

Recent Advances in Audio/Video Networking

Fengkun Yang (A paper written under the guidance of [Prof. Raj Jain](#))

Abstract

Audio and video technology based on Internet protocols (AV over IP) based on network protocol for audio and video signal transmission is being adopted by more and more audio-visual industry customers and even radio and television producers. New retail, enterprise collaboration, hospitals, security monitoring, transportation, radio, and television production, and other professional fields gradually open up the use of network-based audio and video transmission programs. AV over IP technology has a long history, but in recent years, from the audio-visual industry to the radio and television industry, the IP distributed projects are favored and gradually form a "hybrid" workflow.

In the past decade, consumer demand for multimedia applications on Ethernet has been increasing. The traditional Ethernet can not fundamentally meet the voice, multimedia, and other dynamic content and other real-time data transmission needs. So Ethernet audio and video bridging technology (AVB) came into being.

Keywords

AV, AV over IP, AVB, BMCA, Stream Reservation Protocol, Multi-Registration Protocol, Audio Video Transport Protocol, Connection Management, and Control Protocol

Table of Contents

- [1. Introduction 3](#)
- [1.1. AV over IP definition 3](#)
- [1.2. AVB historical reasons 3](#)
- [1.3. Summary of the introduction 4](#)
- [2. AV over IP 4](#)
- [2.1. Current AV technology 4](#)
- [2.2. The same and different points of traditional AV and AV over IP 5](#)
- [2.3. The game of image quality, bandwidth, and latency 8](#)
- [2.4. Selection of standards 8](#)
- [2.5. Security of the AV over IP 9](#)
- [2.6. Diversity of AV over IP 10](#)
- [2.7. AV over IP summary 11](#)

[3. AVB 11](#)

[3.1. The main problem facing audio and video transmission 11](#)

[3.2. Solution 12](#)

[3.3. AVB system 13](#)

[3.4. AVB Summary 19](#)

[4. Summary 20](#)

[5. List of Acronyms 20](#)

[6. References 20](#)

1.Introduction

AV, which originally means Audio and Video, is the two basic internal elements that make up non-abstract media such as film, TV, and video. All movies, TV, and even live media can be called AV because they have one thing in common: sound and synchronized images. This article focuses on audio and video technology based on Internet protocols (AV over IP) and Ethernet audio and video bridging technology (AVB). [\[baidu21\]](#)

1.1.AV over IP definition

AV over IP is an audio-visual protocol that transmits audio-visual data over LAN, WAN, or the Internet. It uses standard network devices to transmit and switch between video and audio, contrary to the traditional AV environment. [\[Matrox21\]](#)

It involves areas such as live video on the Internet, video on demand, professional video in broadcast studios, and so on. Over the past few years, the professional AV community has been gradually replacing the traditional AV infrastructure with IP-based infrastructure. This means that the technology has developed to the present day, the input and output signals are all the time, everywhere. The source can be "any" and "many", including desktop PC output, video cameras, media players, satellite/cable set-top boxes, and more. Terminals can also be "any" and "many", including receiver boxes that connect directly to the TV in Huddle Room or public spaces, video wall controllers in the control room, media streams on laptops, and even software-only decoding and display on mobile devices. [\[Marta19\]](#)

1.2.AVB historical reasons

Traditional audio and video (AV) device configurations were once single-purpose point-to-point one-way links. This dedicated connection mode results in the need for extensive cabling in professional and user applications that are difficult to manage and operate. There are several recognized mechanisms to address this issue. But all of these mechanisms are either unreadable, difficult to operate and configure, or they are inflexible and expensive. Moving to an Ethernet facility is recognized as a way to address the needs of professional AV equipment. However, at

Recent Advances in Audio/Video Networking

present, the configuration mechanism lacks flexibility and interoperability. To speed up the use of Ethernet-based AV configurations that can be operated together, IEEE has developed the IEEE802.1 Audio and Video Bridge (AVB) standard. [\[Cisco21\]](#)

1.3. Summary of the introduction

For one thing, simply put, the so-called "AV over IP" is the transmission of uncompressed audio and video information on a standard IP medium. A little more complicated is to encode the source of the signal, the compression of the reduced bit rate on the IP medium transmission. But in general, it refers to extending and switching video and audio sources on a standard IP network.

Also, since Ethernet was originally designed to handle purely static non-real-time data and ensure its reliability, sequencing and packet latency are not important considerations. Although the traditional two-tier network has introduced the priority mechanism, the three-tier network has the built-in quality of service (QoS) mechanism, but because of the multimedia real-time traffic and ordinary asynchronous TCP traffic there is resource competition, resulting in too many delays (Delay) and Jitter so that traditional Ethernet could not meet the demand, and IEEE developed the Ethernet Audio/Video Bridging technology. [\[IETimes07\]](#)

2. AV over IP

AV over IP technology based on network protocol over audio and video signal transmission is being adopted by more and more industries.

2.1. Current AV technology

AV products and services are designed with the following important performance: 1. Maintain maximum sound quality and images as you move and switch sources. 2. Quickly switch the entire pipeline. 3. Low latency. 4. Perform processing operations, such as multiplying the sources and making them available in multiple places at the same time. 5. Modify the source in real-time. These include converting video signals from one type to another (such as DisplayPort to HDMI), cropping, scaling up or down (such as converting from HD to 4K or 4K to HD), compositing (text overlaying or combining multiple videos), and so on. [\[Matrox21\]](#)

Internet Protocol (IP) is the rule of management of data sent over the Internet or other video and audio networks. Data transmitted through IP is broken down into packets. Each package contains a portion of the original file and additional control information, such as source, destination, and sequence. AV-over-IP technology organizes audio-visual data to comply with these rules. [\[Matrox21\]](#)

2.2. The same and different points of traditional AV and AV over IP

AV over IP technology is developed from the traditional AV technology, both have things in common, but also differences.

2.2.1. Same points

As shown in Figure 1, the main elements and architectures of both AV modes are similar, except that in AV over IP, the "A/V transmitters" becomes the "encoders", the "A/V receivers" becomes the "decoders", and the "A/V switcher" becomes a standard "IP switch". [Matrox21]

