

INTERNATIONAL STANDARD

NORME INTERNATIONALE



**Universal serial bus interfaces for data and power –
Part 1-7: Common components – USB Audio 3.0 device class definition data
formats**

**Interfaces de bus universel en série pour les données et l'alimentation
électrique –
Partie 1-7: Composants communs – Définition de classes de dispositifs USB
Audio 3.0 pour formats de données**



THIS PUBLICATION IS COPYRIGHT PROTECTED
Copyright © 1997-2016 USB-IF

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from IEC, or USB-IF at the respective address given below. Any questions about USB-IF copyright should be addressed to the USB-IF. Enquiries about obtaining additional rights to this publication and other information requests should be addressed to the IEC or your local IEC member National Committee.

IEC Central Office
3, rue de Varembe
CH-1211 Geneva 20
Switzerland
Tel.: +41 22 919 02 11
info@iec.ch
www.iec.ch

USB Implementers Forum, Inc.
3855 S.W. 153rd Drive
Beaverton, OR 97003
United States of America
Tel: +1 503-619-0426
admin@usb.org
www.usb.org

About the IEC

The International Electrotechnical Commission (IEC) is the leading global organization that prepares and publishes International Standards for all electrical, electronic and related technologies.

About IEC publications

The technical content of IEC publications is kept under constant review by the IEC. Please make sure that you have the latest edition, a corrigendum or an amendment might have been published.

IEC publications search - webstore.iec.ch/advsearchform

The advanced search enables to find IEC publications by a variety of criteria (reference number, text, technical committee,...). It also gives information on projects, replaced and withdrawn publications.

IEC Just Published - webstore.iec.ch/justpublished

Stay up to date on all new IEC publications. Just Published details all new publications released. Available online and once a month by email.

IEC Customer Service Centre - webstore.iec.ch/csc

If you wish to give us your feedback on this publication or need further assistance, please contact the Customer Service Centre: sales@iec.ch.

Electropedia - www.electropedia.org

The world's leading online dictionary on electrotechnology, containing more than 22 000 terminological entries in English and French, with equivalent terms in 16 additional languages. Also known as the International Electrotechnical Vocabulary (IEV) online.

IEC Glossary - std.iec.ch/glossary

67 000 electrotechnical terminology entries in English and French extracted from the Terms and Definitions clause of IEC publications issued since 2002. Some entries have been collected from earlier publications of IEC TC 37, 77, 86 and CISPR.

IECNORM.COM : Click to view the full PDF © IEC 62689-1-1:2019

INTERNATIONAL STANDARD

NORME INTERNATIONALE



**Universal serial bus interfaces for data and power –
Part 1-7: Common components – USB Audio 3.0 device class definition data
formats**

**Interfaces de bus universel en série pour les données et l'alimentation
électrique –
Partie 1-7: Composants communs – Définition de classes de dispositifs USB
Audio 3.0 pour formats de données**

INTERNATIONAL
ELECTROTECHNICAL
COMMISSION

COMMISSION
ELECTROTECHNIQUE
INTERNATIONALE

ICS 33.160; 35.100.20

ISBN 978-2-8322-7243-5

**Warning! Make sure that you obtained this publication from an authorized distributor.
Attention! Veuillez vous assurer que vous avez obtenu cette publication via un distributeur agréé.**

INTERNATIONAL ELECTROTECHNICAL COMMISSION

UNIVERSAL SERIAL BUS INTERFACES FOR DATA AND POWER –

**Part 1-7: Common components –
USB Audio 3.0 device class definition data formats**

FOREWORD

- 1) The International Electrotechnical Commission (IEC) is a worldwide organization for standardization comprising all national electrotechnical committees (IEC National Committees). The object of IEC is to promote international co-operation on all questions concerning standardization in the electrical and electronic fields. To this end and in addition to other activities, IEC publishes International Standards, Technical Specifications, Technical Reports, Publicly Available Specifications (PAS) and Guides (hereafter referred to as “IEC Publication(s)”). Their preparation is entrusted to technical committees; any IEC National Committee interested in the subject dealt with may participate in this preparatory work. International, governmental and non-governmental organizations liaising with the IEC also participate in this preparation. IEC collaborates closely with the International Organization for Standardization (ISO) in accordance with conditions determined by agreement between the two organizations.
- 2) The formal decisions or agreements of IEC on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC National Committees.
- 3) IEC Publications have the form of recommendations for international use and are accepted by IEC National Committees in that sense. While all reasonable efforts are made to ensure that the technical content of IEC Publications is accurate, IEC cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.
- 4) In order to promote international uniformity, IEC National Committees undertake to apply IEC Publications transparently to the maximum extent possible in their national and regional publications. Any divergence between any IEC Publication and the corresponding national or regional publication shall be clearly indicated in the latter.
- 5) IEC itself does not provide any attestation of conformity. Independent certification bodies provide conformity assessment services and, in some areas, access to IEC marks of conformity. IEC is not responsible for any services carried out by independent certification bodies.
- 6) All users should ensure that they have the latest edition of this publication.
- 7) No liability shall attach to IEC or its directors, employees, servants or agents including individual experts and members of its technical committees and IEC National Committees for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication, use of, or reliance upon, this IEC Publication or any other IEC Publications.
- 8) Attention is drawn to the Normative references cited in this publication. Use of the referenced publications is indispensable for the correct application of this publication.
- 9) Attention is drawn to the possibility that some of the elements of this IEC Publication may be the subject of patent rights. IEC shall not be held responsible for identifying any or all such patent rights.

International Standard IEC 62680-1-7 has been prepared by technical area 18: Multimedia home systems and applications for end-user networks, of IEC technical committee 100: Audio, video and multimedia systems and equipment.

The text of this standard was prepared by the USB Implementers Forum (USB-IF). The structure and editorial rules used in this publication reflect the practice of the organization which submitted it.

The text of this International Standard is based on the following documents:

CDV	Report on voting
100/3159/CDV	100/3229/RVC

Full information on the voting for the approval of this International Standard can be found in the report on voting indicated in the above table.

Copyright © 1997-2016 USB Implementers Forum, Inc. All rights reserved.

The committee has decided that the contents of this document will remain unchanged until the stability date indicated on the IEC website under "<http://webstore.iec.ch>" in the data related to the specific document. At this date, the document will be

- reconfirmed,
- withdrawn,
- replaced by a revised edition, or
- amended.

IMPORTANT – The 'colour inside' logo on the cover page of this publication indicates that it contains colours which are considered to be useful for the correct understanding of its contents. Users should therefore print this document using a colour printer.

IECNORM.COM : Click to view the full PDF of IEC 62680-1-7:2019

INTRODUCTION

The IEC 62680 series is based on a series of specifications that were originally developed by the USB Implementers Forum (USB-IF). These specifications were submitted to the IEC under the auspices of a special agreement between the IEC and the USB-IF.

This standard is the USB-IF publication USB Device Class Definition for Audio Data Formats Release 3.0.

The USB Implementers Forum, Inc.(USB-IF) is a non-profit corporation founded by the group of companies that developed the Universal Serial Bus specification. The USB-IF was formed to provide a support organization and forum for the advancement and adoption of Universal Serial Bus technology. The Forum facilitates the development of high-quality compatible USB peripherals (devices), and promotes the benefits of USB and the quality of products that have passed compliance testing.

ANY USB SPECIFICATIONS ARE PROVIDED TO YOU "AS IS, "WITH NO WARRANTIES WHATSOEVER, INCLUDING ANY WARRANTY OF MERCHANTABILITY, NON-INFRINGEMENT, OR FITNESS FOR ANY PARTICULAR PURPOSE. THE USB IMPLEMENTERS FORUM AND THE AUTHORS OF ANY USB SPECIFICATIONS DISCLAIM ALL LIABILITY, INCLUDING LIABILITY FOR INFRINGEMENT OF ANY PROPRIETARY RIGHTS, RELATING TO USE OR IMPLEMENTATION OR INFORMATION IN THIS SPECIFICATION.

THE PROVISION OF ANY USB SPECIFICATIONS TO YOU DOES NOT PROVIDE YOU WITH ANY LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS.

Entering into USB Adopters Agreements may, however, allow a signing company to participate in a reciprocal, RAND-Z licensing arrangement for compliant products. For more information, please see:

<https://www.usb.org/documents>

IEC DOES NOT TAKE ANY POSITION AS TO WHETHER IT IS ADVISABLE FOR YOU TO ENTER INTO ANY USB ADOPTERS AGREEMENTS OR TO PARTICIPATE IN THE USB IMPLEMENTERS FORUM."

UNIVERSAL SERIAL BUS DEVICE CLASS DEFINITION FOR AUDIO DATA FORMATS

IECNORM.COM : Click to view the full PDF of IEC 62680-1-7:2019

Release 3.0
September 22, 2016

SCOPE OF THIS RELEASE

This document is the Release 3.0 of this device class definition.

CONTRIBUTORS

Joe Scanlon
Rhoads Hollowell
Girault Jones
Matthew X. Mora
Tzung-Dar Tsai
Brad Lambert
Dan Bogard
Pete Burgers
David Roh
Leng Ooi
Pierre-Louis Bossart
David Hines
Abdul Rahman Ismail (Co-Chair)
Devon Worrell
Chandrashekhar Rao
Terry Moore
Alex Lin
Bala Sivakumar
Geert Knapen (Co-Chair & Editor)

Advanced Micro Devices
Apple Inc.
Apple Inc.
Apple Inc.
C-Media Electronics, Inc.
Cirrus Logic, Inc.
Conexant Systems, Inc.
DisplayLink (UK), Ltd.
Dolby Laboratories, Inc.
Google, Inc.
Intel Corporation
Intel Corporation
Intel Corporation
Intel Corporation
Logitech, Inc.
MCCI Corporation
MediaTek, Inc.
Microsoft Corporation
NXP Semiconductors

PL Mobile Audio
411 E. Plumeria drive
San Jose, CA 95134, USA
E-mail: geert.knapen@nxp.com

James Goel
Andre Schevciw
Jin-Sheng Wang
Morten Christiansen

Qualcomm, Inc.
Qualcomm, Inc.
Qualcomm, Inc.
Synopsys

REVISION HISTORY

Revision	Date	Filename	Description
1.0	Mar. 18, 98	Frmts10.pdf	Release 1.0
2.0	May 31, 06	Frmts20 final.pdf	Release 2.0
3.0	Sep. 22, 16	Frmts30.pdf	Release 3.0

**Copyright © 1997-2016 USB Implementers Forum, Inc.
All rights reserved.**

INTELLECTUAL PROPERTY DISCLAIMER

A LICENSE IS HEREBY GRANTED TO REPRODUCE THIS SPECIFICATION FOR INTERNAL USE ONLY. NO OTHER LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, IS GRANTED OR INTENDED HEREBY.

USB-IF AND THE AUTHORS OF THIS SPECIFICATION EXPRESSLY DISCLAIM ALL LIABILITY FOR INFRINGEMENT OF INTELLECTUAL PROPERTY RIGHTS RELATING TO IMPLEMENTATION OF INFORMATION IN THIS SPECIFICATION. USB-IF AND THE AUTHORS OF THIS SPECIFICATION ALSO DO NOT WARRANT OR REPRESENT THAT SUCH IMPLEMENTATION(S) WILL NOT INFRINGE THE INTELLECTUAL PROPERTY RIGHTS OF OTHERS.

THIS SPECIFICATION IS PROVIDED “AS IS” AND WITH NO WARRANTIES, EXPRESS OR IMPLIED, STATUTORY OR OTHERWISE. ALL WARRANTIES ARE EXPRESSLY DISCLAIMED. USB-IF, ITS MEMBERS AND THE AUTHORS OF THIS SPECIFICATION PROVIDE NO WARRANTY OF MERCHANTABILITY, NO WARRANTY OF NON-INFRINGEMENT, NO WARRANTY OF FITNESS FOR ANY PARTICULAR PURPOSE, AND NO WARRANTY ARISING OUT OF ANY PROPOSAL, SPECIFICATION, OR SAMPLE.

IN NO EVENT WILL USB-IF, MEMBERS OR THE AUTHORS BE LIABLE TO ANOTHER FOR THE COST OF PROCURING SUBSTITUTE GOODS OR SERVICES, LOST PROFITS, LOSS OF USE, LOSS OF DATA OR ANY INCIDENTAL, CONSEQUENTIAL, INDIRECT, OR SPECIAL DAMAGES, WHETHER UNDER CONTRACT, TORT, WARRANTY, OR OTHERWISE, ARISING IN ANY WAY OUT OF THE USE OF THIS SPECIFICATION, WHETHER OR NOT SUCH PARTY HAD ADVANCE NOTICE OF THE POSSIBILITY OF SUCH DAMAGES.

NOTE: VARIOUS USB-IF MEMBERS PARTICIPATED IN THE DRAFTING OF THIS SPECIFICATION. CERTAIN OF THESE MEMBERS MAY HAVE DECLINED TO ENTER INTO A SPECIFIC AGREEMENT LICENSING INTELLECTUAL PROPERTY RIGHTS THAT MAY BE INFRINGED IN THE IMPLEMENTATION OF THIS SPECIFICATION. PERSONS IMPLEMENT THIS SPECIFICATION AT THEIR OWN RISK.

Dolby™, AC-3™, Pro Logic™ and Dolby Surround™ are trademarks of Dolby Laboratories, Inc.

