Realtime
publishers

The Shortcut Guide™ *To*

Implementing Virtualization in the Small Environment

sponsored by

vmware®

Greg Shields

Copyright Statement

Chapter 3: What Do I Need to Get Started with Virtualization?

"All these benefits are great. But if the technology I already own doesn't support virtualization, it won't work for my business!"

Nowhere is this statement more important than with small businesses and environments. The goal of your business is to support your customers and ultimately make a profit. As such, implementing technology for technology's sake is always a losing bet, especially when your business isn't necessarily a technology company. For a technology such as virtualization to make sense for your small business or environment, it must be capable of arriving with a minimal or zero cost outlay. It must run atop your existing infrastructure, and it must provide a direct benefit to your goals of supporting your customers and making profit. If it doesn't accomplish these three things, it doesn't make sense for your business.

This chapter's goal is to provide you the technical information you need to ultimately make this decision. Whereas the first two chapters in this guide focused specifically on the cost and operational benefits associated with virtualization, this chapter will focus on a technical discussion of its hardware and software requirements. The intent here and in the next chapter is to provide a comprehensive guide to getting started. Although the content may be of a fairly high technical level, rest assured that the steps necessary to get started are actually pretty easy. Consider the information in these next chapters as your guide for ensuring that start is as successful as possible.

To begin, Figure 3.1 details what will be your grocery list of hardware and software components. These are minimally required to build a lightweight and no-added-cost virtualization infrastructure. You'll see in this figure that specific server components are required as well as elements at the network and storage layers. In addition to the right hardware, certain software components are required or optionally suggested. Much of the rest of this chapter will discuss each of these requirements in detail.

Your environment is likely to have most of these components already in place. Thus, repurposing them for use with your virtualization environment will be the first step in your implementation.

Physical Servers

- Two or more Servers preferred
- Two or more Processors
  - 2.0GHz or Faster Intel or AMD
  - x64 architecture preferred
- Two gigabytes or more of Memory
  - Four or more preferred
- Onboard SAS/SATA disk storage
  - Two gigabytes required for server installation
  - Additional space equal to sum of virtual machines required
- Two or more Network Interface Cards
  - Four or more preferred
  - Gigabit speed highly recommended

Networking

- Business class network hardware required
  - Gigabit speed highly recommended
  - Support for NIC teaming highly recommended
  - VLAN support recommended

Storage

- Direct-Attached Storage required
  - Capacity equal to the sum of VM disk size
- Storage Area Network Hardware optionally recommended
  - iSCSI-based SAN is often most cost effective.
  - Support for multipathing highly recommended
  - Support for concurrent host access highly recommended
  - Support for storage virtualization recommended
  - Support for storage-level backups recommended

