



Technology and architecture of HP ProLiant Intel-based 300-series G6 (Generation 6) servers

Technology brief

Abstract.....	3
Introduction.....	3
Intel Xeon 5500 Series processor technology	3
Multi-level caches.....	3
Integrated memory controller.....	4
QuickPath Interconnect controller.....	4
Hyper Threading.....	5
Turbo Boost technology.....	6
Memory.....	6
DDR-3.....	6
DIMM Choices.....	7
Memory Mirroring with DDR3.....	8
Memory channel interleaving.....	9
I/O technologies	9
PCI Express technology.....	9
HP Smart Array and SAS/SATA technology.....	10
SAS-2.....	10
New generation HP Smart Array controllers.....	11
Battery backed write cache.....	11
Zero Memory RAID.....	11
Smart Array Advanced Pack (SAAP).....	11
Software RAID.....	12
Networking technologies.....	12
Power and thermal technologies.....	13
Efficient power delivery.....	13
Common slot power supplies.....	13
Increasing power efficiency in redundant power operation.....	14
Voltage regulation.....	14
Improved thermal sensors and fan control.....	14
Dynamic Power Capping.....	15
Phase shedding.....	16
Phase shedding based on the installed processor.....	16
Memory phase shedding.....	16
Dynamic CPU phase shedding.....	17
Managing processor technologies.....	17
QuickPath Interconnect power.....	17
Disabling processor cores.....	17
Minimum processor idle power state.....	17
Managing memory technologies.....	17
Memory channel interleaving.....	17
Maximum memory data rates.....	17

Managing I/O technologies	18
Disable PCIe 2.0	18
I/O Power Provisioning	18
Power profiles	18
Power Performance Benchmarks	18
Security	19
Trusted Platform Module	19
BitLocker Drive Encryption	19
Systems management and monitoring	19
HP Insight Control Environment	20
HP Insight Dynamics – VSE	20
HP ProLiant Onboard Administrator powered by iLO 2	20
OS support	20
Summary	20
For more information	21
Call to action	22

Abstract

This technology brief describes the key technologies implemented in Intel-based HP ProLiant 300-series G6 servers.

Introduction

HP ProLiant 300-series G6 servers have been the focus of extensive engineering and development. Characterized by increased performance, better power efficiency, and more powerful management tools, the servers include these new technologies:

- The Intel® Xeon® Processor 5500 Series
- Double Data Rate-3 (DDR-3) DIMMs
- Thermal sensors incorporated throughout the ProLiant 300-series G6 servers
- The ProLiant Onboard Administrator powered by Integrated Lights-Out 2 (iLO 2) that delivers power and temperature management through multiple sensors and fan control
- Dynamic Power Capping that limits peak power consumption without impacting performance and without risk of overloading data center branch circuits
- Input/output (I/O) technologies such as PCIe generation 2 (PCIe 2.0) and faster Smart Array controllers that incorporate common form factor components
- “Right Size” common slot power supplies in multiple sizes to provide the required amount of power and improve power efficiency

The technologies discussed in this paper are implemented in all Intel-based ProLiant 300-series G6 servers. Exceptions are noted where different levels of technology implementation or service exist among individual ProLiant 300-series G6 platforms.

For complete specifications of all ProLiant 300-series servers, see the HP website:
www.hp.com/products/servers/platforms.

Intel Xeon 5500 Series processor technology

ProLiant Intel-based 300-series G6 servers contain the Xeon 5500 Series processors. The Xeon 5500 Series microarchitecture is built on hafnium-based, 45 nanometer, high-k metal gate silicon technology. Xeon 5500 Series processors use a new material combination of high-k gate dielectrics and conductors which reduce electrical leakage—enabling smaller, more energy-efficient, and higher performance processors. This new silicon technology, along with technologies such as distributed shared memory, simultaneous multi-threading, Turbo Boost, and dynamic power management, represent the technological heart of the HP ProLiant G6 servers.¹

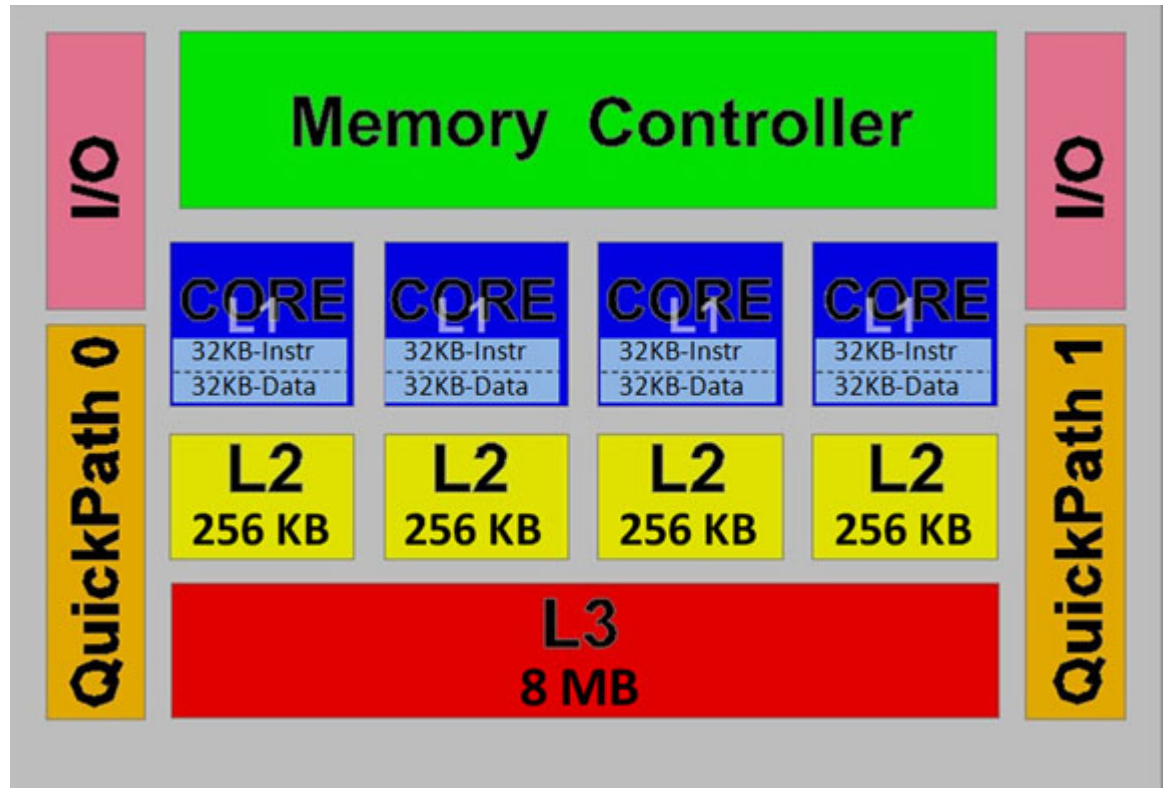
Multi-level caches

Xeon 5500 Series processors have a three-level cache hierarchy (Figure 1):

- An on-core 64-kilobyte Level 1 cache, split into two 32 kilobyte caches: one for data and one for instructions
- 256-kilobyte, Level 2 cache for each core to reduce latency
- A Level 3 cache of up to 8 megabytes shared by all cores

¹ For additional information about Intel processors, see the HP technology brief titled “The Intel processor roadmap for industry-standard servers”:
<http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00164255/c00164255.pdf>.

