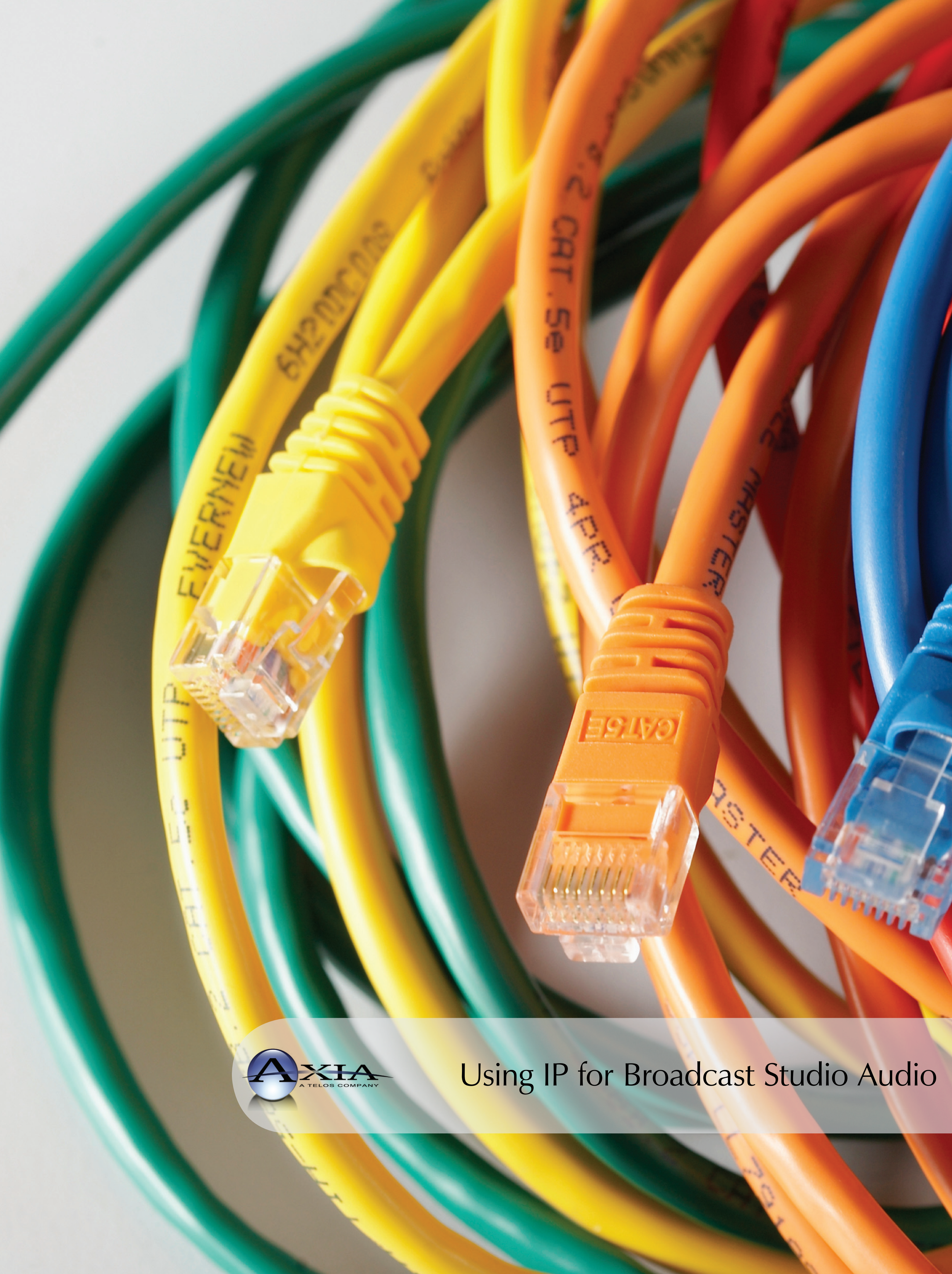

AXIA
A TELOS COMPANY

Using IP for Broadcast Studio Audio

Using IP for Broadcast Studio Audio

Skip Pizzi

Executive Summary

A revolution is currently taking place in the field of audio studio design. It involves a fundamental rethinking of the way signals are distributed and managed throughout the broadcast facility.