All other product names are trademarks, registered trademarks, or service marks of their respective owners.

Please send comments via electronic mail to audio-chair@usb.org

TABLE OF CONTENTS

Scope of This Release	6
Contributors	6
Revision History	6
Table of Contents	8
List of Tables	9
List of Figures	10
1 Introduction	10
1.1 Related Documents	11
1.2 Terms and Abbreviations	11
2 Audio Data Formats	13
2.1 Transfer Delimiter	14
2.2 Service Interval and Service Interval Packet Definitions	14
2.3 Simple Audio Data Formats	14
2.3.1 Type I Formats	14
2.3.2 Type III Formats	18
2.3.3 Type IV Formats	19
2.4 Extended Audio Data Formats	19
2.4.1 Extended Type I Formats	20
2.4.2 Extended Type III Formats	21
2.5 Class-specific AS Interface Descriptor	21
3 Auxiliary Protocols	23
3.1 HDCP Protocol	23
4 Adding New Audio Data Formats	24
5 Adding New Side Band Protocols	25
Appendix A. Additional Audio Device Class Codes	26
A.1 Audio Data Formats Bit Allocations	26
A.2 SubHeader Codes	27
A.3 Audio Format General Constants	27

LIST OF TABLES

Table 2-1: Packetization	16
Table 2-2: SIPDescriptor Layout	20
Table 2-3: Class-Specific AS Interface Descriptor	22
Table 3-1: HDCP SubHeader Layout	23
Table A-2: Audio Data Formats Bit Allocations in the bmFormats Field and Usage	26
Table A-3: SubHeader Codes	27
Table A-4: General Constants	27

IECNORM.COM : Click to view the full PDF of IEC 62680-1-7:2019

LIST OF FIGURES

Figure 2-1: Type I Audio Stream.....	13
Figure 2-2: Extended Type I Format.....	21
Figure 2-3: Extended Type III Format.....	21

IECNORM.COM : Click to view the full PDF of IEC 62680-1-7:2019

1 INTRODUCTION

The intention of this document is to describe in detail all the Audio Data Formats that are supported by the Audio Device Class. This document is considered an integral part of *the Audio Device Class Specification*, although subsequent revisions of this document are independent of the revision evolution of the main *USB Audio Specification*. This is to easily accommodate the addition of new Audio Data Formats without impeding the core *USB Audio Specification*.

1.1 RELATED DOCUMENTS

- *Universal Serial Bus Specification*, Revision 2.0 (referred to in this document as the *USB Specification*). In particular, see Chapter 5, “USB Data Flow Model” and Chapter 9, “USB Device Framework.”
- Universal Serial Bus Device Class Definition for Audio Devices (referred to in this document as USB Audio Device Class).
- Universal Serial Bus Device Class Definition for Terminal Types (referred to in this document as USB Audio Terminal Types).
- ANSI S1.11-1986 standard.
- MPEG-1 standard ISO/IEC 11172-3 1993. (available from <http://www.iso.ch>)
- MPEG-2 standard ISO/IEC 13818-3 Feb. 20, 1997. (available from <http://www.iso.ch>)
- Digital Audio Compression Standard (AC-3), ATSC A/52A Aug. 20, 2001. (available from <http://www.atsc.org>)
- Windows Media Audio (WMA) specification. (available from <http://www.microsoft.com>)
- ANSI/IEEE-754 floating-point standard.
- ISO/IEC 60958 International Standard: *Digital Audio Interface and Annexes*.
- ISO/IEC 61937 standard.
- ITU G.711 standard.
- ETSI Specification TS 102 114, “DTS Coherent Acoustics; Core and Extensions”. (Available from http://webapp.etsi.org/action%5CPU/20020827/ts_102114v010101p.pdf)

1.2 TERMS AND ABBREVIATIONS

This section defines terms used throughout this document. For additional terms that pertain to the Universal Serial Bus, see Chapter 2, “Terms and Abbreviations,” in the *USB Specification*.

AC-3	Audio compression standard from Dolby Labs.
Audio Slot	A collection of audio subslots, each containing a PCM audio sample of a different physical audio channel, taken at the same moment in time.
Audio Stream	A concatenation of a potentially very large number of audio slots ordered according to ascending time.
Audio Subslot	Holds a single PCM audio sample.
DTS	Acronym for Digital Theater Systems.
DVD	Acronym for Digital Versatile Disc.
Encoded Audio Bit Stream	A concatenation of a potentially very large number of encoded audio frames, ordered according to ascending time.

Encoded Audio Frame	A sequence of bits that contains an encoded representation of audio samples from one or more physical audio channels taken over a fixed period of time.
MPEG	Acronym for Moving Pictures Expert Group.
PCM	Acronym for Pulse Coded Modulation.
Service Interval	A grouping of USB (micro)frames or Bus Intervals that are related.
Service Interval Packet	A packet that contains all the audio slots that are transferred over the bus during a Service Interval.
Transfer Delimiter	A unique token that indicates an interruption in an isochronous data packet stream. Can be either a zero-length data packet or the absence of an isochronous transfer in a certain USB frame.
WMA	Acronym for Windows Media Audio.

IECNORM.COM : Click to view the full PDF of IEC 62680-1-7:2019

2 AUDIO DATA FORMATS

Audio Data formats can be divided in two main groups:

- Simple Audio Data Formats
- Extended Audio Data Formats

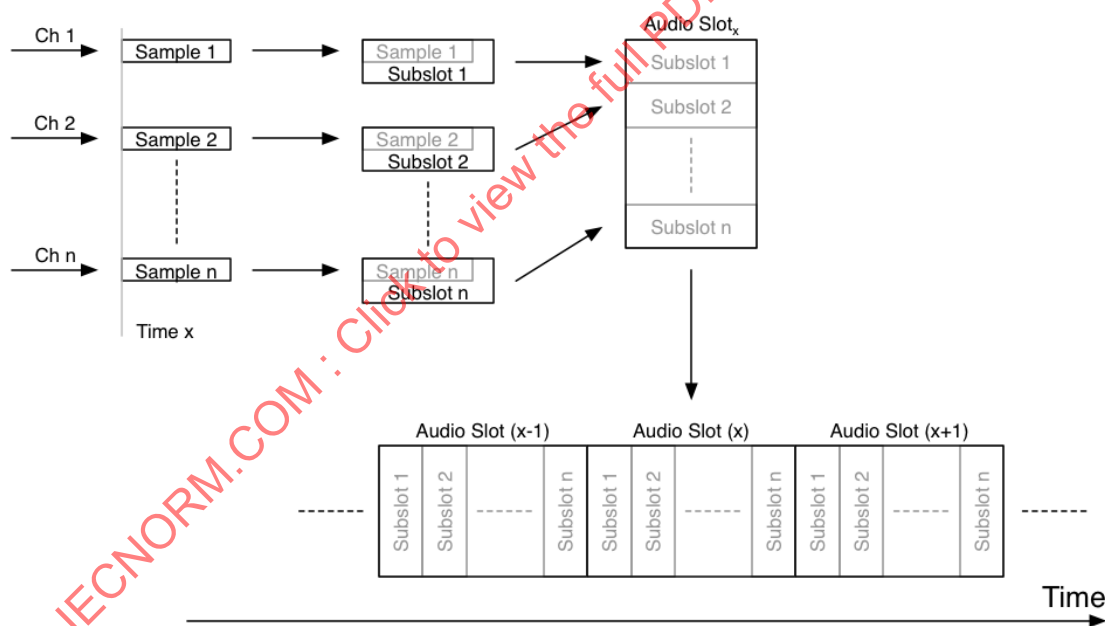
Simple Audio Data Formats can then be subdivided into three groups according to type.

The first group, Type I, deals with audio data streams that are transmitted over USB and are constructed on a sample-by-sample basis. Each audio sample is represented by a single independent symbol, contained in an audio subslot. Different compression schemes may be used to transform the audio samples into symbols.

Note: This is different from encoding. Compression is considered to take place on a per-audio-sample base. Each audio sample generates one symbol (e.g. A-law compression where a 16-bit audio sample is compressed into an 8-bit symbol).

If multiple physical audio channels are formatted into a single audio channel cluster, then samples at time x of subsequent channels are first contained into audio subslots. These audio subslots are then interleaved, according to the cluster channel ordering as described in the main *USB Audio Specification*, and then grouped into an audio slot. The audio samples, taken at time $x+1$, are interleaved in the same fashion to generate the next audio slot and so on. The notion of physical channels is explicitly preserved during transmission. A typical example of Type I formats is the standard PCM audio data. The following figure illustrates the concept.

Figure 2-1: Type I Audio Stream



The second group, Type III, contains Audio Data Formats that use encapsulation as described in the ISO/IEC 61937 standard before being sent over USB. One or more non-PCM encoded audio data streams are packed into “pseudo-stereo samples” and transmitted as if they were real stereo PCM audio samples. The sampling frequency of these pseudo samples (transport sampling frequency, as reported by the Clock Frequency Control of the associated Clock Source Entity) either matches the sampling frequency of the original non-encoded PCM audio data streams (native sampling frequency) or there is an integer ratio relationship between them. Therefore, clock recovery at the receiving end is relatively easy. The drawback is that unless multiple non-PCM encoded streams are packed into one pseudo stereo stream, more bandwidth than necessary is consumed.

The third group, Type IV, deals with audio streams that are not transmitted over USB. Instead, they interface with the Audio Function through an AudioStreaming interface that does not contain a USB isochronous IN or OUT endpoint. These streams typically connect via a digital interface like S/PDIF (or some other means of connectivity) and may require interaction from the Host before they enter or leave the Audio Function. A

typical example is an external S/PDIF connector that can accept an AC-3 encoded audio stream. This stream is first processed by an AC-3 decoder before the (decoded) logical audio channels enter the Audio Function through the Input Terminal that represents this S/PDIF connection.

In addition to the Simple Audio Data Formats described above, Extended Audio Data Formats are defined. These are based on the Simple Audio Data Formats Type I and III definitions but they provide an optional packet header and for the Extended Audio Data Format Type I, an optional synchronous (i.e. sample accurate) control channel. Type IV Audio Data Formats do not have an Extended Audio Data Format definition.

The following sections explain the different Audio Data Formats and Format Types in more detail.

2.1 TRANSFER DELIMITER

Isochronous data streams are continuous in nature, although the actual number of bytes sent per packet may vary throughout the lifetime of the stream (for rate adaptation purposes for instance). To indicate a temporary stop in the isochronous data stream without closing the pipe (and thus relinquishing the USB bandwidth), an in-band Transfer Delimiter needs to be defined. This specification considers two situations to be a Transfer Delimiter. The first is a zero-length data packet and the second is the absence of an isochronous transfer in a USB (micro)frame that would normally have an isochronous transfer. Both situations are considered equivalent and the Audio Function is expected to behave the same. However, the second type consumes less isochronous USB bandwidth (i.e. zero bandwidth). In both cases, this specification considers a Transfer Delimiter to be an entity that can be sent over the USB.

2.2 SERVICE INTERVAL AND SERVICE INTERVAL PACKET DEFINITIONS

Note: The *USB Audio 2.0 Specification* used the terms Virtual Frame and Virtual Frame Packet to describe the same concepts that were called Service Interval and Service Interval Packet in the *USB 3.0 and higher Specifications*. In this specification, the terminology has been consolidated to use the more general terms of Service Interval and Service Interval Packet instead of the USB Audio 2.0-specific terminology of Virtual Frame and Virtual Frame Packet. Also, the term Bus Interval is used instead of USB (micro)frame, wherever applicable.

In the following paragraphs, the packetizing process for audio is described in terms of Service Interval and Service Interval Packets. This provides a consistent model of ‘one Service Interval Packet (SIP) per Service Interval (SI)’, irrespective of the actual transactions on the USB and the version of USB used.

A Service Interval is defined as:

$$\text{Service Interval} = \text{Bus Interval} * 2^{(b\text{Interval}-1)}$$

where Bus Interval has a value of 1 ms for full-speed isochronous endpoints and 125 μs for high-speed, SuperSpeed, and Enhanced SuperSpeed isochronous endpoints and where bInterval is the value specified in the bInterval field of the standard endpoint descriptor.

A Service Interval Packet is defined as the amount of isochronous data that is transported during an entire Service Interval. For high-speed high-bandwidth endpoints, the Service Interval Packet is the concatenation of the two or three physical packets that are transferred over the bus in a Bus Interval.

Note: The *USB Specification* already considers the 2 or 3 transactions of a high-speed high-bandwidth transfer to be part of a single packet.

2.3 SIMPLE AUDIO DATA FORMATS

2.3.1 TYPE I FORMATS

The following sections describe the Audio Data Formats that belong to Type I. A number of terms and their definition are presented.

2.3.1.1 USB PACKETS

Audio data streams that are inherently continuous shall be packetized when sent over the USB. The quality of the packetizing algorithm directly influences the amount of effort needed to reconstruct a reliable sample clock at the receiving side.

Furthermore, the chosen size of the Service Interval has a direct impact on the amount of buffer memory needed on both sides of the pipe and also on the incurred latency over the pipe. Shorter Service Intervals minimize buffer requirements and therefore also latency at the potential expense of higher power consumption. Indeed, longer Service Intervals potentially allow the bus (and parts of the sender and receiver's hardware) to enter lower power states for longer periods of time, thus conserving more power.

