



NETAPP TECHNICAL REPORT

# NetApp and VMware vSphere Storage Best Practices

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

June 2009 | TR-3749 | Version 1.0

## TABLE OF CONTENTS

<b>1</b>	<b>EXECUTIVE SUMMARY .....</b>	<b>5</b>
<b>2</b>	<b>VMWARE STORAGE OPTIONS.....</b>	<b>5</b>
2.1	STORAGE OVERVIEW: VMFS DATASTORES.....	6
2.2	STORAGE OVERVIEW: NAS DATASTORES .....	7
2.3	STORAGE OVERVIEW: RAW DEVICE MAPPINGS.....	8
2.4	DATASTORE COMPARISON TABLES.....	9
<b>3</b>	<b>NETAPP FAS CONFIGURATION AND SETUP .....</b>	<b>11</b>
3.1	DATA PROTECTION.....	11
3.2	NETAPP ARRAY CONFIGURATION .....	12
3.3	NETAPP STORAGE CONFIGURATION .....	12
<b>4</b>	<b>ESX FC, FCOE, AND ISCSI STORAGE CONFIGURATION.....</b>	<b>14</b>
4.1	LUN SIZING FOR VMFS DATASTORES .....	14
4.2	CLUSTER SIZING CONSIDERATIONS WHEN USING LUNS .....	14
4.3	FC, FCOE, AND ISCSI LUN PROVISIONING .....	14
4.4	CONNECTING FC AND FCOE DATASTORES.....	17
4.5	CONNECTING ISCSI DATASTORES.....	19
<b>5</b>	<b>VMWARE NATIVE MULTIPATHING .....</b>	<b>28</b>
5.1	DEFAULT NMP SETTINGS .....	28
5.2	ENABLING ALUA .....	29
5.3	DEFAULT NMP SETTINGS WITH ALUA ENABLED.....	30
5.4	CONFIGURING THE ROUND ROBIN PSP .....	30
<b>6</b>	<b>NFS STORAGE RECOMMENDATIONS .....</b>	<b>35</b>
6.1	INCREASING THE NUMBER OF NFS DATASTORES.....	35
6.2	FILE SYSTEM SECURITY .....	36
6.3	ESX NFS TIMEOUT SETTINGS.....	37
6.4	NFS STORAGE NETWORK BEST PRACTICE.....	38
6.5	CONNECTING NFS DATASTORES .....	39
<b>7</b>	<b>THE NETAPP ESX HOST UTILITIES .....</b>	<b>43</b>
7.1	INSTALLING THE EHU IN ESX.....	43
7.2	MANUAL CONFIGURATION OF FC HBAS IN ESX .....	44
<b>8</b>	<b>FC AND FCOE STORAGE NETWORKING BEST PRACTICES .....</b>	<b>45</b>
8.1	HOST BUS AND CONVERGED NETWORK ADAPTERS .....	45
8.2	NETAPP IGROUPS (LUN MASKING) .....	45
8.3	FC AND FCOE ZONING.....	45

<b>9</b>	<b>ETHERNET STORAGE NETWORKING BEST PRACTICES .....</b>	<b>46</b>
9.1	10 GIGABIT ETHERNET .....	46
9.2	VIRTUAL LANS (VLANS) .....	46
9.3	FLOW CONTROL .....	46
9.4	SPANNING TREE PROTOCOL .....	47
9.5	BRIDGE PROTOCOL DATA UNITS .....	47
9.6	NETAPP VIRTUAL INTERFACES .....	47
9.7	ETHERNET SWITCH CONNECTIVITY .....	48
<b>10</b>	<b>CONFIGURING ETHERNET STORAGE NETWORKS .....</b>	<b>49</b>
10.1	HIGHLY AVAILABLE STORAGE DESIGNS WITH TRADITIONAL ETHERNET SWITCHES .....	49
10.2	VMKERNEL CONFIGURATION WITH TRADITIONAL ETHERNET .....	52
10.3	A STORAGE ARCHITECTURE WITH TRADITIONAL ETHERNET .....	54
10.4	DATASTORE CONFIGURATION WITH TRADITIONAL ETHERNET .....	56
10.5	VMKERNEL CONFIGURATION WITH MULTI-SWITCH TRUNKING .....	57
10.6	A STORAGE ARCHITECTURE WITH MULTISWITCH TRUNKING .....	59
10.7	DATASTORE CONFIGURATION WITH MULTISWITCH TRUNKING .....	60
<b>11</b>	<b>INCREASING STORAGE UTILIZATION .....</b>	<b>61</b>
11.1	DATA DEDUPLICATION .....	62
11.2	ZERO-COST VIRTUAL MACHINE CLONING .....	64
11.3	STORAGE THIN PROVISIONING .....	64
<b>12</b>	<b>VIRTUAL MACHINE BEST PRACTICES .....</b>	<b>68</b>
12.1	WINDOWS VM FILE SYSTEM PERFORMANCE OPTION .....	68
12.2	ENSURING OPTIMUM VM AVAILABILITY .....	68
12.3	ENSURING OPTIMAL STORAGE PERFORMANCE .....	69
12.4	THE IMPACT OF PARTITION MISALIGNMENT .....	69
12.5	IDENTIFYING PARTITION ALIGNMENT .....	70
12.6	CORRECTIVE ACTIONS FOR VMS WITH MISALIGNED PARTITIONS .....	71
12.7	CREATE PROPERLY ALIGNED PARTITIONS FOR NEW VMS .....	72
<b>13</b>	<b>VIRTUAL MACHINE STORAGE LAYOUT .....</b>	<b>74</b>
13.1	DEFAULT VIRTUAL MACHINE LAYOUT .....	74
13.2	VIRTUAL MACHINE LAYOUT WITH NETAPP SNAP* TECHNOLOGIES .....	74
<b>14</b>	<b>STORAGE MONITORING AND MANAGEMENT .....</b>	<b>80</b>
14.1	MONITORING STORAGE UTILIZATION WITH NETAPP OPERATIONS MANAGER .....	80
14.2	STORAGE GROWTH MANAGEMENT .....	80
<b>15</b>	<b>DISK-BASED SNAPSHOT BACKUPS FOR VMWARE .....</b>	<b>87</b>
15.1	COMPLEMENTARY SNAPSHOT TECHNOLOGIES .....	87
15.2	IMPLEMENTING NETAPP SNAPSHOT BACKUPS FOR VMWARE VSPHERE .....	88

<b>16 TECHNICAL REPORT SUMMARY</b> .....	<b>89</b>
<b>APPENDIX A: CONFIGURING SSH ON ESX SERVERS AND NETAPP ARRAYS</b> .....	<b>90</b>
<b>A.1 CONFIGURING SSH ON ESX SERVERS AND NETAPP ARRAYS</b> .....	<b>90</b>
<b>APPENDIX B: RELOCATING THE PAGEFILE IN WINDOWS VMS</b> .....	<b>92</b>
<b>APPENDIX C: DOCUMENT REFERENCES</b> .....	<b>93</b>
<b>APPENDIX D: VERSION TRACKING</b> .....	<b>95</b>

## 1 EXECUTIVE SUMMARY

NetApp® technology enables companies to extend their virtual infrastructures to include the benefits of advanced storage virtualization. NetApp provides unified storage solutions that provide industry-leading technologies in the areas of storage efficiencies, instantaneous VM and datastore cloning for virtual servers and virtual desktops, and virtual data center backup and business continuance solutions.

This technical report reviews the best practices for implementing VMware® vSphere with NetApp unified storage arrays. NetApp has been providing advanced storage features to VMware ESX solutions since 2001. During this 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 data center.

The target audience for this paper is familiar with concepts pertaining to VMware ESX/ESXi Server 4.0, vCenter Server 4.0, and NetApp Data ONTAP® 7.X.

## 2 VMWARE STORAGE OPTIONS

VMware ESX supports three types of storage configurations when connecting to shared storage arrays: VMFS datastores, NAS datastores, and raw device mappings. It is assumed that customers understand that shared storage is required to enable high-value VMware features such as HA, DRS, VMotion®, and Fault Tolerance. The goal of the following sections is to provide customers information to consider when designing their virtual data center.

VMware virtualization technology makes it easy for customers to leverage all of these storage designs at any time or simultaneously. The following section reviews these storage options and summarizes the unique characteristics of each architecture. For information regarding deploying with VMFS, NFS, and RDMs, see the VMware ESX and ESXi Server Configuration Guide.

## 2.1 STORAGE OVERVIEW: VMFS DATASTORES

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

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

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

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

NetApp enhances the use of VMFS datastores through many technologies, including array-based 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, the ESX Host Utilities, and SANscreen® VMI Insight. This last tool address the challenges stated in the previous paragraph.

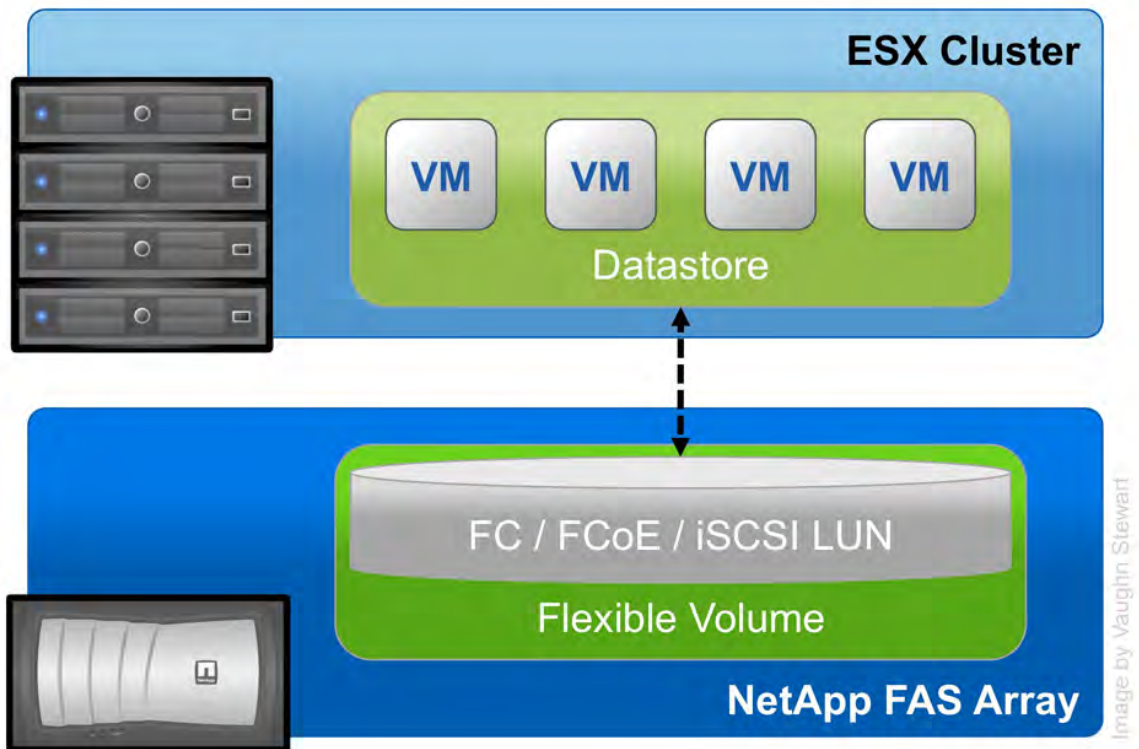


Figure 1) ESX Cluster connected to a VMFS datastore using FC or iSCSI.

## 2.2 STORAGE OVERVIEW: NAS DATASTORES

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

Deploying VMware with NetApp NFS datastores is the easiest means to integrate VMware virtualization technologies directly with WAFL®, NetApp's advanced data management and storage virtualization engine.

Examples of this transparent integration include production-use data deduplication, immediate zero-cost VM and datastore clones, array-based thin provisioning, and direct access to array-based Snapshot™ copies. NetApp provides additional VMware integrated tools for NFS such as SnapManager for Virtual Infrastructure, the Rapid Cloning Utility, and the ESX Host Utilities.

Figure 2 displays 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. Combining NetApp's advanced NFS servers with VMware's high-performance NFS implementation can provide I/O to shared datastores that is on par with that of other storage protocols such as Fibre Channel.

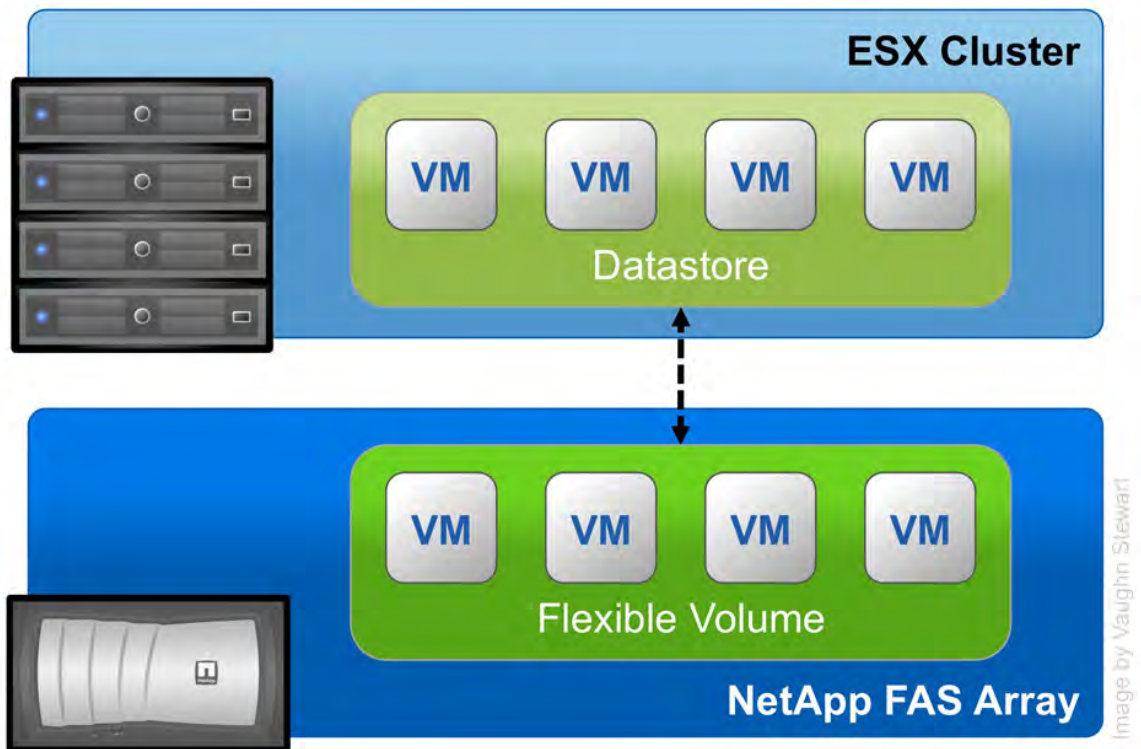


Figure 2) ESX Cluster connected to an NFS datastore.

### 2.3 STORAGE OVERVIEW: RAW DEVICE MAPPINGS

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

Unlike VMFS and NFS, RDMs are not used to provide shared datastores.

RDMs are an enabling technology for solutions such as virtual machine and physical-to-virtual-machine host-based clustering, such as with Microsoft® Cluster Server (MSCS). RDMs provide traditional LUN access to a host. So they can achieve high individual disk I/O performance, and they can be easily monitored for disk performance by a storage array.

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

The challenges of this solution are that VMware clusters may have to be limited in size, and this design requires ongoing interaction between storage and VMware administration teams. Figure 3 shows an example of this configuration.

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

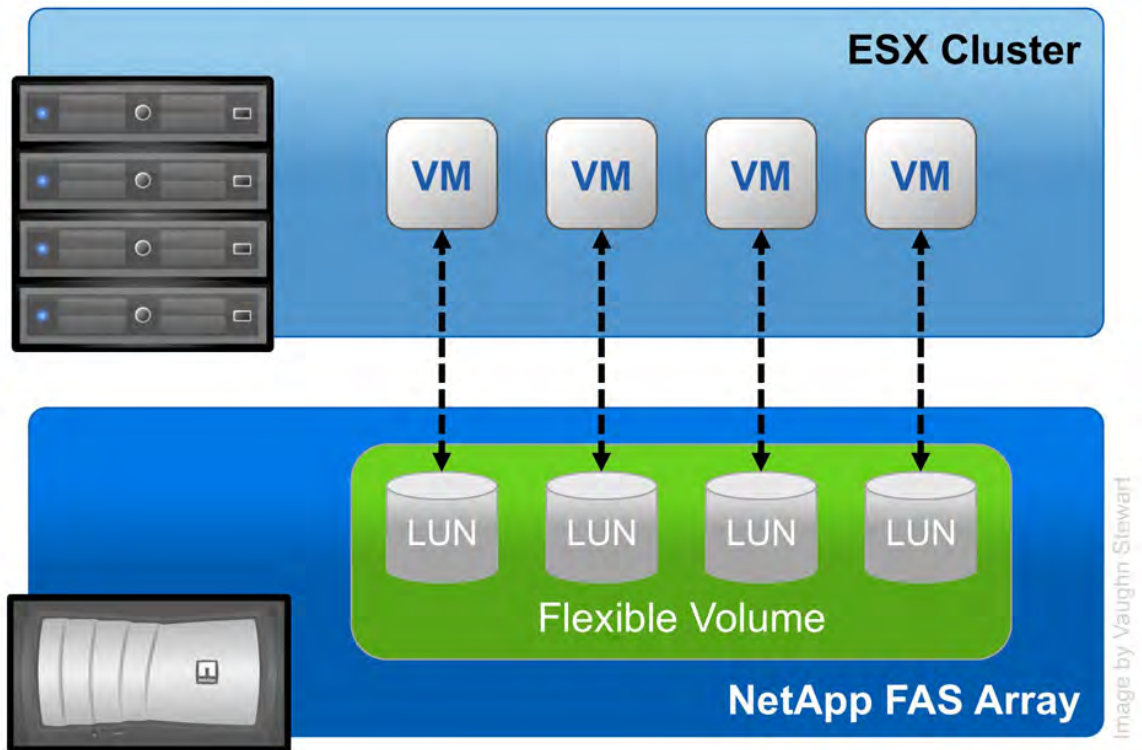


Figure 3) ESX Cluster with VMs connected to RDM LUNs using FC or iSCSI.



## 2.4 DATASTORE COMPARISON TABLES

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 supported features.

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

Table 2) VMware supported functionality.

Capability/Feature	FC/FCoE	iSCSI	NFS
VMotion	Yes	Yes	Yes
Storage VMotion	Yes	Yes	Yes
VMware HA	Yes	Yes	Yes
DRS	Yes	Yes	Yes
VCB	Yes	Yes	Yes
MSCS within a VM	Yes, using RDM	Not supported	Not supported
Fault Tolerance	Yes, VMFS only	Yes, VMFS only	Yes
Site Recovery Manager	Yes*	Yes*	Yes*
Thin Provisioned VMDK	Yes	Yes	Yes
VMware NMP	Yes	Yes	N/A

\* Available when support for vSphere is announced for SRM.

Table 3) NetApp supported storage management features.

Capability/Feature	FC/FCoE	iSCSI	NFS
Data deduplication	Savings in the array	Savings in the array	Savings in the datastore
Thin provisioning	Datastore or RDM	Datastore or RDM	Datastore
Resize datastore	Grow only	Grow only	Grow, Auto-grow, and Shrink
Data deduplication	Savings in the array	Savings in the array	Savings in the datastore
Thin provisioning	Datastore or RDM	Datastore or RDM	Datastore
Resize datastore	Grow only	Grow only	Grow, Auto-grow, and Shrink
NetApp Rapid Cloning Utility	Yes*	Yes*	Yes
SANscreen VMInsight	Yes	Yes	Yes
SnapDrive	Yes	Yes using GOS initiator	No
SnapManager for VI	Yes	Yes	Yes
ESX Host Utilities	Yes	Yes	Yes

\* Available with the release of RCU 2.2.

Table 4) Supported backup features.

Capability/Feature	FC/FCoE	iSCSI	NFS
Snapshot backups	Yes	Yes	Yes
SMVI replicated backup storage supports SRM	Yes*	Yes*	Yes*
SnapMirror®	Datastore or RDM	Datastore or RDM	Datastore or VM
SnapVault®	Datastore or RDM	Datastore or RDM	Datastore or VM
VMDK image access	VCB	VCB	VCB, VIC File Explorer
VMDK file level access	VCB, Windows® only	VCB, Windows only	VCB and 3 <sup>rd</sup> party apps
NDMP granularity	Datastore	Datastore	Datastore or VM

\* Available when support for vSphere is announced for SRM.

## 3 NETAPP FAS CONFIGURATION AND SETUP

### 3.1 DATA PROTECTION

#### RAID AND 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 a virtual data center. These technologies include physical server clustering using VMware HA, application load balancing with DRS, and the ability to nondisruptively move running VMs and data sets between physical 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<sup>®</sup> 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.

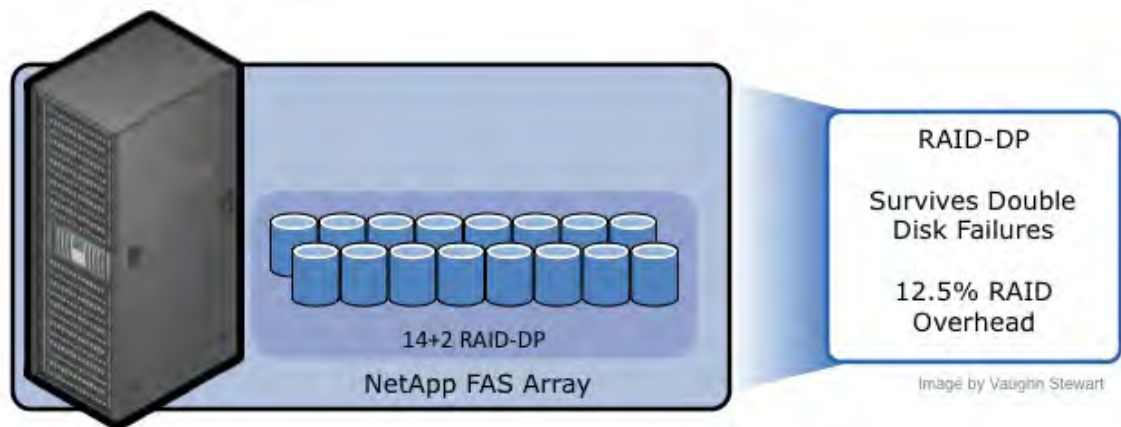


Figure 4) NetApp RAID-DP.

#### 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 I/O 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 CONFIGURATIONS

NetApp HA arrays ship configured 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. If you are deploying ESX on an older HA array with FC or FCoE, then make sure the `cfmode` is set to SSI. To verify the current `cfmode`, follow these steps.

1	Connect to the FAS system console (using either SSH, Telnet, or Console connection).
2	Enter <code>fc show cfmode</code> .

To set the `cfmode`, follow these steps.

1	Connect to the FAS system console (using either SSH, Telnet, or Console connection).
2	If <code>cfmode</code> needs to be changed, enter <code>FC set cfmode single_image</code>

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: <code>snap sched &lt;vol-name&gt; 0 0 0</code>
3	Set the volume Snapshot reserve: <code>c</code> <code>snap reserve &lt;vol-name&gt; 0</code>

## 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, FCOE, AND ISCSI STORAGE CONFIGURATION

### 4.1 LUN SIZING FOR VMFS DATASTORES

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

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

Advanced storage technologies such as thin provisioning, which is available with VMware VMDKs and NetApp datastores, 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.

### 4.2 CLUSTER SIZING CONSIDERATIONS WHEN USING LUNS

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

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

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

#### RDM EXAMPLE

The formula for 2 RDMs per VM with 20 VMs per ESX Server would be:

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

### 4.3 FC, FCOE, AND ISCSI LUN PROVISIONING

When provisioning LUNs for access using 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 multiple block-based protocols, separate igroups must be created for each.

