

# Best Practices for VMware vSphere With NexentaStor 4.0.x

## At a Glance

**Overview**

**Architectural Design Recommendations**

**NFS Storage Configuration**

**iSCSI Storage Configuration**

## Table of Contents

<b>Preface .....</b>	<b>3</b>
<b>About this document .....</b>	<b>3</b>
<b>Audience .....</b>	<b>3</b>
<b>How to contact Nexenta.....</b>	<b>3</b>
<b>External References .....</b>	<b>4</b>
<b>Document History.....</b>	<b>4</b>
<b>Versions Tested .....</b>	<b>4</b>
<b>Overview .....</b>	<b>5</b>
<b>NexentaStor Advantages.....</b>	<b>5</b>
<b>Document Highlights.....</b>	<b>5</b>
<b>Prerequisites.....</b>	<b>6</b>
<b>Additional Information.....</b>	<b>6</b>
<b>Architectural Design Recommendations .....</b>	<b>7</b>
<b>Storage Recommendations .....</b>	<b>7</b>
<b>Network Recommendations.....</b>	<b>8</b>
<b>NFS Storage Configuration .....</b>	<b>9</b>
<b>About NFS Storage Configuration .....</b>	<b>9</b>
<b>NFS Configuration in NexentaStor .....</b>	<b>9</b>
<b>Creating a Volume &amp; NFS Share in NexentaStor .....</b>	<b>10</b>
<b>NFS Protocol Settings within vSphere .....</b>	<b>14</b>
<b>Adding NFS Storage to the vSphere Server using traditional client.....</b>	<b>17</b>
<b>iSCSI Storage Configuration.....</b>	<b>19</b>
<b>About iSCSI Configuration .....</b>	<b>19</b>
<b>Adding Remote Initiators to NexentaStor .....</b>	<b>19</b>
<b>Create an Initiator Group .....</b>	<b>20</b>
<b>Create a Target Group.....</b>	<b>20</b>
<b>Creating a ZVOL .....</b>	<b>21</b>
<b>Configuring iSCSI in vSphere Traditional Client.....</b>	<b>23</b>
<b>Adding iSCSI Servers to ESX Hosts.....</b>	<b>26</b>
<b>Adding iSCSI Storage to the vSphere Server using traditional client .....</b>	<b>27</b>
<b>Using VAAI for Block Storage.....</b>	<b>30</b>
<b>Validate VAAI is Enabled .....</b>	<b>35</b>
<b>Conclusion .....</b>	<b>36</b>

## PREFACE

### About this document

This best practices guide describes how to integrate VMware vSphere 5.x or later with NexentaStor 4.0.x, utilizing NexentaStor as a backend storage for virtual machine storage.

### Audience

This document is intended for system administrators that want to integrate NexentaStor with an existing VMware vSphere installation.

It is expected that the reader is experienced with VMware vSphere and has basic knowledge of NexentaStor.

### How to contact Nexenta

For general questions or sales inquiries please use [sales@nexenta.com](mailto:sales@nexenta.com), or the contact forms at <http://nexenta.com/contact-us>.

For support and services inquiries, please use [support@nexenta.com](mailto:support@nexenta.com).

## External References

For additional information, please see the following documents:

Reference	Location
Nexenta website	<a href="http://www.nexenta.com">www.nexenta.com</a>
VMware website	<a href="http://www.vmware.com">www.vmware.com</a>
NexentaStor Download and Documentation	<a href="http://nexenta.com/products/downloads/nexentastor">http://nexenta.com/products/downloads/nexentastor</a>
VMware vSphere Downloads	<a href="https://www.vmware.com/go/download-vsphere">https://www.vmware.com/go/download-vsphere</a>

## Document History

Version	Date	Changes
1.0	December, 2015	Initial version
1.1	December, 2015	Updated Versions Tested

## Versions Tested

NexentaStor	VMware ESX
4.0.4	5.5, 6.0
4.0.4-FP1	5.5, 6.0
4.0.4-FP2a	5.5, 6.0

## OVERVIEW

This guide lists recommendations and guidelines for deploying and configuring NexentaStor storage in your high availability VMware vSphere environment.

Nexenta Systems provides enterprise class storage within the complexities of the enterprise environment. Business requirements constantly challenge IT management to provide a level of high availability, while simultaneously reducing costs. To this end, storage is a commodity, one where costs must decrease without negatively impacting availability.

Server virtualization is the de-facto standard for reducing server hardware costs. VMware is the leader in server virtualization with their vSphere product suite. As VMware is recognized as the leader in server virtualization, Nexenta is recognized as the leader in open software defined storage.

NexentaStor is a cost effective storage solution for VMware vSphere environments. It provides storage administrators with enhanced flexibility so they can meet their customer requirements. The basic NexentaStor product includes iSCSI, NFS and SMB support. You can enable Fibre Channel (FC) support through an optional licensed plug-in.

### NexentaStor Advantages

NexentaStor improves the vSphere environment by providing:

- Virtually Unlimited snapshots
- Multiple levels of RAID support (Single (RAID5), Double (RAID6), Triple Parity, 2-way mirrors (RAID10))
- Integrated replication
- User-friendly Web GUI
- Reporting

NexentaStor enhanced data protection features include:

- Data integrity
- Thin provisioning
- Fault management

Optional NexentaStor plugins support:

- High Availability
- FC Target

### Document Highlights

- Best practices and recommendations for deploying VMware vSphere 5.x with a NexentaStor 4.x Array
- Tuning options and deployment methods for NFS and iSCSI protocols in production environments
- VMware cluster recommendations for high availability and load balancing

- Utilizing vSphere API Array Integration with NexentaStor

## Prerequisites

This section lists the prerequisites that must be set up before you can implement the suggested solutions in this documentation.

- **VMware vSphere Infrastructure** — Set up a VMware vSphere Hypervisor and configure all appropriate management software to include VMware vCenter Server as Windows or Linux Based Install.
- **NexentaStor Array** — Set up a Highly Available NexentaStor Array with at least 2 available 10GbE interface per controller
- **Network** – 10GbE Network Infrastructure capable of supporting Link Aggregation (LACP)

## Additional Information

- VMware Documentation
  - <https://www.vmware.com/support/vsphere/>
- NexentaStor Installation Guide
  - <https://nexenta.com/products/downloads/nexentastor>

## ARCHITECTURAL DESIGN RECOMMENDATIONS

This chapter lists recommendations for using storage and network hardware with NexentaStor.

### Storage Recommendations

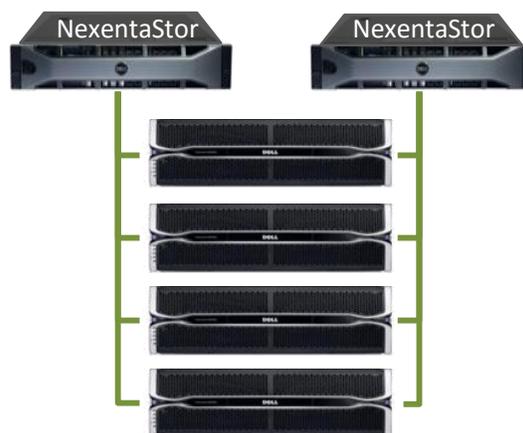
Virtualized environments often mask the specific storage needs of an application, however this is not the case for some specific workloads. In light of this standardization Nexenta recommends the following minimum configurations for VMware environments. Nexenta maintains an extensive list of **Reference Architectures** that utilize industry standard x86 servers and JBODs. The current list can be found at:

<http://info.nexenta.com/rs/059-GUR-914/images/NexentaStor-HCL.pdf>

The following combination and sizing is critical to meet the demands of data reduction operations as well as overall performance in large deployments of a virtual infrastructure. The minimum recommended configuration is a **Reference Architecture** build with at least:

- Dual Controllers (HA Configuration)
- At least two 2.4GHz or faster Intel® Xeon CPUs per NexentaStor head
- At least two 2.4GHz or faster Intel® Xeon CPUs per NexentaStor head
- At least 128GB of DRAM memory per head with a maximum of 256GB
- A mirrored disk pool of (at least) Twenty 10K or 15K RPM SAS Hard Drives or Forty 7.2K RPM SAS disk drives
  - Include an additional 2 drives for hot spares
- At least two write-intensive SSD for Log devices in a mirrored configuration
- Two 10GbE Network interfaces
- Latest NexentaStor version
- Use the Performance based (mirrored) Profile

Figure 1: Typical HA Configuration with Multiple JBODs



The high random I/O patterns of virtual infrastructures demand high storage performance and is best satisfied with a mirrored storage profile. This performance based dataset will produce high availability as well as fast access distributing the data load across multiple redundancy groups.

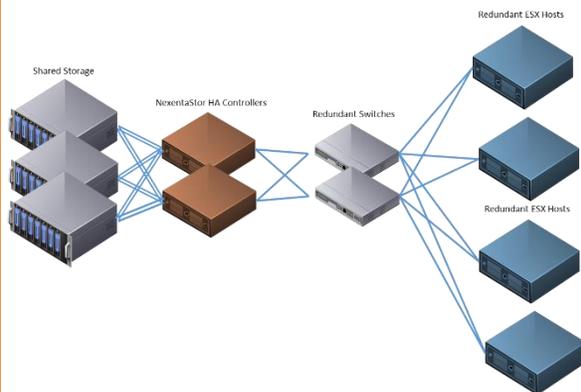
For the highest level of resiliency it is advised to spread the pool across multiple JBODs to eliminate single points of failure.

## Network Recommendations

The following recommendations will ensure that NFS and iSCSI traffic will be designed to achieve high availability and present no single point of failure

- 10GbE Physical Switches
- Separate the backend storage NFS and iSCSI network from any client traffic. This can be done using VLANs, network segmentation, or dedicated switches. This will allow for maximum performance and prevent unwanted access to your storage environment.
- Do not over-subscribe the network connection between the LAN switch and the storage array.
- Retransmitting of dropped packets can affect the performance of systems that are using VMDKs (Virtual Machine Disks).
- Check the current physical port settings on the used switches, and if STP is deployed make sure the ports are configured as PortFast. This enables the switch to set the forwarding state immediately, instead of going through the listening, learning, and forwarding states.
- Use LACP for port aggregation and redundancy. This provides better overall network performance and link redundancy.

