



Using EMC[®] VNX[®] Storage with VMware vSphere

Version 4.0

TechBook

P/N H8229
REV 05

EMC²

Copyright © 2015 EMC Corporation. All rights reserved. Published in the USA.

Published January 2015

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.

EMC², EMC, and the EMC logo are registered trademarks or trademarks of EMC Corporation in the United States and other countries. All other trademarks used herein are the property of their respective owners.

For the most up-to-date regulatory document for your product line, go to EMC Online Support (<https://support.emc.com>).

CONTENTS

Preface

Chapter 1

Configuring VMware vSphere on VNX Storage

Technology overview	18
EMC VNX family	18
FLASH 1st	18
MCx multicore optimization	19
VNX performance	20
VNX compounded efficiencies	20
VNX protection	20
File system management	21
VNX availability	21
VNX family software	22
Management tools	23
EMC Unisphere	23
EMC VSI for VMware vSphere	25
VMware vStorage APIs for Array Integration	29
VMware vStorage APIs for Storage Awareness	29
Installing VMware vSphere on VNX	29
Installing the ESXi hypervisor	31
Installing vSphere	31
Configuring vSphere with VNX	35
Host connectivity	35
Configuring the Network	41
iSCSI connectivity options for ESXi and VNX	43
Provisioning VNX storage for vSphere	51
Creating an NFS datastore using VSI	51
Provisioning block storage for VMFS datastores and RDM volumes	54
Unified storage considerations	56
Datastore virtual machine density	57
Best practices for extending a datastore	57
Solid state volumes for VNX OE for File	58
General recommendations for storage sizing and configuration	58
Storage multipathing	59
vSphere storage considerations	70
Dead space reclamation (Unmap)	70
VMFS-5	72
vStorage API for Storage Awareness	80
Network considerations	89
Network I/O Control	89
Virtual machine considerations	91
Virtual machine disk partition alignment	91
Virtual machines resiliency over NFS	98
Monitoring and managing storage	99
Monitoring datastores using vCenter	99
Configuring VNX file-system storage usage notification	103
Thinly provisioned storage	108
LUN compression	115
File deduplication and compression	116

	VNX storage options.....	117
	VNX supported disk types	118
	Disk grouping.....	118
Chapter 2	Cloning Virtual Machines	
	Introduction	128
	Using EMC VNX cloning technologies	128
	Replicating virtual machines with VNX SnapView	129
	Replicating virtual machines on VMFS datastores with SnapView clones	129
	Replicating virtual machines on VMFS datastores with SnapView Snapshot	131
	ESXi volume signatures	132
	Replicating virtual machines with SnapView clones of RDM LUNs	134
	Cloning virtual machines on VNX NFS datastores with VNX SnapSure ..	135
	Cloning virtual machines with native vCenter cloning and VAAI	136
	Cloning individual virtual machines on NFS datastores.....	137
	Summary	138
Chapter 3	Backup and Recovery Options	
	Introduction	142
	Virtual machine data consistency	142
	VNX native backup and recovery options	143
	File system logical backup and restore using VNX SnapSure.....	144
	Physical backup and restore using VNX File Replicator	145
	Snapshot backup and recovery of a VMFS datastore	145
	Backup and recovery of RDM volumes	148
	AppSync	148
	VSI AppSync Management	151
	Replication Manager	154
	Backup and recovery of a VMFS with VNX Advanced Snaps.....	157
	vStorage APIs for Data Protection	164
	Backup and recovery using VMware Data Protection	165
	Local data protection with vSphere Data Protection.....	167
	Installing VDP.....	168
	Creating a backup job	170
	Local Data Protection with vSphere Data Protection Advanced	172
	Configuring replication sessions	173
	VDP Emergency Restore.....	182
	Backup and recovery using Avamar	183
	Architectural view of the Avamar environment	184
	Backing up data using Avamar	184
	Recovering data using Avamar	186
	Backup and recovery using NetWorker.....	189
	VNX storage devices for NetWorker.....	190
	Using NetWorker to backup and restore VNX NAS file system NDMP	192
	Summary	193
Chapter 4	Using VMware vSphere in Data Restart Solutions	
	Introduction	196
	Definitions and Considerations	196
	Design considerations for DR and data restart	197

Testing the solution.....	197
Geographically distributed vSphere environments	197
EMC remote replication technology overview.....	198
EMC Replicator.....	199
EMC MirrorView.....	202
Failover MirrorView LUNs to a remote site using CLI.....	208
EMC RecoverPoint	209
RDM volume replication	212
Configuring remote sites for vSphere virtual machines with RDM.....	213
Starting virtual machines at a remote site after a disaster.....	214
Configuring remote sites for virtual machines using VMFS	214
EMC Replication Manager.....	215
Automating site failover with SRM and VNX	217
SRM testing.....	217
EMC Storage Replication Adapter	218
SRM protection groups at the protected site.....	219
SRM recovery plan.....	220
Summary	225

Chapter 5

Data Vaulting and Migration

Introduction	228
Using SAN Copy with VMware file systems.....	228
Using SAN Copy with RDM virtual disks	229
Using SAN Copy for data vaulting	229
Using SAN Copy for data vaulting of VMware file systems.....	230
Using SAN Copy for data vaulting of virtual machines configured with RDMs	233
Importing Storage into the remote environment	234
Configuring remote sites for virtual machines using VMFS	234
Configuring remote sites for vSphere virtual machines with RDM.....	234
Using SAN Copy to migrate data to VNX arrays.....	235
Migrating devices used as RDM.....	237
Summary	237

	Title	Page
1	EMC Unisphere interface	24
2	Viewing LUN properties	25
3	VSI Feature Manager	26
4	Unified Storage Access Control workflow.....	27
5	Storage Viewer NFS datastore details	28
6	Storage Viewer VNX block storage details.....	28
7	Configuration workflow	30
8	Unisphere LUN assignment for ESXi boot device.....	33
9	iBFT interface for VNX target configuration.....	34
10	VNX iSCSI port management interface	34
11	VNX storage with VMware vSphere	36
12	ESXi topology with FC/FCoE/iSCSI/NFS connectivity to VNX.....	38
13	VNX configuration of host initiator.....	40
14	VMkernel port configuration	42
15	VNX multiple subnet iSCSI target port addressing	43
16	iSCSI Software initiator configuration in vSphere 5.....	46
17	Topology diagram for multiple-subnet iSCSI configuration.....	47
18	Recommended configuration for VNX iSCSI targets.....	49
19	Disable Delayed Acknowledgement setting on storage adapter.....	50
20	File storage provisioning with Unified Storage Management.....	52
21	Creating a new NFS datastore with Unified Storage Management	53
22	File storage provisioning with Unified Storage Management.....	54
23	Creating a new VMFS datastore with Unified Storage Management.....	55
24	LUN ownership.....	60
25	LUN trespass.....	61
26	Stripe locking service	62
27	VMkernel pluggable storage architecture	62
28	Esxcli command output.....	63
29	VSI Path Management feature	65
30	Storage Viewer LUNs view	65
31	Elements of a multipathing configuration for NFS	67
32	Unisphere interface.....	67
33	Data Mover link aggregation for NFS server.....	68
34	vSphere networking configuration	69
35	VMkernel Properties window	69
36	Virtual machine configured on a Thick LUN.....	71
37	Virtual machine migrated to a thin LUN	72
38	Create File System.....	75
39	Plug-in installation	75
40	NFS Hardware Accelerate datastore property	76
41	Vmkfstools disk utilization option	76
42	SDRS datastore cluster.....	77
43	SDRS advanced policy configuration	78
44	SDRS I/O metric enablement setting	79
45	VASA datastore storage capability of VNX Flash drive LUN.....	80
46	Storage profile assignment	82
47	Compatibility/incompatibility with SAS Fibre storage profile	83
48	Enabling virtual machine storage profiles.....	84
49	Associating datastores with a user-defined storage profile.....	85
50	Associating the virtual machine with a user defined storage capability.....	85
51	VASA configuration	86
52	Virtual disk shares configuration.....	87
53	NFS SIOC congestion window	88

	Title	Page
54	Network Resource Allocation interface	89
55	vSphere 5 Datastore removal wizard	90
56	Select the disk	92
57	Guest disk alignment validation	93
58	NTFS data partition alignment (wmic command).....	93
59	Output of 1 MB aligned Linux partition	94
60	Output for an unaligned Linux partition (starting sector 63).....	94
61	Host Cache configuration on VNX EFD storage	95
62	Enable NPIV for a virtual machine after adding an RDM volume	97
63	Manually register virtual machine (virtual WWN) initiator records	97
64	Data Alarm Settings-Actions window	100
65	Storage Viewer\Datastores window-VMFS datastore.....	101
66	Adjustable percent full threshold for the storage pool	102
67	Create Storage Usage Notification window	103
68	User-defined storage usage notifications	104
69	User-defined storage projection notifications	105
70	VNX Monitoring and Reporting: Capacity Planning Report.....	105
71	VNX Monitoring and Reporting - Performance report	106
72	vCenter Operations Manager Dashboard	107
73	vCenter Operations Manager: VNX Storage Analytics	108
74	Thick or ZeroedThick virtual disk allocation	111
75	Thin virtual disk allocation	111
76	Virtual machine disk creation wizard	112
77	Virtual machine out-of-space error message.....	113
78	File system High Water Mark in the Unified Storage Management feature of VSI	114
79	Provisioning policy for an NFS virtual machine virtual disk.....	115
80	LUN compression property configuration.....	115
81	VNX FAST VP reporting and management interface	122
82	Disk Provisioning Wizard.....	125
83	Unisphere clone LUN management.....	130
84	Performing a consistent clone fracture operation.....	131
85	Creating a SnapView session to create a copy of a VMware file system	132
86	Device signature assignment	133
87	Selecting virtual machine configuration files in the datastore browser.....	134
88	Adding the new virtual machine to the ESXi host inventory.....	135
89	Creating a writeable NAS datastore checkpoint	136
90	Cloned NFS datastore in vSphere	138
91	Viewing ChildFsRoot parameter properties in Unisphere.....	144
92	Snapshot Configuration Wizard	146
93	Snapshot Configuration Wizard (continued)	147
94	Adding a VNX Storage System to AppSync	149
95	Editing Service Plan settings	150
96	Subscribing to a Service Plan	150
97	Overview of protected and unprotected datastores	150
98	VSI AppSync registration	151
99	Protect datastore.....	152
100	Using VSI AppSync Feature to Configure datastore protection.....	152
101	Subscribe Datastores to a AppSync Service Plan	153
102	Resignaturing a copy of an existing datastore.....	153
103	Using vCenter to manually copy a virtual machine	154
104	Replication Manager Job Wizard	155
105	Viewing Replica Properties in Replication Manager.....	156
106	Restoring a virtual machine using Replication Manager.....	156

	Title	Page
107	Viewing a read-only copy of the datastore in vSphere Client	157
108	Advanced Snapshot Basic Configuration	158
109	Viewing Snapshot Mount Points	159
110	Using the Snapshot Mount Point Configuration Wizard	160
111	Creating a snapshot consistency group	161
112	Creating a consistency group snapshot	162
113	Attaching a consistency group snapshot	163
114	VADP flow diagram	164
115	VMware Data Recovery	165
116	vSphere Data Protection	166
117	Allocation of virtual disks on one or more datastores	169
118	vSphere Getting Started menu	169
119	Backup with selected virtual disks	171
120	VDP backup schedule	171
121	Retention policy form	171
122	License Key window	172
123	In-guest agents	173
124	Creating a new replication job	174
125	Replication backup schedule	174
126	Avamar replication target	175
127	Scheduling the replication session	175
128	Retention policy for the Avamar target	175
129	Restore of a powered on virtual machine	176
130	Restoring a virtual machine	177
131	Restoring a backup with VDP	177
132	Setting Restore Options	177
133	Selecting the restore destination	178
134	Set Restore Options dialog box	178
135	Selecting an individual disk	179
136	Setting the Restore Options for restoring an individual disk	179
137	Selecting the restore destination	180
138	A virtual disk after a restore	180
139	Login options for the vSphere Data Protection Restore Client	181
140	Selecting individual disks from which to restore files	181
141	Restoring selected files or directories	182
142	Monitor Restores menu	182
143	Performing a VDP Emergency Restore of a virtual machine	183
144	Performing an emergency restore to an individual ESXi host	183
145	Sample Avamar environment	184
146	Sample proxy configuration	185
147	Avamar backup management configuration options	186
148	Avamar virtual machine image restore	187
149	Avamar browse tree	188
150	NetWorker-virtualization topology view	190
151	Viewing a VADP snapshot	190
152	NetWorker configuration settings for VADP	191
153	NDMP recovery using NetWorker	192
154	Backup with integrated checkpoint	193
155	Data replication wizard	200
156	Replication wizard (continued)	201
157	Preserving dependent-write consistency with MirrorView consistency groups	203
158	EMC Unisphere interface	204
159	Enable MirrorView between VNX systems	205

	Title	Page
160	MirrorView Wizard - select source LUNs	206
161	MirrorView Wizard - select remote storage.....	207
162	Promote mirrored LUN	207
163	Business continuity solution using MirrorView/S in a virtual infrastructure with VMFS	208
164	Synchronize MirrorView LUNs.....	209
165	RecoverPoint architecture overview	210
166	Disabling VAAI support on an ESXi host.....	211
167	RM protection for NFS datastores and virtual machines	216
168	Using the vSphere client to register a virtual machine with ESXi	217
169	SRM recovery plan summary	218
170	VMware vCenter SRM configuration	219
171	Creating an SRM protection group	220
172	Testing the recovery plan	221
173	Recovery plan Cleanup	222
174	SRM recovery plan with EMC MirrorView	223
175	SRM reprotect	224
176	Data vaulting with Incremental SAN Copy	230
177	Using Unisphere or Storage Viewer to identify source LUNs	231
178	Creating an Incremental SAN Copy session.....	232
179	Creating an Incremental SAN Copy session (continued).....	233
180	Creating a SAN Copy session to migrate data to a VNX.....	236

TABLES

	Title	Page
1	VNX family software	22
2	Multiple Subnet Configuration IPs/IQNs	47
3	Recommended Native Multipathing Plug-in path selection	64
4	NFS VAAI features.....	74
5	Supported SDRS LUN configurations	80
6	VASA storage capability mapping to VNX LUNs.....	81
7	VNX OE for Block 5.32 storage capability mapping to VNX LUNs	82
8	SIOC congestion windows	88
9	VNX Connector metrics.....	107
10	Command line descriptions for vSphere 5.5,vSphere 5, and vSphere 4	109
11	Virtual machine disk allocation policies	109
12	Disk types supported by VNX.....	118
13	Pool capabilities	119
14	VNX RAID options	119
15	Thin LUNs versus Thick LUNs	124
16	VNX-based technologies for virtual machine cloning	139
17	VDP and VDPZ feature comparison	166
18	Capacities for VDP deployments.....	168
19	vSphere Data Protection tabs	170
20	Resources for VDPA-Deployments	173
21	Backup and recovery options	194
22	EMC replication options for VMware environments.....	198
23	VNX MirrorView limits.....	202
24	Minimum revision levels for VAAI support with VNX RecoverPoint splitter	211
25	EMC RecoverPoint feature support.....	212
26	VNX to virtual machine RDM	213
27	Data replication solutions	225

PREFACE

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

Contact your EMC technical support professional if a product does not function properly or does not function as described in this document.

Note: This document was accurate at publication time. Go to EMC Online Support (<https://support.emc.com>) to ensure that you are using the latest version of this document.

Purpose

This document describes how VMware vSphere works with the EMC VNX family of storage systems. Although this document focuses on VNX2 storage, much of the content also applies when using vSphere with EMC Celerra or EMC CLARiiON storage systems.

Audience

This document is intended for storage administrators, system administrators, and VMware administrators. Individuals involved in acquiring, managing, or operating EMC VNX storage arrays and host devices and readers with knowledge of the following topics will benefit from this document:

- ◆ EMC VNX family
- ◆ EMC Unisphere
- ◆ EMC Virtual Storage Integrator (VSI) for VMware vSphere
- ◆ VMware vSphere

Related documentation

The following EMC publications provide additional information:

- ◆ *EMC VSI for VMware vSphere Web Client: Product Guide*
- ◆ *EMC VSI for VMware vSphere Web Client: 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 Management Product Guide*
- ◆ *EMC VSI for VMware vSphere: Storage Management Release Notes*
- ◆ *EMC VSI for VMware vSphere: Path Management Product Guide*
- ◆ *EMC VSI for VMware vSphere: Path Management Release Notes*
- ◆ *EMC VSI for VMware vSphere: AppSync Management Product Guide*
- ◆ *EMC VSI for VMware vSphere: AppSync Management Release Notes*

- ◆ *EMC VSI for VMware vSphere: EMC RecoverPoint Management Product Guide*
- ◆ *EMC VSI for VMware vSphere: EMC RecoverPoint Management Release Notes*
- ◆ *EMC VSI for VMware vSphere: SRA Utilities Product Guide*
- ◆ *EMC VSI for VMware vSphere: SRA Utilities Release Notes*

The following VMware websites provide more information about VMware products:

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

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



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



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

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

Typographical conventions

EMC uses the following type style conventions in this document:

Bold	Use for names of interface elements, such as names of windows, dialog boxes, buttons, fields, tab names, key names, and menu paths (what the user specifically selects or clicks)
<i>Italic</i>	Use for full titles of publications referenced in text and for variables in body text.
Monospace	Use for: <ul style="list-style-type: none"> • System output, such as an error message or script • System code • Pathnames, file names, prompts, and syntax • Commands and options
<i>Monospace italic</i>	Use for variables.
Monospace bold	Use for user input.

[]	Square brackets enclose optional values
	Vertical bar indicates alternate selections — the bar means “or”
{ }	Braces enclose content that the user must specify, such as x or y or z
...	Ellipses indicate nonessential information omitted from the example

Where to get help

EMC support, product, and licensing information can be obtained as follows:

Product information — For documentation, release notes, software updates, or information about EMC products, go to EMC Online Support at:

<https://support.emc.com>

Technical support — Go to EMC Online Support and click Service Center. You will see several options for contacting EMC Technical Support. Note that to open a service request, you must have a valid support agreement. Contact your EMC sales representative for details about obtaining a valid support agreement or with questions about your account.

Your comments

Your suggestions will help us continue to improve the accuracy, organization, and overall quality of the user publications. Send your opinions of this document to:

techpubcomments@emc.com

CHAPTER 1

Configuring VMware vSphere on VNX Storage

This chapter presents the following topics:

◆ Technology overview	18
◆ Management tools	23
◆ Installing VMware vSphere on VNX	29
◆ Configuring vSphere with VNX	35
◆ Provisioning VNX storage for vSphere	51
◆ Unified storage considerations	56
◆ vSphere storage considerations	70
◆ Network considerations	89
◆ Virtual machine considerations	91
◆ Monitoring and managing storage	99
◆ VNX storage options.....	117

Technology overview

This section provides information about the technologies associated with the next generation of EMC® VNX® storage systems in the context of a VMware vSphere virtualized environment.

EMC VNX family

EMC VNX unified storage systems deliver platform efficiency and innovation with enterprise capabilities for file, block, and object storage in a scalable, easy-to-use, high-performance solution. The VNX family of storage systems is ideal for mixed workloads in virtual environments. It combines extremely powerful and flexible hardware with advanced multicore optimization, management, and data protection software to meet the most demanding needs of today's enterprises.

VNX is multiprotocol compatible (file, block, object), giving you the flexibility to deploy expandable, future-ready storage. The powerful controller architecture of VNX storage helps ensure that you won't run out of storage processor power for transactions or bandwidth during the service life of the storage system. Capacity- and performance-scaling capabilities enable on-demand capacity management in combination with built-in capacity efficiency features such as thin provisioning, block-level deduplication, EMC Fully Automated Storage Tiering (FAST™) technology, and compression.

FLASH 1st

The general availability of NAND flash storage media has fundamentally changed storage system design and IT best practices. Flash drive storage delivers the highest I/O performance with the lowest latency of any storage media. The need to store all active data on flash drives is driven by the multicore/multisocket designs of modern servers. Such servers typically run multiple virtualized server instances—anywhere from 10 to 100 virtual machines. In addition, a VNX system typically serves 10 to 100 or more servers. Therefore, a single VNX system must be able to meet the transactional demands of several thousand virtual machines. Flash is the only way to accomplish this without using up to a quarter million 15K disk drives. Flash memory is higher performing and delivers lower transaction costs.

EMC's FLASH 1st strategy takes advantage of FAST technology. As application workloads shift to new and highly active data, less active data is identified and automatically moved from flash to lower-cost, high-capacity disks. FAST technology ensures that adequate flash capacity is available for the active portion of all the workloads on the VNX system.

The FAST Suite consists of FAST-VP and FAST Cache. While FAST Cache works to handle any transitory exceptions, FAST-VP moves data to the appropriate tier based on persistent trending. FAST-VP moves slices between tiers over a broader sample time. FAST Cache caches any data that is not already on a Flash drive onto a separate set of flash drives. The FAST Suite enables the servicing of workloads requiring a large number of 15K drives with a mix of tiers and lower drive count.

Typically a small amount of flash memory can service a high percentage of the overall IOPS. The target is to serve at least 80 percent of the IOPS, with flash memory making up about 5 percent of total storage capacity. For overall value in flash performance, hybrid arrays are the best platform choice.

MCx multicore optimization

The advent of flash technology has completely changed the requirements of mid-range storage systems. EMC redesigned the mid-range storage platform to efficiently optimize multicore CPUs and provide the highest-performing storage system at the lowest cost. The EMC MCx™ multicore optimization architecture, which comprises Multicore Cache, Multicore FAST Cache, and Multicore RAID, provides the market-leading foundation of mid-range storage for the next decade.

MCx distributes all VNX data services across all cores. MCx dramatically improves file performance for transactional applications, such as databases, and virtual machines over network-attached storage (NAS).

The advances in the VNX platform ensure that cache management and back-end RAID management processes take full advantage of multicore CPUs. This allows cache and back-end processing software to scale in a linear fashion, enabling EMC to take full advantage of the power of the latest multicore CPUs. The VNX has a scaling factor of 97 percent, providing the ability to add cores as needed for future scaling.

Multicore Cache

The cache is the most valuable storage asset in the storage subsystem, and its efficient use is essential to the overall efficiency of a platform handling variable and changing workloads. With Multicore Cache, the cache engine is modularized to take advantage of all the cores available in the system, enabling the VNX storage to scale seamlessly.

In addition, Multicore Cache eliminates the need to manually reserve separate space for read and write cache, so no management overhead is required to ensure that the cache is working most effectively.

The VNX system uses dynamic and adaptive cache management through advanced caching algorithms to optimally handle varying sequential and random read and write workloads. Highly intelligent write flushing tracks arrival rates as well as the ability of the back end-disks or SSDs to write the data out of cache. It provides adaptive queueing to throttle write rates in place of forced flush situations. All cache pages are dynamically assigned depending on the needs of the system. New pages are allocated immediately, as needed, for reads, writes, thin metadata, and so forth, which ensures that the system resources are constantly being tuned to match the changing workloads.

Multicore RAID

The MCx architecture improves the handling of I/O to the permanent back-end storage (HDDs and SSDs). The significant performance improvements in the VNX come from the modularization of the back-end data management processing, which allows it to seamlessly scale across all processors.

In addition, with RAID processing that is highly mature, the VNX system delivers improved flexibility, ease of use, and reliability.

VNX performance

The VNX with the MCx architecture takes advantage of the FLASH 1st strategy and provides unprecedented overall performance, optimizing transactional performance, bandwidth performance, and capacity efficiency. Compared to the previous generation, the VNX provides:

- ◆ Up to 4 times more file transactions—best in the industry compared to dual controller arrays
- ◆ Improved file performance for transactional applications (for example, Exchange on VMware over NFS) by up to 3 times with 60 percent better response time
- ◆ Up to 4 times more Oracle and SQL OLTP transactions
- ◆ Support for up to 4 times more virtual machines—a greater than 3 times improvement
- ◆ Up to 3 times more bandwidth for Oracle and SQL data warehousing

VNX compounded efficiencies

Compounded efficiencies provide a way to easily save money by combining VNX out-of-band block-based deduplication and improved FAST Suite tiering technologies to dramatically lower the costs of the flash tier and reduce the cost per gigabyte. VNX provides compounded efficiencies through the following technologies:

- ◆ **FAST Suite improvements**—The VNX series increases FAST Cache capacities, delivers 4 times better FAST VP granularity (256 MB versus 1 GB), and supports new FAST VP SSDs based on enterprise multilevel cell (eMLC) technology to lower the cost per gigabyte.
- ◆ **Fixed block deduplication**—While only a few flash drives—typically between 1 percent and 5 percent of total capacity—are needed to improve performance and lower costs, you can further reduce capacity with deduplication, which is ideal for virtual machine environments.

VNX protection

VNX storage systems provide enhanced data protection through the following improved technologies:

- ◆ **Online file migrations**—The VNX Virtual Data Mover (VDM) technology has been enhanced to perform automated, highspeed file system migrations between systems.
- ◆ **Improved application uptime**—VNX active/active for block improves application uptime by allowing clients to access a classic LUN (non-pooled LUN) simultaneously through both storage processors for improved reliability, ease of management, and improved performance.
- ◆ **More affordable remote replication**—EMC RecoverPoint® 4.0 enhancements include synchronous replication over IP, failover to any point in time with VMware SRM, and a new virtual EMC RecoverPoint appliance to reduce acquisition cost.

File system management

File system management includes:

- ◆ Transactional NAS—Encompasses workloads such as server virtualization and databases, which have traditionally been directed toward block storage.
- ◆ Capacity NAS—Designed for large-capacity file data (“big data”) and delivers high aggregate performance. It is scaled out NAS and is characterized by providing a very large single namespace.
- ◆ Traditional NAS—Typically handles smaller file stores but still needs comprehensive features such as deduplication, snapshots, quotas, and WORM.

VNX availability

This section describes the asymmetric active/active—also known as asymmetrical logical unit access (ALUA)—and symmetric active/active access models.

Asymmetric active/active

In the first-generation VNX models, EMC implemented an asymmetric access model where a single path is optimized and all I/O is sent down this path unless an issue (failure/congestion) occurs, after which the I/O is sent via the secondary path to the backup storage processor (SP). In the ALUA model a single SP “owns” the host LUN and any access from the secondary path is serviced via the inter-SP links from the owning SP to the secondary SP. If the failed path to the owning SP is offline for a period of time, the LUN is trespassed and the secondary SP becomes the owning SP. Although this is a satisfactory access model, the host still uses a single storage controller to access the LUN unless the ESXi host multipath settings have been changed.

Symmetric active/active

With the VNX systems and the introduction of MCx, EMC offers the first phase of an active/active access model that allows a host’s multipathing software to access the LUN in a balanced manner across both SPs.

This design eliminates the concept of LUN ownership and redirection throughout the stack. It provides reduced failover times and reduces data outages caused by trespass storms, resulting in significant performance improvement. In the past, due to CPU bottlenecks, the systems could not take advantage of the back-end resources. With symmetric active/active, access to the remaining back-end resources is available through the peer SP. This model allows clients to access a classic LUN (non-pooled) through both SPs simultaneously for improved reliability, ease-of management, and improved performance.

VNX family software

Table 1 describes the contents of the VNX software components.

Table 1 VNX family software

Software	Contents
Unisphere® Management Suite (required)	Unisphere, Unisphere Remote, Unisphere Analyzer, Unisphere Quality of Service Manager (QoS), VNX Monitoring and Reporting (storage-only version)
Operating Environment (required)	All protocols, thin provisioning, compression, deduplication, SAN Copy, and ODX Enabler
EMC Storage Analytics for VNX	VNX version of VMware vCenter Operations Manager, EMC Adapter for VNX
FAST Suite	FAST CAche and FAST VP
Security and Compliance Suite	VNX Host Encryption, File Level Retention, Common Event Enabler (CEE) Anti-Virus Agent, and Event Publishing Agent
Local Protection Suite	SnapSure, SnapView, VNX Snapshots, EMC RecoverPoint SE CDP
Remote Protection Suite	Replicator, MirrorView A/S, RecoverPoint SE CRR
Application Protection Suite	AppSync, Replication Manager

The VNX software components are described as follows:

- ◆ **Unisphere Management Suite**—Extends Unisphere's easy-to-use interface to include VNX monitoring and reporting for validating performance and anticipating capacity requirements. The suite also includes Unisphere Remote for centrally managing up to thousands of VNX and VNXe systems.
- ◆ **Operating Environment**—Enables management of resources and services on the array and consists of VNX OE for File and VNX OE for Block.
- ◆ **EMC Storage Analytics for VNX**—Delivers a single, end-to-end view of virtualized infrastructures (servers to storage) powered by the VMware vCenter Operations Management Suite analytics engine. EMC Storage Analytics for VNX (ESA) delivers actionable performance analysis and proactively facilitates increased insight into storage resource pools to help detect capacity and performance issues so they can be corrected before they cause a major impact. ESA provides increased visibility, metrics, and a rich collection of storage analytics for VNX storage infrastructures in VMware virtual environments.
- ◆ **FAST Suite**—Enables deployment of a FLASH-1st strategy using a small number of flash drives. This strategy automatically tiers and serves data from the most cost-effective drive type based on the data need, which lowers storage costs and delivers higher performance levels for important applications.
- ◆ **Security and Compliance Suite**—Helps ensure that data is protected from unwanted changes, deletions, and malicious activity. You can encrypt data, maintain data confidentiality, and enforce data retention to meet compliance requirements.

- ◆ **VNX Local Protection Suite**—Includes snapshots for point-in-time recovery and continuous data protection with EMC RecoverPoint local replication. RecoverPoint has a new virtual appliance that can reduce costs by 60 percent.
- ◆ **VNX Remote Protection Suite**—Includes RecoverPoint Continuous Remote Replication with DVR-like recovery. RecoverPoint includes a new virtual RecoverPoint appliance to reduce acquisition cost by 60 percent and provide synchronous replication over IP and failover to any-point-in-time with VMware SRM.
- ◆ **VNX Application Protection Suite**—Makes application-consistent replicas. AppSync provides a simple, self-service, and SLA-driven approach for protecting virtualized Microsoft SQL/Exchange and VMware over file and block.

Management tools

Administrators can use the tools discussed in this section to view and manage VNX storage in a VMware vSphere environment.

EMC Unisphere

EMC Unisphere is the next-generation unified storage management platform that provides intuitive user interfaces for the newest range of unified platforms including the EMC VNX family. Unisphere continues to support existing EMC CLARiiON®, EMC Celerra®, and EMC RecoverPoint SE systems. The Unisphere approach to storage management features simplicity, flexibility, self-help, and automation—all key requirements for the journey to the cloud. Unisphere can be customized to the needs of a mid-size company, a department within large enterprises, or a smaller remote office or branch office environment. With its plugable architecture, Unisphere is easily extensible and continues its seamless support for additional EMC offerings, including integration with data protection and security.

The following Unisphere capabilities greatly simplify mid-tier and entry-level systems management:

- ◆ Single sign-on automatically discovers all VNX, CLARiiON, Celerra, and EMC RecoverPoint SE systems in the environment.
- ◆ Unisphere Remote provides centralized multibox monitoring for hundreds of VNXe and VNX systems deployed in remote and branch offices.
- ◆ The dashboard is a single screen for at-a-glance management, providing administrators with instant and actionable knowledge about the status of their virtualized and physical infrastructures.
- ◆ Unisphere enables administrators to drill down and troubleshoot the most critical issues from a single view.
- ◆ Ease of use helps administrators improve IT staff productivity by reducing the amount of time spent on critical storage administrative tasks.

Unisphere is an interface based on Java that allows you to quickly provision, manage, and monitor storage assets. Customizable dashboard views provide realtime details on the health of the environment, as illustrated in [Figure 1](#).

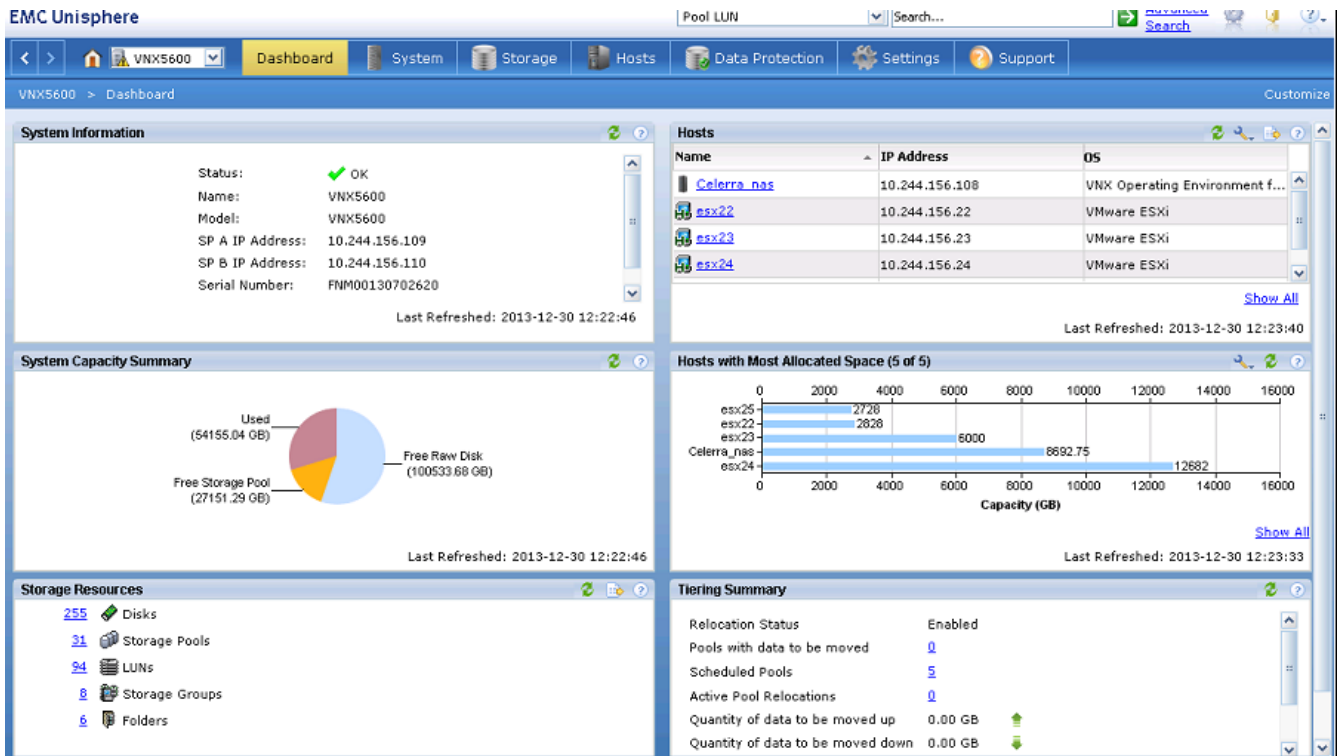


Figure 1 EMC Unisphere interface

Unisphere includes discovery capabilities for VMware environments. The SP displays details from vSphere about the virtual storage as it relates to the VNX storage resources. The Unisphere virtualization management interface is accessible from the Unisphere **Hosts** pane. The discovery agent requires vSphere credentials to establish a web services session with one or more hosts or vCenter servers. Unisphere allows administrators to understand how VNX storage is being utilized within the vSphere environment.

Figure 2 illustrates the properties of a sample LUN.

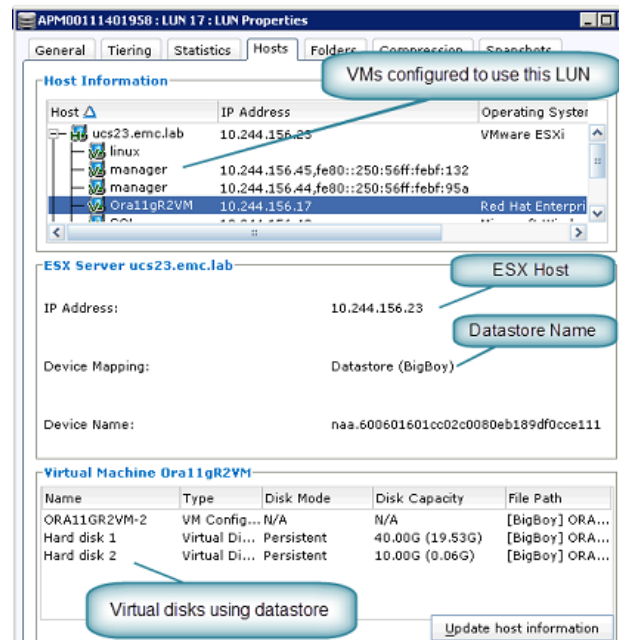


Figure 2 Viewing LUN properties

The interface shows that LUN 17 is assigned to ESX host ucs23.emc.lab and has been formatted as a VMFS datastore. It also shows that the datastore (named Big Boy) is being used to support virtual disks on several virtual machines including an Ora11gR2VM. Unisphere displays the two virtual disks and their capacities and attributes.

You can use information presented in this interface to monitor the environment and validate virtual disk placement when you configure storage system replication and data protection policies.

EMC VSI for VMware vSphere

Virtual Storage Integrator (VSI) is a no-charge EMC plug-in for VMware vCenter that is available to all VMware users who have EMC Symmetrix VMAX family, EMC VNX family, or VPLEX storage in their environment. VSI enables IT organizations to achieve simplicity and efficiency in data center operations.

VSI dramatically simplifies management of virtualized storage. It provides management integration of EMC storage platforms with vCenter, enabling server administrators to easily handle storage-related tasks. VMware administrators can gain visibility into EMC storage using the familiar vCenter interface. The familiar interface, along with the ability to self-provision EMC storage, configure FAST-VP tiering policies, clone virtual machines, and manage compression from within vCenter, makes managing virtual storage as easy as ever.

With VSI, IT administrators can do more in less time. VSI offers unmatched access control that enables you to efficiently manage and delegate storage tasks with confidence. You can perform daily management tasks with up to 90 percent fewer clicks and up to 10 times higher productivity. This increase in efficiency also applies to the management of virtual desktop environments. The EMC VSI plug-in enables you to easily deploy Citrix XenDesktop and VMware Horizon View virtual desktops from vCenter. To do this, create a

fast clone of a master image, and then register the fast clone image with virtual desktop clients. Using this process, you can deploy multiple XenDesktop or Horizon View virtual desktops from a single image in minutes.

VSI is a modular framework that allows management features to be added or removed in support of specific EMC products. This section describes the VSI Unified Storage Management, Storage Viewer, and Path Management features that are most applicable to VNX systems.

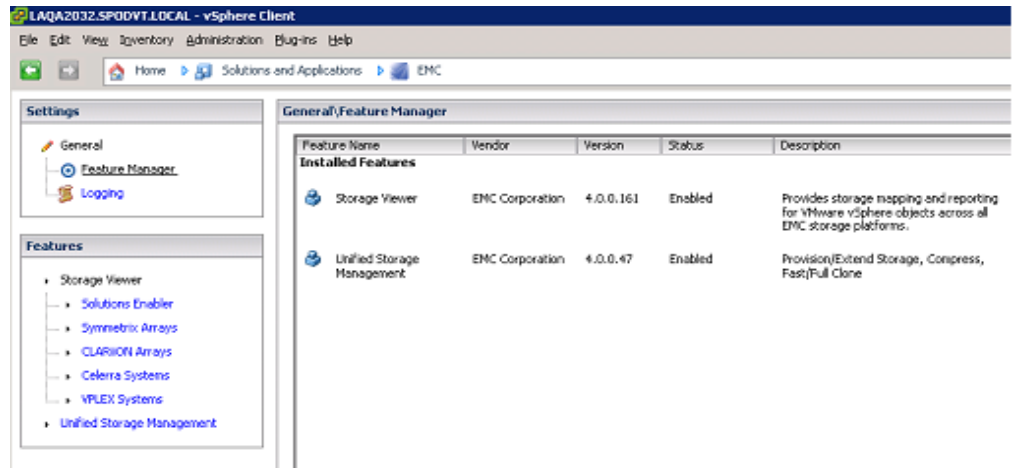


Figure 3 VSI Feature Manager

VSI: Unified Storage Management

The Unified Storage Management feature of VSI enables vSphere administrators to manage VNX storage using the vSphere Client.

With the Unified Storage Management feature you can:

- ◆ Create VMFS datastores and Raw Device Mapping (RDM) devices by provisioning a new storage device on an ESXi host or data center cluster
- ◆ Provision datastores in accordance with EMC best practices
- ◆ Create multiple LUNs and masking for RDMs or Virtual Machine File Systems (VMFS)
- ◆ Rapidly provision full virtual machine clones or space-efficient fast clones within NFS datastores
- ◆ Enable deduplication on selected datastores
- ◆ Compress virtual NFS disk files

EMC Unified Storage Access Control Utility

Unified Storage Management requires administrative credentials to access and manage the storage system. The EMC Unified Storage Access Control Utility (ACU) enables the storage administrator to assign VNX management access to authorized users. It operates under an explicit deny model with view (read) and management (modify) entitlements granted at the RAID group, storage pool, or NFS file-system level.

ACU credentials can be exported as an encrypted key. Storage administrators import the key into other systems running the vSphere Client. [Figure 4](#) illustrates the steps to create an access profile.

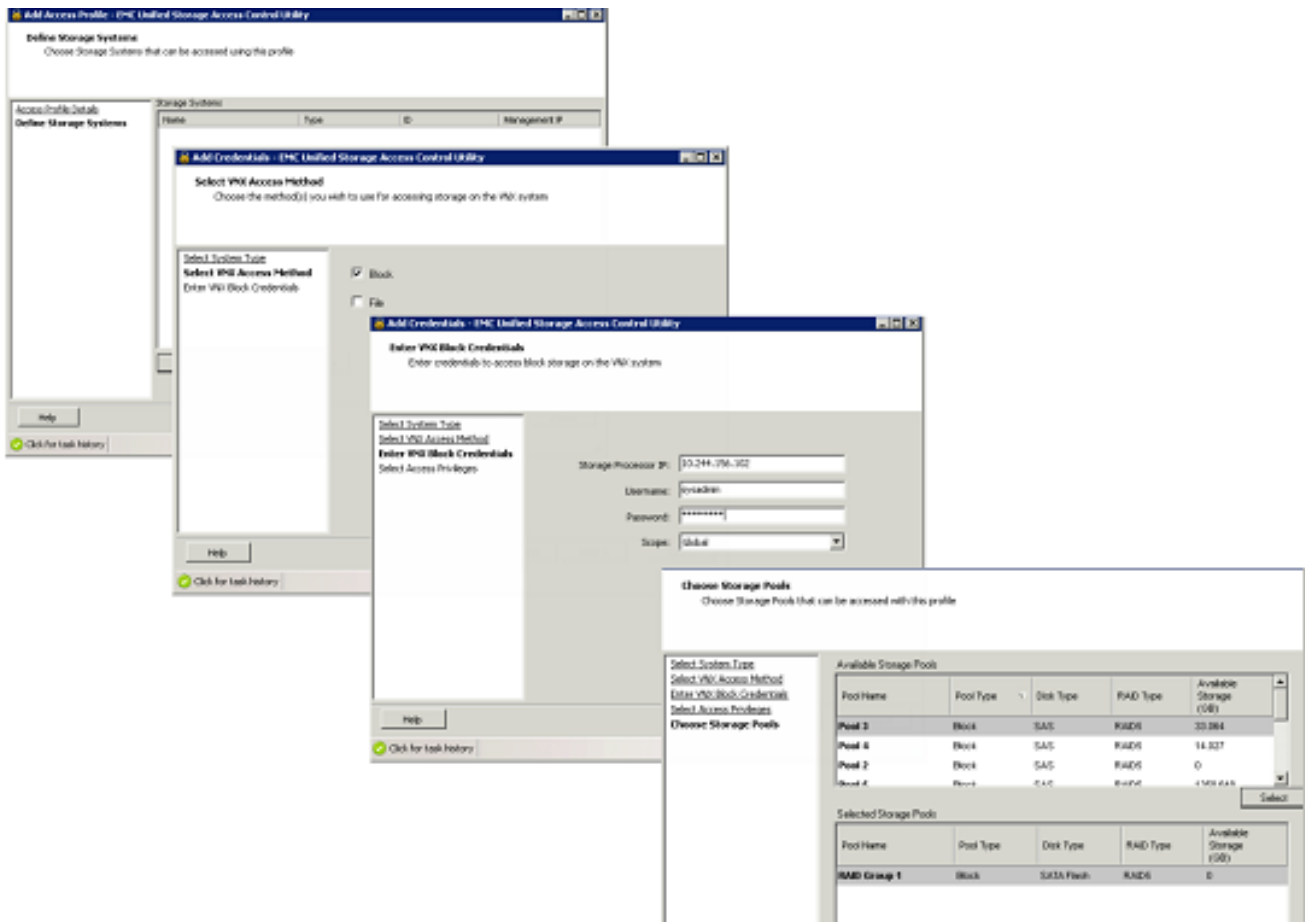


Figure 4 Unified Storage Access Control workflow

VSI: Storage Viewer

The VSI Storage Viewer feature displays VNX storage device details for LUNs, file systems, and data paths in the ESXi datastore. You can view VNX device information by selecting an ESXi storage object in Storage Viewer. This information can help to isolate a particular storage device when you are troubleshooting the environment. Storage Viewer provides the following functions and benefits:

- ◆ 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 such as the VNX System ID, file system, RAID type, storage pool, and Data Mover.



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.

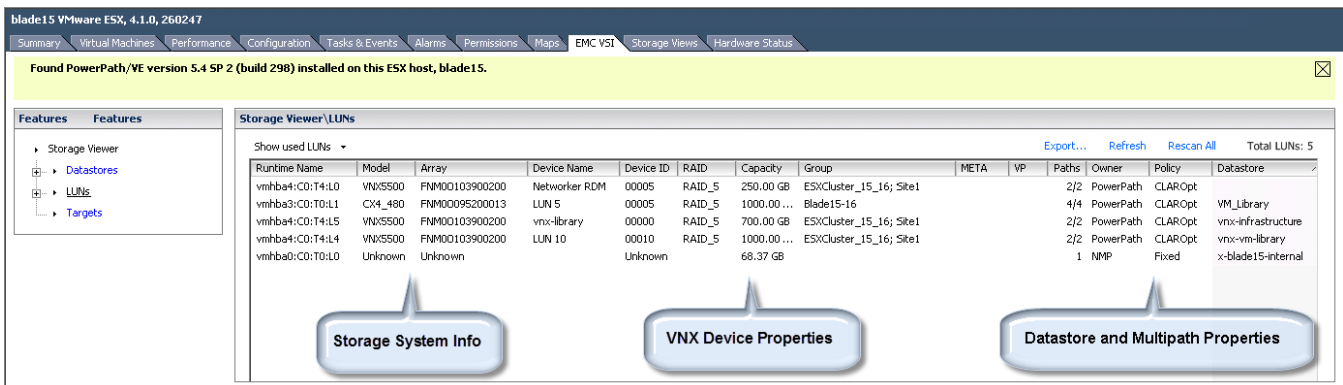


Figure 6 Storage Viewer VNX block storage details

VSI: Path Management

The VSI Path Management feature includes support for the VMware Native Multipathing Plug-in and for PowerPath/VE.

The Path Management feature enables you to change the multipath policy and manage multiple paths from within the VMware vSphere Client. You can change the multipath policy for several devices based on storage class and virtualization object.

For example, you can manage all Symmetrix devices for a given ESXi host or all CLARiion devices for a given VMware vSphere cluster. You can manage multipath policies for devices owned by both the Native Multipathing Plug-in and PowerPath/VE.

This feature is a valuable asset to administrators who need to maintain consistent multipath policies across a large virtual datacenter containing a large number or wide variety of storage devices.

VMware vStorage APIs for Array Integration

VMware vStorage APIs for Array Integration (VAAI) offloads VMware storage-related functions from the server to the storage system, enabling more efficient use of server and network resources for increased performance and consolidation. Letting the VNX series software perform common data management tasks, such as vMotion migration, results in greater network IOPS, support for more virtual machines, and faster response time. Other examples of offloaded tasks include:

- ◆ Thin Provisioning (block)
- ◆ Thin Provisioning Stun (block)
- ◆ Full Clone (file)
- ◆ Extended Statistics (file)
- ◆ Space Reservations (file)
- ◆ Hardware Accelerated Locking (block)
- ◆ Hardware Accelerated Zero (block)
- ◆ Hardware Accelerated Copy (block)

VMware vStorage APIs for Storage Awareness

VMware vStorage APIs for Storage Awareness (VASA) is a VMware API that enables the display of storage information through vCenter. Integration between VASA technology and VNX makes storage management in a virtualized environment a seamless experience. Administrators can use the familiar vSphere interface to view details of virtual and physical resources, provision storage, integrate replication, and offload storage functions to the storage system. VASA enables the VMware administrator to view basic storage components including arrays, storage processors, I/O ports, and LUNs.

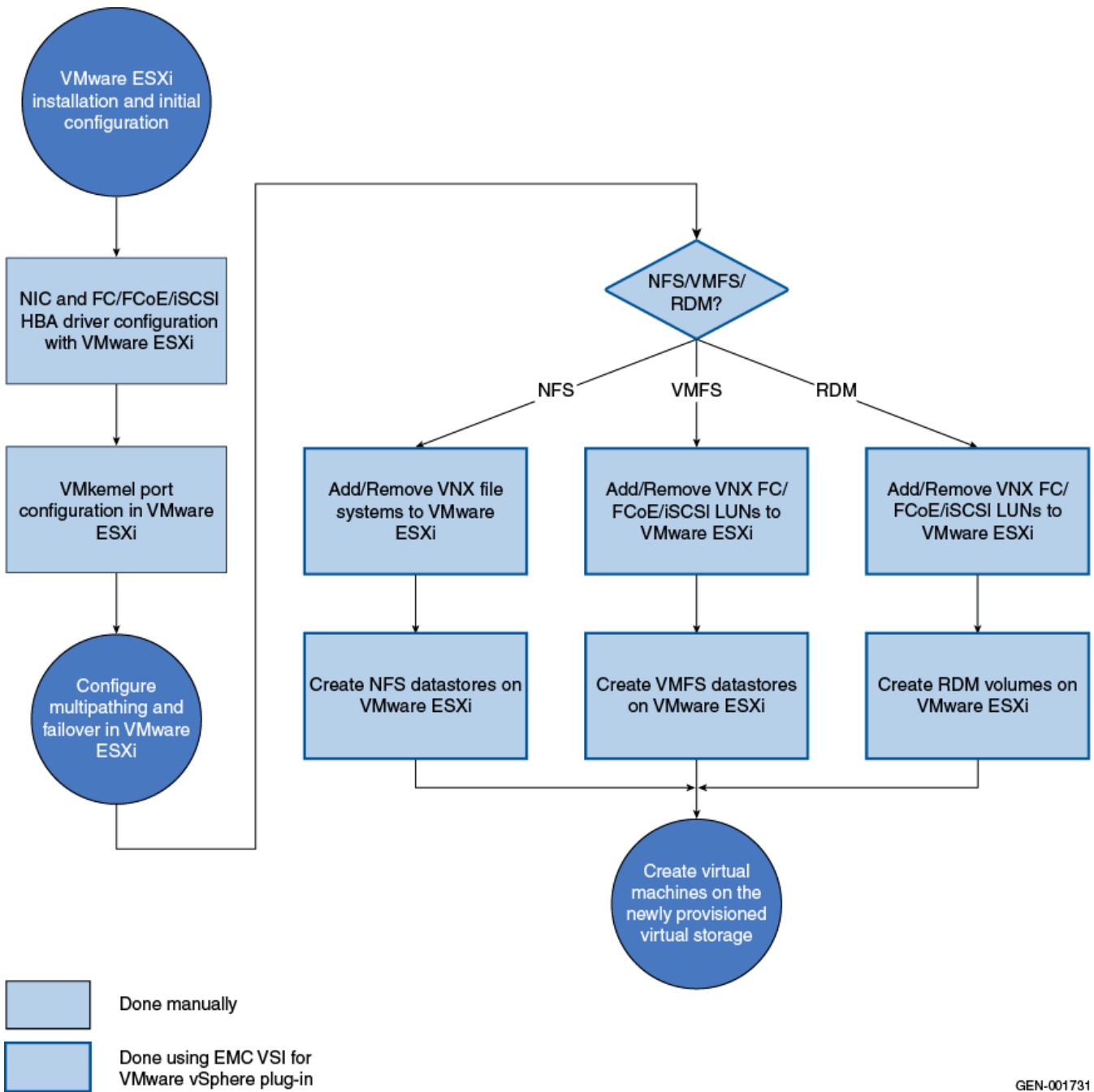
Installing VMware vSphere on VNX

This section describes the steps required to build the vSphere environment. You can use the vSphere Auto Deploy and Host Profile features to automate and accelerate multihost deployments.

The following sections provide configuration guidelines and recommendations:

- ◆ [“Installing the ESXi hypervisor”](#)
- ◆ [“Installing vSphere”](#)

Figure 7 illustrates the workflow for installing and configuring the ESXi systems with a VNX storage system.



GEN-001731

Figure 7 Configuration workflow

Installing the ESXi hypervisor

Use the instructions on the [VMware website](#) to install the ESXi hypervisor. You can install the ESXi image on the following storage types:

- ◆ A local server disk
- ◆ A USB storage device
- ◆ A storage area network (SAN) SCSI LUN in a boot-from-SAN configuration

Installing vSphere

This section provides instructions for installing vSphere. Installing the image on the SAN improves performance and reliability through:

- ◆ RAID-protected storage to reduce the risk of downtime potential of a local disk failure
- ◆ I/O distribution across multiple spindles and I/O channels
- ◆ Simplified host replacement in the event of a hardware failure

Note: vSphere 5.x includes Auto Deploy to reduce installation time for larger environments. The *vSphere Installation and Setup Guide* provides details on Auto Deploy.

Choose from the following storage protocols, depending on your infrastructure and environment requirements:

- ◆ **vSphere boot from SAN FC LUNs**—ESXi supports booting either through a Fibre Channel host bus adapter (HBA) or a Fibre Channel over Ethernet (FCoE) converged network adapter (CNA). Make sure you enable and correctly configure the adapter, so it can access the boot LUN. See your vendor documentation and the *vSphere Storage* document for more information.
- ◆ **vSphere boot from SAN iSCSI LUNs**—ESXi version 4.1 and later supports iSCSI software boot firmware (iBFT). This configuration requires that the network card supports software initiator boot and at least 1 Gb of throughput. Consult the *VMware Compatibility Guide* to verify that the device is supported before beginning this procedure.

Both protocols provide similar benefits and the configuration tasks are nearly identical.

Configuring a vSphere boot from SAN FC LUNs

1. Cable the hosts. Zone the HBAs to ensure that the host initiators log in to the VNX SPs when the host is powered on.
2. Gather the following information to configure the environment to use the selected front-end ports on the array:
 - ESXi hostname
 - IP addresses
 - HBA World Wide Name (WWN)—Obtain the WWN from the Unisphere Host Connectivity page after the initiators log in to the SPs, or from within ESXi.
 - VNX management IP address and credentials

Note: If storage zoning is not complete, obtain the HBA WWN from the SAN switch.

3. Power on the ESXi host.
4. Modify the following host BIOS settings to establish the proper boot order.
 - Ensure that the SAN boot device appears immediately after the peripheral devices.
 - Unless explicitly required, disable the local RAID SCSI controller on the host to reduce boot times.
 - For software iSCSI, enable iSCSI boot support on the network card.
5. Enable the FC, FCoE, or iSCSI adapter as a boot device, and scan the bus to initiate a Port Login.
6. Display the properties of the Array Controllers to verify that the adapter can access the VNX.
7. Access Unisphere to view the **Host Connectivity Status**. Verify that the adapters are logged in to the correct SP ports.
8. Boot from SAN requires manual registration of the HBAs. Select the new initiator records and manually register them using the fully qualified domain name of the host. Set the failover mode to **Asymmetrical Logical Unit Access (ALUA) mode 4** for support of vStorage API for Array Integration (VAAI).

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

9. Create a LUN of 20 GB or less on which to install the boot image. Do not store virtual machines within the datastore created from this LUN.
10. Create a storage group and add the host record and the new LUN to it.
11. 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.
12. Reserve a specific Host LUN ID to identify the boot devices.