A new approach to this process utilizes a digital transmission format similar to that used on the Internet for purposes of transport and distribution of content around an audio production environment.

This format is known as the *Internet Protocol*, generally referred to as "IP," and it is the very core of the Internet. IP is the common format used for any kind of data that flows on the Internet (including streaming media), and on private extensions of the Internet, such as the LANs employed in enterprise networks and small office/home office (SOHO) networks, both wired and wireless. It sets the rules for this entire data networking infrastructure, both hardware and software, which has emerged and been so broadly embraced over the last quarter century or more. Since most audio facilities have already converted to digital audio, it therefore makes sense to now examine the use of IP there, as well.

Given that IP is the *lingua franca* of contemporary data networking, it can provide significant economies of scale for specialized applications such as professional digital audio distribution. This exploits the same process that has made the general-purpose desktop computer an efficient and cost-effective platform for the creation and storage of professional audio content. IP Audio distribution is simply an extension of that thinking and technology, replacing the purpose-built (and relatively expensive) mixers, routers and switchers that have traditionally been used by audio studios for managing multiple audio signals as they pass through a production or broadcast facility. IP also allows the full and continuing force of Moore's Law to be applied to audio distribution, just as the PC has done for other audio processes.

Beyond simple cost-effectiveness, however, IP Audio offers other important benefits. These include high scalability (i.e., the ability to easily accommodate growth and other configuration changes), convenience (i.e., easy and fast installation) and "future-proofing" (i.e., high likelihood of fitting well into any scenario for future facility requirements). Putting all these elements together creates a value proposition that is hard to ignore when considering options for new facility designs or existing studio upgrades.

Many other industries have already converted their legacy communications processes to IP-based replacement systems. Studio audio systems using IP-based technology are now sufficiently mature to allow audio producers and broadcasters to do the same, providing them with substantial savings while simultaneously positioning them well to accommodate indeterminate future needs.

Introduction

This paper considers the replacement of traditional forms of audio signal transport in the broadcast studio with networked audio carried via Internet Protocol (IP). It examines the value and process of such conversion, the challenges to doing so, and the likely future of this environment.

This study was commissioned by Axia Audio in early 2008.

A Brief History

The broadcast audio studio has a long legacy relationship with the telecommunications world. The earliest audio facilities and standard practices were developed by Bell System and Western Electric engineers in the early 20th century, and the two worlds have never strayed far from each other since.

In particular, broadcast audio has retained a close connection to the telecom environment, since so much of broadcasting's content comes and/or goes from the studio via telco-provided paths.

Thus it is not surprising that the next generation of studio audio technology should once again follow a path blazed by telecommunications technologies.

In this case, the technology involves digital networking – specifically using the popular and now nearly ubiquitous *Internet Protocol* (IP). This protocol, coupled with either the *Transmission Control Protocol* (TCP) or *Universal Datagram Protocol* (UDP) data transport format, provides the bulk of data communication carried on the Internet. Its popularity for that purpose has also made it useful for other localized, non-Internet data networking purposes. This is primarily due to the volume of development and resulting hardware and software that supports TCP/IP or UDP/IP, and the great economy of scale that results.

For a little background, note that a "protocol" in the data networking context is simply an agreement on the format of data that will be passed between devices. Therefore it specifies a set of rules for various parameters of that data, such as the data rates allowed, the error checking algorithms employed, any data compression formats that might be used, how the start and end of individual messages will be determined, how confirmation that a message has been successfully delivered will be communicated, and so on.

It's also helpful to review some of the early development that led to IP's particular popularity. In a nutshell, the IP approach is a simplification of the canonical seven-layer networking architecture[1] down to a stack consisting of only four layers, as shown in Table 1.

[1] The *Open Systems Interconnection* (OSI) reference model, established in the late 1970s, included Application, Presentation, Session, Transport, Network, Data Link, and Physical layers.

