VIRTUALIZATION: THE INS AND OUTS

Most enterprise IT/network professionals have been tinkering with one form of virtualization or another in non-production environments for quite some time. Now many of those tinkerers are readying for the big move into live deployments, while those already there are tossing an increasing number of applications onto the virtual infrastructure. Still, with all its various nuances and its role not only in the server but also desktop, storage and network realms, virtualization can be confusing. In these articles, Network World and its sister publications CIO, Computerworld and InfoWorld provide the lowdown on virtualization technologies.

IN THIS eGUIDE

2 Seven Half-Truths About Virtualization
‘Virtualization will make my life easier’ and other caveat-laced claims

5 Virtualization Technologies Compared
In this book excerpt, virtualization experts provide a vendor-neutral look at the technology

11 Emulation or Virtualization?
Don’t be confused! Virtualization and emulation aren’t the same. Here’s how they differ

13 Virtualization or Cloud Computing: Which Way to Go?
The decision between virtualization and cloud computing should be made on a case-by-case basis

16 Desktop Virtualization: Popular, But Confusing
Here’s a look at the five most fashionable flavors

19 Deep Dive: SAN and NAS Virtualization
How storage virtualization is proving its worth

23 Seven Required Topics to Study for Virtualization Certification
An expert blogger’s view
SEVEN HALF-TRUTHS ABOUT VIRTUALIZATION

By Jon Brodkin • Network World

‘Virtualization will make my life easier’ and other caveat-laced claims

Virtualization is revolutionizing the data center, mostly for the better. But no technology is without potential pitfalls. Problems related to management, security, ROI and power use can all trip up a virtualization deployment that isn’t planned properly.

“Virtualization has the potential to deliver immense cost savings and technical benefits” through the consolidation of servers and reduction of space and power needs, notes Laura DiDio, lead analyst with Information Technology Intelligence. “However, these savings don’t automatically happen.”

Here are seven—oh, let’s call them half-truths—to consider fully before your virtualization project is implemented.

1. Virtualization will make my life easier

Virtualizing servers will greatly reduce the time it takes to spin up new workloads. Some IT shops have reported being able to deploy new virtual machines (VM) in as little as 30 minutes, as opposed to weeks for their physical counterparts. The promise that virtualization simplifies IT is real in many respects.

But virtualization simultaneously introduces management challenges that can’t be ignored. IT shops need strict policies and perhaps third-party automation tools to prevent virtual server sprawl, the unchecked spread of VMs. Even if you end up with fewer physical servers, Burton Group analyst Chris Wolf says the overall number of managed objects can increase, because of the hypervisors and sheer number of VMs.

Many users assume administrative time will be lessened, but in reality the virtual infrastructure itself has to be managed and may require a new centralized storage system, says Martijn Lohmeijer, managing consultant with TriNext, an IT outsourcing and consulting firm.

Frustratingly, many software vendors don’t offer the same levels of support for applications running in VMs as they do for applications running on bare metal. Microsoft eased up on some licensing restrictions last year, but analysts are still criticizing Oracle and other vendors for restrictive policies related to support in virtualized environments. The calculation of software licensing fees also can be more complicated in a virtual data center.

“Not all server virtualization licenses are the same,” DiDio says. “You have to really study the terms and conditions of your licensing contracts from the various vendors.”

2. Consolidating onto fewer servers will be simple

The first goal of a server virtualization project is often consolidation. If you can run the same number of workloads on 10 servers that you were running on 100, why not consolidate as fast as
possible? Unfortunately, many IT shops that plan to consolidate end up doing so much slower than they expect, says George Pradel, director of strategic alliances for virtualization management vendor Vizioncore. It’s easy to say “every new workload has to go on a virtual machine,” but moving old workloads from a physical box to a virtual one is not always a simple task, he says.

“Doing P-to-Vs, or physical-to-virtual conversions, there’s a black art to it, if you will,” Pradel says. The conversions “don’t happen in a vacuum. That happens in conjunction with different business units’ schedules and the ability to withstand downtime.”

3. Virtualization automatically reduces power use

If you have consolidated onto fewer servers, it might be tempting to say “I’ve solved my power use problems.” Not so fast. While you now have fewer servers using up watts, each server is running at a higher CPU capacity and has greater power needs. At Brandeis University in Waltham, Mass., a virtualization project has actually increased overall power use, reports network and systems director John Turner. Although Brandeis dramatically reduced its number of servers, it is now offering more services to users because spinning up new VMs is so easy. Each new workload increases power use.

“If you walk behind the racks of virtualized servers, the heat is just pouring out of those guys,” Turner says. “We’re seeing heat dump into these rooms like never before.”

Another issue to consider: If you’re shutting off lots of servers, a data center has to be reconfigured to prevent cooling from being directed to empty space, explains APC CTO Jim Simonelli. “The need to consider power and cooling alongside virtualization is becoming more and more important,” he says. “If you just virtualize, but don’t alter your infrastructure, you tend to be less efficient than you could be.”

4. Virtualization makes me safer

The ability to clone VMs and move them from one physical box to another opens up great possibilities for disaster recovery—and that in turn protects your business from data loss and downtime. But virtualization, if not managed properly, also brings new security risks that could threaten the safety of data and continuity of business systems.

People and processes are often not ready for virtualization and the security risks it introduces, IBM security expert Joshua Corman has argued. Virtualization brings new attack surfaces and various operational and availability risks. Consolidating many applications onto a single server “gives you a single point of failure,” Di Dio notes. If you’re suffering from virtual server sprawl, it may be difficult to keep track of all your VMs, and it may thus be difficult to ensure that all of them are properly patched. Also, hypervisors do not perform encryption, leaving open the possibility of man-in-the-middle attacks such as Xensploit, which intercepts unencrypted data when VMs are migrated between physical servers.

That doesn’t mean you should avoid virtualization altogether, but it’s often best to start with minor systems and work your way up to mission-critical applications.

5. Desktop virtualization will save me money right away

Virtualization should make it easier to deploy new desktops to users, apply patches and perform other management tasks. Desktop virtualization can also save money in the long run. A Denver transportation agency is expecting a $619,000 ROI over eight years by purchasing thin clients that will last longer than traditional PCs.

