IEC 61937-11

Edition 2.0 2021-07

INTERNATIONAL STANDARD

Digital audio – Interface for non-linear PCM encoded audio bitstreams applying IEC 60958 –
Part 11: MPEG-4 AAC and its extensions and MPEG-D USAC in LATM/LOAS

IEC 61937-11:2021-07(en)

IEC Central Office
3, rue de Varembé
CH-1211 Geneva 20
Switzerland

Tel.: +41 22 919 02 11
info@iec.ch
www.iec.ch

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 It also gives information on projects, replaced and withdrawn publications.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.

IEC online collection - oc.iec.ch
Discover our powerful search engine and read freely all the publications previews. With a subscription you will always have access to up to date content tailored to your needs.

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 18 additional languages. Also known as the International Electrotechnical Vocabulary (IEV) online.

IEC 61937-11

Edition 2.0 2021-07

INTERNATIONAL STANDARD

Digital audio – Interface for non-linear PCM encoded audio bitstreams applying IEC 60958 –
Part 11: MPEG-4 AAC and its extensions and MPEG-D USAC in LATM/LOAS

INTERNATIONAL
ELECTROTECHNICAL
COMMISSION

ICS 33.160.30; 33.170

ISBN 978-2-8322-9950-0

CONTENTS

INTERNATIONAL ELECTROTECHNICAL COMMISSION

____

DIGITAL AUDIO – INTERFACE FOR NON-LINEAR PCM ENCODED AUDIO BITSTREAMS APPLYING IEC 60958 –

Part 11: MPEG-4 AAC and its extensions and MPEG-D USAC in LATM/LOAS

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.

IEC 61937-11 has been prepared by technical area 20: Analogue and digital audio, of IEC technical committee 100: Audio, video and multimedia systems and equipment. It is an International Standard.

This second edition cancels and replaces the first edition published in 2010, and Amendment 1:2018. This edition constitutes a technical revision.

This edition includes the following significant technical changes with respect to the previous edition:

a) MPEG-D USAC has been added.

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

Draft	Report on voting
100/3523/CDV	100/3582/RVC

Full information on the voting for its approval can be found in the report on voting indicated in the above table.

The language used for the development of this International Standard is English.

A list of all parts in the IEC 61937 series, published under the general title *Digital audio – Interface for non-linear PCM encoded audio bitstreams applying IEC 60958*, can be found on the IEC website.

This document was drafted in accordance with ISO/IEC Directives, Part 2, and developed in accordance with ISO/IEC Directives, Part 1 and ISO/IEC Directives, IEC Supplement, available at www.iec.ch/members_experts/refdocs. The main document types developed by IEC are described in greater detail at www.iec.ch/standardsdev/publications.

The committee has decided that the contents of this document will remain unchanged until the stability date indicated on the IEC website under 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.

INTRODUCTION

Modern digital video broadcasting standards, such as DVB, include support for the MPEG-4 HE AAC and/or HE AAC v2 audio codecs specified in ISO/IEC 14496-3. An increasing number of countries are adopting these new codecs for their standard-definition and high-definition digital video broadcasting services and have started with implementations.

For MPEG-2 AAC audio (ISO/IEC 13818-7), the specified framing format for the audio bit stream is ADTS and its transport over an IEC 60958 interface is specified in IEC 61937-6.

However, the MPEG-4 (ISO/IEC 14496-3) and MPEG-D (ISO/IEC 23003-3) audio codecs introduce new features and capabilities that require a framing format that supports more flexible signalling and delivery mechanisms. Therefore, MPEG-2 systems (ISO/IEC 13818-1) specify the MPEG-4 LATM/LOAS framing format for MPEG-4 audio codecs to overcome the limitations of ADTS.

In order to be able to pass the MPEG-4 or MPEG-D audio bit stream from a set-top box to an A/V receiver connected via the IEC 60958 interface without needing to reframe the audio bit stream within ADTS, the MPEG-4 LATM/LOAS framing format needs to be supported by IEC 61937, including the high-speed transmission protocol where the interface does not carry an embedded sampling frequency clock.

