

STORAGE

FOCUS ON

Storage for Virtual Machines

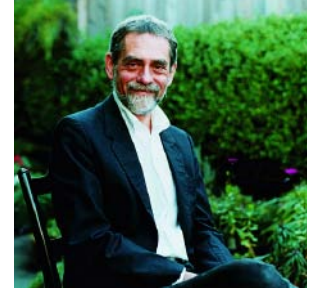


Virtualization puts new demands on storage, so whether you opt for FC, iSCSI or NAS, storage will be a key factor in your virtual environment.



INSIDE

- 4** Ten things to know about storage for VMs
- 11** iSCSI and vSphere: A good match?
- 16** Choosing NAS for Virtual Machines



Plenty of alternatives, plenty of confusion about storage for virtual machines

SAN, NAS or DAS? They're all valid types of storage for virtual servers or desktops, now your job is to find out which one will work best in your environment.

VIRTUALIZING SERVERS and desktops can make maintenance and management a lot easier for sys admins but not so much for storage managers. Consolidating servers or desktop PCs can put new strains on storage systems, but it's a lot more complicated than just needing more capacity to accommodate all those centralized virtual machines which can spawn at an alarming rate.

It's a foregone conclusion that networked storage is a basic requirement, but there are still plenty of decisions that must be made when configuring storage for virtual environments.

Our research shows that Fibre Channel arrays are, by far, the most popular type of storage used in support of virtual servers. On our latest survey, 43% of the respondents said they're using FC systems for the virtual server storage, a figure that's down a bit from the 50%-plus marks we've seen in the past. There's a considerable drop-off after FC, with second-place iSCSI registering only 17%; that's only a few percentage points ahead of the NAS and direct-attached storage. FC's predominance may be at least partly attributable to its legacy status—it supported pre-virtualization physical servers and was kept in place to handle the virtualized environment. And the fact that the other forms of storage are essentially bunched up in a dead heat for the runner-up spot indicates that there are not only uncertainties about which type of storage best serves a virtual environment, but that there are ample alternatives.

There are many considerations that figure into evaluating storage systems for virtual environments, such as whether block storage is required or if a file system will suffice. Network infrastructure may be a decisive factor, too, especially if a Fibre Channel fabric is already in place or if 10 Gig Ethernet is in the plans.

The performance of a storage system will certainly be a key factor, as consolidation tends to translate into server growth and introduces new phenomena like boot storms. And as virtualization software gets smarter with increased integration into storage systems, the capabilities of the storage array will become even more critical to the well being of your virtual environment.

The one basic rule to keep in mind is that there are few—if any—hard and fast rules when it comes to choosing storage to support your virtual machines. The trick is to carefully match the needs of your organization with the particular capabilities of a networked storage system. ☉

Rich Castagna is the editorial director of TechTarget's Storage Media Group.

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources



10 TIPS FOR MANAGING STORAGE FOR VIRTUAL SERVERS AND VIRTUAL DESKTOPS

Server virtualization and virtual desktops can make configuring and managing storage systems a lot tougher. These 10 tips will help ease some of the stress caused by managing storage in a virtual environment. By Eric Siebert

SERVER AND DESKTOP VIRTUALIZATION have provided relatively easy ways to consolidate and conserve, allowing a reduction in physical systems. But these technologies have also introduced problems for data storage managers who need to effectively configure their storage resources to meet the needs of a consolidated infrastructure.

Server virtualization typically concentrates the workloads of many servers onto a few shared storage devices, often creating bottlenecks as many virtual machines (VMs) compete for storage resources. With desktop virtualization this concentration becomes even denser as many more

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources

desktops are typically running on a single host. As a result, managing storage in a virtual environment is an ongoing challenge that usually requires the combined efforts of desktop, server, virtualization and storage administrators to ensure that virtualized servers and desktops perform well. Here are 10 tips to help you better manage your storage in virtual environments.

1. Know your storage workloads. Virtual desktop workloads are very different from virtual server workloads, and the workloads imposed by individual desktops and servers can also vary dramatically. Blindly placing VMs on hosts without regard for their disk I/O usage can create instant resource bottlenecks.

You should have a general idea of how much disk I/O a VM will generate based on the applications and workloads it will host. Therefore, you should try to balance high disk I/O VMs among both physical hosts and

data resources. If you have too many VMs with high disk I/O on a single host it can overwhelm the host's storage controller; likewise, having too many high disk I/O VMs accessing a single storage system or LUN may also create a performance bottleneck. Even if you have a good idea of your virtual machine's disk I/O workloads, it's still a good idea to use performance monitoring tools to get detailed statistics such as average and peak usage.

