DATA CENTER STRATEGIES

In-Depth Research Overview

Let's Get Virtual: A Look at Today's Server Virtualization Architectures

v1, May 14, 2007

AUTHOR:

Chris Wolf
(cwolf@burtongroup.com)

TECHNOLOGY THREAD:

Server Virtualization

CONCLUSION:

Increasing power demands and space limitations in the data center have begun to transition server virtualization technologies from luxuries to necessities. Server virtualization provides a path toward server consolidation that results in significant power and space savings, while also offering high availability and system portability. Today, vendors are building hardware and software platforms that can deliver virtualization solutions at near-native performance. To get the most out of virtualization technologies, keep in mind that the answer to every consolidation or availability problem may not be a single virtualization technology, but instead a combination of complementary solutions.

BURTON GROUP 7090 Union Park Center · Suite 200 · Midvale, Utah · 84047 P 801.566.2880 F 801.566.3611

www.burtongroup.com

Publishing Information

Burton Group is a research and consulting firm specializing in network and applications infrastructure technologies. Burton works to catalyze change and progress in the network computing industry through interaction with leading vendors and users. Publication headquarters, marketing, and sales offices are located at:

Burton Group
7090 Union Park Center, Suite 200
Midvale, Utah USA 84047-4169
Phone: +1.801.566.2880
Fax: +1.801.566.3611
Toll free in the USA: 800.824.9924
Internet: info@burtongroup.com; www.burtongroup.com

Burton Group's *Identity and Privacy Strategies* service provides objective analysis of networking technology, market trends, vendor strategies, and related products. The information in Burton's *Identity and Privacy Strategies* service is gathered from reliable sources and is prepared by experienced analysts, but it cannot be considered infallible. The opinions expressed are based on judgments made at the time, and are subject to change. Burton offers no warranty, either expressed or implied, on the information in Burton's *Identity and Privacy Strategies* service, and accepts no responsibility for errors resulting from its use.

If you do not have a license to Burton's *Identity and Privacy Strategies* service and are interested in receiving information about becoming a subscriber, please contact Burton.

Table of Contents

Synopsis 5
Analysis 6
  Benefits 6
    Power and Hardware Savings 7
    Consolidation of Logical Resources 7
    Server Portability 8
    Application Failover 8
  Lay of the Land 9
    Host-Based Server Virtualization 9
      Full Virtualization 10
      Paravirtualizaton 11
      Hardware-Assisted Virtualization 12
    OS Virtualization 13
    Comparing Virtualization Architectures 14
    Challenges to the Virtual Infrastructure 16
      Migration 16
      Management 17
      Disk Performance 19
      Impact on Network Infrastructure 19
      Impact on Disk Storage 20
      Impact on the SAN 21
      Backup and Recovery Complexity 22
        The I/O Dilemma 22
        Change in Backup Strategy 23
  Recommendations 23
    Virtualization Application Selection 24
    Candidate Workload Selection 24
      Getting Started 24
      Testing, Training, and Development 25
      Converting Production Systems 25
    Deployment 26
    Management 26
    No Single Solution 26
  Future Challenges 27
    Common Virtual Hard Disk Format 27
    Standardized Management 27
The Details 29
  Server Virtualization Architectures 29
    Host-Based Server Virtualization 30
      Full Virtualization 30

BURTON GROUP 7090 Union Park Center · Suite 200 · Midvale, Utah · 84047 · P 801.566.2880 · F 801.566.3611 · www.burtongroup.com

Virtual Machine Monitor 31
Host OS/Hypervisor 33
Paravirtualization 34
Hardware-Assisted Virtualization 35
OS Virtualization 38
Partitioning Approach 39
Common Architecture 40
Disk Resources 40
Virtual Disks 41
Physical Disks 43
Virtual Network Adapters 44
Bridged 44
Host Only 45
NAT-ed 47
VLAN Support 47
Virtual Hardware 47
Migration Terminology 48
Vendor Reference 49
Conclusion 51

Synopsis

In recent years, server virtualization has evolved from a technology with significant usage in development, training, and test environments to one that also has a viable place in the data center. With these changes, it's no coincidence that information technology (IT) staffers no longer hum to the tune of Olivia Newton John's "Let's Get Physical." Space and power limitations in the data center have fueled a large consolidation movement, with server virtualization and clustering at the forefront.

While virtualization allows organizations to run multiple unique operating systems (OSs) on the same physical host simultaneously, it also offers other benefits in high availability and system portability. Naturally, the benefits come with tradeoffs including potential performance degradation, as well as challenges associated with network and storage area network (SAN) integration, backup, and system management.

The tradeoffs for server virtualization solutions often vary both by virtualization architecture and product vendor. Host-based server virtualization (fueled by market leader VMware) provides excellent options for system consolidation, portability, storage integration, and automated failover. Evolving host-based virtualization architectures such as paravirtualization and hardware-assisted virtualization have led to significant performance improvements. OS virtualization offers superior consolidation ratios, yet does not provide the same flexibility in system portability offered by host-based server virtualization.

Ultimately, deciding on a virtualization platform boils down to how well a product can take advantage of virtualization-enabled hardware as well as integrate with the existing network infrastructure. Support for 802.1Q virtual local area network (VLAN) trunking and N_Port ID Virtualization (NPIV) SAN integration is dependent on both the selected virtualization application and the hardware, whether existing or planned, that will be used to connect virtualized hosts to an organization's resources.

Virtualization adoption in the data center offers several benefits: efficient hardware resource utilization, server portability, and high availability for any system (including applications that do not natively support clustering). To realize those benefits, careful evaluation of prospective virtualization products is paramount. With numerous moving parts (Fibre Channel or Internet Small Computer Systems Interface [iSCSI] SAN, network infrastructure, and data protection and management applications) affected by any conversion to virtualized resources, awareness of each virtualization technology's limitations is essential. There is little room for error when it comes to managing data center resources. Understanding where each virtualization technology is best suited in the data center allows organizations to realize the benefits of virtualization without falling victim to its weaknesses. ●

Analysis

Server virtualization has quickly jumped from a niche to a mainstream technology over the last couple of years. Organizations faced with limited power and space in the data center have looked to new technologies to enable data center expansion while reducing hardware and power requirements at the same time. Today, there are two primary approaches to solving the problems of growth while consolidating: clustering and server virtualization.

Many organizations have turned to clustering as a means to consolidate high-performance applications while also providing high availability and failover support. While clustering is proven in the data center, it has its limitations, like all other technologies. A primary restriction of clustered solutions is application and service support. Deployment complexity and management requirements also preclude clustering solutions from being viewed as viable for some applications and services. Given that many third-party applications do not support clustering, nor have a proprietary cluster model, clustering in itself is often not enough to solve every consolidation problem.

Server virtualization provides a means to consolidate multiple servers onto one or more physical systems. Virtualization technology may also allow the relocation of servers to systems with completely different hardware without any downtime. With server virtualization, each virtualized host becomes a unique virtual machine (VM), with major elements of the VM's hardware emulated. For example, motherboard emulation allows a VM to be copied or moved to a system with a different physical motherboard than the VM's original host. To the operating system (OS) running inside the VM, no difference exists between the hardware seen on one physical host versus another.

On the surface, server virtualization holds tremendous potential for consolidation. Advanced features such as automatic failover, dynamic relocation, load balancing, and consolidated backup have pushed many organizations to embrace and use this technology in production. However, virtualization is not a cure-all solution, and it does have its place. Knowing where to apply virtualization as well as using the correct virtualization architecture for a given situation is critical in ensuring server virtualization's success with mission-critical applications.

Benefits

Ask IT staff members why they are adopting server virtualization in their organization, and they will likely answer, "Because we have to." When that reason is insufficient, an organization should consider the following as justification for migrating server resources to VMs:

- Power and hardware savings
- Consolidation of logical resources
- Server portability
- Application failover

Power and hardware savings has been the driving factor behind most consolidation projects, so that's a good starting point.

Power and Hardware Savings

When IT staffers state that they consolidate via virtualization "Because we have to," that usually implies they have either run out of space or available additional power in their data centers or server rooms.

Independent hardware vendors (IHVs) have chipped in to assist with power and hardware savings by offering blade systems that take up less space, and they have worked to optimize the density of their one rack unit (1U) and 2U server offerings as well. Power improvements such as IBM's Calibrated Vectored Cooling (CVC), which varies the speed of each cooling fan based on system temperature, reduces both system noise and power consumption (up to 37% depending on the product). In addition to cooling fan efficiency, improvements in power supply efficiency result in less wasted energy dissipated as heat and in lower server room cooling costs.

Server density comes with tradeoffs such as limited expansion options. Blade servers, for example, have limited network and storage controller expansion options. Network and disk bottlenecks are the most prevalent in server virtualization. When several virtual machines have to share one or two network or storage controllers, performance bottlenecks are likely. To alleviate bottleneck potential, requisitioned servers must have substantial room available for additional storage and network controllers. For example, a 2U server with dual onboard Gb network interface cards (NICs) and four PCI Express expansion slots could be configured with two quad port Fibre Channel host bus adapters (HBAs) and two quad port Gb NICs. With this configuration, the server would have eight Fibre Channel ports and 10 Gb NIC ports.

When an organization is consolidating to more powerful servers, running a legacy system on a new piece of hardware makes little sense and sometimes is not possible due to driver limitations. This is where server virtualization has left its mark. When multiple VMs are allowed to concurrently run on one physical server unit, the number of physical servers is reduced, and each VM host server more efficiently uses its resources. Instead of a single OS with an average CPU utilization of 10% running on one box, consolidation may allow six or seven similar servers to run on the same system and thus make better use of the available hardware. Of course, consolidation numbers can be higher depending on the average load of each system to be consolidated.

Aside from the cost savings on power, fewer servers will also likely mean lower hardware maintenance, procurement, and support costs over time.

Consolidation of Logical Resources

Adopting server virtualization results in the reduction of physical devices where logical OS resources can be placed in the physical environment. Therefore, some see the consolidation of logical resources

as another benefit to consolidation via virtualization. Having resources on fewer systems may mean fewer hardware dependencies for operating systems and thus fewer places to troubleshoot when operating systems fail.

While consolidation offers hardware savings, it means that the number of managed logical resources will likely remain the same, if not increase. While OS virtualization can reduce the number of operating systems and in turn OS licenses, with server virtualization the number of licensed operating systems will remain the same. In addition, the IT staff will have another OS to worry about—the VM hypervisor or host OS—which may require an additional backup agent, for example, in order to back up the VM configuration files on each VM host. So while the number of physical systems is reduced by consolidation, the number of managed systems (in the case of server virtualization) is not.

Another consideration when consolidating OS resources onto a single system is the possibility that single points of failure on a physical host could impact beyond a dozen VMs. Some have learned this lesson in the past from deploying a nonredundant storage area network (SAN). Those who were less fortunate had to experience the failure of a SAN switch simultaneously taking down multiple servers in the process.

Given lessons of the past, redundancy must remain an important part of any server-virtualized environment. To achieve resiliency, redundant hardware should exist on physical host systems, and a server virtualization product that provides dynamic VM failover should be deployed.

Server Portability

Server portability is another advantage to server virtualization. Without physical hardware dependencies, a VM can be copied and run on multiple physical host systems, or it can be configured to fail over to another physical host in the event that its primary physical host fails.

With server portability, the hardware at a VM's disaster recovery (DR) center does not have to match that of its primary production center. Several methods exist for getting VM data to the DR center, including traditional backup and recovery as well as scheduled or real-time replication. Because it is unlikely that two geographically dispersed production centers would fail simultaneously, virtualization can add to the agility of a common DR center that can respond to the failure of a number of production sites.

Application Failover

Application failover is a benefit that is a result of high availability server virtualization solutions and VM portability. Many information technology (IT) shops have trouble managing critical third-party applications that do not offer any type of failover support. For applications that cannot be clustered, virtualization is a logical choice. Server virtualization platforms that support VM failover allow any

VM to dynamically fail over to another physical host when a problem is encountered. Problems that generate a failover include virtual machine hangs as well as failures on a VM's physical host and its underlying hypervisor.

Some server virtualization solutions support dynamic VM failover using their own proprietary failover intelligence, while others provide failover support via their host operating system's native clustering service. When evaluating server virtualization products, if application failover is a critical issue, the organization should collect information on how each potential product enables dynamic VM failover. Other failover considerations include shared storage requirements (Fibre Channel, Internet Small Computer Systems Interface [iSCSI], etc.) and failover latency, which differ by server virtualization product.

Lay of the Land

There are two primary approaches to server virtualization:

- Host-based server virtualization
- OS virtualization

Host-based server virtualization is used in the vast majority of server virtualization deployments and is described in the next section.

Host-Based Server Virtualization

Host-based (also known as machine-based) server virtualization is the most commonly deployed server virtualization technology today. People often use the term "server virtualization" to refer to host-based server virtualization, which allows multiple virtual machines (VMs) with dissimilar operating systems to concurrently run on the same physical host system.

