IEC 62365

Edition 2.0 2009-05

INTERNATIONAL STANDARD

NORME INTERNATIONALE

Digital audio – Digital input-output interfacing – Transmission of digital audio over asynchronous transfer mode (ATM) networks

Audionumérique – Interface numérique d'entrée-sortie – Transmission de l'audionumérique sur les réseaux à mode de transfert asynchrone (ATM)

IEC 62365:2009

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

Tel.: +41 22 919 02 11
Fax: +41 22 919 03 00
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 corrigenda or an amendment might have been published.

Useful links:

IEC publications search - www.iec.ch/searchpub

The advanced search enables you 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It also gives information on projects, replaced and withdrawn publications.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 on-line and also once a month by email.

Electropedia - www.electropedia.org

The world's leading online dictionary of electronic and electrical terms containing more than 30 000 terms and definitions in English and French, with equivalent terms in additional languages. Also known as the International Electrotechnical Vocabulary (IEV) on-line.

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: csc@iec.ch.

A propos de la CEI

La Commission Electrotechnique Internationale (CEI) est la première organisation mondiale qui élabore et publie des Normes internationales pour tout ce qui a trait à l'électricité, à l'électronique et aux technologies apparentées.

A propos des publications CEI

Le contenu technique des publications de la CEI est constamment revu. Veuillez vous assurer que vous possédez l’édition la plus récente, un corrigendum ou amendement peut avoir été publié.

Liens utiles:

Recherche de publications CEI - www.iec.ch/searchpub

La recherche avancée vous permet de trouver des publications CEI en utilisant différents critères (numéro de référence, texte, comité d’études,…).
Elle donne aussi des informations sur les projets et les publications remplacées ou retirées.

Just Published CEI - webstore.iec.ch/justpublished

Restez informé sur les nouvelles publications de la CEI. Just Published détaille les nouvelles publications parues. Disponible en ligne et aussi une fois par mois par email.

Electropedia - www.electropedia.org

Le premier dictionnaire en ligne au monde de termes électroniques et électriques. Il contient plus de 30 000 termes et définitions en anglais et en français, ainsi que les termes équivalents dans les langues additionnelles. Egalement appelé Vocabulaire Electrotechnique International (VEI) en ligne.

Service Clients - webstore.iec.ch/csc

Si vous désirez nous donner des commentaires sur cette publication ou si vous avez des questions contactez-nous: csc@iec.ch.

IEC 62365

Edition 2.0 2009-05

INTERNATIONAL STANDARD

NORME INTERNATIONALE

Digital audio – Digital input-output interfacing – Transmission of digital audio over asynchronous transfer mode (ATM) networks

Audionumérique – Interface numérique d'entrée-sortie – Transmission de l'audionumérique sur les réseaux à mode de transfert asynchrone (ATM)

INTERNATIONAL ELECTROTECHNICAL COMMISSION

COMMISSION ELECTROTECHNIQUE INTERNATIONALE

PRICE CODE
CODE PRIX U

ICS 35.200, 33.160

ISBN 978-2-83220-241-8

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.

CONTENTS

INTERNATIONAL ELECTROTECHNICAL COMMISSION

DIGITAL AUDIO – DIGITAL INPUT-OUTPUT INTERFACING – TRANSMISSION OF DIGITAL AUDIO OVER ASYNCHRONOUS TRANSFER MODE (ATM) NETWORKS

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 provides no marking procedure to indicate its approval and cannot be rendered responsible for any equipment declared to be in conformity with an IEC Publication.

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 62365 has been prepared by technical area 4: Digital systems interfaces and protocols, of IEC technical committee 100: Audio, video and multimedia systems and equipment.

This bilingual version (2012-08) corresponds to the monolingual English version, published in 2009-05.The text of this standard is based on the following documents:

FDIS	Report on voting
100/1517/FDIS	100/1550/RVD

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

The French version of this standard has not been voted upon.

This publication has been drafted in accordance with the ISO/IEC Directives, Part 2.