APPLICATION	HTTP, RTP, FTP, SMTP, TELNET
TRANSPORT	TCP, UDP
NETWORK	IP
LINK	Ethernet, WiFi

Table 1: The four layers of Internet data transmission, with some examples of each layer's protocols.

The Internet process also includes an addressing protocol for each of its data packets, the *IP address*. Any device attached to an IP network is assigned an IP address. Until recently – i.e., using IPv4 – an IP address was specified as a numeric string of four one-byte numbers (or *octets*, since one byte is eight bits), each expressed in decimal form (from 0 to 255) and separated by periods (e.g., 169.10.206.2). [2] This implies that the number of possible addresses in IPv4 is that expressed by a 32-bit number (4 x 8 bits), meaning that 2^{32}, or approximately 4.3 billion (4.3×10^9), unique IP addresses are available. This may sound like a lot, but many of these are reserved for specific uses (more on this below).

Today, the IP world is converting to IPv6, which specifies its IP addresses using 128-bit rather than 32-bit numbers. [3] The numerical expression of IPv6 addresses also differs from IPv4's, in that it generally uses hexadecimal numbers in the form hhhh:hhhh:hhhh:hhhh:hhhh:hhhh:hhhh:hhhh, where each byte (or octet) is represented by a hexadecimal pair of numbers (from 00 to ff, e.g., e7), and each pair of bytes is separated from the next pair by a colon. An example is 30c1:0ab6:0000:0000:0000:8a2e:0370:2f8e. For awhile, there may be a lot of zeros in IPv6 addresses, and they can be skipped with the insertion of a double-colon, as in this notation of the previous address: 30c1:0ab6::8a2e:0370:2f8e.

One of the primary improvements of IPv6 over IPv4 is its allowance of far more IP addresses. This is a real issue given the expectation that so many devices in the future will require unique IP addresses. IPv6's 128-bit range provides more than 3.4×10^{38} possible addresses, or more than 5,000 addresses per square micrometer of the Earth's surface – probably enough to last for awhile.

Nevertheless, it is expected that IPv4 will remain the standard format of the Internet for some time to come, while IPv6 support is gradually deployed worldwide.

The reason that IP addresses are important to this discussion is because they essentially replace the audio crosspoints in a traditional, circuit-switched environment. In an IP Audio system, all traffic flows along a single, serialized path, and each packet of data gets to its intended destination via the IP address in its header, rather than by its being sent along a dedicated wire by a switcher.

[2] IPv4 has been in use since 1981, established with the publication by DARPA of the seminal RFC 791 document, generally cited as the original specification for the Internet. Although other protocols preceded it, for most of us, IPv4 is the only version of IP the Internet has ever used.

[3] If you're wondering what happened to IPv5, it was ascribed to a version that was originally intended to be used for connection-based (rather than packet-based) streaming media on the Internet, but work was abandoned on it as streaming media became possible with the development of new protocols over IPv4.

IP Beyond the Internet

As noted earlier, IP is used today on many local area networks that are not connected to the Internet. This is why there are a large number of IP addresses that are reserved for non-Internet uses on private networks.[4] A number of IP address ranges are internationally agreed to be reserved for this purpose, the largest contiguous group of which spans from 10.0.0.0 to 10.255.255.255. This group alone provides some 16 million possible addresses that are not accessible from the Internet (routers are programmed to ignore the addresses on incoming Internet traffic), and are only available from within a local network. This also implies that a private IP address has no need to be globally unique, and so these same addresses can be used by any entity on its internal network, thereby conserving the number of IP addresses required worldwide.

Devices assigned such private addresses can still access the Internet if necessary, via a *proxy server* or *Network Address Translation* (NAT) device. [5]

The private address space is particularly useful for studio audio applications of IP, since the devices so interconnected are typically not intended to be accessible directly via the public Internet.

