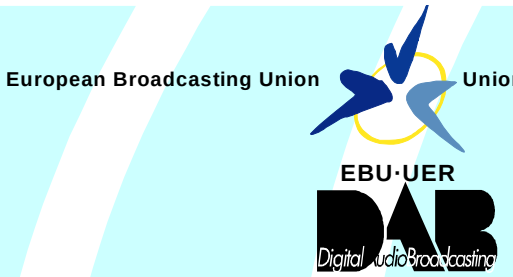
ETSI TS 102 428 V1.1.1 (2005-06)

Technical Specification

Digital Audio Broadcasting (DAB); DMB video service; User Application Specification



Reference
DTS/JTC-DAB-39

Keywords
broadcasting, digital, DAB, video

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights 4
Foreword 4
1 Scope 5
2 References 5
3 Definitions and abbreviations 6
3.1 Definitions 6
3.2 Abbreviations 6
4 Architecture 7
4.1 Video service transmission architecture 7
4.2 Video multiplexer architecture 8
5 Contents description 9
5.1 Composition of contents 9
5.2 Packetization of content 10
5.3 Audio Object 11
5.4 Video Object 11
5.5 Auxiliary Data Object 11
5.5.1 Scene description specification 11
5.5.2 Graphics data specification 11
6 Transport stream specification 12
6.1 Transport stream packet specification 12
6.2 PES packet specification 13
7 User Application signalling 15
8 DMB video service profiles 16
8.1 Profile 1 16
8.1.1 Audio object 16
8.1.2 Video Object 16
8.1.2.1 Profile and levels supported 16
8.1.2.2 Specification related to the transport of a video stream 17
8.2 Profile 2 17
8.2.1 Audio object 17
8.2.1.1 List of Tools/Functionalities 17
8.2.1.2 Comparison with existing profiles and object types 18
8.2.1.3 Level definition 18
8.2.2 Video Object 18
Annex A (informative): An example of the IOD/OD/BIFS 19
A.1 IOD (binary syntax and field values) 20
A.2 OD (binary syntax and values) 22
A.2.1 The case of a single audio object in a broadcast stream 22
A.2.2 The case of a single audio object and a single video object in a broadcast stream 23
A.2.3 The case where multi-lingual audio is supported (audio 1/ audio 2/ video) 25
A.3 BIFS 27
A.3.1 The case of a single audio object within a broadcast stream 27
A.3.1.1 Syntax 27
A.3.1.2 Coded data 27
A.3.2 The case of a single audio object and a single video object within a broadcast stream 28
A.3.2.1 Syntax 28
A.3.2.2 Coded data 28
Annex B (informative): Content Access Procedure 29
History 30

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://webapp.etsi.org/IPR/home.asp).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Specification (TS) has been produced by Joint Technical Committee (JTC) Broadcast of the European Broadcasting Union (EBU), Comité Européen de Normalisation ELECtrotechnique (CENELEC) and the European Telecommunications Standards Institute (ETSI).

NOTE 1: The EBU/ETSI JTC Broadcast was established in 1990 to co-ordinate the drafting of standards in the specific field of broadcasting and related fields. Since 1995 the JTC Broadcast became a tripartite body by including in the Memorandum of Understanding also CENELEC, which is responsible for the standardization of radio and television receivers. The EBU is a professional association of broadcasting organizations whose work includes the co-ordination of its members' activities in the technical, legal, programme-making and programme-exchange domains. The EBU has active members in about 60 countries in the European broadcasting area; its headquarters is in Geneva.

European Broadcasting Union
CH-1218 GRAND SACONNEX (Geneva)
Switzerland
Tel: +41 22 717 21 11
Fax: +41 22 717 24 81

The Eureka Project 147 was established in 1987, with funding from the European Commission, to develop a system for the broadcasting of audio and data to fixed, portable or mobile receivers. Their work resulted in the publication of European Standard, EN 300 401 [1], for DAB (see note) which now has worldwide acceptance. The members of the Eureka Project 147 are drawn from broadcasting organizations and telecommunication providers together with companies from the professional and consumer electronics industry.

NOTE 2: DAB is a registered trademark owned by one of the Eureka Project 147 partners.

1 Scope

The present document specifies the user application for video services carried via DAB. It also includes profile definitions for the application.

The user application is delivered using the MSC stream mode (EN 300 401 [1]) including additional error protection (TS 102 427 [2]).

The present document defines the components of the video services; the content compression, the synchronization mechanism and multiplexing mechanism. The components of the service are the video object, the audio object, and the auxiliary data object. All the objects are packetized and synchronized using MPEG-4 SL (ISO/IEC 14496-1 [4]). The present document also specifies the mechanism for the multiplexing of the multimedia data using MPEG-2 TS (ISO/IEC 14496-1 [4]). For efficiency, some appropriate restrictions to MPEG-4 SL and MPEG-2 TS are specified.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication and/or edition number or version number) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

[1] ETSI EN 300 401: "Radio Broadcasting Systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers".

[2] ETSI TS 102 427: "Digital Audio Broadcasting (DAB); Data Broadcasting - MPEG 2 TS streaming".

[3] ISO/IEC 13818-1: "Information technology - Generic coding of moving pictures and associated audio information: Systems".

[4] ISO/IEC 14496-1: "Information technology - Coding of audio-visual objects - Part 1: Systems".

[5] ISO/IEC 14496-3: "Information technology - Coding of audio-visual objects - Part 3: Audio".

[6] ISO/IEC 14496-10: "Information technology - Coding of audio-visual objects - Part 10: Advanced Video Coding".

[7] ETSI TS 101 756: "Digital Audio Broadcasting (DAB); Registered tables".

[8] ISO/IEC 10918-1: "Information technology - Digital compression and coding of continuous-tone still images: Requirements and guidelines".

[9] ITU-T Recommendation H.264: "Advanced video coding for generic audiovisual services".

[10] ISO/IEC 14496-3 (2001/AMD2:2004): "Parametric coding for high-quality audio".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

audio object: data encoded by the audio compression algorithm used for video service

audio service: service composed of an audio component, and optionally, additional audio or data components, as defined in EN 300 401 [1]

auxiliary data object: data encoded by the auxiliary data compression algorithm used for video service

elementary stream: consecutive flow of compressed mono-media data, i.e. one of the coded audio, coded video or other coded bitstream

interactive service: service in which, the users can select, respond to or control the broadcast content

random access: capability of receiving a service from an arbitrary point in its timeline rather than being confined to progressively receiving it from its beginning

Synchronization (Sync) layer: Sync Layer (SL), as defined in ISO/IEC 14496-1 [4], provides a combined transmission of timing and synchronization information with elementary streams (ESs)

transport rate available for video services: maximum bit-rate of the outer-coded TS that can be fed as an input to the stream mode data path defined in EN 300 401 [1]

Transport Stream (TS) multiplexer: device that combines all the ESs that belong to a single program into a Transport Stream (TS) defined in ISO/IEC 13818-1 [3] for the delivery through a single transmission path

video object: data encoded by the video compression algorithm used for the video service

video service: service composed of the video object, the audio object associated with the video and optional auxiliary data object

