Characterizing x86 processors for industry-standard servers: AMD Opteron and Intel Xeon

Technology brief

Abstract

The HP ProLiant server portfolio includes systems using the Intel® Xeon™ family of x86 processors and systems using AMD Opteron x86 processors. To help customers understand the implications of these processor architectures on the system platform, this paper compares and contrasts the Xeon processors with the Opteron processor. It discusses both the 32-bit and 64-bit computing space for those x86 processors. The paper also briefly compares the Intel Itanium® 64-bit architecture to the Xeon and Opteron 64-bit architectures.

This paper is intended for IT professionals familiar with industry-standard server technology.

Acronyms in text

The following acronyms are used in the text of this document.

Table 1. Acronyms

Acronym abbreviation	Acronym expansion
2-way	Containing two microprocessors
4-way	Containing four microprocessors
8-way	Containing eight microprocessors
AGU	Address generation unit
ALU	Arithmetic logic unit
CPU	Central processing unit, or microprocessor
DDR	Double data rate
DIMM	Dual inline memory module
DRAM	Dynamic random access memory
FSB	Front-side bus
Gb	Gigabit
GB	Gigabyte
GPR	General purpose register
KB	Kilobyte
I/O	Input/output – commonly referring to devices such as keyboards, mice, video graphics, etc.
IT	Information technology
MHz	Megahertz
ns	Nanosecond
OS	Operating system
PCI	Peripheral interconnect bus
SDRAM	Synchronous dynamic random access memory
USB	Universal serial bus

Introduction

Since the early 1980's, industry-standard computers have used the x86 instruction set architecture. During the last 20 years, the x86 architecture has been expanded with more instructions and additional registers to enable easier floating point and multimedia calculations. Throughout these expansions, the x86 architecture has maintained a high level of compatibility with 16-bit and 32-bit software, providing important backward compatibility with software applications. Furthermore, the x86 architecture has offered continually increasing performance at decreasing cost levels.

There are many versions of Intel x86 processors, configured for different markets. As of this publication, the most recent models of the Intel x86 processors for industry-standard servers are the Pentium® 4, Xeon™, and the Xeon MP processors (Table 1). All the Intel processors listed in Table 1 use the NetBurst® Architecture and Hyper-Threading Technology, but they have varying core frequencies, system bus frequencies, and amounts of cache. Hyper-Threading Technology[1] provides multi-thread level parallelism in a single processor core. The newest Pentium 4 and Xeon processors (previously code-named Prescott and Nocona) also support Extended Memory 64-bit technology (EM64T) that allows 64-bit operating systems and applications to run natively. The use of 64-bit extension technology will enable IT organizations to deploy common platforms for both 32-bit and 64-bit computing and to move to 64-bit computing gradually as it benefits their businesses.

Table 1. Processor Specifications (maximum values are given)

Processor	Core frequency (GHz)	L1 cache (KB)	L2 cache (KB)	L3 cache (MB)	Front side bus speed	Feature size (nm)	Supports 64-bit?	Supports Hyper-Threading?
Pentium 4 Supporting Hyper-Threading technology (previously code-named Prescott)	3.6 GHz	16 K data 12K µop instruction	1024	None	800 MHz	90	Yes	Yes
Xeon (previously code-named Nocona) Announced in mid-2004	3.6 GHz	16 K data 12K µop instruction	1024	None	800 MHz	90	Yes	Yes
Xeon MP (Gallatin core)	3.0	8 K data 12K µop instruction	512	4	400 MHz	130	No	Yes
Opteron 2xx (Hammer core —for up to 2-way servers)	2.4	64K data 64 K instruction	1024	None	N/A - Integrated memory controller	130	Yes	No
Opteron 8xx (Hammer core —for up to 8-way servers)	2.4	64K data 64 K instruction	1024	None	N/A - Integrated memory controller	130	Yes	No

[1] Additional information about Hyper-Threading is available in the technology brief titled "The Intel® processor roadmap for industry-standard servers," TC040504TB, at http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00164255/c00164255.pdf

In 2003, AMD introduced its eighth-generation x86 processor, the Opteron. The Opteron was the first processor in the industry to deliver 64-bit extensions (the AMD64 instruction set) for running 64-bit operating systems and applications.

A three-digit model number, Zxx, identifies Opteron processor models. The "Z" indicates the maximum number of processors that can be connected in a system, and "xx" indicates the relative performance within the series. The Opteron 2xx series processor is designed for 2-way servers; the Opteron 8xx series supports up to 8-way servers.[2] Both processors use a 0.13-micron process and operate at up to 2.4 GHz. Each processor includes an integrated, 128-bit DDR memory controller, an integrated 64-KB L1 instruction cache, 64-KB integrated L1 data cache, and a 1-MB L2 cache. Because the memory controller is integrated and the Opteron processors use the HyperTransport architecture (a point-to-point connection to other processors and I/O devices), there is no need for a front-side bus system such as the Xeon uses.

Micro-architectural similarities

Both the Xeon family processors and the Opteron processors adhere to the x86 instruction set architecture in order to be compatible with the wealth of 32-bit software applications available. Therefore, both processors perform in the same way at the programming level. In other words, at the software/hardware interface, each processor's software interface remains the same with regard to the memory addressing size, the instruction sets, and the register designs for the x86 architecture.

32-bit operations

A 32-bit processor has general purpose registers (GPRs) that are 32 bits wide and can operate on an integer data stream that is 32 bits wide. In addition, and what is most commonly understood when discussing 32-bit architectures, a 32-bit processor can hold 32 bits of memory address data in a single register, for a maximum of 4 GB of addressable memory.

The x86 architecture also supports physical addressing extensions (PAE), which extend the address space to allow addressing to 36 bits for a maximum of 64 GB of physical addressable memory. However, this requires the OS and applications to take advantage of the additional memory addressing.[3] The Xeon family processors and the Opteron processors support 32-bit addressing as well as the 36-bit PAE.

As shown in Table 2, the x86, 32-bit instruction set common to both the Xeon family processors and Opteron processors includes:

- Standard x86, which are general purpose arithmetic functions
- Single Input Multiple Data (SIMD) Instructions, which allow one command to work simultaneously on multiple data items. This includes MMX, Streaming SIMD Extensions (SSE) and SSE2.
- x87 floating point instructions

[2] For more information about naming of the AMD processors, see www.amd.com/us-en/Processors/ProductInformation/0,,30_118_8796_9240,00.html.

[3] For more information, see "36-Bit Physical Addressing Using the PAE Paging Mechanism" in Chapter 3 of the IA-32 Intel Architecture Software Developer's Manual, Volume 3., available at http://www.intel.com/design/pentium4/manuals/index_new.htm

Table 2. 32-bit x86 instructions common to both Intel and AMD processors

Instruction name	Description	Register type	Size of registers	Number of registers
standard X86	Instructions for logical and arithmetic operations, address calculations, and holds memory pointers	GPR	32-bit	8
MMX	Multimedia instructions that allow the processor to do 64-bit SIMD operations on integers (packed integer data types)	MMX	64-bit	8
x87	Instructions for floating point calculations	FP	80-bit*	8
SSE and SSE2	SSE improved upon the MMX instructions and allowed processors to do 128-bit SIMD floating-point operations. SSE2 brought 64-bit parallel floating point numeric support to the IA-32 architecture. It also extended old instructions and added new ones to support 128-bit SIMD integer	XMM	128-bit	8

* According to the article "Introduction to 64-bit Computing and x86-64," the "x87 uses 80-bit registers to do double-precision floating point. The floats themselves are 64-bit, but the processor converts them to an internal, 80-bit format for increased precision when doing computations." [4]

There are two differences to note between the 32-bit instruction sets: support for SSE3 and support for 3Dnow! The latest Pentium 4 and Xeon processors support SSE3 instructions. SSE3 instructions include 13 instructions that accelerate performance of SSE technology, SSE2 technology, and x87-FP math capabilities. It is expected that AMD will support SSE3 in future versions of the Opteron processor. [5]

The Opteron supports the AMD 3Dnow! instructions. When Intel introduced the MMX instructions, they were not widely used, and AMD developed its own version of multimedia instructions with the 3DNow! instructions. The 3DNow! set added SIMD instructions to improve the vector-processing (floating point) requirements of graphic-intensive and multimedia applications. The 3DNow! instructions use the same registers as the MMX instructions.

64-bit operations

64-bit architectures have registers and arithmetic logic units that are capable of manipulating 64 bits of data during a single processing step. Because the registers store addresses to memory, this also means that a 64-bit architecture has a much larger amount of directly addressable memory than a 32-bit processor. Therefore, 64-bit architectures can provide performance advantages by their ability to use large amounts of memory (for example, in data mining operations) and by their ability to manipulate large amounts of numbers, such as with intensive floating-point calculations used in scientific and engineering modeling programs.

