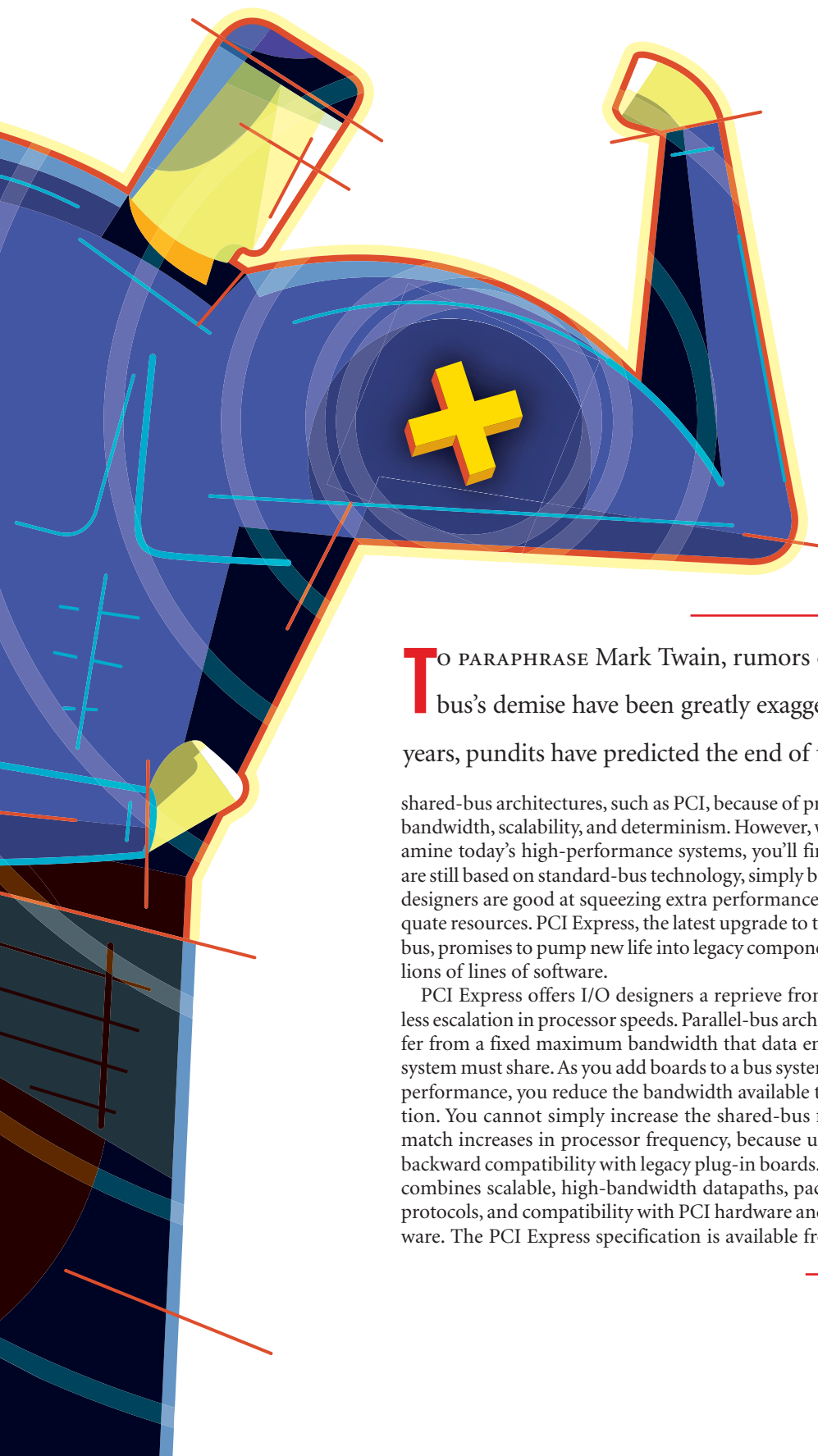


# Pump it up

ILLUSTRATION BY STEVEN LYONS

WITH A NEW PCI EXPRESS  
SPECIFICATION IN HAND,  
DESIGNERS ARE CREATING  
THE INTERCONNECTING  
BUILDING BLOCKS  
OF TOMORROW'S  
HIGH-PERFORMANCE  
EMBEDDED SYSTEMS.



