Volume Multi-Processor Systems: Part 3

By Chris Rijk – July 2002

Introduction

Welcome to the third and final part of Volume Multi-Processor Systems, a series generally concentrating on current and upcoming two to eight-way multi-processor systems. In the first part, we covered uniprocessor performance, workstation and server sizing, and CPU and system design considerations. For Part 2, we took a look at the specific architectural implementations of several volume multi-processor systems, including those based around the Pentium III, Pentium 4, Athlon, PowerPC, Itanium, and UltraSPARC architectures. We also covered architectural improvements in CPUs due to appear in the future, including integrated northbridges/memory controllers and improvements in thread-level parallelism (TLP) through on-chip multiprocessing (CMP) and fine-grained multithreading. This time around, we'll take a look into the future of volume multi-processor systems (or the very immediate present in the case of the recently introduced Itanium 2). Of course, we'll look at the significantly improved SPECfp2000-leading Itanium 2, Sun's "low-end" UltraSPARC IIIi with on-board 128-bit DDR SDRAM memory controller and 1 MB L2 cache, and the first 64-bit x86 processor: AMD's Opteron. I'll also be discussing some more general trends of the market as well as making some observations of customer and vendor requirements alike. Finally, there will be an overview of where all the major CPU architectures and companies are at the moment and where they're going.

The primary goal of this article is to look into the volume workstation and server markets to determine what the implications are for the hardware design and usage, for software and the marketplace in general, concentrating on multi-processor systems with 2-8 CPUs. Though various aspects of system design, performance and cost are looked into, this article is not intended to be a buyers' guide. Some basic working knowledge of CPU and system design is assumed, with the target audience being computer systems professionals and computer enthusiasts.

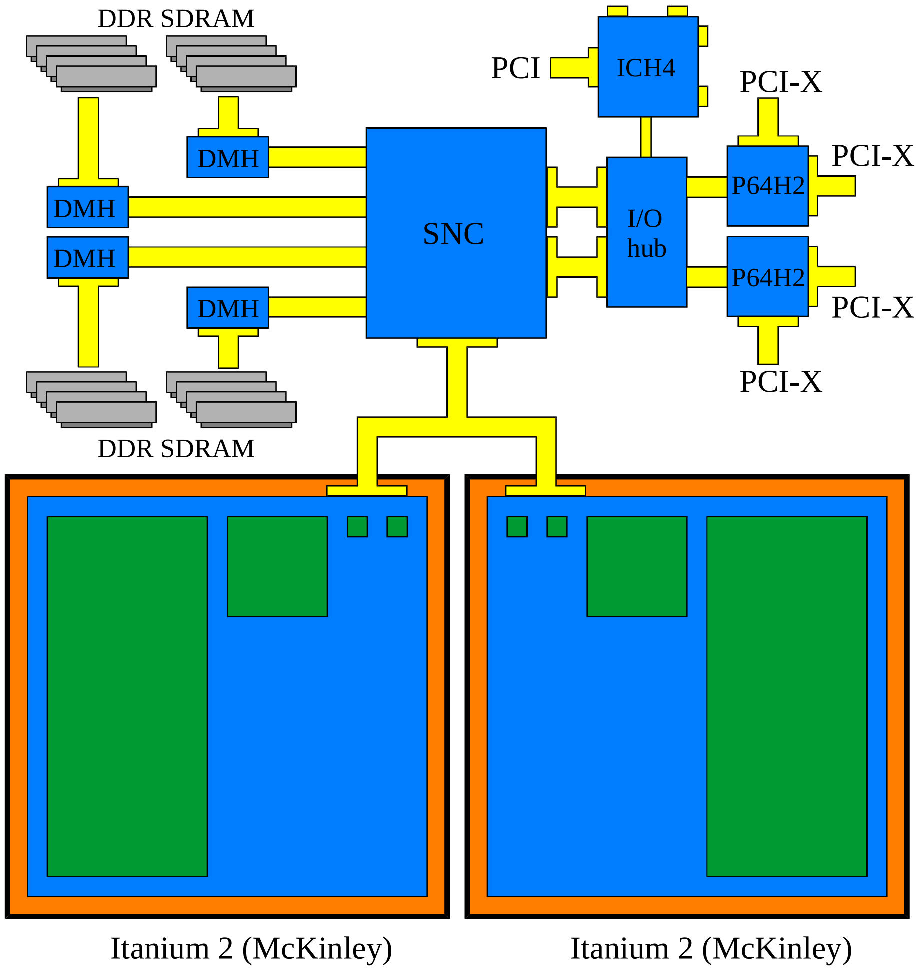
Mostly to help reduce the scope of the article, the main focus is on CPUs and systems designed to be at least reasonably cost effective and for decent volume. This certainly includes systems based on the architectures above from Intel, AMD, Sun, and others.

Since we'll be jumping right into the systems analysis, if you are unfamiliar with how to interpret the system diagrams, please refer to Part 2 for a key.

Intel 0.18µ 2 and 4-way Itanium 2 Systems

Availability: Late Q3 2002

Introduction: McKinley is the second generation Itanium CPU ("Itanium 2") though development has mostly paralleled the Merced design and was begun and mostly implemented by HP's CPU division. HP's CPU design division has a tradition of making very large CPUs and McKinley is no exception, the die size being 421mm^2. This is partly because of the on-die 3MByte Level 3 cache, which is also why the CPU die has 221m transistors, though the logic takes up a significant area too.



Chipset architecture: Intel's chipset for McKinley (the E8870) is similar to the one for Merced, but the address and data ASICs have been integrated into one, called the SNC - Scalable Node Controller. The SNC is rather like a northbridge, except that memory requests and data go via 4 little DMH (DDR Memory Hub) ASICs, each of which handles up to 8 DDR SDRAM DIMMs running at 100MHz DDR, supporting up to 32GBytes of memory with 1GByte DIMMs. The SNC also has 2 "scalability ports", which in 2 and 4-way systems connect to the southbridge I/O hub. In 8-way and larger systems the two "scalability ports" connect to two switch ASICs which route requests and data between groups of 4 CPUs and their respective SNCs, and the I/O hubs.

Bandwidth and latency: Each of the DMHs has one 100MHz DDR SDRAM channel (though uses PC2100 memory modules), for a total of 6.4GBytes/s of bandwidth. This is quite a lot of bandwidth for a single CPU, and even shared it should be enough for most server applications. The route to main memory from the CPUs goes though 6 hops (CPU to SNC to DMH to memory and back again) compared to 4 for a normal northbridge. But as the connection from the SNC to the DMH runs at 800MT/s (probably 16-bit 400MHz DDR or 200MHz QDR) this would save some latency. However, overall memory latency would still be somewhat worse than recent northbridge style Pentium and Athlon systems.

Scalability: 6.4GByte/s of bandwidth with 200MHz DDR signaling shared between 4 CPUs will make for impressive low latency high bandwidth cache snooping. The large amount of bandwidth and the 3MByte cache would also help scalability considerably for 4-way systems. However, this chipset is expected to be used for 0.13um Itaniums as well, and as a basic part for 8 and 16-way Itanium systems, which may strain scalability somewhat by the end of its lifetime. For bandwidth intensive applications (one of the main target markets initially), scalability will be poor since a single CPU can consume all the available bandwidth.

I/O: From the I/O hub, there are up to 4 connections using the same 1GBytes/s P64H2 I/O chips that Pentium 4 Xeon systems have. This allows for up to 8 PCI-X channels, and a lot of I/O bandwidth for a 4-way system - 2-way systems would likely only use a sub-set, as shown in the diagram above.

CPU and cache: The Merced has a decent cache system (the 4MB off-die high-speed SRAM in a MCM was the best bit though) but seemed to be held back by the CPU core. With 16KByte 1 cycle latency dual-ported level 1 caches, an on-die 256KByte level 2 cache and a 1.5/3MByte on-die level 3 cache, the McKinley has a very impressive cache system. The CPU core is more efficient and runs faster, despite having an 8 stage pipeline, compared to 10 for Merced. With the significantly improved system interface, together this should help both integer and floating-point applications significantly. Maximum clock speed at 0.18um will be 1GHz at launch; with possibly faster CPUs coming later, and maximum power consumption at 1GHz is 130W. Prices for a 0.9GHz Itanium 2 with 1.5MBytes of level 3 cache will be $1338 (probably when buying 1000+ units though) and $4226 for the 1GHz 3MByte level 3 cache part.

Summary: The EPIC/IA-64 architecture was supposed to be superior to other architectures to date, but the Merced design failed to show this, and has been re-labeled as test design. Itanium 2 lives up to the promise far better than its predecessor ever could. On a 4-way motherboard using HP's Itanium chipset (which has twice the memory bandwidth as Intel's chipset) a 1GHz Itanium 2 gets a SPECint_base of 810 and a SPECfp_base of 1356, which are the best for any 0.18um system reported so far. However, in the SAP 2-tier SD benchmark, a 4-way system is just 11% faster than a 4-way 1GHz Alpha system. It seems likely that the SPEC codes can be much more highly optimized for the Itanium than a database application benchmark - the SPEC code can be optimized exactly for the benchmark but not a database application, which requires general-purpose optimizations, as do most real-world applications. Many such applications also feature integer-dominant workloads, and though Itanium 2 is hardly weak in this area, the highly-clocked “speed demon” 0.13um Pentium 4 does manage to outperform it in SPECint.

Prices for systems will probably be in a similar range to Merced systems, or about the same as a similarly sized UltraSPARC-III system, though there will probably be some "low cost" (and lower performance) versions. 1, 2 and 4-way systems are expected on release, with 8-way systems coming late in 2002 and much larger systems coming in 2003. Systems larger than 4-way are based on connecting multiple 4-way boards together in a NUMA architecture, so such systems will not likely have good performance until the OS has NUMA optimizations.

However, whether McKinley systems (and Itanium as an architecture) succeed or not is highly unlikely to be determined by a price or performance difference of 10%, 20% or 30% for example. For a start, a platform is only useful to end-users if it has applications available for it that users want or need. The more applications there are available and supported (including sub-modules for the large applications), the more users can consider that platform as a viable alternative. About the biggest problem Itanium 1 (Merced) systems had was lack of software. There are thousands of programs that are important to business customers as a whole, and though Intel are putting a lot of effort into this, it will still take a lot of time for Itanium to build up a comprehensive software portfolio. The main problem is that the Itanium is not backwards compatible with anything so will go through the growing pains all new platforms face - the x86 backwards compatibility will almost certainly be too slow to be of serious use.

