ISS Technology Update

Volume 8, Number 7

Keeping you informed of the latest ISS technology

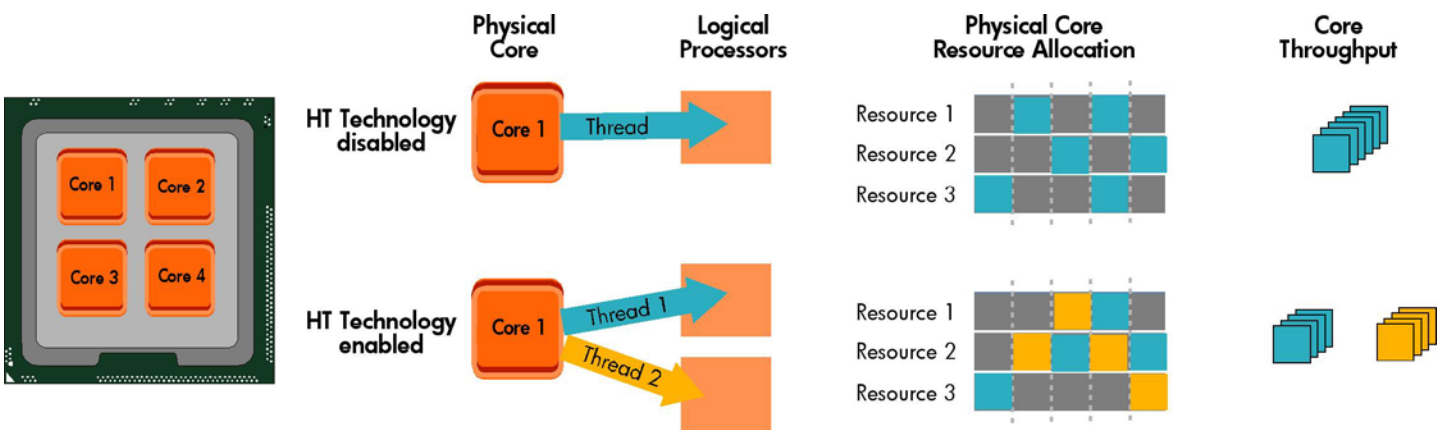
Recently published industry standard server technology papers

Title	URL
Configuring SATA and SAS in HP ProLiant SL6000 servers	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01893303/c01893303.pdf
HP ProLiant SL6000 Scalable System technology	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01938587/c01938587.pdf
Implementing Microsoft® Windows Server® 2008 R2 Foundation on HP ProLiant servers	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01923164/c01923164.pdf
Implementing Microsoft® Windows Server® 2008 R2 Hyper-V™ and Microsoft Hyper-V Server 2008 R2 on HP ProLiant servers	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01925882/c01925882.pdf
HP ProLiant DL585 G5/G6 server technology	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01895559/c01895559.pdf
The Intel processor roadmap for industry-standard servers, 10th Edition	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00164255/c00164255.pdf
The AMD processor roadmap for industry-standard servers, 6th Edition	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00428708/c00428708.pdf
HP BladeSystem c-Class SAN connectivity	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01096654/c01096654.pdf
Implementing UPS power management applications with HP Systems Insight Manager, 3rd edition	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00383560/c00383560.pdf

Using Intel® Hyper-Threading Technology in HP ProLiant G6 servers

Intel® Xeon® four-core processors based on the Intel Microarchitecture (Nehalem) re-introduce support for Hyper-Threading (HT) Technology, also known as simultaneous multi-threading. Each core has its own physical resources (architectural state, cache, registers, and execution units), enabling it to appear as two logical processors to the operating system (OS). HT Technology allows the OS to schedule two instruction streams (threads) to each core to share its physical resources (see Figure 1). For example, the OS scheduler in Windows 2008 R2 prefers to send a software thread to an idle physical core to avoid using two logical processors on a single core. First, the OS scheduler checks to see if an idle core is available. If an idle core is available, the scheduler assigns a new thread to that core. If an idle core is not available, the scheduler assigns the thread to a core's logical processor. However, if an idle physical core becomes available after the operation starts, to increase processor performance, the OS scheduler uses time-slicing to migrate the thread to the idle core.