**T**O PARAPHRASE Mark Twain, rumors of the PCI bus's demise have been greatly exaggerated. For years, pundits have predicted the end of traditional

shared-bus architectures, such as PCI, because of problems with bandwidth, scalability, and determinism. However, when you examine today's high-performance systems, you'll find that most are still based on standard-bus technology, simply because clever designers are good at squeezing extra performance from inadequate resources. PCI Express, the latest upgrade to the aging PCI bus, promises to pump new life into legacy components and millions of lines of software.

PCI Express offers I/O designers a reprieve from the relentless escalation in processor speeds. Parallel-bus architectures suffer from a fixed maximum bandwidth that data endpoints in a system must share. As you add boards to a bus system to increase performance, you reduce the bandwidth available to each function. You cannot simply increase the shared-bus frequency to match increases in processor frequency, because users demand backward compatibility with legacy plug-in boards. PCI Express combines scalable, high-bandwidth datapaths, packetized data protocols, and compatibility with PCI hardware and driver software. The PCI Express specification is available from the PCI-

*At a glance*.....46  
*Another view on high-performance interconnections* .....46  
*For more information* .....50



SIG (PCI Special Interest Group), which maintains the spec.

In addition to bandwidth limitations, real-time problems occur with shared-bus systems. External events happen at random times, and a bus system can respond to only one event at a time. If a data transfer is in progress and a higher priority transfer needs the bus, circuitry or software must suspend the lower priority data until the high priority finishes and then retransmit the blocked data. This contention for the bus is proportional to the number of nodes in the system. The latency of suspended data may become intolerable in some real-time situations. PCI Express eliminates these problems by providing a separate interconnect path for each point-to-point data-transfer possibility so that multiple data transfers can occur simultaneously.

PCI's developers introduced the specification more than 10 years ago as a high-performance alternative to the ISA bus for desktop computers. The initial bus frequency was 33 MHz with a 32-bit-

### AT A GLANCE

- ▶ As data bandwidths escalate, the conventional-PCI shared-bus technology is too slow for many high-performance I/O applications.
- ▶ PCI Express promises to extend the life of PCI by providing a low-cost, scalable, serial interconnection that is compatible with legacy hardware and software.
- ▶ The basic PCI Express data lane transfers data at 2.5 Gbps and scales to as many as 32 parallel lanes to form a single high-bandwidth interconnect.
- ▶ Advanced Switching options extend the PCI Express signal-routing capabilities for high-availability, dynamic scalability, and system management.

wide parallel datapath, yielding a 133-Mbyte/sec data rate (**Table 1**). In 1995, the PCI-SIG released Version 2.1 of the specification, which increased the bus

frequency to 66 MHz and path widths to 64 bits and maintained backward compatibility with earlier hardware and software. In the late 1990s, the developers added PCI-X to the mix to initially boost clock frequencies to 133 MHz and then finally to as much as 533 MHz. Although every generation is interoperable, each device on the bus must operate at the speed of the slowest installed adapter. In addition, buses in some higher frequency PCI-X configurations have only a single slot. These problems, along with ever-increasing data-transfer rates, led designers to a develop PCI Express, a new architecture for interconnecting I/O devices.

