HP ProLiant AMD-based 300–series G7 servers

Technology brief

Abstract

This technology brief describes the key technologies implemented in HP ProLiant 300-series G7 servers based on AMD™ processors. As of this writing, the AMD-based 300--series G7 server platforms are limited to the ProLiant DL385. For detailed information about this server, refer to the QuickSpec link listed at the end of this technology brief.

Introduction

The HP ProLiant DL385 G7 server includes these key technologies:

- AMD Opteron™ eight- and twelve-core 6100 Series processors
- Thermal sensors incorporated throughout the ProLiant 300-series G7 servers
- BIOS controlled memory and processor management capabilities
- Integrated Lights-Out 3 (iLO 3) remote server management and control
- I/O technologies such as PCI Express generation 2 (PCIe 2.0) and Smart Array controllers that incorporate common form factor components
- Flash-backed write cache for Smart Array controllers
- Common-slot power supplies in multiple sizes to provide the required amount of power and improve power efficiency

HP servers and balanced architecture

HP designs cost-competitive, power-efficient servers that use a balanced architecture to address performance requirements and provide value.

HP servers achieve a balanced architecture through superior engineering of fundamental elements such as mechanical infrastructure, power, cooling, processor, memory, IO devices, storage, boot, networking, and interconnecting components. A balanced architecture includes the following aspects:

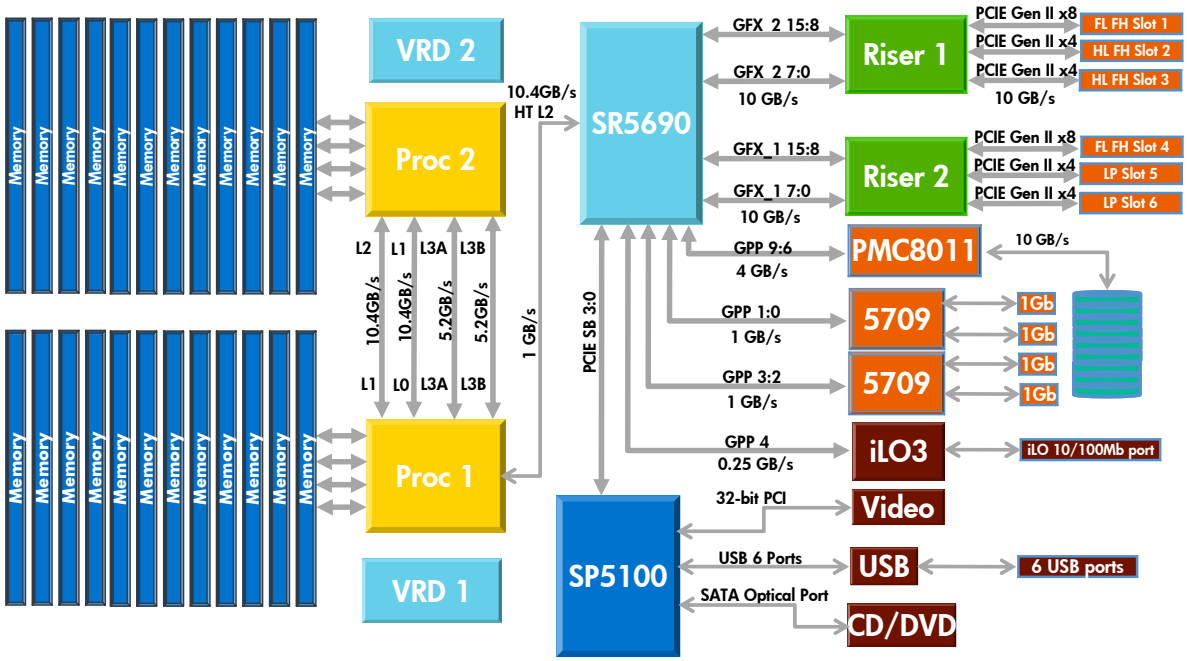
- Compute capability, processor core count, cache size per processor, and processor socket count
- Low-latency processor-to-memory bandwidth commensurate with core count
- Memory footprint and capacity that maximizes bandwidth and capacity with power efficiency and performance without compromising quality or reliability
- Application-appropriate IO devices
- Closely-coupled and balanced processor-to-memory and processor-to-I/O ratios
- Mechanical design that ensures optimum levels of cooling, stability, and serviceability through space-efficient, modular partitioning across the server

By designing a balanced architecture, HP ensures that all subsystems can be used effectively under a broad range of applications and workloads. For example, increasing memory capacity asymmetrically will not increase performance as effectively as distributing the same amount of memory across processors and IO devices. Inefficient memory distribution yields diminishing returns on power consumption and cost. A virtual machine (VM), for example, benefits from memory closely coupled to the processor responsible for that VM. Furthermore, a server needs to have appropriate levels of I/O bandwidth and CPU capabilities to ensure that memory can be used effectively by every VM.