Figure 1. HT Technology allows the OS to schedule two threads to each core to share the core's physical resources.



Some workloads, such as Internet and e-commerce applications and multi-tasking, may receive significant (20% to 30%) performance gains from HT Technology. Performance gains vary based on the hardware and software used. Alternatively, applications that routinely use high percentages of processor resources, such as high performance computing (HPC), may not receive a measurable performance increase from HT Technology. HPC applications make better use of processor resources by minimizing memory latencies through cache optimizations. So, although the OS may present two threads to a core as if there are two logical processors, both threads will not have access to all of the core's resources, resulting in a minor performance gain. If thread contention for the core's resources is high, processor performance may actually decrease. Intel® Turbo Boost Technology complements HT Technology by increasing the performance of multi-threaded workloads. Turbo Boost Technology increases the clock frequency of all active cores when some cores are idle or the processor is operating below power and thermal design points set by the user.

Disabling HT Technology

Test applications thoroughly, or contact the software manufacturer, to determine whether HT Technology can provide a measurable performance gain. If using HT Technology degrades application performance, it may be beneficial to disable it through the ROM Based Setup Utility (RBSU).

Additional resources

For additional information on the topics discussed in this article, visit these websites:

Resource	URL
How to Determine the Effectiveness of Hyper-Threading Technology with an Application	http://software.intel.com/en-us/articles/how-to-determine-the-effectiveness-of-hyper-threading-technology-with-an-application/
Dual Processors, Hyper-Threading Technology, and Multi-Core Systems	http://www.intel.com/cd/ids/developer/asmo-na/eng/dc/hpc/choice/200677.htm?page=2

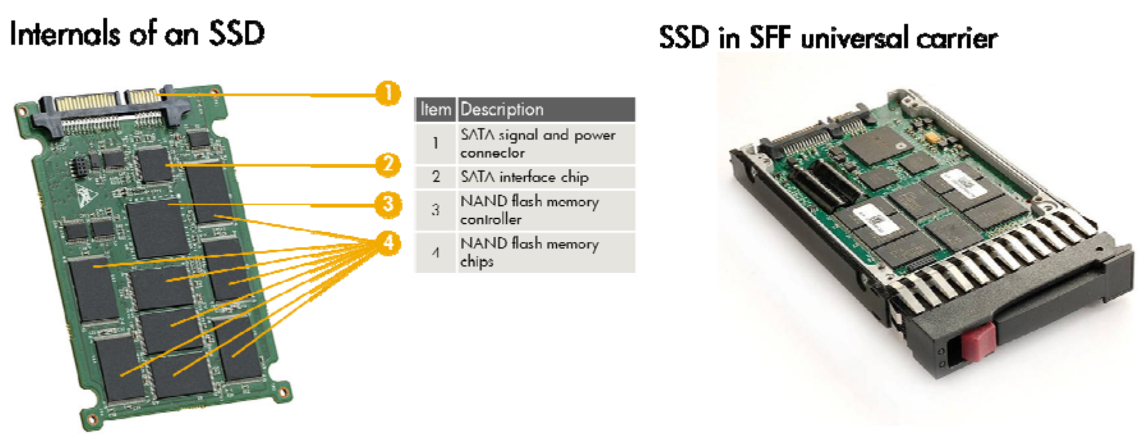
2nd generation HP 3Gb/s SATA solid state drive technology expands SSD functionality and performance

Second generation HP 3Gb/s SATA solid state drives (SSDs) for ProLiant servers are the next step in the evolution of SSDs that can be used in server environments. These SSDs feature several significant advances from the non-hot plug first-generation SSDs that were introduced for selected ProLiant BladeSystem servers in 2008:

- Increased capacities of 60 GB and 120 GB
- SFF and LFF hot plug supported across most of the ML, DL, and BL G5 and G6 ProLiant servers
- Advances in controller technology that provide significant improvements in read and write performance

In addition to these improvements, the new SSDs retain the original ruggedness that made them well suited for environments where traditional disk drives cannot operate. HP 3G SATA SSDs can tolerate significantly higher shock and vibration levels than traditional disk drives, as well as temperatures up to 60° C.