For example, assign 0 as a Host LUN number to LUNs that contain the boot volume. This approach makes it easy to differentiate the boot volume from other LUNs assigned to the host.

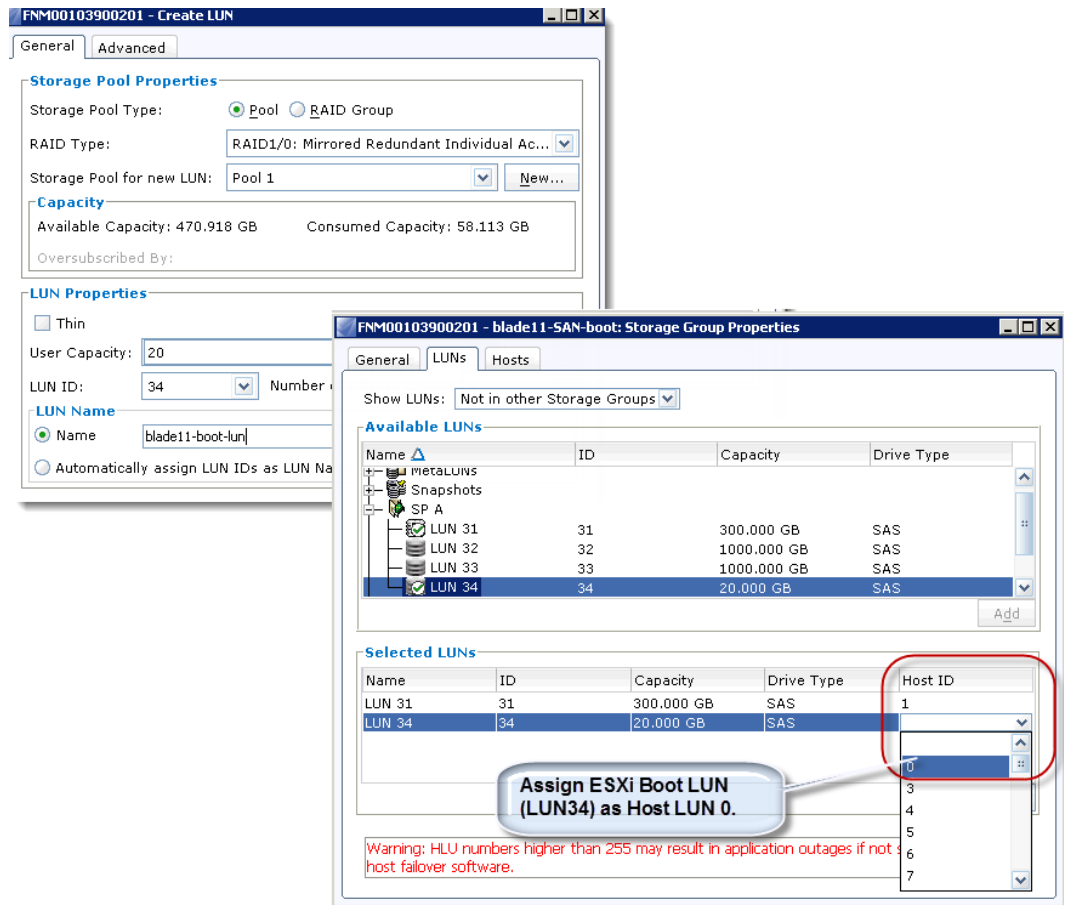


Figure 8 Unisphere LUN assignment for ESXi boot device

13. 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.

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

Configuring a vSphere boot from SAN iSCSI LUNs

During the system boot, access the iSCSI adapter configuration utility and configure the HBA as follows. Refer to the vendor documentation for instructions to enable and configure the following for the iSCSI adapter:

1. Set the IP address and IQN name of the iSCSI initiator.
2. Define the VNX iSCSI target address.
3. Scan the target.
4. Enable the boot settings and the target device.

A sample from a Broadcomm BIOS is provided in [Figure 9](#), illustrating the IQN Assignment for the VNX iSCSI target.

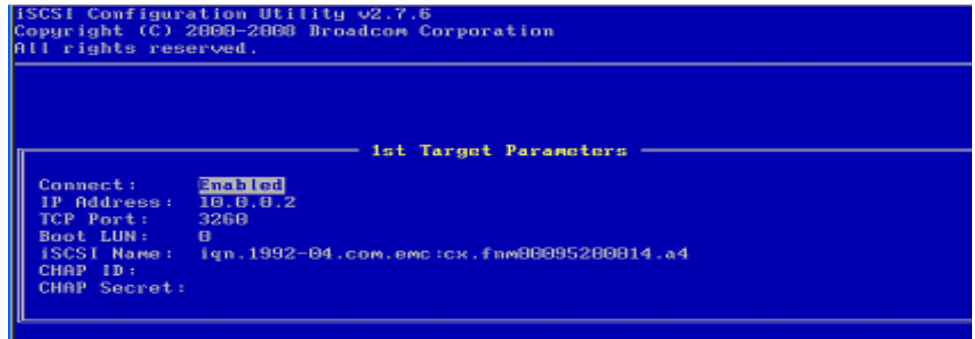


Figure 9 iBFT interface for VNX target configuration

5. Use Unisphere to configure an iSCSI portal on the VNX platform, as shown in [Figure 10](#).

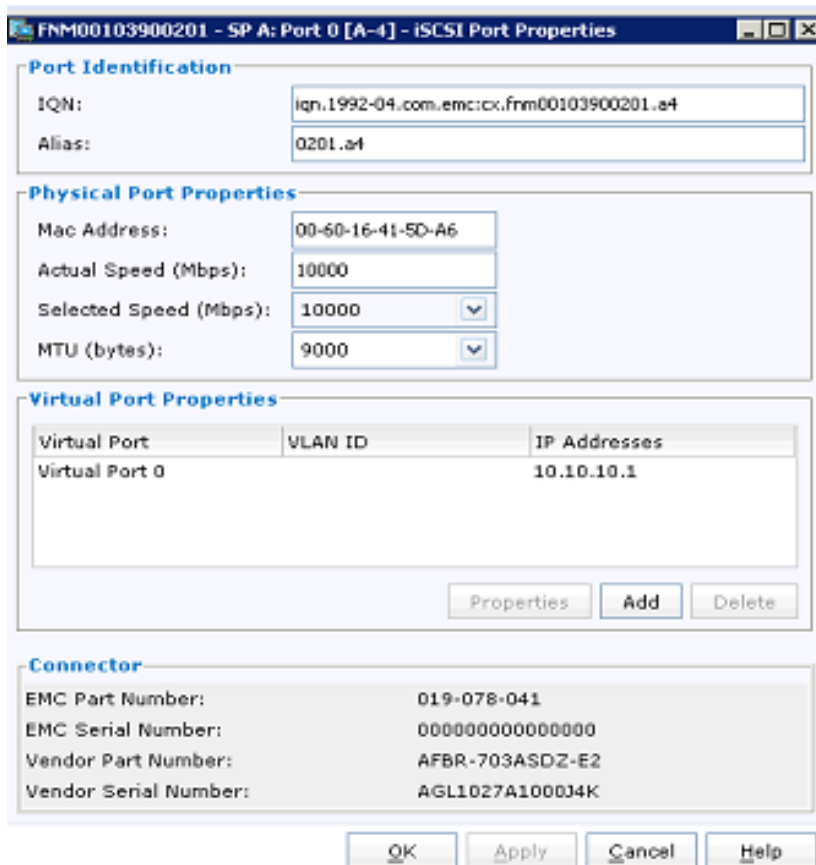


Figure 10 VNX iSCSI port management interface

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

6. Specify the IP address and IQN name of the iSCSI port from the previous step to configure the iSCSI target.

7. Optionally, enable CHAP for to enforce initiator authentication to the iSCSI target.
8. Configure the secondary target with the address information for the iSCSI port on VNX SP-B.
9. Open Unisphere and 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.
10. Proceed with the ESXi image installation.

Configuring vSphere with VNX

VNX storage systems are scalable to satisfy shared storage requirements in mid- to high-end vSphere environments. The VNX system addresses a broad range of application and scalability requirements, which makes it an ideal platform for vSphere. The following topics in this section provide recommendations to consider when you use vSphere with VNX storage:

- ◆ [“Host connectivity”](#)
- ◆ [“Configuring the Network”](#)
- ◆ [“iSCSI connectivity options for ESXi and 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 techniques
- ◆ Port zoning or WWN zoning
- ◆ Host adapter settings
- ◆ Storage port configuration

ESXi and VNX provide common support for Fibre Channel, FCoE, iSCSI, and NFS storage protocols as shown in Figure 11.

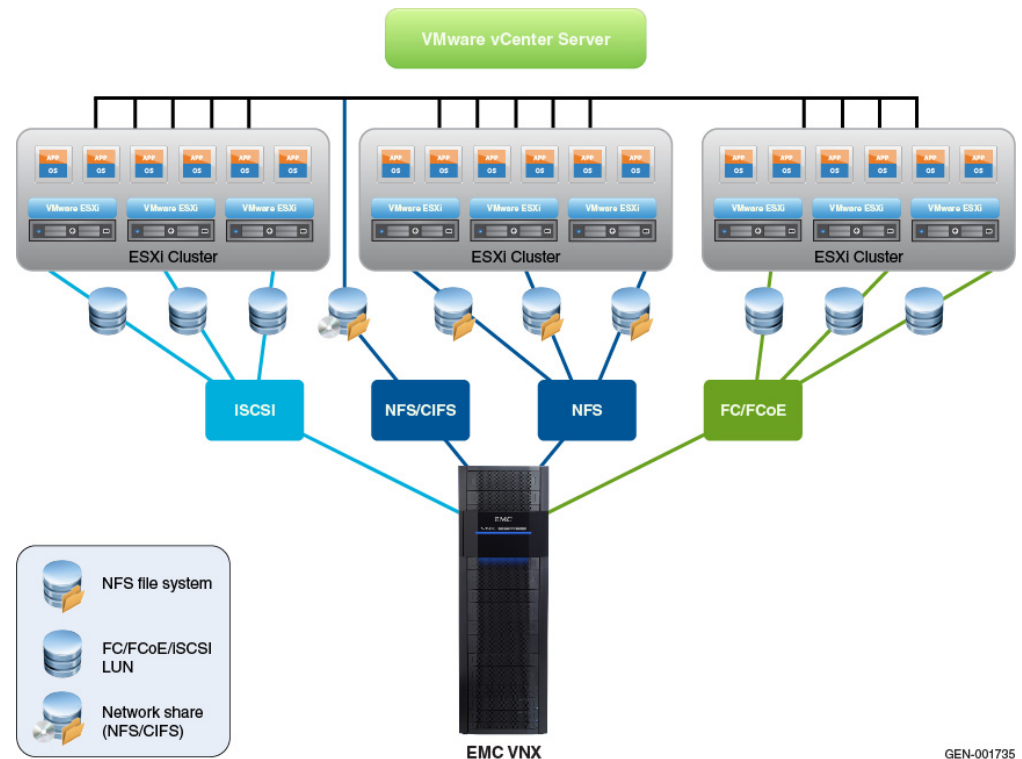


Figure 11 VNX storage with VMware vSphere

VNX also offers the CIFS/SMB3.0 file sharing protocol for sharing file systems on Windows virtual machines.

Note: You can configure ESXi hosts with multiple SCSI transports, such as Fibre Channel HBAs, and iSCSI software adapters, but accessing the same LUN with multiple SCSI transport types is not supported. For example, a host can access LUN 0 using FC and LUN 1 using iSCSI, but it cannot access LUN 2 using iSCSI and FC.

This section discusses the following topics:

- ◆ “Physical configuration”
- ◆ “Port configuration”
- ◆ “ESX HBAs queue depth”
- ◆ “Fibre Channel zoning”
- ◆ “Virtual local area networks”
- ◆ “Manual initiator registration”
- ◆ “Fibre Channel over Ethernet”

Physical configuration

Use the following general best practices when connecting ESXi hosts to the VNX storage system:

- ◆ Configure each ESXi host with at least two physical host adapters for device path redundancy between the host and the storage system.
- ◆ Cable each physical path through a separate switch for redundancy and fault tolerance.
- ◆ Logically, create a separate switch zone for each initiator-target pair, with each initiator zoned to a separate SP target port.
- ◆ Add all of the host initiators to a single storage group on the VNX.

Port configuration

VNX storage systems include a minimum of 8 Gb FC ports with expansion slots to accommodate additional FC, FCOE, iSCSI, and Ethernet I/O connectivity modules. VNX systems can be customized with connectivity options that scale to match host requirements and distribute host I/O to the storage system. EMC recommends the following:

- ◆ Ensure that ESXi hosts have a minimum of two physical paths to the storage system. Each path (or pair of paths for path counts greater than two) should 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 and achieve the best overall response times.
- ◆ Make note of port requirements for EMC MirrorView™ and EMC RecoverPoint® when planning port configurations.

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

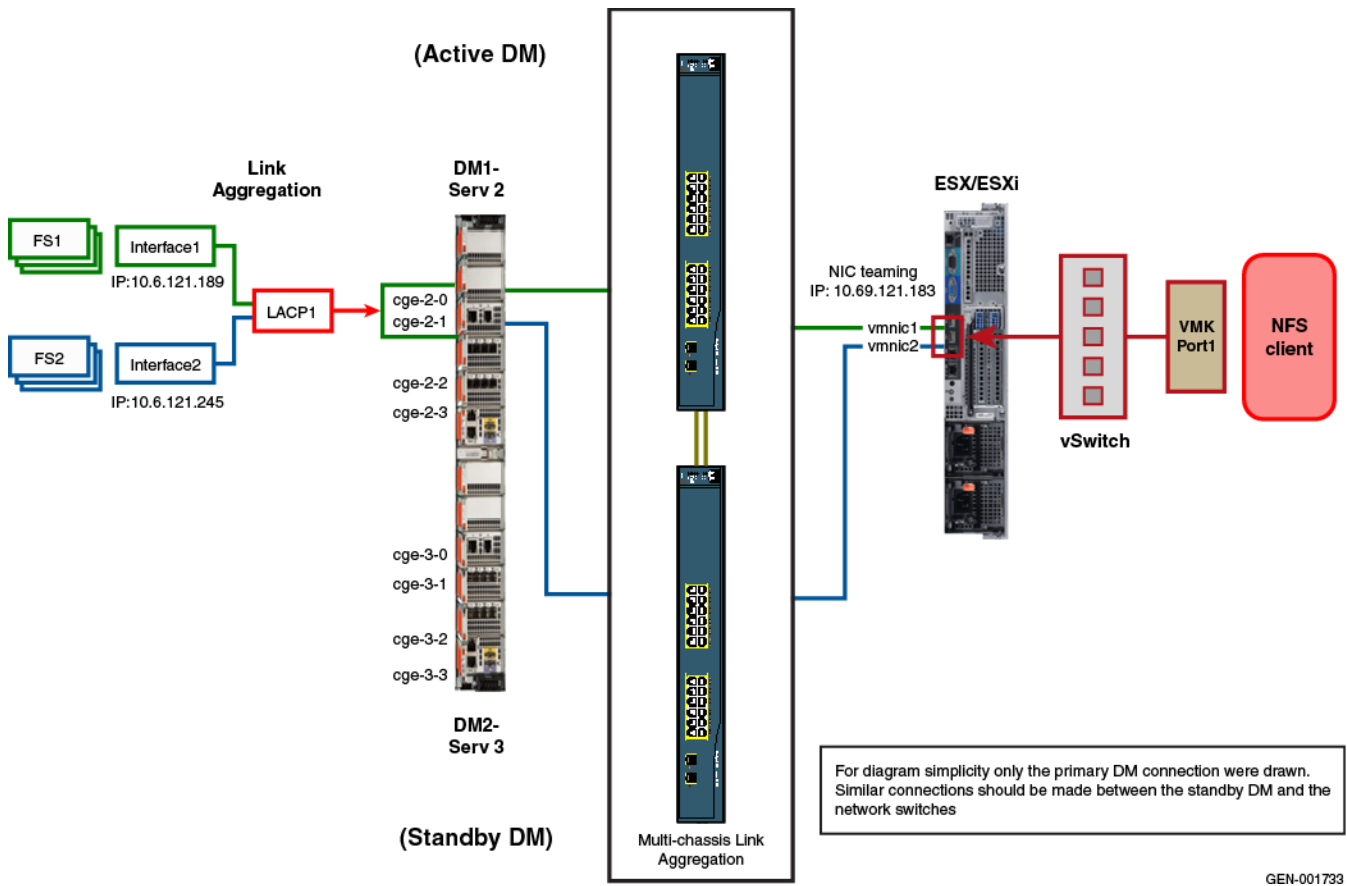


Figure 12 ESXi topology with FC/FCoE/iSCSI/NFS connectivity to VNX

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 on the ESXi or VNX system when the HBA is installed.

One possible exception is the HBA queue depth. The default maximum queue depth value of an ESXi5 FC HBAs is 64. That means the VMkernel will allow up to 64 outstanding I/Os at any given time. This value is sufficient for most workloads, particularly when more than three hosts are accessing a device.

Within the VNX, relevant I/O queues that could impact ESXi host performance are the front-end host port queue and the LUN queue that is being used to support the datastore. EMC recommends that you limit the maximum number of I/Os per front-end port to 1,600. You can add I/O modules to provide additional I/O paths to the storage devices.

The LUN queue is the most important consideration when tuning host adapter queues. In most cases, the only time you might consider modifying the maximum HBA queue depth is when all of the following are true:

- ◆ The LUN queue depth is larger than the cumulative queue depth of all host adapters accessing the LUN.
- ◆ The esx top value of the device queue used percentage (%USD) is continuously at 100.
- ◆ Queued commands (WQLEN) for the device are greater than 0.

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

Note: When altering the maximum queue depth, set **Disk.SchedNumRequestsOutstanding** to match this value.

If the multiple ESXi hosts are configured in a datastore cluster, the cumulative queue depth can quickly surpass the LUN queue. VMware Storage I/O Control (SIOC) helps avoid a situation where the host queue depths are set too high; however, this example is provided as an illustration of how to optimize a particular configuration. Unless you are familiar with the explanation provided, EMC recommends that you leave the queue depth at the default value of 64.

Fibre Channel zoning

VNX uses single-initiator, single-target zoning. For best results, configure only one active path between an initiator and the VNX SP. 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

While IP storage systems do not use the term *zoning*, a similar Ethernet concept is applied through virtual local area networks (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 network traffic isolation between ESXi IP storage adapters and the VNX IP storage adapters used to provide iSCSI and NFS connectivity.

Note: EMC recommends that ESXi VMkernel NICs be configured to access the VNX iSCSI using a separate broadcast domain to isolate network traffic between each ESXi iSCSI initiator port and VNX iSCSI target port. Separate broadcast domains can be established using separate nonrouted subnets or VLANs.

Manual initiator registration

In certain cases, such as boot from SAN, you must add host initiators to the VNX 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.

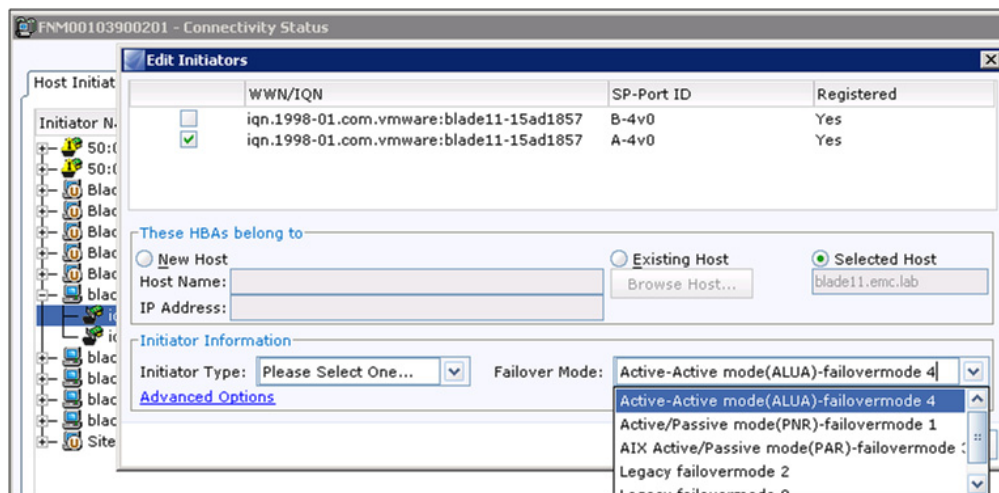


Figure 13 VNX configuration of host initiator

Relevant parameters for the initiator are:

- ◆ ESXi hostname—User provided
- ◆ ESXi management IP address—User provided
- ◆ Initiator type—CLARiiON/VNX
- ◆ Failover mode—Failover mode 4 (ALUA)

A failover mode defines the scope of SCSI commands that a host can use when connected to the VNX to support features such as LUN Trespass and multipath behavior. VNX offers four host failover modes, two of which are applicable for ESXi host initiators:

- ◆ **Asymmetrical LUN Unit Access (ALUA) or 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 that 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. EMC recommends using mode 4 for all ESXi initiators connected to VNX.
- ◆ **Active/passive mode (failover mode 1)**—This mode uses a single optimal or preferred path to the SP that was assigned to the LUN 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 earlier EMC CX[®] platforms.

ESXi version 4.0 and later are compliant with ALUA. When configured with failover mode 4, the ESXi host sends I/O to the VNX LUN using the active/optimized path. If an active/optimized path becomes unavailable, the host attempts to use another active/optimized path on the SP that owns the LUN. If all active/optimized paths are unavailable 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 becomes the LUN owner and satisfies all subsequent I/O requests. The Native Multipathing Plug-in [“Storage multipathing” on page 59](#) provides more details on path trespass and restore.

For vSphere version 5.1 and later, all LUNs are trespassed back to the default owner when the paths are restored.

Fibre Channel over Ethernet

Native Fibre Channel over Ethernet (FCoE) support, included with the VNX platform, offers a consolidated cabling option between servers, switches, and storage subsystems. FCoE connectivity allows general server network traffic and storage I/O 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.

Configuring the Network

Equipment

Consider the following recommendations when configuring 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 for NFS” on page 66](#) provides more details.
- ◆ Consider FCoE hardware adapters with 10 GbE converged network switches for consolidated storage networks. [“Fibre Channel over Ethernet” on page 41](#) provides more details.
- ◆ Select a switch vendor that includes 10 GbE support for NFS, iSCSI, or FCoE.

Configuring Ethernet networks

Consider the following when configuring IP storage networks:

- ◆ Use a dedicated physical switch or isolated VLAN to provide the most efficient network.
- ◆ On network switches that are also used for the storage network:
 - Enable flow control.
 - Enable spanning tree protocol with either RSTP or port-fast enabled.
 - Restrict bridge protocol data units (PDUs) on storage network ports.
- ◆ In general, ESXi host I/O is small and random; in most cases, jumbo frames (Ethernet MTU greater than 1,500) provide minimal benefit.

Large block I/O and sequential workloads can benefit from larger frame sizes. If configuring jumbo frames, ensure that each network interface (host, switch, VNX) in the I/O path uses the same MTU value.

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

Configuring the VMkernel port in ESXi

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

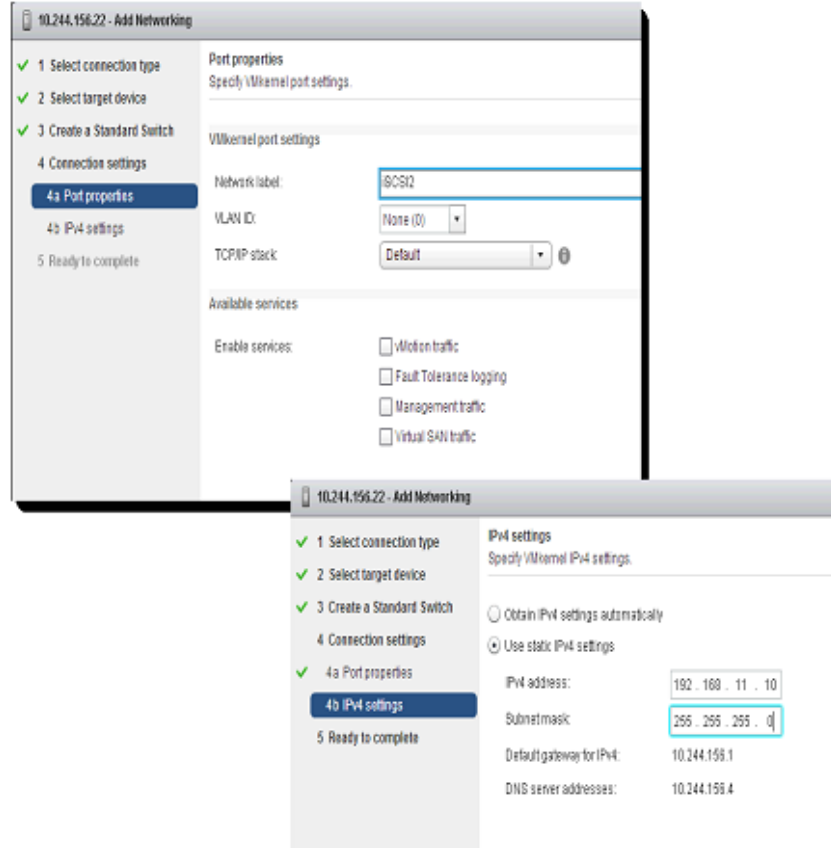


Figure 14 VMkernel port configuration

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

1. Create a new virtual switch to support the IP storage interfaces.
2. Assign a network label that describes what the interface is used for, such as NFS-DM1, iSCSI1, iSCSI-Oradb, and so on.
3. For NFS, assign a network adapter.

For best performance use a nonrouted interface on the same broadcast domain as the VNX NFS server.

4. In the **VMkernel - IP Connection Settings** dialog box, specify the following VMkernel IP settings:

- IP address
- Subnet mask
- Default network gateway

Notes:

- Do not use DHCP.
 - ESXi management and VMkernel interfaces share the same routing table of the ESXi host. As a result, the management interface routes storage I/O when the NFS server is configured to use the same subnet. To prevent this from happening, 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 connectivity options for ESXi and VNX

VNX iSCSI architecture provides a flexible IP storage system that can be customized for almost any network topology. Each iSCSI network port supports multiple virtual interfaces, and each virtual interface can reside on a different address and/or VLAN. Furthermore, each iSCSI interface presents a unique iSCSI target, allowing initiators to be configured to connect to one or more targets in support of high availability and multipath technology.

- ◆ “VNX iSCSI target”
- ◆ “ESXi iSCSI initiator”
- ◆ “VMkernel port group assignment”
- ◆ “Subnet configuration”

VNX iSCSI target

Figure 15 illustrates a basic iSCSI network design for a VNX that has a single SLIC with four Ethernet ports. Each port can support multiple virtual interfaces and be addressed across subnets and VLANs. In this illustration, a one-to-one relationship exists between the physical port and the subnet, and each port is configured with a separate subnet address.

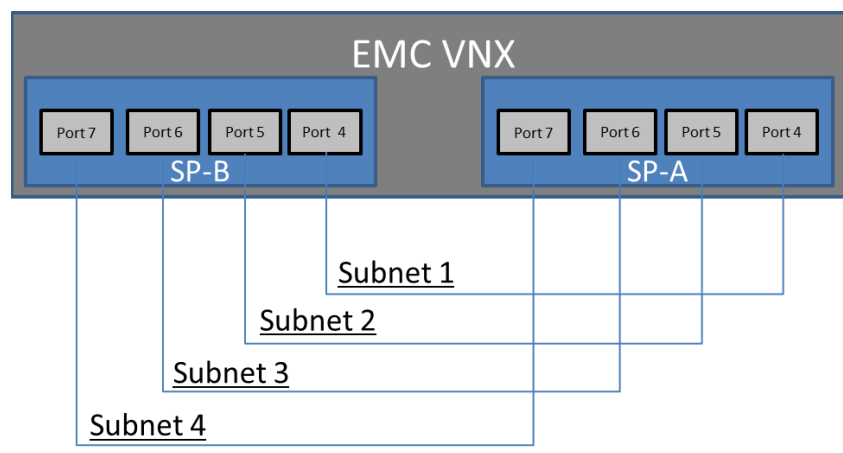


Figure 15 VNX multiple subnet iSCSI target port addressing

ESXi iSCSI initiator

ESXi offers two iSCSI connectivity options:

- ◆ **Hardware initiator**—This configuration uses a dedicated storage adapter card and Ethernet network to access the iSCSI target. Most adapters provide BIOS-level utilities to configure the initiator settings, including the VNX target addresses. Thereafter, the adapter appears and operates as a host storage adapter similar to those within the vSphere environment.
- ◆ **Software initiator**—This option is implemented as a VMkernel software driver that is dependent on one or more ESXi 1 GbE or 10 GbE network interfaces to transport the iSCSI traffic. Configuration of the software initiator is provided through vCenter or ESXi command line utilities.

This document covers the generic configuration of iSCSI with VNX. It focuses heavily on the software initiator, because proper configuration depends on understanding some topology nuances.

ESXi software initiator

The ESXi software initiator is implemented as a VMkernel device driver. It relies on one or more VMkernel ports (vmknics) on the ESXi host to support the iSCSI sessions and transport data to the VNX. Each VMkernel port depends on one or more physical network interfaces (vmnics).

- ◆ The software initiator should be configured with two or more physical paths (vmnics) to the storage system for redundancy and scaling.
- ◆ For performance reasons, iSCSI interfaces should not be shared with other ESXi host services.
- ◆ Use dedicated physical interfaces for iSCSI and provide additional interfaces for vSphere services such as NFS, vMotion, and Virtual Machine networks.

VMkernel port groups can be configured with a single physical adapter or grouped together into “NIC teams” for increased availability.

- ◆ NIC teaming provides increased reliability at the iSCSI session or path level; however, it does present the potential to reserve network resources that could be used to process I/O.
- ◆ Performance-centric environments using VNX might be better served by configuring separate port groups for each physical adapter and using the multipath and failover capabilities of VNX and the Native Multipathing Plug-in to mitigate a single vmnic failure.

VMkernel port group assignment

VMkernel port groups are associated with the software initiator in one of two ways: explicitly through port binding or implicitly by configuring the initiator with a discovery address.

Explicit assignment (port binding)

This configuration method is required for iSCSI storage systems that use a single iSCSI target. This option provides multipathing by binding multiple VMkernel port groups to the software initiator.

The following characteristics apply to port binding:

- ◆ The VMkernel port group is dedicated to the iSCSI initiator.
- ◆ All VMkernel port groups must use the same subnet address.

Note: VNX is a multitarget architecture. While technically possible, port binding is not supported on VNX.

Implicit assignment (recommended for VNX)

This option is used when the iSCSI storage system supports multiple targets or network portals. With this configuration, no VMkernel port groups are bound to the software initiator. The VMkernel determines which VMkernel port groups to use based on the port group's ability to access and establish an iSCSI session with the VNX iSCSI targets.

The following characteristics apply to configurations that do not use port binding:

- ◆ The VMkernel port group is shared with other VMkernel services.
- ◆ NFS servers configured on the same network may impact iSCSI service.
- ◆ Single subnet configurations will not distribute I/O across all adapters.

During the ESXi initiator-iSCSI target nexus, the initiator uses the dynamic discovery address to access the target and issue a SCSI “send targets” inquiry command. The iSCSI target responds with a list of all configured iSCSI targets and network (portal) addresses.¹

The VMkernel performs a network connectivity test between each VMkernel port group and VNX network portal address returned from the SCSI send targets command. Validated network paths are used to establish iSCSI sessions with the VNX. Paths that result in network or login failure are not retried.

1. The addresses are visible in the static address tab of the vCenter software initiator properties window.

Figure 16 illustrates the steps required to configure the iSCSI initiator using implicit assignment.

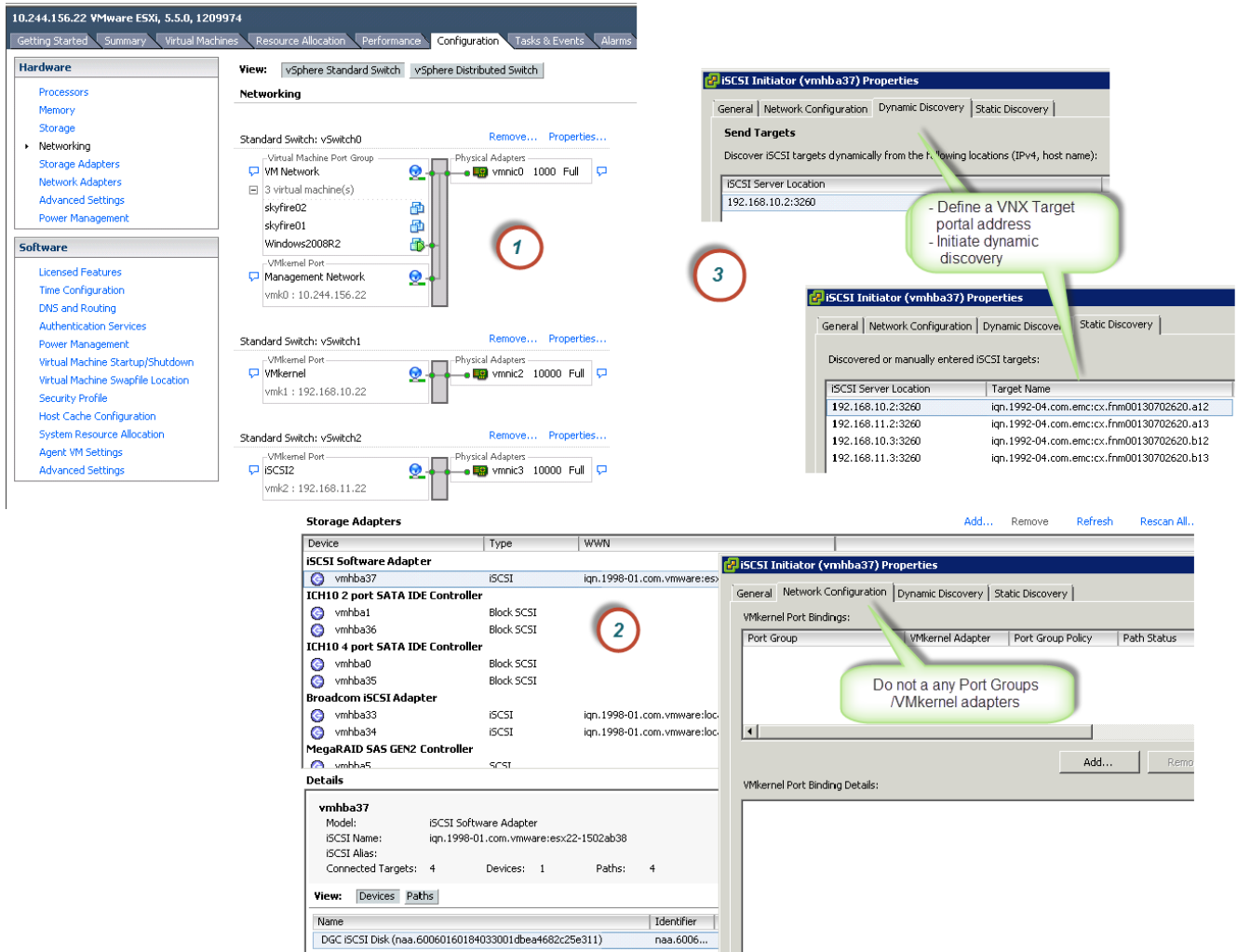


Figure 16 iSCSI Software initiator configuration in vSphere 5

The VNX network configuration is extremely flexible. The VNX iSCSI targets can be presented with a single network or multiple networks as illustrated in Figure 17. Basic configurations use a single subnet which simplifies the configuration of the software initiator.

Subnet configuration

Single subnet configuration

In this example, the VNX system has two iSCSI targets configured (one on SPA and one on SPB). This results in one optimal and one non-optimal path to each storage pool LUN.

Multiple Subnet Configuration

In this example the VNX system has four iSCSI targets configured (two on SP-A and two on SP-B). Two networks are used for the host and storage addresses as illustrated in Table 2. Assuming the networks are non-routable, the configuration results in four iSCSI sessions with two optimal and two non-optimal paths to each storage pool LUN.

Note: If the VMkernel network ports can route I/O to all four VNX target ports, each LUN will have 32 potential I/O paths resulting in 8 active/optimal paths and 8 non-optimal paths. However, the additional sessions to the targets do not provide a performance improvement.

Table 2 Multiple Subnet Configuration IPs/IQNs

Network	Target IQN
192.168.10.x	iqn.1992-04.com.emc:cx.fnm00103900200.a4 iqn.1992-04.com.emc:cx.fnm00103900200.b4
192.168.11.x	iqn.1992-04.com.emc:cx.fnm00103900200.a5 iqn.1992-04.com.emc:cx.fnm00103900200.b5

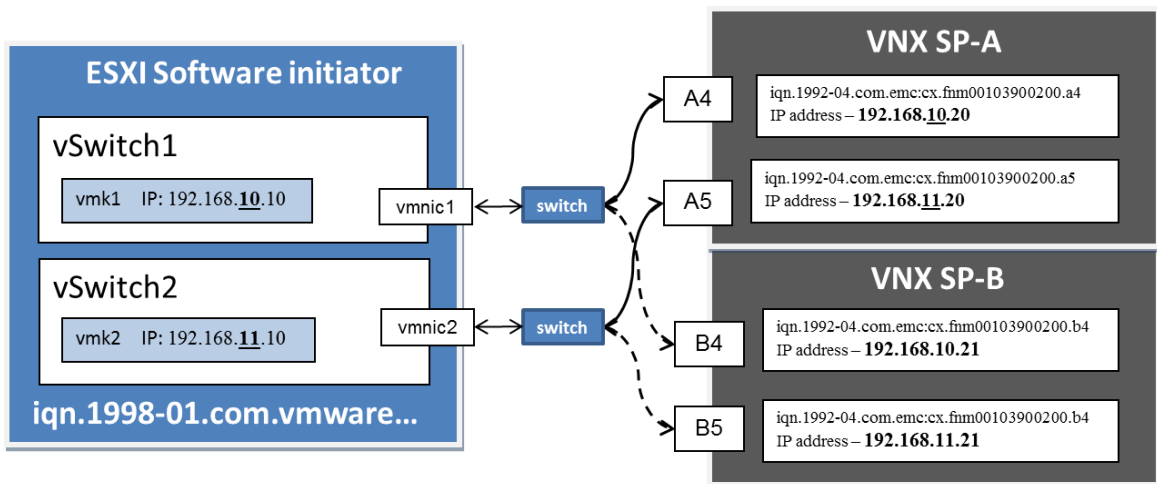


Figure 17 Topology diagram for multiple-subnet iSCSI configuration

Single network address space

All of the ports are on the same subnet: 192.168.10.0/24. Based on the Flare 30 changes, this configuration should be supported from a technical standpoint so that you can configure multiple VMkernel adapters and all VNX iSCSI FE ports using the same subnet.

Advantages and disadvantages of using a single network address space are:

- ◆ **Pros:** Simple to configure, scalable
- ◆ **Cons:** Each initiator establishes a session with all VNX target ports.
 - Results in eight iSCSI sessions in this example
 - Results in 16 sessions when 4 VMkernel and target ports are used
 - Requires NIC teaming or VLANs to take advantage of all paths

Multiple subnet configuration (VNX Config)

The second option is to create multiple subnets or broadcast domains. This is done by defining separate network addresses for each VMkernel NIC and VNX SP port pair. [Figure 17](#) illustrates a configuration that uses two subnets (subnet 192.168.10 and 192.168.11). To simplify this example, the endpoint or host component of the network address is the same for all interfaces.

Advantages and disadvantages of using multiple subnets are:

- ◆ **Pros:** 1:1 source to target/LUN session subscription
Performs I/O distribution across all paths when the host and target LUN are configured with native multipathing or PowerPath.
- ◆ **Cons:** Requires additional network setup
The host can still establish eight iSCSI sessions if subnets are routable.

Configure separate broadcast domains for each iSCSI initiator and VNX target pair (that is, SP-A4 and SP-B4). A broadcast domain can be created in either of the following ways:

- ◆ Configure a network subnet and place all of the nodes in that subnet.
- ◆ Configure a VLAN that is port-based or one that uses packet tagging. With port-based VLANs all switch ports are assigned to a particular VLAN ID. Network traffic is restricted to ports that are assigned to that VLAN. Packet tagging uses a field in the Ethernet frame to carry the VLAN ID. Only systems that have tagging configured for that VLAN ID will see those packets.

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 that will be used in the event that both active/optimized paths become unavailable.

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

For additional connectivity, configure additional VMkernel port groups and use network addresses to distribute them across additional VNX iSCSI network portals. The sample configuration illustrated in Figure 18 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.

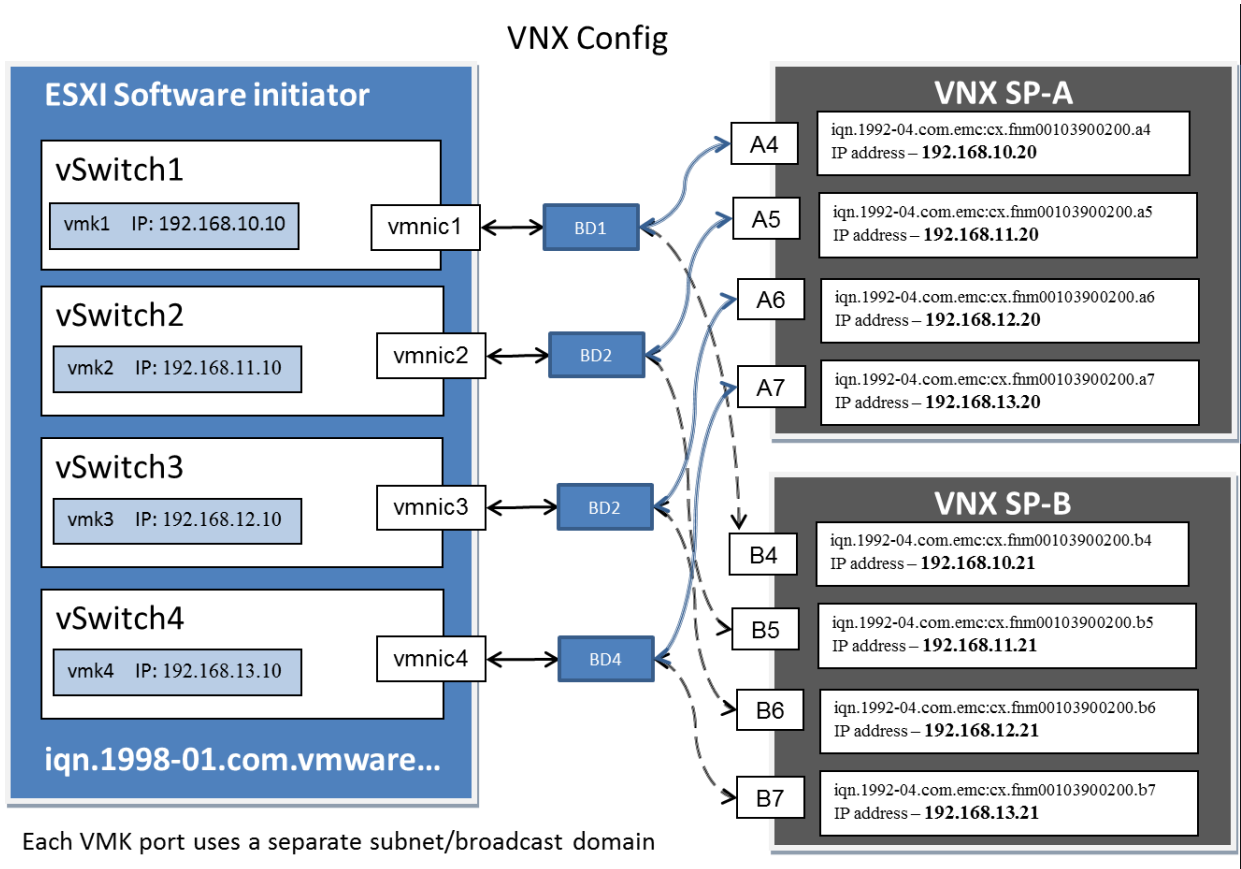


Figure 18 Recommended configuration for VNX iSCSI targets

Delayed acknowledgement settings for iSCSI

In some circumstances the VNX and ESXi hosts have encountered suboptimal iSCSI performance when using 10 GbE interfaces with the delayed acknowledgment (Nagle's algorithm) enabled.

Delayed acknowledgement is a TCP optimization intended to reduce network packets by combining multiple TCP acknowledgements into a single response. This optimization works well when many TCP packets are being transmitted between the ESXi host and the VNX. However, in some cases, such as when a single virtual machine or ESXi host performs sequential writes, the VNX does not have any data to return. In this case, the host waits for an acknowledgement to its write, and, due to Nagle, the VNX groups multiple acknowledgments and waits for the timeout (200 ms) to elapse before sending them all in a single response. This behavior has the potential to insert 200 ms delays into the I/O stream.

Disable the iSCSI Delayed Acknowledgement setting on the 10 GbE NIC in cases where performance delays are observed.

Figure 19 illustrates how to disable this setting.

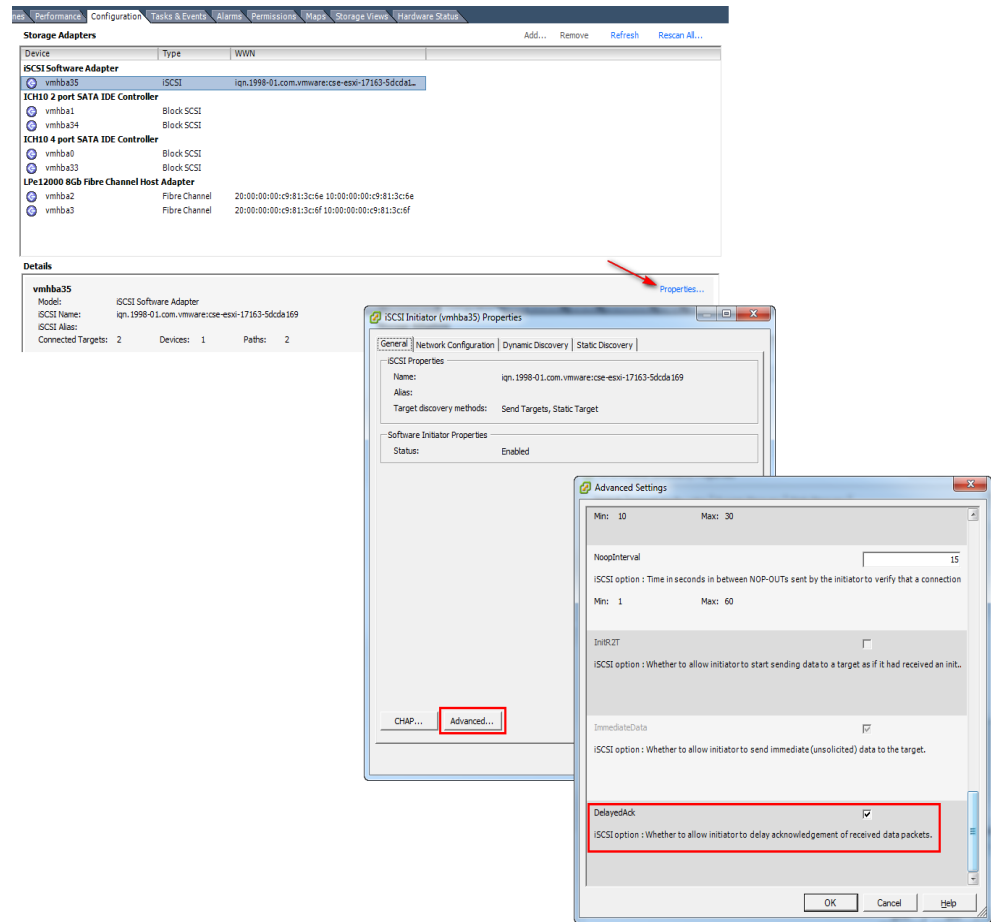


Figure 19 Disable Delayed Acknowledgement setting on storage adapter

VMware provides the following steps in Knowledge Base article 1002598 to disable delayed acknowledgement in ESX/ESXi hosts that might experience read/write performance issues with certain storage arrays:

1. Log in to the vSphere Client and select the host.
2. Select the **Configuration** tab.
3. Select **Storage Adapters**.
4. Select the iSCSI vm HBA to be modified.
5. Click **Properties**.
6. Do one of the following:
 - Modify the Delayed ACK setting on the discovery address (recommended):
 - a. On a discovery address, select the **Dynamic Discovery** tab.
 - b. Select the **Server Address** tab.
 - c. Click **Settings**.
 - d. Click **Advanced**.

- Modify the Delayed ACK setting on a specific target:
 - a. Select the **Static Discovery** tab.
 - b. Select the target.
 - c. Click **Settings**.
 - d. Click **Advanced**.
 - Modify the Delayed ACK setting for all targets:
 - a. Select the **General** tab.
 - b. Click **Advanced**.
7. Reboot the host.

Provisioning VNX storage for vSphere

VNX storage is presented to ESXi hosts in two forms: NFS exported file systems and SCSI LUNs. While NFS file systems are used only 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.

Creating an NFS datastore using VSI

EMC provides vCenter integration tools to automate and simplify the creation of storage devices and datastores using the VSI Unified Storage Management feature.

The high-level steps to configure VNX NFS file systems for vSphere are:

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

You can use Unisphere to complete these steps manually or you can use the Unified Storage Management feature, as follows:

1. From the vSphere Client right-click 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**. and provide the information for the new VNX in the **Add Credentials** wizard.
6. In the **Datastore Name** field, type the datastore name, and then click **Next**.

The **Data Mover Details** dialog box appears, as shown in [Figure 20](#).

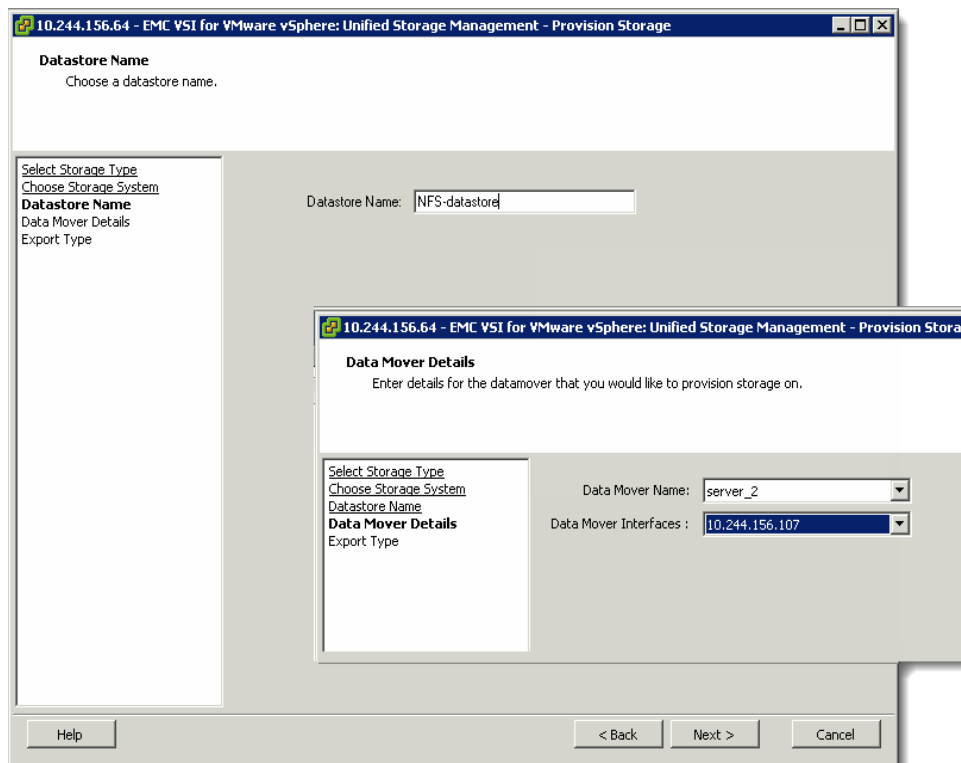


Figure 20 File storage provisioning with Unified Storage Management

7. From the **Data Mover Name** list box, select a Data Mover.
8. From 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**
10. From the **Storage Pool** list box, select a storage pool.

Note: All available storage within the storage pool is displayed. Ensure that the storage pool you select is designated by the storage administrator for use by VMware vSphere.

11. 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.
12. If thin provisioning is 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. In the **New NFS Export** window, type the values for the initial capacity and maximum capacity of the datastore.

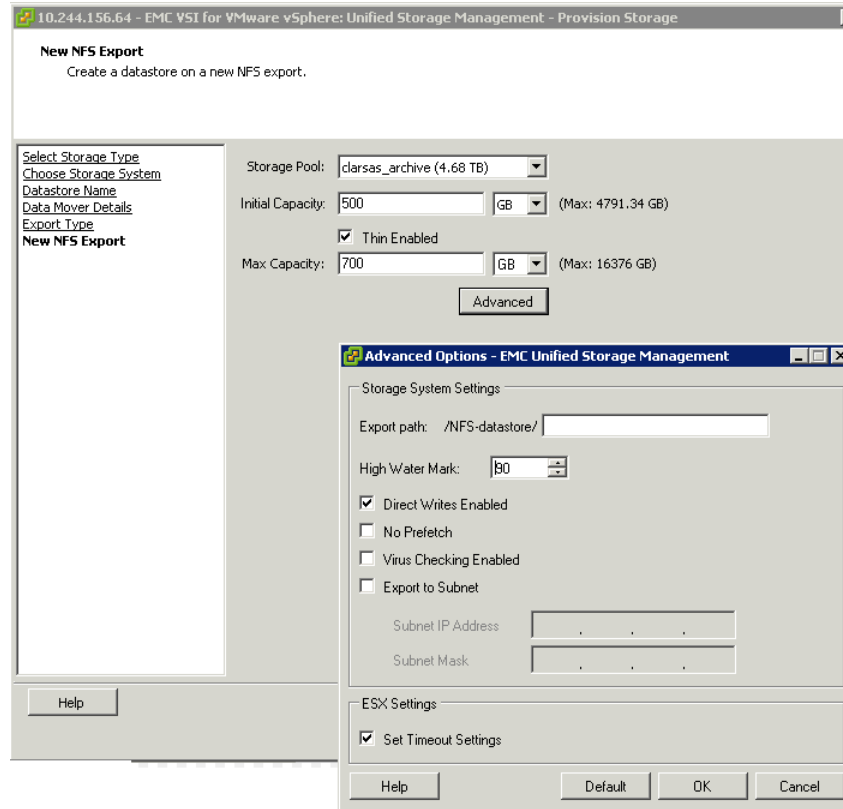


Figure 21 Creating a new NFS datastore with Unified Storage Management

13. Select the unit of measurement from the list box to the right.

If **Virtual Provisioning** is enabled for the file system, the maximum capacity is required. [Figure 21](#) shows an initial capacity entered in the **Max Capacity** field for the NFS export.

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

To add a VMFS datastore to a vSphere environment, you must create a LUN, unmask the LUN, rescan the host, and create a VMFS datastore.

The Unified Storage Management feature of VSI provides the following benefits:

- ◆ Offers an integrated workflow to automate LUN creation, LUN unmasking, host rescan, and VMFS datastore creation.
- ◆ 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 that the LUN can be surfaced as an RDM disk to a virtual machine.

Use the Unified Storage Management feature to provision storage as follows:

1. Right-click a vSphere object, such as a host, cluster, folder, or data center 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.

2. Select **EMC > Unified Storage**.
3. Select **Provision Storage**.

The **Provision Storage** wizard appears, as shown in [Figure 22](#).

