



Running VMware vSphere 4 on an HP P4000 SAN

White paper



Introduction	2
Initial iSCSI setup of the vSphere 4 server	2
Networking for the software initiator.....	2
Network port configurations	2
Two network ports	2
Four network ports.....	3
Six network ports	3
More than six ports	4
Enabling the iSCSI software adapter	4
HBA connectivity and networking	4
Multi-Pathing iSCSI for vSphere 4	5
Native iSCSI multi-pathing in vSphere 4 provides superior bandwidth performance by aggregating network ports better. Configuring iSCSI multi-pathing requires at least two network ports on the virtual switch. The following steps must be performed on each ESX or ESXi server individually	5
Connecting and using iSCSI volumes.....	6
Creating the first iSCSI volume on the SAN.....	6
Enabling VIPLB for performance	7
Discovery of the first iSCSI volume.....	7
Discovering additional volumes	7
Disconnecting iSCSI volumes from ESX or ESXi hosts.....	8
Troubleshooting volume connectivity.....	8
Creating a new datastore on the iSCSI volume	8
Expanding an HP P4000 volume and extending a datastore on the iSCSI volume	8
Snapshots, Remote Copy, and HP SmartClone volumes.....	10
Resignaturing.....	10
HP P4000 snapshots of vSphere 4 raw devices	10
HP P4000 snapshots of vSphere 4 VMFS datastores	10
HP P4000 Remote Copy volumes and SRM	10
HP P4000 Multi-Site SANs and vSphere 4	10
HP P4000 SmartClone volumes	10
VMotion, Clustering, HA, DRS, FT, and VCB	12
Choosing datastores and volumes for virtual machines.....	12
Best practices	13
Frequently asked questions	14
For more information.....	15

Introduction

If you're an administrator implementing VMware vSphere 4 on an HP P4000 storage area network (SAN), you should read this document in its entirety. You'll find important configuration guidelines, best practices, and frequently asked questions that will help you to accelerate a successful deployment.

Initial iSCSI setup of the vSphere 4 server

Networking for the software initiator

Before SAN connectivity via the iSCSI software adapter can be enabled, a specific network configuration is required. To correctly enable connectivity, a VMkernel network must be created with sufficient network access to the iSCSI SAN. As a best practice, use at least two Gigabit network adapters teamed together for performance and failover. As a best practice, the VMkernel network for iSCSI should be separate from the management and virtual networks used by virtual machines. If enough networks are available, VMotion and FT should use a separate VMkernel network.

Network port configurations

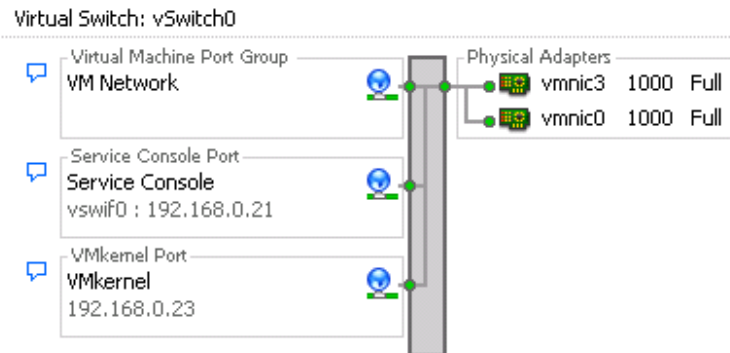
The ideal networking configuration for iSCSI depends on the number of network connections available to a vSphere 4 server. The most common number connections—for 2, 4, and 6 ports—are described here.

Two network ports

VMware vSphere 4 servers with only two Gigabit network ports are not ideal for iSCSI SANs connected by the software initiator, because they lack sufficient network bandwidth to help ensure good performance. If performance is not a concern, vSphere 4 servers can function well with only a pair of Gigabit ports. As shown in Figure 1, vSphere 4 servers with only two Gigabit network ports should be configured with:

- A single virtual switch that comprises both Gigabit ports (teamed together) and contains:
 - A network for virtual machines
 - A service console network (ESX) or management network (ESXi)
 - A VMkernel network with VMotion and FT enabled

Figure 1. Configuring a vSphere 4 server with only two Gigabit network ports



As a best practice, the VMkernel network's failover order should be reversed from the rest of the port groups on the switch. Reversing the order favors the second adapter for iSCSI, VMotion, and

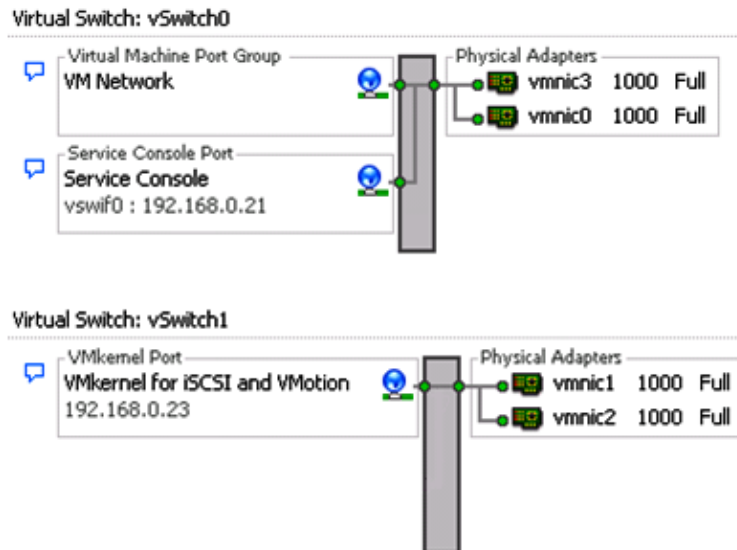
FT as it favors the first adapter for VM network traffic and management access, for best use of bandwidth.