Server virtualization today contains three distinct, but sometimes overlapping, approaches:

- **Full virtualization:** Supported by market leaders VMware and Microsoft
- **Paravirtualization:** Supported by VMware and select Xen vendors
- **Hardware-assisted virtualization:** Supported by VMware, Microsoft, and Xen vendors on virtualization-aware (Intel Virtualization Technology [VT] or Advanced Micro Devices Virtualization [AMD-V]) hardware platforms

Each of these three architectures is described in the next three sections.

Full Virtualization

Full virtualization represents the first-generation offering of x86/x64 server virtualization. At the start of 2007, full virtualization maintained the lion's share of the market, but its dominance is being challenged by hardware-assisted virtualization. Full virtualization provides complete hardware emulation, which offers a major advantage—total VM portability. This allows a VM running on a Dell server to be relocated to a Hewlett-Packard server without any problems. Normally, hardware and driver incompatibilities on the new system would likely cause a boot failure and possibly several system hangs as the OS tries to load on the new hardware. By emulating a consistent set of system hardware, VMs have the ability to transparently move between hosts with dissimilar hardware without any problems.

The VM portability provided by full server virtualization offers the following benefits:

- Simple DR staging, as no strict hardware requirements exist for DR VM
- Support for VM failover between hosts with dissimilar hardware
- Ability to share a preconfigured VM between multiple hosts
- Simple base VM image support—a single baseline VM image can be maintained for each deployed OS, regardless of the uniqueness of each host system's hardware configuration (many traditional imaging solutions require a unique baseline image for each group of systems with dissimilar hardware)
- Simple deployment of virtual appliances by independent software vendors (ISVs)

Of course, while full server virtualization has many benefits, it also has drawbacks. Hardware emulation comes with a performance price as the virtual machine monitor (VMM) translates instructions between the emulated hardware and the actual system device drivers. Any instruction from within a virtual machine will pass through two sets of drivers: the VM's device drivers and the device drivers of the host system. Performance degradation for emulated devices such as RAM is nominal (usually less than 2%). However, for input/output (I/O)–intensive devices such as network cards and hard disks, emulation comes at a much higher price, with latency ranging anywhere from 8% to 20%, depending on the virtualization application and the requirements of the internal applications running inside the virtual machine. Another drawback of full virtualization involves the processing of privileged instructions. In traditional x86 architectures, OS kernels expect to run privileged code in Ring 0. However, because Ring 0 is controlled by the host OS or hypervisor, virtual machines are forced to execute at Ring 1, which requires the VMM to trap and emulate privileged instructions from the VM, again inducing latency.

The high latency of full virtualization has placed limits on its usage, especially with high I/O applications. For applications that are less I/O intensive, performance degradation due to driver translation remains negligible to users.

While negligible (less than 10%) disk and network I/O latency may be unnoticeable from a user's perspective, the latency will have a direct impact on backup completion time. Due to these

performance limitations, paravirtualization and hardware-assisted virtualization were developed to improve the inherent weaknesses of full virtualization.

Paravirtualizaton

At its inception, paravirtualization required compiling operating systems to be virtualization aware, which allowed them to detect and work with an underlying hypervisor, if one was present. Today, VM guest operating systems are paravirtualized using two different approaches:

- Recompiling the OS kernel
- Installing paravirtualized kernel mode drivers

Adopting paravirtualization by using operating systems with recompiled kernels comes with tradeoffs. Because paravirtualization drivers and application programming interfaces (APIs) must reside in the guest operating system kernel, operating system vendors need to create OS builds that are compatible with paravirtualization. Some vendors (such as Novell) have embraced paravirtualization and have provided paravirtualized OS builds, while other vendors (such as Microsoft) have not. An OS with paravirtualization drivers will be able to detect the existence of a paravirtualization hypervisor or run on native hardware as well.

VMware established much of the groundwork for a common paravirtualization interface with its 2006 release of the open Virtual Machine Interface (VMI) specification. By following the VMI specification, paravirtualization-enabled guest operating systems can detect a hypervisor, regardless of the hypervisor vendor. This level of support enables OS vendors to create a single paravirtualized OS build and have that OS run on any hypervisor that supports paravirtualization.

Because hardware-assisted virtualization offers the same CPU performance of paravirtualized recompiled operating systems without requiring OS kernel modifications, many vendors are leaving paravirtualization behind and focusing their full attention on hardware-assisted virtualization products. However, hardware-assisted virtualization has yet to solve the disk and network latency issues that are common in full virtualization. Disk and network latency can be reduced by using paravirtualized kernel mode storage and network drivers in VMs running on hardware-assisted virtualization platforms.

The first iteration of paravirtualization solved problems in CPU performance, but because hardware-assisted virtualization can solve the same problem, some media pundits are quick to pronounce paravirtualization a dying technology. However, paravirtualization proponents feel that the CPU performance improvements realized with paravirtualization are just the tip of the iceberg. The collaboration between Novell and Microsoft in early 2007 that resulted in paravirtualized network and storage drivers for Windows guest operating systems on Novell Xen platforms is evidence that paravirtualization is here to stay. While too many variables exist (host hypervisor, host hardware, guest OS, guest applications) to make a blanket statement about the performance

improvements of paravirtualization, some vendors have noticed disk and network latency, typically ranging from 8% to 20%, drop to under 2%.

Based on the recent innovations in paravirtualization solutions, paravirtualization should not be looked at as competing with hardware-assisted virtualization; instead, it should be viewed as a complementary technology that provides needed performance enhancements.

Hardware-Assisted Virtualization

Intel and AMD have both aggressively worked to make their processors virtualization aware, resulting in hardware that improves virtual machine performance. As of early 2007, the most significant aspect of hardware-assisted virtualization was that system hardware could interact with virtualization hypervisors and in turn allow the hypervisor's VMM to run at Ring -1. This allows virtualized guest operating systems to process privileged instructions without the need for any translation on the part of the VMM, and it removes the requirement for a paravirtualization-enabled OS to eliminate privileged instruction processing latency.

With Intel and AMD developing hardware-assisted virtualization for their 64-bit platforms, several virtualization vendors have focused new development on virtualization engines that will only run on hardware-assisted, virtualization-enabled x64 platforms. To ensure compatibility with both current and future virtualization platforms, all new server requisitions should support x64 hardware-assisted virtualization. Note that while the industry is moving toward virtualization engines that will only run on x64 systems, the virtualization platforms will continue to support both x86 and x64 virtual machines.

Hardware-assisted virtualization is very likely to emerge as the standard for server virtualization well into the future. While the first-generation hardware that supports hardware-assisted virtualization offers better CPU performance and improved virtual machine isolation, future enhancements promise to extend both performance (such as memory) and isolation on the hardware level. The key to isolation and memory performance lies in dedicating hardware space to virtual machines. This will come in the form of dedicated address space that is assignable to each VM. AMD-V's forthcoming nested paging support will remove the paging bottleneck found in the current shadow paging methodology and in turn improve memory performance. Note that Intel will offer the same functionality, referred to as Extended Page Tables (EPT), in future enhancements to its VT chips.

A fine line will continue to exist between fully virtualizing resources and giving virtual machines direct access to system hardware. While direct access to resources offers native-level performance, portability is sometimes sacrificed. In order to continue to provide VM failover and portability, virtualization vendors will further develop paravirtualized guest operating system drivers that are hypervisor-aware. These drivers will represent synthetic virtual devices such as network cards and storage controllers. The use of synthetic device drivers will allow a guest operating system to see a consistent set of hardware resources even when moved to a different host, while still providing near-native performance.

While AMD and Intel have received most of the accolades for hardware-assisted virtualization, other IHVs have done significant work in this area as well. Storage vendors today ship virtualization-aware Fibre Channel HBAs, for example, and network hardware vendors are also working toward solutions to provide virtualized network interfaces that can run at near-native performance.

Access to virtual hard disks will continue to be vendor centric. The leading server virtualization vendors, VMware and Microsoft, have yet to agree on a common virtual disk format. So while I/O improvements in networking and storage controllers will increase write performance to virtual hard disks, interoperability will remain an issue. With an industry-standard virtual hard disk format, complete virtual disk interoperability across storage controllers and server virtualization platforms would be realized.

With across-the-board participation between server virtualization vendors and IHVs, the failover and portability benefits of server virtualization will remain while the I/O bottlenecks that plagued server virtualization at its onset will continue to fade away.

OS Virtualization

OS virtualization should not be viewed as a competing technology to host-based server virtualization, but rather as complementary. Because OS virtualization is application centric and allows multiple virtual environments (VEs) to share a common operating system, each environment can run with significantly less overhead than a fully virtualized host. Note that with server virtualization, the term "VM" is used to define each virtual server instance, whereas with OS virtualization, the term "VE" is used. Understand that vendor naming varies by product; for example, Sun Microsystems refers to a VE as a "Solaris Container," while SWsoft's Virtuozzo calls a VE a "Virtual Private Server" (VPS).

With OS virtualization, each virtual environment does not require a separate installed OS. This results in no direct OS overhead such as disk space or RAM for the VE.

For memory alone, a VM's requirements can be substantial. For example, suppose that eight virtual machines are running on a single host server and that each of the guest operating systems uses 512 MB of RAM. This means that without counting the application or VMM overhead, the cost of virtualization will be 8 x 512 MB, or 4 GB of RAM. Assuming that each individual OS installation requires 4 GB of disk space, then 32 GB of disk space would be needed on the physical host system to store all of the guest operating systems. When an organization evaluates resource requirements for virtualization, the requirements to roll out host-based server virtualization can quickly add up. This is what has allowed OS virtualization to emerge as another virtualization technology to root itself in the data center.

Another benefit to OS virtualization is that it does not require any drivers or full hardware emulation within the virtual environment. This allows I/O within the virtual environment to run at

near-native performance. Because VEs run as application shells, they offer the same portability as VMs, with no required dependence on host system hardware.

OS virtualization is not perfect and does have drawbacks that will not allow it to fully replace host-based server virtualization in the data center. Because all VEs connect to a shared OS, more than one OS type can never exist on a physical host system. For OS virtualization vendors, this has meant that OS support is limited, with supported operating systems restricted to specific versions of Windows and Linux. It's highly unlikely that OS virtualization vendors will add support for legacy operating systems, making host-based server virtualization the only choice for legacy OS consolidation projects.

Another key question that arises with OS virtualization concerns isolation. Hardware-assisted server virtualization, for example, can communicate with a hypervisor to carve out dedicated address space for each VM. This provides for additional isolation between VMs, as well as between each VM and the host. OS virtualization achieves isolation at the process level by restricting the amount (by percentage) of access that a single VE can have to a specific host OS process or resource, such as the CPU, RAM, or network. These safeguards are in place to provide quality of service (QoS) guarantees of minimum levels of server resources for critical VEs. With OS virtualization, each VE runs as an application at Ring 3; therefore, isolation is provided by the host OS at the application level.

Change control is another issue facing OS virtualization. When patches are applied to the host OS, their impact on each VE must be considered. For example, a patch not supported by an application in a VE may cause the VE to stop responding. OS virtualization vendors have worked to control patch management by deploying operating system templates, which define the system files seen by each VE. This provides the ability to update a host OS and still isolate the changes from the running VEs. VEs could then be duplicated and tested against the OS changes prior to applying the changes to each VE. With OS template support, VEs can be updated according to the organization's existing change control guidelines.

In spite of its differences from host-based virtualization, OS virtualization has made significant inroads in both the high-performance web server and database space. With the light overhead of OS virtualization, organizations have been able to run up to 100 (or more in some cases) isolated web server instances on one physical server. With no reliance on hardware, each web server instance is portable and can be moved to another host in the event of a system failure. Because of the performance benefits and flexibility afforded to OS virtualization, it has found its place alongside host-based server virtualization in the data center.

Comparing Virtualization Architectures

Table 1 compares the four server virtualization architectures.

	Host-based server virtualization			
	Full	Para	Hardware-assisted	OS virtualization
Common roles	• Legacy server consolidation • Training • Testing • Development	• Production servers that run paravirtualized operating systems • Training • Testing • Development	• Production servers • Training • Testing • Development	• High-performance web and database servers that require full isolation and high consolidation ratios
Limitations	• Reduced performance due to higher virtualization overhead	• Limited OS support • No support for legacy operating systems	• Requires server hardware that supports Intel VT or AMD-V	• Does not support concurrently running different operating systems on the same host • No support for legacy operating systems
Isolation	• Each VM runs its own OS • Hardware resource isolation on the physical host is not supported	• Provides the same isolation as full virtualization	• Improved over full and paravirtualization • Supports hardware isolation (address space, memory) for VM resources	• Isolation achieved by running each VE as an application on a shared OS
Performance	• Low: Noticeable degradation in CPU-intensive operations	• Good: No CPU degradation	• Better: No CPU degradation • Disk and network I/O expected to run at near-native performance by the end of 2007	• Best: No CPU, network, or disk overhead
Management	• Point-level tools available by each virtualization vendor for management and monitoring of VMs and physical hosts • Plug-ins available for enterprise management software such as IBM Director and HP OpenView	• Same as full virtualization	• Same as full virtualization	• Few integration options available for enterprise management tools due to low market share of OS virtualization vendors
Patching	• Distributed: Enterprise patch management software required to simplify management	• Same as full virtualization	• Same as full virtualization	• Centralized: Patching for the host and all virtual environments can be applied simultaneously

Table 1: *Comparing Server Virtualization Architectures*

Host virtualization (full, para, and hardware-assisted) defines complete virtual machines, with each VM having an installed operating system and assigned hardware resources. OS virtualization, on the other hand, partitions a single shared OS into numerous virtual environments, with each virtual environment running as an application instance.

