# INTERNATIONAL STANDARD

# IEC 62365

First edition
2004-11

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



Reference number
IEC 62365:2004(E)

## Publication numbering

As from 1 January 1997 all IEC publications are issued with a designation in the 60000 series. For example, IEC 34-1 is now referred to as IEC 60034-1.

## Consolidated editions

The IEC is now publishing consolidated versions of its publications. For example, edition numbers 1.0, 1.1 and 1.2 refer, respectively, to the base publication, the base publication incorporating amendment 1 and the base publication incorporating amendments 1 and 2.

## Further information on IEC publications

The technical content of IEC publications is kept under constant review by the IEC, thus ensuring that the content reflects current technology. Information relating to this publication, including its validity, is available in the IEC Catalogue of publications (see below) in addition to new editions, amendments and corrigenda. Information on the subjects under consideration and work in progress undertaken by the technical committee which has prepared this publication, as well as the list of publications issued, is also available from the following:

- **IEC Web Site ([www.iec.ch](www.iec.ch))**

- **Catalogue of IEC publications**

  The on-line catalogue on the IEC web site (www.iec.ch/searchpub) enables you to search by a variety of criteria including text searches, technical committees and date of publication. On-line information is also available on recently issued publications, withdrawn and replaced publications, as well as corrigenda.

- **IEC Just Published**

  This summary of recently issued publications (www.iec.ch/online_news/ justpub) is also available by email. Please contact the Customer Service Centre (see below) for further information.

- **Customer Service Centre**

  If you have any questions regarding this publication or need further assistance, please contact the Customer Service Centre:

  Email: custserv@iec.ch
  Tel: +41 22 919 02 11
  Fax: +41 22 919 03 00

# INTERNATIONAL STANDARD

# IEC 62365

First edition
2004-11

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





International Electrotechnical Commission, 3, rue de Varembé, PO Box 131, CH-1211 Geneva 20, Switzerland
Telephone: +41 22 919 02 11 Telefax: +41 22 919 03 00 E-mail: inmail@iec.ch Web: www.iec.ch

Commission Electrotechnique Internationale
International Electrotechnical Commission
Международная Электротехническая Комиссия

PRICE CODE U

*For price, see current catalogue*

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.

The text of this standard is based on the following documents:

| CDV | Report on voting |
|---|---|
| 100/753/CDV | 100/838/RVC |

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

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

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*

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 65536 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 ensure interoperability between equipment designed for different applications, source equipment shall be capable of transmitting at least one of the formats indicated in Table 2, and destination equipment shall be capable of receiving all of the formats indicated in Table 2.

**Table 2 – Combinations of subframe format and packing scheme**

| **Code (NOTE 1)** | **Status** | **Subframe length** bytes | **Audio sample word length bits** | **Ancillary data bits** | **Protocol overhead bits** | **Grouping** | **Number of audio channels** |
|---|---|---|---|---|---|---|---|
| 56 02 | NOTE 2 | 4 | 24 | 4 | 4 | Temporal | 2 |
| 56 01 | NOTE 2 | 4 | 24 | 4 | 4 | Not applicable | 1 |
| 06 02 | NOTE 2 | 3 | 24 | 0 | 0 | Temporal | 2 |
| 06 01 | NOTE 2 | 3 | 24 | 0 | 0 | Not applicable | 1 |
| 56 85 | NOTE 3 | 4 | 24 | 4 | 4 | Multi-channel | 60 |

NOTE 1 The code column shows the encoding (in hexadecimals) of the second and third bytes of the AAL parameters IE as specified in 5.2.2.1 and is informative only.

NOTE 2 Required for all equipment.

NOTE 3 Required for equipment that can convey at least 56 audio channels and has sufficient network capacity to do so.

**4.3.3** When conveying 56-channel MADI data, the format with 60 channels in Table 2 shall be used, with the last four channels being unused.

