4th Generation 64-bit PowerPC-Compatible Commercial Processor Design

Abstract

IBM's NorthStar superscalar RISC microprocessor integrates high-bandwith and short pipe depth with low latency and zero cycle branch mispredict penalty into a fully scalable 64-bit PowerPC-compatible symmetric multiprocessor (SMP) implementation. Based on PowerPC architecture, the first in the Star Series of microprocessors, the NorthStar processor contains the fundamental design features used in the newly available RS/6000 and AS/400 server systems targeted at leading edge performance in commercial applications as characterized by such industry standard benchmarks as TPC-C, SAP, Lotus Notes and SpecWeb.

This paper provides an overview of how the processor microarchitecture, silicon technology, packaging technology, and systems architecture have been leveraged to produce outstanding high-performance in commercial applications and server markets. Follow on microprocessors in the Star Series are also described. The design point and performance attributes of the currently available NorthStar microprocessor and planned enhancements are reviewed.

Design Point

The first microprocessor to launch in the Star Series, internally referred to by the code name "NorthStar", is used in AS/400 and RS/6000 systems. In these systems it is referred to as the PowerPC A50 and RS64-II processor, respectively. The NorthStar microprocessor design objectives were to provide more performance, reliability, and functional robustness than previous 64-bit PowerPC commercial/server processor designs while reducing product and development costs.

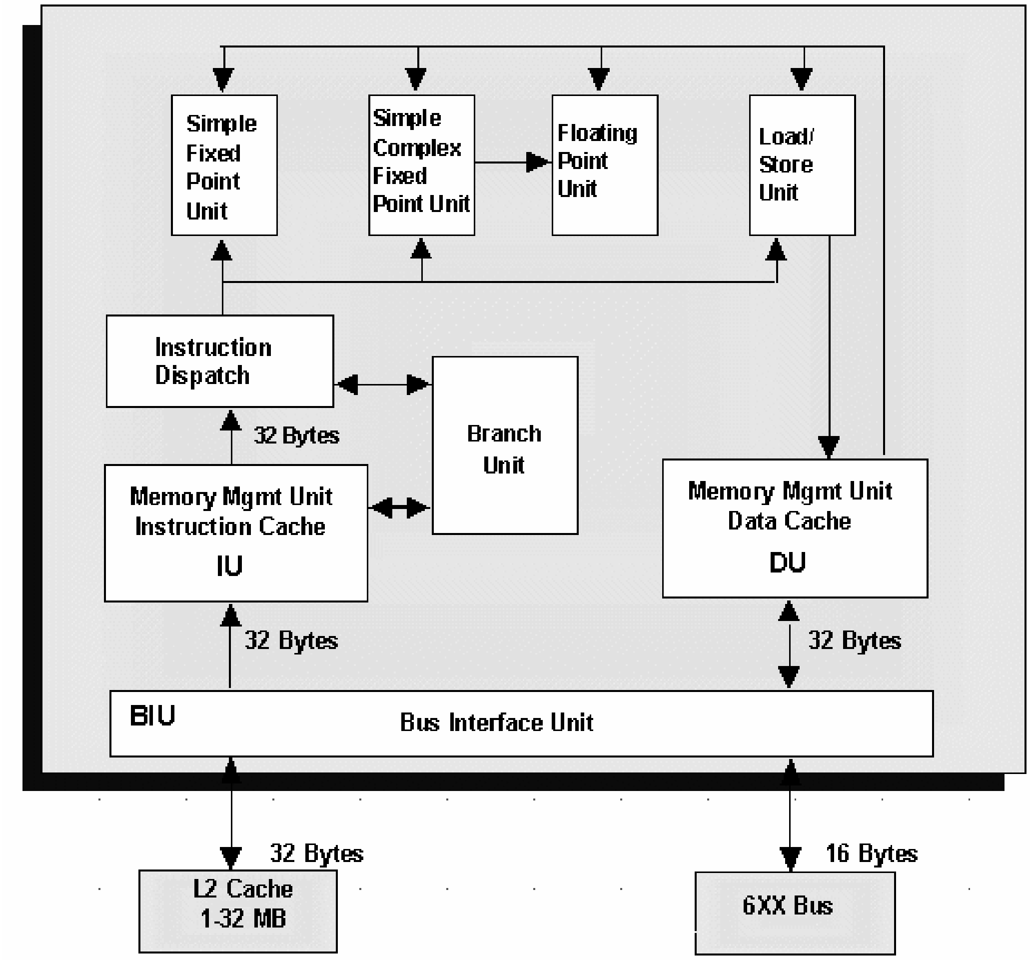
The basic design philosophy was to reuse as much as possible from the previous design point [1], adding enhancements only if they were simple or resulted in significant improvements to the design objectives. The result of this approach is a microprocessor with the following attributes:

- 4 way superscalar

- 5 stage deep pipeline
- Branch mispredict penalty of zero or one cycle
- 64 Kilo Byte (KB) on-chip Level One (L1) instruction cache
- 64 KB on-chip L1 data cache with one cycle load-to-use latency
- Support for up to a 32 Mega Byte (MB) Level Two (L2) cache with a 5 cycle load-to-use latency
- 8.4 Giga Byte (GB) per second L2 cache bandwidth
- 32 byte wide on-chip busses
- 262 MHz operating frequency
- 162 mm2 die size
- 27 Watts maximum power

Major differences from the previous design point include switching from BiCMOS technology into CMOS technology, consolidating the processor from 5 chips into 1, reducing the power by a factor of 5, and adding support for an external L2 cache.

The technology strategy of the NorthStar design was to produce a high performance low cost microprocessor in an advanced, but well-established technology (CMOS6S2). A well defined interface was developed between the processor core logic and the Bus Interface Unit, allowing for future reuse of the processor core with various upgrades and enhancements to the memory subsystem from the level two cache and beyond. The first derivation of NorthStar is called Pulsar. Pulsar leverages IBM's cutting-edge semiconductor copper technology (CMOS7S) to increase the operating frequency to 450 MHz. The higher densities available with this technology also permitted doubling the on-chip L1 instruction and data cache sizes to 128 KByte each. The next step in the roadmap is to map the Pulsar design with its large caches into IBM's newest technology breakthrough called SOI (Silicon On Insulator) to create a microprocessor with a frequency of 550 MHz [2]. Laboratory bring-up is currently underway for both Pulsar and the first SOI based systems. Plans exist for another star series microprocessor that uses a future IBM SOI technology with a target product frequency of 675 MHz.

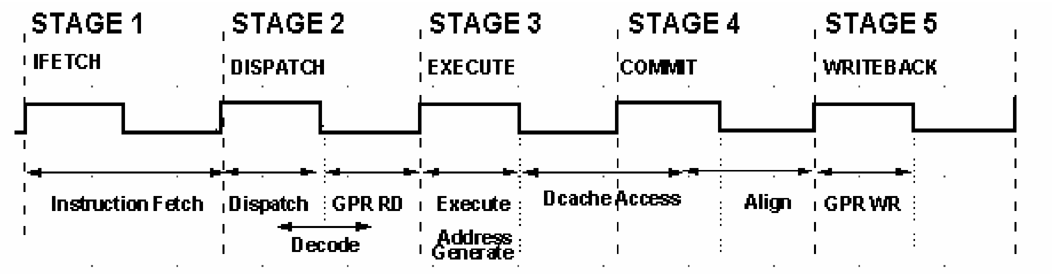


Processor Overview

The NorthStar processor block diagram shown in Figure 1 focuses on server performance with emphasis on conditional branches with zero or one cycle mispredict penalty, contains 64 KB L1 instruction and data caches, has a one cycle load-to-use penalty on the L1 data cache, enhanced string support, four superscalar fixed point pipelines and one floating point pipeline. There is an on board bus interface unit (BIU) that controls both the L2 cache interface and the main memory bus interface.

Description of Pipe Stages

Figure 2 is a pictorial view of the five NorthStar pipe stages described in the following text.



Instruction Fetch stage: In the instruction fetch stage, the L1 instruction cache array is accessed with the address generated by the branch unit and 32 bytes of instructions are output and written into either the 16 entry sequential instruction buffer or the 8 entry branch buffer. The sequential instruction buffer and the branch buffer are used for conditional branch processing, explained in the section, "Branches and Instruction Cache Address Generation."