**Figure 2: High Availability Network Configuration**



## NFS STORAGE CONFIGURATION

*This chapter includes the following topics:*

- About NFS Storage Configuration
- NFS Configuration in NexentaStor
- Creating a Volume & NFS Share in NexentaStor
- NFS Protocol Settings within vSphere
- Adding NFS Storage to the vSphere Server using traditional client

### About NFS Storage Configuration

This chapter describes adding NFS storage from a NexentaStor Array to the VMware vSphere environment. It also contains configuration settings for using NFS storage with NexentaStor.

### NFS Configuration in NexentaStor

#### Configuring the NFS Server version

Up until version 6.0 VMware only supported NFS version 3. To ensure interoperability with all versions we recommend that you reduce the maximum Client and Server version

- Click **Data Management > Shares**
- In the NFS Server pane, click **Configure**
  - Select **3** for the Highest Server Version
  - Select **3** for the Highest Client Version

#### NexentaStor 4 NFSD Default Setting

We have determined that the default "**Concurrent NFSD Servers**" and "**Concurrent LOCKD Servers**" settings for NexentaStor versions 4.0.1 through 4.0.3-FP4 may be too small for typical NFS workloads. These settings limit the maximum number of NFS server and lock daemon threads that may be concurrently executing. The defaults for both of these settings have been increased to **256** in NexentaStor 4.0.4.

We strongly recommend that all NFS users still using the default **Concurrent NFSD Servers** and **Concurrent LOCKD Servers** settings in NexentaStor versions 4.0.1 through 4.0.3-FP4 increase the settings immediately. Increasing these settings does not require a reboot, but may result in a small delay or interruption to NFS clients. The interruption will be 5 seconds or the duration of the client timeout, whichever is greater. We recommend that any system configurations changes follow change control procedures and be performed in a maintenance window.

To configure the settings, use NMV:

- Click **Data Management > Shares**
- In the NFS server panel, click **Configure**
- Review the **Concurrent NFSD Servers** setting; if it is 16, change the value to **256**.
- Review the **Concurrent LOCKD Servers** setting; if it is 16, change the value to **256**.

- Click **Save**, the changes will take effect immediately

## Create Network Aggregate

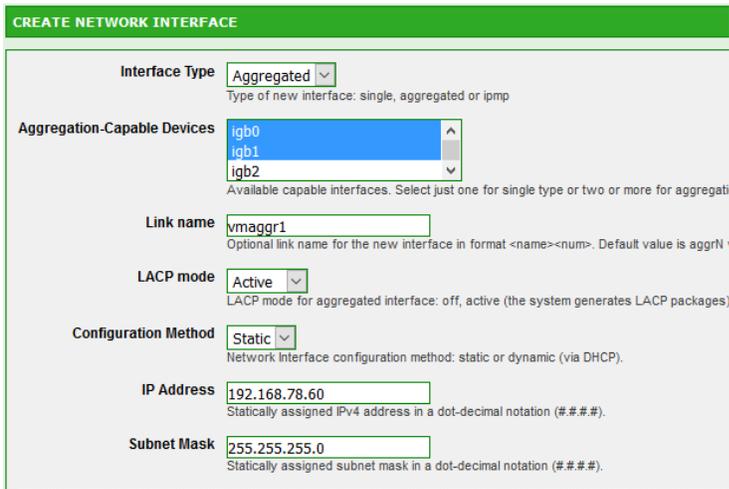
By creating a network aggregate the NexentaStor array will be able to provide better overall network performance and link redundancy.

Click **Settings > Network**

In the **Network** pane, click **Create**

This will bring up the Create Network Interface

Figure 3: Create Network Interface



**CREATE NETWORK INTERFACE**

Interface Type:  (Type of new interface: single, aggregated or ipmp)

Aggregation-Capable Devices:  (Available capable interfaces. Select just one for single type or two or more for aggregated)

Link name:  (Optional link name for the new interface in format <name><num>. Default value is aggrN)

LACP mode:  (LACP mode for aggregated interface: off, active (the system generates LACP packages))

Configuration Method:  (Network Interface configuration method: static or dynamic (via DHCP))

IP Address:  (Statically assigned IPv4 address in a dot-decimal notation (###.###.###.###))

Subnet Mask:  (Statically assigned subnet mask in a dot-decimal notation (###.###.###.###))

- **Interface Type:** Aggregated
  - **Aggregation Capable Devices:** Select the two network interfaces you wish to aggregate together. Use the CTRL key to multi-select
  - **Link Name:** This is optional, if you want to provide a name it must be in the format of <name><num> (vmaggr1)
  - **LACP Mode:** Active (Make sure the switch is set to active mode)
  - **Configuration Method:** Static or DHCP
  - If **Static** is selected enter **IP Address** and **Subnet Mask**
- Click **Add Interface**

This process needs to be performed on both controllers within the HA configuration

## Creating a Volume & NFS Share in NexentaStor

A Volume is a collection of Redundancy Groups that data is striped across. When creating a Volume you will want to create a number of Redundancy Groups (mirrors, single, double and triple parity) and then create the Volume across those. The following steps outline the process to create Volumes. For VMware deployments you will want to use the Performance Based (Mirrors) Profile.

## Create an volume

- Click **Data Management > Datasets**.
- In the Volumes pane, click **Create**.
- This will bring up the dataset creation wizard

Figure 4: Dataset creation wizard

CREATE NEW VOLUME

Name:

Failure domain:

Profile:  Performance  
 Balanced  
 Capacity

Data disks:   30/42

Spare:  2/6

Cache disks:   4/4

Log disks:   2/2



■ Data 13.6TB ■ Parity 13.6TB  
 ■ Spare 1.82TB ■ Log 7.45GB  
 ■ Cache 373GB

- **Name** : Provide a name for your volume
  - **Failure Domain** : If the NexentaStor Array has more than one JBOD the volume creation wizard will try to split the volume across those JBODs evenly. By distributing the volume across the redundant JBODs NexentaStor provides more storage resource and improved performance, load balancing and overall better throughput.
  - **Profile** : Performance
  - **Data Disks** : Select disk capacity and use the slider to indicate how many of the available devices to use
  - **Spare** : Use the slider to select how many drives to use for Spares. This can be any number up to 6. At least two drives should be selected.
  - **Cache Disks** : Use the slider to select how many drives to use for Cache. Based on the minimum requirements Nexenta recommends not using Cache drives for virtualized environments
  - **Log Disks** : Use the slider to select how many drives to use for Log. This needs to be at least two drives. NexentaStor creates a mirror of these drives to ensure performance even in the event of a log device failure.
- Click **Create**

## Create a share

- Click **Data Management > Shares**
- In the Folder pane, click **Create**
  - Select a **Volume** from the dropdown
- **Recordsize:** For **high I/O** environments Nexenta recommends a setting of **32K**. For **high throughput** environments Nexenta recommends this be set to **128K**
  - Set **Compression** to **LZ4**. LZ4 is considered a very high performance compression algorithm with very little overhead on the system
  - Deduplication: **OFF**

Figure 5: Folder creation screen

CREATE NEW FOLDER

! For advanced options, please consider using Nexenta Management Console (NMC). With NMC you will be able to t

**Volume**  Folder's volume.

**Folder Name**  Each folder pathname's component delimited by backslash (\) can only contain alphanumeric characters (a-z, A-Z). Folder pathname must begin with an alphanumeric character and not contain a percent sign (%).

**Description**  Human-readable description for this folder.

**Record Size**  Specifies a suggested block size for files in the folder. Default is 128K.

**Log Bias**  Provide a hint to ZFS about handling of synchronous requests in this dataset. If logbias is set to latency (the default) synchronous operations for global pool throughput and efficient use of resources.

**Deduplication**  Controls the deduplication option for this dataset. If enabled, it will optimize use of duplicate copies of data. Default

**Compression**  Controls the compression algorithm used for this dataset. Default is "on". Setting compression to "on" uses the lzjb

**Number of Copies**  Controls the number of copies of data stored for this dataset. Default is "1".

**Case Sensitivity**  Indicates whether the file name matching algorithm used by the file system should be case-sensitive, case-insens

**Unicode Only**  Enable it if you want to exclude non-Unicode file names creation for this folder. If set, this option will ensure better

**Sync**  Controls synchronous requests (standard - ensure all synchronous requests are written to stable storage; alway

- Accept the defaults or enter appropriate values for the remaining options.
- Click **Create**

## Share the folder via NFS

- Click **Data Management > Shares**
  - Click the **NFS** checkbox
  - Confirm the NFS enablement
  - Set "**nbmand**" property to "**on**"

## Editing Folder Options

- Click **Data Management > Shares**

- Click the Name of the folder you wish to edit
  - This will bring up the Edit Folder Screen

Figure 6: Folder Properties

EDIT FOLDER: VMWARE_VOL/VMWARE_NFS_SHARE	
Read-Only Parameters:	
Name	Value
name	vmware_vol/vmware_nfs_share
creation	Mon Nov 2 10:31 2015
used	31K
available	17.8T
referenced	31K
compressratio	1.00x
mountpoint	/volumes/vmware_vol/vmware_nfs_share
casesensitivity	mixed

**Note:** The mountpoint property will be used when setting up the NFS share on the ESX hosts

- You can modify such options as Record Size, Compression, Quota or Read-Only properties

## Editing NFS Option

By default, anyone can access the newly created NFS share. To prevent unwanted access the sharing properties need to be edited.