Processor technology

The HP AMD-based 300-series G7 servers use the AMD Opteron 6100 series eight- and twelve-core processors allowing up to 24 cores in a two-processor platform. These processors are based on AMD's 45 nanometer process and have a core standard wattage of 75W Average CPU Power. In addition, these processors use Direct Connect™ architecture 2.0, HyperTransport™ 3.0, HT Assist™, and 8x ECC error correction. The 6100 processors fit into the G34 socket infrastructure and feature DDR3 memory, four 64 bit memory channels per processor, the AMD SR5690/SP5100 chipset, and quad HyperTransport (HT) links as shown in Figure 1. The processors operate at speeds of up to 2.4 GHz with eight cores and 2.3 GHz with twelve cores. They access 512 KB of L2 cache memory per core and share a 12 MB L3 cache. The AMD Opteron 6100 series processor includes performance-optimized power and thermal controls. See the Power and thermal technologies section for more information.

Figure 1. DL385 G7 server architecture using the AMD 6100 series processor and chipset

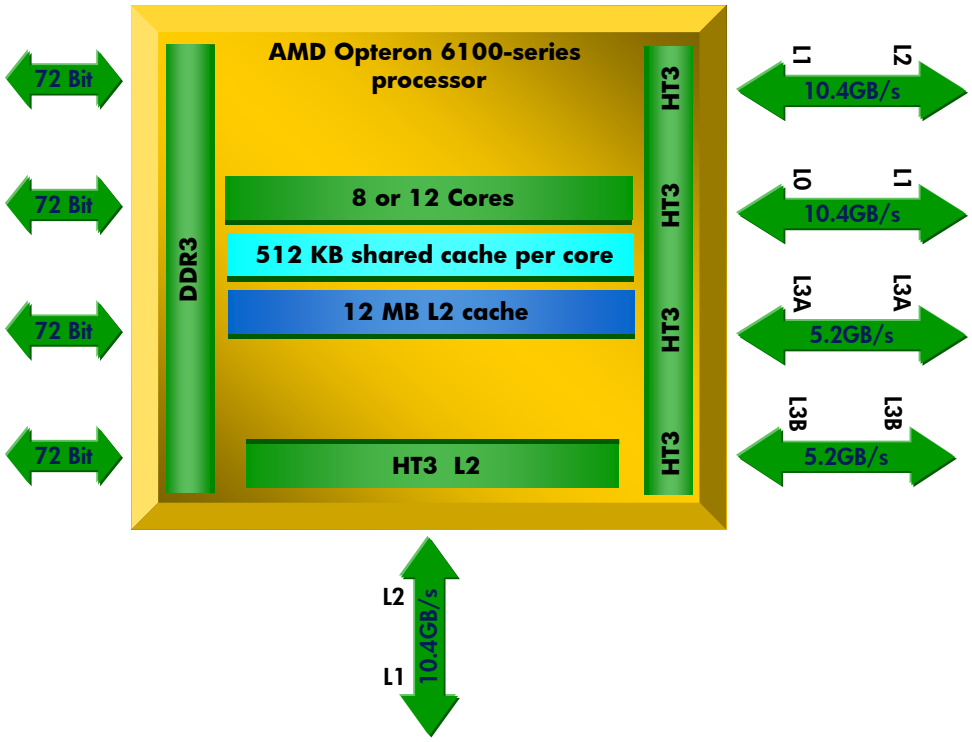


NOTE:
In this AMD Opteron 6100-series architecture, the SR5690 (sometimes referred to as the Northbridge chip) is responsible for I/O and PCI devices, while the SR5100 (sometimes referred to as the Southbridge chip) is responsible for USB, CD/DVD, and video functions

Direct Connect architecture 2.0

Unlike front-side bus architecture, Direct Connect integrates the memory controller into the processor and directly connects CPUs to the I/O subsystem and other processors (Figure 2). Direct Connect architecture uses direct HyperTransport links between CPUs, between CPU and I/O, and between CPU and memory. Direct Connect architecture currently scales up to 12 cores to provide superior memory and I/O capability, near native virtualization performance, and a range of power bands[1] that place a priority on low power consumption.

Figure 2. Block diagram of Direct Connect 2.0 architecture in the AMD 6100 series processors



HyperTransport technology

HyperTransport is a point-to-point interconnect that is designed to connect the processors directly and to connect each processor to its dedicated memory banks, as well as to other I/O chipsets.[2] Compared to a shared, parallel front-side bus, the advantages of HyperTransport include no overhead for bus arbitration and easier signal integrity maintenance, which results in a scalable, high-bandwidth architecture.

