AoIP and AES67 and Their Impact on Future Audio Systems Postprint

Authors: gradual accumulation

Date: 2023-10-08T00:00:00+00:00

Abstract

Owing to the rapid development of digital audio technology, the transmission efficiency of audio signals has been significantly improved. Presently, networked transmission has become the most popular transmission method in contemporary society, gaining widespread public acceptance and high recognition. This paper mainly introduces the advantages of Audio over IP (AoIP), also provides an introduction to AES67, expounds upon the content of the AES67 standard, and on this basis, analyzes the future impact and development trends of audio systems.

Full Text

AoIP and AES67 and Their Impact on Future Audio Systems

Abstract: Due to rapid developments in digital audio technology, audio signal transmission efficiency has improved significantly. Today, networked transmission has become the most popular means of transmission in modern society, gaining widespread acceptance. This paper primarily introduces the advantages of networked audio (AoIP), provides an overview of AES67, elaborates on the content of the AES67 standard, and analyzes its future impact and development trends for audio systems.

Keywords: AES67; AoIP; transmission protocol; digital audio; network audio

Classification: TN931.2
Document Code: A
Article ID: 1671-0134(2018)02-051-02
DOI: 10.19483/j.cnki.11-4653/n.2018.02.017
Author: Cun Lei

1. Network Audio AoIP

Due to network technology development, information processing efficiency has significantly improved. AoIP technology effectively integrates digital audio with information network technology, offering numerous advantages. First, based on TCP/IP networks, professional audio transmission can utilize ethernet cables, fiber optics, and network switches. Second, a single cable can simultaneously transmit audio, video, and control signals, facilitating remote device control. Moreover, in the latest protocol formats, PTP clocks enable device network entry and synchronization. Third, through data streaming, numerous bidirectional data streams can be transmitted over a single cable, ensuring reliable operation of large-scale IP audio networks. Fourth, switches enable packet replication and distribution, achieving point-to-multipoint transmission that simplifies the process and improves efficiency.

In terms of connection management, IP audio networks demonstrate clear advantages over traditional point-to-point connections, primarily through enhanced routing flexibility. AES67 employs multicast mode through IGMP V2 group management protocols, allowing audio data streams to be replicated by switches and transmitted to N devices. System construction is relatively simple and bandwidth-efficient, with transmission bandwidth largely dependent on the switches.

2. Introduction to AES67

2.1 Objectives of AES67

In recent years, numerous networked audio transmission protocols have emerged with distinct characteristics. However, most are proprietary protocols with poor compatibility and lack openness at the infrastructure level, typically operating within isolated domains with dedicated support teams. This environment, particularly with newly developed protocols, has significantly hindered AoIP development. Theoretically, establishing a universal standard compatible with all protocols would allow manufacturers to avoid revolutionary changes while achieving mutual compatibility, thereby jointly advancing IP audio technology. It was this need that gave rise to AES67, which defines appropriate standards enabling digital audio signals to be transmitted and shared over LANs and the Internet through TCP/IP protocols.

2.2 Core Functions of AES67

AES67's core functions include several key aspects. First, for transmission, encoding, and streaming, AES67 clearly defines the application of RTP and UDP protocols for transmitting standard IP format data packets, with each packet containing 1 ms of audio data. Second, media clock synchronization ensures all devices achieve synchronization after connecting to the audio network. Third, connection management leverages IP audio networks' flexible routing

advantages over traditional point-to-point connections. By employing multicast mode through IGMP V2 group management protocols, audio data streams can be replicated by switches and distributed to N devices, simplifying system construction while conserving bandwidth. Fourth, packet replication and distribution via switches enable efficient point-to-multipoint transmission.

2.3 AES67 Packet Encoding

AES67 specifies the use of 16/24-bit linear PCM encoding at a 48 kHz sampling rate, with each data stream limited to a maximum of 8 channels. For audio data transmission, the data can be divided into multiple packets, with each packet's payload not exceeding 1440 bytes—a limitation primarily due to historical constraints, as Ethernet switches support a maximum packet length of 1500 bytes; exceeding this would cause packet fragmentation. To ensure real-time audio stream transmission, each packet should be limited to 1440 bytes.

2.4 AES67 Data Stream Bandwidth

A stereo AES67 data stream requires approximately 3 Mbit/s of bandwidth, while an 8-channel stream requires 10 Mbit/s. Both stream types share the same packet transmission frequency, with each packet accommodating up to 48 samples of 24-bit audio across 8 channels, requiring only one thousand packets per second. Currently, a 10 Mbit/s data volume is relatively modest for IP networks, aligning with AES67's core concept of ensuring real-time audio data transmission even on low-end network equipment. However, multiple AES67 data streams may be transmitted and received over the same ethernet cable or port, and excessive data streams can significantly impact network device performance.

