

Advanced Switching: Traffic converging ahead

By *Tim Miller*

The recent announcements from the Advanced Switching Interconnect Special Interest Group (ASI-SIG) on the release of v1.0 of the Advanced Switching specification and the expected availability of prototype silicon in 2005 signal a major change in the computing and communications interconnect world. The announcements follow more than two years of specification development effort, initially within the Arapahoe Working Group and completed by the ASI-SIG.

With technical contributions and support from a broad base of interested parties from the semiconductor, equipment, and software industries, the ASI-SIG has defined the standard that, in conjunction with PCI Express, will be the successor to PCI and will provide the interconnect technology for inter-chip to inter-chassis communication for the next generation of converged communications and computing systems.

PCI Express and Advanced Switching address the limitations of PCI as a system interconnect. Wider and faster incarnations of the PCI parallel bus standards did little to address PCI's most serious inadequacies. The shared-parallel bus's single point of failure was an obstacle to increasing reliability and providing higher levels of system availability. This, combined with a limited system scale, shared bandwidth, and strictly hierarchical data routing, affected system design, performance, reliability, fault tolerance, and scalability. Advanced Switching leverages PCI Express to provide switch fabric functionality for multiprocessor and peer-to-peer communications and compute applications.

This major effort from the key industry stakeholders is beginning to make an impact, as comparing PCI bandwidth from first generation through PCI Express shows (Figure 1).

PCI Express

The PCI Express Base specification v1.0 was released to members of the PCI-SIG in July 2002. It defines a serial bus structure for chip-to-chip and add-in card applications, where each lane consists of a differential receive and transmit pair running at 2.5 Gbps. Lanes can bundle up to 32x to increase bandwidth, and the clock is encoded in the data using 8b/10b encoding, which eliminates clock skew issues. On the software front, PCI Express is fully compatible with PCI at the application level. PCI Express components will be available in sample quantities in mid 2004, and sys-

tems will be available in the first half of 2005. PCI Express, with its point-to-point architecture, will offer many advantages over the various PCI parallel buses in bandwidth, performance, and scalability. However, PCI Express developers did not intend for PCI Express to address all of PCI's inadequacies. A follow-on standard, Advanced Switching (AS), would be needed to fully realize the benefits of a next generation interconnect.

