Virtualization: Architectural Considerations And Other Evaluation Criteria

vmware®

Table of Contents

Virtualization: Architectural Considerations And Other Evaluation Criteria

Introduction

Of the many approaches to x86 systems virtualization available in the market today, the **hypervisor** architecture—in which virtual machines are managed by a software layer that is installed on bare metal—has gained the greatest market acceptance. This fact has translated into rapid growth and a large and expanding customer base for VMware, which pioneered x86 hypervisors in 2001 with the launch of VMware® ESX Server. It is no wonder, then, that the hypervisor market has attracted attention recently from Microsoft[1] and the usual assortment of venture-funded startups, including the recently created XenSource.

Competitive marketing notwithstanding, the facts in this market are as follows:

- **The VMware product architecture is rooted in its experience in solving real-world customer problems.** The choices VMware has made in its hypervisor-based ESX Server reflect the practical focus on offering the highest levels of performance, reliability and compatibility. In contrast, both XenSource and Microsoft have chosen architectural paths that allow them to get products to market more quickly. These products may satisfy a limited set of use cases, but have yet to grapple with the architectural issues of building an enterprise-class hypervisor. As they attempt to broaden their applicability, they will encounter the same real-world issues that VMware did when it first entered the market. The difference, of course, is that VMware solved these problems long ago.

- **VMware offers a wide range of production-tested solutions**, and provides a comprehensive set of innovative technologies to augment the basic partitioning functions of its hypervisor. While an architectural comparison is of interest to those trying to predict the long-term direction of virtualization technology, what ultimately matters to users are the solutions that they deploy based on virtualization. Today VMware offers products that customers are actively using in production deployments to meet their business demands.

- **VMware is the only enterprise-ready hypervisor available.** Product features aside, vendors must answer questions such as: How well will the products work with what the customer already has? How well supported is it? And how manageable is it? Users rightfully demand a certain level of enterprise readiness before they broadly deploy a technology in production. As with solutions, enterprise readiness is a function of product maturity. VMware has customer references that attest to the maturity of VMware based solutions.

- **This paper examines these issues—architecture, solution support and enterprise readiness—in greater detail.** Comparisons focus primarily on XenSource. Microsoft has not released their hypervisor product and therefore can not be compared at this time.

Architectural Considerations

The x86 architecture was never designed for virtualization. Consequently, high-performance virtualization is difficult to achieve. There are three ways to address the problem of virtualizing the x86 architecture:

- **Transparent virtualization** allows operating systems, or particular components of the operating system, to run inside virtual machines without modification.

- **Paravirtualization** requires the operating system to be modified before it can run inside the virtual machine. Depending on the part of the operating system being changed, the modification may be expected and supported by the operating system vendor (e.g., new drivers) or not (e.g., changes to the kernel).

- **Hardware** can provide explicit support for virtualization. To date such support is comparatively rare, but as virtualization has become a standard layer inside enterprise data centers, hardware vendors have responded with roadmaps promising such support.

These methods are not mutually exclusive. In fact, it is a combination of all three, applied to the basic elements of the hypervisor—CPU, memory management and I/O—that will ultimately provide the greatest performance, reliability and compatibility to the end user.

---

[1] *In fact, in an unusual move for the industry giant, Microsoft has gone so far as to virtually abandon its existing non-hypervisor product, Virtual Server.*

CPU virtualization

The CPU virtualization used by VMware employs **direct execution** and **binary translation**. These transparent virtualization techniques ensure that the vast majority of CPU instructions are executed directly, with zero or low performance overhead.[2]   This use of transparent virtualization also ensures that the guest operating systems are run without modification, resulting in a high degree of operating system compatibility.

XenSource's **kernel-based paravirtualization** helps to reduce the overhead that comes from virtualizing certain privileged CPU instructions, by modifying the operating system internals to avoid them altogether. Currently the industry is working on but have not shipped Linux distributions that support paravirtualization. When these are released customers will have to retest and requalify their solutions on these new distributions.

Although hardware support in the form of Intel's Virtualization Technology (VT) and AMD's Pacifica hold the promise of hardware assist for CPU virtualization, it remains to be seen whether the silicon delivered can provide the same optimized performance that software already provides. In other words, although XenSource has announced plans to rely on hardware assistance to run Windows (as there is no paravirtualized version of Windows), it is unlikely to match the performance of the transparent CPU virtualization used by VMware for some years.

