SMPTE STANDARD

Unidirectional Transport of Variable Bit Rate MPEG-2 Transport Streams on IP Networks

Table of Contents

Approved
April 30, 2010

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 2022-3 was prepared by Technology Committee 32NF.

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

IP-based networks have become increasingly important for delivery of compressed content which is contained within MPEG-2 Transport Streams. However, existing transport protocols do not fully meet the user requirements. This standard describes modifications to existing transport protocols which can be used for the unidirectional carriage of piecewise constant variable bit rate MPEG-2 Transport Streams over IP networks.

This standard is intended for real-time audio/video applications for contribution services over an IP network to professional broadcast equipment. The applications addressed by this standard may employ any compression scheme that is supported by the MPEG-2 Transport Stream.

This standard also provides for signal recovery from limited network errors through a forward error correction scheme.

1 Scope

This standard defines a transport protocol for the carriage of real-time piecewise constant variable bit rate (VBR) MPEG-2 Transport Streams over IP networks, either with or without Forward Error Correction for recovery from network transmission errors. A piecewise constant VBR transport stream can only vary at the PCRs of the program under consideration per ISO/IEC 13818-1:2007 § 2.4.2.2.

This standard covers the encapsulation and transmission of MPEG-2 transport streams but does not cover other processes such as MPEG-2 encoding or multiplexing. The scope of this standard is shown in Figure 1. below.

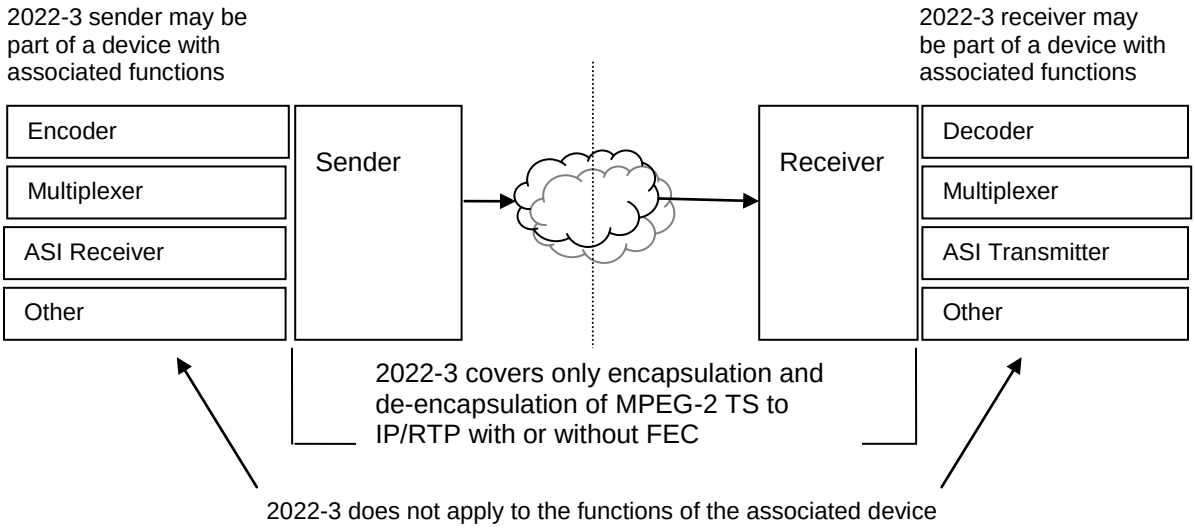

Figure 1 – SMPTE 2022-3 Scope Boundaries

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 standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this standard are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

SMPTE 2022-1-2007, Forward Error Correction for Real-Time Video/Audio Transport Over IP Networks

SMPTE 2022-2-2007, Unidirectional Transport of Constant Bit Rate MPEG-2 Transport Streams on IP Networks

ISO/IEC 13818-1:2007, ITU-TH.222.0-2006, Generic Coding of Moving Pictures and Associated Audio Information Systems

4 Acronyms (Informative)

CBR: Constant Bit Rate

D: Number of Rows in an FEC matrix

FEC: Forward Error Correction

IP: Internet Protocol (see IETF STD 05)

L: Number of Columns in an FEC matrix

PCR: Program Clock Reference

RTP: Real Time Protocol (see IETF RFC 3550)

TS: Transport Stream (see ISO/IEC 13818-1)

UDP: User Datagram Protocol (see IETF STD 06)

VBR: Variable Bit Rate

XOR: Exclusive OR

5 Definitions (Normative)

Block Aligned FEC Operation: Block Aligned FEC operation is a correction scheme using a two dimensional matrix where the media datagrams are a contiguous group of L x D datagrams. The media datagrams are protected as follows: Level A protection is provided by L FEC datagrams derived from each column for the FEC matrix. Optional level B protection is achieved by D FEC datagrams derived from each row for the FEC matrix. The level A FEC stream shall protect all media packets exactly once. The optional level B FEC stream shall protect all media packets exactly once.

