INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

H.222.1

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

(03/96)

TRANSMISSION OF NON-TELEPHONE SIGNALS

MULTIMEDIA MULTIPLEX AND SYNCHRONIZATION FOR AUDIOVISUAL COMMUNICATION IN ATM ENVIRONMENTS

ITU-T Recommendation H.222.1

(Previously “CCITT Recommendation”)

FOREWORD

The ITU-T (Telecommunication Standardization Sector) is a permanent organ of the International Telecommunication Union (ITU). The ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Conference (WTSC), which meets every four years, establishes the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

The approval of Recommendations by the Members of the ITU-T is covered by the procedure laid down in WTSC Resolution No. 1 (Helsinki, March 1-12, 1993).

ITU-T Recommendation H.222.1 was prepared by ITU-T Study Group 15 (1993-1996) and was approved under the WTSC Resolution No. 1 procedure on the 19th of March 1996.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

CONTENTS

SUMMARY

This Recommendation describes the multiplexing and synchronization of multimedia information for audiovisual communications in ATM environments. It specifies the H.222.1 Program Stream and the H.222.1 Transport Stream by choosing necessary coding elements from the generic H.222.0 specifications and by adding items for use in ATM environments. It also covers methods for accommodating ITU-T defined elementary streams.

MULTIMEDIA MULTIPLEX AND SYNCHRONIZATION FOR AUDIOVISUAL COMMUNICATION IN ATM ENVIRONMENTS

(Geneva, 1996)

1 Scope

This Recommendation describes the multiplexing and synchronization of multimedia information, for audiovisual communication in ATM environments. This Recommendation specifies the peer-to-peer syntax, semantics and procedures, and the interactions with the AAL.

2 References

The following Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; all users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published.

[1] CCITT Recommendation G.711 (1988), *Pulse Code Modulation (PCM) of voice frequencies*.

[2] CCITT Recommendation G.722 (1988), *7 kHz audio-coding within 64 kbit/s*.

[3] ITU-T Recommendation G.723.1 (1996), *Dual rate speech coder for multimedia communication transmitting at 5.3 and 6.3 kbit/s*.

[4] CCITT Recommendation G.728 (1992), *Coding of speech at 16 kbit/s using low-delay code excited linear prediction*.

[5] ITU-T Recommendation H.222.0 (1995) | ISO/IEC 13818-1:(1996), *Information technology – Generic coding of moving pictures and associated audio information: Systems*.

[6] ITU-T Recommendation H.233 (1995), *Confidentiality system for audiovisual services*.

[7] ITU-T Recommendation H.245 (1996), *Control protocol for multimedia communication*.

[8] ITU-T Recommendation H.261 (1993), *Video codec for audiovisual services at p × 64 kbit/s*.

[9] ITU-T Recommendation H.262 (1995) | ISO/IEC 13818-2:(1995), *Information technology – Generic coding of moving pictures and associated audio information: Video*.

[10] ITU-T Recommendation H.263 (1996), *Video coding for low bit rate communication*.

[11] ITU-T Recommendation I.311 (1993), *B-ISDN general network aspects*.

[12] ITU-T Recommendation I.361 (1995), *B-ISDN ATM layer specification*.

[13] ITU-T Recommendation I.362 (1993), *B-ISDN ATM Adaptation Layer (AAL) functional description*.

[14] ITU-T Recommendation I.363 (1993), *B-ISDN ATM Adaptation Layer (AAL) specification*.

[15] ITU-T Recommendation T.120 (1996), *Data protocols for multimedia conferencing*.

[16] ITU-T Recommendation Q.2931 (1995), *Broadband Integrated Services Digital Network (B-ISDN) – Digital Subscriber Signalling System No. 2 (DSS 2) – User-Network Interface (UNI) layer 3 specification for basic call connection control*.

3 Terms and definitions

For the purposes of this Recommendation, the following definition applies:

3.1 subchannel: A logical channel in Recommendation H.222.1 formed from packets having a unique multiplex identifier field value. A subchannel carries one H.222.1 elementary stream. A subchannel is unidirectional. There may be many subchannels within one ATM Virtual Channel.

4 Abbreviations

For the purposes of this Recommendation the following abbreviations are used:

AAL	ATM Adaptation Layer
ACELP	Algebraic Codebook Excited Linear Prediction
ATM	Asynchronous Transfer Mode
CA	Conditional Access
CDV	Cell Delay Variation
CPCS	Common Part Convergence Sublayer
DP	Data Partitioning
DTS	Decoding Time Stamp
EOB	End_of_Block
GOP	Group of Picture
LD-CELP	Low-Delay Codebook Excited Linear Prediction
MBA	Macro Block Address
MP-MLQ	Multi-Pulse Maximum Likelihood Quantization
PCR	Program Clock Reference
PDU	Protocol Data Unit
PES	Packetized Elementary Stream
PS	Program Stream
PSI	Program Specific Information
PSM	Program Stream Map
PTS	Presentation Time Stamp
QOS	Quality of Service
RTI	Real Time Interface
SCR	System Clock Reference
SDU	Service Data Unit
STC	System Time Clock
STD	System Target Decoder
TS	Transport Stream

5 General

This Recommendation deals with the multiplexing and synchronization of multiple multimedia signals, for use in audiovisual communications in ATM environments. The multimedia signals may be coded audio or video or other data signals.

This Recommendation is suitable for various applications such as conversational services, distributive services, retrieval services and messaging services.

This Recommendation is applicable to both unidirectional and bidirectional physical connections. A bi-directional connection may be symmetrical or asymmetrical.

This Recommendation addresses constant bit rate coded video. No specific support is provided for variable bit rate coded video. Variable bit rate coded video is however not precluded.