This specification defines two possible modes to transport audio data streams over the USB:

- Continuous Mode
- Burst Mode

Continuous Mode occurs when the Service Interval is smaller or equal to one USB Frame time of 1 ms. This mode minimizes buffer requirements and latency at the potential expense of higher power consumption. Note however that choosing a Service Interval that is smaller than one USB Frame time may result in excessive system level interrupts.

Burst Mode occurs when the Service Interval is larger than one USB Frame time. This mode provides for opportunities to save more power by allowing various system components to enter low power states for extended periods of time at the expense of larger buffer sizes and increased latency.

Devices may choose to expose multiple Alternate Settings of their AudioStreaming interface(s) with different Service Interval settings for each Alternate Setting, thus allowing the Host to choose a setting that best fits the desired use case. However, all devices shall expose at least one Alternate Setting (besides the zero bandwidth Alternate setting 0) that supports Continuous Mode (Service Interval ≤ 1 ms).

2.3.1.1.1 SERVICE INTERVAL PACKET SIZE CALCULATION

The goal shall be to keep the instantaneous number of audio slots per SI, n_i as close as possible to the average number of audio slots per SI, n_{av} . The average n_{av} shall be calculated as follows:

$$n_{av} = \frac{T_{SI}}{\Delta t}$$

where T_{SI} is the duration of an SI and Δt is the sample time ($1/F_s$). In most cases, n_{av} will be a number with a fractional part.

If the sampling rate is a constant, the allowable variation on n_i is limited to one audio slot, that is, $\Delta n_i = 1$. This implies that all SIPs shall either contain $INT(n_{av})$ audio slots (small SIP) or $INT(n_{av}) + 1$ (large SIP) audio slots. For all i :

$$n_i = INT(n_{av}) \text{ or } INT(n_{av}) + 1$$

Note: In the case where $n_{av} = INT(n_{av})$, n_i may vary between $INT(n_{av}) - 1$ (small SIP), $INT(n_{av})$ (medium SIP) and $INT(n_{av}) + 1$ (large SIP).

Furthermore, a large SIP shall be generated as soon as it becomes available. Typically, a source will generate a number of small SIPs as long as the accumulated fractional part of n_{av} remains < 1 . Once the accumulated fractional part of n_{av} becomes ≥ 1 , the source shall send a large SIP and decrement the accumulator by 1.

Due to possible different notions of time in the source and the sink (they might each have their own independent sampling clock), the (small SIP)/(large SIP) pattern generated by the source may be different from what the sink expects. Therefore, the sink shall be capable to accept a large SIP at all times.

Example:

Assume $F_s = 44,100$ Hz and $T_{VF} = 1$ ms. Then $n_{av} = 44.1$ audio slots. Since the source can only send an integer number of audio slots per SI, it will send small SIPs of 44 audio slots. Each SI, it therefore sends '0.1 slot' too few and it will accumulate this fractional part in an accumulator. After having sent 9 small SIPs of 44 audio slots, at the tenth SI it will have exactly one audio slot in excess and therefore can send a large SIP containing 45 audio slots. Decrementing the accumulator by 1 brings it back to 0 and the process can start all over again.

The source will thus produce a repetitive pattern of 9 small SIPs of 44 audio slots followed by 1 large SIP of 45 audio slots. The following table illustrates the process:

Table 2-1: Packetization

#SI	n_{av}	n_i	Fraction	Accumulator
n	44.1	44	0.1	0.1
n+1	44.1	44	0.1	0.2
n+2	44.1	44	0.1	0.3
n+3	44.1	44	0.1	0.4
n+4	44.1	44	0.1	0.5
n+5	44.1	44	0.1	0.6
n+6	44.1	44	0.1	0.7
n+7	44.1	44	0.1	0.8
n+8	44.1	44	0.1	0.9
n+9	44.1	45	0.1	1.0 -> 0
n+10	44.1	44	0.1	0.1
n+11	44.1	44	0.1	0.2
...	...	...	...	...

2.3.1.2 PITCH CONTROL

If the sampling rate can be varied (to implement pitch control), the allowable variation on n_i is limited to one audio slot per SI. For all i :

$$n_{i+1} = n_i | n_i \pm 1$$

Pitch control is restricted to adaptive endpoints only. AudioStreaming interfaces that support pitch control on their isochronous endpoint are required to report this in the class-specific endpoint descriptor. In addition, a Set/Get Pitch Control request is required to enable or disable the pitch control functionality.

2.3.1.3 AUDIO SUBSLOT

The basic structure used to represent audio data is the audio subslot. An audio subslot holds a single audio sample. An audio subslot always contains an integer number of bytes.

This specification limits the possible audio subslot sizes (**bSubslotSize**) to 1, 2, 3, 4, or 8 bytes per audio subslot. An audio sample is represented using a number of bits (**bBitResolution**) less than or equal to the total number of bits available in the audio subslot, i.e. **bBitResolution** ≤ **bSubslotSize***8.

AudioStreaming endpoints shall be constructed in such a way that a valid transfer can take place as long as the reported audio subslot size (**bSubslotSize**) is respected during transmission. If the reported bits per sample (**bBitResolution**) do not correspond with the number of significant bits actually used during transfer, the device will either discard trailing significant bits ([actual_bits_per_sample] > **bBitResolution**) or interpret trailing zeroes as significant bits ([actual_bits_per_sample] < **bBitResolution**).

2.3.1.4 AUDIO SLOT

An audio slot consists of a collection of audio subslots, each containing an audio sample of a different physical audio channel, taken at the same moment in time. The number of audio subslots in an audio slot equals the number of logical audio channels in the audio channel cluster. The ordering of the audio subslots in the audio slot obeys the rules set forth in the *USB Audio Specification*. All audio subslots shall have the same audio subslot size.

2.3.1.5 AUDIO STREAMS

An audio stream is a concatenation of a potentially very large number of audio slots, ordered according to ascending time. Streams are packetized when transported over USB whereby SIPs can only contain an integer number of audio slots. Each packet always starts with the same channel, and the channel order is respected throughout the entire transmission. If, for any reason, there are no audio slots available to construct a SIP, a Transfer Delimiter shall be sent instead.

2.3.1.6 TYPE I SUPPORTED FORMATS

The following paragraphs list all currently supported Type I Audio Data Formats. The bit allocations in the **bmFormats** field of the class-specific AS interface descriptor for the different Type I Audio Data Formats can be found in Appendix A.1, “Audio Data Formats Bit Allocations.” Note that an Alternate Setting of an AudioStreaming interface can only support one Type I Audio Data Format. Consequently, only one Type I Audio Data Format bit shall be set in the **bmFormats** field of the class-specific AS interface descriptor. It is allowed to support one or more Type III Audio Data Formats in the same AudioStreaming interface if the interface is able to independently distinguish between the Type I Audio Data Format and any other Type III Audio Data Format.

2.3.1.6.1 PCM FORMAT

The PCM (Pulse Coded Modulation) format is the most commonly used audio format to represent audio data streams. The audio data is not compressed and uses a signed two’s-complement fixed point format. It is left-justified (the sign bit is the Msb) and data is padded with trailing zeroes to fill the remaining unused bits of the subslot. The binary point is located to the right of the sign bit so that all values lie within the range [-1, +1).

2.3.1.6.2 PCM8 FORMAT

The PCM8 format is introduced to be compatible with the legacy 8-bit wave format. Audio data is uncompressed and uses 8 bits per sample (**bBitResolution** = 8). In this case, data is unsigned fixed-point, left-justified in the audio subslot, Msb first. The range is [0,255].

2.3.1.6.3 IEEE_FLOAT FORMAT

The IEEE_FLOAT format is based on the ANSI/IEEE-754 floating-point standard. Audio data is represented using the basic single-precision format. The basic single-precision number is 32 bits wide and has an 8-bit exponent and a 24-bit mantissa. Both mantissa and exponent are signed numbers, but neither is represented in two’s-complement format. The mantissa is stored in sign magnitude format and the exponent in biased form (also called excess-n form). In biased form, there is a positive integer (called the bias) which is subtracted from the stored number to get the actual number. For example, in an eight-bit exponent, the bias is 127. To represent 0, the number 127 is stored. To represent -100, 27 is stored. An exponent of all zeroes and an exponent of all ones are both reserved for special cases, so in an eight-bit field, exponents of -126 to +127 are possible. In the basic floating-point format, the mantissa is assumed to be normalized so that the most significant bit is always one, and therefore is not stored. Only the fractional part is stored. Denormalized (exponent = 0) values are considered to be zero.

The 32-bit IEEE-754 floating-point word is broken into three fields. The most significant bit stores the sign of the mantissa, the next group of 8 bits stores the exponent in biased form, and the remaining 23 bits store the magnitude of the fractional portion of the mantissa. For further information, refer to the ANSI/IEEE-754 standard.

The data is conveyed over USB using 32 bits per sample (**bBitResolution** = 32; **bSubslotSize** = 4).

2.3.1.6.4 ALAW FORMAT AND μ LAW FORMAT

Starting from 12- or 16-bits linear PCM samples, simple compression down to 8-bits per sample (one byte per sample) can be achieved by using logarithmic companding. The compressed audio data uses 8 bits per sample (**bBitsPerSample** = 8). Data is signed fixed point, left-justified in the subslot, Msb first. The compressed range is [-128,128]. The difference between Alaw and μ Law compression lies in the formulae used to achieve the compression. Refer to the ITU G.711 standard for further details.

2.3.1.6.5 DSD FORMAT

The Direct-Stream Digital (DSD) format uses pulse-density modulation encoding—a technology primarily used to store audio signals on SACD (Super Audio CD) digital storage media.

Audio data is stored as single-bit delta-sigma modulated digital audio; i.e. a sequence of single-bit values at a sampling rate of 2.8224 MHz (64 times the CD audio sampling rate of 44.1 kHz) for basic sampling rate DSD64, 5.6448 MHz for DSD128 (2X-rate DSD), 11.2896 MHz for DSD256 (4X-rate DSD), 22.5792 MHz for DSD 512 (8X-rate DSD), and 45.1584 MHz for DSD1024 (16X-rate DSD). 48 kHz-based DSD streams are also in existence. In that case, the bitstream sampling rates are 3.072 MHz, 6.144 MHz, 12.288 MHz, 24.576 MHz, and 49.152 MHz respectively.

No matter what sampling rate the DSD stream uses, the audio subslot size is fixed to 64 bits (**bBitResolution** = 64; **bSubslotSize** = 8) so that, at the transport layer, the DSD stream always looks like 64-bit PCM data. Therefore, the transport sampling rate of the DSD stream, packetized as 64-bit PCM samples, is always 1/64 of the DSD sampling rate and the Clock Source Entity, connected to the Terminal, representing the DSD AudioStreaming interface shall always advertise this transport sampling rate:

- 44.1 kHz, 88.2 kHz, 176.4 kHz, 352.8 kHz, or 705.6 kHz for 44.1 kHz based DSD streams
- 48 kHz, 96 kHz, 192 kHz, 384 kHz, or 768 kHz for 48 kHz based DSD streams

2.3.1.6.6 TYPE I RAW DATA

This audio format is included to allow transport of data (audio or other) over a USB AudioStreaming interface in the form of PCM-like audio slots when the actual format or even the meaning of the transported data is unknown. The USB pipe simply acts as a pass-through. As a consequence, such data can never be interpreted inside the Audio Function and can only be routed from an Input Terminal to one or more Output Terminals. From a USB standpoint, the data is packed as if it were Type I formatted audio data, but the data is never to be interpreted as being audio data.

2.3.2 TYPE III FORMATS

These formats are based upon the IEC61937 standard. The IEC61937 standard describes a method to transfer non-PCM encoded audio bit streams over an IEC60958 digital audio interface, together with the transfer of the accompanying “Channel Status” and “User Data.”

The IEC60958 standard specifies a widely used method of interconnecting digital audio equipment with two-channel linear PCM audio. The IEC61937 standard describes a way in which the IEC60958 interface shall be used to convey non-PCM encoded audio bit streams for consumer applications.

The same basic techniques used in IEC61937 are reused here to convey non-PCM encoded audio bit streams over a Type III formatted audio stream. From a USB transfer standpoint, the data streaming over the interface looks exactly like two-channel 16 bit PCM audio data.

2.3.2.1 TYPE III SUPPORTED FORMATS

Refer to the ISO/IEC 60958 and ISO/IEC 61937 (several parts) specifications for detailed format information. The bit allocations in the **bmFormats** field of the class-specific AS interface descriptor for the different Type III Audio Data Formats can be found in Appendix A.1, “Audio Data Formats Bit Allocations.”

Note that an Alternate Setting of an AudioStreaming interface can support multiple Type III Audio Data Formats. Consequently, multiple bits can be simultaneously set in the **bmFormats** field of the class-specific AS interface descriptor. It is then up to the decoder to inspect the encoded data stream to determine exactly which format is currently being used over the pipe. Alternatively, a device may expose multiple Alternate Settings on an AudioStreaming interface, each advertising support for only a single Type III Audio Data Format. The Host can then select an appropriate Alternate Setting that corresponds to the format currently in use.

The following is a list of formats that is covered by the above specifications.