As previously noted, transparent virtualization, paravirtualization and hardware assist are not mutually exclusive. VMware has already announced its planned support for Intel's VT and AMD's Pacifica, and was in fact the first vendor to provide a working demonstration using VT at the Intel Developer Forum in April 2005. VMware has also announced its intent to support paravirtualized operating systems when they are available from RedHat, Novell/Suse, and, eventually, Microsoft.[3]   Doing so ensures that VMware customers gain the performance benefits of paravirtualized operating systems like RHEL 5 and SLES 10 (both targeted in mid- to late 2006), all without sacrificing the stability and performance of the ESX Server platform.
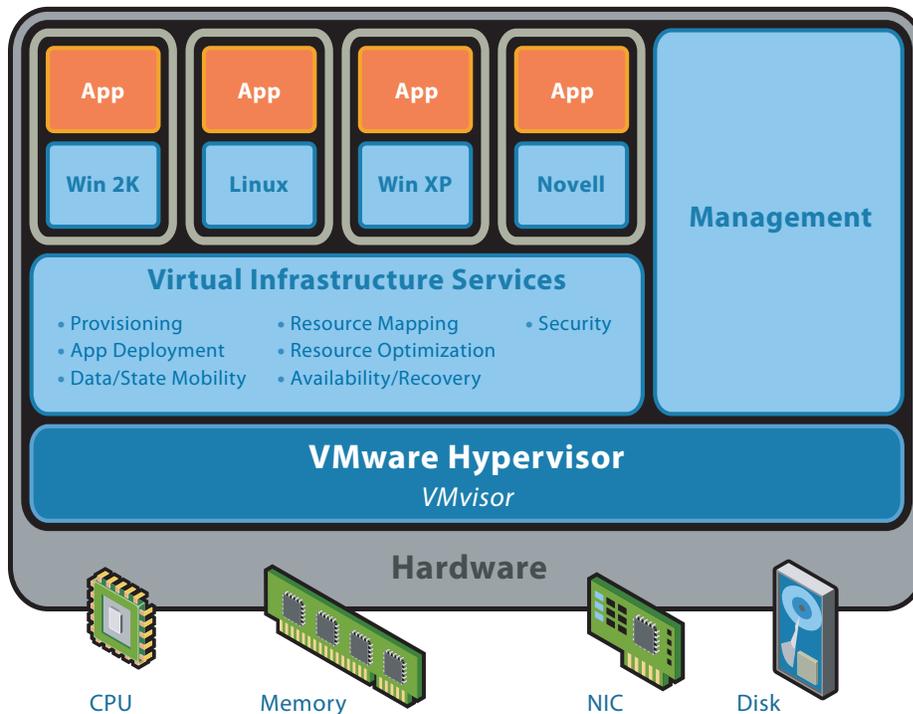

Figure 1: Virtualization Infrastructure Architecture

[2] For the class of privileged CPU instructions that cannot be directly executed, VMware employs a form of binary translation that is small, compact, and serves a well-defined purpose. This simplicity can be contrasted with the complexity of 'optimized binary translators', such as the Dynamo system or the Transmeta processor.

[3] For a vendor that uses transparent virtualization, these paravirtualized operating systems merely represent additional guest OSes that need to be supported. VMware's long tradition of heterogeneous operating system support ensures that adding this support will be straightforward.

Memory Management

As with the CPU, the fundamental question in virtualization of memory is the extent to which the operating system needs to be changed to account for the fact that it is being virtualized, and the extent to which hardware support is required.

The approach used by VMware makes use of **shadow page tables,** which provide a map between the guest operating system's virtual memory pages and the underlying physical machine pages. The shadow page tables are maintained by ESX Server and are used to efficiently virtualize memory access. As a result, neither operating system modification nor hardware support is a requirement, although the latter has some potential for performance enhancement. The implementation of shadow page table used by VMware optimizes caching of guest virtual address translations so that expensive hardware faults resulting from memory access are minimized. Dynamic binary translation is also used to keep the cost of MMU-related operations low. With shadow page tables, the virtual machine's linear addresses can be mapped directly to the real machine addresses, and accesses to memory can occur at native hardware speed.