- In the NFS column, click **Edit**
- Deselect **Anonymous** and **Anonymous Read-Write**
- Grant **Read-Write** access to ESX servers or Networks (e.g. @192.168.78.40, @192.168.78.0/24)
- Grant **Root** privileges to ESX servers or Networks (e.g. @192.168.78.40, @192.168.78.0/24)

For additional information on all of these options please refer to the User Guide located at <https://nexenta.com/products/downloads/nexentastor>

Figure 7: Editing NFS Access Rights

SUMMARY INFORMATION: FOLDERS						
Folder	Auto Service	Refer	Used	Avail	CIFS	NFS
<input type="checkbox"/> vmware_vol/vmware_nfs_share	-	31.00 KB	31.00 KB	17.80 TB	<input type="checkbox"/>	<input checked="" type="checkbox"/> Edit

Figure 8: Disabling Anonymous Access

**Auth Type**  Use AUTH\_SYS authentication. The user's UNIX user-id and group-ids are passed in the clear on the network, unauthenticated by the NFS server.

Use null authentication (AUTH\_NONE). NFS clients using AUTH\_NONE have no identity and are mapped to the anonymous user 'nobody' by NFS server.

Use a Diffie-Hellman public key system (AUTH\_DES)

Use Kerberos V5 protocol to authenticate users before granting access to the shared filesystem.

Use Kerberos V5 authentication with integrity checking (checksums) to verify that the data has not been tampered with.

Use Kerberos V5 authentication, integrity checksums, and privacy protection (encryption) on the shared filesystem.

Sharing uses one of the specified security modes. Current: sys.

**Anonymous**   
Grant access to unknown users and recognize them as anonymous user 'nobody'. Set this option to 'false' to deny anonymous (unknown user) access.

**Anonymous Read-Write**   
Allow anonymous access to this share. Shared top-level directory will be granted read-write access for anonymous user 'nobody'. If recursive share mode is set to true, to future sub-folders will not be allowed until explicitly requested. The anonymous user name is 'nobody'.

**Read-Write**   
Specify Read-Write access list. Overrides the 'Read-Only' field for the clients specified. The list contains zero or more colon-delimited netgroups, hosts and/or networks/s. Read-Write access for all.

**Read-Only**   
Specify Read-Only access list. Overrides the 'Read-Write' field for the clients specified. The list contains zero or more colon-delimited netgroups, hosts and/or networks/s.

**Root**   
Only root users of the hosts specified in this list have root access. For this option to work, appliance domain name MUST match client's domain name. By default, no host is allowed root access. Specify netgroups, hosts and/or networks/subnets, for instance: netgroup-engineering:@10.16.16.92:nas1.local.net:@192.168.0.1:@192.168.1.0/24.

## NFS Protocol Settings within vSphere

Configuring NFS protocol within vSphere is a critical component of a successful deployment with NexentaStor. To ensure high availability of your storage with a vSphere environment it is critical to follow the knowledge base articles and settings below.

### VMware NFS connectivity issues on NexentaStor with ESXi 5.x and 6.0

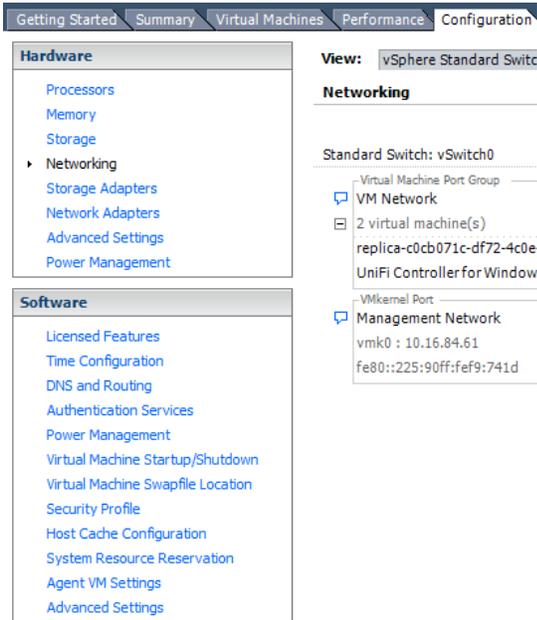
ESXi default setting may result in NFS connectivity issues and "all path down" (APD) events.

<http://kb.vmware.com/selfservice/microsites/search.do?cmd=displayKC&docType=kc&externalId=2133025>

## vSphere Advanced Settings

Table 1 below shows the Advanced Settings that should be modified prior to putting any production data on the array. To alter the parameters listed, go to the VMware vCenter 5.x server and select a VMware server. Select the **Configuration** tab and within the **Software** section click on **Advanced Settings**

Figure 9: Accessing ESX Advanced Settings



**Note: These steps need to be performed on all hosts that will be accessing the Nexenta storage. This change requires a reboot**

Figure 10: VMware NFS Advance Settings

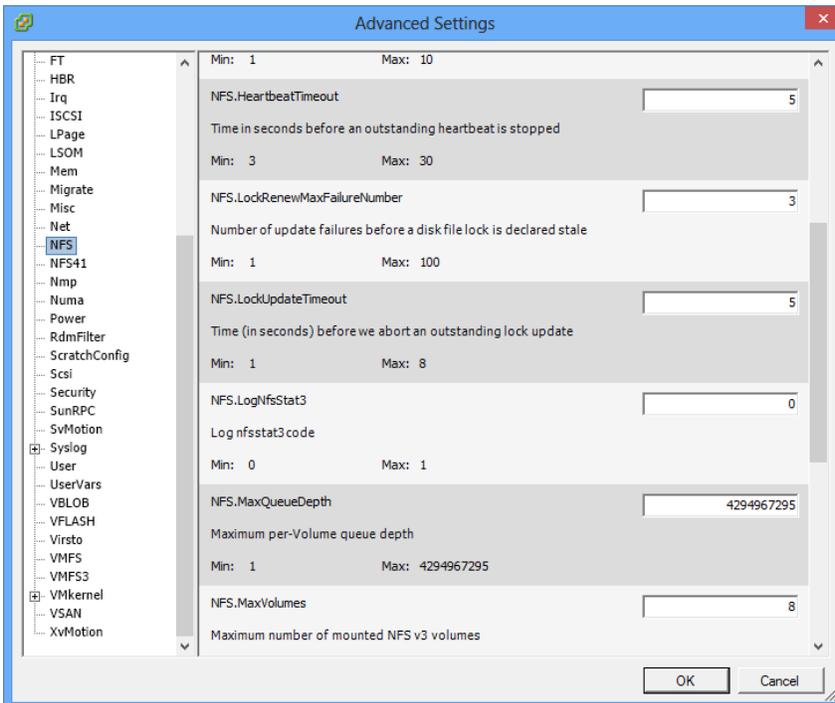


Table 1: Recommended NFS and TCP/IP Settings for vSphere

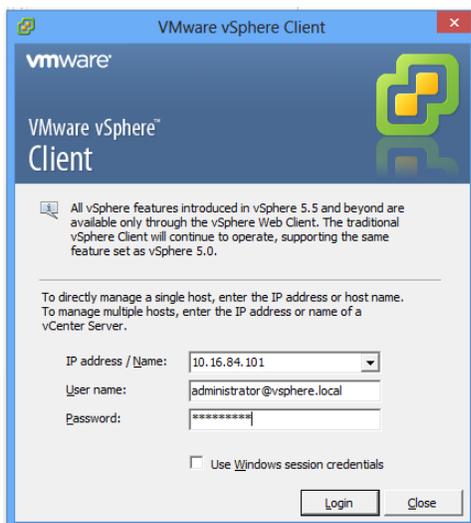
OPTION	VALUE
Nfs.heartbeatfrequency	20

Nfs.heartbeatdelta	12
Nfs.heartbeatmaxfailures	10
NFS.HeartbeatTimeout	5
Nfs.Sendbuffersize	264
Nfs.Receivebuffersize	256
Nfs.MaxVolumes	256
Net.TcpipHeapMax	128
Net.TcpipHeapsize	32

## Adding NFS Storage to the vSphere Server using traditional client

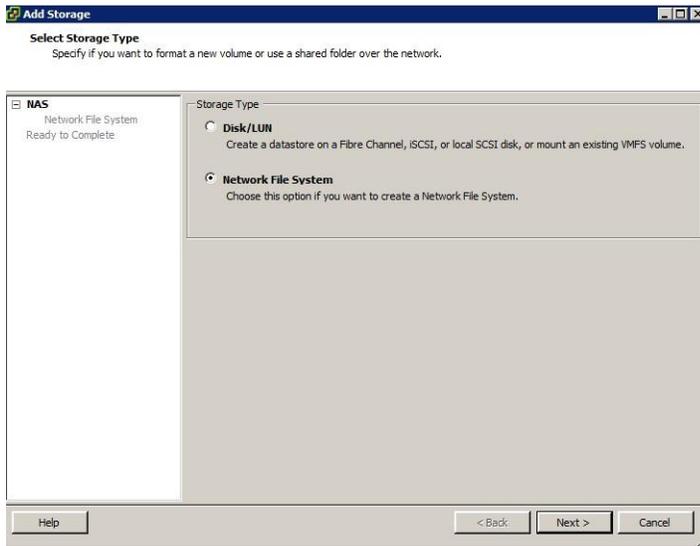
- Connect to the vCenter server through the vSphere Client.
  - Select the IP address for the Server

Figure 11: vSphere Client login



- Click **View > Inventory > Datastore and Datastore Clusters**
  - **Right-Click** on your **Datacenter** and select **Add Datastore**
    - Select the Host you want to have access to the NFS datastore
    - Click **Next**
    - Select **Network File System** as the Storage Type
    - Click **Next**