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.

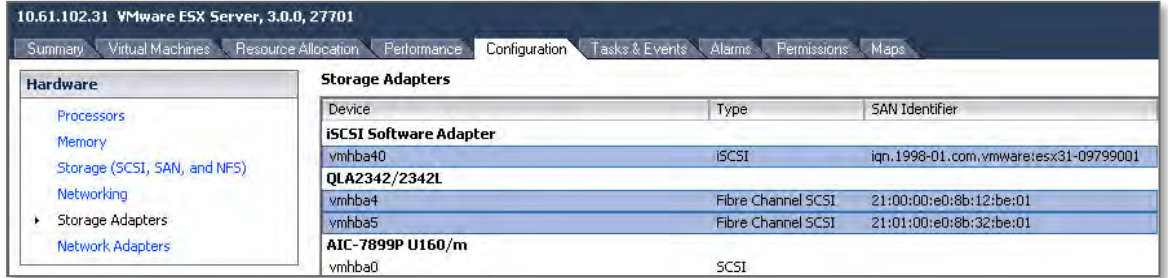


Figure 5) Identifying WWPN and IQN numbers in the Virtual Infrastructure Client.

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.



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.



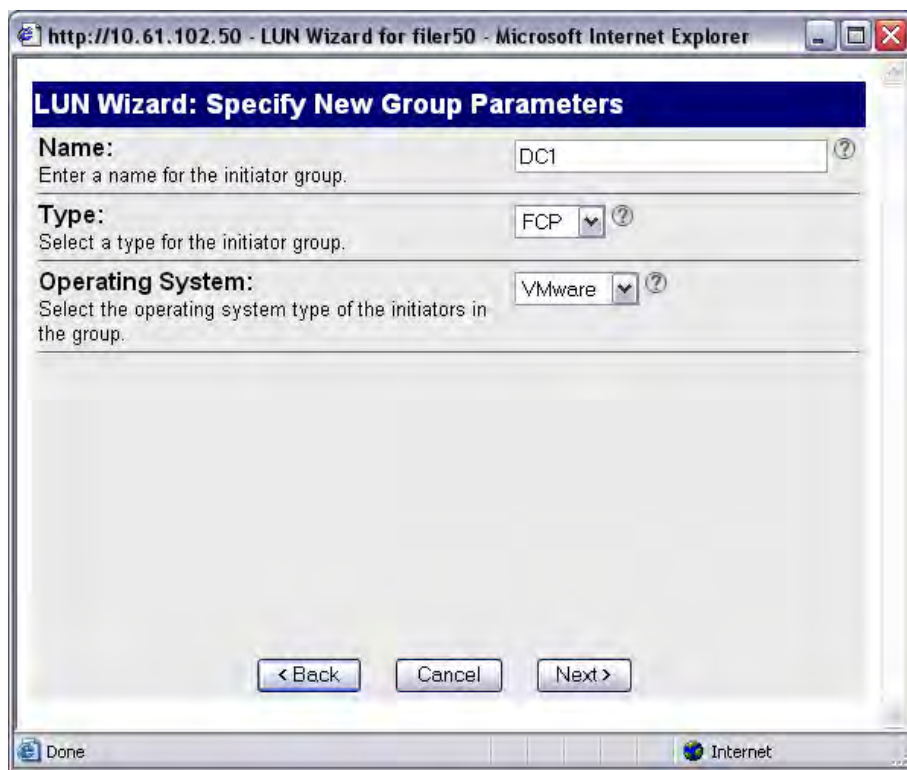


Figure 7) Assigning an igroup to a LUN.

#### 4.4 CONNECTING FC AND FCOE 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, FCoE, and iSCSI) to discover changes in the storage available to the ESX Server.

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.

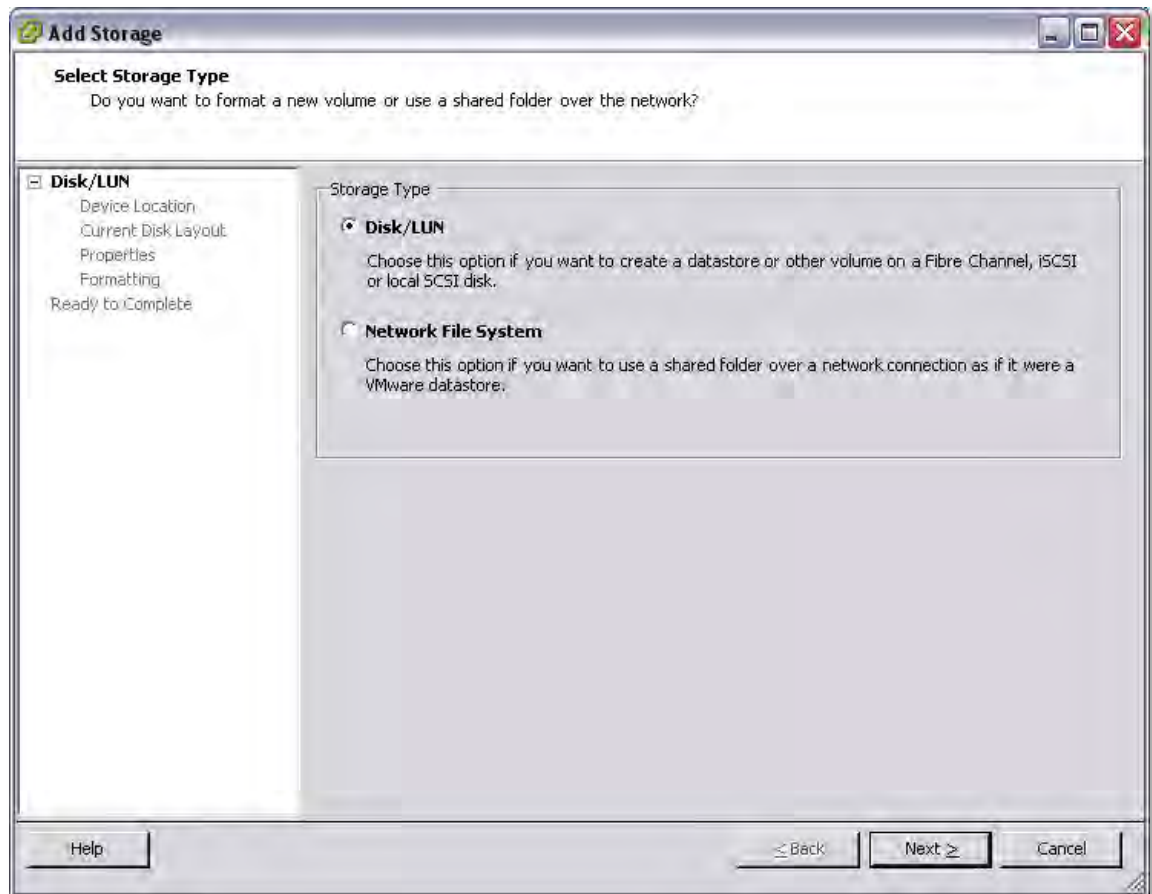


Figure 8) VMware Add Storage wizard.

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).

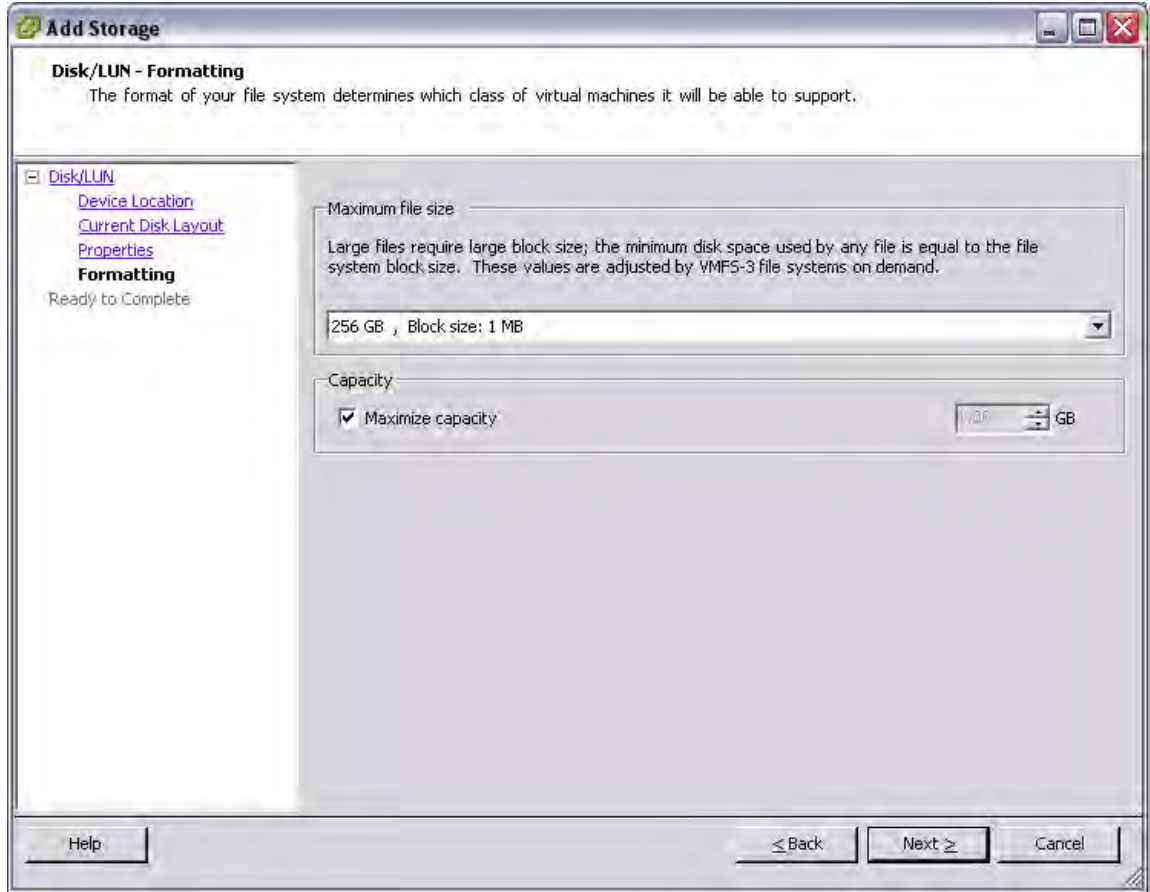


Figure 9) Formatting a LUN with VMFS.

## 4.5 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. New with ESX and ESX 4.0 is the ability to enable multiple TCP sessions with iSCSI datastores. Enabling multiple TCP sessions with the ESX Round Robin PSP will allow iSCSI datastores to send I/O over every available path to an iSCSI target.

For more information on configuring NMP for iSCSI, see 5.4 Configuring the Round Robin PSP

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 recommended to not allow routing of data between these networks. In other words, 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>.

#### ENABLE ISCSI COMMUNICATIONS

1	Open vCenter Server.
2	Select an ESX host.
3	In the right pane, select the Configuration tab.
4	In the Configuration tab, left pane, select Security Profile.
5	In the right pane, select the Properties link to open the Firewall Properties window.
6	Select the Software iSCSI Client checkbox and then click OK to close the Firewall Properties window (see Figure 10).

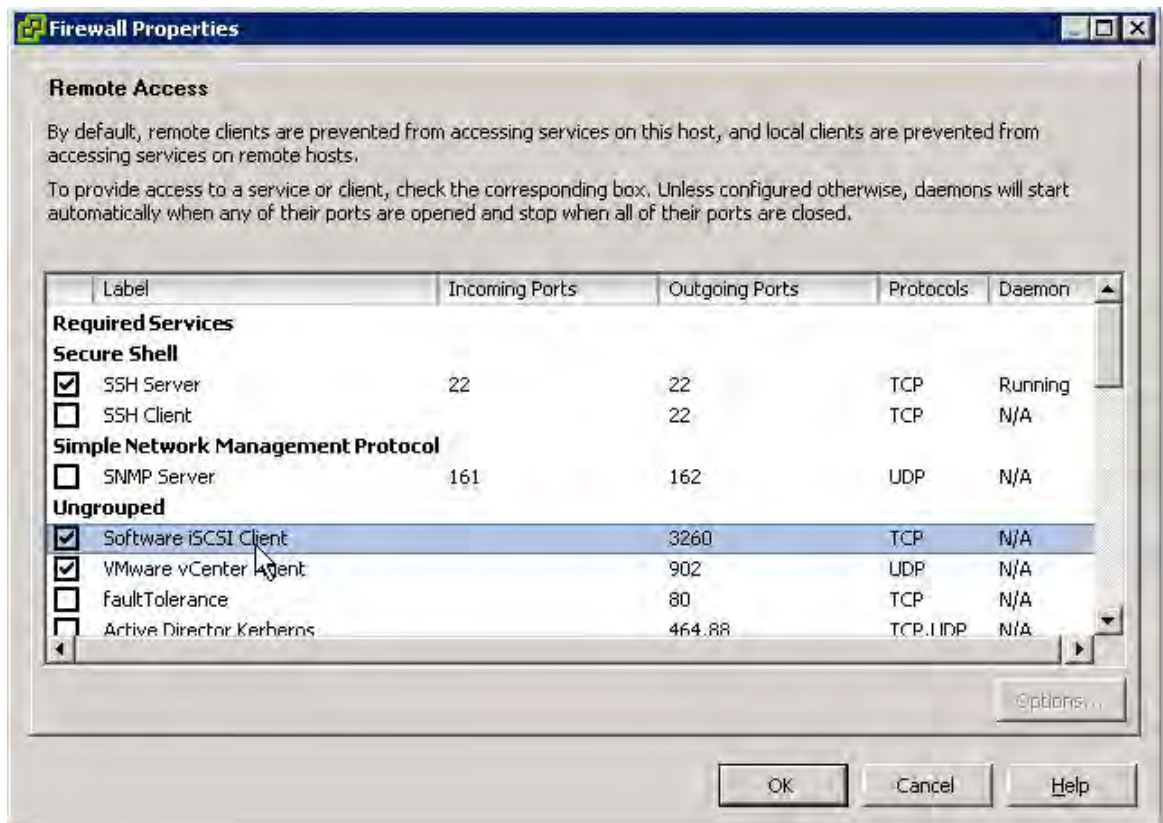


Figure 10) Configuring the firewall in ESX.

## CREATE MULTIPLE ISCSI VMKERNELS

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
6	Select the VMkernel radio button and click Next.
7	Note: You will create a VMkernel for every Ethernet link that you want to dedicate to iSCSI traffic. Note that VMkernels can be on different IP subnets. This configuration is required if combining iSCSI with NFS datastore access.
8	Configure the VMkernel by providing the required network information. A default gateway is not required for the VMkernel IP storage network.
9	Each VMkernel must be configured to use with a single active adapter (such as VMNIC0) that is not used by any other iSCSI VMkernel. Also, each VMkernel must not have any standby adapters (see Figures 11 and 12).
10	<p>The software iSCSI daemon is required to be bound to each VMkernel. This step can only be completed using the CLI.</p> <p>Connect to an ESX or ESXi console and run the following:</p> <pre>esxcli swiscsi nic add -n &lt;VMkernel ID&gt; -d &lt;Virtual HBA ID&gt;</pre> <p>As an example:</p> <pre>esxcli swiscsi nic add -n vmk0 -d vmhba33</pre> <pre>esxcli swiscsi nic add -n vmk1 -d vmhba33</pre>
11	<p>Verify the iSCSI to VMkernel bindings. Connect to an ESX or ESXi console and run the following:</p> <pre>esxcli swiscsi nic list -d &lt;Virtual HBA ID&gt;</pre> <p>As an example:</p> <pre>esxcli swiscsi nic list -d vmhba33</pre> <p>See Figure 13.</p>

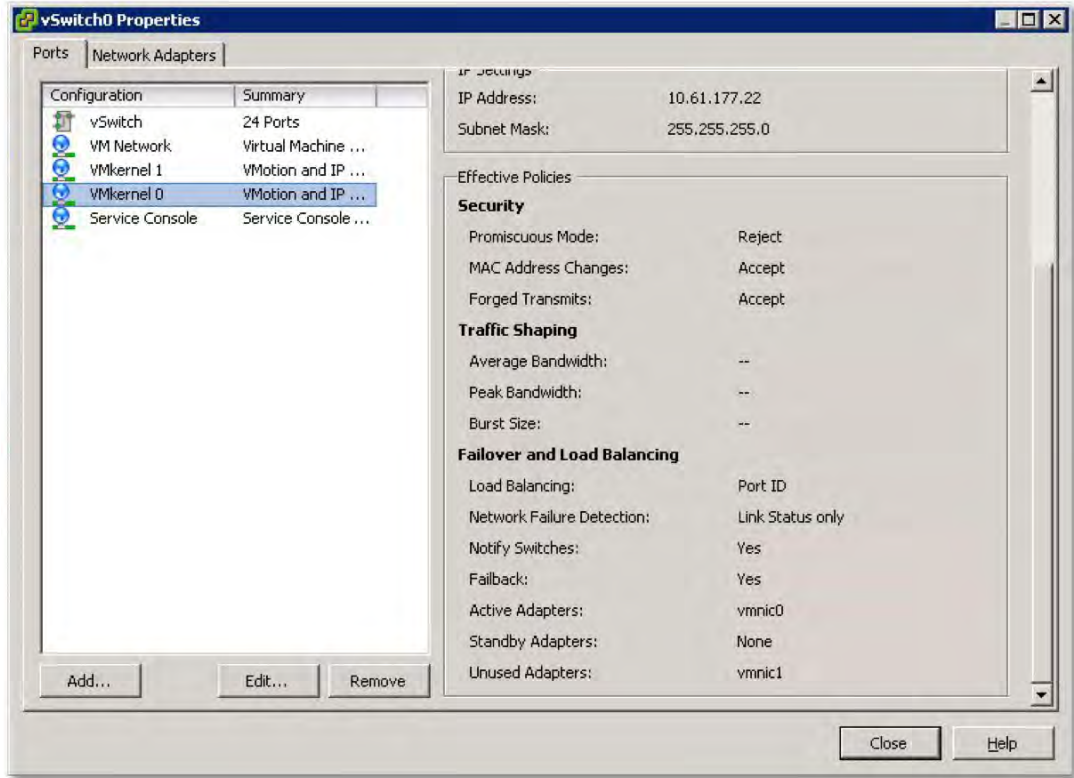


Figure 11) iSCSI VMkernel 0: Note active adapter vmnic0.

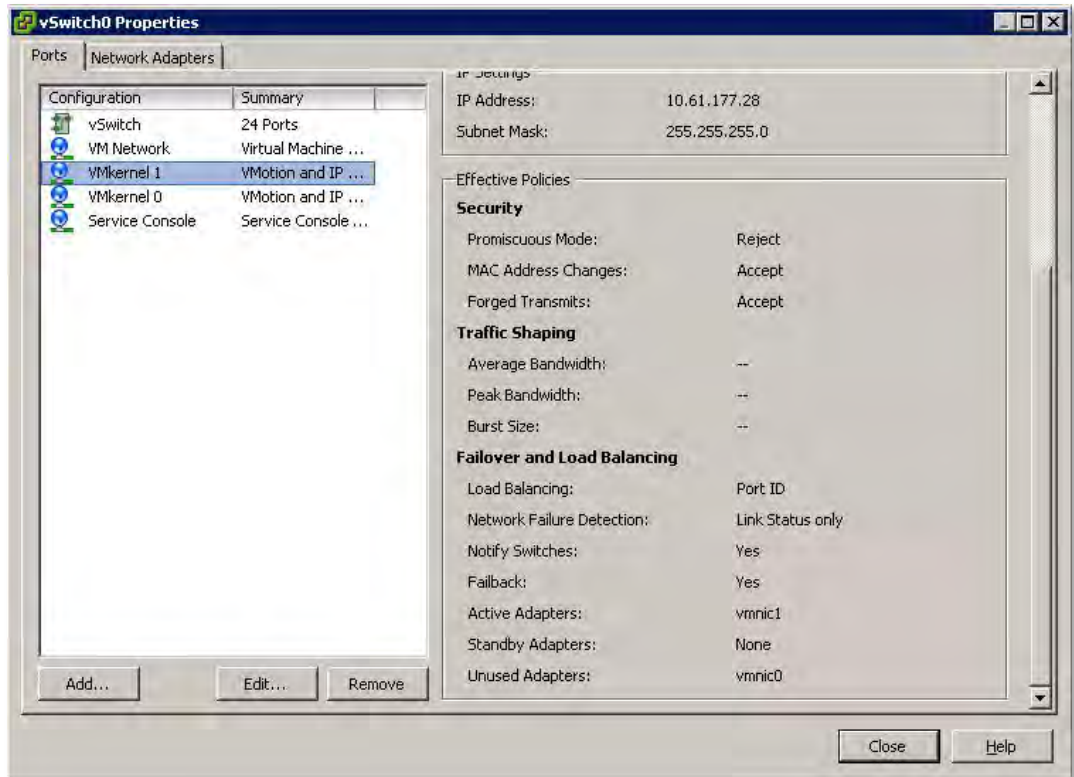
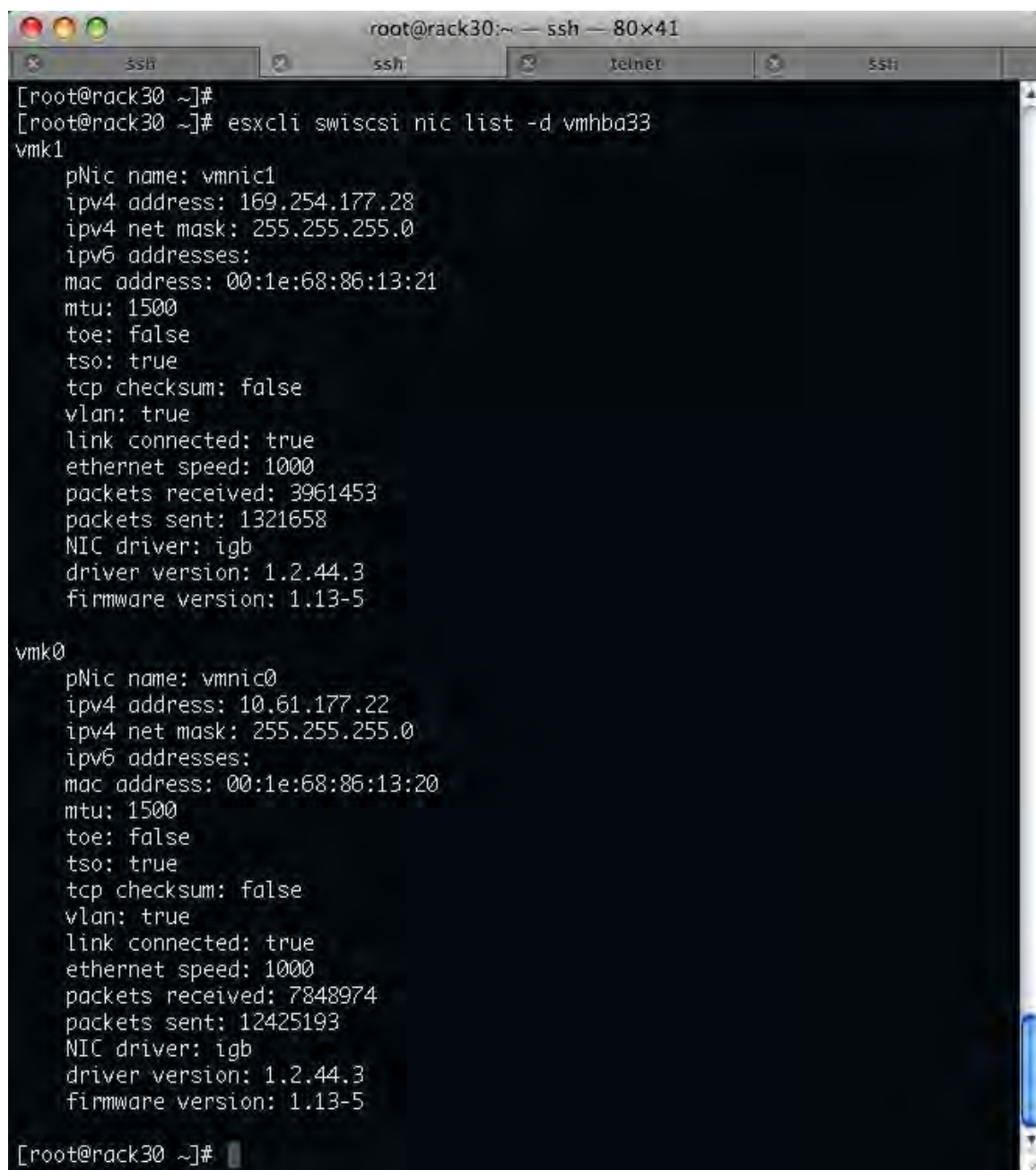


Figure 12) iSCSI VMkernel 1: Note active adapter vmnic1.