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

BIFS	BInary Format for Scene
CTS	Composition Time Stamp
DAB	Digital Audio Broadcasting
DMB	Digital Multimedia Broadcasting
DTS	Decoding Time Stamp
ES	Elementary Stream
ID	IDentifier
IOD	Initial Object Descriptor
IPI	Intellectual Property Identification
IPMP	Intellectual Property Management and Protection
OCR	Object Clock Reference
OD	Object Descriptor
PAT	Program Association Table
PCR	Program Clock Reference
PES	Packetized Elementary Stream
PID	Program IDentifier
PMT	Program Map Table
PSI	Program Specific Information
PTS	Presentation Time Stamp
RS	Reed-Solomon
SL	Synchronization Layer
TS	Transport Stream

URL Universal Resource Locator

4 Architecture

The video service user application described in the present document is delivered through the MSC stream mode (EN 300 401[1]). In order to maintain extremely low bit error rates, the service uses the additional error protection mechanism described in TS 102 427 [2]. The video service is composed of three layers; content compression layer, synchronization layer, and transport layer. In the content compression layer, a specific video compression method is employed for the video contents, a specific compression method for the audio compression, and MPEG-4 BIFS for the auxiliary interactive data services. To synchronize the audio and video content, both temporally and spatially, MPEG-4 SL is employed in the synchronization layer. In the transport layer, MPEG-2 TS with some appropriate restrictions is employed for the multiplexing of the compressed audio and video content.

This clause defines the system architecture for the video service user application. This architecture is illustrated in figure 1.

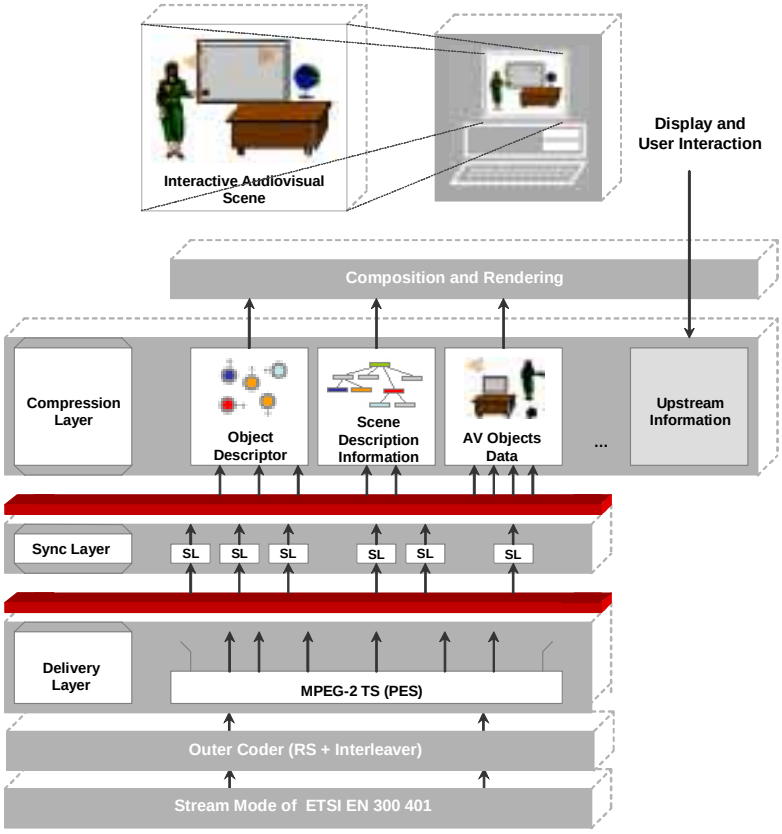


Figure 1: Conceptual architecture for the video service

4.1 Video service transmission architecture

The conceptual transmission architecture for the video service is shown in figure 2.

- The video, audio, and the auxiliary data information which makes up a video service, are multiplexed into an MPEG-2 TS and further outer error correction is added (TS 102 427 [2]). The video multiplexer is described in clause 4.2.
- The multiplexed (and outer-coded) stream is transmitted by the MSC stream mode data channel of DAB defined in EN 300 401 [1].

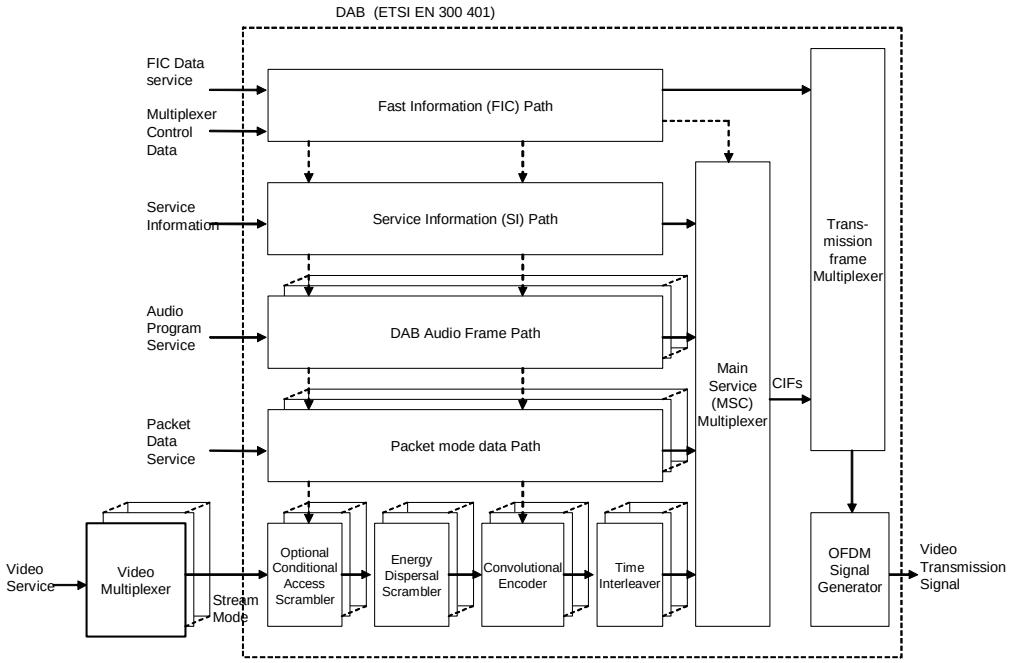


Figure 2: Conceptual transmission architecture for the video services

4.2 Video multiplexer architecture

The conceptual architecture of the video multiplexer for a video service is shown in figure 3.