CBR Transport Stream: A MPEG-2 compliant Transport Stream constructed such that the rate of departure of packets from a sender is constant over time.

Contribution Services: Unidirectional transmission of high quality media content to a media processing facility. These services require high quality transmission such that the signal maintains sufficient quality to support further processing prior to final distribution.

Fill Datagram: An RTP Datagram consisting of a header and no payload.

FEC Matrix: A set of Media datagrams ordered in a matrix with L columns and D rows. The datagrams are entered into the matrix to fill each row sequentially with incremented RTP sequence numbers.

Media Datagram: An RTP Datagram consisting of a header and data payload composed of an integer number of MPEG-2 TS packets

Null Packet: An MPEG-2 Transport Stream packet consisting of a PID value of 0x1FFF and an undefined payload.

Piecewise Constant: A VBR transport stream that can change rate at packets containing PCRs such that the rate of departure of packets from a sender is constant between successive PCR packets, according to ISO/IEC 13818-1 §2.4.2.2.

RTP Datagram: An RTP Packet as defined in RFC3550. A self-contained, independent entity of data carrying sufficient information to be routed from the source to the destination computer without reliance on earlier exchanges between this source and destination computer and the transporting network.

TS Packet: A MPEG-2 Transport Stream packet of 188 or 204 bytes in length. (The use of 188 and 204 byte packets are defined as class of service in SMPTE 2022-2.)

VBR Transport Stream: A MPEG-2 compliant Transport Stream such that the rate of departure of packets from a sender is not constant.

6 Transmission Protocols (Normative)

6.1 Relationship with other SMPTE 2022 Standards

This standard requires compliance with SMPTE 2022-1 and SMPTE 2022-2, with the exceptions that are defined in this standard. If a conflict occurs between this standard and either SMPTE 2022-1 or SMPTE 2022-2, this standard shall prevail.

The Payload Type field of the RTP header shall be set to 33 indicating MPEG-2 TS.

6.2 TS Packets per Media Datagram

Senders and receivers shall use a maximum of 1, 4 or 7 Transport Stream packets per Media datagram for the duration of a session as defined in Section 6.6 below. This maximum number (1, 4 or 7) of Transport Stream packets per Media datagram is the Packet_per_Datagram_max. The sender shall include between 0 and Packet_per_Datagram_max MPEG-2 Transport Stream packets per Media datagram.

Note: Long-length Media datagrams can cause longer latency in building FEC Matrices. A larger number of short Media datagrams results in more network encapsulation overhead, and a subsequent higher bit rate. So, the value chosen for Packet_per_Datagram_max will be a compromise between these factors.

This standard allows for two modes of operation:

Mode 1 – In Mode 1 Media Datagrams shall always be sent with a size of Packet_per_Datagram_max. The rate at which the Media Datagrams are sent varies to accommodate the variation in Bit Rate.

Mode 2 – In Mode 2 Media Datagrams shall always be sent at a constant rate. The number of TS packets per Media Datagram varies from 0 to Packet_per_Datagram_max to accommodate the variation in Bit Rate.

FEC datagrams shall always contain Packet_per_Datagram_max transport packets per RTP datagram for a session. The size of the FEC RTP datagram can be used by the receiver to determine Packet_per_Datagram_max for this session.

Note: See Annex B for further discussion of these two modes.

6.3 Timing Recovery

RTP header Timestamps do not need to be locked to any PCR in the VBR or CBR MPEG-2 transport stream being carried.

The PCR shall not be modified (same value and same position) by the sending or receiving device.

There are a number of ways to achieve clock recovery in an MPEG-2 transport stream. A standardized method of clock recovery is not specified in this document. However, Annex C. describes one clock recovery method based upon the fact that the bit rate of a piecewise constant single program MPEG-2 TS signal is constant between two PCRs.

6.4 FEC Operation

FEC operation shall be as defined in SMPTE 2022-1, with the following clarifications and exceptions.

