Working Draft:
Amendment to ATSC Digital Television Standard, Doc. A/53D, to Revise Annex C and Define E-VSB Transport

The Advanced Television Systems Committee, Inc., is an international, non-profit organization developing voluntary standards for digital television. The ATSC member organizations represent the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries.

Specifically, ATSC is working to coordinate television standards among different communications media focusing on digital television, interactive systems, and broadband multimedia communications. ATSC is also developing digital television implementation strategies and presenting educational seminars on the ATSC standards.

ATSC was formed in 1982 by the member organizations of the Joint Committee on InterSociety Coordination (JCIC): the Electronic Industries Association (EIA), the Institute of Electrical and Electronic Engineers (IEEE), the National Association of Broadcasters (NAB), the National Cable Television Association (NCTA), and the Society of Motion Picture and Television Engineers (SMPTE). Currently, there are approximately 140 members representing the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries.

ATSC Digital TV Standards include digital high definition television (HDTV), standard definition television (SDTV), data broadcasting, multichannel surround-sound audio, and satellite direct-to-home broadcasting.

Replace Annex C of A/53D with the following:

Annex C:
Service Multiplex and Transport Subsystem Characteristics
(Normative)

1. SCOPE

This Annex constitutes the normative specification for the transport subsystem of the Digital Television Standard. The syntax and semantics of this specification conform to ISO/IEC 13818-1 [C3], with additional constraints and conditions specified in this Standard. Within this context, other ATSC Standards may further constrain and/or supplement the transport subsystem specification.

2. NORMATIVE REFERENCES

The following documents contain provisions that in whole or in part, through reference in this text, constitute provisions of this Standard. At the time of publication, the editions indicated were valid. For references not including a publication date, the most recent published version shall apply. All standards are subject to revision and amendment, and parties to agreements based on this Standard are encouraged to investigate the possibility of applying the most recent editions of the documents listed below.

2.1   Normative References

[C1]   ATSC A/52B: "Digital Audio Compression (AC-3, E-AC-3) Standard," Advanced Television Systems Committee, Washington, D.C.

[C2]   ATSC Standard A/65C: "Program and System Information Protocol for Terrestrial Broadcast and Cable," Advanced Television Systems Committee, Washington, D.C.

[C3]   ISO/IEC IS 13818-1:2000 (E), International Standard, Information technology – Generic coding of moving pictures and associated audio information: systems.

2.2   Informative References

[C4]   ATSC A/90: "Data Broadcast Standard," Advanced Television Systems Committee, Washington, D.C.

3. DEFINITION OF TERMS

PAT-E – A table with the same syntax as Program Association Table as defined by [C3] transmitted using an enhanced VSB mode defined in Annex D.

PMT-E – A table with the same syntax as Program Map Table as defined by [C3] transmitted using an enhanced VSB mode defined in Annex D.

PMT-E_PID – A PID that identifies the Transport Stream packets that carry TS_program_map_section()s in a TS-E.

Program – Program shall mean the collection of all elements within the emission that have the same program number, independent of the methods used to propagate the program elements.

Linked – Alternative elements of a Program are 'Linked' when they have identical values in the linked_component_tag field of their Enhanced Signaling Descriptors.

TS-M – The portion of TS-R that contains only all Transport Stream packets transmitted by the main mode (see Annex D) (see Figure C3).

TS-R – The recombined Transport Stream containing all Transport Stream packets delivered by all transmission modes (main, one-half rate and one-quarter rate) (see Annex D) (see Figure C3).

TS-E – The portion of TS-R that contains only all Transport Stream packets transmitted by one-half rate and/or one-quarter rate modes (see Annex D) (see Figure C3).

TS-Ea – The portion of TS-E that contains only all Transport Stream packets transmitted by one-half rate mode (see Annex D) (see Figure C3).

TS-Eb – The portion of TS-E that contains only all Transport Stream packets transmitted by one-quarter rate mode (see Annex D) (see Figure C3).

4. COMPLIANCE NOTATION

As used in this document, "shall" denotes a mandatory provision of the standard. "Should" denotes a provision that is recommended but not mandatory. "May" denotes a feature whose presence does not preclude compliance that may or may not be present at the option of the implementer. Note: this compliance notation applies only to this annex.

5. SYSTEM OVERVIEW

The transport format and protocol for the Digital Television Standard is a compatible subset of the MPEG-2 Systems specification defined in ISO/IEC 13818-1 [C3]. It is based on a fixed-length packet transport stream approach that has been defined and optimized for digital television delivery applications.

As illustrated in Figure C1, the transport subsystem resides between the application (e.g., audio or video) encoding and decoding functions and the transmission subsystem. The encoder's transport subsystem is responsible for formatting the coded elementary streams and multiplexing the different components of the program for transmission. It also is responsible for delivering packets intended for transmission using coding methods defined in Annex D. A receiver recovers the elementary streams for the individual application decoders and for the corresponding error signaling. The transport subsystem also incorporates other higher protocol layer functionality related to properly timing the packets to enable receiver synchronization.
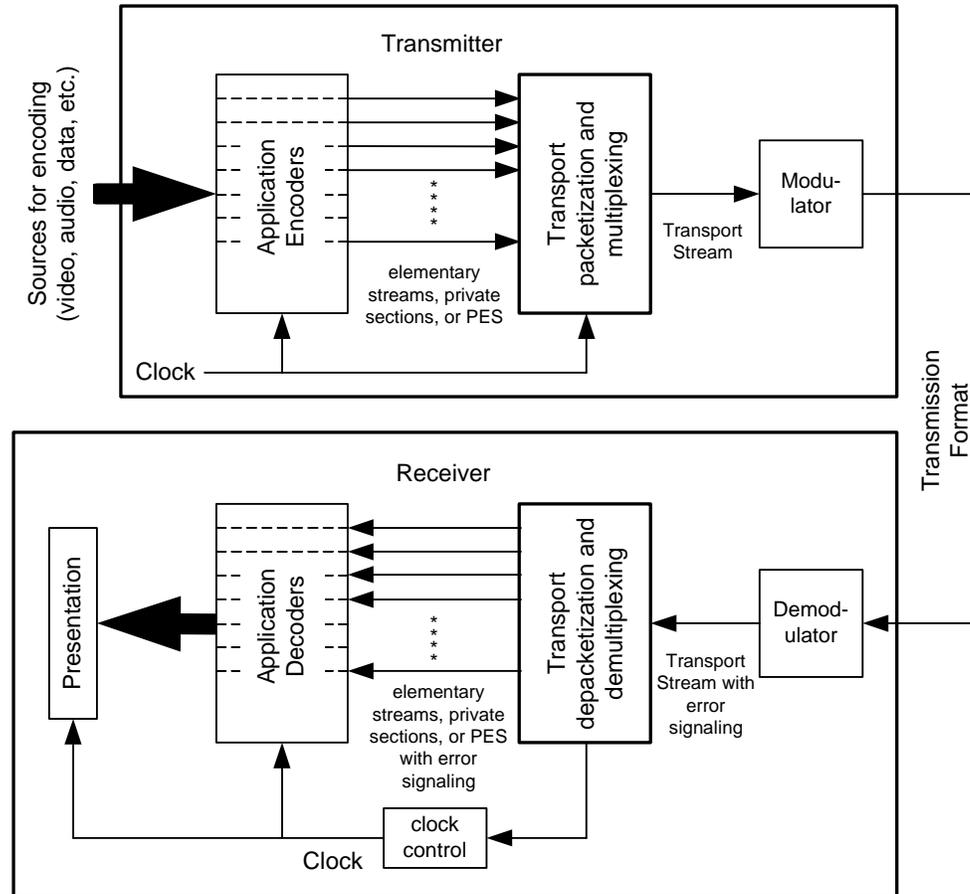
Figure C1 Sample organization of functionality in a transmitter-receiver pair for a single program.