Dispatch stage: The dispatch stage is responsible for decoding and dispatching up to four 4-byte PowerPC instructions each cycle. Instructions are dispatched in order from either the sequential instruction buffer or the branch buffer. During dispatch, operands are read from architected registers, completion buffers, and result busses. The instruction cache branch target address is generated in the dispatch stage unless there are dependencies on instructions that have not yet executed.

Execute stage: During the execute stage, the arithmetic, rotate and data cache address generation functions are performed. All results are bypassable to all execution units for use in the next cycle by subsequent instructions. Condition register lookahead logic based on the arithmetic and rotate zero detect and sign bits bypasses into the conditional branch logic. If any of the input operands are invalid due to dependencies, then the execute stage for that pipeline stalls.

Commit stage: The commit stage holds execution results for each pipeline. Taken branches, exceptions, and page faults can cause commit (and execute) results to be discarded. Commit stage results are bypassable to the execution stage. Cache fetch data is bypassable to the execution stage through an aligner.

Writeback stage: The writeback stage of the pipeline writes the instruction results into architected registers once all branch and exception conditions have been resolved.

Branches and Instruction Cache Address Generation

When it comes to keeping the pipeline full of instructions, conditional branches pose a special problem. Many processor designs solve this problem with branch prediction logic. Software code executed in the commercial environment has fewer code loops, making branches harder to predict with high accuracy. The Star Series of processors takes a different approach in solving the conditional branch problem by minimizing the branch mispredict penalty to zero or one cycle. This is accomplished with a combination of techniques.

The wide 32 byte instruction fetch path from the instruction cache paired with the 16 entry sequential instruction buffer and the 8 entry branch instruction buffer allow instructions at the branch target to be prefetched while instructions are being executed out of the sequential instruction buffer.

The first cycle of branch processing begins prior to the dispatch stage. The branch logic looks ahead, up to six instructions, into the dispatch queue to find a branch instruction. The first branch instruction found is decoded and its branch target address is generated. In the second cycle of branch processing, the instruction cache array is accessed and the aligned output is written into an instruction buffer.

By default, branches are assumed "not taken", that is, instructions are dispatched and executed down the "not taken" or sequential path prior to the outcome of the branch instruction being known. This is equivalent to predicting branches as not taken. Once the outcome of the branch instruction is known and if the branch is taken, the instructions dispatched after the branch instruction are canceled. The "branch taken" logic switches dispatch from the sequential instruction buffer to the branch instruction buffer and cancels instructions dispatched down the "not taken" path. The "branch taken" logic is some of the most timing critical logic in the processor. The zero detect and sign bits in the execution stage of the fixed point units are bypassed into the "branch taken" generation logic.

Branch penalty is defined as the time from dispatch of a branch instruction to the dispatch of the target of a branch instruction. When the instruction cache branch address is generated ahead of the dispatch stage and the branch condition is known at the end of the dispatch stage, there is no branch penalty. This is known as a "zero cycle" branch. When the Instruction cache address is generated during dispatch stage (instead of earlier) or when an instruction modifying the condition register is dispatched in parallel with the conditional branch, a one cycle branch penalty is incurred.

Fixed and Floating Point Units

Two of the four superscalar units are fixed point units (FXUs) and have single cycle execution for the bulk of the integer arithmetic instructions. One of the two FXUs is specialized to also execute multi-cycle integer instructions, such as multiply and divide.

Although a commercial processor, it was deemed necessary to implement a simple and efficient pipeline for floating point arithmetic. The floating point unit (FPU) is fully independent and contains hardware for square root and division as well as for the fused multiply-add instruction. The FPU is fully pipelined with four cycle latency, single cycle throughput.

The load store unit includes a custom dynamic adder to allow for high speed cache address generation.

L1 Data Cache, L2 Cache and Bus Interface Unit

Minimizing L1 data cache latency is key to high performance without complexity. The Star Series of processors are microarchitected so that the L1 data cache access has a 1 cycle load-to-use penalty. The L1 data cache was designed to be as large as possible without increasing the load-to-use penalty to 2 cycles.