"IP but not Internet" is also the case for another major emerging technology called *IPTV*, which uses IP for distribution of television programming, but over dedicated networks operated by telcos (providing a multichannel service competitive with digital cable and satellite TV), not the open Internet.[6]

Streaming Changes All

Although originally developed for standard data communication, subsequent enhancements to the Internet allowed it to be used for media transmission, as well, which is well known as the process called *streaming media*.

This development fundamentally altered the usage of the Internet, and has subsequently had significant impact on all facets of the media industry, as they struggle to cope with the changes it brings, and to take advantage of the new opportunities it engenders.

Besides spawning many currently burgeoning on-line media businesses, streaming technology is also the basis for IP studio audio. This allows the audio studio environment to leverage several key properties of an IP-based environment, which provide substantial improvement over more traditional approaches:

---

[4] These networks use addresses in the private IP address space, as specified in the Internet Engineering Task Force's (IETF) RFC 1918, and administered by the Internet Assigned Numbers Authority (IANA).

[5] Note that in IPv6 there will be no private address space or NAT, given the far greater number of globally unique IP addresses it provides.

[6] Meanwhile, TV content sent via IP that *does* travel via the open Internet ("Internet TV") is also soaring in popularity, increasingly used for distribution of broadcast content as well as content created by consumers (*user-generated content*, or "UGC").

- **Scalability:** Perhaps the most fundamental change between IP-based audio systems and traditional approaches – analog or digital – is the ability of IP architectures to adapt to change and growth. For example, a traditional audio environment must have its spatial or imaging format (e.g., mono, stereo or surround) predetermined, along with the number of simultaneous audio channels it will requires (e.g., one, two or more). An IP Audio environment has no such requirement, and can easily adapt to any format. Similarly, a traditional "crosspoint" audio routing switcher must have its input and output (I/O) configuration fixed in its hardware design. In this way, such a device reflects *circuit switching* and parallel design, whereas IP Audio systems implement *packet switching* and serial design. This allows great flexibility and responsiveness in accommodating changes in I/O configuration. Just as telcos have moved away from the circuit-switched paths of their earlier years for similar reasons, studio audio systems can now enjoy the same advantages of scalability and flexibility to implement expansion in any dimension. This comes not a moment too soon, given the competitive pressures coming to bear on broadcasters for increased content and listener choice.

- **Cost-effectiveness:** At almost any reasonable size, an IP-based audio system will compare favorably with the cost of a traditional system – both in terms of hardware/materials pricing and installation cost. The reduction in wire alone provides substantial economy.[7] Maintenance expenses are generally also lower. These cost differentials increase with the size of the facility, which is why so many larger installations have already moved to IP-based solutions as their needs have called for new technical plants.

- **Convenience:** The small physical footprint, low operating cost, ease of reconfiguration or upgrade, and fast installation of IP Audio systems make them extremely convenient for engineering and operations alike at the audio studio facility. From initial design to implementation to daily operation, IP-based systems make life easier.

- **Future-proofing:** Nothing strikes fear in the heart of the engineer or manager more than making a poor major purchasing decision. Moving to an IP-based audio architecture takes a lot of the pressure off, since it offers such flexibility and allows broad ability for reconfiguration down the road. Provisioning for unforeseen changes is much less problematic and cheaper.

Note that the above advantages only fully apply to systems that utilize standard IP in their architecture. Not all audio systems that use computer networking (over Ethernet and/or on RJ45 connectors) for interconnection are necessarily true IP-based systems. Some simply use Ethernet as a physical layer with a proprietary data format above it,[8] while others may use more IP-like formats but with non-standard

[7] Remember that a packet-switched system like IP does not require individual wiring paths to each input and output of every device. For example, an audio mixing console or multitrack recording device can have all of its inputs and outputs interfaced to the rest of the facility via a single cable in an IP Audio environment.
[8] e.g., *Cobranet*

protocol variations. Some of these non-standard approaches may have offered some value in the past (such as reduced overhead and latency over standard IP networking), but given the capacity, speed and performance of a properly configured, standard IP system today, the penalties paid by working in a non-standard environment generally far outweigh any advantages that such variations might provide, particularly when considered over the long term.