Shadow page tables also provide another important benefit. Because they insulate guest operating systems from their dependence on specific machine memory, they allow the hypervisor to optimize the use of that memory far beyond what an individual operating system is capable of doing. This "memory overcommitment" is not required for static partitioning, but is a key enabler for all other virtualization-based solutions. (See below for more details.)

The XenSource approach does not use shadow page tables, except on a temporary basis when executing their version of VMotion, the ability to migrate a running virtual machine from one physical host to another without downtime. Instead, it provides partial access for the guest operating system directly to physical memory page tables, through kernel modifications. XenSource claims to have chosen this paravirtualized approach for performance reasons. The reality is probably somewhat different: without either binary translation or hardware virtualization assist, it is not possible to implement shadow page tables in a high-performance manner.[4] As already noted, XenSource has not announced plans to provide binary translation, and neither Intel nor AMD have announced plans for the necessary type of hardware assist.

---

[4] *It is possible, of course, to implement shadow page tables in a low performance manner without binary translation or hardware assist.*

I/O Virtualization

In I/O virtualization, there are two key decisions: where the drivers for the physical network and storage hardware reside, and what virtual hardware is presented to the guest operating system. Hardware support for I/O virtualization has yet to be announced from either Intel or AMD.

The direct I/O architecture used by VMware places drivers for high-performance I/O devices directly into the hypervisor, and uses a privileged domain (called the Service Console) for devices that are not performance-critical. To protect the hypervisor from driver faults, VMware employs a number of mechanisms such as private memory heaps for individual drivers.[5] For virtual hardware, ESX Server uses transparent virtualization for storage devices (that is, it presents accurate virtual SCSI devices), but uses a paravirtualized network driver approach (vmxnet).[6]

XenSource's indirect I/O architecture uses a privileged virtual machine(called domain 0) for all drivers. It terms its model a split driver model, with front end drivers inside the guest and back end drivers in domain 0.[7] Because this design requires both disk and network I/O to traverse a lengthier path, XenSource virtual machines suffer from performance degradation. HP Labs' measurements[8] indicate that XenSource's I/O performance is about 30 percent of native.

In addition, XenSource provides a limited and non-standard view of virtual devices. This paravirtualized approach works well for the network driver, and in fact mimics the vmxnet driver developed by VMware. In the case of storage, however, XenSource's approach compromises storage I/O compatibility from the guest. In practice, this means that clustering, multipathing, tape backup and other SCSI command-based applications inside guests often fail in XenSource virtual machines. The VMware model, on the other hand, presents to the guest an accurate representation of SCSI devices, thereby ensuring that the overwhelming majority of in-guest software 'just works'.
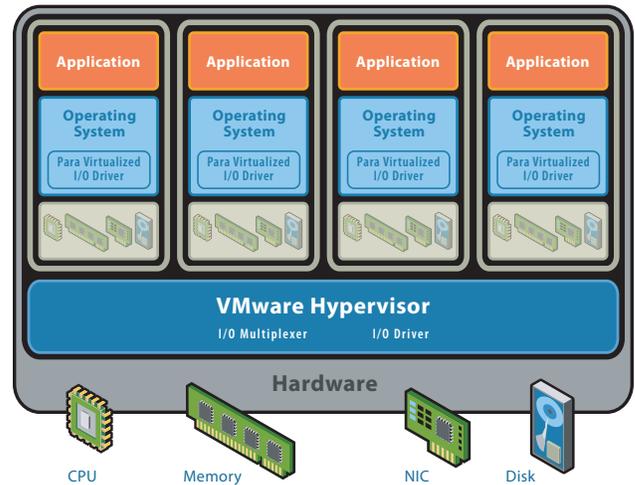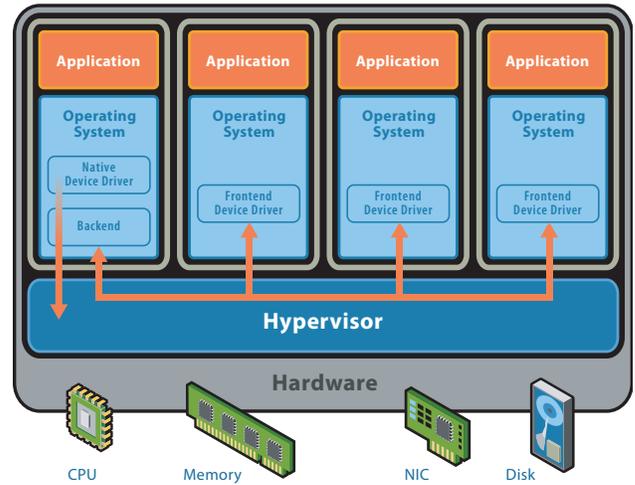