But IT shops have to remember desktop virtualization requires significant upfront costs, from purchasing user devices such as thin clients to back-end infrastructure such as servers, PC blades and networked storage to support VMs.
Anecdotally, Forrester Research analysts have found that enterprises spend about $860 per user, plus network upgrades, to get a desktop virtualization project up and running in the first year. A well-done desktop virtualization deployment can certainly cut long-term costs. It just might take a few years to achieve ROI.

6. Virtualization is the same as cloud computing

Virtualization is a key enabler of cloud computing. But installing VMware on a few servers doesn’t mean you’ve built a private cloud. In addition to virtualization, a private cloud requires service automation technologies and a self-service interface for provisioning new resources, says IBM cloud software chief Kristof Kloeckner.

In a blog post titled “Virtualization isn’t cloud computing,” cloud start-up Enomaly’s founder Reuven Cohen says virtualization is a building block for cloud computing, but the real key is abstraction at every level of the IT stack.

“The key to cloud infrastructure is abstraction to the point that it ‘just doesn’t matter,’” Cohen wrote. “Your infrastructure is always available and completely fault tolerant. Think more along the lines of the iPhone application delivery model (App Store), and less like the desktop application models of the past. The companies that will succeed are the ones who embrace this new hybrid, Internet-centric model. More simply, the cloud is the computer.”

A cloud doesn’t necessarily even need virtualization. This has been proven by none other than Google officials, who have said they do not virtualize production hardware and instead use a job scheduling system of Google’s own design to manage its many thousands of servers.

7. Virtualization is all about technology

This one won’t surprise any longtime IT veteran: Sometimes, it’s the people and not the technology that gets in the way. As IBM’s Corman has noted, people and processes are often not ready for the new challenges raised by virtualization.

Even if your virtualization project is a hit, you might become a victim of your own success. Once users realize how easy it is to spin up a VM, they may become more demanding, making it harder for IT to focus on other tasks. Conversely, there may be resistance from users who prefer to stick with physical servers.

Pradel likes to call politics the “eighth layer” of the network stack.

“The political layer is the most difficult thing you have to deal with as far as virtualization goes,” Vizioncore’s Pradel says. “You have members of your business community who I like to call server huggers. They do not want to get rid of their physical machines, even though they will benefit from going to virtual.”

“You have members of your business community who I like to call server huggers. They do not want to get rid of their physical machines, even though they will benefit from going to virtual.”

—George Pradel, Director of Strategic Alliances, Vizioncore

EMBRACEABLE YOU

The key to cloud infrastructure is abstraction to the point that it ‘just doesn’t matter,’” Cohen wrote. “Your infrastructure is always available and completely fault tolerant. Think more along the lines of the iPhone application delivery model (App Store), and less like the desktop application models of the past. The companies that will succeed are the ones who embrace this new hybrid, Internet-centric model. More simply, the cloud is the computer.”

A cloud doesn’t necessarily even need virtualization. This has been proven by none other than Google officials, who have said they do not virtualize production hardware and instead use a job scheduling system of Google’s own design to manage its many thousands of servers.

**Virtualization is all about technology**

This one won’t surprise any longtime IT veteran: Sometimes, it’s the people and not the technology that gets in the way. As IBM’s Corman has noted, people and processes are often not ready for the new challenges raised by virtualization.

Even if your virtualization project is a hit, you might become a victim of your own success. Once users realize how easy it is to spin up a VM, they may become more demanding, making it harder for IT to focus on other tasks. Conversely, there may be resistance from users who prefer to stick with physical servers.

Pradel likes to call politics the “eighth layer” of the network stack.

“The political layer is the most difficult thing you have to deal with as far as virtualization goes,” Vizioncore’s Pradel says. “You have members of your business community who I like to call server huggers. They do not want to get rid of their physical machines, even though they will benefit from going to virtual.”

“You have members of your business community who I like to call server huggers. They do not want to get rid of their physical machines, even though they will benefit from going to virtual.”

—George Pradel, Director of Strategic Alliances, Vizioncore
VIRTUALIZATION TECHNOLOGIES COMPARED

By Kenneth Hess and Amy Newman

In this book excerpt, virtualization experts provide a vendor-neutral look at the technology


This article gives you a vendor-neutral, technical overview of the types of virtualization available. We approach the various types of virtualization from an application and performance perspective—in other words, a practical look at each technology and its implication for you. Each section also includes at least two representative examples of that technology.

**Guest OS/Host OS**

Virtualization aficionados perhaps know Guest OS/Host OS as classic or hosted virtualization. This type of virtualization relies on an existing operating system (the host operating system), a third-party virtualization software solution and creation of various guest operating systems. Each guest runs on the host using shared resources donated to it by the host. Guests usually consist of one or more virtual disk files and a VM definition file. VMs are centrally managed by a host application that sees and manages each VM as a separate application.

Guest systems are fully virtualized in this scenario and have no knowledge of their virtual status. Guests assume they are standalone systems with their own hardware. They are also not aware of other guests on the system unless it’s via another guest’s network services.

The greatest advantage of this kind of virtualization is that there are a limited number of devices and drivers to contend with. Each VM (guest) possesses a consistent set of hardware. The major disadvantage is that disk I/O suffers greatly in this particular technology. Nondisk operation speed, however, is near native. Therefore, we tell those who use hosted virtualization to interact with their VMs over the network using Windows Terminal Services (RDP) for Windows VMs or SSH for Unix and Linux systems.

**VMWARE SERVER**

VMware Server is used throughout this book to illustrate virtualization techniques and technologies. It is a free offering from VMware and is considered an introductory package for use in small environments, testing, or for in-
VIRTUALIZATION: THE INS AND OUTS

THE ART OF XEN

Xen 4.x products ... have converted me heart and soul. The graphical interface is intuitive, fast and extremely well thought out. The template engine in the new product is a pleasure to use, and provisioning a new VM with it is fast, fast, fast.

individuals. It has limited usefulness in large environments because of its memory limitations for VMs and sluggish disk performance. VMware Server supports 64-bit machines as hosts and guests.

SUN xVM (VIRTUALBOX)

VirtualBox, which is now Sun xVM VirtualBox, is one of my favorite virtualization packages. Like VMware Server, it is free and cross-platform, but unlike VMware Server, it is open source. With adjustable video memory, remote device connectivity, RDP connectivity, and snappy performance, it may well be the best hosted virtualization package in your arsenal.

VirtualBox is best suited for small networks and individuals for the same reasons as VMware Server.