And don't forget that VMs are usually mobile and may not always be on the same host; they may be moved to another physical host using technologies like VMware vMotion. Having a group of busy Exchange servers ending up on the same

host could bring the disk subsystem to its knees. If you're using VMware's Distributed Resource Scheduler (DRS) to balance workloads among hosts, keep in mind that it doesn't take VM disk I/O usage into account, it only balances based on CPU and memory usage. To compensate for that, use DRS rules that will always keep specific virtual machines on different hosts.

2. Avoid intense disk I/O. Certain scenarios with your VMs may create periods of very intense disk I/O, which can create such high resource contention that all of your VMs will slow to a crawl. For virtual desktops this can be caused by time-specific events, like all of your users turning on their desktops at approximately the same

Even if you have a good idea of your virtual machine's disk I/O workloads, it's still a good idea to use performance monitoring tools to get detailed statistics such as average and peak usage.

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources

time each morning—often referred to as a boot storm. While that kind of situation may be unavoidable, there are ways to deal with it, such as by adding large cache controllers like NetApp’s Flash Cache (formerly called Performance Acceleration Module or PAM) to your storage device, or by using automated storage tiering technologies that can leverage faster storage devices like solid-state drives during periods of high disk I/O.

Other scenarios—like virtual machine backup windows and scheduled VM activities such as antivirus scans or patching—are controllable. Having concurrent backups running on multiple VMs on a host or data store can cause high disk I/O that will impact the performance of other VMs running on the host or data store. Try to schedule your backups evenly so you don’t have too many occurring simultaneously on the same host or storage resource. You should also consider backup applications that avoid using host resources by accessing the VM data stores directly to back up VM disks. And some specific virtualization disk-to-disk backup products can shorten backup windows and allow tape backups of the disk repositories to occur afterwards without impacting hosts and virtual machines. For scheduled operations like patching and antivirus scanning, enable randomization or create staggered schedules to spread the operations over a period of time so they don’t run simultaneously. You should also be careful when running disk defragmentation operations; defrag generates high disk I/O and can cause thin disks to rapidly increase in size.

3. Use space efficiently. It’s easy to use up disk space with virtual machines, but there are ways to control and limit the amount of space they take up on your storage devices. For virtual desktops or lab-type server environments, using linked clones can save a great deal of disk space. Linked clones are similar to VM snapshots where a virtual machine’s virtual disk file is made read-only and a smaller delta disk is created for any disk writes that may occur. Linked clones work by creating a master virtual disk image that’s read by many VMs but all writes occur on each virtual machine’s own delta disk. For example, if you create 100 VMs with 40 GB virtual disks, they would consume 4 TB of disk space without linked clones. If you used linked clones, however, you would have a single 40 GB virtual disk for all VMs to read from and smaller 1 GB to 2 GB virtual disks for writes—a huge space savings.

Thin provisioning can also help save space. It can be implemented at the virtualization layer or the storage layer. Almost all VMs are given more disk space than they usually need; thin provisioning allows you to over-provision storage by allowing virtual disk files to only take up the space they’re actually using and not the full disk space they were allocated. The use of thin provisioning can greatly reduce the amount of disk space your virtual machines consume and will give you more control over costly storage capacity upgrades.

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources

4. Avoid unnecessary I/O operations. Why generate excessive disk I/O if you don't have to? You should always try to limit the amount of disk I/O that virtual servers and virtual desktops create. This includes disabling any Windows services that aren't needed, uninstalling unnecessary applications, disabling file indexing, and limiting the amount of logging that both the operating system and applications generate. There are many other smaller things that can be tweaked and they can add up to greatly reduced disk I/O across your VMs. You can use end-point management tools or Active Directory group policy to help manage and control the configurations. You'll not only reduce virtual machine disk I/O, you'll reduce consumption of other host resources.

Reducing the amount of unnecessary disk I/O that VMs generate is always a smart move as it allows your storage subsystem to operate at maximum efficiency.

Why generate excessive disk I/O if you don't have to? You should always try to limit the amount of disk I/O that virtual servers and virtual desktops create.

5. Use the right storage for your workloads. Most hosts have local storage available in addition to being connected to shared storage for virtual machines. The types of storage available to your hosts will often have different performance characteristics, such as an 8 Gb Fibre Channel SAN and a 1 Gb iSCSI or NFS storage device.

