Server consolidation with VMware Virtual Infrastructure and HP ProLiant technology

technology brief

Abstract

The explosive growth of client/server computing infrastructures using industry-standard servers has led to a corresponding growth in administration, hardware, space, and cooling costs. HP has developed a unified partitioning strategy that offers multiple solutions to this problem, depending on a customer's needs. An important part of this strategy is the use of VMware virtualization software on ProLiant servers. Hardware partitioning has been a solution for many years for mainframe-class servers; however, this solution has been limited on Intel-based, industry-standard servers. VMware has developed virtualization software that brings these mainframe-class partitioning techniques to industry-standard servers.

Using VMware software, customers can reduce costs, improve resource efficiency, and benefit from improved management and provisioning by consolidating multiple logical servers onto a more powerful ProLiant servers or server blades. VMware technology dramatically reduces the number of servers needed to support applications while maintaining application isolation and encapsulation. It allows multiple virtual machines, each dedicated to an enterprise application, to reside side by side within a single physical server. VMware GSX Server, in which the virtualization software resides on a host operating system (OS), is ideal for departmental applications. VMware ESX Server, which provides higher performance and resides directly on the server hardware, is well suited for datacenter applications. With VMware VirtualCenter, customers can centrally manage multiple ESX Servers as well as dynamically allocate resources to virtual machines without any interruption to users. The VMotion technology within VirtualCenter allows customers to dynamically migrate virtual machines across the virtual infrastructure with zero downtime.

HP has partnered with VMware to sell and support VMware virtualization software. HP provides a one-stop shop that includes customer services and support for the VMware software, the guest operating systems, ProLiant hardware, and associated storage resources.

This technology brief describes the underlying architecture of the VMware software, the advantages of consolidating workloads using VMware and HP technologies, and some example scenarios in which customers can benefit from server consolidation.

Need for server consolidation

Information technology (IT) consolidation takes many forms and is often described using different terms. For example, IDC[1] outlines several types of IT consolidation, one of which includes combining multiple servers that support different workloads or operating systems onto a single platform. IDC calls this *application consolidation*. This technology brief uses the term *server consolidation* to refer to consolidating multiple applications and operating systems onto a single server.

As enterprises have grown—especially during the Internet boom—a common practice has been to add another server whenever a new application is needed: Need a firewall? Add a server. Need a mail application? Add a server. Need a load-balancing application? Add a server. However, this practice has led to many servers that are using only a fraction of their capabilities, while still requiring the full amount of cooling, space, power, and management resources. As a result, administrators are struggling with the management of tens or hundreds of servers, along with the resulting costs, yet are not fully utilizing these servers.

To minimize these management challenges, administrators are increasingly streamlining their businesses by consolidating servers. Server consolidation is becoming more attractive not only because it reduces costs, but also because of the increased performance and availability levels of

---

[1] IDC provides technology intelligence, industry analysis, market data, and strategic and tactical guidance to builders, providers, and users of information technology. Visit http://www.idc.com/ for more information. Also review their paper *IT Consolidation Solutions from Hewlett-Packard*, available at: http://h30046.www3.hp.com/solutions/3521_rev3.pdf

industry-standard servers. For example, the current Intel 32-bit (IA-32) processors have more powerful architectures and higher clock speeds than previous generations. The seventh-generation IA-32 Intel® Xeon™ family performs approximately 30 percent faster than the previous Intel Pentium® III family. Advanced memory protection techniques such as Online Spare Memory and Hot-plug RAID memory provide levels of availability previously not found in industry-standard servers. TPC-C transactions per minute[2] have increased from approximately the 10,000 TPM range in 1996 to the 70,000 range for 4-way servers in 2003. Thus, hardware advances are making server consolidation much more attractive than it has been in the past.

For more information about server consolidation, see the technology brief titled *Successful server consolidation: it's all in the preparation*.[3]

HP consolidation strategies

Consolidating multiple operating systems or applications onto a single server requires some level of isolation, or partitioning, between them. Isolation helps prevent a fatal error on one application from affecting other applications, ensures an adequate level of resources for each application, and helps maintain security policies. When considering a server consolidation solution, administrators must balance the need for isolating applications against the need for dynamically allocating resources. HP provides multiple solutions for consolidation to allow IT administrators to meet the degree of isolation, cost, and availability needed (Figure 1). For customers that require the highest degree of isolation, HP Integrity Superdome servers can maintain electrically-controlled hardware partitions.[4] While hardware partitions yield the highest degree of isolation, they are often not as flexible as virtualization software or resource partitioning techniques.

Figure 1. HP consolidation continuum
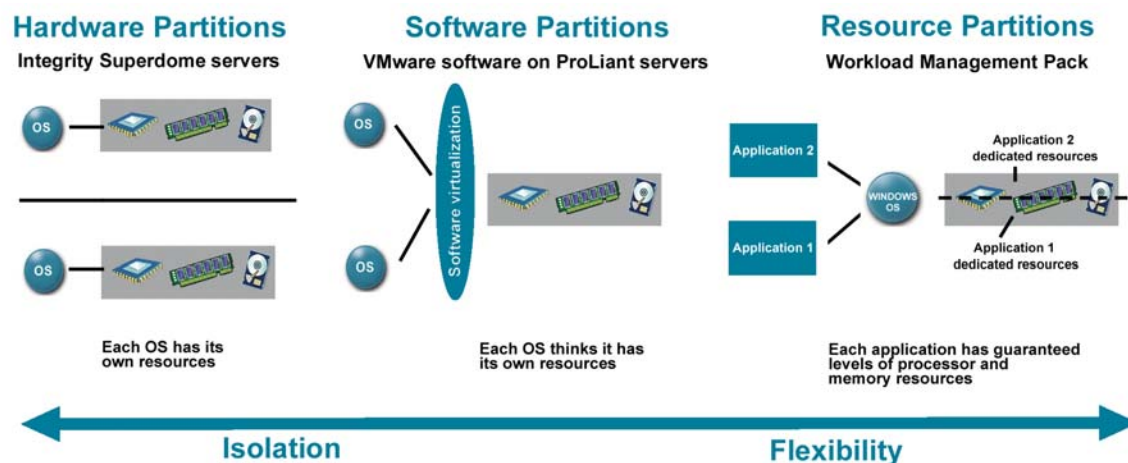
VMware ESX Server and VMware GSX Server software enable administrators to create virtual machines with one or two processors, allowing multiple operating systems to run simultaneously and