DIGITAL AUDIO – INTERFACE FOR NON-LINEAR PCM ENCODED AUDIO BITSTREAMS APPLYING IEC 60958 –

Part 11: MPEG-4 AAC and its extensions and MPEG-D USAC in LATM/LOAS

1 Scope

This part of IEC 61937 describes the method to convey non-linear PCM bitstreams encoded in accordance with the MPEG-4 AAC format and its extensions (spectral band replication, parametric stereo and MPEG surround), and non-linear PCM bitstreams encoded in accordance with the MPEG-D USAC format, framed in MPEG-4 LATM/LOAS.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

IEC 60958-3:2021, *Digital audio interface – Part 3: Consumer applications*

IEC 61937-1:2021, *Digital audio – Interface for non-linear PCM encoded audio bitstreams applying IEC 60958 – Part 1: General*

IEC 61937-2:2021, *Digital audio – Interface for non-linear PCM encoded audio bitstreams applying IEC 60958 – Part 2: Burst-info*

ISO/IEC 14496-3:2019, *Information technology – Coding of audio-visual objects – Part 3: Audio*

ISO/IEC 23003-3:2020, *Information technology – MPEG audio technologies – Part 3: Unified speech and audio coding*

3 Terms and definitions

For the purposes of this document, the terms, definitions and abbreviated terms of IEC 61937-1 and IEC 61937-2, and the following apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- IEC Electropedia: available at http://www.electropedia.org/
- ISO Online browsing platform: available at http://www.iso.org/obp

3.1 Terms and definitions

3.1.1
access unit
smallest entity to which timing information can be attributed

Note 1 to entry: An access unit is the smallest individually decodable unit.

Note 2 to entry: A decoder consumes access units.

3.1.2
AudioMuxElement
LATM element that carries payload data for at least one audio elementary stream, related payload length information, and multiplex configuration information

Note 1 to entry: This element carries payload data in form of PayloadMux elements. If the term is followed by a number in parentheses, a "1" indicates that the multiplexing configuration (StreamMuxConfig) is multiplexed into the AudioMuxElement, i.e. the multiplexing configuration (StreamMuxConfig) is transmitted "in-band". A "0" indicates that the multiplexing configuration (StreamMuxConfig) is not present in the AudioMuxElement and needs to be transmitted by other means ("out-of-band").

3.1.3
AudioSpecificConfig
configuration structure used to convey parameters to initialize the MPEG-4 audio decoder

3.1.4
LATM
low overhead MPEG-4 audio transport multiplex
multiplexing layer defined by ISO/IEC 14496-3 used for multiplexing of audio elementary streams

3.1.5
LOAS
low overhead audio stream
synchronisation layer defined by ISO/IEC 14496-3

Note 1 to entry: Three different formats of LOAS are defined, each of which is designed to address the specific characteristics of the underlying transmission layer.

3.1.6
MPEG-4 AAC profile
contains only the MPEG-4 AAC low complexity audio object type

Note 1 to entry: The MPEG-4 AAC low complexity object type is the counterpart to the MPEG-2 AAC low complexity profile. In addition to the MPEG-2 AAC LC profile, the MPEG-4 AAC low complexity object type enables the usage of the PNS tool.

Note 2 to entry: The MPEG-4 AAC low complexity object type is used when there are restrictions on the usage of RAM and processing complexity.

3.1.7
MPEG-4 high-efficiency AAC profile
contains the spectral band replication object type in conjunction with the MPEG-4 AAC low complexity object type

Note 1 to entry: For further information, see ISO/IEC 14496-3. The MPEG-4 high-efficiency AAC profile is a superset of the MPEG-4 AAC profile.

3.1.8
MPEG-4 high-efficiency AAC profile version 2
contains the parametric stereo object type and the spectral band replication object type in conjunction with the AAC low complexity object type

Note 1 to entry: The MPEG-4 high-efficiency AAC profile version 2 is a superset of the MPEG-4 high-efficiency AAC profile.

3.1.9
MPEG-D Baseline USAC profile
profile that contains the Unified Speech and Audio Coding object type

3.1.10
MPEG-D extended high-efficiency AAC profile
profile that contains the parametric stereo object type and the spectral band replication object type in conjunction with the AAC low complexity object type, as well as the USAC object type

Note 1 to entry: The MPEG-D extended high-efficiency AAC profile is a superset of the MPEG-4 high-efficiency AAC profile version 2 and the MPEG-D Baseline USAC profile.

