AES67-2013:
AES standard for audio applications of networks - High-performance streaming audio-over-IP interoperability

Andreas Hildebrand
ALC NetworX GmbH
Munich, Germany

Abstract - *This paper accompanies the respective NAB BEC 2013 presentation on the AES67 standard on high-performance streaming audio-over-IP interoperability. While the conference presentation can provide just a brief overview on intention, scope and content of the standard, this paper provides a more in-depth view on the technical ingredients of AES67 and its applicability.*

NOTE: *Any information contained in this paper reflects the author's interpretation of the official AES67 standards document and personal perception of the related AES X192 Task Group work. This is not an official AES Standards Committee publication.*

INTRODUCTION / MOTIVATION

High-performance media networks support professional quality audio (16 bit, 48 kHz and higher) with low latencies (less than 10 milliseconds) compatible with live sound reinforcement. The level of network performance required to meet these requirements is available on local-area networks and is achievable on enterprise-scale networks. In the past, a number of networked audio systems have been developed to support high-performance media networking, but despite being based on IP and using alike synchronization and transport mechanisms and payload formats, none of these systems were interoperable, and no recommendations for operating these systems in an interoperable manner were in existence.

STANDARDS WORK

In late 2010 the AES inaugurated a Task Group to address this issue. The Task Group was called SC-02-12-H, the standards project had been designated X192. Under management of Task Group leader Kevin Gross, the Task Group started its work at the end of 2010. After two and a half years of sustained work in bi-weekly web conferences and several face-to-face meetings the Task Group concluded on a proposed standards document, which had been handed over to the hosting SC-02-12 Work Group. Within this period, the Task group membership list grew up to more than 100 individuals, representing a prestigious spectrum of companies and organizations from the professional audio community.

Following a public commenting period, on September 11th, 2013 the AES finally published the *AES67-2013: AES standard for audio applications of networks - High-performance streaming audio-over-IP interoperability.*

SCOPE AND APPLICATIONS

This standard defines an interoperability mode for transport of high-performance audio over networks based on the Internet Protocol suite (OSI layer 3 and above). For the purposes of the standard, high-performance audio refers to audio with full bandwidth and low noise. These requirements imply linear PCM coding with a sampling frequency of 48 kHz and higher and resolution of 16 bits and higher. High performance also implies a low-latency capability compatible with live sound applications. The standard therefore considers latency performance of 10 milliseconds or less.

Definitions, recommendations and guidelines given by this standard cover the areas of synchronization, media clock identification, network transport, encoding and streaming, stream description and connection management. No new protocols are developed to achieve this; the standard rather defines how existing protocols are used to create an interoperable system. However, it is expected that individual media networking solutions will require some modifications and extensions to match or improve interoperability based on this standard.

The standard is expected to be useful for commercial audio applications including fixed and touring live sound reinforcement. It is also expected to be useful for distribution within broadcast, music production and post-production facilities.

STANDARD "INGREDIENTS"

To achieve a sound interoperability definition, AES67 addresses these functional areas:
- Synchronization – defines the mechanism for a common clock system
- Media clocks – defines which media clocks have to be supported and how they are related to the common clock system
- Transport – describes how media data is transported across the network
- Encoding and streaming – describes the means in which audio is digitized and formatted into

the sequence of packets that constitutes a stream

- Stream description – required for connection management, specifies relevant stream information such as network addressing, encoding format and origination information
- Connection management – the procedure and protocols used to establish a media stream connection between a sender and one or more receivers

Device and stream discovery is intentionally excluded from standard works, as a number of different discovery mechanisms are available, with certain preference for one or the other in various network environments and application scenarios. However, some recommendations for suitable discovery mechanisms are given within the annex of this standard, but interoperability can also be achieved without a common discovery mechanism by distributing the relevant parameters for connection management by other (i.e. administrative) means. Future extensions or supplemental standards may address this topic.