Figure 2: VMware ESX I/O architecture

Figure 3: Xen I/O architecture

[5] Applications do not function well without disk or network access, so whether the driver that experiences a fault is inside the hypervisor or within domain 0 is an academic question. What is really required in this case is rigorous qualification of the drivers against the storage and network I/O devices.

[6] Unlike in the case of CPU and memory, where paravirtualization implies deep operating system changes, in the case of I/O there is a well-established precedent for third-party modification of the operating system, via the driver stack.

[7] Part of XenSource's rationale for using the split driver model is that it helps with isolation: with its drivers in a virtual machine, XenSource hints that it can exert control over how the I/O devices access memory. However, I/O devices regularly make use of direct memory access (DMA), bypassing any software intermediary like the XenSource hypervisor. XenSource itself notes that hardware support in the form of an IOMMU, not its split driver model, is the key requirement. If that hardware support existed, any virtualization vendor could take advantage of it, including VMware and Microsoft.

[8] "Diagnosing Performance Overheads in the Xen Virtual Machine Environment," Menon et al., presented at First ACM/USENIX Conference on Virtual Execution Environments (VEE'05), Chicago, Illinois, 11-12 June 2005.

Feature	VMware ESX Server	XenSource
CPU Virtualization	• Transparent virtualization (binary translation) for maximum guest operating system compatibility today • Support for para-virtualized operating systems and Intel VT/AMD virtualization technologies announced	• Based on kernel-level para-virtualization. Compatible only with operating systems with modified kernels • Future support for Intel/AMD CPU-based virtualization technologies promised
Memory Management	• Transparent virtualization (shadow page tables) to enable most efficient use of memory • Exploits a wide range of advanced memory management techniques	• Para-virtualized approach provides partial access for the guest operating system directly to physical memory page tables • No ability to use advanced memory resource management techniques
I/O Virtualization	• Direct I/O architecture with drivers for high-performance I/O devices in hypervisor • Devices that are not performance-critical managed through Service Console (privileged Linux domain) • Para-virtualized network driver used in virtual machines (vmxnet) • Transparent virtualization of storage devices enables maximum virtual machine compatibility	• Split-driver model puts 'front-end drivers' in virtual machines and 'back-end drivers' in domain 0 (privileged Linux domain) • Para-virtualized network driver used in virtual machines • Block-level storage devices used in virtual machines compromise SCSI compatibility (including clustering, tape backup, multipathing)

Table 1. Summary of architectural differences between VMware ESX Server and XenSource.

A Note on Performance

Although the discussion above describes the performance advantages of particular techniques, it is important to note that performance must be considered in the bigger picture. Whichever method is chosen, there is always some level of performance overhead imposed by virtualization. The overhead comes from CPU virtualization, memory management, and I/O virtualization. The exact proportions and whether they represent a perceptible performance hit depends upon:

• The workload being run—if it can be run at all (see "Workload Suitability" section below). Paravirtualization is effective for situations in which certain non-virtualizable instructions are being called, but its benefits can be outweighed by inadequate memory management or by poor I/O throughput.

• Under which load and target performance assumptions that workload is being run. In other words, is it a multiple virtual machine scenario, in which case resource management capabilities (see QoS section below) become a primary determinant of performance? What level of performance degradation is acceptable, if degradation means downtime versus continuing operation?

• With what other data center infrastructure that workload is being run (see "Enterprise Readiness" section below). Are Fibre Channel SANs required? Multipathing? Clustering? VLANs?

