

Hal Packs Sparc64 Onto Single Chip

Hal's Single-chip Sparc64 Offers Compatibility to Sun's Hardware, Software



by Peter Song

Packing its initial six-chip design onto a 240-mm² die, Hal rolled out the Sparc64-III processor at

October's Microprocessor Forum. Hisashige Ando of Hal Computer Systems described the company's latest processor, which offers much better compatibility with Sun's hardware and software products than do its previous two designs. In another first for Hal, the chip offers multiprocessor capability, which has become mandatory for today's server-class processors. If systems using the new chips are available by 3Q98, as Hal plans, they could fill a gap in high-end SPARC servers left open by Sun.

Hal's Sparc64-III (HS-3) delivers better overall performance than Sun's UltraSparc-2 (see MPR 11/13/95, p. 20) at comparable clock speeds. It has an out-of-order execution engine that can process 63 instructions at once, more than twice the number of instructions possible with the UltraSparc processor's in-order core (see MPR 10/3/94, p. 1). It also has two floating-point and two load-store units, delivering twice as many floating-point results per cycle as UltraSparc-2 (US-2). In addition, its instruction and data caches are four times larger and use a four-way set-associative organization, incurring lower miss rates than the caches in US-2. Furthermore, the HS-3 uses separate buses for the L2 cache and system interfaces, providing greater sustainable memory bandwidth than US-2 at comparable clock speeds.

The Sparc64-III chips are likely to top out at 250 MHz, however, losing much of their performance edge to the faster UltraSparc-2 processors that are already shipping at 300 MHz and are likely to reach 400 MHz in 2H98. For scientific and engineering server applications, which are floating-point intensive and use huge amounts of data, HS-3 systems may still outperform the faster US-2 systems. In addition, the HS-3 chips use ECC or parity in TLB and cache arrays, making them more reliable and suitable in mission-critical applications than the US-2 chips, which do not. Hal must deliver these robust systems, however, before UltraSparc-3 systems become available, which we expect to ship in 1H99 (see MPR 10/27/97, p. 29).

According to Hal, Fujitsu is shipping only a few percent of its SPARC systems using Hal's multichip processors, recouping only a fraction of its investment in Hal. Fujitsu—a \$40 billion company—is not deterred by lack of direct return on its investment, however, and values highly the side benefits Hal brings by advancing various VLSI technologies within Fujitsu. To remain competitive in the systems business, Fujitsu may have little choice but to rely on Hal.

Hal Uses an Existing Core, Redesigns the MMU

The feature set of the Sparc64-III is very similar to that of Hal's first two Sparc64 processors. Hal's Sparc64-I (see MPR 3/6/95, p. 1) has nearly 22 million transistors, 87% of which are in the MMU (memory management unit) and four identical cache chips. To provide flexible sharing and protection on variable-size memory segments, the original MMU chip supports a proprietary two-level translation scheme. It uses a view lookaside buffer to translate from virtual to logical addresses and a TLB (translation lookaside buffer) to translate from logical to physical addresses. It also implements in hardware the tablewalk mechanism—the algorithm that searches page tables in memory and reloads them into TLBs—and buffers to cache page tables, resulting in a complex memory-management scheme that is incompatible with Sun's Solaris. Instead, Hal developed its own operating system and software-development tools.

Hal's second processor combines similar MMU and cache chips with an improved CPU chip. That CPU chip has an 8K instruction cache and a 2K BHT (branch history table), twice as large as in the original design. The number of register windows is also increased from four to five. Fabricated in Fujitsu's 0.34-micron CS-60 process, the Sparc64-II operates at speeds up to 161 MHz and is currently shipping in Hal 385 systems.

