

IEEE Standard for a Transport Protocol for Time-Sensitive Applications in Bridged Local Area Networks

IEEE Computer Society

Sponsored by the
Microprocessor Standards Committee

IEEE
3 Park Avenue
New York, NY 10016-5997
USA

IEEE Std 1722™-2016
(Incorporates
IEEE Std 1722-2016 Cor1-2016
Revision of
IEEE Std 1722-2011)

IEEE Standard for a Transport Protocol for Time-Sensitive Applications in Bridged Local Area Networks

Sponsor

Microprocessor Standards Committee
of the
IEEE Computer Society

Approved 7 December 2016

IEEE-SA Standards Board

Abstract: 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.

Keywords: advanced encryption standard (AES), bridged local area network (LAN), clock recovery, elliptic curve cryptography (ECC), encryption, IEC 61883, IEEE 802.1™ AVB protocols, IEEE 802.1AS™, IEEE 802.1BA™-2011, IEEE 802.1Q™-2014, IEEE 1722™-2011, National Institute of Standards and Technology (NIST), QoS, serial digital interface (SDI), SMPTE ST 2022-6, time-sensitive media streaming, Time-Sensitive Networking (TSN), time synchronization

The Institute of Electrical and Electronics Engineers, Inc.
3 Park Avenue, New York, NY 10016-5997, USA

Copyright © 2016 by The Institute of Electrical and Electronics Engineers, Inc.
All rights reserved. Published 16 December 2016. Printed in the United States of America.

IEEE is a registered trademark in the U.S. Patent & Trademark Office, owned by The Institute of Electrical and Electronics Engineers, Incorporated.

PDF: ISBN 978-1-5044-0802-8 STD20865
Print: ISBN 978-1-5044-0803-5 STDPD20865

IEEE prohibits discrimination, harassment, and bullying.

For more information, visit <http://www.ieee.org/web/aboutus/whatis/policies/p9-26.html>.

No part of this publication may be reproduced in any form, in an electronic retrieval system or otherwise, without the prior written permission of the publisher.

Important Notices and Disclaimers Concerning IEEE Standards Documents

IEEE documents are made available for use subject to important notices and legal disclaimers. These notices and disclaimers, or a reference to this page, appear in all standards and may be found under the heading “Important Notice” or “Important Notices and Disclaimers Concerning IEEE Standards Documents.”

Notice and Disclaimer of Liability Concerning the Use of IEEE Standards Documents

IEEE Standards documents (standards, recommended practices, and guides), both full-use and trial-use, are developed within IEEE Societies and the Standards Coordinating Committees of the IEEE Standards Association (“IEEE-SA”) Standards Board. IEEE (“the Institute”) develops its standards through a consensus development process, approved by the American National Standards Institute (“ANSI”), which brings together volunteers representing varied viewpoints and interests to achieve the final product. Volunteers are not necessarily members of the Institute and participate without compensation from IEEE. While IEEE administers the process and establishes rules to promote fairness in the consensus development process, IEEE does not independently evaluate, test, or verify the accuracy of any of the information or the soundness of any judgments contained in its standards.

IEEE does not warrant or represent the accuracy or content of the material contained in its standards, and expressly disclaims all warranties (express, implied and statutory) not included in this or any other document relating to the standard, including, but not limited to, the warranties of: merchantability; fitness for a particular purpose; non-infringement; and quality, accuracy, effectiveness, currency, or completeness of material. In addition, IEEE disclaims any and all conditions relating to: results; and workmanlike effort. IEEE standards documents are supplied “AS IS” and “WITH ALL FAULTS.”

Use of an IEEE standard is wholly voluntary. The existence of an IEEE standard does not imply that there are no other ways to produce, test, measure, purchase, market, or provide other goods and services related to the scope of the IEEE standard. Furthermore, the viewpoint expressed at the time a standard is approved and issued is subject to change brought about through developments in the state of the art and comments received from users of the standard.

In publishing and making its standards available, IEEE is not suggesting or rendering professional or other services for, or on behalf of, any person or entity nor is IEEE undertaking to perform any duty owed by any other person or entity to another. Any person utilizing any IEEE Standards document, should rely upon his or her own independent judgment in the exercise of reasonable care in any given circumstances or, as appropriate, seek the advice of a competent professional in determining the appropriateness of a given IEEE standard.

IN NO EVENT SHALL IEEE BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO: PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE PUBLICATION, USE OF, OR RELIANCE UPON ANY STANDARD, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE AND REGARDLESS OF WHETHER SUCH DAMAGE WAS FORESEEABLE.

Translations

The IEEE consensus development process involves the review of documents in English only. In the event that an IEEE standard is translated, only the English version published by IEEE should be considered the approved IEEE standard.

Official statements

A statement, written or oral, that is not processed in accordance with the IEEE-SA Standards Board Operations Manual shall not be considered or inferred to be the official position of IEEE or any of its committees and shall not be considered to be, or be relied upon as, a formal position of IEEE. At lectures, symposia, seminars, or educational courses, an individual presenting information on IEEE standards shall make it clear that his or her views should be considered the personal views of that individual rather than the formal position of IEEE.

Comments on standards

Comments for revision of IEEE Standards documents are welcome from any interested party, regardless of membership affiliation with IEEE. However, IEEE does not provide consulting information or advice pertaining to IEEE Standards documents. Suggestions for changes in documents should be in the form of a proposed change of text, together with appropriate supporting comments. Since IEEE standards represent a consensus of concerned interests, it is important that any responses to comments and questions also receive the concurrence of a balance of interests. For this reason, IEEE and the members of its societies and Standards Coordinating Committees are not able to provide an instant response to comments or questions except in those cases where the matter has previously been addressed. For the same reason, IEEE does not respond to interpretation requests. Any person who would like to participate in revisions to an IEEE standard is welcome to join the relevant IEEE working group.

Comments on standards should be submitted to the following address:

Secretary, IEEE-SA Standards Board
445 Hoes Lane
Piscataway, NJ 08854 USA

Laws and regulations

Users of IEEE Standards documents should consult all applicable laws and regulations. Compliance with the provisions of any IEEE Standards document does not imply compliance to any applicable regulatory requirements. Implementers of the standard are responsible for observing or referring to the applicable regulatory requirements. IEEE does not, by the publication of its standards, intend to urge action that is not in compliance with applicable laws, and these documents may not be construed as doing so.

Copyrights

IEEE draft and approved standards are copyrighted by IEEE under U.S. and international copyright laws. They are made available by IEEE and are adopted for a wide variety of both public and private uses. These include both use, by reference, in laws and regulations, and use in private self-regulation, standardization, and the promotion of engineering practices and methods. By making these documents available for use and adoption by public authorities and private users, IEEE does not waive any rights in copyright to the documents.

Photocopies

Subject to payment of the appropriate fee, IEEE will grant users a limited, non-exclusive license to photocopy portions of any individual standard for company or organizational internal use or individual, non-commercial use only. To arrange for payment of licensing fees, please contact Copyright Clearance Center, Customer Service, 222 Rosewood Drive, Danvers, MA 01923 USA; +1 978 750 8400. Permission to photocopy portions of any individual standard for educational classroom use can also be obtained through the Copyright Clearance Center.

Updating of IEEE Standards documents

Users of IEEE Standards documents should be aware that these documents may be superseded at any time by the issuance of new editions or may be amended from time to time through the issuance of amendments, corrigenda, or errata. An official IEEE document at any point in time consists of the current edition of the document together with any amendments, corrigenda, or errata then in effect.

Every IEEE standard is subjected to review at least every ten years. When a document is more than ten years old and has not undergone a revision process, it is reasonable to conclude that its contents, although still of some value, do not wholly reflect the present state of the art. Users are cautioned to check to determine that they have the latest edition of any IEEE standard.

In order to determine whether a given document is the current edition and whether it has been amended through the issuance of amendments, corrigenda, or errata, visit the IEEE-SA Website at <http://ieeexplore.ieee.org/xpl/standards.jsp> or contact IEEE at the address listed previously. For more information about the IEEE-SA or IEEE's standards development process, visit the IEEE-SA Website at <http://standards.ieee.org>.

Errata

Errata, if any, for all IEEE standards can be accessed on the IEEE-SA Website at the following URL: <http://standards.ieee.org/findstds/errata/index.html>. Users are encouraged to check this URL for errata periodically.

Patents

Attention is called to the possibility that implementation of this standard may require use of subject matter covered by patent rights. By publication of this standard, no position is taken by the IEEE with respect to the existence or validity of any patent rights in connection therewith. If a patent holder or patent applicant has filed a statement of assurance via an Accepted Letter of Assurance, then the statement is listed on the IEEE-SA Website at <http://standards.ieee.org/about/sasb/patcom/patents.html>. Letters of Assurance may indicate whether the Submitter is willing or unwilling to grant licenses under patent rights without compensation or under reasonable rates, with reasonable terms and conditions that are demonstrably free of any unfair discrimination to applicants desiring to obtain such licenses.

Essential Patent Claims may exist for which a Letter of Assurance has not been received. The IEEE is not responsible for identifying Essential Patent Claims for which a license may be required, for conducting inquiries into the legal validity or scope of Patents Claims, or determining whether any licensing terms or conditions provided in connection with submission of a Letter of Assurance, if any, or in any licensing agreements are reasonable or non-discriminatory. Users of this standard are expressly advised that determination of the validity of any patent rights, and the risk of infringement of such rights, is entirely their own responsibility. Further information may be obtained from the IEEE Standards Association.

Participants

At the time this IEEE standard was completed, the Audio/Video Bridging Layer2 Transport Working Group had the following membership:

David Olsen, *Chair*
Matthew Xavier Mora, *Vice Chair*
Don Pannell, *Secretary*
Aaron Gelter, *Editor*
Jeffrey Daniel Koftinoff, *Editor*

Clause Editors
Gordon Bechtel, *Clause 9, Clause 10*
Rob Silfvast, *Clause 11, Clause 12*
Ethan Grossman, *Clause 8, Annex K*
Willem-Jan Dirks, *Clause 12, Clause 13, Annex J*
Ashley Butterworth, *Clause 15, Annex I*

Frank Bähren
Gordon Bechtel
Alex Belits
Robert Boatright
Kyle Carbone
Tim Carmean
Rodney Cummings
Rob Currie
Jim Czekaj
Edison David
Rebeca Delgado
Jan Eveleens
Nandhini Ganesan
Jan-Willem de Groot
Bob Edge
Thomas Edwards
Don Franks

Amanda Ge
Craig Gunther
Robbie Gurdan
Oliver Hartkopp
Chris Healy
Joern Henneberg
Clint Herran
Mark Hu
Jeffrey Hutchings
Girault Jones Jr.
Haysam Kadry
Andreas Kern
Jochen Klaus-Wagenbrenner
Andre Labelle
Paul Lambert
Jim Lawlis
Herbert Leuwer
Steve Matovski

Greg Nelson
Josef Nöbauer
Chris Pane
Greg Schlechter
Harkirat Singh
Jose Solis-Marquez
Simon Song
Kevin Stanton
Thilo Streichert
Ryan Tabrah
Michael Johas Teener
Max Turner
Stephen Turner
Kieran Tyrrell
Yu-Ting Wu
Mike Ye
Helge Zinner

The following members of the individual balloting committee voted on this standard. Balloters may have voted for approval, disapproval, or abstention.

Stefan Aust
Gordon Bechtel
Christian Boiger
Ashley Butterworth
Juan Carreon
Yesenia Cevallos
Minho Cheong
Keith Chow
Rodney Cummings
Richard Edgar
Yukihiro Fujimoto
Aaron Gelter
James Gilb
David Gregson
Ethan Grossman
Randall Groves
Craig Gunther

Marco Hernandez
Werner Hoelzl
Noriyuki Ikeuchi
Atsushi Ito
Raj Jain
Michael Johas Teener
Adri Jovin
Piotr Karocki
Stuart Kerry
Yongbum Kim
Jeff Koftinoff
Arthur H. Light
Michael Lynch
Syam Madanapalli
Roger Marks
Arthur Marris
Matthew Mora

David Olsen
Chris Pane
Arumugam Paventhan
Venkatesha Prasad
Alon Regev
Bartien Sayogo
Stephen Schwarm
Thomas Starai
Eugene Stoudenmire
Walter Struppler
David Thompson
Kieran Tyrrell
Dmitri Varsanofiev
Khurram Waheed
Hung-Yu Wei
Oren Yuen
Daidi Zhong

The following members of the individual balloting committee voted on the corrigendum. Balloters may have voted for approval, disapproval, or abstention.

Thomas Alexander
Stefan Aust
Gordon Bechtel
Ashley Butterworth
Juan Carreon
David Chalupsky
Keith Chow
Richard Doyle
Marc Emmelmann
Yukihiro Fujimoto
Aaron Gelter
James Gilb
Joel Goergen
Ethan Grossman
Randall Groves
Craig Gunther

Marco Hernandez
Werner Hoelzl
Noriyuki Ikeuchi
Atsushi Ito
Michael Johas Teener
Piotr Karocki
Ralph Kearfott
Stuart Kerry
Yongbum Kim
Jeff Koftinoff
James Lepp
Jon Lewis
Arthur Marris
Michael Montemurro
Matthew Mora
Michael Newman

Paul Nikolich
David Olsen
Stephen Palm
Venkatesha Prasad
Alon Regev
Maximilian Riegel
Frank Schewe
Veselin Skendzic
Thomas Starai
Walter Struppler
David Thompson
Kieran Tyrrell
Stephen Webb
Colin Whitby-Stevens
Oren Yuen
George Zimmerman

When the IEEE-SA Standards Board approved this standard on 7 December 2016, it had the following membership:

John D. Kulick, *Chair*
Jon Walter Rosdahl, *Vice Chair*
Richard H. Hulett, *Past Chair*
Konstantinos Karachalios, *Secretary*

Masayuki Ariyoshi
Ted Burse
Stephen Dukes
Jean-Philippe Faure
J. Travis Griffith
Gary Hoffman
Michael Janezic

Joseph L. Koepfinger*
David J. Law
Hung Ling
Andrew Myles
T. W. Olsen
Glenn Parsons
Ronald C. Petersen
Annette D. Reilly

Stephen J. Shellhammer
Adrian P. Stephens
Yatin Trivedi
Philip Winston
Don Wright
Yu Yuan
Daidi Zhong

*Member Emeritus

Introduction

This introduction is not part of IEEE Std 1722-2016, IEEE Standard for a Transport Protocol for Time-Sensitive Applications in Bridged Local Area Networks.