This topic is discussed in more detail later in this article, but at the most basic level it can be likened to launching a new games console - the more quickly a games console can build up a large customer base and a large software portfolio the better, and can fail if this doesn't happen fast enough. Customers of heavy-duty systems for businesses are mostly quite resistant to change, and the bigger the change (and the higher the risk), the more resistant they are. To make matters worse for Itanium OEMs, in the current business client with weak IT budgets, customers are going to be less willing to take risks - any new platform is a risk. It is generally expected that demand for Itanium 2 systems will not start to become serious until about a year after launch - by this time cheaper 0.13um versions will be available and that platform will have started to become "proven".

A subtler problem with persuading customers to buy Itanium is that many are quite happy with their Intel x86 systems, which are also in higher volume, and more competitively priced. At the lower-end (2-8 way) the x86 systems will likely have reasonably similar performance for server applications. In addition, they will likely be cheaper to buy and use as they don't require buying Itanium versions of all their current software, and some level of re-training will be required. The higher-end x86 systems can use up to 64GBytes of RAM (though not very efficiently), so at the 2-4 way level the fully 64-bit Itaniums will only have a clear advantage in a few applications, mostly workstation ones. With competition from AMD increasing, Intel can't afford to let their x86 MP chips and systems fall behind, so x86 CPUs are the biggest obstacle to the success of Itanium.

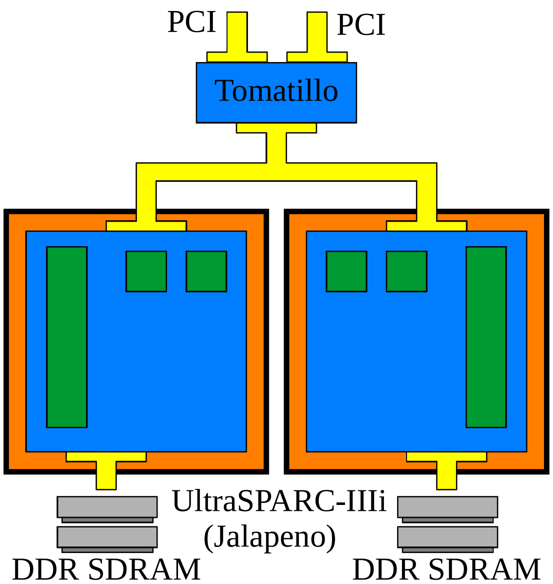
References:

- Introduction to the E8870 chipset
- E8870 chipset white paper
- Itanium 2 white paper

Sun 0.13µ 2-way UltraSPARC-IIIi Systems

Expected: Sometime in 2002.

Introduction: The UltraSPARC-IIIi (codename "Jalapeno") is Sun's first "low cost" multi-processor UltraSPARC - the UltraSPARC-IIi was Sun's first "low cost" single-processor UltraSPARC. It is based on the UltraSPARC-III core, but with 1MByte of on-die Level 2 cache (instead of 8MB of external cache), has a full on-chip memory connection, using PC2100 DDR SDRAM. At 0.13um, it's expected to run at least 1-1.1GHz, and reach 1.6GHz, probably in 2003.



Chipset architecture: The FSB for UltraSPARC-IIIi is called JBus which is a bus connecting the CPUs and the I/O bridge ASIC (codename "Tomatillo"). As each CPU has a build-in memory controller there is no need for a northbridge, making for a very simple design. Each CPU can control a maximum of 4 DIMMs, and maximum memory per CPU is 16GBytes, when 4GByte DDR DIMMs become available.

Bandwidth and latency: The local memory load-use latency is just 75ns, and will likely be the lowest on any high-performance CPU at launch by a significant amount, and bandwidth of 4.2GB/s per CPU will be impressive too. A 2-way system would have an aggregate bandwidth of 8.4GByte/s. Non-local memory access is 120ns, which is comparable to modern northbridge designs. If memory accesses were evenly spread between local and non local, then the average latency for a lightly loaded system would be the average of 75 and 120, or 97.5ns.

Scalability: Using the high-speed bus design allows for low latency cache snooping, though remote memory access (and I/O) has to use the same bus, so it may become a bit of a bottleneck in very intensive applications. Non-local memory requests take 60% longer, so the more memory requests are local, the better performance and scalability will be. Though optimizing this would be relatively simple for benchmarks like SPECint_rate and SPECfp_rate, it's hard to do for programs, which have a lot of shared memory, such as databases. Sun has the advantage of designing the OS and several common applications that are used for their systems, which makes optimization easier. However, the current local memory optimizations planned for Solaris 9 seem to be for higher-end (8-way and higher) UltraSPARC-III systems only. Still, 1MBytes of cache is a good amount for a 2-way system, so cache hit rates should be enough for most applications that the main memory location isn't too significant.

I/O: The JBus supports up to two I/O ASICs, each of which have 2 64-bit 66MHz PCI controllers, one of which can also be used as a UPA64S connection, which is the Sun equivalent to AGP. Each PCI controller can sustain over 500MByte/s simultaneously, for 2GByte/s of I/O for a system with 2 I/O ASICs - which is more than what the Sun Fire V880 supports. The JBus supports 7 devices in total - 4 CPUs, 2 I/O chips and an unspecified other device, which would most likely be a high-speed interface to a graphics board as there aren't really any other sensible uses.

CPU and cache: The UltraSPARC-III core is a pretty general-purpose 64-bit design and so is UltraSPARC-IIIi, though 1MBytes of cache would be a rather lean for heavy-duty databases, but in a similar range to its intended competition - Intel's Xeon CPUs. Interestingly, the CPU itself will be mounted on a small daughter card, which also has 4 DDR SDRAM slots for the local memory, and that card plugs into the motherboard. Those same cards can be used in other UltraSPARC-IIIi system and produced in volume (which reduces costs) and also makes the motherboard design significantly simpler as they would just have some I/O chips and a bunch of PCI and JBus slots.

The UltraSPARC-IIIi has 87.5m transistors, comes in a 950-pin package, and runs at 1.3V on a 0.13um process. It consumes up to 60W of power and it also has some special low power modes where the CPU and JBus clock can be reduced to 1/2, 1/32 or 1/64th of the normal operating speed (down to about 15MHz), including in multi-processor systems. This is primarily for the rack-mount market - most servers only have average utilization of about 20%, so when demand is low, the clock rate can be reduced to conserve power. A die size hasn't been specified but the UltraSPARC-IIIi will probably be about 160-190mm^2.

Summary: The UltraSPARC-IIIi system design is certainly quite simple compared to the 2-way UltraSPARC-III systems, which should help reduce costs significantly. With double the main memory bandwidth per CPU, half the main memory latency, about double the level 2 cache bandwidth and half the cache latency (though a poorer cache hit rate), the UltraSPARC-IIIi could see an IPC improvement of 20% or more in many target applications, as well as running at higher clock rates. So it should make a good workstation and entry-level server, though this somewhat depends on when it gets to market. How much systems will cost is not clear, but Sun have low specification workstations and servers for under $1000 with the UltraSPARC-IIe chip, and one of the design goals with UltraSPARC-IIIi was to compete with 2-way Intel systems on price. The 2-way servers are code-named Enchilada and will likely be called the Sun Fire V240, and there are also expected to be 1-way workstations and servers.

To compete against 2-way Intel systems, the UltraSPARC-IIIi systems will have to be at least half the price of the current UltraSPARC-III systems (which seems quite reasonable given the design differences) while offering more performance than its predecessor. With half the price seeming likely and 25-50% more performance, that'd be 2.5-3 times the price/performance over the current generation. Interesting Sun will soon start selling general-purpose 2-way rack-mount x86/Linux servers using Pentium IIIs, which will probably fit below the 2-way UltraSPARC-IIIi systems in terms of prices and performance. This seems to be more for Sun customers who specifically want x86 Linux systems.

At such prices they'll make popular workstations, though the cost of graphics cards will likely be a problem - Sun might well need to use AGP cards to be competitive. For servers, 1U 2-way systems are also expected, a big improvement on the current 4U systems. Though likely to be popular with Sun's current customers, and highly attractive to Unix customers in general, few companies who use only Windows will be interested due to the high initial costs of having to buy new software, and do re-training and possibly migrations. However, most large companies have a mix of Windows and Unix systems, and Unix systems are particularly popular in data-centers.

References:

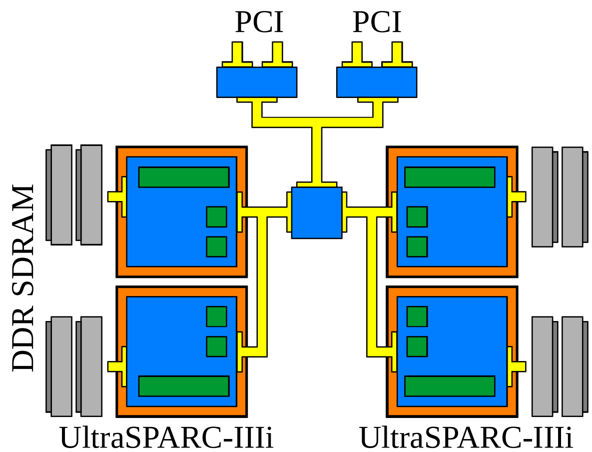
- "Dash of Jalapeño" talk and paper by Kevin Normoyle, given at 2001 Microprocessor Forum

Sun 0.13µ 4-way UltraSPARC-IIIi Systems

Expected: Sometime in 2002

Introduction: The 4-way UltraSPARC-IIIi system designs are very similar to the 2-way ones, except one extra ASIC is needed - a simple chip that splits the JBus into 3 domains, where all can be active at the same time. The exact same UltraSPARC-IIIi CPUs will be used, though the systems come with 2 I/O ASICs - the same as for the 2-way systems.

Performance: The simple extra ASIC adds one cycle of latency (5ns at 200MHz) as it simply copies the data between ports and arbitration is handled by the JBus controllers on the CPUs and I/O ASICs. Assuming the figure of 120ns for remote memory access is for 2-way systems (the documentation isn't clear), remote memory access over the extra chip would add 10ns of latency, so average latency for a 4-way system would be the average of 75, 120, 130 and 130ns, or 114ns, which will likely be lower than all competing Intel designs. It should be noted that it's not clear if the local memory latency for the UltraSPARC-IIIi is always 75ns, but it seems likely given that the UltraSPARC-III can start a memory request before a cache snoop completes. Aggregate bandwidth for a 4-way system would be 16.8GByte/s.



