The Gibson digital guitar seems to be a conventional musical instrument ... until you notice the RJ45 port.

DO YOU WANT TO ROUTE HIGH-BIT-RATE DIGITAL AUDIO OVER LONG EXPANSES WITH MINIMAL LATENCY AND "FIVE-NINES" RELIABILITY? FIND OUT WHETHER A CONVENTIONAL ETHERNET SCHEME WILL MEET YOUR NEEDS OR WHETHER YOU NEED AN AUDIO-OPTIMIZED VARIANT.

CAT5 tracks: Audio goes the distance, reliably and on time

BY BRIAN DIPERT • SENIOR TECHNICAL EDITOR

The next time you're at a music concert, take a look at the cabling that connects on-stage gear to the mixing board in the middle of the audience. Odds are good that you'll see one or several thick "snakes" containing bundles of wire, with each wire carrying a unique analog signal to or from the stage. Heavy? Bulky? EMI-radiating? Sensitive to environmental noise? You bet. All of these issues, and more, hamper the traditional analog-interconnect scheme. So, why is this seemingly archaic approach, mimicked in many recording studios and other audio-equipped environments (including churches, offices, and school auditoriums) still in use?

The wheels of progress in the commercial-, industrial-, and professional-audio industries move at a snail's pace. Gear is designed for long life in spite of rough handling, new-gear budgets are slim to nonexistent, and replacement and upgrade opportunities are consequently few and far between. And when these opportunities do occur, potential customers face a bewildering assortment of often-incompatible approaches, as vendors attempt to lock them into sole-sourced equipment suites. Still, the analog-to-digital conversion is under way, and de facto standardization driven by market leadership and attrition, aided by industry-standards bodies, is whittling down the diversity of options (see Brian's blog entry "Audio over CAT5: Proprietary alternatives and standardization efforts" at www.edn.com/briansbrain).

The "diversity-of-options" claim may baffle those of you who listen to streaming audio from the Internet or who route your server-housed music libraries to various pieces of LAN-connected gear in your homes and offices. Isn't conventional, commodity, and cheap TCP/IP-based Ethernet over CAT5 (Category 5) cable specified to a guaranteed 100m drive distance or, even better yet, no-cables-required WiFi (Wireless Fidelity) good enough? (See **sidebar** "Cabling

AT A GLANCE

- Audio-tuned CAT5 (Category 5)-cable configurations travel into applications where TCP/IP fears to tread.
- CobraNet touts its compatibility with conventional Ethernet traffic, along with the backing of a major semiconductor supplier.
- MaGIC (media-accelerated global-information carrier) is slow out of the starting gate; manufacturers are folding some of its ideas into industry-standard committee activities.
- EtherSound delivers tight tolerances and minuscule latencies at the expense of Ethernet compatibility.
- AudioRail chooses a cost-optimized path that differs from its competitors' trails.

choices.") In some cases—more and more as technology progresses—it is. But in demanding audio environments, its time-insensitive, bulk-file-transfer heritage hinders its applicability.

Press "play" on your LAN- or WAN-connected audio receiver, and there's a several-second delay as the local buffer memory fills up. The buffer is there to compensate for protocol overhead and the inevitable packet-routing delays to come, resulting in out-of-order packet reception, along with the even more egregious network congestion, collisions, and other factors that result in packet loss and force retransmission or error concealment. Most home networks are, at any given time, lightly loaded—a situation that's quickly changing in the era of video streaming. However, you can't make the same assumption of commercial, industrial, and professional networks, particularly those handling multiple channels' worth of high-resolution audio, along with control *and* general data traffic (see Brian's blog entry "Audio over CAT5: Conventional counterparts" at www.edn.com/briansbrain).

Detectable latency is intolerable in a live-performance setting, and it's also unacceptable in most recording environments. What's "detectable"? According to Gibson's chairman and chief executive officer, Henry Juszkiewicz, whereas "university guys" might claim that the human eyes, ears, muscle, and brain can combine forces and detect picosecond-level latency, audio technologists should shoot for a half-millisecond latency through any network "hop." Professional musicians might be able to detect a total latency (that also encompasses analog-to-digital and digital-to-analog conversions, analog- and digital-audio processing algorithms, and other delays) on the order of 2 msec, but anything less than 5 msec is difficult for the average listener to perceive (see Brian's Blog entry "Audio over CAT5: Divide and conquer" at www.edn.com/briansbrain).

