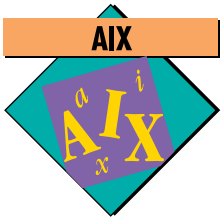
POWER2 Server Performance Analysis



By Maurice T. Franklin, Jacob Thomas, and Sohel R. Saiyed

This article compares the relative performance strengths of two IBM POWER2 Architecture servers so that our customers can make more informed comparisons between these systems.

This article compares the performance of two POWER2 Architecture™ implementations: the RISC System/6000 (RS/6000) 590 and RS/6000-59H. Although the two servers operate at the same clock frequency, their designs are different. Figure 1 shows the key characteristics of these servers.

Both servers (590 and 59H) are identical at the processor level, making them completely binary compatible. Both can execute up to six instructions (one branch, one conditional register, two fixed-point, and two floating-point) per cycle and have the same Instruction cache (I-cache) size (32 KB). The key differences between the two are in memory hierarchy: Data cache (D-cache) size, D-cache line length, presence of an L2 (secondary) cache, and memory bus bandwidth (the 590 D-cache is configured for a 256 KB and 8-word memory interface only if a system has at least four memory cards; otherwise, it will operate with a 128 KB cache and a 4-word memory interface). Another difference between the two systems is cost; the 59H has a higher list price than the 590.

POWER2 Performance Monitor

Every system based on the POWER2 Architecture includes an embedded hardware performance monitor. The monitor consists of 22 hardware counters that can count particular events as they occur on the POWER2 processor, and a control register that allows users to select which events to monitor at any given time. The monitor provides many useful performance measurements, including the number of elapsed cycles and executed instructions, counts of cache and TLB misses, and utilizations of execution elements. The counters can be accessed using a special AIX kernel extension to obtain a precise view of the system behavior when a program is executing—while having a negligible impact on the performance of the program.

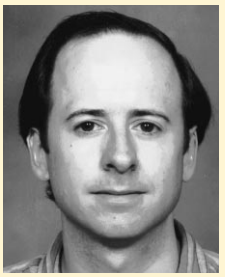
Benchmarks Measured

We analyzed the performance of three classes of workloads: floating point, integer, and Online Transaction Processing (OLTP). Three industry-standard benchmarks—SPECfp92™, SPECint92™, and the TPC-C benchmark—were selected to represent these three workloads. We believe these benchmarks represent the real workloads found in major segments of the server market worldwide.

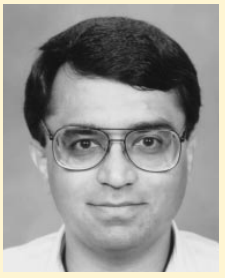
The SPECfp92 benchmark suite, published by the Standard Performance Evaluation Corporation (SPEC™), has five single-precision (32-bit) and nine double-precision (64-bit) floating-point programs. The suite represents the floating-point intensive workloads frequently encountered in workstation environments. Most codes in the suite consist of highly repetitive loops that operate on large sets of data accessed systematically.

A well-designed memory subsystem is key to the performance of this workload on the superscalar POWER2 processor, which can execute up to four floating-point operations per cycle. The fixed-point units, which also handle floating-point memory loads and stores, can perform two operations per cycle. The double-precision codes put an additional load on the memory subsystem since fewer of the 8-byte operands can be held in the D-cache, possibly requiring more frequent memory accesses.

SPECint92, an integer benchmark published by SPEC, consists of six integer workloads chosen from applications such as compilers and text pro-


Maurice T. Franklin


Jacob Thomas


Sohel R. Saiyed

Characteristic	RS/6000 Model 590	RS/6000 Model 59H
Clock frequency	66.7 MHz	66.7 MHz
Processor interface	8 Words	8 Words
Instruction cache size	32 KB	32 KB
Data cache/line size	256 KB/256 B*	128 KB/128 B
L2 cache	N/A	1 MB
Memory interface	8 Words*	4 Words

* Assumes four or eight memory cards

Figure 1. Key characteristics of the RS/6000 models measured

The design differences between the two POWER2 servers center on the memory hierarchy.

cessing. Unlike the SPECfp92 workload, in which the combination of floating-point and fixed-point units can rapidly consume the data, SPECint92 workloads are less sensitive to memory design since the two fixed-point units alone must execute all the arithmetic and memory load/store instructions, limiting the rate of memory references. Also, since integer data is only four bytes wide, the stress on the memory subsystem is less than for double-precision floating-point data.