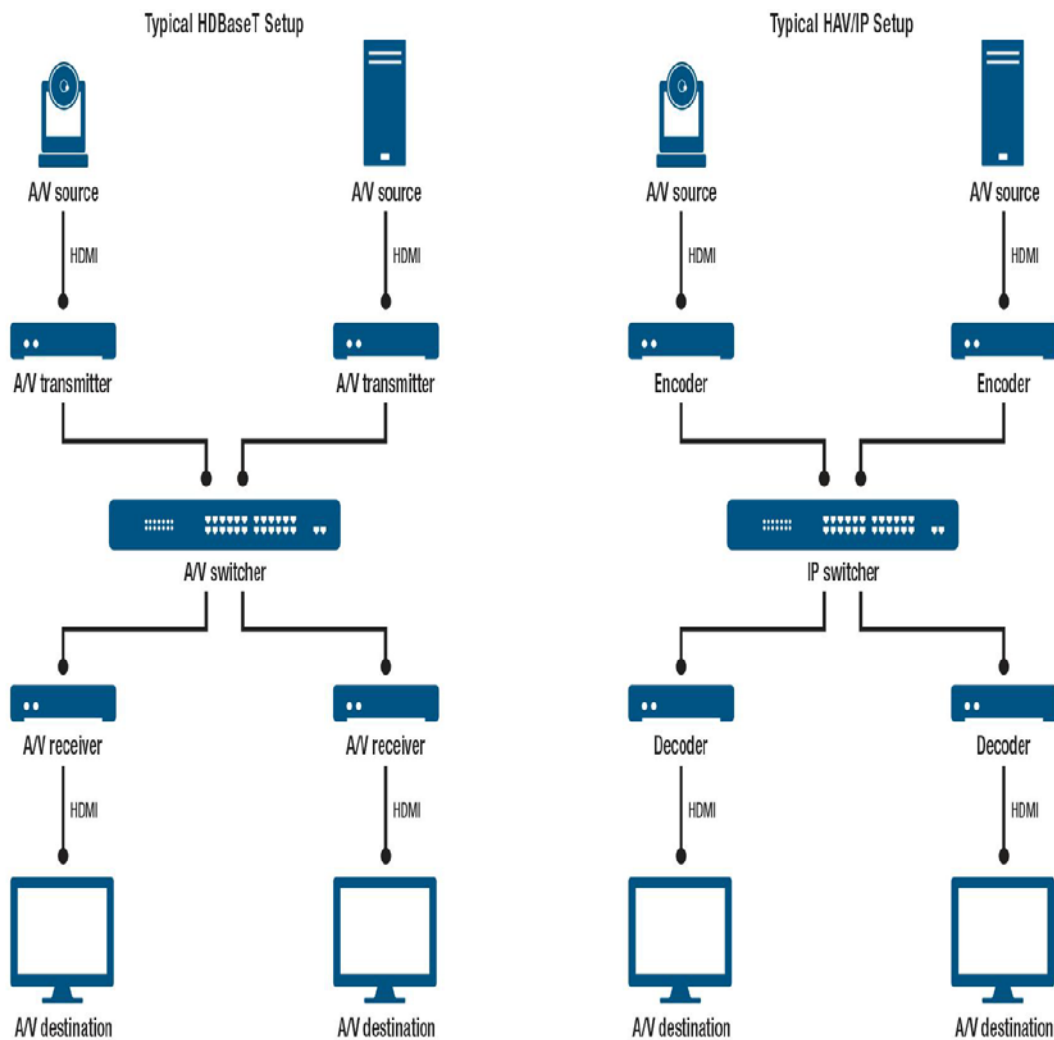


Figure1 Traditional AV and AV over IP architectures [Matrox21]

2.2.2. Different points

The differences can be divided into four:

1. More ports, scalable switching.
 2. Break the distance limit and increase the ratio of input to output.
 3. Extend video standards beyond local facilities.
 4. Converged data and communications. [\[Matrox21\]](#)
-

2.2.2.1. Switching

First of all, hard-wired, circuit-based switching is point-to-point technology. The video matrix switch is a destination and a source that occurs at the same time. All combinations from transmitter to receiver are resolved inside the matrix switch, and depending on the number of ports available on the video matrix switch, any source can be used at any destination. For example, an 8x8 matrix switch allows eight sources to be used in any of the eight destinations. [\[Matrox21\]](#)

Second, there is the ability to process operations, you can display any input on any output or multiple outputs. For example, the source is PPT on a PC, routed from the AV transmitter box to the video matrix switcher, and then to multiple AV receiver boxes, enabling real-time display of PPT. [\[Matrox21\]](#)

Third, the number of sources connected to IP switches is unlimited. When the physical port is exhausted, it can be expanded by connecting multiple IP switches, and the number of ports can be expanded as needed to continue adding sources and destinations. No major changes to the center of the video matrix switcher are required. [\[Matrox21\]](#)

Finally, the input-output ratio of the AV over IP can be customized as well. There may be multiple inputs or outputs, or there may be many, but the quantities vary widely. [\[Matrox21\]](#)

2.2.2.2. Distance

Traditional AVs are limited by the actual distance between boxes. Short-distance wiring, which is only a few feet long, is cheap, and when the distance becomes a few meters high, costs start to rise. Installation and expansion can be extremely expensive when wires are placed in hundreds of meters of facilities. The first one in Figure 2 is traditional AV switching. [\[Matrox21\]](#)

From traditional AVs to AV over IPs, there is no impact. AV over IP uses a packet-switching technology for conversion and series repeat and can be easily transmitted over copper cables and fiber optics. For example, YouTube videos watched at home are spread over considerable distances. It overcomes the limits of the number of sources and destinations, as well as the limitations of distance, for increased flexibility. The second and the third in Figure 2 are AV over IP. [\[Matrox21\]](#)

Recent Advances in Audio/Video Networking

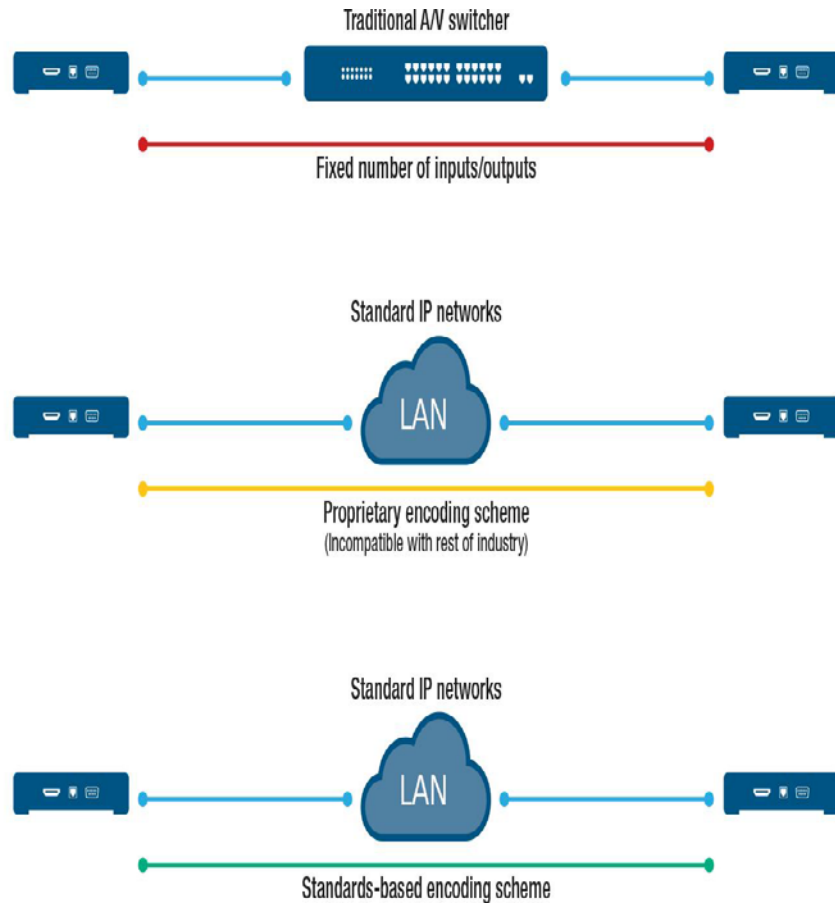


Figure 2 Traditional AV switching and AV over IP [\[Matrox21\]](#)

2.2.2.3. Standard

Standards are balanced from many different technology perspectives and provide greater infrastructure migration benefits than single-vendor-based products. Using standards-based packet AV-over-IP products to transmit over IP networks and to be compatible with IP switches allows products from different vendors to operate with each other. [\[Matrox21\]](#)

2.2.2.4. Interoperability

Grouping scenarios determine interoperability. There are three different possibilities.