IP-Anything

In terms of critical mass behind this trend, studio audio is certainly not alone in its movement toward IP-based processes, as already noted. Numerous other industries have already embraced a transition to IP for similar reasons to those cited above for studio audio.

One such development that is closely related is *Voice-over-IP* (VoIP), which is rapidly gaining ground in the telephony space as a replacement for traditional voice service, in both consumer and enterprise applications. Again the leveraging of IP as a mechanism to use *generalized* systems and transport paths for various *specific* tasks has undeniable appeal, and this argument is also finding favor in a wide range of other industries, from hotel TV systems to health care. Emerging digital TV systems (including new mobile variants) are also favoring an IP distribution model.

IP is fast becoming the *lingua franca* of digital technology and content, allowing anything expressed in its terms to be carried and processed through increasingly available and cost-effective infrastructures. Just as the PC became the general purpose computing platform (delivering unprecedented processing power, speed and cost-effectiveness), IP has become the general purpose data transport format.

For engineers, familiarity with digital networking technologies, including IP, has become a near-requirement of the job anyway (for implementing the on-line services of a radio station), so why not apply this knowledge to studio audio, too?

It's becoming clear that IP is truly the way of the digital media world, particularly for any industry that values connectedness, agility and cost effectiveness. In the radio environment, it's not an overstatement to say that IP Audio is the future of studio audio signal flow. Arguing otherwise is difficult: There is and will continue to be so much development within the IP environment, it only makes sense to harness the power of that effort, while also letting Moore's Law have its full effect on the hardware side for studio audio equipment, just as these forces' effects are being enjoyed by so many other industries today.

What's the catch?

This is not to say that there aren't some challenges to the optimal use of IP for studio audio transport. Primary among these is the latency that the encapsulation process of audio data into IP packets can cause, and their serial routing through a packet-switched network prone to data collisions. As mentioned earlier, various methods have been put forth by developers to ameliorate this, but proper configuration of available devices used on today's high-speed Ethernet networks is usually adequate to resolve any such difficulties in IP Audio systems.

Such configuration of standard IP equipment (e.g., Ethernet routers and switches, buffer sizes and network speeds) can be set to optimally serve the specific needs of a studio audio system, rather than generalized Internet data traffic. For this reason – as well as for obvious security purposes – it is important to use a separate, dedicated network (either physical or virtual) for all studio audio IP applications. This network can carry all audio content, control signals and metadata related to production, but should be isolated from the general data network of the facility. In addition, there generally should not be a direct connection of the studio audio system to the public Internet. When Internet connections are required for access to off-site audio sources or destinations, they should be routed through a proxy server or other isolating path. Some vendors will also set IP packet prioritization at a higher level for audio content packets than for general network data. This helps IP Audio performance even on networks that are not dedicated to IP Audio usage only.

Another issue is a simple one of connector standards. Since IP Audio generally travels on CAT5 or better[9] Ethernet cables, the RJ45 connector is used for all terminations. But some IP Audio system implementers also use RJ45 for analog or AES3 digital audio I/O and patching. While this can minimize the number of different connector types used in a facility, and reduce the physical space required for connector panels, there is no standard for the configuration of an RJ45 connector for this purpose. So individual IP Audio implementers have designed their own unilateral formats, again possibly limiting broadcasters to the use of only certain manufacturers' products in an IP Audio facility, or requiring the use of adapters.

Of course, the need to retain compatibility with analog and AES3 digital audio will likely remain for some time to come at any IP-based audio facility. At the very least, live microphone signals will need to be converted from their native analog audio (or in some cases, perhaps AES3) format, and in some cases analog or digital audio from other devices or remote sources will also have to be accommodated. How this conversion process is accomplished is a key factor in the design of an IP Audio system.

For the time being at least, the above optimization processes argue for the selection of a single vendor for the supply of IP Audio equipment to a given broadcast facility, or at least the verification that compatibility among different vendors' IP Audio equipment is assured. Working with a single supplier also ensures that updates and upgrades will be delivered in a timely fashion, and even special fixes for a particular problem encountered at a given facility can be quickly provided.