- PCM_IEC60958
- AC-3
- MPEG-1_Layer1

- MPEG-1_Layer2/3 or MPEG-2_NOEXT
- MPEG-2_EXT
- MPEG-2_AAC_ADTS
- MPEG-2_Layer1_LS
- MPEG-2_Layer2/3_LS
- DTS-I
- DTS-II
- DTS-III
- ATRAC
- ATRAC2/3
- WMA
- E-AC-3
- MAT
- DTS-IV
- MPEG-4_HE_AAC
- MPEG-4_HE_AAC_V2
- MPEG-4_AAC_LC
- DRA
- MPEG-4_HE_AAC_SURROUND
- MPEG-4_AAC_LC_SURROUND
- MPEG-H_3D_AUDIO
- AC4
- MPEG-4_AAC_ELD

2.3.3 TYPE IV FORMATS

Type IV formats can only be used on external connections to the Audio Function that do not use a USB pipe for their data transport but that do need an AudioStreaming interface to advertise their presence. A typical example of such a connection is an S/PDIF connector that is capable of handling both PCM stereo audio data streams (IEC60958) in one Alternate Setting and encoded data streams (IEC61937) in another Alternate Setting of the interface. Note however that the external connection could also be vendor specific (like a parallel data interface).

2.3.3.1 TYPE IV SUPPORTED FORMATS

This specification supports all Audio Data Formats on an external connection that are defined on a USB pipe (Type I and Type III). See Section 2.3.1.6, “Type I Supported Formats”, and Section 2.3.2.1, “Type III Supported Formats”.

The bit allocations in the **bmFormats** field of the class-specific AS interface descriptor for the different Type IV Audio Data Formats can be found in Appendix A.1, “Audio Data Formats Bit Allocations.”

Note that an Alternate Setting of an AudioStreaming interface can support multiple Type IV Audio Data Formats. Consequently, multiple bits can be simultaneously set in the **bmFormats** field of the class-specific AS interface descriptor. It is then up to the decoder to inspect the encoded data stream to determine exactly what format is currently being used over the connection.

2.4 EXTENDED AUDIO DATA FORMATS

In addition to the Simple Audio Data Formats described above, Extended Audio Data Formats Type I and III are defined. These are based on the Simple Audio Data Formats Type I and III definitions but they provide an optional Header and for the Extended Audio Data Format Type I, an optional synchronous (i.e. sample accurate) Control Stream.

Each SIP shall start with a SIPDescriptor, followed by the following optional components:

- Header
- Audio data, formatted according the Simple Audio Data Formats Type I or III
- Synchronous vendor-specific Control Stream (only for Extended Audio Format Type I)

These three components can be optionally present on a per-SIP basis.

The SIPDescriptor is exactly 4 bytes long and has the following layout:

The **bmFlags** field indicates which of the components are present in the SIP as follows:

- Bit D0 indicates whether a Header is present in the SIP (D0=0b1) or not (D0=0b0). When the Header is not present, the **wHeaderLength** field shall be set to zero (0x0000).
- Bit D1 indicates whether an AudioSlot is present in all of the Extended AudioSlots of the SIP (D1=0b1) or not (D1=0b0).
- Bit D2 indicates whether a Control Stream is present. In other words, it indicates whether a Control Word is present in all of the Extended AudioSlots of the Extended Type I SIP (D2=0b1) or not (D2=0b0). This bit shall be set to zero (D2=0b0) for Extended Type III formats.
- Bits D15..3 are reserved.

The **wHeaderLength** field indicates the total length of the Header in bytes. This includes all the SubHeaders that together make up the total Header.

The presence of the SIPDescriptor allows for maximum flexibility. For example, it is possible to create an Audio Stream consisting of a Control Stream only, without Header or audio data. It is also possible to create an Audio Stream where Headers are only present once in a while.

If present, the Header immediately follows the SIPDescriptor. There is only one Header allowed per SIP. The Header can consist of one or more SubHeaders, each having a different purpose. Each SubHeader shall start with a **wSubHeaderID** field, uniquely identifying the purpose of the SubHeader. The length of each SubHeader and the semantics of the remaining fields in the SubHeader are defined by this specification. The structure and composition of the Header and the order in which the SubHeaders appear in the Header can change from SIP to SIP.

Table 2-2: SIPDescriptor Layout

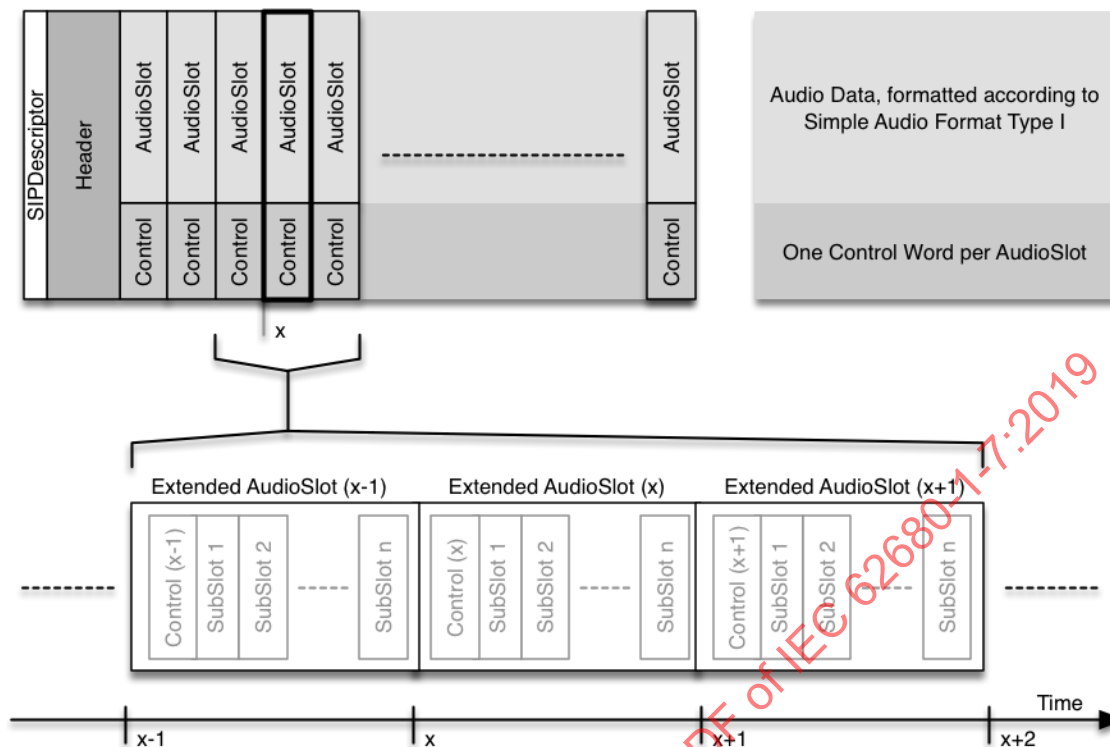
Offset	Field	Size	Value	Description
0	bmFlags	2	Bitmap	D0: Header present. D1: AudioSlot present. D2: Control Stream present. D15..3: Reserved.
2	wHeaderLength	2	Number	Total length of the Header, in bytes.

2.4.1 EXTENDED TYPE I FORMATS

The following figure illustrates the possible layout of a SIP. The greyed parts are optional.

Note: In order to have a meaningful stream, at least one of the optional components shall be present.

Figure 2-2: Extended Type I Format



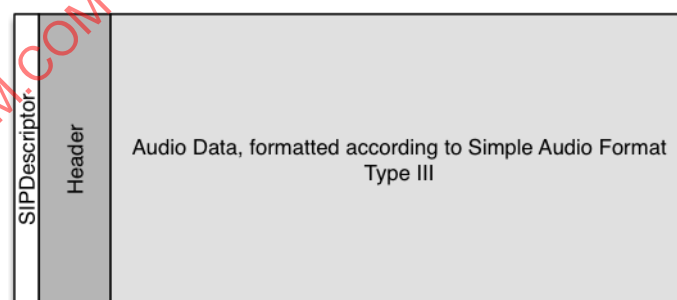
An Extended AudioSlot is the concatenation of a Control Word, followed by the Type I AudioSlot. The Control Stream therefore consists of a sequence of Control Words, where each Control Word is synchronous to its associated AudioSlot. There are as many Control Words per SIP as there are AudioSlots in the SIP. The byte size of the Control Words is independent of the AudioSubSlot size and is the same for each AudioSlot.

2.4.2 EXTENDED TYPE III FORMATS

The following figure illustrates the possible layout of a SIP. The greyed parts are optional.

Note: In order to have a meaningful stream, at least one of the optional components should be present.

Figure 2-3: Extended Type III Format



The Header is optionally followed by the actual encoded audio frame data.

2.5 CLASS-SPECIFIC AS INTERFACE DESCRIPTOR

The first eight fields of this descriptor are described in the USB Audio Device Class Definition document.

The **bSize** field indicates the length in bytes of the **bmFormats** field. The **bmFormats** field consists of an array of bits where each bit indicates a particular Audio Data Format is supported ($D_x = 1$) or not ($D_x = 0$).

It is allowed to support both Type I and Type III formats in the same Alternate Setting of an AudioStreaming interface as long as the interface is able to unambiguously determine which exact format is currently in use by inspecting the data in the audio stream. Type IV formats can never be mixed with either Type I or Type III

formats in the same AudioStreaming interface. Type IV formats are distinguished from Type I/Type III formats by the absence of a USB endpoint in the AudioStreaming interface.

A complete list of currently supported Audio Data Formats, their bit allocation in the **bmFormats** field and their usage can be found in Appendix A.1, “Audio Data Formats Bit Allocations.”

For Simple and Extended Type I formats, the **bSubslotSize** field can be set to 1, 2, 4, or 8, and the **bBitResolution** field indicates how many bits of the total number of available bits in the audio subslot are effectively used to convey audio information.

For Simple and Extended Type III formats, the **bSubslotSize** field shall be set to 2 and the **bBitResolution** field shall be set to 16.

For Type IV formats, the **bSubslotSize** and **bBitResolution** fields are not used and shall be set to zero.

For Simple Type I and Type III formats, the **bmAuxProtocols** and **bControlSize** fields are not used and shall be set to zero.

For Extended Type I and Extended Type III formats, this specification defines a number of Auxiliary Protocols (See Section 3, “Auxiliary Protocols”). The **bmAuxProtocols** field contains a bitmap identifying which Auxiliary Protocols this AudioStreaming interface’s Alternate Setting requires. For each Auxiliary Protocol, the AudioStreaming interface may offer up to two Alternate Settings, one in which the Auxiliary Protocol is required and the other in which it is not. For example, an AudioStreaming Interface may offer two Alternate Settings, one indicating the required use of HDCP and the other indicating that it operates without HDCP.

If the stream does not contain actual audio data, the **bNrChannels**, **bmChannelConfig** and **iChannelNames** in the class-specific AS Interface descriptor (See the *USB Audio Device Class* document) shall all be set to zero.

For Extended Type I formats, the **bControlSize** field indicates the size in bytes of each Control Channel Word in the stream. It shall be set to zero for all other formats.

Table 2-3: Class-Specific AS Interface Descriptor

Offset	Field	Size	Value	Description
0	bLength	1	Number	Size of this descriptor in bytes: 23.
1	bDescriptorType	1	Constant	CS_INTERFACE descriptor type.
2	bDescriptorSubtype	1	Constant	AS_GENERAL descriptor subtype.
3	bTerminalLink	1	Number	The Terminal ID of the Terminal to which this interface is connected.
4	bmControls	4	Bitmap	D1..0: Active Alternate Setting Control. D3..2: Valid Alternate Settings Control. D5..4: Audio Data Format Control. D31..6: Reserved.
8	wClusterDescrID	2	Number	ID of the cluster descriptor of the AS Interface.
10	bmFormats	8	Bitmap	The Audio Data Format(s) that can be used to communicate with this interface.
18	bSubslotSize	1	Number	The number of bytes occupied by one audio subslot.
19	bBitResolution	1	Number	The number of effectively used bits from the available bits in an audio subslot.
20	bmAuxProtocols	2	Bitmap	Bitmap, indicating which Auxiliary Protocols are required.
22	bControlSize	1	Number	Size of the Control Channel Words, in bytes.

3 AUXILIARY PROTOCOLS

3.1 HDCP PROTOCOL

The HDCP protocol uses the HDCP SubHeader to convey the periodic Link Synchronization information, required by [HDCP2.2]. [HDCP2.2] requires that a 32-bit *streamCtr* value and a 64-bit *inputCtr* value be sent periodically from the content transmitter to the content receiver. The Header construct in the Extended Format Types shall be used to convey that information periodically as follows.

The HDCP SubHeader has the following layout:

Table 3-1: HDCP SubHeader Layout

Offset	Field	Size	Value	Description
0	bLength	1	Number	Size of this SubHeader, in bytes. 16.
1	bSubHeaderID	1	Number	HDCP_ENCRYPTION
2	bOffset	1	Number	Offset to first full 16-byte encrypted data block in the SIP.
3	bReserved	1	Number	Reserved for future extensions.
4	dStreamCtr	4	Number	The <i>streamCtr</i> value as assigned by the HDCP transmitter.
8	qInputCtr	8	Number	The <i>inputCtr</i> value associated with the first full 16-byte encrypted data block in the SIP.

The **bLength** field contains the size of this SubHeader, expressed in bytes.

The **bSubHeaderID** field contains the HDCP_ENCRYPTION constant to uniquely identify this SubHeader as an HDCP SubHeader.