One approach to describe the system multiplexing approach is to consider it as a combination of multiplexing at two different layers. In the first layer, single program transport streams are formed by multiplexing Transport Stream (TS) packets from one or more Packetized Elementary Stream (PES) and/or private section (ISO/IEC 13818-1 [C3] Table 2-30) sources. In the second layer, one or more single program transport streams are combined to form a service multiplex of programs (also known as a multi-program transport stream in the MPEG-2 Systems standard, and a Digital Television Standard multiplexed bit stream in this ATSC standard). Program Specific Information (PSI), carried within Transport Stream packets, relates to the identification of programs and the components of each program.

Not shown explicitly in Figure C1, but essential to the practical implementation of this Standard, is a control system that manages the transfer and processing of the elementary streams from the application encoders. The rules followed by this control system are not a part of this Standard. The output of the control system implementation shall conform to the MPEG-2 Transport Stream coding as specified in ISO/IEC 13818-1 [C3] with the additional constraints specified in this Standard.

In specifying the characteristics of the signal processing required of a system based upon this Standard, the functions to be performed may be described in terms of encoders and multiplexers at the transmitting end and in terms of a Reference Receiver at the receiving end. A Reference

Receiver is defined such that it performs the functions described in the way that they are described herein, which may or may not be the way in which actual receivers, built to this Standard, are designed to operate. Some mandatory transmission requirements are stated herein in terms of the resulting transmission stream elements that are required to be produced, through use of the Reference Receiver functional definition.

6. SPECIFICATION

This section of the Standard describes the coding constraints that apply to the use of the MPEG-2 systems specification [C3] in the digital television system including mandatory main and optional enhanced services.

6.1   MPEG-2 Systems Standard

The transport subsystem shall comply with the transport stream definition of the MPEG-2 Systems standard as specified in ISO/IEC 13818-1 [C3] and shall be further constrained as specified herein. Program shall mean the collection of all elements within the emission that have the same value of MPEG-2 program_number, independent of the methods used to propagate the program elements.

6.1.1   Video T-STD

Video streams in TS-R (Figure C3) shall conform to the T-STD as defined in Sections 2.4.2.2 and 2.4.2.3 of ISO/IEC 13818-1 [C3] and shall follow the constraints for the level encoded in the video elementary stream. When there is a video stream of stream_type 0x02 in TS-R, the T-STD buffer Bn defined in ISO/IEC 13818-1 [C3], Section 2.4.2 shall apply[1] for such a stream.

When a TS-E is present, video streams of stream_type 0x02 may need to be constructed with constraints on the buffer size of the T-STD as defined in Section 2.4.2 of ISO/IEC 13818-1 [C3] so that the variable time delay caused by the E-VSB system does not result in a violation of the requirement that the TS-R conform per the preceding paragraph.

When there is a video stream of stream_type 0x02 in TS-E, the T-STD buffer Bn defined in ISO/IEC 13818-1 [C3] Section 2.4.2 for the stream in TS-E may underflow when the low_delay flag in the video sequence header is set to '1' (per ISO/IEC 13818-1 [C3], Section 2.4.2.6). When the low_delay flag in the video sequence header is set to '1' for this TS-E stream, [C3] at Section 2.4.2.6 requires that the delay of data through the STD buffer defined by $td_n(j) - t(i)$ is <= 1 second except for still picture video data where $td_n(j) - t(i)$ is <= 60 sec, for all bytes contained in access unit j. The terms $td_n(j)$, $t(i)$, and j are those defined in ISO/IEC 13818-1 [C3], Section 2.4.2. Any elementary stream containing Still Picture data shall include a video_stream_descriptor() in accordance with ISO/IEC 13818-1 [C3] Section 2.6.2 and shall have the value of the field still_picture_flag set to '1' and the interval between I frames shall not be greater than 60 seconds.

6.1.2   Audio T-STD

6.1.2.1   Audio T-STD When TS-E is Not Present

The audio T-STD shall comply with Section 3.6 of Annex A of ATSC Standard A/52B [C1].

---

[1] Section 7.2 also constrains video streams of stream_type 0x02 in TS-M.

6.1.2.2   Audio T-STD When TS-E is Present

Audio streams of stream_type 0x81 in TS-R (Figure C3) shall conform to the T-STD as defined in Sections 2.4.2 and 2.4.2.3 of ISO/IEC 13818-1 [C3], with the buffer size BSn = 2592 bytes.

Audio streams of stream_type 0x87 (enhanced AC-3) shall conform to the T-STD as defined in Sections 2.4.2 and 2.4.2.3 of ISO/IEC 13818-1 [C3], with the buffer size BSn = 5184 bytes.

6.1.3   Program Constraints

This section standardizes how to carry Programs[2] in the ATSC system. Each Program shall be constrained to contain certain standardized elementary streams, which are dependent on the service_type[3] associated with that Program and the use of TS-E. Programs may also contain private elementary streams as defined herein and/or by other ATSC Standards.

6.1.3.1   ATSC Digital Television (service_type 0x02)

For Programs of service_type 0x02, if there is a video program element in the TS-E, the stream_type for that video program element shall be 0x02 and the video Elementary Stream shall be based on the same video content[4] as a video program Elementary Stream in TS-M. In addition, for any given Program of service_type 0x02:

- There shall be at most one video program element present in TS-M.
- There shall be at most one video program element present in TS-Ea.
- There shall be at most one video program element present in TS-Eb.

If there is one or more audio program element in the TS-E, the stream_type for each audio Elementary Stream in TS-E shall be 0x81 or 0x87 and each program element shall be based upon the same audio source as an audio element of the Program in TS-M.

6.1.3.2   ATSC Audio (service_type 0x03)

For Programs of service_type 0x03 that include audio program elements in both TS-M and TS-E, the stream_type for each audio Elementary Stream in TS-E shall be 0x81 or 0x87 and each program element shall be based upon the same audio source as an audio element of the Program in TS-M.

6.1.3.3   Unassociated/Small Screen Service (service_type 0x06)

An additional type of television service is reserved for use in TS-E called "unassociated/small screen service," which shall use service_type 0x06. An unassociated/small screen service carries programming whose contents are not based on the same schedule of programming as that carried in TS-M. Programs of service_type 0x06 shall be announced in TS-E and all program elements shall be carried only in TS-E. Further constraints on program elements for service_type 0x06 are to be defined at a future date.