Software

- Virtualization platform software required
  - Virtualization management software recommended
  - External monitoring software recommended
- Virtual backup software optionally recommended
- Replication software optionally suggested for disaster recovery
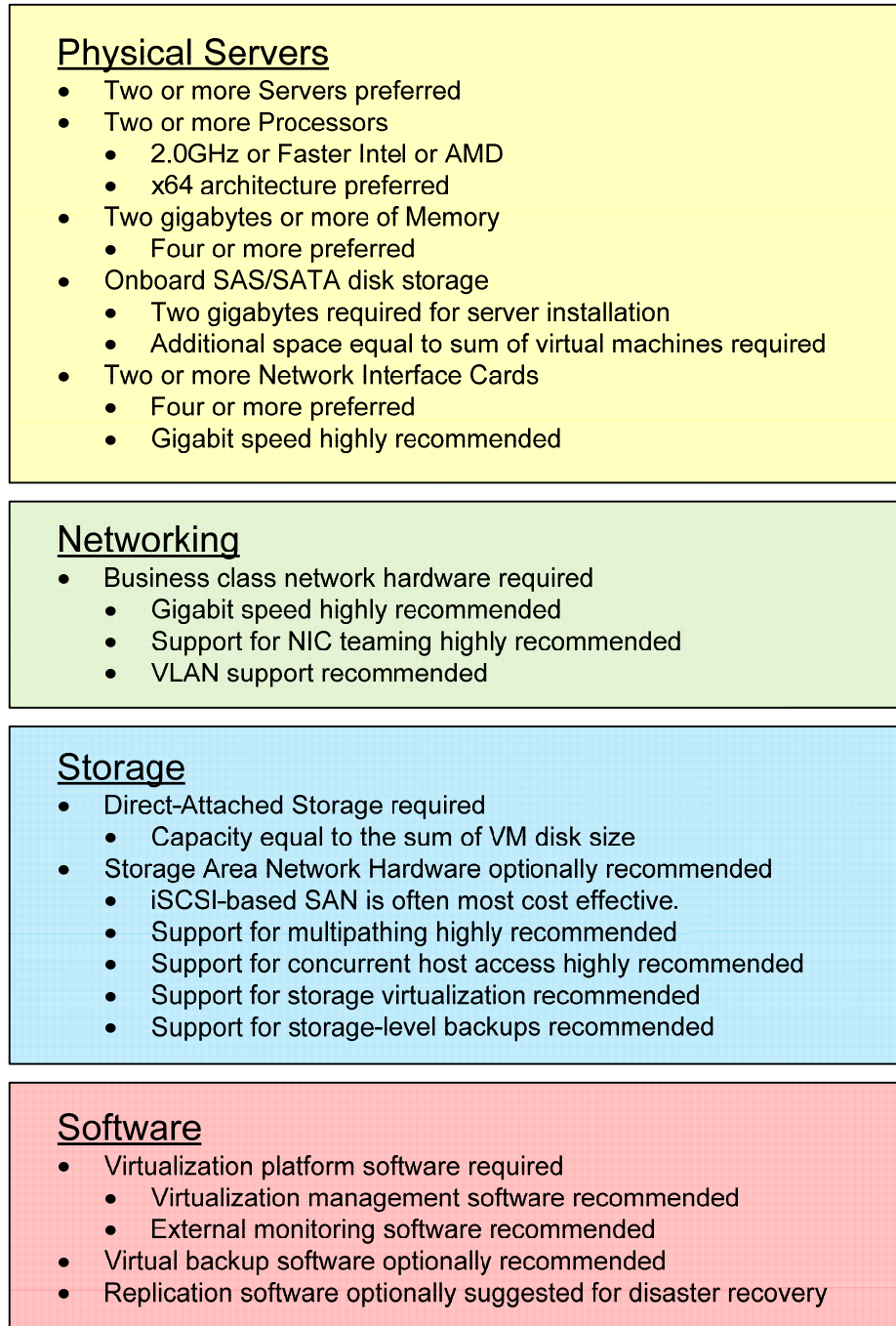
Figure 3.1: Your grocery list for a basic virtualization implementation with optional additional hardware and software.

Physical Servers

Obviously, to create a virtual host, you first need a physical host. Unlike many other IT services, virtualization requires large quantities of physical resources to fulfill its mission. Such is the case because a virtualization platform has the ability to scale the running of concurrent virtual machines to essentially any level. The limiting factors are the number of hosts you have as well as the resources on each of those hosts. Greater numbers of physical processors at higher speeds can support the processing needs of more virtual machines. Higher levels of onboard RAM memory mean more virtual machines can be supported by a single physical host. With this "more is greater" concept in mind, let's analyze each of the classes of server resources that you should consider.

Processors

Today's server-class equipment typically arrives with a minimum of two processors. Many servers now include the support for four, eight, or more processors, leveraging multiple processor cores per socket. From the perspective of overall virtualization performance, the number of processor cores versus the number of sockets is irrelevant; however, the overall count of processors is important.

Licensing Processors and Cores

Although a processor versus a core may be insignificant from a raw performance perspective, be aware that virtual platform licensing may make one configuration's price different than another. For example, some virtual platform software may be priced by the number of sockets rather than the total number of processors.

A major function of the hypervisor layer is to schedule the processing of virtual machine workloads across available processors. Although a processor can process only a single instruction at a time, the rapid swapping of instructions in and out of a server's processors enables the appearance of concurrent functionality across every virtual machine. Because of this scheduling process, the more processors available for targeting by the hypervisor's scheduler means that more virtual machines can be serviced at the same time.

There is no established technical precedent associated with the number of virtual machines per onboard physical processor. Instead, effective performance monitoring is key to determining the best fit. Virtual machines with light resource requirements will use fewer processor resources than those with higher resource requirements. Understanding the anticipated resource needs of your virtual machines is important to knowing how many processors to bring to bear in a virtualization environment.

With this in mind, consider using or purchasing physical servers that have a minimum of two onboard processors for your initial implementation. Servers with four or more processors will support more virtual machines at a higher purchase cost. Going even further, servers with eight or more processors are available today for even greater expansion potential; however, these servers tend to be of much greater cost.

Realtime
publishers

vmware

Memory

The second major factor in optimizing the maximum number of virtual machines per physical host relates to the amount of physical RAM memory in each server. Each running virtual machine requires the use of RAM memory that is roughly equivalent to its assigned virtual memory. Thus, if you create a virtual machine that has been assigned 2GB of virtual memory, that virtual machine when booted will consume an equal amount of physical memory.

