HP ProLiant DL585 G2 server technology

technology brief, 4th edition

Abstract

The second generation (G2) HP ProLiant DL585 server offers the performance and efficiency of AMD Opteron® processors, enhanced by improvements to all major subsystems in the server. This paper explains the technologies implemented in the HP ProLiant DL585 G2 server that make it an excellent choice for enterprise customers needing increased performance and investment protection with a 64-bit migration path. This paper is intended for IT professionals familiar with system administration and HP ProLiant industry-standard servers. It discusses only key technologies of the ProLiant 500-series servers. For complete specifications of each server, see the HP website at www.hp.com/go/ProLiant.

Introduction

For the ProLiant DL585 G2 server, HP improved upon the first-generation design by integrating higher performance components and enhancing all major subsystems in the server. The DL585 G2 provides excellent system performance by using powerful processors, on-chip memory controllers, a large memory footprint, and fully redesigned I/O and disk subsystems. The server is well suited for many computing environments:

- High-performance technical computing (HPTC)
- Electronic design automation (EDA)/semiconductor
- Financial applications
- Petrochemical applications
- Enterprise Resource Planning (ERP)/Customer Relationship Management (CRM) applications
- Large database applications
- Life science and material science applications
- Video rendering applications

Operating system support

The DL585 G2 supports the following operating systems and virtualization software:

- Microsoft® Windows® Server 2003 Standard, Enterprise, or R2 Edition (32- and 64-bit)
- Microsoft Windows Server 2003 Datacenter Edition
- Microsoft Windows Compute Cluster Server 2003
- Microsoft Windows Server "Longhorn"
- Red Hat Enterprise Linux 4 or Linux 5
- SuSE Linux Enterprise Server 9 or Server 10
- VMWare ESX virtualization software
- Microsoft Virtual Server
- Xen virtualization software[1]

For a complete and up-to-date listing of supported OSs and versions, please visit the ProLiant OS Support Matrix at http://www.hp.com/go/supportos

[1] Xen virtualization software is an open source project that allows multiple virtual machines to share the physical hardware of a single host machine. Some virtualization products run virtual machines as guests on top of a host operating system (which requires a large amount of overhead). By contrast, Xen operates as a thin layer directly on top of the underlying hardware. Xen coordinates the interaction of the virtual machines and the physical hardware. HP is a key participant in the development of Xen. For more information about the Xen project, see the xensource website at www.xensource.com

DL585 G2 architecture

The DL585 G2 is a powerful, 4U enterprise server incorporating technologies that extend the capabilities of industry-standard x86 computing. This high-performance server is capable of simultaneously running both 32-bit and 64-bit applications with no performance degradation when using an operating system that supports 64-bit extensions. This is possible because the AMD64 instruction set is a super-set containing the x86 instruction set architecture.

The DL585 G2 architecture includes these essential features:

- AMD Opteron 8200 series dual-core or 8300-series quad-core processors
- NVIDIA nForce Professional 2200 and 2050 chipsets, and the AMD 8132 chipset
- Two 100MHz PCI-X slots (one half-length, one full-length)
- Four PCI Express x4 slots (one half-length)
- Three PCI Express x8 slots
- Two embedded multifunction gigabit network adapters with TCP/IP offload engine, including support for accelerated iSCSI
- HP Integrated Lights Out 2 (iLO 2) management

Quad-core processor support

A quad-core processor includes four separate microprocessor cores on one physical die. Quad-core technology significantly increases performance. When the number of processors available to the operating system is increased, multiple threads can be executed more efficiently. The on-die memory controller gives quad-core processors optimum performance in virtual machine (VM) environments. Moreover, quad-core processors operate within similar power and thermal envelopes as dual-core processors.

The ProLiant DL585 G2 offers investment protection since a dual-core system is upgradeable to a quad-cord system by means of a quad-core processor option kit. Complete server replacement is not required to migrate to quad-core technology.

