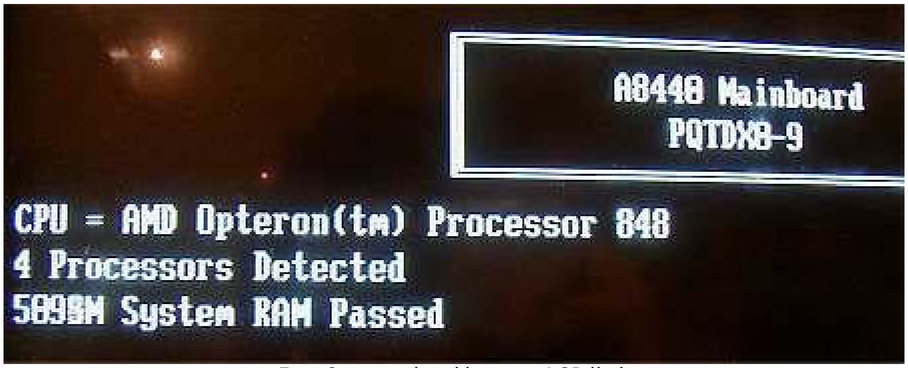
Ace's Server Guide: Dual Xeon, Dual Opteron, and Quad Opteron

By Johan De Gelas – December 2003

Introduction

In **our last article about x86 server CPUs**, the Opteron 244 made a quite an impression when compared to the Xeon DP. In many of our real-world server benchmarks, the 1.8 GHz Opteron bested the 2.8 GHz Xeon DP. Of course, neither Intel nor AMD have been sitting still since that time, and to make the Xeon DP more competitive, Intel equipped its midrange server CPU with a 1 MB L3-cache. Correspondingly, 2.2 GHz Opteron 248 and 848 processors are available this month. This article serves as an update on how these server platforms compare, but it's also much more. We've benchmarked AMD's Quartet, a four-headed Opteron hydra with up to 8.8 GHz (4 x 2.2 GHz) of processing power. Unfortunately, the quad Xeon MP was not yet able to join the party, but rest assured that this is not our last server CPU article.



Four Opterons breaking past 4 GB limit

So, let's get down to business. The following is what you may expect from this article:

- Quad Opteron, Dual Opteron and Dual Xeon benchmarks on a real-world dynamic Java Application Server
- ApacheBench comparison: 64-bit versus 32-bit, NUMA versus non-NUMA
- Datamining MySQL Benchmarks
- Professional 3DSMax Rendering
- Quick comparison with the Apple's Dual G5 (Cinebench 2003)
- Quick comparison with a 1.3 GHz Itanium and a SGI Origin 3800 supercomputer rack (Chess - AI benchmarking)

The prime objective is to show you what these systems are really capable of. All our benchmarks are based on applications that are used in the real-world. Manufacturers did not get a chance to optimize the machine, the compiler or setup for our benchmarks...

Opteron Platform: Very Promising

AMD's Athlon MP has never impressed in the server market. The CPU was too fragile, lacked several important RAS features (thermal protection, ECC on L1) and didn't get any support from the Tier One OEMs. The Opteron **does not have any of these disadvantages**.

Data Protection	Opteron	Athlon MP	Xeon
L1-data Cache	ECC	Parity	ECC
L2-Cache	ECC	ECC	ECC
Memory Controller	ECC + Chipkill	n/a	ECC + Chipkill
Thermal Protection	Opteron	Athlon MP	Xeon
Heatspreader	yes	no	yes
Shutdown on overheat (reaction time)	On-die thermal diode (slow)	On-die thermal diode (slow)	Clock Throttling (fast)
Temperature damage prevention	Thermtrip (fast)	no	Clock Throttling (fast)

Indeed, since the launch back in April, the Opteron is gaining mindshare, slowly but surely. IBM markets the eServer 325 only to the HPC market, but the 1U Dual Opteron server could easily target other markets as well. The lack of solid ISV support and certification for running software on Opteron Servers, although the Opteron has no trouble pretending to be a Xeon, is probably the reason why IBM is targeting solely the HPC market for the moment. Still, support for the Opteron is growing steadily with the biggest development thus far coming in the form of a **strategic alliance with Sun** that involves plans for Opteron servers with 2, 4, and more processors.

Of course it's also pretty easy to build your own single or dual-processor Opteron server. MSI, Gigabyte, IWill, Newisys and Tyan offer dual Opteron boards. MSI even has three different boards for the Dual Opteron: two based on the AMD 8111 + 8131 chipset, and one based on the VIA K8T800.

Tyan **goes a step further** with four dual Opteron motherboards, two single-processor boards, and one quad board. Newisys is the other manufacturer that has designed a Quad board for the Opteron. Leadtek and ASUS have so far only disclosed single CPU boards. Nevertheless, Gigabyte and MSI are ranked number three and four in the motherboard market, and Tyan is a big player in the market of multi-processor motherboards. Motherboard support for the Opteron is nothing short of excellent.

According to IDC, 10,746 Opteron servers have already been shipped in the third quarter. Considering that Opteron servers have only been available for a few months and AMD is not well known in this market, the Opteron is off to a very good start.

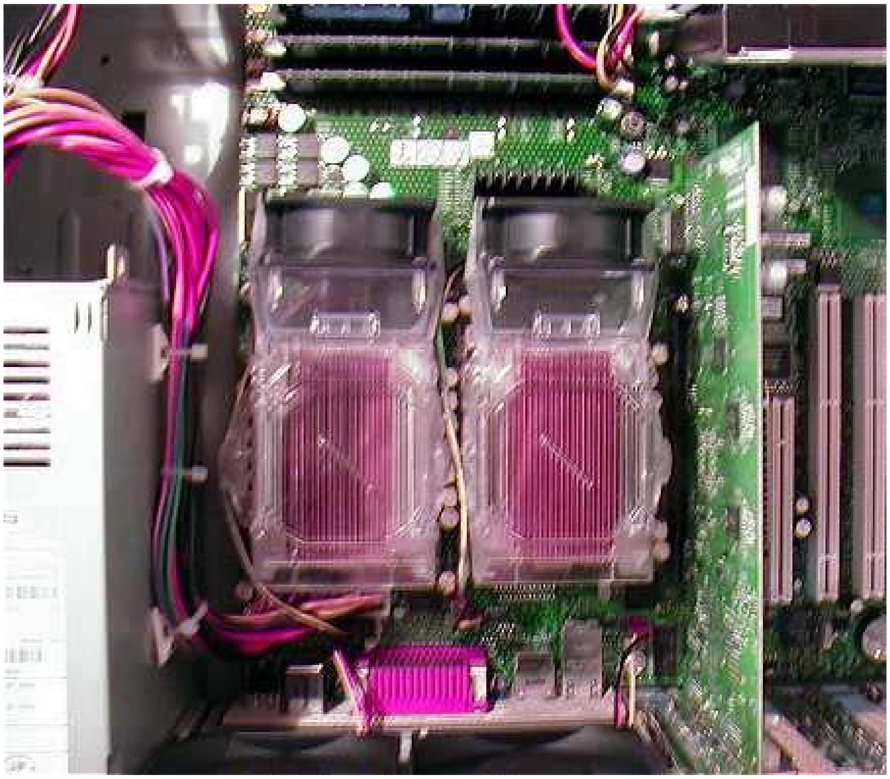
Meet the Improved Xeon DP

In the spirit of "only the paranoid survive," Intel anticipated the clockspeed ramp of the Opteron. Major price cuts in October 2003, higher clockspeeds and extra cache in the form of a 1MB on-die L3 cache should make the Pentium 4 Xeon a lot more competitive than it was at the time of the Opteron launch. However, the benefit of the large L3-cache is limited as you will see in our benchmarks.

One of reasons for this is that **Xeon DP's 1 MB L3-cache** is mostly inclusive. This means that most lines in cachelines of the L2 are duplicated in the L3 but not always. Eric Bron pointed out that there might be several cases where the cachelines are not duplicated:

- Data with high access frequency : kept in the L2 by the LRU but evicted from L3 to serve other misses (as seen from the L3 these data have no recent use)
- Data prefetched explicitly with the PREFETCHNTA instruction : 1/8 way of the L2 is dedicated to handle such "non-temporal" data, it is possible that the L3 is bypassed when they are read (like the bypassing the L2 on the PIII)
- Hardware prefetcher could fetch directly to L2

It might seem strange that the 512 KB L2-cache is almost always fully mirrored in the L3-cache and that this renders about half of the L3-cache "useless," but it simplifies the cache controller design and reduces latency on average. Intel claims a latency of 14 cycles for the L3-cache, in addition to the 7 and 2 cycles of the L2 and L1 caches, respectively. However, it must be said that in the case of the L2-cache, a total of 9 cycles is the best case and we assume that that 23 (2 + 7 + 14) cycles for the L3-cache is also a best case number. We measured a 30 cycle latency with ScienceMark.



The cache designs of the Xeon MP and DP are identical: both Level 3 cache designs are full speed, 8-way associative with ECC capability. It is interesting to point out that the Xeon DP is moving ever closer to the Xeon MP from an architectural standpoint. The latter is much more expensive because it can be used in quad configurations, but that advantage is negated somewhat by the higher clockspeeds at which the Xeon DP is available. With the Itanium moving in from above, the Xeon MP is increasingly relegated to a market niche. Who wants to pay $8000 for four 2.5 GHz Xeon MPs (1 MB L3) when you can get two powerful 3.2 GHz Xeon DPs with 1 MB L3 for $1900? Of course, we'll save the final judgment on the Xeon MP when it is available in our labs... But right now, the Xeon DP with 1 MB L3 seems like a very good deal.

Quad Opteron Power

At four rack-units high and 25 inches deep, the quad Opteron server, called "Quartet", is an impressive beast. Besides having four nodes of CPU and memory connected to each other, it also comes with all the features that differentiate a heavy duty server from the rest of the pack. The 4 CPUs are cooled by 8 fast, 120 mm hot-plug fans, for example. The 4 SCSI drives and 2 133 MHz PCI-X slots are also hot-pluggable, as are the 2+1 redundant 500 Watt power supplies (and a third can be added for redundancy).

This is not a "only for reviewers" system, it is the basically **the same machine that is being sold by seven vendors**: Appro, CCSI, Colfax Int, CSI Labs, MicroWay, Penguin Computing, and Racksaver. A system with four 844 CPUs (best Opteron price/performance wise), three 10,000 RPM SCSI disks in RAID-5 and 4 GB of memory, you run you around $14,000.