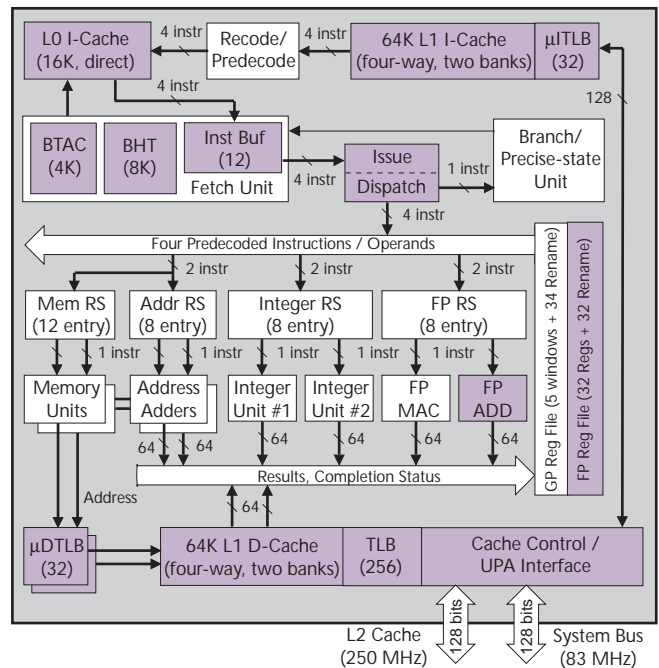


Figure 1. Hal packs its existing CPU and cache chips, with minor improvements, and a new memory subsystem onto a single chip.

Hal's latest design integrates the second-generation CPU, two cache chips with minor improvements, and a new memory subsystem onto a single die. Integrating two of the four 4.6-million-transistor cache chips reduces the L1 (level-one) caches to 64K each, as Figure 1 shows, but allows the entire design to fit on a 240 mm² die. Due to the external cache, integrating the two remaining cache chips would have gained only 1–2% on SPEC95 performance while increasing the die size to about 310 mm². The two-million-transistor MMU chip is replaced with TLBs, an L2 cache interface, and a system bus that is compatible with Sun's UPA (UltraSparc Port Architecture) bus.

Relying on TLBs to translate from virtual to physical addresses and software to manage the TLBs, the new MMU design is not only simpler but is also compatible with Solaris. In addition to the 8K, 64K, 512K, and 4M page sizes, which UltraSparc-2 supports, the new MMU supports 12 more page sizes, ranging from 4K to 4G. The page sizes larger than 4M are added to better support huge applications, such as operating systems or database programs. The HS-3 has three micro-TLBs to translate three addresses in each cycle—one for instruction fetch and two for load/store accesses. Each micro-TLB has 32 entries and, to support 16 different page sizes, each entry is fully associative. An access that misses the micro-TLB takes four extra cycles to access the main TLB, which has 256 fully associative entries.

UPA Bus Leverages Sun's Chip Sets

The original MMU chip uses proprietary interfaces for main memory and I/O devices, requiring Hal to design its own chip sets. In contrast, the HS-3 uses a UPA-compatible system interface, enabling it to work seamlessly with chip sets from Sun Microelectronics (see MPR 5/30/95, p. 12). The new interface makes it easier for Hal to design systems using the HS-3 chips.

Unlike UltraSparc-2, which shares a 128-bit-wide data bus between the L2-cache and system interfaces, the HS-3 has separate 128-bit-wide data buses for the two interfaces. The cache interface operates at the full speed or half the speed of the processor, delivering 4 Gbytes/s of peak bandwidth using 250-MHz SRAMs. The system bus can operate at 1/2, 1/3, 1/4, or 1/5 the speed of the processor, up to a maximum of 150 MHz. The dual buses eliminate the need for the external UDB (UltraSparc Data Buffer) chips required by US-2 but increase the Hal processor's pin count.

The Sparc64-III not only provides more sustainable memory bandwidth but also uses the available bandwidth more efficiently than UltraSparc-2. Due to its write-through data cache, US-2 can spend a significant portion of its memory bandwidth on individual stores to the L2 cache. The data

cache in the HS-3, in contrast, uses a write-back policy, eliminating much store traffic to the L2 cache. The combination of write-back and write-allocate policies effectively replaces all store traffic, which tends to use only 4 of the 16 bytes of the data bus, with the cache lines that are being evicted. Although UltraSparc-2 can combine two stores into one, the stores must be successive and to consecutive addresses, limiting the effectiveness to special cases.

Designed to Enhance System Reliability

Eliminating individual stores to the external cache makes using ECC (error-correcting code) on the L2-cache and system interfaces easier, enhancing the reliability of HS-3 systems. Because ECC requires multiple bits to encode a group of bytes (i.e., eight bits per eight bytes), individual stores that modify fewer than the number of bytes in a group require read-modify-write operations on the error-correcting codes. Since the HS-3 updates an entire cache line at a time, it can generate the 16-bit ECC for each 16-byte transfer without using a read-modify-write operation. In contrast, US-2 systems use parity on the L2-cache interface, avoiding complex read-modify-write operations on the interface but compromising system reliability.