The emergence of hardware assist for virtualization in the form of processor support (Intel's VT and AMD's Pacifica Technology) may help improve performance in some cases. However, it is important to note that hardware-assisted virtualization alone cannot eliminate performance overhead. In some cases a software-based approach still provides superior performance to a purely hardware-based approach. It is important for an enterprise-class hypervisor to possess both software-based and hardware-assisted virtualization support to provide the best virtualization performance under any circumstance.

The key takeaways here are that the advantage of any one technique for any single element of virtualization overhead may be outweighed by a variety of other factors, and that real-world use cases—best discussed in the context of virtualization solutions—are what should matter most to users.

Solution Support

As of the end of 2004, the worldwide customer base of VMware had grown to over 10,000 server customers, with more than 70 percent of the ESX Server user base reporting use in production with a variety of workloads. Based on this experience, VMware has observed that customers typically start by implementing virtualization as the basis for one of the following solutions:

• Server consolidation and containment

• Disaster recovery and business continuity

• Enterprise hosted desktops

Over time, customers tend to branch out in their use of virtualization, to the point where it becomes a standard part of the production data center infrastructure. This standardization on virtual infrastructure provides tremendous value to the customers, because it enables greatly improved resource utilization, superior manageability and flexibility, and increased application availability. Standardization on virtual infrastructure is also the basis for utility computing.

These benefits are not achieved through the hypervisor alone. The essential function provided by a hypervisor is the partitioning of a single physical server. While important, partitioning is but a small subset of the functionality required for production-grade solutions based on server virtualization. Because of this fact, a large part of the engineering talent at VMware over the past several years has been devoted to the following facets of ESX Server:

• Consolidation ratio and server utilization

• Quality of service (QoS) for individual virtual machines

• Workload suitability

• Rapid restart and provisioning

The previously mentioned solutions are dependent to differing degrees on these factors. Table 2 shows the relative importance of these factors by solution.

	Consolidation Ratio	Virtual Machine QoS	Workload Suitability	Rapid Restart and Provisioning
Server consolidation	Med – High	Very High	High	Med
DR / business continuity	Very High	Med	Med	Very High
Hosted desktops	Very High	High	Med	Med
Virtual infrastructure	Depends on app	High – Very high	Very High	High

Table 2. Key success factors for solutions based on virtualization.

Consolidation ratio and server utilization

The number of virtual machines that can be run at target performance levels on a given physical machine, usually referred to as the consolidation ratio, is often the primary rationale for initial deployment of virtualization. In part, the consolidation ratio is determined by architecture. For example, a split driver model consumes CPU cycles that could be better spent on the virtualized workloads. More importantly, however, the ratio is determined largely by the efficiency with which the virtualization layer utilizes the underlying hardware.

In the case of ESX Server from VMware, utilization is greatly enhanced by memory overcommitment. Memory overcommitment is the ability to have the total memory available to virtual machines be greater than the actual physical memory on a system. For example, with ESX Server it is possible to have 10 virtual machines, each with 3.6GB of RAM available to their operating systems and applications, on a 2-CPU system with 16GB of physical RAM. Without memory overcommitment, there is a strict limitation on the number of high-performance virtual machines that can be run on a given physical server. (In the previous example, the maximum number of virtual machines would have been closer to 5, compared to the 10 that is regularly achieved by the customers of VMware .) More importantly, solutions which depend on high consolidation ratios and the associated hardware savings for their economic justification—including n+1 disaster recovery, server containment and hosted desktops—are not possible.

ESX Server includes several techniques invented at VMware for memory overcommitment, including transparent page sharing, memory ballooning, transparent swap, page remapping across NUMA nodes, and the idle memory tax.[9] Of these techniques, XenSource has followed VMware in using ballooning, but even that requires manual intervention reclaim memory. XenSource can not implement any of the other methods for improving the efficiency of memory usage, due to their architectural limitations (that is, they do not have shadow page tables).

In addition to its memory management capabilities, ESX Server also uses a number of other special techniques to increase the efficiency of its resource utilization. Advanced CPU scheduling capabilities help to optimize the utilization of all of the processors on a physical server, including the cases where those processors use hyperthreading or have NUMA architectures. VMware also provides a lightweight, protected environment for running virtualization assists and third-party software directly on the ESX Server hypervisor. XenSource requires such software to reside within a full, heavyweight operating system running in domain 0.