Because OS virtualization requires virtual environments to share a common OS, operating system support is limited to OSs released within the past four years. The limited OS support removes OS virtualization as an option for legacy system consolidation. To consolidate several servers with unique operating system requirements to the same physical system, host-based server virtualization is required.

Today, the lower market share of OS virtualization has resulted in limited support from enterprise management tool vendors such as Hewlett-Packard, IBM, and Dell. This means that any centralized management of virtual environments is accomplished by using the OS virtualization vendor's proprietary application.

Challenges to the Virtual Infrastructure

As a newer technology in the IT landscape, server virtualization still faces growing pains. The most challenging issues affecting server virtualized environments today involve:

- Migration
- Management
- Disk performance
- Network and storage integration
- Backup and recovery complexity

Like other developing technologies, the core platform for server virtualization has evolved ahead of the management infrastructure.

Migration

Deciding on virtualization candidates requires more than just drawing server names out of a hat to determine the VM physical host on which to place each virtualized server. When servers are consolidated to VMs on newer hardware, the performance relationship between the new hardware and existing hardware is far from 1:1. In turn, the physical hardware differences further complicate the math involved in sizing systems for consolidation.

When an organization decides to migrate systems to VMs, the first question IT staffers must ask is, "What can we migrate?" followed by, "How much resource consumption should we expect of the migrated VM on the new hardware?" The answers to these questions are not always easy to determine. With that in mind, several software vendors, including PlateSpin, Leostream, VMware,

and Microsoft, have spent considerable time developing tools that can help identify VM candidates and in turn project the number of VMs that can run on a given system. Recommendations are based on collected performance characteristics of the VM candidates. Another method of VM sizing is to take an application-centric approach, which involves determining required host system resources based on the VM's guest OS, installed applications, and expected client load.

Once the VM candidates have been selected and physical host systems staged, the next part of the process is to convert the physical servers into virtual machines. Some organizations create virtual machines by manually recreating each server as a VM. This involves installing the OS and applications, then synchronizing the VM's data with that of the live server. Another approach is to use a physical to virtual (P2V) migration tool to automate the migration process. Because the emulated hardware on a VM will not match the hardware of the physical VM candidate, the OS on the VM candidate must first be prepared to run on the new hardware before its disk data can be copied. While it is possible to manually uninstall system-specific drivers and remove driver references in the OS kernel, the process can be extremely difficult. Accordingly, most organizations use enterprise P2V migration tools such as PlateSpin PowerConvert or Leostream P>V Direct to assist in the migration. Many P2V tools can not only clone a physical machine to a virtual machine, but can also clone one virtual machine format to another. Several tools also allow the administrator to convert a VM back to a physical box. This is useful when staging a common OS for several systems with different hardware. For example, a P2V tool can perform a virtual to physical (V2P) migration to copy a client baseline image to workstations with different hardware configurations.

Similar to deploying a replacement server, P2V cloning allows administrators to load a physical server image onto a virtual machine and then fully test the image to ensure that it is stable. If problems occur, the physical server can remain online until it is cloned to a VM that remains in a stable state.

Management

As mentioned in the "Consolidation of Logical Resources" section of this overview, consolidation to VMs running on fewer hosts will reduce the number of physical systems, but the number of managed systems remains the same. For managing patches and software updates, existing enterprise system management software can still be used. The large IHVs, including IBM, Hewlett-Packard, and Dell, have developed tools for automating VM deployment, as well as plug-ins to their existing management frameworks to provide VM monitoring and maintenance.

The key issues affecting VM management today include:

- Centralized VM monitoring and alerting
- VM sprawl
- Patch and update maintenance

With products that support VM monitoring and failover, alerting of failover events and host server failures is a critical concern. The level of available alerting will likely vary with both the organization's preferred management product and its server virtualization product.

Many organizations that have adopted virtualization employ products from a range of vendors, such as VMware, Microsoft, SWsoft, XenSource, and Novell. One of the challenges of VM host monitoring is having a proprietary management interface for each product. Because of this problem, several vendors have begun to work on management tools that can centrally manage all of the major virtualization products. For example, this level of support is available as part of Novell's ZENworks Orchestrator. Using separate management interfaces for each virtualization product results in added training costs and complexity. With each interface using proprietary architecture and terminology, it could be easy for administrators to confuse processes or make assumptions about techniques that are valid for one virtualization platform but not another.

Aside from patching and updating the guest operating systems, the administrator will also need to update the VM physical host operating systems and hypervisors from time to time. Because the hypervisor controls each VM on a given physical host system, failure of the hypervisor kernel or compromise due to a security vulnerability could be catastrophic. While it's true that with so little code in a hypervisor, there is less software available to compromise, the fact that compromise is possible cannot be ignored. This is why software maintenance and updating of hypervisors will continue to be equally important as patch management for virtualized operating systems. Vendors have come forward with automated updating and patch management for enterprise operating systems and applications. Given this development, server virtualization vendors should be expected to provide the same level of automated patch management for their respective platforms.

In addition to patch management, anti-virus software management remains a significant concern. The existence of host-based server virtualization engines is transparent to anti-virus software, so each virtual machine instance should have anti-virus software installed locally. If the server virtualization software is running on an OS such as Microsoft Windows instead of a dedicated hypervisor, then anti-virus software will be required to run on the physical host system in addition to being installed in each VM. The anti-virus software on the physical host will need to be configured to not scan virtual hard disk (.vmdk, .vhd) files. Inclusion of virtual disk files in real time and scheduled anti-virus scans on a physical host will result in a significant performance hit. Anti-virus auto-protection services offer the greatest threat to virtual disk performance, because they will attempt to scan a virtual disk file each time a write operation to the virtual disk files occurs.

The considerations for anti-virus support on virtual environments running on an OS virtualization engine vary by both OS virtualization vendor and installed anti-virus software program. With SWsoft's Virtuozzo, for example, administrators only need to install anti-virus software on the physical host. Anti-virus scanning of files written to virtual environments will occur as each file is interpreted by the physical host OS.

Disk Performance

Hardware-assisted virtualization and paravirtualization are doing much to address the performance issues that had plagued full virtualization. Still, disk and network I/O latency will remain a primary concern for server virtualization.

For many shops with large I/O-intensive databases and applications, clustering will remain the architecture of choice. While technologies such as OS virtualization can operate within an existing cluster (offering greater portability), it should not be viewed as a replacement but rather as a way to complement the existing cluster architecture.

A crucial design issue with server virtualization involves the placement of virtual disk files. Ideally, virtual disks should be striped across redundant storage such as a redundant array of independent disks 5 (RAID 5) or RAID 6 array. Note that another layer of disk virtualization may also exist, as RAID logical unit numbers (LUNs) are also commonly referred to as "virtual disks."

Striping virtual hard disks across a RAID 5 or RAID 6 offers benefits in both performance and availability. Without striping, storage of multiple virtual disks on a single physical drive will result in noticeable I/O bottlenecks for disk I/O–intensive applications. Also, trying to push access to too many virtual disks through the same controller (Fibre Channel, SCSI, Serial Advanced Technology Attachment [SATA], etc.) can also cause a performance bottleneck. In these instances, use of multiple storage controllers on the VM host is recommended.

For servers with less intensive performance demands, server virtualization today is a logical choice. By the end of 2007, server virtualization software vendors are expected to offer products that provide near pass-through I/O for both disk and network access. I/O improvements are fueled by two distinct, yet complementary innovations: synthetic device drivers (Windows) or paravirtualized drivers (Linux), and virtualization-aware hardware (Fibre Channel HBAs today and NICs in the future). Virtualization-enabled network and storage devices, combined with synthetic or paravirtualized device drivers, will enable I/O-intensive applications to run within host-based server virtualization environments in the very near future. Note that I/O-intensive applications can successfully run in OS virtualized environments today because OS virtualization does not induce any performance degradation from device translation.

Impact on Network Infrastructure

Upon initial inspection, integrating VMs into an organization's network infrastructure probably seems simple enough. Virtual machines can be bridged to one or more NICs on their physical host system and communicate directly with other systems on the network.

However, when one looks past the physical NIC on the host and sees that the VMs are bridged to the network via one or more virtual switches on the host system, the problem becomes clearer. Most

networks have a well-managed virtual local area network (VLAN) architecture. When more virtual switches are added to the network fabric, several questions are raised, including:

- Who manages the switch?
- How is the switch managed?
- Can the switch be integrated with the existing VLANs?

Answers to these questions will vary with each virtualization product vendor, primarily due to the differences in virtual switch support between each vendor. Some server virtualization vendors emulate nothing more than an unmanaged Layer 2 switch. Other vendors such as VMware have added 802.1Q VLAN trunking support to their switches.

Without 802.1Q support, virtual switches are little more than unmanaged Layer 2 switches, with little to no configuration options. When VLAN integration, logical VM isolation, and security are concerns, server virtualization platforms that support 802.1Q VLAN trucking should be considered a requirement. Without 802.1Q support, the only method for providing network isolation for VMs is to configure multiple virtual switches, with each virtual switch connected to a dedicated physical network interface. The boundaries of VM security would then be determined by their virtual switch assignment.

Impact on Disk Storage

Most server virtualization applications provide support for virtual machines that can use virtual hard disks or take ownership of a physical hard disk attached to the physical host. Virtual hard disks are the most popular disk format type. Because virtual disks exist as individual files, virtual disk duplication is as simple as copying a set of virtual disk files from one location to another.

When a virtual disk file is created, administrators have the option to allocate all space to the virtual disk file at once or allow it to expand over time. For optimal performance, administrators should allocate all space to a virtual disk file at the time it is created. For example, if a 16 GB virtual disk is created, the virtualization software would create a single 16 GB file. Alternatively, administrators could keep virtual disk files to a more manageable size by configuring each virtual disk to be split into 2 GB files. In that case, the 16 GB virtual disk would comprise eight 2 GB files. If the virtual hard disk is stored on a legacy file system that supports a maximum file size of 2 GB, then splitting a hard disk into 2 GB files would be a requirement.

When a virtual hard disk is configured to dynamically expand as content is added to it, the disk will become fragmented on the host system's hard drive. Over time, the fragmentation will continually degrade performance of the virtual disk. To regain performance, an administrator would need to power down the VM and then defragment the physical disk on the host. The major server virtualization vendors also provide command-line disk management tools, so after the disk is defragmented, the administrator could use a disk management tool to convert a dynamically expanding disk into a fixed-size (also known as pre-allocated) virtual hard disk. Due to the

performance cost of dynamic growth, dynamically expanding virtual hard disks should never be used in mission-critical virtual machines.

Keep in mind that, although creating fixed-size or pre-allocated virtual hard disks will prevent them from becoming fragmented on the physical host system, fragmentation can still occur within the virtual disk file. File writes and deletions will fragment a virtual hard disk, like any physical hard disk, over time. Running scheduled disk defragmentation operations at periodic intervals or using a disk defragmenter application that defragments disks in real time is the best approach to prevent fragmentation-related disk performance degradation. As is the case with physical servers, a system backup should be run on the VM prior to the start of any defragmentation operation.

When virtualizing applications that need to see resources as specific LUNs on a SAN, it will be necessary to configure a VM to use physical disks instead of virtual disks. Keep in mind that when assigning physical disks to VMs, administrators will need to use LUN assignment and zoning to ensure that no other hosts attempt to mount a physical disk used by a VM. Letting two nonclustered hosts write to the same physical disk, for example, is an easy way to corrupt the data on the physical drive.

Impact on the SAN

Server virtualization integration with the SAN has been rapidly gaining momentum in recent years. By early 2007, many server virtualization software vendors offered some level of Fibre Channel and iSCSI SAN support.

The easiest aspect of SAN integration has been the inclusion of support for SAN drives by the server virtualization vendors. When using drives on the SAN, the initial concern of most administrators centers on data integrity. When assigning physical disks to VMs, administrators must use protective measures to ensure that no other hosts attempt to mount a physical disk used by a VM. Segregation is accomplished using traditional SAN isolation methods such as LUN assignment or zoning.

Additional challenges to SAN integration arise when high availability architectures are implemented that allow for VM failover between multiple hosts. When a failover occurs, any LUNs associated with a particular VM must relocate with the VM. If the VM does not support virtual Fibre Channel HBAs, then the only association between a VM and the SAN exists through its assigned physical host. Therefore, any physical server that could potentially host a specific VM will need access to any LUNs associated with that VM.