---

[2] Program is defined in Section 3.

[3] This field is defined in A/65C [C2], Section 6.3.1 and Table 6.7.

[4] The video in the TS-E may be a different aspect ratio, frame rate, and/or bitrate than that in TS-M.

6.2   Identification of MPEG-2 Private Ranges

ATSC defines code points in the MPEG-2 user private range and may define code points private to ATSC users within this range.

6.2.1   MPEG-2 Registration Descriptor

Under circumstances as defined below, this Standard uses the MPEG-2 Registration Descriptor described in Sections 2.6.8 and 2.6.9 of ISO/IEC 13818-1 [C3] to identify the contents of programs and program elements to decoding equipment. No more than one MPEG-2 Registration Descriptor shall appear in any given descriptor loop.

The presence of an MPEG-2 Registration Descriptor in any descriptor loop shall not affect the meaning of any other descriptor(s) in the same descriptor loop. The ATSC Private Information Descriptor (defined in Section 6.8.4) shall be the method to carry descriptor-based information associated with a private entity.

6.2.1.1   Program Identifier

Programs that conform to ATSC standards may be identified by the use of an MPEG-2 Registration Descriptor (as defined in Sections 2.6.8 and 2.6.9 of ISO/IEC 13818-1 [C3]). When present for this purpose, the MPEG-2 Registration Descriptor shall be placed in the descriptor loop immediately following the program_info_length field of the TS_program_map_section() describing this program and the format_identifier field of this MPEG-2 Registration Descriptor shall have a value 0x4741 3934 ("GA94" in ASCII).

6.2.1.2   Audio Elementary Stream Identifier

Audio elementary streams of stream_type 0x81 that conform to ATSC standards may use an MPEG-2 Registration Descriptor (as defined in Sections 2.6.8 and 2.6.9 of ISO/IEC 13818-1 [C3]). When present for this purpose, the MPEG-2 Registration Descriptor shall be placed in the descriptor loop immediately following the ES_info_length field in the TS_program_map_section() for each program element of stream_type 0x81 (AC-3 audio) and the format_identifier field of the MPEG-2 Registration Descriptor shall have a value 0x4143 2D33 ("AC-3" in ASCII).

6.2.1.3   Other Program Element Identifiers

Any program element carrying content not described by an approved ATSC standard shall be identified with an MPEG-2 Registration Descriptor (as defined in Sections 2.6.8 and 2.6.9 of ISO/IEC 13818-1 [C3]). The format_identifier field of the MPEG-2 Registration Descriptor shall be registered with the SMPTE Registration Authority, LLC[5]. The descriptor shall be placed in the descriptor loop immediately following the ES_info_length field in the TS_program_map_section() for each such non-standard program element.

6.3   Audio Constraints

For each Program of service_type 0x02 delivered in TS-M, if audio is present, at least one audio component shall be a complete main audio service (CM) as defined in Section 6 of Annex B.

For each Program of service_type 0x03 delivered in TS-M, at least one audio component shall be a complete main audio service (CM) as defined in Section 6 of Annex B.

---

[5] The ISO/IEC-designated registration authority for the format_identifier is SMPTE Registration Authority, LLC. See (http://www.smpte-ra.org/mpegreg.html).

For each Program of service_type 0x02 completely delivered in TS-E, if audio is present, at least one audio component shall be a complete main audio service (CM) as defined in Section 6 of Annex B.

For each Program of service_type 0x03 completely delivered in TS-E, at least one audio component shall be a complete main audio service (CM) as defined in Section 6 of Annex B.

Program elements within one Program with the same value of linked_component_tag (defined in Section 6.7.5) shall be of the same audio type (for example CM, VI, or HI) and shall be one of the services defined in section 6 of Annex B.

6.4 Constraints on PSI

All Program elements in the Transport Stream shall be described in the PSI.

6.4.1 Constraints on Main Services' Program Specific Information

There are the following constraints on the PSI information in TS-M:

- Transport Stream packets identified by a particular PMT_PID value shall be constrained to carry only one program definition, as described by a single TS_program_map_section(). For terrestrial broadcast applications, these Transport Stream packets shall be further constrained to carry no other kind of PSI table.

- The Transport Stream shall be constructed such that the time interval between the byte containing the last bit of the TS_program_map_section() containing television program information and successive occurrences of the same TS_program_map_section() shall be less than or equal to 400 milliseconds.

- The program numbers are associated with the corresponding PMT_PIDs in the Program Association Table (PAT). The Transport Stream shall be constructed such that the time interval between the byte containing the last bit of the program_association_section() and successive occurrences of the program_association_section() shall be less than or equal to 100 milliseconds. However, when program_association_section()s, CA_section()s, and TS_program_map_section()s are approaching their maximum allowed sizes, the potential exists to exceed the 80,000 bps rate specified in ISO/IEC 13818-1 [C3] Section 2.4.2.3. In cases where the table section sizes are such that the 100 millisecond repetition rate of the program_association_section() would cause the 80,000 bps maximum rate to be exceeded, the time interval between the byte containing the last bit of the program_association_section() may be increased but in no event shall exceed 140 milliseconds, so that under no circumstances the limit of 80,000 bps is exceeded.

- When an Elementary Stream of stream_type 0x02 (MPEG-2 video) is present in the Transport Stream, the data_stream_alignment_descriptor() (described in Section 2.6.10 of ISO/IEC 13818-1 [C3]) shall be included in the descriptor loop immediately following the ES_info_length field in the TS_program_map_section() describing that Elementary Stream. The descriptor_tag value shall be set to 0x06, the descriptor_length value shall be set to 0x01, and the alignment_type value shall be set to 0x02 (video access unit).

- Adaptation headers shall not occur in Transport Stream packets identified by a program_map_PID value for purposes other than for signaling with the discontinuity_indicator that the version_number (Section 2.4.4.9 of ISO/IEC 13818-1 [C3]) may be discontinuous.

- Adaptation headers shall not occur in Transport Stream packets identified by PID 0x0000 (the PAT PID) for purposes other than for signaling with the discontinuity_indicator that the version_number (Section 2.4.4.5 of ISO/IEC 13818-1 [C3]) may be discontinuous.

- This Standard does not define a Network Information Table (NIT) as specified in MPEG-2 Systems. The use of program_number 0x0000 should be avoided as MPEG-2 Systems reserves this value for the network_PID, which in turn is used to identify the TS packets of a NIT.

6.4.2   PAT-E

Sections of the PAT-E shall be carried in one or more program_association_section()s within TS-E packets and shall be identified by the PID 0x1FF7 and shall only reference PMT_PID values for TS_program_map_section()s present in Transport Stream packets in TS-E. Sections of the PAT-E shall be present only in the TS-Eb when TS-Eb is present.

6.4.3   PMT-E

Sections of PMT-E shall be carried in TS_program_map_section()s within TS-E packets identified by the PAT-E and shall only reference PID values for packets that are present in the TS-E. Sections of PMT-E shall be present only in the TS-Eb when TS-Eb is present.