Scalability: Compared to the 2-way systems, the 4-way UltraSPARC-IIIi systems may well struggle a bit, as latency rises and contention of the FSB increases. In addition, for real applications, larger systems typically have larger and more complex data, so cache hit rates would likely be lower for the same applications. Certainly the benefit of local memory optimizations are larger, the biggest the systems, and Sun have estimated that maximum sustained bandwidth can double for UltraSPARC-IIIi with this.

Summary: If the 2-way UltraSPARC-IIIi system prices are reasonably close to 2-way Pentium 4 Xeon prices, then the 4-way UltraSPARC-IIIi systems should be very competitive. This may undercut the 4-way Pentium systems, though much depends on how prices may change. For 4-way and above, the Xeon MP CPUs have a significant price premium, as do the systems, and a 4-way capable Xeon MP system with 2 CPUs costs over double that of a 2-way Pentium 4 Xeon system with the same amount of CPUs and memory. A 4-way UltraSPARC-IIIi system with 2 CPUs would likely have only a small extra cost over a 2-way system with the same amount of CPUs and memory, as there are few differences. In addition, the Xeon MPs run more slowly than their 2-way counterparts, though have more cache, and also lag in terms of using the latest process technology - the UltraSPARC-IIIi will most likely be at 0.13um before the Xeon MPs.

A leading edge UltraSPARC server that's cheaper than leading edge 4-way Intel x86 servers would certainly be popular in the Unix market. But it's a hard sell to companies which only have Windows systems, as they would face extra costs in getting new staff, software and so on. Also, though Solaris has 10,000 applications available for it, there are a number of packages that are only available for Windows. This is an area in which Sun can benefit from the growing uptake of Linux systems, as if a customer which previously only bought Windows systems buys a few Linux systems and is happy, they're much more likely to then buy a Sun system, as they've already done a Unix migration.

Technical footnote: An interesting feature of the CPU design from a technical point of view (it's purely a performance enhancing feature) is the asynchronous logic in the UltraSPARC-IIIi's memory controller. Although with some standard northbridge designs, the memory controller supports "asynchronous" speeds (that is, the memory bus clock speed is independent of the FSB speed), this is altogether different. Previously for their system designs, Sun run the memory at below the top specified speed for the memory, to guarantee solid reliability - for example, the UltraSPARC-IIe also has a full on-die memory controller but even though it uses PC133 memory, it's only run at a maximum of 100MHz. The asynchronous logic in the UltraSPARC-IIIi is just for the bit on the edge of the CPU that does the I/O signaling with the memory and because of its self-timed nature (completes as soon as it has the signal, but only when it has the signal properly) it absorbs small variations in the signal timing, allowing the signal to be run at a higher clock rate with more reliability than possible before. A small FIFO (first in first out) queue synchronizes the memory timing with the rest of the CPU.

This feature also allows the CPU to be run at a lower clock rate than the main memory - DDR SDRAM doesn't support changing the clock speed dynamically.

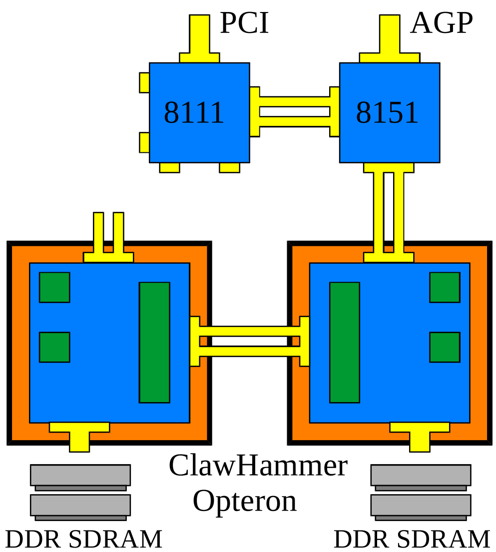
AMD 0.13µ 2-way Opteron Systems

Expected: First Half of 2003

Introduction: The code-named Hammer family of CPUs from AMD will replace the current Athlon range eventually, with all MP capable Hammers being released under the brand name "Opteron". The 1-way only Hammer CPUs will use the Athlon brand name in some form. Although the Hammer CPUs are a 64-bit CPU design, in many ways it has a "keep it simple stupid" approach as instead of doing a completely new architecture, AMD simply extended the x86 architecture again, like with the migration from 16 to 32-bits Intel did many years ago.

There are two Opteron variants aimed at the 2-way market, one a lower-cost model with smaller cache and lower memory system bandwidth, probably a MP version of the desktop CPUs (ClawHammer), which is the one illustrated in the diagram below. The other looks likely to have more cache and twice the bandwidth (SledgeHammer), being designed pretty much for MP and business systems only.

All the Hammer CPUs have an integrated memory controller, and use uni-directional point-to-point links between CPUs and I/O ASICs, called HyperTransport. The links on SledgeHammer are 16-bits wide in each direction, and 8-bits wide in each direction on ClawHammer, running at 800MHz DDR. Two uni-directional 16-bit wide links at 800MHz DDR would be a bit slower overall on average than a single dual-directional 32-bit wide 800MHz DDR link, but the later would be much harder to engineer.



Chipset architecture: Each ClawHammer has two HyperTransport links and each SledgeHammer has 3. For a dual processor system, one link connects the two CPUs with another connecting to a southbridge-like I/O ASIC. Though in theory multiple CPUs could directly connect to multiple I/O ASICs, in practice this is troublesome, as the CPUs are not fixed to the motherboard.

Bandwidth and latency: With 64-bit wide memory running up to PC2700 (166MHz DDR) and an on-die memory controller, ClawHammer should have about 40% lower latency and 30-40% more bandwidth than Athlon systems with PC2100 memory. AMD claim that with the same memory (PC2100), first word latency drops from 160ns, which seems high, to 80ns in single CPU systems. Minimum latency in a 2-way system would be a bit higher (maybe 90-95ns) due to the wait for a cache snoop. This is a bit slower than the UltraSPARC-IIIi (for the same memory), and perhaps the UltraSPARC-IIIi CPU/system design is more optimized for latency than bandwidth, relative to the Opteron and HyperTransport.

SledgeHammer has a 128-bit wide memory interface, running at up to 166MHz DDR (5.4GBytes/s) with PC2700 memory. The ClawHammer HyperTransport links are 1.6GBytes/s in each direction (3.2GByte/s aggregate) and they are 3.2GBytes/s in each direction for SledgeHammer (6.4GBytes/s aggregate). In practice DDR SDRAM's sustained bandwidth is somewhat below the maximum, so the HyperTransport links between the CPUs would likely not limit delivered bandwidth when one CPU is accessing memory via the other CPU.

Scalability: With ClawHammer and SledgeHammer, the connection between the two CPUs can sustain more than enough bandwidth for both CPUs to be simultaneously streaming data from their local memory at top speed while copying it over to the other CPU. So effective bandwidth should scale very well from 1-way to 2-way systems. The latency will be low and stay low as well, thanks to the HyperTransport connection, though like with the UltraSPARC-IIIi, local memory optimization would reduce average latency and increase maximum sustained bandwidth for some applications.

I/O: The speed of the HyperTransport links is more than enough for the I/O ASICs, and AMD have an I/O ASIC that support AGP 8X, another that supports two 64-bit wide 133MHz PCI-X channels, and other which supports lower-end I/O needs. Though this means I/O latency is not even between the CPUs (unlike all the other systems in this article), this likely won't make much difference in practice.

CPU and cache: From a high-level view, the CPU core looks rather like that of the Athlon, but with all-round incremental improvements. The branch prediction seems to be the most improved, with a much larger branch prediction table and probably more accuracy. SSE2 instructions are also supported, which will boost workstation applications. The core also supports a new group of general purpose 64-bit instructions, and have double the number of general purpose registers (compared to the 32-bit versions), which puts it more in the league of RISC CPUs. Re-compilations simply with these new instructions will help general-purpose code a bit, when using these new instructions, perhaps by about 5-10% on average - depends on the algorithm a lot. The main performance boost will come from the integrated memory controller and lower latency, though clock speeds are expected to be higher than the Athlons as well.

The cache structure seems rather like that Athlon, with 64KByte instruction and data caches. The on-die Level 2 cache is 256KBytes in size for ClawHammer and 0.5-1MBytes for SledgeHammer. The expected die-size is 106mm^2 for ClawHammer, compared to 80mm^2 for the 0.13um Athlons with 256KBytes of Level 2 cache. SledgeHammer die size will probably be about 150mm^2. The 1-way ClawHammers are expected to run at 2GHz or more on release, and perhaps up to 2.6GHz on 0.13um, while SledgeHammer CPUs will probably be running 20% slower, perhaps starting at 1.6GHz and running at up to 2GHz on 0.13um. Power consumption is not particularly clear either, but probably not significantly different from 0.13um Athlons. The ClawHammer Opterons come in a 754-pin package and the SledgeHammer ones come in a 940-pin package.

Summary: In the short term, the success of Opteron CPUs, particularly ClawHammer, will be judged on their performance in current 32-bit x86 code. On that score they should do well, with improvements in all areas, particularly the memory system and with SSE2 instructions. ClawHammer will probably be between 15 and 30% faster than a same speed Athlon XP on current applications, and a dual processor ClawHammer system should have a bigger advantage over a dual processor Athlon MP system due to having more memory bandwidth. With up to 5.4GByte/s of memory bandwidth, dual-processor ClawHammer systems will make good workstations. With a low latency memory system and good general-purpose core, it should make a good low-end server system too, though 256KBytes of Level 2 cache would be a bit poor. With twice the memory bandwidth and more cache, the SledgeHammer Opterons should be particularly popular as more heavy-duty workstations and higher-level 2-way servers.

