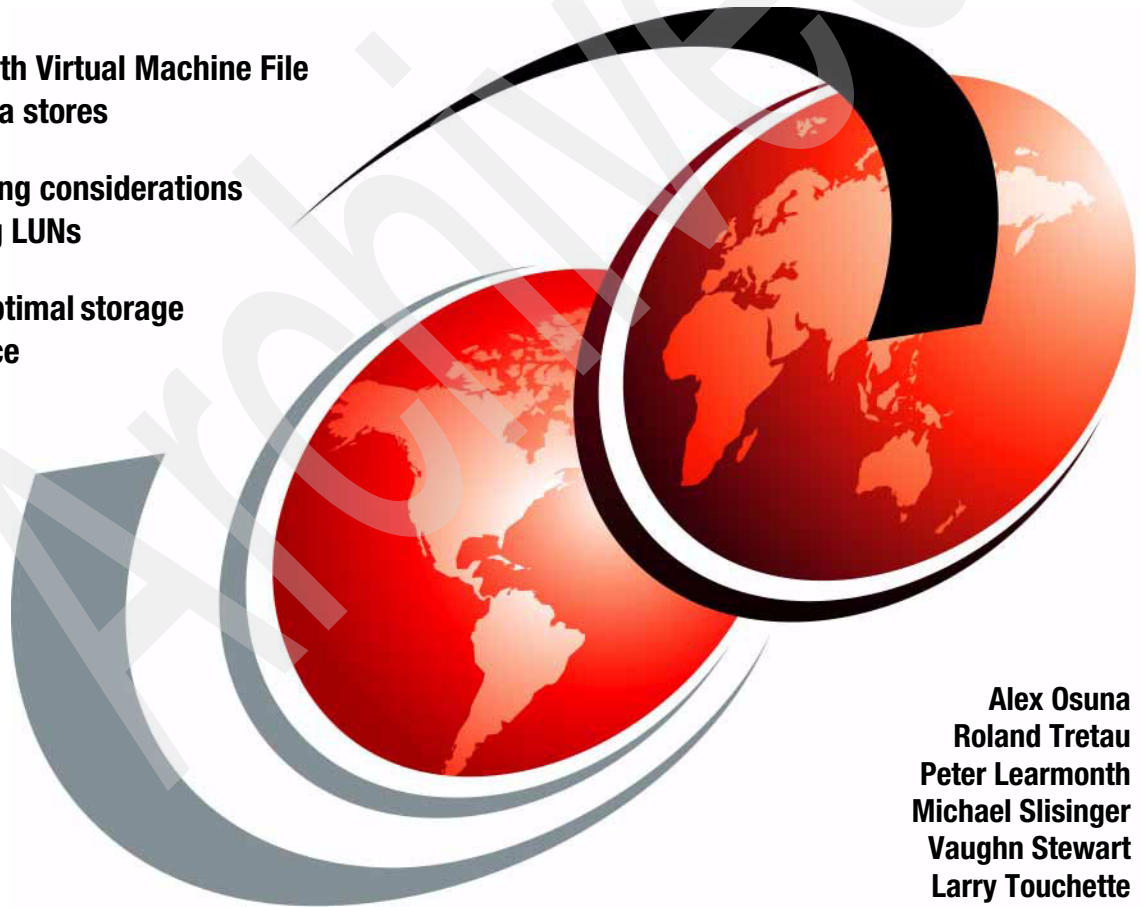


# IBM System Storage N series and VMware vSphere Storage Best Practices

Working with Virtual Machine File  
System data stores

Cluster sizing considerations  
when using LUNs

Ensuring optimal storage  
performance



Alex Osuna  
Roland Tretau  
Peter Learmonth  
Michael Slisinger  
Vaughn Stewart  
Larry Touchette





International Technical Support Organization

## **IBM System Storage N series and VMware vSphere Storage Best Practices**

July 2010

Archived

**Note:** Before using this information and the product it supports, read the information in “Notices” on page vii.

### **First Edition (July 2010)**

This edition applies to Data ONTAP 7.2.4 and later, VMware ESX 4.0, and VMware ESXi 4.0. It also applies to IBM System Storage N series.

**© Copyright International Business Machines Corporation 2010. All rights reserved.**

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

# Contents

<b>Notices</b> .....	vii
Trademarks .....	viii
<b>Preface</b> .....	ix
The team who wrote this book .....	ix
Now you can become a published author, too! .....	x
Comments welcome .....	xi
Stay connected to IBM Redbooks .....	xi
<b>Chapter 1. Concepts</b> .....	1
1.1 VMware ESX Server architecture .....	2
1.2 VMware storage options .....	3
1.2.1 Storage overview: VMFS datastores .....	3
1.2.2 Storage overview: NAS datastores .....	4
1.2.3 Storage overview: raw device mappings .....	5
1.3 Datastore comparison tables .....	6
<b>Chapter 2. N series configuration and setup</b> .....	9
2.1 Data protection .....	10
2.1.1 RAID and data protection .....	10
2.1.2 Aggregates and flexible volumes .....	11
2.2 HA mode for FC configurations .....	12
2.3 Storage configuration .....	13
2.3.1 Flexible volumes .....	13
2.3.2 Snapshot reserve .....	13
2.3.3 LUNs .....	14
2.3.4 Storage naming conventions .....	14
2.4 vSphere in a N series MetroCluster environment .....	14
<b>Chapter 3. VMware ESX FC, FCoE, and iSCSI storage configuration</b> ...	17
3.1 LUN sizing for VMFS datastores .....	18
3.2 Cluster sizing considerations when using LUNs .....	18
3.3 FC, FCoE, and iSCSI LUN provisioning .....	19
3.4 Connecting FC and FCoE datastores .....	21
3.5 Connecting iSCSI datastores .....	23
3.5.1 Enabling iSCSI communications .....	24
3.5.2 Creating multiple iSCSI VMkernels .....	25
3.5.3 Connecting to iSCSI targets .....	28
3.5.4 Restricting iSCSI targets to preferred interfaces .....	30

<b>Chapter 4. VMware native multipathing</b>	33
4.1 Default NMP settings	34
4.2 Enabling ALUA	35
4.3 Default NMP settings with ALUA enabled	36
4.4 Configuring the Round Robin PSP	36
4.4.1 Setting the default PSP for ALUA to Round Robin	36
4.4.2 Manually setting the PSP for a datastore	37
<b>Chapter 5. NFS storage best practices</b>	41
5.1 Increasing the number of NFS datastores	42
5.2 File system security	43
5.3 VMware ESX NFS timeout settings	44
5.4 NFS storage network best practice	44
5.5 Connecting NFS datastores	45
<b>Chapter 6. The N series VMware ESX Host Utilities</b>	51
6.1 Installation	52
6.1.1 Prerequisites	52
6.1.2 Installation	52
6.1.3 EHU assisted multipathing	53
6.2 Manual configuration of FC HBAs in VMware ESX	53
<b>Chapter 7. FC and FCoE storage networking best practices</b>	55
7.1 Host bus and converged network adapters	56
7.2 N series igroups (LUN masking)	56
7.3 FC and FCoE zoning	56
<b>Chapter 8. Ethernet storage networking best practices</b>	59
8.1 The 10 Gigabit Ethernet (10 GbE) standard	60
8.2 Virtual LANs (VLANs)	60
8.3 Flow control	60
8.4 Spanning Tree Protocol (STP)	61
8.5 Bridge Protocol Data Unit (BPDU)	62
8.6 Virtual Interfaces	62
8.7 Ethernet switch connectivity	63
<b>Chapter 9. Configuring Ethernet storage networks</b>	65
9.1 Highly available storage designs with traditional Ethernet switches	66
9.2 VMware ESX server adapter failover behavior with iSCSI	67
9.2.1 VMware ESX server adapter failover behavior with NFS	67
9.2.2 Reviewing link aggregation within VMware ESX server	68
9.2.3 Switch failure	68
9.2.4 Connecting to datastores	68
9.2.5 Scalability of VMware ESX server network connections	70

9.2.6 Configuring ESX/ESXI VMkernel storage network ports. . . . .	70
9.3 A storage architecture with traditional Ethernet. . . . .	72
9.4 Datastore configuration with traditional Ethernet. . . . .	75
9.5 VMkernel configuration with multiswitch trunking . . . . .	76
9.6 Storage network architecture with multiswitch link aggregation . . . . .	78
<b>Chapter 10. Increasing storage utilization . . . . .</b>	<b>81</b>
10.1 Introduction . . . . .	82
10.2 Data deduplication. . . . .	82
10.2.1 Deduplication considerations with VMFS and RDM LUNs . . . . .	84
10.2.2 Deduplication considerations with NFS. . . . .	85
10.3 Storage thin provisioning. . . . .	85
<b>Chapter 11. Virtual Machine best practices . . . . .</b>	<b>89</b>
11.1 Optimizing Windows VM file system performance . . . . .	90
11.2 Ensuring optimum VM availability . . . . .	90
11.3 Ensuring optimal storage performance . . . . .	91
11.3.1 Datastore alignment . . . . .	91
11.3.2 VM partition alignment . . . . .	91
11.3.3 Identifying partition alignment . . . . .	92
11.3.4 MBRtools: identification of partition alignment status . . . . .	92
11.4 Creating properly aligned partitions for new VMs . . . . .	93
11.4.1 Creating a properly aligned VMDK for a new VM with diskpart . . . . .	93
11.4.2 Creating a properly aligned VMDK for a new VM with fdisk . . . . .	94
<b>Chapter 12. Virtual Machine storage layout . . . . .</b>	<b>97</b>
12.1 Default virtual machine layout. . . . .	98
12.2 Virtual machine layout with N series Snap technologies. . . . .	98
12.2.1 Layout option 1: Implement a central virtual swap datastore . . . . .	99
12.2.2 Layout option 2: Locate VM swap or pagefile on a second datastore. . . . .	102
<b>Chapter 13. Storage monitoring and management . . . . .</b>	<b>105</b>
13.1 Monitoring storage utilization with N series Operations Manager. . . . .	106
13.2 Storage growth management . . . . .	106
13.2.1 Growing VMFS datastores . . . . .	106
13.2.2 Growing a virtual disk (VMDK) . . . . .	108
13.2.3 Growing a raw device mapping (RDM) . . . . .	109
13.2.4 Growing a file system within a guest OS (NTFS or EXT3) . . . . .	111
13.2.5 Growing bootable volumes within a guest operating system . . . . .	112
<b>Chapter 14. Disk-based Snapshot backups for VMware . . . . .</b>	<b>113</b>
14.1 Complementary Snapshot technologies . . . . .	114
14.2 Implementing snapshot backups. . . . .	114

<b>Appendix A. Configuring SSH on VMware ESX servers and N series systems</b>	117
N series system SSH configuration	118
ESX system SSH configuration	119
<b>Appendix B. Relocating the pagefile in Windows Virtual Machines</b>	121
<b>Abbreviations and acronyms</b>	123
<b>Related publications</b>	125
IBM Redbooks	125
Other publications	125
Online resources	126
IBM resources	126
VMware resources	126
How to get Redbooks	127
Help from IBM	127
<b>Index</b>	129



# Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:  
*IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.*

**The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:** INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

## COPYRIGHT LICENSE:


This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

# Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at <http://www.ibm.com/legal/copytrade.shtml>

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

IBM®  
Redbooks®  
Redpaper™

Redbooks (logo) ®  
System Storage®  
System x®

Tivoli®

The following terms are trademarks of other companies:

Snapshot, RAID-DP, WAFL, SyncMirror, SnapVault, SnapMirror, SnapManager, SnapDrive, FlexVol, FlexClone, FilerView, Data ONTAP, NetApp, and the NetApp logo are trademarks or registered trademarks of NetApp, Inc. in the U.S. and other countries.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

# Preface

IBM® System Storage® N series technology enables companies to extend their virtual infrastructures to include the benefits of advanced storage virtualization. IBM offers unified storage solutions that provide industry-leading technologies in the areas of storage efficiencies, instantaneous virtual machine (VM) and datastore cloning for virtual servers, virtual desktops, and virtual data center backup and business continuance solutions.

This IBM Redbooks® publication reviews the best practices for anyone who is implementing VMware® vSphere with IBM System Storage N series unified storage arrays. The book describes operational guidelines for the N series systems and VMware ESX Server. These techniques have been documented and are referred to as best practices.

These practices are only recommendations, not requirements. Not following the practices does not affect the support provided to your implementation by IBM and VMware. Not all practices apply to every scenario. We believe that customers can benefit from considering these practices before making any implementation decisions.

VMware vSphere offers several storage methods to virtual machines. All storage methods provide flexibility in infrastructure design, which in turn provides cost savings, increased storage utilization, and enhanced data recovery.

This book is not intended to be a definitive implementation or solutions guide. Expertise might be required to solve user-specific deployments. If you need solution and implementation support, consider contacting IBM services.

## The team who wrote this book

This book was produced by a team of specialists from around the world working at the International Technical Support Organization, San Jose Center.

**Alex Osuna** is a Project Leader at the International Technical Support Organization, Tucson Center. He writes extensively about all areas of storage. Before joining the ITSO in 2005, Alex worked in the Tivoli® Western Region as a Principal Systems Engineer. Alex has 32 years in the IT industry with the majority of them spent focusing on storage. He holds certifications from IBM, Microsoft®, RedHat, and the Open Group.

**Roland Tretau** is an Information Systems professional with IBM in Germany. He has over 15 years experience in the IT industry. He holds engineering and business master degrees, and is the author of many storage-related IBM Redbooks publications. Roland has a solid background in project management, consulting, operating systems, storage solutions, enterprise search technologies, and data management.

**Peter Learmonth** is a Technical Marketing Engineer for NetApp®. His primary focus is on Virtualization Solutions.

**Michael Slisinger** is a Technical Marketing Engineer with NetApp; his focus is server virtualization. He is the coauthor of several white papers and best practices guides about integrating VMware products with NetApp storage. Previously, he was a Consultant with NetApp Professional Services, and has over 10 years of experience with NetApp products.

**Vaughn Stewart** is the Infrastructure Virtualization Consultant for NetApp and is the coauthor of several white papers about deploying virtual infrastructures with NetApp systems. Vaughn has been with NetApp since 2000 and has worked with hundreds of customers and their implementations. Vaughn presented at the previous three VMworld® events, has been published in several articles including the NetApp blog: the Virtual Storage Guy. Vaughn currently holds industry certifications in solutions that are offered by Microsoft, Cisco, Sun Microsystems, IBM, NetApp, and VMware.

**Larry Touchette** is a Technical Marketing Engineer in the NetApp Server Virtualization and Grid Infrastructure Business Unit. He moved into this role three years ago. Prior to that he was on the NetApp Professional Services team for over six years, as both a PS Engineer and a PS Consultant. He started with NetApp in November 2000. Prior to NetApp was a Systems Administrator and Engineer in the automotive and IT industries.

## Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author - all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

[ibm.com/redbooks/residencies.html](http://ibm.com/redbooks/residencies.html)

## Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- ▶ Use the online **Contact us** review Redbooks form found at:

[ibm.com/redbooks](http://ibm.com/redbooks)

- ▶ Send your comments in an e-mail to:

[redbooks@us.ibm.com](mailto:redbooks@us.ibm.com)

- ▶ Mail your comments to:

IBM Corporation, International Technical Support Organization  
Dept. HYTD Mail Station P099  
2455 South Road  
Poughkeepsie, NY 12601-5400

## Stay connected to IBM Redbooks

- ▶ Find us on Facebook:

<http://www.facebook.com/IBMRedbooks>

- ▶ Follow us on twitter:

<http://twitter.com/ibmredbooks>

- ▶ Look for us on LinkedIn:

<http://www.linkedin.com/groups?home=&gid=2130806>

- ▶ Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

<https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm>

- ▶ Stay current on recent Redbooks publications with RSS Feeds:

<http://www.redbooks.ibm.com/rss.html>





# Concepts

This chapter introduces the architecture of VMware vSphere and how storage devices are connected.

## 1.1 VMware ESX Server architecture

VMware ESX Server is virtual infrastructure partitioning software that is designed for server consolidation, rapid deployment of new servers, increased availability, and simplified management. The software helps to improve hardware utilization, save space, IT staffing, and hardware costs.

You might have had earlier experience with VMware's virtualization products in the form of VMware Workstation or VMware GSX Server®. VMware ESX Server differs from other VMware products in that it runs directly on the hardware, offering a mainframe-class virtualization software platform that enables the deployment of multiple, secure, independent virtual machines on a single physical server.

VMware ESX Server allows several instances of operating systems, such as Windows® Server 2003, Windows Server 2008, Red Hat and (Novell) SUSE Linux®, and more, to run in partitions that are independent of one another; therefore, this technology is a key software enabler for server consolidation that provides the ability to move existing, unmodified applications and operating system environments from a large number of older systems onto a smaller number of new high-performance System x® platforms.

The architecture of VMWare ESX Server is shown in Figure 1-1.

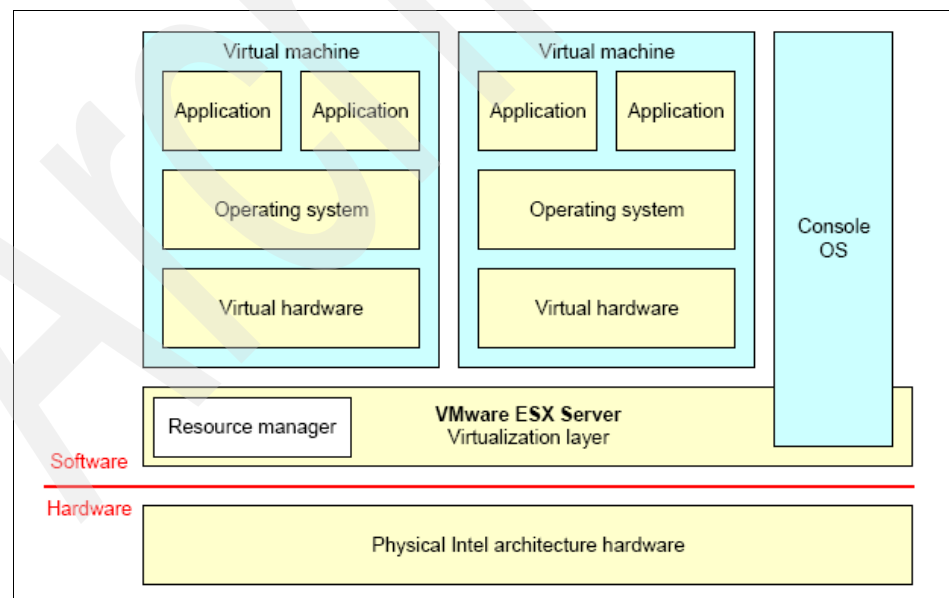


Figure 1-1 VMware ESX Server architecture



Instead of deploying multiple servers that are scattered around a company and running a single application on each, they can be consolidated together physically, as they simultaneously enhance system availability. VMware ESX Server allows each server to run multiple operating systems and applications in virtual machines, providing centralized IT management. Because these virtual machines are completely isolated from one another, if one is down, the others are unaffected. This feature means that VMware ESX Server software is both helpful for optimizing hardware usage, and also offers added benefits of higher availability and scalability.

## 1.2 VMware storage options

VMware ESX supports three types of storage configurations when connecting to shared storage arrays: VMFS datastores, NAS datastores, and raw device mappings. An assumption is that you understand that shared storage is required when enabling high-value VMware features such as High Availability (HA), Distributed Resource Scheduler (DRS), vMotion, and Fault Tolerance. The goal of the following sections is to provide information to consider when you are designing a virtual data center.

VMware virtualization technology helps you use all of these storage designs at any time or simultaneously. This section reviews these storage options and summarizes the unique characteristics of each architecture. For information regarding deploying with VMFS, NFS, and RDMs, see the *ESXi Installable and vCenter Server Setup Guide*, located at:

[http://www.vmware.com/pdf/vsphere4/r40/vsp\\_40\\_esxi\\_i\\_vc\\_setup\\_guide.pdf](http://www.vmware.com/pdf/vsphere4/r40/vsp_40_esxi_i_vc_setup_guide.pdf)

### 1.2.1 Storage overview: VMFS datastores

The VMware Virtual Machine File System (VMFS) is a high-performance clustered file system that provides datastores, which are shared storage pools. VMFS datastores can be configured with LUNs accessed by Fibre Channel, iSCSI, or Fibre Channel over Ethernet. VMFS allows traditional logical unit numbers (LUNs) to be accessed simultaneously by every VMware ESX Server in a cluster.

VMFS provides the VMware administrator with a fair amount of independence from the storage administrator. By deploying shared datastores, the VMware administrator is free to provision storage to virtual machines as needed. In this design, most data management operations are performed exclusively through VMware vCenter™ Server.

Applications traditionally require storage considerations, to be sure that their performance can be virtualized and served by VMFS. With these types of deployments, deploy the virtual disks on a datastore that is connected to all nodes in a cluster, but is accessed only by a single VM.

This storage design can be challenging in the area of performance monitoring and scaling. Because shared datastores serve the aggregated I/O demands of multiple VMs (Figure 1-2), this architecture does not natively allow a storage array to identify the I/O load generated by an individual VM. This issue can be exacerbated by spanning VMFS volumes across multiple LUNs.

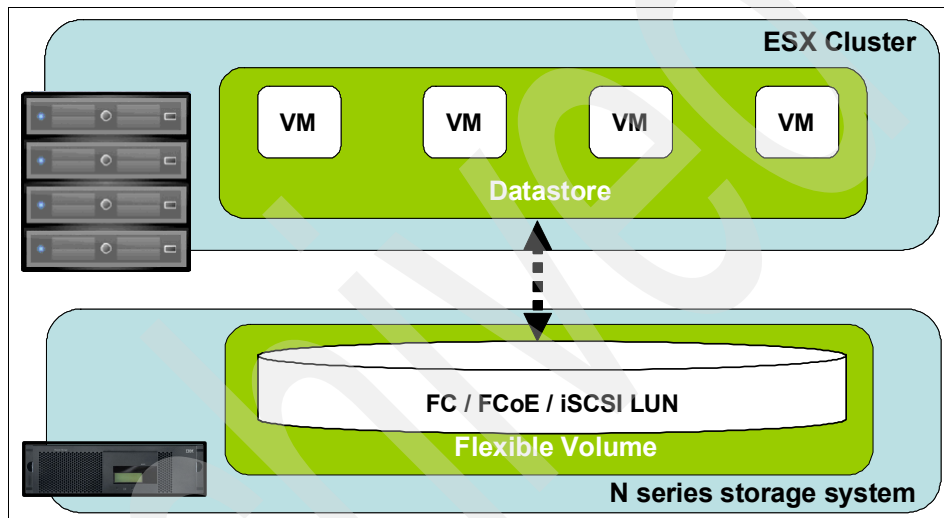


Figure 1-2 ESX server connected to a VMFS datastore using FC or iSCSI

## 1.2.2 Storage overview: NAS datastores

In addition to VMFS, vSphere allows the use of enterprise-class NFS servers, to provide datastores with concurrent access by all nodes in a VMware ESX cluster. This method of access is very similar to that with VMFS. NFS provides high performance, the lowest per-port storage costs (as compared to Fibre Channel solutions), and several advanced data management capabilities.

Deploying VMware with N series NFS datastores is the easiest means to integrate VMware virtualization technologies directly with Write Anywhere File Layout (WAFL®).

Examples of this transparent integration include production-use data deduplication, immediate zero-cost VM and datastore clones, array-based thin provisioning, and direct access to array-based Snapshot™ copies.

IBM provides additional VMware integrated tools for NFS such as SnapManager® for Virtual Infrastructure.

Figure 1-3 shows an example of this configuration. The storage layout is similar to that of a VMFS datastore, but each virtual disk file has its own I/O queue directly managed by the N series system. Combining N series advanced NFS servers with VMware's high-performance NFS implementation can provide I/O to shared datastores that is comparable with that of other storage protocols such as Fibre Channel.

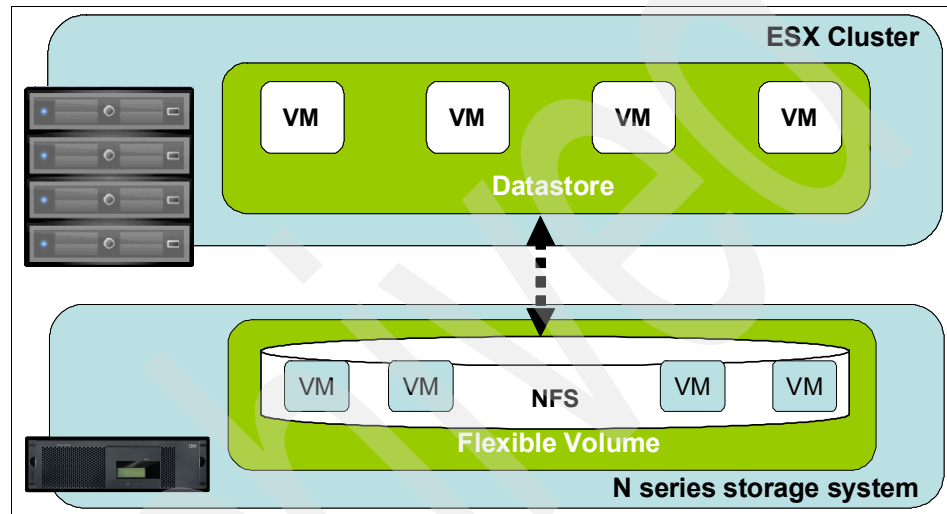


Figure 1-3 ESX server connected to an NFS datastore

### 1.2.3 Storage overview: raw device mappings

VMware ESX allows for virtual machines to have direct access to LUNs for specific use cases such as P2V clustering or storage vendor management tools. This type of access is referred to as a *raw device mapping* and can be configured with Fibre Channel, iSCSI, and Fibre Channel over Ethernet. In this design, VMware ESX acts as a connection proxy between the VM and the storage array.

Unlike VMFS and NFS, RDMs are not used to provide shared datastores. RDMs are an enabling technology for solutions such as virtual machine, and physical-to-virtual-machine host-based clustering such as with Microsoft Cluster Server (MSCS). RDMs provide traditional LUN access to a host, so they can achieve high individual disk I/O performance, and they can be easily monitored for disk performance by a storage array.

The N series can enhance the use of RDMs by providing array-based LUN-level thin provisioning, production-use data deduplication, advanced integration components such as SnapDrive®, VM granular Snapshot copies, and FlexClone® zero-cost cloning of RDM-based data sets.

The challenges of this solution are that VMware clusters might have to be limited in size, and this design requires ongoing interaction between storage and VMware administration teams. Figure 1-4 shows an example of this configuration.

RDMs are available in two modes: physical and virtual. Both modes support key VMware features such as vMotion and can be used in both HA and DRS clusters. The key difference between the two technologies is the amount of SCSI virtualization that occurs at the VM level. This difference results in certain limitations regarding MSCS and VMware Snapshot use case scenarios.

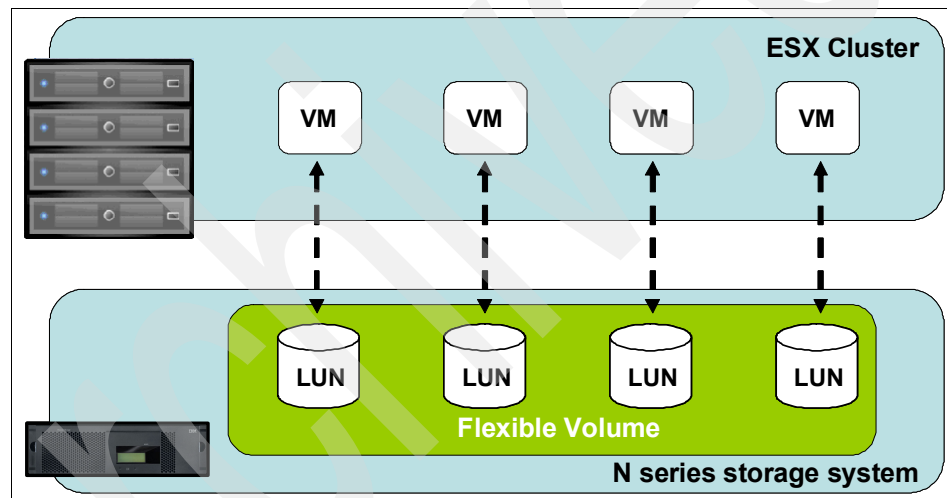


Figure 1-4 ESX Cluster connected to a VMFS datastore using FC or iSCSI

## 1.3 Datastore comparison tables

Differentiating what is available with each type of datastore and storage protocol can require the consideration of many aspects. The following tables compare the features that are available with each storage option. Table 1-1 on page 7 shows supported datastore features.

