

Using EMC VNX Storage with VMware vSphere

Version 3.0

- Configuring VMware vSphere on VNX Storage
- Cloning Virtual Machines
- Establishing a Backup and Recovery Plan for VMware vSphere on VNX Storage
- Using VMware vSphere in Data Restart Solutions
- Using VMware vSphere for Data Vaulting and Migration

Jeff Purcell

Copyright © 2013 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED "AS IS." EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date regulatory document for your product line, go to the Technical Documentation and Advisories section on the EMC Online Support website.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com.

All other trademarks used herein are the property of their respective owners.

h8229.4

Chapter 1	Configuring VMware vSphere on VNX Storage	
	Introduction	16
	Management tools	18
	VMware vSphere installation.....	25
	VMware vSphere configuration.....	31
	Provisioning VNX Storage for vSphere	50
	Unified storage considerations	58
	vSphere storage configuration	77
	Network considerations.....	105
	Virtual machine considerations	108
	Monitor and manage storage	119
	Storage efficiency	132
	VNX storage options	145
Chapter 2	Cloning Virtual Machines	
	Introduction	160
	Using EMC VNX cloning technologies.....	162
	Summary	176

Chapter 3	Backup and Recovery Options	
	Introduction	178
	Virtual machine data consistency	179
	VNX native backup and recovery options	181
	Snapshot backup and recovery of a VMFS datastore	183
	Backup and recovery of RDM volumes	186
	Replication Manager	187
	Backup and recovery of a VMFS with VNX Advanced Snaps	192
	vStorage APIs for Data Protection	200
	Backup and recovery using VMware Data Recovery	201
	Backup and recovery using Avamar	204
	Backup and recovery using NetWorker	213
	Summary	219
Chapter 4	Using VMware vSphere in Data Restart Solutions	
	Introduction	222
	EMC Remote Replication technology overview	225
	RDM volume replication	247
	EMC Replication Manager	251
	Automating site failover with SRM and VNX	254
	Summary	264
Chapter 5	Data Vaulting and Migration	
	Introduction	266
	SAN Copy interoperability with VMware file systems	267
	SAN Copy interoperability with RDM virtual disks	268
	Using SAN Copy for data vaulting	269
	Importing Storage into the remote environment	276
	SAN Copy for data migration to VNX arrays	279
	Summary	283

	Title	Page
1	EMC Unisphere interface	18
2	LUN properties.....	19
3	VSI Feature Manager	20
4	Unified Access Control workflow	22
5	Storage Viewer NFS datastore details.....	23
6	Storage Viewer VNX block storage details.....	23
7	Configuration workflow	24
8	Unisphere LUN assignment for ESXi boot device	27
9	VNX iSCSI port management interface	29
10	iBFT interface for VNX target configuration.....	30
11	VNX storage with VMware vSphere.....	32
12	ESXi topology with FC/FCoE/iSCSI/NFS connectivity to VNX.....	34
13	VNX configuration of host initiator.....	36
14	VMkernel port configuration.....	40
15	VMkernel adapter binding in vSphere 5	42
16	Minimum configuration for VNX iSCSI targets	44
17	Recommended configuration for VNX iSCSI targets.....	45
18	Bind iSCSI Initiator with VMkernel network adapter	46
19	Disable Delayed Acknowledgement setting on storage adapter	48
20	File storage provisioning with USM.....	51
21	Creating a new NFS datastore with USM.....	52
22	File storage provisioning with USM.....	54
23	Creating a new VMFS datastore with USM	56
24	LUN ownership	63
25	LUN trespass.....	64
26	VMkernel pluggable storage architecture	65
27	Esxcli command output.....	66
28	VSI Path Management feature.....	68
29	Storage Viewer LUNs view	69
30	Elements of a multipathing configuration for NFS.....	71

31	Unisphere interface.....	72
32	Data Mover link aggregation for NFS server.....	73
33	vSphere networking configuration	74
34	VMkernel Properties window.....	75
35	Virtual machine configured on a Thick LUN	78
36	Virtual machine migrated to a Thin LUN	79
37	Plug-in Installation.....	83
38	NFS Hardware Accelerated Datastore Property	84
39	Create File System.....	84
40	Vmkfstools disk utilization option	85
41	Storage DRS datastore cluster	86
42	SDRS advanced policy configuration.....	88
43	SDRS I/O metric enablement setting.....	89
44	VASA datastore storage capability of VNX Flash drive LUN.....	91
45	Storage profile assignment	94
46	Compatible or incompatible with SAS Fibre storage profile.....	95
47	Creating a user-defined profile	96
48	Creation of a user-defined virtual machine storage profile.....	97
49	Associating datastores with a user-defined storage profile.....	98
50	Associating the virtual machine with a user defined storage capability	98
51	VASA configuration	100
52	Virtual disk shares configuration	102
53	NFS SIOC congestion window.....	104
54	Network Resource Allocation interface.....	105
55	vSphere 5 Datastore removal wizard.....	107
56	Select the disk	109
57	Guest disk alignment validation.....	111
58	NTFS data partition alignment (wmic command).....	111
59	Output of 1 MB aligned Linux partition.....	112
60	Output for an unaligned Linux partition (starting sector 63).....	112
61	Host Cache configuration on VNX EFD storage	114
62	Enable NPIV for a virtual machine after adding an RDM volume	116
63	Manually register virtual machine (virtual WWN) initiator records...	117
64	Data Alarm Settings—Actions window	120
65	Storage Viewer\Datastores window—VMFS datastore	121
66	Adjustable percent full threshold for the storage pool.....	123
67	Create Storage Usage Notification window	124
68	User-defined storage usage notifications	125
69	User-defined storage projection notifications.....	126
70	VNX Monitoring and Reporting - Capacity Planning Report.....	127
71	VNX Monitoring and Reporting - Performance report	128
72	vCenter Operations Manager Dashboard	129
73	vCenter Operations Manager - VNX Storage Analytics.....	131

74	Thick or zeroedthick virtual disk allocation.....	134
75	Thin virtual disk allocation.....	135
76	Virtual machine disk creation wizard.....	136
77	Virtual machine out-of-space error message.....	137
78	File system High Water Mark in the EMC VSI: USM feature.....	139
79	Provisioning policy for an NFS virtual machine virtual disk.....	140
80	LUN compression property configuration.....	141
81	VNX FAST VP reporting and management interface.....	151
82	Disk Provisioning Wizard.....	157
83	Unisphere clone LUN management.....	164
84	Performing a consistent clone fracture operation.....	165
85	Creating a SnapView session to create a copy of a VMware file system.....	167
86	Device signature assignment.....	169
87	Selecting virtual machine configuration files in the Datastore Browser.....	170
88	Adding the new virtual machine to the ESXi host inventory.....	170
89	Creating a writeable NAS datastore checkpoint.....	171
90	Cloned NFS datastore in vSphere.....	174
91	ShowChildFsRoot parameter properties in Unisphere.....	181
92	Snapshot Configuration Wizard.....	184
93	Snapshot Configuration Wizard (continued).....	185
94	Replication Manager Job Wizard.....	188
95	Replica Properties in Replication Manager.....	189
96	Replication Manager virtual machine restore.....	190
97	Read-only copy of the datastore view in the vSphere client.....	191
98	Advanced Snapshot Basic Configuration.....	193
99	Snapshot Mount Point.....	194
100	Mount Point configuration wizard.....	195
101	Snapshot consistency group creation.....	196
102	Consistency group snapshot creation.....	197
103	Consistency group snapshot attach.....	198
104	VADP flow diagram.....	200
105	VMware Data Recovery.....	201
106	Sample Avamar environment.....	205
107	Sample proxy configuration.....	207
108	Avamar backup management configuration options.....	208
109	Avamar virtual machine image restore.....	210
110	Avamar browse tree.....	211
111	NetWorker-virtualization topology view.....	214
112	VADP snapshot.....	214
113	NetWorker configuration settings for VADP.....	215
114	NDMP recovery using NetWorker.....	217

115	Backup with integrated checkpoint.....	218
116	Replication Wizard	229
117	Replication Wizard (continued).....	230
118	Preserving dependent-write consistency with MirrorView consistency group technology	233
119	EMC VMware Unisphere interface	235
120	Enable MirrorView between VNX systems	236
121	MirrorView Wizard — select source LUNs	237
122	MirrorView Wizard — select remote storage	238
123	Promote mirrored LUN.....	239
124	Business continuity solution using MirrorView/S in a virtual infrastructure with VMFS	240
125	Synchronize MirrorView LUNs.....	241
126	RecoverPoint architecture overview	242
127	Disabling VAAI support on an ESXi host	245
128	RM protection for NFS datastores and virtual machines.....	252
129	Using the vSphere client to register a virtual machine with ESXi	253
130	SRM recovery plan summary.....	255
131	VMware vCenter SRM configuration	256
132	Create an SRM protection group	257
133	Recovery plan test.....	259
134	Recovery plan cleanup	260
135	SRM recovery plan with EMC MirrorView	261
136	SRM reprotect.....	262
137	Data vaulting with Incremental SAN Copy.....	270
138	Using Unisphere or Storage Viewer to identify source LUNs	271
139	Creating an Incremental SAN Copy session.....	273
140	Creating an Incremental SAN Copy session (continued)	274
141	Creating a SAN Copy session to migrate data to a VNX.....	280

	Title	Page
1	Recommended NMP path selection plug-in	67
2	NFS VAAI features.....	82
3	Supported SDRS LUN configurations	90
4	VASA storage capability mapping to VNX LUNs	92
5	VNX OE for Block 5.32 storage capability mapping to VNX LUNs	93
6	SIOC congestion windows.....	103
7	VNX Connector metrics	130
8	Command line descriptions for vSphere 4 and vSphere 5.....	132
9	Virtual machine disk allocation policies	133
10	VNX supported disk types	146
11	Pool capabilities.....	148
12	VNX RAID options	149
13	Thin LUNs versus Thick LUNs.....	155
14	VNX-based technologies for virtual machine cloning	176
15	Backup and recovery options	220
16	EMC replication options for VMware environments	226
17	VNX MirrorView limits.....	232
18	Minimum revision levels for VAAI support with VNX RecoverPoint splitter	244
19	EMC RecoverPoint feature support.....	246
20	VNX to virtual machine RDM.....	248
21	Data replication solutions	264

As part of an effort to improve and enhance the performance and capabilities of its product lines, EMC periodically releases revisions of its hardware and software. Therefore, some functions described in this document may not be supported by all versions of the software or hardware currently in use. For the most up-to-date information on product features, refer to your product release notes.

If a product does not function properly or does not function as described in this document, please contact your EMC representative.

Note: This document was accurate as of the time of publication. However, as information is added, new versions of this document may be released to the EMC Online Support website. Check the EMC Online Support website to ensure that you are using the latest version of this document.

Audience

This TechBook describes how VMware vSphere works with the EMC VNX series. The content in this TechBook is intended for storage administrators, system administrators, and VMware vSphere administrators.

Note: Although this document focuses on VNX storage, most of the content also applies when using vSphere with EMC Celerra or EMC CLARiiON storage.

Note: In this document, ESXi refers to VMware ESX Server version 5.0. Unless explicitly stated, ESXi 5.x and ESXi are synonymous.

Individuals involved in acquiring, managing, or operating EMC VNX storage arrays and host devices can also benefit from this TechBook. Readers with knowledge of the following topics will benefit:

- ◆ EMC VNX series
- ◆ EMC Unisphere
- ◆ EMC Virtual Storage Integrator (VSI) for VMware vSphere
- ◆ VMware vSphere 4.0, 4.1, and 5.0

Related documentation

The following EMC publications provide additional information:

- ◆ *EMC CLARiiON Asymmetric Active/Active Feature (ALUA)*
- ◆ *EMC VSI for VMware vSphere: Path Management—Product Guide*
- ◆ *EMC VSI for VMware vSphere: Path Management—Release Notes*
- ◆ *EMC VSI for VMware vSphere: Unified Storage Management—Product Guide*
- ◆ *EMC VSI for VMware vSphere: Unified Storage Management—Release Notes*
- ◆ *EMC VSI for VMware vSphere: Storage Viewer—Product Guide*
- ◆ *EMC VSI for VMware vSphere: Storage Viewer—Release Notes*
- ◆ *Migrating Data From an EMC CLARiiON Array to a VNX Platform using SAN Copy - white paper*

The following links to the VMware website provide more information about VMware products:

- ◆ <http://www.vmware.com/products/>
- ◆ http://www.vmware.com/support/pubs/vs_pubs.html

The following document is available on the VMware web site:

- ◆ *vSphere iSCSI SAN Configuration Guide*

Conventions used in this document

EMC uses the following conventions for special notices:



DANGER indicates a hazardous situation which, if not avoided, will result in death or serious injury.

⚠ WARNING

WARNING indicates a hazardous situation which, if not avoided, could result in death or serious injury.

⚠ CAUTION

CAUTION, used with the safety alert symbol, indicates a hazardous situation which, if not avoided, could result in minor or moderate injury.

NOTICE

NOTICE is used to address practices not related to personal injury.

Note: A note presents information that is important, but not hazard-related.

IMPORTANT

An important notice contains information essential to software or hardware operation.

Typographical conventions

EMC uses the following type style conventions in this document.

Normal	Used in running (nonprocedural) text for: <ul style="list-style-type: none"> Names of interface elements (such as names of windows, dialog boxes, buttons, fields, and menus) Names of resources, attributes, pools, Boolean expressions, buttons, DQL statements, keywords, clauses, environment variables, functions, utilities URLs, pathnames, filenames, directory names, computer names, filenames, links, groups, service keys, file systems, notifications
Bold	Used in running (nonprocedural) text for: <ul style="list-style-type: none"> Names of commands, daemons, options, programs, processes, services, applications, utilities, kernels, notifications, system calls, man pages <p>Used in procedures for:</p> <ul style="list-style-type: none"> Names of interface elements (such as names of windows, dialog boxes, buttons, fields, and menus) What user specifically selects, clicks, presses, or types

<i>Italic</i>	Used in all text (including procedures) for: <ul style="list-style-type: none"> • Full titles of publications referenced in text • Emphasis (for example a new term) • Variables
<code>Courier</code>	Used for: <ul style="list-style-type: none"> • System output, such as an error message or script • URLs, complete paths, filenames, prompts, and syntax when shown outside of running text
Courier bold	Used for: <ul style="list-style-type: none"> • Specific user input (such as commands)
<i>Courier italic</i>	Used in procedures for: <ul style="list-style-type: none"> • Variables on command line • User input variables
< >	Angle brackets enclose parameter or variable values supplied by the user
[]	Square brackets enclose optional values
	Vertical bar indicates alternate selections - the bar means “or”
{ }	Braces indicate content that you must specify (that is, x or y or z)
...	Ellipses indicate nonessential information omitted from the example

We’d like to hear from you!

Your feedback on our TechBooks is important to us! We want our books to be as helpful and relevant as possible, so please feel free to send us your comments, opinions and thoughts on this or any other TechBook:

TechBooks@emc . com

Configuring VMware vSphere on VNX Storage

This chapter includes the following topics:

◆ Introduction	16
◆ Management tools.....	18
◆ VMware vSphere installation.....	25
◆ VMware vSphere configuration.....	31
◆ Provisioning VNX Storage for vSphere	50
◆ Unified storage considerations	58
◆ vSphere storage configuration	77
◆ Network considerations.....	105
◆ Virtual machine considerations	108
◆ Monitor and manage storage	119
◆ Storage efficiency	132
◆ VNX storage options	145

Introduction

VMware virtualization and EMC® VNX® storage systems are ever present in today's data centers. VMware offers the number one virtualization platform, and the VNX series delivers uncompromising scalability and flexibility for virtual environments while providing market-leading simplicity and efficiency to minimize total cost of ownership.

Customers can benefit from the following VNX features:

- ◆ Unified storage, optimized for virtualized applications.
- ◆ Industry-leading performance with the latest Intel multicore CPUs.
- ◆ VNX allows administrators to combine Flash, SAS, and Near-Line SAS drives to meet any needs within the environment, and scale-out-storage to satisfy future requirements.
- ◆ 6 Gb/s SAS back end with the latest Flash, SAS, and NL-SAS drive technologies.
- ◆ Highly reliable storage system with five 9s of availability.
- ◆ EMC UltraFlex™ I/O connectivity-Fibre Channel (FC), Internet Small Computer System Interface (iSCSI), Common Internet File System (CIFS), Network File System (NFS) including parallel NFS (pNFS), Multi-Path File System (MPFS), and Fibre Channel over Ethernet (FCoE) connectivity.
- ◆ Extended LUN cache using Flash drives.
- ◆ Multiprotocol support for file, block, and object with object access through EMC Atmos® Virtual Edition (Atmos VE).
- ◆ Simplified storage management interface with EMC Unisphere®.
- ◆ VMware management integration through VMware-aware Unisphere, EMC Virtual Storage Integrator plug-in for VMware vCenter™, and VNX VC Operations Manager adapter.

The VNX series is ideal for VMware vSphere with product integration features for storage management and product capabilities that are beneficial for virtual environments.

VMware administrators can take advantage of the following features to manage virtual storage:

- ◆ **Thin provisioning** — Block and File storage conservation and simplified management.
- ◆ **File compression** — NFS efficiency by compressing virtual machine disk files.
- ◆ **File deduplication** — Elimination of redundant files within an NFS file system.
- ◆ **LUN compression** — Condenses data to improve storage utilization in Storage Pools.
- ◆ **FAST VP and FAST Cache** — Automated relocation of subLUN elements to optimize and balance application needs with storage resources.
- ◆ **NFS Virtual Data Mover** — Isolation of NFS services for additional security and replication of NFS environments.
- ◆ **vStorage APIs for Array Integration (VAAI)** — SCSI and NFS storage integration to reduce I/O between the host and the storage system.
- ◆ **Advanced Snapshots** — Up to 3,000 space efficient snapshots with up to 256 snapshots of each source LUN. This feature is available in VNX OE for Block version 5.32 and later.
- ◆ **EMC Replication Manager** — A single interface to manage application-consistent virtual machine replicas on VNX.

Management tools

EMC provides two VMware-centric administrative options for VNX storage management, EMC Unisphere, and EMC Virtual Storage Integrator vSphere Client plug-in.

EMC Unisphere

Unisphere is an easy-to-use, web-enabled interface for remote management of VNX systems. It offers an intuitive interface to manage and monitor the storage system. The customizable dashboard views provide real time details on the health of the environment as illustrated in [Figure 1](#).

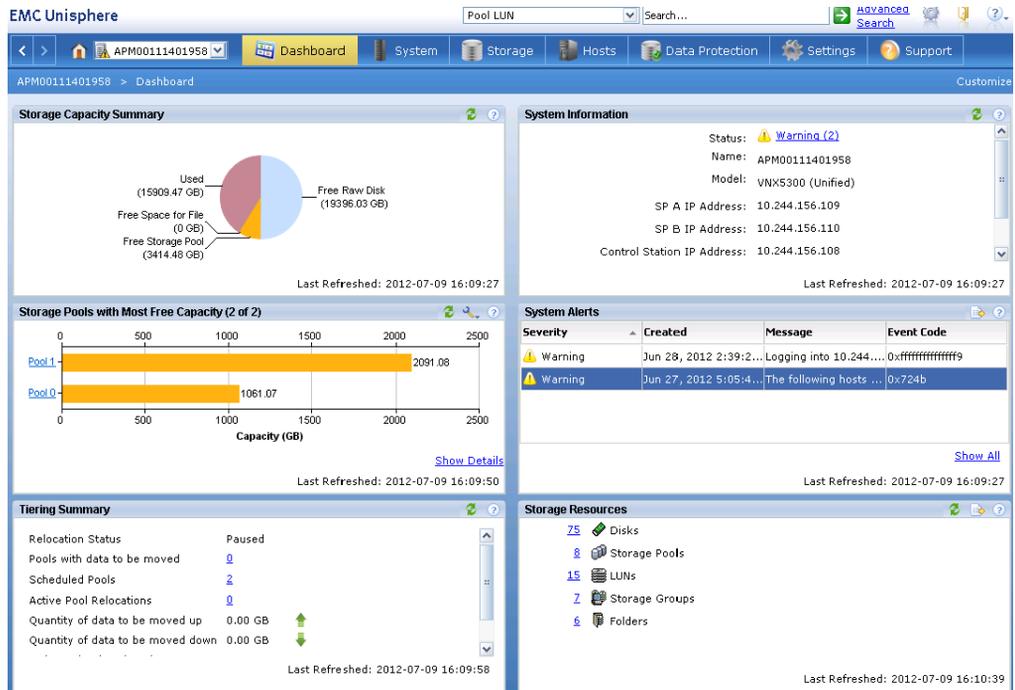


Figure 1 EMC Unisphere interface

Unisphere includes VMware discovery capabilities to collect virtual machine and datastore storage details from vSphere and display them in the context of VNX storage system devices. This integration allows Unisphere administrators to understand how VNX storage is used within the vSphere environment. [Figure 2 on page 19](#) illustrates the properties of LUN number 17.

The interface identifies that the LUN is assigned to host ucs23.emc.lab and is being used by a virtual machine named Ora11gR2VM. Unisphere also provides additional information about the virtual disks and the datastore used to support the virtual machine.

Use information presented in this interface to monitor the environment and validate virtual disk placement when configuring storage system replication and data protection policies.

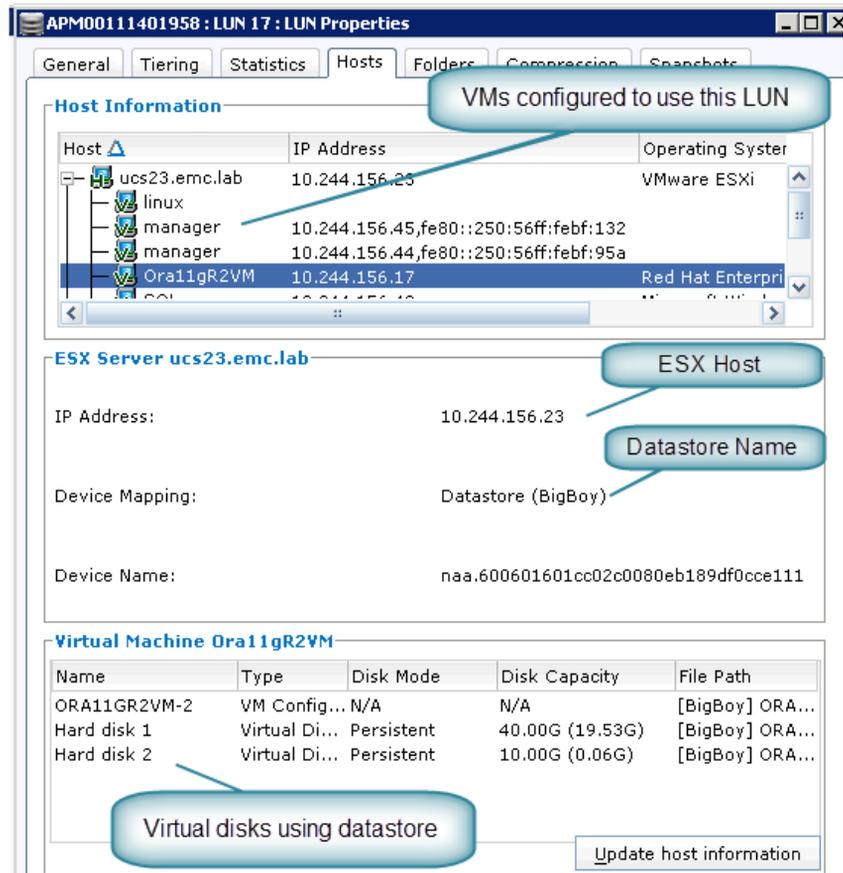


Figure 2 LUN properties

EMC VSI for VMware vSphere

Virtual Storage Integrator (VSI) is a vSphere Client plug-in framework that extends storage management capabilities to vCenter. VSI has a modular framework that allows management features to be added or removed in support of specific EMC products installed in the environment. This section describes the VSI Unified Storage Management (USM), Storage Viewer, and Path Management features that are most applicable to the VNX.

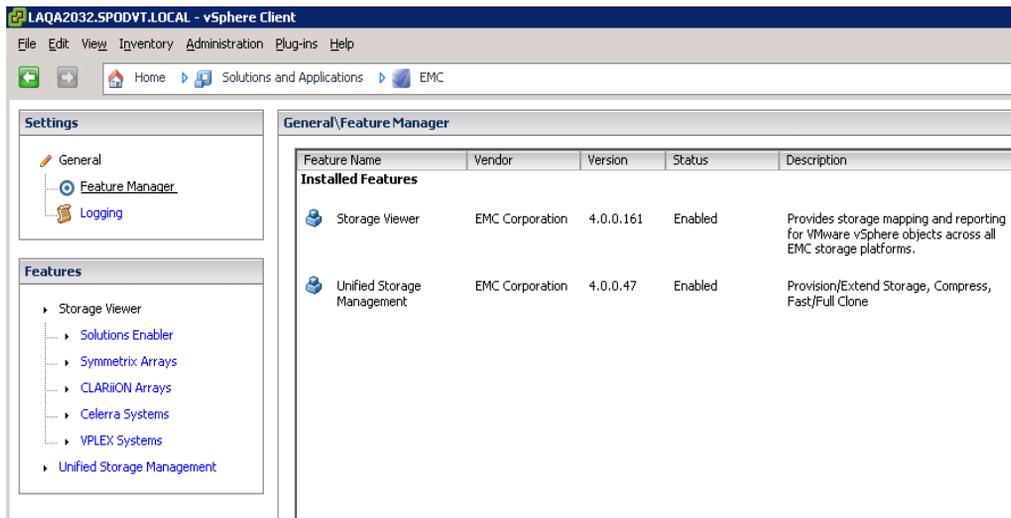


Figure 3 VSI Feature Manager

VSI: Unified Storage Management

The Unified Storage Management (USM) feature allows vSphere administrators to manage VNX storage through the vCenter interface. It automates datastore and RDM creation by performing the vSphere and VNX provisioning tasks required to provision storage to an ESXi™ host or datacenter cluster.

USM functionality includes:

- ◆ End-to-end datastore provisioning in accordance with EMC best practices.
- ◆ MultiLUN creation and masking for Raw Device Mapping (RDM) or Virtual Machine File System (VMFS) use.
- ◆ Rapid provisioning of full virtual machine clones or space-efficient fast clones within NFS datastores.
- ◆ NFS datastore deduplication
- ◆ Compression of virtual disk files.

VSI Unified Access Control

USM requires administrative or delegated administrative rights to access and manage the storage system. VSI Unified Access Control (UAC) for VNX is a utility that grants VNX management access to authorized users. It operates under an explicit deny model with view and management entitlements granted at the RAID group, storage pool, or NFS file system level.

UAC rights are also exported as an encrypted key. Storage administrators import the key into other systems running the VI client. Figure 4 illustrates the steps to create an access profile.

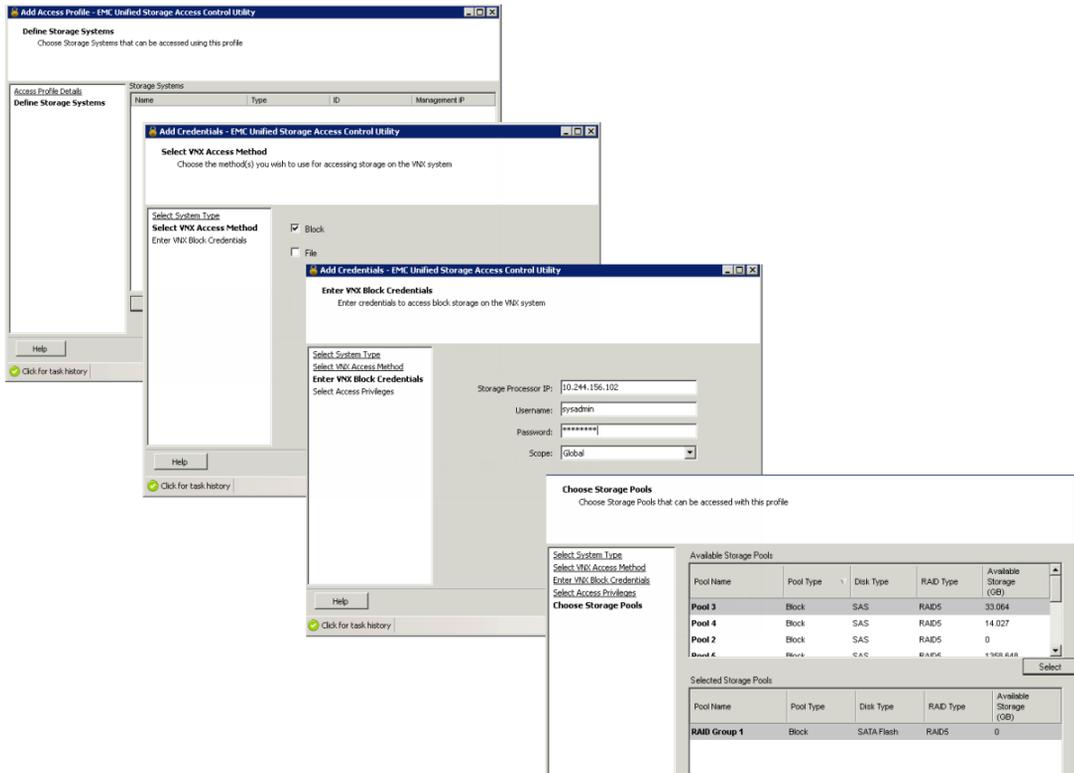


Figure 4 Unified Access Control workflow

VSI: Storage Viewer

VSI Storage Viewer presents VNX storage details of LUNs, File Systems, and data paths in the context of the ESXi datastore. Contextual device information is presented within the VI client when a datastore or LUN is selected. Information listed within this interface is useful for identifying device details to troubleshoot the environment and perform the following storage administration tasks:

- ◆ Presents storage information in a common view within the vSphere Client.
- ◆ Enables VMware administrators to identify VNX storage properties of VMFS, NFS, and RDM storage.

- ◆ Presents LUN connectivity and device details for VNX storage.

Figure 5 provides an example of Storage Viewer for VNX file devices. This view provides details for the VNX System ID, file system, RAID type, storage pool, and so on.



Figure 5 Storage Viewer NFS datastore details

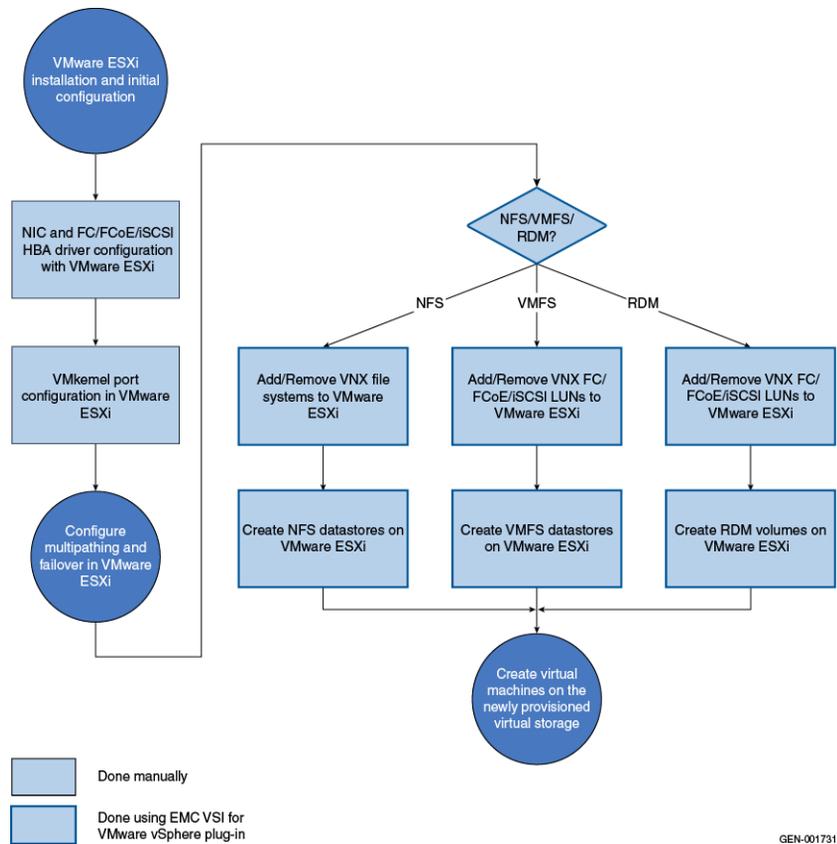
Figure 6 provides an example of Storage Viewer for VNX block devices. This view provides details for the VNX System ID, LUN ID, RAID type, LUN type, and so on.



Figure 6 Storage Viewer VNX block storage details

vSphere installation and configuration on VNX

The configuration of the ESXi environment begins with the installation of the ESXi hypervisor. Supported storage devices for installation of the ESXi image are: a local server disk, a USB storage device, or a SAN SCSI LUN in a boot from SAN configuration. [Figure 7](#) illustrates a workflow to get the ESXi systems installed and configured with a VNX storage system. The remaining sections of this chapter mirror the steps outlined in the workflow with the intent of providing a logical process to build out the vSphere environment. Some of these tasks can be automated with the vSphere autodeploy and host profile features to accelerate deployment of additional hosts.



GEN-001731

Figure 7 Configuration workflow

VMware vSphere installation

Installing the ESXi image on the SAN provides improved performance and reliability through:

- ◆ RAID-protected Tier 1 storage to eliminate downtime that results from a local disk failure.
- ◆ I/O distribution across multiple spindles and multiple I/O channels.
- ◆ Simplified host replacement in the event of a hardware failure.

Note: With vSphere 5, the installation process is automated to significantly reduce installation time for larger environments. See the *vSphere Installation and Setup Guide* for details on Auto Deploy.

vSphere boot from SAN LUNs

Cable the hosts. Zone the HBAs and LUNs to ensure the host initiators log in to the VNX storage processors (SPs) when the host is powered on.

1. Gather the following information to configure the environment to use the selected front-end ports on the array:
 - ESXi hostname
 - IP addresses
 - The HBA WWN
 - Obtain the WWN from the **Unisphere Host Connectivity** page after the initiators log into the SPs, or from within ESXi.
 - VNX management IP address and credentials

Note: If storage zoning is not complete, obtain the HBA World Wide Names (WWNs) from the SAN switch.

2. Power on the ESXi host.

3. Modify the host BIOS settings to establish the proper boot order. Ensure the SAN boot device appears immediately after the peripheral devices:
 - Unless explicitly required, disable the local RAID SCSI controller on the host.
 - Virtual floppy or CD-ROM device.
 - Local device.

Note: Even though this is a SAN boot, the VNX LUN BIOS identifies it as a local device.

- For software iSCSI, enable iSCSI boot support on the network card.
4. Enable the FC, FCoE, or iSCSI adapter as a boot device, and scan the bus to initiate a Port Login.
 5. Display the properties of the Array Controllers to verify that the adapter can access the VNX.
 6. Access Unisphere to view the Host Connectivity Status. Verify that the adapters are logged in to the correct SP ports.
 7. Boot from SAN requires manual registration of the HBAs. Select the new initiators and manually register them using the fully qualified domain name of the host. Set the failover mode to Asymmetrical Logical Unit Access (ALUA) mode for support of vStorage API for Array Integration (VAAI) and Native Multipathing Plug-in (NMP) autorestore.

Note: In some servers, the host initiators may not appear until the host operating system installation starts. An example of this is ESXi installations on Cisco UCS, which lacks an HBA BIOS probe capability.

8. Create LUN on which to install the boot image. The LUN does not need to be any larger than 20 GB. Do not store virtual machines within the datastore created from this LUN.
9. Create a storage group and add the host record and the new LUN to it.
10. Rescan the host adapter to force the host to discover the new device. If the LUN does not appear, or still appears as LUN Z, recheck the configuration and rescan the HBA.

- It is a good practice to reserve a specific Host LUN ID to identify the boot devices. For example, assign a Host LUN number of 0 to LUNs that contain the boot volume. This approach makes it easy to differentiate the boot volume from other LUNs assigned to the host as shown in Figure 8.

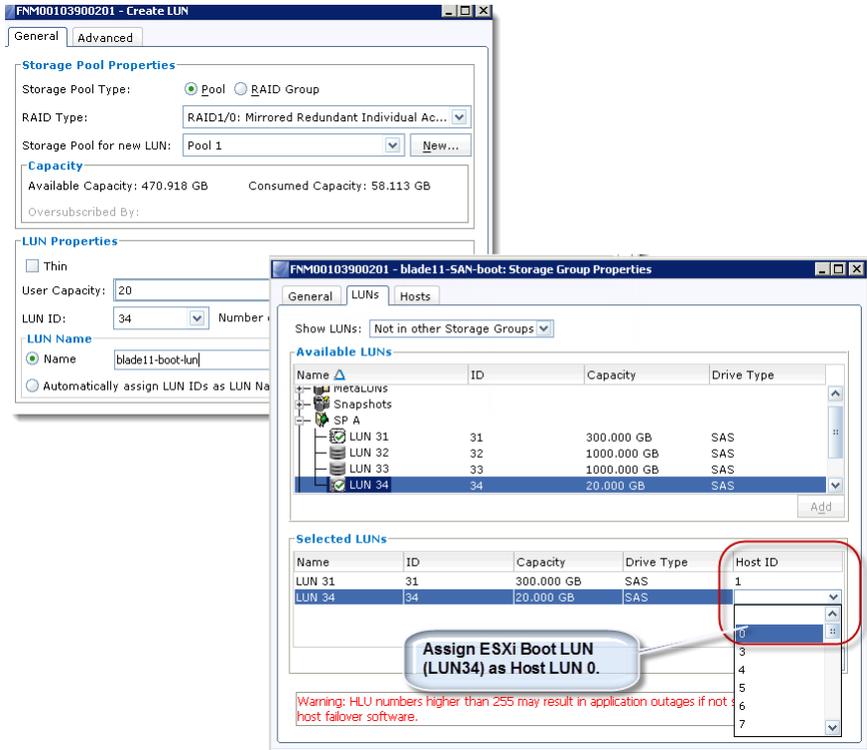


Figure 8 Unisphere LUN assignment for ESXi boot device

- Ensure the CD-ROM/DVD-ROM/USB/virtual media is in the caddy and precedes the local device in the boot order.

Note: The BIOS does not differentiate a SAN boot device from a local disk.

- Begin the ESXi installation. Select the DGC device, and follow the installation steps to configure the host.

vSphere boot from SAN iSCSI LUNs

ESXi 4.1 and later includes support for iSCSI software boot firmware (iBFT). iBFT from VNX iSCSI storage provides similar benefits to those described in boot from FC and the configuration tasks are nearly identical. The choice of storage protocol is largely a matter of preference and infrastructure.

The network card used must support software initiator boot and it should support gigabit or greater throughput for this configuration to work correctly. Consult the *VMware Compatibility Guide* to verify that the device is supported before beginning this procedure.

Access the iSCSI adapter configuration utility during the system boot to configure the HBA:

- ◆ Set the IP address and IQN name of the iSCSI initiator.
- ◆ Define the VNX iSCSI target address.
- ◆ Scan the target.
- ◆ Enable the boot settings and the target device.

Refer to the vendor documentation for instructions to enable and configure the iSCSI adapter:

1. Each initiator requires a unique IQN for storage group assignment on the VNX platform. Specify a unique IQN for each iSCSI initiator in the environment.
2. Use Unisphere to configure an iSCSI portal on the VNX platform.

VNX iSCSI supports jumbo frames with MTU values of 1488-9000. When enabling Jumbo Frames, verify that all components in the I/O path (host, network switch port, and storage interface) support jumbo frames, and that their MTU sizes are consistent.

- Specify the IP address and IQN name of the iSCSI port from the previous step to configure the iSCSI target. There is an option to configure CHAP for additional iSCSI session security.

The screenshot displays the 'iSCSI Port Properties' dialog box for 'Port 0'. It is organized into several sections:

- Port Identification:** Contains fields for 'IQN' (iqn.1992-04.com.emc:cx.fnm00103900201.a4) and 'Alias' (0201.a4).
- Physical Port Properties:** Includes 'Mac Address' (00-60-16-41-5D-A6), 'Actual Speed (Mbps)' (10000), 'Selected Speed (Mbps)' (10000), and 'MTU (bytes)' (9000).
- Virtual Port Properties:** A table listing virtual ports and their IP addresses.

Virtual Port	VLAN ID	IP Addresses
Virtual Port 0		10.10.10.1
- Connector:** Lists hardware identifiers: 'EMC Part Number' (019-078-041), 'EMC Serial Number' (0000000000000000), 'Vendor Part Number' (AFBR-703ASDZ-E2), and 'Vendor Serial Number' (AGL1027A1000J4K).

Buttons for 'OK', 'Apply', 'Cancel', and 'Help' are located at the bottom of the dialog.

Figure 9 VNX iSCSI port management interface

```
iSCSI Configuration Utility v2.7.6
Copyright (C) 2000-2008 Broadcom Corporation
All rights reserved.

----- 1st Target Parameters -----
Connect:      Enabled
IP Address:   10.0.0.2
TCP Port:    3260
Boot LUN:     0
iSCSI Name:  iqn.1992-04.com.emc:cx.fnm00095200014.a4
CHAP ID:
CHAP Secret:
```

Figure 10 iBFT interface for VNX target configuration

4. Configure the secondary target with the address information for the iSCSI port on VNX SP B.
5. Open Unisphere to complete the following tasks:
 - Register the new initiator record
 - Create a new storage group
 - Create a new boot LUN
 - Add the newly registered host to the storage group
6. Proceed with the ESXi image installation.

VMware vSphere configuration

VNX is a scalable storage system that satisfies shared storage requirements in mid- to high- end vSphere environments. The VNX architecture addresses a broad range of application and scalability requirements making it an ideal platform for vSphere. This section discusses the relationship of vSphere features and notes considerations when used with VNX.

Host connectivity

Proper host storage connectivity is a key element to obtaining the most value from the vSphere and VNX systems. Host connectivity consists of physical cabling, port, or WWN zoning, host adapter settings, and storage port configuration.

ESXi and VNX provide common support for Fibre Channel, FCoE, iSCSI, and NFS storage protocols as shown in Figure 11. VNX also offers the CIFS file sharing protocol for shared file system access by Windows virtual machines. This section describes considerations when establishing ESXi host connectivity to a VNX storage environment.

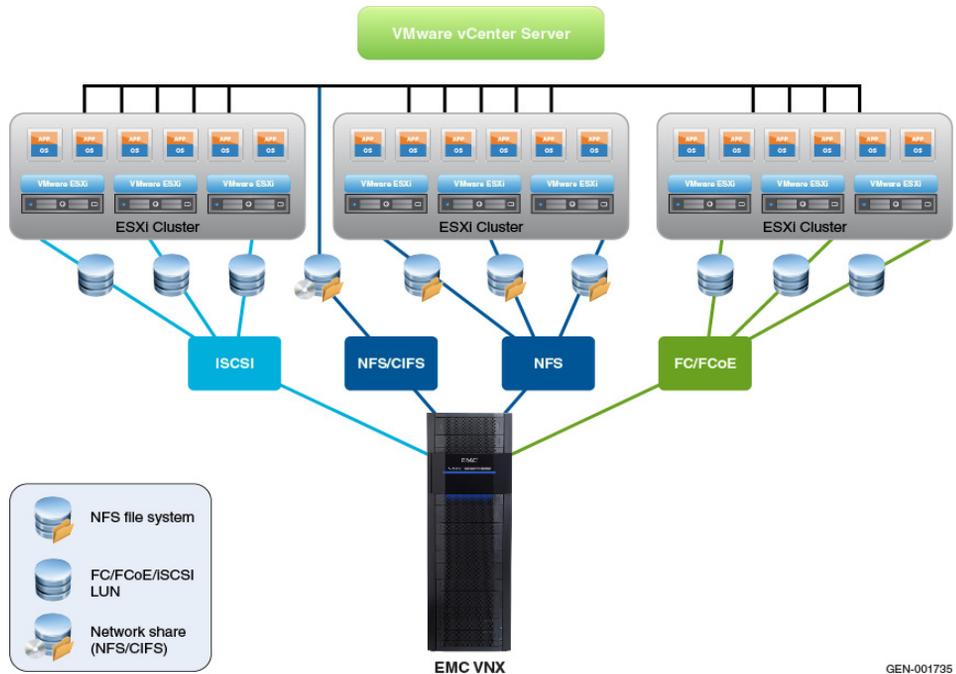


Figure 11 VNX storage with VMware vSphere

Note: VNX and ESXi support one SCSI transport type at a time. An ESXi host can access a LUN using an FCoE, FC, or iSCSI interface. However, accessing the same device using different SCSI transport protocols is unsupported.

Physical configuration

The following steps offer recommendations for general ESXi host connectivity to the VNX storage system:

1. Configure each ESXi host with at least two physical host adapters for device path redundancy between the host and the storage system.
2. Cable each physical path through a separate switch for redundancy and fault tolerance.

3. Logically, a separate switch zone should be created for each initiator-target pair with each HBA zoned to separate ports of each SP.
4. All of the hosts' initiators should be added to a single storage group on the VNX.

Port configuration

VNX storage systems include four on-board 8 Gb FC ports with expansion slots to accommodate additional I/O modules for FC, FCoE, iSCSI, and Ethernet connectivity. VNX systems can be customized with connectivity options that match host requirements and distribute host I/O to the storage system.

ESXi hosts should have a minimum of two physical paths to the storage system. Ideally, the cabling for each path (or pair of paths for path counts greater than two) will be connected to separate physical switches.

Distribute ESXi host adapter connections across all available SP I/O ports to increase parallelism to the target device through multipathing and achieve the best overall response times. Make note of port requirements for MirrorView and RecoverPoint when planning port configurations.

Figure 12 illustrates basic FC/FCoE and iSCSI topologies for connectivity to the VNX.

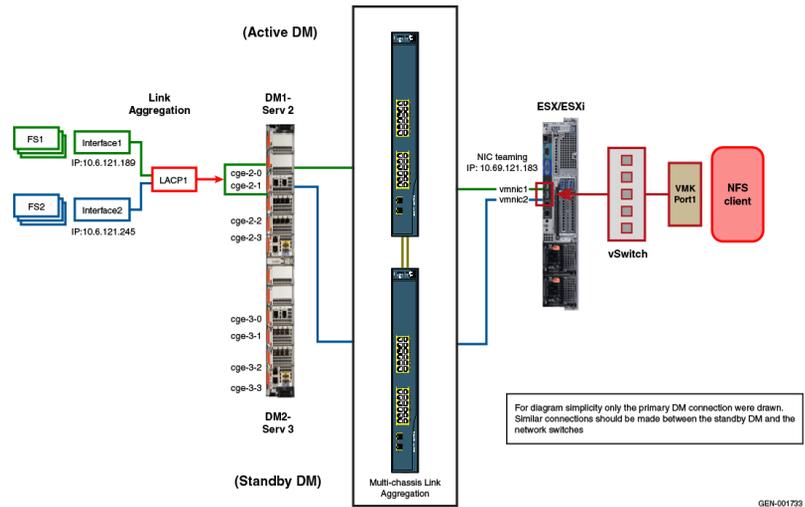


Figure 12 ESXi topology with FC/FCoE/iSCSI/NFS connectivity to VNX

Note: The iSCSI hardware initiator configuration is similar to the FC HBA configuration.

ESX HBAs queue depth

The ESX host adapters provide connectivity to the storage system. In most cases the default adapter settings are sufficient and no additional configuration is required at the ESXi or VNX system when the HBA is installed.

One potential exception is in the case of the HBA queue depth. The default value for max queue depth an ESXi 5 HBAs is 64. That means the VMkernel will only have 64 outstanding I/Os at any given time. In general, this value need not be altered, particularly when there are more than three hosts accessing a device.

Within the VNX, relevant I/O queues that could impact ESXi host performance are the front host port queue and the LUN queue that is being used to support the datastore. It is recommended that the maximum number of I/Os per front end port be limited to 1600 or less and with the front end connectivity options provided with VNX additional SLICs allow for scaling host connectivity.

The LUN queue is the most important consideration when tuning host adapter queues. In most cases, the only time you consider modifying the maximum HBA queue depth is when the LUN queue depth is larger than the cumulative queue depth of all host adapters accessing the LUN, and the `esx_top` value of the device queue used percentage (`%USD`) is continuously at 100 and queued commands (`WQLEN`) for the device are greater than 0.

For example, a LUN created disk from a 20 disk VNX Pool in VNX OE for Block 5.32 has an approximate queue depth of 224. The host adapter queue depth is 64. If the host is part of a 2 node cluster, the cumulative maximum queue depth is 128 which means the host adapter may be limiting the I/O capabilities of the application.

Set the `Disk.SchedNumRequestsOutstanding` to match this value. If the multiple ESXi hosts are configured in a datastore cluster, the cumulative queue depth can surpass the LUN queue fairly quickly. VMware Storage I/O Control (SIOC) helps avoid a situation where the host queue depths are set too high; however, it is suggested to keep the queue depth at the default of 64.

Fibre Channel Zoning

VNX uses single initiator-single target zoning. For best results limit the number of active paths between an initiator and the VNX SP to one. Create two zones per initiator with one zone configured for the host initiator and one Storage Processor A (SP A) port, and the other zone configured with the host initiator and one Storage Processor B (SP B) port.

In cases where I/O is asynchronous or reliability is favored over performance, an initiator can be zoned to two ports per SP. This could limit I/O throughput during active periods.

Virtual Local Area Networks (VLANs)

While IP storage systems do not use the term "zoning", a similar Ethernet concept is applied through virtual local area networks or VLANs on Ethernet switches. VLANs limit the broadcast domain to switch ports or host adapters that are configured with the same VLAN ID. VLANs provide a method of port isolation between ESXi IP storage adapters and the VNX IP storage adapters used to provide iSCSI and NFS connectivity.

Note: EMC has traditionally recommended the use of separate subnets for network isolation between VNX iSCSI ports. In a non-routed network, iSCSI ports on the ESXi host and VNX system are configured to use the same network addresses as long as they are in separate VLANs.

Manual initiator registration

In certain cases, such as boot from SAN, configure host initiators on the VNX in order to create storage groups for the boot LUN. For these cases, use the Unisphere host initiator interface to create the new initiator records. [Figure 13](#) shows how this registration works.

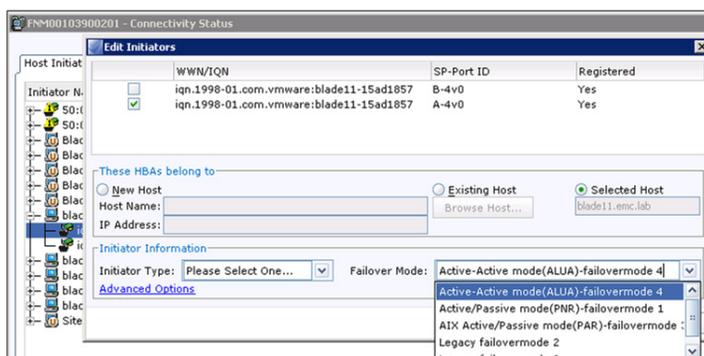


Figure 13 VNX configuration of host initiator

Relevant parameters for the initiator are:

- ◆ ESXi hostname - user provided
- ◆ ESXi management IP address - user provided
- ◆ Initiator type and - CLARiiON/VNX
- ◆ Failover mode - Select failover mode 4 (ALUA).

VNX provides four failover modes; however, only two are applicable when you configure the ESXi host initiators. ESXi is ALUA aware; therefore, configure failover mode 4 in all cases. However, failover mode 1 is described here because it is the default mode configured in CX platforms.

- ◆ **Asymmetrical Active/Active mode (failover mode 4)** — When configured in ALUA mode, the host issues I/O to either VNX SP. The LUN is owned by one SP that provides an optimal I/O path. The peer SP provides a non-optimal path which is used only when all optimal paths have failed or are otherwise unavailable. Failover mode 4 is required for support of VAAI operations on VNX.
- ◆ **Active-Passive mode (failover mode 1)** — This mode uses a single optimal or preferred path to the SP to which the LUN was assigned when it was created. The LUN remains active on that SP unless a disruption occurs at the SP or host level. This mode was used in older CX platforms.

ESXi 4.0 and later are ALUA compliant. This means the ESXi hosts sends I/O to VNX using the active/optimized LUN path(s). If an active/optimized path becomes unavailable, the host attempts to use another active/optimized path on the SP that owns the LUN. If there are no active/optimized paths available, and the host has active paths to the non-optimized SP, it issues a trespass request to the LUN via the peer SP. The peer SP will become the LUN owner and satisfy all subsequent I/O requests. More details on path trespass and restore are in the NMP configuration section.

In vSphere 5.1, all paths are restored to the default owner when the paths are restored.

Fibre Channel over Ethernet (FCoE)

Native FCoE support, included with the VNX platform, offers a simplified physical cabling option between servers, switches, and storage subsystems. FCoE connectivity allows the general server IP-based traffic and I/O to the storage system to be transmitted to and from the server through fewer high-bandwidth, IP-based physical connections.

Converged Network Adapters (CNAs) and FCoE software initiator support in vSphere 5 reduce the physical hardware footprint requirements to support the data traffic and provide a high flow rate through the consolidated network.

High-performance block I/O, previously handled through a separate FC-based data traffic network, can be merged into a single IP-based network with CNAs or 10 GbE adapters that provide efficient FCoE support.

VNX expansion modules add 10 GbE FCoE connectivity with minimal configuration.

Network Considerations

Network equipment

Consider the following items for Ethernet storage networks:

- ◆ Use CAT 6 cables to connect to copper Ethernet networks.
- ◆ Use network switches that support a MultiChassis Link Aggregation technology such as cross-stack EtherChannel or Virtual Port Channeling. [“Multipathing considerations - NFS” on page 69](#) provides more details.
- ◆ Consider FCoE hardware adapters with 10 GbE converged network switches for consolidated storage networks. [“Fibre Channel over Ethernet \(FCoE\)” on page 37](#) provides more details.
- ◆ Select a switch vendor that includes 10GbE support for NFS, iSCSI, or FCoE.

Ethernet configuration considerations

When you configure IP storage networks consider the following:

- ◆ To increase network and I/O efficiency, use a dedicated physical switch or isolated VLAN.
- ◆ On network switches that are also used for the storage network:
 1. Enable flow control.
 2. Enable spanning tree protocol with either RSTP or port-fast enabled.
 3. Restrict bridge protocol data units (PDUs) on storage network ports.
- ◆ In general, ESXi host I/O is random, and in most cases, Jumbo Frames provide minimal benefit. Large block I/O and sequential workloads can benefit from larger frame sizes and VNX supports frames up to 9000 bytes. To improve the performance of I/O-intensive workloads, configure jumbo frames with a consistent MTU size on each network interface (host, switch, VNX) in the I/O path.

- ◆ vSphere 5 supports an FCoE software initiator. Consider software initiator FCoE with 10 GbE network switches to consolidate storage and switching equipment.

VMkernel port configuration in ESXi

ESXi uses VMkernel ports for systems management and IP storage. VMkernel IP storage interfaces provide access to one or more VNX iSCSI network portals or NFS servers.

To configure a VMkernel interface for NFS, use [Figure 14 on page 40](#) as a guide and complete the following steps:

1. Create a new virtual switch to support the IP storage interface(s).
2. Assign network label that describes what the interface is used for, such as NFS.

3. Assign a network adapter from the same physical or logical network as the VNX NFS Server then click **Next**.

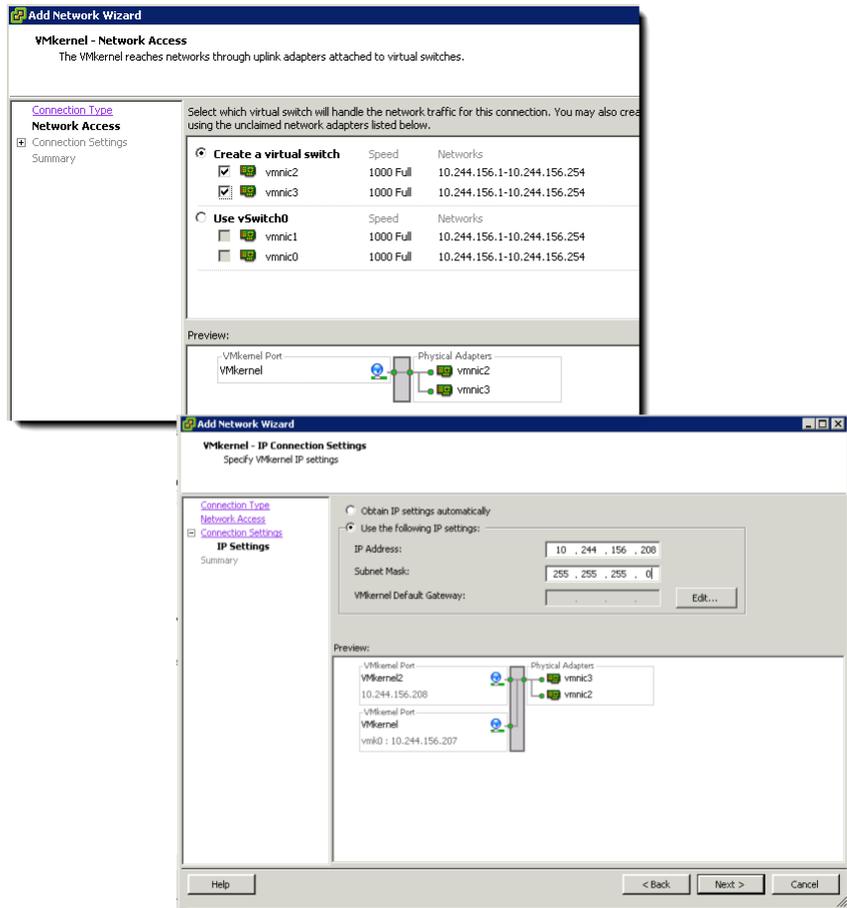


Figure 14 VMkernel port configuration

4. In the **VMkernel - IP Connection Settings** dialog box, specify the following VMkernel IP settings:
 - a. IP address
 - b. Subnet mask
 - c. Default network gateway

Note: Avoid the use of DHCP.

Note: ESXi management and VMkernel interfaces share the default routing table of the ESXi host. As a result, the management interface can inadvertently route storage I/O when an NFS server is configured to use the same subnet. To avoid this scenario, use separate subnets or VLANs for the management and storage networks.

5. Click **Next**. The **Ready to Complete** dialog box appears.
6. Verify the settings and then click **Finish** to complete the process.

iSCSI port binding

iSCSI port binding associates the ESXi iSCSI software initiator with a host network adapter. vSphere 5 and later includes iSCSI management of VMkernel adapters in vCenter, and gives administrators the ability to bind up to eight VMkernel adapters to the software initiator of the ESXi host as shown in [Figure 15](#).

