NETAPP TECHNICAL REPORT

NetApp and VMware Virtual Infrastructure 3 Storage Best Practices

Vaughn Stewart, Larry Touchette, Michael Slisinger & Peter Learmonth | NetApp

July 2009 | TR-3428 | Version 4.5.2

TABLE OF CONTENTS

1 EXECUTIVE SUMMARY

NetApp® storage virtualization technologies enables virtual data centers to virtualize more systems, faster and with less storage than traditional legacy storage architectures. VMware on NetApp solutions focus in the areas of: increasing data protection; reducing infrastructure costs via storage efficiencies; providing virtual data center backup and restore with simple and flexible business continuance, and instantaneous zero cost VM cloning for virtual desktops and servers.

This technical report reviews the best practices for implementing VMware® Virtual Infrastructure on NetApp fabric-attached storage (FAS) systems. NetApp has been providing advanced storage features to VMware ESX solutions since the product began shipping in 2001. During that time, NetApp has developed operational guidelines for the FAS systems and ESX Server. These techniques have been documented and are referred to as *best practices.* This technical report describes them.

Note: These practices are only recommendations, not requirements. Not following these recommendations does not affect the support provided to your implementation by NetApp and VMware. Not all recommendations apply to every scenario. NetApp believes that their customers will benefit from thinking through these recommendations before making any implementation decisions. In addition to this document, professional services are available through NetApp, VMware, and our joint partners. These services can be an attractive means to enable optimal virtual storage architecture for your virtual infrastructure.

The target audience for this paper is familiar with concepts pertaining to VMware ESX Server 3.5 and NetApp Data ONTAP® 7.X. For additional information and an overview of the unique benefits that are available when creating a virtual infrastructure on NetApp storage see TR-3515; NetApp and VMware: Building a Virtual Infrastructure from Server to Storage.

2 VMWARE STORAGE OPTIONS

VMware ESX supports three types of configuration when connecting to shared storage arrays: VMFS Datastores, NAS Datastores, and raw device mappings. The following sections review these options and summarize the unique characteristics of each storage architecture. It is assumed that customers understand that shared storage is required to enable high-value VMware features such as HA, DRS, and VMotion™. The goal of the following sections is to provide customers information to consider when designing their Virtual Infrastructure.

Note: No deployment is locked into any one of these designs; rather, VMware makes it easy to leverage all of the designs simultaneously.

2.1 VMFS DATASTORE OVERVIEW

Virtual Machine File System (VMFS) Datastores are the most common method of deploying storage in VMware environments. VMFS is a clustered file system that allows LUNs to be accessed simultaneously by multiple ESX Servers running multiple VMs.

The strengths of this technology are that it is robust, mature, and well understood. For shared infrastructures VMFS provides the VMware administrator with a fair amount of independence from the storage administrator, because once storage has been provisioned to the ESX Servers, the VMware administrator is free to use the storage as needed. With shared datastores most data management operations are performed exclusively through VMware vCenter Server.

Scalability needs to be carefully planned when deploying many virtual machines on this type of datastore. Shared datastores fill the aggregated I/O demands of many VMs and as a result this design doesn't allow a storage array to identify the I/O load generated by an individual VM. Monitoring of I/O load is best monitored within vCenter. For datasets with demanding I/O requirements it is recommended to deploy a virtual disk on a VMFS datastore that is not shared by other VMs.

NetApp can enhance VMFS datastores by providing array-based LUN level thin provisioning, production-use data deduplication, immediate zero cost datastore clones, and integrated tools such as Site Recovery Manager, SnapManager for Virtual Infrastructure, the Rapid Cloning Utility, and VMInsight; which maps virtual storage object to the physical storage infrastructure including initiators, targets, switches, LUNs and arrays.

For information about deploying VMFS see the VMware ESX and ESXi Server Configuration Guide.

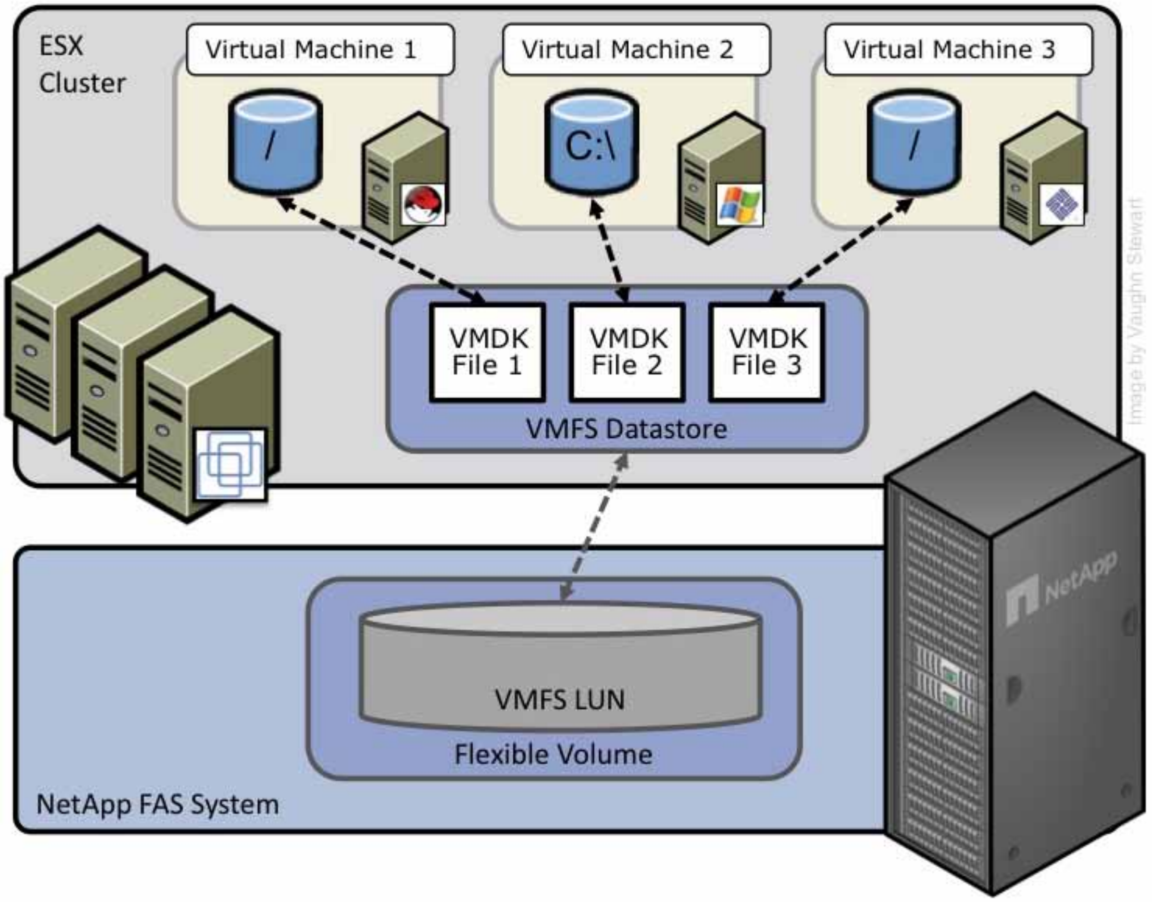


Figure 1) ESX Cluster connected to a VMFS Datastore via FC or iSCSI.

2.2 NAS DATASTORE OVERVIEW

Support for storing virtual disks (VMDKs) on a network file system (NFS) has been gaining popularity since its introduction with VMware ESX 3.0 in June of 2006.

Like VMFS NFS allows datastores to be concurrently accessed by all of the nodes in an ESX cluster which runs multiple VMs. The strengths of this solution include those of VMFS Datastores where once storage has been provisioned to the ESX Servers the VMware administrator is free to use the storage as needed. Additional benefits of NFS include the lowest per-port costs (as compared to Fibre Channel solutions), high performance, and storage savings provided by VMware thin provisioning, which is the default format for VMDKs created on NFS.

NFS Datastores integrates VMware virtualization technology with NetApp's advanced data management and storage virtualization features of WAFL. Examples of this level of transparency include production-use data deduplication, immediate zero cost VM and datastore clones, array-based thin provisioning, and direct access to array-based Snapshot™ copies. NetApp also provides integrated tools such as SnapManager for Virtual Infrastructure, and the Rapid Cloning Utility.

Customers looking for a business continuance solution with this design should be aware that at present there is no support for NFS with VMware Site Recovery Manager (version 1 update 1). Customers will have to continue to use manual DR practices with NFS solutions until a future Site Recovery Manager update is released.

Figure 2 shows an example of this configuration. Note that the storage layout appears much like that of a VMFS Datastore, yet each virtual disk file has its own I/O queue directly managed by the NetApp FAS system. For more information about storing VMDK files on NFS, see the VMware ESX and ESXi Server Configuration Guide.

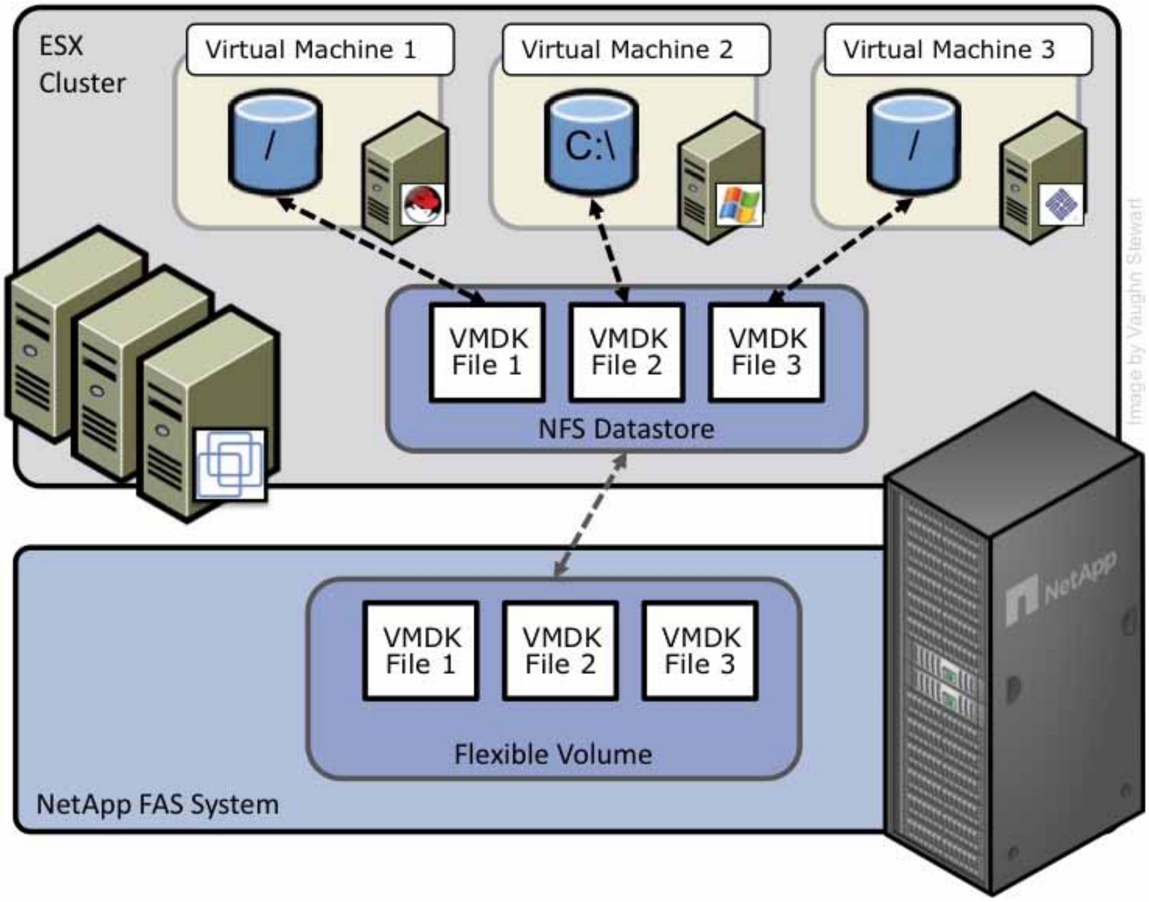


Figure 2) ESX Cluster connected to a NFS datastore.

2.3 RAW DEVICE MAPPING OVERVIEW

RDMs provide VMs with direct access to LUNs for specific use cases such as Physical to Virtual (P2V) clustering or storage vendor management tools. Note: RDMs are not used to create shared datastores.

In this design, ESX acts as a connection proxy between the VM and the storage array. The core strength of this solution is support for virtual machine and physical-to-virtual-machine host-based clustering, such as Microsoft® Cluster Server (MSCS). In addition, RDMs provide high individual disk I/O performance; easy disk performance measurement from a storage array.

NetApp can enhances the use of RDMs by providing array-based LUN level thin provisioning, production-use data deduplication, advanced integration components such as SnapDrive®, VM granular snapshots, and FlexClone® zero cost cloning of RDM based datasets.

VMware clusters have a limited to the number of LUNs which can be address and connected. Customer interested in this design should be aware such limit and consider RDMs only for use case scenarios and not as the predominant mode of storage connectivity to a VM. Figure 3 shows an example of this configuration. Note that each virtual disk file has a direct I/O to a dedicated LUN. This storage model is analogous to providing SAN storage to a physical server, except for the storage controller bandwidth, which is shared.

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 some limitations around MSCS and VMware snapshot use case scenarios. For more information about raw device mappings over Fibre Channel and iSCSI, see the VMware ESX and ESXi Server Configuration Guide.

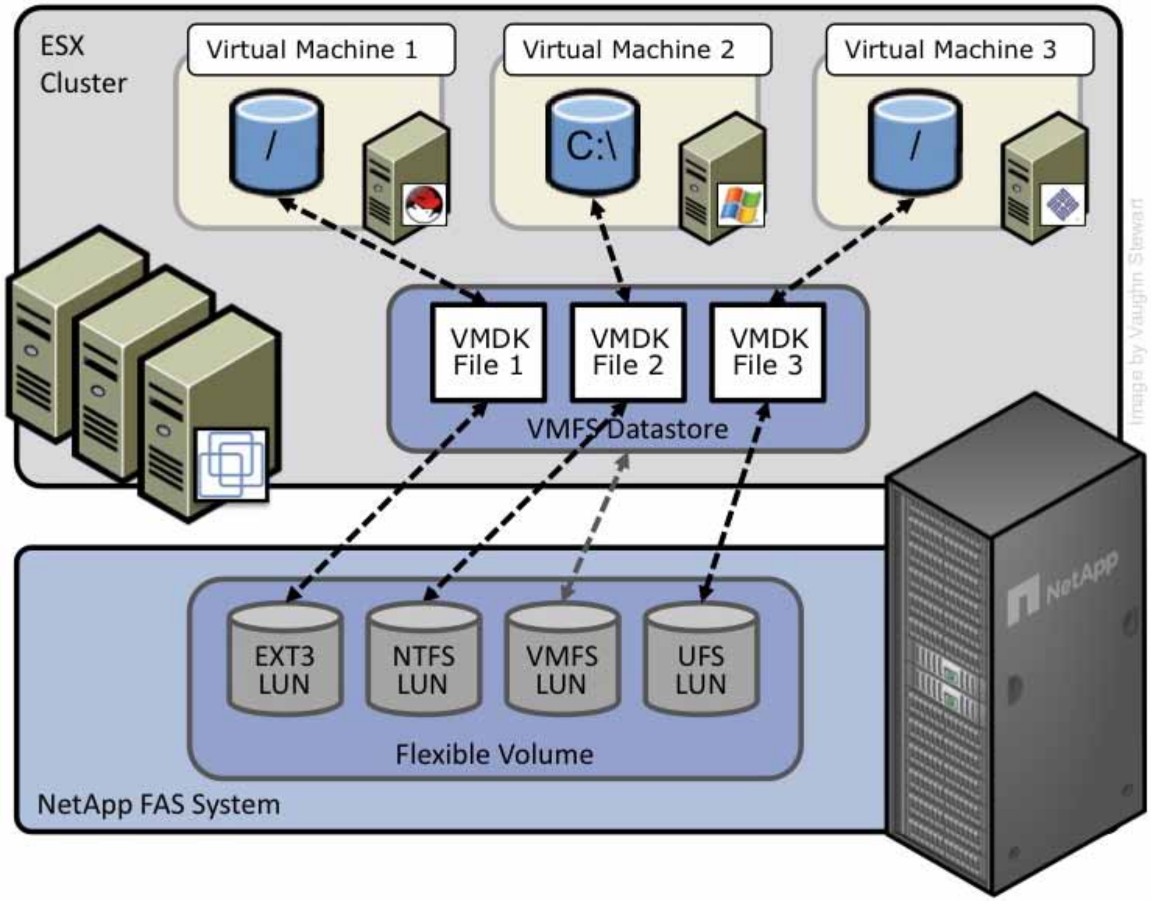


Figure 3) ESX Cluster with VMs connected to RDM LUNs via FC or iSCSI.

2.4 DATASTORE COMPARISON TABLE

Differentiating what is available with each type of Datastore and storage protocol can require considering many points. The following table compares the features available with each storage option. A similar chart for VMware is available in the VMware ESX and ESXi Server Configuration Guide

Table 1) Datastore comparison.

Capability/Feature	**FC**	**iSCSI**	**NFS**
Format	VMFS or RDM	VMFS or RDM	NetApp WAFL®
Max Datastores or LUNs	256	256	32
Max Datastore size	64TB	64TB	16TB
Recommended VMs per datastore	16 Per VMware SAN Design and Deployment Guide	16 Per VMware SAN Design and Deployment Guide	250
Available link speeds	1, 2, 4, 8 Gb	1, 10 Gb	1, 10 Gb
Backup Options			
VMDK image access	VCB	VCB	VCB, VIC File Explorer
VMDK file level access	VCB, Windows® only	VCB, Windows® only	VCB & 3rd party apps
NDMP granularity	Datastore	Datastore	Datastore or VM
VMware Feature Support			
VMotion	Yes	Yes	Yes
Storage VMotion	Yes	Yes	Experimental Support
VMware HA	Yes	Yes	Yes
DRS	Yes	Yes	Yes
VCB	Yes	Yes	Yes
MSCS within a VM	Yes, via RDM	Not supported	Not supported
Resize Datastore	Grow only	Grow only	Grow, Auto-grow, & Shrink
NetApp Integration Support			
Snapshot copies	Yes	Yes	Yes
SnapMirror®	Datastore or RDM	Datastore or RDM	Datastore or VM
SnapVault®	Datastore or RDM	Datastore or RDM	Datastore or VM
Data Deduplication	Savings in the array	Savings in the array	Savings in the datastore
Thin provisioning	Datastore or RDM	Datastore or RDM	Datastore
FlexClone	Datastore or RDM	Datastore or RDM	VM and Datastore
MultiStore®	No	Yes	Yes
SANscreen	Yes plus VMInsight	Yes plus VMInsight	Yes
Open Systems SnapVault	Yes	Yes	Yes

3 NETAPP FAS CONFIGURATION AND SETUP

3.1 DATA PROTECTION

RAID & DATA PROTECTION

A byproduct 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 could be catastrophic. Fortunately, VMware provides multiple technologies that enhance availability of the virtual infrastructure. These technologies include physical server clustering via VMware HA, application load balancing with DRS, and the ability to non-disruptively move running VMs and data sets between physical 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 HBAs, deploying redundant storage networking and network paths, and leveraging storage arrays with redundant controllers. A deployed storage design that meets all of 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. NetApp RAID-DP® is an advanced RAID technology that is provided as the default RAID level on all FAS 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. NetApp recommends using RAID-DP on all RAID groups that store VMware data.

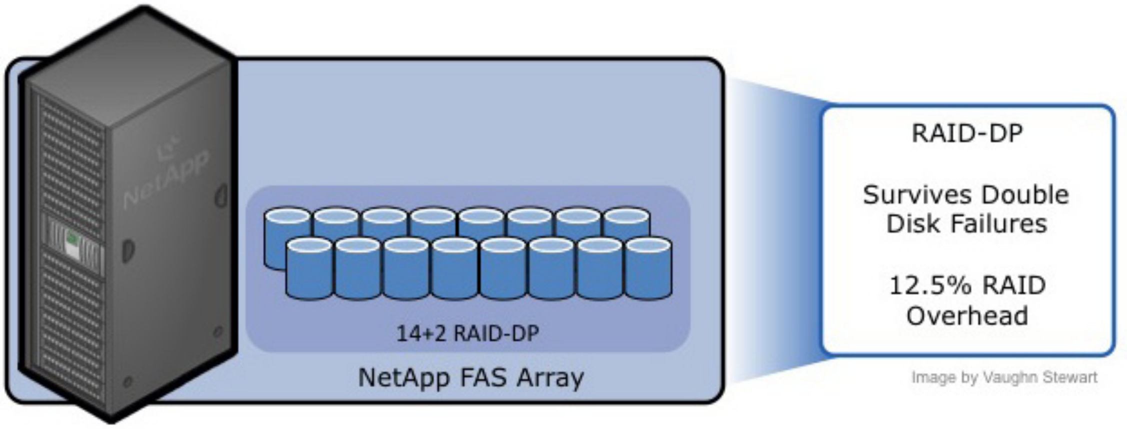


Figure 4) NetApp RAID-DP

AGGREGATES

An aggregate is NetApp's virtualization layer, which abstracts physical disks from logical data sets that are referred to as *flexible volumes.* Aggregates are the means by which the total IOPs available to all of the physical disks are pooled as a resource. This design is well suited to meet the needs of an unpredictable and mixed workload. NetApp recommends that whenever possible a small aggregate should be used as the root aggregate. This aggregate stores the files required for running and providing GUI management tools for the FAS system. The remaining storage should be placed 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 IO requests. On smaller FAS arrays, it may not be practical to have more than a single aggregate, due to the restricted number of disk drives on the system. In these cases, it is acceptable to have only a single aggregate.

3.2 NETAPP ARRAY CONFIGURATION

NETAPP HA MODE FOR FC CONFIGURAITONS

NetApp HA arrays ship confugred with an option known as cfmode, which controls the behavior of the system's Fibre Channel ports if a controller failover occurs. This setting should be set as Single System Image. This is the default and recommended setting.

If you are deploying ESX on an older HA array with FC, then please ensure the cfmode is set to SSI. To verify the current cfmode, follow these steps.

1	Connect to the FAS system console (via either SSH, Telnet, or Console connection).
2	Enter `fcp show cfmode`.

To set the cfmode, follow these steps. Note: changing cfmodes with existing VMware environments may impact you access to existing datastores if the version of ESX is less than 3.5. If this is the case please see NetApp KB33990 before implementing this change.

1	Connect to the FAS system console (via either SSH, Telnet, or Console connection).
2	If cfmode needs to be changed, enter FC `set cfmode single_image`

For more information about the different cfmodes available and the impact of changing a cfmode, see section 8 in the Data ONTAP Block Management Guide.

3.3 NETAPP STORAGE CONFIGURATION

FLEXIBLE VOLUMES

Flexible volumes contain either LUNs or virtual disk files that are accessed by VMware ESX Servers.

NetApp recommends 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 FAS array. This mapping model also makes it easy to implement Snapshot backups and SnapMirror replication policies at the Datastore level, because NetApp implements these storage side features at the flexible volume level.

SNAPSHOT RESERVE

NetApp flexible volumes should be configured with the snap reserve set to 0 and the default Snapshot schedule disabled. All NetApp Snapshot copies must be coordinated with the ESX Servers for data consistency. NetApp Snapshot copies are discussed in section 10.1, "Implementing Snapshot Copies." To set the volume options for Snapshot copies to the recommended setting, enter the following commands in the FAS system console.

1.	Log into the NetApp console.
2.	Set the volume Snapshot schedule: `snap sched <vol-name> 0 0 0`
3.	Set the volume Snapshot reserve: c `snap reserve <vol-name> 0`

LUNS

LUNs are units of storage provisioned from a FAS system directly to the ESX Servers. The 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. For more information, see the VMware Storage/SAN Compatibility Guide for ESX Server 3.5 and ESX Server 3i

STORAGE NAMING CONVENTIONS

NetApp 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.

NetApp suggests the following naming guidelines:

- **FlexVol name:** Should match the name of the Datastore.
- **LUN name for VMFS:** Should match the name of the Datastore.

  LUN name for RDMs: Should include both the hostname and volume label or name.

4 ESX FC & ISCSI STORAGE CONFIGURATION

4.1 LUN SIZING FOR VMFS DATASTORES

VMFS Datastores offer the simplest method of provisioning storage; however, it's necessary to balance the number of Datastores to be managed with the possibility of overloading very large Datastores with too many VMs. In the latter case, 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. It is common for large VMFS Datastores to have hit their IO performance limit before their capacity limit has been reached. Advanced storage technologies like thin provisioning can return provisioned but unused storage back to the FAS 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. Although there is no definitive recommendation, a commonly deployed size for a VMFS Datastore is somewhere between 300 and 700GB. The maximum supported LUN size is 2TB. For more information, see the *VMware Storage/SAN Compatibility Guide for ESX Server 3.5 and ESX Server 3i.*

4.2 CLUSTER SIZING CONSIDERATIONS WHEN USING RDMS

A VMware cluster is collectively bound to the same limits of an indivlidual ESX server. Currently the maximum number of LUNs which can be connected to a cluster is 256 LUNs. This limitation typically comes into consideration only with RDM-based deployments. With RDMs, you must plan for the use of a VFMS datastore to store virtual machine configuration files. Note that the VMDK definition file associated with RDMs is reported to be the same size as the LUN. This is the default behavior in the vCenter Server; the actual VMDK definition file consumes only a few megabytes of disk storage.

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

> 254 / (number of RDMS per VM) / (planned number of VMs per ESX host) = number of ESX nodes in a data center

RDM EXAMPLE

Say you plan to run 2 RDMs per VM with 20 VMs per ESX Server, then formula would be:

> 254/2/20 = 6.35 rounded up = 7 ESX Servers required in the cluster

4.3 VMWARE PATCH FOR FIBRE CHANNEL AND ISCSI

VMware has identified a bug which impacts the stability of hosts connect via FC & iSCSI, and has released patch ESX350-200808402-BG, which addresses this bug. At present time, this patch applies to ESX releases prior to ESX version 3.5 update 3.

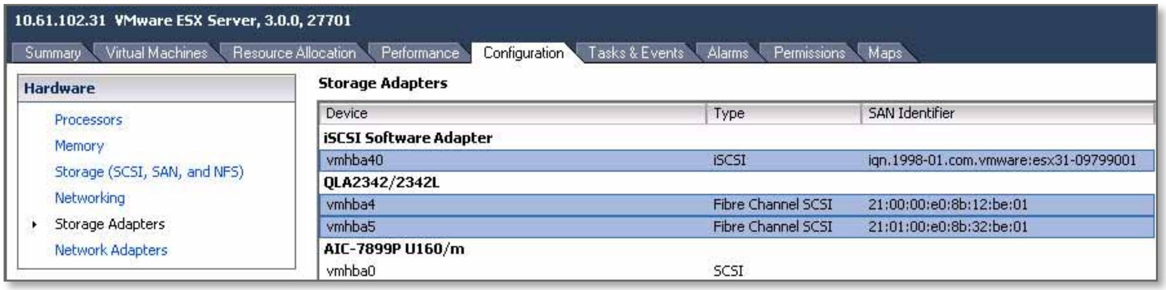
VMware patches can be found at: http://support.vmware.com/selfsupport/download/

4.4 FIBRE CHANNEL AND ISCSI LUN PROVISIONING

When provisioning LUNs for access via FC or iSCSI, the LUNs must be masked so that the appropriate hosts can connect only to them. With a NetApp FAS system, LUN masking is handled by the creation of initiator groups. NetApp recommends creating an igroup for each VMware cluster. NetApp also recommends 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 igroups by reducing the total number created. It also means that all 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 ESX Servers in the VMware cluster.

Note: If a cluster will use both Fibre Channel and iSCSI protocols, separate igroups must be created for Fibre Channel and iSCSI.

For assistance in identifying the WWPN or IQN of the ESX Server, select Storage Adapters on the Configuration tab for each ESX Server in vCenter Server and refer to the SAN Identifier column.



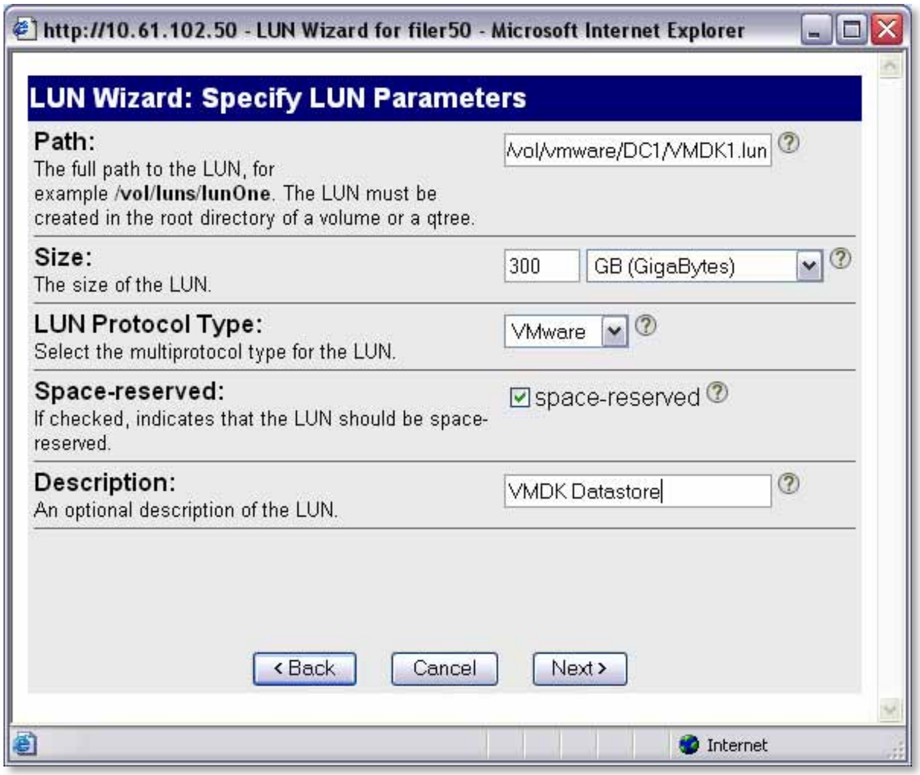
Storage Adapters

Device	Type	SAN Identifier
iSCSI Software Adapter		
vmhba40	iSCSI	iqn.1998-01.com.vmware:esx31-09799001
QLA2342/2342L		
vmhba4	Fibre Channel SCSI	21:00:00:e0:8b:12:be:01
vmhba5	Fibre Channel SCSI	21:01:00:e0:8b:32:be:01
AIC-7899P U160/m		
vmhba0	SCSI	

Figure 5) Identifying WWPN and IQN numbers in the Virtual Infrastructure client.

LUNs can be created by using the NetApp LUN Wizard in the FAS system console or by using the FilerView® GUI, as shown in the following procedure.