This second edition of IEC 62365 cancels and replaces the first edition published in 2004.

The main changes with respect to the previous edition (2004) are listed below.

– Second, third, and fourth required formats in 4.3 removed.
– 4.3 reformatted, eliminating Table 2, and subsequent Tables renumbered.

The committee has decided that the contents of this publication will remain unchanged until the maintenance result date indicated on the IEC web site under "http://webstore.iec.ch" in the data related to the specific publication. At this date, the publication will be

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

INTRODUCTION

This International Standard describes means for the transmission of professional audio across digital networks, including metropolitan- and wide-area networks, to provide the best performance with regard to latency, jitter, and other relevant factors.

Current-generation wide-area communications are based on two very similar systems, synchronous optical network (SONET) and synchronous digital hierarchy (SDH), SONET being used in the United States and SDH in Europe. On top of them are run integrated services digital network (ISDN), asynchronous transfer mode (ATM), and Internet protocol (IP).

ISDN provides telephone call connections of a fixed capacity that carry one 8-bit value per 125 μs; when a call is set up, its route through the system is chosen, and the switches that route the data are configured accordingly. Each link, between switches or between switch and end equipment, is formatted into frames that take 125 μs to transmit, and each data byte is identified by its position in the frame.

ATM, also called broadband ISDN, provides a service similar to ISDN, but with the capacity of each call being specified by the caller. Links are formatted into cells, which consist of a header and 48 data bytes; the header is typically 5 bytes long, and most of it is taken up with the virtual channel identifier (VCI) that shows to which call the cell belongs. Call set-up, routeing, and switching are done in the same way as in ISDN, but with calls not being restricted to 1 byte every 125 μs.

IP provides a very different service, not designed for continuous media such as audio and video. There is no call set-up, and each packet contains enough information within itself to allow it to be routed to its destination. This means that the header is much larger than in the case of ATM, typically 74 bytes, and packets will also typically be much larger, if only because otherwise the overheads would be excessive. Each packet is liable to be routed separately, so two packets that are part of the same flow may well take different routes. This can mean that the one that was sent first does not arrive first.

For many professional audio applications, a round-trip time from the microphone through the mixing desk and back to the headphones of no more than 3 ms is required. Allowing 0,5 ms each for conversion from analog to digital and back again, it follows that the network connections to and from the mixing desk must have a latency of less than 1 ms each. For distances of more than about 200 km, the transmission delay alone will exceed 1 ms, but within a metropolitan area the transmission delay should be no more than 0,25 ms (equivalent to about 50 km), leaving 0,75 ms for packetization, queuing within switches, and resynchronization within the receiving equipment.

Packetization delays are proportional to the size of the transmission unit (frame, cell, or packet), and resynchronization delays depend on how evenly spaced the transmission units are when they arrive at their destination. Both classes of delay are thus small for ISDN and large for IP. Using the format specified in this standard to carry dual-channel IEC 60958-4 audio with a 48 kHz sampling frequency over ATM results in an inter-cell time of 125 μs, at which ATM will have similar delays to ISDN. A higher sampling frequency or a larger number of channels would reduce the inter-cell time and hence also the delays.

The queuing time within each ISDN switch is likely to be around one frame time or 125 μs. The ATM documents limit the queuing time in an ATM switch to approximately the inter-cell time for the call, which, as with the other delays, translates into performance similar to that of ISDN for dual-channel 48 kHz IEC 60958-4 audio and better for higher sampling frequencies or larger numbers of channels.

The queuing time within an IP router for normal, best effort, Internet traffic is unbounded, and if the router is congested, packets may simply be thrown away. Resource reservation protocol

(RSVP) (see Annex A) allows capacity to be reserved for a particular traffic flow, but it does not guarantee that the packets will actually be routed over the links on which the capacity has been reserved; if the flow is re-routed, it will only get a best effort service until a reservation has been made on the new route, and it may not even be possible to make a reservation on the new route at all.

ATM has therefore been chosen as providing a more convenient service than ISDN and significantly better performance than IP, even when RSVP is used.