The **bOffset** field contains an offset value ranging from 0 to 15, indicating at which byte position, measured from the start of the audio data in the SIP, the first full 16-byte encrypted data block (HDCP Cipherblock) starts. The value in the **qInputCtr** field shall pertain to that same 16-byte encrypted data block.

Note: Since an HDCP Cipher Block is always 16 bytes or 128 bits long, the start of a full Cipher Block is guaranteed to occur within the first 16 bytes of the actual audio data payload. Also, when a Control Stream is present, the **bOffset** field value only pertains to the actual audio data bytes (the Control Stream is not encrypted) and the Control Stream should be separated from the actual audio data bytes before the offset is applied.

The **bReserved** field is reserved and shall be set to zero.

The **dStreamCtr** field contains the *streamCtr* value associated with this particular stream as assigned by the HDCP transmitter.

The **qInputCtr** field contains the *inputCtr* value associated with the first full 16-byte encrypted audio data block in the SIP.

The HDCP SubHeader shall be present at least every HDCP_PACKET_HEADER_TIME but may occur more frequently. The presence of the HDCP SubHeader indicates to the HDCP receiver that the content is HDCP-encrypted.

4

ADDING NEW AUDIO DATA FORMATS

Adding new Audio Data Formats to this specification is achieved by proposing a fully documented Audio Data Format to the Audio Device Class Working Group. Upon acceptance, they will register the new Audio Data Format (attribute a unique bit position in the **bmFormats** field of the class-specific AS interface descriptor) and update this document accordingly. This process will also guarantee that new releases of generic USB audio drivers will support the newly registered Audio Data Formats.

It is always possible to use vendor-specific definitions if the above procedure is considered unsatisfactory.

IECNORM.COM : Click to view the full PDF of IEC 62680-1-7:2019

5 ADDING NEW SIDE BAND PROTOCOLS

Adding new Side Band Protocols to this specification is achieved by proposing a fully documented Side Band Protocol to the Audio Device Class Working Group. Upon acceptance, they will register the new Side Band Protocol (attribute a unique SideBandProtocol constant) and update this document accordingly. This process will also guarantee that new releases of generic USB audio drivers will support the newly registered Side Band Protocols.

It is always possible to use vendor-specific definitions if the above procedure is considered unsatisfactory.

IECNORM.COM : Click to view the full PDF of IEC 62680-1-7:2019

Appendix A. Additional Audio Device Class Codes

A.1 AUDIO DATA FORMATS BIT ALLOCATIONS

Table A-1: Audio Data Formats Bit Allocations in the bmFormats Field and Usage

Name	bmFormats	Usage		
		Type I	Type III	Type IV
PCM	D0	X		X
PCM8	D1	X		X
IEEE_FLOAT	D2	X		X
ALAW	D3	X		X
MULAW	D4	X		X
DSD	D5	X		X
RAW_DATA	D6	X		X
PCM_IEC60958	D7		X	X
AC-3	D8		X	X
MPEG-1_Layer1	D9		X	X
MPEG-1_Layer2/3 or MPEG-2_NOEXT	D10		X	X
MPEG-2_EXT	D11		X	X
MPEG-2_AAC_ADTS	D12		X	X
MPEG-2_Layer1_LS	D13		X	X
MPEG-2_Layer2/3_LS	D14		X	X
DTS-I	D15		X	X
DTS-II	D16		X	X
DTS-III	D17		X	X
ATRAC	D18		X	X
ATRAC2/3	D19		X	X
WMA	D20		X	X
E-AC-3	D21		X	X
MAT	D22		X	X
DTS-IV	D23		X	X
MPEG-4_HE_AAC	D24		X	X
MPEG-4_HE_AAC_V2	D25		X	X
MPEG-4_AAC_LC	D26		X	X
DRA	D27		X	X
MPEG-4_HE_AAC_SURROUND	D28		X	X
MPEG-4_AAC_LC_SURROUND	D29		X	X
MPEG-H_3D_AUDIO	D30		X	X
AC4	D31		X	X
MPEG-4_AAC_ELD	D32		X	X

A.2 SUBHEADER CODES

Table A-2: SubHeader Codes

SubHeader Code	Value
HEADER_UNDEFINED	0x00
HDCP_ENCRYPTION	0x01
Reserved	0x02..0xFF

A.3 AUDIO FORMAT GENERAL CONSTANTS

Table A-3: General Constants

Constant Identifier	Value
HDCP_PACKET_HEADER_TIME	512 (ms)

IECNORM.COM : Click to view the full PDF of IEC 62680-1-7:2019

COMMISSION ÉLECTROTECHNIQUE INTERNATIONALE

INTERFACES DE BUS UNIVERSEL EN SÉRIE POUR LES DONNÉES ET L'ALIMENTATION ÉLECTRIQUE –

Partie 1-7: Composants communs – Définition de classes de dispositifs USB Audio 3.0 pour formats de données

AVANT-PROPOS

- 1) La Commission Electrotechnique Internationale (IEC) est une organisation mondiale de normalisation composée de l'ensemble des comités électrotechniques nationaux (Comités nationaux de l'IEC). L'IEC a pour objet de favoriser la coopération internationale pour toutes les questions de normalisation dans les domaines de l'électricité et de l'électronique. A cet effet, l'IEC – entre autres activités – publie des Normes internationales, des Spécifications techniques, des Rapports techniques, des Spécifications accessibles au public (PAS) et des Guides (ci-après dénommés "Publication(s) de l'IEC"). Leur élaboration est confiée à des comités d'études, aux travaux desquels tout Comité national intéressé par le sujet traité peut participer. Les organisations internationales, gouvernementales et non gouvernementales, en liaison avec l'IEC, participent également aux travaux. L'IEC collabore étroitement avec l'Organisation Internationale de Normalisation (ISO), selon des conditions fixées par accord entre les deux organisations.
- 2) Les décisions ou accords officiels de l'IEC concernant les questions techniques représentent, dans la mesure du possible, un accord international sur les sujets étudiés, étant donné que les Comités nationaux de l'IEC intéressés sont représentés dans chaque comité d'études.
- 3) Les Publications de l'IEC se présentent sous la forme de recommandations internationales et sont agréées comme telles par les Comités nationaux de l'IEC. Tous les efforts raisonnables sont entrepris afin que l'IEC s'assure de l'exactitude du contenu technique de ses publications; l'IEC ne peut pas être tenue responsable de l'éventuelle mauvaise utilisation ou interprétation qui en est faite par un quelconque utilisateur final.
- 4) Dans le but d'encourager l'uniformité internationale, les Comités nationaux de l'IEC s'engagent, dans toute la mesure possible, à appliquer de façon transparente les Publications de l'IEC dans leurs publications nationales et régionales. Toutes divergences entre toutes Publications de l'IEC et toutes publications nationales ou régionales correspondantes doivent être indiquées en termes clairs dans ces dernières.
- 5) L'IEC elle-même ne fournit aucune attestation de conformité. Des organismes de certification indépendants fournissent des services d'évaluation de conformité et, dans certains secteurs, accèdent aux marques de conformité de l'IEC. L'IEC n'est responsable d'aucun des services effectués par les organismes de certification indépendants.
- 6) Tous les utilisateurs doivent s'assurer qu'ils sont en possession de la dernière édition de cette publication.
- 7) Aucune responsabilité ne doit être imputée à l'IEC, à ses administrateurs, employés, auxiliaires ou mandataires, y compris ses experts particuliers et les membres de ses comités d'études et des Comités nationaux de l'IEC, pour tout préjudice causé en cas de dommages corporels et matériels, ou de tout autre dommage de quelque nature que ce soit, directe ou indirecte, ou pour supporter les coûts (y compris les frais de justice) et les dépenses découlant de la publication ou de l'utilisation de cette Publication de l'IEC ou de toute autre Publication de l'IEC, ou au crédit qui lui est accordé.
- 8) L'attention est attirée sur les références normatives citées dans cette publication. L'utilisation de publications référencées est obligatoire pour une application correcte de la présente publication.
- 9) L'attention est attirée sur le fait que certains des éléments de la présente Publication de l'IEC peuvent faire l'objet de droits de brevet. L'IEC ne saurait être tenue pour responsable de ne pas avoir identifié de tels droits de brevets et de ne pas avoir signalé leur existence.

La Norme internationale IEC 62680-1-7 a été établie par le domaine technique 18: Systèmes multimédias domestiques et applications pour réseaux d'utilisateurs finaux, du comité d'études 100 de l'IEC: Systèmes et équipements audio, vidéo et services de données.

Le texte de cette norme a été élaboré par l'USB Implementers Forum (USB-IF). Les règles structurelles et rédactionnelles utilisées dans la présente publication reflètent les pratiques en vigueur au sein de l'organisme responsable de sa soumission.

Le texte de cette Norme internationale est issu des documents suivants:

CDV	Rapport de vote
100/3159/CDV	100/3229/RVC

Le rapport de vote indiqué dans le tableau ci-dessus donne toute information sur le vote ayant abouti à l'approbation de cette Norme internationale.

Le comité a décidé que le contenu de ce document ne sera pas modifié avant la date de stabilité indiquée sur le site web de l'IEC sous "<http://webstore.iec.ch>" dans les données relatives au document recherché. A cette date, le document sera

- reconduit,
- supprimé,
- remplacé par une édition révisée, ou
- amendé.

IMPORTANT – Le logo "colour inside" qui se trouve sur la page de couverture de cette publication indique qu'elle contient des couleurs qui sont considérées comme utiles à une bonne compréhension de son contenu. Les utilisateurs devraient, par conséquent, imprimer cette publication en utilisant une imprimante couleur.

INTRODUCTION

La série IEC 62680 est issue d'une série de spécifications initialement établies par l'USB Implementers Forum (USB-IF). Ces spécifications ont été soumises à l'IEC dans le cadre d'un accord particulier conclu entre l'IEC et l'USB-IF.

La présente norme est la publication relative à la définition de classes de dispositifs USB pour formats de données, version 3.0, publiée par l'USB-IF.

L'USB Implementers Forum, Inc. (USB-IF) est un organisme à but non lucratif fondé par le groupe de sociétés qui a développé la spécification du bus universel en série. L'USB-IF a été créé dans le but de proposer un organisme et un forum à même de favoriser la progression et l'adoption de la technologie USB. Le forum facilite le développement de périphériques (dispositifs) USB compatibles et de haute qualité et promeut les avantages de la technologie USB et la qualité des produits qui ont été validés par des essais de conformité.

TOUTES LES SPÉCIFICATIONS USB VOUS SONT FOURNIES "EN L'ÉTAT", SANS GARANTIE D'AUCUNE SORTE, Y COMPRIS TOUTE GARANTIE DE QUALITÉ MARCHANDE, DE NON-VIOLATION OU D'ADÉQUATION À UN USAGE PARTICULIER. L'USB IMPLEMENTERS FORUM ET LES AUTEURS DE L'ENSEMBLE DES SPÉCIFICATIONS USB DÉCLINENT TOUTE RESPONSABILITÉ, Y COMPRIS TOUTE RESPONSABILITÉ RELATIVE À LA VIOLATION DE DROITS DE PROPRIÉTÉ, EN CE QUI CONCERNE L'UTILISATION OU LA MISE EN ŒUVRE DES INFORMATIONS CONTENUES DANS LA PRÉSENTE SPÉCIFICATION.

LA MISE À DISPOSITION D'UNE SPÉCIFICATION USB, QUELLE QU'ELLE SOIT, N'IMPLIQUE L'OCTROI D'AUCUNE LICENCE, EXPRESSE OU IMPLICITE, PAR PERCLUSION OU AUTRE, SUR AUCUN DROIT DE PROPRIÉTÉ INTELLECTUELLE.

La conclusion des accords des adoptants de l'USB peut toutefois permettre à une société signataire de participer à un accord de licence réciproque RAND-Z pour les produits conformes. Pour plus d'informations, voir:

<https://www.usb.org/documents>

L'IEC NE PREND AUCUNE POSITION QUANT À SAVOIR S'IL VOUS EST CONSEILLÉ DE CONCLURE UN QUELCONQUE ACCORD DES ADOPTANTS DE L'USB OU DE PARTICIPER À L'USB IMPLEMENTERS FORUM.

DÉFINITION DE CLASSES DE DISPOSITIFS DE BUS UNIVERSEL EN SÉRIE POUR FORMATS DE DONNÉES

IECNORM.COM : Click to view the full PDF of IEC 62680-1-7:2019

Version 3.0
22 septembre 2016

DOMAINE D'APPLICATION DE LA PRÉSENTE VERSION

Le présent document est la Version 3.0 de la présente définition de classes de dispositifs.