[9] CAT5e or CAT6

Implementation and integration

This brings us to actual IP Audio system considerations. Given the advantages of scale provided by these systems, it makes sense to make the IP Audio domain as large as possible within a given facility. That implies that audio signals in other forms should be converted to IP packets as soon as possible.

The best place to do this in most studio configurations is at the central patch bay and/or the studio mixing console(s). Microphone outputs and signals from other “legacy” audio sources can be immediately converted to digital audio form (if they aren’t already) and packetized as IP. Once in the IP domain, these signals can be addressed and routed to any other location on the network. This can include destinations within the confines of a facility via LAN, or anywhere in the world via WAN.

Another advantage of this approach is that a mixing console can act as a router. In other words, because any input on the console can have a unique IP address, it can be connected to any IP Audio source on the network. (Even more amazing to veteran audio engineers is that this can be accomplished even though the entire console is connected to the network via a single cable.) A central switching control unit (typically a PC) can assign these I/O connections, or the mixing console itself can have a control interface for this purpose. In addition, standalone hardware switch controllers can be distributed around the facility, essentially duplicating the appearance and function of traditional router-control panels. In all cases, however, all that is being done to achieve the routing is the addition of appropriate IP addresses to the headers of IP Audio packets on the network.[10]

Of course, the mixing console can also be equipped with traditional analog (mic/line) or AES3 inputs on some inputs as well, to accommodate local audio source devices. These sources are in turn converted to IP by the console, and via their routing to the console bus or direct outputs, these signals may in turn be made available as IP sources to the rest of the network – again via the single cable connecting the console to the network. (See Figure 1.)

Consider also how PC-based audio playout/automation systems can be interfaced to such a system. Rather than their audio outputs being directed through PC sound cards to traditional audio inputs, the automation system can be fitted with an IP driver that provides a software interface between the PC audio and the IP network directly in the IP Audio domain. This not only maintains high audio quality, but cuts costs in the automation system since no (or at least fewer) sound cards are required. The IP interface can also carry control data and PAD as well, eliminating the need for separate data links between devices.

[10] Ideally, the IP feeds to/from any mixing console in an IP Audio system are used only for those items that need access to the console’s audio manipulation features. Simple input-to-output connections can be handled by the core IP Audio architecture without taking up console resources (or even using up a crosspoint, in traditional switching terms). The console can be considered as an IP appliance or “audio engine,” which can host any sort of audio level-adjustment, mixing, processing or coding/transcoding that the facility requires, applying all such processes efficiently and within the IP domain.

Moreover, a *single* IP driver interface between an automation system and an IP routing architecture can carry many independent audio channels (perhaps up to 24), whereas a traditional switching system would require a crosspoint (plus wiring) for each sound card input and output. The combined hardware savings (soundcards + crosspoints + wire + installation) in a large facility could be substantial.

Any IP Audio system today will also include one or more standard Ethernet *switches*, which are used to interconnect the Ethernet connections from each facility or device that provides audio I/O via IP. These switches are standard telecom devices, which replace the more common Ethernet *hubs* or *routers* used for similar RJ45 terminations in typical home or small office situations. Switches are preferred in IP Audio applications because they have more intelligence than hubs, which allows them to inspect data packets to determine their source and destination, and forward them appropriately. Through this intelligent control, each message is only sent to the intended device, thereby conserving network bandwidth and providing better performance (including reduced latency).