---

[2] Information obtained from the Intel white paper titled *Twenty-to-One Consolidation on Intel Architecture*, available at http://www.vmware.com/pdf/Intel_Consolidation.pdf. For more information about TPC benchmarks, see the HP website at http://h18020.www1.hp.com/products/servers/benchmarks/index.html and the TPC website at http://www.tpc.org/
[3] Available at ftp://ftp.compaq.com/pub/supportinformation/papers/tc030505wp_rev0_us.pdf
[4] For more information about Integrity Superdome servers, see http://www.hp.com/products1/servers/integrity/superdome_high_end/index.html

independently on the same industry-standard server. VMware products use a software layer that abstracts the hardware, so that each OS thinks it is accessing its own set of hardware resources—when in fact, the resources are shared among multiple virtual machines. The administrator controls the resources committed to each virtual machine. When one of the virtual machines is idle, the other virtual machines on the server are able to use the free resources. Each virtual machine is independent of the other virtual machines. Thus, user error, software failure, or complete shutdown of one of the virtual machines has no effect on other virtual machines.

Finally, in the HP partitioning continuum, for customers that need to run multiple applications under a single Windows operating system, the ProLiant Essentials Workload Management Pack (WMP) is available.[5] WMP is a software solution that controls and dynamically allocates system resources such as microprocessors and memory. Using WMP, administrators can consolidate multiple applications on Microsoft® Windows® 2000 or 2003 server platforms. WMP provides the tools for a customer to allocate processor and memory resources to specific applications so that the applications can fully utilize those resources.

Customers can use more than one solution to meet their needs. For example, an administrator might want to use VMware software to isolate a Linux® OS from a Windows OS and then use WMP on the Windows OS to optimize the assigned system resources for three different applications running on Windows.

VMware and HP

HP and VMware have a strong relationship initiated in 2000. HP is a charter member in the VMware Hardware Alliance Program, which allows HP and VMware to test compatibility between ProLiant platforms and VMware software. In 2002, HP developed a worldwide alliance agreement with VMware.

Further building on this agreement, HP and VMware announced an enhanced resale and customer support agreement in May of 2003. This agreement enables HP to provide a one-stop virtualization solution by reselling VMware Virtual Infrastructure Software along with HP-branded technical support and subscription services for the VMware environment. HP provides support of the VMware environment, the operating systems within VMware virtual machines, and the HP hardware. The partnership agreement allows HP to provide customers with a complete turn-key virtualization solution.

In 2004, EMC Corporation acquired VMware. EMC is committed to supporting VMware's ongoing product and partnering strategy and will operate VMware as a wholly owned subsidiary. HP will continue to work with VMware to provide premier solutions to our customers.

VMware architecture

The key function of VMware software is to abstract, or virtualize, the hardware resources of a server so that multiple operating systems and applications can each access an independent set of virtual hardware resources. The software maps the resources of the virtual machines to the hardware resources of the physical server, so that neither the guest OS (the OS running on each virtual machine) nor the individual application knows which processor, memory, disk drive, or network adapter it is physically mapping to. VMware GSX Server uses a software layer that sits on top of the existing host OS to virtualize the hardware (Figure 2). The VMware virtualization layer provides hardware independence, isolation, and encapsulation.

[5] For more information about WMP, visit http://h18013.www1.hp.com/products/servers/proliantessentials/rpmdescription.html
Readers may also refer to the technology brief titled *HP Resource Partitioning Manager: A Server Consolidation and Optimization Solution for Microsoft Windows 2000,* available at ftp://ftp.compaq.com/pub/supportinformation/papers/tc020102tb_rev1_us.pdf