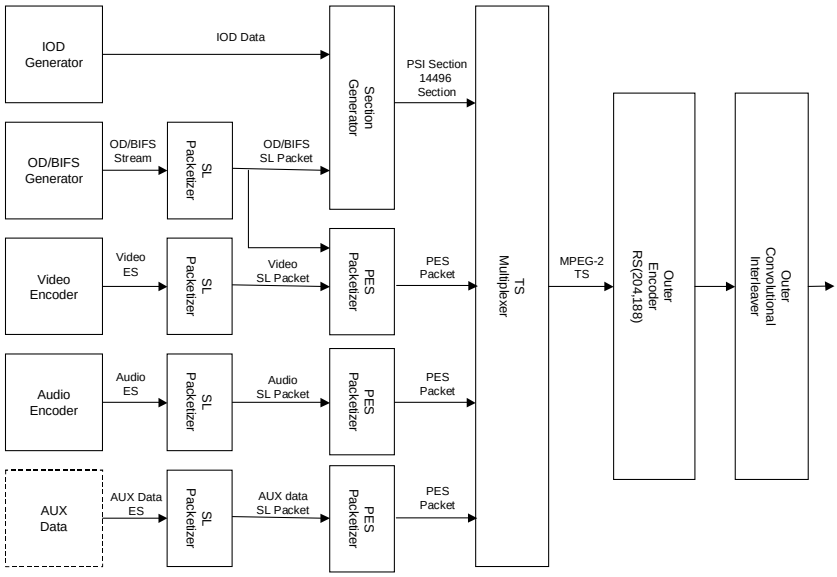


Figure 3: Conceptual architecture of the video multiplexer

- The IOD generator creates IODs that comply with the ISO/IEC 14496-1 [4].
- The OD/BIFS generator creates OD/BIFS streams that comply with the ISO/IEC 14496-1 [4].

- The audio encoder generates an encoded bitstream compliant with the audio specification defined in one of the video service profiles defined in clause 8, by performing data compression processing of the input audio signal (see clause 5.3).
- The video encoder generates an encoded bitstream compliant with the video specification defined in one of the video service profiles defined in clause 8, by performing data compression processing of the input video signal (see clause 5.4).
- Each SL packetizer generates a SL packet stream, compliant with ISO/IEC 14496-1 [4] system standard for each input media stream (see clause 5.2).
- The section generator creates sections compliant with ISO/IEC 13818-1 [3] for the input IOD/OD/BIFS (see annex B).
- Each PES packetizer generates a PES packet stream compliant with ISO/IEC 13818-1 [3] for each SL packet stream (see clause 6.2).
- The TS multiplexer combines the input sections and PES packet streams into a single MPEG-2 transport stream complying with ISO/IEC 13818-1 [3] (see clause 6.1).
- The outer encoder generates Reed-Solomon (RS) codes for error correction to each packet in the MPEG-2 TS multiplexed data stream (see TS 102 427 [2]).
- The outer-coded data stream is interleaved by the outer interleaver, which is a convolutional interleaver (see TS 102 427 [2]), and is output into the DAB sub-channel as a MSC stream data service component.

5 Contents description

5.1 Composition of contents

Among several OD profiles defined in ISO/IEC 14496-1 [4], the tools defined in the "Core Profile" are used for the composition of the contents in the video services. The IPMP tool is not used.

The following restrictions are imposed on the MPEG-4 descriptors that are used for the composition of the contents of the video service.

The following descriptors shall always be used:

- OD.
- IOD.
- ES Descriptor.
- Decoder Config Descriptor.
- SL Config Descriptor.

The following descriptors shall **not** be used:

- IPI Descriptor Pointer.
- IPMP Descriptor Pointer.
- IPMP Descriptor.

The object type of the contents for the video service shall be determined as the objectTypeIndication values of DecoderConfigDescriptor. The values and the object types are listed in ISO/IEC 14496-1 [4] as table 1 objectTypeIndication values.

The permitted values for this application are listed in table 1.

Table 1: objectTypeIndication values

values	objectTypeIndication Description
0x02	Systems ISO/IEC 14496-1 [4]
0x21	Visual ISO/IEC 14496-10 [6]
0x40	Audio ISO/IEC 14496-3 [5]
0x6C	Visual ISO/IEC 10918-1 [8]
0xC0 – 0xFE	user private

The permitted values for the stream types are listed in table 2.

For examples of the IOD/OD/BIFS in the cases where a single video object and a single or multiple audio objects are broadcast, see annex B.

Table 2: Permitted stream types

streamType value	Stream Type
0x01	ObjectDescriptorStream
0x02	ClockReferenceStream
0x03	SceneDescriptionStream
0x04	VisualStream
0x05	AudioStream
0x20 - 0x3F	user private

For the content access procedure within receiving terminals playing a video service, see annex B. For the video service, only one video object and only one audio object shall be rendered in a scene.

5.2 Packetization of content

Content shall be packetized as Sync Layer (SL) packets as defined in ISO/IEC 14496-1 [4]. The following rules are applied to SL packet headers:

- The "useAccessUnitStartFlag" field has no restriction on its value.
- The "useAccessUnitEndFlag" field has no restriction on its value, but shall always be used with the "useAccessUnitStartFlag" field.
- The "useRandomAccessPointFlag" field has its value restricted to "0".

NOTE 1: Random access is supported through use of the "random_access_indicator" field within the TS packet.

- The "hasRandomAccessUnitsOnlyFlag" field has its value restricted to "0".
- The "usePaddingFlag" field has its value restricted to "0".

NOTE 2: Padding is employed in PES packets.

- The "useTimeStampsFlag" field has its value restricted to "1".
- The "useIdleFlag" field has its value restricted to "1".
- The "durationFlag" field has no restriction on its value.
- The "timeScale" field shall always be used if the "durationFlag" field has the value of "1".
- The "accessUnitDuration" field shall always be used if the "durationFlag" field has the value of "1".
- The "compositionUnitDuration" field shall always be used if the "durationFlag" field has the value of "1".
- The "timeStampResolution" field shall be set to 90,000 Hz.
- The "OCRResolution" field shall be set to 90,000 Hz.

- The "timeStampLength" field shall be less than or equal to 33 bits.
- The "OCRLength" field shall be less than or equal to 33 bits.
- The "AU_Length" field has its value restricted to "0".
- The "instantBitrateLength" field has no restriction on its value.

NOTE 3: This field shall be used if an OCR is encoded within an SL packet header since the "instantBitrate" field shall also be encoded in the case.

- The "degradationPriorityLength" field has its value restricted to "0".
- The "AU_seqNumLength" field has its value restricted to "0".
- The "packetSeqNumLength" field has its value restricted to "0".

The recovery and usage of timing information shall comply with the following:

- Clauses 2.11.3.3, 2.11.3.4 and 2.11.3.6 in ISO/IEC 13818-1 [3].
- The OCR defined in ISO/IEC 14496-1 [4] shall synchronize all the objects necessary for the description of a scene.

5.3 Audio Object

In the video service user application, the audio object type that needs to be supported by the audio decoder shall be determined as an objectTypeIndication value within DecoderConfigDescriptor. The detailed specifications for the audio decoding are defined in clause 8.

5.4 Video Object

In the video service user application, the video object type that needs to be supported by the video decoder shall be determined as an objectTypeIndication value within DecoderConfigDescriptor. The detailed specifications for the video decoding are defined in clause 8.