3.1.11
MPEG surround
technology used for coding of multichannel signals based on a downmixed signal of the original multichannel signal, and associated spatial parameters

Note 1 to entry: MPEG surround is defined in ISO/IEC 23003-1.

3.1.12
PayloadMux
payload data chunk in an AudioMuxElement that contains potentially multiplexed payload data for multiple audio elementary streams

Note 1 to entry: In general, PayloadMux elements can be concatenated inside AudioMuxElements.

3.1.13
SpatialSpecificConfig
configuration structure used to initialize the MPEG surround decoder

3.1.14
StreamMuxConfig
configuration structure that describes the structure of the LATM payload multiplex

3.1.15
MDCT
modified discrete cosine transformation
transformation schema used by AAC

3.1.16
transformation length
number of audio samples or corresponding MDCT lines that are processed as a block per each audio frame

Note 1 to entry: An MDCT line is a spectral component described by frequency, amplitude and phase.

3.1.17
USAC frame length
number of PCM audio samples per USAC frame

Note 1 to entry: USAC can operate in several modes using 1 024, 2 048, 4 096 or 768 linear PCM samples per USAC frame.

3.2 Abbreviated terms

AAC	Advanced Audio Coding
AAC LC	MPEG-4 AAC Low Complexity
HE AAC	MPEG-4 High-Efficiency AAC and MPEG-4 High-Efficiency AAC Version 2
ADTS	Audio Data Transport Stream
DVB	Digital Video Broadcasting
MDCT	modified discrete cosine transformation
MPEG	Moving Picture Experts Group
MPS	MPEG Surround

PNS	Perceptual Noise Substitution
PS	parametric stereo
SBR	spectral band replication
USAC	Unified Speech and Audio Coding
LT	AAC transformation length

4 Mapping of the audio bit stream on to IEC 61937-1

4.1 General

The coding of the bit stream and data-burst is in accordance with IEC 61937-1 and IEC 61937-2.

4.2 Burst-info for MPEG-4 AAC and its extensions and MPEG-D USAC in LATM/LOAS

The 16-bit burst-info contains information about the data found in the data-burst (see Table 1).

Table 1 – Values for data-type bits 0-4 and data-type bits 5-6

Data-type bits 0-4 according to IEC 61937-2 Value of Pc bits 0–4	Data-type bits 5-6 Value of Pc bits 5–6	Contents	Reference point R	Repetition period of data-bursts in IEC 60958 frames
0–22	0–3	According to IEC 61937		
23	0	According to IEC 61937		
	1	AAC LC	Bit 0 of Pa	960 / 1 024
	2	HE AAC	Bit 0 of Pa	1 920 / 2 048
	3	According to IEC 61937		
24	0–3	According to IEC 61937		
25	0–2	According to IEC 61937		
	3	USAC	Bit 0 of Pa	768 / 1 024 / 2 048 / 4 096
26-31	0–3	According to IEC 61937		

Bits 0–4 of the burst-info (Pc) signal the data-type bits 0-4 used for transmission. For MPEG-4 AAC-based audio in LATM/LOAS, the signalled data-type bits 0-4 is 23 (for AAC LC and HE AAC) or 25 (for USAC). Annex C gives a brief overview of MPEG-4 AAC, its extensions, and MPEG-D USAC.

If the Pc bits 0-4 are equal to 23, the Pc bits 5–6 indicate if the transmitted data stream contains audio encoded in AAC LC or HE AAC (including high-efficiency AAC version 2). Only values 1 and 2 refer to the transmission of AAC LC or HE AAC based audio. The values 0 and 3 are used for indication of codec types which are described by other or future parts of IEC 61937.

If the Pc bits 0-4 are equal to 25, the Pc bits 5–6 indicate if the transmitted data stream contains audio encoded in USAC. Only value 3 refer to the transmission of USAC based audio. The values 0, 1 and 2 are used for indication of codec types which are described by other or future parts of IEC 61937.

5 Format of data-burst for MPEG-4 AAC and its extensions and MPEG-D USAC in LATM/LOAS

5.1 General

Clause 5 specifies the data-burst for MPEG-4 AAC audio and its extensions and MPEG-D USAC in LATM/LOAS. Specific properties such as reference points, repetition period, the method of filling stream gaps, and decoding latency are specified.