Instruction set and registers

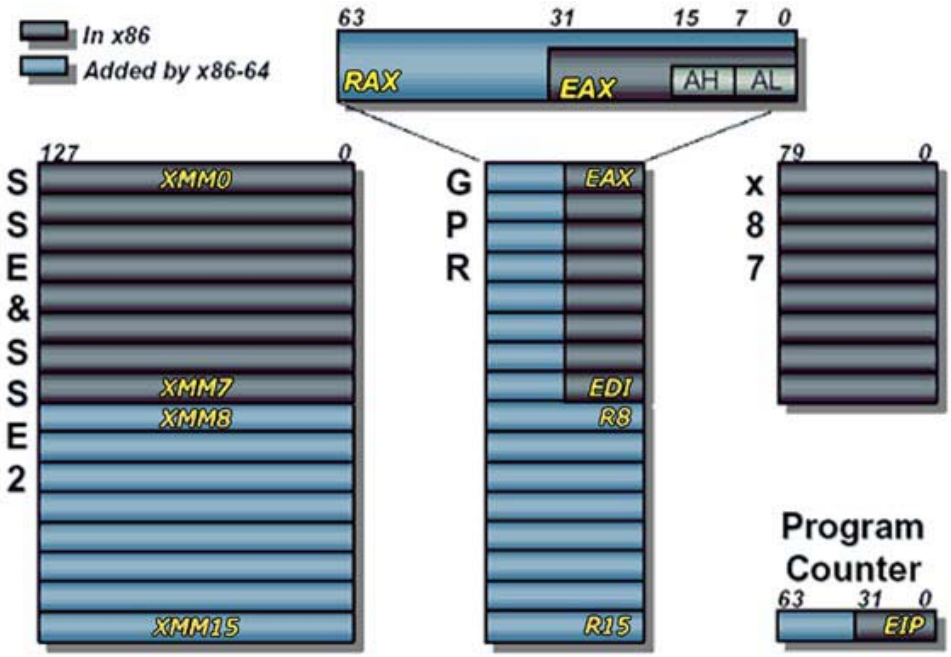
As already noted, AMD was the first to develop the 64-bit extensions with their AMD64 instructions. Intel then delivered their EM64T instruction set that is broadly compatible with AMD64 for 64-bit computing. Both AMD64 and EM64T instructions can take advantage of the 64-bit wide registers in Opteron and the latest Xeon processors. These registers are used by the applications only when

[4] "Introduction to 64-bit Computing and x86-64," Jon Stokes, available at http://arstechnica.com/cpu/03q1/x86-64/x86-64-1.html
[5] As reported by the website TechReport, "Future Athlon 64s to include SSE3 support ", Scott Wasson, March 2, 2004, http://techreport.com/onearticle.x/6363

running the processors in 64-bit long mode. To support the AMD64 or EM64T instructions, the registers expand to include:

- 8 new 64-bit GPRs
- Extensions of the 8 original, 32-bit GPRs to 64 bits
- 8 new 128-bit registers for SSE and SSE2 instructions (and SSE3 instructions for Xeon)

Figure 1. Difference between 32-bit and 64-bit registers (source: "AMD Opteron Coverage – Part 1: Intro to Opteron/K8 Architecture,"[6] April 23, 2003)



Operating modes

The processors use three different operating modes: 64-bit long mode; 64-bit compatibility mode; and 32-bit legacy mode. The 64-bit long mode requires a 64-bit OS and an application recompiled to use the 64-bit registers. In other words, the full capabilities of the expanded register set are available only when both the OS and the application support 64 bits. The 64-bit compatibility mode requires a 64-bit OS, but can use a 32-bit application. The additional registers are available to the OS, but not to the 32-bit application, since it cannot make use of them. When running in legacy mode, the processor acts just like a 32-bit processor and the extra registers are not available (Table 3).

[6] Available at www.anandtech.com/cpuchipsets/showdoc.aspx?i=1815

Table 3. Operating modes for Xeon family and Opteron processors[7]

Mode	OS required	Application recompile required?	Register extensions available?	GPR width (bits)
64-bit long mode	64-bit OS	Yes	Yes	64
64-bit compatibility mode	64-bit OS	No	Yes – to OS No – to application	32
32-bit legacy mode	32-bit or 16-bit OS	No	No	32

Memory addressability

Neither the Xeon family processors nor the Opteron processors use all 64 bits in the 64-bit registers to address memory. Because this is such an extremely large number (2^{64} bits equals 16 exabytes)[8], current applications cannot support this much memory in either local memory or page files. In addition, the use of all 64 bits to address memory can cause the architecture to be more complex than it needs to be, wasting space and resources. Since the majority of memory accesses would not require such a large number to identify the address, having 64-bit memory addresses would result in lots of extra bits taken up by zeros. Therefore, when operating in 64-bit long mode, the Xeon family processors support up to 48 bits of virtual memory (256 terabytes) and 36 bits of physical memory (64 GB). Opteron processors, when operating in 64-bit long mode, support up to 48 bits of virtual memory (256 terabytes) and 40 bits of physical memory (1 terabyte).

Micro-architectural differences

As just discussed, the fundamental design of the Xeon family and Opteron processors are similar at a programming level – in other words, the instruction/register/addressing schemes are essentially the same at the software/hardware interface. Of course, there are many micro-architectural differences in the implementation of those processors (how the processor decodes instructions, how the branch prediction buffers work, how many branch predictions can be held in those buffers, etc.). However, it is important to understand two main differences between the processors:

- Overall design of the pipelines and how this relates to processor frequency. Xeon family processors optimize frequency at the expense of pipeline execution efficiency, while Opteron processors optimize the pipeline efficiency at the expense of frequency.
- Implementation of multi-threaded parallelism. Xeon family processors includes Hyper-Threading technology, which allows different threads to operate in parallel, in addition to instruction-level parallelism. Opteron offers only instruction-level parallelism.

Pipelines and clock frequency

The pipeline in a processor is analogous to an assembly line in a factory: to execute program code (or to build a widget), the work is split into multiple "stages," with each stage comprising a small part of the whole job. The idea, of course, is that splitting up the work into stages keeps the processor (or factory worker) busy at all times, allowing the processor to execute more code during a certain period of time.

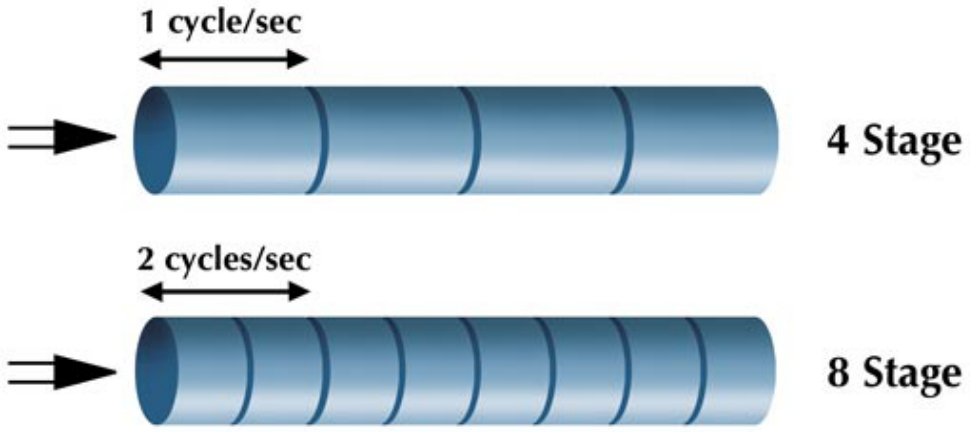
It is important to understand the amount of activity that can occur within each stage of the pipeline. The processor must complete the operation for each stage within a single clock cycle. If the processor

[7] From the documents titled "Intel Extended Memory 64 Technology Software Developer's Guide," Vol. 1, available at ftp://download.intel.com/technology/64bitextensions/30083402.pdf; and from "AMD64 Architecture Programmer's Manual, Vol. 1: Application Programming", available at www.amd.com/us-en/assets/content_type/white_papers_and_tech_docs/24592.pdf

[8] Appendix B defines some of the less familiar engineering prefixes such as exabytes.

reduces the task size by splitting it into two or more smaller tasks, each stage can be shorter, but there will be more stages (Figure 2). Thus, each stage can be completed more quickly, allowing the processor to have a higher clock frequency. This does not necessarily mean that more work is being done in the pipeline; it just means that the clock frequency can be higher.

Figure 2. Decreasing the amount of work done in each stage allows the clock frequency to increase.

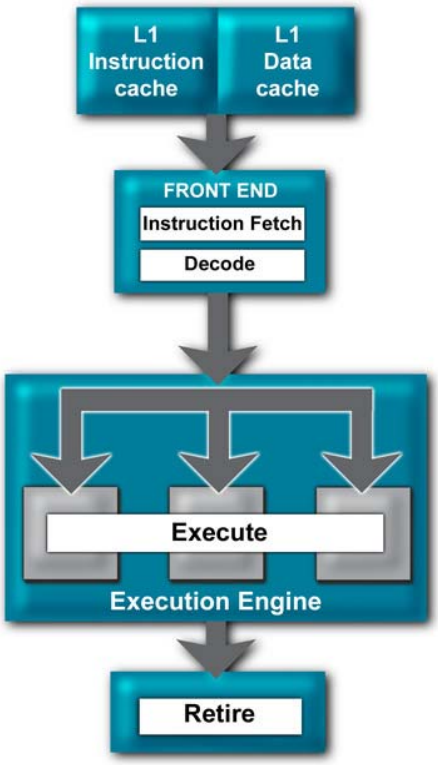


A basic structure for a computer pipeline consists of the following four steps, which are performed repeatedly to execute a program:

1. Fetch the next instruction from the address stored in the program counter.
2. Store that instruction in the instruction register and decode it, and increment the address in the program counter.
3. Execute the instruction currently in the instruction register.
4. Write the results of that instruction from the execution unit back into the destination register.