COBRANET STRIKES

Although conventional TCP/IP-based CAT5 cabling may not currently address demanding audio-application needs, the volume-driven low cost of CAT5 hardware *is* attractive to hardware and software suppliers. Cirrus Logic's CobraNet technology, the fruit of the company's mid-2001 acquisition of Peak Audio, is one of the more mature contenders for the CAT5-audio throne and a market-share leader. CobraNet is also notable for its ability to coexist with other Ethernet traffic on a common set of network resources (**Figure 1**). However, its capabilities aren't *completely* unbounded; Michael Johas Teener—currently a networking architect at Broadcom and formerly the chief architect of FireWire at Apple—observes that CobraNet employs a "limited-topology" concept to ensure its claimed capabilities are achievable. "One extra switch, or a legal but 'non-CobraNet' configuration," his analysis states, and quality will be poor (see **sidebar** "Additional protocols," and **Reference 1**).

CobraNet's protocol combines one or more audio channels into an Ethernet packet, along with identifying data, such as the sample size and rate. It also enables you to send control information by bridging RS-232 data over Ethernet or by generic Ethernet packets containing user-defined control data. The first-generation CobraNet transceiver modules contained digital logic housed in FPGAs alongside Motorola (now Freescale) DSPs and analog circuitry. Second-generation modules combine analog and digital circuits within a single CS1810xx ASIC , alongside a Cirrus Logic DSP. (Variants with 2×2, 8×8, and 16×16 channels are available.) Cirrus also launched DSP-inclusive CS4961xx product variants at January's Consumer Electronics Show. A multidimensional matrix of sample size, sample rate, number of audio channels to support, and desired latency—5.33, 2.66, and 1.33 msec—defines which chip variant your application will require. (See Brian's blog entry "Audio over CAT5: interesting lit and other bits" at www.edn.com/briansbrain for a downloadable spreadsheet from Cirrus Logic that will allow you to make the necessary calculations.)

CobraNet supports 48- and 96-kHz sample rates but, oddly, not the 44.1-kHz standard that Red Book Audio CDs and common audio peripherals employ. It also tackles samples as large as 32 bits. One showcase CobraNet installation, according to Senior Marketing Manager David Parker, is Tokyo Disney Park, which combines a gigabit Ethernet optical-fiber backbone and 250 nodes' worth of 100-Mbit Ethernet transceivers. This unified network handles data traffic related to lighting, ride controls, and point-of-sale cash registers, as well as distributed- and live-audio broadcasts. "The park is open seven days per week, 365 days per year," says Parker, "and any network failure would result in a complete park evacuation."

MAGIC TUNES UP

Back when Gibson was contemplating a move to a digital-guitar-to-amplifier-and-mixer interface, CobraNet still required upfront and per-channel royalty payments. And, pragmatically, the decision-makers at the company probably had visions of incoming royalty-revenue streams dancing in *their* heads, as well. So, Gibson went its own way, and MaGIC (media-accelerated global information carrier) was the result. (Zipi, a token-ring-based approach from the company, predates MaGIC.) Along with a FAQ, a full specification, and other materials, Gibson's Web site offers an approximately 10-minute video in Windows Media format that, among other things, outlines Gibson's sweeping vision for MaGIC as a control, multimedia, and general data-transfer protocol for the home. However, Gibson produced the video in 2001; since then, other companies and industry-standards bodies have produced alternative, now-dominant schemes for the consumer-electronics market, including Apple's Bonjour, UPnP (Universal Plug and Play), and Microsoft's PlaysForSure and Windows Media Extenders.

Nonetheless, Gibson still believes its approach has merit in commercial-, industrial-, and professional-audio settings. MaGIC provides as many as 32 channels of 32-bit bidirectional audio over 100-Mbit Ethernet with sample rates as high as 192 kHz. (As the number of channels increases, the allowable sample rate decreases, and visa versa.) Gibson claims that data and control transport occurs 30 to 30,000 times faster than MIDI (musical-instrument digital interface). Additional cable features include phantom power, automatic clocking, and network synchronization. And, notably, the MaGIC specification touts 250-µsec point-to-point latency times across 100m CAT5 spans.

These impressive performance specifications, however, come at the price of less-than-full Ethernet compatibility. MaGIC *does* conform to the IEEE 802.3 PHY (physical) layer and uses standard CAT5 cable and RJ45 connectors. (Ruggedized EtherCon connectors from Neutrik are also available.) However, it employs unconventional packet sizes and requires a unique MAC (media-access-control) implementation. The FAQ on Gibson's Web site states, "The footprint of our packet is the same as Ethernet UDP [User Datagram Protocol]. We differ from standard Ethernet because our packet size and transmission rate do not change ... We route on the MAC layer, which is Layer 2. We refer to data as frames on this layer instead of packets. Some parts of the MAC layer can be incorporated into software" (**Reference 2**).