This Recommendation uses the services provided by the AAL. The AAL is specified in Recommendation I.363 [14]. The use of AAL type 1 and type 5 is specified here.

This Recommendation may also be suitable for use in environments other than ATM.

This Recommendation specifies two separate and independent protocols. They are:

- H.222.1 Program Stream.
- H.222.1 Transport Stream.

These two protocols are based upon the Program Stream and Transport Stream respectively, defined in Recommendation H.222.0 [5]. This Recommendation specifies the necessary coding elements in H.222.0 plus additional functionalities to form the H.222.1 Program Stream protocol and the H.222.1 Transport Stream protocol. An overview of H.222.1 and its relation to Recommendation H.222.0 is shown in Figure 1.

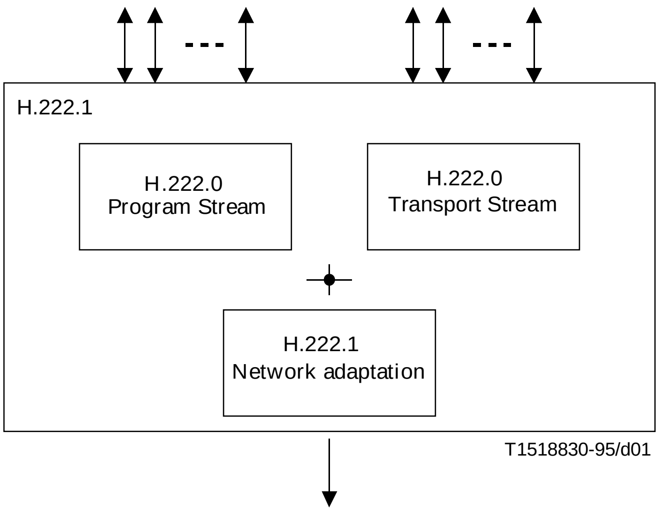

FIGURE 1/H.222.1

H.222.1 overview

6 Services provided by Recommendation H.222.1

The following services are provided by H.222.1 to the H.222.1 user:

a) *Multiplexing*

Multiplexing is based on a sequence of PDUs, each of which carry consecutive data from one and only one media source type, i.e. audio, video or other data signals. In the H.222.1 Program Stream these PDUs may be of variable length and relatively large in size. In the H.222.1 Transport Stream these PDUs are of fixed length and of relatively small size. The H.222.1 Transport Stream has a large multiplex capacity.

b) *Timebase recovery*

A program is a collection of associated media, all of which refer to a common timebase. The timebase is referred to as the System Time Clock (STC). The H.222.1 Program Stream supports one and only one program. The H.222.0 Transport Stream supports multiple programs. H.222.1, however, constrains the number of programs in the Transport Stream to be one.

The send side and receive side each have their own timebases. Time stamps attached to specific PDUs identify the intended time of arrival of the PDU at the receive side. Synchronization of the receive side timebase with the send side timebase may be achieved using these time stamps. H.222.1 provides optional additional timebase recovery information, which may be useful in jittered environments.

c) *Media presentation synchronization*

Additional time stamps identify times at which entities in each media are to be presented to the end user.

d) *Timing jitter removal*

Recommendation H.222.1 may offer the capabilities to remove the effects of time delay variation on the encoded bit stream at the receiver.

e) *Buffer management*

Rules are specified so as to avoid underflow and overflow of receive side buffers. This is achieved by a hypothetical receive side timing model, which specifies timing relationships between outgoing PDUs at the send side.

f) *Security and access control*

Security and access control services are provided by media encryption. Support for entitlement control and management messages is also provided.

g) *In-band signalling*

The multiplexing service provides multiple connection end points at the user/H.222.1 service boundary. Protocol is provided that signals to the receive side the association between a PDU and a connection end point. The nature of the information carried by the connection is also described.

h) *Error reporting*

Protocol at the receive side reports error conditions to the H.222.1 user.

i) *Trick mode*

Mechanisms to support video recorder like control functionality, e.g. fast forward, rewind etc. are included.

j) *Network maintenance*

A network maintenance service, which monitors channel errors, is available.

k) *Remultiplex support*

In the H.222.1 Transport Stream mechanisms to assist in the addition and removal of individual elementary streams are provided. This service only has meaning at a network element. This service is not explicitly supported in the H.222.1 Program Stream.

l) *Local program insertion*

In the H.222.1 Transport Stream mechanisms are provided to assist in the replacement of one elementary stream with another elementary stream. This service only has meaning at a network element. This service is not explicitly supported in the H.222.1 Program Stream.

m) *Priority*

In the H.222.1 Transport Stream, one of two priorities may be indicated for each PDU. This service only has meaning at a network element. This service is not explicitly supported in the H.222.1 Program Stream.

It is not stated here as to which services are mandatory. An application decides what services are appropriate.

7 H.222.1 Program Stream

7.1 H.222.1 Program Stream syntax and semantics

The H.222.0 Program Stream syntax and semantics apply, as defined in 2.5.3/H.222.0.

The H.222.0 Program Stream Map is defined in 2.5.4/H.222.0. See clause 11 for details of H.222.1 usage of the H.222.0 PSM.

The H.222.0 Program Stream Directory is defined in 2.5.5/H.222.0. Use of the H.222.0 Program Stream Directory is optional in the H.222.1 Program Stream.

7.2 H.222.1 Program Stream timing model

The H.222.0 Program Stream System Target Decoder as defined in 2.5.2/H.222.0 applies.

8 H.222.1 Transport Stream

The H.222.0 Transport Stream supports multiple programs, where a program is as defined in Recommendation H.222.0. Recommendation H.222.1 supports only single program Transport Streams. Multiple program Transport Streams are not supported.