The decoding latency (or delay), indicated for the data-type bits 5-6, should be taken into account by the transmitter to schedule data-bursts as necessary to establish synchronisation between picture and decoded audio.

5.2 Pause data-bursts for MPEG-4 AAC and its extensions in LATM/LOAS

Pause data-bursts for MPEG-4 AAC and its extensions in LATM/LOAS are defined in Table 2.

Table 2 – Repetition period of pause data-bursts

Data-type of audio data-burst	Repetition period of pause data-burst	
	Mandatory	Recommended
Data-type bits 5-6 for MPEG-4 audio in LATM/LOAS based on MPEG-4 AAC core codec	—	64 IEC 60958 frames

If regular audio data-bursts are not being transmitted due to for example a pause condition, it is recommended to use pause data-bursts to fill such stream gaps. The repetition period of the pause data-bursts should be selected in accordance with Table 2. If other repetition periods are necessary to precisely fill the stream gap length, or to meet the requirement on audio data-bursts spacing (see IEC 61937), pause data-bursts may have other lengths which can or not be an integer multiple of 64 IEC 60958 frames.

When a stream gap in an audio stream is filled by a sequence of pause data-bursts, the Pa of the first pause data-burst shall occur after exactly that amount of IEC 60958 frames as indicated by the AAC transformation length in conjunction with the codec type information from Table 3. It is recommended that the sequence(s) of pause data-bursts which fill the stream gap should continue from this point up to the Pa of the first audio data-burst which follows the stream gap, or as close as possible considering the specific IEC 60958 frame length of the pause data-burst with respect to the AAC core codec transformation length. The repetition-period-length parameter contained in the pause data-burst is intended to be interpreted by the receiver as an indication of the number of decoded PCM samples that are missing (due to the resulting audio gap).

5.3 Audio data-bursts

5.3.1 MPEG-4 AAC and its extensions in LATM/LOAS

The stream of data-bursts, as shown in Figure 1, consists of sequences of MPEG-4 AAC and its extensions in LATM/LOAS frames. Each data-burst consists of a preamble followed by the payload and stuffing. The data-type of a data-burst according to this specification is 23.

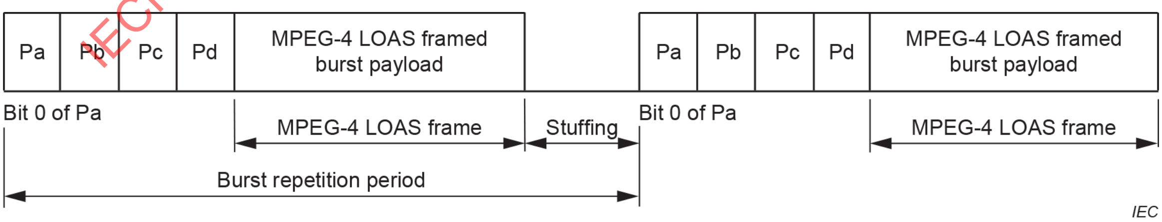


Figure 1 – Data-burst structure

The length of the audio payload data in the data-burst depends on the bit rate and other parameters of the encoded audio. The size of the data-burst payload is indicated by the Pd preamble word and is measured in bits.

P_{AD} = 4 × 16 bit (P_{AD} is the size of the preamble words Pa _ Pd measured in bits)

B_S = 4 × 16 bit (B_S is the size of the burst spacing measured in bits)

L_T = 1 024 or 960 lines (L_T is the used MDCT transformation length in MDCT lines)

The maximum data-burst payload size for AAC not utilizing SBR is calculated according to the following equation:

$2 \times 16 \text{ bit} \times L_T - (P_{AD} + B_S)$ = maximum payload size in bits.

If HE AAC is used, the maximum data-burst payload size is calculated according to the following equation:

$4 \times 16 \text{ bit} \times L_T - (P_{AD} + B_S)$ = maximum payload size in bits.

The data-type-dependent information for MPEG-4 AAC and its extensions in LATM/LOAS is given in Table 3. Bits 8–12 of Pc contain information about the audio codec used and about the LATM configuration.