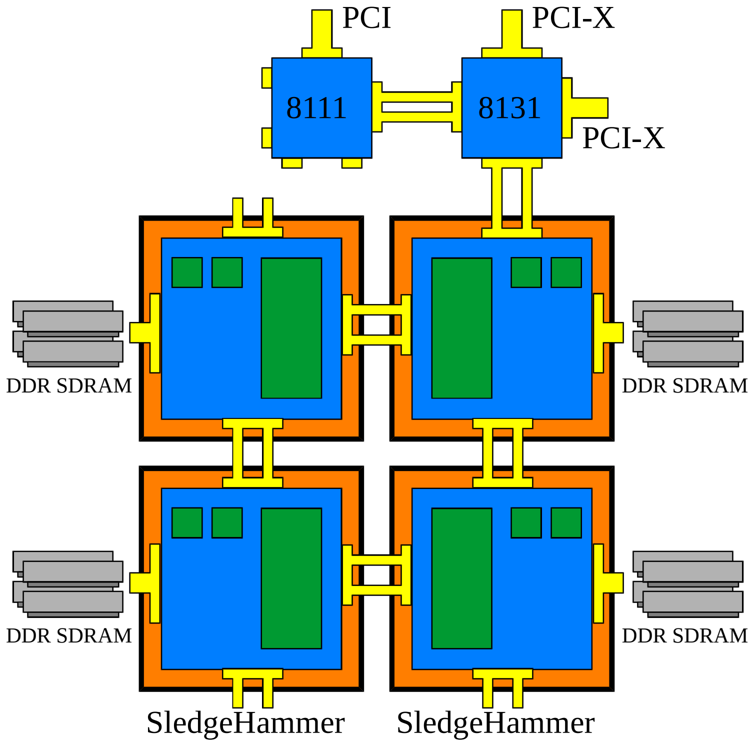
The full 64-bit support will be particularly popular in some of the larger workstation (or HPC) tasks, and there seems to be a lot of customer anticipation in this area, particularly from workstation users. This will require OS and compiler support, which appears to be good so far, though application support will take longer. The crucial difference between Hammer and Itanium is that current x86 customers can use all their existing applications at full performance. In addition, software vendors can support the 64-bit side much more easily - by issuing "patches" and alternative libraries to their normal 32-bit versions for example.

AMD 0.13µ 4 and 8-way Opteron servers

Expected: Q2 2003 or Later

Introduction: Since SledgeHammer has 3 HyperTransport links integrated into the core, as well as a main memory connection and logic to handle multi-processing, 4 and 8-way systems don't need anything extra.

Chipset Architecture: 4 and 8-way systems become a network of connections, with requests and data now having to travel through 1 or 2 CPUs, which adds to latency. The alternative to such a setup, apart from a bus, would be to have a central switch ASIC which is connected to every CPU. This means all connections between CPUs always take 1 "hop", instead of being variable, but such a switch ASIC would become very expensive very quickly - practical for 4-way, but not for 8-way, at any reasonable cost at least. I/O ASICs would be connected to spare HyperTransport links as before.



Bandwidth and latency: For each CPU, when it has a cache miss, it first does a cache snoop of all the CPUs in the system, and if no hits are found, it issues a main memory request, locally or remotely. With a 2-way design, the snoop request and data would be over just one HyperTransport link each way, with 4-way the worst case is to go over 2 links, and 3 links for 8-way. As a snoop request doesn't complete until all CPUs can respond, the snoop latency increases with the maximum number of hops it takes for the request to travel over, so overall memory latency gets steadily worse. Local memory latency in a 4-way system is around 100ns and remote memory (on a diagonally opposite CPU, i.e. worst case) is around 150ns. Latency at 8-way would probably be 10-15% worse.

Predicting bandwidth is a bit trickier - maximum system bandwidth increases with every extra CPU (21.6GBytes/s for a 4-way system with PC2700 memory), but so does usage as there's more CPUs making requests and more hops needed for any signals. Like with the UltraSPARC-IIIi systems, local memory optimization would help increase bandwidth efficiency. However, the 4-way Opteron systems will probably be able to sustain more bandwidth than 4-way UltraSPARC-IIIi systems, which are limited by the shared 3.2GBytes/s FSB.

Scalability: Though absolute latency is better than a northbridge style system, scalability is worse because latency (for a lightly loaded system anyway) gets steadily worse, while for a northbridge style it would be flat. However, a northbridge style system (or shared bus system like UltraSPARC-IIIi) can start to run out of bandwidth on the FSB, which can really hurt performance, like with the Pentium 3 Xeon and the SPECcpu_rate tests. However, most server benchmarks seem to need much less bandwidth than SPECcpu.

Summary: The SledgeHammer system designs are quite different to anything else today, making it difficult to predict real world performance. The design does look good on paper, particularly from a cost perspective as the same CPUs, memory, I/O chips (all the significant chips) can be identical for 1-way to 8-way, which gets better economies of scale. Though performance and price/performance may well be very good, the cache size is a bit small for 8-way systems and by the time SledgeHammer is out the main competing 8-way systems will have 2MBytes or more of cache, and for higher-end systems, cache size is taken very seriously. 2MBytes of level 2 cache would add perhaps 20-30% to costs, but such CPUs could probably be sold for 50% more, for 4 and 8-way systems at least. However, it would reduce demand a fair bit for the lower-end SledgeHammer systems, which is more price sensitive. The more demand AMD have for SledgeHammer (making use for 1-way and 2-way systems for example), the more cost effective it can be, due to economies of scale.

However, the real problem in particular 4-way and 8-way SledgeHammer systems face, is getting support from the top-tier OEMs - as they have the vast majority of the 4-way and larger server market. Though some white box manufacturers, and specialists, would likely ship such systems, they account for only a tiny amount of the higher-end market, though a growing amount of the 1-way and 2-way market. AMD claims to have a top-tier OEMs lined up for 4-way SledgeHammer systems already, which would at least be useful for doing a reference system, and also benchmarks - AMD will definitely need to show several serious server benchmarks.

On the software side, Hammer already has OS and compiler support in Linux and GCC, and the same from Microsoft for Windows and their compilers. AMD will absolutely need 64-bit OS support in Windows (to clear the 4GByte main memory barrier that 32-bit has) to stand any chance of succeeding in the 4-way and above market - a typical system with 8 CPUs would likely have 16GByte/s or more of main memory. If sales become decent and good compiler support is available, then AMD can start to attract ISVs support for the x86-64 instruction set too. Only when SledgeHammer has good OS, application and top-tier OEM support will it stand much chance of affecting Intel's server sales.

References:

- Hammer architecture white paper
- Hammer presentation from Microprocessor Forum
- Opteron systems presentation

Other New CPUs in 2002 and 2003 (All Markets)

Apple 0.13µ PowerPCs

Some rumors suggest that the PowerPC G4, which just reaches 1GHz at 0.18um, will reach 1.5GHz at 0.13um later in 2002, though this may be optimistic for 2002 unless the core design has been tweaked significantly. The 0.13um G4 is expected to have 0.5MB on-die level 2 cache, and support for up to 4MBytes of external cache. The new workstations will certainly support DDR SDRAM given that the servers do.

According to rumor, in early 2003, the PowerPC G5 will start shipping, and having a longer pipeline than the G4 (the current low-end embedded only G5s have a 50% or so longer pipeline) it is expected to run at higher clock rates. 64-bit support is also expected as well as RapidIO interconnect technology, which is similar to HyperTransport. However, from the available documentation on the G5, the core doesn't look particularly aggressive - 2-way issue out-of-order execution, and a similar cache structure to the G4. Perhaps the high-performance versions intended for Apple will be rather different though. In addition, Apple currently doesn't appear to be preparing for the "blade server" market, even though the PowerPC (which is more of an embedded CPU) is particularly well suited with its low power consumption.

AMD 0.13µ 2nd-generation Athlon MPs "Barton"

Expected: Early 2003

The desktop versions are expected by the end of 2002, but the Athlon MP versions will follow a bit afterwards. The only certain difference over the initial 0.13um "Thoroughbred" versions is that Barton has twice the level 2 cache (512KBytes), though it may well have a faster FSB (probably 166MHz DDR) and use PC2700 memory.

Fujitsu-Siemens 0.13um SPARC64-V

Expected: Maybe in 2003

Summary: SPARC International is a non-profit organization setup by Sun and others in the early days of SPARC, and a license to create a compatible architecture costs under $100. However, of the SPARC vendors, apart from Sun only Fujitsu-Siemens is still actively developing high-end systems. Their next design is the SPARC64-V, which is expected to run at 1GHz on a 0.13um process, and takes a rather kitchen-sink approach to features - 8-way super-speculative out-of-order issue with trace cache, three levels of cache an so on. Details are scarce but there was a presentation at Stanford (90 minute video available), which had a lot of details.

HP 0.13µ PA-RISC 8800 "Mako"

Expected: Sometime 2003

Summary: Based on the PA-RISC 8700 design, it has two CPUs on one die with 0.75MBytes of instruction and data cache per core, and a shared 32MByte external level 2 cache, and running at up to 1GHz. It also uses the same FSB as the Itanium II so HP will do systems that support both their PA-RISC CPUs and the Itanium. There is expected to be a PA-RISC 8900 as well (probably a 0.09um version of the 8800) but this will probably be the end of the line for PA-RISC, and HP plan to migrate customers to Itanium systems.

HP 0.18µ Alpha EV7 "Marvel"

Expected: Sometime 2003

Summary: Based on the current EV6 core, the level 2 cache and system components are completely different. Overall, it's very similar to the Hammer systems, each CPU having on-die memory controller (8 Rambus channels for 12.8GByte/s per CPU), 4 pairs of uni-directional connections to other CPUs and 1.75MBytes of on-die level 2 cache. Given the initial version will be on 0.18um, the die size will be large, making systems rather expensive. Clock rates are expected to reach about 1.2GHz with power consumption up to 150W. The main target markets are HPC systems - on the low end (2-8 way) I think competition from 64-bit higher-volume lower-cost 0.13um CPUs with high memory bandwidth will provide tough competition. The EV7 is expected to be moved to 0.13um eventually but development is being wound down - HP plans to replace Alpha systems with Itanium ones.

Intel 0.13µ Xeon MP "Gallatin"

Expected: Early 2003.

Summary: The 0.13um high-end Pentium 4 Xeons will feature 2MBytes of on-die Level 3 cache, run at 2-2.2GHz, will support Hyperthreading, and is intended for 4-16-way systems. It seems Intel itself will not be doing chipsets specifically for 8-way and 16-way systems, but is working with vendors such as IBM and chipset companies like ServerWorks to help support such configurations. From some of the proposed solutions, it seems the 8-way and 16-way systems will be use a 4-way base chipset with a 128-bit wide DDR SDRAM memory system. 8 and 16-way systems using 2 and 4 of the base chipsets respectively, connected together by a custom bus or set of point-to-point connections, rather like the Itanium systems. Though some of the third-party chipset solutions are interesting there are rather too many (and not enough information) to examine here.

