

The Future of High-Performance Computing: Direct Low Latency CPU-to-Subsystem Interconnect

Mario Cavalli
General Manager
HyperTransport Consortium

For system and subsystem manufacturers, delivering low-latency high-performance computing solutions at affordable prices has been an insurmountable barrier. Although processor speeds and bandwidth have taken quantum leaps over the last decade, the last few inches between the adapter slot and system CPU create a bottleneck that restricts the development of cost-effective high-performance computing solutions. For complex modeling, high-end transactional systems, commercial data centers and large system clusters, high latency or "wait time" is a stumbling block that is holding back the development of systems with supercomputer-level performance using off-the-shelf components.

Fortunately, a new expansion interconnect brings ultra-low latency and high performance to an expansion slot, enabling direct communication between the system CPU and high performance subsystems. HTX™, an expansion connector specification that leverages industry-standard HyperTransport™ technology, overcomes the latency barrier common with standard systems. This white paper will address some of the challenges related to developing low-cost high-performance computing solutions and introduce the advantages of HTX.

Market Drivers

Latency deficiencies, the limitations of traditional interconnects and the need for instant responses for high-performance transaction systems are driving the need for low-latency CPU-to-subsystem direct interconnects

Every microsecond in improved latency can result in hours, days or weeks of processing improvements for complex and transaction-intensive applications. Over the last few decades, processor speed and network bandwidth have improved dramatically, while latency has scarcely advanced. As computer science professor David Patterson of the University of California at Berkeley explains, "In the time that bandwidth doubles, latency improves by no more than a factor of 1.2 to 1.4."¹ "Had latency kept up with bandwidth, we would have 0.1-nanosecond processor latency today", Patterson added.

¹ Patterson, David. (October 2004). Why Latency Lags Bandwidth and What It Means to Computing.

PCI and PCI-X interconnects have also been pushed to their performance limits as serious I/O bottlenecks were introduced between system processor, memory controller and the system I/O interconnects. With the addition of advanced I/O functions to the south bridge, such as high-speed drives and peripherals, PCI and PCI-X interconnects have become major performance bottlenecks.

As systems continue to advance, the need for faster system interconnects is becoming more critical. For example, 16-core system processors of tomorrow will be best leveraged by low-latency connections between the CPU and compute-intensive subsystem. As the network infrastructure rockets forward, system latency will be less tolerable. High-performance transaction systems, like those processing online airline reservations, are becoming more dependent on low-latency solutions in their quest to deliver rapid feedback to customers. Consumers and businesses alike are coming to enjoy and expect instant transactional results. In these types of applications, high latency can negatively impact end-user instant gratification.

Given the need for instant response, the limitations of existing buses and dramatic bandwidth and processor speed improvements, latency is the final holdout preventing low-cost commodity systems from advancing to high-performance compute level.

System and Card Manufacturer Challenges

Manufacturers of high-performance systems and peripheral devices are held back by the cost of delivering low latency, the restraints of existing CPU-to-peripheral connections and the inherent limitations of standard buses.

The Cost of Low Latency

Off-the-shelf technology and components are the staple of high-performance computing solutions, yet latency is deterred by interconnect limitations

The cost of developing high-performance systems is becoming increasingly prohibitive. The expense of utilizing more silicon and engineering resources for small production volumes are deterrents for next generation system development. Simulation, modeling and high-end transaction processing are the domain of costly, proprietary large scale systems.

As markets are increasingly commoditized, manufacturers seek to leverage off-the-shelf components to rapidly deliver cost-effective solutions to market. This is especially critical for manufacturers that produce low volumes of high-performance systems. The capital costs of developing high-performance proprietary systems from scratch are major market limitations. Now high-performance systems can be built for much less, more quickly and without the need for massive sales volume to regain profits. However, the challenge is to reduce the processing latency of these inexpensive server-based systems so they may successfully and profitably target high-performance applications.

CPU-to-Subsystem Limitations

The data path from system CPU to subsystem slots is slowed as data passes from chip to chip and through complex control logic

Crossing data over conventional peripheral interconnects - such as PCI, PCI-X or PCI Express - to system CPUs involves forcing data through the many paths and stops that the system's chipset logic imposes, akin to a city bus making multiple stops while carrying eager passengers to their destinations. Latency can be a significant issue when data travels from chip to chip and passes through layers of intermediate IC controller functions.

