AVT A. Begen
Internet-Draft E. Friedrich
Intended status: Standards Track Cisco
Expires: March 31, 2011 September 27, 2010

RTP Payload Format for MPEG2-TS Preamble
draft-begen-avt-rtp-mpeg2ts-preamble-06

Abstract

Demultiplexing and decoding an MPEG2 Transport Stream (MPEG2-TS) requires the knowledge of specific information about the transport stream, which we refer to as the MPEG2-TS Preamble. While this information is spread over different locations throughout the transport stream and can be eventually assembled after some time a receiver started receiving the MPEG2-TS, the time it takes to retrieve all this information (especially in multicast environments) may be long. Instead, having this information readily available as a Preamble and sending the Preamble to a receiver that will shortly start receiving the transport stream will virtually eliminate the waiting time and let the receiver start processing/decoding the MPEG2-TS sooner. In this document, we give an overview of the MPEG2-TS and the delay components in video systems, and motivate the need for constructing and using the MPEG2-TS Preamble for rapidly acquiring the source stream in RTP multicast sessions. We also define and register the RTP payload format for the MPEG2-TS Preamble.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

MPEG2 Transport Stream (MPEG2-TS) [MPEG2TS] is an encapsulation method and transport that multiplexes digital video and audio content, together with ancillary metadata, and produces a synchronized multiplexed stream that is tailored for transport over packet or cell-oriented networks. MPEG2-TS is ubiquitous in broadcast applications over both terrestrial and satellite networks. Both Advanced Television Systems Committee (ATSC) in North America and Digital Video Broadcasting (DVB) in Europe use MPEG2-TS in their standards. MPEG2-TS has been standardized by both ISO and ITU [MPEG2TS]. While MPEG2-TS was originally limited to carry MPEG-2 encoded content, the specification was later extended to cover MPEG-4/AVC audio/video encoding standards as well.

Due to the inherent design of MPEG2-TS, a receiver must first acquire certain information before demultiplexing and decoding an incoming MPEG2-TS. As will be explained in Section 4.1, this information resides in the transport stream. However, it is often not contiguous and is usually dispersed over a large period. Thus, a receiver starting to receive an MPEG2-TS at a random location will have to wait until the whole required information shows up in the received data. In multicast applications, since the joining receivers do not have any control over what point in the flow is currently being transmitted, their waiting times will vary. This problem has been identified and examined in detail in RAMS [I-D.ietf-avt-rapid-acquisition-for-rtp], where the time lag before a receiver can usefully consume the multicast data is referred to as the Acquisition Delay. Section 4 provides an overview of the delay components in video systems that contribute to the acquisition delay.

[I-D.ietf-avt-rapid-acquisition-for-rtp] refers to the information that must first be acquired before starting to process any data sent in the multicast session as the Reference Information. In this document, we refer to the subset of the Reference Information that is related to the MPEG2-TS as the MPEG2-TS Preamble.

For multicast applications running over RTP, [I-D.ietf-avt-rapid-acquisition-for-rtp] proposes an approach where an auxiliary unicast RTP session is established between a retransmission server and the joining RTP receiver. Over this unicast RTP session, the retransmission server provides the Reference Information the RTP receiver needs to rapidly acquire the multicast session. If the source stream in the RTP multicast session is carrying an MPEG2-TS, the Reference Information will comprise the MPEG2-TS Preamble as well. For its proper transmission from the retransmission server to the joining RTP receiver, a new RTP payload format has to be defined for the MPEG2-TS Preamble. This document

defines and registers this payload format.

Since the RTP packet(s) carrying the MPEG2-TS Preamble will not be able to fed to the decoder in the form they are received, a post-processing is required at the RTP receiver. This document also explains this process in Section 7.

2. Requirements Notation

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. Acronyms

This document uses the following acronyms frequently:

CAT: Conditional access table.

DTS: Decoding timestamp.

ECM: Entitlement control message.

EMM: Entitlement management message.

ES: Elementary stream.

GoP: Group of pictures.

IDR: Instantaneous decoding refresh.

MPEG2-TS: MPEG2 transport stream.

MPTS: Multi program transport stream.

PAT: Program association table.

PCR: Program clock reference.

PMT: Program map table.

PSI: Program specific information.