6.4.4   Multiple PCR_PIDs per Program and a Common Time Base

For every Program that includes program elements in both TS-M and TS-E, there is a TS_program_map_section() carried in TS-M and a TS_program_map_section() carried in TS-E with a common value of program_number. For such a Program, each of the PCRs referenced by the PCR_PID fields in the two TS_program_map_section()s shall be samples from a common time clock. Note that a PCR_PID value of 0x1FFF has the same meaning as in ISO/IEC13818-1 [C3] Section 2.4.4.9.

6.4.5   Constraints on Optional Enhanced Services' Program Specific Information

There are the following constraints on the PSI information in TS-E:

- Only one Program shall be described in a transport bit stream for TS_program_map_section()s carried in packets with a particular PMT-E_PID value. Transport stream packets containing a TS_program_map_section() and identified with a particular PID value shall not be used to transmit any other kind of PSI table (identified by a different table_id).

- TS-E should be constructed such that the time interval between the byte containing the last bit of the TS_program_map_section() containing television program information and successive occurrences of the same TS_program_map_section() should be less than or equal to 1600 milliseconds.

- The program numbers referenced in the PAT-E shall be associated with the corresponding PMT-E_PIDs in the collection of program_association_section()s, which are sections of the PAT-E. The sections that make up the PAT-E shall be carried in packets identified with the PID 0x1FF7. The TS-E should be constructed such that the time interval between the byte containing the last bit of the program_association_section() and successive occurrences of the program_association_section() should be less than or equal to 800 milliseconds.

- When the video elementary stream_type is equal to 0x02 the descriptor loop immediately following   ES_info_length   in   the   TS_program_map_section()   shall   contain   the

data_stream_alignment_descriptor() described in Section 2.6.10 of ISO/IEC 13818-1[C3] and the alignment_type field shown in Table 2-47 of ISO/IEC 13818-1 shall be 0x02.

- Adaptation headers shall not occur in Transport Stream packets of the PMT-E_PID for purposes other than for signaling with the discontinuity_indicator that the version_number (Section 2.4.4.5 of ISO/IEC 13818-1) may be discontinuous.

6.5   PES Constraints

Packetized Elementary Stream syntax and semantics shall be used to encapsulate the audio and video elementary stream information. The Packetized Elementary Stream syntax is used to convey the Presentation Time-Stamp (PTS) and Decoding Time-Stamp (DTS) information required for decoding audio and video information with synchronism. This section describes the coding constraints on this MPEG-2 Systems layer.

Within the PES packet header, the following restrictions shall apply:

- PES_scrambling_control shall be coded as '00'.
- ESCR_flag shall be coded as '0'.
- ES_rate_flag shall be coded as '0'.
- PES_CRC_flag shall be coded as '0'.

Within the PES packet extension, the following restrictions shall apply.

- PES_private_data_flag shall be coded as '0'.
- pack_header_field_flag shall be coded as '0'.
- program_packet_sequence_counter_flag shall be coded as '0'.
- P-STD_buffer_flag shall be coded as '0'.

6.5.1   MPEG-2 Video PES Constraints (for streams of stream_type 0x02)

Each PES packet shall begin with a video access unit, as defined in Section 2.1.1 of ISO/IEC 13818-1 [C3], which shall be aligned with the PES packet header. The first byte of a PES packet payload shall be the first byte of a video access unit. Each PES header shall contain a PTS. Additionally, it shall contain a DTS as appropriate. For terrestrial broadcast, the PES packet shall not contain more than one coded video frame, and shall be void of video picture data only when transmitted in conjunction with the discontinuity_indicator to signal that the continuity_counter may be discontinuous.

Within the PES packet header, the following restrictions apply:

- The PES_packet_length shall be coded as '0x0000'
- data_alignment_indicator shall be coded as '1'

6.5.2   AC-3 Audio PES Constraints (for Streams of stream_type 0x81)

The AC-3 audio decoder may be capable of simultaneously decoding more than one elementary stream containing different program elements, and then combining the program elements into a complete program. In this case, the audio decoder may sequentially decode audio frames (or audio blocks) from each elementary stream and do the combining (mixing together) on a frame or (block) basis. In order to have the audio from the two elementary streams reproduced in exact sample synchronism, it is necessary for the original audio elementary stream encoders to have encoded the two audio program elements frame synchronously; i.e., if audio program 1 has

sample 0 of frame *n* at time *t*0, then audio program 2 should also have frame *n* beginning with its sample 0 at the identical time *t*0. If the encoding is done frame synchronously, then matching audio frames should have identical values of PTS.

If PES packets from two AC-3 audio services that are to be decoded simultaneously contain identical values of PTS then the corresponding encoded audio frames contained in the PES packets should be presented to the audio decoder for simultaneous synchronous decoding. If the PTS values do not match (indicating that the audio encoding was not frame synchronous) then the audio frames that are closest in time may be presented to the audio decoder for simultaneous decoding. In this case, the two services may be reproduced out of sync by as much as one-half of a frame time (which is often satisfactory; e.g., a voice-over does not require precise timing).

The value of stream_id for AC-3 shall be '1011 1101' (private_stream_1).

6.5.3   Audio PES Constraints for Enhanced AC-3 (stream_type = 0x87)

The value of stream_id for enhanced AC-3 shall be '1011 1101' (private_stream_1).

If an audio service is delivered in TS-E, and is Linked to an audio service delivered in TS-M, constraints are required on the content of the corresponding audio access units, and on the PTS values in the corresponding PES packets, of the two Linked audio services in order to enable a receiver to perform sample time synchronous switching between the two audio services (e.g., fallback audio).

The constraints are:

1) That the audio access units of AC-3 in TS-M, and the corresponding audio access units of enhanced AC-3 in TS-E shall be constructed from sets of input audio samples that are time synchronous; and

2) That the presentation times corresponding to the corresponding audio access units shall be identical.

Note that not every audio access unit has an explicit value of PTS in a PES because there is only one PTS per PES while there can be multiple audio access units per PES. However, there is a unique presentation time for each audio access unit, and it is the presentation times that must be identical to enable a receiver to switch between the two audio streams with exact sample synchronism.

6.6   Services and Features

6.6.1   System Information and Program Guide

Transport Streams shall include system information and program guide data formatted according to the structure and syntax described in ATSC Standard A/65C "Program and System Information Protocol for Terrestrial Broadcast and Cable" [C2]. System information and program guide data shall be conveyed in Transport Stream packets of PID 0x1FFB, which shall be reserved exclusively for this purpose. System information provides data necessary for navigation among digital service offerings. The program guide database allows a receiver to build an on-screen grid of program information for the various services that may be available.

6.6.1.1   SI base_PID for TS-M

System information and program guide data shall be conveyed in Transport Stream packets of PID 0x1FFB, which shall be reserved exclusively within TS-R (Figure C3) for this purpose.

6.6.1.2　SI base_PID for TS-E

System information and program guide data shall be conveyed in TS-E packets of PID 0x1FF9, which shall be reserved exclusively within TS-R (Figure C3) for this purpose.

