SMPTE STANDARD

MPEG TS Packets Carried in VANC

Table of Contents

Approved
June 29, 2011

Foreword

SMPTE (the Society of Motion Picture and Television Engineers) is an internationally-recognized standards developing organization. Headquartered and incorporated in the United States of America, SMPTE has members in over 80 countries on six continents. SMPTE's Engineering Documents, including Standards, Recommended Practices, and Engineering Guidelines, are prepared by SMPTE's Technology Committees. Participation in these Committees is open to all with a bona fide interest in their work. SMPTE cooperates closely with other standards-developing organizations, including ISO, IEC and ITU.

SMPTE Engineering Documents are drafted in accordance with the rules given in Part XIII of its Administrative Practices.

SMPTE ST 2056 was prepared by Technology Committee 24TB.

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential to the implementation of this Standard. However, attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. SMPTE shall not be held responsible for identifying any or all such patent rights.

Introduction

This section is entirely informative and does not form an integral part of this Engineering Document.

In many digital systems, subsystems which historically were physically co-located needs to be now separated by sometimes a significant distance. Many of these subsystems produced MPEG-2 Transport Stream (TS) packets, carried over ASI or IP, which were then multiplexed with other compressed video and audio signals for emission to home viewers.

Examples of these subsystems include DVB bitmapped subtitle generators and EBIF carousels. In both cases the earliest implementations assumed close proximity of the generator to the emission multiplexer. Over time it has become necessary in some cases to have the generator either remotely located or to take an original compressed signal, decode it to baseband, switch/route that signal, and re-compress it for emission. In either case, the system design will be greatly simplified by carriage of the low-bandwidth MPEG-2 TS packets in the VANC of their associated video.

1 Scope

This Standard defines a standardized wrapper for the carriage of low bit rate MPEG-2 TS packets, in the 10-bit vertical ancillary data (VANC) space of a standard definition or high definition serial digital link in accordance with SMPTE ST 291. Examples of such low bitrate services are EBIF and DVB bitmapped subtitles.

Note: Low bit rate means typically 1-20 TS packets per video frame.

2 Conformance Notation

Normative text is text that describes elements of the design that are indispensable or contains the conformance language keywords: "shall", "should", or "may". Informative text is text that is potentially helpful to the user, but not indispensable, and can be removed, changed, or added editorially without affecting interoperability. Informative text does not contain any conformance keywords.

All text in this document is, by default, normative, except: the Introduction, any section explicitly labeled as "Informative" or individual paragraphs that start with "Note:"

The keywords "shall" and "shall not" indicate requirements strictly to be followed in order to conform to the document and from which no deviation is permitted.

The keywords, "should" and "should not" indicate that, among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required; or that (in the negative form) a certain possibility or course of action is deprecated but not prohibited.

The keywords "may" and "need not" indicate courses of action permissible within the limits of the document.

The keyword “reserved” indicates a provision that is not defined at this time, shall not be used, and may be defined in the future. The keyword “forbidden” indicates “reserved” and in addition indicates that the provision will never be defined in the future.

A conformant implementation according to this document is one that includes all mandatory provisions ("shall") and, if implemented, all recommended provisions ("should") as described. A conformant implementation need not implement optional provisions ("may") and need not implement them as described.

Unless otherwise specified, the order of precedence of the types of normative information in this document shall be as follows: Normative prose shall be the authoritative definition; Tables shall be next; followed by formal languages; then figures; and then any other language forms.

3 Normative References

The following standards contain provisions which, through reference in this text, constitute provisions of this recommended practice. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this recommended practice are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

SMPTE ST 291:2010, Ancillary Data Packet and Space Formatting

ISO/IEC 13818-1:2007, Information Technology – Generic Coding of Moving Pictures and Associated Audio Information: Systems

ISO/IEC 13818-6:1998, Information Technology – Generic Coding of Moving Pictures and Associated Audio Information – Part 6: Extensions for DSM-CC

4 Definitions and Acronyms

carousel: A term referring to a technique defined by MPEG and other standards where one or more TS packets are multiplexed into a TS by cyclically repeating the contents of those packets. See Section 5.5.

