ProLiant ML530 High-Performance technologies

technology brief

Abstract

The ProLiant ML530 Generation 2 (G2) server features new technologies that improve on the performance, scalability, fault tolerance, and manageability of the first generation ProLiant ML530 server. A discussion of all these technology improvements is beyond the scope of this document. This paper focuses on the synergy of the server's high-performance technologies that provide the balanced system architecture of this mid-range departmental server.

Introduction

The ProLiant ML530 G2 server (Figure 1) is a high-performance 2-way server with optimized system resources for intensive data center and remote office environments. The server is designed with a balanced system architecture (2.8-GHz Intel® Xeon™ processors, Double Data Rate (DDR) SDRAM, and PCI-X technology) to maximize application performance and user workload. The system architecture is balanced by an enterprise-class chipset (the ServerWorks Grand Champion HE) that supports up to 16 gigabytes (GB) of memory and seven 64-bit, 100-MHz PCI-X slots.

The performance and scalability of the ProLiant ML530 G2 server make it a flexible solution for applications such as:

- Server consolidation
- Remote site or branch office server
- High-performance, low-cost database engine
- Mail and messaging
- Dedicated application server

Figure 1. ProLiant ML530 dual-processor server

First, this paper describes the overall system architecture of the ProLiant ML530 G2 server. Then it describes the high performance features of the individual processor, memory, and input/output (I/O) subsystems in more detail.

System architecture

Figure 2 illustrates the balanced system architecture of the ProLiant ML530 G2 server. At the heart of the server architecture is the enterprise-class ServerWorks GC HE chipset, which controls 3.2 GB/s of data transfer between the processor, memory, and input/output (I/O) subsystems. The processor subsystem contains up to two 2.8-GHz Intel Xeon Processors with 512-KB L2 cache and new features

2

such as Intel NetBurst™ microarchitecture and Hyper-Threading technology. The memory subsystem features

200-MHz DDR SDRAM with 2-way memory interleaving that doubles the performance of the PC133 SDRAM used in the first generation of the server. The I/O subsystem features a quad-peer PCI-X architecture that boosts I/O peak bandwidth to four times that of conventional PCI. The following sections describe the performance features of the three major subsystems in more detail.

**Figure 2.** ProLiant ML530 system architecture
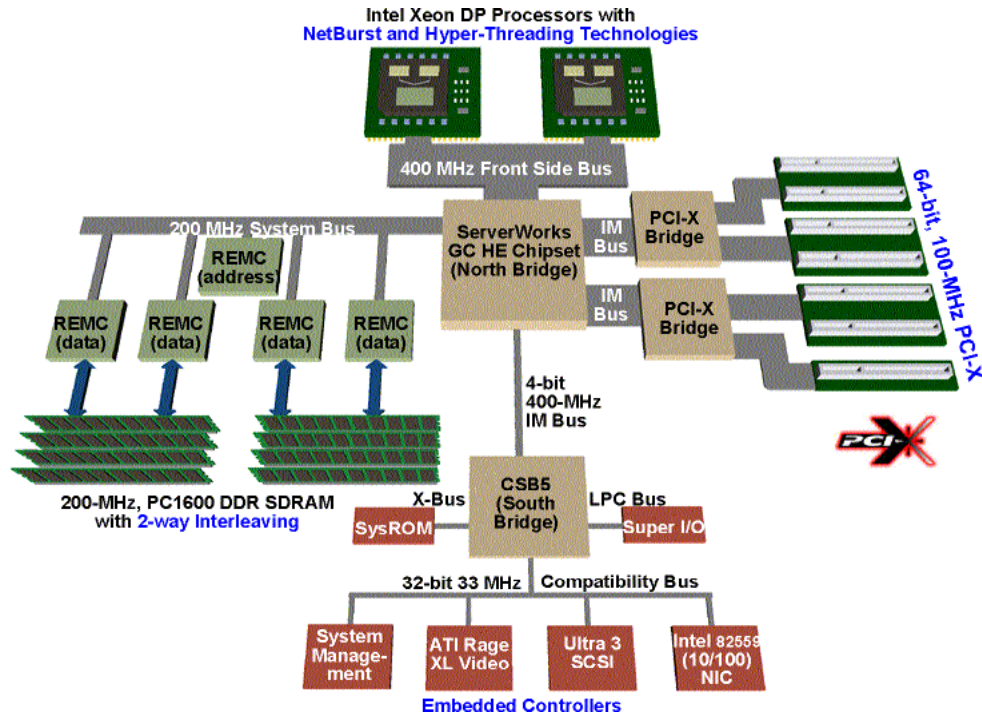
Processor subsystem

In the ProLiant ML530 G2 server, the 2.8-GHz Intel Xeon (Prestonia) processor replaces the 1-GHz Pentium® III Xeon processor that was used in the first generation of the server. Tower and rack models of the ProLiant ML530 G2 server come with one or two 2.8-GHz Intel Xeon processors and a 400-MHz front side bus (FSB). The higher core frequency is made possible with the Intel NetBurst microarchitecture, which doubles the pipeline depth in the processor.