Typical processor architectures split the pipeline into segments that perform those basic steps: the "front end" of the microprocessor, the execution engine, and the retire unit, as shown in Figure 3. The front end fetches the instruction and decodes it into smaller instructions (commonly referred to as micro-ops). These decoded instructions are sent to one of the three types of execution units (integer, load/store, or floating point) to be executed. Finally, the instruction is retired and the result is written back to its destination register.

Figure 3. Basic 4-stage pipeline schematic



Processor stalls due to cache misses

To keep the pipeline busy requires that the processor begin executing a second instruction before the first has traveled completely through the pipeline. However, suppose a program has an instruction that requires summing three numbers:

X = A + B + C

What happens if the processor already has A stored in a register, and B stored in a register, but needs to get C from memory? This causes a "bubble" or a stall in the pipeline, in which the processor cannot execute the instruction until it obtains the value for C from memory. This bubble must propagate all the way through the pipeline, forcing each stage that contains the bubble to sit idle, wasting execution resources during that clock cycle.

Clearly, the longer the pipeline, the more this is a problem.

Processor stalls due to branch misprediction

Processor stalls often occur as a result of one instruction being dependent on another. If the program has a branch, such as an IF… THEN loop, the processor has two options. The processor either waits for the critical instruction to finish (stalling the pipeline) before deciding which program branch to take; or it predicts which branch the program will follow.

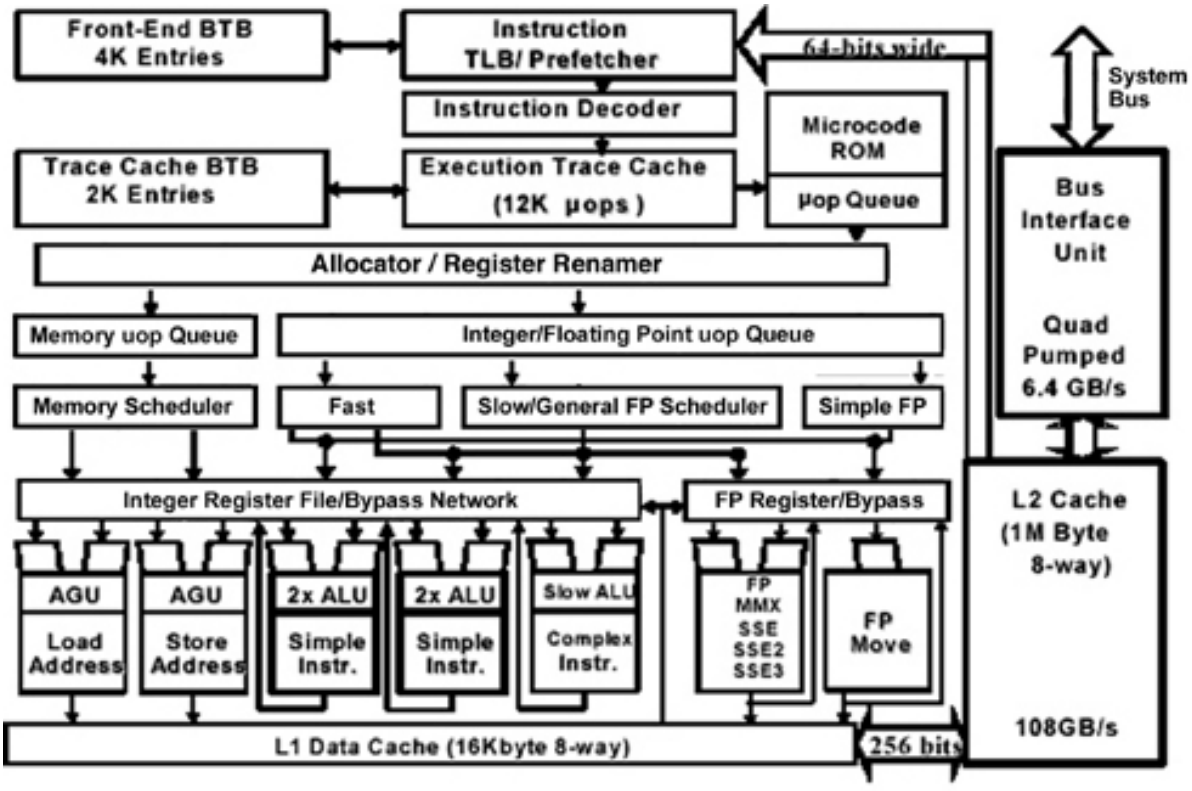
If the processor predicts the wrong code branch, it must flush the pipeline and start over again with the IF… THEN statement using the correct branch. The longer the pipeline, the higher a performance cost for branch mispredicts. For example, the longer the pipeline, the more the processor must execute speculative instructions which must be discarded when a mispredict occurs.

Xeon

As stated in documentation by Intel, the NetBurst architecture used in the Xeon provides a deep pipeline in order to enable industry-leading clock rates.[9] In other words, Intel designs their processors for raw clock speed. The most recent versions of the Xeon family processors use a *31-stage integer pipeline*, with two "fast" arithmetic logic units (ALUs), one "slow" ALU, two address generation units (AGUs), and two floating-point execution units (Figure 4). The long pipeline allows the Xeon to operate at frequencies up to 3.6 GHz.

One of the ways that Intel has compensated for such a long pipeline is the 12KB instruction cache known as the Execution Trace Cache. Intel reports that this 12K cache has a hit rate similar to a conventional 8 – 16 KB instruction cache.[10] It stores translated and decoded instructions, and puts them into traces, or "mini-programs". Whenever there is an L1 cache hit, the processor executes these traces without having to translate and decode the instructions each time it executes that particular trace – thus, reducing the amount of work that the main pipeline must do. Furthermore, the Execution Trace Cache includes its own branch prediction algorithms so it can store translated micro-ops in speculative order.

Figure 4. Schematic of the Xeon processor microarchitecture (source: The MicroArchitecture of the Intel Pentium 4 Processor on 90nm Technology, Intel Technology Journal, Vol 8, Issue 1, Feb. 2004.)[11]



[9] "IA-32 Intel Architecture Software Developer's Manual," Volume 1:Basic Architecture, pg. 2-7, available at http://developer.intel.com/design/pentium4/manuals/index_new.htm

[10] Intel Technology Journal, Vol. 8, Issue 1, 2004, "MicroArchitecture of the Intel Pentium 4 Processor on 90 nm Technology."

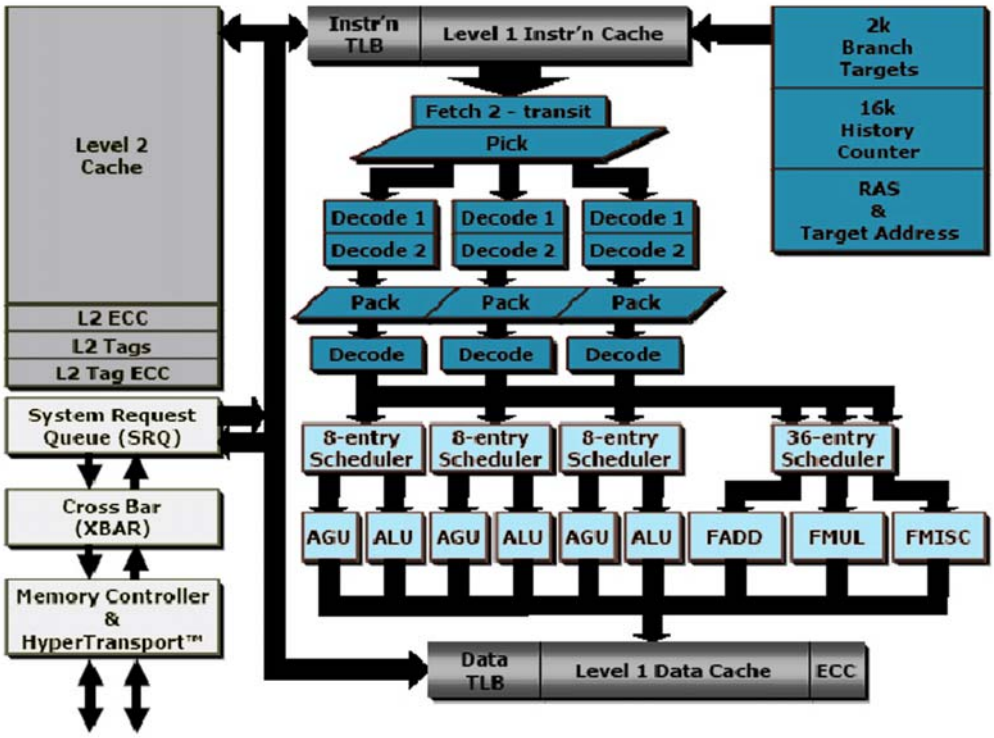
[11] Available at http://www.intel.com/technology/itj/2004/volume08issue01/art01_microarchitecture/p01_abstract.htm

Opteron

Unlike the Xeon family processors, the Opteron processor has been optimized to deliver a highly efficient pipeline. According to AMD,[12] this is due to the pipeline's front end instruction fetch and decode logic, which has been optimized to pack multiple decoded, micro-op instructions together to execute them in parallel. Opteron has a *12-stage integer pipeline*, much shorter than the pipeline for Xeon processors. This requires the frequency be slower: The Opteron currently operates in a range of 1.6 to 2.4 GHz. However, the shorter pipeline reduces the risk of delays due to branch mispredictions and cache misses. The shorter pipeline also requires less extensive branch prediction algorithms and target buffers.

