

HotRail Rides With New Core Logic

New SMP Chip Set to Support AMD Athlon, Alpha, and More

by Peter N. Glaskowsky

After six years as Poseidon Technology, and four years since the release of its only previous product, newly renamed HotRail is preparing to release a new chip set for symmetric multiprocessing (SMP) servers and workstations. HotRail (www.hotrail.com) is working with AMD to develop the first version of its chip set for Athlon (see MPR 7/12/99, p. 1), to be followed by versions for Alpha and other processors.

The new core logic, which has not yet been named and is still some months away from being taped out, uses a new cache-coherent switch-based architecture to achieve levels of throughput unmatched in PC-based commodity servers. HotRail's (and AMD's) primary competition for the planned four-way and eight-way configurations will come from Intel's standard high-volume (SHV) server designs based on the four-way 450NX chip set (see MPR 7/13/98, p. 11) and the eight-way Profusion chip set (see MPR 9/16/96, p. 9), but HotRail appears to have some substantial advantages over these alternatives.

Switching Architecture Eliminates Bus Arbitration

At the heart of the new core logic is a single-chip multiport switching fabric that can route requests from any port to the appropriate agent on any other port. As Figure 1 shows, the fabric is nearly symmetrical; the only asymmetry is that the ports assigned to memory channels cannot initiate transactions.

Each port supports one 3.2-GByte/s point-to-point HotRail Channel (HRC) interface. Each port may be connected to HRC-equipped bridges for CPU, memory, or I/O, though memory bridges will usually be connected to only

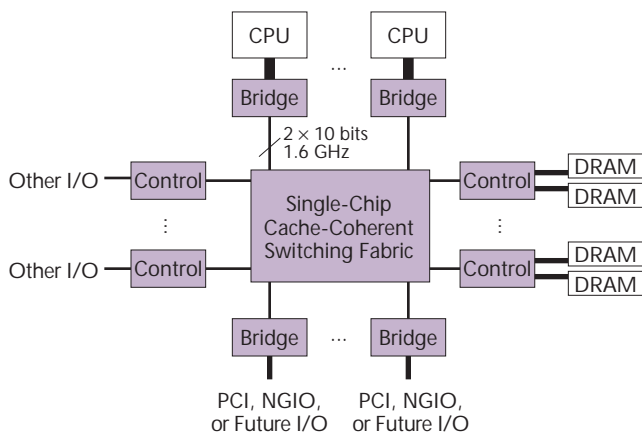


Figure 1. Four channels in HotRail's switching fabric are assigned to memory. The remainder may be used for either CPUs or I/O.

the predefined memory ports. Memory cache coherency is maintained by the switch fabric itself, issuing coherency requests to the CPUs as needed. The other ports may be assigned to processors or I/O devices as desired, providing valuable flexibility for server OEMs. For example, an OEM can use the HotRail chip set to create an application server with the maximum number of processors, or to create a file server with more I/O and fewer CPUs.

HotRail is designing the bridge chips needed for the first systems, but it plans to license the HRC specification in hopes that peripheral-chip makers may design products for it. Each time the system powers up, the switch chip configures its ports according to the devices connected to it.

The HotRail switching chip contains internal buffering to accommodate simultaneous requests from multiple initiators to a single target, a common concern for all switching architectures. A feedback mechanism can delay requests to prevent buffer overflow.

Fast Channels Provide Bandwidth Headroom

Each HotRail Channel is implemented as a pair of unidirectional 10-bit differential buses with separate source-synchronous differential clocks for each bus—44 signal pins total. Each of these signals operates at 1.6 Gbits/s. The channel carries packetized data. The packets contain 64 bits of data, eight bits of control information, and an eight-bit Hamming-code error-detection field. These 80 bits are transmitted in eight clock phases, yielding a throughput of 1.6 GBytes/s for each direction or 3.2 GBytes/s total per port.

This is roughly twice the throughput needed for the initial 200-MHz speed of an AMD Athlon bus interface and about 30% faster than the 2.5-GByte/s sustained throughput of the initial Future I/O interface. HotRail plans to support SDRAM and double-data-rate (DDR) SDRAM on its first memory controllers. Each port will connect to two 64-bit memory arrays at speeds of at least 100 MHz. HotRail says that two 64-bit, 100-MHz DDR SDRAM arrays will provide close to 3.2 GBytes/s of sustained throughput, matching the bandwidth of the channel. Memory may be interleaved on power-of-two multiples of cache-line boundaries, or each memory array may be assigned to a single contiguous range of addresses.

As processors, memory, and I/O devices speed up over time, HotRail is prepared to boost the speed of the HotRail Channel to compensate.