Each16-bit (2-byte) HyperTransport link is double-pumped to perform two data transfers per clock cycle. From HyperTransport 1.0 (HT1) in 2001 to HyperTransport 3.1 in 2008, the maximum clock

[1] Power bands refer to a new metric developed by AMD to reflect power consumed by the processor and its integrated memory controller during peak workloads. This metric is based on AMD's measurement of Average CPU Power (ACP). For more information on ACP, see the whitepaper at www.amd.com/us-en/assets/content_type/white_papers_and_tech_docs/43761C_ACP_WP.pdf

[2] HyperTransport Technology was invented at AMD with contributions from industry partners and is managed and licensed by the HyperTransport Technology Consortium, a Texas non-profit corporation.

speed and transfer rate increased from 800 MHz (1.6 MT/s[3]) to a maximum of 3.2 GHz (6.4 GT/s) in each direction.

The G34 socket infrastructure features four, 16-bit HT3 links operating at up to 6.4 GT/s per link. Of the four HT3 links, three are shared processor to processor links and one link on processor one is used for I/O communication.

HT Assist

HT Assist reduces cache coherence traffic on the HT links. By tracking where data is stored in cache and guiding the processor directly to the cache of other processors, HT Assist reduces cache probe traffic between processors, especially in 4- socket servers. Therefore, HT Assist results in faster queries that can increase performance for cache-sensitive applications such as database, virtualization, and compute-intensive applications.

Memory technologies

In the AMD Opteron 6100-series architecture, the memory controller is integrated into the processor chip to optimize memory performance and bandwidth per CPU. The memory controller reduces latency inherent in front side bus architectures by eliminating the bus contention between memory and I/O cycles. A server's overall memory bandwidth increases as processors are added to a configuration, unlike legacy designs that scale poorly because access to main memory is limited by external Northbridge chips. HP designed specific BIOS and RBSU functions to manage memory configurations, letting the customer optimize configurations for maximum performance while reducing power consumption and cooling requirements. These management options also include memory protection and latency reduction. Because of the increased reliability of DDR3 on-DIMM thermal sensors, HP incorporates DIMM thermal data into the algorithms controlling thermal and power states within the server.

DDR3 DIMM choices

HP ProLiant 300-series G7 servers with Opteron 6100 series processors support double data rate (DDR3) DIMMs. DDR3 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 DDR3, 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 1333 MT/s data rate. 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.

HP DDR3 DIMM modules incorporate an integrated thermal sensor that signals the processor to throttle memory traffic to the DIMM if its temperature exceeds a programmable critical trip point. Using the data from these thermal sensors, ProLiant AMD-based 300-series G7 servers can reduce fan speed when memory is idle, which reduces power consumption. The BIOS in ProLiant G7 servers verifies the presence of the thermal DIMM sensor during POST. Some third-party DIMMs may not include this thermal sensor. If it 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 speeds (requiring more power).

DDR3 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).

[3] MT/s, or megatransfers per second, equals the speed of the link in millions of cycles per second times the number of transfers per cycle.

There are three types of standard voltage DDR3 available for ProLiant 300 series G7 servers:

- PC3-8500R- DDR3 (RDIMM, ECC compliant)—1333 or 1066 MT/s data rate depending on memory configuration.
- PC3-10600R DDR3 (RDIMM, ECC compliant)—1333 or 1066 MT/s data rate depending on memory configuration.
- PC3-10600E DDR3 (UDIMM, ECC compliant)—1333 MT/s data rate

Low-voltage DDR3

Customers can take advantage of the HP low voltage (LV) DDR3 memory option. LV memory can operate at 1.35 V, reducing power and cooling requirements. Some LV DIMM configurations can affect performance. The difference occurs in 2 DIMM per channel (DPC) configurations with dual-rank DIMMs. In these configurations, the memory bus runs at 1.5 V with 2 DPC at 1333 MT/s, or 1.35 V with 2 DPC at 1066 MT/s. This results in a 20% reduction in bandwidth for the lower voltage. In all other HP LV configurations, the LV data rate is the standard 1.5V data rate. The BIOS determines the operating voltage and data rate capability from the DIMMs and the DIMM population from the system. The BIOS then sets the data rate based on that information.

NOTE:
Although the bandwidth reduction from DDR3-1333 to DDR3-1066 is 20%, the measured reduction in throughput is 10%

Single rank DIMMs configured at 1, 2, and 3 DPC; and dual rank DIMMs configured at 1 and 3 DPC run at the 1333 MT/s data rate at both voltages.

ProLiant G7 server models with the AMD 6100-Series processor support LV memory:

- PC3L-10600R—1333 1066 or 800 MT/s data rate, depending on memory configuration.

