



## **Sizing Up the Super Heavyweights**

By Paul DeMone, September 17, 2004

### **Leapfrog**

For more than 40 years, competition in the computer industry has been characterized as a game of leapfrog. Major new systems take years to conceive, design, verify and put into production, while at the same time the basic electronic technology underlying computers has improved in performance and dropped in cost exponentially. Because the product cycles of different vendors are often out of phase, it is common for the performance lead to exchange hands whenever new systems are introduced. This was the case in mid-2003 when Intel introduced the Madison 6M version of the Itanium 2 processor, which quickly took the lead in most commercial and technical computing workloads. The wheel turned quite convincingly a year later when IBM introduced its POWER5 processor. What was originally represented as basically a POWER4+ augmented with simultaneous multithreading (SMT) capabilities turned out to be effectively a top to bottom redesign with substantial improvements in both the processor and system level architecture.

Although moderate increases in commercial and throughput oriented performance were expected from the addition of SMT to POWER5, greatly improved latency and bandwidth in its cache and memory system not only amplified the increase in throughput but also raised single thread performance far above that of the POWER4+ and allowed it to take the lead in technical computing benchmarks like SPECfp2000 which the Madison I2 had previously held unchallenged. The POWER5's margin of leadership can be expected to increase in the near future, as larger CPU count systems are gradually introduced.

But the game of leapfrog never ends. Intel's response will come in two parts; the first is the imminent release of the Madison 9M. However, this device was originally a modest incremental improvement over the existing Madison and a recently disclosed reduction in its maximum clock rate target will reduce its impact even more. The real battle will come about a year from now when Intel introduces the Montecito, a 90nm dual core IPF processor with multithreading and massive, fast on-chip caches. This article examines the POWER5 and Montecito, and describes the various ways each device is a substantial improvement over its predecessor and speculates how Montecito, the POWER5, and its anticipated 90nm shrink follow-on, POWER5+, will compare in the arena of super heavyweight class of server MPUs and systems.

## POWER5 – Major Improvements From Top to Bottom

Prior to its release, the aspect of the POWER5 that IBM directed most attention to was the addition of SMT to a POWER4 style microarchitecture. This technique exploits the natural inefficiencies and stalls that occur when programs are executed on a fast and wide processor to allow the CPU to opportunistically interleave the execution of two threads. Although this does not speed up an individual program, it allows a single CPU to get more work done from a number of threads in the same amount of time. In addition to SMT, the POWER5 greatly improved system level functional partitioning compared to POWER4 as shown below in Figure 1.

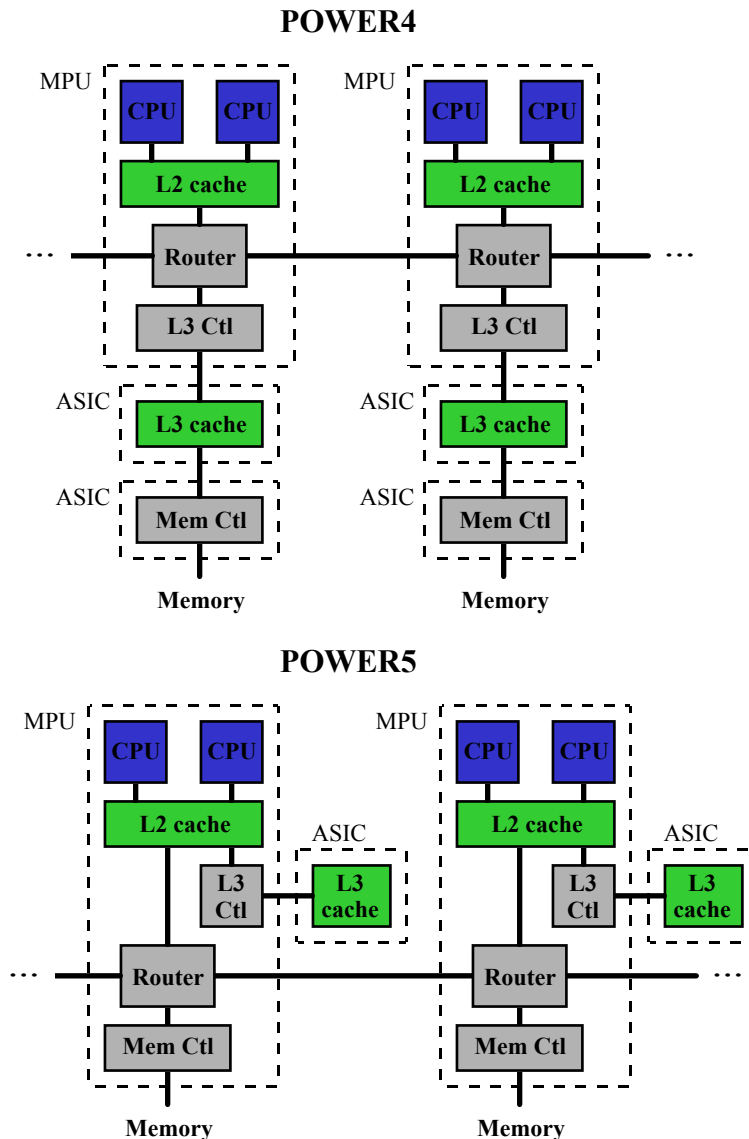
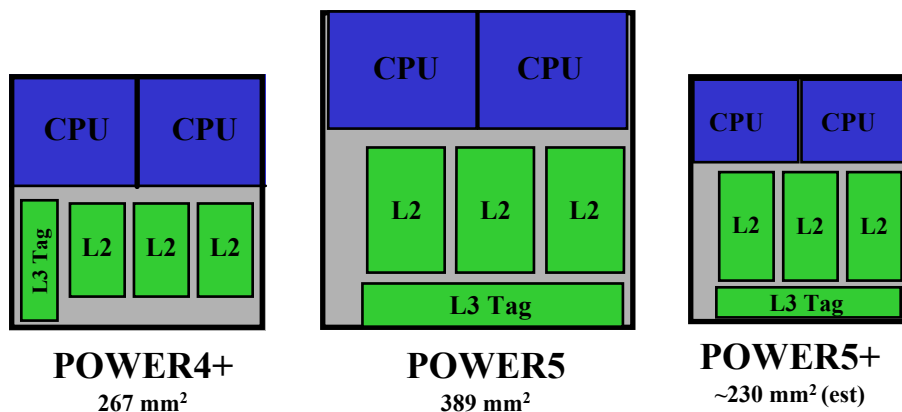