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

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

[root@rack30 ~]#
```

Figure 13) Verifying iSCSI to VMkernel bindings.

## CONNECTING TO ISCSI TARGETS

1	Open vCenter Server.
2	Select an ESX host.
3	In the right pane, select the Configuration tab.
4	In the right pane, Hardware box, select Storage Adapters.
5	Highlight the iSCSI adapter and click the Properties link in the Details box (see Figure 14).
6	Select the Dynamic Discovery tab in the iSCSI Initiator Properties box.
7	Click Add and enter the IP address of the iSCSI-enabled interface on the NetApp FAS system (see Figure 15).
8	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.

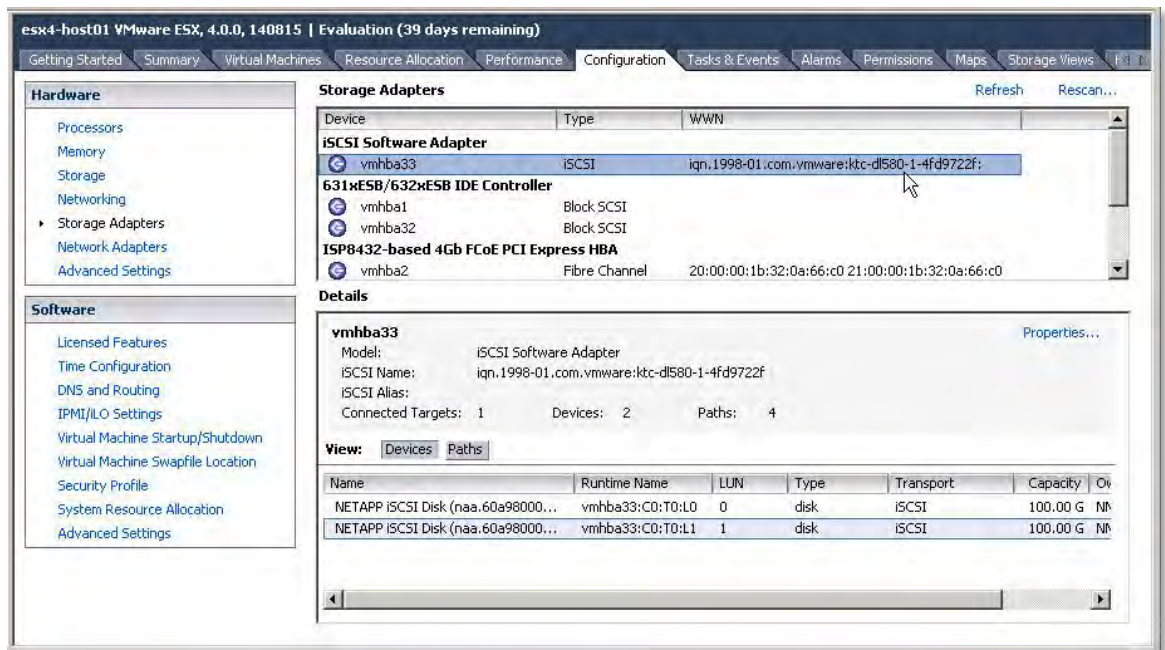


Figure 14) Selecting an iSCSI initiator.



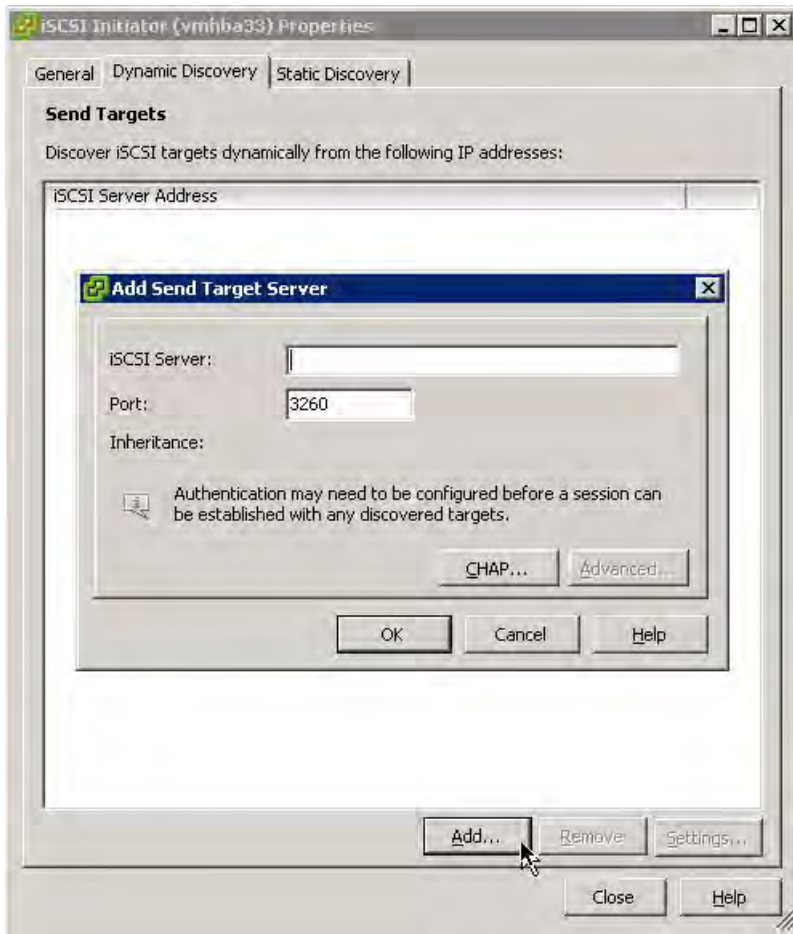


Figure 15) Configuring iSCSI dynamic discovery.

## RESTRICTING ISCSI TARGETS TO PREFERRED INTERFACES

By default NetApp storage arrays provide iSCSI access over every network interface. This default configuration might not be optimal as it can lead to conditions where ESX servers attempt to communicate to interfaces that are unreachable. It is recommended that you disable iSCSI on NetApp network interfaces over which you do not want to send iSCSI traffic.

Data ONTAP allows this filtering to be accomplished either on a host-by-host basis using iSCSI access lists or on a global basis by unbinding iSCSI to a specific interface or set of interfaces. It is recommended that you configure iSCSI access restrictions with one of these two methods.

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

1	Connect to the FAS system console (using either SSH, Telnet, or Console connection).
2	To create an iSCSI access list type: iscsi interface accesslist add <ESX iqn address>
3	Repeat step 2 for each ESX host in the data center.
4.	To verify the iSCSI access list type: iscsi interface accesslist show

Globally disabling iSCSI traffic on a set of network interfaces is less granular than iSCSI access lists; however, it is much simpler to configure. In order to do so, complete the following process.

1	Connect to the FAS system console (using either SSH, Telnet, or Console connection).
2	To disable iSCSI on an interface type: iscsi interface disable <interface hw address>
3	To verify the iSCSI bindings type: iscsi interface show

Alternatively iSCSI restrictions can be configured in within ESX 4.0 and ESXi 4.0. This method provides for another way to accomplish optimal I/O communications in the event that an appropriate restriction cannot be configured with Data ONTAP.

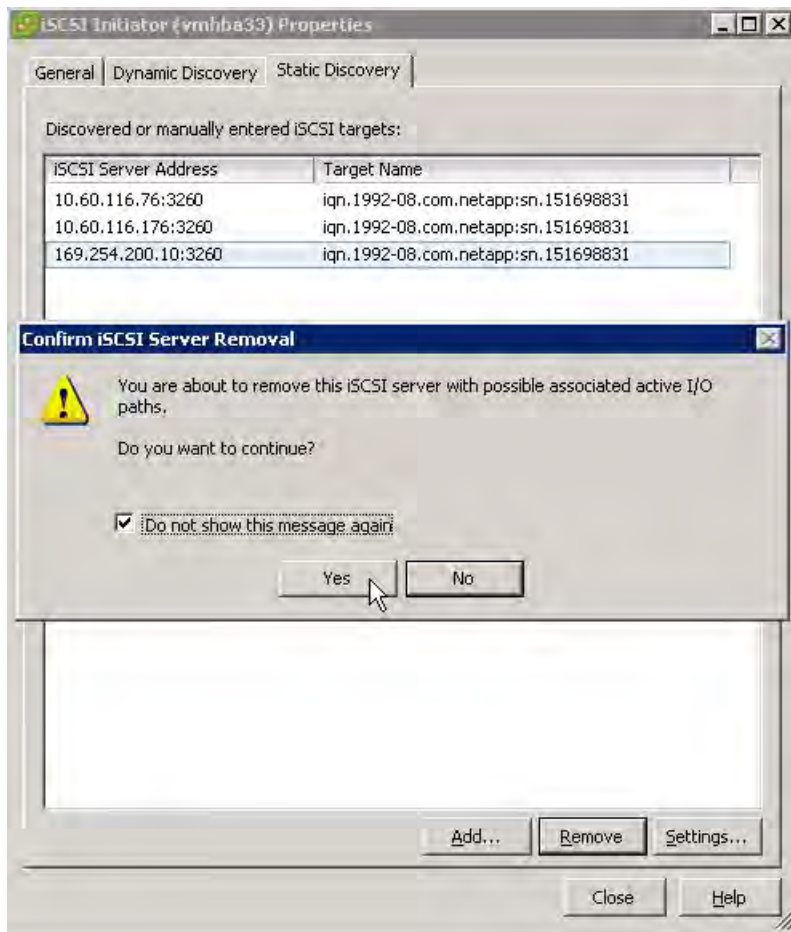


Figure 16) Configuring iSCSI to filter or restrict access to undesired iSCSI targets.

## 5 VMWARE NATIVE MULTIPATHING

VMware ESX servers ship with a native multipathing solution for FC, FCoE, and iSCSI storage networks which enable high-performance data access and enhanced storage resiliency. With the release of ESX and ESXI 4.0, VMware has introduced the concept of a Pluggable Storage Architecture (PSA) that introduced several new concepts to its Native Multipathing (NMP). This section will review leveraging the Storage Array Type Plug-in (SATP) and the Path Selection Plug-in (PSP) along with the Asymmetric Logical Unit Access protocol (ALUA).

### 5.1 DEFAULT NMP SETTINGS

Connecting a NetApp array to an ESX 4.0 server will result in the array being identified as an active-active storage controller, and the VMware native multipathing path selection policy will apply the "Fixed" multipathing policy. This configuration is identical to the default behavior with ESX 3.5.

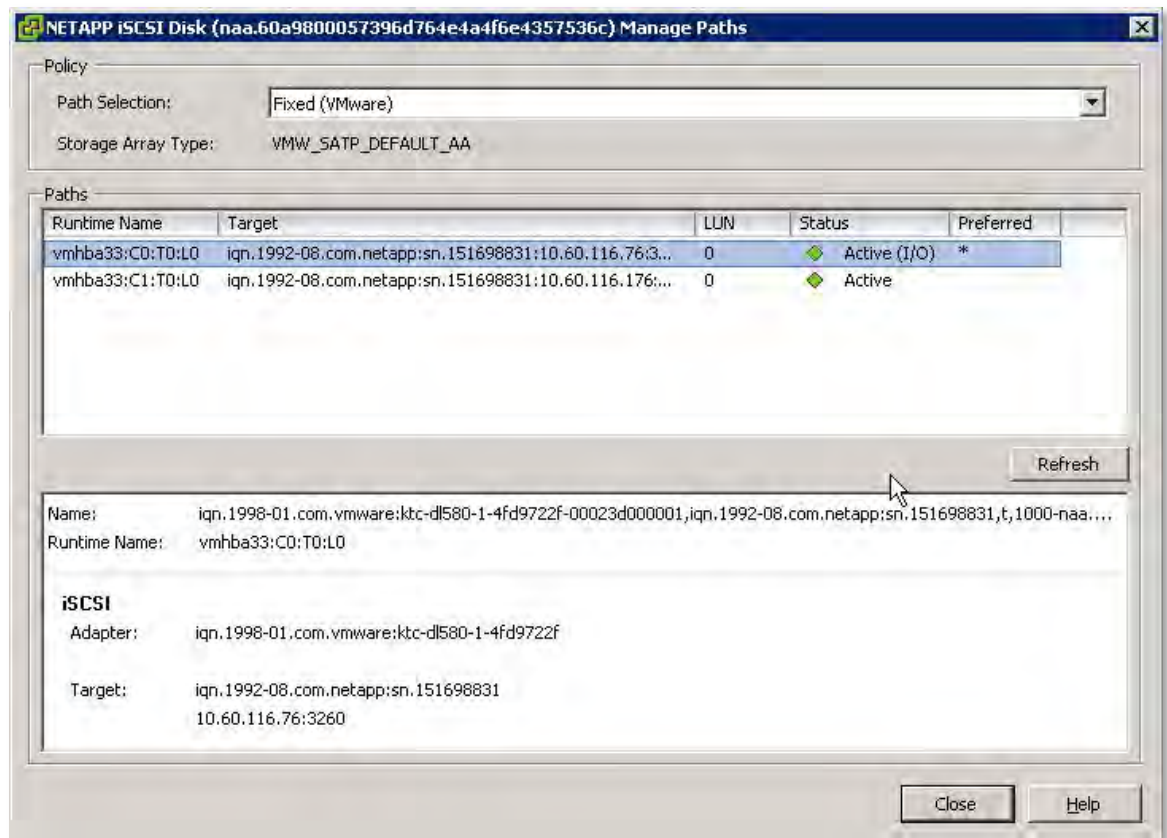


Figure 17) Default path selection policy.

Deployments that leverage the "Fixed" multipathing policy will be required to manually identify and set the I/O to traverse the primary FC paths. In addition, users of this configuration will be required to manually load balance I/O across the primary paths. The NetApp EHU can automate this process for environments that prefer the NMP "Fixed" PSP.

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

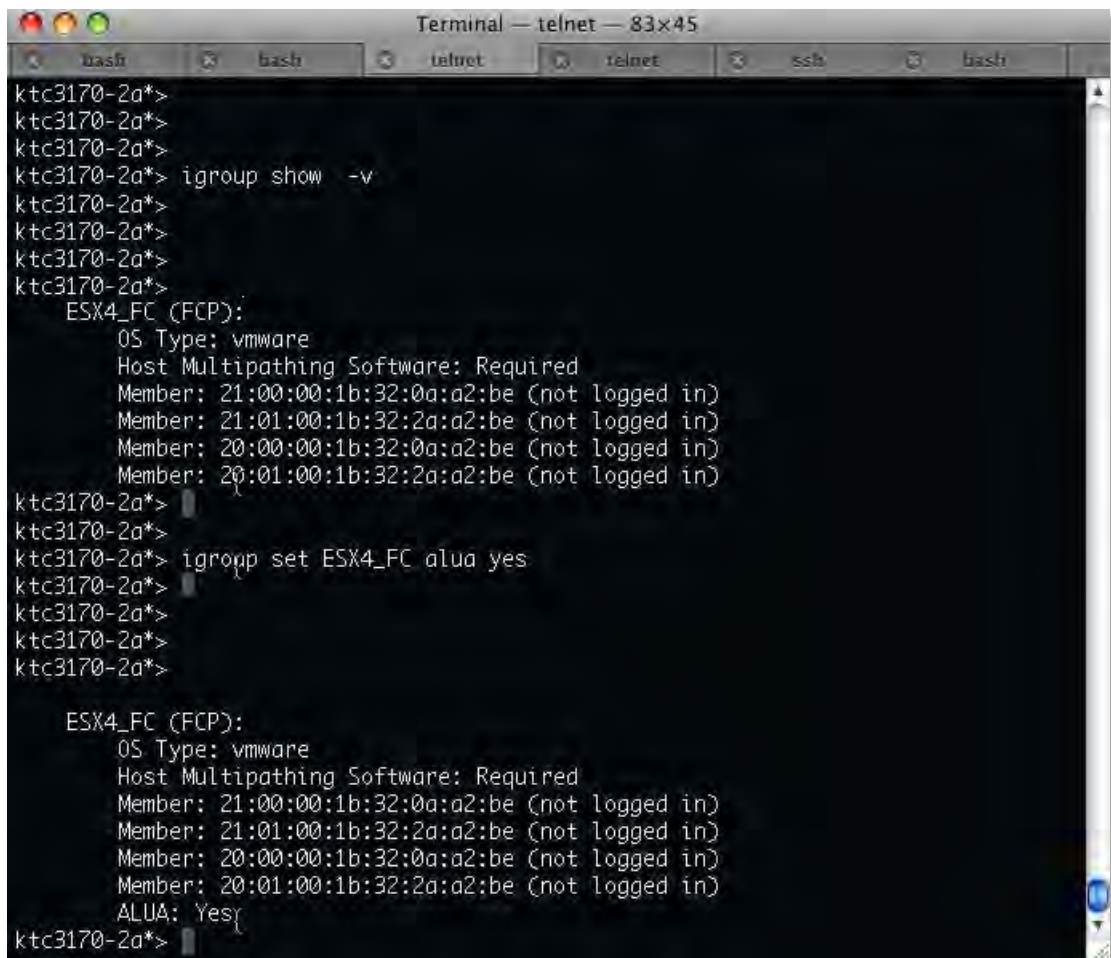
## 5.2 ENABLING ALUA

NetApp and VMware support the Asymmetric Logical Unit Access protocol. ALUA allows for the auto-negotiation of paths between SCSI target devices and target ports enabling dynamic reconfiguration. Enabling ALUA on NetApp initiator groups will result in a more dynamic, or plug-n-play like, architecture.

Note: ALUA is supported with ESX and ESXI for FC and FCoE. Support for iSCSI is not required as iSCSI addresses this functionality natively within the protocol.

ALUA is enabled on ESX 4.0 by default. To enable ALUA on a NetApp storage array, complete the following steps.

1	Log in to the NetApp console.
2	From the storage appliance console, run <code>igroup set &lt;igroup-name&gt; alua yes</code>
3	Repeat step 2 for each LUN accessed by ESX.
4	Results can be verified by running <code>igroup show -v &lt;igroup-name&gt;</code>