SYNCHRONIZATION

The ability for network participants to share an accurate common clock distinguishes high-performance media streaming from its lower-performance relatives such as Internet radio and IP telephony. A common clock allows for an identical packet rate among all senders and receivers and thus for a fixed and determinable latency between sender and receivers. A common clock assures that all streams are sampled and presented at exactly the same rate. Streams running at the same rate may be readily combined in receivers. This property is critical for efficient implementation of networked audio devices such as digital mixing consoles.

Synchronization in AES67 is achieved by using IEEE 1588-2008 Precision Time Protocol (PTPv2) [1] to establish a network-wide common wall clock. IEEE 1588-2008 defines a protocol for distributing absolute time between a Grandmaster and any number of slave clocks (Ordinary Clocks).

The IEEE 1588-2008 time domain is TAI (International Atomic Time), and has its epoch origin at 01.01.1970 00:00:00 UTC. Unlike UTC, TAI does not feature leap seconds, thus providing unambiguous monotonic time information. However, actual UTC time may be calculated from IEEE1588-2008 taking all leap second occurrences since 01.01.1970 into account, if required.

The PTP protocol includes a Best Master Clock algorithm (BCMA), which would result in the most accurate clock being elected to serve as the Grandmaster. Although not required, a Grandmaster clock is typically synchronized to an external traceable reference (i.e. GPS).

The protocol utilizes multicast SYNC messages to transport the absolute time information from the Grandmaster to all slave clocks. To increase the achievable accuracy, a delay request mechanism is employed to calculate for individual link offsets through the network. The achievable precision is depending on various factors like network segment size, number of clients, traffic conditions and other network-specific parameters, but may well reach into the dual-digit nanoseconds range.

IEEE 1588-2008 defines various operating profiles which are designated for specific applications (i.e. Telecom profile, Power profile etc.). For AES67, support of the Default profile is mandatory. In addition, it is recommended to also support a new Media profile which is proposed in annex A of the AES67 standard. Application of this Media profile would yield faster synchronization in standard (non-PTP capable) IP network environments. Note: SMPTE Technology Committee 33TS-20 is also working on a new PTP profile definition; a liaison between SMPTE TC33 and AES had been established during X192 Task Group work to achieve a close alignment between both profile definitions. However, SMPTE TC33 has not yet concluded its work.

MEDIA CLOCKS

Media clocks are used by senders to sample and by receivers to play digital media streams. AES67 mandates for support of 48 kHz sampling frequency; to further expand interoperability and application support AES67 recommends to also support 44.1 and 96 kHz. The media clocks have a fixed relationship to the network-wide wall clock, i.e. the media clock for an audio stream sampled at 48 kHz advances exactly 48,000 samples for each elapsed second on the wall clock.
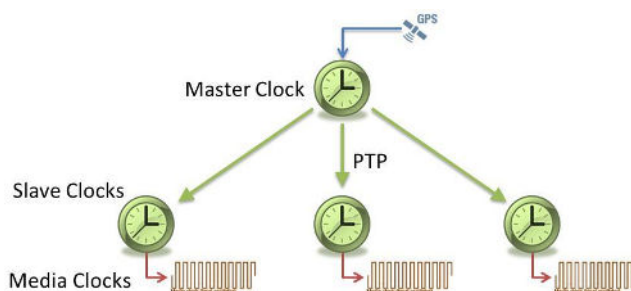

FIG 1: SYNCHRONIZATION & MEDIA CLOCK MODEL

To stay compliant with RFC 3550 [2], the RTP clocks (representing sample counts) operate with a constant offset with respect to the media clock. The offset is conveyed through stream description on a per-stream basis.

TRANSPORT

Transport aspects describe how media data, once encoded and packetized, is transported across the network.

I. Network Layer

Media packets are transported using IP version 4 as defined in RFC 791 [3]. Although care has been taken in design so as to facilitate future support for IPv6, it is currently outside the scope of AES67.

II. Quality of Service

On a network shared with unregulated non-real-time traffic, time-critical media traffic generally requires prioritized handling known as QoS. In order to facilitate the implementation of suitable QoS in the network, devices shall implement the DiffServ method as described in RFC 2474 [4]. DiffServ uses the DSCP field in each IP packet header to mark packets according to their traffic class so that the network can easily recognize packets that need to be treated preferentially.