Increasingly, entertainment media are digitally transported. Streaming audio/video and interactive applications over local area networks is becoming more common.

This standard builds on the work done by the IEEE 802.1 AVB task group by providing a common audio/video transport protocol capable of supporting the needs of both consumer and professional audio/video applications.

Contents

1. Overview	1
1.1 Scope	1
1.2 Purpose	2
1.3 Notes about this revision	2
2. Normative references	3
3. Terms, definitions, and notation	5
3.1 Definitions	5
3.2 Reserved fields	8
3.3 Numerical values	8
3.4 Notation of fields and values taken from other documents	8
3.5 Acronyms and abbreviations	11
4. AVTP base protocol	13
4.1 Overview	13
4.2 AVTP network requirements	13
4.3 Timing and synchronization	14
4.4 Header formats	20
5. IEC 61883/IIDC Format	29
5.1 Overview	29
5.2 IEC 61883/IIDC stream data encapsulation	29
5.3 IIDC encapsulation	32
5.4 IEC 61883 encapsulation	39
6. MMA Format	48
7. AVTP Audio Format	48
7.1 Overview	48
7.2 AAF common stream data encapsulation	48
7.3 AAF PCM stream data encapsulation	51
7.4 AAF AES3 stream data encapsulation	56
7.5 AAF timestamps	60
7.6 AAF packetization	61
7.7 AAF and SRP	61
8. Compressed Video Format	63
8.1 Overview	63
8.2 AVTP Presentation Time handling	63
8.3 Compressed Video Stream data encapsulation	64
8.4 MJPEG Video Format	66
8.5 H.264 Compressed Video Format	69
8.6 JPEG 2000 Compressed Video Format	73
9. AVTP Control Format	77
9.1 Overview	77
9.2 Non-Time-Synchronous Control Format header	78
9.3 Time-Synchronous Control Format header	79
9.4 ACF messages	81

10. Clock Reference Format	110
10.1 Overview	110
10.2 Background.....	110
10.3 Audio example.....	110
10.4 Clock Reference Format Data encapsulation.....	112
10.5 CRF transmission rates	117
10.6 Interaction between CRF timestamps and Media Stream timestamps	118
10.7 Requirements of CRF timestamp values.....	119
10.8 Relationship between CRF timestamps and stream timestamps.....	120
11. SDI Video Format	121
11.1 Overview	121
11.2 SDI Video Stream data encapsulation	121
12. Raw Video Format	129
12.1 Overview	129
12.2 Raw Video Stream data encapsulation	130
13. AES Encrypted Format.....	139
13.1 Overview	139
13.2 AEF data encapsulation	140
13.3 AES-SIV	141
14. Vendor Specific Format.....	142
14.1 Overview	142
15. Experimental Format	144
15.1 Overview	144
16. ECC Signed Control Format.....	144
16.1 Overview	144
16.2 ECC Signed Control Data encapsulation.....	145
16.3 ECSSA DSA EMSA1 SHA256 signature variant	147
17. ECC Encrypted Control Format	148
17.1 Overview	148
17.2 ECC Encrypted Control AVTPDU.....	148
17.3 Elliptic curve variant 1 algorithm	149
Annex A (informative) Bibliography	151
Annex B (normative) MAC Address Acquisition Protocol.....	153
B.1 Overview.....	153
B.2 Protocol Message Format.....	154
B.3 Protocol specification.....	157
B.4 Reserved MAAP MAC addresses	163
Annex C (informative) IEEE 802.3 Media-specific encapsulation	164
C.1 Introduction.....	164
C.2 IEEE Std 802.3 encapsulation.....	164
C.3 Other considerations	168
Annex D (normative) MA-L usage	169
Annex E (informative) Diagnostics and counters.....	170
E.1 Audio/Video Transport Protocol (AVTP) stream counters	170
E.2 AVTP counter definitions	171

Annex F (normative) Protocol Implementation Conformance Statement (PICS) proforma.....	173
F.1 Introduction	173
F.2 Abbreviations and special symbols	173
F.3 Instructions for completing the PICS proforma.....	174
F.4 PICS proforma for IEEE Std 1722	176
F.5 Major capabilities and options.....	177
F.6 AVTP Protocol	177
F.7 61883 Protocol	178
F.8 AVTP Audio Format	179
F.9 MMA Format	180
F.10 Compressed Video Format.....	181
F.11 AVTP Control Format.....	182
F.12 Clock Reference Format.....	184
F.13 SDI Video Format	185
F.14 Raw Video Format	186
F.15 AES Encrypted Format.....	187
F.16 Vendor Specific Format	187
F.17 ECC Signed Control Format	187
F.18 ECC Encrypted Control Format	188
F.19 MAC Address Acquisition Protocol.....	189
F.20 IP Encapsulation.....	189
Annex G (informative) Example event field usage	190
G.1 Overview	190
G.2 Synchronized format changes in a automotive environment	190
G.3 Synchronized external event indication	191
Annex H (informative) Content protection.....	192
Annex I (normative) AVDECC stream formats	193
I.1 Overview	193
I.2 Stream formats	193
Annex J (normative) IP encapsulation.....	210
J.1 Overview	210
J.2 Encapsulation styles.....	210
J.3 IP AVTPDU Format	210
J.4 Source port.....	211
J.5 Destination port.....	211
J.6 Destination address	212
J.7 Bandwidth reservation	212
J.8 Clock synchronization.....	212
Annex K (informative) AES3 and AVTP Audio Format	213
K.1 Definitions	213
K.2 AES3 block start.....	215
K.3 Bit ordering in AES3/AES10 compared with AES47/AAF.....	215
K.4 Audio sample, preamble, and ancillary data field order in AAF	216
K.5 The Validity (V) bit	216
K.6 The Parity (P) bit	216
K.7 Bit ordering of ancillary data	217
K.8 SMPTE ST 337 data bursts.....	217
K.9 Using AVDECC to convey AES3 stream data types.....	217
K.10 Other considerations	219
K.11 Other references.....	220

IEEE Standard for a Transport Protocol for Time-Sensitive Applications in Bridged Local Area Networks

IMPORTANT NOTICE: IEEE Standards documents are not intended to ensure safety, security, health, or environmental protection, or ensure against interference with or from other devices or networks. Implementers of IEEE Standards documents are responsible for determining and complying with all appropriate safety, security, environmental, health, and interference protection practices and all applicable laws and regulations.

This IEEE document is made available for use subject to important notices and legal disclaimers. These notices and disclaimers appear in all publications containing this document and may be found under the heading “Important Notice” or “Important Notices and Disclaimers Concerning IEEE Documents.” They can also be obtained on request from IEEE or viewed at <http://standards.ieee.org/IPR/disclaimers.html>.

1. Overview

Increasingly, entertainment media are digitally transported. Streaming audio/video and interactive applications over bridged local area networks (LANs) need to have comparable real-time performance with legacy analog distribution. There is significant end-user and vendor interest in defining a simple yet common method for handling real-time audio/video suitable for consumer electronics, professional audio/video applications, etc. Technologies such as IEEE Std 1394™,¹ ² Bluetooth®, and USB exist today, but each has their own encapsulation, protocols, timing control, etc., such that building interworking functions is difficult. The use of a common audio/video transport over multiple IEEE 802™ network types will realize operational and equipment cost benefits. By ensuring that all IEEE 802 wired and wireless devices share a common set of transport mechanisms for time-sensitive audio, video, and control streams, the effort of producing interworking units between IEEE 802 devices and other digital networks is lessened.

1.1 Scope

This standard specifies the protocol, data encapsulations, and synchronization procedures used to enable interoperability between time-sensitive audio, video, and control applications using the quality of service capabilities provided by IEEE 802 Time-Sensitive Networking standards.

¹The IEEE standards or products referred to in this clause are trademarks of The Institute of Electrical and Electronics Engineers, Inc.

²IEEE publications are available from The Institute of Electrical and Electronics Engineers, Inc., 445 Hoes Lane, Piscataway, NJ 08854, USA (<http://standards.ieee.org/>).

1.2 Purpose

This standard facilitates interoperability between end stations that transport time-sensitive media across LANs providing time synchronization, latency, and bandwidth services by defining additional packet format protocols, synchronization mechanisms, and diagnostic counters.

1.3 Notes about this revision

1.3.1 Compatibility

Although some field definitions have changed from those defined in the previous version of this standard, this revision of the document is fully backward-compatible with the previous version. For example, the **cd** bit of the Audio Video Transport Protocol Data Units (AVTPDU) common header (see 4.4.3) has been removed, but formats that used this bit in the previous revision have been updated to have the most-significant bit of the **subtype** field set, retaining binary compatibility.

1.3.2 Use of TSN and AVB

With the expansion of use cases to industrial control and other applications that are not audio or video in nature, updated standards have been created that collectively use the moniker Time-Sensitive Networking (TSN). In this document, AVB is considered a subset of the TSN standards, and as such, both names are used. The term AVB is used for audio and video specific applications for which the term Audio Video Bridging (AVB) was originally envisioned, and TSN is used for formats that are not audio or video in nature.

1.3.3 New additions

In addition to the formats defined in the previous version of this standard, this revision adds new formats that add features that were not present in the previous revision or allow for easier implementation of common use cases. A summary of the additions present in this revision is given in Table 1.

Table 1—New additions in this revision

Addition	Description	Reference
AES-SIV Encrypted Format	Provides a method for encrypting Audio/Video Transport Protocol (AVTP) streams using Advanced Encryption Standard–Synthetic Initialization Vector (AES-SIV).	Clause 11
Audio/Video Discovery, Enumeration, Connection Management, and Control (AVDECC) stream formats	Annex that describes stream formats that were not available when IEEE Std 1722.1™-2013 was published.	Annex I
AVTP Audio Format	Provides a mechanism for transporting PCM and AES3 audio data over an AVTP network that is optimized for minimum overhead and maximum utilization of stream reservations. AVTP audio also allows for higher channel counts in streams due to lower header overhead compared to the IEC 61883-6 method (see Clause 5).	Clause 7
Clock Reference Format	Allows for the dissemination of timing information to devices using TSN without the use of a media stream.	Clause 10
Compressed Video Format	Makes use of video payload definitions that are defined in Internet Engineering Task Force (IETF) request for comments (RFCs).	Clause 8
Content protection	Annex that gives an overview of using High-bandwidth Digital Content Protection (HDCP) over an AVTP network.	Annex H
Control formats	Allows for both synchronous and non-synchronous delivery of a variety of control messages over TSN.	Clause 9
Diagnostics and counters	Annex that provides suggestions on counters that could be useful for debug.	Annex E
ECC Encrypted Format	Provides a method for encrypting AVTP streams using ECC.	Clause 17
ECC Signed Control Format	Provides a method for signing AVTP streams using ECC.	Clause 16
Event field example usage	Annex that describes possible usages of the application-specific event field bits.	Annex G
Experimental format	A mechanism for testing and developing future enhancements to the IEEE 1722 family of standards.	Clause 15
IP encapsulation	Annex that describes a method for encapsulating any AVTPDU in a User Datagram Protocol (UDP) datagram on an Internet Protocol (IP) network.	Annex J
MA-L usage	Annex that outlines the allocation of addresses within the AVTP-assigned MA-L.	Annex D
PICS proforma	A list of requirements that a device must adhere to in order to be deemed compliant with the standard.	Annex F
Raw Video Format	Provides a direct method for transporting video data (active video only) over an AVTP network.	Clause 12
SDI Video Format	Provides a method for transporting direct mappings of commonly used SDI data streams over an AVTP network.	Clause 11
Vendor Specific Format	Provides a method for vendors to transport proprietary stream formats over AVTP.	Clause 14

2. Normative references

The following referenced documents are indispensable for the application of this document (i.e., they must be understood and used, so each referenced document is cited in text and its relationship to this document is explained). For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments or corrigenda) applies.

AES3-1-2009, AES Standard for Digital Audio—Digital Input-Output Interfacing—Serial Transmission Format for Two-Channel Linearly-Represented Digital Audio Data—Part 1: Audio Content.

AES3-2-2009, AES Standard for Digital Audio—Digital Input-Output Interfacing—Serial Transmission Format for Two-Channel Linearly-Represented Digital Audio Data—Part 2: Metadata and Subcode.

AES3-3-2009, AES Standard for Digital Audio—Digital Input-Output Interfacing—Serial Transmission Format for Two-Channel Linearly-Represented Digital Audio Data—Part 3: Transport.

AES47-2006 (r2011), AES Standard for Digital Audio—Digital Input-Output Interfacing—Transmission of Digital Audio over Asynchronous Transfer Mode (ATM) Networks.

IEC 61883-1:2008, Consumer Audio/Video Equipment—Digital Interface—Part 1: General.³

IEC 61883-4:2004, Consumer Audio/Video Equipment—Digital Interface—Part 4: MPEG2-TS Data Transmission.

IEC 61883-6:2014, Consumer Audio/Video Equipment—Digital Interface—Part 6: Audio and Music Data Transmission Protocol.