Figure 1 – System Partitioning: POWER4 and POWER5

To support SMT, the POWER5's architects increased the number of GPR physical rename registers from 80 to 120 and the number of FPR rename registers from 72 to 120 [1][2]. The additional FP rename capability had the benefit of increasing POWER5's single thread performance on some HPC workloads compared to the original POWER4 microarchitecture; this gain stemmed from an enhanced ability to execute critical code sections out of order.

Despite the significant internal and external architectural changes that separate the POWER4 and POWER5, their device floorplans are remarkably similar. The original 180nm POWER4 device was 412mm<sup>2</sup>. It was later shrunk to 130nm and became known as the POWER4+. Because the POWER4+'s L2 capacity was barely increased (~500kb, from 1.44MB) its die size fell to 267mm<sup>2</sup>. The POWER5 is manufactured in the same 130nm SOI CMOS process as the POWER4+, but enhancements to the CPU (including SMT), modestly larger L2 capacity, and increased integration of system level functions including the memory controller, swell the POWER5's die size to 389mm<sup>2</sup>. Next year IBM will likely release the POWER5+, a shrink of the existing POWER5 device to the 90nm process used to currently manufacture the PowerPC 970FX. The relative die size of the POWER4+, POWER5, and an estimate for the POWER5+ presuming no major changes to L2 capacity are shown in Figure 2.



**Figure 2 – Relative size of POWER4+, POWER5, and estimated POWER5+**

Although eight processor/quad device MCMs similar to those used for POWER4 and POWER4+ are planned for POWER5, current systems use a dual chip module (DCM) that integrates a single POWER5 device along with its associated 36 MB L3 EDRAM ASIC. Although this packaging configuration is quite reminiscent of the Pentium Pro, the POWER5 DCM lacks a conventional system bus and instead has separate wide interfaces for connection to memory, other DCMs, and I/O devices. Despite a 46% increase in die area and an increase in complexity from 184m transistors to 276m transistors, IBM engineers were able to increase POWER5's clock frequency over the POWER4+. Until quite recently, the POWER4+ topped out at 1.7 GHz. The POWER5 was introduced at 1.65 GHz and 1.9 GHz speed grades. IBM may even be able to coax even slightly higher clock frequencies from the 130nm POWER5 before the emphasis shifts to 90nm.

## It's the Memory, Stupid

But higher clock frequency and SMT, are not the primary performance differentiator between the POWER4 and POWER5 architectures. Nearly ten years ago when DEC MPU designers were considering their third generation RISC design, chief Alpha architect Dick Sites reportedly summarized his analysis of the biggest short coming of the second generation EV5 chip with the concise phrase “It’s the memory, stupid”. It is clear now that IBM’s server MPU architects consider Site’s advice as relevant today as it was a decade ago. The most distinguishing characteristics of the POWER5 are the extent that IBM improved upon the POWER4+’s cache and memory system hierarchy [3][4]. These changes are listed below in Table 1.

	<b>POWER4</b>	<b>POWER5</b>
<b>L1 Dcache</b>	2-way associative	4 way associative
<b>L1 Icache</b>	direct mapped	2 way associative
<b>L2 cache</b>	1.44 MB	1.92 MB
<b>L3 cache</b>	8-way associative 32 MB	10-way associative 36 MB
<b>Memory</b>	8-way associative 123 cycle latency 1/3 CPU frequency ~4 GB/s per device 351 cycle latency	12-way associative 87 cycle latency ½ CPU frequency ~16 GB/s per device 220 cycle latency

**Table 1 – Improvements to POWER5 Cache and Memory Hierarchy**

The three levels of cache in the POWER5 all have higher degrees of associativity than their counterparts in the POWER4 to reduce miss rates. In addition the on-chip L2 and external L3 are slightly larger. The biggest improvement to cache in the POWER5 is in the latency and bandwidth of the L3. Although it is still external to the processor, it now operates at 1/2 the processor clock rate in the POWER5 compared to 1/3 the processor clock rate in the POWER4. This change increases bandwidth by 50% and reduces latency by roughly a third.