Four network ports

VMware vSphere 4 servers with four Gigabit network ports can perform better if you separate management and virtual machine traffic from iSCSI, VMotion, and FT traffic. As illustrated in Figure 2, vSphere 4 servers with four Gigabit network ports should be configured with two virtual switches, each comprising two Gigabit ports teamed together. If possible, use one port from each of two separate Gigabit adapters. For example, if using two onboard Gigabit adapters and a dual-port Ethernet card, team together port 0 from the onboard adapter and port 0 from the Ethernet card, and then team together port 1 from the onboard adapter and port 1 from the Ethernet card. This process provides protection from some bus or card failures.

- The first virtual switch should have:
 - A service console (ESX) or management network (ESXi)
 - A virtual machine network
- The second virtual switch should have:
 - A VMkernel network with VMotion and FT enabled

Figure 2. Configuring four Gigabit network ports using two virtual switches



Six network ports

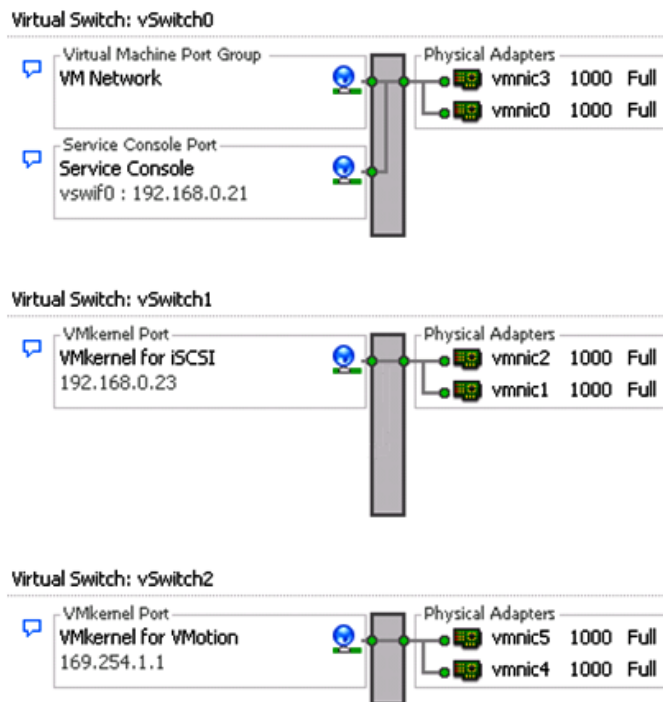
VMware vSphere 4 servers with six Gigabit network ports are ideal for delivering performance with the software iSCSI initiator. The improvement over four ports is achieved by separating VMotion and FT traffic from iSCSI traffic so that they don't have to share bandwidth. iSCSI, VMotion, and FT will perform better in this environment.

To configure vSphere 4 servers with six Gigabit network ports, use three virtual switches, each comprising two Gigabit ports teamed together, as shown in Figure 3. If possible, one port from each of the separate Gigabit adapters should be used in each team to prevent some bus or card failures from affecting an entire virtual switch.

- The first virtual switch should have:
 - A service console (ESX) or management network (ESXi)
 - A virtual machine network
- The second virtual switch should have:

- A VMkernel network with VMotion and FT disabled
- The third virtual switch should have a VMkernel network, with VMotion and FT enabled

Figure 3. Configuring six Gigabit network ports by using use three virtual switches



More than six ports

If more than six network ports are available, you can add more ports to the iSCSI virtual switch to increase available bandwidth, or you can use them for any other network services desired.