It is for this reason that Chapter 1 discussed the fallacy of assigning too much memory to server instances. Consider the situation in which a physical server is running a very light workload, such as a file server or DNS server. In this case, if the physical server is outfitted with 4GB of memory but only uses one for its daily processing, its extra memory goes unused. Although a waste in and of itself, this situation does not create a problem for other physical servers because their memory is contained within their server chassis and not shared.

The oversubscription of memory in the virtualization context has substantially more effect. Consider a virtual host that contains 8GB of memory. If 4GB of memory are assigned to a virtual machine that only in fact needs one, the extra memory is both wasted and unavailable for use by other virtual machines on the same host. It is for this reason that virtual machine memory consumption must be carefully monitored and adjusted based on the resource needs of the virtual machine.

Physical Memory Sharing

Although this consumption of physical memory is the case with virtual machines, some virtualization platforms leverage the use of memory page table sharing. This feature allows physical memory to be shared between multiple virtual machines when the contents of each page table are equivalent. The use of memory page table sharing reduces the net effect of virtual machine memory requirements and results in an overall reduction of consumed memory when multiple virtual machines are simultaneously running.

There are, however, downsides to the use of this feature. The resource overhead associated with managing the sharing of memory page tables takes away from available resources on the host, which can result in a reduction in overall performance. Additionally, when virtual machines power on, they initially do not have the right information necessary to begin memory sharing. Thus, recently powered-on virtual machines cannot participate. As such, even with this feature, make sure that ample physical memory is available on your virtual hosts.

It is also worth mentioning that the amount of physical memory installed to a virtual host should be maximized as much as possible. If your business can afford to augment an existing virtual host with additional memory, its addition is usually the best and least expensive change that will net the greatest benefit to the virtual environment.

Onboard Storage

Virtual hosts also require onboard disk storage for the hosting of the virtual platform itself. It is to this onboard disk storage that you will install the virtual platform's operating system (OS) and runtime files. It is also to this onboard disk storage where you may install virtual machines and their disk files as well as any supporting data such as software media.

The onboard storage requirements for most virtual platforms are relatively low in comparison with the quantity of disk storage available in today's servers. Figure 3.1 shows that a minimum of 2GB of disk storage are suggested for the virtual platform itself. This quantity is for the installation of its files and supporting software.

Due to its low cost and high levels of performance, direct-attached storage (DAS) is where many small environments also store their virtual machine disk files. If you are repurposing existing equipment and do not have a storage-area network (SAN) in your environment, DAS will be your only option for virtual machine storage. The correct amount of DAS required for your virtual environment will be based on a number of factors:

- **Disk size for running virtual machines.** You will need enough storage to support the full disk size of every concurrently-running virtual machine. Thus, if you plan to host 10 virtual machines and each virtual machine is configured with a 40GB disk, your host will need a minimum of 400GB of DAS spread across your virtual hosts.

- **Disk size for non-running virtual machines.** You will find that virtual machines that are running are only one part of the total number of virtual machines in your environment. Over time, you will begin to accumulate non-running virtual machine instances as well. These instances may be related to test and evaluation systems created for one-time or irregular use. They may be virtual machines that are powered on only once in a while for specific uses. These virtual machines and their disks also require storage and will need to be factored into your calculations for disk storage.

- **Disk size for future expansion.** Because of the ease in deploying new virtual machines, many environments see a resulting explosion in virtual server instances once they have made the jump to virtualization. Your plans should include the additional disk space necessary to support your future expansion.

- **Disk size for virtual machine templates.** As discussed in the previous chapters, an important value-add of virtualization is the ability to quickly create new server instances that are based on virtual machine templates. These templates are effectively virtual machines themselves that are not actively run as servers. Instead, they are used as the source for copying and pasting new server instances.

- **Disk size for software media.** Virtualization environments tend to make heavy use of ISO files that are created from original CD and DVD media. These files are convenient for use, as administrators do not need to locate the media and load it into the correct server's CD/DVD tray. Your additional disk space requirements for this media will depend on how much media you want to make available for use by your virtual machines.

As you can see, the disk space requirements for virtual machines can grow to become extremely large. In fact, many environments quickly find that the limitations of DAS soon become a barrier to virtual environment expansion. With the concurrent running of even 10 virtual machines requiring a half-terabyte or more of DAS, you will quickly find that many servers are limited in the number of virtual machine disks they can host. In this case, it is likely that you will eventually decide to move to network-based storage through an iSCSI, Fibre Channel, NAS, or NFS connection.

Cross-Reference

Later on, this chapter will discuss some of the options and decisions you may consider when making the jump to SAN storage.