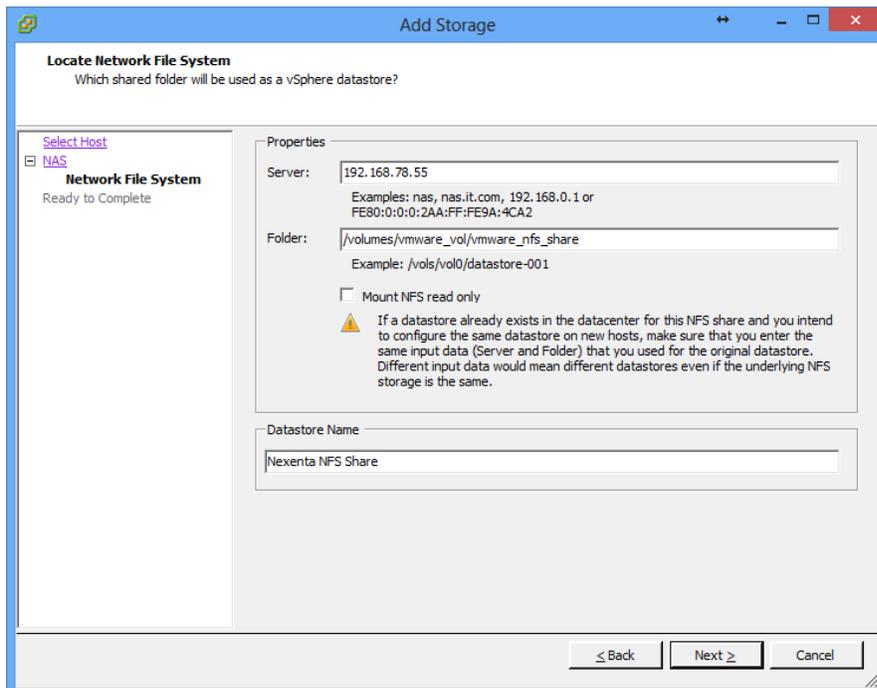
Figure 12: NFS Storage Type



Enter the following information

- **Server** : This is the IP or the DNS resolvable name of the NexentaStor Array
- **Folder** : This is the full path to the NFS share we created earlier. This path will be /volumes/<volume\_name>/<share\_name> (e.g. /volumes/vmware\_vol/vmware\_nfs\_share). (See [Editing Folder Options](#) for this information)
- **Datastore Name** : This is a name used inside of vCenter to identify the storage

Figure 13: NFS Connection Properties



- Click **Finish**

## ISCSI STORAGE CONFIGURATION

This chapter includes the following topics:

- About iSCSI Configuration**
- Adding Remote Initiators to NexentaStor**
- Create an Initiator Group**
- Create a Target Group**
- Creating a ZVOL**
- Configuring iSCSI in vSphere Traditional Client**
- Adding iSCSI Servers to ESX Hosts**
- Adding iSCSI Storage to the vSphere Server using traditional client**
- Using VAAI for Block Storage**
- Validate VAAI is Enabled**

### About iSCSI Configuration

The following items represent the best practices and recommendations when using iSCSI to access storage.

- Ensure that each ESX host has at least one dual 10GbE NIC
- Use at least two physical IP network switches.
- On the NexentaStor Storage Array side make sure you have at a minimum of one LACP pair of 10GbE NICs, attached to physical switches and a working port-channel group
- Ensure that your iSCSI network has enough throughput and low latency to handle iSCSI traffic
- Isolate the iSCSI traffic through different VLANs and/or VMware vSwitches for iSCSI
- Configure VMware iSCSI in port binding mode to achieve best performance. This also provides load balancing of the I/O traffic between paths and failover
- Change the storage array type to VMW\_SATP\_ALUA
- Change path selection policy VMW\_PSP\_MRU to VMW\_PSP\_RR, and verify that all NICs are balancing the I/O traffic

### Adding Remote Initiators to NexentaStor

A remote initiator permits the vSphere server to access storage on the NexentaStor Array.

- Click **Data Management > SCSI Target**
  - In the **iSCSI** pane, click **Remote Initiators**
  - Click **Create** to create a new remote initiator
  - Type a **name** for the remote initiator (This is the IQN of the iSCSI software adapter, iqn.1998-01.com.vmware:esx1-49850a8e)
  - Click **Create**

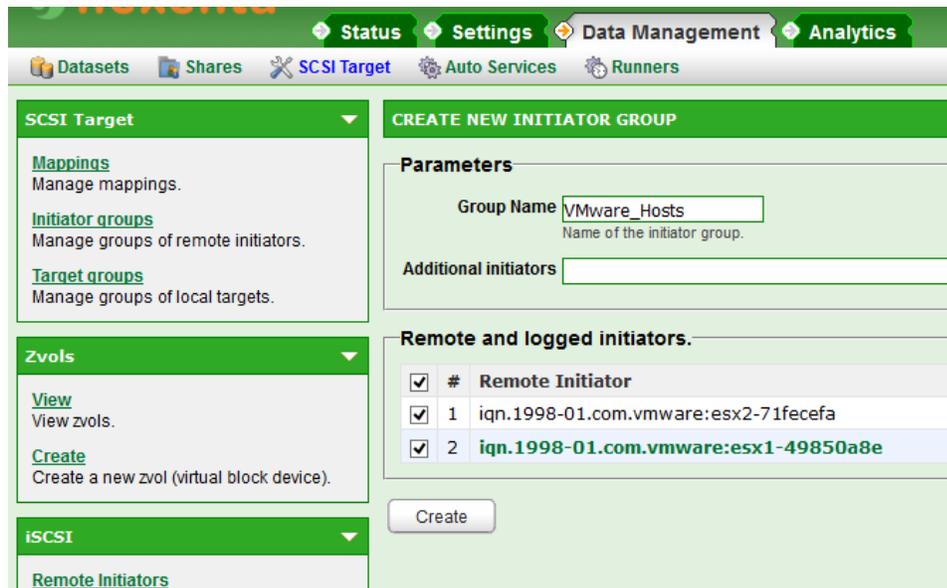
Repeat this process for all ESX HOSTS that will be accessing the iSCSI storage

## Create an Initiator Group

An Initiator Group controls who can see iSCSI devices (ZVOLs) on the network. By using Initiator Groups the administrator can control access to iSCSI presented luns.

- Click **Data Management > SCSI Target**
  - In the **SCSI Target** pane, click **Initiator Groups**
  - Type a **name** for the group
  - Select the **Remote Initiators** for the ESX hosts you want to add to this group

Figure 14: Creating Initiator Groups



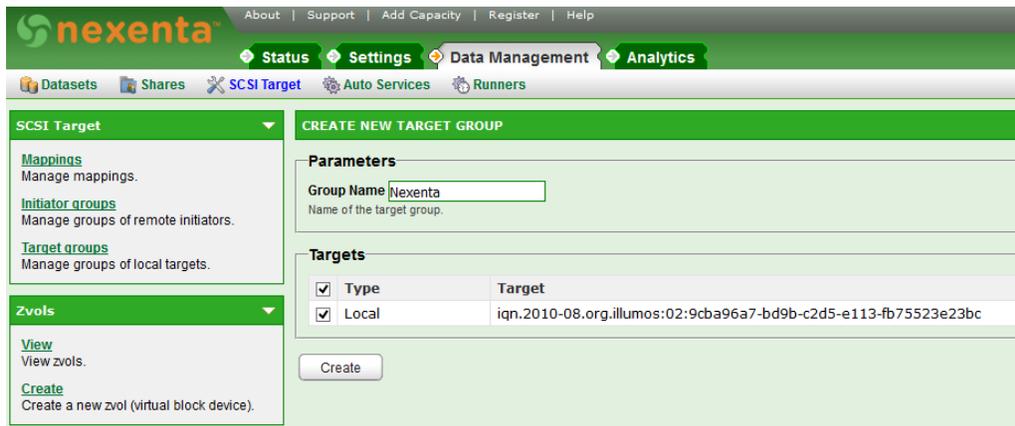
- Click **Create**

## Create a Target Group

A Target Group allows you to specify which IQN an iSCSI device is presented from. This provides another layer of segregation on the storage array. We are going to create a single Target Group that includes the default IQN of the NexentaStor Array

- Click **Data Management > SCSI Target**
  - In the **SCSI Target** pane, click **Target Groups**
  - Type a **name** for the group
  - Select the **Target**
  - Click **Create**

Figure 15: Creating Target Groups



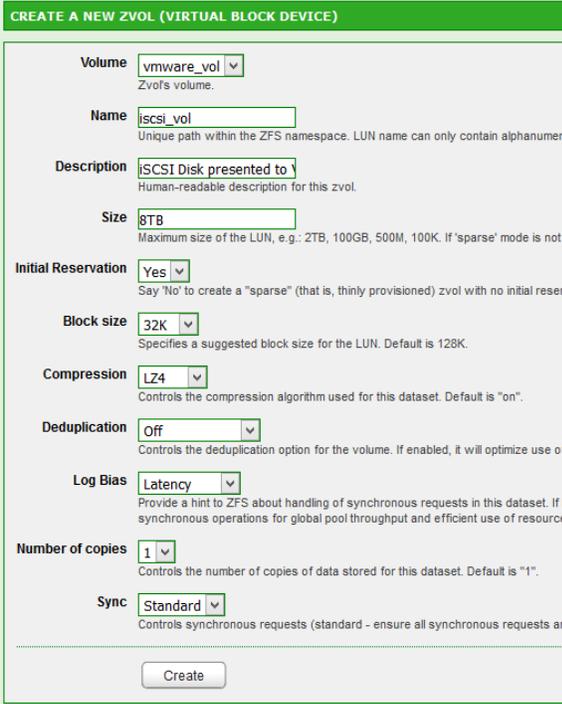
## Creating a ZVOL

The ZVOL is a virtual block device that you access through iSCSI.

- Click **Data Management > SCSI Target**
- In the **ZVOLs** pane, click **Create**
- Use the following settings when creating a new ZVOL:
  - **Name:** This is something that can help identify the iSCSI device on the NexentaStor Array
  - **Size:** XXTB (GB,TB) This is the size of the iSCSI disk you are going to present
  - **Initial Reservation:** **No** – This is for thin provisioning.
  - **Block size:** **32K (High I/O) – 128K (High Throughput)**
  - **Compression:** **LZ4**
  - **Deduplication:** **Off**

For all other settings, accept the default or edit the value as appropriate for your network.

Figure 16: ZVOL Creation Options



**CREATE A NEW ZVOL (VIRTUAL BLOCK DEVICE)**

**Volume**   
Zvol's volume.

**Name**   
Unique path within the ZFS namespace. LUN name can only contain alphanumeric characters.

**Description**   
Human-readable description for this zvol.

**Size**   
Maximum size of the LUN, e.g.: 2TB, 100GB, 500M, 100K. If 'sparse' mode is not selected, the size must be a multiple of 128K.