Enabling the iSCSI software adapter

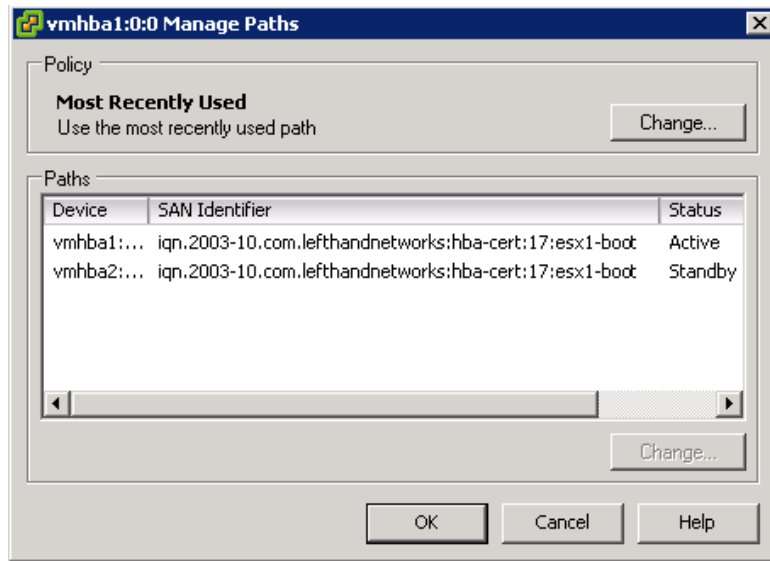
You will need to enable the vSphere 4 iSCSI software adapter on each ESX or ESXi server. The iSCSI software adapter is managed from each server's storage adapters list. Here are some guidelines:

- Enable the iSCSI adapter on each ESX or ESXi server.
- Copy or write down the iSCSI qualified name (IQN) that identifies each vSphere 4 server; it will be needed for authentication later on the SAN.
- Reboot the ESX or ESXi server after enabling the iSCSI software adapter if prompted to do so.

HBA connectivity and networking

SAN connectivity via iSCSI HBAs enables both offloading of the iSCSI processing from the vSphere 4 server and booting of ESX itself from the iSCSI SAN. HBAs do not require licensing or special networking within vSphere 4 servers as they provide a dedicated network connection for iSCSI only. The physical network for HBAs should be a Gigabit network dedicated to the SAN, just as it is for software initiators. As a best practice, use two HBA initiators (a dual port or two single ports), each configured with a path to all iSCSI targets for failover. Configuring multiple HBA initiators to connect to the same target requires authentication for each initiator's IQN to be configured on the SAN. (See Figure 4.) Typically this is configured in HP P4000 software as two servers (one for each HBA initiator), each with permissions to the same volumes on the SAN.

Figure 4. Managing paths to the SAN



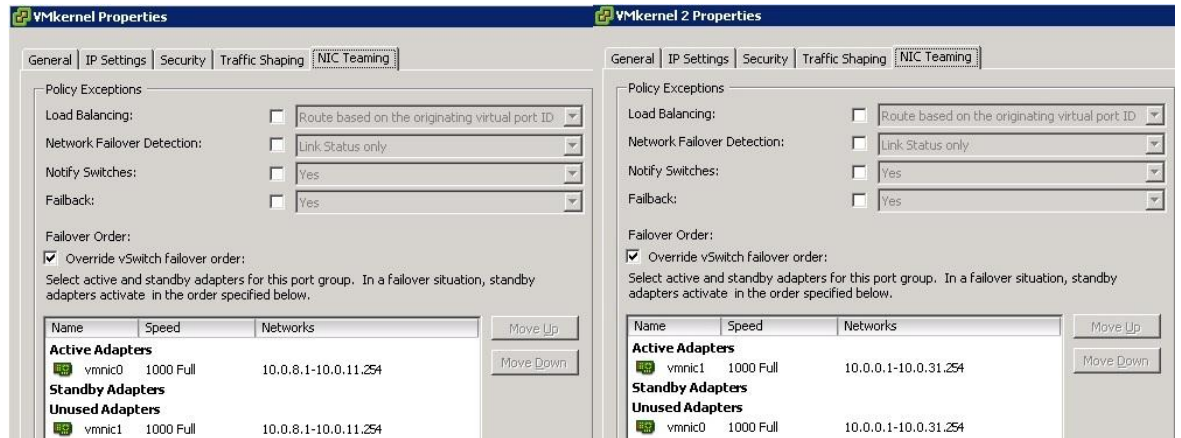
Multi-Pathing iSCSI for vSphere 4

Native iSCSI multi-pathing in vSphere 4 provides superior bandwidth performance by aggregating network ports better. Configuring iSCSI multi-pathing requires at least two network ports on the virtual switch. The following steps must be performed on each ESX or ESXi server individually.

- Create a second vmkernel port on the virtual switch for iSCSI
- For each vmkernel port on the virtual switch assign a different physical network adapter as the active adapter. This ensures the multiple vmkernel ports use different network adapters for their IO. Each vmkernel port should use a single physical network adapter and not have any standby adapters.
- From command line bind both vmkernel ports to the software iSCSI adapter. The vmk# and vmhba## must match the correct numbers for the ESX or ESXi server and virtual switch you are configuring
 - > vmkiscsi-tool -V -a vmk0 vmhba36
 - > vmkiscsi-tool -V -a vmk1 vmhba36

Once configured correctly perform a rescan of the iSCSI adapter. An iSCSI session should be connected for each vmkernel bound to the software iSCSI adapter. This gives each iSCSI LUN two iSCSI paths using two separate physical network adapters.

Figure 5. Physical Network Adapter Failover Order for vmkernel MPIO



To achieve load balancing across the two paths datastores should be configured with a path selection policy of round robin. This can be done manually for each datastore in the vSphere client or ESX can be configured to automatically choose round robin for all datastores. To make all new datastores automatically use round robin, configure ESX to use it as the default path selection policy from command line.