Intel 0.13µ Itaniums "Madison" and "Deerfield"

Expected: Around Q2 2003

Summary: Madison and Deerfield, which is the McKinley core with a 6MByte and 3MByte Level 3 cache respectively, are expected to run at between 1.2 and 1.6GHz. The CPUs will be purely an upgrade, and no change to the FSB is expected, though naturally second generation chipsets and systems will likely have changes to I/O and scalability beyond 4-way.

IBM 0.13µ POWER4

IBM is expected to start shipping 0.13um POWER4-IIs later in 2002, which is expected to mostly be just a faster, cheaper POWER4 running at 1.5-1.8GHz. Shortly after the previous section of this article was published (on current 2-8 way systems), IBM announced some 4-way 1GHz POWER4 systems. For a similar configuration (same CPU count and amount of main memory) the IBM systems are a bit more expensive than same sized Sun UltraSPARC-III systems, though the former have more disk storage. It will be interesting to see how aggressive IBM get with POWER4 systems at the low end, the main competitors for them being Sun UltraSPARC-IIIs and Itanium 2.

SGI MIPS

It's not entirely clear what SGI's long-term strategy is, but they're planning at least one new core, the MIPS R18000, planned for 2003. It will run at up to 800MHz on a 0.13um process, with 2 full floating-point units. For more information, see the presentation at HotChips 2001.

Sun UltraSPARCs

Summary: Sun give little details about their roadmaps, but do drop enough hints - the UltraSPARC-III is expected to reach 1.2GHz or more at 0.13um. Most comments from Sun suggest that the UltraSPARC-IV will arrive in 2003, starting at 1.2GHz, eventually hitting 1.5GHz. It has two UltraSPARC-III cores on the same die, and will likely come with 16-32MBytes of external cache. Some of Sun's technical documentation says that the UltraSPARC-IV will use the same "Fireplane" chipset as the UltraSPARC-III, meaning it's likely to be an upgrade, and the proposed CPU speeds are all multiples of 150MHz, just like the UltraSPARC-III. With 2 cores running at up to 50% faster than current UltraSPARC-IIIs, maximum core performance would be 3x higher. Scalability is mostly about having enough system resources, and with UltraSPARC-IV putting much higher demands on the system, scalability will likely be a fair bit lower, but overall performance a lot higher, making for a nice upgrade.

Sun's roadmaps also show an UltraSPARC-IVi probably appearing in 2003, and at speeds of up to 1.6GHz or more, though it may be a 0.09um part. It seems quite unlikely that the UltraSPARC-IVi will be dual core, and is more likely to be an incremental upgrade over the UltraSPARC-IIIi - perhaps with more on-die cache or faster memory. Sun also have a MP capable CPU planned in their "embedded" range - below the UltraSPARC-IIIi, but no details are known.

General Aspects of the Market

This final section deals with less technical, and more operational and managerial aspects of owning and using workstations and servers in particular. This section is also less concentrated on "volume" systems, and rather the whole workstation and server market.

What Customers Want and Need

First and foremost, customers need a solution for their particular problems, which means having software applications that can help solve them. Likewise, a VCR isn't much use without tapes and a games console isn't much use without games. For most situations, it's necessary that the applications be supported properly, as well as simply be available, otherwise customers can be left in a bad situation if they have a problem, which can be very costly. This is not always so simple even for well-established platforms - sometimes-new hardware requires the most up to date version of a particular OS to run it, and often applications are slow to get qualification and support for the most recent OSs, even if little or no changes are required. For example, Sun's UltraSPARC-III is only supported on Solaris 8 and beyond. Also, IBM's POWER4 needs the latest version of AIX which also requires a recompile, and because of that Oracle won't be supported until May or June 2002, 9 months after the launch. Finally, for a large software suite like high-end databases or engineering applications, there are often high-end modules that are only available for particular platforms.

At the higher-end, good reliability and availability is vital for "business critical" (company may go bankrupt if it's not good enough) and "life critical" (people may die if it's not good enough) situations. If the solution isn't good enough and this doesn't look likely to get solved soon enough, the customer will quickly start to look elsewhere - because the cost of problems can easily exceed the price of the hardware. It's not uncommon for customers to deliberately use older hardware or software because it will have been tested and patched a lot more - more "proven". Some vendors will offer claims of high reliability, but include a large number of get-out clauses in any agreement, generally making it worthless, which just tends to cause ill will between the customer and the vendor. Vendors with bad reputations will be avoided - I've seen customers refuse 100% discounts for systems because the previous purchases were consistently unreliable. It's not uncommon for customers to over-specify uptime requirements though.

Providing good quality service and support is a lot harder than it might seem, on the high-end certainly. Real enterprise and data-center customers will expect the vendors to provide real support, which can mean getting engineers on-site within a few hours of a problem occurring to try to diagnose and fix the problem. It's also becoming increasingly common for higher-end systems to have a dedicated line to "talk" to the vendor for remote diagnosis and fixing. For major applications some systems vendors will provide joint service centers with the application company, giving customers "one throat to choke" if anything goes wrong - customers get very irritated when the systems vendor and application vendor blame each other, and providing a joint support facility helps give customers peace of mind, as well as better support. Customers very much appreciate good support - it can seem a bit perverse but a company that has first-rate support but second-rate products may be better appreciated than a company with second-rate support and first-rate products.

Performance is important of course - the faster the better, as time is money. However, the system has to be fast enough for the customers' particular situation and doing well on benchmarks but poorly in real situations is not useful to the customer. Still, a common use for benchmarks by customers is to help eliminate systems with consistently poor performance. Scalability is secondary to raw performance but could be viewed as performance over the long term - as time goes by, the complexity of the application and size of the data will tend to increase, demanding more performance from the system. More scalable systems will also be able to use upgrades better, prolonging the system's usable life.

Other significant factors customers take into consideration are compatibility with older existing systems, multi-platform interoperability and the general reputation of the vendor. Vendors need to make customers aware of their products of course, which leads to advertising and also having a large sales force.

There are other factors that can override some of the above though. Probably the most common reason for a customer to buy from a particular vendor is "because we bought from them last time". This saves costs of re-training, testing, porting and so on. However, there are political reasons as well - it's common that whoever made the previous purchase recommendation (in the customer's internal order process) will make the next one, and if they recommend something different then that suggests that they made a mistake the last time. Another common way of choosing is to simply choose whatever system is the most popular (or perceived to be the industry leader) for their particular problem, or to take what's considered the option with the lowest risk, which is often the industry leader. A lot of high-end enterprise systems look like this: database application from SAP, PeopleSoft or Oracle, running on Sun (database hardware king) and Oracle database (SQL king), with EMC (storage king) with SAN (Storage Area Network) switches from Brocade and LAN (Local Area Network) switches from Cisco.

It's also common for customers to accept the hardware vendor their chosen software vendor recommends, for the main application. For example, for a long time database application companies like SAP recommended Oracle as the database vendor, and Oracle's recommended hardware vendor is Sun. Oracle and Sun have had a close working relationship for a long time, and Sun concentrate more on the database market than just about anybody which is probably why they have the best ISV support of the commercial Unix companies. It's worth noting that once Oracle started offering competing database applications (instead of just database server software), companies like SAP started recommending IBM's DB2 instead.

Putting Costs Together - TCO

There is much more to costs than the initial purchase price - Total Cost of Ownership (TCO) is much more important. TCO can cover many things including service and support costs, networking and storage costs, replacement costs, upgrade costs, administration, training, migration costs, electricity costs and harder to measure costs such as productivity or the cost of any downtime. TCO is basically "the big picture" for the costs (and potential costs) for owning a system for its operational lifetime. For example, a number of companies buy Apple PowerPC based portables despite using PCs for desktops, because the Apple portables last longer and cost less to support. As a general rule, the larger the system (or solution) the less important the initial purchase price becomes.

Customers may well attempt to summarize their purchase decisions as "which solution that meets our objectives has the lowest TCO" or "for a given TCO which solution has the greatest benefits". In practice, there are a number of problems with this though, the main difficulty is that the future is always uncertain, and estimating how often problems may occur and of what magnitude is nearly impossible. Going to great lengths to estimate TCO for cheap systems is also rather pointless - the cost of estimating the TCO itself can get quite high. For these reasons, many companies don't really bother with a formal analysis of the costs/benefits of a particular purchase, particularly smaller ones. At large companies, significant purchases may require a full TCO analysis to get approval.

Estimating TCO for a potential project (or calculating it for an old project) can have some quite subtle elements. For example, that old Unix favorite - the command line, compared to GUI based administration tools. Many would accept that GUI tools can present information more clearly, make it easier to make configuration changes, and are basically quicker to learn. However, Unix style configuration files have several advantages - those configuration files can easily be copied to many other systems (reduces scope for error as well as reducing time to make changes), can be put in revision control systems (reduces scope for errors) and it's much easier to find changes or differences. With a lot of administration done by command-line, it's much easier to do this remotely (using SSH - secure access to command lines on other systems with SSL encryption), and an increasing number of packages have GUI based front-ends to the configuration files as well. With the X11 protocol that most Unix systems use for graphics, it's much more practical to have a remote GUI interface as well - the Windows equivalents have limitations. With many large websites using a large number of low-end servers, reducing administration costs is critical, and the more automated administration can be made, the fewer staff members are needed, and the better service can be. Naturally, any such features are only useful if they're used, which is why training and support can make such a big difference.

Another thorny issue is reliability and uptime. Most servers and most recent operating systems can manage to get 100% uptime. How easy this is to achieve is quite varied as some operating systems (particularly older versions of Windows) are quite brittle, but a problem all OSs face is reducing the number of situations in which a reboot is required, and the more configuration changes or software fixes/additions (including bug fixes and security patches) that can be made "live" (without requiring a reboot) the better. This is a major cause of "planned downtime", which is often much larger than "unplanned downtime" (system rebooted or became unusable by itself). Better system design (OS and hardware) can improve both types of downtime, but overall the main cause of downtime is that the administration isn't good enough. An increasing number of vendors offer services to configure systems (including software packages and storage) to a high quality level - the better the initial configuration, the better the system will run in the long term.

In the end though, customers like "easy" choices - going to all the effort of estimating TCO for current systems and again for potential future systems is a lot of effort, and quite hard. On the low-end (1-4 way systems), the price and CPU performance of x86 systems has been very competitive for many years, and support works well enough with a mass-market approach for most of these systems. It also helps that most corporate desktops use Windows, which makes development and administration simpler for the servers - fewer platforms to support. With a range of similar systems available from many vendors, customers have little fear in being locked to a single vendor, so x86 systems have become an easy choice at the low-end. However, the same approach is almost the opposite of what's needed at the higher end (the real "enterprise" space), where initial purchase price (particularly of the CPUs) is a minor part of TCO, where system performance is more important than CPU performance and where customers require real support.