The L1 data cache data bypasses directly into the execution units. It is 2 way set associative with a 16 byte interface to the execution units and a 32 byte interface for cache line replacement.

The L1 data cache was designed with four single port arrays. Chip area was saved by using single port arrays instead of multi-port arrays to increase the number of entries in the cache while minimizing cache access latency. Cache line replacements and stores normally done with a second cache port are accomplished by queuing them in a line fill buffer and a store buffer. The fills and stores are done either during background cycles when the instruction stream is not accessing the data cache or simultaneously with instructions that operate on 8 bytes of data or less. The majority of instructions operate on 8 bytes of data or less and these instructions use at most one half of the available L1 data cache interface. The line fill buffer holds 7 cache lines and has the characteristics of an L0 (Level Zero) cache in that any portion of an incoming line can be stored to or read from. A high speed bypass path around the line buffer exists for the first data transfer coming from L2 cache or main store going directly to the execution units.

The on-chip BIU contains the interface logic to support up to a 32 MB L2, 6XX system bus protocols, and dedicated hardware to hide latency to memory. While the NorthStar chip supports up to 32 MB of L2, a maximum 8 MB of L2 was installed on the first systems using the NorthStar processor due to considerations external to processor support. L2 latency is right behind L1 latency when it comes to impact on performance, so various innovative techniques were used to minimize the L2 load-to-use latency to a total of 5 cycles. L1 data cache accesses are speculatively forwarded to the L2 and canceled if an L1 cache hit is detected. The L2 SRAM clocking logic on NorthStar is designed to tolerate a wide range in access delays caused by SRAM process variation without adding latency to the access path. In the first incantation of the Star Series, both the L2 cache and L2 cache directory are implemented with SRAMs external to the processor chip. The external L2 cache is 4-way set associative. Associativity in the L2 results in higher L2 cache hit rates for most commercial workloads.

Instructions that gate SMP performance, such as those related to locks, TLB (Translation Lookaside Buffer), cache management, and synchronizing, are optimized for performance in the storage control microarchitecture. Lock information is bypassed between pipeline stages to prevent pipeline stalls. The TLB table walk routine is implemented in circuits instead of microinstructions to reduce table walk time. The cache coherency scheme implemented by the Star Series processors does not require synchronizing instructions to be broadcast on the system bus, minimizing the performance impact due to synchronization.