At least three traffic classes have to be supported by the network. AES67-compliant devices should tag outgoing traffic with an appropriate DSCP value according to table 1:

Class name	Traffic type	DiffServ class (DSCP value)
Clock	IEEE 1588-2008 *Announce, Sync* and all delay request & response messages	EF (46)
Media	RTP and RTCP media stream data	AF41 (34)
Best effort	IEEE 1588-2008 signaling and management messages; connection management messages; other (non-X192) network traffic	DF (0)

TABLE 1: QoS CLASSES AND DIFFSERV ASSOCIATIONS

The participating network switches need to be configured in an appropriate manner through administrative means to support adequate forwarding behavior required by the real-time nature of AES67-related traffic. Other traffic classes / DSCP markings may be used in certain network situations; however, these situations require specific knowledge about network transport policies and requirements of other traffic services in order to establish a suitable system configuration.

III. Transport Layer

The transport layer provides end-to-end communications between devices on a network.

The Real-time Transport Protocol (RTP) in accordance with the AVT profile as defined in RFC 3550 / 3551 [5] is used for media data transport. RTP packets are transported with UDP as defined in RFC 768 [6]. Use of RTCP according to RFC 3550 clause 6 is recommended, but not required. For maximum payload size, the standard 1500 byte maximum Ethernet MTU is assumed. However, to assure future compatibility with IPv6, for AES67 the maximum allowed RTP payload is 1440 bytes.

IV. Multicasting

Multicast messaging, such as that used for synchronization or media streaming, is accomplished using IP multicasting as described in RFC 1112 [7]. To ensure that desired multicasts are received and to allow the network to filter multicasts from devices that do not need them (to prevent flooding), all devices must support IGMPv2 as defined in RFC 2236 [8].

ENCODING AND STREAMING

Encoding describes the means in which audio is digitized and formatted into the sequence of packets that constitutes a stream.

I. Payload Formats & Sampling Rate

Payload format defines audio sample encodings. The following payload formats are supported:

L16 16-bit linear format defined in RFC 3551, clause 4.5.11

L24 24-bit linear format defined in RFC 3190 [9], clause 4

All devices must support 48 kHz sampling rate and are recommended to support 44.1 and 96 kHz sampling rate. With 48 kHz, receivers must support both L16 and L24, senders just need to support L16 or L24 (or both). When operating at 44.1 kHz, devices must support L16; at 96 kHz, support of L24 is mandatory. Devices may support other sampling rate / encoding combinations, but are not forced to support any combination at any given time, i.e. it may require configurative means to switch between different sampling rates.

II. Packet time

Packet time is the real-time duration of the media data contained in a media packet. Given the sampling rate and packet time, the number of samples per packet can be calculated. Shorter packet times allow for lower latency but introduce overhead and high packet rates that may overtax some devices or networks. Longer packet times imply higher latency and require additional buffering which may not be available on memory-constrained devices.

The actual packet time of a given stream is determined by the sender but may be negotiated through connection management. Interoperability is addressed by the requirement that devices must support a 1-millisecond packet time; in order to ease implementation, at 44.1 and 48 kHz devices need to support 48 samples of audio per packet, when operating at 96 kHz devices need to support 96 samples per packet. Further interoperability is encouraged through additional packet time recommendations (see table 2).

Packet time	Samples @ kHz			Notes
	48	**96**	**44.1**	
Required:				
"1 ms"	48	96	48	Required common packet time for all devices
Recommended:				
"125 µs"	6	12	6	Compatible with class A AVB transport
"250 µs"	12	24	12	High-performance, low-latency operation. Interoperable with class A and compatible with class B AVB transport.
"333 µs"	16	32	16	Efficient low-latency operation
"4 ms"	192	n.a.[1]	192	For applications desiring interoperability with EBU Tech 3326 [10], transport over wider areas or on networks with limited QoS capability

TABLE 2: REQUIRED AND RECOMMENDED PACKET TIMES

Senders and receivers may support additional packet times. Maximum packet time is limited by network MTU (see transport layer).