Table 3 – Data-type-dependent information for MPEG-4 AAC audio and its extensions in LATM/LOAS

Bits of Pc LSB...MSB	Value	Definition	Description
8	0	AAC Transformation Length	1 024 lines
	1		960 lines
9	0	PS	PS data not present
	1		PS data present
10–11	0	MPS	MPS data not present
	1		Embedded MPS data present / LATM single layer transport mode (implicit MPS signalling)
	2		Embedded MPS data present / explicit signalling of MPS in second LATM layer
	3		Do not use until further definition
12	0	Reserved	Set to "0" until further definition
	1		Do not use until further definition

The Pc bit 8 indicates the transformation length of the AAC core codec which is used to encode the transmitted audio stream. Information from Pc bit 8 does not define the repetition period of data-bursts on its own. This information is required in conjunction with the codec signalled by the data-type bits 5-6 to calculate the data-burst repetition period. Receivers shall read the data-type bits 5-6 as well as the data-type-dependent information in order to compute the repetition period of data-bursts. Examples can be found in Annex A.

The Pc bit 9 indicates whether PS data is present in the encoded audio stream.

The two Pc bits 10–11 indicate the presence and transport configuration of MPS data in the encoded audio stream. The value 0 indicates that no MPS data is present. Values 1 and 2 indicate that MPS data is present in the audio bit stream. Pc bits 10-11 signalling 1 indicate that the MPS payload as well as the MPS SpatialSpecificConfig are embedded inside the payload of the first LATM layer, which conveys the AAC LC or HE AAC data stream. Pc bits 10–11 signalling 2 indicate that the MPS payload is also embedded inside the payload of the first layer. But in this case, the MPS SpatialSpecificConfig is signalled as being explicitly associated to the second layer inside the LATM StreamMuxConfig. The value 3 signalled by the Pc bits 10–11 is reserved for future use.

This specification does not allow the transmission of MPS payload that is not embedded inside the AAC LC or HE AAC payload, but resides separated from the AAC LC or HE AAC payload inside another LATM layer.

The presence of the PS or MPS extensions does not influence the data-burst repetition rate or the calculation of the transmission and decoding latency as described in 5.3.3.

The Pc bit 12 is reserved for future use. This bit shall be set to 0.

One complete AAC access unit represents a time interval of 1 024 or 960 audio samples embedded into the data-burst payload. When transmitting MPEG-4 HE AAC encoded audio programmes, SBR is used as an extension to AAC. In this case, the sampling frequency of the MPEG-4 AAC core component is usually half the sampling frequency of the SBR tool and audio program. One complete HE AAC access unit represents a time interval of 2 048 or 1 920 audio samples embedded into the data-burst payload.

HE AAC bit streams with downsampled SBR shall be transmitted signalling AAC LC in the data-type bits 5-6. Therefore, one access unit corresponds to 1 024 or 960 AAC encoded audio samples.

If HE AAC is signalled by the data-type bits 5-6, the IEC 60958 frame rate shall be equal to the sampling frequency of the SBR tool. If AAC LC is signalled, the IEC 60958 frame rate shall correspond to the sampling frequency of AAC. Annex B details the signalling of different IEC 60958 framerates and audio sample rates.

The reference point of a data-burst is bit 0 of Pa and occurs exactly once every number of IEC 60958 sampling periods, which is computed using the information from Table 2 and Table 3. The data-burst containing one LATM/LOAS audio frame shall occur at a constant rate. The intervals for data-bursts sharing the same bit-stream number shall correspond exactly to the amount of IEC 60958 frames calculated by using the information from Table 2 and Table 3.

It is not allowed to transmit audio data streams using IEC 60958 frame rates below 32 kHz.

5.3.2 USAC in LATM/LOAS

The stream of data-bursts as shown in Figure 2 consists of sequences of USAC in LATM/LOAS frames. Each data-burst consists of a preamble followed by the payload and stuffing. The data-type of a data-burst according to this specification is 25.