- > esxcli corestorage claiming unclaim --type location
- > esxcli nmp satp setdefaultpsp --satp VMW_SATP_DEFAULT_AA --psp VMW_PSP_RR
- > esxcli corestorage claimrule load
- > esxcli corestorage claimrule run

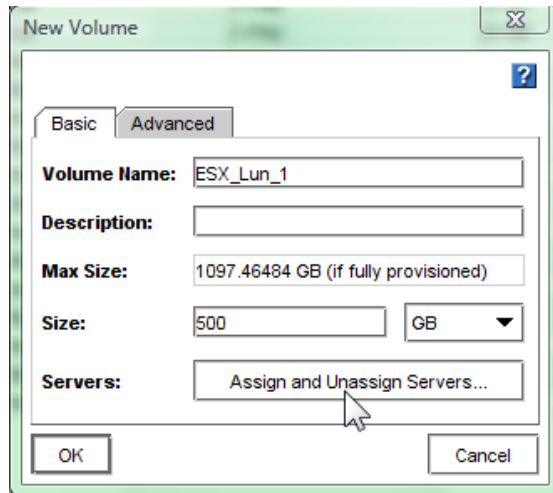
It is important to note that native vSphere 4 multipathing cannot be used with HP P4000 multi-site SAN configurations that utilize more than one subnet and VIP. Multiple paths cannot be routed across those subnets by the ESX 4 initiator.

Connecting and using iSCSI volumes

Creating the first iSCSI volume on the SAN

Before a vSphere server can mount a new iSCSI volume, that volume must be created on the SAN and must have authentication configured to allow the server to access the volume. Use the HP Centralized Management Console (CMC), as shown in Figure 5, or use the command-line interface (CLI) to create a new volume. Then, using the HP CMC, create a new server to represent the ESX or ESXi server, using the IQN(s) copied from the vSphere server's initiator(s), or use the Challenge Handshake Authentication Protocol (CHAP) as required. SAN volumes are simply assigned to these servers within the CMC or CLI.

Figure 6. Creating an iSCSI volume on the SAN



Enabling VIPLB for performance

Virtual IP load balancing (VIPLB)—or just load balancing—is a setting on each server in the HP CMC that allows for the distribution of iSCSI sessions throughout the cluster of storage nodes in order to maximize performance and bandwidth utilization. Most initiators, including vSphere 4 software initiators and HBAs, support this feature, so it is on by default in HP CMC software version 8 or higher SANs. All vSphere 4 iSCSI initiators connected to a HP P4000 SAN should have this feature enabled.

Discovery of the first iSCSI volume

To discover volumes, the software initiator of each ESX or ESXi host must have added the Virtual IP (VIP) address of the HP P4000 cluster that contains the volume to its dynamic discovery list. New targets can then be discovered by simply rescanning the iSCSI software adapter on the ESX or ESXi host.

The iSCSI session status of an ESX server can be viewed in the HP CMC by selecting the server in question. Volumes that are connected show an IP address in the Gateway Connection field on the server's Volumes & Snapshots tab. (See Figure 6.) If an IP address is not listed, look for an incorrect configuration on the ESX host or on the server in the HP CMC.

Figure 7. Viewing connected volumes

Name	Status	Provisioned Space	Utilization	Permission	Gateway Connect...
ESX_Lun_1	Normal	1 GB	0%	Read/Write	10.0.11.62
ISO_Images	Normal	1 GB	10%	Read/Write	10.0.11.63
Test_VMs	Normal	1 GB	0%	Read/Write	10.0.11.100
WinXP_Image	Normal	200 GB	10%	Read/Write	10.0.11.62

Discovering additional volumes

You can log in to and discover additional volumes by simply performing a rescan of the iSCSI software adapter or all adapters. HBAs can also add targets manually by configuring static targets.

Disconnecting iSCSI volumes from ESX or ESXi hosts

The vSphere 4 iSCSI software adapter does have a session disconnect feature. Before an iSCSI volume can be disconnected, it must first be unassigned from the "server" within the HP CMC or the next rescan will automatically connect it again. Unassigned volumes are not forcibly disconnected by the HP CMC. Instead the server simply will not be allowed to log in again. Rebooting an ESX or ESXi server will clear all unassigned volumes. Individual iSCSI sessions can be reset forcefully from the HP P4000 CLI by the "resetsession" command. Before an iSCSI session is forcibly reset from the CLI, all VMs accessing the volume through that session should be powered off or, using VMotion, reassigned to another server that will continue to access the volume.

