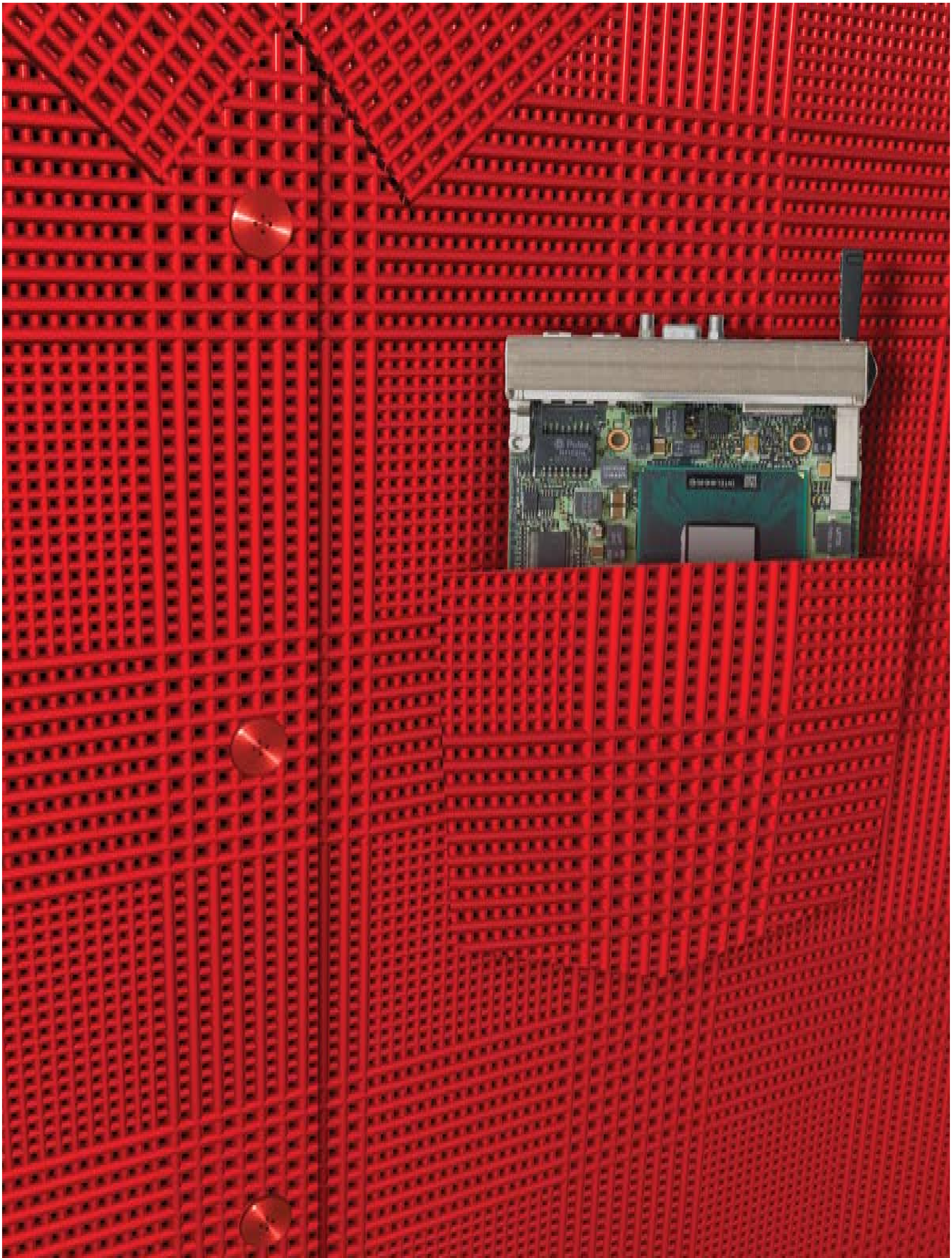
BY WARREN WEBB • TECHNICAL EDITOR

EMBEDDED TECHNOLOGY

DESIGNERS CHOOSE FROM THE LATEST FABRICS

As high-performance embedded systems stretch the limits of technology, board-level interconnection strategies require constant updating to match soaring data-transfer requirements. Designers can choose from a series of evolving and sometimes competing fabric-technology standards to extend system performance and preserve budget constraints. Medical instrumentation, military systems, communications installations, and process automation are just a few areas in which rising bandwidth, increased processing requirements, and escalating application complexity continue to expand the fabric-technology envelope.