For sender and receiver FEC calculations, all datagrams are filled to Packet_per_Datagram_max size with binary zeros before FEC calculations are performed, and all sender and receiver FEC calculations are performed using Packet_per_Datagram_max size media datagrams and Packet_per_Datagram_max FEC datagrams.

Note: The timing recovery and subsequent jitter buffer assignment does not define the recovery mechanism for FEC buffer sizing. When FEC is used, the receiver will normally wait until an FEC datagram is received to define the maximum_latency and maximum_bit_rate of the signal and will then define the FEC buffer size. However, if FEC is not used, the receiver must determine an acceptable interval to wait until it recognizes the lack of need for a FEC buffer. Since this is a receiver implementation choice and has no impact on compatibility, it is left to the receiver manufacturer to define the best practice for the intended use.

6.4.1 Mode 1 operation

Only Block Aligned FEC Operation shall be supported.

A sending device shall support a maximum_latency time. This maximum_latency time is used to limit the latency in building the FEC matrix in both the sending and receiving devices. Maximum_latency time shall be configurable only in the sending device. The receiver shall not require configuration of maximum_latency time for a session. While a longer latency is not prohibited, a SMPTE 2022-3 compliant sender and receiver shall support a minimum FEC buffer size of 1MB. A 1MB buffer allows a maximum_latency of 500 ms at a Maximum_bit_rate of 16 Mbs.

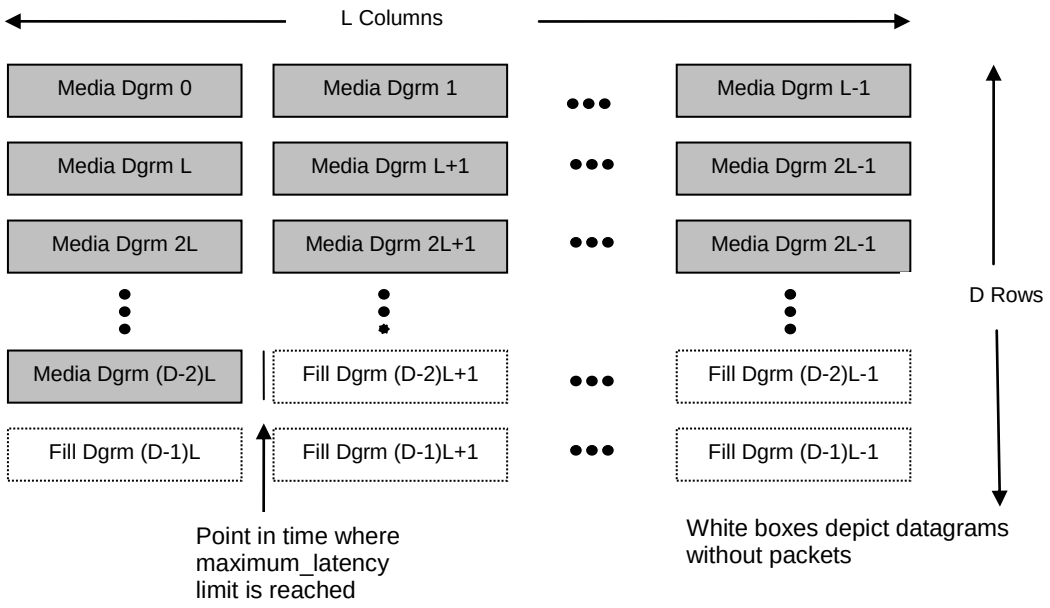

Figure 2 – Filling a FEC matrix after maximum_latency for Mode 1 Operation

Using the example in Figure 2 above, looking at only Row (D-2), when the time limit occurs before the current Media Datagram has reached the Packet_per_Datagram_max the Media Datagram is removed from the matrix, replaced with a Fill Datagram, and the partially filled Media Datagram is placed as the first datagram in the next matrix