Firstly, the tightly coupled AV-over-IP encoder and decoder are from the same vendor, providing customers with guaranteed specifications and performance, as well as a very controlled out-of-the-box setup and ease of use experience. [\[Matrox21\]](#)

Secondly, the vendor provides fully open compatibility, either by producing encoders that work

Recent Advances in Audio/Video Networking

with other vendors' decoders, or by producing decoders that work with other vendors' encoders, or both. These products place great emphasis on interoperability and the ability to leverage the features and capabilities of a large number of hardware and software vendors. [\[Matrox21\]](#)

Finally, vendors provide products that support both tight coupling and interoperability. Provides all the features and benefits of tight coupling, while making good use of features and functionality from a wide range of third-party hardware and software vendors. [\[Matrox21\]](#)

2.2.3. Summary of the similarities and differences between traditional AV and AV over IP

The main elements and structures of traditional AV and AV over IP are similar. The main differences are that AV over IP has scalable switches, no distance limitations, a wider range of video standards involved, and greater interoperability.

2.3. The game of image quality, bandwidth, and latency

The AV industry has found the next best way to transmit audio and video signals. IP provides a standardized architecture, but today, the way it is delivered on the architecture is no longer sufficient for a variety of customer needs, and we need to determine what image quality, bandwidth, and latency are allowed based on the application.

For low-bandwidth applications, H.265 video coding standard is the answer, but due to high levels of compression, there is a cost of latency and image quality. If enough bandwidth is available, choose low-latency, loss-free compression methods such as JPEG 2000, TICO, and others that provide higher-quality images at lower latency levels.

When bandwidth is not a problem, it is important without latency and compression, especially in broadcasting, video conferencing, and so on. Such applications can now be implemented in the 10G network topology. Using standard IP media and standard IP protocol transmission, audio and video information can come and go like computer data, enjoy the same management tools, stream access rights, and security protocols. However, most vendors that use standard IP media compress the source audio and video information in some form, so that the data network can be properly balanced. [\[LAWO21\]](#)

The lighter the AV over IP packet, the faster the transmission. In addition to compression, you can also look forward to purchasing unlimited bandwidth for dedicated signals, enjoying the freedom of the "virtual signal path" provided by the AV over IP solution, without being limited by distance, and enabling full HD video and minimized latency. Unlike traditional transport methods, AV over IP solutions provide dynamic and real-time content, creating more engagement and lowering the cost of ownership for users. Especially in the business environment, IP is the ideal way to transmit signals, which is the ultimate trend of AV distribution.

2.4. Selection of standards

The choice of AV over IP technology also depends on the user's tolerance for compression. If you don't want to compress video, you can use standards like HDBaseT or HDMI, which are designed to handle these high bandwidths. Understanding the main compression techniques and making trade-offs is key to a successful design if you want to transmit AV over IP signals with lower latency over long distances. There are two main compression technologies in the professional audiovisual field, H.264 and JPEG2000, both of which enable high-quality images, except that the H.264 optimizes bandwidth and JPEG 2000 optimizes latency. [\[Earon21\]](#)

Low-bandwidth H.264 traffic performs well in most network environments, allowing multiple traffic streams on each link and interoperability on the corporate network. For example, most web-based video uses H.264 compression. It is also commonly used by streaming cameras and mobile devices, and H.264 streaming solutions in the professional AV market can interact frequently with these devices. The price is latency, which can easily add half a second to the video stream. This also makes H.264 more used in long-distance applications but does not apply to room extensions, where the delay of the mouse keyboard is abrupt. [\[Matrox21\]](#)

The JPEG 2000 solution means higher bandwidth, almost 1Gb of Ethernet links. As a result, they often require specialized network cables to run, much like the standard non-compression technologies of HDMI and HDBaseT. But the advantage of the JPEG 2000 is that the latency is lower than H.264. [\[Earon21\]](#)

Delays are superimposed, so consider delays from other devices, such as display multipliers and video conferencing systems. At the same time, the synchronization of audio and video is also important, people are often sensitive to out-of-sync sound painting. Even small video delays can become a big problem if you use a single audio distribution system. [\[Wikipedia20\]](#)

2.5. Security of the AV over IP

A common misconception is that AV over IP presents more security risks than traditional AV. This is not correct. The following two methods are applied to AV over IP protection, which can guarantee its security. [\[Matrox21\]](#)

2.5.1. Digital content protection

Some AV over IP products support third-party devices that use digital key exchange or encryption. For example, high-bandwidth digital content protection (HDCP) that protects digital copyright content that propagates between devices limits the ability to extend, copy, modify, or view protected content. When a cable or satellite receiver box or media player with HDMI output

Recent Advances in Audio/Video Networking

plays HD or 4K content of protected content, it is locked and can only be viewed by properly certified HDCP-compatible products. [\[Matrox21\]](#)

2.5.2. Encryption

Encryption reduces openness, locks encoder and decoder compatibility, and provides a good user experience to customers by limiting the scope of their quality assurance testing, technical support, and customer care options only by using its brand. [\[Matrox21\]](#)

2.5.3. Summary of the security of AV over IP

Digital content protection and Encryption protection could guarantee the security of AV over IP.

2.6. Diversity of AV over IP

The latest IP-based AV products have demonstrated a wealth of versatility. In addition to mobile control signals, distributing audio, and managing video conferencing, many new solutions add more and more differentiation to network applications. From microphones to PTZ cameras, hardware can not only be connected, but they can also run on multiple platforms, collect more audio and video data, and enable further collaboration from desktops to coffee shops. [\[Sohu16\]](#)

As we can see, IP has created smarter microphones, and it's not hard to build an end-to-end network audio system now, and a Dante-enabled microphone doesn't require a separate audio interface. For example, the smallpox-mounted microphones in the Delta, the ATND931 six-inch gooseneck, and the ATND933 suspension microphones, all of which use standard ATND8734 microphone smallpox-mounted power modules with Dante interfaces, can be powered by PoE, and the RJ45 connectors of these microphones provide plug-and-play access to Dante audio networks on the ultra-five lines. [\[Sohu16\]](#)

With the introduction of Q-SYS I/O-USB bridges and two PTZ IP conference cameras, QSC provides a networked AV example for the soft codec integration of medium and large conference room systems by integrating IP cameras into USB bridging solutions. The bridge solution of Q-SYS AV-to-USB leverages USB 2.0 for webcam video analog and AEC speaker loudspeaker audio. You don't need additional software or drives to support the latest PC operating systems and soft codec applications. [\[QSC16\]](#)