8.1 H.222.1 Transport Stream syntax and semantics

The H.222.0 Transport Stream syntax and semantics apply, as defined in 2.4.3/H.222.0.

The H.222.0 Program Specific Information (PSI) is defined in 2.4.4/H.222.0. See clause 11 for details of H.222.1 usage of the H.222.0 PSI.

8.2 H.222.1 Transport Stream timing model

The Transport Stream System Target Decoder as defined in 2.4.2/H.222.0 applies.

9 H.222.1 Network adaptation functions

Recommendation H.222.1 may offer the capabilities to remove the effects of time-delay variation on the encoded bit stream at the receiver. No specific syntax is provided. The actual mechanism used is application dependent.

NOTE – The Cell Delay Variation (CDV) at the input of the H.222.1 decoder depends on the network size, the volume and the characteristics of accommodated traffic, amongst other factors. It is assumed that the CDV value may be of the order of about 1-3 ms for CBR traffic at 156 Mbit/s data rate. If the H.222.0 decoder implementation can only accept CDV less than the above value at its input, then the H.222.1 network adaptation should provide the necessary CDV reduction. Allocation of the allowable CDV to the H.222.0 decoder and the H.222.1 network adaptation as well as the method to reduce the CDV in the H.222.1 network adaptation are left to the choice of implementation.

10 Interaction with the AAL

The Program Stream and the Transport Stream may use the services of AAL type 1 and AAL type 5.

10.1 AAL type 1

The Program Stream and the Transport Stream may use the service provided by AAL type 1 at the AAL-SAP. Two primitives are used between the AAL and the AAL user being AAL-UNITDATA-REQUEST and AAL-UNITDATA-INDICATION. The AAL-SDU length may be one bit or one byte according to the application.

The AAL type 1 Convergence Sublayer offers a number of functions to deal with cell delay variation, lost and misinserted cells, timing relationship and correction of bit errors and lost cells.

It is not specified here as to which AAL type 1 functions should be used. This is left to the application to specify.

10.2 AAL type 5

The Program Stream and the Transport Stream may use the services provided by the AAL type 5 Common Part Convergence Sublayer. Two signals are used between AAL type 5 CPCS and the AAL type 5 CPCS user, being CPCS-UNITDATA invoke and CPCS-UNITDATA signal.

Recommendation H.222.1 uses the error detection function offered by the AAL type 5 CPCS layer. This, in combination with simple error concealment in the video decoder, should provide sufficient error resilience.

It is an implementation option as to whether a corrupted AAL type 5 CPCS-PDU is discarded, or whether a corrupted AAL type 5 CPCS-PDU is passed to H.222.1.

10.2.1 AAL type 5 for one way audiovisual transmission

The Transport Stream

H.222.1 Transport Stream packets shall be mapped into AAL type 5 with a NULL Service Specific Convergence Sublayer. In the mapping, one to N Transport Streams packets are mapped into an AAL type 5 DDU.

For Switched Virtual Circuits (SVCs), the value of N is established via ATM user-to-network call/connection control at call setup using the AAL type 5 maximum CPCS-SDU size negotiation procedure. The AAL type 5 maximum CPCS-SDU size that is signalled is N ∗ 188 bytes, where N is the number of TS packets in the AAL type 5 SDU. This procedure is defined in Recommendation Q.2931 [16].

For Permanent Virtual Circuits (PVCs), the default value of N is two (maximum CPCS-SDU size = 376 bytes). Other values of N may be selected by bilateral agreement via network provisioning.

Furthermore, in order to ensure a base level of interoperability, all equipment shall support N = 2 (CPCS-SDU size = 376 bytes).

In summary, the mapping shall be:

- Each AAL type 5 CPCS-SDU shall contain N Transport Stream packets, unless there are fewer than N packets left in the Transport Stream, in which case the final CPCS-SDU contains all of the remaining packets.
- The value of N is established via ATM signalling and is equal to the AAL type 5 CPCS-SDU size divided by 188. The default AAL type 5 CPCS-SDU size is 376 octets, which is two Transport Stream packets (N = 2).
- In order to ensure a base level of interoperability, all equipment shall support the value N = 2 (AAL type 5 CPCS-SDU size of 376 octets).

Specifically when N = 2, the mapping shall be as follows:

- An AAL type 5 PDU shall contain two Transport Stream packets unless it contains the last Transport Stream packet of the Transport Stream, in which case an AAL type 5 PDU shall contain one or two Transport Stream packets.

When an AAL type 5 PDU contains two Transport Stream packets, which have a length of 188 octets, the AAL type 5 CPCS-SDU has a length of 376 octets. This AAL type 5 CPCS-SDU, together with the CPCS-PDU trailer of 8 octets, requires 384 octets and maps into eight ATM cells with zero CPCS padding octets. This is illustrated in Figure 2.

The Program Stream

The mapping method for Program Streams is for further study.

10.2.2 AAL type 5 for two way audiovisual communications

The method described in 10.2.1 may be used. Other methods are for further study.