6.6.1.3　System Information and Program Guide STD Model

The STD model for program guide and system information is specified in ATSC Standard A/65C [C2].

6.6.2　Specification of ATSC Private Data

Within the ATSC set of standards, private data may be transported by various means:

1) **Data services** – Carriage of ATSC data services including system information shall be as documented in applicable ATSC Standards. See for example the ATSC A/90 Data Broadcast Standard [C4].

2) **Private program elements** – The stream_type codes in the range 0xC4 to 0xFF shall be available for stream types defined privately (not described by ATSC Standards). Such privately-defined program elements are associated with an MPEG-2 Registration Descriptor (see Section 6.2.1.3).

3) **Adaptation fields** – Private data may be transmitted within the adaptation field of Transport Stream packets (Sections 2.4.3.4 and 2.4.3.5 of ISO/IEC 13818-1 [C3]). Program elements that include private data in the adaptation fields of their Transport Stream packets shall be associated with an MPEG-2 Registration Descriptor (see Section 6.2.1.3).

6.7　Assignment of Identifiers

In this section, those identifiers and codes that shall have a fixed value are summarized. These include PES Stream IDs and Descriptors. Stream_type codes for program element types managed by the ATSC Code Points Registrar (currently assigned or available for future assignment) shall be in the range 0x80 to 0xC3. Descriptor_tag code 0x81 has already been assigned within the Digital Television Standard (see Section 6.7.1). Those descriptor_tag codes managed by the ATSC Code Points Registrar (currently assigned or available for future assignment) shall be in the range 0x40 to 0xEF.

6.7.1　AC-3 Audio Stream Type

The stream_type value for AC-3 audio program elements shall be 0x81.

6.7.2　MPEG-2 Video Stream Type

The stream_type value for MPEG-2 video program elements shall be as defined in ISO/IEC 13818-1 [C3] which value is 0x02.

6.7.3　Enhanced AC-3 Audio Stream Type

The stream_type value for the enhanced AC-3 audio program element defined in A/52B [C1] Annex E shall be 0x87.

6.8　Descriptors

Unless explicitly stated to the contrary for a given descriptor, no more than one descriptor with a given value of descriptor_tag shall appear in any descriptor loop.

6.8.1   AC-3 Audio Descriptor

When an Elementary Stream of stream_type 0x81 (AC-3 audio) or stream_type 0x87 (E-AC-3) is present in the digital television Transport Stream, an AC-3 Audio Descriptor (AC-3_audio_stream_descriptor()) shall be included in the descriptor loop immediately following the ES_info_length field in the TS_program_map_section() describing that Elementary Stream. The syntax shall be as given in Table A2 of Annex A of ATSC Standard A/52B [C1]. The following constraints shall apply to the AC-3 Audio Descriptor:

1) The value of the descriptor_tag shall be 0x81.

2) The 6-bit field for Bit Rate Code shall have a value in the range '000000' through '001111' or '100000' through '101111'; i.e., signaling a bit rate less than or equal to 448 kbps.

3) The num_channels field shall have a value in the range 1 to 13.

4) The langcod field is a reserved field. Audio language shall be indicated using an ISO-639 Language Descriptor (see Section 6.8.3).

5) The descriptor shall identify the type of the audio service in the bsmod field, which shall be the same as the bsmod field in the elementary stream associated with this descriptor.

6) The descriptor may optionally carry a 3-byte language code that is represented per ISO-639. If this language code is present in the AC-3_audio_stream_descriptor(), it shall match the language code carried in the ISO_639_language_descriptor(), if present.

Effective *<two years from date of publication by ATSC, rounded to nearest month>*, audio language, shall be indicated by including the optional ISO-639 Language bytes within the AC-3_audio_stream_descriptor().

Effective *<two years from date of publication by ATSC, rounded to nearest month>*, the use of the ISO-639 Language Descriptor to indicate language shall be optional.

> *Informative note:* Receiving devices are expected to use the bsmod (bit stream mode) field in the AC-3_audio_stream_descriptor() to determine the type of each AC-3 or E-AC-3 audio stream rather than the audio_type field in the ISO_639_language_descriptor().

6.8.2   Program Smoothing Buffer Descriptor

The TS_program_map_section() of each program shall contain a smoothing buffer descriptor pertaining to that program in accordance with Section 2.6.30 of ISO/IEC 13818-1. During the continuous existence of a program, the value of the elements of the smoothing buffer descriptor shall not change.

The fields of the smoothing buffer descriptor shall meet the following constraints:

- The field sb_leak_rate shall be allowed to range up to the maximum transport rates specified in Section 7.2.

- The field sb_size shall have a value less than or equal to 2048. The size of the smoothing buffer is thus ≤ 2048 bytes.

6.8.3   ISO-639 Language Descriptor

In the ATSC Digital Television System, the ISO_639_language_descriptor() defined in ISO/IEC 13818-1 [C3] Section 2.6.18 shall be used to indicate the language of AC-3 audio (stream_type 0x81) or enhanced AC-3 (stream_type 0x87) Elementary Stream components.

Effective *<two years from date of publication by ATSC, rounded to nearest month>*, audio language, when indicated, shall be indicated by including the ISO-639 Language bytes within the AC-3_audio_stream_descriptor().

Effective *<two years from date of publication by ATSC, rounded to nearest month>*, the use of the ISO-639 Language Descriptor to indicate language shall be optional.

When used, the ISO_639_language_descriptor() shall be included in the descriptor loop immediately following the ES_info_length field in the TS_program_map_section() for each Elementary Stream of stream_type 0x81 (AC-3 audio) or 0x87 (enhanced AC-3 audio) or when the number of AC-3 or enhanced AC-3 audio Elementary Streams in the TS_program_map_section() having the same value of bit stream mode (bsmod in the AC-3 Audio Descriptor) is two or more.

> *Informative note:* As an example, consider an MPEG-2 program that includes two AC-3 audio ES components: a AC-3 Complete Main (CM) audio track (bsmod = 0) and a Visually Impaired (VI) audio track (bsmod = 2). Inclusion of the ISO_639_language_descriptor() is optional for this program. If a second CM track were to be added, however, it would then be necessary to include ISO_639_language_descriptor()s in the TS_program_map_section().

The audio_type field in any ISO_639_language_descriptor() used in this Standard shall be set to 0x00 (meaning "undefined").

An ISO_639_language_descriptor() may be present in the TS_program_map_section() in other positions as well, for example to indicate the language or languages of a textual data service program element.

6.8.4   ATSC Private Information Descriptor

The ATSC_private_information_descriptor() provides a method to carry and unambiguously label private information. More than one ATSC_private_information_descriptor() may appear within a single descriptor loop. Table C1 defines the bit-stream syntax of the ATSC_private_information_descriptor().