Traditionally, a key problem with SAN integration has been the means in which SAN devices are treated by virtual machines. With Fibre Channel SANs, VM access to SAN devices has typically been translated. For example, a Fibre Channel disk connected to the physical host could be mounted as a virtual SCSI disk within the VM. Without full SAN integration from the perspective of the VM, organizations have been limited in the ways that they can address storage and data availability of virtualized environments.

While broadened support for iSCSI has added flexibility in architecting storage for virtual environments, the adoption of N_Port ID Virtualization (NPIV) support by Fibre Channel HBA vendors is opening the door to simpler SAN integration for virtual machines. By early 2007, major HBA vendors such as QLogic and Emulex had begun to add support for NPIV to their HBAs. When deploying VMs in a data center that includes a Fibre Channel SAN, NPIV support by server virtualization vendors should ultimately be viewed as a requirement.

While newer SAN innovations are assisting server virtualization growth, not all organizations have the flexibility to buy the latest and greatest SAN hardware. When administrators plan to connect VM physical hosts to the existing SAN infrastructure, they should take care to note the specific equipment in place, including storage units, switches, and HBAs. Server virtualization support for SAN storage varies by virtualization software vendor. Some vendors will only work with specific SAN hardware and only have drivers for that hardware loaded into their hypervisor. For some hypervisors, manual installation of third-party drivers is not supported. Other hypervisors that run as part of an operating system will support whatever hardware is supported by the OS. If the OS can see it, then so can the hypervisor. Because SAN hardware support varies with each vendor, administrators should take care to inventory the SAN devices they plan to connect to VM physical hosts and ensure that the devices will be supported by the server virtualization software the organization selects.

Backup and Recovery Complexity

While consolidating and virtualizing servers in the data center will save space and server hardware resources, it will almost certainly require an organization to rethink its backup strategy. Before virtualization, servers were backed up through the SAN or via a high-speed LAN using their own dedicated Fibre Channel or network interfaces. Planning server backups usually involved scheduling around backup storage availability and low server utilization.

The I/O Dilemma

In a nonvirtualized environment, the server has full access to disk and network resources. But in a virtualized environment, where multiple VMs share a single server, VMs compete for hardware resources; disk and network I/O may quickly become a bottleneck, especially during backups.

When determining VM placement, the administrator should consider the backup window along with peak production load times. Many server virtualization vendors have developed products that allow VMs to be automatically relocated to another host in the event of system failure or be relocated at scheduled intervals. So while concurrent backups of four servers on a single physical host may not be feasible, moving one or two VMs to another physical host system to facilitate the backups of all four virtual systems may be an alternative.

Change in Backup Strategy

As backup scheduling and VM placement for backups are significant issues for backup operations, another concern of many IT shops centers on how to back up VMs. One clear-cut method for backups that often mirrors the organization's current backup architecture is to install and run backup agent software inside each VM. This allows the VMs to be backed up just as if they were physical servers. Keep in mind that the VM physical host servers will need to be backed up as well. Each physical host server will contain the configuration files that define the hardware and OS settings for each virtual machine. Maintaining consistent backups of VM configuration data will ensure the quick recovery of VMs in the event of a disaster.

Virtualization software vendors have also added support for centralized backups of VMs from the host server. This could allow a single backup job to run on the physical host that captures the data for every VM on the host. Alternatively, enterprise server virtualization software such as VMware ESX Server supports running a consolidated backup from another "proxy" server that has SAN connectivity to virtual machine disk files. When the proxy is used, no CPU cycles on a VM's physical host are used to back up virtual machine data.

Before deploying any centralized backup approach, the organization should check with both its server virtualization software vendor and its backup vendor to determine if the centralized backup fully supports the organization's applications. Not all operating systems or applications (such as database) support live VM backups, which means that the centralized backup program might need to momentarily suspend (freeze) a VM in order to obtain a consistent copy of its data. If the VM cannot have any downtime, this approach is likely not an option. Because support varies by virtualization product, OS, and installed applications, it's important to check with vendors before making final decisions on backup.

VM replication products that offer replication of VM data in real time to another host have provided one other backup alternative. These products (such as Double-Take Software) allow organizations to replicate VM data to an alternate host and then run DR backups from the near-line host.

As this section illustrates, server virtualization will cause an organization to rethink and likely retest its current data protection strategies.

Recommendations

The rapid growth of server virtualization and its acceptance for mission-critical applications has led to a flood of server virtualization products on the market. The growth of available products has resulted in a landscape of products, each of which is at a different maturity level.

Virtualization Application Selection

VMware clearly has the longest tenure in the host-based server virtualization space, with its first server-class product shipping in 2001. This established history has made VMware's products an easy and safe choice when virtualizing mission-critical applications. OS virtualization vendor SWsoft has a similar story, with its virtualization seeds planted in 1999 and a customer base surpassing 10,000 in 2006. Other virtualization software vendors with less mature software, such as Microsoft, Novell, XenSource, and Virtual Iron Software, will need to prove their platform's reliability to ensure its suitability for mission-critical applications. Otherwise, these products should be first evaluated in training, development, and testing environments, with first production adoptions involving non-mission-critical servers.

Candidate Workload Selection

Server virtualization offers several benefits, including system portability, automated failover, and easier application deployment. However, care still must be taken to apply server virtualization to systems where any latency introduced by the virtualization overhead does not impact organizational operations.

Until server virtualization vendors are shipping products that keep network and disk latency to below 1% (expected by the end of 2007), systems that require high I/O performance should not be virtualized. For systems where a full dedicated server is required for performance, load balancing and high availability can be ensured via an OS-based or third-party clustering solution.

Getting Started

Any transition to server virtualization should occur on an incremental basis. Almost all virtualization platforms are easily acquired. Commercial vendors like VMware offer workgroup-class versions of their product at zero cost. Other solutions, such as Xen and Virtuozzo (OpenVZ), are available as open source projects or as part of Linux distributions. While the free products have limitations and may be more suited for small business or departmental roles, they offer the benefit of allowing an organization's staff members to familiarize themselves with virtualization (for more information on free and open source virtualization products, see the "Vendor Reference" section of this overview).

The process of virtualization begins with IT staff becoming sufficiently comfortable with server virtualization technologies. The free and open source products offer an excellent means for staff to gain this experience. To gain experience in data center-specific issues, staff can complete user training and evaluation of an enterprise-class product such as VMware's Virtual Infrastructure.

Testing, Training, and Development

An organization's initial foray into virtualization should begin with virtualizing test and development systems. The simple portability of virtual machines will help to stage baseline test and development systems. Given that VMs can be easily duplicated and replaced within minutes, testers should not have to fear taking chances with any test exercise.

Converting Production Systems

When an organization is looking to virtualize production systems, selecting virtualization products that offer high availability and dynamic failover should be considered a requirement. With server virtualization vendors writing code specific for hardware-assisted virtualization, virtualization physical host server procurement should involve only servers with 64-bit platforms that support hardware-assisted virtualization.

Low-priority production systems should be the first systems converted to VMs, which will allow an organization to become sufficiently comfortable with virtualized production resources with the least amount of risk. Besides gaining comfort with virtualization, other groups (such as storage, backup, and network administrators) responsible for the information system will gain experience in integrating physical and virtual resources.

Low I/O software applications that are required to run on a dedicated system are also ideal initial candidates. Small, under-utilized web and print servers could be considered for initial VM conversion as well. Virtualizing these systems can give them the ability to run on any physical host and also afford the organization the opportunity to consolidate several low-priority application servers onto a single host. P2V conversion tools are highly recommended in converting the physical systems to virtual machines. Whenever possible, converting highly repeated systems is always recommended. For example, converting a set of 10 identical application servers would allow an organization to take a cookie-cutter approach to migration, not to mention enable a server to switch personalities by booting another "identical" server's virtual hard disk. With cookie-cutter server migrations, administrators will gain experience with P2V migration while engendering little to no risk and very predictable results.

Following the initial VM conversion, each production system should be thoroughly tested against existing practices, including software and patch management, backup and recovery, and DR.

Once the enterprise is comfortable with the initial conversion, it can consider more essential systems for VM conversion. This assumes that the benefits of virtualization will increase the availability and reliability of the systems in question.

In hosting environments where an identical service (such as a web service) is hosted for several clients, OS virtualization should be considered, as it offers a better consolidation ratio than server virtualization.

Deployment

As mentioned in the "Converting Production Systems" section of this overview, conversion of physical systems to virtual resources can be a difficult process. Differences in driver requirements between the physical source system and virtualized hardware on the target VM must be taken into account prior to cloning physical disk resources into virtual disks. The time and testing involved in converting each individual host into a virtualized system usually removes manual cloning as a conversion option. Instead, look to automate candidate selection and cloning using a P2V conversion tool. For small-scale conversions, another option is to manually recreate the production server as a VM (install OS, applications, etc.) and restore the production server's data to the VM from backup.

Management

VM management involves not only system updating and monitoring, but also management between virtual and physical resources. If an organization employs 802.1Q VLANs, it should consider only virtualization products that fully support 802.1Q integration.

While single-interface management consoles are in development that promise to manage all major virtualization applications, any "end to end" VM management tool should be carefully compared to vendor-specific tools to ensure that all critical management features are available. Although a single interface for managing an entire virtual infrastructure is ideal, tools that promise this level of integration were in their infancy as of this writing in early 2007 and will need time to reach maturity.

Integration with enterprise data protection tools should also be a consideration in virtualization product selection. Some backup products today can integrate with the centralized backup features of the most popular virtualization engines. Backup compatibility, as well as compatibility with existing or planned SAN hardware, should be thoroughly checked against any potential server virtualization product.

No Single Solution

Several technologies exist to assist in data center consolidation while also offering increased availability. Clustering should remain the solution of choice for high I/O database and e-mail applications. Server virtualization is ideal in areas where better system availability is desired, as well as for consolidating legacy application servers onto fewer host systems.

OS virtualization has to date been proven in environments where service providers are required to host the same application for dozens of clients, but as a technology it is still maturing. OS virtualization vendors are currently working toward realizing the same hardware-level isolation that server virtualization takes advantage of via hardware-assisted virtualization. Once OS virtualization can offer the same level of hardware isolation, and as long as it provides dynamic failover, it will be ready for critical applications in the data center. Until that time, critical applications should remain

on clusters (where supported), or should be provided with expanded availability by deploying a mature server virtualization product.

Future Challenges

Server virtualization has rapidly evolved over the past seven years, but to meet all management, availability, deployment, and integration concerns, more work remains.

Common Virtual Hard Disk Format

Today, two proprietary virtual hard disk formats dominate the market: VMware's .vmdk format and Microsoft's .vhd format. Both VMware and Microsoft list their virtual disk formats as "open," yet maintain complete control of their development. Microsoft's partnerships with XenSource and Novell allow Windows virtual machines stored on .vhd disk images to run on Xen virtualization platforms without any required modifications. While Microsoft may tout this as evidence that it is developing a virtual hard disk standard, precluding market leader VMware from participating in any discussion of a standard does not lead to one format, but two.

The existence of two virtual disk formats places a greater burden on industry resources that support virtualization than is necessary. For example, if a software vendor chooses to distribute its software as a virtual machine appliance, it will have to distribute two versions of its appliance (one as a .vhd and one as a .vmdk).

Another benefit of a common virtual hard disk format would be in virtual disk performance. With a single format, the processing needed to read and write from a virtual disk could be offloaded to iSCSI and Fibre Channel HBAs. Doing so would improve the performance of the physical host system and practically eliminate any of the disk I/O bottlenecks associated with server virtualization.

Backup and data protection software vendors are currently building out support for protecting virtual hard disks, but are limited in the virtualization products they support. With a common disk format, data protection vendors could develop data protection methods that would be viable for any virtual hard disk, regardless of its underlying virtualization engine. When multiple server virtualization products are used in the same environment, backup complexity would be significantly reduced if no differences existed in how each VM could be backed up.

Standardized Management

One of the greatest weaknesses facing server virtualization today is the lack of available management and monitoring tools. Innovations in server virtualization management could grow at a substantial rate if all server virtualization vendors adopted a Common Information Model (CIM)–based

standard for their metadata. At the start of 2007, a virtualization CIM-based standard was under development by the Distributed Management Task Force (DMTF) System, Virtualization, Partitioning, and Clustering (SVPC) Working Group. Vendor acceptance and adoption of this standard would go far toward speeding the development of better virtualization management solutions. ●

The Details

Products that provide virtualization can be viewed as those that abstract the physical boundaries or dependencies of a technology. This allows administrators to manage servers, storage, and network devices without intricate knowledge of their physical makeup.

Storage virtualization has been around for a number of years, beginning with the use of redundant array of independent disks (RAID). RAID provided a means to logically group physical disks and present those groupings as one or more virtual disks to the operating system (OS). Administrators managing the OS did not need to know the underlying components of a RAID volume in order to format and partition it. While today storage virtualization extends well beyond RAID and into Fibre Channel and Internet Small Computer Systems Interface (iSCSI) storage area networks (SANs), the general purpose of storage virtualization—easing the administrative burden of storage and data, remains the same.

