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

Published by
Audio Engineering Society, Inc.
Copyright ©2013 by the Audio Engineering Society

Abstract

High-performance media networks support professional quality audio (16 bit, 44,1 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. A number of networked audio systems have been developed to support high-performance media networking but until now there were no recommendations for operating these systems in an interoperable manner. This standard provides comprehensive interoperability recommendations in the areas of synchronization, media clock identification, network transport, encoding and streaming, session description and connection management.

An AES standard implies a consensus of those directly and materially affected by its scope and provisions and is intended as a guide to aid the manufacturer, the consumer, and the general public. The existence of an AES standard does not in any respect preclude anyone, whether or not he or she has approved the document, from manufacturing, marketing, purchasing, or using products, processes, or procedures not in agreement with the standard. Prior to approval, all parties were provided opportunities to comment or object to any provision. Attention is drawn to the possibility that some of the elements of this AES standard or information document may be the subject of patent rights. AES shall not be held responsible for identifying any or all such patents. Approval does not assume any liability to any patent owner, nor does it assume any obligation whatever to parties adopting the standards document. This document is subject to periodic review and users are cautioned to obtain the latest edition. Recipients of this document are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Audio Engineering Society Inc. 60 East 42nd Street, New York, NY., US.
www.aes.org/standards standards@aes.org

2014-03-24 printing

Contents

0 Introduction **5**
0.1 General 5
0.2 Patents 5

1 Scope **6**

2 Normative references **6**

3 Definitions and abbreviations **7**

4 Synchronization **12**
4.0 General 12
4.1 IP network synchronization 12
4.2 IEEE 1588 network synchronization 13
4.3 AVB network synchronization 13

5 Media clocks **13**

6 Transport **13**
6.0 General 13
6.1 Network layer 14
6.2 Quality of service 15
6.3 Transport layer 16

7 Encoding and streaming **16**
7.0 Introduction 16
7.1 Payload format and sampling rate 16
7.2 Packet time 17
7.2.0 General 17
7.2.1 Required packet time 17
7.2.2 Recommended packet times 18
7.3 Stream channel count 18
7.4 Link offset 19
7.5 Sender timing and receiver buffering 20
7.6 Multicasting 20

8 Session description **21**
8.0 General 21
8.1 Packet time 21
8.2 Clock source 22
8.3 RTP and media clock 22
8.4 Payload types 23
8.5 Example descriptions 23
8.5.1 Multicast session description example 23
8.5.2 Unicast session description example 23

9 Discovery **24**

10 Connection management **24**
10.0 General 24
10.1 Unicast connections 24
10.1.1 SIP URI 24
10.1.2 Server and serverless modes 24
10.1.3 User-Agent 25

10.1.4 Format negotiation 25
10.1.5 Packet time negotiation 25
10.2 Multicast connections 25

Annex A (Normative) - Media profile 26
A.0 General 26
A.1 Media profile description 26
A.2 Media profile 26
A.2.1 Identification 26
A.2.2 PTP attribute values 26
A.2.3 PTP options 29
A.2.4 Clock physical requirements 29

Annex B (Informative) - Network QoS configuration recommendations 30
B.0 General 30
B.1 DiffServ network configuration 30
B.1.1 Clock 30
B.1.2 Media 31
B.1.3 Best effort 32

Annex C (Informative) – AVB network transport 33
C.0 General 33
C.1 AVB network transport 33
C.1.1 Interoperable media as AVB time-sensitive streams 33
C.1.2 Interoperable media as other traffic 34

Annex D (Informative) - Interfacing to IEEE 802.1AS clock domains 36
D.0 General 36
D.1 Boundary clock interface 36
D.2 Ordinary clock interface 36
D.3 Traceable reference 36
D.4 AVB network as a boundary clock 37

Annex E (Informative) – Discovery systems 38
E.0 General 38
E.1 Bonjour 38
E.2 SAP 38
E.3 Axia Discovery Protocol 38
E.4 Wheatstone WheatnetIP Discovery Protocol 38

Annex F (Informative) - Informative references 39

Annex G (Informative) - Glossary 40

Foreword

This foreword is not part of the AES67-2013 *AES standard for audio applications of networks - High-performance streaming audio-over-IP interoperability*

This document was developed in project AES-X192, in the SC-02-12-H task group on high-performance streaming audio-over-IP interoperability, under the leadership of Kevin Gross.