Table 1-1 Datastore support features

Capability or feature	FC/FCoE	iSCSI	NFS
Format	VMFS or RDM	VMFS or RDM	N series WAFL
Maximum number of datastore LUNs	256	256	64
Maximum datastore size	64TB	64TB	16TB
Maximum LUN/NAS file system size	2TB	2TB	16TB
VMDKs per LUN/NAS file system	16	16	250
Optimal queue depth and LUN / file system	64	64	N/A
Available link speeds	4 and 8 Gb FC and 10 GbE	1 and 10 GbE	1 and 10 GbE

Table 1-2 shows VMware-supported functionality when using the N series as the storage system.

Table 1-2 VMware supported functionality

Capability or feature	FC/FCoE	iSCSI	NFS
vMotion	Yes	Yes	Yes
Storage vMotion	Yes	Yes	Yes
VMware HA	Yes	Yes	Yes
DRS	Yes	Yes	Yes
VCB	Yes	Yes	Yes
MSCS within a VM	Yes, using RDM	not supported	not supported
Fault Tolerance	Yes, VMFS only	Yes, VMFS only	Yes
Thin provisioned VMDK	Yes	Yes	Yes
VMware NMP	Yes	Yes	N/A

Table 1-3 shows supported N series storage management features.

*Table 1-3 N series supported storage management features*

Capability or Feature	FC/FCoE	iSCSI	NFS
Data deduplication	Savings in the array	Savings in the array	Savings in the datastore
Thin provisioning	Datastore or RDM	Datastore or RDM	Datastore
Resize Datastore	Grow only	Grow only	Grow, auto-grow, and shrink
SANscreen VMinsight	Yes	Yes	Yes
SnapDrive	Yes	Yes, using GOS initiator	No
SnapManager for Virtual Infrastructure (SMVI)	Yes	Yes	Yes
VMware ESX host utilities	Yes	Yes	Yes

Table 1-4 shows supported N series backup features.

*Table 1-4 Supported backup features*

Capability or Feature	FC/FCoE	iSCSI	NFS
SnapShot backups	Yes	Yes	Yes
SnapMirror®	Datastore or RDM	Datastore or RDM	Datastore or VM
SnapVault®	Datastore or RDM	Datastore or RDM	Datastore or VM
VMDK image access	VCB	VCB, Windows only	VCB, VIC file explorer
VMDK file level access	VCB, Windows only	VCB, Windows only	VCB and third-party applications
NDMP granularity	Datastore	Datastore	Datastore or VM



## **N series configuration and setup**

This chapter offers guidance for setting up and configuring N series storage systems.

## 2.1 Data protection

This section describes data protection concepts.

### 2.1.1 RAID and data protection

A by-product of any consolidation effort is increased risk if the consolidation platform fails. As physical servers are converted to virtual machines and multiple VMs are consolidated onto a single physical platform, the impact of a failure to the single platform can be catastrophic. Fortunately, VMware provides multiple technologies that enhance availability of a virtual data center. These technologies include physical server clustering using VMware HA, application load balancing with DRS, and the ability to nondisruptively move running VMs and data sets between physical VMware ESX Servers with vMotion and Storage vMotion, respectively.

When focusing on storage availability, many levels of redundancy are available for deployments, including purchasing physical servers with multiple storage interconnects or host bus adapters (HBAs), deploying redundant storage networking and network paths, and using storage arrays with redundant controllers. A deployed storage design that meets all these criteria can be considered to have eliminated all single points of failure.

The reality is that data protection requirements in a virtual infrastructure are greater than those in a traditional physical server infrastructure. Data protection is a paramount feature of shared storage devices. N series RAID-DP™ (see Figure 2-1 on page 11) is an advanced RAID technology that is provided as the default RAID level on all N series systems. RAID-DP protects against the simultaneous loss of two drives in a single RAID group. It is very economical to deploy; the overhead with default RAID groups is a mere 12.5%. This level of resiliency and storage efficiency makes data residing on RAID-DP safer than data stored on RAID 5 and more cost effective than RAID 10. Use RAID-DP on all RAID groups that store VMware data.





## FlexVol™ dynamic virtualization: *Designed to offer dramatic improvement in storage management*

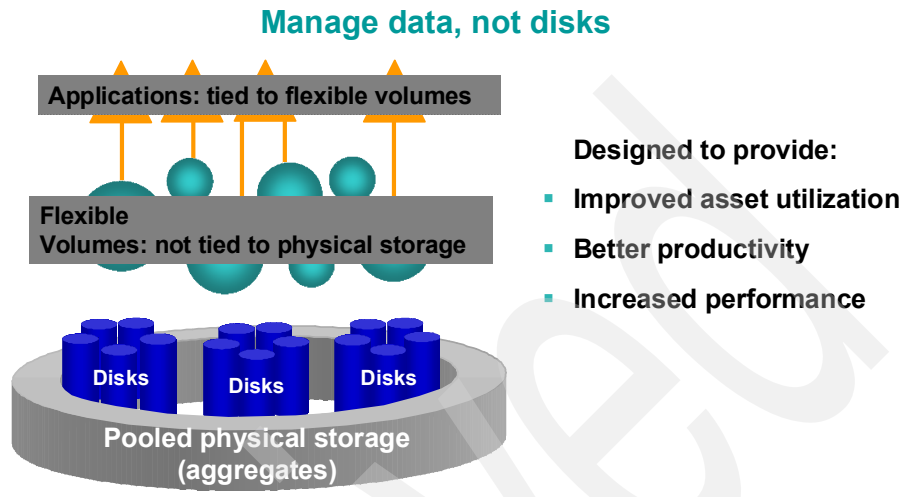


Figure 2-2 Aggregates and flexible Volumes

Place the remaining storage into a small number of large aggregates. The overall disk I/O from VMware environments is traditionally random by nature, so this storage design gives optimal performance because a large number of physical spindles are available to service I/O requests. On smaller N series storage systems, having more than a single aggregate might not be practical because of the restricted number of disk drives on the system. In these cases, having only a single aggregate is acceptable.

## 2.2 HA mode for FC configurations

N series HA systems are configured with an option known as *cfmode*, which controls the behavior of the system's Fibre Channel (FC) ports if a controller failover occurs. This option must be set as Single System Image (SSI). If you are deploying VMware ESX on an older HA array with FC or Fibre Channel over Ethernet (FCoE), be sure that *cfmode* is set to SSI.

To verify the current *cfmode*, perform the following steps:

1. Connect to the N series system console (using either SSH, Telnet, or Console connection).
2. Enter **fcp show cfmode** command.

To set the `cfmode`, perform the following steps:

1. Connect to the N series system console (using either SSH, Telnet, or Console connection).
2. If `cfmode` must be changed, enter the FC `set cfmode single_image` command.

## 2.3 Storage configuration

This sections describes N series storage configurations.

### 2.3.1 Flexible volumes

Flexible volumes contain either LUNs or virtual disk files that are accessed by VMware ESX Servers. Use a one-to-one alignment of VMware datastores to flexible volumes.

This design offers an easy means to understand the VMware data layout when viewing the storage configuration from the N series storage system. This mapping model also makes the implementing of Snapshot backups and SnapMirror replication policies at the datastore level easy, because these storage side features are implemented at the flexible volume level.

### 2.3.2 Snapshot reserve

Flexible volumes should be configured with the snap reserve set to 0 (zero) and the default Snapshot schedule disabled. All Snapshot copies must be coordinated with the VMware ESX Servers for data consistency. Snapshot copies are described in Appendix A, “Configuring SSH on VMware ESX servers and N series systems” on page 117.” To set the volume options for Snapshot copies to the recommended setting, perform the following steps:

1. Log into the N series storage system console.
2. Set the volume Snapshot schedule:  
`snap sched <vol-name> 0 0 0`
3. Set the volume Snapshot reserve:  
`snap reserve <vol-name> 0`

### 2.3.3 LUNs

LUNs are units of storage provisioned from a N series storage system directly to the ESX Servers. The VMware ESX Server can access the LUNs in two ways:

- ▶ The first and most common method is as storage to hold virtual disk files for multiple virtual machines. This type of usage is referred to as a VMFS datastore.
- ▶ The second method is as a raw device mapping (RDM). With RDM, the ESX Server accesses the LUN, which in turn passes access directly to a virtual machine for use with its native file system, such as NTFS or EXT3.

### 2.3.4 Storage naming conventions

N series storage systems allow human or canonical naming conventions. In a well-planned virtual infrastructure implementation, a descriptive naming convention aids in identification and mapping through the multiple layers of virtualization from storage to the virtual machines. A simple and efficient naming convention also facilitates configuration of replication and disaster recovery processes.

Consider the following naming guidelines:

- ▶ FlexVol® name: This name should match the name of the datastore.
- ▶ LUN name for VMFS: This name should match the name of the datastore.
- ▶ LUN name for RDMs: This name should include both the host name and volume label or name.

## 2.4 vSphere in a N series MetroCluster environment

N series MetroCluster configurations consist of a pair of active-active storage controllers configured with mirrored aggregates and extended distance capabilities to create a high-availability solution. The primary benefits are as follows:

- ▶ Higher availability with geographic protection
- ▶ Minimal risk of lost data, easier management and recovery, and reduced system downtime
- ▶ Quicker recovery when a disaster occurs
- ▶ Minimal disruption to users and client applications

A MetroCluster (either Stretch or Fabric) behaves in most ways similar to an active-active configuration. All of the protection provided by core N series technology (RAID-DP, Snapshot copies, automatic controller failover) also exists in a MetroCluster configuration. MetroCluster adds, however, complete synchronous mirroring and the ability to perform a complete site failover, from a storage perspective, with a single command.

The following N series MetroCluster types exist and work seamlessly with the complete VMware vSphere and ESX server portfolio:

- ▶ Stretch MetroCluster (sometimes referred to as *nonswitched*)  
This type is an active-active configuration that can extend up to 500 m, depending on speed and cable type. It also includes synchronous mirroring (SyncMirror®) and the ability to do a site failover with a single command.
- ▶ Fabric MetroCluster (also referred to as *switched*)  
This type uses four Fibre Channel switches in a dual-fabric configuration and a separate cluster interconnect card to achieve an even greater distance (up to 100 km depending on speed and cable type) between primary and secondary locations.

The integration of the MetroCluster and VMware vSphere provides seamless integration, and storage and application redundancy. In addition to connecting to the vSphere environment using FCP, iSCSI, or NFS, the solution is able to serve other network clients with CIFS, HTTP, and FTP at the same time. The solution shown Figure 2-3 on page 16 provides redundant VMware vSphere server, redundant N series heads, and redundant storage.

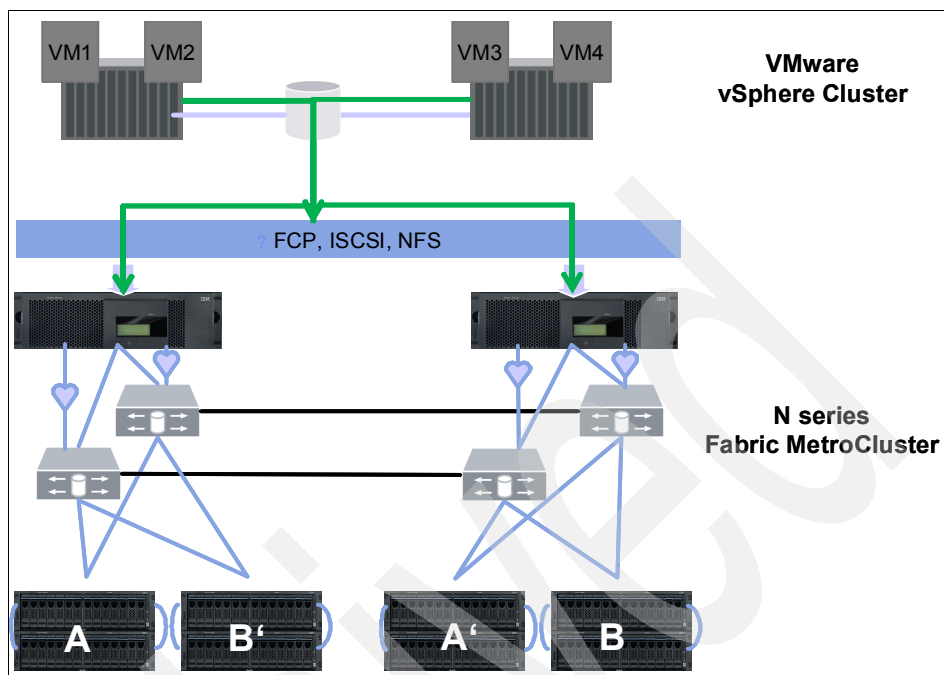


Figure 2-3 MetroCluster and VMware vSphere integrated solution

For more information about N series MetroCluster, see the “MetroCluster” chapter in *IBM System Storage N series*, SG24-7129.



## **VMware ESX FC, FCOE, and iSCSI storage configuration**

This chapter describes block-based storage configurations.

## 3.1 LUN sizing for VMFS datastores

VMFS datastores offer a simple method for provisioning shared storage pools, with any storage architecture, to implement a design that can address the performance needs of the infrastructure. A common issue we see is customers overloading very large datastores with too many VMs. In this scenario, the I/O load must be leveled. VMware provides Storage vMotion as a means to redistribute VM storage to alternative datastores without disruption to the VM. Large VMFS datastores commonly reach their I/O performance limit before their capacity limit has been reached.

Although there is no definitive recommendation, a commonly deployed size for a VMFS datastore is somewhere in the range of 300 - 700 GB. The maximum supported LUN size is 2 TB. Larger datastores can be created through VMFS spanning. VMFS spanning uses VMFS extents to concatenate multiple partitions into a single datastore.

Advanced storage technologies such as thin provisioning, which is available with VMware Virtual Machine Disks (VMDKs) and N series datastores, can return provisioned but unused storage back to the N series storage pool for reuse. Unused storage does not include storage that contains data that has been deleted or migrated as part of a Storage vMotion process.

## 3.2 Cluster sizing considerations when using LUNs

A VMware cluster is collectively bound to the same limits of an individual VMware ESX server. Currently, the maximum number of LUNs that can be connected to a cluster is 256 LUNs. This limitation typically comes into consideration with VMFS spanned datastores or RDM-based deployments.

Based on LUN limits, the following formula can be used to determine the maximum number of VMware ESX nodes per VMware ESX cluster. This formula implies that all nodes in a cluster are connected to all shared LUNs:

$$254 / (\text{number of RDMS per VM}) / (\text{planned number of VMs per VMware ESX host}) = \text{number of VMware ESX nodes in a data center}$$

For example, the formula for 2 RDMs per VM with 20 VMs per VMware ESX Server is as follows:

$$254/2/20 = 6.35 \text{ rounded up} = 7 \text{ VMware ESX Servers required in the cluster}$$

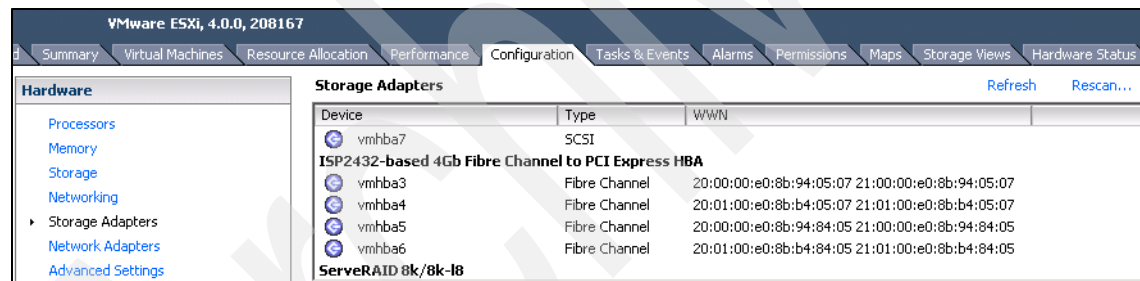


### 3.3 FC, FCoE, and iSCSI LUN provisioning

When provisioning LUNs for access by using FC or iSCSI, the LUNs must be masked so that the appropriate hosts can connect only to them. With N series systems, LUN masking is handled by the creation of initiator groups. Create an igroup for each VMware cluster. We also recommend including in the name of the igroup the name of the cluster and the protocol type (for example, DC1\_FC and DC1\_iSCSI). This naming convention and method simplify the management of initiator groups (igroups) by reducing the total number created. It also means that all VMware ESX Servers in the cluster see each LUN at the same ID. Each initiator group includes all of the FC worldwide port names (WWPNs) or iSCSI qualified names (IQNs) of the VMware ESX Servers in the VMware cluster.

**Note:** If a cluster will use multiple block-based protocols, separate igroups must be created for each.

For assistance in identifying the WWPN or IQN of the VMware ESX Server, select **Storage Adapters** on the Configuration tab for each VMware ESX Server in vCenter Server, and see the WWN column (Figure 3-1).



Device	Type	WWN
vmhba7	SCSI	
<b>ISP2432-based 4Gb Fibre Channel to PCI Express HBA</b>		
vmhba3	Fibre Channel	20:00:00:e0:8b:94:05:07 21:00:00:e0:8b:94:05:07
vmhba4	Fibre Channel	20:01:00:e0:8b:b4:05:07 21:01:00:e0:8b:b4:05:07
vmhba5	Fibre Channel	20:00:00:e0:8b:94:84:05 21:00:00:e0:8b:94:84:05
vmhba6	Fibre Channel	20:01:00:e0:8b:b4:84:05 21:01:00:e0:8b:b4:84:05
<b>ServeRAID 8k/8k-l8</b>		

Figure 3-1 Identifying WWPN and IQN numbers using the vSphere client

LUNs can be created by using the N series LUN Wizard in the N series system console or by using the FilerView® GUI, as follows:

1. Log in to FilerView.
2. Select **LUNs**.
3. Select **Wizard**.
4. In the Wizard window, click **Next**.
5. Enter the path (see Figure 3-2 on page 20).
6. Enter the LUN size.
7. Enter the LUN type (for VMFS select VMware; for RDM select the VM type).
8. Enter a description and click **Next**.

**LUN Wizard: Specify New Group Parameters**

**Name:**  
Enter a name for the initiator group.  ?

**Type:**  
Select a type for the initiator group.  ?

**Operating System:**  
Select the operating system type of the initiators in the group.  ?

Figure 3-2 LUN Wizard

The next step in the LUN Wizard is LUN masking, which is accomplished by assigning an igroup to a LUN. With the LUN Wizard, you can either assign an existing igroup or create a new igroup.

**Important:** The VMware ESX Server expects a LUN ID to be the same on every node in a VMware ESX cluster. Therefore, create a single igroup for each cluster rather than for each VMware ESX Server.

To configure LUN masking on a LUN created in the FilerView GUI, perform the following steps

1. Select **Add Group**.
2. Do *one* of the following steps:
  - Select **Use Existing Initiator Group** and click **Next**. Select the group from the list and either assign a LUN ID or leave the field blank (the system will assign an ID). Click **Next** to complete the task.
  - Select **Create a New Initiator Group** and click **Next**. Supply the igroup parameters, including name, type of connectivity (FC or iSCSI), and operating system type (VMware), and then click **Next** (see Figure 3-3 on page 21)

**LUN Wizard: Specify New Group Parameters**

**Name:**  
Enter a name for the initiator group.  ?

**Type:**  
Select a type for the initiator group.  ?

**Operating System:**  
Select the operating system type of the initiators in the group.  ?

Figure 3-3 Assigning an igroup to a LUN

3. For the systems that will connect to this LUN, enter the new SAN identifiers or select the known identifiers (WWPN or IQN).
4. Click **Add Initiator**.
5. Click **Next** to complete the task.

## 3.4 Connecting FC and FCoE datastores

The Fibre Channel service is the only storage protocol that is running by default on the VMware ESX Server. Be sure that each VMware ESX Server has two FC HBA ports available for storage path redundancy. To connect to FC LUNs provisioned on an N series system, perform the following steps:

1. Open vCenter Server.
2. Select a VMware ESX host.
3. In the pane on the right, select the **Configuration** tab.
4. In the Hardware box, select the **Storage Adapters** link.
5. In the upper-right corner, select the **Rescan** link. Selecting Rescan forces the rescanning of all HBAs (FC, FCoE, and iSCSI) to discover changes in the storage available to the VMware ESX Server.
6. Repeat these steps for each VMware ESX Server in the cluster.

To add a LUN as a datastore, perform the following steps:

1. Open vCenter Server.
2. Select a VMware ESX host.
3. In the pane on the right, select the **Configuration** tab.
4. In the Hardware box, select the **Storage** link and then click **Add Storage**. The Add Storage wizard opens, as shown in Figure 3-4.
5. Select **Disk/LUN** and click **Next**.

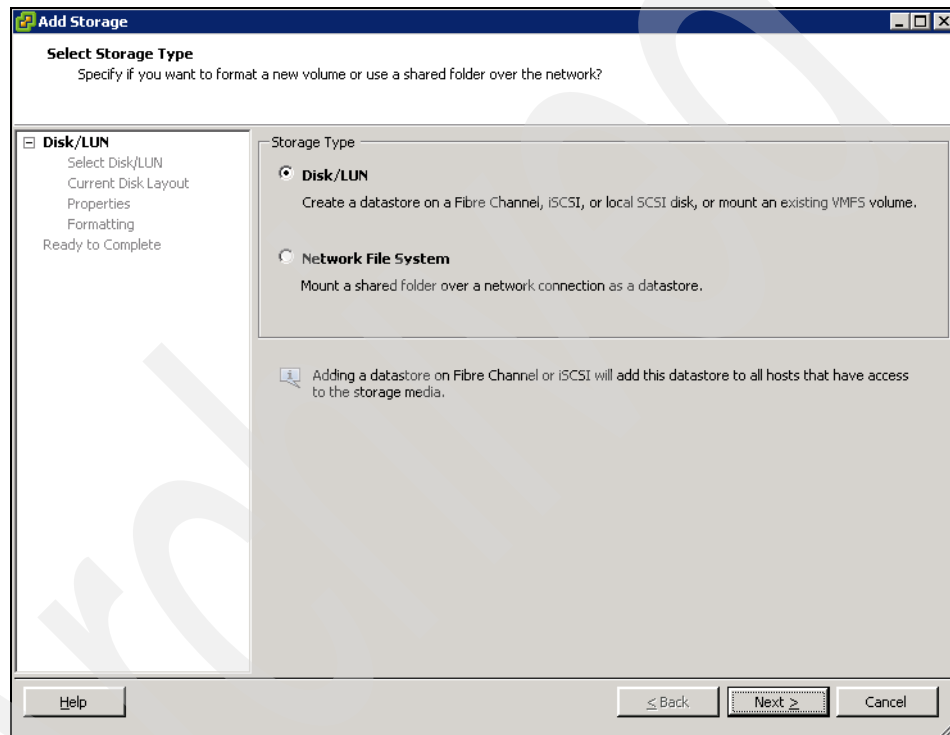


Figure 3-4 VMware Add Storage wizard

6. Select the LUN to use and click **Next**.
7. Enter a name for the datastore and click **Next**.
8. Select the block size, click **Next**, and click **Finish**.

The default block size of a virtual machine file system is 1 MB. This block size supports storing virtual disk files up to a maximum of 256 GB in size. If you plan to store virtual disks larger than 256 GB in the datastore, you must increase the block size to be greater than the default (see Figure 3-5 on page 23).

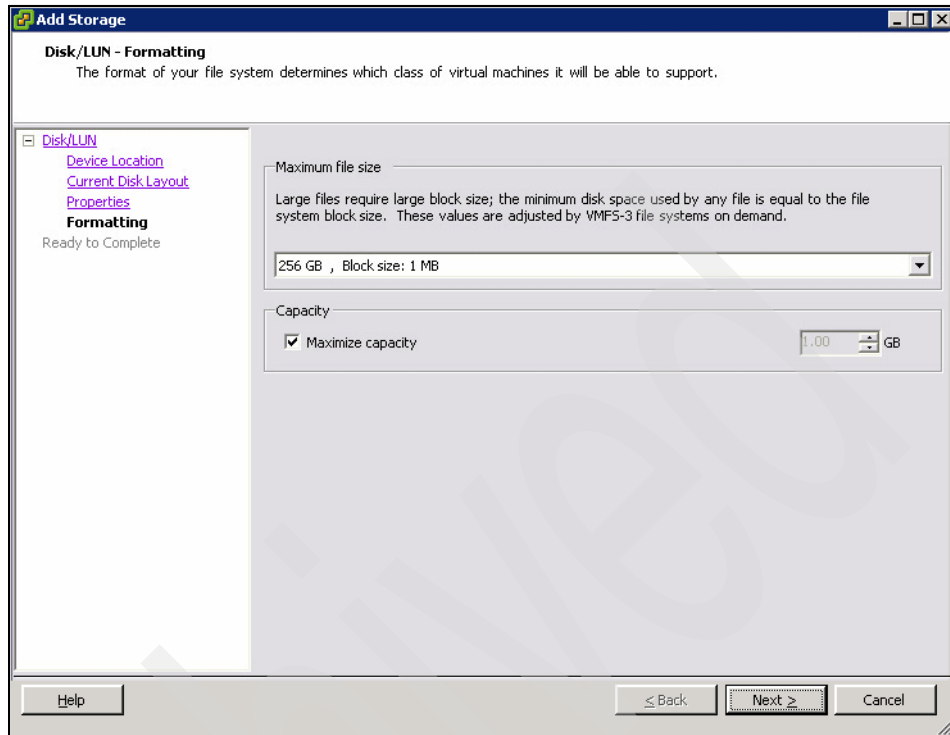


Figure 3-5 Formatting LUN with VMFS

## 3.5 Connecting iSCSI datastores

As a best practice, separate the IP-based storage traffic from the public IP network traffic by implementing separate physical network segments or VLAN segments. This design follows the architecture of SCSI and FC connectivity. New with VMware ESX and VMware ESX 4.0 is the ability to enable multiple TCP sessions with iSCSI datastores. Enabling multiple TCP sessions with the VMware ESX Round Robin Path Selection Plug-in (PSP) can allow iSCSI datastores to send I/O over every available path to an iSCSI target.

For more information about configuring NMP for iSCSI, see 4.4, “Configuring the Round Robin PSP” on page 36.

Creating a second network in a VMware ESX server requires the creation of a second vSwitch so that public traffic can be diverted to other physical NICs. The VMware ESX Server will require a VMkernel port to be defined on the new vSwitch.

Each VMware ESX Server should have a service console port defined on the vSwitch that transmits public virtual machine traffic and on the vSwitch configured for IP storage traffic. This second service console port adds the redundancy in VMware ESX HA architectures and follows VMware ESX HA best practices.

With this design, do not allow routing of data between these networks (do not define a default gateway for the iSCSI storage network). With this model, iSCSI deployments require that a second service console port be defined on the VMkernel storage virtual switch within each VMware ESX server.

IP storage network, or VMkernel, connectivity can be verified by the use of the **vmkping** command. With iSCSI-connected LUNs, test connectivity by using the **vmkping <iSCSI target>** syntax.

### 3.5.1 Enabling iSCSI communications

To enable iSCSI communications, perform the following steps:

1. Open vCenter Server.
2. Select a VMware ESX host.
3. In the pane on the right, select the **Configuration** tab.
4. In the Configuration tab, left pane, select **Security Profile**.
5. In the pane on the right, select the **Properties** link to open the Firewall Properties window (see Figure 3-6 on page 25).
6. Select the **Software iSCSI Client** check box and then click **OK** to close the Firewall Properties window.