The Quartet server - A view of the four CPU nodes which are connected by 16/16-bit HT links

Some of the features of the Quartet platform are as follows:

- Four AMD Opteron processors
- Each processor supports four registered DIMM slots of memory.
- DRAM controllers support ECC, bank, and node interleaving.
- Three 64/66-MHz PCI-X slots
- Two 64/133-MHz hot-plug PCI-X slots
- Four hot-swap SCSI drives supported by an SAF-TE backplane
- Two Ultra320 SCSI ports—one internal and one routed through the back panel
- Three 500-W hot-swap power supplies in a 2+1 configuration
- Two copper gigabit Ethernet ports
- One 10/100 Ethernet port from the AMD-8111™ HyperTransport™ I/O hub
- Dedicated management LAN port
- PS/2 mouse and keyboard port out the front panel
- Optional graphics out the front panel (through small form factor PCI)
- Single USB 1.1 out the front panel
- Slimline CD-ROM drive and floppy disk drive
- Legacy serial port
- IPMI-based server management

The Security features:

- Power-on password
- Remote control security settings
- Selectable drive startup
- System management security
- User login password
- Administrator password
- Read-only or read/write access
- Diskette boot override and control, write control
- Configuration lock

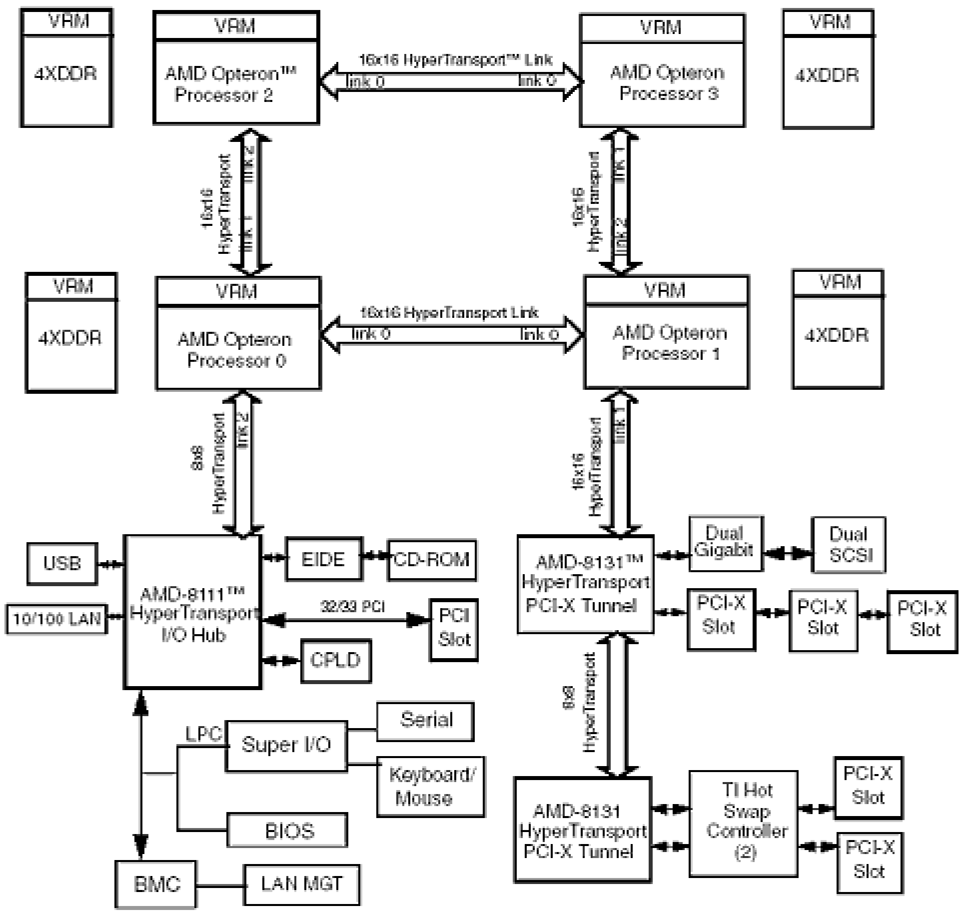
As the Opteron requires two 64-bit DIMMs to complete a single 128-bit memory interface, there are of course a few restrictions on how you may populate the DIMM sockets:

- A module placed in the DIMM3 socket is assumed to be identical to the one placed in DIMM2. The memory controller creates 128-bit banks from these adjacent pairs.
- If a given node is to receive two DIMMs that are not identical, then they should be placed in DIMM0 and DIMM2. This creates two 64-bit wide banks that cannot be formed into a single 128-bit bank.
- It is illegal to install only three DIMM slots on a given node. Memory controllers can be in either 128-bit or 64-bit mode, not both.

With 16 DIMM slots, you can have up to 32 GB of memory....

Overview of the Quartet System

Below you can see a block diagram of the Quad Opteron that provides you with a good idea of the system configuration at a glance.



At first I wanted to test the performance scaling of the Opteron, by testing with one, two, and four CPUs. This diagram makes it clear, however, that this is not possible: if you do not populate the second socket (CPU 1), you can not use the Gigabit ethernet, PCI-X, and SCSI ports. You have to use at least two CPUs to get the machine working. Of course, nobody is going to invest in such a server and use only one CPU.

Notice also that each CPU has its own Voltage Regulator Module (VRM), a guarantee for ripple free DC current to the CPUs and DIMMs.

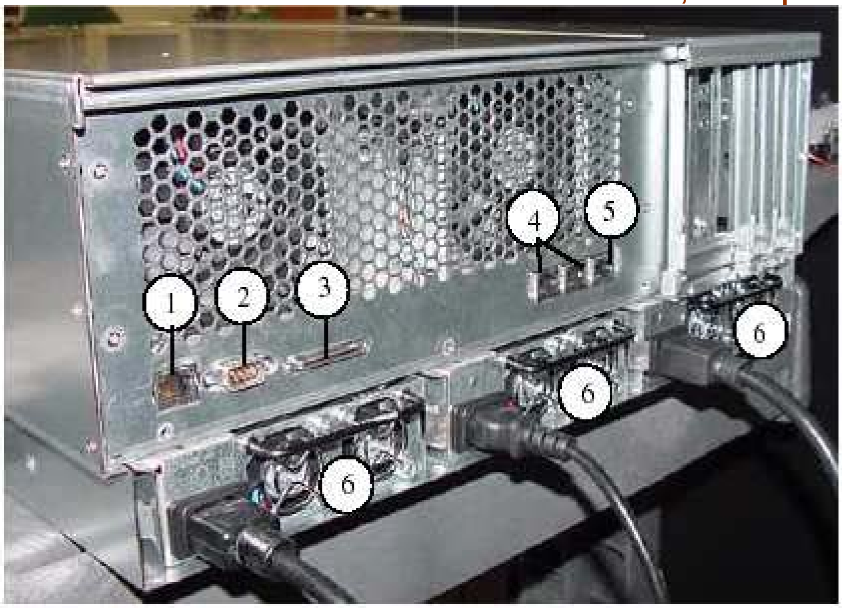
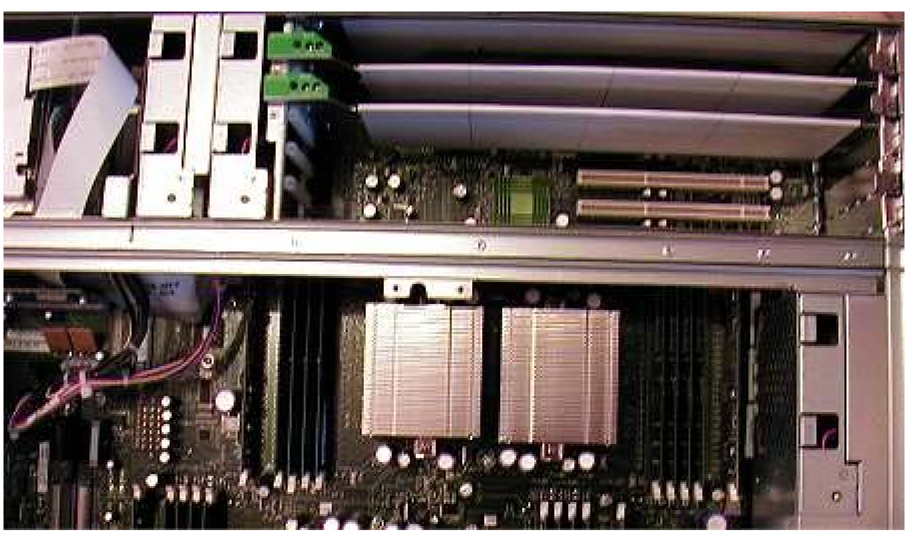


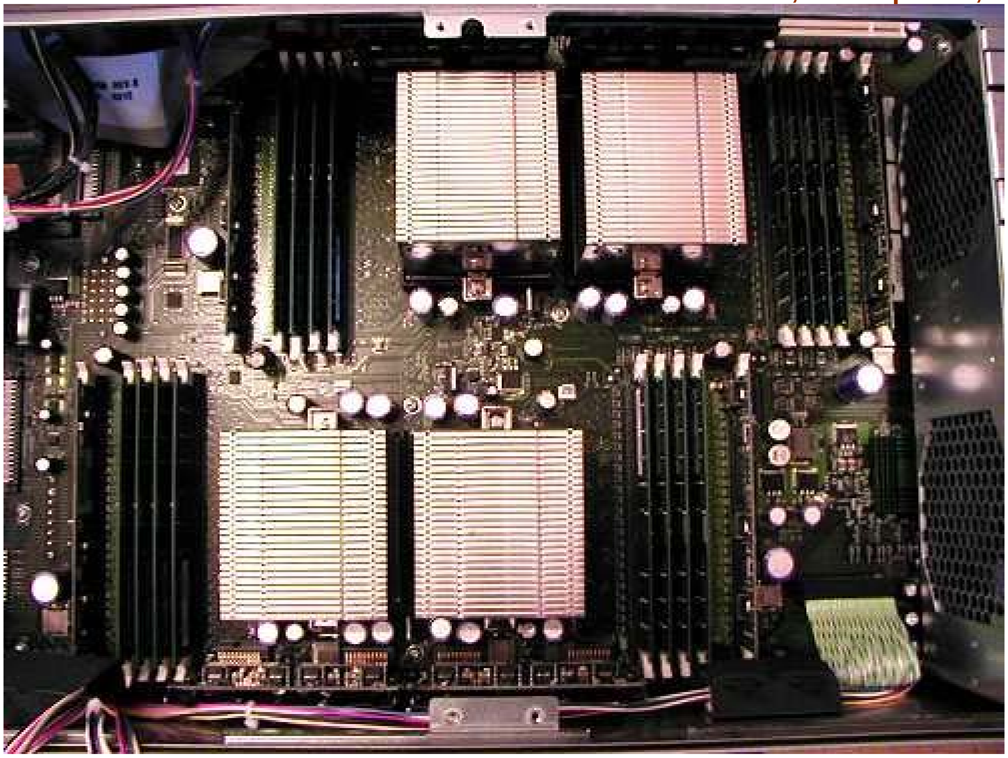
Figure 3. Location of Major Quartet Platform Components (Back Panel)

No.	Component
1	BMC manager port
2	Serial port
3	External Ultra320 SCSI port
4	Gigabit Ethernet port
5	10/100 Ethernet port
6	2+1 redundant hot-swap power supplies

Three 64/66-MHz PCI-X slots and two 64/133-MHz hot-plug PCI-X slots are available. You can recognize the hot plug PCI-X card as a diagnostic "Pearl board" is visible in the picture. Each Pearl board is attached to a cable assembly that connects to a switch mounted on the back of the chassis in the PCI area. The switch indicates if a PCI card is populated in one of the hot-plug slots.



All CPUs were cooled with aluminum heatsinks which made excellent contact with the heat-spreader as the heatsinks were torqued tightly to the chassis and motherboard.



Above is a view of the four CPUs with heatsinks, below we can see the four CPUs without the heatsinks:

Note the two VRMs for CPU2 and 3 in the background of the picture. There is also a VRM for every 4 DIMMs.

Processors and heatsinks are covered in pairs by aluminum shrouds with two high-speed 120 mm fans. In the picture above I have lifted the shroud and fans.

Intel/AMD Processor Comparison

For your reference, you can find a table with the most important features of each CPU below. Before the Opteron can engage the markets where the Itanium is active, it needs a lot more OEM support, 64-bit applications and ISV support. But as the Opteron might be able to compete with Intel's Itanium in the future, the Itanium is included here.