Members of the writing group that contributed to this document in draft are: R. Abraham, J. Amate, M. Barbour, C. Becker-Foss, J. Berryman, M. Bishop, H. Blum, J. Boqvist, T. Borland, N. Borthwick, L. Bradshaw, P. Briscoe, L. Brito, J. Britton, C. Broad, N. Brunsgaard, D. Brutzman, E. Bukont Jr., A. Calvanese, R. Camprodon, M. Carter, A. Cedronius, A. Clark, H. Clarke, M. Coinchon, P. Cyrta, S. de Jaham, P. Demuytere, J. Dibley, A. Dickens, P. Dietrich, C. Dodds, S. Dove, A. Eales, E. Echols, R. Economaki, T. Edwards, A. Elder, L. Ellison, K. Fitzke, J. Fletcher, S. Flock, R. Foss, P. Foulkes, F. Gierlinger, R. Goforth, J. Grant, M. Graubner, D. Gravereaux, S. Gray, H. Hansen, B. Harshbarger, S. Heinzmann, A. Hildebrand, D. Hoch, M. Holtmann, Os. Igumbor, H. Jahne, T. Johansson, M. J. Teener, S. Johnson, L. Jonsson, A. Karamustafaoglu, K. Kearney, P. Keller, J. Koftinoff, P. Koftinoff, D. Koss, S. Langhans, M. Lave, D. Lazecko, S. Ledergerber, C. Lefebvre, J. Lindsay, G. Linis, S. Loehberg, A. Louko, A. Makivirta, J. A. Martinez, A. Mayo, W. McQuay, C. Merienne, A. Metz, J. Meunier, J. Meyer, R. Michl, S. Millan, D. O'Gwynn, N. O'Neill, H. Okai-Tettey, K. Parker, J. Passaniti, J. Peavey, J. Perez, W. Peters, S. Price, S. Pro, M. Quaix, F. Ragenard, R. Rayburn, D. Rice, T. Rohwedder, G. Rosenboom, M. Saito, M. S. Carreres, J. Sauter, R. Schoonbroodt, V. Schueppel, P. Schwizer, S. Scott, G. Shay, T.Shuttleworth, D. Silver, J. Simpson, M. Sims, A. Smimite, J. Snow, T. Staros, P. Stevens, J. M. Strawn, N. Sturmel, M. Szlapka, T. Thompson, P. Treleaven, A. Trevena, S. Turner, R. van der Zalm, B. van Kempen, I. Vysick, J. P. Waddell, Y. Wang, P. Warrington, H. Weibel, A. Williams, M. Willsher, A. Witham, J. Wood, J. A. Yeary, J. Yoshio, P. Yurt, U. Zanghieri, D. Zimmermann, R. Zwiebel.

This document was edited by Kevin Gross.

Richard Foss
Chair, working group SC-02-12
2013-07-18

Addendum 2014-03-24
The Introduction has been updated with a clause 0.2 to show that a patent statement has been received under our patent policy.

Note on normative language

In AES standards documents, sentences containing the word “shall” are requirements for compliance with the document. Sentences containing the verb “should” are strong suggestions (recommendations). Sentences giving permission use the verb “may”. Sentences expressing a possibility use the verb “can”.

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

0 Introduction

0.1 General

High-performance media networks support professional quality audio (16 bit, 44,1 kHz and higher) with low latencies (less than 10 ms) 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 but is generally not available on wide-area networks or the public internet.

The most recent generation of these media networks use a diversity of proprietary and standard protocols. Despite a common basis in Internet Protocol, the systems do not interoperate.

This standard provides specific recommendations for interoperability. The standard focuses on defining how existing protocols are used to create an interoperable system. No new protocols have been developed to achieve this.

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.

0.2 Patents

The Audio Engineering Society draws attention to the fact that it is claimed that compliance with this AES standard or information document may involve the use of a patent.