### FAST SERIAL LINKS

Formerly known as 3GIO (Third Generation Input Output), PCI Express is based on LVDS (low-voltage differential signaling) for maximum bandwidth between nodes. The basic PCI Express link comprises two signal paths that use small differential voltage swings and constant-

## ANOTHER VIEW ON HIGH-PERFORMANCE INTERCONNECTIONS

By Sam Fuller, RapidIO Trade Association

The RapidIO interconnect architecture solves the unique connectivity and reliability challenges of high-performance embedded systems. Embedded equipment today usually implements a distributed-processing model using SOC (system-on-chip) components that integrate system-processing and peripheral functions. These SOC devices, however, are still but subcomponents in larger systems. Therefore, the need exists for a technology to connect these SOC devices to accomplish the larger system task. (This scenario is unlike that of the desktop market, in which a set of relatively simple peripheral devices typically support one microprocessor. PCI adequately serves that segment.) In these distributed-processing systems, the hierarchical host/peripheral single address space model that PCI and PCI Express use place unwelcome restrictions on the

embedded-system architecture.

For example, one DSLAM (digital-subscriber-line-access-multiplexer) system uses multiple control processors, network processors, and DSPs. Control, communications, and storage processors all work together to produce the functions of a storage-area-network switch.

For these types of systems, RapidIO's peer-to-peer model provides a better system architecture. RapidIO's network layer is similar to the Ethernet-network layer but, unlike Ethernet, RapidIO offers low-overhead hardware-based support for memory-mapped and message transactions among devices. RapidIO's processing elements can communicate with each other through an efficient peer-to-peer intrasystem network at performance levels of 1 to 60 Gbps per link using just a small fraction of the CPU performance.

In addition to offering a system-connectivity model and bandwidth in line with the requirements of embedded-system vendors, the RapidIO technology also offers hardware-based reliability suitable for 99.999-class systems, an important requirement in the embedded-storage and -communications markets.

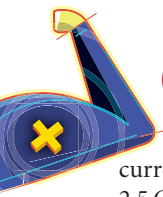
It is easier to optimize each physical-layer interface to its unique requirements in a point-to-point-switched environment, unlike a shared-bus topology, in which all devices must operate with exactly the same interface parameters. Taking advantage of this characteristic, RapidIO provides optimized physical layers for the various application requirements. Again, like Ethernet, RapidIO provides common medium-access-control and upper protocol layers that physical layers support. You can optimize these layers for the reach,

bandwidth, latency, and power dissipation that a system in that application requires.

RapidIO is a relatively mature technology. Its developers started work on the standard in 1997 and completed it in 2001. 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. Despite the maturity of the technology, work continues within the RapidIO Trade Association to further extend the technology. The association is also working on higher speed physical layers as well as protocol enhancements to broaden the market opportunities for the technology within the embedded-system market.

#### AUTHOR'S BIOGRAPHY

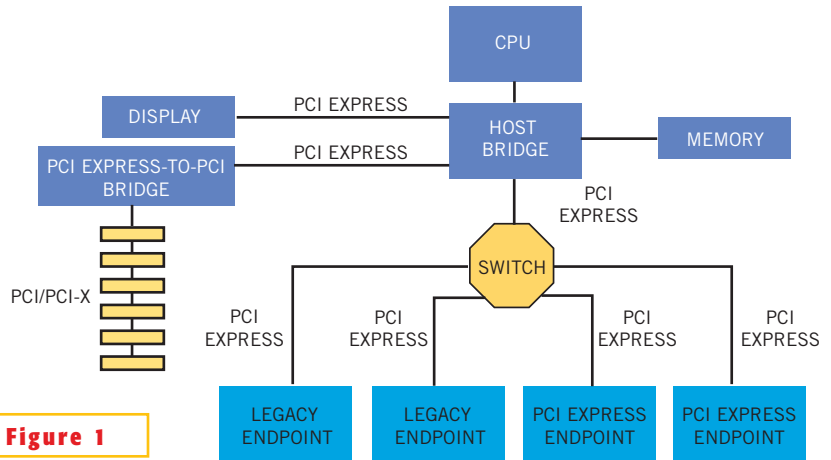
*Sam Fuller is president of the RapidIO Trade Association.*



current line drivers to communicate at 2.5 Gbps in each direction. Designers expect this data rate to increase to 10 Gbps in each direction as silicon-production technology improves. Low-voltage swings deliver low-noise signals at low power consumption. You can easily increase the bandwidth of a PCI Express link by simply adding signal pairs, or “lanes,” until you reach the desired performance level. The PCI Express specification supports  $\times 1$ -,  $\times 2$ -,  $\times 4$ -,  $\times 8$ -,  $\times 16$ -, and  $\times 32$ -lane widths. The basic 2.5-Gbps direction-transfer rate plus encoding overhead results in a single-lane performance of approximately 200 Mbytes/sec.