Gibson is offering 10-year royalty-free MaGIC licenses to all interested parties, hoping to stimulate interest in the protocol as an industry-standard interconnect scheme. A $2500 software-development kit and a $600 evaluation board are available, as are VHDL implementations of the CobraNet transceiver targeting Altera and Xilinx FPGAs. Juszkiewicz acknowledges that a modern ASIC implementation will be necessary to hit cost targets in high-volume applications and is negotiating for MaGIC protocol support in future CobraNet silicon spins that target Gigabit Ethernet transfer rates. (Cirrus Logic has no comment on Juskiewicz's claims.) And where's the Gibson digital guitar that was supposed to be in production by the end of 2002? (See Brian's blog entry "Audio over CAT5: Hands-on jams" at www.edn.com/briansbrain). Prototypes have existed for several

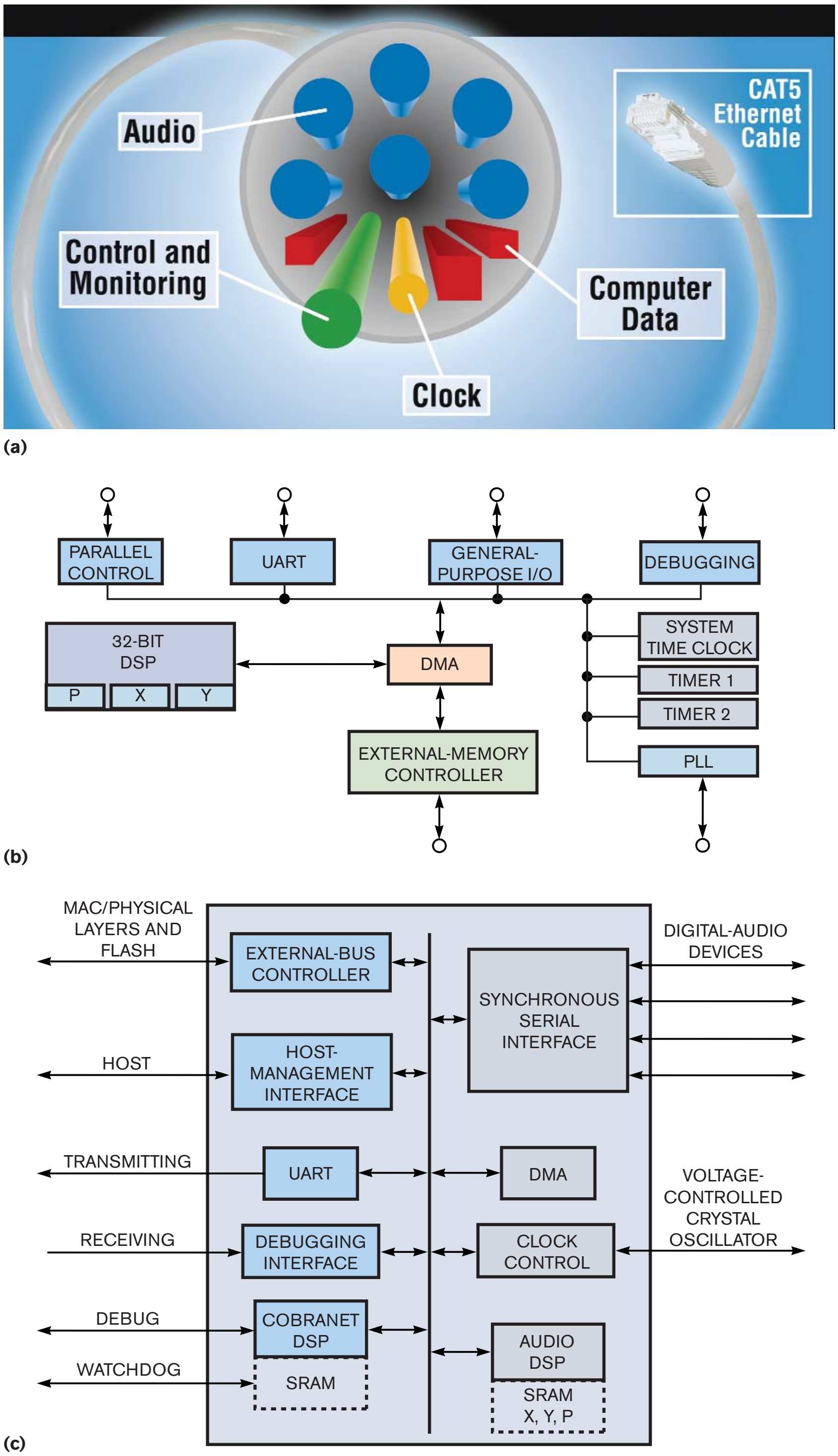

Figure 1 Cirrus Logic touts CobraNet's interoperability with other Ethernet traffic (a), and latest-generation ASSPs (application-specific standard products) come in conventional (b) and audio-DSP-core-inclusive (c) variants.

years, but the company is waiting for a partner-developed pickup that delivers a standard of excellence reflecting MaGIC's capabilities, which include the ability to place each string on a unique audio channel.

A "SOUND" ALTERNATIVE

Digigram's EtherSound is currently CobraNet's dominant competitor. Like MaGIC, EtherSound is compatible with Ethernet at the PHY layer but requires a unique MAC. It also doesn't support traditional bus and ring topologies; instead, it requires a daisy-chain or star-interconnect architecture (based on Layer 2 switches; higher-layer switches and routers are not supported), or combinations of these two approaches (**Figure 2**). The upside of these restrictions, however, is low and deterministic latencies; an audio transfer between a master module's serial input and the next slave module's serial output is six samples—125 μsec at a 48-kHz sampling rate. Each time you add another daisy-chained EtherSound