4. Select **Disk/LUN**, and then click **Next**.

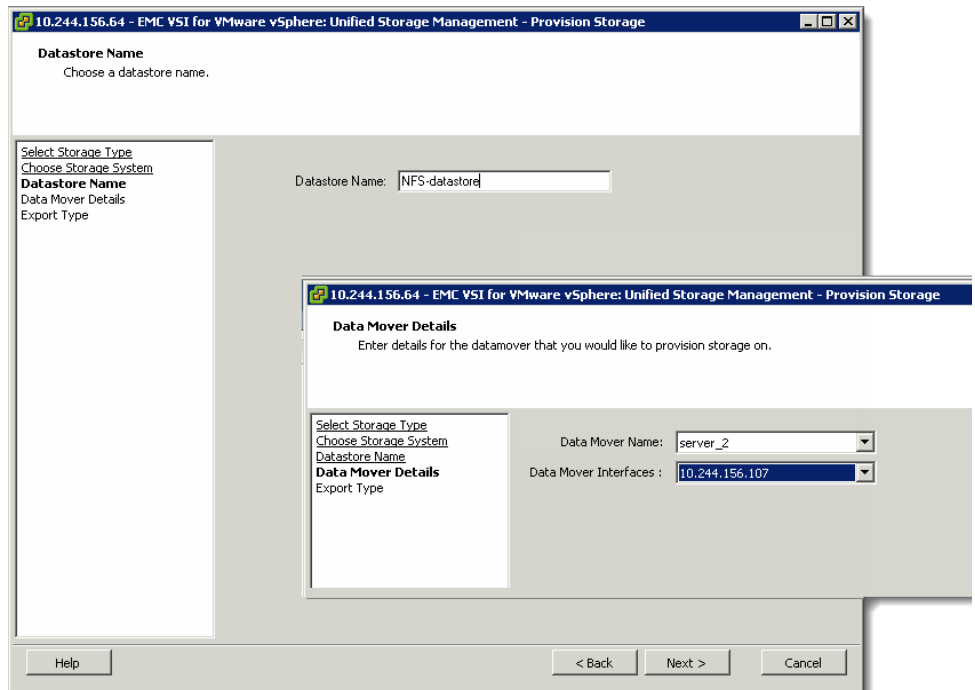


Figure 22 File storage provisioning with Unified Storage Management

5. In the **Storage System** table, select a VNX. If a VNX does not appear in the Storage Array table, click **Add** and add the VNX storage system using the Add Credentials wizard.
6. Select the storage pool or RAID group on which you want to provision the new LUN, and then click **Next**.
7. Select the datastore volume format as VMFS-3 or VMFS-5, and then click **Next**.
8. Select **VMFS Datastore** or **RDM Volume**.
9. Select an SP to own the new LUN and select **Auto Assignment Enabled**. Click **Next**.

Notes:

- Install and correctly configure failover software for failover of block storage.
 - 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.
-

10. For VMFS datastores, complete the following steps:
 - a. In the **Datastore Name** field, type a name for the datastore.
 - b. From the **Maximum File Size** list box, select a maximum file size.
11. From the **LUN ID** list box, select a LUN number.
12. From the **Default Owner** list box, select the SP that will own the new LUN.
13. In the **Capacity** field, type an initial capacity for the LUN and select the unit of measurement from the list box.

Figure 23 illustrates this action.

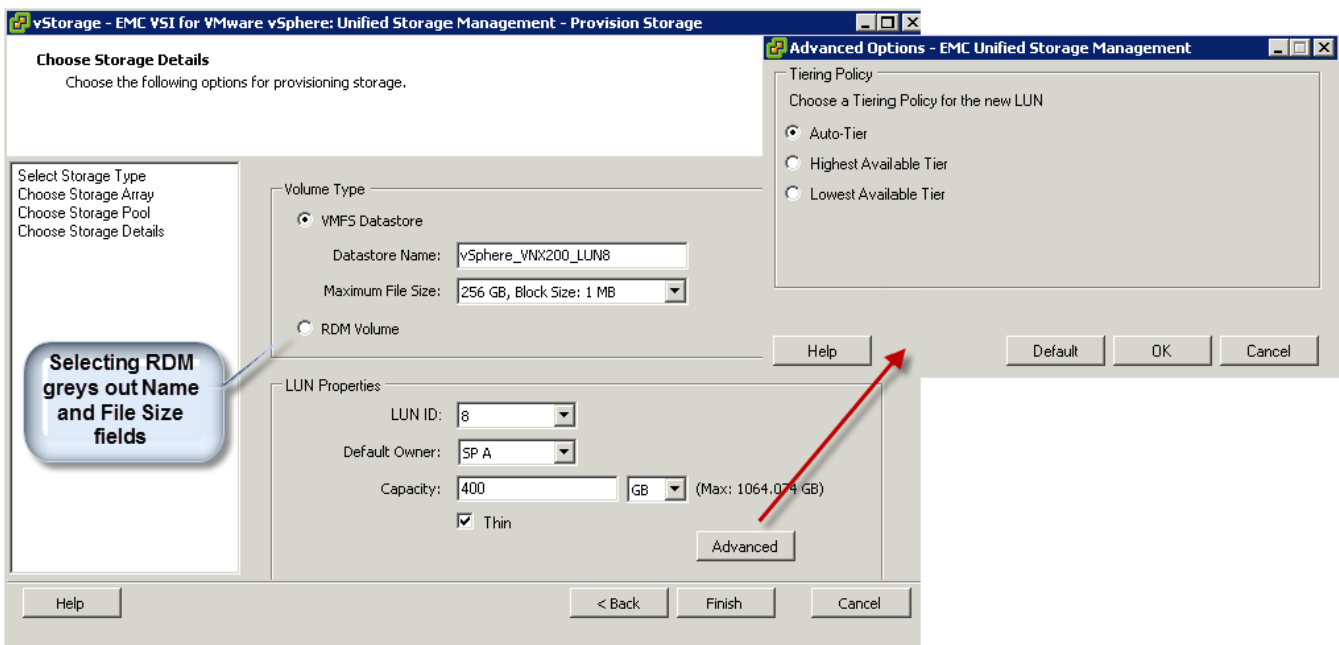


Figure 23 Creating a new VMFS datastore with Unified Storage Management

14. Click **Advanced** to configure the VNX FAST VP policy settings for the LUN.

The tiering policy options are:

- **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.

15. Click **Finish**.

When these steps are completed, Unified Storage Management initiates 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

16. 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.

Begin storage planning with an assessment of the application requirements. The primary factors that determine the storage configuration are:

- ◆ 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.

The following sections include general considerations and recommendations that apply to VNX storage arrays:

- ◆ [“Datastore virtual machine density”](#)
- ◆ [“Best practices for extending a datastore”](#)

- ◆ [“Solid state volumes for VNX OE for File”](#)
- ◆ [“General recommendations for storage sizing and configuration”](#)
- ◆ [“Storage multipathing”](#)

Datastore virtual machine density

When sizing the number of virtual machines per datastore, the following points are useful:

- ◆ vSphere support for VAAI and Storage I/O Control (SIOC) alleviates many of the historical factors that limited virtual machine scalability in a datastore. vSphere 5 also includes features such as Storage Distributed Resource Scheduler (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 correct 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 hundreds of thousands of IOPS. EMC has produced results for VDI that illustrate support for hundreds of virtual machines with a datastore with medium workloads (5 IOPS per virtual machine).
- ◆ For performance-sensitive environments where ESXi host clusters generate significant IOPS, create multiple LUNs to distribute the I/O across multiple LUN queues.
- ◆ For non-SIOC environments, the VMkernel serializes and queues I/O 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” on page 87](#).
- ◆ If SIOC is not enabled, this control falls to a number of other ESXi host parameters including **Disk.SchedNumReqOutstanding**, which, by default, limits to 32 the number of requests the host sends to a LUN. The value limit ensures that no single virtual machine monopolizes the LUN queue.

Best practices for extending a datastore

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

Using multi-LUN volumes adds unnecessary complexity and management overhead. If the VNX has enough free space, use the following best practices for multi-LUN extents:

- ◆ 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 SDRS to create a datastore cluster and allow it to manage virtual machine placement.

Solid state volumes for VNX OE for 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 enterprise flash drives (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.
- ◆ Volumes can be striped across EFD LUNs from the same RAID Group.

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, follow these 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.
- ◆ The VM boot virtual disk can reside on either an NFS or VMFS datastore. 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.
- ◆ Do not create more than three virtual machine snapshots, and do not keep them for an extended period of time. Instead, use a 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. VNX Snapshots offload the overhead on the ESXi host and provide a suitable alternative to vSphere snapshots.
- ◆ 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 such as booting, virus scanning, and logging on in large numbers.
 - Use LUNs enabled with FAST Cache, 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 candidates 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 establishes two or more I/O paths between a host and a LUN to address two important areas:

- ◆ **Reliability**—Multiple I/O paths ensure that access to application data is maintained in the event of a component failure.
- ◆ **Scalability**—Hosts can parallelize I/Os across multiple storage adapters to increase efficiency and balance storage resource utilization.

VNX storage systems have two SPs or storage targets. Access to a LUN is provided through all front-end host ports on the SPs. VNX provides several target access methods that determine how the LUN IO is processed. Pool LUNs configured for vSphere default to asymmetric active/active or ALUA mode as illustrated in [Figure 24](#).

VNX storage systems define a set of target port groups for each new LUN. One port group is assigned for the SP that owns the LUN and one for the nonowner. The assigned owner is also the default owner for that LUN, which means any time an event occurs relating to ownership, the VNX will attempt to place the LUN back on that SP. The port group of the owning SP provides the optimal path to the LUN through all of its front-end I/O ports. The peer SP provides a non-optimal path that can also satisfy I/O; however, the I/O must traverse an internal bus to satisfy the request.

[Figure 24](#) illustrates the concept of LUN ownership and I/O paths. When a LUN is owned by SP-A, optimal paths to the LUN are through the I/O ports of SP-A.

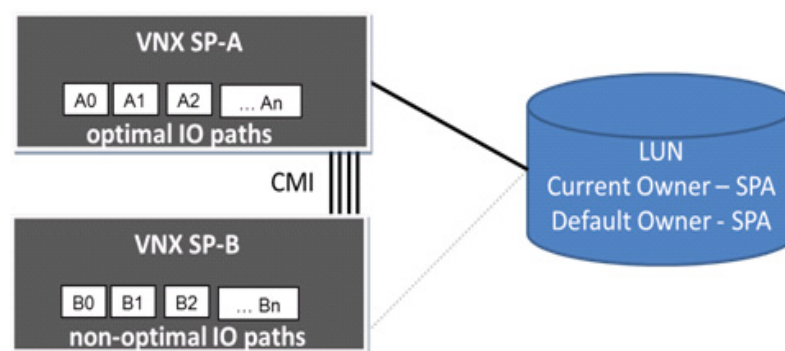


Figure 24 LUN ownership

LUN trespass

VNX transfers LUN ownership to the peer SP when certain host or storage system states are encountered. Examples include instances when:

- ◆ A host loses access to the SP due to an environmental or hardware failure.
- ◆ The VNX storage processor is undergoing a software update.
- ◆ A significant number of I/Os are being received down the non-optimized path.

Under these conditions LUN ownership is transferred to the peer SP and hosts must use the paths to that SP to perform I/O to the LUN. 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 40](#). 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 versions.

ESXi hosts are ALUA-aware. When accessing a VNX LUN in ALUA mode, the VMkernel can issue ALUA commands to determine the state of the LUN path, including the optimal path. The default LUN configuration for ESXi issues I/O only to the optimal paths unless they all become unavailable and only non-optimal paths remain. If all active optimized paths become unavailable the host issues an explicit trespass to the peer SP and resumes I/O using that SP.

During a host boot the Native Multipathing Plug-in module uses ALUA commands to identify the default owner and optimal paths to its LUNs. If a LUN has been trespassed by the host and the path to the default owner is available, the host issues a command to VNX to restore it to the default owner.

These processes are illustrated in Figure 25.

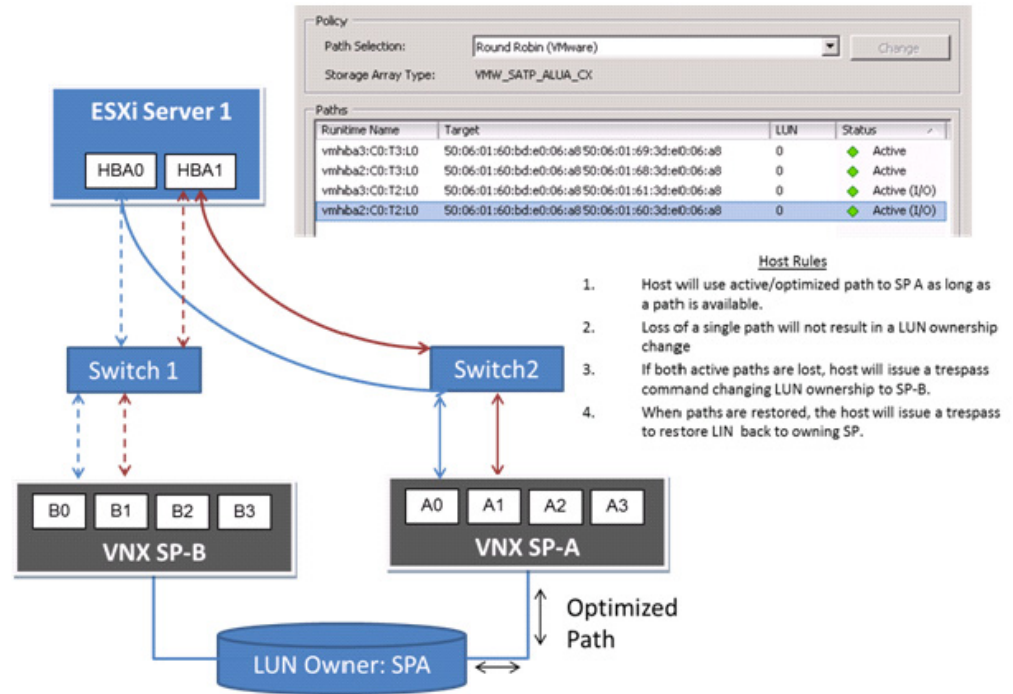


Figure 25 LUN trespass

VNX active/active LUN assignment

The VNX provides the capability to assign a classic LUN to both SPs and thus service I/O to the LUN through both SPs. With true active/active, a LUN can be simultaneously read from and written to using both storage processors.

Active/active applies to classic LUNs, which are LUNs created from RAID group storage. LUNs created from storage pools continue to provide ALUA mode, which is described in “VNX availability” on page 21.

Active/active is possible because MCx introduced a locking technology for LUN read and write access. SPs communicate with each other using the Communication Manager Interface (CMI). The new VNX hardware has a PCI Gen3 bus connecting both SPs, over which CMI runs between SPs, as illustrated in [Figure 26](#).

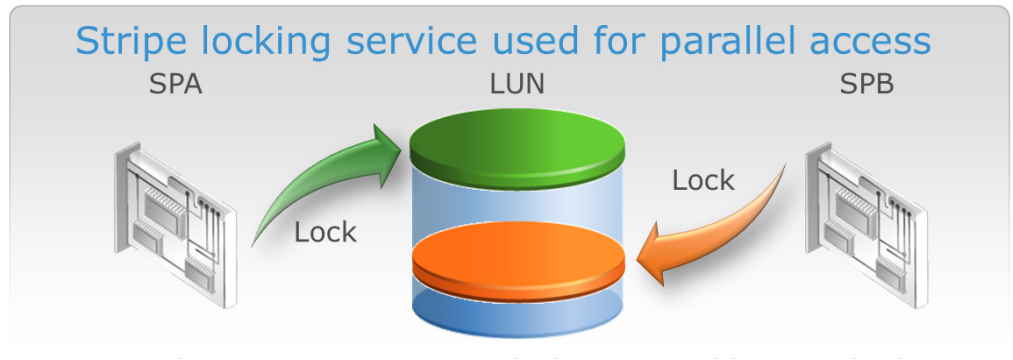


Figure 26 Stripe locking service

An SP must request an exclusive lock for a logical block addressing (LBA) range on a LUN prior to proceeding with the write from a locking service. A locking service communication between SPs is done over the CMI bus. The SP proceeds with the write on a successful lock grant.

Similarly, a shared-read lock must be taken out for reads. The SP needs to know whether to read from the disk or from a peer SP's cache, if it has newer data than the disk (dirty cache pages have not been flushed from the peer SP to the disk yet). Once the I/O is complete, the lock is released.

Active/active LUNs provide the potential for significant performance improvements in most application environments.

vSphere native multipath

The ESXi VMkernel provides a pluggable storage architecture (PSA) to support different multipath modules. [Figure 27](#) illustrates the PSA used with vSphere.

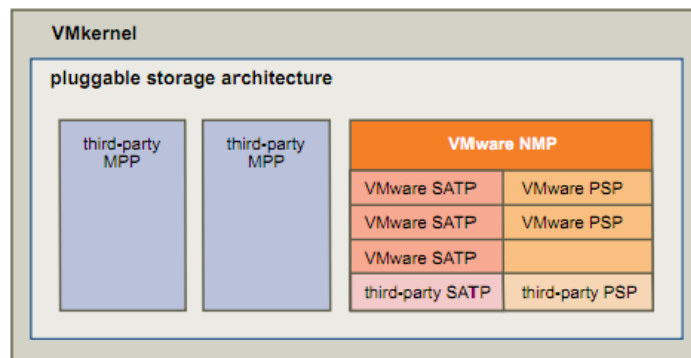


Figure 27 VMkernel pluggable storage architecture

The default module is Native Multipathing Plug-in, 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

The Native Multipathing Plug-in provides the framework to discover new LUNs and identify the storage array type plug-in (SATP), the initiator mode, and LUN properties such as the default SP that owns the LUN.

The Native Multipathing Plug-in uses the SATP to assign a PSP to the LUN. Run the `esxcli storage nmp satp list` command to view the rules.

Native Multipathing Plug-in PSPs

vSphere has four native PSPs:

- ◆ **Fixed Path Array Preference (AP)**—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 versions. 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. This 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. This is the default vSphere 5.0 and 5.1 PSP for VMW_SATP_CX.

Each SATP uses a predefined PSP agreed upon by VMware and the storage vendor.

Figure 28 illustrates the `esxcli` command output to identify the PSPs used for VNX systems in vSphere 5.1.

```

10.244.156.23 - default* - SSH Secure Shell
File Edit View Window Help
~ # 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 28 Esxcli command output

ESX 5.1 contains an enhancement to the Round-Robin Native Multipathing Plug-in PSP that allows auto restore to the preferred (default) VNX SP when a fabric failure (failed HBA/NIC/CNA, switch, or SP front-end port) to that preferred SP is repaired. [Table 3](#) shows the recommended Native Multipathing Plug-in PSP.

Table 3 Recommended Native Multipathing Plug-in path selection

ESX version	VNX software version	Recommended Native Multipathing Plug-in 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 that allows the Round-Robin Native Multipathing Plug-in PSP to auto restore to the preferred (default) VNX SP after the preferred SP reboots, whether the reboot is manual, due to failure, or is part of an array software upgrade (non disruptive 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. This environment provides the benefits of multiple active/optimized paths for I/O scalability and auto restore to the preferred SP after any fabric failure or SP reboot.

Use Round-robin when using Native Multipathing Plug-in.

Third-party multipathing - EMC PowerPath Virtual Edition

EMC provides a multipath plug-in called PowerPath Virtual Edition (PowerPath/VE) to enhance the 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 device 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 auto restore 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 a feature of VSI that simplifies the LUN path policy configuration. Figure 29 shows how administrators assign global Native Multipathing Plug-in or PowerPath path configuration preferences to VNX LUNs and maintain consistent policies across all hosts in a virtual data center.

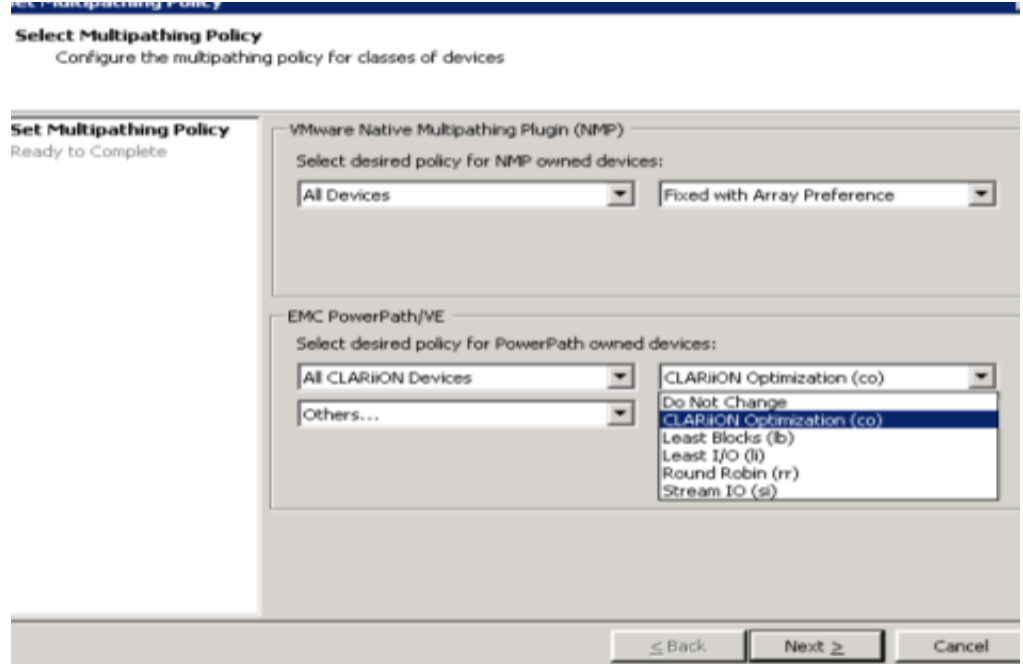


Figure 29 VSI Path Management feature

Use the VSI: Storage Viewer feature or the vCenter device properties to verify the existing multipath policy for each LUN (Datastore).

Figure 30 illustrates the LUNs properties page within Storage Viewer. This view includes the pluggable storage architecture that owns the LUN and the current PSPs. Use the VSI Path Management feature to modify these properties if required.

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:L0	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 30 Storage Viewer LUNs view

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

Multipathing considerations for 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. Elements of a multipathing configuration for NFS illustrates the recommended configuration for high availability and load balancing. Use the following guidelines to achieve high availability and load balancing for NFS:

- ◆ Ensure that no single points of failure are 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 31](#).

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.

Ensure that Link Aggregation is enabled on the switch ports, VNX Data Mover, and ESXi network interfaces. Complete the following steps in EMC Unisphere and the vSphere Client to create the multipath NFS configuration:

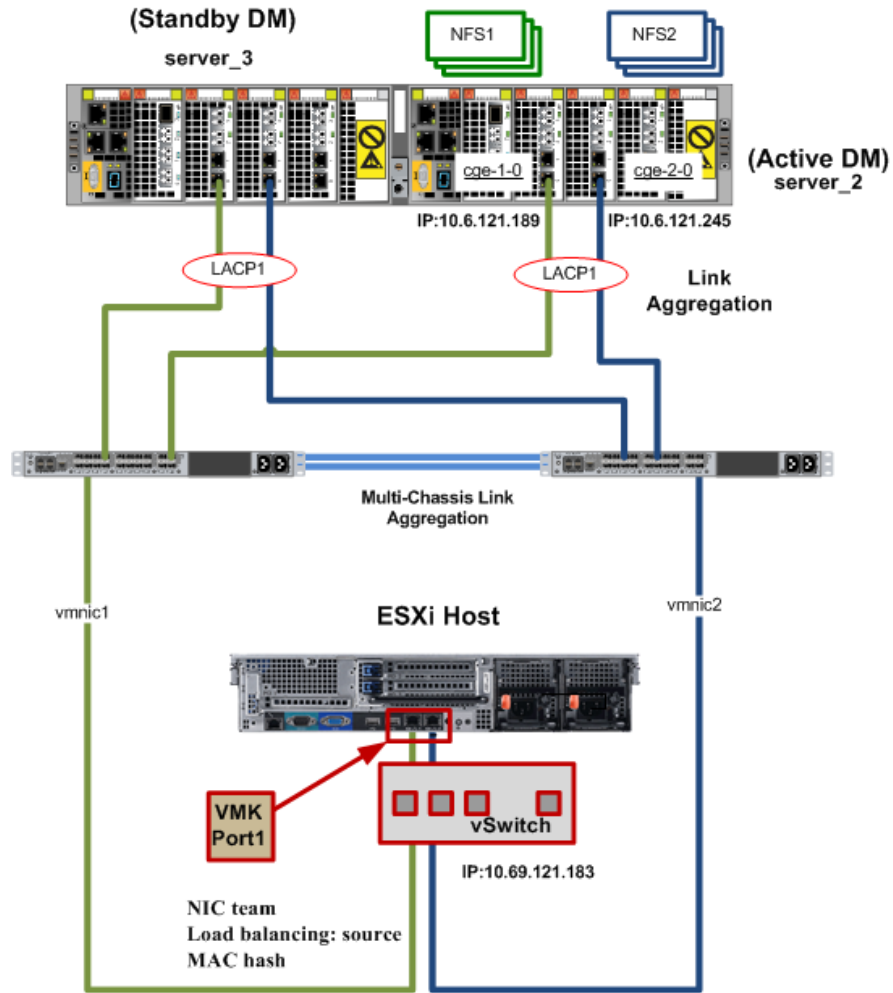


Figure 31 Elements of a multipathing configuration for NFS

1. Log in to Unisphere.
2. Select the VNX system to manage. Select **Settings > Network > Settings For File**. The **Settings For File** window appears as shown in Figure 32.

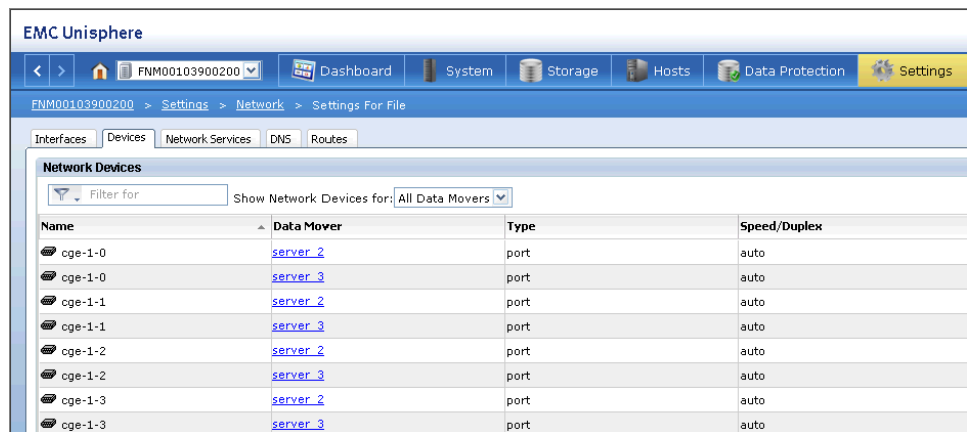


Figure 32 Unisphere interface

3. 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.
4. From the **Settings For File** window, select the **Interfaces** tab.
5. Click **Create** to create a new network interface.

Figure 33 Data Mover link aggregation for NFS server

6. Complete the following steps:
 - a. From 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 **Data Mover link aggregation for NFS server**, ensure that the IP address is set to 10.244.156.102 and the interface name is set to DM2_LACP1.
7. Click **Apply** to create the first network interface and keep the **Create Network Interface** window open.
8. 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.
9. Login to the vSphere Client and complete the following steps for each ESXi host:
 - a. Create a vSwitch for all the new NFS datastores in this configuration.
 - b. 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 **Data Mover link aggregation for NFS server**, ensure that the VMkernel IP address is set to 10.244.156.183, with physical NICs VMnic0 and VMnic1 connected to it.

- c. Click **Properties**. The vSwitch1 Properties dialog box, vSphere networking configuration, appears.

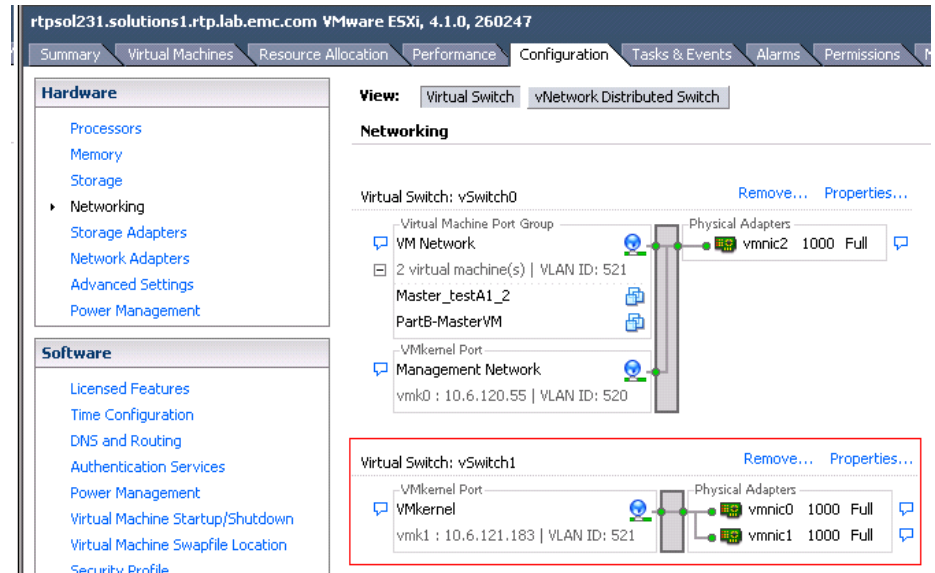


Figure 34 vSphere networking configuration

- d. Select **vSwitch**, and then click **Edit**. The **VMkernel Properties** window appears as shown in Figure 35.

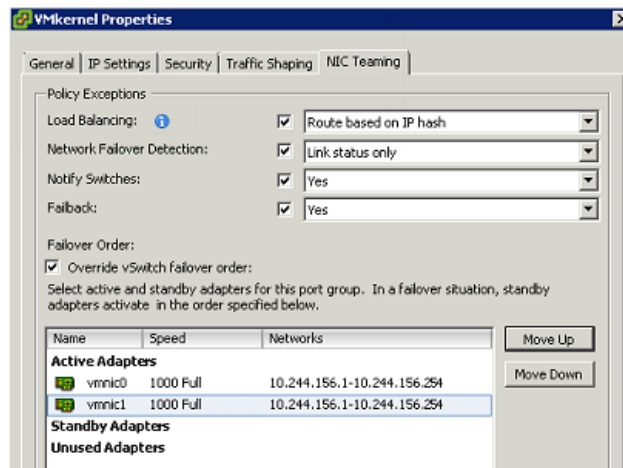


Figure 35 VMkernel Properties window

- e. 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.

- f. Use the VSI: Unified Storage Management feature to provision NFS datastores.
 - i. From the **Data Mover Name** list box, select the primary Data Mover.
 - ii. From the **Data Mover Interface** list box, select the IP address of the first network interface that was created.
- g. Create and distribute the virtual machines evenly across datastores, or use SDRS to automate the virtual machine placement within the datastore cluster.

vSphere storage considerations

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, as indicated.

Dead space reclamation (Unmap)

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

Unmap works with VMFS datastores provisioned using VNX thin LUNs.

Most environments are dynamic: 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 are still 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 return the unused space to the global storage pool so that it can be used by other datastores or RDM LUNs.

In ESXi 5.1, the unmap process is initiated through an extension to the vmkfstools. Passing the `-y` argument with a numerical value instructs the ESXi host to perform an UNMAP command on the corresponding VNX device.

The unmap command is contextual, which means it must be run from the command of the ESXi host and the current directory must be within the directory path of the datastore. The numerical value provided to the `-y` argument is the percentage of allocated space that you intend to reclaim.

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 the LUN blocks that correspond to the freed blocks within the VMFS datastore. Freeing the space within the thin LUN returns the slices to the storage pool.

Note: This process does incur overhead that could impact other systems. It should be run during a maintenance window.

In [Figure 36](#), 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 it 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, approximately 2 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 [Figure 36](#), 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 Pool 0.

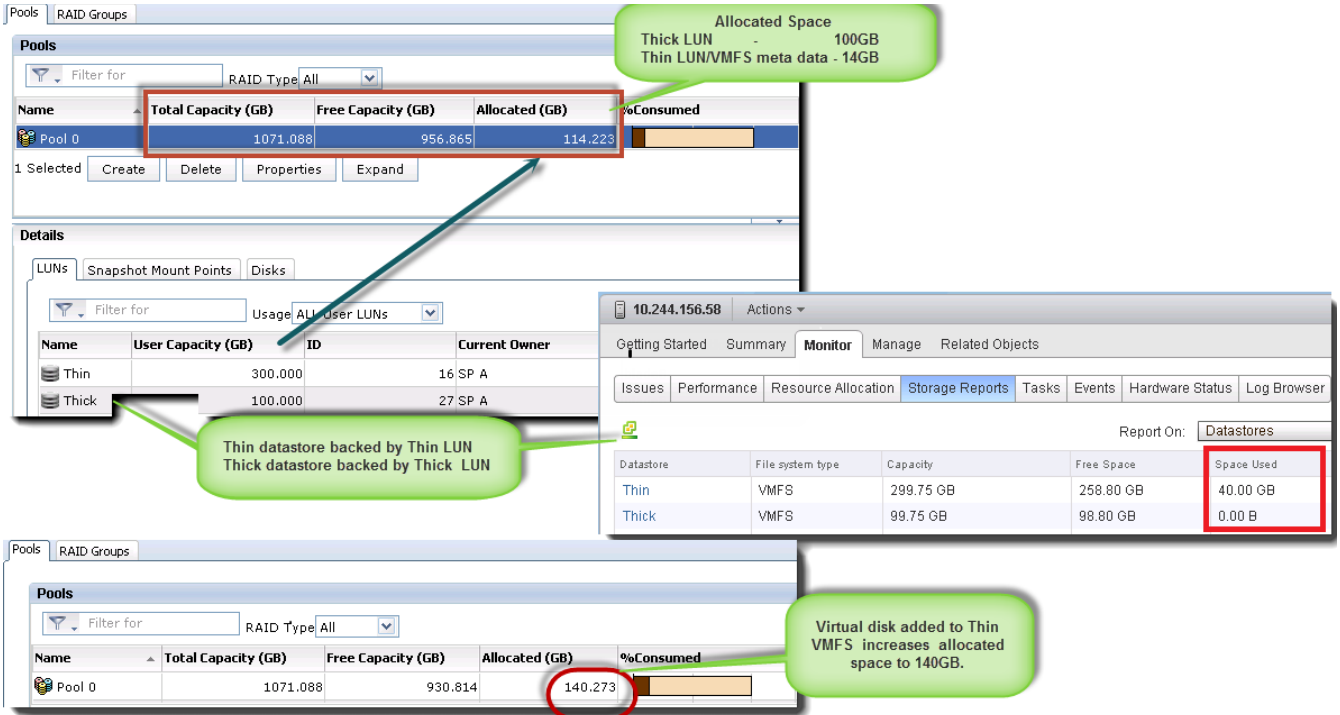


Figure 36 Virtual machine configured on a Thick LUN

The top half of [Figure 36](#) 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.

Use the **vmkfstools** command to reclaim unused space within ESXi datastores provisioned from thin LUNs. The example in [Figure 37](#) illustrates how to identify the space.

Note: The vSphere/ESXi5.5 release provides the unmap control through the esxcli utilities. The syntax for that command is
`# esxcli storage vmfs unmap -I (datastore UUID) -n (# blocks to free)`

The screenshot shows the vSphere Storage Reports interface. The 'Storage Reports' tab is active, displaying 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

Below this is the 'Pools' section, showing a table for storage pools:

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

A terminal window shows the execution of the `vmkfstools` command to reclaim space:

```

~ # cd /vmfs/volumes/Thin
/vmfs/volumes/50a27784-615b1f80-81f3-00219bcad60c # 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 .vmfsBallooniHvumJ of size 295.8 GB to reclaim free blocks.
Done.
/vmfs/volumes/50a27784-615b1f80-81f3-00219bcad60c #
    
```

Finally, the 'Pools' section is shown again, indicating that the pool consumption is back to the original size:

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

Figure 37 Virtual machine migrated to a thin LUN

VMFS-5

In vSphere 5, VMware introduced a new version of the VMware file system. 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/Hardware Accelerated Locking) is enabled for all SCSI devices.
 - ESXi hosts always attempt to use ATS on VMFS-5 volumes.
 - If an ATS operation fails, the LUN processes the request using SCSI-2 commands. Subsequent requests revert to ATS.
 - Small block allocation configurations.

Datstores created with vSphere 5 use the VMFS-5 format; however, VMFS-3 is still a supported option. Although the preceding list may not seem extensive, VMFS-5 volumes should be used for all new datstores.

An option to perform an upgrade from VMFS-3 to VMFS-5 through vmkfstools is provided; however, upgraded VMFS-3 file systems do not take advantage of all VMFS-5 features.

The limitations of upgraded VMFS-3 datstores are that they:

- ◆ Use the VMFS-3 block sizes, which might 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 datstores.

Instead of upgrading from VMFS-3 to VMFS-5, 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 storage integration improves ESXi host resource utilization by off loading storage-related tasks to the VNX. The storage system performs the select storage tasks, freeing host resources for application processing and other tasks.

Storage vMotion is a core feature of SDRS in vSphere 5 that takes advantage of the capabilities of VAAI. During a Storage vMotion task, the ESXi host issues SCSI extended copy (XCOPY) commands that include the source and destination block address information required for the VNX to perform the block copies. The commands include the information required for the VNX to copy the data blocks from the source LUN to the target device.

With VNX OE for Block version 5.32 and VNX, VAAI is executed in less time and consumes significantly fewer CPU, memory, and SAN I/O resources than when the task is executed on the host.

The primary VAAI functions are:

- ◆ **Hardware Accelerated Zeroing** (Block Zero)—Uses SCSI WRITE SAME commands to initialize blocks on a virtual disk. It is used to zero out newly created virtual disks during new virtual disk or virtual disk clone tasks. When a thick eager zeroed virtual disk is created, this primitive feature creates the virtual disk file and rapidly initializes all of the blocks.
- ◆ **Hardware Accelerated Locking** (Atomic Test and Set [ATS])—Alleviates VMFS contention resulting from metadata operations such as virtual disk creation and virtual machine boot. 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 (Full Copy)**—Uses SCSI XCOPY commands to perform block movements within the array. The primitive is initiated by vSphere Clone, Storage vMotion (SDRS), and Deploy Virtual Machine from Template tasks.
- ◆ **NFS Clone Offload**—Off loads 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 versions provide support for VAAI operations on NFS datastores. Working in conjunction with storage vendors such as EMC, VMware has integrated VAAI with VNX through a software interface installed on each ESXi host. With this software or host plug-in installed, ESXi hosts can use the VNX Data Mover to perform the tasks listed in NFS VAAI features.

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

[Table 4](#) provides a list of the NFS VAAI features.

Table 4 NFS VAAI features

Feature	VNX OE
Full clone	7.31 and later
Extended stats	7.31 and later
Space reservation	7.31 and later
Snap of snap (VMware 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 replicas and thick full-clone replicas of virtual machines on the NFS datastores. Thin clones are created instantaneously using a few file system blocks that reference the original virtual machine data. 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 that allocates additional blocks within the file system as needed.

When creating full clones, the VNX Data Mover/X-Blade creates an exact replica of an existing virtual machine using an identical amount of storage as the source virtual machine.

Nested clones (snap-of snap)

VNX File OE version 7.1 and later versions include a VAAI capability to create a thin-clone virtual machine from an existing thin clone virtual machine. This functionality is referred to as nested clones or snap-of-snap. The functionality uses the VNX snapshot architecture to instantaneously create lightweight virtual machine clones. VMware products such as View Composer and vCloud Director use View Composer Array Integration (VCAI) to initiate virtual machine clones using the VNX Data Mover API for nested snapshots. The clone operations are off loaded to the VNX Data Mover. To take advantage of this feature, enable the nested clone support property when you create the file system.

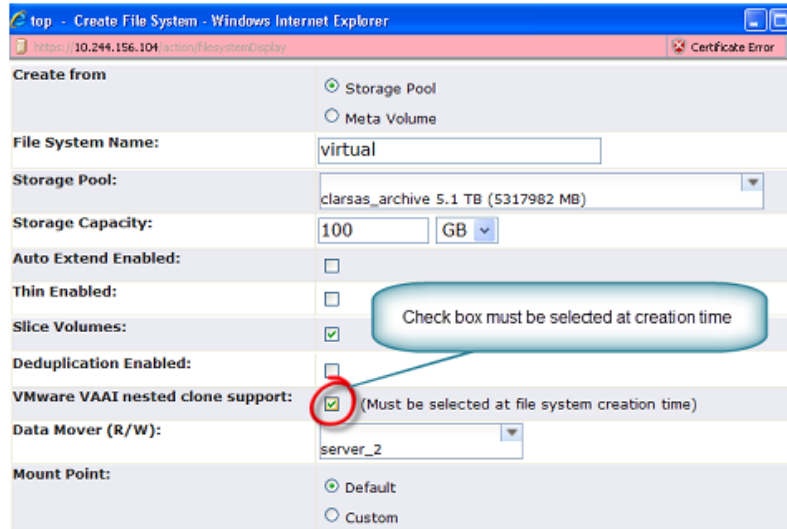


Figure 38 Create File System

EMC NAS plug-in installation

EMC provides a software module or plug-in to enable VNX NFS VAAI functionality on ESXi hosts. The software is provided as a VMware installation bundle (VIB) that is installed through the command line of the ESXi host or through VMware vCenter Update Manager.

Figure 39 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  [2]

```

Figure 39 Plug-in installation

vCenter displays the Hardware Accelerate property of the NFS datastores in the Datastores tab, illustrated in [Figure 40](#).

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	VMFS5	Supported	Enabled
VNX958-4	Normal	DGC Fibre Channel Disk (naa.6006...	Non-SSD	329.75 GB	120.79 GB	VMFS5	Supported	Enabled
VNX958-3	Normal	DGC Fibre Channel Disk (naa.6006...	Non-SSD	329.75 GB	184.71 GB	VMFS5	Supported	Enabled
datastore1	Normal	Local SEAGATE Disk (naa.5000c50...	Non-SSD	63.25 GB	62.29 GB	VMFS5	Unknown	Enabled

Figure 40 NFS Hardware Accelerate datastore property

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

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

Vmkfstools extended stats

vSphere 5 includes additional vmkfstools command-line arguments to display the disk utilization for virtual machine disks configured on NFS datastores. 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 argument reports all values in bytes, as shown in [Figure 41](#). This helps when creating automated reports or custom provisioning scripts.

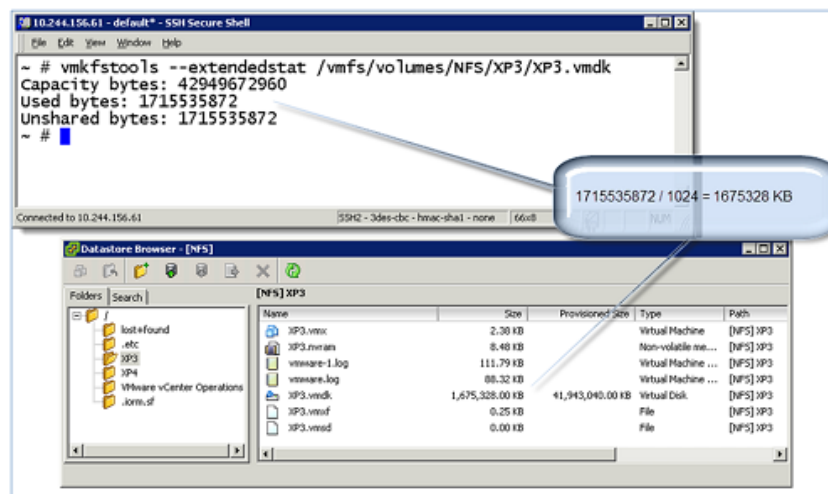


Figure 41 Vmkfstools disk utilization option

Storage Distributed Resource Scheduler

SDRS is a vSphere 5 feature that provides the ability to group multiple datastores into a single managed storage object called a datastore cluster. The cluster is integrated into vCenter management storage devices as a storage abstraction with some of the same placement and migration capabilities that make SDRS so valuable for CPU and memory. It allows VMware administrators to manage storage as a pool of resources or datastore cluster instead of a series of individual datastores.

SDRS relies on a new storage object called a datastore cluster. These clusters consist of multiple VMFS or NFS datastores, as shown in [Figure 42](#).

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

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

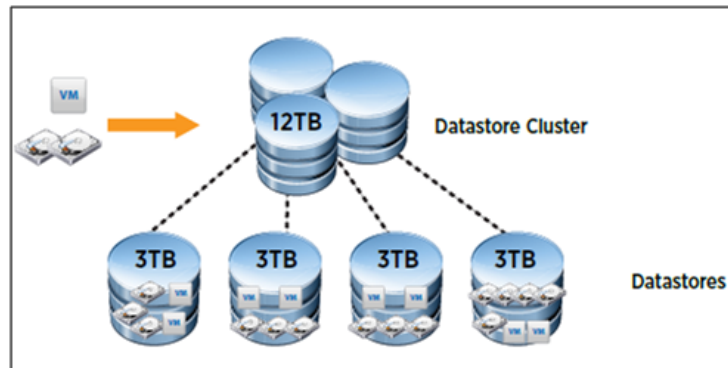


Figure 42 SDRS datastore cluster

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

SDRS is not meant to be a highly reactive solution. It can be tuned for aggressive relocations, but the default relocation policy requires 8 to 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 work together; however, do not use FAST VP relocation when SDRS I/O metrics are in use. Disable I/O metrics on FAST VP LUNs.

SDRS policy configuration

SDRS provides the following automation policies, as shown in [Figure 43](#):

- ◆ **Fully Automated**—Performs initial placements and virtual machine relocation without user intervention
- ◆ **No Automation**—Presents a recommendation each time a virtual machine relocation is triggered

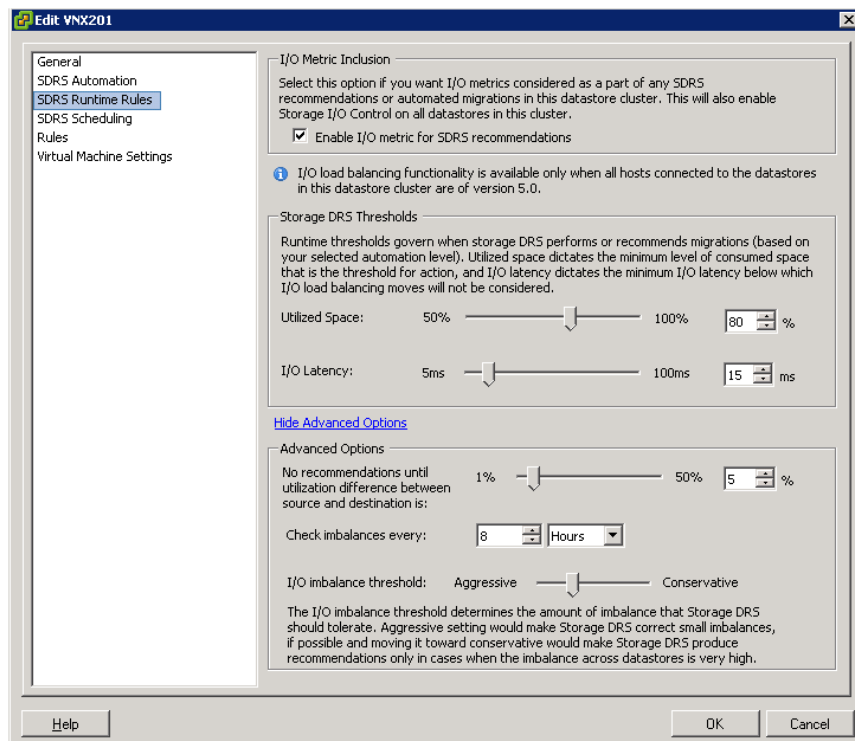


Figure 43 SDRS advanced policy configuration

Policy metrics are:

- ◆ **Utilized Space**—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**—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**—Interval for applying the SDRS 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. SDRS and FAST VP perform a similar function, although at a different level of granularity, to rebalance resources. Use either SDRS or FAST VP for workload rebalancing across storage resources. Do not use both services at the same time.

Figure 44 shows the interface to disable I/O metrics and apply policy based on capacity utilization. Clear the check box for enabling I/O metrics for SDRS recommendations.

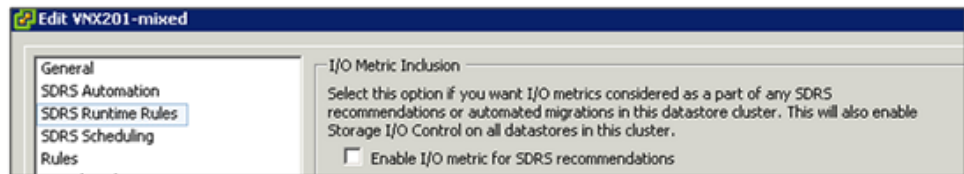


Figure 44 SDRS I/O metric enablement setting

Notes:

- ◆ SDRS I/O load balance does not work if the datastore cluster is not configured for all hosts that share the datastores.
 - ◆ VAAI operations do not span storage systems. The host processes migrations of virtual machines between clustered datastores from different storage systems.
-

VNX storage recommendations for SDRS

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

When 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 an 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 results in their sharing 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 5 shows supported SDRS LUN configurations.

Table 5 Supported SDRS LUN configurations

VNX feature	Initial placement	Migration recommendation
Thin, Thick, FLARE LUN	X	X
FAST VP	X	No, manual mode
FAST Cache	X	No, manual mode
Replication	X	No
LUN sanpshots	X	No
Dedupe	X	No
Thin	X	Supported with VASA

vStorage API for Storage Awareness

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 45 illustrates the storage capabilities of a datastore cluster using SAS drives with Fast Cache enabled for the LUN.

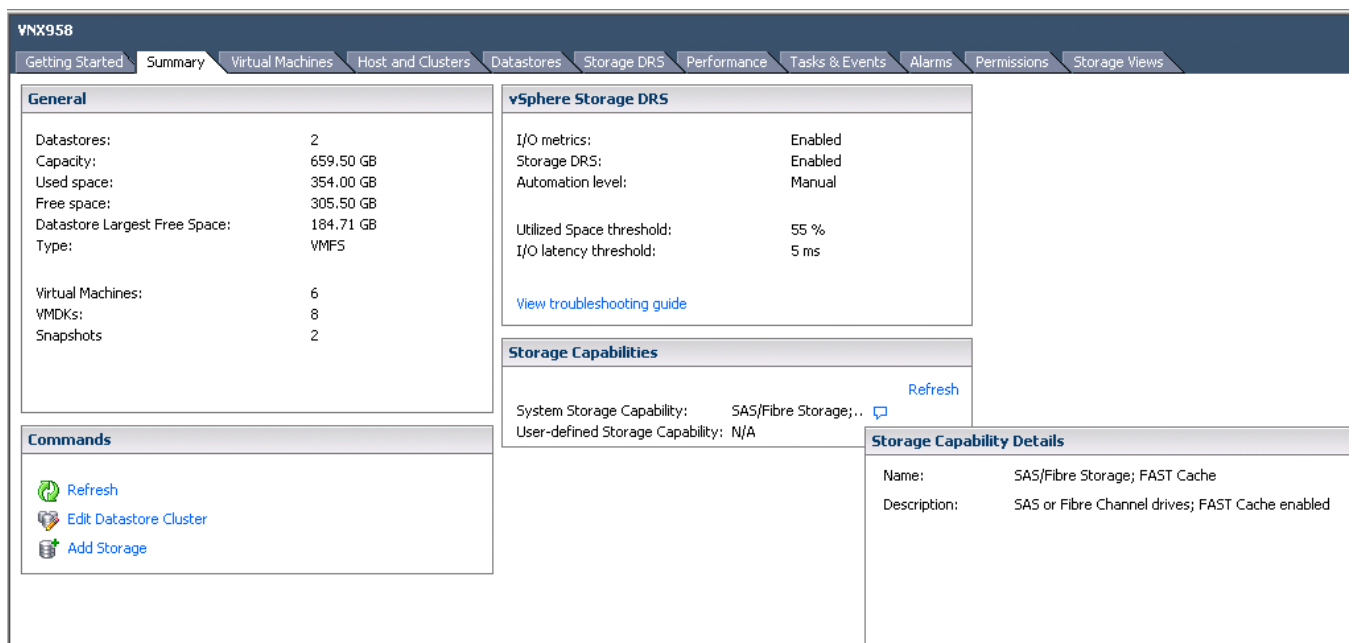


Figure 45 VASA datastore storage capability of VNX Flash drive LUN

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 use storage capabilities to identify appropriate datastores for Storage vMotion operations.

VNX OE for Block provides native VASA support in versions 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. Support for VNX File OE was unavailable prior to release 7.1. [Table 6](#) lists capabilities for the initial implementation.

Table 6 VASA storage capability mapping to VNX LUNs

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

The VASA service is provided 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.

Datastore capabilities are reported based on the device type used to create the datastore. The disk properties are listed in column 1 of [Table 7](#) 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 VASA datastore storage capability of VNX Flash drive LUN 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 7 shows that a thin pool LUN with mixed drive types has a storage capability of Automated Storage Tiering, Thin Pool LUN.

Table 7 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
VNX Block Provider 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
VNX File Provider 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 46 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.

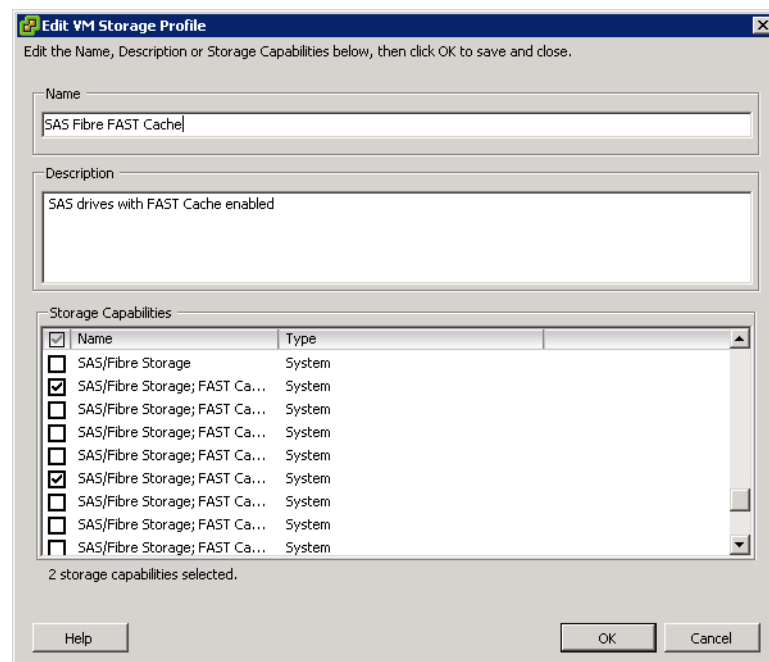


Figure 46 Storage profile assignment

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 47](#), 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.

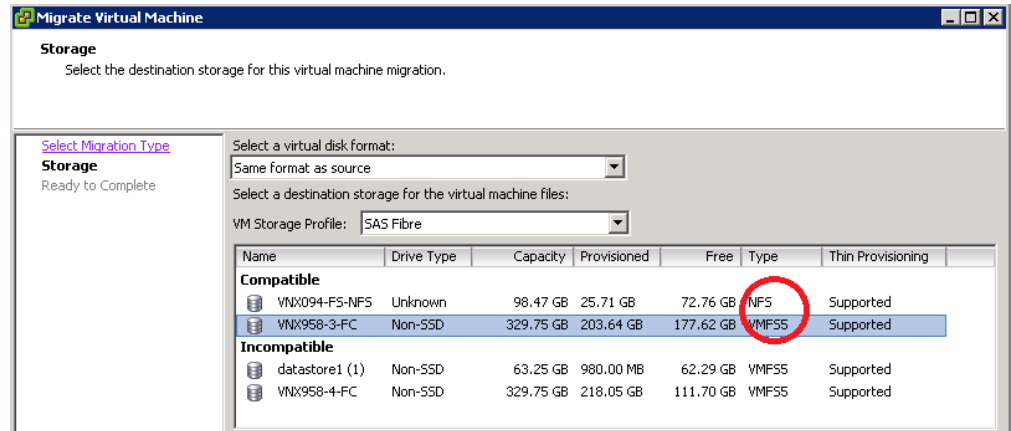


Figure 47 Compatibility/incompatibility with SAS Fibre storage profile

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 might not work correctly.