Opteron also has more execution units and decode units than Xeon, to make operations more parallel. The Opteron includes three ALUs, three AGUs, and three floating-point execution units (Figure 5). Although Opteron has more individual execution units than Xeon, the maximum effective throughput of these execution units is the same as for Xeon—three integer operations per cycle.

Figure 5. AMD Opteron architectural block diagram (source: AMD Eighth-generation Processor Architecture paper).



Pipeline and frequency comparison

In contrast to the Xeon processors which have high frequencies and long pipelines, processors that are designed to achieve high performance—such as the Alpha EV68, HP-PA RISC, and the Itanium—tend to be designed with shorter pipelines and to operate at relatively lower frequencies. For

[12] "AMD Eighth-generation Processor Architecture," available at www.amd.com/us-en/assets/content_type/white_papers_and_tech_docs/Hammer_architecture_WP_2.pdf

example, the Alpha EV68/21264 uses a 7-stage pipeline and operates at approximately 1.25 GHz; and the Itanium 2 uses an 8-stage pipeline and operates at approximately 1.6 GHz. This does not mean that Xeon is automatically lower performing because it has a longer pipeline. However, it does mean that the Xeon processors must compensate for the long pipeline by techniques such as more efficient branch prediction algorithms and larger look-aside buffers. It also means that Xeon processors may be best suited for applications where raw clock speed is important, such as applications using a linear programming style.

Power considerations

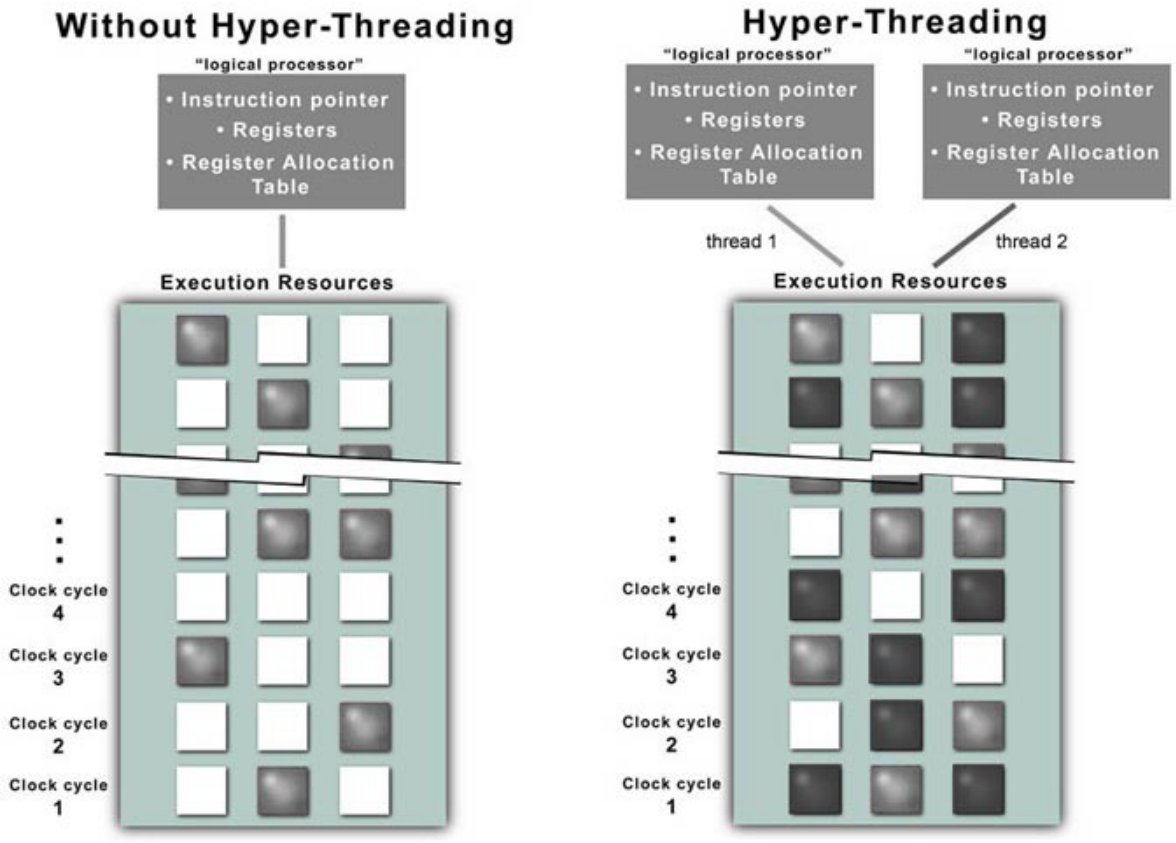
One final note on the difference between the Xeon and Opteron pipelines: Because power is directly related to the operating frequency of the processor, as the frequency increases, the power also increases. Therefore, because Opteron runs at lower operating frequencies than Xeon, it will also require less power and run cooler than the Xeon processors. Power consumption is increasingly becoming an issue for large datacenters and customers using densely packed racks.

Hyper-Threading in the Xeon family processors

The Intel NetBurst architecture incorporates Hyper-Threading as another method to keep its pipeline full. Normally, a processor executes instructions from only a single instruction stream. Whenever a switch occurs between threads, or streams of instructions, the processor must pay an overhead penalty for the context switch. Because modern processors are superscalar (have parallel execution units and out-of-order execution), the processors re-arrange the instructions and execute them out-of-order to perform some instructions in parallel (instruction-level parallelism). Both the Xeon and the Opteron processors use instruction-level parallelism to execute up to three instructions per clock.

In addition, the Xeon family processors use Hyper-Threading technology to execute two separate threads in parallel (multi-threaded parallelism). The Xeon family processors replicate certain components of the processor such as the instruction pointer, register allocation tables, and other architectural registers. Intel refers to these components as the "architectural state." The architectural states, or logical processors, schedule instructions for the processor execution resources, which are shared between the two logical processors. When the processor identifies a bubble in one thread stream, it shifts the second thread stream into those execution resources. This allows the processor to execute incoming micro-op instructions from different threads in a time-multiplexed manner to keep execution resources as busy as possible (Figure 6).

Figure 6. Comparison of processors with and without Hyper-Threading technology.[13] In the Hyper-Threading example on the right, the light gray indicates instructions from thread 1, and the dark gray indicates instructions from thread 2. The white squares are idle execution resources.



Hyper-Threading can improve system performance for applications and operating systems that can take advantage of multi-threading. Intel reports maximum gains of up to 30 percent,[14] but this is highly dependent on the application. Average performance improvements are more likely to be approximately 5 to10 percent.[15] The actual performance increase depends on how well the OS and applications take advantage of multiple threads. In other words, applications and operating systems that are not multi-threaded may not gain any performance benefits from Hyper-Threading.

[13] This figure was created from information provided in the Intel white paper "Hyper-Threading Technology on the Intel Xeon Processor Family for Servers."
[14] Source. "Hyper-Threading Technology on the Intel Xeon Processor Family for Servers," available at www.intel.com/business/bss/products/hyperthreading/server/ht_server.pdf
[15] The Anandtech website reports 3 to 5 percent gains in database performance testing; "AMD Opteron vs. Intel Xeon: Database Performance Shootout", March 2, 2004, www.anandtech.com/

System architecture differences

Server performance depends not only upon the processor performance itself, but also upon the memory subsystem, I/O subsystem, and the types of applications running. The Xeon family processors and Opteron architectures are vastly different in their memory and I/O subsystems:

- Memory controller implementation. The Xeon family processors have a north bridge that links to a memory controller; the Opteron processors have an integrated memory controller.
- I/O transport. The Xeon family processors use a parallel, shared, front-side bus. Opteron processors use the point-to-point HyperTransport links that can offer significantly higher bandwidth and lower latencies.

Memory controller

Today's processors operate at much faster speeds than the memory subsystem. Improving the memory subsystem, then, yields overall system performance benefits that cannot be achieved simply by increasing the speed or performance of the processor core. Both the Xeon family processors and Opteron processors use industry-standard DDR SDRAM. However, the Xeon processors use the well-known architecture of a front-side bus to connect to a separate memory controller, while the Opteron uses a memory controller integrated into the processor itself.

Xeon front side bus

There are three important considerations when discussing the memory architecture: bandwidth, latency, and scalability.