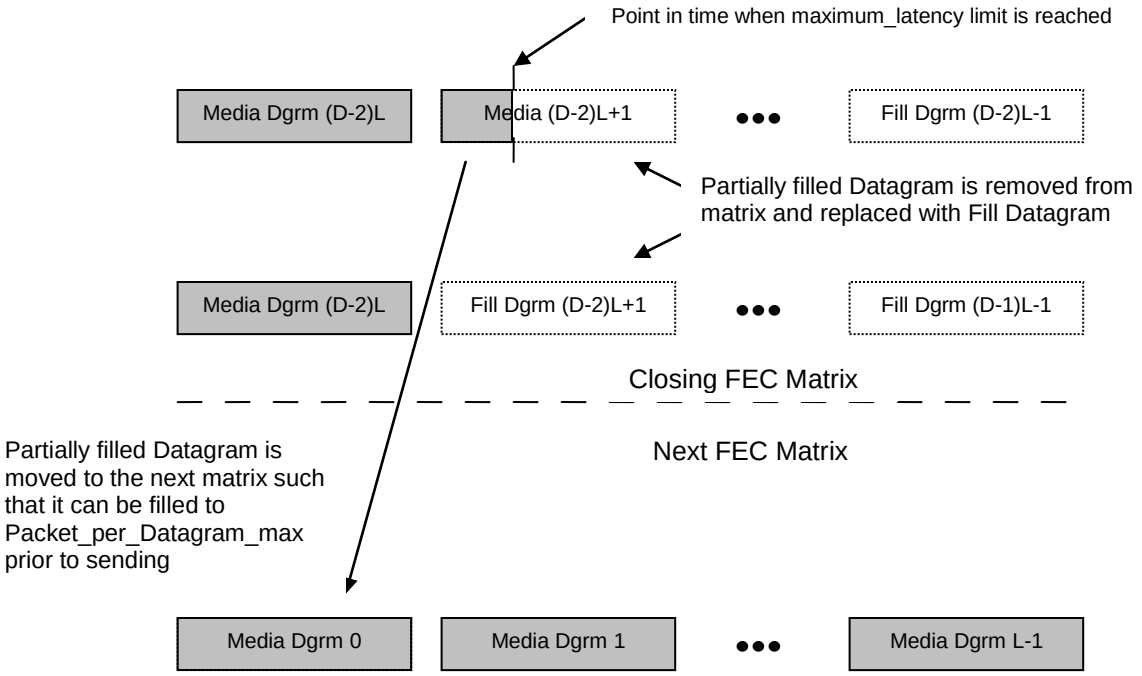

Figure 3 – Accommodating Datagrams not filled prior to maximum_latency limit

The latency timer is initialized at the time the last media datagram of an FEC block is sent. It is possible for the FEC matrix to be partially filled after a configured maximum_latency time. When this occurs, the sender shall immediately complete the FEC matrix. The sender shall fill the remainder of the matrix with Media datagrams containing zero MPEG-2 transport packets (Fill Datagrams). Figure 3 above shows an example of a partially filled FEC Matrix when the maximum_latency time expires. In the figure grey blocks show Media Datagrams that are filled with MPEG-2 TS packets. Only Media Datagrams filled to the Packet_per_datagram Max size are included in the matrix and sent to the network. If there is a partially filled Media Datagram, this Media Datagram remains at the sender and is neither sent nor included in an FEC matrix until, at a later time, it is filled to Packet_per_Datagram Max size. Once this Media Datagram is filled it becomes the first Media Datagram in the next FEC matrix. Fill datagrams containing zero MPEG-2 transport stream packets are used to complete the FEC matrix.

6.4.2 Mode 2 Operation

There are no exceptions to SMPTE 2022-1 operation

6.5 Media Datagram RTP Numbering

RTP sequence numbers of all Media datagrams shall be numbered sequentially incrementing by “1”. Where Fill Datagrams are required, the RTP numbering will sequentially step from the last Media datagram to the first Fill datagram without any gap, such that the first Media datagram in any FEC Matrix has the sequence

number equal to the sequence number of the first Media datagram in the previous FEC matrix + the value of L * D as defined in SMPTE 2022-1.

6.6 Session

A session is defined as a continuing data stream while the parameters defined in Section 7 remain constant. A decoder may produce an interruption in its MPEG-2 transport stream output at the start or end of a session.

7 Compliance

A SMPTE 2022-3 sender and receiver shall output from the receiver an exact bit copy of the MPEG-2 transport stream when compared to the MPEG-2 transport stream input to the sender, while operating within the FEC correction capability of the system. The sender and receiver shall not drop or add MPEG-2 null packets. The sender and receiver shall not adjust PCR time stamps.

A SMPTE 2022-3 sender and receiver shall support values of L and D according to SMPTE 2022-1, with a limitation that:

$$L*D \leq 256$$
$$1 \leq L \leq 50$$
$$4 \leq D \leq 50$$

A SMPTE 2022-3 receiver shall be able to receive and process FEC SMPTE 2022-1 FEC datagram streams and SMPTE 2022-3 datagram streams without a priori knowledge of the stream type or configuration.