User-defined storage capabilities

In some cases, VASA does not have a profile that matches the properties of a datastore or a need exists 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 do not support NFS datastores. Create a user-defined profile to use storage profiles.

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.

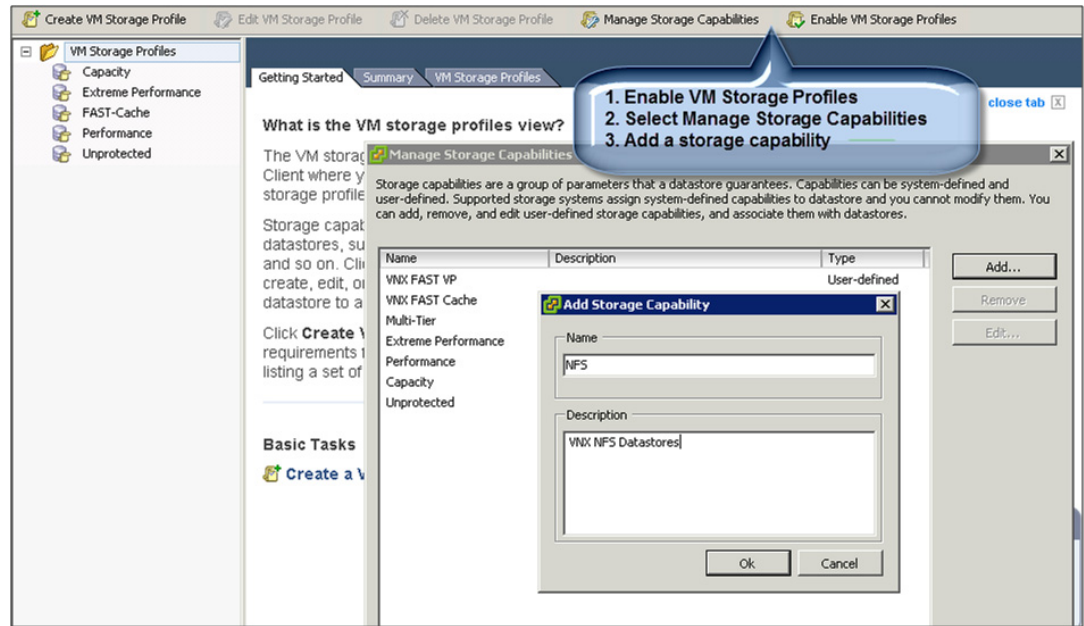
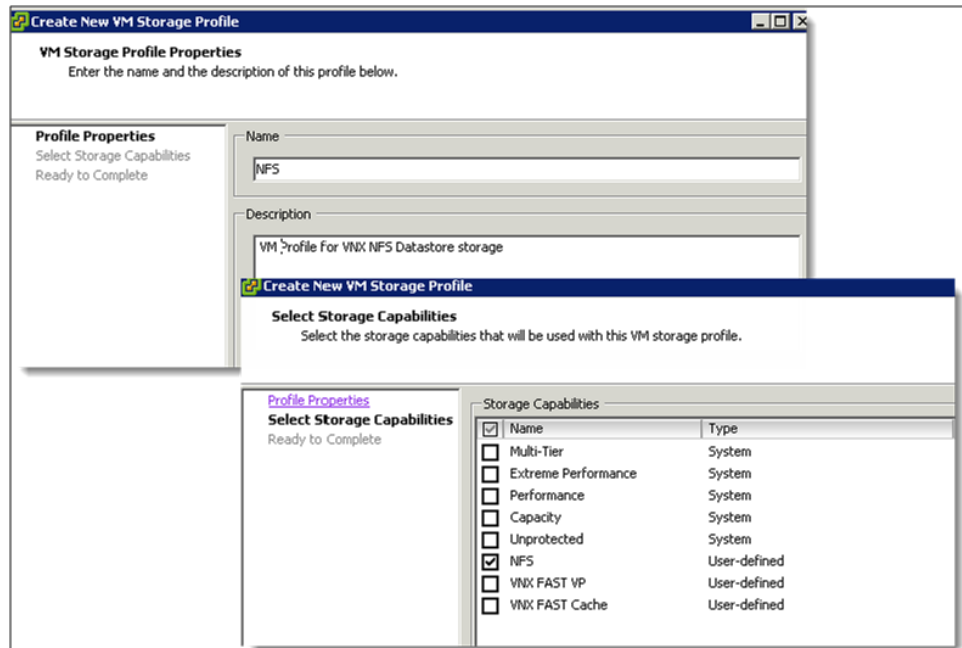


Figure 48 Enabling virtual machine storage profiles

3. Add a virtual machine storage profile. The 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.

- Assign the new profile to existing datastores, as shown in Figure 49. and Figure 50.

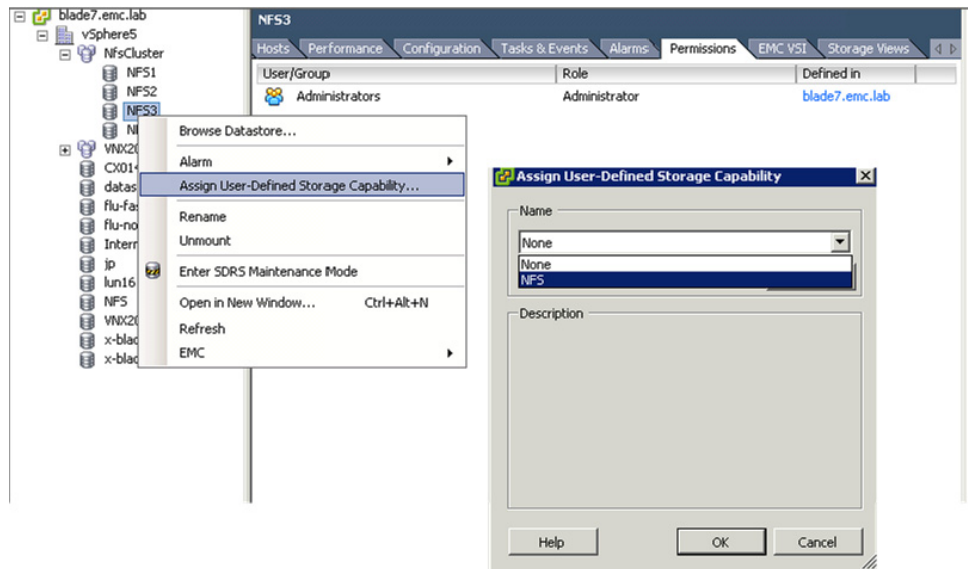


Figure 49 Associating datastores with a user-defined storage profile

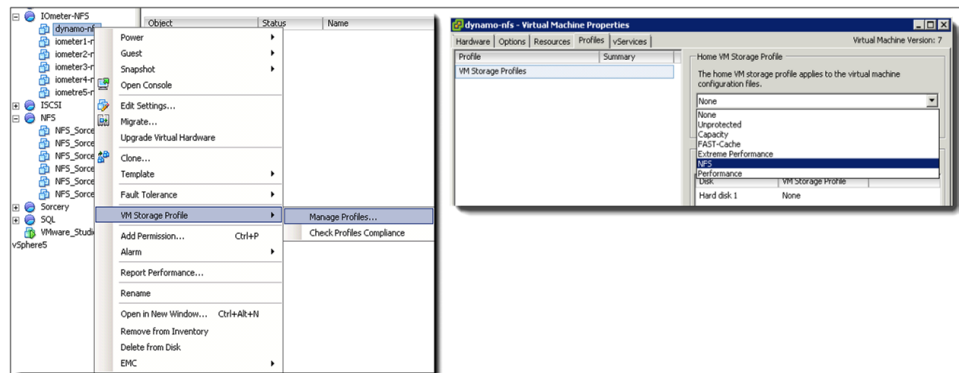


Figure 50 Associating the virtual machine with a user defined storage capability

- Associate the virtual machine virtual disks with this profile to ensure compliance.
- 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.

To configure the vCenter VASA service:

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

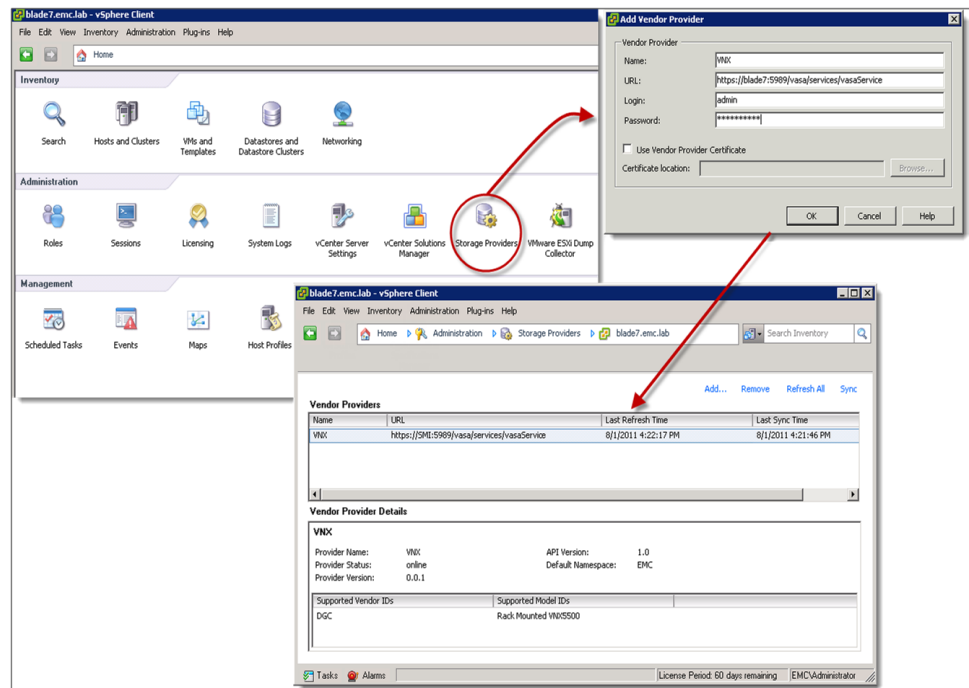


Figure 51 VASA configuration

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.
3. Provide the following information 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 version 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.

- ◆ SP 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 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, with settings of low (500) and high (2,000) share values. 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.

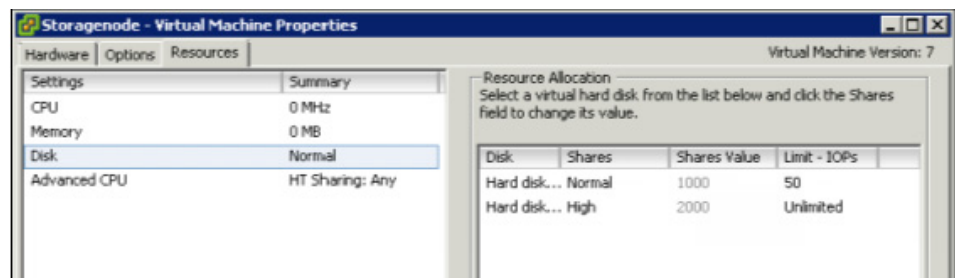


Figure 52 Virtual disk shares configuration

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. 30 ms is the default value.

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

- ◆ 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 8 lists the recommendations for setting the congestion threshold.

Table 8 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 workloads on a shared storage device that are external to VMware. 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 5.0 and later versions 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.

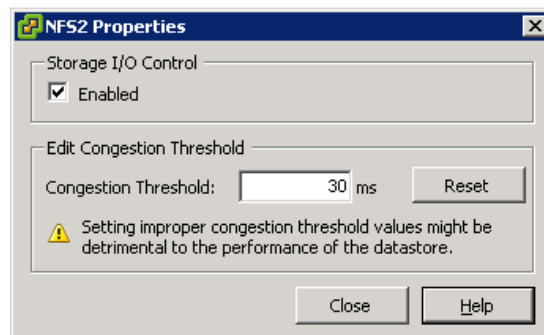


Figure 53 NFS SIOC congestion window

Notes:

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

Network considerations

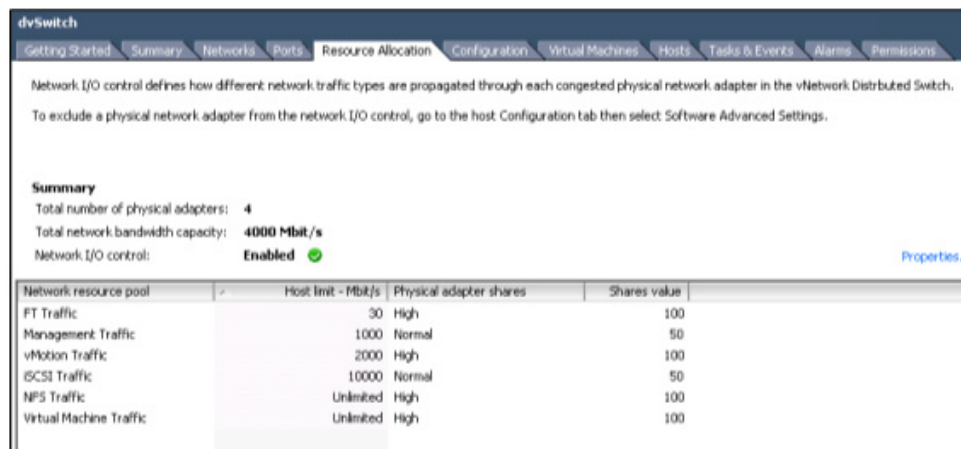
The VNX platform supports a wide range of network topologies and capabilities for VMware vSphere. This section lists items to consider when planning an IP storage network for vSphere servers.

Network I/O Control

vSphere Network I/O Control (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 4.1 and later versions.

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.



The screenshot shows the vSphere Network Resource Allocation interface for a vNetwork Distributed Switch. It includes a summary section and a table of network resource pools.

Summary

- Total number of physical adapters: 4
- Total network bandwidth capacity: 4000 Mbit/s
- Network I/O control: Enabled

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 such as a switch failure, or an administrative action such as removing a LUN from a storage group. In ESXi releases prior to ESXi 5, the host could not differentiate between these two states.

In ESXi 5 and later versions, 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 targets 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](#).

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.

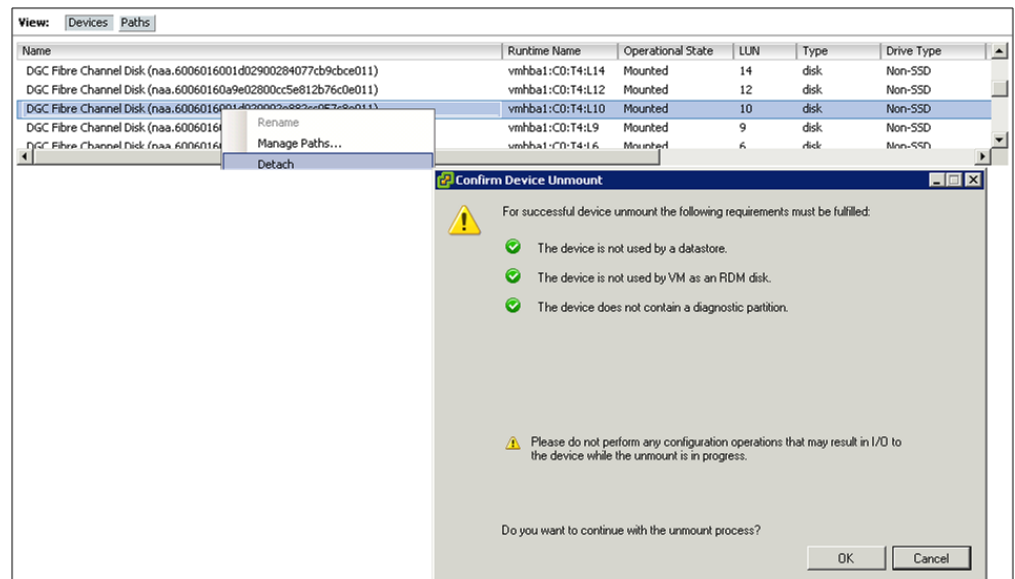


Figure 55 vSphere 5 Datastore removal wizard

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 partition 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 can 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 VSI Unified Storage Management.
- ◆ The benefits of aligning boot partitions are generally marginal. If only a single virtual disk exists, 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 unit size allocation.
- ◆ EMC also provides a free tool called UBERAlign that identifies and corrects misaligned virtual disks. [Everything VMware At EMC](#) provides more information on this tool.

Align virtual machine disk partitions

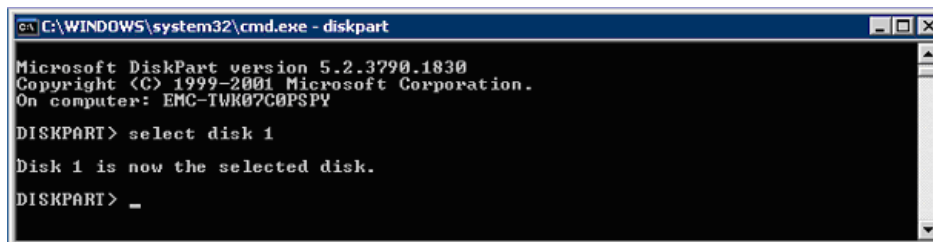
The disk partition alignment within virtual machines is affected by a long-standing issue with the x86 processor storage configuration. As a result, external storage devices are not always aligned in an optimal manner. This is true for vSphere environments, in most cases. The following examples illustrate 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 as follows; the 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-TWK07C0PSPY

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](#).

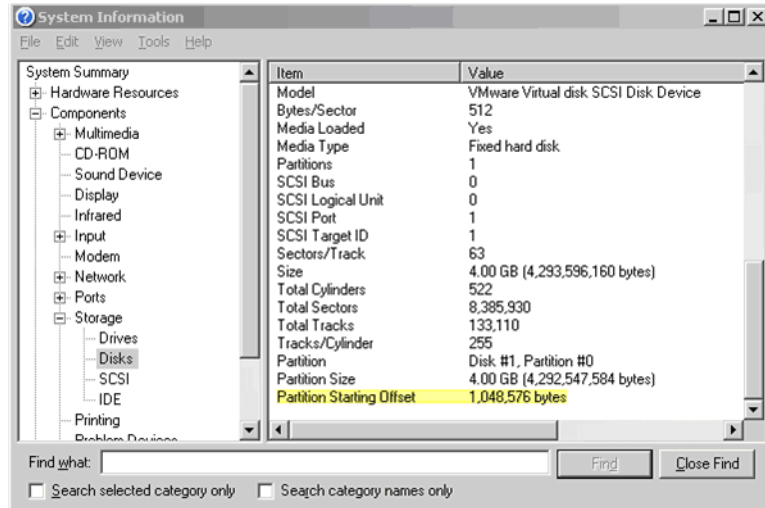


Figure 57 Guest disk alignment validation

2. Locate the **Partition Starting Offset** property and verify the value is **1,048,576** bytes as shown in [Figure 58](#).

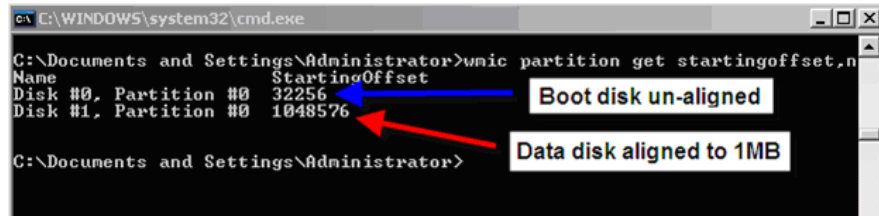


Figure 58 NTFS data partition alignment (wmic command)

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.

Partition allocation unit size

Run the **fsutil** command to identify the allocation unit size of an existing data partition.

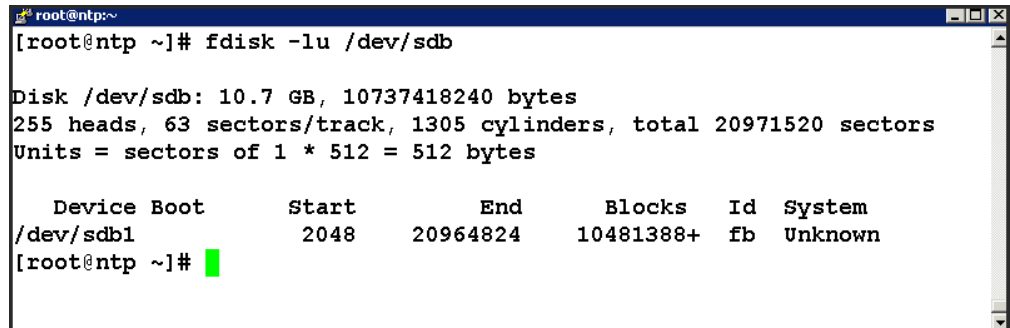
At the command prompt, type **fsutil fsinfo ntfsinfo <drive_letter>**.

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 [Figure 59](#), /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:~# fdisk -lu /dev/sdb

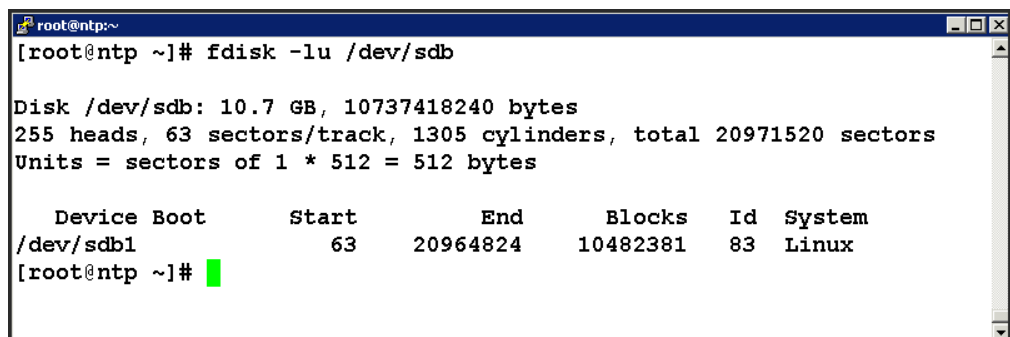
Disk /dev/sdb: 10.7 GB, 10737418240 bytes
255 heads, 63 sectors/track, 1305 cylinders, total 20971520 sectors
Units = sectors of 1 * 512 = 512 bytes

   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, as shown in [Figure 60](#).



```

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 a lot of concurrent I/O 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 off load 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 use 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 situation. 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 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.

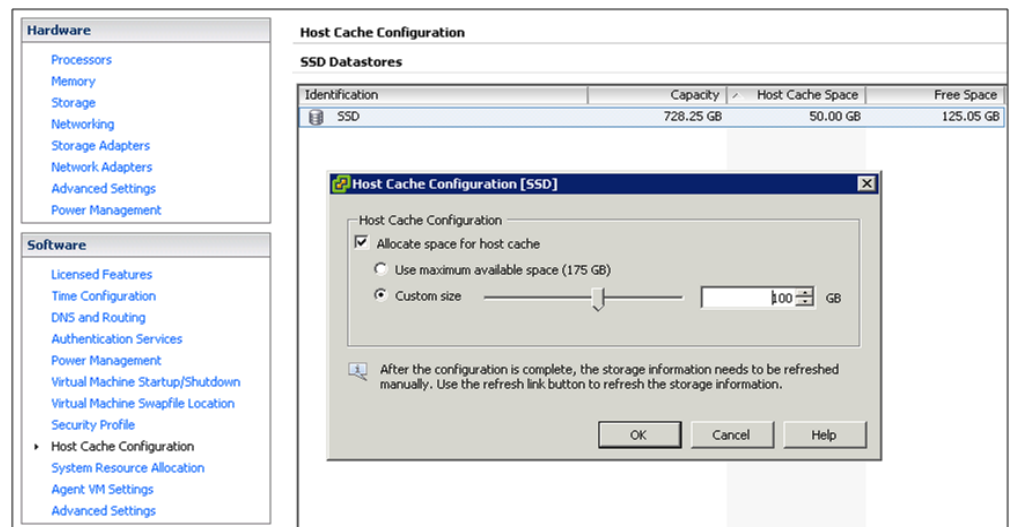


Figure 61 Host Cache configuration on VNX EFD storage

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. It also allows a storage administrator to provision a storage directory to the virtual adapter on the VM bypassing the hypervisor storage stack.

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.

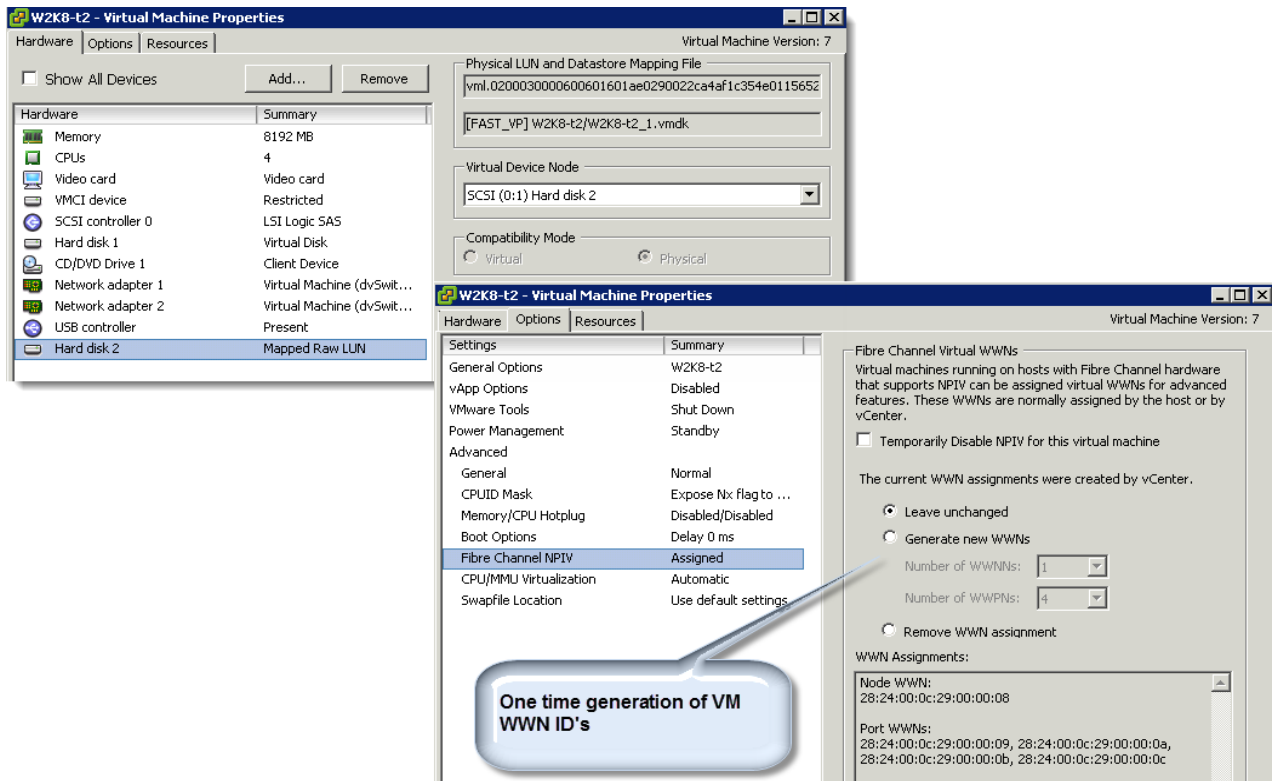


Figure 62 Enable NPIV for a virtual machine after adding an RDM volume

The virtual WWNs must be zoned in order to create a storage group for the virtual machine on the VNX. If the virtual WWNs are not listed in the switch utilities, manually add them with an alias for the virtual machine, 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.

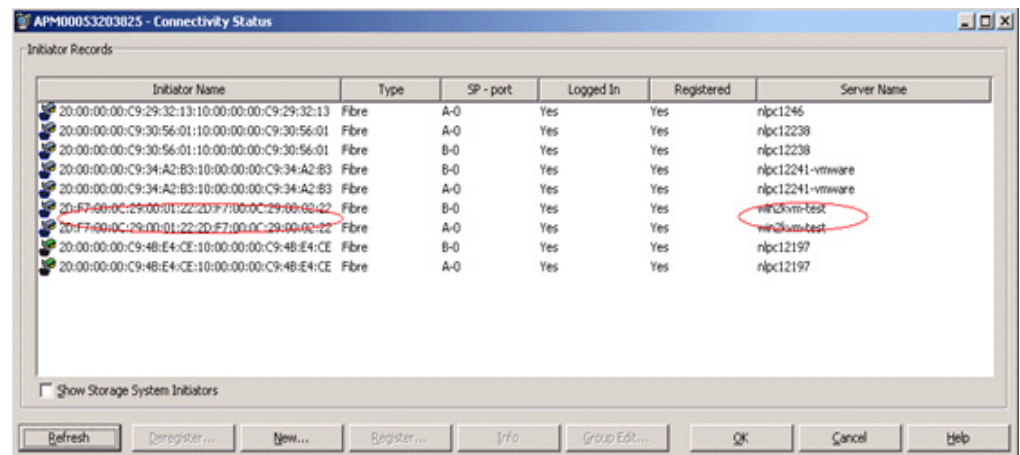


Figure 63 Manually register virtual machine (virtual WWN) initiator records

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:
 - a. Mask the LUNs to the ESXi hosts and the virtual machine storage group.
 - b. Assign the LUNs the same host LUN number (HLU) as the ESXi hosts.
 - c. 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 crashes.

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 OS to keep the application and virtual machines available during VNX Data Mover outage events and avoid downtime:

- ◆ 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
```

Monitoring and managing 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 the Storage Viewer feature of VSI. 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 25](#), Storage Viewer exposes the datastore and VNX storage details. Use the information presented in [“VSI: Storage Viewer” on page 27](#) to configure VNX file system and LUN monitoring through Unisphere.

Monitoring datastores using vCenter

Use 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](#), 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.

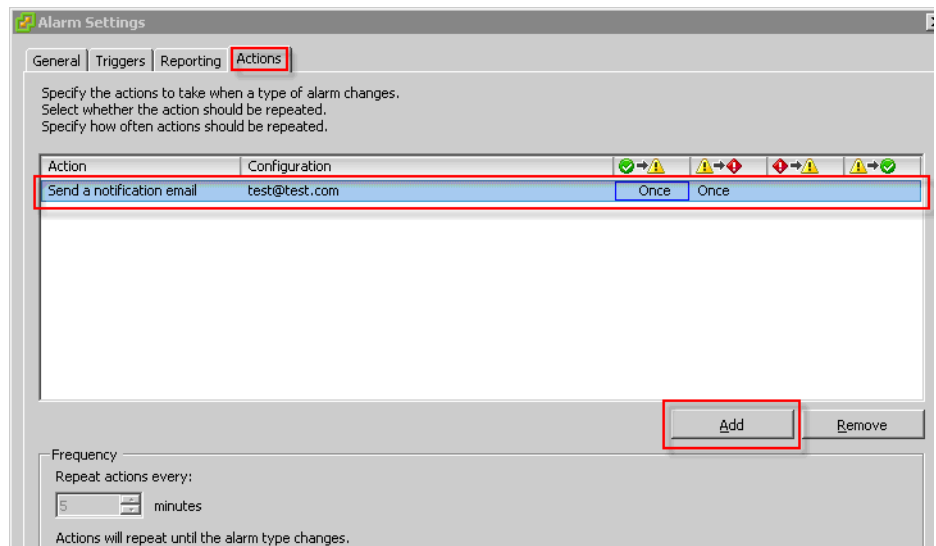


Figure 64 Data Alarm Settings-Actions window

When VNX Thin Provisioning is in use, you should correlate the storage information presented in vCenter with the storage utilization from the storage array. 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: Datastores, LUNs, and Targets.

3. Click **Datastores**.

The **Storage Viewer\Datastores** information is displayed.

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 displays **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 following data:

Datastores							Refresh	
Identification	Status	Device	Capacity	Free	Type			
vnx-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.600601601ae02900de195124a54ae011):1	99.75 GB	99.20 GB	vmfs3			
vnx-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	240.41 GB	55.03 GB	NFS			

Storage Details													Export...		Total LUNs: 1
Runtime Name	Model	Array	Device Name	Device ID	RAID	Capaci...	Group	META	VP	Paths	Owner	Policy			
vmhba4:C0:T4:L7	VNX5500	FNM00103900200	LUN 16	00016	RAID_5	100.00 GB	ESXCluster_15_16		Yes	2/2	PowerPath	CLAROpt			

Figure 65 Storage Viewer\Datastores 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 percent.
- ◆ 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.
6. 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.

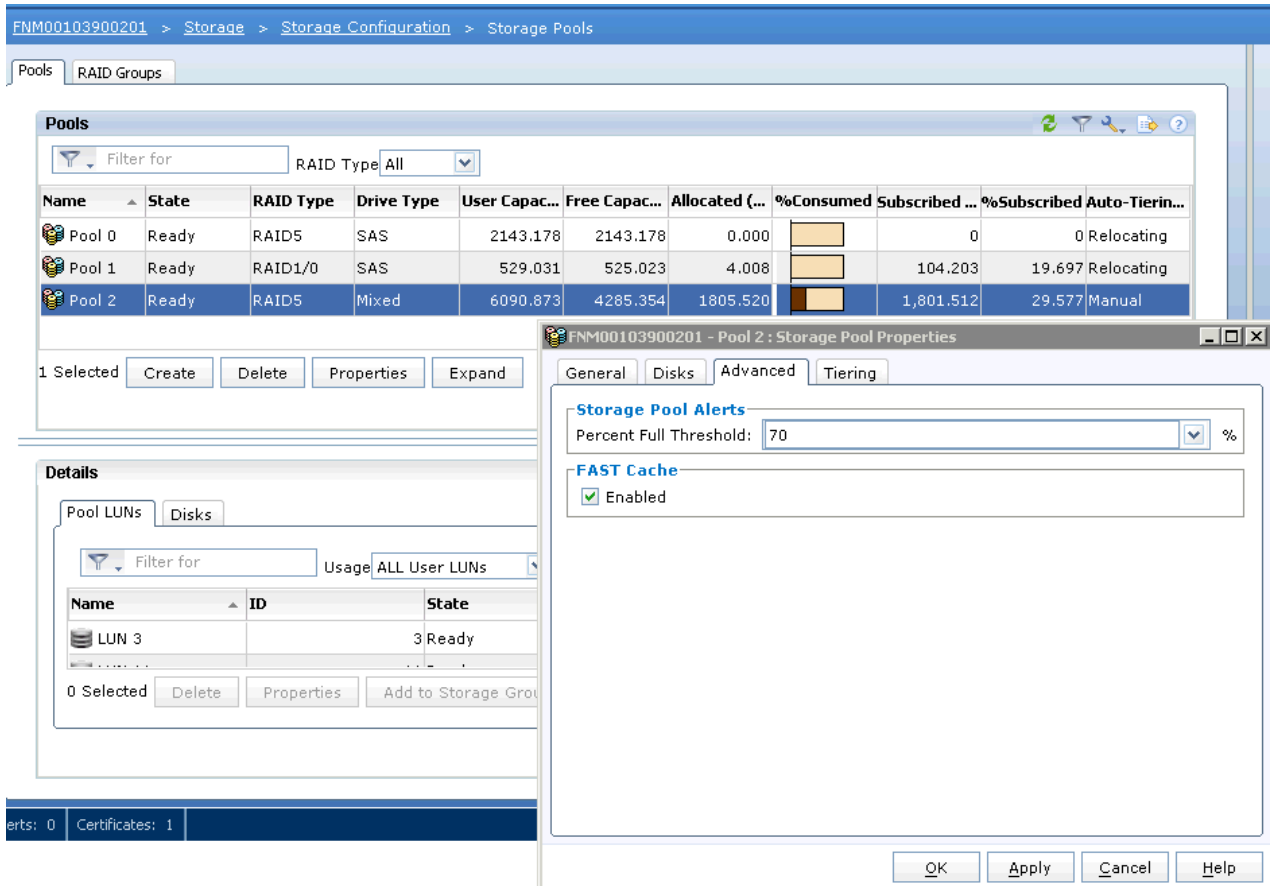


Figure 66 Adjustable percent full threshold for the storage pool

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.

Configuring 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 is displayed, appears as shown in [Figure 67](#).

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.

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 the condition is met.

- e. Type the email or SNMP address, which consists of an IP address or hostname and community name. Use commas between multiple email addresses or trap addresses.
- f. Click **OK**.

The configured notification appears in the **Storage Usage** window as shown in [Figure 68](#).

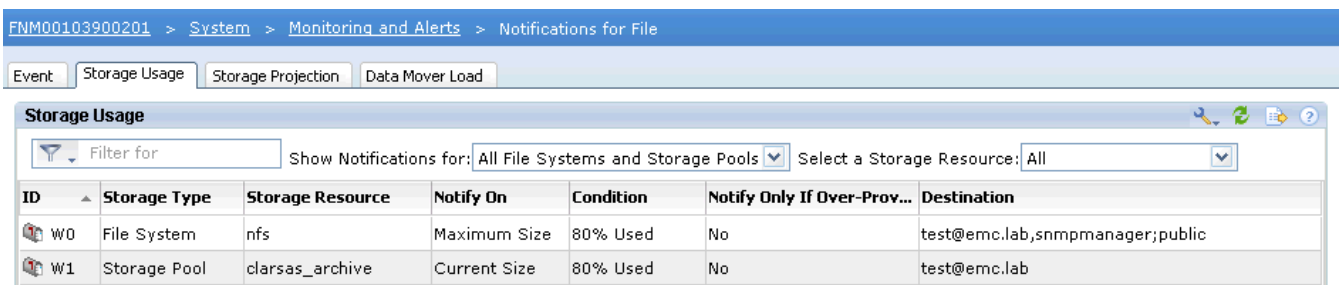


Figure 68 User-defined storage usage notifications

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](#).

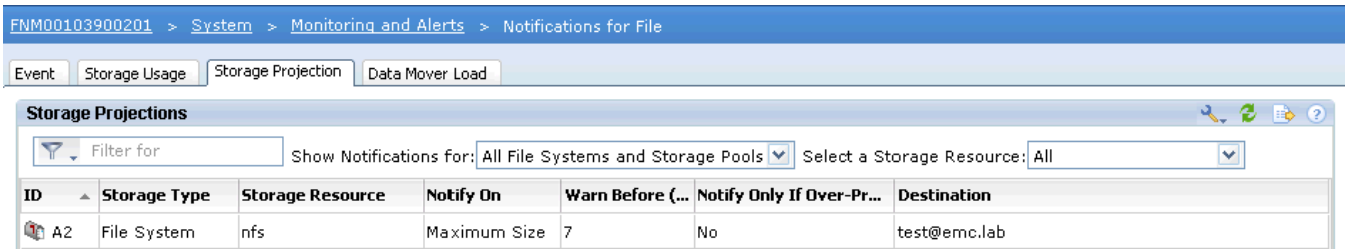


Figure 69 User-defined storage projection notifications

Note: Unisphere has no comparable capability 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, EMC has provided new options for monitoring the resource utilization of the VNX storage system.

EMC VNX Monitoring and Reporting

The EMC VNX Monitoring and Reporting product provides an extended set of tools and system views to help understand storage utilization and workload patterns. It presents resource information in a several formats, including color-coded graphical views to quickly identify and isolate storage resource constraints.

The product collects data from one or more VNX systems and stores it in a database to be used for problem diagnosis, trend analysis, and capacity planning.

VNX Monitoring and Reporting includes a web interface to view VNX storage system inventory, performance information, capacity planning metrics, and health information.

[Figure 70](#) shows an example of a capacity planning report that 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

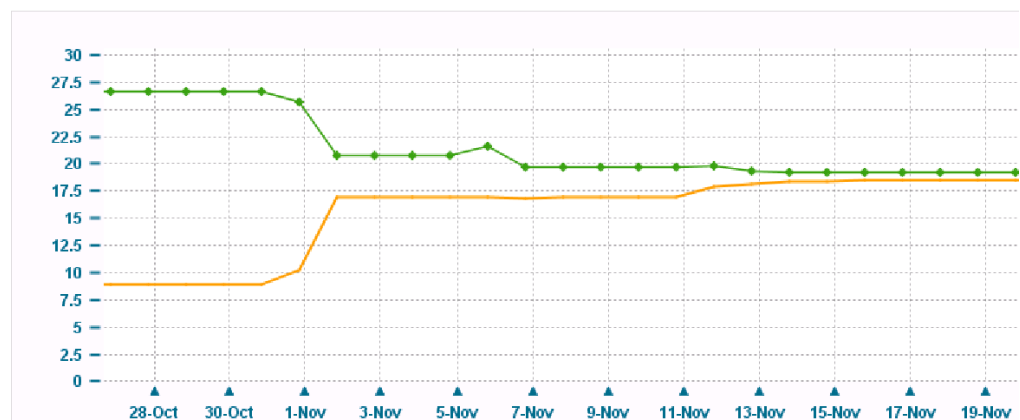


Figure 70 VNX Monitoring and Reporting: Capacity Planning Report

Figure 71 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.

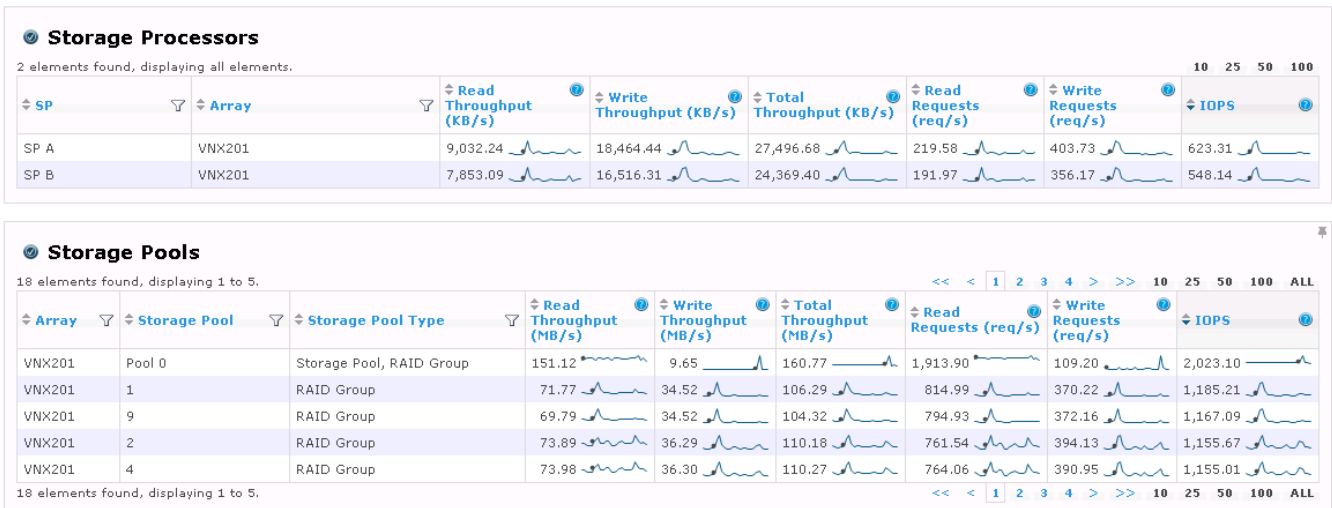


Figure 71 VNX Monitoring and Reporting - Performance report

This product 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 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 in 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 9](#).

Table 9 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, enables 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 balance resources on the storage system or vSphere environment.

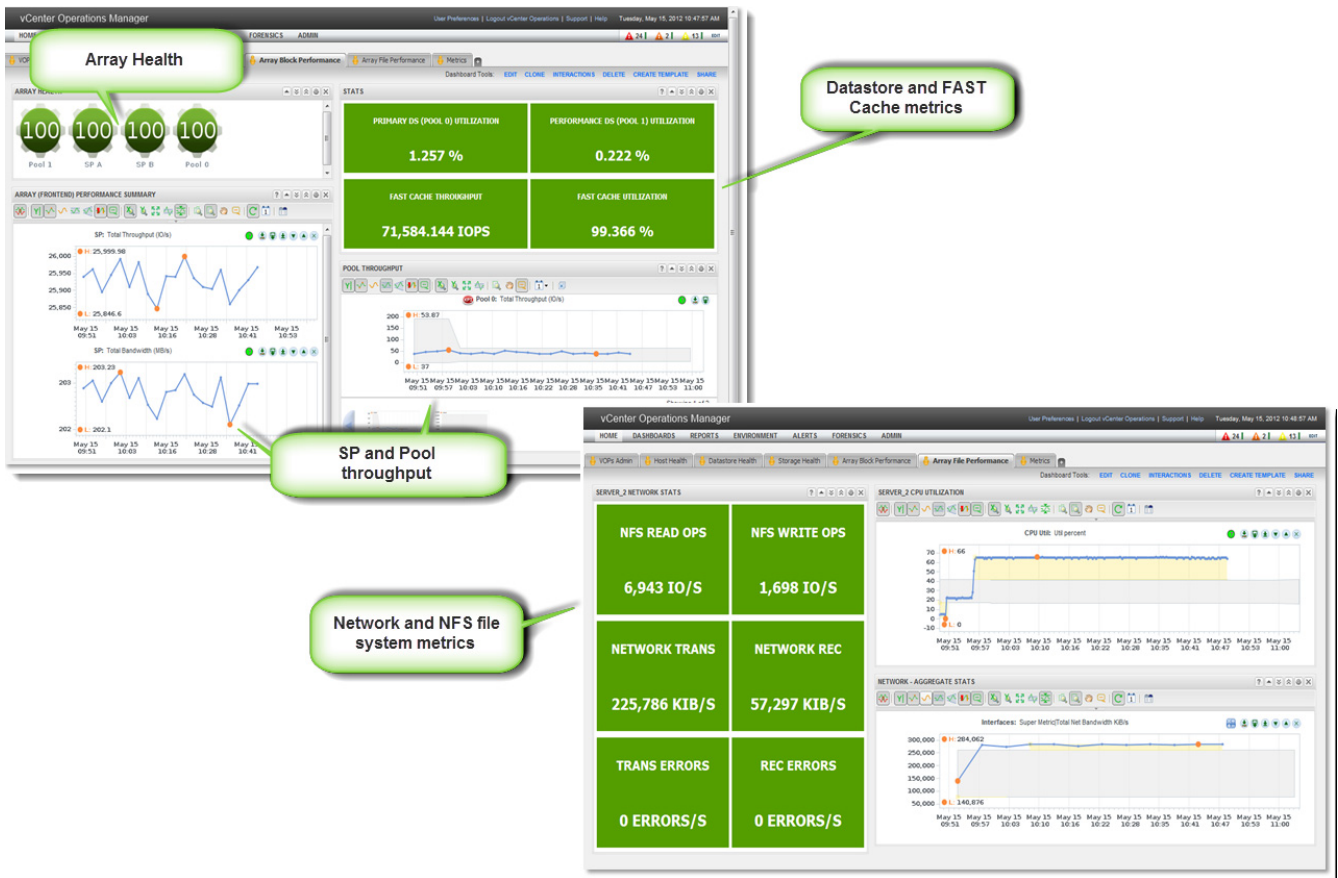


Figure 73 vCenter Operations Manager: VNX Storage Analytics

As cloud computing evolves, understanding how resources are consumed across the environment and how one resource impacts another become increasingly important. Products such as EMC Monitoring and Reporting and VNX Analytics Suite provide the information to assist you in quickly identifying the root cause of a performance problem and relate that to applications and services in the environment. 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 has renamed the virtual disk identifiers between versions. This document uses the current naming conventions. [Table 10](#) provides the reference to describe virtual disks in vSphere 4.

Table 10 Command line descriptions for vSphere 5.5, vSphere 5, and vSphere 4

vSphere 5.5	vSphere 5	vSphere 4	Command line
Thick lazy zeroed	Flat	Thick	ZeroedThick
Thick eager zeroed	Thick	Fault Tolerant	EagerZeroedThick
Thin	Thin	Thin	Thin

VMware offers the following options to provision a virtual disk: thin, flat (ZeroedThick), and thick (EagerZeroedThick). [Table 11](#) provides a description of each option, 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 11 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
Thick Lazy Zeroed/Flat (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.

Table 11 Virtual machine disk allocation policies (continued)

Allocation mechanism (virtual disk format)	VMware kernel behavior	Impact on VNX pool
Thick Eager Zeroed	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

Thinly provisioned storage for VNX is available only when using storage pools, and more specifically when using thin LUNs. Thick LUNs reserve space and classic LUNs are fully allocated on RAID groups. Thin LUNs preserve storage pool space by deferring block allocations until the ESXi host or a guest virtual machine allocates new blocks within the VMFS data store (or virtual disk). This architecture provides the ability to oversubscribe the storage pool when using thin LUNs.

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 virtual VMFS virtual disk format is thick lazy zeroed (also referred to as Flat in some releases). Lazy zeroed disks do not initialize or claim all the space during creation. The virtual disk space is reserved within the VMFS datastore, but blocks are not zeroed until they are allocated by the virtual machine.

[Figure 74](#) illustrates the creation of a 500 GB virtual disk with 100 GB of data 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. This disk format provides faster allocation time and space conservation within thin LUNs. However, blocks allocated within this virtual disk format cannot be returned to the pool after allocation.

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.

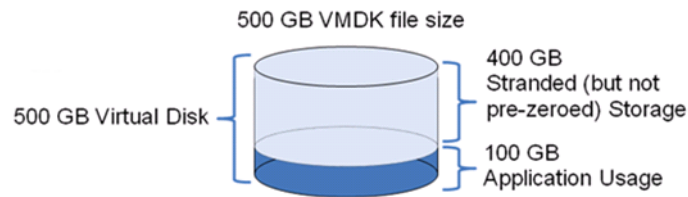


Figure 74 Thick or ZeroedThick virtual disk allocation

Use thin virtual disks to preserve space within the VMFS datastore. They consume only the space that is required or consumed by the guest OS or application. When configured on thick LUNs or datastores, thin virtual disks limit consumption of the LUN space. When thin virtual disks are configured on a thin LUN, this benefit is extended to the storage pool.

[Figure 75](#) illustrates the same 500 GB virtual disk within a VMFS volume.

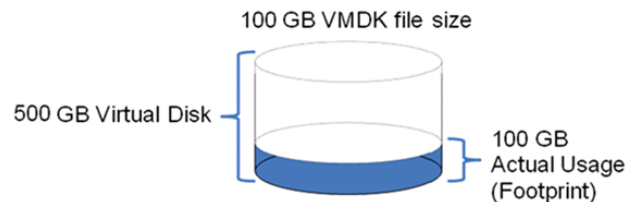


Figure 75 Thin virtual disk allocation

This time the virtual disk is created using the 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, 8k, or 32k block size used by the guest operating system of the virtual machine, ESXi performs an allocation of 1 MB, which is the default block size for VMFS-5. This allocation improves locality, which is beneficial for a thin-on-thin configuration.

The default format for virtual machine disks is Flat (ZeroedThick). [Figure 76](#) shows that you should select one of the other options if you need a thick or thin disk.

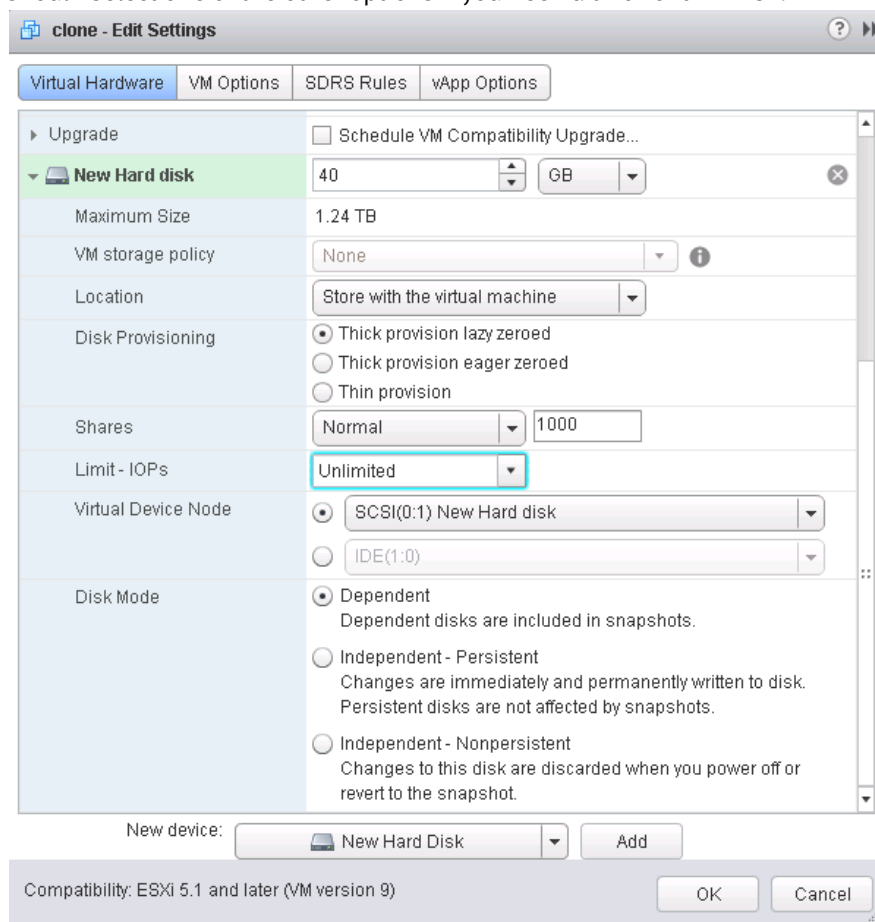


Figure 76 Virtual machine disk creation wizard

Selecting the thick lazy zeroed option for virtual disks on VMFS volumes can provide space savings. If the guest performs a file system format on the virtual disk, it will consume the equivalent amount of VMFS space.

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. When using thin format, monitor the VMFS datastore's free capacity to avoid an out-of-space condition within the datastore. vSphere provides a simple alert for this condition.

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, an out-of-space condition can occur 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 an alert in the vSphere Client, shown in [Figure 77](#).

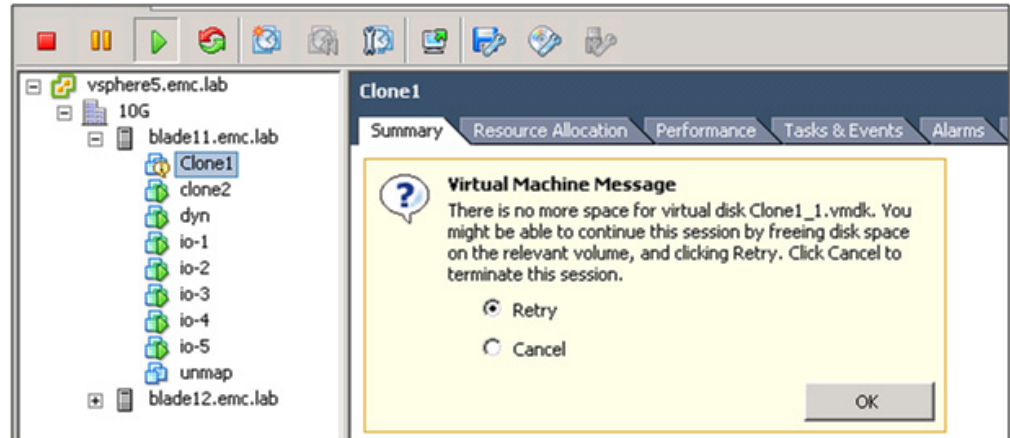


Figure 77 Virtual machine out-of-space error message

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 by expanding a storage pool or removing space from another virtual machine or LUN consumer. Do one of the following:

- ◆ Select **Cancel** to power off the virtual machine.
- ◆ Select **Retry** to resume the virtual machine after adding or reclaiming additional storage capacity.
- ◆ 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 Unified Storage Management 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](#). 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 using the Unified Storage Management feature of 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, but you can view the currently allocated capacity of the file system in the Storage Viewer feature of VSI.