III. Stream Channel Count

Receivers must support reception of streams with 1 to 8 audio channels. Senders need to offer at least one stream with 8 channels or fewer. Streams with more than 8 channels may optionally be supported; however, the maximum number of channels per stream is limited by the packet time, encoding format and network MTU. Table 3 indicates the maximum channels count per stream at different encoding settings:

Format, sampling rate	Packet time	Maximum channels per stream
L24, 48 kHz	125 µs	80
L16, 48 kHz	250 µs	60
L24, 48 kHz	250 µs	40
L24, 48 kHz	333.33 µs	30
L24, 96 kHz	250 µs	20
L24, 48 kHz	1 ms	10
L24, 48 kHz	4 ms	2

TABLE 3: EXAMPLES: MAXIMUM CHANNEL CAPACITIES PER STREAM

IV. Link Offset & Receiver Buffering

Link offset describes the latency through a media network. It is defined as the difference in time between when audio enters the sender (*ingress time*) and when it leaves the receiver (*egress time*).

Ingress time is referenced at ingress to the *sender network system*. RTP packets are marked with origination timestamps

in the *timestamp* field (RFC 3550, clause 5.1) based on this reference point. *Egress time* is referenced at egress from the *receiver network system*. *Link offset* is therefore the time difference between ingress at the sender and egress at the receiver. Link offset and ingress and egress reference points are illustrated in figure 2.

Link offset is determined at the receiver and is dependent on multiple factors, including packet time, packet handling algorithms in the devices and buffering at the receiver. The required buffering at the receiver is implementation and network specific. Minimally buffering must accommodate packet time, packet handling delays in the devices, network delivery delay and delay variation and synchronization skew between sender and receiver. If buffering is too short, data may not arrive in time to be played resulting in audio dropouts. Longer buffering improves robustness but introduces additional latency. Under AES67, receivers must have a minimum buffer capacity of at least 3 times the packet time and are recommend providing a buffer capacity of at least 20 times the packet time or 20 ms, whichever is smaller. These requirements are designed to allow a range of implementations from hardware to applications running on desktop operating systems.

A receiver should attempt to maintain a constant link offset, recognizing that unexpected changes to network conditions may require a change in link offset.

NOTE: The scope of AES67 does not include specific requirements for link offset settings. Future work may specify a latency management mechanism which is expected to require reporting of link offset to a central latency management server on the network and means for receivers to adjust link offset based on commands from the latency management server. See IETF *draft-ietf-avtcore-idms* [11] for possible details.

---

[1] NOTE: 96 kHz is not discussed in EBU Tech 3326. MTU restrictions limit a 96 kHz audio stream using 4-ms packet time to a single channel.
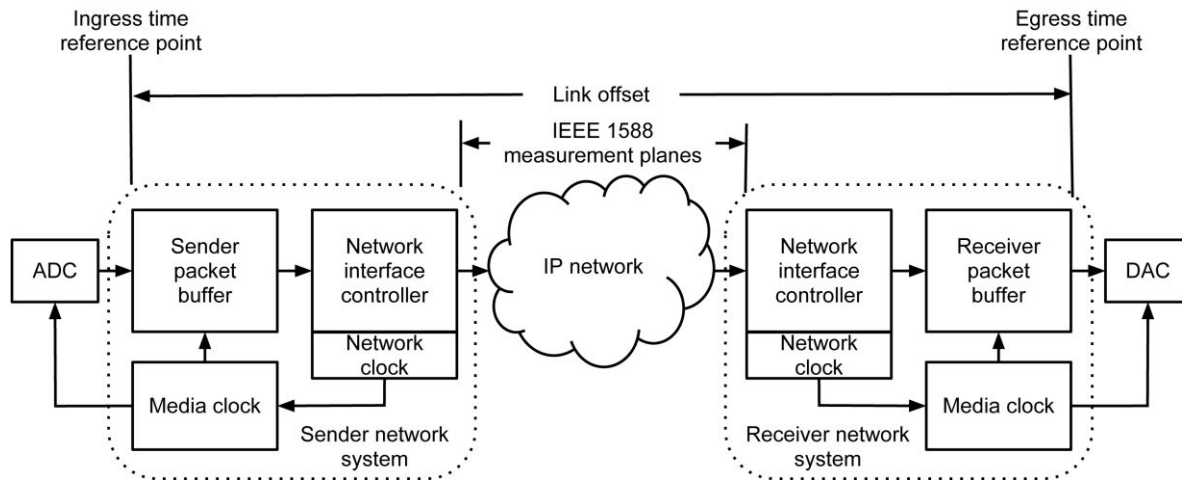
FIG 2: LINK OFFSET AND INGRESS AND EGRESS REFERENCE POINTS