Vendor Vs Vendor, Customer Vs Vendor

"List prices" are used for price comparisons, but actually for any significant purchase order, the customer can almost always negotiate with the potential suppliers, whether the order involves a large number of cheap PCs or a few large servers and a ton of storage. The customer will often make the suppliers bid against each other, so that they can get a better price. For customers, it's simpler to have fewer architectures for production systems, but many make sure they have at least two, particularly for higher-end systems, because then they don't get dependent on a particular vendor.

The whole game between customers and vendors can get rather political at times - some customers will threaten to issue press releases saying they'll drop a particular supplier if they feel strongly about a particular issue, which can be very effective for large customers. It's also fairly common for customers to demand that they can get memory upgrades from alternative suppliers (often Kingston) so that they're not dependent on a single vendor. Customers will also "threaten" to move to a cheap alternative to help push prices down.

Vendors like to use customer success stories and wins to show how well they're doing, which helps convince new potential customers that they're "for real", and will give customers a discount to get approval. Overall, in most cases the customers have significantly more power than the vendors. In cases where customers have little choice but to buy from a single vendor, even if the vendor abuses them somewhat over prices or licenses, they will continue to buy in general. But they won't forget and a lot of ill-will gets built up, so if a viable alternative appears, sometimes the market can change unexpectedly rapidly.

Different Market Categories

To the market research companies, the "high-end" for servers is for systems typically costing over $1m (which covers 25% of the whole server market by revenue sales), the "mid-range" is $100,000-$1m, and the "low-end" is for systems typically costing under $100,000. Systems that are clearly "enterprise" level (capable of running mission critical applications for large enterprises) could be vaguely said to be those costing over $100,000, though the minimum feature and performance requirements increase every year. Marketing also has its fun with the terms "enterprise" and "data-center", but the general idea with the "enterprise" label is for solutions that are well suited for large companies in general - e.g. the top 100/500/1000 companies in a particular country or the whole world. For Sun, they classify their 8-way Sun Fire V880 as "entry level", while Intel and x86 vendors would classify 8-way systems to be "enterprise" and 2-way to be "entry level".

Systems companies like to "position" their products against competing products (including their own) based on price, performance, features and marketing. However, customers will choose according to their own needs, and basically the market as a whole decides what products "compete" with each other, not system sellers. So a customer might well choose a 2-way high-spec system over a 4-way lowish-spec system, or vice versa.

The Blessing and Curse of Backwards Compatibility

Having a next generation OS, application or hardware platform that is highly backwards compatible with the previous generation is useful to customers and ISVs as it significantly helps software availability and support, as well as reducing the costs and risks of migrating. Though in theory there should be no problems running older applications on a new OS with supposedly full-backwards compatibility, it is a lengthy process to fully check and certify that this is the case for both ISVs and customers, particularly for enterprise applications as the cost of any mistakes could be huge. Customers with business critical systems and ISVs will often have long beta and testing programs before moving to any new platform. The worse backwards compatibility is, the slower ISVs and customers are to move to the new platform. Putting it another way, customers resist changes (particularly ones they feel are unnecessary) and the bigger the change (which includes a "leap of faith" when the changes are recent), the more customers will resist them.

For any system, there is little point in upgrading it or replacing it with a faster new one unless the higher performance would actually be useful (profitable). There's only two reasons to upgrade if performance is "good enough": that support is being dropped for the current system (which really annoys customers) or server consolidation - replacing many small servers with a few big servers, to reduce running costs. For customers migrating an existing project to a new system, the fewer differences there are between the old and new system (hardware, operating system and applications) the less the chance is anything will go wrong, even if in theory nothing should go wrong due to backwards compatibility. If the new system isn't compatible (on the hardware, OS or application side), then the project will have to be "ported", which adds to costs. For a new project, backwards compatibility still has some indirect issues, with regards to training and experience (of developers and administrators) and interoperability with other projects and systems. So for customers, backwards compatibility reduces costs and risks. For vendors though, backwards compatibility adds to costs (more development and testing) and can often get in the way of speed optimizations and feature additions.

There is a tendency for ISVs to only "port" the latest versions of their products when certifying products for a new OS release, as this reduces complexity and hence costs. New hardware, particularly new CPU features, is often only supported on the latest OS releases as well. These both tend to slow down migration of customers and ISVs to support and use new hardware and OSs, even if in theory it's simple to migrate. For customers it's simpler, cheaper and less of a risk to have as few variations in hardware and software being used as possible, particularly for higher-end projects, so in general they are reluctant to jump to new platforms and new software versions. They will also often delay migrating to a new OS, application or hardware generation until they're ready to migrate all production systems, sometimes even skipping a whole "generation".

Windows 2000 (a significant improvement on NT 4) and Solaris 8 (a decent improvement on Solaris 7 and also free for 1-8 way systems) were both launched early in 2000, but the majority of all customers' systems are still using older OS versions, 30 months later. Almost all new systems come with the new OSs though. MacOS X launched early in 2001 but many major applications have yet to be ported or have only just become available. Desktop systems are generally the quickest to change (fairly low risk) though if user management and similar servers need to be upgraded significantly to support this, it can become slower. Workstations and low-end servers tend to be next with paranoia increasing as the system size (and generally risk) increases. However, for the highest-end systems for a particular architecture or product range, there's often some customers (business and HPC) who desperately want more performance, and they migrate quickly.

Microsoft was doing DOS with x86 in the early 1980s, while Sun and Apple started with 68000 CPUs in a similar time-frame. Sun started moving to its in-house SPARC design in the late 1980s - Sun was very much a workstation vendor at the time and this migration worked well. Sun moved to Solaris (from SunOS) in the early 1990s (which was quite unpopular early on) and moved to 64-bit in the mid-late 1990s, while providing full backwards compatibility for 32-bit applications. The other high-end RISC vendors who started at 32-bits also added 64-bit extensions in a backwards-compatible way. Apple moved to the PowerPC platform in the early 1990s (which went pretty well), and a significant OS change with MacOS X in 2001 with reaction somewhat mixed so far. Microsoft began pushing Windows on the consumer side and Windows NT for small business tasks in the early 1990s, but it wasn't until 1995 that Windows started to become significant for consumers, and it wasn't until 2001 that Windows NT really became a consumer OS, with Windows XP.

Entirely new OSs and CPU architectures are both fairly rare, but it's interesting to note that the major general purpose desktop and server OSs today all started about 10 years ago. The successful general-purpose CPU architectures all started in the 1980s or earlier, except the PowerPC from the early 1990s, though that's based on IBM's POWER architecture. An entirely new OS can at least be installed on old architectures, so its potential market is large, while a new CPU architecture has to wait for new buyers and builds up volume slowly.

For the RISC architectures from the 1980s, the companies designing the new RISC CPUs had low volume anyway, and mostly went for the workstation market which has a lot of demand for more performance, particularly floating-point performance. The x86 architecture started on the low-end which gave it a volume disadvantage. The Apple PowerPC systems could also handle old 68000 code faster than real 68000 CPUs, so software availability wasn't really a problem. The Alpha was 64-bit from the start in 1993 and had great performance, but was an entirely new architecture, and failed to get enough ISV support, which is possibly the main reason why DEC couldn't convert their success with mini-computers to the Alpha.

For those vendors migrating from CISC designs to RISC designs, it was often a move from 32-bit CISC to 32-bit RISC, where the RISC CPUs also had a significant performance advantage at the time, the systems were mostly workstations (which is simpler), and where all levels of the platform were migrated at the same time. However, for Itanium, the OSs and systems are provided by completely different companies, with comparatively little performance difference, and with Itanium's main extra feature being 64-bit addressing, which is a higher-end and lower-volume part of the market, though becoming more necessary every year.

Companies are always worried about taking risks with critical applications, and new platforms are always a risk. Lower-end systems can support higher risks but it seems likely that Itanium will have few advantages over existing 1-2 way systems for some time, which means small volumes, and volumes are a good way to attract ISVs. If Itanium could offer competitive x86 performance, it would have a much better chance of success as customers wouldn't need to wait for significant ISV support.

Back when x86, SPARC, POWER, PowerPC, MacOS, Solaris, Windows NT and Linux first came out, the computer industry was much smaller, with many more competitors, with fewer and less complicated applications. Though new platforms have been created since then, they have mostly been evolutionary - backwards compatible. MacOS X is a rare recent exception, but given that almost its entire market share is in applications that don't have significant backwards compatibility issues (low-end workstation applications), it's hard to use as a reference.

After the Internet "Gold Rush"

Particularly in the last 5 years, the Internet caused a "gold rush" that strongly affected the server and server software markets. For example, several years ago the market for 1U servers was almost non-existent. But in 2000 the market for rack-mount servers in general became a significant slice of the overall market, though was probably down about 60% or more in 2001. This market is still very active though, with products and demand for even denser "server blades" is just starting to warm up.

There was also a huge explosion in demand for ever more powerful and flexible web-servers starting before this, mostly to run on dedicated single and dual-processor servers. For more complex, and particularly with e-commerce orientated sites, came database servers, which were then accessed by programs running on the web-servers, or by an application server. The demand for large servers with large databases also spurred the growth for dedicated storage systems.

The demand for custom web-server applications (websites) also spurred a rush by developers, for the generous pay packages being offered by companies keen to get to market quickly. Developer platforms, software and experience also became critical - most companies wanted to get their websites to market as quickly as possible, so there was little time for training. Back in 1995, the Unix OSs had a significant head start over Windows, as they came with a TCP/IP stack as standard, and also some typical internet software like e-mail (servers and clients) and software that made remote administration fairly easy (telnet and SSH).

Apache became the web-server of choice before too long on most Unix OSs (being free, stable and well supported by additional packages, if not the most scalable) and to a large extent drowned out the competing Unix web-servers, despite many of them having commercial backing. Sun were an early mover on developing "internet applications", but perhaps their biggest advantage was that they had been supplying rack-mount servers to telcos for quite some time (their "Netra" range), and were perhaps the first major company to bring out a 1U server, using the UltraSPARC-IIi. When demand for rack-mount servers for the telco and service provider markets exploded, Sun were in a lead position, and had 30% of the single CPU 1U market in 2001, but didn't have a low-end MP-capable UltraSPARC CPU, and pretty much entirely missed out on the 2-way 1U and 2U markets.