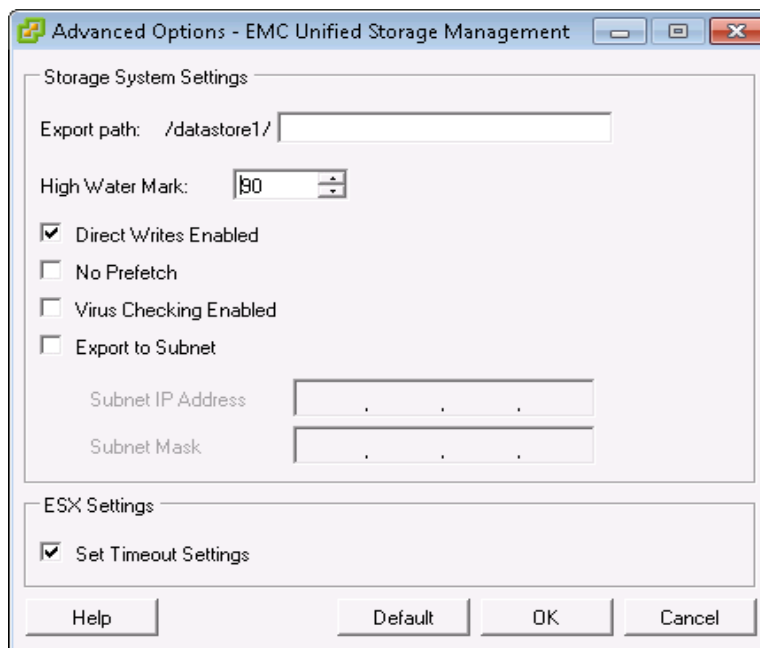


Figure 78 File system High Water Mark in the Unified Storage Management feature of VSI

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: [“Monitoring and managing storage” on page 99](#) 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.

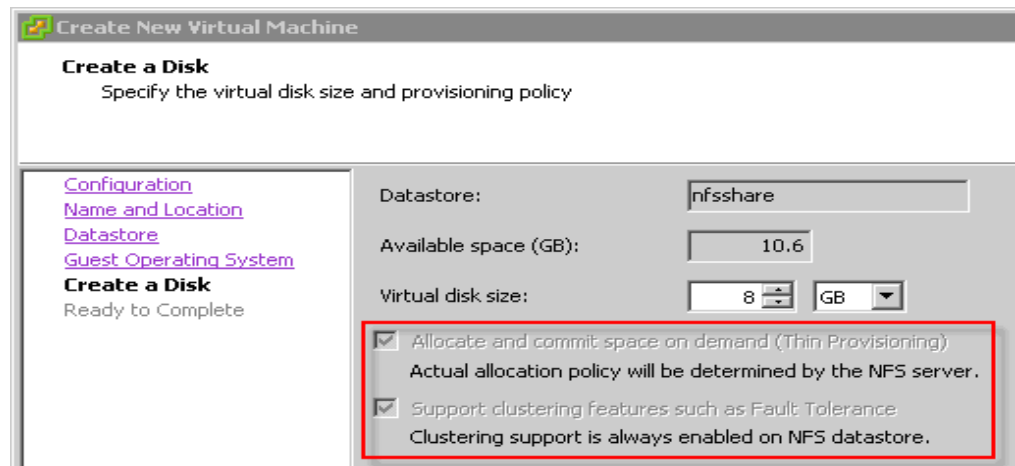


Figure 79 Provisioning policy for an NFS virtual machine virtual disk

LUN compression

VNX LUN compression offers capacity savings for data types with lower performance requirements. LUNs presented to the VMware ESXi host are compressed or decompressed as needed. [Figure 80](#) shows that compression is a LUN attribute that you can enable or disable for each individual LUN. When enabled, data on the 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.

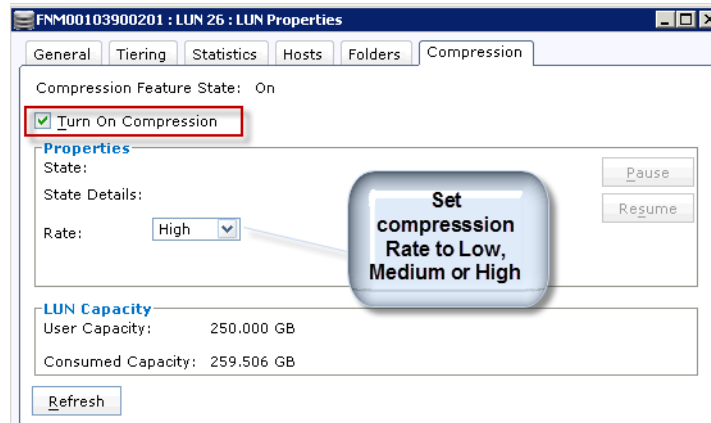


Figure 80 LUN compression property configuration

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), nonproduction clones of databases, and 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, which improves file-storage efficiency. Deduplication eliminates redundant files in a file system with minimal impact to the end user. The use of these technologies results in lower cost-per-megabyte and improved TCO of the VNX system.

VNX file deduplication and compression provide data-reduction cost-savings capabilities in these 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. 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. Some CPU overhead is 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 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 Unified Storage Management and VNX file deduplication and compression.

NFS and CIFS network shares

This section discusses 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 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 12 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 12 Disk types supported by VNX

Type of drive	Available size	Benefit	Suggested Use	Notes
Flash	<ul style="list-style-type: none"> • 100 GB • 200 GB • 400 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 • 900 GB • 10k rpm • 15k rpm 	<ul style="list-style-type: none"> • Cost effective • High 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 • 4 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 	

Disk grouping

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.

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 and deduplication.
- ◆ Storage pools are segmented into 256 MB slices. Pool LUNs are created using multiple slices within the pool.

Table 13 lists the capabilities of RAID groups and storage pools.

Table 13 Pool capabilities

Pool type	LUN type	Allocation	Maximum no. of disks	Expansion	Compression	Deduplication	Unmap	Shrink	Auto tiering
RAID group	FLARE LUN	Full	18	✗	✓	✗	✗	✗	✗
Storage pool	Thin (TLU)	No allocation	121-996 determined by platform	✓	✓	✓	✓	✓	✓
	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 groups

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

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 14 lists the RAID protection options.

Table 14 VNX RAID options

Algorithm	Description	RAID group	Pools	Considerations
RAID 0	Striped RAID.	✓		No data protection
RAID 1	Data is striped across all spindles.	✓		Uses 1 minor disk for each data disk
RAID 1/0	Data is mirrored and striped across all spindles.	✓	✓	uses 1 minor disk for each data disk. Consumes more disk space than distributed parity.

Table 14 VNX RAID options

Algorithm	Description	RAID group	Pools	Considerations
RAID 3	Data is striped, with a dedicated parity disk.	✓		
RAID 5	Data is striped with distributed parity among all disks.	✓	✓	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.	✓	✓	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 pools

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)
- ◆ Block Deduplication
- ◆ 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 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 pool features

FAST VP

Fully Automated Storage Tiering for Virtual Pools (FAST VP) is configured using a combination of two or more of the disk types listed in [Table 14](#). 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 requirement. 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 256 MB 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 and later 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.

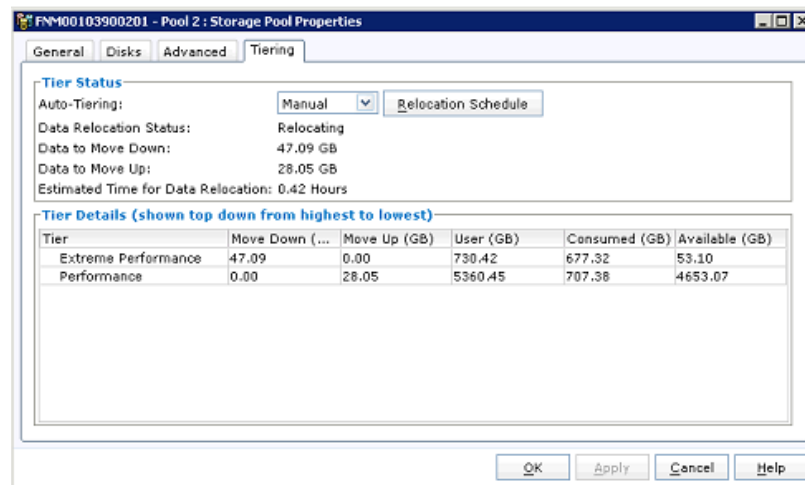


Figure 81 VNX FAST VP reporting and management interface

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.

With advanced snapshots, you can:

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

Advanced snapshots do not perform copy-on-write operations, which means write operations require little overhead. 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 13](#) shows the 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 on how full a pool is, a LUN may not be striped across all disks that make up the pool. 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 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 for the LUNs you are configuring to all fill at the same time, they might not be a good candidate for TLUs. Oversubscribed storage pools should be monitored to detect when pool space is getting low. [“Monitoring and managing storage” on page 99](#) provides more details on how to monitor VNX storage.

Thin LUNs versus thick LUNs

Table 15 illustrates the major differences between thick and thin LUNs.

Table 15 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	

Classic LUNs

Classic LUNs (RAID group LUNs) are the traditional devices created from fixed disk groups. All disk blocks associated with a classic LUN are allocated in a contiguous manner when the LUN is created. Classic LUNs have a fixed limit of 16 drives 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 among more spindles when using RAID groups.

Pool LUN versus RAID group LUN performance

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.

With VNX OE for Block version 5.32 and later, thick LUNs also perform all block allocation when created. This provides locality and performance similar to RAID group LUNs.

Thick LUNs created prior to VNX OE for Block 5.32 do not perform an initial allocation and can have spatial locality and response times that 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 (dvols). VNX OE for File volume manager is used to create aggregate, stripe, and slice dvols for file systems that are presented to ESXi as NFS datastores.

Most LUN properties and features described in this document also apply to file system storage for NFS datastores.

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.

MVM enables the storage administrator to select the components to be used to create the volume, providing 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, 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.

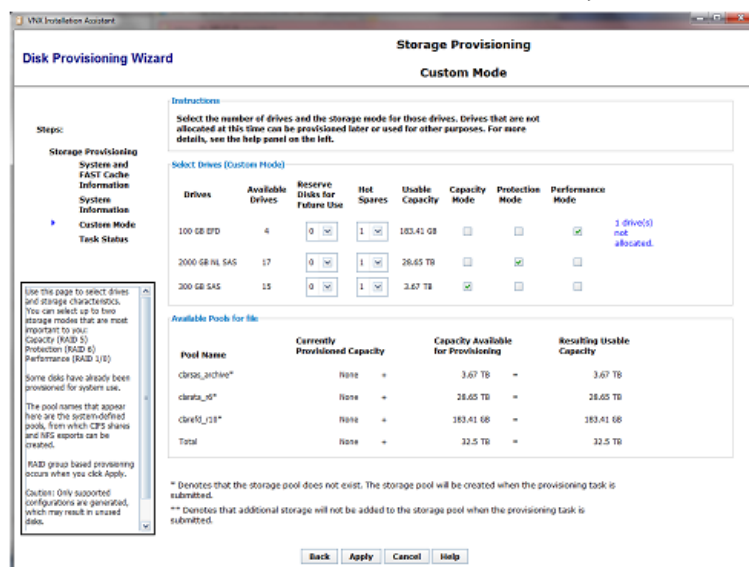


Figure 82 Disk Provisioning Wizard

CHAPTER 2

Cloning Virtual Machines

This chapter presents the following topics:

- ◆ Introduction 128
- ◆ Using EMC VNX cloning technologies 128
- ◆ Summary 138

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 the 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 clones can be used to replicate the VMFS datastore by creating an identical block-for-block replica of a LUN used by ESXi.

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.

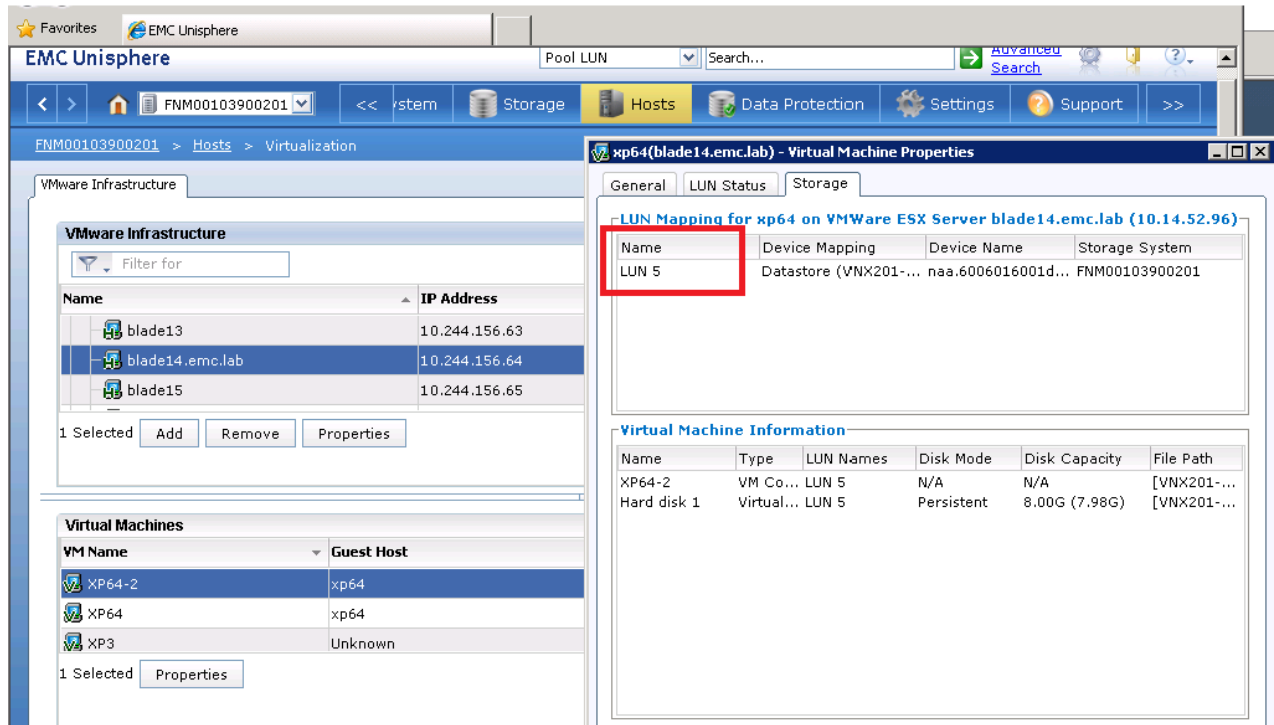


Figure 83 Unisphere clone LUN management

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.

Adding 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.

You can 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.

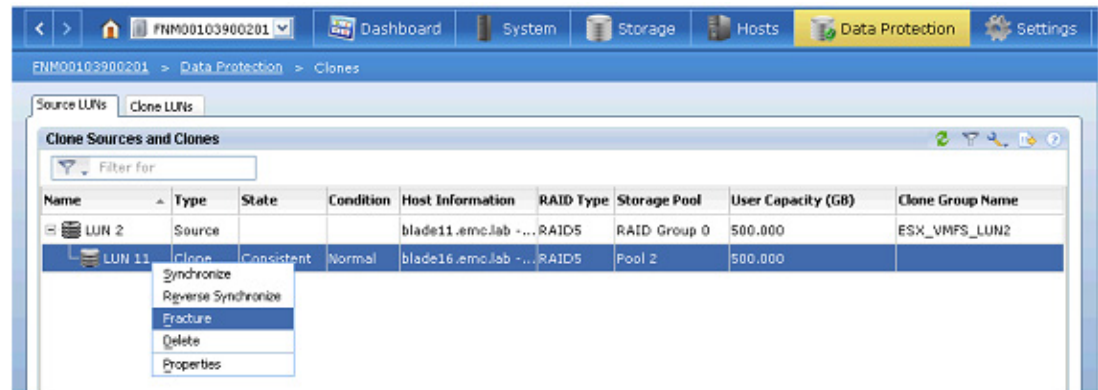


Figure 84 Performing a consistent clone fracture operation

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.

- Use either Unisphere or Navisphere CLI, as shown in [Figure 85](#), to start a SnapView session on the source device. This step initiates the copy-on-write activity.

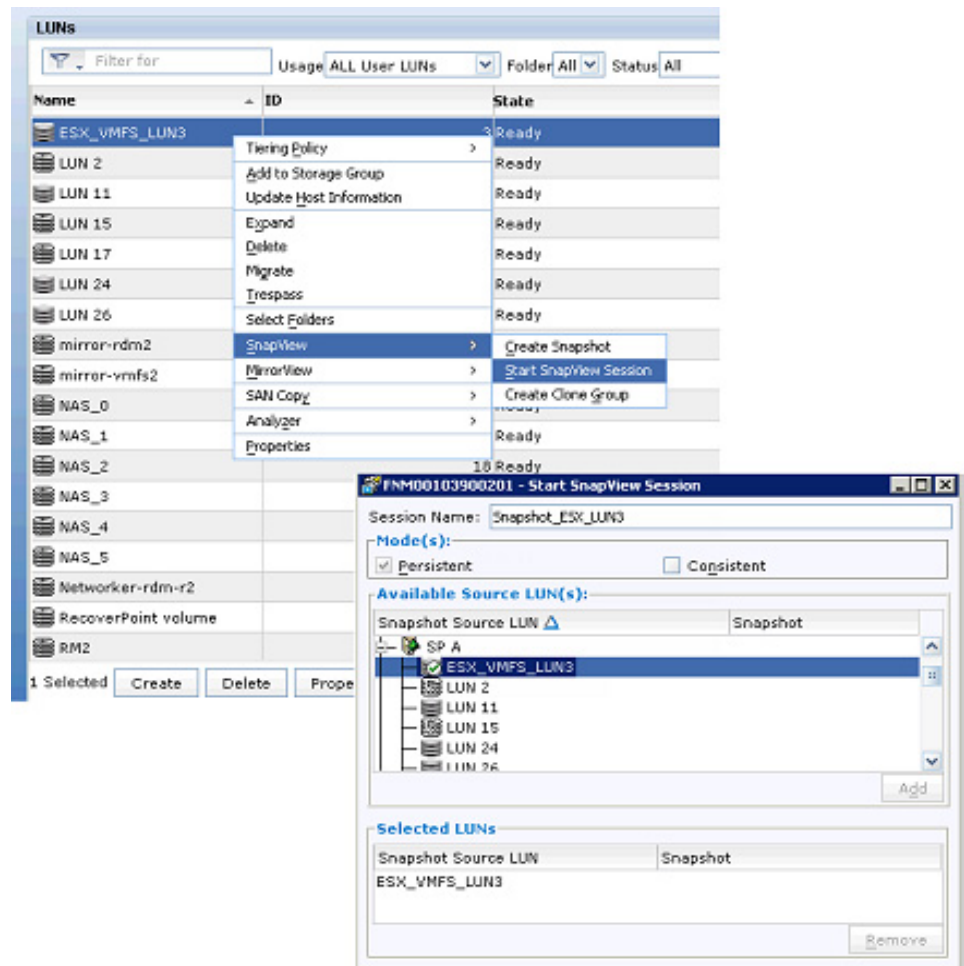


Figure 85 Creating a SnapView session to create a copy of a VMware file system

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, do one of the following:

- ◆ 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:

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 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**.

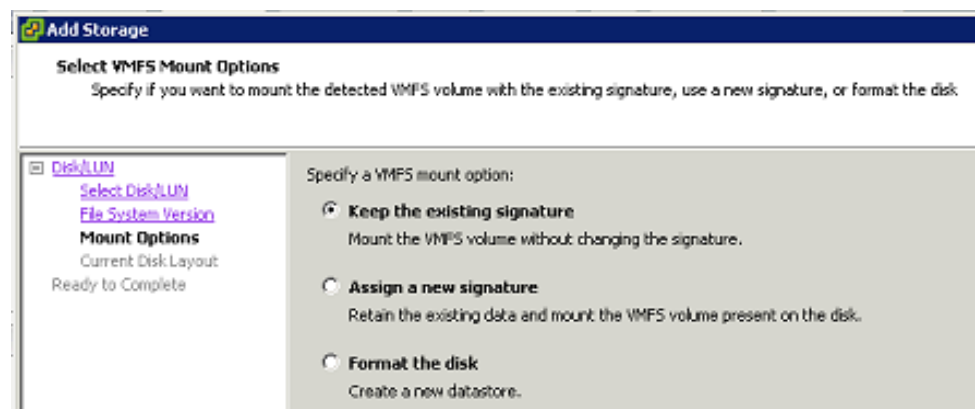


Figure 86 Device signature assignment

8. Review the datastore configuration information, and then click **Finish**.
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.
4. 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.

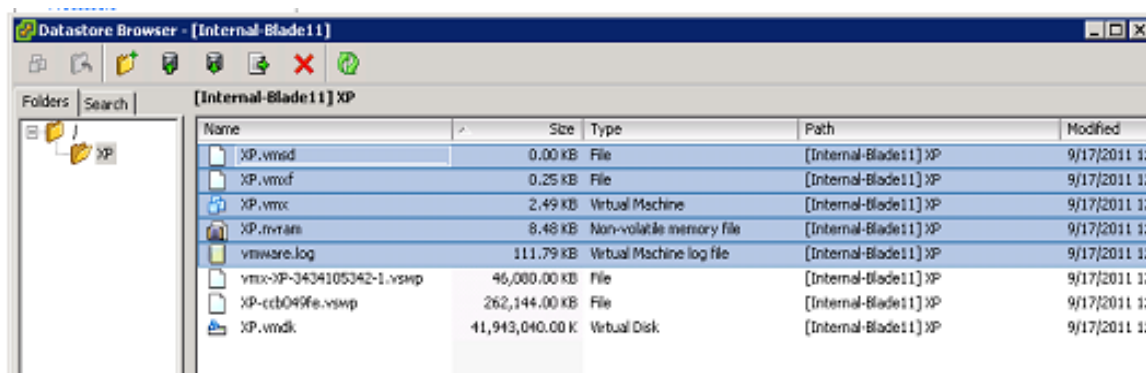


Figure 87 Selecting virtual machine configuration files in the datastore browser

5. 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](#).

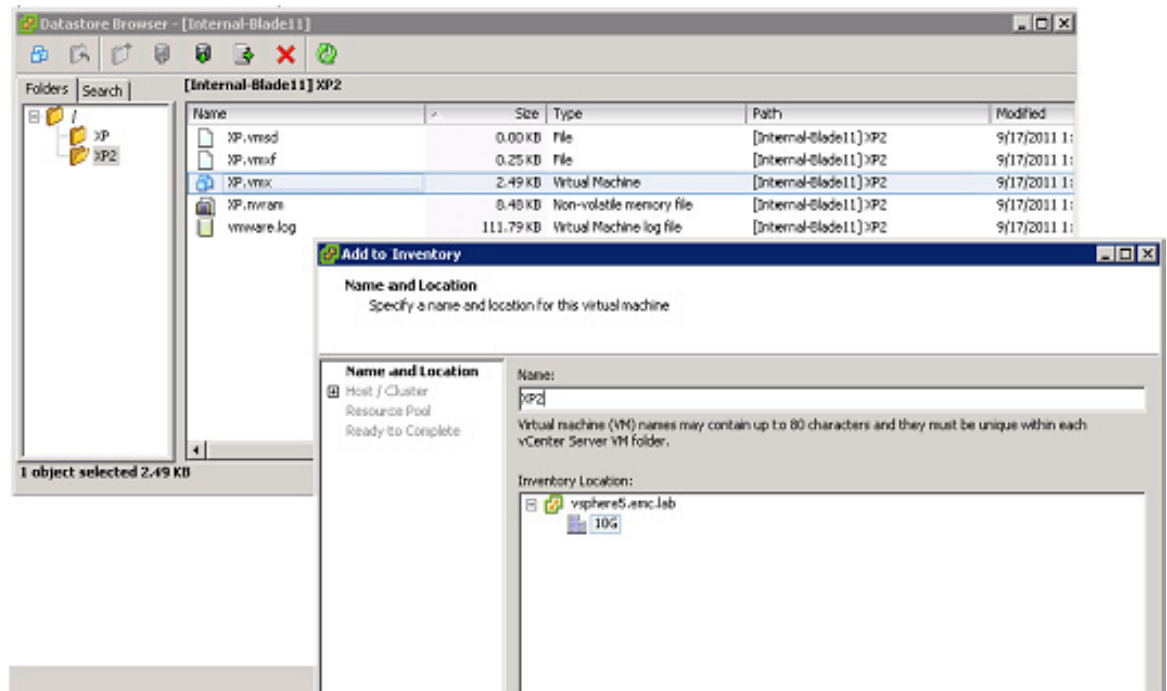


Figure 88 Adding the new virtual machine to the ESXi host inventory

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](#).

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 VSI](#)” on page 51 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 in the vCenter UI:

1. Select an ESXi host with access to the Checkpoint File System.
2. Select **Configuration** and click **Add Storage**.
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 version 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. To verify that EMC NAS VIB is installed, use the vSphere Client or run the following command:

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

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

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

View: Datstores Devices

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

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

Figure 90 Cloned NFS datastore in vSphere

NFS VAAI clones do 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 Unified Storage Management feature performs cloning operations directly within the storage system using a separate management approach from the VAAI cloning operations.

The utilities in Unified Storage Management include full clones 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 about the Unified Storage Management 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 16](#).

Table 16 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

CHAPTER 3

Backup and Recovery Options

This chapter presents the following topics:

◆ Introduction	142
◆ Virtual machine data consistency	142
◆ VNX native backup and recovery options	143
◆ Snapshot backup and recovery of a VMFS datastore	145
◆ Backup and recovery of RDM volumes	148
◆ AppSync	148
◆ Replication Manager	154
◆ Backup and recovery of a VMFS with VNX Advanced Snaps.....	157
◆ vStorage APIs for Data Protection	164
◆ Backup and recovery using VMware Data Protection	165
◆ Backup and recovery using Avamar	183
◆ Backup and recovery using NetWorker.....	189
◆ Summary	193

Introduction

The combination of EMC data protection and VMware vSphere technologies offers several backup and recovery options for virtual environments.

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.

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.

Virtual machine data consistency

When RDM is used in VMware environments, all virtual disks reside on a datastore or a LUN. In a basic configuration, all of the virtual disks are stored on the same datastore and you can use native VNX features to provide crash consistency of a virtual machine by creating a replica of the LUN or NFS file system that supports the datastore.

Application vendors, especially database vendors, recommend distributing applications across multiple virtual disks and datastores to achieve better performance. In this type of configuration, all datastores that support the application must be managed as a single entity. VNX consistency groups can be used with VMware snapshots to provide crash consistency of block storage devices in these cases.

A VMware snapshot is a log-based protection mechanism in which changes made to a virtual disk are applied to a journal file. The hypervisor suspends all Virtual Machine I/O prior to the creation of a virtual machine snapshot. The snapshot captures the entire virtual machine state, including configuration settings, virtual disk contents, and the contents of the virtual machine memory.

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 use VMware snapshots to ensure that the virtual machines are in a consistent state prior to creation of an NFS SnapSure checkpoint or SnapView LUN snapshot. When a virtual machine spans datastores, the backup set consists of EMC snapshots of all the datastores containing the virtual machine's disks.

A backup must contain all virtual machine files in order to preserve the system state when the snapshot is created. Ensure that the backup set includes all datastores where the virtual machine disks and files might reside, taking into account features like Storage DRS that might have relocated a virtual machine from its initial datastore placement.

Performing a backup

Note: EMC Replication Manager is used to automate these steps and provide application integration and application consistency. [“Replication Manager” on page 154](#) provides more information.

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 27](#) provides more details.

4. Delete the VMware snapshot.

Restoring a virtual machine from a snapshot

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.

VNX native backup and recovery options

VNX offers native utilities to create replicas of file systems and LUNs. While these utilities can be used to protect ESXi Datastores in a vSphere environment, EMC provides comprehensive solutions like AppSync and Replication Manager with application-level integration for enterprise-level backup and recovery.

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

File system logical backup and restore using VNX SnapSure

Use VNX SnapSure to create near-linear 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 in the table, and a detailed dialog box is open for it. The dialog shows the parameter name, data mover, facility, and value (1). The description explains that setting showChildFsRoot=1 makes checkpoints visible to NFS clients as subdirectories of the root directory.

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

showChildFsRoot Properties:

- 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 Viewing ChildFsRoot parameter properties in Unisphere

Recovering a virtual machine from an NFS checkpoint

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, perform 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 allowing for selective virtual machine recovery or complete file system restoration through Unisphere.

Selective virtual machine recovery

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.

Best practices:

- ◆ 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 management or service level requirements should be placed in separate file systems.

Note: If virtual machine snapshots exist before the creation of a backup, vCenter Snapshot Manager might not report them correctly when a virtual machine is restored. In this case, 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!*

File system recovery

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 can be accessed using Unisphere, Unisphere Snapshot Configuration Wizard, or the `admsnap` utility.

For simplified configuration, automation, and monitoring of replication jobs in enterprise environments, you can control LUN protection using Replication Manager or AppSync Manager. 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 data blocks of 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 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 is provisioned from a single LUN.

Best practices:

- ◆ 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 restored 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](#).

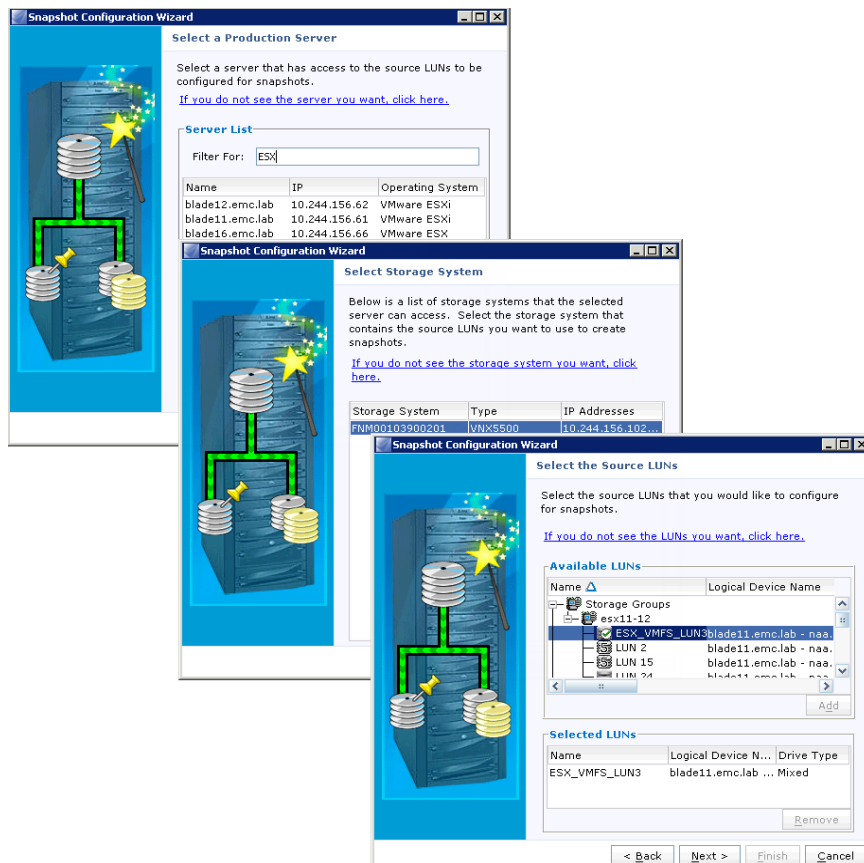


Figure 92 Snapshot Configuration Wizard

3. Select the number of copies for each source LUN.
4. Optionally, assign the snapshot to other ESXi hosts as shown in [Figure 93](#).

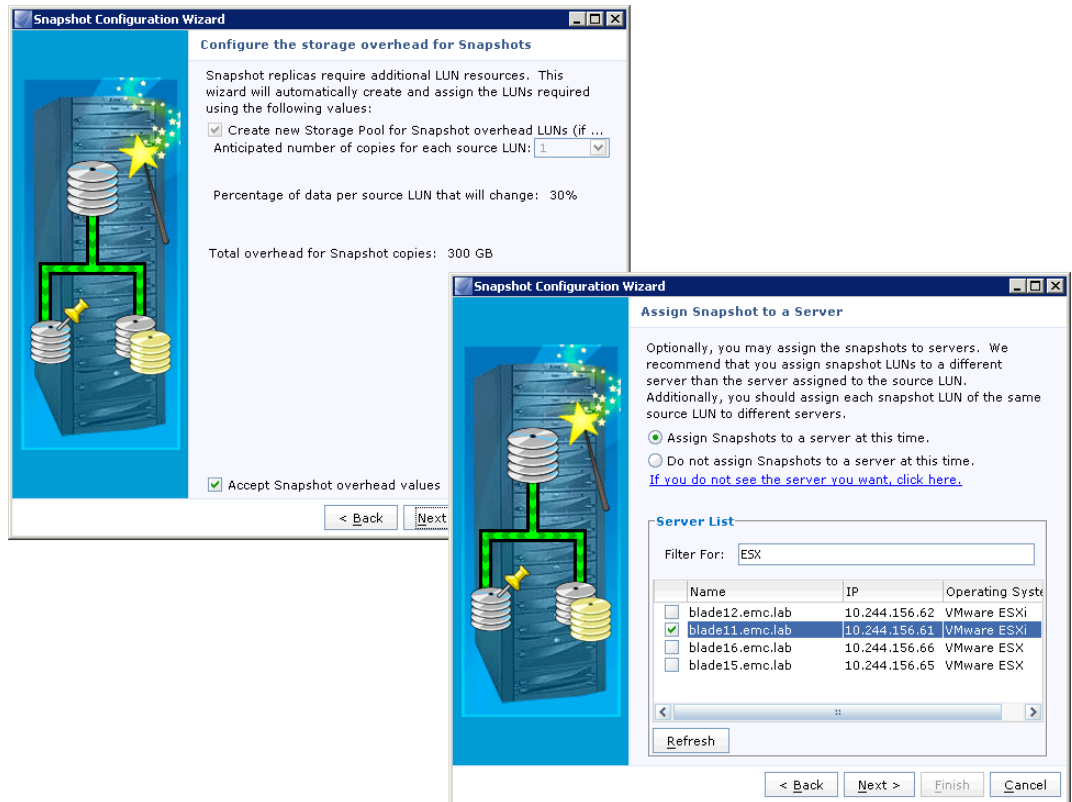


Figure 93 Snapshot Configuration Wizard (continued)

5. Type the snapshot name.
6. Select the host to which you want to add the snapshot image.
Unisphere assigns the LUN to the host's storage group.
7. Review the configuration information and click **OK** to create and mount the snapshots.
8. Use Unisphere to start the snapshot session and activate the snapshot for access by another host.
9. Rescan the ESXi hosts to verify that the storage appears in the correct location.
10. If required, select **Assign a new signature** to automatically resignature the device.
“[ESXi volume signatures](#)” on [page 132](#) provides more information on device signatures.

When the snapped VMFS LUN is accessible from the ESXi host, you can copy virtual machine files from the snapped datastore to the original VMFS datastore to recover the virtual machine.

Backup and recovery of RDM volumes

VNX LUNs provisioned to ESXi hosts are formatted as VMFS file systems or presented directly to a virtual machine as RDM (raw device map) devices. RDM volumes possess similar properties to VMFS virtual disks, while retaining the capabilities of a physical device. Storage administrators can take full advantage of storage array-based data protection technologies at the LUN level. EMC data protection products such as SnapView and AppSync provide native protection of RDM devices.

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

For RDM volumes, snapshots or clones can be created as follows:

- ◆ By using the **admsnap** command or the Unisphere Snapshot Configuration Wizard
- ◆ By using Replication Manager or AppSync 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.

AppSync

AppSync is an EMC data protection product that simplifies replication of ESXi datastores. Appsync is integrated with EMC VNX Snapshots, EMC SnapSure, EMC RecoverPoint, and IP replicator to create local or remote replicas of VNX datastores.

AppSync identifies all on-line virtual machines and initiates the creation of virtual machine snapshots prior to creating local replicas. The VMware snap attempts to quiesce all I/O to the virtual machine before the snap is created, providing a higher level of consistency than simply snapping the datastore at the device level.

AppSync uses a proxy (physical or virtual) host to communicate with the vCenter and VNX storage systems. It performs device enumeration within both environments and initiates the necessary management tasks to establish consistent copies of the datastores and virtual machine disks. The AppSync Job Wizard simplifies the definition of the backup task providing options to select the replica type and expiration options.

Note: AppSync version 1.5 is required for NFS datastore support.

Configuring AppSync

To configure AppSync:

1. Log into AppSync using the desktop shortcut.
2. Select **Storage Infrastructure** from the settings.
3. Click **Add VNX Storage System**.

4. Type the IP Addresses of the SP and the administrative user credentials.

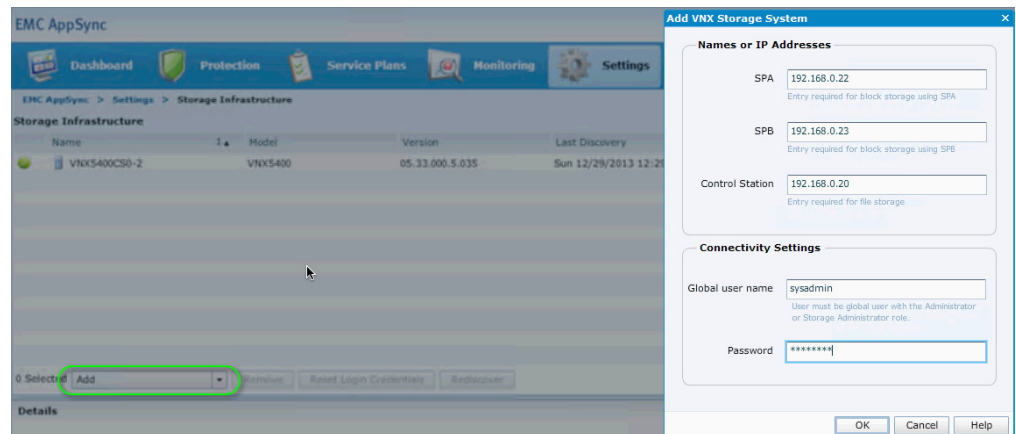


Figure 94 Adding a VNX Storage System to AppSync

5. Add a vCenter-Server:
 - a. Go to **Protection > VMware Datacenter** and click **Discover Datacenter**.
 - b. Click **Add vCenter Server**.
 - c. Type the system properties and select **Run Discovery Now**.

AppSync is authenticated to vCenter and lists all existing datastores.

AppSync provides Service Plans to map an application's Service Level Objectives (SLO) to the datastore. To simplify configuration, AppSync provides templates for VMware Datacenters, Microsoft Exchange, and Microsoft SQL.

Adding a Service Plan

To add a new Service Plan:

1. From the AppSync management interface, select **Service Plans > VMware Datacenter**.
2. Click **Create** and enter the system properties.
3. Choose the service level and provide a name.

To adjust this Service Plan to your own needs, click on its name.

4. Select a **Recovery Point Objective**.

For example, the service plan named “four hours” creates a new snapshot of each datastore in the plan, every four hours.

5. Confirm your changes and click **Apply**.

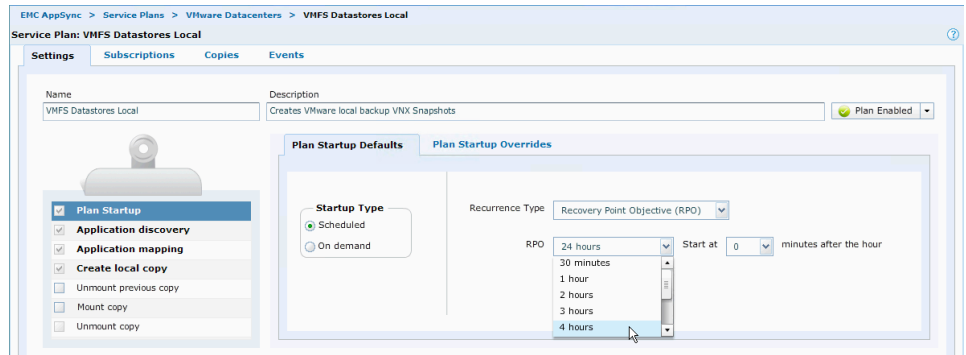


Figure 95 Editing Service Plan settings

Once the protection plan has been created, you can use it to protect datastores in the environment. The following steps and illustration in [Figure 96](#) illustrate the protection of a datastore using the protection plan created in “[Adding a Service Plan.](#)”

Protecting the datastores

1. Click **Protection > VMware Datacenter**.
2. Select the datacenter to view a list of all datastores.
3. Select a VMFS Datastore on the VNX system and select **Protect** to assign it to the newly created Service Plan.
4. Select the Service Plan to which you want to assign the datastore.

In this example VMFS Datastore Local is assigned.

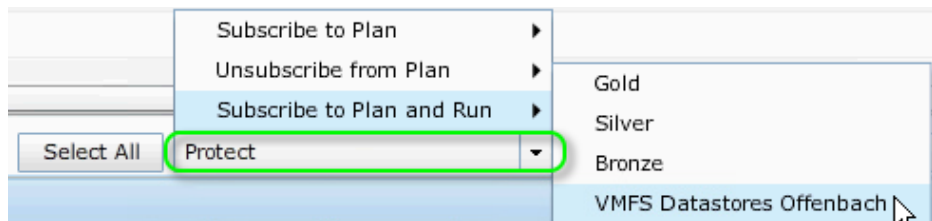


Figure 96 Subscribing to a Service Plan

AppSync provides immediate protection. A check mark appears next to datastores that are protected.

EMC AppSync > Protection > VMware Datacenters > DR-Site				
VMware Datastores				
Name	Type	Service Plan	ESX Servers	
VNX5400CS0-2-VMFS-Datastore-02	VMFS		esx06.demo.ffm,esx05.demo.ffm	
VNX5400CS0-2 NFS Datastore	NFS		esx06.demo.ffm,esx05.demo.ffm	
VNX5400 VMFS Datastore	VMFS	VMFS Datastores Local	esx06.demo.ffm,esx05.demo.ffm	

Figure 97 Overview of protected and unprotected datastores

VSI AppSync Management

The AppSyncManagement feature of VSI offers protection for VMFS and NFS datastores. Both AppSync and Replication Manager allow you to take application and virtual machine-consistent snapshots or replications. ReplicationManager supports all of EMC's diverse data protection capabilities, while AppSync is tailored for VNX and EMC RecoverPoint. Both products feature a VSI plugin.

Configuring AppSync in VSI

1. Login to a vSphere Client where VSI is installed.
2. Select **AppSyncManager** from the VSI **Features** window.
3. Click **Register**.
4. Enter the AppSync system credentials and click **Next**.
5. Enter the vCenter credentials and click **Finish** to complete the configuration and exit the wizard.

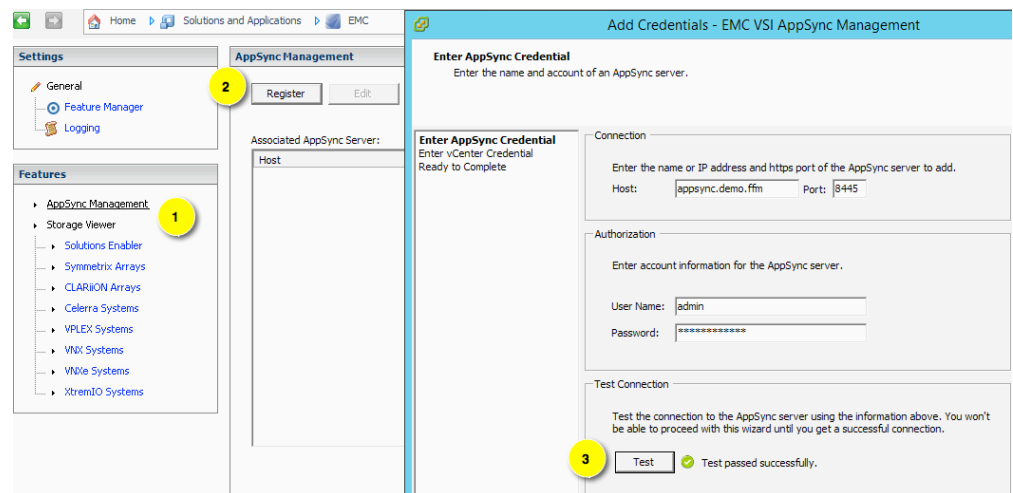


Figure 98 VSI AppSync registration

Protecting a datastore using AppSync Management

1. From the vSphere Client, click **vSphere > Datastores and Datastore Clusters**.

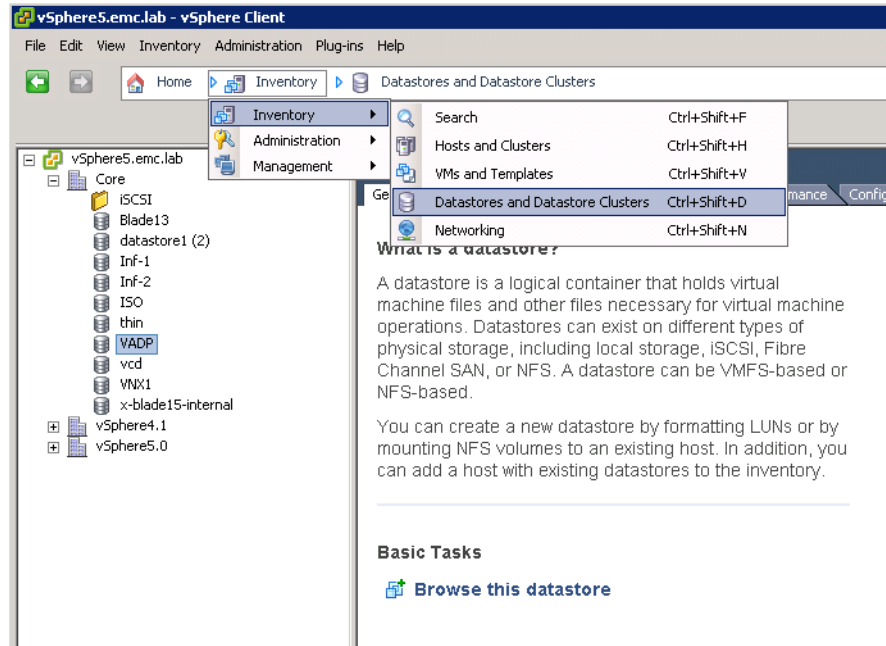


Figure 99 Protect datastore

2. Right-click the datastore you want to protect, and then select **EMC > AppSync > Configure Protection**.

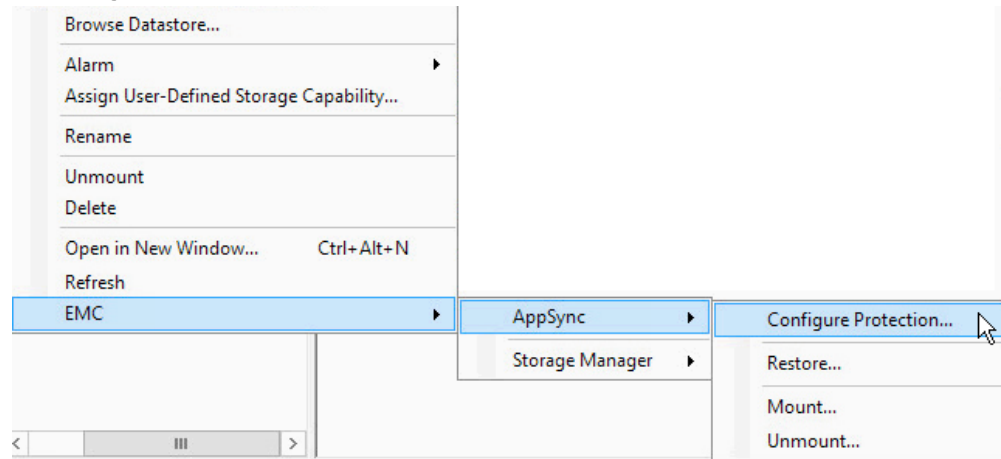


Figure 100 Using VSI AppSync Feature to Configure datastore protection

3. In the AppSync wizard, perform the following steps to complete the configuration:
 - a. Select **Create and Subscribe**.
 - b. Select the VMFS datastore local protection plan you created.
 - c. In **Plan Detail**, type a descriptive name and click **Next**.
 - d. Click **Set Schedule** to define the RPO.

e. Click **Finish**.

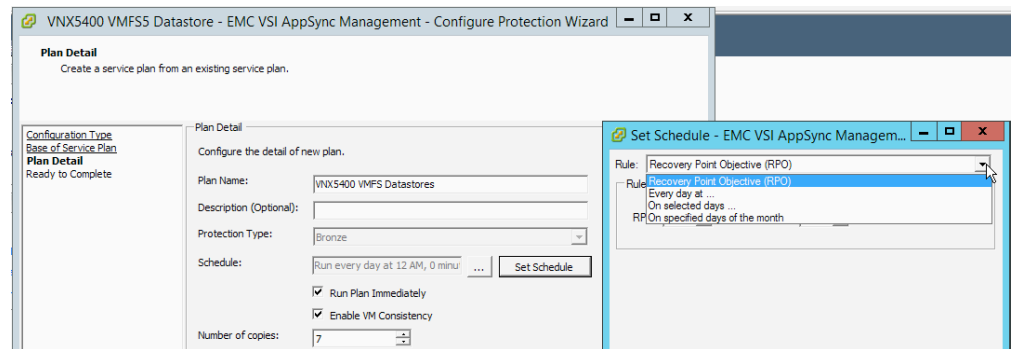


Figure 101 Subscribe Datastores to a AppSync Service Plan

f. Validate your settings and exit the wizard.

The VMFS Datastore is immediately protected.

Mounting replicas

Using VSI AppSync Manager, you can manually create copies or restore virtual machines by presenting the replica to any ESXi Host that is configured on the VNX storage system where the replica was created.

1. Select the datastore in the VSI AppSync Manager menu.
2. Click **Mount**.
All the snapshot copies for that device are listed.
3. Select a copy and click **Next**.
4. Select an ESXi host on which to mount the snapshot copy.

Note: to avoid a device signature conflict between the source device and the snapshot, ensure that **Resignature** is selected, as shown in [Figure 102](#).

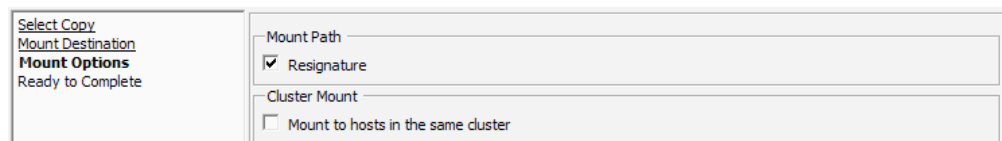


Figure 102 Resignaturing a copy of an existing datastore

5. Complete the wizard to complete the mount task.

Restoring a virtual machine

Once the datastore is mounted it can be browsed using the datastore browser. This is a useful approach to performing a selective virtual machine restore. You can restore a failed virtual machine by mounting a protected copy of the datastore from either AppSync or the VSI AppSync Manager feature.

1. Power off the target virtual machine you want to recover.
2. Delete it from the disk.

3. Browse to the replica datastore to identify the virtual machine to restore.
4. Copy the virtual machine folder and its contents from the replica to the root of the production datastore.
5. Locate the VMX file and register the virtual machine.

Figure 103 illustrates these steps.

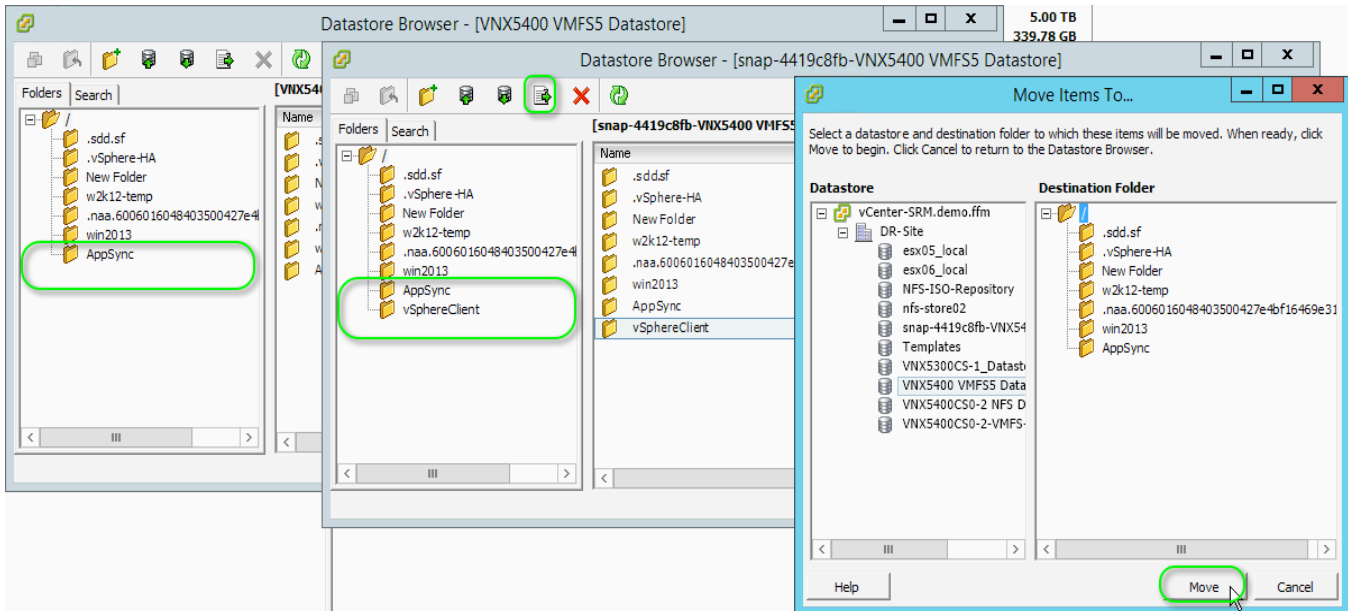


Figure 103 Using vCenter to manually copy a virtual machine

6. After the virtual machine is fully recovered, unmount the snapshot copy of the datastore using VSI AppSync Manager.

Note: Using AppSync, you can also restore the entire datastore. This is a much quicker and more direct restore method. However, a complete restore overwrites all virtual machines on the source datastore. There is potential for inadvertent loss, if you restore a datastore with active virtual machines that you did not want to change.

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. Including virtual machine snapshots 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 initiate 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.

Restoring the replicas

Use the Replication Manager Job Wizard, as shown in [Figure 104](#), to select the replica type and expiration options. Replication Manager 5.2.2 is required for datastore support.

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

Figure 104 Replication Manager Job Wizard

3. In Replication Manager, select **Restore** to restore the entire datastore.
4. Restore the replica.
5. When the restore is complete, import the virtual machines to the vCenter Server inventory.
6. Revert to the VMware snapshot taken by Replication Manager to obtain an OS-consistent replica, and delete the snapshot.
7. 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 105](#).

- Verify the contents of the restored replica, and then delete the rollback snapshot.

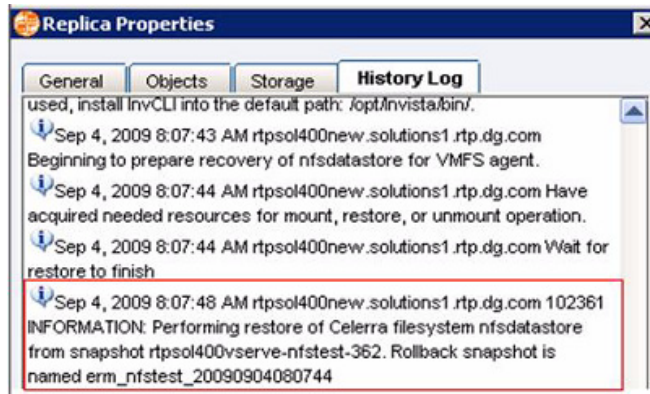


Figure 105 Viewing Replica Properties in Replication Manager

Restoring a virtual machine

Replication Manager version 5.3 and later provides the ability to selectively restore a virtual machine, as shown in Figure 106.

