Online:

SYSTEM BOTTLENECKS HAVE MOVED FROM CPU TO MEMORY AND NOW TO THE I/Os.

Brian Holden
Principal Engineer
PMC-Sierra's MIPS Processor Division
brian_holden@pmc-sierra.com

HyperTransport

With communications and networking technologies advancing so rapidly, it becomes increasingly clear how much each component depends on the links to other components in the system.

As each component achieves new performance plateaus, the other system elements must keep pace, or the value of the advancements is lost to the systems designer. For example, the current lagging performance of I/O bus architectures is actually limiting the impact of recent, significant gains made in microprocessor performance. Without a comparable I/O bus solution that can handle these high-performance processors, the industry will not be able to realize the full potential of recent advances.

The answer to this dilemma may be HyperTransport I/O technology. This scalable architecture delivers dramatically increased bandwidth over existing bus architectures. Also, it simplifies in-the-box connectivity by replacing legacy buses and bridges. It's immediately available, scalable, flexible, fast, and cost-effective. In addition, it's compatible with legacy PCI buses and emerging standards. So, HyperTransport has been heralded as the solution the industry has been seeking.

The problem becomes evident with some simple math. While microprocessor performance doubles every 18 months, I/O bus architecture performance only doubles about every three years. The resulting I/O bottleneck constrains overall system performance by hindering the gains made in the processor and memory subsystem arenas.

Further complicating processor system designs, a number of legacy buses—including ISA, VL-Bus, AGP, LPC, PCI-32/33, and PCI-X—are connected to support a wide range of devices. This increases system complexity, adds active devices for bus arbitration and bridge logic, and in the end, provides substandard performance. The many legacy buses need increasing numbers of signal pins to connect to the chip packages. These additional pins require additional power and ground pins to deliver adequate electrical current return paths. The extra power also generates increased heat, further diminishing system utility.

Additional bandwidth demands from increasingly complex 3D graphics processing, high-speed networking, wireless communications, and software applications are straining today's already overworked PCI bus. Furthermore, added interface functions such as MP3 audio, v.90 modems, USB, 1394, and 10/100 Ethernet all compete for any remaining bandwidth. Together, the bandwidth needs quickly exceed the capabilities of the PCI bus.

A HyperTransport link is a high-speed, high-performance, point-to-point connection between ICs that provides a high-performance link for embedded applications and enables highly scalable multiprocessing systems. It offers a universal bus type to reduce the number of buses within the system. HyperTransport's scalable architecture vastly increases bus transaction throughput over PCI, PCI-X and AGP, and other currently available I/O bus architectures.

HyperTransport delivers the higher data rates required for high-performance embedded applications such as networking and communications, supporting up to a peak aggregate bandwidth of 12.8 Gbytes/s for a 32-bit wide link.

HyperTransport empowers the chips deployed in PCs, servers, networking, and communications devices to communicate with each other much faster than current technologies will allow, enabling highly scalable multiprocessing systems.

For instance, PCI transfers data at a rate of 133 Mbytes/s. PCI-X transfers data at 1 Gbyte/s. And, InfiniBand transfers data at rates from 1.25 to 4 Gbytes/s. HyperTransport's 12.8-Gbyte/s data-transfer rate is 50 times faster than PCI 64/66 MHz, 12 times faster than PCI-X, and 10 times faster than a four-channel InfiniBand solution. HyperTransport serves as a complementary technology for both InfiniBand and 1G/10G Ethernet solutions, providing an "in-the-box" solution for connecting these high-bandwidth devices within a system.

The flexible HyperTransport I/O bus architecture is a comprehensive solution for embedded systems. Clock rates range from 200 to 800 MHz, with 2 bits transferred per clock period. Standard bus widths of 2, 4, 8, 16, and 32 bits permit customized I/O bus characteristics for each specific application, and asymmetric bus widths support both upstream and downstream bandwidth requirements.