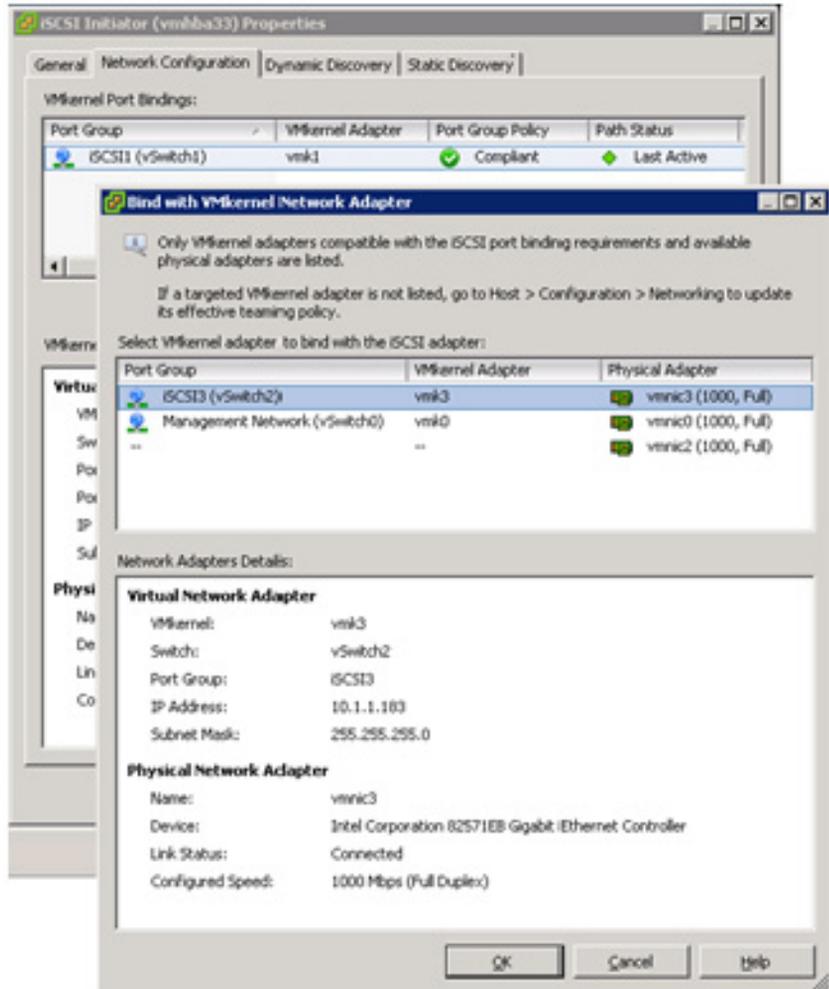


Figure 15 VMkernel adapter binding in vSphere 5

vSphere 5 enforces configuration rules requiring all iSCSI interfaces to be configured with a single physical adapter. There are two configuration options for VNX systems:

- ◆ Configure each adapter with an IP address from a separate network subnet.
- ◆ Use a separate Ethernet switch path to the VNX iSCSI Targets/Network Portals.

[Figure 16 on page 44](#) illustrates the minimum configuration for an ESXi host with two network cards. The network interface for vmk1 is configured with an IP address on the 17.24.110.0/24 subnet. The iSCSI targets on ports A4 and B4 are also configured with addresses on the 17.24.110.0 subnet. ESXi network interfaces for vmk2 and the iSCSI targets on VNX ports A5 and B5 use IP addresses on the 10.1.1.0/24 subnet.

Each VMnic has two paths to the array for a total of four paths.

If the ESXi host uses Native Multipathing, and the LUN is configured in fixed mode, one path, the SP active/optimized path for the LUN is used for I/O. The other paths are set to standby in case the fixed path fails.

If the ESXi host uses PowerPath/VE, or Native Multipathing round-robin, then the host has two active/optimized paths to each LUN and two standby paths in case both active/optimized paths fail.

In both scenarios, if the LUN is owned by SP A, the SP B paths are not used unless there is a failure of both SP A paths.

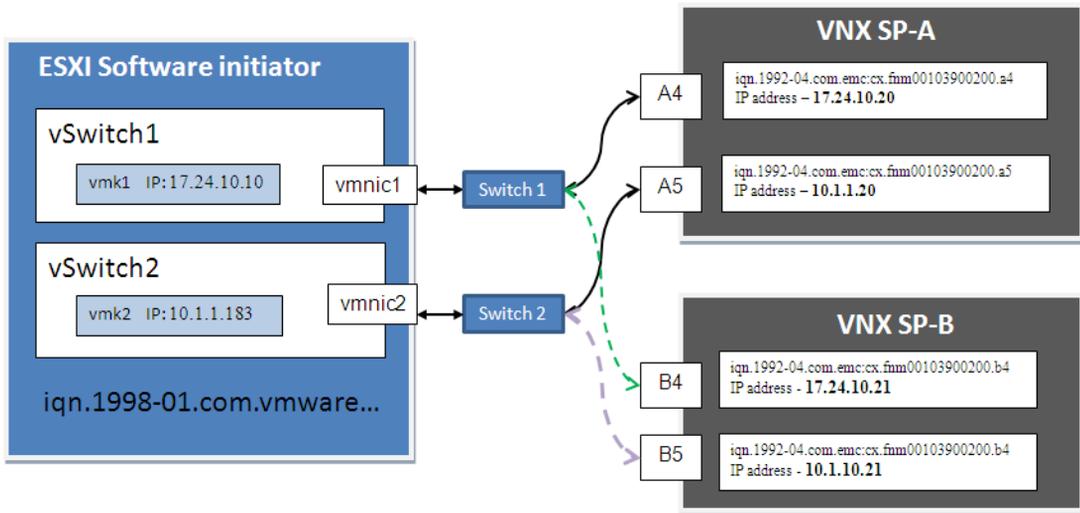


Figure 16 Minimum configuration for VNX iSCSI targets

In an environment where optimum host throughput is required, configure additional ESXi adapters to establish a dedicated path to the VNX iSCSI network portals. The sample configuration illustrated in Figure 17 provides additional dedicated I/O paths for four VNX iSCSI target ports. In this configuration, two dedicated paths are available to each SP. This provides increased bandwidth to any LUNs presented to the host. If the environment requires additional bandwidth or increased availability, configure additional ESXi and VNX iSCSI ports.

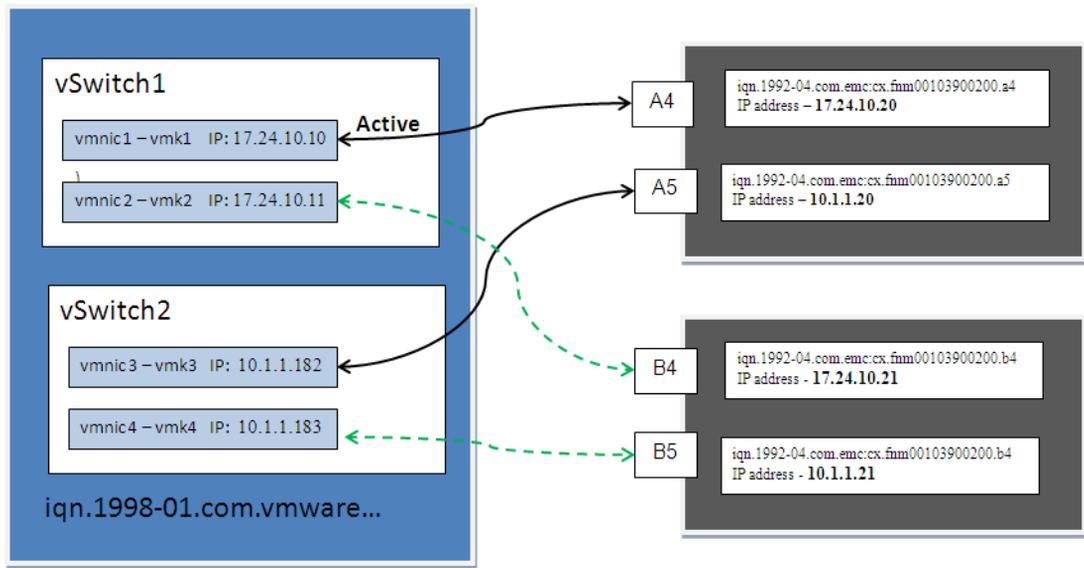


Figure 17 Recommended configuration for VNX iSCSI targets

vSphere 5 provides UI support for iSCSI VMkernel port binding. For earlier releases of vSphere, configure the IP Storage interface using `esxcli` commands. vSphere 5.x provides the option to configure iSCSI through the vSphere Client as shown in Figure 18.

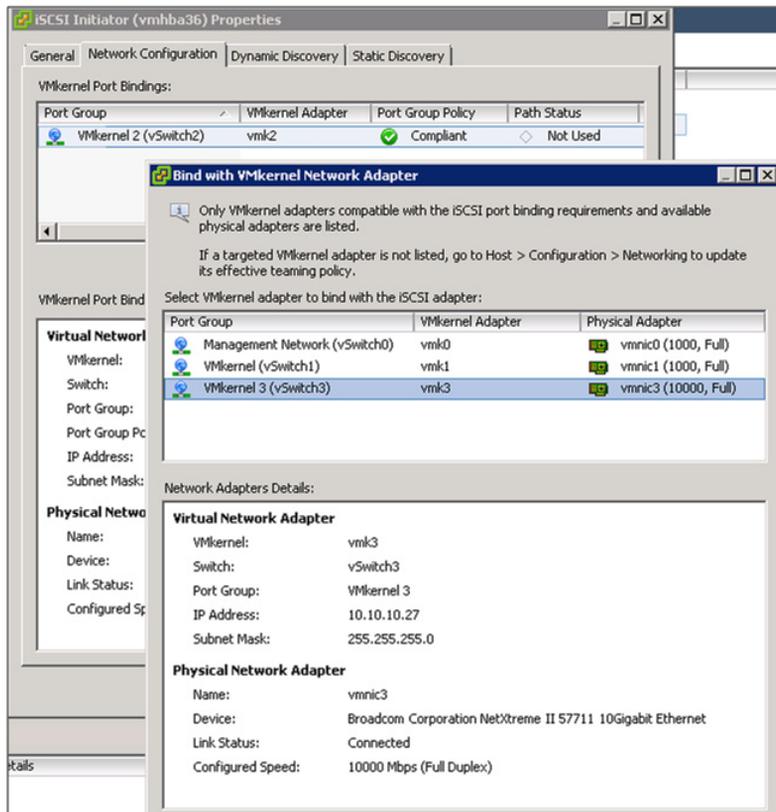


Figure 18 Bind iSCSI Initiator with VMkernel network adapter

Note: For vSphere 4 hosts, run `esxcli` commands on the ESXi host to configure port binding.

Run the following `esxcli` command to activate iSCSI multipathing:

```
# esxcli swiscsi nic add -n <port_name> -d <vmhba>
```

Run the following command to verify that the ports are added to the software iSCSI initiator:

```
# esxcli swiscsi nic list -d <vmhba>
```

Delayed acknowledgement settings for iSCSI

In some circumstances, ESXi hosts encounter suboptimal performance when accessing VNX iSCSI LUNs over a 10GbE interface. The issue results when the ESXi host is configured to use TCP Delayed Acknowledgement. Delayed Acknowledgement is a TCP optimization intended to reduce network packets by combining multiple TCP acknowledgements into a single response to the ESXi host. This works as expected when there are a lot of TCP packets being transmitted between the host and the VNX, however, there are cases such as when a single virtual machine or ESXi host performs a sequential write. In this case the host may write a series of I/Os and wait for acknowledgment. If the VNX has multiple outstanding requests, they are grouped with the acknowledgements and sent as a single packet. However, if the VNX has nothing to respond to except an acknowledgement, it waits for more data. If there are no other packets to send, it waits until the delayed acknowledgement timeout value is reached (200 ms) and then sends the acknowledgement. This behavior has the potential to insert 200 ms delays into the I/O stream. Disable the software iSCSI Delayed Acknowledgement setting on the 10 GbE NIC in cases where performance delays are observed.

Figure 19 illustrates how to disable this setting.

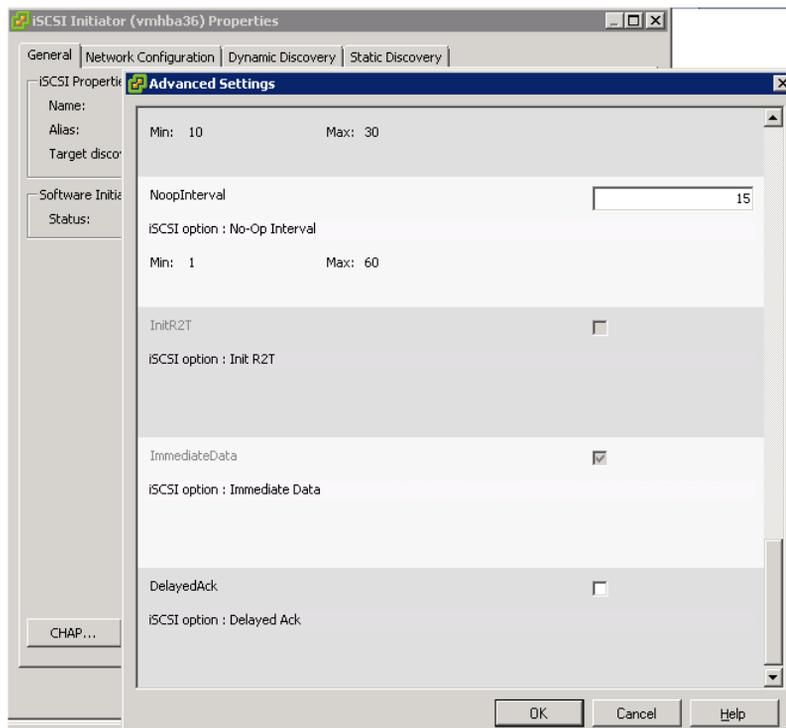


Figure 19 Disable Delayed Acknowledgement setting on storage adapter

VMware provides the following steps to disable Delayed Ack in ESX/ESXi hosts might experience read/write performance issues with certain storage arrays (1002598), available in the VMware Knowledge Base:

1. Log in to the vSphere Client and select the host.
2. Select the **Configuration** tab.
3. Select **Storage Adapters**.
4. Select the iSCSI vmhba to be modified.
5. Click **Properties**.

6. Modify the Delayed ACK setting, using the option that best matches site's needs:
 - a. Modify the Delayed ACK setting on a discovery address (recommended):
 1. On a discovery address, select the **Dynamic Discovery** tab.
 2. Select the **Server Address** tab.
 3. Click **Settings**.
 4. Click **Advanced**.

OR
 - b. Modify the Delayed ACK setting on a specific target:
 1. Select the **Static Discovery** tab.
 2. Select the target.
 3. Click **Settings**.
 4. Click **Advanced**.

OR
 - c. Modify the Delayed ACK setting globally:
 1. Select the **General** tab.
 2. Click **Advanced**.
7. In the **Advanced Settings** dialog box, scroll down to the **Delayed ACK** setting.
8. Clear **Inherit From parent**.
9. Clear **Delayed Ack**.
10. Reboot the host.

Provisioning VNX Storage for vSphere

VNX storage is presented to ESXi hosts in two forms: NFS exported file systems or SCSI LUNs. While NFS file systems are only used as vSphere datastores, LUNs can be formatted for datastore use or assigned to a virtual machine as a RDM virtual disk.

RDM disks are assigned directly to a virtual machine without VMFS formatting. The VMkernel generates a vmdk mapping file for the RDM with LUN information including the unique device id. The virtual machine issues I/Os directly to the VNX LUN using the UUID. RDMs reduce file system overhead and device contention that can be introduced when multiple virtual machines share a VMFS volume.

EMC provides vCenter integration tools to automate and simplify storage device and datastore creation using EMC Unified Storage Management plug-in.

Creating an NFS datastore using EMC Unified Storage Management

Use these steps to configure VNX NFS file systems for vSphere:

1. Create a VNX file system.
2. Export the file system to the ESXi host through VSI or Unisphere.
3. Add the file system as an NFS datastore in ESXi.

These steps can be completed manually using Unisphere or completed through the Unified Storage Management plug-in described here and seen in [Figure 20 on page 51](#).

To provision an NFS datastore through USM:

1. From the vSphere Client right-click on a host or cluster object.

Note: If you choose a cluster, folder, or data center, all ESXi hosts within the object are attached to the newly provisioned storage.

2. Select **EMC > Unified Storage**.
3. Select **Provision Storage**. The **Provision Storage** wizard appears.
4. Select **Network File System** and then click **Next**.

5. In the **Storage System** table, select a VNX. If a VNX does not appear in the Storage System table, click **Add**. The **Add Credentials** wizard appears. Add the VNX storage system.
6. In the **Datastore Name** field, type the datastore name, and then click **Next**.
7. In the **Data Mover Name** list box, select a Data Mover.
8. In the **Data Mover Interfaces** list box, select a Data Mover interface, and then click **Next**.
9. Select **Create New NFS Export** and then click **Next**.

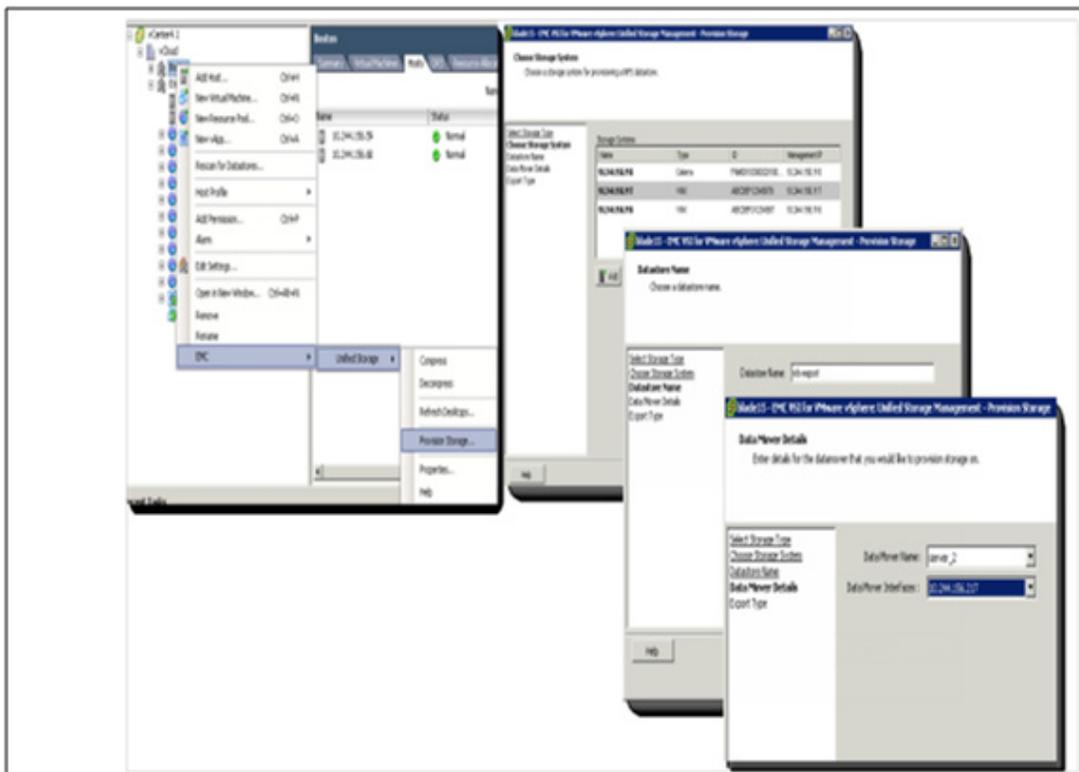


Figure 20 File storage provisioning with USM

- In the **Storage Pool** list box, select a storage pool.

Note: The user sees all available storage within the storage pool. Ensure that the storage pool selected is designated by the storage administrator for use by VMware vSphere.

- In the **Initial Capacity** field, type an initial capacity for the NFS export and select the unit of measurement from the list box at the right.
- If required, select **Thin Enabled** to indicate the new file systems are thinly provisioned.

Note: When a new NFS datastore is created with EMC VSI, Thin Provisioning, and Automatic File system extension are automatically enabled. On the **New NFS Export** window, type the values for the desired initial capacity and maximum capacity of the datastore.

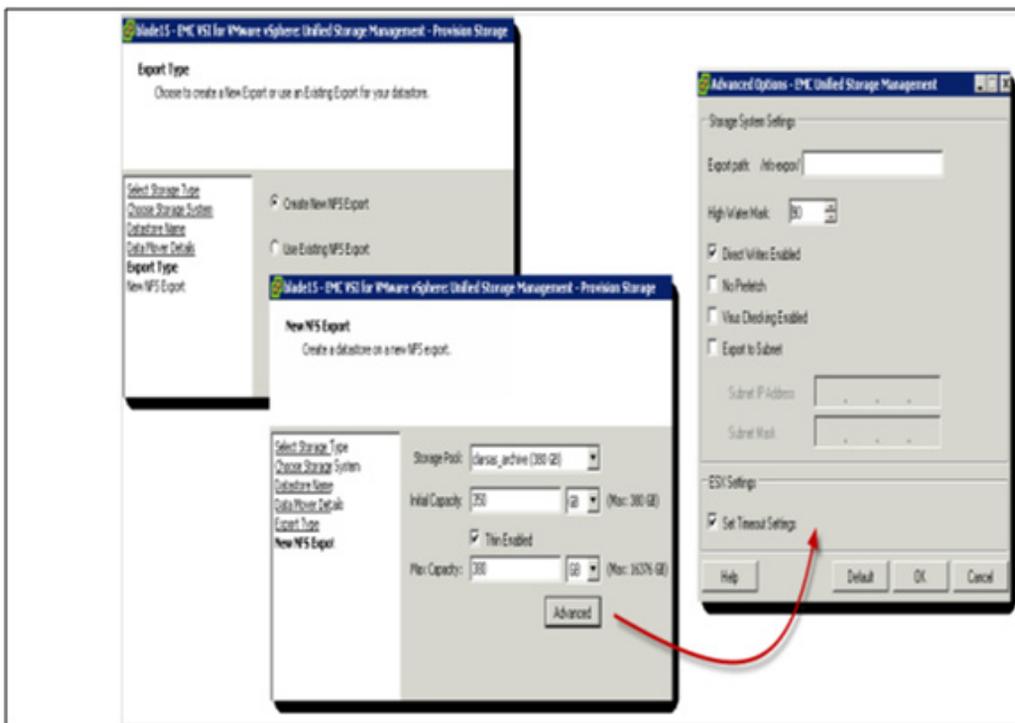


Figure 21 Creating a new NFS datastore with USM

13. If Virtual Provisioning is enabled for the file system, the maximum capacity is required. [Figure 21 on page 52](#) shows an initial capacity entered in the **Max Capacity** field for the NFS export. Select the unit of measurement from the list box to the right.
14. Click **Advanced**. The **Advanced Options** dialog box appears.

The following settings are important for optimal VNX with VMware vSphere performance:

- **High Water Mark** — Specifies the percentage of consumed file system space at which VNX initiates automatic file system extension. Acceptable values are 50 to 99. (The default is 90 percent.)
 - **Direct Writes** — Enhances write performance to the VNX file system. This mechanism enables well-formed NFS writes to bypass the Data Mover cache. The Direct Writes mechanism is designed to improve the performance of applications with many connections to a large file, such as virtual disk files. When replication is used, Direct Writes are enabled on the secondary file system as well.
15. Review the settings, click **OK**, and then click **Finish**.

Provisioning block storage for VMFS datastores and RDM volumes

The following tasks are required to add a VMFS datastore to a vSphere environment:

- ◆ LUN creation
- ◆ LUN unmasking
- ◆ Host rescan
- ◆ VMFS datastore creation

The USM feature of VSI:

- ◆ Offers an integrated workflow to automate the manual provisioning tasks listed above.
- ◆ Allows the administrator to create one or more VMFS volumes and ensures that each volume is correctly aligned on 64 KB boundaries.
- ◆ Performs LUN creation and assignment without formatting so the LUN can be surfaced to a virtual machine as an RDM disk.

After USM is installed, right-click a vSphere object, such as a host, cluster, folder, or datacenter in vCenter:

Note: If you choose a cluster, folder, or data center, all ESXi hosts within the object are granted access to the newly provisioned storage.

1. Select **EMC > Unified Storage**.
2. Select **Provision Storage**. The **Provision Storage** wizard appears as shown in [Figure 22](#).
3. Select **Disk/LUN**, and then click **Next**.

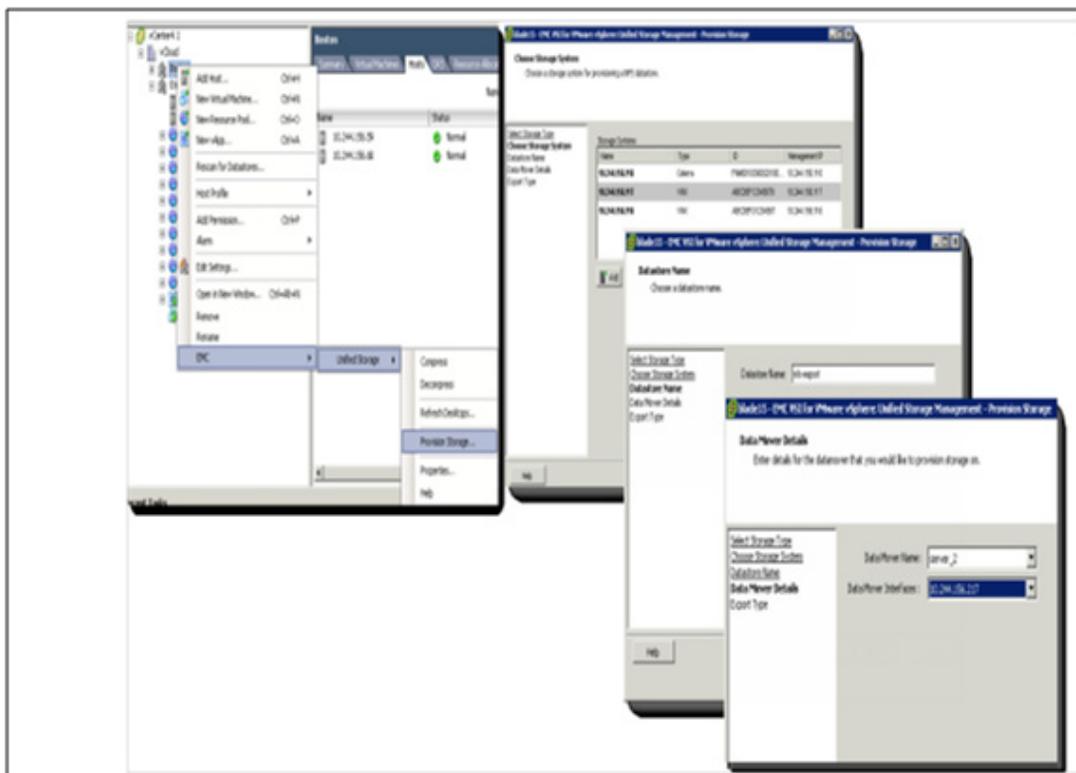


Figure 22 File storage provisioning with USM

4. In the **Storage System** table, select a VNX. If a VNX does not appear in the **Storage Array** table, click **Add**. The **Add Credentials** wizard appears. Add the VNX storage system.

5. Select the storage pool or RAID group on which you want to provision the new LUN and then click **Next**.
6. Select the datastore volume format as **VMFS-3** or **VMFS-5**, and then click **Next**.
7. Select **VMFS Datastore** or **RDM Volume**.
8. Select a SP to own the new LUN and select **Auto Assignment Enabled**. Click **Next**.

Note: Install and correctly configure failover software for failover of block storage.

Note: Unlike VMFS datastores, RDM LUNs are bound to a single virtual machine and cannot be shared across multiple virtual machines, unless clustering is established at the virtual machine level. Use VMFS datastores unless a one-to-one mapping between physical and virtual storage is required.

9. For VMFS datastores, complete the following steps:
 - In the **Datastore Name** field, type a name for the datastore.
 - In the **Maximum File Size** list box, select a maximum file size.
10. In the **LUN ID** list box, select a LUN number.
11. In the **Default Owner** list box, select the SP that will own the new LUN.

12. In the **Capacity** field, type an initial capacity for the LUN and select the unit of measurement from the list box to the right. [Figure 23](#) illustrates this action.

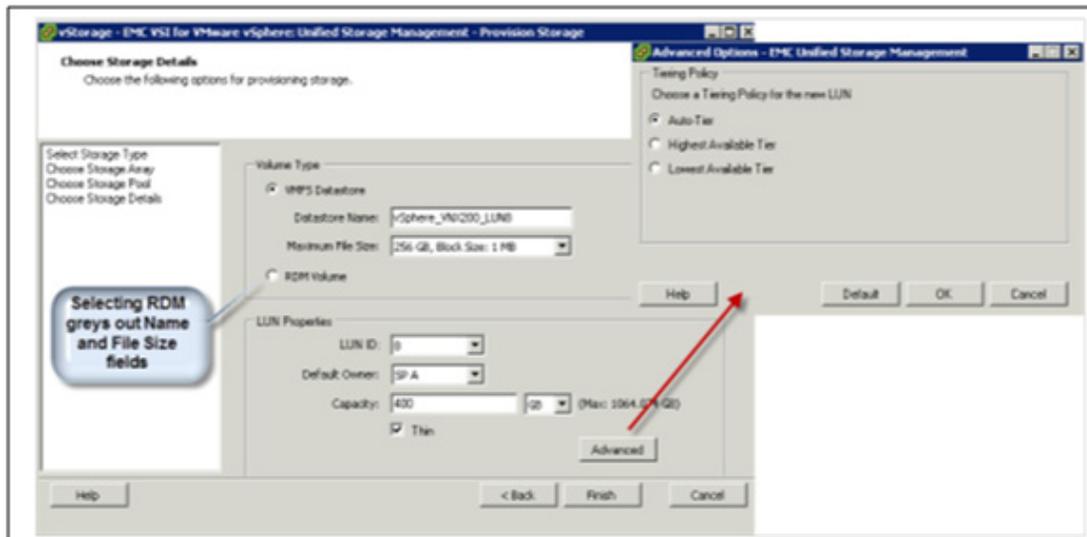


Figure 23 Creating a new VMFS datastore with USM

13. Click **Advanced** to configure the VNX FAST VP policy settings for the LUN. There are three tiering policy options:
- **Auto-Tier** — Distributes the initial data placement across all drive types in the pool to maximize tier usage for the LUN. Subsequent data relocation is based on LUN performance statistics to ensure data is relocated among available tiers according to I/O activity.
 - **Highest Available Tier** — Sets the preferred tier for initial data placement and subsequent data relocation (if applicable) to the highest performing disk drives with available space.
 - **Lowest Available Tier** — Sets the preferred tier for initial data placement and subsequent data relocation (if applicable) to the most cost-effective disk drives with available space.

14. Click **Finish**.

When these steps are complete, USM does the following tasks:

- Creates a LUN in the selected storage pool.
- Assigns the LUN to the designated SP.
- Adds the LUN to the storage group associated with the selected ESXi hosts, making it visible to the hosts.
- Creates the VMFS datastore on the newly created LUN if VMFS is chosen.

15. Select **Configuration > Storage** to see the newly provisioned storage.

Unified storage considerations

A good storage configuration starts with a plan. A proper plan makes configuration easier and documented configuration plans provide a useful reference for validation and support.

The recommendations in this section provide general guidance. Specific configuration suggestions are driven by the actual workload.

The best way to design a storage environment is to understand the requirements. Begin storage planning with an assessment of the application requirements. There are three primary factors that determine the storage configuration:

- ◆ Required throughput measured in IOPS or bandwidth in MB/s
- ◆ Response time or latency measured in milliseconds
- ◆ Storage capacity

Understand the application profile and response time requirements, and translate them into storage resource requirements.

Datastore virtual machine density

With vSphere support for VAAI and Storage I/O Control (SIOC) many of the historical factors that limited virtual machine scalability in a datastore are alleviated. vSphere 5 also added features such as SDRS to balance virtual machine workloads across storage resources.

VNX includes VAAI support, multiple classes of storage devices, and support for an increased number of storage devices, ports, and LUNs.

The right number of virtual machines to add to the datastore is determined by I/O workload, response time, and capacity.

Depending on the configuration, VNX LUNs are capable of delivering tens of thousands of I/Os; and EMC has produced results for VDI which illustrate support for hundreds of virtual machines with a datastore with medium (5 IOPS per virtual machine) workloads.

For performance-sensitive environments where ESXi host clusters are generating significant I/O, create multiple LUNs to distribute the I/O across multiple LUN queues.

For non-SIOC environments, the VMkernel serializes and queues I/Os from all virtual machines that use the LUN. The potential exists for a long LUN queue that can result in longer response times.

SIOC alleviates this condition by throttling the LUN queue depth when response times exceed the defined congestion parameter. Enable and configure SIOC based on the recommendations provided in [“Storage I/O Control \(SIOC\)” on page 101](#).

If SIOC is not enabled, this control falls to a number of other ESXi host parameters including, `Disk.SchedNumReqOutstanding` which, by default, limits the number of requests the host sends to a LUN to 32. That value is used to limit the number of requests the host sends and to ensure that no single virtual machine monopolizes the LUN queue.

Expanding a datastore

VMFS supports the use of multiLUN or multiextent volumes. Adding a new extent increases the capacity for a VMFS datastore that grows short on free space.

The use of multiLUN volumes is generally discouraged because it adds unnecessary complexity and management overhead. If the VNX has enough free space, a preferred approach to multiLUN extents is:

- ◆ Extend the LUN and grow the VMFS volume within vSphere.
- ◆ Create a new device and use LUN migration to migrate data to it. This also provides the ability to change the underlying storage type since LUN migration to any LUN of the same or larger capacity is possible.
- ◆ Use Storage DRS™ to create a datastore cluster and allow it to manage virtual machine placement.

Solid state volumes for VNX File

Follow these general configuration recommendations for Flash drives with VNX OE for File:

- ◆ Use Automatic Volume Management (AVM) pools for general NFS datastores.

AVM templates for EFDs are RAID 5 (4+1 or 8+1) and RAID 1/0 (1+1)

- ◆ Create four LUNs per EFD storage pool and distribute LUN ownership among SPs.

Note: This recommendation does not apply to other storage pools.

- ◆ Use Manual Volume Management (MVM) for custom volume configurations not available with AVM.
- ◆ Due to the lack of mechanical head movement in EFDs, striping across LUNs on a RAID group configured from EFDs is supported.

General recommendations for storage sizing and configuration

VNX enables administrators with an understanding of the I/O workload to provide different service levels to virtual machines. This is done primarily through the storage class and advanced LUN capabilities.

If workload details are not available, use the following general guidelines:

- ◆ Allow for overhead in the datastore for snapshots, swap files, and virtual machine clones. Try to limit datastores to 80 percent of their capacity. This enables administrators to quickly allocate space, create VMware snapshots, clone virtual machines, and accommodate virtual machine swap files.
- ◆ A virtual machine boot disk generates a limited number of IOPS. For example, during boot a standard Windows XP desktop generates about 350 IOPS for a period of about 30 seconds. The boot volume can reside on either an NFS or VMFS virtual disk.
- ◆ Do not create more than three virtual machine snapshots, and do not keep them for an extended period of time. Instead use virtual machine clone to get a point-in-time image of a virtual machine to avoid the logging activity within the datastore that results from change tracking.
- ◆ Enable SIOC to control periods of high I/O traffic, and monitor SIOC response times within vSphere. If response times are consistently high, rebalance the virtual machines with VMware vSphere Storage vMotion®, or configure an SDRS cluster to automate redistribution.

- ◆ Use FAST Cache with the appropriate workload. FAST cache is beneficial for random I/O workloads that are frequently accessed. Sequential workloads typically read or write data once during the operation. Sequential data access patterns often require a longer period of time to warm the FAST cache and are better handled by SP read cache.
- ◆ Monitor the amount of data relocated on FAST VP LUNs. If the FAST VP Pools consistently rebalance a large percentage of data, consider increasing the number of disks in the highest tier.

The following recommendations are specific to workload size:

- ◆ Low Workload
 - Virtual desktop environments have relatively low I/O requirements with occasional bursts caused by operations like booting, virus scanning, or logging on in large numbers.
 - Use FAST Cache-enabled LUNs or Host Cache to reduce the impact of I/O bursts within the virtual machines.
 - Use Host Cache on SSD for linked clone VDI environments. Consider the use of Host Cache on EFDs to support virtual swap files.
 - Use RAID 5 FAST VP pools with a combination of SAS and NL-SAS drives for file servers with static files.
 - Medium-size SAS drives, such as the 300 GB, 10k RPM drive, may be appropriate for these virtual machines.
 - Use 1 and 2 TB NL-SAS drives for datastores that are used to store archived data.
 - Use RAID 6 with NL-SAS drives greater than 1 TB.
 - Infrastructure servers, such as DNS Servers, are primarily processor-based with relatively little I/O. Those virtual machines can be stored on NFS or a FAST VP Pool consisting of SAS and NL-SAS drives.
- ◆ Medium Workload
 - Medium DB application workloads are good candidate for SAS datastores. FAST Cache or FAST VP configured with as few as two SSDs provides a good option for heavily used tables within the database. Use a separate RAID 10 datastore for DB log virtual disks.

- ◆ High Workload
 - Applications with hot regions of data benefit from the addition of FAST Cache.
 - Store DB log files on separate virtual disks in RAID 10 VMFS, NFS, or RDM devices.
 - Allocate RAID 10 protected volumes, EFDs, or FAST Cache to enhance the performance of virtual machines that generate high small block, random I/O read workload. Consider dedicated RDM devices for these virtual machines.
 - Use RAID 1/0 LUNs or file systems for virtual machines that are expected to have a write-intensive workload.

Storage multipathing

Multipathing provides two or more I/O paths between a host and a LUN to address two important requirements:

1. Reliability: multiple I/O paths ensure that access to application data is maintained in the event of a component failure.
2. Scalability: hosts can parallelize I/Os across multiple storage adapters to increase efficiency, and ideally balance for optimal storage-resource utilization.

Before moving to ESXi host multipath configuration, we review VNX LUN ownership and ALUA as illustrated in [Figure 24 on page 63](#). VNX storage systems have two storage processors, identified as SP A and SP B. At any given time, a LUN is owned by only one SP.

When a LUN is created, it is assigned to an SP which also becomes the LUN's default owner. Since LUN can only be owned by one SP at a time, the SP that owns the LUN provides the optimal paths to it through all of its front end I/O ports. The peer SP can also satisfy I/O, however, the I/O must traverse an internal bus in order to satisfy the request and it is therefore non-optimal.

Figure 24 illustrates the concept of LUN ownership and I/O paths. When a LUN is owned by SP A, optimal paths are through the I/O ports of SP A.

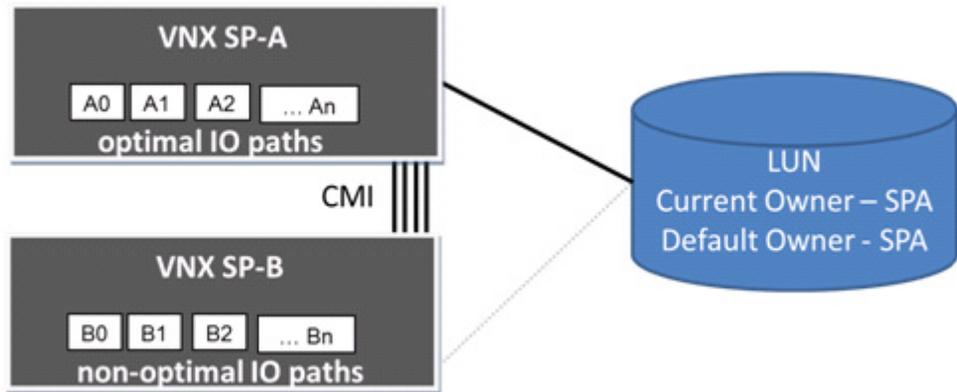


Figure 24 LUN ownership

LUN trespass

VNX provides the ability to transfer LUN during various host and storage system states such as when all paths from a host become unavailable, or when the VNX storage processor is undergoing a software update.

Under these conditions LUN ownership is passed to the peer SP and hosts use that SP to provide optimal I/O to the LUNs. A LUN is considered to be trespassed when its current owner is different from its default owner.

LUN failover modes are introduced in [“Manual initiator registration” on page 36](#). VNX provides multiple failover modes including Active/Standby (mode 1) and Active/Active (ALUA) which is the recommended failover mode for vSphere 4 and later.

When configured in ALUA mode, a host issues an I/O request to either SP and the VNX services them. However, I/O received on the non-owning SP has to traverse an internal system bus in order to service the request.

ESXi hosts are ALUA compliant and they use the optimal paths whenever they are available.

If the optimal paths become unavailable, the host issues a request to the peer SP to transfer ownership of the LUN providing an optimal I/O path to the SP it has access to. When an ESXi host boots, the NMP module performs an inquiry on each discovered LUN and uses the default SP owner to establish the preferred path to the LUN. All of these processes are illustrated in Figure 25.

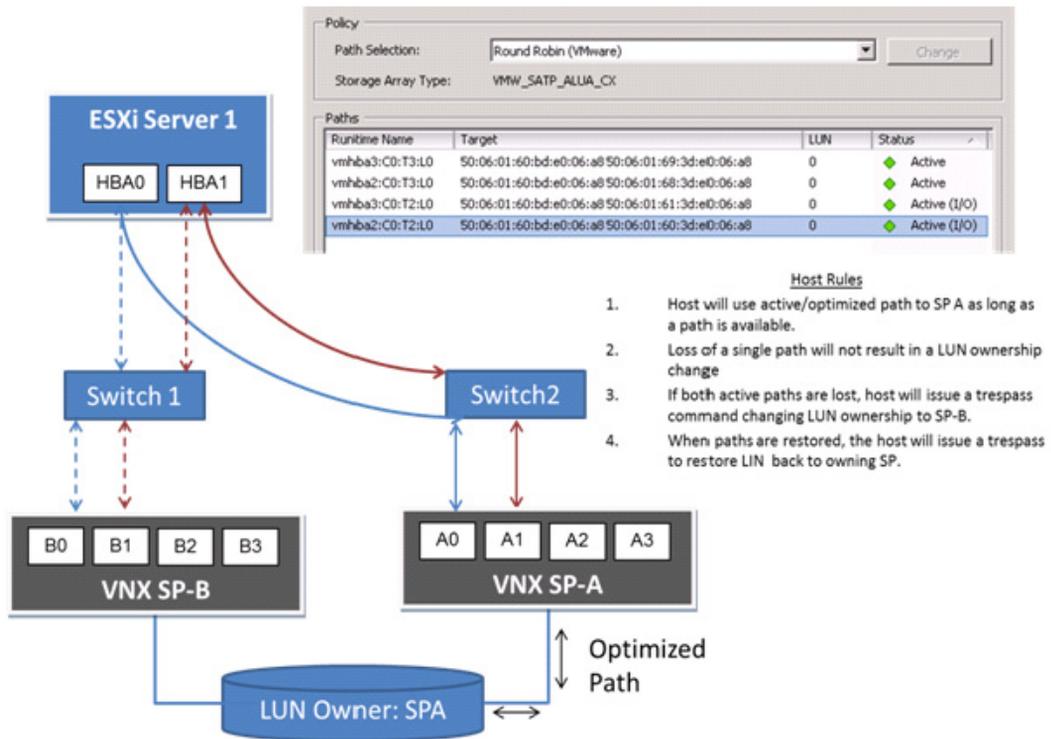


Figure 25 LUN trespass

vSphere Native Multipath

The ESXi VMkernel provides a pluggable storage architecture to support different multipath modules. Figure 26 illustrates the pluggable storage architecture (PSA) used with vSphere.

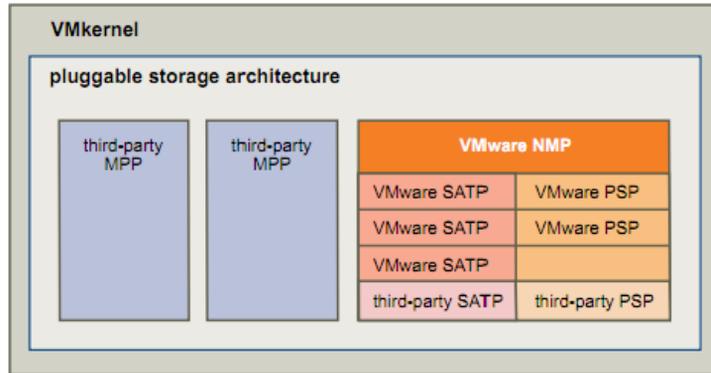


Figure 26 VMkernel pluggable storage architecture

The default module is Native Multipathing Plug-in (NMP) which presents several path configuration options to determine:

- ◆ The Path Selection Plug-in (PSP) used when multiple physical paths exist.
- ◆ Path failure and recovery policy.

NMP provides the framework to discover new LUNs, identify the Storage Array Type Plug-in (SATP), the initiator mode, and LUN properties such as the default Storage Processor that owns the LUN.

NMP uses the SATP to assign a PSP to the LUN. Run the `esxcli storage nmp satp list` command to view the rules.

NMP Path Selection Plug-ins (PSPs)

vSphere has four native path selection plug-ins:

- ◆ **Fixed Path Array Preference (AP)** — The `FIXED_AP` plug-in queries the array for the preferred path and uses that path unless a failure occurs. This PSP has been removed in vSphere 5.0 and later. It is the default vSphere 4.1 PSP for `VMW_SATP_ALUA_CX`.

- ◆ **Fixed Path** — Uses the single preferred (active/optimized) I/O path for the VNX LUN. If the preferred path is unavailable, it uses an alternate path. It reverts to the preferred path when it is restored. It is the default vSphere 5.0 PSP for VMW_SATP_ALUA_CX.
- ◆ **Round Robin** — Uses all active/optimized paths between the host and the LUN. The host sends a fixed number of I/Os down the first active/optimized path, followed by a fixed number of I/Os down each subsequent active/optimized path. Non-optimized paths are not used for I/O, unless all active/optimized paths have failed. It is the default, vSphere 5.1 PSP for VMW_SATP_ALUA_CX.
- ◆ **Most Recently Used (MRU)** — This option is used by all vSphere hosts when the failover mode of the host initiator records is set to one. It uses the first LUN path detected when the host boots. The host continues to use that path as long as it remains available. If a path failure occurs, the host attempts to use another path on the same SP or issues a trespass to the peer SP. It is the default vSphere 5.0 and 5.1 PSP for VMW_SATP_CX.

Each SATP uses a predefined Path Selection Policy (PSP) agreed upon by VMware and the storage Vendor. [Figure 27](#) illustrates the `esxcli` command output to identify the PSPs used for VNX systems in vSphere 5.1.

```

~ # esxcli storage nmp satp list
-----
Name                Default PSP      Description
-----
VMW_SATP_CX         VMW_PSP_MRU     Supports EMC CX that do not use the ALUA protocol
VMW_SATP_ALUA_CX   VMW_PSP_RR      Supports EMC CX that use the ALUA protocol
VMW_SATP_ALUA       VMW_PSP_MRU     Supports non-specific arrays that use the ALUA protocol
VMW_SATP_MSA        VMW_PSP_MRU     Placeholder (plugin not loaded)
VMW_SATP_DEFAULT_AP VMW_PSP_MRU     Placeholder (plugin not loaded)
VMW_SATP_SVC        VMW_PSP_FIXED   Placeholder (plugin not loaded)
VMW_SATP_EQL        VMW_PSP_FIXED   Placeholder (plugin not loaded)
VMW_SATP_INV        VMW_PSP_FIXED   Placeholder (plugin not loaded)
VMW_SATP_EVA        VMW_PSP_FIXED   Placeholder (plugin not loaded)
VMW_SATP_SYMM       VMW_PSP_RR      Placeholder (plugin not loaded)
VMW_SATP_LSI        VMW_PSP_MRU     Placeholder (plugin not loaded)
VMW_SATP_DEFAULT_AA VMW_PSP_FIXED   Supports non-specific active/active arrays
VMW_SATP_LOCAL      VMW_PSP_FIXED   Supports direct attached devices

```

Figure 27 Esxcli command output

ESX 5.1 contains an enhancement to the Round-robin NMP PSP which allows autorestore to the preferred (default) VNX storage processor (SP) when a fabric failure (failed HBA/NIC/CNA, switch, or SP front end port) to that preferred SP is repaired. [Table 1](#) shows the recommended NMP path selection plug-in.

Table 1 Recommended NMP path selection plug-in

ESX Revision	VNX Software Revision	Recommended NMP Path PSP
ESX 5.1	05.31.000.5.726 or later	Round-robin
ESX 4.x	05.31.000.5.726 or later	Round-robin or Fixed
Any release of ESX	Elias MR2 SP3 or earlier	Round-robin or Fixed

VNX array software Elias MR2 SP4 contains an enhancement which allows the Round-robin NMP PSP to autorestore to the preferred (default) VNX storage processor (SP) after the preferred SP reboots (whether due to failure, manually, or as part of an array software upgrade (NDU)).

With ESX 5.1 and VNX OE for Block 05.31.000.5.726 or later, Round-robin is the preferred PSP for VNX LUNs. In this environment, users get the benefit of multiple active/optimized paths for I/O scalability as well as the benefit of autorestore to the preferred SP after any fabric failure or SP reboot.

Use Round-robin when using NMP.

Third-party multipathing - EMC PowerPath Virtual Edition

EMC provides a multipath plug-in called PowerPath Virtual Edition or PowerPath/VE to enhance reliability and I/O efficiency of ESXi environments. PowerPath provides the most comprehensive multipathing solution for vSphere environments.

PowerPath/VE is supported for all SCSI configurations and offers the following benefits:

- ◆ Performs adaptive load-balancing and path optimization.
- ◆ Performs proactive monitoring of I/O path for health status.
- ◆ Contains an intuitive CLI that provides end-to-end viewing and reporting of the host storage resources, including HBAs.
- ◆ Applies policy changes at the host level.

- ◆ Uses autorestore to restore LUNs to the optimal SP after NDU or environmental failure to ensure load balancing and performance.
- ◆ Provides the ability to balance queues on the basis of queue depth and block size.

Note: PowerPath provides the most robust functionality and is the recommended multipathing option for VNX.

VSI: Path Management

The Path Management feature is an extension to the VSI that simplifies the LUN path policy configuration. [Figure 28](#) shows how administrators assign global NMP or PowerPath path configuration preferences to VNX LUNs and maintain consistent policies across all hosts in a virtual data center.

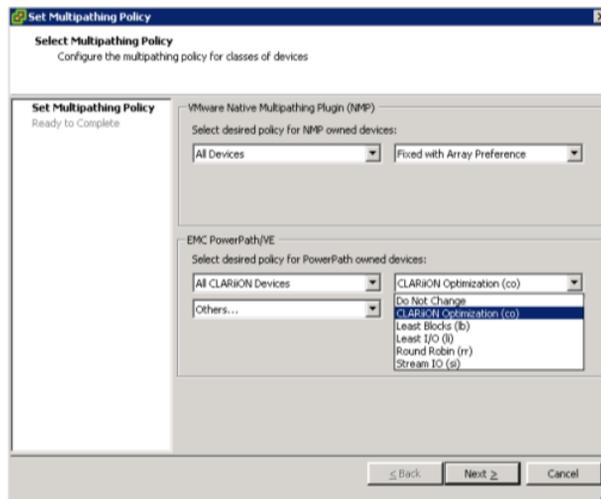


Figure 28 VSI Path Management feature

The Storage Viewer feature VSI or vCenter device properties are used to verify the existing multipath policy for each LUN (Datastore). [Figure 29 on page 69](#) illustrates the LUNs properties page within Storage Viewer. This view includes pluggable storage architecture that owns the LUN and the current PSPs.

Use the VSI Path Management feature to modify it if required.

Note: Individual path modification must be done through vCenter or with the vSphere command line utilities.

Runtime Name	Paths	Capacity	Owner	Policy	Used By	Storage Group	Pool	FAST VP Poli...
vmhba3:C0:T2:L13	4	5.00 GB	NMP	Fixed	VSI			
vmhba3:C0:T2:L5	4	2.00 TB	NMP	Fixed	VNX201-LUN4			
vmhba3:C0:T2:L3	4	2.00 TB	NMP	Fixed	VNX201-LUN3			
vmhba3:C0:T2:L2	4	1.02 TB	NMP	Fixed	VNX201-LUN2			
vmhba3:C0:T2:L1	4	1.02 TB	NMP	Fixed	VNX201-LUN1			
vmhba3:C0:T2:L10	4	1.02 TB	NMP	Fixed	VNX201-LUN0			
vmhba3:C0:T2:L9	4	228.00 GB	NMP	Fixed	SSD-LUN9			
vmhba3:C0:T2:L8	4	274.00 GB	NMP	Fixed	SSD-LUN8			
vmhba3:C0:T2:L7	4	274.00 GB	NMP	Fixed	SSD-LUN7			
vmhba3:C0:T2:L6	4	228.00 GB	NMP	Fixed	SSD-LUN6			
vmhba3:C0:T2:L4	4	1.00 GB	NMP	Fixed	SMI-5			
vmhba4:C0:T2:L0	2	2.00 GB	NMP	Fixed	SMI-5			
vmhba3:C0:T0:L1	4	700.00 GB	NMP	Fixed	CX014-500			
vmhba3:C0:T2:L12	4	20.00 GB	NMP	Fixed				
vmhba3:C0:T0:L0	4	1000.00 ...	NMP	Fixed				
vmhba0:C0:T0:L0	1	68.37 GB	NMP	Fixed				
vmhba3:C0:T2:L10	4	200.00 GB	NMP	Fixed				

Figure 29 Storage Viewer LUNs view

Multipathing considerations - NFS

ESXi hosts access NFS servers using NFS version 3 (NFSv3). The NFSv3 protocol is limited to a single TCP session per network link. Therefore, the only way to balance the I/O load for NFS is to use the physical layer to mount the NFS file system on different ESXi source interfaces, and different destination interfaces on the Data Mover. Configure multiple Data Mover interfaces and distribute NFS TCP sessions between different source and destination network interfaces. The default number of NFS mounts in ESXi4 and ESXi5 is eight and 64 respectively. The number reaches a maximum value of 64 after the NFS.MaxVolumes parameter on the host is modified. [Figure 30 on page 71](#) illustrates the recommended configuration for high availability and load balancing. Use the following guidelines to achieve high availability and load balancing for NFS:

- ◆ Ensure there are no single points of failure at the physical network layer (NIC ports, switch ports, physical network switches, and VNX Data Mover network ports).
- ◆ Balance the workload among all available I/O paths.
- ◆ Data Mover network ports, connections to switch - configure Link Aggregation on VNX Data Movers and network switches for fault tolerance of network ports. LACP supports load balancing among multiple network paths. Configure the Data Mover and ESXi switch ports for static LACP.

Note: When a Cisco Nexus 1000v pluggable virtual switch is used on the ESXi hosts, configure dynamic LACP for the ESXi and Data Mover NIC ports.

- ◆ **ESXi NIC ports** — NIC teaming provides physical network fault tolerance and load balancing for ESXi hosts. Set the NIC teaming load balancing policy for the virtual switch to **Route based on IP hash** for LACP configurations.
- ◆ **Physical network switch** — Use multiple switches and network paths for physical-layer fault tolerance. Configure each Data Mover and ESXi host to use both switches. If the switch supports Multichassis Link Aggregation, configure it to span the switches and offer redundant port termination for each I/O path from the Data Mover and ESXi host.

Note: Use Fail-Safe Network on the VNX Data Movers with switches that do not support Multichassis Link Aggregation technology.

Configure multiple network paths for NFS datastores

This section describes how to build the configuration shown in [Figure 30 on page 71](#).

Create a single LACP network device for the Data Mover through the Unisphere Management UI. LACP devices use two physical network interfaces on the Data Mover, and two IP addresses on the same subnet.

Complete the following steps to create the multipath NFS configuration.

- ◆ Steps 1 through 7 are performed in EMC Unisphere.
- ◆ Steps 8 through 14 are performed in the vSphere Client.

- ◆ Ensure that Link Aggregation is enabled on the switch ports, VNX Data Mover, and ESXi network interfaces.

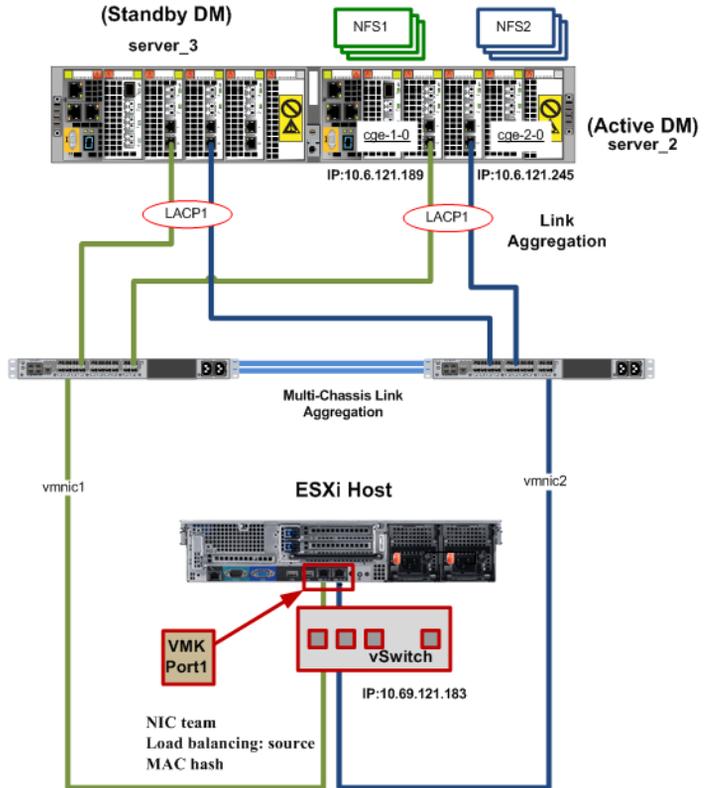


Figure 30 Elements of a multipathing configuration for NFS

Log in to Unisphere to complete the following steps:

1. Select the VNX system to manage. Select **Settings > Network > Settings For File**. The **Settings For File** window appears as shown in [Figure 31](#).

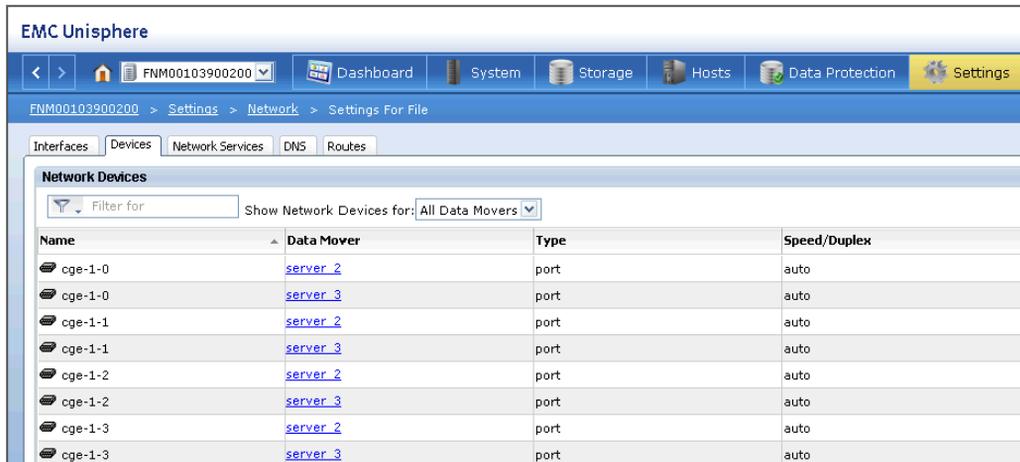


Figure 31 Unisphere interface

2. Select the **Devices** tab, and then click **Create**. The **Network Device** dialog box appears.
 - a. In the **Device Name** field, specify a name for the LACP device.
 - b. In the **Type** field, select **Link Aggregation**.
 - c. In the **10/100/1000/10000 ports** field, select two unused Data Mover ports.
 - d. Click **OK** to create the LACP device.
3. From the **Settings For File** window, select the **Interfaces** tab.
4. Click **Create** to create a new network interface.