The memory system in the POWER5 has been greatly improved by integrating the memory controller into the processor. In the POWER4 architecture, the memory was controlled by a separate ASIC that was accessed by the processor through the external L3 device. This change provides separate and independent paths between the MPU and L3 cache and the MPU and memory. It has the double benefit of increasing potential operational parallelism and bandwidth as well significantly reducing latency. Instead of requiring six chip-to-chip transfers and four extraneous ASIC crossings to perform a memory read, the POWER5 device directly controls up to four independent channels of DDR or DDR2, albeit through intermediary buffer chips. In a sense IBM is following the same path as Alpha EV7 and AMD Opteron architects in extracting performance gains outside of the CPU core at the system level by bringing memory closer to the processor.

The results of these optimizations are astounding – memory bandwidth per device is quadrupled even as memory latency in terms of processor clocks falls by nearly 40%. The cache and memory performance of the Madison 6M, POWER4+, and POWER5 are provided in Table 2, along with speed and throughput performance as measured by Linpack, SPEC, SPEC\_rate2k, and TpmC.

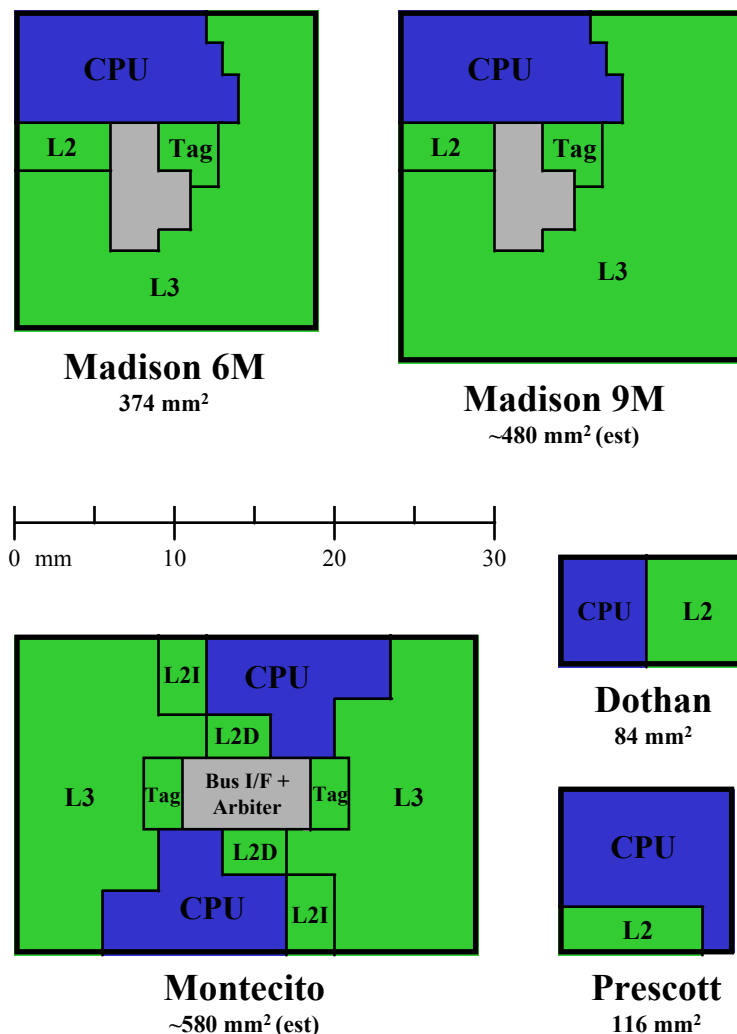
	<b>Madison 6M</b>	<b>POWER4+</b>	<b>POWER5</b>
<b>Frequency (GHz)</b>	1.5	1.7	1.9
<b>L2 Latency</b>	5 cycles	12 cycles	13 cycles
	3.3 ns	7.1 ns	6.8 ns
<b>L3 Latency</b>	14 cycles	123 cycles	87 cycles
	9.3 ns	72.3 ns	45.8 ns
<b>Memory Latency</b>	~224 cycles	351 cycles	220 cycles
	~149 ns	206 ns	116 ns
<b>STREAM Copy BW, 4P</b>	5.07 GB/s	8.37 GB/s	17.9 GB/s
<b>Linpack GFLOP/s, N=1000</b>			
<b>4P</b>	18.2	14.7	21 est
<b>1P</b>	5.43	3.88	5.6 est
<b>SPECint_base2k</b>	1408	1077	1398
<b>SPECfp_base2k</b>	2161	1598	2576
<b>SPECint_rate_base2k, 4P</b>	63.4	48.4	74.4
<b>SPECfp_rate_base2k, 4P</b>	82.2	66.5	125
<b>TPC-C (TpmC), 4P</b>	136k	-	194k

**Table 2 – Comparison of Madison 6M, POWER4+, POWER5**

The effect of lower latency and higher bandwidth is quite evident. Compared to the 1.7 GHz POWER4+ the 1.9 GHz POWER5 achieves a 30% higher SPECint\_base2k score and a 61% higher SPECfp\_base2k score with only a 12% higher clock frequency. The difference in throughput is even more dramatic because SPECrate2k testing on the POWER5 was run with 2N copies of the SPEC CPU suite on N processors to take advantage of its SMT capability. The unexpected magnitude of absolute improvement in cache and memory latency and bandwidth is the primary reason why the POWER5 blew past most of my performance estimates in [The Battle in 64-bit Land: Merchant Chips on the Rise](#).