Troubleshooting volume connectivity

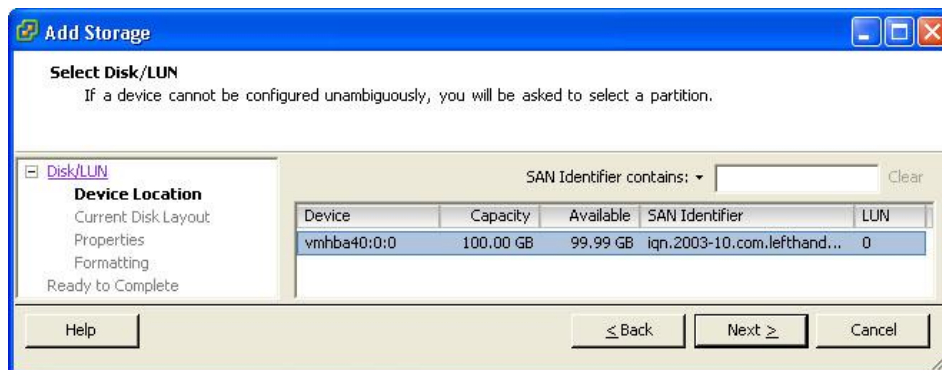
If new volumes are not showing up as expected, try the following troubleshooting steps:

- Ping the virtual IP address from the iSCSI initiator to ensure basic network connectivity.
 - For the software initiator, this can be done by logging into the vSphere 4 service console and executing `vmkping x.x.x.x`. This ping must be successful or networking is not correct for iSCSI.
 - HBAs typically have their own ping utilities inside the BIOS of the HBA, so use the appropriate utility for your HBA.
 - ESXi has a network troubleshooting utility that can be used from a KVM console to attempt a ping to the SAN.
- Double-check all IQNs or CHAP entries. For iSCSI authentication to work correctly, these must be exact. Simplifying the IQN to something shorter than the default can help with troubleshooting.
- Make sure that all servers on the SAN have load balancing enabled in the HP CMC. If there is a mix—if, for example, some have it enabled and some don't—then servers that don't have load balancing enabled may not connect to their volumes.

Creating a new datastore on the iSCSI volume

Now that the vSphere 4 server has an iSCSI SAN volume connected, it can be formatted as a new VMFS datastore or mounted as a raw device mapping (RDM) directly to virtual machines. New datastores are formatted from within the VMware vSphere client.

Figure 8. Creating a new datastore



Expanding an HP P4000 volume and extending a datastore on the iSCSI volume

Both HP P4000 volumes and vSphere 4 datastores can be expanded dynamically. If space is running low on a datastore, you must first expand the volume on the HP P4000 SAN. To increase the size of the HP P4000 volume, simply edit the volume in the HP CMC or CLI and increase its size. The HP P4000 Software immediately changes the LUN size and lays out data across the

cluster accordingly, without affecting the volumes availability. Once the volume has been expanded, you can increase the size of the VMFS datastore from within the vSphere client. Increasing the size of the datastore is preferred over adding extents.

Snapshots, Remote Copy, and HP SmartClone volumes

Resignaturing

In order for vSphere 4 servers to utilize SAN-based snapshots or clones, you need to resignature them. To resignature any clone or snapshot follow the add storage wizard in the vSphere client.

HP P4000 snapshots of vSphere 4 raw devices

HP P4000 snapshots of vSphere 4 raw devices function and are supported in exactly the same way as for physical servers, either booting from the SAN or accessing raw disks on the SAN. Detailed information about HP P4000 snapshots and how they work for specific applications is available on the HP P4000 Web site: www.hp.com/go/p4000

HP P4000 snapshots of vSphere 4 VMFS datastores

HP P4000 snapshots are very useful in a vSphere 4 environment. All virtual machines stored on a single volume can be snapshot and rolled back to at any time. Moreover, HP P4000 snapshots can be mounted to any vSphere 4 server without interrupting access to their source volume.

HP P4000 Remote Copy volumes and SRM

Remote Copy replicates HP P4000 snapshots over WAN links to remote sites for disaster recovery or backup. vSphere 4 environments can be protected by Remote Copy volumes on a scheduled basis and automated by VMware Site Recovery Manager for a simple and complete disaster recovery solution. HP provides a storage replication adapter (SRA) for VMware Site Recovery Manager (SRM) to integrate Remote Copy volumes seamlessly with a vSphere 4 environment. For more information on Remote Copy volumes, review the Remote Copy user guide installed with the CMC.