Features	**Opteron 2xx**	**Opteron 8xx**	**Xeon DP 1 MB L3**	**Xeon MP**	**Itanium II "Mc Kinley"**	**Itanium II "Madison"**
Top clockspeed	2.2 GHz	2.2 GHz	3.2 GHz	2.8 GHz	1 GHz	1.5 GHz
process technology (µm)	0.13 SOI Cu	0.13 SOI Cu	0.13 Cu	0.13 Cu	0.18	0.13
Transistors (million)	105.9	105.9	143	190	221	410
voltage	1.5V	1.5V	1.55V	1.5V	?	?
die size (mm^2)	193	193	169	?	464/421	374
MP and Address Space	**Opteron**	**Opteron 8xx**	**Xeon**	**Xeon MP**	**Itanium II**	**Itanium II**
Typical Multi processor system	2	4-8	2	2-8 (up to 32)	2-8 (up to 64)	2-8 (up to 128)
Max. Physical Address Space	1024 GB flat (40 bit)	1024 GB flat (40 bit)	64 GB PSE (36 bit)	64 GB PSE (36 bit)	1024 TB (50 bit)	1024 TB (50 bit)
Max. Virtual Space	256 TB (48 bit)	256 TB (48 bit)	4 GB	4 GB	1024000 TB (60 bit)	1024000 TB (60 bit)
Cache configuration	**Opteron**	**Opteron 8xx**	**Xeon**	**Xeon MP**	**Itanium II**	**Itanium II**
L1-cache (Data/Instr)	64/64 KB	64/64 KB	8 KB/ +-20 KB**	8 KB/ +-20 KB**	16 KB/ 16KB	16 KB/ 16KB
L1-cache latency (load to use)	3	3	2	2	2	2
L2-cache	1 MB	1 MB	512 KB	512 KB	256 KB	256 KB
L2-cache Width	128 bit	128 bit	256 bit	256 bit	256 bit	256 bit
L2-cache Latency load to use (+L1-latency)	16	16	9-20	9-20	5	5
L3-cache	-	-	1 MB inclusive	1 - 2 MB inclusive	3 MB	6 MB
Memory	**Opteron 2xx**	**Opteron 8xx**	**Xeon**	**Xeon MP**	**Itanium II**	**Itanium II**
Memory configuration	2 x DDR400	2x DDR333	2 x DDR266	2xDDR200/266	4xDDR266	4x DDR266
Max. Memory Bandwidth to CPU	6.4 GB/s	5.4 GB/s	4.2 GB/s	3.2 GB/s	6.4 GB/s	6.4 GB/s

*** 12,000 Micro-ops, which is probably comparable to about 20 KB of x86 instruction cache*

Below you'll find the pricing of the different CPUs. We tried to compare the competing CPUs in the same performance league based on industry standard benchmarks such as SPECjbb and SPECint, but it is not scientific at all. There's a lot of guesswork right now, as we were not able to test systems based on the Xeon MP and Itanium CPUs. This table is just for your information, it is not meant to be precise.

Opteron CPU	Price	Xeon CPU	Price	Itanium	Price
244	$455	Xeon DP 2.8 GHz - no L3	$316		
246	$794	Xeon DP 3.06 GHz - 1 MB L3	$690		
248	$913	Xeon DP 3.2 GHz - 1 MB L3	$851		
842	$999			Itanium 2 1.4 GHz - 1.5 MB	$1172
844	$1299	Xeon MP 2.5 GHz - 1 MB L3	$1980	Itanium 2 1.5 GHz - 3 MB	$1338
846	$2149	Xeon MP 2.8 GHz - 2 MB L3	$3692	Itanium 2 1.4 GHz - 4 MB	$2247
848	$3199			Itanium 2 1.5 GHz - 6 MB	$4227

For example, the Opteron 244 might seem pricey compared to the 2.8 GHz Xeon, but our previous testing shows that the Opteron CPU outperforms the 2.8 GHz Xeon in MySQL. With the same reasoning, the Xeon DP is a bargain if you run MS SQL server, as it proved to be stronger than the 244 in this software.

It is pretty hard to compare the Itanium to the Opteron. Oracle and DB2 are available for the Opteron, but enterprises need full support and certification for an Opteron server. At this point support is still in the "beta" stage.

What we can say with more certainty is that you have to pay a big price premium for the fact that the Xeon MP platform is an established, proven and well supported platform with tons of certified (x86) applications. Both Intel and AMD reward the IT managers that dare to take a risk to move towards a more powerful, but less mature 64-bit platform.

Benchmarked Configurations: Hardware

We will discuss the software settings for each of the tests since we used different configurations for HPC, web, and database server tests. The desktop was set in Windows at a resolution of 1024x768x32bpp with at a 75 Hz refresh rate. In Linux, we used strictly the console.

A very special thanks to Shelley Baldiga (**Crucial**), who made this review possible by sending us 2x 512 MB of the best quality **Crucial PC2100R-ECC Buffered** RAM.

Server 1: AMD Quartet: Dual 844, Dual 848, Quad 844 and Quad 848

- Quartet motherboard, Zildjian personality board, Tobias backplane board and Rivera power distribution board.
- Quad configurations: 4 GB: 8x512 MB infineon PC2700 Registered, ECC enabled
- Dual configurations: 2 GB: 4x512 MB infineon PC2700 Registered, ECC enabled
- NIC: Broadcom NetExtreme Gigabit

Server 2: Dual Xeon DP 3.06 GHz 1 MB L3-cache

- Intel SE7505VB2 board - Dual DDR266
- 2 GB: 4x512 MB Crucial PC2100R - 250033R (2.5-3-3-6)
- NIC: 1 Gb Intel RC82540EM - Intel E1000 driver.

Server 3: Dual AMD Opteron 244 (1.8 GHz)

- Newisys Khepri
- 2 GB: 4x512 MB Infineon PC2700R - 250033R (2.5-3-3-6)
- NIC: Broadcom 5703, bcm5700 driver

Server 4: Dual Xeon DP 2.8 GHz - 533 MHz FSB

- Gigabyte GA-8IPXDR-E
- 2 GB 4x512 MB Crucial PC2100R - 250033R (2.5-3-3-6)
- NIC: 1 Gb Intel RC82540EM - Intel E1000 driver.

Client Configuration 1: 1x Pentium 4 3.06 GHz HT on - 533 MHz FSB

- MSI GNB MAX FISR (E7205)
- 2x256 MB Crucial PC2100R - 250033R (2-2-2-6)
- NIC: 1 Gb Intel RC82546EB - Intel E1000 driver.

Shared Components

- 3x Seagate 36 GB - 15000 rpm - 320 MB/s SCSI RAID 5
- Maxtor 80 GB DiamondMax 740X (7200 rpm, ATA-100/133)

Software

- Intel chipset inf update 5.09.1012

We'd like to thank the following helpful people for their support and important contributions to this review:

- Damon Muzny, Bill Robins (**AMD**)
- Mattty Bakkeren, Kristof Semhke (**Intel**)
- Nicole Chia (**Gigabyte**)
- Angelique Berden and Marga Zanders (**MSI**)
- Shelley Baldiga (**Crucial**)
- Robert Pearce (**Corsair**)

Benchmark Details and Notes

We know very well that it is not fair to directly compare a 4-processor Opteron system to a dual-processor Xeon system. As soon as we have access to a 4-way Xeon MP system, you'll see those benchmarks. Our main objective is to see how well the quad Opteron scales when compared to the dual Opteron. It also gives us some insight as to how the Xeon MP with 1 MB cache will perform: if the quad Opteron doubles the performance of the 3.06 GHz Xeon, it is pretty clear that a quad Xeon will not be able to match it. Doubling performance from 2 to 4 CPUs with a shared bus architecture is as good as impossible in a benchmark that does not run strictly within the CPU caches.

Additionally, you might be curious as to why we configured the 4-way systems with 4 GB and the dual systems with only 2 GB. The only reason was that the Opteron needs two DIMMs to access the RAM in 128-bit mode. Resultantly, each CPU needs a minimum of 2 DIMMs, and 512 MB ECC Registered DIMMs were available. As none of our tests demand more than 1.5 GB, the 4 GB in the Quad systems is no advantage.

We also included the results of the Newisys Opteron 244 Server we tested back in April, and compared it to our Quad machine running two 844 processors. This way we can see if the Opteron platform has advanced.

All Xeon benchmarks are done with Hyperthreading on, unless otherwise indicated. Yes, we did test with the fastest Opteron, but not the fastest Xeon DP, which is currently the 3.2 GHz with 1 MB L3-cache. As the latter is clocked only 4.5% faster than the CPU we had available in the lab, it is easy to add 4% to the Xeon results to get an idea how the 3.2 GHz Xeon performs.

ApacheBench

OS: Linux 2.4.20 / 2.4.19
Multi - threaded: Yes
Memory Usage: 180 - 200 MB
Typical Error Margin Between Different Runs: 4-7%
Maximum Network Traffic: N/A

ApacheBench is a simple tool for benchmarking a Apache or other HTTP servers. It can execute concurrent requests on a webserver and measure how many requests per second the webserver is capable of serving. The ApacheBench (ab) tool is bundled with the Apache source distribution.

We've tested the server using the first page of **"A Quick Look at the Fastest Apple PowerMac"** (18.8 KB).

ApacheBench: 64-bit versus 32-bit

Apache 1.3.26 has been compiled and ported to AMD64 on the SuSE SLES distribution. The 32-bit version is tested on 32-bit Linux, while the 64-bit version is tested on 64-bit Linux. First we test with 10 concurrent connections:

- Dual Opteron 844 (64-bit) NUMA support, 64-bit, 8240 req/s
- Dual Opteron 844 (64-bit) - 32-bit, 6921 req/s

Next we test with 100 concurrent requests:

- Dual Opteron 844 (64-bit) NUMA support, 64-bit, 6510 req/s
- Dual Opteron 844 (64-bit) - 32-bit, 5976 req/s

You have probably noticed that Apache is about 9 to 20% faster in 64-bit. However, our previous report which claims that the 64-bit binary is about 10% faster than the 32-bit version,is still more or less accurate.

Original ApacheBench Tests

As we were curious what the 64-bit mode can do for real applications, we tested the quad Opteron (still 4x 844 with DDR333) with Apache 1.3.26 32-bit and 64-bit on SuSE SLES8 AMD64.

Apache Bench: 10 concurrent requests, 100,000 total

- Quad Opteron 844 (32-bit): 8263 requests /s, 1.21 ms per 10 requests
- Quad Opteron 844 (64-bit): 9032 requests /s, 1.11 ms per 10 requests

Apache Bench: 100,000 requests, 100 concurrent requests

- Quad Opteron 844 (32-bit): 8554 requests /s, 11.69 ms per 100 requests
- Quad Opteron 844 (64-bit): 9369 requests /s, 10.67 ms per 100 requests

64-bit Apache is about 9.5% faster than its identically (as far as I could see) configured 32-bit brother.

The difference between our first tests and our new tests is that we enabled a special feature in the BIOS to make the 64-bit NUMA aware Linux more efficient.

ApacheBench: "64-bit NUMA Aware" Versus "64-bit Non-NUMA Aware"

When we updated the BIOS, the README that came with the new release pointed to an SRAT setting which could be enabled, but only for NUMA-aware Operating Systems such as Linux 2.4.20 AMD64 and Windows 2003 server.

> The Static Resource Affinity Table (SRAT) can be used to describe the physical location of processors and memory in large-scale systems (such as CC-NUMA) to the Microsoft® Windows® Server 2003 operating system, allowing threads and memory to be grouped in an optimal manner.

So it seems that this helps the OS to keep thread data local to the CPU, so that they are not located in memory on another CPU node. We test again with ApacheBench:

With 10 concurrent connections

- Dual Opteron 848 (64-bit) NUMA support, 9640 req/s
- Dual Opteron 848 (64-bit) - SRAT disabled, 8433 req/s

With 100 concurrent connections:

- Dual Opteron 848 (64-bit) NUMA support, 8307 req/s
- Dual Opteron 848 (64-bit) - SRAT disabled, 7059 req/s

ApacheBench: Opteron versus Xeon

Let's take a look at all the results. All Opteron results were obtained running 64-bit Linux, with the SRAT table enabled.

Benchmark	Quad Opteron 848	Quad Opteron 844	Dual Opteron 848	Dual Opteron 844	Dual Xeon 3.06 1 MB M3	Single Xeon 3.06 1 MB L3
10 Concurrent Connections (req/s)	12033	10888	8433	8240	4882	3073
10 Concurrent Connections (response time)	0.83	0.92	1.19	1.21	2.28	3.25
100 Concurrent Connections (req/s)	11550	9496.8	7059	6510	5118	3038
100 Concurrent Connections (response time)	8.66	10.53	14.1	15.5	19.54	32.9

Apachebench continues to be a very low level, rather synthetic benchmark. Although we have verified all benchmarks by running them at least 5 times, we might have overlooked a small kernel variable which could impact the results a bit. And ApacheBench tends to be rather inprecise. However, I doubt very much that the overall picture will change.

With the 20% extra performance from a 64-bit NUMA aware OS, the Opteron leaves the Xeon far behind: the Opteron 844 and 848 are 65 to 70% faster, add or substract 5-7%. Notice that the results with 100 concurrent connections tend to be more consistent. The differences between faster-clocked CPUs and between dual and quad configurations becomes more clear.