## Montecito – Make it a Double and Supersize It

At the beginning the 1990s Intel broke the million-transistor threshold with two 32-bit MPUs, the 486 and the 860. When Montecito ships next year, it will have taken Intel only 15 years to have crossed three orders of magnitude and cruise by the billion-transistor mark. The upcoming dual core 90nm IPF MPU packs 1.72 billion transistors on a single monster die largely because of its 26.5MB of integrated L2 and L3 cache. Based on analysis of public relations photos of 300mm Montecito wafers and recently disclosed die microphotographs it appears that Montecito is roughly 20 x 29mm or 580mm<sup>2</sup> in size. The relative die size and floorplan of the Montecito is shown in Figure 3 along with those of its 130nm single core IPF predecessors Madison 6M and Madison 9M as well as fellow 90nm Intel chips Prescott and Dothan.



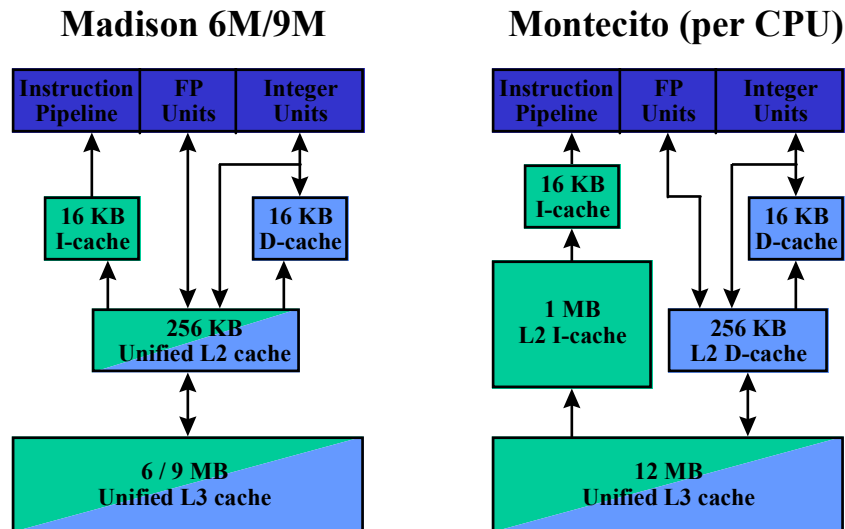
**Figure 3 – Floorplan and Relative Die Size of Montecito and other Intel MPUs**

Despite its size, Montecito will likely cost about the same or less than the estimated ~\$125 of the existing Madison 6M to manufacture. Silicon cost is based on wafer count

and varies weakly with wafer size. Montecito is manufactured on 300mm wafers with nearly 100 candidate dice per wafer, 50% more than the ~63 Madison 6M on a 200mm wafer. Yield is not a specific issue for large memory intensive chips like Montecito, because it is over 70% L2 and L3 cache by area, and regular memory structures are protected against the majority of random point defects using redundant circuit and array elements. The Montecito CPU cores are only about 60mm<sup>2</sup> each. The two CPU cores along with about 60mm<sup>2</sup> of shared logic total about 180mm<sup>2</sup> of non-cache region vulnerable to any defects, about the same critical area as a Willamette based P4 or Celeron.

The Montecito is more than simply dual Itanium 2 CPUs with more cache. Each CPU also incorporates coarse grained multithreading, (CMT) where hardware provides architected processor state for two threads along with logic to automatically switch execution from one thread to the other when a thread relinquishes the CPU under software control or experiences a high latency event, like an L3 miss. The thread switch time is reportedly 15 cycles, which suggests a full pipeline flush is performed when switching threads. Although this sounds like a significant latency, one must keep in mind that for a 2+ GHz processor like Montecito an L3 miss could otherwise stall a CPU for 20 times longer or more. In addition to CMT, each Montecito CPU implements new IA-64 instructions, has extra functional units for shifting and population count, more efficient speculation recovery, and features for processor virtualization and enhanced reliability, availability, and serviceability (RAS) [5].

Although the cache hierarchy of previous IPF MPUs are arguable the most advanced of any processor family in terms of latency, bandwidth, and capacity, this was nevertheless an area of major improvement in Montecito. The changes are shown in Figure 4.



**Figure 4 – Improvements in Montecito Cache Hierarchy**

The biggest change was to split the unified 256KB L2 cache of the McKinley and Madison into separate 1MB L2 instruction cache and 256KB L2 data cache. This was done basically to eliminate instruction stream competition for the bandwidth and capacity of the L2 data cache. The 16KB instruction caches of the Madison and Montecito hold only 1024 instruction bundles which represents about 2.4k useful instructions taking into account the ~20% structural NOP content of a typical IA64 executable. To put that into perspective, that is only about 1/7<sup>th</sup> the instruction capacity of the POWER5's 64KB instruction cache. Obviously for many classes of programs, instruction stream fetching will represent a significant portion of the processor requests on the unified 256KB L2 as well a large portion of its contents.