Networked AVs mean larger AV systems, where a technical manager can handle hundreds of ports of networks. Kramer's Kramer Network platform makes it as easy to operate as traditional smaller AV systems. Using a web-based interface, IT or AV managers can remotely configure, route, control, and manage Kramer's professional AV devices, room environments, and IP streaming devices, regardless of the notebook, PC, or tablet. Kramer Network can be installed on standard, enterprise-class, virtual, or cloud servers to manage and control the manufacturer's

Recent Advances in Audio/Video Networking

entire AV and IP products, as well as traditional AVs, Dante, and other devices. [\[Sara16\]](#)
[\[Dante21\]](#)

More sources and video walk IP also means more flexibility. For example, AMX's SVSI NMX-WP-N1512 window-opening processor, so that the enterprise network AV a step forward. The N1512 is a 1RU rack-mounted device connected to an SVSI Video Over IP network that accepts up to 4 video stream inputs. Each input can be cropped, scaled, and placed according to storage presets, or any user-defined configuration. With the N1512, there is no longer a need to connect directly to the video source because it is all web-based. Encoders are used to stream video content on the network, and window processors can be used anywhere there is a network, including data centers. [\[Sohu16\]](#)

2.7. AV over IP summary

The old walls that once existed between AV and IT are melting. A deep understanding of AV-over-IP technology and networking needs provides maximum value to our customers' audio and video assets. [\[Matrox21\]](#)

The transition of AV over IP from centralized A/V to distributed encoding and decoding emphasizes the value of audio and video assets themselves while allowing for better distribution of AV processing power and the ability to easily scale to meet user needs. In addition, by allowing users to deploy more flexibly, providing the option to use AV assets remotely, and enabling users to pursue new, more powerful features to better reflect evolving technologies and changing work habits. [\[Matrox21\]](#)

3. AVB

With the increasing demand of consumers for Ethernet multimedia applications, traditional Ethernet can not fundamentally meet the needs of real-time data transmission such as voice, multimedia, and other dynamic content, so AVB was born.

3.1. The main problem facing audio and video transmission

Think about it, what are the problems you often encounter when watching online videos? For example Caton, flower screen, (live streaming) time-lapse and audio and video out of sync, and so on. Assuming that the audio and video source is normal, what is the cause of the above problem?

Caton: The network is not smooth, the data can not arrive in time.

Flower screen: video data loss, error. However, this is only the superficial reason, the root cause is that the network is not smooth, data can not arrive in time.

(Live) latency: the network is not smooth, data can not arrive in time.

Recent Advances in Audio/Video Networking

Audio and video out of sync: network is not smooth, data can not arrive in time will cause this phenomenon, in addition, the media clock out of sync will also cause this phenomenon.

[\[coderhuo20\]](#)

In this way, the main reasons affecting the network audio and video viewing experience are the following two points:

1. The network is not smooth, data loss or can not reach
2. The media clock is out of sync. [\[coderhuo20\]](#)

This is also the main problem that AVB has to solve. The difference is that we watch network video that is generally transmitted over the WAN, and AVB solves the problem of audio and video transmission within the LAN. [\[coderhuo20\]](#)

3.2. Solution

How to solve these problems? AVB technology is a good solution to these problems.

3.2.1. Network transmission issues

Traditional Ethernet is a competitive "best-effort transmission" packet messaging network. Some might say that TCP is reliable, yes, but its reliability is based on a retransmitted mechanism, with the side effect of increased latency. In addition, intermediate forwarding nodes (such as switches and routers in the network) have a greater impact on the transmission of messages. For example:

1. During transmission, if the intermediate forwarding node cache queue is full, subsequent packets are lost.
2. Transmission delay is affected by the cache state of the forwarding node, if there is more data in the cache queue, the transfer delay is larger, the data in the cache queue is less, and the transmission delay is small. [\[coderhuo20\]](#)

If such a network can be built that reserves bandwidth for audio and video data (similar to a dedicated emergency lane on a highway), it not only solves the competition problem but also ensures that audio and video data can be transferred from the source to the destination within a specified time (the concept of a time-sensitive network). Yes, bandwidth reservations are an important part of AVB. [\[coderhuo20\]](#)

3.2.2. Media clock synchronization issues

Audio and video transmitted over Ethernet are discrete digital signals, while in the real world they are continuous analog signals. In audio and video acquisition equipment, analog signals need to be sampled, quantified, encoded to convert to digital signals (analog conversion, or A/D conversion). The A/D conversion process involves sampling frequency, i.e. how often the interval is sampled. When the receiving end receives the data, it plays at the same frequency.

Recent Advances in Audio/Video Networking

This is media clock synchronization. [\[Wikipedia21\]](#)

The media clock must have a reference clock so that it knows how long it has been since the last acquisition and whether the next sample should be taken. Some acquisition systems use the system clock as a reference clock, while others use external signals as reference clocks. Different reference clocks are likely to be out of sync, such as 1ms of A clocks may correspond to 0.9ms of the realization world, and 1ms of B clocks may correspond to 1.1ms of real-world amount. In addition, the clock is not stable, due to temperature, running time, and other factors that will produce drift (like watches, some overtime slow down, some overtime become faster). [\[Avnu21\]](#)

The question is, how does the receiver ensure that its playback rate and acquisition frequency are the same as the acquisition frequency on the acquisition side (i.e., how to restore the media clock)? This is difficult to do without a uniform reference clock. If the reference clock is out of sync, it may cause the audio that should have been played first to play later, or the video that should have been played first, which in the view of the audience is audio and video out of sync. [\[Avnu21\]](#)

The precise clock synchronization technology in AVB keeps the clocks of all nodes in sync, thus providing conditions for media clock synchronization. [\[Avnu21\]](#)

3.3. AVB system

AVB is called a time-sensitive network and consists of two main meanings:

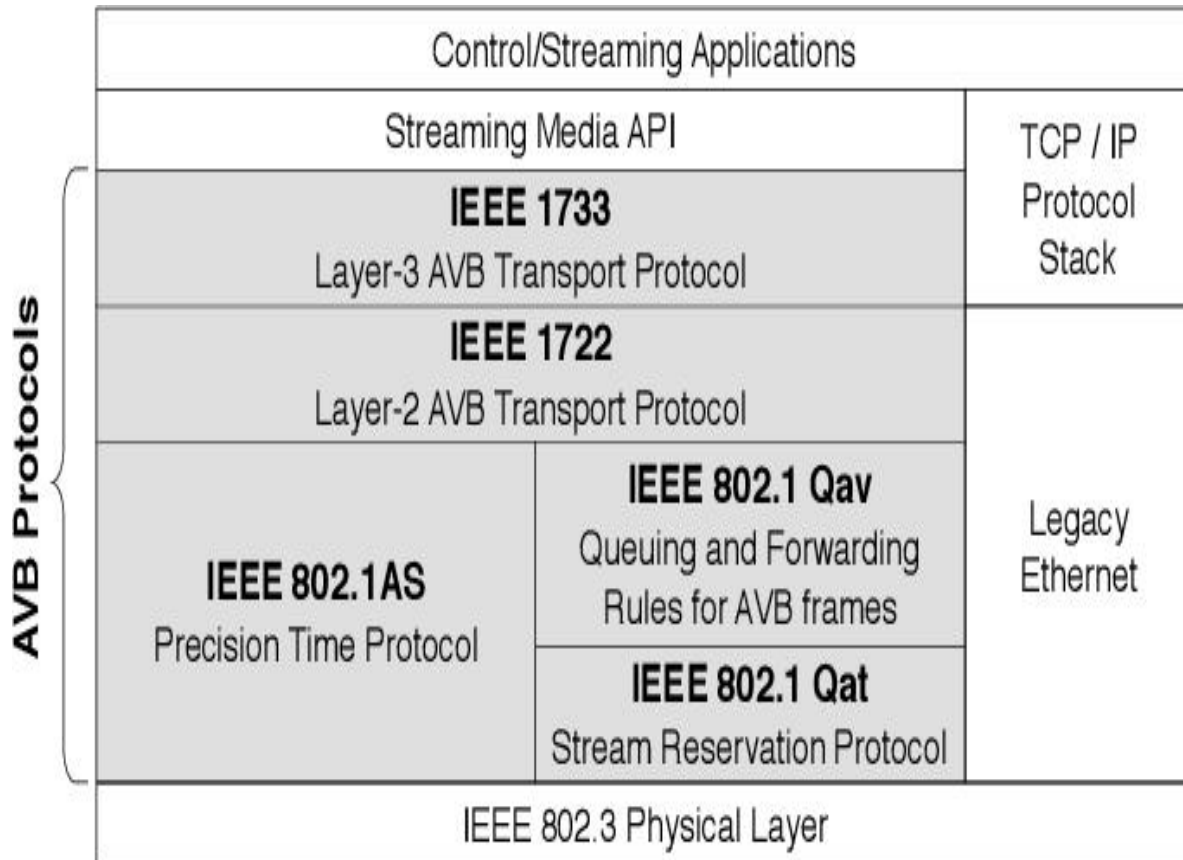
1. Small delay, good real-time.
 2. Delays are controllable and, in the worst case, do not exceed a certain threshold. [\[IEEE802\]](#)
-

3.3.1. Protocol framework

AVB, whose full name is Audio Video Bridging, consists of a series of IEEE standards designed to efficiently transmit audio and video data over a local area network. As shown in Table 1, the AVB stack looks like this:

Table 1 AVB protocol stack [\[coderhuo20\]](#)

Recent Advances in Audio/Video Networking



AVB is mainly a protocol of the link layer, which is related to the traditional TCP/IP protocol stack coexisting side by side. The AVB stack includes not only the bandwidth reservation and clock synchronization protocols mentioned earlier, but also several other auxiliary protocols, the main protocols are as follows. [\[coderhuo20\]](#)

3.3.1.1. IEEE Std 802.1AS-2011

The Generalized Precision Time Protocol (gPTP) is used to synchronize the clocks of all nodes in the network to the same master clock. 1. PTP defines the clock synchronization mechanism of the whole network based on IEEE 1588:2002 protocol. By defining the master clock selection and negotiation algorithm, path delay measurement, and compensation, and the mechanism of clock frequency matching and adjustment, the PTP device exchanges standard Ethernet messages to synchronize the time of each node of the network to a common master clock. [\[AS11\]](#)

As a simplified version of the IEEE 1588 protocol, the biggest difference between IEEE 802.1AS and 1588 is that PTP is a protocol that is based entirely on two-tier networks and is not IP routing. Like IEEE 1588, PTP defines a method of automatically negotiating the network master clock, the optimal master clock algorithm (Best Master Clock Algorithm, or BMCA for short). BMCA defines the underlying negotiation and signaling mechanism to identify the master clock (Grandmaster) within the AVB LAN. Once the master clock is selected, the PTP devices

Recent Advances in Audio/Video Networking

of all LAN nodes will use this master clock as a reference value, and if Grandmaster changes, the entire AVB network can determine the new master clock in the shortest possible time through BMCA, ensuring that the entire network keeps the time in sync. [\[AS11\]](#)

At the heart of 802.1AS is the timestamp mechanism. PTP messages, when entering and exiting ports with 802.1AS capabilities, trigger sampling of the local real-time clock (RTC) according to the protocol, compare their RTC values with information from the corresponding master clock (Master) of the port and match their RTC clock values to the time of the PTP domain using path delay measurement and compensation techniques. When the PTP synchronization mechanism covers the entire AVB LAN, the clock adjustment and frequency matching algorithm can be precisely realized between the network node devices through the exchange of periodic PTP messages. Eventually, all PTP nodes will be synchronized to the same Wall Clock time, the Grandmaster time. In a network environment with a maximum of 7 hops, PTP can theoretically guarantee clock synchronization errors of less than 111/4s. [\[AS11\]](#)

3.3.1.2. IEEE Std 802.1 Qat

Bandwidth Reservation Protocol (Stream Reservation Protocol, SRP) is currently consolidated in Chapter 35 of IEEE Std 802.1Q-2011. [\[Qat10\]](#)

The nature of the traditional IEEE 802 network standard limits its inability to prioritize normal asynchronous traffic with time-sensitive streaming traffic. To provide guaranteed quality of service (QoS), the Flow Reservation Protocol (SRP) ensures end-to-end bandwidth availability between AV-streaming devices. If the required path bandwidth is available, all devices on the entire path, including switches and end devices, lock this resource. [\[Qat10\]](#)

SRP uses the IEEE 802.1ak Multi-Registration Protocol (MRP) to deliver messages in exchange for bandwidth description messages for AV streams and to reserve bandwidth resources. SRP-compliant switches can use 75 percent of the bandwidth resources available across the network for AVB links, leaving 25 percent of the bandwidth left to traditional Ethernet traffic. In SRP, the provider of the streaming service is called Talker, and the recipient of the streaming service is called Listener. [\[Qat10\]](#)

The same Talker streaming service can be received by multiple Listeners at the same time, and SRP allows only one-way traffic from Talker to Listener to be guaranteed. Talker can start offering real-time AV services as long as bandwidth resources on any path from Talker to any of the multiple Listener can be negotiated and locked. The SRP's internal periodic state machine maintains the registration information of Talker and Listener, and can dynamically monitor the state of the network nodes and update its internal registration information database to adapt to the dynamic changes of the network topology. Both Talker and Listener can join or leave the AVB network at any time without an unrecoverable impact on the overall functionality and status of the AVB network. The SRP consists of registration and reservation, with Talker negotiating reservations for bandwidth resources required for AV streams and Listener registering and receiving the required AV streams. [\[Qat10\]](#)

3.3.1.3. IEEE Std 802.1 Qav

The Traffic Shaping Protocol (Forwarding and Queuing for time-sensitive streams, or FQTSS), is responsible for queuing and forwarding traffic and is currently integrated into Chapter 34 of IEEE Std 802.1Q-2011. [\[Qav09\]](#)