IEC 61883-7:2003, Consumer Audio/Video Equipment—Digital Interface—Part 7: Transmission of ITU-R BO.1294 System B.

IEC 61883-8:2013, Consumer Audio/Video Equipment—Digital Interface—Part 8: Transmission of ITU-R Bt.601 Style Digital Video Data.

IEEE Std 754™-2008, IEEE Standard for Floating-Point Arithmetic.^{4, 5}

IEEE Std 802.1AS™-2011, IEEE Standard for Local and Metropolitan Area Networks—Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks.

IEEE Std 802.1Q™-2014, IEEE Standard for Local and Metropolitan Area Networks—Bridges and Bridged Networks.

IEEE Std 1363a™-2004, IEEE Standard Specifications for Public-Key Cryptography—Amendment 1: Additional Techniques.

IEEE Std 1394™, IEEE Standard for High-Performance Serial Bus.

IEEE Std 1722.1™-2013, IEEE Standard for Device Discovery, Connection Management, and Control Protocol for IEEE 1722™-Based Devices.

IIDC2 Digital Camera Control Specification Ver.1.0.0, Jan. 2012.⁶

NIST FIPS-197, Specification for the Advanced Encryption Standard (AES).⁷

RFC 2435, RTP Payload Format for JPEG-Compressed Video.⁸

RFC 5297, Synthetic Initialization Vector (SIV) Authenticated Encryption Using the Advanced Encryption Standard (AES).

RFC 5371, RTP Payload Format for JPEG 2000 Video Streams.

³IEC publications are available from the Sales Department of the International Electrotechnical Commission, Case Postale 131, 3 rue de Varembe, CH-1211, Genève 20, Switzerland/Suisse (<http://www.iec.ch/>). IEC publications are also available in the United States from the Sales Department, American National Standards Institute, 11 West 42nd Street, 13th Floor, New York, NY 10036, USA.

⁴IEEE publications are available from The Institute of Electrical and Electronics Engineers, 445 Hoes Lane, Piscataway, NJ 08854, USA (<http://standards.ieee.org/>).

⁵The IEEE standards or products referred to in this clause are trademarks owned by The Institute of Electrical and Electronics Engineers, Incorporated.

⁶Available at http://jiia.org/wp-content/themes/jiia/pdf/standard_dl/ngcp/CP-001-2011_1394TA_TS2011001.pdf.

⁷Available from <http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf>

⁸RFC documents are a product of the IETF (Internet Engineering Task Force) and are available from <http://www.rfc-editor.org>.

RFC 6184, RTP Payload Format for H.264 Video.

SMPTE ST 259:2008, Television—SDTV Digital Signal/Data—Serial Digital Interface.⁹

SMPTE ST 274:2008, Television—1920 × 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates.

SMPTE ST 292-1:2012, 1.5 Gb/s Signal/Data Serial Interface.

SMPTE ST 296:2012, 1280 × 720 Progressive Image 4:2:2 and 4:4:4 Sample Structure—Analog and Digital Representation and Analog Interface.

SMPTE ST 337:2008, Format for Non-PCM Audio and Data in an AES3 Serial Digital Audio Interface.

SMPTE ST 338:2015, Format for Non-PCM Audio and Data in AES3—Data Types.

SMPTE ST 425-1:2011, Source Image Format and Ancillary Data Mapping for the 3 Gb/s Serial Interface.

SMPTE ST 2022-6:2012, Transport of High Bit Rate Media Signals over IP Networks (HBRMT).

3. Terms, definitions, and notation

3.1 Definitions

For the purposes of this document, the following terms and definitions apply. The *IEEE Standards Dictionary Online* should be consulted for terms not defined in this clause.¹⁰

4:1:1, 4:2:0, 4:2:2, 4:4:4: Denotes the proportion of luma (Y) to chroma (U and V) components. For example in 4:1:1 there is one U and one V for every four Y components.

4:2:2:4, 4:4:4:4: Denotes the proportion of luma (Y) to chroma (U and V) components similar to 4:2:2 and 4:4:4, but with an extra channel for keying.

audio sample frame: An audio sample frame is a set of audio samples in a stream that are intended to be processed at the same instant. For example in a stereo audio stream an audio sample frame consists of two audio samples.

audio/video bridging (AVB): A set of specifications that allow time-synchronized, low-latency streaming services through IEEE 802 networks, as defined by IEEE Std 802.1BA™. Specifically, this includes the generalized Precision Time Protocol (gPTP) as defined in IEEE Std 802.1AS™, as well as the Stream Reservation Protocol (SRP) and the Forwarding and Queuing Enhancements for Time-Sensitive Streams (FQTSS) as defined in IEEE Std 802.1Q™. For the purposes of Audio/Video Transport Protocol (AVTP), AVB and Time-Sensitive Networking (TSN) are equivalent.

audio/video bridging (AVB) bridge: A bridge that implements AVB standards. *See* **audio/video bridging (AVB)**.

⁹SMPTE publications are available at <http://standards.smpte.org>.

¹⁰*IEEE Standards Dictionary Online* subscription is available at:
http://www.ieee.org/portal/innovate/products/standard/standards_dictionary.html.

Audio/Video Discovery, Enumeration, Connection Management, and Control (AVDECC): The protocol defined by IEEE Std 1722.1™-2013. As an adjective, it indicates that the modified noun is specified or interpreted in the context of IEEE Std 1722.1-2013.

Audio/Video Transport Protocol (AVTP): The protocol defined by this standard. As an adjective, it indicates that the modified noun is specified in or interpreted in the context of this standard. This protocol includes transportation of data beyond audio and video.

AVDECC Connection Management Protocol: The Audio/Video Discovery, Enumeration, Connection Management, and Control (AVDECC) Connection Management portion of IEEE Std 1722.1-2013 handles the process of making and breaking connections between stream sinks and stream sources.

AVDECC Discovery Protocol: The Audio/Video Discovery, Enumeration, Connection Management, and Control (AVDECC) Discovery portion of IEEE Std 1722.1-2013 handles advertising and discovery on a network.

AVDECC Enumeration and Control Protocol: The Audio/Video Discovery, Enumeration, Connection Management, and Control (AVDECC) Enumeration and Control portion of IEEE Std 1722.1-2013 handles enumeration and control of capabilities, formats, and controls.

AVTP gateway: A device that transports Audio/Video Transport Protocol (AVTP) streams between Time-Sensitive Network and another type of media, for example, an IEEE 1394™ bus.

AVTP stream: An Audio/Video Transport Protocol (AVTP) stream is between one Talker and one or more Listeners.

Bayer grbg, rggb, bgrg, gbrg: An arrangement of a color filter array where there is 50% green, 25% red, and 25% blue arranged as grbg, rggb, bgrg, or gbrg.

Controller Area Network (CAN): A serial communications protocol commonly used in automotive environments to connect control modules, sensors, and other devices. See ISO 11898 [B17]¹¹ for more information.

Controller Area Network with Flexible Data-Rate (CAN FD): An extension of the CAN specification that allows for data rates higher than 1 MBit/s and payloads longer than 8 octets per frame. See ISO 11898 [B17] for more information.

device: Within the context of this standard, a device is a hardware entity capable of participating in Audio/Video Transport Protocol (AVTP), functioning as a Talker, Listener, and/or Controller.

doublet: Two octets of data.

End of Active Video (EAV): A byte sequence defined by ITU-R BT.656 which indicates the end of the active video on a given video line, while also indicating the current field and whether the current line is part of vertical blanking.

FlexRay™: An automotive network communications protocol developed by the FlexRay Consortium. See ISO 10681-1 [B15], [B16], for more information.

frame: A frame is a complete unit of data transmission as transmitted on the network, including media access control (MAC) layer addressing and cyclic redundancy check (CRC).

¹¹The numbers in brackets correspond to those of the bibliography in Annex A.

free-wheel: A clock that is in free-wheel mode continues at its current frequency and phase with no corrections. A clock can be in this mode due to a lack of incoming correction information or because correction information is being disregarded.

Grayscale: A color space that carries only intensity information.

Listener: An end station that is the destination, receiver, or consumer of an AVTP stream.

Local Interconnect Network (LIN®): A serial communications protocol used in distributed automotive applications. See LIN Specification Package—Revision 2.2A for more information.

Media Oriented Systems Transport (MOST®): A streaming and communications protocol used in automotive applications. See MOST Specification Rev. 3.0 E2 07/2010 for more information.

Musical Instrument Digital Interface (MIDI) Manufacturers Association (MMA): The publisher and source of MIDI specifications. The MMA publishes the MMA Payload Specification for AVTP.

Observation Interval: The Observation Interval is equivalent to the Class Measurement Interval as defined in Clause 34 of IEEE Std 802.1Q-2014.

octet: Eight bits of data.

packet: Refers to the IEC 61883 definition. Packet is used to refer to specific features of 61883 that use this term, such as source packet header (SPH) or common isochronous packet (CIP).

quadlet: Four octets of data.

Real-time Transport Protocol (RTP): A standardized set of packet formats designed for transporting audio and video across IP networks.

RGB: An additive color space based on a red, green, blue color model.

Serial Data Interface (SDI): A family of digital video interfaces standardized by SMPTE for use in professional video applications.

Start of Active Video (SAV): A byte sequence defined by ITU-R BT.656 that indicates the start of the active video on a given video line, while also indicating the current field and whether the current line is part of vertical blanking.

Stream Reservation (SR) Class: As defined by Clause 35 of IEEE Std 802.1Q-2014.

StreamID: A 64-bit field that uniquely identifies a stream, as defined in Clause 35 of IEEE Std 802.1Q-2014.

Talker: An end station that is the source or producer of an AVTP stream.

Time-Sensitive Networking (TSN): *See: audio/video bridging (AVB).*

Time-Sensitive Networking (TSN) Bridge: *See: audio/video bridging (AVB) bridge.*

XYZ: A color space that uses tristimulus values based on the color matching functions of the CIE 1931 [B14] standard observer.

YCbCr: The color space formed by luminance (Y), chrominance blue (Cb), and chrominance red (Cr).

YCgCo: The color space formed by luminance (Y), chrominance green (Cg), and chrominance orange (Co).

YCM: A subtractive color space based on a yellow, cyan, and magenta color model.

3.2 Reserved fields

Any field within a data structure that is defined in this specification as reserved is reserved for future standardization. All reserved fields specified in this standard shall be set to a value of zero (0) on transmit and shall be ignored on receive.

Reserved fields in this standard are designated as one of **reserved**, **rsv**, or **r**.

3.3 Numerical values

Decimal, hexadecimal, and binary numbers are used within this document. For clarity, decimal numbers are generally used to represent counts, hexadecimal numbers are used to represent addresses, and binary numbers are used to describe bit patterns within binary fields.

Decimal numbers are represented in their usual 0, 1, 2, ... format. Hexadecimal numbers are represented by a string of one or more hexadecimal (0 to 9 and A to F) digits, followed by the subscript 16. Binary numbers are represented by a string of one or more binary (0, 1) digits in left to right order, where the left-most bit is the most-significant bit and the right-most bit is the least-significant bit, followed by the subscript 2. Thus, the decimal number “26” may also be represented as “1A₁₆” or “11010₂”.

These notational conventions have one exception: media access control (MAC) addresses and IEEE Organizationally Unique Identifier (OUI)/IEEE Extended Unique Identifier (EUI)/MAC Address Blocks (MA-L, MA-M, MA-S) values are represented as strings of 1-octet hexadecimal numbers separated by hyphens and without a subscript as, for example, “01-80-C2-00-00-15” or “AA-55-11.”

3.4 Notation of fields and values taken from other documents

Since this document references the protocols described in other documents, the notational conventions for referenced fields and values have been preserved where possible to avoid any confusion.

- a) Fields from IEEE Std 802.1Q-2014: Fields are in all uppercase, no underscores (examples: destination MAC address [DA], source MAC address [SA], tagged protocol identifier [TPID], drop eligible indicator [DEI], and VLAN identifier [VID]).
- b) Fields from IEEE Std 1394: Fields in lowercase except for acronyms within the field name with optional underscores (examples: tcode, data_length, and source_ID).
- c) Fields from IEC 61883: Fields always starting in uppercase, acronyms in uppercase, abbreviations with uppercase first followed by lowercase, no underscores (examples: data block count [DBC], data block size [DBS], and reserved [Rsv]).

3.4.1 Bit, octet, doublet, and quadlet ordering

All Audio Video Transport Protocol Data Units (AVTPDUs) consist of an integral number of octets that are numbered starting from 1 and increasing in the order that they are put into a Data Link Service Data Unit. The bits in each octet are numbered from 0 to 7, where 7 is the low-order bit.

When consecutive octets are used to represent a binary number, the lower octet number has the most-significant value. When the encoding of (an element of) an AVTPDU is represented using a diagram in this clause, the following representations are used:

- Octet 1 is shown toward the top of the page, higher numbered octets being toward the bottom.
- Where more than one octet appears on a given line, octets are shown with the lowest numbered octet to the left, higher numbered octets being to the right.
- Within an octet, bits are shown with bit 0 to the left and bit 7 to the right.

Figure 1 represents a decimal value of 1, a hexadecimal value of 01_{16} , and a binary value of 00000001_2 with the only bit in the octet set to one of the least-significant bit, bit 7.

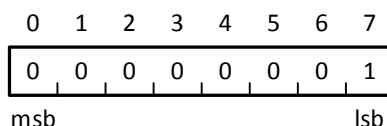


Figure 1—Bit ordering within an octet

A quadlet is a series of four octets. Within a quadlet, the most-significant octet is shown as the left-most octet and the least-significant octet is shown as the right-most, as shown in Figure 2.

