

RapidIO® Interconnect: Best for Communications

A Critical Analysis of the RapidIO Interconnect vs. Advanced Switching Interconnect

Design engineers face a multitude of choices in interconnect technologies. This paper compares the technical merits of the RapidIO standard and the Advanced Switching Interconnect (ASI) for use in embedded, high availability networking and communications applications.

Background on the RapidIO Interconnect and ASI

The RapidIO interconnect is a relatively mature technology. Its developers started work on the standard in 1997, the RapidIO Trade Association was officially formed in 2000, and the base specification was completed in 2001. The RapidIO interconnect architecture was designed to solve the unique connectivity and reliability challenges of high-performance embedded systems. It offers low latency and extremely high bandwidth, requires very low pin count, and is transparent to software. The RapidIO interconnect also addresses the demand for reliability within embedded designs by offering built-in error recovery mechanisms and a point-to-point architecture that helps eliminate single points of failure. Processor, system-logic, FPGA, and ASIC devices have already realized the technology in silicon, and several companies have shipped it in production volumes at the board and system level. In October 2003, the International Standards Organization (ISO) and the International Electrotechnical Committee (IEC) ratified the RapidIO interconnect specification as ISO/IEC DIS 18372 in a unanimous committee vote, making the RapidIO interconnect the only system interconnect technology so certified.

The ASI standards work was born out of the Arapahoe Work Group effort on PCI Express. It was recognized early on that PCI Express would not adequately meet the needs of communications applications. In fact, some of the backward compatibility constraints imposed by the PCI legacy would make meeting those needs fundamentally incompatible with PCI. What was needed was a new standard that was not encumbered by PCI legacy. At the same time, this effort ignored other standards efforts already in progress in the industry. This work started in the Arapahoe Work Group at approximately the same time as the work on PCI Express. The public has, in some cases, been lead to believe that PCI Express and ASI are the same; however, the technologies are very different. That difference necessitated the formation of a new standard body. ASI is being developed as a generic transport technology. It is designed to encapsulate transactions of its native protocols, PCI Express, and other protocols in the future. Except for PHY reuse, the ASI SIG will require further years of work to define or establish the standards for passing other protocols across ASI.

Figure 1 shows various interconnects positioned in terms where they can be used in a system design based on the functionality they can support. The RapidIO interconnect is the only technology that spans from chip-to-chip local interconnect all the way up through high-end fabrics – offering an open fabric standard. ASI supporters have positioned ASI as a low-end backplane interconnect that is not intended to scale to replace the proprietary fabrics that dominate the communications segment today. In this respect, RapidIO interconnect’s applicability within the communications segment is much broader than ASI’s.

Interconnect Use		Characteristics	
LAN/WAN (billions of end-points)	Ethernet	IPv4/IPv6, 48-bit MAC address	
Fabric (50K+ end-points)	RapidIO	Hundreds of classes, millions of flows, end-to-end flow control, interworking	
Switched Interconnect	ASI	Message passing, architectural/topological independence, one flow, protocol tunneling	
Serial Local Bus (10+ end-points)	PCI Express	Serialized Input/Output Transactions/DMA	
Parallel Local Bus	PCI-X	Parallel Input/Output Transactions/DMA	

Figure 1: Pervasiveness of Interconnects in Embedded Communications

Historically, systems tended to separate control and data planes on separate interconnect fabrics. However, the cost of maintaining two or more infrastructures was high from a hardware perspective. In general, the bandwidth requirements of the data plane are orders of magnitude higher than the control plane. It is more economical to provide for control plane on the same interconnect as the data plane, provided the grade of service for the control plane traffic can be protected. (that is, data plane traffic should never preempt control plane traffic). The RapidIO interconnect is the only interconnect to provide the requisite constructs to combine and yet separate these distinct traffic flows.

Addressing and Scalability

Both RapidIO interconnect and ASI provide a large system address space. Both are capable of addressing systems with hundreds to thousands of end-points. ASI provides a more general addressing solution than PCI Express, which has a single unified 64-bit address space and supports a maximum of 32 devices each for 256 buses. ASI and RapidIO interconnect approach packet transport in a very different manner. RapidIO interconnect supports unicast and multicast traffic using table-based routing, which simplifies end-points and requires switches to perform a look-up to map packet destination IDs to ports.

ASI switches must support a table-based routing mechanism for multicast traffic in addition to a path-based routing scheme for unicast traffic. Path-based routing can be problematic in a number of ways:

- Path-based routing requires end-points to retain a real-time database of the entire topology map. In large fabrics, this could be 10,000's of nodes. The need to reprogram the switch topology at all end-points using two methods in the presence of faults will be significantly more complex in ASI. This may make the implementation of fault tolerant systems impractical. In addition, in memory mapped systems, forcing software to understand the topology of the system in order to generate the address of a device is problematic and awkward.
- Path-based routing requires more bits of information to be transported with the packet and thus higher overhead. It is also much more cumbersome to create these extra bits in the source nodes. An advantage claimed for path-based routing is that a packet can switch paths to a destination without having to reprogram routing tables in switches. However, one can accomplish the same effect with table-based routing by giving the destination more than one target ID.

