ProAV ONLINE

hanley▲wood

Out of control

Digital Audio Networking Demystified
The OSI model helps bring order to the chaos of various digital audio network options.

Source: PRO AV Magazine
Publication date: 2007-08-01

By Brent Harshbarger

Networking has been a source of frustration and confusion for pro AV professionals for decades. Fortunately, the International Organization of Standardization, more commonly referred to as ISO, created a framework in the early 1980s called the Open Systems Interconnection (OSI) Reference Model, a seven-layer framework that defines network functions, to help simplify matters.

Providing a common understanding of how to communicate to each layer, the OSI model (Fig. 1) is basically the foundation of what makes data networking work. Although it's not important for AV professionals to know the intricate details of each layer, it is vital to at least have a grasp of the purpose of each layer as well as general knowledge of the common protocols in each one. Let's take a look at the some key points.
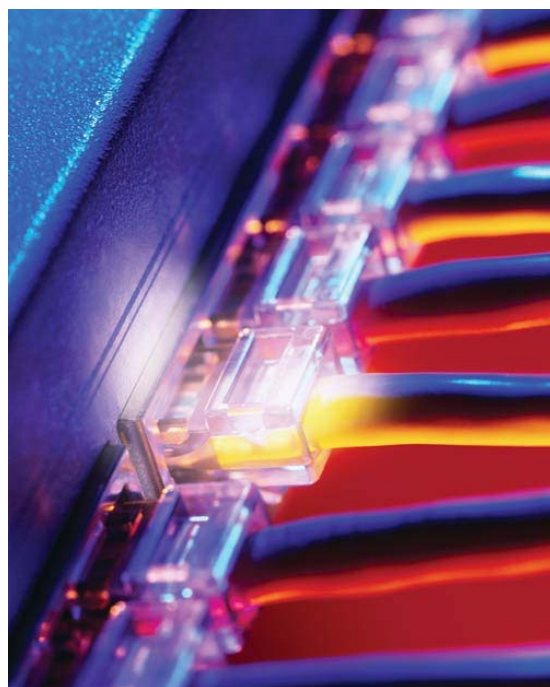
The Seven Layers

Starting from the bottom up, the seven layers of the OSI Reference Model are Physical, Data Link, Network, Transport, Session, Presentation, and Application. The Physical layer is just that — the hardware's physical connection that describes its electrical characteristics. The Data Link layer is the logic connection, defining the type of network. For example, the Data Link layer defines whether or not it is an Ethernet or Asynchronous Transfer Mode (ATM) network. There is also more than one data network transport protocol. The Data Link layer is divided into two sub-layers: the Media Access Control (MAC) and the Logical Link Control (above the MAC as you move up the OSI Reference Model).

Here is one concrete example of how the OSI model helps us understand networking technologies. Some people assume that any device with a CAT-5 cable connected to it is an Ethernet device. But it is Ethernet's Physical layer that defines an electrical specification and physical connection — CAT-5 terminated with an RJ-45 connector just happens to be one of them. For a technology to fully qualify as an Ethernet standard, it requires full implementation of both the Physical and Data Link layers.

The Network layer — the layer at which network routers operate — "packetizes" the data and provides routing information. The common protocol for this layer is the Internet Protocol (IP).

Layer four is the Transport layer. Keep in mind that this layer has a different meaning in the OSI Reference Model compared to how we use the term "transport" for moving audio around. The Transport layer provides protocols to determine the delivery method. The most popular layer four protocol is Transmission Control Protocol (TCP). Many discuss TCP/IP as one protocol, but actually they are two separate protocols on two different layers. TCP/IP is usually used as the data transport for file transfers or audio control applications.

TCP provides a scheme where it sends an acknowledge message for each packet received by a sending device. If it senses that it is missing a packet of information, it will send a message back to the sender to resend. This feature is great for applications that are not time-dependent, but is not useful in real-time applications like audio and video.

Streaming media technologies most common on the Web use another method called User Datagram Protocol (UDP), which simply streams the packets. The sender never knows if it actually arrives or not. Professional audio applications have not used UDP because they are typically Physical layer or Data Link layer technologies — not Transport layer. However, a newcomer to professional audio networking, Australia-based Audinate, has recently become the first professional audio networking technology to use UDP/IP technology over Ethernet with its product called Dante.