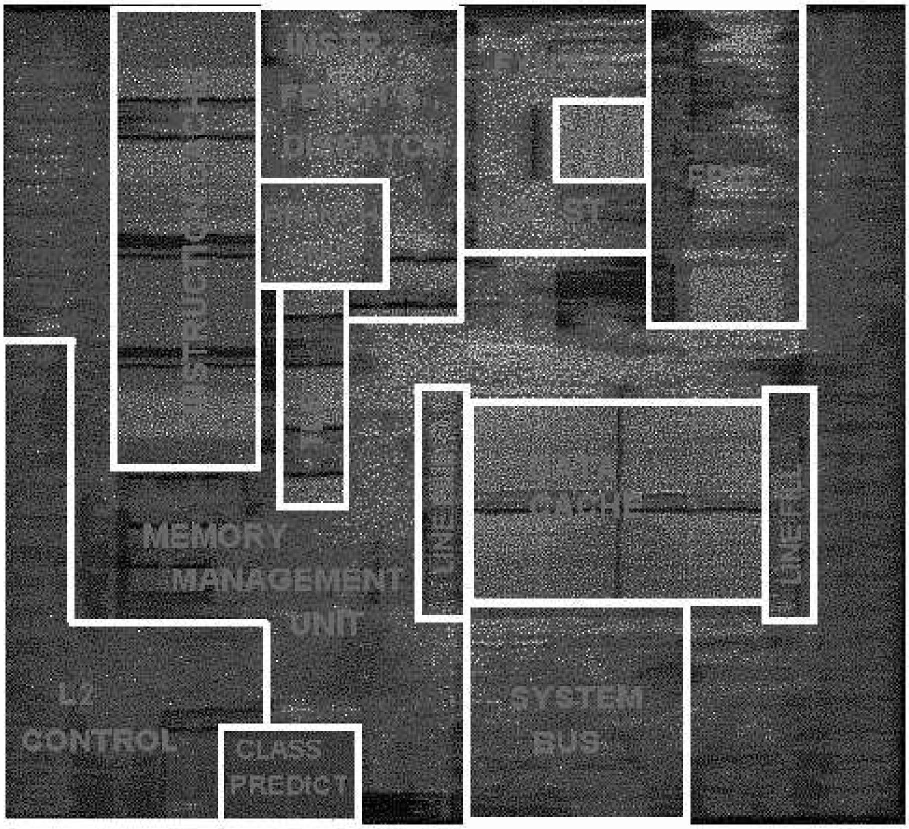


Figure 3. NorthStar Chip

Containing 12.5 million transistors, the NorthStar processor die is shown in Figure 3. It is manufactured in IBM's 0.35 micron hybrid CMOS 6S2 technology, with five levels of interconnect metallurgy.

System Implementation

A key challenge of the NorthStar processor was to design a high bandwidth system interface required to support the high miss rates driven by commercial processing. NorthStar leveraged IBM's advanced packaging technology to implement separate, independent 16 byte memory bus and 32 byte L2 bus, each with separate address, data, and control lines, achieving 8.4 GB (Giga Bytes) per second to the L2 at 262 MHz. This was achieved with a total of 2030 chip I/Os of which 985 are signal I/Os.

The system interface is designed to allow flexibility in system implementation from low cost, bus-based systems to more complex switch-based configurations providing greater address and data bandwidth.

The NorthStar processor design supports Modified Exclusive Shared Invalid (MESI) snoop-oriented SMP cache coherence along with remote processor bus protocols for increased throughput and large system topologies.

One characteristic of transaction processing is a high rate of data sharing between processors. The Star Series of processors provides improved performance in this environment by allowing cache lines to be transferred directly between processors with a technique called intervention. This results in shorter cache miss latencies compared to retrieving all L2 cache miss data from main store.

Scalability

Current NothStar-based AS/400 and RS/6000 systems are limited to a 12 way symmetric multi-processor (SMP). The Star Series processors provide logic support for greater than a 12 way SMP. This capability will be leveraged in future systems.

Error Correction, Detection, and Isolation

The commercial processing environment requires both high data integrity and high availability. On-chip arrays comprise the largest portion of chip area and are also the most susceptible to failures. For this reason, Star Series processors have built in recovery for single bit array failures. If an error is detected in the instruction cache, instruction cache directory, data cache directory, or the TLB, the entry in error is invalidated or marked unusable and its correct contents refetched. Separate L1 data directories for processor use and SMP snooping are implemented to provide adequate bandwidth. These separate L1 data directories are exact copies of each other resulting in built in redundancy that is used to recover from errors in either directory.

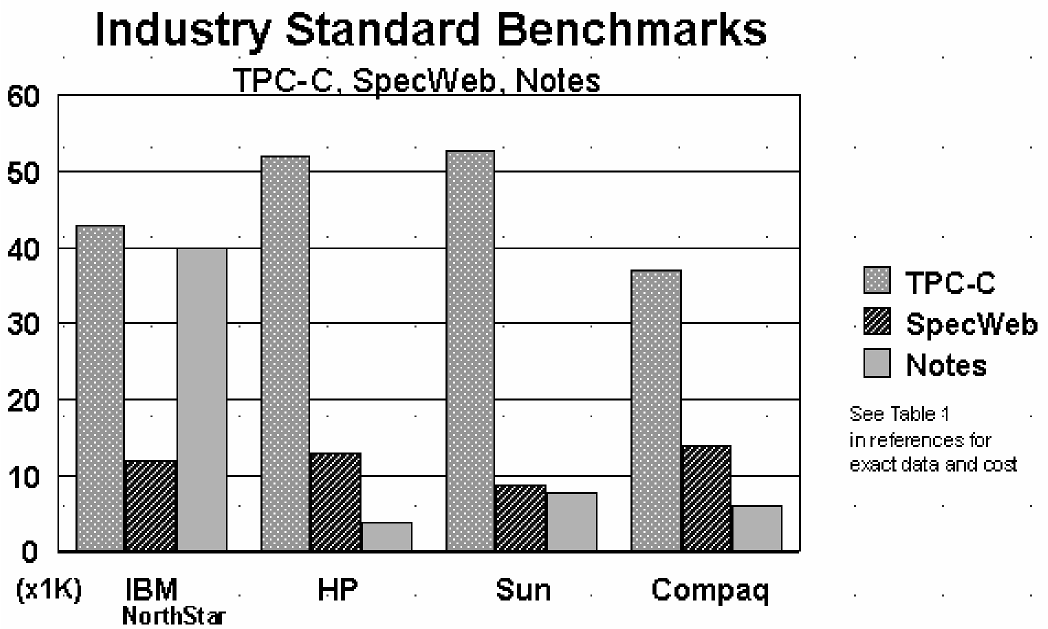
The L1 data cache policy is store-in and may hold the only copy of modified data in the system. For this reason, the L1 data cache is implemented with an ECC scheme that can detect double bit errors and correct single bit errors. The off-chip L2 data cache and the L2 Directory are covered by the same ECC scheme as the L1 data cache.