EMBEDDED-SYSTEM DESIGNERS ARE USING THE LATEST SWITCHED-FABRIC TECHNOLOGY TO NOT ONLY BOOST DATA RATES, BUT ALSO DYNAMICALLY OPTIMIZE PERFORMANCE, BYPASS FAILED SUBSYSTEMS, AND COEXIST WITH LEGACY COMPONENTS.



New embedded designs traditionally rely on COTS (commercial off-the-shelf) processor and peripheral modules from the embedded-computing marketplace to reduce costs and shorten development schedules. Board standards eliminate the trial-and-error design iterations necessary to get optimal cooling performance and mechanical alignment. Standards-based designs also shorten the software-development effort by providing access to compatible operating systems, vendor-supplied drivers, and sample firmware. Most of the widely used embedded-board standards, such as PCI (peripheral-component interconnect), VME (Versa-module Eurocard), CompactPCI, and ATCA (Advanced Telecom Computing Architecture) already have several fabric-interconnection options so that designers can choose to bypass the shared-bus architecture where needed. However, as high-performance-embedded-system requirements stretch the limits of technology, standards bodies and COTS manufacturers are struggling to deliver a steady stream of updated products to meet the needs of a growing list of advanced applications.

Switched-fabric architectures allow datapaths between computing nodes to change dynamically to support multiple simultaneous data transfers. Designers use the term "fabric" to represent this architecture, because you can connect any node to any other node through datapaths that resemble the interwoven threads in cloth material. A major benefit of a switched fabric is that each connection is a direct point-to-point data path. This feature yields better electrical characteristics, allowing higher fre-

AT A GLANCE

- With high-speed fabric technology extensions, off-the-shelf shared-bus architectures continue to support today's high-performance embedded systems.
- Low-voltage swings and point-to-point transmission paths allow serial switched-fabric systems to extend data rates into the multigigabit range.
- Switched-fabric architectures can adapt to system failures by routing data around defective paths or processing modules while repairs are made.
- Multiple serial-fabric communications options in board-level specifications may lead to industry fragmentation and product-interoperability issues.