PTS: Presentation timestamp.

RAP: Random access point.

SPTS: Single program transport stream.

4. Elements of Delay in Video Systems

For typical multicast-based video delivery systems, the multicast switching delay (time required to leave the previous multicast session and join the new session) is not the primary contributor to the overall acquisition delay. The multicast flows are typically already present at the edge or deep in the network, the propagation delays for join operations are modest, and the multicast routers can typically process the Join and Leave messages quickly. Even if the edge multicast router is not currently a member of the requested multicast session, the multicast routing control messages propagate through the network rapidly and trees are built without experiencing large delays. Even in cases where a number of tree branches need to be built to the edge multicast router, this cost is frequently amortized over a large number of receivers such that only the first receiver joining the group experiences the increased delay. Further, this delay can be eliminated at the cost of extra bandwidth in the network core by having the edge routers do static joins for the set of sessions they expect receivers to be interested in. These techniques usually provide a well-bounded multicast switching delay.

Once the join operation completes and a receiver starts receiving media content for the first time in a multicast session, it often experiences a considerable amount of Reference Information latency and buffering delays. In the following subsections, we discuss the details of these delay elements using MPEG2-TS as the motivating use case.

4.1. Overview of MPEG-2 Transport Streams

MPEG2-TS is a container format that describes the schema of the audio and video content and the in-band control information. Prior to multiplexing, an audio and a video encoder output audio and video Elementary Streams (ES), respectively. The ES streams are then packetized to form the Packetized Elementary Streams (PES). The resulting elements are called PES packets. A transport stream (TS) encapsulates several PES streams and other data, and carries them in TS packets. The RTP payload format for carrying TS packets in an RTP stream is specified in [RFC2250]. In addition to the audio and video ES streams, there are ES streams that carry control data.

Program Specific Information (PSI) consists of metadata carried in the transport stream (See Table 2-28 in [MPEG2TS] for details). PSI includes Program Association Table (PAT), Conditional Access Table (CAT) and Program Map Table (PMT). A PAT has information about all the programs carried in the transport stream. It lists the 13-bit Program IDs (PID) for all the PMTs, associating them with the individual programs. Each of the ES streams of a particular program

in the transport stream also has the same PID values. This way, a decoder at the receiving side can extract the desired TS packets from the transport stream by checking their PID values. If the transport stream is not a Multi-Program Transport Stream (MPTS), but rather it is a Single-Program Transport Stream (SPTS), all the ES streams in the transport stream correspond to a single program.

CAT defines the type of the scrambling used (either at the PES or TS level), and identifies all the PID values of the TS packets that contain the Entitlement Management Messages (EMM). In addition to containing the PID values of each ES stream associated with a particular program, the PMT table also includes private data associated with the program such as the PID value of the packet containing the Entitlement Control Messages (ECM). The data contained in the EMM and ECM messages are vital in descrambling encrypted content. Note that PSI is carried in clear and is never scrambled so that a receiver which just started receiving the transport stream can process the PSI. The PAT, CAT and PMT tables must be parsed by the decoder in order to find the ES streams, private data as well as the encryption information for a given program.

MPEG2-TS produces output that is synchronized to a common clock across all the ESs in the multiplex. To assist the audio and video decoders, programs periodically provide a Program Clock Reference (PCR) value in the transport stream. PCR values are embedded in the TS adaptation field headers and are inserted by the encoder at least every 100 ms. A PCR timestamp represents the value of the encoder's system clock when it was sampled. The system clock is driven by a local 27 MHz oscillator.

PCR is extremely important in native MPEG-2 transport to keep the decoders synchronized. For example, the periodically sent Decoding Timestamps (DTS) and Presentation Timestamps (PTS) are specified relative to the PCR value and the decoders use the PCR value as the basis for a master clock during decoding and playout.

4.2. Reference Information Latency in Video Applications

We classify the Reference Information latency into two categories.

4.2.1. PSI (PAT/CAT/PMT) Acquisition Delay

As we discussed in Section 4.1, the video (and the audio as well) in an MPEG2-TS is self describing, and the receiver must parse certain control information in the PAT, CAT and PMT tables (i.e., PSI) contained in the transport stream in order to know how to parse the rest of the stream (i.e., to find the audio and video elementary

streams, private data and the encryption information for a given program).