With the increasing speed of processors and networks, this time lag has become a serious limitation for high-performance computing systems. In addition, each chip function that bridges data between the card slot and the CPU impacts the system's cost, its power consumption and its reliability. Since the overall system reliability is inversely proportional to the number of components used, system reliability is much lower when complex, active components such as chipsets and bridge controllers are used, because they increase the statistical chance of failure.

When applications demand low latency, even a minuscule 100 nanosecond latency reduction becomes a very significant improvement. For a 2-gigahertz chip, a 100 nanosecond latency reduction means that 400 more instructions can be processed per second. Optimizing the path from the north bridge to the south bridge and back is one way to reduce in-system latency. However, the ideal solution is a super-highway that drives data traffic directly from the system CPU to the peripheral subsystem and viceversa, dramatically reducing latency. HTX represents such a solution.

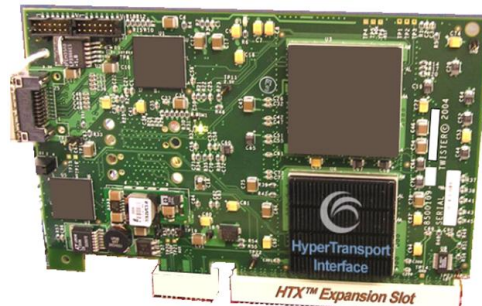
Standard Bus Limitations

Standard interconnect technologies require bridge chips functions that increase latency

Many of today's standard interconnect technologies, such as PCI, PCI-X and PCI Express, require a bridge chip to connect peripheral devices to the system CPU. As stated earlier, adding integrated circuits increases system development costs, increases latency, lowers performance, increases power consumption and lowers reliability.

Although sporting a modern, packet-based interconnect architecture, PCI Express was not intended to be a front side bus directly connected to the CPU. Instead, it was designed to supersede the older-generation PCI and PCI-X buses as a new slot type interconnect. Its legacy support and handling of many possible configuration settings and system verifications burden PCI Express with important latency penalization drawbacks that discourage its use in very latency sensitive applications.

HTX emerges as the most suitable interconnect to enable the highest-speed, lowest-latency communication between CPU and compute-intensive subsystems.



The Solution: HTX – Direct Peripheral-to-CPU Interconnect

HTX enables a high-performance subsystem to directly access the system CPU with the same performance and latency as if its functions were embedded on the motherboard

Designed to enable the development of flexible and powerful high-performance computing systems while leveraging low-cost industry standards, HTX is an expansion interface specification that employs the popular HyperTransport protocol (see "HyperTransport Technology" below). HTX allows high-performance expansion cards to directly plug into HyperTransport-enabled server systems, bypassing the traditional challenges of developing low-latency systems and opening the door for the development of cost-effective and highly powerful transactional computing platforms, application-specific co-processing and other low latency solutions. Now system and peripheral manufacturers can extend the power of HyperTransport to peripheral subsystems, bridging the last few inches that have prevented the development of supercomputing-level performance with commodity components.

Gone is the need for market-specific, costly motherboards. With the HTX interconnect, system manufacturers can design and market a single, powerful system or motherboard platform that can be easily tailored for specific high-performance markets and industry sectors by simply adding an HTX-based peripheral card from a growing portfolio of functionalities and vendors.

Whereas HyperTransport is an in-system chip-to-chip interconnect, HTX enables plug-in subsystems to achieve the same direct-connect performance benefits. Leveraging the same mechanical connector as PCI Express combined with HTX-specific signal allocation, HTX system manufacturers can benefit from the economy of scale and multi-sourcing of industry-standard connectors when adding HTX to their products.

Compared to PCI Express, HTX delivers state-of-the-art latency by eliminating clock recovery circuit logic, adding HyperTransport's Priority Request Interleaving™ and employing a super lean packet payload protocol.

HyperTransport Technology

HyperTransport interconnect technology is an industry-standard high-bandwidth, point-to-point link that provides the lowest possible latency for chip-to-chip and board-to-board communications. HyperTransport technology provides a flexible, scalable interconnect architecture designed to reduce the number of buses within a system. The technology serves a wide array of applications, including embedded systems, networking, personal computers, workstations, servers and high-performance computing (HPC) systems. Major system manufacturers, including, Acer, Cisco, Cray, Dell, Fujitsu, Fujitsu-Siemens, Gateway, Hewlett-Packard, Huawei, IBM, Lenovo, Sun Microsystems and Toshiba are fully invested in HyperTransport technology.