Hypervisor

A hypervisor is a bare metal approach to virtualization. Bare metal refers to the server system hardware without any OS or other software installed on it. The best way to describe hypervisor technology is to draw a comparison between it and hosted virtualization. At first glance, the hypervisor seems similar to hosted virtualization, but it is significantly different.

A hypervisor is virtualization software that runs an operating system. Conversely, hosted virtualization utilizes an operating system and runs virtualization software as an application. The hypervisor software is installed to the bare metal; then the operating system is installed, which is itself, a paravirtualized VM. The host operating system, if you can call it that, is designated as VM zero.

A new product, VMware ESXi, implements a bare-metal hypervisor without a traditional operating system interface. It installs directly to the hardware in an almost impossibly small 32MB footprint. ESXi must be installed onto hardware that is virtualization optimized. VM management is performed via Direct Console User Interface (DCUI), which is the low-level configuration and management interface performed at the physical console of the server system. The VM kernel allows for remote management via a set of APIs and agents.

CITRIX XEN

Xen versions 3.0 and earlier weren’t particularly interesting to me because they were somewhat difficult to use and didn’t seem to perform all that well for my specific applications. Xen 4.x products, however, have converted me heart and soul. The graphical interface is intuitive, fast and extremely well thought out. The template engine in the new product is a pleasure to use, and provisioning a new VM with it is fast, fast, fast. If you have a need for high-end virtualization, you must check it out.
VMWARE ESX/VMWARE ESXi
Enterprise virtualization at its finest is brought to you by the people who breathed life into PC-based virtualization. ESX is a mature product that is rivaled only by Xen at this level of virtualization. Both products require 64-bit architecture, but ESXi has very special hardware requirements beyond those of ESX. ESXi is now a free product.

MICROSOFT HYPER-V
Microsoft steps up to the plate with its Windows 2008 Server family and Hyper-V virtualization solution where Citrix and VMware fall short: a Windows-based Enterprise virtualization product. Both Citrix Xen and VMware are Linux-based, which means that if you aren’t familiar with Linux or Unix commands, you may be better off using the Microsoft product.

This product, when more mature, promises to be a formidable challenge to VMware’s and Xen’s dominance in the Enterprise virtualization world.

Emulation
Emulation refers to the capability to mimic a particular type of hardware for an operating system regardless of the underlying host operating system. For example, using an emulation virtualization solution, you can install a Sparc version of the Solaris operating system on a non-Sparc host computer. The emulation software runs as an application on the host system, but emulates an entire computer of another platform. The guest operating system has no awareness of its status as a guest operating system or that it is running in a foreign environment.

In some cases, hardware emulation can be painfully slow, but newer technology, updated emulation software and drivers, and faster 64-bit host processors make emulation a viable virtualization option—especially for those who need to develop drivers or technologies for other platforms without a large investment in support staff or hardware for them.

The best examples of hardware emulation software are Bochs and QEMU.

BOCHS
Bochs is a free, open-source, Intel architecture x86 (32-bit) emulator that runs on Unix and Linux, Windows and Mac OS X, but only supports x86-based operating systems. Bochs is a very sophisticated piece of software and supports a wide range of hardware for emulating all x86 processors and x86_64 processor architecture. It also supports multiple processors but doesn’t take full advantage of SMP at this time.

QEMU
QEMU is another free, open-source emulation program that runs on a limited number of host architectures (x86, x86_64, and PowerPC) but offers emulation for x86, x86_64, ARM, Sparc, PowerPC, MIPS and m68k guest operating systems.

MICROSOFT VIRTUAL PC AND VIRTUAL SERVER
Virtual PC is a free virtualization software package from Microsoft. Virtual PC uses emulation to provide its VM environment. These are good solutions for hosting a few VMs on a Windows XP Workstation or Windows 2003 Server. It isn’t a large environment solution by any stretch of the imagination, but it can get some VMs up and running cheaply and in very short order.

VM performance on these products is surprisingly good for Windows VMs. It is difficult, if not impossible, to tell that you are using a VM when connecting over the network. Console performance tends to be a little sluggish at times—so whenever possible, minimize the console and use RDP to connect to your virtualized Windows systems.
Kernel-Level

Kernel-level virtualization is kind of an oddball in the virtualization world in that each VM uses its own unique kernel to boot the guest VM (called a root file system) regardless of the host’s running kernel.

KVM

Linux KVM (Kernel Virtual Machine) is a modified QEMU, but unlike QEMU, KVM uses virtualization processor extensions (Intel-VT and AMD-V). KVM supports a large number of x86 and x86_64 architecture guest operating systems, including Windows, Linux and FreeBSD. It uses the Linux kernel as a hypervisor and runs as a kernel loadable module.

USER-MODE LINUX

User-mode Linux (UML) uses an executable kernel and a root file system to create a VM. To create a VM, you need a user-space executable kernel (guest kernel) and a UML-created root file system. These two components together make up a UML VM. The command-line terminal session you use to connect to the remote host system becomes your VM console. UML is included with all 2.6.x kernels.

Shared Kernel

Shared kernel virtualization, also called operating system virtualization or system level virtualization, takes advantage of the unique ability of Unix and Linux to share their kernels with other processes on the system. This shared kernel virtualization is achieved by using a feature called change root (chroot). The chroot feature changes the root file system of a process to isolate it in such a way as to provide some security. It (chroot) is often called a chroot jail or container-based virtualization. A chrooted program, set of programs, or entire system in the case of shared kernel virtualization is protected by setting up the chrooted system to believe that it is a standalone machine with its own root file system.

The chroot mechanism has been enhanced to mimic an entire file system so that an entire system can be chrooted, hence creating a VM. The technical advantages and disadvantages of shared kernel virtualization are listed next:

Advantages
- Enhanced Security and Isolation
- Native Performance
- Higher Density of Virtualized Systems

Disadvantages
- Host Kernel and Guest Compatibility

The chroot system offers much in the way of enhanced security features and isolation; however, the greatest advantages of shared kernel virtualization are not in its security, although that’s certainly important to consider, but in its performance. With this kind of virtualization, you’ll get native performance for each individual system. Not only does each system perform at native speeds, but you can also have more than the standard number of VMs on a host system.

By standard number, we mean the number that you could logically have on a host system if you used memory as the limiting factor—leaving 1GB for the host and taking the rest of the RAM for VMs.