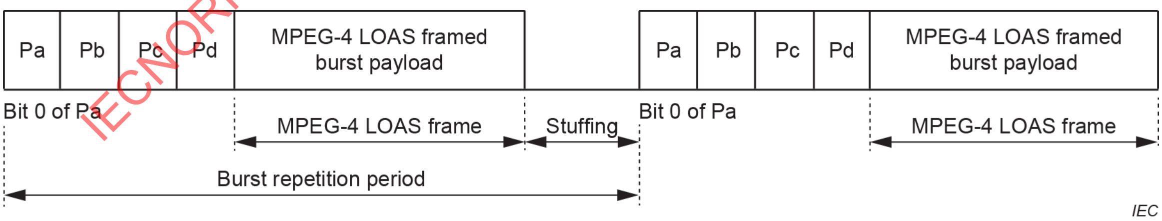


Figure 2 – Data-burst structure

The length of the audio payload data in the data-burst depends on the bit rate and other parameters of the encoded audio. The size of the data-burst payload is indicated by the Pd preamble word and is measured in bits.

The IEC 60958 frame rate for data-type USAC shall be equal to the audio sample rate, if the audio sample rate is between and including 32 kHz and 48 kHz. The IEC 60958 frame rate for data-type USAC shall be equal to twice the audio sample rate, if the audio sample rate is between and including 16 kHz and 24 kHz. The IEC 60958 frame rate for data-type USAC shall be equal to four times the audio sample rate, if the audio sample rate is between and including

8 kHz and 12 kHz. Annex B details the signalling of different IEC 60958 framerates and audio sample rates. The repetition period of data-bursts in IEC 60958 frames shall be determined from the Pc bits 8-10, in accordance with Table 4.

The maximum data-burst payload size in bits is determined from 2 × 16 bit × R_{p} – (P_{AD} + B_{S}), where

P_{AD} = 4 × 16 bit (P_{AD} is the size of the preamble words Pa – Pd measured in bits)

B_{S} = 4 × 16 bit (B_{S} is the size of the burst spacing measured in bits.)

R_{p} (Repetition period of data-bursts in IEC 60958 frames)

The data-type-dependent information for USAC in LATM/LOAS is given in Table 4. Bits 8–10 of Pc contain information about the repetition period.

Table 4 – Data-type-dependent information for USAC audio in LATM/LOAS

Bits of Pc LSB..MSB	Value	Meaning
8-10	0	1 024 IEC 60958 frames repetition period
	1	2 048 IEC 60958 frames repetition period
	2	4 096 IEC 60958 frames repetition period
	3	768 IEC 60958 frames repetition period
	4-7	Reserved
11-12	0-3	Reserved

The reference point of a data-burst is bit 0 of Pa and occurs exactly once every number of IEC 60958 sampling periods, which is computed using the information from Table 1 and Table 4. The data-burst containing one LATM/LOAS audio frame shall occur at a constant rate. The intervals for data-bursts sharing the same bit-stream number shall correspond exactly to the amount of IEC 60958 frames, which is calculated using the information from Table 1 and Table 4.

It is not allowed to transmit USAC audio data streams using IEC 60958 frame rates below 32 kHz.

5.3.3 LATM/LOAS framing

The LOAS frame as described in ISO/IEC 14496-3 shall be mapped directly to the payload section, right after the preamble words of the data-burst. The first bit of the LOAS frame shall always correspond to the first bit after the preamble section in the data-burst.

The payload in a data-burst consists of one complete LOAS frame containing one LATM AudioMuxElement. It is not allowed to convey one LATM/LOAS frame using multiple data-bursts. LOAS frames exceeding the payload capacity of a data-burst shall be dropped and the actual data-burst shall be replaced by a sequence of pause-bursts to match the duration of that data-burst.

The parameter numSubFrames from the LATM StreamMuxConfig shall be 0. The parameter numProgram from the LATM StreamMuxConfig shall be 0. The parameter numLayer from the LATM StreamMuxConfig shall be 0 except for audio streams signalling the availability of MPS with payload embedding and explicit signalling of MPS in the second LATM layer. In such cases, the presence of a second layer in LATM frames is allowed and, therefore, numLayer shall be 1, indicating 2 layers. In this configuration, there exists no payload associated to the second LATM layer and, therefore, the payload size indication for the second layer in LATM is set to zero.

Only the LOAS AudioSyncStream() scheme shall be used in the context of this specification.

The LATM StreamMuxConfig structure shall be conveyed inside the LATM multiplex. This is the main structure that is used by the decoder for configuration. The StreamMuxConfig may not be present in each LATM frame in order to save bandwidth. It may be sent in intervals to allow decoders to tune in to a running stream.