1.	Log in to FilerView.
2.	Select LUNs.
3.	Select Wizard.
4.	In the Wizard window, click Next.
5.	Enter the path (see Figure 6).
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.


http://10.61.102.50 - LUN Wizard for filer50 - Microsoft Internet Explorer
< Back
Cancel
Next >
Internet

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. — /vol/vmware/DC1/VMDK1.lun

Size: The size of the LUN. — 300 GB (GigaBytes)

LUN Protocol Type: Select the multiprotocol type for the LUN. — VMware

Space-reserved: If checked, indicates that the LUN should be space-reserved. — ☑ space-reserved

Description: An optional description of the LUN. — VMDK Datastore

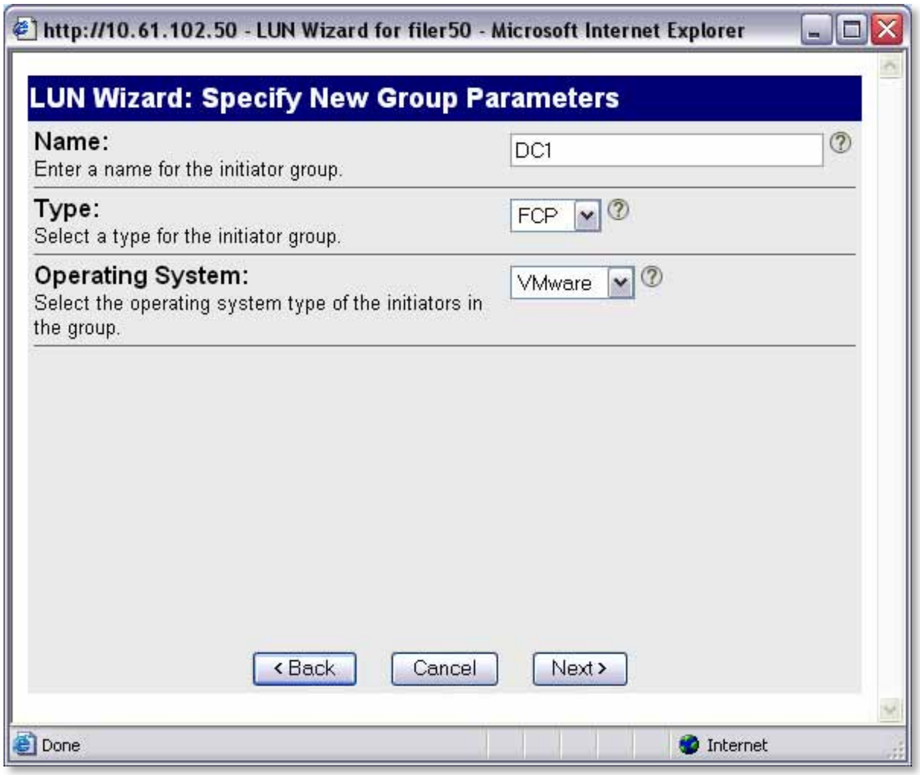
Figure 6) NetApp LUN Wizard.

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 ESX Server expects a LUN ID to be the same on every node in an ESX cluster. Therefore NetApp recommends creating a single igroup for each cluster rather than for each ESX Server.

To configure LUN masking on a LUN created in the FilerView GUI, follow these steps.

1.	Select Add Group.
2.	Select the Use Existing Initiator Group radio button. Click Next and proceed to step 3a. Or Select the Create a New Initiator Group radio button. Click Next and proceed to step 3b.
3a.	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.
3b.	Supply the igroup parameters, including name, connectivity type (FC or iSCSI), and OS type (VMware), and then click Next (see Figure 7).
4.	For the systems that will connect to this LUN, enter the new SAN identifiers or select the known identifiers (WWPN or IQN).
5.	Click the Add Initiator button.
6.	Click Next to complete the task.



http://10.61.102.50 - LUN Wizard for filer50 - Microsoft Internet Explorer

LUN Wizard: Specify New Group Parameters

Name: DC1
Enter a name for the initiator group.

Type: FCP
Select a type for the initiator group.

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

Figure 7) Assigning an igroup to a LUN.

4.5 CONNECTING FIBRE CHANNEL DATASTORES

The Fibre Channel service is the only storage protocol that is running by default on the ESX Server. NetApp recommends that each ESX Server have two FC HBA ports available for storage path redundancy. To connect to FC LUNs provisioned on a FAS system, follow these steps.

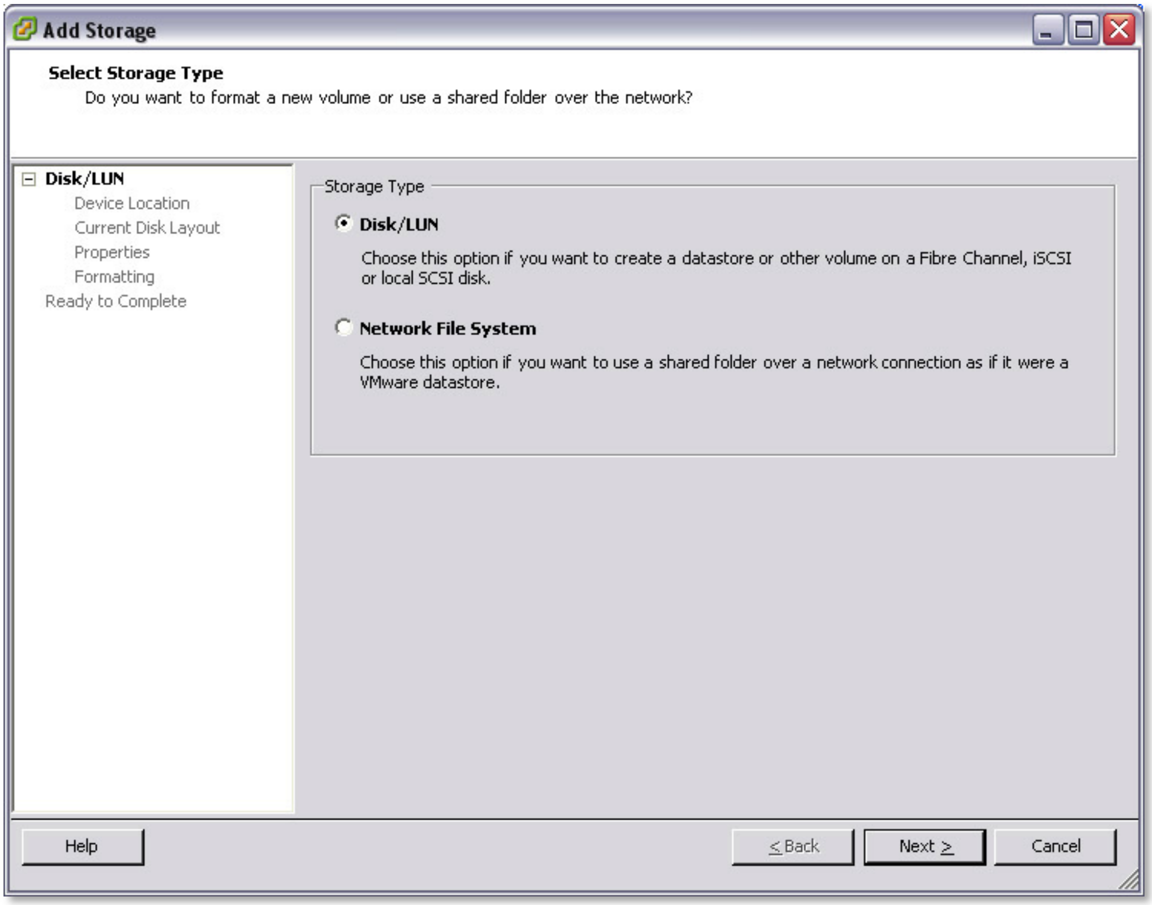
1	Open vCenter Server.
2	Select an ESX host.
3	In the right pane, select the Configuration tab.
4	In the Hardware box, select the Storage Adapters link.
5	In the upper right corner, select the Rescan link.
6	Repeat steps 1 through 5 for each ESX Server in the cluster.

Selecting Rescan forces the rescanning of all HBAs (FC and iSCSI) to discover changes in the storage available to the ESX Server.

Note: Some FC HBAs require you to scan them twice to detect new LUNs (see VMware kb1003988). After the LUNs have been identified, they can be assigned to a virtual machine as raw device mapping or provisioned to the ESX Server as a datastore.

To add a LUN as a Datastore, follow these steps.

1	Open vCenter Server.
2	Select an ESX host.
3	In the right pane, select the Configuration tab.
4	In the Hardware box, select the Storage link and then click Add Storage to open the Add Storage Wizard (see Figure 8).
5	Select the Disk/LUN radio button and click Next.
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.



Select Storage Type

Do you want to format a new volume or use a shared folder over the network?

Storage Type

● **Disk/LUN**

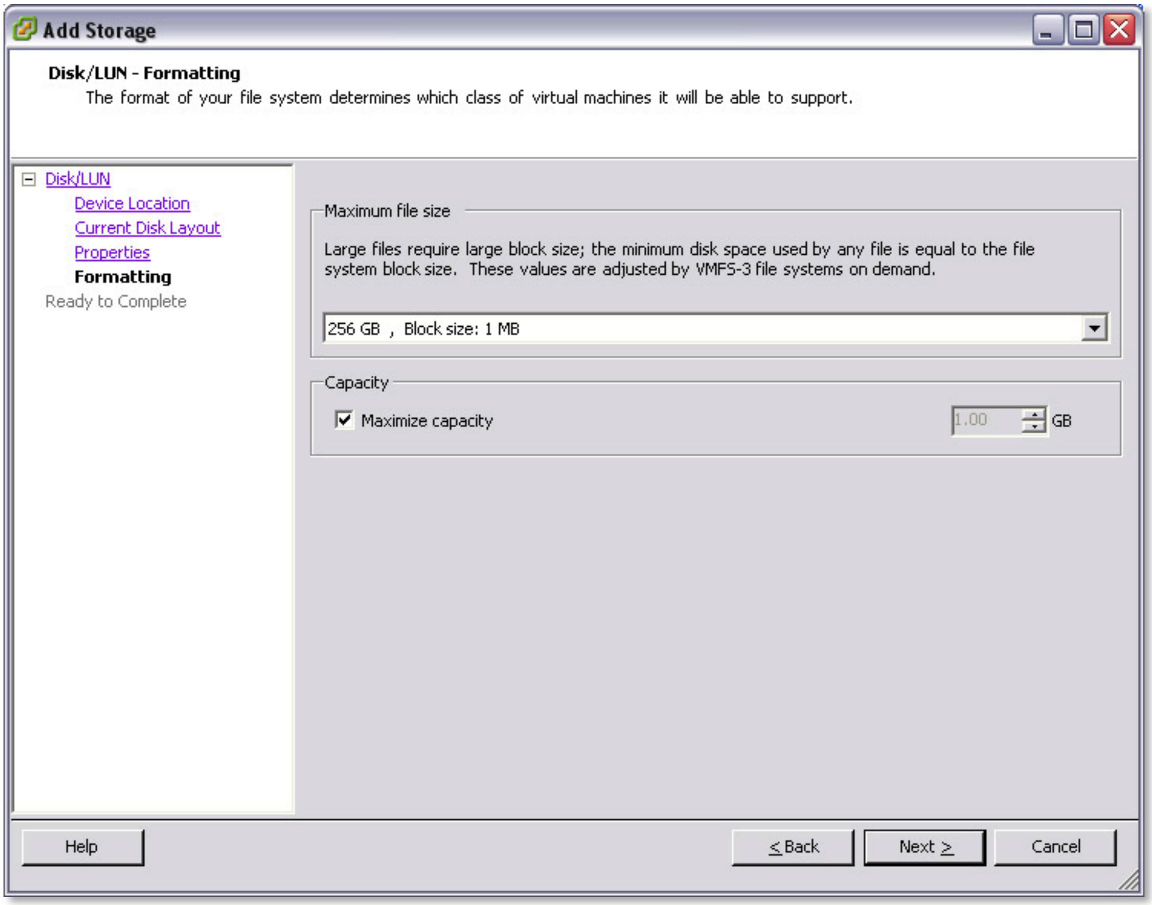
Choose this option if you want to create a datastore or other volume on a Fibre Channel, iSCSI or local SCSI disk.

○ **Network File System**

Choose this option if you want to use a shared folder over a network connection as if it were a VMware datastore.

Figure 8) VMware Add Storage wizard.

The default block size of a Virtual Machine File System is 1MB. This block size supports storing virtual disk files up to a maximum of 256GB in size. If you plan to store virtual disks larger than 256GB in the datastore, you must increase the block size to be greater than the default (see Figure 9).



Add Storage

Disk/LUN - Formatting

The format of your file system determines which class of virtual machines it will be able to support.

Maximum file size

Large files require large block size; the minimum disk space used by any file is equal to the file system block size. These values are adjusted by VMFS-3 file systems on demand.

256 GB , Block size: 1 MB

Capacity

☑ Maximize capacity 1.00 GB

Figure 9) Formatting a LUN with VMFS.

4.6 CONNECTING ISCSI DATASTORES

As a best practice NetApp recommends separating 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.

To create a second network in ESX requires one to create a second vSwitch in order to separate the traffic on to other physical NICs. The ESX Server will require a VMkernel port to be defined on the new vSwitch.

Each 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 ESX HA architectures and follows ESX HA best practices.

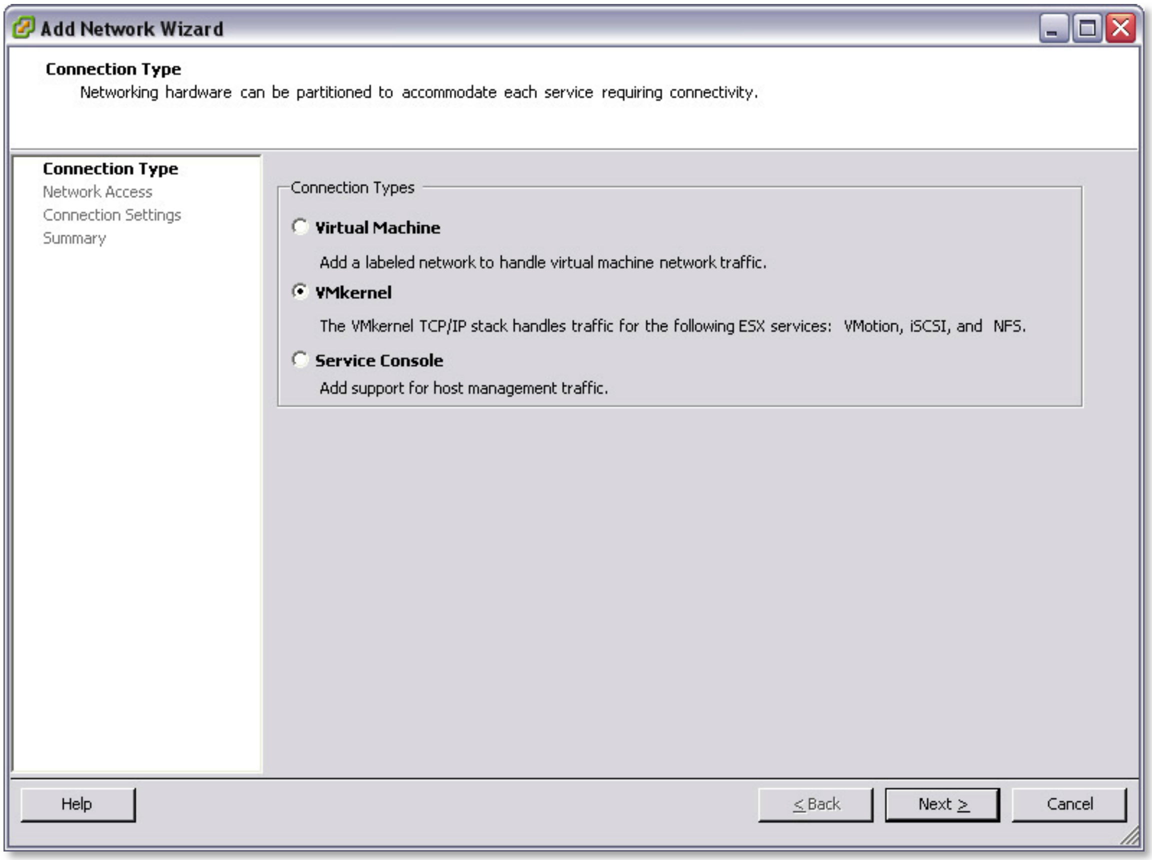
With this design it is recommened to not allow routing of data between these networks. In other word, do not define a default gateway for the iSCSI storage network. With this model iSCSI deployments will require a second service console port be defined on the VMKernel storage virtual switch within each ESX server.

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

It is recommended to disable iSCSI on NetApp interfaces which you do not want to send iSCSI traffic over.

To configure the iSCSI connectivity, follow these steps.

1	Open vCenter Server.
2	Select an ESX host.
3	In the right pane, select the Configuration tab.
4	In the Hardware box, select Networking.
5	In the upper right corner, click Add Networking to open the Add Network Wizard (see Figure 10).
6	Select the VMkernel radio button and click Next.
7	Either select an existing vSwitch or create a new one. **Note:** If a separate iSCSI network does not exist, create a new vSwitch.
8	Click Next.
9	Enter the IP address and subnet mask, click Next, and then click Finish to close the Add Network Wizard (see Figure 11).
10	Optional: A default gateway is not required for the VMkernel IP storage network. (see Figure 11).
11	In the Configuration tab, left pane, select Security Profile.
12	In the right pane, select the Properties link to open the Firewall Properties window.
13	Select the Software iSCSI Client checkbox and then click OK to close the Firewall Properties window (see Figure 12).
14	In the right pane, Hardware box, select Storage Adapters.
15	Highlight the iSCSI Adapter and click the Properties link in the Details box (see Figure 13).
16	Select the Dynamic Discovery tab in the iSCSI Initiator Properties box.
17	Click Add and enter the IP address of the iSCSI-enabled interface on the NetApp FAS system (see Figure 14).
18	For an additional layer of security, select the CHAP tab to configure CHAP authentication. NetApp recommends setting up and verifying iSCSI access before enabling CHAP authentication.



Connection Type

Networking hardware can be partitioned to accommodate each service requiring connectivity.

Connection Types

○ **Virtual Machine**

Add a labeled network to handle virtual machine network traffic.

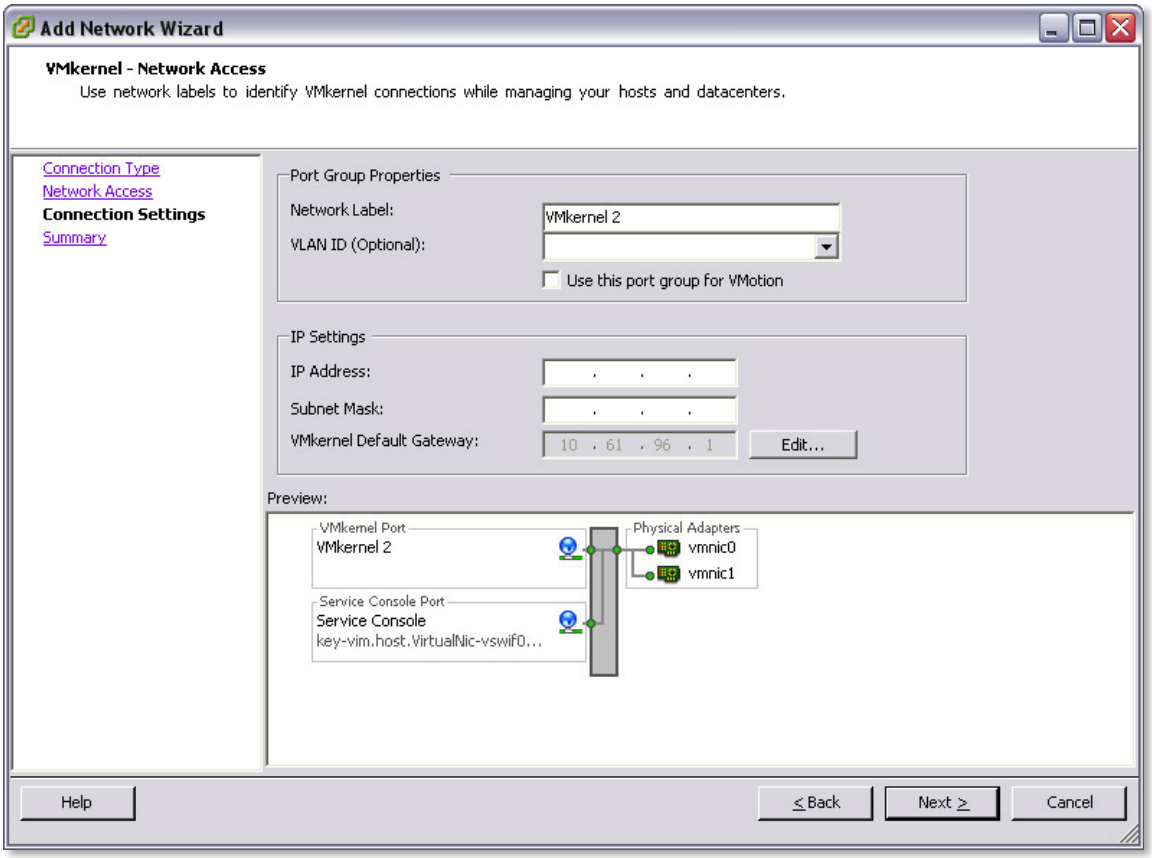
● **VMkernel**

The VMkernel TCP/IP stack handles traffic for the following ESX services: VMotion, iSCSI, and NFS.

○ **Service Console**

Add support for host management traffic.

Figure 10) Adding a VMkernel port.



Add Network Wizard

VMkernel - Network Access

Use network labels to identify VMkernel connections while managing your hosts and datacenters.

Port Group Properties

Network Label: VMkernel 2

VLAN ID (Optional):

☐ Use this port group for VMotion

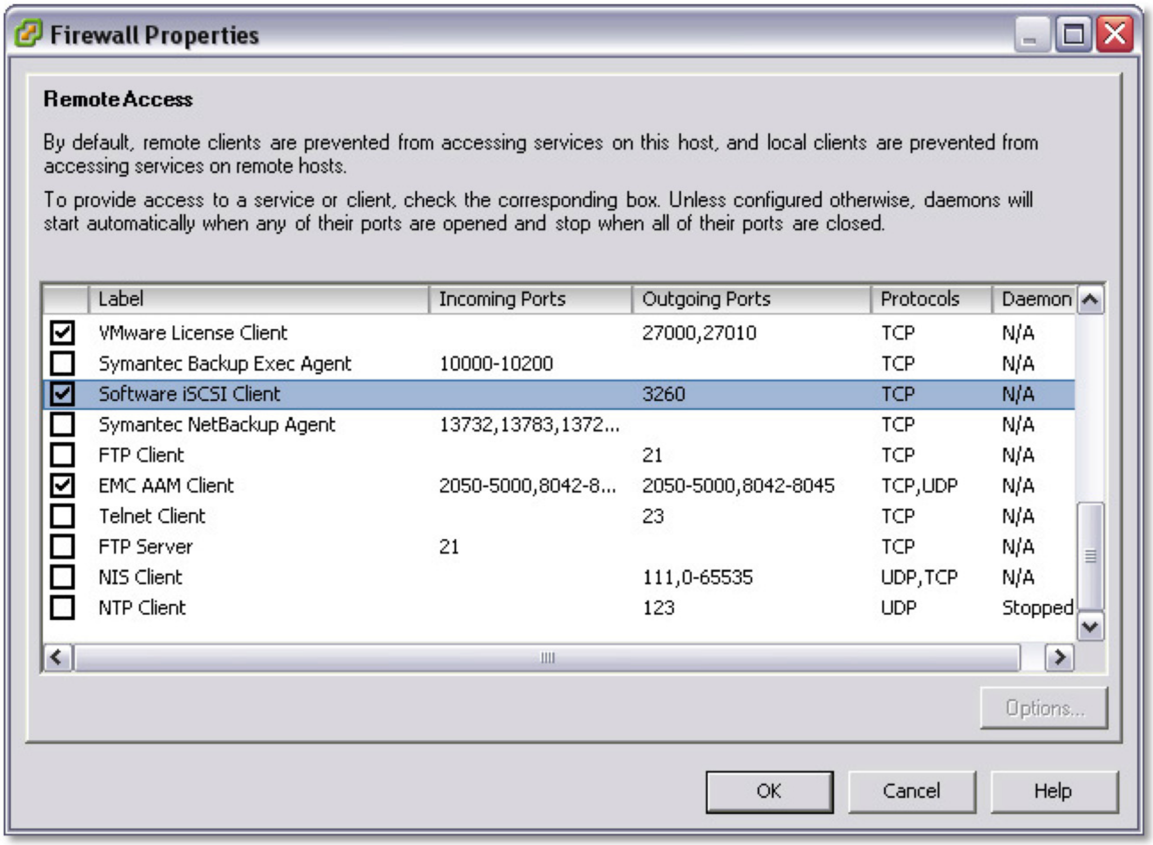
IP Settings

IP Address: . . .

Subnet Mask: . . .

VMkernel Default Gateway: 10 . 61 . 96 . 1

Figure 11) Configuring a VMkernel port.



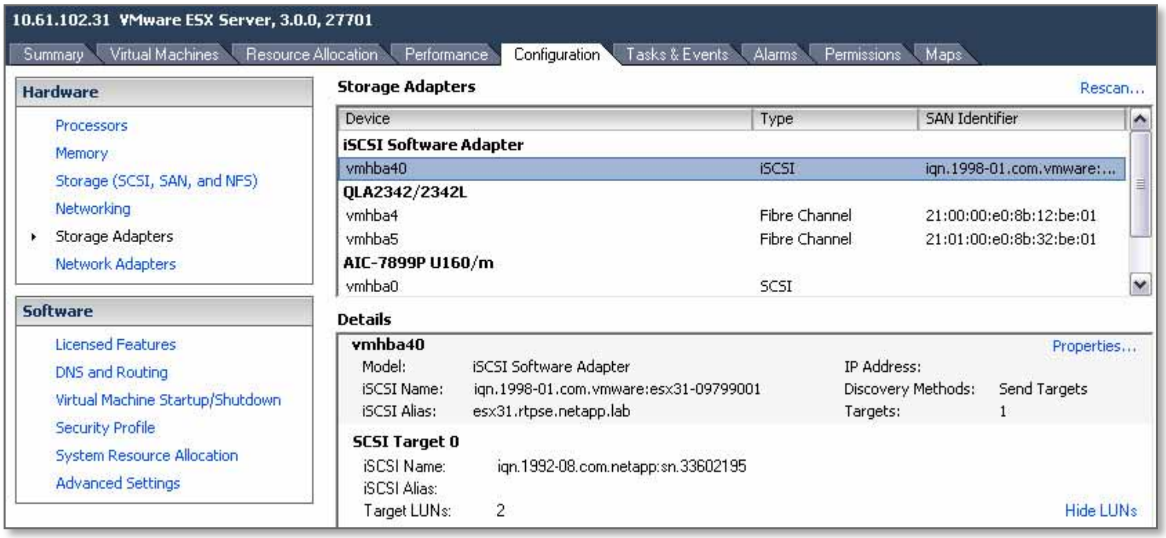
Remote Access

By default, remote clients are prevented from accessing services on this host, and local clients are prevented from accessing services on remote hosts.

To provide access to a service or client, check the corresponding box. Unless configured otherwise, daemons will start automatically when any of their ports are opened and stop when all of their ports are closed.

	Label	Incoming Ports	Outgoing Ports	Protocols	Daemon
☑	VMware License Client		27000,27010	TCP	N/A
□	Symantec Backup Exec Agent	10000-10200		TCP	N/A
☑	Software iSCSI Client		3260	TCP	N/A
□	Symantec NetBackup Agent	13732,13783,1372...		TCP	N/A
□	FTP Client		21	TCP	N/A
☑	EMC AAM Client	2050-5000,8042-8...	2050-5000,8042-8045	TCP,UDP	N/A
□	Telnet Client		23	TCP	N/A
□	FTP Server	21		TCP	N/A
□	NIS Client		111,0-65535	UDP,TCP	N/A
□	NTP Client		123	UDP	Stopped

Figure 12) Configuring the firewall in ESX.



Storage Adapters

Device	Type	SAN Identifier
iSCSI Software Adapter		
vmhba40	iSCSI	iqn.1998-01.com.vmware:...
QLA2342/2342L		
vmhba4	Fibre Channel	21:00:00:e0:8b:12:be:01
vmhba5	Fibre Channel	21:01:00:e0:8b:32:be:01
AIC-7899P U160/m		
vmhba0	SCSI	

Details

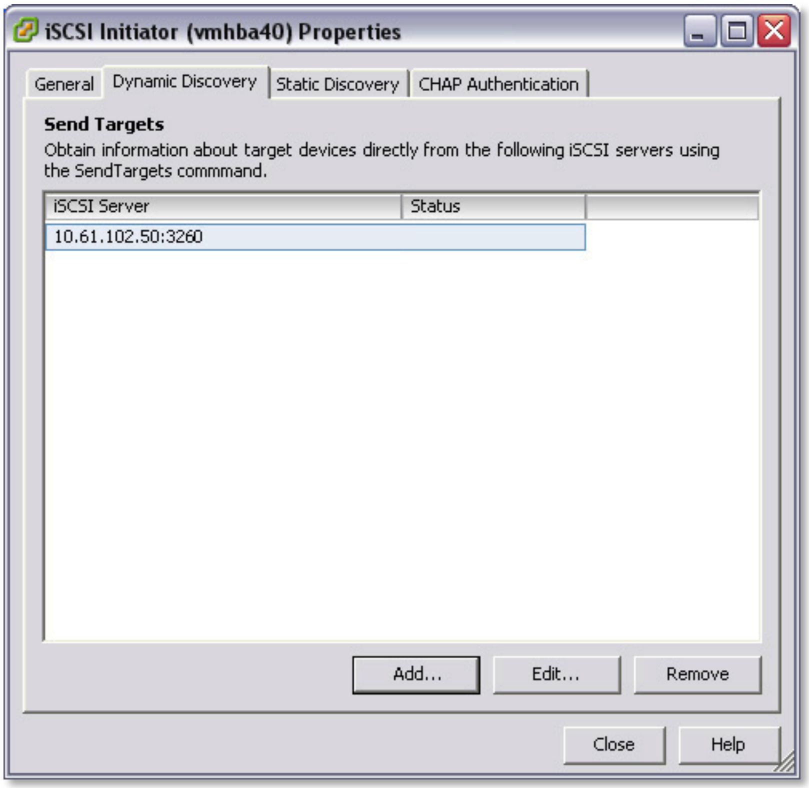
vmhba40

Model: iSCSI Software Adapter
iSCSI Name: iqn.1998-01.com.vmware:esx31-09799001
iSCSI Alias: esx31.rtpse.netapp.lab
IP Address:
Discovery Methods: Send Targets
Targets: 1

SCSI Target 0

iSCSI Name: iqn.1992-08.com.netapp:sn.33602195
iSCSI Alias:
Target LUNs: 2

Figure 13) Selecting an iSCSI initiator.



iSCSI Initiator (vmhba40) Properties

General | Dynamic Discovery | Static Discovery | CHAP Authentication

Send Targets

Obtain information about target devices directly from the following iSCSI servers using the SendTargets commmand.