The limit of the number of chrooted systems you can have on a host system more closely resembles a standalone system supporting multiple applications. If you think of each chroot system as an application instead of a VM, you’ll more accurately allocate resources and enjoy performance that surpasses many other types of virtualization.

The disadvantage of shared kernel virtualization is a big one: All VMs have to be compatible with your running
kernel. In other words, you can’t run Windows operating systems, Solaris, Mac OS X or any other operating system that couldn’t run your system’s kernel on its own. Major Web hosting providers have run this scenario for years so that customers get their own virtual server for their hosting needs. They don’t know that the system is virtual, nor can they contact the host system through their VM.

SOLARIS CONTAINERS (ZONES)
Solaris 10 comes with built-in virtualization. The Solaris 10 operating system, itself, is known as the Global Zone. Solaris Zones are actually BSD jails, each with its own virtual root that mimics a complete operating system and file system. When you create a new zone, a full file system is copied to the new zone directory. Each zone sees only its own processes and file systems. The zone believes that it is a full, independent operating system; only the Global Zone has any knowledge of virtualization.

Each zone essentially creates a clean sandbox in which you may install applications, provide services or test patches. Solaris zones are a scalable, enterprise-level virtualization solution providing ease of use and native performance.

OPENVZ
We use the OpenVZ kernel on my personal Linux server system. The OpenVZ kernel is optimized for virtualization and proves to be extremely efficient at handling VM performance for other virtualization products as well.

On my personal Linux server system, we run VMware Server, Sun’s xVM and QEMU. Before we installed the OpenVZ kernel, we had many CPU-related performance problems with some of my VMs. OpenVZ is similar to Solaris Zones except that you can run different Linux distributions under the same kernel. Various distribution templates are available on the OpenVZ website at openvz.org.

IN THE VIRTUAL TRENCHES
As someone who works with virtualization software on a daily basis, we can give you some pointers, opinions and suggestions for your environment. These are from my experiences; they may be biased, and, as always, your mileage may vary.

For true Enterprise-ready virtualization, you can’t beat Xen or VMware ESX. They are robust, easy to use, well supported, well documented and ready to go to work for you. Hypervisor technology is absolutely the right decision if you need to virtualize multiple operating systems on one host system. They are both costly solutions but well worth the price you pay for the performance you receive. You should use this technology in situations where disk I/O is of major concern.

As to which one of the hypervisor technologies we prefer, we’re afraid that we can’t answer that for you. Either one you choose will serve you well.

Solaris Zones (containers), and any jail-type virtualization, works extremely well for Unix host systems where you want a consistent and secure environment with native performance. Kernel-level virtualization is extremely well suited for isolating applications from each other and the global zone (host operating system). This type of virtualization is an excellent choice for anyone who wants to get acquainted with virtualization for no money, little hassle, and ease of use. We highly recommend this virtualization method for your Solaris 10 systems.

Microsoft Virtual PC and VMware Server are great choices for testing new applications, services, patches, service packs and much more. We use Virtual PC and VMware Server on a daily basis and can’t live without them. We wouldn’t recommend either for heavy production or enterprise use, but for smaller environments, desktops or IT laboratories, you can’t go wrong with these. They’re free, easy to use, du-
rable and can host a wide range of guest operating systems. In this same arena, Sun’s xVM is also very good.

VMware Server and Sun xVM are both available on multiple platforms, whereas Virtual PC is available only for Windows.

We deliberately left out several other virtualization products from this dialog. Either we’ve had less experience with them or less good experience with them than the others mentioned previously, and we don’t want to keep you from investigating them on your own. We are not diminishing their value or importance for viable virtualization solutions, but we just don’t feel qualified to speak for or against them in this context.

**SUMMARY**

This was an overview of virtualization technology from a vendor-neutral perspective. There is always the question of which virtualization software is best. There is no single correct answer to this question unless it is either emotionally based or prejudicial in some way.

All virtualization software does the same thing: virtualize physical machines and the services that they provide. You’ll have to decide what you need from virtualization and then choose the best technology that fits that need—and worry about vendor specifics later. You may also use more than one virtualization solution to solve the various needs within your network.

If you’re going to invest thousands, perhaps hundreds of thousands, in virtualization, you need to experience the software for yourself. Vendors know this and are willing to work with you. Many offer full versions for a trial period. If a trial version won’t work for you, get in touch with the vendor and get the actual licensed software for evaluation.
Don’t be confused! Virtualization and emulation aren’t the same. Here’s how they differ

Virtualization is a buzzword applied to many different technologies: servers, storage, communications, networks. An older, seemingly related term is emulation, but it’s not the same thing. Here we sort out the differences.

Emulation

Emulation is what we do when we try to make one system behave like or imitate a different system. We want to take System A (something we already have) and give it the inputs we would normally use for System B (which we may not have) and have System A produce the same results as System B.

What’s involved is more than a simple translation of commands or machine instructions; compilers and interpreters have done that for years. No, we’re taking complete operating systems, APIs and functions, and we’re having them work on a machine they were never designed for—a machine that may use totally different methods and commands. That this is even possible is sometimes miraculous, but it nearly always carries a high performance price. Emulation entails a high overhead, and it significantly reduces throughput.

If emulation takes such a toll, why bother? Because we might want to do one of the following:

- Run an OS on a hardware platform for which it was not designed.
- Run an application on a device other than the one it was developed for (e.g., run a Windows program on a Mac).
- Read data that was written onto storage media by a device we no longer have or that no longer works.

Emulation is important in fighting obsolescence and keeping data available. Emulation lets us model older hardware and software and re-create them using current technology. Emulation lets us use a current platform to access an older application, operating system or data while the older software still thinks it’s running in its original environment.

The term emulator was coined at IBM in 1957. Before 1980, it referred only to hardware; the term simulation was preferred when talking of software. For example, a computer built specifically to run programs designed for a different architecture would be called an emulator, whereas we’d use the word simulator to describe a PC program that lets us run an older program (designed for a different platform) on a modern machine. Today emulation refers to both hardware and software.
Virtualization

Virtualization is a technique for using computing resources and devices in a completely functional manner regardless of their physical layout or location. This includes splitting a single physical computer into multiple “virtual” servers, making it appear as though each virtual machine is running on its own dedicated hardware and allowing each to be rebooted independently. In storage virtualization, on the other hand, the server regards multiple physical devices as a single logical unit.