Figure 1. Block diagram of three-level cache hierarchy for Intel Xeon 5500 Series processors



The Level 3 cache is shared and inclusive, which means that it duplicates the data stored in the Level 1 and Level 2 caches of each core. This guarantees that data is stored outside the cores and minimizes latency by eliminating unnecessary core snoops to the Level 1 and Level 2 caches. Flags in the Level 3 cache track which core's cache supplied the original data. Therefore, if one core modifies another core's data in Level 3 cache, the Level 1 and Level 2 caches are updated as well. This eliminates excessive inter-core traffic and ensures multi-level cache coherency.

Integrated memory controller

Instead of sharing a single pool of system memory, each processor accesses its own dedicated DDR-3 system memory directly through an integrated memory controller. Three memory channels from each memory controller to its dedicated memory provide a total bandwidth of 32 gigabytes per second. The three memory channels eliminate the bottleneck associated with earlier processor architectures in which all system memory access took place through a single memory controller over the front side bus. In cases in which a processor needs to access the memory of another processor, it can do so through the QuickPath Interconnect.

QuickPath Interconnect controller

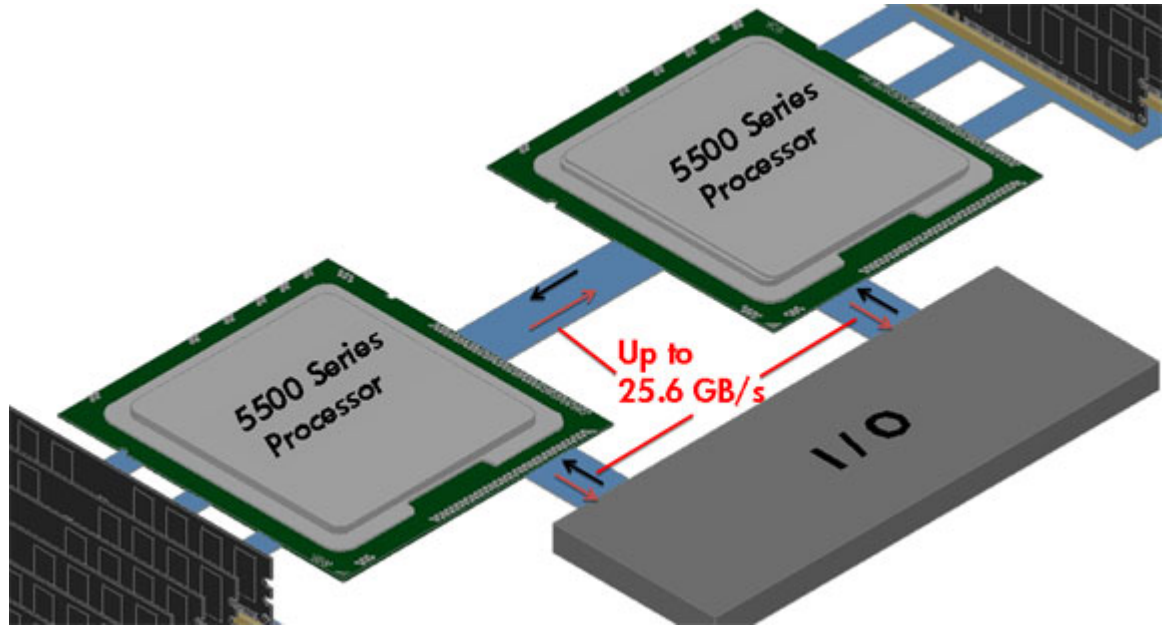
Xeon 5500 Series processors attain their performance potential through the Intel QuickPath Architecture (Figure 2): high-speed, point-to-point interconnects directly connect the processors. The Intel QuickPath Architecture also connects each processor to distributed shared memory and to the I/O chipset.

Each QuickPath Interconnect consists of two unidirectional links that operate simultaneously in each direction using differential signaling. Unlike a typical serial bus, the QuickPath interconnects transmit data packets in parallel across multiple lanes and packets are broken into multiple parallel transfers.

Each link is comprised of twenty 1-bit lanes. A maximum of 16 bits are used to transfer data; the remaining 4 bits are used for the protocol and error correction. The interconnect performs a maximum of 6.4 gigatransfers per second and has a bandwidth of 12.8-gigabytes per second in each direction, for a total bandwidth of 25.6 gigabytes per second.

If an application requests data from the memory of another processor, the QuickPath interconnect uses high-bandwidth inter-processor communication to retrieve the data.

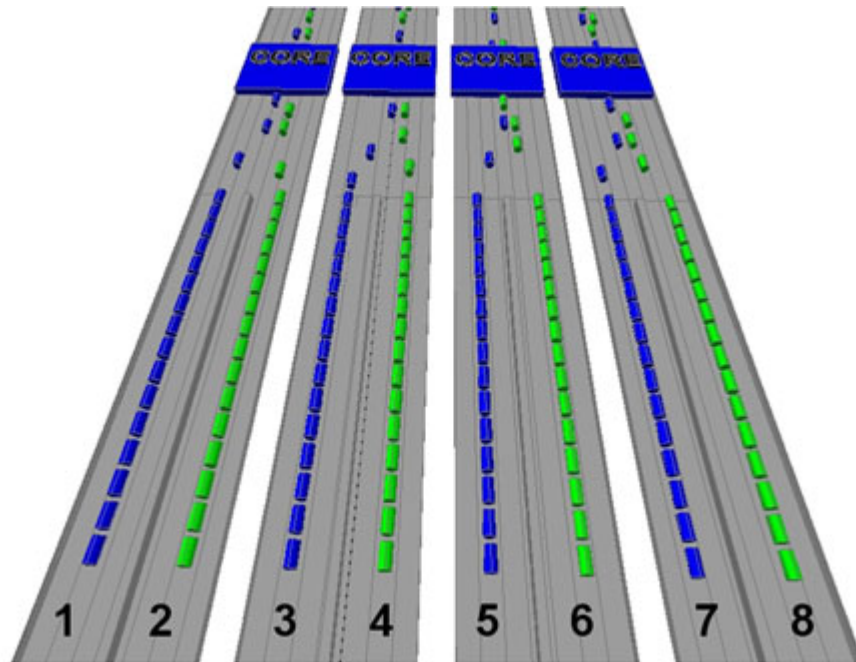
Figure 2. Block diagram of QuickPath architecture



Hyper Threading

Simultaneous Multi-threading Technology, or SMT, is an enhanced version of Intel's Hyper-Threading Technology. SMT lets each core execute two computational threads at the same time. A single processor can execute up to eight threads simultaneously (Figure 3). Compared to the previous Intel core architecture, the high-bandwidth memory subsystem supplies data faster to the two computational processes, and the low-latency cache hierarchy allows more instructions to be processed simultaneously.

Figure 3. Diagram representing SMT technology in the 5500 Series processor



Incorporating simultaneous multithreading improves performance at a proportionately small additional power cost, improving performance per watt (W), and allowing ProLiant G6 servers to do more using the same—or less—power as previous-generation, Intel processor-based servers.

Turbo Boost technology

Intel's Turbo Boost technology complements Simultaneous Multi-threading Technology by increasing the performance of both multi-threaded and single-threaded workloads. For workloads and applications that do not benefit from multi-threading, Turbo Boost technology can provide better performance. Turbo Boost is engaged by default and automatically increases the clock frequency of active cores operating below power and thermal design points determined by the processor. The maximum frequency depends on the number of active cores and varies based on the specific configuration on a per-processor-number basis. Turbo Boost technology is operating system independent and operating systems that are Advanced Configuration and Power Interface-aware (ACPI) require no changes to support Turbo Boost technology.