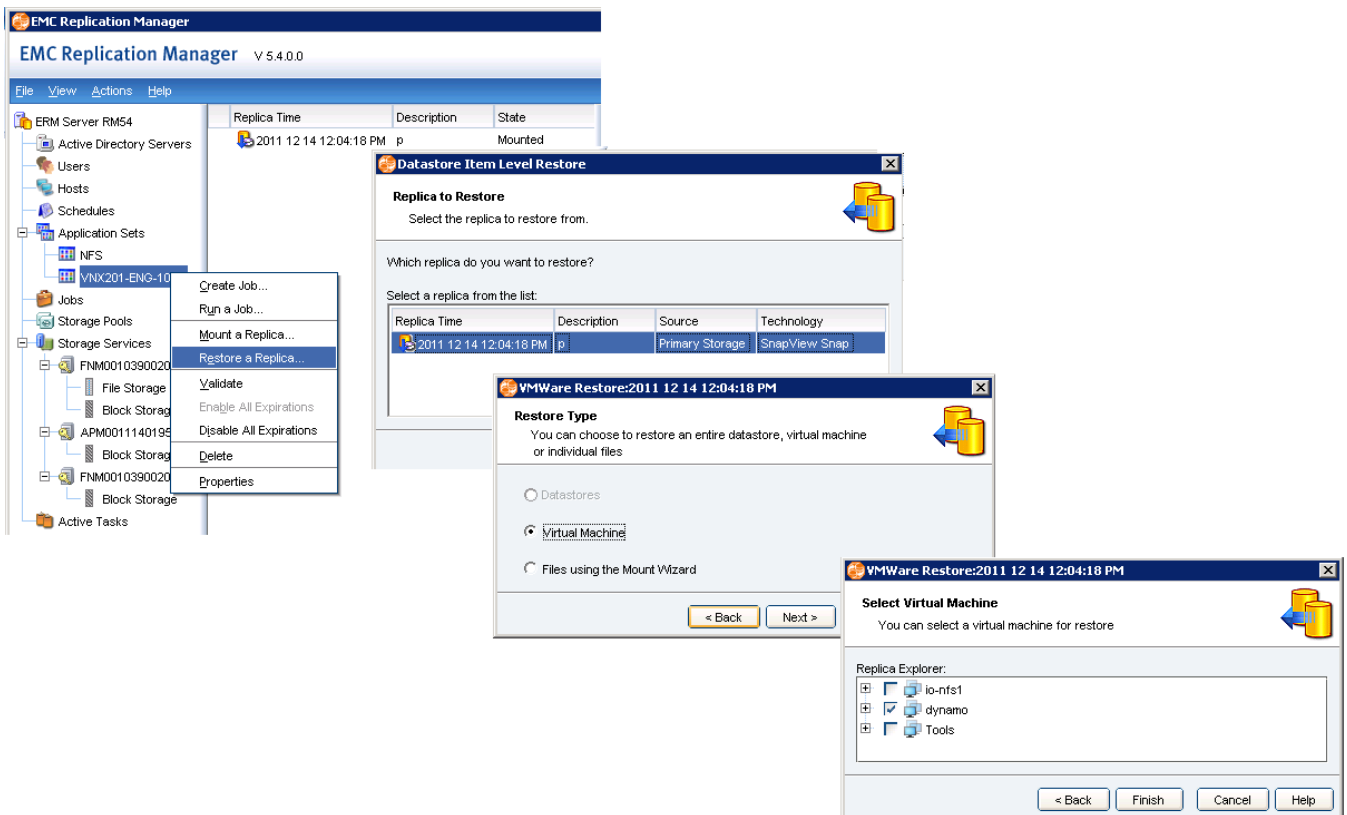


Figure 106 Restoring a virtual machine using Replication Manager

- Select the application set that contains the replica you want to restore.
- Identify the date and time the replica was created.
- Right-click the replica and select **Restore a Replica**, and then click **Next**.
- Select the virtual machine or virtual machines to restore, and then click **Next**.

The Replication Manager status pane displays the progress.

5. Revert to the VMware snapshot taken by Replication Manager to obtain an OS-consistent replica, and then delete the snapshot.
6. Use Replication Manager to unmount the replica.
7. 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 107 Viewing a read-only copy of the datastore in vSphere Client

Backup and recovery of a VMFS with VNX Advanced Snaps

EMC VNX OE for Block release 5.32 and later provides a new snapshot architecture for pool LUNs. Using the new snapshot you can create up to 256 snapshots of the source LUN, including snapshots of existing LUN snapshots.

Snapshots are created from individual LUNs or groups of LUNs defined in a consistency group.

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 GUI enables you to manage advanced snapshots. Additionally a command line utility is available for in-band management when the snapshot mount point is enabled. [Figure 108](#) shows this option selected to allow in-band management of the snapshots assigned to the host.

Using Unisphere to create snapshots of Pool LUNs

1. Right-click the LUN and select **Snapshot > Create Snapshot**.

Use VSI or Unisphere virtualization view to identify the datastore LUN.

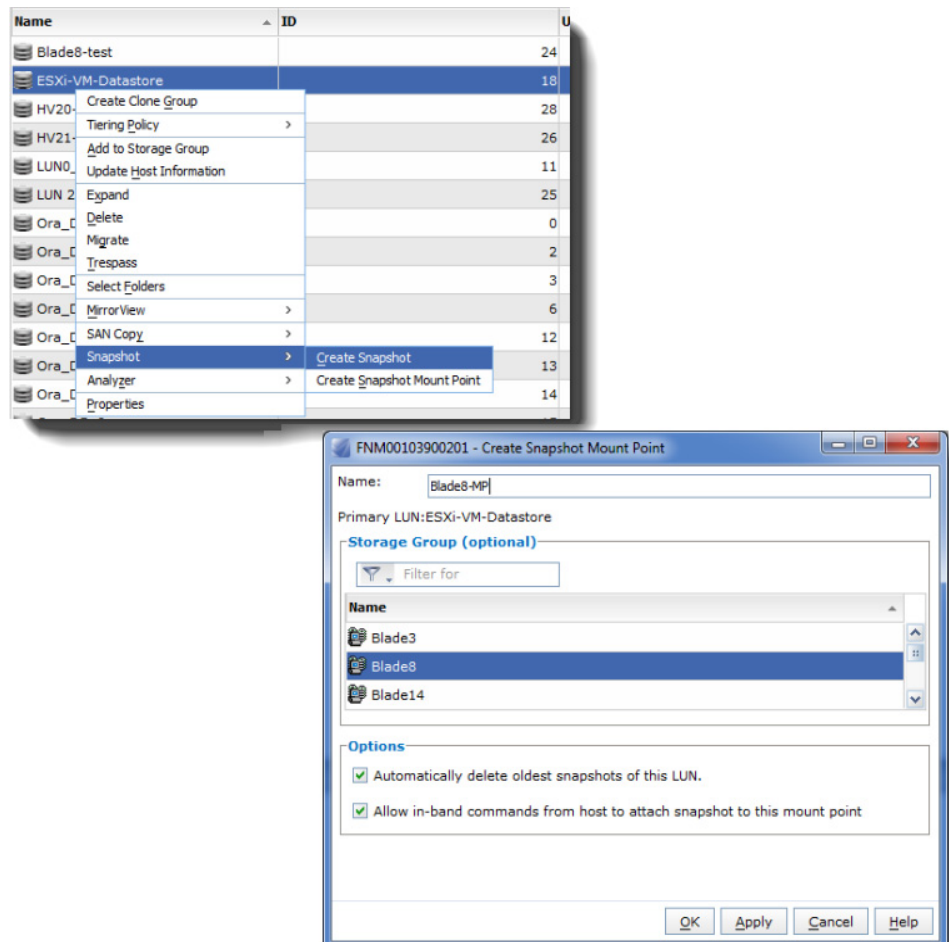


Figure 108 Advanced Snapshot Basic Configuration

2. If a snapshot mount point does not exist, create one and assign it to a storage group so the ESXi host can access the snapshot image.

In the example in [Figure 109](#), a snapshot mount point named **Blade8-MP** is assigned to Blade8. You can use Unisphere to attach and detach snapshots from the mount point.

Select the storage group to display the **Snapshot Mount Points** details, as illustrated in [Figure 109](#).

The screenshot shows the 'Storage Groups' interface. At the top, there is a search bar with 'Blade8' entered. Below it, a table lists storage groups with columns for 'Storage Group Name' and 'WWN'. The 'Blade8' group is selected, showing its WWN as 'A2:13:A1:DA:24:E5:E1:11:AB:6C:00:60:16:41:5D:A7'. Below the table are buttons for 'Create', 'Delete', 'Properties', 'Connect LUNs', and 'Connect Hosts'. The 'Details' section is active, with 'Snapshot Mount Points' selected. A table below shows the details for 'Blade8-MP'.

Name	ID	State	RAID Type	Storage P...	User Capa...	Current O...	Host Infor...	Additional...	Snapshot...	Host LUN ...
Blade8-MP	8082	Ready	Mixed	Pool 2	100.000	SP B	blade8		On	1

Figure 109 Viewing Snapshot Mount Points

3. Specify the snapshot name when the snapshot and mount point are created.

Creating consistency groups with multiLUN configurations

1. In Unisphere, select **Data Protection > Snapshots**.
2. Open the **Snapshot Mount Points Configuration Wizard** as shown in [Figure 110](#).
 - The host requires a snapshot mount point for each LUN in the consistency group.
 - a. Select the system to mount the snapshots.
 - b. Select the storage system containing the LUNs to be part of the consistency group.
 - c. Select all of the LUNs to be part of the consistency group.
 - d. 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.

3. Click **OK** to finish.

The mount points are created for the snapshots from your application LUNs.

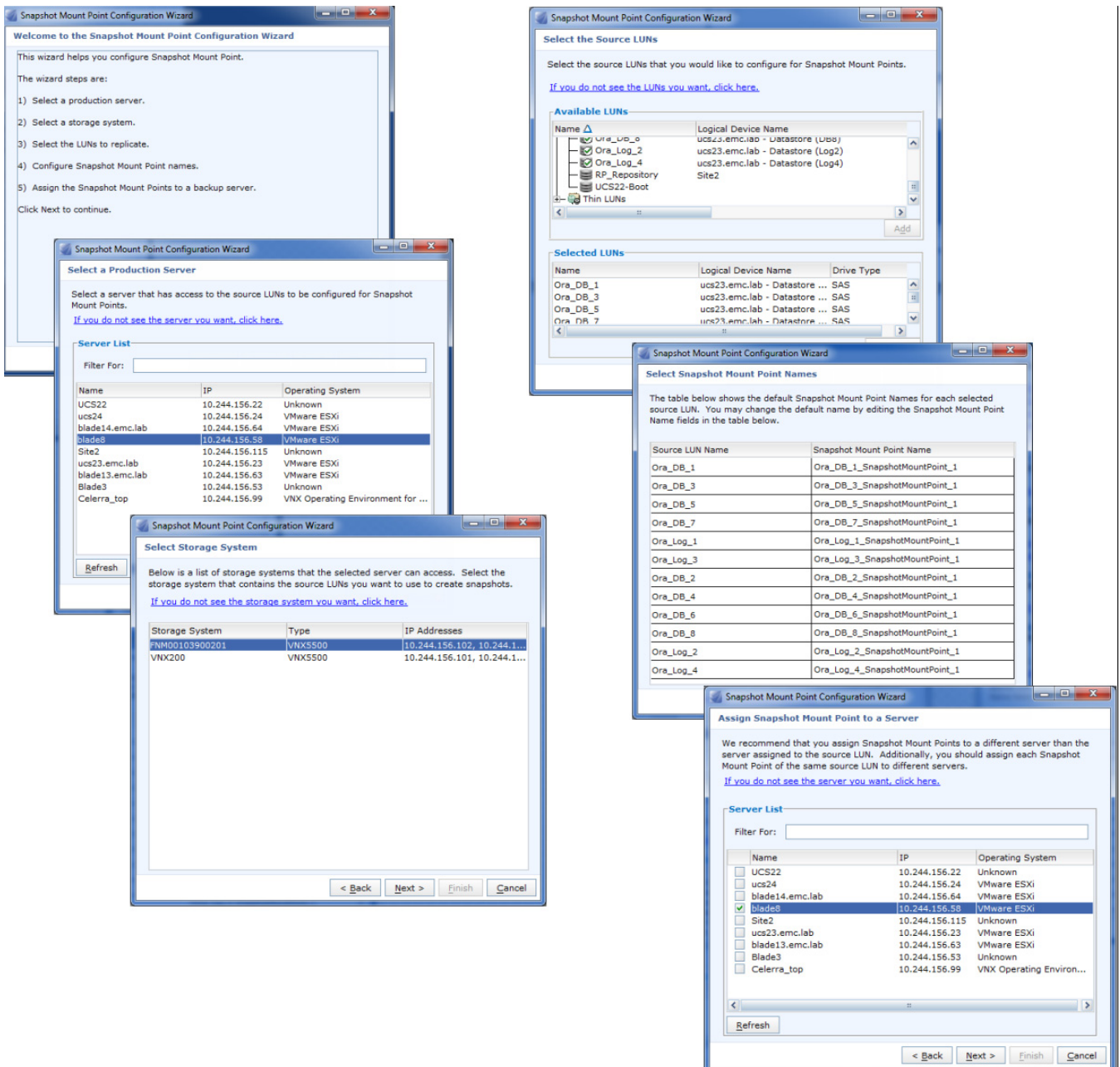


Figure 110 Using the Snapshot Mount Point Configuration Wizard

4. To create the consistency group, click **Data Protection > Snapshots > Create Group**.
5. Type the group name and, optionally, the description.
This example protects 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 111 illustrates creating the consistency group.

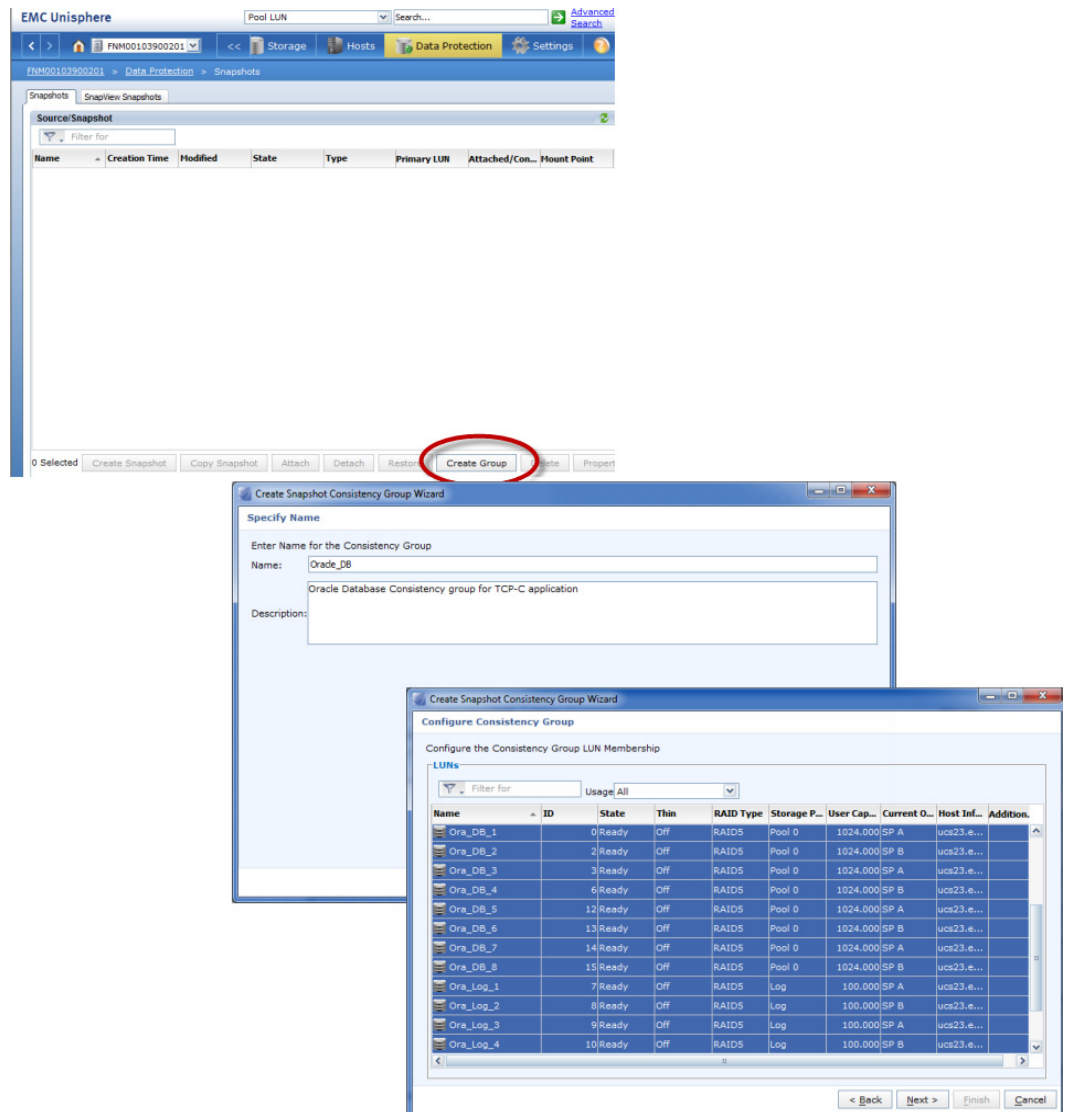


Figure 111 Creating a snapshot consistency group

8. To create a snapshot of all LUNs in the consistency group, select the consistency group.
9. Select a host to add the snapshot image to the host storage group.

Figure 112 shows the process of creating a consistency group snapshot.

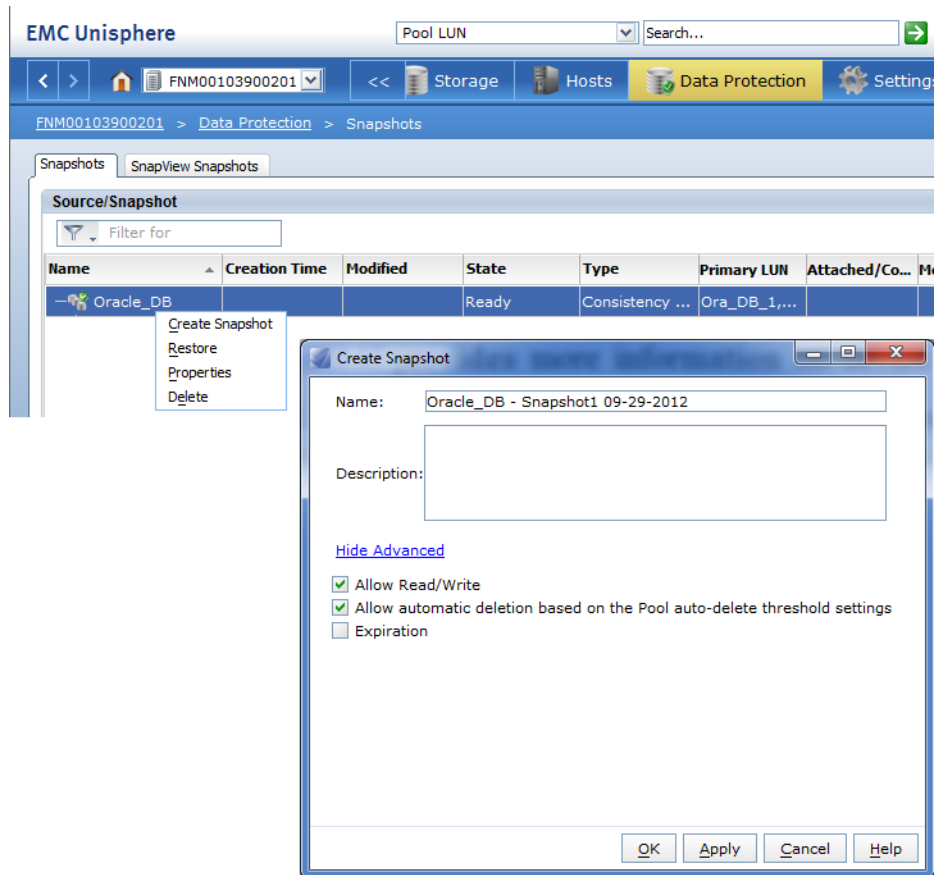


Figure 112 Creating a consistency group snapshot

Figure 113 shows how to attach the snapshots to the mount points to present them to the host.

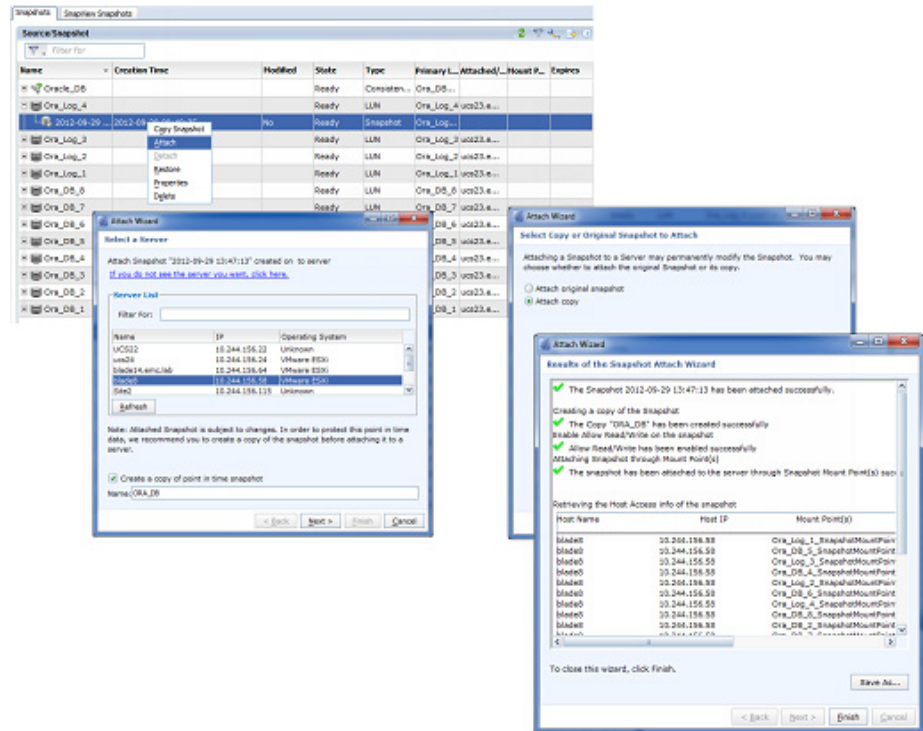


Figure 113 Attaching a consistency group snapshot

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 which you want to attach the snapshots.
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.
15. If required, select **Assign a new signature** to automatically resignature the device.

“ESXi volume signatures” on page 132 provides more information on device signatures.
16. When the snapped VMFS LUN is accessible from the ESXi host, copy the virtual machine files 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. Data protection vendors use VADP 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.

VADP provides the following benefits:

- ◆ Shorter backup times for an environment
- ◆ 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 114](#) shows how VADP works.

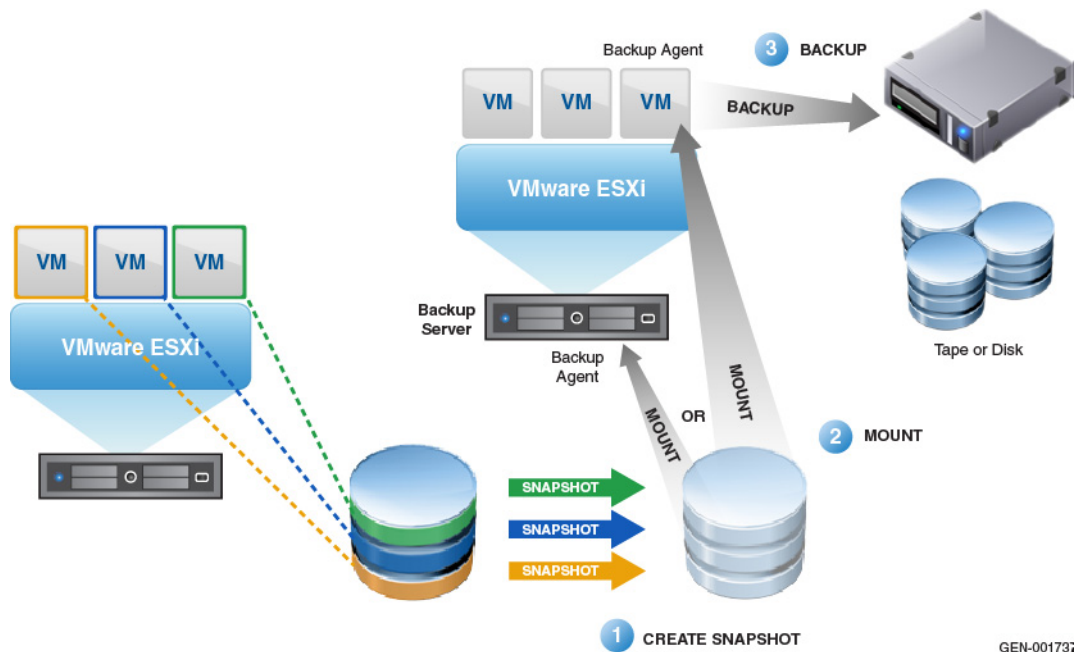


Figure 114 VADP flow diagram

GEN-001731

Backup and recovery using VMware Data Protection

VMware Data Protection (VDP) 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 Protection 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 VDP backups. Each virtual machine backup is stored on a target disk in a deduplicated store.

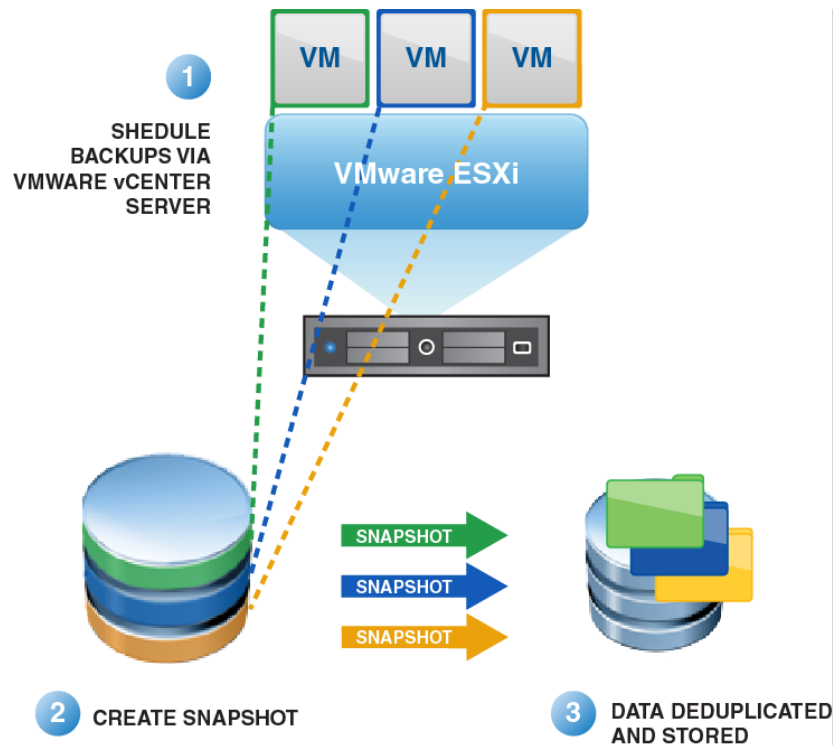


Figure 115 VMware Data Recovery

VMware Data Protection process

1. During the backup, VDP takes a snapshot of the virtual machine and mounts it directly to the VDP virtual machine.
2. The VDP streams blocks of data to the destination storage as shown in [Figure 116](#).
3. During this process, VDP uses the VADP CBT functionality on ESXi hosts to identify the changed blocks and minimize the amount of data to be backed up.
4. VDP deduplicates the stream of data blocks to further eliminate redundant data prior to writing the backup to the destination disk.
5. The deduplicated store creates a virtual full backup based on the last backup image and applies the changes to it.
6. When all the data is written, VMware Data Protection dismounts the snapshot and takes the virtual disk out of snapshot mode.

VMware Data Protection supports only full and incremental backups at the virtual machine level, and does not support backups at the file level. Both VDP and its more sophisticated version, vSphere Data Protection Advanced (VDPA), are based on EMC2 Avamar and are co-developed with EMC.

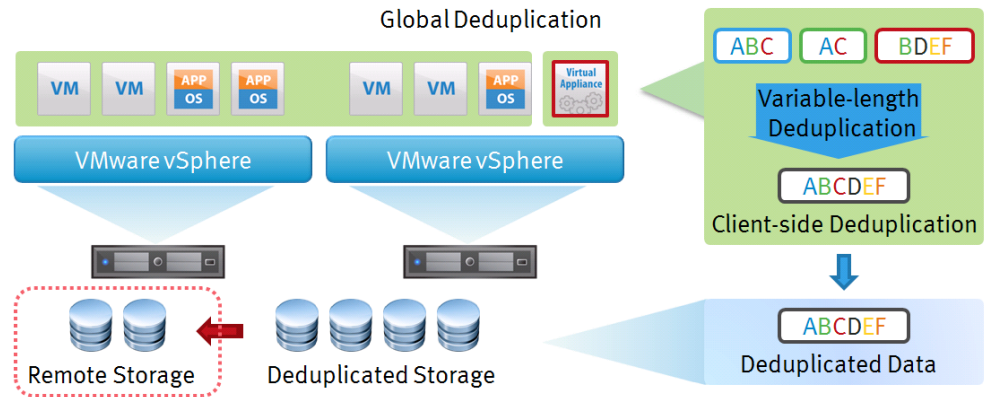


Figure 116 vSphere Data Protection

VDP and VDPA feature comparison

Table 17 provides a feature comparison between VDP and VDPA to help you decide which product is best suited to protect the vSphere infrastructure in your particular environment.

Table 17 VDP and VDPZ feature comparison

Feature	VDP	VDPA
Max # virtual machines	100	400
Appliances per vCenter	10	10
Max backup capacity	2 TB	8 TB
Variable length deduplication	✓	✓
Global deduplication	✓	✓
In-guest deduplication	✗	✓
Change Block Tracking	✓	✓
CBT Restore	✓	✓
Microsoft Exchange	✗	✓
Exchange Single Mailbox Restore	✗	✓
Microsoft SharePoint	✗	✓
Microsoft SQL	✗	✓
Flexible schedule and retention	✗	✓
Replication topologies	1:1	1:1, 1:n, n:1
Replication targets	Avamar	VDPA, Avamar, Data Domain
Encrypt replication	✓	✓

Table 17 VDP and VDPZ feature comparison

Feature	VDP	VDPA
Backup and restore of individual vmdk file	✓	✓
In-guest restore of files (Windows and Linux)	✓	✓
Emergency restore to an ESX- host without vCenter	✓	✓
Flexible storage options	✗	✓
Mount existing BDP datastores to another appliance	✗	✓
Automated backup tests	✗	✓
vSphere Integration	✓	✓
One-step recovery	✓	✓

Local data protection with vSphere Data Protection

VDP is targeted at small businesses and branch or remote offices supporting up to 100 virtual machines, depending on their size.

You must deploy VDP (or VDPA) as a virtual machine in vCenter. It is then managed through vCenter or the vSphere web client. Apart from emergency restores, all actions are managed within vCenter.

VDP Best Practices

EMC recommends the following best practices for VDP:

- ◆ Validate that all virtual machines are on hardware revision 7 to allow for changed block tracking (CBT) backups and restores.
- ◆ Install VMware Tools on every virtual machine to enable the benefit of File Level Restore (FLR).
- ◆ Use VMware Tools for time synchronization.
- ◆ Use a dedicated VMFS5 datastore as a target for the virtual disks to hold your backup data.

VDP installation prerequisites

The following prerequisites must be met before installing VDP:

- ◆ Validate that all ESXi-hosts are time synchronized with vCenter.
- ◆ Register VDP in DNS for forward and reverse lookup.
- ◆ Use a dedicated VMFS Datastore to store VDPS virtual disks (3 to 12 disks with capacities from 256GB to 1.024 GB).
- ◆ For the appliance, identify a fixed IP address and 100 GB capacity, which can be thin provisioned.

- ◆ Plan the capacity needed to protect your virtual machines carefully as these settings are final and cannot be changed in VDP. The only option to change it later on is to upgrade to VDPA.

Table 18 Capacities for VDP deployments

Big Data appliance	CPU	RAM	Hard disk
0.5 TB	4	4.096 MB	3 x 256 GB
1.0 TB	4	4.096 MB	3 x 512 GB
2.0 TB	4	6.144 Mb	3 x 1024 GB

Installing VDP

1. Download the VDP or VDPA OVA file from VMware.
2. Deploy the OVA file.

Recommendations:

 - Use a dedicated VMFS Datastore based on a single LUN of a dedicated Storage Pool to store the VDP data disks. For the storage pool physical disks choose NL-SAS.
 - Format the VDP data disks with "Eager Zero Thick."
3. When the installation completes, a web url is provided with the address of the system. Type the url into a supported web browser to configure the system. Use the credentials **root** and **changeme**.
4. The configuration wizard guides you through the following steps:
 - a. Network configuration
 - b. Time zone
 - c. Password for the appliance (root)
 - d. Register vCenter
 - e. VDP license
 - f. Create Storage
 - g. Allocate devices
 - h. CPU & memory
5. Validate the network settings, especially **Hostname** and **Domain**.
6. In **Create Storage**, select the size of your VDP deployment .

Ensure that you provide the correct settings, because the settings cannot be changed after the configuration is complete.

- Optionally, clear the **Store with appliance** checkbox, to choose a different datastore.

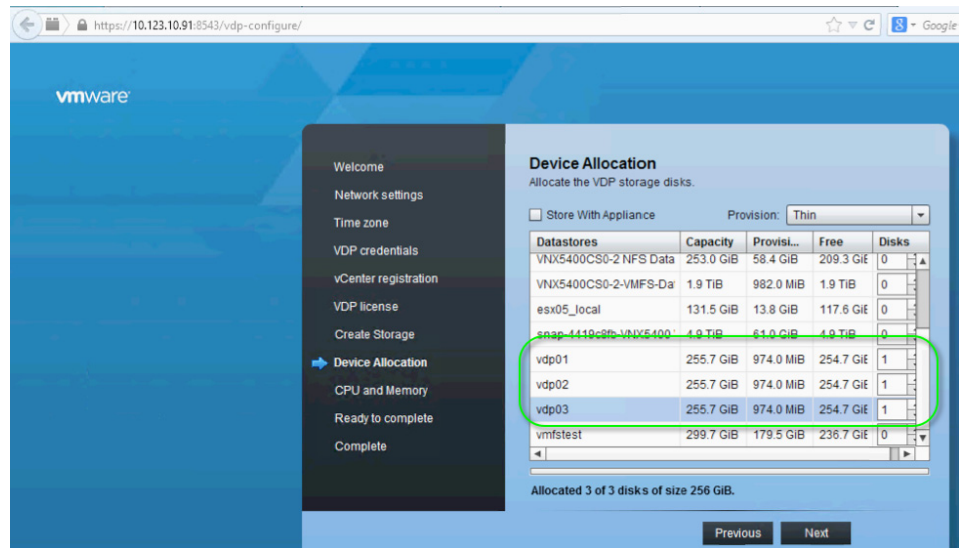


Figure 117 Allocation of virtual disks on one or more datastores

A dialog box appears indicating the resources needed for CPU and Memory.

Note: Changing these values can adversely affect performance.

- Optionally, you can run a performance check.

After the check completes it takes VDP approximately 30 minutes to be ready for production.

- Log on to vSphere Web Client and click **vSphere Data Protection** in the left pane.

The **Getting Started** menu appears, as shown in [Figure 118](#).



Figure 118 vSphere Getting Started menu

vSphere Data Protection functions

The following table describes the functions available in the vSphere Data Protection tabs.

Table 19 vSphere Data Protection tabs

Tab	Function/Description
Backup	Manage backup of virtual machines: Create, clone, edit, enable, disable, and delete a backup job.
Restore	Manage restoring of virtual machines.
Replication	Manage replication sessions with Avamar. See “Replication Manager” on page 154 for more information.
Reports	View all protected virtual machines including status, assigned backup jobs, and replication.
Configuration	<ul style="list-style-type: none"> • Adjust the backup and maintenance window to your needs. • Enter license keys. • Change to VDPA (requires a license key). • Submenus: <ul style="list-style-type: none"> – Log: View log files. – Email: Receive reports from VDP. – Settings: Conduct an integrity check of the virtual disks holding the backup data.

Creating a backup job

1. In the vSphere navigation pane, select **vSphere Data Protection**.
2. Click **All Actions** > **Create a new backup job**.

The **Create a new backup job** wizard appears.

3. For **Data Type**, select the appropriate data type.
4. For **Backup Targets**, choose one of the following:
 - **Full Image**—Backs up all virtual disks and virtual dependent RDMs contained in a virtual machine.
 - **Individual Disks**—Backs up selected virtual disks or dependent virtual RDMs. This method is useful for adhering to SLO, RPO, and RTO definitions.

Note: The restore job is not affected because you can select virtual disks, as shown in [Figure 119](#). However, **Individual Disks** presents only the virtual disks you select and might not contain a boot disk.

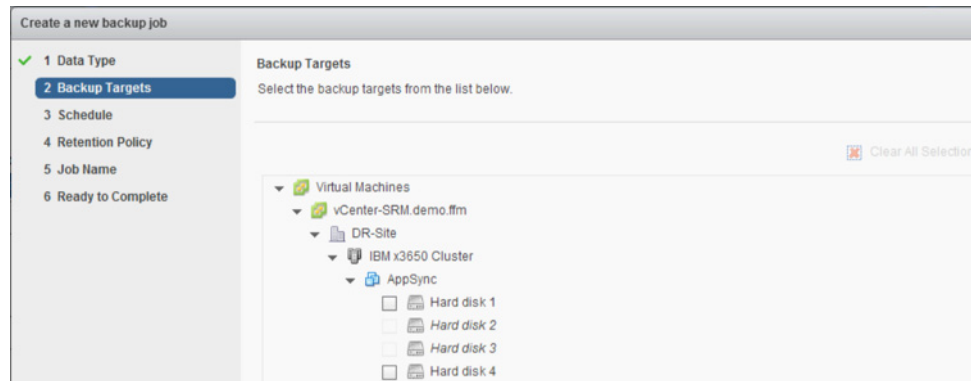


Figure 119 Backup with selected virtual disks

- For **Schedule**, configure when backups occur. This example uses the default, **Daily**.

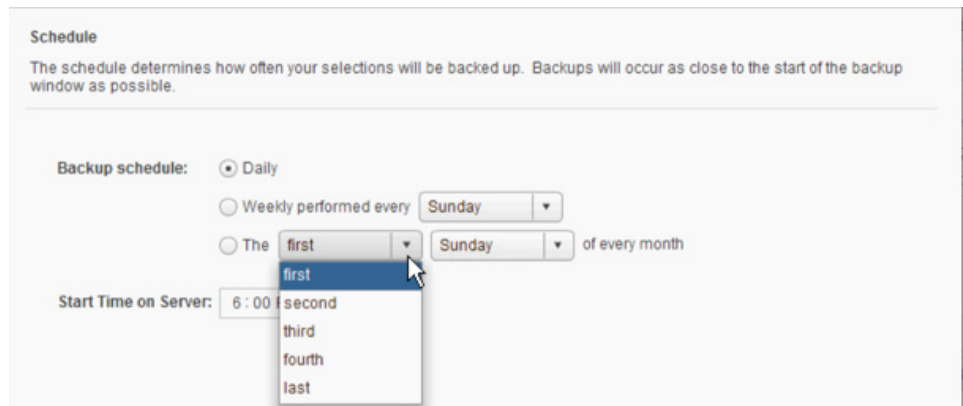


Figure 120 VDP backup schedule

- For **Retention Policy**, select the parameters required by the SLO of the application to be backed up.

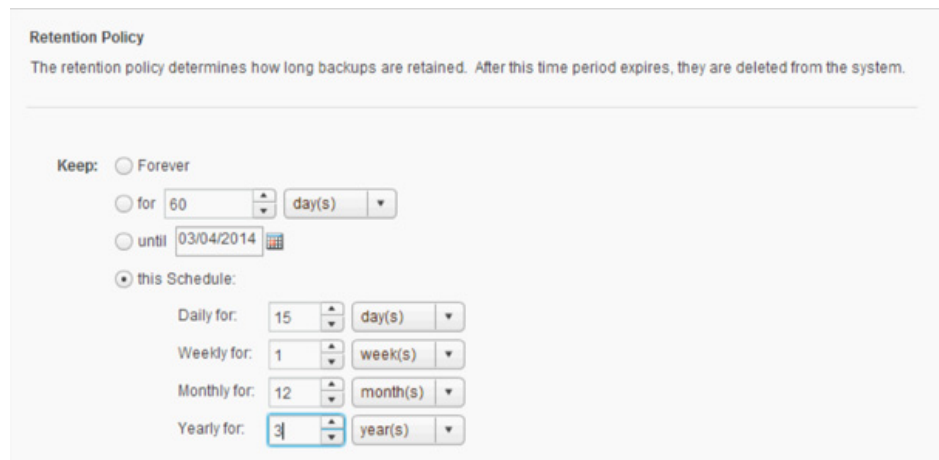


Figure 121 Retention policy form

- For **Job Name**, type a name for the backup job and confirm your settings.
- Optionally, select **Backup Now** to immediately start the backup job.

- In **Ready to Complete**, confirm your selections and click **Finish**.

Local Data Protection with vSphere Data Protection Advanced

You can easily upgrade VDP to vSphere Data Protection Advanced (VDPA). The Advanced edition features the following benefits:

- ◆ Application-consistent protection
- ◆ Automated backup verification
- ◆ Advanced data replication features

Upgrading to vSphere Data Protection Advanced

To upgrade to VDPA, complete the following steps:

- In the vSphere navigation pane, select **vSphere Data Protection**.
- Under **All Actions**, select **Add a VDP Advanced License**.
- In **VDP Advanced License Assignment**, click **Settings** > **Add a New License Key**.
- Type the license key provided by VMware and click **Decode**.
- Click **OK** to accept the message that appears.
- Under **Backup appliance details**, select **Switch to VDPA**.
- Click **OK** to accept the warning.

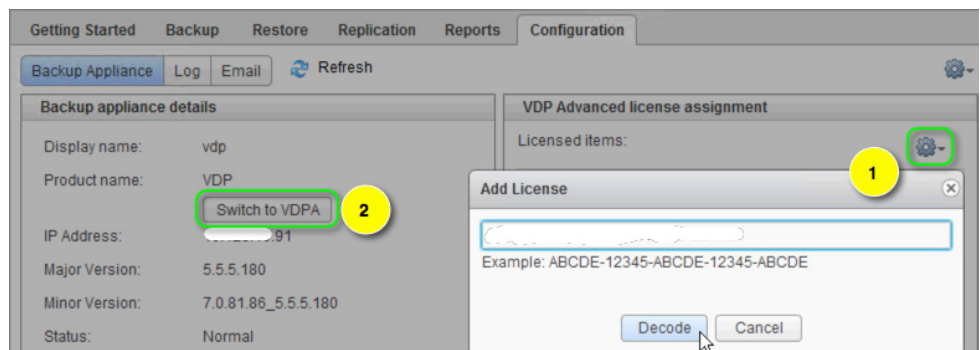


Figure 122 License Key window

- After you have successfully upgraded to VDPA you must reconnect to the appliance.
- Click **Configuration** > **Settings** > **Edit License Assignment** to assign the license key to the ESXi hosts.

VDPA benefits

With VDPA you can upgrade your appliance to up to 8 TB of total capacity. [Table 20](#) provides sizing guidelines. As the capacity grows, more compute resources are needed for the appliance.

Table 20 Resources for VDPA-Deployments

Big Data appliance	CPU	RAM	Hard disks
0.5 TB	4	4.096 MB	3 x 256 GB
1.0 TB	4	4.096 MB	3 x 512 GB
2.0 TB	4	6.144 MB	3 x 1024 GB
4.0 TB	4	8.192 MB	6 x 1024 GB
6.0 TB	4	10.240 MB	9 x 1024 GB
8.0 TB	4	12.288 MB	12 x 1024 GB

VDPA provides application-consistent backups. Click the **Configuration** tab to view the agents for deployment, as shown in [Figure 123](#).

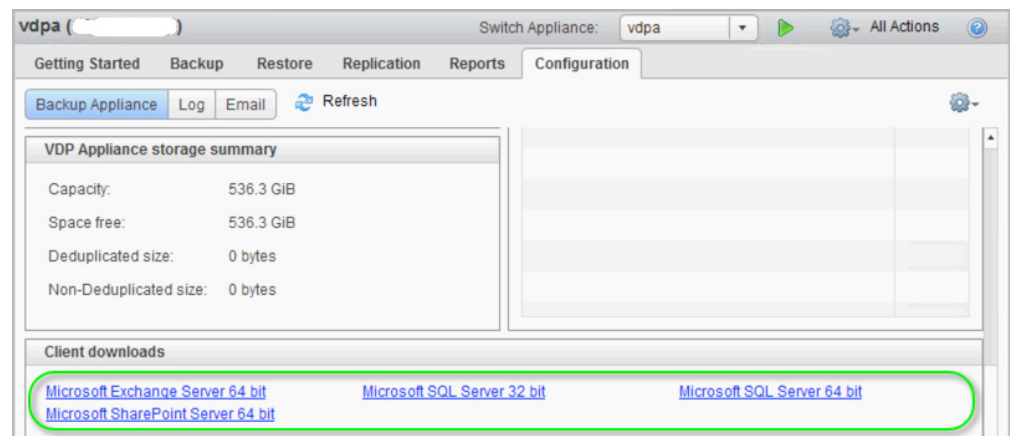


Figure 123 In-guest agents

Configuring replication sessions

Both VDP and VDPA allow you to create replication jobs.

- ◆ VDP supports Avamar only.
- ◆ VDPA supports various replication options, including replicating to a Data Domain system and to another VDPA instance within your data center.

Data replication is encrypted and uses TCP port 29000, exclusively.

Configure and size your target according to your requirements. Creating a replication job takes only a few minutes.

Replicating to an Avamar datastore

1. In the vSphere navigation pane, select **vSphere Data Protection**.
2. Click **Replication**.

- From the menu, select **Replication Job Actions > New**.

The **Create a new replication job** wizard appears. Virtual machines with a valid backup are displayed.

- Select one of the following:
 - All Clients
 - Select clients individually

Note: To meet application SLOs that might be different for some virtual machines, choose **Select clients individually**. You can choose different schedules later in the wizard.

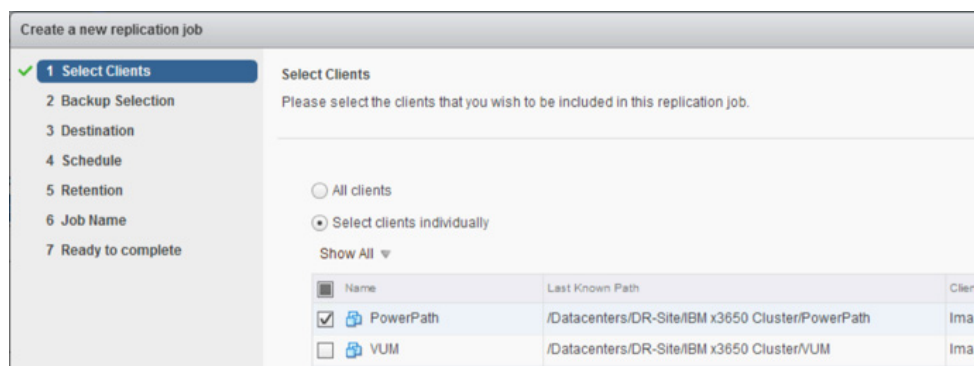


Figure 124 Creating a new replication job

- For **Backup Selection**, select the backup types that you want to replicate, the maximum number of backups to replicate, and the date range limits.

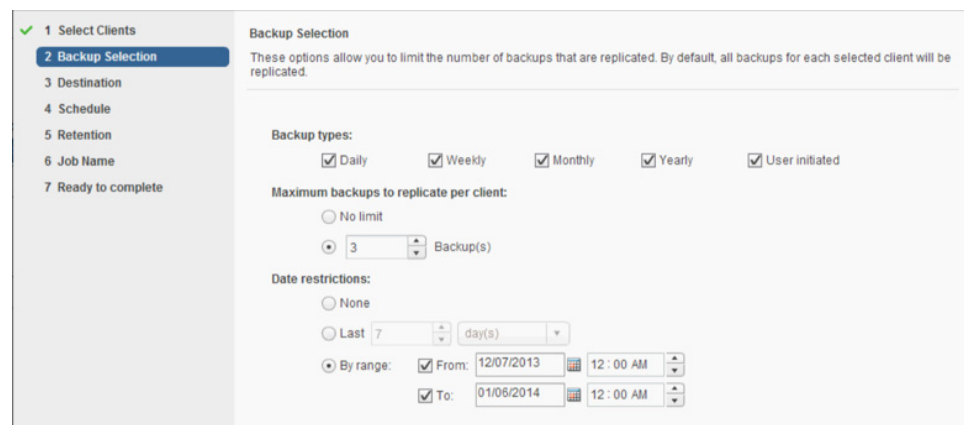


Figure 125 Replication backup schedule

The **Destination** window identifies an Avamar datastore as a replication target.

- Type the Avamar default user, **repluser**, and the password.

7. Click **Verify authentication** to verify the connection, and then click **OK**.

The screenshot shows the 'Destination' configuration page in the Avamar interface. On the left, a navigation pane lists steps 1 through 7, with '3 Destination' highlighted. The main area is titled 'Destination' and contains the instruction: 'Please specify a destination to replicate your backups to.' Below this are four input fields: 'Hostname or IP:' (with a partially filled field ending in '.51'), 'Port:' (with '29000'), 'Username:' (with 'repluser'), and 'Password:' (with masked characters). A 'Verify authentication' button is at the bottom. An 'Info' dialog box is overlaid on the right, showing the message 'Authentication succeeded.' and an 'OK' button.

Figure 126 Avamar replication target

8. For **Schedule**, select the parameters to define when the data replication will occur. It is best to run the replication during off-peak hours.

The screenshot shows the 'Schedule' configuration page. The left navigation pane has '4 Schedule' highlighted. The main area is titled 'Schedule' and contains the instruction: 'The schedule determines how often your selections are replicated and at what time of day they are to be replicated.' Below this are three radio button options for 'Replication schedule:': 'Daily' (selected), 'Weekly performed every' (with a 'Sunday' dropdown), and 'The first' (with a 'Sunday' dropdown) 'of every month'. At the bottom, there is a 'Start time on server:' field set to '10:00 PM' with a time selector.

Figure 127 Scheduling the replication session

9. For **Retention**, select the expiration parameters for the retention policy.

The screenshot shows the 'Retention' configuration page. The left navigation pane has '5 Retention' highlighted. The main area is titled 'Retention' and contains the instruction: 'These options define when the replicated backups will expire from the destination.' Below this are three radio button options: 'Keep the current expiration for each backup', 'Keep forever', and 'Set expiration by backup type:' (selected). Under the selected option, there are five rows of settings: 'Daily backups expire in:' (30 days), 'Weekly backups expire in:' (4 weeks), 'Monthly backups expire in:' (12 months), 'Yearly backups expire in:' (3 years), and 'User initiated backups expire in:' (15 days). Each row has a numeric input field and a unit dropdown menu.

Figure 128 Retention policy for the Avamar target

10. In **Job Name**, type a name for the replication job.
 11. In **Ready to complete**, verify your settings and click **Finish**.

The job starts at the time you specified. Alternatively, you can start it manually.

Recovering data with VDP

Every copy of a virtual machine created with VDP is considered a full backup. However, after the first backup only incremental changes are applied. A restore action is always applied to a full backup. Backing up to disk with deduplication technology provides the following benefits:

- ◆ Eliminates the need to apply incremental backups after a full restore.
- ◆ In VMware environments, CBT is optimized for both backup and restore to significantly improve backup and restore times.

Best practices for restoring a virtual machine:

- ◆ Remove all snapshots of the affected virtual machine prior to the restore action. [VMware-Knowledgebase-article 1025279](#) provides more information.
- ◆ If you want to restore to its original location, power off the corrupted machine.
- ◆ If you restore to another location, you can keep the source virtual machine powered on, but the benefit of a CBT restore is not available.

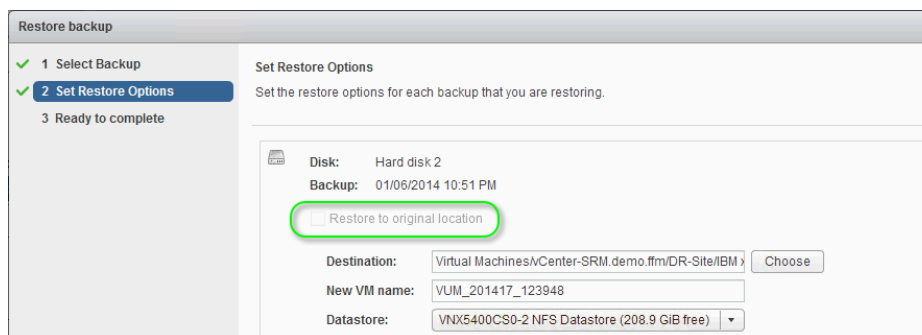


Figure 129 Restore of a powered on virtual machine

You can restore a virtual machine in the following ways:

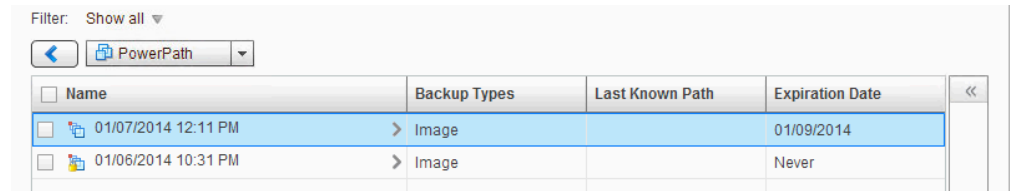
- ◆ Right-click the virtual machine object in the web client and select **VM All VDP 5.5 Actions > Restore Rehearsal**.
- ◆ Use the VDP Appliance-Plug-in in the Web Client
- ◆ Use the VDP Restore Client (for FLR only). Open a browser window and type the URL in the following format: `https://<VDP_IP_address>:8543/flr/`

You can restore virtual machines, virtual machine images, or individual disks.

Restoring a virtual machine using CBT

1. Ensure that your virtual machine is powered off.
2. From the vSphere Web Client select **vSphere Data Protection**.
3. In **Getting Started**, select **Restore Backup > Restore**.
4. In the **Restore** tab, select the check box of the virtual machine to restore.

Use **Filter** to narrow your search, as shown in [Figure 130](#).



Name	Backup Types	Last Known Path	Expiration Date
01/07/2014 12:11 PM	Image		01/09/2014
01/06/2014 10:31 PM	Image		Never

Figure 130 Restoring a virtual machine

- Click **Restore**, as shown in [Figure 131](#).

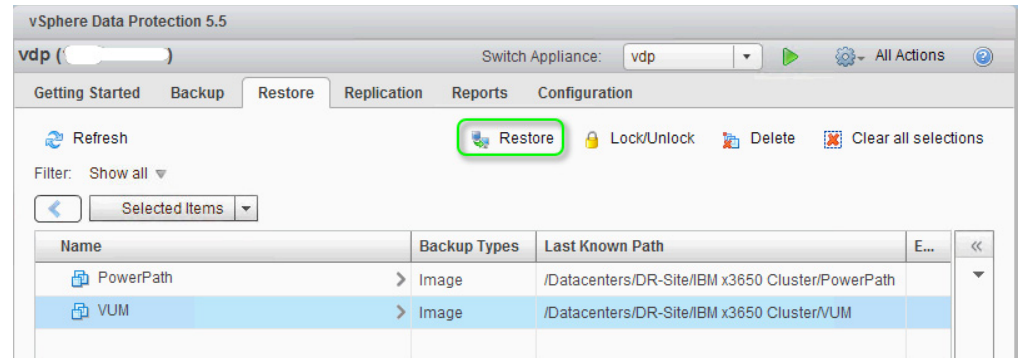


Figure 131 Restoring a backup with VDP

- In the following window, select **Backup** and click **Next**.
- Ensure that **Restore to Original Location** is selected.
- Select **Power On** and **Reconnect NIC**, as shown in [Figure 132](#), to have vCenter power on the virtual machine and connect to the network after the restore completes.