Data Mover:	server_2
Device Name:	LACP1
Address:	10.244.156.102
Name:	DM2_LACP1
Netmask:	255.255.255.0
Broadcast Address:	10.244.156.255
MTU:	
VLAN ID:	
<input type="button" value="OK"/> <input type="button" value="Apply"/> <input type="button" value="Cancel"/> <input type="button" value="Help"/>	

Figure 32 Data Mover link aggregation for NFS server

5. Complete the following steps:
 - a. In the **Device Name** list box, select the LACP device that was created in Step 2.
 - b. Enter the IP address for the first Data Mover LACP interface.
 - c. In [Figure 32](#) the IP address is set to 10.244.156.102 and the interface name is set to DM2_LACP1.
6. Click **Apply** to create the first network interface and keep the **Create Network Interface** window open.
7. In the **Create Network Interface** window, type the details for the second network interface. This information is identical to the information provided in Step 5 with the exception of the IP address.
 - a. Type the IP address for the second LACP connection.
 - b. Click **OK** to create the second network interface.
 - c. Access the vSphere Client and complete steps 7 through 12 for each ESXi host.

8. Create a vSwitch for all the new NFS datastores in this configuration.
9. Create a single VMkernel port connection in the new vSwitch. Add two physical NICs to it and assign an IP address for the VMkernel on the same subnet as the two Data Mover network interfaces.

In [Figure 32 on page 73](#) the VMkernel IP address is set to 10.244.156.183, with physical NICs VMnic0 and VMnic1 connected to it.

10. Click **Properties**. The **vSwitch1 Properties** dialog box, [Figure 33](#), appears.

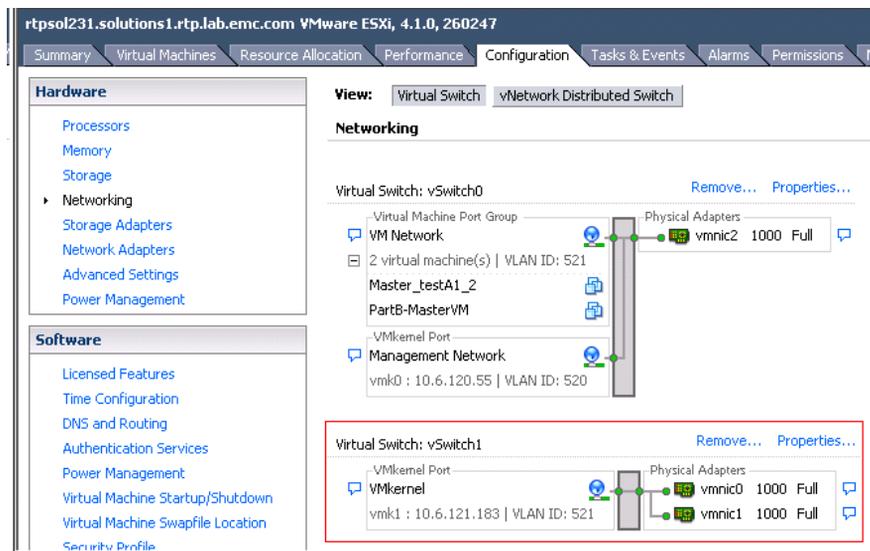


Figure 33 vSphere networking configuration

11. Select **vSwitch**, and then click **Edit**. The **VMkernel Properties** window appears as shown in [Figure 34 on page 75](#).

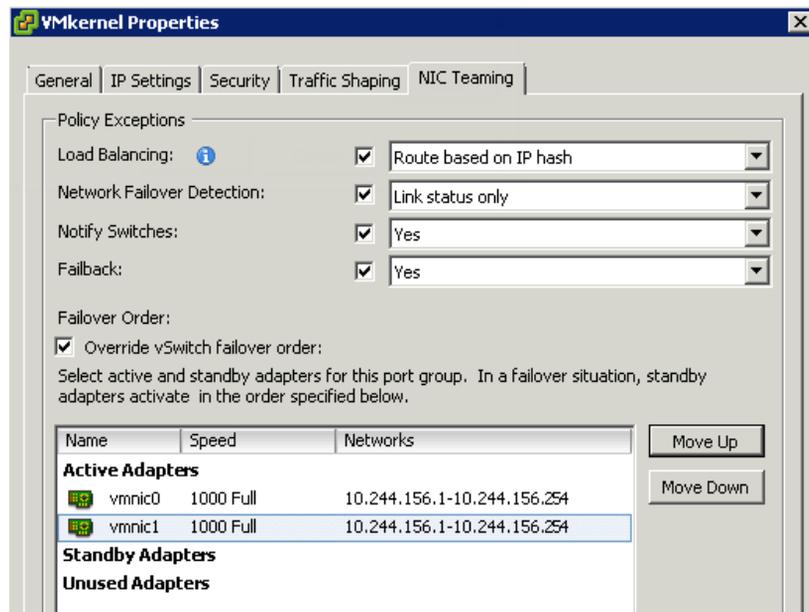


Figure 34 VMkernel Properties window

12. Select the **NIC Teaming** tab, and in the **Load Balancing** list box, select **Route based on IP hash**.

Note: The two VMnics are listed in the **Active Adapters** list for the NIC team. If the corresponding switch ports are enabled for EtherChannel, network traffic is statically balanced by using a unique hash value derived from the source and destination IP addresses. As a result, the host will distribute network packets between the VMkernel and Data Mover network interfaces to provide distribution of I/Os among the available network paths, and yield a higher throughput.

13. Use VSI USM to provision NFS datastores.
 - a. In the **Data Mover Name** list box, select the primary **Data Mover**.
 - b. In the **Data Mover Interface** list box, select the IP address of the first network interface that was created.
14. Create and distribute the virtual machines evenly across datastores, or use Storage DRS to automate the virtual machine placement within the datastore cluster.

vSphere storage configuration

Up to this point, this chapter has presented the configuration options for host connectivity and automated storage provisioning using EMC's management features. This section of the paper describes vSphere storage related features and identifies considerations for using them with VNX storage. Some of these features are version specific and are annotated to highlight that.

Dead space reclamation

Release 5.0 U1 introduced a feature called dead space reclamation to reclaim disk space when a file or virtual machine is deleted within, or moved off of a thin VMFS datastore.

Unmap works with VMFS datastores provisioned using VNX thin LUNs.

As a normal course of business virtual machines, virtual disks, and files are added, removed and migrated within VMFS datastores. After a file is deleted or migrated the ESXi host de-allocates the disk blocks within the VMFS file system. However, those blocks remain allocated within the VNX LUN and are reusable only by virtual machines that share the same VMFS datastore.

Dead space reclamation provides a manual method to instruct the Thin LUN to release the allocated blocks and thus return the unused space to the global storage pool so that space can be used by other datastores or RDM LUNs.

In the current implementation, the unmap process is initiated through the `vmkfstools` command by specifying the `-y` option and a numerical value that represents the percentage of space you want to reclaim. Unmap operates at the datastore level so the command needs to be run from within each datastore. The operational tasks are to change directory to the datastore where unmap will be performed, and then execute commands such as `"vmkfstools -y 90"` to reduce the space by 90%. When the command is initiated, the ESXi host creates a balloon file within the datastore and issues SCSI UNMAP commands (0x42) to the VNX SCSI target to release freed blocks within the thin LUN. The space freed within the Thin LUN results in slices being returned to the Storage Pool. Figure X provides an example of how unmap is used to reclaim space from a datastore configured on a VNX Thin LUN.

There is also a potential performance impact to other systems and it should be run during a maintenance period.

In Figure 35, two LUNs, named Thin and Thick to identify the device type, exist within a VNX storage pool (Pool 0). The thick provisioned LUN is 100 GB in size and the disk slices associated with that LUN are persistently reserved within the pool when the LUN is created. Those blocks are not released to the pool until the LUN is deleted.

The second LUN, called thin, is 300 GB in size. However, since it is thin provisioned, only 3 GB of metadata is allocated from the storage pool. After formatting the VMFS volume the total slice allocation within the pool is 114 GB.

In step 2 of the figure a virtual machine with a virtual disk of 40 GB is created and stored on the Thin LUN resulting in 140 GB of allocated space within the Pool 0..

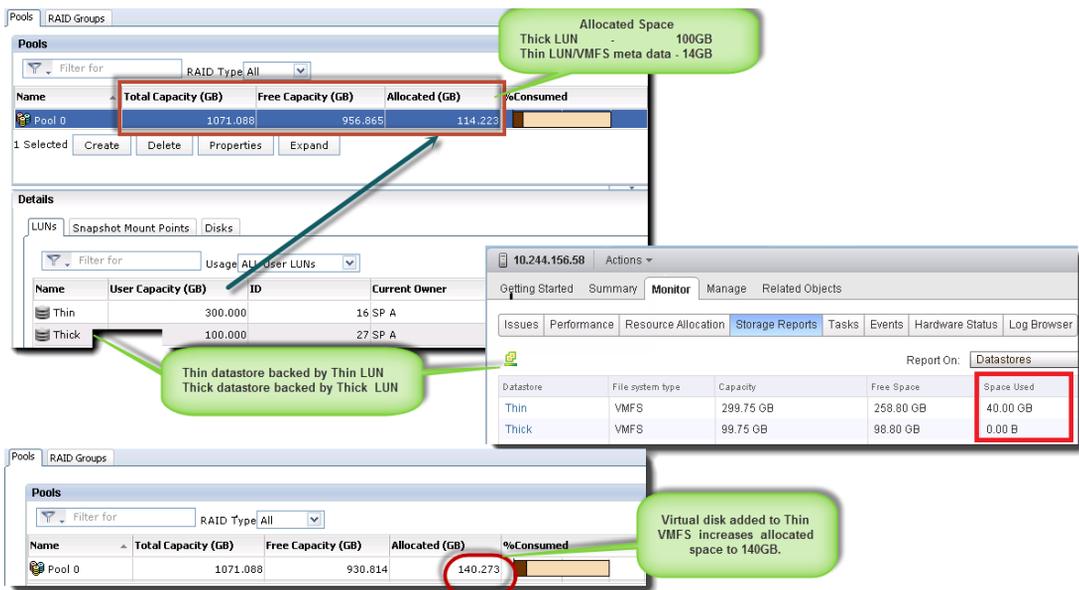


Figure 35 Virtual machine configured on a Thick LUN

The top half of Figure 36 on page 79 illustrates what happens when the virtual machine is migrated from the thin LUN datastore to the thick LUN datastore. The space within each datastore changes, but the thin pool utilization remains the same.

In the example the ESX host command line was used to change directory to the /vmfs/volumes/Thin directory where the vmkfstools command was run with the (-y) and 99 percent arguments. The result is seen in the bottom half of the screen, illustrating that the space was restored to the VNX pool.

The screenshot shows the vSphere Storage Reports interface. The top section displays a table of datastores:

Datastore	File system type	Capacity	Free Space	Space Used
Thin	VMFS	299.75 GB	298.80 GB	0.00 B
Thick	VMFS	99.75 GB	57.71 GB	41.09 GB

A callout bubble points to the 'Thin' row with the text: "VM has been migrated to Thick LUN".

The middle section shows the 'Pools' view for 'Pool 0':

Name	Total Capacity (GB)	Free Capacity (GB)	Allocated (GB)	%Consumed
Pool 0	1071.088	930.814	140.273	

A callout bubble points to the 'Free Capacity' column with the text: "Allocated space is unchanged".

The bottom section shows a terminal window with the following command and output:

```

~ # cd /vmfs/volumes/Thin
/vmfs/volumes/50a27784-615b1f90-81f3-00219bcad60c # vmkfstools -y 99
Attempting to reclaim 99% of free capacity 298.8 GB (295.8 GB) on VMFS-5 file system 'Thin' with max file size 64 TB.
Creating file .vmfsBalloonihUvmJ of size 295.8 GB to reclaim free blocks.
Done.
/vmfs/volumes/50a27784-615b1f90-81f3-00219bcad60c #
    
```

A callout bubble points to the terminal with the text: "Run vmkfstools to reclaim the space".

The bottom section shows the 'Pools' view for 'Pool 0' after the command:

Name	Total Capacity (GB)	Free Capacity (GB)	Allocated (GB)	%Consumed
Pool 0	1071.088	956.865	114.223	

A callout bubble points to the 'Free Capacity' column with the text: "Pool consumption is back to the original size 40GB space for virtual disk".

Figure 36 Virtual machine migrated to a Thin LUN

Use the vmkfstools command to reclaim unused space within ESXi datastores provisioned from Thin LUNs. This example illustrates how to identify the space.

Virtual Machine File System 5 (VMFS-5)

vSphere 5.0 and later include an update to the VMFS file system called VMFS version 5. VMFS-5 provides improved scalability, and interoperability with storage systems with the following properties:

- ◆ Support for larger file systems and volumes.
- ◆ Fixed block size of 1 MB.
 - Eliminates the 1, 2, 4 MB block sizes required to support larger files in previous releases.
 - Supports double and triple block indirect pointers for larger files.
- ◆ Atomic Test and Set (ATS) also known as Hardware Accelerated Locking is enabled for all SCSI devices.
 - ESXi hosts always attempts to use ATS on VMFS-5 volumes.
 - If an ATS operation fails, the LUN processes the request using SCSI-2 commands. Subsequent requests will revert to ATS.
 - Small block allocation configurations.

New datastores created vSphere 5 defaults to VMFS-5 volumes when creating a new datastore; however, VMFS-3 is still available. Although the list above may not seem extensive, VMFS-5 volumes should be used for all new datastores.

The upgrade option available through `vmkfstools` helps upgrade existing VMFS-3 file systems to VMFS-5. However, upgraded VMFS-3 file systems do not take advantage of all VMFS-5 features.

The following explains the limitations of the upgraded VMFS-3 datastores:

- ◆ Use the VMFS-3 block sizes, which may be larger than the unified 1 MB file block size.

Note: VAAI Full Copy operations are not supported between datastore volumes that are formatted with different block sizes.

- ◆ Use 64 KB sub-blocks instead of the new 8 K sub-blocks.
- ◆ Have a file limit of 30,720 instead of the file limit of > 100,000 for new VMFS-5 datastores.

A better alternative to performing a VMFS-3 upgrade to VMFS-5 is to create a new VMFS-5 volume and migrate the virtual machines with Storage vMotion. Use VNX thin provisioned LUNs in conjunction with VMware thin virtual disks to reduce the amount of storage space required to perform this task.

vStorage API for Array Integration (VAAI)

VAAI storage integration improves ESXi host resource utilization by offloading storage-related tasks to the VNX. The storage system processes select storage tasks for the host, freeing resources for application processing and other tasks.

Storage vMotion is a core feature of Storage DRS in vSphere 5 and a good example of the use of VAAI. During a Storage vMotion task, the ESXi host sends SCSI extended copy (XCOPY) commands containing the source and destination LUN on the VNX. The VNX storage processor copies the virtual disk to the target device. With VNX OE for Block 5.32, this operation finishes much faster using significantly fewer CPU, memory, and SAN fabric I/O resource than is required to perform the task on the host.

The primary VAAI functions are:

- ◆ **Hardware Accelerated Zeroing** — Known as Block Zero, it uses SCSI WRITE SAME commands to perform bulk write operations when all blocks contain the same data, zeros. From a practical standpoint, it is used to zero out newly created virtual disks that contain unallocated blocks. When a new flat (eagerzeroedthick) VMDK is created, the feature instantaneously creates a file with the proper allocations and initializes the remaining blocks to zero.
- ◆ **Hardware Accelerated Locking** — Known as Atomic Test and Set (ATS), it alleviates VMFS contention resulting from metadata operations such as virtual machine creation, virtual machine boot, modification to virtual machine property settings. ATS provides extent level locking to the VNX LUN, which enables metadata updates without locking the entire device. ATS alleviates contention during boot storms and other vSphere operations that require considerable metadata updates.

- ◆ **Hardware Accelerated Copy** — Known as Full Copy, it uses SCSI XCOPY commands to perform block movements within the array. The primitive is initiated by vSphere Clone, Storage vMotion (Storage DRS), and Deploy Virtual Machine from Template tasks.
- ◆ **NFS Clone Offload** — Offloads ESXi clone operations to the VNX Data Mover. This produces results similar to those for Hardware Accelerated Copy. The ESXi host achieves a reduction in CPU and network resource utilization.

VAAI improves host efficiency by using host CPU, memory, and SAN to satisfy application servers. They enable dense datastore configurations with improved operational value.

With the exception of NFS, ESXi hosts use these features by default. Use these features with VNX storage.

EMC NAS Plug-in for NFS

vSphere 5.0 and later provides support for VAAI operations on NFS datastores. Working in conjunction with storage vendors like EMC, VMware has integrated VAAI with VNX through a software interface installed onto each ESXi host. With this software or host plug-in installed, ESXi hosts can leverage the VNX X-Blade to perform the tasks listed in [Table 2](#).

VMware View 5.1 provides the ability to deploy new virtual machines using VNX Fast Clones. The View 5.1 product leverages the NFS plug-in to create thin virtual disks within the virtual machine directory on the NFS datastore.

[Table 2](#) includes a summary of the NFS VAAI features and the supported VNX OE for File version.

Table 2 NFS VAAI features

Feature	VNX OE
Full Clone	5.31 and later
Extended Stats	5.31 and later
Space Reservation	5.31 and later
Snap of Snap (Tech Preview in View 5.1)	7.31 and later

Virtual machine clones

VAAI for NFS leverages the VNX Data Mover to create thin fast clone and thick full clone replicas of virtual machines on the NFS datastores. Thin clones are created instantaneously using a few file system blocks to link to the source virtual machine. This option preserves space by using a single image. All of the virtual disk blocks that make up the virtual machine exist as references to the source virtual machine. The exception is in modified data which allocates additional blocks within the file system as needed.

When creating full clones, create an exact replica of an existing virtual machine and use the same amount of storage as the source virtual machine. The VNX File OE provides the ability to create thin fast clones or thick provisioned full clones. Full clones are copied using the Data Mover resources.

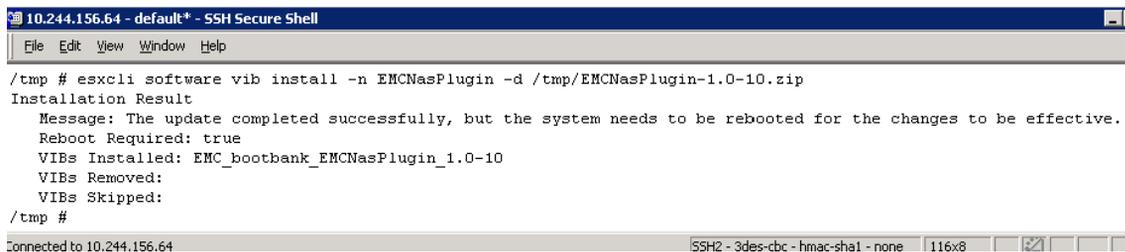
Nested Snaps (Snaps-of Snaps)

VNX OE for File 7.31 includes new NFS capabilities when creating virtual machine clones.

EMC NAS Plug-in Installation

In order for the VMkernel to use the NFS VAAI features a software module needs to be installed on each host. The software bundle is provided as a VMware installation bundle that is installed through the command line of the ESXi host or through VMware vCenter Update Manager.

Figure 37 illustrates the `esxcli` command issued to install the VIB after copying it to the `/tmp` directory of the ESXi host. Place each host in maintenance mode before installing the plug-in.



```

10.244.156.64 - default* - SSH Secure Shell
File Edit View Window Help
/tmp # esxcli software vib install -n EMCNasPlugin -d /tmp/EMCNasPlugin-1.0-10.zip
Installation Result
  Message: The update completed successfully, but the system needs to be rebooted for the changes to be effective.
  Reboot Required: true
  VIBs Installed: EMC_bootbank_EMCNasPlugin_1.0-10
  VIBs Removed:
  VIBs Skipped:
/tmp #
connected to 10.244.156.64  SSH2 - 3des-cbc - hmac-sha1 - none  116x8

```

Figure 37 Plug-in Installation

vCenter displays the **Hardware Accelerate** property of the NFS datastores within the datstores tab illustrated in [Figure 38](#).

View: Datastores Devices

Datstores Refresh Delete Add Storage... Rescan All...

Identification	Status	Device	Drive Type	Capacity	Free	Type	Hardware Accelerat...	Storage I/O
home	Normal	10.244.156.107:/home	Unknown	2.95 TB	2.80 TB	NFS	Supported	Enabled
Auto	Normal	DGC Fibre Channel Disk (naa.6006...	Non-SSD	399.75 GB	398.80 GB	VMF55	Supported	Enabled
VNX958-4	Normal	DGC Fibre Channel Disk (naa.6006...	Non-SSD	329.75 GB	120.79 GB	VMF55	Supported	Enabled
VNX958-3	Normal	DGC Fibre Channel Disk (naa.6006...	Non-SSD	329.75 GB	184.71 GB	VMF55	Supported	Enabled
datastore1	Normal	Local SEAGATE Disk (naa.5000c50...	Non-SSD	63.25 GB	62.29 GB	VMF55	Unknown	Enabled

Figure 38 NFS Hardware Accelerated Datastore Property

If the VNX NFS datastore property is not set to "Supported", run the following command on the ESX host to verify the plug-in is correctly installed:

```
esxcli software vib list | grep EMCNasplugin
```

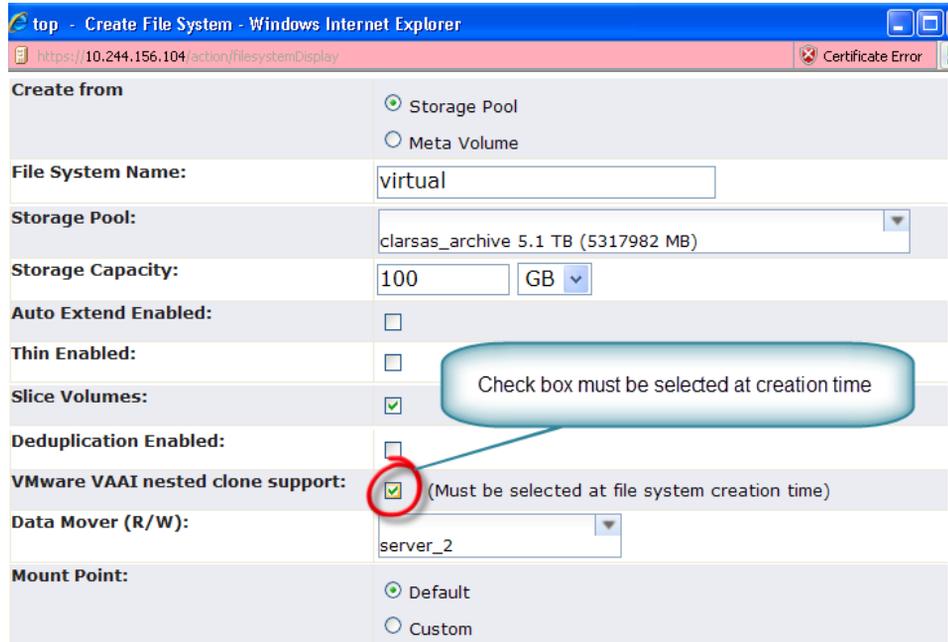


Figure 39 Create File System

Vmkfstools extended stats

vSphere 5 includes additional vmkfstools command line argument to display the disk utilization for virtual machine disks configured. The `-extendedstat` argument provides disk details for the virtual disks using NFS storage. The command reports virtual disk size, used space, and unshared space. The `-extendedstat` reports all values in bytes, as shown in Figure 40. This helps when creating automated reports or custom provisioning scripts.

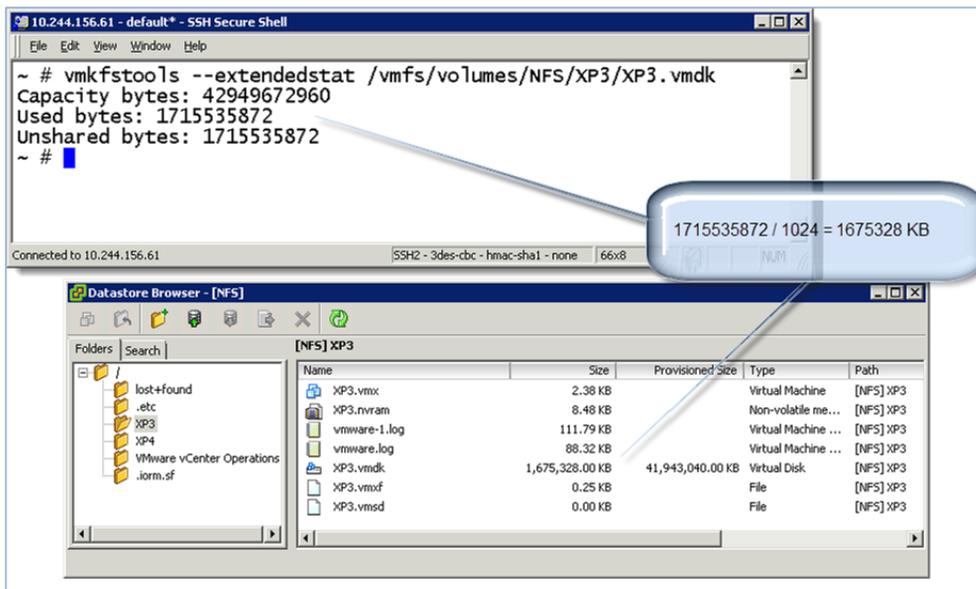


Figure 40 Vmkfstools disk utilization option

Storage Distributed Resource Scheduler (SDRS)

Storage DRS (SDRS) is a vSphere 5 feature that allows VMware administrators to apply Distributed Resource rules to storage in a similar manner to the way vSphere manages CPU, and Memory Resources in DRS. Independent datastores are grouped together and placed under SDRS control to simplify virtual disk management and improve storage resource utilization in vSphere environments.

SDRS relies upon a new storage object called a datastore cluster. These clusters consist of multiple VMFS or NFS datastores as shown in Figure 41 on page 86.

An SDRS cluster is configured by adding existing VMFS or NFS datastores; however, each cluster must contain either NFS or VMFS volumes, and not both in the same cluster. Clusters are resized quickly by adding or removing datastores through the vCenter SDRS management.

Datastore clusters can include LUNs from multiple VNX systems, although this is not recommended. However, VAAI only works with LUNs accessed from the same storage system. The performance of Storage vMotion is impacted due to the lack of VAAI support if LUNs reside on different systems.

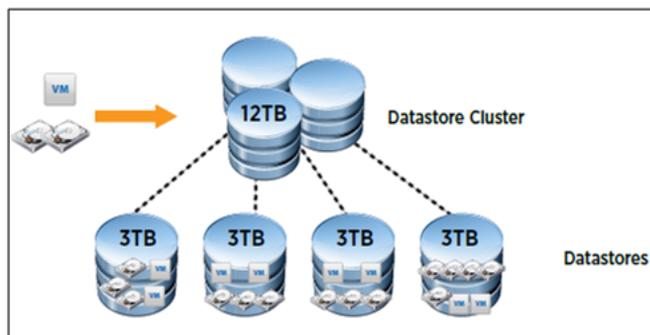


Figure 41 Storage DRS datastore cluster

SDRS monitors the capacity and response time of each datastore within the cluster. It applies policy rules to determine virtual machine initial placement and relocation within the clustered datastores.

Virtual machine placement simplifies resource planning, which has traditionally required performance monitoring and analysis. Instead of running tools to identify hot spots and perform manual migrations, create an SDRS cluster. Use datastores with similar performance characteristics and establish a policy to specify capacity and latency requirements for virtual machine disks. SDRS will continuously monitor the storage resources and provide recommendations to distribute the virtual machines between the datastores.

Relocation moves the virtual machine from the existing datastore to one of the other datastores in the cluster. SDRS relocation recommendations can be configured for manual or automated execution.

SDRS monitors available capacity (free space) and, optionally, device latency for each datastore within the cluster. SDRS makes recommendations for virtual machine relocation when:

- ◆ An individual datastore exceeds its defined capacity threshold.
- ◆ A change occurs in the environment.
- ◆ The administrator selects the SDRS button.
- ◆ A capacity- or service-level imbalance exists between the datastore where the virtual machine resides and another datastore in the cluster.

Storage DRS is not meant to be a highly reactive solution. It can be tuned for aggressive relocations, but the default relocation policy requires 8 - 24 hours of activity. SDRS continuously collects datastore capacity and, optionally, I/O latency information. At user-defined intervals, the datastore information is assessed against existing policy rules to determine if virtual machine relocation is warranted.

Note: VNX FAST VP is also a periodic task that can be automated or run manually to rebalance the blocks within a Pool LUN. The two features will work together; however, do not use FAST VP when I/O metrics are in use. I/O metrics are to be disabled on FAST VP LUNs.

SDRS policy configuration

Storage DRS provides two automation policies as shown in [Figure 42 on page 88](#):

- ◆ **Fully Automated** performs initial placements and virtual machine relocation without user intervention.
- ◆ **No Automation** presents a recommendation each time a virtual machine relocation would be triggered.

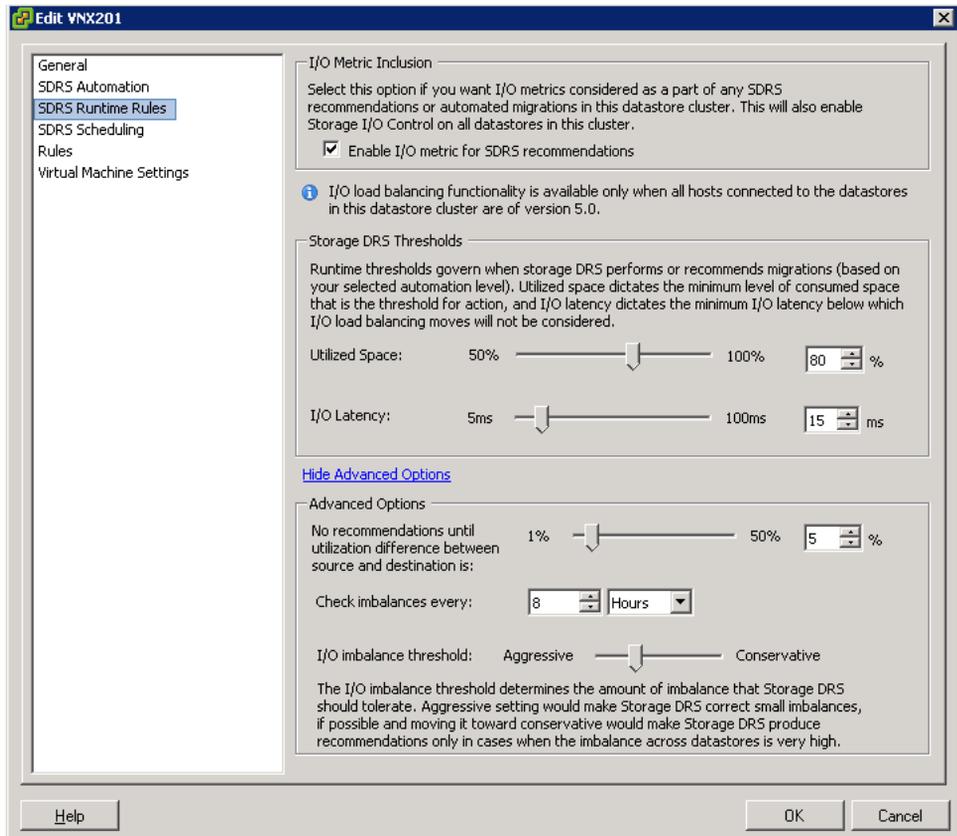


Figure 42 SDRS advanced policy configuration

Policy metrics are:

- ◆ **Utilized Space** — the amount of space consumed within a datastore. The default value for this parameter is 80 percent. This means that SDRS does not evaluate migration policy until the datastore exceeds that capacity threshold.
- ◆ **I/O Latency** — the datastore response time measured in milliseconds. The default value is 15 ms. SDRS does not evaluate migration policy until the datastore exceeds a 15 ms response time, and the imbalance rules are also satisfied.

- ◆ **Imbalance timer value** — this value defines the interval for applying the DRS policy to the datastore cluster. The default value is eight hours. vSphere collects data at standard intervals and reconciles resource utilization every eight hours.

Do not complete latency assessments for FAST VP LUNs, because variability in the application workload can distort the results. Although performed at a different level of granularity, SDRS and FAST VP perform a similar function to rebalance resources. Use either SDRS or FASTVP for workload rebalancing across storage resources. Do not use both services at the same time.

Figure 43 shows the interface to disable I/O metrics and apply policy based on capacity utilization. Clear **Enabled I/O metric for SDRS recommendations**.

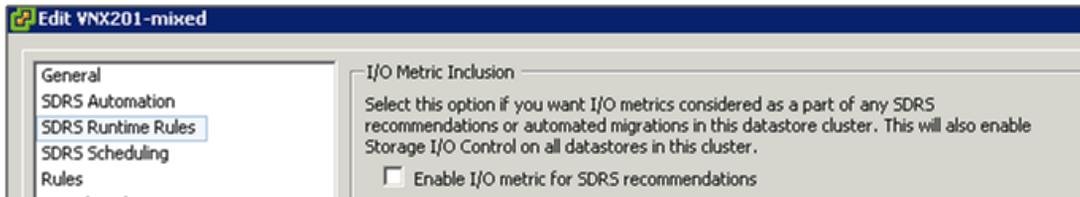


Figure 43 SDRS I/O metric enablement setting

Note: SDRS I/O load balance does not work if the datastore cluster is not configured for all hosts that share the datastores.

Note: VAAI operations do not span storage systems. The host processes virtual machine migrations between clustered datastores from different storage systems.

VNX storage recommendations for SDRS

Create a datastore cluster from LUNs that has the similar storage characteristics such as capacity, drive type, latency, and tiering policy. This configuration allows SDRS to balance virtual machine capacity and I/O requirements evenly.

When vStorage APIs for Storage Awareness (VASA) and virtual machine storage profiles are configured, each datastore must have the same capability for automated migrations and virtual machine evacuations.

Avoid using LUNs from the same RAID group or storage pool within a SDRS cluster. The intent of SDRS is to distribute the I/O between the storage resources within VNX. Creating multiple LUNs from the same RAID group will share the same set of spindles, which could negate the benefits to SDRS. The following list identifies several actions to complete for SDRS:

- ◆ Use LUNs of equal size and storage type.
- ◆ Add LUNs in pairs and distribute LUN ownership between the VNX storage processors.
- ◆ Disable I/O metrics when using FAST VP pool LUNs.
 - Set migration policy to manual when using FAST VP configurations.
- ◆ Configure the migration policy to manual mode until you have assessed the environment for a period of time.
- ◆ Assign multiple Storage vMotion connections to reduce migration times.
- ◆ Do not use SDRS with datastore LUNs that are protected with VNX synchronous replication technologies such as MirrorView.
 - Virtual machine relocations can significantly impact synchronous replication. To use synchronous replication, set the SDRS migration policy to manual to limit unnecessary data replication from virtual machine migration.

Table 3 shows supported SDRS LUN configurations.

Table 3 Supported SDRS LUN configurations

VNX Feature	Initial Placement	Migration Recommendations
Thin, Thick, FLARE LUN	X	X
FAST VP	X	No, manual mode
FAST Cache	X	No, manual mode
Replication	X	No
LUN snapshots	X	No
Dedupe	X	No
Thin	X	Supported with VASA

vStorage API for Storage Awareness (VASA)

VASA, introduced in vSphere 5.0, is implemented as a vCenter service that communicates with the storage system to discover the storage capabilities of the VNX devices. vCenter presents these storage capabilities in various management interfaces related to datastores, datastore clusters, and virtual machine disks. [Figure 44](#) illustrates the storage capabilities of a datastore cluster using SAS drives with Fast Cache enabled for the LUN.

The screenshot displays the vSphere Storage DRS configuration for a VNX958 datastore cluster. The interface includes a navigation bar with tabs for Getting Started, Summary, Virtual Machines, Host and Clusters, Datastores, Storage DRS, Performance, Tasks & Events, Alarms, Permissions, and Storage Views. The main content area is divided into several sections:

- General:**
 - Datastores: 2
 - Capacity: 659.50 GB
 - Used space: 354.00 GB
 - Free space: 305.50 GB
 - Datastore Largest Free Space: 184.71 GB
 - Type: VMFS
 - Virtual Machines: 6
 - VMDKs: 8
 - Snapshots: 2
- Commands:**
 - Refresh
 - Edit Datastore Cluster
 - Add Storage
- vSphere Storage DRS:**
 - I/O metrics: Enabled
 - Storage DRS: Enabled
 - Automation level: Manual
 - Utilized Space threshold: 55 %
 - I/O latency threshold: 5 ms
 - [View troubleshooting guide](#)
- Storage Capabilities:**
 - System Storage Capability: SAS/Fibre Storage;.. [Refresh](#)
 - User-defined Storage Capability: N/A
- Storage Capability Details:**
 - Name: SAS/Fibre Storage; FAST Cache
 - Description: SAS or Fibre Channel drives; FAST Cache enabled

Figure 44 VASA datastore storage capability of VNX Flash drive LUN

Awareness of the storage capabilities of each datastore allows the vSphere administrator to make informed decisions when performing administrative tasks. For example, knowing that the target datastore for a virtual machine migration has the same capabilities as the source ensures that the task does not impact the virtual machine service level.

Additionally, virtual machine storage profiles leverage storage capabilities to identify appropriate datastores for Storage vMotion operations.

VNX OE for Block provides native VASA support in releases 5.32 and later. In versions prior to 5.32, VASA support is provided through the EMC Solutions Enabler VASA Provider.

The initial VASA release included a basic set of properties to identify the capabilities of each LUN. There was no support for NFS capabilities prior to VNX OE for Block version 5.32. [Table 4](#) lists capabilities for that implementation.

Table 4 VASA storage capability mapping to VNX LUNs

VNX LUN type	vCenter storage capability
FLARE SAS LUN	Capacity
FLARE EFD LUN	Extreme performance
Pool LUN	Multitier storage
Fully Automated Storage Tiering LUN	Multitier storage
FAST Cache LUN	Multitier storage
NFS export	Unsupported

VNX OE for Block version 5.32 provides the VASA service through the VNX storage processor and Control Station. When using VASA on a VNX OE for Block version 5.32 or later, configure the vCenter VASA service with direct access to the VNX controllers.

The 5.32 release reports the datastore capabilities based on the disk type used to create the datastore. The disk properties are listed in column 1 of [Table 5 on page 93](#) as SAS, NL-SAS, Solid State, or Automated Storage Tiering when the LUN is created from multiple disk types using VNX FAST VP technology.

Additional properties of the device are appended to the basic storage element to differentiate the capabilities. Those are listed in the LUN properties column. The end result as shown in [Figure 44 on page 91](#) is that the LUN will include a single storage type and can include zero or more properties.

For example, a SAS RAID group LUN without FAST Cache enabled has a storage capability of SAS/Fibre Channel.

Table 5 shows that a thin pool LUN with mixed drive types has a storage capability of "Automated Storage Tiering, Thin Pool LUN."

Table 5 VNX OE for Block 5.32 storage capability mapping to VNX LUNs

VNX LUN type	LUN properties	vCenter filters include one or more item listed below
<u>VNX Block Provider</u> NL-SAS/SATA SAS/Fibre Channel Solid State Auto-Tier	FAST Cache enabled LUN Replication LUN Compression Thin Pool LUN	FAST Cache Remote Replication Space Efficiency Thin
<u>VNX File Provider</u> NL-SAS/SATA SAS/Fibre Channel Solid State Auto-Tier	FAST Cache enabled File Replication (RepV2) File Dedeuplication Thin Pool LUN	FAST Cache Storage Efficiency Thin Replication

Virtual machine storage profiles

Virtual machine storage profiles provide the ability to associate each virtual machine disk with a particular storage capability. Virtual machine storage profiles are defined by associating the profile with one or more VNX storage capabilities. [Figure 45 on page 94](#) shows a new user-defined profile name called "SAS Fibre FAST Cache." This profile includes all SAS LUNs that have FAST Cache enabled, and no other LUN capabilities enabled. All datastores that possess the SAS and FAST Cache capabilities are candidates for virtual machine disks that are assigned to this storage profile.

Note: A storage capability can be assigned to multiple storage profiles. Use caution when creating new profiles to ensure that the policy performs as intended.

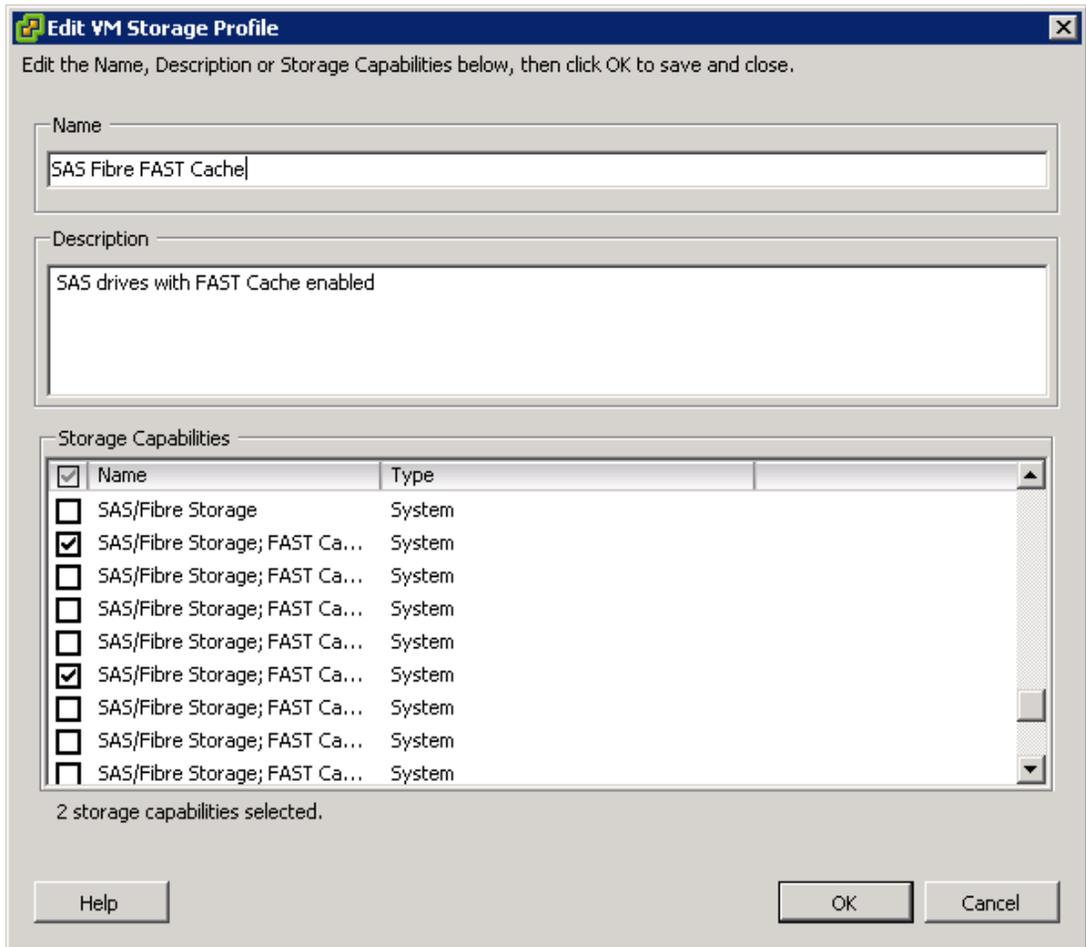


Figure 45 Storage profile assignment

Virtual machine storage profiles are assigned to each virtual disk. They enforce virtual disk to datastore compliance, and virtual disk migration for tasks such as Storage vMotion. When a migration or Storage vMotion is initiated, the Migration wizard identifies the datastores that are compatible for the current virtual machine storage profile.

In Figure 46, two datastores are compatible with the SAS Fibre storage profile. In this example, both datastores are using SAS disks; however, one is an NFS datastore and the other is a VMFS datastore. The VASA service highlights the recommended datastore, but presents both as compatible options. Use the **Type** field in the list to identify the transport protocol and ensure that the correct one is selected.

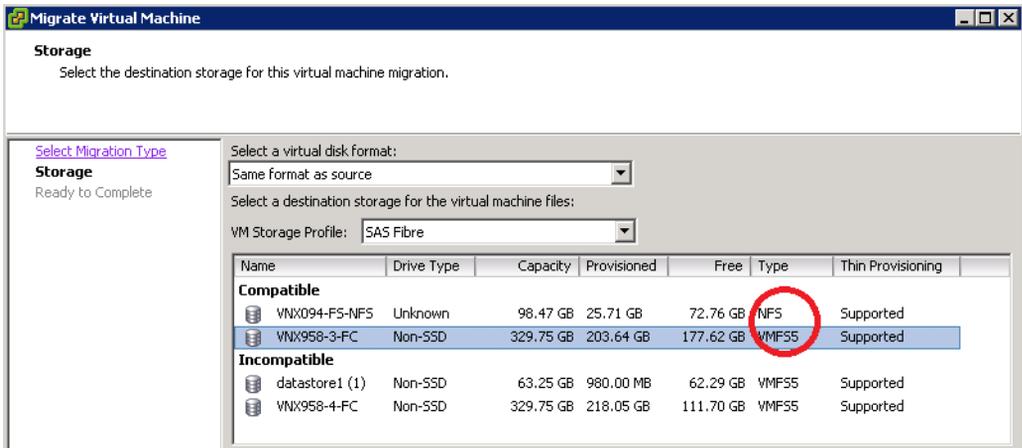


Figure 46 Compatible or incompatible with SAS Fibre storage profile

Virtual machine storage profiles are also used by datastore clusters when SDRS is enabled. SDRS controls virtual disk placement and uses profiles for migrations and evacuations when a datastore is placed in maintenance mode.

Note: If SDRS and Storage Profiles are used, ensure that the datastores support the storage capabilities, otherwise automated migrations may not work correctly.

User-defined storage capabilities

In some cases, VASA does not have a profile that matches the properties of a datastore, or there is a need to define a profile for specific datastores in the environment. For example, vSphere 5.0 and VNX OE for Block version 5.31 provide a limited set of VMFS capabilities, and does not support NFS datastores. Create a user-defined profile to use storage profiles.

Complete the following steps to configure a user-defined storage profile for NFS datastores from VNX storage.

1. Log in to vSphere and select the **VM Storage Profiles** icon.



2. Enable virtual machine storage profiles for the hosts in the cluster:
 - a. Select **Manage storage capabilities**
 - b. Add a storage profile with a user-defined name

In [Figure 47](#), the storage capability is defined as NFS and includes a description of the storage in this profile.

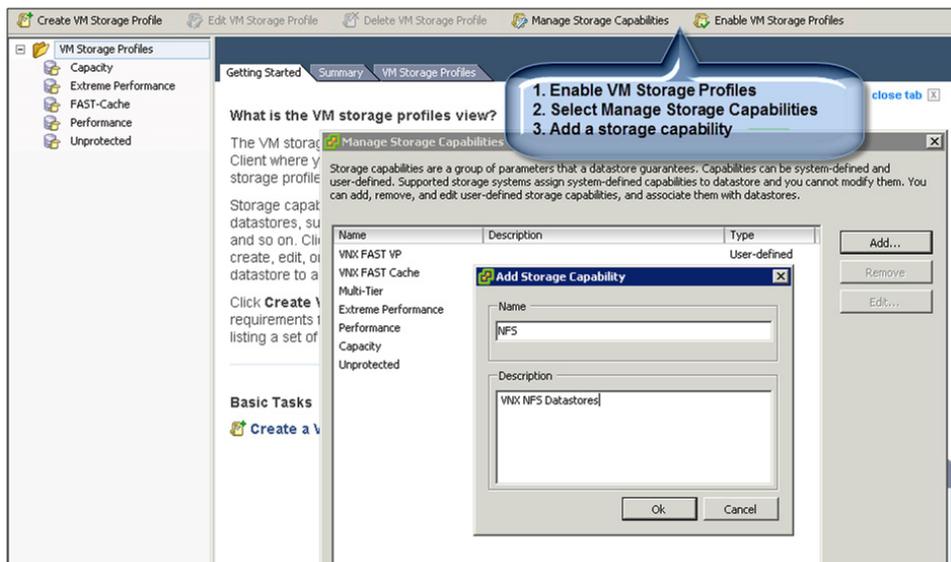


Figure 47 Creating a user-defined profile

3. Add a virtual machine storage profile as shown in [Figure 48](#) on [page 97](#):

This virtual machine profile can use the same name as the storage profile.

4. Select the user-defined storage profile from step 2 to associate the virtual machine profile with the storage profile.

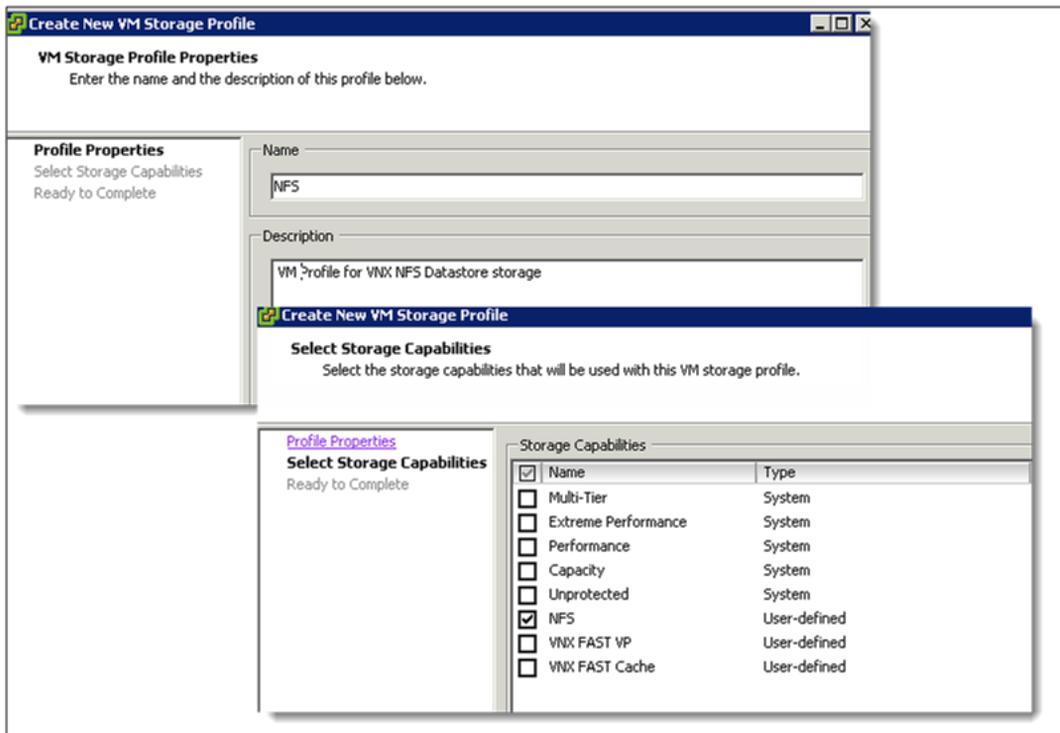


Figure 48 Creation of a user-defined virtual machine storage profile

5. Assign the new profile to existing datastores as shown in [Figure 49 on page 98](#). and [Figure 50 on page 98](#).

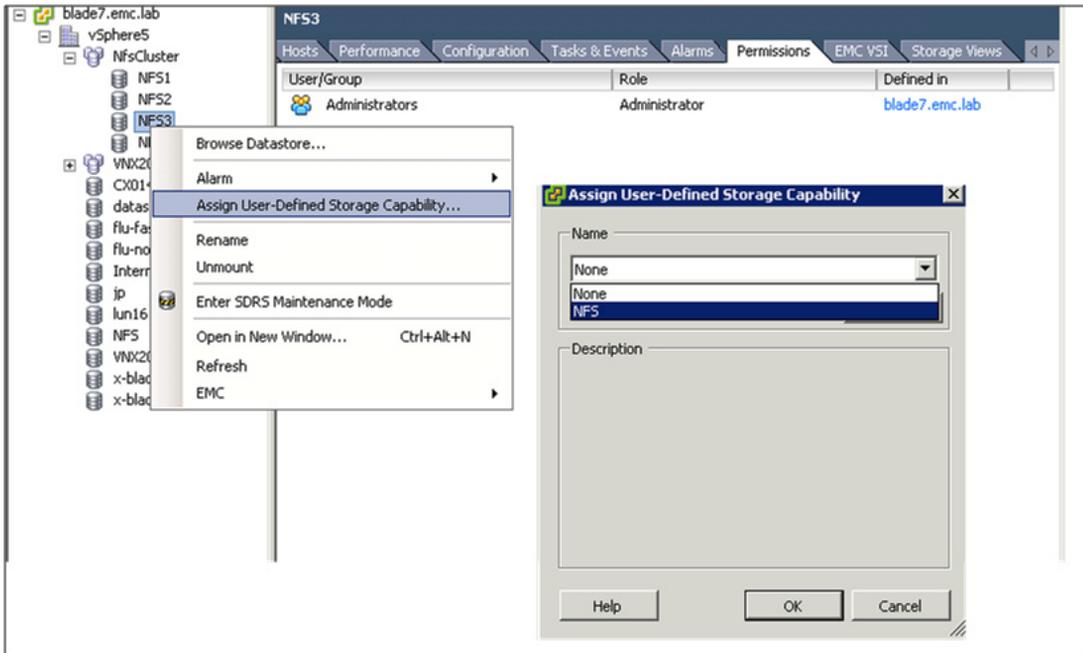


Figure 49 Associating datastores with a user-defined storage profile

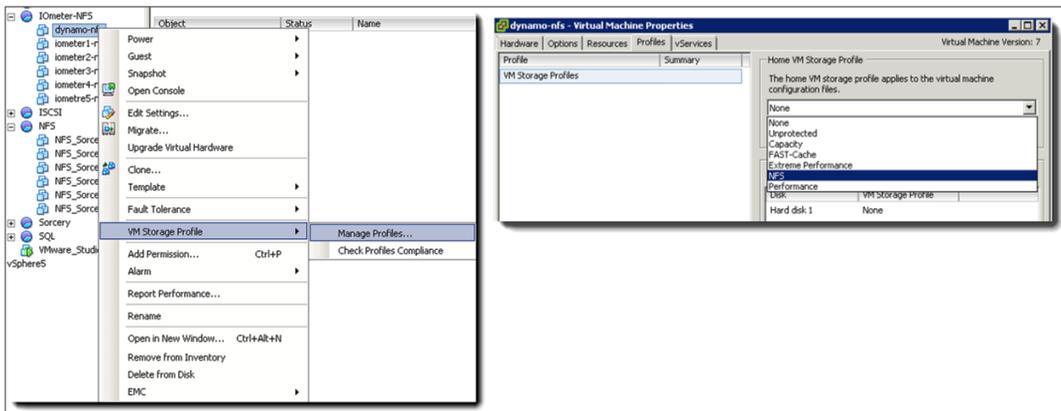


Figure 50 Associating the virtual machine with a user defined storage capability

6. Associate the virtual machine virtual disks with this profile to ensure compliance.
7. Introduce this profile as part of the virtual machine storage management tasks.

vCenter storage provider configuration

VASA runs as a client service called vSphere Profile-Driven Storage on the vCenter server. The service interacts with an EMC Provider running on either a Windows system running Solutions Enabler, the VNX storage processor, or on the VNX Control Station.

Note: VNX OE for Block version 5.31 requires an SMI-S proxy service to communicate with the storage processor. Install and configure the EMC VASA provider on a Windows system, or deploy the VASA provider virtual appliance. The Windows system can be the same host that runs vCenter or a stand-alone system.

The vSphere storage provider communicates with the EMC provider over secure http and an administrative SMI-S-authorized user account.

1. Select the **Storage Providers** icon in the vSphere management screen to start the configuration interface illustrated in [Figure 51](#).

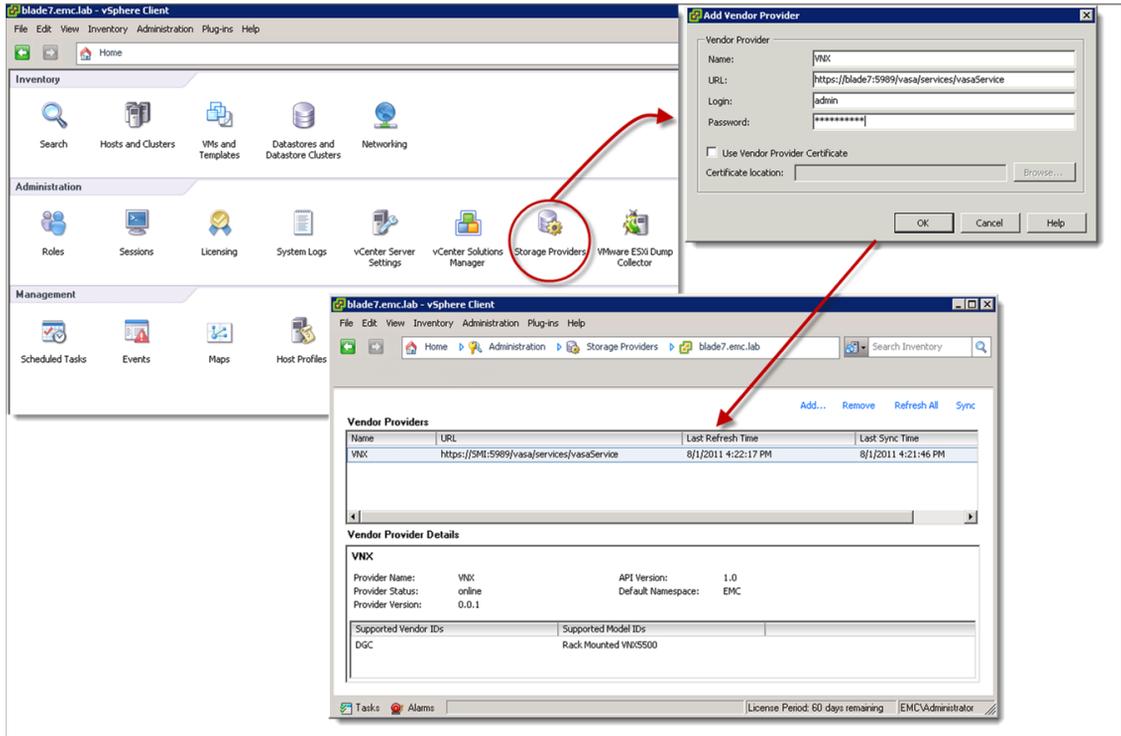


Figure 51 VASA configuration

2. Click **Add** to configure the vCenter VASA service with an existing EMC SMI-S VASA provider service that is configured to monitor the VNX system in the vSphere environment.

The following information is required to assign the new service in the vSphere Client to an SMI-S Server:

- ◆ User-defined name
- ◆ VASA Service Uniform Resource Locator on the SMI-S system in the following format:
https://<smi_server_name>:5989/vasa/services/vasaService

- ◆ Login credentials for the SMI-S service: These are the credentials defined for the VASA service within SMI-S.

VNX OE for Block versions 5.32 and later have embedded the Block provider onto the storage processor. A File provider is available on the VNX Control Station for File and Unified systems. With the release of version 5.32, an external SMI-S service is not required. Configure the VASA service to communicate directly to the storage processor for block, and the Control Station for file. This service does not require an external SMI-S server. As of VNX OE for File 7.1, the Control Station supports a VNX Provider for File storage.

VNX OE for Block version 5.32 and VNX OE for File version 7.1 environments use the following URL syntax and the IP address of the storage processor.

URL to the VASA service in the following format:

- ◆ Storage Processor configuration
 - `https://<storage processor IP Address>/vasa/services/vasaService`
 - Login credentials for the Control Station:
 - user id: vmadmin
 - password: <vmadmin password>
- ◆ Control Station configuration
 - `https://<Control Station IP address>:5989/vasa/services/vasaService`
 - Login credentials for the Control Station:
 - user id: vmadmin
 - password: <vmadmin password>

Storage I/O Control (SIOC)

SIOC offers storage resource management capability for virtual disks and datastores. It provides a way to govern virtual disk utilization within a clustered datastore. SIOC uses virtual machine disk shares and disk IOPS settings to establish precedence, and apportions the virtual machine storage resources when the datastore response time exceeds predefined levels.