DTS: An abbreviation for Decode Time Stamp. (Note: See ISO/IEC 13818-1.)

EBIF: An abbreviation for Enhanced TV Binary Interchange Format (Note: See CableLabs OC-SP-ETV-BIF1.0-I04-070921.)

IRD: An abbreviation for Integrated Receiver Decoder.

NOT: The logical function NOT is the logical inverse of the bit or bits designated.

PCR: An abbreviation for Program Clock Reference. (Note: See ISO/IEC 13818-1.)

PSI: An abbreviation for Program Specific Information tables. (Note: See ISO/IEC 13818-1.)

PTS: An abbreviation for Presentation Time Stamp. (Note: See ISO/IEC 13818-1.)

SI: An abbreviation for Service or System Information tables. (Note: See relevant ATSC, DVB, and SCTE standards.)

STB: An abbreviation for Set-Top Box.

TS: An abbreviation for (MPEG-2) Transport Stream. (Note: See ISO/IEC 13818-1.)

TSCD: An abbreviation for TS Carriage Data.

UDW: An abbreviation for “user data words.” (Note: See SMPTE ST 291.)

uimsbf: An abbreviation for Unsigned integer, most significant bit first. Note: The byte order of multi-byte uimsbf words is most significant byte first.

VANC: An abbreviation for Vertical ANCillary. As used in this document it is an acronym for vertical ancillary data space.

5 System Overview (Informative)

A number of system architectures have assumed co-location of some type of data generator (such as a bitmapped subtitle generator) with the system multiplexer. In the particular case of subtitles, this was due to the necessity of obtaining the MPEG-2 Systems’ PCR timestamp samples from the associated video bitstream, thus permitting correct values of PTS in the subtitle TS packets.

As “digital turn-around” architectures became commonplace, the need to “bridge” data services between an IRD and a downstream encoder arose. System operators desire a “one-wire” system architecture, and thus the requirement for placing the subtitle or EBIF TS packets into the VANC of their associated video frame. This standard provides that mechanism.

5.1 Use Cases for MPEG-2 TS in VANC Services (Informative)

There are three general types of services using this Standard: one is for TS packets which bear a definitive relationship to a given frame of video (which might be termed “frame synchronous”); the second is where the

relationship is to a frame in the future (which might be termed “loosely coupled” to the video), not in the present, and the third is where there is a much looser relationship, if any.

In the first case, the transmitter for this system (which, creating potential confusion, is often an MPEG decoder, also called an IRD) needs to establish that relationship using the clock sample timestamps provided by MPEG-2 Systems (PCR, PTS, DTS), decode the video and audio as expected by the user plus ensure that the associated MPEG-2 TS packet carrying the ancillary service is placed into the VANC of the designated video frame and no other. This permits the downstream receiver (which, also potentially confusing) is typically an encoder of audio/video/data baseband (which produces a MPEG-2 TS output) to ensure that the clock sample timestamps in the conveyed TS packet are modified to match those of the new TS.

In the second case (“loosely coupled”), the transmitter (IRD) needs to perform calculations using PTS values (see Section 5.2) as well as getting the TS packet into the VANC of the closest frame of video to its original placement in the arriving MPEG-2 TS. The transmitter also needs to support situations where several different arriving TS packets carry the same PTS values, and are simply passed through without buffering or re-ordering.

In the third case, the transmitter (IRD) needs not worry about the presence of PTS or the VANC of the exact video output frame into which to place the TS packet. Rather, the transmitter is concerned with whether or not the packet needs to be carouseled or will be simply placed into the ultimate output TS once. If the packet needs to be carouseled, then the transmitter needs to set the repetition rate and duration in the header of the message.

5.2 Breaking the Prior MPEG TS Clock Domain (Informative)

An example of a frame synchronous service is AMOL data (see ANSI/SCTE 127) which is tied to the video frame most closely associated with the PTS value in the MPEG-2 TS PES packet header. An example of a loosely coupled service is DVB subtitles (see ETSI 300 743), where there may be several TS packets carrying subtitle data with the same PTS value. Each of these services uses the MPEG-2 PES packet structure to carry a PTS value. The values are handled differently for each case.