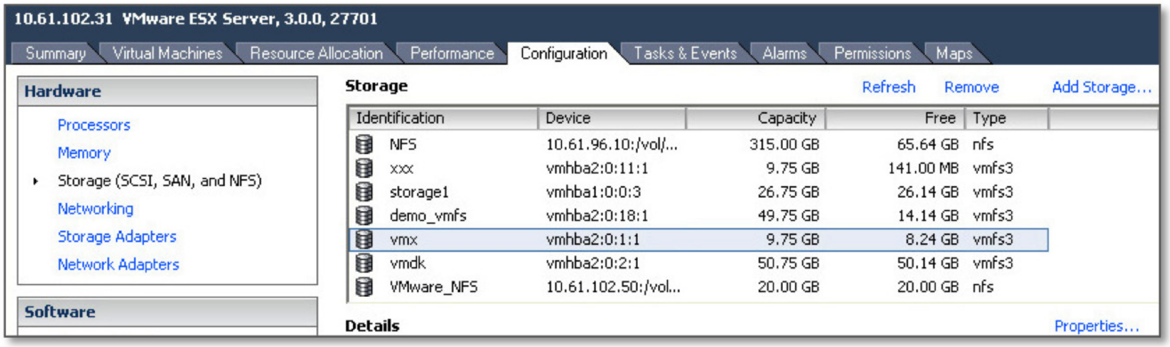
iSCSI Server	Status
10.61.102.50:3260	

Figure 14) Configuring iSCSI dynamic discovery.

4.7 VMWARE NATIVE MULTIPATHING FOR FIBRE CHANNEL AND ISCSI

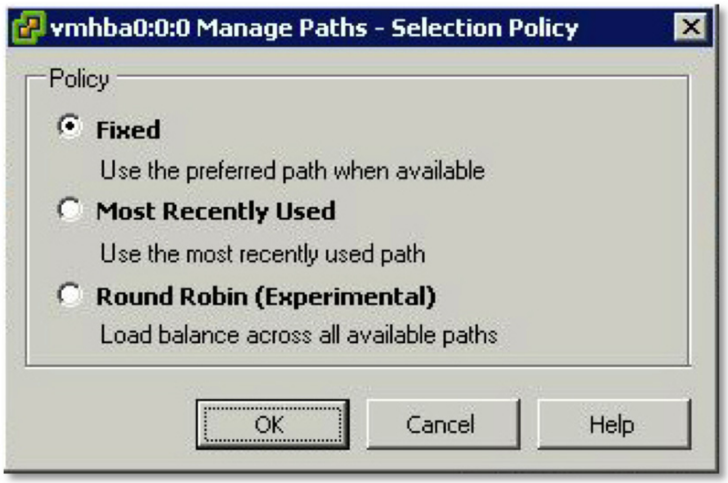
If you have implemented Single System Image cfmode, then you must configure ESX multipathing. When you are using multipathing, VMware requires the default path to be selected for each LUN connected on each ESX Server. To set the paths, follow these steps.

1	Open vCenter Server.
2	Select an ESX Server.
3	In the right pane, select the Configuration tab.
4	In the Hardware box, select Storage.
5	In the Storage box, highlight the storage and select the Properties link (see Figure 15).
6	In the Properties dialog box, click the Manage Paths button.
7	Identify the path to set as the primary active path and click the Change button (See Figure 16).
8	In the Change Path State window, select the path as Preferred and Enabled and click OK (See Figure 17).



Identification	Device	Capacity	Free	Type
NFS	10.61.96.10:/vol/...	315.00 GB	65.64 GB	nfs
xxx	vmhba2:0:11:1	9.75 GB	141.00 MB	vmfs3
storage1	vmhba1:0:0:3	26.75 GB	26.14 GB	vmfs3
demo_vmfs	vmhba2:0:18:1	49.75 GB	14.14 GB	vmfs3
vmx	vmhba2:0:1:1	9.75 GB	8.24 GB	vmfs3
vmdk	vmhba2:0:2:1	50.75 GB	50.14 GB	vmfs3
VMware_NFS	10.61.102.50:/vol...	20.00 GB	20.00 GB	nfs

Figure 15) Selecting a Datastore.

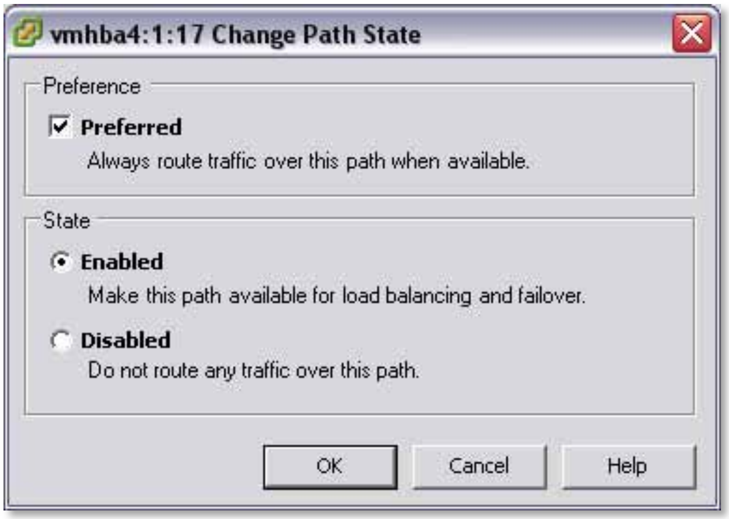


● **Fixed**
Use the preferred path when available

○ **Most Recently Used**
Use the most recently used path

○ **Round Robin (Experimental)**
Load balance across all available paths

Figure 16) VMware Manage Paths dialog box.



☑ **Preferred**
Always route traffic over this path when available.

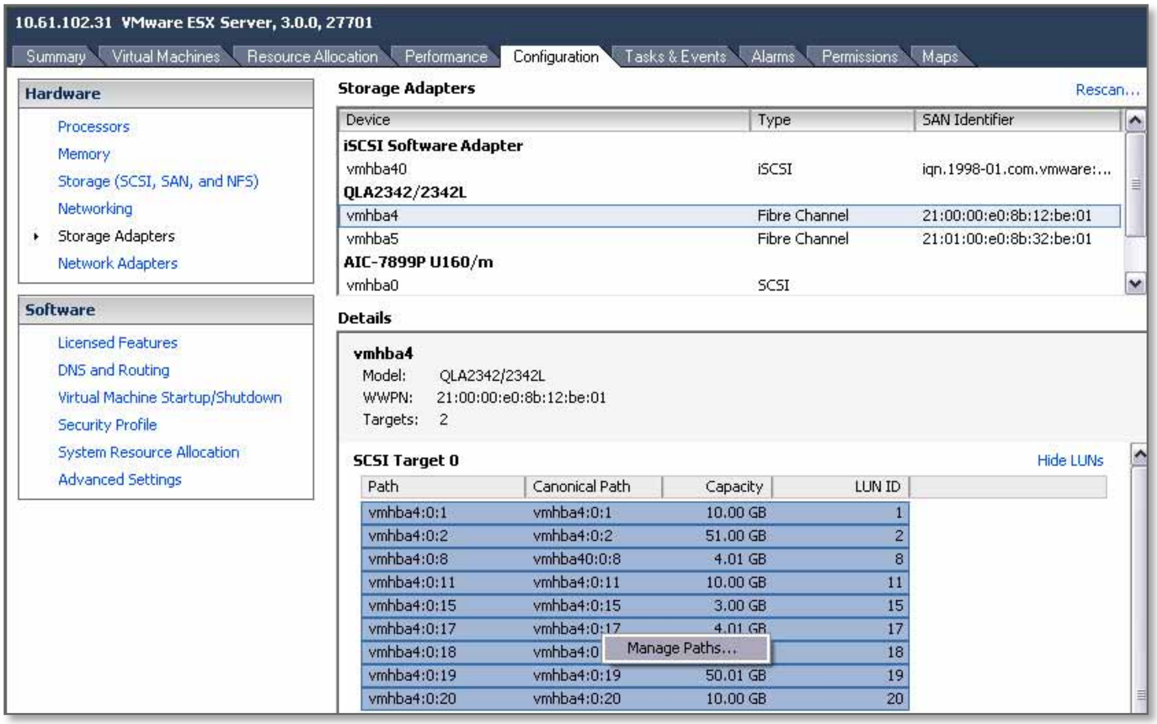
● **Enabled**
Make this path available for load balancing and failover.

○ **Disabled**
Do not route any traffic over this path.

Figure 17) Setting path preferences.

An alternative method for setting the preferred path for multiple LUNs is available in vCenter Server. To set the path, follow these steps.

1	Open vCenter Server.
2	Select an ESX Server.
3	In the right pane, select the Configuration tab.
4	In the Hardware box, select Storage Adapters.
5	In the Storage Adapters pane, select a host bus adapter.
6	Highlight all of the LUNs to configure.
7	Right-click the highlighted LUNs and select Manage Paths (see Figure 18).
8	In the Manage Paths window, set the multipathing policy and preferred path for all of the highlighted LUNs (see Figure 19).



Storage Adapters

Device	Type	SAN Identifier
iSCSI Software Adapter		
vmhba40	iSCSI	iqn.1998-01.com.vmware:...
QLA2342/2342L		
vmhba4	Fibre Channel	21:00:00:e0:8b:12:be:01
vmhba5	Fibre Channel	21:01:00:e0:8b:32:be:01
AIC-7899P U160/m		
vmhba0	SCSI	

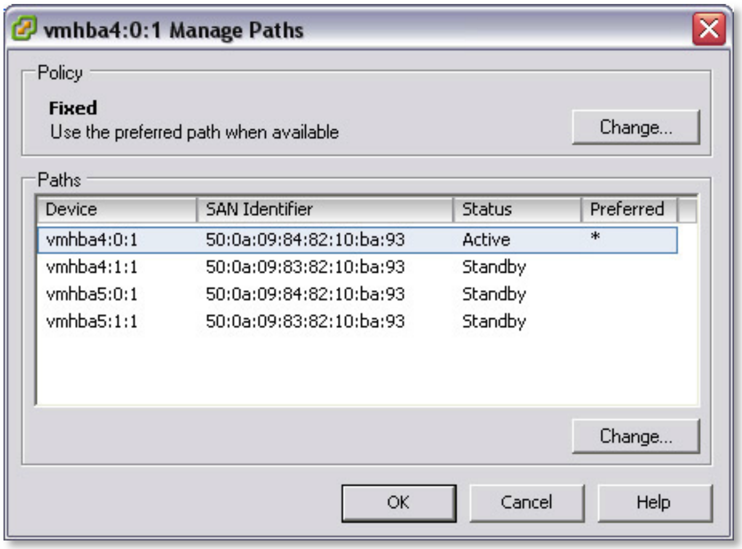
Details

vmhba4
Model: QLA2342/2342L
WWPN: 21:00:00:e0:8b:12:be:01
Targets: 2

SCSI Target 0

Path	Canonical Path	Capacity	LUN ID
vmhba4:0:1	vmhba4:0:1	10.00 GB	1
vmhba4:0:2	vmhba4:0:2	51.00 GB	2
vmhba4:0:8	vmhba40:0:8	4.01 GB	8
vmhba4:0:11	vmhba4:0:11	10.00 GB	11
vmhba4:0:15	vmhba4:0:15	3.00 GB	15
vmhba4:0:17	vmhba4:0:17	4.01 GB	17
vmhba4:0:18	vmhba4:0		18
vmhba4:0:19	vmhba4:0:19	50.01 GB	19
vmhba4:0:20	vmhba4:0:20	10.00 GB	20

Figure 18) Bulk selecting SCSI targets to set their properties.



vmhba4:0:1 Manage Paths

Policy

Fixed

Use the preferred path when available

Device	SAN Identifier	Status	Preferred
vmhba4:0:1	50:0a:09:84:82:10:ba:93	Active	*
vmhba4:1:1	50:0a:09:83:82:10:ba:93	Standby	
vmhba5:0:1	50:0a:09:84:82:10:ba:93	Standby	
vmhba5:1:1	50:0a:09:83:82:10:ba:93	Standby	

Figure 19) Setting a preferred path to a target.

5 NFS STORAGE RECOMMENDATIONS

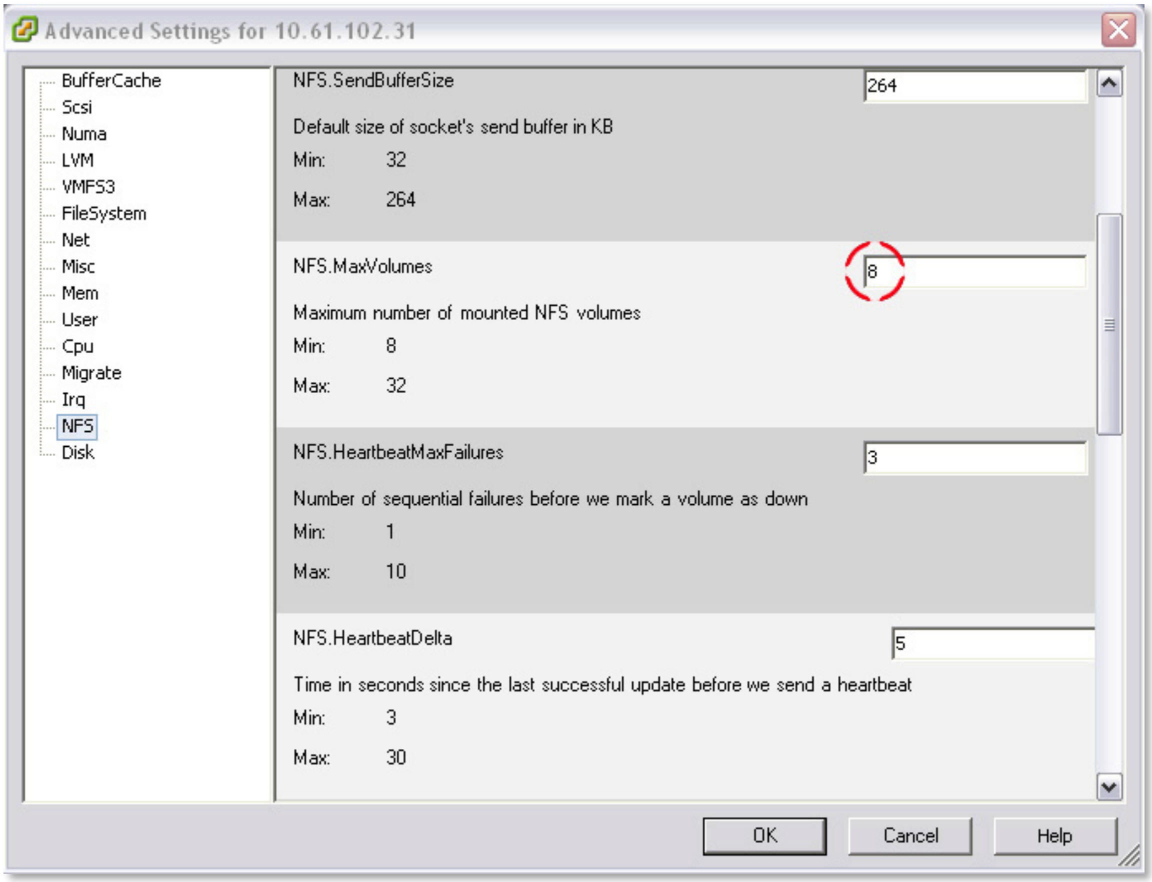
5.1 NFS DATASTORE LIMITS

By default, VMware ESX allows 8 NFS datastores; however, this limit can be increased to 32 in order to meet the needs as the Virtual Infrastructure grows. While the maximum number of NFS datastores (32) is significantly less than what is available with VMFS datastores (256) this difference is offset by the density available to NetApp NFS datastores.

In order to ensure availability NetApp recommends that you increase the maximum number of datastores available when deploying an ESX Server as preconfiguring this setting ensures that NFS datastores can be dynamiacly added at any time without disruption or effort.

To make this change, follow these steps from within the Virtual Infrastructure client. For more information on this process, including the requirement to reboot the ESX server, please refer to VMware KB 2239.

1	Open vCenter Server.
2	Select an ESX host.
3	In the right pane, select the Configuration tab.
4	In the Software box, select Advanced Configuration.
5	In the pop-up window, left pane, select NFS.
6	Change the value of NFS.MaxVolumes to 32 (see Figure 20).
7	In the pop-up window, left pane, select Net.
8	Change the value of Net.TcpIpHeapSize to 30.
9	Change the value of Net.TcpIpHeapMax to 120.
10	Repeat the steps for each ESX Server.



NFS.SendBufferSize: 264

Default size of socket's send buffer in KB

Min: 32

Max: 264

NFS.MaxVolumes: 8

Maximum number of mounted NFS volumes

Min: 8

Max: 32

NFS.HeartbeatMaxFailures: 3

Number of sequential failures before we mark a volume as down

Min: 1

Max: 10

NFS.HeartbeatDelta: 5

Time in seconds since the last successful update before we send a heartbeat

Min: 3

Max: 30

Figure 20) Increasing the maximum number of NFS Datastores.

5.2 FILE SYSTEM SECURITY

NetApp storage arrays allow customers 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 additon, security style can also be made on a file-by-file basis using the MIXED permissions setting

For VMware deployments it is highly recommended to set the security style of all datastores to UNIX. The secutiry setting of the root volume will be the security setting when a new volume is created.

It is common for customers who run VMware on NFS to want to access their datastores from Windows systems in order to complete administrative functions. With this use case in mind set the volume security style to UNIX and ensure that the FAS usermapping is setup correctly in order to enable windows user access to this data. For more information on this subject please review the section File Sharing Between NFS & CIFS in the ONTAP File Access and Protocol Management Guide.

If you need to change the file system security type follow these steps.

1	Log in to the NetApp console.
2	From the storage appliance console, run `vol options <vol-name> no_atime_update on`
3	From the storage appliance console, run `qtree security <volume path> UNIX`
4	Repeat steps 2 & 3 for each NFS accessed volume.

5.3 OPTIONAL NETWORKING RECOMMENDATION

When deploying VMware on NFS, the following setting has shown to increase maximum IO to a datastore for Gigabit Ethernet networks which are not running on low-latency switches (which is defined by running with a 5 microsecond response times).

1	Log in to the NetApp console.
2	From the storage appliance console, run `options nfs.tcp.recvwindowsize 64240`
3	Disconnect and remount each NFS datastore to each ESX Server

5.4 ESX NFS TIMEOUT SETTINGS

When connecting to NFS datastores NetApp recommends adjusting a few NFS options around connection monitoring and resiliency. These settings can be automatically set for you should you decide to install the NetApp ESX Host Utilities. The EHU is only supported with ESX, so if you are running ESXi or should you opt to not install the EHU the steps for mupdating these settingare listed below.

ESX 3.5 HOST

For optimal availability with NFS Datastores, NetApp recommends making the following changes on each ESX 3.5 host.

1	Open vCenter Server.
2	Select an ESX host.
3	In the right pane, select the Configuration tab.
4	In the Software box, select Advanced Configuration.
5	In the pop-up window, left pane, select NFS.
6	Change the value of NFS.HeartbeatFrequency to 12.
7	Change the value of NFS.HeartbeatMaxFailures to 10.
8	Repeat for each ESX Server.

ESX 3.0 HOST

For optimal availability with NFS Datastores, NetApp recommends making the following changes on each ESX 3.0.x host.

1	Open vCenter Server.
2	Select an ESX host.
3	In the right pane, select the Configuration tab.
4	In the Software box, select Advanced Configuration.
5	In the pop-up window, left pane, select NFS.
6	Change the value of NFS.HeartbeatFrequency to 5 from 9.
7	Change the value of NFS.HeartbeatMaxFailures to 25 from 3.
8	Do not change the value for NFS.HeartbeatTimeout (the default is 5).
9	Repeat for each ESX Server.

5.5 NFS STORAGE NETWORK BEST PRACTICE

As a best practice NetApp recommends separating 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.

To create a second network in ESX requires one to create a second vSwitch in order to separate the traffic on to other physical NICs. The ESX Server will require a VMkernel port to be defined on the new vSwitch.

Each 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 ESX HA architectures and follows ESX HA best practices.

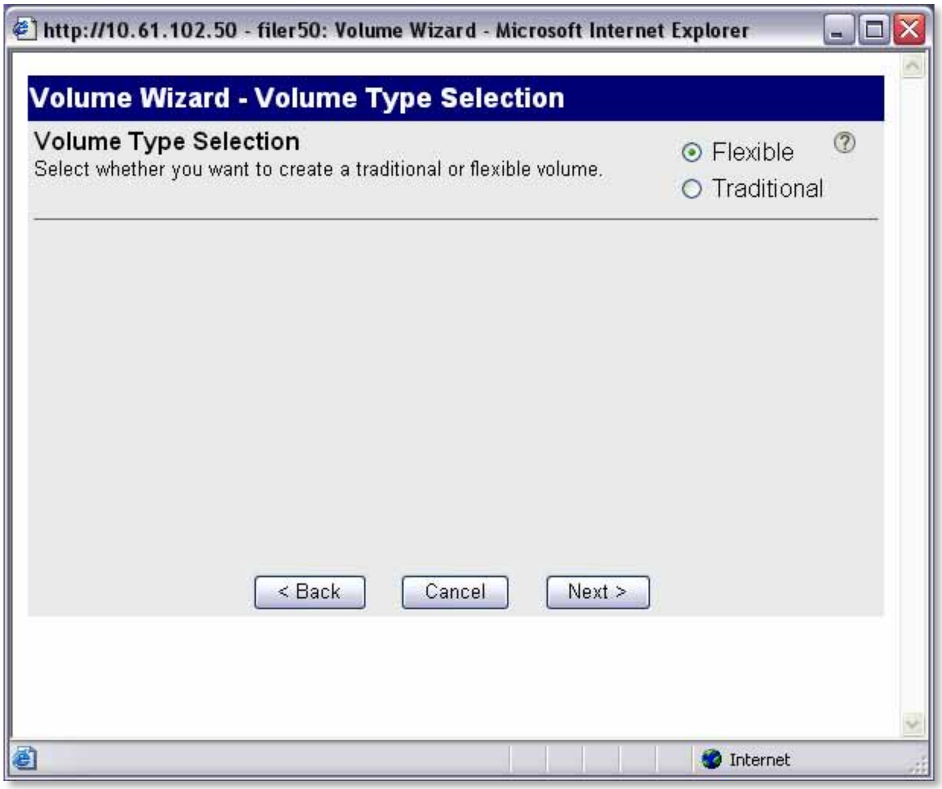
With this design it is recommened to not allow routing of data between these networks. In other word, do not define a default gateway for the NFS storage network. With this model NFS deployments will require a second service console port be defined on the VMKernel storage virtual switch within each ESX server.

IP storage network, or VMkernel, connectivity can be verified by the use of the vmkping command. With NFS connected datastores the syntax to test connectivity is vmkping <NFS IP address>.

5.6 CONNECTING NFS DATASTORES

TO create a file system for use as an NFS Datastore, follow these steps.

1	Open FilerView (http://filer/na_admin).
2	Select Volumes.
3	Select Add to open the Volume Wizard (see Figure 21). Complete the Wizard.
4	From the FilerView menu, select NFS.
5	Select Add Export to open the NFS Export Wizard (see Figure 22). Complete the wizard for the newly created file system, granting read/write and root access to the VMkernel address of all ESX hosts that will connect to the exported file system.
6	Open vCenter Server.
7	Select an ESX host.
8	In the right pane, 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 23).
11	Select the Network File System radio button and click Next.
12	Enter a name for the storage appliance, export, and Datastore, then click Next (see Figure 24).
13	Click Finish.



http://10.61.102.50 - filer50: Volume Wizard - Microsoft Internet Explorer

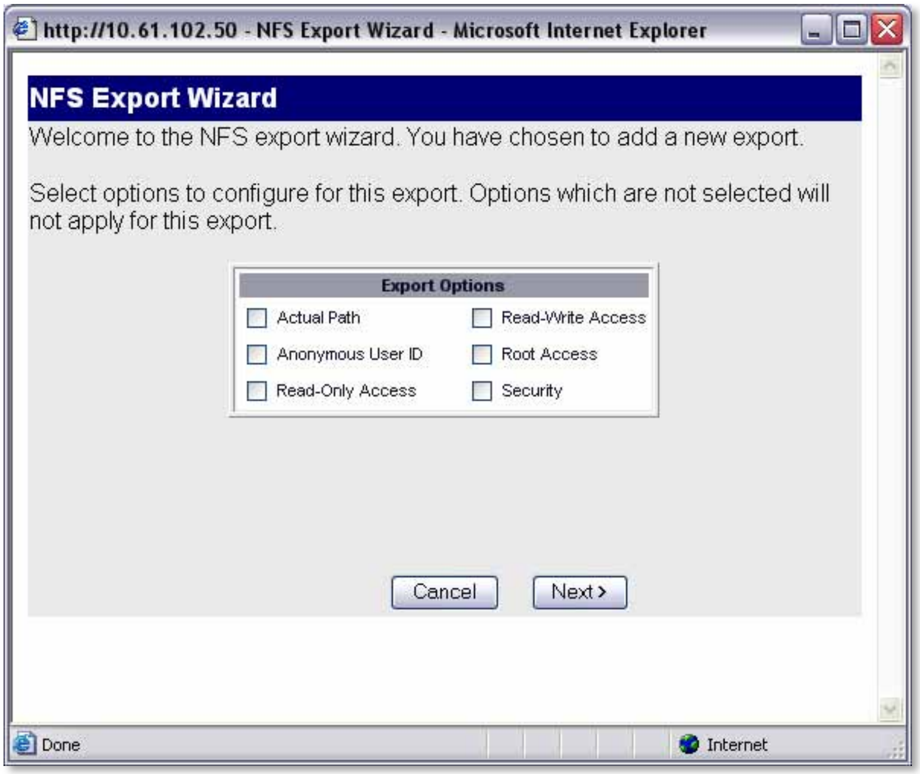
Volume Wizard - Volume Type Selection

Volume Type Selection

Select whether you want to create a traditional or flexible volume.

● Flexible
○ Traditional

Figure 21) NetApp Volume Wizard.



http://10.61.102.50 - NFS Export Wizard - Microsoft Internet Explorer

NFS Export Wizard

Welcome to the NFS export wizard. You have chosen to add a new export.

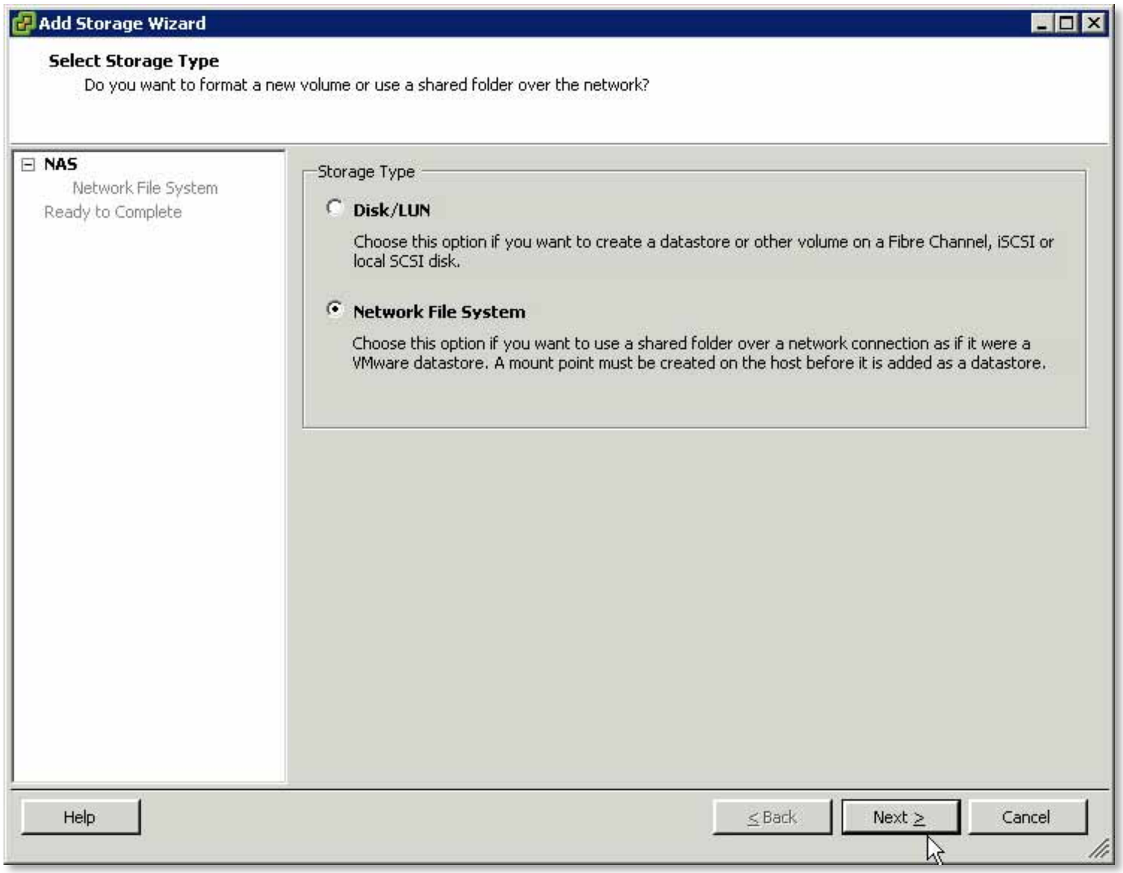
Select options to configure for this export. Options which are not selected will not apply for this export.

Export Options

□ Actual Path	□ Read-Write Access
□ Anonymous User ID	□ Root Access
□ Read-Only Access	□ Security

Cancel
Next >
Done
Internet

Figure 22) NetApp NFS Export Wizard.



Select Storage Type

Do you want to format a new volume or use a shared folder over the network?