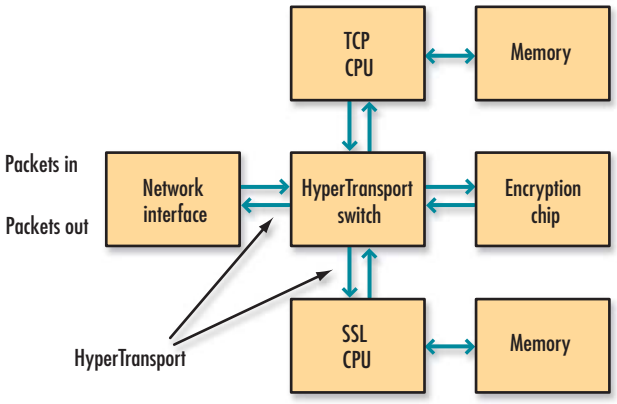
HyperTransport provides a broad selection of bus widths and speeds to meet the power, performance, and cost requirements of current and future embedded systems. It can enhance performance for any application requiring high speed, low latency, and scalability, including networking, telecommunications, computer, and high-performance embedded applications *(Fig. 1)*.

The HyperTransport programming model is compatible with existing models, requiring minimal changes to existing operating-system and driver software. It enables system designers to develop very complex, high-performance, scalable networking topologies through switching technology while maintaining and improving the scalability and performance of their existing legacy PCI infrastructures. Also, it complements PCI and emerging technologies such as the InfiniBand and 3GIO standards.

The PCI bus has been considered the universal socket for years. HyperTransport's backwards compatibility with PCI software lets developers maintain their PCI-compatible driver software while taking advantage of the performance gains enabled by HyperTransport. Able to serve as a mezzanine bus for PCI 66/64 and PCI-X based systems, HyperTransport lets network equipment manufacturers extend the number of ports and the bandwidth of their systems with minimal architecture changes. Its scalable network fabric can handle the needs of switching equipment into the future.

fig. 1

This general view of a network shows where the various modules can communicate via HyperTransport links.



HyperTransport is based on dual point-to-point unidirectional links consisting of datapaths, control signals, and clock signals. Each datapath can be from 2 to 32 bits wide. Commands, addresses, and data share the datapath. A link consists of the datapath, a control signal, and one or more clock signals. A complete HyperTransport-based system consists of a processor with a HyperTransport port, an input link and an output link, and any I/O channels connected to the HyperTransport bus.

The HyperTransport architecture is divided into five layers, with a structure similar to the Open System Interconnection (OSI) reference model. The physical layer includes data, control, and clock lines. The data link layer includes the initialization and configuration sequence, periodic cyclic redundancy check (CRC), disconnect/reconnect sequence, information packets for flow control and error management, and doubleword framing for other packets. The protocol layer features the commands, the virtual channels in which they run, and the ordering rules that govern their flow. The transaction layer uses the elements the protocol layer provides to perform actions, such as reads and writes. And, the session layer includes rules for negotiating power-management state changes, as well as interrupt and system management activities.

The signaling technology used in HyperTransport technology is a form of low-voltage differential signaling (LVDS) that requires fewer pins and wires. Allowing decreased pin counts enables compact profiles at lower costs and simplifies board design, layout routing, and signal-integrity issues. HyperTransport features an enhanced LVDS technique developed to evolve with the performance of future process technologies and ensure that the HyperTransport technology standard has a long lifespan. This design also reduces overall system cost and power requirements, because the transceivers are built into the controller chips.

HyperTransport technology has been licensed to chip vendors, including PMC-Sierra, a supplier of 64-bit MIPS-based microprocessors to the networking and communications markets. PMC-Sierra has integrated the HyperTransport bus into its new RISC processors designed for servers and communications systems. The company's first implementation of HyperTransport is in its next-generation high-performance processor. The 1-GHz RM9000x2 dual-CPU IC is based on the MIPS instruction-set architecture.