Table C1 ATSC Private Information Descriptor

Syntax	No. of Bits	Format
ATSC_private_information_descriptor() {		
descriptor_tag	8	0xAD
descriptor_length	8	uimsbf
format_identifier	32	uimsbf
for (i = 0; i < N; i++) {		
private_data_byte	8	bslbf
}		
}		

descriptor_tag – This 8-bit field is set to 0xAD.

descriptor_length – This 8-bit field specifies the number of bytes of the descriptor immediately following the descriptor_length field.

format_identifier – The format_identifier is a 32-bit field as defined in ISO/IEC 13818-1 [C3], Section 2.6.9 for the registration_descriptor(). Only format_identifier values registered and recognized by the SMPTE Registration Authority, LLC shall be used (see http://www.smpte-ra.org/mpegreg.html)[6]. Its use in this descriptor shall scope and identify only the private information contained within this descriptor.

private_data_byte – The syntax and semantics of this field is defined by the assignee of the format_identifier value.

6.8.5 Enhanced Signaling Descriptor

The Enhanced Signaling Descriptor identifies the method of terrestrial broadcast transmission (main, one-half -rate, one-quarter rate)[7] of a program element, and when a program element is an alternative to another program element, is Linked to the alternative and indicates the broadcaster's preference. For example, an Enhanced Signaling Descriptor may indicate that although an audio stream is the same language and content as another, it is preferred for some reason.

The Enhanced Signaling Descriptor may be included in either descriptor loop in the TS_program_map_section() structure.

When the Enhanced Signaling Descriptor is only in the descriptor loop in the TS_program_map_section() just following the program_info_length field, the linkage_preference field shall have the value '00'. In this case the Enhanced Signaling Descriptor shall indicate the transmission method (tx_method) for every program element described in the TS_program_map_section() that does not have a separate instance of the Enhanced Signaling Descriptor included in the descriptor loop immediately following the ES_info_length field (see precedence rule below).

The Enhanced Signaling Descriptor shall be present for each program element that is Linked (linkage_preference not equal to '00') to another program element. Therefore the Enhanced Signaling Descriptor appears in TS_program_map_section()s in the TS-M, TS-Ea and TS-Eb to establish linkage.

When the Enhanced Signaling Descriptor is in the descriptor loop immediately following the ES_info_length field in the TS_program_map_section() then:

- No more than one alternative program element shall be referenced for any given program element.
- If the program element has the linkage_preference value of '00,' then the program element is not Linked.
- If the program element has the linkage_preference value of '01,' then the Linked alternative program element shall have the linkage_preference value of '01' and the two shall have identical values for program_number and shall have identical values for linked_component_tag.

---

[6] SMPTE Registration Authority, LLC, 595 West Hartsdale Avenue, White Plains, NY 10607 USA

[7] See Annex D for definitions.

- If the program element has the linkage_preference value of '10,' then the Linked alternative program element shall have the linkage_preference value of '11' and the two shall have identical values for program_number and shall have identical values for linked_component_tag.
- If the program element has the linkage_preference value of '11,' then the Linked alternative program element shall have the linkage_preference value of '10' and the two shall have identical values for program_number and shall have identical values for linked_component_tag.

If the Enhanced Signaling Descriptor is in both descriptor loops, the Enhanced Signaling Descriptor in the descriptor loop immediately following the ES_info_length field shall take precedence for the particular program element described therein.

An Enhanced Signaling Descriptor shall be present for each program element transmitted using one-half rate or one-quarter rate coding in either descriptor loop in the TS_program_map_section(), or both descriptor loops in the TS_program_map_section(), subject to the above precedence rule. The bit stream syntax for the Enhanced Signaling Descriptor is shown in Table C2.

Table C2 Bit Stream Syntax for the Enhanced Signaling Descriptor[8]

Syntax	No. of Bits	Format
enhanced_signaling_descriptor () {		
descriptor_tag	8	0xB2
descriptor_length	8	uimsbf
linkage_preference	2	uimsbf
tx_method	2	uimsbf
if (linkage_preference != '00')		
linked_component_tag	4	uimsbf
else		
reserved	4	'1111'
}		

descriptor_tag – This 8-bit unsigned integer shall have the value 0xB2, identifying this descriptor as an enhanced_signaling_descriptor.

descriptor_length – This 8-bit unsigned integer that shall indicate the length (in bytes) immediately following this field up to the end of this descriptor.

linkage_preference – This 2-bit bit field indicates whether or not the associated program element is Linked to another program element in the transmission. If so, it identifies the broadcaster's preference for the program element associated with this descriptor. The linkage_preference field shall be coded as defined in Table C3.

reserved – This 4 bit field shall contain all '1's.

---

[8] This descriptor may be extended in the future to provide additional capability by adding a byte loop after the last defined byte. Proposals to do so are expected to be considered in the future.

Table C3 Linkage Preference Values

linkage_preference	Meaning
'00'	**not_linked** – This element is not Linked to any other program element in the transmission
'01'	**linked_no_preference** – There is no preference between Linked program elements
'10'	**linked_preferred** –This program element is preferred over the Linked program element
'11'	**linked_not_preferred** – The Linked program element is preferred over this program element

tx_method – This 2-bit field shall identify the VSB transmission method used to transmit the associated program element.

Table C4 Transmission Method Values

tx_method	Meaning
'00'	**main** – This program element transmitted using main coding
'01'	**half_rate** – This program element transmitted using rate-one-half enhanced coding
'10'	**quarter_rate** – This program element transmitted using rate-one-quarter enhanced coding
'11'	Reserved for signaling future transmission methods

linked_component_tag – A 4-bit unsigned integer that links the associated program element to an equivalent, preferred, or less preferred alternative. The Linked program element is the program element labeled with the same value of linked_component_tag in a TS_program_map_section labeled with an equivalent value of program_number as the TS_program_map_section that carries this descriptor. If the program element has the linkage_preference value equal to '00', then the program element is not Linked and the linked_component_tag value shall be reserved. See Table C2.

6.9   PID Value Assignments

In order to avoid collisions with fixed PID values and ranges already established in this and other international standards, transport_packet() PID field values are restricted as follows:

- TS packets identified with PID values in the range 0x1FF0 – 0x1FFE shall only be used to transport data compliant with ATSC-recognized standards specifying fixed-value PID assignments in that range. (Informative note: One such use is A/65, which requires the use of 0x1FFB to identify packets containing certain tables defined in that Standard.)

- In order to avoid collisions with fixed PID values and ranges already established in this and other international standards, PID values used to identify Transport Stream packets carrying TS_program_map_section() or program elements shall not be set below 0x0030. (Informative note: One such use is in ETS 300 468, which requires the use of 0x0011 to identify packets containing certain tables defined in that Standard.)

6.10   Extensions to the MPEG-2 Systems Specification

This Section covers extensions to the MPEG-2 Systems specification.

6.10.1 Scrambling Control

The scrambling control field within the packet header allows all states to exist in the digital television system as defined in Table C5.

Table C5 Transport Scrambling Control Field

transport_scrambling_ control	Function
'00'	packet payload not scrambled
'01'	not scrambled, state may be used as a flag for private use defined by the service provider.
'10'	packet payload scrambled with "even" key
'11'	packet payload scrambled with "odd" key

Elementary Streams for which the transport_scrambling_control field does not exclusively have the value of '00' for the duration of the program must carry a CA_descriptor in accordance with Section 2.6.16 of ISO/IEC 13818-1 [C3].