This is a shot across the bow for Intel: don't underestimate the Opteron, especially if it gets 64-bit ISV support. Even if we can add another 10% of performance to the Xeon, and ApacheBench reported Xeon performance 5% too low, it is clear that the Opteron is not within reach of the Xeon in Apache.

Luckily for Intel, this benchmark represents static HTML page serving. Very few people need this kind of performance to simply serve up HTML pages as fast as theses monster can. Let us see how the Xeon and Opteron perform on a highly interactive, dynamic and database-based Java webserver platform.

Benchmark Methods: Java Servlets and Java Server Pages

The webservers were tested in two configurations: with 6 clients (client 1 configuration: Duron 1300 PC) connected to one of the eight 100Mbit ports of our gigabit switch. This means that we should be able to push about 50 to 60 MB/s of client traffic to our server which is hooked up to the gigabit uplink.

Our second configuration consisted of one client (client 2: P4 3.06 GHz PC) connected on the integrated gigabit Intel RC82540EM via a UTP5 crossover cable to our servers and their gigabit connectors. In this configuration it was theoretically possible to push about 100 MB/s of client traffic to the server. This way we could see whether or not extra clients or higher network bandwidth could push our servers higher.

The 3.06 GHz Pentium 4 client was loaded with a Debian Linux kernel version 2.4.20-686-smp, which is a Pentium 4 Hyperthreading-optimized version.

For benchmarking, **httperf** was used in conjunction with **Autobench**, a Perl script written by Julian T. J. Midgley, designed to run httperf against a server several times, with the number of requests per second increasing with each iteration. The output from the program enables us to see exactly how well the system being tested performs as the workload is gradually increased until it becomes saturated.

Autobench and httperf in action

In each case, the server was benchmarked with 5 requests per connection. The benchmark was configured to timeout any request that takes longer than 5 seconds to complete, 10 seconds in case of the "compressed message board index." However we had to recompile httperf, as it was running on the clients with only 1024 file descriptors, an old UNIX per-process limitation. Our own httperf does not have that limitation.

The clients were therefore set to: ulimit -n 10000 (set number of open files to 10000) (default 1024)

The dual Xeon Debian Linux kernel version was 2.4.20-686-smp. We also tried out 2.4.22-1-686-smp, but we could not detect any difference. The Opteron used the SuSE SLES 8 kernel 2.4.19 - 32-bit, non NUMA aware. As these kernels are the most stable ones which are optimized for each of our server CPUs, we believe this is a fair comparison. The same kernel on different distribution will perform very similar if you tune the right parameters such as we have done.

The servers were tweaked with the following parameters:

- ulimit -n 64000, set number of open files to 64000 (default 1024)
- shmmax = 512288000 or 512 MB, shared memory maximum (default 33 MB)

When comparing our benchmark methods with others, you'll see that these 3 following parameters are turned off.

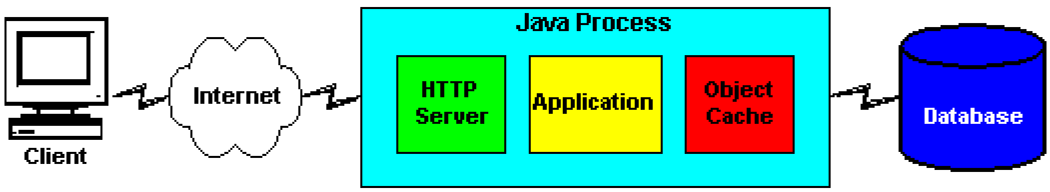
- net.ipv4.tcp_timestamps=1, turns TCP timestamp support on, (default)
- net.ipv4.tcp_sack=0, turn SACK support on, (default)
- net.ipv4.tcp_window_scaling=0, turn TCP window scaling support off, (default on)

We decided to leave them on as as these features allow a webserver to perform well on a high latency WAN. Yes, we were testing on low latency LAN, but webservers are mostly deployed on the Internet, one of the highest latency WANs, or as Intranet servers on very large corporate LANs.

The benchmarked software includes:

- **Caucho Technology's Resin 2.1.6**
- **Java HotSpot Server VM 1.4.2_02 SDK**
- **Sybase ASE 11.9.2 for Linux**

These benchmarks test the production Ace's Hardware application with the exact same dataset. The full platform is described in the following diagram:

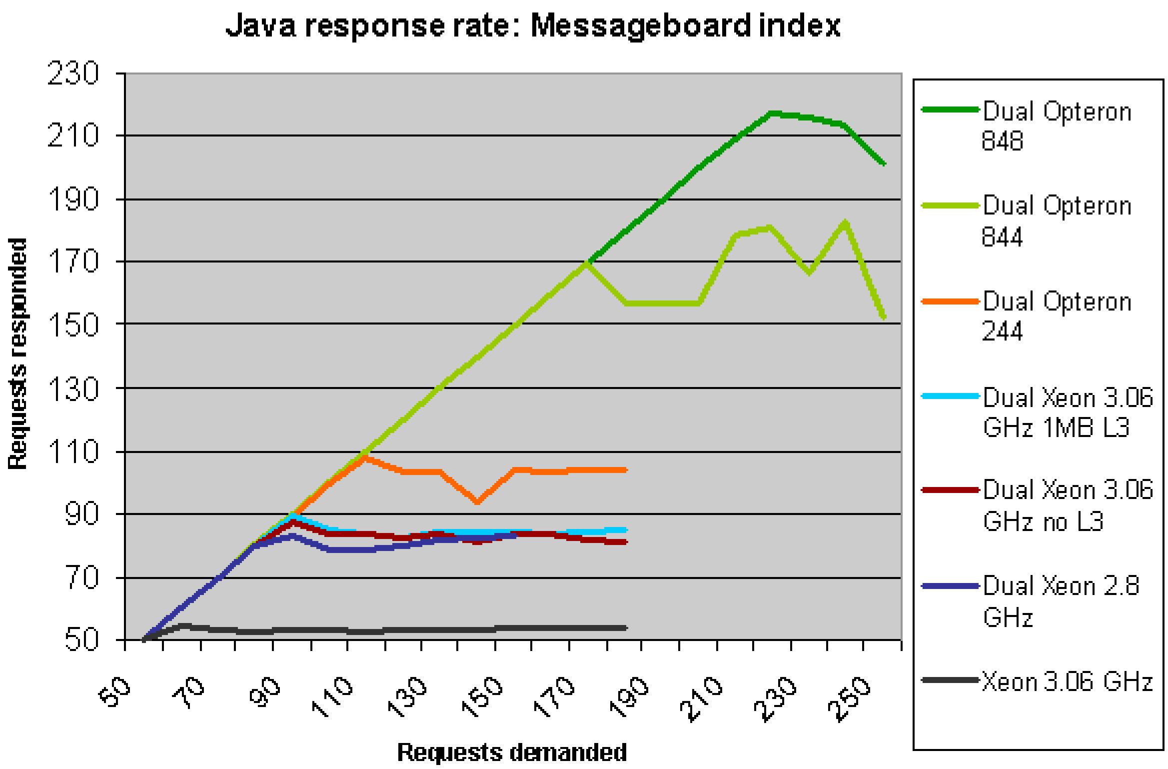


To understand this better, I recommend your read **Brian's article "Scaling Server Performance."** Basically, the clients run httperf, which generates a lot of "HTTP GET" requests. Looking at the diagram below, the "Internet" is in our case a crossover cable, or the gigabit switch. Resin or another webserver (green) accepts the requests and transmits the response back to the client. The Java application (yellow) fetches data from the object cache (red) and generates the response to be sent back to the client via the HTTP server. The cache consists of data contained in the database (blue).

Cached Java Webserver Benchmark: Message Board Index

OS: Linux 2.4.20 / 2.4.19
Multi-threaded: Yes
Memory Usage: 800 - 850 MB
Typical Error Margin Between Different Runs: 3-4%
Maximum Network Traffic: up to 44 MB/s

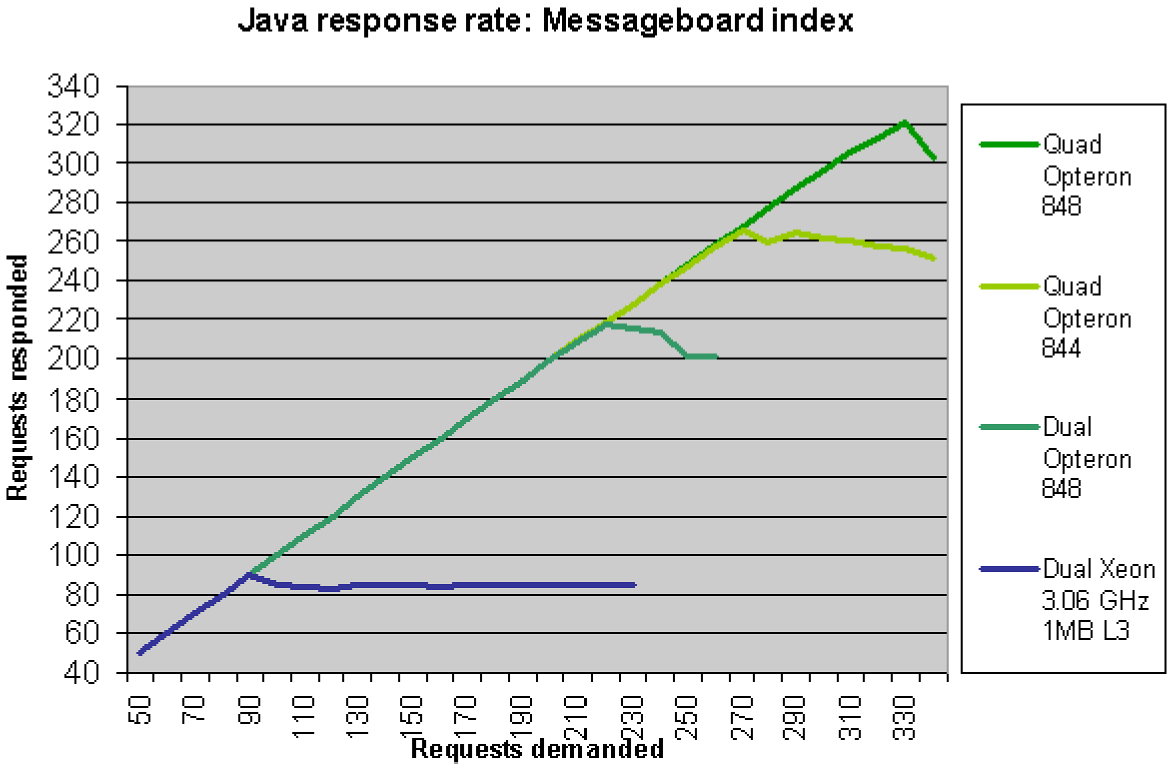
Our first test is very CPU intensive one, as it requests the main index of our **General Message Board**. In this benchmark, the index is a threaded tree structure consisting of approximately 1000 elements. As the content is not compressed, it also generates a lot of network traffic.



On the Quad Opteron, we were approaching the limits of our Gigabit ethernet card, which is nothing short of amazing. Theoretically, a gigabit card should be able to deliver around 100 MB/s of bandwidth, but we noticed that the real limit (well, measured with httperf) is somewhere between 45-55 MB/s. Notice the huge difference between the original Newisys 244 and the Dual Opteron 844 in our Quartet system. We have yet to pin down why the performance has improved so spectacularly. It might be a combination of improved AMD64 support in the 2.4.19 kernel and a better Broadcom Ethernet Driver (we believe this might be one of the prime reasons). Additionally, we noticed 20% better performance once we upgraded the BIOS from the original version to the Pqtdx0-9 version.