Bandwidth

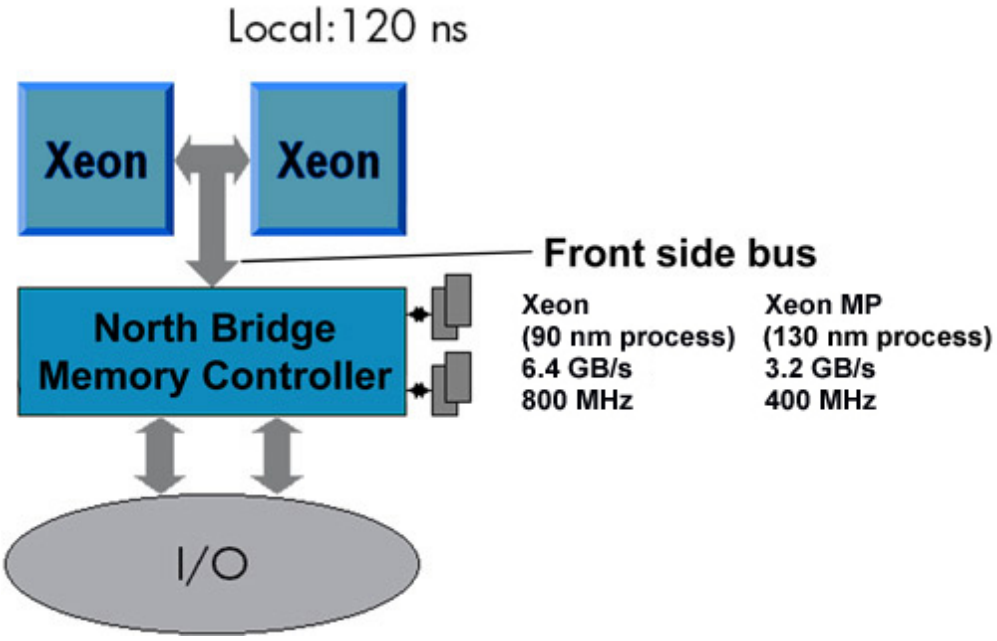
The Xeon front-side bus (FSB) is a parallel, 64-bit, multi-drop technology that shares bandwidth between all processors and the memory controller. The memory controller is in a separate chipset from the processor, and the amount and type of memory that can be used depends on the chipset design (determined by the platform designer). Typically, HP has used either Intel or Serverworks chipsets with the Xeon family processors. They support up to 32 GB of dual-channel DDR SDRAM.

The most recent Xeon family processors use an 800-MHz front-side bus to connect the processors to the memory and I/O devices (Figure 7). This gives a total maximum bandwidth of 6.4 GB/s. Earlier versions of the Xeon family processors use the 400-MHz or 533-MHz front-side bus, which limits the total bandwidth to 3.2 GB/s or 4.3 GB/s.

Latency

The Xeon family processors are designed for use in a symmetric multi-processing environment, in which every processor has equal access to the memory. The memory latency—the time it takes for a processor to request data from memory—is uniform across all processors. HP measurements show that the memory latency for a Xeon family processor is about 120 ns.

Figure 7. Block diagram of a typical 2-way server design using the Xeon family processors



Scalability

For single-processor systems, the FSB architecture may be an advantage because memory resources can be added without adding more processors. However, as customers add more processors, sharing the bandwidth between multiple processors (2 or 4) becomes an increasing problem. The shared nature of the front side bus results in frequent arbitration and thus higher latencies, which may limit the performance of the Xeon family processors with applications that are memory intensive.

Opteron integrated memory controller

Unlike the Xeon FSB architecture, the Opteron processor integrates the memory controller into the processor. This gives distinct benefits in all three critical memory concerns: It increases the bandwidth, reduces memory latency, and improves scalability compared to the Xeon family processors.

Bandwidth

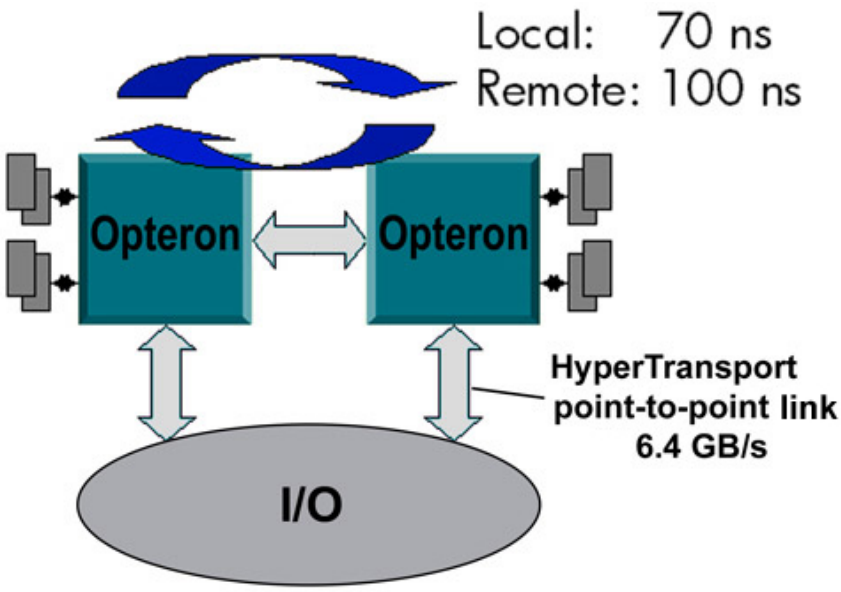
The Opteron incorporates a dual-channel DDR SDRAM controller with a 128-bit interface that is capable of supporting up to eight DDR DIMMs (four per channel).

The controller is designed to support PC1600, PC2100, PC2700, and PC3200 DDR memory using registered DIMMs. For PC3200 memory, which operates at an effective transfer rate of 400 MHz, this gives a bandwidth of 3.2 GB/s per channel, or a total of 6.4 GB/s for both channels to the processor. For a customer using systems with older versions of Xeon processors, this could mean that systems using Opteron have up to 100 percent greater bandwidth.

Latency

Figure 8 shows how the Opteron processor relates to the memory in the system architecture. Each processor has memory attached to it locally. The other processors can access that memory by means of a crossbar switch internal to the processor and a fast, point-to-point interconnect between processors (the HyperTransport interconnect).

Figure 8. Example of how the Opteron relates to memory in a 2-way system.



Some customers may be concerned that this architecture, which splits the memory into a processor's local memory and memory that is remote to the processor, will result in large latencies. However, the latency difference between local and remote accesses is actually very small because the memory controller is integrated into and operates at the core speed of the processor, and because of the fast interconnect between processors. In a dual-processor system, the memory latency for a local access is about 70 ns, while a remote access is about 100 ns—which are both less than the comparable Xeon latency (Table 4).[16] Because the difference between local and remote memory accesses is so small, AMD refers to it as "sufficiently uniform" memory. Although the memory subsystem design is not uniform, the speed of the interconnects allows it to appear uniform.

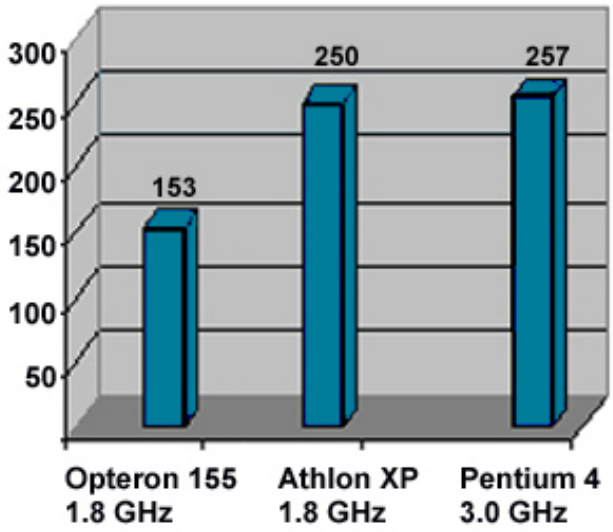
Table 4. Comparison of memory latencies for 2-way servers

Type of memory access	Xeon	Opteron	Percent difference
local access	120 ns	70 ns	42 %
remote access	120 ns	100 ns	17 %

If one compares the number of clock cycles that are spent during memory accesses, the difference in clock cycles is even greater than the actual time difference (Figure 9). Because the clock speed runs faster in a Xeon processor, the higher memory latency results in an even greater number of clock cycles that the CPU is potentially waiting for memory accesses.

[16] "AMD Hammers Multiprocessor Hyperchannel," William Wong , ED Online ID #1680, April 01, 2002, www.elecdesign.com/Articles/ArticleID/1680/1680.html

Figure 9. Number of clock cycles spent waiting for memory accesses (Source: "AMD Opteron Coverage – Intro to Opteron/K8 Architecture," April 23, 2003)[17]



In addition, with newer versions of operating systems, the OS has the ability to take advantage of local and remote memory. HP ProLiant platforms such as the ProLiant DL585 have a setting in the ROM-Based Setup Utility to exploit local memory on the server nodes rather than going through the crossbar switch in the processor to remote memory.

Scalability

Because customers can add memory with each processor, total memory can scale linearly with the number of processors. For example, a 4-way Opteron system could have up to 64 GB of memory, while a comparable Xeon system could have only 32 GB. For applications that can use this amount of memory, an Opteron-based system could provide significant performance advantages.

Furthermore, because each Opteron processor contains its own memory controller and HyperTransport links, the available memory bandwidth scales linearly with the number of processors. The integrated memory controllers also allow multiple memory requests to be made in parallel. The parallelism increases the effective bandwidth to memory and decreases the average memory latency.