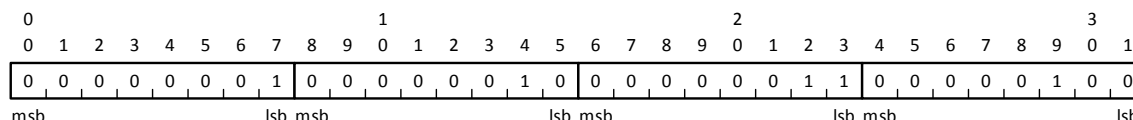


Figure 2—Octet ordering within a quadlet

Figure 2 represents a quadlet holding a four-octet array with values of 1, 2, 3, and 4 decimal for each octet.

A quadlet may contain bit fields of any length between one bit and four octets. When a field spans more than one octet, the point where it spans the octet is shown as a large tick mark as shown in Figure 3.

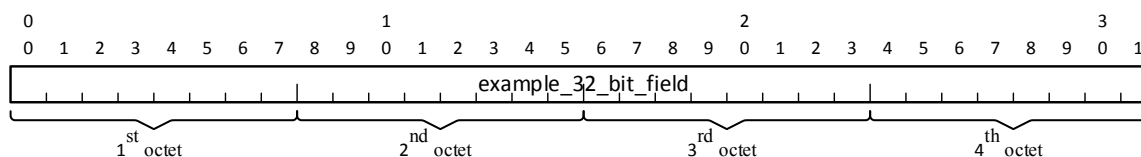


Figure 3—Example 32-bit field diagram

The 64-bit fields that need to be contained in more than one quadlet are shown as a series of eight octets in Figure 4.

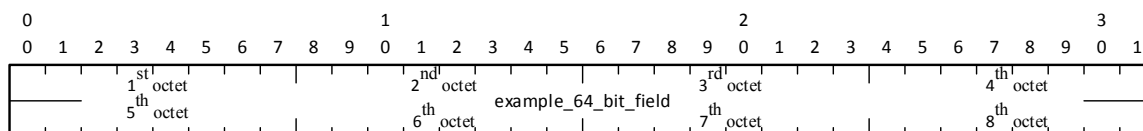


Figure 4—Example 8-octet field diagram

3.4.2 Field value conventions

This document describes values of fields. For clarity, names can be associated with each of these defined values, as illustrated in Table 2. A symbolic name, consisting of uppercase letters with underscore separators, allows other portions of this document to reference the value by its symbolic name, rather than by a numerical value.

Table 2—Wrap field values

Value	Name	Description
0	WRAP_AVOID	Frame is discarded at the wrap point
1	WRAP_ALLOW	Frame passes through wrap points
2 to 3	—	Reserved

A field value of TRUE shall always be interpreted as being equivalent to a numeric value of one (1) unless otherwise indicated. A field value of FALSE shall always be interpreted as being equivalent to a numeric value of zero (0) unless otherwise indicated.

3.4.3 AVTPDU diagram conventions

This document uses black and gray coloring in the AVTPDU diagrams to indicate the focus of discussion within the diagram. Figure 5 shows an example of the usage of black and grayed coloring.

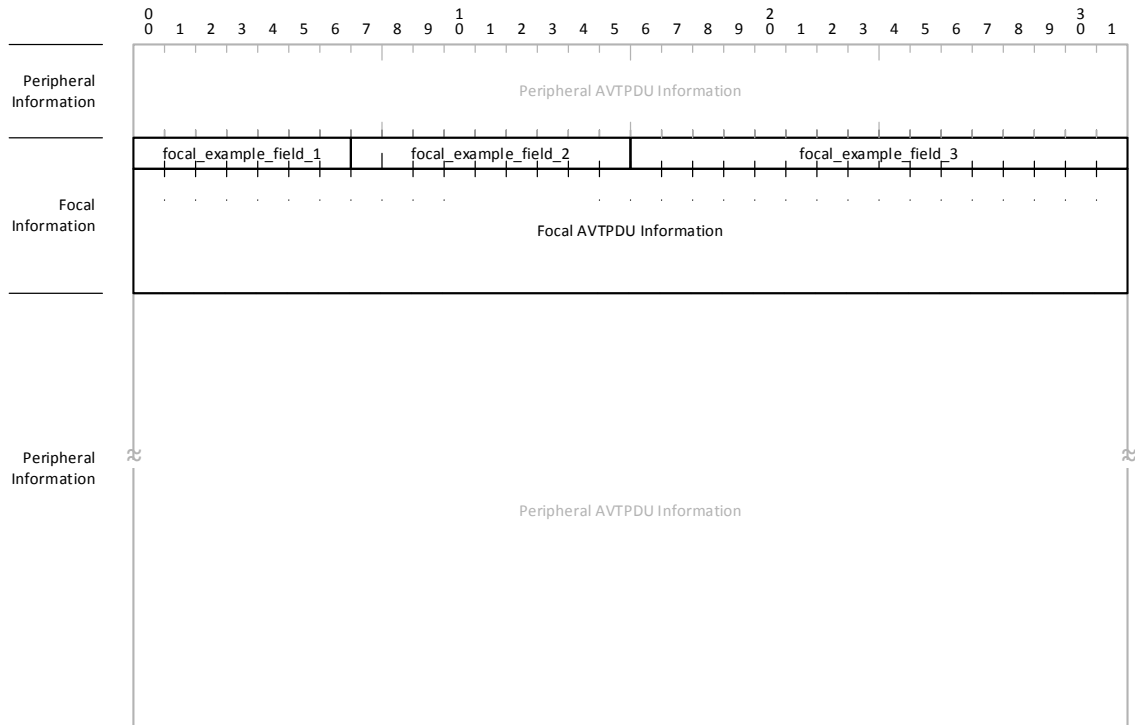


Figure 5—Example diagram showing grayed fields

In Figure 5 items that are focal to the text of the section are shown in black. In this example, the fields **focal_example_field_1**, **focal_example_field_2**, and **focal_example_field_3** would be described and discussed in the section's text. The items labeled "Peripheral AVTPDU Information" are grayed-out to show that they are described elsewhere. The diagram shows them to provide context to the focal items.

3.5 Acronyms and abbreviations

AAF	AVTP Audio Format
ADP	AVDECC Discovery Protocol
AECp	AVDECC Enumeration and Control Protocol
ACMP	AVDECC Connection Management Protocol
AVB	audio/video bridging
AVDECC	Audio/Video Discovery, Enumeration, Connection Management, and Control
AVTP	Audio/Video Transport Protocol
AVTPDU	Audio/Video Transport Protocol Data Unit
CAN	Controller Area Network
CAN FD	CAN with Flexible Data-Rate
CIP	common isochronous packet
DA	destination MAC address
DBC	data block count
DBS	data block size

DEI	drop eligible indicator
DTCP	digital transmission content protection
DVCR	digital videocassette recorder
EAV	End of Active Video
EUI	IEEE Extended Unique Identifier
FDF	format dependent field
FQTSS	Forwarding and Queuing for Time-Sensitive Streams (Clause 34 of IEEE Std 802.1Q-2014)
FMT	stream format
FN	fraction number
gPTP	generalized Precision Time Protocol (IEEE Std 802.1AS)
HANC	Horizontal Ancillary Data
HDCP	High-bandwidth Digital Content Protection
IIDC	Instrumentation and Industrial Digital Camera
kHz	kilohertz (thousand cycles per second)
LAN	local area network
LIN®	Local Interconnect Network
MAAP	MAC Address Acquisition Protocol
MAC	media access control
MA-S	MAC Address Block Small (formerly OUI-36)
MA-M	MAC Address Block Medium
MA-L	MAC Address Block Large (formerly OUI)
MHz	megahertz (million cycles per second)
MIDI	Musical Instrument Digital Interface
MMA	MIDI Manufacturers Association, Inc. ¹
MMRP	Multiple MAC Registration Protocol
MOST®	Media Oriented Systems Transport
MTU	maximum transmission unit, for this document refers to the maximum length of the MAC Client Data field defined in 3.2.7 of IEEE Std 802.3™-2012
MVRP	Multiple VLAN Registration Protocol (subclause 11.2 of IEEE Std 802.1Q-2014)
NALU	Network Abstraction Layer Unit. The basic transport layer unit for H.264 compressed video. (See 1394 Trade Association, 2004006 [B1].)
OUI	IEEE organizationally unique identifier
PCP	priority code point
PDU	protocol data unit
QPC	quadlet padding count
r	reserved

¹MIDI Manufacturers Association, Inc., P.O. Box 3173, La Habra, CA 90632-3173, <http://www.midi.org>.

rsv	reserved
RTP	Real-time Transport Protocol
SAV	Start of Active Video
SDI	Serial Digital Interface
SID	source identifier
S/PDIF	Sony/Philips Digital Interconnect Format
SRP	Stream Reservation Protocol (Clause 35 of IEEE Std 802.1Q-2014)
SPH	source packet header
sRGB	standard RGB
SYT	synchronization timing
TLV	Type-Length-Value
TPID	tagged protocol identifier
TSN	Time-Sensitive Networking
TSP	Transport Stream Packet
VANC	Vertical Ancillary Data
VDSPC	Video Data Source Packet Count
VID	VLAN identifier
VLAN	Virtual Local Area Network

4. AVTP base protocol

4.1 Overview

Audio Video Transport Protocol (AVTP) specifies a protocol for audio, video, and control data to be transported on a Time-Sensitive Networking (TSN) capable network. AVTP is designed to take advantage of TSN capable networks that support generalized Precision Time Protocol (gPTP), Stream Reservation Protocol (SRP), and Forwarding and Queuing Enhancements for Time-Sensitive Streams (FQTSS).

AVTP specifies methods to transport data and timing information so that audio, video, or control content sent by a Talker can be reproduced on Listeners. AVTP streams are sent from a Talker and can be received by one or more Listeners.

AVTP stream data is typically carried directly on an underlying MAC layer such as IEEE Std 802.3 or IEEE Std 802.11™; however, it is also possible to encapsulate AVTP in IP as described in Annex J.

This clause contains the basic assumptions and operation requirements of the protocol.

4.2 AVTP network requirements

All AVTP end stations that send or receive audio, video, or time-sensitive control data are required to support the TSN services provided by gPTP, SRP, and FQTSS. AVTP relies on these services being available to interoperate. AVTP end stations that only support data formats that are not time-sensitive are

not required to support TSN services. The behavior of a Talker, Listener, or bridge on the network that forwards AVTP audio, video or time-sensitive control data and does not support the TSN protocols is out of the scope of this standard.

AVTP makes use of gPTP to provide a network-wide time base that can be used to convey timing information from a Talker to Listener(s).

AVTP makes use of SRP and FQTSS to provide reliable network delivery with bounded network latency for transporting data from Talker to Listener(s).

AVTP data can be sent from Talker to Listener(s) either directly or forwarded by TSN bridges to the Listener(s).

4.3 Timing and synchronization

4.3.1 General

AVTP defines a presentation time to achieve timing synchronization between a Talker and Listener(s); it represents in nanoseconds the gPTP time when the data contained in the AVTPDU is to be available to the AVTP Listener(s).

AVTP Presentation Time is used as a reference to synchronize any necessary media clocks and to determine when the payload of a stream is to be presented to the time-sensitive application. Media clocks vary with audio/video types therefore the exact usage of the AVTP Presentation Time is media format dependent.

AVTPDUs are subject to variable transit times between Talker and Listener. The AVTP Presentation Time is used to account for this variability and facilitate synchronization of each AVTPDU's payload data at the Listener.

4.3.2 AVTP Presentation Time

The AVTP Presentation Time represents the gPTP time at which a designated media sample or event within the AVTPDU payload is transferred to the time-sensitive application within each Listener. This enables multiple Listeners to present data at the same time, regardless of the location of the Listeners in the network. It tells a Listener when to start processing (e.g., playing) the stream's data and is also used to recover the stream's media clock.

The AVTP Presentation Time may not be valid in every AVTPDU. If an AVTPDU contains a valid timestamp, then the **tv** (avtp_timestamp valid) bit shall be set to one (1).

4.3.3 AVTP Presentation Time-measurement points

This subclause defines the time-measurement points that are referenced in the calculation of the AVTP Presentation Time.

The AVTP Presentation Time is shown as the Presentation Time Reference Plane in Figure 6.