64-bit architecture

A 64-bit architecture has a much larger amount of directly addressable (flat) memory space than a 32-bit processor. Using the AMD64 instruction set, the OS can access a flat memory address space greater than 4 GB without enabling Physical Address Extensions (PAE) and incurring the overhead of PAE. Because they can use large amounts of memory, 64-bit architectures can have performance advantages for intensive floating-point calculations used in scientific and engineering modeling programs. For additional information about AMD 64-bit technology, see the technology brief[2] titled "The AMD Processor Roadmap for Industry-Standard Servers."

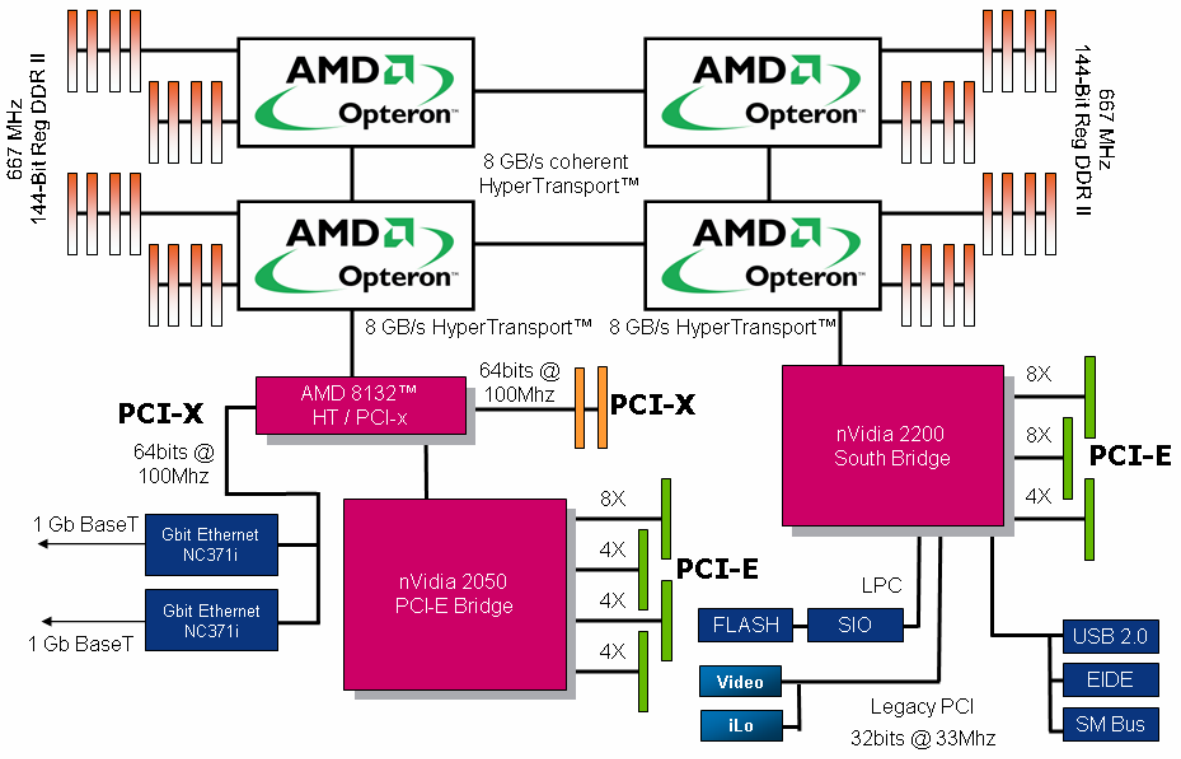
AMD Opteron processors and the AMD-8132 chipset

As an alternative to traditional front-side bus technology, AMD Opteron processors use Direct Connect Architecture and HyperTransport® technology. HyperTransport is a parallel, point-to-point interconnect that replaces parallel front-side bus technology. Direct Connect architecture is AMD's designation for the coherent HyperTransport connection between processors. It eliminates the bottlenecks inherent in front-side bus technology by integrating the memory controller into the processor and directly connecting CPUs to the I/O subsystem and other processors.