HotRail developed the HRC electrical interface for this chip set. The company describes it as a constant-current, controlled-swing signaling technology that uses low voltage levels and consumes little power. No external terminating

resistors are required. The driver circuits in each chip offer adjustable output impedance over a range of 47 to 62 ohms, and the receivers can adjust to remove skew. Both adjustments are made automatically during chip initialization.

Each channel can span up to 18 inches of total trace length with one connector pair in the path. Higher frequencies will be possible at shorter distances, while greater distances may be achieved by reducing the operating frequency from the standard 1.6-GHz rate. HotRail is currently qualifying connectors for the channel but believes some existing commodity connectors will work.

The eight-way version of HotRail's switch chip will have 14 such channels—4 for memory plus 10 for CPUs and I/O. The theoretical peak throughput of a fully configured system is seven times the 3.2-GByte/s figure, since each transaction involves two endpoints. Our analysis suggests that the peak throughput of the eight-way configuration is achieved only when two CPUs and the two I/O channels are accessing memory while the remaining six CPUs perform cache-coherency transactions among themselves. This combination of operations will be seen rarely, however.

The practical bandwidth limit for HotRail systems is imposed by the memory subsystem, which will participate in almost all transactions in a server. At up to 12.8 GBytes/s, the effective throughput of the HotRail chip set is well above what we expect to see in Profusion-based systems, which are expected to have just two DRAM channels.

Switch Architecture Forces Some Compromises

Perhaps the most significant problem with the HotRail architecture is the extra latency added by the relatively long path each transaction must take through the chip set. Processor memory reads pass through the CPU bridge, the switching chip with its internal buffers, a memory controller, and out to the DRAM array, which adds additional latency. The read data must return to the CPU over the same path. By the time the data gets back to the CPU, the total transaction has passed through the equivalent of seven chips and four instances of the HotRail Channel.

Each HRC transit imposes significant latency. One complete packet must be prepared at the sending side before the packet can be placed on the channel. A minimum latency of eight channel clocks plus flight time (about 7.5 ns total) is imposed by each transit. Additional clock periods (at the unspecified but presumably lower internal operating frequency of the HotRail chips) will be required to translate HRC packets to and from the native formats of the Athlon, SDRAM, and PCI interfaces.

HotRail estimates this overhead will impose some 30 ns of extra delay compared with Intel's 450NX four-way server chip set. The company points out—correctly, we believe—that its advantage in sustained throughput for the whole system is much more important for most server applications.

Device and system complexity provides another challenge for HotRail. Switched architectures naturally require

more pins than bused architectures. The 14-port HotRail switch chip has a million gates of logic and 616 signal pins just for its HotRail Channel interfaces. A complete eight-way server will use at least 15 HotRail chips.

OEMs are not likely to be dissuaded by these issues, however. Most of the cost in a HotRail server will come from the Athlon processors and other components, and the performance potential of the HotRail solution should be very attractive to server makers.

Roadmap Includes Several Variations

HotRail and AMD will have to deliver impressive products to gain entry to the Intel-dominated server market. AMD has no presence in high-end servers today, and server buyers are a conservative lot. HotRail's plans cover two of the key issues for this market: compatibility and scalability.

To ensure its products work as planned, HotRail has licensed the Athlon and Alpha EV6 buses from AMD and is validating its products against AMD's Irongate chip set. Phoenix provided an Athlon BIOS as well as PCI and AGP core designs. HotRail plans to use IBM's 0.25-micron CMOS-6SF foundry process to manufacture its chips.

HotRail expects to release its four-way (10-channel) and eight-way (14-channel) chip sets close together, possibly followed by a simpler two-way chip set for Athlon workstations. The company is considering a 16-way design for further in the future. HotRail will offer its own reference designs for its chips, and it plans to sell these chips directly to OEMs rather than through AMD.

HotRail is focused on Athlon today, the company is otherwise processor neutral. Once it has proved the value of its unique architecture, it plans to approach Intel and the major RISC processor vendors, hoping to achieve additional design wins. Though Intel's Corollary group continues work on an eight-way solution of its own, Intel is surely pragmatic enough to consider alternatives—especially if HotRail's chip set can outperform Profusion.

San Jose-based HotRail is also looking ahead to other applications. The company believes that its switch-based architecture is a natural match for the needs of high-performance network controllers such as ATM edge switches, Internet firewalls, and encrypting routers. In such applications, the HotRail switching fabric might be used to connect multiple local-area networks to the Internet or to connect a single server farm to multiple Internet gateways. Each HotRail Channel is more than adequate to support an OC-192 (10-Gbit/s) ATM connection, which is four times faster than the OC-48 links that provide the fastest Internet connections today. Conventional CPUs might be replaced by dedicated routing engines in such a system, but little else in the HotRail architecture would need to change.

Switching technology is a natural fit for many applications in computing and data communications. If it can execute on its ambitious plans, HotRail is likely to find at least one such application to call its own. 