module to the path, you incur an additional approximately 1.5-μsec delay (see **sidebar** "Whither wireless?"). EtherSound is a synchronous network: The primary master module generates the network audio clock from which each downstream EtherSound device's audio clock is derived, synchronized whenever an application requires phase consistency.

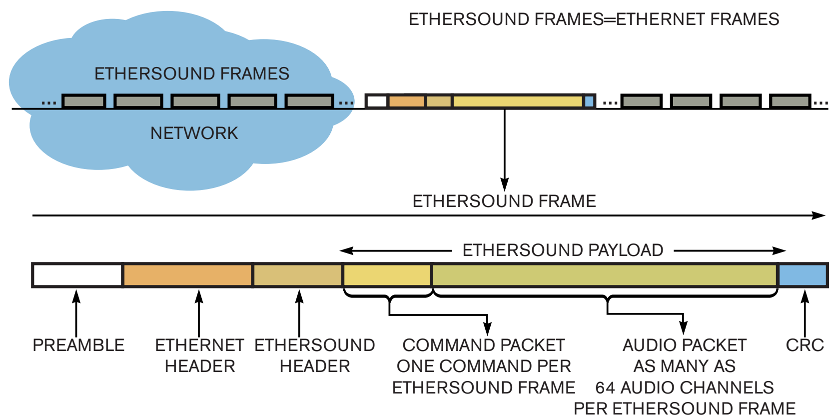

Figure 2 Each EtherSound frame combines a command packet with as many as 64 channels' worth of audio data.

Version 1.0 of the EtherSound protocol is a two-packet-per-frame approach; the first packet contains command information, and the packet that follows it contains as much as 64 channels' worth of audio information, with each channel supporting 24-bit sample sizes. EtherSound natively authorizes 44.1- and 48-kHz sample rates; audio streams at higher sampling rates—88.2, 96, and 192 kHz, for example—employ multiple channels within one EtherSound frame. A 100-Mbps EtherSound network can carry a variety of combinations of audio streams—64 audio streams at 48 kHz, for

Whither wireless?

Given that many on-stage guitar players–and, for that matter, singers and other musicians–have "gone wireless" in recent years, the idea of tethering a guitar over CAT5 (Category 5) cable seems antiquated. David Mayne, Gibson's vice president of business development for MaGIC (media-accelerated global-information carrier), points out that the company's implementation focuses on delivering low latency and maintaining 100% QOS (quality of service). Nothing inherent in the design prevents wireless transport, but wireless options fall short of the company's QOS expectations. Gibson recognizes the benefits of wireless and expects continued adoption of untethered options, but the company's first priority is quality. Chief Executive Officer and Chairman Henry Juszkiewicz refers to UWB (ultrawideband) as the "big hope," due to its comparative immunity to interference and consequent latency predictability versus today's 2.4- and 5.2-GHz WiFi (Wireless Fidelity) approaches.