Other new processor features include:

- Rapid execution engine — The two integer Arithmetic Logic Units (ALUs) in the processor run at twice the core frequency, which increases performance by allowing many integer instructions to execute in one half of the internal core clock period.
- Execution trace cache — Reduced decoder latency speeds up instruction throughput, which improves response times.

Figure 3. ProLiant ML530 G2 processor subsystem architecture
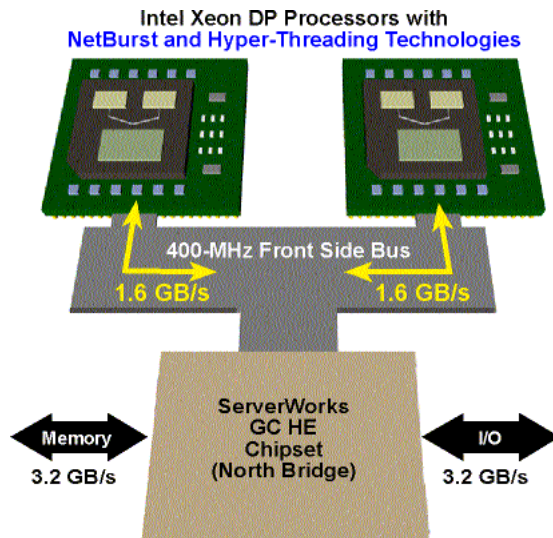
Smaller feature size

The Intel Xeon Processor is built with a 130-nanometer (0.13-micron) process to allow higher frequencies and better performance. The manufacturing term 0.13 micron refers to the circuit (feature) size. Feature size is a major limiting factor in processing speed. The smaller the feature size, the more transistors are packed into the circuit. As the feature size decreases, the processing speed increases and the power requirements decrease. The 0.13-micron Xeon processor has a smaller feature size and faster circuitry than the 0.18-micron Intel Foster processor.
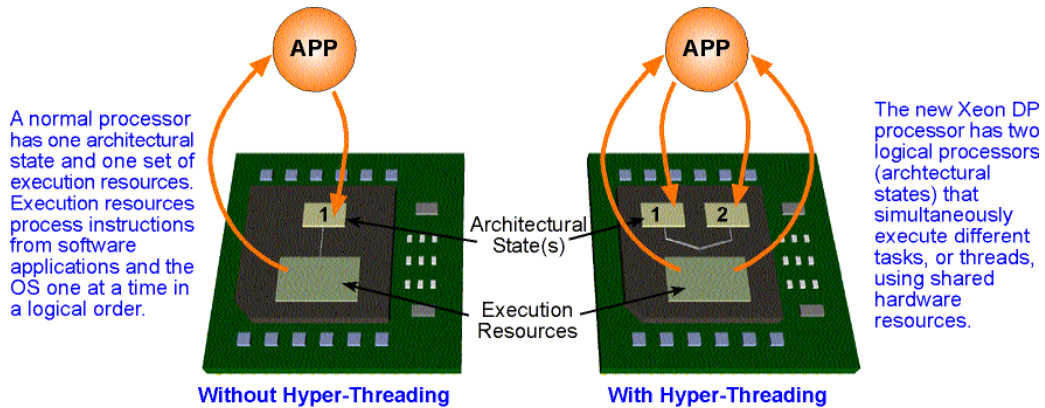
Hyper-Threading (Jackson) technology

Hyper-Threading technology lets a single processor execute two applications or processes at one time by handling instructions in parallel.

A processor without Hyper-Threading technology has one architectural state and one set of execution resources on the processor core (see Figure 4 left). The architectural state is a set of registers that track program execution, and it is viewed by the operating system (OS) as one logical processor. The execution resources process instructions from the OS and applications one at a time in a logical order. During each clock cycle, a typical operation uses only a fraction of the execution resources while the rest are idle. Hyper-Threading technology addresses this low processor utilization by using as many execution resources as possible during each clock cycle.

The OS views a processor with Hyper-Threading technology as if it were two logical processors—two architecture states sharing one set of execution resources. This allows the processor to simultaneously execute incoming instructions from different software applications by using out-of-order instruction scheduling to keep execution resources as busy as possible. As a result, a processor with Hyper-Threading technology can execute as many instructions as 1.5 processors. The result is a performance boost during multi-threading and multi-tasking operations. The actual performance increase depends on the independent operations being executed and the execution resources required to complete the operation.