Storage Type

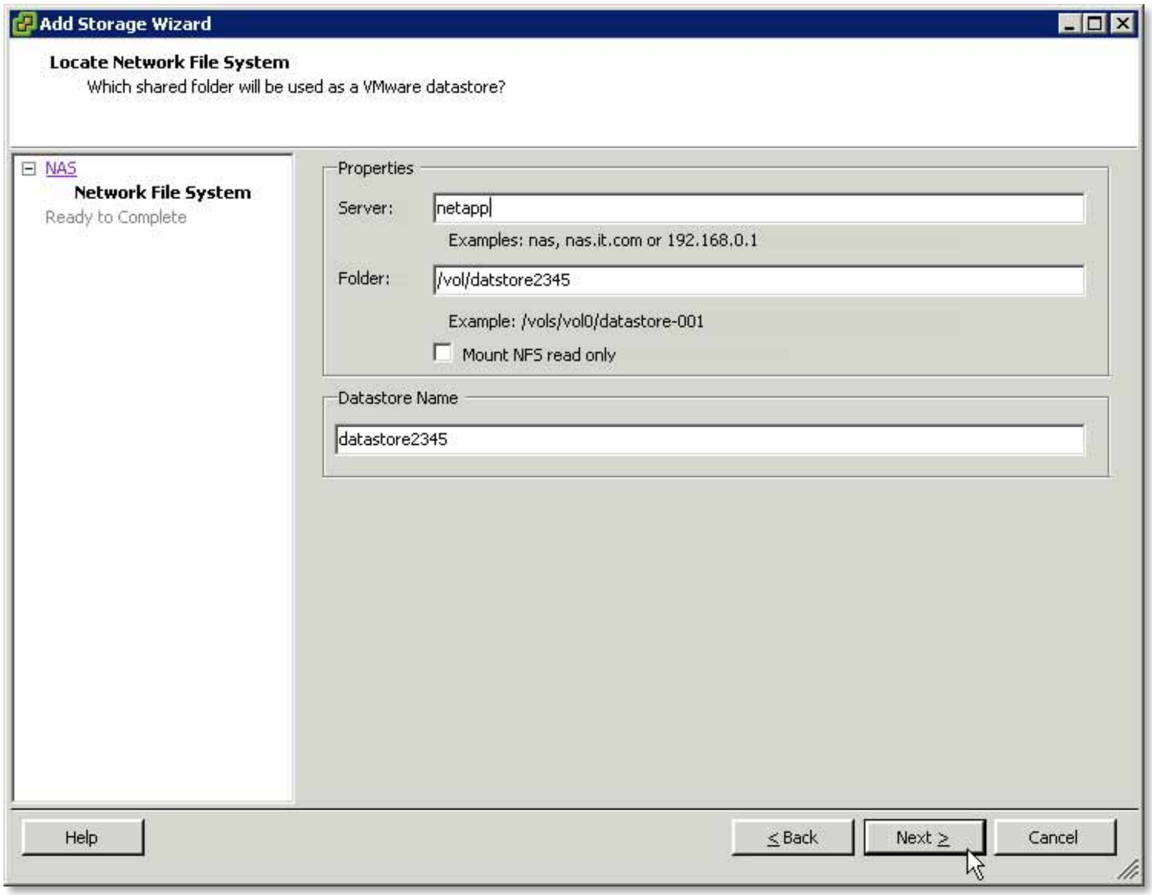
○ **Disk/LUN**

Choose this option if you want to create a datastore or other volume on a Fibre Channel, iSCSI or local SCSI disk.

● **Network File System**

Choose this option if you want to use a shared folder over a network connection as if it were a VMware datastore. A mount point must be created on the host before it is added as a datastore.

Figure 23) VMware Add Storage Wizard.



Locate Network File System

Which shared folder will be used as a VMware datastore?

Properties

Server: netapp

Examples: nas, nas.it.com or 192.168.0.1

Folder: /vol/datstore2345

Example: /vols/vol0/datastore-001

□ Mount NFS read only

Datastore Name

datastore2345

Figure 24) VMware Add Storage Wizard NFS configuration.

6 NETAPP ESX HOST UTILITIES

NetApp provides a utility for simplifying the management of ESX nodes running on NetApp storage arrays. With the ESX Host Utilities a customer receives a collection of tools to automate and simplify the configuration of HBAs, sets ESX Server options for NFS, optimizes Guest OS (VM) SCSI settings, and provide diagnostic utilities should a support case be opened.

The ESX Host Utilities installs on ESX 3.5 systems, and currently is not supported for ESXi 3.5 systems. The ESX Host Utilities can be downloaded from NOW.

6.1 INSTALLING THE EHU IN ESX

EHU PREREQUISITES

- FC, iSCSI, or NFS is licensed on the storage system.
- You have root access to each ESX Server
- All storage systems have DNS-resolvable names.
- SSL is set up on every storage controller before installing the Host Utilities 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 Host Utilities, it is recommended that you enable options httpd.admin.hostsequiv on each controller (option httpd.admin.hostsequiv.enable on) and that all VMware ESX host names are added to the /etc/hosts.equiv file on each controller. This will prevent connection problems between the controller(s) and the host(s).

EHU INSTALLATION

To install the ESX Host Utilities complete the following steps:

1	Download the EHU.
2	Copy the EHU to a location accessible to the ESX
3	Extract the EHU by running tar –zxf <name of EHU file>.tar.gz
4	Migrate running VMs to other ESX nodes
5	Place the ESX Server in maintenance mode.
6	Complete the EHU installation wizard
7	Run ./install
8	Complete the EHU installation wizard
9	Reboot the ESX Server and return to normal operations

EHU ASSISTED MULTIPATHING

One of the components of the Host Utilities is a script called config_mpath. This script reduces the administrative overhead of managing SAN LUN paths by using the procedures previously described. The config_mpath script determines the desired primary paths to each of the SAN LUNs on the ESX Server and then sets the preferred path for each LUN to use one of the primary paths. Simply running the config_mpath script once on each 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, the script is simply run an additional time to update the multipathing configuration based on the changes to the environment.

6.2 MANUAL CONFIGURATION OF FC HBAS IN ESX

While NetApp highly recommends using the ESX Host Utilities to configure optimal settings for deployments these setting can also be scripted for customers who deploy ESX Server via scripts or deployment appliances.

Note: If your storage arrays are not running Data ONTAP 7.2.4 or later in Single System Image, then replace the value of 10 in the following steps with 120.

QLOGIC FC HBAS

1	Connect to the ESX console.
2	Migrate running VMs to other ESX nodes
3	Place the ESX Server in maintenance mode.
4	Run the command in step 5.
5	/usr/sbin/esxcfg-module -s `qlport_down_retry=5' qla2300_707_vmw
6	Verify the settings by running the command in step 7
7	/usr/sbin/esxcfg-module -q qla2300_707_vmw
8	The value of 5 should be returned.
9	Reboot the ESX Server and return to normal operations

EMULEX FC HBAS

1	Connect to the ESX console.
2	Migrate running VMs to other ESX nodes
3	Place the ESX Server in maintenance mode.
4	Run the command in step 5.
5	/usr/sbin/esxcfg-module -s "lpfc_nodev_tmo=10" lpfc_740
7	Verify the settings by running the command in step 7
8	/usr/sbin/esxcfg-module -q lpfc_740
9	The value of 10 should be returned.
10	Reboot the ESX Server and return to normal operations

6.3 MANUAL CONFIGURATION OF FC HBAS IN ESXI

As mentioned earlier in this section, The ESX Host Utilities is currently not supported on ESXi systems; however, many of the configuration setting must be applied. With ESXi, these settings can be set using the VMware Remote Console. Note: for this article, the Windows Remote CLI was used.

When installing the remote console, make sure the location has been added to your path. Otherwise, to run the RCLI commands, change to Program Files\VMware\VMware VI Remote CLI\bin.

Note: If your storage arrays are not running Data ONTAP 7.2.4 or later in Single System Image, then replace the value of 10 in the following steps with 120.

QLOGIC FC HBAS

1	Connect to the ESXi host via the remote console.
2	Migrate running VMs to other ESXi nodes
3	Place the ESXi Server in maintenance mode.
4	Run the command in step 5.
5	perl vicfg-module.pl --server <hostname> --username root --password <password> --set-options "qlport_down_retry=5" qla2300_707_vmw
6	Verify the settings by running the command in step 7
7	perl vicfg-module.pl --server <hostname> --username root --password <password> --get-options qla2300_707_vmw
8	The value of 5 should be returned.
9	Reboot the ESX Server and return to normal operations

EMULEX FC HBAS

1	Connect to the ESXi host via the remote console.
2	Migrate running VMs to other ESXi nodes
3	Place the ESXi Server in maintenance mode.
4	Run the command in step 5.
5	perl vicfg-module.pl --server <hostname> --username root --password <password> --set-options "lpfc_nodev_tmo=10" lpfc_740
7	Verify the settings by running the command in step 7
8	perl vicfg-module.pl --server <hostname> --username root --password <password> --get-options lpfc_740
9	The value of 10 should be returned.
10	Reboot the ESX Server and return to normal operations

7 FC STORAGE NETWORKING BEST PRACTICES

Fibre Channel storage networks make up the largest percentage of shared storage infrastructures host ESX. This market share is attributed to FC being the first networked attached storage protocol supported by ESX in version 2.0. While FC is a well-known and mature technology this section will cover best practices for deploying VMware on Fibre Channel with NetApp storage arrays.

7.1 HOST BUS ADAPTERS

ESX servers and NetApp storage arrays connect to a SAN fabric by the SAN Host Bus Adapters (HBAs). Each HBA can run as either an initiator (ESX) or as a target (NetApp), and each HBAs has a global unique address referred to as a World Wide Port Name (WWPN). Each WWPN is required to be known in order to configure LUN access on a NetApp storage array.

Both NetApp and VMware highly recommend that as a best practice each ESX server should have at least two FC HBA ports, preferably with each on a separate physical card. For more information on VMware FC best practices and recommendations please see VMware Fibre Channel SAN configuration Guide.

7.2 NETAPP IGROUPS (LUN MASKING)

LUN (Logical Unit Number) Masking is an authorization process that makes a LUN available to a host or set of hosts in a cluster. On a NetApp array LUN Masking is implemented by assigning HBA (Host Bus Adapter) addresses to initiator groups (igroups). Once an igroup has been defined then LUNs can be assigned the igroup for access to the LUN.

Implementation best practices for LUN masking is covered in the storage provisioning section for both FC & iSCSI.

7.3 FC ZONING

Many devices and nodes can be attached to a SAN, and 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 leverages its name servers to either allow or block access to particular World Wide Names (WWNs) in the fabric. A major advantage of WWN zoning is the ability to re-cable the fabric without having to redo the zone information.

ZONING RECOMMENDATION

NetApp & VMware highly recommend customer implement single initiator multiple storage target zones. This design offers an ideal balance of simplicity and availability with FC deployments.

8 IP STORAGE NETWORKING BEST PRACTICES

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

CONFIGURATION OPTIONS FOR PRODUCTION IP STORAGE NETWORKS

One of the challenges of configuring VMware ESX networking for IP storage is that the network configuration should meet these three goals simultaneously:

- Be redundant across switches in a multi switch environment
- Use as many available physical paths as possible
- Be scalable across multiple physical interfaces

8.1 10 GIGABIT ETHERNET

VMware ESX 3 and ESXi 3 introduced support for 10 Gb Ethernet. 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 ESX I/O compatibility guide.

8.2 VIRTUAL LANS (VLANS)

When segmenting network traffic with VLANs, interfaces can either be dedicated to a single VLAN or they 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 an ESX server with physical link redundancy. By adding multiple VLANs, one can group common IP traffic onto separate VLANs for optimal performance. It is recommended to group Service console access with the Virtual Machine Network on one VLAN, and on a second VLAN the VMkernel activities of IP Storage and VMotion should reside.

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. NetApp storage appliances also allow the restriction of the iSCSI protocol to specific interfaces and/or VLAN tags. 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, make 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 ESX servers, FAS storage arrays, and network switches. It is recommended to configure the end points, ESX servers & NetApp arrays with flow control set to ‘send on’ and ‘receive off’

For network switches it is recommended to set the switch ports connecting to ESX hosts and FAS storage arrays to either ’Desired’, or if this mode isn’t available, set these ports to ‘send off’ and ‘receive on.’ Note the switch ports are configured with the opposite settings of the end points, the ESX & FAS systems. See Figure 25.

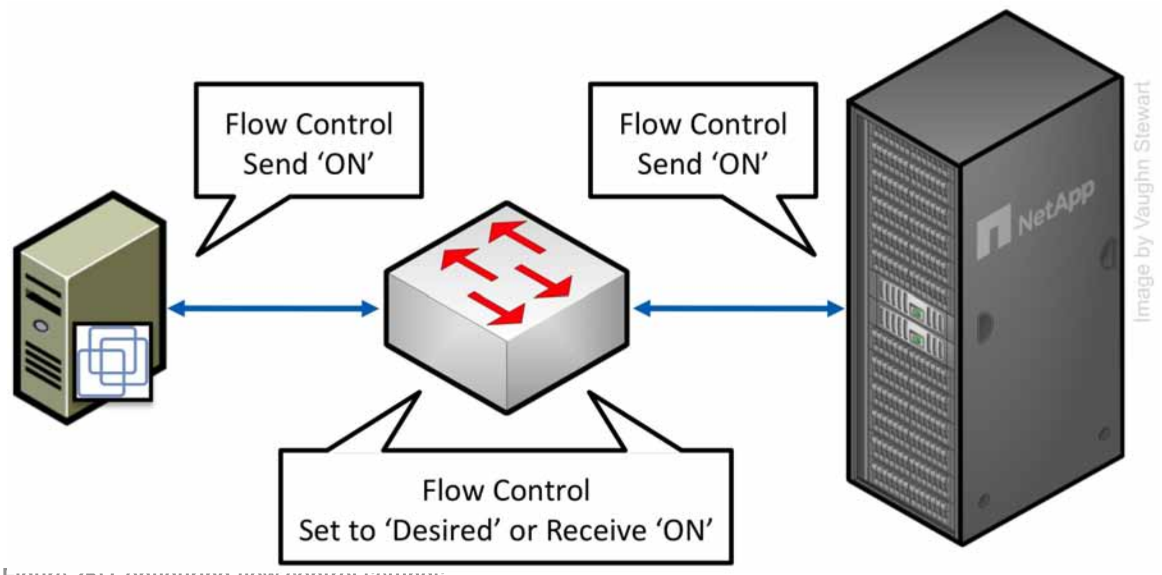


Figure 25) Configuring flow control settings.

8.4 SPANNING TREE PROTOCOL

The Spanning Tree Protocol (STP) is a network protocol that ensures 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 ESX & NetApp storage arrays to Ethernet storage networks it is highly recommended that the Ethernet ports that these systems connect to be configured with either RSTP or portfast enabled.

8.5 BRIDGE PROTOCOL DATA UNITS

Bridge Protocol Data Units (BPDUs) exchange information about bridge IDs and root path costs within STP. When connecting ESX & NetApp storage arrays to Ethernet storage networks it is highly recommended that the Ethernet ports which these systems connect to are configured with BPDU disabled.

8.6 NETAPP VIRTUAL INTERFACES

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

NetApp enables the use of two types of load-balancing VIFs: Multimode and Dynamic Multimode

Multimode VIFs are static etherchannels. 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 VIFs are LACP (IEEE 802.3ad) compliant VIFs. In a dynamic multimode VIF, all of the physical connections are simultaneously active and carry traffic as with multimode VIFs, described above. Dynamic Multmode VIFs introduce the use of LACP signaling transmissions between the FAS array and the remote switch. This signaling informs the remote channeling partner of link status and if a failure or inability to transmit data on a link is observed the device identifying this problem will inform 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 a interface to be removed from a standard multimode VIF in 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. It is recommended to use IP based source and destination load-balancing especially when the network is designed to route storage traffic. This is because during a transmission of a routed packet a host will transmit the packet to the default router IP address. Upon arriving at the router, the router will change the MAC address of the routed packet to the MAC address of the local router inerface the packet is transmitted out on. 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 ESX hosts and FAS arrays 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.

It is also possible to create second-level single or multimode VIFs. By using second-level VIFs it is possible to 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 different 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 in the event of an interface or switch failure, 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 different 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 different physical switches are combined into an 802.3ad Etherchannel trunk. The advantage of multi-switch Etherchannel trunks is that they can eliminate the need for additional passive links that are accessed only during failure scenarios in some configurations.

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

9 CONFIGURING ETHERNET STORAGE NETWORKS

9.1 HIGH AVAILABLE IP STORAGE DESIGN WITH TRADITIONAL ETHERNET SWITCHES

INTRODUCING MULTIPLE VMKERNELS PORTS

In order to simultaneously use multiple paths while providing high availability with traditional Ethernet switches, each ESX server must be configured with a minimum of two VMkernel ports in the same VSwitch. This VSwitch is configured with a minimum of two network adapters. Each of these VMkernel ports supports IP traffic on a different subnet. Because the two VMkernel ports are in the same VSwitch they can share the physical network adapters in that VSwitch.

EXPLAING ESX SERVER ADAPTER FAILOVER BEHAVIOR

In case of ESX server adapter failure (due to a cable pull or NIC failure), traffic originally running over the failed adapter is rerouted and continues via 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

REVIEWING LINK AGGREGATION WITHIN ESX SERVER

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

In this document any reference to Etherchannel in the terms of configuring an ESX server is actually referring to a static etherchannel. ESX server does not support the use of LACP 802.3ad..

SWITCH FAILURE

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

CONNECTING TO DATASTORES

With Ethernet based storage networking protocols VMware datastores are mounted by IP addresses. Both iSCSI & 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 ESX/ESXi host. Using NFS to connect to the same volume multiple times from a single ESX/ESXi host should be avoided, as ESX/ESXi and vCenter will consider these connections to be different datastores.

Note 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 ESX/ESXi host may use a different active IP path.

Regarding NFS datastores each datastore should be connected only once from each ESX/ESXi server, and using the same netapp target IP address on each ESX/ESXi server.

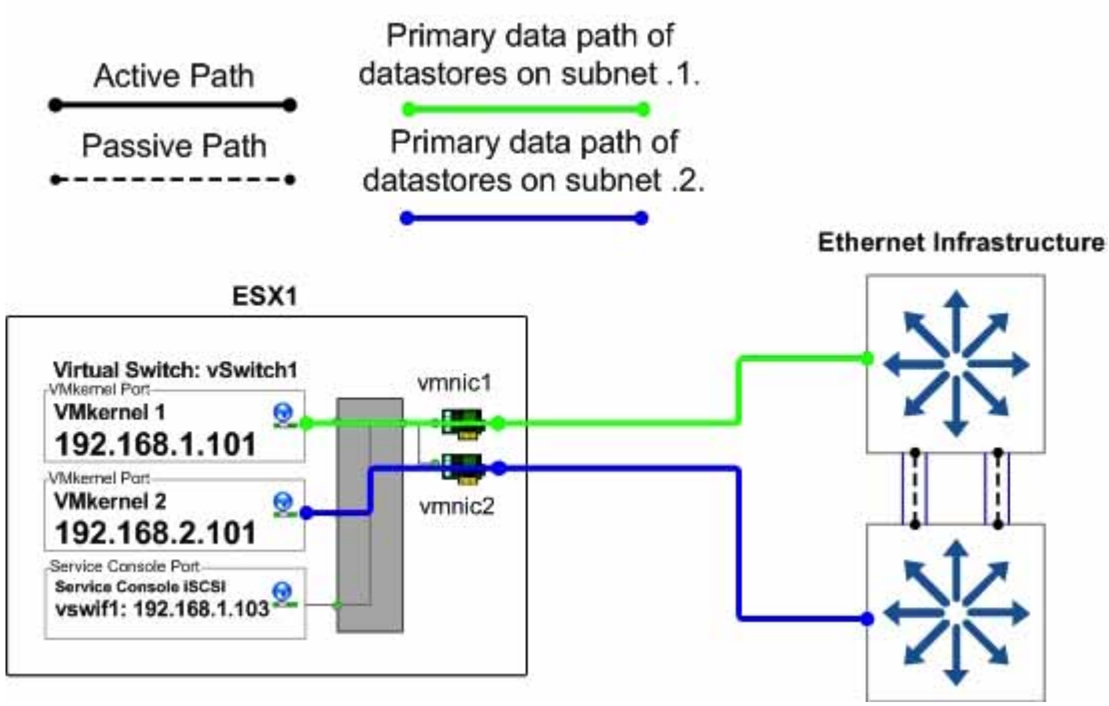


Figure 26) ESX vSwitch1 normal mode operation.

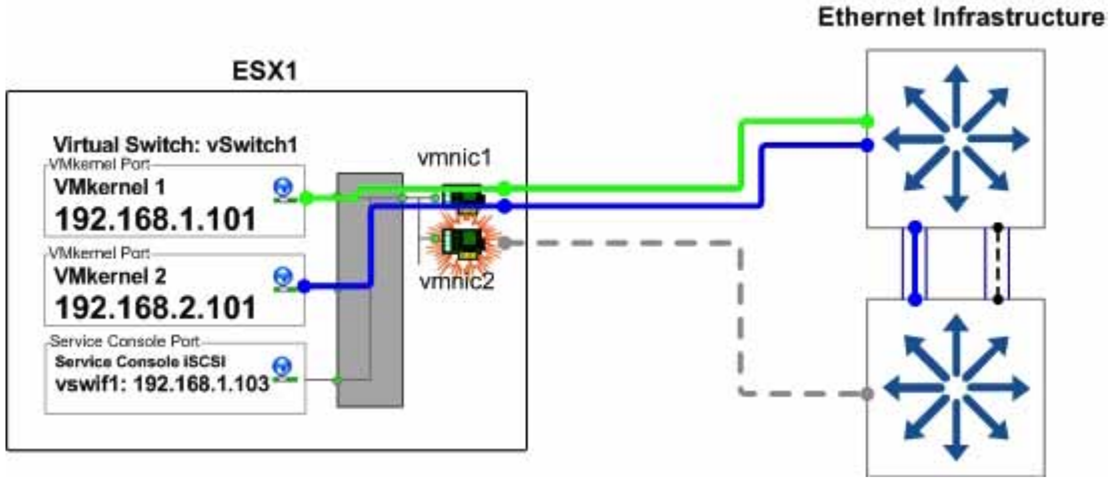


Figure 27) ESX vSwitch1 failover mode operation.

SCALABILITY OF ESX SERVER NETWORK CONNECTIONS

Although the configuration shown in the figure above uses two network adapters in each ESX Server, it could be scaled up to use additional adapters, with another VMkernel port, subnet, and IP address added for each additional adapter.

Another option would be 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 could 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 VMKERNEL CONFIGURATION WITH TRADITIONAL ETHERNET

If the switches to be used for IP storage networking do not support multi-switch Etherchannel trunking, or virtual port channeling, then the task of providing cross-switch redundancy while making active use of multiple paths becomes more challenging. To accomplish this, each ESX Server must be configured with at least two VMkernel IP storage ports addressed on different subnets. As with the previous option, multiple Datastore connections to the storage controller are necessary using different target IP addresses. Without the addition of a second VMkernel port, the VMkernel would simply route all outgoing requests through the same physical interface, without making use of additional VMNICs on the vSwitch. In this configuration, each VMkernel port is set with its IP address on a different subnet. The target storage system is also configured with IP addresses on each of those subnets, so the use of specific VMNIC interfaces can be controlled.

ADVANTAGES

- Provides two active connections to each storage controller (but only one active path per Datastore).
- Easily scales to more connections.
- Storage controller connection load balancing is automatically managed Virtual Port load balancing policy. This is a non-Etherchannel solution.

DISADVANTAGE

- Requires the configuration of at least two VMkernel IP storage ports.

In the ESX Server configuration shown in Figure 28, 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 different physical switch.

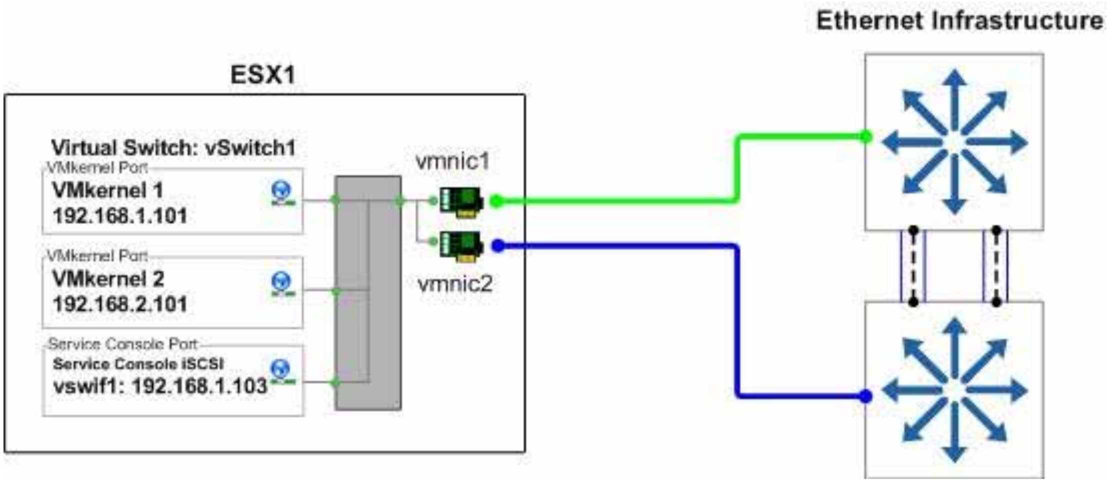
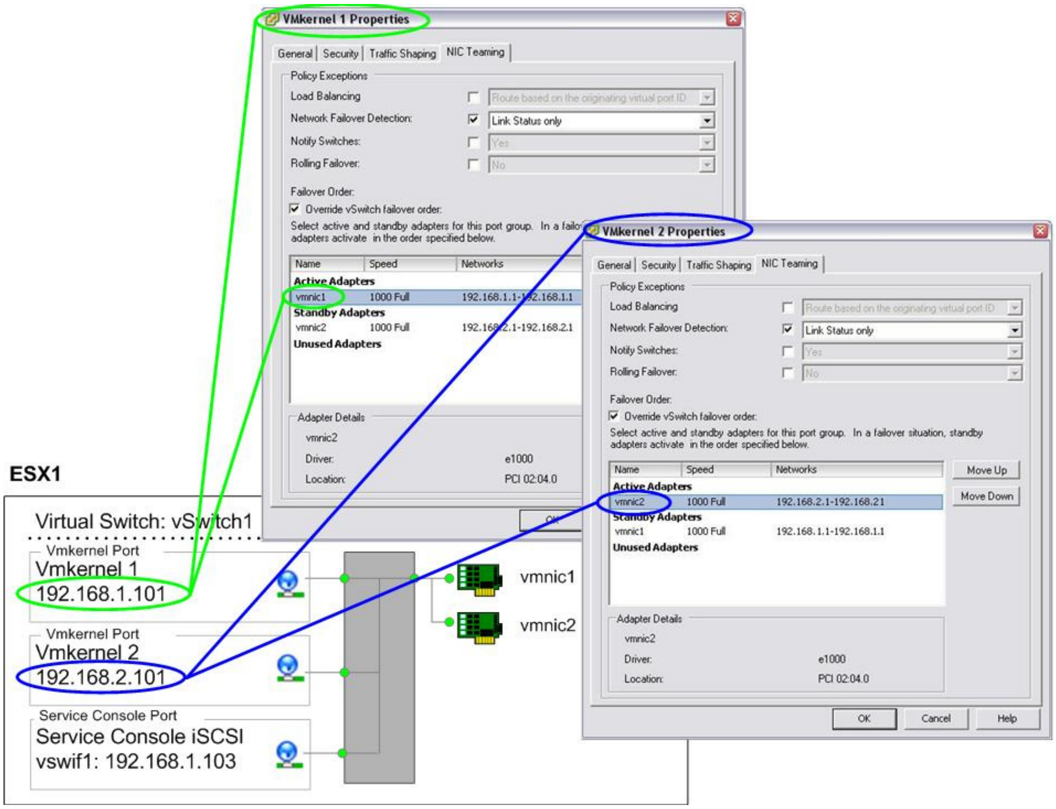


Figure 28) ESX Server physical NIC connections with traditional Ethernet.

In vSwitch1, two VMkernel ports have been created (VMkernel 1 and VMkernel 2). Each VMkernel port has been configured with an IP address on a different subnet, and the NIC Teaming properties of each VMkernel port have been configured as follows.

- **VMkernel 1:** IP address set to 192.168.1.101.
- **VMkernel 1 Port Properties:**
  - o Enable the Override vSwitch Failover Order option.
  - o Set Active Adapter to vmnic1.
  - o Set Standby Adapter to vmnic2.
- **VMkernel 2:** IP address set to 192.168.2.101.
- **VMkernel2 Port Properties:**
  - o Enable the Override vSwitch Failover Order option.
  - o Set Active Adapter to vmnic2.
  - o Set Standby Adapter to vmnic1.



VMkernel 1 Properties

General | Security | Traffic Shaping | NIC Teaming

Policy Exceptions

Load Balancing: □ Route based on the originating virtual port ID

Network Failover Detection: ☑ Link Status only

Notify Switches: □ Yes

Rolling Failover: □ No

Failover Order:

☑ Override vSwitch failover order:

Select active and standby adapters for this port group. In a failo adapters activate in the order specified below.

Name	Speed	Networks
Active Adapters		
vmnic1	1000 Full	192.168.1.1-192.168.1.1
Standby Adapters		
vmnic2	1000 Full	192.168.2.1-192.168.2.1
Unused Adapters		

Adapter Details

vmnic2

Driver: e1000

Location: PCI 02:04.0

VMkernel 2 Properties

General | Security | Traffic Shaping | NIC Teaming

Policy Exceptions

Load Balancing: □ Route based on the originating virtual port ID

Network Failover Detection: ☑ Link Status only

Notify Switches: □ Yes

Rolling Failover: □ No

Failover Order:

☑ Override vSwitch failover order:

Select active and standby adapters for this port group. In a failover situation, standby adapters activate in the order specified below.

Name	Speed	Networks
Active Adapters		
vmnic2	1000 Full	192.168.2.1-192.168.21
Standby Adapters		
vmnic1	1000 Full	192.168.1.1-192.168.1.1
Unused Adapters		

Adapter Details

vmnic2

Driver: e1000

Location: PCI 02:04.0

Figure 29) ESX Server VMkernel port properties with traditional Ethernet.

9.3 IP STORAGE ARCHITECTURE WITH TRADITIONAL ETHERNET

In this configuration, the IP switches to be used do not support multi-switch Etherchannel trunking, so each storage controller requires four physical network connections. This design is available in two options (represented in Figure 30 & Figure 31). Both designs are very similar. They both provide multiple active links to each storage controller, provides a means to scale throughput by simply adding more links, require multiple IP addresses per controller and each utilize two physical links for each active network connection in order to achieve path high availability.

THE MULTI-MODE DESIGN

The multi-mode design 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. NetApp refers 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 MULTI MODE VIFS

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

DISADVANTAGES OF USING MULTI MODE VIFS

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

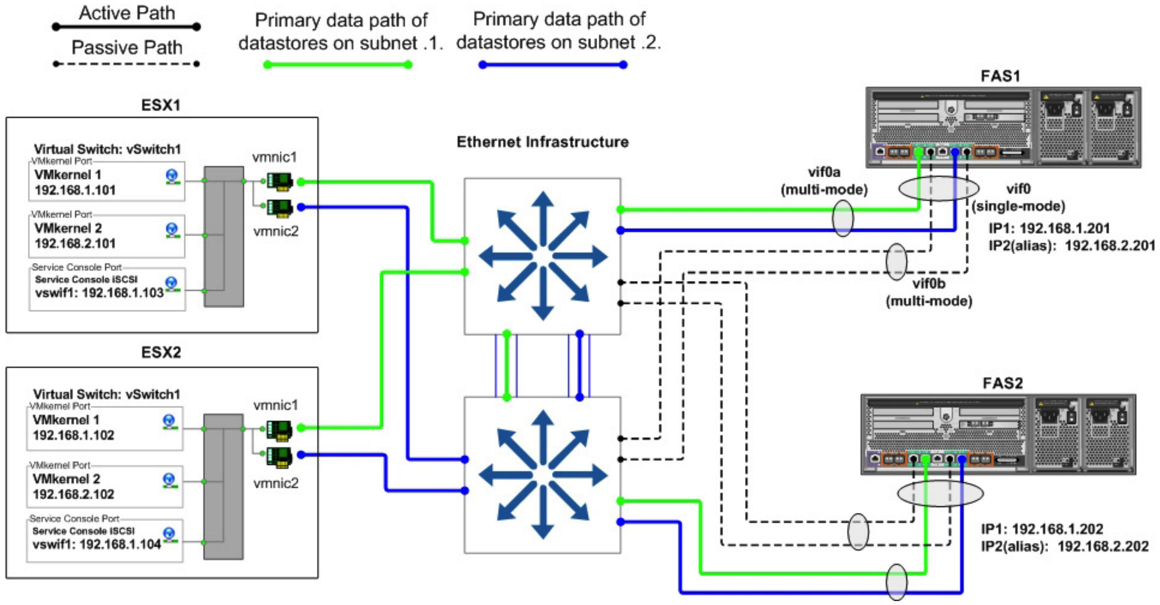


Figure 30) Storage side multimode VIFs.