DIMM configuration guidelines

ProLiant AMD-based 300-series G7 servers support 3 DIMMs per channel up to 24 DDR3 memory DIMMs. DDR3 DIMM speeds will vary depending on number of DIMMs per channel. Consult the server QuickSpecs to determine DIMM speeds for given configurations. Administrators can configure Opteron-based ProLiant 300-series G7 servers using either RDIMMs or UDIMMs, but RDIMM and UDIMM memory cannot be mixed within a single server. Low voltage DIMMs can provide up to 10% DIMM power savings. Low and standard voltage DIMMs are compatible; systems automatically adjust voltage based on DIMMs installed.

NOTE:
DDR3 DIMM speeds will vary depending on number of DIMMs per channel. Consult the server QuickSpec to determine DIMM speeds for given configurations.

For detailed memory configuration rules and guidelines, please use the Online DDR3 Memory Configuration Tool at www.hp.com/go/ddr3memory-configurator.

Memory management technologies

Memory interleaving on the AMD Opteron 6100-series processors can occur between the processor memory banks, memory channels, and between processor nodes in a multiprocessor system. ProLiant

AMD-based 300-series G7 servers support all three memory interleaving technologies. Even though these technologies are independent of each other, they can operate simultaneously.

Memory bank interleaving

With memory bank interleaving engaged, data is routed alternately to memory banks through the common memory channel connecting the DIMM banks and the integrated memory controller. However, memory bank interleaving does increase the probability that more DIMMs need to be kept in an active state (requiring more power) since the memory controller alternates between memory banks and therefore between DIMMs.

The processor node memory controller automatically enables memory bank interleaving under the following conditions:

- Two single-rank DIMMs per channel result in two way-bank interleaving
- Two dual-rank DIMMs per channel result in four way-bank interleaving
- Two quad-rank DIMMs per channel results in eight-way bank interleaving
- Two dual-rank DIMMs and one quad-rank DIMM results in eight-way bank interleaving (in servers using three DIMMs per channel)

Memory Channel Interleaving

With memory channel interleaving, the processor memory controller routes data alternately through the two available memory channels. The result is that when the memory controller needs to access a block of logically contiguous memory, the requests are distributed more evenly across the two channels rather than potentially stacking up in the request queue of a single channel. This alternate routing decreases memory access latency and increases performance. As with memory bank interleaving, memory channel interleaving increases the probability that more DIMMs need to be kept in an active state.

Memory channel interleaving is always active on the AMD 6100-series processor.

Memory node interleaving

With node interleaving, memory can be interleaved across any subset of nodes in the multiprocessor system. Node interleaving breaks memory into 4 KB addressable entities. Addressing starts with address 0 on node 0 and assigns sequential addresses through address 4095 to node 0, addresses 4096 through 8191 to node 1, addresses 8192 through 12287 to node 2, and addresses 12888 through 16383 to node 3. Address 16384 is assigned to node 0, and the process continues until all memory has been assigned in this fashion. An application that uses a common allocation thread will benefit from node interleaving.

Memory node interleaving is disabled by default. Administrators can activate node interleaving using the RBSU. Node interleaving can only be configured if the memory footprint for both processors identical.

X8 error correction

In AMD Opteron 6100-series processors, the memory controller supports error correction circuitry (ECC) for both x4 and x8 DIMMs.

I/O technologies

ProLiant 300-series G7 servers incorporate PCI Express, Serial-Attached SCSI (SAS), and Serial ATA (SATA) I/O technologies. This server architecture lets administrators add PCI Express-compliant

expansion cards 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 G7 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. 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 1 shows the level of interoperability between PCIe cards and PCIe slots.

Table 1. 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 latest 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.[4]

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 — SAS-2 compliance lets administrators deploy and 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 bundles all four lanes together. SAS wide ports enhance performance by balancing SAS traffic across the links. 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 tolerance for link failures is possible because wide port connections are established from Phy[5] to Phy, and multiple, simultaneous connections to different destinations are possible. 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.

[4] 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 .

[5] The mechanism that contains a transceiver which electrically interfaces to a physical link. Phy is a common abbreviation for the physical layer of the OSI model.

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 speed [6] 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.

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 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.

For the most up-to-date listing of HP Smart Array controllers that support the SAS-2 specification, see the Smart Array controller matrix: www.hp.com/products/smartarray.

HP Smart Array controllers based on PCIe 2.0

The Smart Array PCIe 2.0-based controllers are modular solutions with a common form factor, hardware, and firmware. Any of the ProLiant 300-series G7 servers can use PCIe 2.0-based controllers. As a standard entry level RAID, HP designed the Smart Array 410i with a unique Zero Memory RAID (ZMR) capability. Administrators can choose the cache size and whether to include the battery backed write cache (BBWC) or the flash-backed write-cache (FBWC). These options allow users to upgrade from ZMR to 512 MB BBWC or 1GB FBWC.