- Path-based routing requires a pointer change while the packet is in flight, which means that the Cyclic Redundancy Check (CRC) must be regenerated because the packet has changed. This has inherent reliability issues because the packet must carry two CRCs – one that is variant and one that is invariant.
- Finally, path-based routing is problematic to system debug in that the routing address to an endpoint will be different based on the source. This makes it more difficult to figure out where a packet is headed when sitting mid fabric.

With its simple elegance, RapidIO interconnect scales to much higher-end fabrics than does ASI yet this flexibility incurs no penalty in added cost or power consumption. An ASI block is expected to be about the same size as a PCI Express block, both of which are significantly larger than a RapidIO block.

Protocol Encapsulation and Tunneling

The Advanced Switching community has based its architecture on an encapsulation and tunneling mechanism. At this time, the encapsulation protocol has not been defined for ASI. Encapsulation alone is of limited use unless it is accompanied by protocol interworking. The ASI specification does define interworking and transaction mapping between PCI Express and the AS native Protocol Interfaces (PI) such as Simple Load Store (SLS) and Simple Queuing (SQ). This mapping requires bridging hardware to be implemented within the switch. This will lead to ASI products that are bilingual with PCI Express being more complex than RapidIO products. Some application diagrams suggest that this bridging hardware is implemented in both switches and end-points. Figure 2 shows the format of a basic ASI packet.

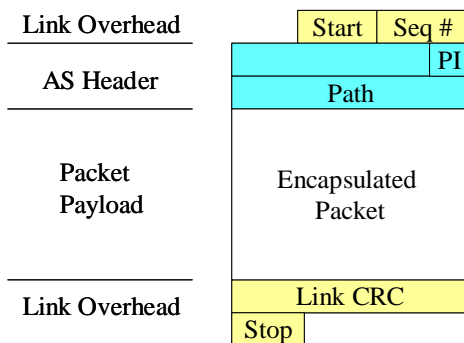


Figure 2: ASI Packet Format

The PI field defines which Protocol Interface is to be employed. PI #8 is defined for PCI Express. It is the presence of this bridge that allows PCI Express to operate in an ASI environment. However, the effort to design a bridge to ASI is really no simpler than the effort to design a bridge between PCI Express and, say, Infiniband, or PCI Express and RapidIO interconnect. PCI Express is not an ASI native protocol and there is little to support the assertion that the migration from PCI Express to ASI is somehow more straightforward than bridging it to other technologies.

RapidIO interconnect provides a simple method for bridging PCI transactions. It is defined in the interworking specification. The set of PCI transactions that are routinely used in embedded systems map very cleanly to basic RapidIO transactions. Products that bridge from RapidIO interconnect to PCI are already available in the marketplace. The basic load store and messaging transactions defined in ASI are similar to the RapidIO memory mapped I/O transactions, and the RapidIO messaging transactions. Advanced Switching must provide several addressing mechanisms, the ASI way and the PCI Express method. However SLS and PCI Express are not identical protocols. Have all of the issues here really been solved?

Of greater importance in the communications world is the mapping to protocols other than PCI. Protocols such as Ethernet, ATM, HDLC, Tcap, H.100, Q.2934 are all important. None of these interworkings or encapsulations are defined in ASI. RapidIO interconnect does a much better job of defining a framework for interworking with these communications protocols.

Data Integrity

RapidIO interconnect has support for Atomic transactions as part of the memory mapped IO logical layer. In addition, RapidIO interconnect has full support for hardware based cache and memory coherency using the Global Shared Memory logical layer.

A glaring weakness of ASI is the lack of support for barrier transactions, and memory coherency. Systems that do not provide a hardware solution to memory coherency should at least provide a software solution. Barrier transactions are required to implement software solutions. Without this, mutual exclusion between software processes cannot be guaranteed. Data coherency in such a system cannot be guaranteed. This limits ASI's usefulness in supporting distributed compute fabrics.

ASI relies upon the PCI Express physical and link layers. The PCI Express link layer uses implied acknowledgement for packet transmission. When an acknowledgement Data Link Layer Packet (DLLP) is sent, every packet that preceded it is assumed to be acknowledged. In contrast, RapidIO interconnect acknowledges every packet and ensures hardware enforced packet delivery at the physical layer. Designers who build high reliability products will clearly understand the value of explicit packet acknowledgement. One of the keys to fault tolerance is to not assume anything about failure modes. Implicit protocols are inherently dangerous in fault tolerant systems.