SIOC can be used along with FAST VP.

Virtual machine disk shares are assigned when the virtual disk is created. The default share value is normal or 1,000 shares. It is customizable, and there are settings of low (500) and high (2,000) share value. SIOC works at the host and cluster level. It aggregates the virtual disk share values of all powered-on virtual machines on the host and uses that value as a percentage of all other host disk shares when it needs to throttle the device queue among hosts in the cluster.

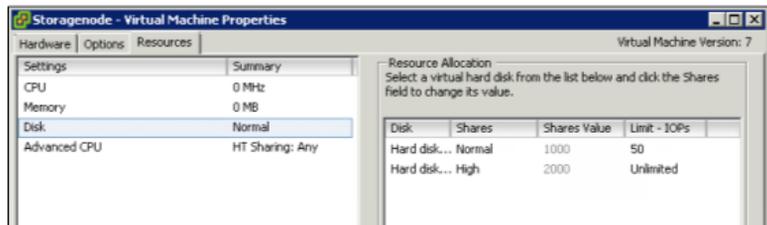


Figure 52 Virtual disk shares configuration

SIOC uses a latency value called a congestion threshold. This value, specified in milliseconds (ms), defines the acceptable latency of the device that supports the datastore. Valid settings range from 5 ms to 100 ms. Thirty ms is the default value.

The appropriate congestion control value for a datastore depends on multiple factors:

- ◆ The type of device
- ◆ Number of disks supporting the LUN
- ◆ Other consumers of the spindles

Define an IOPS limit per virtual machine to avoid a situation where a single virtual machine monopolizes the datastore. For example, limit the amount of IOPS per virtual machine to 1,000.

Table 6 lists the recommendations for setting the congestion threshold.

Table 6 SIOC congestion windows

Datastore storage type	Congestion window (ms)	Notes
Enterprise Flash drive	10-20	
SAS drive	20-30	
NL-SAS	35-50	
FAST VP/Tiered LUN	35-50	View the storage distribution within the pool.
NFS	30	<ul style="list-style-type: none"> Response time includes any latency that exists in the network. Increase the congestion window by any latency that exists in the network.

Note: SIOC detects non-VMware workloads on a shared storage device. If the SIOC LUN is accessed for some other purpose, such as replication or storage system cloning, ESXi generates an error that states that an external workload is detected. *Unmanaged I/O workload detected on shared datastore running Storage I/O Control (SIOC) for congestion management (1020651)*, available in the VMware Knowledge base, provides more information.

SIOC for NFS

vSphere versions 5.0 and later provide SIOC support for NFS datastores mounted on ESX host clusters. SIOC for NFS uses the same framework as VMFS by applying a synthetic queue depth for NFS file systems. The SIOC driver throttles I/O by adjusting the host queue depth to the NFS datastore file systems when contention is encountered. Each configured datastore inherits a default host response time value of 30 ms.

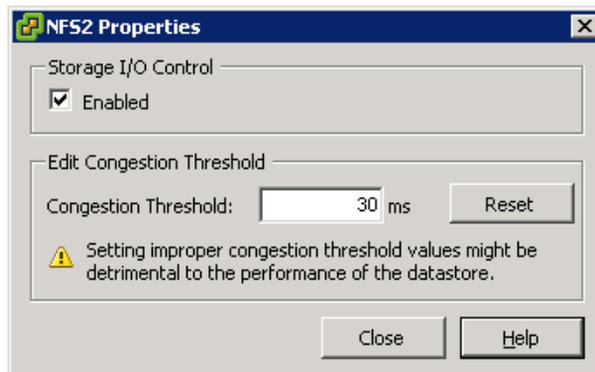


Figure 53 NFS SIOC congestion window

Note: NFS datastore response time includes network latency. Ensure the IP storage network does not contribute latency of more than a few milliseconds, or adjust the congestion threshold setting for network overhead.

Note: Workloads that compete for the NFS datastore I/O can impact SIOC. Do not share the NFS datastore or file system disks.

Network considerations

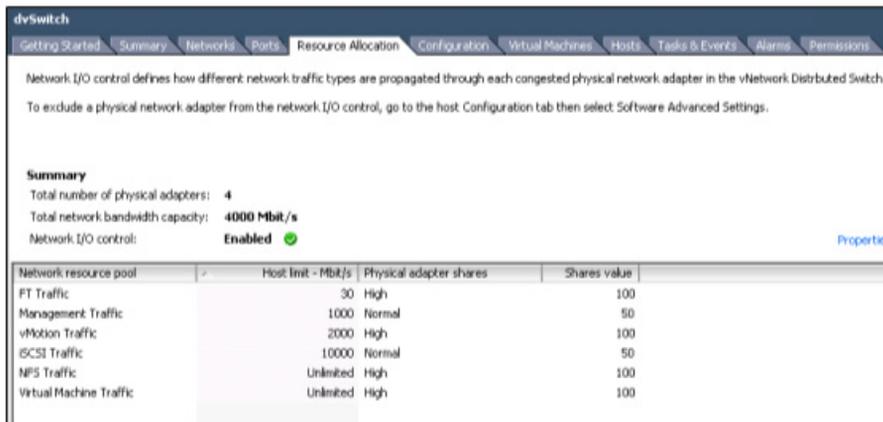
The VNX platform supports a wide range of network topologies and capabilities for VMware vSphere. This section lists items to consider when planning an IP storage network for vSphere servers.

Network I/O Control (NIOC)

NIOC provides a way to manage and prioritize network resources at the cluster level. NIOC is an advanced networking feature of vNetwork distributed switches for vSphere versions 4.1 and later.

vNetwork distributed switches provide an efficient way to centralize, manage, and share datacenter network resources. NIOC enables the virtual administrator to classify network traffic. Each network type is configured with a share value which applies a weighting factor to prioritize network traffic.

Figure 54 shows that NIOC has several default network classes that enable finer control of the network resources within each network resource pool. A throughput value can also be assigned to limit the resource utilization in Mb/s for each host that shares that resource.



dvSwitch

Getting Started Summary **Network** Ports Resource Allocation Configuration Virtual Machines Hosts Tasks & Events Warnings Permissions

Network I/O control defines how different network traffic types are propagated through each congested physical network adapter in the vNetwork Distributed Switch.
To exclude a physical network adapter from the network I/O control, go to the host Configuration tab then select Software Advanced Settings.

Summary

Total number of physical adapters: **4**

Total network bandwidth capacity: **4000 Mbit/s**

Network I/O control: **Enabled**  [Properties](#)

Network resource pool	Host limit - Mbit/s	Physical adapter shares	Shares value
FT Traffic	30	High	100
Management Traffic	1000	Normal	50
vMotion Traffic	2000	High	100
iSCSI Traffic	10000	Normal	50
NFS Traffic	Unlimited	High	100
Virtual Machine Traffic	Unlimited	High	100

Figure 54 Network Resource Allocation interface

The ability to adjust network prioritization offers some flexibility to tune the network for particular applications. With the trend toward converged networks, NIOC provides the ability to establish fairness for storage and virtual machine network adapters. Monitor the environment to ensure that the VMkernel resources are set to normal or high, and are not artificially limited by the network resource pool configuration.

LUN removal (All Paths Dead)

Prior to vSphere 5, a condition known as All Paths Dead (APD) occurs when an ESXi host loses access to a shared storage device. The device loss can be due to a temporary environmental issue like a switch failure, or an administrative action like removing a LUN from a storage group. In pre-ESXi 5 releases, the host could not differentiate between these two states.

In ESXi5 and later, the VMkernel performs additional SCSI commands to detect the state of the device and determine whether a device is in an All Paths Dead state, or a Permanent Device Loss (PDL) state.

All Paths Dead results when none of the HBAs on an ESXi host can establish a session with the VNX SCSI target that supports the datastore LUNs. In this state, the host continues to retry the connection for a period of time before marking the device unavailable.

PDL is a different state in which the host initiator has an active session with the SCSI target(s) on the storage processor. The host issues SCSI commands to the target and uses the SCSI sense codes returned by the VNX to determine the state of the missing device. If the host determines that the device is removed, it flags the device as PDL and performs the necessary steps to clean up the vCenter storage objects that were dependent on the storage device.

vSphere does not remove virtual machines that were stored within a datastore on the missing LUN. If a LUN is blindly removed, the virtual machines remain in an orphaned state.

To prevent orphan virtual machines, vSphere 5 provides a datastore workflow option to detach or unmount a datastore from the host as illustrated in [Figure 55 on page 107](#).

The feature provides a graceful device removal and ensures that the datastore removal does not violate any dependent relationships between the virtual disks and the datastore. Remove the device from the host storage group in vSphere after it is detached or unmounted.

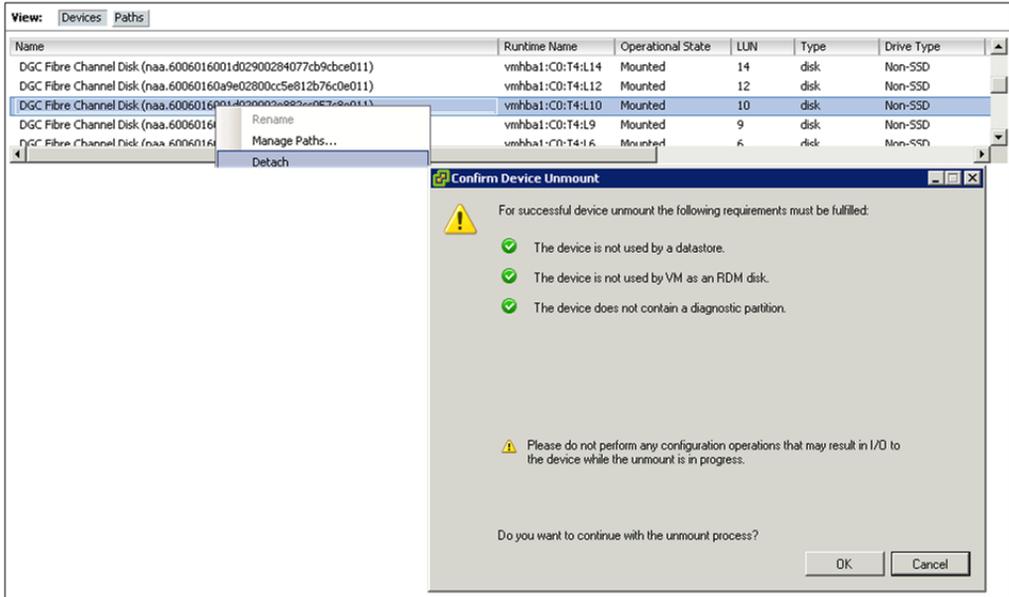


Figure 55 vSphere 5 Datastore removal wizard

Virtual machine considerations

Consider the following items to achieve optimal performance and functionality in virtual machines on VNX storage:

- ◆ Virtual machine disk partition alignment
- ◆ Virtual machine swap file location
- ◆ Paravirtualized SCSI adapter (PVSCSI)
- ◆ N Port ID virtualization (NPIV)
- ◆ Virtual machine resiliency over NFS

Virtual machine disk partitions alignment

The alignment of virtual machine disk partitions improves application performance and the efficiency of the storage system. Because a misaligned disk partition in a virtual machine may lead to degraded performance, align virtual machines that are deployed over any storage protocol. The following recommendations provide the best performance for the environment:

- ◆ Create the datastore in the vSphere Client or USM.
- ◆ The benefits of aligning boot partitions are generally marginal. If there is only a single virtual disk, consider adding an app/disk partition.
- ◆ It is important to align the app/data disk partitions that sustain the heaviest I/O workload. Align the partitions to a 1 MB disk boundary in both Windows and Linux.

Note: Windows 2008, Windows Vista, and Windows 7 disk partitions are aligned to 1 MB by default.

- ◆ For Windows, use the allocation unit size recommended by the application. Use a multiple of 8 KB, if no allocation unit size is recommended.
- ◆ For NFS, use the Direct Writes option on VNX file systems. This option helps with random write workloads and virtual machine disks formatted with a 4 KB allocation unit size.
- ◆ EMC also provides a free tool called UberAlign that identifies and corrects misaligned virtual disks. The Everything VMware at EMC website provides more information on this tool.

Align virtual machine disk partitions

The disk partition alignment within virtual machines is affected by a long-standing issue with the x86 processor storage configuration. As a result, external storage devices are not always aligned in an optimal manner. This is true for VMware in most cases. The following examples illustrate how to align data partitions with VNX storage for Windows and Linux virtual machines.

Aligning Windows virtual machines

Note: This step is not required for Windows 2008, Windows Vista, Windows 7, and Windows 8 which align partitions on 1 MB boundaries for disks larger than 4 GB (64 KB for disks smaller than 4 GB).

To create an aligned data partition, use the `diskpart.exe` utility. This example assumes that the data disk to be aligned is disk 1:

1. At the command prompt, type **diskpart**.
2. Type **select disk 1**, as shown in [Figure 56](#).

```

C:\WINDOWS\system32\cmd.exe - diskpart
Microsoft DiskPart version 5.2.3790.1830
Copyright (C) 1999-2001 Microsoft Corporation.
On computer: EMC-TWR07C0PSPV
DISKPART> select disk 1
Disk 1 is now the selected disk.
DISKPART> _

```

Figure 56 Select the disk

3. Type **create partition primary align=1024** to create a partition to align to a 1 MB disk boundary.
4. Type **Exit**.

Set the allocation unit size of a Windows partition

Use Windows Disk Manager to format an NTFS partition. Select an allocation unit that matches your application needs.

Note: The default allocation unit is 4 KB. However, larger sizes such as 64 KB can provide improved performance for volumes that store large files.

Aligning Linux virtual machines

Use the **fdisk** command to create an aligned data partition:

1. At the command prompt, type **fdisk /dev/sd<x>** where **<x>** is the device suffix.
2. Type **n** to create a new partition.
3. Type **p** to create a primary partition.
4. Type **1** to create partition number 1.
5. Select the defaults to use the complete disk.
6. Type **t** to set the partition system ID.
7. Type **fb** to set the partition system ID to fb.
8. Type **x** to go into expert mode.
9. Type **b** to adjust the starting block number.
10. Type **1** to choose partition 1.
11. Type **2048** to set the starting block number to 2048 for a 1 MB disk partition alignment.
12. Type **w** to write the label and partition information to disk.

Identify the alignment of virtual machines on Windows

Complete the following steps to identify virtual disk alignment:

1. From the **Start** menu, select **Programs > Accessories > System Tools > System Information**. The **System Information** window appears as shown in [Figure 57 on page 111](#).

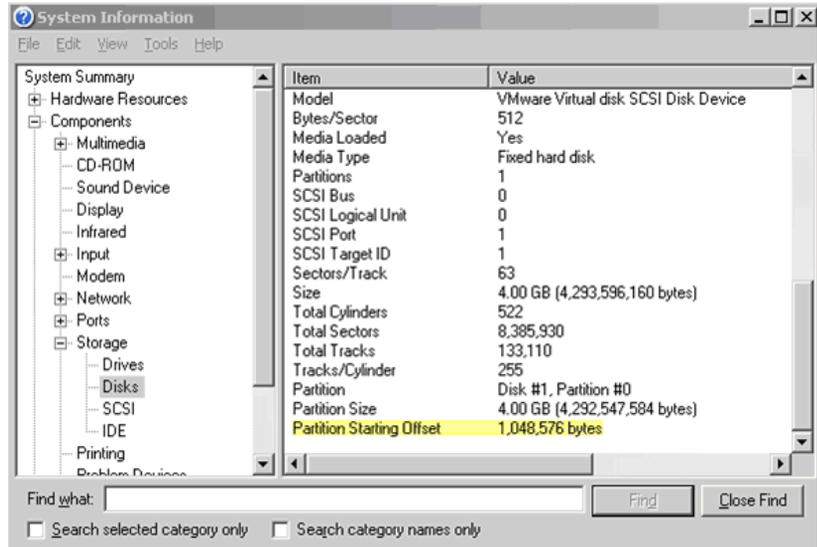


Figure 57 Guest disk alignment validation

2. Locate the **Partition Starting Offset** property and verify the value is **1,048,576** bytes as shown in Figure 58. This value indicates alignment to a 1 MB disk boundary.

Note: Type `wmic partition get StartingOffset, Name` at the command prompt to display the partition starting offset.

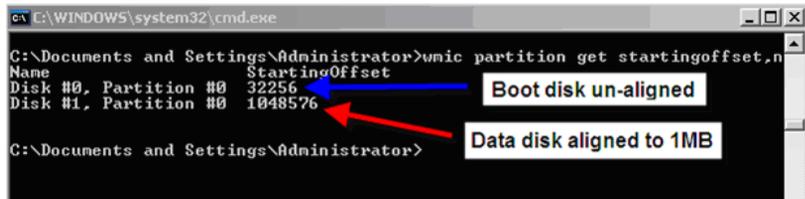


Figure 58 NTFS data partition alignment (wmic command)

Partition allocation unit size

Run the `fsutil` command to identify the allocation unit size of an existing data partition. In the following example, the E: drive is an NTFS data partition that is formatted with an allocated unit size of 8 KB.

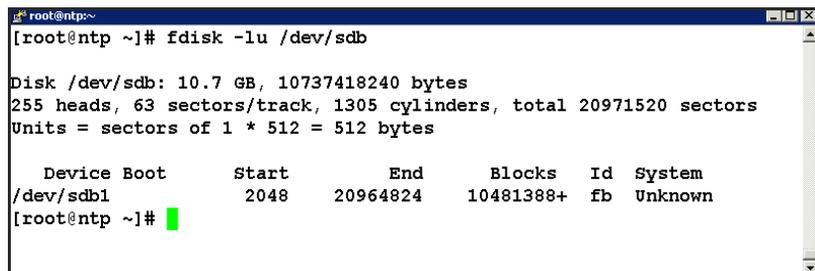
At the command prompt, type `fsutil fsinfo ntfsinfo <drive_letter>`.

The Bytes Per Cluster value identifies the allocation unit size of the data partition.

Identify Linux virtual machine alignment

Run the `fdisk` command to identify the current alignment of an existing Linux data partition. In the following example, `/dev/sdb` is a data partition that was configured on a Linux virtual machine.

In the terminal session, type `fdisk -lu <data_partition>`.



```

root@ntp:~
[root@ntp ~]# fdisk -lu /dev/sdb

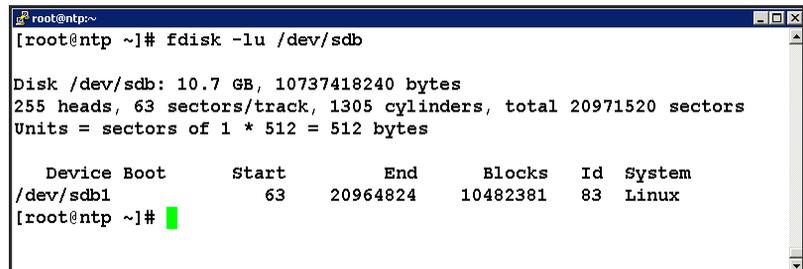
Disk /dev/sdb: 10.7 GB, 10737418240 bytes
255 heads, 63 sectors/track, 1305 cylinders, total 20971520 sectors
Units = sectors of 1 * 512 = 512 bytes

   Device Boot      Start         End      Blocks   Id  System
/dev/sdb1             2048     20964824     10481388+  fb  Unknown
[root@ntp ~]#

```

Figure 59 Output of 1 MB aligned Linux partition

The unaligned disk shows the starting sector as 63.



```

root@ntp:~
[root@ntp ~]# fdisk -lu /dev/sdb

Disk /dev/sdb: 10.7 GB, 10737418240 bytes
255 heads, 63 sectors/track, 1305 cylinders, total 20971520 sectors
Units = sectors of 1 * 512 = 512 bytes

   Device Boot      Start         End      Blocks   Id  System
/dev/sdb1             63     20964824     10482381    83  Linux
[root@ntp ~]#

```

Figure 60 Output for an unaligned Linux partition (starting sector 63)

Virtual machine swap file location

Each virtual machine is configured with a swap file that stores memory pages under certain conditions, such as when the balloon driver is inflated within the guest OS. By default, the swap file is created and stored in the same folder as the virtual machine.

When the swap file is stored on a SAN device it can have an adverse impact on virtual machine performance if there is a lot of concurrent I/O that results from paging activity.

Use proper virtual machine memory and resource configuration to avoid swapping. Do not unnecessarily reserve or artificially cap memory resources for virtual machines. These configurations contribute to swapping conditions.

The best way to avoid the impact of swapping is to use low latency, high throughput devices such as local or SAN EFD storage. This alleviates the contention that results from swapping activity.

It is possible to use a local device to offload up to 10 percent of the network traffic that results from the page file I/O. The trade-off for moving the swap file to the local disk is that it may result in additional I/O when a virtual machine is migrated through Storage vMotion or DRS. In such cases, the swap file must be copied from the local device of the current host to the local device of the destination host. It also requires dedicated local storage to support the files.

A better solution is to leverage high-speed, low-latency devices such as EFDs to support the swap files.

If each virtual machine has 100 percent of its memory reserved from host physical memory, it is possible to use SATA drives to support page files. Implementations for virtual desktop environments are examples of this scenario. Reserve the virtual machine desktop memory to allow the applications and OS to take advantage of client-side caching by using DD RAM within the ESXi host instead of the slower SAN storage. This approach yields sustained application performance.

If this configuration option is unavailable, use EFDs for page files where performance is a concern. vSphere 5 provides a feature called Host Cache to assist with the configuration of virtual machine swap files with EFD storage.

Host Cache

vSphere 5 simplifies the configuration of virtual swap through a new feature called Host Cache. Host Cache recognizes EFD storage assigned to the host, and allows a portion of that storage to be used to support virtual swap files. This feature configures virtual swap files within the datastore and provides them to the virtual machine to complement the existing swap configuration.

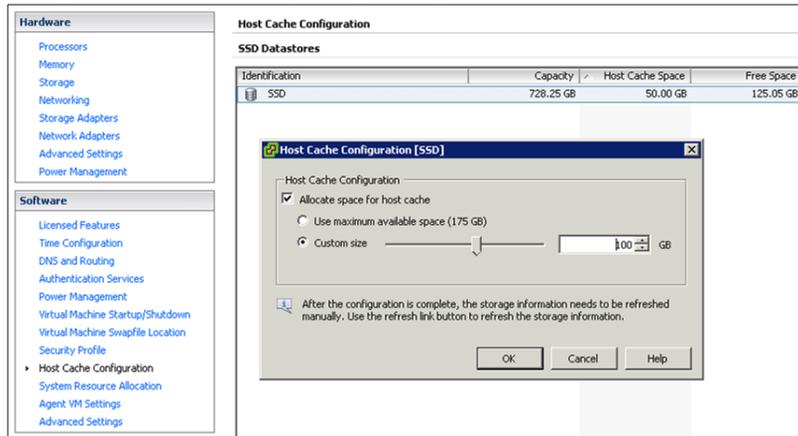


Figure 61 Host Cache configuration on VNX EFD storage

Paravirtual SCSI adapters

Paravirtual SCSI (PVSCSI) adapters are high-performance storage adapters that can provide greater throughput and lower CPU utilization. PVSCSI is best suited for SAN environments where hardware or applications drive very high throughput.

PVSCSI adapters combine I/O requests to reduce the cost of virtual interrupts. vSphere 4 Update 1 and later support the PVSCSI adapter for the virtual machine boot disk in addition to virtual data disks.

In tests run with Windows 2003 and Windows 2008 guest operating systems, the PVSCSI adapter has been found to improve the resiliency of virtual machines running on NFS-based storage.

The following guest operating systems support the PVSCSI adapters:

- ◆ Windows Server 2003 and 2008
- ◆ Red Hat Enterprise Linux (RHEL) 5

PVSCSI adapters have the following limitations:

- ◆ Hot-add or hot-remove requires a bus rescan from the guest.
- ◆ PVSCSI may not provide performance gains when the virtual disk has snapshots, or the ESXi host memory is overcommitted.
- ◆ If RHEL 5 is upgraded to an unsupported kernel, data may not be accessible from the virtual machine's PVSCSI disks. Run `vmware-config-tools.pl` with the `kernel-version` parameter to regain access.
- ◆ Booting a Linux guest from a disk attached to a PVSCSI adapter is not supported.
- ◆ Booting a Microsoft Windows guest from a disk attached to a PVSCSI adapter is not supported in ESXi prior to ESXi 4.0 Update 1.

Configuring disks to use VMware Paravirtual SCSI (PVSCSI) adapters (1010398), available in the VMware Knowledge Base, provides detailed information.

Note: Hot-adding a PVSCSI adapter to a virtual machine is not supported. Configure PVSCSI on the storage controller when the virtual machine is created.

N-Port ID Virtualization for RDM LUNs

N-Port ID Virtualization (NPIV) within the FC protocol enables multiple virtual N-Port IDs to share a single physical N-Port. This feature provides the ability to define multiple virtual initiators through a single physical initiator. It enables SAN tools that provide Quality of Service (QoS) at the storage-system level to guarantee service levels for virtual machine applications.

NPIV does have some restrictions. Adhere to the following guidelines to enable NPIV support:

- ◆ VMware NPIV support is limited to RDM volumes.
- ◆ Both the host HBAs and the FC switch must support NPIV.
- ◆ Enable NPIV on each virtual machine.
- ◆ Each virtual machine must have at least one RDM volume assigned to it.

- ◆ Mask LUNs to both the ESXi host and the virtual machine where NPIV is enabled.

Within VMware ESXi, NPIV is enabled for each virtual machine so that physical HBAs on the ESXi host assign virtual initiators to each virtual machine. As a result, a virtual machine has virtual initiators (WWNs) available for each HBA. These initiators can log in to the storage like any other host to provision block devices directly to the virtual machine through Unisphere.

Figure 62 shows how to enable NPIV for a virtual machine. To enable the NPIV feature, present an RDM volume through the ESXi host to the virtual machine. Virtual WWNs are assigned to that virtual machine after NPIV is enabled.

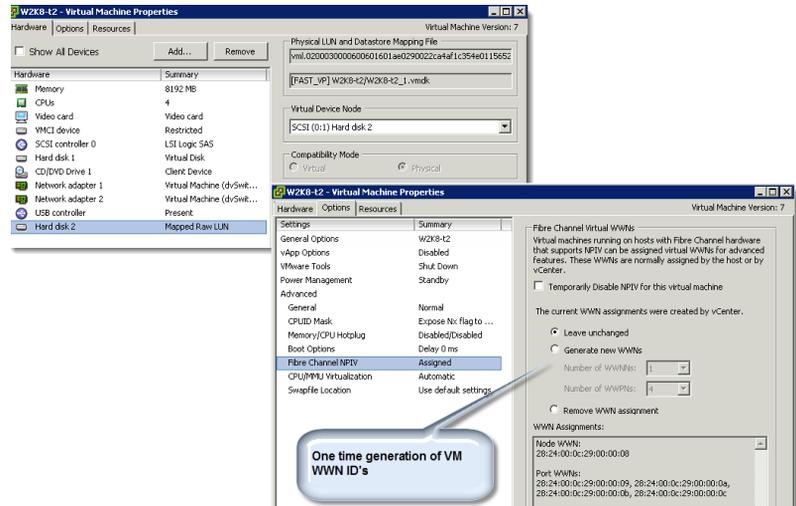


Figure 62 Enable NPIV for a virtual machine after adding an RDM volume

For some switches, manually type the virtual WWN names from the switch interface and then zone them to the storage system ports. The virtual machine initiator records then appear within the **VNX Connectivity Status** window for registration as shown in [Figure 63](#). Create a separate storage group for each NPIV-enabled virtual machine. In addition, present any LUNs assigned to the virtual machine storage group to the ESXi storage group.

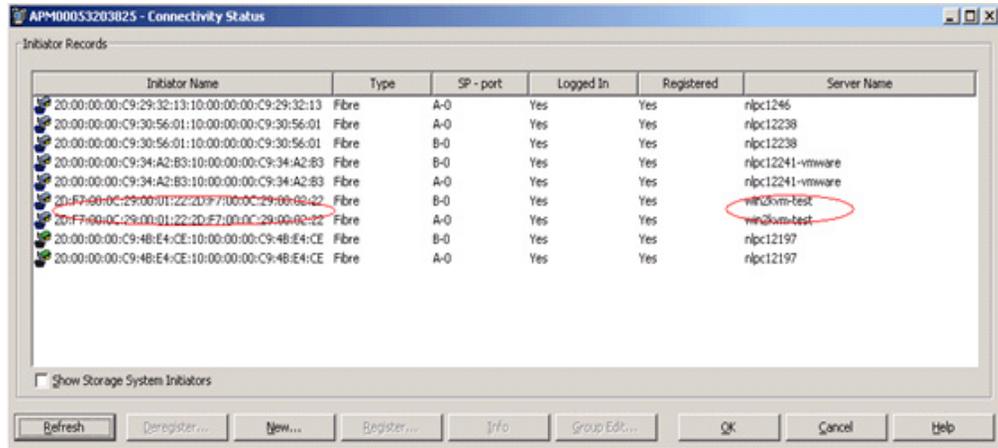


Figure 63 Manually register virtual machine (virtual WWN) initiator records

Complete the following steps to configure NPIV:

1. Ensure that the HBA and the FC switch support NPIV.
2. Assign an RDM volume to the ESXi host, and then to the virtual machine.
3. Enable NPIV to allow the virtual machine to create virtual WWNs.
4. Manually type in the virtual WWNs in the switch interface.
5. Zone the virtual WWNs to the VNX platforms in the switch interface. Add them to the same zone that contains the ESXi HBA and VNX storage ports.
6. Use Unisphere to manually register the initiator records for the virtual machine, and set the virtual machine to failover mode 4 (ALUA)
7. Create a new virtual machine storage group and assign the virtual machine records to it.

8. Add LUNs to the virtual machine:
 - c. Mask the LUNs to the ESXi hosts and the virtual machine storage group.
 - d. Assign the LUNs the same host LUN number (HLU) as the ESXi hosts.
 - e. Assign the LUNs to each virtual machine as RDM volumes.

Virtual machines resiliency over NFS

VNX Data Mover disruption in vSphere environments can result in application unavailability, and guest operating system crash.

In the event of a Data Mover disruption, the guest OS loses its connection to the NAS datastore on the VNX file system. Virtual machine I/O requests to virtual disks in the NAS datastore experience Disk SCSI Timeout errors in the OS system event viewer.

Use the following best practices on the guest OSs to keep the application and virtual machines available during VNX Data Mover outage events:

To avoid the downtime caused by the VNX Data Mover outage events:

- ◆ Configure the environment with at least one standby Data Mover to avoid a guest OS crash and application unavailability.
- ◆ Configure the Data Mover and ESX host to take advantage of DNS round-robin for NFS path fault tolerance.
- ◆ Install the VMware tools for the guest OS.
- ◆ Set the disk timeout value to at least 60 seconds in the guest OS.
 - For a Windows OS, modify the **HKEY_LOCAL_MACHINE/System/ControlSet/Services/DISK** and set the TimeoutValue to 120. The following command performs the same task and can be used for automation on multiple virtual machines:

```
reg.exe add
\\%1\HKLM\SYSTEM\CurrentControlSet\Services\Disk /V
TimeoutValue /t /REG_DWORD /d 120 /f"
```

Monitor and manage storage

vSphere makes it possible to proactively monitor storage utilization through vCenter datastore alarms. Datastore monitoring is particularly useful when using thin provisioned VNX storage. It helps prevent out-of-space conditions when thin virtual disks are provisioned on thin LUNs.

This section explains how to proactively monitor the storage utilization of vSphere datastores within vCenter and use EMC VSI for VMware vSphere Storage Viewer. It also explains how to monitor the utilization of the underlying VNX file system LUNs when they are thinly provisioned through Unisphere.

Note: As described in [“EMC VSI for VMware vSphere” on page 20](#), Storage Viewer exposes the datastore and VNX storage details. Use the information presented in Storage Viewer to configure VNX file system and LUN monitoring through Unisphere.

Monitor datastores using vCenter

Use the vSphere Client to display the current utilization information for NFS and VMFS datastores. Configure vCenter to trigger datastore alarms that occur in response to events, conditions, and state changes of datastores within the inventory. Create and modify the alarms from a vSphere Client connected to a vCenter Server. Datastore alarms, as shown in [Figure 64 on page 120](#), can be set for a single datastore, a host, or an entire datacenter.

Complete the following steps to create a datastore alarm:

1. From vSphere Client, select the datastore to monitor.
2. Right-click the datastore and then select **Add Alarm**.
3. Click **General** and then type the required properties:
 - a. Type the alarm name and description.
 - b. In the **Monitor** list box, select **Datastore**.
 - c. Select **Monitor for specific conditions or state**, for example thin LUN utilization.
 - d. Add a trigger to warn at 80 percent capacity, and to alert at 90 percent capacity.

- e. Add an action to generate email notifications when the condition occurs.

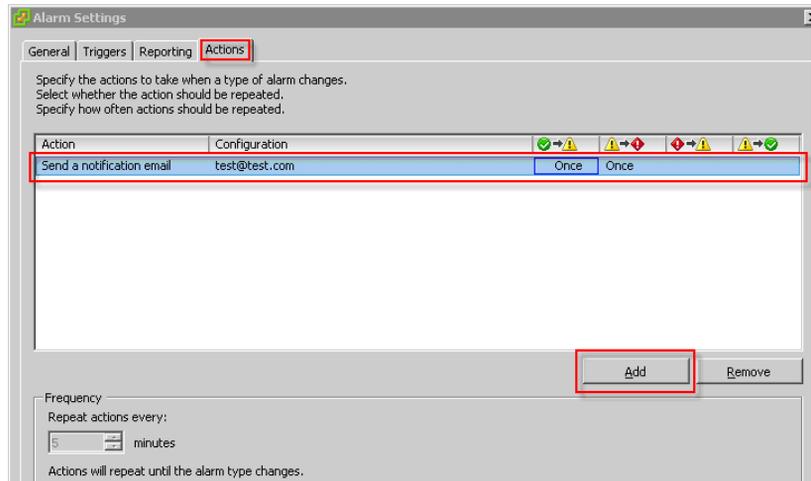


Figure 64 Data Alarm Settings—Actions window

When VNX Thin Provisioning is in use, it is important to correlate the storage information presented in vCenter with the storage utilization from the storage array. EMC Storage Viewer feature does this from within the vSphere Client.

To accomplish this task, complete the following steps:

1. From vSphere Client, select an ESXi host.
2. Click the **EMC VSI** tab. This tab lists three subviews of EMC storage information in the **Features Navigation** panel: **Datstores**, **LUNs**, and **Targets**.
3. Click **Datstores**. The Storage Viewer Datstores information appears on the right.

4. Select a datastore from the list of datastores. The **Storage Details** window lists the storage devices or the NFS export that back the selected datastore.

Note: The highlighted **VP** column in the **Storage Details** pane has a value of **Yes** if Thin Provisioning is enabled on the LUN. [Figure 65](#) shows the information that appears in Storage Viewer for a VMFS datastore provisioned on a VNX LUN.

The screenshot shows the Storage Viewer interface with the 'Datstores' pane selected. The 'Storage Details' pane is expanded for the 'VNX-Thin' datastore, showing a 'VP' (Thin Provisioning) status of 'Yes'.

Identification	Status	Device	Capacity	Free	Type
vnv-vm-library	Warning	DGC Fibre Channel Disk (naa.600601601ae02900debc23ec3547e011):1	999.75 GB	352.41 GB	vmfs3
VNX-Thin	Normal	DGC Fibre Channel Disk (naa.600601601ae029000e195124a54ae011):1	99.75 GB	99.20 GB	vmfs3
vnv-infrastructure	Alert	DGC Fibre Channel Disk (naa.600601601ae02900d43a1fce3e47e011):1	699.75 GB	165.46 GB	vmfs3
VM_Library	Warning	DGC Fibre Channel Disk (naa.60060160b3402200fa447ca34f89df11):1	999.75 GB	169.93 GB	vmfs3
nfs	Normal	10.244.156.216:/nfs		55.03 GB	NFS

Instance Name	Model	Array	Device Name	Device ID	RAID	Capacity	Group	META	VP	Paths	Owner	Policy
nhba4:C0:T4:L7	VNX5500	FNMM00103900200	LUN 16	00016	RAID_5	100.00 GB	ESXCluster_15_16		Yes	2/2	PowerPath	CLAROpt

Figure 65 Storage Viewer\Datstores window—VMFS datastore

Thin Provisioning enables physical storage to be over-provisioned. The expectation is that not all users or applications require their full storage allotment at the same time. They can share the pool and conserve storage resources. However, it is possible that applications may grow rapidly and request storage from a storage pool with insufficient capacity. This section describes a procedure to avoid this condition with VNX LUNs.

Unisphere monitors storage pool utilization and displays the current space allocations. Administrators can add alerts to objects to be monitored with the Event Monitor, and send alerts via email, page, or SNMP traps. Unisphere provides the following:

- ◆ Usable pool capacity is the total physical capacity available to all LUNs in the storage pool.
- ◆ Allocated capacity is the total physical capacity currently assigned to all thin LUNs.
- ◆ Subscribed capacity is the total host-reported capacity supported by the pool.

When LUN allocations begin to approach the capacity of the pool, the administrator receives alerts. Two non-dismissible pool alerts are provided:

- ◆ A warning event is triggered when the pool exceeds a user-defined value between 1 and 84.
- ◆ A critical alert is triggered when the pool reaches 85 percent.

Both alerts trigger a user-defined, associated secondary notification.

Complete the following steps to configure a user-defined alert on the storage pool:

1. Access EMC Unisphere.
2. In the **Systems** list box, select the VNX platform.
3. Select **Storage > Storage Configuration > Storage Pools for Blocks**. The **Pools** window appears.
4. Select the storage pool for which to set the alert. Click **Properties** to display the **Storage Pool Properties** window.
5. Click the **Advanced** tab.

In the **Percent Full Threshold** list box, type or select a value as the threshold at which to generate an alert.

In Figure 66, the **Percent Full Threshold** value in the **Advanced** tab of the **Storage Pool Properties** dialog box is set to 70 percent. Alerts are sent when the utilization of the storage pool reaches 70 percent.

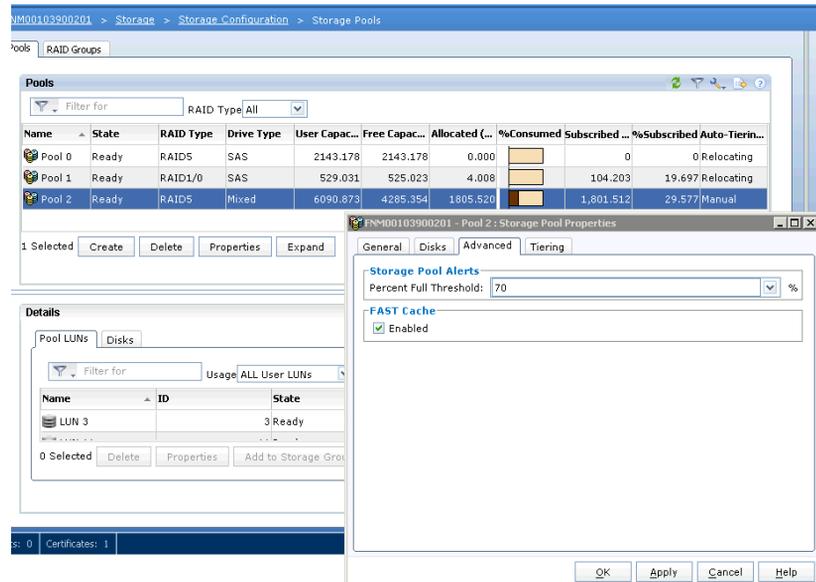


Figure 66 Adjustable percent full threshold for the storage pool

Adding drives to the storage pool non-disruptively increases the available usable pool capacity.

Note: Allocated capacity is only reclaimed by the pool when LUNs are deleted. Removing files or freeing space within a virtual machine disk does not free space within the pool. Monitor thinly provisioned file storage on VNX with EMC Unisphere.

Administrators must monitor the space utilization in over-provisioned storage pools and thinly provisioned file systems to ensure that they do not become full and deny write access. Configure and customize notifications based on the file system, storage pool usage, and time-to-fill predictions. Notifications are particularly important when over-provisioned resources exist in the environment.

Use VNX file system notifications to proactively monitor VNX file systems used for NFS datastores and generate SMTP (email) or SNMP (network management) alerts when an event occurs.

Multiple notification settings can be applied to the same resource to provide information about a trend or a worsening condition.

Configure VNX file system storage usage notification

Complete the following steps to configure a notification based on the percentage used of the maximum capacity:

1. Access EMC Unisphere to select the VNX platform.
2. Select **System > Monitoring and Alerts > Notifications for Files**.
3. Click **Storage Usage** and then click **Create**. The **Create Storage Usage Notification** window appears as shown in [Figure 67](#).

The screenshot shows the 'Create Storage Usage Notification' window with the following configuration:

- Storage Type:** File System (selected)
- Storage Resource:** nfs
- Storage Capacity:** 244.1 GB
- Auto Extend Enabled:** Yes
- Thin Enabled:** Yes
- High Water Mark:** 80 %
- Maximum Capacity:** 281.4 GB
- Frequency:** 24 hours
- Notify On:** Maximum Size (selected)
- Condition:** 80 % Used
- Notify Only If Over-Provisioned:**
- E-mail Addresses for Notification:** test@emc.lab
- SNMP Traps for Notification:** snmpmanager; <community name>

Figure 67 Create Storage Usage Notification window

4. Specify the storage information:
 - a. In the **Storage Type** field, select **File System**.
 - b. In the **Storage Resource** list box, select the name of the file system.

 Note: Notifications can be added for all file systems.

- c. Select **Maximum Size**.

 Note: Maximum Size is the autoextension maximum size and is valid only for file systems with autoextend enabled.

- d. In the **Condition** field, type the percentage of storage (percent used) and then select **% Used** from the list box to the right.

 Note: Select **Notify Only If Over-Provisioned** to trigger the notification only if the file system is over provisioned. If this is not selected, a notification is sent every time when the condition is met.

- e. Type the email or SNMP address, which consists of an IP address or hostname and community name. Separate multiple email addresses or trap addresses with commas.
 - f. Click **OK**. The configured notification appears in the **Storage Usage** window as shown in [Figure 68](#).

ID	Storage Type	Storage Resource	Notify On	Condition	Notify Only If Over-Prov...	Destination
W0	File System	nfs	Maximum Size	80% Used	No	test@emc.lab,snmpmanager;public
W1	Storage Pool	clarsas_archive	Current Size	80% Used	No	test@emc.lab

Figure 68 User-defined storage usage notifications

Configure VNX file system storage projection notification

Complete the following steps to configure notifications for the projected file-system-full time:

1. Access EMC Unisphere and select the VNX platform.
2. Select **System > Monitoring and Alerts > Notifications for Files**.
3. Click **Storage Usage** and then click **Create**.

4. Specify the storage information:
 - a. In the **Storage Type** field, select **File System**.
 - b. In the **Storage Resource** list box, select the name of the file system.

Note: Notifications can be added for all file systems.

- c. In the **Warn Before** field, type the number of days to send the warning notification before the file system is projected to be full.

Note: Select **Notify Only If Over-Provisioned** to trigger this notification only if the file system is over provisioned.

- d. Specify optional email or SNMP addresses.
 - e. Click **OK**. The configured notification is displayed in the **Storage Projection** window as shown in [Figure 69](#).

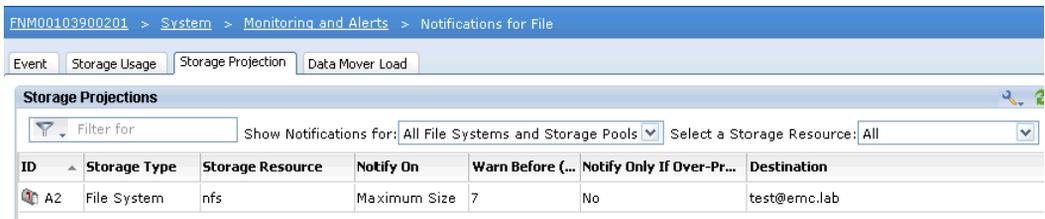


Figure 69 User-defined storage projection notifications

Note: There is no comparable capability in Unisphere for block storage. VSI provides a useful way to monitor space utilization for block storage.

VNX storage system resource monitoring

Coincident with the release of vSphere 5.1, two new options are provided for monitoring the resource utilization of the VNX storage system.

EMC VNX Monitoring and Reporting

EMC introduced the VNX Monitoring and Reporting product to help customers quickly identify and understand storage utilization and workload patterns. The product collects data from one or more VNX systems and stores it into a database to be used for problem diagnosis, trend analysis, and capacity planning.

VNX Monitoring and Reporting includes a web interface that users can access to view VNX storage system Inventory, Performance information, Capacity Planning metrics, and Health information.

Figure 70 provides an example of a capacity planning report which illustrates the storage system utilization over the past month.

Space Capacity Planning / Usable Capacity by Array

2012, October » November, the 26 at 3:42 PM EST | Last 1 Month: average on 1 day

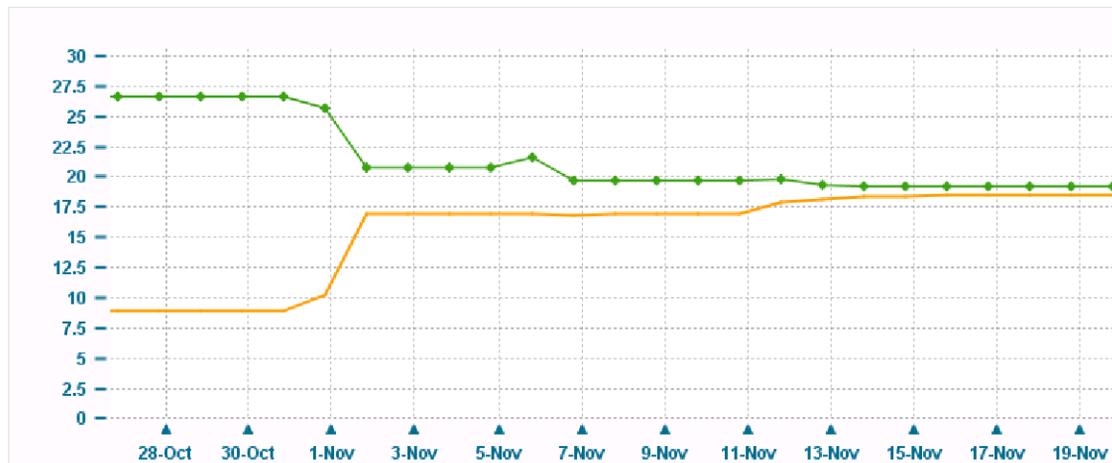


Figure 70 VNX Monitoring and Reporting - Capacity Planning Report

Figure 71 on page 128 illustrates a performance report for the same system. The output is generated by the Top IOPs report. This report lists the top 5 consumers. The N represents a user selectable value that defines how many entries you want to display on each page. Top N IOPs for the storage pools lists the top 6 storage pools and RAID groups in the system along with the current values for throughput and bandwidth.

TopN & Exceptions / TopN IOPS

November 2012, Wednesday 14 → Thursday 15, 1:59 PM EST | Last 1 Day

A list of different storage components, ordered by their I/O rate in descending order.

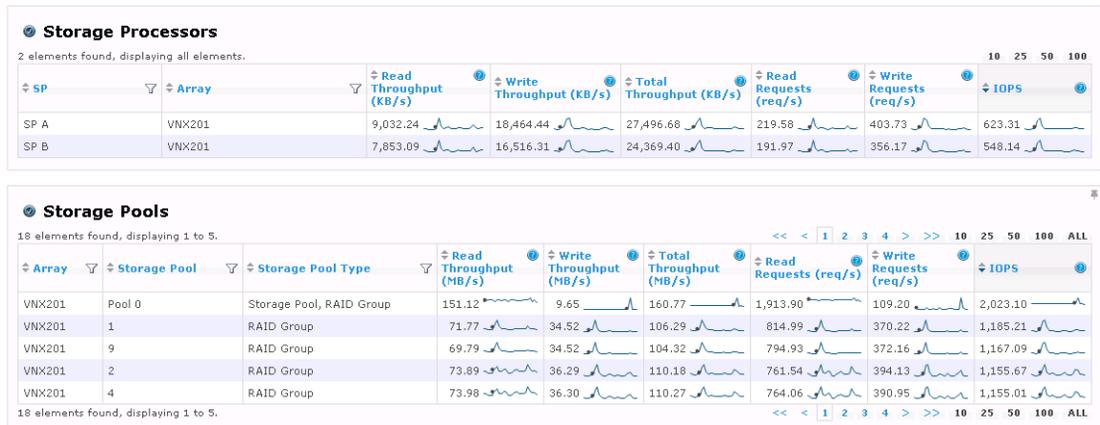


Figure 71 VNX Monitoring and Reporting - Performance report

The value of this product is that it simplifies the collection and presentation of performance information. Data is presented in easily understandable graphs with views of global and isolated resources so that the user easily identifies a potential resource imbalance or utilization problem.

The product provides key performance indicators that are normally obtained through Unisphere Analyzer. Users of Analyzer will be familiar with the metrics and use VNX Monitoring and Reporting to complement analyzer through automated collection and reporting.

VNX Analytics for vCenter Operations Manager

The second monitoring feature is an extension to VMware vCenter Operations Manager. vCenter Operations (vC Ops) provides a comprehensive view into the resources within the vSphere environment. It offers comprehensive monitoring of host, virtual machine, Network, and storage utilization metrics. It applies patented analytics to establish normal conditions and infer a health score for each resource. [Figure 72](#) illustrates the The vCenter Operations Manager dashboard interface that quickly identifies the state of the environment. Each component is identified assigned a numerical value and color (green, yellow, red) to indicate its health state.



Figure 72 vCenter Operations Manager Dashboard

EMC has developed an adapter for vCenter Operations Manager connector to allow vC Ops to collect and store information about the VNX storage system. vC Ops polls the VNX for utilization metrics at five minute intervals and stores the results within the vCenter Operations Manager database for up to 30 days.

In addition to monitoring the array status, the VNX connector provides metrics for the resource types as shown in [Table 7](#).

Table 7 VNX Connector metrics

Resource types for block	Resource types for file
Storage Processor	Data Mover
FAST Cache	NFS export
Storage Pool	File System
RAID group	File Pool
LUN	Disk volume (dVol)
Disks	

vCenter Operations Array Block and File performance interfaces illustrated in Figure 73 allow administrators to view performance metrics in real time. Administrators use the information to identify potential resource imbalance or over-utilization conditions and take measure to adjust or rbalance resources on the storage system or vSphere environment.

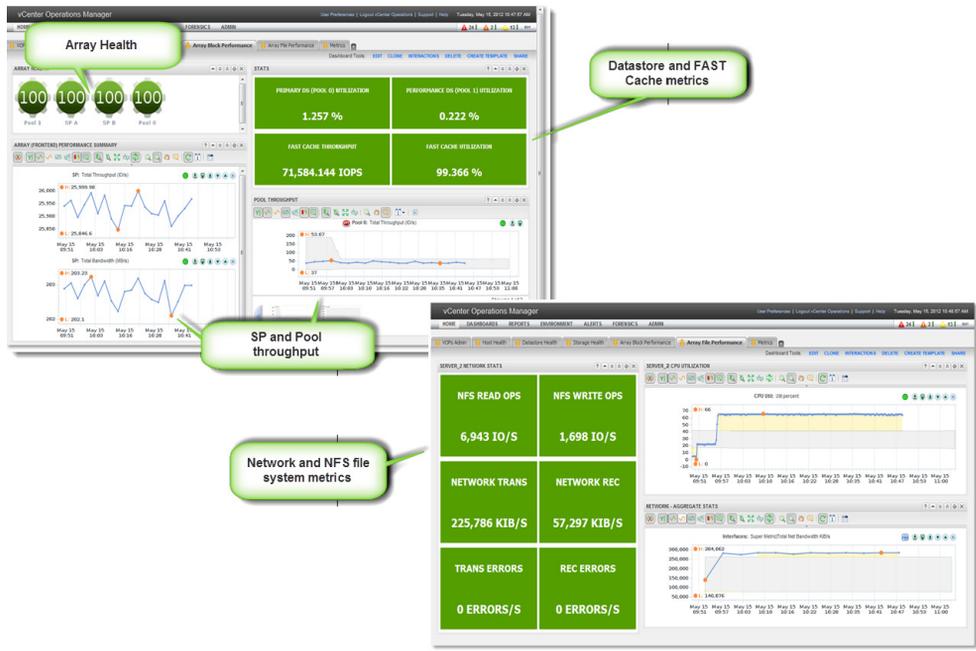


Figure 73 vCenter Operations Manager - VNX Storage Analytics

As cloud computing evolves, understanding how resources are consumed across the environment and how one resource impacts another will become increasingly important. Products such as EMC Monitoring and Reporting and VNX Analytics Suite will provide the information to assist in quickly identifying the root cause of a performance problem and allow you to relate that to applications and services in the environment. For more information on these products, check the EMC website (<http://emc.com>)

Storage efficiency

Thin Provisioning and compression are practices that administrators use to store data more efficiently. This section describes how to use these technologies in an environment with vSphere and VNX.

Thinly provisioned storage

Thin Provisioning is a storage efficiency technology that exists within VMware vSphere and EMC VNX. With Thin Provisioning, the VNX presents the host with a storage device that is not fully allocated. VNX performs an initial allocation with a portion of the device capacity. Additional space is consumed on an as-needed basis by the user, applications, or operating system. When using vSphere with VNX, the following Thin Provisioning combinations exist:

- ◆ On ESXi, through ESXi Thin Provisioning
- ◆ On VNX file systems, through thinly provisioned VNX file systems
- ◆ On VNX block LUNs, through VNX thin LUNs.

Monitor the storage utilization to prevent an accelerated out-of-space condition when Thin Provisioning is in use. For thin virtual disks on thin LUNs, the storage pool is the authoritative resource for storage capacity. Monitor the pool to avoid an out-of-space condition.

Virtual machine disk allocation

vSphere 5 renamed the disk identifiers used to provision new virtual disks. This document covers vSphere 4 and 5, however, the new names will be used and [Table 8](#) provides the reference to describe virtual disks in vSphere 4.

Table 8 Command line descriptions for vSphere 4 and vSphere 5

vSphere 5	vSphere 4	Command line
Flat	Thick	ZeroedThick
Thick	Fault Tolerant	EagerZeroedThick
Thin	Thin	Thin

VMware offers three options to provision a virtual disk. They are thin, flat (ZeroedThick), and Thick (Eagerzeroedthick). Table 9 provides a description of each, along with a summary of their impacts on VNX storage pools. Any supported VNX storage device (thin, Thick, VNX OE, or NFS) can provision any of the options.

Table 9 Virtual machine disk allocation policies

Allocation mechanism (virtual disk format)	VMware kernel behavior	Impact on VNX pool
Thin Provisioned (NFS default)	Does not reserve any space on the VMware file system on creation of the virtual disk. The space is allocated and zeroed on demand.	Minimal initial VNX pool allocation. Allocation is demand-based.
Flat (Zeroedthick) VMFS default	All space is reserved at creation, but is not initialized with zeros. The allocated space is wiped clean of any previous contents on the physical media. All blocks defined by the block size of the VMFS datastore are initialized on the first write.	Reserves .vmdk size within the LUN or pool. Allocation occurs when blocks are zeroed by the virtual machine.
Thick Provisioned (Eagerzeroedthick)	Allocates all the space and initializes every block with zeros. This allocation mechanism performs a write to every block of the virtual disk.	Full allocation of space in the VNX storage pool. No thin benefit.
RDM	Creates a virtual disk as a mapping file that contains the pointers to the blocks of the SCSI disk it maps. The SCSI INQ information of the physical media is virtualized. This format is commonly known as the "Virtual compatibility mode of raw disk mapping."	Allocation depends on the type of file system or application.
pRDM	Similar to the RDM format except that the SCSI INQ information of the physical media is not virtualized. This format is commonly known as the "Pass-through raw disk mapping."	Allocation depends on the type of file system or application.

Thinly provisioned block-based storage

Thin LUNs are the only devices that support oversubscription. Thin LUNs are created from storage pools that delay block allocation until an application or guest operating system needs the blocks to preserve space. The space for a thick LUN is always reserved so there are no thin-provisioning benefits. Similarly, the blocks assigned for VNX OE LUNs are always allocated within RAID groups with no option for thin provisioning.

In this section, the discussion of block-based thin provisioning focuses exclusively on VNX thin LUNs for VMFS or RDM volumes.

VMFS datastores are thin-friendly, which means that a VMware file system on a thin LUN uses a minimal number of extents from the storage pool. A VMFS datastore reuses previously allocated blocks, and thereby benefits from thinly provisioned LUNs. For RDM volumes, the file system of the guest OS dictates whether the RDM volume is thin-friendly.

Virtual machine disk provisioning options with block storage

The default .vmdk format with vSphere is flat. This format does not initialize or zero all blocks and claim all the space during creation. RDM volumes are formatted by the guest OS. Therefore, virtual disk options such as **flat**, **thin**, and **thick** apply only to VMFS volumes.

From an allocation standpoint, space is reserved at the VMFS level, but it is not allocated until the blocks within the .vmdk are zeroed. [Figure 74](#) shows that a 500 GB .vmdk is created and 100 GB is written to the disk. These actions result in 500 GB of file space reserved from the VMFS file system and 100 GB of space allocated in the VNX storage pool. The flat option provides some performance benefits in allocation time and potential space utilization within the storage pool. The blocks cannot be compressed again after they are allocated.

Note: Quick Format helps to preserve storage space. If a Windows file system is formatted with NTFS, each block is zeroed, which performs a full allocation at the storage pool level.

Use the Quick Format option for NTFS volumes to preserve space.

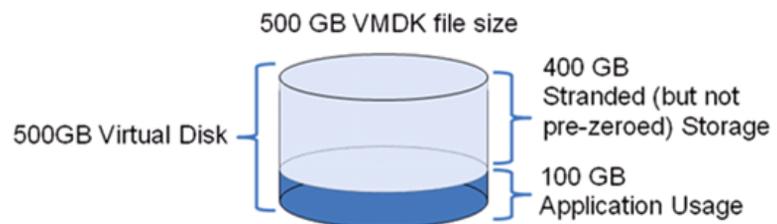


Figure 74 Thick or zeroed-thick virtual disk allocation

Use thin virtual disks to preserve space within the VMFS datastore. The thin .vmdk only allocates VMFS blocks that the virtual machine needs for guest OS or application use. Create thin .vmdks on a thick LUN to preserve space within the file system, or on a thin LUN to extend that benefit to the storage pool. [Figure 75](#) shows the same 500 GB virtual disk within a VMFS volume. This time the disk is created in a thin-provisioned format. With this option, the VMFS uses only 100 GB within the file system and 100 GB within the VNX storage pool. Additional space is allocated when the virtual machine needs it. The allocation unit is the equivalent of the block size for the VMFS datastore. Instead of allocating at the 4k or 8k block that the virtual machine uses, the minimum allocation size for ESXi is 1 MB, which is the default block size for a VMFS volume, and scales up to 4 MB, which is the maximum block size used by VMFS. This is beneficial for a thin-on-thin configuration.

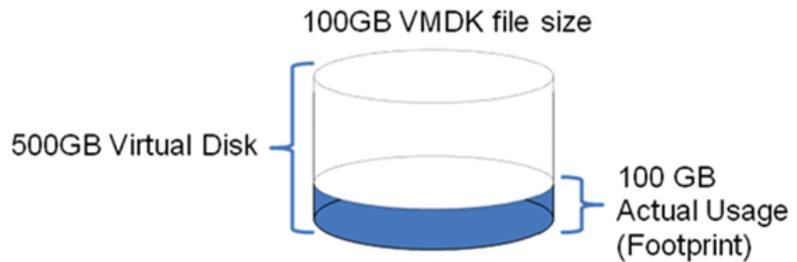


Figure 75 Thin virtual disk allocation

Flat virtual disks are the default option to create a SCSI virtual disk. Accepting the default creates a flat or zeroedthick virtual disk. Figure 76 shows that you should select one of the other options if you need a thick or thin disk.