A SMPTE 2022-3 receiver shall support reordering RTP datagrams that are up to 10 datagrams out of order.

For Mode 1 Operation, a SMPTE 2022-3 sender shall supply two transport parameters that shall remain constant during a session.

1. maximum_latency
2. maximum_bit_rate

The FEC header described in the SMPTE 2022-1 is originally 16 bytes. To allow for an extension which includes the values maximum_latency and maximum_bit_rate, the FEC header shall be defined as detailed in the text below and as shown in Figure 4.

Maximum_latency (Mode 1 only): This field shall contain a 10-bit number defining the multiplier of 10 ms equal to the maximum FEC latency up to 10.24 seconds

Maximum_bit_rate (Mode 1 only): This field shall contain a 10-bit number where the first seven bits contain the mantissa value and the last 3 bits contain an exponent value defining the multiplier of 10 Kbs up to 1,280.0 Gbs.

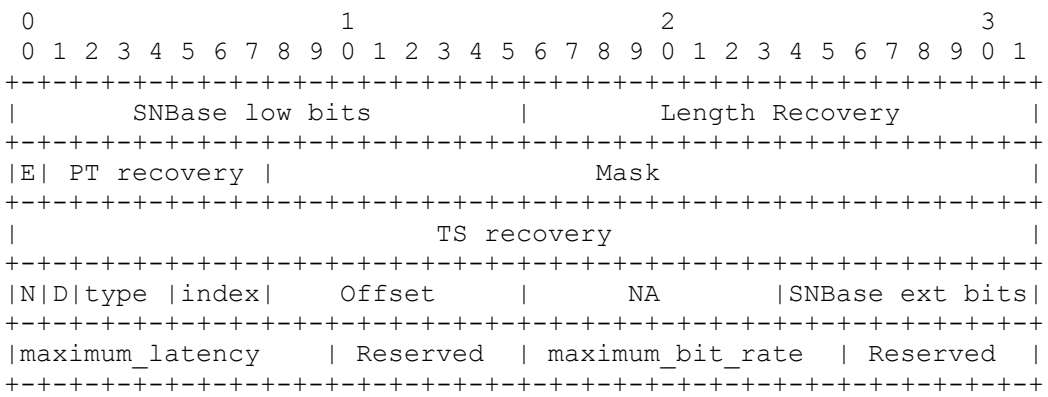

Figure 4 – Definition of the FEC header

The following fields shall be used as defined in RFC2733 and SMPTE 2022-1:

- SNBase low bits: minimum sequence number of the packets associated to the FEC datagrams, and where 16 bit sequence numbers are sufficient, this parameter shall contain the entire sequence number. For transport protocols with longer sequence numbers this field shall contain the least significant 16 bits of the sequence number.
- Length Recovery: this field should be used to determine the length of any media datagrams associated with the FEC datagrams.
- PT recovery: this field should be used to determine the Payload Type of any media datagrams associated with the FEC datagrams.
- TS Recovery: this field should be used to recover the timestamp of any media datagrams associated with the FEC datagrams.

The additional fields below shall be modified from RFC2733, or are new. The definition of these is:

- E: This shall be set to ‘1’ to indicate that the RFC2733 header is extended.
- Mask: The mask field shall be set to zero for implementations supporting this Standard.
- N: This shall be set to ‘1’ to indicate that the SMPTE 2022-1 header is extended for Mode 1 operation. For Mode 2 operation N shall be set to ‘0’ to indicate that the header is not extended.
- D: This bit shall be set to 0 for FEC datagrams from the first FEC stream (FEC over Columns), and set to 1 for FEC datagrams from the second FEC stream (FEC over Rows).
- Note: This bit is provided as an additional means of determining which FEC datagrams are associated with which FEC stream.
- Type: This field indicates which error-correcting code is chosen and shall be set to zero. Receivers shall ignore packets with an unrecognized type value.
- Index: This field is used for more complex error protection codes and shall be set to zero.