Quality of Service for Individual Virtual Machines

With multiple virtual machines on the same host, quality of service(QoS) guarantees for the virtual machines become extremely important, especially in a production environment. At the most basic level, the requirement is for performance isolation, or the ability to ensure that virtual machines do not negatively affect the resource availability of other virtual machines.[10] More than that, however, it is important to be able to guarantee a minimum level of resources for a given virtual machine, and then to let the virtual machine use other resources that may be available on the physical machine.

The proportional share mechanism used by VMware, coupled with min/max guarantees and admission control, enables this type of control for CPU and memory resources. Similarly, for network bandwidth management, VMware employs traffic shaping.[11] As mentioned above, VMware also has production-tested mechanisms for rebalancing CPU and memory utilization, so that virtual machines can take advantage of resources that become freed up as other virtual machines are powered down or migrated off. This stands in stark contrast to the current state of XenSource resource management. XenSource has yet to even provide basic load balancing across CPUs within a single server—a capability that was available in the VMware ESX Server 1.0 beta release. Put simply, advanced resource management for virtual machines is a basic requirement for the flexible IT environment.

Workload Suitability

As the range of applications being considered for virtualization continues to increase, the criteria of if they run, how well they run, and whether they are supported will increasingly become key for customers deciding what virtual platform to use. The suitability of a given workload as a candidate for virtualization is determined by:

• Operating system support

• Virtual hardware features

• Application-specific performance tuning

• ISV support

---

[9] *For more detail, see "Memory resource management in VMware ESX Server," Carl A. Waldsburger, in Proc. of the 5th Symposium on Operating Systems Design and Implementation, Boston, MA, December 9-11 2002.*

[10] *In fact, performance isolation should be regarded as the basic enabler of the flexibility of virtual infrastructure; without it, virtual machines continue to be constrained by physical requirements.*

[11] *More details on each of these can be found in the product documentation, available online at http://www.vmware.com/support/esx25/doc/admin/index.html*

Operating System Support

VMware ESX Server fully supports a wide range of guest operating systems. XenSource, on the other hand, supports only Xenolinux (its privately modified Linux kernel), which to date is not available in any supported enterprise Linux distribution. Microsoft, while threatening to support Linux, has yet to support anything other than Windows. Table 3 below provides a more detailed comparison of what is available, and what has been announced.[12]

It is important to define what each company means by "support". For VMware products, it is not merely a matter of being able to boot up the operating system and run applications at a minimal performance level. VMware optimizes the performance of its hypervisor for each operating system, tests enterprise-class workloads against them (see below for the list of applications), and runs extensive, vendor-supported certification suites for each operating system. Customers should demand the same from whichever virtualization vendor they choose.

VMware is also working with the Linux community and other partners to define the Hypercall standard. Hypercall leverages the considerable expertise of VMware in heterogeneous guest operating system support to ensure that customers can use the same operating system whether it is running on top of a virtualization layer or not, and that they are not always be forced to migrate to a single version of the operating system. As an example, without this standard customers would have to run two different versions of RHEL 5 or SLES 10 (one paravirtualized, one not) in their data centers, assuming they did not move their entire data center to run on virtualization. This problem would only be exacerbated over time, as the next versions of these operating systems become available.

Operating System Support	VMware	XenSource	Microsoft
Unmodified Linux	Yes (Since 1999)	No	Announced
Microsoft Windows	Yes (since 1999)	Announced[13]	Yes (since 2004)[14]
Solaris x86	Announced	No[15]	No
Novell NetWare	Yes (since 2002)	No	No
Para-virtualized guest operating systems	Announced	Yes (since 2004)	No

Table 3. Operating system support available from server virtualization vendors.

[12]*See http://www.vmware.com/pdf/esx_systems_guide.pdf for more detailed info.*
[13]*Requires hardware virtualization assist.*
[14]*http://www.microsoft.com/presspass/press/2004/sep04/09-13AvailabilityVS2005PR.mspx*
[15]*Sun has a prototype of Solaris booting under Xen*

Virtual Hardware Features

The virtual hardware presented to the operating system in the VMware virtual machines is extremely accurate. Part of the motivation for this is to ensure that the vast majority of functionality that is available to physical servers works without change in virtual machines. For example, clustering software can run without modification inside VMware virtual machines.

