TechTarget SearchServer Virtualization

E-guide

PRO+ Content

Essential Guide to Choosing Virtualization Hardware

Introduction

To say that virtualization changed the way businesses use and purchase hardware is an understatement. Simply spinning up a new VM on existing hardware is worlds easier than purchasing and provisioning a physical server to match a new workload. However, even though vitalization has allowed us to abstract workloads from the underlying hardware, those physical resources are still the backbone of any data center. Without a solid foundation of hardware resources, bottlenecks and resource contention problems can cripple a data center. The ease with which admins can provision a new VM can lead to VM sprawl, causing major networking headaches as admins struggle to track networked devices and avoid network bottlenecks.

Modern servers are shipping with massive amounts of memory, multiple network interface cards and support for solid-state storage. With all the options available, it's hard to know what you need. This guide on choosing the best hardware for virtualization can help.

CPU and RAM

Processor and memory requirements

Choosing the best hardware for virtualization begins with a server's memory and computing resources. More often than not, memory will be the limiting factor in the number of virtual machines a server can host. And, of course, a shortage in either RAM or processing power can directly affect performance. In many cases, organizations choose to repurpose existing hardware for virtualization, but do you know what to look for when it's time to replace aging servers?

Selecting CPU, processors and memory for virtualized environments

There are several aspects to choosing CPU and memory for your virtual environment. In the selection process, processor speed, the number of cores needed and the type of DIMMs you use are all critical considerations. An expert outlines what you should look for.

In the first part of this series on hardware selection for virtualized environments, we discussed the choice between blade and rackmount servers.

Now, this tip delves further into the hardware selection process and covers hardware purchasing considerations for CPUs, processors and memory.

Selecting a CPU for a virtualization deployment

When you purchase a CPU, the first decision is which brand: AMD or Intel? Over the years, many performance studies have compared the two. With constant changes in processor architecture, AMD is sometimes ahead of Intel, then vice versa. Both Intel and AMD have integrated virtualization extensions, Intel Virtualization Technology (Intel VT) and AMD Virtualization (AMD V), respectively, into their latest processors in an attempt to speed up instruction execution in virtual servers.

The major difference between Intel and AMD processors is physical architecture. Intel uses a front-side bus model to connect processors to the memory controller, while AMD uses an integrated memory controller for each processor that is interconnected though a hyper-transport. And depending on the processor family, these processors tend to have different power consumption levels.

In terms of performance, when you compare processors from each vendor with similar speed, features and number of cores, Intel and AMD tend to be equal. Some performance studies show that Intel processors have an edge in performance, and others show the inverse. Both Intel and AMD processors work well in VMware ESX hosts, so it's a matter of brand preference when it comes to choosing one. Because Intel and AMD continually release new processor families, you should check which currently has the latest technology as you make a decision between the two.

So which CPU should you choose? In general it's a good idea to choose a brand and stick with it, especially if the majority of your current servers already use a particular brand. The reason for this is that you can't move running virtual machines (VMs) from one host to another if the hosts run on different processors (though see AMD demos live migration between Intel and AMD processors. For example, a VM started on a host with an Intel processor typically crashes if it's moved while running to a host with an AMD processor. If you decide to use different brands, it's best to isolate hosts of the same brand processor into separate clusters for compatibility purposes.

Selecting a processor: Virtualization extensions

When purchasing a processor, choose a model that's optimized for virtualization, such as those with AMD-V or Intel-VT extensions. To grasp why extensions are important, you need to understand how rings work on a CPU.

X86 operating systems use protection rings that provide levels of protection where code can execute jobs. These rings are arranged in a hierarchy, from the most privileged (Ring 0) to the least privileged (Ring 3), and are enforced by the CPU that places restrictions on processes. On nonvirtualized servers, an OS resides in Ring 0 and owns the server hardware and applications that run in

Ring 3. On virtualized systems, a hypervisor or virtual machine monitor (VMM) needs to run in Ring 0 so a VM guest OS is forced into Ring 1 instead. Since most OSes have to run in Ring 0, the VMM fools a guest OS into thinking that it's running in Ring 0 by trapping privileged instructions and emulating Ring 0 to the guest VM.

Unfortunately, this operation can reduce performance, which forced Intel and AMD to develop Intel VT and AMD-V extensions to solve the problem. Both sets of extensions are integrated into their CPUs so a VMM can instead run in a new ring called Ring -1 (minus 1), which allows guest OSes to run natively in Ring 0. These extensions to the CPU enable better performance. VMM no longer needs to fool a VM guest OS into thinking that it's running in Ring 0 because it can operate there and not conflict with VMM, which has moved to the new Ring 1 level. To get the best performance from your virtual hosts, choose a CPU that uses these virtualization optimized extensions.

Also, stay tuned for future processor releases from AMD and Intel that support the new Nested Page Tables (NPT). AMD's version is Rapid Virtualization Indexing (RVI); Intel's is Extended Page Tables (EPT). This new CPU technology helps reduce the performance overhead of virtualizing large applications such as databases.

Selecting multicore CPUs

Another critical choice is the number of physical CPUs (sockets) and the number of cores that a CPU should have. A multicore CPU combines multiple independent cores on a single physical CPU; essentially turning a single physical CPU into multiple CPUs. An example of this is a server with two quad-core CPUs that has eight processors available for use. Depending on the CPU brand and model, these cores sometimes share a single cache, or have separate Level 2 caches for each core. Most virtualization software vendors sell licenses by the socket and not by the number of cores that the socket has, so multicore processors are fantastic for virtualization. For new servers, multicore CPUs are now virtually standard.

You also have to decide between dual- and quad-core CPUs. You may be tempted to choose quad-core over dual-core based on an assumption of the more cores, the better. But dual- and quad-core CPUs feature crucial differences. CPU core increases do not follow the same scale as CPU clock-speed increases.

A 3.2 GHz CPU is twice as fast as a 1.6 GHz CPU, but a quad-core CPU is not four times faster than a single core. A dual-core CPU is roughly 50% faster than a single-core CPU (not 100%, as you might expect), and a quad-core CPU is only about 25% faster than a dual-core CPU. In addition, dual-core CPUs typically have higher clock speeds than do quad-core CPUs. Quad-core CPUs generate excessive heat and, as a result, cannot be clocked as high as single- and dual-core CPUs.

In this e-guide

In general quad-core CPUs are recommended for virtual hosts for two reasons. The first is that most virtualization software is licensed by the number of sockets in a server and not the number of cores you have. This means you get more CPUs per license than you purchase. The second reason is that having more cores in a host server gives the hypervisor CPU scheduler more flexibility when trying to schedule CPU requests that are made by VMs. Having more cores available makes a CPU scheduler's job easier and improves VMs' performance on a host.

In some situations, however, dual-core rather than quad-core CPUs are preferable (if you don't plan on running more then six to eight VMs on hosts, for example). The faster CPU MHz of dual-core hosts increases speed to the VMs running on it. In addition, if you plan on assigning VMs a single virtual processor, dual-core processors can be a better option, because single vCPU VMs are easier for the hypervisor to schedule than multiple vCPU VMs.

Choosing memory for virtualization

You don't want to skimp on memory, because it's often a host's first hardware resource to get used up. Not having enough memory on a host while having plenty of other resources available (CPU, disk, network, etc.) will limit the number of VMs that can put on a host. While a memory overcommit functionality is available with some virtualization software, it isn't recommended to exhaust all physical host memory because VMs' performance will suffer.

A server's memory type is dictated by what a server supports, so check server specifications or use the online purchasing guides to see what is available. Check on how many memory slots are in your server and whether memory needs to be installed in pairs.

As different size dual in-line memory modules (DIMMs) can be used in servers (e.g. 512 MB, 1 GB, 2 GB, etc) you should choose a DIMM size that works with the amount of memory a server needs. Larger-size memory DIMMs (e.g. 4 GB or 8 GB) are more expensive than smaller sizes, but they use fewer memory slots to leave more room for future expansion. Once you choose a DIMM size, stick with it. Mixed DIMM sizes in your server can cause reduced performance. For best results, use the maximum-size DIMM available in memory slots.

In addition to sizes, there are also many different memory types (e.g. PC2100, PC5300) which are based on the memory module's peak data transfer rate. Originally, the number after "PC" that was used to label memory modules stood for the clock rate of the data transfer (e.g. PC133). This was later changed to the peak data transfer rate in Mbps, so memory that is classified as PC5300 has a peak data transfer rate of 5300 Mbps. Most servers can use several different memory types, so choosing the fastest memory available is desirable if you can afford it.

The final memory-related decision that you'll have to make is between single, dual or quad-rank DIMMs. A memory rank is defined as a block of 64 bits, or 72 bits for error-correcting code (ECC) memory, created by using the DRAM chips on a DIMM. For example, single-rank DIMMs combine all chips into a single block, while dual-rank DIMMs split the chips into two blocks. Dual-rank DIMMs improve memory density by placing the components of two single-rank DIMMs in the space of one module, typically making them cheaper than single-rank DIMMs.

Unfortunately, in some instances, a server chipset can support only a certain number of ranks. If a memory bus on a server has four DIMM slots, for example, the chipset may be capable of supporting only two dual-rank DIMMs or four single-rank DIMMs. If two dual-rank DIMMs are installed, then the last two slots should not be populated.

If the total number of ranks in the populated DIMM slots exceeds the maximum number of loads the chipset can support, the server may not operate reliably.

So which DIMM type should you choose? Single-rank DIMMs allow a server to reach its maximum memory capacity and highest performance levels, but they cost more because of higher density. Dual-rank DIMMs are cheaper but can limit overall system capacity and restrict future upgrade options. If you can afford the more expensive single-rank DIMMs opt for them. If you can't, dual-rank DIMMs work equally well. On some servers, you can mix single and dual-rank DIMMs if they're not in the same bank (though this method isn't recommended). And for best results, try to stick with the same rank type in all slots.

Finally, the market features several memory manufacturers, and it's best not to mix brands within a server. Buy OEM memory if that's what is present in a server or replace it all with memory from another vendor. Memory configurations and selections can be complicated so always consult with the server hardware vendor to make sure you've made the right choices for your server.