Network virtualization is another form of virtualization that has been around for decades. Virtual local area networks (VLANs) have long provided a means to logically subdivide physical network switches. So a host's view of its local switch is not constrained to the physical makeup of the switch in which it is connected, but instead is determined by the logical presentation of the VLAN itself. As with storage virtualization, network virtualization has moved well beyond simple VLANs to many other elements of network infrastructure management.

Today, network virtualization exists on the host as well as the network infrastructure. Network interface card (NIC) teaming has been around for a number of years as a means to logically group multiple physical NICs for the purpose of improving fault tolerance and performance. Sun Microsystems' Crossbow initiative is working to further extend host-based network virtualization so that individual NICs or NIC teams can be subdivided into multiple logical NICs. Each logical NIC can be assigned priority and configured with bandwidth restrictions. Individual services and applications or virtual machines (VMs) can then be associated with any logical network card. This level of control allows administrators to prioritize and control network input/output (I/O) at the host level, thus providing a methodology to guarantee network I/O to critical applications or virtual machines.

While this overview focuses on server virtualization, understanding the evolution of virtualization technologies in other areas of the information system should allow readers to better appreciate the emergence of server virtualization as a whole.

Server Virtualization Architectures

Like network and storage virtualization, server virtualization has existed for several decades. The initial iterations of server virtualization began with IBM's VM/370 mainframe in the 1980's. The VM/370's underlying virtual machine operating system allowed individual systems, or virtual machines, to be accessible to network users and developers.

Failover and load-balanced clusters have also long offered server virtualization. With clusters, a virtual server object can reside on one or more physical cluster nodes simultaneously. When part of a cluster, a virtual server is not defined by its physical hosts, but rather by its logical resource assignments, such as its host name, IP address, disks, and services.

So while by definition, clustering does provide a form of virtualization, server virtualization is commonly accepted as a technology that is used to simultaneously run one or more isolated virtual hosts on a physical host system. There are two general server virtualization architectures:

- Host-based server virtualization
- OS virtualization

The unique elements of these architectures are detailed in the next two sections.

Host-Based Server Virtualization

Host-based server virtualization, which allows one or more VMs to concurrently run on the same physical host, is far and away the most common form of server virtualization today. When information technology (IT) administrators discuss "server virtualization" or "machine virtualization," they are ultimately describing host-based server virtualization. With host-based virtualization, an independent operating system resides in a defined virtual machine.

The logical resource that defines a virtualized server, the VM includes all of the resources typically found in a server such as RAM, a CPU, hard disks, and network cards. Once the VM is defined, deployment of a virtual server is similar to that of a physical server, beginning with the installation of an operating system.

Ultimately, a virtual machine needs access to resources on its host system in order to process operations and to communicate with other systems on the network. The methods in which a VM processes operations on its host vary based on several server virtualization architectures. Today's prevalent architectures include:

- Full virtualization
- Paravirtualization
- Hardware-assisted virtualization

Full Virtualization

Full virtualization, the original x86 virtualization architecture, in early 2007 was the predominant architecture on the market. The general architecture associated with full virtualization is shown in Figure 1.

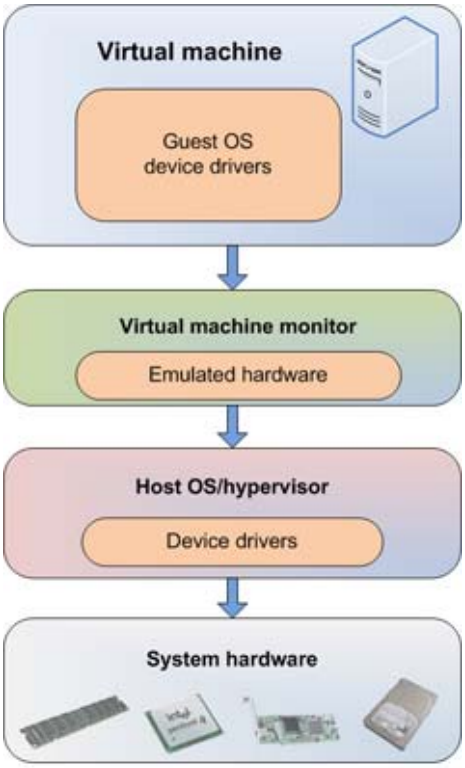


Figure 1: *Full Server Virtualization Architecture*

At the top of the architecture is the virtual machine (VM). As mentioned in the "Host-Based Server Virtualization" section introduction, the VM will contain an installed guest operating system and use the device drivers loaded on the guest operating system to communicate with resources on the physical host system. This communication is facilitated via the virtual machine monitor.

Virtual Machine Monitor

Each virtual machine interfaces with its host system via the virtual machine monitor (VMM). Being the primary link between a VM and the host OS and hardware, the VMM provides a crucial role. The VMM primarily:

- Presents emulated hardware to the virtual machine
- Isolates VMs from the host OS and from each other
- Throttles individual VM access to system resources, preventing an unstable VM from impacting system performance
- Passes hardware instructions to and from the VM and the host OS/hypervisor

When full virtualization is employed, the VMM will present a complete set of emulated hardware to the VM's guest operating system. This includes the CPU, motherboard, memory, disk, disk controller, and network cards. For example, Microsoft Virtual Server 2005 emulates an Intel 21140 NIC card and Intel 440BX chipset. Regardless of the actual physical hardware on the host system, the emulated hardware remains the same.

A significant tradeoff with full virtualization is the performance overhead induced by emulation. This starts with differences in ring architecture in virtual environments (VEs), shown in Figure 2.

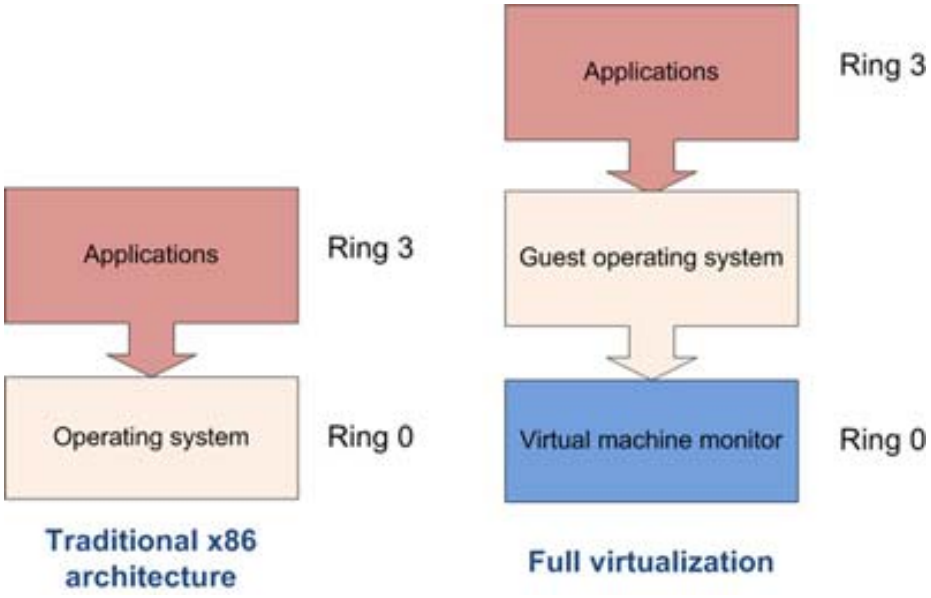


Figure 2: *Comparing Traditional x86 Architecture and Virtualized Resource Access*

Operating systems designed for x86/x64 environments are written to have full access to Ring 0, which is where they run privileged OS instructions. Privileged instructions include OS kernel and device driver access to system hardware. Applications run at Ring 3.

In a virtualized environment, the VMM runs at Ring 0 along with the host operating system's kernel and device drivers. Each VM cannot be given full access to Ring 0 without inducing conflicts, so the VMM runs all VMs at Ring 1. Because privileged instructions within the guest expect to run at Ring 0, the VMM must provide translation in order to "trick" the guest into believing that it has Ring 0 access. If the guest OS kernel did not demand Ring 0 access in the first place, then the translation would not be necessary and thus performance would improve substantially. This is where paravirtualization (described in the next section) comes into play.

The next significant role of the VMM is to provide isolation. The VMM has full control of the physical host system's resources, leaving individual virtual machines with access only to their

emulated hardware resources. The VMM contains no mechanisms for inter-VM communication, thus requiring that two virtual machines wishing to exchange data do so over the network. As the "Hardware-Assisted Virtualization" section of this overview explains, additional isolation can also be provided at the hardware level.

Another major role of the VMM is to throttle host system resource access. This is important, as it can prevent over-utilization of one VM from starving out the performance of other VMs on the same host. Through the system configuration console, system hardware resources such as the CPU, network, and disk access can be throttled, with maximum usage percentages assigned to each individual VM. This allows the VMM to properly schedule access to host system resources as well as to guarantee that critical VMs will have access to the amount of hardware resources they need to sustain their operations.

Host OS/Hypervisor

The primary role of the host operating system or hypervisor is to work with the VMM to coordinate access to the physical host system's hardware resources. This includes scheduling access to the CPU as well as the drivers for communication with the physical devices on the host, such as its network cards. The host OS or hypervisor will also provide management services and either coordinate with a dedicated management server or serve up its own management webpage.

The term *hypervisor* is used to describe a lightweight operating shell that has the sole purpose of providing VM hosting services. The hypervisor differs from a traditional OS in that the OS may be designed for other roles on the network. As it is tailored to VM hosting, a hypervisor solution generally offers better performance and should have fewer security vulnerabilities because it runs few services and contains only essential code. As discussed in the "Hardware-Assisted Virtualization" section of this overview, hypervisors written for hardware-assisted virtualization can embed themselves much deeper into the system architecture and offer superior performance improvements as a result.

Like any traditional OS, a hypervisor-based OS still contains its own operating system code; therefore, maintaining security updates is still important. Unlike a traditional OS, hypervisors are vendor specific, so any needed hypervisor patches or security updates will come directly from the virtualization software vendor.

Because hypervisors are vendor-centric, individual device support often comes directly from the virtualization vendors. Hence, it is important for the organization to ensure that any planned virtualization products are compatible with its existing or planned system hardware. When hosting VMs on a traditional OS such as SUSE Linux Enterprise Server or Windows Server "Longhorn," the organization will find that while the host OS has a larger footprint than a hypervisor, it does provide additional flexibility with hardware devices. With SAN integration, for example, if the host OS does not recognize a Fibre Channel host bus adapter (HBA), the administrator can download the

appropriate driver from the vendor's website. With a hypervisor, the administrator will need to get the driver from the virtualization software vendor, or learn that the device is not supported.

Both hypervisors and operating systems have their strengths and weaknesses. Operating systems provide greater device support than hypervisors, but also require attention to ensure that they are current on all patches and security updates. Hypervisors run on minimal disk and storage resources, but patches and device drivers must come directly from the virtualization software vendor.

One method for combating the latency of full virtualization is paravirtualization, which is described in the next section.

Paravirtualization

Paravirtualization was developed as a means to overcome the emulation requirement of privileged instructions from virtual machines. With paravirtualization, virtualization application programming interfaces (APIs) and drivers are loaded into the kernel of guest operating systems. This allows the guest operating systems to run while fully aware of the virtualization architecture and thus run kernel-level operations at Ring 1. The end result is that privileged instruction translation is not necessary. The architectural differences between paravirtualization and full virtualization exist between the VM and the VMM, as shown in Figure 3.

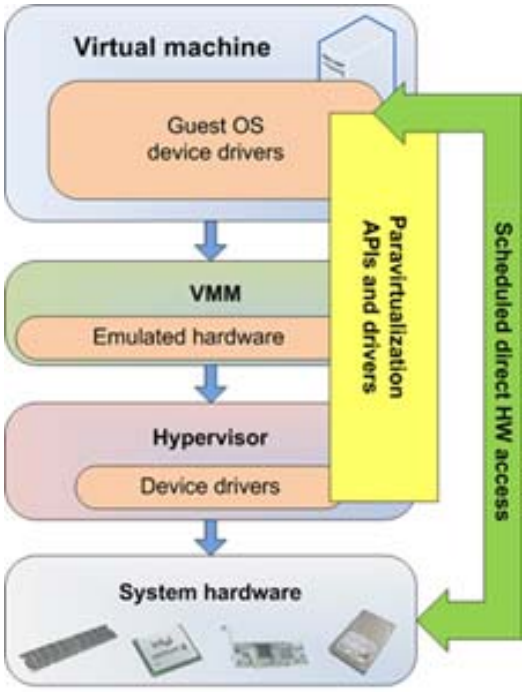


Figure 3: *Paravirtualization Architecture*

Paravirtualization requires the existence of paravirtualization device drivers in the guest VM, the guest VM's OS, the VMM, and the hypervisor. By including paravirtualization APIs within the guest OS kernel, the guest is fully aware of how to process privileged instructions; thus, privileged instruction translation by the VMM is no longer necessary. Furthermore, paravirtualized device drivers such as for network and storage devices are written to communicate with the VMM and hypervisor drivers. Hence, the VMM does not have to present a legacy device to the guest OS and then translate its instructions for the physical host operating system or hypervisor. Removing the heavy emulation requirements from the VMM reduces its workload to merely isolating and coordinating individual VM access to the physical host's hardware resources.