**4.3.4** Sampling frequencies shall be as specified in AES5.

**4.3.5** Destination equipment shall support the 48 kHz sampling frequency.

## 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.

**4.5.3** For the first block transmitted after a clock tick, the `ATM-user-to-ATM-user indication` (`UI` bit) in the cell header shall be set to `1` in the first and last cells and to `0` in all other cells.

For all other blocks, the `ATM-user-to-ATM-user indication` shall be set to `1` in the last cell and to `0` in all other cells.

NOTE If the state of the `UI` bit is latched as each cell is unpacked, the resulting signal can be a pulse train with the leading edges of the pulses being evenly spaced with frequency $f/8n$, where $f$ is the sampling frequency and $n$ the number of samples from the same channel in a cell, and with a double-width pulse once each second.

## 5 Switched virtual circuits

### 5.1 Addresses

**5.1.1** The distinction between audio and other circuits, and between different types of ports, shall be made using protocol and other information conveyed as specified in 5.2 and 8.1.

**5.1.2** Source and destination ports may be different types; the calling party number (or subaddress) shall identify a source port, and the called party number (or subaddress) shall identify a destination port.

**5.1.3** The distinction between different ports of the same type within an interface shall be made using the `Selector` value.

**5.1.4** An audio port may carry more than one audio channel; the protocol information specifies how many channels are to be received. Ports with different numbers of channels may be considered to be of different types, and the same physical port may be addressed in more than one way; for instance, an AES3 output port may be addressed as a single stereo port, two mono ports, a single mono port using `double sampling frequency` mode, or as one of a group of three carrying a `5.1-channel` signal.

### 5.2 Audio call connection: SETUP and ADD PARTY messages

#### 5.2.1 Restrictions on connection requests

**5.2.1.1** All audio connections shall be point to multipoint and shall be originated by the equipment that contains the source port.

NOTE When connecting a new call, the caller sends a SETUP message and the destination equipment receives a SETUP message; when adding a new destination, the caller sends an ADD PARTY message and the destination equipment receives a SETUP message unless it already has a port that is a destination for that call, in which case it receives an ADD PARTY message.

**5.2.1.2** The AAL parameters, broadband high-layer information, and calling party number information elements (IE), which are optional in the ATM signalling specification, shall be required when connecting an audio call.

**5.2.1.3** If the called party number is not in the network service access point (NSAP) format conforming to Table 4-12 of ITU Q.2931, the called party subaddress IE shall be required.

If the calling party number is not in the NSAP format, the calling party subaddress IE shall be required.

### 5.2.2 Information elements in the SETUP and ADD PARTY messages

#### 5.2.2.1 ATM adaptation layer parameters IE

Within the ATM adaptation layer parameters IE, the following coding shall be used.

a) The AAL type shall be coded as user-defined AAL ($10_{16}$ in octet 5).

b) The first octet of the User Defined AAL Information (octet 6) shall encode the qualifying information specified in 6.1.

c) The second octet of the User Defined AAL Information (octet 6.1) shall encode the subframe format as specified in 6.2.

d) The third octet of the User Defined AAL Information (octet 6.2) shall encode the packing of subframes into cells as specified in 6.3.

e) The fourth octet of the User Defined AAL Information (octet 6.3) shall encode the sampling frequency as specified in 6.4.

#### 5.2.2.2 Broadband high-layer information IE

Within the broadband high-layer information IE, the following coding shall be used.

a) The high-layer information type shall be coded as vendor-specific application identifier ($83_{16}$ in octet 5).

b) The OUI value shall be coded as $00_{16}$ in octet 6, $0B_{16}$ in octet 7, and $5E_{16}$ in octet 8.

c) The first octet of the application identifier (octet 9) shall be coded as zero to indicate the first edition of this standard.

d) The second octet of the application identifier (octet 10) shall be coded as zero to indicate audio data.

The remaining octets of the application identifier (octets 11 to 12) are reserved. They shall be coded as zero but ignored by the recipient unless specified by a means specified outside this standard.