For frame synchronous services, the transmitting IRD needs to not only reconstruct the encoded video and audio, but also ensure that each incoming frame synchronous TS packet is placed into the correct output video frame’s VANC space. The transmitting IRD does no adjustment to the original PTS value. This provides the receiving encoder with the precise input TS packet to video frame relationship to ensure that the TS packet can be extracted from the VANC and placed into the outgoing TS simply by overwriting the PTS value in the extracted TS packet with the PTS value matching the compressed video frame to which it is associated.

For loosely coupled services, the transmitting IRD needs to adjust the PTS value by subtracting the current value of the program’s PCR as maintained by the IRD from the PTS value in the arriving TS packet. The IRD then places that relative value into the PES packet header, replacing the original value. The receiving encoder needs to then further adjust the PTS value by adding that relative value to its current PCR clock sample to become the final PTS value, as it places the TS packet in the outgoing TS. The IRD needs to place each arriving TS packet for the service into the VANC of a corresponding output video frame using a constant processing delay.

This process (as done by both devices) results in the following equation:

newPTS = oldPTS - inputPCR + outputPCR

See the semantics for **PTS_processing_flag** in Section 6.2 for the specific details.

For the third case outlined above, there is no requirement to perform any adjustment of timestamps or careful alignment with output video.

5.3 Additional Considerations for DVB Subtitles (Informative)

The bit rate allocated to DVB subtitles within a TS is typically small and can be as low as 64-128 kbps per language. Subtitle data is typically “pre-loaded” to reduce the bitrate per frame. There is no upper range, but for example, 8 languages might run as high as 1.6 Mbps total for HD. This leads the origination equipment to pre-load subtitles up to 3 seconds ahead of the display time to give the STB time to decode the subtitle. The space available within the HD-SDI is relatively large and is able to handle 32 lines of SMPTE ST 2031 data within a single vertical interval line for example. Video encoding delays allows the subtitle data to be sent early if required.

The DVB subtitle specification (see ETSI EN 300 743) details a “buffer model” for the decoder plus a “decoder rendering bandwidth” model. Adherence to these models can be broken if subtitle data TS packets are not delivered at roughly the same data rate as that originally intended. The relative temporal spacing between the PES packets being delivered in VANC and placed into the receiving encoder’s outgoing TS needs to be maintained.

Various geographical regions have deployed STBs which are sensitive to the above features. Implementation experience has led some implementers to build local ‘profiles’ into their equipment to cope with known limitations with the STB population.

5.4 Table Handling (Informative)

This Standard also provides a mechanism to pass MPEG/ATSC/DVB/SCTE table structures (“PSI/SI”), such as a TS packet carrying a PMT section. One use is to permit changes in PID values without a human needing to intervene. Other uses relate to the different types of services hosted on this link but are beyond the scope of this document.

5.5 Data Carousels (Informative)

As noted earlier, this Standard provides mechanisms to pass data carousels, which are well documented by other Standards. As the name suggests, the data is sent repeatedly through the transport system. Since the data is repeated, the IRD (knowing that the data is part of a carousel) can pass it through the IRD-to-inserter (encoder) “bridge” without the need to send each repetition seen in the incoming TS. See Section 9 for specifics and the helpful tutorials in ETSI TR 101 202 Section 4 and ATSC A/90 Section 5 for further background.

6 Format of TS Carriage Data Packets

Each TS Carriage Data (TSCD) packet shall comply with the format defined in SMPTE ST 291 for a 10-bit type 2 ancillary data packet. It consists of the ancillary data flag (ADF), the data ID (DID), the secondary data ID (SDID), the data count (DC), the user data words (UDW), and the checksum (CS).

The UDW shall be formatted as shown in Table 1, consisting of a header structure followed by the **transport_packet( )** as defined in Table 2-2 of ISO/IEC 13818-1. The format of the TS data contained within each **transport_packet( )** structure is arbitrary.

The DID word in the TSCD packet shall be set to the value 41h. The SDID word in the TSCD packet shall be set to the value of 09h. DC shall be the size of the structure defined in Table 1.

Note: 188 is the size of the **transport_packet( )** structure defined in Table 2-2 of ISO/IEC 13818-1. The present document specifies two sizes for the entire structure in Table 1: 190 bytes (TS_placement_flag != 2) and 196 bytes (TS_placement_flag == 2).