The TPC-C benchmark simulates OLTP workloads. TPC-C, like the commercial workloads it represents, consists primarily of integer operations (such as pointer manipulation) with few repetitive loops and many branches to parts of the code not recently referenced. The OLTP executables are large (typically greater than 1 MB) and complex (two to three million instructions per transaction). As a result, the set of unique instruction addresses that are referenced with every transaction (the instruction footprint) is large. The combination of branching behavior and code path length makes the I-cache design a significant performance factor when executing the TPC-C workload.

These benchmark codes have a compilation aspect to their performance, which for many codes can often be improved by code optimization. Customers can use the RS/6000 compilers (C Set++ for AIX/6000 and AIX XL Fortran Compiler/6000) to perform advanced optimizations. Preprocessors such as KAP™ (Kuck and Associates, Inc.) and VAST™ (from Pacific-Sierra Research Corporation) are also available to optimize source programs. We used these products to increase the performance of the benchmarks within the rules specified by SPEC and the TPC-C council.

Processor Design Impact on Performance

The design differences between the two POWER2 servers center on the memory hierarchy. The Model 590 has a high-capacity Level-1 (L1) D-cache and an 8-word interface to memory, but no L2 cache. Model 59H has a large L2 cache (1 MB) with a smaller L1 D-cache and a 4-word interface to the memory. The L2 cache on the 59H is a combined instruction and data cache.

One key aspect of processor performance is the cache miss rate and miss penalty. Cache miss rates for a particular cache design (set associativity, cache size, and line size) vary between workloads depending on their memory reference pattern (for both data and instructions). For example, frequent non-localized branches increase the I-cache miss rate, while noncontiguous data references increase the D-cache miss rate. The cache miss penalty is the number of cycles to service a cache miss. The penalty depends on whether the missing item (instruction or data) is fetched from the L2 cache (requiring only a few cycles) or memory (a dozen or more cycles).

Another key aspect of processor performance is the degree to which a processor can perform instructions in a few cycles with as much parallelism as possible. This aspect of performance depends on the sequence of instructions encountered and the processor's superscalar design. The memory hierarchy and superscalar design aspects determine the average number of Instructions Per Cycle (IPC) completed by the processor during the execution of a workload. Since the 590 and 59H are identical at the processor level, only the changes in the memory hierarchy design affect the IPC. Lower miss rates or reduced penalties (or both) increase the IPC. Since the 590 and 59H have the same clock frequency, a higher IPC results in faster execution of the workload.

Performance Analysis

This section analyzes the relative cache miss rates and IPC of the three workloads gathered using the performance monitor. To highlight the performance differences, we show the relative increase/decrease in cache miss rates and IPCs.

Figure 2 compares the relative D-cache miss rates for the 59H and 590 for each of the SPECfp92 programs. The first column lists the name of the benchmark program. The second column shows the 59H to 590 L1 D-cache miss ratio: values greater than 1.0 indicate a 59H L1 D-cache miss rate higher than the 590 L1 D-cache miss rate. Column 3 shows the 59H L2 to 59H L1

D-cache miss ratio. Although the 59H L2 cache is a combined instruction and data cache, only the data miss rate is included in the ratio presented in column 3 (the I-cache miss rate is insignificant, generally for SPECfp92). Those ratios that are close to 1.0 in column 3 imply high 59H L2 data miss rates, with most of the misses from the L1 also missing in the L2 and going to memory. The IPC ratios of the two servers are presented in the fourth column. For benchmarks with IPC ratios less than 1.0, the 590 is better than the 59H.