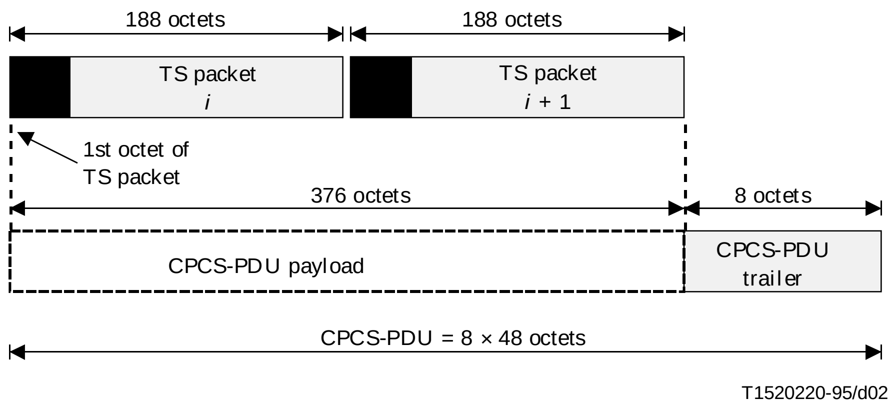

FIGURE 2/H.222.1

Representation of AAL type 5 CPCS-PDU containing two Transport Stream packets

11 Subchannel signalling

Subchannel signalling is the process by which a subchannel is established and released between peer send and receive H.222.1 entities. Subchannel signalling involves management of a unique multiplex identifier value at the send side and transmission of this value, together with information about the audiovisual data to be carried within the subchannel, to the receiver.

Subchannel signalling may use acknowledged or unacknowledged procedures. In addition subchannel signalling may be performed in-band or out-of-band with respect to the H.222.1 multiplex.

A unidirectional audiovisual call, with no return signalling channel, may only use unacknowledged procedures.

A unidirectional or bidirectional audiovisual call, where a return signalling channel exists, may use unacknowledged or acknowledged procedures. Acknowledged signalling procedures offer improved call phase synchronization.

11.1 In-band subchannel signalling

Unacknowledged in-band subchannel signalling in the H.222.1 Program Stream is coded using the H.222.0 Program Stream Map, as defined in 2.5.4/H.222.0.

Unacknowledged in-band subchannel signalling in the H.222.1 Transport Stream is coded using the Program Specific Information (PSI), as defined in 2.4.4/H.222.0.

Subchannel signalling using unacknowledged procedures achieves reliable operation through the repeated transmission of signalling information. The frequency of repetition of signalling information is not specified here.

No in-band acknowledged procedures are specified in Recommendation H.222.0 or H.222.1.

11.2 Out-of-band subchannel signalling

An out-of-band channel may have the advantage of being able to offer a reliable transport service for subchannel signalling.

When out-of-band subchannel signalling is used, it is not mandatory to code in-band signalling information, i.e. Program Stream Map/Program Specific Information.

However, all of the signalling information that would have otherwise been coded in-band in the PSM/PSI tables must be made available to the H.222.1 decoder through the out-of-band channel.

11.3 Default subchannels

Default subchannels are defined as shown in Table 1. These subchannels are assumed to be established as soon as call establishment has been completed. All applications must recognize these subchannel numbers.

TABLE 1/H.222.1

Default subchannels

Application	Program Stream		Transport Stream PID
	stream_id	stream_id_extension	
Rec. H.245	ITU-T Rec. H.222.1 type C	0x10	0x00010
G.711 A-law audio	ITU-T Rec. H.222.1 type B	0x10	0x00011
G.711 μ-law audio	ITU-T Rec. H.222.1 type B	0x20	0x00012

During the call the default subchannels may be released using subchannel signalling procedures. The subchannel numbers may then be reused for other purposes.

12 H.222.1 stream_id

For type A-E stream_ids defined in Table 2-18/H.222.0, in both Program Streams and Transport Streams, the first byte of the PES packet payload is the stream_id_extension field, as shown in Table 2. In Table 2, PES_packet_payload() are the bytes following the PES packet header, where PES packet header is as defined in Recommendation H.222.0. N is the number of bytes in the PES packet payload.

These stream_ids have an associated stream_type value of 0x09 as indicated in Table 2-29/H.222.0.

TABLE 2/H.222.1

stream_id_extension

Syntax	No. of bits	Identifier
PES_packet_payload() {		
stream_id_extension	**8**	**uimsbf**
for(i=0;i<N–1;i++){		
PES_packet_data_byte	**8**	**bslbf**
}		
}		

Semantics of PES_packet_payload() fields

stream_id_extension: The stream_id_extension field is coded according to the stream_id value, as shown in Table 3.

TABLE 3/H.222.1

Coding of stream_id_extension field

stream_id	Class	stream_id_extension: $b_7b_6b_5b_4$ $b_3b_2b_1b_0$
ITU-T Rec. H.222.1 type A	video	$b_7b_6b_5b_4$ – coded as per Table 8 (Note 2) $b_3b_2b_1b_0$ – stream number (Note 1)
ITU-T Rec. H.222.1 type B	audio	$b_7b_6b_5b_4$ – coded as per Table 11 (Note 2) $b_3b_2b_1b_0$ – stream number
ITU-T Rec. H.222.1 type C	data	$b_7b_6b_5b_4$ – coded as per Table 13 (Note 2) $b_3b_2b_1b_0$ – stream number
ITU-T Rec. H.222.1 type D	any	$b_7b_6b_5b_4$ $b_3b_2b_1b_0$ – stream number
ITU-T Rec. H.222.1 type E	–	reserved

NOTES

1 The stream number is allocated using the stream id extension descriptor (see 14.2).

2 It is the four lower order of the bits of the values defined in Tables 8, 11 and 13, which are mapped to $b_7b_6b_5b_4$, when the stream_id indicates types A, B and C, respectively.