The purpose of Qav queueing and forwarding protocols is to ensure that traditional asynchronous Ethernet data traffic does not interfere with AVB's live audio and video streams. Time-sensitive AV stream forwarding uses Pseudo-Synchronous mode, a mechanism that relies on the SRP to provide reserved bandwidth along the route, and an 8Khz clock for PTP. In each 125us time slot, isochronous frames such as 802.3 Ethernet, which contains data, are forwarded. At the same time, to avoid the competition of network resources between ordinary data traffic and AVB traffic, the time-sensitive AV stream and ordinary data stream are handled differently within the AVB switch, queuing the equal-time frame and asynchronous frame respectively, and giving the equal-time frame the highest priority. [\[Qav09\]](#)

The service of normal asynchronous transmission continues to be provided under the condition of priority guarantee and equal-time frame transmission, which is Qav's Priority Management (Prioritize) and Traffic Shaping. While both terminals and switch devices require mechanisms to guarantee 75% of bandwidth resources for AVB applications, most implementations of 802.1Qav will be the responsibility of AVB switches. [\[Qav09\]](#)

3.3.1.4. IEEE Std 1722-2016

Audio-Video Transport Protocol (AVTP). AVBTP defines the two-tier package format required to provide real-time audio and video streaming services within the local area network, and the establishment, control, and shutdown protocol of AV streams. AVBTP establishes a virtual link with low latency between physically separated audio and video codecs that transmit data and synchronize AVs using the same streaming format as IEEE 1394. [\[IEEE16\]](#)

The IEC 61883 format used by AVBTP is as follows:

IEC 61883-2: SD-DVCR data transmission

IEC 61883-4: MPEG2-TS data transmission

IEC 61883-6: Audio and music data transmission protocol

IEC 61883-7: Transmission of ITU-R BO.1294 System B

IEC 61883-8: Transmission of ITU-R BT.601 style Digital Video Data

IIDC 1394-based Digital Camera Specification [\[IEEE16\]](#)

Various compressed and non-compressed raw audio and video streams are packaged through the AVBTP protocol (filling the stream ID retained by the SRP, timestamped with PTP, and related information such as media type), grouped through AVBTP-specific Ethernet frame types, emitted from the streaming service (Talker), forwarded by the AVB switch, and then received and unpacked, decoded, decoded, and exported by the recipient (Listener) who has registered for the AV streaming service. [\[IEEE16\]](#)

3.3.1.5. IEEE Std 1722.1-2013

Audio Video Discovery, Enumeration, Connection Management, and Control Protocol, AVDECC, is responsible for service discovery, node capability traversal, link management, and other control functions. [\[IEEE13\]](#)

3.3.1.6. IEEE Std 1733-2011

AVB-based RTP/RTCP Transport Protocol (Transport Protocol for Time-Sensitive Applications in Area Networks) is primarily used to coordinate multiple devices in a network, such as the following scenario: There are multiple Talker devices in the network and they are located on different devices. The receiving end needs to synchronize these Talker audio and video data. [\[IEEE11\]](#)

RTP is one of the most widely used real-time streaming protocols, unlike IEEE 802.1 AVB, which is based entirely on two-tier standards, and RTP is a three-tier UDP/IP network-based protocol. To take advantage of the performance of the two-tier AVB in IP-based three-tier applications, IEEE 1733 extends RTP to provide time synchronization, latency assurance, and bandwidth reservation services within the local area network through bridges and routing to provide real-time audio and video stream transmission. It involves the packet format, the establishment, control, synchronization, and closure of the flow and other protocols. [\[IEEE11\]](#)

3.3.2. Network topology

As shown in Figure 3, the network topology of the AVB looks like this:

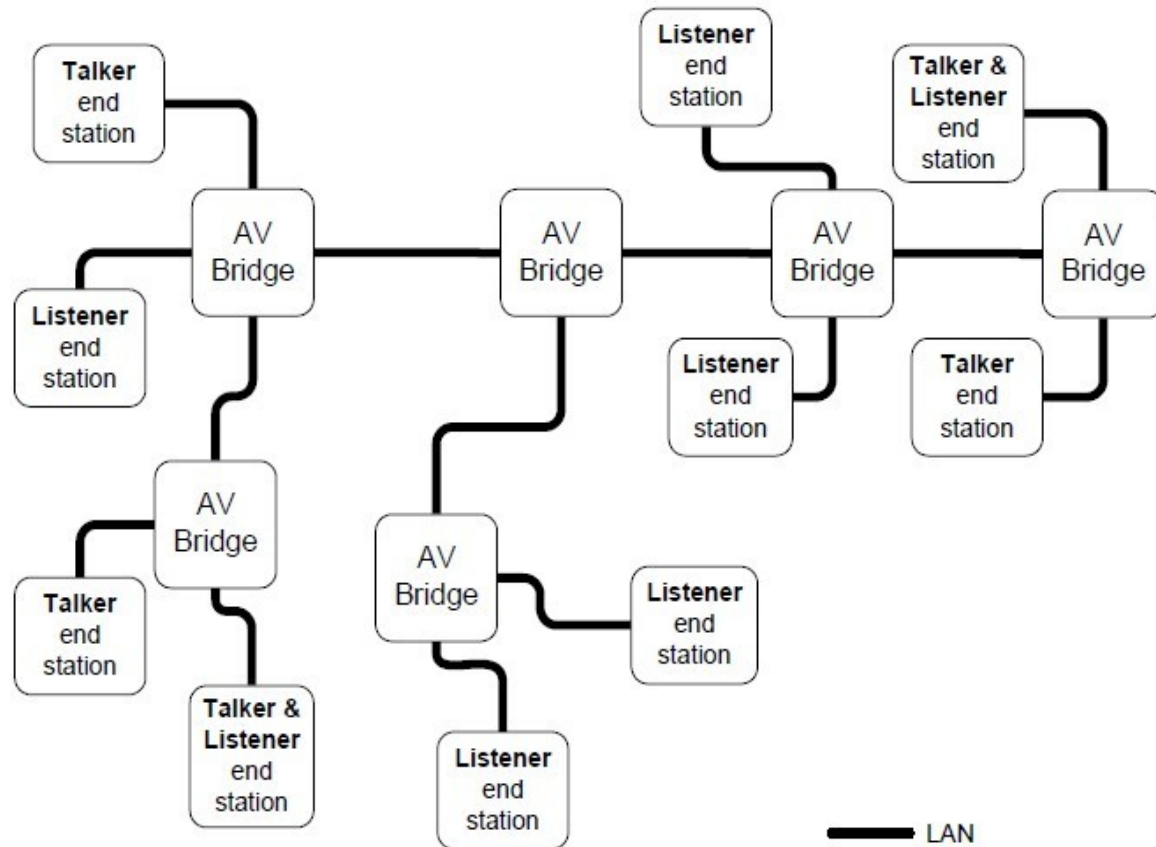


Figure 3 AVB network topology [\[coderhuo20\]](#)

It consists of two main types of nodes:

1. End Station: Sometimes referred to as endpoint, it is divided into Listener and Talker: The listener is the node that receives audio and video data. Talker is the node that outputs audio and video data.
2. AV Bridge: Audio and Video Gateway. [\[coderhuo20\]](#)

There are two characteristics:

1. A node can contain more than one Talker, i.e. produce multiple audio and video streams.
2. A node can have the dual identity of Talker and Listener, both external output and receiving other people's data. [\[coderhuo20\]](#)

3.3.3. Typical scenario

At present, AVB is mainly used in time-sensitive scenes, such as in-vehicle entertainment systems. Here is a simple example.