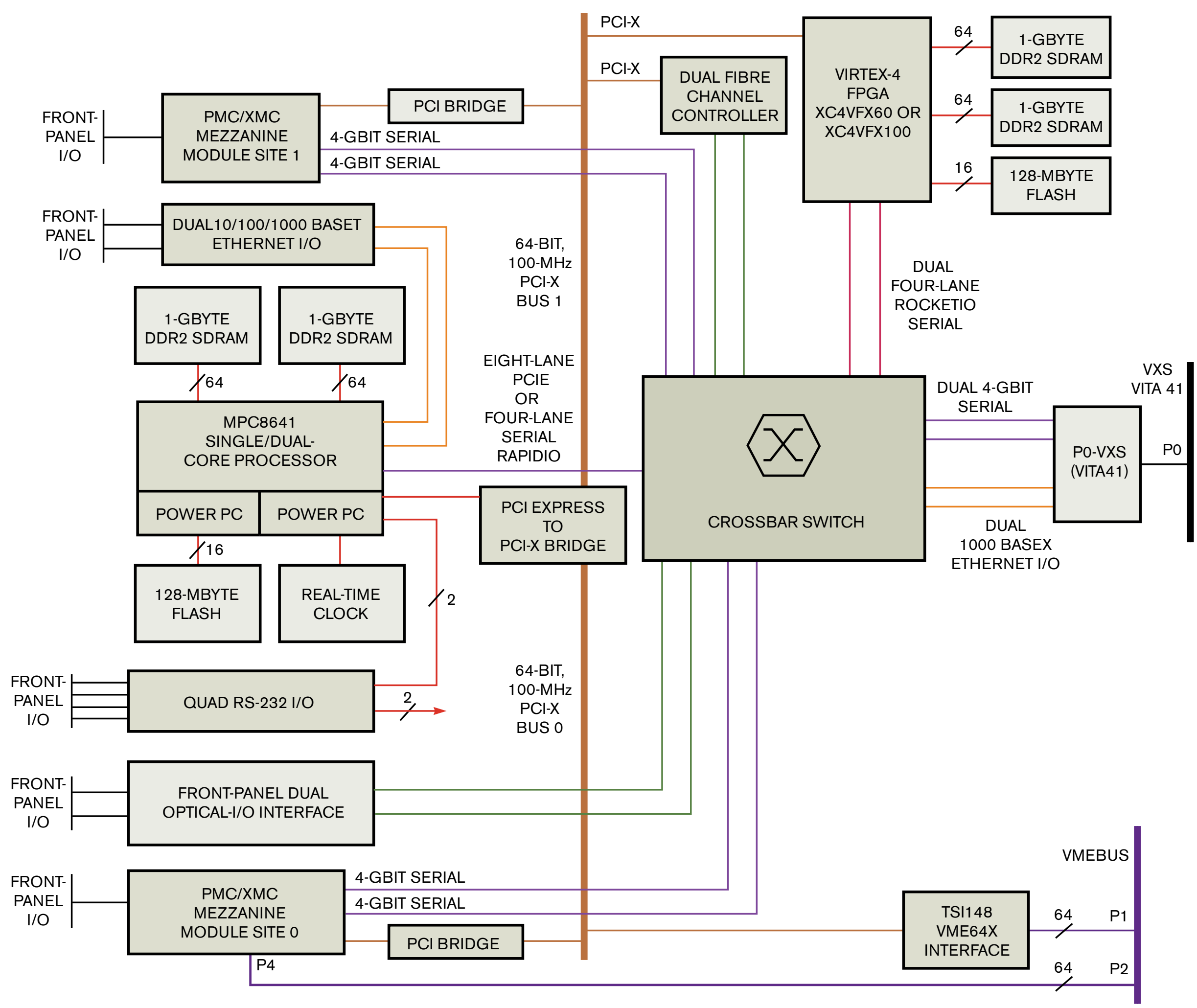


Figure 1 The Pentek Model 4207 VME/VXS digital-signal-processing and data-acquisition board features a fabric-transparent crossbar switch and multiple gigabit serial resources.

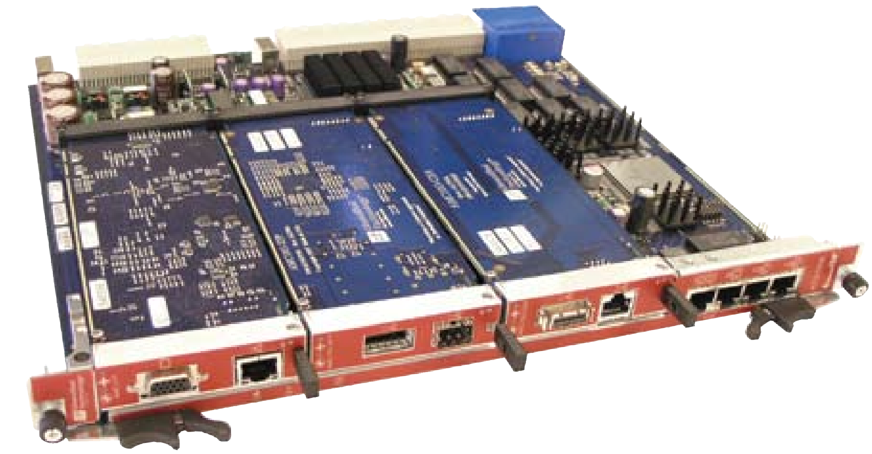


Figure 2 The ATS1936 ATCA Switch Blade features separate base and fabric gigabit-GbE networks plus three onboard AdvancedMC module sites.

quencies and bandwidth than bus architectures. A typical switching fabric uses multiple stages of switches to route transactions between a source and a target.

DIFFERENTIAL SPEED

Most of the switched-fabric specifications call for LVDS (low-voltage differential signaling) for maximum bandwidth between nodes. LVDS uses voltage swings of approximately 350 mV to communicate over PCB (printed-circuit-board) traces or a balanced transmission cable at thousands of megabits per second. With this much bandwidth, most systems use few parallel lines, opting instead for serial-data streams. This approach leaves plenty of room for future performance enhancements by simply adding parallel datapaths. Low-voltage swings and constant-current-line drivers deliver low-noise signals at low power consumption. The ANSI (American National Standards Institute)/TIA (Telecommunications Industrial Association)/EIA (Electrical Industries Association)-644, and IEEE 1596.3 specifications detail LVDS.