Figure 1. HP 3Gb/s SATA solid state drives tolerate higher shock, vibration, and temperatures than traditional disk drives.



SSD Performance

HP 3Gb/s SATA SSD read and write performance is significantly improved over first-generation SSDs, as shown in Table 1. The new SSDs are particularly appropriate for applications that require storage systems capable of performing large numbers of random reads per second. Traditionally, solutions designers have used RAID configurations with large numbers of disk drives to achieve this. Using the new HP solid state drives, a single SSD can deliver two to six times as many random reads per second (IOPs) as an entire RAID 0 array of 12 SAS disk drives. This makes the new drives ideal for read-intensive, high-IOPs applications that do not require the large storage capacities that RAID configurations can offer.

Table 1. Performance comparison of SATA/SAS drives with SSDs for servers

	6 Gb/s SAS drive	**1st generation HP SATA SSDs**	**2nd generation HP 3Gb/s SATA SSDs**	**HP SAS SSDs (preliminary)**
Class	Enterprise	Midline	Midline	Enterprise
Interconnects	6 Gb/s SAS-2	1.5 Gb/s SATA	3.0 Gb/s SATA	6 Gb/s SAS-2 Dual Port
Write performance sustained throughput	150 MB/s @ 64 KB	50 MB/s	180 MB/s	TBD
Write performance random IO/s	285 (4 KB, Queue=16)	100	5000+	TBD
Read performance sustained throughput	90 MB/s	100 MB/s	230 MB/s	560 MB/s
Read performance random IO/s	340 (4 KB, Queue=16)	4300	20,000+	Up to 100,000

HP SSD futures

The new HP 3Gb/s SATA SSDs are the first SSD products designed to operate across most of the ProLiant server line. They continue to use the more robust single-level cell NAND flash technology and are suitable for use in Midline (MDL) drive environments where unconstrained workloads and a 100% duty cycle are not required. The NAND flash memory used in SSDs has a limited life cycle for writes. Although advanced techniques such as wear-leveling and over-provisioning are used to help overcome this, HP 3G SATA SSDs are still better suited for read-intensive environments where writes are 30% or less of the total IO load.

SSD technology is still rapidly evolving. HP hopes to introduce its first Enterprise class SAS SSDs in 2010. These drives will be engineered for unconstrained read and write workloads and 100% duty cycles. This will match the basic reliability characteristics of Enterprise class disk drives while providing the IO performance advantages of SSDs. HP may also introduce server SSDs based on multi-level cell NAND flash memory. These products will provide a solution for application environments that need higher capacity SSDs, but do not require support for unconstrained IO workloads.

Additional resources

For additional information on the topics discussed in this article, visit the following websites:

Resource	**URL**
HP Solid State Storage	www.hp.com/go/solidstate
Technology brief - Solid state drive technology for ProLiant servers	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01580706/c01580706.pdf
HP ProLiant drives (including solid state drives)	www.hp.com/products/harddiskdrives

Deploying VMware ESXi on HP ProLiant BL400 Series server blades

Because of their robust memory and network capabilities, the HP ProLiant BL490c and HP ProLiant BL495c server blades are ideal systems for deploying virtualized infrastructures. The ESXi variant of VMware vSphere 4 is being adopted by many customers as a next generation hypervisor due to the stronger security that results from the absence of a console operating system. Many customers are interested in the best practices for deploying this basic building block of their vSphere infrastructure on HP BladeSystem.

NOTE
This article refers to the ProLiant BL490c G6 server blade for simplicity, but the principles are also applicable to the ProLiant BL495c server blade.