Clones Reduce Disk Consumption

Many virtualization platforms use a feature called *virtual machine cloning* to reduce the overall impact of virtual machine storage. With clones, the disk files of virtual machines are linked. Here, a source virtual machine is used as the starting point for creating additional virtual machines. Linked virtual machine disk files are much smaller in size because they contain only the differences between the source and their unique configuration.

This technology reduces the overall disk space because servers tend to be highly similar in nature. Two freshly-built server instances can be almost completely equal in their disk configuration—with the obvious differences associated with their names and other personality elements.

Cloning technology, however, comes at a cost. The extra resources required to manage the clones and their linkages can reduce the overall performance of the virtual machine. Although clones can often be created from other clones, this multiple-hop linking reduces performance further still. Also, although freshly-built virtual machines are highly similar in configuration, that similarity will tend to diverge over the life of the virtual machine, reducing the disk savings efficacy of cloning technology.

Network Interface Cards

Servers that operate as virtual hosts should be outfitted with a minimum of two network interface cards (NICs) attached to your production network. These network cards are usually operated in a "teamed" configuration, with each network card providing redundancy for the other. In this configuration, should one network card fail, the other can take over and prevent a loss of service.

Although the teaming of network cards is not new technology, this network card configuration is particularly important with virtual hosts. In a physical environment, the loss of networking to a server can result in the loss of a network service. However, with virtual hosts, the loss of networking can result in the loss of numerous network services at once. Therefore, ensuring a high level of network redundancy is critical for assuring high levels of virtual machine and network service redundancy.

Using servers with two network cards is, however, only the start. More network cards are extremely useful and are often required when using advanced technologies such as iSCSI and Live Migration. It is not unheard of for virtual servers to support six network cards or more, depending on their requirements. Figure 3.2 shows an example of how a server with six network cards can be typically provisioned in a virtual environment. This server makes use of redundant iSCSI connections, redundant production network connections, and separate connections that are dedicated to Live Migration and virtual host management.

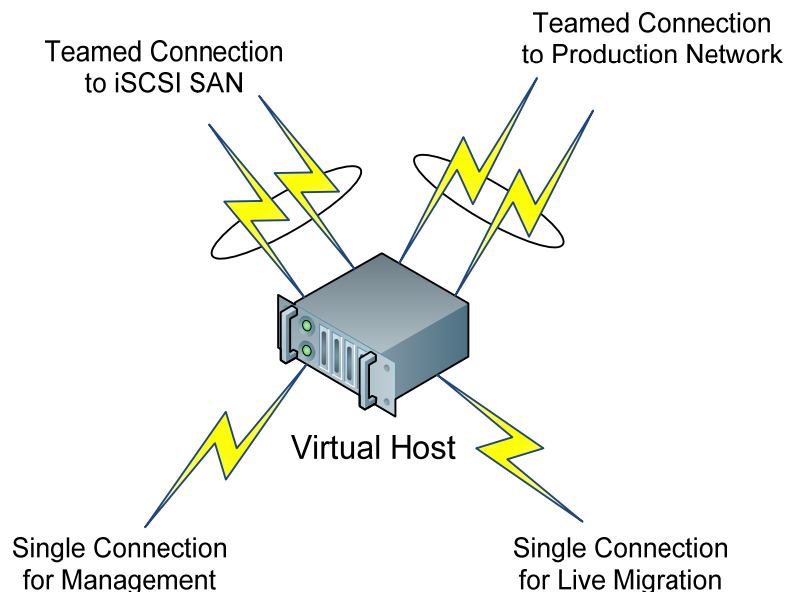**Figure 3.2: Virtual hosts often use multiple and teamed network connections for various purposes.**

Additional network cards can be further used in situations in which complex networking arrangements are required. For example, a virtual host may need to support virtual machines that are in different network domains: One virtual machine may reside in your company's DMZ, while another must reside in your internal LAN.