In addition to higher data rates, switched fabrics offer designers several system advantages. A major benefit is that each connection is a direct point-to-point datapath, thereby eliminating the multiple connections of a parallel-bus structure. Another obvious benefit of serial connections is the reduced connector sizes due to fewer signal lines. An integrated switched-fabric interconnection can also dynamically increase system availability by routing critical data around defective paths or nodes. Multistage switching allows the designer to easily scale interconnections as requirements change. You can install new switch elements along with new I/O components so that the system retains its interconnection flexibility as the system grows.

Product compatibility is one of the most complicated issues when upgrading or extending the performance level of board standards. VME and PCI have undergone multiple upgrades to increase the shared-bus data-transfer rate and allow legacy products to communicate at their original speed. Switched-fabric upgrades send high-speed data across the backplane through edge-card connections that are unused in the shared-bus configuration. Long-term availability is a prime requirement for many high-performance embedded products. Although the average life of desktop components is about 18 months, users expect typical embedded products to remain in service for five years or more. Some military projects ask for a 15-year life cycle.

PCIe (PCI Express) is one of the more popular fabric extensions and sees wide use in commercial desktop computers in addition to embedded designs. The basic PCIe link consists of two signal paths that use LVDS swings and constant-current-line drivers to communicate at 2.5 GT (gigatransfers)/sec in each direction. Standards bodies have approved

INFINIBAND HAS BECOME THE FABRIC OF CHOICE FOR MANY HIGH-PERFORMANCE COMPUTING APPLICATIONS.

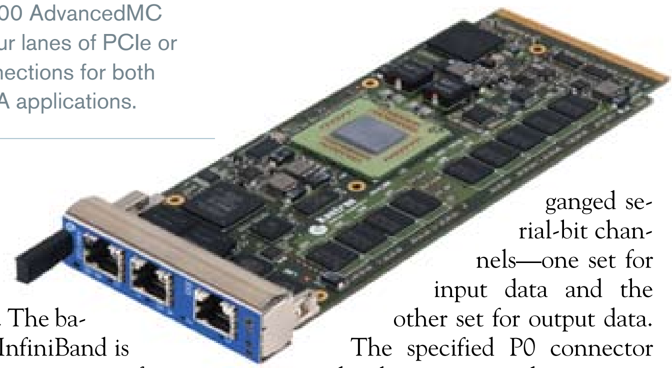
Version 2.0 of the PCIe to increase this data rate to 5 GT/sec. However, because of the 8B/10B encoding scheme used to increase the number of transitions, the maximum effective data rate is 4 Gbps. You can easily increase the bandwidth of a PCIe link by simply adding signal pairs, or lanes, until you reach the desired performance level. The PCIe specification supports one-, two-, four-, eight-, 16-, and 32-lane widths. In August 2007, the PCI-SIG (special-interest-group) announced that the next version of PCIe will support a bit rate of 8 GT/sec.

RAPID RESULTS

RapidIO, which the RapidIO Trade Association defines, is also a popular point-to-point interconnect technology. Originally conceived by Motorola and Mercury Computer Systems, the RapidIO packet-switched architecture transmits data and control information between computing and peripheral nodes in embedded systems. Full-duplex point-to-point links have one or four high-speed serial lanes and 8B/10B-encoded data transmission at signaling rates of 1.25, 2.50, or 3.125 Gbaud for peak bandwidth of 20 Gbps. The association based the initial RapidIO specifications on bit-parallel clock and data, but subsequent specifications have adopted serialized-clock and -data transmission to reduce pin requirements and extend signal reach.

InfiniBand, another popular high-speed interconnect, has become the fabric of choice for many high-performance computing applications. Like the channel model that most mainframe computers use, all InfiniBand transmissions begin or end at a channel adapter. Each processor contains a host channel adapter, and each peripheral has a target channel adapter. These adapters can also exchange information for security or quality of service. InfiniBand uses the 128-bit Internet Protocol Version 6 to uniquely identify each node and provide Internet compatibility. InfiniBand transmissions are either packet- or connection-based to support blocks or continu-

Figure 3 The AM4100 AdvancedMC module supports four lanes of PCIe or RapidIO fabric connections for both ATCA and MicroTCA applications.



ous streams of data. The basic technology for InfiniBand is a bidirectional link consisting of two LVDS pairs, providing one transmitting and one receiving path, each operating at 2.5 Gbps. A usable bandwidth of 80% of the signal rate yields a 2-Gbps or a 250-Mbps data rate in each direction. Higher bandwidth connections are possible by grouping four or 12 links for each transmission path. With 12 links, the effective throughput is 48 Gbps.