NOTE A different value for octet 10 is specified in 8.1.2.2. Encodings in the broadband high-layer information IE are summarized in Annex C.



### 5.2.3 Destination response to SETUP and ADD PARTY messages

#### 5.2.3.1 Destination response to SETUP message

**5.2.3.1.1** A SETUP message received from the network shall be processed according to the provisions of this standard for audio calls if it meets the following three criteria.

a) It contains a broadband high-layer information IE which conforms to 5.2.2.2 or does not contain a broadband high-layer information IE but is in other respects consistent with being an audio call conforming to this standard.

b) It is for a point-to-multipoint connection.

c) It contains ATM adaptation layer parameters IE indicating a user-defined AAL ($10_{16}$ in octet 5).

**5.2.3.1.2** If the first octet of the User Defined AAL Information (octet 6) does not contain an encoding recognized by the equipment, the destination equipment shall reject the call with cause value call rejected (octet 6 = $95_{16}$ in the cause IE), rejection reason user specific and condition permanent (octet 7 = $81_{16}$), and the user specific diagnostic (octet 7.1) coded as a single octet with the value zero.

**5.2.3.1.3** If the remaining octets of the user defined AAL information (octets 6.1 to 6.3) indicate a format or sampling frequency which the equipment does not support, the destination equipment shall reject the call with cause value call rejected (octet 6 = $95_{16}$), rejection reason user specific and condition permanent (octet 7 = $81_{16}$), and the first user specific diagnostic octet coded as 01.

NOTE The destination equipment should take the sampling frequency from the AAL parameters IE and not calculate it from the cell rate in the ATM traffic descriptor IE.

**5.2.3.1.4** If the Selector value, in the context of the format indicated by the ATM adaptation layer parameters IE, does not correspond to an output port or corresponds to an output port which is disabled, the destination equipment shall reject the call with cause value call rejected (octet 6 = $95_{16}$), rejection reason user specific and condition permanent if the port does not exist, transient if it is disabled (octet 7 = $81_{16}$ or $82_{16}$, respectively), and the first user specific diagnostic octet coded as $02_{16}$.

**5.2.3.1.5** If the Selector value (in the context of the format indicated by the ATM adaptation layer parameters IE) corresponds to an output port which would conflict with an existing connection, or completing the connection would require some other resource that has been used up for other calls, the destination equipment shall reject the call with cause value user busy (octet 6 = $91_{16}$).

NOTE Receiving equipment can include two cause IEs in the RELEASE message to give information on two different aspects of the reason for rejection of the call. If two cause IEs are included, the cause value specified in this subclause may be in either of them.

#### 5.2.3.2 Destination response to ADD PARTY message

**5.2.3.2.1** An ADD PARTY message received from the network shall only be processed according to this standard if it relates to an audio call conforming to this standard.

**5.2.3.2.2** The receiving equipment may process the ATM adaptation layer parameters IE in the same way as for a SETUP message (see 5.2.3.1), or it may check that it is the same as that received when the call was connected and reject the call with cause invalid information element contents (octet 6 = $E4_{16}$) citing the ATM adaptation layer parameters IE (octet 7 = $58_{16}$) if it is not.

**5.2.3.2.3** The Selector value is processed in the same way as for a SETUP message (see 5.2.3.1).

### 5.3 Call disconnection

Disconnection may be initiated by either the source or the destination equipment without giving any warning to the other party.

NOTE Disconnection can also be initiated by the network in the event of various kinds of error or if it detects that the other party has been switched off, reset, or unplugged from the network.

The cause shall be specified as normal call clearing ($90_{16}$ in octet 6 of the cause IE).

NOTE This value is not included in the table in 5.4.5.15 of the UNI 3.0 specification. There are no diagnostics associated with it.

# 6 Coding of audio formats

NOTE In this clause, bits are numbered in the same way as in the ATM UNI specifications. Thus, bit 8 is the most significant bit of the octet and bit 1 the least significant.