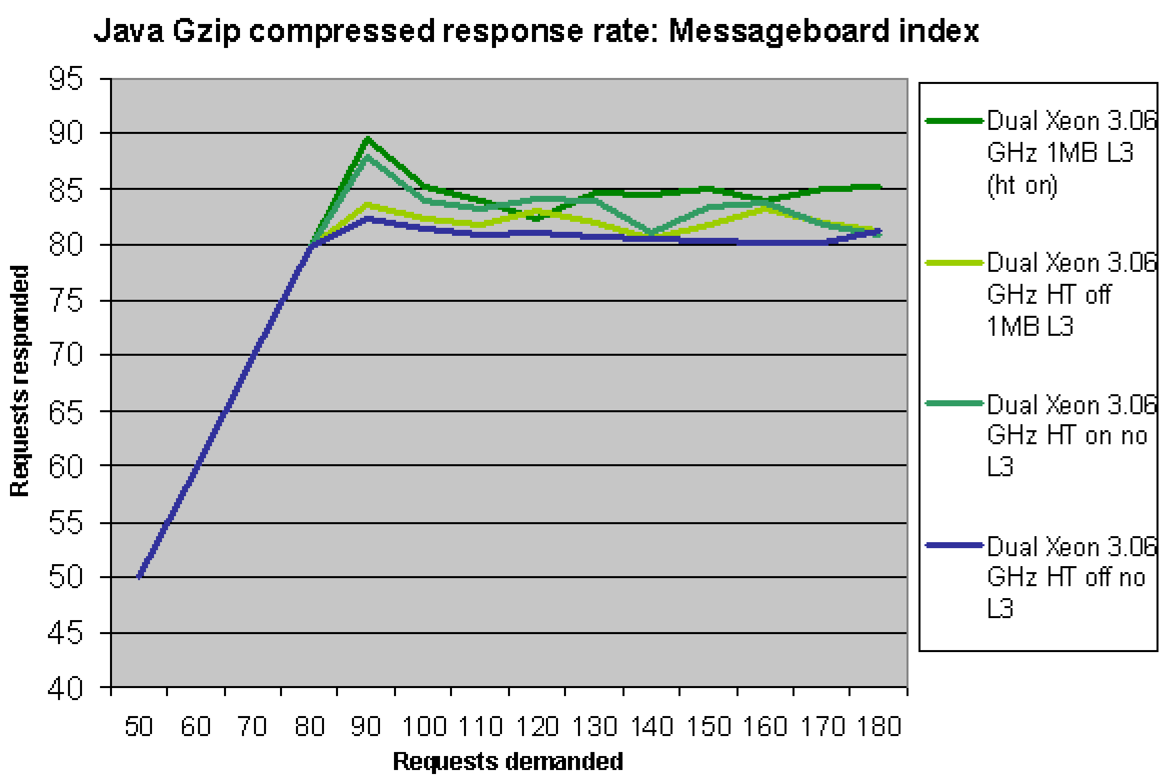
The performance of the Dual 3.06 GHz Xeon is in the line of our expectations: about 8% faster than its 9% lower-clocked 2.8 GHz brother. At this point, the message board request generated about 13 MB/s of bandwidth. When we tested with a simple article, we noticed that the Intel Gigabit had no trouble pushing 40 MB/s and more, so we are 100% sure that the NIC or its driver is not the bottleneck here. No, the real question is: what made the Opteron perform 55% better? We know that the new BIOS improved performance 15-20%. But where did the rest come from? Is the memory controller more aggressive? We asked AMD, we will ask SUSE, but right now, we don't have the answer. Note that all Opterons are using registered ECC DDR333 CAS 2.5.

Let's see how the Opteron scales.



A 22% increase in clockspeed results in a 20% performance increase and adding two extra CPUs is a 60% performance boost. Not really fantastic, but only when we enable compression will things begin to really get interesting. At more than 260 req/s and thus almost 40 MB/s, a lot of work is being done in the network driver, and this probably explains the mediocre scaling with more CPUs. Clock scaling is of course excellent, however.

Next, we take a more detailed look at the performance of the Xeon.



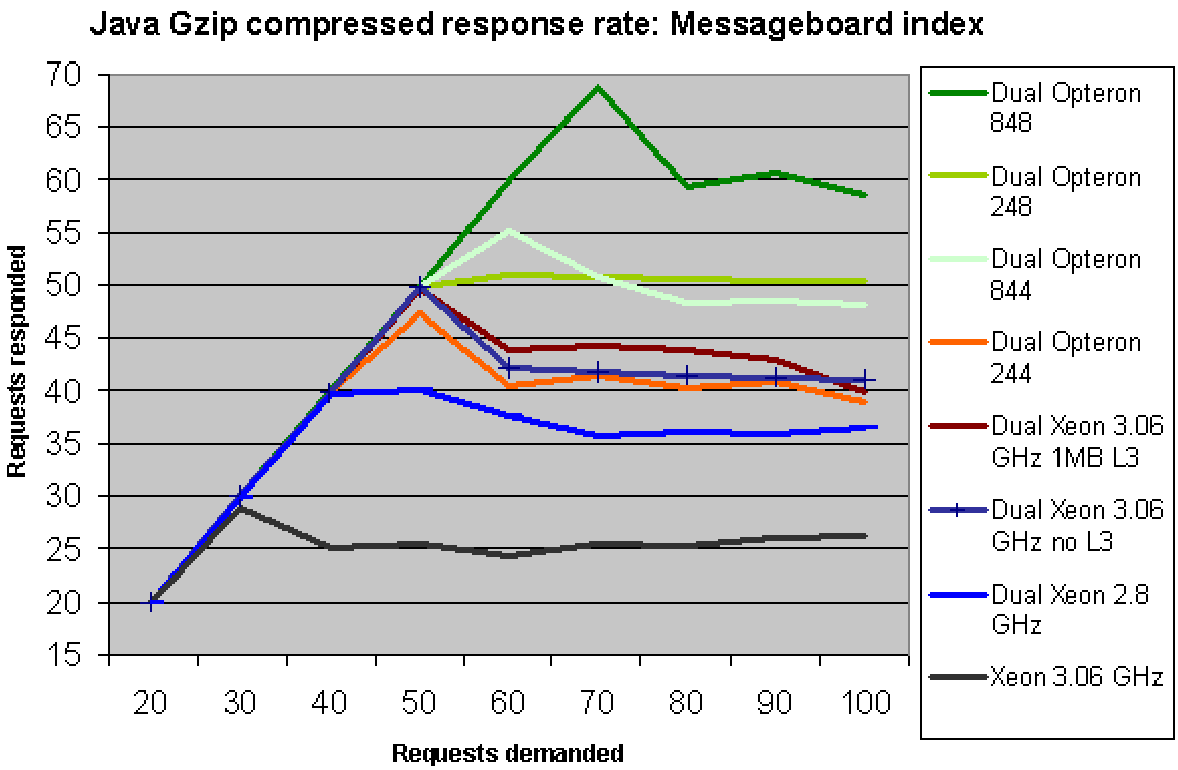
The L3-cache pushes the peak performance of the Xeon about 3% higher, but Hyperthreading boosts performance by 7%. To resume: the Opteron simply destroys the competition.

Cached and Compressed Java Webserver Benchmark: Message Board Index

OS: Linux 2.4.20 / 2.4.19
Multi - threaded: Yes
Memory Usage: 850- 900 MB
Typical Error Margin Between Different Runs: 2-3%
Maximum Network Traffic: up to 1 MB/s

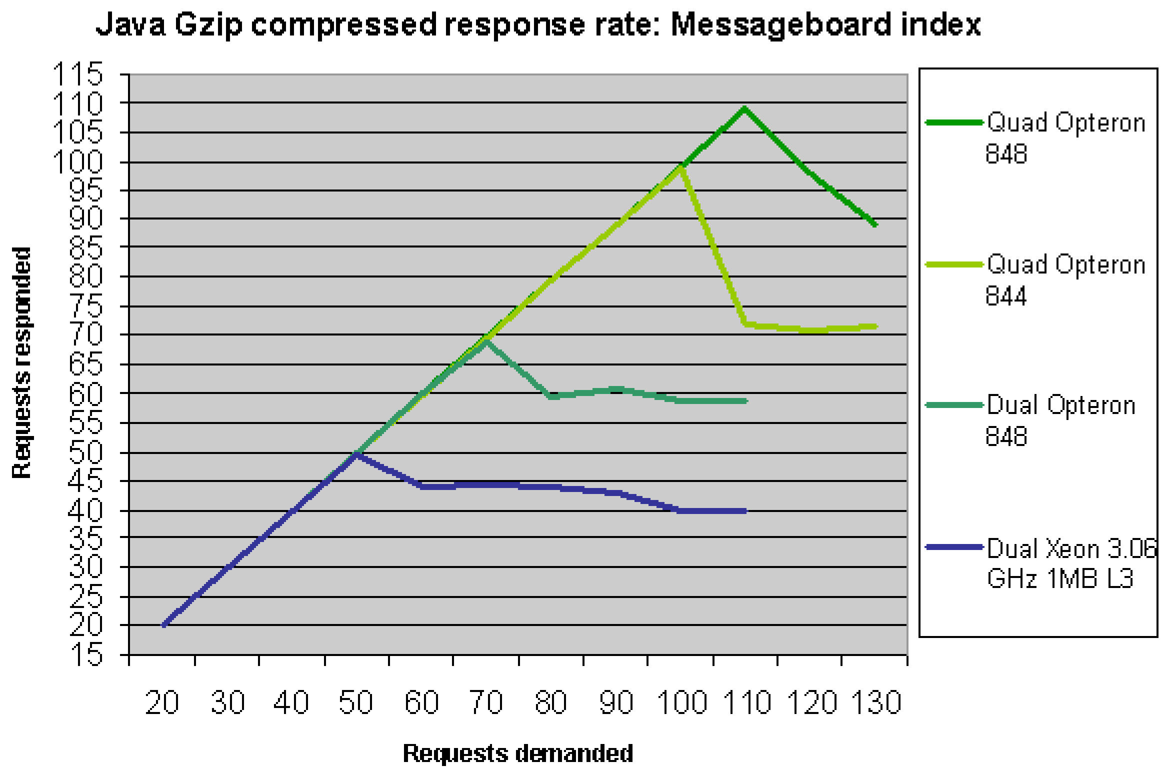
Fast Java performance is great, but one of the main concerns of a webmaster is bandwidth. Costs scale sometimes higher than linear with bandwidth consumption, so excessive bandwidth usage can be an expensive proposition. So, why not compress content on the server before sending it to the client? Almost every browser out there supports **gzipped** content.

The added computational demand of the compression reduces performance by half, but it also reduces the network I/O to 1/15th of its uncompressed size. Performance may be lower, but the bandwidth savings make up for it ten times over, and we can always make up the difference in performance by buying faster hardware with all the money saved on bandwidth charges. In other words, expensive hardware can pay itself back in a very short time.

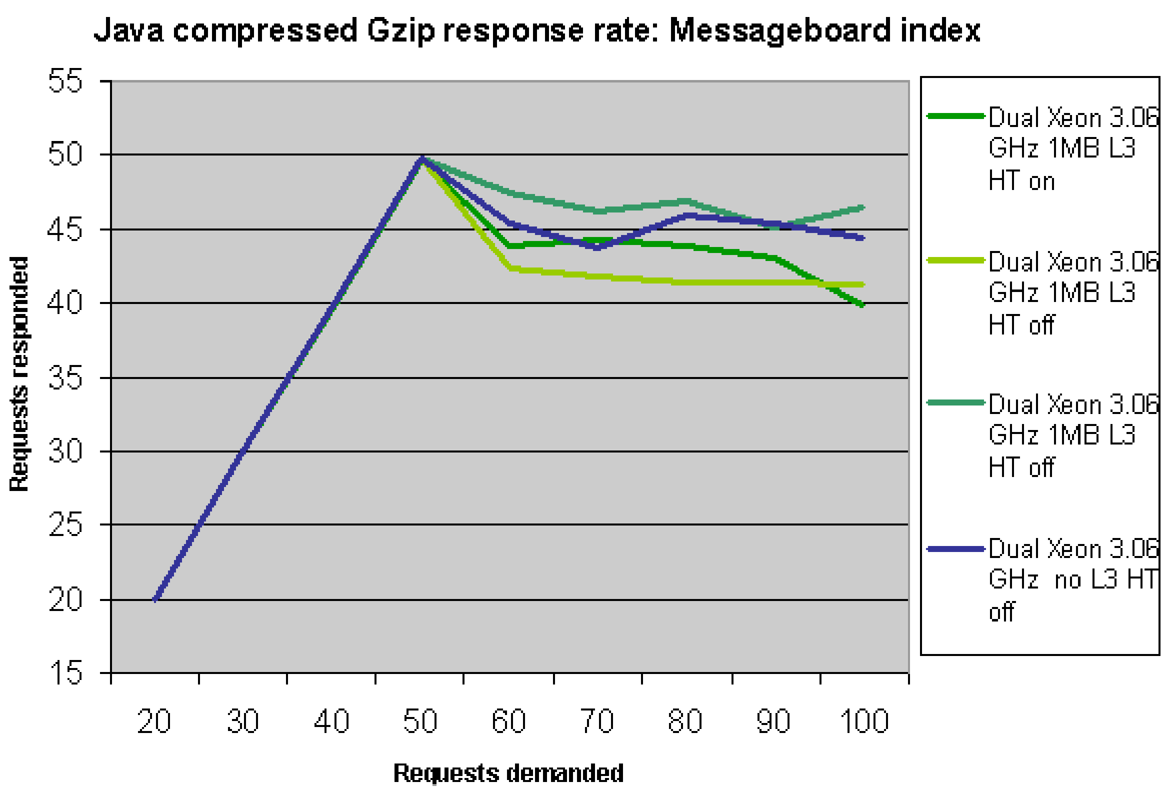


The Xeon being based on (the Pentium 4) "Netburst" architecture likes most compression schemes, and gzip is no exception. The Xeon can keep up with the most of the dual Opterons, but notice what happens once you overload your server. The Xeon starts to give up, and its performance lowers as we stress the server even more. The Opteron dual 844 / dual 248 does not suffer that much, and this is probably where the extra bandwidth comes in. The first Opteron CPU does not have to share its bandwidth with the second.