A virtual server is a carefully isolated software “container” with its own software-based CPU, RAM, hard disk and network connection. An operating system or application—even the virtual server itself, or other computers on the same network—can’t tell the difference between a virtual machine and a physical machine.

Virtual machines offer the following advantages:

- They’re compatible with all Intel x86 computers.
- They’re isolated from one another, just as if they were physically separate.
- Each is a complete, encapsulated computing environment.
- They’re essentially independent of the underlying hardware.
- They’re created using existing hardware.

IBM developed virtualization in the 1960s so big, expensive mainframes could run multiple applications and processes concurrently. During the 1980s and ’90s, virtualization was largely abandoned in favor of client/server applications and distributed computing. Today’s servers and PCs, however, face many of the same underutilization problems as those 1960s mainframes.

VMware invented virtualization for the x86 platform in the late 1990s. Recently it introduced a product, called Fusion, that lets Windows applications run concurrently on Macintosh computers that use OS X.

Kay is a Computerworld contributing writer in Worcester, Mass. You can contact him at russkay@charter.net.
VIRTUALIZATION OR CLOUD COMPUTING: WHICH WAY TO GO?

By Cindy Waxer, Computerworld

The decision between virtualization and cloud computing should be made on a case-by-case basis

For the past three years, HotSchedules.com Inc., an Austin-based provider of online labor-scheduling services, has experienced annual revenue growth of over 100%—a boom that could have cost it $60,000 in server hardware purchases.

Yet HotSchedules.com has managed to support an ever-expanding network, enhance its in-house server security and deliver maximum uptime to nearly 375,000 users—all while cutting hardware expenditures and stabilizing electricity costs.

The answer was virtualization, an increasingly popular technology as cash-strapped companies aim to cut capital spending on server hardware. Essentially, virtualization is a layer of software that lets companies consolidate several of their in-house servers onto a single piece of hardware. The upshot: the power of dozens of servers for a fraction of the price and space.

But virtualization wasn’t HotSchedules.com’s first stab at boosting its computing power. Back in 1999, it rented server power and storage for a monthly fee from a third-party provider—essentially what marketers today call cloud computing. However, poor customer service, mounting costs and limited capacity prompted the company to finally make the switch to virtualization.

“The vendor provisioned a couple of servers and some memory for us, but that’s all we got,” says Matt Woodings, HotSchedules.com’s CTO. “We received poor customer service, and it just wasn’t the best experience.”

Companies such as HotSchedules.com are fast discovering that there’s no such thing as a one-size-fits-all approach to bolstering your server and storage capabilities. Some argue that virtualization requires high-priced in-house expertise in exchange for greater security. Others applaud cloud computing’s same-day scalability while at the same time questioning its overall reliability.

Even the cost savings loudly touted by well-respected vendors are being hotly debated. A controversial March 2009 study by consulting firm McKinsey & Co. concluded that for large businesses, shifting IT work to the cloud can be more costly (and less reliable) than owning the hardware in-house.
So how can a company decide whether virtualization or cloud computing is a better bet for its needs?

Mark Tonsetic, a program manager at The Corporate Executive Board’s Infrastructure Executive Council in Washington, says the key is to make a selection “on a project-by-project basis, based on the nature of the application or data that’s being supported.” Each project, he adds, should be evaluated using criteria ranging from server workload demands and disaster recovery requirements to security risks and vendor.

In the case of HotSchedules.com, the deciding factor was the need to accommodate growth while keeping expenses down. On the brink of signing a large new client, the company realized that it would have to purchase $60,000 worth of server hardware, doubling the size of its data center, in order to accommodate the staggering computing demands of a sizable restaurant chain.

“That’s a huge expense to outlay right out of the gate,” says Woodings. “So I thought, ‘Well, there’s got to be a better way to do this. I’ve got to get more bang for my buck.’”

Since deploying Microsoft’s Hyper-V virtualization technology in 2008, HotSchedules.com has consolidated 42 physical servers down to four. By multiplying capacity by 10 times, the company has been “able to take on a considerable amount of clients without seeing any kind of increase in monthly expenses,” says Ray Pawlikowski, CEO at HotSchedules.com. For example, the company pays an average of $12,000 per month in energy bills—a figure that Pawlikowski estimates would have spiked to $20,000 a month by adding brand-new servers to the data center.

**Scalability Issue**

Cloud computing may offer low rates for basic packages, but Woodings warns that “once you start adding the bells and whistles that our clients request, then suddenly that price range becomes prohibitive.” These more sophisticated features include iPhone scheduling apps, automated English-Spanish phone service and instant notification of shift changes.

But external cloud services can be a boon to businesses that don’t want the hassles of keeping up with burgeoning data growth.

That was the case for FreshBooks, an online invoicing and time-tracking unit of 2ndSite Inc. in Toronto. The company could either farm out the storage of photos, design logos and spreadsheets for its nearly 1 million customers or continue to contend with huge “infrastructure management headaches,” says CEO Mike McDerment. He knew that constantly scaling FreshBooks’ servers to meet the fluctuating storage needs of its primarily small-business clientele could only spell trouble. So the 33-person outfit signed on to Cloud Files, a service from Rackspace US Inc.

“While we are a technically sound company, scaling vast amounts of data quickly is not core to our business,” says McDerment. “The storage of these [document] files is painful and costly. By taking that piece away from us, [Rackspace] enables us to focus on writing our applications.”

That wouldn’t have been the case with virtualized servers, according to McDerment, who says maintaining the project in-house would have “left us with the same problem of scaling all kinds of different servers to achieve the needs of our customer base.”

Better yet, by dipping its toes into cloud computing with a low-risk application such as document storage and management, FreshBooks reaps the rewards of cloud computing without losing sleep over the odd outage or technical snafu. After all, much-publicized service outages have plagued cloud computing from the beginning. That’s why McDerment carefully weighed the technology’s benefits against the amount of downtime expected in today’s cloud.

14 of 25
Choosing between cloud computing and virtualization is as much about looking to the future as it is about assessing present-day needs. After all, companies need to be able to grow with their computing power and storage capacity for years to come.

SMART CHOICES

service-level agreements.