Each **transport_packet( )** shall begin in a new TSCD packet. If there are several **transport_packet( )** structures to be carried with a given frame of video, and another TSCD packet will fit in the remaining space

in a given line, then more than one TSCD packet may be placed in that line. The TSCD packets shall be placed into the VANC space conforming to the rules defined in SMPTE ST 291.

6.1 TSCD Structure Syntax

Table 1 – UDW format – TSCD Structure

Syntax	No. of bits	Format
zero_bits	4	'0000'
sequence_number	4	uimsbf
TS_placement_flag	4	uimsbf
PTS_processing_flag	4	uimsbf
If (TS_placement_flag == 2) {		
stream_target_bitrate	8	uimsbf
zero_bit	1	'0'
num_ts_packets	15	uimsbf
zero_bit	1	'0'
ts_packet_index	15	uimsbf
version	4	uimsbf
zero_bits	4	'0000'
}		
else {		
}		
transport_packet( )	188*8	

Each field in Table 1 shall be placed into bits 0 to 7, packing against each neighbor field to fill complete bytes, with the low order bit going into bit 0 of the 10-bit word. Fields which are larger than 8 bits are divided between two or more bytes. Each byte of the **transport_packet( )** data shall be placed into bits 0 to 7 of the 10-bit UDW, with the low order bit going into bit 0 of the 10-bit word. See Figure 1 for additional information.

Bit b8 is the even parity for bits b7 through b0 and bit b9 is NOT b8.

Note: The bytes which proceed **transport_packet( )** in Table 1 are considered the "header" structure referred to above.