## 6.1 Qualifying information

| 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | See 6.1.2 | reserved | | |

**6.1.1** Bits 8 to 5 shall be zero to indicate the formats specified in this edition of this standard.

NOTE Other formats, with bits 8 to 5 not all 0, are reserved.

**6.1.2** Bit 4 shall be coded as

0 no information about the exact frequency of the sample clock is provided;

1 sample clock is frequency-locked to a global reference.

NOTE This standard does not specify to what reference the sample clock is locked if bit 4 is a 1.

**6.1.3** Bits 3 to 1 are reserved and shall be coded as zero unless otherwise specified.

## 6.2 Subframe format

| 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |
|---|---|---|---|---|---|---|---|
| Ancillary | | Overhead | | Sample length | | | |

**6.2.1** The ancillary field shall be coded as

00 no ancillary data bits;

01 4 ancillary data bits as specified in 4.1.3;

1x reserved.

**6.2.2** The overhead field shall be coded as

00 no protocol overhead bits;

01 four protocol overhead bits as specified in 4.1.4;

1x reserved.

**6.2.3** The sample length field shall be coded to indicate the number of bits in the audio sample word as

000x reserved

0010 8

0011 12

0100 16

0101 20

0110 24

0111 28

1000 32

1001 reserved

1010 40

1011 reserved

11xx reserved

### 6.3 Packing of subframes into cells

| 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |
|---|---|---|---|---|---|---|---|
| Packing | | Number of channels or cells | | | | | |

**6.3.1** The packing field shall be coded as

00 temporal grouping, bits 6 to 1 show number of audio channels;

01 grouping by channel, bits 6 to 1 show number of audio channels;

10 multi-channel, bits 6 to 1 show number of cells per sample time;

11 reserved.

NOTE In the multi-channel case, the number of samples per cell must be deduced from the subframe format and multiplied by the number in bits 6 to 1 to discover the number of channels.

**6.3.2** Where the number of channels is 1, or equal to the number of samples per cell, the coding for temporal grouping shall be used.

### 6.4 Sampling frequency

| 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |
|---|---|---|---|---|---|---|---|
| Basic | | Scale factor | | | Multiplier | | |

**6.4.1** The basic field shall be coded as

00 reserved

01 44,1 kHz

10 48 kHz

11 32 kHz

**6.4.2** The scale factor field shall be coded as

000 0,25

001 0,5

010 1

011 2

100 4

101 8

11x reserved

**6.4.3** The multiplier field shall be coded as

000 1;

001 1000/1001;

010 1001/1000;

011 varispeed: multiplier may vary between 0,875 and 1,125;

1xx reserved.

**6.4.4** The sampling frequency shall be the product of the basic frequency, the scale factor, and the multiplier encoded by the individual fields.

**6.4.5** In the absence of any explicit indication of tolerance or synchronization, destination equipment shall be designed to the tolerances specified in 5.2.2.2 of AES11.

# 7 Permanent virtual circuits

Equipment may support the use of permanent virtual circuit connections for digital audio data formatted as specified in Clause 4.

Default VPI-VCI values are specified in Table 3 for a unit with up to 128 single-channel audio ports numbered 0 to 127; these apply separately to the audio-input-to-ATM-transmission and ATM-reception-to-audio-output directions. In all cases, a subframe consists of 24 audio bits, 4 ancillary data bits as specified in 4.1.3, and 4 protocol overhead bits as specified in 4.1.4; and samples are grouped by channel as specified in 4.2.3.

**Table 3 – Default port number and packing for certain VCIs**

| VPI | VCI range | Number of audio channels | Port number(s) ($n$ = VCI) |
|---|---|---|---|
| 0 | 128 to 254 | 2 | $n$-128, $n$-127 |
| 0 | 256 to 383 | 1 | $n$-256 |
| 0 | 384 to 500 | 12 | $n$-384 to $n$-373 |
| 0 | 512 to 634 | 6 | $n$-512 to $n$-507 |

How VPI-VCI values are associated with audio ports other than those in Table 3, and formats other than that specified above, is outside the scope of this standard. However, the mechanisms described in Clause 8 should be used as far as is appropriate.