Garth Wiebe, founder of AudioRail Technologies, also shares his thoughts on wireless transmission in commercial, industrial, and professional audio: "Wireless is full of pitfalls. Enter RF, and you must account for signal dropouts and retransmissions of data. This is no problem for data, but live sound cannot tolerate these. To account for them, in theory, would require large amounts of buffering, which would incur large amounts of latency, which is also unacceptable."

Expanding the discussion to video, Michael Johas Teener, a networking-system architect at Broadcom, points out, "The quality of service for wireless is not adequate for HD-quality video. Latency is excessive–tens of milliseconds–in a single-attachment-point domain, much worse in a mesh, and normal home environments can result in momentary packet loss" (Reference A). That's the reason that Belkin, for example, chose a proprietary wireless-transmission scheme based on a Magis Networks chip set for its RemoteAV transmitter-and-receiver set.

REFERENCE

A. http://grouper.ieee.org/groups/802/3/tutorial/mar05/tutorial_1_0305.pdf.

example, or 62 audio streams at 48 kHz plus one audio stream at 96 kHz, all the way up to 16 audio streams at 192 kHz.

Materials available under license to speed your time to production include a software-development kit, source and object code, an evaluation board, and two reference designs targeting Xilinx Spartan-IIe FPGAs. Digigram also sells a number of EtherSound transmitters and receivers with various numbers and types of analog inputs and outputs, along with a the miXart 8 ES PC sound card. Planned 100-Mbit enhancements include the ability to define bidirectional daisy chains anywhere on a network, support for Xilinx Spartan-3 FPGAs, ubiquity of network management, ring-based network redundancy, and the ability to generate the network clock from any device. And, like Cirrus Logic and Gibson, Digigram is eyeing Gigabit Ethernet as a future enhancement, enabling the support of standard Ethernet components along with additional media transports, such as video and generic data.

FOCUS ON SIMPLICITY

Garth Wiebe founded AudioRail Technologies based on the premise that most people in live sound have little money to spend. He explains, "You have schools, churches, small musical organizations, theaters, conference centers, and so on. There are large sound companies and large venues, for sure, but they are dwarfed by the majority of smaller, lower budget organizations." As a result, he created the elementary M11 protocol, a 4-bit, 25-MHz data stream based on time-division multiplexing. AudioRail uses Ethernet transceivers but replaces the Ethernet MAC with a simpler FPGA-based programmable-logic design.

The result, according to Wiebe, is a reliable product that costs only a fraction of what you'd pay for other real-time-networked-audio approaches. Says Wiebe, "We have yet to have one fail in the field. Period. No customer complaints of malfunction. Period. And we have them in countries all over the world in addition to here in the United States." You can see for yourself how AudioRail's cost and reliability claims stack up to scrutiny; the company's Web site details both its and its competitors' technologies and products. Each $500 ADAT (Alesis digital-audio tape) rx32tx32 rack-mountable unit provides eight ADAT "light-pipe" optical connections, encompassing four inputs and four outputs (**Figure 3**). Each light-pipe port can carry eight 16- to 24-bit, 48-kHz audio channels for a sum total of 64 channels' worth of 24-bit, 48-kHz audio through each AudioRail dual-CAT5 link (see Brian's blog entry "Audio over CAT5: Other approaches" at www.edn.com/briansbrain).

AudioRail, like its competitors, delivers low latency: approximately 5 µsec in

Cabling choices

Is CAT5's (Category 5's) 100m specification too short for the distance you need to route audio? Consider fiber-optic cable. Cirrus Logic's CobraNet literature, for example, claims that it's possible to reliably drive a 2-km distance over multimode fiber or 100 km over single-mode fiber. Fiber-optic cable, unlike all electrical-based cabling schemes (to a greater or lesser extent), also doesn't emit an EMF and isn't susceptible to interference from other EMF-generating sources.

Figure 3 AudioRail transceivers bridge ADAT and the company's proprietary, cost-optimized MII protocol.

Additional protocols

CAT5 (Category 5) cabling's low cost and long-distance drive capability make it an attractive option for audio-data formats beyond the Ethernet-packetized variety. The MIDI (musical-instruments digital interface) Manufacturers Association, for example, is pursuing standardization of MIDI transport over CAT5, according to Analog Devices' DSP Marketing Programs Manager (and formerly corporate marketing manager at MIDI software provider Staccato Systems) Denis Labrecque, although association representatives didn't respond to requests for information, and the association's Web site currently lists only MIDI-over-IEEE-1394 specifications.