5.5 Auxiliary Data Object

This part of the specification is selectively used only when auxiliary information is transported or synchronized interactive services are provided.

5.5.1 Scene description specification

Scene description specification complies with Core2D@Level 1 defined in ISO/IEC 14496-1 [4].

5.5.2 Graphics data specification

Graphics data specification complies with Core2D@Level 1 defined in ISO/IEC 14496-1 [4].

The character codes used for "Text" nodes are the basic characters used in DAB (EN 300 401 [1]).

6 Transport stream specification

The MPEG-2 transport stream layer multiplexes video, audio, and auxiliary data to form a single program. It does not support the conditional access scheme defined in ISO/IEC 13818-1 [3] because CAT is not used in the PSI. It uses PCR for system clock recovery. The ISO/IEC 14496-1 [4] MPEG-4 system layer provides synchronization among ESs using OCR, CTS, and DTS together with the PCR described above. In addition, it provides linkage among ESs that constitutes a video service and uses scene description information for the composition of a video service. It uses the SL packetization, but does not utilize FlexMux multiplexing.

6.1 Transport stream packet specification

A TS packet shall have the structure shown in table 3. Mandatory restrictions are detailed where applicable.

Table 3: Structure of a TS packet

Syntax	Number of bits	Restrictions
Transport_packet(){		
sync_byte	**8**	
transport_error_indicator	**1**	
payload_unit_start_indicator	**1**	
transport_priority	**1**	
PID	**13**	
transport_scrambling_control	**2**	**"00"**
adaptation_field_control	**2**	
continuity_counter	**4**	
if(adaptation_field_control = = '10' || adaptation_field_control = = '11'){		
adaptation_field()		
}		
if(adaptation_field_control = = '01' || adaptation_field_control = = '11') {		
for (i=0; i<N; i++){		
Data_byte	**8**	
}		
}		
}		

The adaptation field within a TS packet shall have the structure shown in table 4. Manditory restrictions are detailed where applicable.

Table 4: Structure of the adaptation field of a TS packet

Syntax	Number of bits	Restrictions
adaptation_field() {		
adaptation_field_length	**8**	
if (adaptation_field_length>0) {		
Discontinuity_indicator	**1**	
random_access_indicator	**1**	
elementary_stream_priority_indicator	**1**	
PCR_flag	**1**	
OPCR_flag	**1**	**"0"**
splicing_point_flag	**1**	
transport_private_data_flag	**1**	
adaptation_field_extension_flag	**1**	**"0"**
if (PCR_flag = = '1') {		
program_clock_reference_base	**33**	
Reserved	**6**	
program_clock_reference_extension	**9**	
}		
if (OPCR_flag = = '1') {		**not used**
}		
if (splicing_point_flag = = '1') {		
splice_countdown	**8**	
}		
if (transport_private_data_flag = = '1') {		
transport_private_data_length	**8**	
for (i=0; i<transport_private_data_length; i++) {		
Private_data_byte	**8**	
}		
}		
if (adaptation_field_extension_flag = = '1') {		**not used**
}		
for (i=0; i<N; i++) {		
stuffing_byte	**8**	
}		
}		
}		

6.2 PES packet specification

A PES packet shall have the structure shown in table 5. Manditory restrictions are detailed where applicable.

Table 5: Structure of a PES packet

Syntax	Number of bits	Restrictions
PES_packet() {		
packet_start_code_prefix	**24**	
stream_id	**8**	**0xFA**
PES_packet_length	**16**	
if (stream_id !=program_stream_map		
&& stream_id !=padding_stream		
&& stream_id !=private_stream_2		
&& stream_id !=ECM		
&& stream_id !=EMM		
&& stream_id !=program_stream_directory		
&& stream_id !=DSMCC_stream		
&& stream_id !=ITU-T Rec. H.222.1 type E stream) {		
'10'	**2**	
PES_scrambling_control	**2**	**"00"**
PES_priority	**1**	
data_alignment_indicator	**1**	
Copyright	**1**	
original_or_copy	**1**	
PTS_DTS_flags	**2**	**"10" or "00"**
ESCR_flag	**1**	**"0"**
ES_rate_flag	**1**	**"0"**
DSM_trick_mode_flag	**1**	**"0"**
additional_copy_info_flag	**1**	**"0"**
PES_CRC_flag	**1**	**"0"**
PES_extension_flag	**1**	**"0"**
PES_header_data_length	**8**	
if (PTS_DTS_flags = = '10') { (see note)		
'0010'	**4**	
PTS [32..30]	**3**	
marker_bit	**1**	
PTS [29..15]	**15**	
marker_bit	**1**	
PTS [14..0]	**15**	
marker_bit	**1**	
}		
if (PES_extension_flag = = '1') {		**not used**
}		
for (i=0; i<N1; i++) {		
stuffing_byte	**8**	
}		
for (i=0; i<N2; i++) {		
PES_packet_data_byte	**8**	
}		
}		
}		
NOTE: The PTS field is included in a PES header only when the encapsulated SL packet header includes an OCR. Otherwise, the PTS field is not used.		

The following rules are applied at the transmitting side in order to allow random access at the receiving side.

- A PAT (Program Association Table) shall describe only one program, and its transmission period shall be no greater than 500 ms.
- A PMT (Program Map Table) shall have the structure shown in table 6 and adhere to the following rules:
  - A group of descriptors with Restriction "A" in the table shall include an IOD_descriptor.
  - A group of descriptors with Restriction "B" in the table shall include an SL_descriptor for an ES_ID.
  - The transmission period of a PMT shall be no greater than 500 ms.

Table 6: Structure of a PMT

Syntax	Number of bits	Restrictions
TS_program_map_section() {		
table_id	**8**	
section_syntax_indicator	**1**	
'0'	**1**	
Reserved	**2**	
section_length	**12**	
program_number	**16**	
Reserved	**2**	
version_number	**5**	
current_next_indicator	**1**	
section_number	**8**	
last_section_number	**8**	
Reserved	**3**	
PCR_PID	**13**	
Reserved	**4**	
program_info_length	**12**	
for (i=0; i<N; i++) {		
descriptor()		A
}		
for (i=0; i<N1; i++) {		
stream_type	**8**	"0x12" or "0x13"
Reserved	**3**	
elementary_PID	**13**	
Reserved	**4**	
ES_info_length	**12**	
for (i=0; i<N2; i++) {		
descriptor()		B
}		
}		
CRC_32	**32**	
}		

- The transmission period for scene description information and object description information shall be no greater than 500 ms.
- The transmission period of a PCR within a transport stream shall be no greater than 100 ms.
- The transmission period of an OCR in the ISO/IEC 14496-1 [4] SL layer shall be no greater than 700 ms.
- The transmission period of a CTS in the ISO/IEC 14496-1 [4] SL layer shall be no greater than 700 ms.