PCIe, RapidIO, and InfiniBand are among a number of similar fabric architectures now integrated into and spreading throughout standards-based, high-performance-embedded-system products. For example, CompactPCI, the rugged-form-factor equivalent of low-cost PCI-based desktop hardware, now defines serial-fabric interconnections with a series of optional addenda to the original specification. CompactPCI boards are based on the Eurocard industry standard defining both 3U and 6U board sizes and allow for front-loading and removal from a card cage. The more popular 6U version has as many as five connectors on the rear of the card, allocating two for the traditional CompactPCI bus; the remaining three connectors provide additional pins suitable for optional connections to backplane fabrics. CompactPCI and CompactPCI Express are open specifications that the PICMG (PCI Industrial Computer Manufacturers Group) supports and controls.

Similarly, The VITA (VMEbus International Trade Association) 41 VXS (switched serial extensions) append fabric technology to the venerable VMEbus standard and preserve product compatibility. The VXS specification defines a payload card, a switch card, and a new high-bandwidth P0-backplane connector and retains the standard P1 and P2 parallel VMEbus connectors. Each P0 fabric port consists of two sets of four ganged serial-bit channels—one set for input data and the other set for output data. The specified P0 connector technology supports data rates as fast as 10 Gbps for each serial channel. Payload cards are simply standard VMEbus processor, memory, or I/O boards with the addition of the new VXS-fabric interface. With no P1 and P2 connectors, switch cards have the same form factor as payload cards and include as many as 18 full-duplex serial connectors plus a power connector. The switch card contains the fabric switching necessary to route serial data between payload cards. To remain fabric-agnostic, VITA 41 subspecifications define switch and payload-card definitions for InfiniBand, serial RapidIO, GbE, and PCIe.

ONE SIZE FITS ALL

Pentek recently released its Model 4207 digital-signal-processing and data-acquisition system, featuring multiple high-speed serial interfaces in a VXS form factor (**Figure 1**). This new board employs Freescale's MPC8641D dual-core PowerPC AltiVec processor and a Xilinx Virtex-4 FX series FPGA. In addition, a fabric-transparent crossbar switch bridges multiple gigabit-serial resources, including the PowerPC and FPGA, two XMCs (express mezzanine cards), dual VXS ports, dual Fibre Channel ports, and two optical serial transceivers. The Model 4207 combines many standard interfaces and protocols including VXS, PMC (PCI mezzanine card), XMC, PCI-X (PCI extended), PCIe, GbE, RocketIO, RapidIO, Fibre Channel, Xilinx Aurora, and VME64x technologies, all of which the crossbar switch accommodates. Developers can take advantage of unused FPGA resources to implement custom signal-processing algorithms and to process captured data in real time. Prices for the immediately available Pentek Model 4207 start at $14,725.

With a larger form factor, high-availability features, and high-speed fabric interconnections, the ATCA provides a viable off-the-shelf alternative to the proprietary equipment common in the telecom industry. The ATCA specification provides hot-swap capability for all boards and active modules, allowing systems to achieve and even exceed the elusive "five-nines," or 99.999%, availability. The fabric interface provides a full-mesh interconnection in which each slot has a direct connection to every other slot. Because it has a built-in fabric interface, ATCA lacks the legacy-compatibility problems of other board standards. However, the basic ATCA specification does not call out a specific fabric technology for data transport. Instead, a series of subsidiary specifications define backplane details for the various fabrics, such as Ethernet, Fibre Channel, InfiniBand, StarFabric, PCIe, and RapidIO. Although this approach allows designers to build conforming boards with any fabric technology, it creates interoperability issues and promises to fragment the specifications.

Due to ATCA's telecommunications focus, GbE is the fabric of choice for many system designs, and the ATS1936 Switch Blade from Diversified Technology is a good example (**Figure 2**). The ATS1936 is a relatively low-cost ATCA switch complying with the core PICMG 3.0 ATCA specification plus the 3.1 Ethernet-fabric option. The blade features three AdvancedMC (advanced-mezzanine-card) sites for operations, administration, and management; firewall; and encryption applications. By separating the base and fabric networks, the ATS1936 provides a separate control plane and data plane and provides 1-Gbit Ethernet switching on the base fabric; the expansion paths provide 1- or 10-Gbit Ethernet switching. The single-piece price for the ATS1936 is $5245.