Besides different storage protocols, you may have hard drives with different speeds (e.g., 10K rpm, 15K rpm) and interfaces (e.g., SAS, SATA, solid state). With so many different storage options to choose from, it makes sense to fit the VM to the right type of storage. Place less-critical VMs on the slower storage tiers and your more critical VMs with higher I/O requirements on the faster tiers. You can also use an automated storage tiering system like Compellent Technologies Inc.'s Fluid Data architecture or EMC Corp.'s Fast technology that moves data between storage tiers based on demand.

You can go a step further by splitting a VM into multiple disk partitions whose virtual disk files reside on multiple storage tiers according to their performance needs. One common way to do this is to create separate disk partitions for the operating system, Windows pagefile, applications and data. The faster storage tiers can be used for the data's higher I/O requirements, while slower tiers can be used for everything else. Even if you don't do that, you can still specify slower or local storage for the large virtual machine virtual swap file created for each VM and used when a host exhausts its physical memory. This also helps ensure that your VM uses less disk space on the more expensive storage tiers.

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources

6. Don't forget to monitor. People usually pay attention to storage statistics when problems occur, but data storage requires attention on a continuous basis. If you don't monitor your storage performance on an ongoing basis you might not know of potential problems or bottlenecks, or be able to spot trends or patterns that may allow you to act proactively. It's particularly important when using network-based iSCSI and NFS storage because network health can impact storage performance. Storage performance should therefore be monitored at both the virtualization layer and storage layer, as a problem may be visible from one viewpoint but not the other. Monitoring a virtual environment is not as simple as monitoring a physical environment. Products designed for virtual environments that monitor end-user or application experiences can help pinpoint exactly which resource or component may be causing a bottleneck.

7. Watch out for storage threats that can grow. Virtual machine snapshots and thin provisioned virtual disks represent a double threat as they have the potential to consume all of the disk space on your VM data stores, which can potentially crash or shut down your running VMs. If you plan to overcommit your storage using thin disks, you need to closely monitor their growth. Don't rely completely on thin disks to address disk space woes; try rightsizing VM disks when you create them and don't give them a lot more disk than they need.

Snapshots are an even bigger threat, as VMs can have multiple snapshots with their combined space much larger than the original virtual disk file size. While VM snapshots can be a handy tool, you should never use them in lieu of traditional backups. Not only do snapshots take up additional space, they can reduce data storage performance. That's especially true when you delete a snapshot and the delta disks are written back to the original disks causing intense disk I/O while the operation is occurring. For very large snapshots, try scheduling their deletion during off hours when the extra I/O will have less of an impact.

Don't rely on manual methods to monitor thin provisioning and snapshots. For thin disks, set alarms for specific overcommitment percentages so you'll know when your data stores are becoming overcommitted. For snapshots, use alarms to look for snapshots that grow beyond a certain size. You should also use alarms to monitor data store free space to alert you when space is low. Be sure to set your alarms high enough, as thin disks and snapshots can sometimes grow very quickly and there might not be much time to respond. Don't rely completely on alarms for snapshots; use regular reporting tools to identify snapshots so they don't run longer than needed.

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources

8. Integrate server virtualization with storage management. More and more storage vendors are integrating server virtualization and storage so they can be managed and monitored using a single console. Examples include plug-ins developed for VMware vCenter Server from NetApp (Virtual Storage Console) and EMC (Virtual Storage Integrator) that allow storage arrays to be managed from vCenter Server.

This type of integration allows for much simpler management of the storage devices used by virtual hosts because monitoring, provisioning, replication and other storage operations can be done directly from vCenter Server.

Storage vendors are also leveraging virtualization APIs to provide very tight integration between the storage layer and the virtualization layer. Using the VMware vStorage APIs, storage tasks traditionally handled by the virtual host (e.g., block zeroing, Storage vMotion) can be offloaded to the storage array, thereby freeing up

host server resources. The APIs also provide more intelligent multipathing to achieve better I/O throughput and failover, and offer replication integration for products like VMware's vCenter Site Recovery Manager.

More and more storage vendors are integrating server virtualization and storage so they can be managed and monitored using a single console.

9. Traditional methods might not cut it. Moving from a physical to a virtual environment also requires a change in thinking. Things like backups, server provisioning, monitoring and management are all very different once servers are virtualized. Applications written specifically to monitor and manage physical environments typically aren't effective in virtual environments because they're not aware of the virtualization layer between the server hardware and the guest operating system.