V. Stream Data Multicasting

Multicasting of stream data allows for efficient one-to-many distribution of audio. Multicasting is also an important component in simplified connection management (see below) in which a sender multicasts a stream and receivers can join in at any time and simply listen to the stream in progress.

The multicast destination address used for a particular stream is configured through the management interface at the sender. It is assumed that each stream will be assigned a unique destination address within the administratively scoped multicast addresses in the range `239.0.0.0` to `239.255.255.255`. The nature of the management interface and the allocation scheme used for multicast addresses are outside the scope of the AES67 standard.

STREAM DESCRIPTION

Stream description is used by connection management to specify relevant information about each stream including network addressing, encoding and origination information.

SDP as specified in RFC 4566 [12] is used to represent the relevant stream information for connection management. However, interoperability under this standard imposes additional SDP requirements and recommendations:

I. Packet Time

Packet time is described with two SDP attributes defined in RFC 4566:

    a=ptime <milliseconds>
    a=maxptime <milliseconds>

The ptime attribute indicates the preferred packet time. If more than one packet time is supported, maxptime indicates the maximum permitted packet time. If more than two packet times are supported, the description should use the capability negotiation attributes of RFC 5939 [13] to enumerate the supported packet times in order of preference.

The ‹milliseconds› parameter needs to be given with enough precision so that the equivalent number of samples per packet rounds to the correct integer value when the calculation is done in floating point arithmetic. In some cases, this requires the ‹milliseconds› parameter itself to be given as a floating point number (i.e. a=ptime:0.33)

NOTE: The non-integral-millisecond descriptions may not be correctly understood by non-AES67 devices; in this case the description may need to be confined to integer ‹millisecond› values when attempting connection to such devices.

II. RTP Clock, Media Clock and Clock Source

The clock source and the relationship of media clock to RTP clock are specified for each stream with additional attributes as specified in *draft-ietf-avtcore-clksrc* [14]*:*

    a=ts-refclk:<parms>
    a=mediaclk:<parms>

The ts-refclk attribute specifies the network clock source for a given stream. Ideally, all streams refer to the same reference clock, as only streams with the same reference clock can be synchronously related to each other. Receivers need to reference themselves to the referred clock source in order to properly receive that stream.

In a more complex setup, different clock masters referenced to the same traceable source (i.e. GPS) may be used. This may be signaled by different clock IDs (GMID), but matching PTP domains; in this case, receivers may attempt to connect to senders, but should be prepared for possible synchronization failure. Likewise, in case the clock reference changes due to grandmaster change, receivers should reassess their ability to connect to a given stream.

The mediaclk attribute specifies the relationship between media clock and RTP clock for a given stream. Since the media clock is synchronous to the network clock referenced, the mediaclk attribute usually contains a number describing the offset between network time origin and RTP sample count origin for a given stream.

III. Payload Types

As none of the encoding formats specified within AES67 are defined as static payload types within RFC 3551, an rtpmap attribute as described in RFC 4566 is required for each stream to describe its current encoding formats:

    a=rtpmap:<parms>

IV. Example Descriptions

SDP description for a 24-bit, 48 kHz, 2-channel stream, transmitted as a multicast stream with 1 ms packet time:

    v=0
    o=- 1311738121 1311738121 IN IP4 192.168.1.1
    c=IN IP4 239.0.0.1/32
    s=Stereo PGM
    t=0
    m=audio 5004 RTP/AVP 96
    i=Channels 1-2
    a=rtpmap:96 L24/48000/2
    a=sendonly
    a=ptime:1
    a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:0
    a=mediaclk:direct=963214424