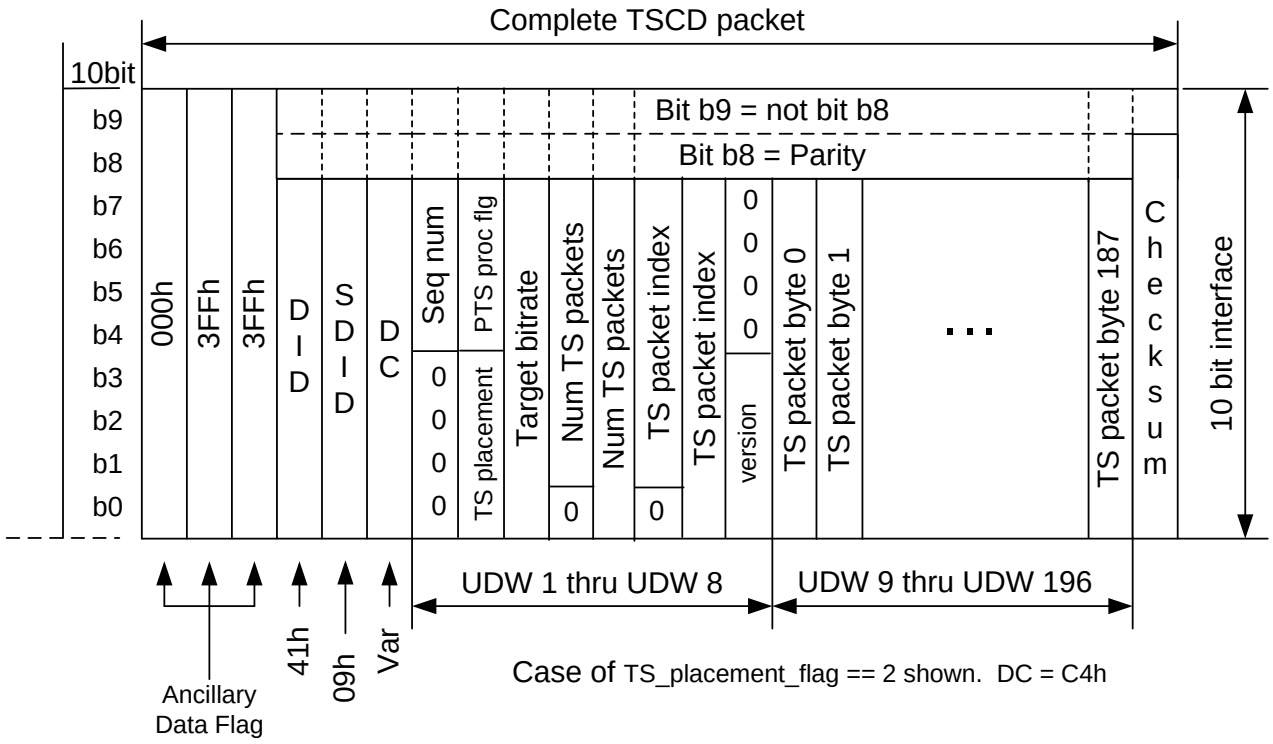

Figure 1 – TS Carriage Data Structure illustrated

6.2 TS Carriage Data Structure Semantics

A single transport_packet( ) Structure as shown in Table 1 shall only contain a single PTS value.

The **sequence_number** is provided to permit unambiguous placement of TS packets within a single frame of video where such is important. Otherwise this field should be set to 0.

The values for **stream_target_bitrate**, **num_ts_packets**, **ts_packet_index**, and **version** are specified in Section 9.

The values of **TS_placement_flag** shall be one of those given in Table 2:

Table 2 – TS_placement_flag values

Value	Meaning
0	Immediate Placement
1	Frame Aligned Placement
2	Cyclic Placement
3	PSI/SI Section
4 – 15	Reserved

A value of 0 in the **TS_placement_flag** indicates that the TS packet should be output within the outgoing TS as soon as possible making best use of any transcoding delays which may delay the video and audio path. See Section 7 for additional information.

A value of 1 in the **TS_placement_flag** indicates that the TS packet should be output synchronized as close as possible to the current frame. See Section 8 for additional information.

A value of 2 in the **TS_placement_flag** indicates that the TS packet should be output per the specifications in Section 9 of this document.

A value of 3 in the **TS_placement_flag** indicates that the TS packet should be processed per the specifications in Section 10 of this document.

Table 3 – PTS_processing_flag values

Value	Meaning
0	No PTS adjustment
1	Match video frame
2	Relative to video frame
3	Reserved

6.2.1 PTS_processing_flag: Cases 0 and 1

A value of 0 for the **PTS_processing_flag** indicates that the TS packet shall be placed in the outgoing TS without any adjustment to the value of PTS, if any exist, within the packet.

A value of 1 for **PTS_processing_flag** indicates that the TS packet shall be placed in the outgoing TS, with the inserting device (encoder) replacing the value of PTS within the packet with that of the associated frame of video.

Note: A bit-for-bit match of the two values is not expected, but matching values should be within 1/2 of a video frame-time of each other.

6.2.2 PTS_processing_flag: Case 2

A value of 2 for **PTS_processing_flag** indicates that the TS packet contains a PTS field and that field contains a relative value. The IRD shall replace the original value of the PTS within the TS packet using the following calculation:

UpdatedPacketPTS = *OrginalPacketPTS* – *incomingStreamPCR*.

The encoder shall calculate the resultant value of the output PTS timestamp using the following calculation:

ResultantPTS = *outgoingStreamPCR* + *UpdatedPacketPTS.*

In the above calculations, the value *OrginalPacketPTS* shall be the value of the PTS in the incoming TS packet, the value *incomingStreamPCR* shall be the value of the IRD's local instantiation of the PCR-locked System Time Clock, the value of *outgoingStreamPCR* shall be the value of the encoder's local System Time Clock (from which it generates PCR timestamps in the output MPEG-2 TS).

If the value of *OrginalPacketPTS* is greater than *incomingStreamPCR* or more than 7 seconds in advance of the value of *incomingStreamPCR,* then the IRD shall place a value of 0 into **PTS_processing_flag** for that TS packet.

Note: CEA-CEB-20 has a useful discussion of decoder clock instantiation and maintenance of PCR lock.

7 Immediate Placement Messages

UDW containing **TS_placement_flag** of value 0 shall indicate an Immediate Placement Message. The contents of TS packets shall pass through the IRD-to-inserter (encoder) “bridge” unchanged.

The placement within the stream may be variable, but order shall be maintained.