Various parity schemes are integrated into the control and data flow logic on the processor chip. Whenever a recoverable or non-recoverable error is detected, information pertaining to the error is recorded by the hardware and made available to the system diagnostics to isolate the failing circuits.

Performance

The NorthStar processor excels in real applications precisely because its many facilities combine to cover the wide spectrum of demands that characterize commercial computing. Applications may be limited by the rate of computational speed or by the rate of data delivery to the computational units. They are primarily fixed point intensive yet floating point was not ignored and performs quite acceptable for commercial applications without being a burden to the chip's cost and area. NorthStar's well balanced design handles these challenges with its five execution units, wide data paths and short pipe length.

Four major benchmarks show the remarkable performance of server systems using the NorthStar processor. NorthStar achieves top ten rank for a twelve way AS/400 system on the TPC-C Non-clustered and Non-NUMA subset. The majority of competitors in this category required 24, 32 and 64 way solutions to make top ten rank. The outstanding Lotus Notes Bench performance shows RS/6000 and AS/400 systems implemented with the NorthStar processor owning the top three performance positions unsurpassed. On SPECweb96 its performance stands out by repeatedly filling 3 out of the 10 top positions with RS/6000 systems using NorthStar. And to complete the suite, NorthStar systems take 2 out of the 5 top positions on SAP benchmark for Two-tier Client/Server in AS/400 systems. Figure 4 shows the results of these benchmarks compared to HP, Sun and Compaq. The results depict the balanced performance available with NorthStar-based system. In summary, NorthStar-based systems represent simplicity as its best by giving maximum performance without complexity, resulting in a small (162 mm2) and low power (27 Watt max) chip with a big punch.



Continued Performance Enhancements

Figure 5 shows the future road map of the Star series processor family. The Second design point implements IBM's industry leading CMOS7S process, which provides technology performance gains associated with shrinking channel lengths to 0.18 micron and a reduction in RC delay with the copper interconnect.

In addition to mapping technology, commercial performance will be improved in the Star series of processors by using NorthStar design as a base, doubling the instruction and data L1 caches, and moving the L2 cache directories on-chip. Additional functionality for commercial and server applications is also planned to be added. The technology map and performance tuning are planned to rapidly scale the Star family of processors up to 675 MHz in order to sustain top performance benchmarks in the commercial environment.

Work is already underway to apply IBM's recently announced SOI technology to the Star series roadmap of products. SOI technology is projected to give higher frequencies while at the same time reduce power requirements. The ISSCC98 Conference paper listed in reference 2 provides technical details on the first of the Star series processors to leverage SOI technology. A planned second generation SOI design will further push performance up to 675 MHz.