Simple SDP description for a 24-bit, 96 kHz, 8-channel stream, transmitted as a unicast stream with 250 µs packet time:

    v=0
    o=audio 1311738121 1311738121 IN IP4 192.168.1.1
    c=IN IP4 192.168.1.1/32
    s=Stage left I/O
    t=0
    m=audio 5004 RTP/AVP 96
    i=Channels 1-8
    a=rtpmap:96 L24/96000/8
    a=sendonly
    a=ptime:0.250
    a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:0
    a=mediaclk:direct=2216659908

Connection management is the procedure and protocols used to establish a media stream between a sender and one or more receivers. Within AES67, for unicast streams connection management is accomplished using the Session Initiation Protocol (SIP) as defined in RFC 3261 [15]. SIP is widely used in IP telephony and by codec devices utilizing the ACIP protocol as specified in EBU Tech 3326 [10].

However, in a multicast situation, a simple connection mode without using SIP is possible by using IGMP protocol, given that the relevant stream parameters (i.e. multicast address) are known by the receiver (i.e. through a discovery mechanism or administrative means).

I. SIP

Under SIP, audio devices are SIP *user agents* with an associated SIP URI. SIP allows user agents to locate and make connections to other user agents by referencing its SIP URI.

SIP is conventionally used with the assistance and participation of SIP servers. Different types of servers perform different tasks for a SIP network. Servers may be located anywhere on the network where they are reachable by end stations. The use of servers creates a flexible and scalable connection management system.

For modest installations (i.e. with limited scale), where servers produce minimal benefit and the overhead of installing and configuring SIP servers introduces unnecessary complication, the server-less mode is appropriate. Server-less mode is used to perform connection management between user agents in direct peer-to-peer fashion.

In order to perform peer-to-peer connection management, the caller must have some means of determining network contact information (that is, host name or IP address) of the callee. In peer-to-peer connection management, all SIP messages are directed to the target device instead of the server. AES67-compliant devices are required to support server-less operation. However, this does not release devices from requirements to operate in a normal SIP environment featuring servers. Specifically, devices will still need to attempt to discover, and register with, SIP registration servers and respond to messages originating from servers.

NOTE: Server-less mode is discussed in more detail in IETF *draft-lee-sip-dns-sd-uri* [16].

II. Format Negotiation

Where appropriate, the standard offer/answer model as described in RFC 3264 [17] is to be used to negotiate the encoding format for a connection. The offer/answer model does not address negotiation of other attributes such as

packet time. Where multiple packet times are supported, the ptime and maxptime attributes shall be used for negotiation.

OUTLOOK

It is now up to market forces (i.e. customer demand, manufacturers decisions) to accelerate the commercial introduction and acceptance of the AES67 standard. While support of AES67 will facilitate interoperability between different systems or technologies, it is not meant to be a solution on its own. Existing solutions usually offer performance, functionality and flexibility exceeding the ASES67 interoperability definitions, positioning them far from being rendered obsolete. But from observing X192 Task Group participation and contribution, it can be expected that several existing IP-based solutions will