The implementation of a digital television delivery system that employs conditional access will require the specification of additional data streams and system constraints.

7. FEATURES OF 13818-1 NOT SUPPORTED BY THIS STANDARD

The transport definition is based on the MPEG-2 Systems standard, ISO/IEC 13818-1; however, it does not implement all parts of the standard. This section describes those elements that are omitted from or constrained by this Standard.

7.1 Program Streams

This Annex does not include those portions of ISO/IEC 13818-1[C3] and Annex A of ATSC Standard A/52B that pertain exclusively to Program Stream specifications.

7.2 Still Pictures

This Annex does not include those portions of ISO/IEC 13818-1 Transport Stream specification which pertain to the Still Picture model for the main service (TS-M).

8. TRANSPORT SUBSYSTEM INTERFACES AND BIT RATES

8.1 Transport Subsystem Input Characteristics

The MPEG-2 Systems standard defines system coding at two hierarchical layers: The packetized elementary stream (PES) and the systems stream, either in Transport Stream or Program Stream format (the ATSC only uses the Transport Stream format). Under this standard and by common industry usage, private_section encapsulated data is a parallel layer to PES. Physical implementations may include the PES packetizer within a video, audio, or other data encoder; and a private_section encapsulator within a data encoder; and not as part of the transport subsystem. Therefore, the inputs to the transport subsystem may be elementary streams, PES packets, or private_section encapsulated data.

8.2 Transport Subsystem Output Characteristics

Conceptually, the output from the transport subsystem is a continuous MPEG-2 Transport Stream as defined in this Annex at a constant rate of T_r Mbps when transmitted in an 8 VSB system and $2T_r$ when transmitted in a 16 VSB system where

$$T_r = 2 \times \left(\frac{188}{208}\right)\left(\frac{312}{313}\right)\left(\frac{684}{286}\right) \times 4.5 = 19.39... \text{ Mbps}$$

The symbol rate S_r in Msymbols per second for the transmission subsystem (see Section 5 of Annex D) is

$$S_r = \left(\frac{684}{286}\right) \times 4.5 = 10.76... \text{ Msymbols per second}$$

T_r and S_r shall be locked to each other in frequency.

All Transport Streams conforming to this Standard shall conform to the ISO/IEC 13818-1 [C3] T-STD (Transport System Target Decoder) model.

9. PACKET DELIVERY TO THE E-VSB AND VSB MODULATION SYSTEM

This section describes a reference model for interfacing with the enhanced transmission system model as described in Annex D and a receiver reference model. These models are used in specifying system constraints and requirements.

9.1   Head End Reference Model

The E-VSB exciter defined in Annex D can accept three separate Input Streams (IS), one each for Transport Stream packets destined to be sent via each of the one-half rate (IS-Ea), one-quarter rate (IS-Eb) and main rate (IS-N) modes. As evinced by Annex D, only one of the three inputs has a Transport Stream packet at any one time.
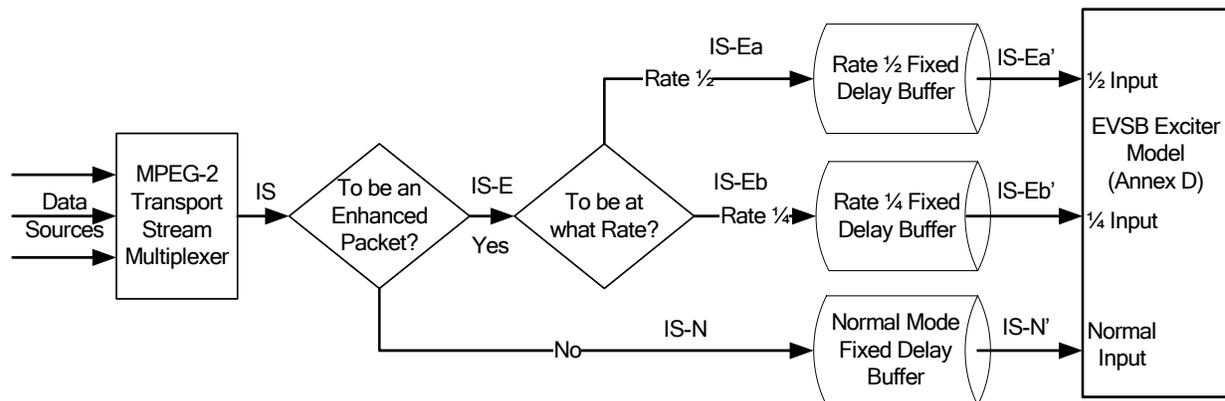
Figure C2 Head end reference model.

The above Head End Reference Model (Figure C2) models the data flow for Transport Stream packets from the final MPEG-2 multiplexer through logic and buffers and into the three exciter inputs (Annex D).

The decisions about whether a packet is enhanced or not, and, if enhanced, which of one-quarter-rate or one-half-rate is to be used (as shown by diamond boxes in Figure C2), are made

based on the PID value in each Transport Stream Packet. Special rules for certain streams (like PSIP-E data) may apply (see A/65C).

The three fixed delay buffers shown in Figure C2 are utilized to effect a delay on each stream (IS-Ea, IS-Eb, IS-N) to compensate for the fixed variance of delivery time due to the additional interleaving and error correction applied to streams in IS-E. The magnitude of the delay on each stream (IS-Ea, IS-Eb, IS-N) that must be applied to the various streams varies by which MAP is being utilized.

The stream labeled "IS" in Figure C2 may be a MPEG-2 Transport Stream, but need not be. A Transport Stream has constraints and requirements on it as to timing that need not be met at IS, but shall be met at TS-R in the Receiver Reference Model. It is, however, at least "similar" to a MPEG-2 Transport Stream.

9.2   Receiver Reference Model

The modeled Reference Receiver (as shown in Figure C3) has zero processing time, has a packet selector and has additional de-interleaving for the enhanced data.
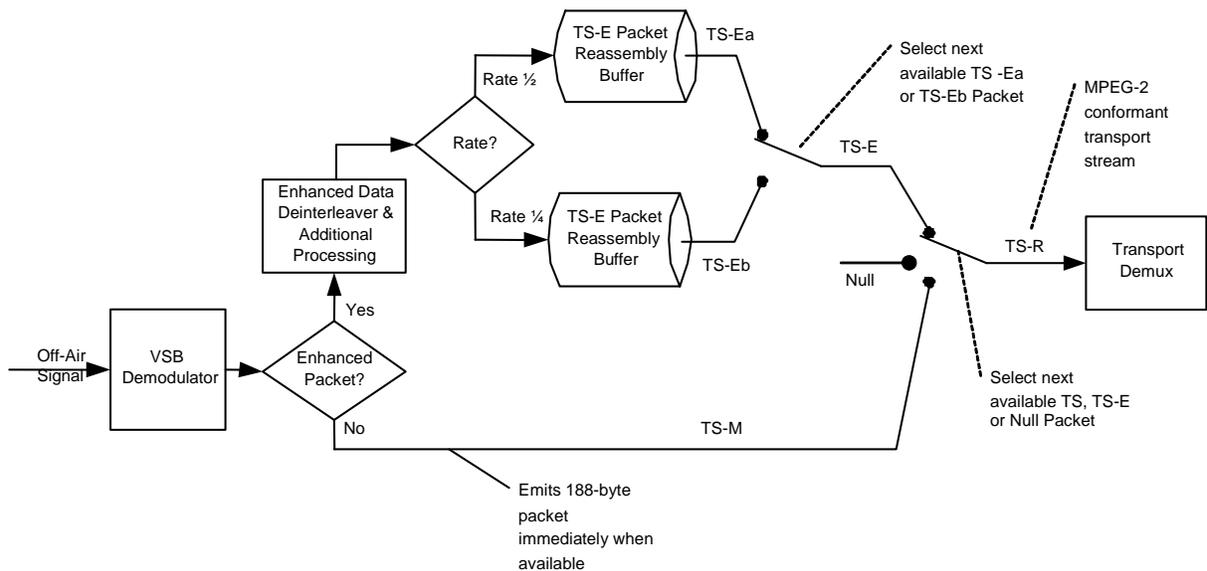
Figure C3 Reference receiver model.