Memory

The Xeon 5500 Series processor connects directly to memory rather than through a chipset. The Xeon 5500 Series processor supports only DDR-3 dual inline memory modules (DIMMs). DDR-3 is part of the synchronous dynamic random access memory (SDRAM) group of technologies. HP manages all processor functionality options through the G6 server Basic Input/Output System (BIOS) and ROM Based Setup Utility (RBSU). These options include Memory Mirroring and are discussed later in this section.

DDR-3

DDR-3 has several key enhancements including an 8-bit prefetch buffer for storing data before it is requested. By comparison, DDR-2 has a 4-bit buffer. For DDR-3, the data signal rate can increase to

1333 Megatransfers per second (MT/s). While this is commonly referred to as having a speed of 1333 MHz, the maximum clock speed for the DIMMs is actually 667 MHz and the signal is double-pumped to achieve the data rate of 1333 MT/s. DDR3-1333 DIMMs can operate at clock speeds of 667 MHz, 533 MHz and 400 MHz; with corresponding data rates of 1333, 1066, and 800 MT/s. The three memory channels between the processor's integrated memory controller and the dedicated DDR-3 memory provide a total bandwidth of 32 gigabytes per second.

Energy efficiency is improved over previous generations. HP DDR-3 DIMM modules incorporate an integrated thermal sensor that signals the chipset to throttle memory traffic to the DIMM if its temperature exceeds a programmable critical trip point. By using the data from these thermal sensors, the ProLiant G6 servers can reduce fan speed and therefore reduce power consumption when memory is idle. Some third-party DIMMs may not include this DIMM thermal sensor. The BIOS in ProLiant G6 servers verifies the presence of the thermal DIMM sensor during POST. If the thermal sensor is absent, a POST message will warn that the DIMM does not have a thermal sensor, and the fans will be forced to run at higher speed (requiring more power).

Operating voltage for DDR-3 has been reduced. DDR-2 operates at 1.8V, while DDR-3 currently operates at 1.5V, consuming less power than DDR-2 DIMMs at the same capacity and speed. DDR-3 DIMMs have the same number of pins and are the same physical size as DDR-2 DIMMs, but they are electrically incompatible and have a different key notch location.

DIMM Choices

DDR-3 is available as both Unbuffered Dual In-line Memory Modules (UDIMMs) and Registered (buffered) Dual In-line Memory Modules (RDIMMs). Both RDIMMs and UDIMMs support error correcting code (ECC).

There are three types of DDR-3 available for ProLiant G6 servers:

- PC3-8500R (RDIMM, ECC compliant) — 1066 or 800 MT/s data rate, depending on memory configuration and processor installed.
- PC3-10600E (UDIMM, ECC compliant) — 1333, 1066 or 800 MT/s data rate, depending on memory configuration and processor installed
- PC3-10600R (RDIMM, ECC compliant) — 1333, 1066 or 800 MT/s data rate, depending on memory configuration and processor installed.

Administrators can configure ProLiant 300-series G6 servers using either RDIMMs or UDIMMs, although RDIMM and UDIMM memory cannot be mixed within a single server. ProLiant 300-series G6 servers have up to 18 DIMM slots, allowing larger memory capacities than with the platforms that used DDR-2. ProLiant G6 servers optimize memory performance by operating DDR-3 memory at the maximum rate possible, based on the memory configuration and processor installed.

When choosing memory configurations for 300-series G6 systems, the following guidelines should prove helpful:

- UDIMM configurations are limited to a maximum of two UDIMMs per memory channel because the memory controller must drive the address and command signals to each DRAM chip on a channel. This results in a 24 GB maximum configuration in ProLiant 300-series G6 servers. Because they require fewer components, UDIMMs are typically less expensive than RDIMMs.
- RDIMM configurations can provide larger memory capacity configuration because the memory controller only drives the address and command signals to a single register chip, thereby reducing the electrical load on the memory controller. Users requiring large memory footprints can install three 8-GB RDIMMs per channel for a total of 144 GB.
- For smaller memory configurations, installing only one or two DIMMs per memory channel can potentially increase memory performance. In many instances this allows administrators to clock the memory channel at a higher data rate.

Processor SKU determines the ability of ProLiant G6 servers to run DDR-3 memory at a top speed of 1333 MT/s. The processor SKU also dictates the range of speeds possible in different DIMM per channel (DPC) configurations. DIMM operating speeds are also subject to memory slot configuration and whether the server in question has 12 or 18 slots. Table 1 displays DIMM speeds possible in 300-series G6 servers with a Xeon X5500 (95W) or W5500 (130W) processor.

Table 1. DDR-3 memory options on 300-series G6 servers supporting Xeon 5500 Series processors running at 95W or 130W

DIMM	Type	Capacity (GB)	Rank	Max. data rate in MT/s DIMMs per channel (DPC)			
				1 DPC	2 DPC in 18 slots	2 DPC in 12 slots	3 DPC in 18 slots
PC3-8500R	RDIMM	4	quad-rank	1066	800	800	N/A
PC3-8500R	RDIMM	8	dual-rank	1066	1066	1066	800
PC3-10600R	RDIMM	2	dual-rank	1333	1333	1333	800
PC3-10600R	RDIMM	4	dual-rank	1333	1333	1333	800
PC3-10600R	RDIMM	8	dual-rank	1333	1333	1333	800
PC3-10600E	UDIMM	1	single-rank	1333	1066	1333*	N/A
PC3-10600E	UDIMM	2	dual-rank	1333	1066	1333*	N/A

* UDIMMs will not run at 1333 MT/s data rates with two DIMMs per channel on systems that support three DIMM sockets per channel. Support for 1333 MT/s data rate with two DIMMs may require a ROM upgrade.

A BIOS setting on select ProLiant 300-series G6 servers allows the PC3-10600R memory modules to run at 1333 MT/s with two DIMMs per channel. HP design and testing have produced system boards with signal integrity that can operate at 1333 speeds in this configuration without degradation.

Only HP branded DIMMs have been fully validated to operate at this speed. Therefore, HP does not recommend configuring the BIOS setting for 1333 MT/s data rates with two DIMMs per channel when using third-party DIMMs, because they may not meet HP's stringent design requirements.

With two DIMMs per channel, operating the memory at 1333 MT/s is supported only on select 300-series G6 servers using the Xeon 5500 Series processors operating at 95W or 130W. The HP ProLiant DL320 G6 cannot take advantage of this unique HP feature as it does not support the 95W or 130W processors.

Memory Mirroring with DDR3

ProLiant 300-series G6 servers support memory mirroring, which protects the system against uncorrectable memory errors that would otherwise result in a system hang or crash. Mirroring occurs when all memory is written to both sets of physical memory in channels one and two. Administrators can configure memory mirroring through RBSU. To implement mirroring with DDR-3, the two memory channels must be populated identically. The third memory channel must be empty.

If an uncorrectable error occurs, the system automatically directs the read to the mirrored location to obtain the correct data. The OS does not revert to Advanced ECC Mode until the DIMM is replaced and the server rebooted. Since each mirrored DIMM is one of a pair, one DIMM can be protected by mirroring while another is degraded. As a result, even after mirroring is degraded by a DIMM failure, the other DIMM in the mirrored pair is still protected by Advanced ECC.