Battery backed write cache

The BBWC system continues to be an option for capacity expansion (adding one or more physical disks to an existing array). The Smart Array 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 can power the cache for up to 48 hours before recharging becomes necessary.

Flash-backed write cache

HP introduced the flash-backed write-cache (FBWC) system in the fourth quarter of 2009. The FBWC uses NAND[7] flash devices to retain cache data and super-capacitors (Super-caps) instead of batteries to provide power during a power loss. The FBWC offers significant advantages over the HP Battery-backed write-cache (BBWC) system. Since the FBWC writes the contents of memory to flash devices, there is no longer a 48-hour battery life limitation and the data will be posted to the disk drive on the next power up.

---

[6] 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.

[7] Non-volatile semiconductor memory that can be electronically erased and reprogrammed. No power is needed to maintain data stored in the chip

The FBWC DDR2 mini-DIMM cache module is specifically designed for the present generation of PCIe2.0, SAS-based Smart Array controllers based on the PMC PM8011 SAS SRC 8x6G RAID on a chip (RoC). The primary FBWC components consist of the cache module, Super-caps with integrated charger, and RoC located on the system board.

At the time of this writing, the FBWC cache is supported on the Smart Array P410, P410i, P411, P212, P812, and P712m.

For more information on the flash-backed write cache, see the "HP Smart Array Controller technology brief" at
http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00687518/c00687518.pdf

Zero Memory RAID

Using Zero Memory RAID (ZMR), administrators can create a RAID 0-1 configuration without additional memory. Smart Array P410, 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. ZMR does not include caching; however, all systems can be upgraded to a BBWC or FBWC memory module that can significantly increase performance.

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

Smart Array Advanced Pack

HP Smart Array Advanced Pack (SAAP) firmware provides advanced functionality within Smart Array controllers. This firmware further enhances performance, reliability, and data availability. The Smart Array controller hardware firmware stack supports SAAP. 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 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 will be available for data. ADG can tolerate multiple simultaneous drive failures without downtime or data loss and is ideal for applications requiring large logical volumes, because it can safely protect a single volume of up to 56 disk drives. RAID ADG also offers lower implementation costs and greater usable capacity per U than RAID 1.
- RAID 60 allows administrators to split the RAID storage across multiple external boxes. It requires a minimum of eight drives, but only four will be available for data.
- Advanced Capacity Expansion (ACE) automates higher capacity migration using capacity transformation to remove logical drives by shrinking and then expanding them online. Standard drive migration and expansion remain unchanged.
- Mirror Splitting and Recombining in Offline Mode breaks a RAID 1 configuration into two RAID 0 configurations. This is similar to a scaled down rollback functionality that requires two disk drives.
- Drive Erase completely erases physical disks or logical volumes. This capability is useful when decommissioning, redeploying, or returning hard drives.
- Video On Demand Performance Optimization decreases latency and improves video streaming.

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

NOTE:
At a minimum, a 256 MB cache and battery kit is required to enable the SAAP license key. SAAP is not available on Zero Memory Configurations.

Solid state drives

HP has introduced the second generation of solid state drives (SSD) for ProLiant servers. These solid state drives are 3Gb/s SATA interface in both 60GB and 120GB capacities. The product, based on NAND Single Level Cell flash technology, are implemented as SFF and LFF hot plug devices on the HP universal drive carrier for general use across the ProLiant portfolio. These drives deliver higher performance, lower latency, and low power solutions when compared with traditional rotating media.

You can use the HP second generation SSDs with the present generation Smart Array controllers based on the PM8011 SRC MIPS processor on select ProLiant G6 and G7 servers. See the server QuickSpecs to confirm that SSDs are supported.

For more information on HP second generation SSDs, download the "Drive technology overview" technology brief at
http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01071496/c01071496.pdf

Power and thermal technologies

HP engineers have developed a robust set of power and thermal technologies and components to manage power within ProLiant 300-series G7 servers. The following technologies improve power efficiency throughout the power delivery chain:

- Efficient power delivery
- Thermal sensors and fan control
- Dynamic Power Capping
- Processor Management technologies

Administrators can disable certain components and capabilities within ProLiant 300-series G7 servers or reduce capabilities to bring the components to a lower power state.

Efficient power delivery

The ProLiant DL385 G7 server uses HP common slot power supplies. To see the power supply options offered with other HP servers, consult the server QuickSpecs or go to the HP ProLiant Server Compatibility Guide at www.hpproliantoptions.com/intro/.

Common Slot power supplies

The HP Common Slot power strategy provides power supply commonality across supported ProLiant G7 servers. Three different sized common slot power supplies are available so customers can choose the most effective power supply to match their power needs. "Right sizing" power supplies lets customers more closely match the power supply to the server power requirements in specific environments, significantly reducing wasted power. The HP Common Slot power strategy has also reduced the number of power supply designs, which in turn reduces the number of spares the customer must keep in the data center.