Figure 4. Hyper-Threading technology

A normal processor has one architectural state and one set of execution resources. Execution resources process instructions from software applications and the OS one at a time in a logical order.

The new Xeon DP processor has two logical processors (archtectural states) that simultaneously execute different tasks, or threads, using shared hardware resources.

Architectural State(s)

Execution Resources

Without Hyper-Threading

With Hyper-Threading

In a multiprocessing system, the OS manages the tasks performed by all processors in the system. To take advantage of multiple processors, applications must be multi-threaded, which means they must be designed to be split into multiple streams of instructions, or threads. The OS can allocate various software threads to run on more than one processor simultaneously, which results in improved performance. But first, the OS needs to know the number of available processors so it can distribute the optimum number of threads among the processors.

The system BIOS counts the number of processors so the OS can create the optimum number of software threads for better load balancing. A table in the system BIOS records the number of processors and tags each one as a physical or logical processor. Figure 5 illustrates the counting order. The system BIOS counts the first logical processor on each physical processor. Then, in the same sequence, the system BIOS counts the second logical processor on each physical processor. This ensures that the OS uses separate physical processors as often as possible to maximize performance.

Figure 5. The system BIOS counts processors

Logical Processors (4)

Architectural States

Execution Resources

Physical Processors (2)
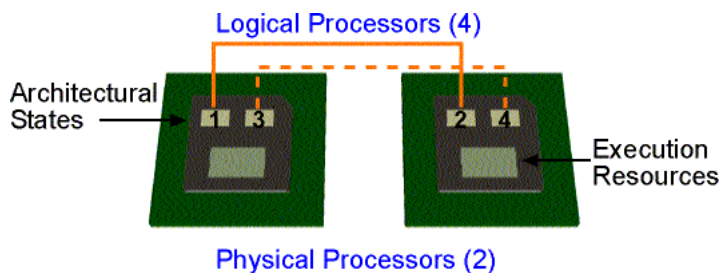
T The counting of physical and logical processors can also be used to determine per-processor license compliance. Using the example in Figure 5, the system with two processors would exceed the license limit for a two-processor OS if the OS cannot differentiate between physical and logical processors. For example, Microsoft Windows 2000 Server products counts the logical processors, so it will not use subsequent logical processors once it reaches the license limit. On the other hand, Windows Server 2003 products count the physical processors and use all their logical processors. For example, Windows Server 2003 Standard Edition has a two-processor licensing limit. However, in a 2P system using Xeon processors with Hyper-Threading technology, Windows Server 2003 can get the benefit of four logical processors. The table that records the processors in the BIOS allows Windows Server 2003 to resolve logical processors to their associated physical processors.

In summary, OSs that support Hyper-Threading include:

- Microsoft Windows 2000 Server (counts logical processors)
- Microsoft Server 2003 (uses all logical processors, regardless of physical count)
- Sun Solaris 8

These OSs will support Hyper-Threading, but they will need drivers:

- NetWare v 5.0
- NetWare v 5.1
- NetWare v 6.0
- NetWare v 6.5

OSs that will not support Hyper-Threading include any Linux distribution.

OSs aware of Hyper-Threading schedule application threads to run on logical processors in the same way they manage physical processors. With Hyper-Threading technology, OSs schedule threads not only to separate processors, but also to separate logical processors on a single physical processor.

Because of the way the processors are counted, and subsequently identified by the OS, threads are always scheduled to logical processors on different physical processors before multiple threads are scheduled to the same physical processor. This optimization allows software threads to use different physical execution resources when possible.

The second logical processor can also be turned off when it is not needed. A HALT instruction is issued to the inactive logical processor. Without this instruction, an OS may execute on the idle logical processor a sequence of instructions that repeatedly checks for work to perform. This so-called "idle loop" can consume significant execution resources that could otherwise be used by the active logical processor.

Note:
Hyper-Threading can be turned off in the ROM-Based Setup Utility (RBSU). This may be necessary for testing or verifying performance gains for enterprise applications. Also, it is possible that some applications not designed for Hyper-Threading may not perform as well with Hyper-Threading turned on.

Level 2 advanced transfer cache

The principle behind caching is based on the probability that a processor will need information it has recently accessed in system memory more often than a random piece of information it has not accessed. Just as a carpenter uses a tool belt, the processor uses the cache to hold the most recently used information closer for faster and more efficient operation.

Typically, there are two levels of cache memory: primary Level 1 (L1) cache and secondary Level 2 (L2) cache. The L1 cache resides within the processor core and holds 8 kB of recently accessed data. The L2 cache stores recently accessed data that is not held in the L1 cache. When the processor needs data, it first looks in the L1 cache. If the information is found in the L1 cache (known as a cache hit), the processor uses it without a performance delay. If the information is not in the L1 cache, the processor searches the 512-kB data store in the L2 cache. The data store is organized in columns and rows. Each row, or cache line, contains 64 bytes (512 bits) of data. To optimize performance, data is written to or read from the L2 cache as a complete 512-kB cache line. The

512-bit cache line size in the Intel Xeon Processor is twice the size of the cache line in the Pentium III processor. As a result, there is greater chance of a cache hit for any given memory request.