OSI Model

Layer 7	Application
Layer 6	Presentation
Layer 5	Session
Layer 4	Transport
Layer 3	Network
Layer 2	Data Link
Layer 1	Physical

The Session and Presentation layers are not commonly used in professional audio networks; therefore, they will not be covered in this article. Because these layers can be important to some
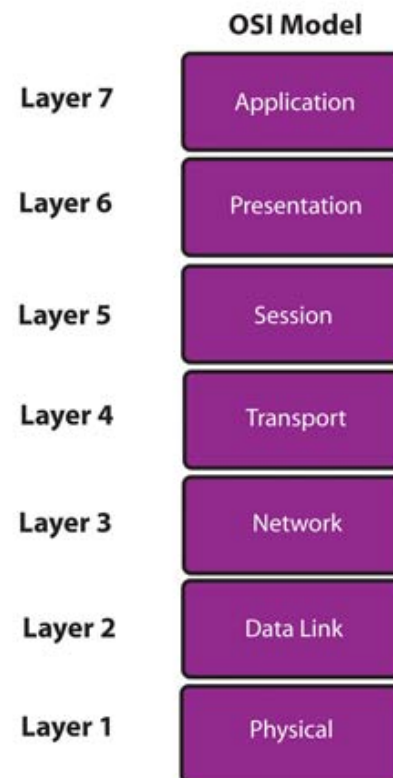
integration projects, you may want to research the OSI model further to complete your understanding of this useful tool.

The purpose of the Application layer is to provide the interface tools that make networking useful. It is not used to move audio around the network. It controls, manages, and monitors audio devices on a network. Popular protocols are File Transfer Protocol (FTP), Telnet, Hypertext Transfer Protocol (HTTP), Domain Name System (DNS), and Virtual Private Network (VPN), to name just a few.

Now that you have a basic familiarity with the seven layers that make up the OSI model, let's dig a little deeper into the inner workings of a digital audio network.

Breaking Down Audio Networks

Audio networking can be broken into in two main concepts: control and transport. Configuring, monitoring, and actual device control all fall into the control category and use several standard communication protocols. Intuitively, getting digital audio from here to there is the role of transport.

Control applications can be found in standard protocols of the Application layer. Application layer protocols that are found in audio are Telnet, HTTP, and Simple Network Management Protocol (SNMP). Telnet is short for TELetype NETwork and was one of the first Internet protocols. Telnet provides command-line style communication to a machine. One example of Telnet usage in audio is the Peavey MediaMatrix, which uses this technology, known as RATC, as a way to control MediaMatrix devices remotely.

SNMP is a protocol for monitoring devices on a network. There are several professional audio and video manufacturers that support this protocol, which provides a method for managing the status or health of devices on a network. SNMP is a key technology in Network Operation Center (NOC) monitoring. It is an Application layer protocol that communicates to devices on the network through UDP/IP protocols, which can be communicated over a variety of data transport technologies.

Control systems can be manufacturer-specific, such as Harman Pro's HiQnet, QSC Audio's QSControl, or third party such as Crestron's CresNet, where the control software communicates to audio devices through TCP/IP. In many cases, TCP/IP-based control can run on the same network as the audio signal transport, and some technologies (such as CobraNet and Dante) are designed to allow data traffic to coexist with audio traffic.

The organizing and managing of audio bits is the job of the audio Transport. This is usually done by the audio protocol. Aviom, CobraNet, and EtherSound are protocols that organize bits for transport on the network. The transport can be divided into two categories: logical and physical.

Purely physical layer technologies, such as Aviom, use hardware to organize and move digital bits. More often than not, a proprietary chip is used to organize and manage them. Ethernet-based technologies packetize the audio and send it to the Data Link and Physical layers to be transported on Ethernet devices. Ethernet is both a logical and physical technology that packetizes or "frames" the audio in the Data Link layer and sends it to the Physical layer to be moved to another device on the network. Ethernet's Physical layer also has a Physical layer chip, referred to as the PHY chip, which can be purchased from several manufacturers.

Comparing Digital Audio Systems

The more familiar you are with the OSI model, the easier it will be to understand the similarities and differences of the various digital audio systems. For many people, there is a tendency to gloss over the OSI model and just talk about networking-branded protocols. However, understanding the OSI

model will bring clarity to your understanding of digital audio networking (Fig. 2).

OSI Model	TCP/IP Model	Streaming Media	Aviom	EtherSound	CobraNet	Dante
Application	FTP					
Presentation						
Session		RTP				
Transport	TCP	UDP				UDP
Network	IP	IP				IP
Data Link	Ethernet	Ethernet		Ethernet	Ethernet	Ethernet
Physical	Copper/Fiber Wireless	Copper/Fiber Wireless	Copper	Copper/Fiber	Copper/Fiber	Copper/Fiber

This chart gives a visual representation of how the technologies compare to the OSI Reference Model

The FTP shows that the File Transport Protocol is at the applications layer and then the TCP/IP over Ethernet.
The Streaming Media shows the Real Time Protocol using UDP/IP over Ethernet Protocols.
Aviom is a proprietary Physical Layer technology.
EtherSound is a fully compliant 802.3 Ethernet technology but doesn't use the Logical Link Control sub-layer of the Data Link Layer.
CobraNet is a fully compliant 802.3 Ethernet technology and uses the both sub-layers of the Data Link Layer.
Dante is a UDP/IP over Ethernet technology.