5.3.4 Latency

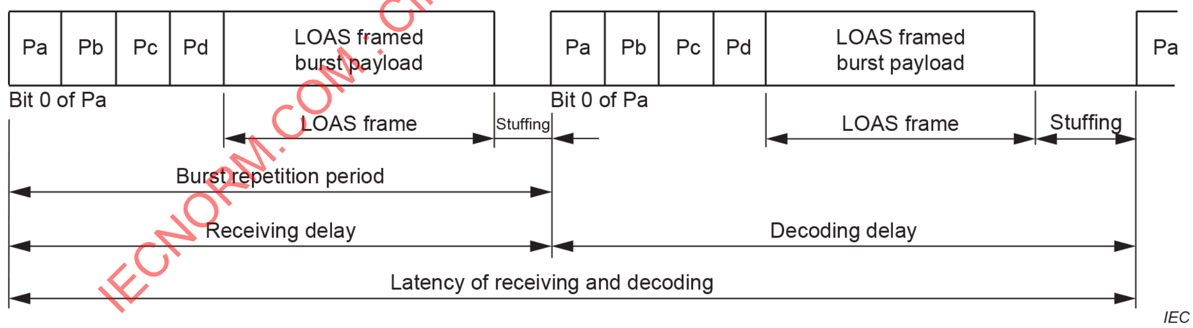
The latency of an external audio decoder to decode MPEG-4 AAC and its extensions in LATM/LOAS is defined as the sum of the receiving time of the audio payload in one data-burst and the time used for decoding of one access unit.

Each data-burst contains a minimum of 4 stuffing words (Pz of 16 bits). The repetition period of data-bursts in IEC 60958 frames is computed according to information from Table 1 and Table 3. The reception delay for one audio access unit is calculated as the time elapsed counting from the first bit of the data-burst until the last bit of the actual audio payload inside the data-burst received. Subsequent stuffing is not taken into account. After a complete frame is received, immediate decoding and subsequent rendering of the audio frame is not recommended as the size of the next audio frame, and therefore the time required for receiving it completely, cannot be determined accurately.

In order to simplify the timing mechanism for receiving and decoding of content of data-bursts, the receiving delay should be calculated as the time necessary to receive the complete data-burst including the stuffing. The maximum time available to decode (the decoding delay) should be selected to correspond to the length of one full data-burst. This results in an overall delay corresponding to two complete data-bursts for reception and decoding.

For synchronisation (for example with video), the recommended value for latency corresponds to the time necessary to receive two complete data-bursts. Figure 3 shows the simplified and recommended method for calculating the latency for reception and decoding.

A shorter latency may be acceptable if synchronisation is not required.



This diagram shows the recommended method.

Figure 3 – Latency diagram for burst reception and decoding

Annex A
(informative)

Calculation of delay and data-burst repetition rates – guidelines

A.1 Examples

Some examples for the calculation of data-burst-repetition rates and latencies for AAC LC and HE AAC can be found in Table A.1.

Table A.1 – Examples – Calculation of delay and data-burst repetition rates for AAC_LC and HE AAC

Bits 5-6 of Pc codec indication	**Bit 8 of Pc TL** lines	**AAC sampling rate** kHz	**SBR sampling rate** kHz	**IEC 60958 frame rate** kHz	**Data-burst repetition rate** IEC 60958 frames	**Overall latency** ms
AAC LC	1 024	32	n/a	32	1 024	64
		44,1	n/a	44,1		46,44
		48	n/a	48		42,67
		96	n/a	96		21,33
	960	32	n/a	32	960	60
		44,1	n/a	44,1		43,54
		48	n/a	48		40
		96	n/a	96		20
HE AAC	1 024	16	32	32	2 048	128
		22,05	44,1	44,1		92,88
		24	48	48		85,33
		48	96	96		42,67
	960	16	32	32	1 920	120
		22,05	44,1	44,1		87,07
		24	48	48		80
		48	96	96		40

The presence of the PS or MPS extensions signaled by Pc bits 9–11 does not influence the data-burst repetition rate or the calculation of the transmission and decoding latency as described in 5.3.4.