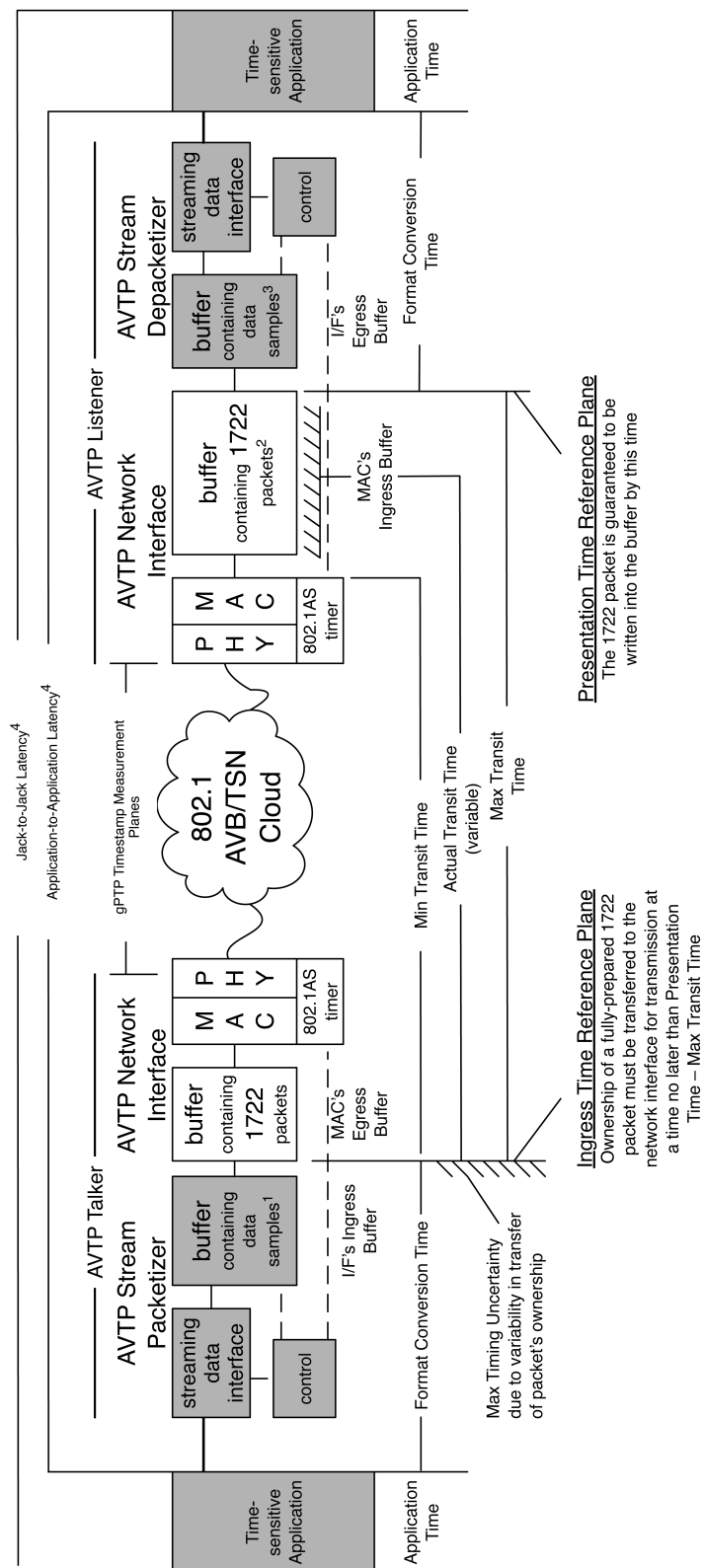


Figure 6—AVTP Timing Reference Planes

Notes to the figure:

- ¹Optional buffer if samples are not put directly into the 1722 packets
- ²1722 Buffer size is required to be at least the length of Max Transit Time + Max Timing Uncertainty to avoid possible data loss.
- ³Optional buffer if samples are not pulled directly from the 1722 packets
- ⁴Jack-to-Jack latency and Application to Application Latency are outside the scope of AVTP, please refer to other standards such as IEEE Std. 1722.1-2013

The Max Transit Time is the time between the Ingress Time Reference Plane and the Presentation Time Reference Plane. The Max Transit Time is a function of the speed, size (number of bridges), and configuration of the audio/video bridging (AVB) network.

The default Max Transit Times for SR classes A and B are given in Table 3 (see subclause 6.5 of IEEE Std 802.1BA-2011).

Table 3—AVTP default Max Transit Time

SR class	Max Transit Time (ms)
A	2.0
B	50.0

It is possible to determine an alternate Max Transit Time value for use in a given stream. Guidelines for determining alternate Max Transit Time values are given in 4.3.4. However, the actual mechanism for calculating and assigning this value is outside the scope of this standard.

In most cases, the AVTP Presentation Time corresponds to the first sample in the AVTPDU. The Talker sets the AVTP Presentation Time such that the time at which it transfers ownership of the AVTPDU to the network interface on the Talker (Ingress Time Reference Plane in Figure 6) is at or before the AVTP Timestamp minus the Max Transit Time and within the limits of the Max Timing Uncertainty (Equation (1), where $T_{\text{samples}} = 0$).

In cases where the AVTP Presentation Time corresponds to a sample that is not the first sample in the AVTPDU, the packet must be transferred to the network interface at the AVTP Presentation Time minus the Max Transit Time, minus the number of samples before the AVTP Presentation Time in the AVTPDU (Equation (1), where $T_{\text{samples}} \neq 0$).

$$PT - TT_{\text{max}} - TU_{\text{max}} - T_{\text{samples}} < T_T < PT - TT_{\text{max}} - T_{\text{samples}} \quad (1)$$

where

T_T	AVTPDU transfer time requirement
PT	AVTP Presentation Time
TT_{max}	Max Transit Time
TU_{max}	Max Timing Uncertainty
T_{samples}	Time required to buffer all samples before the sample associated with the AVTP Presentation Time

The Talker shall transfer ownership of stream AVTPDUs to the MAC layer within the constraints of T_T , as defined in Equation (1). If the Talker fails to transfer a stream AVTPDU before T_T , then the AVTPDU may not arrive at the Listener before it is required for audio/video playback. This may result in loss of portions of the media data stream. In any design, there is uncertainty about when this Ingress Time Reference Plane crossing will occur. It is the Talker's responsibility to know the Timing Uncertainty of its own design and to begin the transfer of ownership such that, in the worst case, the crossing of the Ingress Time Reference Plane will occur within the constraints given in Equation (1).

Timing Uncertainty is the potential delay introduced by the transfer of ownership of the frame from the AVTP layer to the media interface. The Max Timing Uncertainty is the maximum amount of transfer delay allowed. Max Timing Uncertainty is different for different traffic priorities. Since higher-priority traffic

may be queued ahead of lower-priority traffic, there will be more uncertainty in the lower-priority traffic. A Talker's Timing Uncertainty is a combination of many factors, including its own local timer resolution and its worst-case response time to a timer event. The Timing Uncertainty of a Talker shall be no more than the Max Timing Uncertainty amount defined in Table 4.

Table 4—AVTP Max Timing Uncertainty

Class	Max Timing Uncertainty (μ s)
A	125
B	1000

The required buffering within a Listener is sufficient to hold the stream's events for the amount of time given in Equation (2) for the class of the stream. If the Listener does not have sufficient buffering for a given stream, a buffer overflow could occur, resulting in data loss.

$$T_{\text{buf}} \geq TT_{\text{default}} + TU_{\text{max}} + PST_{\text{max}} \quad (2)$$

where

- T_{buf} Listener Buffering Time
- TT_{default} Largest of the Default Max Transit Times for the SR classes supported
- TU_{max} Max Timing Uncertainty
- PST_{max} Max gPTP Skew Time (as defined in subclause B.3 of IEEE Std 802.1AS-2011)

Figure 6 also shows Format Conversion and Reception/Transmission Time on the Talker and Listener. These interfaces are optional media interfaces that could connect to physical devices. These media interfaces are outside the scope of this standard. Any format conversion time required for these interfaces (e.g., buffering time, audio/video codec delay times, etc.) are outside of and in addition to T_{buf} (see Equation (2)).

4.3.4 Calculation of Max Transit Time

In many cases it is desirable to use a Max Transit Time other than the default values listed in Table 3. In the case of an AVTP stream between a Talker and known Listeners it is possible to optimize Max Transit Time to accommodate low-latency applications or to accommodate a large number of intermediate bridges.

The basic method to calculate an appropriate Max Transit Time is to take the worst-case transit time from the Talker to a Listener and choose a Max Transit Time that is greater than or equal to the largest worst-case transit time as shown in Equation (3).

$$\text{MaxTransitTime} \geq \max(\text{WorstCaseTransitTime}_1, \text{WorstCaseTransitTime}_2, \dots, \text{WorstCaseTransitTime}_n) \quad (3)$$

Factors that affect the calculation of the worst-case transit time include both internal latency in the Talker and Listener and network delays. Network delays are typically related to the physical network topology and number of bridges or other network elements between the Talker to the desired Listeners. Worst-case network delays can be discovered using SRP and the Accumulated Latency as reported within a Talker Advertise attribute.

Once the worst-case transit time has been determined for each Talker/Listener pair, a Max Transit Time for the stream can be determined.

Figure 7 shows an example of a Talker with two Listener devices on a network, and illustrates the worst-case timing components discussed in this subclause.

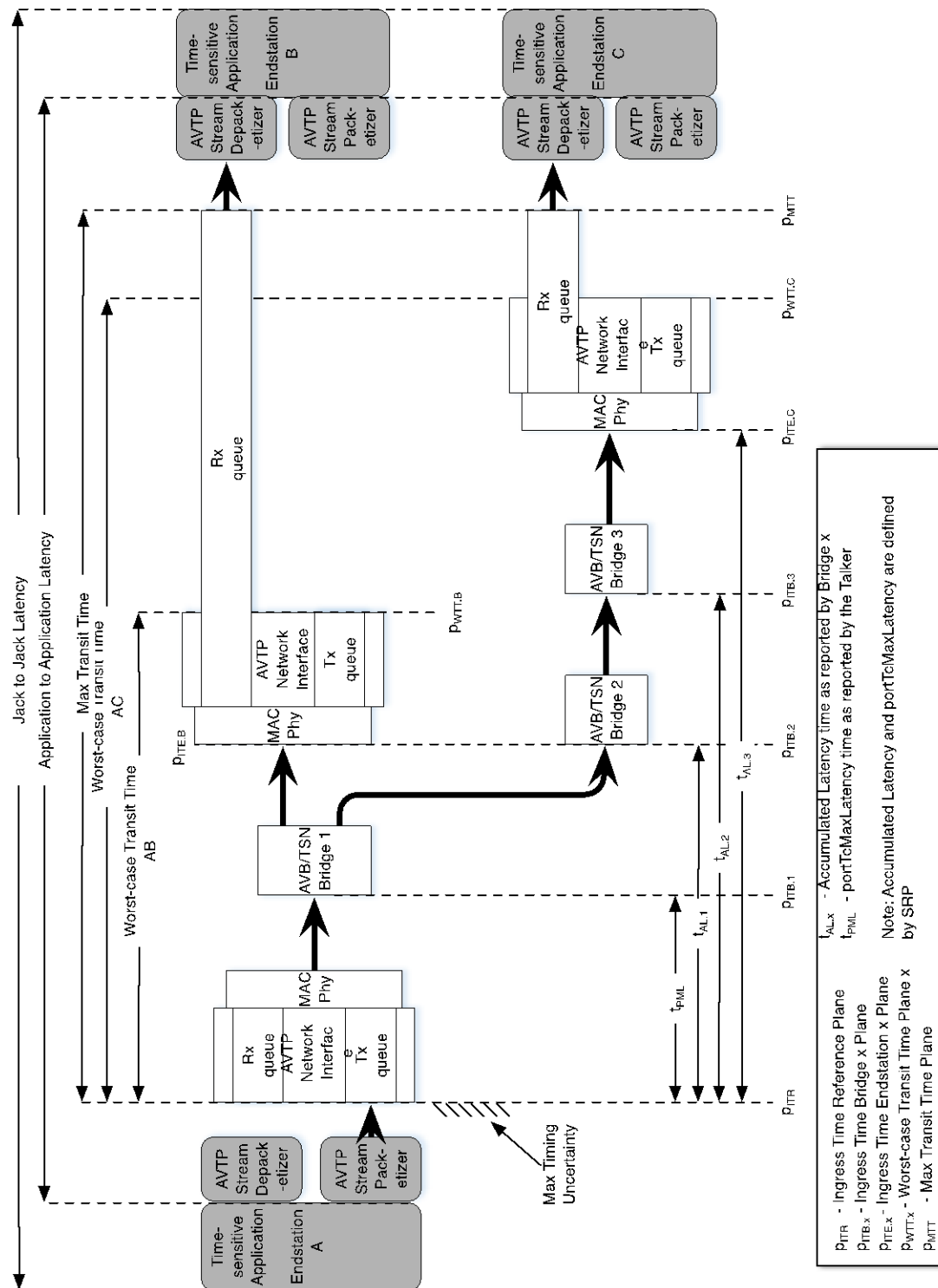


Figure 7—Determining Max Transit Time with multiple Listeners

Equation (4) shows the calculations for worst-case transit time for an AVTP stream from Talker A to Listener B and Listener C in Figure 7. Equation (5) shows the calculation for the Max Transit Time for an AVTP stream from Talker A to Listener B and Listener C in Figure 7.

$$\begin{aligned} WTT_{AB} &= t_{AL,1} + (P_{WTT,B} - P_{ITE,B}) \\ WTT_{AC} &= t_{AL,3} + (P_{WTT,C} - P_{ITE,C}) \end{aligned} \quad (4)$$

$$TT_{max} \geq \max(WTT_{AB}, WTT_{AC}) \quad (5)$$

where

WTT_{AB}	worst-case transit time from Talker A to Listener B
WTT_{AC}	worst-case transit time from Talker A to Listener C
$t_{AL,1}$	Accumulated Latency as reported by SRP for Bridge 1
$t_{AL,3}$	Accumulated Latency as reported by SRP for Bridge 3
$P_{WTT,B}$	Worst-case Timing Plane for Endstation B
$P_{ITE,B}$	Ingress Timing Plane for Endstation B
$P_{WTT,C}$	Worst-case Timing Plane for Endstation C
$P_{ITE,C}$	Ingress Timing Plane for Endstation C
TT_{max}	Max Transit Time

The time $t_{AL,x}$ as reported by SRP includes all queuing and transmission delays for the Talker and Bridges. The reported Accumulated Latency for a Talker is the portTcMaxLatency of the Talker. The reported Accumulated Latency for a Bridge is the worst-case latency from the first bit received by the bridge to the first bit received by the next network device. The reported time is accumulated at each device, so for example, the Accumulated Latency reported by Bridge 3 will include worst-case latency for Talker A, Bridge 1, Bridge 2, and Bridge 3.

The time represented by the equation $(P_{WTT,x} - P_{ITE,x})$ represents the worst-case time from the first bit of the AVTP frame arriving at the Listener until the stream data will be available to the next function in the Listener.

4.3.5 Other timing considerations

There are use cases where it is desirable to synchronize the presentation of media transported in AVTP Streams from multiple Talkers. This can be accomplished by the Talkers subscribing to a single stream that provides the media clocking information for the AVTP Streams that are generated. (See Clock Reference Format in Clause 8.)