Economic Considerations

One assumed benefit of PCI Express and ASI is the economies of scale due to the reuse of the physical and link layer technologies. While there is value in reusing existing technology, it is not clear that these economies will be achieved. ASI devices will have much higher prices than volume PC products like motherboard chipsets. ASI devices, by their nature will sell in much lower volume and will have much higher performance, reliability and support expectations imposed upon them. Potential users should be clear that ASI and PCI Express are different technologies that will be deployed on different classes of devices. There are no clear device reuse scenarios between volume desktop computers and embedded equipment. In general, the vendors servicing these markets are quite distinct from one another.

The RapidIO technology has leveraged existing IEEE standard physical layer technologies for its physical and data encoding layers. The RapidIO technology is able to benefit from the economies of scale and targeted engineering focus of this technology for communications backplane applications.

Performance

The other major differences between RapidIO interconnect and ASI are performance related. The RapidIO header is significantly smaller in size than the ASI header. Many existing applications have small transaction sizes. A small packet header means that efficiency is increased. The telecommunications service providers are extremely vigilant about managing efficiency in the network because every bit in the data path that is wasted on header or overhead is essentially a lost revenue opportunity.

The ASI physical layer uses a new 2.0 Gbps physical layer technology, which was developed for PCI Express. This PHY offers lower performance and is of higher complexity to implement than the more mature 2.5 Gbps XAUI PHY that is used in Ethernet and RapidIO interconnect. The ASI PHY must implement several of the PCI Express power management features. These features define multiple levels

of power down based on link activity. These many power levels are appropriate for notebook computer applications, but are inappropriate for embedded systems for which a link tends to be either on or off.

Figure 3 shows the performance of RapidIO interconnect and ASI as a function of packet size. All other things being equal, a RapidIO link will be 20-50% higher performance than an equivalent ASI link. Conversely, if a system's performance requirements targets are fixed, then an ASI link will require twice the number of lanes as an equivalent RapidIO link. This translates into double the cost, double the power dissipation and more routing complexity.

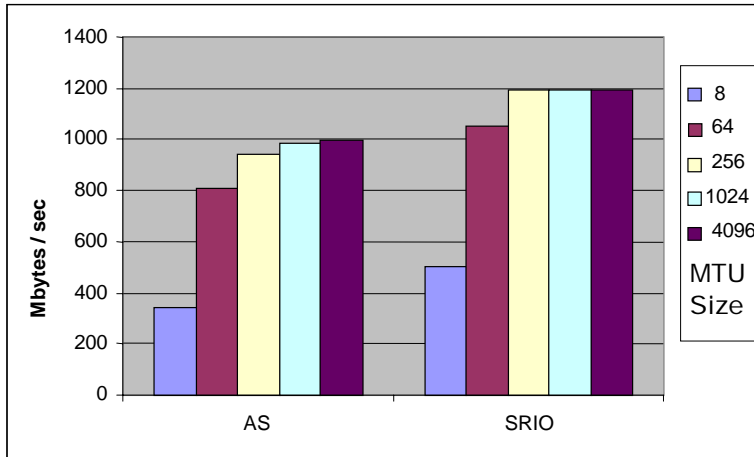


Figure 3: Performance of RapidIO Interconnect and ASI as a Function of Packet Size

Data Streaming

Interconnects such as PCI have been used in control plane applications. It is no longer acceptable to treat all data as equal; it is more important, for example, to process streaming data first because late real-time data is as useless as no data. Data needs to be "tagged" so that an I/O system can prioritize its flow throughout the communications fabric.

ASI only supports 24 traffic classes on 20 VCs. In contrast, RapidIO interconnect supports 256 traffic classes. The RapidIO solution for data streaming is more complete and comprehensive than the hooks provided in the ASI specification. Clearly RapidIO interconnect is far more scalable than ASI for data streaming applications.

Conclusion

RapidIO interconnect with its initial release in 2001 is years ahead of ASI in terms of maturity. The RapidIO standard, unlike other standards, has incorporated essential features and performance attributes that embedded system designers look for when building high availability systems. The leading vendors of microprocessors, communications processors and digital signal processors to the embedded systems market support RapidIO interconnect. RapidIO interconnect also benefits from strong, continuing participation and guidance for leading embedded systems vendors.

RapidIO interconnect is a general-purpose streaming fabric that can effectively span chip-to-chip, board-to-board and shelf-to-shelf. It is not a one-size-fits-all technology. Specific features to support a breadth of applications have been embedded in the definition of the standard and RapidIO interconnect may be implemented in a very small footprint. This ensures world-leading performance for microprocessors, DSPs, fabrics, communications processors and I/O end-points. The critical mass of investment in each of these applications propels RapidIO interconnect to the forefront as the technology of choice for embedded and communications applications.