“People need to go in with open eyes and ask, ‘If cloud computing is supporting my company’s core service, can 45 minutes of downtime be acceptable?’” says McDerment. The occasional cloud outage is a luxury Bill Gillis couldn’t afford, however. Gillis is the manager of clinical application services at Beth Israel Deaconess Medical Center. The Boston-based teaching hospital uses technology from virtualization giant VMware that lets more than 200 private physician practices throughout Massachusetts access electronic health records. Given the highly confidential nature of the medical data, Gillis says the choice was obvious.

“Because it’s patient records, we wanted to own the space where that information resides. Sure, you could have a vendor sign whatever confidentiality agreements you need, but it opens up a risk,” he says. “If a [vendor’s] rogue employee sells all of our AIDS patient data to some medical research company, we’re responsible for that breach of security without having any real recourse. At least with virtualization, it’s our own staff, and we own that data.”

In the end, though, choosing between cloud computing and virtualization is as much about looking to the future as it is about assessing present-day needs. After all, companies need to be able to grow with their computing power and storage capacity for years to come.

“In today’s environment, where capital comes at a premium,” Tonsetic says, “organizations have to look very hard at their capacity needs in the next two or three years, and whether the capital investment they make today is the right decision.”

Waxer is a freelance writer in Toronto. Contact her at cwaxer@sympatico.ca.
VIRTUALIZATION: THE INS AND OUTS

DESKTOP VIRTUALIZATION: POPULAR, BUT CONFUSING

By Kevin Fogarty, CIO

Here’s a look at the five most fashionable flavors

One of the big questions in technology for the last three years has been how end users will adopt desktop virtualization. The answer, at least from some early adopters, seems to be “how won’t we do it?”

A survey released by Enterprise Management Associates in September 2009 found that companies with desktop virtualization projects in place or underway were almost all using more than one method of delivery, ranging from traditional terminal services to server-based applications accessed through a Web browser, according to Andi Mann, vice president of research for the Boulder, Colo., consultancy.

A typical Citrix XenDesktop virtual desktop connection takes up between 56Kbit/sec and 100Kbit/sec of network bandwidth, and can satisfy the needs of many users by running shared operating systems and applications on back-end servers, according to George Thornton, network operations manager for Texas’ Montgomery Independent School District, which standardized two of the three schools it opened this fall on Citrix virtual desktops.

His users who need more power for graphics-intensive or number-crunching applications can use the same thin-client hardware, but connect to a virtual machine on the server that supports only that one user, along with the additional memory or processing power required. That requires more like 2 megabits of network bandwidth per second, so Thornton kept those connections to a minimum, he says.

With all of the market competition around desktop virtualization, some users are confused about the pros and cons of the various options. Here’s a snapshot of the major desktop virtualization approaches and of the types of situations for which they might be appropriate.

1. REMOTE HOSTED DESKTOPS: This is what most people think of when they think “terminal services.” A server runs one image of an operating system or application and many clients log in to it using connection broker software that is the only part of the software hosted on the client machine. Client machines operate only to show an image on the monitor of the application that user is sharing, and to transmit keyboard and mouse input back and forth.

Advantages: Low cost, high degree of control over data and apps.
**Disadvantages:** Performance depends on the quality of the network connection; display protocols often can’t handle complex graphics; some applications designed for desktops can’t run in shared mode on the server; inflexible for end users, who can’t store data locally, use most peripherals or move data back and forth using thumb drives; doesn’t work when disconnected.

**Example vendor offerings**

**Software:** Citrix XenDesktop; Wyse ThinOS; Microsoft Remote Desktop Services; Microsoft Enterprise Desktop Virtualization (MED-V); VMware View Manager.

**Hardware:** Pano Logic Device, Remote; nComputing thin clients; Wyse thin clients; Sun Ray Ultra-Thin client; Symbiont Network Terminal; Rangee Thin Client

2. **REMOTE VIRTUAL APPLICATIONS:** What you get in every Web application you’ve ever used. Differs from shared desktops in that the only thing required is a browser and standard Web protocols (HTTP, HTTPS, SSL, etc.) to create secure connections and transmit graphics and data. Depending on design of the applications (think Flash downloads) the end-user’s machine may process some of the application’s logic or graphics, or may only light up the monitor and send clicks to the server.

**Advantages:** Doesn’t require that IT control the hardware or software environment of the end user.

**Disadvantages:** Doesn’t allow IT to control the hardware or software environment of the end user, which could affect performance; doesn’t work when disconnected.

**Example vendor offerings:** Citrix XenApps; Microsoft Remote Desktop Services; VMware View; VMware ThinApps.

3. **REMOTE HOSTED DEDICATED VIRTUAL DESKTOPS:** The next step up in power for end users and step down in cost and resource conservation for IT from Web apps or terminal services. Rather than having many users share one instance of the same application or operating system, the server hosts an entire operating system and set of applications within a virtual machine (VM) that is accessible only to that user. The VM could run on a server, sharing resources with other dedicated VMs, or could run by itself on a blade PC. It can either be hosted remotely or streamed. In the streamed scenario, both applications and operating systems can be streamed to the client—downloading parts of the software as the user requires them, and executing on the client machine, using its processing power but not local storage.

**Advantages:** (hosted remotely): Can run applications that balk at running in shared mode; isolates activity of each user to prevent resource constraints.

**Disadvantages:** (hosted remotely): Uses far more bandwidth than shared desktops, and far more hardware on the server. Performance still depends on the quality of the network connection and ability of the display protocol to handle graphics; doesn’t work when disconnected.

**Example vendor offerings:** Citrix XenDesktop; Wyse ThinOS; VMware View; Microsoft Remote Desktop Services; Microsoft Enterprise Desktop Virtualization (MED-V)

**Advantages:** (streamed): Often gives the end user better performance because demanding graphic or other operations execute locally.

**Disadvantages:** (streamed): Requires more powerful client hardware, reducing the cost benefit of virtual desktops;
doesn't work when disconnected.

**Example vendor offerings:** Citrix XenDesktop, XenApp, Xen Provisioning; Wyse TCX; VMware View Manager, Thin Apps, Composer; Microsoft VDI suite.

4. **LOCAL VIRTUAL APPLICATIONS:** Think “Java.” Applications download from the server to the client machine and run there, using local memory and processing power. But they run within a “sandbox” that enforces a set of rules on what the local machine can do and to what it can connect.

**Advantages:** More computing resources and sometimes better performance than remotely hosted applications; less bandwidth consumption; can be used offline.