[2] Available on the HP website at http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00428708/c00428708.pdf

The AMD-8132 HyperTransport PCI-X Tunnel provides two PCI-X bridges (see Figure 1). One 16-bit, 1-GHz HyperTransport interface connects to the CPUs, while the other 16-bit, 1-GHz HyperTransport interface connects to the downstream NVIDIA nForce® Professional 2050 PCI-E Bridge. The AMD-8132 supports two PCI-X bridges operating at 100 MHz. One PCI-X bridge supports two 100-MHz expansion slots (one full length and the other half length). The other PCI-X bridge supports two NC371i gigabit Ethernet interfaces.

Figure 1. Block diagram of DL585 G2 architecture



NVIDIA nForce Professional 2200 and 2050 chipsets

The DL585 G2 makes use of both the NVIDIA nForce Professional 2200 and 2050 chipsets to offer a large number of expansion slots for enhanced flexibility. The 2050 chipset connects to the AMD-8132 HyperTransport tunnel with a 16-bit, 1-GHz HyperTransport interface. The 2050 offers 20 lanes of PCI Express, which are divided in the DL585 G2 into three x4 slots and one x8 slot. The 2200 connects directly to the processors with an additional 16-bit, 1-GHz HyperTransport connection. In addition to supporting system devices and USB ports, the 2200 offers an additional PCI Express x4 slot and two PCI Express x8 slots.

Memory technologies

Typical multiprocessor PC server architecture connects IA-32 processors to memory DIMMs by means of a north bridge chip. The north bridge is a memory controller and bridge to the I/O expansion interface. AMD Opteron architecture differs from typical multiprocessor server architecture in that the memory controller is integrated into the processor to boost performance. Performance is enhanced by eliminating bus contention created when memory and I/O paths pass through a typical, non-integrated north bridge. Because the memory controller is integrated onto the processor chip, memory latency is greatly reduced.

In the DL585 G2, processors access the memory controllers at core speed. Each on-chip memory controller directly accesses the DIMMs on the same processor/memory board at the DIMM speed. The aggregate bandwidth for system-accessible memory increases with the number of processors.

The Opteron chipset supports dual-channel memory, which reduces memory latency by increasing the bus width from 64 bits to 128 bits. Each memory controller has two 64-bit-wide memory channels, and the channels operate in parallel to support the 128-bit interface. Because the DL585 G2 has dual-width memory channels, DIMMs must be installed in pairs.

On-die memory controllers result in faster memory accesses by each processor. The memory controller on each processor has direct access to the memory attached directly to that controller. However, for a process thread running on one processor to access memory attached to another processor, extra coordination is required.

To provide optimum performance for a wide variety of applications, the DL585 G2 can support either of two methods of organizing memory access: linear, non-uniform memory access (NUMA) or node interleaving, sufficiently uniform memory accessing (SUMA).

A node consists of one or more processors, its embedded memory controller, and the attached DIMMs. The total memory attached to all the processors is divided into 4-KB segments. In the case of linear addressing (NUMA), consecutive 4-KB segments are on the same node. In the case of node interleaving (SUMA), consecutive 4-KB segments are on different or adjacent nodes.

Linear memory accessing (NUMA) defines the memory on all nodes sequentially. It assigns sequential addresses to all memory locations on node 0, then to all the memory locations on node 1, and so on until memory locations on all nodes have been assigned.

Node interleaving (SUMA) breaks memory into 4-KB addressable entities. Addressing starts with address 0 on node 0. Sequential addresses through address 4095 are assigned 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.

To take advantage of NUMA architecture, the operating system and the applications that run on the system must assign memory on a per-thread basis. In general, a NUMA-aware operating system such as Microsoft Windows and a NUMA-aware application such as Microsoft SQL Server will benefit from the NUMA organization. A NUMA-aware operating system and applications that allocate and de-allocate memory at the thread level will benefit from the NUMA organization because the allocation and the thread will have a tendency to run on the same node. If an application uses a common allocation thread, it will benefit from node interleaving.