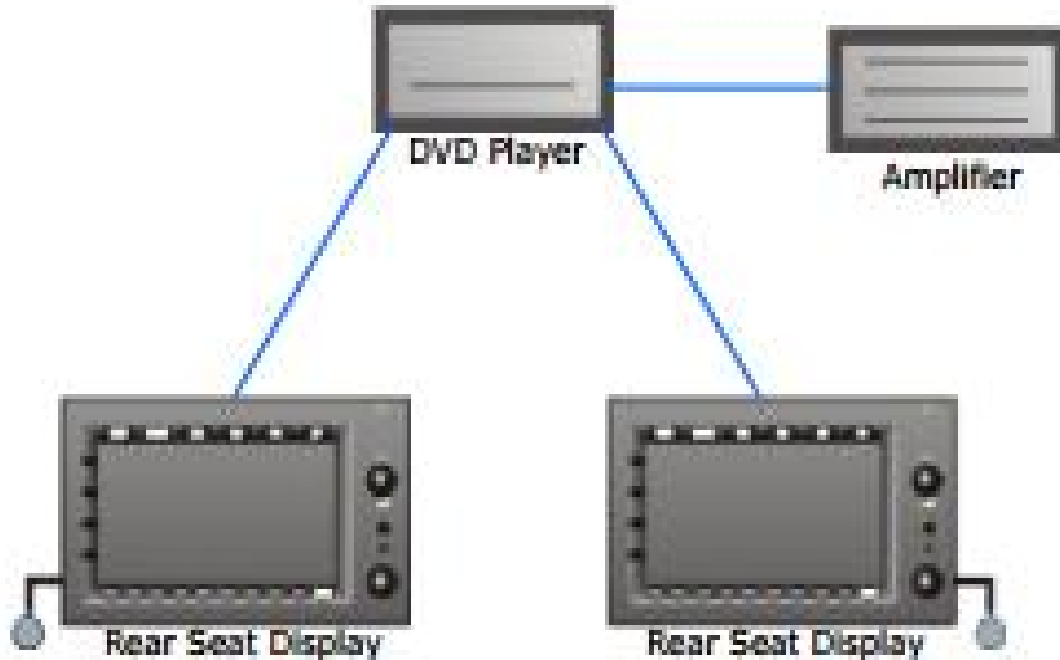


Figure 4 In-car entertainment system [\[coderhuo20\]](#)

As shown in Figure 4, the DVD Player in the car plays audio and video files, which are transmitted to the speakers and video to two monitors in the rear seats. How do I keep my speakers and two monitors in sync for audio and video synchronization? [\[coderhuo20\]](#)

First, the gPTP protocol ensures that the clocks of all nodes in the network are synchronized. Secondly, the SRP protocol reserves bandwidth for audio and video data to ensure that the destination can be reached within a certain time. So, if the DVR Player is agreed, will audio and video data from different destinations be played at the same time in the future? This is what timestamps are shown in the AVTP protocol. [\[coderhuo20\]](#)

3.4. AVB Summary

AVB is a collection of IEEE 802.1 standards developed by the Time-SensiTive Network Task Group to increase the capacity, industry support, and standardization of AV products for information exchange. Audio and video bridges are often referred to as switches that are designed to provide time-synchronized, low-latency, and bandwidth-guaranteed streaming capabilities for audio and video data. [\[Wiki21\]](#)

The goal of AVB is to use "smart" switches that can handle AVB traffic without compromising

Recent Advances in Audio/Video Networking

network consolidation to transmit audio, video, and other forms of data. With AVB, IT managers can adopt a hybrid data network that is more time-saving, lower-cost, and more efficient than parallel stand-alone systems to perform their duties as AV system managers. Integrating AVB technology into today's IT systems is a huge step forward and will lead to a fundamental shift in the principles that most IT and AV administrators design systems. All these features show that it will be one of the next generations of streaming file transmission standards, whether professional or civilian, will show its strong charm.

4. Summary

The AV over IP architecture significantly increases the throughput of transferred data by adding network switching devices, comparable to the traditional way of using cable-to-point transmissions such as HDMI/SDI, and is not limited by the number of physical interfaces of the device. In the AV over IP architecture, the traditional audio and video signal transmitter become an IP encoder, the audio and video signal receiver becomes an IP decoder, for compressed or non-compressed audio and video transmission, and through the codec to complete part of the video processing work, such as up and down, color space conversion, picture flipping and so on. [\[LEDinside20\]](#)

The IEEE 802.1 AVB Working Group has developed a series of new standards to extend the capabilities of existing Ethernet-based on maintaining full compatibility with existing Ethernet systems, providing a variety of LAN supporting solutions for ordinary data and real-time audio and video streaming by establishing high-quality, low-latency, time-synchronized audio and video Ethernet networks.

6. References

[\[baidu21\]](#) -- "AV", <https://baike.baidu.com/item/AV/20385223?fr=aladdin>, [AV(industry terminology)]

[\[Matrox21\]](#) -- "Fundamentals of AV over IP", <https://www.matrox.com/en/video/media/guides-articles/fundamentals-of-av-over-ip>, [Fundamentals of AV over IP]

[\[Marta19\]](#) -- "AV over IP: The benefits of remote live video production", <https://www.epiphan.com/blog/av-over-ip-remote-video-production/>, July 19th, 2019, Marta Chernova, [AV over IP could be used in different workflows, including webcasting/streaming, VOD environments, or a broadcast setup. Additionally, AV over IP makes the concept of remote video production possible.]

[\[Cisco21\]](#) -- "Understand AVB in Catalyst 3K and Catalyst 9000 Series Switches", <https://www.cisco.com/c/en/us/support/docs/ip/multicast/216531-overview-of-avb-in-cat3k-and->

Recent Advances in Audio/Video Networking

cat9k-switc.html, October 15, 2021, Eduardo Cruz, Luis Fernando Pastrana, [This document describes how to configure and troubleshoot Audio Video Bridging (AVB) in the Catalyst 3650, 3850, 9300, and 9500 platforms.]

[[EETimes07](#)] -- "Ethernet and Multimedia Applications - The History and the Future - Part 1", <https://www.eetimes.com/ethernet-and-multimedia-applications-the-history-and-the-future-part-1/#>, Neeraj Parik, Xilinx, 11.21.2007, [Of course, the fact that FPGAs are increasing being used in Ethernet-based systems means that it would be useful for me to discover more.]

[[LAWO21](#)] -- "IP NETWORKING GUIDE FOR VIDEO AND AUDIO APPLICATIONS", https://lawo.com/Downloaery/89_Reference_Material/IP_Networking_Guide_for_Video_and_Audio_Applications_EN.pdf

[[Earon21](#)] -- "Understanding & Evaluating AV-over-IP", S. Ann Earon, https://www.crestron.com/getmedia/b6a586ce-9c00-4e50-ac64-9d844881ef6d/wm_av-over-ip_whitepaper

[[Wikipedia20](#)] -- "Audio-to-video synchronization", https://en.wikipedia.org/wiki/Audio-to-video_synchronization, 5 August 2020, [Audio-to-video synchronization (AV synchronization, also known as lip sync, or by the lack of it: lip-sync error, lip flap) refers to the relative timing of audio (sound) and video (image) parts during creation, post-production (mixing), transmission, reception and play-back processing. AV synchronization can be an issue in television, videoconferencing, or film.]