The next part of this series will cover choosing network and storage I/O adapters, as well as storage options for your virtual host.

Dual-core qua-processor systems, or quad-core dual processor systems?

We are looking at hardware for our new ESX farm and are wondering if we should go dual-core quad-processor systems, or quad-core dual-processor systems. Essentially we will have the same number of cores in either case, but the dual-processor servers are much cheaper. Thoughts?

Ahh, multiple cores, a great thing to happen in virtualization. The reality with the multi-core systems (for sure in the Intel space) is that you are right now trading clock speed for cores. Most quad core systems with a reasonable price are 1.6 or 1.8 GHz, while you can get dual core systems at 3.0 GHz all day long for cheap. Your question comes down to two answers. 1- What is best for the virtualization systems, speed or cores? And what hits the pocket book the least (i.e. gives you a better return on investment (ROI)/lowest cost per VM)?

What is better for the virtualization system (in my opinion) is cores vs. speed. Comparing quad processors to dual processors is another argument. But we'll get there in a minute. The scheduling mechanism in use today allows for only a single VM (virtual machine) to access a core at any given time. Essentially VMs are treated like processes and the host schedules them according to priority. This means that the more cores you have the more VMs you can have simultaneously executing on a processor/core.

Trading clock speed on the processor for more cores is OK in my book. The reason being is that most (90%+) of the x86 servers out there do nothing with the processor. Sure there are spikes, but sustained utilization above 1 or 2 GHz is very rare as a percentage of the environment. Get more cores when possible.

On the money/ROI side, multi-core systems are the best thing to happen to virtualization hosts since… well, since VMware hit the market hard. Essentially you can now get four and eight cores in a dual processor chassis. If you use standard industry consolidation ratios on multiple core systems and get three VMs per core, you have a 12:1 ratio on four cores and a 24:1 ratio for eight. The question then becomes about cost of memory and memory expandability. Can the DP system handle enough memory to support all your VMs, or do you wind up with an eight core system that due to memory limits only handles 12 VMs and the processor remains unused? Best bet: Compare like to like. If comparing the quad to dual, assume the quad will host MORE VMs and therefore you need to add (really double) the memory you need. Run the costs down, and use a standard like 3:1 per core to figure out your cost per VM in different configurations. The cheapest one that provides the same levels of hardware redundancy should win.

Sizing server hardware for virtual machines

How many virtual machines can you put on a server? Depends upon whether your hardware and virtualization environment are well-matched. This tip offers four criteria for selecting the right VM server hardware.

How many virtual machines can you fit on a host server? That's a frequently-asked question when IT pros consider which hardware options to purchase for their virtual hosts. In this tip, I share things I've learned when choosing servers for different types of virtual machines and to fit the current and future needs of a virtual environment.

You may be able to fit as many as 100 VMs on a single host, or as few as two. The types of applications that you are running on your virtual machines will largely dictate how many you can put on your host server: for example, servers that have very light resource requirements, such as web, file and print servers, versus servers with medium to heavy resource requirements, such as SQL and Exchange servers. Overall, one should analyze the performance usage of a current environment to get a better understanding of the virtual environment requirements.

Four criteria for sizing up host servers

There are four major criteria to consider when sizing up virtual server hardware: memory, CPU, network and disk resources. Let's start with memory which is typically used up on host servers first.

Memory

When it comes to figuring out how much RAM to put in a host server, I would recommend installing the maximum amount if possible.

Yet, the opposite mentality should be taken when it comes to allocating memory for virtual servers, by only giving a VM the exact amount of memory it needs. Usually with physical servers, more memory than what is needed is installed and much of it ends up being wasted. With a VM, it is simple to increase the RAM at any time, so start out with the minimum amount of memory that you will think it will need and increase it later if necessary. It is possible to over-commit memory to virtual machines and assign more RAM to them then the physical host actually has. By doing this you run the risk of having your VMs swapping to disk when the host memory is exhausted which can cause decreased performance.

CPU

With the advent of multi-core CPUs, it has become easier and cheaper to increase the number of CPUs in a host server. Nowadays, almost all servers come with two or four cores per physical CPU. A good rule of thumb is that four single CPU VMs can be supported per CPU core. This can vary by as much as 1-2 per core, and up to 8-10 per core based on the average CPU utilization of applications running on VMs.

A common misconception with virtual server hardware is that a VM can utilize as much CPU megahertz as needed from the combined total available. For example, a four CPU, quad core 2.6 GHz would have a combined total of 20,800 megahertz (8 x 2.6 GHz). A single vCPU VM however, can never use more megahertz then the maximum of one CPU/core. If a VM has 2 vCPUs, it can never use more megahertz than the maximum of each CPU/core. How many cores needed will also depend on whether multiple vCPU VMs are used or not.

You should always have at least one more core than the maximum number of vCPUs that will be assigned to a single VM. For example, don't buy a two processor dual-core server with a total of four cores and try to run a four vCPU VM on it. The reason being that the CPU scheduler of the hypervisor needs to find 4 free cores simultaneously each time the VM makes a CPU request. If there are only a total of four available cores, the performance will be very slow.

My recommendation would be to use quad core CPUs because more cores provide the CPU scheduler with more flexibility to process requests.

Network

The number of network interface cards (NICs) needed in a virtual server will vary based on how much redundancy is desired, whether or not network storage will be used and which features will be selected. Using 802.1Q VLAN tagging provides the flexibility of using multiple VLANs on a single NIC, thus eliminating the need to have a separate NIC for each VLAN on a host server. For smaller servers, you can get away with using two NICs, but it is best to a have a minimum of four NICs on your host server. If you are using network storage, such as iSCSI, it would be wise to have more than four NICs, especially if you are going to use features like VMware's vMotion. When creating vSwitches it's best to assign multiple NIC's to them for redundancy and increased capacity available to VMs.

Disk

Finally, disk resources need to be evaluated. There are many choices available, and which one you should choose will largely be dictated by your budget and if you have a storage-area network, or SAN, available in your environment. Local disk is the cheapest option, but does not allow for advanced features that require shared storage amongst host servers like vMotion.

SAN (Fibre Channel) disk is typically the best performing disk solution, but usually one of the most expensive. Network disk is a good alternative and has come close to matching SAN performance. Also, using 15K hard drives will give a performance gain when compared to 10K drives, but it is also important to have larger RAID groups available to help spread the disk I/O across as many drive spindles as possible.

When determining how much disk to buy, make sure you have enough available for all your virtual machines to use, plus an additional 10-20% for additional VM files and snapshots. If you plan on making heavy use of snapshots, you may want to allow for even more disk space. In many cases, a combination of disk resources is used with your hosts, for example, storing development and test VMs on local disk, while keeping production VMs on shared storage.

Typically, you want your virtual machines using at least 80% of the capacity of your host server to maximize your investment. However, leave enough spare capacity for future growth and ensure that enough resources will be available to support additional virtual machines in case of a host failure. It is better to have too much capacity than not enough so that you can avoid constraining your resources and prevent the need to purchase additional host servers.

Handle changing server workloads by employing high-performance GPUs

Graphics capabilities have long been an important feature of PCs, but today's data centers are also deploying high-performance GPUs for servers.

As computing needs in the modern data center change, high-performance GPUs have become an important part of server virtualization. Server workloads such as big data analytics employ more powerful data visualization and rendering capabilities to express complex data. Servers also need the graphics capabilities to handle endpoint-type tasks that are now virtualized.

Until recently, server vendors largely overlooked graphics, because such rendering and visualization features were not needed by traditional server workloads such as transactional databases or file and Active Directory servers. System designers opted to forego a graphics processing unit (GPU), lowering the server's cost and minimizing the system's energy demands.

But the push to virtualization, along with increased reliance on multimedia and visualization tools, has prompted businesses to reconsider the need for server-based graphics hardware.

As server technology continues to evolve, vendors have begun to offer servers with GPUs directly incorporated onto the hardware.

Before you deploy high-performance GPUs, however, be sure to plan and test in advance because servers don't always provide the same slot space and power cabling as desktop PCs and workstations.

The role played by high-performance GPUs

A GPU serves the same role on a server that it does on a client-side computer: The GPU offloads an application's graphics instructions from the main processor. This process frees up the main processor for other tasks and executes the application's graphics instructions in hardware, providing for the high level of sophisticated, life-like rendering, video processing and visualization we expect today. Without a GPU, graphics instructions would need software emulation and would tie up the main processor, yielding unacceptable levels of performance.

Application virtualization, for example, might allow a server to host an application shared by multiple users. If that shared application demands graphics functionality, such as a video-rendering tool, then the server must provide that capability. Additionally, virtual desktop infrastructure (VDI) allows endpoints to be hosted on centralized servers.

In this case, 3-D modeling software and other graphics tools that might normally run in a desktop PC would now run in a virtual machine hosted on a server, which also requires the addition of graphics functionality.

Installing GPUs on a virtualized server

High-performance GPUs are typically deployed on traditional servers through a highly specialized graphics adapter card, such as NVIDIA's Tesla, installed in an available PCI Express (PCIe) slot at the server. This is the easiest and most common way to retrofit an existing server with no onboard GPU, but there are still challenges to consider.

These GPU cards are often large, power-hungry devices, but servers typically provide only one or two PCIe slots, one of which may already be filled by another PCIe expansion device, such as a multiport NIC or an I/O accelerator. Even if a suitable expansion slot is free, a GPU card, complete with a large heat sink and fan assembly, simply might not fit in the space available.

You should also keep in mind that a GPU card can require several hundred additional watts of system power. This requirement can cause problems for server platforms designed with small power supply modules for high efficiency and minimum power use.

Some systems may need to upgrade power supplies and provide additional power cables to accommodate the GPU card. The PCIe bus cannot deliver that much power to a load device.

Ultimately, adding an after-market GPU card should always be approached as a proof-of-principle project. IT professionals will need to evaluate GPU card deployment techniques very carefully and verify the server's ability to support a GPU load across a range of operating conditions.