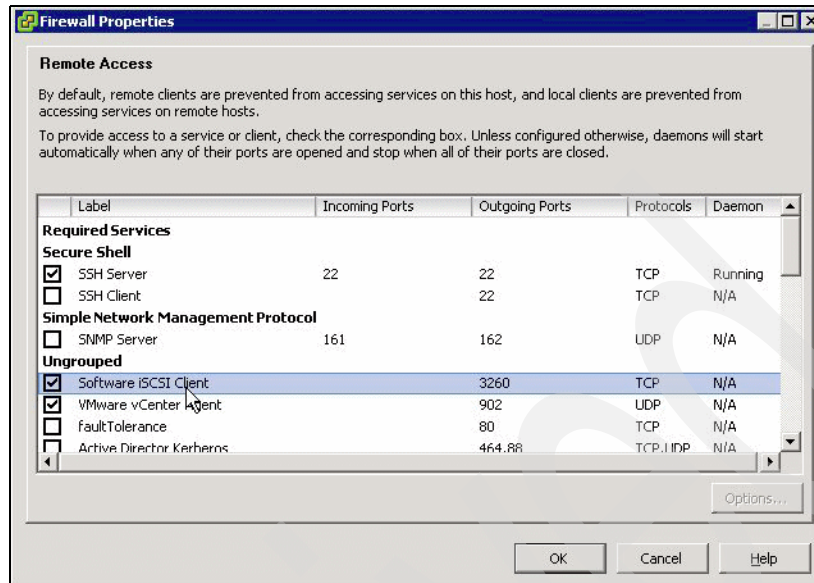


Figure 3-6 Configuring the firewall in VMware ESX

### 3.5.2 Creating multiple iSCSI VMkernels

To create multiple iSCSI VMkernels, perform the following steps:

1. Open vCenter Server.
2. Select a VMware ESX host.
3. In the pane on the right, select the **Configuration** tab.
4. In the Hardware box, select **Networking**.
5. In the upper-right corner, click **Add Networking**. The Add Network wizard opens.
6. Select **VMkernel**, and click **Next**.
7. Create a VMkernel for every Ethernet link that you want to dedicate to iSCSI traffic. Be aware that VMkernels can be on separate IP subnets. This configuration is required if combining iSCSI with NFS datastore access.

Configure the VMkernel by providing the required network information. A default gateway is not required for the VMkernel IP storage network.

Each VMkernel must be configured to use with a single active adapter (such as VMNIC0) that is not used by any other iSCSI VMkernel. Also, each VMkernel must not have any standby adapters. Figure 3-7 shows iSCSI VMkernel 0; the active adapter is vmnic0.

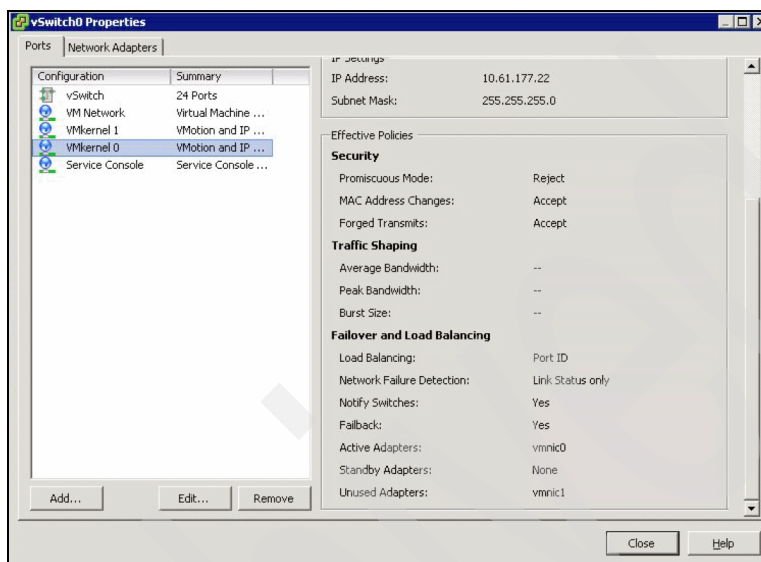


Figure 3-7 iSCSI VMkernel 0; active adapter is vmnic0



Figure 3-8 shows iSCSI VMkernel 1; active adapter is vmnic1.

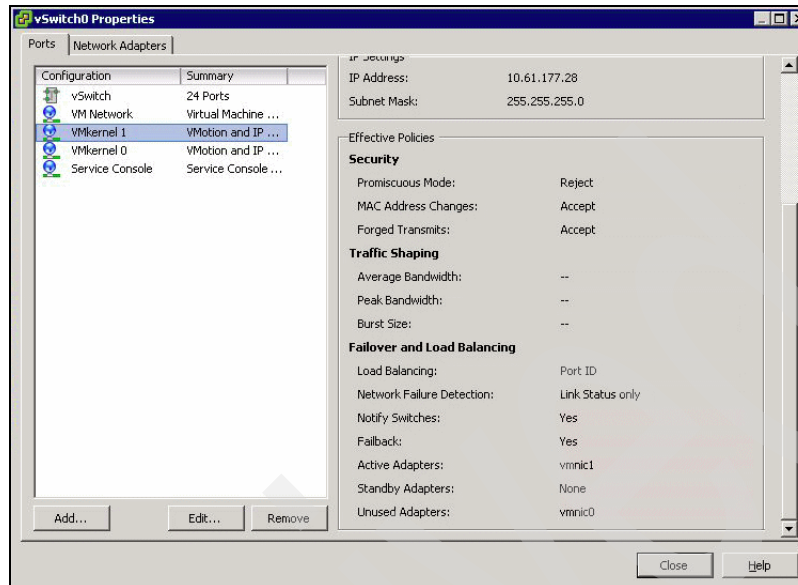


Figure 3-8 iSCSI VMkernel 1; active adapter is vmnic1

8. (The software iSCSI daemon must be bound to each VMkernel. This step can be completed only by using the CLI. See Figure 3-9 on page 28.) Connect to a VMware ESX or VMware ESXi console and run the following command:

```
esxcli swiscsi nic add -n <VMkernel ID> -d <Virtual HBA ID>
```

Examples of the command are as follows:

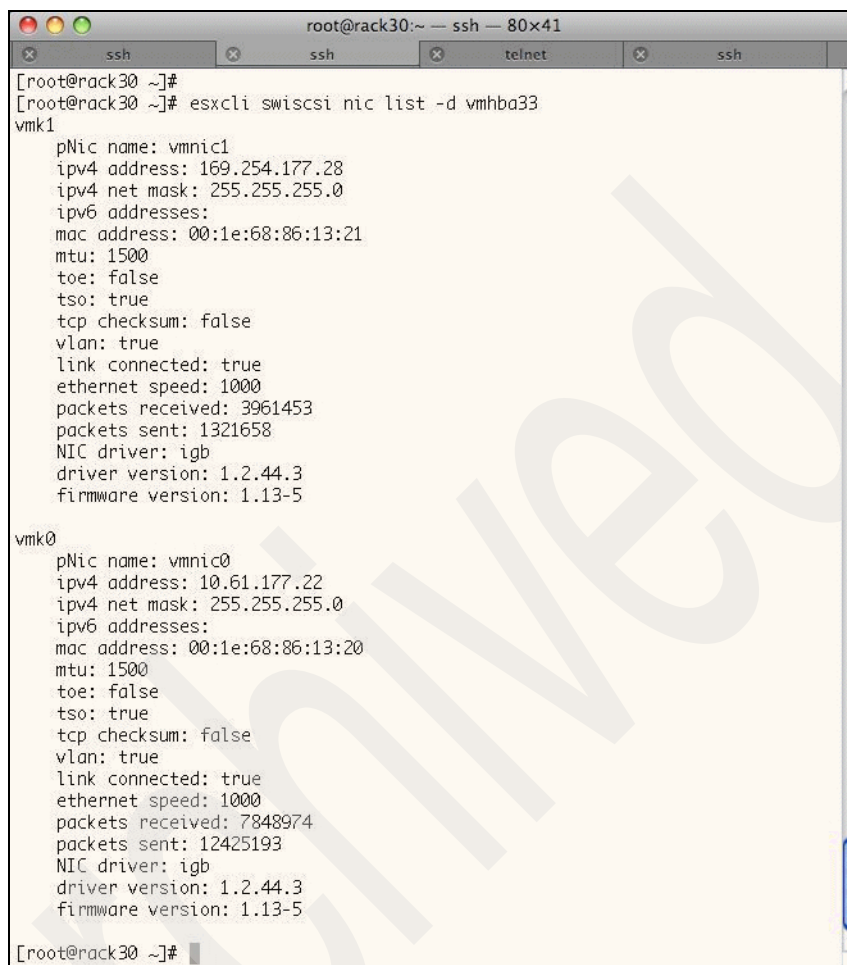
```
esxcli swiscsi nic add -n vmk0 -d vmhba33
esxcli swiscsi nic add -n vmk1 -d vmhba33
```

9. Verify the iSCSI to VMkernel bindings. Connect to a VMware ESX or VMware ESXi console and run the following command:

```
esxcli swiscsi nic list -d <Virtual HBA ID>
```

An example of the command is as follows:

```
esxcli swiscsi nic list -d vmhba33
```



```
root@rack30:~ — ssh — 80x41
[ssh] [ssh] [telnet] [ssh]
[root@rack30 ~]#
[root@rack30 ~]# esxcli swiscsi nic list -d vmhba33
vmk1
  pNic name: vmnic1
  ipv4 address: 169.254.177.28
  ipv4 net mask: 255.255.255.0
  ipv6 addresses:
  mac address: 00:1e:68:86:13:21
  mtu: 1500
  toe: false
  tso: true
  tcp checksum: false
  vlan: true
  link connected: true
  ethernet speed: 1000
  packets received: 3961453
  packets sent: 1321658
  NIC driver: igb
  driver version: 1.2.44.3
  firmware version: 1.13-5

vmk0
  pNic name: vmnic0
  ipv4 address: 10.61.177.22
  ipv4 net mask: 255.255.255.0
  ipv6 addresses:
  mac address: 00:1e:68:86:13:20
  mtu: 1500
  toe: false
  tso: true
  tcp checksum: false
  vlan: true
  link connected: true
  ethernet speed: 1000
  packets received: 7848974
  packets sent: 12425193
  NIC driver: igb
  driver version: 1.2.44.3
  firmware version: 1.13-5

[root@rack30 ~]#
```

Figure 3-9 Verifying iSCSI to VMkernel bindings

### 3.5.3 Connecting to iSCSI targets

To connect to iSCSI targets, perform the following steps:

1. Open vCenter Server.
2. Select a VMware ESX host.
3. In the pane on the right, select the **Configuration** tab.
4. In the Hardware box, select **Storage Adapters**.
5. Highlight the iSCSI adapter, and click the **Properties** link in the Details box (Figure 3-10 on page 29).

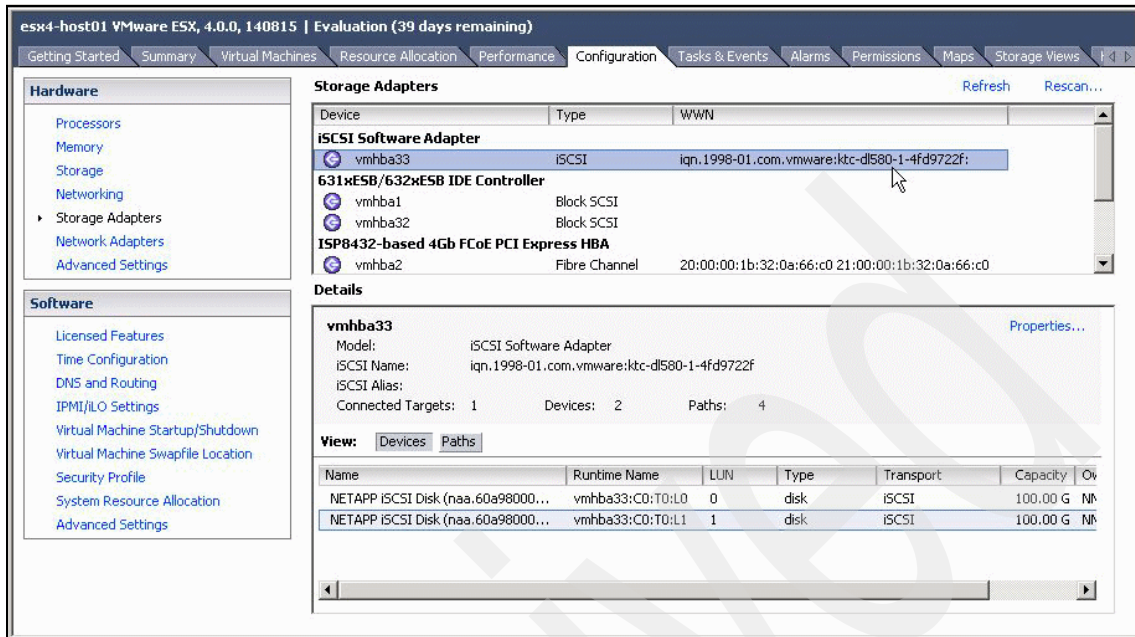


Figure 3-10 Selecting an iSCSI initiator

6. Select the **Dynamic Discovery** tab in the iSCSI Initiator Properties box.
7. Click **Add**, and enter the IP address of the iSCSI-enabled interface on the N series storage system (see Figure 3-11 on page 30).

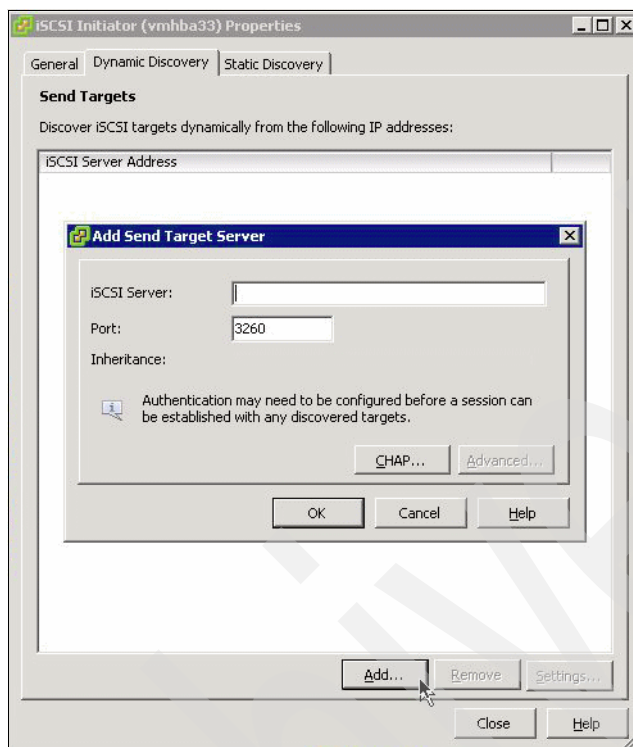


Figure 3-11 Configuring iSCSI dynamic discovery

8. For an additional layer of security, select the CHAP tab to configure CHAP authentication. Be sure to set up and verify iSCSI access *before* enabling CHAP authentication.

### 3.5.4 Restricting iSCSI targets to preferred interfaces

By default, N series storage systems provide iSCSI access over every network interface. This default configuration might not be optimal because it can lead to conditions where VMware ESX servers attempt to communicate to interfaces that are unreachable. Be sure to disable iSCSI on N series network interfaces over which you do not want to send iSCSI traffic.

Data ONTAP® allows this filtering to be accomplished by either of the following methods (use one of these methods to configure iSCSI access restrictions):

- ▶ On a host-by-host basis using iSCSI access lists
- ▶ On a global basis by unbinding iSCSI to a specific interface or set of interfaces

Host restricted iSCSI access lists currently require each IQN of a VMware ESX server to be configured on the array. This process is more granular and might lead to additional tasks each time a new host is introduced into the data center. To configure iSCSI access lists, perform the following steps:

1. Connect to the N series system console (using either SSH, Telnet, or Console connection).
2. Create an iSCSI access list by using the following command:  
`iscsi interface accesslist add <ESX iqn address>`  
Repeat this step for each VMware ESX host in the data center.
3. Verify the iSCSI access list type by using the following command:  
`iscsi interface accesslist show`

Globally disabling iSCSI traffic on a set of network interfaces is less granular than iSCSI access lists; however, it is much simpler to configure. To configure it, perform the following steps:

1. Connect to the N series system console (using either SSH, Telnet, or Console connection).
2. Disable iSCSI on an interface type by using the following command:  
`iscsi interface disable <interface hw address>`
3. Verify the iSCSI bindings type by using the following command:  
`iscsi interface show`

Alternatively iSCSI restrictions can be configured within VMware ESX 4.0 and VMware ESXi 4.0. This method provides for another way to accomplish optimal I/O communications if an appropriate restriction cannot be configured with Data ONTAP (see Figure 3-12).

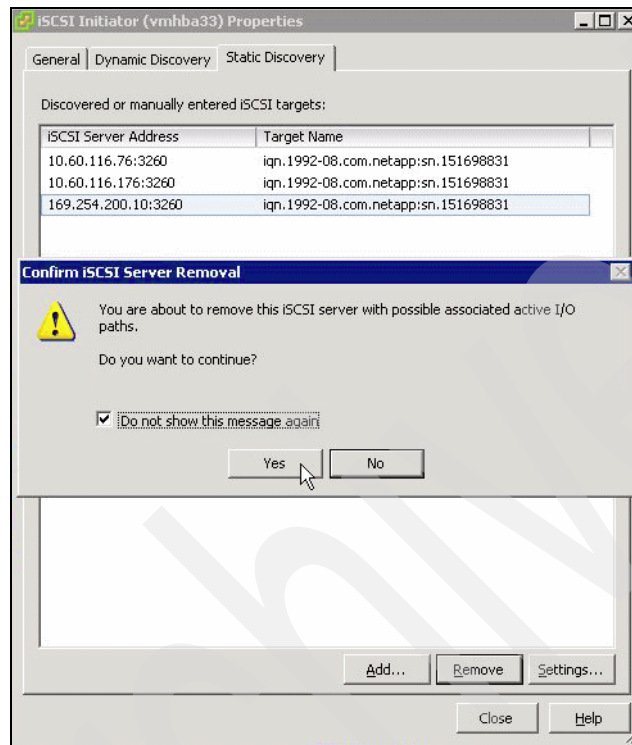


Figure 3-12 Configuring iSCSI to filter or restrict access to undesired iSCSI targets

## VMware native multipathing

VMware ESX servers include a native multipathing solution for FC, FCoE, and iSCSI storage networks, which enable high-performance data access and enhanced storage resiliency. With the release of VMware ESX and VMware ESXi 4.0, VMware has introduced the concept of a Pluggable Storage Architecture (PSA), which in turn introduced several concepts to its Native Multipathing (NMP).

This chapter describes the use of the Storage Array Type Plug-in (SATP), the Path Selection Plug-in (PSP), and the Asymmetric Logical Unit Access (ALUA) protocol.

## 4.1 Default NMP settings

Connecting an N series storage system to a VMware ESX 4.0 server results in the array being identified as an active-active storage controller, and the VMware native multipathing path selection policy applies the *Fixed* multipathing policy. This configuration is identical to the default behavior with VMware ESX 3.5 (see Figure 4-1).

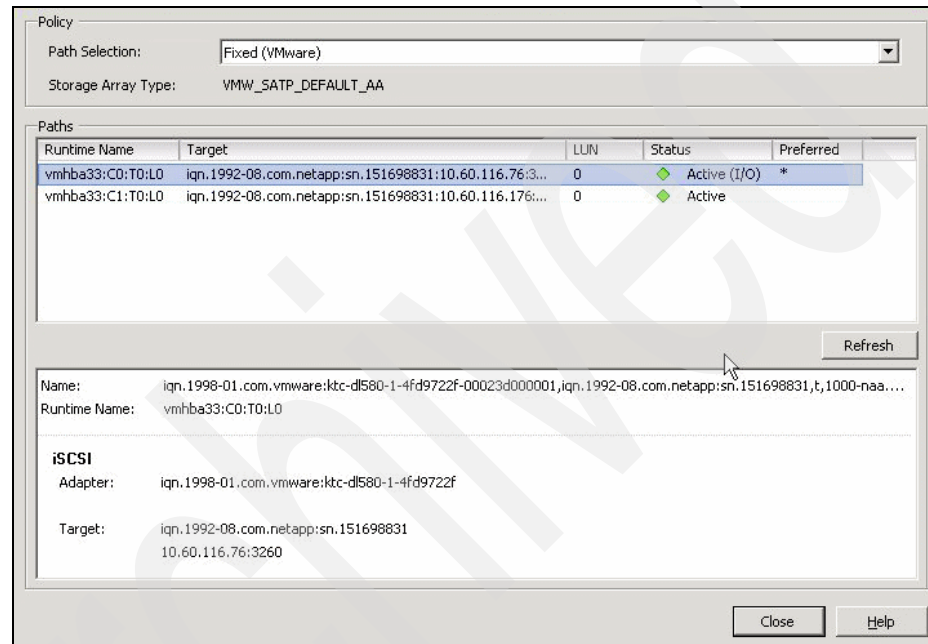


Figure 4-1 Default path selection policy

Deployments that use the Fixed multipathing policy are required to manually identify and set the I/O to traverse the primary FC paths. In addition, users of this configuration are required to manually load-balance I/O across the primary paths. The N series ESX Host Utilities (EHU) can automate this process for environments that prefer the NMP Fixed PSP.

For deployments that prefer a complete *plug-n-play* architecture, enable ALUA on the N series storage array and configure the Round Robin PSP.



## 4.2 Enabling ALUA

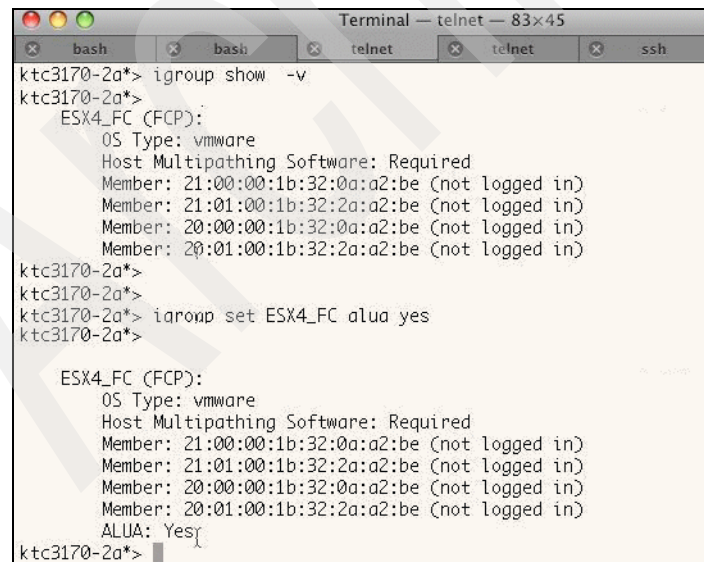
N series and VMware ESX support the Asymmetric Logical Unit Access (ALUA) protocol. ALUA allows for the auto negotiation of paths between SCSI target devices and target ports enabling dynamic reconfiguration.

Enabling ALUA on N series initiator groups can result in a more dynamic, or plug-n-play-like, architecture.

**Note:** ALUA is supported with VMware ESX and VMware ESXi for FC and FCoE. Support for iSCSI is not required because iSCSI does address this functionality, natively within the protocol.

ALUA is enabled on VMware ESX 4.0 by default. To enable ALUA on a N series storage array, perform the following steps:

1. Log in to the N series console.
2. From the storage appliance console, run the following command:  
`igroup set <igroup-name> alua yes`  
Repeat this step for each LUN accessed by VMware ESX.
3. Verify the results by running the following command (Figure 4-2):  
`igroup show -v <igroup-name>`



```
Terminal — telnet — 83x45
x bash x bash x telnet x telnet x ssh
ktc3170-2a*> igroup show -v
ktc3170-2a*>
  ESX4_FC (FCP):
    OS Type: vmware
    Host Multipathing Software: Required
    Member: 21:00:00:1b:32:0a:a2:be (not logged in)
    Member: 21:01:00:1b:32:2a:a2:be (not logged in)
    Member: 20:00:00:1b:32:0a:a2:be (not logged in)
    Member: 20:01:00:1b:32:2a:a2:be (not logged in)
ktc3170-2a*>
ktc3170-2a*>
ktc3170-2a*> igroup set ESX4_FC alua yes
ktc3170-2a*>
  ESX4_FC (FCP):
    OS Type: vmware
    Host Multipathing Software: Required
    Member: 21:00:00:1b:32:0a:a2:be (not logged in)
    Member: 21:01:00:1b:32:2a:a2:be (not logged in)
    Member: 20:00:00:1b:32:0a:a2:be (not logged in)
    Member: 20:01:00:1b:32:2a:a2:be (not logged in)
    ALUA: Yes
ktc3170-2a*>
```

Figure 4-2 Enable and verify ALUA settings

## 4.3 Default NMP settings with ALUA enabled

Connecting an N series storage system to a VMware ESX 4.0 server with ALUA enabled results in the array and server being able to negotiate which paths are primary for I/O and which are to be used for failover. By enabling ALUA, the array will be identified as an ALUA-enabled storage controller, and the VMware ESX native multipathing path selection policy will apply the Most Recently Used or MRU multipathing policy.

Deployments that use ALUA along with the MRU multipathing policy are required to manually load-balance I/O across the primary paths. The result of only enabling ALUA is a reduction in several configuration requirements. For deployments that prefer a complete *plug-n-play* architecture, enable ALUA on the N series storage system and configure the Round Robin PSP.

## 4.4 Configuring the Round Robin PSP

The two ways to configure a PSP are as follows:

- ▶ The preferred way is to set the VMware ESX system default PSP for the VMware Default ALUA SATP to use the Round Robin PSP.
- ▶ Alternatively, you can manually manage datastore and LUN policies as was done in VMware ESX 3.5, inside the virtual infrastructure client.

### 4.4.1 Setting the default PSP for ALUA to Round Robin

To set the default PSP for ALUA to Round Robin, perform the following steps:

1. Connect to the CLI of a VMware ESX or VMware ESXi server, by running the following command from the console (see the example in Figure 4-3 on page 37):

```
esxcli nmp satp setdefault -psp <PSP type> -satp <SATP type>
```

The following PSP types are available:

- VMW\_PSP\_RR
- VM\_PSP\_FIXED
- VM\_PSP\_MRU

The following SATP types for N series arrays are available:

- VMW\_SATP\_DEFAULT\_AA
- VM\_SATP\_ALUA

```

root@k1c-d1580-1:~ — ssh — 83x45
[roo@k1c-d1580-1 ~]# esxcli nmp satp setdefaultpsp --psp VMW_PSP_RR --satp VMW_SAT
P_ALUA
Default PSP for VMW_SATP_ALUA is now VMW_PSP_RR
[roo@k1c-d1580-1 ~]#

```

Figure 4-3 Setting the RR PSP to the ALUA SATP

2. Verify the results by typing the following command (see Figure 4-4):

esxcli nmp satp list

```

root@k1c-d1580-1:~ — ssh — 83x45
[roo@k1c-d1580-1 ~]# esxcli nmp satp list
Name                Default PSP      Description
VMW_SATP_ALUA_CX    VMW_PSP_FIXED   Supports EMC CX that use the ALUA protocol
VMW_SATP_SVC        VMW_PSP_FIXED   Supports IBM SVC
VMW_SATP_MSA        VMW_PSP_MRU     Supports HP MSA
VMW_SATP_EQL        VMW_PSP_FIXED   Supports EqualLogic arrays
VMW_SATP_INV        VMW_PSP_FIXED   Supports EMC Invista
VMW_SATP_SYMM       VMW_PSP_FIXED   Supports EMC Symmetrix
VMW_SATP_LSI        VMW_PSP_MRU     Supports LSI and other arrays compatible with
the SIS 6.10 in non-AVT mode
VMW_SATP_EVA        VMW_PSP_FIXED   Supports HP EVA
VMW_SATP_DEFAULT_AP VMW_PSP_MRU     Supports non-specific active/passive arrays
VMW_SATP_CX         VMW_PSP_MRU     Supports EMC CX that do not use the ALUA
protocol
VMW_SATP_ALUA       VMW_PSP_RR      Supports non-specific arrays that use the ALUA
protocol
VMW_SATP_DEFAULT_AA VMW_PSP_RR      Supports non-specific active/active arrays
VMW_SATP_LOCAL      VMW_PSP_FIXED   Supports direct attached devices
[roo@k1c-d1580-1 ~]#

```

Figure 4-4 Listing the active PSP to SATP configurations

Repeat these steps on each VMware ESX or VMware ESXi server.