The modeled Reference Receiver (as shown in Figure C3) has two 188-byte buffers that are filled by rate-one-quarter and rate-one-half data as they are recovered from the rate-one-quarter and rate-one-half E-VSB segments, respectively. The streams TS-Ea and TS-Eb contain the series of bytes from the respective TS-E reassembly buffers for each packet after it is assembled. At any time a buffer has an entire Transport Stream packet available for output (188 bytes), that buffer is selected by the switches and output immediately. The Reference Receiver does not reorder packets identified with a particular PID.

Note that due to the transmission system and the instantaneously modeled Reference Receiver behavior, it is not possible for more than one data source (TS-Ea, TS-Eb, or TS-M) to be emitting data at the same time. However, as it is possible for no data to be available at a given

packet time, the Reference Receiver selects a Null packet for insertion, so that the output is at a constant packet rate[9] to deliver 19.39… Mbps.

The emitted E-VSB stream shall be ordered and coded such that when processed by the Reference Receiver, the output from the switch (TS-R) shall be an MPEG-2 Transport Stream compliant with Sections 1 through 7 of this annex. The PCRs shall be adjusted as described in Section 9.2.

TS-E is the portion of the E-VSB transmission that results in TS-E in the Reference Receiver of Figure C3.

9.3   Stream Delays

Some streams shall be delayed relative to other streams by a number of segment times to compensate for the additional processing of enhanced-mode streams. The actual values needed are dependent on the particular mix of the enhanced and main rates defined in Annex D Section 5.6. Note that buffers of the below sizes are sufficient to accomplish this delay:

- One-half rate stream: 128 kB
- One-quarter rate stream: 128 kB
- Main stream: 3 MB

This delay compensates for the E-VSB interleaver delay and 164/188 byte packing delay.

9.4   PCR Correction

As the main, one-half rate, and one-quarter rate Transport Stream packet streams are combined for transmission, the order that the Transport Stream packets are received by the Reference Receiver from each of the Transport Stream packet streams will be different from original generation (and therefore they arrive at different times relative to one another). The time difference is dependent on the packet cadence resulting from a particular rate selection per Table D5.3 in Annex D. Any Transport Stream packets that are received by the Reference Receiver in an order different from original generation and that carry PCRs require adjustment due to this time displacement and the packet placement within each data frame.

The PCRs of all packets shall be adjusted by the emission system to compensate for the variations in the packet delivery time as compared to the time multiplexed. The adjustment is based on the timing of a virtual model stream containing the combined packets of all three streams after recovery in the receiver. (Note that when no E-VSB stream is present, PCR correction for absolute packet timing occurs prior to the exciter.) This virtual stream represents the order in which the packets are provided to a reference decoder. The virtual stream is modeled as making each one-quarter rate and/or one-half rate 188-byte TS packet (recovered from 2 or 3 enhanced packets) available with no processing delay. The Reference Receiver has instantaneous processing.

The PCRs shall be adjusted by the transmission subsystem such that when processed by the Reference Receiver, the output from the switch (TS-R) has PCRs compliant with ISO13818-1 [C3].

---

[9] This models a system that produces a constant rate packet stream output, however this may not be necessary in a real receiver.

9.5   Packet Ordering

Within each packet stream that is identified with a particular PID, the packets identified by that PID shall not be reordered by the transmission subsystem.

9.6   Main Stream Packet Jitter Handling (Informative)

The packet timing issues described in this section are solely related to the data packing structure (at the packet level).

The Transport Stream system target decoder buffer model described in ISO/IEC 13818-1 [C3] §2.4.2 has two different buffers of differing sizes and speeds (three for video). The first order buffer, TBn, is the smallest and fastest. Care should be taken during the multiplexing process to ensure meeting the requirements that TBn and T-STD not be violated for any stream at TS-R.

Adding E-VSB packets to the mix will cause each main stream Transport Stream packet to incur delay with respect to their original packet timing. The amount of delay for each main stream Transport Stream packet varies as a function of the MAP number and position in the frame (indeed, this is true for one-half rate and one-quarter rate Transport Stream packets as well). As the ratio of E-VSB to main increases, the maximum delay that may be imposed on main stream packets generally increases due to the increasing limitation on the available packet position in the VSB frame. At the physical layer, the main stream packets tend to get bunched together near the bottom of the data frame.

The E-VSB emissions system induces additional variable delay (jitter) to Transport Stream packets. This delay is knowable in real-time, with sufficient integration between the MPEG encoding/multiplexing system and the emission system.

Absent such integration, it may be possible to meet the T-STD requirement (and other timing/buffering-based requirements) by limiting the usage of T-STD buffers by the encoding system, allowing for some additional jitter.

In particular, for video packets, it is possible that the packet delay can be accommodated by increasing the average amount of data in the video T-STD buffer or other means.

However, audio is not forced to reduce its bit rate in step with the video. In fact, even when the available bit rate will not support MPEG-2 video, it can still carry full rate audio. The audio decoder buffer was not designed to tolerate a significant added packet delay. A consequence of this is that the audio buffer could underflow unless care is taken in the utilization of the audio buffer by the system.

It may be necessary to constrain the video/audio encoders and/ or the transmission system in order to meet the requirement to conform with the MPEG-2 T-STD model under various circumstances.

In particular, insuring that no overflow of TBn occurs may require particular implementation attention to the time between TS-M packets in a TS multiplex as compared to the time between TS-M packets in the transmitted frame. Under certain circumstances, the number of consecutive packets at TS-R for any particular stream of TS packets with the same PID must be limited, in order to not overflow TBn.

For example, when the chosen E-VSB mix provides only a very low rate main stream that is occupied entirely by an AC-3 audio program, it may not be possible to prevent TBn from

overflowing after insertion of the E-VSB packets. Avoidance of this overflow condition may require changes to the original allocation of packet rate to TS-M services or eschewing certain combinations in the multiplex.

-- End of document --