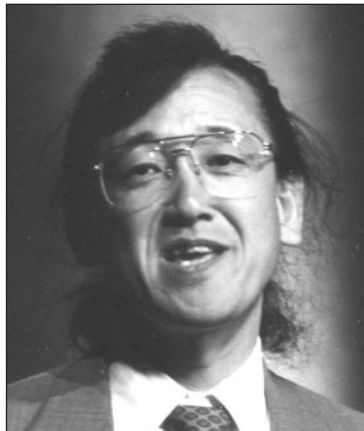
All on-chip cache and TLB arrays in the HS-3 are also protected with parity or ECC. The data arrays are protected by ECC, since they may contain the only valid data in the system. The tag arrays are protected by parity, since the L2 cache—being inclusive of the on-chip caches, as required by the UPA protocol—is guaranteed to have another copy of the tags.

In contrast, the on-chip arrays in Sun's UltraSparc-2 do not use parity or ECC protection. Furthermore, its merely parity-protected L2 cache makes US-2 less acceptable in mission-critical applications. Sun plans to remedy this problem with UltraSparc-3, which uses parity or ECC in all arrays larger than 1K. Using a 2K write cache, UltraSparc-3 also eliminates individual stores to the L2 cache, making it easier to support ECC on the L2 cache.

Adding Dispatch Stage Improves Speed

In the original Sparc64-I, instructions are decoded and issued to their respective execution pipelines in one cycle. Although the instructions are recoded—certain bits are rearranged and even encoded differently than defined in the instruction set—and four extra predecode bits are used for each instruction, the issue stage could not process four instructions in the target cycle time of four nanoseconds. In the Sparc64-III, a dispatch stage is added to relieve critical timing paths in the issue stage.

In each cycle, up to four instructions are fetched from the L0 (level-zero) instruction cache and placed into the



MICHAEL MUSTACCHI

Hal Microprocessor Group's general manager Hisashige Ando describes the Sparc64-III design.

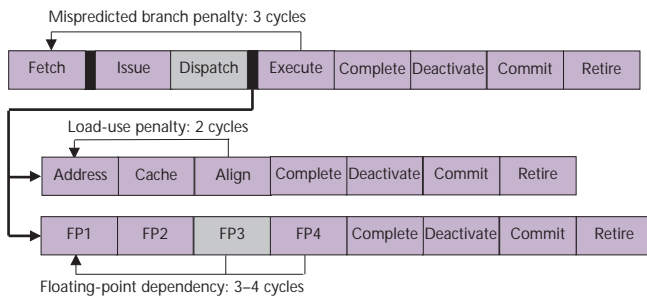


Figure 2. The Sparc64-III adds the dispatch stage, improving speed but incurring an extra cycle of branch-misprediction penalty.

12-entry instruction buffer. The 16K direct-mapped L0 instruction cache provides the recoded instructions and the predecoded bits. In the issue and dispatch stages, four instructions from the buffer are decoded, their destination registers are mapped to rename registers, and the rename registers for their source operands are identified. The source operands are read from the register file or the rename registers or forwarded from the execution units via result buses. The instructions and their operands are sent to the appropriate RSs (reservation stations) by the end of the dispatch stage. Each RS can accept two instructions per cycle, provided it has enough available entries. The load/store instructions are sent to both the address-generation and load-store RSs.

Adding the dispatch stage also adds a cycle to the branch-misprediction penalty, as Figure 2 shows. To reduce performance loss, however, the HS-3 uses a larger BHT and a two-level prediction scheme, known as Gselect (see MPR 11/17/97, p. 22), improving the branch-prediction accuracy. Its 8K BHT is indexed using the 5-bit global history register concatenated with eight bits of the branch instruction address, providing two bits of branch history. In contrast, UltraSparc-2 provides two bits of branch history for each instruction pair, which corresponds to having a 2K-entry BHT, and it does not maintain the global branch history.

In the HS-3, the global history register and the BHT entry are accessed during the fetch stage. Instead of waiting until the branch condition is known, the HS-3 speculatively updates the global history register and the BHT entry one cycle after the branch instruction is fetched. Since the branch is likely to repeat its last branch direction, its BHT entry is updated to the strongly-taken or strongly-not-taken state if the current state is weakly-taken or weakly-not-taken, respectively. When a branch is mispredicted, its BHT entry and the global history register are repaired.