**Initial Reservation**   
Say 'No' to create a "sparse" (that is, thinly provisioned) zvol with no initial reservation.

**Block size**   
Specifies a suggested block size for the LUN. Default is 128K.

**Compression**   
Controls the compression algorithm used for this dataset. Default is "on".

**Deduplication**   
Controls the deduplication option for the volume. If enabled, it will optimize use of space.

**Log Bias**   
Provide a hint to ZFS about handling of synchronous requests in this dataset. If latency is selected, synchronous operations for global pool throughput and efficient use of resources are prioritized.

**Number of copies**   
Controls the number of copies of data stored for this dataset. Default is "1".

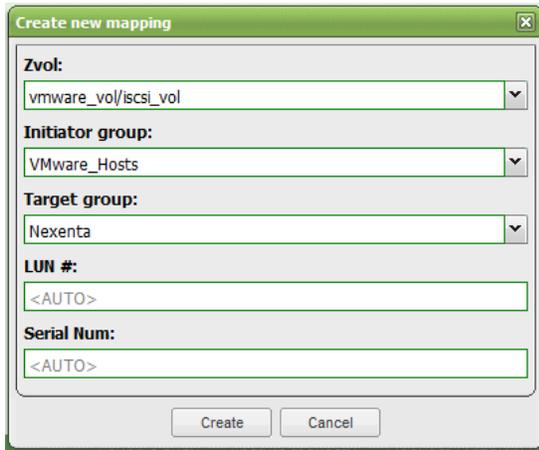
**Sync**   
Controls synchronous requests (standard - ensure all synchronous requests are completed before returning).

- Click **Create**.

## Mapping an iSCSI Target

- Click **Data Management > SCSI Target**
  - In the **SCSI Target** pane, click **Mappings**
  - If no mappings currently exist click the **HERE** link to create a new one
  - Select the **Zvol** you want to share from the drop down menu
  - Select the **Initiator Group** you want to have access to the Zvol
  - Select the **Target Group** you wish to present from
  - Leave **Lun#** and **Serial Num** to **Auto**
  - Click **Create**
  - **Repeat** this process for all iSCSI devices you wish to share

Figure 17: How to map a ZVOL



## Configuring iSCSI in vSphere Traditional Client

- To configure iSCSI in vSphere:
  - Log in to the vSphere server
  - Select the **vSphere server**, (not a VM)
  - Click **Configuration > Storage Adapters**
  - Click **Add**
  - Click **OK** to add the software iSCSI Adapter, if it is not already present.
  - Write down the IQN associated with the newly created iSCSI adapter

Figure 18: Adding the iSCSI Software Adapter

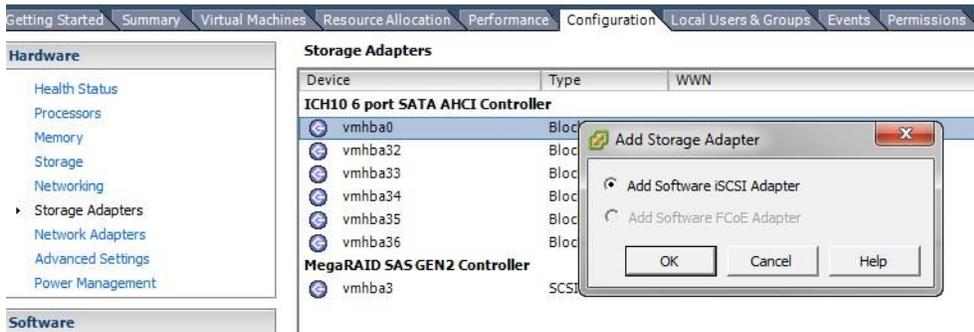


Figure 19: Finding the IQN of the iSCSI Software Adapter

### Storage Adapters

Device	Type	WWN
<b>iSCSI Software Adapter</b>		
vmhba37	iSCSI	iqn.1998-01.com.vmware:esx1-49850a8e:
<b>ICH10 6 port SATA AHCI Controller</b>		
vmhba0	Block SCSI	
vmhba32	Block SCSI	

## Modify Advanced Properties for iSCSI Interface

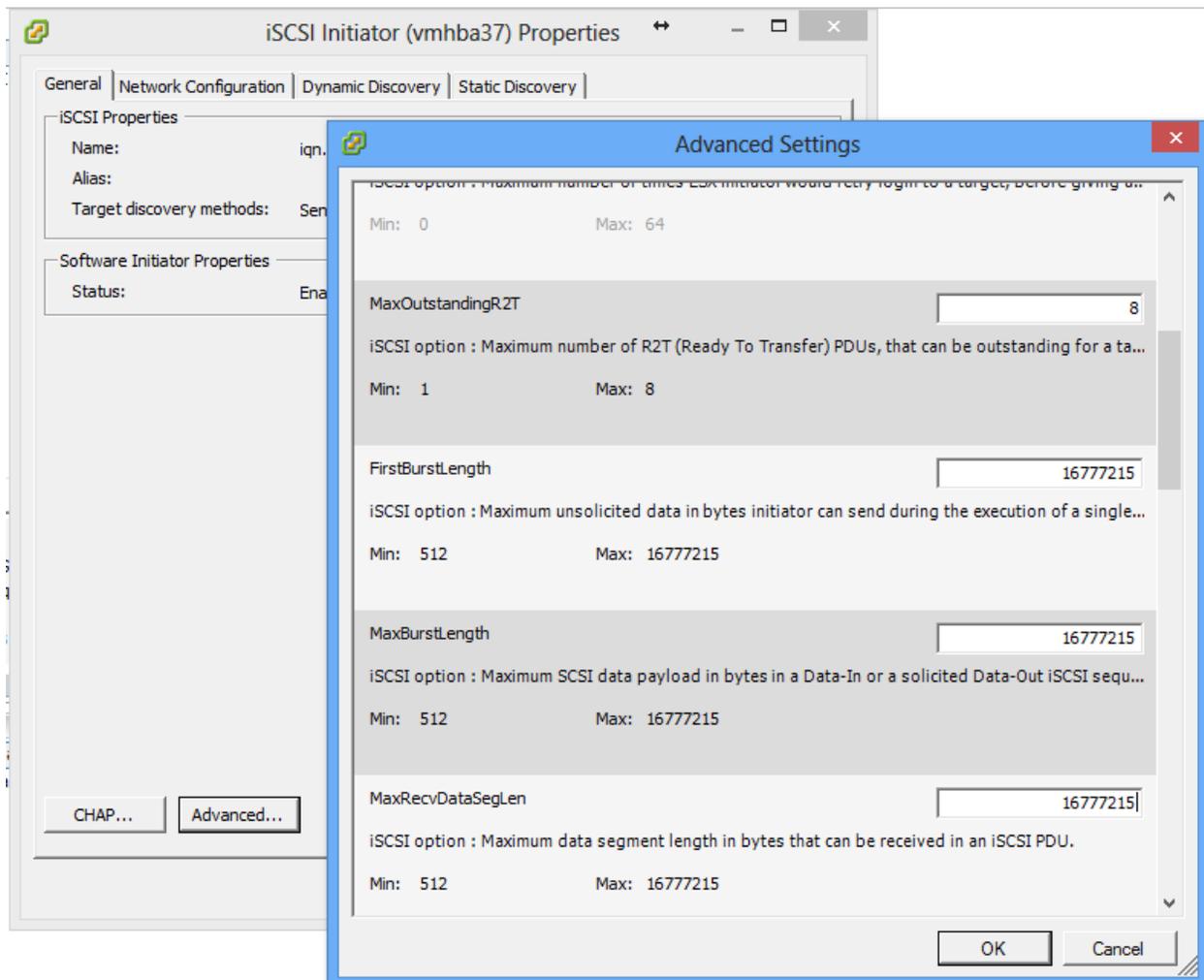
Alter the following iSCSI software parameters listed in the following table.

Table 2: iSCSI Software Parameters

ISCSI ADVANCED SETTINGS OPTION	VALUE
MaxOutstandingR2T	8
FirstBurstLength	16777215
MaxBurstLength	16777215
MaxRecvDataSegLen	16777215
RecoveryTimeout	120

To perform this task, **right-click** on your **iSCSI interface**, then click on **Properties** and **Advanced options**

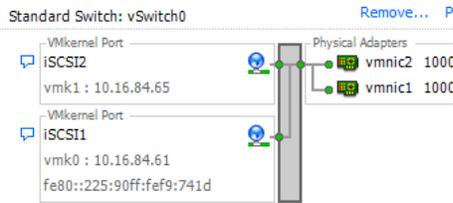
Figure 20: Advanced Settings for iSCSI Parameters