## 4.4.2 Manually setting the PSP for a datastore

To set up manually the PSP for a datastore, perform the following steps:

1. Open vCenter Server.
2. Select a VMware ESX Server.
3. Select the **Configuration** tab (see Figure 4-5 on page 38).

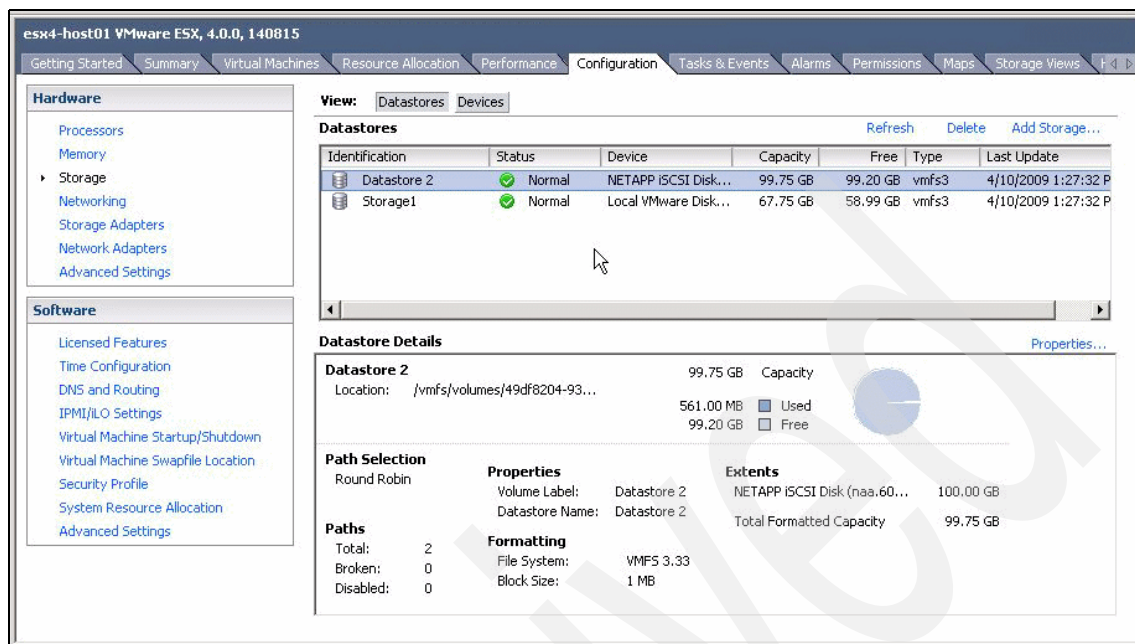


Figure 4-5 Selecting a datastore and displaying its details

4. In the Hardware box, select Storage.
5. In the Storage box, select the storage and then click the **Properties** link.

6. In the Properties dialog box, click **Manage Paths**.
7. For the path selection, set the multipathing policy to Round Robin (see Figure 4-6).

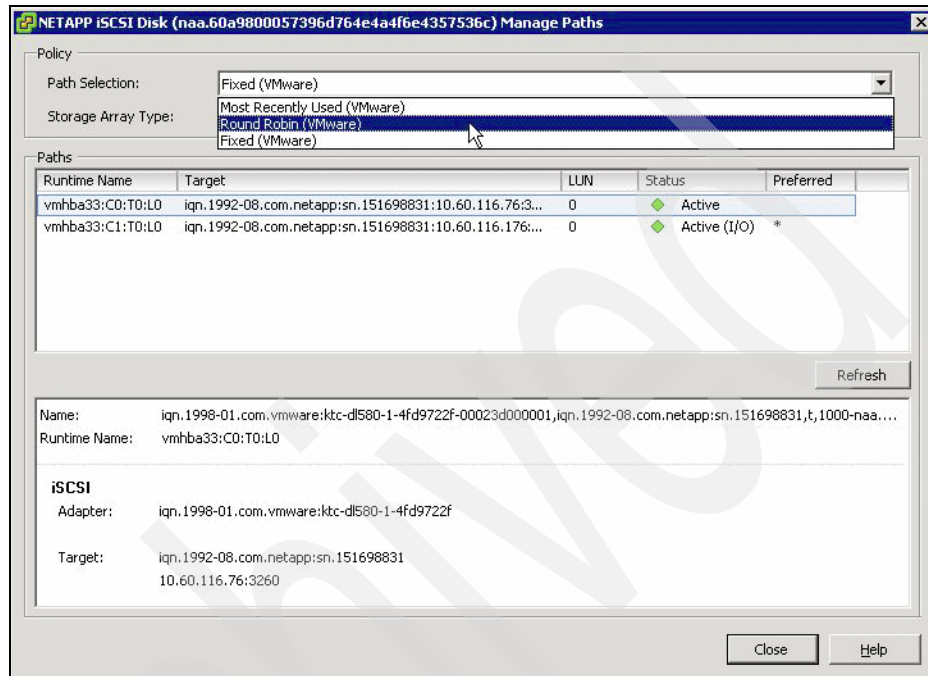


Figure 4-6 Setting the PSP for an iSCSI target

Archived

# NFS storage best practices

This chapter includes practices for integrating NFS data storage with VMware vSphere.

## 5.1 Increasing the number of NFS datastores

By default, VMware ESX is configured with eight NFS datastores; however, this limit can be increased to 64 in order to meet the needs as the virtual infrastructure grows. Although the maximum number of NFS datastores (64) is less than what is available with VMFS datastores (256), this difference is offset by the density available to N series NFS datastores.

To be sure of availability, increase the maximum number of datastores available when deploying a VMware ESX Server because preconfiguring this setting ensures that NFS datastores can be dynamically added at any time without disruption or effort.

To make this change, perform the following steps from within the virtual infrastructure client:

1. Open vCenter Server.
2. Select a VMware ESX host.
3. Select the **Configuration** tab.
4. In the Software box, select **Advanced Configuration**.
5. In the dialog box (left pane), select **NFS**.
6. Change the value of NFS.MaxVolumes to 64 (see Figure 5-1 on page 43).
7. In the dialog box (left pane), select **Net**.
8. Change the value of Net.TcplpHeapSize to 30.
9. Change the value of Net.TcplpHeapMax to 120.

Repeat the steps for each VMware ESX Server.



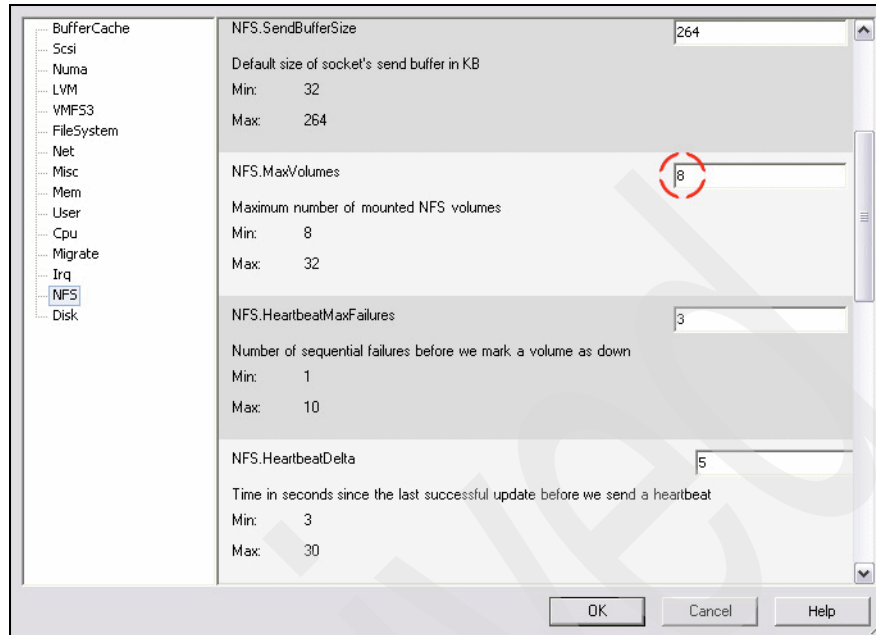


Figure 5-1 Increasing the maximum number of NFS datastores

## 5.2 File system security

N series storage systems permits you to set the security style of each Flexible Volume (or file system) to use UNIX® permissions or NTFS permissions. File system security can be mixed and matched with share or export security. As an example, a UNIX share (or export) can allow access to a file system with NTFS permissions and vice versa. In addition, security style can also be made on a file-by-file basis using the MIXED permissions setting

For VMware ESX deployments, be sure to set the security style of all datastores to UNIX. The security setting of the root volume will be the security setting when a new volume is created.

Typically, customers who run VMware ESX on NFS want to access their datastores from Windows systems, to be able to complete administrative functions. Considering this situation, set the volume security style to UNIX and make sure that the N series user mapping is set up correctly to enable Windows users access to this data.

If you must change the file system security type, perform the following steps:

1. Log in to the N series console.
2. From the storage appliance console, run the following command:  
`vol options <vol-name> no_atime_update on`
3. From the storage appliance console, run the following command  
`qtree security <volume path> UNIX`

Repeat steps 2 and 3 for each NFS accessed volume.

## 5.3 VMware ESX NFS timeout settings

When connecting to NFS datastores we recommend adjusting a few NFS options around connection monitoring and resiliency. These settings can be automatically set for you if you decide to install the N series VMware ESX Host Utilities. The VMware ESX Host Utilities are supported only with VMware ESX. Therefore, if you are running VMware ESXi do not install the VMware ESX Host Utilities but update these settings by performing the following steps:

**Note:** For optimal availability with NFS datastores, make the following changes on each VMware ESX 4.0 host.

1. Open vCenter Server.
2. Select a VMware ESX host.
3. Select the **Configuration** tab.
4. In the Software box, select **Advanced Configuration**.
5. In the dialog box (left pane) select **NFS**.
6. Change the value of NFS.HeartbeatFrequency to 12.
7. Change the value of NFS.HeartbeatMaxFailures to 10.

Repeat these steps for each VMware ESX Server.

## 5.4 NFS storage network best practice

As a best practice, be sure to separate IP-based storage traffic from public IP network traffic by implementing separate physical network segments or VLAN segments. This design follows the architecture of SCSI and FC connectivity.

Creating a second network in VMware ESX requires the creation of a second vSwitch to separate the traffic on to other physical NICs. The VMware ESX Server requires a VMkernel port to be defined on the new vSwitch.

Each VMware ESX Server must have a service console port defined on the vSwitch that transmits public virtual machine traffic and on the vSwitch that is configured for IP storage traffic. This second service console port adds the redundancy in VMware ESX HA architectures and follows VMware ESX HA best practices.

With this design, do not allow routing of data between these networks (do not define a default gateway for the NFS storage network). With this model, NFS deployments require a second service console port be defined on the VMkernel storage virtual switch within each VMware ESX server.

IP storage network, or VMkernel, connectivity can be verified by the use of the **vmkping** command. With NFS-connected datastores, the following syntax tests connectivity:

```
vmkping <NFS IP address>
```

## 5.5 Connecting NFS datastores

To create a file system for use as an NFS datastore, perform the following steps:

1. Open FilerView ([http://filer/na\\_admin](http://filer/na_admin)).
2. Select **Volumes**.
3. Select **Add** to open the Volume Wizard (see Figure 5-2 on page 46). Complete the wizard.

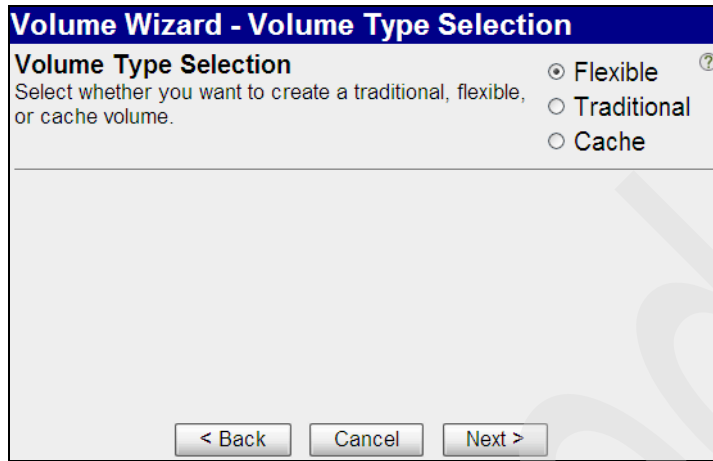


Figure 5-2 Volume Wizard

4. From the FilerView menu, select **NFS**.
5. Select **Add Export** to open the NFS Export Wizard (see Figure 5-3). Complete the wizard for the newly created file system, granting read/write and root access to the VMkernel address of all VMware ESX hosts that will connect to the exported file system.



Figure 5-3 NFS Export Wizard

6. Open vCenter Server.
7. Select a VMware ESX host.
8. Select the **Configuration** tab.

9. In the Hardware box, select the **Storage** link.
10. In the upper-right corner, click **Add Storage** to open the Add Storage wizard (see Figure 5-4).

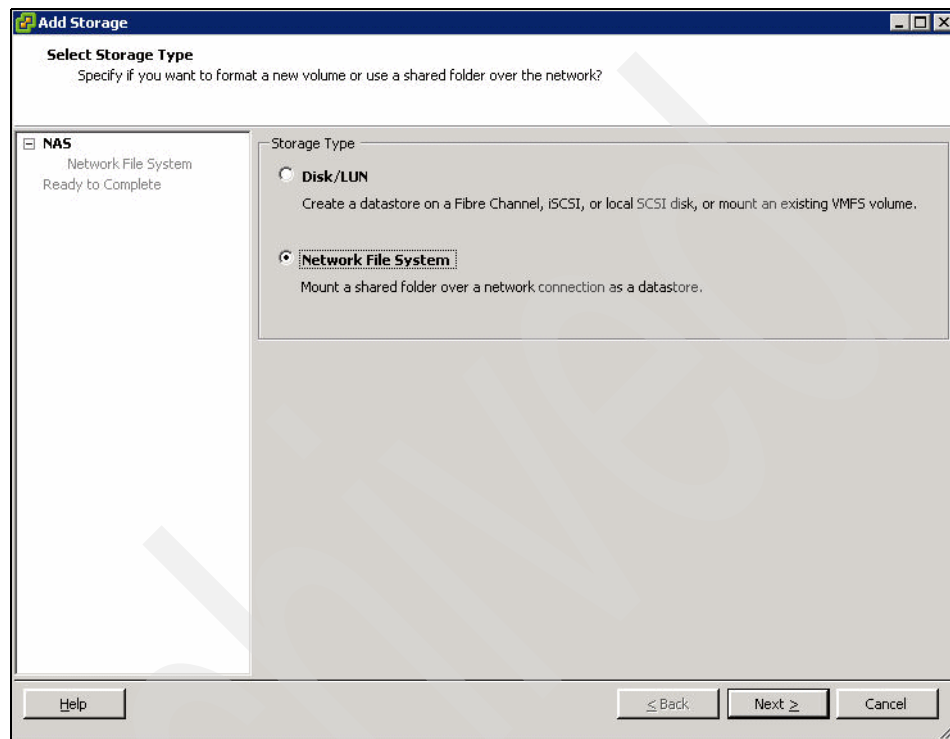


Figure 5-4 VMware Add Storage Wizard

11. Select **Network File System** and click **Next**.

12. Enter a name for the storage appliance, export, and datastore, then click **Next** (see Figure 5-5).

**Add Storage**

**Locate Network File System**  
Which shared folder will be used as a VMware datastore?

**NAS**  
Network File System  
Ready to Complete

**Properties**

Server: 9.11.218.143  
Examples: nas, nas.it.com, 192.168.0.1 or FE80:0:0:0:2AA:FF:FE9A:4CA2

Folder: /vol/vol2/nfs  
Example: /vols/vol0/datastore-001

☐ Mount NFS read only

**Datastore Name**  
Tucson1 NFS

Help < Back Next > Cancel

Figure 5-5 VMware NFS configuration

13. Click **Next** to verify the summary.

14. Click **Finish**.

15. Verify the NFS datastore attachment using the vSphere client (see Figure 5-6).

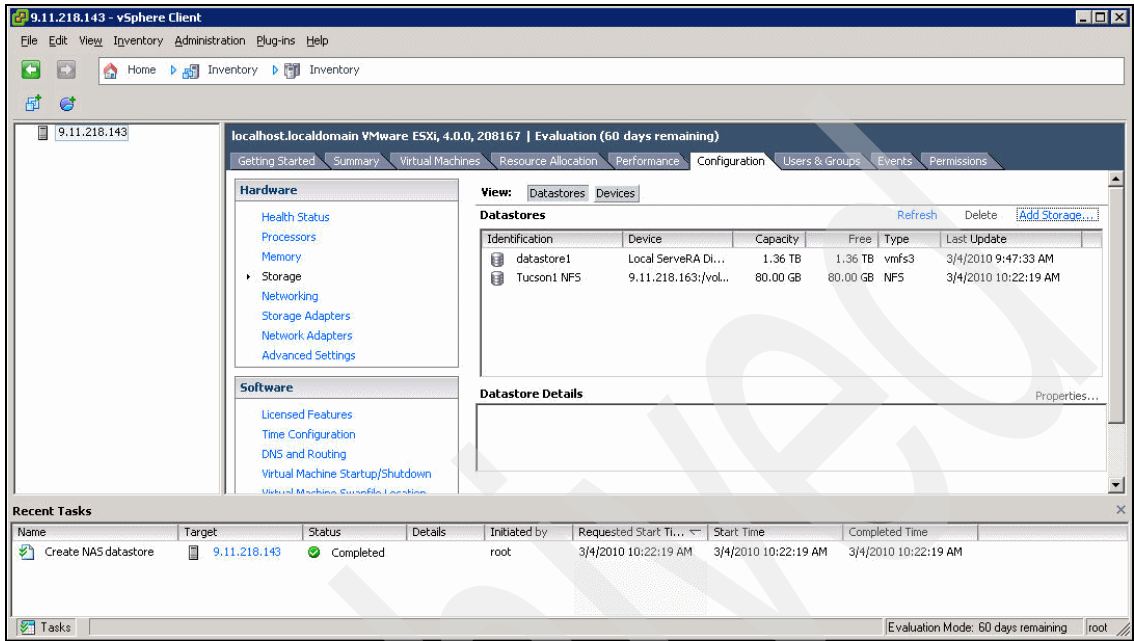


Figure 5-6 NFS storage attachment





## The N series VMware ESX Host Utilities

IBM provides the ESX Host Utilities (EHU) for simplifying the management of VMware ESX nodes running on N series storage systems. With the EHU, you receive a collection of tools to automate and simplify the configuration of HBAs. EHU sets optimized VMware ESX/VMware ESXi options for NFS, enhances Guest OS (VM) SCSI settings, and provides diagnostic utilities if a support case is opened.

The VMware ESX Host Utilities installs on VMware ESX 4.0 systems, and currently is not supported for VMware ESXi 4.0 systems.

After the N series is registered, the VMware ESX Host Utilities can be downloaded by using the IBM support processes.

## 6.1 Installation

This section lists the prerequisites and outlines the installation for the EHU.

### 6.1.1 Prerequisites

Before installing the EHU, be sure the following requirements are met:

- ▶ FC (which includes FCoE), iSCSI, or NFS is licensed on the storage system.
- ▶ You have root access to each VMware ESX Server.
- ▶ All storage systems have names that can be resolved by DNS.
- ▶ Secure Sockets Layer (SSL) is set up on every storage controller before installing EHU (if you plan to use SSL to securely communicate with the controller).
- ▶ If you do not want to enter a user name or password when running the EHU, be sure to enable option `httpd.admin.hostsequiv` on each N series (for example, option `httpd.admin.hostsequiv.enable on`) and that all VMware ESX host names are added to the `/etc/hosts.equiv` file on each N series. This step helps prevent connection problems between the N series and the hosts.

### 6.1.2 Installation

To install the VMware ESX Host Utilities, perform the following steps:

1. Download the EHU.
2. Copy the EHU to a location accessible to the VMware ESX server.
3. Extract the EHU by running the following command:  

```
tar -zxvf <name of EHU file>.tar.gz
```
4. Migrate running VMs to other VMware ESX nodes.
5. Place the VMware ESX Server in maintenance mode.
6. Run the following command:  

```
./install
```
7. Complete the EHU installation wizard.
8. Reboot the VMware ESX Server and return to normal operations.

### 6.1.3 EHU assisted multipathing

One of the components of the EHU is a script named `config_mpath`. This script reduces the administrative overhead of managing SAN LUN paths. The `config_mpath` script determines the preferred primary paths to each of the SAN LUNs on the VMware ESX Server and then sets the preferred path for each LUN to use one of the primary paths. Running the `config_mpath` script once on each VMware ESX Server in the cluster can complete multipathing configuration for large numbers of LUNs quickly and easily. If changes are made to the storage configuration, run the script again to update the multipathing configuration, based on the changes to the environment.

## 6.2 Manual configuration of FC HBAs in VMware ESX

In previous versions of VMware ESX, manually configuring the settings on the HBAs was required. This requirement is eliminated when VMware ESX 4.0 and VMware ESXi 4.0 servers are connected to N series storage systems (running Data ONTAP 7.2.4 or later). If your storage arrays are not running a release of Data ONTAP version 7.2.4 or later, consider upgrading them. Alternatively you must install the VMware ESX Host Utilities to adjust values of the HBAs in your VMware ESX servers.



## FC and FCoE storage networking best practices

Fibre Channel (FC) storage networks make up the largest percentage of shared storage infrastructures that host VMware ESX. This market share is attributed to FC being the first networked-attached storage protocol that is supported by VMware ESX in version 2.0. Although FC is a well-known and mature technology, this chapter covers best practices for deploying VMware ESX on Fibre Channel with N series storage systems.

## 7.1 Host bus and converged network adapters

VMware ESX servers and N series storage systems connect to a SAN fabric using host bus adapters (HBAs). Connectivity to FCoE fabrics is enabled through converged network adapters (CNAs). Each HBA/CNA can run as either an initiator (VMware ESX) or as a target (N series). Each adapter has a global unique address referred to as a *worldwide port name (WWPN)*. Each WWPN is required to be known in order to configure LUN access on a N series.

N series and VMware best practices indicate that each VMware ESX server have at least two adapter ports. For more information about VMware FC best practices, see *VMware Fibre Channel SAN Configuration Guide*:

[http://vmware.com/pdf/vsphere4/r40/vsp\\_40\\_san\\_cfg.pdf](http://vmware.com/pdf/vsphere4/r40/vsp_40_san_cfg.pdf)

## 7.2 N series igroups (LUN masking)

Logical unit number (LUN) masking is an authorization process that makes a LUN available to a host or set of hosts in a cluster. On an N series storage system, LUN masking is implemented by assigning HBA addresses to initiator groups (igroups). After an igroup has been defined, LUNs can be assigned the igroup for access to the LUN.

Implementation best practices for LUN masking is covered in Chapter 3, “VMware ESX FC, FCOE, and iSCSI storage configuration” on page 17.

## 7.3 FC and FCoE zoning

Many devices and nodes can be attached to a SAN; a way to secure access to these devices is by implementing zones. SAN zoning is a method of arranging Fibre Channel devices into logical groups over the physical configuration of the fabric or Fibre Channel network.

Zoning is available in hardware (hard zoning) or in software (soft zoning). An option available with both implementations is port zoning, where physical ports define security zones. A host's access to a LUN is determined what physical port connects it to. With port zoning, zone information must be updated every time a user changes switch ports. In addition, port zoning does not allow zones to overlap.

Another form of zoning is WWN zoning, where the fabric uses its name servers to either allow or block access to particular worldwide names (WWNs) in the fabric. A major advantage of WWN zoning is the ability to recable the fabric without having to redo the zone information.

A good practice is to implement single initiator multiple storage target zones. This design offers an ideal balance of simplicity and availability with FC and FCoE deployments.

Archived





## Ethernet storage networking best practices

Use dedicated resources for storage traffic when possible. With Ethernet storage networks, this task can be achieved with separate physical switches, or logically by implementing VLAN segments for storage I/O on a shared, switched IP infrastructure.

One of the challenges of configuring VMware ESX networking for IP storage is that the network configuration meets all of the following goals:

- ▶ Be redundant across switches in a multiswitch environment.
- ▶ Use as many available physical paths as possible.
- ▶ Be scalable across multiple physical interfaces.

## 8.1 The 10 Gigabit Ethernet (10 GbE) standard

VMware ESX 4 and ESXi 4 include support for 10 GbE standard. An advantage of 10 GbE is the ability to reduce the number of network ports in the infrastructure, especially but not limited to, blade servers. To verify support for your hardware and its use for storage I/O, see the VMware Compatibility Guide:

<http://www.vmware.com/resources/compatibility/search.php?action=base&deviceCategory=io>

## 8.2 Virtual LANs (VLANs)

When segmenting network traffic with VLANs, interfaces either can be dedicated to a single VLAN or can support multiple VLANs with VLAN tagging.

For systems that have fewer NICs, such as blade servers, VLANs can be very useful. Channeling two NICs together provides a VMware ESX server with physical link redundancy. By adding multiple VLANs, you can group common IP traffic on separate VLANs for optimal performance, as follows, for example:

- ▶ Group the Service console access with the virtual machine network on one VLAN.
- ▶ The VMkernel activities of IP Storage and vMotion can reside on a second VLAN.

VLANs and VLAN tagging also play a simple but important role in securing an IP storage network. NFS exports can be restricted to a range of IP addresses that are available only on the IP storage VLAN. N series storage systems also allow the restriction of the iSCSI protocol to specific interfaces, VLAN tags, or both.

These simple configuration settings have an enormous effect on the security and availability of IP-based datastores. If you are using multiple VLANs over the same interface, be sure that sufficient throughput can be provided for all traffic.

## 8.3 Flow control

Flow control is the process of managing the rate of data transmission between two nodes to prevent a fast sender from over running a slow receiver. Flow control can be configured for VMware ESX servers, N series storage systems, and network switches. Be sure to configure the end points: VMware ESX servers and N series storage systems with flow control set to *Send ON* and *Receive OFF*.

For network switches, set the switch ports connecting to VMware ESX hosts and N series storage systems to `Desired`, or if this mode is not available, set these ports to `Send OFF` and `Receive ON`. See Figure 8-1

**Note:** the switch ports are configured with the opposite settings of the end points, the VMware ESX and N series systems.

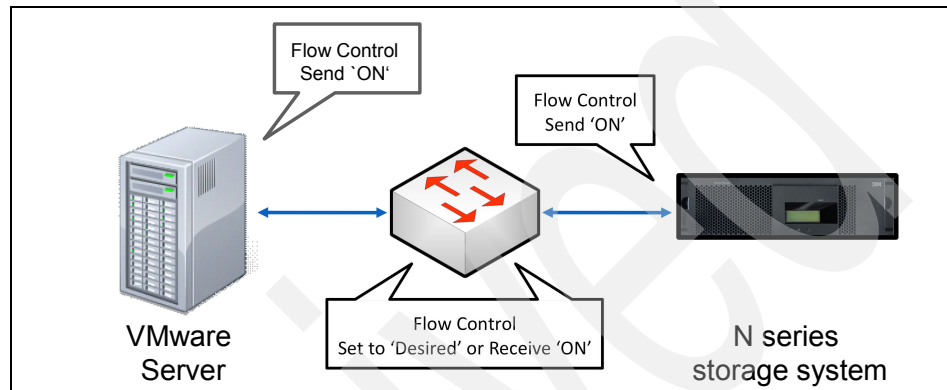


Figure 8-1 Configuring flow control setting

## 8.4 Spanning Tree Protocol (STP)

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

When connecting VMware ESX and N series storage systems to Ethernet storage networks, be sure that the Ethernet ports, which these systems connect to, are configured with either Rapid Spanning Tree Protocol (RSTP) or are portfast-enabled.