Configuration alternatives

There are several alternatives for configuring an ESXi installation on the BL490c G6 server, including installing the image on one of the following media:

- Flash media
- Solid state drive
- Shared storage

Flash media

Many customers are interested in using the most widely known technique that installs the image on an internal USB flash drive or SD card. This eliminates the need for any internal disk drives to host operating system files. The virtual machines (VM) hosted on this infrastructure are stored on external networked storage, which can range from low-cost, high-performance shared SAS to iSCSI to Fibre Channel SAN.

Using a server with an integrated ESXi hypervisor on USB or SD flash memory provides a boot and run environment almost out of the box. However, customers should be aware that this type of flash memory uses multi-level cell (MLC) technology, which is most typically used for light duty applications such as photo storage and occasional file transfer. Despite elaborate wear-level techniques, these devices have a failure rate that is somewhat higher than HP enterprise SAS drives. However, failure of the flash device typically does not cause a server to crash because the ESXi image is not really critical after the system boots. Eventually, the server will have to be taken offline and repaired to ensure continued operation.

Solid state drive

Another alternative is to use the ESXi installable kit to install the image on a solid state drive (SSD) within the ProLiant BL490c G6 server blade chassis. While it is possible to create a Virtual Machine File System (VMFS) partition on the SSD to host VMs, this typically would only be done for the smallest of test environments. In general, the preferred model would be to store the VMs on a high-performance shared SAS, iSCSI, or Fibre Channel storage array to take advantage of capabilities such as VMotion and high availability.

Customers who want a higher level of reliability than that attainable with USB or SD flash media may wish to use the SSD alternative. The flash memory in HP server SSDs uses single-level cell technology, which has a much higher reliability than the MLC technology used in USB and SD card flash media. In first generation SSD drives, this is comparable to the reliability of the HP mid-line rotating media drives used in HP servers. However, failures will occur, and the customer environment must tolerate the ability to take a server offline from time to time to change a failed drive. It is important to note that the SSDs in the current ProLiant BL490c G6 server blades cannot be combined into a mirrored RAID set due to the integrated SATA controller used within these systems. Thus, a single SSD is the most typical configuration within ESXi servers unless there is a specific customer-defined purpose for including a second drive.

Shared storage

A third alternative would be to install the ESXi installable image on shared storage, which could be the same as that used for VM files or a lower-cost dedicated array. Boot from SAN deployments are the most commonly used configuration of this type.

Failure recovery

Failure certainly sounds like a bad word, but in a well-designed vSphere infrastructure based on HP BladeSystem, the effects may not be all that serious. Deploying a clustered environment of ESXi hosts provides greater high availability and fault tolerance. With virtual machines hosted on shared storage, VMotion can easily be used to transfer VMs from a server with a failing flash-based system device to another vSphere host in the cluster. The malfunctioning server can then be pulled from the enclosure and repaired. However, with an infrastructure based on Virtual Connect Flex-10 technology, a new server can simply be inserted into the empty bay, and it will assume the profile of the failed system. In this way, a fully functional vSphere host can be restored in minutes.

For customers who cannot tolerate even this level of down time due to a failing flash-based ESXi install, the shared storage boot scenario is probably best. For example, a BladeSystem configuration using SAS switch interconnect modules and an HP MSA2000sa G2 SAS shared storage array provides a robust environment that can be configured to tolerate multiple levels of failure and provide maximum uptime for all vSphere host servers. Another alternative might be to consolidate the system drives for a number of blade servers on an HP StorageWorks MDS600 Disk System. More elaborate configurations based on Fibre Channel SAN arrays are also possible.

Summary

VMware ESXi on HP BladeSystem with ProLiant BL490c or ProLiant BL495c server blades provides an ideal infrastructure for a robust and secure vSphere 4 deployment. HP provides system configurations that meet a wide range of customer requirements. Customers are encouraged to review their specific needs and expectations with their HP sales or reseller contact to determine the best configuration for their application.