support AES67 in the near future, by either adopting AES67 fundamentals as their native operating mode or by just offering a special stream interchange option; potential technologies and solutions include *RAVENNA* by ALC NetworX, *Livewire* by Axia, *QLAN* by QSC, *WheatNet* by Wheatstone and *Dante* by Audinate. Fig. 3 is illustrating a typical scenario for AES67 application. Limited interoperability can also be expected with devices compliant to the next revision of the EBU Tech 3326 protocol suite, which is currently work in progress in the EBU ACIP2 Work Group. AES67 also include informal notes on how interoperability with AVB networks can be facilitated and specifically honors AVBs IEEE802.1AS synchronization mechanism.
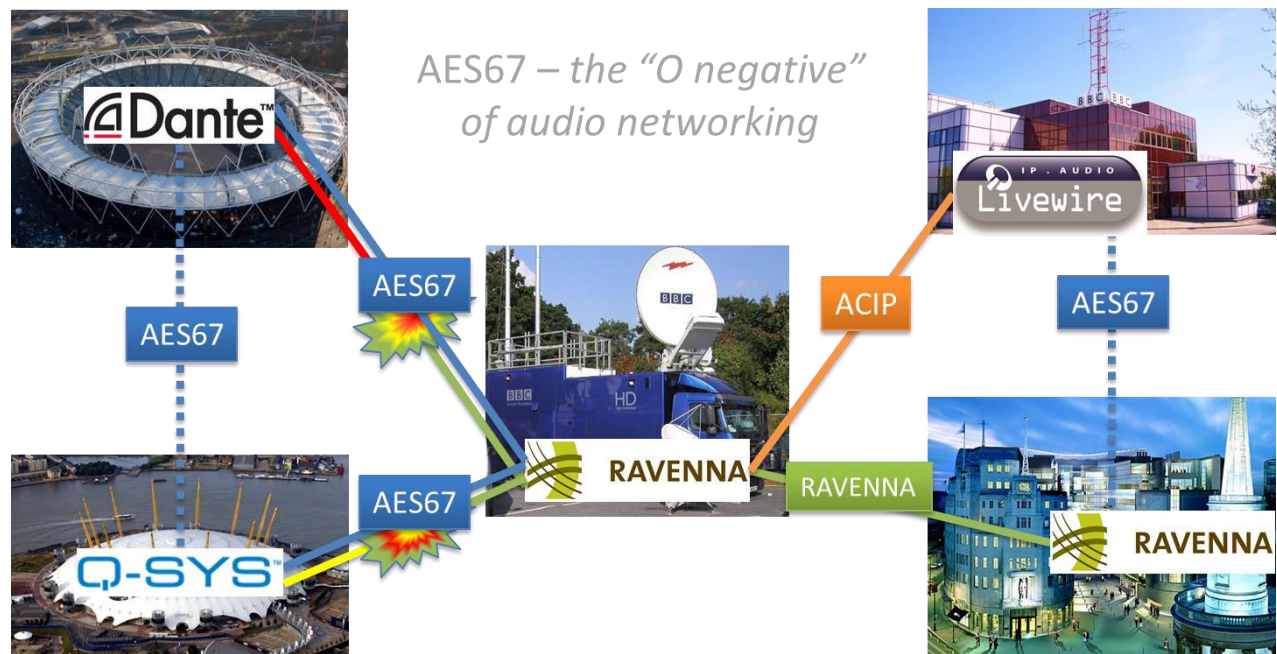

FIG 3: TYPICAL AES67 APPLICATION SCENARIO

GLOSSARY

ACIP – Audio Contribution over IP, a protocol suite defined by EBU Tech 3326

AES – Audio Engineering Society

AVB – Audio Video Bridging, enhanced Ethernet protocol suite for real-time media transport, specified in IEEE 802.1BA, IEEE 802.1Q-2011 and IEEE 802.1AS

BEC – Broadcast Engineering Conference, an educational conference accompanying the NAB show, featuring technical papers addressing the most recent developments in broadcast technology

DiffServ – Differentiated Services, a QoS scheme for classification and management of IP-based network traffic

Discovery – a mechanism for advertisement and detection of network devices and / or services

DSCP – Differentiated Services Code Point, a 6-bit field in the IP packet header that is used for traffic

classification purposes within the Differentiated Services architecture

EBU – European Broadcasting Union

GPS – Global Positioning System

Grandmaster - master source of synchronization for clock distribution via IEEE1588-2008

IEEE – Institute of Electrical and Electronics Engineers

IETF – Internet Engineering Task Force

IGMP – Internet Group Management Protocol, a communications protocol used by hosts and adjacent routers on IP networks to establish multicast group memberships

IPv4 – Internet Protocol version 4, the fourth revision in the development of the Internet Protocol (IP)

IPv6 – Internet Protocol version 6, the latest revision of the Internet Protocol, intended to replace IPv4