```
Terminal — telnet — 83x45
kct3170-2a*>
kct3170-2a*>
kct3170-2a*>
kct3170-2a*> igroup show -v
kct3170-2a*>
kct3170-2a*>
kct3170-2a*>
kct3170-2a*>
kct3170-2a*>
  ESX4_FC (FCP):
    OS Type: vmware
    Host Multipathing Software: Required
    Member: 21:00:00:1b:32:0a:a2:be (not logged in)
    Member: 21:01:00:1b:32:2a:a2:be (not logged in)
    Member: 20:00:00:1b:32:0a:a2:be (not logged in)
    Member: 20:01:00:1b:32:2a:a2:be (not logged in)
kct3170-2a*>
kct3170-2a*>
kct3170-2a*> igroup set ESX4_FC alua yes
kct3170-2a*>
kct3170-2a*>
kct3170-2a*>
kct3170-2a*>
  ESX4_FC (FCP):
    OS Type: vmware
    Host Multipathing Software: Required
    Member: 21:00:00:1b:32:0a:a2:be (not logged in)
    Member: 21:01:00:1b:32:2a:a2:be (not logged in)
    Member: 20:00:00:1b:32:0a:a2:be (not logged in)
    Member: 20:01:00:1b:32:2a:a2:be (not logged in)
    ALUA: Yes
kct3170-2a*>
```

Figure 18) Enable and verify ALUA settings.

### 5.3 DEFAULT NMP SETTINGS WITH ALUA ENABLED

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

Deployments that leverage ALUA along with the MRU multipathing policy will be required to manually load balance I/O across the primary paths. The result of only enabling ALUA is a reduction in some of the configuration requirements. For deployments that prefer a complete "plug-n-play" architecture, enable ALUA on the NetApp storage array and configure the Round Robin PSP.

### 5.4 CONFIGURING THE ROUND ROBIN PSP

There are two ways to configure a PSP. The recommended way is to set the ESX system default PSP for the VMware Default ALUA SATP to use the Round Robin PSP. Alternatively, one can manually manage datastore and LUN policies as they did in ESX 3.5 inside of the virtual infrastructure client.

#### SETTING THE DEFAULT PSP FOR ALUA TO ROUND ROBIN

1	Connect to the CLI of an ESX or ESXi server.
	<p>From the console, run</p> <pre>esxcli nmp satp setdefault -psp &lt;PSP type&gt; -satp &lt;SATP type&gt;</pre> <p>Available PSP types are:</p> <pre>VMW_PSP_RR VM_PSP_FIXED VM_PSP_MRU</pre> <p>Available SATSP types for NetApp arrays are:</p> <pre>VMW_SATP_DEFAULT_AA VM_SATP_ALUA</pre>
2	An example of executing this command is displayed below.
	<p>Verify the results of this command by typing</p> <pre>esxcli nmp satp list</pre>
3	An example of executing this command is displayed below.
4	Repeat steps 1–3 on each ESX or ESXi server.

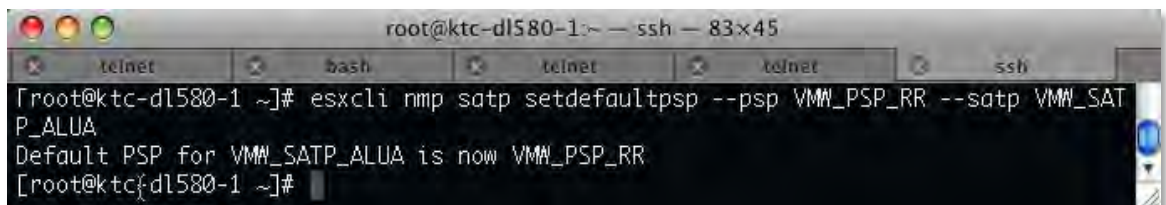


Figure 19) Setting the RR PSP to the ALUA SATP.

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

Figure 20) Listing the active PSP to SATP configurations.

## MANUALLY SETTING THE PSP FOR A DATASTORE

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 21).
6	In the Properties dialog box, click the Manage Paths button.
7	Set the multipathing policy to Round Robin.

The screenshot shows the VMware ESX Configuration console for host 'esx4-host01'. The 'Configuration' tab is active, and the 'Storage' section is expanded. The 'Datastores' table lists two datastores: 'Datastore 2' (NETAPP iSCSI Disk, 99.75 GB capacity, 99.20 GB free) and 'Storage1' (Local VMware Disk, 67.75 GB capacity, 58.99 GB free). 'Datastore 2' is selected, and its details are shown below. The 'Datastore Details' section includes a capacity gauge, location, and a 'Path Selection' dropdown set to 'Round Robin'. The 'Properties' section shows 'Volume Label: Datastore 2' and 'Datastore Name: Datastore 2'. The 'Extents' section shows 'Total Formatted Capacity: 99.75 GB'. The 'Formatting' section shows 'File System: VMFS 3.33' and 'Block Size: 1 MB'.

Identification	Status	Device	Capacity	Free	Type	Last Update
Datastore 2	Normal	NETAPP iSCSI Disk...	99.75 GB	99.20 GB	vmfs3	4/10/2009 1:27:32 P
Storage1	Normal	Local VMware Disk...	67.75 GB	58.99 GB	vmfs3	4/10/2009 1:27:32 P

Path Selection		Properties		Extents	
Round Robin		Volume Label:	Datastore 2	NETAPP iSCSI Disk (naa.60...	100.00 GB
		Datastore Name:	Datastore 2	Total Formatted Capacity	99.75 GB

Paths		Formatting	
Total:	2	File System:	VMFS 3.33
Broken:	0	Block Size:	1 MB
Disabled:	0		

Figure 21) Selecting a datastore.



## MANUALLY SETTING THE PSP FOR A LUN

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 22).
8	In the Manage Paths window, set the multipathing policy and preferred path for all of the highlighted LUNs (see Figure 23).

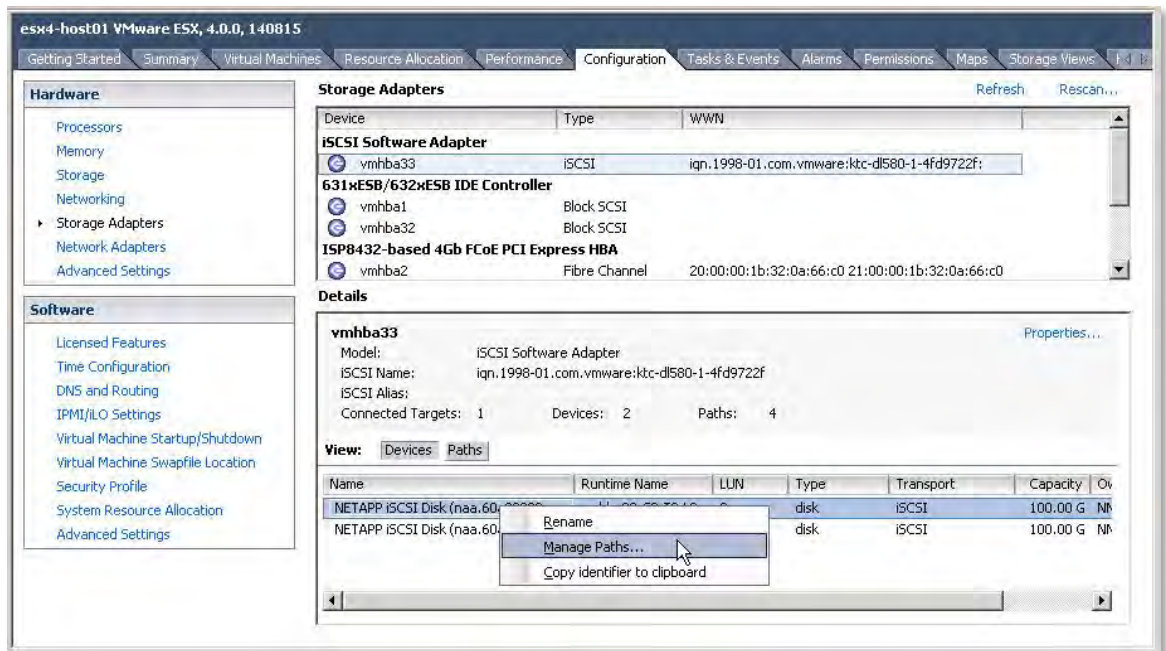


Figure 22) Selecting a SCSI target to manage its paths.

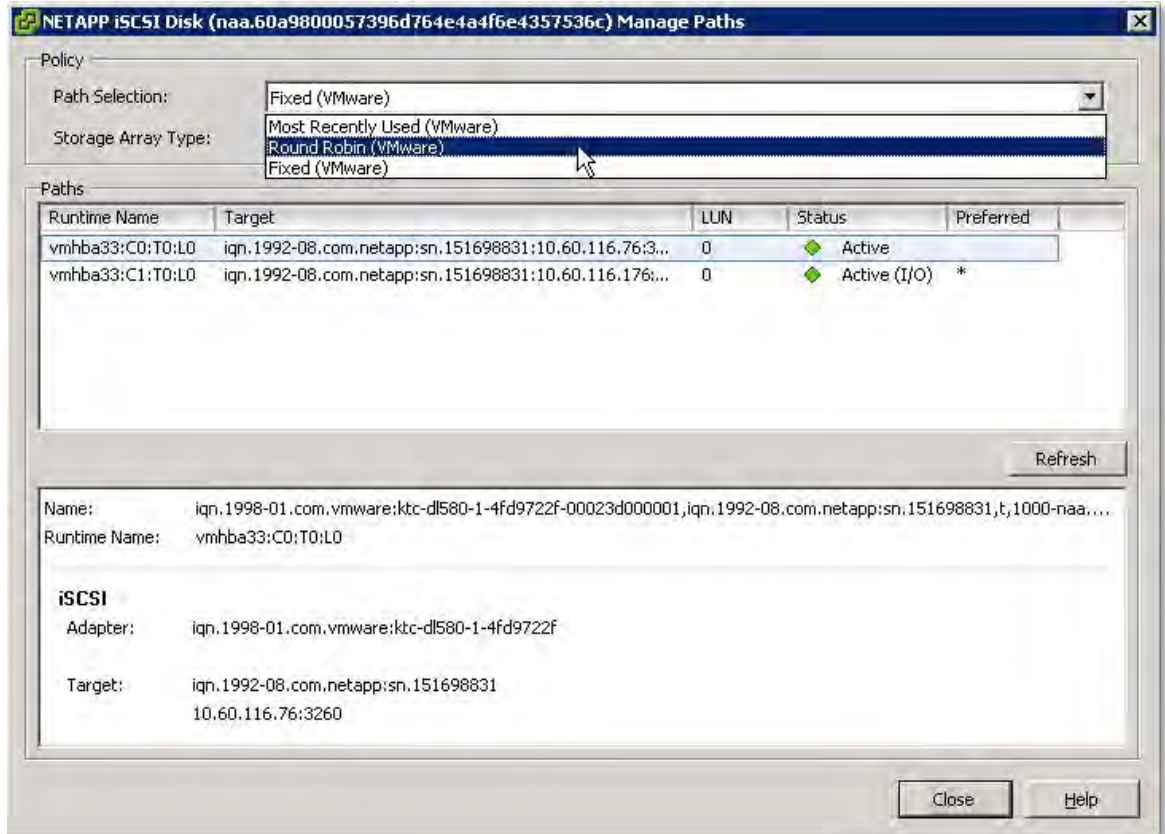


Figure 23) Setting the PSP for a SCSI target.

## 6 NFS STORAGE RECOMMENDATIONS

### 6.1 INCREASING THE NUMBER OF NFS DATASTORES

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

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

To make this change, follow these steps from within the virtual infrastructure client.

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 64 (see Figure 24).
7	In the pop-up window, left pane, select Net.
8	Change the value of Net.TcplpHeapSize to 30.
9	Change the value of Net.TcplpHeapMax to 120.
10	Repeat the steps for each ESX Server.

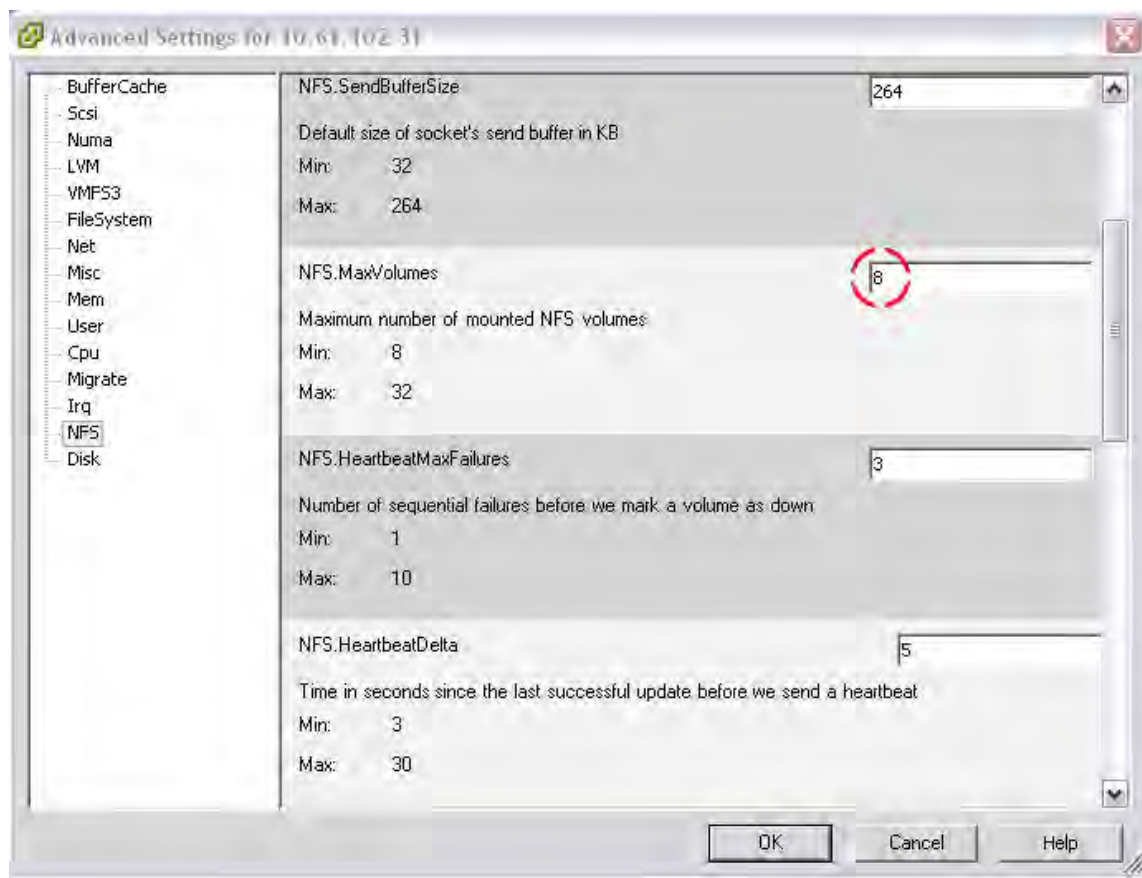


Figure 24) Increasing the maximum number of NFS datastores.

## 6.2 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 addition, 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 security 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 make sure that the FAS user mapping is setup correctly in order to enable windows user access to this data. For more information on this subject review the section File Sharing Between NFS and CIFS in the Data 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 <code>vol options &lt;vol-name&gt; no_atime_update on</code>
3	From the storage appliance console, run <code>qtree security &lt;volume path&gt; UNIX</code>
4	Repeat steps 2 and 3 for each NFS accessed volume.

### 6.3 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 7 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 updating these setting are listed below.

#### ESX 4.0 HOST

For optimal availability with NFS datastores, NetApp recommends making the following changes on each ESX 4.0 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.

## 6.4 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 recommended 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>`.

## 6.5 CONNECTING NFS DATASTORES

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

1	Open FilerView ( <a href="http://filer/na_admin">http://filer/na_admin</a> ).
2	Select Volumes.
3	Select Add to open the Volume Wizard (see Figure 25). Complete the Wizard.
4	From the FilerView menu, select NFS.
5	Select Add Export to open the NFS Export Wizard (see Figure 26). 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 27).
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 28).
13	Click Finish.



Figure 25) NetApp Volume Wizard.

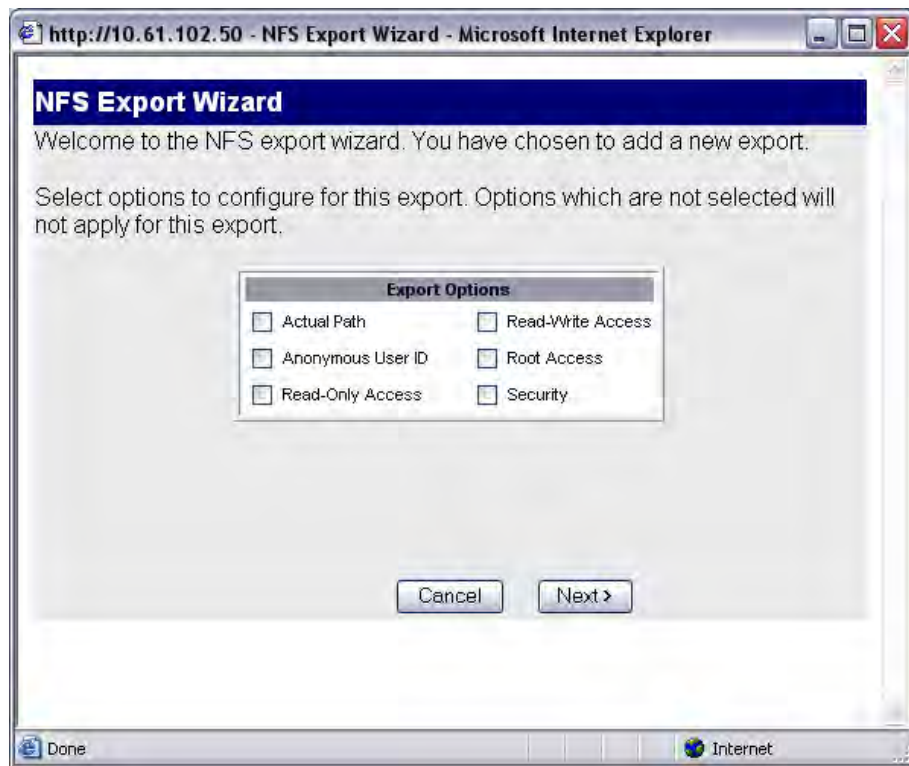


Figure 26) NetApp NFS Export Wizard.



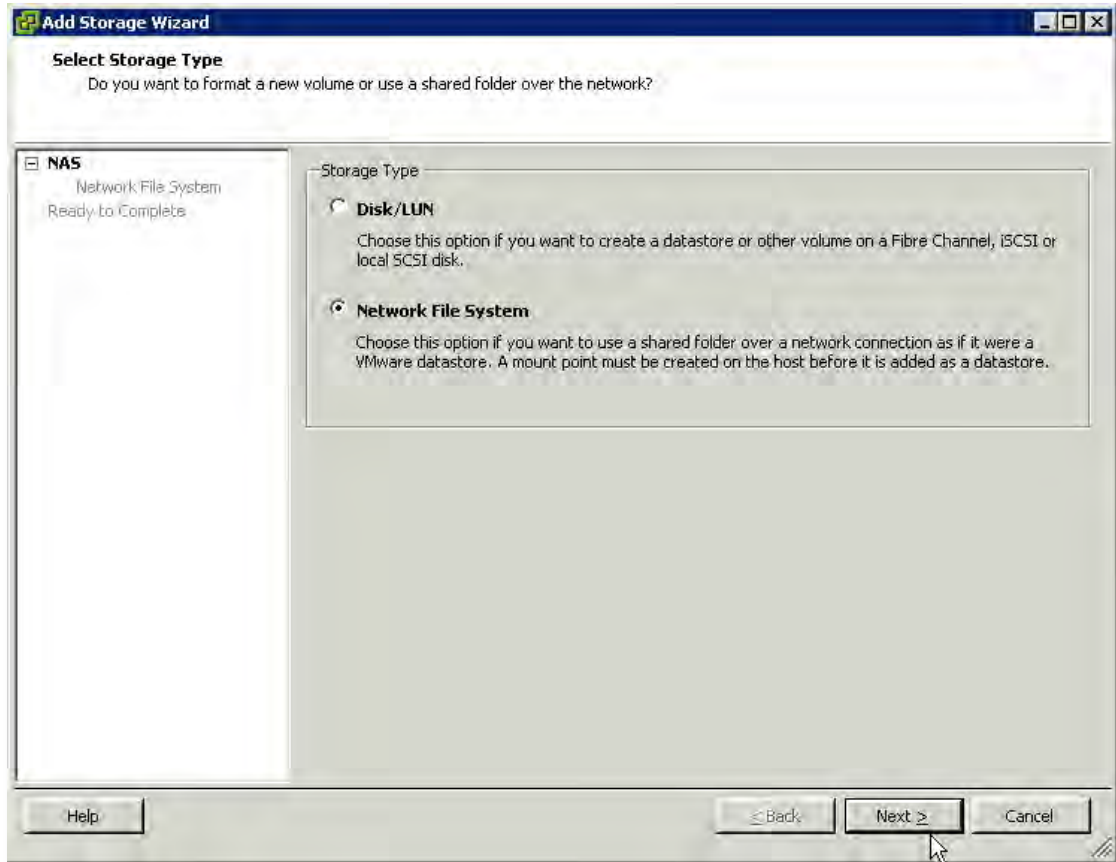


Figure 27) VMware Add Storage Wizard.

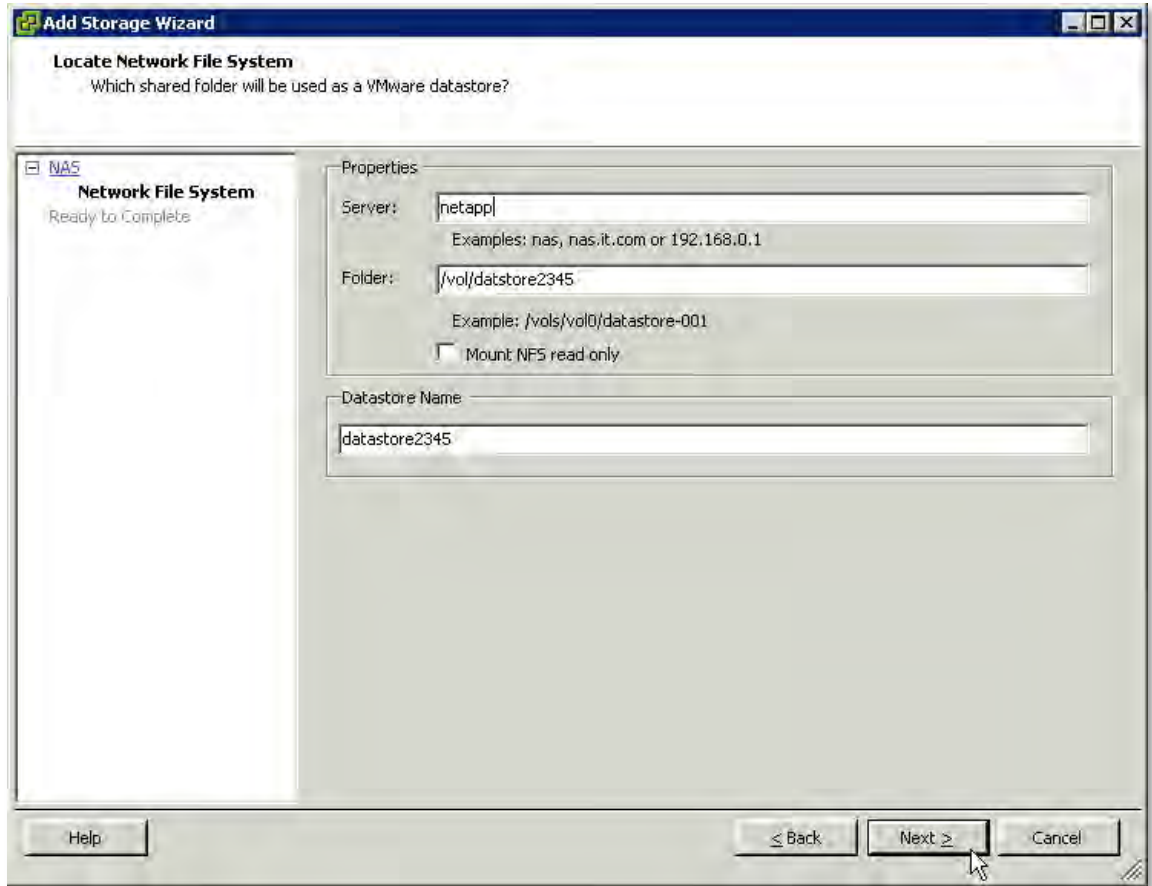


Figure 28) VMware Add Storage Wizard NFS configuration.

## 7 THE NETAPP ESX HOST UTILITIES

NetApp provides the ESX Host Utilities or EHU for simplifying the management of ESX nodes running on NetApp storage arrays. With the EHU customers receive a collection of tools to automate and simplify the configuration of HBAs, sets optimized ESX/ESXi options for NFS, enhances Guest OS (VM) SCSI settings, and provides diagnostic utilities should a support case be opened.

The ESX Host Utilities installs on ESX 4.0 systems, and currently is not supported for ESXi 4.0 systems. The ESX Host Utilities can be downloaded from NOW™ (NetApp on the Web).