## 8.5 Bridge Protocol Data Unit (BPDU)

BDUs exchange information about bridge IDs and root path costs within STP. When connecting VMware ESX and N series storage systems to Ethernet storage networks, be sure that the Ethernet ports, which these systems connect to, are configured with BPDUs disabled.

## 8.6 Virtual Interfaces

A Virtual Network Interface (VIF) is a mechanism that supports aggregation of network interfaces into one logical interface unit. After the VIF is created, it is indistinguishable from a physical network interface. VIFs are used to provide fault tolerance of the network connection and, in certain cases, higher throughput to the storage device.

The N series enables the use of two types of load-balancing VIFs:

- ▶ **Multimode VIF:** This type is a static-configured Ethernet trunk. In a multimode VIF, all of the physical connections in the VIF are simultaneously active and can carry traffic. This mode requires that all of the interfaces be connected to a switch that supports trunking or aggregation over multiple port connections. The switch must be configured to understand that all the port connections share a common MAC address and are part of a single logical interface. In the event of a physical interface failure resulting in the loss of link the VIF will automatically transmit traffic on the surviving links in the VIF without loss of connectivity.
- ▶ **Dynamic multimode VIF:** This type is an LACP-compliant (IEEE 802.3ad) VIF. In a dynamic multimode VIF, all of the physical connections are simultaneously active and carry traffic as with multimode VIFs, described previously. Dynamic multimode VIFs introduce the use of LACP signaling transmissions between the N series storage system and the remote switch. This signaling informs the remote channeling partner of link status. If a failure or inability to transmit data on a link is observed, the device identifying this problem informs the remote channeling partner of the failure, causing the removal of the interface from the VIF. This feature differs from standard multimode VIFs in that there is no signaling between channel partners to inform the remote partner of link failure. The only means for an interface to be removed from a standard multimode VIF is loss of link.

Multimode and dynamic multimode VIFs each use the same algorithm for determining load-balancing. This algorithm is based on source and destination IP or MAC address. Use IP-based source and destination load-balancing especially when the network is designed to route storage traffic. The reason is, because during a transmission of a routed packet, a host transmits the packet to the default router IP address. When arriving at the router, the router changes the MAC address of the routed packet to the MAC address of the local router interface on which the packet is transmitted out. The changing of the source MAC address can produce situations where traffic arriving from other subnets is always load-balanced to the same physical interfaces in the VIF. IP addresses are not changed unless Network Address Translation (NAT) is used. NAT is rarely used within the data center, where communications between VMware ESX hosts and N series systems occur.

In a single-mode VIF, only one of the physical connections is active at a time. If the storage controller detects a fault in the active connection, a standby connection is activated. No configuration is necessary on the switch to use a single-mode VIF, and the physical interfaces that make up the VIF do not have to connect to the same switch. Note that IP load balancing is not supported on single-mode VIFs.

Another possibility is to create second-level single or multimode VIFs. By using second-level VIFs, you can take advantage of both the link aggregation features of a multimode VIF and the failover capability of a single-mode VIF. In this configuration, two multimode VIFs are created, each one to a separate switch. A single-mode VIF is then created composed of the two multimode VIFs. In normal operation, traffic flows over only one of the multimode VIFs; but if an interface or switch fails, the storage controller moves the network traffic to the other multimode VIF.

## 8.7 Ethernet switch connectivity

An IP storage infrastructure provides the flexibility to connect to storage in separate ways, depending on the needs of the environment. A basic architecture can provide a single non-redundant link to a datastore, suitable for storing ISO images, various backups, or VM templates. A redundant architecture, suitable for most production environments, has multiple links, providing failover for switches and network interfaces. Link-aggregated and load-balanced environments make use of multiple switches and interfaces simultaneously to provide failover and additional overall throughput for the environment.

More modern Ethernet switch models support *cross-stack EtherChannel* or *virtual port channel* trunks, where interfaces on separate physical switches are combined into an IEEE 802.3ad EtherChannel trunk. The advantage of multiswitch EtherChannel trunks is that they can eliminate the need for additional passive links that are accessed only during failure scenarios in certain configurations.

All IP storage networking configuration options covered here use multiple switches and interfaces to provide redundancy and throughput for production VMware environments.



# Configuring Ethernet storage networks

This chapter contains information about applying best practices for VMware Ethernet storage networks.

## 9.1 Highly available storage designs with traditional Ethernet switches

This section provides details about storage solutions designs that use traditional Ethernet switches. The section also introduces multiple VMkernel ports.

To simultaneously use multiple paths while providing high availability with traditional Ethernet switches, each VMware ESX server must be configured with a minimum of two VMkernel ports in the same vSwitch.

Depending on storage protocol, this vSwitch may be configured with multiple network adapters. For iSCSI datastores, each VMkernel is configured with a single vmnic. No standby vmnics (physical network adapters on the ESX server) may exist in the VMkernel.

For NFS datastores each VMkernel is configured with a single active vmnic, with one or more standby vmnics defined.

A good practice is to define a separate VMkernel for each storage protocol. Doing so simplifies the configuration of iSCSI with NFS. As an example, see Figure 9-1 on page 67. Each of these VMkernel ports supports IP traffic on a separate subnet. Because the two VMkernel ports are in the same vSwitch, they can share the physical network adapters in that vSwitch.



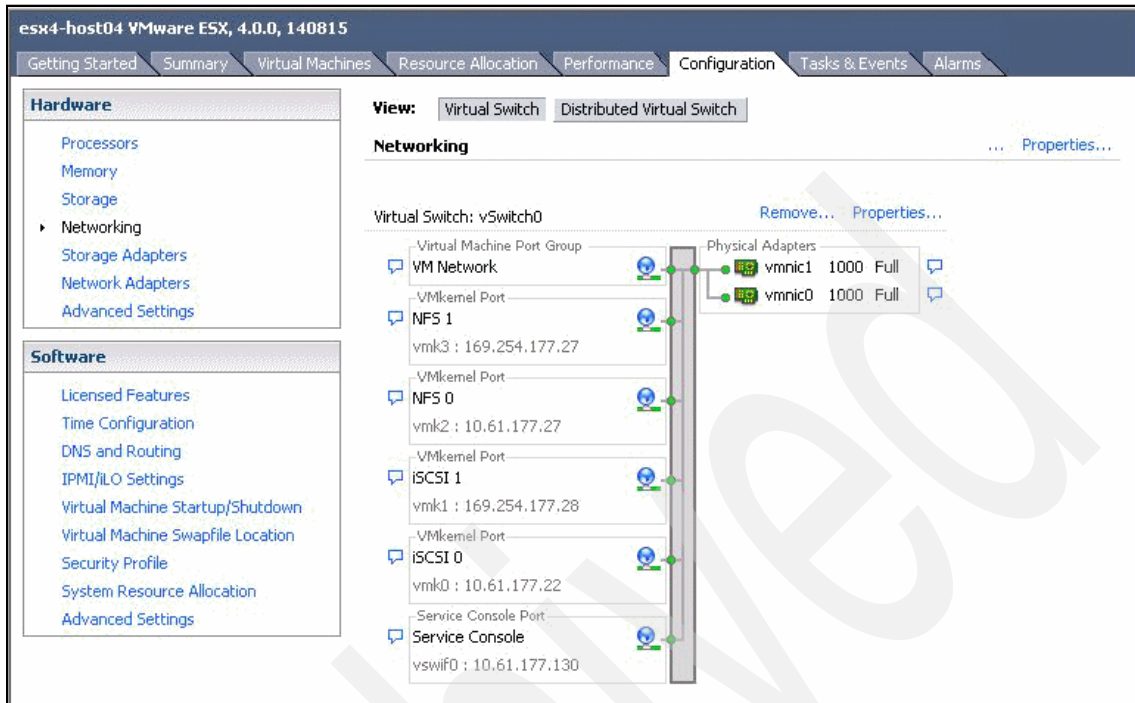


Figure 9-1 Displaying VMkernel ports for iSCSI

## 9.2 VMware ESX server adapter failover behavior with iSCSI

In case of VMware ESX server adapter failure (because of a cable pull or NIC failure), traffic that was originally running over the failed adapter is rerouted and continues to use the second adapter. This failover is managed by VMware's native multipathing, therefore network failover configuration on the switch or VMkernel is not necessary. Traffic returns to the original adapter when service to the adapter is restored.

### 9.2.1 VMware ESX server adapter failover behavior with NFS

In case of VMware ESX server adapter failure (because of a cable pull or NIC failure), traffic that was originally running over the failed adapter is rerouted and continues using the second adapter, but on the same subnet where it originated. Both subnets are now active on the surviving physical adapter. Traffic returns to

the original adapter when service to the adapter is restored. In this scenario EtherChannel provides the network failover.

## 9.2.2 Reviewing link aggregation within VMware ESX server

VMware ESX server supports static link aggregation. Link aggregation provides the means to channel multiple network ports, and the channeling of ports provides a means to distribute traffic based on source and destination and to increase link redundancy for higher availability.

In this document, any reference to *EtherChannel* in the terms of configuring a VMware ESX server is actually referring to a static EtherChannel. VMware ESX server does not support the use of Link Aggregation Control Protocol (LACP) 802.3ad.

## 9.2.3 Switch failure

Traffic originally running to the failed switch is rerouted and continues using the other available adapter, through the surviving switch, to the N series storage controller. Traffic returns to the original adapter when the failed switch is repaired or replaced.

## 9.2.4 Connecting to datastores

With Ethernet based storage networking protocols, VMware datastores are mounted by IP addresses. Both iSCSI and NFS access datastores by a single IP address.

With both iSCSI and NFS datastores, multiple datastores are required to make use of multiple IP paths simultaneously on each VMware ESX/ESXi host. Avoid using NFS to connect to the same volume multiple times from a single VMware ESX/ESXi host because VMware ESX/ESXi and vCenter consider these connections to be separate datastores.

With iSCSI, this design is represented as multiple IP paths to a single SCSI target, only one IP path is active per datastore, however each VMware ESX/ESXi host may use a separate active IP path. This behavior can be changed to send I/O traffic over multiple paths by enabling the Round Robin PSP (described in 3.5, “Connecting iSCSI datastores” on page 23). This design results in the aggregation of multiple links and providing fault tolerance for a link, switch, or NIC.

With NFS datastores, each datastore must be connected only once from each VMware ESX/ESXi server, and must be using the same N series target IP address on each VMware ESX/ESXi server.

Figure 9-2 shows an Ethernet connection in normal operation.

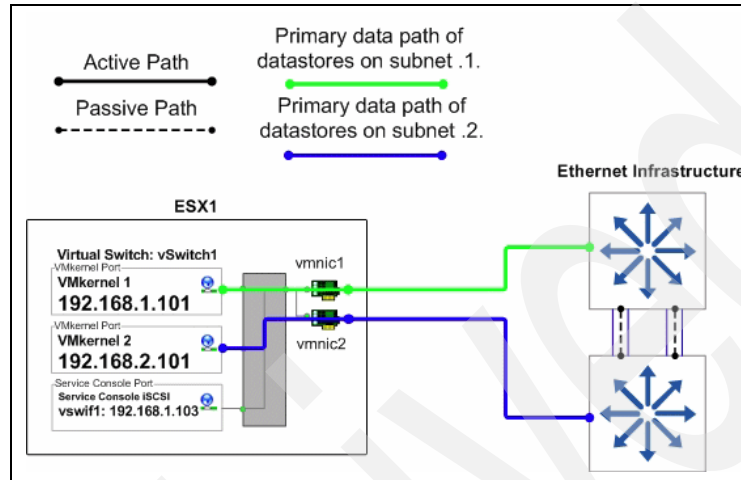


Figure 9-2 VMware ESX vSwitch1 normal mode operation

When a network card fails, the system operates in failover mode as shown in Figure 9-3.

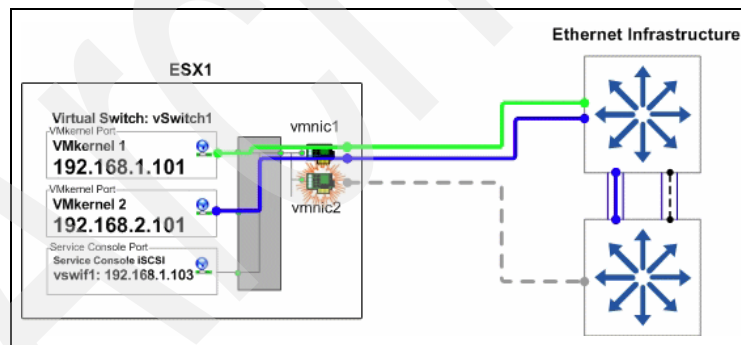


Figure 9-3 VMware ESX vSwitch1 failover mode operation

## 9.2.5 Scalability of VMware ESX server network connections

Although the configuration shown in Figure 9-3 on page 69 uses two network adapters in each VMware ESX Server, it can be scaled up to use additional adapters, with another VMkernel port, subnet, and IP address added for each additional adapter.

Another option is to add a third adapter and configure it as an N+1 failover adapter. By not adding more VMkernel ports or IP addresses, the third adapter can be configured as the first standby port for both VMkernel ports. In this configuration, if one of the primary physical adapters fails, the third adapter assumes the failed adapter's traffic, providing failover capability without reducing the total amount of potential network bandwidth during a failure.

## 9.2.6 Configuring ESX/ESXI VMkernel storage network ports

If the switches used for IP storage networking support multi-switch EtherChannel trunking, or virtual port channeling, each ESX Server requires one physical connection to each switch in the stack with IP load balancing enabled. One VMkernel port with one IP address is required. To use each of the available physical links, multiple datastore connections to the storage controller using separate target IP addresses are necessary.

The benefits of this configuration are as follows:

- ▶ Offers low complexity.
- ▶ Provides two active connections to each storage controller.
- ▶ Easily scales by using more connections.
- ▶ Storage controller connection load balancing is automatically managed by IP load balancing policy.
- ▶ Requires only one VMkernel port for IP storage to make use of multiple physical paths.

In the ESX Server configuration, shown in the Figure 9-4 on page 71, a vSwitch (named vSwitch1) has been created specifically for IP storage connectivity. Two physical adapters have been configured for this vSwitch (in this case vmnic1 and vmnic2). Each of these adapters is connected to a separate physical switch; the switch ports are configured into a cross-stack EtherChannel trunk.

**Note:** At this time, VMware does not support LACP, or IEEE 802.3ad, which is the dynamic negotiation of Ethernet trunks.

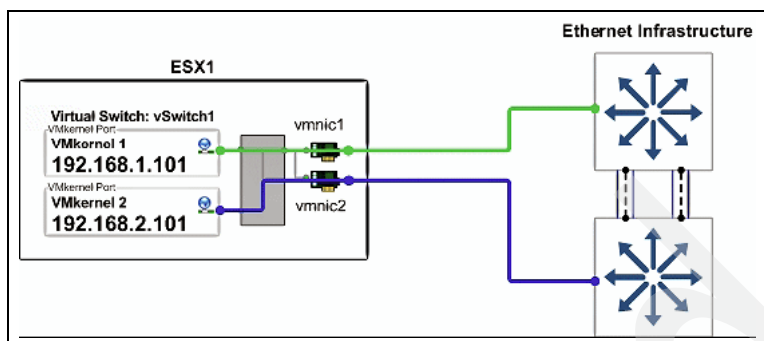


Figure 9-4 VMware ESX Server physical NIC connections with traditional Ethernet

In vSwitch1, one VMkernel port has been created (VMkernel 1) and configured with one IP address, and the NIC Teaming properties of the VMkernel port have been configured as follows (see Figure 9-5 on page 72):

- ▶ VMkernel 1: IP address is set to 192.168.1.101.
- ▶ VMkernel 1 port properties: Load Balancing policy is set to “Route based on ip hash.”

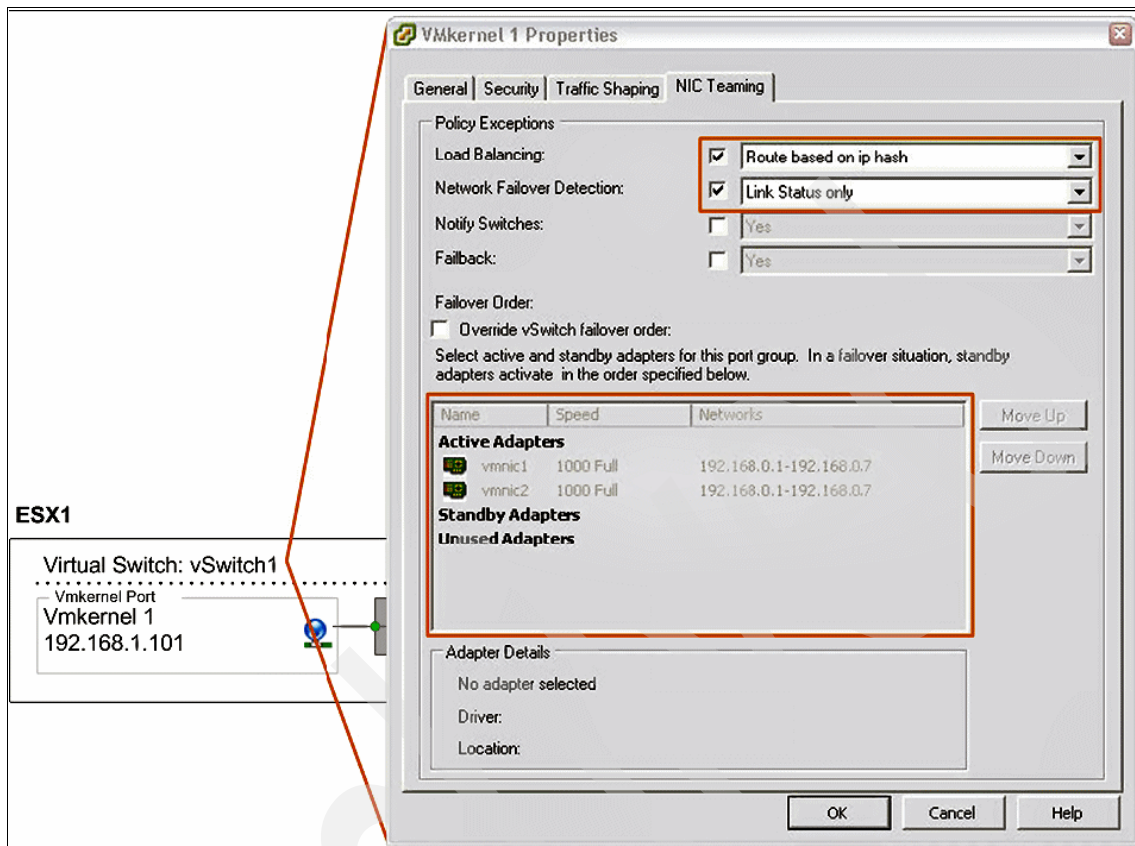


Figure 9-5 VMware ESX Server VMkernel port properties with traditional Ethernet

## 9.3 A storage architecture with traditional Ethernet

In this configuration, the IP switches to be used do not support multiswitch EtherChannel trunking, so each storage controller requires four physical network connections. This design is available in two options (represented in Figure 9-6 on page 73 and Figure 9-7 on page 74). Both designs are very similar. They both provide multiple active links to each storage controller, provide a means to scale throughput by simply adding more links, and require multiple IP addresses per controller. Each design uses two physical links for each active network connection in order to achieve path high availability.

## The multi-multimode design

The multi-mode design (see Figure 9-6) requires each storage controller to have at least four physical network connections (depicted). The connections are divided into two multimode (active-active) VIFs with IP load balancing enabled, one VIF connected to each of the two switches. These two VIFs are then combined into one single mode (active-passive) VIF. We refer to this configuration as a *second-level VIF*. This option also requires multiple IP addresses on the storage appliance. Multiple IP addresses can be assigned to the single-mode VIF by using IP address aliases or by using VLAN tagging.

## Advantages of using multimode VIFs

The advantages of using multimode VIFs are as follows:

- ▶ Data I/O to a single IP is aggregated over multiple links
- ▶ Storage controller connection load balancing is automatically managed by the EtherChannel IP load balancing policy.

## Disadvantages of using multimode VIFs

The disadvantages of using multimode VIFs are as follows:

- ▶ Some switch-side configuration is required.
- ▶ Some storage traffic can cross the uplink between the two switches.

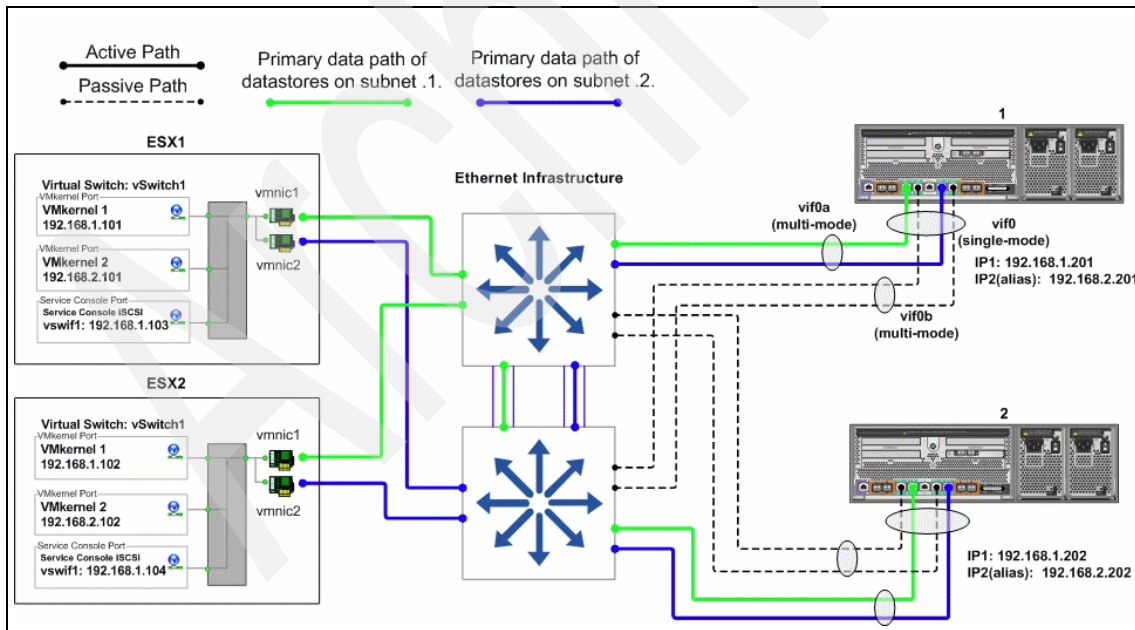


Figure 9-6 Storage-side multimode VIFs

## The single-mode design

The single-mode design (see Figure 9-7) requires each pair of network links to be configured as a single mode (active passive) VIF. Each VIF has a connection to both switches and has a single IP address assigned to it, providing two IP addresses on each controller. The `vif favor` command is used to force each VIF to use the appropriate switch for its active interface. This option is preferable because of its simplicity and the lack of any special configuration in the network switches.

## Advantage of using single mode VIFs

The main advantage of using a single mode VIF is simplicity, because no switch side configuration is required.

## Disadvantage of using single mode VIFs

The main disadvantage of using a single mode VIF is that data I/O to a single IP is not aggregated over multiple links without adding more links.

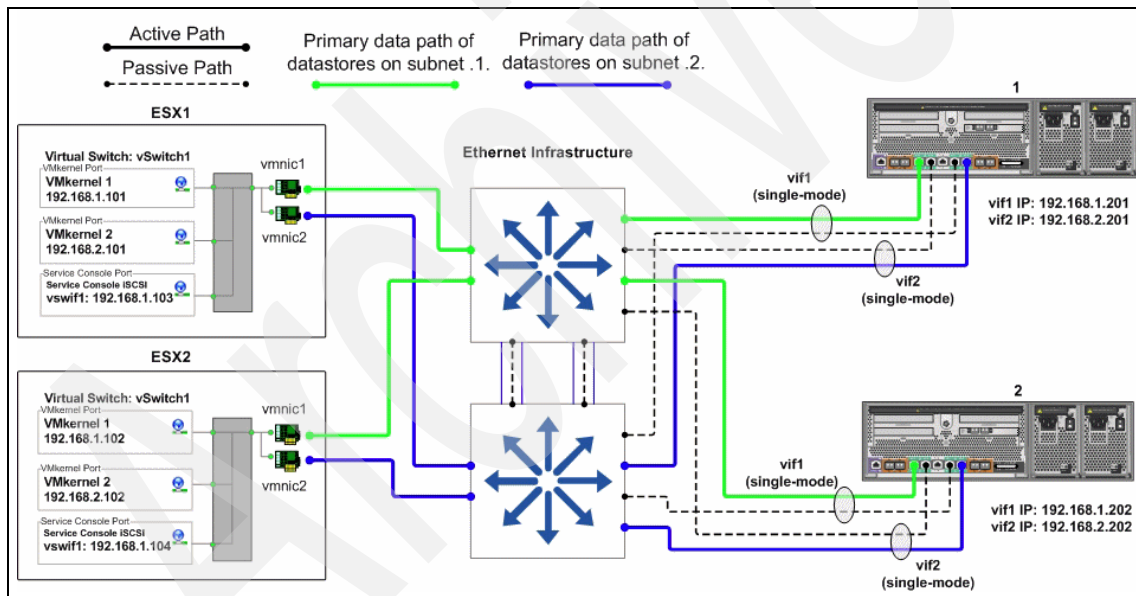


Figure 9-7 Storage side single-mode VIFs



## 9.4 Datastore configuration with traditional Ethernet

In addition to properly configuring the vSwitches, network adapters, and IP addresses, using multiple physical paths simultaneously on an IP storage network requires connecting to multiple datastores, making each connection to a separate IP address.

In addition to configuring the VMware ESX Server interfaces, the N series storage controller has been configured with an IP address on each of the subnets that are used to access datastores. This task is accomplished by the use of multiple teamed adapters, each with its own IP address or, in certain network configurations, by assigning IP address aliases to the teamed adapters, allowing those adapters to communicate on all the required subnets.

When connecting a datastore to the VMware ESX Servers, the administrator configures the connection to use one of the IP addresses assigned to the N series storage controller. When using NFS datastores, this assignment is accomplished by specifying the IP address when mounting the datastore.

Figure 9-8 on page 76 shows the storage traffic flow when using multiple VMware ESX Servers and multiple datastores. With iSCSI, this design is represented as multiple IP paths to a single SCSI target; only one IP path is active per datastore, however each VMware ESX/ESXi host may use a separate active IP path. This behavior can be changed to send I/O traffic over multiple paths by enabling the Round Robin multipathing plug-in (described in 3.5, “Connecting iSCSI datastores” on page 23). Regarding NFS datastores, each datastore must be connected only once from each VMware ESX/ESXi server, and must be using the same N series target IP address on each VMware ESX/ESXi server.