New server designs can, however, incorporate GPUs directly onto the server's motherboard. The SuperServer 1027GR-TRFT from Super Micro Computer Inc. incorporates a Matrox G200eW GPU on board, which gives you the advantage of integration simplicity. The GPU will not need a PCIe slot, and the power supplies will already be sized to run the additional GPU chips.

Software requirements for a server GPU

Graphics platforms are extremely demanding subsystems for any computer, both in terms of physical space and power supplies, but the GPU must also be compatible with the server's operating system.

NVIDIA's Tesla for servers currently only supports 32-bit and 64-bit Linux. Depending on its intended use, the GPU may require driver support from Windows Server 2012, as well as a hypervisor such as vSphere or Hyper-V. In short, there must be some mechanism for CPU cores to share GPUs. This is particularly important to VDI deployments where many desktop instances require graphics functionality.

For decades, server vendors have avoided the use of graphics capabilities, preferring to relegate high-performance rendering and visualization tasks to endpoint systems with their own individual graphics subsystems. As virtualization enables application and endpoint consolidation within the data center, graphics functionality must also shift to the server's hardware. IT professionals, however, will need to take great care to avoid storage or power bottlenecks and compatibility issues when retrofitting enterprise-class GPUs onto current servers.

RISC technology, ARM processors trending in server virtualization

New 64-bit ARM processors are better equipped to support virtualization than before, but not all workloads are good candidates for the technology.

RISC technology has been a staple of mobile devices for decades, but it is now finally poised to take on a serious role in data center servers and server virtualization. The latest RISC processors support virtualization and will change the way computing resources scale to meet workload demands.

Virtualization is tailor-made for complex instruction set computing (CISC), allowing x86 and other CISC processors to reach utilization levels that would have been almost impossible without a hypervisor.

But reduced instruction set computing (RISC) and RISC chips now play a bigger role in servers, thanks in large part to the efforts of ARM and its reference architectures. The advent of 64-bit ARM processors means admins must adapt their virtualization strategies to maximize system utilization and scalability.

Support for virtualization in 64-bit ARM processors

For virtualization to provide businesses with the benefits of optimal resource and hardware performance, servers need to support several critical processor technologies.

First, processors must provide an instruction set that allows hypervisors to manage and allocate computing resources without overlapping resources in ways that would jeopardize workload security.

Early deployments relied on software for this management and allocation, but that software actually impaired performance and seriously limited virtualization on older systems. Conventional x86 processors adopted virtualization extensions, and ARM has followed suit, introducing hardware acceleration for virtualization with ARMv7.

Second, memory is a major concern in virtualized environments. Conventional 32-bit processors can only address up to 4 GB of system RAM, limiting the total number of VMs the system can support.

This limitation prompted the move to 64-bit processors, which can theoretically address 16 exabytes of memory, far more than you can practically implement in a current server system.

Older ARM processors overcame the 32-bit memory limitation by mapping the 32-bit address into a 40-bit address range, called the large physical address extension. But ARM processors must also move to 64-bit designs in order to support data center workloads and run enterprise-class operating systems (OSes). The ARMv8 architecture, such as ARM's Cortex-A50 series, is a true 64-bit design that should accommodate 64-bit addressing.

When to use ARM processors for server virtualization

Intel Xeon, AMD Opteron and other x86 processors evolved to handle a wide variety of instructions and can now contain billions of transistors operating at 2-3 GHz. This huge number of transistors processes the extensive set of complex instructions, but puts pressure on power demands and causes processing efficiency to suffer.

The goal of RISC technology -- via ARM processors -- is improved power and performance through a narrowed instruction set. RISC reduces the number of instructions and thereby reduces the total number of transistors. Fewer transistors demand less power and having shorter pathways through the processor can boost performance. This simpler processor can achieve better performance with clock speeds of 1 GHz or less.

RISC applications and the underlying OS tend to be simpler and more task-specific than workloads deployed on x86 servers. Web server workloads, for example, mainly exchange data and serve Web pages from storage to the network. Such tasks demand a relatively small suite of instructions, making them ideal candidates for RISC and complementary services such as Java and ActiveX. The performance boost from ARM can also speed Web page delivery and enhance the user experience.

By contrast, applications with heavy or diverse processing demands, or those with large or variable memory footprints, are generally not candidates for RISC deployment. This means businesses must match the server to the task.

A RISC server must use ARMv8 or similar processors that include virtualization extensions. The hardware then has the acceleration needed to make server virtualization efficient. The server also needs enough processors and memory to support the intended number of workloads. Major virtualization vendors provide hypervisor software that functions on ARM processors and OSes, such as Ubuntu Server and Desktop for ARM 12.04.

Considering the fledgling nature of RISC at the server level, any ARM server deployment and use of virtualization, such as an HP Moonshot platform, should only follow extensive lab testing and in-house proof-of-principle projects.

Allocating resources with 64-bit ARM processors

Virtualized resource allocation for 64-bit ARM systems works in a similar way to conventional x86 systems: You must adjust resources to optimize the workload's performance. The way computing resources are scaled, however, differs on each system.

In traditional x86 systems, processor scalability is limited because OSes and workloads are generally not designed to scale between multiple processors. A single x86 processor usually provides enough computing cycles that it's rare for a workload to scale across multiple cores or processors.

In RISC processor systems, on the other hand, individual RISC cores don't provide the sheer processing power of their x86 counterparts. Shaving speed and complexity help the processors save energy and run efficiently. Administrators will still need to change the way they perceive RISC processor cores, however. For example, an Intel Xeon may provide 10 cores, while a 64-bit Tilera TILEGx-8072 provides 72 cores. Tilera's S2Q multimode cloud server, a complete RISC server, provides eight 64-core processors totaling 512 cores in a 2U rack server package.

This paradigm shift means that RISC applications are designed to more readily scale across cores and processors than conventional x86 platforms.

This opens the door for greater workload computing scalability and can be a suitable platform for some cloud-based applications where users demand large scalability.

Initiatives like HP's Project Moonshot underscore the growing importance of RISC and RISC processor architectures. The growth of RISC in server virtualization will take a firm adoption of 64-bit architecture and the careful deployment of virtualization for optimum workload flexibility and scalability.

Buying the next generation of hardware for virtualization

Virtualization is still changing how we use hardware. Here's what you should know about buying the next generation of hardware for virtualization.

Virtualization has precipitated a lot of changes in data centers around the world. Most notably has been how it has changed the IT pro's relationship with hardware. By detaching workflows from physical servers, virtualization opened new doors and changed the way businesses buy and use hardware. When businesses first adopt virtualization, most simply repurpose existing servers. Nowadays, hardware vendors were shipping specialized servers with more memory and computing power. What's in the future of hardware for virtualization?

This month, we ask our Advisory Board members about how virtualization is continuing to change this relationship with hardware and what you should keep in mind when looking to buy new hardware for virtualization.

Maish Saidel-Keesing, NDS Group Ltd.

Once upon a time, your average server had 2 GB, perhaps 4 GB, or even more rarely 8 GB of RAM. Today even your basic VDI desktop uses more than that.

So how have vendors changed their hardware to accommodate this ever-hungry monster? They are continuously evolving and pushing their hardware to accommodate more and more RAM. In today's blade servers you can put 768 GB. DIMM sizes are becoming even larger -- up to 32 GB now, but that will not be the end.

When looking for suitable hardware for virtualization, there are several things that should stand out.

1. How much RAM can you cram into the server? If you were to ask any virtualization admin, the first and most limiting factor for their servers is RAM, not I/O, not processing power. The more RAM, the better. VMware's decision to introduce (and later eliminate) what is now known as the vRAM tax was a mistake that rightly upset businesses that saw the importance of more RAM.

2. Does the server have the capability to boot from a flash device? With today's hypervisor, there is not real need to have local disks.

3. How much solid-state storage can I put into the server? Local solid-state drives accelerate storage access.

4. Does the underlying technology provide you the possibility to rip and replace a physical server and reconfigure all the unique identifiers (MAC addresses and WWNs, for example) on the fly?

   Applying a set of configurations to a blank piece of hardware should be the way to go.

5. Can your server benefit from some kind of I/O flash card that will accelerate storage access and lower your storage costs?

These are some of the things to consider when getting the best of the breed for your virtual machines.

Jack Kaiser, Focus Technology Solutions

This month I went to our CTO, Bill Smeltzer, to answer the questions.

"Server virtualization created a hardware abstraction layer within the x86 environment. As the software-defined data center matures, it promises to abstract more data center elements, such as firewalls, storage and other hardware devices.

As the functionality of these devices move to software in the form of virtual appliances and native components within the hypervisors, propriety hardware devices will become less important. The problem with this approach usually shows up in terms of lower performance and scalability. The elements of the software-defined data center run on general-purpose x86 hardware, while many hardware devices utilize hardware-specific design, such as encryption and graphic chips.

Because of the advantages of application-specific integrated circuits, specific virtualization-aware devices that can integrate seamlessly into the environment can provide the ability to leverage hardware scalability while not hampering the software-defined data center functionality."

Jason Helmick, Concentrated Technologies

Virtualization released all of us from being constrained to our hardware -- it even seems silly today to think back into the past when we would need to purchase hardware to add a new server. Today, we scale our server on the virtualization platform of our choice, making better -- and smaller -- hardware purchases. There may come a day when the data center is completely empty!

Take a good look at Office 365 and the amazing amount of customers dumping their local hardware for Exchange and SharePoint in the cloud.

The new release of Windows Server 2012 R2 will blend integration with Windows Azure so that local on-premises VMs can be easily transitioned to Azure, again allowing you to dump local hardware (and hardware management resources) for a more cost-effective model.

This isn't going to happen overnight, but look inside your own data center and notice how virtualization has allowed you to consolidate and squeeze the best return on investment from your hardware.

Going forward, losing the hardware altogether is the next logical step. Until then, make smart purchasing decisions. Realize that the hardware you purchase will be repurposed for a variety of virtualization needs, not a specific line of business application. Maximize the flexibility of your hardware and minimize the cost. Who knows? Maybe your next server deployment will be to the public cloud, where someone else can worry about the spinning gears.