## Configure port binding in VMware

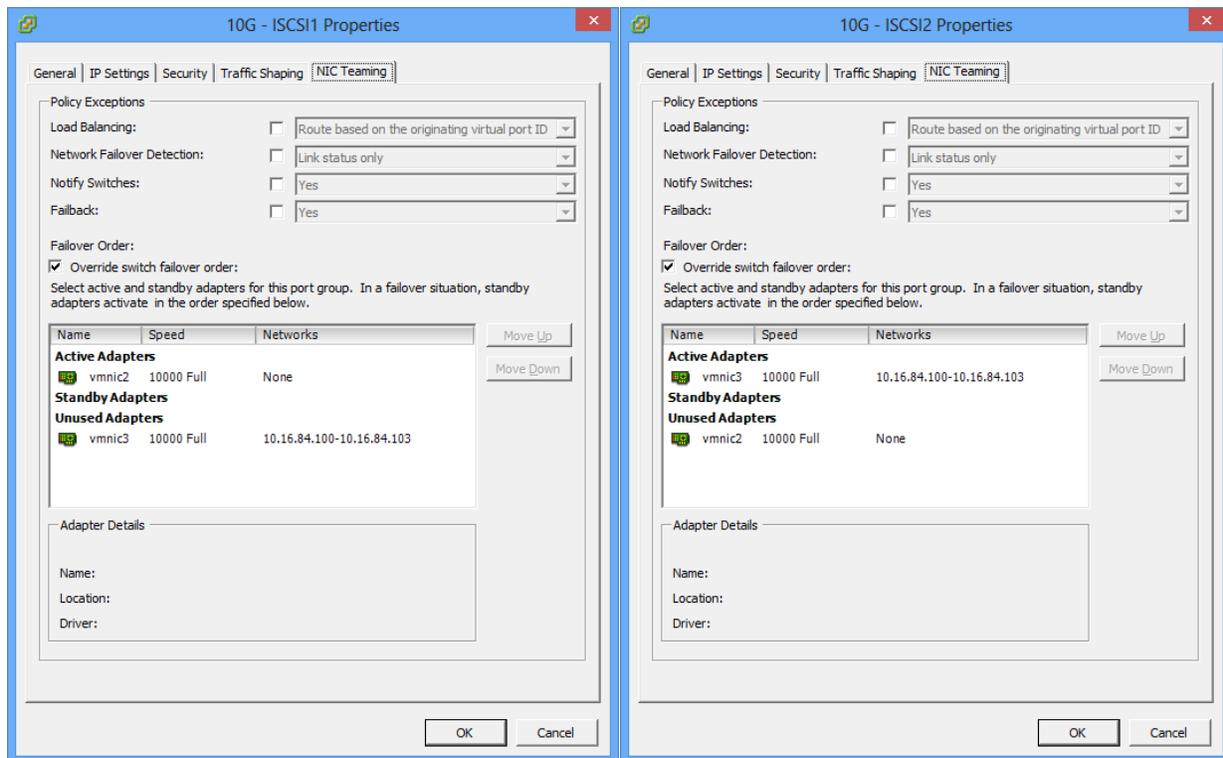
- Create a new vSwitch with two or more VMkernel ports and two 10GbE interfaces

Figure 21: vSwitch with two VMkernel ports



For each VMkernel port, enable the **Override switch failover mode** and make sure that only one 10GbE adapter is enabled per port group. Additional cards must be moved to Unused Adapters. To perform this task, select the ESXi5.x host, then select the **Configuration** tab, **Networking** and then the **Properties** of your iSCSI vSwitch. Select the **iSCSI port group**, click on **Edit**, and then select the **NIC Teaming** tab.

Figure 22: VMkernel properties for iSCSI1 & iSCSI2



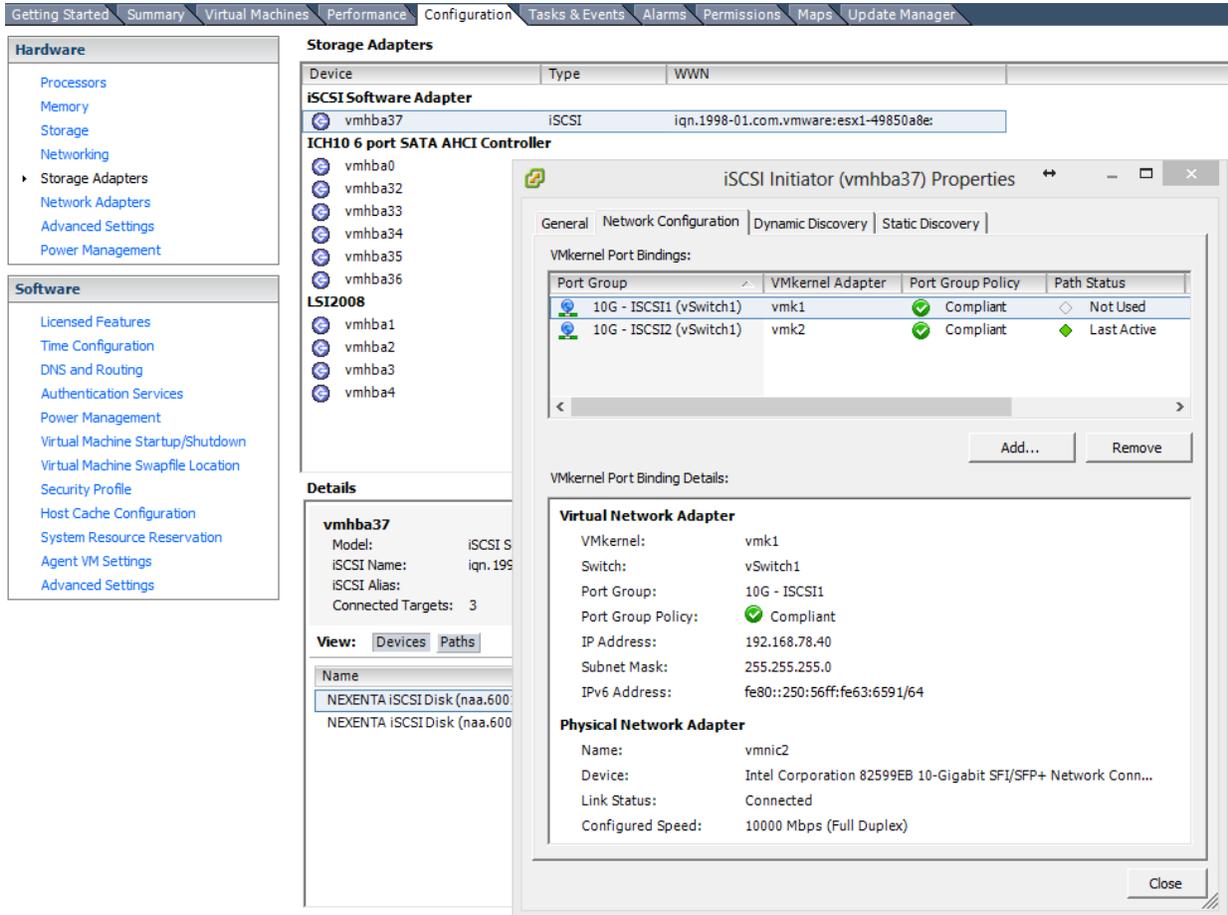
**Note:** Each port group should have only one active adapter. Move all other adapters to the Unused Adapters section.

## Adding Port Groups to iSCSI Software Adapter

To perform this task, select the ESXi5.x host, then select the **Configuration** tab, **Storage Adapters** and then the **Right-Click** on the **iSCSI Software Adapter** and select **Properties**

Select the **Network Configuration** tab, click on **Add**, and then select **iSCSI01** and **iSCSI02** port groups. Click **OK**.

Figure 23: Properly added port groups on an iSCSI adapter

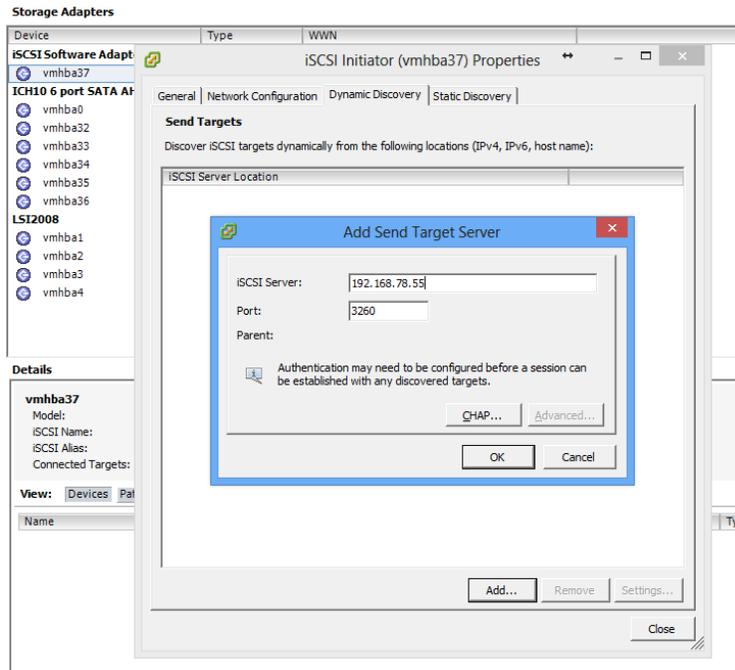


## Adding iSCSI Servers to ESX Hosts

In vSphere Client, you must add the iSCSI Server to each vSphere host.

- Log in to the vSphere Server.
- Select the **vSphere server**, (not a VM).
- Click **Configuration > Storage Adapters**
- **Right-click the iSCSI Storage Adapter** and click **Properties**.
- Click **Dynamic Discovery > Add**.

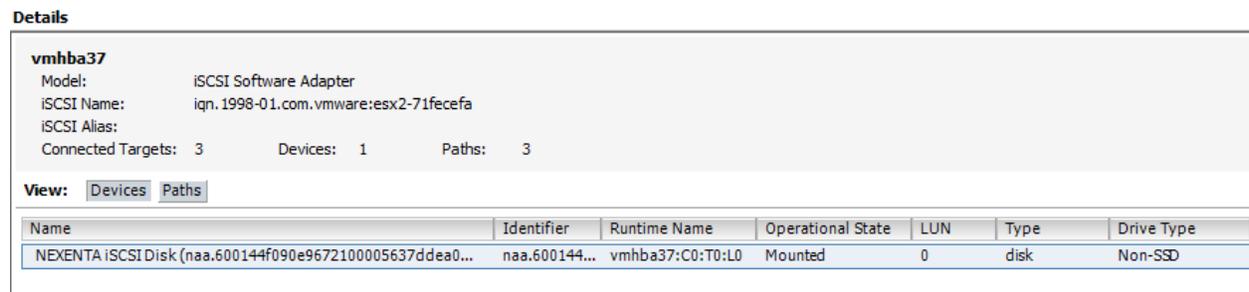
Figure 24: Adding a Target Server within Dynamic Discovery



Type the **IP address** of the iSCSI server. Click **OK**.

Click **Yes** to rescan and discover the iSCSI device.