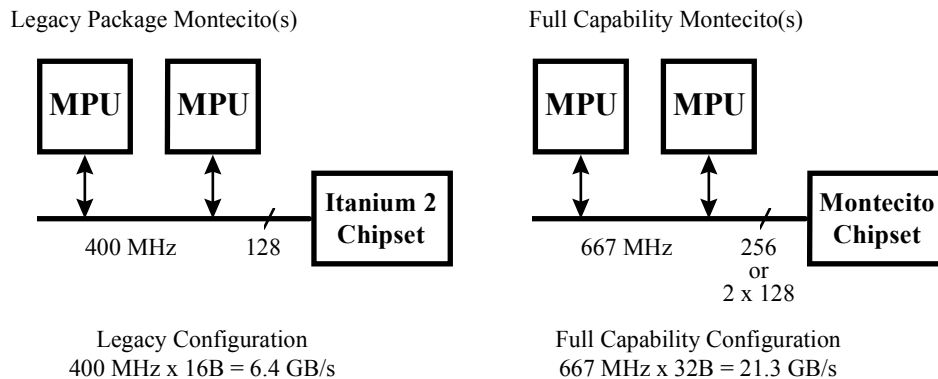
By splitting the L2 caches in Montecito a lot of good things happen. From the data stream perspective, the 256KB L2 suddenly has one less port, and its entire 256KB capacity is available for data. This means less contention and stalls and fewer capacity and conflict misses. This adds up to more predictable memory hierarchy behavior, a very important feature for an architecture that relies heavily on static instruction scheduling. From the instruction stream perspective, the L2 I-cache can be located physically close to the L1 I-cache and its design optimized for the task. It doesn't need to be multi-ported or support sub-word access. As a result the 1MB L2 I-cache in Montecito likely has little or no latency penalty over the 256KB L2 D-cache, despite having four times its capacity. The combination of a very fast latency (1 cycle) L1 I-cache and large and fast L2 I-cache has operational characteristics are impossible to duplicate in a single level cache. For example, if 90% of instruction stream accesses that hit in the L1 or L2 hit in the L1, and the L2 has a latency of 6 cycles, than the L1/L2 combination performs like as a single level 1MB instruction cache with average latency of  $0.9*1 + 0.1*6 = 1.5$  cycles. This is half the latency of the 64KB instruction cache in the Alpha EV6/7 and AMD K7/8.

Information about the L3 caches in Montecito is encouraging but ambiguous. Although L3 capacity is doubled per CPU (12 MB) and quadrupled per device (24 MB) compared to Madison 6M, L3 latency was said be the same as in Madison 6M and 9M [5]. The ambiguous part is that the Madison 6M's latency can be described either in absolute terms (9.3 ns) or in processor clock cycles (14). Given the size of the Montecito and the fact that it uses only seven layers of interconnect, just maintaining 9.3ns latency while doubling cache size is a good accomplishment. Keeping L3 latency at 14 cycles while clocking 50+% faster than Madison 6M (i.e. ~6 ns) would be an astounding accomplishment. Rounding out the improvements in Montecito's cache hierarchy are more efficient L2 and L3 queuing logic as well an increase in the number of L2 and L3 cache line victim buffers.

## Bottleneck

The 400MHz, 6.4GB/s Itanium 2 system interface was fully adequate for the original McKinley and only a moderate constraint for Madison generation MPUs. But it is clear that maintaining backward compatibility would impose a major bottleneck on the 2+ GHz dual core multithreaded Montecito that would prevent the device from reaching its full performance potential on many types of applications despite the ample on-chip caches. But even a Montecito hobbled by backward compatibility still represents a substantial upgrade path for existing Itanium 2 systems and chipsets.

To solve this dilemma Intel will offer at least two versions of the Montecito. The legacy version will implement a 400MHz system interface and be packaged in the existing PAC611 module. It will serve as a drop-in replacement for the Madison 6M and the upcoming Madison 9M. The full capability, high performance version of Montecito will reportedly implement a 667MHz system interface. That transfer rate represents a 67% increase in system bandwidth, yet Intel has indicated that Montecito will offer over three times higher bandwidth than Madison [6]. That suggests that in addition to operating at the higher transfer rate the data path width will either be doubled to 256 bits or that two separate 128 bit wide data paths will be provided. The legacy package and full capability Montecito system interfaces are shown in Figure 5.