The RM9000x2 is a scalable multiprocessing architecture that solves the industry-wide problem of slow data transfers between processors in cache coherent systems. It offers greater throughput efficiencies over single-CPU products. Anchored by dual CPU cores running at 1 GHz, the RM9000x2 achieves maximum performance while drawing only 5 W. The RM9000x2's high-speed I/O connectivity—which includes a 500-MHz, 8-bit

HyperTransport bus interface—supports the demanding performance requirements of high-speed devices, such as edge routers, core routers, and enterprise servers.

In addition to the HyperTransport interface, the RM9000x2's integrated high-speed bus interfaces include double-data-rate (DDR) SDRAM, SysAD, and a local bus, providing low latency accesses into the main memory and high bandwidth to the I/O devices. The RM9000x2 also has dual 256-kbyte L2 caches as well as a cache coherency system to permit Symmetric Multiprocessing applications. The HyperTransport I/O bus delivers 2 Gbytes/s of bus bandwidth, enabling the processor to achieve maximum performance. The HyperTransport interface provides easy connections to a wide range of high-speed networking peripherals *(Fig. 2)*.

More than 50 companies have licensed the HyperTransport specification. Many have already announced products that incorporate this technology. In addition to PMC-Sierra's RM9000x2, HyperTransport technology is being integrated with Altera's APEX II FPGA, AMD's Opteron, Broadcom's BCM1250, NVidia's nForce, Teradyne's Tiger, and Xilinx's Virtex II FPGA.

BUILDING ON INDUSTRY SUPPORT

Because HyperTransport is seen as an essential solution for the future of chip-to-chip communication and high-speed connectivity, industry leaders have joined together to support its continued growth. The HyperTransport Consortium was founded in July 2001 to drive the future development and adoption of the HyperTransport specification as the industry-wide communication and networking bus standard.

Membership is open. The consortium currently has nine promoter members, which constitute the board of directors, and a growing list of contributors and adopters. It anticipates having 70 or more new members on board by the end of the year. The nine promoters are Advanced Micro Devices, Apple Computer, Broadcom, Cisco Systems, NVidia Corp., PMC-Sierra, SGI, Sun Microsystems, and Transmeta Corp. For details, visit *www.hypertransport.org*.

HyperTransport can scale from speeds of 200 to 800 MHz. It delivers more bandwidth than other available alternatives, uses low-latency responses, minimizes pin counts, ensures compatibility with legacy PC buses and transparency to operating systems, extends to new SNA buses, and causes minimal impact on peripheral drivers. These extensive capabilities position it as the most comprehensive I/O bus solution for the next generation of embedded systems.

Rather than disrupt traditional design components, HyperTransport technology complements externally visible bus standards like PCI, as well as emerging technologies like InfiniBand and Gigabit Ethernet, making it a universal solution for in-the-box connectivity. Currently, HyperTransport technology can be used for high-performance networking, telecommunications, computer, and embedded applications, as well as in any application in which high speed, low latency, and scalability are essential.

In addition to boosting performance for integrated microprocessors, a variety of components with native HyperTransport interfaces is currently being developed to change the face of communication systems designs. These components include PCI-X bridges, Gigabit Ethernet controllers, InfiniBand communication adapters, encryption/decryption engines, and packet classification engines. Future applications may involve advanced error management for high-availability systems, enhanced networking features, and even higher-speed physical interfaces.

System design will continue to grow more complex and challenging as component manufacturers strive to enhance performance. In this environment, system designers will depend on the innovation and drive of industry groups such as the HyperTransport Technology Consortium to develop the technology required to complete designs and bring successful products to market. ▢

***Brian Holden** is the principal engineer with PMC-Sierra's MIPS Processor Division in Santa Clara, Calif. He is the technical chair of the HyperTransport Consortium as well as the chair of its Network Task Force. He also has 16 allowed U.S. patents. He can be contacted via e-mail at brian_holden@pmc-sierra.com.*

fig. 2

This detailed view of a netronics system implementation shows the various types of functions where HyperTransport interconnects the function blocks.