For help with configuring DDR-3 memory in ProLiant G6 servers, use the DDR-3 Memory Configuration Tool found at http://h18004.www1.hp.com/products/servers/options/tool/hp_mementool.html.

Memory channel interleaving

The Xeon 5500 Series processor retrieves data from the memory DIMMs in 64-byte chunks. With channel interleaving, the system is set up so that each consecutive 64-byte chunk in the memory map is physically transferred by means of alternate routing through the three available data channels.

The result is that when the memory controller needs to access a block of physically contiguous memory, the requests are distributed more evenly across the three channels rather than potentially stacking up in the request queue of a single channel. This alternate routing decreases memory access latency and increases performance. Memory channel interleaving, does, however, increase the probability that more DIMMs need to be kept in an active state (requiring more power) since the memory controller alternates between channels and therefore between DIMMs. This is discussed further in the “Power and thermal technologies” section.

For additional information about DDR-3 memory, see the paper titled “Memory technology evolution: an overview of system memory technologies” at

<http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00256987/c00256987.pdf>.

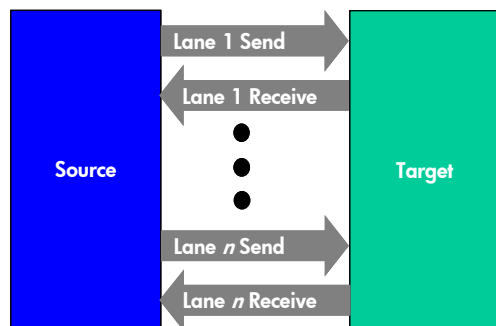
I/O technologies

ProLiant 300-series G6 servers incorporate PCI Express, Serial-Attached SCSI (SAS), and Serial ATA (SATA) I/O technologies. PCI Express lets administrators add expansion cards with various capabilities to the system. SAS is a serial communication protocol for direct attached storage devices such as SAS and SATA hard drives.

PCI Express technology

All ProLiant G6 servers support the PCIe 2.0 specification. PCIe 2.0 has a per-lane signaling rate of 5 Gb/s —double the per-lane signaling rate of PCIe 1.0 (Figure 4).

Figure 4. PCIe data transfer rates



Link size	Max. bandwidth (Send or receive)		Total (Send and receive)	
	PCIe 1.0	PCIe 2.0	PCIe 1.0	PCIe 2.0
x1	250 MB/s	500 MB/s	500 MB/s	1 GB/s
x4	1 GB/s	2 GB/s	2 GB/s	4 GB/s
x8	2 GB/s	4 GB/s	4 GB/s	8 GB/s
x16	4 GB/s	8 GB/s	8 GB/s	16 GB/s

PCIe 2.0 is completely backward compatible with PCIe 1.0. A PCIe 2.0 device can be used in a PCIe 1.0 slot and a PCIe 1.0 device can be used in a PCIe 2.0 slot. Table 2 shows the level of interoperability between PCIe cards and PCIe slots.

Table 2. PCIe device interoperability

PCIe device type	x4 Connector x4 Link	x8 Connector x4 Link	x8 Connector x8 Link	x16 Connector x8 Link	x16 Connector x16 Link
x4 card	x4 operation	x4 operation	x4 operation	x4 operation	x4 operation
x8 card	Not allowed	x4 operation	x8 operation	x8 operation	x8 operation
x16 card	Not allowed	Not allowed	Not allowed	x8 operation	x16 operation

HP Smart Array and SAS/SATA technology

The newest serial, PCIe 2.0 capable Smart Array controllers use Serial Attached SCSI (SAS) technology, a point-to-point architecture in which each device connects directly to a SAS port rather than sharing a common bus as with parallel SCSI devices. Point-to-point links increase data throughput and improve the ability to locate and fix disk failures. More importantly, SAS architecture solves the parallel SCSI problems of clock skew and signal degradation at higher signaling rates.²

The latest Smart Array controllers are compatible with Serial Advanced Technology Attachment (SATA) technology and include the following features to enhance performance and maintain data availability and reliability:

- SAS and SATA compatibility — The ability to use either SAS or SATA hard drives lets administrators deploy drive technology that fits each computing environment. HP Smart Array controllers can manage both SAS arrays and SATA arrays. Smart Array configuration utilities help administrators configure arrays correctly so that data remains available and reliable.
- SAS wide port operations — Wide ports contain four single lane (1x) SAS connectors and the cabling has all four lanes bundled together. SAS wide ports allow balanced SAS traffic distribution across the links for enhanced performance. In addition, wide ports provide redundancy by tolerating up to three physical link failures while maintaining the ability to communicate with the disk drives. The most common use of these wide links is to a JBOD or to an internal server expander connecting to large numbers of drives. No special configuration is required for this functionality.
- SAS expanders — Low-cost, high-speed switches called expanders can combine multiple single links to create wide ports and increase available bandwidth. SAS expander devices also offer higher system performance by expanding the number of hard drives that can be attached to an HP Smart Array controller. SAS expanders are an aggregation point for large numbers of drives or servers providing a common connection. By cascading expanders, administrators can chain multiple storage boxes together. For more information on the HP SAS Expander Card, go to <http://h18004.www1.hp.com/products/servers/proliantstorage/arraycontrollers/sas-expander/index.html>.

SAS-2

SAS-2 and PCIe 2.0 are among the technologies responsible for a significant increase in performance over past generations of Smart Array controllers. The second-generation SAS (SAS-2) link speed³ of 6.0 Gb/s is double the SAS-1 transfer rate. Operation at SAS-2 link speeds requires SAS-2 compliant hard drives. SAS-2 eliminates the distinction between fanout and edge expanders by replacing them with self-configuring expanders. SAS-2 enables zoning for enhanced resource deployment, flexibility, security, and data traffic management. SAS-2 is also backward compatible with SAS-1.

² For more information about SAS technology, refer to the HP technology brief titled “Serial Attached SCSI storage technology” available at <http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01613420/c01613420.pdf>.

³ The Serial Attached SCSI - 2 (SAS-2 or SAS 2.0) is a draft standard, and is the product of the Technical Committee 10t Organization. SAS 2.0 is second generation of SAS and is based upon SAS - 1.1. The SAS-2 specification is available from the 10t website, <http://www.t10.org>.

Beginning with HP product releases in the first quarter of 2009, Smart Array controllers are SAS-2 capable. In fully supported controllers, 6-Gb/s SAS technology allows Smart Array controllers to deliver peak data bandwidth up to 600 megabytes per second (MB/s) per physical link in each direction. SAS devices are capable of sending and receiving data simultaneously across each physical link (full duplex mode). When running full duplex, 6-Gb/s SAS technology can deliver peak data bandwidth up to 1200 MB/s.

The SAS-2 specification is compatible with both Serial SCSI and Serial ATA protocols for communicating commands to SAS and SATA devices. SAS-2 compliant controllers are fully compatible with 1.5-Gb/s and 3.0-Gb/s SATA technology.