Power supply efficiency relates to the level of effective transfer and delivery of power through the power chain. Table 2 shows that HP power supplies have achieved efficiency ratings of up to 94%, meeting the Climate Savers Platinum requirements.

Table 2. HP power supply efficiency and Climate Savers rating

Power supply	Efficiency	Rating
460W AC	up to 92% efficiency	Climate Savers Gold
750W AC	up to 92% efficiency	Climate Savers Gold
1200W AC	up to 94% efficiency	Climate Savers Platinum
1200W 48VDC	up to 90% efficiency	Climate Savers Silver

Use the HP Power Advisor to help determine which power supplies will best meet your needs: www.hp.com/go/hppoweradvisor.

Redundant power operation

With ProLiant G7 servers, customers can select a power supply operation mode for redundant power systems. Through the RBSU, administrators can select "High Efficiency" mode or "Balanced" mode. 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. In High Efficiency mode, 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 uses voltage regulators that maintain greater than 90% efficiency over a broad range of power requirements. The net result is near an 8% gain in DC power efficiency, which results in almost 10% 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 AMD-based 300-series G7 servers include many more thermal sensors —referred to as a "sea of sensors" — these sensors are located on DIMMs, hard drives, and elsewhere 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, DDR3 DIMMs, as used in the ProLiant G7 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 G7 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 firmware, system firmware, and RAID storage controller firmware. The 300-series G7 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 management processor in the G7 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 G7 servers change fan speed as the situation dictates. In ProLiant AMD-based servers prior to G7, if one fan failed, all the other fans were set to high speed to assure the server remained within thermal specifications. ProLiant 300-series G7 servers

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.

HP Dynamic Power Capping

Because Dynamic Power Capping lets you keep server power consumption below a power threshold in real time, you can use it as a tool for planning and managing both electrical provisioning and cooling requirements in the data center. You can electrically provision a PDU or a rack to less than the full faceplate power rating of all the servers loaded and be assured it will not exceed the set limits.

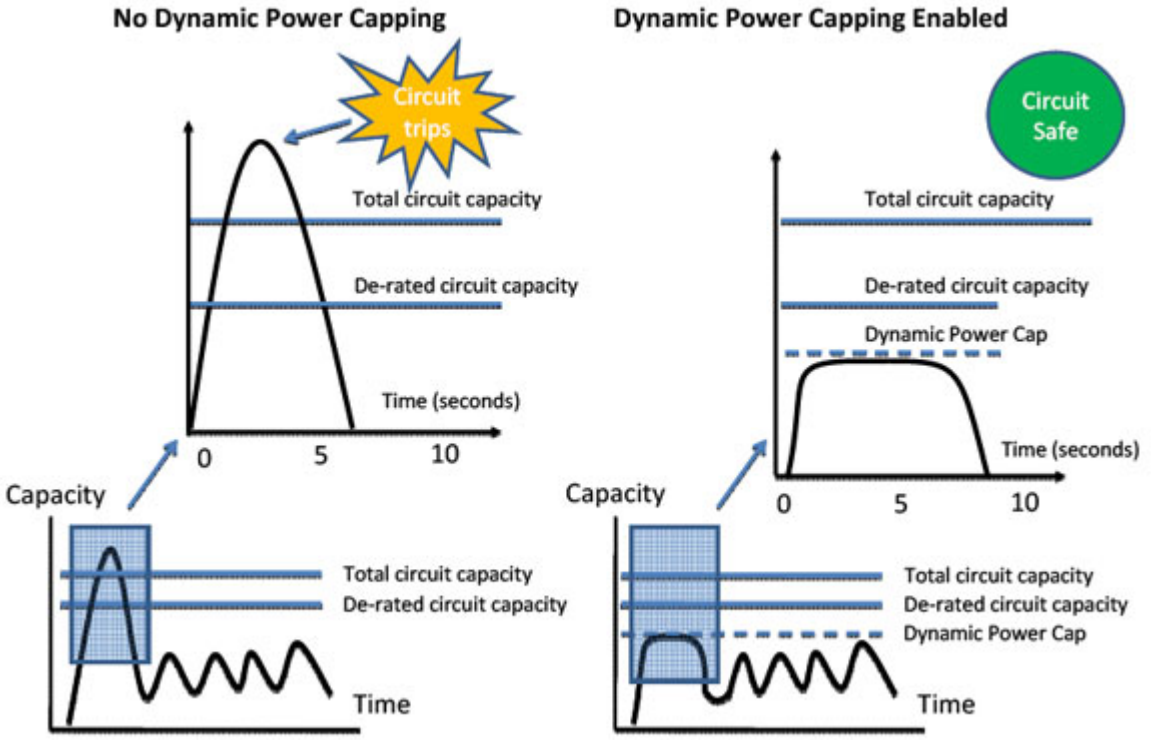
To implement Dynamic Power Capping, the iLO management processor works in conjunction with a power microcontroller both to measure and to control power use. When enforcing the Dynamic Power Cap, the power microcontroller keeps the processor's performance and power use under the set cap. This process is illustrated in Figure 4. You can set a Dynamic Power Cap for an individual server from the iLO Advanced user interface. For multiple rack-mount servers, you can set the Dynamic Power Caps 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 HP designed the hardware to be independent of the OS.