HyperTransport's aggregate bandwidth of 51.2 gigabytes per second represents more than a 170-fold increase in data throughput over legacy PCI buses. While providing far greater bandwidth, HyperTransport technology complements legacy I/O standards like PCI as well as the latest interconnects such as PCI Express. HyperTransport employs a packet-based protocol and clock-forwarding technique that eliminates the need for many control and command signals. HyperTransport supports asymmetric, variable-width data paths and operates with low-voltage differential signaling (LVDS) point-to-point links, delivering increased data throughput, minimized signal crosstalk and lowered electromagnetic interference. In short, HyperTransport combines the best features of parallel and serial interconnects.

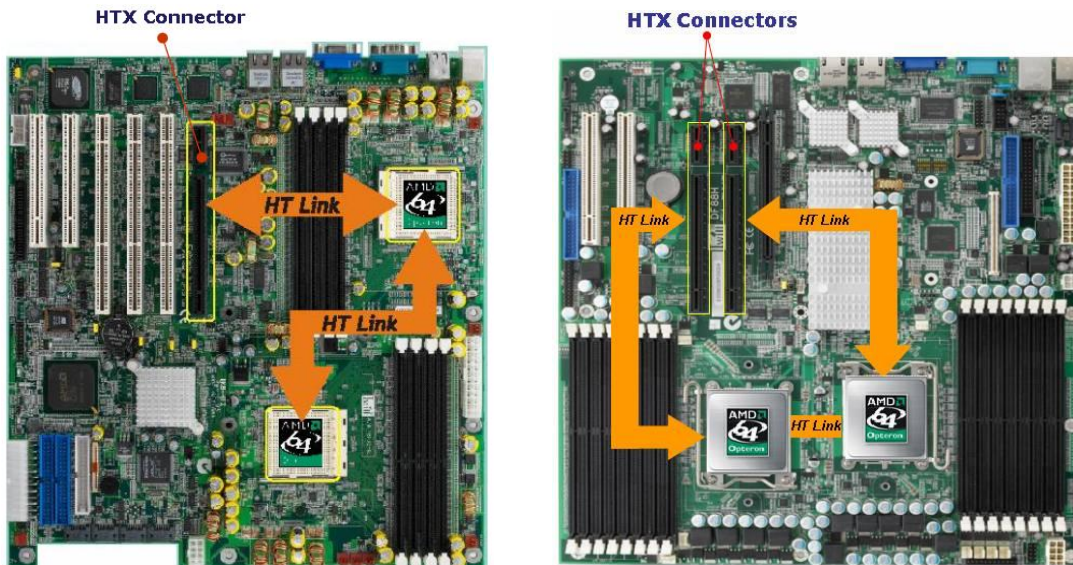


Figure 1: This picture shows possible dual-CPU system designs – one with single HTX connector (left) and the other with two HTX connectors (right). Each HTX connector interfaces directly with the CPU via a dedicated HyperTransport link. The HTX slot is mechanically the same as a PCI Express connector, although HyperTransport signals and pin allocation are completely redesigned. The HTX connector slot is reverse-installed compared to PCI Express slots in order to prevent board misplacements.