Advanced Switching

As they developed the PCI Express Base specification, the architects determined

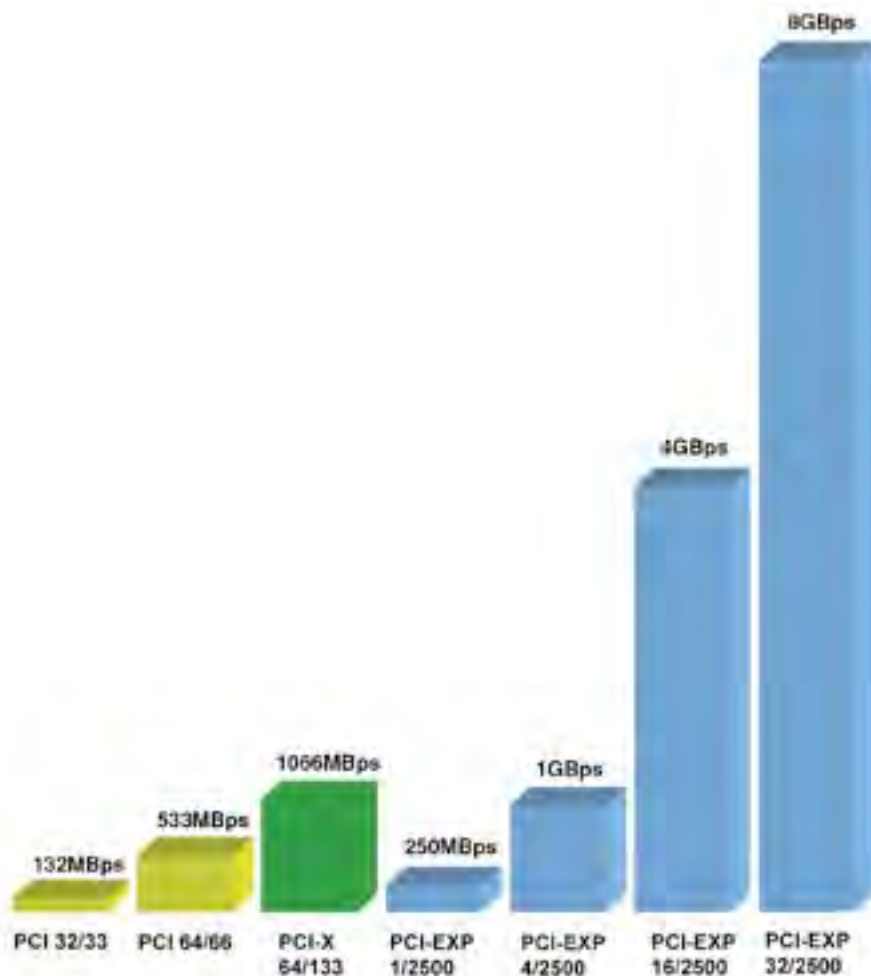


Figure 1

that a class of applications would be best served by a superset of the PCI Express features. Serial bus architectures open the door to features that enable the highest levels of reliability through redundant topologies and to distributed processing through peer-to-peer transactions. However, not all levels of systems necessarily require these features. The ASI-SIG was organized to formalize a specification that would build this functionality on top of the PCI Express Base Physical and Data Link layers.

The founding Board of Directors, which included Agere, Alcatel, Huawei, Intel, Siemens, Vitesse, and Xilinx, was complemented by the involvement of significant players in the communications and compute markets as well as by additional semiconductor vendors with expertise in developing advanced serial interconnects. More than 40 companies are involved in developing the architecture under the auspices of the ASI-SIG. The broad support and participation from these market leaders in the systems and silicon business led to a specification that will result in components that meet the most demanding performance, feature, and cost goals. Advanced Switching is poised to become the dominant multi-point, peer-to-peer interconnect in its performance range.

It will displace proprietary solutions with a family of modular building blocks with wide availability and support. The release of the specification enables designers of silicon, software, test equipment, and design tools to move ahead with product implementation. Figure 2 shows the PCI Express and Advanced Switching layers.

Advanced Switching protocol interfaces

Advanced Switching provides a high level of flexibility for system architects, allowing a number of different I/O protocols to share the fabric. A header attached to the data packet identifies the protocol. In addition, the data payload can contain either a native AS packet or encapsulate a packet in its native format (such as Ethernet, SONET, TCP/IP, or PCI Express).

In February 2004, the ASI-SIG announced four data transport Protocol Interfaces (PIs) that will serve in a wide variety of applications. Additional PIs will be developed by the SIG and made available as industry standards over time, ensuring

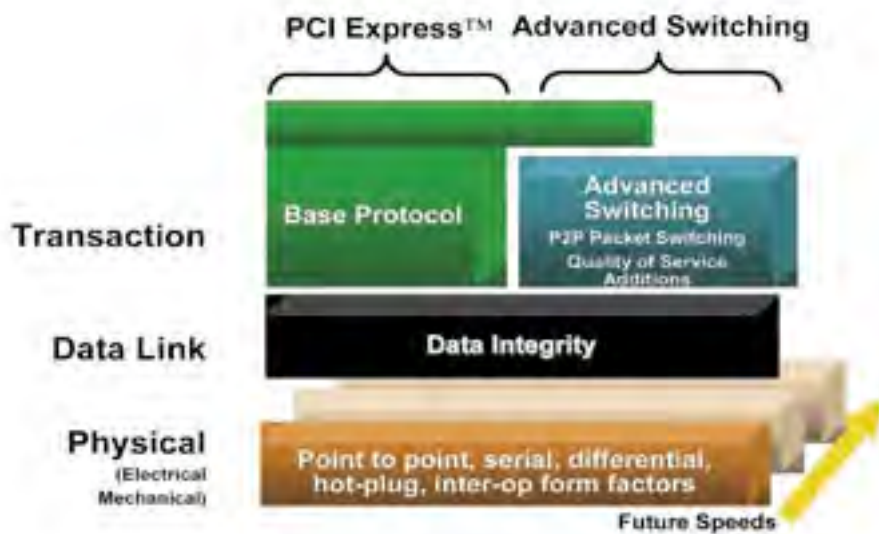


Figure 2

against obsolescence and providing for use in unforeseen applications. Finally, vendor-specific PIs are allowed in the specification for proprietary systems running on industry standard AS hardware. The announced PIs are:

- **PCI Express encapsulation:** This is the standard tunneling scheme for passing native PCI Express packets through the AS fabric, offering the simplicity of complete software compatibility with PCI Express peripherals within an AS environment. A system can contain a mix of PCI Express and AS components to offer the best features of both technologies.
- **Simple Load/Store (SLS):** An extension of the PCI load/store model that offers a low-overhead model for transporting data across the fabric. SLS provides a simple load/store abstraction that allows PCI, PCI-X, PCI Express, HyperTransport, RapidIO, and virtually any other interconnect that uses a load/store model to interoperate within an AS fabric via translation of their native protocol into the common SLS protocol. SLS is a trusted communication model that delivers the advantages of efficiency, low overhead, and low latency.
- **Simple Queuing (SQ):** A simple messaging protocol that uses queues in place of specific addresses to move messages across an AS fabric. SQ allows multiple endpoints to share a single queue resource, thus minimizing the context required for concurrent communication.
- **Socket Data Transport (SDT):** A low-overhead protocol that enables direct hardware implementation of the well-known socket inter-process-

or communication interface's read/readv/readn and write/writev/written data movement model. SDT will move massive amounts of data with minimal processor overhead.

PCI Express and Advanced Switching road map


PCI Express and Advanced Switching are complementary technologies that share physical and data link layers. Their mutual success is tied to the development of the industry standard switch fabric infrastructure. AS is dependent on PCI Express to create the commodity marketplace for silicon, peripheral cards, cables, connectors, and tools, while AS completes the fabric solution with advanced features that enable converged data communication solutions. PCI Express components will show up in prototype quantities in mid 2004 about a year ahead of Advanced Switching silicon. Initial components will be bridges to the various flavors of PCI and switches with assorted port counts and lane widths. CPUs, chipsets, and NPUs with PCI Express ports will show up in late Q2 2004. Systems that use PCI Express instead of a parallel PCI bus to connect to peripherals may ship as early as Q4 2004.

The first half of 2005 will see the introduction of a number of AS chips. Initial AS products will include switches, bridges to PCI and PCI Express, as well as CPU chipsets and I/O processors. Additional announcements will continue throughout the year, and AS based systems will be in production in 2006.

Advanced Switching adoption

Target applications for AS lie in the general areas of communications servers, advanced storage, and blade servers.

These applications have not been well served by industry standard interconnects, relying instead on proprietary solutions for the combination of high availability, distributed processing, QoS features, and multi-Gbps performance. The economic advantages inherent in participation in an industry standard market and the ability to utilize PCI Express peripherals will drive rapid AS penetration in these markets. The development of software and hardware tools to simplify design and manufacturing will further lower the cost of implementing switch fabric technology.

Fundamental shifts in standard interconnect technology occur only rarely. The last such shift saw the emergence of PCI as a widely deployed technology starting in the early 1990s. The market is in the midst of another shift. This one holds the promise for the first time of unifying the compute and communications markets on a single standard interconnect. AS, in conjunction with PCI Express, is emerging as this next standard bearer, and the ASI-SIG is chartered to help the industry in this transition. 

***Tim Miller** is Marketing Working Group Chair, ASI-SIG. He co-founded StarGen in 1999 and served as vice-president of marketing until taking on the CEO position in 2004. He has more than 20 years of experience as a computer system and semiconductor business manager. Tim served in the U.S. Navy as an officer in the Civil Engineer Corps, and has an engineering undergraduate degree from Cornell University, an MS in Computer Science from University of Pennsylvania, and an MBA from the Wharton School. Additionally, Tim serves as president of the StarFabric Trade Association.*

For further information, contact Tim at:

ASI-SIG
5440 SW Westgate Drive
Suite 217
Portland, OR 97221
E-mail: miller@stargen.com
Website: www.asi-sig.org