- Note: For the XOR method, only one FEC datagram protects each group of media datagrams and hence the index field will always contain 0.
- Offset: This 1-byte field is the period used to select the media datagrams associated with this FEC datagram, and shall be the L parameter as defined in SMPTE 2022-1 § 7.1 for FEC datagrams computed over columns (the first FEC stream). For FEC datagrams computed over rows (the second FEC stream) this parameter shall always be one. This field shall be kept constant by the sender during a transmission for each FEC stream.
- NA: This 1-byte field indicates the number of media datagrams associated with this FEC datagram, and shall be the D parameter as defined in SMPTE 2022-1 § 7.1 for media datagrams belonging to the first FEC stream, and shall correspond to the L parameter as defined in SMPTE 2022-1 § 7.1 for media datagrams belonging to the second FEC stream. This field shall be kept constant by the sender during a transmission for each FEC stream.
- SNBase ext bits: This field is for use with protocols which require extended sequence numbers longer than 16 bits. Where 16 bit sequence numbers are sufficient, this parameter shall be set to zero. For protocols with longer sequence numbers this field shall contain the next eight most significant bits of the sequence number beyond those contained in the SNBase.

The media datagrams protected by any given FEC datagrams shall be defined as those with sequence numbers given by the formula:

$$SNBase + j \times Offset$$

$$0 \leq j < NA$$

Annex A Bibliography (Informative)

IETF Standard 5, Internet Protocol

IETF Standard 6 (RFC 768), UDP: User Datagram Protocol

IETF RFC 3550, RTP: A Transport Protocol for Real Time Applications

Annex B Sending Device Options (Informative)

This annex is intended to provide guidance in building and configuring a sending device that is compliant to this standard.

A sending device has two options in how it builds its output Media Datagram stream to adjust to the varying transport stream rate of a VBR MPEG-2 transport stream. The sending device can adjust its output Media Datagram stream according to the VBR MPEG-2 transport stream rate by continually changing the number of MPEG-2 transport stream packets carried in a media datagram. Alternatively, a sending device can adjust its output Media Datagram stream according to the VBR MPEG-2 transport stream rate by changing the period between sent Media Datagrams. There is a trade off between complexity and FEC overhead bit rate in the method used to adjust to the VBR MPEG-2 transport stream rate. Both methods are supported by this standard.

Figure B.1 shows a datagram stream from a sending device that adjusts the sending rate of media datagrams. This VBR implementation of SMPTE ST 2022-3 fills every media datagram with maximum_packets_per_datagram. This method yields a minimum FEC bit rate overhead in the system and is defined as Mode 1 in Section 6.2.

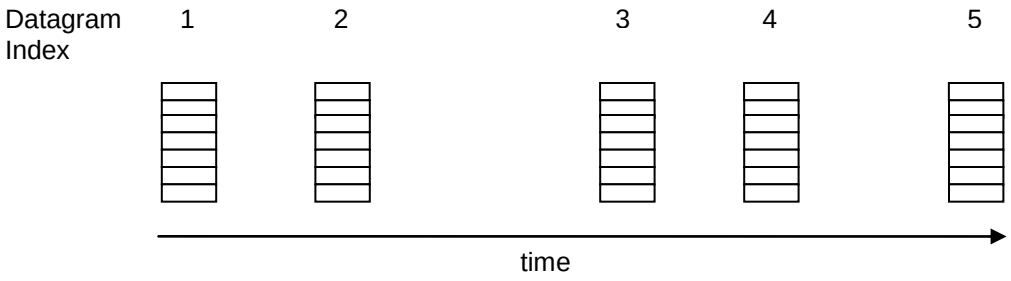

Figure B.1 – Example of Mode 1 datagram stream

Figure B.2 shows a datagram stream from a sending device that adjusts the number of MPEG-2 transport stream packets per Media Datagram. This VBR implementation of SMPTE ST 2022-3 is very similar to a CBR implementation of SMPTE 2022-1, in that Media Datagrams are sent at a constant rate, and FEC matrices are built and sent at a constant rate. This method yields a minimum FEC delay in the system and is defined as Mode 2 in Section 6.2.