13 Use of multiple ATM Virtual Channel Connections

13.1 General

This clause addresses applications which use coded audiovisual data in multiple ATM Virtual Channel Connections (VCC) in one call[1]. When this is the case the following conditions apply:

- for coded audiovisual data, there is one instance of H.222.1 per VCC. This means that for coded audiovisual data there is one Program Stream or Transport Stream per VCC. There may be additional VCCs which are part of the call, but which do not require use of the Program Stream or Transport Stream;
- a call shall consist of only one H.222.0 defined program in each communication direction, i.e. all Program Streams/Transport Streams used in that call refer to one STC, for each direction of communication;
- it is mandatory to code SCR/PCR in at least one Program Stream/Transport Stream respectively used in the call. It is optional as to whether SCR/PCR is coded in more than one Program Stream/Transport Stream respectively;

  the notion here is that since there is only one STC in the decoder, there is redundancy in sending SCR/PCRs on multiple VCCs;
- where in-band signalling using PSM/PSI tables is used, then PSM/PSI tables must be coded and must be valid, in all Program Streams/Transport Streams used in the call.

  Where out-of-band signalling is provided PSM/PSI tables are optional in each Program Stream/Transport Stream as stated in 11.2.

[1] Virtual Channel Connection is defined in Recommendation I.311.

13.2 Hierarchical coded video and multiple ATM Virtual Channel Connections

Some applications may require that the hierarchical layers of the scalability profiles in Recommendation H.262 be mapped to separate virtual channels. This allows:

- the video layer to be routed only where it is required; and
- to select a network Quality of Service (QOS) appropriate for each layer.

In order to improve error resilience, especially during severe error bursts, the video information may be optionally encoded according to a modified and restricted usage of the Data Partitioning (DP) specification of Recommendation H.262.

In this usage, the base layer is not a DP video stream, it does not contain the DP sequence_scalable_extension() and does not contain priority break points. The base layer contains a profile_and_level id specified in Recommendation H.262. All End_of_Blocks (EOBs) are in the base layer, as usual.

The (modified) DP enhancement layer consists only of Sequence, GOP, Picture and Slice headers that are an exact copy of the headers in the base layer.

In this modified usage, the Hierarchy descriptor, defined in 2.6.6/H.222.0, must be transmitted. It defines both base and enhancement layers using the profile_and_level_id of the base layer. The enhancement layer is specified as hierarchy_type DP.

As usual, decoders that cannot utilize enhancement layers should discard them. Decoders that utilize this modified DP can identify the enhancement layer by its designation as hierarchy_type DP, but with a non DP Profile.

14 Descriptors

14.1 ITU-T Recommendation H.222.0 descriptor priority

This clause defines priority rules for an H.222.1 decoder in the case of semantic conflict of Recommendation H.222.0 descriptors.

Appendix I recommends H.222.1 encoder rules for H.222.0 descriptors. When applied these rules result in bit streams that avoid semantic conflict.

H.222.0 program level descriptors and elementary stream level descriptors may sometimes be in conflict. To deal with this conflict four priority rules are defined. These rules are as follows:

a) The descriptor at the elementary stream level overrides a descriptor with similar meaning at the program level. However, the descriptor at the program level applies to all of the elementary streams of the program (descriptors that are not relevant are ignored).

b) The descriptor used at the program level is applied only to that program and has no meaning for an individual elementary stream. Similarly, an elementary stream descriptor has meaning only for that elementary stream and has no meaning at the program level.

c) The descriptor is ignored when it is present at the program level. It is applied to the associated elementary stream only when it is present at the elementary stream level.

d) The descriptor is ignored when it is present at the elementary stream level. It is applied to the associated program only when it is present in program level.

In general programs and elementary streams descriptors are ignored where the meaning of that descriptor conflicts with some other semantic, e.g. an audio descriptor when contained in an elementary stream with a video stream_type value.

The priority rules for the H.222.0 defined descriptors are shown in Table 4.

TABLE 4/H.222.1

H.222.1 decoder priority rules for H.222.0 defined descriptors

Tag no.	Descriptor	Priority rule
2	video_stream_descriptor	A
3	audio_stream_descriptor	A
4	hierarchy_descriptor	C
5	registration_descriptor	NS
6	data_stream_alignment_descriptor	A
7	target_background_grid_descriptor	A
8	video_window_descriptor	A
9	CA_descriptor	NS
10	ISO_639_language_descriptor	A
11	system_clock_descriptor	D
12	multiplex_buffer_utilization_descriptor	A
13	copyright_descriptor	B
14	maximum_bitrate_descriptor	NS
15	private_data_indicator_descriptor	NS
16	smoothing_buffer_descriptor	B (Note)
17	STD_descriptor	A
18	IBP_descriptor	A

NS Not specified in this Recommendation

NOTE – When the descriptor at the program level conflicts with a descriptor of similar meaning at the elementary stream level, the descriptor at the program level takes precedence.

14.2 ITU-T H.222.1 descriptors

Table 5 defines the descriptor_tag values for descriptors defined in this Recommendation.

TABLE 5/H.222.1

H.222.1 descriptor_tag values

descriptor_tag	Identification
0-63	ITU-T Rec. H.222.0 defined
64	reserved
65	ITU-T_video_stream_descriptor
66	ITU-T_audio_stream_descriptor
67	ITU-T_data_stream_descriptor
68	stream_id_extension_descriptor
69	ITU-T_timing_descriptor
70-255	reserved

The allowed application of the H.222.1 descriptors is specified in Table 6.

TABLE 6/H.222.1

Allowed use of H.222.1 descriptors

Identification	Program level	Elementary stream level
ITU-T_video_stream_descriptor	.	X
ITU-T_audio_stream_descriptor	.	X
ITU-T_data_stream_descriptor	.	X
stream_id_extension_descriptor*	.	X
ITU-T_timing_descriptor**	X	X
X Presence of the descriptor is permitted . Presence of the descriptor is disallowed * Not used in Transport Streams ** Not mandatory		

For H.222.1 descriptors, when a descriptor is used at the program level it is applied only to that program and has no meaning for an elementary stream within that program. Similarly, when a descriptor is used at the elementary stream level, it has meaning only for that elementary stream and has no meaning at the program level [same as item b) defined in 14.1].