### 7.1 INSTALLING THE EHU IN ESX

#### EHU PREREQUISITES

- FC (which includes FCoE), 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 controllers and the hosts.

#### 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 <code>tar -zxf &lt;name of EHU file&gt;.tar.gz</code>
4	Migrate running VMs to other ESX nodes
5	Place the ESX Server in maintenance mode.
6	Complete the EHU installation wizard
7	Run <code>./install</code>
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.

## **7.2 MANUAL CONFIGURATION OF FC HBAS IN ESX**

In previous versions of ESX it was required to manually configure the settings on one's HBAs. This requirement has been eliminated when ESX 4.0 and ESXi 4.0 servers are connected to NetApp storage arrays (running Data ONTAP 7.2.4 or later). If your storage arrays are not running a release of Data ONTAP that is version 7.2.4 or later, consider upgrading them. Alternatively you will have to install the ESX Host Utilities to adjust values of the HBAs in your ESX servers.

## 8 FC AND FCOE 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.

### 8.1 HOST BUS AND CONVERGED NETWORK ADAPTERS

ESX servers and NetApp storage arrays connect to a SAN fabric using host bus adapters (HBAs). Connectivity to FCoE fabrics is enabled through converged network adapters (CNAs). Each HBA/CNA can run as either an initiator (ESX) or as a target (NetApp). Each adapter has a global unique address referred to as a World Wide Port Number (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 adapter ports. For more information on VMware FC best practices and recommendations, see VMware Fibre Channel SAN Configuration Guide.

### 8.2 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 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 FC, FCoE, and iSCSI.

### 8.3 FC AND FCOE 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 recable the fabric without having to redo the zone information.

#### ZONING RECOMMENDATION

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

## 9 ETHERNET STORAGE NETWORKING BEST PRACTICES

NetApp recommends using dedicated resources for storage traffic whenever possible. With Ethernet 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

### 9.1 10 GIGABIT ETHERNET

VMware ESX 4 and ESXi 4 include 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.

### 9.2 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.

### 9.3 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 and 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 and FAS systems. See Figure 29.

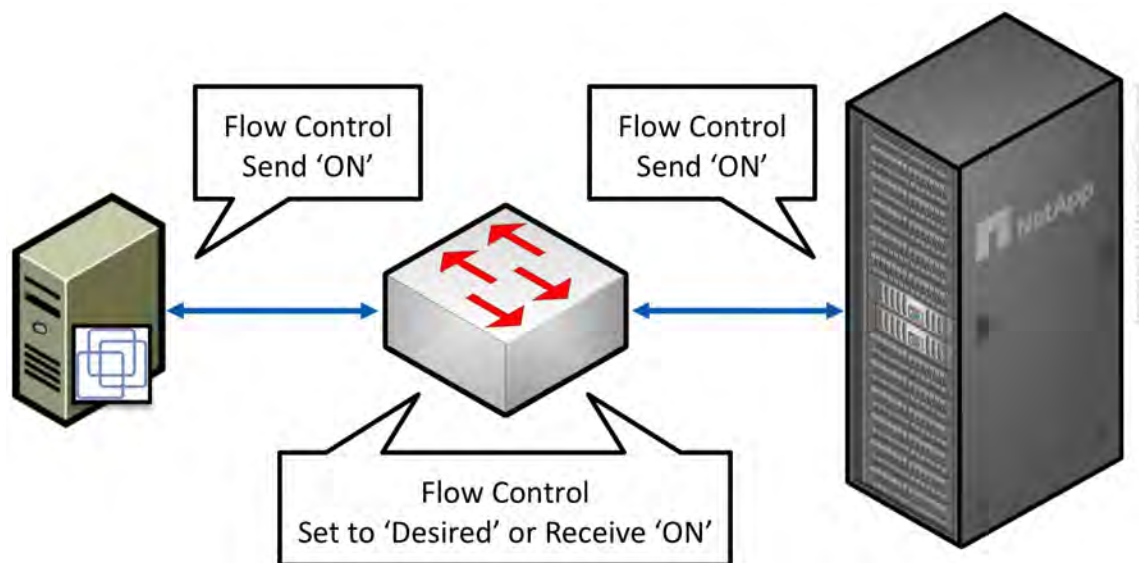


Figure 29) Configuring flow control settings.

## 9.4 SPANNING TREE PROTOCOL

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

When connecting ESX and 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.

## 9.5 BRIDGE PROTOCOL DATA UNITS

Bridge Protocol Data Units (BPDUs) exchange information about bridge IDs and root path costs within STP. When connecting ESX and 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.

## 9.6 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 configured Ethernet trunks. 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 Multimode 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 an interface to be removed from a standard multimode VIF is loss of link.

Multimode and Dynamic multimode VIFs each use the same algorithm for determining load-balancing. This algorithm is based on source and destination IP or Mac address. 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 interface 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.

## **9.7 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 nonredundant 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 multiswitch 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.



## 10 CONFIGURING ETHERNET STORAGE NETWORKS

### 10.1 HIGHLY AVAILABLE STORAGE DESIGNS 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. Depending on storage protocol this vSwitch may be configured with multiple network adapters.

For iSCSI datastores each VMkernel is configured with a single vmnic. No standby vmnics may exist in the VMkernel.

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

NetApp recommends defining a separate VMkernel for each storage protocol. Doing so makes the configuration of iSCSI with NFS very simple. As an example, see Figure 30. 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.

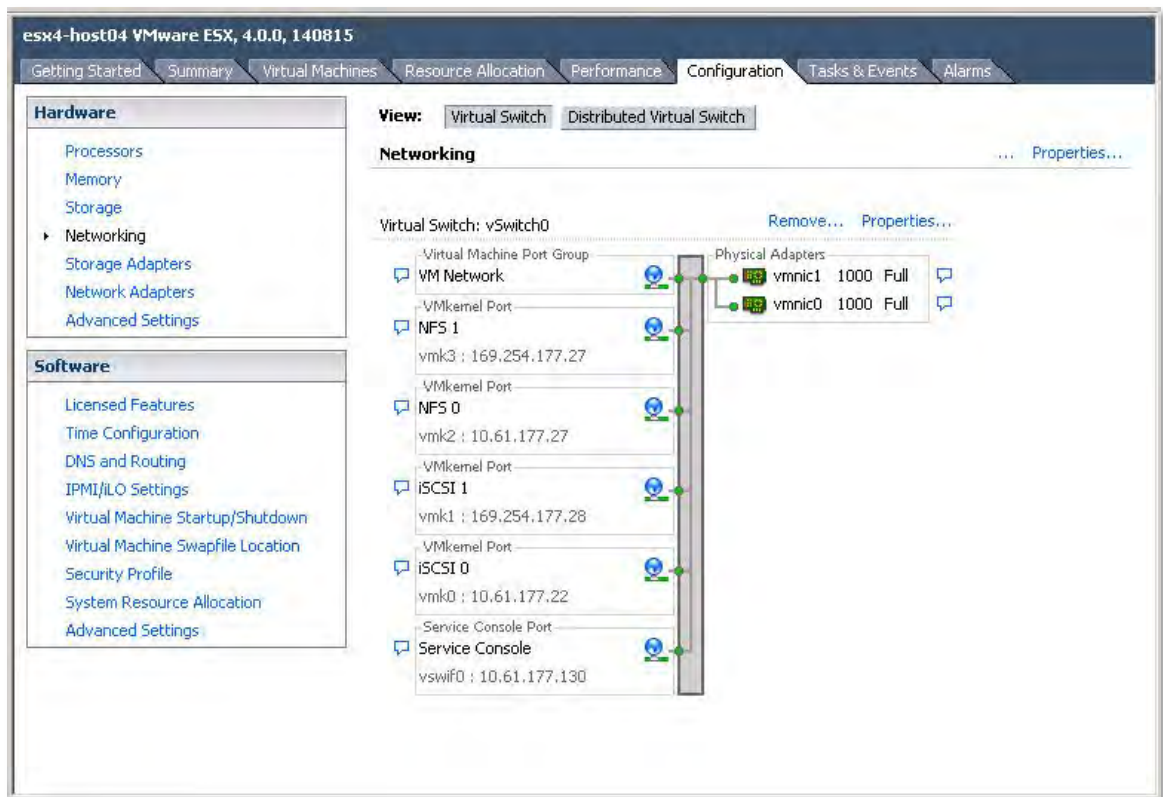


Figure 30) Displaying VMkernel ports for iSCSI.

### **ESX SERVER ADAPTER FAILOVER BEHAVIOR WITH ISCSI**

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 using the second adapter. This failover is managed by VMware's native multipathing, thus there is no need for network failover configuration on the switch or VMkernel. Traffic returns to the original adapter when service to the adapter is restored.

### **ESX SERVER ADAPTER FAILOVER BEHAVIOR WITH NFS**

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 using the second adapter but on the same subnet where it originated. Both subnets are now active on the surviving physical adapter. Traffic returns to the original adapter when service to the adapter is restored. In this scenario Etherchannel provides the network failover.

### **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 increase 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 using 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. This behavior can be changed to send I/O traffic over multiple paths by enabling the Round Robin PSP (as covered in section 4.5). This design results in the aggregation of multiple links while providing fault tolerance for a link, switch, or NIC.

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.

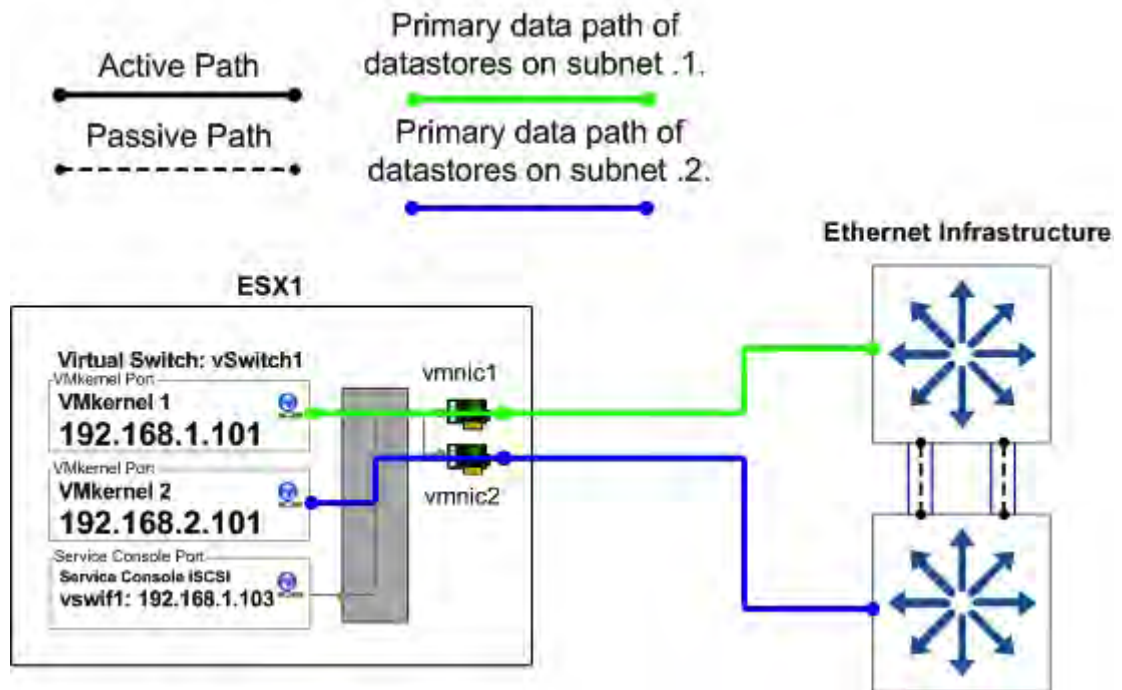


Figure 31) ESX vSwitch1 normal mode operation.

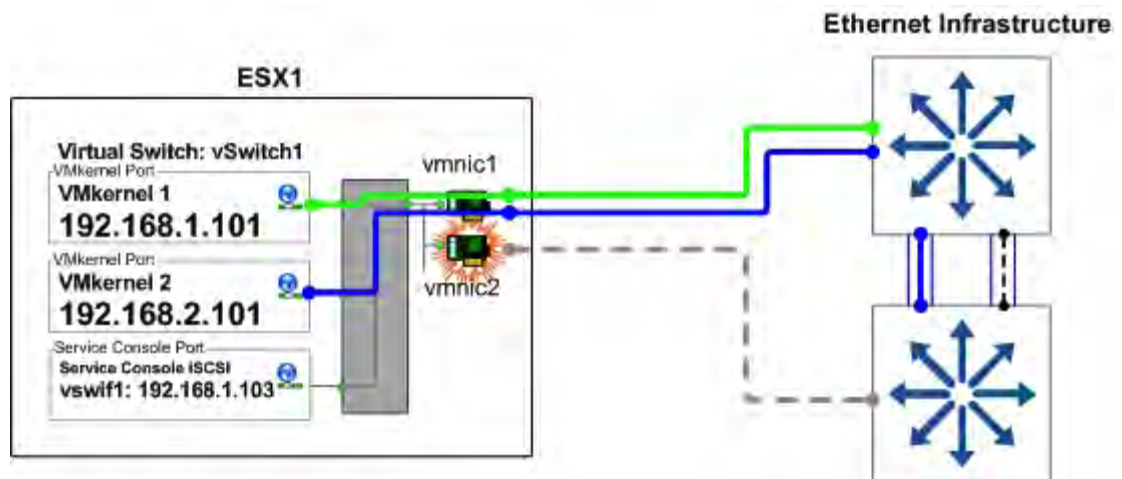


Figure 32) ESX vSwitch1 failover mode operation.

### 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.

## 10.2 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 33, 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.

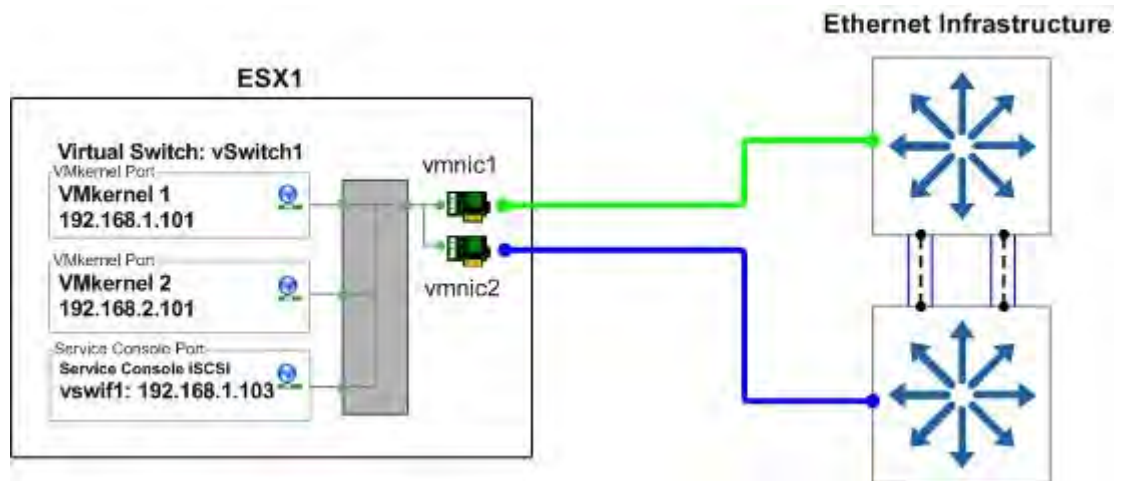


Figure 33) ESX Server physical NIC connections with traditional Ethernet.

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:**
  - Enable the Override vSwitch Failover Order option.
  - For NFS and iSCSI set Active Adapter to vmnic1.
  - For NFS set Standby Adapter to vmnic2.
  - For iSCSI set no Standby Adapters.
- **VMkernel 2:** IP address set to 192.168.2.101.
- **VMkernel2 Port Properties:**
  - Enable the Override vSwitch Failover Order option.
  - For NFS and iSCSI set Active Adapter to vmnic2.
  - For NFS set Standby Adapter to vmnic1.
  - For iSCSI set no Standby Adapters.

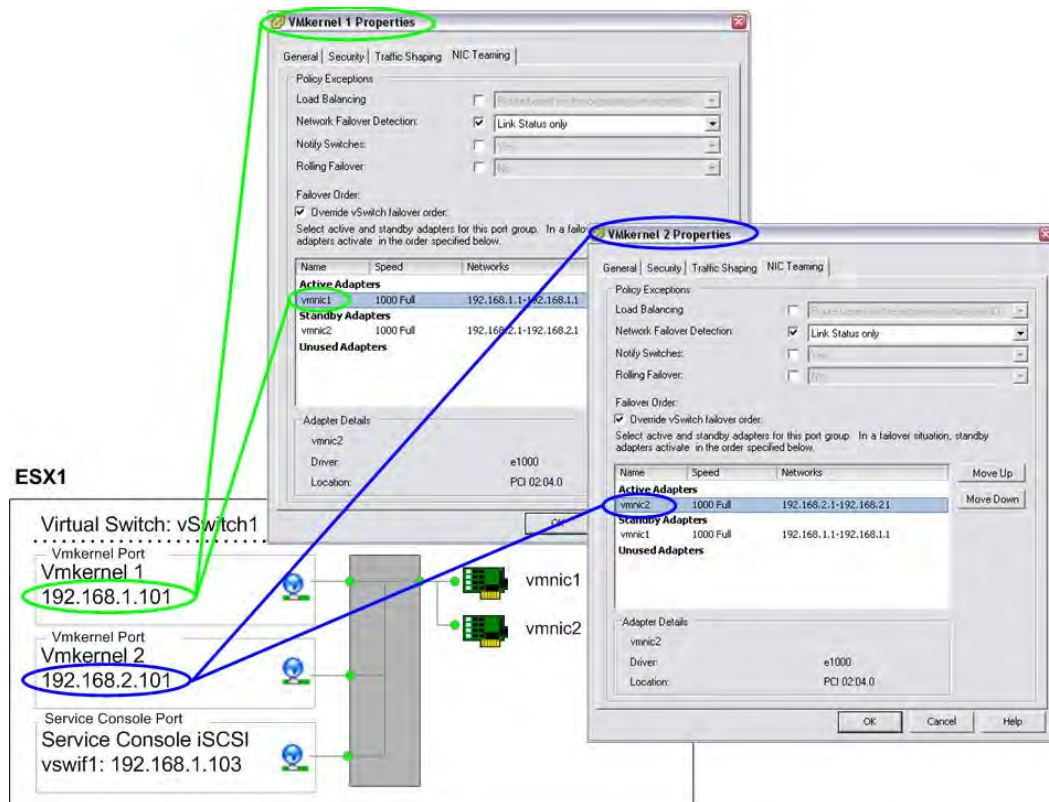


Figure 34) ESX Server VMkernel port properties with traditional Ethernet.

### 10.3 A 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 Figures 35 and 36). 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-MULTIMODE 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.

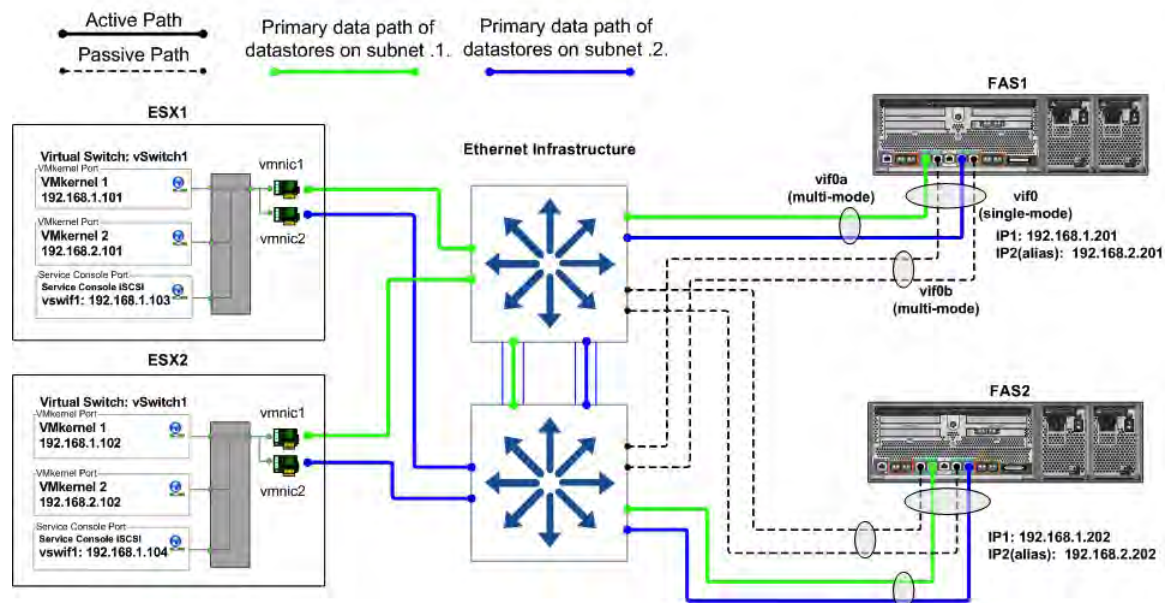


Figure 35) Storage side multimode VIFs.

## 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.

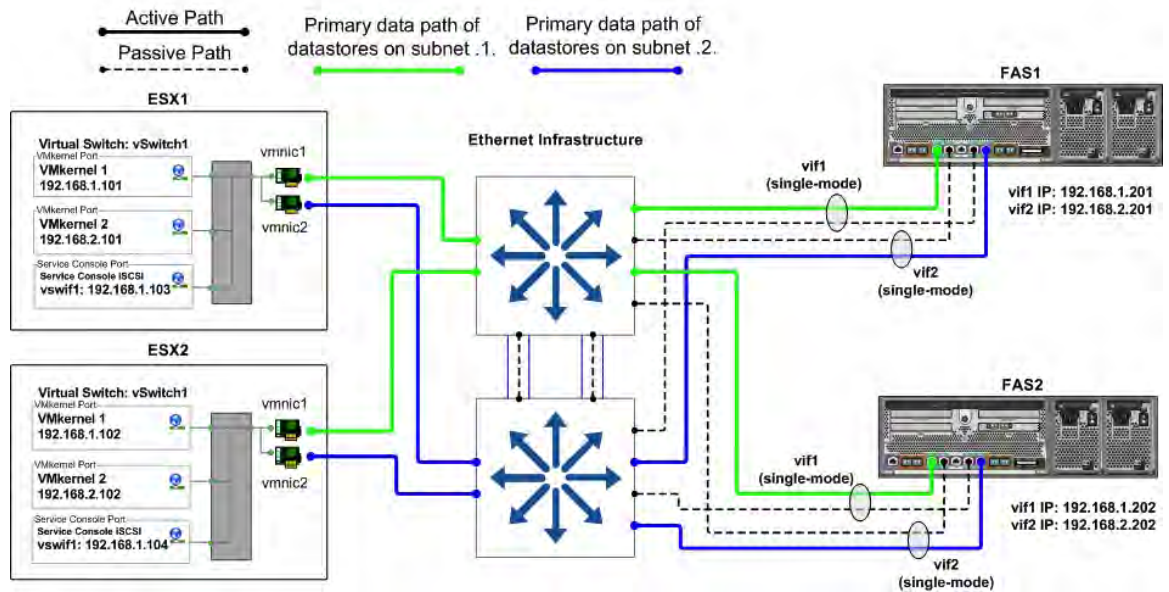


Figure 36) Storage side single-mode VIFs.

## 10.4 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. This behavior can be changed to send I/O traffic over multiple paths by enabling the Round Robin multipathing plug-in (as covered in section 4.5). 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.

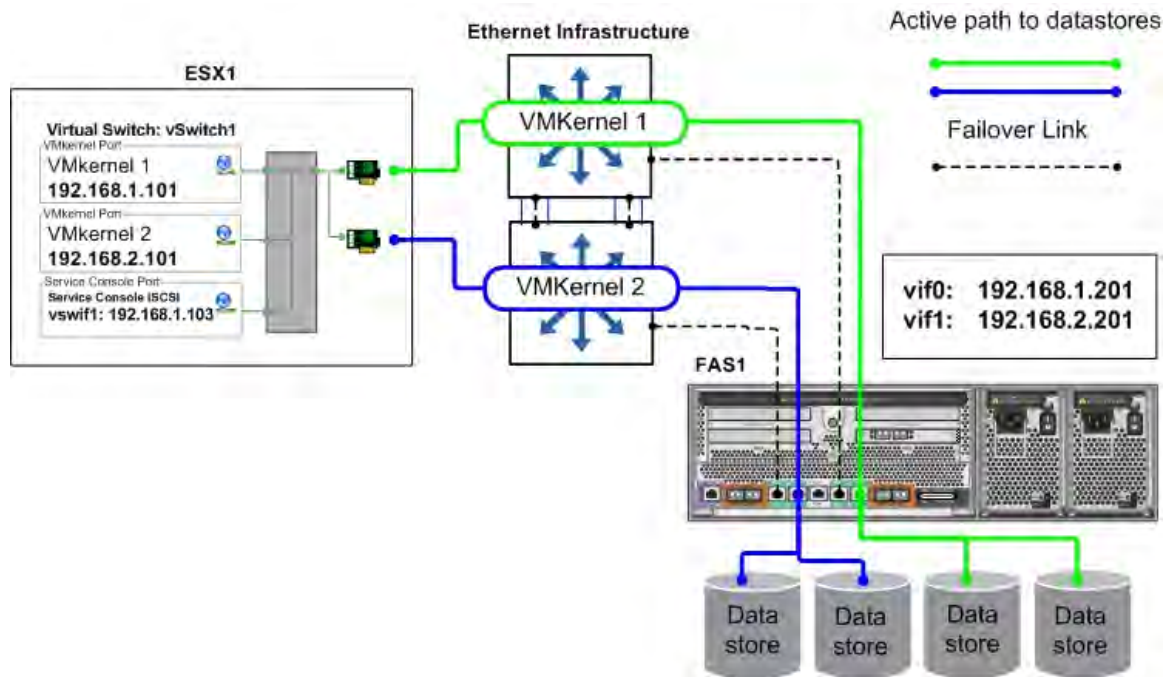


Figure 37) Datastore connections with traditional Ethernet.



## 10.5 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 using more connections.
- Storage controller connection load balancing is automatically managed by IP load balancing policy.