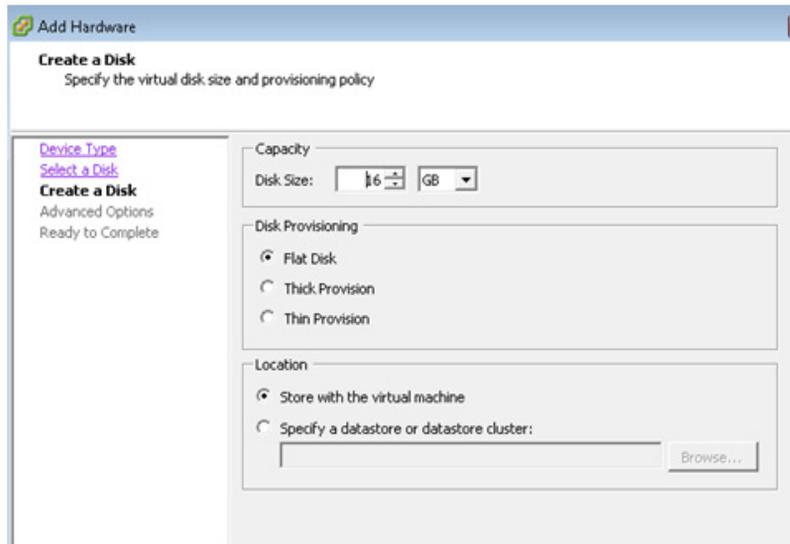


Figure 76 Virtual machine disk creation wizard

Selecting the flat option for virtual disks on VMFS volumes affects the space allocated to the guest file system, or the write pattern of the guest OS device. If the guest file system initializes all blocks, the virtual disk needs all the space to be allocated up front. When the first write occurs on a flat virtual disk, it writes zeros on the region defined by the VMFS block size, not just the block that was written to by the application. This behavior affects the performance of array-based replication software because more data, which is not required, must be copied based on the VMFS block size. However, it also alleviates some of the concerns about fragmentation with a thin-on-thin configuration.

In ESXi, configure a virtual machine disk as flat or thin. With the thin virtual disk format, the VMFS datastore is aware of the space the virtual machine consumes. However, continue to monitor the VMFS datastore's free capacity to avoid an out-of-space condition; vSphere provides a simple alert when a datastore reaches its threshold.

In addition, with ESXi, the zeroedthick or thin format remains intact on the destination datastore after the use of vCenter features such as Cloning, Storage vMotion, Cold Migration, and Deploying a Template. The consumed capacity of the source virtual disk is preserved on the destination virtual disk, and is not fully allocated.

Because the virtual machine is not thin-aware, the potential exists to encounter an out-of-space condition when the storage pool that backs a thin LUN reaches its full capacity. If the thin LUN cannot accommodate a new write request from the virtual machine due to an out-of-space error, ESXi pauses the virtual machine I/O and generates a pop-up message in the vSphere Client that alerts the user to the problem. [Figure 77](#) shows the alert.

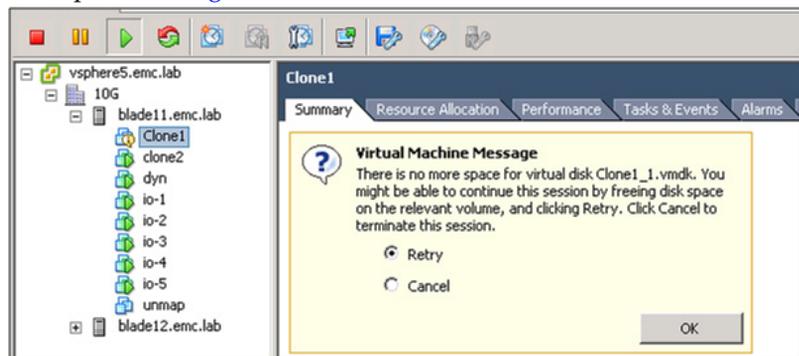


Figure 77 Virtual machine out-of-space error message

The virtual machine does not generate any I/O while in this state.

Do not select **Retry** as this results in repeated failures until additional capacity is added for the thin LUN via a storage pool expansion or removing space from another virtual machine or LUN consumer.

- a. Select **Cancel** to power off the virtual machine.
- b. Select **Retry** to resume the virtual machine after adding or reclaiming additional storage capacity.
- c. Restart any applications that time out while waiting for storage capacity to become available. Select **Cancel** to power off the virtual machine.

File-based thinly provisioned storage

File-based thin provisioning with VNX is available by using VNX Thin Provisioning for file systems. Both USM and Unisphere can set up Thin Provisioning on a file system.

Thin Provisioning and Automatic File System Extension are enabled by default.

Automatic File System Extension on the file system is controlled by the High Water Mark (HWM) value in the **Advanced** window for provisioning NFS datastores on new NFS exports, as shown in [Figure 78 on page 139](#). This value (percentage) determines when to extend the file system. By default, VSI sets the HWM to 90 percent. This means that the file system extends itself when 90 percent of the capacity is consumed. The NFS datastore is created by VSI, and presented to the VMware ESXi host with the file system maximum capacity.

The ESXi host is unaware of the currently allocated capacity in the file system. However, the Storage Viewer feature of EMC VSI makes it possible to view the currently allocated capacity of the file system.

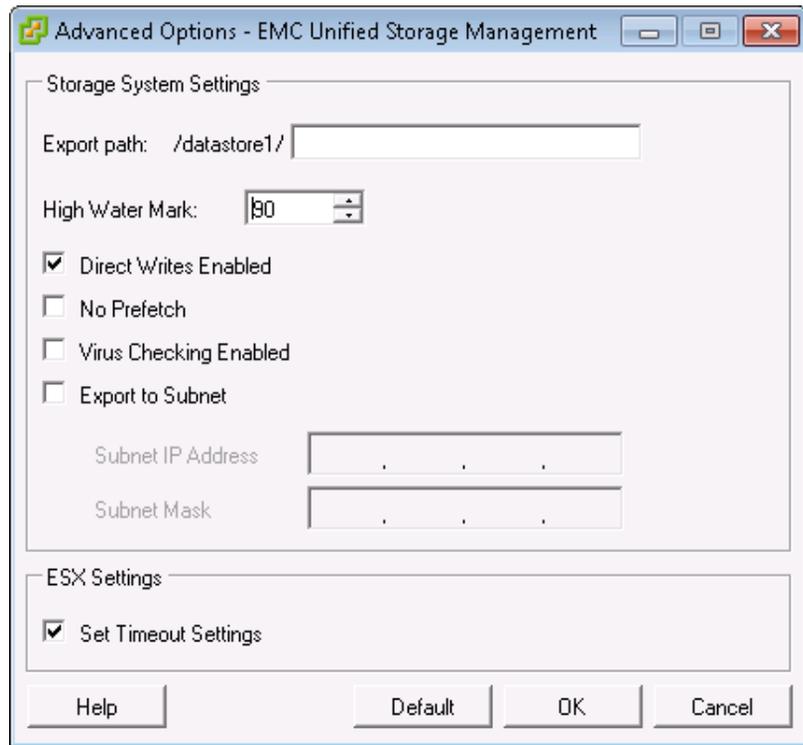


Figure 78 File system High Water Mark in the EMC VSI: USM feature

Additional virtual machines can be created on the datastore even when the aggregated capacity of all their virtual disks exceeds the datastore size. Therefore, it is important to monitor the utilization of the VNX file system to identify and proactively address upcoming storage shortages.

Note: [“Monitor and manage storage” on page 119](#) provides further details on how to monitor the storage utilization with VMware vSphere and EMC VNX.

The thin provisioned virtual disk characteristics are preserved when a virtual machine is cloned or migrated to another datastore, or when its virtual disk is extended.

VNX-based block and file system operations that affect a datastore are transparent to the virtual machine disks stored in them. Virtual-provisioning characteristics of the virtual disk are preserved during all the operations listed above.

VMware vSphere virtual disks based on NFS storage are always thin provisioned. [Figure 79](#) shows the virtual disk provisioning policy settings for NFS.

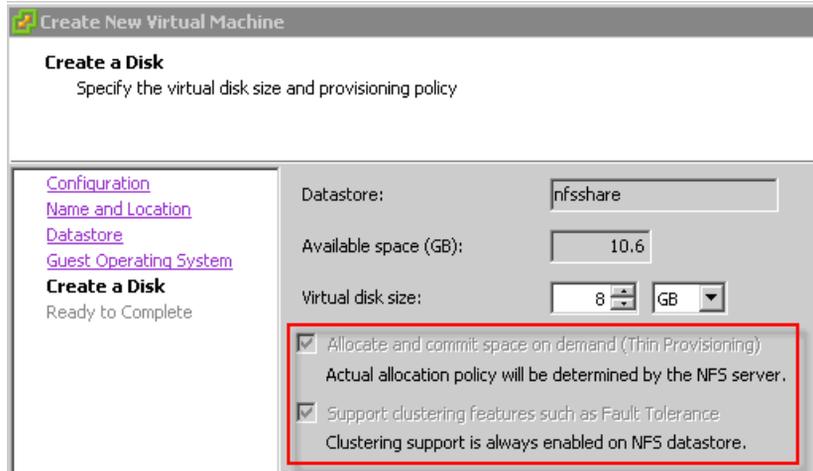


Figure 79 Provisioning policy for an NFS virtual machine virtual disk

LUN compression

VNX LUN compression offers capacity savings to the users for data types with lower performance requirements. LUNs presented to the VMware ESXi host are compressed or decompressed as needed. [Figure 80 on page 141](#) shows that compression is a LUN attribute that users enable or disable for each individual LUN. When enabled, data on disk is compressed in the background. If the source is a RAID group LUN or thick pool LUN, it undergoes an online migration to a thin LUN when compression is enabled. Additional data written by the host is initially stored uncompressed, and system-defined thresholds are used to automatically trigger asynchronous compression of any new data.

Hosts decompress data in memory to read it, but the data remains compressed on disk. These operations are largely transparent to the end user, and the system automatically processes new data in the background when compression is in use.

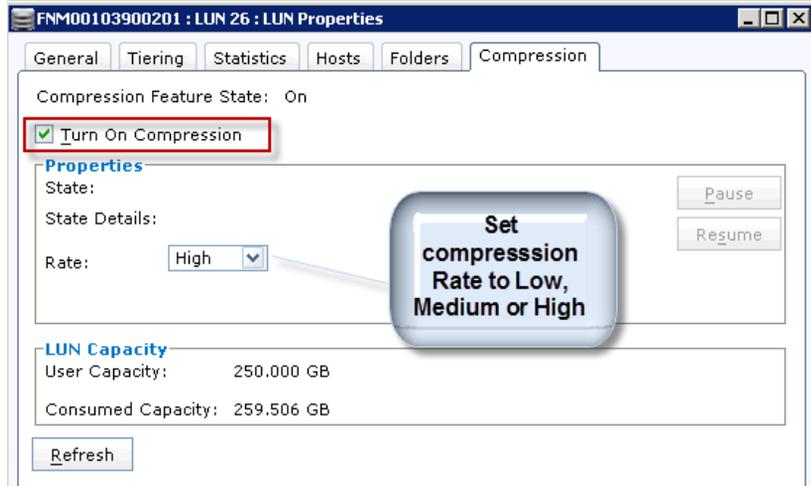


Figure 80 LUN compression property configuration

The inline read and write operations of compressed data affect the performance of individual I/O threads. Do not compress in the following cases:

- ◆ I/O-intensive or response-time-sensitive applications
- ◆ Active database or messaging systems

Compression is successfully applied to more static data sets such as archives (virtual machine templates), non-production clones of databases, or messaging system volumes that run on virtual machines.

If compression is disabled on a compressed LUN, the entire LUN is processed in the background. When the decompression process completes, the LUN remains a thin LUN and remains in the same pool. Capacity allocation of the thin LUN after decompression depends on the original pre-compression LUN type.

File deduplication and compression

The VNX file deduplication and compression feature provides data reduction for files through data compression and data deduplication. The main objective of VNX file compression is to improve file-storage efficiency by compressing files stored on a VNX file system. Deduplication eliminates redundant files in a file system with minimal impact to the end user. The use of these technologies result in a lower cost-per-megabyte, and an improved total cost of ownership of the VNX.

VNX file deduplication and compression provide data-reduction cost savings capabilities in two usage categories:

- ◆ Efficient deployment and cloning of virtual machines that are stored on VNX file systems over NFS.
- ◆ Efficient storage of file-based business data stored on NFS or CIFS network shares accessed by virtual machines.

Deployment of virtual machines stored on NFS datastores

VNX file deduplication and compression targets active virtual disk files to compress. This feature is available for VMware vSphere virtual machines that are deployed on VNX-based NFS datastores.

Virtual machine compression with VNX file deduplication and compression

With this feature, the VMware administrator compresses a virtual machine disk at the VNX level to reduce the file system storage consumption by up to 50 percent. There is some CPU overhead associated with the compression process, but VNX includes several optimization techniques to minimize this performance impact.

Virtual machine cloning with VNX file deduplication and compression

VNX file deduplication and compression provides the ability to perform efficient, array-level cloning of virtual machines. Two cloning alternatives are available:

- ◆ **Full clone** — This operation creates a full virtual machine clone that is comparable to a native VMware vSphere clone operation. A full VNX virtual machine clone operation is performed on the storage system instead of the ESXi host to save the ESXi CPU cycles required to perform the native cloning operation. The result is an efficient virtual machine clone operation that is up to two or three times faster than a native vSphere virtual machine clone operation.
- ◆ **Fast clone** — This operation clones only the blocks that are changed between the replica and the source virtual machine. This is very similar to a VNX LUN snapshot operation, except that in this case the operation is done at the file level instead of at the LUN-level. A fast clone resides in the same file system as the source virtual machine. The source files satisfy unchanged block reads, and the fast clone files deliver the updated blocks. Fast Clone creation is an almost instantaneous operation because no data needs to be copied from the source virtual machine to the target device.

All of the compression and cloning operations available in VNX file deduplication and compression are virtual machine-based rather than file-system-based. This provides the administrator with the flexibility to use VNX file deduplication and compression with VMware vSphere to further increase VNX storage efficiency.

The *EMC VSI for VMware vSphere: Unified Storage Management Product Guide* provides further information on how to efficiently compress and clone virtual machines with USM and VNX file deduplication and compression.

Efficient storage of file-based business data stored on NFS/CIFS network shares that are mounted or mapped by virtual machines

VNX file deduplication and compression also eliminates redundant files to provide a high degree of storage efficiency with minimal impact on the end user experience. This feature also compresses the remaining data.

VNX file deduplication and compression automatically targets files that are the best candidates for deduplication and subsequent compression in terms of the file-access frequency and file size. In combination with a tiered storage architecture, VNX file deduplication and compression can also run on the secondary tier to reduce the size of the archived dataset.

With VMware vSphere, VNX file deduplication and compression run on file systems that are mounted or mapped by virtual machines that use NFS or CIFS. This is most suitable for business data such as home directories and network-based shared folders. Similarly, use VNX file deduplication and compression to reduce the space consumption of archived virtual machines to eliminate redundant data and improve the storage efficiency of the file systems.

VNX storage options

VNX provides a wide range of configuration options to meet the needs of any vSphere environment. VNX is flexible enough to support basic configurations for general environments, and, advanced capabilities for specific configurations required in some environments. This section provides an overview of the different storage components and configuration options.

VNX supported disk types

[Table 10 on page 146](#) illustrates the current drive types offered for the VNX platform and includes general recommendations for suggested use. The drives within the system are organized into storage pools and RAID groups. Solid state drives provide an additional option as an extended SP cache when FAST Cache is configured on the system.

Table 10 VNX supported disk types

Type of drive	Available size	Benefit	Suggested Usage	Notes
Flash	<ul style="list-style-type: none"> • 100 GB • 200 GB 	<ul style="list-style-type: none"> • Extreme performance • Lowest Latency 	Virtual machine applications with low response time requirements	EFDs are not recommended for small block sequential I/O, such as log files
Serial Attached SCSI (SAS)	<ul style="list-style-type: none"> • 300 GB • 600 GB • 10k rpm • 15k rpm 	<ul style="list-style-type: none"> • Cost-effective • Better performance 	<ul style="list-style-type: none"> • Large-capacity, high-performance VMware environments • Most tier 1 and 2 business applications, such as SQL and Exchange 	
NL-SAS	<ul style="list-style-type: none"> • 1 TB • 2 TB • 3 TB • 7200 rpm 	Performance and reliability equivalent to SATA drives	<ul style="list-style-type: none"> • High-capacity storage • Archived data, backups, virtual machine template, and ISO images area • Good solution for tier 2/3 applications with low throughput and medium response-time requirements, such as infrastructure services DNS, AD, and similar applications 	

Storage pools

VNX provides two types of disk grouping; RAID groups and storage pools. Both options organize physical disks into logical groups, however, they support different LUN types with different functional capabilities.

RAID groups offer the traditional approach to storage management that predates storage pools. The key characteristics of RAID groups are, they support up to 16 disks and RAID group LUNs that reserve and allocate all disk blocks at creation time.

Storage pools offer more flexible configuration options in terms of the number of disks, space allocation, and LUN types. Pools provide advanced features such as cost-effective thin provisioning and self-adjusting tiered storage options. Pools can be created to support single or multitiered storage configurations created from any supported drive type.

Pool LUNs support the following features:

- ◆ Thick or thin provisioned LUNs
- ◆ Expansion without metaLUNs
- ◆ LUNs that can be shrunk
- ◆ Block Compression (with compression enabler)
- ◆ Auto-tiered (with FAST enabler installed)
- ◆ Dead space reclamation

Pool storage results in more fluid space utilization within each pool. Free space within the pool is dynamic and fluctuates along with the storage requirements of the virtual machines and applications. FAST VP simplifies LUN configuration, allowing the pool to support different service levels and workloads with multiple tiers of storage.

Storage pools versus RAID groups

The primary differences between storage pools and RAID groups are:

- ◆ RAID groups are limited to 16 disks. Larger disk configurations are possible using metaLuns.
- ◆ Pools can be created with higher disk counts for simplified storage management.
- ◆ Pools support thin LUNs (TLUs).
- ◆ When configured for FAST VP, pools can use a combination of any disk types.
- ◆ Pools support LUN compression.
- ◆ Storage pools are segmented into 1 GB slices. Pool LUNs are created using multiple slices within the pool.

Table 11 lists the capabilities of RAID groups and storage pools.

Table 11 Pool capabilities

Pool Type	Types	Allocation	Max Disks	Expandable	Compression	Unmap	Shrink	Auto Tiering
RAID group	FLARE LUN	Full	16	N	Y	N	N	N
Storage pool	Thin (TLU)	No allocation	71 - 996 Determined by platform	Y	Y	Y	Y	Y
	Thick (DLU)	No allocation space is reserved						

Note: MetaLUNs provide the ability to extend RAID groups. Enabling LUN compression converts the existing LUN to a thin pool LUN. FLARE LUNs can be shrunk when running Windows 2008 with Solutions Enabler.

Although pools are introduced to provide simplicity and optimization, VNX preserves RAID groups for internal storage devices used by data protection technologies, and environments or applications with stringent resource requirements.

RAID configuration options

VNX provides a range of RAID protection algorithms to address the performance and reliability requirements of VMware environments. All block and file devices use VNX RAID protection. [Table 12](#) lists the RAID protection options.

Table 12 VNX RAID options

Algorithm	Description	RAID group	Pools	Considerations
RAID 0	Striped RAID	X		No data protection
RAID 1	Data is striped across all spindles	X		Uses 1 mirror disk for each data disk.
RAID 1/0	Data is mirrored and striped across all spindles	X	X	Uses 1 mirror disk for each data disk. Consumes more disk space than distributed parity.
RAID 3	Data is striped, with a dedicated parity disk	X		
RAID 5	Data is striped with distributed parity among all disks	X	X	Parity RAID provides the most efficient use of disk space to satisfy the requirements of the applications.
RAID 6	Data is striped, with distributed double parity among all disks.	X	X	Additional parity computation results in additional write latency.

Note: Current configurations for NL-SAS devices suggest the use of RAID 6, limiting their use with mixed pools.

Choose the storage and RAID algorithm based on the throughput and data protection requirements of the applications or virtual machines. The most attractive RAID configuration options for VMFS volumes are RAID 1/0, and RAID 5. Parity RAID provides the most efficient use of disk space to satisfy the requirements of the applications. In tests conducted in EMC labs, RAID 5 often provides the broadest coverage of storage needs for virtual machines. An understanding of the application and storage requirements in the computing environment will help identify the appropriate RAID configuration.

Storage pool features

FAST VP

Fully Automated Storage Tiering for Virtual Pools (FAST VP) is configured using a combination of two or more disk types listed in [Table 12 on page 149](#). FAST VP identifies the drive type by performance tier. Tier names are:

- ◆ Extreme Performance (Solid State Disks)
- ◆ Performance (SAS)
- ◆ Capacity (NL-SAS)

Flash provides the highest performance with the lowest capacity. NL-SAS provides the best capacity and lowest cost, and SAS disks provide a performance tier that is a blend of both.

Note: Rotational speed is not differentiated within a FAST VP tier. Therefore, disks with different rotational speeds such as 10k and 15k RPM SAS drives are assigned to the same pool tier. EMC does not recommend this configuration.

LUNs created within the pool are distributed across one or more storage tiers. FAST VP operates at a subLUN level using a one GB segment called a slice. When a LUN is created slices are distributed across the available tiers within the pool. The policy assigned to the pool and existing tier utilization, determines the slice distribution for the LUN.

FAST VP pools perform slice relocation to align the most frequently used storage with the highest tier, and the less frequently used storage with the lowest tier. Slice rebalancing occurs automatically at scheduled periods of the day, or is manually completed by an administrator.

VNX OE for Block version 5.32 performs slice rebalancing within a tier when a pool is expanded or when the software that monitors the slices identifies hot spots on private LUNs within the storage pool. The slice rebalance at EMC labs showed minimal performance impact during pool expansion, and improved performance benefits when the slice rebalancing is completed.

FAST VP is beneficial because it adjusts to the changing data access patterns in the environment as block usage patterns change within the vSphere environment.

Unisphere provides configuration guidance for all pool creation tasks. FAST VP pools are bound in multiples of five disks for RAID 5 pools, eight disks for RAID 1/0 pools, and eight disks for RAID 6 pools.

Pool expansion should adhere to these configuration rules, and grow in similar increments to the existing configuration to avoid parity overhead and unbalanced LUN distribution. For example, if the existing pool configuration is made up of 20 disks, the pool should be expanded with 20 disks for even extent distribution of LUNs within the pool.

Figure 81 shows the Unisphere tiering window. The window indicates that 47 GB of data is identified for migration to the Performance tier, and 28 GB will be moved to the Extreme Performance tier. In this example, the pool-tiering policy is set to Manual. The administrator must manually initiate the relocation for the migration to occur.

Block relocation with FAST VP is not generally performed in real time. Depending on the workload, it is best to schedule the relocation to occur during periods of lower use, or off hours.

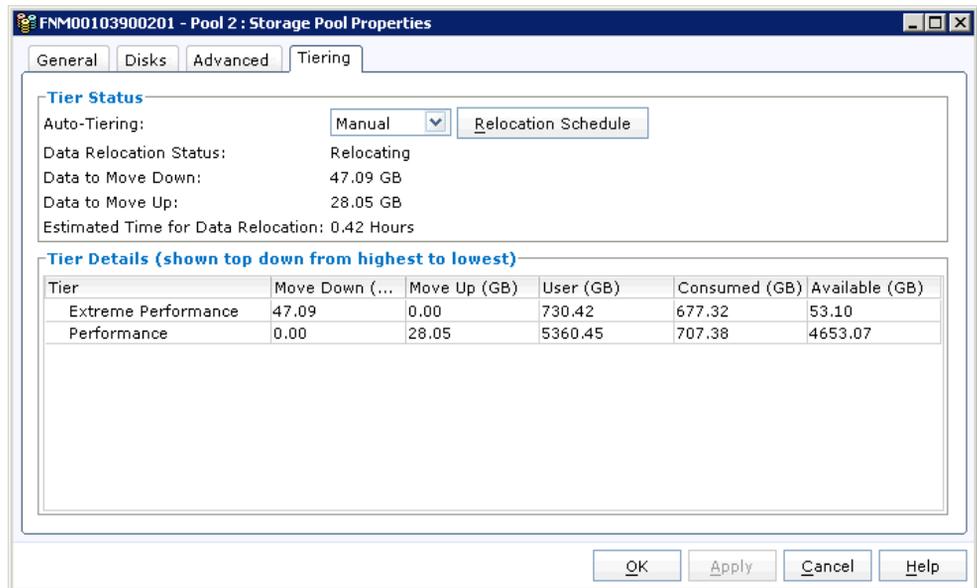


Figure 81 VNX FAST VP reporting and management interface

FAST Cache

FAST Cache is an optimization technology that greatly improves the performance of applications such as databases within VMware environments. FAST Cache uses Solid State disks to store the most frequently used data within the system. FAST Cache operates at a 64 KB extent size. If a block within an extent is accessed multiple times within a system-defined interval, the extent is promoted to the Flash disks where subsequent access requests result in a significant performance improvement.

As access to data blocks within cached extents becomes less frequent, or block priorities change, they are de-staged to HDD and replaced with the higher priority extents. FAST Cache operates in real time, which results in more frequent migration of extents to match the access requirements of the virtual machines.

Note: FAST Cache is best suited for data reuse. Applications with heavy reuse or hot regions achieve more benefit than those that perform sequential reads or writes. If your application is more sequential in nature, configure the SSDs as part of a FAST VP pool to achieve better performance.

Advanced Snapshots

VNX OE for Block version 5.32 supports a new LUN snapshot capability called advanced snapshots. Advanced snapshots are used to create instantaneous snapshot images of storage pool LUNs.

Features of the advanced snapshots are:

- ◆ Provide the ability to create up to 256 copies of any source LUN.
- ◆ Create a snapshot of an existing snapshot
- ◆ Delete snapshots at any time, in any order.
- ◆ Support consistency groups for application consistent images of multiLUN storage devices.

Advanced snapshots do not perform copy-on-write operations which means there is very little overhead for write operations. They perform "allocate on writes" operations to write updated data to a new area within the storage pool.

VNX Snapshots do not require additional setup or reserved LUNs. Snapshots use available space within the storage pool.

VNX LUNs

[Table 11 on page 148](#) shows the different LUN types supported by each storage pool. The following comparison describes the LUN options available in VNX.

Thick LUNs

The default LUN created in a storage pool is called a Thick LUN. These LUNs consist of 1 GB slices which are distributed across storage groups within the pool.

Thick LUNs require three 1 GB slices for metadata. Based on the version of VNX OE for Block running on the system, the remaining slices are either reserved or allocated.

In releases of VNX OE for Block prior to 5.32, the remaining slices are reserved within the pool and additional slice allocation is performed when the virtual machine or host requires additional space within the LUN.

In VNX 5.32, Thick LUN space is allocated at creation time. This change improves the locality of blocks within the LUN.

A pool LUN uses a number of disks based on the size of the pool, available slices, and the LUN size. Pools perform initial slice placement based on available space. Depending upon how full a pool is, a LUN may not be striped across all disks that make up the pool, however, VNX OE for Block version 5.32 monitors slice activity and rebalances them to adjust the distribution of slices within the pool.

Thin LUNs

Thin LUNs (TLUs) are created within storage pools when the Thin enabler is installed. Thin LUNs require three 1 GB slices for metadata. Since the goal of thin LUNs is to preserve space, block allocation is deferred until a virtual machine or host requires additional space and new space is allocated at a more granular eight KB size.

To limit the amount of space consumed by Thin LUNs, their space is not reserved within the storage pool. This capability allows the storage pool to be oversubscribed with many TLUs whose configuration size may exceed the pool capacity. Thin LUNs should be used with "thin friendly" storage and applications. Thin LUNs work best when they either are not filled on a regular basis, or their capacity is dynamic filling for a period and then releasing the space back to the pool. If the potential exists that the LUNs you are configuring will all fill at the same time, they may not be a good candidate for TLUs. Oversubscribed storage pools should be monitored to detect when pool space is getting low. [“Monitor and manage storage” on page 119](#) provides more details on how to monitor VNX storage.

Thin LUNs versus Thick LUNs

Table 13 illustrates the major differences between thick and thin LUNs.

Table 13 Thin LUNs versus Thick LUNs

Thin LUNs	Thick LUNs
Allocate space at a more granular level using 8 KB increments to conserve storage space	Reserve and allocate (blocks in VNX 5.32) all of the required slices
Provide no reservation which means that all of the TLUs in the pool are sharing the free space of that pool	Favor performance
Favor space reuse, particularly with the Dead space reclamation functionality included in ESXi 5.0 U1 and later	

RAID group LUNs

RAID group LUNs are the traditional devices created from fixed disk groups. All disk blocks associated with an RG LUN are allocated in a contiguous manner when the LUN is created. RAID-group LUNs have a fixed drive limit of 16 with no thin LUN option.

VNX metaLUNs

A metaLUN is an aggregate LUN created by striping or concatenating multiple LUNs from different RAID groups. They allow VNX to present a single RAID group device that spans more than 16 disks to provide more resources for capacity or distribute the workload amongst more spindles when using RAID groups.

Pool LUN versus RAID group LUN performance

As described above, each LUN type uses a different allocation approach.

RAID LUNs allocate all blocks when the LUN is created, providing a higher probability that the LUNs will have good spatial locality or skew. This layout usually results in better LUN response times. RAID LUNs can use MetaLUNs to create aggregate devices with linear scalability. LUNs created from a RAID group offer the most predictable LUN performance.

As of VNX OE for Block version 5.32 Thick LUNs also perform all block allocation when created. This provides similar locality and performance to RAID group LUNs.

Thick LUNs created prior to VNX OE for Block 5.32 do not perform an initial allocation and may have spatial locality and response times which are marginally different than the RAID group LUNs. Depending on the configuration, thick LUNs have up to 10 percent performance overhead when compared to RAID-group LUNs.

Thin LUNs preserve space on the storage system by deferring block allocation until the space is required. This can impact the response time for thin LUNs, and could result in a difference of 20 percent or more when compared with a RAID-group LUN.

VNX File volumes

VNX OE for File version 7.1 uses the same LUNs and LUN types described in the introduction section, to create NFS file systems for vSphere environments. VNX LUNs are imported into the file environment as disk volumes or dvols. VNX OE for File volume manager is used to create aggregate, stripe, and slice dvols to create file systems that are presented to ESXi as NFS datastores.

Therefore most LUN properties and features described in this document apply to file system storage for NFS datastores as well.

VNX provides two approaches to volume creation, Automated Volume Management (AVM) and Manual Volume Management (MVM). AVM provides templates to automate the creation of volumes and VNX file systems. It simplifies the creation by applying best practice algorithms to the existing storage resources.

The second option, MVM, enables the storage administrator to select which components are used to create the volume for additional flexibility and precise configuration of an NFS volume.

VNX volume management allows administrators to:

- ◆ Create customized volumes for file system storage.
- ◆ Group, combine, and slice volumes to meet specific configuration needs.
- ◆ Manage VNX volumes and file systems and LUNs through a single interface.

AVM generated volumes meet the requirements for most VMware deployments.

MVM is best suited to file system configurations with specialized application requirements. MVM provides an added measure of control for precise selection and layout of the storage configuration. The MVM interface allows the creation of file systems with different characteristics.

Unisphere exposes a set of configuration wizards that allow the administrator to reserve LUNs exclusively for the file environment. The **Disk Provisioning Wizard** illustrated in Figure 82 allows the storage administrator to define pools of storage for file provisioning.

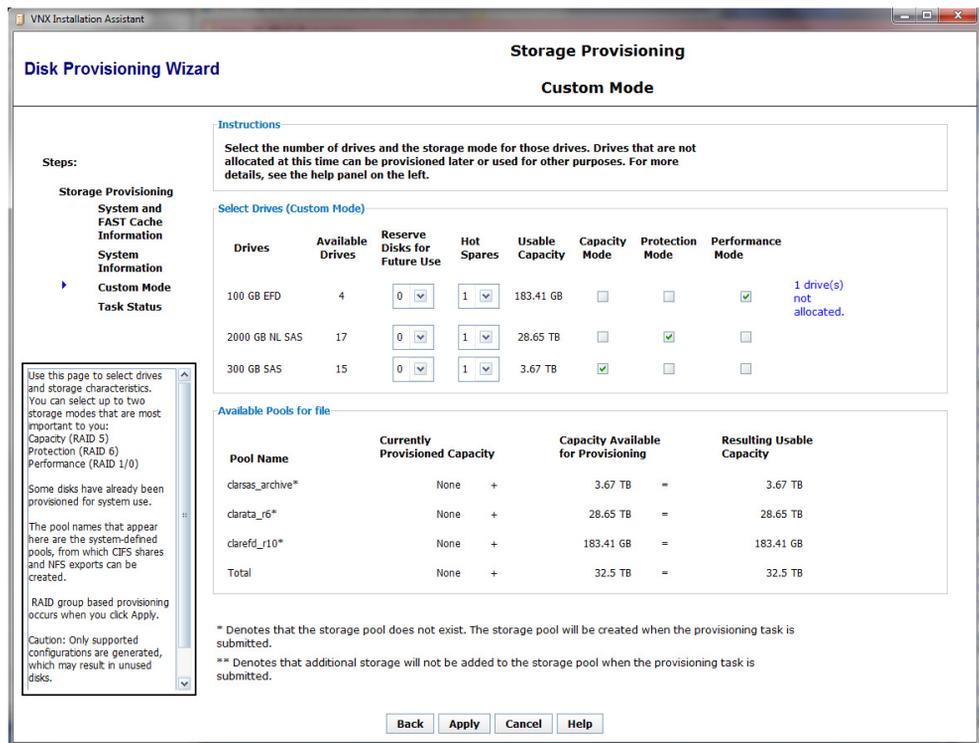


Figure 82 Disk Provisioning Wizard

This chapter includes the following topics:

- ◆ Introduction 160
- ◆ Using EMC VNX cloning technologies..... 162
- ◆ Summary 176

Introduction

Virtualized environments benefit from the ability to quickly create replicas of existing virtual machines. The two types of vCenter initiated virtual machine replicas are:

- ◆ Full virtual machine replicas or clones that are block-for-block copies of a virtual machine and its virtual disks
- ◆ Snapshot replicas that are typically thin journal file images, or block/file system pointer-based images of the files that constitute the virtual machine and its virtual disks

VMware provides the following native replication capabilities to clone virtual machines through the Clone Virtual Machine wizard in vCenter, and the VMware vCenter Converter Standalone utility:

- ◆ **Clone Virtual Machine wizard** — Enables users to create a clone of an existing virtual machine and store it on any supported datastore accessible to the ESXi host. The wizard also provides the option to clone the virtual disks using a different allocation policy, such as thin, to preserve the amount of space within a datastore.
- ◆ **vCenter Converter** — Enables users to convert any Windows system to a virtual machine on an ESXi host. It also provides the ability to clone an existing virtual machine, and optionally, to resize existing virtual disks. This tool is invaluable for resizing operating system disks with minimal downtime and administrative effort.

In most cases the native snapshot and replication wizards within vCenter provide the best virtual machine replication option. They offer integrated vCenter functionality to automate and register the virtual machine replicas.

EMC provides alternative replication options to create and register virtual machine replicas on NFS datastores, and create datastore replicas on VNX storage devices.

VNX provides the following features for virtual machine clones:

- ◆ VNX SnapView™ for block storage when using the FC, iSCSI, or FCoE protocols.
- ◆ VNX SnapSure™ for file systems when using the NFS protocol.
- ◆ VMware VAAI technology for block storage to accelerate native virtual machine cloning.
- ◆ VAAI plug-in for NFS to perform space-efficient FAST virtual machine clones on NFS datastores.
- ◆ VSI Unified Storage Management for individual virtual machine cloning.

Using EMC VNX cloning technologies

This section explains how to use the EMC VNX software technologies to clone virtual machines. The VNX platform-based technologies produce exact copies of the storage devices that back the vSphere datastores and RDM virtual disks.

To produce reliable storage system clones, take the following precautions prior to creating a clone of a VNX storage device:

- ◆ Shut down or quiesce applications running on the virtual machines to commit all data from memory to the virtual disk.
- ◆ Use Windows System Preparation tool, Sysprep, or a comparable tool to place the virtual machine in a deployable state.

Assign a unique virtual machine hostname and network address to avoid identity conflicts with other virtual machines. For Windows virtual machines, run Sysprep within the guest operating system to automatically generate a new security identifier and network address upon system boot.

Replicating virtual machines with VNX SnapView

VNX SnapView technology creates copies of VMFS datastores or RDM LUNs that support virtual machines.

SnapView enables users to create LUN-level copies for testing, backup, and, recovery operations. SnapView includes three flexible options:

- ◆ **Pointer-based, space-saving snapshots** — SnapView snapshots use pointer-based technology to create point-in-time images of existing LUNs. SnapView maintains the snapshot image contents by copying source LUN blocks before updates are applied to the source LUN. A single source LUN can have up to eight snapshots to capture the contents of the LUN over a period of time.

- ◆ **VNX advanced snapshots** — VNX OE for Block version 5.32 supports advanced snapshots to create up to 256 snapshots of pool-based LUNs. An advanced snapshot is a pointer-based copy of the source LUN, however, modified blocks are not written to the snapshot. Advanced snapshots maintain the pointers to the original blocks and new blocks are allocated to accommodate block changes to the source LUN. Advanced snapshots write updates to the LUN within the storage pool and do not require a separate reserved LUN pool.
- ◆ **Full-volume clones** — SnapView clones are full-image copies of a source LUN that can be used for almost any business purpose. SnapView tracks the block changes of the device. This resynchronizes a clone device with the source with changes from a prior synchronized state. A LUN can have up to eight simultaneous target clones.

Replicating virtual machines on VMFS datastores with SnapView clones

VNX LUNs are formatted as VMFS datastores or surfaced to virtual machines as RDM volumes. SnapView Clone can be used to replicate the VMFS datastore by creating an identical block-for-block replica of a LUN used by ESXi to create a VMFS datastore.

SnapView cloning is managed through Unisphere or Navisphere® CLI. [Figure 83](#) illustrates the interface used to create and present a cloned LUN to an ESXi host.

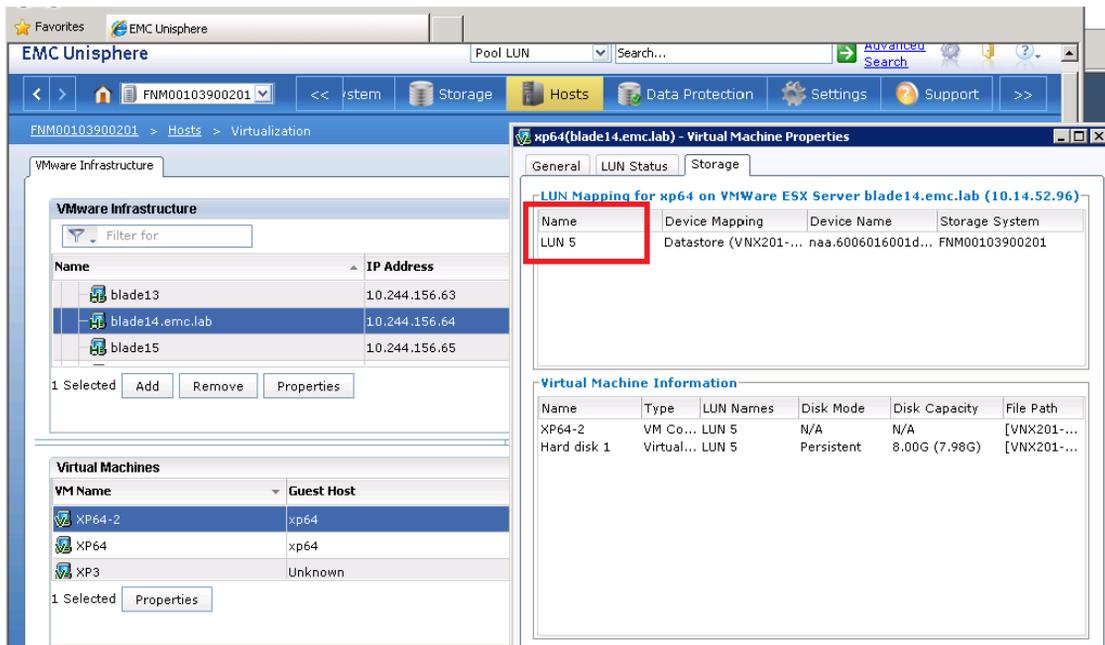


Figure 83 Unisphere clone LUN management

Complete the following steps to create and present a cloned LUN:

1. Use Unisphere Host Virtualization interface or the EMC VSI Storage Viewer feature to identify:
 - a. The VNX LUN that supports the VMFS datastore
 - b. The virtual machines contained within the datastore
2. Define a clone group for each VNX LUN to be cloned.
3. Add clone target LUNs to each clone group.

The addition of the target devices automatically starts the SnapView clone synchronization process.

4. Fracture the clone volumes from the source volumes after they have synchronized. This step preserves the current LUN state and sets the LUNs to a read/write state so the LUNs can be accessed by an ESXi host.

It is possible to create multiple VNX clones of the same source LUN. To make use of the clone, fracture it from the source LUN and present it to a storage group as shown in [Figure 84](#). Any ESXi host that is part of the storage group is presented with a consistent read/write copy of the source volume at the time it was fractured.

Note: To perform this task with the Navisphere CLI utility (`naviseccli`), specify the `-consistent` switch to perform a consistent fracture.

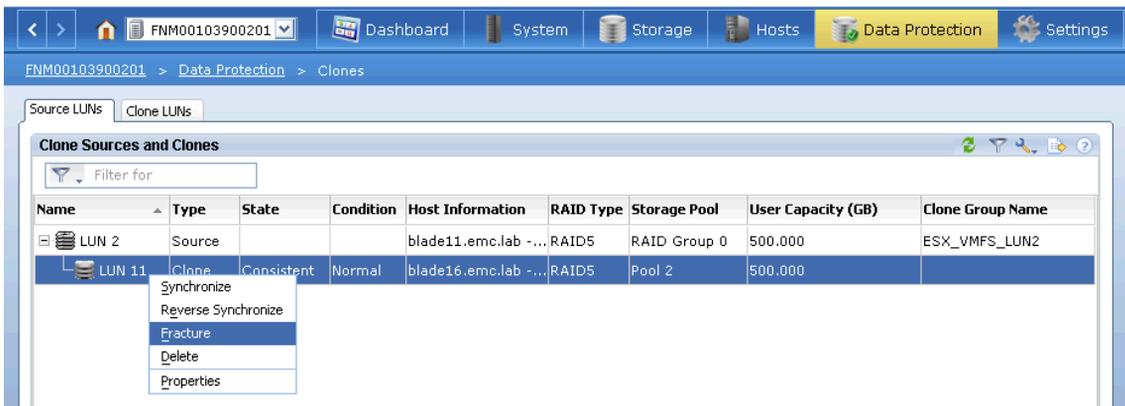


Figure 84 Performing a consistent clone fracture operation

Replicating virtual machines on VMFS datastores with SnapView Snapshot

To create and present SnapView snapshots, complete the following steps:

1. Use the Unisphere Host Virtualization interface to identify the source devices to snap.
2. Use Unisphere to create a SnapView snapshot of the source devices.

A Snapshot establishes the necessary storage resources for the snapshot LUN.

3. Use either Unisphere or Navisphere CLI, as shown in [Figure 85 on page 167](#), to start a SnapView session on the source device.

This step initiates the copy-on-write activity.

4. Access the SnapView session by activating the SnapView snapshot device session that was previously created.

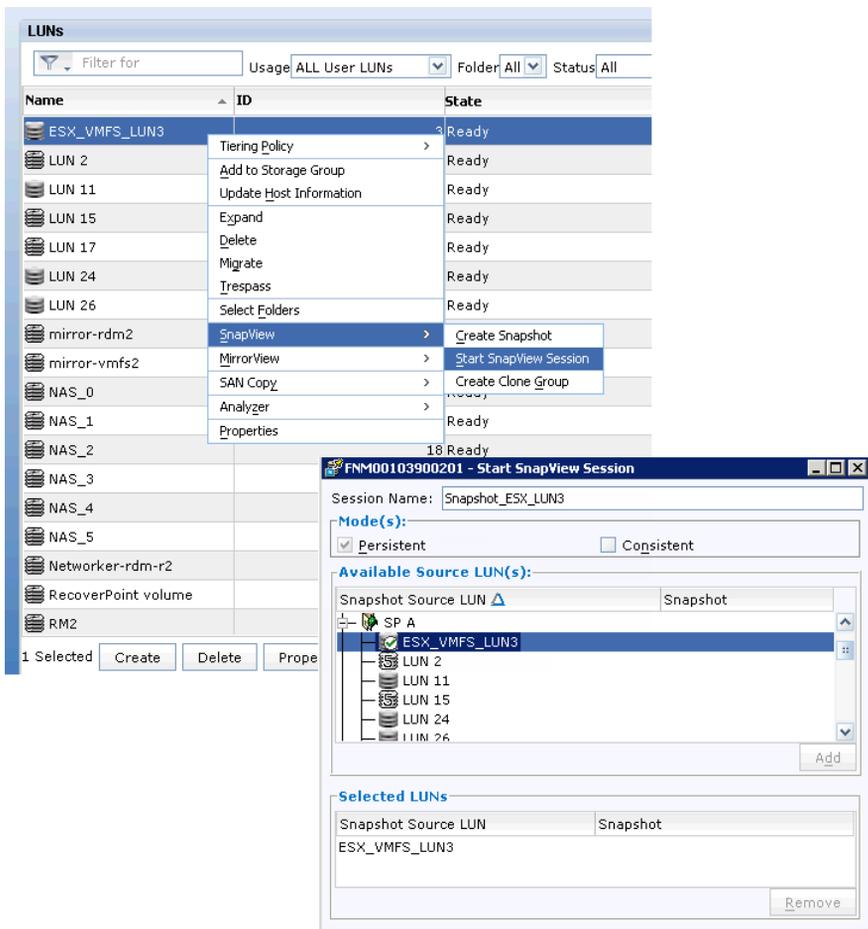


Figure 85 Creating a SnapView session to create a copy of a VMware file system

ESXi volume signatures

The ESXi VMkernel assigns a unique signature to all VMFS-formatted disks. The signature is based on the device ID of the LUN. It also includes user-assigned properties such as the datastore/volume name. A replicated VNX storage device is an exact block-for-block copy that includes the unique signature and volume details.

The VMkernel performs a SCSI device inquiry on all devices accessible to the host to discover the properties of the device and determine if there is an existing device signature. If vSphere detects that the device contains a signature of an existing device, it prevents it from being mounted and presents the option to use the LUN by assigning a new signature to the device. When presenting the replica to a host that is not part of the same cluster, keep the existing signature to mount the device.

After a rescan, the user can either keep the existing signature of the LUN replica or resignature the LUN replica if needed:

- ◆ **Keep the existing signature** — Presents the copy of the data with the same label name and signature as the source device. ESXi does not surface a replica when a signature conflict exists. Assign a new signature to activate the replica on the same host as the source LUN.
- ◆ **Assign a new signature** — Assigns a new signature to the VMFS volume replica. The new signature is computed using the UID and LUN number of the replica LUN. The default format of the new label assigned to the datastore is `snap-<snap_ID>-<old_label>`, where `<snap_ID>` is an integer and `<old_label>` is the label of the original datastore.

To resignature a SnapView clone or snapshot LUN, complete the following steps:

1. Rescan storage on the ESXi host to perform device discovery and update the SCSI device list.
2. Select the host from the **Inventory** area.
3. Select **Configuration**, and then click **Storage** in the **Hardware** area.
4. Click **Add Storage**.
5. Select the Disk/LUN storage type and then click **Next**.
6. Select the LUN, from the list of LUNs, that displays a datastore name in the **VMFS Label** column, and then click **Next**. The **Select VMFS Mount Options** dialog box appears.

Note: The name presented in the VMFS Label column indicates that the LUN is a copy of an existing vStorage VMFS datastore.

7. Select **Keep the existing signature** or **Assign a new signature**, as shown in [Figure 86](#), and then click **Next**.

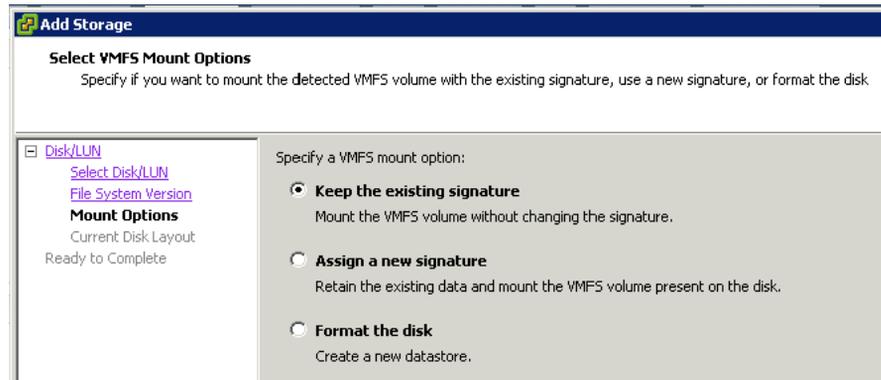


Figure 86 Device signature assignment

8. Review the datastore configuration information, and then click **Finish** to complete the task and add the datastore.
9. Browse the new datastore to locate the virtual machine's configuration (.vmx) file and import it to the vCenter inventory.

Replicating virtual machines with SnapView clones of RDM LUNs

Replicating an RDM volume requires a copy of the source virtual machine configuration files to facilitate access to the replicated RDM volumes. SnapView technology creates a logical, point-in-time copy of the RDM volume. In turn, the copy is presented to a virtual machine.

An RDM volume has a one-to-one relationship with a virtual machine or virtual machine cluster.

To replicate virtual machines with SnapView clones of RDM LUNs, complete the following steps:

1. Create a SnapView clone or snapshot of the RDM LUN.
2. Within vCenter, identify the ESXi host where the clone image will be created.
3. Create a folder within an existing datastore to hold the copy of the virtual machine configuration files.

- Use the **Datastore Browser** in the vSphere Client, as shown in [Figure 87](#), to copy the configuration files of the target virtual machine to the directory created in step 3.

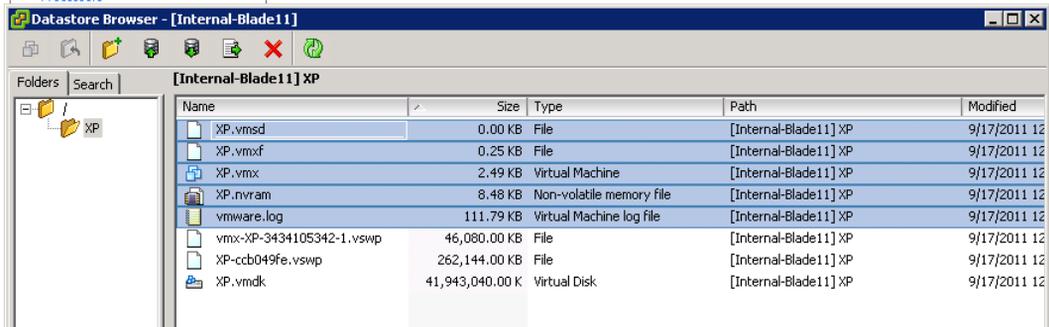


Figure 87 Selecting virtual machine configuration files in the Datastore Browser

- Identify the copy of the virtual machine configuration file (.vmx) and use it to add the new virtual machine to the inventory of the ESXi host, as shown in [Figure 88](#).

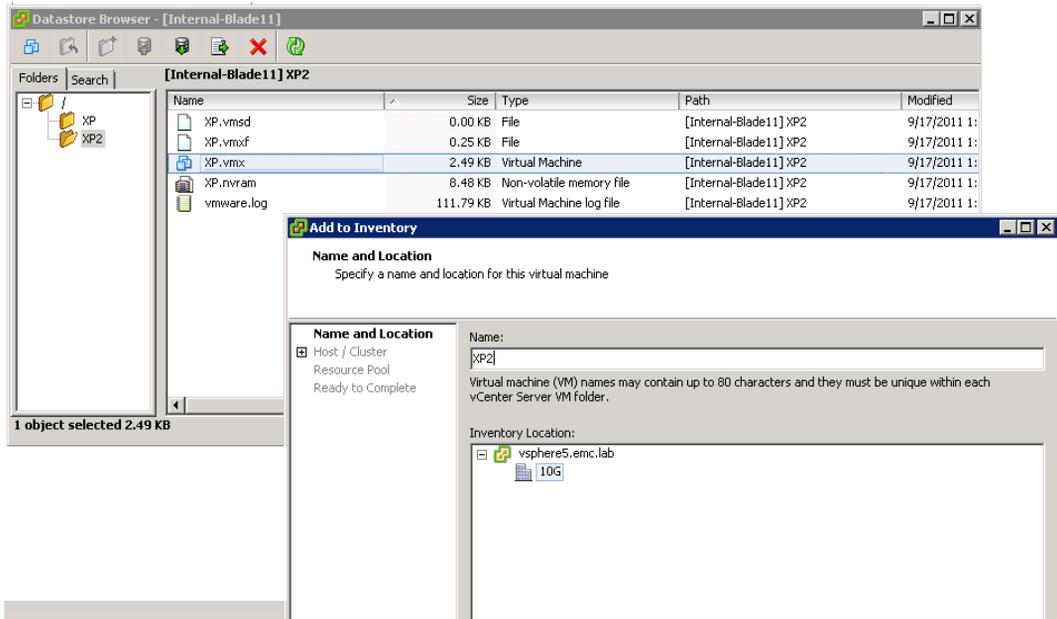


Figure 88 Adding the new virtual machine to the ESXi host inventory

6. Edit the following virtual machine settings:
 - a. Remove the existing Hard Disk entry referring to the source RDM.
 - b. Add a new hard disk as type RDM, and specify the cloned RDM device.
7. Power on the cloned virtual machine from the vSphere Client.

Cloning virtual machines on VNX NFS datastores with VNX SnapSure

The VNX SnapSure feature creates a logical, point-in-time image (checkpoint) of an NFS file system that supports an NFS datastore that contains virtual disks and virtual machine configuration files. The ESXi host requires the file system to be in read/write mode in order to boot the virtual machine. A writeable Checkpoint File System is created in Unisphere as shown in [Figure 89](#).

The screenshot shows the 'Create Checkpoint' dialog box in Unisphere. The 'Writeable Checkpoint' checkbox is checked and highlighted with a red box. The 'Production File System' is 'vSphere_NFS_Datastore'. The 'Data Movers' is 'server_2'. The 'Writeable Checkpoint Name' is 'vSphere_NFS_Datastore_ckpt'. The 'Current Storage System' is 'VNX VNX5500 FNM00103900201'. The 'Storage Pool' is 'clarsas_archive 335.9 GB (343918 MB)'. The 'Storage Capacity (MB)' is '20000'. The 'High Water Mark' is '90 %'. The 'Maximum Storage Capacity (MB)' is empty.

Figure 89 Creating a writeable NAS datastore checkpoint

Execute the following command in the CLI to create writeable Checkpoint File Systems:

```
# fs_ckpt <NAS_file_system_checkpoint> -Create -readonly n
```

To start the virtual machine, the VMkernel requires read/write and root access to the Checkpoint File System. “[Creating an NFS datastore using EMC Unified Storage Management](#)” on page 50 provides more details. Export the checkpoint file system to the ESXi hosts to provide them with root-level access.

To import multiple virtual machines on a Checkpoint File System, complete the following steps within the vCenter UI:

1. Select an ESXi host with access to the Checkpoint File System.
2. Select the **Configuration** tab, and start the **Add Storage Wizard**.
3. Add the writeable Checkpoint File System to the ESXi host as an NFS datastore.
4. Browse for the new datastore and add the .vmx files of the virtual machines to the vCenter inventory.

Cloning virtual machines with native vCenter cloning and VAAI

This section explains how vCenter virtual machine cloning works with VAAI-enabled VNX block storage. The VAAI operations preserve ESXi resources that are consumed if the host performs the clone. The resources used are proportional to the amount of data to be copied.

VAAI allows VMware vSphere 4.1 and later to take advantage of efficient disk-array storage functions as an alternative to ESXi host-based functions. These vStorage APIs enable close integration between vSphere and storage hardware to:

- ◆ Provide better quality of service to applications running on virtual machines.
- ◆ Improve availability through rapid provisioning.
- ◆ Increase virtual machine scalability.

vStorage API supports VMFS datastores, RDM volumes, and NFS systems with the VNX platform. The minimum VNX release versions for VAAI offload are VNX OE for Block 5.31 and VNX OE for File 7.0.45. The Full Copy feature of the VAAI suite offloads virtual machine cloning operations to the storage system.

Note: VAAI support is provided with VNX storage systems running VNX OE for Block version 5.31 and later.

ESXi hosts issue the **XCOPY** command to the array supporting the source and destination devices. The array performs internal data copy operations to create virtual disk replicas. The host issues copy operations to the array which performs the data movement. SCSI status messages are exchanged between the storage system for flow control and copy completion. The array copy offload results in a significant reduction of host I/O traffic and CPU utilization. The full copy feature is supported only when the source and destination LUNs belong to the same VNX platform.

Administrators find the full copy feature useful to:

- ◆ Create multiple copies of a virtual machine within or across LUNs on the same storage system.
- ◆ Storage vMotion virtual machines from one VMFS datastore to another when the LUNs reside on the same storage system.
- ◆ Deploy virtual machines from a template using VNX LUNs.

Cloning individual virtual machines on NFS datastores

vSphere 5.0 introduced VAAI support for NFS copy operations when cloning virtual machines on NFS datastores.

ESXi hosts configured with the EMC NAS software package offload copy operations to the VNX Data Mover. All replication or cloning is performed within the storage environment to minimize consumption of host and network resources.

The EMC NAS software package is required for this functionality. It is available to EMC customers and partners as a VMware Installation Bundle (VIB) from EMC Online Support.

VAAI offload for NFS reduces the amount of ESXi host resources required to perform the clone tasks. It also reduces network resource utilization on ESXi and VNX systems.

Install the EMC NAS VIB package from the ESXi console, or as an autodeploy image in vSphere. Use the vSphere Client, or run the following command, to verify that EMC NAS VIB is installed:

```
esxcli software vib list |grep EMCNas
```

Figure 90 illustrates the datastore properties of a VAAI-enabled NFS datastore from VNX that has been configured with the NFS plug-in.

Note: The datastore list denotes that Hardware Acceleration is supported.

The screenshot shows the vSphere Datastores table with the following data:

Identification	Status	Device	Drive Type	Capacity	Free	Type	Hardware Acceleration	Storage I/O Control
NFS	Normal	10.10.10.30:/NFS	Unknown	345.42 GB	313.55 GB	NFS	Supported	Enabled
thin-2	Warning	DGC iSCSI Disk (naa...	Non-SSD	49.75 GB	8.78 GB	VMFS5	Supported	Disabled
iSCSI	Normal	DGC iSCSI Disk (naa...	Non-SSD	1,023.75 G	403.04 GB	VMFS5	Supported	Disabled
Internal-Blade11	Normal	Local SEAGATE Disk ...	Non-SSD	63.25 GB	62.30 GB	VMFS5	Unknown	Disabled

Figure 90 Cloned NFS datastore in vSphere

NFS VAAI clones may not always result in a faster execution time than host-based clone operations. This is particularly true when tests are performed in isolation with no other load on the environment. The benefit of the offload operations is in the resource utilization and cumulative benefit when these operations are performed under contention for host resources, and not when the host is idle.