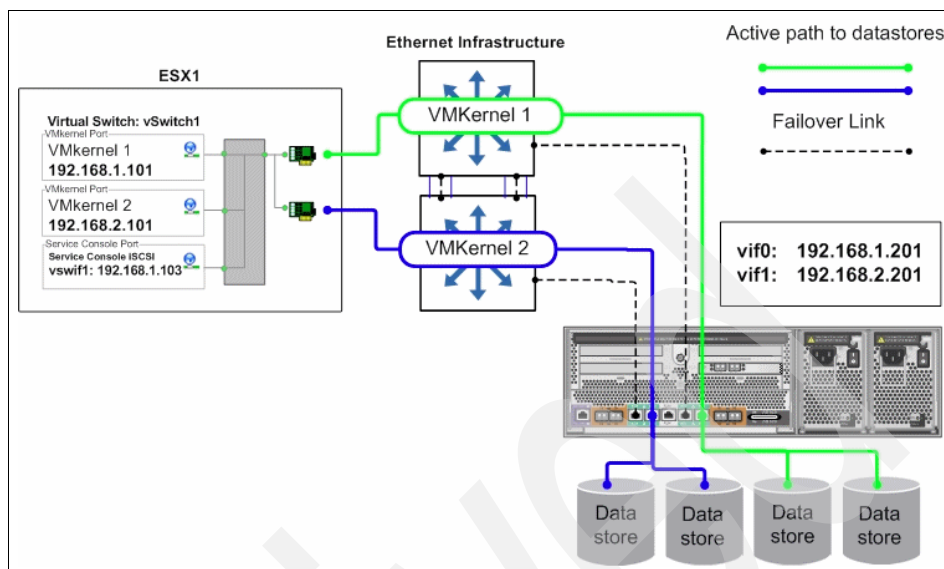


Figure 9-8 Datastore connections with traditional Ethernet

## 9.5 VMkernel configuration with multiswitch trunking

If the switches that are used for IP storage networking support multiswitch EtherChannel trunking or virtual port channeling, each VMware ESX Server requires one physical connection to each switch in the stack, with IP load-balancing enabled. One VMkernel port with one IP address is required. Multiple datastore connections to the storage controller using separate target IP addresses are necessary to use each of the available physical links.

### Advantages

The advantages of this configuration are as follows:

- ▶ Offers low complexity.
- ▶ Provides two active connections to each storage controller.
- ▶ Easily scales using more connections.
- ▶ Storage controller connection load balancing is automatically managed by IP load-balancing policy.
- ▶ Requires only one VMkernel port for IP storage to make use of multiple physical paths.

## Disadvantages

A disadvantage of this configuration is that it requires multiswitch EtherChannel capability such as stackable switches or virtual port channeling.

In the VMware ESX Server configuration, shown in the Figure 9-9, a vSwitch (named vSwitch1) has been created specifically for IP storage connectivity. Two physical adapters have been configured for this vSwitch (in this case vmnic1 and vmnic2). Each of these adapters is connected to a separate physical switch and the switch ports are configured into a cross-stack EtherChannel trunk. Note at this time, VMware ESX does not support LACP, or IEEE 802.3ad, which is the dynamic negotiation of Ethernet trunks.

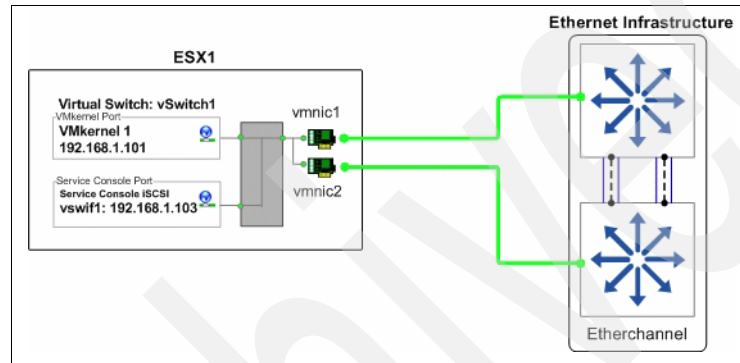


Figure 9-9 VMware ESX Server physical NIC connections with multiswitch EtherChannel

In vSwitch1, one VMkernel port has been created (VMkernel 1) and configured with one IP address, and the NIC teaming properties of the VMkernel port have been configured as follows (see Figure 9-10 on page 78):

- ▶ VMkernel 1, IP address is set to:  
192.168.1.101
- ▶ VMkernel 1 Port Properties, load-balancing policy is set to:  
Route based on ip hash

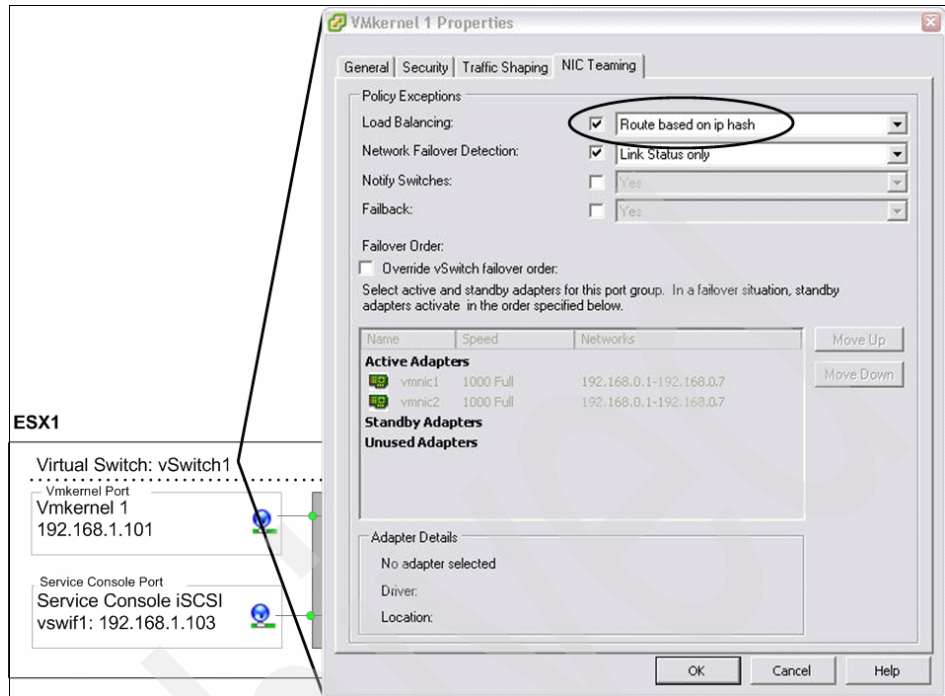


Figure 9-10 VMware ESX Server VMkernel port properties with multiswitch EtherChannel

## 9.6 Storage network architecture with multiswitch link aggregation

In this configuration (Figure 9-11 on page 79), the IP switches to be used for the Ethernet storage network support multiswitch link aggregation. As such, each storage controller requires one physical connection to each switch; the two ports that are connected to each storage controller are then combined into one multimode LACP VIF with IP load balancing enabled. This design provides multiple active links to each storage controller, provides a means to scale throughput by simply adding more links, and requires multiple IP addresses per controller. Each uses two physical links for each active network connection so that path high availability can be achieved.

The advantages of MSLA are as follows:

- Provides multiple active connections to each storage controller.
- Easily scales to more connections by adding NICs and aliases.
- Storage controller connection load balancing is automatically managed by the EtherChannel
- Offers an IP load balancing policy.

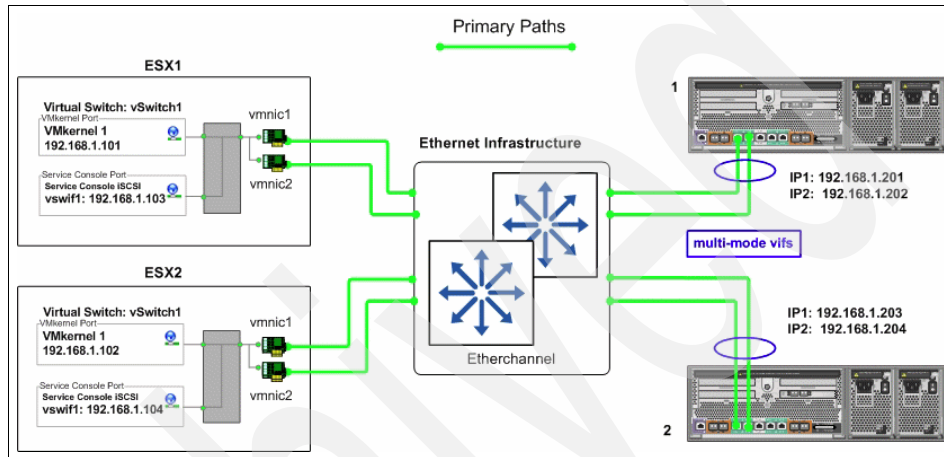


Figure 9-11 Storage side multimode VIFs using multiswitch EtherChannel

## Storage Load Balancing

Using multiple physical paths simultaneously on an IP storage network requires EtherChannel ports and multiple IP addresses on the storage controller. This model results in a design that balances datastore connectivity across all interfaces.

Figure 9-12 shows an overview of storage traffic flow when using multiple ESX Servers and multiple datastores.

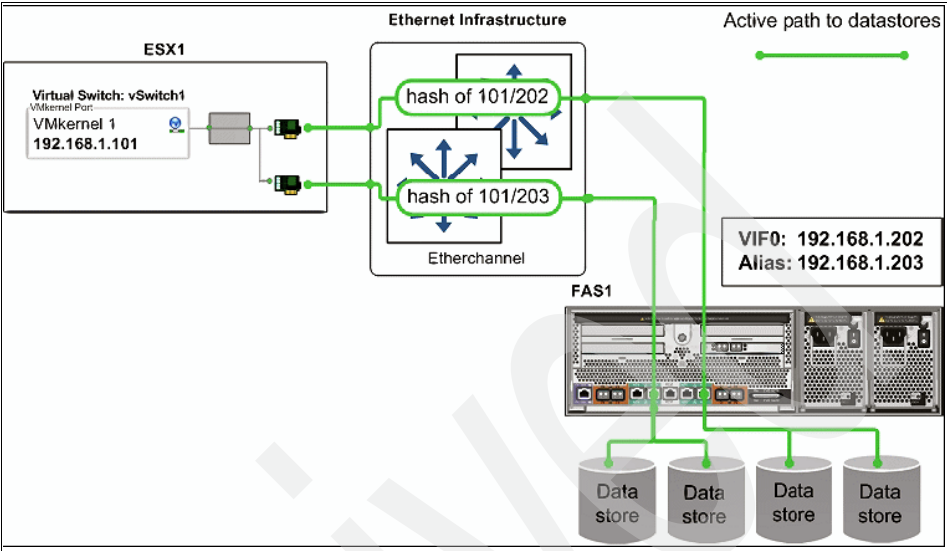


Figure 9-12 Datastore connections with multi-switch EtherChannel.



## Increasing storage utilization

This chapter provides information about increasing storage utilization.

## 10.1 Introduction

VMware vSphere provides an excellent means to increase the hardware utilization of physical servers. By increasing hardware utilization, the amount of hardware in a data center can be reduced, lowering the cost of data center operations. In a typical VMware Vsphere environment, the process of migrating physical servers to virtual machines does not reduce the amount of data stored or the amount of storage provisioned. By default, server virtualization does not affect the improving of storage utilization (and in many cases may have the opposite effect).

By default in VMware ESX 3.5, virtual disks preallocate the storage they require; in the background, they zero out all of the storage blocks. This type of Virtual Machine Disk (VMDK) format is called a *zeroed thick VMDK*. VMware ESX provides a means to consume less storage by provisioning VMs with thin-provisioned virtual disks. With this feature, storage is consumed on demand by the VM. VMDKs, which are created on NFS datastores, are in the thin format by default.

With VMware ESX 4.0, thin-provisioned VMDKs are now available to be created in the virtual infrastructure client with VMFS datastores. By using VMware vSphere thin-provisioning technology, the amount of storage consumed on a VMFS datastore can be reduced. VMDKs that are created as thin-provisioned disks can be converted to traditional zero-thick format; however, you cannot convert an existing zero-thick format to the thin-provisioned format.

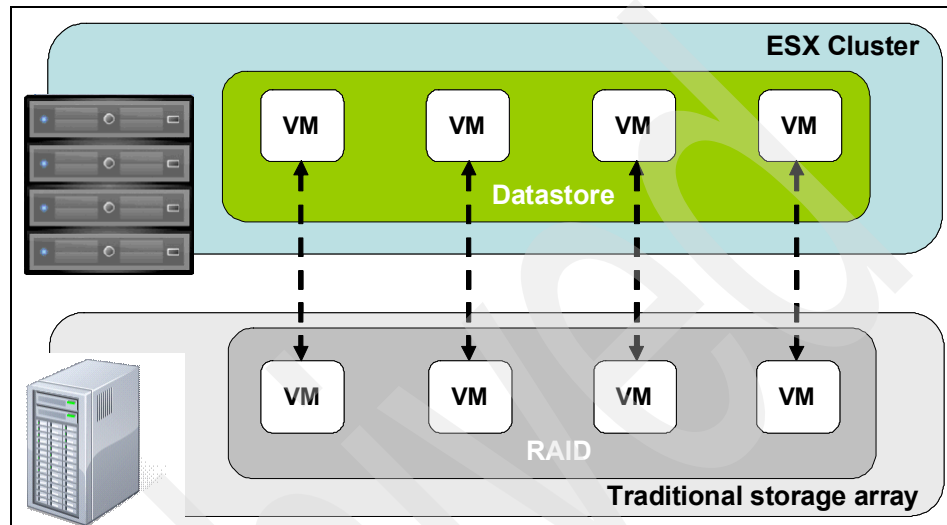
The IBM System Storage N series offers storage virtualization technologies that can enhance the storage savings that is provided by VMware vSphere thin provisioning. N series data deduplication and the thin provisioning of VMFS datastores and RDM LUNs offer considerable storage savings by increasing storage utilization of the N series system. Both of these technologies are native to N series storage system and do not require any configuration considerations or changes to be implemented within the VMware ESX Servers.

## 10.2 Data deduplication

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



Deploying from templates saves administrative time by copying the configuration and virtual disk files and registering this second copy as an independent VM. By design, this process introduces duplicate data for each new VM deployed. Figure 10-1 shows an example of typical storage consumption in a vSphere deployment.



*Figure 10-1 Storage consumption with a traditional array*

IBM N series offers a data deduplication technology called Advanced Single Instance Storage (A-SIS). With N series deduplication, a VMware vSphere deployment can eliminate the duplicate data in its environment, enabling greater storage utilization. Deduplication virtualization technology enables multiple virtual machines to share the same physical blocks in an N series storage system in the same manner that VMs share system memory. It can be seamlessly introduced into a virtual data center without having to make any changes to VMware Vsphere administration, practices, or tasks. Deduplication runs on the N series system at scheduled intervals and does not consume any CPU cycles on the VMware ESX Server.

Figure 10-2 shows the impact of deduplication on storage consumption in a VMware vSphere deployment.

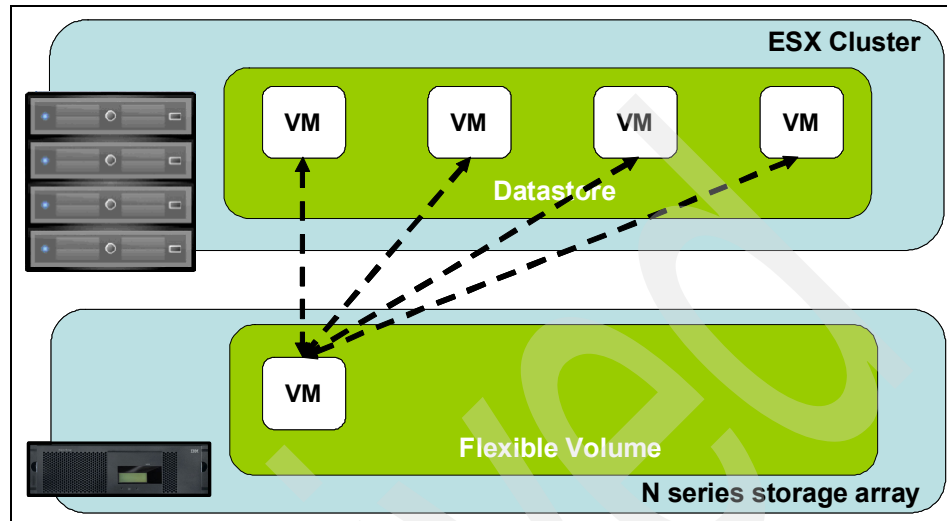


Figure 10-2 Storage consumption after enabling data deduplication

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

**Note:** If you enable data deduplication and have no effect in terms of the number of VMs that reside on a datastore, be sure to size a datastore density the same way you do it without deduplication.

### 10.2.1 Deduplication considerations with VMFS and RDM LUNs

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

To recognize the storage savings of deduplication with LUNs, you must enable N series LUN thin provisioning. In addition, although deduplication reduces the amount of consumed storage, the VMware vSphere administrative team does not see this benefit directly, because the team's view of the storage is at a LUN layer, and LUNs always represent their provisioned capacity, whether the LUNs are traditional or thin-provisioned.

### 10.2.2 Deduplication considerations with NFS

Unlike with LUNs, when deduplication is enabled with NFS, the storage savings are immediately available and recognized by the VMware Vsphere administrative team. No special considerations are required for its usage.

## 10.3 Storage thin provisioning

You should be very familiar with traditional storage provisioning and with the manner in which storage is preallocated and assigned to a server, or in the case of VMware ESX, a virtual machine. A common practice is for server administrators to over provision the storage to avoid running out of storage and the associated application downtime when expanding the provisioned storage. Although no system can be run at 100% storage utilization, certain methods of storage virtualization exist that allow administrators to address and oversubscribe storage in the same manner as with server resources (such as CPU, memory, networking, and so on). This form of storage virtualization is referred to as *thin provisioning*.

Traditional provisioning preallocates storage; thin provisioning provides storage on demand. The value of thin-provisioned storage is that storage is treated as a shared resource pool and is consumed only as each individual VM requires it. This sharing increases the total utilization rate of storage by eliminating the unused but provisioned areas of storage that are associated with traditional storage. The drawback to thin provisioning and oversubscribing storage is that (without the addition of physical storage) if every VM requires its maximum possible storage at the same time, not be enough storage will be available to satisfy the requests.

N series system storage thin provisioning extends VMware ESX thin provisioning for VMDKs and allows LUNs that are serving VMFS datastores to be provisioned to their total capacity, but consume only as much storage as is required to store the VMDK files (which can be of either thick or thin format). In addition, LUNs connected as RDMs can be thin provisioned.

To create a thin-provisioned LUN, perform the following steps.

1. Open FilerView ([http://filer/na\\_admin](http://filer/na_admin)).
2. Select LUNs.
3. Select **Wizard**. The wizard opens.
4. In the Wizard window, click **Next**.
5. Enter the path. See Figure 10-3.
6. Select the LUN type (for VMFS select **VMware**; for RDM select the **VM type**).
7. Enter a description and click **Next**.
8. Deselect the **Space-reserved** check box.
9. Click **Next** and then click **Finish**.

**LUN Wizard: Specify LUN Parameters**

**Path:**  
The full path to the LUN, for example /vol/luns/lunOne. The LUN must be created in the root directory of a volume or a qtree.

**Size:**  
The size of the LUN.

**LUN Protocol Type:**  
Select the multiprotocol type for the LUN.

**Space-reserved:**  
If checked, indicates that the LUN should be space-reserved. ☐ space-reserved

**Description:**  
An optional description of the LUN.

< Back   Cancel   Next >

Figure 10-3 Enabling thin provisioning on a LUN

When you enable N series thin provisioning, be sure to also configure storage management policies on the volumes that contain the thin-provisioned LUNs. These policies aid in providing the thin-provisioned LUNs with storage capacity, as required. The policies include automatic sizing of a volume, automatic Snapshot copy deletion, and LUN fractional reserve.

Volume Auto Size is a policy-based space management feature in Data ONTAP that allows a volume to grow in defined increments up to a predefined limit when the volume is nearly full. For VMware vSphere environments, be sure to set this value to on. Doing so requires setting the maximum volume and increment size options.

To enable these options, perform the following steps:

1. Log in to the N series console.
2. Set the Volume Auto Size policy:

```
vol autosize <vol-name> [-m <size>[k|m|g|t]] [-i <size>[k|m|g|t]] on
```

Snapshot Auto Delete is a policy-based space-management feature that automatically deletes the oldest Snapshot copies on a volume when that volume is nearly full. For VMware vSphere environments, be sure to set this value to delete Snapshot copies at 5% of available space. In addition, be sure to set the volume option to have the system attempt to grow the volume before deleting Snapshot copies. To enable these options, perform the following steps:

1. Log in to the N series console.
2. Set the Snapshot Auto Delete policy:

```
snap autodelete <vol-name> commitment try trigger volume  
target_free_space 5 delete_order oldest_first
```

3. Set the Volume Auto Delete policy:

```
vol options <vol-name> try_first volume_grow
```

LUN Fractional Reserve is a policy that is required when you use N series Snapshot copies on volumes that contain VMware ESX LUNs. This policy defines the amount of additional space reserved to guarantee LUN writes if a volume becomes 100% full. For VMware vSphere environments where Volume Auto Size and Snapshot Auto Delete policies are in use and you have separated the swap, pagefile, and other transient data onto other LUNs and volumes, set this value to 0%. Otherwise, leave this setting at its default of 100%.

To enable this option, perform the following steps:

1. Log in to the N series console.
2. Set the Volume Snapshot Fractional Reserve:

```
vol options <vol-name> fractional_reserve
```





## Virtual Machine best practices

This chapter describes best practices for integrating VMware virtual machines (VMs) with N series storage.

## 11.1 Optimizing Windows VM file system performance

If your virtual machine is not acting as a file server, consider implementing the following change to your virtual machines. This change disables the access-time update process in NTFS, and reduces the amount of IOPs occurring within the file system. To make this change, perform the following steps:

1. Log in to a Windows VM.
2. Select **Start**  **Run**, enter **CMD** mode, and then enter the following command:

```
fsutil behavior set disablelastaccess 1
```

## 11.2 Ensuring optimum VM availability

In Chapter 6, “The N series VMware ESX Host Utilities” on page 51, we covered the ESX Host Utilities (EHU). One component of the host utilities is the guest operating system (GOS) timeout scripts, which are a collection of ISO images that can be mounted by a VM to be able to configure its local SCSI to values that are optimal for running in a virtual infrastructure. To install the GOS timeout scripts and optimize the SCSI bus, perform the following steps:

1. Download the EHU.
2. Copy the EHU to a location that is accessible to the ESX server.
3. Extract the EHU by running the following command:  

```
tar -zxvf <name of EHU file>.tar.gz.
```
4. From within vCenter Server, right-click a VM to upgrade it, and select **Edit Settings**.
5. Select CDROM and then select the **ISO** button.
6. Select the appropriate ISO, matching the operating system of the VM you are configuring.
7. Click **OK**.
8. Connect to the VM console.
9. Run the script for the operating system of the VM.
10. Exit and unmount the ISO image.

Repeat these steps as necessary for each VM.



## 11.3 Ensuring optimal storage performance

Alignment of VM partitions and VMFS to storage arrays, virtual machines store their data on virtual disks. As with physical disks, these virtual disks contain storage partitions and file systems, which are created by the VM's guest operating system. To make sure of optimal disk I/O within the VM, align the partitions of the virtual disks to the block boundaries of VMFS and the block boundaries of the storage array. Failure to align all three items can result in a dramatic increase of I/O load on a storage array and negatively affects the performance of all virtual machines being served on the array.

Be sure that the partitions of VMs and the partitions of VMFS datastores are to be aligned to the blocks of the underlying storage array. For more information about VMFS and GOS file system alignment, see the following documents:

- ▶ IBM: *Storage Block Alignment with VMware Virtual Infrastructure*, NS3593
- ▶ VMware: *Recommendations for Aligning VMFS Partitions*

See “Related publications” on page 125 for links to IBM and VMware resources.

### 11.3.1 Datastore alignment

N series systems automate the alignment of VMFS with N series iSCSI, FC and FCoE LUNs. This task is automated during the LUN provisioning phase of creating a datastore, when you select the LUN type **VMware** for the LUN. If you deploy VMware over NFS, aligning the datastore is not necessary. With any type of datastore, VMFS or NFS, the virtual disks contained within the datastore, should have the partitions aligned to the blocks of the storage array.

### 11.3.2 VM partition alignment





When aligning the partitions of virtual disks for use with N series systems, the starting partition offset must be divisible by 4096. As an example, the starting partition offset for Microsoft Windows 2000, 2003, and XP operating systems is 32256. This value does not align to a block size of 4096.

Virtual machines running a clean installation of Microsoft Windows 2008, 7, and Vista operating systems automatically have their starting partitions set to 1048576. By default this value does not require any adjustments.

**Note:** if your Windows 2008 or Vista VMs were created by upgrading an earlier version of Microsoft Windows to one of these versions, a high probability is that these images require partition alignment.

Finally, storage arrays can be over taxed and as the virtual data center grows the storage array can require hardware upgrades to meet the additional I/O load that is generated by this misalignment. Simply stated, you can save a company a significant amount of money by optimizing the I/O of their VMs.

### 11.3.3 Identifying partition alignment

To verify the starting partition offset for a Windows-based virtual machine, log on to the VM and run the System Information utility (or `msinfo32`). The utility shows the Partition Starting Offset value (Figure 11-1). To run `msinfo32`, select **Start**  **All Programs**  **Accessories**  **System Tools**  **System Information**.

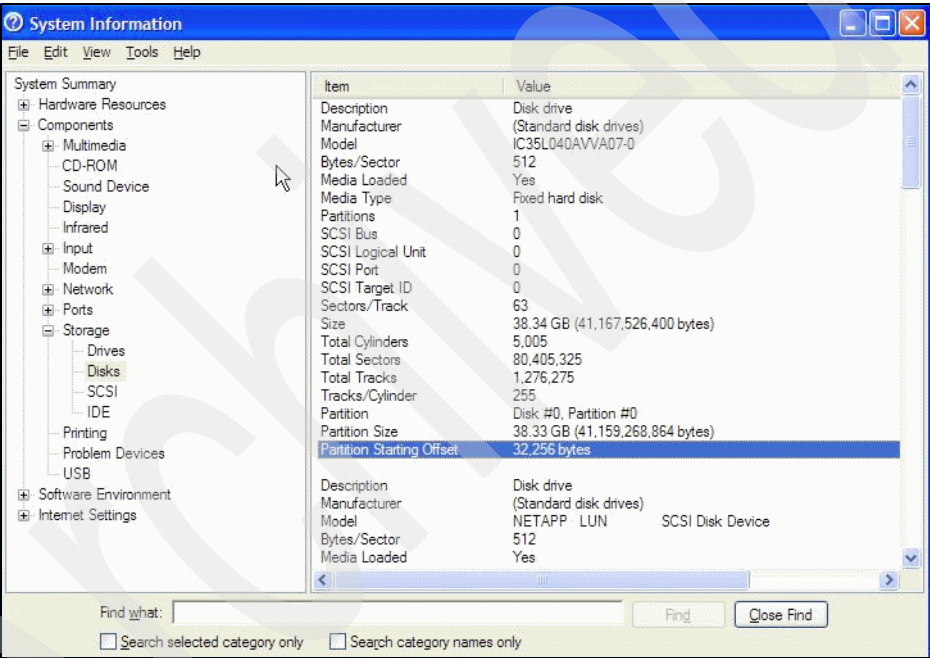


Figure 11-1 Using system information to identify the starting partition offset

### 11.3.4 MBRtools: identification of partition alignment status

A tool named MBRScan is an integrated component of the VMware ESX Host Utilities. MBRScan runs on a VMware ESX host and can identify whether partitions are aligned with Windows and Linux virtual machines running within VMFS and NFS datastores. MBRScan is run against the virtual disk files that comprise a virtual machine. Although this process requires only a few seconds per VM to identify and report on the status of the partition alignment, each VM must be powered off.

For this reason, identifying the file system alignment from within each VM might be easier because this action is nondisruptive.