When a cache hit occurs in the L2 cache, the data is transferred at 2.8 GHz to the processor core along a 32-byte interface on each core clock cycle. As a result, the
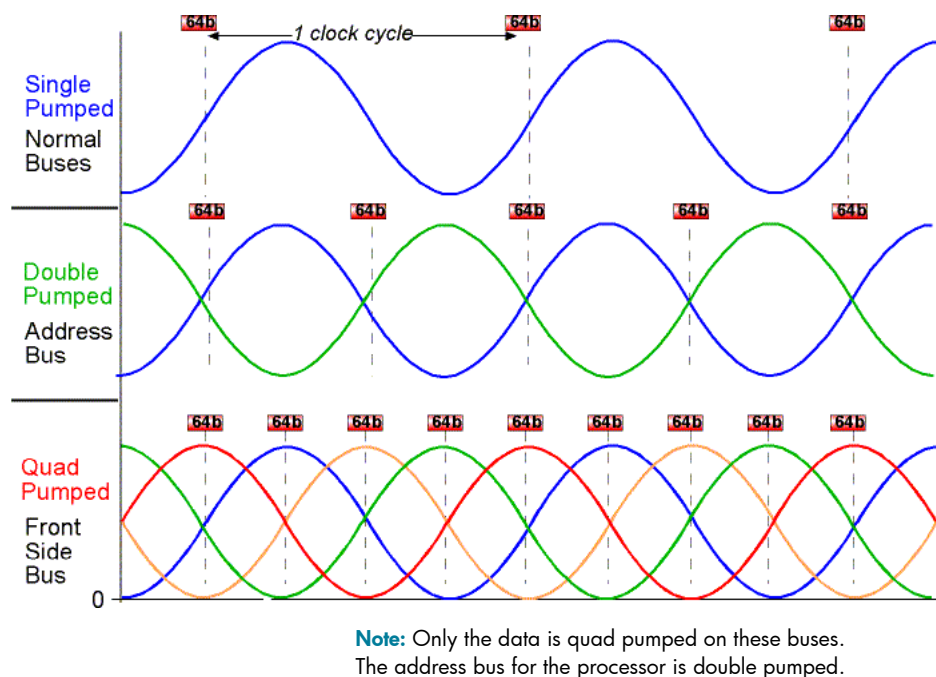
512-kB L2 Advanced Transfer Cache can deliver a data transfer rate of 89.6 GB/s to the processor so that it can keep executing instructions instead of sitting idle. This compares to a transfer rate of 16 GB/s for the 1-GHz Intel® Pentium® III processor.

If the requested information is not in L1 or L2 cache, the processor must issue a request to read it from the system memory.

400-MHz front side bus

All data transfers go to and from the processor over the FSB. The Intel Xeon processor's FSB is a 64-bit, quad-pumped bus running at 100 MHz. A normal (single-pumped) bus sends, or latches, data out once per clock cycle on the rising or falling edge of the bus clock signal. A quad-pumped bus latches data at four times the rate of a normal bus (Figure 6). This is accomplished with four overlapping clock strobes, each operating 90 degrees out of phase with the next. Data is sent on the rising edge of each of the four strobes, four times per clock cycle. This makes it possible to transfer 3.2 GB/s of data on a 100-MHz FSB, which is triple the data rate of the Pentium III FSB (1.06 GB/s with a 133-MHz FSB).

Figure 6. Comparison of clock signals for a quad-pumped and single pumped 100-MHz front side bus

Note: Only the data is quad pumped on these buses. The address bus for the processor is double pumped.

Memory subsystem

The memory subsystem of the ProLiant ML530 G2 server is designed for high performance using PC1600 DDR SDRAM, which has an effective data rate of 1.6 GB/s. Combined with two-way interleaving (described below), the memory subsystem provides the bandwidth necessary to keep up with the 3.2-GB/s data transfer rate to and from the processor (Figure 7). This balanced configuration reduces latency of data transfers between memory and processors, further enhancing system performance.

The ProLiant ML530 G2 server comes standard with a single memory board. The memory board has eight Dual Inline Memory Modules (DIMM) sockets for a total capacity of 16 GB, if 2-GB DIMMS are used in Standard Memory mode.[1] The sockets are organized into four banks (A, B, C, and D) with two sockets in each bank (Figure 7). The memory board contains five Reliability-Enhanced Memory Controllers (REMCs). One REMC is dedicated to addressing. It identifies the specific location of the data in memory. The other four REMCs control the data transfers to and from the DIMMs. They serve as the bridge between the DDR memory bus and the system bus.