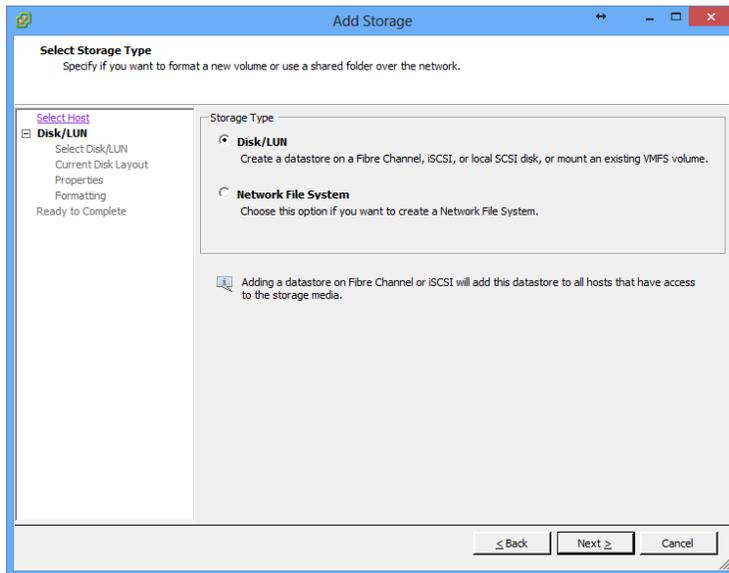
Figure 25: iSCSI Device presented to ESX host



## Adding iSCSI Storage to the vSphere Server using traditional client

- Click **View > Inventory > Datastore and Datastore Clusters**
  - **Right-Click** on your **Datacenter** and select **Add Datastore**
    - Select the Host you want to have access to the NFS datastore
    - Click **Next**
    - Select **Disk/LUN** as the Storage Type
    - Click **Next**

Figure 26: iSCSI Storage Type



- Select the **Disk** and Click **Next**
- Click **Next**
- Enter **Datastore Name** : This is a name used inside of vCenter to identify the storage
- Select **Maximum** available space
- Click **Next**
- Click **Finish**. The iSCSI datastore is available with hardware acceleration.

## Configure Path Selection Policy for Round Robin

The following steps will configure the ESX host to identify all NexentaStor Block Devices and assign them the Path Selection Policy of VMW\_PSP\_RR (Round Robin) and the Storage Array Type Plug-in to VMW\_SATP\_ALUA. We will also show how to modify the Round Robin I/O policy from the default settings to maximize Round Robin behavior. **THE ESX HOST WILL NEED TO BE REBOOTED DURING THIS PROCESS**

SSH to your ESX host and use the following command to create a custom identification rule on **ALL ESX HOSTS**.

- **esxcli storage nmp satp rule add -V NEXENTA -M COMSTAR -P VMW\_PSP\_RR -s VMW\_SATP\_ALUA -e "NexentaStor Block Devices"**

This command will add a SATP rule that will set the Path Selection Policy to **VMW\_PSP\_RR** for any devices that have a Vendor name of "NEXENTA" and a model of "COMSTAR". This will also set the Storage Array Type Plug-In to **VMW\_SATP\_ALUA**

- **REBOOT THE ESX HOST**

Identify all the device IDs currently presented to the ESX host from the NexentaStor Array and confirm that the Path Selection Policy is **VMW\_PSP\_RR**

- **esxcli storage nmp device list | grep -i 'NEXENTA iSCSI Disk'**  
This will return a list of devices with an output similar to this:
  - Device Display Name: NEXENTA iSCSI Disk  
(naa.600144f090e9672100005637ddea0001)
- **esxcli storage nmp device list -d naa.600144f090e9672100005637ddea0001 | egrep -i "Nexenta | Policy:"**
  - Device Display Name: NEXENTA iSCSI Disk  
(naa.600144f090e9672100005637ddea0001)
  - Path Selection Policy: **VMW\_PSP\_RR**

## Modify default Round Robin I/O Limits

To change the Round Robin I/O operation limit, use the steps shown in the following ESXi command lines.

To view the current Round Robin policy for a given device use the following commands

- **esxcli storage nmp device list | grep -i 'NEXENTA iSCSI Disk'**  
This will return a list of devices with an output similar to this:
  - Device Display Name: NEXENTA iSCSI Disk  
(naa.600144f090e9672100005637ddea0001)
- **esxcli storage nmp psp roundrobin deviceconfig get -d naa.600144f090e9672100005637ddea0001**
  - Byte Limit: 10485760
  - Device: naa.600144f090e9672100005637ddea0001
  - IOOperation Limit: 1000
  - Limit Type: Default
  - Use Active Unoptimized Paths: false

As you can see the default policy has the paths being used every 1,000 IOPS. If we want to have a truly equally balanced Round Robin we need to modify this to every 1 IOP. These changes get applied directly to the device so if you add additional iSCSI disks to the ESX environment you will need to rerun the following command.

An easy way to loop through all the NexentaStor iSCSI devices that are currently presented to this ESX host and modify the I/O Limit and type of policy would be the following script:

```
for disk in `esxcli storage nmp device list | awk '/NEXENTA iSCSI Disk/i {print $7}' | cut -c 2-37`
do
esxcli storage nmp psp roundrobin deviceconfig set -d $disk -l 1 -t iops
done
```

Use the following script to check that the changes were applied

```
for disk in `esxcli storage nmp device list | awk '/NEXENTA iSCSI Disk/i {print $7}' | cut -c 2-37`  
do  
esxcli storage nmp psp roundrobin deviceconfig get -d $disk  
done
```

## Using VAAI for Block Storage

While NexentaStor 4.0.4 is fully VAAI Block certified, it also ships with VAAI support disabled by default. In order to use VAAI Block on a fresh NexentaStor 4.0.4 system, the user will have to first enable the functionality on NexentaStor.

The process follows the steps below:

- Enable ATS on NexentaStor<sup>1</sup>
- Upgrade NexentaStor
- Re-enable ATS on existing ESXi servers

The procedures and background information provided here summarize information provided in VMware knowledge base articles. You are recommended to consult the original documents. The following KB articles were reviewed by Nexenta in consultation with VMware:

- [1033665](#)
- [2037144](#)
- [2006858](#)
- [2030416](#)
- [2094604](#)

In case of any issues with these procedures, you should first open support cases with VMware and then, as needed, with Nexenta as joint support.

## Warning about upgrading without preparation

VMware and Nexenta both recommend against an upgrade that changes defaults without first configuring ATS locking for affected datastores to be consistent with the new defaults. As per VMware KB [2037144](#), datastores configured to use ATS-only locking fail to mount after an upgrade that changes defaults to disable and do not show up in the vSphere client datastore view. In this situation Nexenta recommends that you revert to the previously running snapshot, thereby reverting the change in defaults, performing the preparatory steps outlined below, then returning to the upgrade checkpoint.

---

<sup>1</sup> To change ATS settings on NexentaStor requires a reboot. You can minimize reboots by applying configuration changes before upgrading your system, as they will take effect if the settings are on the system at upgrade. A `vaactl` script, provided by Nexenta, may be used to enable VAAI features. Download the `vaactl` script from <https://nexenta.com/products/downloads/nexentastor> in the NexentaStor Utilities section. See section Enabling ATS on NexentaStor for more details.

If you are running ESXi 6.0 with multi-extent datastores mounted by multiple hosts, you should consult KB [2094604](#) and open a support case with VMware as necessary, given the following notice in that KB:

- The combination of one host using ATS-only and another host using SCSI Reserve/Release might result in file system corruption.

This can result from other procedures to disable ATS documented by VMware but not recommended by Nexenta for this situation, including others from the listed KBs.

## Planning for upgrade

You should schedule a maintenance window to prepare for upgrades on ESXi hosts and complete NexentaStor upgrades. As per KB [2030416](#), the datastore must be inactive (guests must either be migrated off the datastore or powered off) before disabling ATS:

- All virtual machines must be migrated off the affected datastore, or powered off, prior to running the below steps.

## Enabling ATS on NexentaStor

To identify VMFS datastores using NexentaStor storage and check their ATS status, see the section [Identifying VMFS datastore extents and checking ATS status](#).

To change ATS settings on NexentaStor requires a reboot. You can minimize reboots by applying configuration changes before upgrading your system, as they will take effect if the settings are on the system at upgrade.

The `vaaictl` script may be used to enable VAAI features. Download the `vaaictl` script from <https://nexenta.com/products/downloads/nexentastor> in the NexentaStor Utilities section. Copy the script to admin's home directory, and make sure that it is executable by setting the execution mode of the file:

- `# -bash-4.2$ chmod 555 ./vaaictl`
- To apply changes, ssh into the device as admin and run the following script:
  - `# -bash-4.2$ sudo ./vaaictl --enable`
- If you need to disable VAAI thereafter, you can follow the previous steps, changing the `vaaictl` invocation either to restore defaults, allowing the system to follow product defaults again:
  - `# -bash-4.2$ sudo ./vaaictl --default`
- or to disable explicitly:
  - `# -bash-4.2$ sudo ./vaaictl --disable`

All settings changes require reboot. You can check current settings using this invocation, which will warn you if the config file has been changed without a system reboot, which means that settings on the running system are uncertain:

- `# -bash-4.2$ ./vaaictl --status`

## Identifying VMFS datastore extents and checking ATS status

VMFS datastores use one or more extents. All procedures assume that a datastore uses extents backed exclusively by NexentaStor LUNs. In case of datastores using extents backed by storage from more than one vendor, you should open a support case with VMware to confirm appropriate procedures and identify possible further sources of risk, opening joint support cases with Nexenta and other storage vendors as appropriate.