- Requires only one VMkernel port for IP storage to make use of multiple physical paths.

### DISADVANTAGES

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

In the ESX Server configuration shown in the Figure 38, 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.

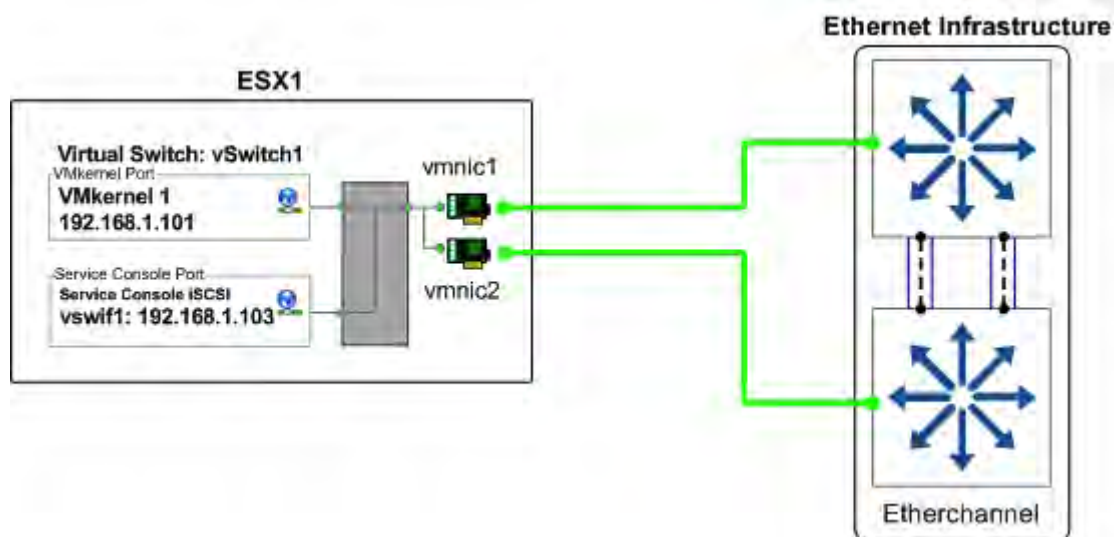


Figure 38) ESX Server physical NIC connections with multi-switch Etherchannel.

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.”

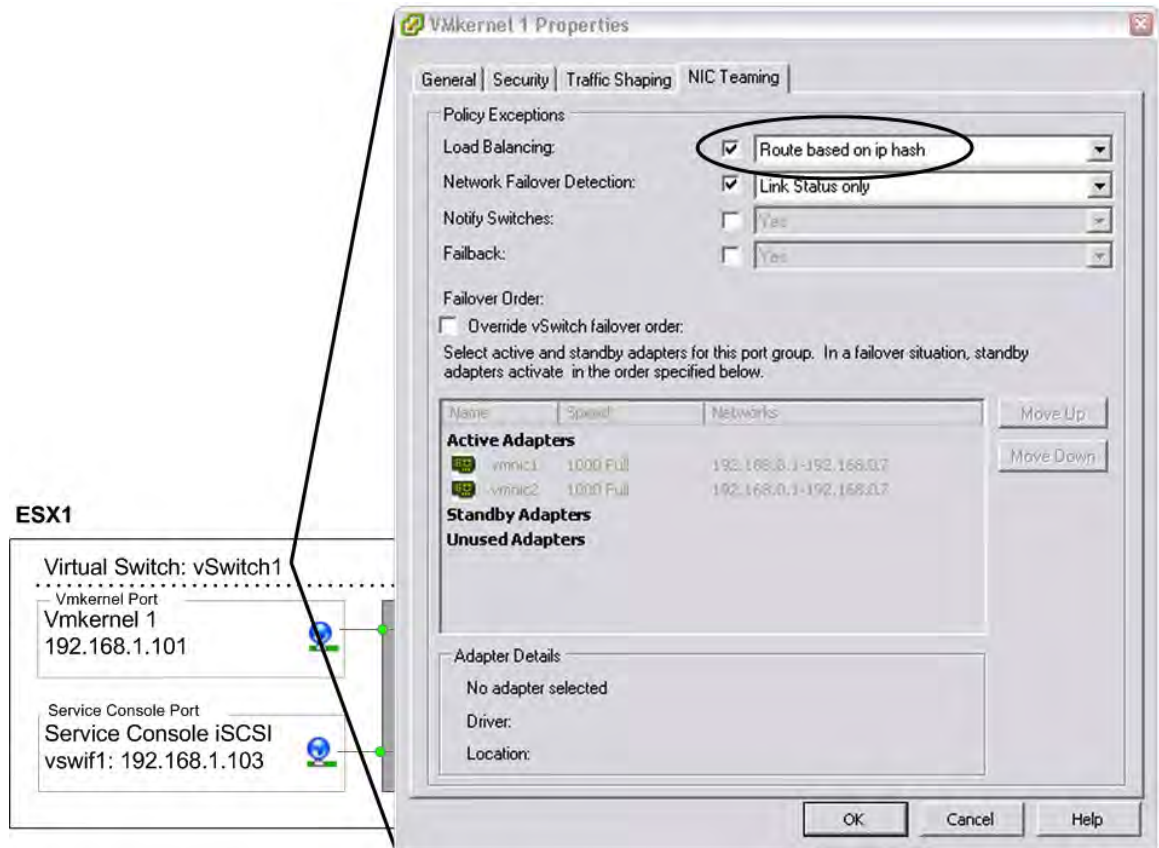


Figure 39) ESX Server VMkernel port properties with multi-switch Etherchannel.

## 10.6 A STORAGE ARCHITECTURE WITH MULTISWITCH 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.

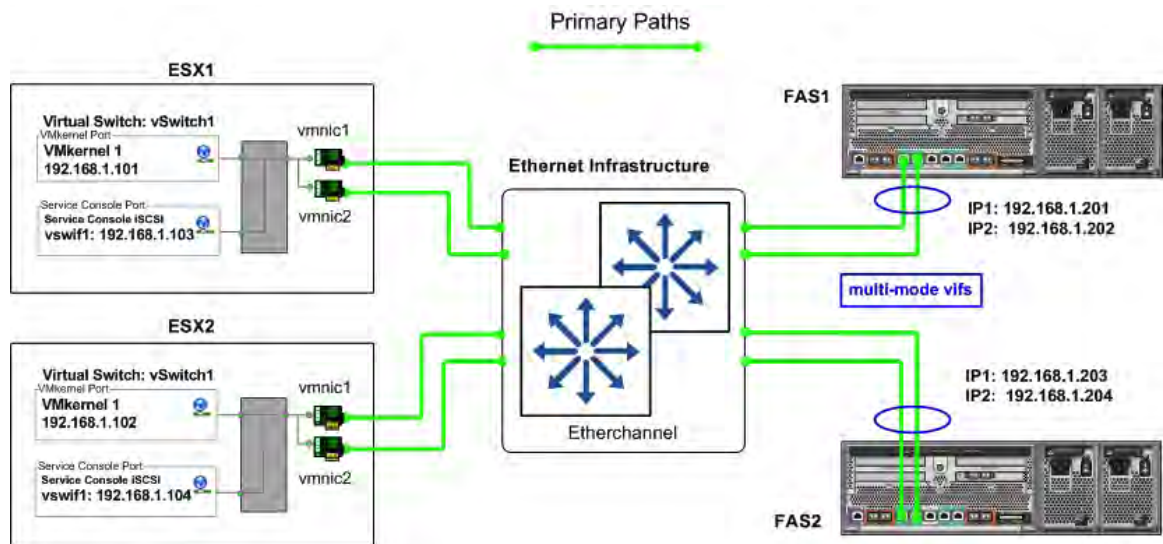


Figure 40) Storage side multimode VIFs using multi-switch Etherchannel.

## 10.7 DATASTORE CONFIGURATION WITH MULTISWITCH 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. This behavior can be changed to send I/O traffic over multiple paths by enabling the Round Robin multipathing plug-in (as covered in section 4.5). 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.

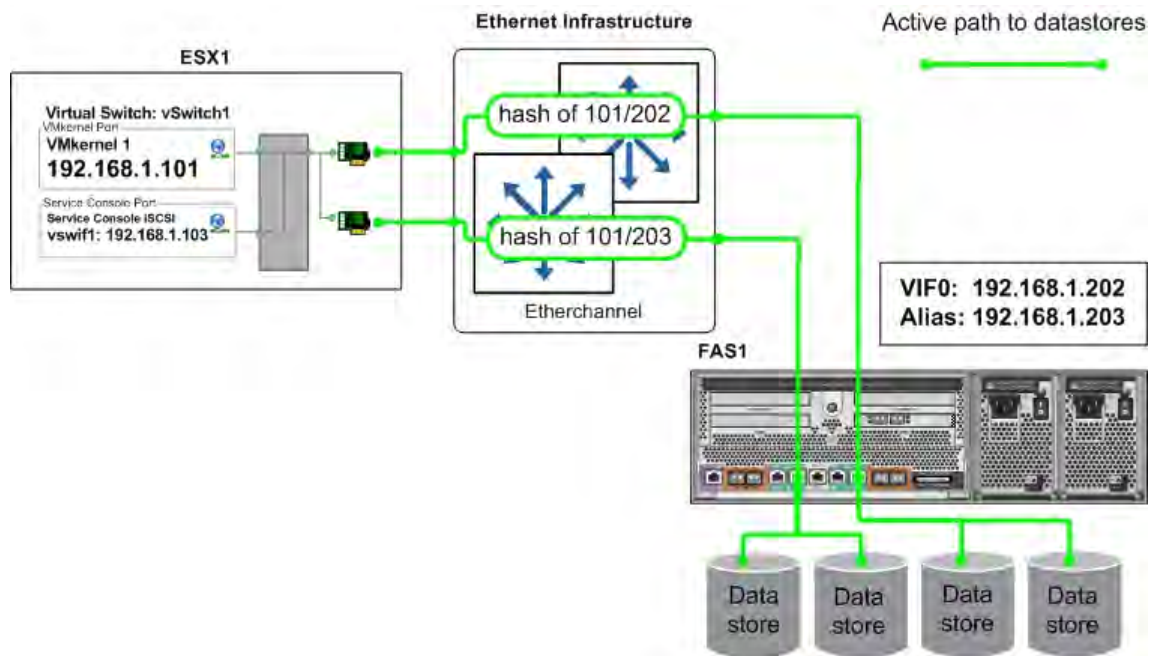


Figure 41) Datastore connections with cross-stack Etherchannel.

## 11 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.

With ESX 4.0, thin-provisioned VMDKs are now available to be created in the virtual infrastructure client with VMFS datastores. By using VMware thin-provisioning technology, one 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.

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.

## 11.1 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 42 shows an example of typical storage consumption in a vSphere deployment.

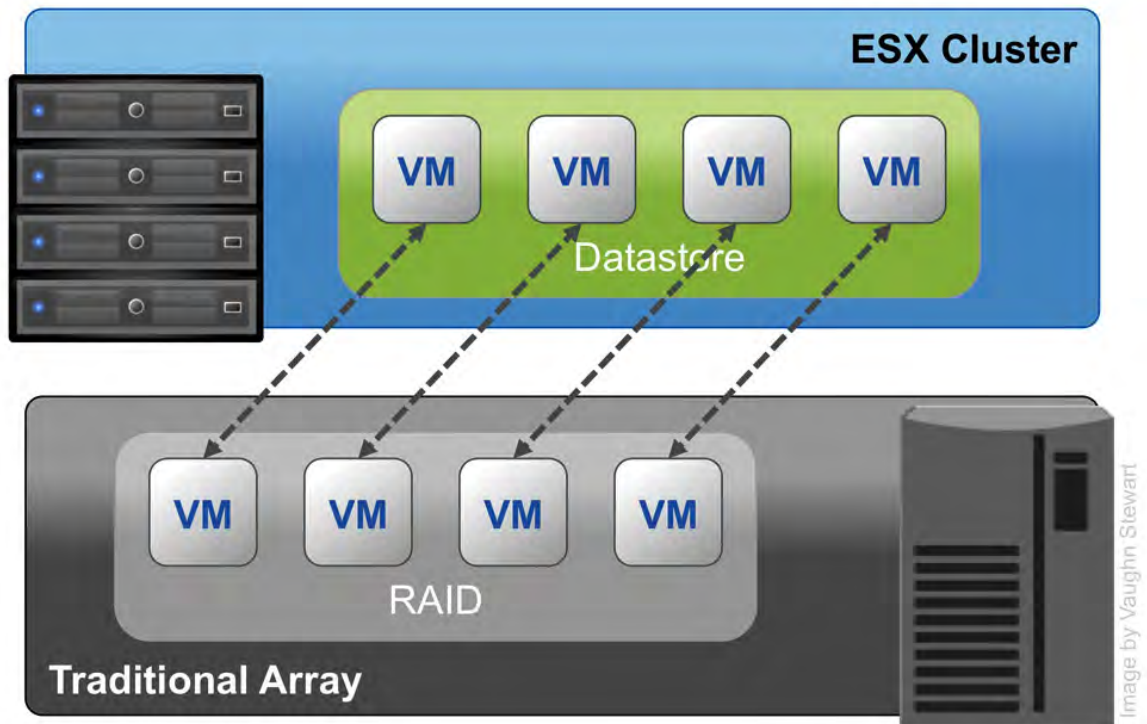


Figure 42) Storage consumption with a traditional array.

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 data center 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 43 shows an example of the impact of deduplication on storage consumption in a vSphere deployment.

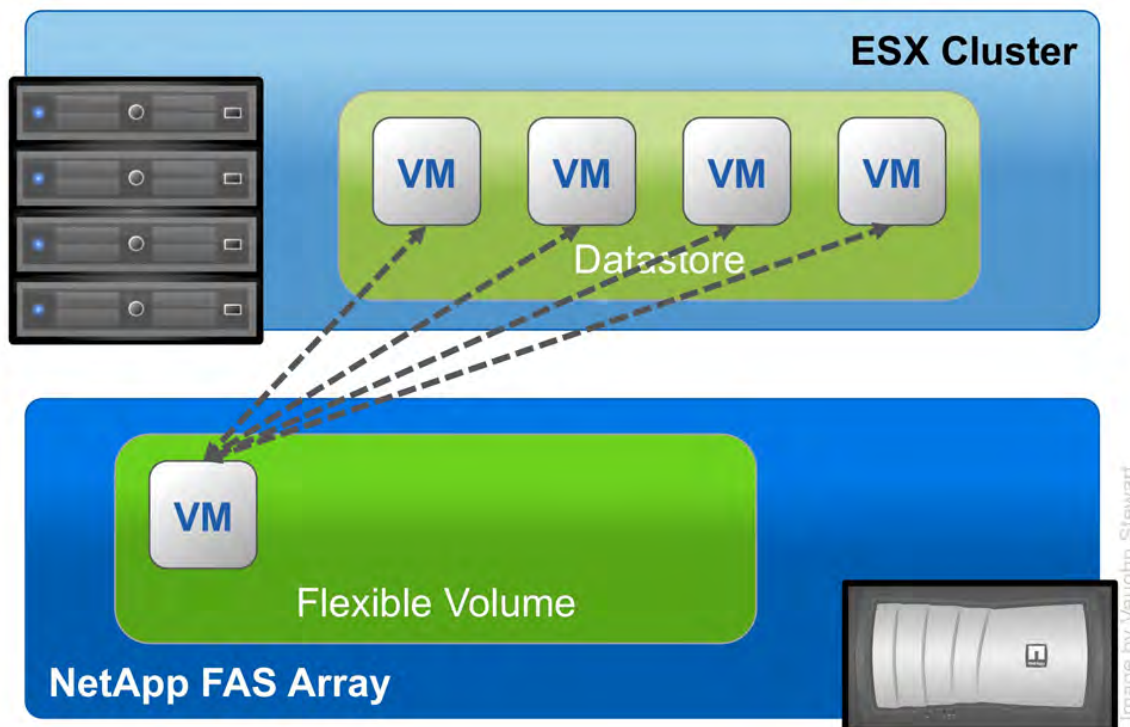


Figure 43) Storage consumption after enabling FAS data deduplication.

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 7.2, "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.

## 11.2 ZERO-COST VIRTUAL MACHINE CLONING

Customers who deploy VMware on NetApp can leverage NetApp's patented FlexClone technology using NetApp's Rapid Cloning Utility. The RCU is a vCenter Server Plug-in that provisions zero-cost, or pre-duplicated, VMware virtual machines and datastores. The RCU is available with VI3, vSphere support will be available with the 2.2 release.

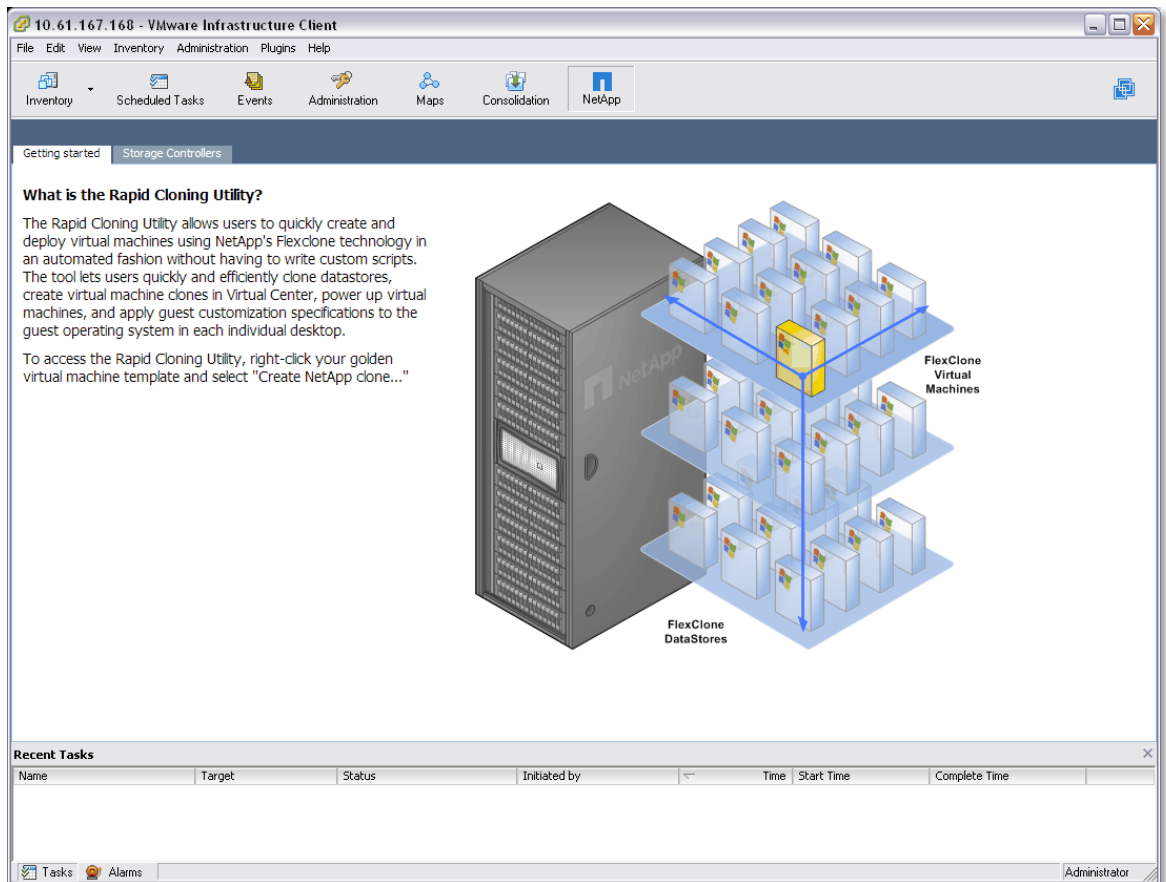


Figure 44) The NetApp Rapid Cloning Utility version 2.0.

## 11.3 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 ( <a href="http://filer/na_admin">http://filer/na_admin</a> ).
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 45).
10	Click Next and then click Finish.

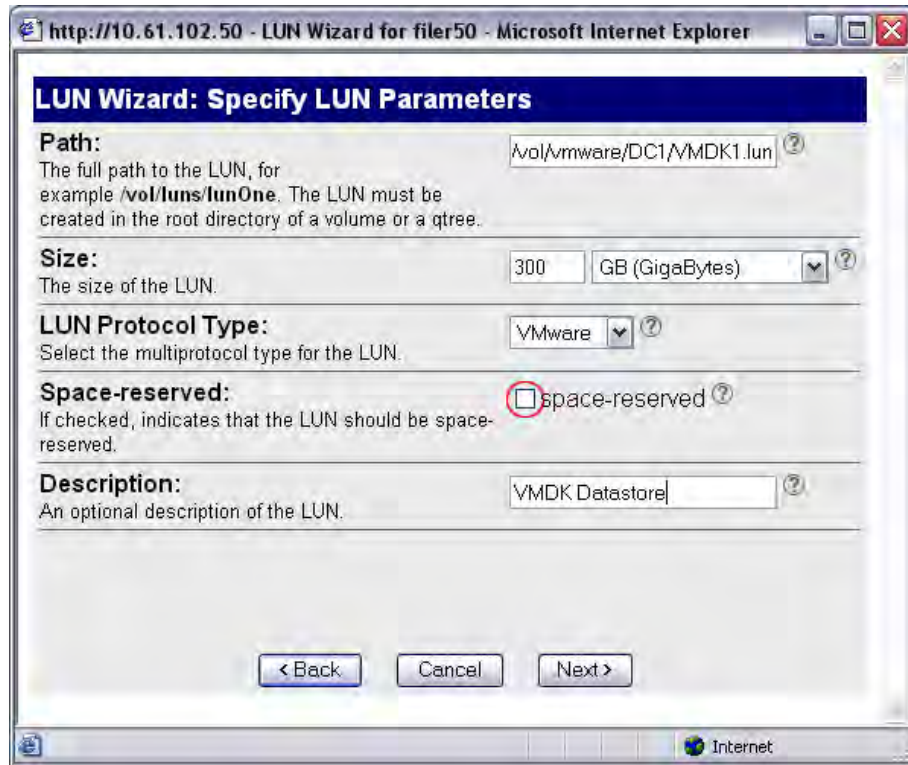


Figure 45) Enabling thin provisioning on a LUN.

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 copy deletion, and LUN fractional reserve.

Volume Auto Size is a policy-based space management feature in Data ONTAP that allows a volume to grow in defined increments up to a predefined limit when the volume is nearly full. For VMware 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: <code>vol autosize &lt;vol-name&gt; [-m &lt;size&gt;[k m g t]] [-i &lt;size&gt;[k m g t]] on.</code>

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: <code>snap autodelete &lt;vol-name&gt; commitment try trigger volume target_free_space 5 delete_order oldest_first.</code>
3	Set Volume Auto Delete Policy: <code>vol options &lt;vol-name&gt; try_first volume_grow.</code>

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 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: <code>vol options &lt;vol-name&gt; fractional_reserve 0.</code>

## 12 VIRTUAL MACHINE BEST PRACTICES

### 12.1 WINDOWS VM FILE SYSTEM PERFORMANCE OPTION

#### OPTIMIZING WINDOWS FILE SYSTEM FOR OPTIMAL I/O 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 <code>CMD</code>
3.	Enter <code>fsutil behavior set disablelastaccess 1</code>

### 12.2 ENSURING OPTIMUM VM AVAILABILITY

#### OPTIMIZING VM SCSI BUS

In Section 7 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 that can be mounted by a VM in order to configure its local SCSI to values that are optimal for running in a virtual infrastructure.

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 <code>tar -zxf &lt;name of EHU file&gt;.tar.gz</code> .
4	From within vCenter Server select a VM to upgrade, right-click 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 to the 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.

## 12.3 ENSURING OPTIMAL STORAGE PERFORMANCE

### ALIGNMENT OF VM PARTITIONS AND VMFS TO STORAGE ARRAYS

Virtual machines store their data on virtual disks. As with physical disks, these virtual disks contain storage partitions and file systems, which are created by the VM's guest operating system. In order to make sure of 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 Partners 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 and FCoE LUNs. This task is automated during 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, and XP operating systems is 32256. This value does not align to a block size of 4096.

Virtual machines running a clean installation of Microsoft Windows 2008, 7, and Vista operating systems automatically have their starting partitions set to 1048576. By default this value 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, then it is highly probable that these images require partition alignment.

## 12.4 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 data center 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.

## 12.5 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 46).

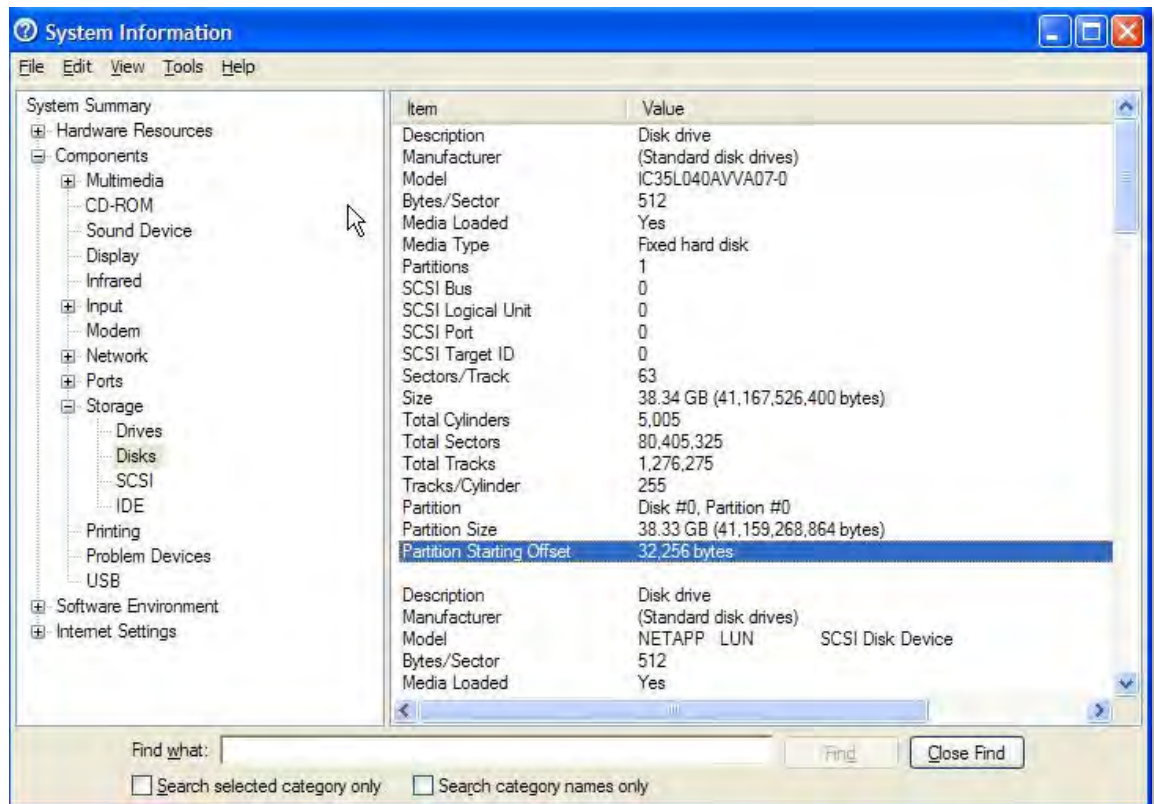


Figure 46) Using system information to identify the starting partition offset.

### NETAPP MBRTOOLS: IDENTIFICATION OF PARTITION ALIGNMENT 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 only requires 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 nondisruptive.

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

## 12.6 CORRECTIVE ACTIONS FOR VMS WITH MISALIGNED PARTITIONS

### 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 make sure that any newly created VM will be properly aligned and will not add to the added I/O load on the storage array.

### CORRECTING PARTITION MISALIGNMENT WITH NETAPP MBRTTOOLS

NetApp provides a tool named MBRAAlign, which runs on an ESX host and can correct and misaligned primary and secondary master boot record-based partitions. MBRAAlign requires the virtual machine that is undergoing the corrective action to be powered off.

MBRAAlign 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 thick vmdk. It is highly recommended to create a NetApp snapshot prior to executing MBRAAlign. Once a VM has been corrected, powered on, and the results verified, then this snapshot can be safely discarded.

MBRAAlign can be 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 that boot using the GRUB boot loader require the following steps after MBRAAlign has been run.

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

## 12.7 CREATE PROPERLY ALIGNED PARTITIONS 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 47.

1.	Boot the VM with the 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: OSTEWARD01-LXP

DISKPART> select disk 0
Disk 0 is now the selected disk.
DISKPART> create partition primary align=32_
```