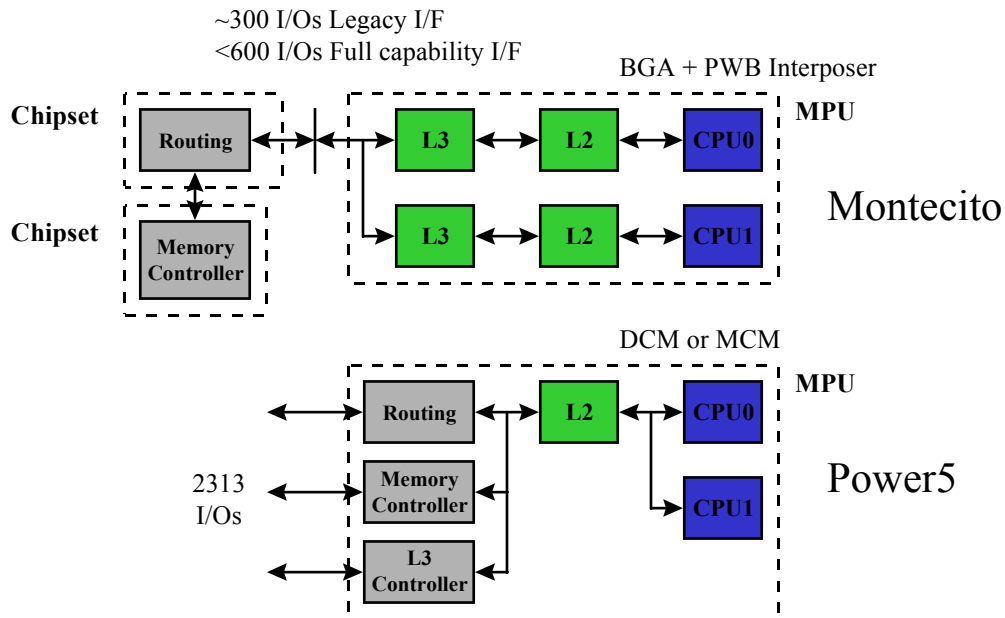


**Figure 5 – Montecito Configurations**

Although the 21GB/s system interface of the full capability Montecito will go a long way to close the massive per device bandwidth gap that POWER5 has opened over Madison, the new IPF chip clearly leaves system building – memory control, routing and scalability architecture entirely in the hands of others. Incredibly, the cancellation of its “Bayshore” chipset means that Intel can’t offer OEMs a standardized means to use the full capability Montecito. The two year old 870 chipset, which is used by Dell and value added resellers (VARs) of Intel white box systems, only supports the 6.4 GB/s legacy interface. Whether a deliberate tactic or a strategic mistake, Intel’s decision to cancel Bayshore means that the full performance capability of Montecito will be available only in the systems of OEMs which aggressively develop their own proprietary IPF chipsets.

## How You Slice It

There is a clear difference in design strategy between Intel's EPIC and IBM's RISC high-end server processors. Intel chips are system agnostic – they embed capabilities but no choices and few assumptions about system level architecture in silicon. The interface between the MPU and system is a clean line of demarcation designed with as few signal crossings as possible to avoid the need for exotic and expensive packaging technology. Large multi-level integrated caches are provided to minimize data traffic across the interface and more efficiently allow 2 or 4 MPUs to share a bus segment (a capability sometimes mistaken as a requirement). In contrast, IBM chips can be considered as the silicon realization of a large-scale system architecture partitioned down to the granularity of two CPU cores. Because IBM sells both the MPU and the system it goes into, it has more freedom to partition in ways that greatly increases MPU signal count and the sophistication of packaging technology required at both the silicon and system level. This difference is shown in Figure 6.



**Figure 6 – Intel and IBM System Partitioning**

These two choices are basically driven by the clear difference in the business models of Intel and IBM. IBM is the greatest, and the last, of the vertically integrated computer OEMs. It designs and manufactures high-end server processors with the sole intention of selling them to customers safely wrapped within an IBM designed and manufactured system. As such, it doesn't need to worry about creating silicon usable by other OEMs with inferior technical capabilities. In contrast, Intel is a component vendor whose end customers are most of the world's computer OEMs, big and small, including IBM. Intel must design its MPUs to appeal to this large and diverse assortment of system designers and integrators, which differ tremendously in their wealth, engineering resources, and the markets they intend to address. Intel's challenge is to create the highest possible performance MPU family that is still usable by a full cross section of OEMs with widely

divergent system engineering capabilities; this necessitates conventional, relatively low pin count packaging. To maximize IPF performance, Intel exploits its inherent semiconductor design and manufacturing strength through the use of large die size devices combining compact full custom processor cores with massive fast caches.

## Tale of the Tape – Sizing up the Contestants

The primary characteristics and features of the legacy and full capability Montecito device are summarized and compared with those of the POWER5 and future POWER5+ in Table 1. Data not directly or indirectly known from public disclosure are my own estimates and are indicated as such with “est”.

	<b>Montecito (Legacy)</b>	<b>Montecito (Full)</b>	<b>POWER5</b>	<b>POWER5+</b>
<b>Process</b>	90 nmBulk, 7LM	90nm Bulk, 7LM	130nm SOI, 8LM	90nm SOI, 10 LM
<b>Die size (mm2)</b>	~580 est	~580 est	389	~230 est
<b>Transistors</b>	1720m	1720m	276m	~300m est
<b>Multiproc.</b>	2 way	2 way	2 way	2 way
<b>Multithread.</b>	2-way CMT	2-way CMT	2-way SMT	2-way SMT
<b>L1 cache (per CPU)</b>	16 KB I 16 KB D	16 KB I 16 KB D	64 KB I 32 KB D	64 KB I 32 KB D
<b>L2 cache</b>	2 x 1.0 MB I 2 x 256 KB D	2 x 1.0 MB I 2 x 256 KB D	1.92 MB	2.25 MB est
<b>L3 cache</b>	2 x 12 MB	2 x 12 MB	external	external
<b>Signal I/O</b>	296	~600	2313	2313
<b>Packaging</b>	PAC611	?	DCM, MCM	DCM, MCM
<b>Max Freq. (GHz)</b>	2.0 - 2.3 est SW dependent	2.0 - 2.5 est SW dependent	2.0 est	2.7 est
<b>TDP (W)</b>	100	120 est	200 est	170 est
<b>Max Mem BW</b>	6.4 GB/s	21.3 GB/s	~16 GB/s	~20 GB/s est