COLLABORATEURS

Joe Scanlon	Advanced Micro Devices
Rhoads Hollowell	Apple Inc.
Girault Jones	Apple Inc.
Matthew X. Mora	Apple Inc.
Tzung-Dar Tsai	C-Media Electronics, Inc.
Brad Lambert	Cirrus Logic, Inc.
Dan Bogard	Conexant Systems, Inc.
Pete Burgers	DisplayLink (UK), Ltd.
David Roh	Dolby Laboratories, Inc.
Leng Ooi	Google, Inc.
Pierre-Louis Bossart	Intel Corporation
David Hines	Intel Corporation
Abdul Rahman Ismail (Co-président)	Intel Corporation
Devon Worrell	Intel Corporation
Chandrashekhar Rao	Logitech, Inc.
Terry Moore	MCCI Corporation
Alex Lin	MediaTek, Inc.
Bala Sivakumar	Microsoft Corporation
Geert Knapen (Co-président et éditeur)	NXP Semiconductors
	PL Mobile Audio
	411 E. Plumeria drive
	San Jose, CA 95134, USA
	E-mail: geert.knapen@nxp.com
James Goel	Qualcomm, Inc.
Andre Schevciw	Qualcomm, Inc.
Jin-Sheng Wang	Qualcomm, Inc.
Morten Christiansen	Synopsys

HISTORIQUE DES REVISIONS

Révision	Date	Nom de fichier	Description
1.0	18 mars 1998	Frmts10.pdf	Version 1.0
2.0	31 mai 2006	Frmts20 final.pdf	Version 2.0
3.0	22 septembre 2016	Frmts30.pdf	Version 3.0

Copyright © 1997-2016 USB Implementers Forum, Inc.
Tous droits réservés.

DENI DE RESPONSABILITÉ CONCERNANT LA PROPRIÉTÉ INTELLECTUELLE

LE PRÉSENT DOCUMENT ACCORDE UNE LICENCE POUR LA REPRODUCTION DE LA PRÉSENTE SPÉCIFICATION POUR UN USAGE INTERNE UNIQUEMENT. LE PRÉSENT DOCUMENT N'ACCORDE NI NE VISE À ACCORDER AUCUNE AUTRE LICENCE, EXPRESSE OU IMPLICITE, PAR PERCLUSION OU AUTRE.

L'USB-IF ET LES AUTEURS DE LA PRÉSENTE SPÉCIFICATION DÉCLINENT EXPRESSÉMENT TOUTE RESPONSABILITÉ RELATIVE À LA VIOLATION DE DROITS DE PROPRIÉTÉ INTELLECTUELLE EN CE QUI CONCERNE LA MISE EN ŒUVRE DES INFORMATIONS CONTENUES DANS LA PRÉSENTE SPÉCIFICATION. EN OUTRE, L'USB-IF ET LES AUTEURS DE LA PRÉSENTE SPÉCIFICATION NE GARANTISSENT NI NE DÉCLARENT QUE LA OU LESDITES MISES EN ŒUVRE NE VIOLERONT PAS LES DROITS DE PROPRIÉTÉ INTELLECTUELLE DE TIERS.

LA PRÉSENTE SPÉCIFICATION EST FOURNIE "EN L'ÉTAT", SANS GARANTIE, EXPRESSE OU IMPLICITE, LÉGALE OU AUTRE. TOUTES LES GARANTIES SONT EXPRESSÉMENT EXCLUES. USB-IF, LES MEMBRES DE L'USB-IF ET LES AUTEURS DE LA PRÉSENTE SPÉCIFICATION N'OFFRENT AUCUNE GARANTIE DE QUALITÉ MARCHANDE, AUCUNE GARANTIE DE NON-VIOLATION, AUCUNE GARANTIE D'ADAPTATION À UN USAGE PARTICULIER NI AUCUNE GARANTIE ÉMANANT D'UNE PROPOSITION, D'UNE SPÉCIFICATION OU D'UN ÉCHANTILLON QUELCONQUES.

L'USB-IF, LES MEMBRES DE L'USB-IF ET LES AUTEURS NE POURRONT EN AUCUN CAS ÊTRE REDEVABLES À UN TIERS DU COÛT D'ACQUISITION DE BIENS OU DE SERVICES DE REMPLACEMENT, D'UN MANQUE À GAGNER, D'UNE PRIVATION DE JOUISSANCE, D'UNE PERTE DE DONNÉES OU DE TOUT DOMMAGE ACCESSOIRE, CONSÉCUTIF, INDIRECT OU PARTICULIER, EN VERTU D'UN CONTRAT, D'UN DÉLIT, D'UNE GARANTIE OU AUTRE, ÉMANANT DE QUELQUE MANIÈRE QUE CE SOIT DE L'UTILISATION DE LA PRÉSENTE SPÉCIFICATION, QUE LEDIT TIERS AIT OU NON ÉTÉ AVISÉ AU PRÉALABLE DE LA POSSIBILITÉ DE TELS DOMMAGES.

NOTE: PLUSIEURS MEMBRES DE L'USB-IF ONT PARTICIPÉ À L'ÉLABORATION DE LA PRÉSENTE SPÉCIFICATION. CERTAINS DE CES MEMBRES PEUVENT AVOIR REFUSÉ DE CONCLURE UN ACCORD SPÉCIFIQUE QUANT À L'OCTROI D'UNE LICENCE CONCERNANT LES DROITS DE PROPRIÉTÉ INTELLECTUELLE POUVANT FAIRE L'OBJET D'UNE VIOLATION LORS DE LA MISE EN ŒUVRE DE LA PRÉSENTE SPÉCIFICATION. LES PERSONNES METTANT EN ŒUVRE LA PRÉSENTE SPÉCIFICATION EN ASSUMENT LES RISQUES.

Dolby™, AC-3™, Pro Logic™ et Dolby Surround™ sont des marques déposées de Dolby Laboratories, Inc.

Tous les autres noms de produits sont des marques, des marques déposées ou des marques de service de leurs propriétaires respectifs.

Envoyer les commentaires par courrier électronique à audio-chair@usb.org

SOMMAIRE

Domaine d'application de la présente version	32
Collaborateurs.....	32
Historique des révisions	32
Sommaire	34
Tableaux	35
Figures	36
1 Introduction	37
1.1 Documents connexes	37
1.2 Termes et abréviations.....	37
2 Formats de données audio	39
2.1 Délimiteur de transfert	40
2.2 Définitions d'intervalle de service et de paquet d'intervalle de service	40
2.3 Les formats de données audio simples	41
2.3.1 Formats de Type I.....	41
2.3.2 Formats de Type III.....	45
2.3.3 Formats de Type IV.....	46
2.4 Les formats de données audio étendus	47
2.4.1 Formats de Type I étendus.....	48
2.4.2 Formats de Type III étendus.....	49
2.5 Descripteur d'interface AS spécifique à la classe	49
3 Protocoles auxiliaires	51
3.1 Protocole HDCP	51
4 Ajout de nouveaux formats de données audio.....	52
5 Ajout de nouveaux protocoles à bande latérale	53
Annexe A. Codes de classes de dispositifs audio supplémentaires	54
A.1 Allocations de bits des formats de données audio	54
A.2 Codes SubHeader	55
A.3 Constantes générales du format audio	55

TABLEAUX

Tableau 2-1: Mise en paquets.....	43
Tableau 2-2: Topologie du SIPDescriptor	48
Tableau 2-3: Descripteur d'interface AS spécifique à la classe	50
Tableau 3-1: Topologie du SubHeader HDCP	51
Tableau A-1: Allocations de bits des formats de données audio dans le champ bmFormats et utilisation	54
Tableau A-2: Codes SubHeader	55
Tableau A-3: Constantes générales.....	55

IECNORM.COM : Click to view the full PDF of IEC 62680-1-7:2019

FIGURES

Figure 2-1: Flux audio de Type I	39
Figure 2-2: Format de Type I étendu.....	48
Figure 2-3: Format de Type III étendus	49

IECNORM.COM : Click to view the full PDF of IEC 62680-1-7:2019

1 INTRODUCTION

Le présent document a pour objectif de décrire en détail tous les formats de données audio pris en charge par la classe des dispositifs audio. Le présent document est vu comme faisant partie intégrante de la *Spécification de classe de dispositifs audio*, bien que les révisions ultérieures du présent document soient indépendantes de l'évolution des révisions de la *Spécification USB audio* principale. L'objectif est de faciliter la prise en compte de l'ajout de nouveaux formats de données audio sans affecter la *Spécification USB audio* de base.

1.1 DOCUMENTS CONNEXES

- *Spécification du bus universel en série*, révision 2.0 (appelée *Spécification USB* dans le présent document). Voir en particulier le Chapitre 5, "Modèle de flux de données USB" et le Chapitre 9, "Cadre de dispositif USB".
- Universal Serial Bus Device Class Definition for Audio Devices (disponible en anglais seulement) (appelée Classes de dispositifs audio USB dans le présent document).
- Universal Serial Bus Device Class Definition for Terminal Types (disponible en anglais seulement) (appelée Types de terminaux audio USB dans le présent document).
- Norme ANSI S1.11-1986.
- Norme MPEG-1 ISO/IEC 11172-3 1993 (disponible sur <http://www.iso.ch>).
- Norme MPEG-2 ISO/IEC 13818-3, 20 février 1997 (disponible sur <http://www.iso.ch>).
- Norme de compression audionumérique (AC-3), ATSC A/52A, 20 août 2001 (disponible sur <http://www.atsc.org>).
- Spécification WMA (Windows Media Audio) (disponible sur <http://www.microsoft.com>).
- ANSI/IEEE-754, norme de virgule flottante.
- Norme internationale ISO/IEC 60958: *Interface audionumérique et annexes*.
- Norme ISO/IEC 61937.
- Norme ITU G.711.
- Spécification ETSI TS 102 114, "DTS Coherent Acoustics; Core and Extensions" (disponible en anglais seulement) (disponible sur http://webapp.etsi.org/action%5CPU/20020827/ts_102114v010101p.pdf).

1.2 TERMES ET ABRÉVIATIONS

La présente section définit les termes utilisés dans le présent document. Pour les termes supplémentaires concernant le Bus universel en série, voir le Chapitre 2, "Termes et abréviations", de la *Spécification USB*.

AC-3	Norme de compression audio de Dolby Labs.
Emplacement audio	Ensemble de sous-emplacements audio, contenant chacun un échantillon audio MIC d'un canal audio physique différent, prélevés au même instant.
Diffusion audio	Concaténation d'un nombre potentiellement très important d'emplacements audio ordonnés de façon croissante en fonction du temps.
Sous-emplacements audio	Contient un échantillon audio MIC unique.
DTS	Acronyme de Digital Theater Systems.
DVD	Acronyme de Digital Versatile Disc (disque numérique polyvalent).

Diffusion de bits audio codés	Concaténation d'un nombre potentiellement très important de trames audio codées, ordonnées de façon croissante en fonction du temps.
Trame audio codée	Séquence de bits contenant une représentation codée d'échantillons audio provenant d'un ou de plusieurs canaux audio physiques, prélevés sur une période fixe.
MPEG	Acronyme de Moving Pictures Expert Group.
MIC	Acronyme de Modulation par Impulsion et Codage.
Intervalle de service	Regroupement de (micro)trames USB ou d'intervalles de bus liés.
Paquet d'intervalle de service	Paquet contenant tous les emplacements audio transférés sur le bus pendant un intervalle de service.
Délimiteur de transfert	Jeton unique indiquant une interruption dans un flux isochrone de paquets de données. Il peut s'agir d'un paquet de données de longueur nulle ou de l'absence de transfert isochrone dans une trame USB donnée.
WMA	Acronyme de Windows Media Audio.

2 FORMATS DE DONNÉES AUDIO

Les formats de données audio peuvent être divisés en deux groupes principaux:

- les formats de données audio simples;
- les formats de données audio étendus.

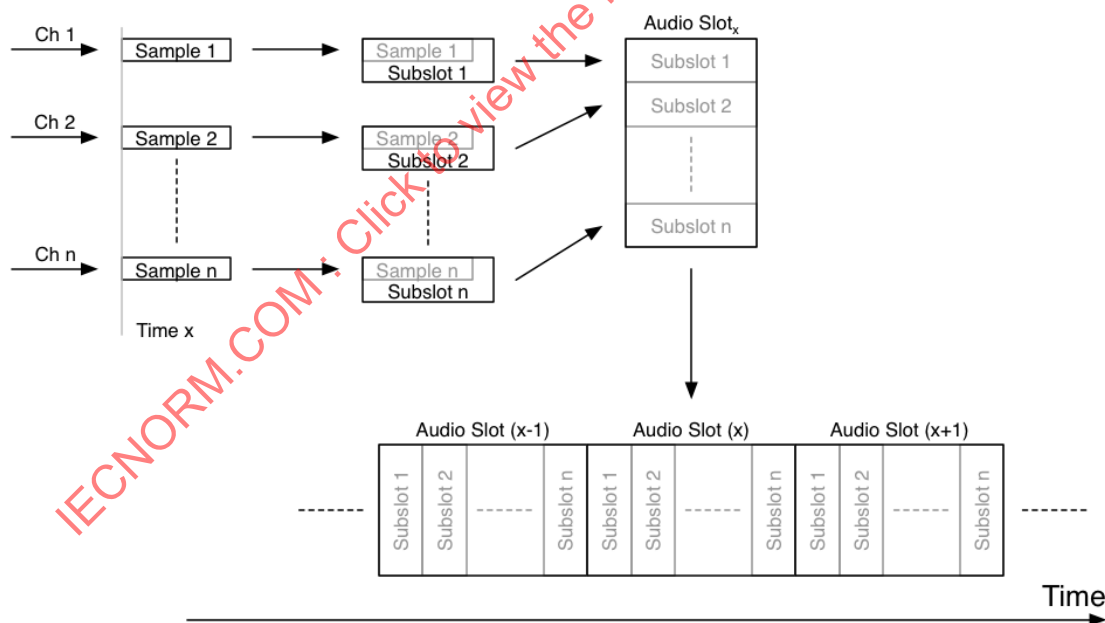
Les formats de données audio simples peuvent ensuite être divisés en trois groupes selon le type.