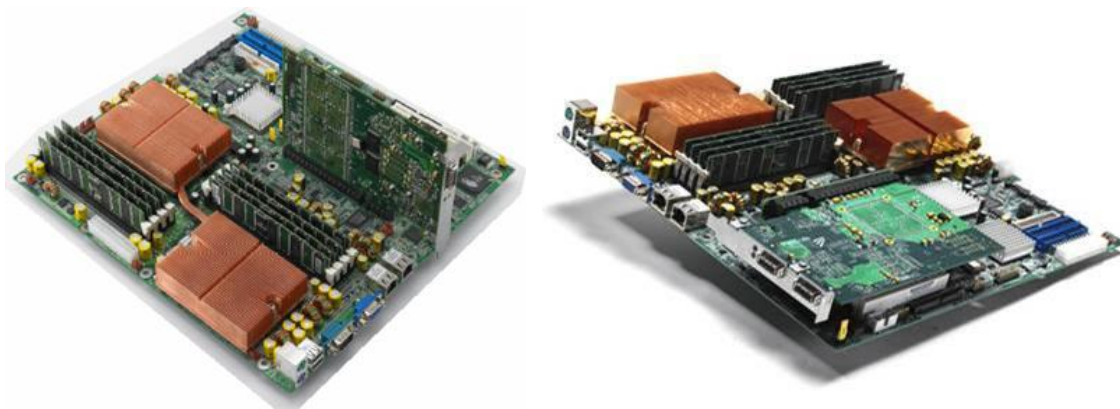


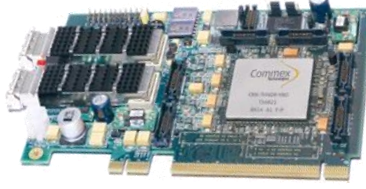
Figure 2: HTX boards can be installed vertically on the system motherboard (left) – configuration suitable for 3U+ system chassis – or horizontally via riser card and parallel to the system motherboard (right) – configuration of choice for 1U system chassis. HTX riser cards are system-dependent and sold by the HTX system manufacturer. They may come configured with HTX-compatible connectors at both ends (male and female), or with proprietary connectors on the motherboard side and standard HTX connector toward the HTX peripheral board.



Data / Algorithm
Accelerator



Content-Aware 10GE
Routing Processor



NETLT
Security Processor



Tarari Grand Prix 9000
Security Processor



10G Ethernet
Reference Design



UNIVERSITY
OF
HEIDELBERG



Production-level design kit offered at no cost
through HyperTransport Consortium

Universal HTX Board Reference Design

UNIVERSITY
OF
HEIDELBERG



Universal HTX Board Reference Design

UNIVERSITY
OF
HEIDELBERG



Universal HTX3 Board Reference Design

Figure 3: Examples of HTX subsystem cards



System x3455



System x3655



System x3755



ProLiant DL 785-G5



ProLiant DL 145-G3



ProLiant DL 165-G5



AS4040C-8R



AS1040C-8



Figure 4: Examples of HTX-enabled systems

HTX Business Benefits for Manufacturers

HTX provides system and subsystem manufacturers numerous business benefits as follows:

HTX offers system and subsystem manufacturers an inexpensive way to quickly deliver low-latency, high-performance solutions

State-of-the-art performance potential: The ultra-low latency available with HTX enables the development of a new class of very high-performance systems and subsystems that can drive transactional systems, complex modeling applications, security processing, storage control, application-specific co-processing and other types of compute-intensive applications to new performance heights without the need for costly investments.

Extremely cost effective: System manufacturers can tap the full performance of HTX at the mere cost of the HTX connector. The HTX standard leverages the economy of scale of widely available, multi-sourced PCI Express-type connectors.

Easy integration: HTX integration does not require the use of any control logic or bridge controllers, effectively extending the range of performance possibilities for conventional systems and motherboards, and making HTX integration easy and cost effective.

Market and platform universality: HTX's low cost enables system manufacturers to implement high-performance connectivity on a greater number of motherboards. Alternatively, it allows manufacturers to design fewer, HTX-equipped universal motherboards that can target high-performance markets while reducing bill of materials (BOM) and stocking requirements.

Increased market share: Due to HTX's extreme cost effectiveness and the inherent universality of any motherboard implementing it, system manufacturers can capture high-performance market segments and maximize their revenue and profit opportunities with minimum investment.

Accelerated technology adoption: By leveraging the HTX connectivity standard, manufacturers can rapidly develop products without the need for custom development. HTX allows manufacturers to quickly and efficiently deliver disruptive innovations to a market with limited investment.

No need for dedicated system designs: HTX eliminates the need for designing ad-hoc motherboards from scratch. Gone are costly product designs, complex engineering, tooling development and prototype debugging.

Multi-core, multi-processor support: HTX fully supports multi-core, multi-processor product architectures.

High-performance, low-cost solutions: By dramatically reducing the system latency of off-the-shelf systems, HTX is key to designing an entirely new class of commoditized HPC platforms as well as cost-effective, fast time-to-market application-specific acceleration solutions.

Product leadership: The HTX standard enables high-performance peripheral manufacturers to deliver product excellence, to secure a competitive edge and to attain market leadership.

Complements general-purpose interconnects: HTX is designed to complement general-purpose interconnects like PCI, PCI-X and PCI Express, while meeting the needs of low latency, compute-intensive applications that could not be met otherwise. For system manufacturers, HTX is a powerful complement to general-purpose interconnects, enabling cost effective, easily configurable universal system platforms targeting multiple market sectors.

Technical Advantages

HTX offers ultra-low latency, high performance, reduced chip count and excellent scalability

HTX provides a number of competitive advantages over alternatives such as PCI Express:

Lowest possible latency: The direct interconnect between the HTX card slot and the system processor, combined with the inherent low-latency features of HyperTransport technology deliver the lowest possible latency and performance of all standard interconnects.