This standard does not specify a physical interface to the network because one of the features of ATM is its ability to make a seamless connection between interfaces operating at a wide variety of data rates and with different ways of encoding the ATM cells. Commonly used interfaces provide 25,6 Mbit/s over category 3 structured wiring and 155,52 Mbit/s over category-5 structured wiring or fibre-optic cable.

The physical layer section description and unique ATM abbreviations can be found in ATM forum approved specifications. See the Bibliography.

DIGITAL AUDIO – DIGITAL INPUT-OUTPUT INTERFACING – TRANSMISSION OF DIGITAL AUDIO OVER ASYNCHRONOUS TRANSFER MODE (ATM) NETWORKS

1 Scope

This International Standard specifies a means to carry multiple channels of audio in linear PCM or IEC 60958-4 format over an ATM layer service conforming to ITU-T Recommendation I.150. It includes a means to convey, between parties, information concerning the digital audio signal when setting up audio calls across the ATM network.

It does not specify the physical interface to the network.

2 Normative references

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

IEC 60958-1, *Digital audio interface – Part 1: General*

IEC 60958-4, *Digital audio interface – Part 4: Professional applications (TA4)*

ITU-T Recommendation I.150: *B-ISDN asynchronous transfer mode functional characteristics*

ITU-T Recommendation I.363.5, *B-ISDN ATM Adaptation Layer specification: Type 5 AAL*

ITU-T Recommendation Q.2931: *Digital Subscriber Signalling System No. 2 – User-Network Interface (UNI) layer 3 specification for basic call/connection control*

3 Terms and definitions

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

3.1
asynchronous transfer mode
ATM
networking technology in which data are carried in 48-o cells

NOTE Octet (unit symbol, o) is defined as an 8-bit data element by IEC 60027-2, which is synonymous with byte (unit symbol, B) whenever the term, byte, is restricted to 8-bit elements.

3.2
ATM adaptation layer
AAL
protocol layer that allows different services, such as packet transfer, to be provided on an ATM network

3.3
ATM signaling
protocol that conveys connection management and other messages between an ATM network and equipment attached to it

3.4
audio channel
path that carries one monophonic digital audio signal

3.5
audio port
physical or virtual connector that carries a fixed number of audio channels

3.6
information element
IE
component of an ATM signalling message

3.7
MADI
serial multi-channel audio digital interface

3.8
organizationally unique identifier
OUI
3-o code issued by a designated agency to form globally consistent bit strings as described in *OUI and company_id assignments*

3.9
user-to-user indication
UI
single bit in the ATM cell header that can be used by the ATM adaptation layer as a marker for certain cells

3.10
virtual channel
communications channel that provides for the sequential unidirectional transport of ATM cells on a link between two pieces of equipment

3.11
virtual channel identifier
VCI
numerical tag occupying a 16-bit field in the ATM cell header that identifies the virtual channel over which the cell is to travel

3.12
virtual circuit
route through a network formed by concatenating virtual channels

3.13
virtual path
group of up to 65 536 virtual channels

3.14
virtual path identifier
VPI
numerical tag occupying an 8-bit field in the ATM cell header that identifies the virtual path which contains the virtual channel over which the cell is to travel

4 Format of audio data in ATM cells

4.1 Format of audio samples

4.1.1 Subframes

4.1.1.1 Each audio sample shall be encoded in a subframe that consists of a whole number of octets. The subframe shall be stored in the cell in consecutive octets, with the first bit of the subframe in the most significant bit of the first octet.

4.1.1.2 A subframe shall consist of the fields listed in Table 1, in the order in which they appear.

Table 1 – Fields contained in a subframe

Field	Specified in
Audio sample word	4.1.2
Ancillary data	4.1.3
Protocol overhead	4.1.4

4.1.2 Audio sample word

4.1.2.1 The audio sample shall be represented in linear 2's complement form, with the most significant bit first. If the source provides fewer bits than the size of this field, the unused least significant bits shall be set to zero.