Another part of the rationale for accurate virtual hardware is to allow VMware to take advantage of physical server innovations. For instance, VMware has incorporated a BIOS in the virtual machine. Without a BIOS, BIOS-controlled functions such as power management have no natural home inside the virtualization platform. Hyperthreading and NUMA support are two other examples of how VMware has taken advantage of CPU innovations.

Yet another area where virtual hardware matters is providing support for multi-processor virtual machines. VMware already includes support for 2-way virtual SMP (VSMP) as part of its virtual infrastructure node (VIN) bundle, and plans to advance VSMP as users require it. Although XenSource has announced similar support it is not yet available.

Application-Specific Performance Tuning

As with physical servers, it is possible to tune a virtual machine to better suit one application or another. ESX Server is tuned to support the following list of applications:

• Citrix Metaframe

• File servers

• Microsoft SQL Server

• J2EE application servers

• Oracle Database

• Web servers

• Compile/build

Customers have also used ESX Server for a much wider variety of applications. As with QoS functionality, improvements in application performance are gradual, and are very much a function of product maturity.

ISV Support

Support from other software vendors comes from having an established track record of widespread, stable production usage and strong evidence of customer demand. As VMware ESX Server is the only virtualization product with these qualities. It is also the only virtual platform with a strong degree of ISV support. An up-to-date list is available at http://www.vmware.com/partners/sw/alliances/, but a sampling includes BMC Software, Citrix, Computer Associates, Hewlett-Packard, IBM, Novell, Oracle, RedHat, and Veritas—essentially all of the key vendors in the x86 software market.

Rapid Restart and Provisioning

Superior manageability and flexibility are among of the primary benefits of server virtualization. As users move towards data centers built on the principles of utility computing, the rapid configuration of IT resources in response to changing business requirements is viewed as a core competency. This requirement extends to virtual machines, which are typically superior to physical machines in terms of restart and provisioning times.

While users largely view the VirtualCenter management suite and VMotion™ technology as the primary offerings from VMware in this area, it is important to note that ESX Server itself includes a key enabler: the Virtual Machine File System (VMFS). VMFS is a specialized file system that acts as a storage virtualization router from the perspective of the virtual machine, allowing the aggregation of disk arrays and provisioning of LUNs to the virtual machines. From the hypervisor perspective, VMFS is a distributed, clustered file system that enables a variety of functions to be supported with production-grade reliability and speed:

- For VMotion, the source and destination servers need to see the same storage simultaneously. Network attached storage (NAS) is one option that enables this visibility; a clustered file system like VMFS on Fibre Channel or iSCSI storage networks is another. NAS—the only option that works with XenSource's VMotion-like capabilities—is functional, but slower.

- Similarly, to rapidly restart a virtual machine on a different physical server, all servers need to see the same storage at the same time. Among the many mechanisms that can meet the requirement are VMFS on SAN; NAS (which are slower), manual restart (by changing the LUN masking and such), dedicated failover servers without dynamic placement; or non-N+1 cluster configurations. As with VMotion, VMFS on SAN is the superior option.

- While virtual machine provisioning does not specifically require a clustered file system, it is substantially enhanced by VMFS. Non-specialized file system access is too slow, and so the only remaining option is to use raw LUNs, as XenSource requires. However, raw LUNs have the drawback of being inflexible, thereby partly negating the value proposition of flexibility.

Enterprise Readiness

Enterprise readiness can be roughly summarized by the following question: how well can a product fit into the existing IT infrastructure, especially in production? More specifically:

- To what extent is the virtual infrastructure software certified and supported against the servers and storage on which it depends?

- How well do the customer's existing mechanisms for availability, data protection, load balancing, and such work in conjunction with virtual machines?

- What tools are available for managing this product, and to what extent does the market provide alternatives?