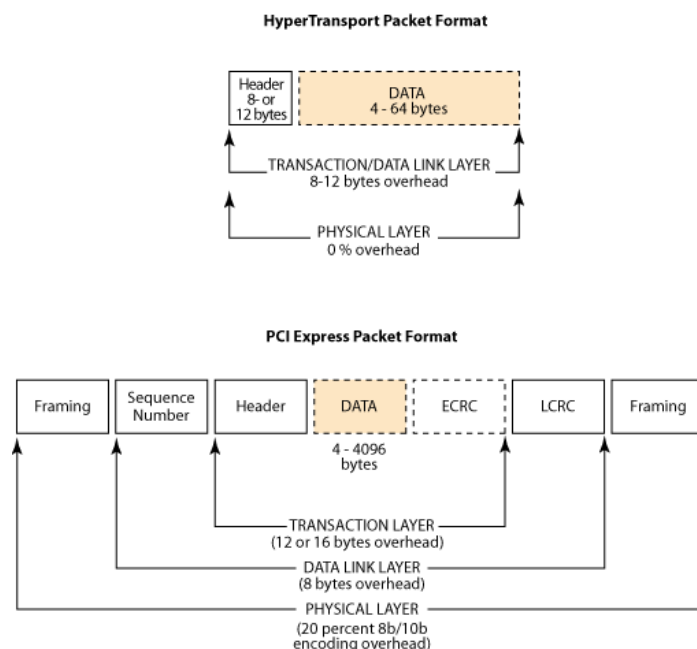


Figure 5: The HyperTransport packet format is lean compared to the much larger PCI Express packet format.

Best of parallel and serial interconnects: HTX and HyperTransport technology employ a packet-based protocol combined with a clock-forwarding technique that eliminates the need for many control and command signals of older-generation buses like PCI and PCI-X. Also, HTX and HyperTransport have no imposed use of clock recovery logic, a major latency burden common with interconnects like PCI Express. With Priority Request Interleaving™ and an extremely lean packet protocol, HTX truly combines the best features of parallel and serial interconnects.

Supports variable width links: For peripheral devices that receive more data than they transmit, such as in graphics applications, HTX allows system manufacturers to implement asymmetric traffic paths - e.g. 16 lanes receiving and 2 lanes sending data - that optimize motherboard real estate and cost.

HTX Application Examples

HTX is ideal for high-performance applications that benefit from low latency

HTX is best leveraged in low-latency, high-performance applications in which latency improvements create exponential processing performance. Co-processing subsystems, security processors, media processors, real-life animation systems, server clusters, multi-gigabit/s networks, as well as transactional modeling and storage systems, are excellent platform candidates for the benefits of HTX. Specific application examples include climate modeling, computational chemistry, medical imaging, 3D rendering and animation, algorithm acceleration, pattern matching, financial modeling and risk assessment, molecular modeling, weapons simulations, security processing, storage management, data encryption/decryption, XML processing and real-time data analysis among others.

HTX Summary

HTX is an industry-standard interconnect that allows direct CPU to card subsystem communications, speeding performance and dramatically reducing latency. For system and peripheral card manufacturers, it offers the ability to deliver leading-edge performance at the lowest achievable cost using industry standard, off-the-shelf components.

To learn more about the HTX interconnect standard, its market positioning, performance advantages, specifications and reference material, visit the Connectors pages in the HyperTransport Technology section of the HyperTransport Consortium web site (www.hypertransport.org).

About HyperTransport and the HyperTransport Consortium

HyperTransport is a widely adopted, royalty-free, industry-standard, processor-to-processor and processor-to-I/O interconnect technology. HyperTransport technology is currently deployed in high-performance personal computers, mobile computers, servers, network and communication equipment, embedded products, HPC systems and supercomputers.

The HyperTransport Technology Consortium is a membership-based non-profit organization in charge of managing and promoting HyperTransport technology. The Consortium was founded in 2001 by leading technology innovators like AMD, Apple, Broadcom, Cisco, NVIDIA, PMC-Sierra and Sun Microsystems and counts more than 60 industry-leading members worldwide, including industry leaders AMD, Broadcom, Cisco, Dell, HP, IBM, Nvidia, Renesas and Sun Microsystems. Consortium membership is based on a yearly fee and it is open to companies interested in licensing the royalty-free use of HyperTransport technology and intellectual property. Consortium members have full access to the HyperTransport technical documents database, they may attend Consortium meetings and events and may benefit from a variety of technical and business promotion services that HTC offers at no cost to its members. For more information on the HyperTransport Consortium and Consortium membership, please click [here](#).