Virtual hosts with extra onboard network cards can uniquely and inexpensively host virtual machines across trust domains through the creation of multiple virtual switches. In this situation, one virtual switch may be physically connected to your internal network while another is physically connected to your DMZ network (see Figure 3.3 for an example of this setup). The use of network VLANs makes the logical configuration of this environment inexpensive to implement and easy to manage.
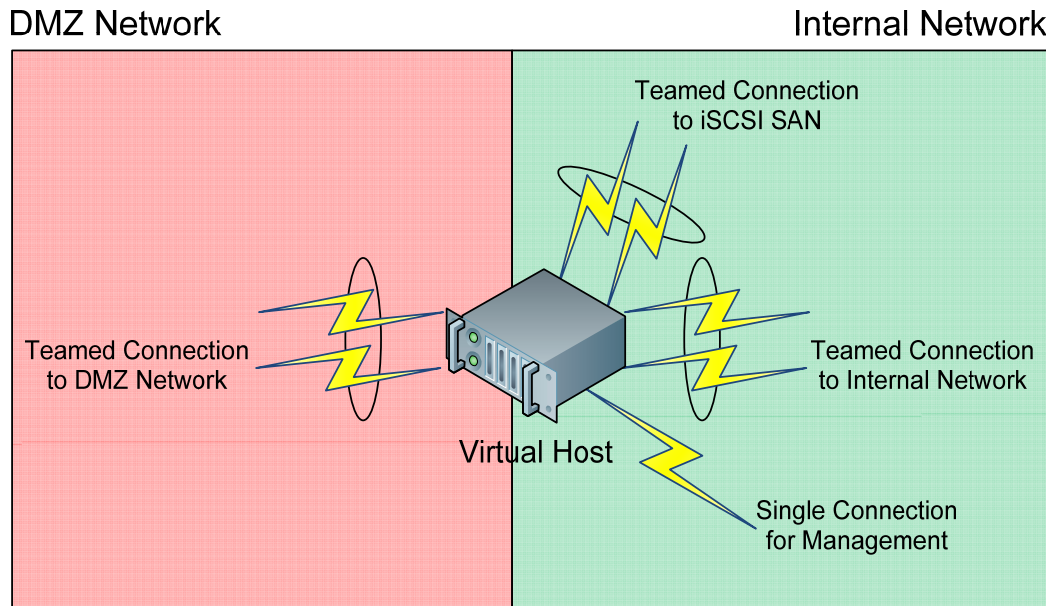
DMZ Network · Internal Network

Teamed Connection
to iSCSI SAN

Teamed Connection
to DMZ Network

Teamed Connection
to Internal Network

Virtual Host

Single Connection
for Management

Figure 3.3: Additional network cards in a server can be used to bridge network domains.

An added benefit of this configuration is that its virtual machines can all be managed from a single location, irrespective of their assigned network. By bridging network domains in this way, you can administer your DMZ servers through the same management interface that you do your internal servers. This eliminates the otherwise problematic step of physically working on servers in your DMZ or extranet because of their network limitations.

Networking

Obviously, the networking infrastructure used by your company's IT environment must be able to support the networking arrangements required by virtualization. Without getting into the specifics of networking devices and protocols, your network hardware should be able to support gigabit networking. It should additionally be able to support the teaming of network connections from a single server. This capability is available on virtually all business-class networking equipment available today and is a requirement if you plan to team the network cards on your servers.

You may also consider creating VLANs for the purpose of segregating network traffic. One example of where this is particularly useful is in segregating iSCSI network traffic from production network traffic. iSCSI network traffic tends to be intolerant of network conditions such as latency and bandwidth contention. Because your virtual machines' hard disks are being connected to over the network, the segregation of this traffic into its own VLAN goes far in preventing transient network issues from affecting iSCSI connections.

Lastly, if you want to use Live Migration features in your virtualization infrastructure, your networking equipment must support the rapid re-convergence of network routes after a Live Migration event occurs. This is commonly supported on most business-class networking equipment today. When a Live Migration occurs, the physical path that clients use to connect to a virtual machine will change. When that path changes, your network and networking equipment must be able to quickly identify the change and reconfigure their internal routing tables to support the new path.

Storage

It was discussed earlier in this chapter that most environments that move to virtualization eventually find the need to segregate their storage from their servers. The physical limitations of DAS quickly become a limiting factor in the number of virtual machines that an environment can simultaneously support. Also, the sheer amount of data storage that is required for a virtualization environment often forces many organizations to move to separate storage. Although your small environment may not initially need this level of storage, you should be conscious of your storage limitations and be aware that an augmentation may be necessary down the road.

Making the move to a SAN is usually not a decision that is made lightly by any small business or environment. Although SAN costs have dropped dramatically in the past few years, any SAN purchase is likely a big decision for a small business. With this in mind, you should be aware of a few questions the answers to which you need to know should you decide to purchase SAN storage for your virtual machines:

- **How much storage do you need? How much will you need?** In purchasing a SAN, you must first know how much storage you plan to peel off your physical servers and re-host on the SAN. Although most of your server instances and their OSs will likely remain on DAS, your virtual machines and other critical data will likely relocate to the SAN. In making this decision, you should also consider your future storage needs and how much extra data you believe you will need in the medium-term future. This extra storage should relate to not only virtual machines but also templates, snapshots, and other metadata that tends to grow in size as the environment ages.