# 8 Management interface

## 8.1 Call connection: SETUP messages

### 8.1.1 Restrictions on connection requests

**8.1.1.1** All management connections shall be point-to-point and shall be originated by the controlling entity.

**8.1.1.2** The AAL parameters, broadband high-layer information, and calling party number IEs, which are optional in the ATM signalling specification, shall be required when connecting a management call.

**8.1.1.3** If the called party number is not in the NSAP format, the called party subaddress IE shall be required.

If the calling party number is not in the NSAP format, the calling party subaddress IE shall be required.

**8.1.1.4** No option shall be selected which adds requirements to the called party's process of sending and receiving messages. Such requirements would include traffic shaping and support for available bit rate connections.

### 8.1.2 Information elements in the SETUP message

#### 8.1.2.1 ATM adaptation layer parameters IE

Within the ATM adaptation layer parameters IE, the following coding shall be used.

The AAL type shall be coded as AAL type 5 ($05_{16}$ in octet 5).

If the service specific convergence sublayer (SSCS) type, as specified in octet group 8 in part 4 of Figure 4-12 of ITU-T Q.2931 (02/95), is present, it shall be encoded as null (0 in octet 8.1 of the same ITU citation).

#### 8.1.2.2 Broadband high-layer information IE

Within the broadband high layer information IE, the following coding shall be used.

a) The high-layer information type shall be coded as a vendor-specific application identifier ($83_{16}$ in octet 5).

b) The OUI value shall be coded as $00_{16}$ in octet 6, $0B_{16}$ in octet 7, and $5E_{16}$ in octet 8.

c) The first octet of the application identifier (octet 9) shall be coded as zero to indicate the first edition of this standard.

d) The second octet of the application identifier (octet 10) shall be coded as 1 to indicate a management connection.

e) The remaining octets of the application identifier (octets 11 to 12) are reserved; they shall be coded as zero but ignored by the recipient unless specified outside this standard.

NOTE A different value for octet 10 is specified in 5.2.2.2. Encodings in the broadband high-layer information IE are summarized in Annex B.

### 8.1.3 Destination response to SETUP message

**8.1.3.1** A SETUP message received from the network shall be processed conforming to the specifications of this standard for management calls if it meets the following three criteria.

a) It contains a broadband high-layer information IE which conforms to 8.1.2.2 or does not contain a broadband high-layer information IE but is in other respects consistent with being a management call conformant with this standard.

b) It is for a point-to-point connection.

c) It contains an ATM adaptation layer parameters IE indicating AAL type 5 ($5_{16}$ in octet 5).

**8.1.3.2** If the Selector value (in the context of the format indicated by the ATM adaptation layer parameters IE) does not correspond to a management port, the destination equipment shall reject the call with cause value call rejected (octet 6 = $95_{16}$), rejection reason user specific, and condition permanent (octet 7 = $81_{16}$), and the first user specific diagnostic octet coded as $02_{16}$.

**8.1.3.3** `Selector` value zero shall always correspond to a management port.

NOTE This value ensures that a management connection can be made to a device about which nothing is known except its ATM address; however, if another controlling entity is already connected, the call may be rejected with cause value `user busy`.

If the `Selector` value corresponds to a management port which would conflict with an existing connection, or completing the connection would require some other resource that has been used up for other calls, the destination equipment shall reject the call with cause value `user busy` (octet 6 = $91_{16}$).

NOTE 1 The destination equipment can be assumed to have one or more management ports, each of which can accept one call at a time; it need not associate a different `Selector` value with each port and may associate all ports with `Selector` value zero (or simply ignore the `Selector` value) and connect an incoming call to the first management port that is free at the time.