This clocking information stream could define an audio media clock, video line or frame clock, or an industrial working clock.

In a system where streams are to be mixed or combined at an endpoint and one or more streams are generated from a recovered media stream, it is required that the difference between the AVTP Presentation Times for the generated stream and the recovered stream be an integer multiple of the nominal period of the stream media clock.

4.3.6 Protocol layering and other required protocols

This clause documents other required protocols and standards required, but not specified, in this transport-level specification.

All AVTP end stations that transmit or receive time-sensitive streams on a plug-and-play network shall support the following standards:

- Generalized Precision Timing Protocol
- Stream Reservation Protocol
- Forwarding and Queuing Enhancements for Time-Sensitive Streams

4.4 Header formats

4.4.1 Overview

This subclause documents the AVTP header format. The AVTP header includes the basic information such as subtype and version numbers required to decode the AVTPDU. AVTP defines a common header and several specialized header formats.

The AVTPDU common header (4.4.3) contains the basic fields that all formats share, such as the subtype definition and the version of the format.

The AVTPDU common stream header (4.4.4) expands on the AVTPDU common header (4.4.3) and contains the fields that all stream-based formats share. Among other shared fields, this header includes the StreamID for stream identification and a timestamp field for indicating when the included samples are to be presented to the Listener (4.3.2).

The AVTPDU common control header (4.4.5) expands on the AVTPDU common header (4.4.3) and contains the fields that all control formats share. Among other shared fields, this header includes the StreamID for stream identification.

The AVTPDU alternative header (4.4.6) is for formats that do not exhibit the commonalities shared between formats that use the stream or control headers.

4.4.2 AVTP EtherType

The EtherType of all AVTP frames shall be the EtherType given in Table 5.

Table 5—AVTP EtherType

AVTP Ethertype
22F0 ₁₆

4.4.3 AVTPDU common header format

4.4.3.1 Overview

This subclause documents the first 12 bits of header that are common to all AVTP stream data and control AVTPDUs. These 12 bits consist of the following fields:

- **subtype**: 1 octet
- **h** (header specific): 1 bit
- **version**: 3 bits

Figure 8 shows these fields with offset zero (0) shown as the first octet of the AVTPDU.

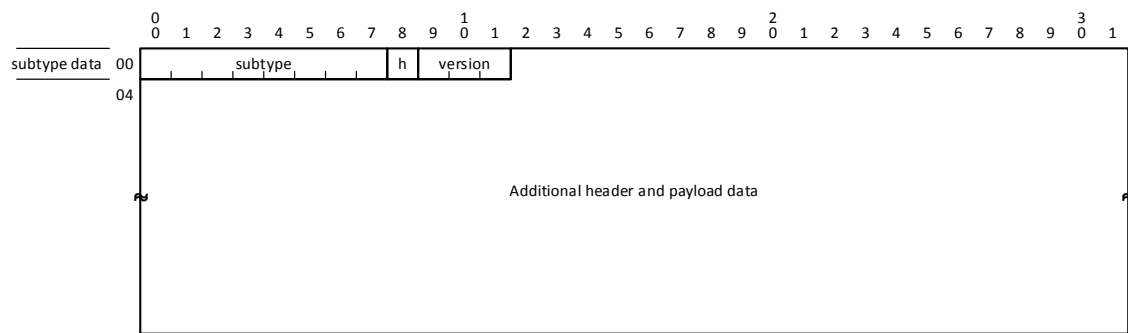


Figure 8—AVTPDU common header fields

4.4.3.2 subtype field

The 1-octet **subtype** field is used to identify the format being carried by AVTP as defined by Table 6. Each format defines its use of AVTP encapsulation. Each subtype is also assigned an *encapsulation style*. Subtypes that are typically composed of a continuous stream of related AVTPDUs are assigned an encapsulation style of *continuous*. Subtypes that are typically composed of discrete AVTPDUs, which have no relationship to each other, are assigned an encapsulation style of *discrete*.

NOTE—The distinction between continuous and discrete stream types is useful in that continuous streams can be extremely high bandwidth and the traffic can overwhelm a network stack. In many cases a special purpose network stack is desired to process continuous-type streams.¹

¹Notes in text, tables, and figures of a standard are given for information only and do not contain requirements needed to implement this standard.

Table 6—AVTP stream data subtype values

Value	Name	Description	Reference	Header type	Encapsulation style
00 ₁₆	61883_IIDC	IEC 61883/IIDC format	Clause 5	Stream (4.4.4)	Continuous
01 ₁₆	MMA_STREAM	MMA streams	Clause 1	Stream (4.4.4)	Continuous
02 ₁₆	AAF	AVTP Audio Format	Clause 7	Stream (4.4.4)	Continuous
03 ₁₆	CVF	Compressed Video Format	Clause 7	Stream (4.4.4)	Continuous
04 ₁₆	CRF	Clock Reference Format	Clause 8	Alternative (4.4.6)	Continuous
05 ₁₆	TSCF	Time-Synchronous Control Format	Clause 9.3	Stream (4.4.4)	Continuous
06 ₁₆	SVF	SDI Video Format	Clause 8	Stream (4.4.4)	Continuous
07 ₁₆	RVF	Raw Video Format	Clause 12	Stream (4.4.4)	Continuous
08 to 6D ₁₆	Reserved	Reserved	—	—	—
6E ₁₆	AEF_CONTINUOUS	AES Encrypted Format Continuous	Clause 11	Alternative (4.4.6)	Continuous
6F ₁₆	VSF_STREAM	Vendor Specific Format Stream	Clause 11	Stream (4.4.4)	Continuous
70 to 7E ₁₆	Reserved	Reserved	—	—	—
7F ₁₆	EF_STREAM	Experimental Format Stream	Clause 15	Stream (4.4.4)	Continuous
80 to 81 ₁₆	Reserved	Reserved	—	—	—
82 ₁₆	NTSCF	Non-Time-Synchronous Control Format	Clause 9.2	Alternative (4.4.6)	Discrete
83 to ED ₁₆	Reserved	Reserved	—	—	—
EC ₁₆	ESCF	ECC Signed Control Format	Clause 16	Alternative (4.4.6)	Discrete
ED ₁₆	EECF	ECC Encrypted Control Format	Clause 17	Alternative (4.4.6)	Discrete
EE ₁₆	AEF_DISCRETE	AES Encrypted Format Discrete	Clause 11	Alternative (4.4.6)	Discrete
EF to F9 ₁₆	Reserved	Reserved	—	—	—
FA ₁₆	ADP	AVDECC Discovery Protocol	IEEE Std 1722.1™, 2013 Clause 6	Control (4.4.5)	Discrete
FB ₁₆	AACP	AVDECC Enumeration and Control Protocol	IEEE Std 1722.1-2013 Clause 9	Control (4.4.5)	Discrete
FC ₁₆	ACMP	AVDECC Connection Management Protocol	IEEE Std 1722.1-2013 Clause 8	Control (4.4.5)	Discrete
FD ₁₆	Reserved	Reserved	—	—	—
FE ₁₆	MAAP	MAAP Protocol	Annex B	Control (4.4.5)	Discrete
FF ₁₆	EF_CONTROL	Experimental Format Control	Clause 15	Control (4.4.5)	Discrete

4.4.3.3 h (header specific) field

The **header_specific_bit** is a 1-bit field whose meaning is defined by the type of header being used (see 4.4.4.2, 4.4.5.2, and 4.4.6).

4.4.3.4 version field

The **version** field specifies the version of the format. Unless explicitly defined by a format definition, this field shall be set to zero (0) on transmit and verified on receive. If the Listener receives an AVTPDU with a version that it does not support, it shall discard the AVTPDU.

4.4.4 AVTPDU common stream header

4.4.4.1 Overview

This subclause documents common fields that are used in addition to the common header (see 4.4.3) for stream-related formats.

- **mr** (media clock restart): 1 bit
- **f_s_d** (format-specific data): 2 bits
- **tv** (avtp_timestamp valid): 1 bit
- **sequence_num** (sequence number): 1 octet
- **format_specific_data_1**: 7 bits
- **tu** (timestamp uncertain): 1 bit
- **stream_id**: 8 octets
- **avtp_timestamp**: 4 octets
- **format_specific_data_2**: 4 octets
- **stream_data_length**: 2 octets
- **format_specific_data_3**: 2 octets
- **stream_data_payload**: 0 to n octets

The **h** (header_specific) bit (4.4.3.3) of the common header is used to define the following:

- **sv** (stream_id valid): 1 bit

These fields are shown in Figure 9.

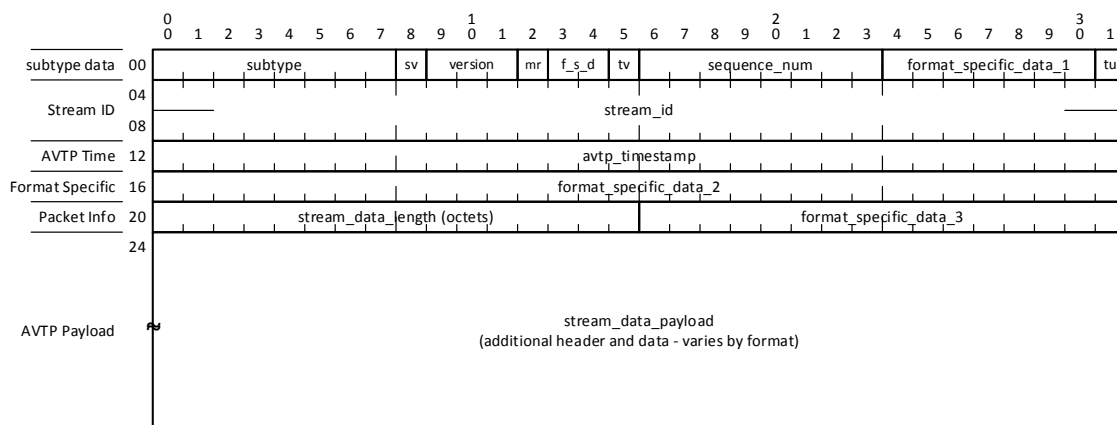


Figure 9—AVTPDU common stream header

4.4.4.2 sv (stream_id valid) field

The **sv** (stream_id valid) field is used to indicate whether the 8-octet **stream_id** field contains a valid StreamID. Since all formats that use the common stream header require a valid StreamID, the value of the **sv** field shall always be one (1).

4.4.4.3 mr (media clock restart) field

The **mr** (media clock restart) field is a 1-bit field used by the Talker to indicate that there has been a change in the source of the media clock. An example of this would be a change from one audio/video input to a different audio/video input where the media clock is supplied from the input (i.e., switching from one Sony/Philips Digital Interconnect Format (S/PDIF) input to a different S/PDIF input). This bit may be used by the Listener to adjust quickly to the new media clock. This kind of change need not be seamless but allows the Listener to take any appropriate action to minimize the disruption to the stream.

This bit is toggled by the Talker each time a media clock restart is needed, and it stays at its new value until a new media clock restart is needed. This toggle helps reduce the likelihood that the media clock restart is missed if it were to appear in a single AVTPDU that just happens to be dropped. Once this bit toggles, it shall remain in its new state for a minimum of eight (8) AVTPDUs for a given continuous stream.

If a Clock Reference Format (CRF) stream (see Clause 8) is being received by a Talker, then any streams deriving timestamps from the CRF stream shall toggle the **mr** bit if a disruption of the CRF stream occurs or if the **mr** bit in the CRF stream has been toggled.

If a Listener is receiving multiple streams that are all derived from a common media clock, the Listener uses only the **mr** bit from the common media clock stream for recovering the media clock.

4.4.4.4 Format-specific data fields

The **f_s_d**, **format_specific_data_1**, **format_specific_data_2**, and **format_specific_data_3** fields are available for use by each given format as specified by the **subtype** field. The specific use of these fields is defined by the format.

4.4.4.5 **tv (avtp_timestamp valid) field**

The **avtp_timestamp valid (tv)** field is a 1-bit field used to indicate the validity of the **avtp_timestamp** field time value.

The AVTP Talker shall set the **tv** bit to one (1) if the **avtp_timestamp** field contains a valid value, or zero (0) otherwise.

The AVTP Listener shall ignore the value of the **avtp_timestamp** field if **tv** is set to zero (0).

For how the **avtp_timestamp** field is interpreted and processed, see 4.4.4.9.

4.4.4.6 **sequence_num field**

The sequence number field (**sequence_num**) is a 1-octet field used to indicate the sequence of AVTPDUs in a stream by the Talker. This field may start at any value for the first AVTPDU of a stream and shall be incremented by one (1), with wrapping, on each subsequent AVTPDU of a stream. This field is to wrap from 11111111_2 to 00000000_2 (FF_{16} to 00_{16}).

The value of this field is determined by the Talker. Listeners can use the sequence number to detect AVTPDUs lost in transit.

A Listener may join a stream at any time after the stream has started, and therefore, a Listener shall be able to tolerate a starting **sequence_num** of any value.

4.4.4.7 **tu (timestamp uncertain) field**

Although gPTP time is intended to be stable, it is possible for discontinuities to occur in gPTP time. These can be caused by events such as changes in the identity of the gPTP grandmaster clock, changes in the timing source of the grandmaster clock, or other events encountered by a Talker.

When a Talker detects a discontinuity in gPTP time, it should set the **tu** (timestamp uncertain) field to one (1). This indicates to the Listener(s) of the stream that the AVTP Presentation Times contained in the **avtp_timestamp** field may for a limited period of time not correspond to gPTP time. The **tu** bit may be used by Listeners, possibly in conjunction with knowledge of the status of the gPTP clock, to prevent unacceptable disturbances in recovered media streams. Once the Talker has determined that the gPTP clock has returned to a normal state, the **tu** field should be reset to zero (0).