We included a 248 system for a quick comparison. This system is based on MSI's K8D Master 2 (VIA K8T) which has only one node to access the DDR400 DRAM memory. It also uses the older 2.4.19 kernel. Let's see how the quad Opterons are doing.



The two extra CPUs deliver 60% more performance. However, the Quad Opteron's performance drops more rapidly than the Dual Opterons. It is possible that the non-NUMA kernel has more trouble with four CPUs than with two as the statistical chance that a piece of information is on the right node is lesser. In other words, we speculate that more data has to come from non-local, higher latency accesses. At 1.8 GHz, a full access to the local memory takes about 90 ns, while it takes 145 ns to get the information from over the HT link. Nevertheless pushing almost 110 request per second, with each request containing a fully compressed 985 message index is very impressive. Let's focus on the Xeon.

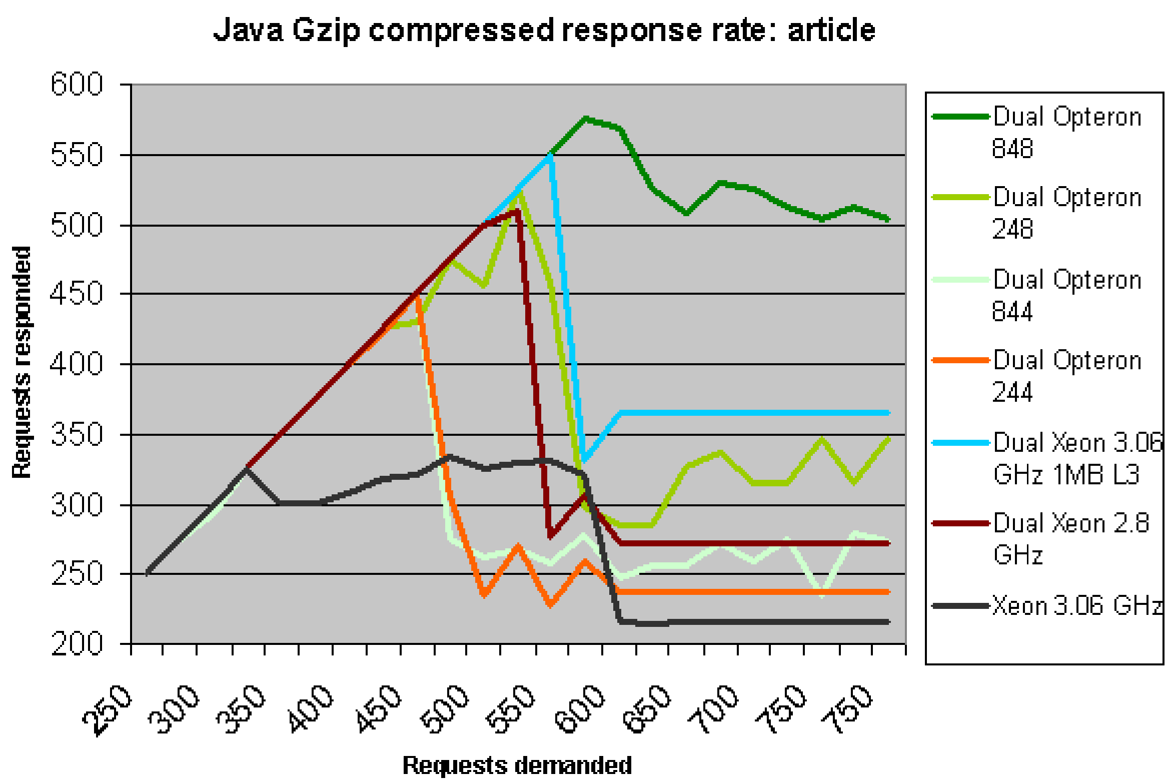


Interestingly, as we noticed in previous tests, performance decreases when Hyperthreading is enabled. The extra L3-cache helps though.

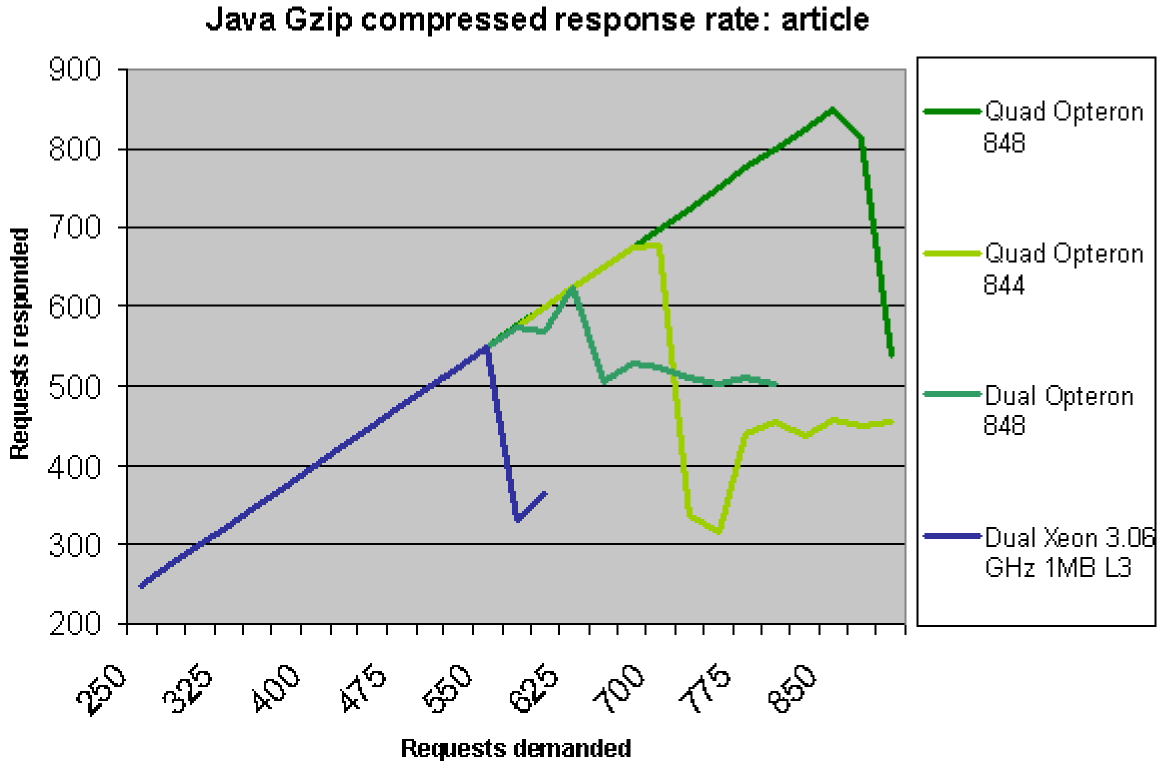
Cached and Compressed Java Webserver Benchmark: Dual Mac Article

OS: Linux 2.4.20 / 2.4.19
Multi - threaded: Yes
Memory Usage: 750 - 800 MB
Typical Error Margin Between Different Runs: 5-7%
Maximum Network Traffic: up to 2.5 MB/s

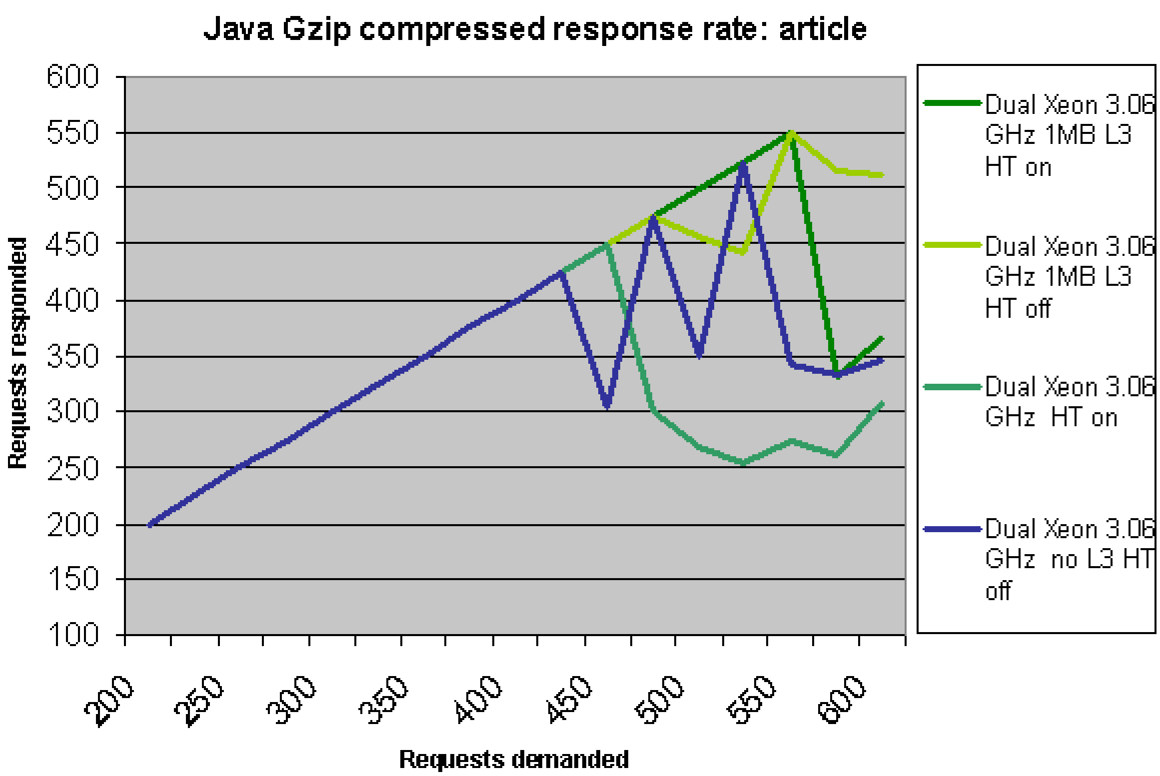
In the next benchmark we do not have the index/tree structure to generate, instead we simply request an article page, namely **A Quick Look at the Fastest Apple PowerMac**.



The revenge of the Xeon! As the CPU does not have to generate the message board tree anymore, the number of branches and the time spent iterating through array structures is significantly diminished. As a result the impact of the gzip compression increases, and the Xeon likes it. Nevertheless, it still has difficulty outperforming the Dual 848, which is running on the new improved BIOS and kernel.



Clockspeed scaling is as good as perfect, but the two extra CPUs deliver only 48% extra performance. Nevertheless, we'll have to compare with the quad Xeon before we can say that this is mediocre.