**Disadvantages:** Less control by IT over the hardware and security of the data.

**Example vendor offerings:** Citrix XenApp, Wyse TCX, VMware Thin App, Microsoft Application Virtualization.

5. **LOCAL VIRTUAL OS:** Present in two major versions. Option one: A client-side hypervisor can create a virtual machine within a laptop or desktop computer, which can function as a completely standalone unit that keeps itself separate from hardware and software on the client machine outside of the VM. Option two: A hypervisor runs on the machine’s BIOS, allowing the user to run multiple operating systems with no “host” OS at all.

**Advantages:** Multiple OSes on a single system; no concerns about OS compatibility, can run on non-traditional VM clients such as smartphones or PDAs.

**Disadvantages:** Potential conflict for resources, relative immaturity of client-side hypervisors leaves security unproven.

**Example vendor offerings:** Citrix Dazzle and Receiver, Wyse PocketCloud, TCX, VMware View Client Virtualization with Offline Desktop (Experimental); Microsoft VDI suite.
How storage virtualization is proving its worth

In just a few short years, storage virtualization, also known as block virtualization, has proven its worth in the large enterprise and traveled that well-worn path from pricey boutique solution to affordable commodity. As a standard feature in all but the most modest mid-tier storage systems, storage virtualization soothes a wide range of storage management woes for small and mid-size organizations. At the same time, dedicated solutions from top-tier vendors deliver the greatest ROI to large shops managing large SANs with intense data availability requirements.

Storage virtualization creates an abstraction layer between host and physical storage that masks the idiosyncrasies of individual storage devices. When implemented in a SAN, it provides a single management point for all block-level storage. To put it simply, storage virtualization pools physical storage from multiple, heterogeneous network storage devices and presents a set of virtual storage volumes for hosts to use.

In addition to creating storage pools composed of physical disks from different arrays, storage virtualization provides a wide range of services, delivered in a consistent way. These stretch from basic volume management, including LUN (logical unit number) masking, concatenation, and volume grouping and striping, to thin provisioning, automatic volume expansion, and automated data migration, to data protection and disaster recovery functionality, including snapshots and mirroring. In short, virtualization solutions can be used as a central control point for enforcing storage management policies and achieving higher SLAs.

Perhaps the most important service enabled by block-level virtualization is non-disruptive data migration. For large organizations, moving data is a near-constant fact of life. As old equipment comes off lease and new gear is brought online, storage virtualization enables the migration of block-level data from one device to another without an outage. Storage administrators are free to perform routine maintenance or replace aging arrays without interfering with applications and users. Production systems keep chugging along.

Virtualization also can help you achieve better storage utilization and faster provisioning. The laborious processes for provisioning LUNs and increasing capacity are greatly simplified—even automated—through virtualization. When provisioning takes 30 minutes instead of six hours and capacity can be reallocated almost on the fly, you can make much more efficient use of storage hardware. Some shops have increased their storage utilization from between 25% and 50% to more than 75% using storage virtualization technology.

SAN: Four Architectural Approaches

In a virtualized SAN fabric, there are four ways to deliver stor-
age virtualization services: in-band appliances, out-of-band appliances, a hybrid approach called split path virtualization architecture, and controller-based virtualization. Regardless of architecture, all storage virtualization solutions must do three essential things: maintain a map of virtual disks and physical storage, as well as other configuration metadata; execute commands for configuration changes and storage management tasks; and, of course, transmit data between hosts and storage. The four architectures differ in the way they handle these three separate paths or streams—the metadata, control and data paths—in the I/O fabric. The differences hold implications for performance and scalability.

An in-band appliance processes the metadata, control and data path information all in a single device. In other words, the metadata management and control functions share the data path. This represents a potential bottleneck in a busy SAN, because all host requests must flow through a single control point. In-band appliance vendors have addressed this potential scalability issue by adding advanced clustering and caching capabilities to their products. Many of these vendors can point to large enterprise SAN deployments that showcase their solution’s scalability and performance. Examples of the in-band approach include DataCore SANsymphony, FalconStor IPStor and IBM SAN Volume Controller.

An out-of-band appliance pulls the metadata management and control operations out of the data path, offloading these to a separate compute engine. The hitch is that software agents must be installed on each host. The job of the agent is to pluck the metadata and control requests from the data stream and forward them to the out-of-band appliance for processing, freeing the host to focus exclusively on transferring data to and from storage. The sole provider of an out-of-band appliance is LSI Logic, whose StoreAge product can be adapted to both out-of-band or split path usage.

A split path system leverages the port-level processing capabilities of an intelligent switch to offload the metadata and control information from the data path. Unlike an out-of-band appliance, in which the paths are split at the host,

**SAN VIRTUALIZERS**

Most storage virtualization solutions today take the in-band, appliance-based approach. The split-path architecture is catching on, while HDS virtualizes internal and external storage in the array controller.

<table>
<thead>
<tr>
<th>VENDOR</th>
<th>PRODUCT</th>
<th>ARCHITECTURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DataCore Software</td>
<td>SANsymphony</td>
<td>In-band appliance</td>
</tr>
<tr>
<td>EMC</td>
<td>Invista</td>
<td>Split path architecture (Brocade and Cisco switches)</td>
</tr>
<tr>
<td>FalconStor Software</td>
<td>IPStor</td>
<td>In-band appliance</td>
</tr>
<tr>
<td>Hitachi Data Systems</td>
<td>Universal Storage Platform</td>
<td>In-band array controller</td>
</tr>
<tr>
<td>IBM</td>
<td>SAN Volume Controller</td>
<td>In-band appliance</td>
</tr>
<tr>
<td>Incipient</td>
<td>Incipient Network Storage Platform</td>
<td>Split path architecture (Cisco switches)</td>
</tr>
<tr>
<td>LSI Logic</td>
<td>StoreAge SVM</td>
<td>Out-of-band appliance or split path architecture</td>
</tr>
</tbody>
</table>

20 of 25
split path systems split the data and the control paths in the network at the intelligent device. Split path systems forward the metadata and control information to an out-of-band compute engine for processing and pass the data path information on to the storage device. Thus, split path systems eliminate the need for host-level agents.

Typically, split path virtualization software will run in an intelligent switch or a purpose-built appliance. Providers of split path virtualization controllers are EMC (Invista), Incipient and LSI Logic (StoreAge SVM).