Figure 2. GSX Server architecture diagram. The operating system layer shows some of the possible operating systems.
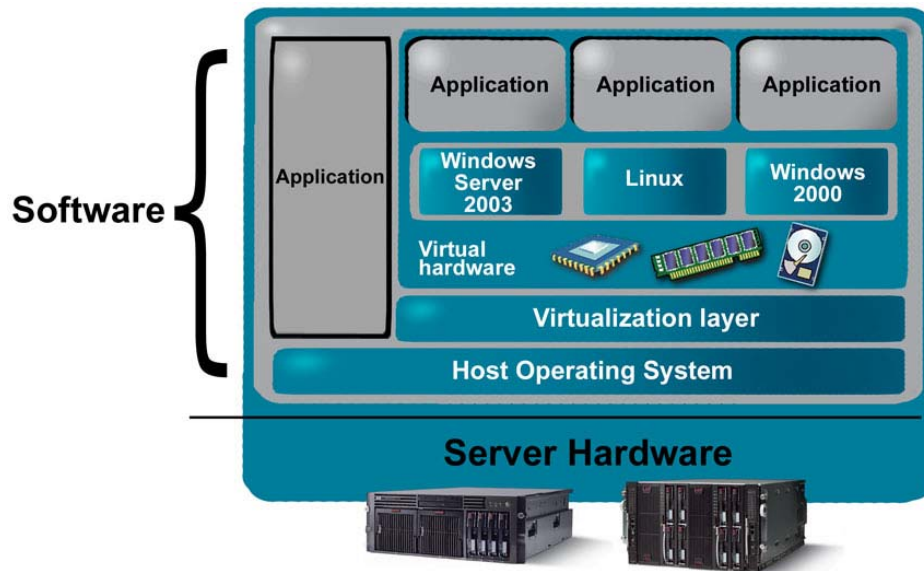
VMware ESX Server, on the other hand, runs the software virtualization layer directly on the server hardware (Figure 3). The VMware virtualization layer running on bare hardware improves performance and enables fine-grain resource control. In addition, the virtualization layer provides hardware independence, isolation, and encapsulation.

Figure 3. ESX Server architecture diagram. The operating system layer shows some of the possible operating systems.
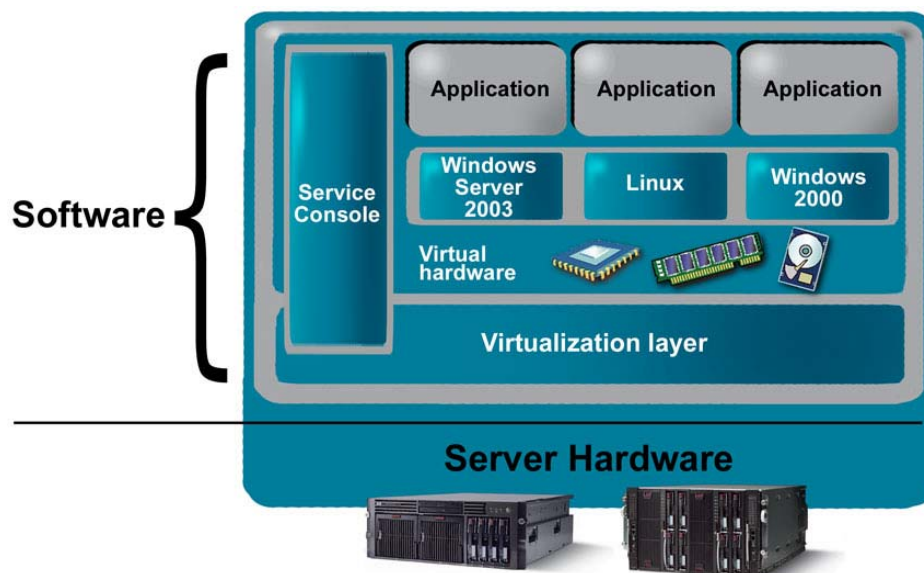
Hardware independence

The VMware virtualization layer performs all the necessary translations to mask the physical hardware resources from the guest OS. The guest OS sees a consistent set of virtual hardware, regardless of what type of physical hardware is on the server. This means that whether the physical server is using, for example, an Intel network controller or an HP network controller, the guest OS will see only the network controller defined by the standardized virtual hardware set in ESX Server and GSX Server.

Because applications interact only with the associated guest OS and not with the underlying virtual hardware, once OS compatibility with the virtual hardware is established, application compatibility is not an issue.

This hardware independence enhances reliability due to the small number of device drivers required for the virtual hardware and, thus, for each virtual machine OS. This also makes virtual machines simple to transport between physical servers, even if those servers vary significantly in their hardware configuration.

Isolation

The virtualization layer also provides the ability to completely isolate one virtual machine from another. Each virtual machine includes a completely separate OS, registry, applications, and data files. Because each OS is isolated, one OS cannot communicate with or leak data to any other OS, other than through networking mechanisms similar to those used to connect separate physical machines.

Errors or user actions that interfere with the operation of the OS and applications running in one virtual machine have no effect on the continuing operation of other virtual machines on the same server. This increases the availability for multiple applications in comparison to running them on a single server.

Encapsulation

VMware encapsulates the entire state of a virtual machine—the memory, disk images, and I/O device state—into as few as two files. The two files include a short text file defining the configuration of the virtual machine and a virtual disk file that contains its data. Thus, the virtual machine can be copied, saved, and moved like a file. Customers can move entire virtual machines among development, testing, and production environments simply by moving the virtual machine state file from one physical system to another. This allows customers to deploy and redeploy virtual machines easily and flexibly.

VMware GSX Server

VMware GSX Server installs and runs as an application on a Microsoft Windows or Linux host OS.[6] GSX Server 2.5 supports up to 32-way multiprocessor servers and can run a maximum of 64 virtual machines concurrently on a single host server. As GSX Server loads each virtual machine, it pre-allocates the memory for each. Therefore, administrators will need to ensure that enough memory is available to run the host OS plus the memory required for each guest OS and applications. GSX Server can support up to 64 GB of main memory and a maximum of 2.0 GB of memory per virtual machine.