The AES holds no position concerning the evidence, validity and scope of this patent right.

The holder of this patent right has assured the AES that it is willing to negotiate licenses under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statement of the holder of this patent right is archived with the AES.

Information may be obtained from

Audinate Pty Ltd.
PO Box 855
Ultimo, NSW 2007
Australia

Attention is drawn to the possibility that some of the elements of this AES standard or information document may be the subject of patent rights other than those identified above. AES shall not be held responsible for identifying any or all such patent rights.

1 Scope

This standard defines an interoperability mode for transport of high-performance audio over networks based on the Internet Protocol. 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 44,1 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 considers latency performance of 10 milliseconds or less.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

AES11 - *AES recommended practice for digital audio engineering - Synchronization of digital audio equipment in studio operations*; Audio Engineering Society, New York, NY., US.

IEEE 1588-2008 - *IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems*, July 2008, Institute of Electrical and Electronics Engineers (IEEE), US.

RFC 768 – *User Datagram Protocol*", Internet Engineering Task Force

RFC 791 – *Internet Protocol*, Internet Engineering Task Force

RFC 1112 – *Host Extensions for IP Multicasting*, Internet Engineering Task Force

RFC 2236 - *Internet Group Management Protocol, Version 2*, Internet Engineering Task Force

RFC 2474 – *Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers*, Internet Engineering Task Force

RFC 2616 - *Hypertext Transfer Protocol - HTTP/1.1*, Internet Engineering Task Force

RFC 2974 – *Session Announcement Protocol*, Internet Engineering Task Force

RFC 3190 – *RTP Payload Format for 12-bit DAT Audio and 20- and 24-bit Linear Sampled Audio*, Internet Engineering Task Force

RFC 3261 - *SIP: Session Initiation Protocol*, Internet Engineering Task Force

RFC 3264 - *An Offer/Answer Model with the Session Description Protocol (SDP)*, Internet Engineering Task Force

RFC 3376 - *Internet Group Management Protocol, Version 3*, Internet Engineering Task Force

RFC 3550 – *RTP: A Transport Protocol for Real-Time Applications*, Internet Engineering Task Force

RFC 3551 - *RTP Profile for Audio and Video Conferences with Minimal Control*, Internet Engineering Task Force

RFC 4566 – *Session Description Protocol*, Internet Engineering Task Force

draft-ietf-avtcore-clksrc - *RTP Clock Source Signalling* (work in progress) , Internet Engineering Task Force. https://datatracker.ietf.org/doc/draft-ietf-avtcore-clksrc/

NOTE The publication of this IETF draft as an IETF standard is anticipated in the next month or two. When it is published, and assuming no changes that would affect its application in this document, this reference will be updated with a corrigendum.

3 Definitions and abbreviations

For the purposes of this document, the following terms, definitions, and abbreviations apply.

3.1
Audio stream
See RTP stream.

3.2
Audio Video Bridging
AVB
describes enhanced Ethernet networks specified in IEEE 802.1BA, IEEE 802.1Q-2011 and IEEE 802.1AS.

3.3
Boundary Clock
A clock that has multiple Precision Time Protocol (PTP) ports in a domain and maintains the timescale used in the domain. It may serve as the source of time, that is, be a master clock; and may synchronize to another clock, that is, be a slave clock. See IEEE 1588-2008.

3.4
Byte
A unit comprising 8 bits of data. Over IP networks, data is transported in units of bytes.

3.5
Digital Audio Reference Signal
DARS
an audio clock signal defined in AES11.

3.6
CSRC
The contributing source (CSRC) is the source of a stream of RTP packets that has contributed to the combined stream produced by an RTP mixer

3.7
DiffServ
Differentiated services (DiffServ) is a system for classifying traffic and providing quality of service (QoS) on an IP network.

3.8
DSCP
The differentiated services code point (DSCP) is a 6-bit field in the IP packet header that is used for classification purposes. DSCP is part of the differentiated services architecture.

3.9
End-to-end Transparent Clock
A transparent clock that supports the use of the end-to-end delay measurement mechanism between slave clocks and the master clock. See IEEE 1588-2008.