Although the configuration topologies are unlimited, a typical PCI Express system comprises a CPU, a host bridge, memory, and multiple I/O devices (Figure 1). The switch replaces the conventional shared bus and provides communications channels between the CPU/memory and individual I/O-device endpoints. The switch can also set up multiple, simultaneous device-to-device data transfers that do not pass through the host bridge. These independent interconnections greatly increase the effective bandwidth of the system compared with a shared-bus topology with sequential data transfers. Although the figure shows the switch as a separate block, you could incorporate it into the same silicon as the host bridge to reduce costs. You can configure device endpoints to accept PCI, PCI-X, or PCI Express adapters. An endpoint may also include a PCI Express to PCI bridge to drive a conventional, multidrop, legacy PCI bus.

The base specification defines PCI Express as having a transaction layer, a data-link layer, and a physical layer. These logical layers further divide into transmitting and receiving sides to process both inbound and outbound information (Figure 2). Although these layers do not represent a particular physical implementation, they are helpful in describing the functions of PCI Express. On the transmitting side, the transaction and data-link layers divide information into packets and append addressing, error-checking, and sequence information to ensure proper transfer through the interconnecting datapath. The physical layer includes parallel-to-serial conversion, path drivers, and impedance-match-



**Figure 1**

**A sample PCI Express system topology includes a CPU, a host bridge, memory, and multiple I/O devices.**

ing circuitry. If an interconnection link has more than one lane width, the physical layer is responsible for distributing data evenly across all paths. On the receiving side, the three layers decompose packets to match the original data.

The basic premise of PCI Express is that the host PCI software remains compatible with and can talk to and receive data from an endpoint device without new drivers or operating-system software. PCI host software tasks include both initialization and runtime functions. During initialization, the operating system must be able to discover all of the endpoint devices and their characteristics with the same protocol a shared-bus system uses. During runtime, the PCI

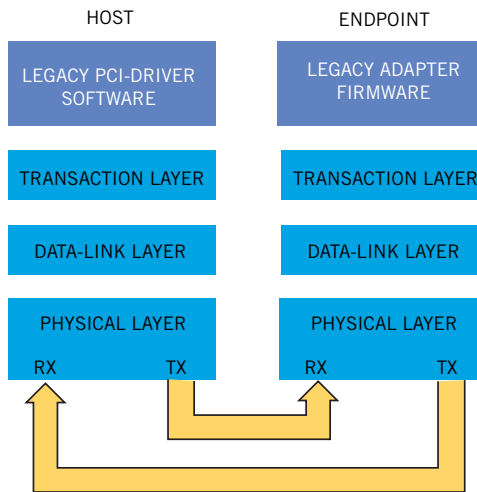
Express layers must automatically packetize the information, send it over the serial link, and reconstruct it on the other end without the need to modify the original software or the endpoint-adaptor hardware. This transparency allows designers to upgrade high-data-rate sections of their systems for improved performance and retain other low-cost but adequate I/O adapters.

### LOW-LATENCY DATA

The PCI Express architecture also simplifies the implementation of other high-performance system requirements. The high-bandwidth serial links ensure low-latency transfers for isochronous delivery of audio- and video-data streams.

You can independently scale individual data links to as wide as  $\times 32$  lanes for high bandwidth where necessary. PCI Express allows system designers to regulate the quality of service across the interconnection for improved signal integrity by adjusting packet size and datapaths. Hot-plug and -swap sequences are also simpler to implement with a packet-based, low-pin-count interconnection architecture.