The DL585 G2 has, by default, NUMA memory configuration. For those applications that cannot take advantage of the NUMA architecture, performance may be improved by activating node interleaving (SUMA). System administrators can activate node interleaving using the HP ROM-Based Setup Utility (RBSU) that is provided as part of the HP ProLiant Essentials Foundation Pack.

The DL585 G2 supports PC2-5300 DDR2 memory at 533 MHz and 667 MHz. When more than four DIMMS are installed on a single memory controller, the memory bus will be clocked down to 533 MHz. Table 1 lists available memory configurations for the DL585 G2.

Table 1. Memory configuration options in the ProLiant DL585 G2 server

Memory type	Maximum capacity	DIMM size	Maximum DIMMs per processor	Memory speed
PC2-5300	128 GB	512 MB 1 GB 2 GB 4 GB	8	533 MHz
PC2-5300	64 GB	512 MB 1 GB 2 GB 4 GB	4	667 MHz

I/O technologies

The DL585 G2 supports two types of I/O expansion slots: PCI Express and PCI-X. PCI Express (also known as PCI-E or PCIe) is a bus technology that builds upon earlier PCI technology, but it uses a serial interface rather than the traditional parallel PCI bus. PCI Express connections are composed of one or more lanes. Devices that use multiple lanes use a total of 2, 4, 8, 12, 16, or 32 lanes. Multiple lane connections are described as x4 for a 4-lane connection, x8 for an 8-lane connection, and so forth.

Each lane is a separate serial connection comprising four wires and has a bandwidth of 250 MB/s in each direction (for a total bandwidth of 500 MB/s per single PCIe lane link). PCI Express links automatically negotiate the highest number of lanes that the slot and the card support. Therefore, a PCI Express expansion card can be installed into any slot in which it will physically fit and it will work correctly. Also, a long slot wired with fewer lanes than it can physically support can work correctly.

The DL585 G2 has seven PCI Express slots. Three of the PCI Express slots are full length x8 slots. The other four PCI Express slots are x4: Three of the four are full length slots and the other one is half length. The server also has two 64-bit, 100-MHz, PCI-X slots (one full-length and one half-length).

When installing PCI Express expansion cards, optimal performance can be achieved by balancing cards across HyperTransport links. Table 2 identifies the preferred order of installation.

Table 2. Preferred order for populating PCI Express expansion slots in the ProLiant DL585 G2 server

Installation sequence	Slot	Operational support
First	7	x8
Second	5	x8
Third	8	x8
Fourth	3	x4
Fifth	4	x4
Sixth	9	x4
Seventh	6	x4

Storage technologies

The DL585 G2 has an updated disk subsystem and includes the Smart Array SAS P400 Controller.

Serial Attached SCSI (SAS)

The peak data transfer rate for the current generation of SAS drives is 3 gigabits per second in full duplex mode. In a serial system, the devices do not share a common bus, and each device can take full advantage of the bandwidth available to it. The SAS backplane connector is compatible with SATA drives, and the chipset on a SAS controller can communicate with SATA drives. This enables administrators to store data that needs to be available at a high performance standard on SAS disk drives, and to use the same enclosure to store data on slower but higher capacity SATA drives.

Note, however, that it is not possible to use a SAS drive on a SATA controller.

For more information about SAS technology, refer to the HP technology brief titled "Serial Attached SCSI technology."[3]

Small Form Factor (SFF) hard drives

SFF drives offer several advantages over 3.5-inch drives. The smaller physical size of the drives increases the number of gigabytes per U that can be implemented in a server rack. SFF drives have been shown to be more reliable than their larger counterparts, largely due to the use of smaller parts and better vibration control.

Using SFF drives reduces power consumption and heat generation. SFF SAS drives consume approximately half of the power used by a 3.5-inch drive of comparable capacity and therefore run cooler than 3.5-inch drives.

Using SFF drives also improves performance. The smaller platter size results in reduced seek times because the heads have a shorter distance to travel. RAID performance naturally benefits from having more spindles.

[3] Available at http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00302340/c00302340.pdf

Smart Array SAS P400 Controller

The DL585 G2 includes a Smart Array P400 Controller which supports SAS storage. Some configurations of the DL585 G2 include a battery-backed write cache (BBWC) on the P400 controller. For models that do not include the BBWC, it is available as an option. The battery will last two days without receiving any power from the computer.