VNX also provides individual virtual machine cloning capabilities when the virtual machine resides on an NFS datastore. The VSI USM feature performs cloning operations directly within the storage system using a separate management approach from the VAAI cloning operations.

USM provides a set of utilities that include Full and Fast clones:

- ◆ **Full clone** — Full clone operations are performed across file systems within the Data Mover. By removing the ESXi host from the process, the virtual machine clone operation can complete two to three times faster than a native vSphere virtual machine clone operation.
- ◆ **Fast clone** — Fast clone operations are performed within a single file system. Fast clones are near-instantaneous operations executed at the Data Mover level with no external data movement. Unlike Full clones, Fast clone images only contain changes to the cloned virtual machines and reference the source virtual machine files for unchanged data. They are stored in the same folder as the source virtual machine.

The *EMC VSI for VMware vSphere: Unified Storage Management—Product Guide*, available on EMC Online Support, provides more information on the USM feature.

Summary

The VNX platform-based technologies provide an alternative to conventional VMware-based cloning. VNX-based technologies create virtual machine clones at the storage layer in a single operation. Offloading these tasks to the storage systems provides faster operations with reduced vSphere CPU, memory, and network resource consumption.

VNX-based technologies provide options for administrators to:

- ◆ Clone a single or small number of virtual machines and maintain the granularity of individual virtual machines.
- ◆ Clone a large number or all of the virtual machines with no granularity of individual virtual machines on a datastore or LUN.

Options for the VNX-based technologies are listed in [Table 14](#).

Table 14 VNX-based technologies for virtual machine cloning

Storage type	Individual virtual machine granularity for a small number of virtual machines	No granularity for a large number of virtual machines
Block storage (VMFS datastores or RDM)	VMware native cloning with VAAI Full Copy	VNX SnapView
Network-attached storage (NFS datastores)	VNX File Data Deduplication using the VSI Unified Storage Management feature	VNX SnapSure

Backup and Recovery Options

This chapter includes the following topics:

◆ Introduction	178
◆ Virtual machine data consistency	179
◆ VNX native backup and recovery options	181
◆ Snapshot backup and recovery of a VMFS datastore	183
◆ Backup and recovery of RDM volumes	186
◆ Replication Manager.....	187
◆ Backup and recovery of a VMFS with VNX Advanced Snaps..	192
◆ vStorage APIs for Data Protection.....	200
◆ Backup and recovery using VMware Data Recovery	201
◆ Backup and recovery using Avamar	204
◆ Backup and recovery using NetWorker.....	213
◆ Summary	219

Introduction

The combination of EMC data protection technologies and VMware vSphere offers several backup and recovery options for virtual environments. When considering backup solutions, determine a recovery point objective (RPO) and a recovery time objective (RTO) to ensure that an appropriate method is used to meet service-level requirements and minimize downtime.

This chapter discusses two types of data protection available at the storage layer: logical backup and physical backup.

A logical backup (snapshot) establishes a point-in-time image of the VNX file system or LUN. Logical backups are created rapidly and require very little storage space, allowing them to be created frequently. Restoring from a logical backup can also be accomplished quickly, dramatically reducing the mean time to recover. Logical backups protect against events such as file system corruption and accidental deletion of files.

A physical backup creates a full copy of the file system or LUN. The full backup provides a complete and independent copy of the source data. It can be managed and stored on devices that are separate from the source device.

A logical backup cannot replace a physical backup. Although full backup and recovery may require more time, a physical backup provides a higher level of protection because it guards against hardware failures.

Virtual machine data consistency

In ESXi environments supported by VNX storage, administrators can use the technologies described in this chapter to generate crash-consistent backups. In a simplified configuration all of the virtual machines and virtual disks are stored on a single datastore. Crash consistency is achieved by creating a replica of the LUN or file system supporting the datastore.

However, many application vendors, especially database vendors, recommend separating data and log files and distributing them across separate storage devices for better performance. When following these practices, treat all datastores that support the application as a single entity. VNX provides a method to achieve multidevice management through consistency groups. Consistency groups are used with VMware snapshots to provide crash consistency of block storage devices in these scenarios.

A VMware snapshot is a software-based virtual machine protection mechanism that uses a journal or log file to track changes made to the source virtual disk. The hypervisor quiesces all I/O from the guest operating system (OS) before the VMware snapshot is created. The snapshot captures the entire state of a virtual machine, including its configuration settings, virtual disk contents, and optionally, the contents of the virtual machine memory.

Virtual disk I/O is paused while a new snapshot virtual device is created. When I/O resumes, the virtual machine writes are applied to the snapshot virtual disk, or delta file, leaving the source disk unchanged. Because updates are not applied to the original virtual disk, the virtual machine can be restored to the pre-snapshot state by discarding the delta files. If the snapshot is deleted, the delta file and virtual disk files are merged to create a single-file image of the virtual disk.

EMC backup technologies leverage VMware snapshots to ensure the virtual machines are in a consistent state before an NFS SnapSure checkpoint or a LUN snapshot is created. The backup set consists of EMC snapshots of all datastores that contain virtual disks of the virtual machines being backed up. All files related to a particular virtual machine are backed up and restored together to establish the system state of the virtual machine when the snapshot is created. Organize virtual machines within datastores so they are backed up and restored together easily. Otherwise, restoring a LUN is not possible without impacting other virtual machines in the datastore.

If the backup set is intact, crash consistency is maintained even if the virtual machine has virtual disks provisioned across different storage types or protocols (VMFS, NFS, or RDM Virtual Mode).

To perform backup operations, complete the following steps:

Note: EMC Replication Manager is used to automate these steps and provide application integration and application consistency. [“Replication Manager” on page 187](#) provides more information about Replication Manager.

1. Initiate a VMware snapshot.
2. Set the flags to quiesce the file systems. Optionally capture the memory state.
3. Create a VNX NFS file system checkpoint or LUN snapshot of the datastore device that contains the virtual machine disks to be backed up.

Note: EMC Storage Viewer and Unisphere Virtualization views assist with the identification of the VNX storage devices backing each datastore. [“VSI: Storage Viewer” on page 22](#) provides more details.

4. Delete the VMware snapshot.

To restore virtual machines from a snapshot, complete the following steps:

1. Power off the virtual machine.
2. Initiate the NFS/LUN restores for all datastores containing virtual disks that belong to the virtual machine.
3. Update the virtual machine status within the vSphere UI by restarting the management agents on ESXi host console. Detailed information is available in *Restarting the Management agents on an ESXi or ESX host (1003490)*, available in the VMware Knowledge Base. Wait 30 seconds for the console to refresh.
4. Open the VMware Snapshot Manager and revert to the snapshot taken in the backup operation. Delete the snapshot.
5. Power on the virtual machine.

EMC Replication Manager supports creating VMFS and NFS datastore replicas in a vSphere environment, and provides point-and-click backup and recovery of virtual machine-level images and selective file restore in VNX OE for Block versions 5.31 and later.

VNX native backup and recovery options

VNX provides native utilities to create replicas of file systems and LUNs that support the ESXi environment. While these utilities are used for enterprise management of a vSphere environment, Replication Manager provides a more appropriate solution with application-level integration for enterprise-level backup and recovery of vSphere environments.

File system logical backup and restore using VNX SnapSure

Use VNX SnapSure to create near-line logical backups of individual NFS datastores mounted on ESXi hosts. Unisphere provides an interface to create one-time file system checkpoints and to define a checkpoint schedule to automate the creation of new file system checkpoints on VNX.

Note: SnapSure Checkpoint File Systems are stored in a hidden folder at the root of the source file system. A change in the Data Mover configuration is required to make the folder visible and perform selective copies from the vCenter Datastore Browser. To make the hidden directory visible, set the value of the Data Mover parameter `showChildFsRoot` to `1`, as shown in [Figure 91](#).

The screenshot shows the Unisphere interface for configuring Data Mover Parameters. The 'showChildFsRoot' parameter is selected, and its properties dialog is open. The dialog shows the following details:

Name	Facility	Value	Data Mover	Description
deleteDelay	cfs	1 (Default)	server_2	1= CIFS file deletes are immediate (default:on)
readwritesharing	cfs	0 (Default)	server_2	Valid read/write sharing on Dart
showChildFsRoot	cfs	0 (Default)	server_2	Enables visible checkpoint directories in root
showHiddenCkpt	cfs	1 (Default)	server_2	Enables/disables Celerra Virtual File System(CVFS) V2

The 'showChildFsRoot' parameter properties dialog is shown below:

```

Name: showChildFsRoot
Data Mover: server_2
Facility: cfs
Value: 1
Default Value: 0
Description: Enables visible checkpoint directories in root
Detailed Description: Enables/disables the Celerra Virtual File System (CVFS) version 1 NFS client access to checkpoints in the root directory of the production file system. param cfs showChildFsRoot=1 means that each mounted checkpoint of a production file system will be visible to NFS clients as subdirectories of the root directory of the production file system. param cfs showChildFsRoot=0 means that the checkpoint subdirectories will not appear in the root of the production file system.
  
```

Figure 91 ShowChildFsRoot parameter properties in Unisphere

Virtual machine files within a datastore are backed up and recovered as a single operation. To recover an individual virtual machine from an NFS checkpoint, complete the following steps:

1. Power off the virtual machine.
2. Browse to the Checkpoint File System to locate the folder that contains the virtual machine.
3. Use the **Datastore Browser** to select and copy the files from the Checkpoint File System to the existing datastore location on the ESXi host.
4. Power on the virtual machine.

Physical backup and restore using VNX File Replicator

Use VNX File Replicator to create a physical backup of NFS datastores. Replicator performs local or Remote Replication through the `/nas/bin/nas_replicate` command or through the Unisphere UI.

Replicator creates an independent file system for selective virtual machine recovery or complete file system restore through Unisphere.

Selective virtual machine recovery is performed through a host copy. After the file system copy is complete, stop the replication to transition the target file system to a stand-alone read/write copy. Mount the target file system to any ESXi host and copy the virtual machine files or folders through the datastore browser.

When using file system restore, ensure that all virtual machines within the file system are recovered to the same point in time. Virtual machines with different manage or service level requirements are placed in separate file systems.

Note: If VMware snapshots exist before the creation of a backup, vCenter Snapshot Manager may not report them correctly when a virtual machine is restored. If this happens, remove the virtual machine from the vCenter Inventory, import it again, and verify that the virtual machine is recognized correctly. Do not delete the virtual disks while removing the virtual machine from Inventory!

To recover an entire file system, establish a replication session from the target file system to the production file system with the `nas_replicate` command.

Snapshot backup and recovery of a VMFS datastore

EMC SnapView for VNX provides the functionality to protect VMFS datastores using either logical replicas (snapshots), or full volume copies (clones) of VNX LUNs. This storage system functionality is exposed through Unisphere, Unisphere Snapshot Configuration Wizard, or the **admsnap** utility.

In enterprise environments, LUN protection is controlled by Replication Manager for simplified configuration, automation, and monitoring of replication jobs. The utilities described in this section offer a manual approach to create or restore a replica of a VNX LUN.

When a snapshot is activated, SnapView tracks all the blocks of data for the LUN. As the LUN is modified, original data blocks are copied to a separate device in the reserve LUN pool.

Similarly, a clone private LUN pool is used to maintain various states between the source and target LUNs in a clone relationship. Ensure that the reserved LUN and the clone private LUN pools are configured before performing these operations.

SnapView operates at the LUN level, which means that VNX snapshot replicas are most effective when the datastore of interest is provisioned from a single LUN.

Note: To simplify snapshot management of VMFS datastore LUNs, create the datastore from a single LUN. Use metaLUNs or Pool LUNs for larger single LUN datastores.

If multiple virtual machines share the same VMFS datastore, they are backed up and recovered together as part of the snap or restore operation. While it is possible to perform manual restores of individual virtual machines from a snapshot LUN, it is best to group similar virtual machines within a datastore to avoid inadvertent impact from a restore operation.

To create a snapshot LUN using the Unisphere Snapshot Configuration Wizard, complete the following steps:

1. In Unisphere, launch the wizard and identify the production server where the source LUN exists.
2. Select the required VNX storage system and LUN for the SnapView session as shown in [Figure 92 on page 184](#).

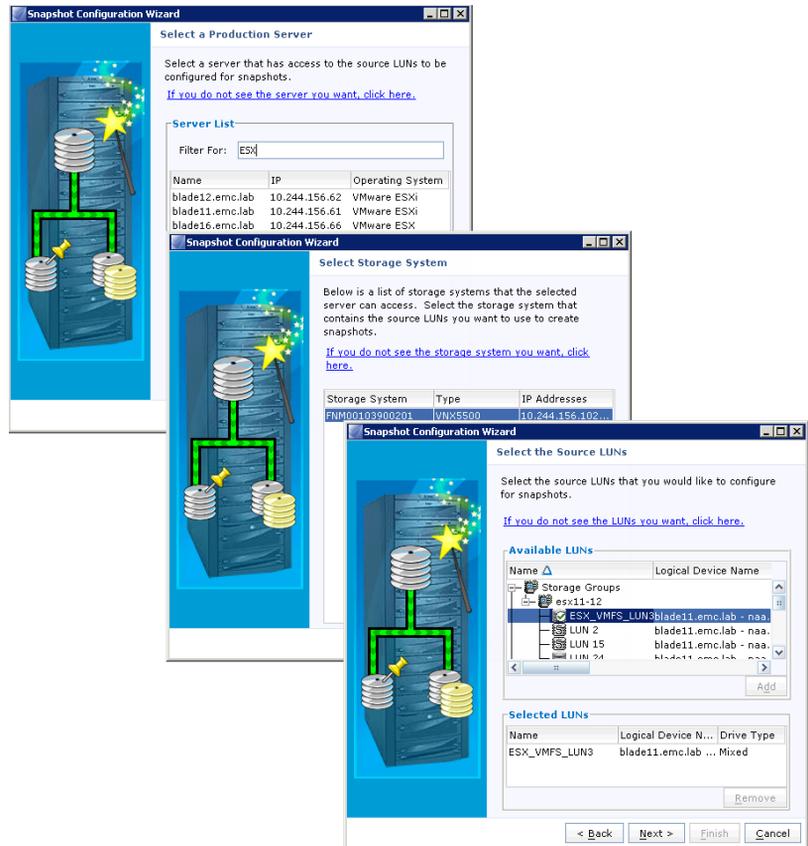


Figure 92 Snapshot Configuration Wizard

3. Select the appropriate number of copies for each source LUN, and optionally assign the snapshot to other ESXi hosts as shown in [Figure 93](#) on page 185.

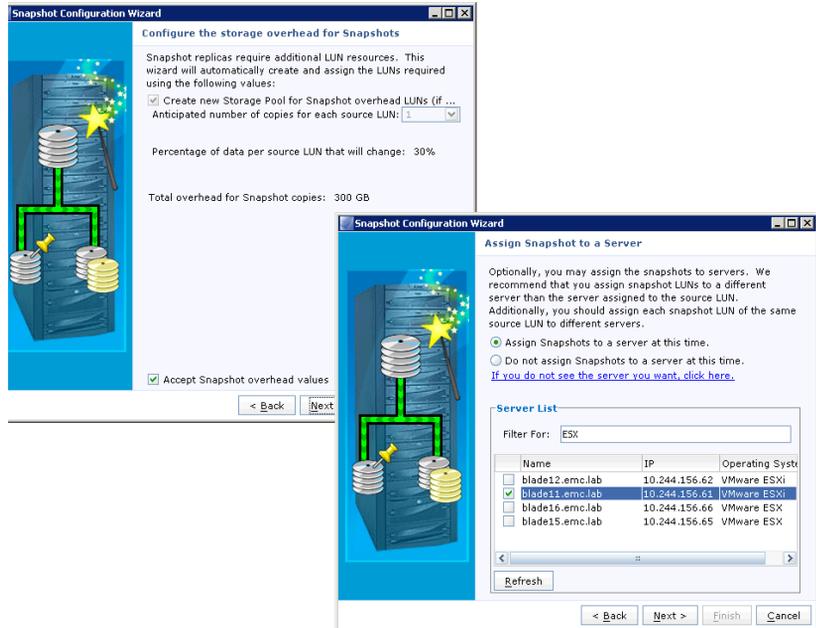


Figure 93 Snapshot Configuration Wizard (continued)

4. Type the snapshot name.
5. Select a host to add the snapshot image to the host storage group.
6. Review the configuration information and click **OK** to create and mount the snapshots.
7. Use Unisphere to start the snapshot session and activate the snapshot for access by another host.
8. Rescan the ESXi hosts and verify that the storage appears in the correct location.

If required, select **Assign a new signature** to automatically resignature the device. “[ESXi volume signatures](#)” on page 167 provides more information on device signatures.

When the snapped VMFS LUN is accessible from the ESXi host, virtual machine files are copied from the snapped datastore to the original VMFS datastore to recover the virtual machine.

Backup and recovery of RDM volumes

VNX LUNs are formatted as VMFS file systems or RDM volumes. An RDM volume is a raw device mapped directly to the virtual machine. RDMs provide capabilities similar to a VMFS virtual disk, while retaining the properties of a physical device. With RDM volumes, administrators take full advantage of storage array-based data protection technologies. EMC SnapView provides logical protection of RDM devices to create snapshot images.

To back up an RDM volume, administrators use a variety of EMC replication technologies to create usable copies of the device.

For RDM volumes, administrators create snapshots or clones in one of the following ways:

- ◆ Use the **admsnap** command or the Unisphere Snapshot Configuration Wizard.
- ◆ Use Replication Manager to integrate with Windows applications or create stand-alone snapshots or clones of the RDM volumes.

Note: Replication Manager only supports RDM volumes created in physical compatibility mode and formatted as NTFS volumes.

Replication Manager

EMC Replication Manager is a software solution that integrates with EMC data protection technologies to simplify and automate replication tasks. Replication Manager uses EMC SnapSure or EMC SnapView to create local or remote replicas of VNX datastores.

Replication Manager works with vCenter to create VMware snapshots of all online virtual machines before creating local replicas. This virtual machine snapshot creation provides a higher level of consistency than simply snapping the datastore. The VMware snap attempts to quiesce all I/O to the virtual machine before the snap is created. Replication Manager uses a physical or virtual machine to act as a proxy host to process all VMware and VNX management tasks. The proxy host is configured to communicate with the vCenter Server and the VNX storage systems. It discovers storage devices in the virtualization and storage environments, and performs the necessary management tasks to establish consistent copies of the datastores and virtual machine disks. Use the Replication Manager Job Wizard, as shown in [Figure 94 on page 188](#) to select the replica type and expiration options. Replication Manager 5.2.2 is required for datastore support.

Complete the following steps before restoring the replicas:

1. Power off the virtual machines that reside within the datastore.
2. Remove those virtual machines from the vCenter Server inventory.

Figure 94 Replication Manager Job Wizard

Select the **Restore** option in Replication Manager to restore the entire datastore:

1. Restore the replica.
2. Import the virtual machines to the vCenter Server inventory after the restore is complete.
3. Revert to the VMware snapshot taken by Replication Manager to obtain an OS-consistent replica, and delete the snapshot.
4. Configure Replication Manager to power on each virtual machine.

Replication Manager creates a rollback snapshot for every VNX file system it restores. The name of each rollback snapshot is available in the restore details as shown in [Figure 95](#). Verify the contents of the restore, and then delete the rollback snapshot.

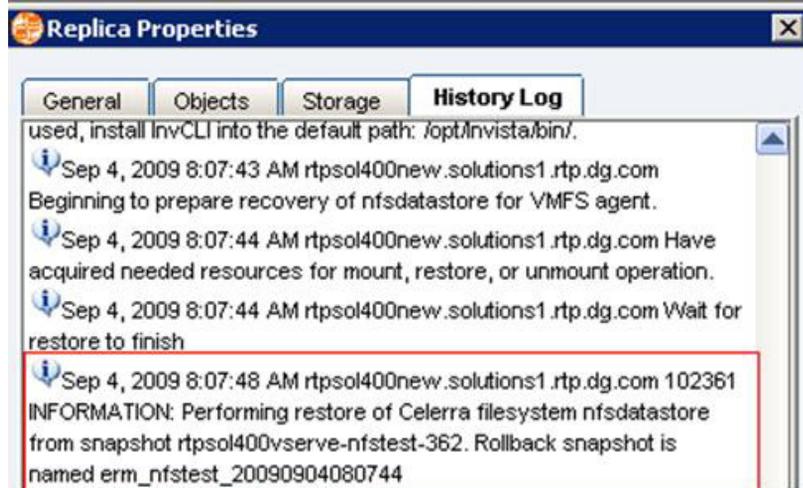


Figure 95 Replica Properties in Replication Manager

Replication Manager version 5.3 and later provides the ability to selectively restore a virtual machine, as shown in [Figure 96](#).

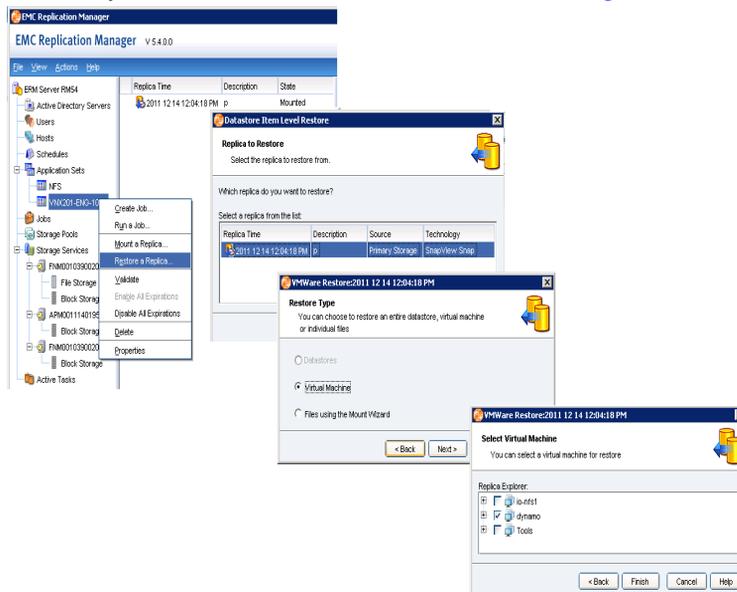
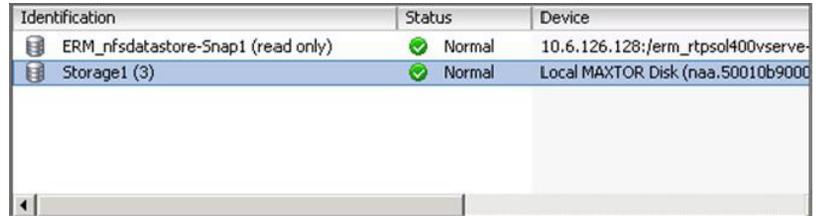


Figure 96 Replication Manager virtual machine restore

To selectively restore a virtual machine, complete the following steps:

1. Select the application set that contains the replica you want to restore.
2. Identify the date and time the replica was created.
3. Right-click to view the management options.
4. Select **Restore a Replica** and then click **Next**.
5. Select the virtual machine or virtual machines to restore and then click **Next**.

6. Monitor the progress through the Replication Manager status window.
7. Revert to the VMware snapshot taken by Replication Manager to obtain an OS-consistent replica, and delete the snapshot.
8. Unmount the replica through Replication Manager.
9. Power on the virtual machine.



Identification	Status	Device
ERM_nfsdatastore-Snap1 (read only)	Normal	10.6.126.128:/erm_rtps0400vserve-
Storage1 (3)	Normal	Local MAXTOR Disk (naa.50010b9000)

Figure 97 Read-only copy of the datastore view in the vSphere client

Backup and recovery of a VMFS with VNX Advanced Snaps

EMC VNX OE for Block release 5.32 introduced a new snapshot architecture for pool LUNs. This new snapshot is used to create up to 256 snapshots of the source LUN, including the ability to create snapshots of other snapshots for that LUN.

Snapshots are created from individual LUNs, or groups of LUNs, defined within a consistency group. Snapshots can be created using an existing snapshot as the source of the new snapshot. A snapshot request creates a crash-consistent version of the selected source LUNs.

A new object called a mount point provides the management object used to present the snap image to a storage group (that is, the host). The mount point appears as a pseudo device within the ESXi host. The device cannot be managed or accessed until an advanced snapshot image is attached to it. Snapshot versions are attached and detached from the mount point to change the content within the device. Advanced snapshots are read/write enabled, which means their content can be modified while a LUN is attached to a mount point.

The Unisphere UI provides the supported interface to manage advanced snapshots. Additionally a command line utility is available for in-band management when the snapshot mount point is enabled. [Figure 98 on page 193](#) shows the check box to select for in-band management of the snapshots assigned to the host. To create a snapshot of a Pool LUN using Unisphere, complete the following steps:

1. Select the LUN to create a snapshot of (use VSI or Unisphere Virtualization view to assist with the identification of the datastore LUN).
2. Right-click the LUN, and then select **Create Snapshot**.

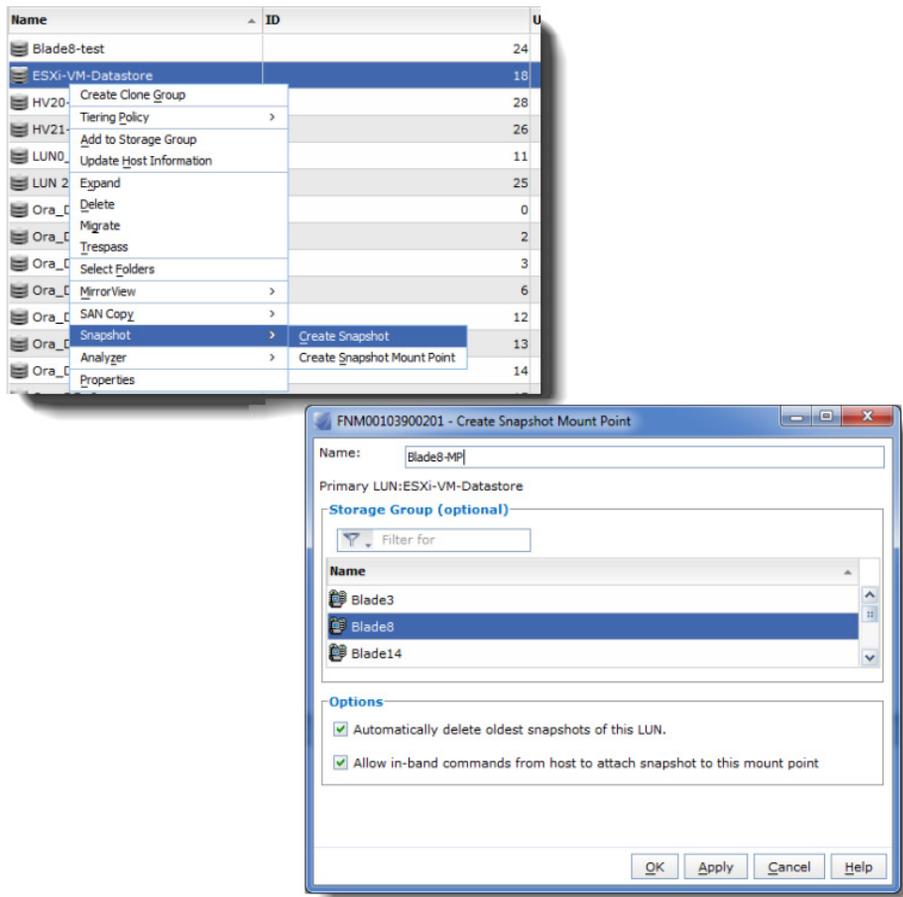


Figure 98 Advanced Snapshot Basic Configuration

3. If a snapshot mount point does not exist, create one and assign it to a storage group for the ESXi host to access the snapshot image.

In the example in [Figure 99](#), a snapshot mount point named Blade8-MP is created and assigned to Blade8. After it is created, snapshots are attached and detached from the mount point through Unisphere.

Note the checkbox option to manage the snapshots from the CLI. If you select the storage group for Blade8 there is a mount point associated with the storage group as illustrated in [Figure 99](#).

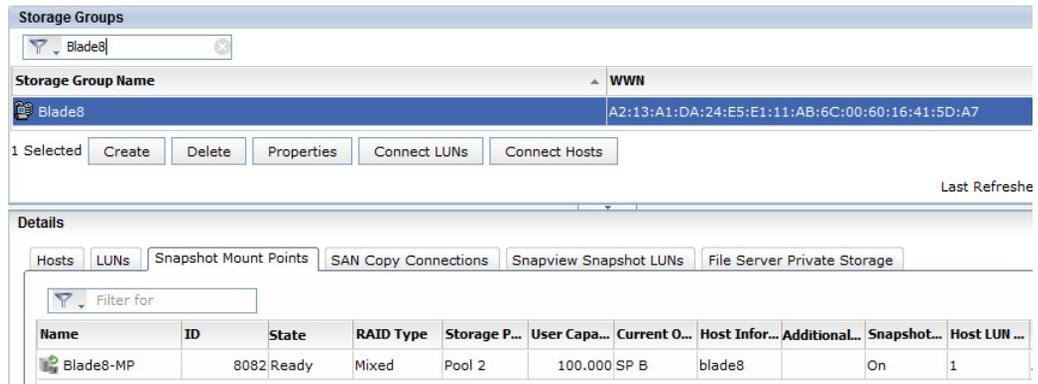


Figure 99 Snapshot Mount Point

- Specify the snapshot name when the snapshot and mount point are created.

Consistency groups

For consistency groups with multiLUN configurations, complete the following steps:

- Select the **Data Protection** tab in Unisphere and select the **Snapshots** option.
- The host requires a snapshot mount point for each LUN in the consistency group. Select the **Create Snapshot Mount Points** wizard as shown in [Figure 100 on page 195](#).
 - Select the system to mount the snapshots.
 - Select the storage system containing the LUNs to be part of the consistency group.
 - Select all of the LUNs to be part of the consistency group.
 - Assign the mount points to the host. After the mount point is created, the host considers the mount point as a logical device. Attempting to mount the device without attaching a snapshot does not yield useful results.

- Click **OK** to finish. You now have the necessary mount points to attach the snapshots from your application LUNs.

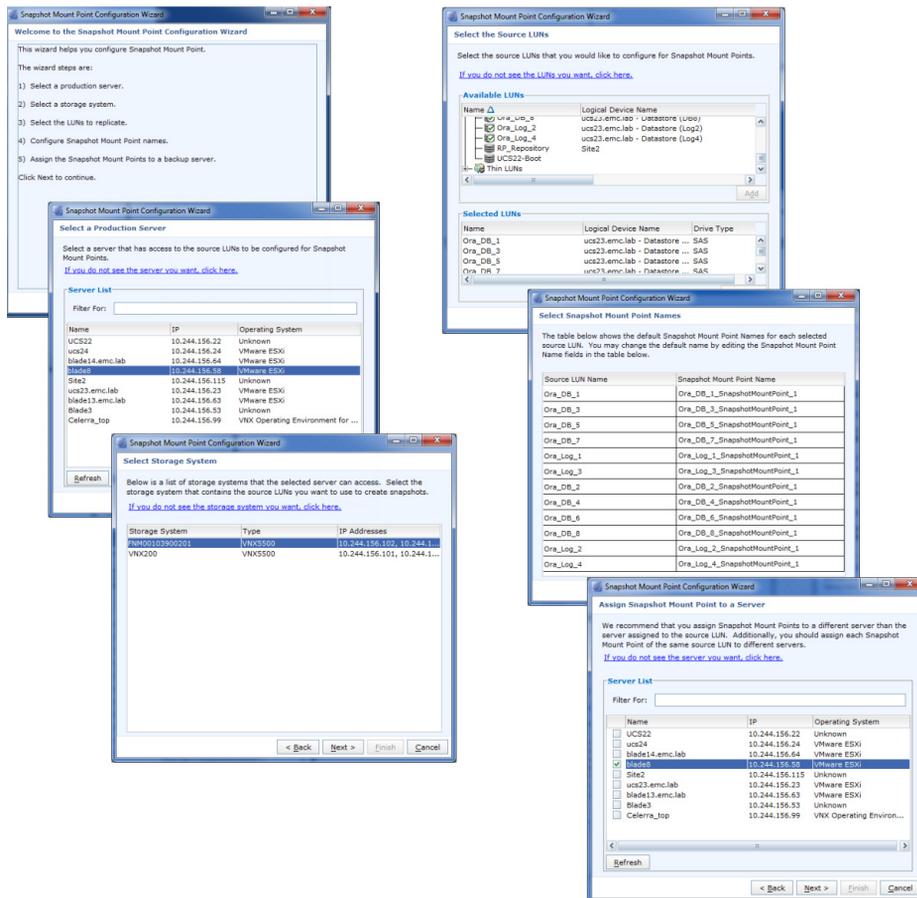


Figure 100 Mount Point configuration wizard

- From the snapshot interface of the **Data Protection** tab, click **Create Group** to create the consistency group.
- Type the group name and the description. The description is optional. (This example is protecting multiple Oracle Database LUNs).

6. Select the LUNs that are part of this consistency group. As soon as a snapshot job is performed, a snapshot for each LUN is created. When one snapshot is attached to a mount point, all LUNs are attached to the mount point.
7. Click **Finish**. Figure 101 shows the complete creation of the consistency group.

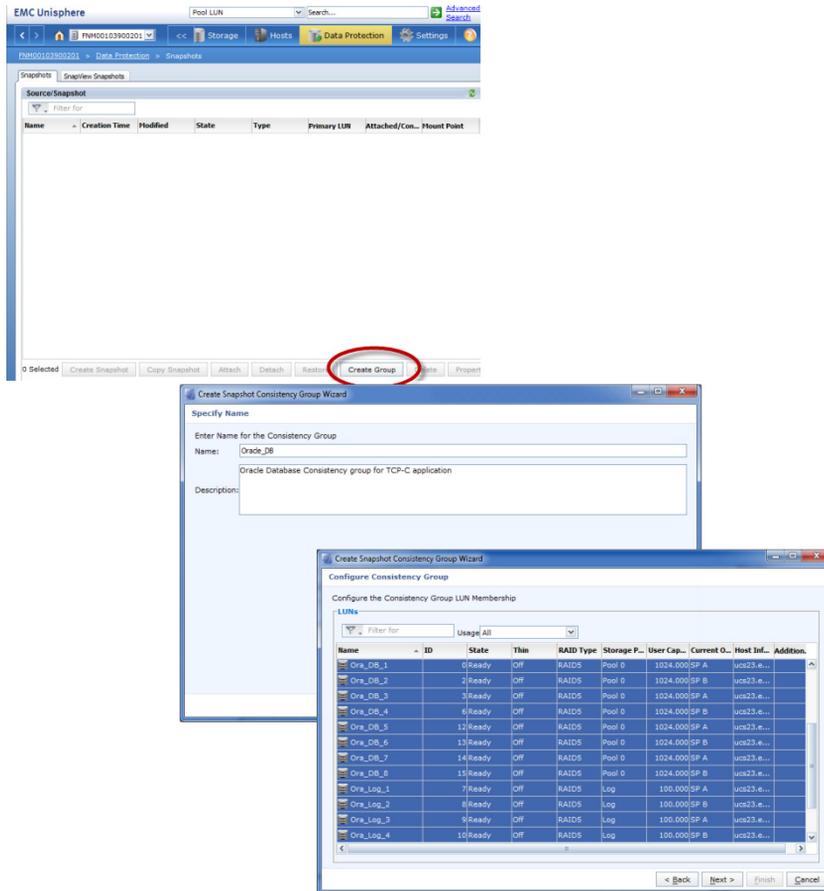


Figure 101 Snapshot consistency group creation

8. Select the consistency group to create a snapshot of all LUNs in the consistency group. Select a host to add the snapshot image to the host storage group. Figure 102 shows the consistency group snapshot creation.

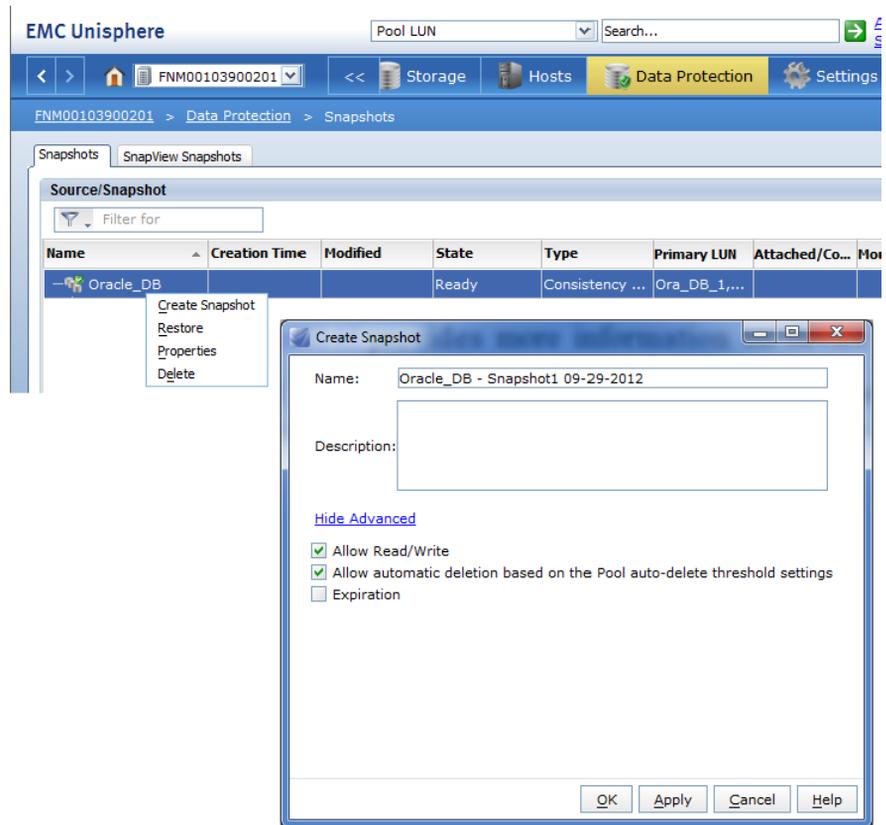


Figure 102 Consistency group snapshot creation

9. Figure 103 shows how to attach the snapshots to the mount points to present them to the host.

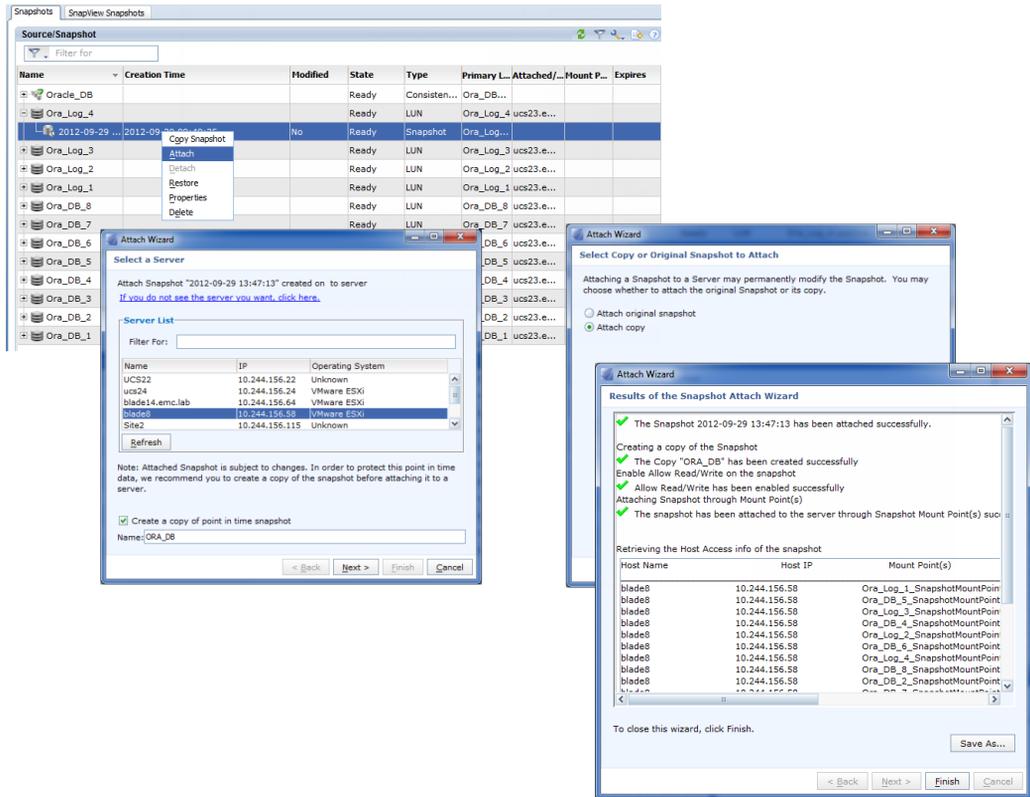


Figure 103 Consistency group snapshot attach

10. Select one of the snapshots created within the consistency group. Do one of the following to attach a snapshot:
 - Right-click the LUN to display management options and select **Attach**.
 - Click **Attach** in the snapshot management window.
11. Select the host to attach the snapshots to.

12. Select from the following options in the wizard:
 - Attach the existing snapshot.
 - Create an additional snapshot copy.
 - Preserve the existing snapshot.
13. Select **Create a new snapshot** to make changes to the snapshot and preserve the existing state, or attach the copy.
14. Identify the host or cluster after logging in to vCenter. Rescan the host adapter(s) to force the host to recognize the new SCSI devices.

If required, select **Assign a new signature** to automatically resignature the device. [“ESXi volume signatures” on page 167](#) provides more information on device signatures.

When the snapped VMFS LUN is accessible from the ESXi host, virtual machine files can be copied from the snapped datastore to the original VMFS datastore to recover the virtual machine.

vStorage APIs for Data Protection

VMware vStorage APIs for Data Protection (VADP) provides an interface into the vCenter environment to create and manage virtual machine snapshots. VADP is leveraged by data protection vendors to automate and streamline non-disruptive, fully recoverable, incremental virtual machine backups. A key feature of VADP is Changed Block Tracking (CBT), which allows a data protection application to identify modified content on the virtual machine based upon a previous VMware snapshot. This reduces the amount of data that needs to be backed up and restored while using differential backups of virtual machines.

The benefits are a reduction in the amount of time required to back up an environment, and storage savings achieved by backing up only the required data blocks instead of the full virtual machine.

VADP integrates with existing backup tools and technologies to perform full and incremental file backups of virtual machines.

Figure 104 shows how VADP works.

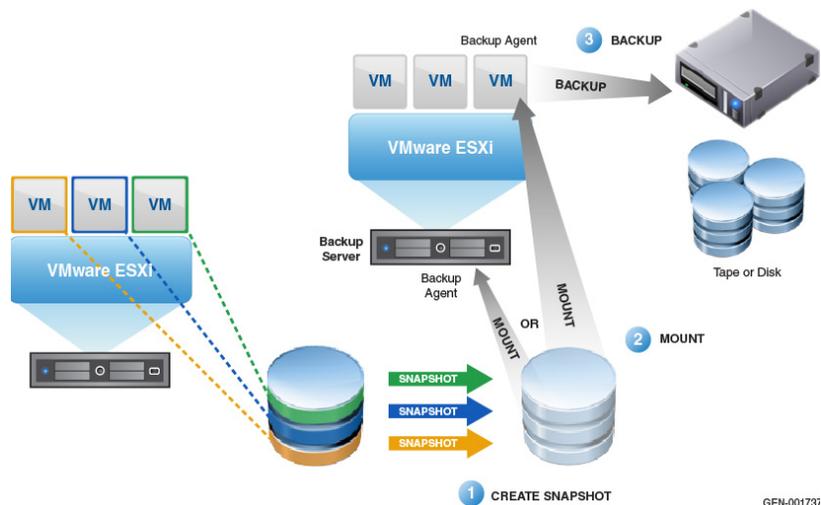


Figure 104 VADP flow diagram

Backup and recovery using VMware Data Recovery

Note: In vSphere 5.1, this feature is known as VMware Data Protection.

VMware Data Recovery (VDR) is a disk-based backup and recovery solution built on the VADP. It uses a virtual appliance and a client plug-in to manage and restore virtual machine backups. VMware Data Recovery can protect any kind of OS. It incorporates capabilities such as block-based data deduplication to perform incremental backups after an initial full backup to maximize storage efficiency. VNX CIFS, iSCSI, and FC storage are used as destination storage for VDR backups. Each virtual machine backup is stored on a target disk in a deduplicated store.

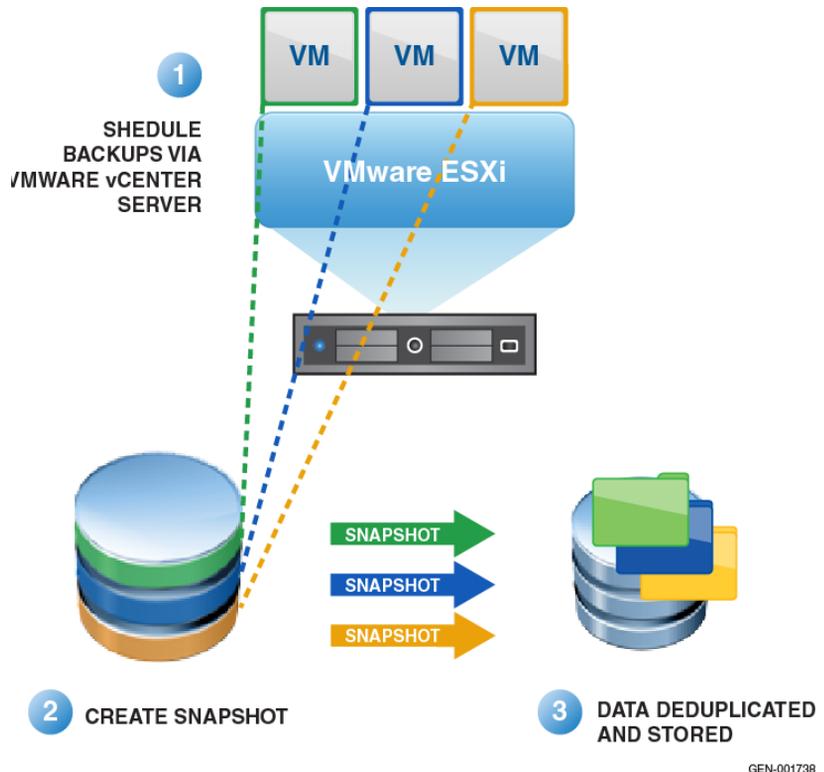


Figure 105 VMware Data Recovery

During the backup, VDR takes a snapshot of the virtual machine and mounts it directly to the VDR virtual machine. The VDR streams blocks of data to the destination storage as shown in [Figure 105 on page 201](#). During this process, VDR uses the VADP CBT functionality on ESXi hosts to identify the changed blocks and minimize the amount of data to be backed up. VDR deduplicates the stream of data blocks to further eliminate redundant data prior to writing the backup to the destination disk. The deduplicated store creates a virtual full backup based on the last backup image and applies the changes to it. When all the data is written, VMware Data Recovery dismounts the snapshot and takes the virtual disk out of snapshot mode. VMware Data Recovery supports only full and incremental backups at the virtual machine level, and does not support backups at the file level.

Adhere to the following guidelines for VMware Data Recovery:

- ◆ A VMware Data Recovery appliance protects up to 100 virtual machines, but it is limited to two simultaneous backup destinations. Schedule the backups serially to overcome this limitation. Stagger VDR backup jobs and ensure the backup destination size does not exceed 1 TB.
- ◆ A VMware Data Recovery appliance cannot use a NFS file system as a backup destination. However, a virtual disk created from a NFS datastore and mounted to the VDR system is a valid backup target. To use NFS, create virtual machine disks within an NFS datastore and assign them to the VDR appliance.
- ◆ VMware Data Recovery supports RDM virtual and physical compatibility modes as backup destinations. Use the virtual compatibility mode for RDM as a backup destination. SAS or NL-SAS devices provide a useful RDM target device for VDR backups.
- ◆ Back up similar virtual machines to the same destination. As VMware Data Recovery performs data deduplication within and across virtual machines, only one copy of the OS is stored if multiple virtual machines use the same OS.
- ◆ The virtual machine must not have a snapshot named **_data recovery_** prior to a backup performed by VMware Data Recovery. VDR creates a snapshot named **_data recovery_** as a part of its backup procedure. If a snapshot with the same name already exists, VDR will delete and re-create it.

- ◆ Backups of virtual machines with RDM can be performed only when the RDM is running in virtual compatibility mode.
- ◆ VMware Data Recovery provides an experimental capability for Windows systems called File Level Restore (FLR). FLR gives users the ability to restore individual files without the need to restore the whole virtual machine.
- ◆ VMware Data Recovery only copies the state of the virtual machine at the time of backup. Pre-existing snaps are not a part of the VMware Data Recovery backup process.

Backup and recovery using Avamar

EMC Avamar® is a backup and recovery software product. Avamar provides an integrated software solution to accelerate backups and restores of virtual machine and application data in a vSphere environment. Avamar provides source and global data deduplication to reduce the amount of backup data that must be copied across the network and stored on disk. Global deduplication means that Avamar stores a single copy of each unique subfile, variable-length data segment for all protected physical and virtual servers in the environment.

After an initial virtual machine backup, Avamar creates full restore backups of virtual machines that require only a fraction of the space and time used to create the original. Avamar integration with vCenter and VMware vStorage APIs allows it to leverage the CBT feature of vSphere to identify data blocks of interest for the backup job. Avamar applies deduplication based on the global view of the stored data, and only copies globally unique blocks to the Avamar Storage Node or Avamar Virtual Edition (AVE) server. This greatly reduces backup times and storage consumption in the backup environment.

Avamar reduces backup times, backup capacity requirements, and ESXi host resource utilization.

Architectural view of the Avamar environment

Avamar Server is a core component that provides management and storage for the virtual machine backup environment. The server provides the management, services, and file system storage to support all backup and administrative actions. Avamar has the following server types:

- ◆ **Avamar Data Grid** — An all-in-one server that runs Avamar software on a preconfigured, EMC-certified hardware platform. The options include single and multinode versions that use either internal or SAN storage.
- ◆ **Avamar Virtual Edition for VMware (AVE)** — A fully functional Avamar Server that installs and runs as a virtual appliance within a vSphere environment.

Both physical and virtual edition products provide the same capabilities. However, AVE is easy to deploy in a vSphere environment. It is backed by VNX block storage for high

performance, Tier 1 protection of virtual machine, application, and user data. AVE also performs significantly better in VMware environments than the Avamar Datastore. [Figure 106](#) shows a sample configuration with a DRS cluster and multiple ESXi hosts with access to VNX block LUNs. These LUNs contain the virtual machines in the environment. The environment illustrates three types of virtual machines: production virtual machines, image proxies, and file-level proxies.

The Production virtual machines can run any VMware-supported OS, and serve any application role or function. In this scenario, the virtual machines do not require an Avamar agent.

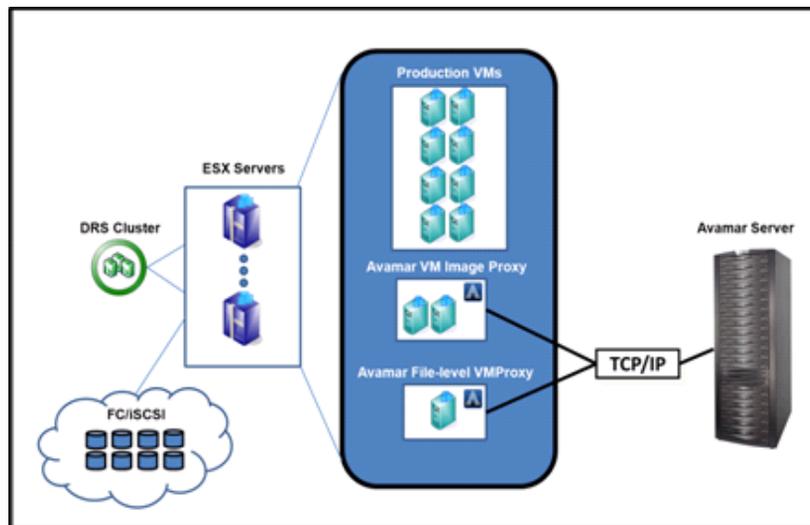


Figure 106 Sample Avamar environment

Avamar backups

Avamar provides the following backup options for vSphere environments:

- ◆ **File Level Backup** — File level backups are enabled by installing the Avamar client inside the guest OS and registering the client with an Avamar Server. This option provides a scheduled backup of all files on the virtual machine, and allows the user to manually backup and restore files to their desktop virtual machine. The client capabilities are the same as when the client is installed in a physical computer environment.

With the Avamar client, backups complete with minimal administrative resource requirements. Scheduled backups occur based on administrative policy. Users also have the ability to manually initiate backups and restores at any time.

The Avamar client runs as a low priority virtual machine process to limit the impact of the backup operation on other processes. From a vSphere standpoint, Avamar can throttle virtual machine CPUs to limit the amount of ESXi host CPU resources consumed during backup operations.

- ◆ **Image Level Backups** — Image Level backups allow the vSphere environment to be backed up without installing a client on each virtual machine. They use one or more Avamar virtual machine Image Proxy servers that have access to the shared VNX storage environment.

The Image Proxy is provided as a downloadable .ova image. It is accessible through the web interface of the AVE server. The Image Proxy server installs as a virtual machine appliance within vCenter. Separate Image Proxy servers are required for Windows and Linux virtual machine image backups.

After installation, the proxy server is configured to protect either Windows or Linux virtual machines. Avamar integrates with vCenter, and provides a similar management interface to import and configure virtual machine protection. [Figure 107](#) shows a sample proxy configuration.

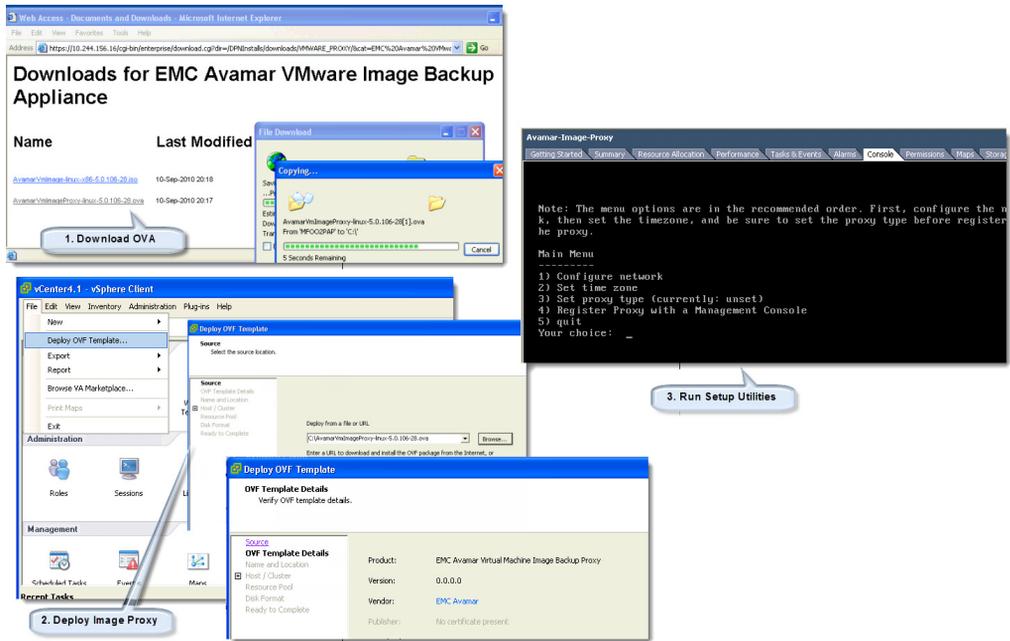


Figure 107 Sample proxy configuration

Avamar Manager can also enable CBT for virtual machines to further accelerate backup processing. With CBT enabled, Avamar easily identifies and deduplicates the blocks that VMware has flagged without the need to perform additional processing. This allows for faster, more efficient backups of the virtual machine image.

Figure 108 provides more details.

Note: CBT is available with virtual machine version 7 and later. Update older virtual machines to version 7 to backup the virtual machine with CBT enabled.

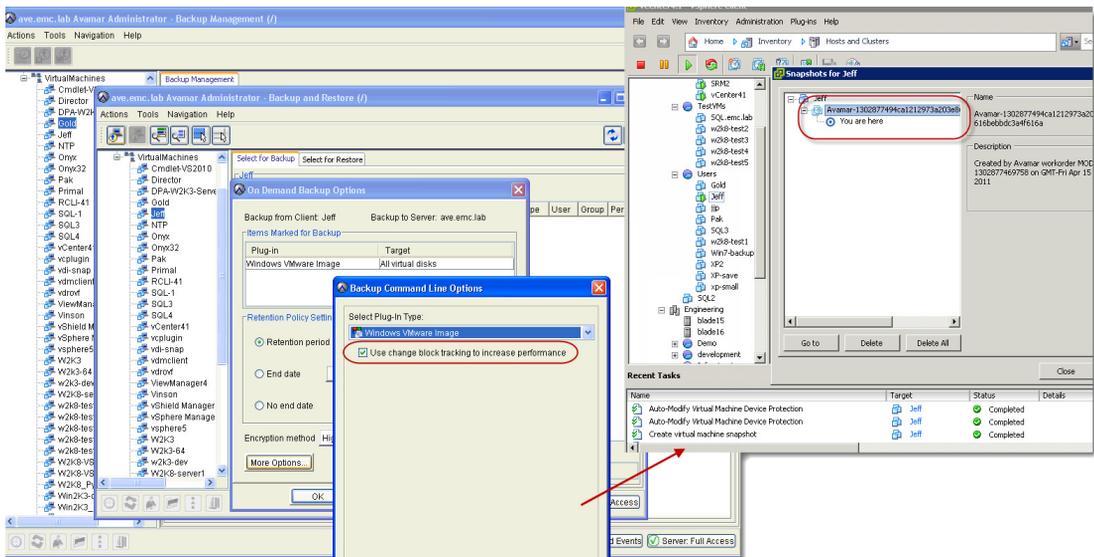


Figure 108 Avamar backup management configuration options

When a backup job starts, Avamar signals the vCenter server to create a new Snapshot image of each VMDK specified in the backup policy. It uses VADP SCSI hot-add to mount the snap to the image proxy. If CBT is enabled, Avamar uses it to filter the data that is targeted for backup. After Avamar establishes a list of blocks, it applies deduplication algorithms to determine if the segments are unique. If they are, it copies them to the AVE server. Otherwise, it creates a new pointer that references the existing segment on disk. The image proxy then copies those blocks to the VNX-backed virtual disks on the Avamar Virtual Appliance.

Unique proxies are required to protect Windows and Linux environments. The administrator can deploy additional proxies to provide scalability, and allow simultaneous backups and recoveries. Avamar provides the ability to configure each image proxy to protect multiple datastores from vCenter, or to load balance backups across all of them in a round-robin fashion, to improve scalability.

Avamar data recovery

Avamar also provides multiple recovery options. The two most common recovery requests made to backup administrators are:

- ◆ **File-level recovery** — Object-level recoveries account for the majority of user support requests. File-level recovery is appropriate for:
 - Deleted files
 - Application recovery
 - Batch process-related erasures

The Avamar client allows users to perform self-service file recovery by browsing the file system and identifying the files they need to restore.