With backups, for example, it's not efficient to back up servers through the OS layer on virtual hosts. Instead, most virtualization-aware backup apps go directly to the virtualization layer, which is quicker and more efficient. Performance monitoring is another example: If you monitor using OS tools that aren't aware of the virtualization layer, the results will often be inaccurate as the OS tools don't have direct access to the underlying host hardware.

10. Prioritize storage traffic. Hosts with many virtual machines running on them can be like the Wild West with all of the VMs fighting for the host's limited resources. You can end up with less-critical VMs impacting the resources of critical virtual machines and

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources

the resources available for host operations. To prevent this kind of contention, consider using storage I/O controls that can provide a Quality of Service (QoS) level for certain critical host functions and VMs. VMware's vSphere 4.1 introduced a new feature called Storage I/O Control (SIOC) that works by measuring storage latency; when a set congestion threshold is reached for at least four seconds, it enforces configurable I/O shares on VMs to ensure the highest-priority virtual machines get the I/O resources they need. SIOC should help restore some order on busy hosts and allow VMs to coexist peacefully by making it less likely that a few rogue VMs will drag down your critical virtual machines. ☉

Eric Siebert is an IT industry veteran with more than 25 years of experience who now focuses on server administration and virtualization. He's the author of *VMware VI3 Implementation and Administration* (Prentice Hall, 2009) and *Maximum vSphere* (Prentice Hall, 2010).

Ten things to
know about
storage for VMs

iSCSI and
vSphere:
A good match?

Choosing
NAS for virtual
machines

Sponsor
resources

Using iSCSI storage with vSphere

To realize the greatest benefits of a vSphere installation, you need networked storage. iSCSI is a good fit for vSphere; here's how to make it work. *By Eric Siebert*

TO TAP INTO SOME of VMware vSphere's advanced features such as vMotion, fault tolerance, high availability and the VMware Distributed Resource Scheduler, you need to have shared storage for all of your hosts. vSphere's proprietary VMFS file system uses a special locking mechanism to allow multiple hosts to connect to the same shared storage volumes and the virtual machines (VMs) on them. Traditionally, this meant you had to implement an expensive Fibre Channel SAN infrastructure, but iSCSI and NFS network storage are now more affordable alternatives.

Focusing on iSCSI, we'll describe how to set it up and configure it properly for vSphere hosts, as well as provide some tips and best practices for using iSCSI storage with vSphere. In addition, we've included the results of a performance benchmarking test for the iSCSI/vSphere pairing, with performance comparisons of the various configurations.

VMware warms up to iSCSI

iSCSI networked storage was first supported by VMware with ESX 3.0. It works by using a client called an initiator to send SCSI commands over a LAN to SCSI devices (targets) located on a remote storage device. Because iSCSI uses traditional networking components and the TCP/IP protocol, it doesn't require special cables and switches as Fibre Channel does.

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources

iSCSI initiators can be software based or hardware based. Software initiators use device drivers that are built into the VMkernel to use Ethernet network adapters and protocols to write to a remote iSCSI target. Some characteristics of software initiators are:

- Use Ethernet network interface cards (NICs) and native VMkernel iSCSI stack
- Good choice for blade servers and servers with limited expansion slots
- Cheaper than using hardware initiators
- Can be CPU-intensive due to the additional overhead of protocol processing
- ESX server can't boot from a software-based initiator; ESXi can by using iSCSI Boot Firmware Table (iBFT)

Hardware initiators use a dedicated iSCSI host bus adapter (HBA) that includes a network adapter, a TCP/IP offload engine (TOE) and a SCSI adapter to help improve the performance of the host server. Characteristics of hardware initiators include:

- Moderately better I/O performance than software initiators
- Uses less ESX server host resources, especially CPU
- ESX server is able to boot from a hardware initiator

iSCSI is a good alternative to using Fibre Channel storage as it will likely be cheaper to implement while providing very good performance. vSphere now supports 10 Gbps Ethernet, which provides a big performance boost over 1 Gbps Ethernet. The biggest risks in using iSCSI are the CPU overhead from software initiators, which can be offset by using hardware initiators, and a more fragile and volatile network infrastructure that can be mitigated by completely isolating iSCSI traffic from other network traffic.