The other benefit of paravirtualization is hardware access. With appropriate device drivers in its kernel, the guest OS is now capable of directly communicating with the system hardware. Note that this doesn't mean that the VM has direct access to all system hardware. In most instances, some system hardware will be available, while other hardware devices will appear as generic representations, as determined by the paravirtualization drivers within the VM. To determine which elements of hardware are paravirtualized and which are available for direct access, consult with the prospective virtualization software vendor.

As with full virtualization, communication with host system resources is scheduled and coordinated by the VMM and hypervisor. Allowing direct I/O operations without emulation and translation will offer significant performance improvements, especially in the greatest traditional bottlenecks to virtualization—network and disk.

As hardware makeup often differs between host systems, it's important to maintain VMM-assisted access to resources via generic drivers as well. This helps to guarantee better VM portability. This is why paravirtualized drivers will remain the best option for coordinating I/O with virtual motherboards, network interfaces, and storage controllers. This allows a VM to relocate to a system with a different physical motherboard and run successfully.

Hardware-Assisted Virtualization

Hardware-assisted virtualization has been fueled by the two leading CPU vendors: Advanced Micro Devices (AMD) and Intel. AMD's version of hardware-enabled virtualization is known as "AMD Virtualization" (AMD-V), while Intel's virtualization support is referred to as "Intel Virtualization Technology" (VT).

Figure 4 shows the general architecture of hardware-assisted virtualization as it compares to full virtualization and paravirtualization.

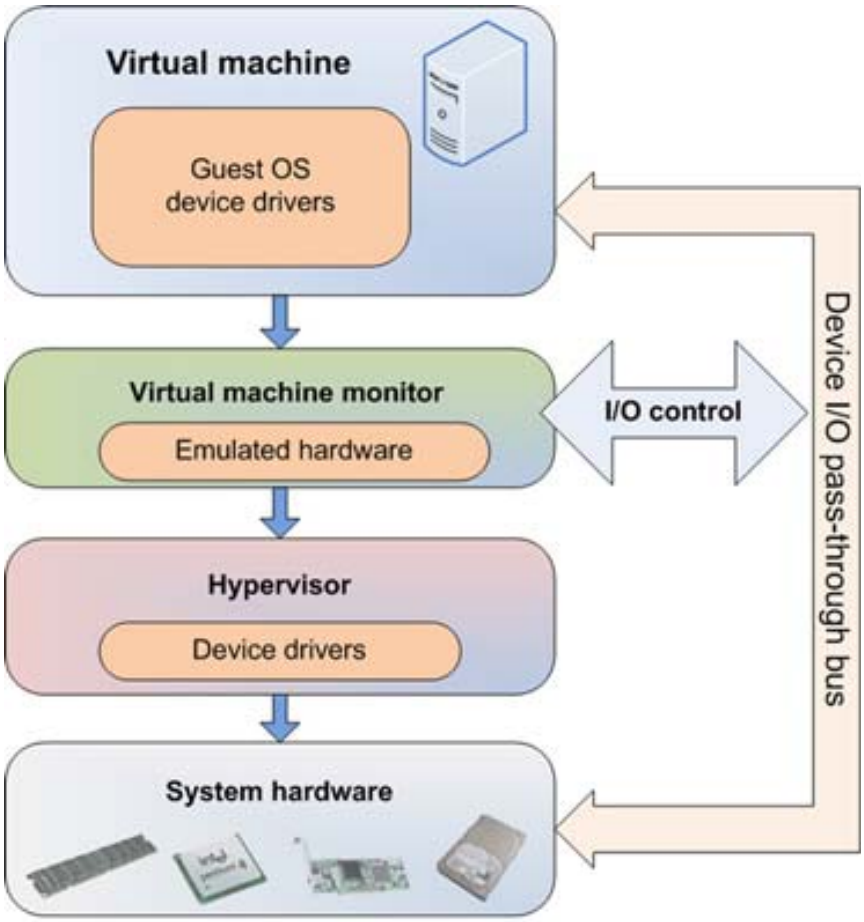


Figure 4: *Hardware-Assisted Virtualization Architecture*

CPUs that support hardware-assisted virtualization are fully aware of the presence of the server virtualization stack. With hardware-assisted virtualization enabled via the system's Complementary Metal Oxide Semiconductor (CMOS) setup, the system will automatically reserve physical address space exclusively for virtual machines. This provides true isolation of virtual machine resources.

Also note the existence of a device I/O pass-through bus in the virtualization stack. This is significant because virtual machines can use this bus to access high I/O devices such as disk and network directly instead of through emulated hardware resources. However, the pass-through bus, also known as the VMBus, is part of the VMM/hypervisor architecture for hypervisors designed to support hardware-assisted virtualization. Keep in mind that while the pass-through bus can provide a clear data path to physical hardware resources, all control information is processed by the VMM, which prevents one VM from taking full control of a hardware resource.

Privileged instruction processing can also operate without the need for emulation or paravirtualization, thanks to the system resource access architecture of hardware-assisted, virtualization-enabled systems (see Figure 5).

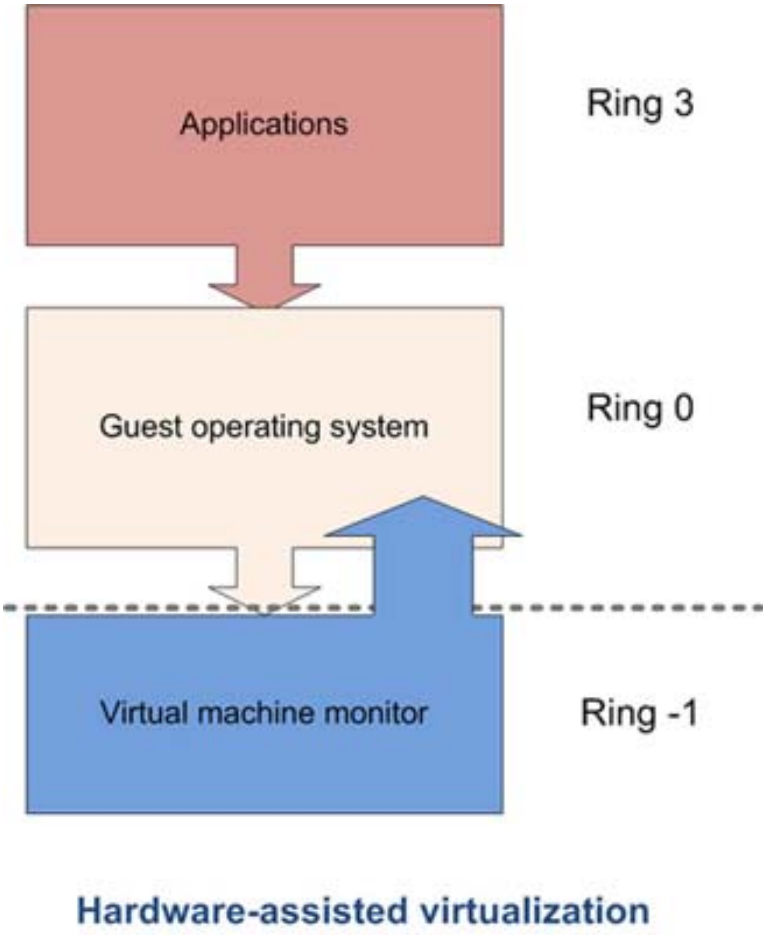


Figure 5: *Guest Operating Systems Provided with Ring 0 Access via Hardware-Assisted Virtualization*

By allowing the hypervisor's VMM to run below Ring 0, guest operating systems can process privileged instructions without the need for any translation on the part of the VMM. This eliminates the previous requirement for privileged instruction translation by the VMM. When an AMD-V or Intel VT platform detects the presence of the VMM, it allows the VMM to run at Ring -1 and in turn runs in super-privileged mode. The VMM maintains control of processor, memory, and system hardware access in order to coordinate access to hardware resources. At the same time, the VMM also allocates specific hardware address space to each VM, thus providing hardware isolation between each VM.

Forthcoming releases of AMD-V and Intel VT chips will improve memory paging support. Full virtualization; paravirtualization; and first-generation, hardware-assisted virtualization rely on Shadow Page Tables (SPT) to translate RAM access for virtual machines. To manage memory, the VMM maintains an SPT for each VM in software. When a VM attempts to write to memory, the VMM intercepts the request, translates it, and stores it in the SPT associated with the VM. The result of the required translation is significant performance overhead (25% to 75%) for memory paging. AMD-V's Nested Page Tables (NPT) support and Intel VT's Extended Page Tables (EPT) support will allow direct translation between guest OS memory addresses and physical host memory addresses. NPT and EPT will enable VM guest operating systems to directly modify their own allocated physical page tables and also handle their own page faults. With the VMM essentially acting as a bridge between a VM and physical memory space on the physical host, the memory performance bottleneck of SPT will no longer exist.

OS Virtualization

OS virtualization is substantially different in approach to host-based server virtualization. With OS virtualization, VEs replace virtual machines. The difference between a VM and a VE is that a VE does not require the installation of an OS, nor does it attempt to emulate any hardware.

This difference in architecture is shown in Figure 6.

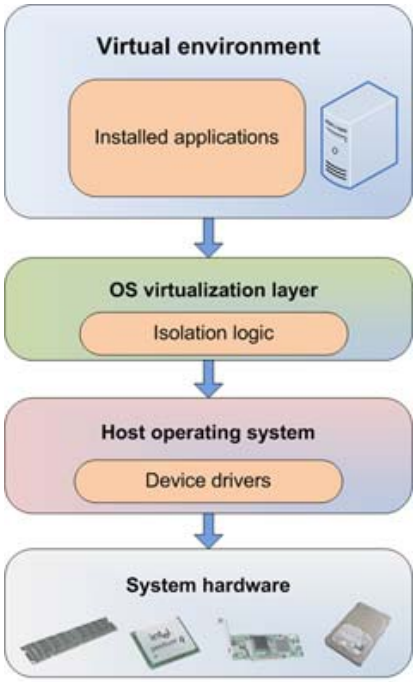


Figure 6: *OS Virtualization Architecture*

The virtual environment has no installed OS, nor any device drivers. Instead, only installed applications and configured network services reside within the VE. A benefit of this architecture is that, without an installed OS, there is no additional OS overhead on the host system. With host-based server virtualization, each VM's OS would require memory, CPU, and disk resources on the physical host system.

The ring architecture of OS virtualization is shown in Figure 7.

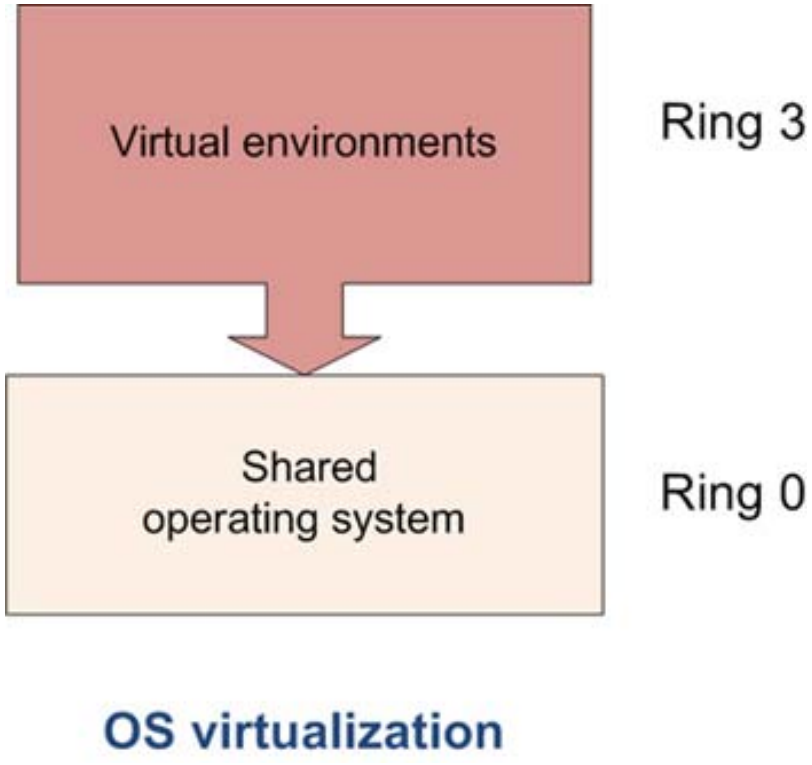


Figure 7: *OS Virtualization Ring Architecture*

Note that with OS virtualization, each virtual environment runs at Ring 3. This allows the host OS to treat each VE as an application, and thus take advantage of the application isolation features that are native to the physical host OS.