Networking

Networking hardware options

Before virtualization, most servers used only a fraction of their network bandwidth and a single network interface card had no trouble keeping up. Now, with a physical server capable of hosting hundreds of virtual machines, a server's network interface card or cards can easily be overwhelmed.

Networking problems, like I/O bottlenecks, are some of the biggest challenges an IT admin will face. Choosing the right network hardware can help ease your problems.

Network I/O virtualization: Making 10 GbE smarter

This tip explains how a 10 GbE card can deliver network I/O virtualization from the evils of low bandwidth, ease management and improve power consumption.

Server virtualization is having the same impact on network I/O that it is having on other parts of the IT infrastructure: exposing weaknesses. In this article, we'll look at how virtualization is making weaknesses more visible, and how lack of bandwidth management is one of the key limitations in a server virtualization rollout. Drilling down, we'll look at one I/O problem-solver, the 10 gigabit Ethernet (or 10 GbE) card

Network I/O weaknesses

Network I/O is not as plentiful as CPU or memory resources. If you follow most of the server virtualization vendors best practices of installing a 1GbE card per virtual machine (VM), you quickly run into a wall as to how many virtual machines you can support on a given virtualization host. After all, there are only so many card slots available.

These weaknesses in network I/O cause the business to react in several ways. First, almost every data center has to quickly break the one card per virtual machine practice.

But doing so means that only certain servers, those with low network I/O requirements, are virtualized. This has the obvious downsides of not virtualizing the entire environment, while at the same time reducing the overall benefit of server virtualization. The second step, and focus of this article, is becoming common practice in virtualization rollouts; installation of a 10 GbE Ethernet card. There is no doubt that 10 GbE greatly increases the amount of bandwidth available to large data sets of an application workload. This is a fairly easy solution for a physical infrastructure, but what if you're dealing with a virtualized server environment?

NICs and network I/O

In this world, multiple virtual servers compete for the same network I/O resources at a very high speed through a single layer of software common to most virtual environments; the hypervisor. The hypervisor is the traffic cop that sits between the physical machines and the physical hardware components of the host server's hardware. Workloads are serviced on a first-come, first-serve basis. As the virtual server count increases on each physical hardware platform, the amount of times the hypervisor is interrupted to handle the network I/O request increases.

This is where one or multiple virtual servers that require high network I/O become such a problem, because they compound the situation by interrupting the hypervisor even more.

You can also see where this leads to the recommendation of network interface card (NIC) per virtual machine . This deployment method divides the shared I/O queue into multiple queues; one for each physical adapter. This results in less hypervisor interruptions and helps to better organize the interruptions that do arise. Using a single NIC for a single virtual machine, or even multiple virtual machines, will cause other issues beyond just running out of card slots, which are already at a premium in a virtualized environment.

Multiple NICs become an obvious cabling management nightmare as well, in addition to increasing power utilization. For example the typical 1GbE card draws about eight watts of power. Installing 10 of these to match the performance of a single 10 GbE card would draw about 80 watts of power while at the same time reducing airflow in the server itself. A reduction in airflow increases cooling requirements that will cause the fans to run more often, hence increasing the power consumption of the server. By comparison, a single 10 GbE card draws only about 15 watts of power and has almost no impact on server airflow.

For these reasons, despite the shared queue problem highlighted above, 10 GbE cards are a wise investment and an excellent way to future-proof the environment.

The problem with a 10 GbE NIC is that it takes advantage of the 10-times increase in bandwidth; a situation that becomes exacerbated in virtualized server environments. In a non-virtualized environment, a high performance server with a 10 GbE card can achieve a throughput of 9.9 Gbps That same server in a virtualized environment can only achieve about 4.0 Gbps. This is due to the fact that in virtualized environments all I/O operations go through a shared queue in the hypervisor and only one I/O channel in the adapter. As the number of virtual machines and network I/O loads increase, the performance of the virtual machines and the applications within those virtual machines cannot be maintained due to I/O contention on the shared queue.

Solving Network I/O issues in a virtualized environment

There are two potential solutions to this problem. First the hypervisor vendors in the industry should develop an advanced I/O queuing system that allows for the bandwidth of the 10 GbE card to be divided up. This will also require a more intelligent 10 GbE card to take advantage of the queuing system. The second part of the solution is for the network card vendors to develop a Quality of Service (QoS) capability similar to what is available currently from switch vendors. This will allow the card to be segmented at a hardware level to deliver a guaranteed level of performance for the I/O channel.

The first hypervisor to deliver this capability is VMware's, although it is expected that other suppliers will follow suit. VMware calls their capability NetQueue and it dramatically improves performance in 10 GbE environments. Support from the NIC supplier is also required and right now NICs from Intel and Neterion support NetQueue. NetQueue offloads the packet routing work that the VMware ESX host performs, therefore freeing up CPU resources and reducing latency.

NetQueue makes broad server consolidation possible by using optimized 10 GbE adapters to allow the 10 GbE adapter with NetQueue support to get much closer to its rated bandwidth potential. Results of 9.8 Gbps are achievable with such a setup. But critical workloads require a guarantee of bandwidth and NetQueue does not deliver a true QoS for specific virtual machines. QoS has long been available from switch manufacturers but there has been a lack of QoS capabilities at the NIC level.

In the days of one app per server, there really was not much need for NIC-based QoS. Unlike a switch that had traffic coming and going to multiple sources a server used to have only one objective; handle compute and processing power for its application. Now because of the advent of virtualization, there is a need to give certain virtual servers network I/O priority and guaranteed bandwidth. Using the queue to service I/O requests on a first-come, first-served basis is unacceptable.

Suppliers like Neterion are helping to complete the solution and broaden virtualization even further by developing a QoS for Network Interface Cards (NICs) called IOQoS.

As part of the coming 802.11 proposals, IOQoS is designed from the ground up for multi-application environments where several applications will compete for I/O access. Using hardware-based channel isolation allows administrators to strictly enforce complete I/O isolation between the different virtual machine data paths. The advantages of true isolation can be seen in security by avoiding interference between shared resources like memory and CPU cores. Most of today's software applications were written on the notion that they would have full access to system resources and that the network card would be all theirs. The sharing of that card is now occurring in virtualized environments, but this reality has not yet been incorporated into minds of the developers. The ability to provision the card to make it appear exactly as the application expects it to be will increase vertical compatibility with existing software solutions.

Finally, there are the bandwidth assurances when consolidating several of the software stacks as individual virtual machines on a physical host. From an I/O perspective the card can be configured to look to the application as a single physical card and that application will have assured access to that bandwidth.

In the past, system administrators had to assure bandwidth to certain application workloads to either keep those servers in an unvirtualized state, or dedicate NIC cards in the virtual server to those workloads.

Each, as was explained earlier, has a negative impact on the savings that a virtualization rollout strategy could potentially help a data center achieve.

The combination of 10 GbE Ethernet with the intelligence of a queuing system from the virtualization vendor, support of that queuing from the NIC vendors and the adoption of IOQoS will allow for significantly broader and denser deployments of server virtualization. It will also optimize the investment in the infrastructure delivering a more significant return on investment.

Maximizing I/O virtualization

Use I/O virtualization to consolidate virtual server workloads and reduce port and NIC overcrowding with this tip.

Input/output (I/O) virtualization gives IT pros ways to maximize the benefits of server virtualization, streamlining provisioning and reducing the number of network interface cards (NICs) and ports used. In this tip, I describe how to work with virtual NICs and other processes, following up on my tip on basic I/O virtualization concepts.

Utilizing virtual I/O

An effective I/O virtualization strategy requires thinking differently about virtualization.

In many ways, this is very similar to the shift in philosophy that was required when organizations first began to embrace server virtualization. In both cases, each creatively utilizes resources by pooling those that are unused.

Organizations and users just getting started with server virtualization have to become comfortable with the idea of consolidating multiple workloads on a single physical server which allows tasks to share physical resources. Similarly, organizations and users investigating I/O virtualization must break away from old ways of thinking with regard to provisioning I/O resources.

Before proceeding any further, it's necessary to define some terminology.

We'll use the term virtual network interface card (vNIC) to denote a virtual NIC presented to the virtualization host. Each of these vNICs will map to a physical network port or to a group of physical network ports. Multiple vNICs may map to the same physical network port(s). Likewise, a virtual host bus adapter (vHBA) is a virtual HBA presented back to the virtualization host. These vHBAs are mapped to physical Fibre Channel ports. As with vNICs, multiple vHBAs may map to the same physical port(s).

In many data centers which use server virtualization, servers are provisioned with six, eight or more network interface cards (NICs). Why so many NICs? In a typical VI3 deployment, NICs might be configured like this:

- 2 NICs for the Service Console (to provide redundancy)
- 2 NICs for the VMotion network (again, to provide redundancy)
- 2 NICs for the virtual machines themselves

This is a fairly common configuration, but are all these NICs necessary? Not really. They're present for redundancy, even if they will never get used in the normal course of operation. The extra Service Console and VMotion NICs in particular are very likely to remain inactive and unused throughout the life of the server in the data center.

Users familiar with server virtualization but unfamiliar with I/O virtualization will start creating vNICs and binding them unnecessarily to physical network ports in much the same way as they would create and configure network ports on a traditional virtualization host. How much traffic does an ESX Service Console really need? How many Service Consoles can I run on pair of Gigabit Ethernet links?

If a typical ESX Service Console generates 50Mbps of traffic, an organization could easily combine ten Service Console connections on a single Gigabit Ethernet link with plenty of bandwidth to spare.

Just as server virtualization is about combining under-utilized workloads across multiple physical servers, I/O virtualization is about combining under-utilized I/O connections across multiple servers.

Implementing I/O virtualization

Let's take a look at a concrete example. Although this example focuses on the use of an Xsigo VP780 I/O Director in conjunction with VMware Infrastructure 3 (VI3), the concepts should be similar for other I/O virtualization solutions and other server virtualization solutions.