[[QSC16](#)] -- "QSC To Unveil New Q-SYS AV-to-USB Bridging Solution at InfoComm 2016", https://www.qsc.com/pl/live-sound/news/detail/?tx_news_pi1%5Bnews%5D=388&tx_news_pi1%5Bcontroller%5D=News&tx_news_pi1%5Baction%5D=detail&cHash=2cf75a794442a303b390a391a835eb37,05/23/2016, [QSC expands its powerful software-based AV/IT solution, Q-SYSa"c, with PTZ-IP Camera and a AV-to-USB Bridging peripheral to meet the needs of the corporate AV market for soft codec camera and audio bridging integration.]

[[Sara16](#)] -- "Kramer Introduces Kramer Network - An AV Over IP System - at ISE 2016", <https://www.ravepubs.com/kramer-network-av-ip/>, February 10, 2016, Sara Abrons, [Kramer Electronics is introducing something called Kramer Network, an enterprise management platform for AV over IP solutions at ISE 2016.]

[[Dante21](#)] -- "Dante Audio and Video, Together at Last", <https://www.audinate.com/products/manufacture-products/dante-av-video>, [Dante AV is a revolutionary new AV solution for standard 1 Gbps networks that allows manufacturers to add networked video to the industry-leading Dante audio-over-IP platform.]

[[Sohu16](#)] -- "Meet the "AV over IP" era", 2016-12-01, https://www.sohu.com/a/120345930_393832

[[coderhuo20](#)] -- "Introduction to AVB - First: Overview", 2020-04-02,

Recent Advances in Audio/Video Networking

<https://cloud.tencent.com/developer/article/1608775>

[[Wikipedia21](#)] -- "Analog-to-digital converter", https://en.wikipedia.org/wiki/Analog-to-digital_converter, 15 November 2021, [An ADC converts a continuous-time and continuous-amplitude analog signal to a discrete-time and discrete-amplitude digital signal.]

[[Avnu21](#)] -- "What is AVB?", <https://avnu.org/faqs/>, [Audio Video Bridging (AVB) replaces both the physical complexity of analog cables and the network complexity of earlier proprietary solutions with an open, standards based approach that enables interoperable platforms to offer more affordable and better market solutions.]

[[IEEE802](#)] -- Michael D. Johas Teener, San Jose, Andre N. Fredette, Christian Boiger, Philippe Klein, Craig Gunther, David Olsen, Sandy, Kevin Stanton, Hillsboro, "Heterogeneous Networks for Audio and Video", <https://www.ieee802.org/1/files/public/docs2013/avb-mjt-et-all-AVB-for-IEEE-Smart-Home-0213.pdf>

[[AS11](#)] -- "IEEE 802.1AS-2011 - IEEE Standard for Local and Metropolitan Area Networks - Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks", https://standards.ieee.org/standard/802_1AS-2011.html, 2011-03-30, [This standard defines a protocol and procedures for the transport of timing over bridged and virtual bridged local area networks.]

[[Qat10](#)] -- "IEEE 802.1Qat-2010 - IEEE Standard for Local and metropolitan area networks--Virtual Bridged Local Area Networks Amendment 14: Stream Reservation Protocol (SRP)", https://standards.ieee.org/standard/802_1Qat-2010.html, 2010-09-30, [This amendment specifies protocols, procedures, and managed objects, usable by existing higher layer mechanisms, that allow network resources to be reserved for specific traffic streams traversing a bridged local area network.]

[[Qav09](#)] -- "802.1Qav - Forwarding and Queuing Enhancements for Time-Sensitive Streams", <https://www.ieee802.org/1/pages/802.1av.html>, Thu, 10 Dec 2009, [The full title of this PAR is "IEEE Standard for Local and Metropolitan Area Networks---Virtual Bridged Local Area Networks - Amendment: Forwarding and Queuing Enhancements for Time-Sensitive Streams.".]

[[IEEE16](#)] -- "IEEE 1722-2016 - IEEE Standard for a Transport Protocol for Time-Sensitive Applications in Bridged Local Area Networks", <https://standards.ieee.org/standard/1722-2016.html>, 2016-12-16, [The protocol, device discovery, connection management, and device-control procedures used to facilitate interoperability between audio- and video-based end stations that use IEEE 1722-based streams on IEEE 802-based networks are specified.]

[[IEEE13](#)] -- "IEEE 1722.1-2013 - IEEE Standard for Device Discovery, Connection Management, and Control Protocol for IEEE 1722(TM) Based Devices", https://standards.ieee.org/standard/1722_1-2013.html, 2013-10-10, [This standard specifies the protocol, device discovery, connection management and device control procedures used to facilitate interoperability between audio and video based End Stations that use IEEE 1722 based Streams on IEEE 802(R) based networks.]

Recent Advances in Audio/Video Networking

[[IEEE11](#)] -- "IEEE 1733-2011 - IEEE Standard for Layer 3 Transport Protocol for Time-Sensitive Applications in Local Area Networks", <https://standards.ieee.org/standard/1733-2011.html>, 2011-04-25, [Protocol, data encapsulations, connection management, and presentation time procedures used to ensure interoperability between audio and video-based end stations that use standard networking services provided by all IEEE 802 networks meeting QoS requirements for time-sensitive applications by leveraging the Real-time Transport Protocol (RTP) family of protocols and IEEE 802.1 Audio/Video Bridging (AVB) protocols is specified in this standard.]

[[Wiki21](#)] -- "Audio Video Bridging", https://en.wikipedia.org/wiki/Audio_Video_Bridging, 29 November 2021, [Audio Video Bridging (AVB) is a common name for the set of technical standards which provide improved synchronization, low-latency, and reliability for switched Ethernet networks.]

[[LEDinside20](#)] -- "AV over IP- The Future of ProAV Integration You Need to Know About", <https://www.ledinside.com/node/31755>, December.8, 2020

5. List of Acronyms

A.AV -- Audio and Video
A.AVB -- Ethernet audio and video bridging technology
A.HDCP -- high-bandwidth digital content protection
A.gPTP -- Generalized Precision Time Protocol
A.BMCA -- Best Master Clock Algorithm
A.SRP -- Stream Reservation Protocol
A.QoS -- quality of service
A.MRP -- Multi-Registration Protocol
A.FQTSS -- Forwarding and Queuing for time-sensitive streams
A.AVTP -- Audio Video Transport Protocol
A.AVDECC -- Audio Video Discovery, Enumeration, Connection Management, and Control Protocol

Last modified on December 15, 2021

This and other papers on recent advances in networking are available online at

<http://www.cse.wustl.edu/~jain/cse570-21/index.html>

[Back to Raj Jain's Home Page](#)