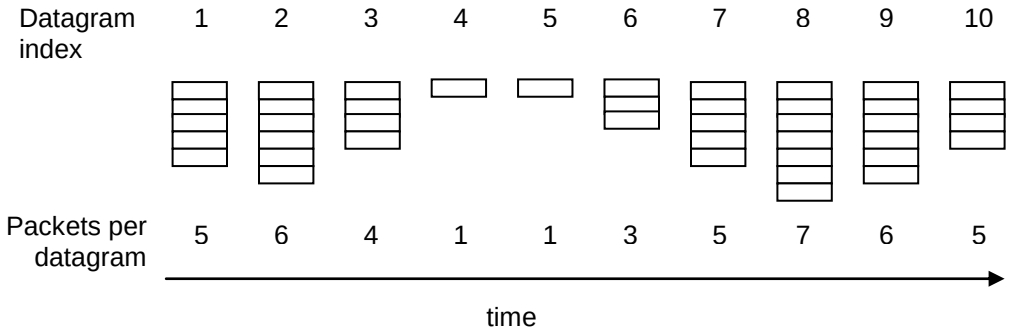

Figure B.2 – Example of Mode 2 datagram stream

Annex C Clock Recovery Method (Informative)

This annex describes one method of receiver clock recovery for piecewise constant MPEG-2 Transport Streams.

- Media Datagrams are sent as soon as they are assembled. As they are sent, Media Datagrams are also written into the FEC array of the sending device to be used in FEC datagrams calculation
- Due to the VBR nature of the transport stream, a varying transmission delay is realized for each PCR packet from when the PCR packet is received by the sender, until the sender outputs the Media Datagram. This delay can be calculated by the receiver by calculating the bit rate between PCRs and knowing the number of TS packets that the sender had to wait for before sending the Media Datagram containing the PCR packet. See step 3 below.
- The receiver must regenerate a copy of the PCR clock, which it uses to clock out transport stream packets at its output. This receiver PCR clock must be phase locked to the arrival times of PCRs into the receiver. These following steps are used for the receiver to PLL its PCR clock. Refer to Figures 1 and 2 for these steps.

    1. The receiver timestamps the time it receives a datagram = DG_t
    2. Receiver calculates the TS bitrate between PCRs = BPS_i
    3. $DG_t - PCR_t = 1/ (BPS_i * (NumPkts_t) * 188 * 8)$ where $NumPkts_t$ are the number of TS packets following a PCR packet in a Datagram
    4. The receiver determines PCR_t and uses this as an input to PLL for the receivers local PCR clock. This PLL must implement an extremely narrow bandwidth filter (time constant of minutes) to filter out IP jitter.

To successfully recover the clock under this method it is imperative that two PCRs be present in the receiver. If the implementation requires two PCRs to be present at all times then the buffer requirement is derived from the formula below.

2 X PCR Interval + Packet Delay Variation = Minimum Buffer Size

Example: If the PCR Interval is given as 100ms, and the packet delay variation is given as 60 ms, the buffer size would be 260 ms. If a shorter buffer time is required to reduce latency, a shorter PCR Interval is required such that the above formula is satisfied.

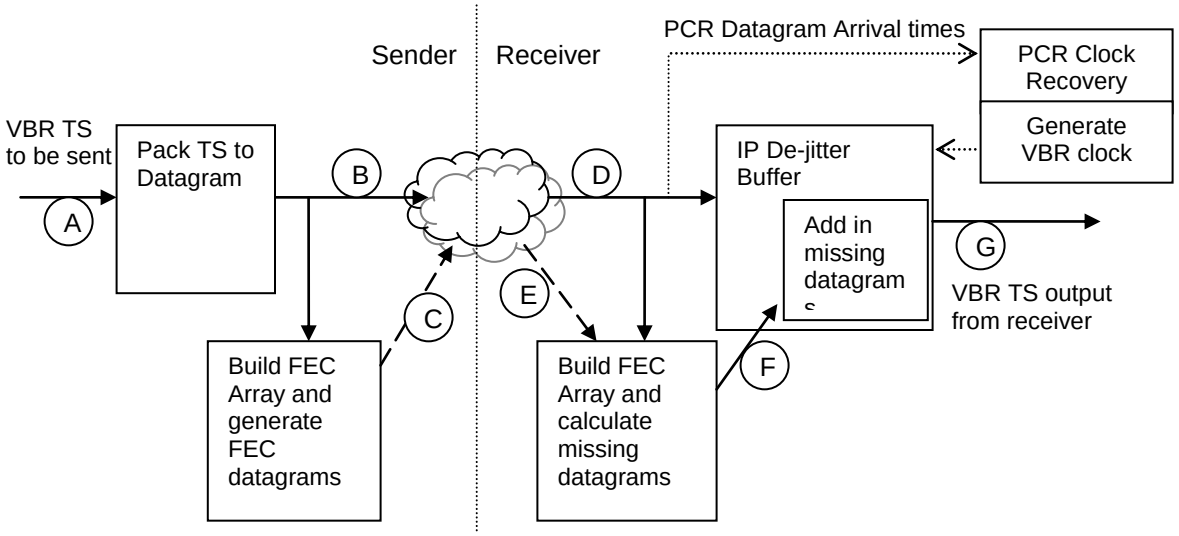

Figure C.1 – Block diagram of VBR MPEG->IP sender and receiver

The reference points A-G in Figure C.1 correspond to the same references in Figure C.2.