NOTE This specification is the same as in IEC 60958-4, except that the bit order is reversed.

4.1.2.2 The number of bits in the audio sample word shall be chosen in such a way that the total number of bits in the subframe is 8, 16, 24, 32, or 48.

4.1.3 Ancillary data

4.1.3.1 This field shall either contain no bits or consist of four bits designated B, C, U, V, in that order.

4.1.3.2 The C, U, and V bits shall be the channel status, user data, and validity bits specified in IEC 60958-1.

4.1.3.3 The B bit shall be a 1 for the first subframe of the block specified in IEC 60958-1, and a 0 for all other subframes.

NOTE The B bit affects the interpretation of the C bit, and possibly also of the U bit, but has no relation to the grouping of samples specified in 4.2.

Where more than one audio channel is carried, the B bit shall be set at the start of the block in every channel, not just in the first channel. The block starts may be unaligned.

4.1.4 Protocol overhead

This field shall either contain no bits or consist of a sequencing bit followed by three bits that provide data protection.

4.1.4.1 Sequencing word

The sequencing word consists of the sequencing bits of all the subframes in a cell, in the order in which the subframes appear in the cell.

4.1.4.1.1 Sequence number

The first four bits of the sequencing word shall contain a sequence number in the form of a binary integer with the least significant bit first.

Except in the first cell transmitted on a virtual circuit, the value of this integer shall be 1 more (modulo 16) than in the previous cell on the same virtual circuit.

The value of this integer in the first cell transmitted on each virtual circuit shall be chosen such that in the first cell of each block, as specified in 4.5, the least significant three bits shall be zero.

The value of the most significant bit in the first cell transmitted on each virtual circuit shall not be defined in this standard.

4.1.4.1.2 Sequence number protection

The fifth to seventh bits of the sequencing word shall contain the 1's complement of the remainder of the division (modulo 2) by the generator polynomial $x^3 + x + 1$ of the product x^3 multiplied by the sequence number. The coefficient of the x^2 term in the remainder polynomial is the fifth bit.

The eighth bit of the sequencing word shall be such that there are an even number of 1's in the first eight bits.

NOTE Additional information is given in Annex A.

4.1.4.1.3 Second number

The ninth to twelfth bits of the sequencing word may contain a second number in the form of a binary integer with the least significant bit first.

a) The value of this integer in the first cell transmitted on each virtual circuit shall be defined in this standard only as in (b). Its value in a cell which is the first cell of a block (as specified in 4.5) and has its user indication bit set to 1 shall be 1 more (modulo 16) than in the previous cell on the same virtual circuit. Its value in each other cell shall be equal to that in the previous cell on the same virtual circuit.
b) Where two ATM virtual circuits carry data from sources that use the same local clock as specified in 4.5, there may be a defined relationship between the second number values on the two connections which can allow co-temporal samples on the two connections to be identified. The method by which the necessary information is conveyed to receiving equipment is not specified in this standard.

   NOTE The second number may be used to identify samples uniquely within a 16-second period.
c) If the sender does not support the inclusion of the second number, these four bits shall be zero in every cell.

4.1.4.1.4 Remainder of sequencing word

Any further bits in the sequencing word shall be reserved and shall be set to zero on transmission and ignored on reception.

4.1.4.2 Data protection bits

If the ancillary data field contains a V bit, the three data protection bits shall contain the 1's complement of the remainder of the division (modulo 2) by the generator polynomial $x^3 + x + 1$ of the sum of the product x^4 multiplied by the most significant nine bits of the audio sample and the product x^3 multiplied by the V-bit.

Otherwise, the three data protection bits shall contain the 1's complement of the remainder of the division (modulo 2) by the generator polynomial $x^3 + x + 1$ of the product x^3 multiplied by the most significant nine bits of the audio sample.

In either case, the coefficient of the x^2 term in the remainder polynomial is the first (most significant) of the three bits.

NOTE This protection scheme is appropriate for linear PCM audio samples. Other data types carried in these streams may need to arrange additional protection within their codecs.