**Figure 7.** Architecture of the ProLiant ML530 memory subsystem
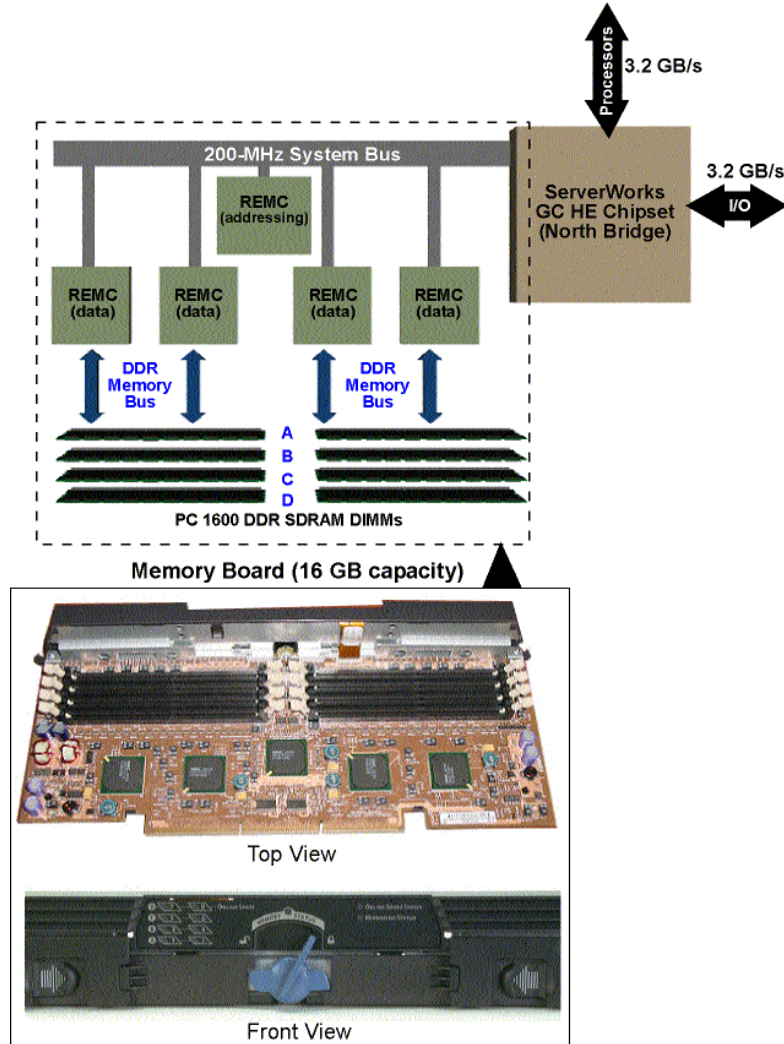
---

Standard configuration

The server comes standard with two 512-MB, DDR SDRAM DIMMs in bank A (Figure 8) for a total of 1 GB of system memory. Because the system uses 2-way interleaving, the DIMMs must be installed in pairs, one bank at a time. The DIMMS in each bank must be of the same type and capacity or the performance of the memory subsystem will be degraded. LEDs on the front panel of the memory board show the operating status of the DIMMs.

Figure 8. ProLiant ML530 G2 system memory banks (top) and front panel of memory board (bottom)
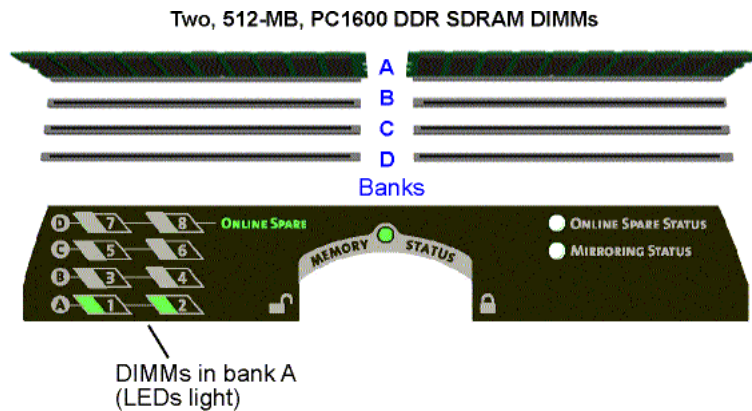
PC1600 DDR SDRAM vs. PCI33