The L3 cache improves peak performance by no less than 20%, HT has a much more limited role.

MySQL 3.23.49 - Open Source Database Performance

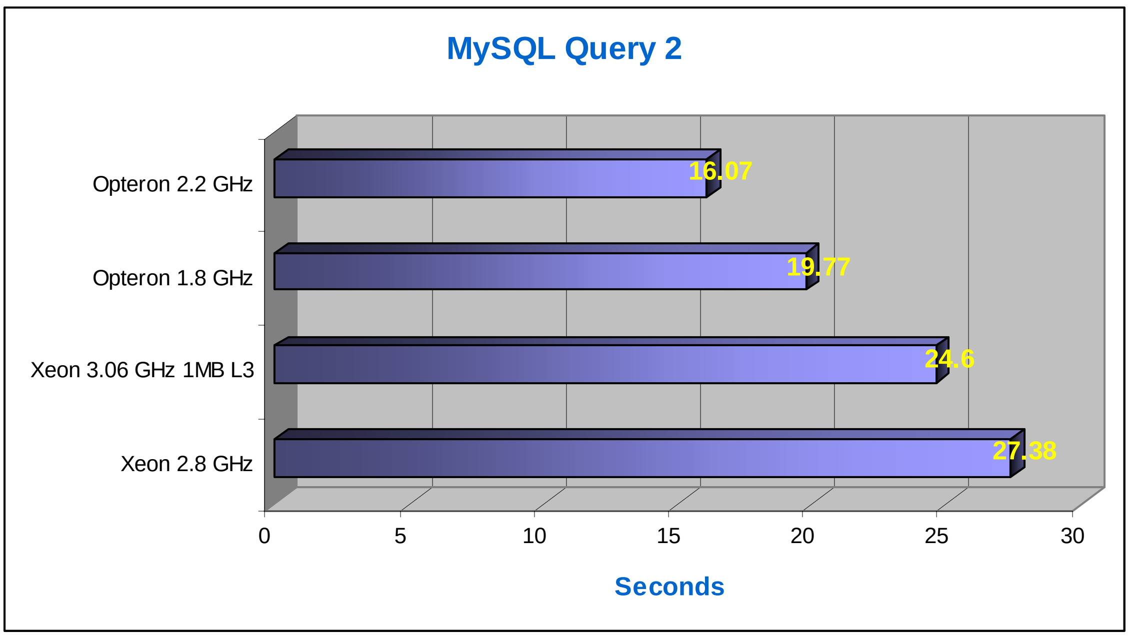
Multi-threaded: No
OS: Linux 2.4.20 / 2.4.19
Memory Usage: 240 MB
Typical Error Margin Between Different Runs: 3-5%
Maximum Network Traffic: N/A, directly on server

For this series of benchmarks, we imported a 400 MB HTTP log from our webserver into a MySQL database. In this test we perform complex "datamining" queries and time them to determine how long they take to run. As this is our own database, we can show you the actual queries.

Note that we are well aware that this kind of benchmarking has its limitations. It only measures the performance when you use your database for reporting. The performance characteristics of inserting and updating records might be totally different. As we launch only one query at a time, this benchmark is also single threaded. It is a good measure to test the kind of response time your server can deliver in a "datamining environment," and of course the most important function of many databases is the information which is available through concise statistical reports. Furthermore, it's important to keep in mind that performance can vary between database servers, as we discovered in our own SQL Server benchmark results.

Query 2

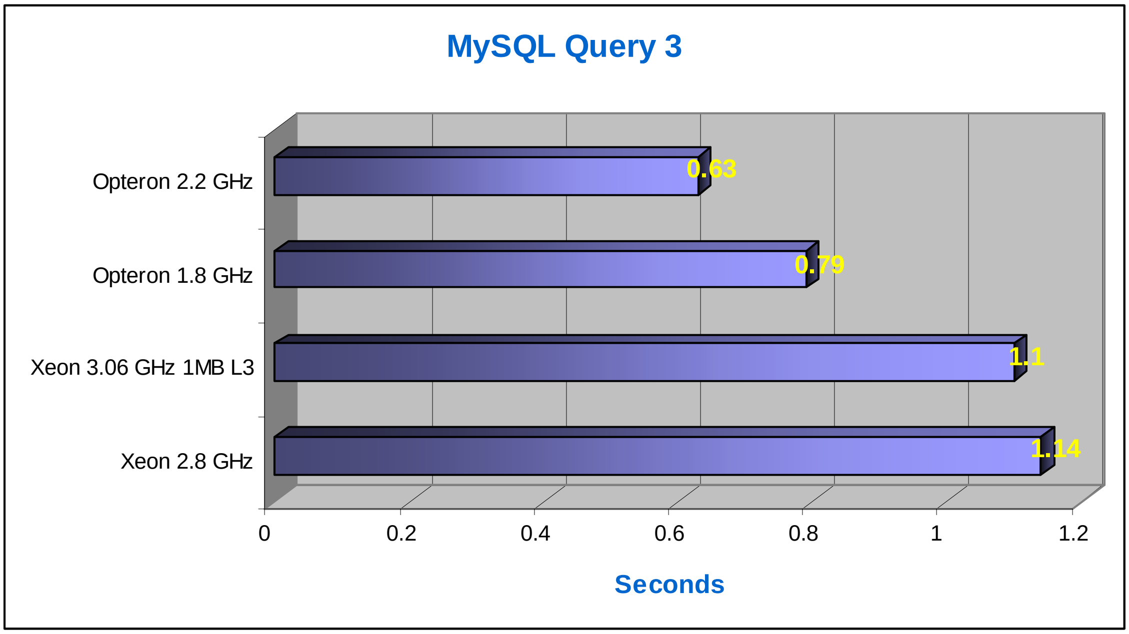
```
SELECT COUNT(*) AS hits,SUM(data_size),f.type FROM files_map f,log l WHERE f.id=file_id
GROUP BY f.type ORDER BY hits DESC
```



We already reported that MySQL runs very fast on the Opteron, and the Opteron x48 extends the lead over the Xeon. Remember that these results are reported in seconds, so lower is better.

Query 3

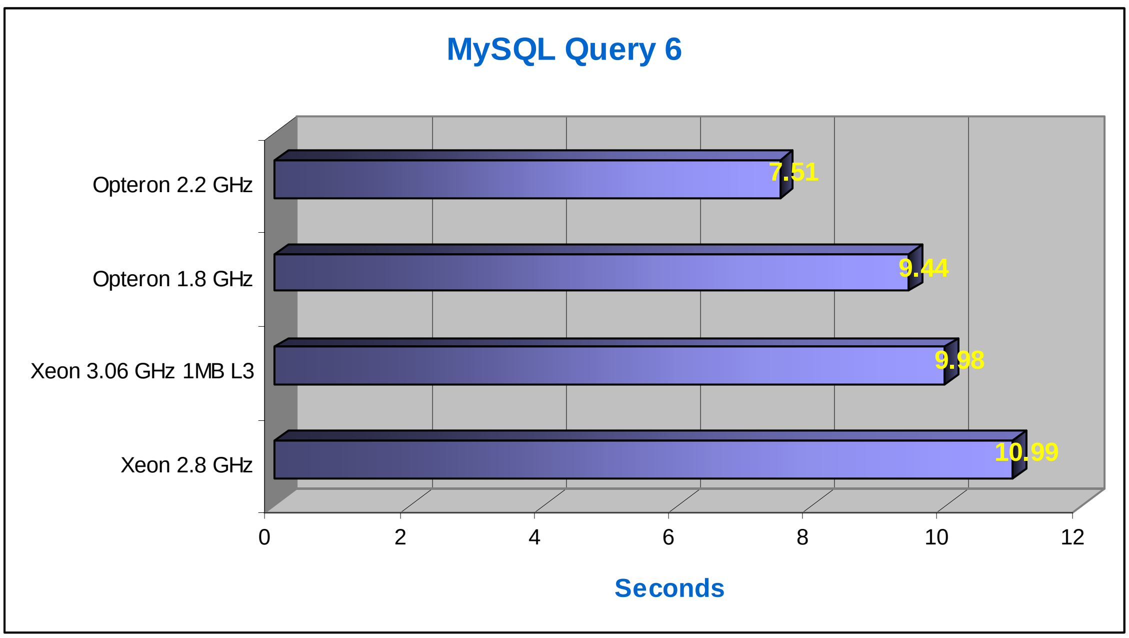
```
SELECT COUNT(*) AS hits,-data_size FROM log WHERE data_size < 0 GROUP BY data_size ORDER
BY hits DESC
```



Again, a landslide victory for the Opteron.

Query 6

```
SELECT COUNT(*) AS hits,f.file FROM files_map f,log l WHERE f.type='' AND f.file LIKE
'%/' AND f.id=l.file_id GROUP BY f.file ORDER BY hits DESCSELECT COUNT(*) AS hits,f.file
FROM files_map f,log l WHERE f.id=l.file_id GROUP BY f.file ORDER BY hits DESC LIMIT 50
```

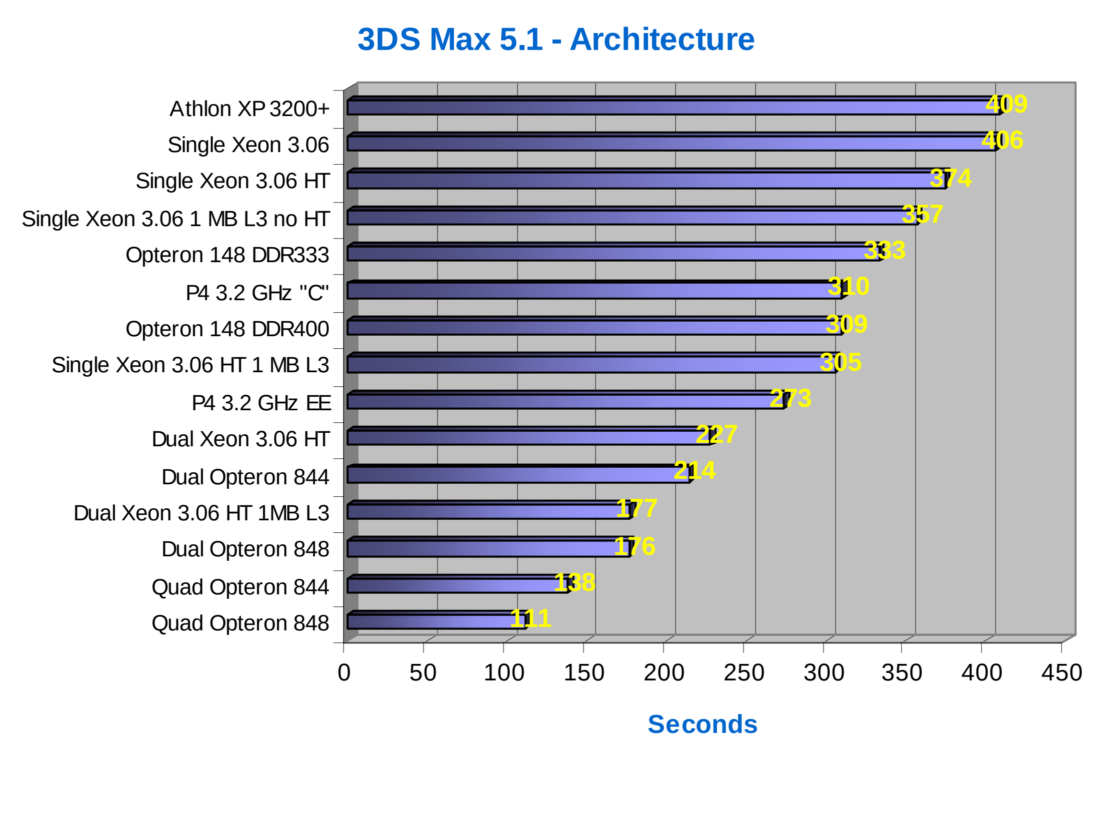


While the Opteron 244 was only 16% faster than its competitor, the Opteron 848 and 248 are no less than 33% faster than the 3.06 GHz Xeon.