- **What type of SAN will you buy?** SANs these days usually come in one of two flavors: iSCSI and Fibre Channel. Although prices can be dramatically different based on any number of factors, iSCSI SANs tend to be of lower cost than those that use fibre channel. This cost difference is primarily due to iSCSI's ability to run over your existing production network. Fibre channel, however, requires its own fibre interconnects and routing/switching hardware, which adds to the overall SAN cost. Also, iSCSI SANs tend to be easier to work with, often being fully manageable by IT generalists. Fibre channel SANs can require specialized training and experience that your IT administrators may not have in-house.

- **What will you relocate to the SAN and why?** It is important when making this determination to understand the total cost of SAN storage in comparison with DAS. SAN storage tends to cost far more per online gigabyte than local storage, yet it tends to be of higher reliability and better performance. Your decision about which virtual machines and other data to move to a SAN will be an important determination for that SAN's size and ultimate cost. You may find that some of your virtual machines and other data do not need the extra availability, resiliency, and features that are provided by the SAN. As such, you may split the storage of your virtual machines and other data so that only your highest-value workloads are stored on higher-dollar SAN storage.

- **Will your SAN support multiple-host access and multipathing?** Virtual environments that want to use Live Migration features for the online relocation of virtual machines must use a SAN that supports multiple and simultaneous host access. Although many virtualization platforms come equipped with their own proprietary file system or file-locking mechanisms to ensure hosts do not overwrite each other's files, your SAN itself must support the ability for more than one machine to see a partition.

- **Will your SAN support storage-level backups?** Chapter 4 will talk more about the specialized backup architectures that are common to virtual environments along with their best practices. But for now, be aware that some SANs support storage-level backups for the rapid block-level backup of your virtual machines.

For most small businesses and environments, the decision to move from DAS to SAN-based storage is usually aligned with a desire to increase virtual machine resiliency. This specifically relates to the ability to use Live Migration to relocate virtual machines from one host to another, a feature that requires shared storage. As your virtual environment grows, this capability will also grow in necessity. This occurs for two reasons:

- **Virtual machine high availability.** With the loss of any single virtual host affecting more than one virtual machine—and thus more than one network service—virtual environments eventually grow to need high-availability features. These features not only allow an administrator to re-host a virtual machine in preparation for a planned outage but also can automatically reboot virtual machines on alternate hosts after a host failure. Figure 1.4 in Chapter 1 shows a graphical representation of how this process works. High availability with virtual machines requires that the virtual machine disk files be stored on external SAN storage, which is why this need is usually the first step in driving a business' SAN purchase.

- **Virtual machine load balancing.** A secondary need for SAN storage relates to the virtual platform's ability to re-host virtual machines when the load on one virtual host grows too great. For example, let's assume that your environment is comprised of three virtual hosts. Each virtual host contains 10 virtual machines. The processing of those 10 virtual machines is not equal at all times. Like what is shown in Figure 3.4, from time to time, one virtual machine may require greater resources to accomplish its stated mission—a mail server on a busy day or a Web server after a new product launch are two examples. In this case, you may find that the load across one of your virtual hosts is substantially greater than on the others. Your virtual platform's Live Migration feature may be able to monitor the load on each host and rearrange the hosting of virtual machines so that each host's load is balanced. This need for the optimized use of available resources is a secondary reason why many environments eventually decide to move to SAN storage for virtual machines.
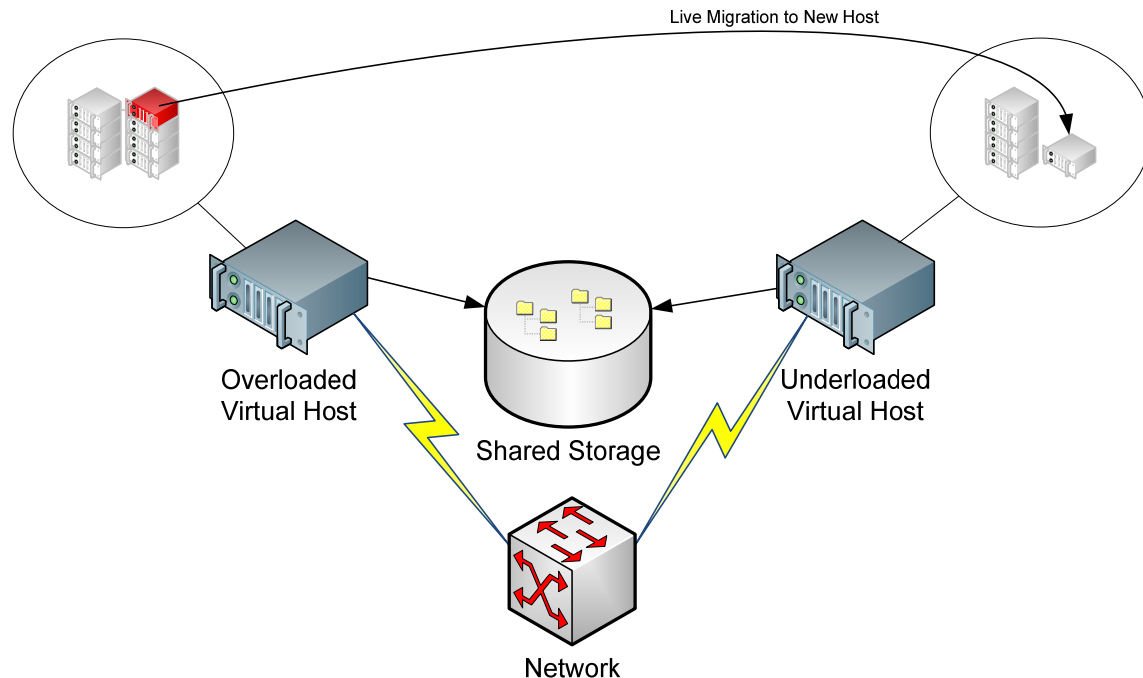
Figure 3.4: Monitoring integrations in your virtual platform can watch for mismatched loads on virtual hosts and relocate virtual machines to re-balance the load.