THE SINGLE-MODE DESIGN

The single-mode design 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 preferred due to its simplicity and the lack of any special configuration on the network switches.

ADVANTAGES OF USING SINGLE MODE VIFS

- Simplicity - No switch side configuration is required.

DISADVANTAGES OF USING SINGLE MODE VIFS

- Data I/O to a single IP is not aggregated over multiple links without adding more links.

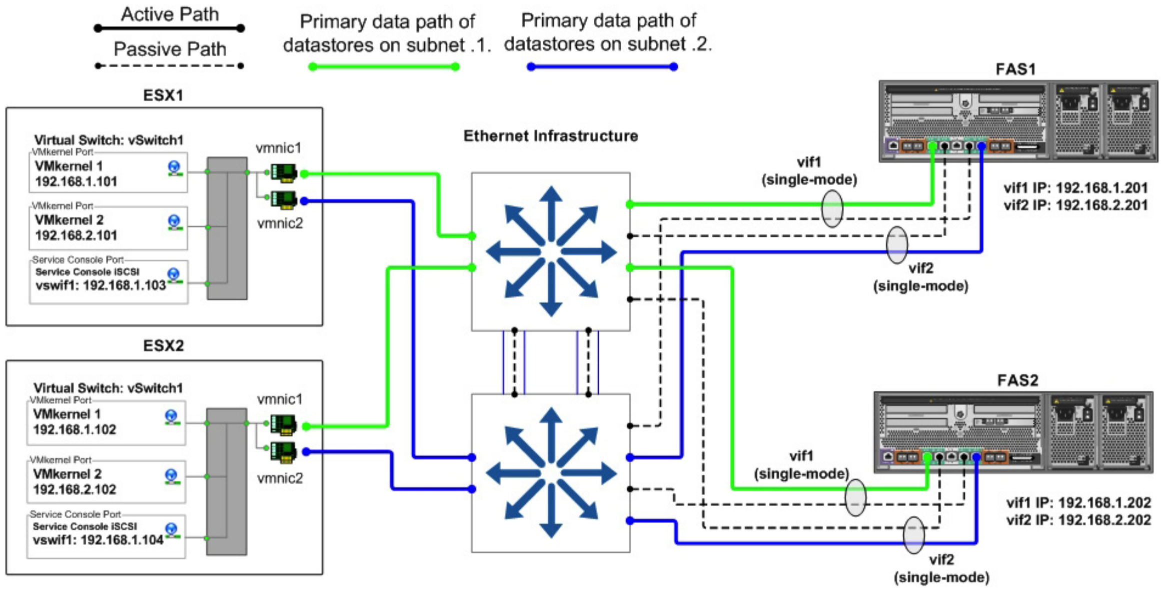


Figure 31) Storage side single-mode VIFs.

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 different IP address.

In addition to configuring the ESX Server interfaces as shown in the examples, the NetApp storage controller has been configured with an IP address on each of the subnets used to access Datastores. This is accomplished by the use of multiple teamed adapters, each with its own IP address or, in some 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 ESX Servers, the administrator configures the connection to use one of the IP addresses assigned to the NetApp storage controller. When using NFS Datastores, this is accomplished by specifying the IP address when mounting the Datastore.

The figure below show an overview of storage traffic flow when using multiple ESX Servers and multiple Datastores. Note 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 ESX/ESXi host may use a different active IP path. Regarding NFS datastores each datastore should be connected only once from each ESX/ESXi server, and using the same netapp target IP address on each ESX/ESXi server.

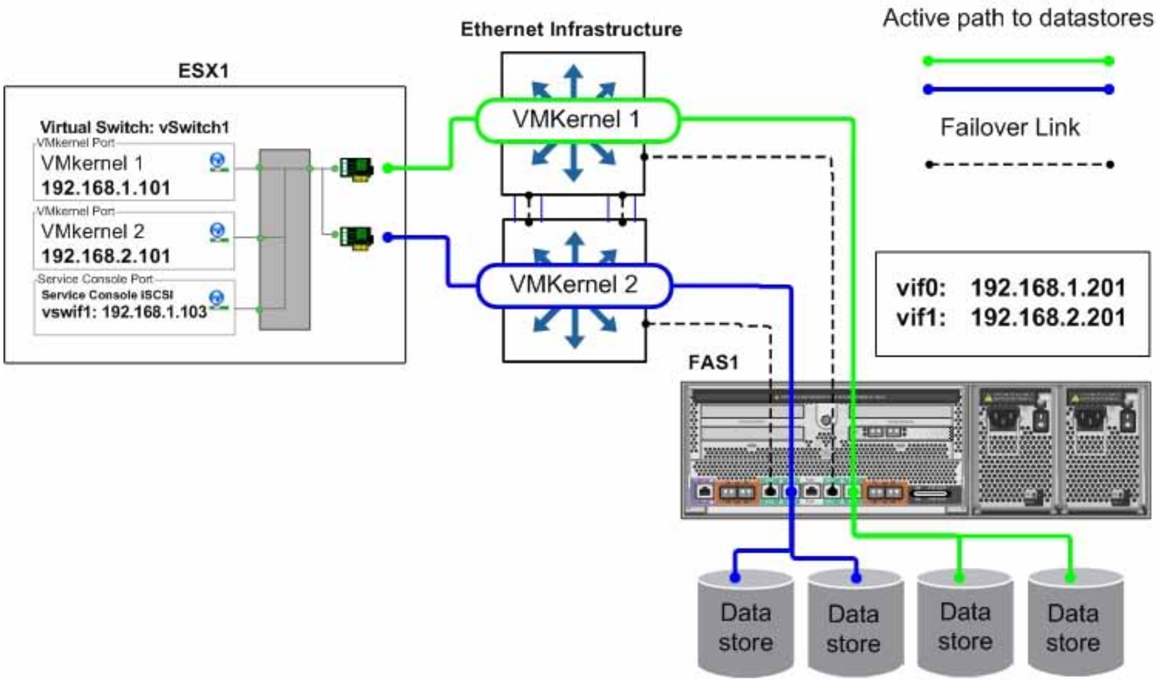


Figure 32) Datastore connections with traditional Ethernet.

9.5 VMKERNEL CONFIGURATION WITH MULTI-SWITCH TRUNKING

If the switches used for IP storage networking support multi-switch Etherchannel trunking, or virtual port channeling, then each ESX Server needs 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 different target IP addresses are necessary to use each of the available physical links.

ADVANTAGES

- Simple
- Provides two active connections to each storage controller.
- Easily scales via 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

- Requires multi-switch Etherchannel capability such as stackable switches or virtual port channeling.

In the ESX Server configuration shown in the Figure 33 below, 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 different physical switch and 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.

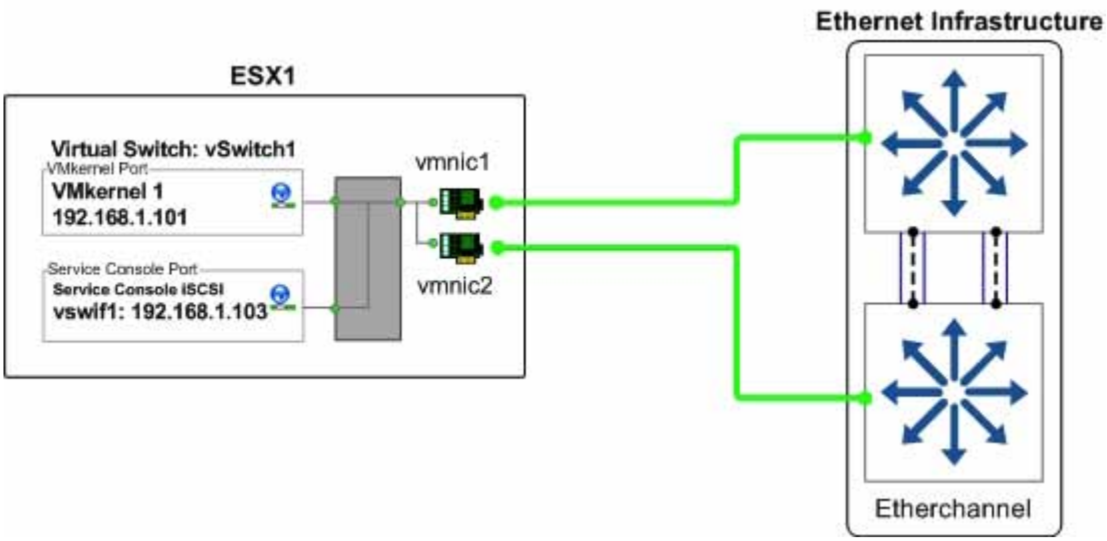
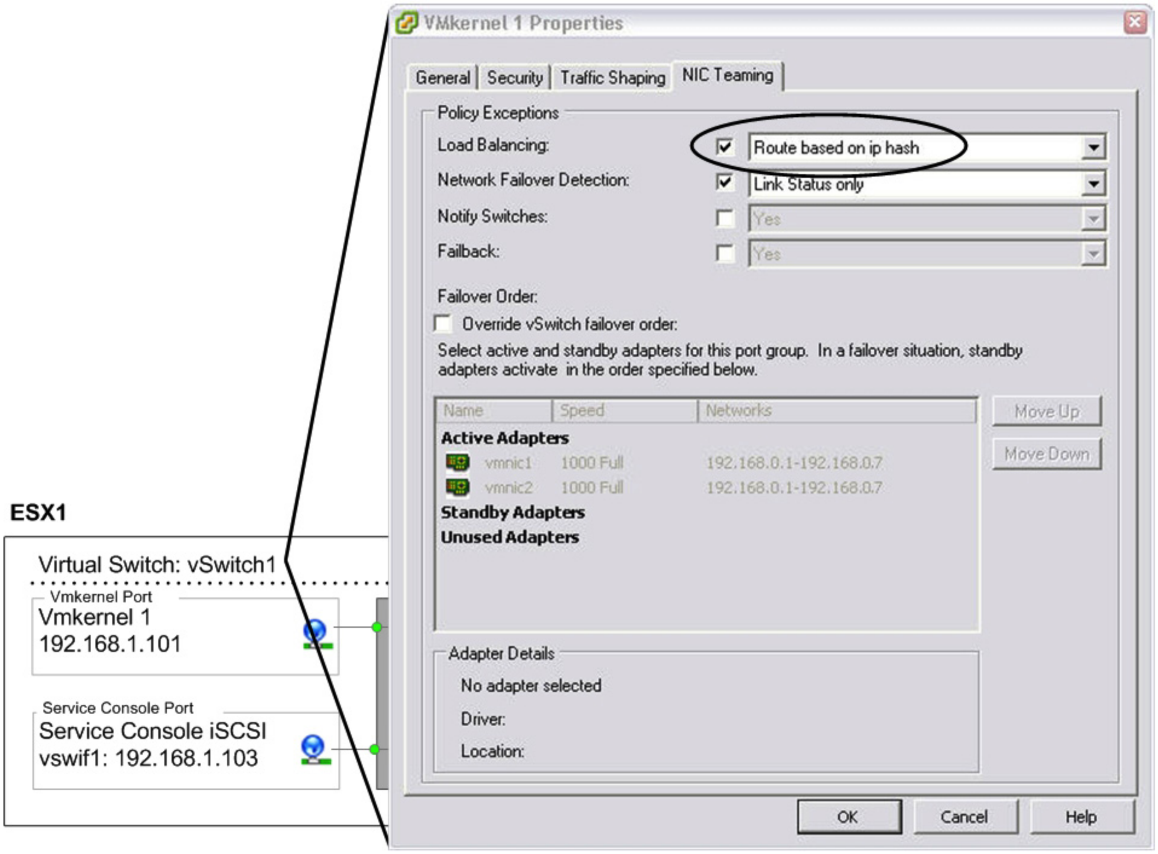


Figure 33) ESX Server physical NIC connections with multi-switch Etherchannel.

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:

- **VMkernel 1:** IP address set to 192.168.1.101.
- **VMkernel 1 Port Properties:** Load-balancing policy set to “Route based on IP hash.”



Policy Exceptions

Load Balancing: ☑ Route based on ip hash

Network Failover Detection: ☑ Link Status only

Notify Switches: □ Yes

Failback: □ Yes

Failover Order:

□ Override vSwitch failover order:

Select active and standby adapters for this port group. In a failover situation, standby adapters activate in the order specified below.

Name	Speed	Networks
Active Adapters		
vmnic1	1000 Full	192.168.0.1-192.168.0.7
vmnic2	1000 Full	192.168.0.1-192.168.0.7
Standby Adapters		
Unused Adapters		

Adapter Details

No adapter selected

Driver:

Location:

Figure 34) ESX Server VMkernel port properties with multi-switch Etherchannel.

9.6 IP STORAGE ARCHITECTURE WITH MULTI-SWITCH TRUNKING

If the switches to be used for IP storage networking support cross-stack Etherchannel trunking, then each storage controller needs only one physical connection to each switch; the two ports connected to each storage controller are then combined into one multimode LACP VIF with IP load balancing enabled. Multiple IP addresses can be assigned to the storage controller by using IP address aliases on the VIF.

ADVANTAGES

- 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 IP load balancing policy.

DISADVANTAGE

- Not all switch vendors or switch models support cross-switch Etherchannel trunks.

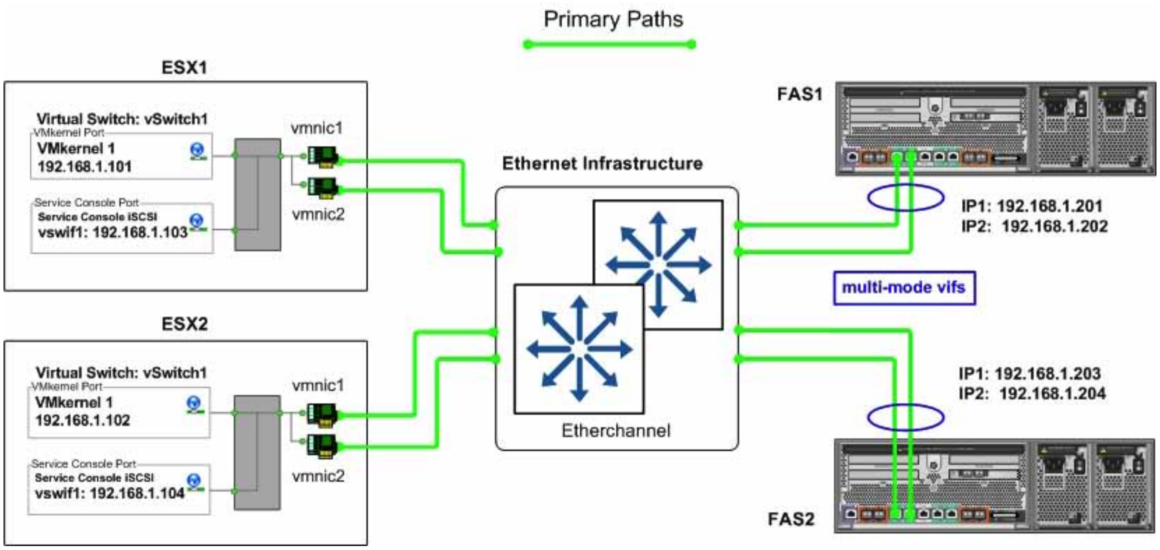


Figure 35) Storage side multimode VIFs using multi-switch Etherchannel.

9.7 DATASTORE CONFIGURATION WITH MULTI-SWITCH TRUNKING

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 different IP address.

In addition to configuring the ESX Server interfaces as shown in the examples, the NetApp storage controller has been configured with an IP address on each of the subnets used to access Datastores. This is accomplished by the use of multiple teamed adapters, each with its own IP address or, in some 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 ESX Servers, the administrator configures the connection to use one of the IP addresses assigned to the NetApp storage controller. When using NFS Datastores, this is accomplished by specifying the IP address when mounting the Datastore.

The figure below show an overview of storage traffic flow when using multiple ESX Servers and multiple Datastores. Note 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 ESX/ESXi host may use a different active IP path. Regarding NFS datastores each datastore should be connected only once from each ESX/ESXi server, and using the same netapp target IP address on each ESX/ESXi server.

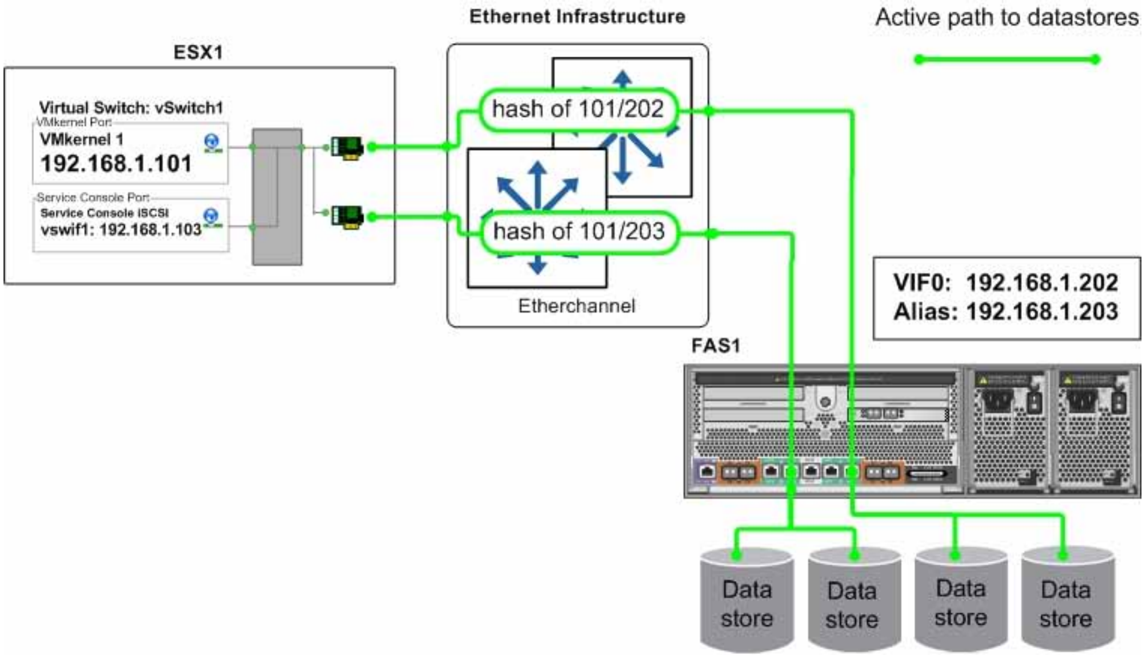


Figure 36) Datastore connections with cross-stack Etherchannel.

10 INCREASING STORAGE UTILIZATION

VMware 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 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 have any impact on improving storage utilization (and in many cases may have the opposite effect).

By default in ESX 3.5, virtual disks preallocate the storage they require and in the background zero out all of the storage blocks. This type of VMDK format is called a *zeroed thick VMDK*. VMware 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.

Thin-provisioned VMDKs are not available to be created in the Virtual Infrastructure client with VMFS Datastores. To implement thin VMDKs with VMFS, you must create a thin-provisioned VMDK file by using the `vmkfstools` command with the `–d` options switch. By using VMware thin-provisioning technology, you can reduce the amount of storage consumed on a VMFS datastore.

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 into the thin-provisioned format, with the single exception of importing ESX 2.x formatted VMDKs into ESX 3.x.

NetApp offers storage virtualization technologies that can enhance the storage savings provided by VMware thin provisioning. FAS data deduplication and the thin provisioning of VMFS Datastores and RDM LUNs offer considerable storage savings by increasing storage utilization of the FAS array. Both of these technologies are native to NetApp arrays and don’t require any configuration considerations or changes to be implemented within the ESX Servers.

10.1 DATA DEDUPLICATION

One of the most popular VMware 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 37 shows an example of typical storage consumption in a VI3 deployment.

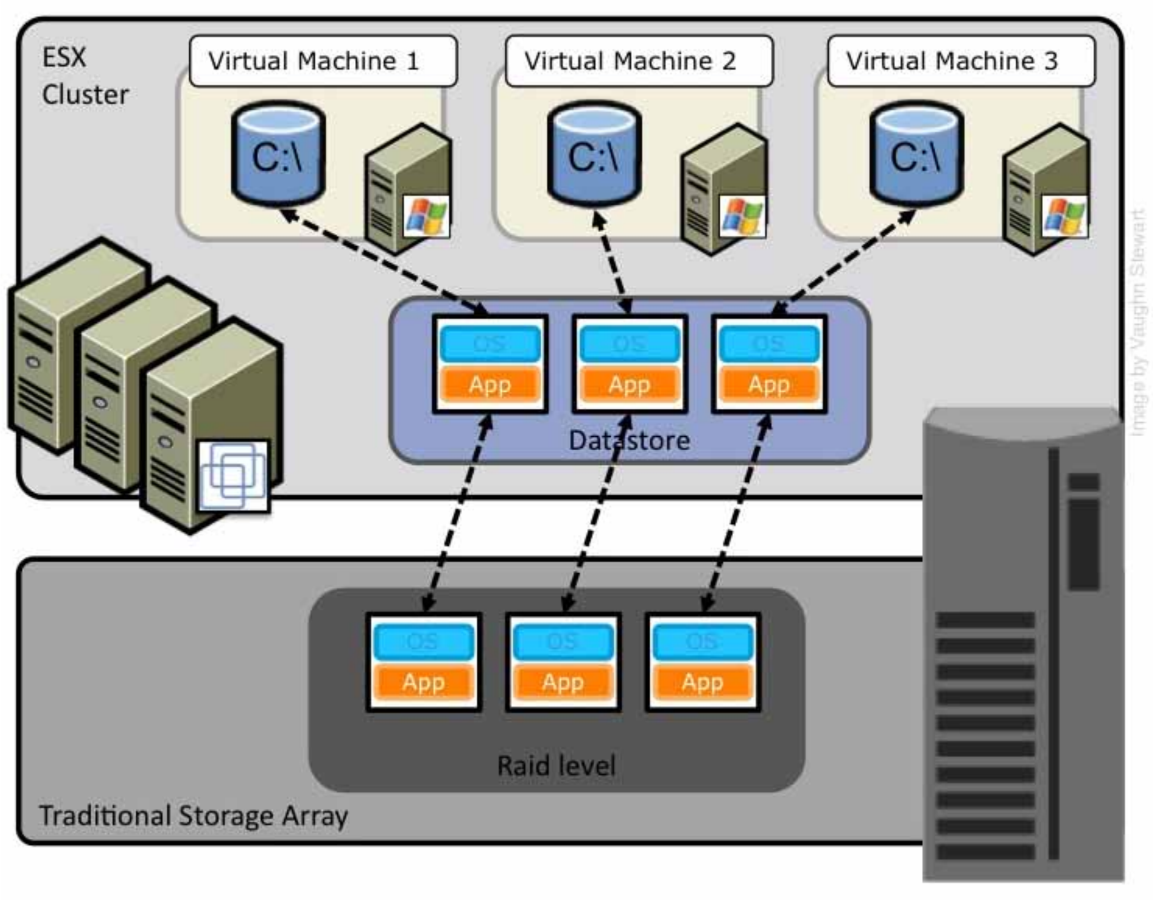


Figure 37) Storage consumption with a traditional array.

NetApp offers a data deduplication technology called *FAS data deduplication.* With NetApp FAS deduplication, VMware deployments can eliminate the duplicate data in their environment, enabling greater storage utilization. Deduplication virtualization technology enables multiple virtual machines to share the same physical blocks in a NetApp FAS system in the same manner that VMs share system memory. It can be seamlessly introduced into a virtual infrastructure without having to make any changes to VMware administration, practices, or tasks. Deduplication runs on the NetApp FAS system at scheduled intervals and does not consume any CPU cycles on the ESX Server. Figure 38 shows an example of the impact of deduplication on storage consumption in a VI3 deployment.

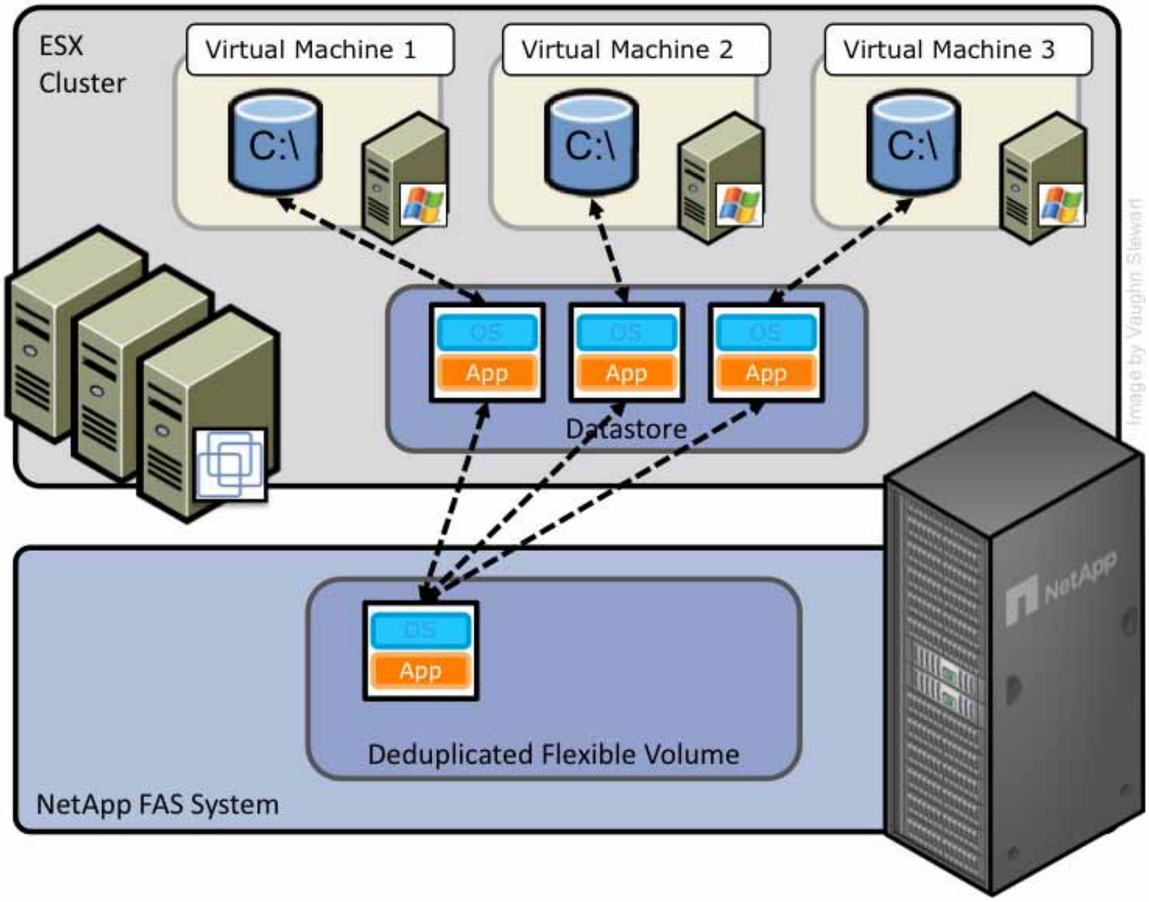


Figure 38) Storage consumption after enabling FAS data deduplication.

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

Note: whether one enables data deduplication or not has no impact in terms of the number of VMs which reside on a datastore. One should size a datastore density as one would without deduplication.

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 NetApp LUN thin provisioning. For details, see section 10.3, “Storage Thin Provisioning.” In addition, although deduplication reduces the amount of consumed storage, the VMware administrative team does not see this benefit directly, because their view of the storage is at a LUN layer, and LUNs always represent their provisioned capacity, whether they are traditional or thin provisioned.

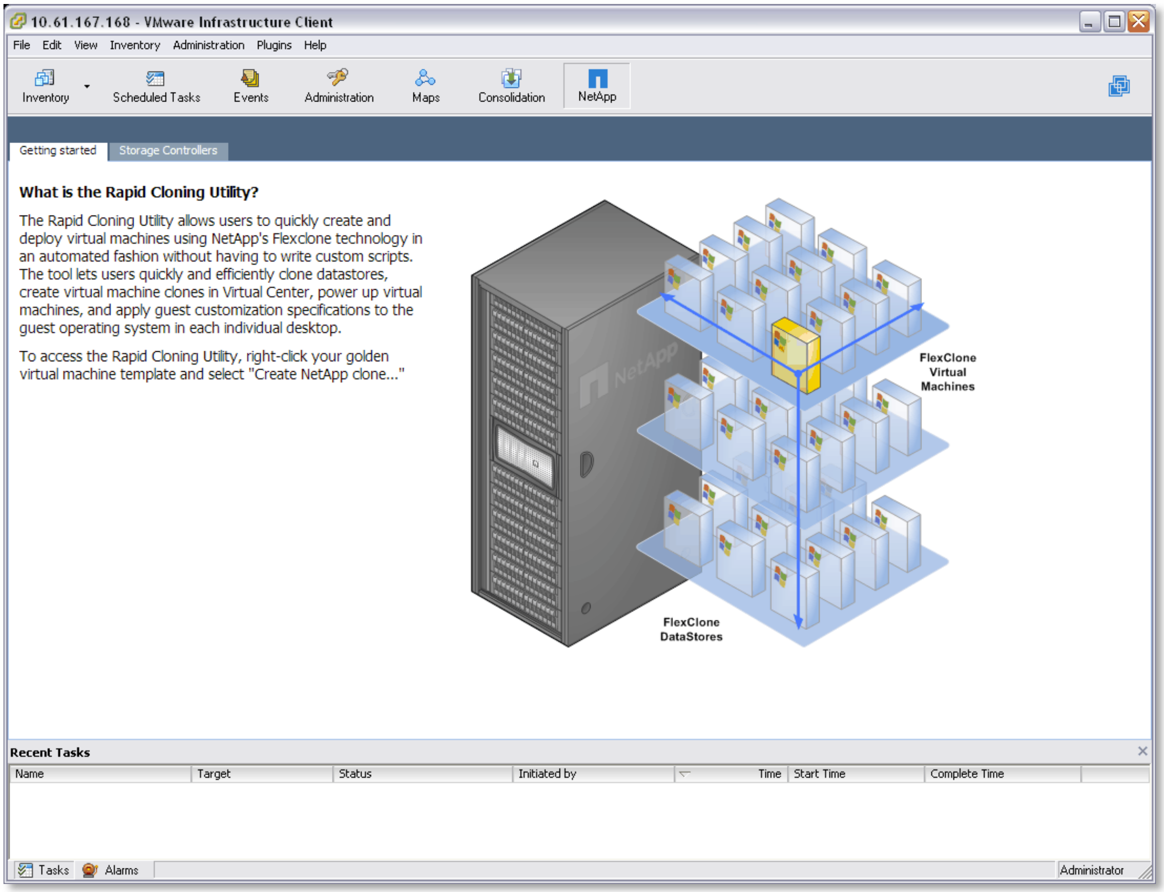
DEDUPLICATION CONSIDERATIONS WITH NFS

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

For deduplication best practices, including scheduling and performance considerations, see TR 3505 NetApp FAS Dedupe: Data Deduplication Deployment and Implementation Guide.