To enumerate mounted datastore and identify which extents they use, log into the ESXi console and type the following from the ESXi console (ssh into the ESXi host(s), using what is also termed “tech support mode”):

Figure 27: Storage Extents

```

~ # esxcli storage vmfs extent list
Volume Name  VMFS UUID          Extent Number Device Name
Partition
-----
-----
ham01-zv01   546fcc0f-d40379dd-5ae5-002590daef96      0
naa.600144f0c140cf6e0000546fca5d0002      1
lrtsesx01-ds-01 53b43e1d-d4ab8871-1a8d-002590daef96      0
t10.ATA_____ST1000NM00332D9ZM173_____Z1W11CAL
3
  
```

To confirm that an extent is backed by a NexentaStor block device, use **esxcli storage core device list -d**, as in our example:

Figure 28: ESX Device List

```

~ # esxcli storage core device list -d naa.600144f0c140cf6e0000546fca5d0002
naa.600144f0c140cf6e0000546fca5d0002
  Display Name: NEXENTA Fibre Channel Disk
(naa.600144f0c140cf6e0000546fca5d0002)
  Has Settable Display Name: true
  Size: 2097152
  Device Type: Direct-Access
  Multipath Plugin: NMP
  Devfs Path: /vmfs/devices/disks/naa.600144f0c140cf6e0000546fca5d0002
  Vendor: NEXENTA
  Model: COMSTAR
  Revision: 1.0
  SCSI Level: 5
  Is Pseudo: false
  Status: on
  
```

```
Is RDM Capable: true
Is Local: false
Is Removable: false
Is SSD: false
Is Offline: false
Is Perennially Reserved: false
Queue Full Sample Size: 0
Queue Full Threshold: 0
Thin Provisioning Status: yes
Attached Filters:
VAAI Status: unknown
Other UIDs:
vml.0200010000600144f0c140cf6e0000546fca5d0002434f4d535441
Is Local SAS Device: false
Is USB: false
Is Boot USB Device: false
No of outstanding IOs with competing worlds: 32
```

Devices exported from NexentaStor are evident because the Vendor field is set to NEXENTA. For each mounted datastore using NextaStor-exported extents, use **vmkfstools -Phv1 /vmfs/volumes/** to confirm that ATS is enabled, as in our example:

Figure 29: ATS Mode Validation

```
~ # vmkfstools -Phv1 /vmfs/volumes/ham01-zv01
VMFS-5.60 file system spanning 1 partitions.
File system label (if any): ham01-zv01
Mode: public ATS-only
Capacity 2 TB, 725.6 GB available, file block size 1 MB, max file size
62.9 TB
Volume Creation Time: Fri Nov 21 23:34:39 2014
Files (max/free): 130000/129619
Ptr Blocks (max/free): 64512/63162
Sub Blocks (max/free): 32000/31911
Secondary Ptr Blocks (max/free): 256/256
File Blocks (overcommit/used/overcommit %): 0/1353841/0
Ptr Blocks (overcommit/used/overcommit %): 0/1350/0
Sub Blocks (overcommit/used/overcommit %): 0/89/0
Volume Metadata size: 814383104
UUID: 546fcc0f-d40379dd-5ae5-002590daef96
Partitions spanned (on "lvm"):
    naa.600144f0c140cf6e0000546fca5d0002:1
Is Native Snapshot Capable: YES
OBJLIB-LIB: ObjLib cleanup done.
```

The "ATS-only" output in the mode line indicates that the datastore is configured to use ATS. If that element is not present in the mode line, the datastore is not configured to use ATS.

## Disabling ATS on Existing Datastores

If you need to disable VAAI, Nexenta recommends disabling ATS on a per-device basis, consistent with VMware's recommendation in KB [2006858](#):

Disabling VAAI entirely on the ESXi host may introduce issues in the environment. Instead of disabling VAAI for all devices, you can be disable it only for the affected LUN without impacting other LUNs.

To identify VMFS datastores using NexentaStor storage and check their ATS status, see the section "Identifying VMFS datastore extents and checking ATS status".

To disable ATS, use **vmkfstools --configATSOOnly 0 /vmfs/devices/disks/**, as in our example:

Figure 30: Disable ATS Command

```
- # vmkfstools --configATSOOnly 0 /vmfs/devices/disks/  
naa.600144f0c140cf6e0000546fca5d002:1
```

The command will produce the following output, including a prompt to confirm the change of settings:

Figure 31: Disable ConfigATSOOnly Result

```
VMware ESX Question:  
VMFS on device naa.600144f0c140cf6e0000546fca5d0002:1 will be upgraded to or  
downgraded from ATS capability. Please ensure that the VMFS-5 volume is not in  
active use by any local or remote ESX 4.x servers.  
  
Continue with configuration of ATS capability?  
  
0) _Yes  
1) _No  
  
Select a number from 0-1: 0  
  
Checking if remote hosts are using this device as a valid file system. This may  
take a few seconds...  
Downgrading VMFS-5 on 'naa.600144f0c140cf6e0000546fca5d0002:1' from ATS  
capability...done
```

In case of any other output, you are recommended to open a support case with VMware, requesting joint support from Nexenta as appropriate.

Once ATS-only mode has been disabled for the datastore, you may proceed with the upgrade, checking guest I/O afterwards. VMware KB [2006858](#) provides a list of symptoms to check in case resulting problems with storage availability are suspected or apparent.

## Enabling ATS on Existing Datastores

To identify VMFS datastores using NexentaStor storage and check their ATS status, see the section "Identifying VMFS datastore extents and checking ATS status". To enable ATS, use **vmkfstools --configATSOOnly 1 /vmfs/devices/disks/**, as in our example:

Figure 32: Enable ATSONly Command

```
~ # vmkfstools --configATSONly 1 /vmfs/devices/disks/  
naa.600144f0c140cf6e0000546fca5d002:1
```

Confirm the disk name on the datastore inventory list within the vCenter Web Client.

The command will produce the following output, including a prompt to confirm the change of settings:

Figure 33: ATS Confirmation Prompt

```
VMware ESX Question:  
VMFS on device naa.600144f0c140cf6e0000546fca5d002:1 will be upgraded to or  
downgraded from ATS capability. Please ensure that the VMFS-5 volume is not in  
active use by any local or remote ESX 4.x servers.  
  
Continue with configuration of ATS capability?
```

Figure 34: Enable ConfigATSONly Result

```
0) __Yes  
1) __No  
  
Select a number from 0-1: 0  
  
Checking if remote hosts are using this device as a valid file system.  
This may take a few seconds...  
Downgrading VMFS-5 on 'naa.600144f0c140cf6e0000546fca5d002:1' from  
ATS capability...done
```

In case of any other output, you are recommended to open a support case with VMware, requesting joint support from Nexenta as appropriate.

Once ATS-only mode has been enabled for the datastore, you may proceed with the upgrade, checking guest I/O afterwards. VMware KB [2006858](#) provides a list of symptoms to check in case resulting problems with storage availability are suspected or apparent.

## Validate VAAI is Enabled

To validate that VAAI has been enabled, log into the vSphere Web Client and navigate to:

- vCenter Inventory Lists > DataStores

**Right-Click** on the table and select Show/Hide Columns and select Hardware Acceleration. The iSCSI datastores will now show “Supported”

Figure 35: VAAI Supported Datastores

vsan3-datastore1	✓ Normal	VMFS5	Unknown	t10.ATA_____WDC_WD2503ABYX2D01WERA1_____...	2
Datastore	✓ Normal	VMFS5	Supported	naa.600144f090e9672100005637ddea0001:1	8
100GBiSCSI	✓ Normal	VMFS5	Supported	naa.600144f090e967210000563917640002:1	9

## CONCLUSION

NexentaStor Arrays provide highly available , performant storage for a VMware vSphere environment. Utilizing either NFS or iSCSI solutions will allow for a robust solution providing thousands of IOPS for your virtualized environment.

## Figure List

Figure 1: Typical HA Configuration with Multiple JBODs.....	7
Figure 2: High Availability Network Configuration .....	8
Figure 3: Create Network Interface .....	10
Figure 4: Dataset creation wizard .....	11
Figure 5: Folder creation screen .....	12
Figure 6: Folder Properties .....	13
Figure 7: Editing NFS Access Rights.....	14
Figure 8: Disabling Anonymous Access.....	14
Figure 9: Accessing ESX Advanced Settings .....	16
Figure 10: VMware NFS Advance Settings.....	16
Figure 11: vSphere Client login .....	17
Figure 12: NFS Storage Type .....	18
Figure 13: NFS Connection Properties.....	18
Figure 14: Creating Initiator Groups .....	20
Figure 15: Creating Target Groups.....	21
Figure 16: ZVOL Creation Options.....	22
Figure 17: How to map a ZVOL .....	23
Figure 18: Adding the iSCSI Software Adapter .....	23
Figure 19: Finding the IQN of the iSCSI Software Adapter .....	23
Figure 20: Advanced Settings for iSCSI Parameters.....	24
Figure 21: vSwitch with two VMkernel ports .....	25
Figure 22: VMkernel properties for iSCSI1 & iSCSI2 .....	25
Figure 23: Properly added port groups on an iSCSI adapter .....	26
Figure 24: Adding a Target Server within Dynamic Discovery.....	27
Figure 25: iSCSI Device presented to ESX host .....	27
Figure 26: iSCSI Storage Type.....	28
Figure 27: Storage Extents .....	32
Figure 28: ESX Device List .....	32
Figure 29: ATS Mode Validation .....	33
Figure 30: Disable ATS Command.....	34
Figure 31: Disable ConfigATSOOnly Result.....	34
Figure 32: Enable ATSOOnly Command .....	35
Figure 33: ATS Confirmation Prompt.....	35
Figure 34: Enable ConfigATSOOnly Result.....	35
Figure 35: VAAI Supported Datastores .....	36

## Abbreviations

Abbreviation	Definition
ATS	Atomic Test & Set - an instruction used to write to a memory location and return its old value as a single atomic (i.e., non-interruptible) operation
LACP	Link Aggregation Control Protocol (LACP)
NFSD	Service that answers NFS requests from client machines.
PSP	Path Selection Plugins
SATP	Storage Array Type Plug-in
VAAI	vStorage API for Array Integration (VAAI) is an application program interface (API) framework from VMware that enables certain storage tasks, such as thin provisioning, to be offloaded from the VMware server virtualization hardware to the storage array.
ZVOL	Block based, iSCSI presented device