3D Studio Max 5.1

Multi-threaded: Yes
OS: Windows 2003 server
Memory Usage: 500 MB
Typical Error Margin Between Different Runs: 1%
Maximum Network Traffic: N/A, directly on server

The 3DSMax test is our classical "architecture" rendering test. We test the Architecture scene from the SPECapc 3DS MAX R4.2 benchmark. This test has a moving camera that shows a complicated building - a virtual tour of a scale model. This complex scene has no less than 600,000 polygons and 7 lights. It runs with raytracing and fog enabled. Frames 20 to 22 were rendered at 500x300 to the virtual frame buffer (memory). To make it more interesting, we reused a few results from our **our Athlon 64 FX and Pentium 4 EE comparison**. The Opteron 148 is a relabeled Athlon 64 FX, as they are otherwise identical.



Our Quartet system was still equipped with DDR333. When registered DDR400 is installed, it can add result in an 8% increase in performance, as we can see from the Opteron 148 values. Again, results are reported in seconds.

Interesting also is the effect of the 1 MB L3 on Hyperthreading. The 1 MB L3 accelerates rendering by 13% if Hyperthreading is off and Hyperthreading accelerates the rendering by 8%. However once we enable both, the Xeon's performance increases by more than 30%. The 2.2 GHz Opteron and Xeon 3.06 GHz (1 MB L3) are equals, which makes the 3.2 GHz Xeon the fastest renderer in dual configurations.

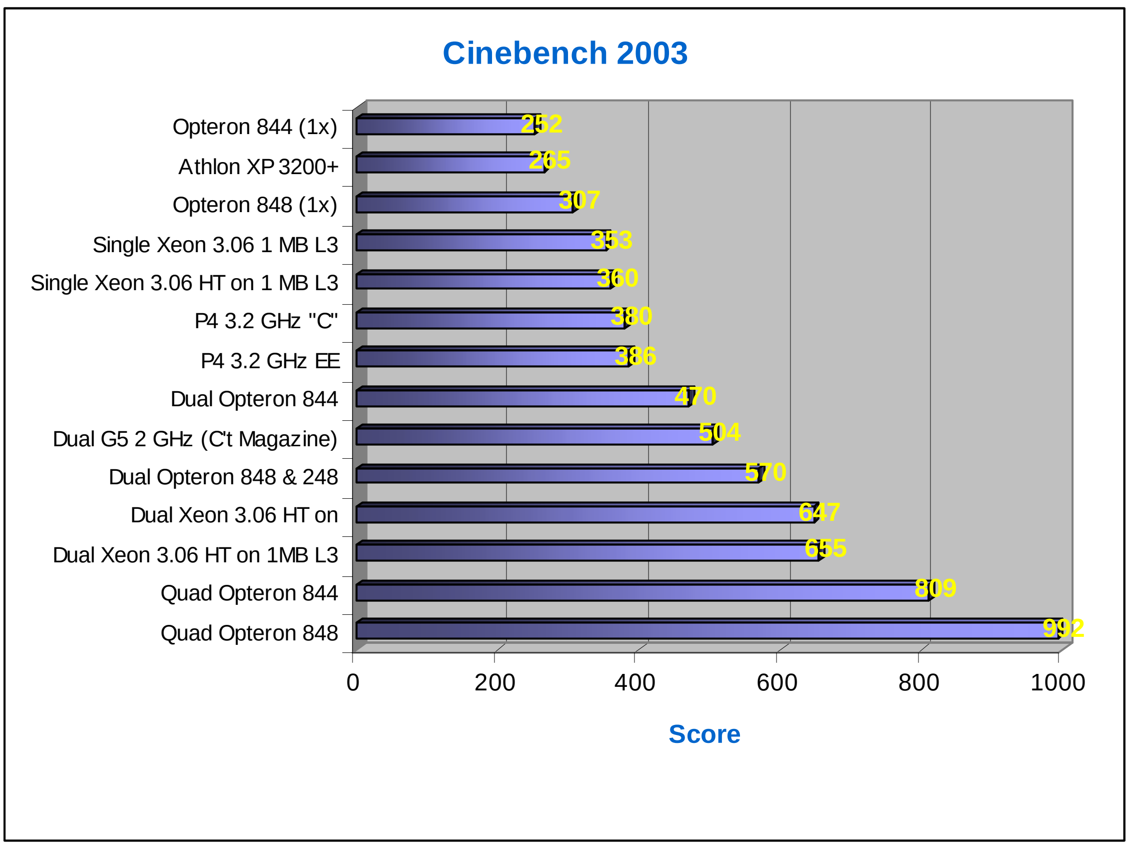
The Opteron scales better, however. Adding a second CPU speeds up rendering by 89%, while a second Xeon makes your rendering machine 72% quicker. Adding a 3rd and 4th CPU is good for another 60% boost.

Cinema4D: Cinebench 2003 (MP)

Multi-threaded: Yes
OS: Windows 2003 server
Memory Usage: 350 MB
Typical Error Margin Between Different Runs: <1%
Maximum Network Traffic: N/A, directly on server

Cinebench is based on Maxon's Cinema 4D modeling, rendering, and animation app. Cinema 4D's renderer is multithreaded and makes very good use of HyperThreading. We report the best rendering scores for all systems (so 2 threads were used for P4).

As we were curious about Apple's PowerPC G5 machines (using the IBM PowerPC 970 processor), we included the Cinebench 2003 result reported by C't Magazine Germany.



The 2 GHz dual PowerPC G5 places itself right between the Opteron x44 and Opteron x48. In other words, in Cinebench 2003, the G5 is clock for clock as fast as the Opteron. Cinebench is well optimized for the G5. For example,

Mathematica show that the Opteron is a lot faster while Lightwave gives a small edge to the Opteron. The Cinebench score is reported here, so higher is better.

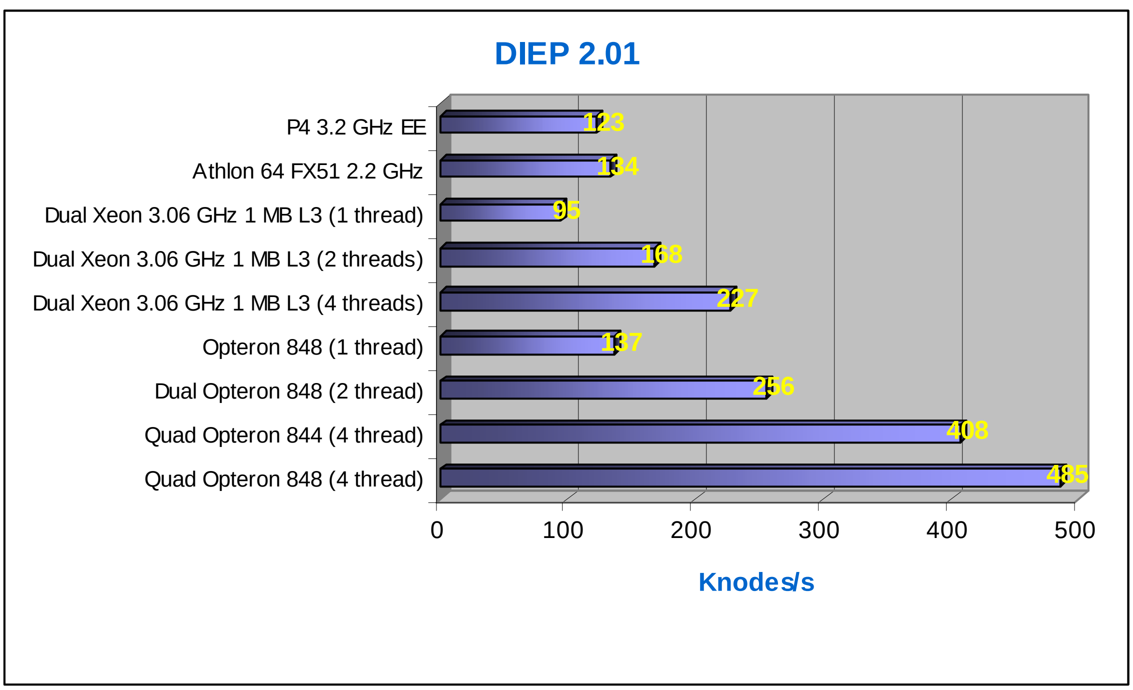
Again, the impact of the L3-cache is rather small. But even the Dual Opteron 848 can not beat the Dual Xeon in Cinema 4D. The excellent HT optimisations give the Xeon the edge and improve performance by up to 22%.

Chess: DIEP

Multi - threaded: Yes
OS: Windows 2003 server
Memory Usage: 500 MB
Typical Error Margin Between Different Runs: 1-2%
Maximum Network Traffic: N/A, directly on server

Vincent Diepeveen is the brain behind DIEP. NUMA systems, whether they run on Itaniums, Opterons, Xeons, or SGI MIPS processors have all been running Vincent's chess program: DIEP. DIEP is an extremely complex and intensive application and it is also a 100% integer program.

The program is only 550KB in size, and has been improved a lot since we introduced it in our first workstation tests. DIEP now makes use of even larger 350MB hash tables (data) in our benchmark, and so the benchmark results we publish today cannot be compared with those from previous tests. While the benchmark depends somewhat on the memory sub-system, pure CPU power is the primary bottleneck. We believe DIEP makes a good addition to our benchmark suite, as good branch prediction is extremely important for the next generation of software that is based on advanced AI algorithms. We tested 13 steps deep:



Thanks to a good friend of ours, we were able to compare this result with the SGI Origin 3800. One computer rack contains **128 500 MHz SGI MIPS R14000 CPUs** with 8 MB L2, and is able to deliver **1681 Knodes/s**. Granted, the Origin

3800 is already an older system, but it still illustrates very well what a well-rounded HPC/Server CPU the Opteron is. The CPU has no real weakness, besides the (theoretical) lower peak SSE-2 FLOPS performance, a result of the slightly lower clockspeed.

A **dual 1.3 GHz Itanium (3 MB)** is able to calculate **about 170 knodes/s**, so we are pretty sure that even a Quad 1.5 GHz Itanium won't beat the Quad 848 machine. We can conclude that the Opteron really shines in these types of A.I. workloads. Scaling is also very good: twice as many CPUs results in almost 90% higher performance.

Conclusion

Since performing and especially verifying server benchmarks take an incredible amount of time, the number of applications we tested was still limited. So we know very well that our conclusions are only valid for a limited part of the Xeon/Opteron market. More precisely, we gained some insight into performance in the webserver, Datamining MySQL, and 3D Rendering markets.

In the 3D Rendering market, the Opteron and Xeon perform similar, with a slight advantage for the Xeon. However, the Opteron scales slightly better with clockspeed and extra CPUs (see 3DSMax benchmarks), while the Xeon still has its Hyperthreading weapon: it's touch and go.

When it comes to (Java) webservers and/or MySQL, the Opteron definitely has the advantage. In some cases, the Opteron simply annihilates the Xeon, but luckily for Intel the latter offers some resistance in our GZIP dominated benchmarks.

Apachebench might be a mostly synthetic benchmark, but it is clearly indicates that the Opteron has a lot of potential left: 64-bit software, AMD64 optimized drivers and NUMA optimized operating systems can push the Opteron far beyond the numbers it is showing today. Especially the last feature could make a real difference for multi-CPU configurations (no wonder AMD and Sun are **working together** on 8-way+ Opteron system architectures).

So far AMD has been advertising the enormous of bandwidth that a Quad or Dual Opteron has available compared to the Xeon, but in most applications the memory latency plays a much bigger role. If the OS keeps all threads local to the processing CPU, latency stays the same, while the Xeon system sees higher latency everytime a CPU has to share the bus.

The Quartet system proves that the Opteron platform is completely ready - with regard to the hardware part - to make headway in the market against the Xeon and Itanium: it has all the RAS features that a modern server CPU must have and it comes with supreme processing power. Now it just needs better support from the software vendors.

Overall, the Opteron is the best server CPU of its class and we doubt very much that the 3.2 Xeon DP with 2 MB L3-cache, coming early next year is going to change that. The Opteron will probably remain the fastest CPU for the server tasks tested here until Intel introduces Nocona, the "Xeon Prescott" at 3.4-3.6 GHz (1 MB L2, 800 MHz FSB) at the end of the 2nd quarter of 2004.