For vSphere, VMware rewrote the entire iSCSI software initiator stack to make more efficient use of CPU cycles; this resulted in significant efficiency and throughput improvements compared to VMware Infrastructure 3. Those results were achieved by enhancing the VMkernel efficiency. Support was also added for the bidirectional Challenge-Handshake Authentication Protocol (CHAP), which provides better security by requiring both the initiator and target to authenticate with each other.

Planning an iSCSI/vSphere implementation

You'll have to make a number of decisions when planning to use iSCSI storage with vSphere. Let's first consider iSCSI storage devices.

You can pretty much use any type of iSCSI storage device with vSphere because the hosts connect to it using standard network adapters, initiators and protocols. But you need to be aware of two things. First, vSphere officially supports only specific models of vendor iSCSI storage devices (listed on the

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources

vSphere Hardware Compatibility Guide). That means if you call VMware about a problem and it's related to the storage device, they may ask you to call the storage vendor for support. The second thing to be aware of is that not all iSCSI devices are equal in performance; generally, the more performance you need, the more it'll cost you. So make sure you choose your iSCSI device

carefully so that it matches the disk I/O requirements of the applications running on the VMs that will be using it.

There are also some network considerations. For optimum iSCSI performance, it's best to create an isolated network. This ensures that no other traffic will interfere with the iSCSI traffic, and also helps protect and secure it. Don't even think of using 100 Mbps NICs with iSCSI; it'll be so painfully slow that it will be unusable for virtual machines. At a minimum, you should use 1 Gbps NICs, and go for 10 Gbps NICs if that's within your budget. If you're concerned about host server resource overhead, consider using hardware initiators (TOE adapters). If you opt for TOE adapters,

make sure they're on VMware's Hardware Compatibility Guide. If you use one that's not supported, there's a good chance vSphere will see it as a standard NIC and you'll lose the TOE benefits. Finally, use multi-pathing for maximum reliability; you should use at least two NICs (not bridged/multi-port) connected to two different physical network switches, just as you would when configuring Fibre Channel storage.

For optimum iSCSI performance, it's best to create an isolated network. This ensures that no other traffic will interfere with the iSCSI traffic, and also helps protect and secure it.

Configuring iSCSI in vSphere

Once your iSCSI environment is set up, you can configure it in vSphere. The method for doing this will differ depending on whether you're using software or hardware initiators. We'll cover the software initiator method first.

Configuring with software initiators: Software initiators for iSCSI are built into vSphere as a storage adapter; however, to use them you must first configure a VMkernel port group on one of your virtual switches (vSwitches). The software iSCSI networking for vSphere leverages the VMkernel interface to connect to iSCSI targets, and all network traffic between the host and target occurs over the NICs assigned to the vSwitch the VMkernel interface is located on. You can have more than one VMkernel interface on a single vSwitch or multiple vSwitches. The VMkernel interface is also used for vMotion, fault-tolerance logging traffic and connections to NFS storage devices. While you can use one VMkernel interface for multiple things, it's highly recommended to create a separate vSwitch and VMkernel interface exclusively for

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources

iSCSI connections. You should also have two NICs attached to the vSwitch for failover and multi-pathing. If you have multiple NICs and VMkernel interfaces, you should make sure you bind the iSCSI VMkernel interfaces to the correct NICs. (See VMware's iSCSI SAN Configuration Guide for more information.)

Once the vSwitch and VMkernel interface is configured, you can configure the software iSCSI adapter. Select Configuration/Storage Adapters in the vSphere Client to see the software iSCSI adapter listed; select it and click Properties to configure it. On the General tab, you can enable the adapter and configure CHAP authentication (highly recommended). On the Dynamic Discovery tab, you can add IP addresses to have iSCSI targets automatically discovered; optionally, you can use the Static Discovery tab to manually enter target names. After entering this information, go back to the Storage Adapters screen and click on the Rescan button to scan the device and find any iSCSI targets.

Configuring with hardware initiators: The process is similar for hardware initiators, but they don't use the VMkernel networking, so that step can be skipped. TOE adapters are technically network adapters, but they'll show up on the Storage Adapters screen instead. Select them, click Properties and configure them in a manner similar to software initiators by entering the appropriate information on the General, Dynamic Discovery and Static Discovery tabs. You'll need to assign IP addresses to the TOEs on the General screen as they don't rely on the VMkernel networking.

Once the initiators are set up and your iSCSI disk targets have been discovered, you can add them to your hosts as VMFS volumes. Select a host, click on the Configuration tab and choose Storage. Click Add Storage and a wizard will launch; for the disk type select Disk/LUN, which is for block-based storage devices. (The Network File System type is used for adding file-based NFS disk storage devices.) Select your iSCSI target from the list of available disks, give it a name and then choose a block size. When you finish, the new VMFS data store will be created and ready to use.