**Table 3 – Device Characteristics**

The thermal design power of the Montecito has been stated as 100W. It is likely that the full capability version will push closer to the long established 130W IPF family power envelope because of higher I/O bandwidth and associated higher instruction throughput. The maximum clock frequency of Montecito is difficult to specify because it is a moving target. Intel’s “Foxton” dynamic power management system automatically adjusts processor clock rate and supply voltage as a function of the “activity factor” of the application being executed to keep the device running near its target power consumption limit. If an activity factor of 1.0 represents the maximum device power that can be extracted for any arbitrary code sequence (i.e. a power virus) then the maximum power of any real application (typically Linpack) is around 0.8, SPECfp2k is about 0.7, SPECint2k about 0.65, and TpmC about 0.6 [7]. Because Foxton adjusts both processor supply

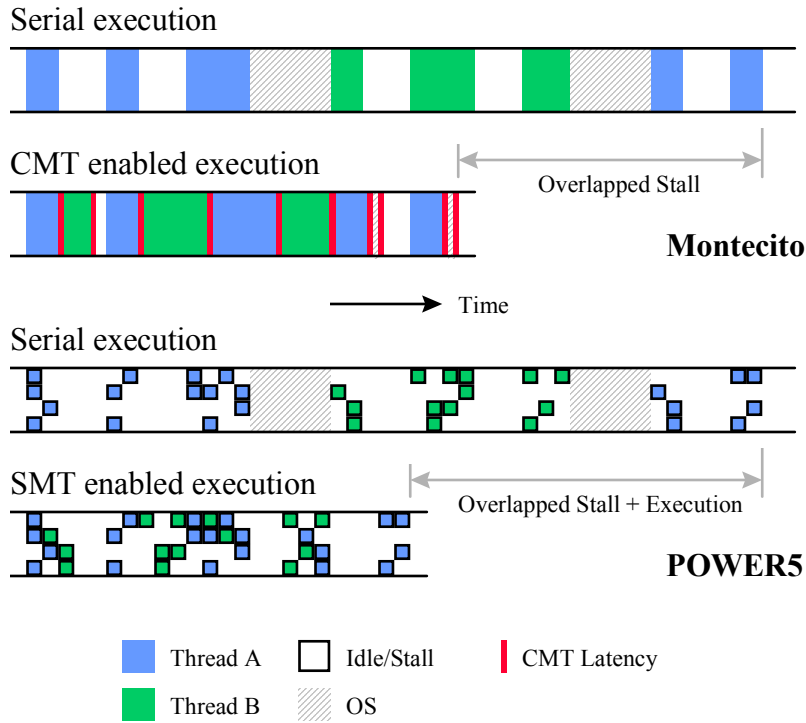
voltage, as well as frequency, Montecito only needs to run Linpack at about 10% lower clock rate than TpmC to keep power constant (i.e. near but not over the programmed TDP limit). Although Intel has disclosed Montecito operating frequency only as over 2 GHz performance data in [8] suggests the highest speed grade (“top bin”) Montecito is expected to run Linpack at 2.5 GHz within a future model of the HP rx2600.

POWER5 is currently shipping at a maximum of 1.9 GHz although it is likely 2.0 GHz or even slightly faster speed grades may be released over its life time. The maximum power estimate of 200W for POWER5 is based on reports of POWER4+ power consumption and system level comparison of the 4U dual socket P5-570 system with the 4U dual socket HP rx2600. Although the POWER5 incorporates an active power monitoring and management system [9], it is much less sophisticated than Foxtan and closer in nature to the thermal protection/throttling scheme Intel has long used in its desktop PC processors. The POWER5+ estimated maximum clock rate of 2.7GHz is based on public statements from IBM concerning process speed scaling at 90nm and beyond, as well as current and expected progress in ramping PPC970FX clock speed beyond that of the PPC970.

## **Two Approaches to Multithreading**

As previously mentioned, both the Montecito and POWER5 microprocessors exploit thread level parallelism (TLP) by integrating two processor cores on each device, i.e. two way chip level multiprocessing (CMP) as well as incorporating two way multithreading within each processor. Thus each device has separate instances of architected processor state to support the execution of four different threads. However Montecito and POWER5 use very different multithreading techniques within their CPUs.

The Montecito uses coarse grained multithreading (CMT) which is sometimes called switched on event multi-threading (CMT/SOEMT). Unlike simultaneous multithreading (SMT), which is used by the POWER5 and some Intel Netburst MPUs, CMT doesn't improve instruction throughput by overlapping instruction execution from two threads to exploit momentarily idle issue slots and functional units. Instead, CMT has the more modest goal of improving instruction throughput by automatically switching back and forth between threads whenever a thread encounters a long latency event such as an L3 cache miss. This difference in operating principle is shown in Figure 7.