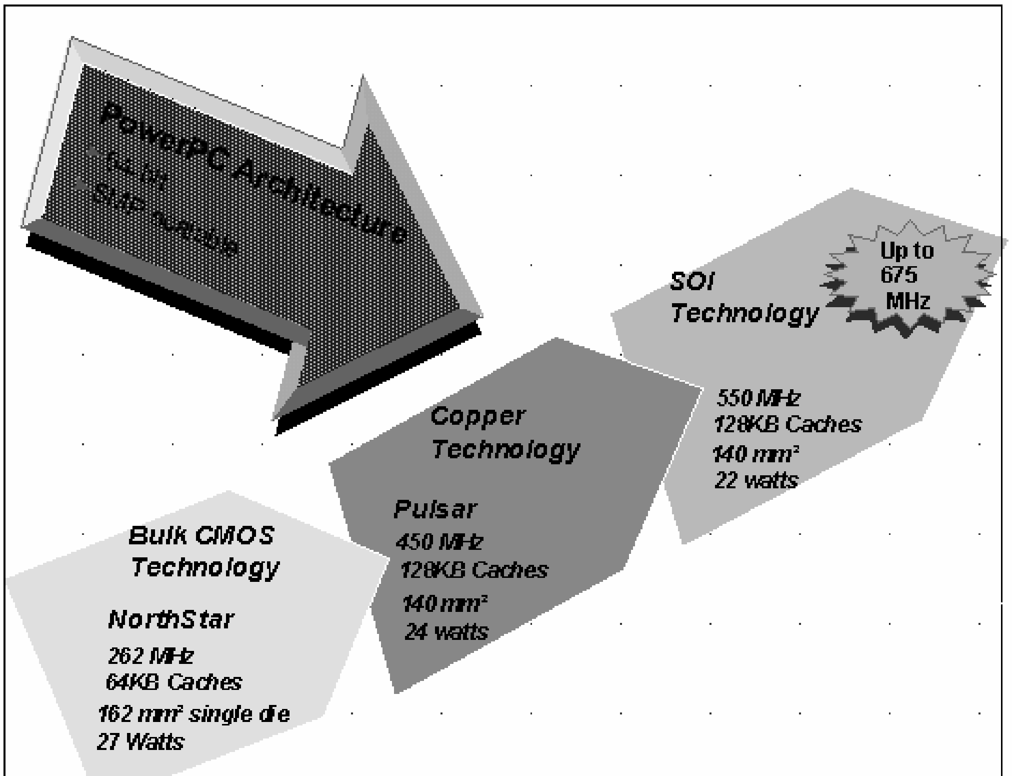


Figure 5. Star Series Roadmap

Summary

In summary, the Star Series processors are very robust, delivering real performance on real applications for the next generation of 64-bit RISC commercial and server processors. All this while retaining optimum chip size and power. It achieves high performance on real application because of its low latency design and IBM's superior silicon technology. The Star Series can be expected to lead the commercial and server benchmarks for years to come

References

[1] J. M. Borkenhagen, et. al., "AS/400 64-bit PowerPC-Compatible Processor Implementation", IEEE ICCD94, pp192-196, Oct. 1994.

[2] D. Allen, et. al., "A 0.2um 1.8V SOI 550Mhz 64b PowerPC Microprocessor with Copper Interconnects", IEEE ISSCC99, Feb., 1999.

Notes
***PowerPC, PowerPC Architecture, IBM RISC System/6000, RS/6000, AS/400 are trademarks of the IBM Corporation.**

Company	Model	TPC-C Non Clustered	Model	SpecWeb	Model	NotesBench Mail only
IBM	AS400e 9406-S40 12W	43,169.85 $128.91	RS/6000 S70 12w	12031	RS/6000 S70 12W	40075 $15.32
HP	HP9000 V2250 16W	52,117.80 $81.17	HP9000 V2250 16W	13811	NetServer LPr	4631 $4.61
SUN	Enterprise 6500 24W	53,049.97 $76.00	Enterprise 450 4W	9115	Enterprise 450 4W	8464 $11.41
COMPAQ	AlphaServ er GS140 6/575 8W	37,540.30 $79.43	AlphaServer GS140 6/575 10W	14263	ProLiant 7000	6766 $13.22

Performance numbers from:
http://www.ideasinternational.com/benchmark/bench.html/
Data taken from website with January 13, 1999 listed as last modification date.

Biographies

John Borkenhagen is a Senior Engineer and Salvatore Storino is an Advisory Engineer in the Commercial MicroProcessor Design Department, IBM Server Group Development, Rochester, Minnesota.