PC1600 DDR SDRAM uses a different naming convention than PC133 SDRAM. The term PC133 signifies DIMMs with memory access times fast enough to work with 133-MHz buses. The emergence of new memory technologies such as Rambus® DRAM and DDR SDRAM, however, made it necessary to develop a different naming convention based on the actual peak data transfer rate in MB/s. For example, PC1600 DDR SDRAM has a data transfer rate of 1,600 MB/s. PC1600 DDR SDRAM has the same data bus width as PC133 SDRAM (64 bits plus ECC bits), but it transfers data twice per clock cycle (on both the rising and falling edges of the clock signal).
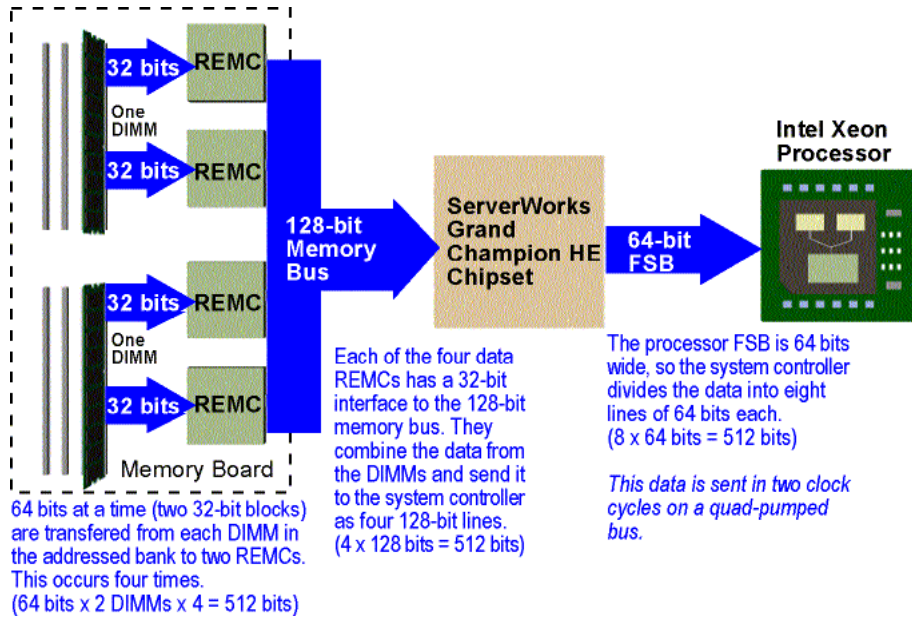
Two-way interleaved memory

The ProLiant ML530 G2 server uses two-way interleaving to improve memory performance. Two-way interleaving works by dividing memory into multiple 64-bit blocks that can be accessed two at a time, thus doubling the amount of data obtained in a single memory access from 64 bits to 128 bits and reducing the required number of memory accesses. Reducing the number of memory accesses also reduces the number of wait states, further improving performance.

When data is written to memory, the memory controller distributes, or interleaves, the data across two DIMMs in a particular bank. When a cache line of data is requested by the processor, the request is sent to the REMC dedicated to addressing. This REMC identifies the specific location of the data on the two DIMMs in the addressed bank. The other four REMCs simultaneously retrieve the 32-bit blocks of data from both of the DIMMs in the addressed bank (Figure 9).

In addition to the requested data, the controllers retrieve data from subsequent sequential memory addresses on both DIMMs in anticipation of future data requests. The retrieved blocks of data are merged together in 128-bit lines on the memory bus. The data is sent to the processor's L2 cache as four 128-bit lines (512 bits) to match the cache line size in the Intel Xeon Processor. The data rate on

the memory bus matches the data rate on the quad-pumped processor bus (3.2 GB/s), which reduces latency in memory reads and writes.

Figure 9. Memory read using interleaving

What are the software application performance benefits of memory interleaving? Dual-interleaved memory fills the processor cache faster than standard, non-interleaved memory systems so that the processors can execute applications faster. This synergy between the processor and memory subsystems boosts the overall system performance of the ProLiant ML530 G2 server well beyond that of 2P servers without Hyper-Threading technology and two-way memory interleaving.

I/O subsystem

The ServerWorks Grand Champion HE chipset ensures that the bandwidth of the I/O subsystem complements the processor and memory bandwidths (see Figure 10). The chipset supports three 400-MHz (double-pumped 200-MHz clock) Inter Module (IM) Buses. Two 32-bit IM Buses are used to provide 3.2 GB/s (1.6 GB/s each) to two PCI-X bridges. Each PCI-X bridge controls two 800-MB/s (100-MHz, 64-bit) PCI X buses. A maximum of two PCI-X slots per bus are used for better load balancing of I/O resources, such as array controllers and network interconnect controllers (NICs). Four of the seven PCI X slots support hot-plug operation. The OSs that support PCI-X hot-plug operation include Windows 2000 Server products, Netware 4.2 and higher, and SCO UNIX.