After VMware GSX Server is installed, it loads the virtualization layer that abstracts the hardware resources. Because GSX Server runs as an application on top of the host OS, it extends services from the host OS to implement the virtual devices. This means that device compatibility is inherited from the host OS. This provides a great advantage for customers that want to run an older OS on a newer server. Suppose an administrator has an application that runs on Windows NT® 4.0 and wants to consolidate that application with others onto a new server. While Windows NT 4.0 might not have driver support for the latest CD drive, disk drive, or network card installed on the new server, the

---

[6] See the documentation on the VMware website at http://www.vmware.com for up-to-date information on OS support.

administrator can still use that server by setting up a GSX Server virtual machine. The GSX Server virtual machine can run Windows NT 4.0 as the guest OS, even though the OS does not have the hardware drivers to directly support the new server peripherals. This gives an administrator freedom from chasing backward compatibility: As long as the host OS runs on that server hardware, VMware GSX Server can perform the necessary translations so that the guest OS can also run on that server hardware.

The tradeoff for this hardware compatibility is the translation overhead for the virtualization layer. Applications that need access to I/O hardware such as disks and networks must go through five layers before "touching" the hardware: the application, the guest OS, the GSX virtualization layer, the host OS, and then the hardware. Therefore, GSX Server is designed for smaller, departmental applications rather than large databases or other extremely large applications. Customers requiring higher performance should consider ESX Server.

The Appendix includes a table summarizing the capabilities of VMware GSX Server and ESX Server.

VMware ESX Server

VMware ESX Server supports up to 16-way multiprocessor servers and can run a maximum of 80 virtual machines concurrently on a single host. ESX Server can run up to eight virtual machines on a single processor. ESX Server supports up to 64 GB of main memory and a maximum of 3.6 GB of memory per virtual machine. Unlike GSX Server, ESX Server does not need to commit memory resources to each virtual machine as it loads. ESX Server includes the capability to have up to eight 64-GB memory swap files. ESX Server can also "over-commit" memory by using virtual memory techniques to take advantage of memory not being used by another virtual machine. ESX Server can over-commit memory by 200 percent—that is, in order to run 6 GB of virtual machines, 2 GB of available memory is required.

In addition, because it is designed for enterprises, VMware ESX Server has several enhancements over VMware GSX Server:

- Runs directly on hardware for improved performance
- Provides dynamic resource allocation
- Supports clustering of virtual machines across different host servers, unlike GSX Server that supports only shared disk clustering on the same host server

VMware has announced two plug-in modules that are available only when using VMware ESX Server 2.x:

- Virtual SMP support that allows a virtual machine to use more than one processor
- VMware VirtualCenter that provides a central management facility for all the virtual machines on the network and allows dynamic allocation of resources. VMotion technology with VirtualCenter allows migration of virtual machines across multiple physical servers.

Direct hardware performance

VMware ESX Server does not use a host OS as GSX Server does. The ESX Server product sits directly on top of the hardware, providing an abstraction layer for the hardware resources (refer to Figure 3). Having the virtualization layer directly on top of the hardware reduces the overhead involved in distributing the physical resources across multiple virtual machines. However, this does mean that ESX Server, unlike GSX Server, must have the device drivers to control the server hardware directly and perform the translations needed to convert the physical hardware to the virtual hardware. Thus, ESX Server requires compatible hardware on which to run. ESX Server architecture supports high performance PCI-based SCSI, RAID, Fibre Channel, Ethernet, and Gigabit Ethernet adapters, as well as internal RAID controllers. The goal of VMware is to support standard storage and network

adapters for a wide variety of server platforms. For up-to-date information on supported platforms and devices, see the current ESX Server compatibility guide.[7]

Because ESX Server runs directly on the hardware and not as an application on the host OS, there is no host OS which can extend its security vulnerabilities into the virtual machine.

Dynamic resource management

ESX Server provides robust resource management capabilities, which allow customers to dynamically reallocate system resources to applications and to control consumption of processor, memory, disk I/O, and network bandwidth by each virtual machine. This prevents one virtual machine from monopolizing any resource to the detriment of another virtual machine. Administrators can dynamically reallocate resources using either a web-based management interface or a scripting language.

The ability to dynamically manage resources allows administrators to define, provide, and guarantee levels of service. ESX Server uses a proportional fair-share algorithm for the processor, memory, and disk bandwidth resources so that each virtual machine is guaranteed a minimum level of those resources. Resources that are unused by one virtual machine can be used by other virtual machines to prevent unnecessary idling. For network bandwidth management, ESX Server uses guarantees based on exact specification of peak and sustained bandwidth and burst sizes.

For a detailed discussion of the ESX Server resource management techniques, see the ESX Server Administration Guide.[8]

Clustering capabilities

ESX Server supports clustering of virtual machines using industry-standard clustering solutions such as Microsoft Cluster Service. ESX Server supports clustering between virtual machines on a single ESX Server, across multiple physical servers running ESX Server, and between virtual machines and physical systems. For example, an administrator can set up a cluster consisting of virtual machines that span two different physical servers. Even if a physical server fails, the cluster software keeps the data and applications available (Figure 4).

Figure 4. Cluster spanning two physical servers

---
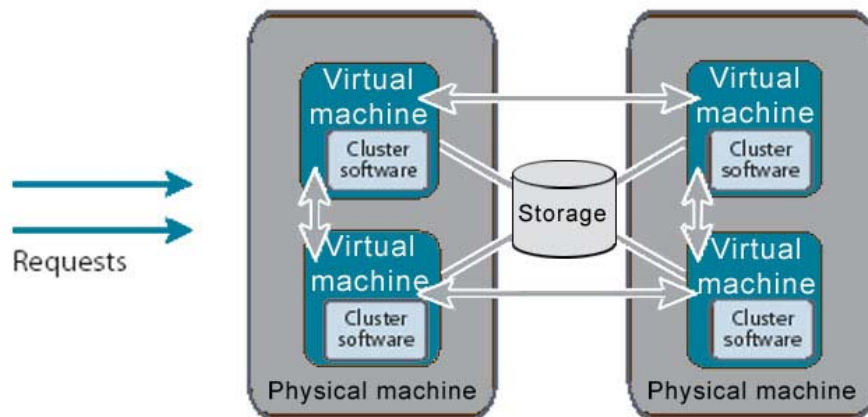
[7] The VMware ESX Server compatibility guides are located at http://www.vmware.com/products/server/esx_specs.html
[8] Available at http://www.vmware.com/pdf/esx2_admin.pdf

VMware Virtual SMP

VMware Virtual SMP is an add-on module to VMware ESX Server 2.x that makes it possible for a single virtual machine to use two physical processors. The Virtual SMP software is an extension of VMware patented virtualization technology that masks the physical resources from the guest OS and applications.