The BBWC buffers disk writes so that disk I/O can be handled efficiently. The battery is needed in case of an unexpected system shutdown so that the information in the buffer will not be lost. In the event of a complete system failure, the controller and disks can be moved to a different server, where the controller will flush out the cache to the disks when power is restored. In the case of a controller failure, the cache module and disks can be removed from the failed controller and installed on a working controller, where the cache will be flushed out to the disks.

The Smart Array P400 supports RAID levels 0, 1, 1+0 and 5. RAID 6 (double parity) is available and does require the BBWC. Mirror splitting is available for RAID 1 arrays. This functionality allows the user to split a RAID 1 mirror into two separate RAID 0 arrays (breaking the mirror).

The BBWC is required for capacity expansion functionality, which allows the user to add a physical disk to an existing array. The controller then recalculates parity and balances the data across the disks. During the expansion, data and logical structures on the array are preserved.

The P400 controller supports a recovery ROM for failing back to the previous valid ROM if an attempt to flash the controller's ROM should fail. Online Drive Flashing is also supported. With Online Drive Flashing, disk drive firmware updates can be pre-loaded onto the controller. At the next reboot, the controller will flash the firmware on the hard drives.

Networking technologies

The DL585 G2 includes an embedded, dual-port, NC371i multifunction, gigabit Ethernet network interface controller (NIC). Technologies enabled by the NC371i include TCP Offload Engine (TOE), including support for Accelerated iSCSI through an optional licensing kit.

TCP Offload Engine

Network bandwidth has improved steadily since the early days of TCP/IP networking. Along with this improvement in speed has come increased the demand for CPU cycles to manage the network protocol stack. A busy, full-duplex gigabit Ethernet connection can consume the available computing power of a 1-GHz Pentium 4 processor. Unfortunately, this means that even a modern, high-powered CPU will show degraded performance in processing application instructions while data is being transferred onto or off of the network. Computers most susceptible to this problem are application, web, and file servers that have a high number of concurrent connections.

To reduce this burden on the CPU, the embedded NC371i network controller in the DL585 G2 is designed with TOE capabilities. TOE NICs are designed with on-board logic to process common and repetitive tasks of TCP network traffic. Because the CPU does not have to devote cycles to processing these tasks, it can be used more efficiently to significantly increase application performance on servers attached to gigabit Ethernet networks.

TOE is supported on Microsoft Windows Server 2003 with the Scalable Networking Pack installed.

iSCSI

iSCSI (internet SCSI) is a standard that implements the SCSI protocol for interacting with storage devices over a TCP/IP network. While iSCSI can be implemented over any TCP/IP network, the most common implementation is over gigabit Ethernet. iSCSI serves the same purpose as Fibre Channel in building Storage Area Networks (SANs), but without the cost, complexity, and compatibility issues associated with Fibre Channel SANs.

In iSCSI terminology, initiators are devices that access storage resources through the iSCSI protocol. A target is the device that the initiator accesses. While the target is usually a hard drive enclosure or another computer, it can also be any other storage device that supports the iSCSI protocol, such as a tape drive.

Initiators can be divided into two categories: software initiators and Host Bus Adapters (HBAs). Software initiators implement the iSCSI protocol in software, and CPU resources are used to manage the protocol stack. A more efficient approach is to offload the management of the protocol to the network adapter. Adapters that implement the iSCSI protocol are known as iSCSI HBAs. These devices appear to the operating system to be a SCSI HBA. The NC371i is an iSCSI HBA and appears to the operating system as a SCSI HBA.

Management technologies

HP ProLiant servers are shipped with management technologies that include Integrated Lights-Out 2 and the ProLiant Essentials Foundation Pack.