Some examples for the calculation of data-burst-repetition rates and latencies for USAC can be found in Table A.2.

Table A.2 – Examples – Calculation of delay and data-burst repetition rates for USAC

Bit 8-10 of Pc **IEC 60958 frames repetition period**	**USAC sampling rate** kHz	**USAC frame length** PCM samples	**IEC 60958 frame rate** kHz	**Data-burst repetition rate** IEC 60958 frames	**Overall latency** ms
1 024	32	1 024	32	1 024	64
	44,1		44,1		46,44
	48		48		42,67
2 048	16	1 024	32	2 048	128
	22,05		44,1		92,88
	24		48		85,33
	32	2 048	32		128
	44,1		44,1		92,88
	48		48		85,33
4 096	8	1 024	32	4 096	256
	11,025		44,1		185,76
	12		48		170,66
	16	2 048	32		256
	22,05		44,1		185,76
	24		48		170,66
	32	4 096	32		256
	44,1		44,1		185,76
	48		48		170,66

A.2 Guidelines

The following guidelines should be taken into account.

a) Receivers that receive an indication in the data-burst-dependent information that signals the presence of MPS, but that are not capable of decoding MPS, should not refuse decoding of that stream. It is highly recommended that non-MPS capable decoders decode just the AAC LC/HE AAC channel configuration as indicated by the downmix codec configuration record and ignore the MPS extension in the bit stream.

b) The IEC 60958 frame rate may be calculated by making use of the audio sampling rate indication from the AudioSpecificConfig inside the LATM StreamMuxConfig. It is highly recommended that the correct codec indication, as well as the matching sampling frequency or IEC 60958 frame rate indication, be available before starting transmission of IEC 61937-11 data-bursts. In the case of signalled audio configuration changes upstream, it is highly recommended that audio data-bursts referring to the new programme are only transmitted after the relevant information (new codec and new sampling frequency) is available to the transmitter and signalled properly.

It is highly recommended that decoders do not attempt to decode an audio stream before they have received the corresponding decoder configuration records (e.g. AudioSpecificConfig).

Annex B
(normative)

High-speed transmission

B.1 Indication

Typically, the transmitting interface frame rate equals the sampling frequency for the IEC 61937-11 protocol. In the case of a mismatch, IEC 61937-11 uses channel status fields to identify their relationship, as shown in Table B.1 and in the following scheme.

Table B.1 – Indication fields

Mode 0 Channel status	IEC 60958-3 (LPCM)	IEC 61937-11 (MPEG-4 AAC)
Bit 24 to 27	Sampling frequency	IEC 60958 frame rate
Bit 30 to 31	Sampling frequency extension	
Bit 36 to 39	Original sampling frequency	
Bit 44 to 47	Audio sampling frequency coefficient	

The "original sampling frequency" identifies reproduction frequency after decoding the IEC 61937-11 bitstream. The IEC 61937-11 bitstream is transmitted at "IEC 60958 frame rate". The "audio sampling frequency coefficient" optionally shows the relationship between "IEC 60958 frame rate" and "original sampling frequency" as follows:

IEC 60958 frame rate = (original sampling frequency) × (audio sampling frequency coefficient)

Values of these fields are given in IEC 60958-3.

IEC 61937-11 uses clock accuracy carried on channel status bits 28 and 29 of the IEC 60958-3 definitions for accuracy of the interface frame rate. The value of "1 1" (interface frame rate not matched to sampling frequency) stands for the condition that the actual interface frame rate in transmitting is not matching the logical IEC 60958 frame rate.

B.2 Example

Table B.2 shows two examples with a 96 kHz interface frame rate and a 48 kHz reproduction frequency, and a 32 kHz interface frame rate and an 8 kHz reproduction frequency. ARIB STD B-32 shows the 96 kHz interface example.

Table B.2 – Signalling example

IEC 60958 frame rate	Original sampling frequency	Audio sampling frequency coefficient
"0 1 0 1 0 0" (96 kHz)	"1 0 1 1" (48 kHz)	"1 1 1 1" (×2) or "0 0 0 0" (No indication)
"1 1 0 0 0 0" (32 kHz)	"0 1 1 0" (8 kHz)	"1 1 1 0" (×4) or "0 0 0 0" (No indication)