New generation HP Smart Array controllers

The Smart Array PCIe 2.0-based controllers are modular solutions with a common form factor, hardware, and firmware. The controllers are compatible with all ProLiant 300-series G6 servers. With the exception of the DL320 G6 and ML330 G6 servers, the ProLiant 300-series G6 servers incorporate embedded HP Smart Array P410i/Zero Memory Controllers (RAID 0/1/1+0). As a standard entry level RAID, the Smart Array 410i incorporates Zero Memory RAID (ZMR). Administrators can choose the cache size and whether to include the battery backed write cache (BBWC). These options allow users to upgrade from ZMR to 512 MB BBWC (this upgrade is included as a standard capability with some 300-series G6 servers).

Battery backed write cache

The BBWC is required for capacity expansion (adding one or more physical disks to an existing array). The controller recalculates parity and balances the data across all the disks. During the expansion, the BBWC preserves data and logical structures on the array. The HP 650 mAh P-Series battery extends battery life up to 48 hours before recharging is necessary.

Zero Memory RAID

Using Zero Memory RAID (ZMR), administrators can create a RAID 0-1 configuration without using any additional memory. Smart Array P410, P410i, P411, and P212 controllers include ZMR. The P212 controller does not include ZMR on the external connector. ZMR supports up to eight drives in Zero Memory Mode, or seven drives and one tape drive. Modular Smart Array (MSA) products are not supported in ZMR mode. ZMR does not include any caching. All systems can be upgraded to a BBWC memory module that can significantly increase performance.

NOTE:

The P212 controller can only be upgraded to 256 MB BBWC and does not support 512 MB BBWC.

NOTE:

Smart Array Advanced Pack is not available on Zero Memory configurations.

Smart Array Advanced Pack (SAAP)

HP Smart Array Advanced Pack (SAAP) is firmware that provides advanced functionality within the newest generation of Smart Array Controllers. These capabilities further enhance performance, reliability, and availability of data. SAAP is hosted on the Smart Array Controller hardware firmware stack. It can be enabled on the P212, P410, P410i, and P411 controllers.

SAAP requires a license key for activation. After activation, administrators can use several standard capabilities:

- RAID 6 with Advanced Data Guarding (ADG) protects against failure of any two drives. It requires a minimum of four drives, but only two would be available.
- RAID 60 allows administrators to split the RAID storage across multiple external boxes. It is supported with a minimum of eight drives, but only four would be available.
- Advanced Capacity Expansion allows users to automate the process of migrating old drives to a new array of drives.
- Mirror Splitting and Recombining in Offline Model breaks a RAID 1 configuration into two RAID 0 configurations. This is similar to a scaled down rollback functionality which requires two disk drives.
- Drive Erase erases physical disks or logical volumes. This capability is useful to decommission, redeploy, or return hard drives.
- Video On Demand Performance Optimization optimizes video on demand performance and improves latency while video streaming.

More information about SAAP is available at www.hp.com/go/SAAP.

Software RAID

HP has developed a software RAID solution based on the Smart Array firmware. The B110i SATA Software RAID supports the Array Configuration Utility (ACU), ACU-CLI (command line tool), Simple Network Management Protocol (SNMP) agents, and Web-Based Enterprise Management (WBEM) providers.

The B110i features an OS-specific driver from HP that uses the embedded ICH10R controller. It can utilize RAID 0, 1, and 1+0 and supports a maximum of two logical drives. The B110i supports up to four 1.5G or 3G SATA drives. Administrators can migrate drives to a Smart Array controller in a seamless procedure that maintains the user data and RAID configuration.

The ProLiant DL320 G6 server supports the B110i. For a listing of the complete feature set and support information for the B110i SATA Software RAID, download the B110i user guide at <http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01706551/c01706551.pdf>

Networking technologies

The NC522SFP and NC524SFP Dual Port 10GbE Modules are eight lane (x8) PCI Express (PCIe) 10 Gigabit network solutions. Both the NC522SFP and NC524SFP NICs include two SFP+ (small form-factor pluggable) connector cages for connectivity in copper or fiber optic environments.

While all ProLiant G6 servers support the NC522SFP, the HP NC524SFP is an exclusive upgrade module for ProLiant ML370 G6 and DL 370 G6 servers. The NC524SFP converts two of the 1GbE embedded ports to 10GbE ports. For complete specifications about HP network adapter products, go to www.hp.com/go/ProLiantNICs.

Power and thermal technologies

HP engineers have developed a robust set of power and thermal technologies, and components to manage power within the ProLiant 300-series G6 servers. The following technologies improve power efficiency throughout the power delivery chain:

- Efficient power delivery
- Improved thermal sensors and fan control
- Dynamic Power Capping
- Phase shedding
- Managing processor technologies
- Managing memory technologies
- Managing I/O technologies
- Power profiles

Administrators can disable certain components and capabilities within the 300-series G6 servers or reduce capabilities to bring the components to a lower power state.

Efficient power delivery

ProLiant G6 servers use common slot “right-size” power supplies and highly efficient DC power regulators to deliver significantly higher power efficiencies.

Common slot power supplies

The HP G6 common slot power strategy provides power supply commonality across ProLiant G6 server lines. HP reduced the number of power supply designs, which reduces the number of spares the customer must keep in the data center.

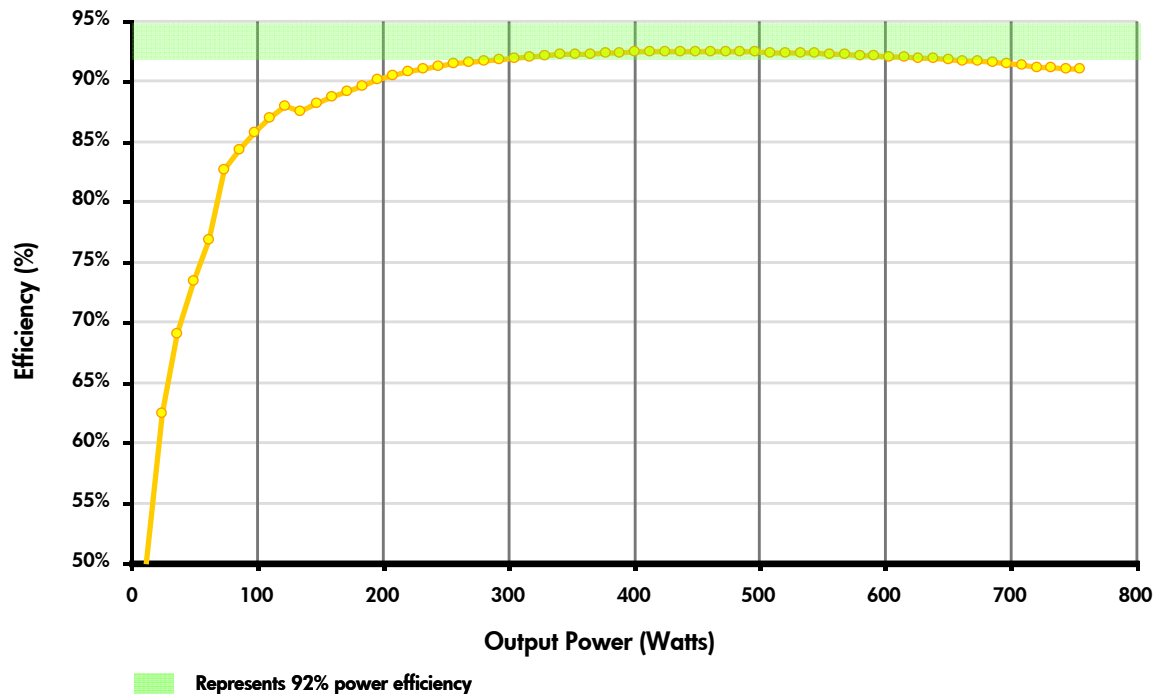
By incorporating “right-sizing” and High Efficiency mode (discussed in the following section), these power supplies have achieved efficiency ratings of up to 92 percent, and they meet Climate Savers Gold requirements.