Many video services employ content encryption and the encryption keys must be parsed as well for decrypting the content. In order to enable various system elements to process video effectively, certain portions of the stream are left unencrypted. The PAT/PMT tables are always in the clear. The structure of the ECMs is also in the clear, although the ECM content which contains keying material is encrypted. The PSI information is repeated periodically in the transport stream, thus, when a receiver joins the multicast session, it needs to wait until the next time PSI is sent in the transport stream.

4.2.2. Random Access Point Acquisition Delay

Conventional MPEG2 video encoders encode the video content in Groups of Pictures (GoP). Each GoP is encoded independently from other GoPs and starts with an intra-coded frame (I-frame) that does not have any reference to other frames in the same GoP, i.e., an I-frame contains the representation of an entire picture and can be decoded independently. Thus, the start of an I-frame is said to be a Random Access Point (RAP).

On the other hand, due to the temporal compression, rest of the frames in the same GoP may have references to the I-frame or to other frames in the same GoP. Due to this interdependency among the frames, one generally has to receive certain elements of the GoP prior to decoding or rendering any part of the GoP. For example, the decoder can decode a frame that is dependent on two other frames only after these two frames are decoded.

Usually, GoP durations are between 500 ms and one second. However, more advanced codecs may use longer GoPs to gain from the encoding efficiency. When a receiver joins the multicast session, it needs to wait until the next RAP shows up in the multicast stream before it can start decoding. Since the frame that is currently being multicast does not depend on the join time, the average time a receiver waits for RAP, i.e., the average RAP acquisition delay, is approximately equal to half of the GoP duration. Hence, for longer GoPs, the RAP acquisition delay is proportionally longer.

Advanced Video Coding (AVC) (also called MPEG4 part 10) compression is very similar to MPEG2 compression. It has a few more compression tools available, including Hierarchical GoPs. In a hierarchical GoP, the dependent frames of a GoP may reference the I-frame at the start of this GoP and the I-frame at the start of the next GoP. This additional dependency causes a longer RAP acquisition delay, as the decoder must receive two I-frames (spread between two logical GoPs)

before decoding can commence. In an Open GoP, a frame in one GoP may refer to a frame in a previous GoP. AVC also has the ability to insert Instantaneous Decoding Refresh (IDR) frames. Frames that follow an IDR frame cannot reference frames that precede an IDR frame. IDR frames are useful for editing AVC streams, but are typically do not appear often enough in streaming video to be useful in a stream acquisition context.

Note that in order for an intermediary network element such as a retransmission server to find the random access points in the video stream (e.g., I-frames), the necessary structural information must be in the clear if the intermediary is not in possession of the necessary keys.

4.3. Buffering Delays in Video Applications

We classify the buffering delays into two categories.

4.3.1. Network-Related Buffering Delays

In general, multicast-based video applications use an unreliable underlying transport protocol such as UDP [RFC0768] to distribute the content to a large number of receivers. This is largely due to the fact that these applications are one way in nature and providing closed-loop reliability does not scale well when the number of receivers is large or the acceptable playout delay is small, or both. Rather, if there is a need for reliability, the applications may employ one or more loss-repair methods to recover the packets missing at each receiver (The Reliable Multicast Transport Working Group has several standardized solutions for this problem. Refer to [RFC5740] for details). For example, Forward Error Correction (FEC) may be used proactively and/or on-demand to provide reliable transmission to a potentially very large multicast group in a scalable manner [I-D.ietf-fecframe-framework]. Similarly, retransmissions may be used in RTP-based multicast sessions where the retransmissions can be handled by local repair servers rather than the source itself [RFC5760]. However, regardless of the type of the loss-repair method(s) adopted by an application, loss-recovery operations always require additional buffering at the receiver side. The amount of buffering increases with the FEC block size when FEC is used, and with the round-trip time between the receiver and the local repair server when retransmission is used.

Audio and video decoders demand almost jitter-free content. If any jitter is introduced during the transmission in the network or due to the loss-repair methods, the jitter has to be smoothed out before the content is fed to the decoder. This is called de-jittering and it usually adds up to the buffering delay.