Le premier groupe (le Type I) traite des flux de données audio transmis par USB et construits échantillon par échantillon. Chaque échantillon audio est représenté par un symbole indépendant unique, contenu dans un sous-emplacement audio. Différents schémas de compression peuvent être utilisés pour transformer les échantillons audio en symboles.

Note: Ce processus est différent du codage. La compression est réputée s'effectuer sur une base par échantillon audio. Chaque échantillon audio génère un symbole (par exemple la compression loi A, où un échantillon audio de 16 bits est compressé en un symbole de 8 bits).

Si plusieurs canaux audio physiques sont formatés en une grappe de canaux audio unique, alors les échantillons à l'instant x des canaux suivants sont d'abord contenus dans des sous-emplacements audio. Ces sous-emplacements audio sont ensuite entrelacés, selon l'ordre des canaux de la grappe, comme décrit dans la *Spécification USB audio* principale, puis regroupés dans un emplacement audio. Les échantillons audio, prélevés à l'instant $x+1$, sont entrelacés de la même manière afin de générer l'emplacement audio suivant, et ainsi de suite. La notion de canaux physiques est explicitement préservée pendant la transmission. Les données audio MIC normalisées constituent un exemple type de format de Type I. La figure suivante représente ce concept.

Figure 2-1: Flux audio de Type I



Anglais	Français
Ch	Canal
Sample	Échantillon
Subslot	Sous-emplacement
Audio Slot	Emplacement audio
Time	Temps

Le deuxième groupe (le type III) contient des formats de données audio qui utilisent l'encapsulation, comme décrit dans la norme ISO/IEC 61937, avant d'être envoyés sur USB. Un ou plusieurs flux de données audio codés non MIC sont regroupés en "échantillons pseudo-stéréo" et transmis comme s'il s'agissait de vrais échantillons audio MIC stéréo. Soit la fréquence d'échantillonnage de ces pseudo-échantillons (fréquence d'échantillonnage de transport, indiquée par la commande de fréquence d'horloge de l'entité source d'horloge associée) correspond à la fréquence d'échantillonnage des flux de données audio MIC non codés d'origine (fréquence d'échantillonnage native), soit il existe une relation de rapport entier entre ces fréquences. Par conséquent, la récupération d'horloge à l'extrémité de réception est relativement aisée. L'inconvénient est qu'à moins que plusieurs flux codés non MIC ne soient regroupés en un flux pseudo-stéréo, davantage de bande passante que nécessaire est consommée.

Le troisième groupe (le type IV) traite des flux audio qui ne sont pas transmis par USB. Au lieu de cela, ils interagissent avec la fonction audio par l'intermédiaire d'une interface AudioStreaming qui ne contient pas de point d'extrémité isochrone USB IN ou OUT. Ces flux se connectent généralement au moyen d'une interface numérique comme S/PDIF (ou autre moyen de connexion) et peuvent exiger une interaction de l'hôte avant qu'ils n'entrent dans la fonction audio ou ne la quittent. Un connecteur S/PDIF externe pouvant accepter un flux audio codé AC-3 constitue un exemple type. Ce flux est d'abord traité par un décodeur AC-3 avant que les canaux audio logiques (décodés) n'entrent dans la fonction audio par le biais du terminal d'entrée qui représente cette connexion S/PDIF.

Outre les formats de données audio simples décrits ci-dessus, des formats de données audio étendus sont définis. Ils reposent sur les définitions des formats de données audio simples de Type I et de Type III, mais ils fournissent un en-tête de paquet facultatif et, pour le format de données audio étendu de Type I, un canal de contrôle synchrone (par exemple correspondant à l'échantillon) facultatif. Les formats de données audio de Type IV n'ont pas de définition de format de données audio étendu.

Les sections suivantes expliquent de manière plus détaillée les différents formats de données audio et types de formats.

2.1 DÉLIMITEUR DE TRANSFERT

Les flux de données isochrones sont de nature continue, bien que le nombre réel d'octets envoyés par paquet puisse varier tout au long de la durée de vie du flux (à des fins d'adaptation du débit, par exemple). Afin d'indiquer un arrêt temporaire du flux de données isochrone sans fermer le canal (et donc céder la bande passante USB), il est nécessaire de définir un délimiteur de transfert dans la bande. La présente spécification envisage deux situations représentant un délimiteur de transfert. La première correspond à un paquet de données de longueur nulle et la seconde à l'absence de transfert isochrone dans une (micro)trame USB ayant normalement un transfert isochrone. Les deux situations sont jugées équivalentes et il est prévu que la fonction audio se comporte de la même façon. Toutefois, le second type consomme moins de bande passante USB isochrone (c'est-à-dire une bande passante nulle). Dans les deux cas, la présente spécification estime qu'un délimiteur de transfert est une entité qui peut être envoyée sur l'USB.

2.2 DÉFINITIONS D'INTERVALLE DE SERVICE ET DE PAQUET D'INTERVALLE DE SERVICE

Note: La *Spécification USB audio 2.0* utilisait les termes de trame virtuelle et de paquet de trame virtuelle pour décrire les mêmes concepts que ceux appelés intervalle de service et paquet d'intervalle de service dans les *Spécifications USB 3.0 et ultérieures*. Dans la présente spécification, la terminologie a été consolidée pour utiliser les termes plus généraux d'intervalle de service et de paquet d'intervalle de service au lieu des termes spécifiques à l'USB Audio 2.0 de trame virtuelle et de paquet de trame virtuelle. En outre, le terme d'intervalle de bus est utilisé à la place de (micro)trame USB, le cas échéant.

Dans les alinéas suivants, le processus de mise en paquets pour l'audio est décrit par l'intervalle de service et les paquets d'intervalle de service. Un modèle cohérent avec "un paquet d'intervalle de service (SIP - Service Interval Packet) par intervalle de service (SI - Service Interval)" est ainsi fourni, indépendamment des transactions réelles sur l'USB et de la version d'USB utilisée.

Un intervalle de service est défini comme suit:

$$\text{Intervalle de service} = \text{Intervalle de bus} * 2^{(bInterval-1)}$$

où l'intervalle de bus a une valeur de 1 ms pour les points d'extrémité isochrones à vitesse intégrale et de 125 µs pour les points d'extrémité isochrones à grande vitesse, SuperSpeed et Enhanced SuperSpeed et où bInterval est la valeur spécifiée dans le champ bInterval du descripteur de point d'extrémité normalisé.

Un paquet d'intervalle de service correspond à la quantité de données isochrones transportées pendant toute la durée de l'intervalle de service. Pour les points d'extrémité à grande vitesse et à large bande passante, le paquet d'intervalle de service est la concaténation des deux ou trois paquets physiques transférés sur le bus au cours d'un intervalle de bus.

Note: La *Spécification USB* estime d'ores et déjà que les deux ou trois transactions d'un transfert à grande vitesse et à large bande passante font partie d'un seul et même paquet.

2.3 LES FORMATS DE DONNÉES AUDIO SIMPLES

2.3.1 FORMATS DE TYPE I

Les sections suivantes décrivent les formats de données audio appartenant au Type I. Plusieurs termes ainsi que leurs définitions sont présentés.

2.3.1.1 PAQUETS USB

Les flux de données audio qui sont intrinsèquement continus doivent être mis en paquets lorsqu'ils sont envoyés sur l'USB. La qualité de l'algorithme de mise en paquets influence directement la quantité d'effort nécessaire pour reconstruire une horloge d'échantillon fiable du côté de la réception.

De plus, la taille choisie pour l'intervalle de service a un impact direct sur la quantité de mémoire tampon nécessaire des deux côtés du canal et sur la latence occasionnée sur le canal. Des intervalles de service plus courts réduisent le plus possible les exigences de mémoire tampon et par conséquent la latence également, au détriment potentiel d'une consommation d'énergie plus élevée. En effet, des intervalles de service plus longs permettent éventuellement au bus (et à des parties du matériel de l'expéditeur et du destinataire) d'entrer dans des états de puissance inférieurs pendant de plus longues périodes, et donc d'économiser plus d'énergie.

La présente spécification définit deux modes de transport possibles des flux de données audio sur l'USB:

- le mode continu;
- le mode en salves.

Le mode continu a lieu lorsque l'intervalle de service est inférieur ou égal à une durée de trame USB de 1 ms. Ce mode réduit le plus possible les exigences de mémoire tampon ainsi que la latence au détriment potentiel d'une consommation d'énergie plus élevée. Noter cependant que le choix d'un intervalle de service inférieur à une durée de trame USB peut entraîner des interruptions excessives au niveau du système.

Le mode en salves a lieu lorsque l'intervalle de service est supérieur à une durée de trame USB. Ce mode permet d'économiser davantage d'énergie en permettant à différents composants du système d'entrer dans des états de faible puissance pendant de longues périodes, au détriment de tailles de mémoire tampon plus importantes et d'une latence accrue.

Les dispositifs peuvent choisir d'exposer plusieurs paramètres alternatifs de leur(s) interface(s) AudioStreaming avec des paramètres d'intervalle de service différents pour chaque paramètre alternatif, ce qui permet à l'hôte de choisir un paramètre correspondant au mieux au cas d'utilisation souhaité. Cependant, tous les dispositifs doivent exposer au moins un paramètre alternatif (en plus du paramètre alternatif de bande passante nulle 0) prenant en charge le mode continu (intervalle de service ≤ 1 ms).

2.3.1.1.1 CALCUL DE LA TAILLE DU PAQUET D'INTERVALLE DE SERVICE

L'objectif doit être de maintenir le nombre instantané d'emplacements audio par intervalle de service n_i aussi proche que possible du nombre moyen d'emplacements audio par intervalle de service n_{av} . La moyenne n_{av} doit être calculée comme suit:

$$n_{av} = \frac{T_{SI}}{\Delta t}$$

où T_{SI} est la durée d'un intervalle de service et Δt est le temps d'échantillon ($1/F_s$). Dans la plupart des cas, n_{av} est un nombre avec une partie fractionnaire.

Si le débit d'échantillonnage est une constante, la variation admissible de n_i est limitée à un emplacement audio, soit $\Delta n_i = 1$. Cela implique que tous les paquets d'intervalle de service doivent contenir $INT(n_{av})$ emplacements audio (SIP de petite taille) ou $INT(n_{av}) + 1$ emplacements audio (SIP de grande taille). Pour tout i :

$$n_i = INT(n_{av}) \text{ ou } INT(n_{av}) + 1$$

Note: Dans le cas où $n_{av} = INT(n_{av})$, n_i peut varier entre $INT(n_{av}) - 1$ (SIP de petite taille), $INT(n_{av})$ (SIP de taille moyenne) et $INT(n_{av}) + 1$ (SIP de grande taille).

En outre, un SIP de grande taille doit être généré dès qu'il devient disponible. Généralement, une source génère un certain nombre de SIP de petite taille tant que la partie fractionnaire accumulée de n_{av} reste < 1 . Lorsque la partie fractionnaire accumulée de n_{av} devient ≥ 1 , la source doit envoyer un SIP de grande taille et réduire l'accumulateur de 1.

En raison de la possibilité de notions de temps différentes entre la source et le destinataire (ils peuvent avoir chacun leur propre horloge d'échantillonnage indépendante), le modèle (SIP de petite taille)/(SIP de grande taille) généré par la source peut être différent de ce que le destinataire attend. Par conséquent, le destinataire doit être capable d'accepter un SIP de grande taille à tout moment.

Exemple:

Prendre pour hypothèse que $F_s = 44\,100$ Hz et $T_{SP} = 1$ ms. Puis que $n_{av} = 44,1$ emplacements audio. Comme la source ne peut envoyer qu'un nombre entier d'emplacements audio par intervalle de service, elle envoie des SIP de petite taille de 44 emplacements audio. Chaque intervalle de service envoie donc "0,1 emplacement" en trop et accumule cette partie fractionnaire dans un accumulateur. Après avoir envoyé 9 SIP de petite taille de 44 emplacements audio, il dispose au dixième intervalle de service d'exactly un emplacement audio en plus et peut donc envoyer un SIP de grande taille contenant 45 emplacements audio. Réduire l'accumulateur de 1 le ramène à 0 et le processus peut recommencer. La source produit ainsi un modèle répétitif de 9 SIP de petite taille de 44 emplacements audio suivis d'un SIP de grande taille de 45 emplacements audio. Le tableau suivant décrit le processus:

Tableau 2-1: Mise en paquets

#SI	n_{av}	n_i	Fraction	Accumulateur
n	44,1	44	0,1	0,1
n+1	44,1	44	0,1	0,2
n+2	44,1	44	0,1	0,3
n+3	44,1	44	0,1	0,4
n+4	44,1	44	0,1	0,5
n+5	44,1	44	0,1	0,6
n+6	44,1	44	0,1	0,7
n+7	44,1	44	0,1	0,8
n+8	44,1	44	0,1	0,9
n+9	44,1	45	0,1	1,0 -> 0
n+10	44,1	44	0,1	0,1
n+11	44,1	44	0,1	0,2
...	...	...	...	...