- 460W AC up to 92% efficiency
- 750W AC up to 92% efficiency
- 1200W AC up to 90% efficiency
- 48Vdc 1200W up to 90% efficiency

The power loading efficiency curve for the 750 W power supply shown in Figure 5 illustrates the high levels of power efficiency achieved by power supplies used in the ProLiant G6 servers.

The HP Power Advisor is available at the following link to help users define the right power supply for their needs: <http://h71028.www7.hp.com/enterprise/cache/600307-0-0-0-121.html>.

Figure 5. Power/Efficiency curve for the 750 W HP power supply



Through the introduction of these different wattage power supplies, HP allows users to choose a power supply that meets, but does not exceed their needs. Oversized and lightly loaded power supplies do not run as efficiently as those that are highly loaded.

Increasing power efficiency in redundant power operation

With ProLiant G6 servers, customers can select a power supply operation mode for redundant power systems. This ROM-based setup provides the user with 'High Efficiency' mode and 'Balanced' mode selections. In Balanced mode, both power supplies provide power equally. This mode ensures full redundancy but can result in higher power consumption when power supplies are operating with reduced loads and lower power efficiency. The High Efficiency mode means the system will only use one power supply until system load exceeds a certain threshold. The second power supply stays online maintaining redundancy but does not supply power until needed. Either selection still provides full power redundancy.

Voltage regulation

Voltage regulators convert the 12V DC supplied from the server power supply into a variety of lower voltages used by the different system components. HP has developed voltage regulators that have higher peak efficiencies and maintain greater than 90 percent efficiency over a broad range of power requirements. The net result is about an eight percent gain in DC power efficiency, which results in almost a ten percent efficiency gain in AC input power.

These efficiency gains come with no loss in performance and require no configuration by the user.

Improved thermal sensors and fan control

The ProLiant 300-series G6 servers include many more embedded thermal sensors compared to the previous ProLiant server generation. Up to 64 thermal sensors —referred to as a "sea of sensors" — are located on DIMMs, hard drives, and throughout the server. The actual number of sensors varies by server platform.

The previous generation of ProLiant servers marked a shift away from processors as the primary producers of heat in the server. As memory modules become denser, they generate more heat. To combat this, DDR-3 DIMMs, as used in the ProLiant G6 servers, incorporate the first reliable on-DIMM thermal sensors

Because hard drive thermal sensors were not directly associated with fans, the fans would often operate at high speeds to prevent hard drives from overheating. ProLiant 300-series G6 servers incorporate hard drive temperature sensors into the body of data used to determine fan speed. This requires collaboration among various pieces of firmware, including the iLO 2 firmware, system firmware, and RAID storage controller firmware. The 300-series G6 servers have “zoned” fans that increase cooling and energy efficiencies in the server by adjusting cooling to those zones when called for by the sensors. This provides improved efficiency and better acoustics for the platform. The iLO 2 management processor in the G6 300 series uses a sophisticated control algorithm to set the speed for each fan zone in the system based on feedback from the appropriate temperature sensors. This allows fans to consume the minimum amount of required power.

The fan control algorithm lets ProLiant 300-series G6 servers change fan speed as the situation dictates. In ProLiant servers prior to G6, if one fan failed, all the other fans were set to high speed to assure the server remained within thermal specifications. ProLiant 300-series G6 servers now include enough sensors to construct an accurate view of the thermal landscape within the server, allowing the sensors and the fan control algorithm to determine if fan speeds need to be increased.

An RBSU option called “increased cooling” is included in the ProLiant 300-series G6 servers. The user can select this option and manually drive fans to higher speeds for use during situations requiring additional cooling.

Dynamic Power Capping

All ProLiant 300-series G6 servers incorporate Dynamic Power Capping, which allows administrators to cap power use and more accurately allocate the amount of power a server will use. Dynamic Power Capping can bring a server experiencing a sudden increase in workload back under its power cap in less than one-half second, preventing any surge in power demand that could cause a typical data center circuit breaker to trip. Dynamic Power Capping has been designed and tested (at 50 degrees C and 150 percent overload) to ensure that it can prevent tripping circuit breakers that have a specified trip time of three seconds or longer.

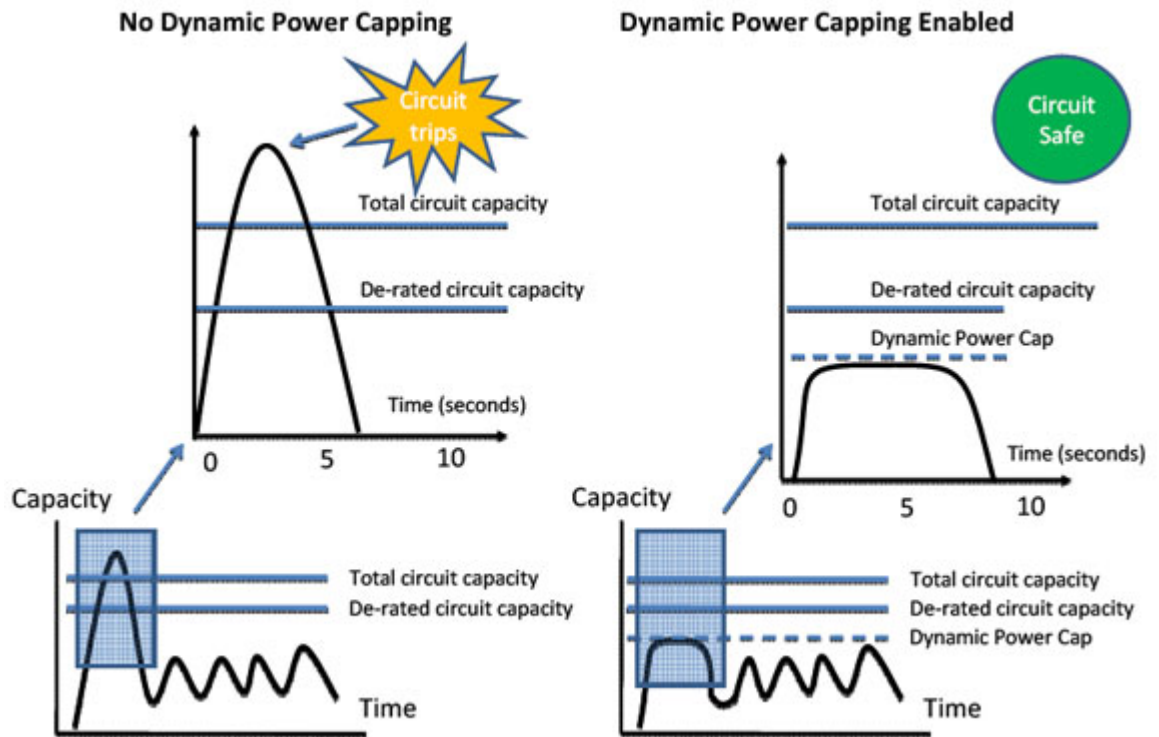
This ability to keep server power consumption below the power cap in real time means that Dynamic Power Capping can be used as an effective tool in planning and managing both electrical provisioning and cooling requirements in the data center. An administrator can electrically provision a PDU or a rack to less than the full faceplate power rating of all the servers loaded.