When a Listener detects that the **tu** field has been set to one (1) or detects a discontinuity, either from an gPTP indication or simple observation, it may stop attempting to correlate AVTP Presentation Time to gPTP time until the source of the discontinuity has ceased and the Listener has allowed the appropriate time for the gPTP clock to stabilize.

The time required for the gPTP clock to stabilize is implementation dependent and therefore outside the scope of this standard.

NOTE—When a Listener is in a mode where the **tu** bit is being set, it will usually ignore media clock recovery data and stop adjusting the media clock. Allowing the media clock to free-wheel in this manner will cause the least amount of corruption to the media data as compared to potentially making extreme adjustments to a malfunctioning clock. Timestamps are taken from the media clock like normal. As soon as the **tu** bit is reset, media clock recovery can begin anew and adjust for any clock wander that has occurred since adjustments were suspended.

4.4.4.8 stream_id field

The **stream_id** field contains the 64-bit StreamID associated with the AVTPDU. The **stream_id** field is used for stream identification.

All stream data AVTPDUs shall contain a valid StreamID in the **stream_id** field, as defined in subclause 35.2.2.8.2 of IEEE Std 802.1Q-2014, and as shown for convenience in Figure 10.

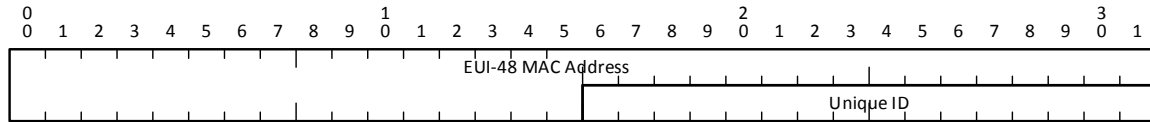


Figure 10—AVTPDU StreamID Format

4.4.4.9 avtp_timestamp field

The 4-octet **avtp_timestamp** field contains the AVTP Presentation Time (see 4.3.2) if the **tv** bit is set to one (1). The **avtp_timestamp** field shall be in units of nanoseconds derived from gPTP time and has a maximum value of $2^{32}-1$ ns. The **avtp_timestamp** field rolls over approximately every 4 s.

The **avtp_timestamp** field is not simply the gPTP nanoseconds field. The **avtp_timestamp** field is given by the formula in Equation (6).

$$\text{avtp_timestamp} = (\text{AS_sec} \times 10^9 + \text{AS_ns}) \bmod 2^{32} \quad (6)$$

where

AS_sec the gPTP seconds field

AS_ns value is the gPTP nanoseconds field

If the **tv** bit is zero (0), then the content of the **avtp_timestamp** field is undefined and is ignored.

4.4.4.10 stream_data_length field

The 2-octet **stream_data_length** field contains the length (in octets) of the **stream_data_payload** field.

The maximum value for this field is determined by the MTU and shall be set to allow all stream data AVTPDUs to fit within a single standard frame payload.

4.4.4.11 stream_data_payload field

The **stream_data_payload** contains the data that is associated with or defined by a format. The actual data encapsulation is defined by the format.

4.4.5 AVTPDU common control header

4.4.5.1 Overview

This subclause documents common fields that are used in addition to the common header (see 4.4.3) for control-related formats.

- **format_specific_data**: 9 bits
- **control_data_length**: 11 bits
- **stream_id**: 8 octets
- **control_data_payload**: 0 to n octets

The **h** (header_specific) bit (4.4.3.3) of the common header is used to define the following:

- **sv** (stream_id valid): 1 bit

These fields are shown in Figure 11.

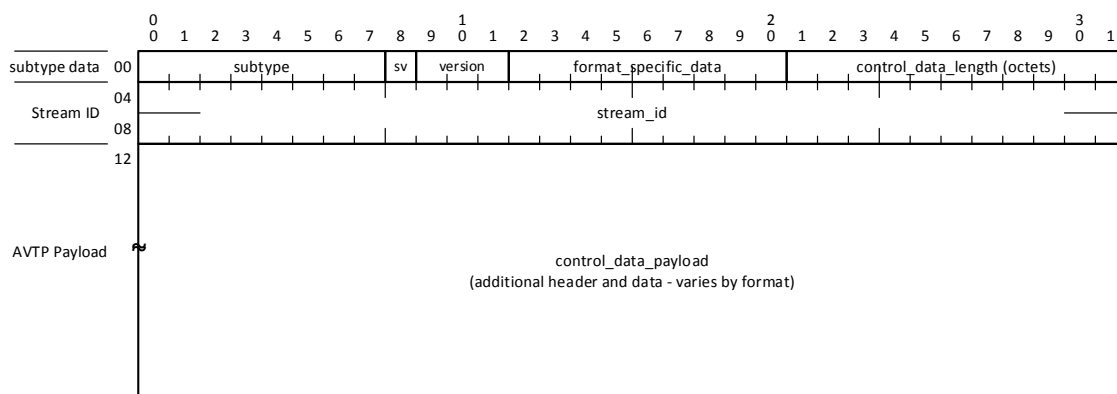


Figure 11—AVTPDU common control header

4.4.5.2 sv (stream_id valid) field

The **sv** (stream_id valid) field is used to indicate whether the **stream_id** field contains a valid StreamID. The options for the value of the **sv** field are defined in Table 7.

Table 7—sv field values

Value	Description
0	The stream_id field is not used or is defined by the format
1	The stream_id field contains a valid StreamID

4.4.5.3 format_specific_data field

The 9-bit **format_specific_data** field is available for use by the given control format as specified by the **subtype** field. The specific use of this field is defined by the format.

4.4.5.4 control_data_length field

The 11-bit **control_data_length** field contains the length (in octets) of the **control_data_payload** field.

The maximum value for this field is determined by the maximum transmission unit (MTU) and shall be set to allow all control AVTPDUs to fit in a single standard frame payload.

4.4.5.5 stream_id field

The **stream_id** field is used for stream identification.

If the **sv** field is set to one (1), the **stream_id** field shall contain a valid StreamID, as defined in subclause 35.2.2.8.2 of IEEE Std 802.1Q-2014, and as shown for convenience in Figure 12.

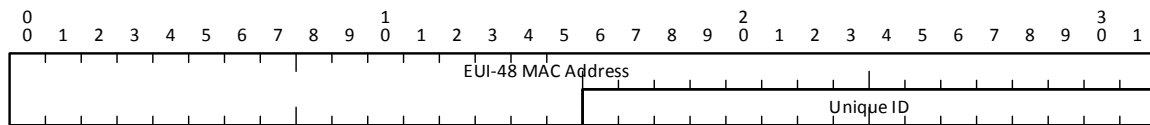


Figure 12—AVTPDU StreamID Format

If the **sv** field is set to zero (0), then the **stream_id** field is reserved unless a specific definition is given by the format.

4.4.5.6 control_data_payload field

The **control_data_payload** contains the data that is associated with or defined by a format. The actual data encapsulation is defined by the format.

4.4.6 AVTPDU alternative header

4.4.6.1 Overview

This subclause documents the fields that are used for formats that do not conform to the common stream or control headers described in 4.4.4 and 4.4.5 respectively.

The usage of the **h** (header_specific) bit (4.4.3.3) of the common header is defined by formats that use the alternative header. If not defined by the format, **h** is reserved.

In addition to the common header (see 4.4.3), one additional field is defined:

- **alternative_data_payload**: 0 to n octets

These fields are shown in Figure 13.

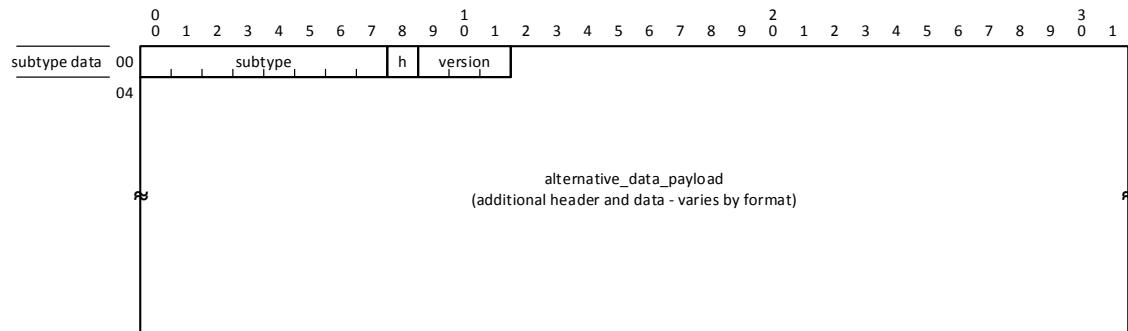


Figure 13—Alternative AVTPDU header

4.4.6.2 alternative_data_payload field

The **alternative_data_payload** field is used by formats to define format-specific headers and, as applicable, a data payload that is used as part of the format. Actual implementation is defined by each format that uses the alternative header.

5. IEC 61883/IIDC Format

5.1 Overview

This clause describes the AVTPDU formats to adapt the IEC 61883 and IIDC protocols to an IEEE 802 network, as well as to adapt the protocols to work with timing and synchronization based on gPTP. This clause does not attempt to document the IEC 61883 protocols, only the changes needed to work in an IEEE 802.1 AVB environment.

AVTP adapts the following IEEE 1394/IEC 61883-type protocols to run in an IEEE 802 environment:

- 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
- IIDC2 Digital Camera Control Specification Ver. 1.0.0, Jan. 2012

5.2 IEC 61883/IIDC stream data encapsulation

5.2.1 Overview

This encapsulation shall use the common stream header and shall have a **subtype** of 61883_IIDC (see 4.4.3.2).

This encapsulation uses the following common stream fields (see 4.4.3 and 4.4.4 for definitions):

- **sv**: 1 bit
- **version**: 3 bits
- **mr**: 1 bit
- **tv**: 1 bit
- **sequence_num**: 1 octet
- **tu**: 1 bit
- **stream_id**: 8 octets
- **avtp_timestamp**: 4 octets
- **stream_data_length**: 2 octets

This encapsulation uses the **f_s_d**, **format_specific_data_1**, **format_specific_data_2**, **format_specific_data_3**, and **stream_data_payload** fields of the common stream header to define the following format-specific fields:

- **r** (reserved): 1 bit
- **gv** (gateway_info valid): 1 bit
- **reserved**: 7 bits
- **gateway_info**: 4 octets
- **tag**: 2 bits
- **channel**: 6 bits
- **tcode**: 4 bits
- **sy**: 4 bits
- **61883_iidc_data_payload**: 0 to *n* octets

These fields are shown in Figure 14.

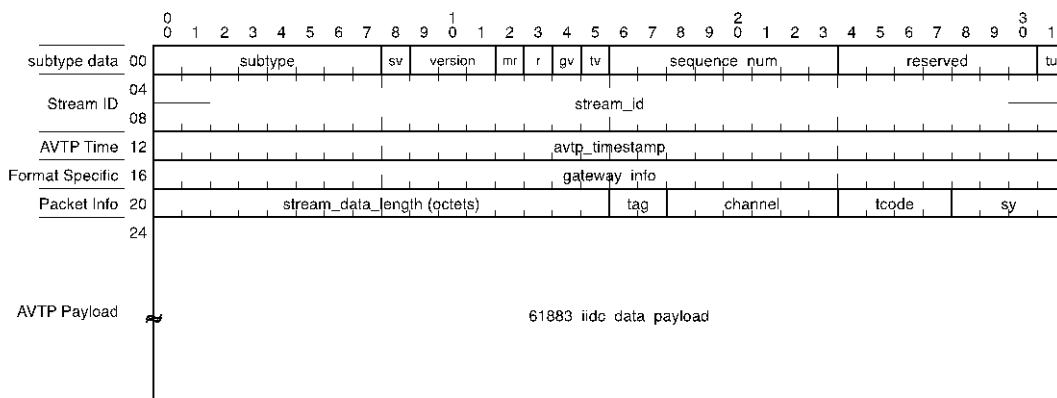


Figure 14—IEC 61883/IIDC common header fields

5.2.2 gv (gateway_info valid) field

The **gv** (**gateway_info** valid) field is reserved for use by AVTP gateways that are not defined by this standard.

AVTP end stations that do not have an AVTP gateway function shall set this field to zero (0) on transmit and ignore this field on receive.

5.2.3 gateway_info field

The 4-octet **gateway_info** field is reserved for use by AVTP gateways that are not defined by this standard.

AVTP end stations that do not have an AVTP gateway function shall set this field to zero (0) on transmit and ignore this field on receive.

5.2.4 tag field

The 2-bit **tag** field follows the same meaning, format, and rules as specified by IEEE Std 1394. Of the four possible combinations for this field, the following values are supported or not supported as specified in this clause.

Supported by AVTP:

- 00₂: No common isochronous packet (CIP) header included (used by Instrumentation and Industrial Digital Camera (IIDC) 1394 trade association specification) (see 5.3)
- 01₂: CIP header included (see 5.4.3.)

