



Fujitsu: Taking Risk Out of RISC

Quick Note

David Freund
11 October 2002

Copyright © 2002
Illuminata, Inc. Licensed to
Fujitsu Technology
Solutions for web posting.
Do not reproduce.

When it unveiled SPARC64 V, the new version of its SPARC-architecture microprocessors last week, Fujitsu, Ltd. moved itself farther out of the shadow of Sun Microsystems and into the very top of the high-end commercial server market.

The last version of Fujitsu's chip, the SPARC64 GP, powers Fujitsu's 128-processor PRIMEPOWER servers, which have produced some eye-popping results on TPC-C and other benchmarks. Though Fujitsu is not widely known in North America as a server vendor, its products include the largest commercially-focused Unix servers available, well above the range of competing HP, IBM, and Sun product lines. That prowess, combined with its compatibility with SPARC/Solaris—which commands the largest user and software base in the commercial Unix market—makes PRIMEPOWER a particularly interesting Unix server option.

RAS features in the company's latest CPUs and servers bring mainframe-level reliability to the Unix server arena that rivals that of any high-end Unix server offering available.

Mainframe Systems Design

The reason Fujitsu can scale to such high processor counts and achieve such good performance is that much of its systems design and expertise comes from its experience designing mainframes. Fujitsu got its start in computers¹ in Japan in the mid 1950s. Then, in the early '70s, Fujitsu and German-based Nixdorf funded Gene Amdahl, a former IBMer who began shipping "plug-for-plug compatible" clones of the IBM mainframe in 1975. After market pressures forced consolidation of the mainframe market, clone-makers Amdahl, ICL, Nixdorf, and Siemens (some of which also manufactured flavors of Unix servers) all eventually become part of Fujitsu.

1. The term "mainframe" wasn't used until the '70s when smaller machines like the Digital Equipment Corporation PDP and VAX series, which were being called "minicomputers," became popular.



While most of its brand names have faded away, with the exception of the still-surviving Fujitsu GlobalServer mainframe, Fujitsu's understanding of Big Iron systems clearly remains.

The company's approach to the Unix-market during the past few years has been to focus on extremely high server scalability. PRIMEPOWER servers can be configured as large as 128 CPUs, giving it an engine with roughly twice as many cylinders as the largest comparable Sun and HP servers,² and four times as many as IBM.

However, the company has also quietly been working on system reliability—again using moves from its mainframe systems playbook.

White-Hot SPARCs Under the Hood

While Sun and its UltraSPARC III processors get the lion's share of the attention from the press and the market, SPARC64 conforms to the same SPARC V9 architecture specifications and runs the same software, including Solaris and Solaris applications.

The SPARC architecture definitions, however, explicitly allow a company to tune its implementation. Fujitsu's extensive tuning has optimized its version for complex, large-footprint applications. Advanced features like out-of-order execution and register renaming give it better performance, even when its microprocessors are fabricated at lower clock speeds than competing designs such as UltraSPARC III.

The new SPARC64 V takes the next step by not only increasing raw chip performance, but also expanding RAS and I/O capabilities. For example, it comes with wide data pipes within the processor, supports up to 16 out-of-order memory operations outstanding at any time, and provides faster bus connections and loads of on-chip cache. The result is

2. While Sun likes to claim that its E15K can be configured with up to 106 CPUs, any processors added after 72 come must be traded off against I/O bandwidth. A hundred-processor 15K would be hugely I/O-starved on anything but a pure compute-bound HPC job.

a SPARC chip that has roughly the same integer and floating-point performance as IBM's POWER4 at similar clock speeds.³

But scaling to the level of a 128-way PRIMEPOWER 2500 requires fatter pipes not just within each processor, but also connecting processors and processor boards to the rest of the system. So Fujitsu not only doubled the capacity of its board—which can now hold up to eight CPUs at an eventual top clock speed as high as 2.5 GHz—but also doubled bandwidth on and off each board, and in the system backplane. The resulting aggregate bandwidth of the system is almost three times that of a Sun Fire 15K, twice that of HP's Superdome, and about 62% greater than IBM's p690. To feed these pipes, Fujitsu also doubled the number of PCI slots, to 320.

Previous PRIMEPOWER machines could only be sliced and diced at system-board level. Just doubling the number of CPUs per board would make partitioning less flexible. So the new machines add configurable interconnect switches on each system board, allowing partitions within a single system board (which Fujitsu calls "extended partitions" or XPARs) that can be as small as two CPUs.⁴

This is an impressive set of specs, but how does the machine perform under real data center back-end workloads? Since the model 2500's predecessor still holds the TPC-C speed title, the new machine is almost certain to maintain that first-rung position (although we won't know for certain until the results are published). Sun's reluctance to participate in the TPC-C race makes it a bit difficult to compare the PRIMEPOWER's database performance with Sun Fire servers. Still, all indications are that PRIMEPOWER servers, especially those using the latest upgrades, will be able to outpace anything Sun is currently selling.

3. According to vendor estimates. SPEC test results have yet to be published for the SPARC64 V.

4. This works even without both processors installed. So a partially-populated system board could have partitions that contain only one CPU.

Keeping the Engine Running

But no matter how fast an engine, it has to run reliably to be of long-term commercial benefit. Here again PRIMEPOWER's mainframe heritage stands it in good stead.

When it comes to RAS smarts, the new SPARC64 V microprocessor clearly inherited even more from its mainframe forebears than the preceding SPARC64 GP. The GP chips had already used ECC protection on internal paths, cache data and cache tags—except for L1 cache tags, which had only parity protection. The new processor adds a redundant copy of the L1 tags, and error detection on the arithmetic units and registers. But the real RAS news is the processor's ability to *recover* from arithmetic or register errors using a new capability called "instruction retry." If an instruction fails, the CPU automatically tries again—eliminating errors caused by intermittent problems like electrical noise, which has become increasingly likely as chip circuitry has shrunk. Though this function has been in mainframes seemingly forever, it's brand new for RISC processors.

Fujitsu's list of other RAS enhancements is long and strong, including redundant and hot-swappable components, redundant I/O paths, extensive error checking, fault isolation, graceful degradation, and live updates at many levels of system design. A system software failure can be isolated to one par-

tion, so the rest of the machine rides through. Cache failure? The system turns the failed cache off. I/O channel problem? It uses the other path. CPU failure? It dynamically removes the glitched chip from the running partition. While no Unix server has completely matched the availability of proprietary mainframe and fault-tolerant systems, progress is being consistently made. Fujitsu's latest processors and servers provide a great example of this, and of the value of bringing Big-Iron experience to bear.

Conclusion

For many U.S. customers, Fujitsu is not the first name that comes to mind when shopping for Unix servers. Those who've heard of it often see it primarily as a second-source for SPARC/Solaris. While it certainly is that, SPARC64 V and ongoing PRIMEPOWER enhancements continue to prove that Fujitsu has the goods to be much more. Beating everyone on TPC-C and other key benchmarks, as PRIMEPOWER has done, demonstrates that Fujitsu has the horsepower and the wherewithal. Beating IBM in transplanting instruction retry from the mainframe into RISC/Unix indicates it also has produced more innovation than a second-choice supplier can usually offer. Clearly, Fujitsu is a Unix-server company to watch.