7 User Application signalling

The use of the video service within a DAB data channel shall be indicated by the use of FIG0/13 (see EN 300 401 [1]) with a UserApplicationType value for the DMB video service (see TS 101 756 [7]). The user application data field shall carry a one byte field - the VideoServiceObjectProfileId - indicating the specific profiles of audio and video object decoder that should be used to decode the audio and video elementary streams of the video service. Receivers shall ignore any user application data following the VideoServiceObjectProfileId field.

The profiles defined for the VideoServiceObjectProfileId field are given in table 7.

Table 7: Video service profiles

VideoServiceObjectProfileId	Description
0x00	Reserved
0x01	Profile 1
0x02	Profile 2
0x03 to 0xff	Reserved

8 DMB video service profiles

8.1 Profile 1

Profile 1 is signalled with VideoServiceObjectProfileId = 0x01.

- audio object: MPEG-4 ER-BSAC.
- video object: ITU-T Recommendation H.264 [9] | MPEG-4 AVC.

8.1.1 Audio object

Audio object specification conforms to the standard relevant to the ER BSAC Audio Object Type with ObjectType ID 22 defined in ISO/IEC IS 14496-3 [5].

Audio object bitstream has the following restrictions.

- In AudioSpecificConfig():
  - epConfig: restricted to 0.
- In GASpecificConfig():
  - frameLengthFlag: restricted to 0.
  - DependOnCoreCoder: restricted to 0.
- In bsac_header():
  - sba_mode: restricted to 0 so that the error resilience tool is not supported.
- In general_header():
  - ltp_data_present: restricted to 0.
- The restrictions in table 8 shall be applied.

Table 8: Restrictions on audio objects

Item	Value
Sampling rate	24,000 Hz; 44,100 Hz; 48,000 Hz
Number of channels	1, 2
Number of objects	1
Maximum bitrate	128 kbps

8.1.2 Video Object

Video objects are based on ITU-T Recommendation H.264 [9] | ISO/IEC 14496-10 [6]. Video bitstreams shall comply with each of the items below.

8.1.2.1 Profile and levels supported

- Profile:
  - Video bitstreams shall comply with the 'Baseline Profile' (ITU-T Recommendation H.264 | ISO/IEC 14496-10 [6], clause A.2.1).
  - 'Arbitrary slice order' shall not be allowed.
  - In the syntax of "Picture Parameter Sets", the 'num_slice_groups_minus1' field has its value restricted to '0'.

- In the syntax of 'Picture Parameter Sets', the 'redundant_pic_cnt_present_flag' field has its value restricted to '0'.
- In the syntax of 'Sequence Parameter Sets', the 'pic_order_cnt_type' field has its value restricted to '2'.
- In the syntax of 'Sequence Parameter Sets', the 'num_ref_frames'field has its value restricted to '3'.

- Level:
  - Level 1,3 in table A-1 of ITU-T Recommendation H.264/AVC annex A shall be used with the following further restrictions.
  - Resolutions supported: the formats listed in table 9.
  - Vertical MV component range (MaxVmvR) shall be [-64,+63,75].
  - Maximum frame rate for each of the resolutions supported shall be 30 fps.
  - MaxDPB shall be 445,5 Kbytes at maximum.

Table 9: Resolutions supported

Format	PicWidthInMbs	FrameHeightInMbs	PicSizeInMbs
QCIF	11	9	99
QVGA	20	15	300
WDF	24	14	336
CIF	22	18	396

8.1.2.2 Specification related to the transport of a video stream

To enable random access at the receiving side, IDR pictures shall be encoded within a video stream at least once every 2 seconds.

The 'Parameter Set' shall be delivered through DecoderSpecificInfo or included in the video stream itself.

The specification related to the transport of a video stream after MPEG-4 SL packetization shall comply with clause 14 of ISO/IEC 14496-1 [4]:2001 Amendment 7.

8.2 Profile 2

Profile 2 is signalled with VideoServiceObjectProfileId = 0x02.

- audio object: MPEG-4 HE AAC V2.
- video object : ITU-T Recommendation H.264 | MPEG-4 AVC.

8.2.1 Audio object

Audio object specification conforms to the standard defined as the Parametric Stereo AOT (Audio Object Type). The PS AOT includes the PS tool, as defined in ISO/IEC 14496-3: 2001/AMD2:2004 [10].

8.2.1.1 List of Tools/Functionalities

The HE AAC v2 profile contains the following audio object types:

- AAC-LC.
- SBR.
- PS.

AAC-LC offers waveform coding over a large bitrate range. In combination with SBR full audio bandwidth is available at all bitrates. The PS is a tool for coding the stereo image of an audio signal at very low bitrates, enabling the reproduction of a stereo signal from a transmitted mono signal given a small amount of side information.

8.2.1.2 Comparison with existing profiles and object types

The AAC-LC object type is part of the AAC profile, and AAC LC and SBR constitutes the HE-AAC profile. These are hierarchical profiles, meaning that the HE-AAC profile decoder of a certain level can handle all AAC profile streams of the corresponding level. The HE-AAC v2 profile is hierarchical w.r.t. the AAC profile and the HE-AAC profile, hence backwards compatibility is ensured so that the HE-AAC v2 profile decoder of a certain level can handle all HE-AAC profile streams of the corresponding level.

8.2.1.3 Level definition

The levels within the HE AAC V2 profile are defined according to table 10.

Table 10: Levels for the HE AAV v2 Profile

Level	Max. channels per object	Max. AAC sampling rate, SBR not present [kHz]	Max. AAC sampling rate, SBR present [kHz]	Max. SBR sampling rate [kHz] (in/out)
1	NA	NA	NA	NA
2	2	48	24	24/48 (note 1)
3	2	48	24/48 (note 3)	48/48 (note 2)
4	5	48	24/48 (note 4)	48/48 (note 2)

NOTE 1: A level 2 HE-AAC v2 Profile decoder implements the baseline version of the parametric stereo tool. Higher level decoders shall not be limited to the baseline version of the parametric stereo tool.
NOTE 2: For level 3 and level 4 decoders, it is mandatory to operate the SBR tool in down-sampled mode if the sampling rate of the AAC core is higher than 24 kHz.
NOTE 3: If Parametric Stereo data is present the maximum AAC sampling rate is 24 kHz, if Parametric Stereo data is not present the maximum AAC sampling rate is either 32 kHz or 48 kHz.
NOTE 4: For one or two channels the maximum AAC sampling rate, with SBR present, is either 32 kHz or 48 kHz. For more than two channels the maximum AAC sampling rate, with SBR present, is 24 kHz.

- The IRD shall support all levels, indicated in table 10. The audio output of the IRD can be restricted to 1 or 2 audio channels only.
- The restrictions in table 11 shall be applied.
- The IRD shall provide functionality for a downmix to either 1-channel mono or 2-channel stereo for an audio object with more than 1 or 2 audio channels present.

Table 11: Restrictions on audio objects within an audio object bitstream

Item	Value
Sampling rate	24,000 Hz; 32,000 Hz; 48,000 Hz
Number of channels	1, 2, 3, 4, 5 plus LFE
Number of objects	1
Maximum bitrate	320 kbps

8.2.2 Video Object

Video objects are based on ITU-T Recommendation H.264 | ISO/IEC 14496-10 [6] as defined in clause 8.1.2.