Note: The most common use of Immediate Placement Messages are for non-carouseled EBIF data.

8 Frame Aligned Placement Messages

UDW containing **TS_placement_flag** of value 1 shall indicate an Frame Aligned Message. The PTS value, if present, may be changed as indicated by the value of **PTS_processing_flag** as it passes through the IRD-to-inserter (encoder) “bridge.”

The most common envisaged use of Frame Aligned Placement messages are the carriage of DVB Subtitle Streams (Note: See ETSI EN 300 743.)

Where TS packets are decoded with a mixture of Immediate Placement and Frame Aligned Placement to ordering of the TS packets with the output transport stream is implementer defined.

8.1 DVB Subtitle Streams

For DVB subtitle packets, the IRD will decode the PES packet header of the subtitle messages and where a PTS is found, the IRD shall set the **PTS_processing_flag** to the value of 2 indicating that the value of the PTS is relative to the PTS of the current video frame. The PTS value within the resulting TS packet shall be changed to include a relative value from the current PCR as specified in Section 6.2.

The UDW contained within a single video line in VANC can contain data for multiple streams identified by different PIDs. If the streams contain multiple subtitle streams, for example, where multiple languages are carried, a single video line in VANC may contain UDW which might require different PTS adjustments.

Where the transport stream packets do not contain PTS values the **PTS_processing_flag** shall be set to the value of 0, indicating no PTS adjustment.

9 Cyclic Placement Messages

UDW containing a **TS_placement_flag** of value 2 shall indicate that the TS packet is part of a carousel which will be only cycled periodically on the IRD–to-inserter (encoder) “bridge,” but is expected to be repeated by other downstream equipment.

Note: The manner in which the TS packets to be carouseled are delivered to the downstream equipment is out of scope of this standard.

The **stream_target_bitrate** shall specify the output bitrate which shall be applied to this carousel, in 5 kbps units.

The **num_ts_packets** field shall be the total number of TS packets which are in this carousel. A value of 0 indicates that the carousel is terminated.

The **ts_packet_index** shall be the packet index of this specific packet with the carousel, taking values from 0 to **num_ts_packets**-1.

version shall be a 4 bit value specifying the version number, which shall range from 1 to 15, with 0 indicating Carousel Removal state (see Section 9.1). A change of version number indicates that the entire carousel has changed version.

9.1 Carousel Removal State

To request that downstream equipment to end the transmission of the current carousel, the value of **version** shall be set to 0 and the value of **num_ts_packets** shall be set to 0. The TS packet should be sent several times to reduce chance of missed reception.

10 PSI/SI Tables

UDW containing **TS_placement_flag** of value of 3 shall indicate that the TS packet contains PSI/SI section data, as defined by an appropriate standard. This data may be placed in the output TS if appropriate to the application.

Note: PSI/SI data is primarily intended for use by the receiving encoder to permit unattended provisioning (such as detecting PID changes in the PMT).

Annex A Bibliography (Informative)

Note: All references in this document to other SMPTE documents use the current numbering style (e.g. SMPTE ST 2031:2007) although, during a transitional phase, the document as published (printed or PDF) may bear an older designation (such as SMPTE 2031-2007). Documents with the same root number (e.g. 2031) and publication year (e.g. 2007) are functionally identical.

SMPTE ST 2031:2007, Carriage of DVB/SCTE VBI Data in VANC

ANSI/SCTE 127-2007, Carriage of Vertical Blanking Interval (VBI) Data in North American Digital Television Bitstreams

ATSC A/90, ATSC Data Broadcast Standard (including Amendment 1 and Corrigendum 1 and Corrigendum 2), 2000

CableLabs OC-SP-ETV-BIF1.0-I04-070921, Enhanced TV Binary Interchange Format 1.0 (EBIF), 2007

CEA-CEB-20, A/V Synchronization Processing Recommended Practice, 2009

ETSI TR 101 202, Digital Video Broadcasting (DVB); Implementation Guidelines for Data Broadcasting, 2003

ETSI EN 300 743, Digital Video Broadcasting (DVB); Subtitling Systems, 2006

ETSI EN 301 192, Digital Video Broadcasting (DVB); DVB Specification for Data Broadcasting, 2008