To adapt PCI Express to the requirements of the communications and embedded-system markets, developers have ensured compatibility with the soon-to-be-released Advanced Switching specification. This new specification allows dynamic rerouting, multiple protocols, and high-avail-



**Figure 2** **PCI Express has a transaction layer, a data link layer, and a physical layer, although these layers do not represent a physical implementation.**



ability applications. Unlike with PCI Express, Advanced Switching's developers based it on a peer-to-peer architecture, and it requires fabric-manager firmware to configure the datapaths and handle fabric-specific events or error messages. Advanced Switching uses the same physical- and data-link layers as the PCI Express architecture, but Advanced Switching's transaction layer provides specialized communications features, such as high-availability, peer-to-peer, or multicast networking; system management; scalability; and support for most networking protocols.

You can mechanically configure PCI Express hardware into form factors that look similar to those of PCI I/O adapters. A general-purpose computer motherboard could have standard PCI slots alongside PCI Express slots. The PCI Express slots has 36 pins for a  $\times 1$  lane width and an increase in the number of pins, depending on the lane width. The PCI-SIG is working on a PCI Express Mini Card specification to produce a successor to the conventional Mini PCI for wired and wireless-communication peripherals for notebook and mobile computers. The PCMCIA Trade Association has released the Newcard specification for a PCI Express-compatible modular-expansion device to replace the PC card standard (Figure 3). As you would expect, PCI Express has also found its way into high-performance embedded sys-



**Figure 3**  
The Newcard specification provides a PCI Express-compatible modular-expansion device to replace the current PC-card standard.



**Figure 4**  
The PETracer ML (multilane) analyzer system from Access Technology features full bidirectional decoding and capturing of  $\times 8$ ,  $\times 4$ ,  $\times 2$ , or  $\times 1$  PCI Express links.

tems, such as the new AdvancedTCA (Advanced Telecom Computing Architecture) from the PICMG (PCI Industrial Computer Manufacturers Group). AdvancedTCA incorporates the latest trends in high-speed interconnection technologies and next-generation processors to

give users improved reliability, manageability, and serviceability.

Although PCI Express chip sets have yet to appear, several manufacturers have development projects in progress. The Intel developer network for PCI Express architecture Web site offers a catalog on which vendors can post their product descriptions. Both Texas Instruments and PLX Technology are pioneers in PCI Express-silicon development, and both expect to soon announce bridge devices. Xilinx offers PCI Express endpoint core IP (intellectual property) with  $\times 1$ -lane width that supports the company's Virtex-II Pro FPGA family. In addition, its Real-PCI Express design kit contains a tested and verified PCI Express  $\times 1$  IP core, a prototyping board, reference designs, and support services. The design kit is available now for \$29,000. In another project, Artisan Components is working on a PCI-Express PHY (physical-layer) IP core for 0.13-micron chip designs. The modular, eight-lane PCI Express 2.5-Gbps core acts a general-purpose I/O interconnection for adapter cards and attachment points for graphics devices or other serial interconnections, such as 1394b, USB 2.0, InfiniBand, and Ethernet. The PCI Express core creates a complete serial link including multiplexer/demultiplexer, data-encode/decode, and clock-recovery circuitry.

#### EXPRESS TEST DEVICES

Because PCI Express is in the early stages of development, its designers are interested in off-the-shelf test and analysis equipment. One company that offers such equipment, Computer Access Technology, recently introduced an advanced PCI Express-verification system. The company's PETracer ML (multilane)

**TABLE 1—DATA-WIDTH, FREQUENCY, AND BANDWIDTH COMPARISONS**

Specification	Data width	Frequency	Maximum bandwidth
PCI 1.0	32 bits	33 MHz	133 Mbytes/sec
PCI 2.1	64 bits	33 to 66 MHz	266 Mbytes/sec to 533 Mbytes/sec
PCI-X 1.0	64 bits	133 MHz	1.06 Gbytes/sec
PCI-X 2.0	64 bits	266 to 533 MHz	2.1 Gbytes/sec to 4.3 Gbytes/sec
PCI Express	Serial	2.5 to 80 GHz	500 Mbytes/sec to 16 Gbytes/sec

### FOR MORE INFORMATION...

For more information on products such as those discussed in this article, contact any of the following manufacturers directly, and please let them know you read about their products in *EDN*.