Virtual SMP technology is useful for compute-intensive applications that can take advantage of multithreading. It enables enterprises to scale computing environments easily without adding new hardware. It also enables administrators to create a more realistic test and development environment by building standard SMP virtual environments.

To use the Virtual SMP technology, the physical server must be using ESX Server 2.x and the guest OS must support multiprocessor systems. Examples include Windows Server 2003, Windows 2000, and Red Hat Enterprise Linux AS 2.1. A complete list of supported guest operating systems[9] is available on the VMware website. The physical server must have at least two physical processors to create a virtual machine with two virtual processors.

The Intel Xeon processor family, used on some server platforms, includes a technology known as Hyper-Threading. Hyper-Threading allows multithreaded applications to take advantage of dual virtual processors within a single physical processor.[10] VMware ESX Server does not recognize or take advantage of Hyper-Threading, so VMware recommends that it be turned off. Administrators using ProLiant servers can turn off Hyper-Threading through the ProLiant ROM-Based Setup Utility.

VMware VirtualCenter and VMotion

VMware VirtualCenter is a complementary product that runs with ESX Server 2.x. It enables administrators to centrally manage and control all of their virtual machines, regardless of which physical server is hosting the virtual machine. The VirtualCenter module allows administrators to move virtual machines from one physical server to another with continuous service availability, using VMotion technology in VMware VirtualCenter. VMotion technology enables live, stateful applications to be moved across distributed systems without service interruption, lost data, or lost transactions.

Advantages of using VMware Virtual Infrastructure on HP ProLiant servers

HP has partnered with VMware because of its outstanding products that complement other HP partitioning solutions such as WMP. This does not mean that a customer that chooses VMware software is obligated to use ProLiant hardware. However, there are numerous benefits to selecting HP as a provider and using ProLiant hardware when consolidating servers with VMware software.

These advantages include:

- High-availability and high-performance server hardware
- Wide breadth of qualified server, server blade, and storage platforms for ESX Server
- Management tools such as Integrated Lights-Out (iLO), Remote Insight Lights-Out Edition II (RILOE II), Remote Deployment Pack (RDP), and Insight Manager 7. This includes the integration between VMware ESX Server and Insight Manager 7.
- VMware resell and support agreements

---

[9] Available at http://www.vmware.com/support/esx2/doc/intro_sysreqs_esx.html - 1002463
[10] For more information about Hyper-Threading, see the technology brief titled *Intel Hyper-Threading Technology*, available at ftp://ftp.compaq.com/pub/supportinformation/papers/c030306tb_rev0_us.pdf

ProLiant hardware advantages

ProLiant industry-standard servers are leaders in innovative designs that include high-availability technologies along with high performance. Therefore, using ProLiant servers, with their numerous high-availability features, allows administrators to have confidence when deploying multiple VMware virtual machines on a single physical server. Furthermore, the high performance of the ProLiant servers allows administrators to effectively deploy multiple machines onto a single, higher-performing machine, saving costs for administration, physical cooling, space, and management.

High availability

High-availability technologies on ProLiant servers include redundant and hot-pluggable features. Redundancy provides instantaneous backup should one component fail, while hot-pluggable devices allow administrators to replace critical or failed components without bringing down the system. Examples of high-availability features within ProLiant servers include redundant ROMs, redundant hot-plug fans, redundant power supplies, redundant NICs and LOMs (LAN-on-motherboard), PCI hot-plug slots, redundant hot-plug disk drives, Online Spare Memory, and hot-plug RAID memory.

The redundant ROM, for instance, provides a disaster recovery mechanism for mission-critical ProLiant servers. In the event that the primary ROM image becomes corrupted, the redundant ROM feature automatically reverts to a backup image of the system ROM firmware.

Online Spare technology copies data to a spare memory bank when a DIMM fails. With Online Spare Memory, the server ROM automatically configures the last populated bank as the spare memory. If a DIMM exceeds a predefined error threshold, the system will copy the data from the entire bank that contains the failed DIMM to the Online Spare Memory bank. The server will deactivate the failed bank and continue operating. The customer can replace the failed DIMM during a scheduled shutdown.

Another HP innovation is hot-plug RAID memory that uses a redundant array of industry-standard DIMMs to provide fault tolerance. It provides the ability to hot replace memory while the server is operating. With hot-plug RAID Memory, the RAID engine provides redundancy to ensure data protection, and the hot-plug abilities allow administrators to replace a DIMM without any downtime.