Annex A (informative): An example of the IOD/OD/BIFS

In annex A, examples of minimum IOD/OD/BIFS is described in the case of a single audio object, a single audio object with a single video object and a single video object with multiple audio objects all within a broadcast stream.

Annex A describes the binary syntax and field values of IOD/OD/BIFS by using the example values of ES_ID, OD_ID, and URL as shown in the following table A.1.

Table A.1: Example values of ES_ID, OD_ID, and URL

	ES_ID	OD_ID	URL
OD stream	1	0	0
BIFS stream	2		
Audio object	101	10	10
Video object	201	20	20

A.1 IOD (binary syntax and field values)

# of bits	Field Name	Value
InitialObjectDescriptor		
8	InitialObjectDescriptor tag	0x02
16	descriptor size	---
10	ObjectDescriptorID	0
1	URL_Flag	0
1	includeInlineProfilesFlag	0
4	Reserved	15
8	ODProfileLevelIndication 1	0x01
8	sceneProfileLevelIndication	0x0C
8	audioProfileLevelIndication	0x23
8	visualProfileLevelIndication	
8	graphicsProfileLevelIndication	0x04
ES_Descriptor(OD)		
8	ES_Descriptor tag	0x03
8	descriptor size	---
16	ES_ID	1
1	streamDependenceFlag	0
1	URL_Flag	0
1	OCRstreamFlag	0
5	streamPriority	0
DecoderConfigDescriptor(OD)		
8	DecoderConfigDescriptor tag	0x04
8	descriptor size	---
8	objectTypeIndication	1
6	streamType	1
1	Upstream	0
1	Reserved	1
24	bufferSizeDB	250
32	maxBitrate	0
32	avgBitrate	0
SLConfigDescriptor(OD)		
8	SLConfigDescriptor tag	0x06
8	descriptor size	---
8	Predefined	0
1	useAccessUnitStartFlag	1
1	useAccessUnitEndFlag	1
1	useRandomAccessPointFlag	0
1	hasRandomAccessUnitsOnlyFlag	0
1	usePaddingFlag	0
1	useTimeStampsFlag	1
1	useIdleFlag	1
1	durationFlag	0
32	timeStampResolution	90000
32	OCRResolution	90000
8	timeStampLength	33
8	OCRLength	33
8	AU_Length	0
8	instantBitrateLength	---
4	degradationPriorityLength	0
5	AU_seqNumLength	0
5	Packet_SeqNumLength	0
2	Reserved	3(0b11)
ES_Descriptor(BIFS)		
8	ES_Descriptor tag	0x03
8	descriptor size	---
16	ES_ID	2
1	streamDependenceFlag	0
1	URL_Flag	0
1	OCRstreamFlag	0
5	streamPriority	0

# of bits	Field Name	Value
DecoderConfigDescriptor(BIFS)		
8	DecoderConfigDescriptor tag	0x04
8	descriptor size	---
8	objectTypeIndication	2
6	streamType	3
1	Upstream	0
1	Reserved	1
24	bufferSizeDB	22
32	maxBitrate	0
32	avgBitrate	0
SLConfigDescriptor(BIFS)		
8	SLConfigDescriptor tag	0x06
8	descriptor size	---
8	Predefined	0
1	useAccessUnitStartFlag	1
1	useAccessUnitEndFlag	1
1	useRandomAccessPointFlag	0
1	hasRandomAccessUnitsOnlyFlag	0
1	usePaddingFlag	0
1	useTimeStampsFlag	1
1	useIdleFlag	1
1	durationFlag	0
32	timeStampResolution	90000
32	OCRResolution	90000
8	timeStampLength	33
8	OCRLength	33
8	AU_Length	0
8	instantBitrateLength	---
4	degradationPriorityLength	0
5	AU_seqNumLength	0
5	Packet_SeqNumLength	0
2	Reserved	3(0b11)

A.2 OD (binary syntax and values)

A.2.1 The case of a single audio object in a broadcast stream

# of bits	Field Name	Value
ObjectDescriptorUpdate		
8	ObjectDescriptorUpdate tag	0x01
8	descriptor size	---
ObjectDescriptor(audio)		
8	ObjectDescriptor tag	0x01
8	descriptor size	---
10	ObjectDescriptorID	10
1	URL_Flag	0
5	Reserved	31
ES_Descriptor(audio)		
8	ES_Descriptor tag	0x03
8	descriptor size	---
16	ES_ID	101
1	streamDependenceFlag	0
1	URL_Flag	0
1	OCRstreamFlag	0
5	streamPriority	5
DecoderConfigDescriptor(audio)		
8	DecoderConfigDescriptor tag	0x04
8	descriptor size	---
8	objectTypeIndication	0x40
6	streamType	5
1	Upstream	0
1	Reserved	1
24	bufferSizeDB	---
32	maxBitrate	---
32	avgBitrate	---
DecoderSpecificInfo(audio)		
8	DecoderSpecificInfo tag	0x05
8	descriptor size	---
8*	DecoderSpecificInfo data	---
SLConfigDescriptor(audio)		
8	SLConfigDescriptor tag	0x06
8	descriptor size	---
8	Predefined	0
1	useAccessUnitStartFlag	1
1	useAccessUnitEndFlag	1
1	useRandomAccessPointFlag	0
1	hasRandomAccessUnitsOnlyFlag	0
1	usePaddingFlag	0
1	useTimeStampsFlag	1
1	useIdleFlag	1
1	durationFlag	0
32	timeStampResolution	90000
32	OCRResolution	90000
8	timeStampLength	33
8	OCRLength	33
8	AU_Length	0
8	instantBitrateLength	---
4	degradationPriorityLength	0
5	AU_seqNumLength	0
5	Packet_SeqNumLength	0
2	Reserved	3(0b11)

A.2.2 The case of a single audio object and a single video object in a broadcast stream

# of bits	Field Name	Value
ObjectDescriptorUpdate		
8	ObjectDescriptorUpdate tag	0x01
8	descriptor size	---
ObjectDescriptor(video)		
8	ObjectDescriptor tag	0x01
8	descriptor size	---
10	ObjectDescriptorID	20
1	URL_Flag	0
5	Reserved	31
ES_Descriptor(video)		
8	ES_Descriptor tag	0x03
8	descriptor size	---
16	ES_ID	201
1	streamDependenceFlag	0
1	URL_Flag	0
1	OCRstreamFlag	1
5	streamPriority	4
16	OCR_ES_ID	101
DecoderConfigDescriptor(video)		
8	DecoderConfigDescriptor tag	0x04
8	descriptor size	---
8	objectTypeIndication	0x21
6	streamType	4
1	Upstream	0
1	Reserved	1
24	bufferSizeDB	---
32	maxBitrate	---
32	avgBitrate	---
DecoderSpecificInfo(video)		
8	DecoderSpecificInfo tag	0x05
8	descriptor size	---
8*	DecoderSpecificInfo data	---
SLConfigDescriptor(video)		
8	SLConfigDescriptor tag	0x06
8	descriptor size	---
8	Predefined	0
1	useAccessUnitStartFlag	1
1	useAccessUnitEndFlag	1
1	useRandomAccessPointFlag	0
1	hasRandomAccessUnitsOnlyFlag	0
1	usePaddingFlag	0
1	useTimeStampsFlag	1
1	useIdleFlag	1
1	durationFlag	0
32	timeStampResolution	90000
32	OCRResolution	90000
8	timeStampLength	33
8	OCRLength	33
8	AU_Length	0
8	instantBitrateLength	---
4	degradationPriorityLength	0
5	AU_seqNumLength	0
5	Packet_SeqNumLength	0
2	Reserved	3(0b11)
ObjectDescriptor(audio)		
8	ObjectDescriptor tag	0x01
8	descriptor size	---
10	ObjectDescriptorID	10
1	URL_Flag	0
5	Reserved	31