Figure 2 clearly shows that the smaller D-cache on the 59H generally results in much higher miss rates than on the 590 for this workload. The high ratios in column 3 indicate that the 1 MB L2 on the 59H is not large enough for many of these benchmarks. Since the penalty to resolve references from the L2 is several cycles (compared to one cycle from the L1), a low miss rate from the L2 (low ratios in column 3) may not completely offset the high L1 miss rate of the 59H. Considering the IPC comparisons, the 59H can almost match and even occasionally exceed the 590's performance on these programs. Nonetheless, the L2 only marginally compensates for the smaller L1; in fact, the 590 is 4% faster on the overall SPECfp92 than the 59H. As a result, Model 590 has a slight performance and a significant price/performance advantage for this type of workload.

Figure 3 shows similar data for SPECint92 benchmarks. SPECint92 is moderately dependent on the L1 D-cache for high performance, but to a lesser extent than SPECfp92.

For the integer workloads, the 59H to 590 data cache miss ratios are smaller than some of the very large ratios (16.00 or 19.50) seen in some floating-point benchmarks (Figure 2). The L2 miss ratios are small for all SPECint92 benchmarks; therefore, most 59H L1 misses are L2 cache hits. For this type of workload, the caching of D-cache lines by the L2 effectively compensates for the smaller L1 D-cache on the 59H system. The overall performance of the two models is very comparable on this type of workload (the 59H is less than 1% faster), but Model 590 clearly leads in price/performance.

Figure 4 shows the same comparative data as the preceding figures for the TPC-C benchmark, with one significant addition. To highlight the importance of the I-cache for this type of workload, column 4 gives the L1 I-cache to L2 Instruction miss ratio on the 59H. The L1 I-cache miss rates of the 59H and the 590 are roughly the same since both servers have the same size (32 KB) I-cache, and so they are not compared. Since the 59H L2 cache is a combined cache, the 59H I-cache misses may be L2 hits. On the 590 system, all I-cache misses are resolved by accessing main memory.

As column 4 in Figure 4 indicates, most I-cache misses on the 59H are resolved by the L2. This implies the 59H design is highly effective in reducing the I-cache miss penalty, thus improving this workload's IPC. Since memory delays dominate TPC-C performance, these improvements significantly increase the IPC and overall performance of this type of workload. The resulting 16% improvement in performance seen in Figure 4 is dramatically higher than the overall through-

	Workload	Data Miss Ratio 59H L1/590 L1	Data Miss Ratio 59H L2/59H L1	IPC Ratio 59H/590
Single Precision	wave5	2.69	0.44	0.96
	alvinn	0.71	0.12	1.15
	ear	1.14	0.59	1.00
	mdljsp2	19.50	0.06	0.99
	swm256	1.91	0.90	0.96
Double Precision	spice	2.47	0.21	0.97
	doduc	16.00	0.14	1.03
	mdljdp2	3.00	0.07	0.99
	tomcatv	2.00	0.83	0.96
	ora	7.60	0.65	1.00
	su2cor	3.18	0.29	0.96
	hydro2d	5.06	0.18	0.97
	nasa7	1.57	0.38	1.13
	fpppp	3.20	0.13	1.01

Figure 2. Floating-point relative cache miss rates and IPC (SPECfp92)

Workload	Data Miss Ratio 59H L1/590 L1	Data Miss Ratio 59H L2/59H L1	IPC Ratio 59H/590
espresso	3.00	0.32	1.02
li	2.50	0.10	1.01
eqntott	1.92	0.22	1.05
compress	2.29	0.20	0.99
sc	4.66	0.23	1.01
gcc	3.00	0.35	1.07

Figure 3. Integer relative cache miss rates and IPCs (SPECint92)

Workload	Data Miss Ratio 59H L1/	Data Miss Ratio 59H L2/ 590 L1	Instruction Miss Ratio 59H L2/ 59H L1	Instructions Per Cycle 59H/590 59H L1
TPC-C	1.96	0.62	0.15	1.16

Figure 4. Transaction processing relative cache miss rates and IPCs

put changes of the other two workloads. When comparing configured system costs for workloads such as TPC-C, the 59H also has a clear price/performance edge over the 590. Therefore, the TPC-C workload—better than either SPECfp92 or SPECint92—illustrates the benefits of the design trade-offs made to the POWER2 implementation in the Model 59H to enhance its commercial performance.