High performance

Because of the careful design of ProLiant servers, they have repeatedly broken TPC and other benchmark records.[11] As an example, HP engineers strive to balance the microprocessor, memory and I/O capabilities in ProLiant servers to achieve optimum performance and reduce bottlenecks across the system. In some cases, HP has developed their own chipsets—such as the F8 chipset for Xeon MP-based servers—to ensure a balanced, high-performance system.[12] The F8 chipset uses industry-standard processors, memory DIMMs, and PCI adapters in an industry-proven, SMP architecture that provides performance gains of up to 4X on the system bus, 4X on the I/O, and 5X on the memory over the previous 8-way Profusion chipset.

The latest models of ProLiant servers include the Xeon processor architectures that provide significant performance increases. The Xeon architecture includes:

- 20-stage pipeline, twice the length of the sixth-generation IA-32 architectures
- Additional execution trace cache and in some processors, an additional level three cache
- Quad-pumped data bus and a double-pumped address bus, for data bus bandwidths of up to 533 MHz.

---

[11] For more information, see the HP website at http://h18020.www1.hp.com/products/servers/benchmarks/index.html
[12] For more information about the F8 chipset, see the website at http://h18004.www1.hp.com/products/servers/technology/whitepapers/adv-technology.html - processor

HP leverages industry-standard technologies and in many cases leads the standards groups to enhance those technologies, such as PCI-X, PCI hot plug, double-data rate (DDR) memory, and future fabric-based I/O protocols such as iSCSI (Internet Small Computer Systems Interface). In the case of PCI-X, the conventional PCI bus bandwidth had been 133 MB/s for many years. In great part due to HP efforts to enhance the PCI protocol, PCI-X technology increases the bus capacity by more than 8X, up to 1066 MB/s with the 64-bit, 133-MHz PCI-X bus.

The use of DDR memory in ProLiant servers provides effective bus frequencies that are twice that of SDRAM because DDR memory transfers data on both the rising and falling edges of the clock signal. DDR memory is another example of HP using evolutionary technology, rather than revolutionary technology, to provide higher-performing systems while maintaining the cost benefits of industry-standard components.

Qualified platforms

As of this writing, HP has 18 qualified platforms for ESX Server, far more than any other vendor. Qualified ProLiant platforms include powerful 8-way servers as well as the dense 2- and 4-way server blades such as the ProLiant BL20p G2 and the BL40p. In all of these servers, the combination of high performance and high availability offer customers a wide choice in meeting their individual solution needs and price points when deploying VMware virtual machines. VMware maintains the complete list of qualified ProLiant platforms, available at: http://www.vmware.com/hp.

Because customers often have consolidated storage area networks (SANs) before they implement a VMware solution, HP has qualified a broad range of StorageWorks SAN solutions. These vary from entry-level to massive, enterprise-class storage arrays to meet the varied needs of customers. As of this writing, HP has qualified the following products with VMware ESX Server 2.0.1:

- StorageWorks Modular SAN Array 1000
- StorageWorks Enterprise Virtual Array 3000
- StorageWorks Enterprise Virtual Array 5000
- StorageWorks DiskArray XP48
- StorageWorks DiskArray XP128
- StorageWorks DiskArray XP512
- StorageWorks DiskArray XP1024

The qualification includes basic connectivity, multipathing with host bus adapter failover,[13] and multipathing with storage port failover[14] to provide continuous access to data. VMware maintains the complete list of qualified SANs, available at: http://www.vmware.com/pdf/esx2_SAN_guide.pdf.

ProLiant management advantages

ProLiant servers have numerous technologies to provide effective management of the server platforms, including Insight Manager 7, the Lights-Out products, and the ProLiant Essentials Rapid Deployment Pack.

Insight Manager 7 provides powerful monitoring and control of ProLiant servers to maximize system uptime and reduce total cost of ownership. Insight Manager 7 automatically discovers devices in the network, provides inventory reporting, proactive notification of problems, version control, and access to hundreds of Management Agents. Insight Manager 7 enables these management functions from anywhere at anytime using its web-based browser access.

---

[13] *Multipathing* refers to the ability of ESX Server to handle multiple paths to the same storage device. *Host bus adapter failover* means that the server will tolerate adapter and switch failure only.
[14] *Storage port failover* refers to a configuration in which the server is attached to multiple storage ports and tolerates storage port failures.

Furthermore, HP and VMware have worked closely to integrate VMware products with Insight Manager 7. When a customer uses VMware ESX Server on a ProLiant server and manages it with Insight Manager 7, Insight Manager 7 identifies that server as running VMware ESX Server and provides the user with a link to the VMware management interface (Figure 5).