Additional resources

For additional information on the topics discussed in this article, visit:

Resource	URL
Virtualization technology communications	http://h18004.www1.hp.com/products/servers/technology/whitepapers/virtualization/index.html
HP ProLiant BL490c G6 Server series	http://h10010.www1.hp.com/wwpc/us/en/sm/WF25a/3709945-3709945-3328410-241641-3328419-3884113.html
HP ProLiant BL495c G6 Server series	http://h10010.www1.hp.com/wwpc/us/en/sm/WF25a/3709945-3709945-3328410-241641-3328419-3948609.html

Meet the Expert—Alan Goodrum

Alan Goodrum is a Fellow with HP Industry Standard Servers (ISS). His career with HP actually began in 1985 with Compaq and over the last several years, his primary focus has been power management both in servers and the data center as a whole. Today, power management is arguably the "hottest" area of the server industry, but it hasn't always been. Dwight Barron, Alan's manager says, "Alan started working on power efficiency of servers long before it was recognized as one of the most critical problems to be solved by the computer industry." Power efficiency appears to be a perfect venue for Alan's skills because "he understands and solves problems by simultaneously working at multiple levels of abstraction and detail," according to Dwight.

Name: Alan Goodrum
Title: Fellow, HP Industry Standard Servers
Years at HP: 24
University: BSEE 1975 University of Houston
MSEE 1985 University of Houston
Patents: 26

A solid foundation

Alan's experience has taught him that the spiritual aspects of life are the most important. Most of his activities outside of work revolve around his church. He teaches a weekly adult Bible study class and visits nursing homes around the Tomball area as part of a group of volunteers. Several years ago, he discovered that residents of nursing homes appreciate a smile, a kind word, and a prayer. He also discovered that "once you have a solid foundation based on who you are and where you want to go, all the other things just seem to fit into place."

In the beginning, there was math and science

His enjoyment of math, science, and problem solving drove him to become an engineer. He liked the challenge of understanding the tools of mathematicians and scientists and then applying those tools to build things that actually work.

His favorite accomplishments

Alan's favorite accomplishment to date is Dynamic Power Capping. Several years ago, he was on the team that conceived the technical possibility of allowing system administrators to manage server power consumption. Dynamic Power Capping controls the maximum power that a server can consume, allowing administrators to fit more servers into the same power and cooling space. The goal of Dynamic Power Capping is not to save energy but to provide predictability in server power consumption. Predictability allows for better planning, which can delay enormously expensive data center expansions. Alan helped with early testing of the power capping concept; nurtured it through a complex joint development process with Intel; drove the detailed functional specifications; and finally saw it launched as a key Thermal Logic Technology in ProLiant G6 servers.

Alan had a similar sense of satisfaction during the development of the PCI-X bus. PCI-X extended the bandwidth of the popular PCI bus to 133 MT/s (megatransfers per second) and later to 266 MT/s. The concept was first proposed by some key ASIC designers at Compaq. And although other companies had proposed abandoning the PCI bus, a coalition of Compaq, HP, and IBM eventually took the idea to the PCI SIG. Alan chaired the committee that published the specification. Eventually, Compaq servers were the first to market with PCI-X slots. Alan noted that for several years a major competitor used our PCI-X ASIC inside their servers but had to put a heatsink on top of it to cover up the Compaq logo.

He has a strong customer focus

Alan believes that customer input is critical, but finding the right solution can be difficult because often, customers can describe only their problems, not the solutions. He says that the key is to talk to enough customers to be able to anticipate the kinds of solutions they will find useful. Dwight believes that Alan's strength resides in understanding the business constraints that customers face in adopting new solutions and then helping them work through those issues.

Industry standard server technical papers can be found at www.hp.com/servers/technology.

Contact us

Send comments about the *ISS Technology Update* to TechCom@HP.com.

To subscribe to the *ISS Technology Update*, click mailto:techcom@hp.com?subject=TechUpdate_subscription

Follow us on Twitter .

TC091109NL