4.2 Packing of sample data into cells

4.2.1 Packing schemes

4.2.1.1 An ATM virtual circuit shall carry either a single audio channel or a group of audio channels. In the latter case, all audio channels in the group shall use the same format and share the same sample clock.

For the purpose of this description, audio channels shall be numbered from 1 upwards. In the examples, the sample times are given letters, so for instance 2a is the first sample on audio channel 2 and 2b is the second.

4.2.1.2 The number of samples per cell shall be 48 divided by the number of octets in a subframe (see 4.1.1).

4.2.1.3 On each ATM virtual circuit, one of the packing schemes specified in 4.2.2, 4.2.3, and 4.2.4 shall be used. To assist interoperability, temporal grouping should be used in preference to grouping by channel.

NOTE Only certain combinations of subframe size and number of audio channels are possible; if necessary, an application may leave some audio channels unused.

4.2.1.4 The audio sample data in every subframe of an unused audio channel shall be 0.

4.2.2 Temporal grouping

The number of samples per cell shall be divisible by the number of audio channels.

Co-temporal samples shall be grouped together. Samples within a group shall be in audio channel number order and groups shall be in temporal order.

A block, for the purposes of 4.5, shall consist of eight cells.

EXAMPLE

2 channels, 12 samples per cell: 1a, 2a, 1b, 2b, 1c, 2c, 1d, 2d, 1e, 2e, 1f, 2f.

4.2.3 Multi-channel

The number of audio channels shall be divisible by the number of samples per cell.

Samples shall be in channel number order.

A block, for the purposes of 4.5, shall consist of eight sets of samples.

EXAMPLE

24 channels, 12 samples per cell: 1a ... 12a in first cell; 13a ... 24a in second; 1b ... 12b in third.

4.2.4 Grouping by channel

The number of samples per cell shall be divisible by the number of channels.

Samples on the same channel shall be grouped together; samples within a group shall be in temporal order, and groups shall be in channel number order.

A block (for the purposes of 4.5) shall consist of eight cells.

EXAMPLE

2 channels, 12 samples per cell: 1a, 1b, 1c, 1d, 1e, 1f, 2a, 2b, 2c, 2d, 2e, 2f.

NOTE If there is just one channel, this scheme is identical to temporal grouping; if the number of channels is equal to the number of samples per cell, all three schemes are identical.

4.3 Formats

4.3.1 Only those subframe formats, packing schemes, and sampling frequencies that are expressible in the notation of Clause 6 shall be used.

NOTE See additional restrictions in 4.1.2 and 4.2.1.

4.3.2 To increase the likelihood that equipment designed for different applications will interoperate successfully, all equipment should support the format with 24 audio data bits, 4 ancillary data bits, and 4 protocol overhead bits, packed as 2 channels with temporal grouping as specified in 4.2.2.

NOTE This format is signaled by the values 56_{16} in the second octet of the AAL Parameters IE and 02_{16} in the third octet; it is the appropriate format for conveying AES3 transparently.

4.3.3 Equipment that can convey at least 56 audio channels should support the format with 24 audio data bits, 4 ancillary data bits, and 4 protocol overhead bits as a multi-channel call carrying 60 channels, as specified in 4.2.3.

NOTE This format is signaled by the values 56_{16} in the second octet of the AAL parameters IE and 85_{16} in the third octet; it is the appropriate format for conveying 56-channel MADI transparently (the last 4 channels being unused).

4.3.4 Sampling frequencies should be as specified in AES5.

NOTE The preferred sampling frequency specified in AES5 is 48 kHz.

4.4 ATM adaptation layer

Audio virtual circuits shall use a user-defined ATM adaptation layer.

4.5 ATM-user-to-ATM-user indication

4.5.1 Cells shall be grouped into blocks as specified in 4.2.

4.5.2 The sender shall include a local clock, which ticks once per second.

NOTE This standard does not specify the accuracy of the local clock, nor to what (if anything) it is synchronized. It need not be related to the audio sample clock.