Best practices for using iSCSI storage with vSphere

Once iSCSI disks have been configured, they're ready to be used by virtual machines. The best practices listed here should help you get the maximum performance and reliability out of your iSCSI data stores.

- The performance of iSCSI storage is highly dependent on network health and utilization. For best results, always **isolate your iSCSI traffic** onto its own dedicated network.
- You can configure only one software initiator on an ESX Server host. When configuring a vSwitch that will provide iSCSI connectivity, **use multiple physical NICs** to provide redundancy. Make sure you bind the VMkernel interfaces to the NICs in the vSwitch so multi-pathing is configured properly.

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources

- Ensure the NICs used in your iSCSI vSwitch connect to **separate network switches** to eliminate single points of failure.
- vSphere supports the use of **jumbo frames** with storage protocols, but they're only beneficial for very specific workloads with very large I/O sizes. Also, your back-end storage must be able to handle the increased throughput by having a large number (15+) of spindles in your RAID group or you'll see no benefit. If your I/O sizes are smaller and your storage is spindle-bound, you'll see little or no increase in performance using jumbo frames. Jumbo frames can actually decrease performance in some cases, so you should perform benchmark tests before and after enabling jumbo frames to see their effect. Every end-to-end component must support and be configured for jumbo frames, including physical NICs and network switches, vSwitches, VMkernel ports and iSCSI targets. If any one component isn't configured for jumbo frames, they won't work.
- Use the new **Paravirtual SCSI (PVSCSI) adapter** for your virtual machine disk controllers as it offers maximum throughput and performance over the standard LSI Logic and BusLogic adapters in most cases. For very low I/O workloads, the LSI Logic adapter works best.
- To set up **advanced multi-pathing** for best performance, select Properties for the iSCSI storage volume and click on Manage Paths. You can configure the Path Selection Policies using the native VMware multi-pathing or third-party multi-pathing plug-ins if available. When using software initiators, create two VMkernel interfaces on a vSwitch; assign one physical NIC to each as Active and the other as Unused; use the esxcli command to bind one VMkernel port to the first NIC and the second VMkernel port to the second NIC. Using Round Robin instead of Fixed or Most Recently Used (MRU) will usually provide better performance. Avoid using Round Robin if you're running Microsoft Cluster Server on your virtual machines.

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources

iSCSI guides available

VMware provides detailed guides for implementing iSCSI storage for vSphere. Two useful guides available from VMware include the iSCSI SAN Configuration Guide and the iSCSI Design Considerations and Deployment Guide. ☉

Eric Siebert is an IT industry veteran with more than 25 years of experience who now focuses on server administration and virtualization. He's the author of *VMware VI3 Implementation and Administration* (Prentice Hall, 2009).

USING NAS FOR VIRTUAL MACHINES

Ten things to
know about
storage for VMs

iSCSI and
vSphere:
A good match?

Choosing
NAS for virtual
machines

Sponsor
resources

Common wisdom says you need block storage for virtual servers, but with most hypervisors supporting the NFS protocol, NAS may work just as well. By Eric Siebert

SHARED STORAGE is a requisite for virtualized servers if you want to use any of the advanced features server virtualization offers, such as high availability or the ability to move a running virtual machine (VM) from one host to another. This typically meant you had to invest in an expensive Fibre Channel SAN (FC SAN). But all server virtualization products also support using network-attached storage (NAS) devices, which can provide a worthy, cost-effective alternative to FC SANs for shared storage.

Another alternative is iSCSI storage which, like NAS, uses TCP/IP over a standard Ethernet network, but iSCSI is block storage like Fibre Channel and tends to be costlier than NAS. NAS systems generally support both NFS and CIFS file-sharing protocols, but server virtualization products prefer—or are limited to—NFS.

Inside NFS

NFS was developed in the late 1980s and has been revised several times over the years; NFS Version 4 (NFSv4) is the most recent version. The NFS architecture consists mainly of three components:

- Remote procedure calls (RPCs)
- External data representation (XDR)
- NFS procedures

The NFS protocol uses an RPC system that allows a server (NFS client) to make a call that's executed on another server (NFS server). XDR is the data encoding standard for NFS and serves as the universal language used between clients and servers. NFS procedures are software instructions used to perform storage-related tasks.