ISO – International Organization for Standardization

ITU-T – Telecommunication Standardization Sector of the International Telecommunication Union, one of the three sectors (divisions or units) of the ITU; it coordinates standards for telecommunications

MTU – Maximum Transfer Unit, the maximum size of an IP packet that can be transferred using a specific data link connection; the MTU for an Ethernet data link is 1500 bytes

Multicast – delivery of a message or information to a group of destination computers simultaneously in a single transmission from the source

NAB – National Association of Broadcasters

OSI – Open Systems Interconnection, an effort to standardize networking started in 1977 by the ISO and ITU-T; the OSI reference model is a layered abstraction model characterizing and standardizing the functions of a communications system

Payload – the cargo information within a data transmission

PCM – Pulse-code Modulation, a method to digitally represent sampled analog signals

QoS – Quality of Service, the ability to provide different priority to different data flows, or to guarantee a certain level of performance to a data flow

RFC – Request for Comments, an IETF memorandum on Internet standards and protocols

RTCP – Real-time Control Protocol, a protocol used for providing statistics such as dropped packets and delayed packets about an RTP transmission.

RTP – Real-time Transport Protocol, a standardized packet format for delivering audio and video over IP networks

SDP – Session Description Protocol, a format for description of multimedia streams

SIP – Session Initiation Protocol, a signaling protocol widely used for controlling communication sessions such as voice and video calls over Internet Protocol

Stream – a sequence of data packets sent at a regular interval

TAI – International Atomic Time, a high-precision time standard based on the notional passage of proper time on Earth's geoid; unlike UTC, TAI does not exhibit leap seconds

UDP – User Datagram Protocol, one of the core members of the Internet protocol suite; a simple transport protocol used to send messages to other hosts on an IP network without prior establishment of a special transmission channel or data path

URI – Uniform Resource Identifier, is a string of characters used to identify a name or a resource

UTC – Coordinated Universal Time, the primary time standard by which the world regulates clocks and time; for most purposes, UTC is synonymous with GMT.

REFERENCES

[1] "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems", *IEEE Std 1588™-2008*, July 2008

[2] "RTP: A Transport Protocol for Real-Time Applications", *IETF RFC 3550*, July 2003

[3] "Internet Protocol – DARPA Internet Program Protocol Specification", *IETF RFC 791*, September 1981

[4] "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", *IETF RFC 2474*, September 1981

[5] "RTP Profile for Audio and Video Conferences with Minimal Control", *IETF RFC 3551*, July 2003

[6] "User Datagram Protocol", *IETF RFC 768*, August 1980

[7] "Host Extensions for IP Multicasting", *IETF RFC 1112*, August 1989

[8] "Internet Group Management Protocol, Version 2", *IETF RFC 2236*, November 1997

[9] "RTP Payload Format for 12-bit DAT Audio and 20- and 24-bit Linear Sampled Audio", *IETF RFC 3190*, January 2002

[10] "Audio contribution over IP - Requirements for Interoperability", *EBU Tech 3326*, April 2008

[11] "Inter-destination Media Synchronization using the RTP Control Protocol (RTCP)", *IETF draft-ietf-avtcore-idms*, October 2012

[12] "SDP: Session Description Protocol", *IETF RFC 4566*, July 2006

[13] "Session Description Protocol (SDP) Capability Negotiation", *IETF RFC 5939*, September 2010

[14] "RTP Clock Source Signalling", *IETF draft-ietf-avtcore-clksrc*, July 2012

[15] "SIP: Session Initiation Protocol", *IETF RFC 3261*, June 2002

[16] "SIP URI Service Discovery using DNS-SD", *IETF draft-lee-sip-dns-sd-uri*, July 2008

[17] "An Offer/Answer Model with the Session Description Protocol (SDP)", *IETF RFC 3264*, June 2002

AUTHOR'S INFORMATION

Andreas Hildebrand is acting as Senior Product Manager for the RAVENNA technology developed by ALC NetworX, Germany. He was a full-time participant in the AES X192 Task Group defining the AES67 standard.

Email: andreas.hildebrand@alcnetworx.de
Web: ravenna.alcnetworx.com