Consider a data center with 10 ESX hosts connected to a VP780 I/O Director. Without an I/O Director, approximately 60 Gigabit Ethernet ports would need to be configured (6 ports per ESX host). But how many would we need with an I/O Director?

Let's assume that the Service Console connection for each ESX host generates an average of 60Mbps of traffic. This means that we could combine all 10 Service Console connections, an aggregate of 600Mbps of traffic, onto a single Gigabit Ethernet port. Add a second Gigabit Ethernet port for redundancy, spread the traffic across the two and each physical port will carry approximately 300Mbps of traffic.

Each ESX server would have two vNICs defined, vNIC1 and vNIC2.

Each vNIC would be mapped to one of the two Gigabit Ethernet ports on the VP780 chassis, and each vNIC would serve as one of two uplinks to vSwitch0, the vSwitch hosting the ESX Service Console connection. Half of the ESX servers would use vNIC1 as the primary uplink for their vSwitch; the other half would use vNIC2.

This configuration meets the performance requirements, provides redundancy and still reduces overall port count by 90% -- down from 20 connections to only two connections.

The network connections for VMotion are handled in much the same fashion, but this time another factor must be considered. Whereas the Service Console traffic was primarily directed from the Service Console connections to other servers on the network, the VMotion traffic is almost exclusively within the ESX server farm. This means that we can take advantage of a feature called "inter-vNIC switching", where traffic between two vNICs on the same I/O card and on the same virtual local area network (VLAN) will be switched internally within the VP780 chassis and won't go out on the network. This means that even fewer Gigabit Ethernet connections are needed. If 75% of the VMotion traffic is between hosts connected to the I/O Director, then we've immediately reduced our needed network connections from 20 (10 servers at 2 connections each) down to only 5 connections -- and that's not even taking into consideration how often VMotion occurs.

This would be implemented as two vNICs, vNIC3 and vNIC4, for each ESX server. These vNICs would be bound across five physical Gigabit Ethernet ports on the VP780 and configured as uplinks to a vSwitch, vSwitch1, which hosts the VMkernel port for VMotion.

Because the vNICs are terminated on the same I/O card and are on the same VLAN, inter-vNIC switching automatically keeps the majority of the traffic off the Gigabit Ethernet uplinks.

Virtual machine uplinks

So far, we reduced our port count from 40 ports -- 20 for Service Console, 20 for VMotion -- down to only seven ports, a reduction of about 82%.

With virtual machine uplinks, even if we assume 50% utilization, which is a very generous utilization figure, we can reduce the port count down from 20 connections to 10 connections, and still provide redundancy. This would be implemented as a pair of vNICs for each ESX server, with vNICs for two different servers bound to the same Gigabit Ethernet ports.

All in all, by understanding the I/O requirements across our ESX server farm and consolidating I/O workloads, we've been able to reduce the total port count from 60 Gigabit Ethernet ports down to only 17 ports, a total reduction of about 72%. That's pretty impressive.

As you can see, the key to effectively using I/O virtualization in your environment means knowing and understanding the I/O requirements of the servers while carefully considering how these I/O workloads can be consolidated. By employing a shift in thinking about the provisioning of I/O resources, organizations can see the same kinds of efficiencies and cost reductions with I/O resources as they have with server virtualization.

How to improve network performance via advanced NIC options

Ethernet networks often struggle with modern data center workloads, but emerging NIC technologies are here to help improve network performance.

A few modern network interface card options can help IT professionals improve network performance for key servers.

Ethernet is the predominant network technology for data centers and business local area networks (LANs), but it often struggles to support modern workloads like storage data, real-time voice and video. Ethernet was designed for a world of simple file transfers and small packets of data that rely on contention to access the wire. Even with vastly expanded bandwidth, Ethernet can be inefficient for time-sensitive traffic that is averse to packet loss.

Network interface cards (NICs), sometimes called network interface controllers, are evolving to include more features and intelligence that will boost network performance, including jumbo frames and offload capabilities, packet tagging, buffer and spacing tweaks, and more. Some of these NIC features have caveats for use in the data center, however.

Efficient CPU use: Jumbo frames vs. offload capabilities

If server performance is lagging, it could be because of network-intensive workloads. Standard Ethernet packets are 1,542 bytes; most files are broken into hundreds, thousands or even millions of packets or frames. These small packets, individually transferred across the network, efficiently use the wire and share the network with a multitude of nodes, but sending and receiving each frame requires CPU overhead.

Most NICs support jumbo frames, which means handling packets, or frames, of up to 9,000 bytes. Jumbo frames contain more data in each packet, so fewer packets are needed to convey data across the network. Throughput improves with less overhead -- packet headers and other packet content -- and CPU overhead shrinks.

Jumbo frames do have disadvantages. The administrator must configure every node on the network to support jumbo frames for nodes to communicate properly.

Jumbo frames are not part of the IEEE standard, so different NICs configure jumbo frame sizes differently; it may take some experimentation to properly configure every node for jumbo frame operation. In addition, larger packets can add latency for some workloads because other nodes wait longer to access the wire, and dropped or corrupted packets will take longer to request and resend.

IT professionals may forgo jumbo frames in favor of NICs with large segment offload (LSO) and large receive offload (LRO) capabilities. LSO and LRO allow the CPU to transfer much larger quantities of data to (outbound) or from (inbound) the NIC with far less processing, essentially providing the same CPU performance benefit as jumbo frames.

Capacity traffic: Adjustable interframe spacing vs Ethernet upgrades

Ethernet waits a set time between each packet send, which is called interframe spacing. This gives other network nodes an opportunity to grab the wire and send a packet. Interframe spacing equals the time it takes to transmit 96 bits on the wire. For example, 1 gigabit Ethernet uses a standard interframe spacing of

0.096 microseconds, and 10 gigabit Ethernet uses one-tenth that gap, or 0.0096 microseconds.

This fixed spacing between transmission attempts isn't always efficient and can degrade network performance under heavy traffic conditions.

NICs that support adaptive interframe spacing can dynamically adjust the interframe spacing based on network traffic, potentially boosting network performance. Adjusting the interframe spacing offers little benefit to network performance unless you're approaching the network's full bandwidth.

Comprehensive network performance benchmarking can reveal utilization patterns. If the Ethernet link frequently reaches capacity, upgrading to a faster Ethernet link or implementing NIC teaming will provide a longer-term fix than adjusting the interframe spacing.

Interrupt throttling for CPU performance

When packets move along the network, NICs generate interrupts for the CPU. At faster Ethernet speeds, the CPU interrupt rate increases, and the CPU must give more attention to network drivers and other software that handles the packets. If traffic levels spike and drop, CPU performance can become erratic. NICs that support interrupt throttling artificially reduce the CPU interrupt rate, freeing the CPU from unlimited NIC interrupts and potentially boosting CPU performance.

More throttling isn't necessarily better. You can slow a CPU's responsiveness with high interrupt throttling; it will take longer for the CPU to get around to handling all of the interrupts being generated.

With a high rate of small packets coming in at close to real-time conditions, throttling degrades performance rather than enhancing it. Test network and CPU performance in various throttling modes until you can establish adequate system responsiveness while smoothing out interrupt demands on the CPU.

Alternatively, consider NICs with TCP/IP offload capabilities. These can handle many of the CPU-intensive tasks onboard, also reducing interrupt demands on the CPU.

Prioritizing time-sensitive data types: Enable packet tagging

Time-sensitive data types like Voice over IP (VoIP) and or video are frequently treated as higher-priority traffic than simple file transfers, but the network treats every packet on the wire equally. Implement packet tagging is enabled. Tagged packets can then be sorted into a traffic queue set up by the operating system (such as Windows Server 2012), pushing these higher-priority VoIP and video packets to the front of the line to be handled before other lower priority packets. Packet tagging is instrumental to Quality of Service (QoS) strategies, and is an essential part of many virtual LAN(VLAN) deployments.

Only apply changes to NICs if network performance falls below defined benchmarks, and always roll out changes after controlled testing with server and NIC benchmarking.

These recommended NIC adjustments won't make the dramatic improvements that a network overhaul accomplishes, but they also aren't limited by budget or logistical concerns. Evaluate network performance changes over time and look for any unintended consequences, such as a boost to one workload but a detriment to others.

Increasing network bandwidth available to virtual machines

These three tips on increasing the amount of network bandwidth available to your VMs, including virtual network creation, can also potentially increase a host server's VM density.

A big server virtualization trend involves achieving the maximum virtual machine (VM) density within the available host servers. As the VM density increases, the hardware cost per VM decreases, and network administrators will often pack physical servers with as many processor cores and as much memory as the server can accommodate. Unfortunately, the server's network ports often limit the server's VM density as each VM on the host server requires network connectivity. While a network interface card (NIC) can be shared among multiple VMs, there is only so much network bandwidth to go around. This can be problematic, especially if some of the virtual servers are running network-intensive applications.

Fortunately, there are ways to increase the amount of network bandwidth available to your VMs, possibly allowing you to increase the VM density on your host servers.

Using additional network hardware

The easiest solution is to install additional NICs in your host server.

The physical hardware only provides one or two integrated network ports on the motherboard and only has a limited number of expansion slots. All expansion slots could be in use, but you may still be able to increase the number of network ports installed in your server by using NICs that have multiple ports. For example, some PCI-X and PCI Express (PCIe) cards contain four separate network ports on a single card. Using these cards can quadruple the number of network ports installed in your server.

Another option may be to take advantage of external NICs. For example, if you run out of network ports on your host server, you could add one or more USB network adapters. USB-based network adapters, however, are not always ideal. Some virtualization platforms will not expose USB devices to VMs. Even if this is the case, you may still benefit by adding a USB network adapter.

A best practice is to reserve a NIC for management traffic. You can't dedicate all of the available network bandwidth solely to VMs -- there must be a way to communicate with the parent partition (machine) to manage it.

Even if your virtualization software exposes a USB NIC to the VM, you may be able to dedicate a USB NIC to the parent partition. If you had previously dedicated one of the other NICs into your system to management traffic, then using a USB NIC instead will free that NIC for use with your VMs.