I/O Interconnect

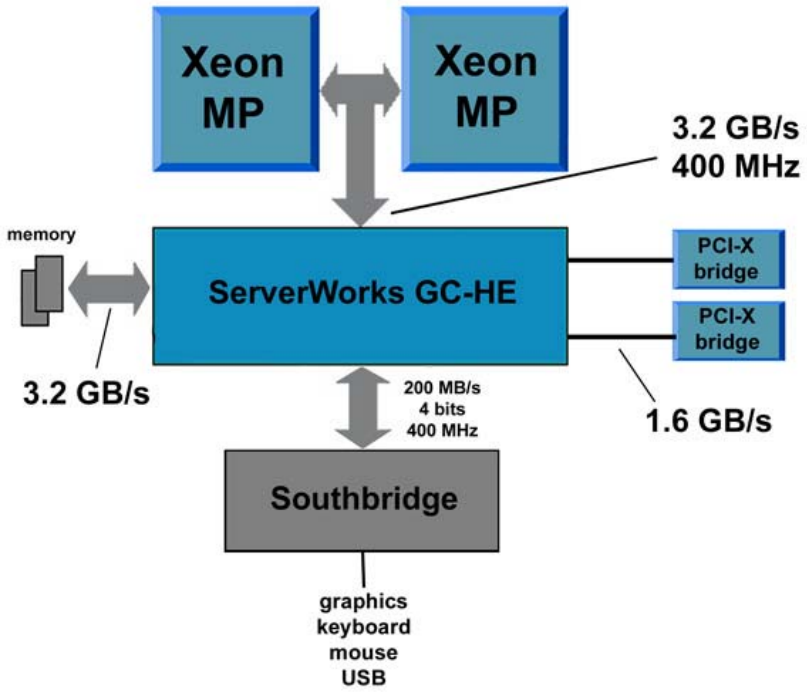
As already discussed, the Xeon processors use the familiar front-side bus technology to connect the processors to a north bridge that links to I/O and memory. The Opteron, on the other hand, uses a point-to-point HyperTransport link. AMD developed the HyperTransport link as a high-performance link to reduce the number of buses within a system and enable highly scalable multiprocessing systems.

Xeon north bridge/south bridge

The I/O subsystem speed and bandwidth depend on which chipset the platform designers select. Figure 10 shows an example of how a Serverworks chipset such as the Grand Champion-HE might be used in a system configuration. This chipset has been used in ProLiant servers and offers an aggregate I/O bandwidth of 3.2 GB/s to the PCI-X bridges, with an additional 200 MB/s bandwidth to the south bridge, which controls the video, networking, storage, and system management devices.

[17] Available at www.anandtech.com/cpuchipsets/showdoc.aspx?i=1815

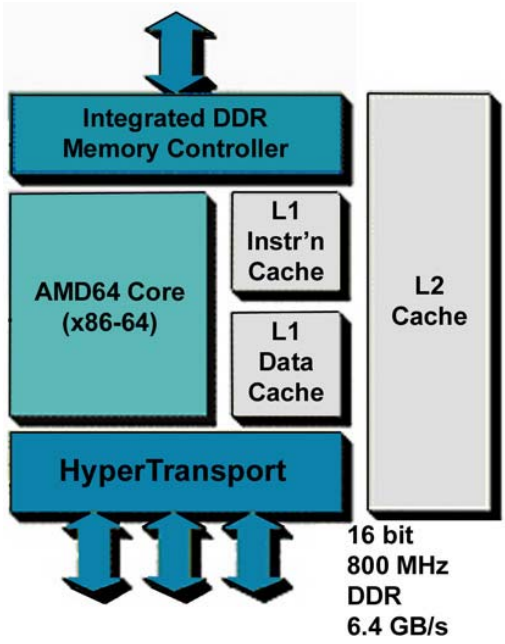
Figure 10. Example of a 2-way system architecture using Xeon MP processors (400-MHz FSB) and a Serverworks chipset



Opteron HyperTransport

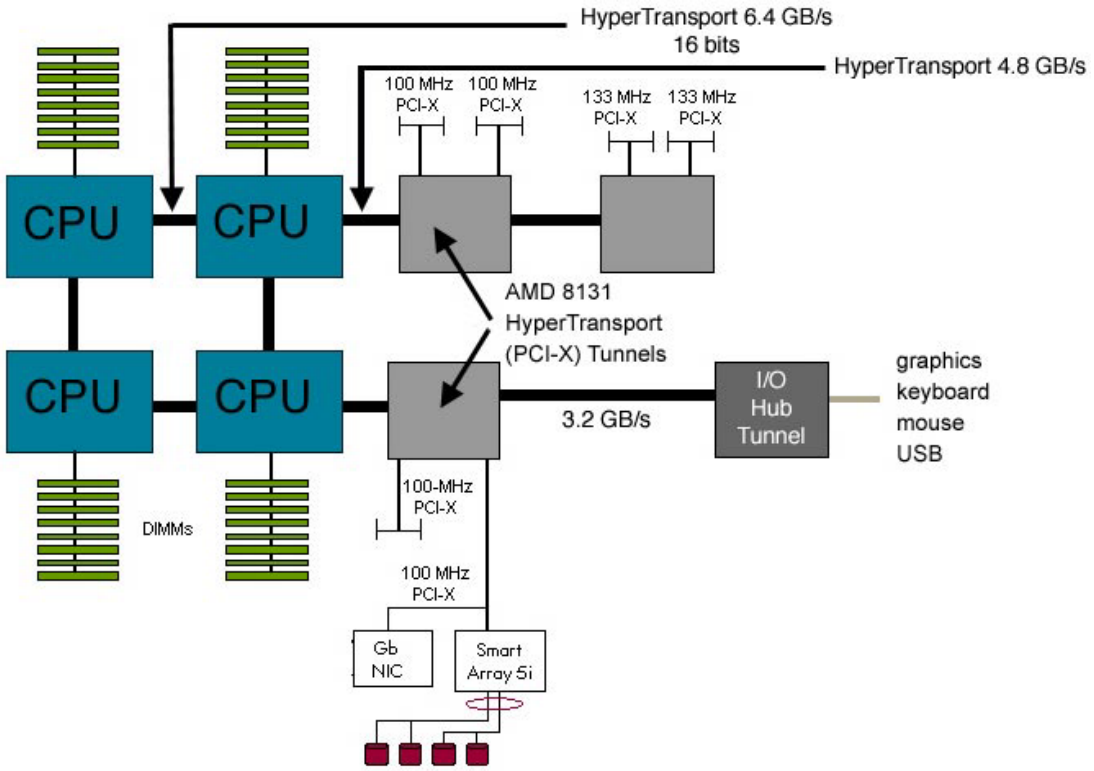
The Opteron processor includes three 16-bit-wide HyperTransport links (Figure 11). HyperTransport is a parallel point-to-point interconnect that replaces parallel front-side bus technology. It uses a double-data rate to transfer two bits of data per clock cycle on a 16-bit link. The HyperTransport clock can vary from 200 MHz up to 800 MHz, providing an effective operating frequency of up to 1600 MT/s (megatransfers per second), and therefore an effective transfer rate of up to 3.2 GB/s in each direction. Since transfers can occur in both directions simultaneously, an aggregate transfer rate of 6.4 GB/s can be achieved in a 16-bit HyperTransport I/O Link. Compared to a shared or bi-directional bus, a point-to-point interconnect has the advantage of no overhead for bus arbitration and easier maintenance of signal integrity. The HyperTransport technology uses low-voltage differential signaling technology to reduce power consumption and to minimize crosstalk and electromagnetic interference.

Figure 11. Schematic showing how Opteron includes HyperTransport technology



The Opteron processor can use the three HyperTransport links to connect to other Opteron processors or to I/O (Figure 12). The primary difference between the 100, 200, and 800 series Opteron processors is the way the processors use the three HyperTransport links. In the 100 series, the three HyperTransport links can only be used to connect to I/O in a non-coherent link. This means that the 100 series Opteron processors are limited to single-processor systems only. In the 200 series, one of the HyperTransport links can be used to connect to one other Opteron processor in a coherent link. The other links can be used to connect to I/O (non-coherent link), thus allowing 200 series Opteron processors to be used in dual-processor systems. With the 800 series Opteron processors, all three HyperTransport links can be used to connect to other Opteron processors or to I/O. This design enables an extremely scalable architecture, so that Opteron processors can be configured easily into 2-way, 4-way, and even 8-way systems.

Figure 12. Example of a 4-way system architecture using Opteron and the AMD 8000 chipset



For example, the ProLiant DL 585 server[18] uses the configuration shown in Figure 12, with the AMD 8131 chip as the HyperTransport Tunnel and the AMD 8111 chip as the I/O Hub. The HyperTransport functionality in these components and the processors provides 4.8-GB/s bandwidth to the HyperTransport Tunnels that support the PCI-X functionality.

The AMD 8131 HyperTransport tunnel provides 3.2-GB/s bandwidth into the I/O subsystem that contains the graphics and storage devices. Even the smaller bandwidth to the I/O hub provides adequate headroom for future expansion and high I/O throughput (Table 5).