Integrated Lights-Out 2

Integrated Lights-Out 2 (iLO 2) is the fourth generation of Lights-Out remote management for HP servers. iLO 2 is hardware and firmware integrated into most ProLiant servers. It provides remote management capabilities over Ethernet. iLO 2 is most useful in situations that require remote server management. It is active even when the OS is not operating. The iLO 2 management processor obtains its power from the auxiliary power plane of the server, so it is always available when the server is plugged into a power source.

There are three levels of licensing for iLO 2: iLO 2 Standard, iLO 2 Select Packs, and iLO 2 Advanced Packs. Each offers different levels of remote access capabilities.

The DL585 G2 can connect to iLO 2 through a dedicated Ethernet port. This port can connect to a dedicated management network that is out of the server's data path and that can be highly secured. On DL585 G2 servers, iLO 2 can assign an additional MAC and IP address to one of the standard embedded network interfaces and share that port with normal traffic. This shared arrangement eliminates the need for an extra network connection for iLO 2.

For more information about iLO 2, refer to the HP technology brief titled "Integrated Lights-Out technology: enhancing the manageability of ProLiant servers."[4]

ProLiant Essentials Foundation Pack software

The ProLiant Essentials Foundation Pack is a full suite of available management software products that support ProLiant servers. Details about these software products are available on the HP website at www.hp.com/servers/proliantessentials.

[4] Available at http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00257345/c00257345.pdf

Mechanical technologies

The DL585 G2 leverages the chassis design of the ProLiant DL580 G3 server. The 4U form factor supports efficient use of space with maximum deployment flexibility.

The DL585 G2 ships with tool-free, snap-in, sliding rails and a cable management system for simple deployment in HP and third-party racks and for in-rack server access. The rail system is ambidextrous; that is, the same rail design can be used on either side of the rack with minor adjustments to the rail system. The rail system provides tool-free support for square-hole, round-hole, and threaded-hole racks. It adjusts to rack depths of 23.75 to 35 inches (60.3 to 88.9 centimeters) and therefore accommodates a much broader range of rack designs than previous rail systems. The cable management arms are also ambidextrous and independent of the hole design in the racks. The cable management arms allow customers to put all the cables down one side of the rack or to switch sides from one server to the next.

Inside the DL585 G2 chassis is a System Insight Display (SID) to provide specific trouble-shooting information. The display includes LEDs for all major subsystems of the server: memory, processors, redundant fans, power supply and Processor Power Modules (PPMs), and overtemp indicator. During normal operations, all LEDs are unlit. An amber LED gives an instant visual indication of a fault condition.

The modular and virtually cable-free interior design of the server provides easy, tool-free access to all of the major subsystems and components. The DL 585 G2 uses plug-in components to eliminate many of the cables. One of the most important aspects of the design is that all of the memory modules, processors, and disk drives are easily accessible from the front of the unit. In addition, the G2 design adds a front accessible video port and two front USB 2.0 ports for accessing the server from the front of the rack.

Conclusion

The HP ProLiant DL585 G2 is a 4U rack-optimized, 4-way server created for large data center deployments requiring enterprise-class performance, uptime, and scalability, plus ease of management and expansion. It offers customers running both 32- and 64-bit applications increased performance and memory speed. Quad-core processor support and improved network and I/O subsystems ensure that the DL585 G2 is scalable in high-performance computing environments.

For more information

For additional information, refer to the resources listed below.

Type of information	Source
Explanation of how HyperTransport technology works and the advantages it offers for system design and performance	AMD white paper *HyperTransport™ Technology: Simplifying System Design*, October 2002 www.hypertransport.org/docs/26635A_HT_System_Design.pdf
Up-to-date information on operating systems and versions supported by the HP ProLiant DL585 G2 server	http://www.hp.com/go/supportos
Latest drivers available for the HP ProLiant DL585 G2 server	www.hp.com/support/files
Details about HP Systems Insight Manager	www.hp.com/go/hpsim
Details about iLO management functionality	Integrated Lights-Out technology: enhancing the manageability of ProLiant servers
Details about HP ProLiant Essentials	www.hp.com/servers/proliantessentials
Information about SAS technology	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00302340/c00302340.pdf

Call to action

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

TC080104TB, January 2008