Software

Your last category of requirements to move to virtualization will be the software that actually runs and manages that virtual environment. Numerous software vendors develop and distribute platforms with a range of features. Some vendors have greater support for enterprise customers, while others focus on small environments. Multiple software platforms from the same vendor may have different focuses as well, with some levels of software being aimed at small businesses and environments and others fulfilling the need of enterprise customers.

The goal of this guide is to provide your business with the information it needs to make an educated decision about which virtual platforms make sense. Its further goal is to assist you with understanding that some no-cost virtual platforms are actually a good start for small businesses and environments. These no-cost platforms provide a low-risk starting point for businesses to make the move into virtualization with little or no initial investment.

The no-cost platforms available today use the same fundamental technologies as found in today's enterprise products. In fact, in many cases, the no-cost technologies—such as the hypervisor and virtual disk architectures to name a few—are actually the very same as those that are found in their more expensive brethren. You will find that the major differences between no-cost and for-cost solutions often relate to the level of management interaction needed by the administrator, the level of availability and resiliency features exposed in its interface, and the tools available for specialized implementations such as computer labs, user self-service, and cloud computing needs.

This being said, implementing a no-cost virtual platform can be only the start for your virtual software needs. You might later find the need to purchase additional software based on your needs, such as:

- **Block-level backup software.** Chapter 1 discussed how one benefit of virtualization is the reconfiguration of backups. These backups represent the entire virtual machine configuration in a single file. The technology used to complete this type of backup is much different than the technology used to back up systems in the traditional file-by-file manner. It tends to use an agentless approach, not requiring backup agents installed to each virtual machine. It is usually OS-agnostic for the virtual machines in each virtual machine, meaning that a Linux virtual machine can be backed up using the same technology as a Windows virtual machine. In addition, it tends to be of higher performance, enabling backups to occur faster with lower levels of stress on the virtual host. Some virtual platforms include basic software for accomplishing these activities, while others do not include them at all. Upon the move to virtualization, a next step will likely be the desire to change your backup methodology to take advantage of block-level backups for enhanced restoration capability. The software used to create and manage these block-level backups is often an extra cost over and above that of the virtual platform.

- **Disaster recovery software.** Enabling cost-effective disaster recovery is a further goal of many environments that make the move to virtualization. In effect, once virtualized and once block-level backups are in place, the disaster recovery process involves little more than the manual or automated transfer of those block-level backups to an alternate site. Virtual disaster recovery software also includes the necessary capabilities to bring servers back online after a disaster and to manage the later failback process once the disaster is over. Your business may later decide that disaster recovery software—and likely additional hardware—is a valuable addition to your virtual infrastructure.

- **Management and monitoring software.** The management tools that arrive with many virtual platforms enable basic management of your virtual servers. However, automation components are often nascent in these native tools. The same holds true for the built-in performance-monitoring capabilities. As your environment grows, you might find the need for additional software that can monitor your environment for known conditions and alert administrators when problems occur. Similar software can assist with automating the management of your environment, aligning your virtual environment with the other components of your IT infrastructure.