Fig. 2. Comparison of four digital audio technologies using the OSI model as a framework. The chart above provides a visual representation of how various technologies compare to the OSI Reference Model. For example, the FTP shows that the File Transport Protocol is at the Applications layer. Next is TCP/IP over Ethernet. The Streaming Media shows the Real Time Protocol using UDP/IP over Ethernet protocols. Aviom is a proprietary Physical layer technology. A fully compliant 802.3 Ethernet technology, EtherSound doesn't use the Logical Link Control sub-layer of the Data Link layer. Also a fully compliant 802.3 Ethernet technology, CobraNet uses the both sub-layers of the Data Link layer. Dante is a UDP/IP over Ethernet technology.

Due to the integration of pro AV systems, true networking schemes are vitally important. A distinction must be made between audio networking and digital audio transports. Audio networks are defined as those meeting the commonly used standard protocols, where at least the Physical and Data Link layer technologies and standard network appliances (such as hubs and switches) can be used. There are several technologies that meet this requirement using IEEE 1394 (Firewire), Ethernet, and ATM technologies, to name a few. However, because Ethernet is widely deployed in applications ranging from large enterprises to the home, this will be the technology of focus. All other technologies that do not meet this definition will be considered digital audio transport systems, and not a digital audio network.

There are at least 15 schemes for digital audio transport systems and audio networking. Three of the four technologies presented here have been selected because of their wide acceptance in the industry based on the number of manufacturers that support it.

Let's compare four CAT-5/Ethernet technologies: Aviom, EtherSound, CobraNet, and Dante. This is not to be considered a "shoot-out" between technologies but rather a discussion to gain understanding of some of the many digital system options available to the AV professional.

As previously noted, Aviom is a Physical layer–only technology based on the classifications outlined above. It does use an Ethernet PHY chip, but doesn't meet the electrical characteristics of Ethernet. Therefore, it cannot be connected to standard Ethernet hubs or switches. Aviom uses a proprietary chip to organize multiple channels of audio bits to be transported throughout a system, and it falls in

the classification of a digital audio transport system.

EtherSound and CobraNet are both 802.3 Ethernet– compliant technologies that can be used on standard data Ethernet switches. There is some debate as to whether EtherSound technology can be considered a true Ethernet technology because it requires a dedicated network. EtherSound uses a proprietary scheme for network control, and CobraNet uses standard data networking methods. The key difference for both the AV and data professional is that EtherSound uses a dedicated network, and CobraNet does not. There are other differences that may be considered before choosing between CobraNet and EtherSound, but both are considered to be layer two (Data Link) technologies.

Dante uses Ethernet, but it is considered a layer four technology (Transport). It uses UDP for audio transport and IP for audio routing on an Ethernet transport, commonly referred to as UDP/IP over Ethernet.

At this point you may be asking yourself why does the audio industry have so many technologies? Why can't there be one standard like there is in the data industry?

The answer to the first question relates to time. Audio requires synchronous delivery of bits. Early Ethernet networks weren't much concerned with time. Ethernet is asynchronous, meaning there isn't a concern when and how data arrives as long as it gets there. Therefore, to put digital audio on a data network requires a way to add a timing mechanism. Time is an issue in another sense, in that your options depend on technology or market knowledge available at the time when you develop your solution. When and how you develop your solution leads to the question of a single industry standard.

Many people don't realize that the data industry does in fact have more than one standard: Ethernet, ATM, FiberChannel, and SONET. Each layer of the OSI model has numerous protocols for different purposes. The key is that developers follow the OSI model as a framework for network functions and rules for communicating between them. If the developer wants to use Ethernet, he or she is required to have this technology follow the rules for communicating to the Data Link layer, as required by the Ethernet standard.

Because one of the key issues for audio involves time, it's important to use it wisely.

Time Management

There are two types of time that we need to be concerned with in networking: clock time and latency. Clock time in this context is a timing mechanism that is broken down into measurable units, such as milliseconds. In digital audio systems, latency is the time duration between when audio or a bit of audio goes into a system until the bit comes out the other side. Latency has many causes, but arguably the root cause in audio networks is the design of its timing mechanism. In addition, there is a tradeoff between the timing method and bandwidth. A general rule of thumb is that as the resolution of the timing mechanism increases, the more bandwidth that's required from the network.