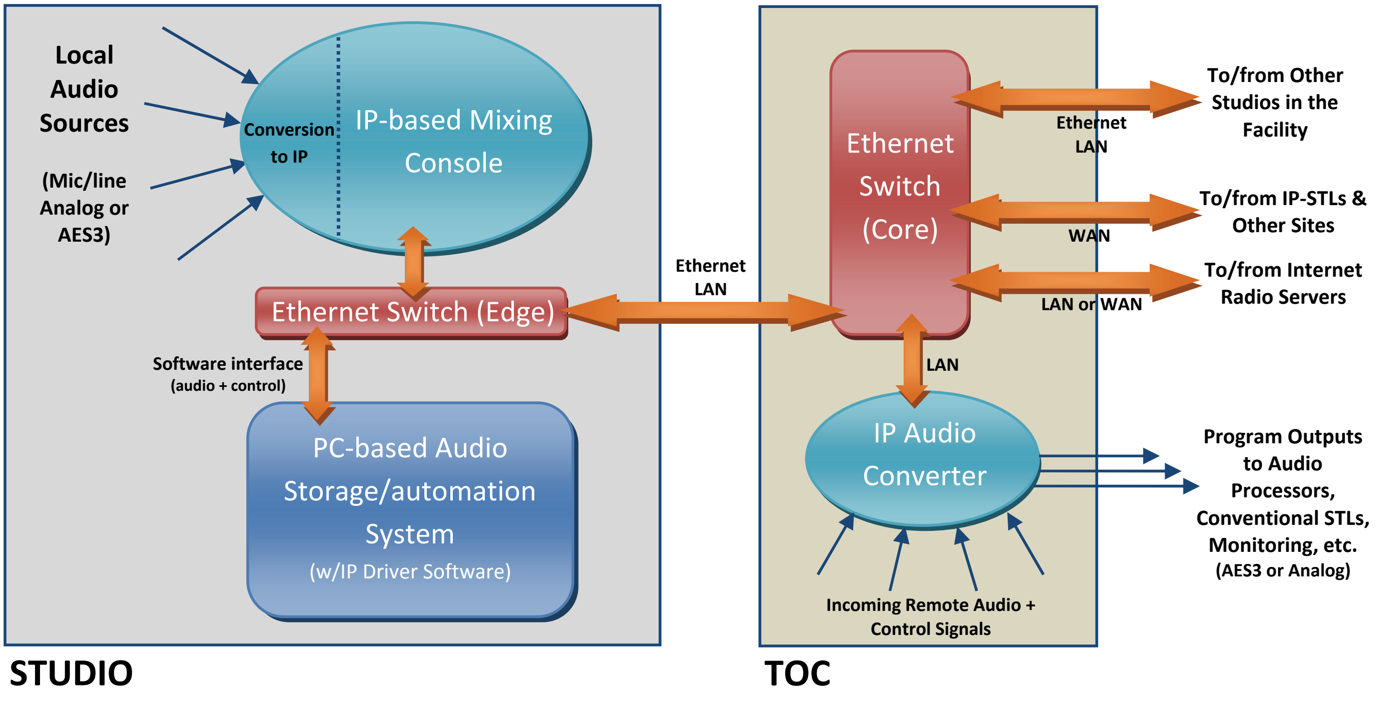

Figure 1: **Conceptual block diagram of a typical IP Audio-based broadcast studio facility, showing one studio and Technical Operations Center (TOC).**

Ethernet switches come in *unmanaged* or *managed* forms, the latter allowing user configuration of various parameters of the switch's operation. In smaller IP Audio facilities, unmanaged switches can perform adequately, while larger facilities may benefit from managed switches. Some switches also include remote administration capability. Today's fastest switches operate at Gigabit speeds, using CAT6 cable, and these are the most appropriate for low-latency IP Audio applications.

As Figure 1 indicates, a typical IP Audio facility includes *multiple* switches, usually arranged with one large ("core") switch in a central room, and smaller ("edge") switches placed as needed in other rooms around the facility. Such distributed routing intelligence improves performance, and also provides redundancy in case of switch failure.

Note also that the proliferation of VoIP and other realtime applications via IP have spawned broad implementation of *non-blocking* architecture in recent Ethernet switches. This approach virtually eliminates data collisions within a switch, while maintaining cost-effectiveness, by ensuring just-adequate capacity for $n \times n$ connectivity[11] through the switching fabric. Mission-critical performance is thereby maintained by using Ethernet switches that implement non-blocking design, and when properly implemented within an IP Audio system, switch capacity should never be exceeded. Most IP Audio vendors recommend only non-blocking Ethernet switches, and this is one reason why users should heed vendors' recommendations for all switches used in their facilities. (Again, the economy of scale driven by broad uptake of these switches today provides such advantages to IP Audio implementers at very attractive prices.)