MEZZANINE FABRICS

To match the bandwidth and performance of the baseboard, AdvancedTCA designers added replaceable plug-in modules with many of the same features as the basic architecture. The resulting AdvancedMC standard offers designers a hot-swappable, field-replaceable module to lower maintenance costs and reduce downtime. AdvancedMC modules feature remote management and switched-fabric technology in an approximately 3×7-in. form factor. Modules come in single- or double-size configurations with compact, midsize, and full-size faceplates. On the interconnect side, AdvancedMC supports high-speed serial interconnects eventually to include all of the switched fabrics that ATCA allows. The basic specification defines a fabric interface with as many as 21 ports or 42 differential pairs, providing full-duplex, point-to-point connectivity between modules or to the baseboard. Rated at 12.5 Gbps per port, AdvancedMC can handle multiple lanes of the modern protocols such as Ethernet, PCIe, RapidI/O, and InfiniBand. Subsidiary specifications map the ports to specific fabric requirements.

MORE AT EDN.COM

For more on designing with PCI Express, go to www.edn.com/article/CA316047.

For more on VMEbus, CompactPCI, PC/104, and AdvancedTCA, go to www.edn.com/article/CA6421375.

Go to www.edn.com/080417cs and click on Feedback Loop to post a comment on this article.

With all the high-power, hot-swap, switched-fabric, and management features of AdvancedMC, designers devised a standard to use these modules to plug directly into a backplane for small, stand-alone systems. The standard, MicroTCA, provides a stand-alone chassis with a backplane that directly accepts AdvancedMC cards, thereby eliminating the ATCA carrier board. The MicroTCA specification defines a minimum system as a collection of interconnected elements consisting of at least one AdvancedMC module, a carrier hub, and a power module, as well as the interconnect, cooling, and mechanical resources to support them. The carrier hub combines the control and management infrastructure and the interconnect-fabric resources to support as many as 12 AdvancedMC modules. The smaller form factor makes the MicroTCA concept viable for lower budget applications in telecom and a range of embedded projects.

Supporting both ATCA plug-in and MicroTCA applications, Kontron offers the AM4100 AdvancedMC module, which the Freescale MPC8641D 1.5-GHz, dual-core PowerPC processor powers (**Figure 3**). The module delivers as much as 2.3-MIPS computing performance, and its 4-Gbit-Ethernet interfaces support real-time-transmission protocol, checksum, quality-of-service, and packet-manipulation activities. The AM4100 includes as much as 2 Gbytes of soldered DDR2-SDRAM, 4 Mbytes of bootable NOR flash, as much as 4 Gbytes of NAND flash, and EEPROM for user and configuration data. Two of the 4-Gbit interfaces are routable to either the module-card-edge connector or the RJ45 connectors on the front panel. Additionally, the board supports four-lane PCI Express or serial RapidIO for high-speed fabric connectivity. The AM4100 is fully hot-swappable, which makes it possible to replace the module without shutting down the ATCA carrier board or the MicroTCA system. A dedicated module-management controller manages the board and supports ATCA-management-interface commands, which allow operators to more quickly detect and eliminate faults at the module level. AM4100 board-support packages are available for Linux and WindRiver's VxWorks operating systems.

Although several architectures still compete for the universal interconnect technology, there is little doubt that switched fabric has become a necessary part of most high-performance embedded systems. Current fabric systems have improved to the point at which additional gains bump into the physical limits of voltage-slew rates, connector characteristics, and transmission-path limitations. Therefore, the addition of parallel lanes will lead to most of the data-rate improvements in the future. As serial-fabric technologies become the limiting factor to system performance, designers will head to the laboratory and come through with the next-generation high-speed data-delivery system. Are we ready for optical paths?EDN

FOR MORE INFORMATION

Diversified Technology
www.dtims.com

Freescale Semiconductor
www.freescale.com

Kontron
www.kontron.com

Mercury Computer Systems
www.mc.com

Motorola
www.motorola.com

Pentek
www.pentek.com

PICMG
www.picmg.org

PCI SIG
www.pcisig.com

RapidIO Trade Association
www.rapidio.org

VITA
www.vita.com

WindRiver
www.windriver.com

Xilinx
www.xilinx.com

You can reach Technical Editor Warren Webb at 1-858-513-3713 and wwebb@edn.com.