HP P4000 Multi-Site SANs and vSphere 4

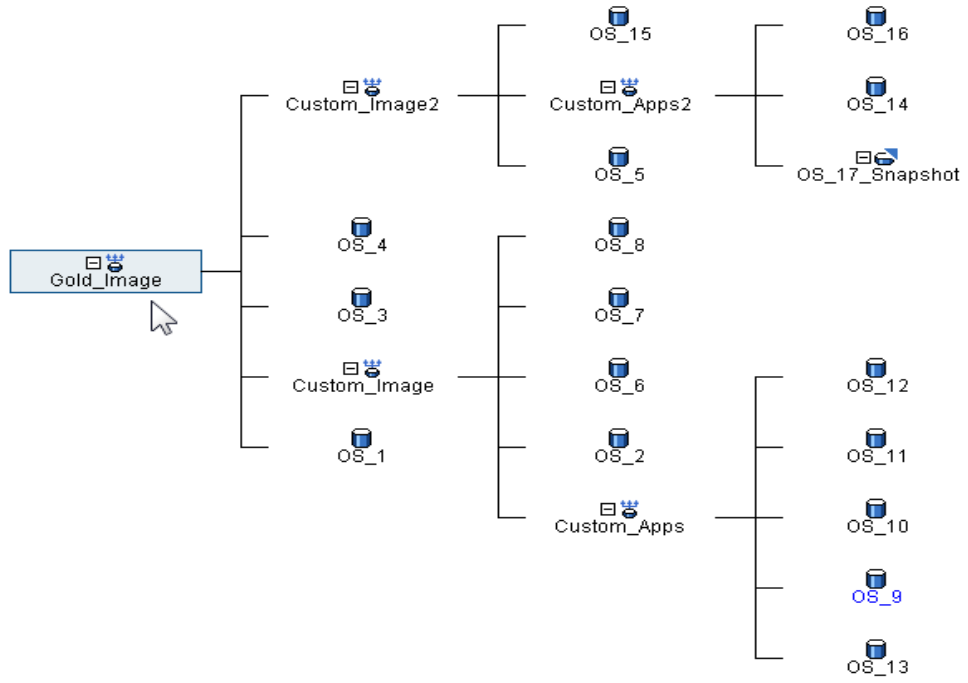
HP P4000 Multi-Site SANs enable vSphere 4 clusters to be stretched across locations to provide multi-site VMotion, HA, DRS, and FT. Multi-Site SAN configurations use synchronous replication in the underlying SAN to create a single SAN that spans both location. This allows vSphere 4 clusters to act exactly the same way they do when physically located in a single location. When connecting ESX or ESXi host to a Multi-Site SAN each of the virtual IP addresses (VIPs) of the SAN from each site should be listed in the discovery list of the ESX or ESXi software initiator. Path selection policy for Multi-Site SAN volumes should be set to fixed (default).

HP P4000 SmartClone volumes

HP SmartClone volumes are also very useful in a vSphere 4 environment. (See Figure 9.) Using HP SmartClone Technology, all the virtual machines stored on a single volume can be cloned instantly and without replicating data. SmartClone volumes consume space for data changes only from the time the SmartClone volume was created. This is the best way to deploy large quantities of cloned virtual machines or virtual desktops.

SmartClone volumes can be used seamlessly with any other HP P4000 software feature, including snapshots or Remote Copy. SmartClone volumes are also very useful for performing tests on virtual machines by quickly reproducing them, without actually using space on the SAN to copy them.

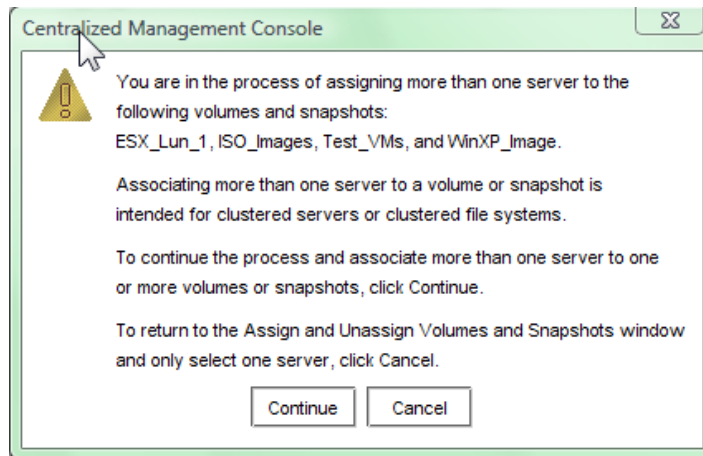
Figure 9. Map view of HP SmartClone volumes in the CMC



VMotion, Clustering, HA, DRS, FT, and VCB

The advanced vSphere 4 features of VMotion, Clustering, HA, DRS, FT, and VCB all require multiple vSphere 4 servers to have simultaneous access to volumes. To enable this on an HP P4000 volume, assign multiple servers (one for each vSphere 4 server initiator) to the same volume. (See Figure 10.) To use multiple servers, simply create one server for each vSphere 4 host that connects to the SAN, using the steps outlined earlier in the section “Creating the first iSCSI volume on the SAN.”

Figure 10. Assigning multiple servers to a single volume



When multiple servers are assigned to the same volume in the HP CMC, a warning indicates that this configuration is intended for clustered servers or clustered file systems such as ESX and ESXi. CHAP authentication can also be used for increased security, or to represent many ESX or ESXi hosts as one server in the HP CMC. Each ESX or ESXi server must be configured to use the correct CHAP credentials. New volumes on each vSphere 4 server must be discovered as described in the section “Discovery of the first iSCSI volume.”