All data elements named **reserved** shall be coded as defined in Recommendation H.222.0, i.e. all reserved bits set to binary ones. Also, extra data contained in descriptors, but undefined by this Recommendation, shall be discarded by decoders conforming to this Recommendation.

14.2.1 ITU-T video stream descriptor

The ITU-T video stream descriptor indicates the H.222.1 defined video elementary stream type. It is coded as shown in Table 7.

TABLE 7/H.222.1

Coding of ITU-T video stream descriptor

Syntax	No. of bits	Identifier
ITU-T_video_stream_descriptor() {		
descriptor_tag	**8**	**uimsbf**
descriptor_length	**8**	**uimsbf**
coding_algorithm	**8**	**uimsbf**
if(coding_algorithm == H.261 || H.263){		
picture_format	**3**	**uimsbf**
minimum_picture_interval	**5**	**uimsbf**
}		
}		

Semantics of the ITU-T video stream descriptor fields

coding_algorithm: The coding_algorithm field indicates the type of H.222.1 defined video elementary stream type. It is coded as shown in Table 8.

TABLE 8/H.222.1

coding_algorithm coding

Value	Description
0x00	forbidden
0x01	Rec. H.261
0x02	Rec. H.261 without FEC
0x03	Rec. H.263
0x04-0xff	reserved

picture_format: The picture_format field indicates the picture format for H.261 and H.263 video. It is coded as shown in Table 9.

TABLE 9/H.222.1

picture_format coding

Value	Description
000	CIF
001	QCIF
010	sub-QCIF
011	4CIF
100	16CIF
101-111	reserved

minimum_picture_interval: The minimum_picture_interval field indicates the minimum interval between pictures for H.261 and H.263 video. The minimum picture interval is calculated from the value contained in the minimum_picture_interval field as (minimum_picture_interval + 1)/29.97.

14.2.2 ITU-T audio stream descriptor

The ITU-T audio stream descriptor indicates the H.222.1 defined audio elementary stream type. It is coded as shown in Table 10.

TABLE 10/H.222.1

Coding of ITU-T audio stream descriptor

Syntax	No. of bits	Identifier
ITU-T_audio_stream_descriptor() {		
descriptor_tag	**8**	**uimsbf**
descriptor_length	**8**	**uimsbf**
coding_algorithm	**8**	**uimsbf**
reserved	**8**	**uimsbf**
}		

Semantics of the ITU-T audio stream descriptor fields

coding_algorithm: The coding_algorithm field indicates the type of H.222.1 defined audio elementary stream type. It is coded as shown in Table 11.

TABLE 11/H.222.1

coding_algorithm

Value	Description
0x00	forbidden
0x01	G.711 A-law
0x02	G.711 μ-law
0x03	Rec. G.722 (mode 1)
0x04	Rec. G.722 (mode 2)
0x05	Rec. G.722 (mode 3)
0x06	Rec. G.723
0x07	Rec. G.728
0x08-0xff	reserved

14.2.3 ITU-T data stream descriptor

The ITU-T data stream descriptor indicates the H.222.1 defined data elementary stream type. It is coded as shown in Table 12.

TABLE 12/H.222.1

Coding of ITU-T data stream descriptor

Syntax	No. of bits	Identifier
ITU-T_data_stream_descriptor() {		
descriptor_tag	**8**	**uimsbf**
descriptor_length	**8**	**uimsbf**
protocol	**8**	**uimsbf**
reserved	**8**	**uimsbf**
}		

Semantics of the ITU-T data stream descriptor fields

protocol: The protocol field indicates the type of H.222.1 defined data elementary stream type. It is coded as shown in Table 13.

14.2.4 stream id extension descriptor

The stream_id_extension descriptor maps the associated H.222.1 defined elementary stream type descriptor to a stream_id_extension field value. The stream_id_extension descriptor is only coded in elementary streams with stream_id value equal to Recommendation H.222.1 type A to type E. The stream id extension descriptor is only coded at the elementary stream level of the Program Stream PSM table. It is not coded in the Transport Stream. The stream id extension descriptor is shown in Table 14.

TABLE 13/H.222.1

Protocol coding

Value	Description
0x00	forbidden
0x01	H.245 subchannel
0x02	video frame synchronous subchannel
0x03	T.120 subchannel
0x04-0xff	reserved

TABLE 14/H.222.1

Coding of the stream id extension descriptor

Syntax	No. of bits	Identifier
stream_id_extension_descriptor() {		
descriptor_tag	**8**	**uimsbf**
descriptor_length	**8**	**uimsbf**
elementary_stream_id_extension	**8**	**uimsbf**
}		

Semantics of the stream_id_extension_descriptor fields

elementary_stream_id_extension: The elementary_stream_id_extension field indicates the value in the PES packet stream_id_extension field which carries the H.222.1 defined elementary stream type.

14.2.5 ITU-T timing descriptor

The ITU-T timing descriptor is shown in Table 15.

TABLE 15/H.222.1

Coding of ITU-T timing descriptor

Syntax	No. of bits	Identifier
ITU-T_timing_descriptor(){		
descriptor_tag	**8**	**uimsbf**
descriptor_length	**8**	**uimsbf**
SC_PESPktR	**24**	**bslbf**
SC_TESPktR	**24**	**bslbf**
SC_TSPktR	**24**	**bslbf**
SC_byte_rate	**30**	**bslbf**
vbv_delay_flag	**1**	**bslbf**
reserved	**1**	**bslbf**
reserved	**32**	**bslbf**
}		

Semantics of the ITU-T timing descriptor fields

SC_PESPktR: "if not equal to" 0xffffff, this parameter indicates that PES packets are passed from the encoder buffer to the multiplexer buffer at a constant rate. Integer value is the ratio of 27 MHz to the PES packet rate. It specifies the interpacket duration for the next PES packet and all following packets up to and including the PES packet following a new valid SC_PESPktR. This parameter is valid only in an elementary stream descriptor. If SC_PESPktR is equal to 0xffffff, the PES interpacket duration is unspecified.