**Agilent Technologies**  
1-650-752-5000  
www.agilent.com

**Artisan Components**  
1-408-734-5600  
www.artisan.com

**Computer Access Technology Corp**  
1-408-727-6600  
www.catc.com

**Mercury Computer Systems**  
1-978-256-1300  
www.mc.com

**PCMCIA (Personal Computer Memory Card International Association)**  
www.picmg.org

**Texas Instruments**  
www.ti.com

**AMD (Advanced Micro Devices)**  
1-408-749-4000  
www.amd.com

**Catalyst Enterprises**  
1-408-365-3846  
www.getcatalyst.com

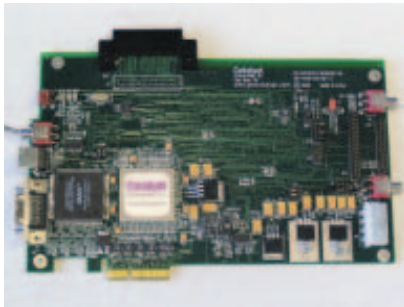
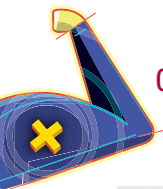
**Intel PCI Express**  
www.intel.com/technology/pciexpress/devnet

**Motorola**  
www.motorola.com

**PICMG (PCI Industrial Computer Manufacturers Group)**  
www.picmg.org

**PLX Technology**  
1-408-774-9060

**Xilinx**  
1-408-559-7778  
www.xilinx.com



**Figure 5**

**The PX-4 PCI Express Exerciser from Catalyst**

**Enterprises is an advanced signal generator for PCI Express systems with links as wide as  $\times 4$ .**

PCI Express-analyzer system features full bidirectional decoding and capture of  $\times 8$ ,  $\times 4$ ,  $\times 2$ , or  $\times 1$  PCI Express links (Figure 4). The system uses a high-impedance, nonintrusive probing technology without affecting data flow. You can tap into PCI Express fabric with a slot-interposer card or a midbus probe. PE-Tracer ML uses the company's CATC Trace software to process, decode, and

analyze captured PCI Express traffic. CATC Trace provides complete decoding of transaction-layer packets, data-link-layer packets, and all PCI Express primitives. For deeper analysis, users can display packet contents as raw 10-bit codes. For automated testing of SOC (system-on-chip) products, Agilent Technologies has integrated protocol analysis of the PCI Express interconnect standard into the popular 93000 SOC-test system.

Catalyst Enterprises offers another test device for early support for PCI Express development. The company's PCI Express Exerciser, the PX-4, is an advanced signal generator for PCI Express systems with link widths to  $\times 4$ . In a PCI Express card form factor, the PX-4 acts as a pattern generator or as a PCI Express agent (Figure 5). You can also set the PX-4 can to act simultaneously as a master and as a local memory device to perform de-

vice-emulation functions. The module can act as a stand-alone development platform for testing PCI Express designs without a motherboard.

Although it has the support of Intel and dozens of other vendors, PCI Express is not the only interconnection standard vying for a place in your future designs. AMD has developed the HyperTransport I/O link architecture with a maximum data rate of 1.6 GHz on each differential wire pair and datapaths as wide as 32 bits. In addition, Motorola and Mercury Computer Systems initially sponsored RapidIO,

which supports both parallel, onboard and serial, backplane connections (see sidebar "Another view on high-performance interconnections"). No matter which architecture gains the largest market share, the high-speed, serial data link should become the next step in your high-performance designs. Optical links will perhaps follow. □

You can reach Technical Editor Warren Webb at 1-858-513-3713, fax 1-858-486-3646, e-mail [wwebb@edn.com](mailto:wwebb@edn.com).