Array controllers have been the most common layer where virtualization services have been deployed. However, controllers typically have virtualized only the physical disks internal to the storage system. This is changing. A twist on the old approach is to deploy the virtualization intelligence on a controller that can virtualize both internal and external storage. Like the in-band appliance approach, the controller processes all three paths: data, control and metadata. The primary example of this new style of controller-based virtualization is Hitachi Universal Storage Platform.

File Virtualization
Just as block virtualization simplifies SAN management, file virtualization eliminates much of the complexity and limitations associated with enterprise NAS systems. We all recognize that the volume of unstructured data is exploding, and that IT has little visibility into or control over that data. File virtualization offers an answer.

File virtualization abstracts the underlying specifics of the physical file servers and NAS devices and creates a uniform namespace across those physical devices. A namespace is simply a fancy term referring to the hierarchy of directories and files and their corresponding meta-
data. Typically with a standard file system such as NTFS, a namespace is associated with a single machine or file system. By bringing multiple file systems and devices under a single namespace, file virtualization provides a single view of directories and files and gives administrators a single control point for managing that data.

Many of the benefits will sound familiar. Like storage virtualization, file virtualization can enable the non-disruptive movement and migration of file data from one device to another. Storage administrators can perform routine maintenance of NAS devices and retire old equipment without interrupting users and applications.

File virtualization, when married with clustering technologies, also can dramatically boost scalability and performance. A NAS cluster can provide several orders of magnitude faster throughput (MBps) and IOPS than a single NAS device. High-performance computing applications, such as seismic processing, video rendering and scientific research simulations, rely heavily on file virtualization technologies to deliver scalable data access.

**NAS: Three Architectural Approaches**

File virtualization is still in its infancy. As always, different vendors’ approaches are optimally suited for different usage models, and no one size fits all. Broadly speaking, you’ll find three different approaches to file virtualization in the market today: platform-integrated namespaces, clustered-storage derived namespaces and network-resident virtualized namespaces.

Platform-integrated namespaces are extensions of the host file system. They provide a platform-specific means of abstracting file relationships across machines on a specific server platform. These types of namespaces are well suited for multisite collaboration, but they tend to lack rich file controls and of course they are bound to a single file system or OS. Examples include Brocade StorageX, NFS v4 and Microsoft Distributed File System (DFS).

Clustered storage systems combine clustering and advanced file system technology to create a modularly expandable system that can serve ever-increasing volumes of NFS and CIFS requests. A natural outgrowth of these clustered systems is a unified, shared namespace across all elements of the cluster. Clustered storage systems are ideally suited for high performance applications and to consolidate multiple file servers into a single, high-availability system. Vendors here include Exanet, Isilon, Network Appliance (Data ONTAP GX) and HP (PolyServe).

Network-resident virtualized name-spaces are created by network-mounted devices (commonly referred to as network file managers) that reside between the clients and NAS devices. Essentially serving as routers or switches for file-level protocols, these devices present a virtualized namespace across the file servers on the back end and route all NFS and CIFS traffic as between clients and storage. NFM devices can be deployed in band (F5 Networks) or out of band (EMC Rainfinity). Network-resident virtualized namespaces are well suited for tiered storage deployments and other scenarios requiring non-disruptive data migration.

File and block storage virtualization may be IT’s best chance of alleviating the pain associated with the ongoing data tsunami. By virtualizing block and file storage environments, IT can gain greater economies of management and implement centralized policies and controls over heterogeneous storage systems. The road to adoption of these solutions has been long and difficult, but these technologies are finally catching up to our needs. You will find the current crop of file and block virtualization solutions to be well worth the wait.
I recently acquired the VCP Certification. Although that certification is centered, of course, on VMWare, I would like to talk more specifically about virtualization technologies for all the major vendors: Microsoft (Hyper-V), VMWare (vSphere) and Citrix (XenServer), rather than vendor-specific information. I think this will help anyone out there new to virtualization, because virtualization changes the game a bit.

Microsoft in particular is ramping up more of its virtualization certifications with Hyper-V, so I thought I’d spend some time on what you should concentrate on both for exams as well as implementation of virtualization if you are new to it:

1. **Host Installation and Configuration** — How to initially configure the host operating system will be a good start. Any knowledge you have of Linux will help you immensely with learning Xen or VMware Installations, since they are Linux-based host operating systems. Additionally, study the user/group/permissions model for the host.

2. **Shared Storage Concepts** — If you’ve never dealt with it, shared storage will be the greatest challenge for you. SAN knowledge and NAS knowledge is crucial. You don’t need to know it well enough to be a SAN admin, but you will need to know private vs. public LUNs, the intricacies of fibre failover, iSCSI concepts, such as software vs. hardware iSCSI initiators, and NAS concepts, including NFS as a file-sharing mechanism, which could prove to be a challenge to MS-only admins.

3. **Guest OS Migration** — This has become a “must have” technology when it comes to balancing dynamic workloads in the data center, so this will take up a lot of your study time. Each of the major vendors has subtle differences with how its guest virtual machines are moved from one host to another.

4. **Overall Virtual Machine Management** — How guest operating systems are configured and managed will be a big part of any certification. For example, what OS’s are officially supported, how to add hardware to an underlying guest OS, how to do mass rollouts of virtual machines, and hardware allocation and management concepts. Also, as part of this, learn how to measure performance and make decisions based on the according data you find.
5. **Virtual Networking and Switching** — All the major vendors have some level of internal switching, and you’ll need to know how that is done. Additionally, you’ll need to look at how external switching is done, such as link aggregation and how NIC hardware failover is done.

6. **Snapshotting** — All of the major virtualization technologies have a way of saving some intermediary form of the guest operating system. This is used for software testing, among other things.  

7. **Backup and Recovery** — Snapshots should not be used for backup and recovery with virtualization. This is another concept that is different than in physical environments, so you will need to look this over as well.

Although the above is not meant to be an all-inclusive list, it is a pretty inclusive list of the most challenging concepts that are unique to virtualization.  

What I have found when it comes virtualization in general, is that it is more about shattering your preconceived notions about physical environments, rather than your ignorance of virtualization. That will work against you in your learning more than anything.

If I were you, even if you don’t plan on getting certified in virtualization, start learning it now because if you don’t you will really be limiting your career options in the future.

And that’s putting it nicely. •

*Sullins is a technical trainer for New Horizons.*