SC_TESPktR: "if not equal to" 0xffffff, this parameter indicates that TS packets for the specified elementary stream (hereinafter called TES packets) are passed from the encoder buffer to the multiplexer buffer at a constant rate. Integer value is the ratio of 27 MHz to the TES packet rate. It specifies the interpacket duration for the next TES packet and all following packets up to and including the TES packet following a new valid SC_TESPktR. This parameter is valid only in an elementary stream descriptor. If SC_TESPktR is equal to 0xffffff, the TES interpacket duration is unspecified.

SC_TSPktR: "if not equal to" 0xffffff, this parameter indicates a constant Transport Stream packet rate. Integer value is the ratio of 27 MHz to the TS packet rate. It specifies the interpacket duration for the current TS packet and all following packets up to and including the TS packet containing a new valid SC_TSPktR. This parameter is valid only in a program descriptor. The program_number indicates to which STC the TS packet rate is locked. If SC_TSPktR is equal to 0xffffff, the Transport Stream interpacket duration is unspecified.

SC_byterate: "if not equal to" 0x3fffffff, this parameter indicates that PES bytes are passed from the encoder buffer to the multiplexer buffer at a constant rate. Integer value is the ratio of 27 MHz to (byte_rate/50) i.e. SC_byterate = 1 350 000 000/byte_rate. It specifies the interbyte duration for the next PES packet and all following packets up to and including the PES packet following a new valid SC_byterate. This parameter is valid only in an elementary stream descriptor. If SC_byterate is equal to 0x3fffffff, the PES interbyte duration is specified.

vbv_delay_flag: This is a 1-bit flag which when set to "1" indicates that the video parameter vbv_delay may be used for timing recovery. The last byte of picture_start_code (see Recommendation H.262 [9]) arrives nominally at time DTS-vbv_delay. This flag is valid only in a video elementary stream descriptor.

All rates mentioned above are locked to the STC, whose frequency is nominally 27 MHz.

In the above defined data, the byte aligned string 0x000001 should be avoided in order to minimize start code emulation.

A detailed explanation of the usage of this descriptor is provided in Appendix II.

15 Synchronization of H.222.1 defined elementary stream types

When H.222.1 defined elementary stream types are carried in a Program Stream or a Transport Stream, they shall have an associated System Time Clock.

In the following subclauses access units for each of the ITU-T stream types are defined. The semantic definition of the PTS field in 2.4.3.7/H.222.0 applies. The PTS field refers to the time of presentation of the presentation unit represented by the first access unit in the PES packet payload.

The DTS field is not coded in PES packets carrying data of the following ITU-T stream types. For these stream types PTS and DTS are always equal.

15.1 H.261 video

For an H.261 video elementary stream type an access unit is defined as the coded representation of a presentation unit, where a presentation unit is a picture. An access unit includes all of the coded data for a picture. The access unit begins with the H.261 picture start code and includes all data, including MBA stuffing up to the next H.261 picture start code.

PTS thus refers to the H.261 picture start code, contained in the PES packet payload.

15.2 H.263 video

For an H.263 video elementary stream type an access unit is defined as the coded representation of a presentation unit, where a presentation unit is a picture. An access unit includes all of the coded data for a picture. The access unit begins with the H.263 picture start code, and includes all data stuffing up to the next H.263 picture start code.

PTS thus refers to the H.263 picture start code, contained in the PES packet payload.

In the case of PB-frames the access unit is the coded representation of two presentation units. In this case PTS refers to the time of presentation of the P-picture and not the B-picture.

15.3 G.711 audio

For a G.711 audio elementary stream type an audio access unit is defined as the coded representation of an audio sample with 8 kHz sampling frequency.

The PTS field in the PES packet header thus refers to the time of presentation of the first byte of the elementary stream in the PES packet payload.

15.4 G.722 audio

For a G.722 audio elementary stream type an access unit is defined as the coded representation of a byte representing the highband audio component (2 bits) and the lowband audio component (4 bits for mode 1, 3 bits for mode 2, 2 bits for mode 3) with 16 kHz sampling frequency. In case of mode 2 or mode 3, the byte also includes 1 or 2 stuffing bits, respectively.

The PTS field in the PES packet header thus refers to the time of presentation of the first byte of the elementary stream in the PES packet payload.

15.5 G.723 audio

For a G.723 audio elementary stream type an access unit is defined as the coded representation of an MP-MLQ/ACELP 30 ms frame. The start of the G.723 frame shall be aligned with the start of the PES packet payload.

The PTS field in the PES packet header thus refers to the time of presentation of the first MP-MLQ/ACELP 30 ms frame in the PES packet payload.

15.6 G.728 audio

For a G.728 audio elementary stream type an access unit is defined as the coded representation of an LD-CELP 2.5 ms frame. The start of the G.728 frame shall be aligned with the start of the PES packet payload.

The PTS field in the PES packet header thus refers to the time of presentation of the first LD-CELP frame in the PES packet payload.

16 System Target Decoder for H.222.1 defined elementary stream types

16.1 Program Streams

When H.222.1 defined elementary stream types are multiplexed into Program Streams, multiplexing shall be performed in such a way that the constraints imposed by the P-STD defined in 2.5.2/H.222.0 with the following parameters, shall be met.