It's usually possible to share a single NIC among multiple VMs. The problem is that a single NIC provides a limited amount of network bandwidth, and bandwidth must be shared among all VMs bound to the NIC. But you can increase the available network bandwidth by installing faster NICs instead of increasing the quantity of NICs installed in your server. For example, switching to 10 Gigabit Ethernet can make the difference in virtual server performance.

Introduce virtual networks

If there isn't a practical way to add additional network bandwidth to your server, there is one more trick to try: decreasing the demands on your existing bandwidth instead of increasing the server's total network bandwidth.

One way to accomplish this is by creating a virtual network segment, a network segment that exists solely within the host server. If you're using Microsoft Hyper-V or VMware, there already is at least one virtual network. Every physical network adapter used by a VM is connected to a virtual switch. That virtual switch is in turn connected to a virtual network adapter on each of the VMs that use the physical NIC.

Virtual networks commonly provide VMs with connectivity to the physical network, but they don't have to access the physical network.

It's possible to create a virtual network segment that services VMs, but doesn't tie into the physical network. You could create such a segment to offload some of the network traffic from your physical network.

For example, if you had a virtualized Web server that frequently used a back-end SQL Server database running on another virtual server, you could ensure that database queries don't consume physical network bandwidth by creating a dedicated virtual network segment between the Web and database servers. By doing so, the network bandwidth available to your virtual servers is increased even without adding additional network ports.

Match network port use to server needs

Creating additional virtual network segments will only help if you can offload a significant amount of network traffic to the virtual network segment. If this isn't possible, by dedicating certain adapters to certain servers, you may still be able to take advantage of how your virtualization platform uses virtual switches for each physical network adapter.

The idea is that some servers inevitably consume more network bandwidth than others. Since you probably don't have enough network ports to dedicate a port to every virtual server, you can decide how NICs should be shared based on each server's bandwidth consumption.

For example, if you have a SQL Server that consumes a lot of network bandwidth, then you could dedicate a NIC to it.

Conversely, you may have domain controllers and a DHCP server that don't use much bandwidth. As such, you could probably get away with sharing a single NIC among these servers.

The availability of network bandwidth has the potential to limit a host server's overall VM density, but there are several ways to increase the amount of bandwidth available to the server and make more efficient use of your server's existing bandwidth.

The iSCSI versus NFS debate: Configuring storage protcols in vSphere

For performance rivaling Fibre Channel, you can bet on iSCSI or NFS. But one protocol has the edge when it comes to ease of configuration in vSphere.

SAN versus NAS and iSCSI versus NFS are long-running debates similar to Mac versus Windows. Many enterprises believe they need an expensive Fibre

Channel SAN for enterprise-grade storage performance and reliability. In reality, your vSphere infrastructure functions just as well whether you use NFS or iSCSI storage, but the configuration procedures differ for both storage protocols.

The block-based vs. file-based storage protocol debate

Whether you use a Windows server, a Linux server or a VMware vSphere server, most will need access to shared storage. With vSphere, the virtual machines (VMs) running in a high availability/distributed resource scheduler cluster must reside on the shared storage, so that if a server goes down, another server can access them.

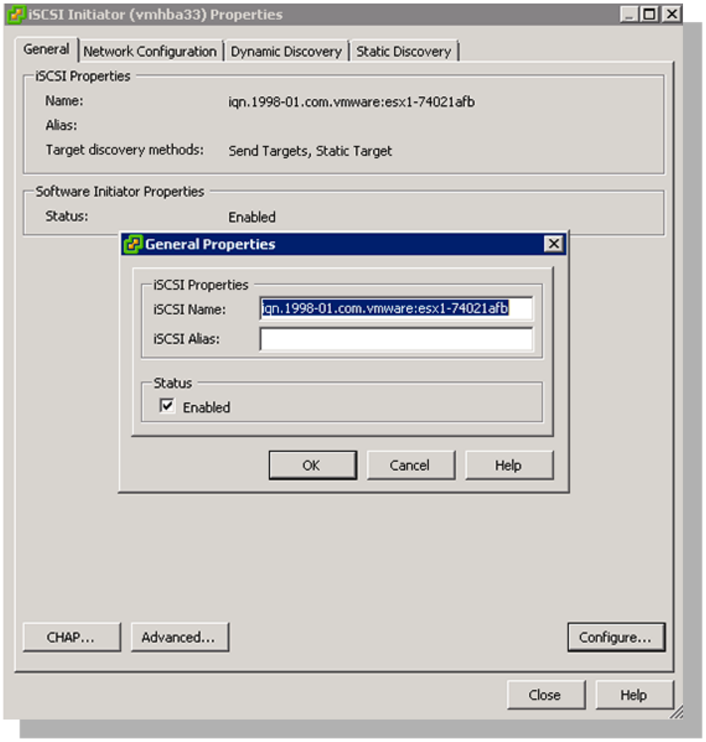
VMware vSphere has an extensive list of compatible shared storage protocols, and its advanced features work with Fibre Channel, iSCSI and NFS storage. Fibre Channel and iSCSI are block-based storage protocols that deliver one storage block at a time to the server and create a storage area network (SAN). NFS, on the other hand, is a file-based protocol, similar to Windows' Server Message Block Protocol that shares files rather than entire disk LUNs and creates network-attached storage (NAS). So which protocol should you use?

The SAN vs. NAS debate

Fibre Channel, unlike iSCSI, requires its own storage network, via the Fibre Channel switch, and offers throughput speeds of 4 Gigabit (Gb), 8 Gb or 16 Gb that are difficult to replicate with multiple-bonded 1 Gb Ethernet connections.

However, with dedicated Ethernet switches and virtual LANs exclusively for iSCSI traffic, as well as bonded Ethernet connections, iSCSI offers comparable performance and reliability at a fraction of the cost of Fibre Channel.

The same can be said for NFS when you couple that protocol with the proper network configuration. Almost all servers can act as NFS NAS servers, making NFS cheap and easy to set up. NFS also offers a few technical advantages.



NFS and iSCSI have gradually replaced Fibre Channel as the go-to storage options in most data centers. Admins and storage vendors agree that iSCSI and NFS can offer comparable performance depending on the configuration of the storage systems in use.

Connecting vSphere to an iSCSI SAN

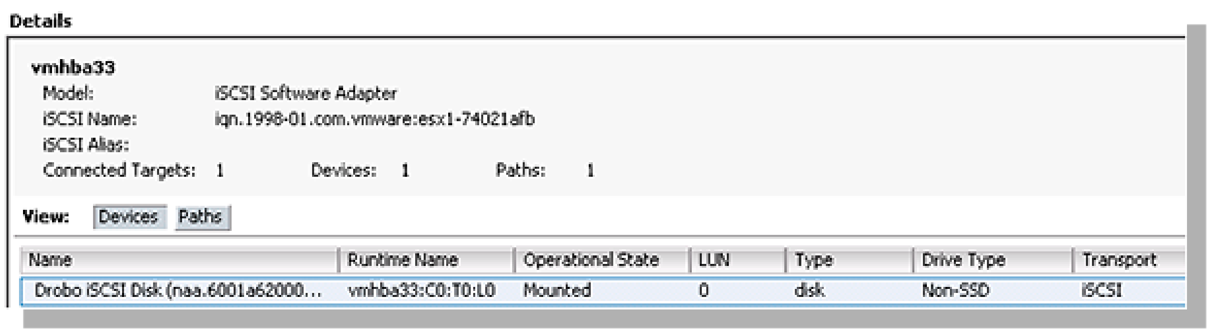
In a vSphere environment, connecting to an iSCSI SAN takes more work than connecting to an NFS NAS.

To demonstrate, I'll connect a vSphere host to my Drobo B800i server that is an iSCSI-only SAN. Then I'll connect the same host to my Synology DS211+

server, which offers NFS, iSCSI and other storage protocols. This comparison gives you a good indication of how to administer connections to each of the storage options.

First, you must enable the iSCSI initator for each ESXi host in the configuration tab, found under storage adapters properties. (See Figure 1.)

Next, you need to tell the host how to discover the iSCSI LUNs. In this example, I use static discovery by entering the IP address of the iSCSI SAN in the static discovery tab.

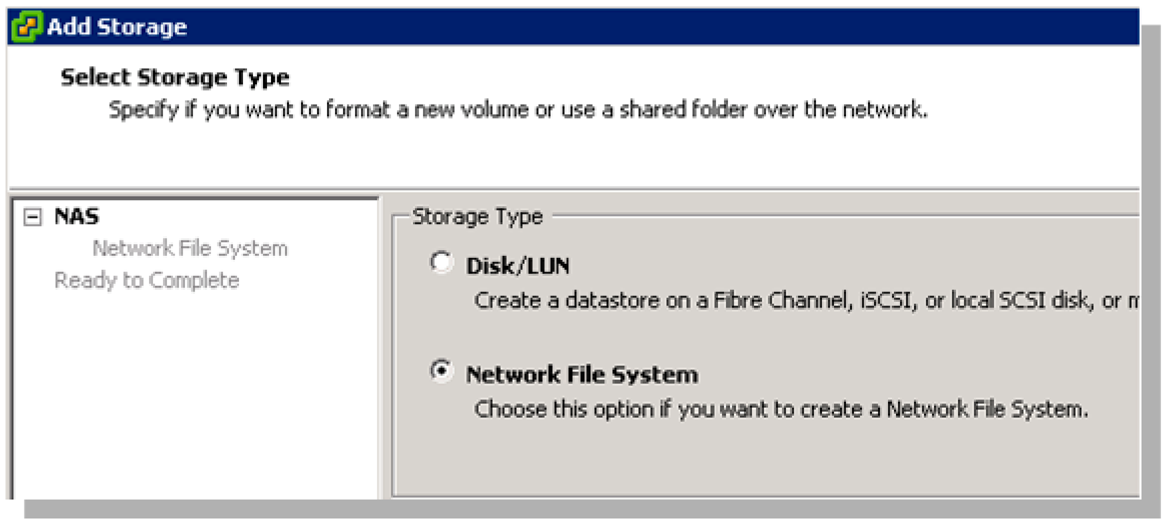


Once you enable the iSCSI initiator, and the host discovers the iSCSI SAN, you'll be asked if you want to rescan for new LUNs. As you see in Figure 2, the host discovered a new iSCSI LUN.