# of bits	Field Name	Value
ES_Descriptor(audio)		
8	ES_Descriptor tag	0x03
8	descriptor size	---
16	ES_ID	101
1	streamDependenceFlag	0
1	URL_Flag	0
1	OCRstreamFlag	0
5	streamPriority	5
DecoderConfigDescriptor(audio)		
8	DecoderConfigDescriptor tag	0x04
8	descriptor size	---
8	objectTypeIndication	0x40
6	streamType	5
1	Upstream	0
1	Reserved	1
24	bufferSizeDB	---
32	maxBitrate	---
32	avgBitrate	---
DecoderSpecificInfo(audio)		
8	DecoderSpecificInfo tag	0x05
8	descriptor size	---
8*	DecoderSpecificInfo data	---
SLConfigDescriptor(audio)		
8	SLConfigDescriptor tag	0x06
8	descriptor size	---
8	Predefined	0
1	useAccessUnitStartFlag	1
1	useAccessUnitEndFlag	1
1	useRandomAccessPointFlag	0
1	hasRandomAccessUnitsOnlyFlag	0
1	usePaddingFlag	0
1	useTimeStampsFlag	1
1	useIdleFlag	1
1	durationFlag	0
32	timeStampResolution	90000
32	OCRResolution	90000
8	timeStampLength	33
8	OCRLength	33
8	AU_Length	0
8	instantBitrateLength	---
4	degradationPriorityLength	0
5	AU_seqNumLength	0
5	Packet_SeqNumLength	0
2	Reserved	3(0b11)

A.2.3 The case where multi-lingual audio is supported (audio 1/ audio 2/ video)

# of bits	Field Name	Value
ObjectDescriptorUpdate		
8	ObjectDescriptorUpdate tag	0x01
8	descriptor size	---
ObjectDescriptor(video)		
8	ObjectDescriptor tag	0x01
8	descriptor size	---
10	ObjectDescriptorID	20
1	URL_Flag	0
5	Reserved	31
ES_Descriptor(video)		
8	ES_Descriptor tag	0x03
8	descriptor size	---
16	ES_ID	201
1	streamDependenceFlag	0
1	URL_Flag	0
1	OCRstreamFlag	1
5	streamPriority	4
16	OCR_ES_ID	101
DecoderConfigDescriptor(video)		
8	DecoderConfigDescriptor tag	0x04
8	descriptor size	---
8	objectTypeIndication	0x21
6	streamType	4
1	Upstream	0
1	Reserved	1
24	bufferSizeDB	---
32	maxBitrate	---
32	avgBitrate	---
DecoderSpecificInfo(video)		
8	DecoderSpecificInfo tag	0x05
8	descriptor size	---
8*	DecoderSpecificInfo data	---
SLConfigDescriptor(video)		
8	SLConfigDescriptor tag	0x06
8	descriptor size	---
8	Predefined	0
1	useAccessUnitStartFlag	1
1	useAccessUnitEndFlag	1
1	useRandomAccessPointFlag	0
1	hasRandomAccessUnitsOnlyFlag	0
1	usePaddingFlag	0
1	useTimeStampsFlag	1
1	useIdleFlag	1
1	durationFlag	0
32	timeStampResolution	90000
32	OCRResolution	90000
8	timeStampLength	33
8	OCRLength	33
8	AU_Length	0
8	instantBitrateLength	---
4	degradationPriorityLength	0
5	AU_seqNumLength	0
5	Packet_SeqNumLength	0
2	Reserved	3(0b11)
ObjectDescriptor(audio1)		
8	ObjectDescriptor tag	0x01
8	descriptor size	---
10	ObjectDescriptorID	10
1	URL_Flag	0
5	Reserved	31

# of bits	Field Name	Value
ES_Descriptor(audio1)		
8	ES_Descriptor tag	0x03
8	descriptor size	---
16	ES_ID	101
1	streamDependenceFlag	0
1	URL_Flag	0
1	OCRstreamFlag	0
5	streamPriority	5
DecoderConfigDescriptor(audio1)		
8	DecoderConfigDescriptor tag	0x04
8	descriptor size	---
8	objectTypeIndication	0x40
6	streamType	5
1	Upstream	0
1	Reserved	1
24	bufferSizeDB	---
32	maxBitrate	---
32	avgBitrate	---
DecoderSpecificInfo(audio1)		
8	DecoderSpecificInfo tag	0x05
8	descriptor size	---
8*	DecoderSpecificInfo data	---
SLConfigDescriptor(audio1)		
8	SLConfigDescriptor tag	0x06
8	descriptor size	---
8	Predefined	0
1	useAccessUnitStartFlag	1
1	useAccessUnitEndFlag	1
1	useRandomAccessPointFlag	0
1	hasRandomAccessUnitsOnlyFlag	0
1	usePaddingFlag	0
1	useTimeStampsFlag	1
1	useIdleFlag	1
1	durationFlag	0
32	timeStampResolution	90000
32	OCRResolution	90000
8	timeStampLength	33
8	OCRLength	33
8	AU_Length	0
8	instantBitrateLength	---
4	degradationPriorityLength	0
5	AU_seqNumLength	0
5	Packet_SeqNumLength	0
2	Reserved	3(0b11)
ObjectDescriptor(audio2)		
8	ObjectDescriptor tag	0x01
8	descriptor size	---
10	ObjectDescriptorID	12
1	URL_Flag	0
5	Reserved	31
ES_Descriptor(audio2)		
8	ES_Descriptor tag	0x03
8	descriptor size	---
16	ES_ID	102
1	streamDependenceFlag	0
1	URL_Flag	0
1	OCRstreamFlag	0
5	streamPriority	5