Sparc64-III Adds a Second Floating-Point Unit

For the HS-3, Hal improves on its second-generation CPU to deliver a core that can outperform US-2 on floating-point-intensive applications. UltraSparc-2 has separate floating-point multiply and add units, allowing an add and a multiply to begin each cycle. It has only one load-store unit, however, often starving the FP units due to inadequate memory bandwidth. In contrast, the Sparc64-II has two load-store units

	Sparc64-III	UltraSparc-2	UltraSparc-3
	250 MHz	400 MHz	600 MHz
64-bit Integer Mul	6 cycles	5-35 cycles	9 cycles
64-bit Integer Divide	2-37 cycles	64 cycles	64 cycles
FP Add (DP)*	3 or 4 cycles	3 cycles	4 cycles
FP Mul (DP)*	4 cycles	3 cycles	4 cycles
FP Divide (SP)	12 cycles	12 cycles	17 cycles
FP Square Root (SP)	12 cycles	12 cycles	24 cycles
FP Divide (DP)	22 cycles	22 cycles	20 cycles
FP Square Root (DP)	22 cycles	22 cycles	24 cycles

Table 1. The Sparc64-III and UltraSparc-2 have comparable execution latencies, but the Sparc64-III can execute two floating-point add operations per cycle compared with one in UltraSparc-2. *pipelined operations. (Source: vendors)

but only one multiply-add unit, limiting the processor to begin at most one floating-point operation per cycle. The Sparc64-III adds a second floating-point unit, increasing to two the number of floating-point operations that can begin in each cycle.

The second FP unit executes only add and subtract instructions, however, occupying only half the area of the multiply-add unit. Since the overwhelming majority of existing SPARC binaries are compiled for systems from Sun, which do not support multiply-add instructions, a second multiply-add unit would offer little or no advantage in executing existing binaries. Keeping the multiply-add unit offers an advantage over replacing it with a multiply unit, since it can execute the second add or subtract instruction in each cycle. The multiply-add unit takes one more cycle to execute an add instruction than does the add unit, however. Table 1 shows that UltraSparc-2 has similar or better floating-point latencies despite its higher clock speed.

In addition to its superior floating-point capability, the Sparc64-III has a highly out-of-order execution core, allowing it to tolerate cache misses better than UltraSparc-2. Each integer, address-generation, and floating-point reservation station has eight entries and dispatches up to two instructions per cycle. Since the operands use rename registers, there is generally no restriction in dispatching instructions out of order, provided that the oldest two are dispatched when more than two are ready. Certain load/store instructions must be executed in program order, however, such as when the programming model requires an in-order execution or when the instructions access the same address. The data cache has two banks, interleaved on eight-byte boundaries, to support two accesses per cycle.

In contrast, UltraSparc-2 executes instructions in program order and does not use rename registers. When an instruction does not have all of its operands and is therefore not ready to execute, US-2 stalls that instruction and those that follow. It resumes the instruction dispatch only when the stalled instruction is ready to execute. To reduce pipeline stalls, it must rely on software to schedule long-latency and data-prefetch instructions well ahead of when their results are needed. Using a more efficient core, better branch prediction,

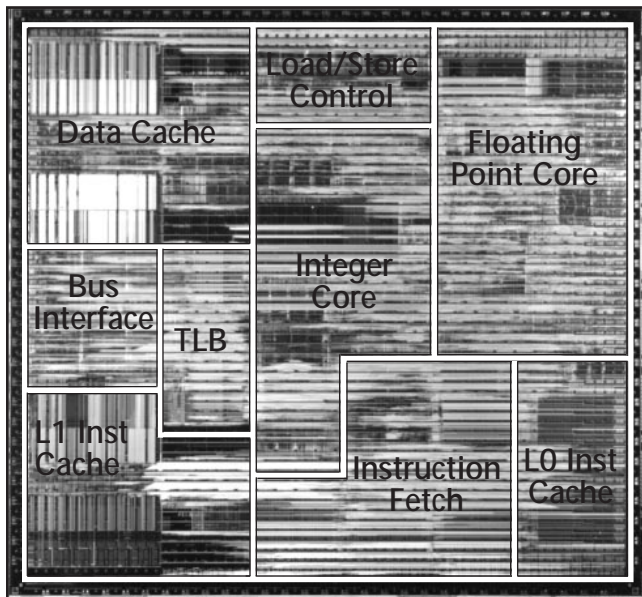


Figure 3. The TLBs and the 144K of caches occupy 40% of the 240-mm² die in a 0.24-micron five-layer-metal process.

larger on-chip caches, a longer cache-line size, and hardware prefetch, the Sparc64-III should perform better than UltraSparc-2 on server applications that use large data sets and intensive floating-point operations.