A formatted iSCSI LUN will automatically be added as available storage, and all new iSCSI LUNs need to be formatted with the VMware VMFS file system in the storage configuration section.

Connecting vSphere to an NFS NAS

With an NFS NAS, there is nothing to enable, discover or format with the Virtual Machine File System because it is already an NFS file share.



To add NFS storage, go to the ESXi host configuration tab under Storage and click Add Storage, then click on Network File System. (See Figure 3.)

You will need to provide the host name of the NFS NAS, the name of the NFS share and a name for the new NFS data store that you are creating.

Within seconds you will be able to create VMs in the NFS share.

Connecting vSphere hosts to either an iSCSI SAN or an NFS NAS provides comparable performance to the underlying network, array configuration and number of disks spindled. Though considered a lesser option in the past, the pendulum has swung toward NFS for shared virtual infrastructure storage because of its comparable performance, ease of configuration and low cost.

Storage

Storage hardware trends

Storage technology has made huge advances in recent years, and keeping up with these trends is essential to having the best hardware for virtualization. Hypervisor-aware storage, with integrated software that gives admins better insight into virtual workloads, storage IOPS and disk use helps prevent performance problems before they start. While most virtualized infrastructures rely on shared storage, there is a growing movement today to outfit servers with local storage, which has improved and dropped in price over recent years. Find out which approach is right for your data center.

Want per-VM reports and statistics? Turn to hypervisor-aware storage

Virtualization admins will have less of an opportunity to blame storage for unidentifiable problems thanks to the advent of hypervisor-aware storage.

Servers, network devices and storage need to be more hypervisor-aware to keep up with today's expanding virtual infrastructures. Fortunately, some storage vendors have recognized this need and developed virtualization-aware storage that could benefit your infrastructure.

Bare-metal hypervisors virtualize everything on the server, and the underlying guests never know the difference. As these hypervisors have matured, however, they have asked a bit more from the server hardware. For example, the latest version of VMware vSphere now requires a 64-bit CPU and either an Intel VT or AMD-V CPU virtualization extension. These CPU features allow the software -- the vSphere or Hyper-V hypervisor -- to work with the hardware in ways not previously possible. Now that the CPU cooperates with the hypervisor, storage should follow suit.

How hypervisor and storage cooperate with VAAI and VASA

Today, VMware's vStorage APIs for Array Integration (VAAI) and Microsoft's Offloaded Data Transfer allow for some cooperation between the storage and the hypervisor. Many enterprise storage vendors have also adopted VMware's vStorage APIs for Storage Awareness (VASA).

VAAI speeds common hypervisor storage-related operations, such as virtual machine (VM) cloning or disk space reclamation. When you perform a snapshot of a VM located on a VAAI-capable array, you'll prevent locking on the entire logical unit number (LUN) unique identifier.

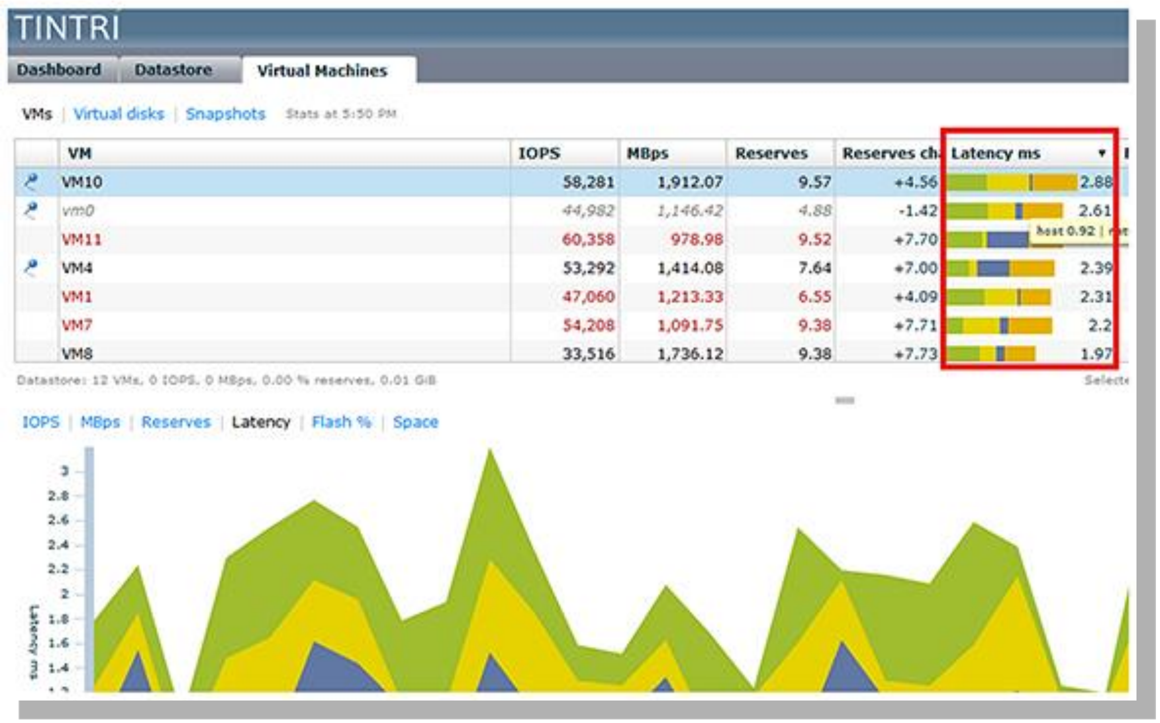
VASA, on the other hand, lets storage report to the hypervisor on the storage capabilities. The hypervisor benefits from these storage reports, but storage could also benefit from VM reports. With that knowledge, the storage could offer the most important VMs the best performance and give the virtualization and storage administrators per-VM statistics.

Let's look at some of the benefits admins can reap from virtualization-aware storage.

The benefits to hypervisor-aware storage

As an admin, you want to know how many IOPS each VM generates, which VMs are write-intensive, how many IOPS the whole virtual infrastructure demands and how many VMs are using that storage. Virtualization-aware storage provides you with that information.

Another benefit to virtualization-aware storage is end-to-end bottleneck identification from the perspective of the storage. Previously, admins blamed the storage for problems they couldn't identify. Now, storage admins and virtualization admins with access to the storage can get more accurate storage statistics. The storage can then combine its knowledge of the virtual infrastructure to quickly solve performance problems and identify bottlenecks. Figure 2 shows latency identification from Tintri Inc.'s storage management interface.



Finally, virtualization-aware storage goes beyond statistics to actually perform operations on virtual machines, thanks to its connection to vCenter. For example, traditional storage can only take a storage snapshot on an entire LUN. Tintri, on the other hand, can take a snapshot or clone an individual VM from the storage layer.

Tintri, Tegile Systems and Nutanix all provide hypervisor-aware storage, but all hardware vendors must create more virtualization-aware products to keep up with evolving and maturing virtual data centers.

VDI storage requirements: Do you need a new storage array?

Implementing VDI could mean a new storage array on your part. Figure out your VDI storage requirements and learn what types of storage arrays are best for VDI.

The wrong storage array can suck the life out of a VDI deployment. Before you virtualize desktops, determine your VDI storage requirements and whether your array can handle the extra load -- otherwise, you may need a whole new system.

Virtual desktop infrastructure (VDI) workloads have different storage requirements than physical server workloads. Physical servers generally have steady storage requirements with occasional peaks and lows. The workload is usually made up of a small number of machines -- often less than 10 servers for every 100 staff members. VDI workloads, on the other hand, can be very quiet one minute, and extremely demanding the next. Plus, VDI servers are made up of many small loads -- possibly as many as there are users.

Adding VDI to your environment could drain the storage performance of existing servers, create VDI storage bottlenecks and bring your VDI deployment to a standstill. VDI storage requirements are much more unpredictable, so you need a storage array that can support more resources and a greater number of workloads. You may need to purchase a dedicated array for VDI storage, but if your environment is small enough, there are other options.

VDI storage depends on deployment size

Small VDI: If your expected VDI workload is smaller than the workload on your physical servers, you could use the same storage for VDI that you have in your existing environment. This will work, for instance, if you only virtualize about 5% to 10% of desktops and your storage array has a good amount of spare capacity.

Since the additional load is light, it won't adversely affect the existing server load. Still, you should re-assess whether you need additional VDI storage if you start to roll out more desktops.

Large VDI: A larger VDI deployment is more difficult to accommodate. If you're virtualizing more than half your desktops, the added VDI load could be too large for your old storage array.

But before you run and buy new storage for VDI, check out your existing array and see if it can support segregation of your physical server load and VDI load. If you can guarantee good storage performance for the physical servers even while the VDI load stresses the storage array, then the two loads can coexist on one array. To segregate the different loads, it's best to have a storage array that's virtualization-aware or provides multi-tenancy.

If you can't segregate the server and VDI storage loads on one array, you'll need a new storage array just for virtual desktops. To get the best VDI storage performance, you may even need to select a completely different model or vendor than you use for your existing storage system. You could also try a different transport protocol, for instance, moving to IP-based VDI storage rather than the traditional Fibre Channel. You should also check out more advanced storage features, such as tiering or modular scale-out approaches.

Having the right storage for VDI is critical to a successful rollout. For a small VDI deployment, the existing storage array may be adequate, but larger implementations may require isolated workloads or a dedicated array for VDI storage. Make sure that the system you have fits your VDI storage requirements and that the storage and virtualization teams' needs are met.

Data center storage virtualization considerations

Before making data center storage virtualization decisions, consider storage performance and redundant storage access, among other issues

Servers get a lot of attention, but successful virtualization deployments really depend on storage. Storage protects virtual machine (VM) states, retains snapshots and provides user access to shared application data. Because of these activities, availability and performance issues place great demand on a storage infrastructure. So administrators need to understand the most critical data center storage virtualization considerations.