FireWire-encoded audio over CAT5 cabling is another possibility, one that FireWire audio pioneer Bob Moses is pursuing with his latest venture as vice president and director of engineering of Wavefront Semiconductor. Moses comments, "As far as I know, I'm the only guy to have built an audio-over-CAT5-using-Firewire box, so far. Actually, I've built two. The first was a reference design called OnRamp using TI silicon. The second reference design is the evaluation module for our new Dice II FireWire audio chip...We plan to send 400 Mbps down CAT6 (or 200 Mbps down CAT5e or 100 Mbps down plain-old CAT5). We believe that this solution is much cheaper than an Ethernet solution, and you get all the benefits of FireWire isochronous transport, which you do not get with the other audio-over-Ethernet solutions."

MORE AT EDN.COM

+ Go to www.edn.com/050707df1 and click on Feedback Loop to post a comment on this article.

+ For more on this and related topics, see the Brian's Brain blog at www.edn.com/briansbrain.

the digital domain—more precisely, 4.5 μsec+0.25 μsec/hop+0.005 μsec/m, end to end. Wiebe warns, "One feature … that provides flexibility and versatility but can cause a nuisance is that AudioRail digital-audio streams are all independently clocked to follow their source. AudioRail can be simplistically modeled as a bundle of digital-audio cables taped together. They do not have a common clock. The flexibility and versatility mean that you can run clocked domains at different sample rates through the same CAT5 cable. The potential nuisance part … is that, if you have only a single digital-audio system, you must make sure that all devices are properly slaved off of a clock master, even though they all come from the same AudioRail box. This can mean a little extra strategizing and perhaps a few extra-short-word clock BNC cables at each end to tie the devices' clocks together." **EDN**

REFERENCES

1. http://grouper.ieee.org/groups/802/3/tutorial/mar05/tutorial_1_0305.pdf.
2. www.gibsonmagic.com/MaGICfaq.pdf.

ACKNOWLEDGMENTS

Thanks to Garth Wiebe, founder of AudioRail Technologies, for his in-depth background on digital-audio-transmission history, to both Wiebe and Michael Johas Teener, a network-systems architect at Broadcom, for their information on current audio-over-CAT5 alternatives, and John Strawn, principal consultant and owner of S Systems, for his efforts in connecting me with various individuals and companies associated with this article's topics.

AUTHOR'S BIOGRAPHY

Technical Editor Brian Dipert longs for a world filled with fiber-optic cable, ultrawideband wireless, and instant-response, blazingly fast servers on the other end of his network connection. Oh, and world peace. A guy can dream, can't he?

FOR MORE INFORMATION

AES (Audio Engineering Society)
www.aes.org

Alesis
www.alesis.com

Altera
www.altera.com

Analog Devices
www.analog.com

Apple
www.apple.com

AudioRail Technologies
www.audiorail.com

Aviom
www.aviom.com

Axia Audio
www.axiaaudio.com

Belkin
www.belkin.com

Broadcom
www.broadcom.com

BSS Audio
www.bss.co.uk

Cirrus Logic
www.cirrus.com

Digidesign
www.digidesign.com

Digigram
www.digigram.com

Edirol
www.edirol.com

Freescale Semiconductor
www.freescale.com

Gibson
www.gibsonmagic.com

Great River Technology
www.greatrivertech.com

GridIron Software
www.gridironsoftware.com

Hear Technologies
www.heartechnologies.com

IEEE
www.ieee.org

Intelligent Media Technologies
www.intelligentmedia.us

Lab.gruppen
www.labgruppen.com

Magis Networks
www.magisnetworks.com

Microsoft
www.microsoft.com

MIDI Manufacturers Association
www.midi.org

MOTU
www.motu.com

Muse Research
www.museresearch.com

Neutrik
www.neutrikusa.com

Nine Tiles Networks
www.ninetiles.com

PortalPlayer
www.portalplayer.com

Rogue Amoeba Software
www.rogueamoeba.com

Roku Labs
www.rokulabs.com

Sony
www.sonyoxford.co.uk

Staccato Systems
www.staccatosys.com

Tascam
www.tascam.com

TC Applied Technologies
www.tcelectronic.com

Wavefront Semiconductor
www.wavefrontsemi.com

Xilinx
www.xilinx.com

Yamaha
www.yamaha.co.jp/tech/1394mLAN/english/

You can reach Technical Editor Brian Dipert at 1-916-760-0159, 1-781-734-8038 (fax), bdipert@edn.com, and www.bdipert.com.