Certification

Because hypervisors represent a new layer of software underneath existing operating systems, certifications and compatibility for those operating systems do not carry over to the hypervisor. Furthermore, if new components are added to the storage stack (such as XenSource's block device emulators for storage, then a fundamental re-certification is required. Even if the hypervisor has the theoretical capability to run unmodified operating system drivers in a privileged domain (that is, the ESX Service Console or Domain 0 in XenSource), in practice those drivers behave entirely differently when put under the strain of a virtual environment with multiple operating systems and shared devices. To run those devices reliably and with high-performance in a virtual environment frequently requires extensive driver modifications and testing to ensure that the devices work in all sorts of enterprise deployments (e.g., with all SAN topologies, with clustered applications, and with layered storage applications, etc.).

As a mature, commercially supported enterprise software product, VMware ESX Server has an extensive hardware and software compatibility list that is maintained through rigorous testing by VMware and its partners. In addition to broad ISV support and certification against a wide variety of server hardware over the past four years, VMware ESX Server includes a certified list of storage hardware and I/O devices that covers the vast majority of networking and storage equipment deployed in data centers today. Customers should be wary of hypervisors that have not undergone the necessary validations to achieve IBM ServerProven status, to be listed on the EMC Support Matrix (ESM), to pass Microsoft's Hardware Compatibility Test (HCT), or any of the other many vendor certification suites.

Operational Fit

Interoperability with existing IT infrastructure goes beyond certification. Most data centers have production requirements that incorporate elements of high availability, data protection, load balancing, and other technologies that enable business continuity. If it is to be accepted for use in production, virtualization must meet those requirements.

On the storage-related front, ESX Server supports multiple methods for improving data availability for virtual machines, including:

- Multipathing

- Clustering

- Array-based snapshot and replication technologies, via raw device mapping technology

- VMotion

- Software-based virtual machine disk snapshots for disaster recovery

- SAN-based logical volume management

XenSource's currently does not support these features. XenSource will need to go through a long and rigorous testing regimen to be customer-ready.[16]

For networking, the virtual switches included as part of ESX Server include:

- VLAN tagging and NIC teaming

- Support for Microsoft network load balancing

- Enhanced layer 2 security

- Mirror ports available if promiscuous mode is required

Analogous technologies are available in the Linux space, but are generally not suitable for production deployments.

Management and the Market

The final piece in the enterprise readiness puzzle is manageability. Within manageability, there are several key areas to consider:

- **Basic management.** Are monitoring, alerting and reporting available? Are virtual machine-specific capabilities like cloning and provisioning supported? Is there a single pane of glass to manage the virtual infrastructure, including the storage and network elements that are hooked into the virtual machines?

- **Advanced management.** Beyond element management for virtual machines, is there a way to manage farms of virtual machines, e.g., for workload management or for aggregate availability?

- **Roles-based access control.** Can these functions be segregated by administrator? Are there adequate audit trails, so that customers can remain in compliance with Sarbanes-Oxley and other regulatory requirements?

- **Partner Ecosystem.** To what extent are the virtual machines supported by other management software vendors?

On all fronts, VMware has by far the strongest offering. VMware VirtualCenter product handles the first area with graphical, easy-to-use tools. The second area is well represented not just on the VMware roadmap, but also on those of its many partners. In addition, the permissions and authentication mechanisms in both VirtualCenter and ESX Server have been created with flexibility as a basic requirement, so that end users can make the management software conform to their organizational structure—not the other way around.

VMware not only enjoys strong relationships with every major systems and storage vendor worldwide, but also actively cultivates a network of software and hardware vendors that contribute to the overall solution stack. In particular, its VMware Community Source initiative will enable industry-wide collaboration and new levels of technological cooperation and innovation between partners. .

Conclusion

In the past several years, server virtualization technology has moved quickly and firmly into the IT mainstream. Of the vendors offering hypervisor-based products, only VMware can be considered production-ready:

- The VMware architectural vision is based on years of experience solving real-world problems in performance, security and compatibility, not unproven academic research.

- VMware's hypervisor has been augmented by a wide array of technologies to enable use in a variety of solutions.

- VMware's products have matured to the level that enterprise customers demand.

End users should expect VMware to continue leading the hypervisor market, bringing to IT organizations a wide range of benefits from virtualization: greatly improved resource utilization, superior manageability and flexibility, and increased application availability.

---

[16]*VMware's VMotion was tested in-depth across a diverse range of server and storage hardware for a full year and a half before being made generally available.*

vmware®