Not supported by AVTP:

- 10₂: Reserved by IEEE Std 1394.1™-2004 clock adjustment [B13]
- 11₂: Global asynchronous stream packet format (used in IEEE Std 1394 for serial bus to serial bus bridges)

5.2.5 channel field

The 6-bit **channel** field follows the same meaning, format, and rules as specified by IEEE Std 1394. The following values are supported as specified:

- 0 to 30 and 32 to 63: Originating channel ID from an IEEE 1394 serial bus via IEEE 1394/IEC 61883 to IEEE 1722/IEC 61883 gateway
- 31: Originating source is on AVTP network (native AVTP)

5.2.6 tcode (type code) field

The 4-bit **tcode** field follows the same meaning, format, and rules as specified by IEEE Std 1394. For AVTP, the only value supported shall be a fixed value of 1010₂ (A₁₆, same as IEEE 1394 isochronous packet format) with the following rules for Talkers and Listeners:

- AVTP Talkers shall set this field to A₁₆ on transmit
- AVTP Listeners shall ignore this field on receive

5.2.7 sy field

Use of the 4-bit **sy** field is application specific and therefore beyond the scope of this standard. The following external standards are known to use this field:

- IIDC 1394-based Digital Camera Specification (used for video start of frame indicator)
- Digital Transmission Content Protection (DTCP)² (DTCP Specification [B7], DTCP Volume 1 Supplement [B8])

5.2.8 61883_iidc_data_payload field

The **61883_iidc_data_payload** field contains the data that is associated with or defined by this format. The actual data encapsulation is defined by the **tag** field and given in 5.3.

5.3 IIDC encapsulation

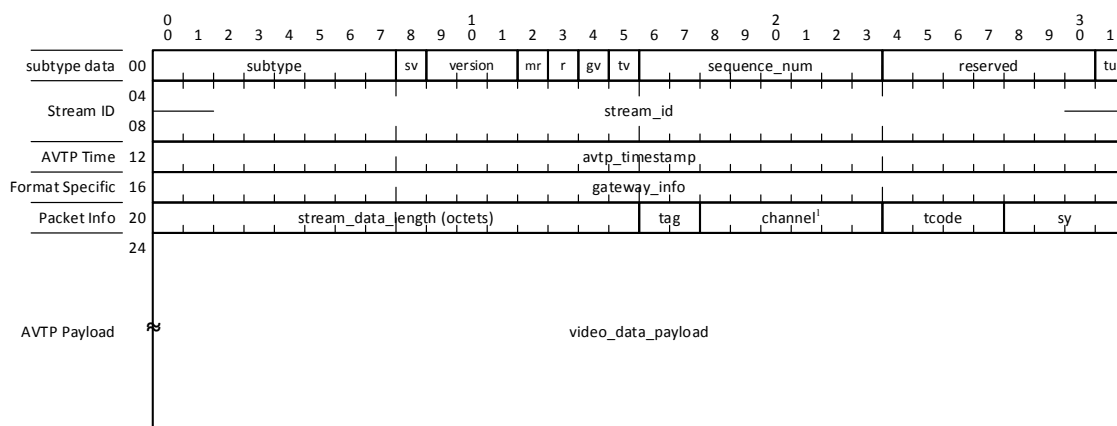
5.3.1 Overview

This subclause describes the AVTPDU format to be used when transporting IIDC data. This option is selected by the **tag** value being set to zero (0).

This encapsulation uses the **61883_iidc_data_payload** field of the IEC 61883/IIDC common header to define the following format-specific field:

- **video_data_payload**: 0 to *n* octets

This is shown in Figure 15.



Notes to figure:

¹Current standard for IIDC restricts channel ID to 0 - 15

Figure 15—IIDC AVTPDU header fields

²Available at <http://www.dtcp.com>.

5.3.2 video_data_payload field

5.3.2.1 Overview

The **video_data_payload** field contains the portion of the video frame data being carried by the AVTPDU. The amount of data within the **video_data_payload** field is defined by Table 8. The IIDC Pixel Encoding Format is defined by Table I.1.

Table 8—IIDC video data payload sizes and rates

IIDC Pixel Encoding Format name	Octets per AVTPDU	AVTPDUs per video frame
VGA_MONO8	1280	240
VGA_RGB8_PACKED	1440	640
VGA_YUV422_PACKED	1280	480
VGA_ARGB32	1280	960
VGA_BGRA32	1280	960
XGA_MONO8	1024	768
XGA_RGB8_PACKED	1152	2048
XGA_YUV422_PACKED	1024	1536
XGA_ARGB32	1024	3072
XGA_BGRA32	1024	3072
HD720P_MONO8	1440	720
HD720P_RGB8_PACKED	1440	1920
HD720P_YUV422_PACKED	1440	1280
HD720P_ARGB32	1440	2560
HD720P_BGRA32	1440	2560
SXGA_MONO8	1280	960
SXGA_RGB8_PACKED	1440	2560
SXGA_YUV422_PACKED	1280	1920
SXGA_ARGB32	1280	3840
SXGA_BGRA32	1280	3840
UXGA_MONO8	1280	1500
UXGA_RGB8_PACKED	1440	4000
UXGA_YUV422_PACKED	1280	3000
UXGA_ARGB32	1280	6000
UXGA_BGRA32	1280	6000
HD1080P_MONO8	1440	1440
HD1080P_RGB8_PACKED	1440	4320
HD1080P_YUV422_PACKED	1440	2880
HD1080P_ARGB32	1440	5760
HD1080P_BGRA32	1440	5760
HVGA_MONO8	1280	60
HVGA_RGB8_PACKED	1440	160
HVGA_YUV422_PACKED	1280	120
HVGA_ARGB32	1280	240
HVGA_BGRA32	1280	240
CUSTOM	—	—

The color space names and their references back to IIDC2 are defined in Clause 7.3.10 of IEEE Std 1722.1-2013.

5.3.2.2 MONO8 color space pixel encoding

The MONO8 color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.1 and IIDC2 Ver 1.0.0 Section 6.6.3.1.

5.3.2.3 MONO12_PACKED color space pixel encoding

The MONO12_PACKED color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.2 and IIDC2 Ver 1.0.0 Section 6.6.3.5.

5.3.2.4 MONO16 color space pixel encoding

The MONO16 color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.3 and IIDC2 Ver 1.0.0 Section 6.6.3.3.

5.3.2.5 MONO16_SIGNED color space pixel encoding

The MONO16_SIGNED color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.4 and IIDC2 Ver 1.0.0 Section 6.6.3.4.

5.3.2.6 YUV411_PACKED color space pixel encoding

The YUV411_PACKED color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.5 and IIDC2 Ver 1.0.0 Section 6.6.3.2.

5.3.2.7 YUV422_PACKED color space pixel encoding

The YUV422_PACKED color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.6 and IIDC2 Ver 1.0.0 Section 6.6.3.2.

5.3.2.8 YUV444_PACKED color space pixel encoding

The YUV444_PACKED color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.7 and IIDC2 Ver 1.0.0 Section 6.6.3.2.

5.3.2.9 RGB_PACKED color space pixel encoding

The RGB_PACKED color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.8 and IIDC2 Ver 1.0.0 Section 6.6.3.1.

5.3.2.10 RGB16 color space pixel encoding

The RGB16 color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.9 and IIDC2 Ver 1.0.0 Section 6.6.3.3.

5.3.2.11 RGB16_SIGNED color space pixel encoding

The RGB16_SIGNED color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.10 and IIDC2 Ver 1.0.0 Section 6.6.3.4.

5.3.2.12 BAYERGR8 color space pixel encoding

The BAYERGR8 color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.11 and IIDC2 Ver 1.0.0 Section 6.6.3.1.

5.3.2.13 BAYERGR12_PACKED color space pixel encoding

The BAYERGR12_PACKED color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.12 and IIDC2 Ver 1.0.0 Section 6.6.3.5.

5.3.2.14 BAYERGR16 color space pixel encoding

The BAYERGR16 color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.13 and IIDC2 Ver 1.0.0 Section 6.6.3.3.

5.3.2.15 BAYERRG8 color space pixel encoding

The BAYERRG8 color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.14 and IIDC2 Ver 1.0.0 Section 6.6.3.1.

5.3.2.16 BAYERRG12_PACKED color space pixel encoding

The BAYERRG12_PACKED color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.15 and IIDC2 Ver 1.0.0 Section 6.6.3.5.

5.3.2.17 BAYERRG16 color space pixel encoding

The BAYERGR8 color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.16 and IIDC2 Ver 1.0.0 Section 6.6.3.3.

5.3.2.18 BAYERGB8 color space pixel encoding

The BAYERGB8 color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.17 and IIDC2 Ver 1.0.0 Section 6.6.3.1.

5.3.2.19 BAYERGB12_PACKED color space pixel encoding

The BAYERGB12_PACKED color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.18 and IIDC2 Ver 1.0.0 Section 6.6.3.5.

5.3.2.20 BAYERGB16 color space pixel encoding

The BAYERGB16 color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.19 and IIDC2 Ver 1.0.0 Section 6.6.3.3.

5.3.2.21 BAYERBG8 color space pixel encoding

The BAYERBG8 color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.20 and IIDC2 Ver 1.0.0 Section 6.6.3.1.

5.3.2.22 BAYERBG12_PACKED color space pixel encoding

The BAYERBG12_PACKED color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.21 and IIDC2 Ver 1.0.0 Section 6.6.3.5.

5.3.2.23 BAYERBG16 color space pixel encoding

The BAYERGB16 color space pixels are encoded as defined in IIDC2 Ver 1.0.0 Section 6.6.2.2 and IIDC2 Ver 1.0.0 Section 6.6.3.3.

5.3.2.24 ARGB32 color space pixel encoding

Each component of the pixel is 8-bits of data and is layed out within the AVTPDU as defined in Figure 16. The data encoding of each component is defined by IIDC2 Ver 1.0.0 Section 6.6.3.1.

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
A[K+0]										R[K+0]										G[K+0]										B[K+0]									
A[K+1]										R[K+1]										G[K+1]										B[K+1]									
A[K+Pn-2]										R[K+Pn-2]										G[K+Pn-2]										B[K+Pn-2]									
A[K+Pn-1]										R[K+Pn-1]										G[K+Pn-1]										B[K+Pn-1]									

Figure 16—AGRB32 pixel encoding

5.3.2.25 BGRA32 color space pixel encoding

Each component of the pixel is 8-bits of data and is layed out within the AVTPDU as defined in Figure 17. The data encoding of each component is defined by IIDC2 Ver 1.0.0 Section 6.6.3.1.

0										1										2										3									
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1								
B[K+0]										G[K+0]										R[K+0]										A[K+0]									
B[K+1]										G[K+1]										R[K+1]										A[K+1]									
B[K+Pn-2]										G[K+Pn-2]										R[K+Pn-2]										A[K+Pn-2]									
B[K+Pn-1]										G[K+Pn-1]										R[K+Pn-1]										A[K+Pn-1]									

Figure 17—BGRA32 pixel encoding

5.3.2.26 YUV420_TRI_FULL color space pixel encoding

The YUV420_TRI_FULL color space pixel encoding is a J420 planar representation. Unlike the other pixel encodings YUV420_TRI_FULL groups all of the luminance (Y) values together followed by all of the U chroma values and then all of the V chroma values. As such the entire encoded video frame, as shown in Figure 18 is split into AVTPDUs at the appropriate octet boundaries. The data encoding of each component is defined by IIDC2 Ver 1.0.0 Section 6.6.3.2.

0								1								2								3							
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1
Y[0]								Y[1]								Y[2]								Y[3]							
Y[N-4]								Y[N-3]								Y[N-2]								Y[N-1]							
U[0]								U[1]								U[2]								U[3]							
U[N/2-4]								U[N/2-3]								U[N/2-2]								U[N/2-1]							
V[0]								V[1]								V[2]								V[3]							
V[N/2-4]								V[N/2-3]								V[N/2-2]								V[N/2-1]							

Figure 18—YUV420_TRI_FULL pixel encoding for whole video frame

5.3.2.27 YUV420_TRI_VIDEO color space pixel encoding

The YUV420_TRI_VIDEO color space pixel encoding is an I420 planar representation. Unlike the other pixel encodings YUV420_TRI_VIDEO groups all of the luminance (Y) values together followed by all of the U chroma values and then all of the V chroma values. As such the entire encoded video frame, as shown in Figure 19, is split into AVTPDUs at the appropriate octet boundaries. The data encoding of each component is defined by IIDC2 Ver 1.0.0 Section 6.6.3.2.

0	1	2	3
0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1
Y[0]	Y[1]	Y[2]	Y[3]
Y[N-4]	Y[N-3]	Y[N-2]	Y[N-1]
U[0]	U[1]	U[2]	U[3]
U[N/2-4]	U[N/2-3]	U[N/2-2]	U[N/2-1]
V[0]	V[1]	V[2]	V[3]
V[N/2-4]	V[N/2-3]	V[N/2-2]	V[N/2-1]

Figure 19—YUV420_TRI_VIDEO pixel encoding for whole video frame

5.3.2.28 YUV420_BI_FULL color space pixel encoding

The YUV420_BI_FULL color space pixel encoding is a J420 like biplanar representation. Unlike the other pixel encodings YUV420_BI_FULL groups all of the luminance (Y) values together followed by all of the U chroma and V chroma value pairs. As such the entire encoded video frame, as shown in Figure 18 is split into AVTPDUs at the appropriate octet boundaries. The data encoding of each component is defined by IIDC2 Ver 1.0.0 Section 6.6.3.2.

0	1	2	3
0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9	0 1
Y[0]	Y[1]	Y[2]	Y[3]
Y[N-4]	Y[N-3]	Y[N-2]	Y[N-1]
U[0]	V[0]	U[1]	V[1]
U[N/2-2]	V[N/2-2]	U[N/2-1]	V[N/2-1]

Figure 20—YUV420_BI_FULL pixel encoding for whole video frame

5.3.2.29 YUV420_BI_VIDEO color space pixel encoding

The YUV420_BI_VIDEO color space pixel encoding is a NV12 biplanar representation. Unlike the other pixel encodings YUV420_BI_VIDEO groups all of the luminance (Y) values together followed by all of the the U chroma and V chroma value pairs. As such the entire encoded video frame, as shown in Figure 19, is split into AVTPDUs at the appropriate octet boundaries. The data encoding of each component is defined by IIDC2 Ver 1.0.0 Section 6.6.3.2.