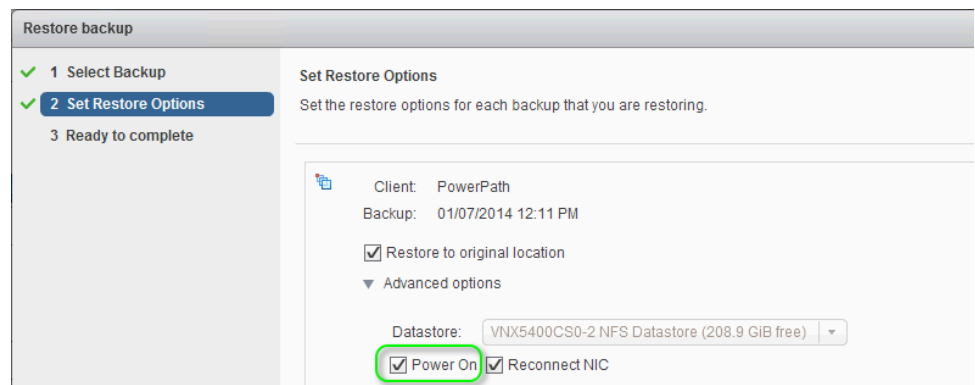


Figure 132 Setting Restore Options

- In the **Ready to Complete** dialog box, click **Finish**.
- To monitor the restore process, select **Reports** > **Task Console**.

Restoring a virtual machine image

When restoring a virtual machine image, the machine can be powered on and connected to the network. This procedure does not use CBT.

- From the vSphere Web Client select **vSphere Data Protection**.

2. In **Getting Started**, select **Restore Backup > Restore**.
3. In the **Restore** tab, select the virtual machine to restore.
Use **Filter** to narrow your search.
4. Click **Restore**.
5. In the dialog box that appears, select **Backup** and click **Next**.
6. Clear the **Restore to Original Location** check box.
7. For **Destination**, click **Choose**.
8. In the dialog box that appears, select an ESXi-Host, a vSphere-Cluster, or a vApp for the restore destination, as shown in [Figure 133](#).



Figure 133 Selecting the restore destination

9. Under **Advanced Options**, select a datastore.
10. Select **Power On** and **Reconnect NIC** to have vCenter power on the virtual machine and connect to the network after the restore completes.

[Figure 134](#) depicts the **Set Restore Options** for restoring a virtual machine image.

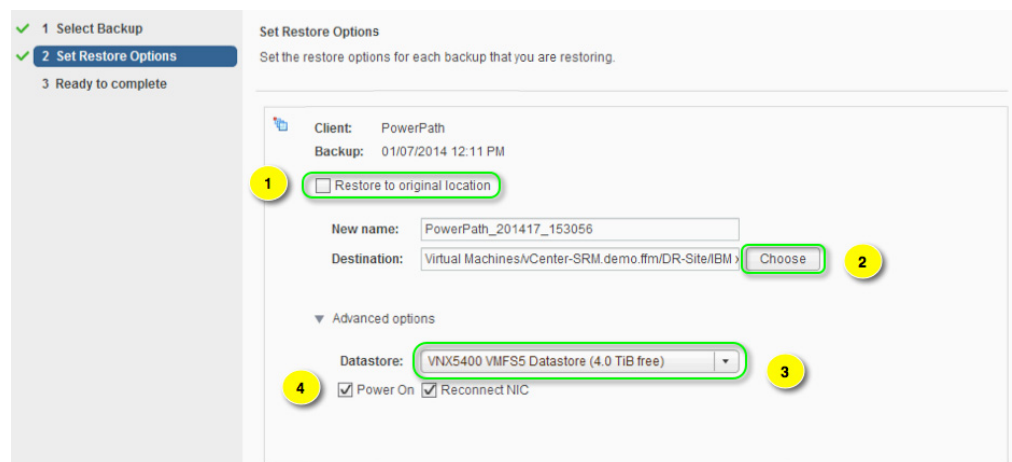


Figure 134 Set Restore Options dialog box

11. Click **Next**.
12. In the **Ready to Complete** dialog box, click **Finish**.

Restoring an individual disk using CBT

1. Ensure that your virtual machine is powered off.
2. From the vSphere Web Client select **vSphere Data Protection**.
3. In **Getting Started**, select **Restore Backup** > **Restore**.
4. Double-click the virtual machine to restore.

The **Backup Type** changes from **Image** to **Disk**.

5. Select the check box of the individual virtual disk to restore, as shown in [Figure 135](#).

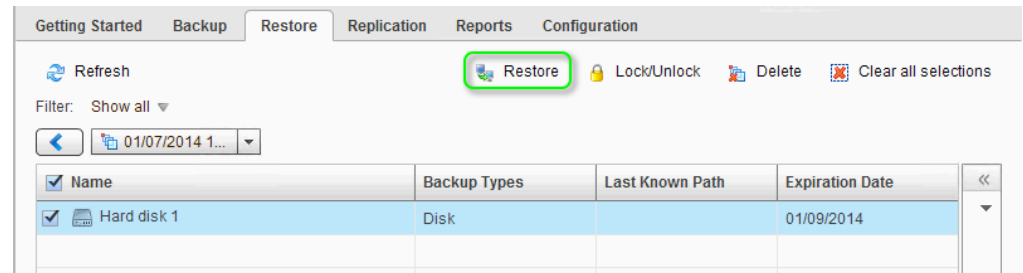


Figure 135 Selecting an individual disk

6. Click **Restore**.
7. Click **Next**.
8. In the **Set Restore Options** dialog box, clear the **Restore to Original Location** check box.
9. For **Destination**, click **Choose**, as shown in [Figure 136](#).



Figure 136 Setting the Restore Options for restoring an individual disk

10. Select an ESXi-Host, a vSphere Cluster, a vApp, or a virtual machine for the restore destination, as shown in [Figure 137](#).

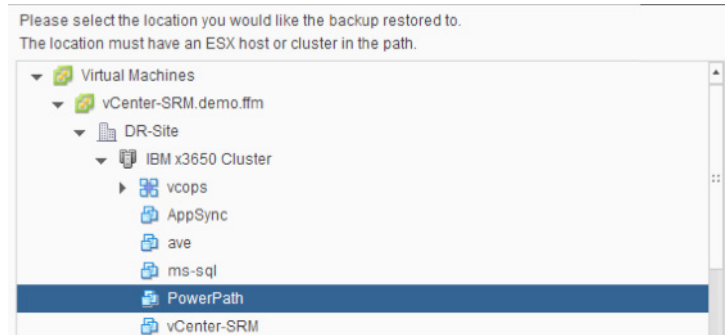


Figure 137 Selecting the restore destination

11. Click **OK**.
12. In the **Ready to Complete** dialog box, click **Finish**.

After recovery, the virtual disk appears as a new SCSI device in the selected virtual machine, as shown in [Figure 138](#).

13. To monitor the restore process, select **Reports > Task Console**.

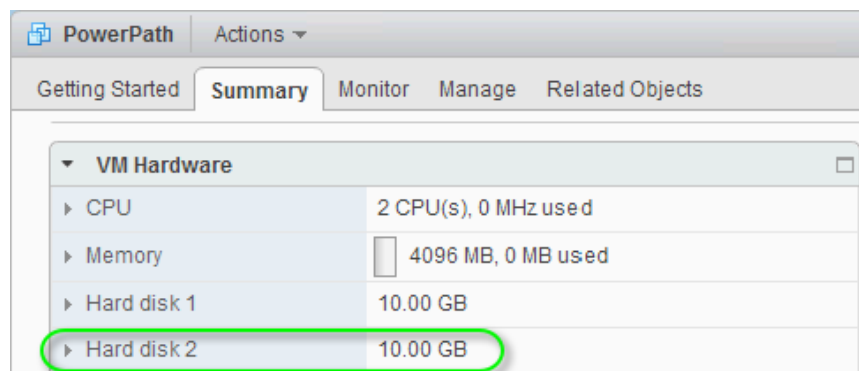


Figure 138 A virtual disk after a restore

Recovering file-level data

Both VDP and VDMA support in-guest file or directory restore of hosted Windows and Linux platforms and dependent vRDMs.

1. Log into the guest OS or a guest OS of the same make from which you want to recover data.
2. Open a browser window and type the URL in the following format:
`https://<VDP_IP_address>:8543/flr/`
3. Choose one of the following login options, as shown in [Figure 139](#):
 - **Basic Login:** On the same host you want to restore data to, log in as local administrator.
 - **Advanced Login:** On another host of the same make, log in as local administrator and vCenter Administrator.

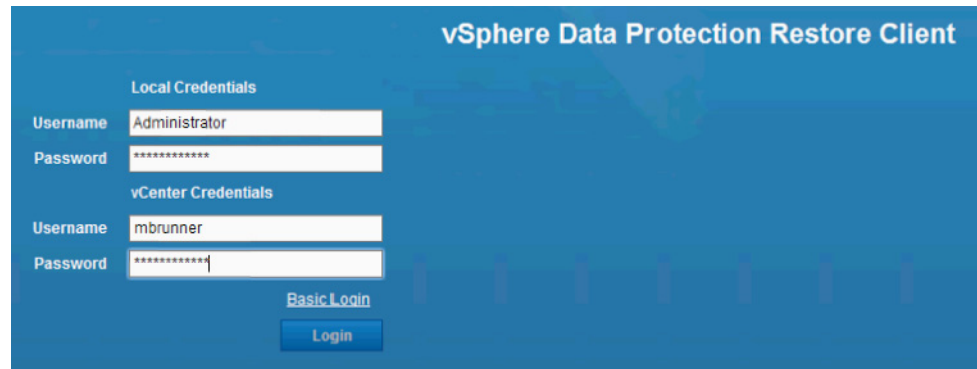


Figure 139 Login options for the vSphere Data Protection Restore Client

4. In the **Manage mounted backups** dialog box, shown in [Figure 140](#), select the virtual disks from which you want to recover files.
5. Click **Close**.

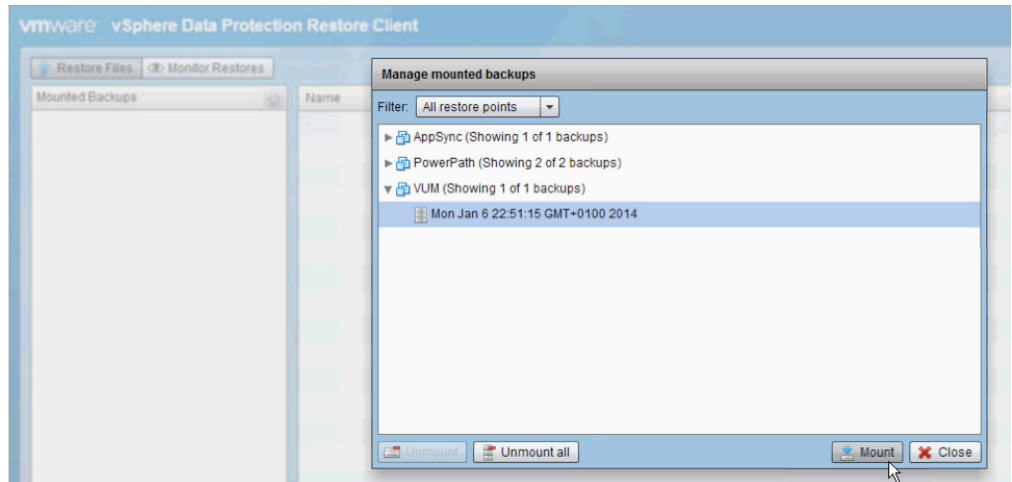


Figure 140 Selecting individual disks from which to restore files

6. Browse to the files you want to recover, select them and click **Restore selected files**.
7. Choose a destination for the selected files.

8. Click **Restore** to recover the selected files and/or directories, as shown in [Figure 141](#).

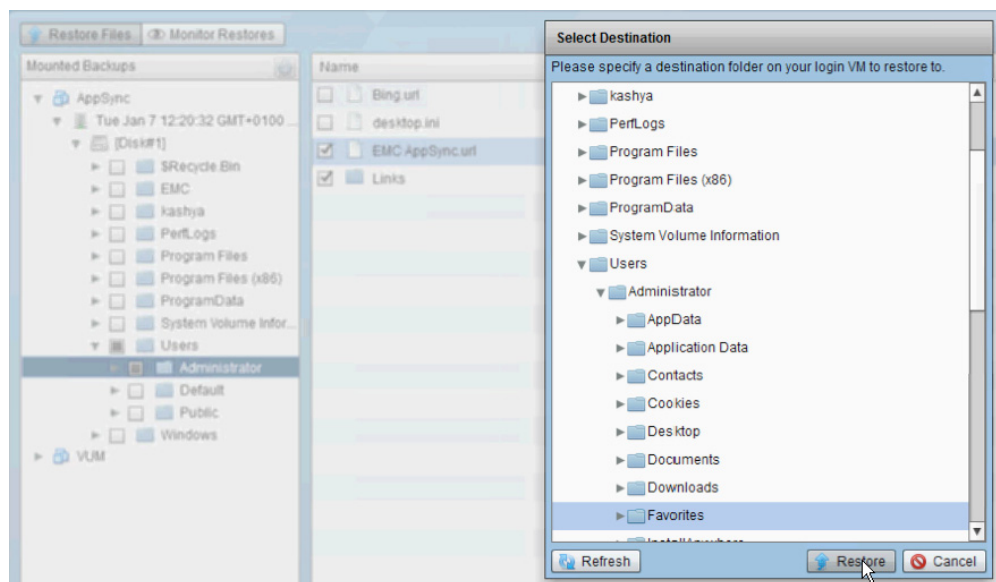


Figure 141 Restoring selected files or directories

9. Click **Monitor restores** to monitor the restore process, as shown in [Figure 142](#).

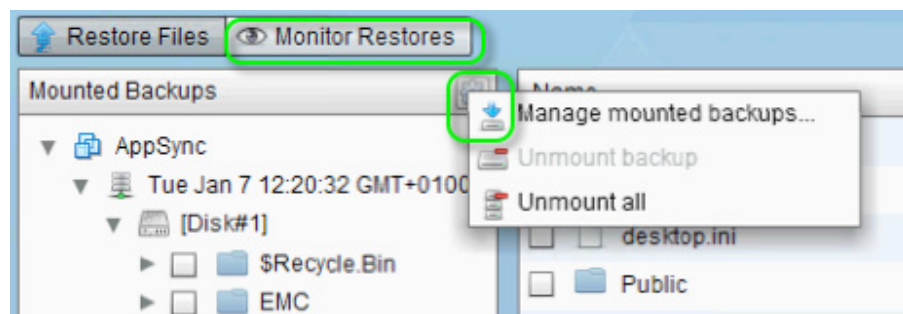


Figure 142 Monitor Restores menu

10. Unmount the drives by selecting **Settings > Mounted backups > Unmount all**.

11. Log off.

VDP Emergency Restore

VDP runs independently of vCenter, which enables you to restore the vCenter itself.

Note: Ensure that the ESXi host is not connected to the vCenter or that the vCenter is powered off.

1. Open a browser window and type the URL in the following format:
`https://<VDP_IP_address>:8543/vdp-configure/`
 and log in as **root**.
2. Click **Emergency Restore**.
3. Select the virtual machine to be restored.

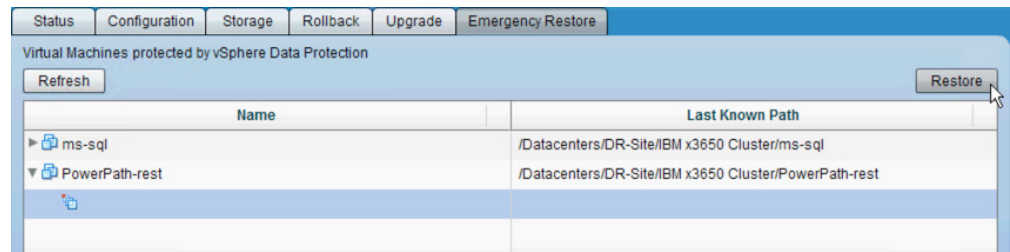
4. Click **Restore**.

Figure 143 Performing a VDP Emergency Restore of a virtual machine

5. In the **Host Credentials** dialog box, type the credentials for the target ESXi host where the virtual machine is registered.

Host Credentials

Enter information about the ESX host on which this VDP resides.

ESX hostname or IP: 192.168.0.208

Port number: 443

Username: root

Password: *****

Figure 144 Performing an emergency restore to an individual ESXi host

6. Click **OK** to restore the virtual machine.

Backup and recovery using Avamar

EMC Avamar® is a backup and recovery software product that 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 enables it to use 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 145 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 sample 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 case, the virtual machines do not require an Avamar agent.

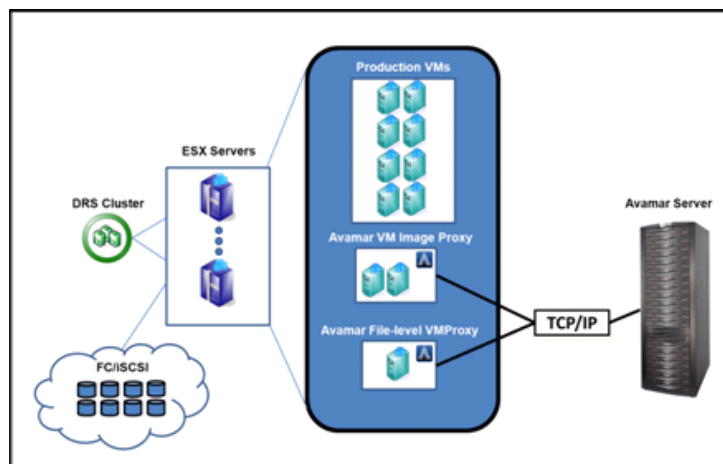


Figure 145 Sample Avamar environment

Backing up data using Avamar

Avamar provides the following backup options for vSphere environments:

- ◆ **File Level Backup**—File level backups are enabled by installing the Avamar client within 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 146](#) shows a sample proxy configuration.

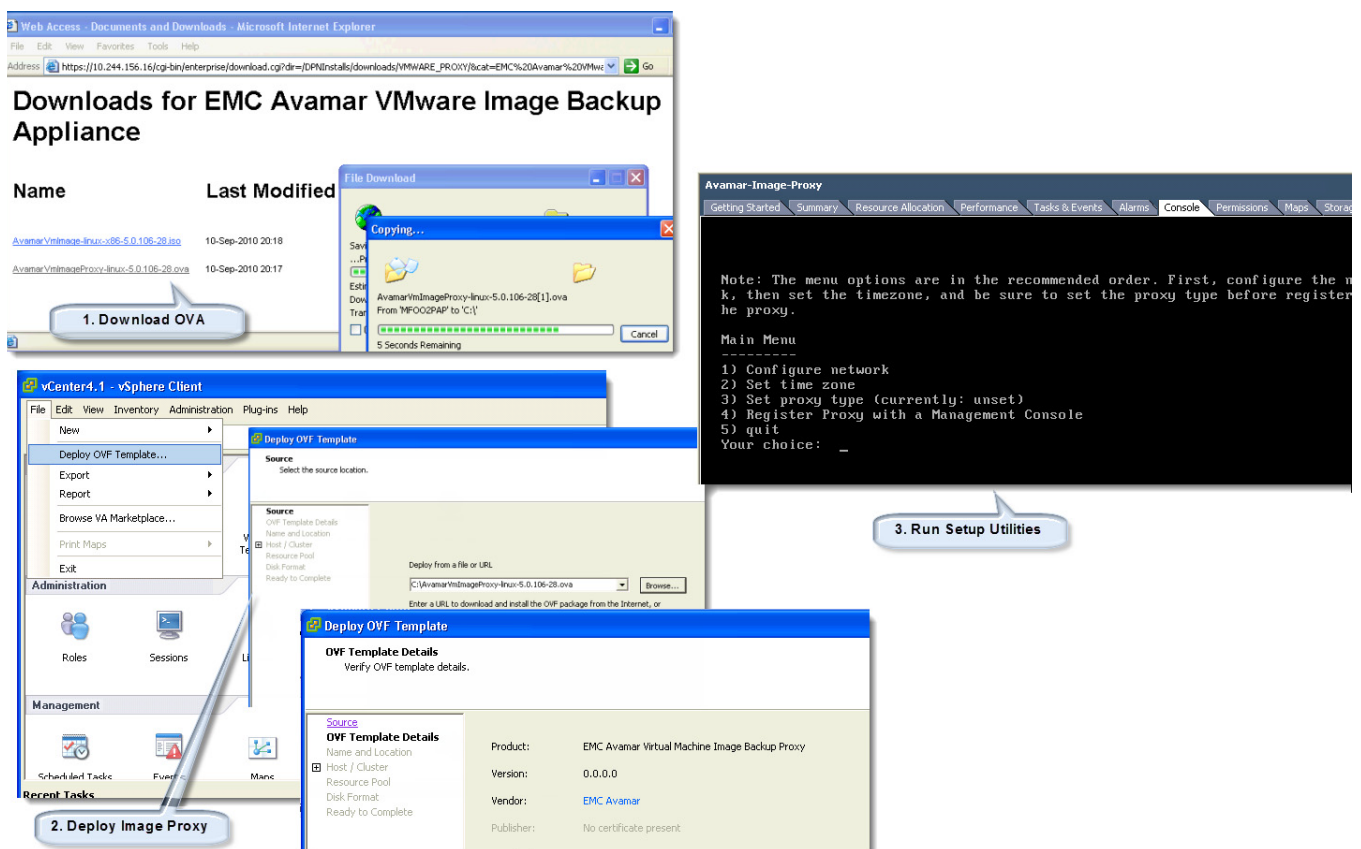


Figure 146 Sample proxy configuration

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 147 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.

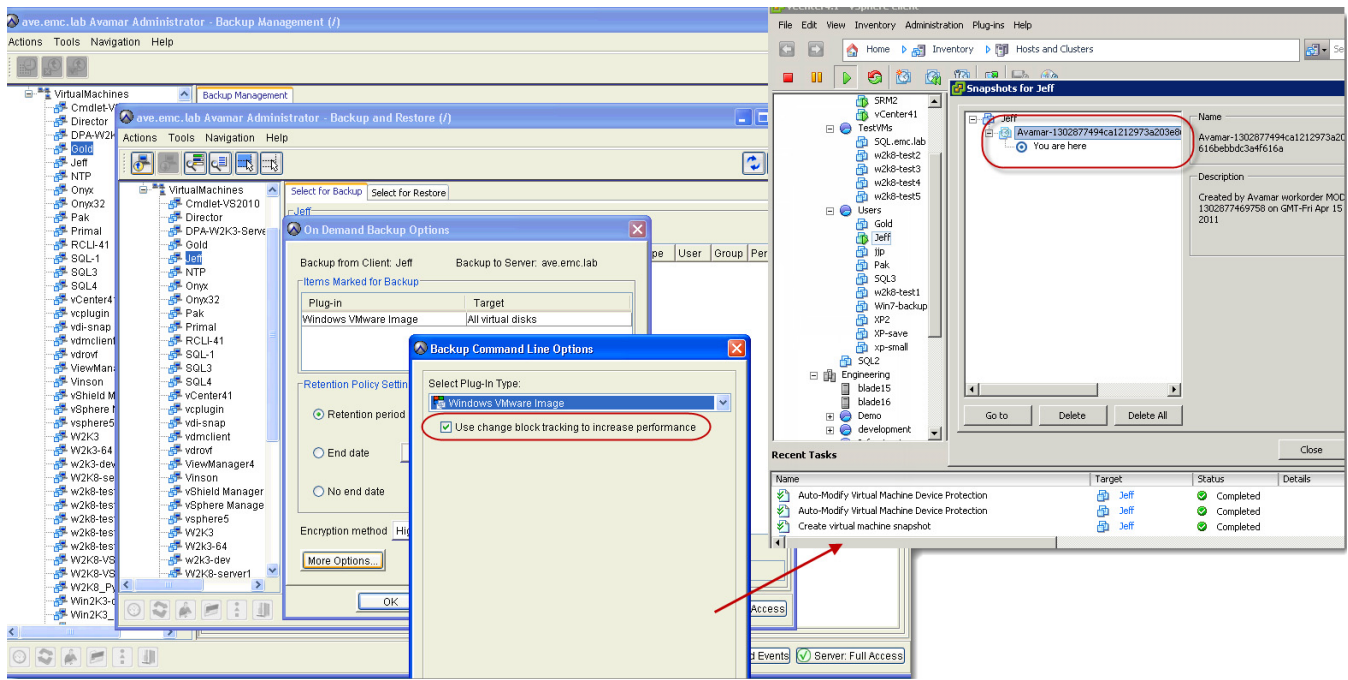


Figure 147 Avamar backup management configuration options

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.

Recovering data using Avamar

Avamar 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

Restoring a virtual machine image

Avamar Image Proxy

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 148](#) 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.

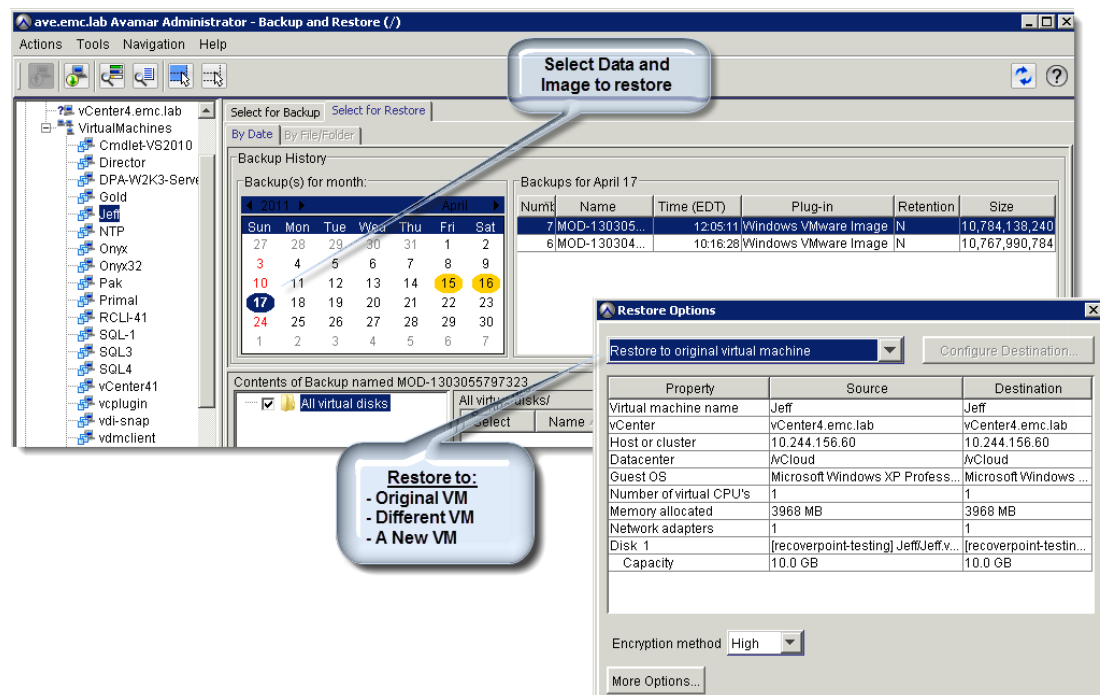


Figure 148 Avamar virtual machine image restore

Avamar file-level proxy

An Avamar file-level recovery (FLR) 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 uses 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 FLR 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 149](#).

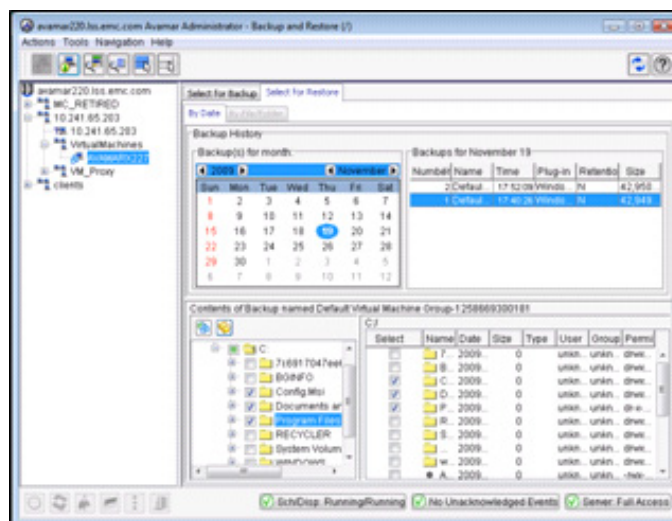


Figure 149 Avamar browse tree

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 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 running virtual machine tools.

Best practices for 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 backups 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 might 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 create flexible backup solutions to improve backup processes, reduce backup windows, and reduce the amount of space required to store backup images.

Figure 150 shows the virtualization topology in an environment with NetWorker.

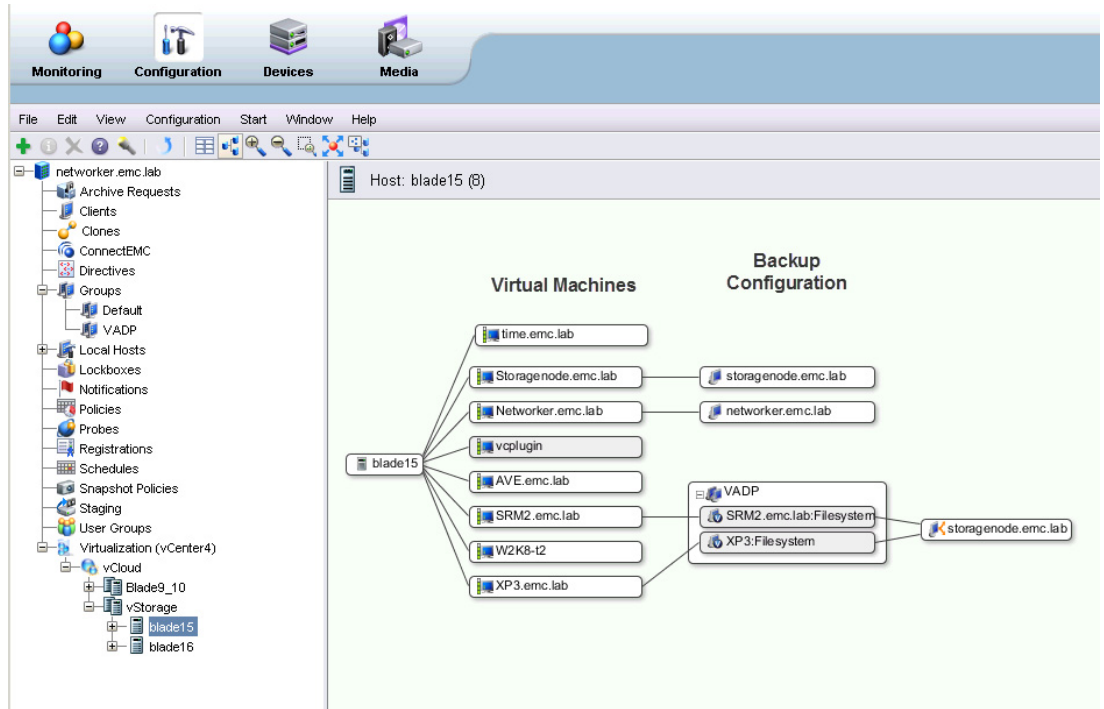


Figure 150 NetWorker-virtualization topology view

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 151.

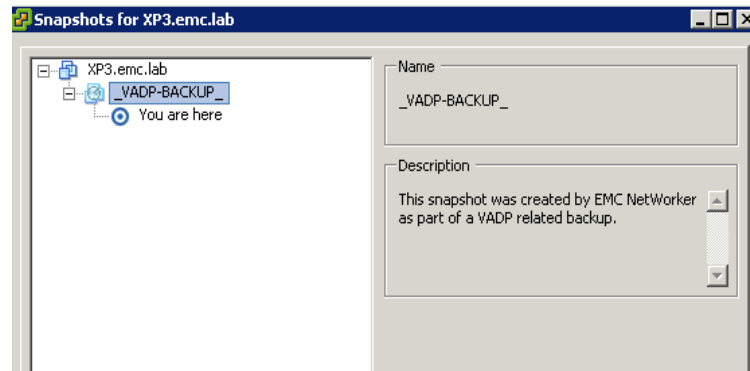


Figure 151 Viewing a VADP snapshot

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.

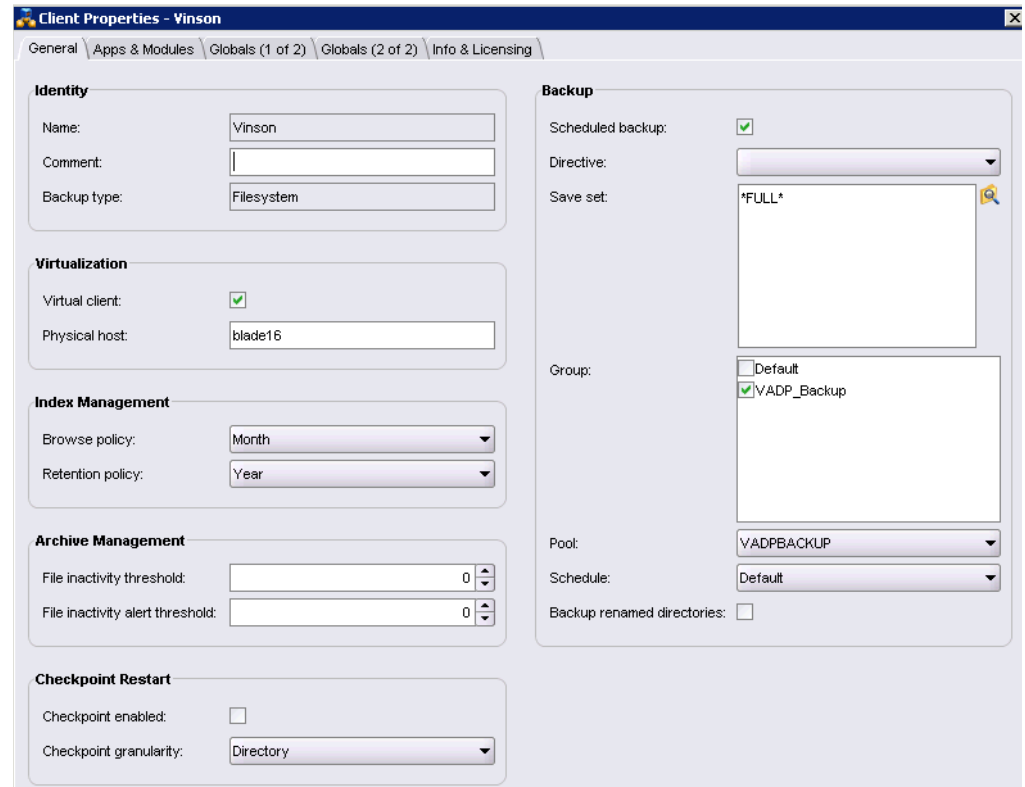


Figure 152 NetWorker configuration settings for VADP

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).

Using NetWorker to backup and restore VNX NAS file system NDMP

NetWorker provides two methods of storage integration with VNX NFS datastores. You can use VNX file systems as

- ◆ Advanced File System Type Devices (AFTD)
 - ◆ Virtual Tape Library Unit (VTLU)
1. Configure a VTLU on the VNX file system.
 2. Configure NetWorker as an NDMP target to back up NFS datastores on the VNX platform.
 3. Configure NetWorker to use VNX File System Integrated Checkpoints to create NDMP backups as follows:
 - a. Create a Virtual Tape Library Unit (VTLU) on VNX NAS storage.
 - b. Create a library in EMC NetWorker.
 - c. Configure NetWorker to create a bootstrap configuration, backup group, and a backup client.
 - d. Run NetWorker backup.
 - e. Execute NetWorker Recover.

The entire datastore or individual virtual machines are available for backup or recovery. [Figure 153](#) shows NetWorker during the process.

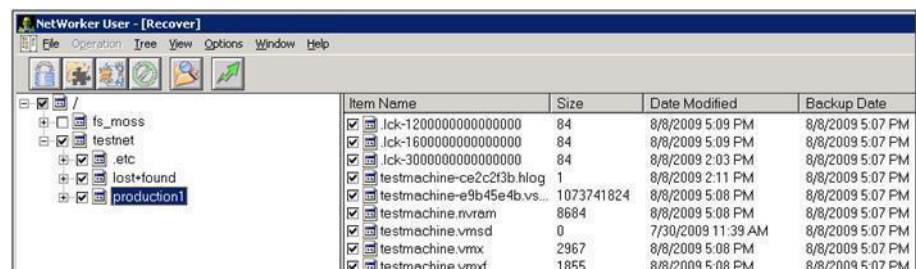


Figure 153 NDMP recovery using NetWorker

- 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 154 shows the SNAPSURE parameter set to create a backup with an integrated checkpoint.

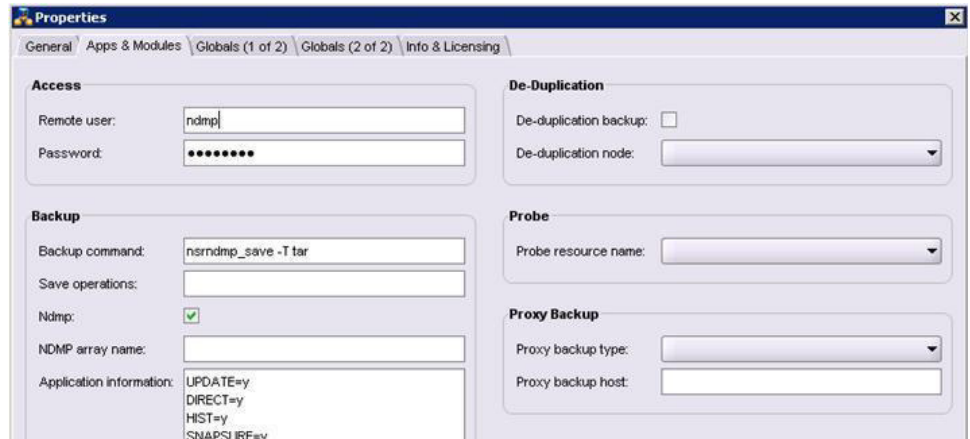


Figure 154 Backup with integrated checkpoint

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, providing 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 21 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 but an example of the data protection options and technologies that exist within EMC VNX and VMware vSphere.

Table 21 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

CHAPTER 4

Using VMware vSphere in Data Restart Solutions

This chapter presents the following topics:

◆ Introduction	196
◆ EMC remote replication technology overview	198
◆ RDM volume replication	212
◆ EMC Replication Manager.....	215
◆ Automating site failover with SRM and VNX	217
◆ Summary	225

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 and 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 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 the following 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 22](#) lists the DR and BC software options available for each storage device type.

Table 22 EMC replication options for VMware environments

Replication technology	NFS	VMFS	RDM
EMC Replicator	✓		
EMC RecoverPoint CRR ¹	✓	✓	✓
EMC MirrorView		✓	✓

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 more than 1,024 separate file system sessions for each 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. In Unisphere, click **Data Protection**.
2. Click **File Replication Wizard - Unisphere**.
The Replication Wizard appears.
3. Complete the following steps, as shown in [Figure 155](#) and [Figure 156](#).
 - 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.

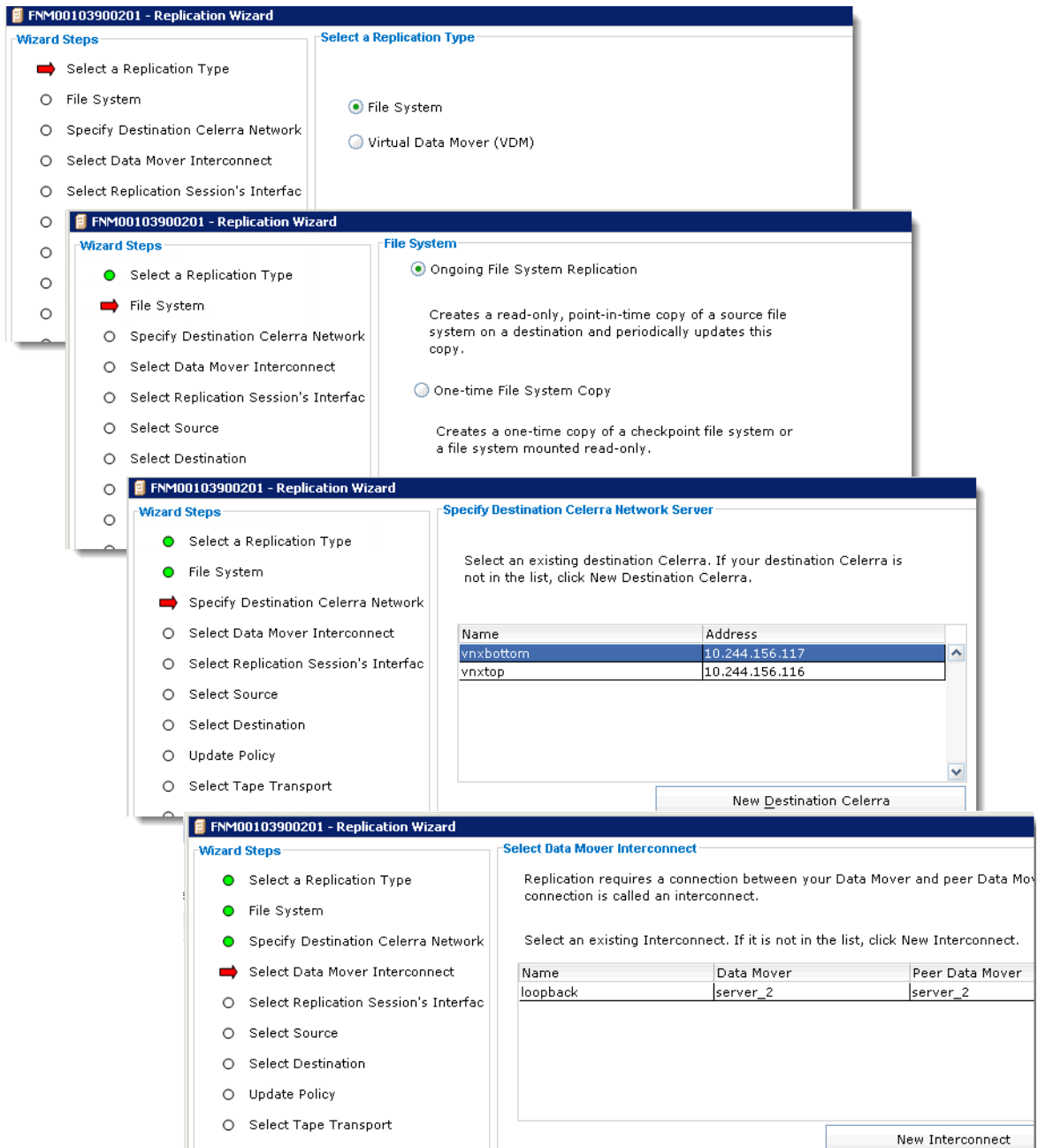


Figure 155 Data replication wizard

e. Specify a name for the replication session.

f. Select the source file system to replicate to the remote location.

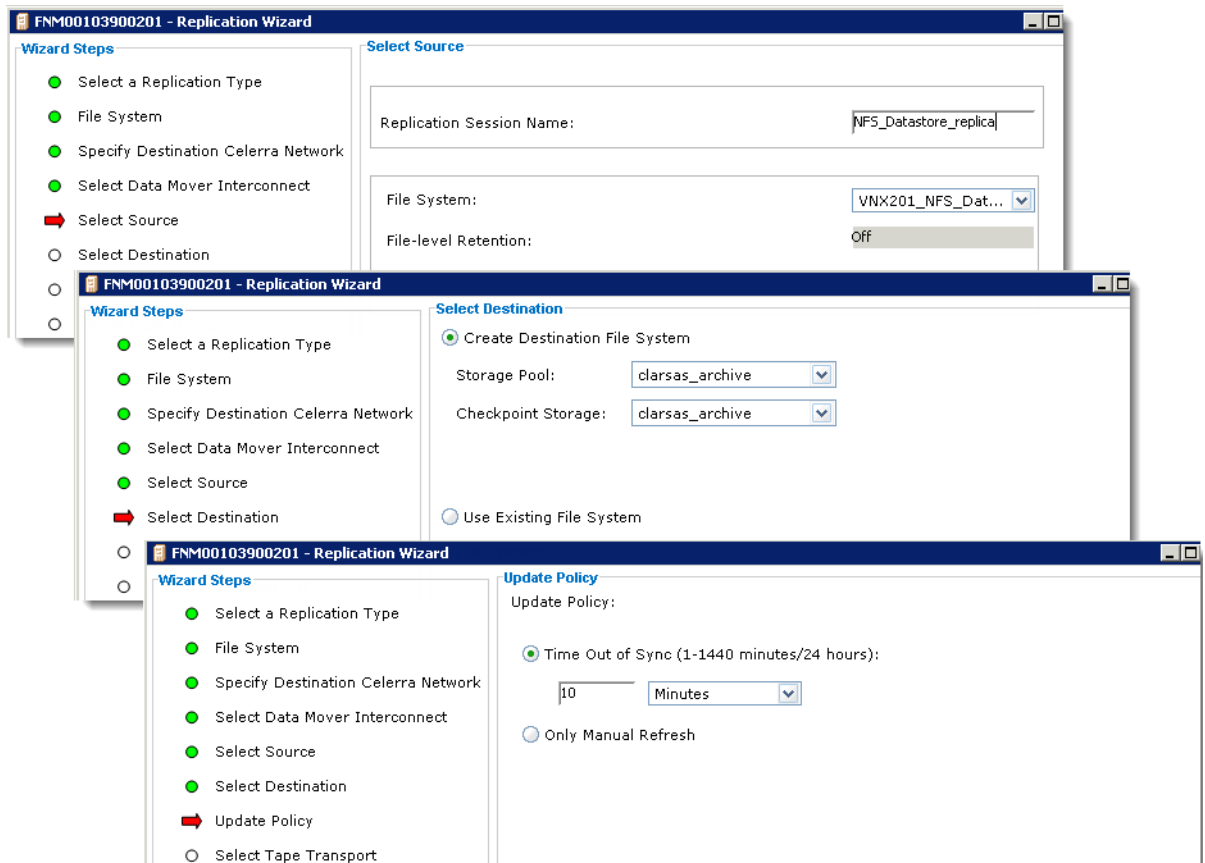


Figure 156 Replication wizard (continued)

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:

- ◆ Initiate a failover
- ◆ Terminate the replication session
- ◆ Reverse 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 23](#) lists the MirrorView limits for the VNX platforms.

Table 23 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 157](#) shows an example of a consistency group with four LUNs. Use MirrorView consistency groups with SRM configurations.

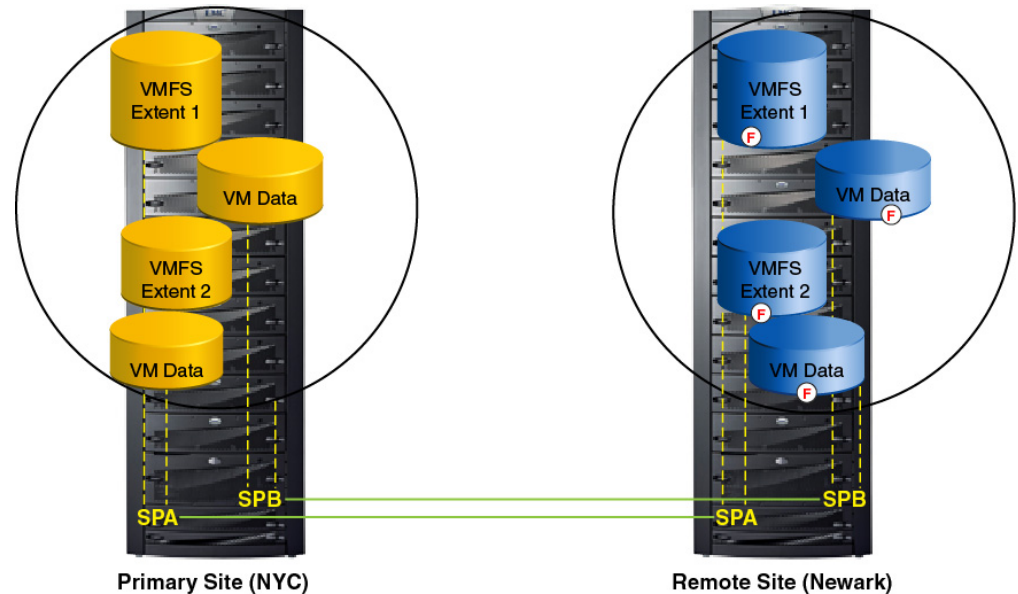


Figure 157 Preserving dependent-write consistency with MirrorView consistency groups

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

Asynchronous MirrorView (MV/A) is a method used to replicate up to 256 LUNs between VNX systems. With MV/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

Synchronous MirrorView (MV/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.

Setting up MirrorView replication

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 158](#).

Note: The process and commands to configure synchronous and asynchronous MirrorView replication are very similar. Specify the `-async` argument for asynchronous replication.

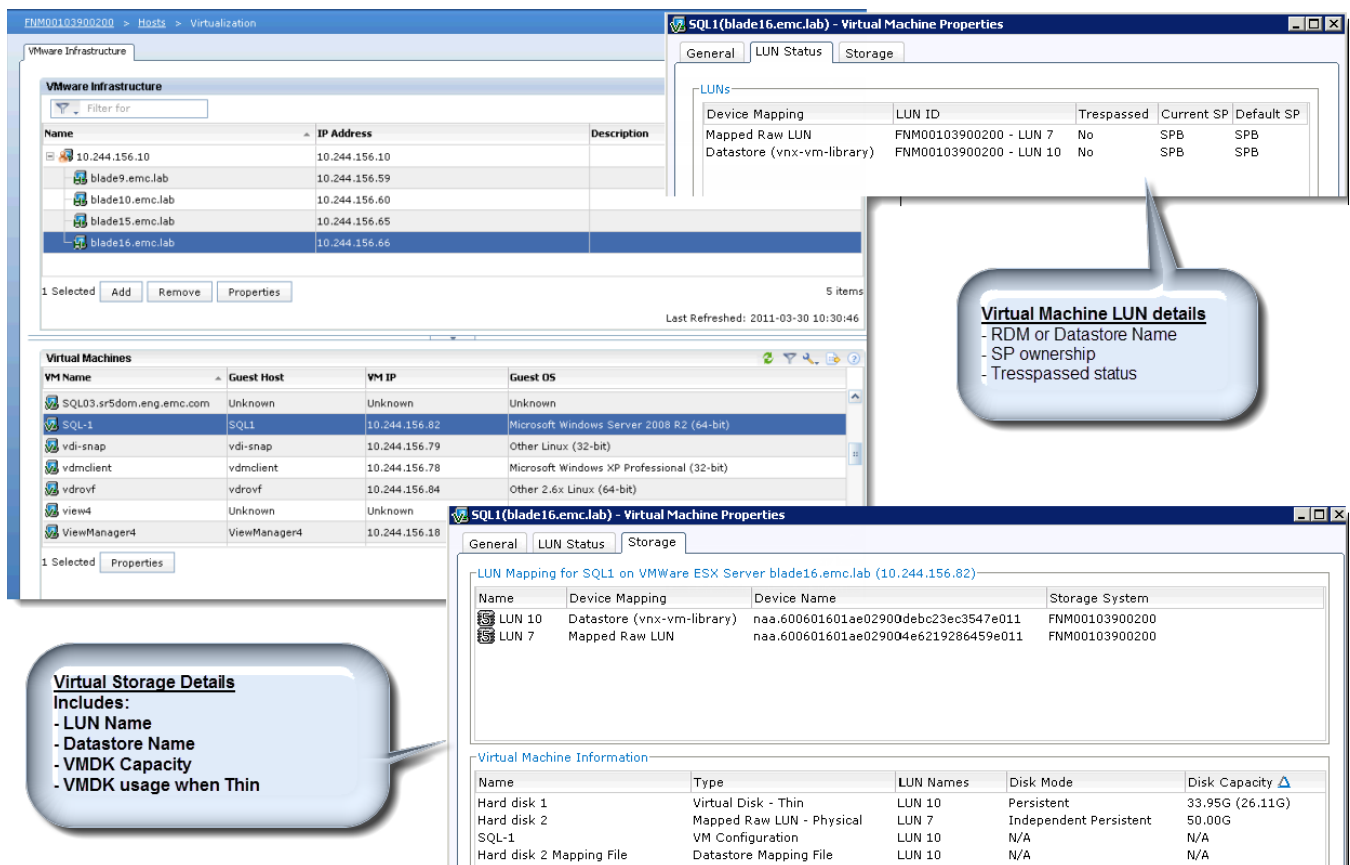


Figure 158 EMC Unisphere interface

1. From the Unisphere Data Protection window, select **Manage Mirror Connections**.

- Identify the Peer Storage System and enable the MirrorView connection between the two systems as shown in [Figure 159](#).

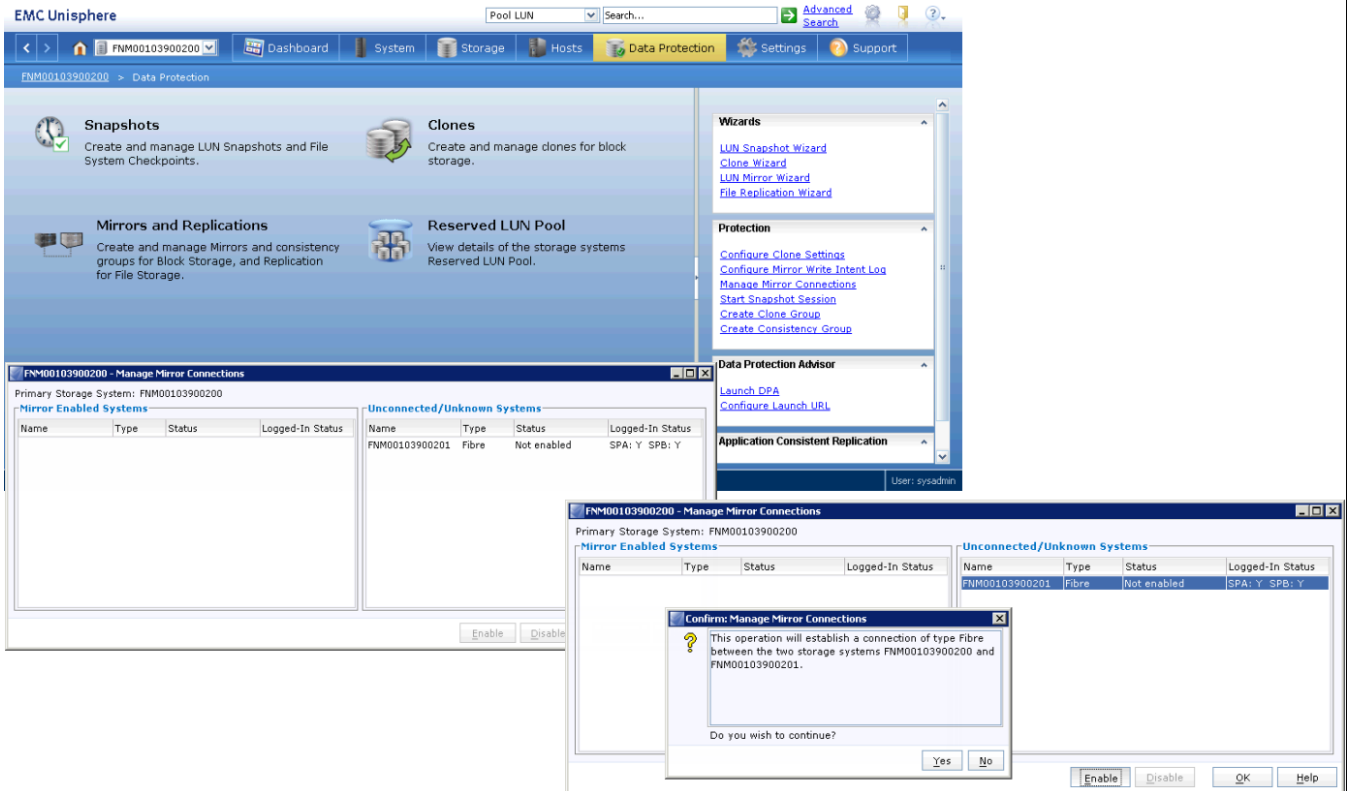


Figure 159 Enable MirrorView between VNX systems

3. Use the Unisphere MirrorView LUN wizard to select the source LUNs and establish a remote mirror at the recovery site as shown in [Figure 160](#).