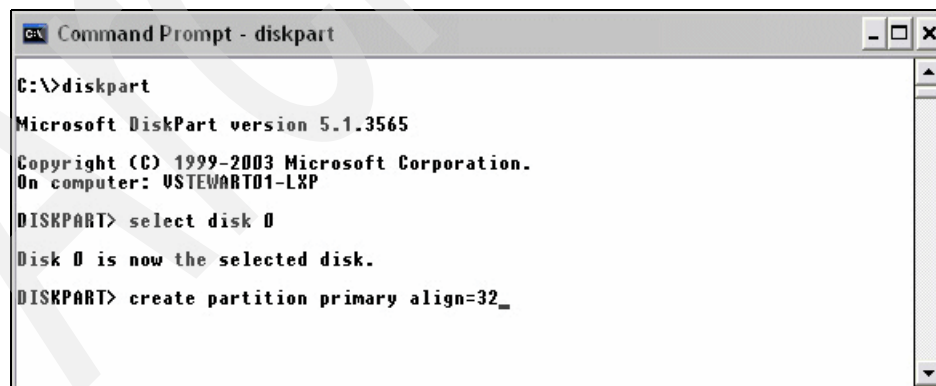
## 11.4 Creating properly aligned partitions for new VMs

In this section, we introduce utilities that can be used for aligning partitions in virtual machines.

### 11.4.1 Creating a properly aligned VMDK for a new VM with diskpart

Virtual disks can be formatted with the correct offset at the time of creation by simply booting the VM before installing an operating system and manually setting the partition offset. For Windows guest operating systems, consider using the Windows Preinstall Environment boot CD or alternative *live dvd* tools. To set up the starting offset, perform the following steps (and see Figure 11-2 for details):

1. Boot the VM with the Microsoft WinPE CD.
2. Select **Start**  $\otimes$  **Run** and enter the following command:  
`diskpart`
3. Enter the text: `select disk 0`
4. Enter the text: `create partition primary align=32`
5. Reboot the VM with WinPE CD.
6. Install the operating system as normal.



```
C:\>diskpart
Microsoft DiskPart version 5.1.3565
Copyright (C) 1999-2003 Microsoft Corporation.
On computer: USTEWART01-LXP
DISKPART> select disk 0
Disk 0 is now the selected disk.
DISKPART> create partition primary align=32_
```

Figure 11-2 Running diskpart to set a proper starting partition offset

## 11.4.2 Creating a properly aligned VMDK for a new VM with fdisk

This procedure works for VMware ESX .vmdk files that are hosted on VMFS or NFS datastores, for both Windows and Linux VMs. This procedure is *not* required for VMs that are running Windows Server 7, 2008, and Vista, because the file systems with these operating systems are aligned by default. To set up the starting offset using the **fdisk** command in the ESX service console, perform the following steps:

1. Log in to the ESX service console.
2. Change to the VM directory and view it by typing the following commands:  

```
cd /vmfs/volumes/<datastore>/<VM home dir>
ls -l
```
3. Identify the number of cylinders in the virtual disk by reading the virtual disk descriptor file. Look for the line `ddb.geometry.cylinders`.  

```
cat <Virtual Disk>.vmdk
```
4. Run **fdisk** on the virtual disk file (the `-flat.vmdk` file) by typing the following command:  

```
fdisk ./<Virtual Disk>.vmdk
```
5. When in **fdisk**, enter Extended Mode by typing `x` and pressing Enter.
6. Select the option to set the number of cylinders. Start by typing `c` and pressing Enter.
7. Enter the number of cylinders that you found from step 3.
8. Type `p` at the expert command screen to look at the partition table.  
The results are a table of all zeros.
9. Return to Regular mode by typing `r`.
10. Create a new partition by typing `n` and then `p` when you are prompted for the partition type.
11. Enter 1 for the partition number, 1 for the first cylinder, and then press Enter for the last cylinder question to force the default value.
12. Go into Extended Mode, to set the starting offset, by typing `x`.
13. Set the starting offset by typing `b` and pressing Enter, selecting 1 for the partition and pressing Enter, and finally entering 64 and pressing Enter.

**Note:** The value 64 represents the number of 512 bytes that are used to create a starting offset of 32,768 KB.

14. Check the partition table by typing `p`. The top row of the output displays disk geometry, including the starting offset of 64.
15. Type `r` to return to the menu.
16. To set the system type to HPFS/NTF type `t`.
17. Enter 7 for the hexcode.
18. Save and write the partition by typing `w`. Ignore the warning, which is the normal message.
19. Start the VM and run Windows setup. During the installation process, you are prompted that a partition exists. Select this partition to format and in which to install Windows.





## Virtual Machine storage layout

This chapter provides information and best practices for VMware ESX virtual machines storage integration and layouts.

## 12.1 Default virtual machine layout

When a virtual machine is provisioned, the VMware ESX administrator must select a datastore to store the files that comprise the VM. The directory that is created is referred to as the VM home directory. By default, all of the files for a single VM reside in the VM home directory. The contents of the home directory include, but are not limited to, the VM's configuration file, virtual disk and virtual disk descriptor files, virtual swapfile, snapshot files, NVRAM, and so on.

From the perspective of simplicity, this design works well where a VM home directory is a virtual machine. See Figure 12-1 for a high-level conceptual view of this layout.

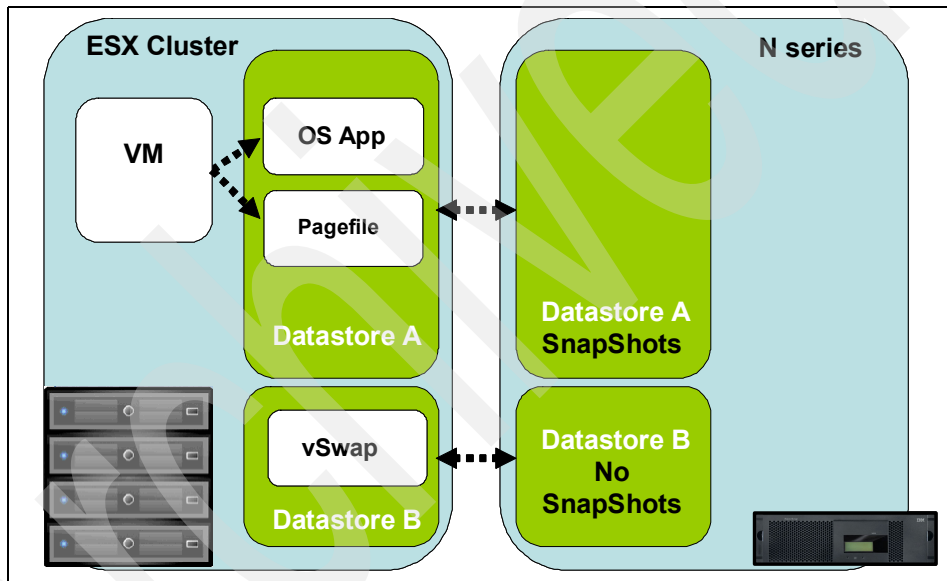


Figure 12-1 VMware's default virtual machine and vswap layout

## 12.2 Virtual machine layout with N series Snap technologies

In this section, we review a data layout design that we suggest using when you integrate VMware vSphere with N series *snap* technologies, such as SnapManager, Snapshot backups or disk-to-disk replication using SnapMirror and SnapVault. In these use case scenarios we recommend separating transient



and temporary data from the production data by implementing architecture that separates these two data types into multiple datastores.

This design is not specific to N series; rather it is an optimal consideration when you deploy VMware ESX on any storage array, providing snapshot backup or disk-based replication. These types of technologies manage the files that make up a VM, not the content inside these files, and as such consume a substantial amount of additional disk and bandwidth if the temporary and transient data is not separated from the production data.

### 12.2.1 Layout option 1: Implement a central virtual swap datastore

VMware ESX Servers create a virtual swap (vswap) file for every running VM. The sizes of these files are considerable, by default the vswap is equal to the amount of memory configured for each VM. Because this data is transient in nature, and not required in the case of recovering a VM from either a backup copy or using VMware Site Recovery Manager, the virtual swap file for every virtual machine must be relocated from the VM home directory to a datastore, on a separate N series volume that is dedicated to storing virtual swap files. See Figure 12-2 for a high-level conceptual view of this layout.

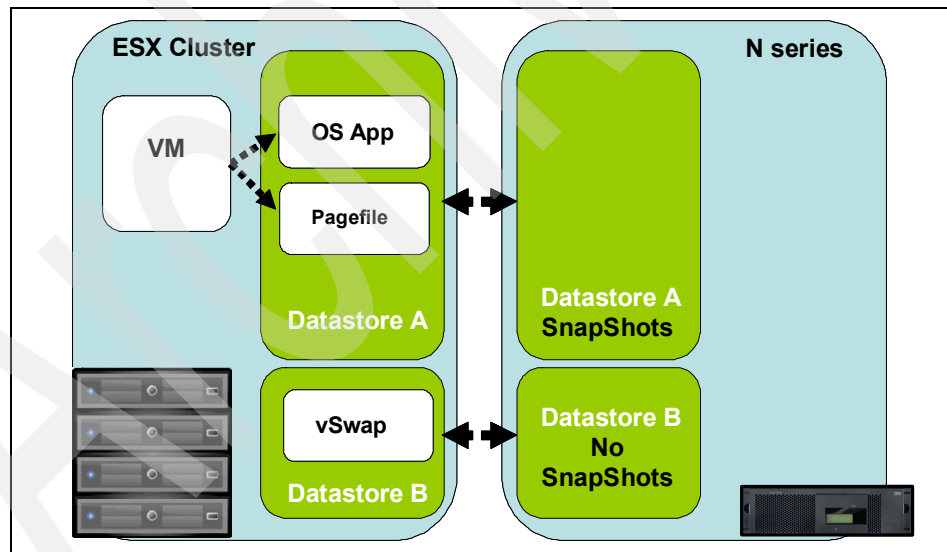


Figure 12-2 A central vSwap datastore for the entire cluster

A prerequisite to making this change is the creation of a datastore to store the swap files. Because the VMware ESX swap file storage requirements are dynamic, we suggest creating either a large thin-provisioned LUN or a FlexVol

volume with the Auto Grow feature enabled. Thin-provisioned LUNs and Auto Grow FlexVol volumes provide a large management benefit when storing swap files. This design removes the need to micromanage the swap space or to reduce the utilization rate of the storage. Consider the alternative of storing VMware ESX swap files on traditional storage arrays. If you undersize the swap space, the VMs fail to start; conversely, if you oversize the swap space, you have provisioned but unused storage.

**Note:** To use a centralized vswap datastore, VMware has documented that the following options must *not* reside in the VMX file:

- ▶ sched.swap.dir
- ▶ sched.swap.derivedName

To configure a central datastore to store the virtual swap files, perform the following steps (and see Figure 12-3 on page 101):

1. Open the vCenter Server.
2. Select a VMware ESX Server.
3. Select the **Configuration** tab.
4. In the Software box, select **Virtual Machine Swapfile Location**.
5. In the pane on the right, select **Edit**. The Virtual Machine Swapfile Location wizard opens
6. Select the datastore that will be the global location.

Repeat steps 2 - 6 for each VMware ESX Server in the cluster.

This process configures the VMware ESX Server and does not affect existing VMs. See the next procedure for configuring existing VMs.

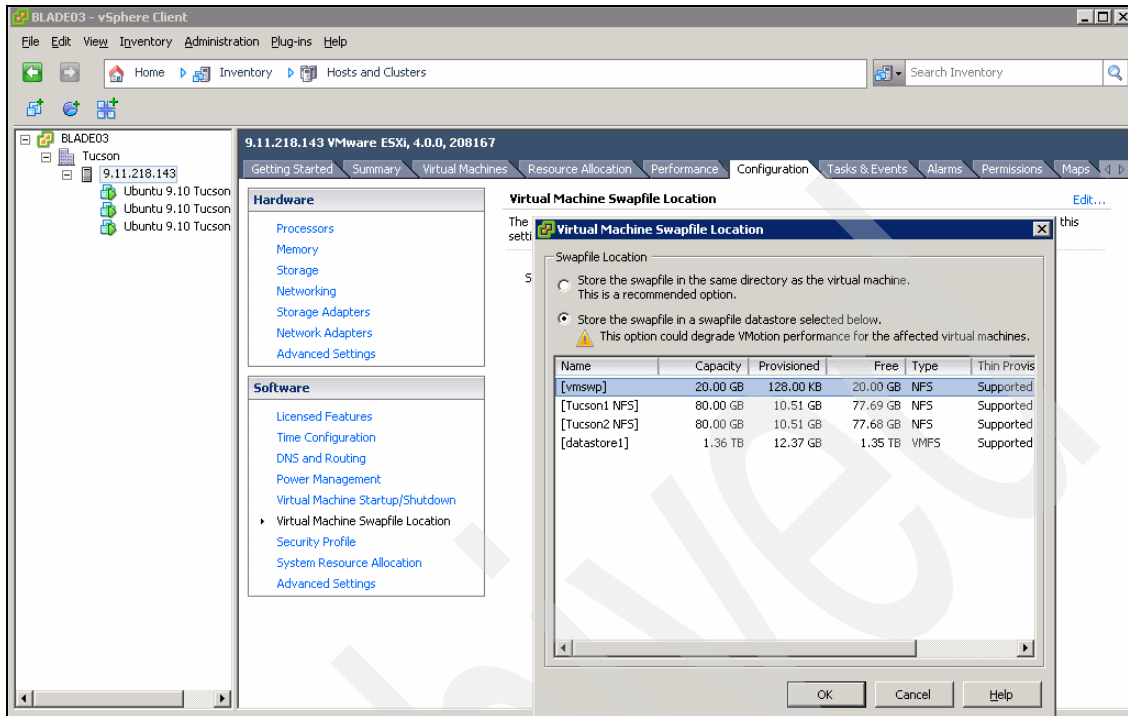


Figure 12-3 Configuring a global location for virtual swap files

To configure a central datastore to store the virtual swap files for VMs that have been deployed, perform the following steps (and see Figure 12-4 on page 102):

1. Open the vCenter Server.
2. Select a virtual machine.
3. Right-click and select **Edit Settings**.
4. Select the **Options** tab.
5. Under **Advanced**, select **Swapfile Location**.
6. To relocate the vswap file for this VM, select **Default**.
7. For this change to take effect, perform *either* of the following steps:
  - Migrate each VM to an ESX Server that is configured with a central vswap datastore.
  - Restart each VM on the existing VMware ESX Server that is configured with a central vswap datastore.

Repeat steps 2 - 7 for each existing VM.

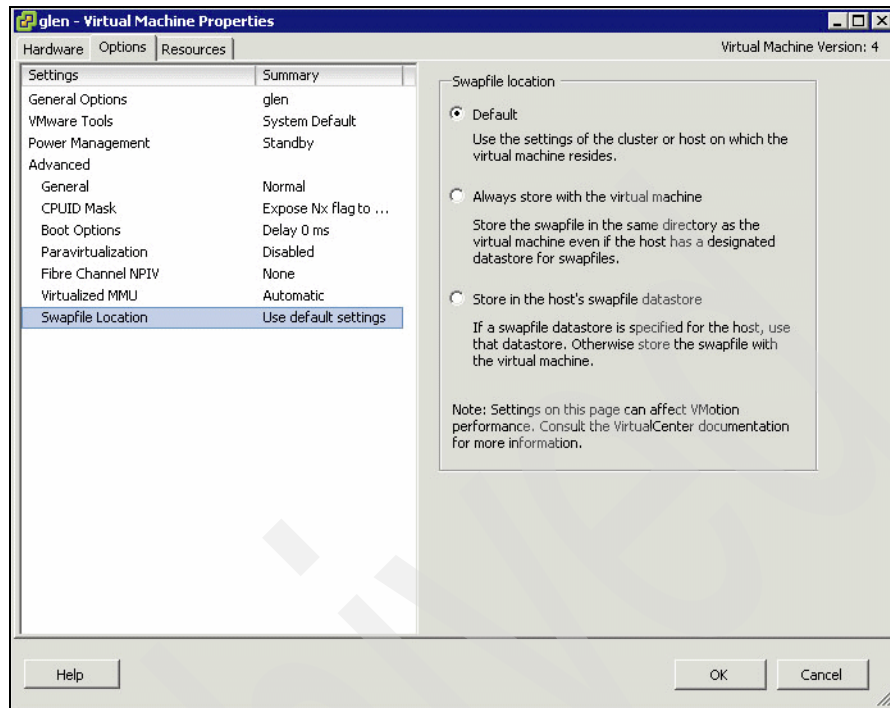


Figure 12-4 Defining the location for virtual swap files for a VM

## 12.2.2 Layout option 2: Locate VM swap or pagefile on a second datastore

In this design layout, although based on layout option 1, we relocate the virtual machine's swap or pagefile in an alternative datastore. This design has advantages and disadvantages that you must consider prior to the implementation. These details are covered after we review the architecture.

Each VM creates a swap or pagefile that is typically 1.5 - 2 times the size of the amount of memory that is configured for each VM. Because this data is transient in nature, we can save an amount of storage and bandwidth capacity by removing this data from the datastore, which contains the production data. To accomplish this design, the VM's swap or pagefile must be relocated to a second virtual disk, stored in a separate datastore, on a separate N series volume. Figure 12-5 on page 103 shows a high-level conceptual view of this layout.

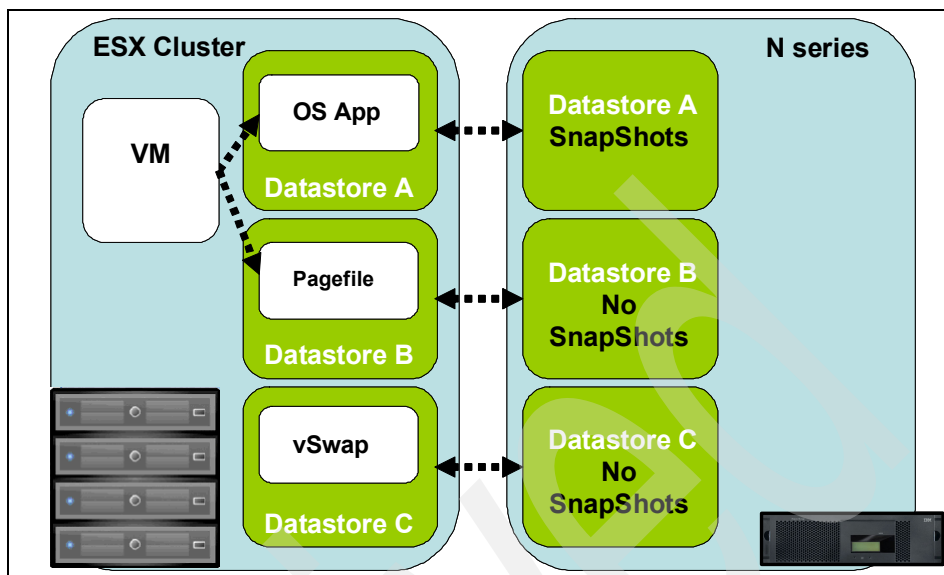


Figure 12-5 Separated VM pagefile to a pagefile datastore and a central vswap datastore

As stated previously, this design has advantages and disadvantages. A benefit is that no temporary and transient data will be contained in either a Snapshot backup or replicated data set, therefore a large amount of storage is conserved.

This design has a negative effect on customers who implement VMware Site Recovery Manager. In this design, the entire VM is not being replicated, so you must configure a second VMDK for each VM in their VMware vCenter Site Recovery Manager recovery plan.

Archived



## Storage monitoring and management

This chapter shows how to monitor and manage storage in VMware vSphere and IBM System Storage N series environments.

## 13.1 Monitoring storage utilization with N series Operations Manager

N series Operations Manager monitors, manages, and generates reports about all the N series storage systems in an organization. When you use N series thin provisioning, be sure to deploy Operations Manager and set up e-mail and pager notifications to the appropriate administrators. With thin-provisioned storage, monitoring the free space available in storage aggregates is very important.

Proper notification of the available free space means that additional storage can be made available before the aggregate becomes completely full.

For more information, see *Managing Unified Storage with IBM System Storage N series Operation Manager*, SG24-7734.

## 13.2 Storage growth management

This section explains how to manage storage growth in VMware environments.

### 13.2.1 Growing VMFS datastores

VMware ESX/ESXi 4 helps you to increase the storage for a VMFS datastore by supporting the dynamic growth of the VMFS file system. Alternatively, adding a VMFS extent can grow a datastore. This second option results in a spanned VMFS datastore. The decision regarding when to use which technology with N series storage arrays is simple: Grow the size of a LUN, then grow VMFS; and add a new LUN, and then add an extent.

Because N series storage systems have array based queue limits, as opposed to the LUN-based queue limits in traditional storage array architectures, be sure to always grow the LUN and VMFS. Extents are necessary only if a datastore requires more than 2 TB of storage. The value of 2 TB is the maximum size of a LUN that can be accessed by ESX/ESXi.

To grow a VMFS file system, perform the following steps:

1. Open FilerView ([http://filer/na\\_admin](http://filer/na_admin)).
2. Select **LUNs**.
3. Select **Manage**.
4. In the left pane, select a LUN from the list.



5. Enter the new size of the LUN in the Size box and click **Apply**.
6. Open the vCenter Server.
7. Select a VMware ESX host.
8. Select the *Configuration* tab.
9. In the Hardware box, select **Storage Adapters**.
10. In the pane on the right, select the HBAs and then select the **Rescan** link.  
This step results in the identification of the additional storage capacity by the VMware ESX/ESXi host.
11. In the Hardware box, select **Storage**.
12. In the pane on the right, select the datastore to grow and then select **Increase Datastore Capacity**.
13. Select the LUN, verify free space availability, click **Next**, and then click **Next** again. See Figure 13-1.
14. Be sure that the **Maximize Space** check box is selected, and then click **Next**, and then **Finish**.

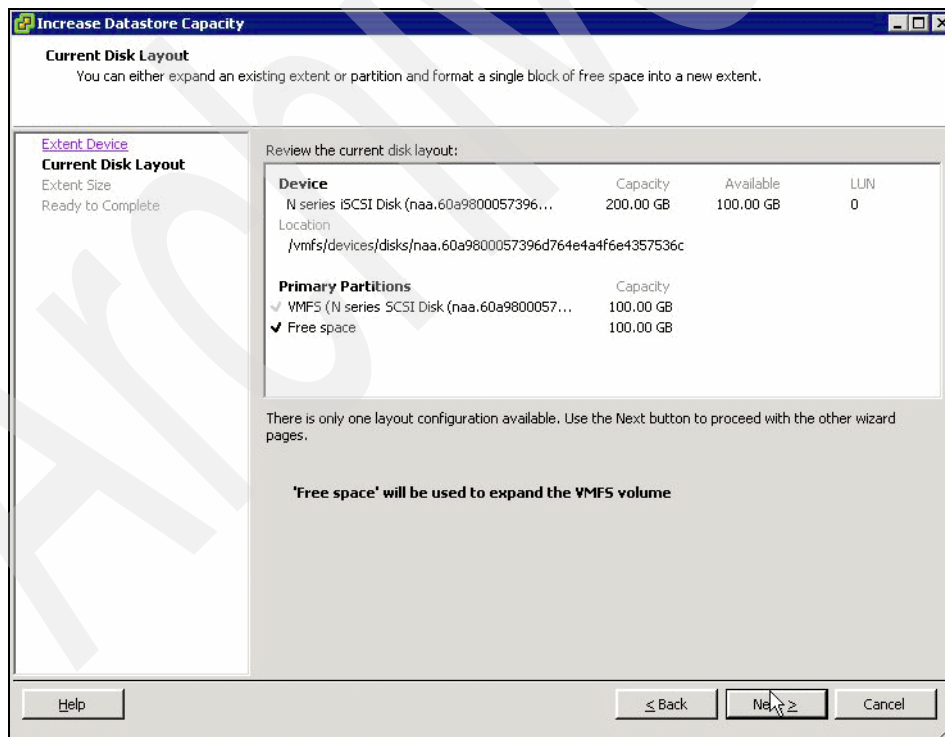


Figure 13-1 Expanding a VMFS partition

## 13.2.2 Growing a virtual disk (VMDK)

Although virtual disks can be extended, the process requires the virtual machine to be powered off. Growing the virtual disk is only part of the equation for increasing available storage; you must also grow the file system after the VM boots. Root volumes such as C:\ in Windows and forward slash (/) in Linux cannot be grown dynamically or while the system is running. For these volumes, see 13.2.5, “Growing bootable volumes within a guest operating system” on page 112. For all other volumes, you can use native operating system tools to grow the volume.

To grow a virtual disk, perform the following steps:

1. Open the vCenter Server.
2. Select a VM and shut it down.
3. Right-click the VM and select **Properties**.
4. Select a virtual disk and increase its size (see Figure 13-2)
5. Start the VM.

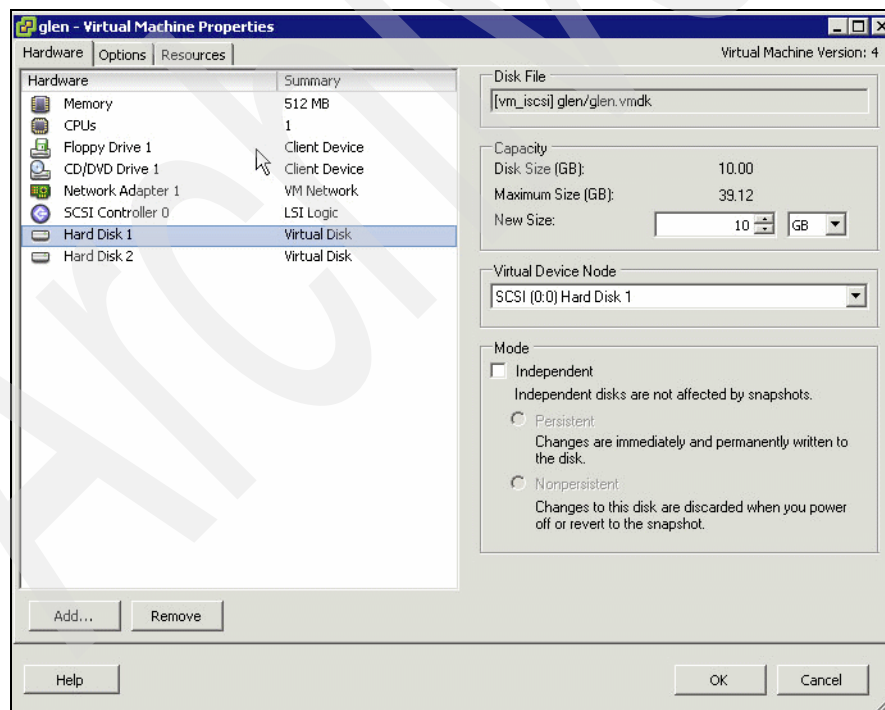


Figure 13-2 Increasing the size of a virtual disk

### 13.2.3 Growing a raw device mapping (RDM)

Growing an RDM has aspects of growing a VMFS and a virtual disk. This process requires the virtual machine to be powered off. To grow RDM-based storage, perform the following steps:

1. Open the vCenter Server.
2. Select a VMware ESX host and power down the VM.
3. Right-click the VM and select **Edit Settings** to open the Edit Settings window.
4. Highlight the hard disk to be resized and click **Remove**. Select **Remove from Virtual Machine and delete files from disk**. This action deletes the mapping file but does not remove any data from the RDM LUN. See Figure 13-3.