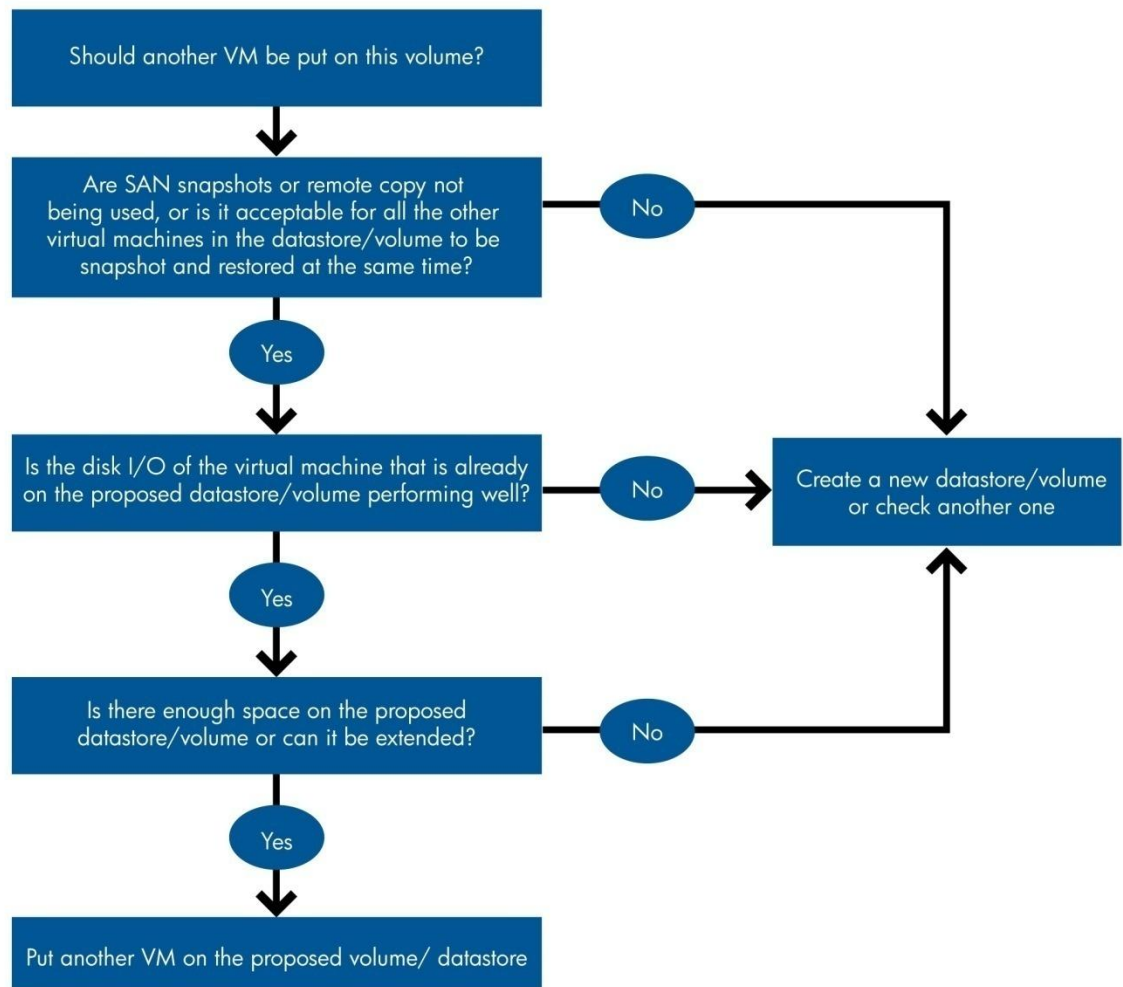
VMware Consolidated Backup (VCB) servers should have their own server authentication configured in the HP CMC, with read/write access to the SAN volumes that ESX or ESXi servers are accessing.

Choosing datastores and volumes for virtual machines

More than one virtual machine can—and generally should—be stored on each datastore or volume. The choice of where to put virtual machines should be driven by backup policies, the relationships of the VMs to each other, and performance. In general, virtual machines that do not have a relationship to one another should not be mixed on the same HP P4000 volume. P4000 features such as snapshots, Remote Copy, and SmartClone volumes are very useful with virtual machines, but will always affect all virtual machines on the same volume simultaneously. If virtual machines that have no relationship are mixed on a single volume, those virtual machines will have to be snapshot, rolled back, remotely copied, and cloned together.

Performance of virtual machines could also be affected if too many virtual machines are located on a single volume. The more virtual machines on a volume, the more I/O and SCSI reservation contention there is for that volume. Up to sixteen virtual machines will function on a single volume—but they might experience degraded performance, depending on the hardware configuration, if they are all booted at the same time. Volumes with four to eight virtual machines are less likely to affect performance.

Figure 11. Determining whether to add a virtual machine



Best practices

Use at least two Gigabit network adapters teamed together for better performance and for failover of the iSCSI connection.

Teaming network adapters provides redundancy for networking components such as adapters, cables, and switches. An added benefit to teaming is an increase in available I/O bandwidth. Network teams on HP storage nodes are easily configured in the CMC by selecting two active links and enabling a bond. Adaptive load balancing (balance ALB) is the most common teaming method used on HP P4000 SAN storage nodes. Network teams on vSphere 4 servers are configured at the virtual switch level. As an added improvement over virtual switch bonding enable native MPIO round robin for iSCSI.

The VMkernel network for iSCSI should be separate from the management and virtual networks used by virtual machines. If enough networks are available, VMotion and FT should also use a separate network.

Separating networks by functionality, iSCSI, VMotion, FT and virtual machines provides higher reliability and performance of those functions.