Figure 47) Running diskpart to set a proper starting partition offset.



## CREATING A PROPERLY 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. Note that this procedure is not required for VMs running Windows Server 7, 2008, and Vista, as the file systems with these operating systems are aligned by default. To set up the starting offset using the fdisk command in the ESX service console, follow these steps.

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):  <pre>cd /vmfs/volumes/&lt;datastore&gt;/&lt;VM home dir&gt; ls -l</pre>
3.	Identify the number of cylinders in the virtual disk by reading the virtual disk descriptor file. Look for the line ddb.geometry.cylinders.  <pre>cat &lt;Virtual Disk&gt;.vmdk</pre>
4.	Run fdisk on the virtual disk file (the –flat.vmdk file) by typing the following command:  <pre>fdisk ./&lt;Virtual Disk&gt;.vmdk</pre>
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.

## 13 VIRTUAL MACHINE STORAGE LAYOUT

### 13.1 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 all 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, and so on.

From the standpoint of simplicity, this design works well where a VM home directory is a virtual machine. See Figure 48 for a high-level conceptual view of this layout.

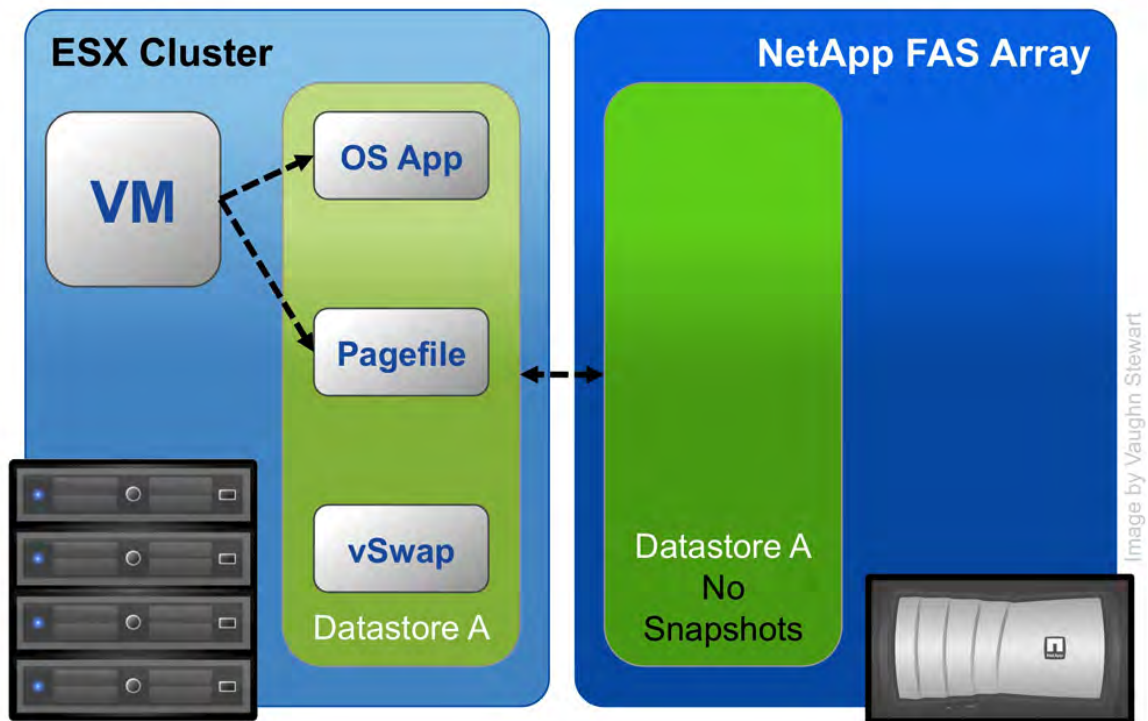


Figure 48) VMware's default virtual machine and vswap layout.

### 13.2 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 using 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 using 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.

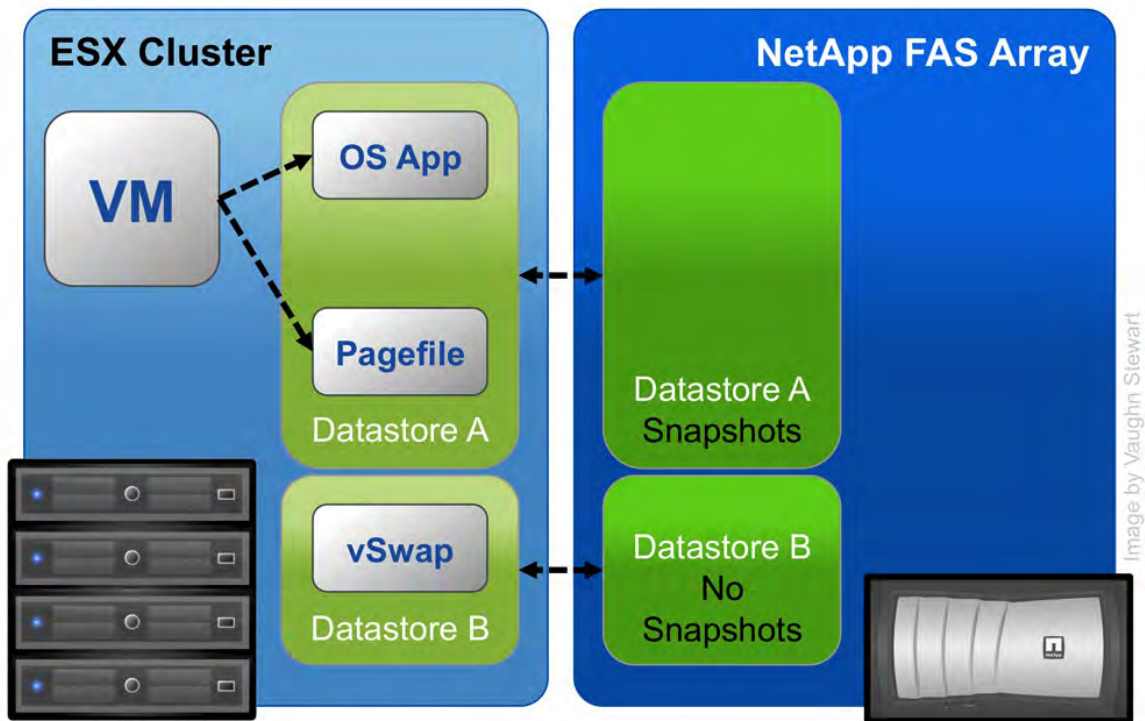


Figure 49) Recommended virtual machine layout option 1: a central vswap datastore for the entire cluster.

A prerequisite to making this change is the creation of a datastore to store the swap files. Because the VMware 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 50).

Steps	
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
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.

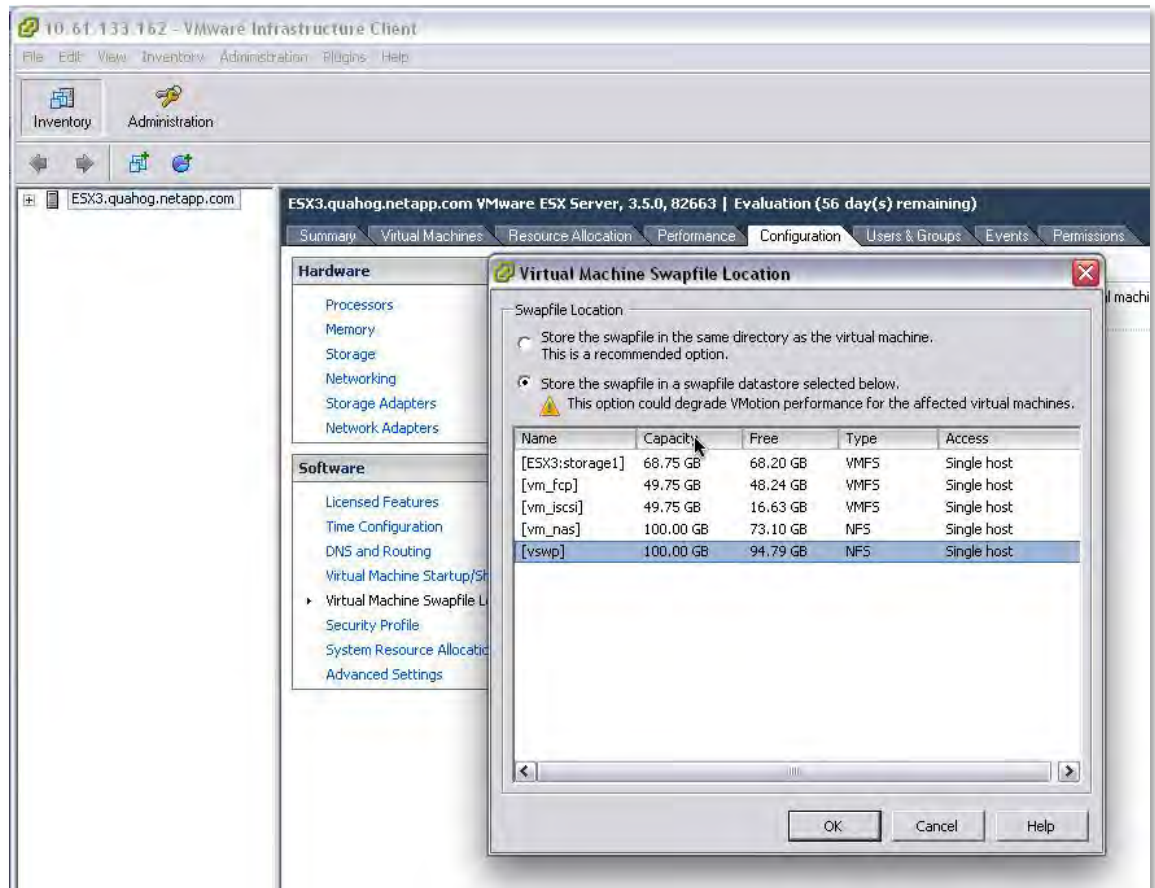


Figure 50) Configuring a global location for virtual swap files.

To configure a central datastore to store the virtual swap files for VMs that have been deployed, follow these steps (and refer to Figure 52).

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
7	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
8	Repeat steps 2 through 7 for each existing VM.

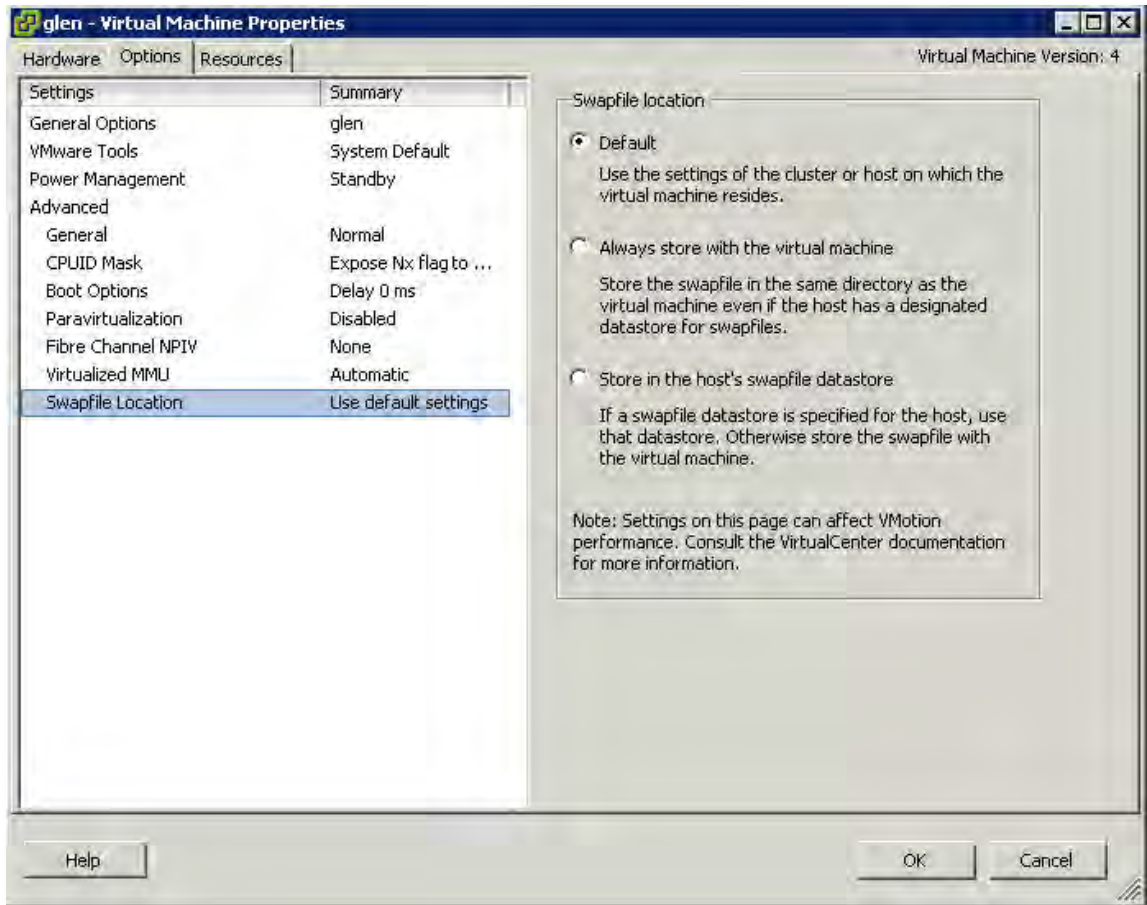


Figure 51) Defining the location for virtual swap files for a VM.

## RECOMMENDED LAYOUT OPTION 2: LOCATE VM SWAP/PAGEFILE ON 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.

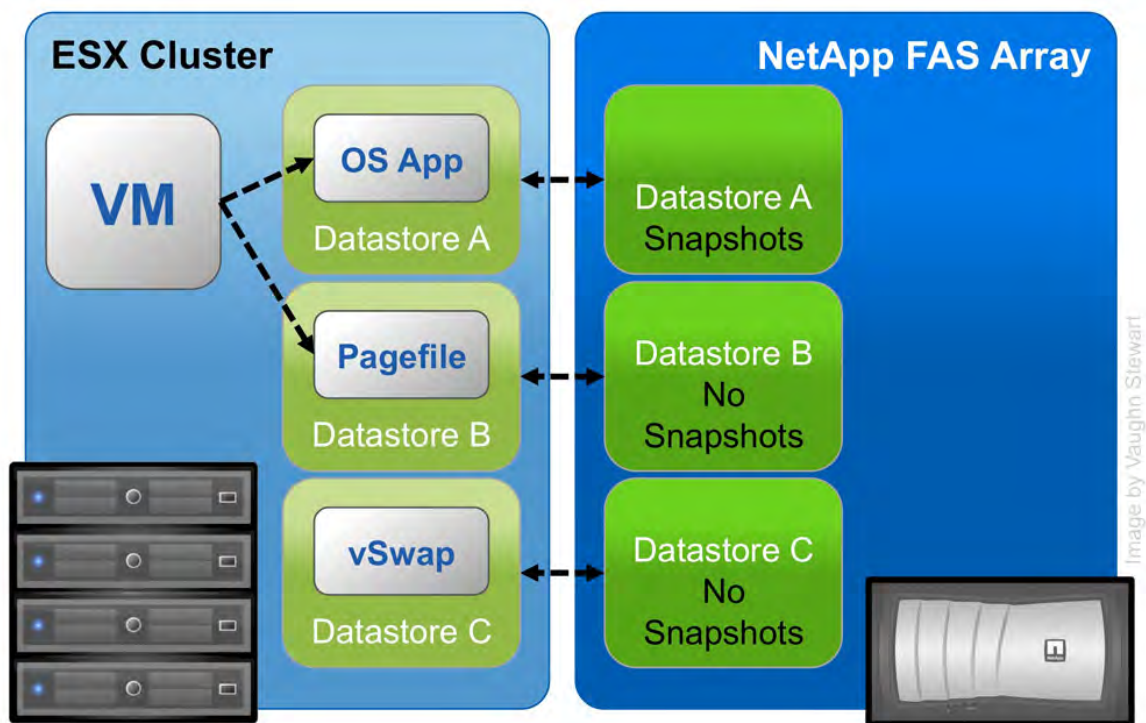


Figure 52) Recommended virtual machine layout option 2: VM pagefile has been separated to a pagefile datastore and a central vswap datastore for the entire cluster.

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

## 14 STORAGE MONITORING AND MANAGEMENT

### 14.1 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.

### 14.2 STORAGE GROWTH MANAGEMENT

#### GROWING VMFS DATASTORES

ESX/ESXi 4 makes it quite easy to increase the storage for a VMFS datastore by supporting the dynamic growth of the VMFS file system. Alternatively, adding a VMFS extent can grow a datastore. This second option results in a spanned VMFS datastore. The decision as to when to use which technology with NetApp storage arrays is simple. Grow the size of a LUN then grow VMFS. Add a new LUN then add an extent.

As NetApp storage arrays have array based queue limits, as opposed to the LUN based queue limits found with traditional legacy storage array architectures, the recommendation is to always grow the LUN and VMFS. Extents should only need to be used should a datastore require more than 2TBs of storage. 2TB is the maximum size of a LUN that can be accessed by ESX/ESXi.

To grow a VMFS file system please complete the following process.



1	Open FilerView ( <a href="http://filer/na_admin">http://filer/na_admin</a> ).
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	In the right pane, select the Configuration tab.
9	In the Hardware box, select the Storage Adapters link.
10	In the right pane, select the HBAs and then select the Rescan link. This step will result in the identification of the additional storage capacity by the ESX/ESXi host.
11	In the Hardware box, select the Storage link.
12	In the right pane, select the datastore to grow and then select Increase Datastore Capacity.
13	Select the LUN and click Next, then click Next again. As long as the window shows free space available on the LUN.
14	Make sure that the Maximize Space checkbox is selected, click Next, and then click Finish. See Figure 53.

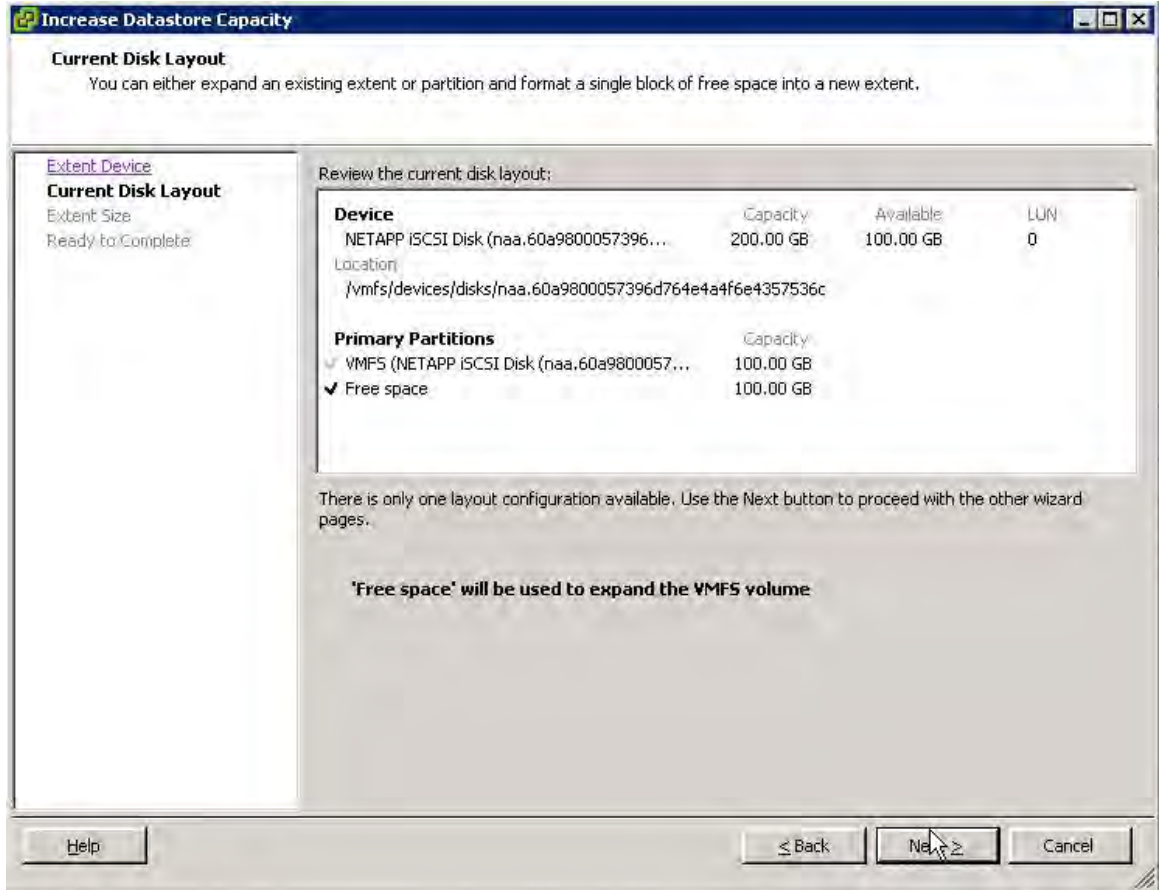


Figure 53) Expanding a VMFS partition.