H.261 and H.263 elementary streams shall be considered to be low delay streams, that is, as though they had a low_delay flag set to "1".

All H.261 and H.263 pictures shall be considered to be P-pictures.

16.2 Transport Streams

When H.222.1 defined elementary stream types are multiplexed into Transport Streams, multiplexing shall be performed in such a way that the constraints imposed by the T-STD defined in 2.4.2/H.222.0 with the following parameters, shall be met.

The rate at which data is removed from the TB buffer, Rxn, shall be 4 000 000 bit/s for H.261 and H.263 elementary streams and shall be 2 000 000 bit/s for G.711, G.722, G.723 and G.728 elementary streams.

For H.261 and H.263 elementary streams, the size of the elementary stream buffer, EBS, shall be equal to the size of the HRD buffer specified in H.261 and H.263 or the size of the HRD buffer negotiated with out-of-band signalling when this is used.

For H.261 and H.263 elementary streams, the size of the multiplexing buffer, MBS, shall be equal to $(1/750 + 0.004) * 4\ 000\ 000$ bits.

For G.711, G.722, G.723 and G.728 elementary streams, the size of the main buffer, BS, shall be equal to 3584 bytes.

Data shall be transferred from the multiplexing buffer to the elementary stream buffer using the leak method, where the leak rate, Rbx, shall be equal to 4 000 000 bit/s.

H.261 and H.263 elementary streams shall be considered to be low delay streams, that is, as though they had a low_delay flag set to "1".

All H.261 and H.263 pictures shall be considered to be P-pictures.

17 Video frame synchronous signalling

Synchronization between a video frame synchronous signal and the video frame is achieved using the PTS field in the PES packet header. The PTS refers to the time of a specific event coded in the PES packet payload. The coded video elementary stream and the video frame synchronous subchannel share the same STC. A video frame synchronous subchannel is identified by the ITU-T_data_stream_descriptor coded in the PSM/PSI tables, with the protocol parameter equal to video frame synchronous subchannel.

The events and the coding of the events, are not specified here.

18 Mode changing

A change of an audio, video or data communications mode is performed using in-band or out-of-band subchannel signalling procedures. Mode changing is achieved by terminating communications on one subchannel and beginning communications in the new mode on another subchannel.

19 Encryption

For further study.

20 H.222.1 demultiplexer errors

At the H.222.1 receive side demultiplexer, it shall be possible to identify the error conditions shown in Table 16. These error conditions may be reported to some management entity for possible transmission to the send side H.222.1 peer.

For the demultiplexer, subchannel establishment is deemed to have occurred when:

- in the case of acknowledged subchannel signalling, when the receiver sends a positive acknowledgment signal to the transmitter that originated the subchannel establishment;
- in the case of unacknowledged subchannel signalling, when the complete signalling information for that subchannel has been received at the receiver.

TABLE 16/H.222.1

H.222.1 receive side demultiplexer error conditions

Error type	Error code	Comment
undefined multiplex identifier value	0	A subchannel has not been established for this multiplex identifier value
stream type error	1	The stream type differs from that agreed to at subchannel establishment

Appendix I

Recommended usage of H.222.0 descriptors

(This appendix does not form an integral part of this Recommendation)

TABLE I.1/H.222.1

Recommended usage of ITU-T defined H.222.0 descriptors

Tag value	Descriptor	Program level	Elementary stream level
2	video_stream_descriptor	.	X
3	audio_stream_descriptor	.	X
4	hierarchy_descriptor	.	X
5	registration_descriptor	NS	NS
6	data_stream_alignment_descriptor	.	X
7	target_background_grid_descriptor	.	X
8	video_window_descriptor	.	X
9	CA_descriptor	NS	NS
10	ISO_639_language_descriptor	.	X
11	system_clock_descriptor	X	.
12	multiplex_buffer_utilization_descriptor	.	X
13	copyright_descriptor	X	X
14	maximum_bitrate_descriptor	X	X
15	private_data_indicator_descriptor	NS	NS
16	smoothing_buffer_descriptor	X	X
17	STD_descriptor	.	X
18	IBP_descriptor	.	X
NS Not specified in this appendix X Presence of the descriptor is permitted . Presence of the descriptor is not recommended			

Appendix II

ITU-T Timing Descriptor usage

(This appendix does not form an integral part of this Recommendation)

In severely jittered or highly lossy networks, timing recovery may be problematic in inexpensive implementations that cannot afford highly stable crystal clocks. Often the encoders can provide information as to which parts of the data stream may be useful for recovering the decoder System Time Clock (STC) frequency and phase. Such information is specified in the ITU-T Timing Descriptor.

The descriptor specifies that one or more data rates are locked to the encoder STC and can be used, for example, in an Adaptive Clock Recovery (ACR) scheme at the decoder. The size of the assumed buffer in the ACR is implementation dependent and will depend on how much jitter is encountered during transmission. The jitter may or may not include multiplexing jitter. If it does, then the assumed ACR buffer will probably be larger.

In an ITU-T H.222.0 multiplexed bit stream, the encoder periodically sends samples of its time clock, which are called Clock References. We denote these Encoder Clock References by PCR. Typically, the PCRs are sent at the rate of a few per second.

With very low jitter, recovery of the time clock at the decoder relies on averaging the jitter effects over many received PCR values. However, in cases of extreme jitter, thousands of PCR values may be needed before high stability operation is achieved, which may require dozens of seconds. In many applications this is unacceptable.

However, if the ITU-T Timing Descriptor specifies an SC_TSPktR value, then additional synthetic PCRs can be calculated for each ensuing TS packet simply by adding SC_TSPktR for each TS packet received. Timing recovery is speeded up considerably by using these synthetic PCRs.