NOTE 2 Receiving equipment can include two cause `IE`s in the `RELEASE` message to give information on two different aspects of the reason for rejection of the call. If two cause `IE`s are included, the cause value specified in this subclause may be in either of them.

## 8.2 Message encapsulation

Each message on a management connection shall be transmitted as an `AAL-SDU` using the service specified in ITU-T Recommendation I.363.5.

## 8.3 Message format and action to be taken by recipient

**8.3.1** The first octet of each message shall be a message type code as specified in 8.4. Remaining octets shall be coded according to the specification for the message type.

**8.3.2** Equipment that has accepted a management call as specified in 8.1 shall recognize message types 1 to 3 as specified in 8.4.1 received on that call and take the action specified in 8.4.1. It may recognize and act upon other message types not specified in this standard. It shall ignore any message whose type it does not recognize.

**8.3.3** A controlling entity that has made a management call as specified in 8.1 shall recognize message type 129 as specified in 8.4.2 received on that call and take the action specified in 8.4.2. It may recognize and act upon other message types not specified in this standard. It shall ignore any message whose type it does not recognize.

**8.3.4** Message type 0 (see 8.4.3) should be used for all messages that are specific to proprietary equipment. Other message types are reserved.

**8.3.5** `AAL-SDU`s with length zero shall be ignored. If an `AAL-SDU` is shorter than the specified length for the message it contains, the recipient shall add sufficient octets containing the value zero to the end of the message to bring it up to the required length. If an `AAL-SDU` is longer than the specified length for the message it contains, the recipient shall ignore the excess octets.

## 8.4 Message types

The contents for each octet or group of octets in the message shall be according to the appropriate table in this subclause, the octets being numbered from zero.

In the coding columns in the tables, = followed by a value indicates that the value shall be fixed for this message type. Otherwise, this column indicates the subclause or subclauses where the coding is specified.

A reply shall be sent on the same management connection on which the message to which it is a reply was received.

### 8.4.1 Messages sent from the controlling entity

#### 8.4.1.1 Status enquiry

**Table 4 – `Status enquiry` message**

| Octet(s) | Contents | Coding reference (see 8.4) |
|---|---|---|
| 0 | Message type | `= 1` |
| 1 | Subsystem code | 8.4.1.1 |
| 2 | Selector | 5.1 |
| 3 | Subframe format | 6.2 |
| 4 | Packing | 6.3 |
| 5 | Sampling frequency | 6.4 |

The subsystem code shall be coded as

`0` input port;

`1` output port;

2 to 255 reserved.

In the cases of subsystem codes 2 to 255, only octets 0 and 1 shall be defined by this standard.

On receipt of the `status inquiry` message, if no audio port fits the data in octets 1 to 5, the recipient shall send, in reply, a status message for the data in octets 1 to 5, otherwise it shall send, in reply, a status message for the audio port which best fits the data in octets 1 to 5.

#### 8.4.1.2 Audio connection request

**Table 5 – Audio connection request message**

| Octet(s) | Contents | Coding reference (see 8.4) |
|---|---|---|
| 0 | Message type | = 2 |
| 1 | Input port selector | 5.1 |
| 2 | Subframe format | 6.2 |
| 3 | Packing | 6.3 |
| 4 | Sampling frequency | 6.4 |
| 5 to 6 | Endpoint reference | 8.4.1.2 |
| 7 to 26 | Destination address | 5.1 |

**8.4.1.2.1** The endpoint reference shall be an integer in the range 1 to 32767 inclusive, which is different for each destination of a call.

**8.4.1.2.2** The destination address shall use the NSAP format.

NOTE This message format does not support the use of an ITU Recommendation E.164 address for the called party address together with a called party subaddress.