The use of Ethernet switches by mission-critical and other high-reliability telecom operations has driven major manufacturers to provide excellent 24x7 and overnight-replacement support. Note also that as a facility grows, it may need to replace older switches with newer models; the fact that Ethernet is a ubiquitous standard means that all upgrades will remain backward compatible. Meanwhile, Moore's Law ensures that as such new hardware becomes available, price/performance ratios will continually improve.

The IP Audio domain is also extending beyond the studio. Figure 1 shows how IP Audio is converted back to AES3 (or even analog) for program outputs' connection to conventional Studio-to-Transmitter Links (STLs), but the diagram also indicates that an STL could carry IP Audio to the transmitter site (via WAN or other dedicated link). Whether leased from telco or using a station-operated RF path, if adequate bandwidth is available, multiple audio channels, control and metadata can all be carried via IP on the link – bidirectionally, if desired – with minimal latency.[12]

IP Audio in Use Today

The advantages of IP Audio have been noticed by broadcasters and studio owners around the world. It is fair to say that the engineers designing every new broadcast studio facility built today (and from this point forward) are at least considering the use of an IP Audio architecture – and many of them are deciding to take the plunge. Speaking with them afterwards will find almost unanimous agreement that

[11] In other words, any input on the switch can always be connected to any output on the switch, under any usage.
[12] Bidirectional links can substantially decrease jitter and latency compared with unidirectional IP paths.

it was the proper choice, and there's no looking back. In many cases you will hear that the transition process was far easier than they expected.

The installation of an IP Audio system makes many people at the operation happy, from the CE to the CFO. The total cost of building and operating an IP Audio facility is significantly reduced, and yet this can be accomplished without giving up flexibility – in fact it, too, is greatly increased. Operations are often minimally interrupted, as well, due to the small footprint and quick installation of IP Audio systems.

This is why broadcasters of all stripes, and with budgets large and small, have moved to IP Audio systems. On the commercial side in the U.S., broadcasters from Clear Channel to Greater Media have recently installed IP Audio systems, while non-commercial operators from Minnesota Public Radio to WYMS, Milwaukee have done the same.

In fact, the clientele for this emerging technology almost defies characterization. It includes small independent stations, college radio (including numerous rural and community colleges), ethnic and religious broadcasters, satellite radio services, radio and telecom network operators, content production and broadcast origination sites, corporate facilities and government agencies – along with some of the largest and most respected stations in the country.

Neither is adoption is by any means limited to the U.S. IP Audio is already well ensconced around the world, from Spain to South Africa, from Italy to Israel, from China to the Czech Republic, and many other places in between.

Clearly the technology has a lot of unique advantages to offer, and these have been noticed by many of the leading purveyors of audio content and delivery throughout the planet.

Conclusions

It's not often that a new technology offers considerable technical improvement, easier installation and maintenance, greatly enhanced flexibility and scalability, *and* reduced cost when compared with its predecessors. Yet these are the attributes of a properly implemented IP Audio system.

Broadcasters have always been a cost-conscious lot, and rightly so, but given today's increasingly competitive landscape, efficiencies in capital expenditures and operating costs have become even more critical and desirable.

Meanwhile, it's become quite clear that the radio industry will face substantial change in the near future, and much of it will likely involve quantitative growth in services. More streams, more audio channels, more data, more responsiveness to audience demands, and probably more still, are all on the path that lies ahead for broadcasters. An IP Audio platform provides a legitimate platform to most realistically accommodate these many challenges.

It behooves any broadcaster faced with the opportunity to confront these uncertainties to seek the counsel of experts in the field of IP Audio, and to benefit from the experience of those who have recently undertaken similar projects. You will likely hear strong encouragement to make the move to IP, just as much of the rest of the world is doing, with good reason.

Fairfax, VA
February 2008