Figure 5. Insight Manager 7 has a direct link to VMware ESX Server on its interface, as identified by the large arrow below.
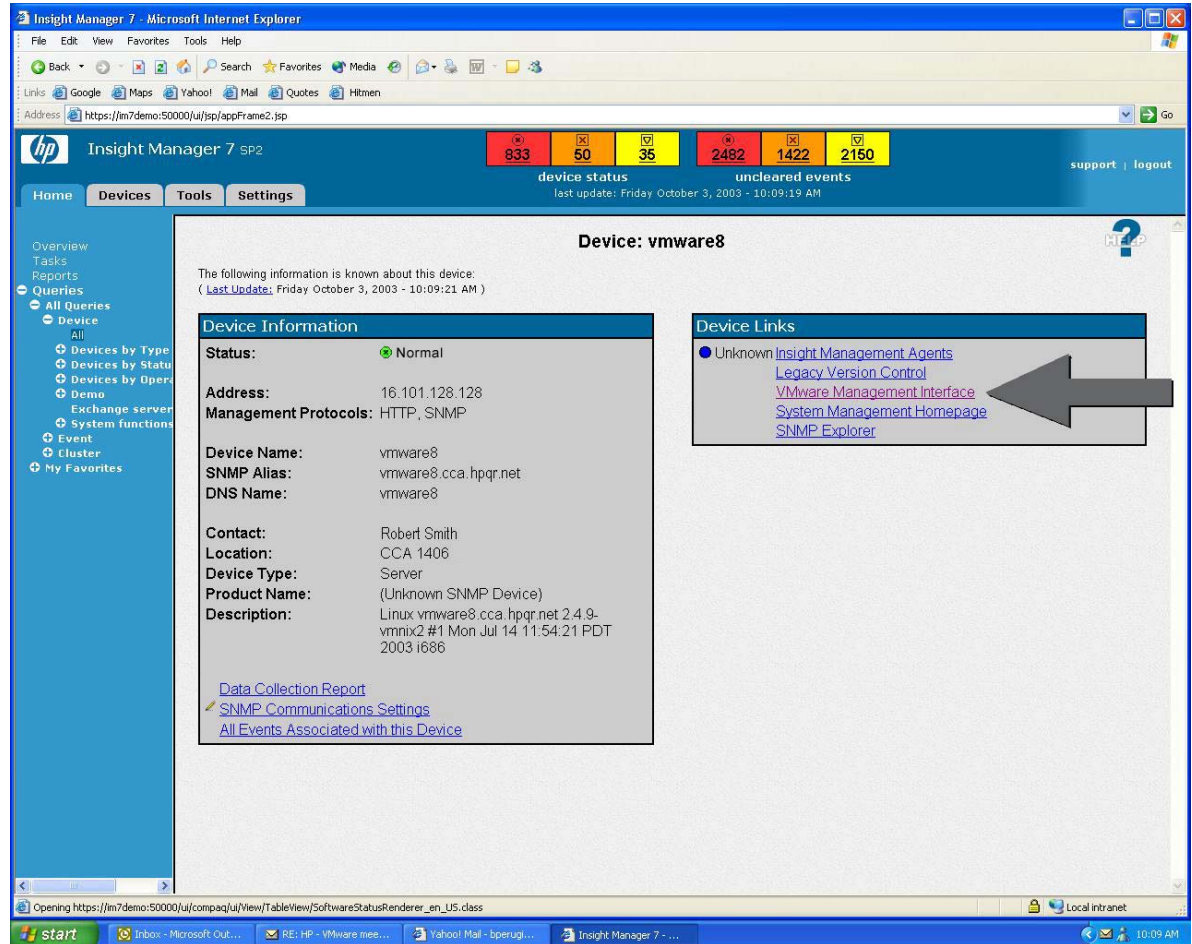
ProLiant servers bring value to any VMware deployment by their Lights-Out technology, which an administrator can use to manage servers remotely.[15] The iLO and RILOE II management processors provide true lights-out abilities; that is, the ability to manage the physical server hardware from any remote location. Administrators can view the server console at any time and through any server state, load operating systems and software over the network, access troubleshooting information about the hardware, reset the server, and view the entire boot sequence from the convenience of the management console.

The ProLiant Essentials Rapid Deployment Pack (RDP) is an integrated HP and Altiris solution that automates the process of deploying and provisioning server software, enabling companies to deploy high volumes of servers quickly and easily. It maximizes a customer's resources by providing a full server build from a remote console, automated server configuration on the fly, and installation of

---

[15] For more information about Lights-Out technology, see the website at
http://h18004.www1.hp.com/products/servers/management/remotemgmt.html

12

standard software sets based on server functions. Customers can use RDP to install ESX Server and GSX Server and to install the virtual machines within ESX Server or GSX Server.

Reseller agreement

A customer buying ProLiant hardware to run VMware is assured of having complete support through HP service and support channels. For customers that are investigating options, HP can provide evaluation licenses and assist with pilot testing of VMware ESX Server, GSX Server, and VirtualCenter products. After selecting the appropriate products, a customer can purchase the VMware software and ProLiant hardware directly from HP. HP provides in-depth support including support and subscription for the VMware software as well as support for the operating systems (Windows, Linux, or Netware) running in a virtual machine. A customer with a VMware virtual machine can call HP for comprehensive support—regardless of whether the issue is with the OS, hardware, or VMware software.

Scenarios for server consolidation

VMware software can be used for multiple reasons and in different ways to consolidate servers, for example:

- IT proof-of-concept
- Web server consolidation
- Multi-tier architectures
- Development sandbox
- Standby server

Consolidating servers may be an extremely large undertaking involving hundreds of servers and applications. To ensure that the consolidation effort runs smoothly, an administrator may set up proof-of-concept servers to test the consolidation process, the manner in which applications are combined, how to partition the servers to achieve the best results, and the server performance level needed for the selected applications. Running a pilot program using VMware software allows companies to determine the benefits of the overall IT consolidation effort in a cost-effective manner.

For many service providers, the best and easiest way to support multiple customers has been to have an individual server for each customer. However, because VMware software can create isolated virtual machines and ProLiant servers provide high availability and reliability, the service provider can reliably consolidate many customer applications onto a single physical server platform. The service provider can greatly increase its hardware utilization rates while still meeting customers' demands for reliability and security. Because many different ProLiant servers are VMware qualified, service providers can pick a 2-way, 4-way, or 8-way platform that best meets their needs. Compared to other vendors that have only high-performance platforms qualified with VMware, HP has a broad range of qualified servers from which to choose.