- **Virtual inventory software.** Some environments as they grow in size also find the widespread distribution of virtual machines to be a problem. Keeping track of virtual machines across multiple platforms, versions, and iterations can be a problematic activity, especially when license counts need to be trued-up. Virtual inventory software can assist with these problems by identifying and reporting on the presence of virtual machines in your IT environment.

It should be obvious that these "extra" software components are likely not necessary for your initial virtualization implementation. However, these add-on tools and many other tools like them are now available as organizations identify the need for them. Your environment will scale over time as your number of virtual machines increases. These extra tools are available for when that size grows unwieldy with native virtual platform tools alone.

So Where Do I Start?

As you can see in this chapter, there is a lot of hardware and software that needs to aggregate to make a successful virtualization implementation. However, you should also see that much of this infrastructure is likely already in your IT environment. You probably already have servers and a network. You may already have a SAN. If you've purchased servers in the last 12 to 24 months, it is likely that those servers already have the right number of processors and amount of RAM to work as virtualization hosts.

The hard part now is in repurposing that equipment to work in a virtual environment. If your business is like many small businesses, you probably don't have excess equipment available for repurposing your existing workloads. If you have 10 servers in your environment and all 10 are currently serving a stated mission for your IT environment, you may find that a new server purchase is necessary.

Your goal with this first step is to make available a single server chassis to be converted for use as a virtual host. One alternate process you can use that replaces the need for an added server purchase is to virtualize the workload of the server you want to convert to use as a virtual host. Doing so with any virtual platform's P2V tools creates a functional copy of that server as a virtual machine. Once created, you have effectively made "a backup" of that original server that can later be hosted on top of the virtual platform.

The Spectrum of Added Cost & Capabilities

$

| Add Memory | Isolate Storage Traffic | Add Processors | For-Cost Virtual Platform |

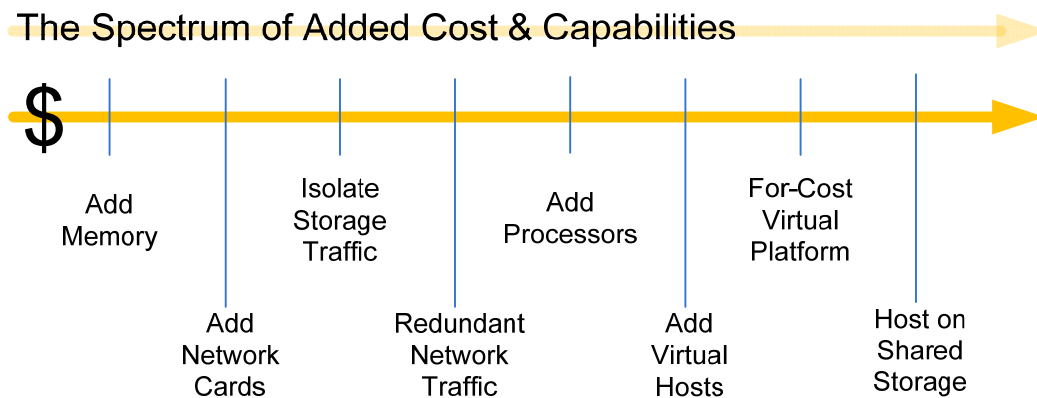| Add Network Cards | Redundant Network Traffic | Add Virtual Hosts | Host on Shared Storage |

Figure 3.5: Starting on the left with low-cost add-ons and moving to the right for more expensive options, there are many ways to enhance your virtual environment.

The other question you may have after reading this chapter relates to where your money is best served for augmenting your existing hardware in preparation for your virtualization implementation. This chapter has discussed a large number of areas in which you can add dollars to gain additional performance, availability, or capacity. Figure 3.5 shows a spectrum of add-on capabilities that you can purchase for your existing hardware to improve any of these areas. On the left of Figure 3.5 are the options with the lowest initial cost, while those on the right will be more expensive. Selecting the right expansions that align with your available funding and the features you want will be the next step in architecting your virtual infrastructure.

To assist with that architecture-development process, Chapter 4 will continue this discussion. In Chapter 4, you'll learn more about the technical best practices that have been developed over the years. These best practices associated with management, backups, and even disaster recovery will assist you with keeping your virtual infrastructure running in good shape after you've made that initial investment.

Download Additional eBooks from Realtime Nexus!

Realtime Nexus—The Digital Library provides world-class expert resources that IT professionals depend on to learn about the newest technologies. If you found this eBook to be informative, we encourage you to download more of our industry-leading technology eBooks and video guides at Realtime Nexus. Please visit http://nexus.realtimepublishers.com.