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.

For a more detailed explanation of HP Dynamic Power Capping, see the technology brief "HP Power Capping and HP Dynamic Power Capping for ProLiant servers" available at http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01549455/c01549455.pdf.

Figure 4. Rapid response of Dynamic Power Capping avoids circuit breaker trips



Managing processor technologies

HP leverages several of the power management technologies built into AMD Opteron 6100 series processor. This suite of technologies includes AMD PowerNow™ Technology, enhanced Advanced Platform Management Link (APML), and AMD Power Cap Manager™.

APML Remote Power Management Interface

The AMD Opteron six-core 6100-series processors incorporate multiple thermal sensors to indicate the hottest part of the processor. Using a systems management device such as iLO, you can remotely monitor and control P-state limits using AMD's APML Remote Power Management Interface. The iLO processor includes the processor thermal data in the data it sends to the fan controller.

Enhanced APML

Enhanced APML is a new capability offered on HP ProLiant G7 platforms with the AMD 6100 series processors. This capability includes a Precision Thermal Monitor feature that provides a more precise digital readout of CPU thermals, monitors power/cooling, and alerts the Baseboard Management Controller (BMC). HP ProLiant thermal policy relies on enhanced APML for accurate processor thermal information.

AMD CoolCore™

This AMD technology reduces processor energy consumption by turning off unused parts of the processor. This in turn helps reduce power and cooling costs for the IT infrastructure. AMD CoolCore is enabled by default and is not a user option.

Cool Speed ™

Cool Speed technology protects processor integrity by reducing power-states when the processor reaches an established temperature limit. This technology allows a server to continue operating if the processor thermal environment exceeds safe operational limits.

It is important for managers to understand that Cool Speed is an automated processor feature. Within a rack environment, some server platforms may initiate this process while other servers do not. Given the same workload, Cool Speed activation will depend on the installed processor model and local environmental conditions. AMD Cool Speed is enabled by default and is not a user option.

C1E ™

In G34 based-systems, the Northbridge chipset detects when all of the processor cores are idle and communicates that command to the Southbridge chipset. After the command is sent, the Northbridge and HT links are powered down and the cores go into a deeper sleep state. Depending on system configuration, this feature can equate to significant data center power savings when the Northbridge and HT links are powered down and cores are at idle. The OS manages C1E and it is enabled or disabled through the RBSU.

PowerNow! Technology

AMD PowerNow! Technology with Independent Dynamic Core technology and Dual Dynamic Power Management™ is technology that allows the processors to run dynamically at different frequencies and voltages depending upon CPU computing demand. As a result, PowerNow! can lower server power consumption without compromising performance.

AMD PowerNow! can be enabled on ProLiant AMD-based 300-series G7 servers through the BIOS-controlled Dynamic Mode of Power Regulator for ProLiant, which does not require an OS driver.

Power Cap Manager

Power Cap Manager allows administrators to set a fixed limit on a server's processor power consumption. In a multi-core environment, administrators control the P-state of individual cores. The caveat is that a single voltage is supplied to all cores in the processor even if different P-states are requested by the user. If an administrator requests different P-states for different cores in the processor, the actual P-state will equal that of the highest voltage required among the selected P-states. Administrators can control this through HP Dynamic Power Capping, accessed from iLO advanced menus or the HP ICE management suite.

AMD Core Select

AMD Core Select lets administrators use the RBSU to select the number of software-visible cores per CPU (minimum of one, up to full number supported). Operating systems and applications can recognize the reduced core count so that fewer software licenses may be required. By reducing core count recognition, Core Select has the potential to reduce software licensing costs. Also, applications can benefit from increased memory bandwidth and cache per core. This feature is offered on HP ProLiant G7 platforms with the AMD 6100 series processors. The feature is also supported on ProLiant G6 servers that are configured with the AMD 2400 or 8400 series processors and the latest System ROM supporting this feature.