An NFS server may be a dedicated NAS appliance such as those sold by NetApp and all major storage vendors, or it can be a common server running an operating system. NFS is commonly used in Unix and Linux systems, but is also available with other operating systems such as Windows. NFS is a stateless protocol, which means the server doesn't store any client information and each RPC event contains all the necessary information to complete the call. In this manner, no open connections between the client and server exist, and crash recovery is as simple as having the client resend requests until the server responds.

NFS in server virtualization

NFS has become increasingly popular for shared storage that's to be shared by multiple virtual hosts. All major server virtualization platforms support the use of NAS NFS storage devices for virtual machines. Because NFS is a widely supported protocol, there are many different options for using NFS storage with your virtual hosts. This can range from taking a standard physical server and converting it into an NFS server, using virtual SAN software or using a dedicated storage appliance. The cost and performance characteristics with each option can vary greatly, but dedicated appliances offer the best performance although at a higher cost. An inexpensive NFS server can be built by putting a bunch of disks in a standard physical server and then loading an operating system like Linux or Windows that has a NFS server, or by using a dedicated storage appliance application like Openfiler open-source shared storage software.

Almost every data storage vendor offers a storage device that supports NFS, including "low-end" devices that support NFS from vendors like NetGear Inc. and Synology Inc. Many storage devices will support both iSCSI and NFS, but allocating storage for iSCSI datastores will consume the full space right away, while with NFS it grows as data is written to it. But with so many devices to choose from, you can easily find a good NFS storage system that will meet

your requirements regardless of your budget.

Because NFS is a file-level protocol, it's configured in a different manner than block storage devices. With block storage devices you have a storage I/O adapter in the host that communicates with the storage device either locally or remotely. This would typically be a SCSI or Fibre Channel adapter, or with iSCSI, a network adapter that serves as either a hardware or software initiator. With NFS you use an NFS client built into the hypervisor that uses a network adapter in the host to communicate with the NFS server. Instead of scanning for storage devices on your I/O adapters as you would with block devices, you simply enter an NFS server name and folder location when adding an NFS storage device to a virtual host. Once you have your NFS data-stores configured, you create virtual machines on them just like you would with block storage devices.

With block storage devices you have a storage I/O adapter in the host that communicates with the storage device either locally or remotely.

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources

The pros and cons of using NAS

For the most part, NAS storage devices in a virtualized server environment function similarly to block storage devices, but there may be some limitations due to their architecture.

- If you don't use local storage on your virtual host and want to boot directly from a shared storage device, you'll need a storage resource other than a NAS system. With Fibre Channel and iSCSI adapters you can boot the hypervisor directly from a shared storage device without using any local storage.
- NFS uses a software client built into the hypervisor instead of a hardware I/O adapter. Because of that, there's CPU overhead as the hypervisor must use a software client to communicate with the NFS server. On a very busy host this can cause degradation in performance as the CPUs are also being shared by the virtual machines.
- In vSphere environments, while you can create VM datastores on NFS devices, they don't use the high-performance VMFS file system. While this doesn't affect the use of most of vSphere's features, you can't use raw device mappings (RDMs) to attach a physical disk directly to a virtual machine.
- Some vendors don't recommend NFS storage for certain sensitive transactional apps (e.g., Exchange and Domino) due to latency that

can occur. But there are many factors that figure into this, such as host resources/configuration and the performance of the NFS device you're using. This shouldn't be a problem for a properly sized NFS system.

- NFS doesn't support using multipathing from a host to an NFS server. Only a single TCP session will be opened to an NFS datastore, which can limit its performance. This can be alleviated by using multiple smaller datastores instead of a few larger datastores, or by using 10 Gb Ethernet (10 GbE) where the available throughput from a single session will be much greater. The multipathing constraint doesn't affect high availability, which can still be achieved using multiple NICs in a virtual switch.

Despite the limitations, there are some good reasons why you might prefer a NAS system over block storage devices.