Conclusion

All POWER2 Architecture implementations are identical at the processor level, differing only in memory hierarchy and bandwidth. These differences have varying impacts on the performance, depending on the characteristics of the workloads. The SPECfp92 workload type relies on a large D-cache and high memory bandwidth, and is handled slightly better by the Model 590. Generally, Model 590 exceeds the performance of the new 59H on traditional floating-point intensive workloads at a considerable cost advantage.

The middle ground for the memory design trade-off is seen in the SPECint92 workloads, in which the D-cache size is less critical, and the addition of an L2 can consistently, though marginally, improve performance. For such workloads, the two systems often have similar performance. For a transaction processing workload, the benefits of adding an L2 cache to the Model 59H are substantial and outweigh the loss due to the smaller D-cache and narrower bandwidth. For TPC-C and similar workloads characterized by large and complex integer manipulation code, the POWER2 Model 59H is clearly the best choice for a high-performance server, delivering higher performance and price/performance.

Together, these two POWER2 Architecture servers—both based on the same superscalar RISC technology and each with memory designs highly tuned to particular customer workload requirements—deliver world-class performance at competitive prices in their respective market segments.

References

- White, S. W. and Dhawan, S. "POWER2: Next Generation of the RISC System/6000 Family," *IBM Journal of Research and Development* 38 (1994): 493-502.
- Saiyed, Sohel R. and Thomas, Jacob. "The IBM POWER2 Architecture Implementations," *AIXpert* (August 1994): 55-57.
- Welbon, E. H.; Chan-Nui, C. C.; Shippy, D. J.; and D. A. Hicks. "The POWER2 Performance Monitor," *IBM Journal of Research and Development* 38 (1994): 545-554.
- Saiyed, Sohel R.; O'Connor, J. Michael; and Franklin, Maurice T. "POWER2: CPU-Intensive Workload Performance," *PowerPC and POWER2: Technical Aspects of the New IBM RISC System/6000* (SA23-2737): 129-136.
- Standard Performance Evaluation Corporation. "CINT92 & CFP92 Benchmark Descriptions," *SPEC Newsletter* 4, No. 4 (December 1992): 9.
- Transaction Processing Performance Council. *TPC Benchmark C Standard Specification Revision 2.0.* October 1993.
- Franklin, M. T.; Alexander, W. P; Jauhari, R.; Maynard, A. M. G; and Olszewski, B. R. "Commercial Workload Performance in the IBM POWER2 RISC System/6000 Processor," *IBM Journal of Research and Development* 38 (1994): 555-561.

Maurice T. Franklin, IBM Corporation, 11400 Burnet Road, Austin, TX 78758. Internet: maurice@austin.ibm.com. Mr. Franklin is a staff programmer in the AIX Software Performance group, currently developing software tools for performance tuning of the Workplace operating system. He has a BA in Computer Science from the University of Texas at Austin.

Jacob Thomas, IBM Corporation, 11400 Burnet Road, Austin, TX 78758. Internet: thomasj@austin.ibm.com. Mr. Thomas is an advisory programmer in processor and system performance in the RISC System/6000 Division. He has an MS in Statistics from Michigan State University and an MS in Physics from Birla Institute of Technology and Science in Pilani, India.

Sohel R. Saiyed, IBM Corporation, 11400 Burnet Road, Austin, TX 78758. Internet: saiyed@hwperform.austin.ibm.com. Mr. Saiyed is an advisory engineer in the RISC System/6000 Division Systems Architecture and Performance group. He has an MS in Computer Engineering from Clemson University and a BTech degree in Electrical Engineering from the Indian Institute of Technology in Kharagpur, India.