**8.4.1.2.3** On receipt of this message, the recipient shall

a) check that octets 1 to 4 identify an input port which can be, or already is, the source of a call with the parameters in octets 2 to 4; if so, proceed to step (b) or (c), otherwise send an input port status message in reply;

b) if the port is already the source of a call, send an `ADD PARTY` message to the network for that call, using the endpoint reference value from octets 5 to 6;

c) if the port is not the source of any call, send a `SETUP` message to the network, and store the endpoint reference value from octets 5 to 6 as the port's `first endpoint reference`.

#### 8.4.1.3 Audio disconnection request

**Table 6 – Audio disconnection request message**

| Octet(s) | Contents | Coding reference (see 8.4) |
|---|---|---|
| 0 | Message type | = 3 |
| 1 | Input port selector | 5.1 |
| 2 | Subframe format | 6.2 |
| 3 | Packing | 6.3 |
| 4 | Sampling frequency | 6.4 |
| 5 to 6 | Endpoint reference | 8.4.1.3 |

**8.4.1.3.1** The endpoint reference value shall be the same as that used when making the connection to the destination that is to be disconnected, or zero which is equivalent to the `first endpoint reference`, or $FFFF_{16}$ to disconnect all destinations.

**8.4.1.3.2** On receipt of this message, the recipient shall

a) check that octets 1 to 4 identify an input port which is the source of a call with the parameters in octets 2 to 4; if so, proceed to step (b), otherwise send an input port status message in reply;

b) if the endpoint reference value in octets 5 and 6 is $FFFF_{16}$, send a `RELEASE` message to the network for that call, otherwise send a `DROP PARTY` message to the network for that call, using the endpoint reference value from octets 5 to 6 unless it is the same as the port's `first endpoint reference` in which case the endpoint reference value, 0, shall be used.

### 8.4.2 Messages sent to the controlling entity

#### 8.4.2.1 Input port status

The `input port status` message shall be sent in reply to a `status enquiry` message, or (unsolicited) to indicate a change in status of an input port.

**Table 7 – Input port status message**

| Octet(s) | Contents | Coding reference (see 8.4) |
|---|---|---|
| 0 | Message type | = 129 |
| 1 | Subsystem code | = 0 |
| 2 | Selector | 5.1 |
| 3 | Input port status | 8.4.2.1 |
| 4 | Subframe format | 6.2 |
| 5 | Packing | 6.3 |
| 6 | Sampling frequency | 6.4 |
| 7 to 8 | First endpoint reference | 8.4.1.2 and 8.4.2.1 |

**8.4.2.1.1** The input port status in octet 3 shall be coded as follows, all values being in hexadecimal:

00 no port with the Selector value in octet 2 exists; octets 4 to 6 are copied from the status enquiry message to which this is a reply;

40 the port identified by octets 2 and 4 to 6 is disabled;

81 to 87 a port with the Selector value in octet 2 exists but does not support the requested format;

C0 the port identified by octets 2 and 4 to 6 is enabled but no audio signal is present;

E0 the port identified by octets 2 and 4 to 6 is enabled and an audio signal with the sampling frequency indicated by octet 6 is present.

Other code points shall be reserved; future definitions of these code points may specify a different format for octets 4 onwards. The recipient shall ignore an input port status message in which it does not recognize the value in octet 3.

In the case of code point 00, octets 4 to 6 shall be copied from the status enquiry message to which the input port status message is a reply.

In the case of code points 81 to 87, octets 4 to 6 shall be copied from the status enquiry message to which the input port status message is a reply, and the aspects of the format that are not supported shall be bitwise encoded in octet 3 as follows:

| 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | F | P | R |

F = 1 if the subframe format indicated by octet 4 is not supported, 0 else;

P = 1 if the packing format indicated by octet 5 is not supported, 0 else;

R = 1 if the sampling frequency indicated by octet 6 is not supported, 0 else.

In the case of code points 40, C0, and E0, octets 4 to 6 shall indicate how the port is currently configured.

**8.4.2.1.2** The first endpoint reference in octets 7 to 8 shall be non-zero if the port is the source of an audio call on the ATM network, and 0 otherwise.