For improved performance, enable load-balancing for iSCSI in the HP P4000 SAN CMC.

The load-balancing feature of the HP P4000 software allows iSCSI connections to be redirected to the cluster's least busy storage node. This keeps the load on storage nodes throughout the cluster as even as possible and improves overall performance of the SAN. This setting is enabled by default in HP P4000 Software version 8 management groups and higher, but had to be enabled explicitly in previous versions.

Virtual machines that can be backed up and restored together can share the same volume.

Because HP P4000 snapshots, Remote Copy volumes, and HP SmartClone volumes work on a per-volume basis, it is best to group virtual machines on volumes based on their backup and restore relationships. For example, a test environment made up of a domain controller and a few application servers would be a good candidate for being put on the same volume. These elements could be snapshot, cloned, and restored as one unit.

Frequently asked questions

I added the HP P4000 cluster virtual IP address to vSphere 4 server's dynamic discovery; why don't I see a new target?

Most likely, either you need to perform a rescan on the iSCSI adapter, or else you have not configured authentication on the SAN correctly. Also, you should confirm all network configurations. Please refer to "Discovery of the first iSCSI volume" on page 7 of this document.

Is virtual IP load balancing supported for vSphere 4?

The vSphere 4 software adapter and hardware adapters are supported by HP P4000 Virtual IP Load Balancing. As a best practice, load balancing should be enabled on the authentication groups of all vSphere 4 initiators. It is enabled by default in HP P4000 Software version 8 management groups.

What initiator should I use for a virtual machine boot volume/LUN?

There are many ways to connect and present an iSCSI volume to a virtual machine on a vSphere 4 server. These include using the vSphere 4 software adapter, a hardware adapter (HBA), or the software iSCSI adapter provided by some guest operating systems.

For VMFS datastores containing a virtual machine's definitions and virtual disk files, you need to use the vSphere server's hardware or software adapter. You can use either adapter—HBA or software. Both give you full vSphere 4 functionality, and both are supported equally. The HBA adapter has the advantages of supporting boot from SAN and offloading iSCSI processing from the vSphere 4 server. If boot from SAN is not necessary, then the software initiator is a good choice. The impact of iSCSI processing on modern processors is minimal, so with either adapter, performance will be related primarily to the quality of physical network, disk quantity, and disk rotation speed of the SAN being attached to. Most vSphere deployments on HP P4000 SANs are done with the ESX or ESXi software initiator included with vSphere.

What initiator should I use for additional raw device mappings or virtual disks?

Aside from the boot LUN/volume, additional volumes should be used for storing application data. Specifically, as a best practice, database and logs volumes for many applications require separate volumes. These should be presented as either raw devices (RDMs) through your chosen vSphere 4 server initiator or connected as iSCSI disks directly through the software initiator of the virtual machine's guest operating system. Using RDMs or direct iSCSI allows these application volumes to be transported seamlessly between physical and virtual servers, as they are formatted in the native file system (NTFS, EXT3, etc.) of the OS. To use the guest operating system initiator successfully, follow these guidelines:

- Make sure the guest operating system initiator is supported by HP P4000 SAN technology.
- For optimal performance and failover, the guest network used by the initiator should have at least dual Gigabit links and be separate from other virtual networks, including VMkernel, VMotion, FT, Service Console, Virtual Machine public networks, etc.
- Make sure that the guest operating system is using the vmxnet NIC driver from VMware tools.

- Make sure that the virtual machine will not be used in conjunction with VMware Site Recovery Manager (SRM). SRM does not work with volumes connected by guest initiators.

How many virtual machines should be on a volume or datastore?

Refer to “Choosing datastores and volumes for virtual machines” on page 12.

Are jumbo frames supported?

Jumbo frames (IP packets configured larger than the typical 1,500 bytes—i.e., up to 9,000 bytes) are supported by all HP P4000 SANs. In order for jumbo frames to be effective, they must be enabled end to end, including in network adapters and switches. For most Gigabit iSCSI implementations jumbo frames are not required for good performance. 10 Gigabit iSCSI implementations do see significant benefits from jumbo frames.

Is VCB (VMware Consolidated Backup) supported on iSCSI?

VMware Consolidated Backup 1.0.3 and higher is fully supported on iSCSI SANs. When using VCB with HP P4000 SANs do not install the HP P4000 DSM for Microsoft MPIO on the VCB server. HP P4000 DSM for MPIO cannot connect to volumes that are in use simultaneously by ESX or ESXi.

Why does my 2 TB or higher iSCSI volume show up as 0 MB to the vSphere 4 server?

The maximum volume size supported by vSphere 4 is 2,047 GB.

What version of P4000 Software supports VMware vSphere 4?

All HP P4000 releases, including HP SAN/iQ Software versions 8 and above, support vSphere 4.

For more information

For more information on HP P4000 SANs, visit: www.hp.com/go/p4000

For more information on the HP P4000 Virtual SAN Appliance, visit: www.hp.com/go/vsa

© Copyright 2008, 2009 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

4AA2-5132ENW, April 2009