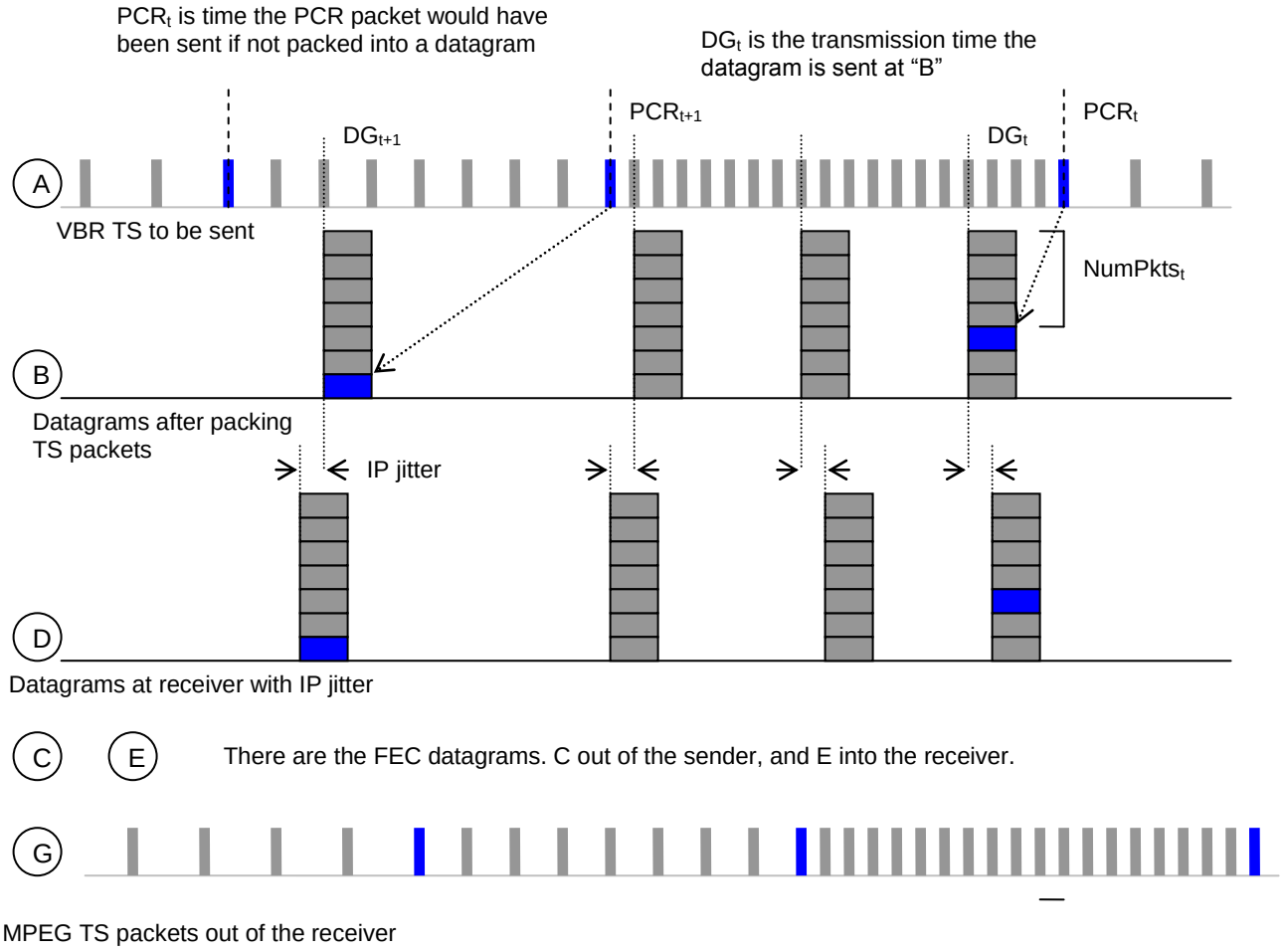

Figure C.2 –MPEG-2 transport stream packets and datagrams timing