To implement Dynamic Power Capping, the iLO 2 management processor works in conjunction with a power microcontroller both to measure and to control power use. When enforcing the Dynamic Power Cap, the power microcontroller keeps the processor’s performance and power use under the set cap. This process is illustrated in Figure 6. Dynamic Power Caps for individual servers can be set from the iLO 2 Advanced user interface. Dynamic Power Caps for multiple rack-mount servers may be set from the power management module within HP Insight Control Environment.

HP Dynamic Power Capping is operating system independent and functions with all operating systems and software applications. HP Dynamic Power Capping will continue to function even if the software fails because it utilizes operating system independent hardware.

Since Dynamic Power Capping can impact server performance if set too aggressively, HP recommends that Dynamic Power Caps be set at values that match or exceed the highest observed power consumption over a representative server workload sample.

Figure 6. Rapid response of Dynamic Power Capping avoids circuit breaker trips



Phase shedding

HP incorporated intelligent phase shedding into ProLiant G6 voltage regulators. Modern digital voltage regulators deliver DC power using up to five different phases of high-speed power pulses that charge capacitors to deliver straight DC power to components at the proper voltage. Each phase delivers its power pulses in a rotating time window with the other phases in the voltage regulator such that the power pulses from one phase do not overlap with those of another. The width of each pulse determines the total power delivered by the particular phase.

Intelligent phase shedding lets the system BIOS turn off one or more of the power phases if it determines that the component power requirements are less than the full amount of power from the voltage regulator. This reduction in phases decreases the maximum power that the voltage regulator can deliver and increases overall efficiency.

Phase shedding based on the installed processor

The G6 server ROM configures the number of phases based on the maximum power consumption of the processor. This capability incurs no performance impact and requires no user configuration.

Memory phase shedding

Memory phase shedding operates much the same way as processor-based phase shedding. At power-up, the ROM BIOS determines the number of phases needed for the memory voltage regulator based on the number of DIMMs installed. Memory phase shedding can save up to 2.5 watts per DIMM socket. This feature is less effective on servers whose DIMM sockets are fully populated since more phases are required to accommodate such a configuration. More phases mean more power consumption.

Dynamic CPU phase shedding

On entry into a low power state (less than 20 W), the Intel Xeon 5500 Series processors will activate the Power Status Indicator (PSI). When PSI is engaged, ProLiant G6 servers turn off voltage regulator phases, thereby saving power and increasing power efficiency.

Managing processor technologies

QuickPath Interconnect power

The Xeon 5500 Series processor will allow the QuickPath Interconnect (QPI) buffers to enter a sleep state to reduce power requirements while the QPI links are not active. HP enables this Intel feature for G6 servers through RBSU. Once this feature is enabled, the Intel processor determines when to put the QPI buffers into a sleep state. It appears that QPI power management has no measureable impact on performance.

Disabling processor cores

Through RBSU, administrators can disable one or more cores in the Xeon 5500 Series processor (per physical processor). When enabled, the command will apply to all physical processors in the server. Engaging this capability saves power and has the potential to improve performance in servers running single workloads or applications with low requirements for threading.

Minimum processor idle power state

The Xeon 5500 Series processor supports C-states for each core within the processor. C-states define the power state of system processors and are an open specification of the ACPI group. The micro-architecture of the Xeon 5500 Series processor supports processor C-states C0, C1, C3, and C6. C-state C0 represents a fully active core that is executing instructions. The other C-states represent increasing levels of power reduction for idle cores. Any core within the processor can change C-states independently from the other cores.

Parameters for the maximum C-state allowed for an idle processor are set through the RBSU and initiated by the OS. The higher the C-state allowed at idle, the more power savings, but only at idle. Also, the higher the allowed C-state, the higher the latency involved when the core returns to activity.

Managing memory technologies

Memory channel interleaving

As described in the memory section, the alternate routing used for channel interleaving decreases memory access latency and increases performance. However, memory channel interleaving increases the probability that more DIMMs must be kept in an active state since the memory controller is alternating between channels and therefore between DIMMs.

Memory interleaving is configured in the RBSU. Disabling memory channel interleaving makes access to contiguous memory addresses revert to one channel. Single-channel access degrades performance, but makes it possible for the memory controller to place less frequently accessed DIMMs into a low power state which saves power. Depending on the application load of the server, memory interleaving can have a negative performance effect. Administrators should perform testing in their application environment to understand the trade-off between power savings and performance.

Maximum memory data rates

The maximum memory data rate is effectively 1333 MHz for ProLiant G6 Intel platforms.⁴ Depending on the memory configuration and installed processor, the system may automatically reduce the Quick Path Interconnect speed. While the "Auto" setting (which equates to 1333 MHz) is the default setting, users have the option to manually lower the effective data rates to 1066 MHz or 800 MHz. This will

⁴ The memory operates in a double-pumped manner so that the effective bandwidth is double the physical clock rate of the memory. Mega-transfers/second describes the data rate.

save power, but may incur some performance penalty. Administrators can configure the maximum memory data rate through the RBSU.

Managing I/O technologies

Disable PCIe 2.0

A ProLiant G6 option allows all expansion slots to run at PCIe 1.0 rather than PCIe 2.0 speed. Enabling this option saves power and provides backward compatibility with cards that may not correctly operate in PCIe 2.0 slots. Administrators can control expansion slot speed through the RBSU.

I/O Power Provisioning

I/O Power Provisioning allows the server to cut power to embedded devices that are not in use. This option is useful for customers who prefer their own storage controllers or NIC devices instead of embedded HP devices. This option saves up to 10W on the embedded storage controller, and between 2 and 6W on embedded NICs. Administrators configure I/O Power Provisioning in the RBSU.

Power profiles

In the ProLiant 300-series G6 RBSU, the HP Power Profile defines three possible configurations of some of the power features identified earlier in this section. The HP Power Profile provides a simple mechanism for users to configure the power management options of their system based on their tolerance to power versus performance without having to individually configure each option.

There are three possible settings for the HP Power Profile: Maximum Performance; Balanced Power and Performance; and Minimum Power Usage. An additional “Custom” setting is simply any combination of user settings that do not match the pre-sets for the three categories listed in Table 3.

Table 3. Power Profile settings

Power Saving Feature	Maximum Performance	Balanced Power & Performance	Minimum Power Usage
Power Regulator	Static High	Dynamic	Static low
Manage QPI power	Off	On	On
Memory Interleave	Full interleave	Full interleave	Disabled
PCIe 2.0	Enabled	Enabled	Off
Memory Speed	Auto	Auto	800 MHz
Minimum processor idle power	No C-states	C6	C6

Power Performance Benchmarks

The Standard Performance Evaluation Corporation (SPEC) is a non-profit corporation formed to establish, maintain, and endorse a standardized set of relevant benchmarks that can be applied to the newest generation of high-performance computers.

SPECpower_ssj2008 is the first industry-standard SPEC benchmark that evaluates the power and performance characteristics of volume server class computers. SPECpower benchmark results for the ProLiant G5 and G6 DL380 are available from the SPEC website. The test results show performance gains achieved over the last generation ProLiant servers.