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 54)
5	Start the VM.

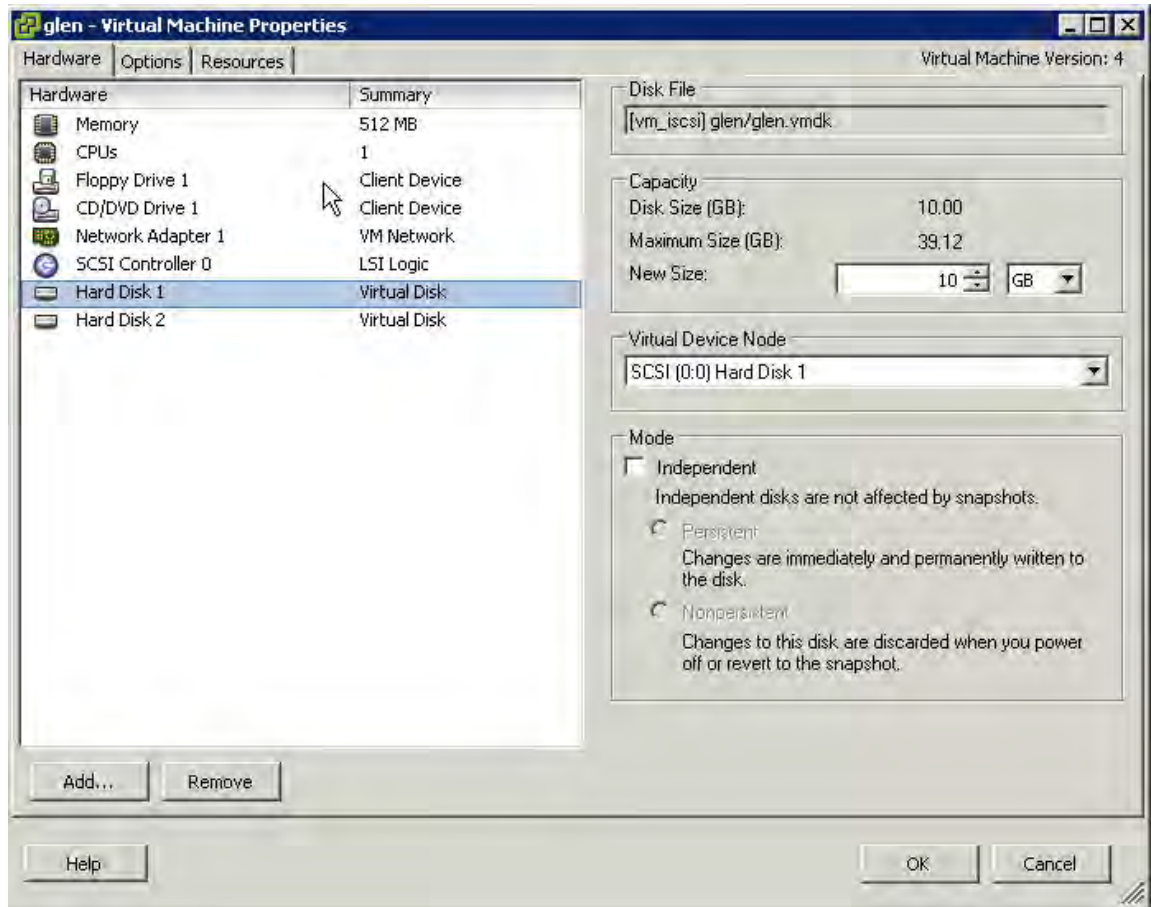


Figure 54) Increasing the size of a virtual disk.

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 55).
5	Open FilerView ( <a href="http://filer/na_admin">http://filer/na_admin</a> ).
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 56).
16	Select the LUN and click Next (see Figure 57).
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.

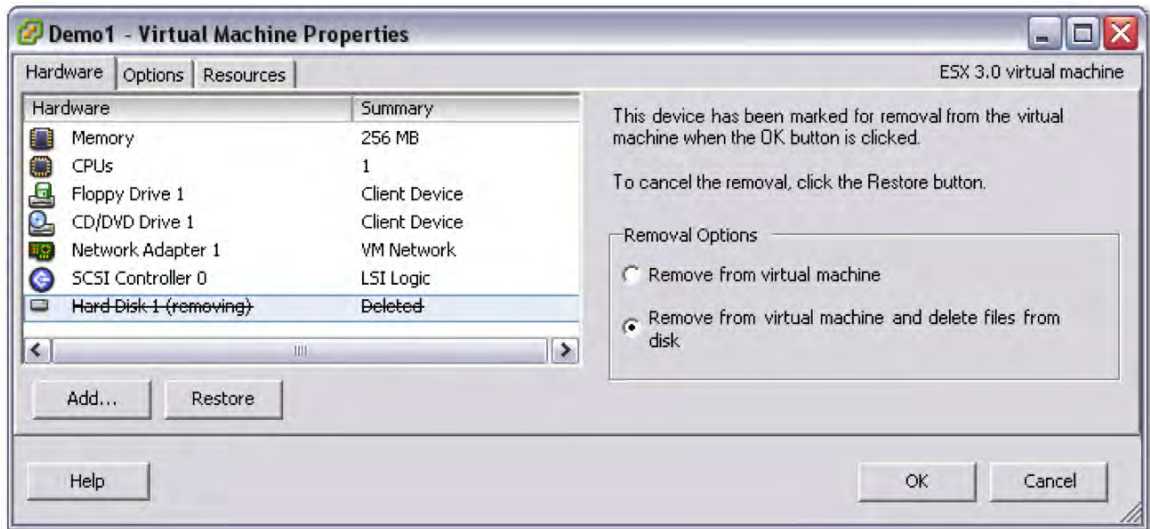


Figure 55) Deleting a VMDK from a VM.

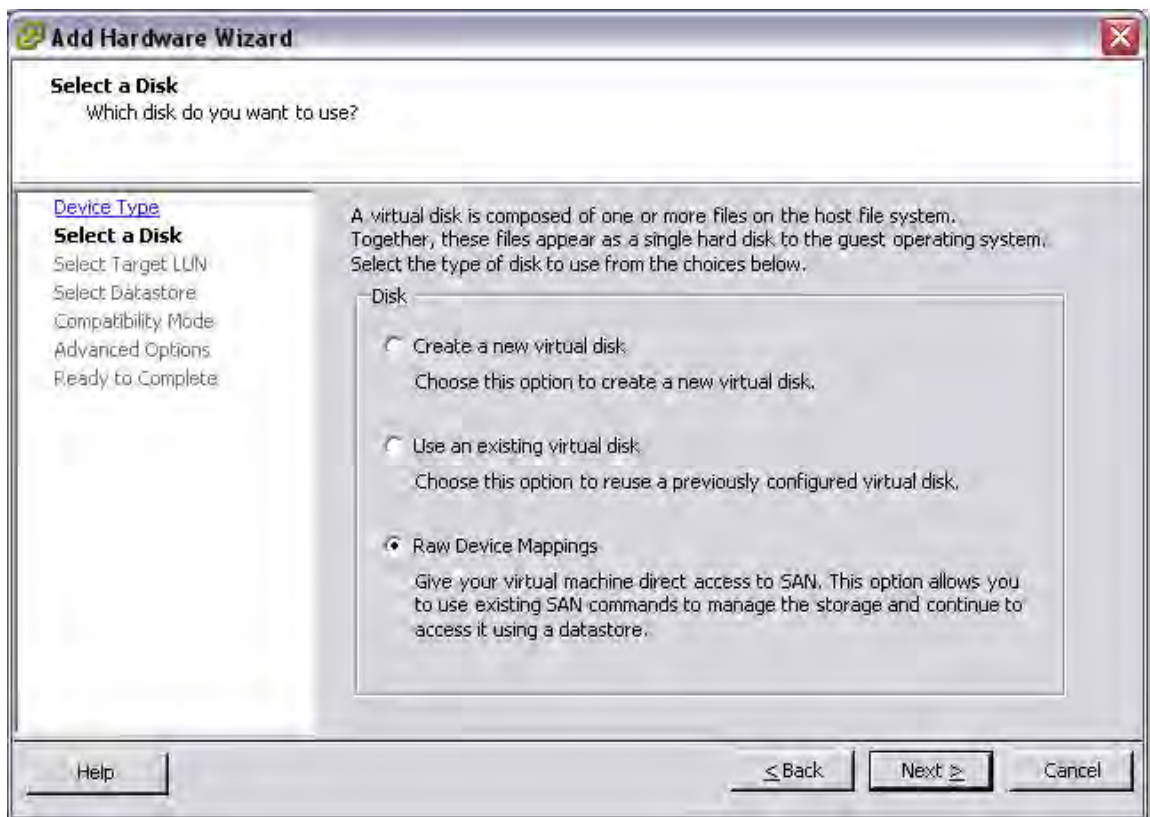


Figure 56) Connecting an RDM to a VM.

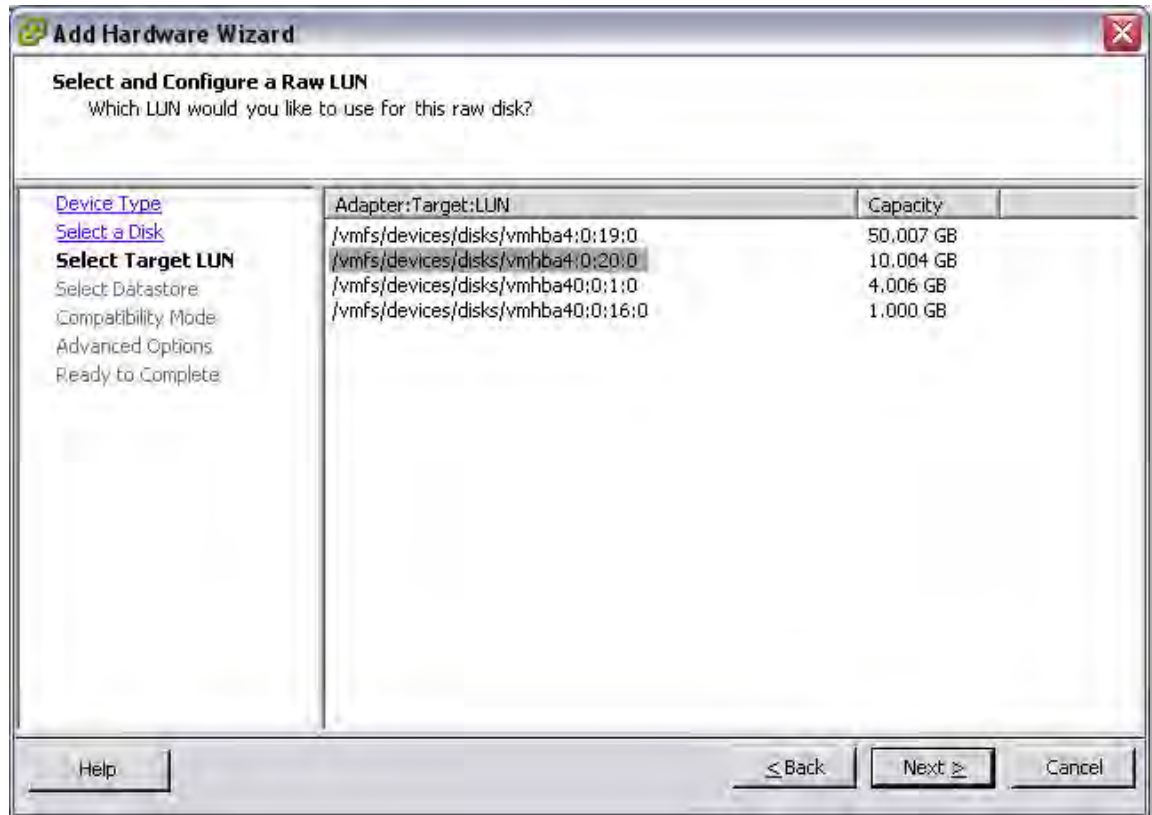


Figure 57) Selecting a LUN to mount as an RDM.

### 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	<p>For Windows VMs, you can use the diskpart utility to grow the file system. For more information, see <a href="http://support.microsoft.com/default.aspx?scid=kb:en-us:300415">http://support.microsoft.com/default.aspx?scid=kb:en-us:300415</a>.</p> <p>Or</p> <p>For Linux VMs, you can use ext2resize to grow the file system. For more information, see <a href="http://sourceforge.net/projects/ext2resize">http://sourceforge.net/projects/ext2resize</a>.</p>

### 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.

## 15 DISK-BASED SNAPSHOT BACKUPS FOR VMWARE

### 15.1 COMPLEMENTARY SNAPSHOT TECHNOLOGIES

VMware vSphere provides 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 copy 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 Snapshot copies provide simple point-in-time versions of VMs, allowing quick recovery. The benefits of VMware Snapshot copies are that they are easy to create and use, because they can be executed and scheduled from within vCenter Server. VMware suggests that the Snapshot technology in ESX should not be leveraged as a means to back up vSphere. For more information about native VMware Snapshot copies, 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. This is the only Snapshot technology that does not have a 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.

## 15.2 IMPLEMENTING NETAPP SNAPSHOT BACKUPS FOR VMWARE VSPHERE

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 the NetApp SnapManager portfolio by providing array-based backups which only consume block level changes to each VM, can provide multiple recovery points throughout the day, and as the backups are an integrated component within the storage array SMVI provides recovery times faster than any other means.

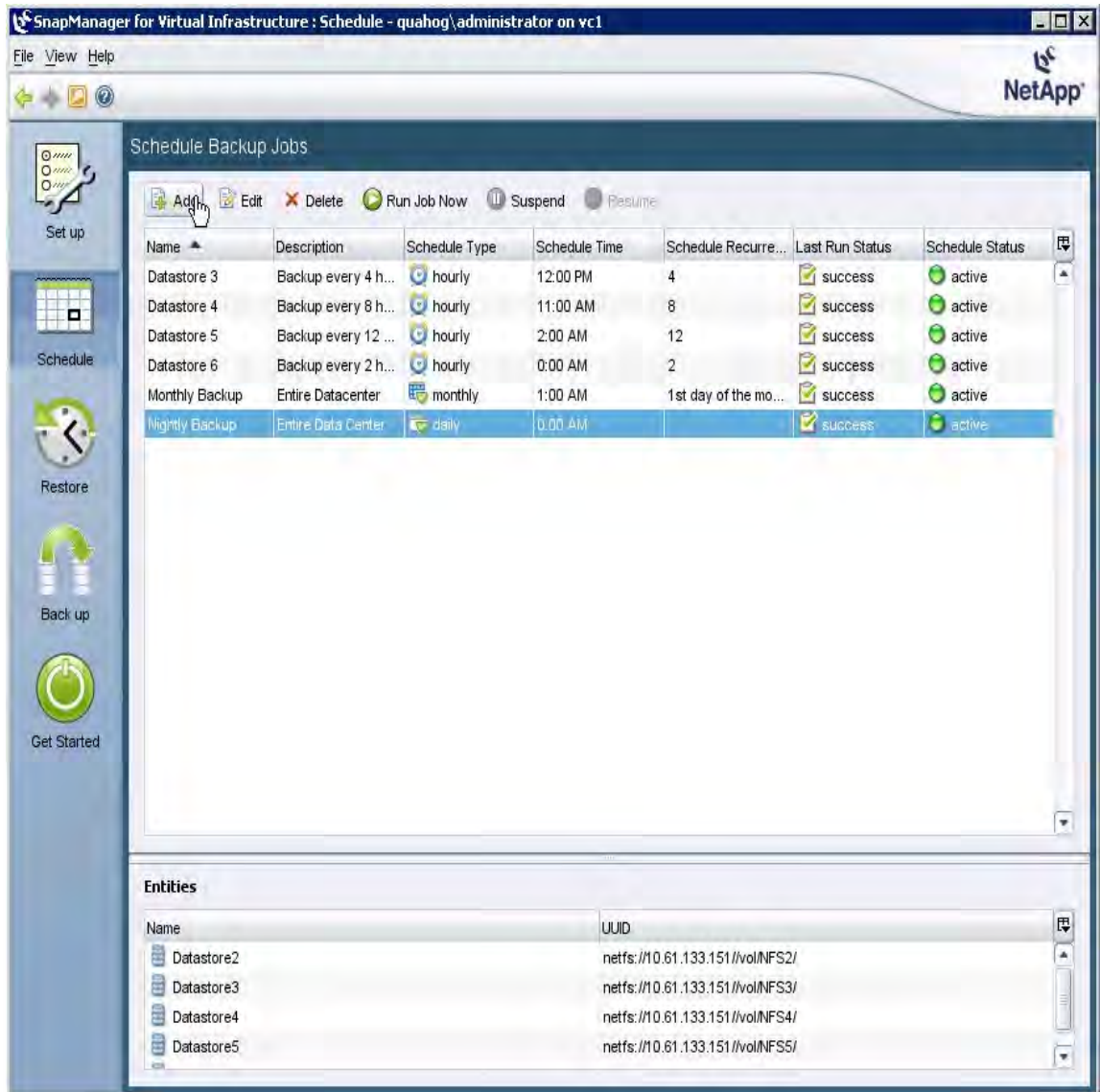


Figure 58) Defining backup policies at a datastore level in SnapManager for Virtual Infrastructure.

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



## 16 TECHNICAL REPORT SUMMARY

VMware vSphere 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 sending an email to [xdl-vgibutmevmtr@netapp.com](mailto:xdl-vgibutmevmtr@netapp.com) and please refer to TR-3749 in the subject line of your email.

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

### A.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 (using either SSH, Telnet, or Console connection).
2	Execute the following commands: <pre>secureadmin setup ssh options ssh.enable on options ssh2.enable on</pre>
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 <code>/etc/ssh/ssh_config</code> file and edit the Ciphers line to read as follows: <pre>Ciphers aes128-cbc, aes256-cbc, 3des-cbc.</pre>
5	Generate a DSA host key. On a Linux or VMware ESX Server, use the following command: <pre>ssh-keygen -t dsa -b 1024.</pre> When prompted for the passphrase, do not enter one; instead, press Enter. The public key is saved to <code>/root/.ssh/id_dsa.pub</code> .
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 <code>/etc/sshd/root/.ssh/authorized_keys</code> file, removing all information except for the key string preceded by the string <code>ssh-dsa</code> and a comment line. See the following example.
8	Test the connectivity from the remote host by issuing the <code>version</code> command on the FAS system. It should not prompt for a password: <pre>ssh &lt;netapp&gt; version NetApp Release 7.2: Mon Jul 31 15:51:19 PDT 2006</pre>

#### Example of the key for the remote host:

```
ssh-dsa AAAAB3NzaC1kc3MAAABhALVbwVyhtAVoaZukcjSTlRb/REO1/ywbQECTAcHijzdzhEJU
z9Qh96HVEwyZDdah+PTxfyitJCerb+1FAnO65v4WMq6jxPVYto6l5Ib5zxfq2I/hhT/6KPziS3LT
ZjKccwAAABUAjkLMwkpiPmg8Unv4fjCsYyhrSL0AAABgF9NsuZxniOOHr8tmW5RMX+M6VaH/nlJ
UzVXbLiI8+pyCXALQ29Y31uV3SzwTdlVOgjJHgv0GBw8N+rvGSBlr60VqggGjSB+ZXA01Eecbnj
vLnUt f0TVQ75D9auagjOAAAAYEJPx8wi9/CaS3dfKJR/tYy7Ja+Mr1D/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.

ESX SYSTEM SSH CONFIGURATION	
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.
```

## APPENDIX B: RELOCATING THE PAGEFILE IN WINDOWS VMS

Following is a registry file example of a simple registry script that sets the pagefile 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 `pagefile.reg`. The Setup Manager has a section where you can add `pagefile.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,00,70,00,61,00,67,00,65,00,66,00,69,00,6c,\

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

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

**End** -----

## APPENDIX C: DOCUMENT REFERENCES

### VMWARE REFERENCES

VMware Introduction to VMware vSphere

[http://vmware.com/pdf/vsphere4/r40/vsp\\_40\\_intro\\_vs.pdf](http://vmware.com/pdf/vsphere4/r40/vsp_40_intro_vs.pdf)

ESXi Server Configuration Guide

[http://vmware.com/pdf/vsphere4/r40/vsp\\_40\\_esxi\\_server\\_config.pdf](http://vmware.com/pdf/vsphere4/r40/vsp_40_esxi_server_config.pdf)

Basic System Administration Guide

[http://vmware.com/pdf/vsphere4/r40/vsp\\_40\\_admin\\_guide.pdf](http://vmware.com/pdf/vsphere4/r40/vsp_40_admin_guide.pdf)

VMware Fibre Channel SAN configuration Guide

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

iSCSI SAN Configuration Guide

[http://vmware.com/pdf/vsphere4/r40/vsp\\_40\\_iscsi\\_san\\_cfg.pdf](http://vmware.com/pdf/vsphere4/r40/vsp_40_iscsi_san_cfg.pdf)

vSphere Upgrade Guide

[http://vmware.com/pdf/vsphere4/r40/vsp\\_40\\_upgrade\\_guide.pdf](http://vmware.com/pdf/vsphere4/r40/vsp_40_upgrade_guide.pdf)

All VMware documents are located at:

[http://www.vmware.com/support/pubs/vs\\_pubs.html](http://www.vmware.com/support/pubs/vs_pubs.html).

## 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-3428: NetApp and VMware ESX Server 3.5

<http://www.netapp.com/library/tr/3428.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](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/>

Data 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](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](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: Data 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](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](http://en.wikipedia.org/wiki/Redundant_array_of_independent_disks)

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\\_EM\\_Celerra\\_IP.pdf](http://www.vmware.com/files/pdf/VMware_VI3_and_EM_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](http://www.vmware.com/pdf/clariion_wp_eng.pdf)

Vizioncore: vOptimizer Pro FAQ

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

## APPENDIX D: VERSION TRACKING

Version 1.0	May 2009
	Original document

© Copyright 2009 NetApp, Inc. All rights reserved. No portions of this document may be reproduced without prior written consent of NetApp, Inc. NetApp, the NetApp logo, Go further, faster, DataFabric, Data ONTAP, FilerView, FlexClone, FlexVol, NOW, RAID-DP, SANscreen, SnapDrive, SnapManager, SnapMirror, Snapshot, SnapVault, and WAFL are trademarks or registered trademarks of NetApp, Inc. in the United States and/or other countries. Microsoft and Windows are registered trademarks of Microsoft Corporation. Linux is a registered trademark of Linus Torvalds. VMware and VMotion are registered trademarks of VMware, Inc. UNIX is a registered trademark of The Open Group. All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such.