PCI-X provides full backward compatibility with PCI 2.2 hardware and software, thereby preserving customer investments as I/O technology continues to evolve.

**Figure 10.** Architecture of the ProLiant ML530 G2 server I/O subsystem
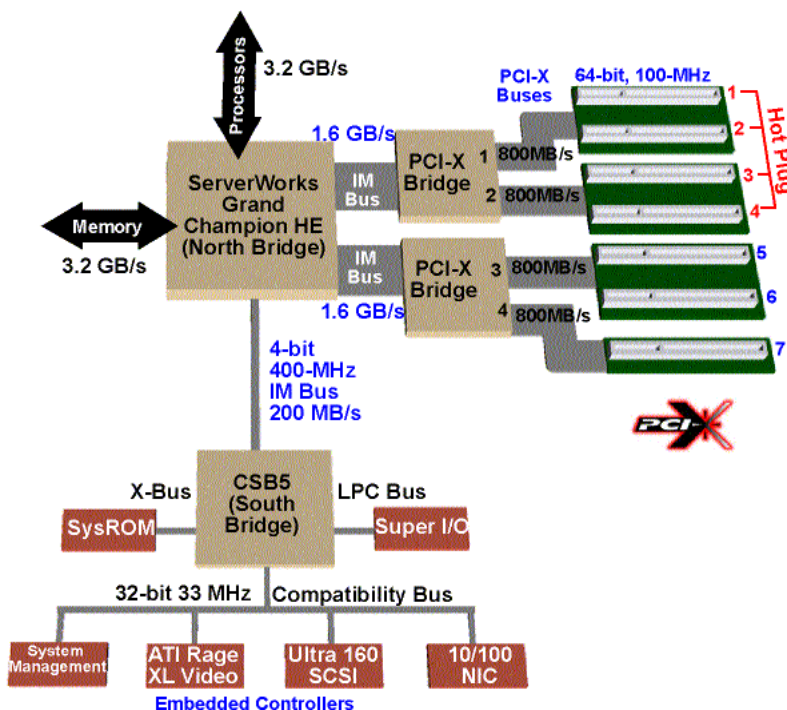
The north bridge uses the third 400-MHz IM Bus (4-bit width) to connect to the ServerWorks CSB 5.0 south bridge. The south bridge provides interfaces to the following buses:

- LPC bus — This bus provides connection to a National NS417 Super I/O controller for diskette, keyboard, mouse, parallel, and serial port support.
- X-Bus — The 2-MB redundant system ROM and bootblock are connected through this bus.
- Compatibility bus — The 33-MHz, 32-bit PCI bus supports the system management controller, ATI Rage XL Video controller, Adaptec 7899 Dual Channel Ultra 160 SCSI controller, and Intel 82559 (10/100) NIC. All controllers are embedded in the system board.

The embedded 10/100 NIC provides high-speed LAN capability while saving a PCI slot for other needs. The configuration utility enables customers to set up NICs for load balancing or failover functions. The Adaptec 7899 Dual Channel Ultra3 SCSI controller operating at 160 MB/s per

11

channel provides twice the data rate of the Ultra2 controller used in the first generation ProLiant ML530 server. The SCSI controller has two ports, which are cabled to two 6-inch x 1-inch hot-plug drive cages in the front of the server to support up to 14 hard drives (Figure 11). A third 2-inch x 1-inch hot-plug SCSI drive cage is optional. It fits in the full-height removable media bay area and requires a dedicated SCSI channel. The SCSI backplanes on the hot-plug drive cages are built to work with Ultra 320 drives and controllers, allowing simple upgrades to faster drive technology when it becomes available.

Figure 11. Internal storage in the ProLiant ML530 G2 server
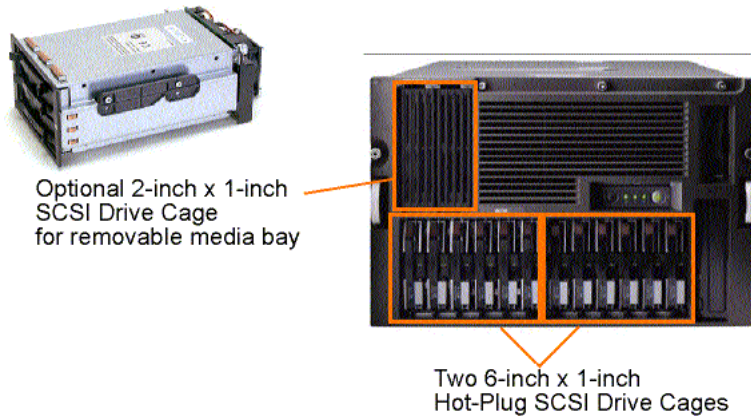
Quad-peer PCI-X architecture

The quad-peer PCI-X architecture consists of four 64-bit, 100-MHz PCI-X bus segments controlled by two PCI-X Bridges. The first PCI-X bridge provides PCI-X Hot Plug capability to slots 1 through 4. The second PCI-X bridge controls the third and fourth PCI-X bus segments. Slots 5 and 6 are connected to the third bus, and slot 7 is connected to the fourth PCI-X bus. Slot 7 should be used for Remote Insight Lights-Out Edition (RILOE) support because it is the closest slot to the virtual power button cable connectors. Slots 1, 3, and 5 should be populated before slots 2, 4, and 6 are populated for two reasons: to populate slots from the center of the server where the best cooling is available and to balance the buses for better system performance.