Data center storage virtualization consideration No. 1: Storage performance

Today, administrators can select from myriad different shared storage platforms, such as iSCSI, Fibre Channel, network-attached storage and Fibre Channel over Ethernet. Each of these platforms works fine in a virtual environment but offers radically different scalability, performance, availability and capacity characteristics.

Storage performance is usually considered chief among equals because multiple VMs residing on a physical host have considerable storage bandwidth requirements. A virtualized server hosting 10 VMs, for example, needs to load all 10 VMs from storage, to periodically snapshot the current state of each VM and to provide the VM data to users.

While storage performance must keep pace with these added demands, it's imperative to select a platform that meets your firm's needs. An iSCSI storage area network (SAN), for example, may offer the characteristics necessary for a small or midsized business, but a large enterprise may require a Fibre Channel SAN.

Data center storage virtualization consideration No. 2: Integration with backup and disaster plans

Storage systems are backed up and protected with some manner of disaster recovery (DR) planning.

Any new storage implementation should be fully compatible with existing backup and DR software, such as local snapshot and remote replication tools. If not, the new storage setup may force changes (and possibly introduce errors) to the existing data protection scheme or add additional tools that unnecessarily complicate data protection. Lab testing can usually confirm a storage system's compatibility.

Data center storage virtualization consideration No. 3: Ensure redundant storage access

Storage disruptions can have devastating consequences on a virtualized data center. When a traditional server is disrupted, usually one application is affected. But when a server with 10 or 20 virtualized workloads is disrupted, it affects far more business applications and users.

Therefore, in virtualized environments, redundant storage pathways must be available to storage systems, SAN (or other network) switches and the virtual servers. This setup requires redundant storage controllers on each server and storage system, along with a well-designed redundant SAN architecture. No matter which storage architecture is in place, though, routinely verifying storage availability should be part of disaster recovery testing.

Data center storage virtualization consideration No. 4: Energy-efficient storage

The energy needed to run larger-capacity and higher-performance storage systems means a larger total cost of ownership. Consider more energy-efficient storage systems, which provide more input/output operations per second and bandwidth per watt of energy for active data. Energy-efficient storage should also provide more capacity per watt for inactive (e.g., archived) data. Achieving more energy-efficient storage is usually accomplished through a combination of controller designs and disk capacity/performance tradeoffs.

Data center storage virtualization consideration No. 5: Support for advanced storage features

Some virtualization platforms can utilize storage features, such as thin provisioning and data deduplication. Generally, storage systems that natively support these features can be more effectively utilized by a virtualization platform. This kind of mutual support often simplifies a storage setup and management requirements.

Similarly, it's important that new storage systems support the migration features in the virtualization platform. It may not be possible for the virtualization engine, for example, to migrate VMs from Fibre Channel to iSCSI to a network file system. This can potentially limit a firm's storage system choices.

But other virtualization platforms, such as VMware vSphere 4, can exchange VMs among all three storage protocols, which can ease your underlying storage system choice.

Local storage for virtualization: Will it catch on?

Local storage is making a comeback. Storage and virtualization vendors are utilizing local disks, providing a cheaper alternative to shared storage.

Using local storage for virtualization is fashionable again. For several years, best practices have dictated that administrators move virtualization storage away from local servers, in the form of storage area networks and network-attached storage. But new virtualization features and products have made local storage a cheap and useful alternative to expensive shared storage.

The trend against local storage began as organizations began moving large databases and file repositories to a storage area network (SAN), while booting servers from local disk. Later, more organization started booting servers from SAN, forgoing local disks altogether.

Then, virtualization exploded -- bringing with it a huge appetite for SANs or network-attached storage (NAS).

With each evolutionary step, local disks became less important. At the same time, local storage became faster and less expensive.

Is it time to take another look at local storage? There are a growing number of vendors that think so. For example, DataCore and Fusion-io are developing products that bring shared-storage closer to the server. Fusion-io's ioDrive and Virtual Storage Layer technologies use local NAND memory on PCIe cards. The Fusion-IO ioDrive cards provide local storage that rivals the already blazing speeds of solid state disks and lower the need for remote calls to the SAN or NAS for data retrieval.

DataCore's SANsymphony-V aggregates local disks from several servers into a pool of shared storage, providing the redundancy and availability you'd expect from an enterprise shared storage solution. It will also offer NAS functionality without a centralized NAS device.

The vendor Nutanix incorporated Fusion-io technology and a file system similar to DataCore's to win the Best of VMworld 2011 award for Desktop Virtualization. The Nutanix server platform uses a combination of ioDrive storage and local disks to create a robust, highly available and highly scalable virtual desktop infrastructure that does not require a separate NAS or SAN back end.

VMware, EMC et al. take second look at local storage

Both EMC and VMware have recognized this local storage trend. VMware introduced the Virtual Storage Appliance (VSA) with vSphere 5. Currently, VSA is similar to DataCore's offerings, grouping local disk into a shared storage pool. This feature is limited to three servers, but VMware also has the CloudFS fling, which doesn't have a server limit.

EMC also has a number of innovations queued up in this category. Project Lightning will ship soon, bringing storage to the server in a new way. Project Lightning is a PCIe card that incorporates a Host Bus Adapter and local flash storage. The flash memory will cache the data retrieved from an EMC SAN or possibly even pre-fetch data and hold it in memory for rapid retrieval.

To go one step further, EMC plans to integrate Lightning cards into its Fully Automated Storage Tiering (FAST) product, coordinating caching activities with the array to ensure the most efficient placement of data among the SAN disk, SAN cache and server-side cache.

In fact, I expect EMC to introduce additional PCIe cards with flash storage in 2012. Where Project Lightning is a read-only caching device, look for future offerings with read-write capabilities.

EMC has also expressed plans to move VMs directly to the storage arrays -- the opposite of moving storage to the server.

With more storage solutions running on hardware that’s very similar to virtual hosts, the idea is to install a lightweight hypervisor on a storage device to allow virtual machines to run directly on the array.

As for other hypervisor vendors, Red Hat’s KVM hypervisor has completely side-stepped the need for shared storage, allowing live migrations between local disks on two servers. Microsoft is also touting the same functionality in the next release of Hyper-V.

I will not say that SAN and NAS devices are here for good, because I have learned not to make definitive, far-reaching statements about technology, and because challengers to shared storage must prove themselves before they can make a serious dent in the storage market.

However, I hope that these emerging solutions will inspire big storage vendors to address the cost and complexity of large storage arrays fueling the return to local disks. And if these products prove reliable, look for the big storage vendors to buy these technologies and add them to existing storage offerings.

VMware SSD updates: Making a solid case for local ESXi storage

Solid-state storage is beginning to make a splash in data centers, and there are a handful of VMware SSD use cases for that can improve ESXi host performance.

As virtualization has grown into a standard in almost all data centers, so too has the use of shared storage. Today, there is very little use for traditional local storage on a vSphere host. In fact, given the small storage footprint of ESXi, most server manufacturers now offer a local SD card slot so admins can install and run ESXi from an SD card to reduce the power, heat and cost of local hard drives. But, the declining cost of solid-state storage and recent VMware SSD updates have caused more organizations to consider using solid-state storage on their hosts. Here are two of the more compelling use cases for local solid-state storage on an ESXi host server.

VMware SSD support: Swap to local cache

A new feature in vSphere 5 allows a host to automatically detect a local solid-state drive (SSD) and allow for part of that disk to be set up as a local cache for VM swap files. In some cases, this VMware SSD configuration can greatly improve performance.

When a host starts to run out of RAM capacity, it takes several steps to free up RAM to meet the needs of the VMs. The host starts by eliminating redundant pages in RAM using Transparent Page Sharing. Next, a host turns to ballooning to force in-guest memory pressure, and then memory compression to reduce the size of pages in RAM. If the host still cannot free up enough RAM, then it will be forced to swap out memory pages to disk. The Swap to Host Cache option, if enabled at the host configuration section of the vSphere client, will allow the swapping to first occur to the local SSD before being swapped to the shared disk. This VMware SSD support results in a significantly reduced performance hit to a VM when compared with swapping to a shared disk, because of the lower latency and faster performance of a local SSD.

VMware View linked clones

In a VMware View environment, disk I/O is often a performance bottleneck. This can lead to expensive shared storage solutions to provide the proper IOPS for the virtual desktops. Utilizing local SSDs is one way to reduce the cost of the shared storage solution. Admins can implement this in one of two ways, both of which require all hosts to have local SSD storage.

1. Place the golden image (replica) on shared storage and the linked clones on the local SSDs. This approach will help reduce the shared disk space needed.

2. Place both replicas and linked clones on the local SSD. This approach increases the amount of SSD space necessary on each host, but drastically reduces the performance requirements of the shared storage array. The reduced cost of the shared storage array can often justify the extra cost of the local SSD space.

In both cases, since the VM is stored on local storage, the virtual desktops cannot be live migrated. Using this approach will require more planning when performing host maintenance since the desktops will have to be powered down in order to reboot host servers.

VMware SSD alternatives

Most server platforms now offer SSDs that fit the form factor of SAS or SATA drives, making it easy to add local SSDs. Alternative solid-state storage options are now appearing on the market offering even higher performance benefits by eliminating the SAS or SATA bus and plugging directly into the PCI bus. VMware currently supports offerings from Fusion-io and LSI.

EMC recently introduced a new product, called VFCache, a PCI solid-state storage device used to extend the SSD-based cache capabilities of their arrays into the host servers. This provides better I/O performance to commonly read blocks of data by caching them locally on the host server.

Although most of these solutions are currently solving niche pain points, the industry seems poised for a renaissance of fast local storage to increase disk performance while also reducing the performance needs of the shared storage arrays.

In this e-guide

Getting more PRO+ exclusive content

This e-guide is made available to you, our member, through PRO+ Offers—a collection of free publications, training and special opportunities specifically gathered from our partners and across our network of sites.

PRO+ Offers is a free benefit only available to members of the TechTarget network of sites.

Take full advantage of your membership by visiting http://pro.techtarget.com/ProLP/

Images; Fotalia