10.2 ZERO-COST VIRTUAL MACHINE CLONING

Customers who deploy VMware on NetApp can leverage NetApp's patented FlexClone® technology via NetApp's Rapid Cloning Utility. The RCU is a vCenter Server Plug-in that provisions zero-cost, or pre-deduplicated, VMware Virtual Machines and Datastores. The RCU is available to customers via NOW.



What is the Rapid Cloning Utility?

The Rapid Cloning Utility allows users to quickly create and deploy virtual machines using NetApp's Flexclone technology in an automated fashion without having to write custom scripts. The tool lets users quickly and efficiently clone datastores, create virtual machine clones in Virtual Center, power up virtual machines, and apply guest customization specifications to the guest operating system in each individual desktop.

To access the Rapid Cloning Utility, right-click your golden virtual machine template and select "Create NetApp clone..."

Figure 39) The NetApp Rapid Cloning Utility version 2.0.

10.3 STORAGE THIN PROVISIONING

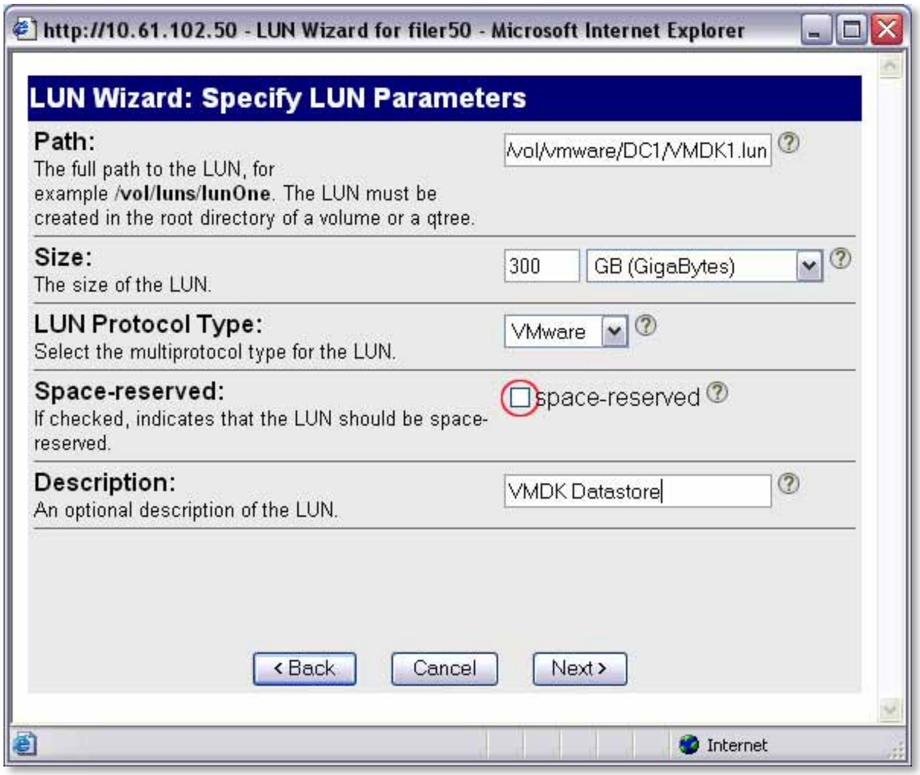
You should be very familiar with traditional storage provisioning and with the manner in which storage is preallocated and assigned to a server—or, in the case of VMware, a virtual machine. It is also a common practice for server administrators to overprovision storage in order to avoid running out of storage and the associated application downtime when expanding the provisioned storage. Although no system can be run at 100% storage utilization, there are methods of storage virtualization 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, there will not be enough storage to satisfy the requests.

NETAPP THIN-PROVISIONING OPTIONS

NetApp thin provisioning extends VMware thin provisioning for VMDKs and allows LUNs that are serving VMFS Datastores to be provisioned to their total capacity yet 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, follow these steps.

1	Open FilerView (http://filer/na_admin).
2	Select LUNs.
3	Select Wizard.
4	In the Wizard window, click Next.
5	Enter the path.
6	Enter the LUN size.
7	Select the LUN type (for VMFS select VMware; for RDM select the VM type).
8	Enter a description and click Next.
9	Deselect the Space-Reserved checkbox (see Figure 40).
10	Click Next and then click Finish.



http://10.61.102.50 - LUN Wizard for filer50 - Microsoft Internet Explorer

LUN Wizard: Specify LUN Parameters

Path: /vol/vmware/DC1/VMDK1.lun
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: 300 GB (GigaBytes)
The size of the LUN.

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

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

Description: VMDK Datastore
An optional description of the LUN.

Figure 40) Enabling thin provisioning on a LUN.

NetApp recommends that when you enable NetApp thin provisioning, you 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 they require it. The policies include automatic sizing of a volume, automatic Snapshot 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 environments, NetApp recommends setting this value to 'on'. Doing so requires setting the maximum volume and increment size options. To enable these options, follow these steps.

1	Log in to NetApp console.
2	Set 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 environments, NetApp recommends setting this value to delete Snapshot copies at 5% of available space. In addition, you should set the volume option to have the system attempt to grow the volume before deleting Snapshot copies. To enable these options, follow these steps.

1	Log in to NetApp console.
2	Set Snapshot Auto Delete Policy: snap autodelete <vol-name> commitment try trigger volume target_free_space 5 delete_order oldest_first.
3	Set Volume Auto Delete Policy: vol options <vol-name> try_first volume_grow.

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

1	Log in to NetApp console.
2	Set Volume Snapshot Fractional Reserve: vol options <vol-name> fractional_reserve 0.

11 VIRTUAL MACHINE BEST PRACTICES

11.1 WINDOWS VM FILE SYSTEM PERFORMANCE OPTION

OPTIMIZING WINDOWS FILE SYSTEM FOR OPTIMAL IO PERFORMANCE

If your virtual machine is not acting as a file server you may want to consider implementing the following change to your virtual machines, which will disable the access time updates process in NTFS. This change will reduce the amount of IOPs occurring within the file system. To make this change complete the following steps.

1.	Log into a Windows VM
2.	Select Start > Run and enter `CMD`
3.	Enter `fsutil behavior set disablelastaccess 1`

11.2 ENSURING VM AVAILABILITY

OPTIMIZING VM SCSI BUS

In Section 6 we covered the ESX Host Utilities. One of the components of the host utilities is the GOS timeout scripts, which are a collection of ISO images which can be mounted by a VM in order to configure its local SCSI to values which are optimal for running in a virtual infrastructure. It is recommended to run this process on your deployed VMs and VM templates.

To Install the GOS Timeout Scripts complete the following steps:

1	Download the EHU.
2	Copy the EHU to a location accessible to the ESX
3	Extract the EHU by running tar –zxf <name of EHU file>.tar.gz
4	From within vCenter Server select a VM to upgrade, right click on it, and select edit settings
5	Select CDROM, and the ISO radio button
6	Select the appropriate ISO, matching the OS of the VM your are configuring
7	Select OK
8	Connect the to VM console.
9	Run the script for the OS of the VM.
10	Exit and unmount the ISO image
11	Repeat as necessary for each VM

11.3 ENSURING VIRTUAL MACHINE DISK I/O

ALIGNMENT OF VM PARTITIONS & 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. In order to ensure optimal disk I/O within the VM one must 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 of these items will result in a dramatic increase of I/O load on a storage array and will negatively impact the performance of all Virtual Machines being served on the array.

It is the recommendation of NetApp, VMware, other storage vendors, and VMware Partnes that the partitions of VMs and the partitions of VMFS datastores are to be aligned to the blocks of the underlying storage array. One can find more information around VMFS and GOS file system alignment the following documents from various vendors:

VMware: Recommendations for Aligning VMFS Partitions

IBM: Storage Block Alignment with VMware Virtual Infrastructure

EMC: Celerra IP Storage with VMware Virtual Infrastructure

Dell: Designing and Optimizing SAN Configurations

EMC: CLARiiON Integration with VMware ESX Server

Vizioncore: vOptimizer Pro FAQ

Links to all documents can be found in the References section of this document.

DATASTORE ALIGNMENT

NetApp systems automate the alignment of VMFS with NetApp iSCSI & FC LUNs. This task is automated suring the LUN provisioning phase of creating a datastore when one selects the LUN type 'VMware' for the LUN. Customers deploying VMware over NFS do not need to align the datastore. With any type of datastore, VMFS or NFS, the virtual disks contained within should have the partitions aligned to the blocks of the storage array.

VIRTUAL MACHINE PARTITION ALIGNMENT

When aligning the partitions of virtual disks for use with NetApp FAS systems, the starting partition offset must be divisible by 4096. As an example, the starting partition offset for Microsoft Windows 2000, 2003, & 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 and Vista operating systems automatically have their starting partitions set to 1048576. By default this value and 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, than it is highly probable that these images require partition alignment.

11.4 THE IMPACT OF PARTITION MISALIGNMENT

Failure to properly align the file systems within Virtual Machines has a negative impact on many aspects of a Virtual Infrastructure. Customers may first notice the impact of misalignment with virtual machines running high performance applications. The reason for this is every I/O operation executed within the VM will require multiple I/O operations on the storage array.

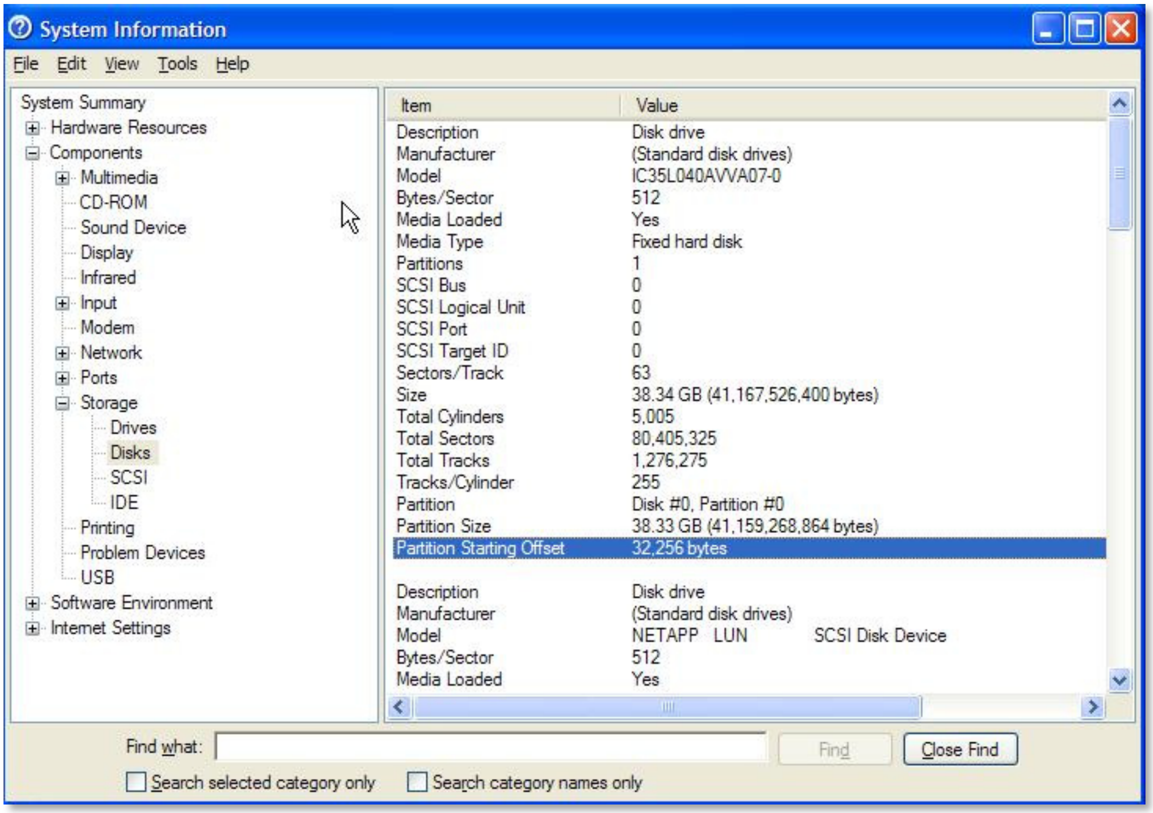
In addition to the negative performance impact storage savings with NetApp Data Deduplication will be negatively impacted, reducing the total amount of storage savings.

Finally, storage arrays will be over taxed and as the Virtual Infrastructure grows the storage array will require hardware upgrades in order to meet the additional I/O load generated by this misalignment. Simply put, one can save their company a significant amount of money by optimizing the I/O of their VMs.

11.5 IDENTIFYING PARTITION ALIGNMENT

VERIFYING PARTITION ALIGNMENT WITH WINDOWS OPERATING SYSTEMS

To verify the starting partition offset for a windows based virtual machine log onto the VM and run the System Information utility (or msinfo32). There you will be able to find this setting). To run msinfo32, select Start > All Programs > Accessories > System Tools > System Information. (see Figure 41)



Item	Value
Description	Disk drive
Manufacturer	(Standard disk drives)
Model	IC35L040AVVA07-0
Bytes/Sector	512
Media Loaded	Yes
Media Type	Fixed hard disk
Partitions	1
SCSI Bus	0
SCSI Logical Unit	0
SCSI Port	0
SCSI Target ID	0
Sectors/Track	63
Size	38.34 GB (41,167,526,400 bytes)
Total Cylinders	5,005
Total Sectors	80,405,325
Total Tracks	1,276,275
Tracks/Cylinder	255
Partition	Disk #0, Partition #0
Partition Size	38.33 GB (41,159,268,864 bytes)
Partition Starting Offset	32,256 bytes
Description	Disk drive
Manufacturer	(Standard disk drives)
Model	NETAPP LUN SCSI Disk Device
Bytes/Sector	512
Media Loaded	Yes

Figure 41) Using system information to identify the starting partition offset.

NETAPP MBRTOOLS: IDENTIFICAITON OF PARTITON ALIGNEMENT STATUS

NetApp provides a tool named MBRScan, which runs on an ESX host and can identify if 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. While this process ronly equires 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 it may be easier to identify the file system alignment from within each VM as this action is non-disruptive.

MBRScan is an integrated component of the VMware ESX Host Utilities 5.0 and is available as a stand-alone utility available in the NOW tool chest.

11.6 CORRECTIVE ACTIONS FOR VMS WITH MISALIGNED PARTITONS

BEGIN BY CORRECTING THE VM TEMPLATES

Once you have identified that you have misaligned partitions with your virtual machines it is recommended that the first corrective action be to correct the partitions in your templates. This step will ensure that any newly created VM will be properly aligned and will not add to the added I/O load on the storage array.

CORRECTING PARTITON MISALIGNMENT WITH NETAPP MBRTOOLS

NetApp provides a tool named MBRAlign, which runs on an ESX host and can correct and misaligned primary and secondary Master Boot Record based partitions. MBRAlign requires the Virtual Machine that is undergoing the corrective action to be powered off.

MBRAlign provides flexible repair options. For example, it can be used to migrate and align a virtual disk as well as change the format from thin to think vmdk. It is highly recommended to take a NetApp snapshot prior to executing MBRAlign. Once a VM has been corrected, powered on, and the results verified then this snapshot can be safely discarded.

MBRAlign can obtained from the NOW tool chest. It is recommended that you contact the NetApp Global Support Center so they can assist as you implement the corrective actions.

Note: Linux VMs which boot using the GRUB boot loader require the following steps after MBRAlign has been ran.

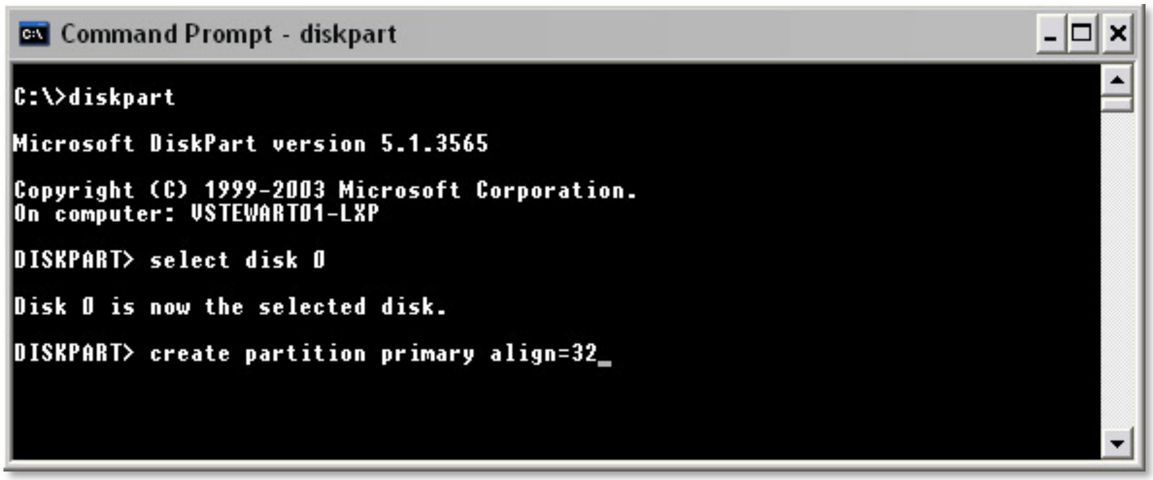
1.	Connect a LINUX CD or CDROM ISO image to the
2.	Boot the VM
3.	Select to boot from the CD
4.	When appropriate execute GRUB setup to repair the boot loader

11.7 CREATE PROPERLY ALIGNED PARTITONS FOR NEW VMS

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, follow these steps and see Figure 42.

1.	Boot the VM with the BartPE or a Microsoft WinPE CD.
2.	Select Start > Run and enter `DISKPART`.
3.	Enter Select `Disk0`.
4.	Enter `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: VSTEWART01-LXP

DISKPART> select disk 0

Disk 0 is now the selected disk.