Hal's First Single-Chip Design to Hit 250 MHz

The Sparc64-III is built in Fujitsu's 0.24-micron five-layer-metal CS-70 process (see MPR 9/16/96, p. 11). The chip first taped out in July of this year and has already gone through a round of bug fixes. According to Ando of Hal, extensive tests indicate that the processor is largely functional and is expected to operate at 250 MHz in systems. Systems using the HS-3 chips are planned to debut in 3Q98.

Figure 3 shows a die photo of the HS-3. The chip has 17.6 million transistors, of which 11.6 million are in the caches and TLBs. Compared with only 1.5 million logic transistors in US-2, the HS-3 uses 6 million transistors for its out-of-order execution engine. The chip is housed in a 957-pin LGA package using flip-chip bonding technology. Using 3.3 V for the I/O and 2.5 V for the core, the chip is expected to have a peak power consumption of 50 W, according to Hal. The MDR Cost Model estimates the manufacturing cost of the Sparc64-III chip to be \$250, three times the cost of the 149-mm² UltraSparc-2.

Because the chip is not yet fully functional, Hal has only simulated performance numbers. At 250 MHz, Hal expects the chip to deliver 13 SPECint95 and 18 SPECfp95 (base), using a 4M L2 cache and 60-ns EDO DRAMs. The simulation model is of an existing system from Fujitsu, which has 10 and 46 cycles of L2-cache and main-memory latencies, respectively. Hal expects better performance for its high-end systems, which will have 250-MHz SRAMs and eight-way interleaved main memory. These numbers place the perfor-

Price & Availability

Hal expects samples of the Sparc64-III chips to be available in 2Q98 and systems to debut in 3Q98. For more information, contact Hal Computer Systems at 408.379.7000 or access the Web site www.hal.com.

mance of 250-MHz HS-3 systems nearly 30% above that of 300-MHz US-2 systems. Sun is already shipping 300-MHz systems today, however, and is expected to offer 400-MHz systems by the time HS-3 systems ship in 3Q98. Hal is also working on a 0.18-micron derivative of the HS-3 design, hoping to reach a clock speed of 500 MHz.

The SPEC95 benchmark, however, is an inadequate measure of value for chips designed to process a huge numbers of floating-point and database operations. On applications that incur high rates of cache misses, systems using Hal's latest processor should deliver better performance than systems using the faster US-2. Competitors to the HS-3 are not US-2 but IBM's Power3 (see MPR 11/17/97, p. 23) and HP's PA-8500 (see MPR 11/17/97, p. 20) in 2H98. These processors offer superior integer and floating-point performance and can tolerate cache misses better than US-2. They also offer a comparable level of reliability expected of mission-critical servers.

Sparc64-III Complements UltraSparc-2

Until UltraSparc-3 is ready for deployment in 1H99, Hal has a chance to establish the Sparc64-III in mission-critical SPARC servers. The HS-3 chips offer performance and features that complement those of UltraSparc-2. In fact, Hal is designing systems using UltraSparc-2 processors and plans to upgrade some of the systems using the HS-3 chips. Offering compatibility with Sun's hardware and software, the HS-3 should ease the planned transition.

Although it hasn't in the past, Hal is willing to sell its latest processors to Sun, which can use the chips to complement its product lines. With its marketing strength and brand recognition, Sun can probably compete with HP and IBM better than Hal/Fujitsu can, establishing a bigger market share for SPARC/Solaris systems. Hal's new chips could enable Sun to establish a foothold in high-end, mission-critical server markets without waiting for UltraSparc-3. Assuming UltraSparc-3 is on schedule, however, Sun may not want to switch to the HS-3 only to switch back to US-3 a few months later.

In less than two years, Fujitsu's needs in mission-critical systems can be met by Sun's UltraSparc-3. In theory, Fujitsu could then disband the Hal team and rely on Sun for all of its SPARC processors. The Japanese vendor counts on Hal for technology, however, not just products. Furthermore, Fujitsu competes with Sun in the systems business, so it expects Hal to continue delivering processors, like the Sparc64-III, that are differentiated from Sun's. □