For more information on these processor management technologies, please consult the HP ROM-Based Setup Utility User Guide for more details on these options.
http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00191707/c00191707.pdf

Systems management and monitoring

HP offers management tools to program and control all aspects of the dynamic server environment. The HP iLO 3 management processor provides remote management with other core-embedded management functions to simplify setup, health monitoring, power and thermal control, and remote administration. The HP Insight Control Environment management suite provides a foundation for deploying, managing, optimizing, and controlling the entire server environment from any location. HP Insight Dynamics for ProLiant delivers comprehensive functions for optimizing and balancing resources and workloads in real time.

HP Integrated Lights-Out 3

ProLiant 300-series G7 servers include the next generation of lights-out management, iLO 3. HP developed iLO 3 to improve performance, streamline the user interface, and enhance support for management standards. Like iLO 2, the iLO3 firmware can be upgraded with HP iLO Advanced for extended functionality.

The iLO 3 hardware includes a faster processor along with increased flash and DDR memory for storing and executing programs. Remote console performance is up to eight times faster than iLO2—and equal to the performance of KVM and software-based remote management solutions. The remote console for Microsoft Windows in iLO 3 is built on the .NET framework using a DirectX control; the console display scales without scroll bars, supports resolutions up to 1600X1200, and supports display across multiple monitors. Users have direct access to the remote console without having to navigate through the iLO 3 web interface. An integrated Linux version provides a single java applet to support remote console, virtual media and virtual power functionality. The faster iLO 3 processor and a USB 2.0 host connection also enable three times better virtual media performance compared to iLO 2.

The web-based graphical user interface has been updated with a look and feel similar to the Onboard Administrator; it uses Javascript Object Notation (JSON) for improved handling of dynamic content.

For environments with elevated security requirements, the iLO 3 ASIC includes an AES encryption engine. AES (Advanced Encryption Standard) is an encryption standard adopted by the U.S. government (FIPS PUBS 197, http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf). The iLO 3 hardware-based solution provides strong AES encryption with better performance than a software-based solution. The latest generation of iLO also provides IPMI over LAN and DCMI management interfaces for environments requiring simple standards-based monitoring and control.

Not all iLO 2 features are supported in iLO 3. For detailed information on standard and iLO Advanced licensed features, see the iLO QuickSpecs available at: www.hp.com/go/iLO.

HP Insight Control Environment

HP Insight Control Environment management suite is the infrastructure management suite that HP recommends for every HP G6 and G7 server. Insight Control Environment provides infrastructure management functions, including a complete set of lifecycle management tools.

For more information about HP Insight Control Environment management suite, go to www.hp.com/go/ice

HP Insight Dynamics

HP provides tools like HP Insight Dynamics for ProLiant to assist you with continuously analyzing and optimizing your data center infrastructure. HP Insight Dynamics delivers four key capabilities for HP ProLiant 300-series G7 servers:

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

For more information about HP Insight Dynamics, go to www.hp.com/go/insightdynamics

Security

The Trusted Platform Module™ (TPM) and Microsoft® BitLocker® technology are supported in all ProLiant 300-series G7 servers by means of the Trusted Platform Module option kit. The Trusted Platform Module v1.2 supported on ProLiant G7 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 100-series G6 and G7 servers. For more information about TPM, go to www.hp.com/go/TPM

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.

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, Citrix XenServer, and VMware hypervisors. The latest information regarding support and deployment can be found online at www.hp.com/go/ossupport.

Summary

The HP ProLiant 300-series G7 servers equipped with AMD processors help administrators increase business performance, lower power costs, and manage their server hardware more easily. To improve performance, these servers use the AMD Opteron 6100-series processor technologies with integrated memory controllers and DDR3 memory. The latest Smart Array controllers and firmware improve RAID performance compared to the previous generation of controllers. Using HP common slot power supplies is another means for customers to refine and constrain server power based on their data center requirements. iLO 3, HP Insight Control Environment management suite, and HP Insight Dynamics all facilitate server management. These servers can be easily deployed with SmartStart, Insight Control server deployment, and RBSU.

For more information

For additional information, refer to the resources listed below.

Resource description	Web address
Dynamic Power Capping TCO and Best Practices White Paper	www2.hp.com/v2/GetPDF.aspx/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 ProLiant DL385 G7 Server QuickSpecs	http://h18004.www1.hp.com/products/quickspecs/13594_na/13594_na.html
HP ROM-Based Setup Utility User Guide	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00191707/c00191707.pdf
HP iLO 3 product information	www.hp.com/go/ilo
HP SAS and SATA technology	www.hp.com/go/serial
HP Smart Array Advanced Pack	http://h18004.www1.hp.com/products/servers/proliantstorage/arraycontrollers/smartarray-advanced/index.html
HP Smart Array controllers	www.hp.com/products/smartarray
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"	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01613420/c01613420.pdf

Call to action

Send comments about this paper to TechCom@HP.com

 Follow us on Twitter: http://www.twitter.com/HPISSTechComm

TC100407TB, June 2010