Note:
Since there are no PCI slots on the compatibility bus with the management controller, the RILOE must be plugged into the management connector (or power button) to enable remote power cycling (virtual power button).

Why PCI-X faster is than conventional PCI

PCI-X technology provides a significant improvement in performance beyond that of conventional PCI systems. The performance improvements are a result of two primary differences between conventional PCI and PCI-X: higher clock frequencies—made possible by the register-to-register protocol—and new protocol enhancements such as the attribute phase and split transactions.

Backward compatibility and bus performance with PCI cards

The ProLiant ML530 G2 server supports the following adapter cards:

- 100-MHz PCI-X
- 66-MHz PCI-X
- 66-MHz PCI
- 33-MHz PCI

The ProLiant ML530 G2 server supports universal adapter cards and 3.3-V PCI cards; however, it does not support 5 V-only PCI cards. The PCI-X slots are keyed so that unsupported adapter cards cannot be inserted.

The PCI-X buses operate at a maximum speed of 100 MHz. The system automatically adjusts the PCI-X bus frequency to match the frequency of the slowest adapter on that bus segment. For example, if one of the bus segments with two slots is populated with a 66-MHz adapter and a 100-MHz adapter, the maximum frequency of that bus segment will be 66 MHz. This means that the slowest adapter card, such as a 33-MHz, 32-bit RILOE card, should be put in slot 7 where it cannot slow down any other adapter cards. To make it easier to determine the speed of each bus segment, a slot speed indicator is located on the backplane over each slot (Figure 12). If no adapter is installed in a slot, the indicator will be off.

Figure 12. PCI-X slot speed indicator located on the backplane over each slot of the ProLiant ML530 G2 server
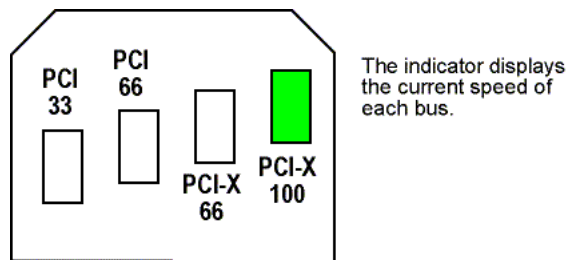
Optimum uses for the ProLiant ML530 G2 server

The ProLiant ML530 G2 server is optimized to function as a dedicated application or database server, as a volume departmental server, and as an ultra-dense Web server.

Database and dedicated applications server

Application servers are typically used to run complex multi-threaded software applications. The ProLiant ML530 server has the built-in redundancy, high-availability, and high-performance needed for distributed application services and support for complex database access. Also, its 512K L2 Advanced Transfer Cache equips the ProLiant ML530 G2 for use in CPU-intensive environments such as database applications.

Remote site server

Large remote sites, such as branch offices, require not only high performance, but also internal expansion capabilities to satisfy increasing user workloads. For example, the ever-increasing volume of e-mail traffic makes the scalability of an e-mail server a major concern, even for a relatively small organization.

The high performance of the Intel Xeon Processor (with Hyper-Threading and NetBurst technologies) allows remote sites to handle a significantly greater end-user workload per processor. This means customers can add more users per server, or more applications per server, and the system will still run significantly faster than Pentium III systems. The ProLiant ML530 G2 server ships standard with 12 hot plug drive bays. Total internal storage can be expanded to 14 drives with an optional two-bay drive cage. With its large internal drive capacity, seven PCI-X slots, embedded 10/100 NIC, and up to 16 GB of memory, the ProLiant ML530 G2 server can handle a large end-user load (500 to 1000 users).

Conclusion

The ProLiant ML530 G2 server is a mid-range departmental server optimized for intensive data center and remote office environments. It incorporates new technologies that improve upon the performance, scalability, fault tolerance, and manageability of the first generation ProLiant ML530 server. This paper has focused on the synergy of the high-performance technologies in this second generation server that provide the balanced system architecture. For more information about the ProLiant ML530 G2 server and the other technologies it incorporates, visit www.hp.com/go/proliant .

Call to action

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