Table 5. Bandwidth comparison between HyperTransport and existing I/O protocols

Protocol	Bandwidth	HyperTransport (8-bit, 3.2 GB/s) is faster by:	HyperTransport (16-bit, 6.4 GB/s) is faster by:
legacy PCI (32 bit, 33 MHz)	133 MB/s	24X	48X
USB 2.0	480 MB/s	6.6X	13X
PCI-X (64 bit, 133 MHz)	1064 MB/s	3X	6X
Infiniband 4X link	10 Gb/s (1.25 GB/s)	2.5X	5X
10 Gb Ethernet	10 Gb/s (1.25 GB/s)	2.5X	5X

[18] For more information about the ProLiant DL585, see the technology brief titled "HP ProLiant DL585 Server Technology," available at http://h200005.www2.hp.com/bc/docs/support/SupportManual/c00180597/c00180597.pdf

Comparing 32-bit performance

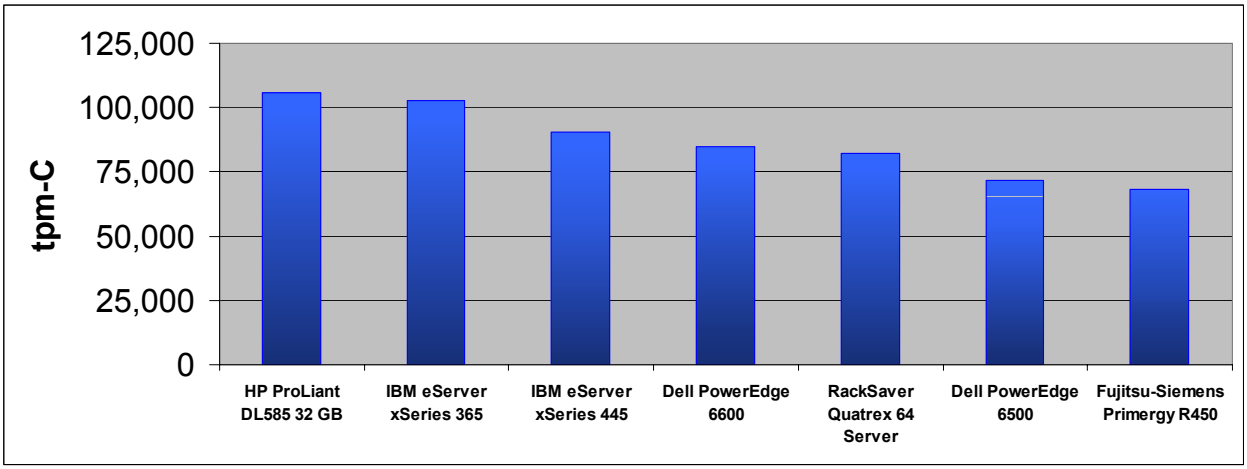
The difficulty in comparing architectures is that there are no absolutes about which processor will have better performance in a given system or application. For example, as already discussed in the previous sections, processor performance is a function of core processor frequency, pipeline design, branch prediction capabilities, cache sizes, and the interface to memory and I/O.

In general, HP believes that systems using the Opteron processor can provide performance improvements when compared to systems using the Xeon family of processors. This is especially true for applications in which there are frequent memory accesses that take advantage of Opteron's integrated memory controller. The performance differences are found in almost all cases in 4-way systems; performance in 2-way systems is much more dependent on the application used.

TPC-C benchmark

Results from the May 2004 TPC-C benchmark show that the HP ProLiant DL585 server, which uses four Opteron processors, took the #1 position in the tpmC rankings for x86 systems (Figure 13). The top result of 105,687 tpm was significantly higher than that of the highest-ranking server using Xeon processors. A second TPC-C benchmark has been delivered using 64 GB of memory, resulting in a 20 percent performance advantage over the Xeon system; however, publication of that benchmark is awaiting production availability of the updated operating system. As already noted, server performance is about more than looking at the speed of the processor (or cache size, or some other individual number on a processor). For example, the RackSaver QuatreX-64 system shown in Figure 13 also uses four Opteron processors. However, because of differences in system design, the RackSaver system has not approached the performance of the ProLiant DL585. Appendix A gives full details of the HP ProLiant DL585 system configuration. Additional information about the TPC benchmark and other systems is available on the TPC website at www.tpc.org.

Figure 13. TPC-C 4-way, 32-bit computing x86 performance



MMB3 benchmark

In July 2004, HP used a four-way ProLiant DL585 server (2.4 GHz/1 MB) to produce a Microsoft® Exchange MAPI Messaging Benchmark (MMB3) result of 9,000 MMB3. This is the highest Microsoft Exchange 2003 performance currently available using Microsoft Windows Server 2003 on a 4-way system.

The MMB3 benchmarking workload and methodology serves as the standard for Exchange 2003 MAPI server comparison. The MMB3 workload is characteristic of a medium corporate mail environment.

The ProLiant DL585 performance parameters included:

- CPU utilization rate of 85 percent on average
- Weighted 95th percentile response-time score of 653 milliseconds
- Average send-queue size of 97 messages for the 4-hour steady-state period

Additional information is available from the white paper titled "HP ProLiant DL585 achieves world-class 4-way, x86 performance results on Microsoft Exchange MAPI Messaging Benchmark 2003 ,"[19] and also from the Microsoft website at www.microsoft.com/exchange/evaluation/performance/default.asp.

SAP benchmark

HP also reported outstanding results on the SAP Sales and Distribution (SD) Standard Application benchmark using a ProLiant DL585 server. As stated on the SAP website, the SAP benchmarks "test and prove the scalability of mySAP Business Suite. The benchmark results provide basic sizing recommendations for customers by testing new hardware, system software components, and Relational Database Management Systems (RDBMS). They also allow for comparison of different system configurations."

Table 6 shows a comparison between 4-way servers using the Windows Server 2003 Enterprise Edition operating system. The ProLiant DL585 with Opteron processors has the #1 and #3 performance for a 4-way system in a 32-bit architecture.

[19] Available at ftp://ftp.compaq.com/pub/products/servers/benchmarks/dl585-mapi.pdf

Table 6. Configuration of SAP R/3 Enterprise 4.7 2-tier platforms

4P Platform	# of tested SD benchmark users	Avg. dialog response time/sec	Dialog steps/hr	OS and DB	SAPS	Fully processed line items/hr	Cert#
HP ProLiant DL585, Opteron 800 series 2.4GHz, 1MB L2 cache, 12GB RAM	**770**	**1.95**	**232,000**	**Microsoft Windows Server 2003 EE, SQL Server 2000**	**3870**	**77,300**	**2004024**
IBM eServer xSeries 365 8862RX, Intel Xeon™ MP 3 GHz, 512KB L2 /4MB L3 cache, 8GB RAM	720	1.93	217,000	Microsoft Windows Server 2003 EE, DB2 UDB 8.1	3620	72,330	2004019
HP ProLiant DL585, Opteron 800 series 2.2GHz, 1MB L2 cache, 16GB RAM	**712**	**1.95**	**214,000**	**Microsoft Windows Server 2003 EE, SQL Server 2000**	**3570**	**71,300**	**2004018**
IBM eServer xSeries 365 8862-3RX, Intel Xeon™ MP 2.8GHz, 512KB L2 /2MB L3 cache, 8GB RAM	562	1.98	169,000	Microsoft Windows Server 2003 EE, DB2 UDB 8.1	2820	56,330	2004003
IBM eServer xSeries 445 8870-4RX, Intel Xeon MP 2.8GHz, 512KB L2/2MB L3 cache, 8GB RAM	525	1.97	158,000	Microsoft Windows Server 2003 EE, DB2 UDB 8.1	2630	52,670	2003035
Dell PowerEdge 6650, SMP, Intel Xeon MP 2.8GHz, 512 KB L2/ 2MB L3 cache, 8GB RAM	420	1.91	127,000	Microsoft Windows Server 2003 EE, SQL Server 2000	2120	42,330	2003055

More information is available from the paper titled "New HP ProLiant DL585 achieves four-way Windows leadership performance result" [20] and the SAP website at www.sap.com/benchmark/.

[20] Available at ftp://ftp.compaq.com/pub/products/servers/benchmarks/dl585-sap2.pdf

64-bit Itanium processor architecture

The focus of this paper is the x86 architecture of the Xeon family and Opteron processors. However, because they support 64-bit extensions, it makes sense to briefly discuss the other 64-bit industry-standard processor, the Itanium. A detailed comparison of the Itanium processor family (IPF) architecture to the x86 architecture is beyond the scope of this paper. For more detailed information about the Itanium and Itanium 2 architecture, see the HP website and the Intel website.[21]

Table 7 gives a general overview of the key features of the Itanium family of processors.

Table 7. Itanium and Itanium 2 processor specifications (maximum values given)

Processor	Itanium	Itanium 2
Core frequency	800 MHz	1.5 GHz
L1 cache, KB	16 K data 16K instruction	16 K data 16K instruction
L2 cache, KB	96	256
L3 cache, MB	4 (off chip)	6 (on chip)
FSB speed	266 MT/s, 64-bits wide	400 MT/s, 128-bits wide
FSbBandwidth	2.1 GB/s	6.4 GB/s
Pipeline stages	10	8
Total number of registers	328	328
Total number of execution units	13	16
Peak instructions per cycle	6	6

Addressing

The Itanium 2 processor expands on the already extremely large addressing capabilities in the Itanium processor. The Itanium 2 supports 64-bit virtual addressing, which allows programs to access up to 16 exabytes (EB) of virtual memory. It also supports 50-bit physical addressing, which allows access for up to 1 petabyte (PB) of physical memory

Instruction set and registers

The Itanium uses an entirely different instruction set architecture than the Intel EMT64 or AMD64. HP and Intel jointly developed the Itanium with Explicitly Parallel Instruction Computing (EPIC). The EPIC approach relies on the compiler to analyze the program source code and embed explicit information in the instruction stream to tell the processor which instructions can be safely executed in parallel for increased performance. The x86 approach requires that the processor decide on the fly what can be executed in parallel. The x86 processors use dedicated hardware for this task that complicates the design and therefore slows performance.