Partitioning Approach

OS virtualization essentially partitions a single OS into separate isolated environments. Each defined environment will maintain the look and feel of a unique operating system, with its own dedicated disk resources and unique IP address on the network. From a management perspective, a virtual

environment may have the look of a unique server, but it is managed more like an application. For example, administrators can adjust the RAM assigned to a VE in real time without having to reboot the VE. Additional disk resources can also be allocated without a reboot.

One key difference between OS virtualization and host virtualization is that all virtual environments must share the same operating system. So each unique operating system that the organization plans to virtualize would require its own dedicated physical host system, which is not a problem in instances where running multiple web or database server instances is desired. However, if the goal is legacy server consolidation, host-based server virtualization is a better (and possibly the only) alternative.

With a shared OS, the key to successful and secure OS virtualization lies in isolation. The major OS virtualization vendors offer a high degree of detail about how to provide isolation for each virtual environment. This is a major concern, given that all environments share a single OS. In general, isolation is achieved by setting quality of service (QoS) levels on a process-by-process basis on the physical host system. In doing so, the organization can set access caps (by percentage) to physical host system processes for each VE. Ultimately, this can prevent a VE that has crashed or that is under a denial of service (DoS) attack from affecting the performance of other VEs on the same host.

So host virtualization and OS virtualization have several similarities. Virtual instances are portable, offering the ability to relocate to a server with dissimilar hardware. Because VEs are primarily defined as service or application containers, the underlying hardware on a host system is transparent to the VE. Also, as in host virtualization, a VE will be seen by applications and clients as an independent server with a unique name and IP address. Once again, the major difference between these two architectures is that host-based server virtualization provides for complete virtual machines, with each VM having its own installed operating system and assigned hardware resources.

Common Architecture

Now that this overview has illustrated the different approaches to server virtualization, it will discuss the similarities in how virtualized resources are defined and managed. While terminology may differ from vendor to vendor, the general concepts that define virtualized resources, such as disk resources, are consistent.

Disk Resources

All virtual instances require dedicated disk resources. Just as a physical server owns physical disks (either locally or via a mapped logical unit number [LUN] on the SAN), virtual server instances do the same.

When virtual machines are created, their disk resources can be defined as either virtual disks or physical disks attached to the physical host.

Virtual Disks

Virtual disks exist as independent files that emulate hard disks. The emulation provided by the virtualization engine will allow administrators to format, partition, and manage a virtual disk "file" the same way that they would manage any other disk. Virtual disk files can be easily identified by their file extension, with VMware's .vmdk and Microsoft's .vhd serving as the two most popular formats. The logical representation of a virtual machine and its associated virtual disk files is shown in Figure 8.

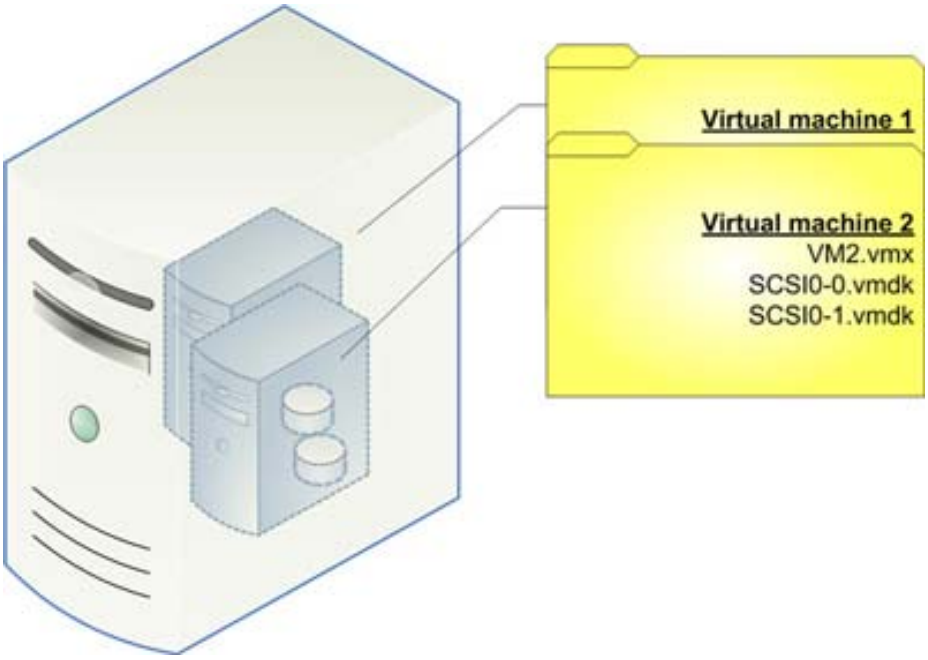


Figure 8: *Virtual Machine 2 Exists as a Single .vmx Configuration File and Two Virtual Hard Disks*

In Figure 8, virtual machine 2 is defined by three files. The VM2.vmx file identifies the virtual machine's configuration, which includes the VM's name and hardware settings, such as defined virtual hard disks. The SCSI0-0.vmdk and SCSI0-1.vmdk files are virtual disk files. So in this example, virtual machine 2's OS would be able to store data on two hard disks. Keep in mind that like a database file, any maintenance to a virtual hard disk file is contained within the file itself and will have no impact on the physical disk where the file is stored. Formatting, partitioning, and file storage within a virtual disk is fully contained within the virtual disk file.

Figure 9 illustrates the relationship between virtual disk files and the physical devices where they reside.

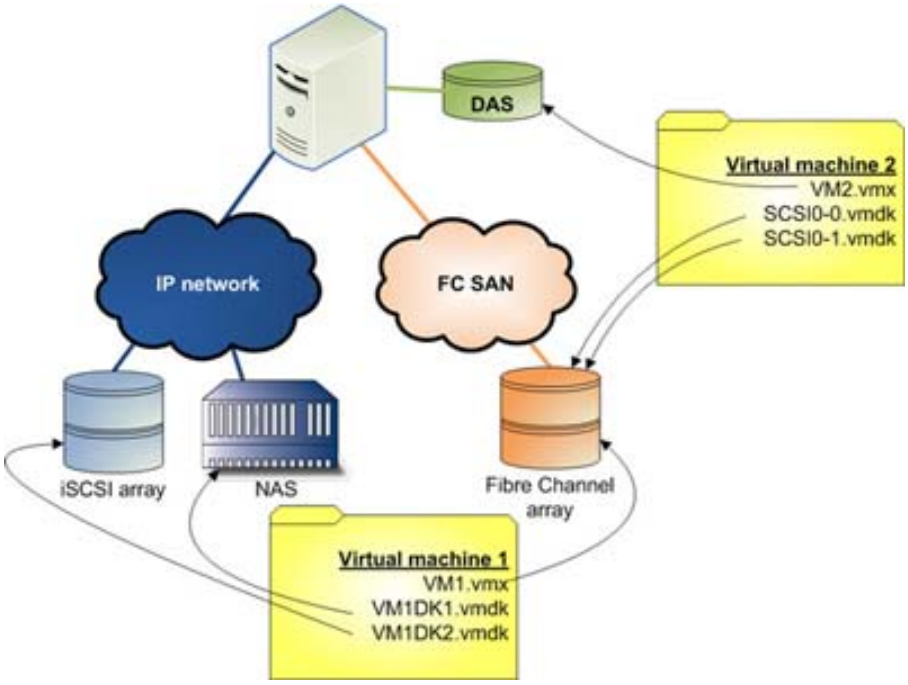


Figure 9: *VM Configuration and Disk Files Stored on DAS, SAN, NAS, and iSCSI Devices*

Specific physical disk support varies by server virtualization software vendor. However, nearly all server virtualization software vendors offer some level of support for direct-attached storage (DAS), network-attached storage (NAS), Fibre Channel SANs (FC SANs), and iSCSI storage. DAS support by most vendors includes the following disk storage interfaces:

- Integrated Drive Electronics (IDE)
- Small Computer Systems Interface (SCSI)
- Serial Attached SCSI (SAS)
- Serial Advanced Technology Attachment (SATA)

To support virtual machine failover and dynamic relocation of virtual machines, virtual machine configuration and disk files must be stored on a shared storage device that's accessible to each virtual machine physical host. In Figure 9, virtual machine 1's configuration file is stored on a Fibre Channel disk, one virtual hard disk file is stored on an iSCSI array, and the other virtual hard disk file is stored via an iSCSI mount to a NAS. Virtual machine 2's files are stored on a locally attached DAS device and on Fibre Channel disks on the SAN. The available storage locations for virtual disk and configuration files are ultimately determined by available storage devices and the limitations of the server virtualization software.

Virtual disks are created at the time a virtual machine is defined and can be added later as well. However, adding a new virtual disk to a virtual machine will require it to be shut down and restarted.

Most server virtualization products allow an administrator to create both IDE and SCSI virtual disks. Virtual disk emulation has yet to be created for other disk types such as SAS and SATA. So while a virtual hard disk can reside on a SAS or SATA disk, the emulated virtual hard disk will be seen by the VM as either an IDE or SCSI hard disk.

With IDE virtual disks, communication is coordinated by the IDE controller on the VM's virtual motherboard. SCSI virtual disks are accessed via an emulated SCSI HBA, such as the LSI Logic PCI-X U320 SCSI HBA. With SCSI disks, the administrator can add more disks to a VM than with emulated IDE virtual disks. For example, if the server virtualization application allows the addition of four virtual SCSI adapters, the administrator could add 15 disks to each adapter (with SCSI ID 7 reserved for the adapter itself). This could allow for a total of 60 virtual SCSI disks on one VM.

The key benefit to a virtual disk is the fact that it is fully portable. Because the disk is nothing more than a collection of files, a virtual disk can be copied over the network or backed up to removable media. A VM can be duplicated by simply copying its virtual disks and related configuration files to another location. For server staging, testing, and training, this provides a major benefit.

Early-generation virtual disks had a high amount of latency due to the virtual disk controller emulation. This often made a physical disk assignment the choice over virtual disks. Today, little to no difference in performance exists between a pre-allocated virtual disk and a physical disk. Note that performance is ultimately determined by how the server virtualization software vendor performs disk emulation and the performance of the underlying hardware. Naturally, a virtual hard disk that is striped in a RAID 6 across five drives would perform better than a virtual disk file located on a single physical disk.

Physical Disks

Many server virtualization applications allow the direct mapping of a virtual machine to a physical hard disk connected to the physical host system. When a physical disk is linked to a virtual machine, it is commonly referred to as a "raw disk" (note that terminology varies by vendor). When bound to a physical disk, a VM is linked directly to a RAID LUN, iSCSI mount, Fibre Channel LUN, or local DAS drive. Depending on the server virtualization software, the physical host system may need to mount a specific LUN and then translate it for the virtual machine. With N_Port ID Virtualization (NPIV) support, it's possible for a virtual machine to directly mount a LUN on a Fibre Channel SAN without having to mount the same LUN to the physical host. NPIV (supported on newer QLogic and Emulex HBAs) allows a single Fibre Channel HBA to be divided into multiple logical HBAs. This lets a logical HBA be transparently bound to a VM, giving each VM its own worldwide port and node name. Virtualization software that supports NPIV enables virtual machines to be fully

aware of LUNs on the SAN and streamline VM failover. Note that VMs can also use iSCSI to directly connect to physical drives on an iSCSI target.

While VMs were originally used to provide better performance than virtual hard disks, organizations today are connecting VMs to physical disks to simplify storage management when VM failover is a concern. When connected to a physical iSCSI target, for example, the intelligence for the connection resides in the iSCSI initiator on the virtual machine. So, as long as network connectivity exists between a VM's new physical host and its iSCSI target, the failover will succeed. Other common uses for binding a VM to a physical disk resource include importing existing operating systems and simplifying migrations by connecting source server physical data disks to a new virtual machine.

Virtual Network Adapters

As with virtual disks, creation of a virtual machine also requires the organization to define how the VM will connect to the network. While terminology varies slightly by vendor, a virtual network interface can be defined as:

- Bridged
- Host only
- NAT-ed

Bridged

When a VM's NIC is set as a bridged NIC, it will bind to its host system's physical NIC. The NIC will have a unique media access control (MAC) address and will need to be assigned an IP address on the production LAN. For this to work, the network virtualization drivers added to the host system's NIC will allow the NIC to behave like a virtual switch. The physical port on the host system's NIC uplinks the virtual switch to the production LAN.

Figure 10 further depicts how bridged networking works.