2.5 AES67 Networking and Transmission

In network transmission, AES67 primarily utilizes Layer 3 IPv4 protocols, enabling audio data streams to be transmitted and shared over the Internet. RTP is one of the transport protocols employed by AES67, commonly used in streaming media systems, video conferencing, and push-to-talk systems, thereby enabling integration with related communication and entertainment services such as telephony, television, video conferencing, and network-based push-to-talk services.

2.6 AES67 Quality of Service (QoS)

Through Quality of Service (QoS) mechanisms, network communication can be prioritized. This security framework effectively handles latency issues caused by sudden increases in network data volume. AES67 implements three QoS priority levels: PTP clock packets receive the highest priority, followed by audio data, with other data ranked last. Priority packets are forwarded preferentially by the system, ensuring real-time transmission of clock and audio signals.

2.7 AES67 Synchronization Signals and Clocking

For synchronization and clocking, AES67 employs PTP (Precision Time Protocol), broadcasting master clock signals over IP networks. Precise clocks are distributed throughout the entire network, enabling all devices to synchronize via these clock signals. Through PTP implementation, AES67 achieves an accuracy of 25 ns, substantially establishing a foundation for system clock precision.

2.8 Comparison and Compatibility Between AES67 and RAVENNA Protocols

The RAVENNA protocol shares many similarities with AES67, but comparative analysis reveals RAVENNA's superior performance in several aspects: (1) Encoding: While AES67 only transmits 24-bit/48 kHz audio, RAVENNA supports 96 kHz and DSD audio encoding; (2) Data streams: AES67 limits each stream to 8 channels, whereas RAVENNA provides 64-channel and 128-channel streams; (3) QoS: AES67 simply divides data into three tiers with all audio data in one tier, while RAVENNA can define multiple priority levels for audio data streams, accommodating both high-priority and low-priority audio to facilitate construction of ultra-large-scale IP audio networks.

From this analysis, it is evident that while AES67 may not represent the pinnacle of technical performance, its strength lies in excellent compatibility with other protocols. Currently, numerous network protocols support AES67 compatibility, including RAVENNA and Q-LAN. Users can freely choose among these protocols for internal architecture while employing AES67 standards for external connections, fundamentally achieving interoperability with other external devices.

The influence of AES67 has been encouraging for both audio professionals and IT practitioners. AES67 implementation improves existing network architectures, reducing both capital and labor costs. In standard networks, multiple data types can be transmitted over a single network without incurring additional expenses. For IT staff, AES67 protocols do not require overhauling existing networks, allowing products with different protocols to be integrated using compatible products. Overall, compared to dedicated network audio transmission, modern IT engineers can monitor and manage all facility data to ensure high-quality audio transmission.

3. Analysis of Future Audio Systems

Networked audio signal transmission offers clear advantages in continuously improving transmission efficiency while meeting diverse user demands—namely, providing convenient, efficient, and user-friendly solutions to practical problems. With the rapid development of information technology, the universality and standardization of AoIP technology are guiding the future of the audio field. Consequently, future digital audio systems will undoubtedly be IP-network-

based, integrated with other network technologies, and will primarily feature the following characteristics.

First, network broadcasting will significantly facilitate signal sharing and rational allocation, while traditional routing equipment such as audio matrices will gradually exit the market, replaced by switches and control software. Second, audio system architecture will become increasingly clear and streamlined, simplifying system construction and operational convenience. However, professional engineers will be required for backend system configuration, making network setup and optimization core elements in audio system construction, debugging, and application. Third, driven by information processing technology and combined with relevant video standards, the integration of audio, video, and related control systems will be further enhanced, unifying them within a single network system architecture. This will establish a unified backend, promoting the integration of image and voice to realize multimedia information systems. Fourth, remote control will become increasingly prevalent, along with remote management of maintenance, signal production, and processing, resulting in substantial cost savings and reduced time and effort.

During the drafting phase of the AES67 protocol, experts noted that its publication would elevate Audio over IP technology to a new level. Users can freely select AoIP devices and exchange audio between different systems. AES67 and AoIP represent the future of networked audio.

From the above discussion, it is evident that educational institutions offering audio engineering programs should incorporate relevant technical courses to ensure future professionals possess knowledge structures aligned with industry trends. Concurrently, current practitioners should further their understanding of network IP technology to remain competitive in this rapidly evolving era.

References [1] Gu Qiyuan, Huang Xiaoge. Analysis of CobraNet Real-time Audio Transmission Protocol[J]. Electroacoustic Technology, 2007, 31(2): 67-68.
[2] Jiang Shaohua. Technical Principles and Practical Applications of CobraNet[J]. Audio Technology, 2012(5): 50-53.
[3] Zhao Hui. Ethersound Network Digital Audio Technology and Application[J]. Electroacoustic Technology, 2004(1): 71-72.

(Author Affiliation: Yunnan Radio and Television Station)

Note: Figure translations are in progress. See original paper for figures.