# of bits	Field Name	Value
DecoderConfigDescriptor(audio2)		
8	DecoderConfigDescriptor tag	0x04
8	descriptor size	---
8	objectTypeIndication	0x40
6	streamType	5
1	Upstream	0
1	Reserved	1
24	bufferSizeDB	---
32	maxBitrate	---
32	avgBitrate	---
DecoderSpecificInfo(audio2)		
8	DecoderSpecificInfo tag	0x05
8	descriptor size	---
8*	DecoderSpecificInfo data	---
SLConfigDescriptor(audio2)		
8	SLConfigDescriptor tag	0x06
8	descriptor size	---
8	Predefined	0
1	useAccessUnitStartFlag	1
1	useAccessUnitEndFlag	1
1	useRandomAccessPointFlag	0
1	hasRandomAccessUnitsOnlyFlag	0
1	usePaddingFlag	0
1	useTimeStampsFlag	1
1	useIdleFlag	1
1	durationFlag	0
32	timeStampResolution	90000
32	OCRResolution	90000
8	timeStampLength	33
8	OCRLength	33
8	AU_Length	0
8	instantBitrateLength	---
4	degradationPriorityLength	0
5	AU_seqNumLength	0
5	Packet_SeqNumLength	0
2	Reserved	3(0b11)
LanguageDescriptor(audio2)		
8	LanguageDescriptor tag	0x47
8	descriptor size	---
24	languageCode	---

A.3 BIFS

A.3.1 The case of a single audio object within a broadcast stream

A.3.1.1 Syntax

```
OrderedGroup {
    children [
        Sound2D {source AudioSource {url 10}}
    ]
}
```

A.3.1.2 Coded data

```
C0 10 12 81   30 2A 05 7C
```

A.3.2 The case of a single audio object and a single video object within a broadcast stream

A.3.2.1 Syntax

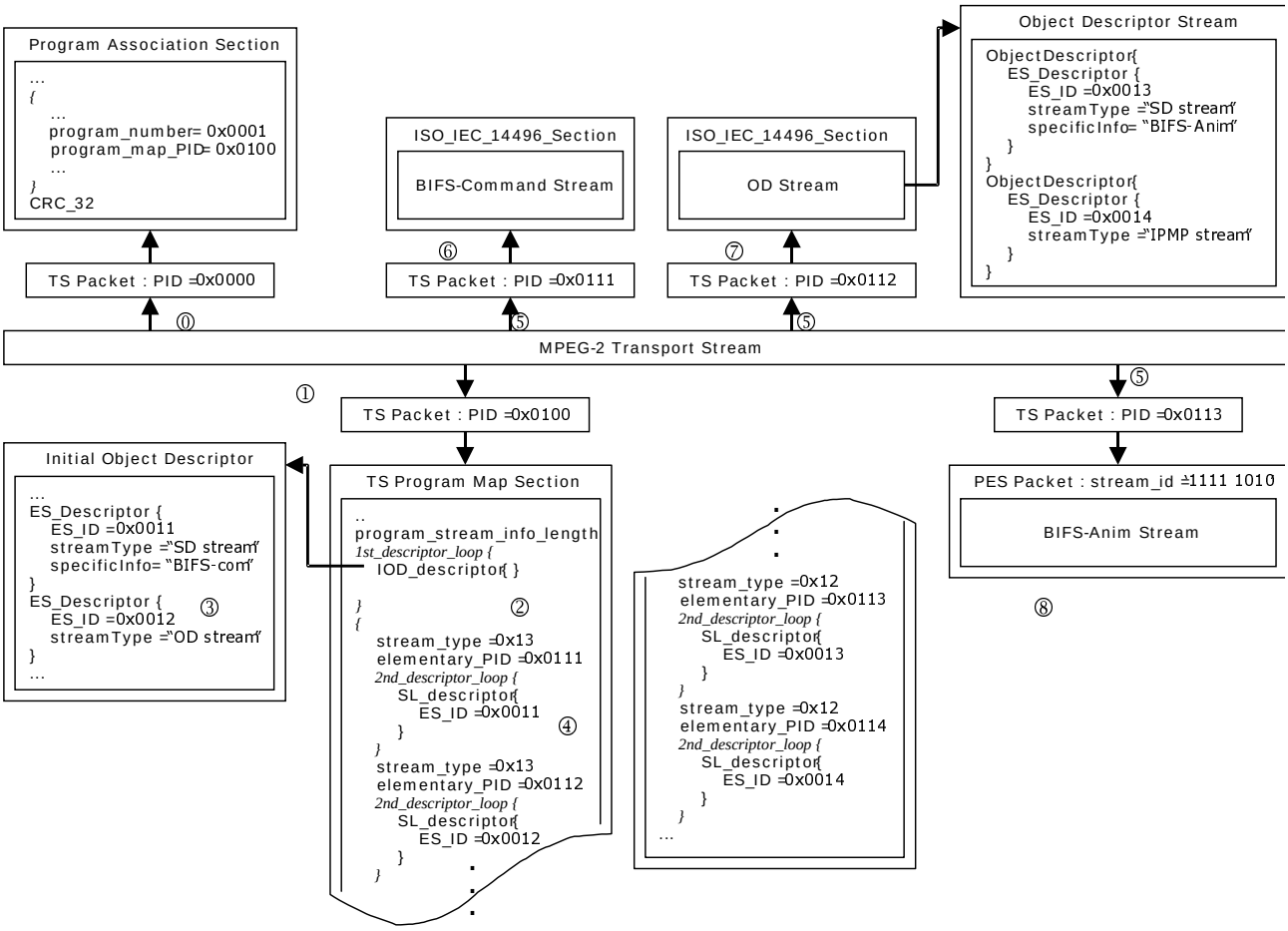
```
OrderedGroup {
    children [
        Sound2D {source AudioSource {url 10}}
        Shape {
            geometry Bitmap {}
            appearance Appearance {texture MovieTexture {url 20}}
        }
    ]
}
```

A.3.2.2 Coded data

```
C0 10 12 81   30 2A 05 72   61 04 88 50   45 05 3F 00
```

Annex B (informative): Content Access Procedure

The procedure by which content is accessed is explained using the example shown in figure B.1.



⓪ From an MPEG-2 TS, the PAT is obtained from the TS packets with PID 0x0000. Then a program number and the corresponding PMT PID are identified. In this case, there exists only one PMT in the PAT.

① Search for the PMT in the TS using the PMT PID obtained above.

② Within the PMT, find the IOD_descriptor that includes the IOD.

③ Within the IOD, find the ES_Descriptors related to the scene description and object description.

④ From the ES_Descriptor information found in ③, obtain the ES_IDs and then search for the elementary stream information corresponding to the ES_IDs through the descriptor loop in the PMT.

⑤ From the elementary stream information found in ④, obtain the PID and stream_type pairs corresponding to the ES_Ids, respectively, and then find packets corresponding to each of these in the transport stream.

⑥,⑦ By using the ObjectDescriptorID contained in the scene description information, obtain the corresponding ObjectDescriptor contained in the object descriptor stream. By using the relationship between the ES_ID contained in the ObjectDescriptor and the PID that can be found in a way similar to ⑤, compose the scene.

⑧ Obtain the animation data using the scene description information.

Figure B.1: MPEG-2 TS example that includes ISO/IEC 14496 contents [4]

History

Document history		
V1.1.1	June 2005	Publication