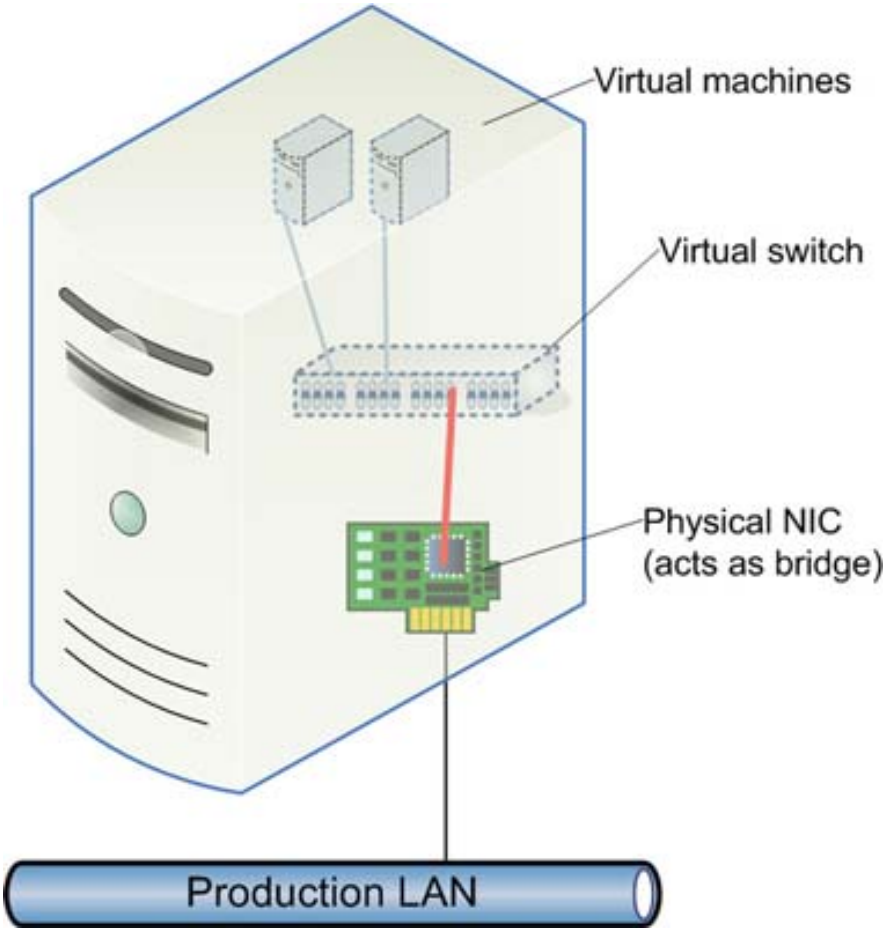


Figure 10: *Bridged Network Configuration*

If the physical NIC in the host is envisioned as a network bridge instead of a network card, the concept of bridged networking is much easier to understand. Frames that are destined for MAC addresses of VMs connected to the host's bridged network are forwarded from the host's physical network interface to the appropriate VM on its internal network. Traffic from the VMs to the production LAN would pass through the virtual switch and out the physical NIC, which again behaves as if it were a network bridge. This configuration provides for relatively simple integration with VMs and the existing network infrastructure.

Host Only

The easiest way to think of the host-only network is that it's like adding a second NIC to any other physical server on the organization's network and using that NIC to connect to hosts on a private

subnet. The only difference with the host-only network is that no other physical devices (NIC, switch, etc.) are needed to facilitate the connections on the private LAN.

When a virtual NIC is defined as a host-only NIC, the NIC communicates to other VMs and the physical host system using a virtual switch on the host. This type of network is common when the organization wants to completely isolate virtual machines from the production network. Reasons to isolate VMs from the production network include:

- Isolating a duplicated VM from its production counterpart
- Running VMs on user workstations while shielding them from the production network
- Securing VMs behind a software firewall running on the host system

If the organization wishes to test an application prior to deployment and wishes to isolate the server from the production network, connecting the VM to a host-only network will provide this level of isolation. Figure 11 illustrates the connections within a host-only network.

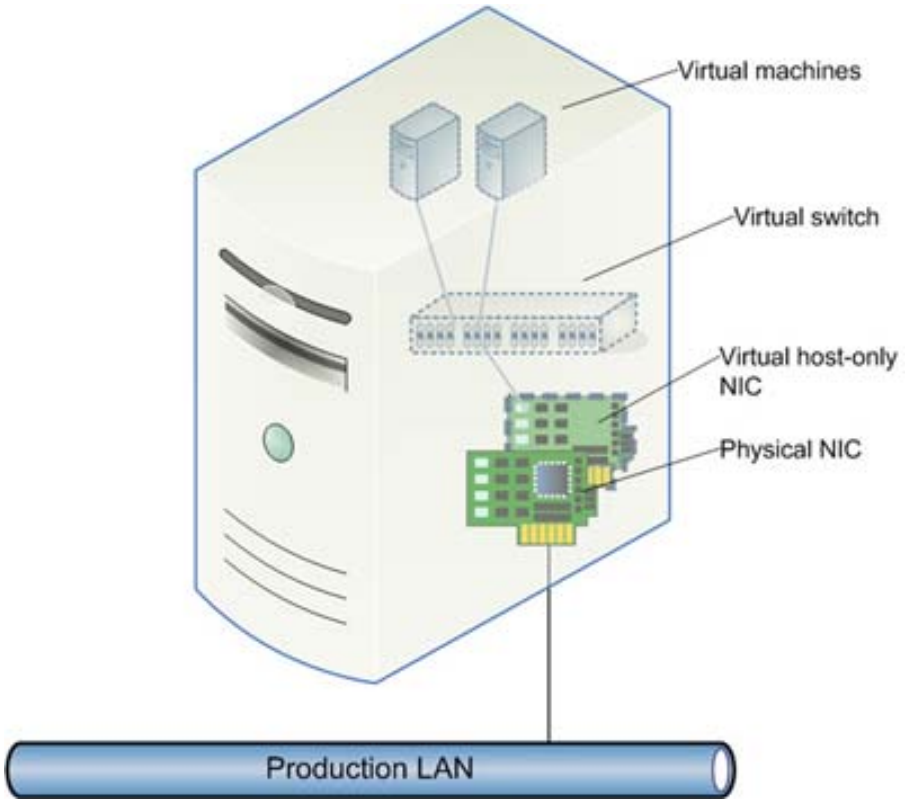


Figure 11: *Host-Only Network Configuration*

When a virtual NIC is installed on a host system, which occurs by default with several virtualization applications, the host treats the virtual NIC like any physical network card. Just like a physical NIC, configuration of the virtual NIC often begins with setting a desired IP address.

Communication over the virtual NIC occurs using an emulated virtual switch. VMs configured to use host-only networking could then communicate with the host system and with each other over the virtual network.

NAT-ed

When a virtual network interface is defined to use Network Address Translation (NAT), VMs are shielded from the production network as they are while on a host-only network. The difference, however, is that the virtualization application provides NAT services for VMs on the virtual network. This will shield a VM from unsolicited access to the production network while allowing the VM to access resources on the production network. With NAT, the VM's private internal IP address is translated to the host system's external public IP address when requesting access to resources outside of the VM's private subnet.

This configuration is useful when the administrator wishes to isolate VMs purely for test purposes but requires VMs to access network resources in order to download updates or install software over the network.

VLAN Support

Support for VLANs varies by server virtualization software vendor. The virtual switches that connect VMs to the LAN are either unmanaged Layer 2 switches or managed Layer 2 switches. With an unmanaged Layer 2 switch, all VMs connected to the same switch will have logical access to one another with no means to logically segment the switch. In order to segregate VMs, administrators need to divide the VMs among two or more internal virtual switches.

Server virtualization applications that offer full 802.1Q VLAN trunking support utilize managed virtual switches that can fully integrate with an organization's VLANs. Having 802.1Q support enables VMs connected to a virtual switch to be recognized and operate within an organization's VLANs.

Virtual Hardware

Of course, any system comprises much more than just a network card or hard disk. Because VMs present a complete machine view to an operating system, all other elements of system hardware must be presented as well. This includes the motherboard, CPUs, video adapter, and all other essential system devices. Most virtualization vendors have chosen to use generic representations for their virtual hardware in order to provide VM portability between hosts with dissimilar hardware.

As virtualization technologies evolve, VMs will continue to see a consistent set of underlying hardware. However, the underlying architecture that moves operations from a VM's virtual hardware to physical host system resources will keep improving.

Often, for an operating system to run within a virtual environment, additional drivers are required for devices such as the virtual video adapter. Implementation of these drivers is normally completed by installing "VM tools" or "VM additions" following the installation of an operating system inside a VM. Installing the VM tools will offer the best performance for the VM and will also enable other advanced features, such as the ability to automatically and gracefully shut down a guest OS inside a VM if the host OS needs to be shut down or rebooted.

Migration Terminology

Three terms are used to describe migration methodologies for virtualization migration products:

- **P2V:** Physical to virtual
- **V2V:** Virtual to virtual
- **V2P:** Virtual to physical

P2V migration is used to convert a physical system to a virtual machine. This process involves creating a VM with virtual hardware (disk, network) similar to the source physical system. Once the VM is staged, the host system's disk data is copied into virtual hard disks.

As with cloning operations such as imaging one physical host and restoring the image on another physical box with different hardware, copying the data is actually the easy part of the process. The most challenging aspect of P2V conversion lies in removing system-specific drivers that will not be compatible in the virtual environment. To help with the conversion process, many third-party imaging vendors have begun to offer P2V conversion solutions. Also, other vendors have emerged that exclusively specialize in VM conversion. Many of the server virtualization product vendors also offer tools that allow users to convert physical systems into virtual machines.

Products that support V2V conversion automate the conversion process between different server virtualization products. For example, an administrator could use a V2V conversion tool to convert a Microsoft .vhd disk image to a VMware .vmdk disk image file. This is helpful when the administrator is evaluating different virtualization products or if he or she would like to evaluate a virtual appliance that is not saved in the product's virtual disk format.

With V2P migration, a virtual machine is clone to a physical system. V2P tools have been a popular resource for administrators who wish to stage and configure client systems as virtual machines and then clone them to physical computers. With V2P, the administrator can maintain a single VM image for each OS type, without having to worry about rebuilding new images each time a new type of system hardware is procured.

Vendor Reference

While "The Details" section of this overview addresses common terms and architecture in the server virtualization space, it does not discuss vendor-specific details. Tables 2 to 5 provide a reference for collecting more information on server virtualization products.

Server virtualization vendors that specialize in host-based server virtualization are listed in Table 2.

Vendor	URL
Microsoft	www.microsoft.com
Novell	www.novell.com
Red Hat	www.redhat.com
Virtual Iron Software	www.virtualiron.com
VMware	www.vmware.com
XenSource	www.xensource.com

Table 2: *Host-Based Server Virtualization Vendors*

Table 3 lists vendors that specialize in OS virtualization.

Vendor	URL
Sun Microsystems	www.sun.com
SWsoft	www.swsoft.com

Table 3: *OS Virtualization Vendors*

Open source and free virtualization platforms are listed in Table 4.

Vendor	Virtualization type	URL
Microsoft	Server	www.microsoft.com/virtualserver
Novell	Server	www.opensuse.org
Red Hat	Server	fedora.redhat.com
Sun Microsystems	OS	opensolaris.org
SWsoft	OS	openvz.org
Virtual Iron	Server	www.virtualiron.com
VMware	Server	www.vmware.com/server
XenSource	Server	www.xensource.com/xen

Table 4: *Free and Open Source Virtualization Products*

Vendors that offer VM conversion and migration tools are shown in Table 5.

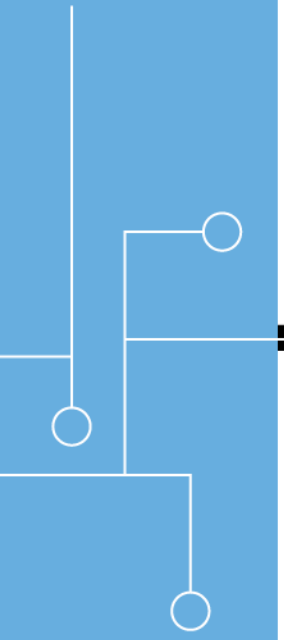
Vendor	URL
Acronis	www.acronis.com
CiRBA	www.cirba.com
Invirtus	www.invirtus.com
Leostream	www.leostream.com
PlateSpin	www.platespin.com
Symantec	www.symantec.com

Table 5: *Virtualization Migration Tool Vendors*

As server virtualization continues to grow, the number of virtualization platform vendors may remain the same, but independent software vendors (ISVs) and independent hardware vendors (IHVs) will very likely offer software that helps in management, conversion, and backup of virtual machines and virtual environments. Consider the references in this section to represent a starting point for research, as they are not meant to capture every product in the rapidly expanding server virtualization landscape. ●

Conclusion

Increasing power demands and space limitations in the data center have begun to transition server virtualization technologies from luxuries to necessities. Server virtualization provides a path toward server consolidation that results in significant power and space savings, while also offering high availability and system portability. Today, vendors are building hardware and software platforms that can deliver virtualization solutions at near-native performance. To get the most out of virtualization technologies, keep in mind that the answer to every consolidation or availability problem may not be a single virtualization technology, but instead a combination of complementary solutions. ●



DATA CENTER STRATEGIES