Visit the SPEC website for ProLiant DL380 G6 test results:

http://www.spec.org/power_ssj2008/results/res2009q2/power_ssj2008-20090325-00136.html

Security

The Trusted Platform Module™ (TPM) and Microsoft® BitLocker® technology are supported in all ProLiant 300-series G6 servers by means of the available Trusted Platform Module option kit.

Trusted Platform Module

The Trusted Platform Module v1.2 supported on ProLiant G6 servers is a microcontroller chip that can create, securely store, and manage artifacts such as passwords, certificates, and encryption keys that are used to authenticate the server platform. The TPM 1.2 chip provides a unique Endorsement Key (EK) and a unique Storage Root Key (SRK). It provides data encryption and uses RSA, SHA-1, RNG cryptographic functions to provide access protection, OS level protection, and stolen disk protection.

The TPM 1.2 chip can also store platform measurements (hashes) to help ensure that the platform remains trustworthy. TPM enables Microsoft BitLocker, part of Windows® Server 2008.

TPM is an option on all ProLiant 300-series G6 servers. For more information about TPM, go to <http://www.hp.com/go/TPM>

BitLocker Drive Encryption

Microsoft BitLocker Drive Encryption (BitLocker) is a data protection feature available in Windows Server 2008. BitLocker uses the enhanced security capabilities of TPM version 1.2 to protect data and to ensure that a server running Windows Server 2008 has not been compromised while the system was offline.

Implementing BitLocker requires the following:

- The Master Boot Record (MBR), a small, encrypted system partition of approximately 50 MB to contain boot utilities
- TPM version 1.2
- Trusted Computing Group (TCG) compliant firmware including support of “Static Root of Trust”
- Two NTFS partitions on the boot drive

During the boot process, the TPM will not release the encryption key until completing a comparison of operating system configuration information (or hash) with an earlier snapshot of the same data. If any part of the hash is compromised (for example by introduction of malicious code), the TPM ensures that the volume encryption key is never released.

Systems management and monitoring

HP offers management tools to program and control all aspects of the dynamic server environment. The HP Insight Control (ICE) suite provides a foundation for deploying, managing, optimizing, and controlling the entire server environment from any location. HP Insight Dynamics – VSE suite for ProLiant delivers comprehensive functions for optimizing and balancing resources and workloads in real time. HP ProLiant Onboard Administrator powered by the iLO 2 management processor, provides remote management with other core embedded management functions to simplify setup, health monitoring, power and thermal control, and remote administration.

HP Insight Control Environment

HP Insight Control Environment (ICE) management suite is the infrastructure management suite that HP recommends for every HP G6 server. ICE provides infrastructure management functions, including a complete set of lifecycle management tools.

HP Insight Dynamics – VSE

HP Insight Dynamics – VSE suite for ProLiant servers assists users in continuously analyzing and optimizing adaptive server infrastructure. HP Insight Dynamics – VSE delivers four key capabilities for HP ProLiant 300-series G6 servers:

- Continuous capacity and power planning
- Balanced physical and virtual resources
- Cost-effective availability
- Consistent infrastructure provisioning

HP ProLiant Onboard Administrator powered by iLO 2

HP ProLiant Onboard Administrator powered by iLO 2 provides core management functionality including server power control, ROM-based setup, embedded health and management logs, power consumption level, SSH and SSL support, and flexible access through a dedicated or shared network port. With ProLiant G6 servers, HP introduces support for viewing configuration and status of CPU, fans, and power supplies. HP plans to introduce the ability to view configuration and status of memory, array controllers, and drives in mid-2009. The iLO 2 management processor is incorporated into all ProLiant 200-series and above servers.

ProLiant Onboard Administrator management capabilities can be extended with HP iLO Advanced. iLO Advanced capabilities include high performance graphical and text console, diagnostic capabilities (including automatic video footage of the server's last boot and last fault), global team collaboration for up to four users, 24-hour power measurement, and single server Dynamic Power Capping. HP Dynamic Power Capping, enabled through iLO Advanced, is available for all ProLiant ML and DL 300-series servers. For additional details, check the HP product support matrix:

<http://www.hp.com/servers/ilo/supportedservers>.

OS support

HP performs extensive testing, qualification, and certification of the latest server operating systems on ProLiant servers to ensure maximum performance and reliability. HP resells and provides full service and support for Microsoft® Windows® operating systems, Red Hat Linux subscriptions, Novell SUSE Linux subscriptions, Sun Solaris subscriptions, Citrix XenServer, and VMware hypervisors. The latest information regarding support and deployment can be found online at www.hp.com/go/ossupport.

Summary

The development and engineering efforts associated with the HP ProLiant 300-series G6 servers help users lower power costs and increase business performance with the new Xeon 5500 Series processor technologies, Dynamic Power Capping, efficient thermal control, common power supplies, DDR-3 memory, and PCIe 2.0-compliant modular Smart Array controllers. The 300-series G6 servers possess superior technology differentiation and management capabilities with ProLiant Onboard Administrator powered by iLO 2, HP Insight Control Environment management suite, and HP Insight Dynamics – VSE. Users also experience the ease of server deployment with SmartStart, Rapid Deployment Pack, and RBSU.

For more information

For additional information, refer to the resources listed below.

Resource description	Web address
DDR3 Memory Configuration Tool	http://h18004.www1.hp.com/products/servers/options/tool/hp_memtool.html
Dynamic Power Capping TCO and Best Practices White Paper	http://h71028.www7.hp.com/ERC/downloads/4AA2-3107ENW.pdf
HP Insight Control Environment	www.hp.com/go/ice
HP Network Adapters for ProLiant DL and ML Servers	http://media.hpvitc.veplatform.com/content/HP_Network_Adapters_for_ProLiant_DL_Family_data_sheet_1237839147.pdf
HP Insight Management WBEM Providers	www.hp.com/go/hpwbem
HP ProLiant DL320 G6 Server series	http://h18004.www1.hp.com/products/quickspecs/13344_na/13344_na.html
HP ProLiant ML350 G6 Server series	http://h18004.www1.hp.com/products/quickspecs/13241_na/13241_na.html
HP ProLiant DL360 G6 Server series	http://h18004.www1.hp.com/products/quickspecs/13235_na/13235_na.html
HP ProLiant ML/DL370 G6 Server series	http://h18004.www1.hp.com/products/quickspecs/13242_na/13242_na.html
HP ProLiant DL380 G6 Server series	http://h18004.www1.hp.com/products/quickspecs/13234_na/13234_na.html
HP Remote Management Lights-Out product information	www.hp.com/go/ilo
HP SAS and SATA technology	www.hp.com/go/serial
HP Smart Array Advance Pack	http://h18004.www1.hp.com/products/servers/proliantstorage/arraycontrollers/smartarray-advanced/index.html
HP Smart Array controllers	www.hp.com/products/smartarray
HP SmartStart	www.hp.com/servers/smartstart
ISS Technology Communications briefs: HP Power Capping and HP Dynamic Power Capping for ProLiant servers	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01549455/c01549455.pdf
ISS Technology Communications briefs: Memory technology evolution: an overview of system memory technologies	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00256987/c00256987.pdf
ISS Technology Communications briefs: Serial Attached SCSI storage technology brief	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01613420/c01613420.pdf

Call to action

Send comments about this paper to TechCom@HP.com.

© 2009 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Microsoft, Windows, and BitLocker are U.S. registered trademarks of Microsoft Corporation.

Linux is a U.S. registered trademark of Linux Torvalds

Intel, Xeon is a trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries and is used under license.

TC090910TB, September 2009