Ethernet, being an asynchronous technology, requires a timing method to be added to support the synchronous nature of audio. The concepts and methodology of clocking networks for audio are key differences among the various technologies.

CobraNet uses a time mechanism called a beat packet. This packet is sent out in 1.33 millisecond intervals and communicates with CobraNet devices. Therefore, the latency of a CobraNet audio network can't be less than 1.33 milliseconds. CobraNet was introduced in 1995 when large-scale DSP-based digital systems started replacing analog designs in the market. Because the "sound system in a box" was new, there was great scrutiny of these systems. A delay or latency in some

time-critical applications was noticed, considered to be a challenge of using digital systems. However, many believe that latency is an overly exaggerated issue in most applications where digital audio systems are deployed. In fact, this topic could be an article unto itself.

A little history of digital systems and networking will provide some insight on the reason why there are several networking technologies available today. In the late '90s, there were two "critical" concerns in the digital audio industry: Year of 2000 compliance (Y2K) and latency. To many audio pros, using audio networks like CobraNet seemed impossible because of the delay —at that time, approximately 5 milliseconds, or in video terms, less time than a frame of video.

Enter EtherSound, introduced in 2001, which addressed the issue of latency by providing an Ethernet networking scheme with low latency and better bit-depth and higher sampling rate than CobraNet. The market timing and concern over latency gave EtherSound an excellent entry point. But since reducing latency down to 124 microseconds limits available bandwidth for data traffic, a dedicated network is required for a 100-MB EtherSound network. Later, to meet the market demands of lower latency requirements, CobraNet introduced variable latency, with 1.33 milliseconds being the minimum. With the Ethernet technologies discussion thus far, there is a relationship between the bit-depth and sample rate to the clocking system.

Audio is not the only industry with a need for real-time clocking schemes. Communications, military, and industrial applications also require multiple devices to be connected together on a network and function in real-time. A group was formed from these markets, and they took on the issue of real-time clocking while leveraging the widely deployed Ethernet technology. The outcome was the IEEE 1588 standard for a real-time clocking system for Ethernet networks in 2002.

As a late entry to the networking party, Audinate's Dante comes to the market with the advantage of using new technologies like IEEE 1588 to solve many of the current challenges in networking audio. Using this clocking technology in Ethernet allows Dante to provide sample accurate timing and synchronization while achieving latency as low as 34 microseconds. Coming to the market later also has the benefit of Gigabit networking being widely supported, which provides the increased bandwidth requirement of ultra-low latency. It should be noted here that EtherSound does have a Gigabit version, and CobraNet does work on Gigabit infrastructure with added benefits but it is currently a Fast Ethernet technology.

Dante provides a flexible solution to many of the current tradeoffs that require one system on another due to design requirements of latency verses bandwidth, because Dante can support different latency, bit depth, and sample rates in the same system. For example, this allows a user to provide a low-latency, higher bandwidth assignment to in-ear monitoring while at the same time use a higher latency assignment in areas where latency is less of a concern (such as front of house), thereby reducing the overall network bandwidth requirement.

The developers of CobraNet and Dante are both working toward advancing software so that AV professionals and end-users can configure, route audio, and manage audio devices on a network. The goal is to make audio networks "plug-and-play" for those that don't want to know anything about networking technologies. One of the advances to note is called "device discovery," where the software finds all of the audio devices on the network so you don't have to configure them in advance. The software also has advance features for those who want to dive into the details of their audio system.

Advances in digital audio systems and networking technologies will continue to change to meet market applications and their specific requirements. Aviom's initial focus was to create a personal monitoring system, and it developed a digital audio transport to better serve this application. Aviom's low-latency transport provided a solution to the market that made it the perfect transport for many live applications. CobraNet provides the AV professional with a solution to integrate audio, video,

and data systems on an enterprise switched network. EtherSound came to the market by providing a low-latency audio transport using standard Ethernet 802.3 technology. Dante comes to the market after significant change and growth and Gigabit networking and new technologies like IEEE 1588 to solve many of challenges of using Ethernet in real-time systems.

Networking audio and video can seem chaotic, but gaining an understanding of the OSI model helps bring order to the chaos. It not only provides an understanding of the various types of technology, but it also provides a common language to communicate for both AV and data professionals. Keeping it simple by using the OSI model as the foundation and breaking audio networking down into two functional parts (control and transport) will help you determine which networking technology will best suit your particular application.

<i>Brent Harshbarger is the founder of m3tools located in Atlanta. He can be reached at **brent@m3tools.com.**</i>