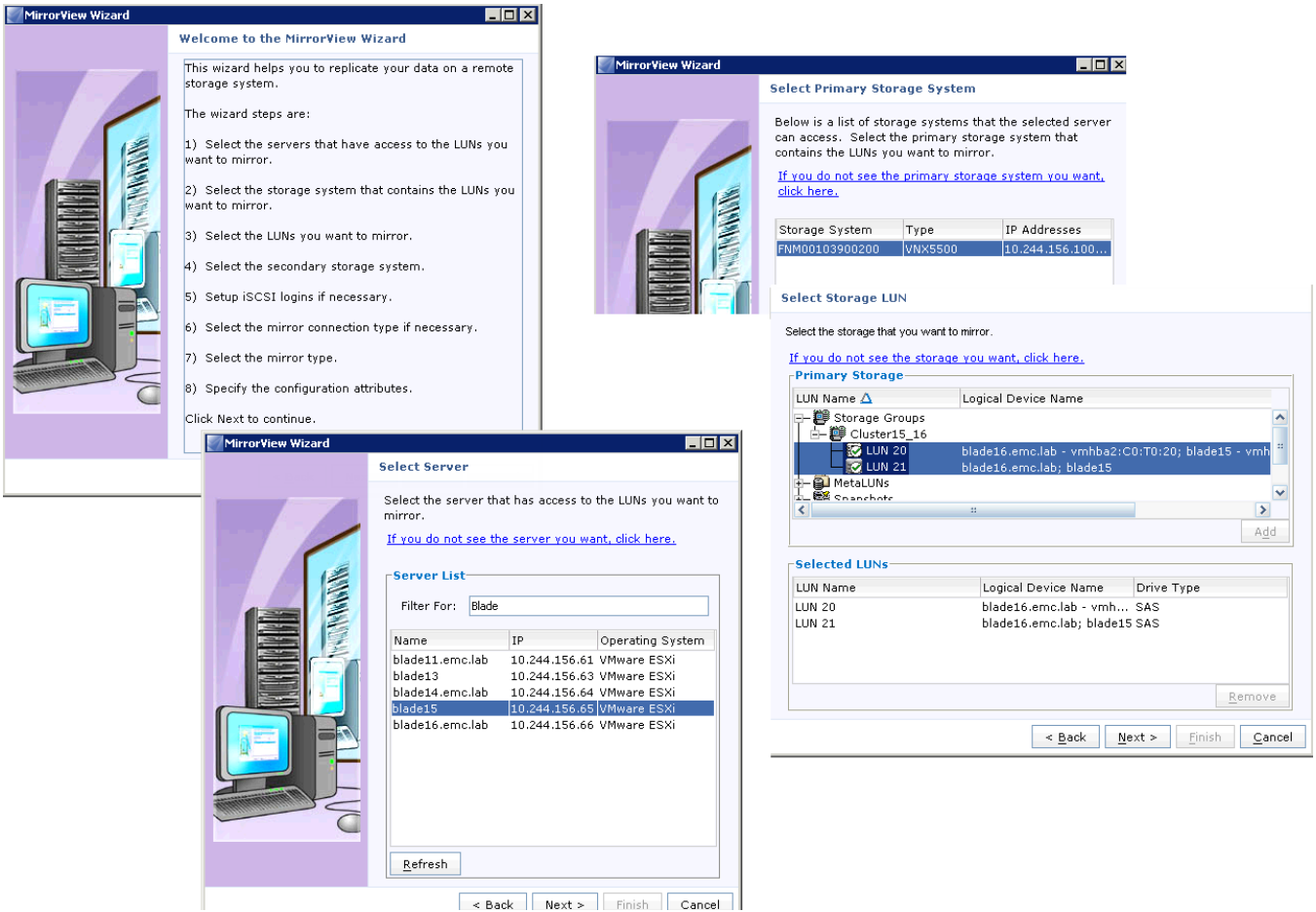


Figure 160 MirrorView Wizard - select source LUNs

- Select the remote storage pools to use for the MirrorView session as shown in Figure 161.

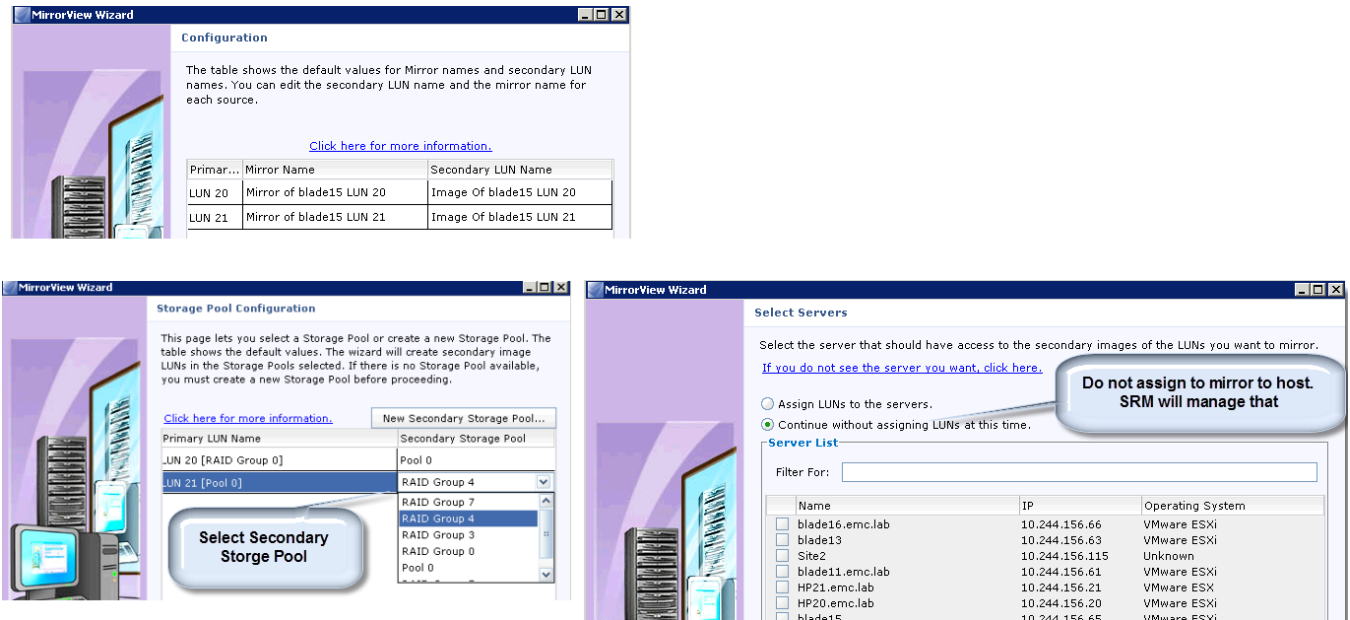


Figure 161 MirrorView Wizard - select remote storage

- Promote the secondary image at the DR site as shown in Figure 162.

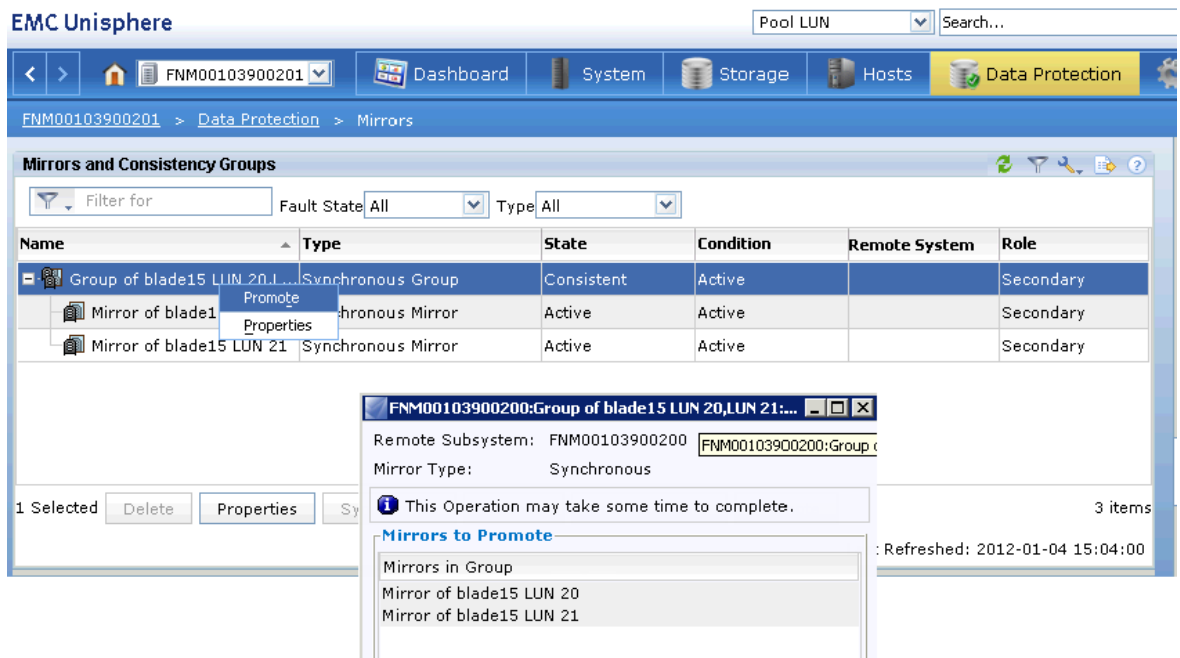


Figure 162 Promote mirrored LUN

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 163 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.

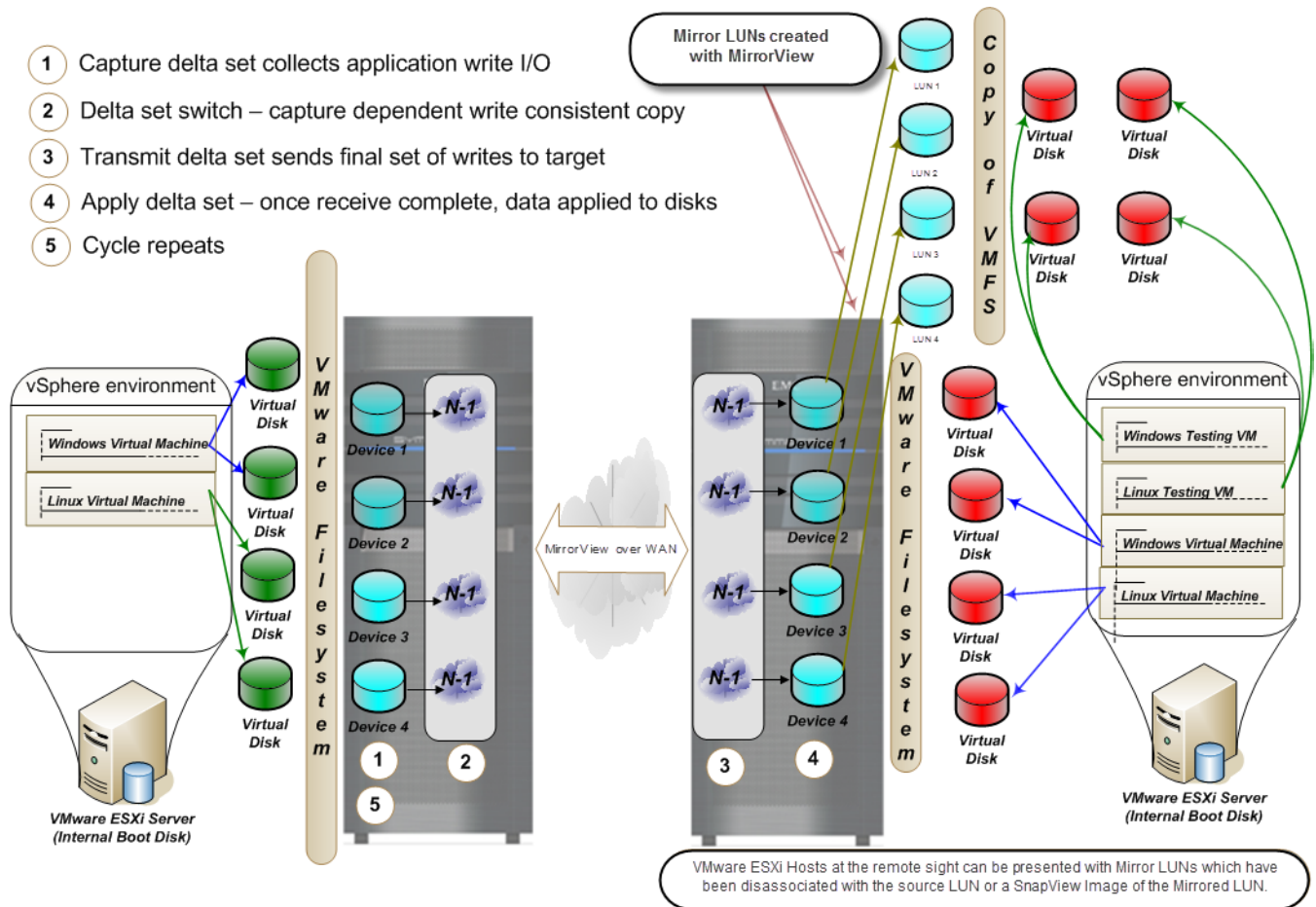


Figure 163 Business continuity solution using MirrorView/S in a virtual infrastructure with VMFS

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.

1. In a planned failover, disable or shut down the VNX at the production site before performing the failover tasks.
2. To prevent data loss, synchronize secondary MirrorView/S LUNs before starting the failover process.
3. 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 164](#).

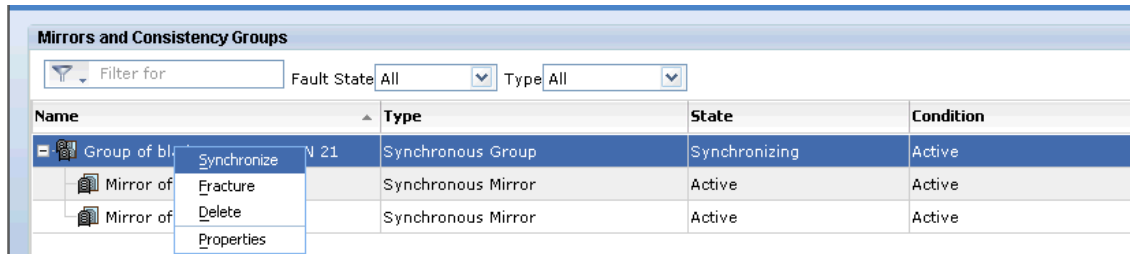


Figure 164 Synchronize MirrorView LUNs

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.

EMC RecoverPoint

EMC RecoverPoint provides local and remote LUN replication.

RecoverPoint consists of the following components:

- ◆ Continuous Data Protection (CDP) for local replication
- ◆ Continuous Remote Replication (CRR) for remote replication
- ◆ Continuous Local and Remote Replication (CLR), which is a combination of the two, for sequential, remote, and local replication of the same LUN.

Figure 165 provides an overview of the RecoverPoint architecture.

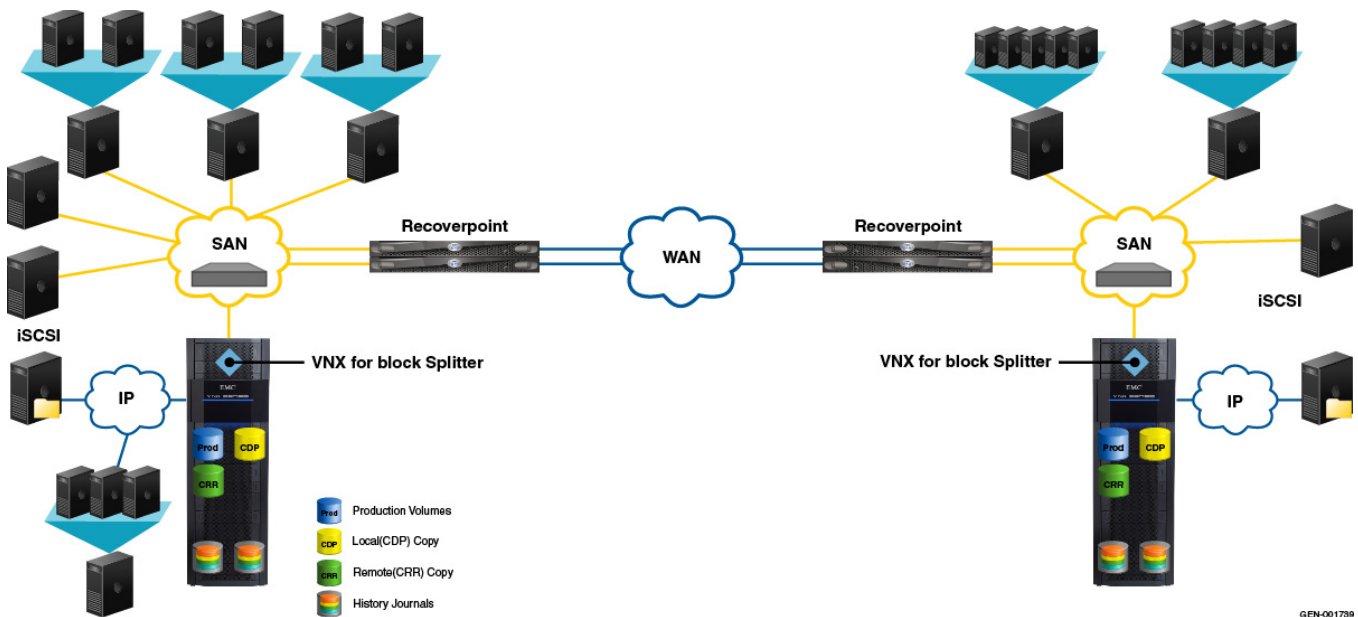


Figure 165 RecoverPoint architecture overview

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 165.
- ◆ 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 environments, 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 and later provides full support for VAAI with the VNX splitter. Table 24 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

support only 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 24](#). If running a prior version, SCSI commands other than ATS are rejected and revert to the host for processing.

Table 24 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 166](#) 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.

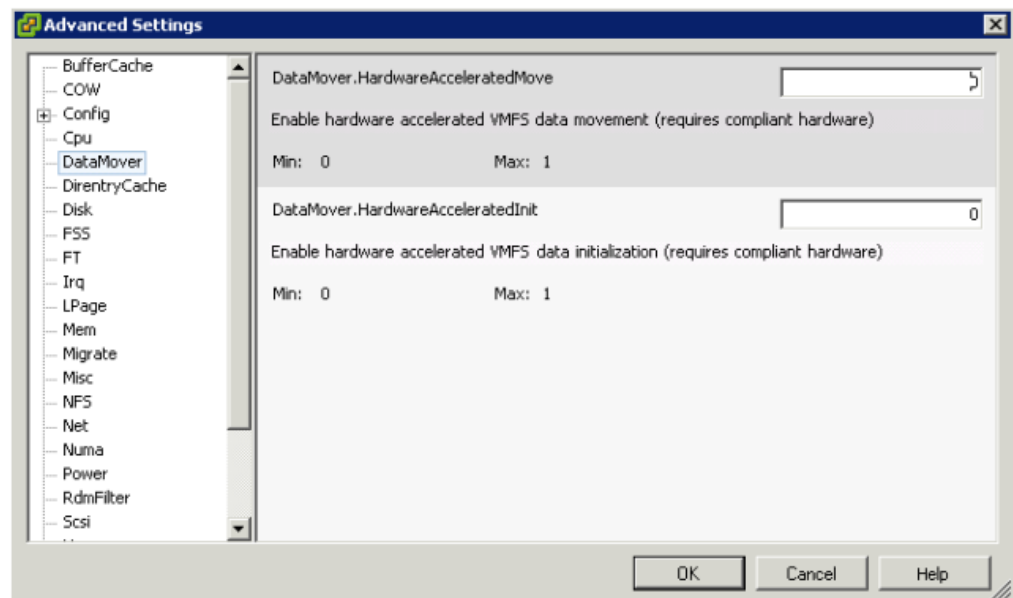


Figure 166 Disabling VAAI support on an ESXi host

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 25 EMC RecoverPoint feature support

Feature	Splitter		
	Windows host write splitter	Array-based write splitter	Brocade/Cisco Intelligent Fabric write splitter
Supports physical RDM	✓	✓	✓
Supports virtual RDM	✗	✓	✓
Supports VMFS	✗	✓	✓
Supports VMotion®	✗	✓	✓
Supports HA/DRS	✗	✓	✓
Supports vCenter Site Recovery Manager	✗	✓	✓
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	✗	✓	✓
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	✗	✓	✓

RDM volume replication

Replication of RDMs requires the completion of management tasks in addition to those for 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 is created that maps the logical virtual machine device to the physical device. 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.

Maintaining device order

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 26](#) shows an example with three application data disks.

Table 26 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.

Creating copies of the production virtual machine

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: Because these tasks 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.

Configuring 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 can be 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 that 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 167](#) shows a NAS datastore replica in the RM.

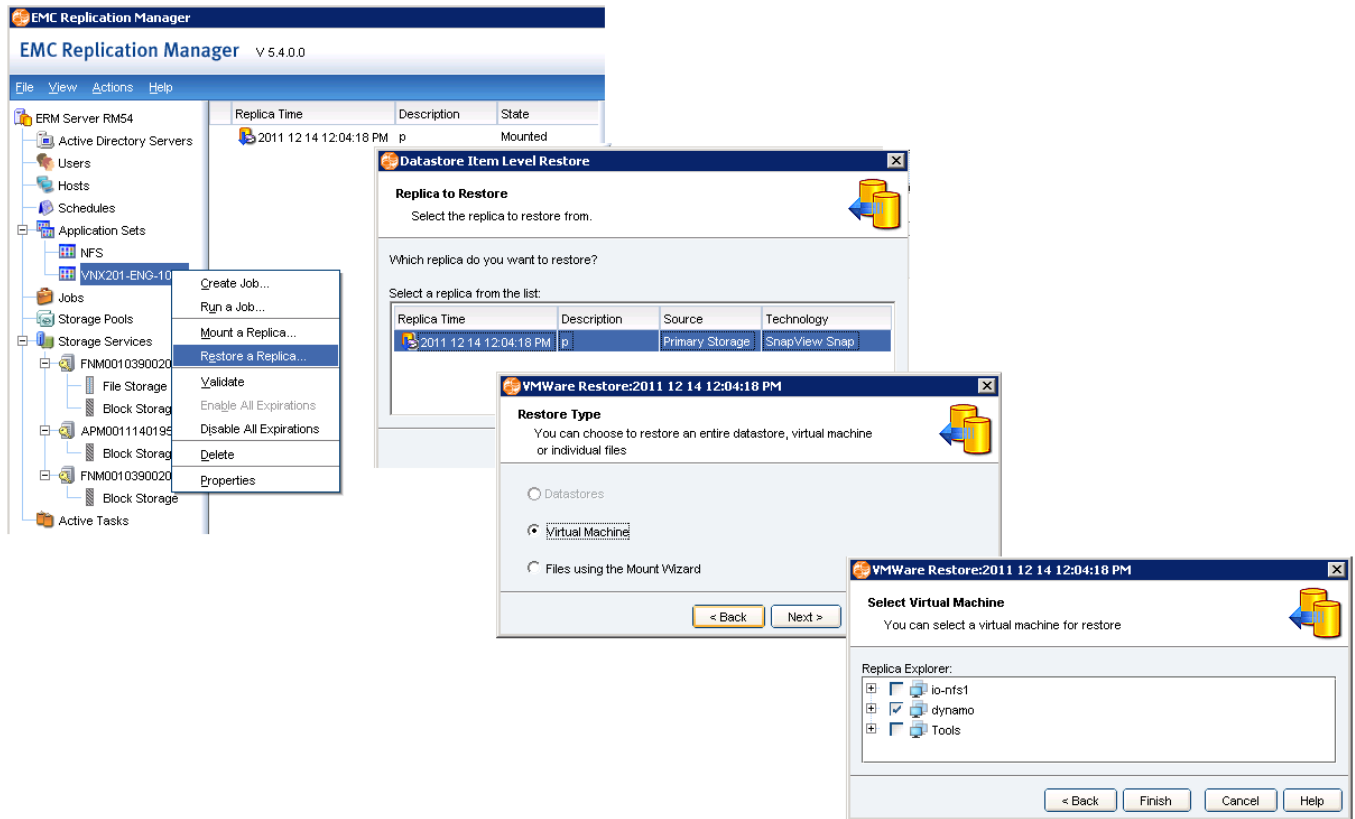


Figure 167 RM protection for NFS datastores and virtual machines

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 failover 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 168](#).

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.

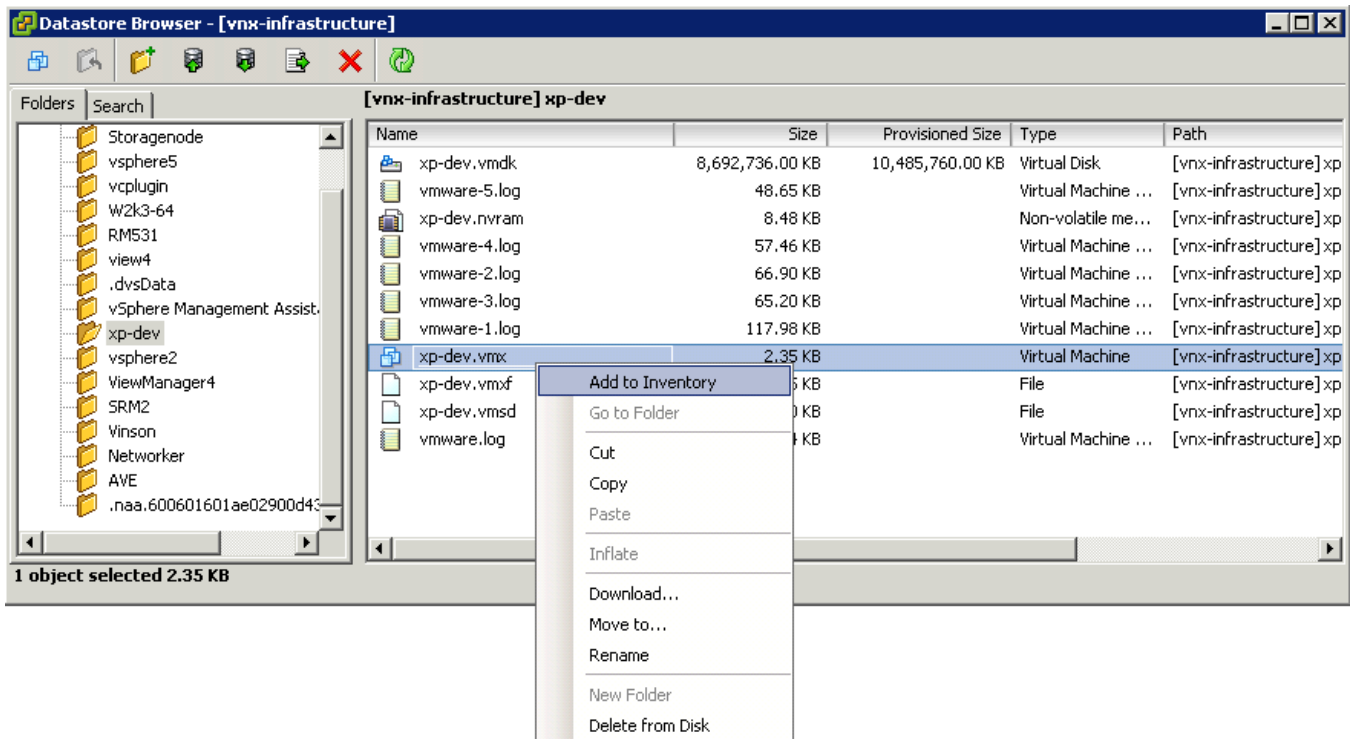


Figure 168 Using the vSphere client to register a virtual machine with ESXi

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 ccommand 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 can add test-specific customization to the workflow for the test failover to handle cases 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 169](#).

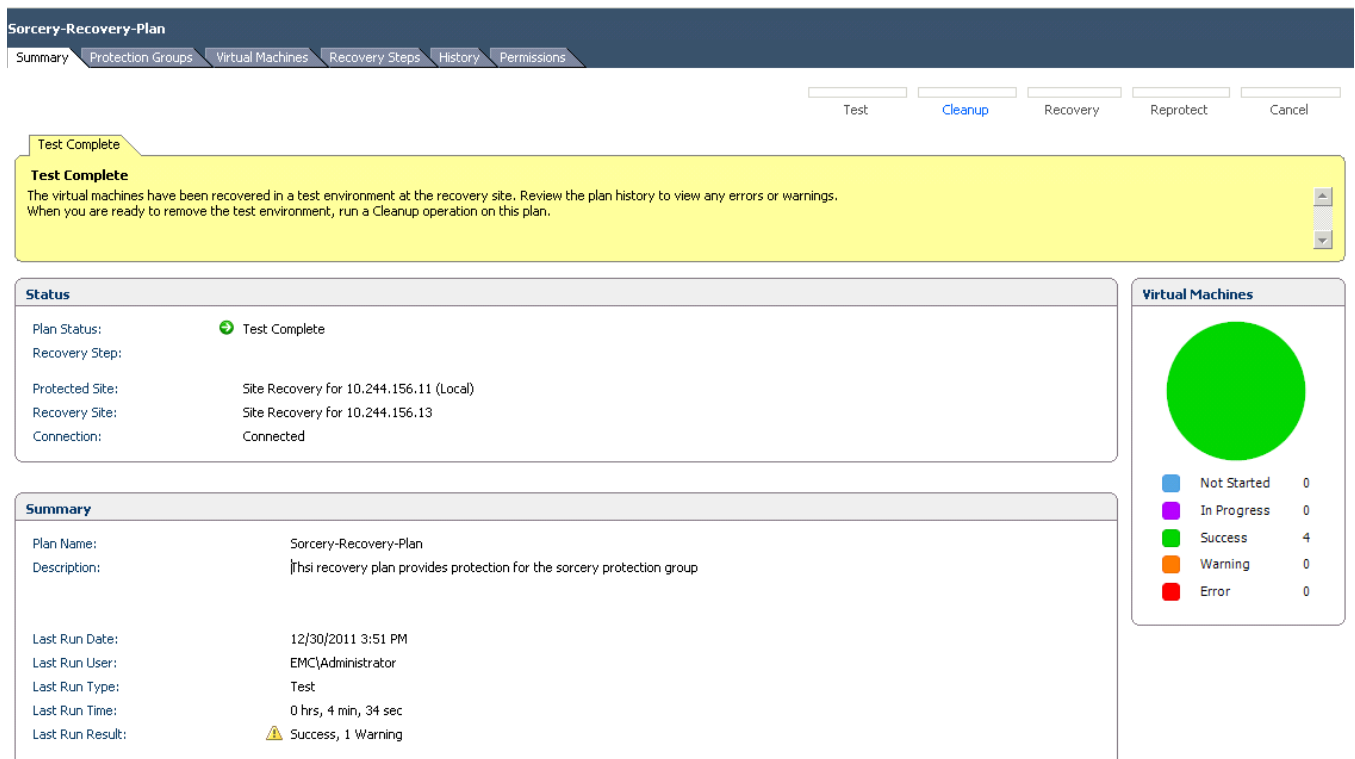


Figure 169 SRM recovery plan summary

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 uses 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.

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 170 shows an example of SRM configuration in vCenter.

The screenshot displays the VMware vCenter SRM configuration interface. The left pane shows the 'Array Managers' tree with 'VNX200' selected. The main pane shows the 'VNX200' configuration page, specifically the 'Array Pairs' tab. The 'Discovered Array Pairs - VNX200' section shows a table with columns: Local Array, Remote Array, Remote Array Manager, Status, and Actions. The status is 'Disabled' and the action is 'Enable | Disable'.

The 'Devices for Enabled Array Pairs' section shows a table with columns: Local Device, Direction, Remote Device, Datastore, Protection Group, and Local Consistency Group. The table contains two rows of data:

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

A callout box labeled 'Replication direction for consistency group' points to the 'Direction' column of the table.

Figure 170 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. A one-to-one

mapping exists between an SRM protection group and a VNX or RecoverPoint consistency group. [Figure 171](#) illustrates the configuration of a protection group that uses a MirrorView LUN consistency group.

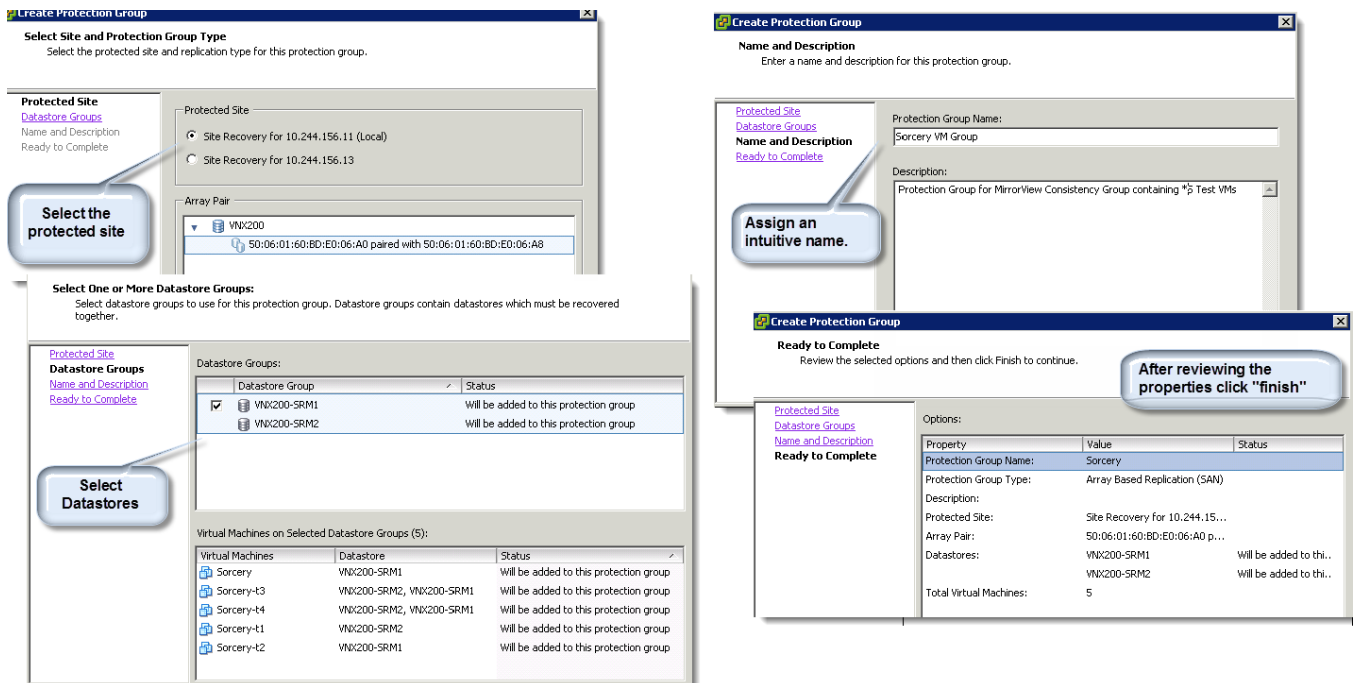


Figure 171 Creating an SRM protection group

Note: In cases that do not use a one-to-one mapping—for example where RecoverPoint is used to protect a database application with separate consistency groups for binaries, user databases, and system databases—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.

Testing the SRM recovery plan at the recovery site

Test the SRM recovery plan to verify that it performs as expected. [Figure 172](#) shows a sample recovery plan.

The screenshot shows the 'Sorcery-Recovery-Plan' interface with tabs for Summary, Protection Groups, Virtual Machines, Recovery Steps, History, and Permissions. The 'Recovery Steps' tab is active, displaying a table of steps. A progress bar at the top indicates the current step is 'Test'. A callout box labeled 'Recovery Steps' points to the table.

Recovery Step	Status	Step Started	Step Completed
1. Synchronize Storage	Success	12/30/2011 3:51:42 PM	12/30/2011 3:52:17 PM
2. Restore hosts from standby	Success	12/30/2011 3:52:17 PM	12/30/2011 3:52:17 PM
3. Suspend Non-critical VMs at Recovery Site	Success	12/30/2011 3:52:17 PM	12/30/2011 3:52:17 PM
4. Create Writeable Storage Snapshot	Running	12/30/2011 3:52:17 PM	45%
5. Power On Priority 1 VMs			
6. Power On Priority 2 VMs			
7. Power On Priority 3 VMs			
8. Power On Priority 4 VMs			
9. Power On Priority 5 VMs			

Figure 172 Testing the recovery plan

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 173 shows the cleanup of a sample recovery plan.

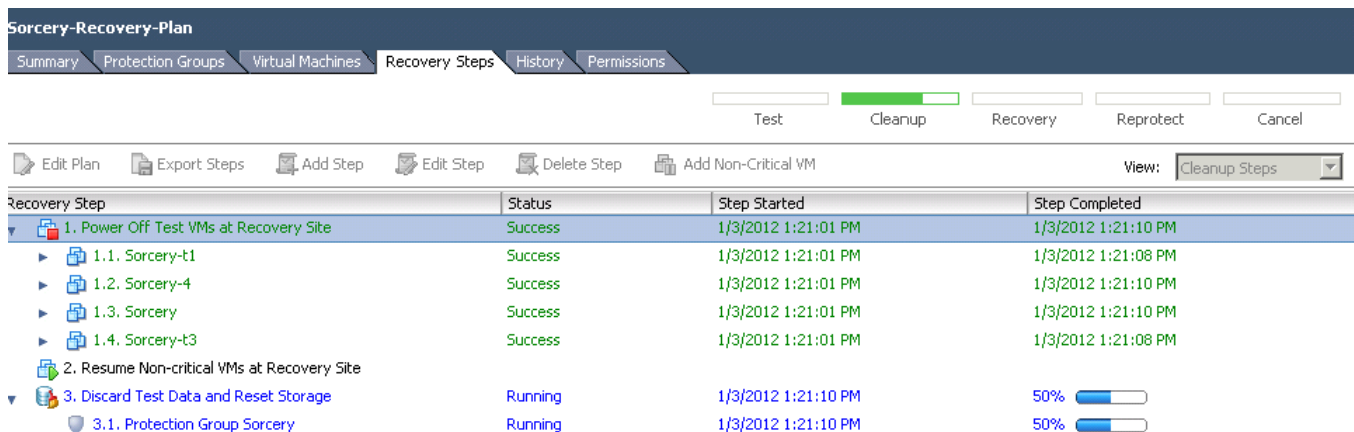


Figure 173 Recovery plan Cleanup

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.

Executing 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 174 shows a completed recovery plan.

Recovery Complete

Recovery Complete

The recovery has completed. Please review the plan history to view any errors or warnings. You may now press Reprotect to configure protection in the reverse direction. Note that if you plan to failback the virtual machines to the original site, you must first run the plan in reprotect mode, then once protection is configured in reverse, you may run the plan in recovery mode to failback the virtual machines to the original site.

Edit Plan Export Steps Add Step Edit Step Delete Step Add Non-Critical VM View: 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			
2.2. Shutdown Priority 4 VMs			
2.3. Shutdown Priority 3 VMs	Already Done		
2.4. Shutdown Priority 2 VMs			
2.5. Shutdown Priority 1 VMs			
3. Resume VMs Suspended by Previous Recovery			
4. Restore hosts from standby	Success	1/3/2012 1:29:41 PM	1/3/2012 1:29:41 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			
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			
10. Power On Priority 2 VMs			
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 1:34:24 PM
11.2. Sorcery-4	Success	1/3/2012 1:32:10 PM	1/3/2012 1:35:18 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			
13. Power On Priority 5 VMs			

Figure 174 SRM recovery plan with EMC MirrorView

SRM failback and reprotect

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 so 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 175](#).

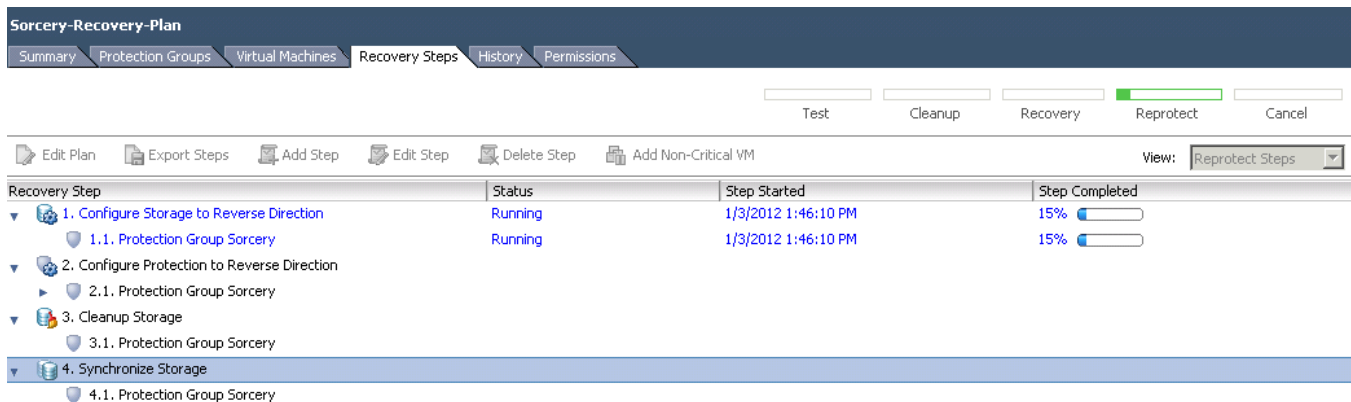


Figure 175 SRM reprotect

Best practices 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 27](#) lists the data replication solutions available for different types of VNX storage presented to an ESXi host.

Table 27 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

CHAPTER 5

Data Vaulting and Migration

This chapter presents the following topics:

◆ Introduction	228
◆ Using SAN Copy with VMware file systems	228
◆ Using SAN Copy with RDM virtual disks	229
◆ Using SAN Copy for data vaulting	229
◆ Importing Storage into the remote environment	234
◆ Using SAN Copy to migrate data to VNX arrays	235
◆ Summary	237

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 might 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 are core components 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, these technologies do not provide an enterprise-level solution for a full-scale migration of datastores from one storage location to another that does not impact 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.

Using SAN Copy with VMware file systems

SAN Copy is a VNX service that enables you 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 and 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.

Using SAN Copy 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 the following 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 176](#). 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.

Use a SnapView Clone LUN 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.

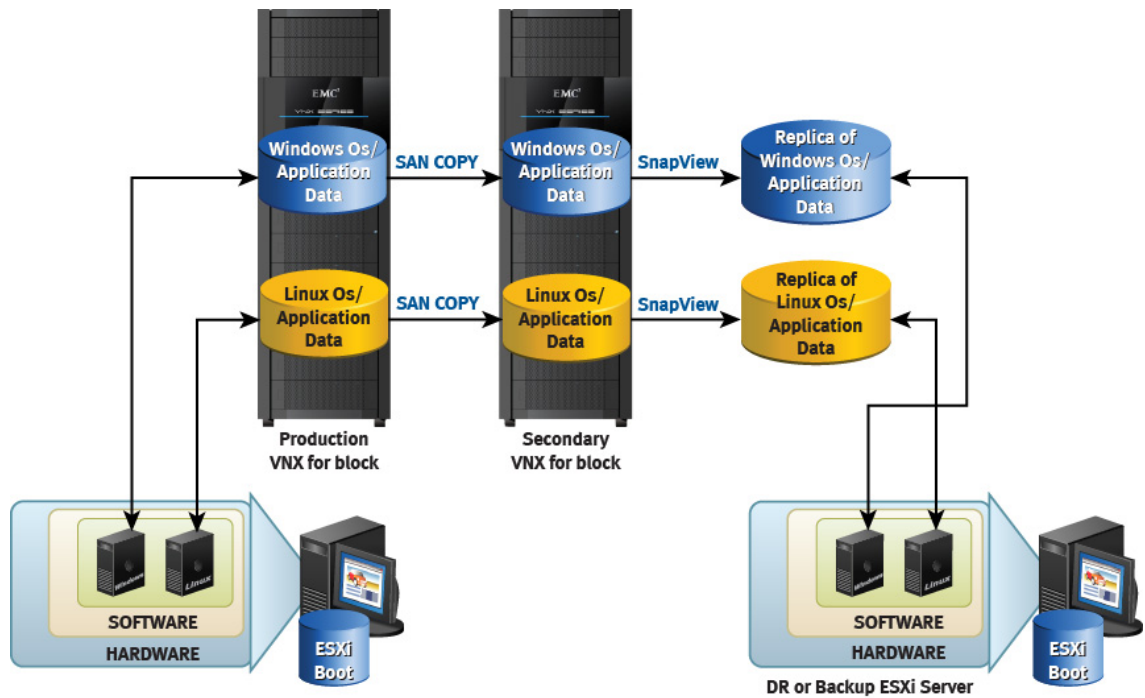


Figure 176 Data vaulting with Incremental SAN Copy

Using SAN Copy for data vaulting of VMware file systems

1. Migrate the LUN:

Note: This process applies to any VMFS or RDM LUN.

- a. Identify all the devices to be copied.
- b. Use Unisphere or the VSI Storage Viewer feature to identify the LUN that supports a VMFS datastore or RDM volume.

- c. 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 177](#).

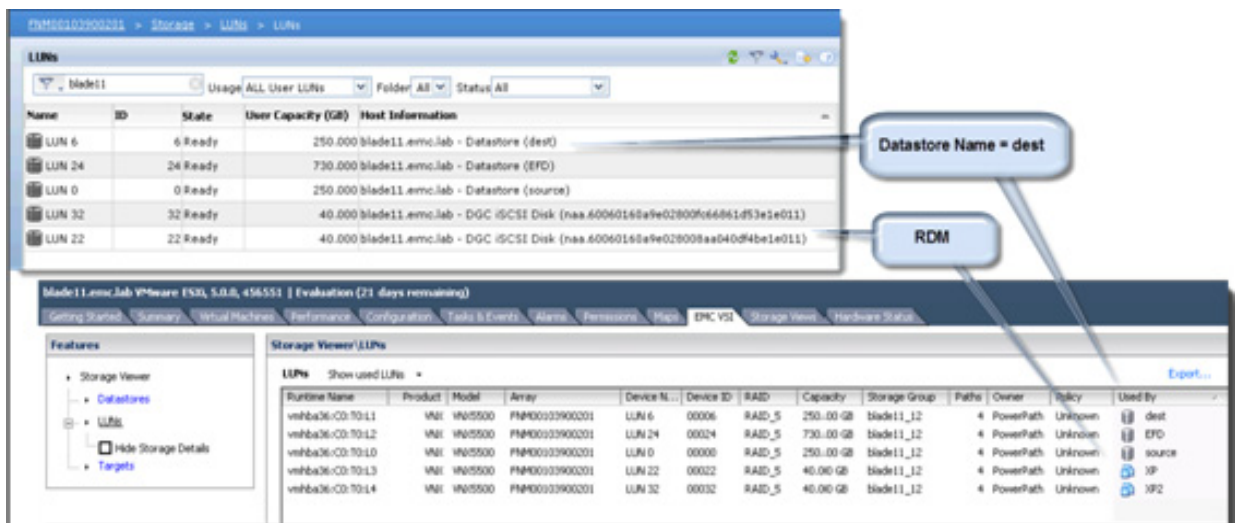


Figure 177 Using Unisphere or Storage Viewer to identify source LUNs

- d. 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).

2. Initiate the migration session and create a data vaulting solution:
 - a. 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.
 - b. 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.
 - c. 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.
 - d. Create an Incremental SAN Copy session between the source and destination LUNs as shown in [Figure 178](#) and [Figure 179](#).

e. 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.

Creating an incremental SAN Copy session is shown in Figure 178 and Figure 179.

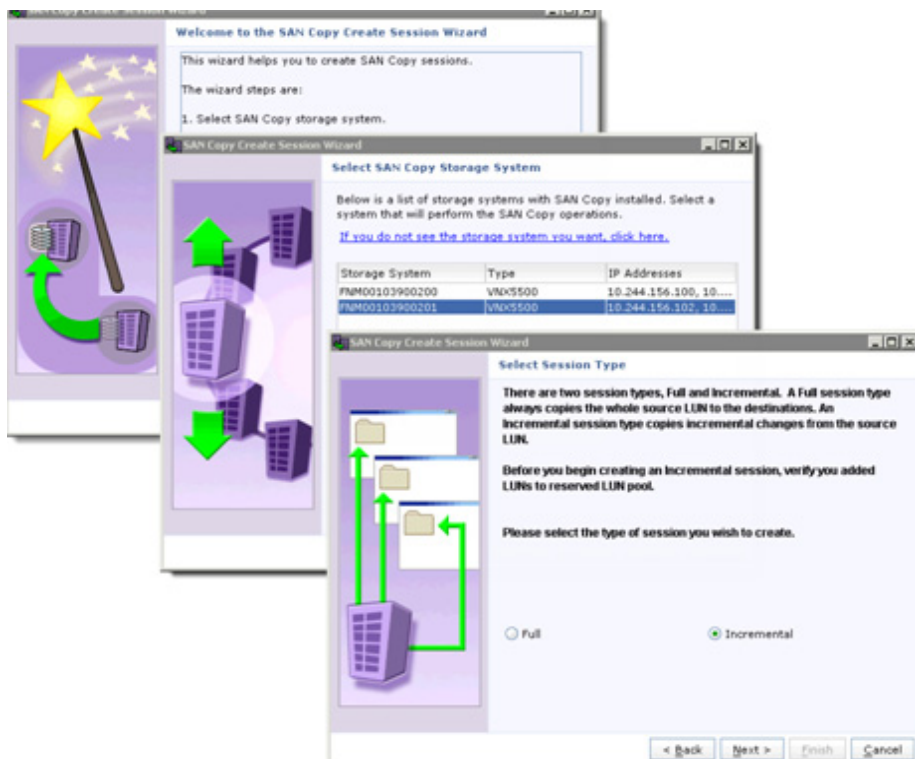


Figure 178 Creating an Incremental SAN Copy session

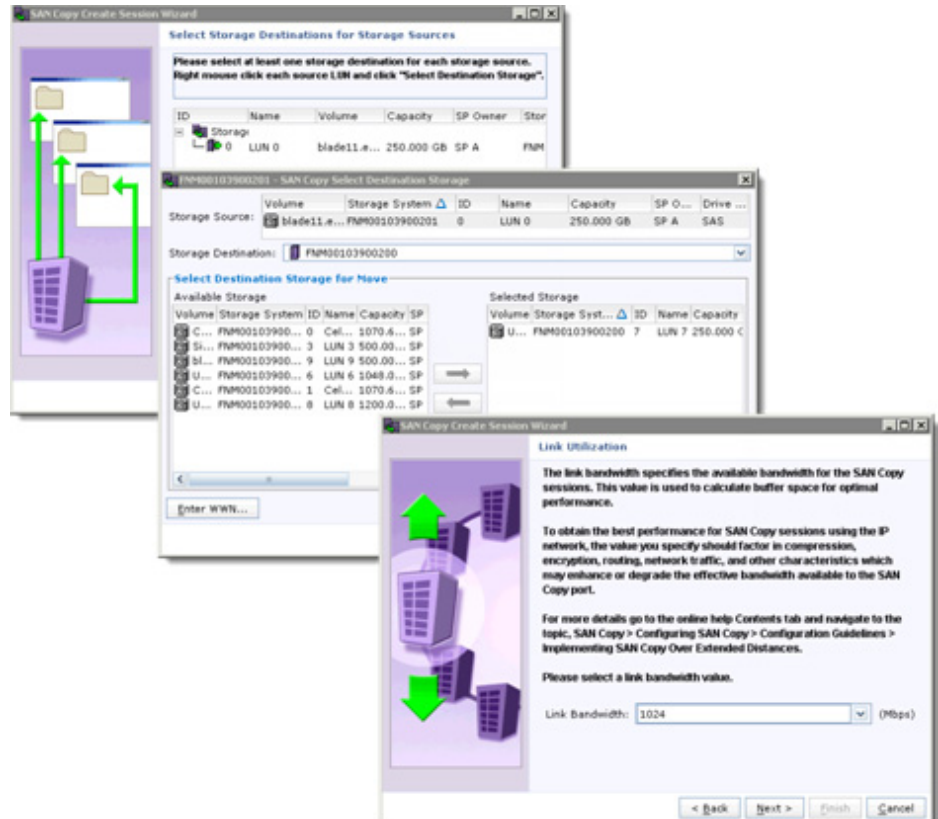


Figure 179 Creating an Incremental SAN Copy session (continued)

- f. 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.

- g. 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 Copy volume has had very little change between updates.

Using SAN Copy for data vaulting of virtual machines configured with RDMs

SAN Copy provides a storage array-based mechanism to create a consistent point-in-time copy of virtual disks stored on VNX LUNs. For RDM LUNs, SAN Copy replicates only the contents of the volume that have been modified by the guest, which is more efficient than replicating multiple virtual disks.

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 [“Using SAN Copy for data vaulting of VMware file systems” on page 230](#).

Importing Storage into the remote environment

You can configure remote sites for vSphere virtual machines using VMFS or RDM.

Configuring 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 vCenter under **Host > Configuration > Storage**.
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 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.

Configuring 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.

Note: The following procedure 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.

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.

Start the virtual machines with the procedure described in [“Starting virtual machines at a remote site after a disaster” on page 214](#).

Using SAN Copy to migrate data 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 are occasions when 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 capacity of the target LUN must be equal to or greater than the source LUN.
3. Create a full SAN Copy session for the clone volume on the remote array. [Figure 180](#) shows the necessary options to create a full SAN Copy session.

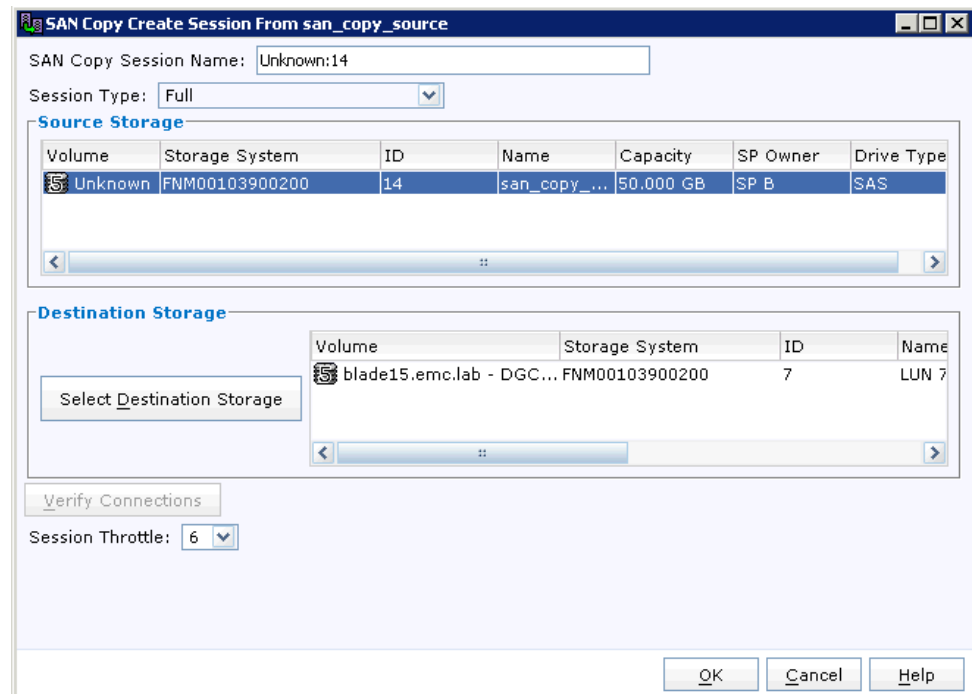


Figure 180 Creating a SAN Copy session to migrate data to a VNX

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 that applications running on the virtual machine function the same way they did on the previous storage system.

SAN Copy provides a convenient mechanism to use 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.

Migrating devices used as RDM

Use the procedure described in [“Configuring remote sites for vSphere virtual machines with RDM” on page 234](#).

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 does 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 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*, available on EMC Online Support, provides more information about data migration with SAN Copy.