Microsoft were somewhat late with developer tools and server software, suffered from a software cost disadvantage against free OSs like Linux and FreeBSD on x86 hardware, had little in the way of remote administration and also Windows NT4 was rather too keen to reboot. However, with many OEMs already selling large numbers of cheap x86 systems, it didn't take them too long to start selling rack-optimized servers with Windows pre-installed. Microsoft significantly improved the OS side with Windows 2000, and the developer tools and server software were also significantly improved by then. Still, by some surveys, only 25-30% of web-servers are running Microsoft's web-server.

On the other hand, Microsoft benefited greatly by the "network effects" created by email for transferring documents, as with Microsoft Office documents being heavily exchanged, it became important for companies to have more up to date software. In particular, companies increasingly needed software that could read Microsoft's file-formats reliably, which basically meant Microsoft Office.

When a market is new and growing fast, there is no "installed base", so the main concerns for any customers wanting systems and software is what hardware, OS and main software packages to use - they don't have to worry too much about fitting it in with what they currently have. Any customer installations in such a new market will likely be in addition to existing ones, quite possibly requiring new developers and administrators, so companies don't necessary have to worry about training and user acceptance of new systems. All in all it makes it much easier for customers to go with a vendor they've never used before, and market share as a whole is less important, making selling easier for "new" vendors. This is why there was such a gold rush - the market was expanding rapidly, and the earlier and quicker a company could build a presence, the more likely they could become a significant player.

With the Internet "gold rush" over, the relative market positions will become more stable, like the rest of the computer systems market. Once established, customers change between systems quite slowly, even to next-generation OSs, hardware and software from what they currently have. Moving to a completely different architecture is even rarer, because it's more of a risk, the initial cost is generally greater than an upgrade, the cost of migrating the software and data, and most importantly, re-training the users, programmers, and administrators. The people side can be particularly difficult, particularly when changing "cultures" - the Windows people don't like the Unix people, while the Unix people don't like the Windows people, and the mainframe people resist changes most of all. Turf war, basically.

Some Future Trends

After the rush for 1U servers, the next-generation focuses on "server blades" - putting most of the components needed for a low-end server on a single board, and have those boards plug into a rack-mount chassis, generally increasing CPU density by a factor of 3 to 4-fold. Despite the downturn, the market for high-density systems is still more than large enough to fund the research and development needed for new setups, though in some ways the hardware side is fairly easy. The software side becomes ever more important though - the better the systems can be managed; the more costs can be reduced due to the higher efficiency. Such systems may also integrate functions like load balancing that were traditionally done by separate systems.

Meanwhile, clusters are back in fashion, though it's still mostly talk, and not action, on the server side at least. Compute farms have become more popular for HPC tasks in general, and with all high-density systems for web-farms, this helped the compute farms become cheaper. Peer-to-peer, grid computing and distributed computing in general are certainly making more headlines recently.

However, using clusters to improve performance is hard in general, particularly for applications such as databases, which are particularly dependent on having low latency between all the CPUs in the system for many tasks. Connections between systems in a cluster like Ethernet, have poor bandwidth and latency compared to the interconnect between CPUs in large servers. Even newer technologies like Infiniband aren't comparable - the latency issue is particularly hard to deal with in a general-purpose connection. To some extent, clustering is rather like an extreme case of a NUMA architecture, and actually in NUMA architectures, the interconnect used between CPU/memory boards is often a bit like a network connection, particularly for SGI's servers. Whether network-like connections can ever compare with low-latency UMA designs in both latency and bandwidth is quite uncertain, and doesn't seem likely to be achieved soon.

The potential value of database clusters for performance depends pretty much entirely on how much cheaper a group of four 4-way systems is compared to a 16-way system, for example. Compared to a single 4-way box, a cluster has extra costs in having a special interconnect between them, cluster server software (which is often more expensive than the non clustered equivalent) and is more complicated to setup and run. Depending on the solution, the storage aspect can also be rather more complicated compared to a non-clustered solution. Worst of all, database clusters have some quite ugly failure modes, significantly increasing the cost of any problems.

Since the performance will almost certainly be lower (compared to a single box solution with similar CPUs), the cluster solution relies entirely on costs, so the lower the premium on 16+ CPU servers becomes, the lower the advantages cluster solutions have. It seems that by 2003, 16-way Intel based servers will become much more common and price/performance competitive compared to 4-way servers, reducing the cost benefit of clusters. Also, with Sun doing products like the Sun Fire V880 and dual-core UltraSPARC-IV, and with IBM's low-end POWER4 systems, it seems unlikely clusters are going to make much headway into the database market in the next few years.

There is more to clustering than performance however - there are many situations in which having a highly reliable low-end server is very useful, and clustering for redundancy is a simple way to get very high reliability with simple servers. With blade servers and high reliability clustering software becoming more common, this market will probably grow significantly in the next few years.

There are other combinations as well - clustering for both performance and reliability. A group of web-servers for a single website isn't a real cluster (normally) as they aren't "aware" of each other - no direct communication between them, unlike in a database cluster. If a single box crashes or is rebooted (or added) it doesn't affect the others directly, except in how much of the load they share. However, being able to share data between the web-server applications is useful both for more advanced features and for performance. Some more advanced web-servers already support clustering for performance and reliability with shared data between the servers.

Looking Ahead

The remainder of this article takes a brief look at the status and issues affecting the future of major companies involved with CPU, system or OS design.

AMD's CPUs

For the rest of 2002, AMD's success will depend almost entirely on how well it migrates its Athlon production to 0.13um. Although the Athlon MP has generated significant amount of interest, it has yet to get a design win from a top-tier OEM, and it's unlikely to make much difference to AMD's finances in 2002. However, to some extent it helps pave the way to ClawHammer DP and SledgeHammer, as well as higher margin CPUs in general. As a replacement for current Athlon CPUs, AMD's success with ClawHammer will depend on both its design and manufacturing, and to some extent on the infrastructure support in terms of chipsets and motherboards - given the support for the Athlon XP, this aspect shouldn't be a problem for low-end systems.

SledgeHammer will be a bit of a first for AMD since the chip itself will be a comparatively low volume product - the Athlon MP core is the same as the Athlon XP core, which is a mass production chip. Also, to succeed against Intel, SledgeHammer will need non-trivial OS, compiler and ISV support, as well as physical infrastructure support - motherboards and OEMs. The biggest problem for AMD is that the vast majority of Intel 4-way and higher systems are from a relatively small number of OEMs - the "top tier" ones. So it's vital for AMD to have a top-name OEM to push SledgeHammer systems, and AMD MP systems in general, and there have been some rumors about Fujitsu-Siemens who are strong in Europe, like AMD. Without significant backing from OEMs above the white-box manufacturers, SledgeHammer is likely to have relatively little direct impact on the market, though it could pressure prices, but for AMD the extra costs and risks of SledgeHammer should be quite small as it just builds on ClawHammer.

AMD's longer term road-map in general is very much to take a "value" approach to its volume CPUs, using process shrinks (0.13um to 0.09um to 0.065um etc) and larger wafers (from 8-inch to 12-inch) to significantly boost the number of CPUs it can manufacture from its current fabs, which helps reduce costs significantly. So it seems AMD's plan for the mass market is to go for price/performance, though what they plan for higher-end markets is less clear.

Apple's Systems

On the hardware side, the PowerPC CPUs and chipsets, Apple is rather like an OEM for Motorola, but they do their own OS and software. Although Apple did benefit from products like the iMac, and extra demand for professional content creation for websites and multi-media in general, Apple's market share is still in the 5% region or lower. It's hard to determine just how much of a difference it would make to Apple if they could increase their performance for general applications to be similar to Athlon XP and Pentium 4 systems, while having similar prices. However, if Motorola continues to deliver PowerPCs to Apple, which are more like high-end embedded designs, then such a scenario is unlikely. For the moment, Apple seem somewhat content to target middle-income families, education markets, content creation professionals and people who don't like PCs or x86 OSs.

If Apple can have highly competitive CPUs in price and performance terms, then that would help increase their share of the workstation and server markets. However, Apple is not a "traditional" workstation or server company, and because of that the amount of workstation and server applications supported on PowerPC systems is small. On the other hand, with a combination of Unix plumbing and stylish front-ends, MacOS X has attracted a lot of new developers, which certainly helps.

An obvious suggestion made to Apple on occasion is to move to a different, more competitive, architecture. The big problem for Apple is not so much the end result, but that the pain and risk of converting could easily exceed the benefits. Apple did move from the 68000 architecture to PowerPC, but this transition was greatly aided by hardware features to help backwards compatibility, and considerable planning on the OS side. Migrating to a completely different architecture with no backwards compatibility assistance would require recompiling and testing all applications, unless Apple develop a highly sophisticated binary run-time translator, which is a lot easier said than done. Instead it's likely better for Apple to continue with PowerPC, perhaps by using its relatively small and cheap design to do dual-core processors for its volume systems for example.

HP's Systems

HP and what was Compaq are both planning to replace their own RISC systems with Itanium over the next several years, while in the process of merging, while also doing one or two more generations of their current RISC CPUs, all of which makes predictions a little tricky. The main reason for the merger is that HP is increasingly interested in services and support, as HP sees x86 systems becoming increasingly low margin, for them at least. Despite this, HP's PC divisions are solid, and both companies had good service organizations separately, and Compaq has particularly good manufacturing and storage, but it leaves them somewhat in the middle between IBM and Dell either way.

On the RISC and Unix side of things, reports from the those in the field indicate that many of HP and Compaq's current RISC customers are far from happy at being forced to transition to Itanium. This causes uncertainty and risk for customers, which is making is hard for HP to get into the bidding process for new business, particularly with new customers. Until the Itanium market becomes large enough and "proven" enough, HP will slowly lose current Unix customers, mostly to IBM and Sun, who are more than willing to help high-end customers migrate to their platforms, the large margins offsetting the extra costs, and both companies have put large migration programs in place. How quickly Itanium sales of systems running HP-UX take off will significantly depend on ISV support and customer confidence, and given the current situation, prospects do not look good. This could quite likely lead to HP losing their 2nd place position in the commercial Unix market to IBM. While the merger goes through, this will distract the combined company at time when it needs a lot of focus, and it's a telling point that none of HP's and Compaq's main rivals objected to the merger - they're cheered it on in fact.