- ◆ **System recovery** — Complete system recovery requests are less frequent than those for file-level recovery, but this bare metal restore capability is vital to the enterprise. Some common root causes for full-system recovery requests include:
 - Viral infestation
 - Registry corruption
 - Unidentifiable, unrecoverable issues

Virtual machine image restore

The image proxy can restore an entire image to the original virtual machine, a new virtual machine, or a pre-existing alternate virtual machine with a configuration similar to the original. Avamar Image Proxy can restore a virtual machine image to the same location where it was created, a different existing virtual machine, or as a new virtual machine to a different location in the environment. Figure 109 shows a virtual machine being restored to its original location. In this example, the virtual machine was deleted from the disk, and restored to the existing datastore.

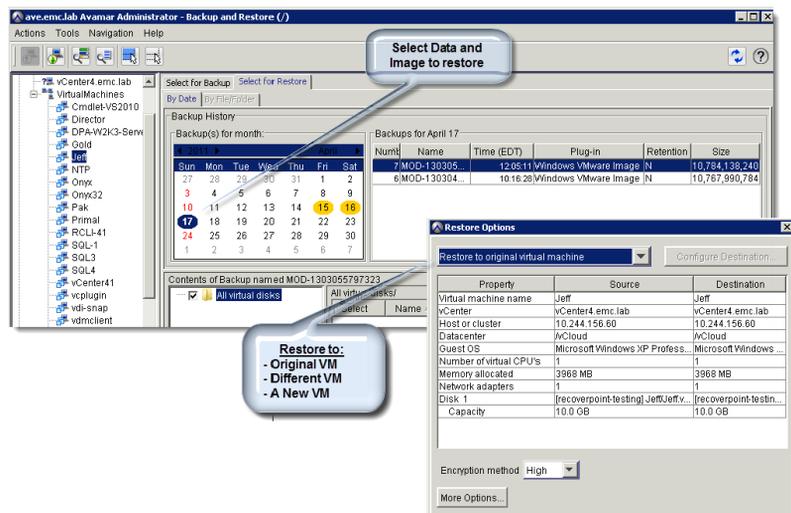


Figure 109 Avamar virtual machine image restore

An Avamar file-level recovery proxy is a virtual machine that allows one or more files to be recovered to a virtual machine from a full image backup. This virtual machine leverages the Avamar Virtual File System (AvFS) to present a view of the virtual machine disk for users to browse. From this view the administrator selects any file or folder to restore to the original location, or to a new location within the same virtual machine. The Avamar file-level proxy feature is available only for Windows virtual machines at this time.

The file-level restore feature uses a Windows proxy client virtual machine. The Avamar and VMware software on the Windows proxy requires a CIFS share, which is exported by the Avamar server.

This CIFS share provides a remote, hierarchical, file system view of the backups stored on the Avamar server. Access the CIFS share to browse and restore the contents of the VMware Image Backups.

When backups are selected for recovery, the FLR proxy server reads the VMDK data from the Avamar system and creates a browse tree that is presented to the administration GUI as shown in [Figure 110](#).

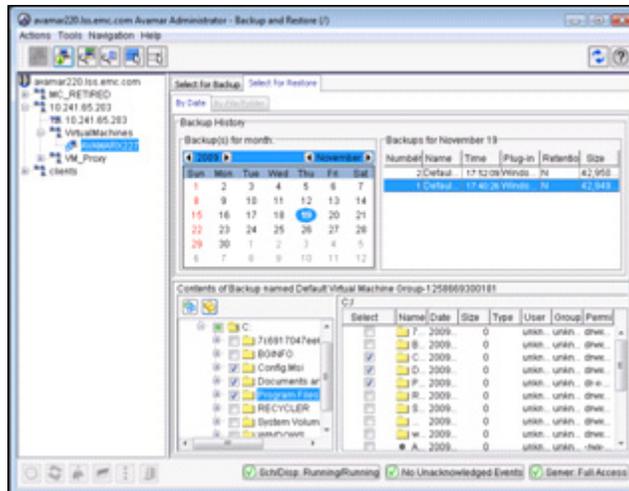


Figure 110 Avamar browse tree

Restore requests pass from the Avamar system, through the Windows FLR proxy, and on to the protected machine. The recovery speed of this operation is governed by the resources of the FLR proxy to read in the data and write it to the virtual machine being recovered. Therefore, large data recoveries through the FLR proxy recovery are not advisable. In this instance, an image-level, out-of-place recovery is more efficient.

Note: FLR requires that target virtual machines be powered on and run virtual machine tools.

Consider the following items while setting up the environment:

- ◆ Avoid using FLR to browse folders or directories with thousands of files or subdirectories. A better alternative is to restore the virtual machine and use the native OS to browse and identify the files you want to restore.
- ◆ Backup of Avamar proxy clients is not required. The proxy client virtual machines are easy to redeploy from the template if necessary.
- ◆ Avamar image backup is dependent on reliable DNS service and time synchronization. Network routing and firewall settings must be correctly configured to allow access to the network hosts that provide these services.
- ◆ SSL certificate must be installed across the vCenter, ESXi hosts, and Avamar proxy virtual machine appliances. However, it is possible to turn off SSL certificate authentication at the Avamar server.
- ◆ Use multiple network interfaces for HA configurations of the Avamar Datastore Node.
- ◆ Backups are a crash-consistent snapshot of the full virtual machine image. Use the Avamar client for OS and application-consistent backups.
- ◆ An image proxy performs one backup at a time. Parallel processing is possible only with multiple proxies in an environment.
- ◆ Virtual machine snapshots are required as part of the image backup process.
- ◆ Image backup supports the following disk types:
 - Flat (version 1 and 2).
 - Raw Device Mapped (RDM) in virtual mode only (version 1 and 2).
 - Sparse (version 1 and 2)

Backup and recovery using NetWorker

EMC NetWorker performs agentless, full image-level backup for virtual machines running any OS and file-level backups for virtual machines running Microsoft Windows. NetWorker consists of the following components:

- ◆ **Agent** — NetWorker Agent architectures are particularly focused on environments that require application consistency. For virtual machine backups that require application integration, the agent is used to place the application and OS into a consistent state before generating a virtual machine snapshot and performing the backup task. The agent configuration requires additional client administration on all of the virtual machines. If crash-consistent or operating system-consistent images are sufficient, VADP may be a better option.
- ◆ **VADP** — NetWorker 7.6 SP2 introduces the integration with VMware environments to support virtual machine protection with VADP. In a NetWorker environment, VADP creates a snapshot copy of a running virtual machine disk. NetWorker offers the ability to architect flexible backup solutions to improve backup processes, reduce backup windows, and reduce the amount of space required to store backup images.

Figure 111 shows the virtualization topology in an environment with NetWorker.

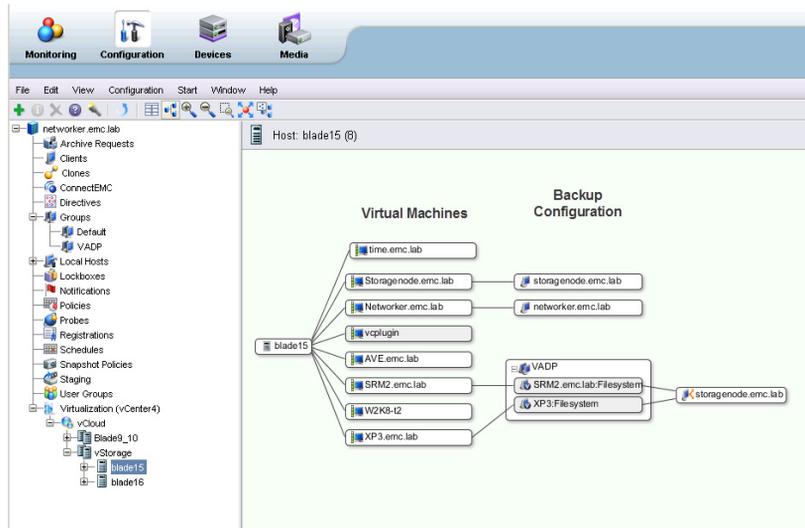


Figure 111 NetWorker-virtualization topology view

NetWorker backups use the VADP API to generate virtual machine snapshots on the vCenter server. The snapshots are hot-added to a VADP proxy host for LAN-free backups. A NetWorker initiated snapshot is identified as `_VADP-BACKUP_` as shown in Figure 112.

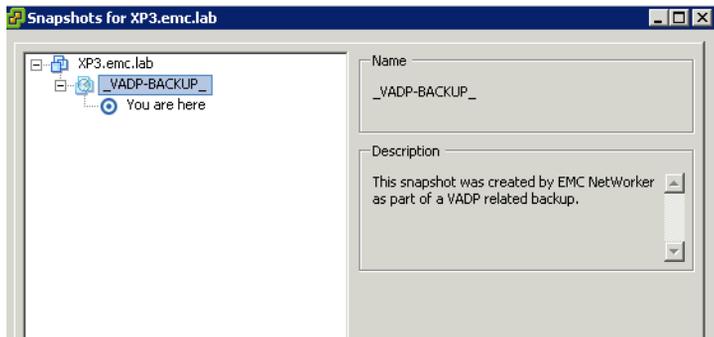


Figure 112 VADP snapshot

VNX storage devices for NetWorker

NetWorker offers the flexibility to use multiple storage types as targets for backup jobs. Supported storage types include standard physical tape devices, virtual tape libraries, and Advanced File Type Devices (AFTD) provisioned on VNX storage. An AFTD can be configured on the NetWorker server or Storage Node using a block LUN, or a NAS file system. NL-SAS LUNs or VNX FAST Pool LUNs that consist of NL-SAS drives are ideal for AFTDs.

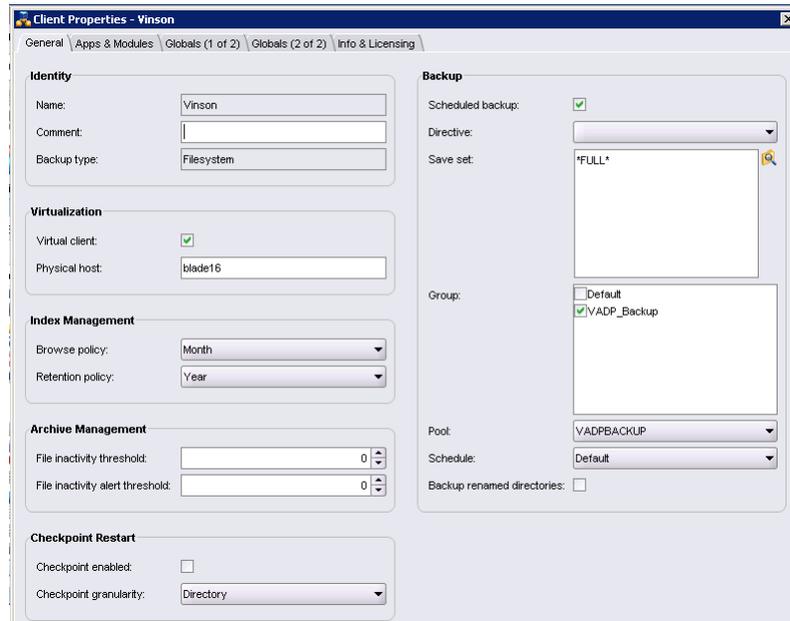


Figure 113 NetWorker configuration settings for VADP

Consider the following guidelines and best practices for VADP with vSphere:

- ◆ The latest version of VMware tools must be installed on all virtual machines. Without VMware tools, the backup created by VADP will be crash-consistent.
- ◆ File-level backup is available only for Windows virtual machines. VADP supports image-level backups for all other OSs.
- ◆ VADP does not support RDM physical compatibility mode.

- ◆ RDMS in virtual compatibility mode are converted to a standard virtual disk format during backup. They are converted to VMFS virtual disks when restored.
- ◆ LAN mode does not allow virtual disks to exceed 1 TB each.
- ◆ SAN is the default backup mode. To perform LAN-based backup, change the **TRANSPORT_MODE** to **nbd**, **nbdssl**, or **hotadd** in the **config.js** file.
- ◆ The hot-add transport mode does not support the backup of virtual disks that belong to different datastores.
- ◆ VADP creates a virtual machine snapshot named **_VADP-BACKUP_** before a file-level backup. A NetWorker backup fails if a snapshot with the same name already exists. Change the **PREEXISTING_VADP_SNAPSHOT** parameter in the **config.js** file to **delete** or to modify the default behavior.
- ◆ Even if a backup job fails, virtual machines remain mounted in the snapshot mode. NetWorker Monitoring Window provides an alert if a snapshot must be manually removed.
- ◆ VADP searches for the target virtual machines by IP address. The virtual machine must be powered on the first time it is backed up, so the virtual disk information is relayed to NetWorker through the vCenter server. This information is cached on the VADP proxy and used for subsequent backup jobs. Change the **VM_LOOKUP_METHOD=name** parameter in the **config.js** file to change this behavior.

Note: The backup will fail if duplicate virtual machine names exist.

- ◆ Beginning with the NetWorker release 7.4.1, users must add each virtual machine to be backed up as a NetWorker client. The NetWorker client software is not required on the virtual machine. With NetWorker release 7.4.1 or later, the VADP method to find virtual machines is based on the virtual machine IP address (default method).

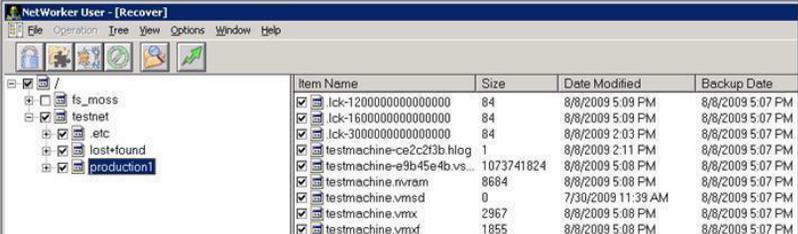
VNX NAS file system NDMP backup and restore using NetWorker

NetWorker provides two methods of storage integration with VNX NFS datastores. VNX provides file systems to use as Advanced File System Type Devices (AFTD), or as a Virtual Tape Library Unit (VTLU).

Configure a VTLU on the VNX file system, then configure NetWorker as an NDMP target to back up NFS datastores on the VNX platform. Configure NetWorker to use VNX File System Integrated Checkpoints to create NDMP backups in the following manner:

1. Create a Virtual Tape Library Unit (VTLU) on VNX NAS storage.
2. Create a library in EMC NetWorker.
3. Configure NetWorker to create a bootstrap configuration, backup group, and a backup client.
4. Run NetWorker backup.
5. Execute NetWorker Recover.

The entire datastore or individual virtual machines are available for backup or recovery. [Figure 114](#) shows NetWorker during the process.



Item Name	Size	Date Modified	Backup Date
<input checked="" type="checkbox"/> lck-1200000000000000	84	8/8/2009 5:09 PM	8/8/2009 5:07 PM
<input checked="" type="checkbox"/> lck-1600000000000000	84	8/8/2009 5:09 PM	8/8/2009 5:07 PM
<input checked="" type="checkbox"/> lck-3000000000000000	84	8/8/2009 2:03 PM	8/8/2009 5:07 PM
<input checked="" type="checkbox"/> testmachine-ce2c2f3b.hlog	1	8/8/2009 2:11 PM	8/8/2009 5:07 PM
<input checked="" type="checkbox"/> testmachine-e9b45e4b.v...	1073741824	8/8/2009 5:08 PM	8/8/2009 5:07 PM
<input checked="" type="checkbox"/> testmachine.rvram	8684	8/8/2009 5:08 PM	8/8/2009 5:07 PM
<input checked="" type="checkbox"/> testmachine.vmsd	0	7/30/2009 11:39 AM	8/8/2009 5:07 PM
<input checked="" type="checkbox"/> testmachine.vmx	2967	8/8/2009 5:08 PM	8/8/2009 5:07 PM
<input checked="" type="checkbox"/> testmachine.vmx	1855	8/8/2009 5:08 PM	8/8/2009 5:07 PM

Figure 114 NDMP recovery using NetWorker

Set the environment variable `SNAPSURE=y` to use VNX file backup with integrated checkpoints. This feature automates checkpoint creation, management, and deletion activities by entering the environment variable in the qualified vendor backup software. [Figure 115](#) shows the `SNAPSURE` parameter set to create a backup with an integrated checkpoint.

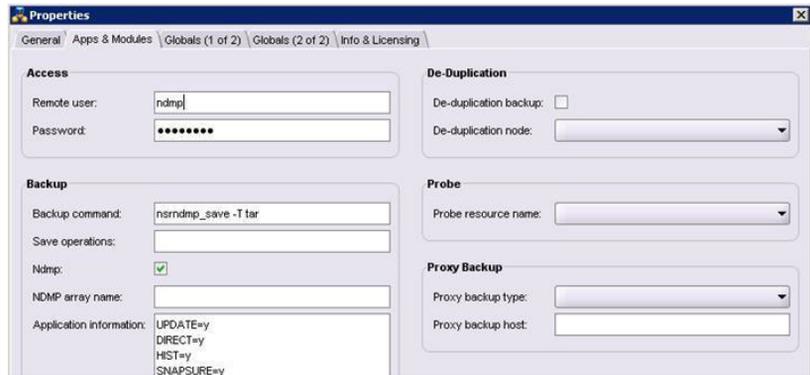


Figure 115 Backup with integrated checkpoint

If the `SNAPSURE` parameter is set to `Y`, a file system checkpoint is automatically created, and mounted as read-only each time particular jobs are run, and before the start of the NDMP backup. This automated process allows production activity to continue without interruption on the file system. The checkpoint is automatically deleted at the end of the backup operation.

Summary

This chapter provides several backup options and examples of virtual machine protection. Native options and tools on the VNX storage system create replicas or snapshots of the storage devices backing the datastores. SnapSure manages point-in-time copies of NFS datastores. LUN clones or snapshots provide similar protection for VNX block environments.

The Virtual Data Recovery appliance is deployed and configured fairly easily and populated with VNX block storage to support up to 100 virtual machines for each appliance.

In larger environments, EMC Avamar scales to significantly improve global data deduplication and reduce resource requirements for all areas of backup. EMC Avamar Virtual Edition for VMware and Avamar Image Proxy virtual appliances are quickly installed and configured with tight vCenter integration for vSphere environments. These products are backed by VNX storage to provide a scalable, efficient data protection solution.

EMC NetWorker offers an image protection option for vSphere, with tight integration with vCenter to create and manage individual virtual machine backup and restore options. NetWorker provides NDMP support for VNX OE for Block, as well as integration with VNX OE for File Virtual Tape Libraries. [Table 15 on page 220](#) summarizes some of the backup technologies and products that are used to establish image- and file-level backup approaches. The VNX storage platform and vSphere are integrated with many data protection solutions.

The information in this section and in the table is not a comprehensive list of qualified products. It is an example of the data protection options and technologies that exist within EMC VNX and VMware vSphere.

Table 15 Backup and recovery options

Storage	Backup/recovery	
	Image-level	File-level
VMFS/NFS datastore	<ul style="list-style-type: none"> • Avamar Image Proxy • NDMP • VDR • EMC NetWorker • EMC SnapSure/SnapClone 	<ul style="list-style-type: none"> • Avamar Client or File Level Recovery • EMC SnapSure/SnapView /Replication Manager
RDM (physical)	Replication Manager	N/A
RDM (virtual)	<ul style="list-style-type: none"> • VDR • Avamar Proxy • NetWorker 	<ul style="list-style-type: none"> • Avamar • NetWorker

This chapter includes the following topics:

- ◆ Introduction 222
- ◆ EMC Remote Replication technology overview 225
- ◆ RDM volume replication 247
- ◆ EMC Replication Manager 251
- ◆ Automating site failover with SRM and VNX 254
- ◆ Summary 264

Introduction

With the increased virtualization of Tier 1 applications, it is critical to have a business continuity (BC) plan for the virtualized data center. EMC VNX systems provide native features to define custom disaster recovery (DR) solutions. EMC replication technologies combine with VMware® vCenter™ Site Recovery Manager™ (SRM) to create end-to-end integrated DR solutions.

This chapter focuses on the use of EMC replication technologies and SRM to create remote DR solutions. These solutions typically include a combination of VMware virtual infrastructure and EMC storage systems located at separate data centers. EMC technologies perform the data replication between them.

This chapter covers:

- ◆ EMC replication configurations and their interaction with ESXi hosts.
- ◆ Integration of guest operating environments with EMC technologies.
- ◆ Use of SRM to manage and automate site-to-site DR with VNX.
- ◆ A review of replication options, such as:
 - EMC VNX Replicator
 - EMC MirrorView™
 - EMC RecoverPoint™

Definitions/Considerations

The following terms are used in this chapter:

- ◆ **Dependent-write consistency** — A state where data integrity is guaranteed by dependent-write I/Os. A dependent-write I/O cannot be issued until a related predecessor I/O is committed to the storage system.

- ◆ **Disaster restart** — Involves the implicit use of active logs during system initialization to ensure transactional consistency. If a database or application is shut down normally, consistency is established quickly. However, if a database or application terminates abnormally, the restart process takes longer, and is dependent on the number and size of the transactions that were in progress at the time of termination.

A replica image created from a running database or application without any preparation is considered to be restartable. This is similar to the state encountered during a power failure. As the application starts, it completes committed transactions and rolls back uncommitted transactions to achieve transactional consistency.

- ◆ **Disaster recovery** — The process of rebuilding data from a backup image, and applying subsequent logs to update the environment to a designated point of consistency. The steps required to establish recoverable copies of data are dependent on the applications being protected.
- ◆ **Roll-forward recovery** — In some cases, it is possible to apply archive logs to a database management system (DBMS) image to roll it forward to a specific point in time. This capability offers a backup strategy that consists of a baseline image backup, and archive logs to establish the recovery point.
- ◆ **Recovery point objective (RPO)** — The consistency point to be established after a failure. It is determined by the acceptable amount of data loss between the time the image was created and the time a failure occurs.
- ◆ **Recovery time objective (RTO)** — The maximum time to recover data after the declaration of a disaster. It includes the time taken to:
 - Provision power and utilities
 - Configure server software and networking
 - Restore data at the new site
 - Roll the environment forward and validate data to a known point of consistency

The following DR preparations made ahead of time reduce or eliminate delays in data recovery:

- Establish a hot site with preconfigured servers.
- Implement a storage replication solution to ensure that applications start with current data.
- Integrate that solution to provide intelligence to recover the entire infrastructure with consideration for boot order, and application and infrastructure dependencies.

Each RTO solution has a different cost profile. It is usually a compromise between the cost of the solution and the potential revenue loss when applications are unavailable.

Design considerations for DR and data restart

The effect of data loss or application unavailability varies from business to business. The tolerance for each determines the metrics and requirements for the DR solution.

When evaluating a solution, ensure that the RPO and RTO requirements of the business are met. In addition, consider the operational complexity, cost, and ability of the solution to return the entire business to a point of consistency. Each of these aspects is discussed in the following sections.

Testing the solution

A DR solution requires tested, proven, and documented procedures. Operational test procedures are often different from disaster recovery procedures.

Operational procedures are clearly documented. They are executed periodically to simulate an actual DR scenario and verify that they are up to date.

Geographically distributed vSphere environments

The integration of VNX storage system replication products and VMware technologies provides cost-effective DR and BC solutions. SRM provides the ability to establish a verifiable runbook to automate and prioritize service recovery after a failover. Some of these solutions are discussed in the following sections.

EMC Remote Replication technology overview

Business continuity solutions

Business continuity solutions for production vSphere environments require offsite or Remote Replication to ensure that reliable copies are created at a secondary location. Active data replication with EMC technologies in conjunction with SRM offers seamless solutions to automate virtual machine failover and resumption of applications and services at the remote location.

VNX offers advanced data replication solutions to help protect file systems and LUNs. In the event of a disaster, an environment failover to the remote location is accomplished with minimal administrator intervention.

EMC replication options allow objects to be grouped together and managed as a single session, or managed independently with different service levels and options for synchronous and asynchronous remote storage updates. WAN bandwidth, RPO, and data change rate drive the update frequency.

EMC provides three replication options for VNX Storage systems:

- ◆ **EMC Replicator** offers native asynchronous replication for NFS datastores.
- ◆ **EMC MirrorView** offers native synchronous and asynchronous replication for VNX Block.
- ◆ **EMC RecoverPoint** offers synchronous and asynchronous out-of-band replication for VNX block and file datastores.

Each replication technology is integrated with Replication Manager and SRM. Table 16 lists the DR and BC software options available for each storage device type.

Table 16 EMC replication options for VMware environments

Replication technology	NFS	VMFS	RDM
EMC Replicator	X		
EMC RecoverPoint CRR ¹	X	X	X
EMC MirrorView		X	X

1. File system replication takes place at the LUN level.

EMC MirrorView and RecoverPoint provide a similar set of LUN and consistency group replication capabilities. There are specific architectural differences, but from a business process standpoint, the primary differences are functional. They relate to the number of supported replicas, manageability, and ease of replica accessibility at the remote site.

EMC Replicator provides the most comprehensive solution to replicate NFS datastore file systems. MirrorView and RecoverPoint support NFS, whereas Replicator is integrated with VNX OE for File and provides the most flexibility for NFS.

Note: Replicator does not offer consistency groups for application consistency across replicated file systems. To improve application consistency, place all virtual machines in a single replicated file system, or replicate VNX OE for File LUNs with MirrorView or RecoverPoint.

The MirrorView driver is integrated with VNX OE for Block. It intercepts I/O sent to a source device and mirrors these writes to a LUN on a remote VNX. MirrorView supports a considerable number of replication sessions for one-to-one replication of many VNX LUNs. It provides a good LUN-level replication solution between storage systems.

RecoverPoint is the most flexible replication technology, and provides a level of granularity that is useful for integration with applications and business processes. RecoverPoint offers a significant number of point-in-time copies (bookmarks), which provide the flexibility to establish precise point-in-time images of the virtual storage devices.

EMC Replicator

EMC Replicator offers native file system replication for NFS datastores. Replicator is an asynchronous replication solution that performs local or remote file system replication within or between VNX systems. Replicator keeps remote file systems consistent with the production environment for upwards of 1024 separate file system sessions per VNX Data Mover.

User-specified update periods define the interval at which Replicator updates the remote file system. By default, a new delta set of accumulated changes is sent to the remote system every 10 minutes. At the remote site, delta sets are played back to update the remote file system. Replication sessions are customized with different update intervals and quality-of-service settings to prioritize updates between NFS datastores.

EMC Replicator operates at the file system level. Therefore, it encapsulates all of the virtual machines and files contained within an NFS datastore. It is a good practice to group virtual machines with similar protection requirements to improve the reliability and efficacy of the DR solution. Organize virtual machines at a file system level to facilitate prioritization of DR policies in accordance with RPOs.

Replicating a NAS file system

Complete the following steps in Unisphere for remote file system replication:

1. Locate and select the **Data Protection** tab from the Unisphere home interface.
2. Click **File Replication Wizard - Unisphere**. The **Replication Wizard** appears.

3. Complete the following steps as shown in [Figure 116 on page 229](#) and [Figure 117 on page 230](#).
 - a. Select **File System** as the replication type.
 - b. Select **Ongoing File System Replication** to display the list of destination VNX network servers.
 - c. Select the destination VNX system to create a read-only, point-in-time copy of a source file system at the destination.

Note: The destination can be the same Data Mover (loop back replication), another Data Mover in the same VNX cabinet, or a Data Mover in a different VNX cabinet.

- d. Select the network interface to transfer the replication delta sets. Replicator requires a dedicated network interconnect between the source and destination Data Movers. The wizard defaults to the first configured interface in the list. Select the most appropriate interface to support replication between Data Movers.

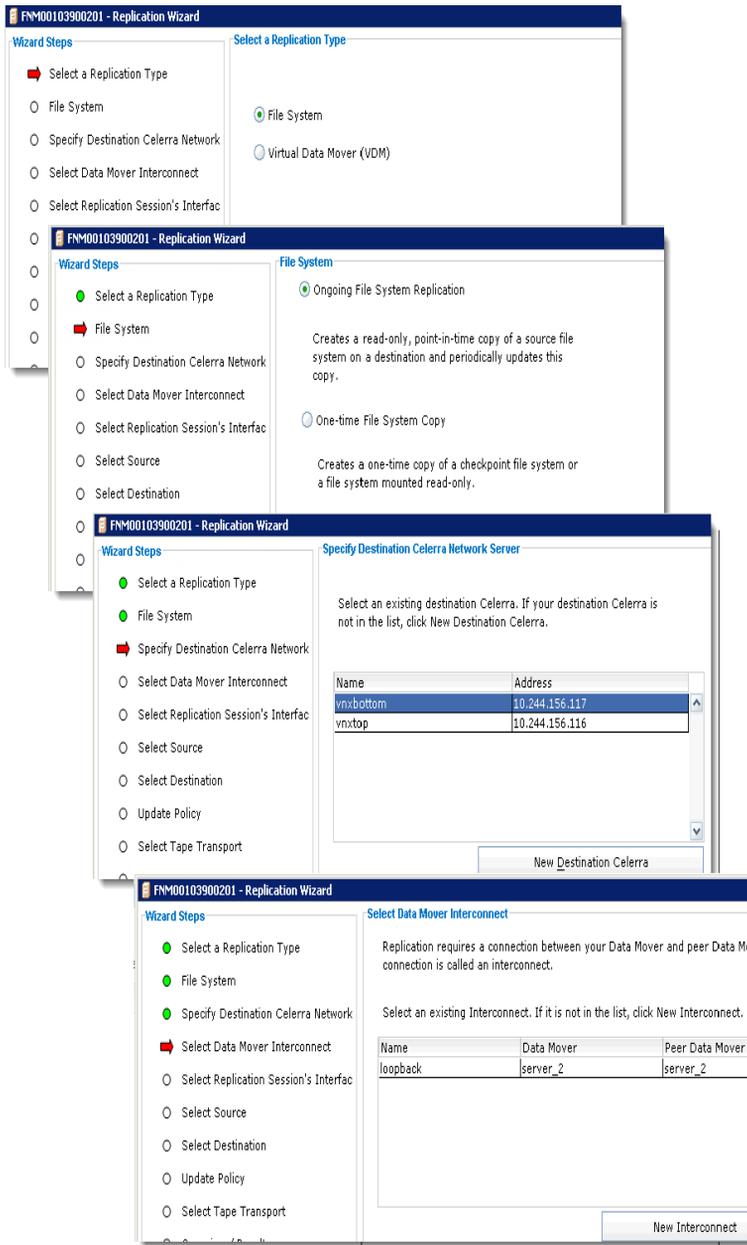


Figure 116 Replication Wizard

- e. Specify a name for the replication session.
- f. Select the source file system to replicate to the remote location.

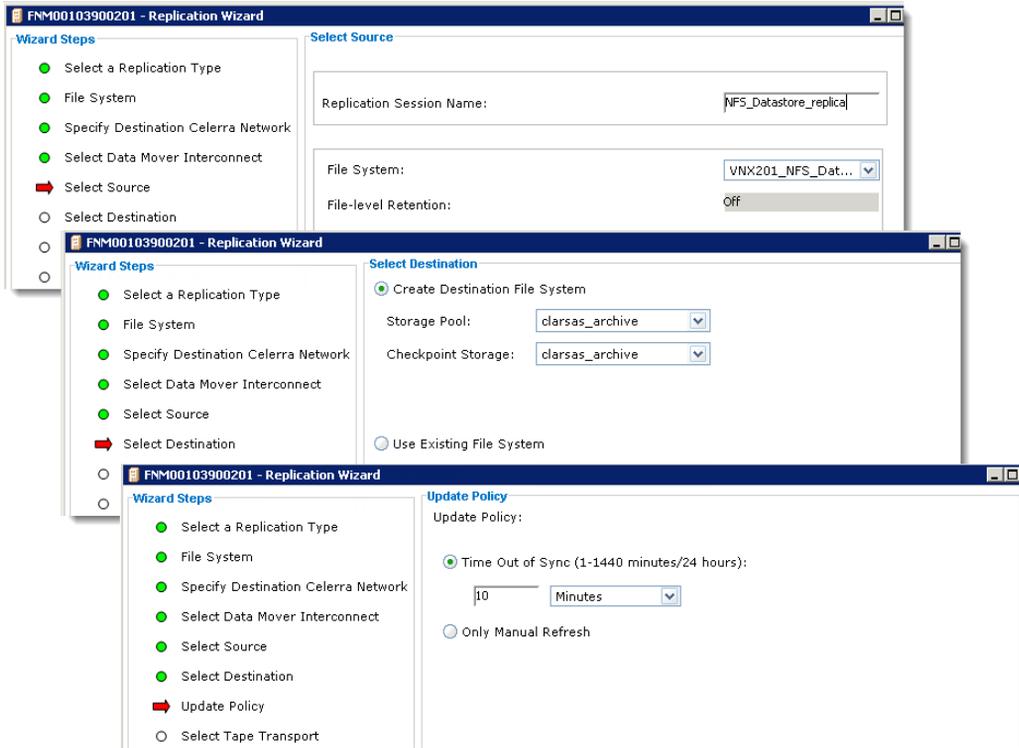


Figure 117 Replication Wizard (continued)

- g. Select a file system at the destination to support the replication session. If a file system does not exist, create one and then click **Next**.

Note: When Replicator is used to create a destination file system, it assigns the name and the size of the destination file system based on the properties of the source file system. Administrators select a storage pool for the destination file system, and a storage pool for checkpoints. Assign a descriptive name with an identifier, such as "DR" to help identify the replication relationship.

- h. Select the interval at which to update the secondary site.

After the file systems are synchronized, the remote image transitions to an operational read-only state. To use an NFS datastore at the remote location, mount the file system as read/write by using any one of the following options:

- ◆ Initiating a failover
- ◆ Terminating the replication session
- ◆ Reversing the replication.

This action promotes the storage devices at the remote location. It collects changes from that environment, and applies them to the previous source location.

After the file system is mounted as read/write, present it to the ESXi host and manually register the virtual machines.

EMC MirrorView

EMC MirrorView supports options for synchronous and asynchronous replication of VNX block storage between separate VNX storage systems. Replication data is transported over Fibre Channel or iSCSI connections established between the storage systems. Protection is assigned to individual LUNs, or to a consistency group.

MirrorView LUN replication

In an ESXi host environment, VMFS datastore LUNs are replicated to establish a synchronous datastore copy at a remote location. Secondary devices undergo an initialization period to establish a block-for-block image of the source device. MirrorView has two usable LUN states, *synchronized*, and *consistent*. In a *synchronized* state, the remote LUN is an identical block-for-block copy of the source LUN. In a *consistent* state, the remote LUN is synchronized, but has changed state because the mirror received updates that are not applied to the LUN. The time period that establishes when a mirror transitions from the *consistent* state to the *synchronized* state after an update is called the *quiesce threshold*. The default value is 60 seconds of no host I/O to the mirror. A LUN or consistency group at the remote location is promoted and used by ESXi when it is in either of these states.

For multiple LUNs, it is a good practice to use a consistency group. [Table 17 on page 232](#) lists the MirrorView limits for the VNX platforms.

Table 17 VNX MirrorView limits

	VNX5100	VNX5300	VNX5500	VNX5700	VNX7500
Maximum number of mirrors	128	128	256	512	1024
Maximum number of consistency groups	64	64	64	64	64
Maximum number of mirrors per consistency group	32	32	32	64	64

MirrorView consistency group

A MirrorView consistency group is a collection of mirrored devices that are treated as a single object within a VNX storage system. Operations such as synchronization, promotion, and fracture, are applied to all components of the consistency group. If an event impacts the state of the consistency group, I/O is suspended to all components of the consistency group to preserve write-ordered I/O to the LUNs and the applications they serve.

All members of a consistency group are owned by different storage processors, but they are on the same VNX storage system.

Although synchronous and asynchronous mirrors are supported on consistency groups, all LUNs in a consistency group are protected by the same replication mode. VNX supports 32 LUNs per consistency group for MirrorView (synchronous and asynchronous). [Figure 118 on page 233](#) shows an example of a consistency group with four LUNs. Use MirrorView consistency groups with SRM configurations.

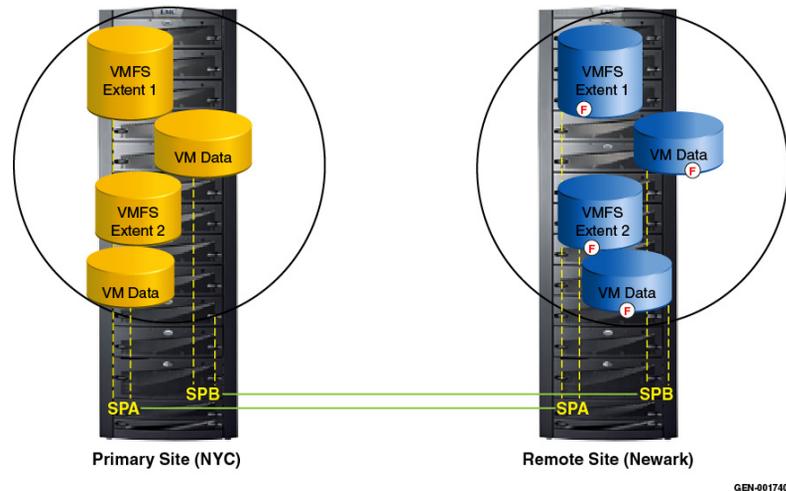


Figure 118 Preserving dependent-write consistency with MirrorView consistency group technology

In this example, a communication failure results in a fracture of the MirrorView link between the storage processors on the local and remote VNX storage systems. At the point of disruption, MirrorView fractures all LUN mirrors in the consistency group. While the secondary images are fractured, updates to the primary volumes are

not propagated to the secondary volumes to preserve data consistency. At this time, the writes to the production LUNs are tracked in a log called a write-intent log. After the error is corrected, all the updates are applied to the consistency group on the remote system.

Asynchronous MirrorView (MV/A)

MirrorView/A is an asynchronous method used to replicate up to 256 LUNs between VNX systems. With MirrorView/A, host writes are acknowledged immediately and buffered at the source VNX. At an administrator-defined interval, MirrorView creates a differential LUN view and copies the changed blocks to the remote VNX to create consistent, write-ordered, point-in-time copies of the production LUN. A gold copy of the target data is created prior to the source or target updates. This copy preserves the data on the target side in case the transfer is disrupted.

The asynchronous nature of MV/A replication implies a non-zero RPO. MV/A is designed to provide customers with an RPO greater than or equal to 30 minutes. There are no distance limitations between the source and target VNX storage systems.

Synchronous MirrorView (MV/S)

MirrorView/S provides synchronous replication for LUNs or consistency groups and ensures that each I/O is replicated to a remote system. Synchronous replication for vSphere maintains lockstep consistency between the primary and secondary storage locations. Write-operations from the virtual machine are not acknowledged until both VNX arrays have a copy of the data in their write caches. These updates incur a propagation delay resulting from the distance and quality of the network. As a result of that delay, MV/S is not suitable for locations separated by distances greater than 100 kilometers.

Complete the following steps to set up MirrorView replication in Unisphere. When configuring MirrorView, use the **Virtualization** tab in Unisphere or the VSI Storage Viewer feature to identify LUN numbers and their relationships to the VMFS datastores and RDM devices, as shown in Figure 119.

Note: The process and commands to configure synchronous and asynchronous MirrorView replication are very similar. Specify the `-async` argument for asynchronous replication.

The screenshot displays the EMC VMware Unisphere interface. The main window shows the 'VMware Infrastructure' view with a list of hosts and virtual machines. A callout box titled 'Virtual Storage Details' lists: LUN Name, Datastore Name, VMDK Capacity, and VMDK usage when Thin.

The 'Virtual Machines' table is as follows:

VM Name	Guest Host	VM IP	Guest OS
SQL03.sr5dm.eng.emc.com	Unknown	Unknown	Unknown
SQL-1	SQL1	10.244.156.82	Microsoft Windows Server 2008 R2 (64-bit)
vdi-snap	vdi-snap	10.244.156.79	Other Linux (32-bit)
vdmclient	vdmclient	10.244.156.78	Microsoft Windows XP Professional (32-bit)
vdrovf	vdrovf	10.244.156.84	Other 2.6x Linux (64-bit)
view4	Unknown	Unknown	Unknown
ViewManager4	ViewManager4	10.244.156.18	Unknown

The 'Virtual Machine Properties' window for 'SQL1(blade16.emc.lab)' is open, showing the 'Storage' tab. It displays 'LUN Mapping for SQL1 on VMWare ESX Server blade16.emc.lab (10.244.156.82)'. A callout box titled 'Virtual Machine LUN details' lists: RDM or Datastore Name, SP ownership, and Tresspassed status.

The 'LUN Mapping' table is as follows:

Name	Device Mapping	Device Name	Storage System
LUN 10	Datastore (vnx-vm-library)	naa.600601601ae02900debc23ec3547e011	FNMM00103900200
LUN 7	Mapped Raw LUN	naa.600601601ae029004e6219286459e011	FNMM00103900200

The 'Virtual Machine Information' table is as follows:

Name	Type	LUN Names	Disk Mode	Disk Capacity
Hard disk 1	Virtual Disk - Thin	LUN 10	Persistent	33.95G (26.11G)
Hard disk 2	Mapped Raw LUN - Physical	LUN 7	Independent Persistent	50.00G
SQL-1	VM Configuration	LUN 10	N/A	N/A
Hard disk 2 Mapping File	Datastore Mapping File	LUN 10	N/A	N/A

Figure 119 EMC VMware Unisphere interface

1. From the **Unisphere Data Protection** window, select **Manage Mirror Connections**.
2. Identify the Peer Storage System and enable the MirrorView connection between the two systems as shown in **Figure 120**.

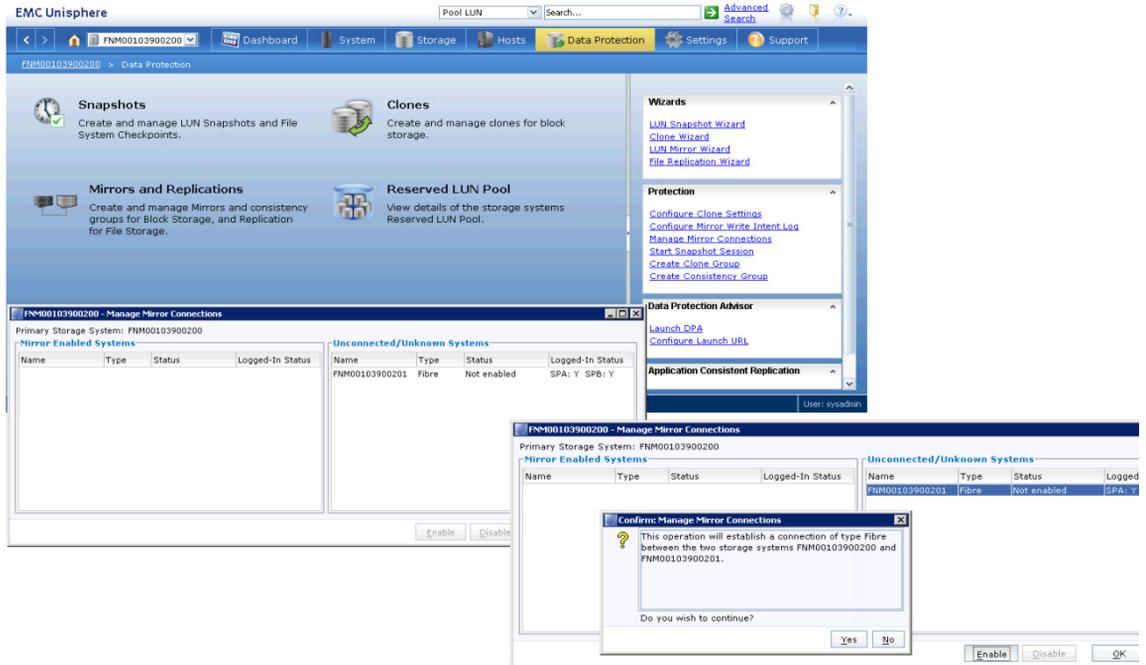


Figure 120 Enable MirrorView between VNX systems

- Use the Unisphere **MirrorView LUN** wizard to select the source LUNs and establish a remote mirror at the recovery site as shown in [Figure 121](#).

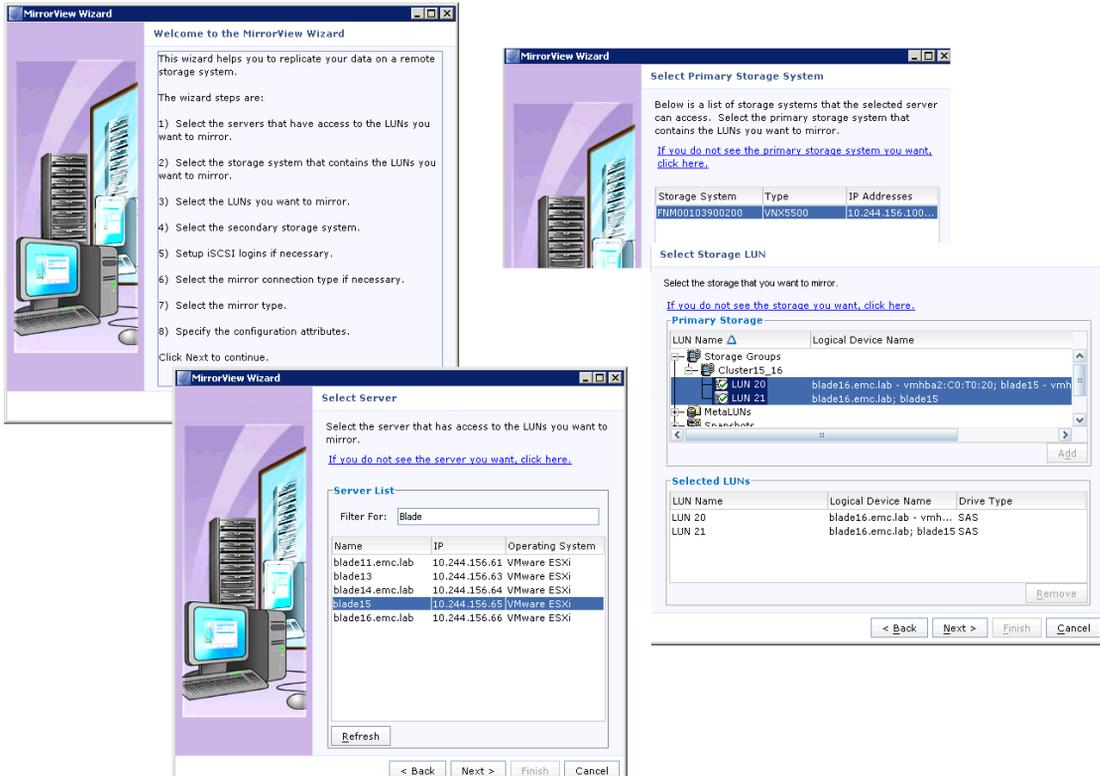


Figure 121 MirrorView Wizard — select source LUNs

4. Select the remote storage pools to use for the MirrorView session as shown in Figure 122.

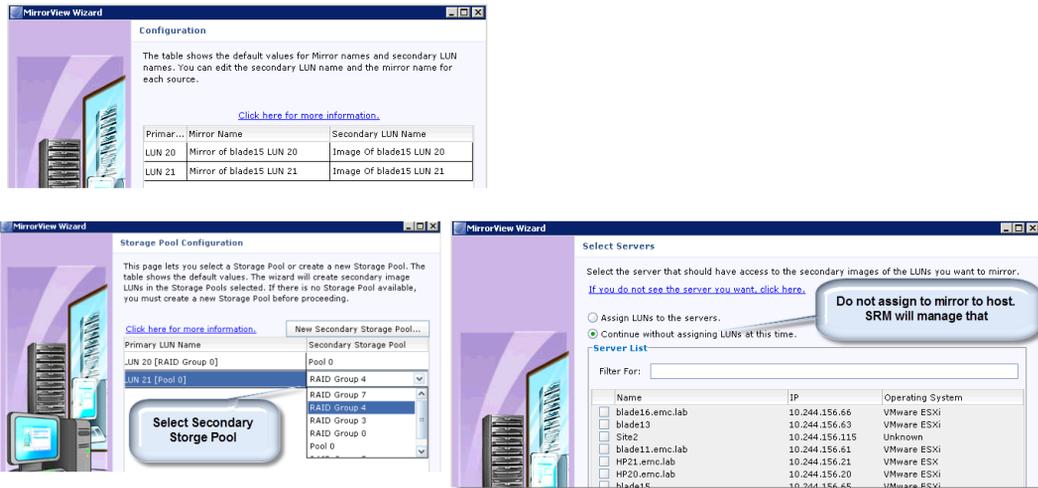


Figure 122 MirrorView Wizard — select remote storage

- Promote the secondary image at the DR site as shown in [Figure 123](#).

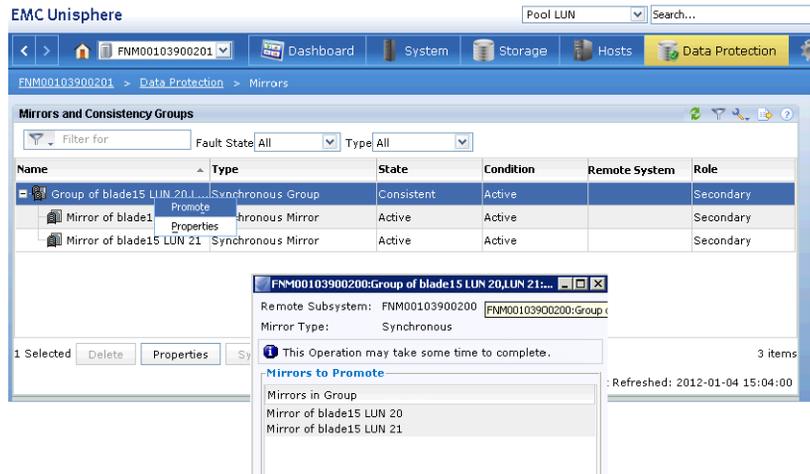


Figure 123 Promote mirrored LUN

Note: When a secondary image is in a synchronized or consistent state, SnapView clones or snapshots provide the ability to create consistent, point-in-time copies of the image without promoting it and disrupting the MirrorView session.

Figure 124 on page 240 shows a schematic representation of a business continuity solution that integrates VMware vSphere and MirrorView. The figure shows two virtual machines accessing VNX LUNs as RDM volumes.

The solution provides a method to consolidate the virtual infrastructure at the remote site. Because virtual machines can run on any ESXi host in the cluster, fewer ESXi hosts are required to support the replicated virtual machines at the remote location.

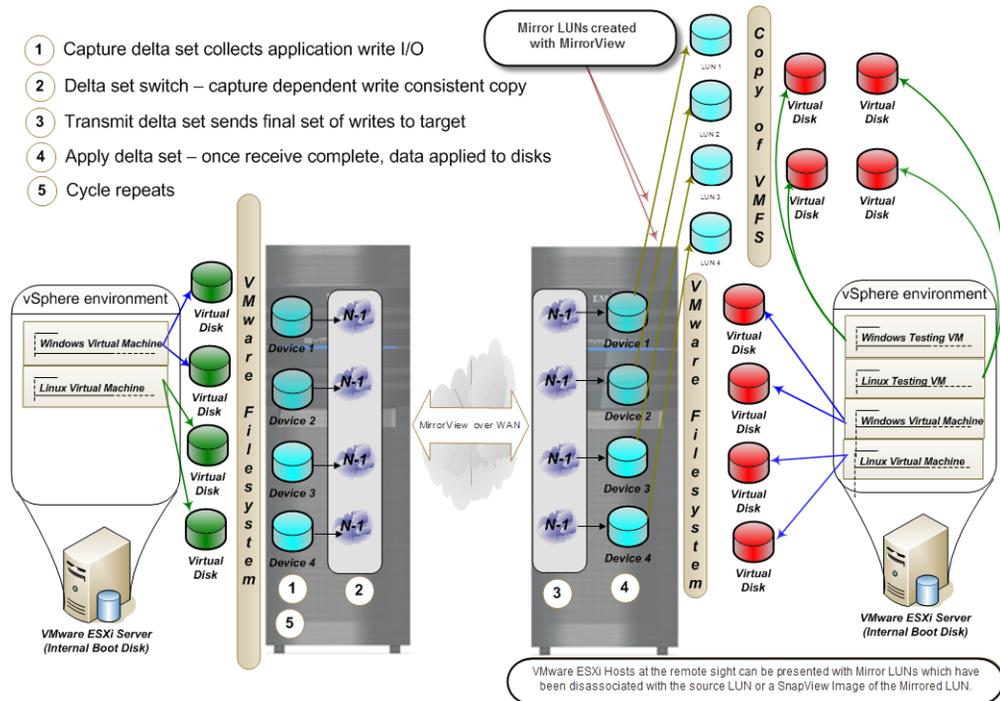


Figure 124 Business continuity solution using MirrorView/S in a virtual infrastructure with VMFS

Failover MirrorView LUNs to a remote site using CLI

MirrorView LUNs or consistency groups are activated at the secondary site during the failover process. The result is that all devices are transitioned to a writeable state and are available to restart applications in that environment. In a planned failover, disable or shut down the VNX at the production site before performing the failover tasks.

To prevent data loss, synchronize secondary MirrorView /S LUNs before starting the failover process. Shut down the applications at the production site, and update the secondary image manually.

Right-click a consistency group and select **Synchronize** to synchronize all LUNs as shown in [Figure 125](#).



Figure 125 Synchronize MirrorView LUNs

MirrorView LUN synchronization performs the following changes:

- ◆ Sets the primary images on the production site to write-disabled.
- ◆ Reverses the mirror relationship of the devices. The devices at the remote site assume the primary role and are set to write-enabled.
- ◆ Resumes the MirrorView link to allow updates to flow from the remote data center to the production data center.
- ◆ Registers and powers on the virtual machine from the vSphere client or command line utilities.

Administrators use RecoverPoint to:

- ◆ Support flexible levels of protection without distance limitations or performance degradation. RecoverPoint offers fine-grain recovery for VMFS and RDM devices that reduce the recovery point through frequent updates to the replica devices.
- ◆ Replicate block storage to a remote location through a cluster of tightly coupled servers.
- ◆ Use write splitters that reside on the VNX arrays or hosts in the SAN fabric. The write splitter copies write I/Os destined for the ESXi datastore volumes and sends them to the RecoverPoint appliance. The RecoverPoint appliance transmits them to the remote location over IP networks as shown in [Figure 126 on page 242](#).
- ◆ Provide a full-featured replication and continuous data protection solution for VMware ESXi hosts. For Remote Replication, RecoverPoint CRR uses small-aperture snapshot images to provide a low RPO, or asynchronous replication with a small RPO to provide VMware protection and guarantee recoverability with little or no data loss.

Virtual machine write splitting

For VMware, RecoverPoint provides a host-based write splitter to support application integration for Windows virtual machines. The driver filters write operations to each protected RDM volume and ensures that each write command is sent to the RecoverPoint appliance. Since the splitter or KDriver runs on the virtual machine, only SAN volumes attached to virtual machine in physical RDM mode (pRDM) are replicated by RecoverPoint.

RecoverPoint VAAI support

vSphere version 5.1 provides full support for VAAI with the VNX splitter. [Table 18](#) illustrates the minimum releases for VAAI support with the VNX RecoverPoint splitter. Versions of the VNX splitter or VNX OE for Block code prior to those listed in the table only support Hardware Accelerated Locking (ATS) for block storage devices. ATS is the only SCSI command supported for VNX and RecoverPoint versions previous to those listed in [Table 18](#). If running a prior version, SCSI commands other than ATS are rejected and revert to the host for processing.

Table 18 Minimum revision levels for VAAI support with VNX RecoverPoint splitter

VAAI Primitive	VNX Revision level	Notes
Hardware Assisted Locking	VNX splitter 3.4 with FLARE 31 and later	Supported
Block Zeroing	VNX splitter 3.4 with FLARE 31 and later	Supported
Full Copy	VNX splitter 3.4 with FLARE 31 and later	Supported, without performance enhancement
Uncopy	VNX splitter 3.5 SP1 and later	Supported

Note: The RecoverPoint SAN splitter Storage Services Interface earlier than version 4.2(3K) does not support VAAI SCSI commands. For SAN splitters prior to SSI 4.2(3K), disable VAAI to use the SAN splitter.

Figure 127 illustrates the Data Mover advanced settings interface for VAAI Hardware Accelerated Move (XCOPY) and Hardware Accelerated Init (Write-Same). Set the value of these parameters to zero to disable XCOPY and Write-Same support on the ESXi host.

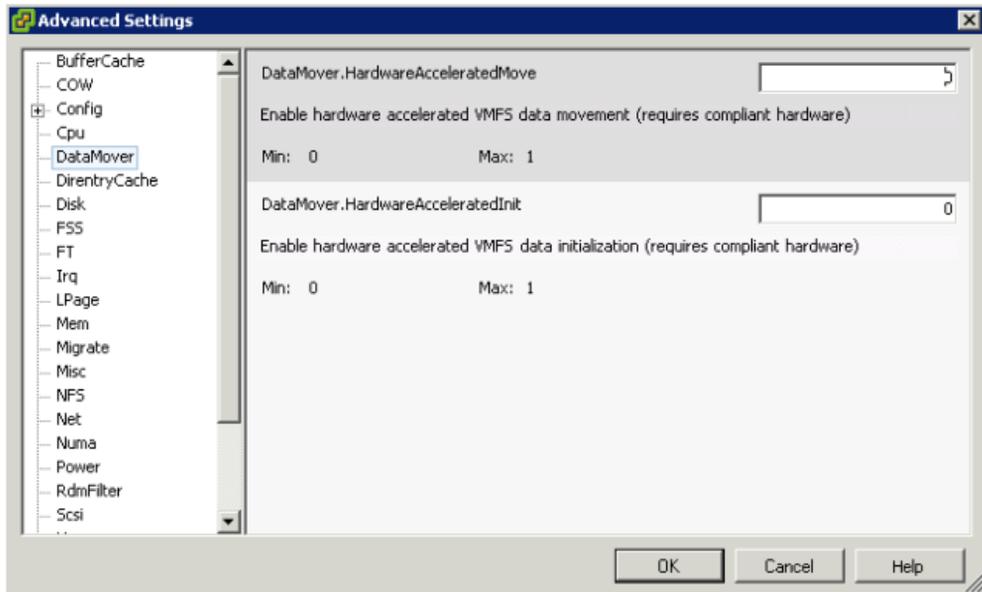


Figure 127 Disabling VAAI support on an ESXi host

RecoverPoint provides consistency groups to assign VNX storage devices to ESXi hosts. Each consistency group is made up of LUNs that are protected. A journal LUN (volume) is also assigned to each consistency group to maintain the bookmarks and the various states provided with RecoverPoint. Separate VNX storage groups are created for the RecoverPoint appliance and ESXi host HBAs. LUNs that require protection are assigned to both storage groups.

Configure the consistency groups, apply policies, and manage storage access through the RecoverPoint management UI or CLI.

Note: All virtual disks that constitute a virtual machine are a part of the same consistency group. If application consistency is required when using RDMS, install the RecoverPoint driver in the Windows guest OS. [Table 19 on page 246](#) summarizes the support options available with RecoverPoint for VNX replication.

Table 19 EMC RecoverPoint feature support

Feature	Splitter		
	Windows host write splitter	Array-based write splitter	Brocade/Cisco Intelligent Fabric write splitter
Supports physical RDM	Yes	Yes	Yes
Supports virtual RDM	No	Yes	Yes
Supports VMFS	No	Yes	Yes
Supports VMotion®	No	Yes	Yes
Supports HA/DRS	No	Yes	Yes
Supports vCenter Site Recovery Manager	No	Yes	Yes
Supports P2V replication	RDM/P only	RDM/P and VMFS	RDM/P and VMFS
Supports V2V replication	RDM/P only	RDM/P and VMFS	RDM/P and VMFS
Supports guest OS Boot from SAN	RDM/P only	RDM/P and VMFS	RDM/P and VMFS
Supports ESXi Boot from SAN	No	Yes	Yes
Maximum number of LUNs supported per ESXi hosts	255 (VMware restriction)	N/A	N/A
Heterogeneous array support	EMC VNX, CLARiiON CX, Symmetrix and, selected third party storage	EMC VNX and CLARiiON CX3/CX4	EMC and third party
Shareable between RecoverPoint clusters	No	Yes	Yes

RDM volume replication

Replication of RDMs requires the completion of additional management tasks than datastore replication of VMFS LUNs. RDM volumes are separate physical devices assigned directly to the virtual machines without the hypervisor I/O path. As a result, the ESXi host does not have a device ID or LUN signature to identify the device on the remote host. The RDM device paths are preserved at the OS level to ensure OS and application integrity.