VMware provides robust and unique networking abilities so that a multi-tier architecture can be built on a single physical server. Administrators can define virtual network adapters within each virtual machine. A virtual network adapter can be mapped to a dedicated NIC on the physical server, or the virtual NIC can be bound to a private network segment implemented in memory. For example, administrators can develop a firewall by configuring one virtual machine (virtual machine 1) with two virtual Ethernet adapters. One of those adapters connects to the physical NIC. The other virtual Ethernet adapter is bound to the private virtual network. A second virtual machine (virtual machine 2) connects only to the private virtual network. By running firewall or filtering software in virtual machine 1, the administrator can construct an effective firewall without the need for additional hardware and with high-performance virtual networking between the machines (see Figure 6).

Figure 6. Multi-tier architecture inside a single physical server using virtual private networks
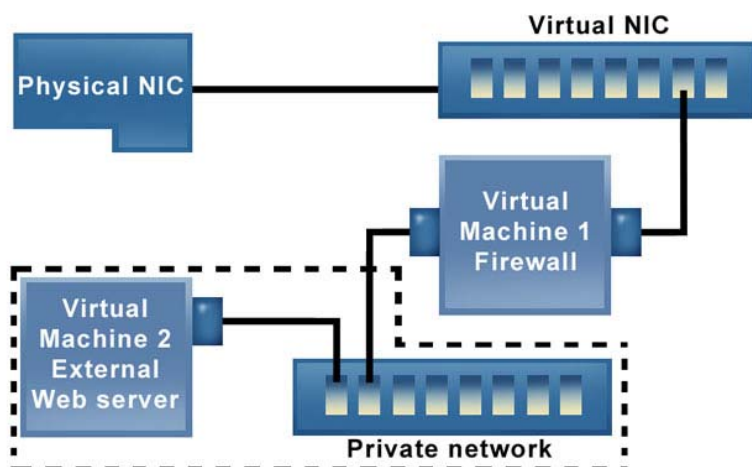
VMware software can simplify development and testing. When software development groups need to test new versions, they can run the applications on virtual machines rather than dedicating a physical server to each test environment. Development groups can use these virtual machines as "sandboxes," or isolated environments, so that they can replicate and exchange software bugs between the development and quality assurance checks. Because the sandbox contains only the software in development, it removes any concern about application interaction or dependencies. The virtual machine environment is encapsulated, making it easy to transport and move the environment to other physical servers, without disturbing the virtual machine configuration. VMware software also enables very fast server provisioning from pre-configured templates, which permits administrators to shorten development and test cycles.

As a final example, using VMware software allows an administrator to keep backup copies of critical server virtual machines in a suspended state, ready to take over for a failed primary server in seconds. VMware software has a suspend/resume feature that saves the memory state of a running virtual machine in a file so that it can be resumed quickly with no boot delay. When a virtual machine is in hot standby mode, it only occupies disk space and does not use any processor or memory resources until powered on. This capability is another way that administrators can provide high availability to their IT environments.

Conclusion

VMware enables administrators to adapt their existing IT infrastructure to use compute resources more efficiently. By allowing multiple applications to reside on a single server, administrators can significantly increase their return on investment, reduce capital expenditures for new equipment, and increase their utilization of existing hardware. By using VMware on ProLiant servers, customers have all the advantages of VMware software with the high availability, high performance, and management features of ProLiant servers.

For customers that desire to keep hardware compatibility for older applications, GSX Server provides the ability to run older applications on newer hardware. ESX Server provides this same capability with better performance, dynamic management capabilities, flexible memory allocations, the ability to run dual processor virtual machines using the Virtual SMP add-in module, and the ability to manage virtual machines across multiple physical servers using VirtualCenter and VMotion.

With the VMware and HP reseller agreement in place, customers can come to HP for solutions help in implementing and deploying any VMware solution—from consulting services to providing evaluation licenses of VMware software. As a customer deploys the VMware solution on ProLiant servers, HP can provide customer support for any issue with a VMware server—whether it is an issue with the ProLiant hardware, the operating system, or the VMware software. The VMware Virtual Infrastructure, coupled with HP ProLiant servers, is a powerful tool that customers can use to build an adaptive infrastructure that is capable of responding quickly and efficiently to fast-changing business needs.

Glossary

guest OS: The operating system running inside a virtual machine is called a guest operating system.

host OS: The operating system running on the physical server hardware on which the VMware GSX Server software is installed.

SMP: symmetric multi-processing

virtualization: The abstraction or idealizing of resources so that physical characteristics are masked.

Appendix

The following table summarizes the differences between ESX Server and GSX Server.

Capability	ESX Server	GSX Server
Requires a host OS	No	Yes
Servers supported	Up to 32-way multiprocessor	Up to 16-way multiprocessor
Maximum virtual machines per server	80	64
Maximum main memory	64 GB	64 GB
Maximum memory per virtual machine	3.6 GB	2.0 GB
Pre-allocates memory resources	No	Yes, pre-allocates memory for each virtual machine
Memory swap files	Up to eight files, 64 GB each	No
Memory over-commitment (virtual memory)	Up to 200%	No
Runs directly on hardware	Yes	No
Provides dynamic resource allocation	Yes	No
Supports clustering across different hosts	Yes	No
Supports virtual SMP	Yes	No
Supports VirtualCenter	Yes	No

For more information

For more information about ProLiant IT consolidation, see the HP website at:
http://h18004.www1.hp.com/products/servers/consolidation/

For more information about the HP and VMware alliance, see:
http://www.hp.com/solutions/vmware

For more information about VMware, see the website at: http://www.vmware.com

Call to action

To help us better understand and meet your needs for ISS technology information, please send comments about this paper to: TechCom@HP.com

TC040101TB, 01/2004