2.3.1.2 COMMANDE DE HAUTEUR

Si le débit d'échantillonnage peut être amené à varier (pour mettre en œuvre la commande de hauteur), la variation admissible de n_i est limitée à un emplacement audio par intervalle de service. Pour tout i :

$$n_{i+1} = n_i n_i \pm 1$$

La commande de hauteur est limitée aux points d'extrémité adaptatifs uniquement. Il est exigé que les interfaces AudioStreaming qui prennent en charge la commande de hauteur sur leur point d'extrémité isochrone l'indiquent dans le descripteur de point d'extrémité spécifique à la classe. De plus, une demande de commande de hauteur "Set/Get" est exigée pour activer ou désactiver la fonctionnalité de commande de hauteur.

2.3.1.3 SOUS-EMPLACEMENT AUDIO

La structure de base utilisée pour représenter les données audio est le sous-emplacement audio. Un sous-emplacement audio contient un échantillon audio unique. Un sous-emplacement audio contient toujours un nombre entier d'octets.

La présente spécification limite les tailles de sous-emplacement audio possibles (**bSubslotSize**) à 1, 2, 3, 4 ou 8 octets par sous-emplacement audio. Un échantillon audio est représenté par un nombre de bits (**bBitResolution**) inférieur ou égal au nombre total de bits disponibles dans le sous-emplacement audio, soit $\mathbf{bBitResolution} \leq \mathbf{bSubslotSize} * 8$.

Les points d'extrémité AudioStreaming doivent être construits de manière qu'un transfert valide puisse avoir lieu tant que la taille de sous-emplacement audio indiquée (**bSubslotSize**) est respectée pendant la transmission. Si le nombre de bits indiqué par échantillon (**bBitResolution**) ne correspond pas au nombre de bits de poids fort réellement utilisés pendant le transfert, le dispositif rejette les bits de poids fort ($[\text{actual_bits_per_sample}] > \mathbf{bBitResolution}$) ou interprète les zéros de fin comme des bits de poids fort ($[\text{actual_bits_per_sample}] < \mathbf{bBitResolution}$).

2.3.1.4 EMBLEMENT AUDIO

Un emplacement audio consiste en un ensemble de sous-emplacements audio, contenant chacun un échantillon audio d'un canal audio physique différent, prélevés au même instant. Le nombre de sous-emplacements audio dans un emplacement audio équivaut au nombre de canaux audio logiques dans la grappe de canaux audio. L'ordre des sous-emplacements audio dans l'emplacement audio obéit aux règles énoncées dans la *Spécification USB audio*. Tous les sous-emplacements audio doivent avoir la même taille.

Copyright © 1997-2016 USB Implementers Forum, Inc. Tous droits réservés.

2.3.1.5 FLUX AUDIO

Un flux audio est une concaténation d'un nombre potentiellement très important d'emplacements audio, ordonnés de façon croissante en fonction du temps. Les flux sont mis en paquets lorsqu'ils sont transportés sur USB et les SIP ne peuvent contenir qu'un nombre entier d'emplacements audio. Chaque paquet commence toujours par le même canal, et l'ordre des canaux est respecté tout au long de la transmission. Si, pour quelque raison que ce soit, il n'existe pas d'emplacement audio disponible pour construire un SIP, un délimiteur de transfert doit être envoyé à la place.

2.3.1.6 FORMATS DE TYPE I PRIS EN CHARGE

Les alinéas suivants répertorient tous les formats de données audio de Type I actuellement pris en charge. Les allocations de bits dans le champ **bmFormats** du descripteur d'interface AS spécifique à la classe pour les différents formats de données audio de Type I peuvent être consultées en A.1, "Allocations de bits des formats de données audio". Noter qu'un paramètre alternatif d'une interface AudioStreaming ne peut prendre en charge qu'un seul format de données audio de Type I. Par conséquent, seul un bit de format de données audio de Type I doit être défini dans le champ **bmFormats** du descripteur d'interface AS spécifique à la classe. La prise en charge d'un ou de plusieurs formats de données audio de Type III dans la même interface AudioStreaming est admise si l'interface est capable de distinguer indépendamment le format de données audio de Type I de tout autre format de données audio de Type III.

2.3.1.6.1 FORMAT MIC

Le format MIC (Modulation par Impulsion et Codage) est le format audio le plus couramment utilisé pour représenter des flux de données audio. Les données audio ne sont pas compressées et utilisent un format à virgule fixe signé en complément à deux. Il est aligné à gauche (le bit de signe est le Msb) et les données sont complétées par des zéros de fin pour remplir les bits inutilisés du sous-emplacement. Le point binaire est situé à droite du bit de signe de sorte que toutes les valeurs se trouvent dans la plage [-1, +1).

2.3.1.6.2 FORMAT MIC8

Le format MIC8 est introduit en vue d'être compatible avec le format WAVE de 8 bits existant. Les données audio ne sont pas compressées et utilisent 8 bits par échantillon (**bBitResolution** = 8). Dans ce cas, les données sont à virgule fixe, non signées, justifiées à gauche dans le sous-emplacement audio, en commençant par Msb. La plage est [0,255].

2.3.1.6.3 FORMAT IEEE_FLOAT

Le format IEEE_FLOAT est fondé sur la norme de virgule flottante ANSI/IEEE-754. Les données audio sont représentées au moyen du format simple précision de base. Le nombre correspondant au format simple précision de base est de 32 bits, avec un exposant de 8 bits et une mantisse de 24 bits. La mantisse et l'exposant sont des nombres signés, mais aucun n'est représenté au format en complément à deux. La mantisse est stockée au format de valeur de signe et l'exposant sous forme biaisée (également appelée décalage binaire). Sous forme biaisée, un entier positif (appelé biais) est soustrait du nombre stocké pour obtenir le nombre réel. Par exemple, dans un exposant de huit bits, le biais est 127. Pour représenter 0, le nombre 127 est stocké. Pour représenter -100, le nombre 27 est stocké. Un exposant de tous les zéros et un exposant de tous les uns sont tous deux réservés pour des cas particuliers, aussi dans un champ de huit bits des exposants de -126 à +127 sont possibles. Dans le format à virgule flottante de base, il est admis par hypothèse que la mantisse est normalisée de sorte que le bit le plus significatif est toujours un et n'est donc pas stocké. Seule la partie fractionnaire est stockée. Les valeurs dénormalisées (exposant = 0) sont vues comme nulles.

Le mot à virgule flottante IEEE-754 de 32 bits est divisé en trois champs. Le bit le plus significatif stocke le signe de la mantisse, le groupe suivant de 8 bits stocke l'exposant sous forme biaisée, et les 23 bits restants stockent la valeur de la partie fractionnaire de la mantisse. Pour plus d'informations, se référer à la norme ANSI/IEEE-754.

Les données sont transportées sur USB à raison de 32 bits par échantillon (**bBitResolution** = 32; **bSubslotSize** = 4).

2.3.1.6.4 FORMAT ALAW ET FORMAT μ LAW

A partir d'échantillons MIC linéaires de 12 ou 16 bits, une compression simple jusqu'à 8 bits par échantillon (un octet par échantillon) peut être obtenue en utilisant la compression-extension logarithmique. Les données audio compressées utilisent 8 bits par échantillon (**bBitsPerSample** = 8). Les données sont à virgule fixe, signées, justifiées à gauche dans le sous-emplacement, en commençant par Msb. La plage compressée est [-128,128]. La différence entre la compression Alaw et la compression μ Law réside dans les formules utilisées pour obtenir la compression. Se référer à la norme ITU G.711 pour plus d'informations.

2.3.1.6.5 FORMAT DSD

Le format Direct-Stream Digital (DSD) utilise le codage de modulation de densité d'impulsions, une technologie principalement utilisée pour stocker des signaux audio sur le support de stockage numérique SACD (Super Audio CD).

Les données audio sont stockées sous forme d'audio numérique modulé delta-sigma à bit unique, soit une séquence de valeurs à bit unique à un débit d'échantillonnage de 2,8224 MHz (64 fois le débit d'échantillonnage audio CD de 44,1 kHz) pour le débit d'échantillonnage de base DSD64, 5,6448 MHz pour DSD128 (débit 2X DSD), 11,2896 MHz pour DSD256 (débit 4X DSD), 22,5792 MHz pour DSD 512 (débit 8X DSD) et 45,1584 MHz pour DSD1024 (débit 16X DSD). Des flux DSD sur 48 kHz existent également. Dans ce cas, les débits d'échantillonnage du flux binaire sont respectivement de 3,072 MHz, 6,144 MHz, 12,288 MHz, 24,576 MHz et 49,152 MHz.

Quel que soit le débit d'échantillonnage utilisé par le flux DSD, la taille du sous-emplacement audio est fixée à 64 bits (**bBitResolution** = 64; **bSubslotSize** = 8) de sorte que, dans la couche de transport, le flux DSD ressemble toujours à des données MIC de 64 bits. Par conséquent, le débit d'échantillonnage de transport du flux DSD, mis en paquets sous forme d'échantillons MIC de 64 bits, est toujours égal à 1/64 du débit d'échantillonnage DSD et l'entité source d'horloge, connectée au terminal, représentant l'interface AudioStreaming DSD doit toujours annoncer ce débit d'échantillonnage de transport.

- 44,1 kHz, 88,2 kHz, 176,4 kHz, 352,8 kHz ou 705,6 kHz pour les flux DSD sur 44,1 kHz;
- 48 kHz, 96 kHz, 192 kHz, 384 kHz ou 768 kHz pour les flux DSD sur 48 kHz.

2.3.1.6.6 DONNÉES BRUTES DE TYPE I

Ce format audio est inclus pour permettre le transport de données (audio ou autre) sur une interface USB AudioStreaming sous la forme d'emplacements audio de type MIC lorsque le format réel ou bien la signification des données transportées est inconnu. Le canal USB sert simplement de passage. Par conséquent, ces données ne peuvent jamais être interprétées au sein de la fonction audio et peuvent uniquement être acheminées à partir d'un terminal d'entrée vers un ou plusieurs terminaux de sortie. D'un point de vue USB, les données sont mises en paquets comme s'il s'agissait de données audio formatées de Type I, mais les données ne doivent jamais être interprétées comme des données audio.

2.3.2 FORMATS DE TYPE III

Ces formats reposent sur la norme IEC61937. La norme IEC61937 décrit une méthode de transfert de flux de bits audio codés non MIC sur une interface audionumérique IEC60958, ainsi que le transfert des "Voies de signalisation" et des "Données utilisateur" associées.

La norme IEC60958 spécifie une méthode largement utilisée d'interconnexion d'équipements audio numériques avec un audio MIC linéaire à deux canaux. La norme IEC61937 décrit une manière dont l'interface IEC60958 doit être utilisée pour transporter des flux de bits audio codés non MIC pour des applications consommateur.

Les mêmes techniques de base utilisées dans l'IEC61937 sont réutilisées ici pour transporter des flux de bits audio codés non MIC sur un flux audio formaté de Type III. D'un point de vue de transfert USB, le flux de données sur l'interface ressemble exactement à des données audio MIC de 16 bits à deux canaux.

2.3.2.1 FORMATS DE TYPE III PRIS EN CHARGE

Se référer aux spécifications ISO/IEC 60958 et ISO/IEC 61937 (plusieurs parties) pour plus d'informations sur les formats. Les allocations de bits dans le champ **bmFormats** du descripteur d'interface AS spécifique à la classe pour les différents formats de données audio de Type III peuvent être consultées en A.1, "Allocations de bits des formats de données audio".

Noter qu'un paramètre alternatif d'une interface AudioStreaming peut prendre en charge plusieurs formats de données audio de Type III. Par conséquent, plusieurs bits peuvent être définis simultanément dans le champ **bmFormats** du descripteur d'interface AS spécifique à la classe. Il revient alors au décodeur d'inspecter le flux de données codé pour déterminer exactement le format actuellement utilisé sur le canal. En variante, un dispositif peut exposer plusieurs paramètres alternatifs sur une interface AudioStreaming, annonçant chacun la prise en charge d'un seul format de données audio de Type III. L'hôte peut ensuite choisir un paramètre alternatif approprié correspondant au format actuellement utilisé.

La liste suivante répertorie les formats couverts par les spécifications susmentionnées.

- MIC_IEC60958
- AC-3
- MPEG-1_Layer1
- MPEG-1_Layer2/3 ou MPEG-2_NOEXT
- MPEG-2_EXT
- MPEG-2_AAC_ADTS
- MPEG-2_Layer1_LS
- MPEG-2_Layer2/3_LS
- DTS-I
- DTS-II
- DTS-III
- ATRAC
- ATRAC2/3
- WMA
- E-AC-3
- MAT
- DTS-IV
- MPEG-4_HE_AAC
- MPEG-4_HE_AAC_V2
- MPEG-4_AAC_LC
- DRA
- MPEG-4_HE_AAC_SURROUND
- MPEG-4_AAC_LC_SURROUND
- MPEG-H_3D_AUDIO
- AC4
- MPEG-4_AAC_ELD

2.3.3 FORMATS DE TYPE IV

Les formats de Type IV ne peuvent être utilisés que sur les connexions externes à la fonction audio qui n'utilisent pas de canal USB pour leur transport de données, mais qui nécessitent une interface AudioStreaming pour annoncer leur présence. Un connecteur S/PDIF capable de gérer à la fois les flux de données audio stéréo MIC (IEC60958) d'un paramètre alternatif et les flux de données codés (IEC61937) d'un autre paramètre alternatif de l'interface constitue un exemple type d'une telle connexion. Noter cependant que la connexion externe peut également être spécifique au fournisseur (comme une interface de données parallèle).