**Figure 7 – Throughput Increase from Multithreading**

Although CMT is less effective than SMT, it is a much better match for the in-order, statically scheduled Montecito processors. For an infinitesimally small increase in processor complexity and area, CMT can be expected to provide Montecito a performance increase on the order of 20% or more on memory/data sharing intensive, cache unfriendly workloads such as on-line transaction processing (OLTP). But in most cases, the benefit will be much less. For example, data in [10] shows that a McKinley running SPECint2k and SPECfp2k is stalled on memory only about 5% and 11% of the time respectively. Given compiler improvements since 2002 and quadrupled L3 capacity in Montecito that suggests that CMT will only increase SPECrate2k throughput by roughly 2% to 5%.

SMT can be implemented as a direct extension of the dynamic scheduling hardware in an aggressively out of order (OOO) execution MPU design with a complexity/area penalty of under 10% in a RISC design [11]. However IBM chose to significantly increase processor resources in POWER5 to maximize the throughput boost from SMT. This increased the cost of introducing SMT to 24% growth in CPU area. Although this is an order of magnitude greater than what Intel incurred in the addition of CMT to Montecito, IBM claims substantial performance improvements from SMT: ~35% in database transaction processing and Websphere workloads, ~28% in SAP, and ~45% in Domino R6 Mail [12]. IBM also claims SMT increases throughput for SPECint\_rate2k and SPECfp\_rate2k by about 21% and 10% respectively [13]. This is significantly better than what CMT can do for Montecito because SMT allows for overlap of thread instruction execution time not just long thread stalls.

## Putting It All Together

The estimated performance of dual socket Montecito, POWER5, and POWER5+ based systems are given below in Table 4. It is impossible to predict how large scale systems based on these MPUs will compare because Intel leaves system architecture choice and design trade-off decisions completely in the hands of OEMs. However if IPF OEMs continue to favor lower cost and complexity over maximum performance and scalability, then IBM will maintain a technical edge in its high end POWERx systems employing large scale multi-chip modules (MCMs) and very high signal count inter-MCM interconnect.

	<b>Montecito (Legacy)</b>	<b>Montecito (Full)</b>	<b>POWER5</b>	<b>POWER5+</b>
<b>Frequency (GHz)</b>	2.3 TpmC 2.2 SPEC 2.1 Linpack	2.5 TpmC 2.4 SPEC 2.3 Linpack	2.0	2.7
<b>Linpack, 4P (GFLOP/s, N=1000)</b>	28	32	22	29
<b>SPECint_base2k</b>	2100	2400	1500	2000
<b>SPECfp_base2k</b>	3000	3700	2700	3500
<b>SPECint_rate_base2k, 4P</b>	96	115	78	104
<b>SPECfp_rate_base2k, 4P</b>	120	170	131	172
<b>TpmC, 4P</b>	190k	230k	203k	255k

**Table 4 – Performance Estimates**

## Summary and Conclusion

Based on extrapolations and estimates derived from publicly available information, Montecito will leapfrog the performance of POWER5 in small and medium scale systems and be very competitive with POWER5+ based upgrades of those systems. The combination of Foxtan technology and an associated demand based switching (DBS) system will likely give Montecito based systems significant thermal management advantages for small to medium scale SMP systems and clusters. POWER5 and POWER5+ hold an advantage in virtualization support, while RAS capability is pretty evenly matched. Too little is known about POWER6 at this time to make quantitative predictions with regards to Montvale, the planned 65 nm shrink of Montecito, but in this industry it is always safe to assume that the game of leapfrog will continue.

## References

- [1] Kalla, R. et al, “Simultaneous Multi-threading Implementation in POWER5”, Technical Digest, Hot Chips 2003, August 2003.
- [2] “POWER5 Overview”, IBM Corporation, July 13, 2004.
- [3] Grassl, C., “POWER5 and HPS Programming Strategies: System Architecture”, IBM Corporation, July 2004.
- [4] “IBM Deep Computing Strategy”, IBM Corporation, 2004.
- [5] McNairy, C., “A Technical Review of Montecito”, Intel Developer Forum, September 2004.
- [6] Otellini, P., “Growth at the Platform Level”, Intel Corporation, May 2004.
- [7] Wolochow, P., “Architecting New IT Enterprise Deployment Models: A Strategic Roadmap”, Intel Developer Forum, September 2004.
- [8] Juling, W., “HPC System for Universities in Baden-Wuerttemberg”, T-systems HPCN Workshop, Braunschweig, April 2004.
- [9] Clabes, J. et al., “Design and Implementation of the POWER5 Microprocessor”, Proceedings of the ACM/IEEE Design Automation Conference, June 2004.
- [10] McCormick, J. and Knies, A., “A Brief Analysis of the SPEC CPU2000 Benchmarks on the Intel Itanium 2 Processor”, Technical Digest, Hot Chips 2002, August 2002.
- [11] Preston, R. et al., “Design of an 8-wide Superscalar RISC Microprocessor with Simultaneous Multithreading”, Digest of Technical Papers, 2002 IEEE International Solid-State Circuits Conference, p. 334.
- [12] Funk, M., “Simultaneous Multi-threading (SMT) on eServer iSeries POWER5 Processors”, IBM Corporation, May 2004.
- [13] “IBM eServer P5 Virtualization Performance Considerations”, Draft Redbook, IBM Corporation, August 24, 2004.