EMC Replication Manager interacts with EMC replication technologies to manage the remote replicas and preserve the device mappings of NTFS-formatted pRDM volumes.

Configuring remote sites for vSphere virtual machines with RDM

When an RDM is added to a virtual machine, a virtual disk file that maps the logical virtual machine device to the physical device is created. The file contains the VNX LUN WWN and LUN number of the device presented to the virtual machine.

The virtual machine configuration is updated with the name of the RDM volume and the label of the VMFS datastore where the RDM volume resides. When the datastore that contains the virtual machine is replicated to a remote location, it maintains the configuration and virtual disk file information. However, the target LUN has a different UUID that results in a configuration error if the virtual machine is powered on.

Snapshots and clone LUNs are used to validate the configuration because they are presented to hosts or virtual machines without disrupting the replication session. They are also beneficial for ancillary purposes such as QA or backup.

The most important consideration for RDM replication is to ensure that SCSI disks maintain the same device order within the Guest OS. This requires precise mapping of the VNX LUNs to the virtual machine at the secondary site.

Determine the device mapping for the ESXi hosts and document the disk order for the devices presented to the virtual machines on the remote site. Table 20 shows an example with three application data disks.

Table 20 VNX to virtual machine RDM

LUN number	Windows disk	Virtual device node
2	\\.\PHYSICALDRIVE2	SCSI (0:1)
3	\\.\PHYSICALDRIVE3	SCSI (0:2)
4	\\.\PHYSICALDRIVE4	SCSI (0:3)

These three VNX LUNs are replicated to a remote VNX. Exclude the boot device that occupies SCSI target 0:0 and configure the virtual machine at the remote site to present the following:

- ◆ Replicated LUN associated with LUN 2 as SCSI disk 0:1
- ◆ Replicated LUN 3 as SCSI disk 0:2
- ◆ Replicated LUN 4 as SCSI disk 0:3

Use a copy of the source virtual machine configuration file instead of replicating the VMware file system. Complete the following steps to create copies of the production virtual machine by using RDMs at the remote site:

1. Create a directory within a cluster datastore at the remote location to store the replicated virtual machine files.

Note: Select a datastore that is not part of the current replication configuration to perform this one-time operation.

2. Copy the configuration file of the source virtual machine to the directory.
3. Register the cloned virtual machine through the vSphere Client or the service console.
4. Configure the ESXi hosts at the remote site to use the secondary MirrorView LUNs as RDM devices.

5. Use the vSphere Client or service console to power on the virtual machine at the remote site.

Note: As the tasks described here present configuration risks, they are best supported with SRM or through an automated Power Shell scripted utility.

Starting virtual machines at a remote site after a disaster

Complete the following steps to restart virtual machines at the remote site with the replicated copy of the data:

1. Verify that the replicas are in a synchronized or consistent state.
2. Promote the replica LUNs, file systems, or consistency groups at the remote site. Promoting a LUN changes the state of the device to write-enabled, which makes it usable by the ESXi hosts in the remote environment.
3. Add the promoted devices to the ESXi storage groups to allow the ESXi hosts access to the secondary images.
4. Rescan the SCSI bus to discover the new devices for block storage.
5. Power on the cloned virtual machines with the vSphere Client or the CLI.

Configure remote sites for virtual machines using VMFS

The management of virtual machines on a replicated VMFS volume is very similar to that of an RDM volume.

Complete the following steps to create virtual machines at the remote site:

1. Promote the secondary LUN images to make them write-enabled and accessible by the VMware ESXi cluster group at the remote data center.
2. Use the vSphere Client to initiate an SCSI bus rescan after surfacing the target devices to the VMware ESXi hosts.

3. Use the vSphere Client **Add Storage** wizard to select the replicated devices that contain the copy of the VMware file systems. Select the **Keep existing signature** option for each LUN copy. After all the devices are processed, the VMware file systems are displayed on the **Storage** tab of the vSphere Client interface.
4. Browse the datastores with the vSphere Client, to identify and register the virtual machines.

Note: Duplicate virtual machine names are unintentionally introduced when using replication services. vCenter does not allow duplicate names within the same datacenter. If a duplicate object name is encountered, assign a new virtual machine name to complete the registration.

5. Verify that the following requirements are met to ensure the virtual machines on the ESXi hosts at the remote site start without any modification:
 - The target ESXi host has the same virtual network switch configuration as the source ESXi host. For example, the name and number of virtual switches are duplicated from the source ESXi cluster group.
 - All VMware file systems used by the source virtual machines are replicated.
 - The minimum resource requirements of all cloned virtual machines are supported on the target ESXi hosts.
 - Peripheral devices such as CD-ROM and floppy drives are attached to physical hardware, or set to a disconnected state on the virtual machines.
6. Power on the cloned virtual machines from vCenter or the command line when required. If vCenter generates a `msg.uuid.altered` message, select the **copied** option to complete the power-on procedure.

EMC Replication Manager

EMC Replication Manager (RM) supports all of the EMC replication technologies. RM simplifies the creation and management of storage device replicas through Application Sets. An Application Set includes the replication job details and any tasks required to place applications running inside the virtual machines in a consistent state prior to creating a replica of a virtual machine or datastore.

In a VMware environment, RM uses a proxy host (physical or virtual) to initiate management tasks on vCenter and VNX. The RM proxy service runs on the same physical or virtual host as the RM server.

Other requirements include:

- ◆ The proxy host is configured with:
 - RM agent
 - EMC Solutions Enabler for VNX Block
 - Navisphere Secure CLI for VNX Block
 - Administrative access to the VNX storage systems
- ◆ If application consistency within the guest virtual machine is required, install the RM agent on the virtual machine.
- ◆ The environment has a proper DNS configuration to allow the proxy host to resolve the hostnames of the RM server, the mount host, and the VNX Control Station.

When an Application Set is initiated on a VNX device containing virtual machines, the RM proxy sends a vCenter request to create VMware snapshots of all online virtual machines that reside on the ESXi datastore. This step ensures that the resulting replica is OS consistent. [Figure 128](#) shows a NAS datastore replica in the RM.

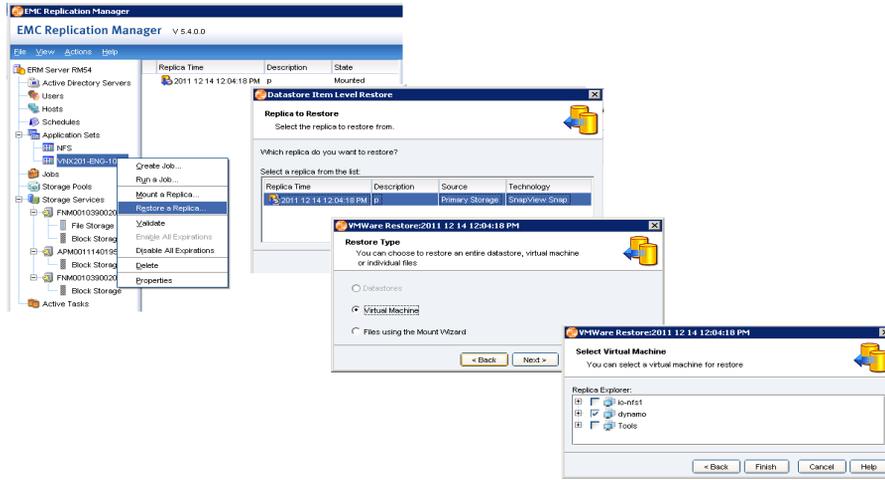


Figure 128 RM protection for NFS datastores and virtual machines

RM includes the option to mount a replicated device to another ESXi host. After a failover operation, RM performs all the necessary steps to change the device state and mount and import the datastore into the ESXi host environment. Additional administrative tasks, such as starting virtual machines and applications, are defined within the Application Set and automated through RM.

Unisphere provides the option to administratively fail over file systems to a remote location. After the failover, the file systems are mounted on the remote ESXi host. Virtual machines that reside in the datastores are optionally registered through the vSphere Client.

Complete the following steps to register virtual machines in the vSphere Client:

1. Use the datastore browser to select a virtual machine folder.
2. Locate and right-click the configuration (.vmx) file, and then select **Add to Inventory** to register the virtual machine with an ESXi host as shown in [Figure 129](#).

Note: The ESXi host names for virtual machine networks, VMkernel, and similar properties are identical to the source. Inconsistent network names result in accessibility issues.

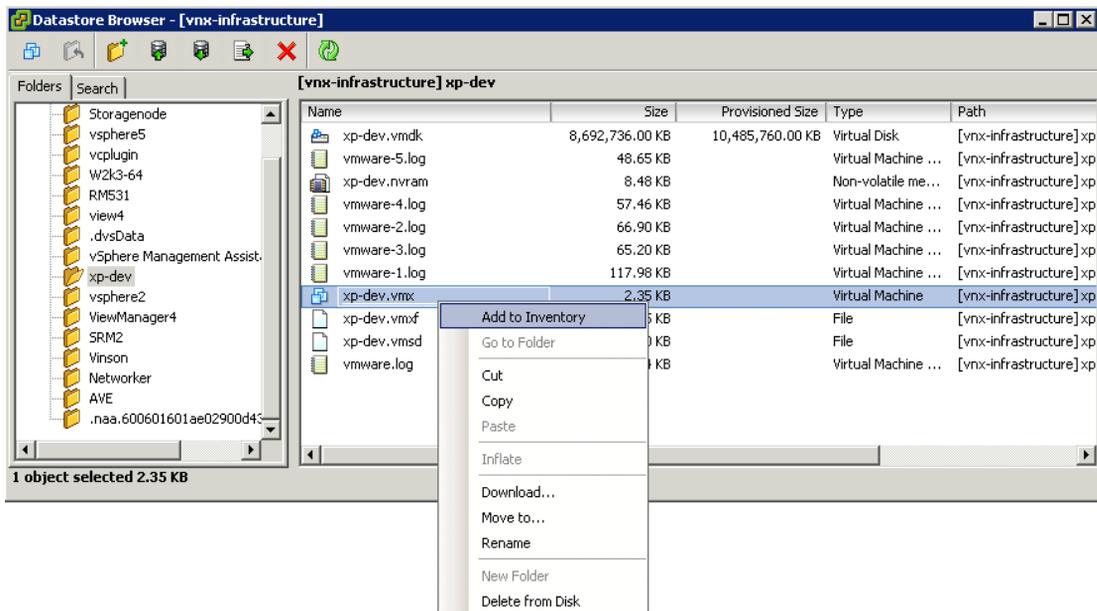


Figure 129 Using the vSphere client to register a virtual machine with ESXi

Automating site failover with SRM and VNX

VMware vCenter Site Recovery Manager (SRM) provides a standardized framework to automate VMware site failover. SRM is integrated with vCenter and EMC storage systems. It is managed through a vCenter client plug-in that provides configuration utilities and wizards to define, test and, execute failover processes called recovery plans. A recovery plan defines which assets are failed over, and the order in which they are restored when the plan is executed. SRM includes capabilities to execute pre- and post-failover scripts to assist in preparing and restoring the environment.

SRM testing

An attractive feature of SRM is provided through recovery plan validation tests which allow a failover to be simulated in advance of an actual site outage. During the recovery plan validation test, production virtual machines at the protected site continue to run, and the replication sessions remain active for all the replicated LUNs or file systems.

When the **test failover** command is run, SRM simulates the storage device failover by issuing commands to the VNX to generate writeable snapshots at the recovery site. The snapshot LUNs or file systems are mounted to the ESXi hosts. Virtual machines are powered on and optional post-power-on scripts are run. The test recovery executes the same steps as a failover does. Therefore, a successful test process increases the likelihood of a successful failover. Companies realize a greater level of confidence when they know that their users are trained on the disaster recovery process, and execute it correctly each time. Administrators have the ability to add test-specific customization to the workflow for the test failover to handle scenarios where the test differs from the actual failover scenario. If the virtual machines are powered on successfully, the SRM test process is complete. If necessary, users can start applications and perform validation tests. Run the Cleanup task to revert the environment to the pretest state and remove any temporary storage devices that were created as part of the test as shown in [Figure 130 on page 255](#).

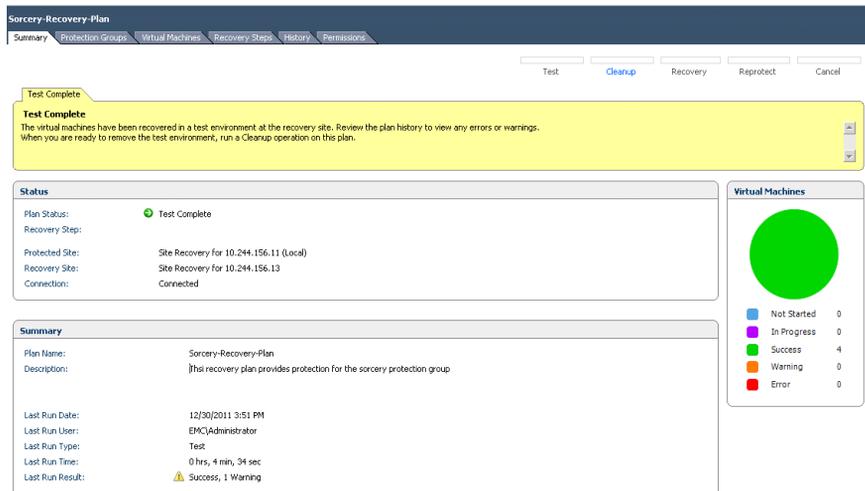


Figure 130 SRM recovery plan summary

Actual failover, or Recovery as it is called in SRM, is similar to the test failover, except that, rather than using snapshots, the actual storage devices are failed over to a remote location. LUNs and file systems at the recovery site are brought online, and the virtual machines are powered on.

During failover, SRM powers off active virtual machines at the protected site to avoid having active virtual machines at both sites. This task will not complete if the protected site is not operational.

EMC Storage Replication Adapter

SRM leverages the data replication capabilities of the underlying storage system through an interface called a Storage Replication Adapter (SRA). SRM supports SRAs for EMC Replicator, EMC MirrorView, and EMC RecoverPoint.

Each EMC SRA is a software package that enables SRM to implement disaster recovery for virtual machines by using VNX storage systems that run replication software. SRA-specific scripts support array discovery, replicated LUN discovery, test failover, failback, and actual

failover. Disaster recovery plans provide the interface to define failover policies for virtual machines running on NFS, VMFS, and RDM storage.

Figure 131 shows an example of SRM configuration in vCenter.

The screenshot displays the SRM configuration in vCenter for an Array Manager named VNX200. The interface is divided into two main sections.

Discovered Array Pairs - VNX200

After an array manager has been added for each site, click Enable to enable array pairs for use with SRM. You only need to enable the array pairs once, and this can be done...

Local Array	Remote Array	Remote Array Manager	Status	Actions
50:06:01:60:BD:ED:06:A0	50:06:01:60:BD:ED:06:A8		Disabled	Enable Disable

Devices for Enabled Array Pairs

Local devices are shown here for each enabled array pair. Remote device information is only available when the remote site is connected.

Devices for Array Pair: 50:06:01:60:BD:ED:06:A0 - 50:06:01:60:BD:ED:06:A8

Local Array Manager:	VNX200
Local Array:	50:06:01:60:BD:ED:06:A0
Remote Array Manager:	VNX201
Remote Array:	50:06:01:60:BD:ED:06:A8
Errors:	None

Replication direction for consistency group

Local Device	Direction	Remote Device	Datastore	Protection Group	Local Consistency Group
Mirror of blade15 ...	→	Mirror of blade15 LUN 21	Local: [VNX200-SRM2]		Group of blade15 LUN 20, LUN 21
Mirror of blade15 ...	→	Mirror of blade15 LUN 20	Local: [VNX200-SRM1]		Group of blade15 LUN 20, LUN 21

Figure 131 VMware vCenter SRM configuration

SRM protection groups at the protected site

A protection group consists of one or more replicated datastores that contain virtual machines and templates. It specifies the items to be transitioned to the recovery site in the event of a disaster. A

protection group establishes virtual machine protection and maps virtual machine resources from the primary site to the recovery site. There is a one-to-one mapping between an SRM protection group and a VNX or RecoverPoint consistency group. Figure 132 illustrates the configuration of a protection group that uses a MirrorView LUN consistency group.

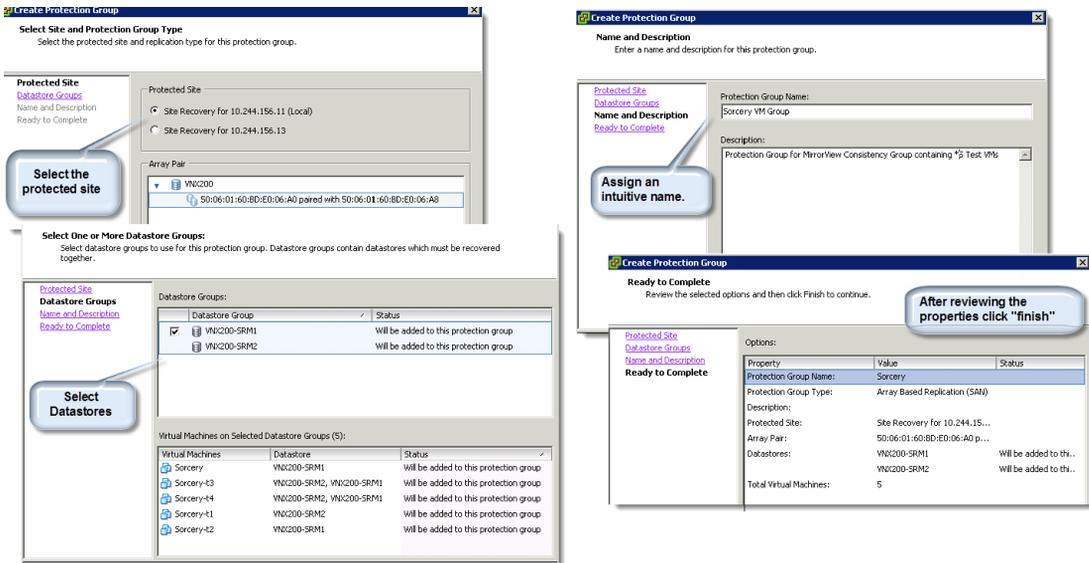


Figure 132 Create an SRM protection group

Note: There are cases that do not use a one-to-one mapping. An example is when RecoverPoint is used to protect a database application with separate consistency groups for binaries, user databases, and system databases. In that case, the SRM protection group consists of multiple consistency groups.

If the VNX model does not support the number of devices being protected within a protection group, create multiple VNX consistency groups for each protection group.

Note: The maximum number of consistency groups allowed per storage system is 64. Both MirrorView/S and MirrorView/A count toward the total.

The *VNX Open Systems Configuration Guide*, available on EMC Online Support, provides the most up-to-date synchronous and asynchronous mirror limits.

SRM recovery plan

The SRM recovery plan is a list of steps required to switch the operation of the datacenter from the protected site to the recovery site. The purpose of a recovery plan is to establish a reliable failover process that includes prioritized application recovery. For example, if a database management server needs to be powered on before an application server, the recovery plan starts the database management server, and then starts the application server. After the priorities are established, test the recovery plan to ensure the order of activities is correctly aligned to continue running the business at the recovery site.

Recovery plans are created at the recovery site, and are associated with one or more protection groups created at the protected site. Multiple recovery plans for a protection group are defined to handle applications and virtual machines with differing recovery priorities.

The options for recovery plan management are:

- ◆ **Test** — Tests the failover of the storage and virtual machine environment using temporary snapshot-based storage devices.
- ◆ **Cleanup** — Reverts the protected and recovery environments back to their pretest states. It also removes the temporary storage created to support the virtual machines at the recovery site.
- ◆ **Recovery** — Provides two options: migration and disaster. The migration option shuts down virtual machines from the protected site and synchronizes the storage between the two VNX systems to perform a graceful migration of virtual machines from the protected site to the recovery site. The disaster option performs the same storage tasks but does not attempt to shut down the virtual machines at the protected site.
- ◆ **Reprotect** — Re-establishes protection of virtual machines after a planned migration. Protection is established at the failover site, and virtual machines are protected at a secondary site that includes the previous production site.

Test the SRM recovery plan at the recovery site

Test the SRM recovery plan to verify that it performs as expected. Figure 133 shows a sample recovery plan.

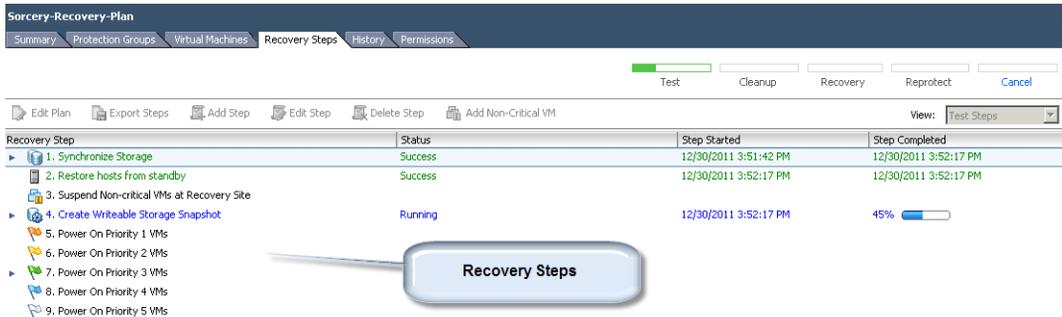


Figure 133 Recovery plan test

Click **Test** to test the recovery plan. During the test, the following events occur:

- ◆ Production virtual machines are shut down.
- ◆ SnapView sessions are created and activated using the existing snapshots.
- ◆ All the resources created within the SRM protection group are re-created at the recovery site.
- ◆ Virtual machines power on in the order defined in the recovery plan.

In SRM release 4, after all tasks in the recovery plan are complete, SRM pauses until the results are verified. After the test results are verified, click **Continue** to revert the environment to its production state.

SRM release 5 provides the cleanup option to revert the recovery environment to the pretest configuration and remove temporary storage devices created as part of the test.

Figure 134 shows the cleanup of a sample recovery plan.

Recovery Step	Status	Step Started	Step Completed
1. Power Off Test VMs at Recovery Site	Success	1/3/2012 1:21:01 PM	1/3/2012 1:21:10 PM
▶ 1.1. Sorcery-t1	Success	1/3/2012 1:21:01 PM	1/3/2012 1:21:08 PM
▶ 1.2. Sorcery-4	Success	1/3/2012 1:21:01 PM	1/3/2012 1:21:10 PM
▶ 1.3. Sorcery	Success	1/3/2012 1:21:01 PM	1/3/2012 1:21:10 PM
▶ 1.4. Sorcery-t3	Success	1/3/2012 1:21:01 PM	1/3/2012 1:21:08 PM
2. Resume Non-critical VMs at Recovery Site			
3. Discard Test Data and Reset Storage	Running	1/3/2012 1:21:10 PM	50% <input type="checkbox"/>
3.1. Protection Group Sorcery	Running	1/3/2012 1:21:10 PM	50% <input type="checkbox"/>

Figure 134 Recovery plan cleanup

The *VMware vCenter SRM Administration Guide*, available on EMC Online Support and on the VMware website, provides more information on SRM recovery plans and protection groups.

Execute an SRM recovery plan at the recovery site

The execution of an SRM recovery plan is similar to testing the environment with the following differences:

- ◆ Execution of the SRM recovery plan is a one-time activity.
- ◆ SnapView snapshots are not involved when the SRM recovery plan runs.
- ◆ The MirrorView/RecoverPoint/Replicator secondary copies are promoted as the new primary production LUNs.
- ◆ Restoring to the production environment requires the execution of the reprotect feature of SRM 5. Reprotect in SRM 5, along with the test, cleanup, and failback features, provide capabilities beyond DR, such as data center load-balancing and migration support.

- ◆ In the absence of any of the failback options listed above, manual steps are required to restore the protected site after executing a recovery plan.

Note: Do not execute an SRM recovery plan unless it is part of a validation test or a disaster has been declared.

Figure 135 shows a completed recovery plan.

The screenshot shows the 'Sorcery-Recovery-Plan' interface. At the top, there are tabs for Summary, Protection Groups, Virtual Machines, Recovery Steps, History, and Permissions. Below the tabs is a 'Recovery Complete' message box with a yellow background, containing instructions on how to reprotect or failback. Below the message box is a toolbar with buttons for Edit Plan, Export Steps, Add Step, Edit Step, Delete Step, and Add Non-Critical VM. Below the toolbar is a table showing the recovery steps.

Recovery Step	Status	Step Started	Step Completed
1. Pre-synchronize Storage	Success	1/3/2012 1:29:03 PM	1/3/2012 1:29:41 PM
1.1. Protection Group Sorcery	Success	1/3/2012 1:29:03 PM	1/3/2012 1:29:41 PM
2. Shutdown VMs at Protected Site	Already Done		
2.1. Shutdown Priority 5 VMs	Already Done		
2.2. Shutdown Priority 4 VMs	Already Done		
2.3. Shutdown Priority 3 VMs	Already Done		
2.4. Shutdown Priority 2 VMs	Already Done		
2.5. Shutdown Priority 1 VMs	Already Done		
3. Resume VMs Suspended by Previous Recovery	Success	1/3/2012 1:29:41 PM	1/3/2012 1:29:41 PM
4. Restore hosts from standby	Success	1/3/2012 1:29:41 PM	1/3/2012 1:30:40 PM
5. Prepare Protected Site VMs for Migration	Success	1/3/2012 1:29:41 PM	1/3/2012 1:30:40 PM
5.1. Protection Group Sorcery	Success	1/3/2012 1:29:41 PM	1/3/2012 1:30:40 PM
6. Synchronize Storage	Success	1/3/2012 1:30:40 PM	1/3/2012 1:31:19 PM
6.1. Protection Group Sorcery	Success	1/3/2012 1:30:40 PM	1/3/2012 1:31:19 PM
7. Suspend Non-critical VMs at Recovery Site	Success	1/3/2012 1:31:19 PM	1/3/2012 1:32:10 PM
8. Change Recovery Site Storage to Writeable	Success	1/3/2012 1:31:19 PM	1/3/2012 1:32:10 PM
8.1. Protection Group Sorcery	Success	1/3/2012 1:31:19 PM	1/3/2012 1:32:10 PM
9. Power On Priority 1 VMs	Success	1/3/2012 1:32:10 PM	1/3/2012 1:35:18 PM
10. Power On Priority 2 VMs	Success	1/3/2012 1:32:10 PM	1/3/2012 1:34:24 PM
11. Power On Priority 3 VMs	Success	1/3/2012 1:32:10 PM	1/3/2012 1:35:18 PM
11.1. Sorcery-t1	Success	1/3/2012 1:32:10 PM	1/3/2012 1:34:51 PM
11.2. Sorcery-t4	Success	1/3/2012 1:32:10 PM	1/3/2012 1:34:51 PM
11.3. Sorcery	Success	1/3/2012 1:32:10 PM	1/3/2012 1:34:51 PM
11.4. Sorcery-t3	Success	1/3/2012 1:32:10 PM	1/3/2012 1:34:24 PM
12. Power On Priority 4 VMs	Success	1/3/2012 1:32:10 PM	1/3/2012 1:34:24 PM
13. Power On Priority 5 VMs	Success	1/3/2012 1:32:10 PM	1/3/2012 1:34:24 PM

Figure 135 SRM recovery plan with EMC MirrorView

SRM failback scenarios

SRM failback is the process of restoring the protected VMware configuration after the protected environment storage infrastructure and vSphere environment are restored to a state that supports the application data.

SRM 5 provides an integrated reprotect feature that re-creates virtual machine and storage resource relationships between the site where the environment was recovered, and the previous protected site that supported the production environment after a failover.

Use the reprotect feature to establish a new relationship between the sites, with the two environments reversing roles. The recovery site becomes the protected site, and the protected site becomes the recovery site.

SRM reprotect works with all EMC storage replication adapters to re-establish or reverse the storage replication sessions between the two sites.

Reprotect provides the functionality to re-establish the protection relationships and storage configuration between the two environments such that the storage devices at recovery site are immediately protected after a failover occurs. After reprotect tasks are complete, SRM recovery plan tests are performed to validate the configuration prior to initiating a recovery to the production site, as shown in [Figure 136](#).

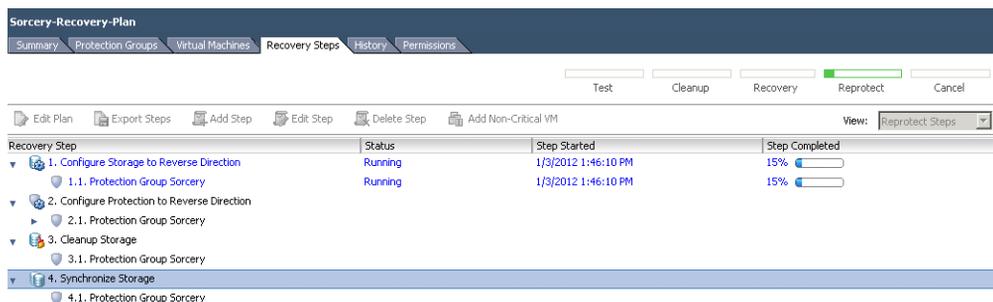


Figure 136 SRM reprotect

Recommendations and cautions for SRM with VNX

Observe the following recommendations and cautions:

- ◆ Install VMware tools on the virtual machines targeted for failover. If the tools are not installed, an error event is generated in the recovery plan when SRM attempts to shut down the virtual machine. Click the **History** tab to view any errors.
- ◆ Enable SnapView on the arrays with snapshots at both the primary and secondary sites to test failover and failback.

- ◆ Create alarms to announce the creation of new virtual machines on the datastore so that the new virtual machines are added to the mirrors in the SRM protection scheme.
- ◆ Complete the VNX-side configurations (MirrorView setup, snapshots creation, and so on) before installing SRM and SRA.
- ◆ Ensure that there is enough disk space configured for both the virtual machines and the swap file at the secondary site so that recovery plan tests run successfully.
- ◆ If SRM is used for failover, use SRM for simplified failback. Manual failback is a cumbersome process where each LUN is processed individually, including selecting the appropriate device signature option in vSphere on primary ESXi hosts. SRM automates these steps.
- ◆ Testing a recovery plan only captures snapshots of the MirrorView secondary image; it does not check for connectivity between the arrays or verify whether MirrorView works correctly. Use the SRM connection to verify the connectivity between the virtual machine consoles. Use SRM Array Manager or Unisphere to check the connectivity between arrays.

Summary

Table 21 lists the data replication solutions available for different types of VNX storage presented to an ESXi host.

Table 21 Data replication solutions

Type of virtual object	Replication
NAS datastore	<ul style="list-style-type: none"> • EMC Replicator • EMC Replication Manager • VMware vCenter SRM
VMFS/iSCSI	<ul style="list-style-type: none"> • EMC RecoverPoint • EMC MirrorView • EMC Replication Manager • VMware vCenter SRM
RDM/iSCSI (physical)	<ul style="list-style-type: none"> • EMC RecoverPoint • EMC MirrorView • VMware vCenter SRM
RDM/iSCSI (virtual)	<ul style="list-style-type: none"> • EMC RecoverPoint • EMC MirrorView • VMware vCenter SRM

This chapter includes the following topics:

- ◆ Introduction 266
- ◆ SAN Copy interoperability with VMware file systems..... 267
- ◆ SAN Copy interoperability with RDM virtual disks 268
- ◆ Using SAN Copy for data vaulting 269
- ◆ Importing Storage into the remote environment 276
- ◆ SAN Copy for data migration to VNX arrays..... 279
- ◆ Summary 283

Introduction

A core value of virtualization is the ability to move applications and data freely throughout the datacenter and networked environment. Data mobility enables you to move your data where it needs to be, when it needs to be there. An application server and its data can be encapsulated and transferred to another location in a relatively short period of time. This capability saves time and IT resources, provides additional measures of data protection, and enables improved collaboration.

The evolution of cloud computing has accelerated the trend toward data and application mobility, and established a need for periodic and cyclical migration processes to satisfy a variety of business purposes.

Regulatory compliance may require that multiple copies of data be retained in a protected facility for a specified period of time. The criticality of business information also imposes strict availability requirements. Few businesses can afford protracted downtime to identify and redistribute data to user groups. Data copy and migration is a core component of virtual datacenter management for tapeless backups, data vaulting, and many other use cases.

These examples highlight the need for technologies and practices to simplify data migration.

VMware provides Storage vMotion and Storage DRS to redistribute and migrate virtual machines between datastores. However, there is still no enterprise-level solution for a full-scale migration of datastores from one storage location to another with no impact to the production environment.

EMC offers technologies to migrate data between storage systems with minimal impact to the ESXi operating environment. This chapter discusses SAN Copy™ and its interoperability in vSphere environments with VNX block storage.

SAN Copy interoperability with VMware file systems

SAN Copy provides a VNX service to create copies of block storage devices on separate storage systems. SAN Copy propagates data from the production volume to a volume of equal or greater size on a remote storage array. SAN Copy provides the ability to:

- ◆ Create one-time LUN replicas on a separate system.
- ◆ Perform LUN migration as part of a system upgrade process.
- ◆ Perform periodic updates between storage systems for centralized data vaulting or archiving.

SAN Copy performs replication at the LUN level. It creates copies of LUNs that support VMFS datastores or RDM volumes.

Like other LUN cloning and replication technologies discussed in [Chapter 2, “Cloning Virtual Machines,”](#) the contents of the file system or the RDM volume are encapsulated within the replica LUN. The replica is presented to another host where the virtual machines and data can be imported into the environment.

Note: Avoid using SAN Copy with multiextent file systems. If a VMFS file system contains multiple extents, then all LUNs must be replicated to the target location and presented in the same device order.

To ensure application consistency, shut down the virtual machines that access the spanned VMware file system before you start the SAN Copy session. If the virtual machines cannot be shut down, use SnapView™ to create crash-consistent LUNs and use the SnapView LUN as the source for the SAN Copy session.

SAN Copy interoperability with RDM virtual disks

RDM volumes configured for physical compatibility mode provide direct VNX LUN access to the virtual machine. The virtual machine I/O bypasses the VMkernel and issues SCSI commands directly to the VNX LUN.

Since the guest operating system can issue SCSI commands to the storage array through an RDM LUN, the virtual machine uses application utilities and storage commands to prepare the LUNs before starting the SAN Copy session. When migrating data from an RDM volume, place applications in a hot standby mode or shut them down to ensure application consistency.

Using SAN Copy for data vaulting

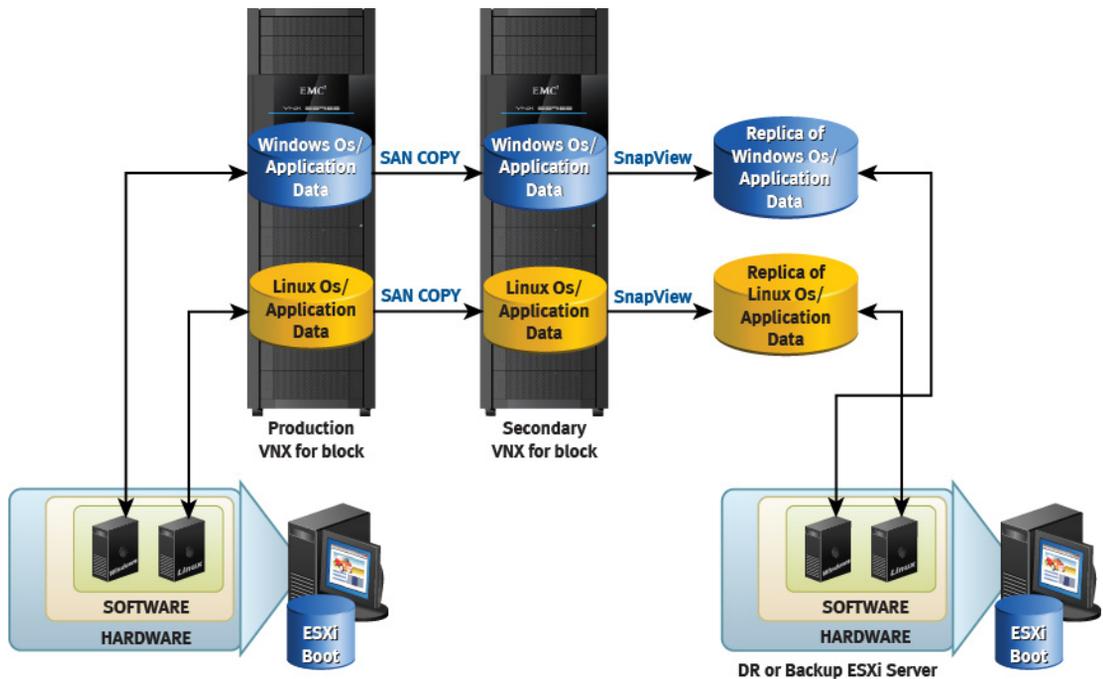
SAN Copy has two modes of operation:

- ◆ **Full mode** performs a complete re-silvering of the target device during each SAN Copy operation.
- ◆ **Incremental mode** performs periodic updates to an existing replica. It provides the foundation for data vaulting solutions. Offsite copies are periodically refreshed to maintain updated content from the production environments.

A schematic representation of the data vaulting solution is shown in [Figure 137 on page 270](#). Incremental SAN Copy uses SnapView technology to establish a consistent image of the production LUN state and to buffer data before it is copied to the target array. SnapView uses copy-on-write processing to maintain image versions.

Note: Consider the amount of I/O overhead when using Incremental SAN Copy in environments with high rates of data change.

A SnapView Clone LUN is used with SAN Copy to eliminate copy-on-write overhead. SnapView Clone establishes an independent replica to alleviate I/O to the production LUN. A clone refresh is required to update the SAN Copy replica LUN.



VNX-000659

Figure 137 Data vaulting with Incremental SAN Copy

Data vaulting of VMware file system using SAN Copy

Complete the following steps to migrate a LUN with SAN Copy. The core process applies to any VMFS or RDM LUN:

1. Identify all the devices to be copied.
2. Use Unisphere or the VSI Storage Viewer feature to identify the LUN that supports a VMFS datastore or RDM volume.

3. Select the SAN Copy target devices on the remote storage system. If multiple VNX systems are configured in a domain, storage devices on the remote storage system are visible in the **SAN Copy Wizard**, as shown in [Figure 138](#).

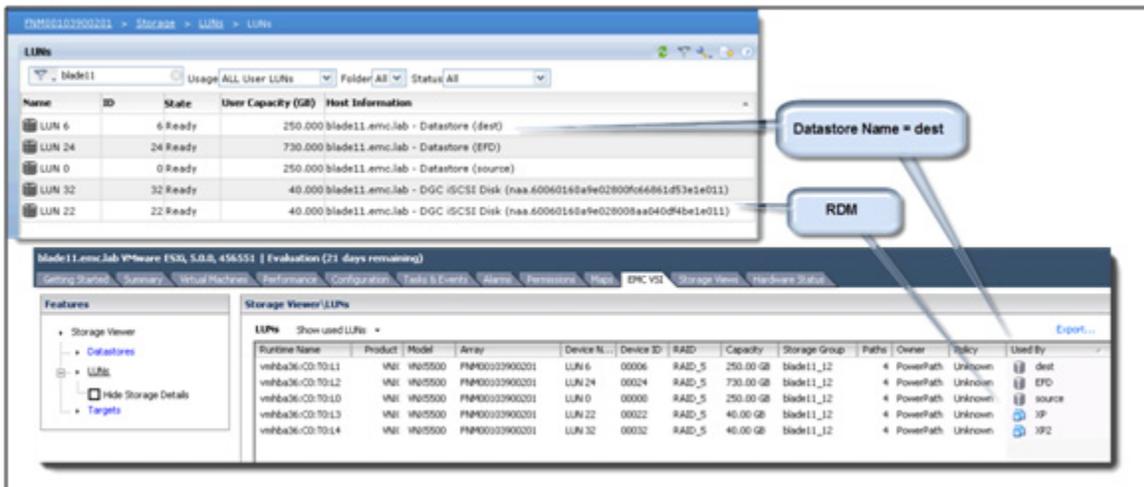


Figure 138 Using Unisphere or Storage Viewer to identify source LUNs

4. For non-VNX storage systems, identify the LUN number and the 128-bit WWN number that uniquely identify the SCSI devices. After you identify the source and destination LUNs, connect to the Unisphere SAN Copy configuration interface.

Note: There are multiple ways to determine the WWN. Use the management software for the storage array and Solutions Enabler to obtain the WWN of devices on supported storage arrays (Symmetrix, HDS, and HP StorageWorks).

Complete the following steps to initiate the migration session and create a data vaulting solution using SAN Copy:

1. In a SAN Copy configuration, VNX storage processor (SP) ports act as host initiators that connect the source VNX to SP ports on the remote VNX system. Create a storage switch zone including VNX SP WWNs from the source and target VNX systems.

2. VNX does not allow unrestricted access to storage. Create a storage group to mask the source VNX initiators with the VNX target LUNs. Use the storage array management utility to give the VNX SP ports access to the appropriate LUNs on the remote storage array.
3. Incremental SAN Copy sessions communicate with SnapView internally to keep track of updates for a SAN Copy session. Before you create an Incremental SAN Copy session, configure the SnapView-reserved LUN pool with the available LUNs. The size and quantity of the reserved LUNs depend on the number of accumulated changes to the source LUN between SAN Copy updates. If the rate of change is very high, or if the updates between the source and destination are infrequent (perhaps due to scheduling or bandwidth), increase the size of the reserved LUN pool.
4. Create an Incremental SAN Copy session between the source and destination LUNs as shown in [Figure 139 on page 273](#) and [Figure 140 on page 274](#).
5. Specify the attributes for the SAN Copy session:
 - SAN Copy session name
 - WWNs of the source and destination LUNs
 - Throttle value, latency, and bandwidth control value of the storage system interconnect.

Note: SAN Copy establishes a latency value by sending test I/O to the target. Do not alter the latency value.

Establishing a SAN Copy session does not trigger data movement. Initiating the session performs a series of validation tests to ensure that the VNX SP ports can access the remote devices, and that the capacity of each remote device is equal to or greater than the source devices.

- ◆ Activating the session establishes a point-in-time copy of the data from the source devices and propagates it to the target devices.
- ◆ SAN Copy provides a throttle parameter to control the rate at which data is copied between the source and target systems. A throttle value of 10 causes SAN Copy to use all available system resources to speed up the transfer. You can adjust the throttle value at any time after a session is created.

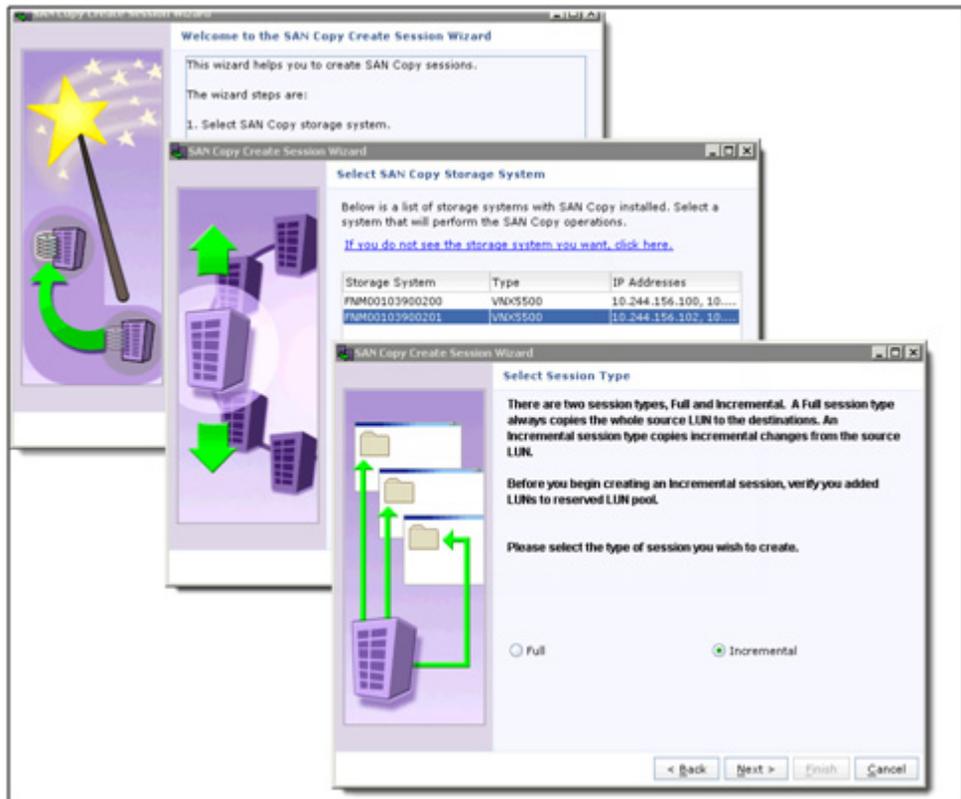


Figure 139 Creating an Incremental SAN Copy session

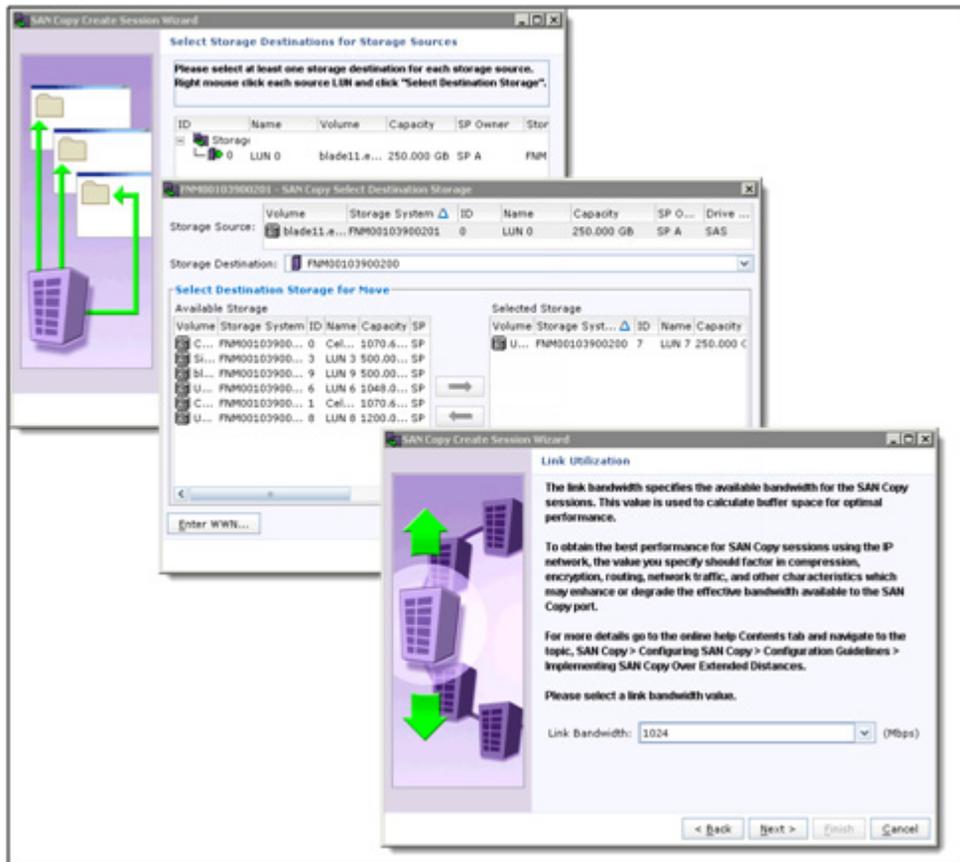


Figure 140 Creating an Incremental SAN Copy session (continued)

- After the copy process is complete, activate the LUNs at the remote site in Unisphere to make them available to the ESXi hosts.

Note: The target devices must remain inactive to continue to perform Incremental updates. Create SnapView LUN snapshots and present them to the ESXi host to allow the remote ESX environment to access the copies of the data.

7. Restart the existing SAN Copy session to perform an incremental update of the remote device. Incremental updates dramatically reduce the amount of data that must be propagated when the source volume has had very little change between updates.

Data vaulting of virtual machines configured with RDMs using SAN Copy

SAN Copy provides a storage array-based mechanism to create a consistent point-in-time copy of virtual disks stored on VNX LUNs. SAN Copy RDM LUN replication provides a more efficient method of virtual disk replication. With RDM volumes, SAN Copy replicates only the contents of the volume that have been modified by the guest, as opposed to multiple virtual disks contained within a VMFS volume.

Virtual machines configured with RDM volumes in physical compatibility mode are aware of the presence of VNX devices when Navisphere CLI/Agent is installed. The virtual machine has the ability to determine the devices to replicate with SAN Copy. Identify the devices that require protection and then configure SAN Copy to perform the replication of raw devices in the same manner as described in [“Data vaulting of VMware file system using SAN Copy”](#) on page 270.

Importing Storage into the remote environment

Configure remote sites for virtual machines using VMFS

Complete the following steps to create virtual machines at the remote site:

1. Enable ESXi host access to the remote LUN copy at the remote datacenter. Use a snapshot of the LUN instead of the actual device to preserve the Incremental SAN Copy capabilities.
2. Use unique virtual machine and datastore names to avoid name collisions. vCenter does not allow duplicate object names (like virtual machine names) within a vCenter datacenter.
3. Activate the LUN, assign it to the storage group of the ESXi cluster at the target site, and perform a host bus rescan to identify the new devices.
4. Use the vSphere Client to add the storage devices where the replicated VMware file system devices reside. Select **Keep existing signature** for each LUN. After all the replica storage has been added, the VMFS datastores appear in the **Host > Configuration > Storage** window of vCenter.
5. Browse the datastores to locate and register the virtual machines.

You can start the virtual machines at the remote site without modification if the following configuration requirements are met:

- ◆ The target ESXi hosts must use the same virtual switch configuration as the source ESXi hosts. For example, the virtual switch and virtual machine network names must be consistent with the source vCenter cluster.
- ◆ All VMware file systems used by the source virtual machines are replicated.
- ◆ The target ESXi host contains sufficient memory and processor resources to satisfy admission control in DRS cluster configurations.
- ◆ Devices such as CD-ROM and floppy drives are attached to physical hardware or disconnected from the virtual machines when they are powered on.

Configure remote sites for vSphere virtual machines with RDM

When a LUN is assigned to a virtual machine as an RDM device, a new virtual disk file is created within a VMware file system. This virtual disk file contains metadata that maps the virtual disk to the physical SCSI device. The file includes information such as the device ID, LUN number, RDM name, and the name of the VMware file system where the mapping is stored. If the datastore that holds the virtual machine configuration and the RDM file is replicated and presented to a different ESXi host, it is likely that the mapping file is not valid because it references an inaccessible device. Therefore, use a copy of the source virtual machine configuration file to reconstruct the virtual machine at the remote location. Use the .vmx file to register the virtual machine, and remap the virtual machine disk to the RDM replica in vCenter.

Complete the following steps to create a remote copy of a virtual machine with RDMs:

1. Create a folder in a datastore that resides on an ESXi host within the cluster at the remote site. This folder contains the virtual machine configuration files for the replicated virtual machine. Use a datastore that is not part of a replication session to avoid the possibility that the files may be overwritten.
2. Copy the configuration files of the source virtual machine to the directory created in step 1. Use a command line utility like scp, or use the vSphere Client Datastore Browser to complete this step.
3. From the remote vCenter environment, register the cloned virtual machine using the .vmx file copied in step 2.
4. Generate RDMs on the target ESXi hosts in the directory created in step 1. Configure the virtual machine RDM virtual disks to use the remote copy of the devices.

5. Power on the virtual machine at the remote site and verify that the devices are accessible within the guest OS.

Note: The procedure listed in this section assumes that the source virtual machine does not have a virtual disk on a VMware file system. The process to clone virtual machines with a mix of RDMs and virtual disks is complex, and beyond the scope of this document.

Start the virtual machines with the procedure described in [“Starting virtual machines at a remote site after a disaster”](#) on page 249.

SAN Copy for data migration to VNX arrays

VMware storage migration is largely accomplished by Storage vMotion, which offers an integrated solution to relocate virtual machines from an existing storage platform to a new system as part of a platform upgrade.

The value of Storage vMotion as a migration solution is that it preserves the virtual machine, datacenter, resource pool, and host configuration within the vCenter environment. Storage vMotion in vSphere 5 includes support for multiple vMotion interfaces, and offers the ability to perform simultaneous migrations between ESXi hosts. In most cases Storage vMotion provides the best approach to system migration. However, there may be occasions where a migration is limited by time and/or process. This is addressed by migrating the virtual machines at the datastore level. For example, large-scale LUN migrations benefit from SAN Copy because it reduces resource utilization of the host.

Storage vMotion does not preserve RDM volumes. When a virtual machine with RDM LUNs is migrated, the virtual disks are converted to VMFS as part of the process.

SAN Copy is frequently used to migrate LUNs to VNX. One of the major advantages of SAN Copy is that it offers Incremental SAN Copy to prepopulate and validate the target environment to limit service disruption during a cutover.

SAN Copy provides various modes of operation. In addition to the incremental copy mode, SAN Copy supports the full copy mode where data from a supported storage system is migrated to the VNX storage system. Complete the following steps to migrate VMware virtual infrastructure data from SAN Copy-supported storage arrays to an EMC VNX storage system:

1. Use the management interface of the source storage array to identify the WWNs of the source devices.
2. Identify the target LUN on the VNX system. The target LUN must be of the same or greater capacity as the source LUN.
3. Create a full SAN Copy session for the clone volume on the remote array. [Figure 141 on page 280](#) shows the necessary options to create a full SAN Copy session.

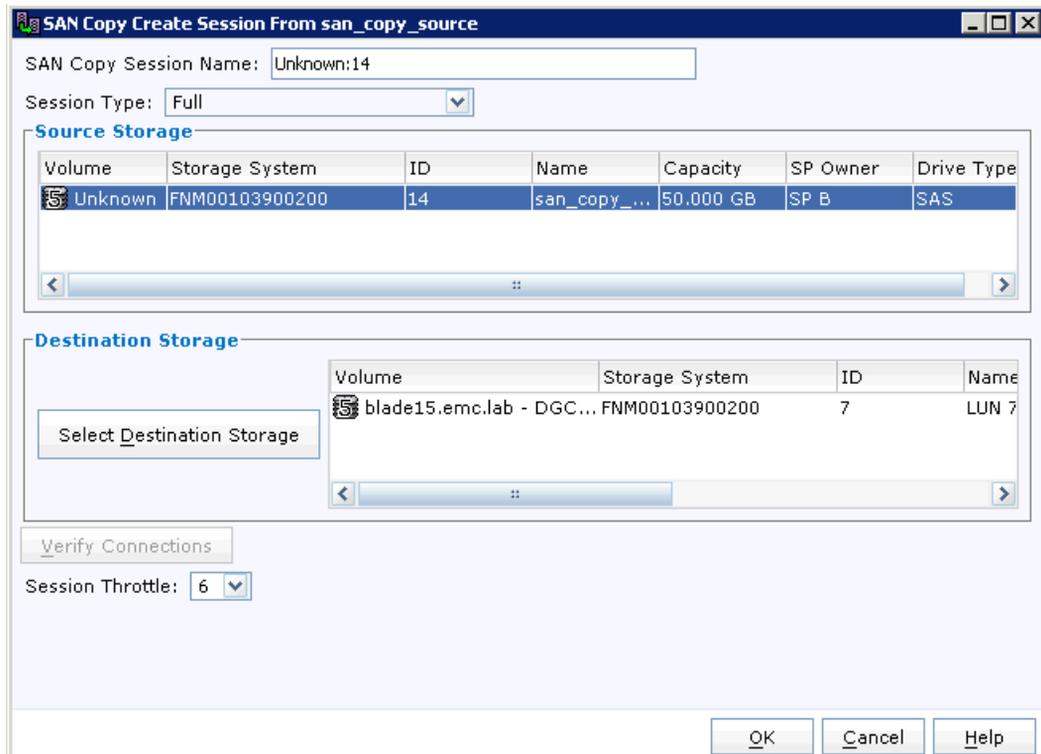


Figure 141 Creating a SAN Copy session to migrate data to a VNX

4. Shut down the virtual machines that use the devices being migrated to ensure application consistency.
5. Start the SAN Copy session to initiate the data migration from the source devices to the VNX devices.

6. Modify the VNX LUN masking to ensure that the ESXi hosts have access to the migrated devices. Update the zoning information to ensure that the ESXi hosts have access to the appropriate front end Fibre Channel ports on the VNX storage system.

Note: It is a good practice to maintain the source environment until the target environment has been thoroughly validated. A convenient way to do that is to remove the ESXi hosts from the storage group, while maintaining the LUN mapping. With this approach, the previous configuration can be quickly restored by adding the hosts back to the storage group if a problem is encountered.

7. After the full SAN Copy session completes, perform an ESXi host bus rescan to discover the VNX devices. The ESXi hosts recognize the VMFS volumes and populate them into the ESXi hosts that are visible from the **Storage** tab in the vSphere Client.
8. Using the vSphere Client Datastore Browser, identify each virtual machine within the migrated LUNs.
9. Register each virtual machine and power it on to ensure that the virtual machine boots correctly, and any applications running on the virtual machine function the same way as they did on the previous storage system.

SAN Copy provides a convenient mechanism to leverage storage array capabilities to accelerate the migration when there is a significant amount of content to migrate. SAN Copy can significantly reduce the downtime due to the migration of data to VNX arrays.

Migrate devices used as RDM

The procedure described in [“Configure remote sites for vSphere virtual machines with RDM” on page 277](#) also applies to this scenario.

RDM volumes contain unique device information that cannot be transferred. When an RDM virtual disk is replicated to a new LUN, the virtual disk configuration is invalidated because the RDM mapping file points to a device UUID that no longer exists for that virtual machine.

Modification of the virtual machine virtual disk configuration impacts applications that rely on the existing device path. RDM replication can be accomplished easily through the vSphere Client if the source and destination device IDs are correctly mapped.

When the data for virtual machines containing RDM volumes is migrated to another VNX, the disk configuration for the virtual machine must be modified to address the RDM replica LUN. Failure to correct the device mapping results in a virtual machine that will not boot correctly. Complete the following steps to ensure this does not occur:

1. Remove the existing RDM LUN from the virtual machine.
2. Disassociate the ESXi host with the LUNs being used as RDM volumes.
3. Re-create the RDM device mapping by using the canonical name of the replica device. Present the device with the same ALU/HLU sequence, and add the device with the same disk ID inside the Guest virtual machine.
4. Rescan the ESXi hosts and establish the correct device mapping by using the vSphere Client to associate the virtual machine with the appropriate migrated LUN.
5. Power on the virtual machines and confirm that the OS and applications function correctly.

Summary

This chapter describes how to use SAN Copy as a data migration tool for vSphere. SAN Copy provides an interface between storage systems for one-time migrations or periodic updates between storage systems.

One of the unique capabilities of SAN Copy is that it is compatible with different storage system types. Therefore, it is a useful tool to migrate data during storage system upgrades, and is a valuable tool to migrate from existing storage platforms to a VNX platform.

The *Migrating Data from an EMC CLARiiON Array to a VNX Platform using SAN Copy* white paper, on EMC Online Support, provides more information about data migration with SAN Copy.