DISKPART> create partition primary align=32_
```

Figure 42) Running diskpart to set a proper starting partition offset.

CREATING A PORPERLY ALIGNED VMDK FOR A NEW VM WITH FDISK

This procedure works for VMware vmdk files hosted on VMFS or NFS datastores, for both Windows and Linux VMs. Please note that this procedure is not required for Windows Server 2008 and Vista VMs, which are aligned by default. To set up the starting offset using the fdisk command in the ESX service console, follow these steps.

Step	Action
1.	Log in to the ESX service console.

2.	CD to the VM directory and view this directory by typing the following commands (shown below):
	`cd /vmfs/volumes/<datastore>/<VM home dir>`
	`ls –l`

3.	Indentify the number of cylinders in the virtual disk by reading the virtual disk descriptor file. Look for the line ddb.geometery.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.	Once 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 should be 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 asked for the partition type.
11.	Enter 1 for the partition number, 1 for the first cylinder, and press Enter for the last cylinder question to make it use 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 entering 64 and pressing Enter. Note the value 64 represents the number of 512 bytes used to create a starting offset of 32,768 KB.
14.	Check the partition table by typing p. If you did this correctly the top row of the output should display`disk geometry including the starting offset of 64.
15.	Type r to return to the regular 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, as this is normal.
19.	Start the VM and run Windows setup. During the installation process you will be prompted that a partition exists. Select this partition to format and install Windows into.

12 VIRTUAL MACHINE STORAGE LAYOUT

12.1 DEFAULT VIRTUAL MACHINE LAYOUT

When a virtual machine is provisioned the VMware 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 al of the files for a single VM will 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, etc.

From the standpoint of simplicity, this design works well where a VM home directory is a Virtual Machine. See the image below for a high level conceptual view of this layout.

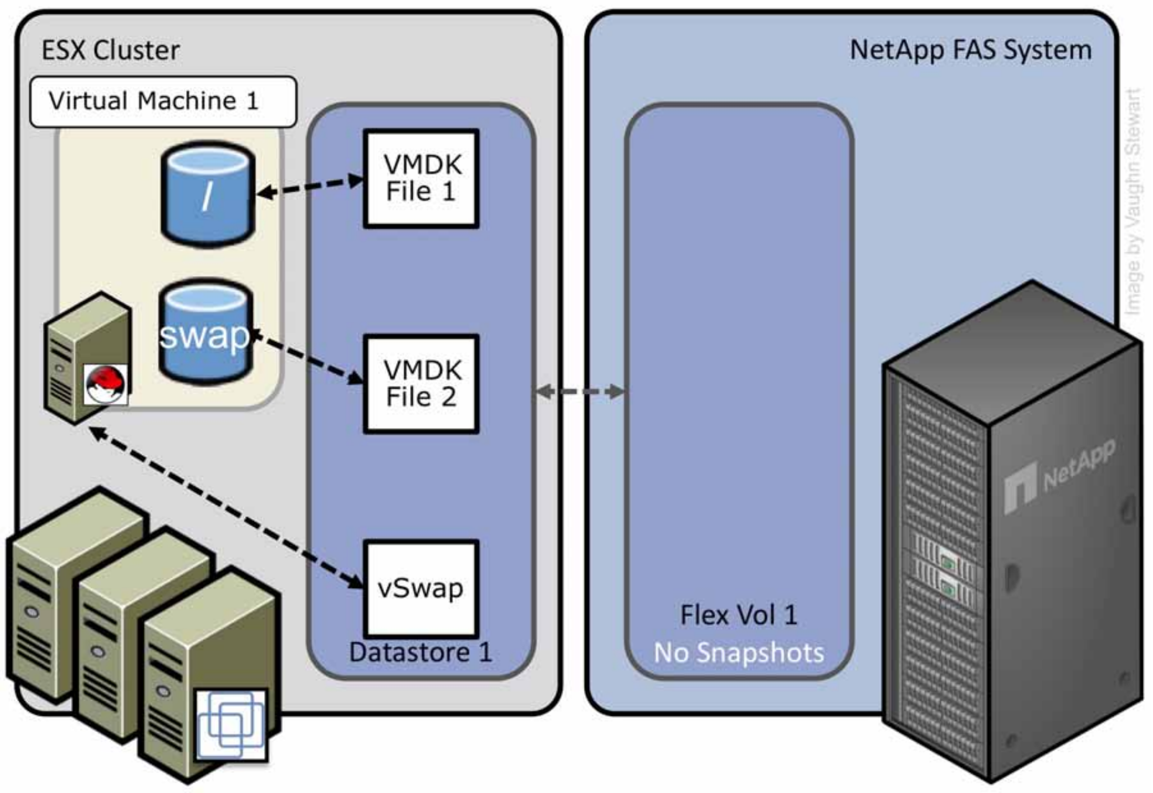


Figure 43) VMware's default Virtual Machine layout

12.2 VIRTUAL MACHINE LAYOUT WITH NETAPP SNAP* TECHNOLOGIES

In this section we will review a data layout design which is recommended when integrating VMware with NetApp snap* technologies such as SnapManager snapshot backups or disk-to-disk replication via SnapMirror and/or SnapVault. In these use case scenarios NetApp recommends separating transient and temporary data from the production data by implementing architecture that separates these two data types into multiple datastores.

It should be noted that this design is not NetApp specific, but instead is an optimal consideration when deploying VMware 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 of these files, and as such will consume a substantial amount of additional disk and/or bandwidth if the temporary and transient data is not separated from the production data.

RECOMMENDED LAYOUT OPTION 1: IMPLEMENT A CENTRAL VIRTUAL SWAP DATASTORE

ESX Servers create a virtual swap or 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 via Site Recovery Manager; NetApp recommends that the virtual swap file for every Virtual machines should be relocated from the VM home directory to a datastore, on a separate NetApp volume which is dedicated to storing virtual swap files. See the image below for a high level conceptual view of this layout.

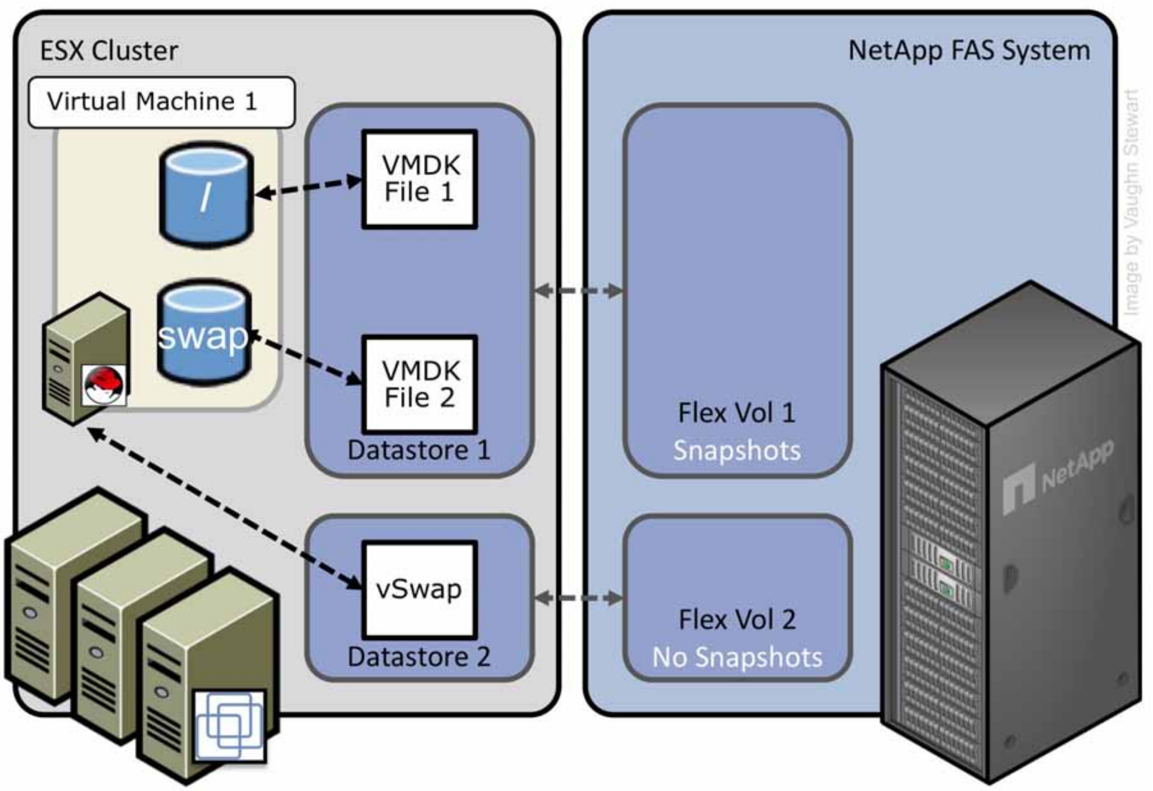


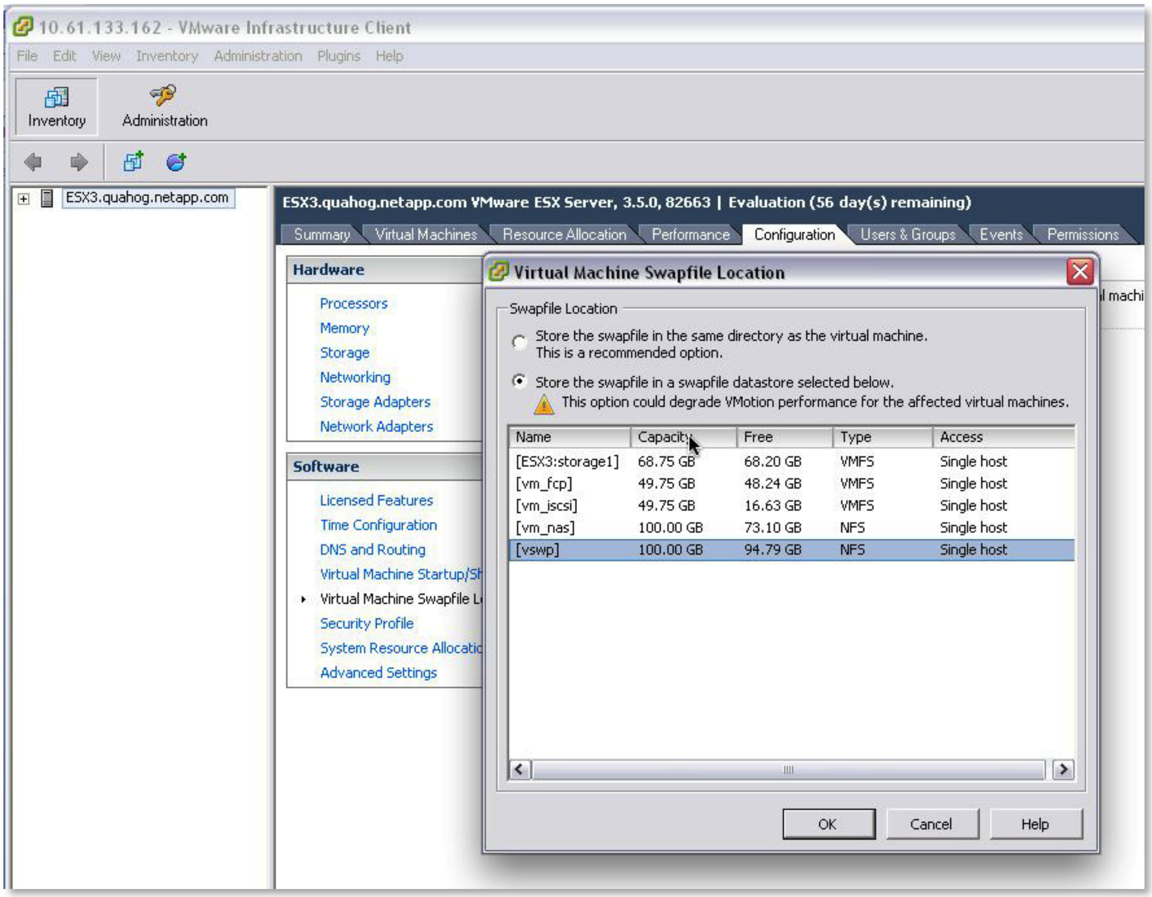
Figure 44) Recommended Virtual Machine layout option 1: a central vswap datastore

A prerequisite to making this change is the creation of a datastore to store the swap files. Because the VMware swap file storage requirements are dynamic, NetApp suggests 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 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: VMware has documented that the following options must not reside in the VMX file in order to use a centralized vswap datastore: sched.swap.dir or sched.swap.derivedName

To configure a central datastore to store the virtual swap files follow these steps (and refer to Figure 45).

1	Open vCenter Server.
2	Select an ESX Server.
3	In the right pane, select the Configuration tab.
4	In the Software box, select Virtual Machine Swapfile Location.
5	In the right pane, select edit.
6	The Virtual Machine Swapfile Location Wizard will open (Figure 46)
7	Select the Datastore which will be the global location
8	Repeat steps 2 through 7 for each ESX Server in the cluster.
9	This process configures the ESX Server and does not affect existing VMs. See the next table for configuring existing VMs.



Swapfile Location

○ Store the swapfile in the same directory as the virtual machine.
This is a recommended option.

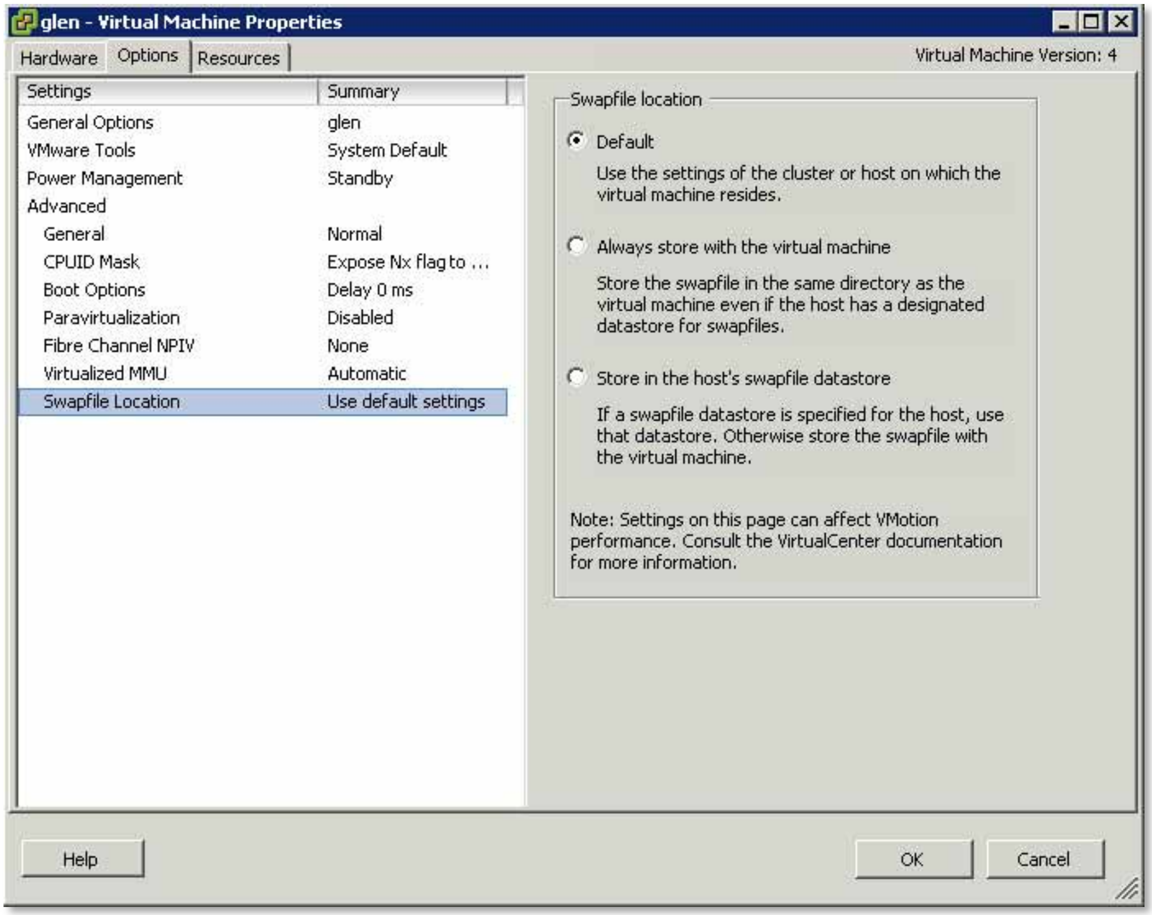
● Store the swapfile in a swapfile datastore selected below.
⚠ This option could degrade VMotion performance for the affected virtual machines.

Name	Capacity	Free	Type	Access
[ESX3:storage1]	68.75 GB	68.20 GB	VMFS	Single host
[vm_fcp]	49.75 GB	48.24 GB	VMFS	Single host
[vm_iscsi]	49.75 GB	16.63 GB	VMFS	Single host
[vm_nas]	100.00 GB	73.10 GB	NFS	Single host
[vswp]	100.00 GB	94.79 GB	NFS	Single host

Figure 45) 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, follow these steps (and refer to Figure 48).

1	Open vCenter Server.
2	Select 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 the Default radio button
9	To make this change take effect either… Migrate each VM to an ESX Server which is configured with a central vswap datastore or… Restart each VM on the existing ESX Server which is configured with a central vswap datastore
10	Repeat steps 2 through 9 for each existing VM.



Virtual Machine Version: 4

Settings	Summary
General Options	glen
VMware Tools	System Default
Power Management	Standby
Advanced	
General	Normal
CPUID Mask	Expose Nx flag to ...
Boot Options	Delay 0 ms
Paravirtualization	Disabled
Fibre Channel NPIV	None
Virtualized MMU	Automatic
Swapfile Location	Use default settings

Swapfile location

● Default

Use the settings of the cluster or host on which the virtual machine resides.

○ Always store with the virtual machine

Store the swapfile in the same directory as the virtual machine even if the host has a designated datastore for swapfiles.

○ Store in the host's swapfile datastore

If a swapfile datastore is specified for the host, use that datastore. Otherwise store the swapfile with the virtual machine.

Note: Settings on this page can affect VMotion performance. Consult the VirtualCenter documentation for more information.

Figure 46) Defining the location for virtual swap files for a VM.

RECOMMENDED LAYOUT OPTION 2: IMPLEMENT A CENTRAL VIRTUAL SWAP DATASTORE AND RELOCATE VM SWAP/PAGEFILE TO A SECOND DATASTORE

This design layout build off of the layout option number 1 except in this design we are relocating the Virtual Machine's swap or pagefile in an alternative datastore. This design has pros and cons, which should be understood prior to implementing. These details are covered after we review the architecture.

Each VM creates a swap or pagefile that is typically 1.5 to 2 times the size of the amount of memory configured for each VM. As this data is transient in nature we can save a fair amount of storage and/or bandwidth capacity by removing this data out of the datastore, which contains the production data. In order 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 NetApp volume. See the image below for a high level conceptual view of this layout.

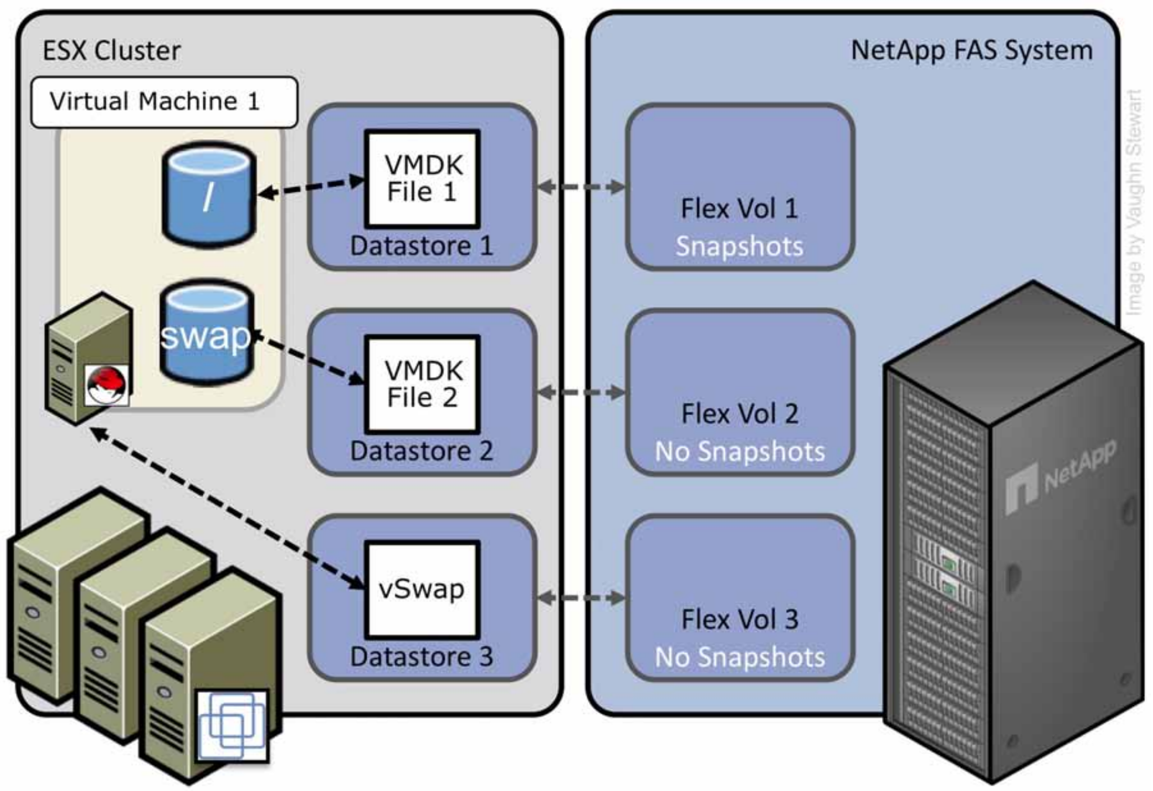


Figure 47) Recommended Virtual Machine layout option 2: separate VM swap/pagefile & a central vswap datastore

As stated earlier, there are pros and cons to this design. The benefits are no temporary and transient data will be contained in either a Snapshot backup or replicated data set, thus conserving a large amount of storage.

This design has a negative impact on customers who implement VMware Site Recovery Manager. In this design the entire VM is not being replicated, so customers will have to configure a second VMDK for each VM in their SRM recovery plan. For more information on the details of this design with SRM please see TR-3671: VMware Site Recovery Manager in a NetApp Environment.

13 STORAGE MONITORING AND MANAGEMENT

13.1 MONITORING STORAGE UTILIZATION WITH NETAPP OPERATIONS MANAGER

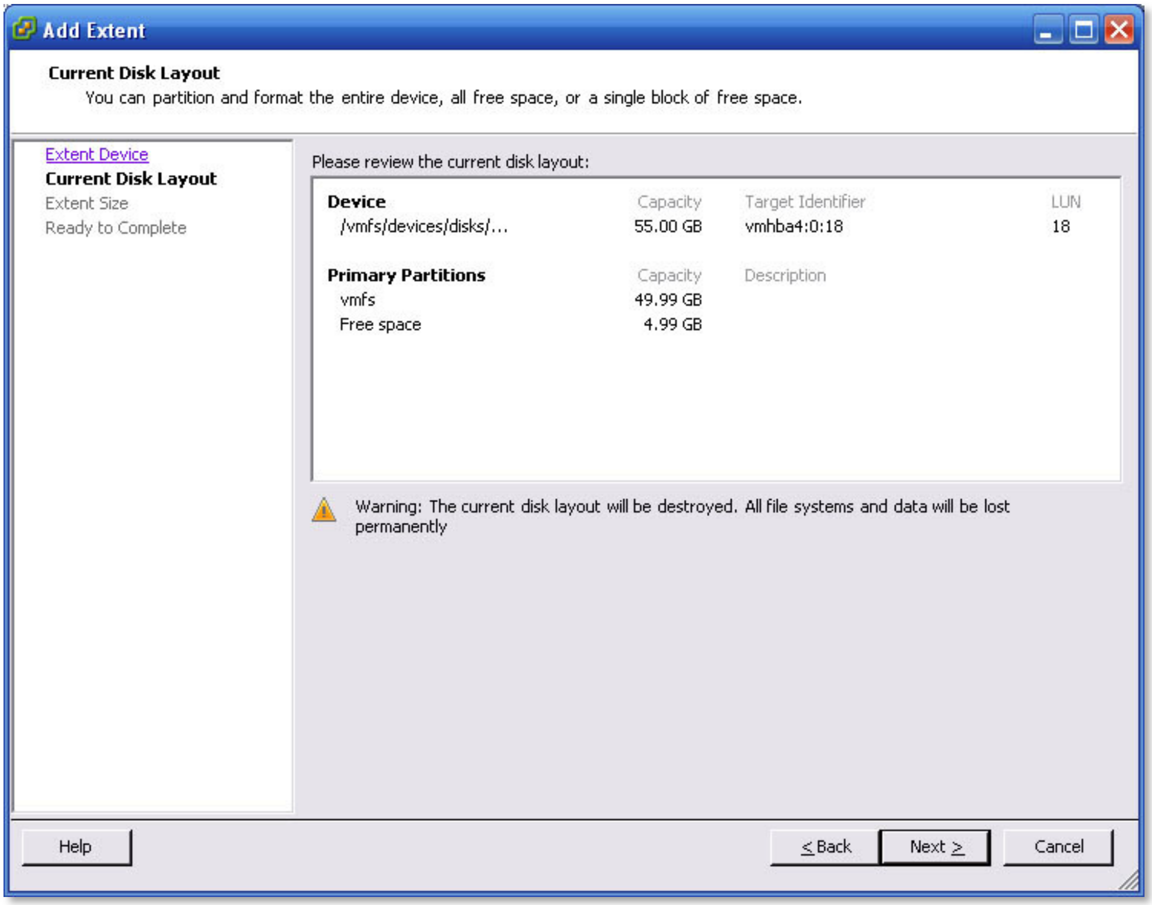
NetApp Operations Manager monitors, manages, and generates reports on all of the NetApp FAS systems in an organization. When you are using NetApp thin provisioning, NetApp recommends deploying Operations Manager and setting up e-mail and pager notifications to the appropriate administrators. With thin-provisioned storage, it is very important to monitor the free space available in storage aggregates. Proper notification of the available free space means that additional storage can be made available before the aggregate becomes completely full. For more information about setting up notifications in DataFabric® Manager Server: Operations Manager Administration Guide.

13.2 STORAGE GROWTH MANAGEMENT

GROWING VMFS DATASTORES

It is quite easy to increase the storage for a VMFS Datastore; however, this process should be completed only when all virtual machines stored on the Datastore are shut down. For more information, see the VMware white paper “VMFS Technical Overview and Best Practices.” To grow a Datastore, follow these steps.

1	Connect to the FAS system console (via either SSH, Telnet, or Console connection).
2	Select LUNs.
3	Select Manage.
4	In the left pane, select the LUN from the list.
5	Enter the new size of the LUN in the Size box and click Apply.
6	Open vCenter Server.
7	Select an ESX host.
8	Make sure that all VMs on the VMFS Datastore are shut down
9	In the right pane, select the Configuration tab.
10	In the Hardware box, select the Storage Adapters link.
11	In the right pane, select the HBAs and then select the Rescan link.
12	In the Hardware box, select the Storage link.
13	In the right pane right, select the Datastore to grow and then select Properties.
14	Click Add Extent.
15	Select the LUN and click Next, then click Next again. As long as the window shows free space available on the LUN, you can ignore the warning message (see Figure 48).
16	Make sure that the Maximize Space checkbox is selected, click Next, and then click Finish.



Current Disk Layout

You can partition and format the entire device, all free space, or a single block of free space.

Please review the current disk layout:

Device	Capacity	Target Identifier	LUN
/vmfs/devices/disks/...	55.00 GB	vmhba4:0:18	18

Primary Partitions	Capacity	Description
vmfs	49.99 GB	
Free space	4.99 GB	

Warning: The current disk layout will be destroyed. All file systems and data will be lost permanently

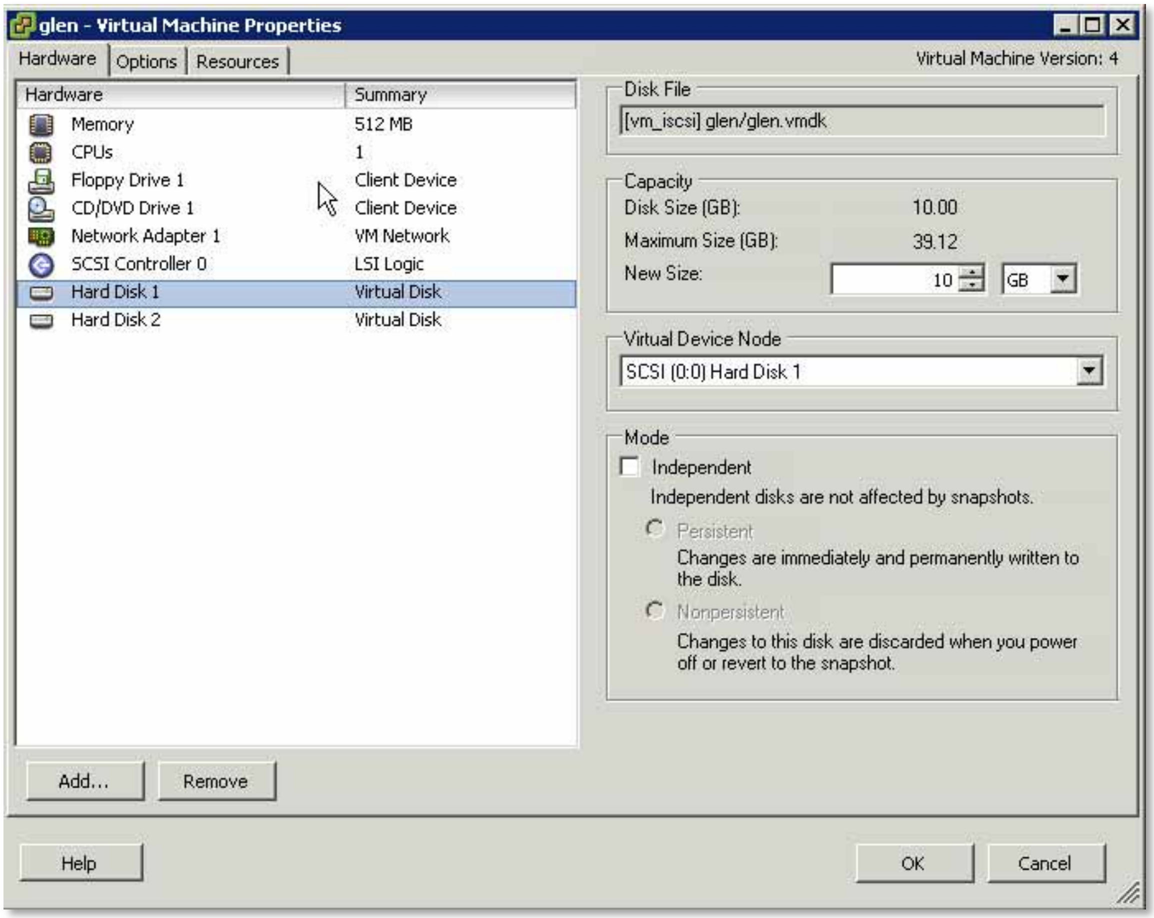
Figure 48) Expanding a VMFS partition.

For more information about adding VMFS extents, see the VMware ESX and ESXi Server Configuration Guide.

GROWING A VIRTUAL DISK (VMDK)

Virtual disks can be extended; however, this process requires the virtual machine to be powered off. Growing the virtual disk is only half of the equation for increasing available storage; you still need to grow the file system after the VM boots. Root volumes such as C:\ in Windows and / in Linux cannot be grown dynamically or while the system is running. For these volumes, see “Growing Bootable Volumes,” later in this report. For all other volumes, you can use native operating system tools to grow the volume. To grow a virtual disk, follow these steps.

1	Open 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 49)
5	Start the VM.



glen - Virtual Machine Properties

Virtual Machine Version: 4

Hardware	Summary
Memory	512 MB
CPUs	1
Floppy Drive 1	Client Device
CD/DVD Drive 1	Client Device
Network Adapter 1	VM Network
SCSI Controller 0	LSI Logic
Hard Disk 1	Virtual Disk
Hard Disk 2	Virtual Disk

Disk File

[vm_iscsi] glen/glen.vmdk

Capacity

Disk Size (GB): 10.00

Maximum Size (GB): 39.12

New Size: 10 GB

Virtual Device Node

SCSI (0:0) Hard Disk 1

Mode

□ Independent

Independent disks are not affected by snapshots.

○ Persistent

Changes are immediately and permanently written to the disk.

○ Nonpersistent

Changes to this disk are discarded when you power off or revert to the snapshot.

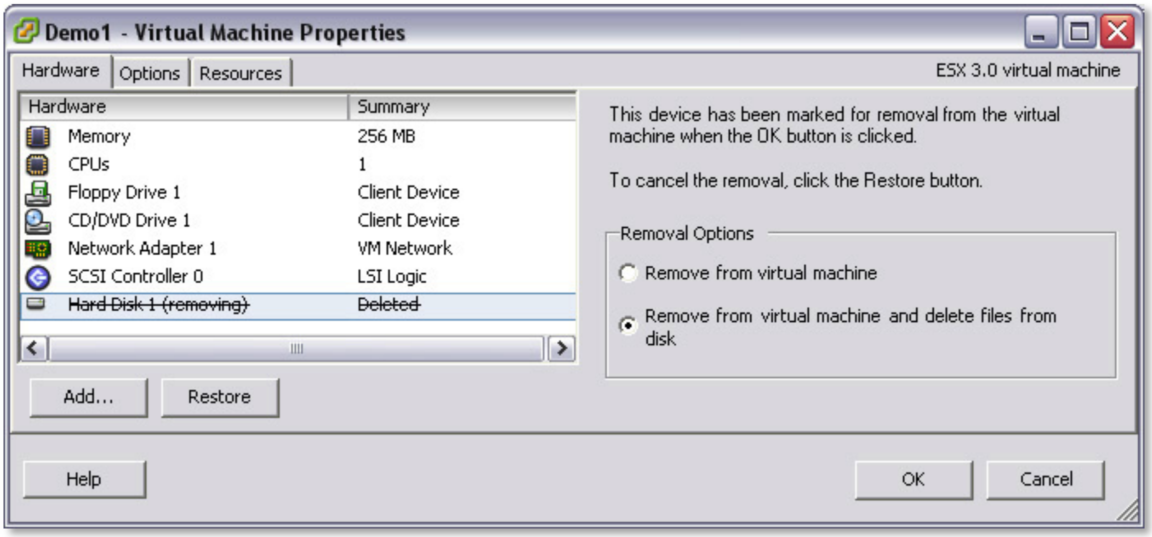
Figure 49) Increasing the size of a virtual disk.

For more information about extending a virtual disk, see VMware ESX and ESXi Server Configuration Guide.

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 base storage, follow these steps.

1	Open vCenter Server.
2	Select an 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 the Remove from Virtual Machine radio button and select Delete Files from Disk. This action deletes the Mapping File but does not remove any data from the RDM LUN (see Figure 50).
5	Open FilerView (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 vCenter Server.
11	In the right pane, select the Configuration tab.
12	In the Hardware box, select the Storage Adapters link.
13	In the right pane, select the HBAs and select the Rescan link.
14	Right-click the VM and select Edit Settings to open the Edit Settings window,
15	Click Add, select Hard Disk, and then click Next. (see Figure 51).
16	Select the LUN and click Next (see Figure 52).
17	Specify the VMFS Datastore that will store the Mapping file.
18	Start the VM. Remember that although you have grown the LUN, you still need to grow the file system within it. Follow the guidelines in “Growing a VM File System,” next.



ESX 3.0 virtual machine

Hardware	Summary
Memory	256 MB
CPUs	1
Floppy Drive 1	Client Device
CD/DVD Drive 1	Client Device
Network Adapter 1	VM Network
SCSI Controller 0	LSI Logic
~~Hard Disk 1 (removing)~~	~~Deleted~~

This device has been marked for removal from the virtual machine when the OK button is clicked.

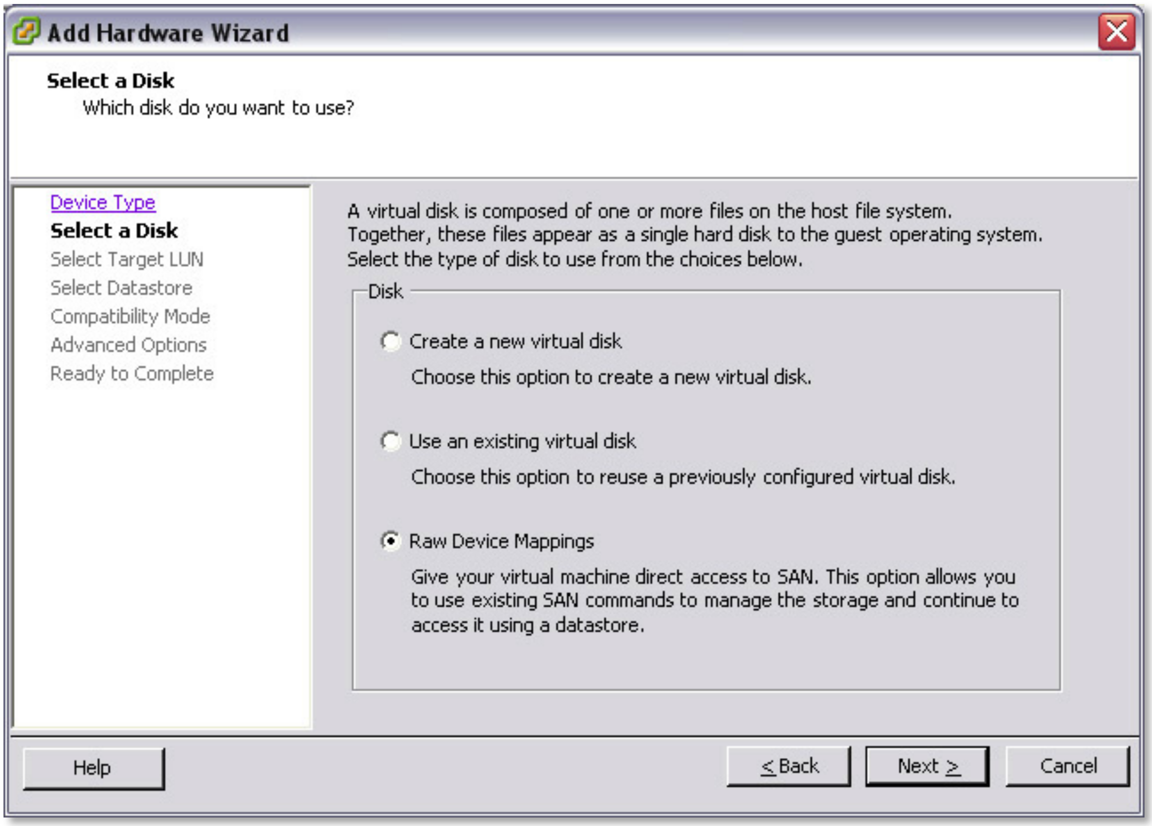
To cancel the removal, click the Restore button.

Removal Options

○ Remove from virtual machine

● Remove from virtual machine and delete files from disk

Figure 50) Deleting a VMDK from a VM.



Add Hardware Wizard

Select a Disk

Which disk do you want to use?

A virtual disk is composed of one or more files on the host file system. Together, these files appear as a single hard disk to the guest operating system. Select the type of disk to use from the choices below.

Disk

○ Create a new virtual disk

Choose this option to create a new virtual disk.

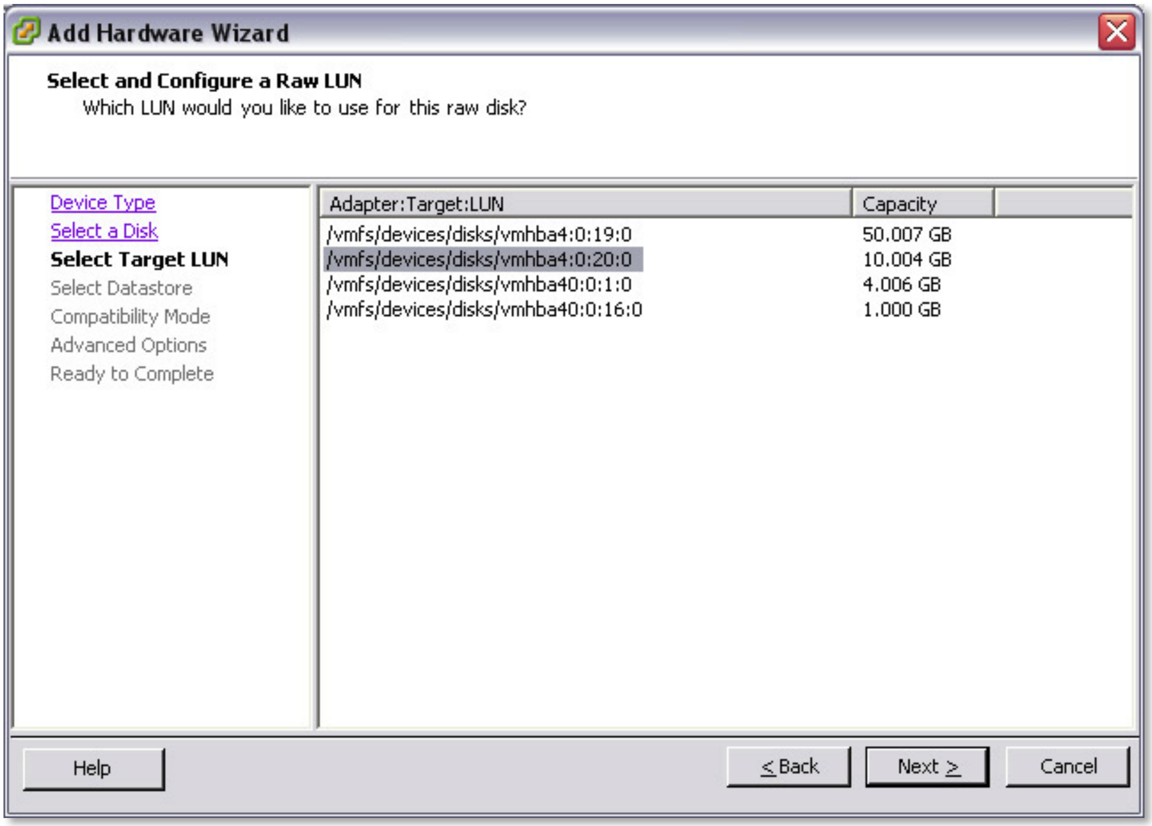
○ Use an existing virtual disk

Choose this option to reuse a previously configured virtual disk.

● Raw Device Mappings

Give your virtual machine direct access to SAN. This option allows you to use existing SAN commands to manage the storage and continue to access it using a datastore.

Figure 51) Connecting an RDM to a VM.



Select and Configure a Raw LUN

Which LUN would you like to use for this raw disk?

Adapter:Target:LUN	Capacity
/vmfs/devices/disks/vmhba4:0:19:0	50.007 GB
/vmfs/devices/disks/vmhba4:0:20:0	10.004 GB
/vmfs/devices/disks/vmhba40:0:1:0	4.006 GB
/vmfs/devices/disks/vmhba40:0:16:0	1.000 GB

Figure 52) Selecting a LUN to mount as an RDM.

GROWING A FILE SYSTEM WITHIN A GUEST OPERATING SYSTEM (NTFS OR EXT3)

When a virtual disk or RDM has been increased in size, you still need to grow the file system residing on it after booting the VM. This process can be done live while the system is running, by using native or freely distributed tools.

1	Remotely connect to the VM.
2	Grow the file system.
3	For Windows VMs, you can use the diskpart utility to grow the file system. For more information, see http://support.microsoft.com/default.aspx?scid=kb;en-us;300415. Or For Linux VMs, you can use ext2resize to grow the file system. For more information, see http://sourceforge.net/projects/ext2resize.

GROWING BOOTABLE VOLUMES WITHIN A GUEST OPERATING SYSTEM

Root volumes such as C:\ in Windows VMs and / in Linux VMs cannot be grown on the fly or while the system is running. There is a simple way to expand these file systems that 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 using the processes defined earlier. Once 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.

14 DISK BASED SNAPSHOT BACKUPS FOR VMWARE

14.1 COMPLEMENTARY SNAPSHOT TECHNOLOGIES

VMware Virtual Infrastructure 3 introduced the ability to 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. NetApp has been providing customers with the ability to create Snapshot copies of their data since 1992, and although the basic concept of a snapshot is similar between NetApp and VMware, you should be aware of the differences between the two, and when you should use one rather than the other.

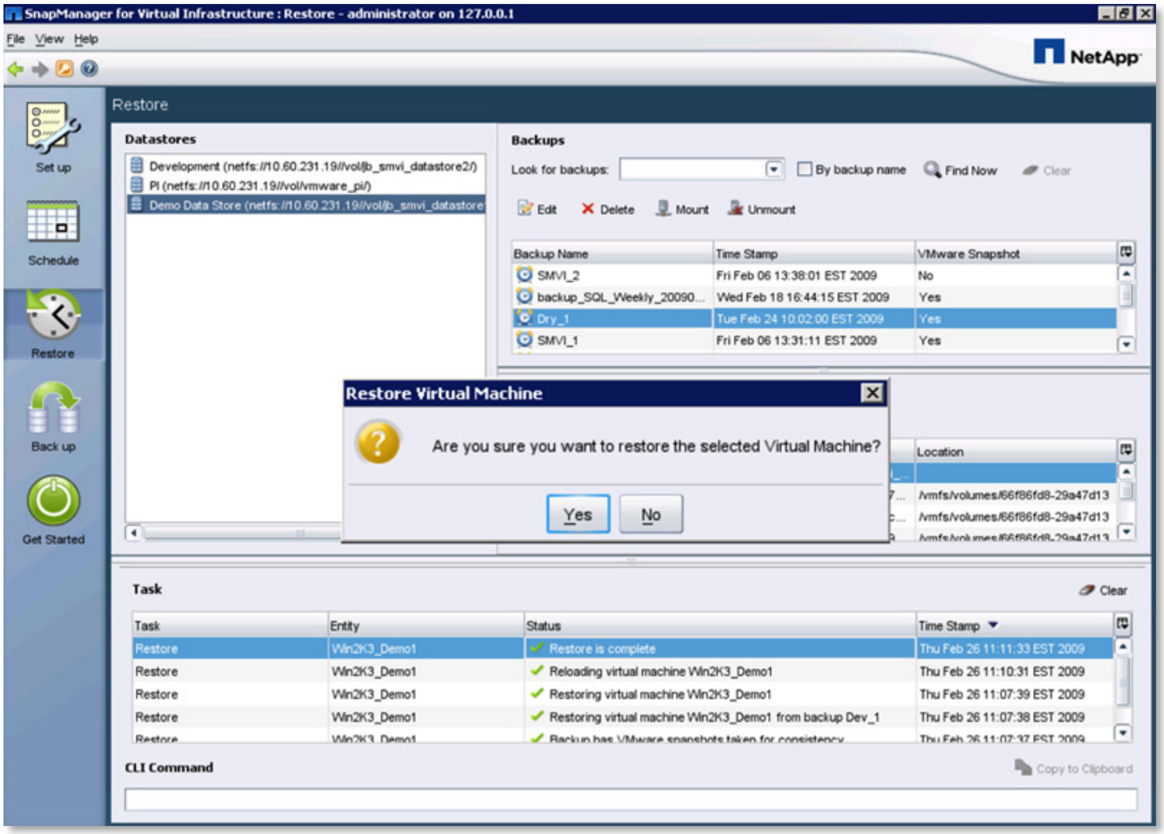
VMware snapshots provide simple point-in-time versions of VMs, allowing quick recovery. The benefits of VMware snapshots are that they are easy to create and use, because they can be executed and scheduled from within vCenter Server. VMware suggests that the snapshot technology in ESX should not be leveraged as a means to back up Virtual Infrastructure. For more information about native VMware snapshots, including usage guidelines, see the VMware Basic System Administration Guide for more information.

NetApp Snapshot technology can easily be integrated into VMware 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. Unlike traditional snapshots the NetApp technology does not have negative impact 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, requiring external scripting and/or scheduling to manage the process. For details, see the VMware Basic System Administration Guide and the VMware ESX and ESXi Server Configuration Guide.

14.2 IMPLEMENTING NETAPP SNAPSHOT BACKUPS FOR VMWARE VIRTUAL INFRASTRUCUTRE

The ability to quickly backup tens of Virtual machines without impact to production operations can accelerate the adoption of VMware within an organization. NetApp offers a means to do this with SnapManager for Virtual Infrastructure (or SMVI). SMVI builds on NetApp's SnapManager portfolio by providing array based backups which only consume block level changes to each VM, can provide multiple recovery points through out the day, and as the backups are an integrated component within the storage array SMVI provides recovery times faster than tradtion solutions which require the data to be copied from a second source.



Development (netfs://10.60.231.19//vol/jb_smvi_datastore2/)
PI (netfs://10.60.231.19//vol/vmware_pi/)
Demo Data Store (netfs://10.60.231.19//vol/jb_smvi_datastore

Backup Name	Time Stamp	VMware Snapshot
SMVI_2	Fri Feb 06 13:38:01 EST 2009	No
backup_SQL_Weekly_20090...	Wed Feb 18 16:44:15 EST 2009	Yes
Dry_1	Tue Feb 24 10:02:00 EST 2009	Yes
SMVI_1	Fri Feb 06 13:31:11 EST 2009	Yes

Are you sure you want to restore the selected Virtual Machine?

Location
/vmfs/volumes/66f86fd8-29a47d13
/vmfs/volumes/66f86fd8-29a47d13
/vmfs/volumes/66f86fd8-29a47d13

Task	Entity	Status	Time Stamp
Restore	Win2K3_Demo1	✓ Restore is complete	Thu Feb 26 11:11:33 EST 2009
Restore	Win2K3_Demo1	✓ Reloading virtual machine Win2K3_Demo1	Thu Feb 26 11:10:31 EST 2009
Restore	Win2K3_Demo1	✓ Restoring virtual machine Win2K3_Demo1	Thu Feb 26 11:07:39 EST 2009
Restore	Win2K3_Demo1	✓ Restoring virtual machine Win2K3_Demo1 from backup Dev_1	Thu Feb 26 11:07:38 EST 2009
Restore	Win2K3_Demo1	✓ Backup has VMware snapshots taken for consistency	Thu Feb 26 11:07:37 EST 2009

Figure 53) Restoring a VM in SnapManager for Virtual Infrastructure.

For more information see the SnapManager for Virtual Infrastructure best practices TR-3737.

14.3 USING VMWARE SNAPSHOTS WITH NFS DATASTORES

As discussed earlier in this section VMware ESX Server provides a snapshot mechanism. This mechanism is used by many technologies including, but not limited to, Storage VMotion, Consolidated Backup, SnapManager for Virtual Infrastructure, The process to delete a VMware snapshot (or commit the VMware snapshot log files) can encounter an issue where the I/O of a GOS is suspended for an extended period of time when the Datastores are connected by NFS.

VMware snapshots are being used for virtual machines on the NFS Datastores. This usage includes, but is not limited to VMware Snapshots, Storage VMotion, Consolidated Backup, NetApp SnapManager for Virtual Infrastructure, etc.

This issue is resolved in ESX 3.5 update 3 and ESX 3.0.3; however, additional steps must be completed in order to enable this fix. The patch, ESX350-200808401-BG, has been created to address this issue with systems running ESX 3.5. When this patch is in use, there is a condition where virtual machines running 3rd party virtual machine management agents may get powered off unexpectedly. In order to avoid this behavior, please consult the support organization of the management agent regarding virtual disk pooling interval tuning. VMware patches available on the VMware support site.

PATCH PROCESS FOR SYSTEMS RUNNING ESX AND ESXI 3.5 UPDATE 3 OR LATER:

1	Using VMotion migrate the running VMs to other ESX nodes
2	Place the ESX server identified in step two into maintenance mode
3	Ensure the advanced configuration option for disabling NFS locks is set to 0
4	Backup the /etc/vmware/config file
5	Insert the following line into the /etc/vmware/config file `prefvmx.consolidateDeleteNFSLocks = "TRUE"`
6	Reboot the ESX server
7	After the reboot completes, login and exit maintenance mode
8	Migrate a VM back to the patched ESX server & verify the creation of a lock file in the root directory of this VM. *Lock files have the following file prefix: .lck*
9	If the lock file is created repeat steps 1 thru 7 on each ESX server
10	If the lock file is not created please ensure you followed every step in this process. If a lock is still not created, the VM ay require to be restarted

The following subset is a single process that streamlines some of the patching process. It requires one to log into an ESX server via the CLI. It replaces steps 2 thru 6 from the previous process. The following 5 lines must be run from the console of the ESX server as root.

1	vimsh -n -e /hostsvc/maintenance_mode_enter
2	esxcfg-advcfg -s 0 /NFS/LockDisable
3	cp -p /etc/vmware/config /etc/vmware/config.bak
4	echo ‘prefvmx.consolidateDeleteNFSLocks = "TRUE"’ >> /etc/vmware/config
5	reboot

After this process complete steps 7 thru 10 from the process to patch ESX 3.5 update 3 systems.

PATCH PROCESS FOR SYSTEMS RUNNING ESX AND ESXI 3.5 PRIOR TO UPDATE 3:

1	Apply patch ESX350-200808401-BG and refer to VMware patch best practices guide.
2	Identify an ESX server(s) to apply the patch to and ensure the systems the requirements of the patch
3	Using VMotion migrate the running VMs to other ESX nodes
4	Place the ESX server identified in step two into maintenance mode
5	Install the patch
6	Ensure the advanced configuration option for disabling NFS locks is set to 0
7	Backup the /etc/vmware/config file
8	Insert the following line into the /etc/vmware/config file `prefvmx.consolidateDeleteNFSLocks = "TRUE"`
9	Reboot the ESX server
10	After the reboot completes, login and exit maintenance mode
11	Migrate a VM back to the patched ESX server & verify the creation of a lock file in the root directory of this VM. *Lock files have the following file prefix: .lck*
12	If the lock file is created repeat steps 2 thru 10 on each ESX server requiring the patch
13	If the lock file is not created please ensure you followed every step in this process. If a lock is still not created, the VM ay require to be restarted

The following subset is a single process that streamlines some of the patching process. It requires one to log into an ESX server via the CLI. It replaces steps 5 thru 9 from the previous process. The following 5 lines must be run from the console of the ESX server as root.

1	vimsh -n -e /hostsvc/maintenance_mode_enter
2	esxcfg-advcfg -s 0 /NFS/LockDisable
3	cp -p /etc/vmware/config /etc/vmware/config.bak
4	echo ‘prefvmx.consolidateDeleteNFSLocks = "TRUE"’ >> /etc/vmware/config
5	reboot

After this process complete steps 10 thru 13 from the process to patch ESX 3.5 update 1 and update 2 systems.

15 TECHNICAL REPORT SUMMARY

VMware Virtual Infrastructure offers customers several methods of providing storage to virtual machines. All of these storage methods give customers flexibility in their infrastructure design, which in turn provides cost savings, increased storage utilization, and enhanced data recovery.

This technical report is not intended to be a definitive implementation or solutions guide. Expertise may be required to solve user-specific deployments. Contact your local NetApp representative to make an appointment to speak with a NetApp VMware solutions expert.

Comments about this technical report are welcome. Feel free to contact the authors by emailing xdl-vgibutmevmtr@netapp.com and please refer to TR-3428 in the subject line of your email.

16 APPENDIX A: CONFIGURING SSH ON ESX SERVERS AND NETAPP ARRAYS

16.1 CONFIGURING SSH ON ESX SERVERS AND NETAPP ARRAYS

The most efficient way to integrate NetApp Snapshot copies is to enable the centralized management and execution of Snapshot copies. NetApp recommends configuring the FAS systems and ESX Servers to allow a single host to remotely execute commands on both systems. This management host must have an SSH client installed and configured.

FAS SYSTEM SSH CONFIGURATION

To configure SSH access on a NetApp FAS system, follow these steps.

1	Connect to the FAS system console (via either SSH, Telnet, or Console connection).

2	Execute the following commands:
	`secureadmin setup ssh`
	`options ssh.enable on`
	`options ssh2.enable on`

3	Log in to the Linux or VMware system that remotely executes commands on the FAS system as root.
4	Add the Triple DES cipher to the list of available SSH ciphers; this is the only cipher recognized by the NetApp FAS 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 `/root/.ssh/id_dsa.pub`.
6	Mount the FAS root file system as root.
7	Copy only the key information from the public key file to the FAS 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. See the following example.

8	Test the connectivity from the remote host by issuing the `version` command on the FAS system. It should not prompt for a password:
	`ssh <netapp> version`
	`NetApp Release 7.2: Mon Jul 31 15:51:19 PDT 2006`

Example of the key for the remote host:

```
ssh-dsa AAAAB3NzaC1kc3MAAABhALVbwVyhtAVoaZukcjSTlRb/REO1/ywbQECtAcHijzdzhEJU
z9Qh96HVEwyZDdah+PTxfyitJCerb+1FAnO65v4WMq6jxPVYto6l5Ib5zxfq2I/hhT/6KPziS3LT
ZjKccwAAABUAjkLMwkpiPmg8Unv4fjCsYYhrSL0AAABgF9NsuZxniOOHHr8tmW5RMX+M6VaH/nlJ
UzVXbLiI8+pyCXALQ29Y31uV3SzWTd1VOgjJHgv0GBw8N+rvGSB1r60VqgqgGjSB+ZXAO1Eecbnj
vLnUtf0TVQ75D9auagjOAAAAYEJPx8wi9/CaS3dfKJR/tYy7Ja+MrlD/RCOgr22XQP1ydexsfYQx
enxzExPa/sPfjA45YtcUom+3mieFaQuWHZSNFr8sVJoW3LcF5g/z9Wkf5GwvGGtD/yb6bcsjZ4tj
lw==
```

ESX SYSTEM SSH CONFIGURATION

To configure an ESX Server to accept remote commands by using SSH, follow these steps.

1	Log in to the 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	Change the following line from
	`PermitRootLogin no`
	to
	`PermitRootLogin yes`

6	Restart the SSH service by running the following command: `service sshd restart`
7	Create the SSH public key: `ssh-keygen -t dsa –b 1024` This command outputs content similar to the following example. Retain the default locations, and do not use a passphrase.
8	Change to the `.ssh` directory: `cd /root/.ssh`

9	Run the following commands:
	`cat id_dsa.pub >> authorized_keys`
	`chmod 600 authorized_keys`

10	Repeat steps 1 through 9 for each ESX Server in the cluster.

Example 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.
```

17 APPENDIX B: EXAMPLE SNAPSHOT SCRIPT

This script is provided as an example of how to create consistent backups of virtual machines in a VMware Virtual Infrastructure 3 environment leveraging NetApp Snapshot technology. It is provided as an example that can easily be modified to meet the needs of an environment. For samples of advanced scripts built from this example framework, check out VIBE, located in the NetApp Tool Chest.

Backing up VMs with this script completes the following process:

- Will quiesce all of the VMs on a given Datastore.
- Takes a crash-consistent NetApp Snapshot copy.
- Applies the Redo logs and restores the virtual disk files to a read-write state.

Example hot backup Snapshot script:

```
#!/bin/sh

# Example code which takes a snapshot of all VMs using the VMware

# vmware-cmd facility. It will maintain and cycle the last 3 Snapshot copies.

#

# This sample code is provided AS IS, with no support or warranties of any

# kind, including but not limited to warranties of merchantability or

# fitness of any kind, expressed or implied.

#

# --------------------------------------------------------------------------


PATH=$PATH:/bin:/usr/bin


# Step 1 Enumerate all VMs on an individual ESX Server, and put each VM in hot
backup mode.
for i in `vmware-cmd -l`
do
  vmware-cmd $i createsnapshot backup NetApp true false
done


# Step 2 Rotate NetApp Snapshot copies and delete oldest, create new,
maintaining 3.
ssh <Filer> snap delete <esx_data_vol> vmsnap.3
ssh <Filer> snap rename <esx_data_vol> vmsnap.2 vmsnap.3
ssh <Filer> snap rename <esx_data_vol> vmsnap.1 vmsnap.2
ssh <Filer> snap create <esx_data_vol> vmsnap.1


# Step 3 Bring all VMs out of hot backup mode,
for i in `vmware-cmd -l`
do
  vmware-cmd $i removesnapshots
done
```

18 APPENDIX C: RECOVERING DATA FROM A NETAPP SNAPSHOT BACKUP

18.1 ESX CONFIGURATION FOR NETAPP SNAPSHOT BACKUPS

In a VMware Virtual Infrastructure, the storage provisioned to virtual machines is stored in either virtual disk files (residing on VMFS or NFS) or in raw device mappings (RDMs). With the introduction of VI3, administrators can mount storage-created Snapshot copies of VMFS LUNs. With this feature, customers can now connect to Snapshot copies of both VMFS and RDM LUNs from a production ESX Server. To enable this functionality, follow these steps.

1	Connect to vCenter Server.
2	Select an ESX Server.
3	In the right pane, select the Configuration tab.
4	In the Software box, select Advanced Settings to open the Advanced Settings window.
5	In the left pane, select LVM.
6	In the right pane, enter the value of 1 in the LVM.EnableResignature box.
7	After the resignature is complete return the value of LVM.EnableResignature to 0

18.2 RECOVERING VIRTUAL MACHINES FROM A NETAPP SNAPSHOT (VMFS DATASTORE)

NetApp Snapshot copies of VMFS Datastores offer a quick method to recover a VM. In summary, this process powers off the VM, leverages FlexClone to copy the backup of the VMFS LUN, copies the VMDK from the Snapshot copy to the production VMFS, and powers on the VM. To complete this process, follow these steps.

1	Connect to vCenter Server.
2	Select an ESX host and power down the VM.
3	Log in to the ESX console as root.
4	Rename the VMDK files: `mv <current VMDK path> <renamed VMDK path>`
5	Connect to the FAS system console (via either SSH, Telnet, or Console connection).

6	Clone the original LUN from a recent Snapshot copy, bring it online, and map it. From the storage appliance console, run:
	`lun clone create <clone LUN path> –b <original LUN path> <Snapshot name>`
	`lun online <LUN path>`
	`lun map <LUN path> <igroup> <ID>`

7	Connect to vCenter Server.
8	Select an ESX host.
9	In the right pane, select the Configuration tab.
10	In the Hardware box, select the Storage Adapters link.
11	In the upper right corner, select the Rescan link. Scan for both new storage and VMFS Datastores. The Snapshot VMFS Datastore appears.
12	Log in to the ESX console as root.

13	Copy the virtual disks from the Snapshot Datastore to the production VMFS:
	`cd <VMDK snapshot path>`
	`cp <VMDK> <production VMDK path>`

14	Connect to vCenter Server.
15	Select the ESX Server and start the virtual machine.
16	Validate that the restore is to the correct version. Log in to the VM and verify that the system was restored to the correct point in time.
17	Connect to the FAS system console (via either SSH, Telnet, or Console connection).
18	Delete the Snapshot copy LUN: `lun destroy –f <LUN path>`
29	In the upper right corner, select the Rescan link. Scan for both new storage and VMFS Datastores.

18.3 RECOVERING VIRTUAL MACHINES FROM AN NETAPP SNAPSHOT (NFS DATASTORE)

NFS offers a quick method to recover a VM from a Snapshot copy. In summary, this process powers off the VM, restores the VMDK, and powers on the VM. The actual process of recovering a VMDK file can either be completed on the FAS array or within ESX 3.5. With ESX 3.5 the restore process can also be completed as a copy operation from within the Virtual infrastructure Client’s Data Browser. To complete this process from the FAS array, follow these steps.

1	Connect to vCenter Server.
2	Select an ESX host and power down the VM.
3	Log in to the ESX console as root.
4	Rename the VMDK files: `mv <current VMDK path> <renamed VMDK path>`
5	Connect to the FAS system console (via either SSH, Telnet, or Console connection).
6	Restore the VMDK file from a recent Snapshot copy: `snap restore –t file -s <snapshot-name> <original VMDK path> <original VMDK path>`
7	Connect to vCenter Server.
8	Select the ESX and start the virtual machine.
9	Validate that the restore is to the correct version. Log in to the VM and verify that the system was restored to the correct point in time.
10	Log in to the ESX console as root.
11	Delete the renamed VMDK files: `rm <renamed VMDK path>`

18.4 RECOVERING RDM BASED VIRTUAL MACHINES FROM A NETAPP SNAPSHOT

RDMs provide the quickest possible method to recover a VM from a Snapshot copy. In summary, this process powers off the VM, restores the RDM LUN, and powers on the VM. To complete this process, follow these steps.

1	Connect to vCenter Server.
2	Select an ESX host and power down the VM.
3	Connect to the FAS system console (via either SSH, Telnet, or Console connection).
4	Clone the original LUN from a recent Snapshot copy: `lun clone create <clone LUN path> –b <original LUN path> <Snapshot name>`
5	Take the current version of the LUN in use off line: `lun offline <LUN path>`

6	Map the cloned LUN and put it on line:
	`lun online <LUN path>`
	`lun map <LUN path> <igroup> <ID>`

7	Connect to vCenter Server.
8	Select an ESX host and power on the VM.
9	Validate that the restore is to the correct version. Log in to the VM and verify that the system was restored to the correct point in time.
10	Connect to the FAS system console (via either SSH, Telnet, or Console connection).

11	Delete the original LUN and split the clone into a whole LUN:
	`lun destroy –f <original LUN path>`
	`lun clone split start <cloned LUN path>`

12	Rename the cloned LUN to the name of the original LUN (optional): `lun mv <cloned LUN path> <original LUN path>`

19 APPENDIX D: RELOCATING THE PAGEFILE IN WINDOWS VMS

Following is a registry file example of a simple registry script that sets the pagefile and temp area (for both user and system) to the D:\ partition. This script should be executed the first time a new virtual machine is created. If the D:\ partition does not exist, the systems 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 `temp.reg`. The Setup Manager has a section where you can add `temp.reg` to the run the first time the virtual machine is powered on. For more information about automating the deployment of cloned Windows servers, see Microsoft Setup Manager.

REGISTRY FILE EXAMPLE

```
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,00,70,00,61,00,67,00,65,00,66,00,69,00,6
   c,\


   00,65,00,2e,00,73,00,79,00,73,00,20,00,32,00,30,00,34,00,38,00,20,00,32,0
   0,\

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

[HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Session
   Manager\Environment]

"TEMP"="D:\\"

"TMP"="D:\\"

[HKEY_CURRENT_USER\Environment]

"TEMP"="D:\\"

"TMP"="D:\\"

[HKEY_USERS\.DEFAULT\Environment]

"TEMP"="D:\\"

"TMP"="D:\\"

End ----------
```

20 DOCUMENT REFERENCES

VMWARE REFERENCES

VMware Introduction to Virtual Infrastructure

http://www.vmware.com/pdf/vi3_35/esx_3/r35u2/vi3_35_25_u2_intro_vi.pdf

VMware ESX Server 3 Configuration Guide

http://www.vmware.com/pdf/vi3_35/esx_3/r35u2/vi3_35_25_u2_3_server_config.pdf

VMware Basic System Administration Guide

http://www.vmware.com/pdf/vi3_35/esx_3/r35u2/vi3_35_25_u2_admin_guide.pdf

VMware Fibre Channel SAN configuration Guide

http://www.vmware.com/pdf/vi3_35/esx_3/r35u2/vi3_35_25_u2_san_cfg.pdf

VMware SAN Design and Deployment Guide

http://www.vmware.com/pdf/vi3_san_design_deploy.pdf

Configuration Maximums for VMware Infrastructure 3

http://www.vmware.com/pdf/vi3_35/esx_3/r35u2/vi3_35_25_u2_config_max.pdf

VMware VMworld Conference Sessions Overview

http://www.vmworld.com/vmworld/home.jspa

VMware Recommendations for Aligning VMFS Partitions

http://www.vmware.com/pdf/esx3_partition_align.pdf

VMware iSCSI SAN Configuration Guide

http://www.vmware.com/pdf/vi3_35/esx_3/r35u2/vi3_35_25_u2_iscsi_san_cfg.pdf

VMware Fibre Channel SAN Configuration Guide

http://www.vmware.com/pdf/vi3_35/esx_3/r35u2/vi3_35_25_u2_san_cfg.pdf

VMware Update Manager Administration Guide

http://www.vmware.com/pdf/vi3_vum_10u2_admin_guide.pdf

VMware Patches can be found at:

http://support.vmware.com/selfsupport/download/

VMware KB: New LUNs Not Detected on QLogic HBAs

http://kb.vmware.com/kb/1003988

VMware VMFS Technical Overview and Best Practices

http://www.vmware.com/pdf/vmfs-best-practices-wp.pdf

VMware KB: Newly extented VMFS datastore fluctuates between the old size and new size

http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1002558

VMware KB: Free space from an extent added to a VMFS3 datastore is not seen on the other ESX Server

http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1002784

Best Practices for Patching VMware® ESX/ESXi
http://www.vmware.com/files/pdf/esx_patching_best_practices.pdf

VMware KB: NFS mounts are restricted to eight by default in ESX 3.5.x

http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2239

NETAPP REFERENCES

NetApp VMInsight with SANscreen

http://www.netapp.com/us/products/management-software/sanscreen-vm-insight.html

NetApp TR-3612: NetApp and VMware Virtual Desktop Infrastructure

http://www.netapp.com/library/tr/3612.pdf

NetApp TR-3515: NetApp and VMware ESX Server 3.0: Building a Virtual Infrastructure from Server to Storage

http://www.netapp.com/library/tr/3515.pdf

NetApp TR-3482: NetApp and VMware ESX Server 2.5.x

http://www.netapp.com/library/tr/3482.pdf

NetApp TR-3348: Block Management with Data ONTAP 7G: FlexVol, FlexClone, and Space Guarantees

http://www.netapp.com/library/tr/3348.pdf

NetApp TR-3737: SnapManager for Virtual Infrastructure best practices

http://www.netapp.com/us/library/technical-reports/tr-3737.html

RAID-DP: NetApp Implementation of RAID Double Parity

http://media.netapp.com/documents/wp_3298.pdf

NetApp TR-3671: VMware Site Recovery Manager in a NetApp Environment

http://media.netapp.com/documents/tr-3671.pdf

NetApp MBRTools

http://now.netapp.com/NOW/download/tools/mbralign/

ONTAP File Access and Protocol Management Guide

http://now.netapp.com/NOW/knowledge/docs/ontap/rel731/html/ontap/filesag/accessing/task/t_oc_accs_file_sharing_between_NFS_and_CIFS.html

NetApp Rapid Cloning Utility

http://now.netapp.com/NOW/download/tools/rcu/

DataFabric® Manager Server 3.7: Operations Manager Administration Guide

http://now.netapp.com/NOW/knowledge/docs/DFM_win/rel371/html/software/opsmgr/index.htm

NetApp KB: VMFS volume resignaturing in a NetApp environment

https://now.netapp.com/Knowledgebase/solutionarea.asp?id=kb33990

NetApp: ONTAP File Access and Protocol Management Guide

http://now.netapp.com/NOW/knowledge/docs/ontap/rel731/html/ontap/filesag/accessing/task/t_oc_accs_file_sharing_between_NFS_and_CIFS.html

TR-3505: NetApp FAS Dedupe: Data Deduplication Deployment and Implementation Guide

http://www.netapp.com/library/tr/3505.pdf

MISCELLANEOUS REFERENCES

Total Cost Comparison: IT Decision-Maker Perspectives on EMC and NetApp Storage Solutions in Enterprise Database Environments

http://www.netapp.com/library/ar/ar1038.pdf

Wikipedia RAID Definitions and Explanations

http://en.wikipedia.org/wiki/Redundant_array_of_independent_disks

Bart's Preinstallation Evniroment Windows live CD

http://www.nu2.nu/pebuilder/

Microsoft Diskpart utility

http://support.microsoft.com/default.aspx?scid=kb;en-us;300415.

Ext2resize

http://sourceforge.net/projects/ext2resize.

IBM: Storage Block Alignment with VMware Virtual Infrastructure

ftp://service.boulder.ibm.com/storage/isv/NS3593-0.pdf

EMC: Celerra IP Storage with VMware Virtual Infrastructure

http://www.vmware.com/files/pdf/VMware_VI3_and_EMC_Celerra_IP.pdf

Dell: Designing and Optimizing SAN Configurations

http://www.dell.com/downloads/global/power/ps4q04-20040149-Mehis.pdf

EMC: CLARiiON Integration with VMware ESX Server

http://www.vmware.com/pdf/clariion_wp_eng.pdf

Vizioncore: vOptimizer Pro FAQ

http://www.vizioncore.com/products/vOptimizerPro/documents/vOptimizerProFAQ.pdf

21 VERSION TRACKING

Version 1.0	May 2006
	Original document
Version 2.0	January 2007 – Major revision update
	Added support for VI3
	Added co-author Mike Slisinger
	Added NFS, FC, & iSCSI connectivity options
	Added datastore and VMDK growth content
	Added storage management content
Version 2.1	May 2007 – Miscellaneous edits
	Updated VM Snapshot script and instructions
Version 3.0	September 2007 – Major revision update
	Enhanced FC connectivity section
	Added section on TOE adapters and HW iSCSI targets within the storage array
	Added NetApp ESX Host Utilities content
	Added A-SIS content
	Added Thin Provisioning content
	Added optimized storage layout content including registry settings for Windows server
Version 3.1	October 2007 – Miscellaneous edits
	Added NFS snapshot configuration requirement
Version 4.0	June 2008 – Major revision update
	Added support for VI3.5
	Added Ethernet networking section
	Added co-author Larry Touchette
	Enhanced NFS connectivity section and removed vswap location restrictions
	Enhanced NetApp ESX Host Utilities content
	Updated A-SIS references to data deduplication
	Update Network Appliance references to NetApp
Version 4.1	July 2008– Miscellaneous edits
	Reorganized Ethernet networking section
Version 4.1.1	August 2008 – Miscellaneous edits
Version 4.2	September 2008 – Miscellaneous edits
	Added information on VMware SR195302591
	Added VMware patch ESX350-200808402-BG
Version 4.3	November 2008 – Miscellaneous edits
	Added VMware KB 2239 update
	Added VMware patch ESX350-200808401-BG
Version 4.4	December 2008 – Miscellaneous edits
	Added support for ESX 3.5 update 3
	Added patch processes for ESX350-200808401-BG
	Added more content to IP storage network segmentation
	Removed references to TOE adapters and HW iSCSI targets within the storage array
	Replaced the backup section with SnapManager for Virtual Infrastructure
	Backup scripts relocated to appendix
	Updated diagrams to reflect ESX clusters
Version 4.5	March 2009 – Miscellaneous edits and document re-organization
	Added co-author Peter Learmonth
	Updated protocol comparison chart content
	Clarified VMkernel and gateway settings with Ethernet storage networking
	Removed CLI editing of VMX file in regards to relocating vswap
	Updated NetApp cfmode settings
	Enhanced FC & IP storage network best practices
	Removed independent disk option for swap/pagefile VMDK
	Added a section on the NetApp ESX Host Utility
	Added a section on the NetApp Rapid Cloning Utility
	Enhanced content on GOS file system alignment
Version 4.5.1	Added GOS timeout settings

	Added IP connectivity clarifications
Version 4.5.2	VMware co-loged the document
	Added NFS snapshot commit patch details to include The ESX/ESXi 3.0.3 release
	Minor edits and corrections