The EPIC design includes the compatibility for IA-32 tacked on as an emulation capability. Typically, 32-bit performance on Itanium lags current x86 architectures; EPIC is designed to run 64-bit operating systems and applications. Recent improvements by Intel include the IA-32 Execution Layer (IA-32 EL).

[21] See the HP website at www.hp.com/products1/itanium/ and the Intel website at www.intel.com/products/server/processors/server/itanium2/index.htm?iid=sr+itanium&

According to Intel, IA-32 EL is "a new technology that executes IA-32 applications on Itanium processor family systems. Currently, support for IA-32 applications on Itanium-based platforms is achieved using hardware circuitry on the Itanium processors. This capability will be enhanced with IA-32 EL-software that will ship with Itanium-based operating systems and will convert IA-32 instructions into Itanium instructions via dynamic translation."[22]

Itanium 2 has eight, 64-bit GPRs and 256 data registers: 128 of these data registers are 64-bit integer; 128 are 82-bit floating point; and 72 data registers are related to branch prediction information. Table 8 summarizes the differences between the Itanium 2 registers and the registers available in IA-32 or EM64T/AMD64.

Table 8. Comparison of registers in x86 and IPF processors

Register size	32-bit x86 (Xeon or Opteron)	64-bit x86 (EM64T or AMD64)	IPF (Itanium 2)
1 bit	—	—	64 for predicate (branch prediction)
32 bit	8	—	—
64 bit	16 (MMX and x87)	32	136 (128 general purpose and 8 branch prediction)
82 bit	—	—	128 (floating point)
128 bit	8	16	—

System architecture

The IPF architecture uses an FSB architecture similar to the Xeon architecture. However, there are several important differences that improve the overall system performance:

- The FSB protocol is more efficient than the Xeon FSB protocol.
- The Itanium2 FSB is twice the width of the Xeon FSB.
- The Itanium processors use much larger cache sizes than Xeon, which reduces the need to access the FSB.

Summary

With the introduction of the ProLiant DL585 and DL145 servers, customers have the choice of ProLiant platforms using either Intel Xeon processors or AMD Opteron processors. Both are 32-bit processors that comply with the x86 architecture. This means that they have the same hardware/software interface at the programming level. In other words, they support the same instruction set, level of memory addressability, and the same underlying register set. For customers that want or need to move to 64-bit applications and operating systems, both the Xeon and Opteron processors enable a smooth transition with 64-bit extensions while remaining compatible with existing 32-bit applications. The IPF architecture, also industry-standard and 64-bits, gives excellent performance while using 64-bit applications. However, the emulation mode that Itanium and Itanium2 processors use to translate 32-bit code is not as efficient as running 32-bit code natively in an x86 processor. This gives the Xeon and Opteron processors a distinct advantage for customers who want to access the extra level of memory addressability available with these processors in 64-bit mode, without the cost and infrastructure changes required when moving to Itanium.

[22] "IA-32 Execution Layer: Technical Whitepaper," available from the Intel website at http://www.intel.com/cd/ids/developer/asmo-na/eng/microprocessors/itanium/93086.htm

The Xeon and Opteron processor architectures differ in their internal design details and platform implementations. The Xeon has a pipeline structure optimized for high clock speeds, while the Opteron is optimized for parallel execution. In the platform architecture, Xeon processors use the familiar front-side bus technology with a single north bridge that acts as a memory controller. The Opteron uses an integrated memory controller that runs at core processor speed. The memory interface enables large bandwidths and low latencies. Similarly, Opteron uses the HyperTransport point-to-point interconnect for the I/O subsystem, which provides high bandwidths. The use of the integrated memory controller and the point-to-point I/O connections appear to greatly improve performance, making the Opteron a higher-performing system than the Xeon in many cases, especially in 4-way systems. Recent testing shows that the 4-way ProLiant DL585 offers outstanding performance in typical benchmarks in transaction processing, messaging, and SAP database configurations.

Appendix A. ProLiant DL585 system configuration for TPC-C Benchmark

The following information is taken from the paper titled "New HP ProLiant DL585 with AMD Opteron processors achieves top 4-way 32-bit computing x86 TPC-C performance results," available on the HP website at http://h18004.www1.hp.com/products/servers/benchmarks/index.html.

In May 2004, the new HP ProLiant DL585 server posted leading 4-way 32-bit computing x86 results for the TPC-C benchmark on Microsoft Windows Server 2003 Enterprise Edition operating system, taking the #1 performance position. The current position was gained with 105,687 tpmC at $3.23/tpmC. A second TPC-C benchmark has been delivered using 64 GB of memory, resulting in a 20 percent performance advantage over the Xeon system; however, publication of that benchmark is awaiting production availability of the updated operating system.

Table A1. Top performing 4-way 32-bit computing x86 TPC-C system configurations

System configuration	TpmC	$/tpmC	Operating system & database	System Availability
HP ProLiant DL585 4P, AMD Opteron Model 848, (2.2 GHz), 1MB L2 cache, 32 GB RAM	**105,687**	**$3.23**	**Microsoft Windows Server 2003 EE, Microsoft SQL Server 2000 EE SP3**	**05-03-04**
IBM eServer xSeries 365 c/s 4P, 3.0 GHz, 4 MB L2 cache, 32 GB RAM	102,667	$3.52	Microsoft Windows Server 2003 EE w QFE, Microsoft SQL Server 2000 EE SP3 w QFE	03-31-04
IBM eServer xSeries 445 4P, 2.8 GHz, 2 MB L2 cache, 32 GB RAM	90,271	$3.97	Microsoft Windows Server 2003 Enterprise Edition, Microsoft SQL Server 2000 Enterprise Ed. SP3 w QFE	12-31-03
Dell PowerEdge 6600 4P, 2.8 GHz, 2 MB L2 cache, 32 GB RAM	84,595	$3.58	Microsoft Windows2003 Enterprise Server, Microsoft SQL Server 2000 SE	12-30-03
RackSaver Quatrex 64 Server, 4P AMD Opteron 844 1.8 GHz with 1MB L2 Cache, 32 GB RAM	82,226	$2.72	Microsoft Windows SQL Server 2000 Enterprise Edition SP3, Windows Server 2003 Enterprise Edition, Windows 2000 Server	10-21-03
Dell PowerEdge 6650, 4P Intel Xeon MP 2.0 GHz with 2MB L2 Cache, 32 GB RAM	71,586	$5.10	Microsoft Windows SQL Server 2000 Enterprise Edition SP3, Windows Server 2003 Enterprise Edition	03-31-03
Fujitsu-Siemens Primergy R450 c/s 4P Intel Xeon MP2.0 GHz with 2MB iL3 Cache, 16 GB RAM	68,264	*	Microsoft Windows SQL Server 2000 Enterprise Edition SP2, Microsoft Windows 2000 Advanced Server SP	05-03-03

*Results for the R450 are in Eurodollars and cannot be directly compared per TPC rules.

Appendix B. Engineering prefixes

Table B1

	Abbreviation	Exponential form	Number of bytes	Relationship to next lowest prefix
Gigabyte	[G/GB]	2^{30} bytes	1,073,741,824 bytes	1024 Megabytes
Terabyte	[T/TB]	2^{40} bytes	1,099,511,627,776 bytes	1024 Gigabytes
Petabyte	[P/PB]	2^{50} bytes	1,125,899,906,842,624 bytes	1024 Terabytes
Exabyte	[E/EB]	2^{60} bytes	1,152,921,504,606,846,976 bytes	1024 Petabytes

For more information

For additional information, refer to the resources listed below.

Resource description	Web address
AMD website	www.amd.com/
Opteron technical documents	www.amd.com/us-en/Processors/DevelopWithAMD/0,,30_2252_739_9003,00.html
AMD64 technical documents	www.amd.com/us-en/Processors/DevelopWithAMD/0,,30_2252_739_7044,00.html
Anandtech website Contains information about general processor architecture, Intel architecture, and AMD architecture	www.anandtech.com/
Ars Technica website: Contains information about general processor architecture, Intel architecture, and AMD architecture	www.arstechnica.com
Hyper-Threading Information (from Intel)	www.intel.com/technology/hyperthread
Hyper-Threading Information (from HP)	
Webpage	http://h18004.www1.hp.com/products/servers/technology/hyper-threading.html
Technology Brief, "Intel Hyper-Threading Technology," March 2003, TC030306TB	ftp://ftp.compaq.com/pub/supportinformation/papers/c030306tb_rev0_us.pdf

Resource description	Web address
HP website	
ProLiant DL servers	http://h18004.www1.hp.com/products/servers/platforms/index-dl.html
Itanium servers	www.hp.com/products1/itanium/infolibrary/
Industry standard server technology papers	www.hp.com/servers/technology
"Intel processor roadmap for industry-standard servers," technology brief	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00164255/c00164255.pdf
Intel	www.intel.com/

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

TC040902TB, 09/2004