IBM's Systems

IBM had a particularly good year in 2001, since it grew server revenue over 2001 in the US at least (worldwide sales fell a bit), while the revenue of all its major competitors fell. However only IBM's mainframe revenues grew (from reduced competition it seems) while its PC and Unix revenues fell, though less than its main competitors. Though IBM is planning to do large (16-way) Itanium and Xeon servers, it's also continuing with its POWER RISC CPUs, recently shown with their POWER-4 systems. They also recently decided to drop development for a port of AIX to Itanium, which would be a necessary first stage if they were ever planning to migrate customers using POWER CPUs to Itanium. For the rest of 2002, it seems IBM will mostly be introducing a 0.13um POWER-4 and lower-end systems that use it, followed by the POWER-5 in 2003 or 2004, which seems to be a heavy-duty SMT design.

For a number of years, IBM's revenue from servers has been flat or negative, and IBM were particularly affected by the "Y2K bug" with all their legacy systems. Their relative position has improved somewhat recently though, and while their highly diversified product line limits how fast they can grow, it does give them a certain amount of protection. Meanwhile, IBM's famous professional service division, IBM Global Services, is becoming increasingly envied, with companies like HP viewing making money from services as a better alternative than the often slim margins on x86 systems.

Intel's CPUs

For Intel, most of 2001 was about transitioning its lower-end product line to the Pentium 4, and 2002 will be mostly about transitioning its server line to the Pentium 4. 2003 will be mostly about finishing off the transition to 0.13um while beginning the transition to 0.09um. Though its new Itanium product line based around the McKinley will be making a lot headlines, it won't make much difference to Intel's revenue in 2002, or 2003. This is mostly because 89% of servers sold use Intel CPUs already and most of the money to be made on the higher-end is from systems not the CPUs. The biggest obstacles to the Itanium's success is its software support and its price/performance compared to its Xeon systems.

Though there are certainly plenty of uncertainties about the Itanium, Intel's 2-way and Xeon product lines are currently solid, with huge software and OEM support, as well as significant economies of scale. Part of the problem with IA-64 is it has to fight IA-32's massive momentum. If none of the large OEMs start pushing MP systems using AMD CPUs, then Intel's only significant competition for x86 MP systems will be Sun and (to a less extent) IBM's low-end RISC systems. But it would be very hard for either IBM or Sun to sell RISC-based systems to customers who just use Windows based Intel systems - customers are never keen to make a significant switch in operating systems.

One difficulty Intel will likely have is that for 2 to 8-way systems, neither AMD, Sun or IBM particularly rely on making good margins at this level (while Intel do expect good margins) and all 3 are being much more aggressive about this market than they have in the past. AMD has also improved its relative position in CPU design and manufacturing terms and this looks set to continue which gives Intel less leeway than they had before. Another difficulty Intel have is that Microsoft is causing a lot of pain to its customers with regards to support and license costs which is causing many customers (particularly government agencies) to consider alternatives. However, a common alternative OS for servers is Linux, which is almost always a x86 system.

Linux and Open Source Software

Given that open source software is free to modify and distribute as well as free to use, in economic terms it should lead to a market with "perfect competition" where no particular service or systems provider can gain any clear long term leadership. This would be because nobody can gain any technical advantage, and this would mean that competition would be intense but profits would be low. In reality however this is not the case, as "momentum" and "network effects" are significant in the computer industry, and for example, Red Hat is the leading Linux distribution for servers with around 85% of the market, and looks likely to continue. The market for servers pre-installed with Linux is rather like that for servers pre-installed with Windows, and today most Linux systems are sold by the big "Wintel" OEMs - they already had the manufacturing and distribution infrastructure in place. There are some rare cases where there is only one supplier for Linux on a particular platform, and for example Linux for IBM mainframes costs about $20,000.

In 2001, total sales of pre-installed Linux servers was $400m, a tiny fraction of the market, but growing fast. With slim profits available for software or hardware sales, the real action for revenue and profit for Linux (and open source software in general) is for services and support. This isn't so easy as one of the main requirements for service and support is to have lots of experienced engineers, particularly those with good OS experience and with skills in finding and solving complex problems. So any experienced "Linux kernel hackers" are going to be in strong demand.

An interesting question is how Linux's rise affects Microsoft and also the commercial Unix companies, particular Sun's Solaris which is the only major commercial Unix with a significant market share in lower-end servers. Linux really affects both - on the hardware side, most systems come from existing OEMs who build, sell and support Windows based systems, so are being sold as an alternative effectively, and this makes it simpler for existing Windows customers to buy Linux systems. On the software side, it's much easier to switch from commercial Unix OSs to Linux, though it's even easier to switch from Linux to commercial Unix OSs. Also, on the people side, programmers and system administrators who use Unix a lot would often be a lot happier to switch to Linux (or other free Unix systems) than to Windows, and vice versa.

In the enterprise market, Microsoft does not currently provide nearly as good support as the commercial Unix companies, due to taking a more mass-market approach, mostly concentrating on making money on selling software. This gives Linux companies an edge in focus as they will be concentrating on service and support as there's little money to be made on the software side, and will be particularly attracted to high margin enterprise customers. Against commercial Unix systems, the Linux specialist companies will only really be able to compete on price, almost certainly having poorer depth, breath and quality, though they may be used by customers to extract discounts from the commercial Unix companies. The commercial Unix companies are also much more paranoid about reliability and uptime that the Linux kernel developers, who are perhaps rather too keen to put in new features for their own good on occasion, which causes pain for higher-end users.

Microsoft's OSs and Software

Like Intel, Microsoft is in a very strong and secure position in the lower-end of the market - up to 4-way and to some extent, 8-way servers, and also workstations, portables, and desktops in general. However, while Windows 2000 is a very significant improvement over previous versions of Windows NT, Microsoft has yet to make any real headway into the enterprise market, partly due to lack of competitive servers in the mid-range market. On the technical side, Microsoft are also new to 64-bit OS and applications, and large systems in general, and haven't yet achieved the same smooth running that most commercial Unix OSs enjoy.

On the political side, Microsoft's software license prices are causing much discontent, and this is starting to be recognized at company board level now too, which means companies may be willing to take more unusual responses. Microsoft pushing customers to upgrade to the latest versions (which is also a common industry tactic), and Microsoft's rather mass-market approach to support and poor interoperability with other platforms (including older versions of their own software), is not helping either. This is not exactly causing customers to stop buying, or making them migrate to a different platform, but it's making customers more willing to do so.

Microsoft's attitude towards competing software providers is also why the US and European anti-trust agencies have become active in trying to curb Microsoft's monopolistic practices, though what actual action this may result in is unclear. Even if Microsoft was compelled to "play nice" this would actually likely lead to greater customer satisfaction in many areas, making it harder for competitors. Several government departments worldwide, particularly in Europe, have also been actively looking into alternatives to Microsoft's Office suite and also Windows, which may do more real financial harm to Microsoft than the anti-trust cases. Meanwhile, Microsoft is actively developing for and selling to lower-end (embedded) markets, new market segments (web services for example), new geographical markets (Asia for example) and the higher-end server market. However, Microsoft may suffer from a lack of trust when entering new markets.

Sun's Systems

For Sun Microsystems 2001 was mostly about transitioning its systems to the UltraSPARC-III (and UltraSPARC-IIe for the bottom-end) and also adjusting to the sudden drop in demand from service providers, financial services markets and "dot.com" firms, as well as the general drop in demand. Sun had done particularly well in these new internet related markets, and built up cash reserves of $6Bn because of it, so 2001 was a particularly difficult transitional year, both because it takes a long time for customer acceptance of new higher-end systems, and because of the rapid market changes. Meanwhile, Sun is more actively targeting more traditional parts of the server market that make up IBM's stronghold, and new markets like biosciences. In some ways, Sun is the opposite of IBM, putting almost all their efforts into SPARC and Solaris, and partner their professional services division with the major services companies, instead of trying to do it all themselves.

Sun have a solid lead in the Unix market (workstations and servers), which is bigger than the x86 or mainframe server markets in revenue terms. To a large extent, Sun are increasingly attacking every part of the server market, from server appliances and server blades at the low-end to $10m systems going against mainframes at the high-end. Apart from the Cobalt line (Linux x86), all Sun's systems are 64-bit on the hardware, OS and most applications, so there's no need for transitions, unlike their competitors. On the developer side, Sun is focusing on many middle-ware products (like web-servers and application servers in general) including its Sun ONE (previously "iPlanet") product line and also its Java development products and JVMs. With the large increase in volumes on the low-end, Sun also seem much more serious about developing for the low-end, hence CPUs like the UltraSPARC-IIe and UltraSPARC-IIIi, and based on their longer-term road-maps this seems likely to continue.

If HP's presence in the higher-end Unix market does weaken significantly, this will pretty much leave the entire high-end market to just IBM and Sun. But on the lower-end Sun's main rivals are effectively Microsoft and Intel, who together with IBM are the three biggest companies in the whole computer industry.

Conclusion

In conclusion, it's very clear that there's quite a bit going on in the volume multi-processor market, and this activity doesn't look to die down anytime soon. In one case, we have two volume desktop microprocessor manufacturers, Intel and AMD, both looking to extend their leadership into the 64-bit multi-processor server range with different approach. Intel's approach with Itanium eschews legacy architecture in favor of a product targeted specifically towards the high-end. AMD, meanwhile, essentially takes a "from the bottom-up" approach, embracing the x86 application base and extending it to 64-bits. At the same time, Sun is attempting to bolster its low-end/workgroup and workstation lineup with upcoming processors like the UltraSPARC IIIi and the $1000 Sun Blade 100 we saw introduced last year, as well as the soon-to-arrive update based on UMC-built 650 MHz UltraSPARC Iii chips with 512 KB integrated L2. But they aren't alone. IBM has made an aggressive push into the UNIX market with its POWER4 series of dual-core chips. And despite the chip's large size and use of massive off-chip caches, IBM has recently moved to extend POWER4 into the entry-level market with the 1-4 way pSeries 630 ("Regatta-LE"). The entry-level systems, which ship with two modules at most (4 CPUs in total), use 1 GHz CPUs and a 64 MB shared L3 cache. This is less than the 1.3 GHz p690 with its 128 MB L3, but the base p630 starts at $12,495, a fraction of its 32-way big brother. And there's much more on the horizon.

In closing, I'd like to take a moment to thank all the people who have helped with the creation of these articles, either through feedback or by providing important information. I'd also like to thank all the readers who have supported this series and read each article from "cover-to-cover," so to speak. Your interest and support both mean a great deal. Thanks for reading.