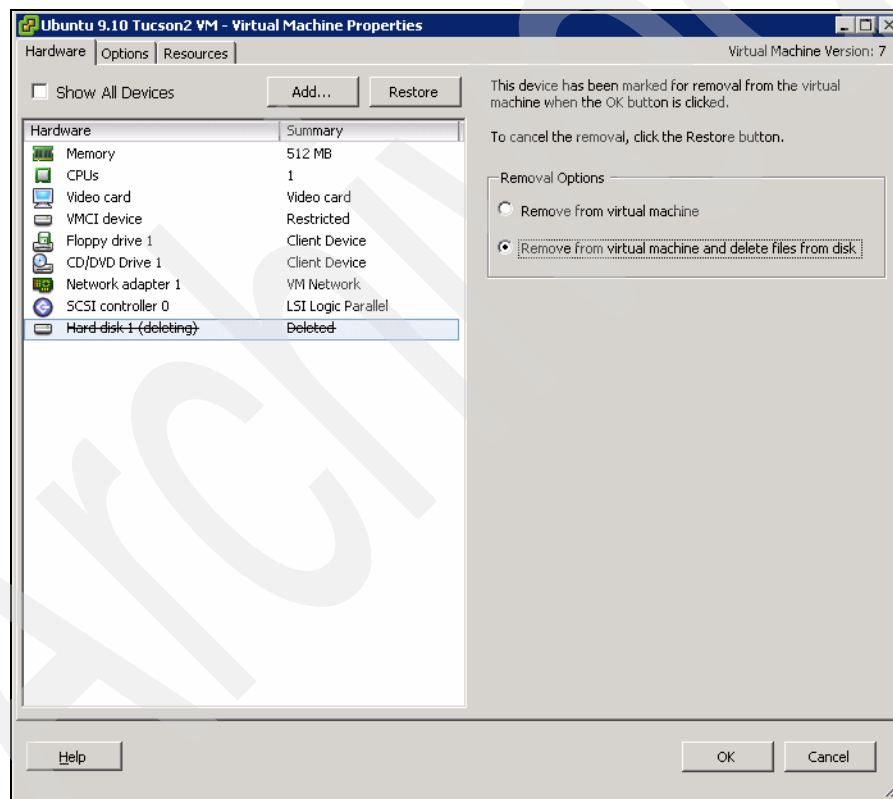


Figure 13-3 Deleting a VMDK from a VM

5. Open FilerView ([http://filer/na\\_admin](http://filer/na_admin)).
6. Select **LUNs**.

7. Select **Manage**.
8. From the list in the left pane, select the LUN.
9. In the Size box, enter the new size of the LUN and click **Apply**.
10. Open the vCenter Server.
11. Select the **Configuration** tab.
12. In the Hardware box, select **Storage Adapters**.
13. In the pane on the right, select the HBAs and select **Rescan**.
14. Right-click the VM and select **Edit Settings** to open the Edit Settings window.
15. Click **Add**, select a disk (see Figure 13-4), and then click **Next**.

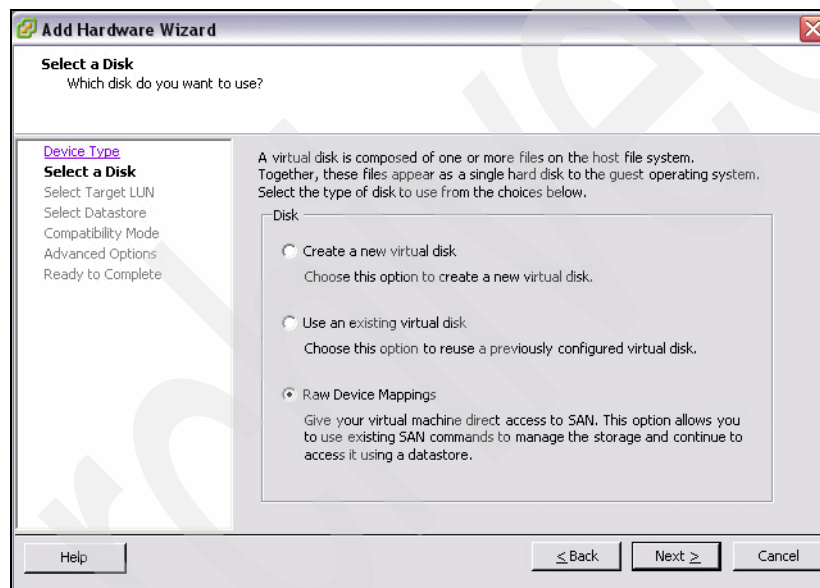


Figure 13-4 Connecting an RDM to a VM

16. Select the LUN and click **Next** (see Figure 13-5 on page 111).

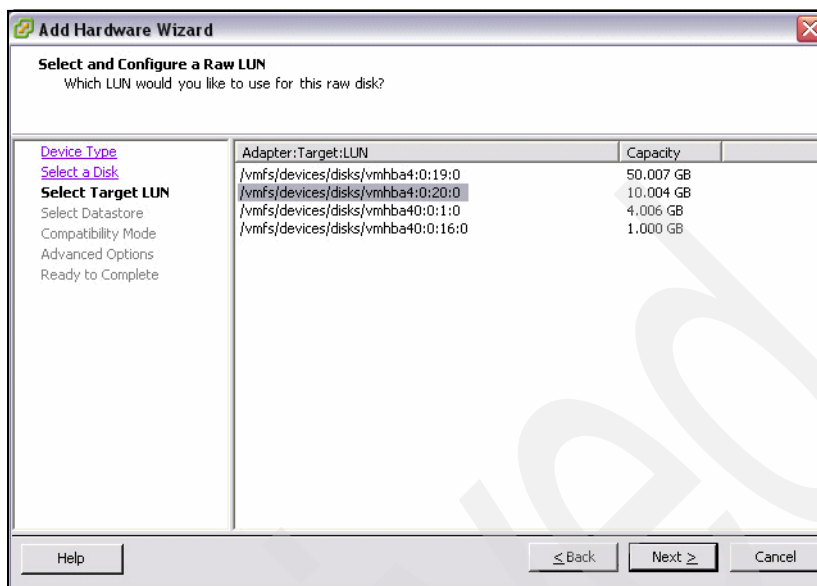


Figure 13-5 Selecting a LUN to mount as an RDM

17. Specify the VMFS datastore that will store the Mapping file.
18. Start the VM. Remember that although you have grown the LUN, you must also grow the file system within it. Follow the guidelines in 13.2.4, “Growing a file system within a guest OS (NTFS or EXT3)” on page 111, next.

### 13.2.4 Growing a file system within a guest OS (NTFS or EXT3)

When a virtual disk or RDM has been increased in size, you still must grow the file system residing on it after booting the VM. You may perform the following process while the system is running, by using native or freely distributed tools:

1. Remotely connect to the VM.
2. Grow the file system, as follows:
  - For Windows VMs, you can use the **diskpart** utility to grow the file system. For information, go to the following location:  
<http://support.microsoft.com/default.aspx?scid=kb;en-us;300415>
  - For Linux VMs, you can use **ext2resize** to grow the file system. For information, go to the following location:  
<http://sourceforge.net/projects/ext2resize>

### 13.2.5 Growing bootable volumes within a guest operating system

Root volumes such as C:\ in Windows VMs and forward slash (/) in Linux VMs cannot be grown concurrently or while the system is running. A simple way exists to expand these file systems, and does not require the acquisition of any additional software (except for **ext2resize**). This process requires the VMDK or LUN, which has been resized, to be connected to another virtual machine of the same operating system type. After the storage is connected, the hosting VM can run the utility to extend the file system.

After extending the file system, this VM is shut down and the storage is disconnected. Connect the storage to the original VM. When you boot, you can verify that the boot partition now has a new size.



## Disk-based Snapshot backups for VMware

This chapter provides a brief overview on how Snapshot technology integrates with VMware vSphere.

## 14.1 Complementary Snapshot technologies

With VMware vSphere, you can create Snapshot copies of virtual machines. Snapshot technologies allow the creation of point-in-time copies that provide the fastest means to recover a VM to a previous point in time. The IBM System Storage N series has been providing customers with the ability to create Snapshot copies of data. Although the basic concept of a Snapshot copy is similar between the N series and VMware, be aware of the differences between the two, and when to use one rather than the other.

VMware Snapshot copies provide simple point-in-time versions of VMs, allowing quick recovery. The benefits of VMware Snapshot copies are that they are easy to create and use, because they can be executed and scheduled from within vCenter Server. VMware suggests not using the Snapshot technology in VMware ESX as a means to back up vSphere.

N series Snapshot technology can easily be integrated into VMware vSphere environments, where it provides crash-consistent versions of virtual machines for the purpose of full VM recovery, full VM cloning, or site replication and disaster recovery. This Snapshot technology does not have a negative effect on system performance.

VMware states that for optimum performance and scalability, hardware-based Snapshot technology is preferred over software-based solutions. The shortcoming of this solution is that it is not managed within vCenter Server, and requires external scripting, scheduling, or both to manage the process.

## 14.2 Implementing snapshot backups

The ability to quickly backup tens of virtual machines without affecting production operations can accelerate the adoption of VMware vSphere within an organization. IBM offers a means to do this with SnapManager for Virtual Infrastructure (or SMVI). SMVI builds on the N series SnapManager portfolio by providing the following benefits:

- ▶ Array-based backups that consume only block-level changes to each VM
- ▶ Multiple recovery points through out the day
- ▶ Recovery times faster than any other means because the backups are an integrated component within the storage array SMVI



Figure 14-1 shows the SMVI interface.

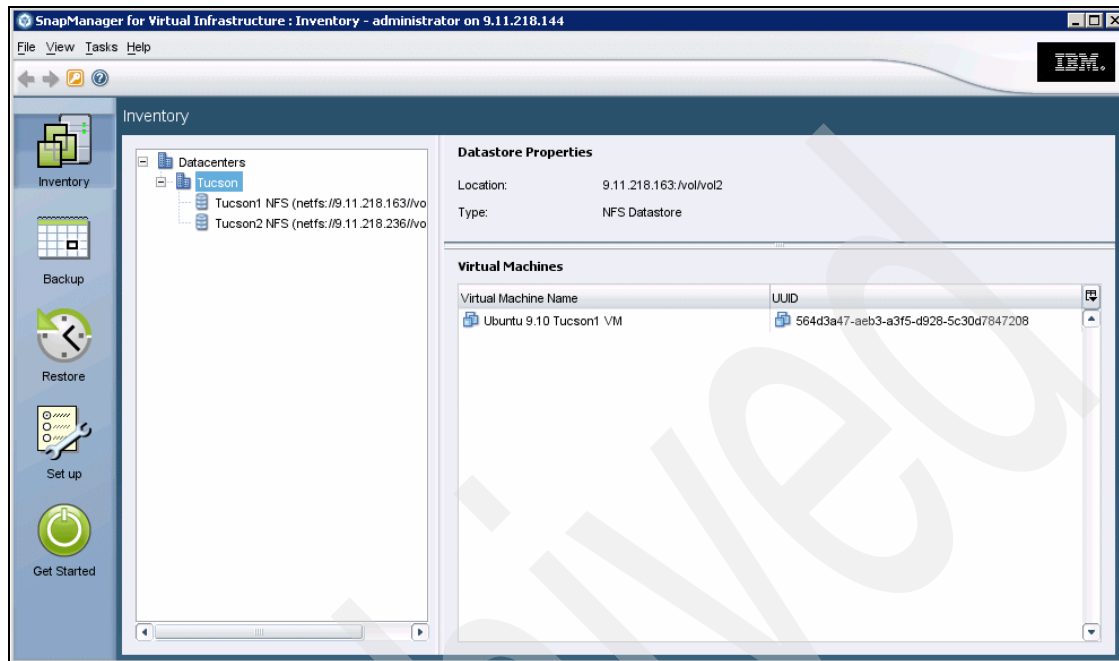


Figure 14-1 SMVI user interface

Archived

# Configuring SSH on VMware ESX servers and N series systems

The most efficient way to integrate N series Snapshot copies is to enable the centralized management and execution of Snapshot copies. A good practice is to configure the N series systems and VMware ESX Servers to allow a single host to remotely execute commands on both systems. This management host must have an Secure Shell (SSH) client installed and configured.

## N series system SSH configuration

To configure SSH access on a N series system, perform the following steps:

1. Connect to the N series system console (use either SSH, Telnet, or Console connection).
2. Execute the following commands:  

```
secureadmin setup ssh
options ssh.enable on
options ssh2.enable on
```

3. As root, log in to the Linux or VMware ESX system that remotely executes commands on the N series system.
4. Add the Triple DES cipher to the list of available SSH ciphers; this is the only cipher recognized by the N series storage system. Edit the `/etc/ssh/ssh_config` file and edit the Ciphers line to read as follows:

```
Ciphers aes128-cbc, aes256-cbc, 3des-cbc
```

5. Generate a DSA host key. On a Linux or VMware ESX Server, use the following command:

```
ssh-keygen -t dsa -b 1024
```

When prompted for the passphrase, do *not* enter one; instead, press Enter. The public key is saved to the following location:

```
/root/.ssh/id_dsa.pub
```

6. Mount the N series root file system as root.
7. Copy only the key information from the public key file to the N series system's `/etc/sshd/root/.ssh/authorized_keys` file, removing all information except for the key string preceded by the string `ssh-dsa` and a comment line.

Example A-1 on page 119 shows the key data (*do not use the example data*, because your key is different).

8. Test the connectivity from the remote host by issuing the version command on the N series system, as follows (it does not prompt for a password):

```
ssh <n series> version
Data ONTAP Release 7.3.2: Thu Oct 15 04:39:55 PDT 2009 (IBM)
```

#### Example A-1 Example key

```
ssh-dsa AAAAB3NzaC1kc3MAAABhALVbwVyhtAVoaZukcJSTlRb/RE01/ywbQECtAcHijzdzhEJU
z9Qh96HVEwyZDdah+PTxfyitJCerb+1FAn065v4WMq6jxPVYto615Ib5zxfq2I/hhT/6KPziS3LT
ZjKccwAAABUAjklMwkpipmg8Unv4fjCsYYhrSLOAAABgF9NsuZxni00HHR8tmW5RMX+M6VaH/nlJ
UzVXbLiI8+pyCXALQ29Y31uV3SzWtd1V0gjJHgv0GBw8N+rvGSB1r60VqgggGjSB+ZXA01Eecbnj
vLnUtF0TVQ75D9auagjOAAAAYEJPx8wi9/CaS3dfKJR/tYy7Ja+Mr1D/RC0gr22XQP1ydxsfYQx
enxzExPa/sPfjA45YtcUom+3mieFaQuWHZSNFr8sVJoW3LcF5g/z9Wkf5GwvGGtD/yb6bcsjZ4rt
```

## ESX system SSH configuration

To configure a VMware ESX Server to accept remote commands by using SSH, perform the following steps:

1. Log in to the VMware ESX console as root.
2. Enable the SSH services by running the following commands:  

```
esxcfg-firewall -e sshServer
esxcfg-firewall -e sshClient
```
3. Change to the SSH server configuration directory:  

```
cd /etc/ssh
```
4. Edit the configuration file:  

```
vi sshd_config
```
5. In the file, locate the following line:  

```
PermitRootLogin no
```
6. Modify the line so it says yes:  

```
PermitRootLogin yes
```
7. Restart the SSH service by running the following command:  

```
service sshd restart
```
8. Create the SSH public key by using the following command, which outputs content similar to Example A-2 on page 120 (retain the default locations, and *do not use* a passphrase):  

```
ssh-keygen -t dsa -b 1024
```

9. Change to the .ssh directory:

```
cd /root/.ssh
```

10. Run the following commands:

```
cat id_dsa.pub >> authorized_keys  
chmod 600 authorized_keys
```

11. Repeat these steps for each VMware ESX Server in the cluster.

Example A-2 shows a VMware ESX system SSH configuration.

*Example A-2 output*

---

Generating public/private dsa key pair.

Enter file in which to save the key (/home/root/.ssh/id\_dsa):

Enter passphrase (empty for no passphrase):

Enter same passphrase again:

Your identification has been saved in /home/root/.ssh/id\_dsa.

Your public key has been saved in /home/root/.ssh/id\_dsa.pub.

The key fingerprint is:

7b:ab:75:32:9e:b6:6c:4b:29:dc:2a:2b:8c:2f:4e:37 root@hostname

Your keys are stored in  
/root/.ssh.

---

## Relocating the pagefile in Windows Virtual Machines

This Appendix shows an example of how to relocate the page in Windows Virtual Machines.

Example B-1 on page 122 shows a Windows registry file, which is a simple registry script that sets the pagefile to the D:\ partition. Execute this script the first time a new virtual machine is created. If the D:\ partition does not exist, the system's default values are used.

The process of launching this script can be automated with Microsoft Setup Manager. To use the values in this example, copy the contents of this section and save it as a text file named `pagefile.reg`.

The Setup Manager has a section where you can add `pagefile.reg` to run the first time the virtual machine is powered on.

### *Example B-1 Registry file*

---

Start-----

Windows Registry Editor Version 5.00

[HKEY\_LOCAL\_MACHINE\SYSTEM\CurrentControlSet\Control\Session Manager\Memory Management]

"PagingFiles"=hex(7):64,00,3a,00,5c,00,70,00,61,00,67,00,65,00,66,00,69,00,6c,\

00,65,00,2e,00,73,00,79,00,73,00,20,00,32,00,30,00,34,00,38,00,20,00,32,00,\

30,00,34,00,38,00,00,00,00,00

End

-----

---



# Abbreviations and acronyms

<b>API</b>	application programming interface	<b>SRA</b>	Site Recovery Adapter
<b>AULA</b>	asymmetric logical unit access	<b>STP</b>	Spanning Tree Protocol
<b>BPDU</b>	Bridge Protocol Data Unit	<b>UI</b>	user interface
<b>CNA</b>	Converged Network Adapter	<b>VCB</b>	VMware Consolidated Backup
<b>DR</b>	disaster recovery	<b>vCenter</b>	Virtual Center
<b>DRS</b>	Distributed Resource Scheduler	<b>VIC</b>	VMware vCenter Server (formerly VMware Virtual Center)
<b>EHU</b>	ESX Host Utilities	<b>VIF</b>	Virtual Network Interface
<b>GUI</b>	graphical user interface	<b>VLAN</b>	virtual LAN
<b>HA</b>	high availability	<b>VM</b>	virtual machine
<b>HBA</b>	host bus adapter	<b>VMFS</b>	Virtual Machine File System
<b>IBM</b>	International Business Machines Corporation	<b>WWPN</b>	worldwide port name
<b>IQN</b>	iSCSI qualified name		
<b>ITSO</b>	International Technical Support Organization		
<b>LAN</b>	local area network		
<b>LUN</b>	logical unit number		
<b>MBR</b>	Master Boot Record		
<b>MSCS</b>	Microsoft Cluster Service		
<b>NDMP</b>	Network Data Management Protocol		
<b>NMP</b>	Native Multipathing		
<b>PSA</b>	Pluggable Storage Architecture		
<b>PSP</b>	Path Selection Plug-in		
<b>RDM</b>	raw device mapping		
<b>SATP</b>	Storage Array Type Plug-in		
<b>SFR</b>	Single File Restore		
<b>SLA</b>	service level agreement		
<b>SMVI</b>	SnapManager for Virtual Infrastructure		

Archived

# Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

## IBM Redbooks

For information about ordering these publications, see “How to get Redbooks” on page 127. Note that some of the documents referenced here might be available in softcopy only.

- ▶ *IBM System Storage N series*, SG24-7129
- ▶ *IBM System Storage N series MetroCluster*, REDP-4259
- ▶ *IBM System Storage N Series SnapMirror*, SG24-7260
- ▶ *IBM System Storage N series with Microsoft Clustering*, SG-7671
- ▶ *IBM System Storage N series with VMware ESX Server*, SG24-7636
- ▶ *Managing Unified Storage with IBM System Storage N series Operation Manager*, SG24-7734

## Other publications

These publications are also relevant as further information sources:

- ▶ *IBM System Storage N series Data ONTAP 7.3 Active/Active Configuration Guide*, GC27-2208
- ▶ *IBM System Storage N series Data ONTAP 7.3 Block Access Management Guide for iSCSI and FCP*, GC52-1282
- ▶ *IBM System Storage N series Data ONTAP 7.3 Data Protection Online Backup and Recovery Guide*, GC27-2204
- ▶ *IBM System Storage N series Data ONTAP 7.3 File Access and Protocols Management Guide*, GC27-2207
- ▶ *IBM System Storage N series Data ONTAP 7.3 Network Management Guide*, GC52-1280

- ▶ *IBM System Storage N series Data ONTAP 7.3 System Administration Guide*, GC52-1279
- ▶ *IBM System Storage N series SnapManager 2.0 for Virtual Infrastructure Installation and Administration Guide*, GC53-1145

## Online resources

These websites are also relevant as further information sources.

### IBM resources

- ▶ IBM System Storage NAS solutions  
[http://www.ibm.com/systems/storage/network/?cm\\_re=masthead\\_-\\_product\\_s\\_-\\_stg-nas](http://www.ibm.com/systems/storage/network/?cm_re=masthead_-_product_s_-_stg-nas)
- ▶ IBM System Storage N series and TotalStorage NAS interoperability matrixes  
<http://www.ibm.com/systems/storage/nas/interophome.html>
- ▶ Support for IBM System Storage and TotalStorage products  
<http://www-947.ibm.com/systems/support/>
- ▶ Support for Data ONTAP  
<http://www-304.ibm.com/jct01004c/systems/support/supportsite.wss/supportresources?brandind=5000029&familyind=5329797&taskind=1>
- ▶ IBM Storage Block Alignment with VMware Virtual Infrastructure and IBM System Storage N series  
<ftp://service.boulder.ibm.com/storage/isv/NS3593-0.pdf>

### VMware resources

- ▶ VMware documentation  
<http://www.vmware.com/support/pubs/>
- ▶ VMware Introduction to VMware vSphere  
[http://vmware.com/pdf/vsphere4/r40/vsp\\_40\\_intro\\_vs.pdf](http://vmware.com/pdf/vsphere4/r40/vsp_40_intro_vs.pdf)
- ▶ VMware ESXi Configuration Guide  
[http://vmware.com/pdf/vsphere4/r40/vsp\\_40\\_esxi\\_server\\_config.pdf](http://vmware.com/pdf/vsphere4/r40/vsp_40_esxi_server_config.pdf)

- ▶ VMware Basic System Administration  
[http://vmware.com/pdf/vsphere4/r40/vsp\\_40\\_admin\\_guide.pdf](http://vmware.com/pdf/vsphere4/r40/vsp_40_admin_guide.pdf)
- ▶ VMware Fibre Channel SAN Configuration Guide  
[http://vmware.com/pdf/vsphere4/r40/vsp\\_40\\_san\\_cfg.pdf](http://vmware.com/pdf/vsphere4/r40/vsp_40_san_cfg.pdf)
- ▶ VMware iSCSI SAN Configuration Guide  
[http://vmware.com/pdf/vsphere4/r40/vsp\\_40\\_iscsi\\_san\\_cfg.pdf](http://vmware.com/pdf/vsphere4/r40/vsp_40_iscsi_san_cfg.pdf)
- ▶ vSphere Upgrade Guide  
[http://vmware.com/pdf/vsphere4/r40/vsp\\_40\\_upgrade\\_guide.pdf](http://vmware.com/pdf/vsphere4/r40/vsp_40_upgrade_guide.pdf)

## How to get Redbooks

You can search for, view, or download Redbooks, Redpapers, Technotes, draft publications and Additional materials, as well as order hardcopy Redbooks publications, at this Web site:

[ibm.com/redbooks](http://ibm.com/redbooks)

## Help from IBM

IBM Support and downloads

[ibm.com/support](http://ibm.com/support)

IBM Global Services

[ibm.com/services](http://ibm.com/services)



# Index

## Numerics

802.3ad 64, 68

## A

active-active configuration 15

aggregates 11–12

Auto Grow FlexVol volumes 100

## C

cfmode 12–13

CIFS 15

client applications 14

Cluster Server 5

clustered file system 3

cp show cfmode 12

## D

data

deduplication 8

protection 10

datastore 6, 8

size 7

disaster recovery 14

DRS 7, 10

dual-fabric configuration 15

## E

ESX Server 2

EtherChannel 68

## F

Fabric MetroCluster 15

fault tolerance 3, 7

FC and FCoE 7

FCP 15

Fibre Channel

ports 12

protocol 5

solutions 4

FlexClone 6

flexible volume

configure 13

definition of 11

level 13

what it contains 13

FlexVol name 14

## G

GUI management tools 11

## H

hard zoning 56

hardware utilization 2

HBAs 10

high-availability solution 14

benefits 14

## I

I/O queue 5

IBM and VMware tools 5

iSCSI 3, 7

## L

link aggregation 68

Link Aggregation Control Protocol (LACP) 68

Linux 2

LUN

definition of 14

maximum number 7

multiples 4

name for RDMS 14

name for VMFS 14

server access 14

## M

mainframe-class 2

MetroCluster

behavior 15

configuration 15

integration 15

reference information 16

mirrored aggregates 14

MSCS 6–7

## N

N series 6, 9

- MetroCluster 14–15

- storage configurations 13

- storage systems 12, 14

- system console 12–13

NDMP 8

NFS

- servers 4–5

- supported backup features 8

NTFS 14

## O

oversubscribe storage 85

## P

P2V clustering 5

physical spindles 12

physical-to-virtual-machine host-based clustering 5

port zoning 56

## R

RAID 10

RAID 10 10

RAID 5 10

RAID groups 10

- overhead 10

RAID-DP 15

- data residing on 10

- definition of 10

- protection 10

raw device mapping 5, 14

RDM

- deploying with 3

- two modes 6

Redbooks Web site 127

- Contact us xi

redundant controllers 10

## S

SAN zoning 56

SANscreen 8

SCSI 6

server consolidation 2

set cfmode single\_image 13

single aggregate 12

single points of failure 10

Single System Image (SSI) 12

small aggregate 11

SMVI 8

snap reserve vol-name 0 command 13

snap sched vol-name 0 0 0 command 13

SnapDrive 6, 8

SnapManager for Virtual Infrastructure (SMVI) 8

SnapMirror 8

- replication 13

SnapShot

- backups 8, 13

- copies 4, 13

- schedule 13

SnapVault 8

soft zoning 56

Spanning Tree Protocol (STP) 61

storage

- array 5

- availability 10

- configuration 13

- thin provisioning 85

- vMotion 7

STP 61

Stretch MetroCluster 15

synchronous mirroring 15

system downtime 14

System x 2

## T

target zones 57

thin provisioning 6–8, 85

thin-provisioned

- LUN 99

- virtual disks 82

## V

VCB 7

virtual

- data center 10

- disk files 14

- disks 82

- infrastructure 5, 14

virtual machine

- See VM

Virtual Machine File System

- See VMFS



- VM 4–5, 14
  - level 6
  - move when running 10
  - multiple 4, 10
- VMDK 82
  - file level 8
  - image 8
  - per LUN 7
- VMFS 3–5
  - datastores 3, 5
- vMotion 4, 7, 10
- VMware 1, 3, 10
  - administration teams 6
  - clusters 6
  - data 10
  - data layout 13
  - features 6
  - virtualization 2–4
- vMware 3
- VMware Consolidated Backup (VCB) 7
- VMware ESX 3, 5, 12
  - cluster 4
  - host utilities 8
- VMware ESX Server 2–3, 10, 13–14
- VMware HA 7, 10
- VMware NMP 7
- VMware vCenter Server 3
- VMware vSphere 15
- VMware Workstation or VMware GSX Server 2
- volume name 13
- volume Snapshot
  - reserve 13
  - schedule 13

## W

- Windows Server 2003 2
- WWN zoning 57

## Z

- zeroed thick VMDK 82
- zones 56
  - hard 56
  - port 56
  - soft 56
  - target 57
  - WWN 57

Archived









# IBM System Storage N series and VMware vSphere Storage Best Practices



## **Working with Virtual Machine File System data stores**

## **Cluster sizing considerations when using LUNs**

## **Ensuring optimal storage performance**

IBM System Storage N series technology enables companies to extend their virtual infrastructures to include the benefits of advanced storage virtualization. The N series offers unified storage solutions that provide industry-leading technologies in the areas of storage efficiencies, instantaneous virtual machine and datastore cloning for virtual servers and virtual desktops, and virtual data center backup and business continuance solutions.

This IBM Redbooks publication reviews the best practices for anyone who is implementing VMware vSphere with N series unified storage arrays.

## **INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION**

## **BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE**

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

**For more information:**  
[ibm.com/redbooks](http://ibm.com/redbooks)