- Many NFS storage devices use thin provisioning by default, which can help conserve disk space because virtual disks don't consume the full amount of space they've been allocated.
- File locking and queuing are handled by the NFS device, which can result in better performance vs. iSCSI/FC where locking and queuing are handled by the host server.
- NFS doesn't have a single disk I/O queue like a block storage device has, so you may get better performance. The performance of NFS is based on the size of the network connection and the capabilities of the disk array.
- Implementing NAS costs a lot less than traditional FC storage. NAS devices require only common NICs instead of expensive HBAs, and use traditional network components rather than expensive FC switches and cables.
- Because NAS takes away a lot of the complexity of managing shared storage, specialized storage administrators aren't necessary in most cases. Managing files on an NFS server is much easier than managing LUNs on a SAN.
- Virtual datastores can be expanded easily by simply increasing the disk on the NFS server; there's no need to increase the size of datastores as they'll automatically increase accordingly.
- Operations like snapshots and cloning are done at the file system level instead of at the LUN level, which can offer greater flexibility and more granular support.

The advantages to using NAS are many and you shouldn't be discouraged by the disadvantages that mainly apply to specific circumstances or with lower quality NAS products. With a properly sized and designed system that

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources

will handle the VM workloads on your hosts, NAS can be as good a choice as any block storage device.

Is NAS performance enough?

Many IT shops considering NAS as an alternative to block storage for their virtual servers are concerned about performance, and with good reason. In most cases, NAS performance won't equal that of an FC SAN, but a properly architected NFS solution can easily meet the performance needs of most workloads.

Some users end up comparing iSCSI to NAS as they're both low-cost alternatives to FC storage and they can each use existing Ethernet infrastructure. VMware Inc. has published test results comparing the performance of virtual machines on NAS, iSCSI and FC storage devices. The results show that the performance of NAS vs. both hardware and software iSCSI is nearly identical. As long as the CPU doesn't become a bottleneck, the maximum throughput of both iSCSI and NFS is limited by the available network band-

Some users end up comparing iSCSI to NAS as they're both low-cost alternatives to FC storage and they can each use existing Ethernet infrastructure.

width. Software iSCSI and NFS are both more efficient than Fibre Channel and hardware iSCSI at writing smaller block sizes (fewer than 16 KB), but with larger blocks more CPU cycles are used, which makes software iSCSI and NFS less efficient than hardware iSCSI and Fibre Channel. The CPU cost per I/O is greatest with NFS; it's only slightly higher than iSCSI, but much higher than hardware iSCSI and FC, but on a host with enough spare CPU capacity this shouldn't be an issue.

Achieving the best performance with NAS comes down to several factors; the first is having enough CPU resources available so the CPU never becomes a bottleneck to NFS protocol processing. It's easy enough to achieve by simply making sure you don't completely overload your virtual host's CPU with too many virtual machines. Unfortunately, there's no way to prioritize or reserve CPU resources for NFS protocol processing, so you need to make sure you adjust your workloads on your hosts accordingly and monitor CPU usage. Using a technology like VMware's Distributed Resource Scheduler will help balance CPU workloads evenly across hosts.

The second factor is network architecture; the performance of NAS storage is highly dependent on network health and utilization. You should isolate your NAS traffic on dedicated physical NICs that aren't shared with virtual machines. You should also ensure that you use a physically isolated storage network that's dedicated to your hosts and NFS servers, and isn't

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources

shared with any other network traffic. Your NICs are your speed limit; 1 Gbps NICs are adequate for most purposes, but to take NFS to the next level and experience the best possible performance, 10 Gbps is the ticket. There are a number of network configuration tweaks you can use to boost

performance, as well as technology like jumbo frames.

The final factor in NFS performance is the type of NAS storage device you're connected to. Just like any storage device, you must size your NAS systems to meet the storage I/O demands of your virtual machines. Don't use an old physical server running a Windows NFS server and expect to meet the workload demands of many busy virtual machines. Generally,

the more money you put into a NAS product the better performance you'll get. There are many high-end NAS systems available that will meet the demands of most workloads.

Just like any storage device, you must size your NAS systems to meet the storage I/O demands of your virtual machines.

NAS has its niche

NAS might not be appropriate for every virtualized server environment—for certain workloads only a FC SAN will do—but it's certainly attractive and effective for most use cases. In past years, NAS wasn't a viable alternative because of limited support by virtualization vendors, but that has changed and NFS is now fully supported. NFS has also matured and improved in all areas, including in the hypervisor, on the network and in the storage device to become a solid storage platform for virtualization. ☺

Eric Siebert is an IT industry veteran with more than 25 years of experience who now focuses on server administration and virtualization. He's the author of *VMware VI3 Implementation and Administration* (Prentice Hall, 2009) and *Maximum vSphere* (Prentice Hall, 2010).

Ten things to know about storage for VMs

iSCSI and vSphere: A good match?

Choosing NAS for virtual machines

Sponsor resources