4.3.2. Application-Related Buffering Delays

The application buffering requirements for MPEG-encoded content are quite rigorous, particularly for the MPEG-based video applications. Video compression devices apply more bits to represent certain scenes than they do for other scenes. A very complex scene (individual picture) requires considerably more information than a simple scene. Furthermore, pictures that are entirely intra-coded, e.g., I-frames, consume more bits compared to pictures that make use of predictive coding. Each scene is shown by the decoder at a certain fixed frame rate (e.g., 24 fps or 30 fps). Since some scenes are comprised more bits than other scenes, the output rate of the decoder buffer is usually variable. However, the network flow is typically Constant Bit Rate (CBR) or Capped Variable Bit Rate (Capped VBR). The net effect is that the input rate to the decoder buffer is close to constant, but the output rate is highly variable.

The video encoders keep track of the decoder buffer size, and use this information to regulate the temporal compression. This forces the decoder buffer to "breathe." In order to avoid underflow, the decoder buffer must fill up to a certain level prior to starting to decode and play the content. The decoder buffer size required to avoid underflow is dependent on the encoder, and the encoder signals the decoder buffering requirements in-band. Typical decoder buffer requirements for MPEG2 content range from a few hundreds of milliseconds to a few seconds. However, AVC/MPEG4 part 10 encoders usually tend to use more temporal compression, and thus require a larger buffer at the decoder side. This consequently increases the buffering delays.

5. Packet Format

This section defines the MPEG2-TS Preamble packet format.

The RTP header is formatted according to [RFC3550] with some further clarifications listed below:

- o Marker (M) Bit: This bit SHALL be used to indicate the last packet carrying the MPEG2-TS Preamble information. It SHALL be set to zero in all RTP packets other than the last packet, where it SHALL be set to one.

- o Payload Type: The (dynamic) payload type for the MPEG2-TS Preamble packets is determined through out-of-band means. Note that this document registers a new payload format for the MPEG2-TS Preamble packets (Refer to Section 6 for details). According to [RFC3550], an RTP receiver that cannot recognize a payload type must discard it. This provides backward compatibility.

- o Sequence Number (SN): The sequence number has the standard definition. It MUST be one higher than the sequence number in the previously transmitted RTP packet. If this RTP packet is the first packet in the session, its sequence number SHOULD be random (unpredictable) [RFC3550].

- o Timestamp (TS): The timestamp SHALL be set to a time corresponding to the packet's transmission time.

- o Synchronization Source (SSRC): Per [RFC3550], the SSRC value SHOULD be chosen randomly with collision detection.

However, in RAMS applications [I-D.ietf-avt-rapid-acquisition-for-rtp] where the MPEG2-TS Preamble packets are payload-type multiplexed with the burst packets in the same RTP session, the following rules apply:

- o The SSRC of the Preamble packets MUST be set to the SSRC of the retransmission packets.

- o The Preamble and retransmission packets share the same sequence number and timestamp space. Thus, the sequence numbers and timestamps MUST be set consistently.

The payload consists of TOLV elements that are defined in Section 5.2. Before defining them, we first introduce the TOLV structure.

5.1. Extensions

The MPEG2-TS Preamble MUST be encoded as TOLV elements as described below and sketched in Figure 1:

- o Type: A single-octet identifier that defines the type of the parameter represented in this TOLV element.

- o Order: A single-octet identifier that specifies the order in which this TOLV element MUST be processed by the receiver when forming a demux/decoder-friendly stream. As explained in Section 7, the order information is crucial for majority of the TOLV elements defined in this document. Yet, for some other TOLV elements, the order may not matter. In that case, the Order value MAY be set to 0.

  Starting from 1, a lower Order value indicates an earlier order in post-processing and the Order values MUST be consecutive. Note that none of the TOLV elements that belong to the same Preamble data MUST have the same Order value except 0.

- o Length: A two-octet field that indicates the length of the TOLV element excluding the Type, Order and Length fields in octets. Note that this length does not include any padding that is required for alignment.

- o Value: Variable-size set of octets that contains the specific value for the parameter. Note that the value field of an TOLV element may contain other TOLVs.

If a TOLV element does not fall on a 32-bit boundary, the last word must be padded to the boundary using further bits set to 0. The support for extensions is OPTIONAL.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |     Order     |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             Value                             /
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: Structure of a TOLV element

The TOLV encoding/decoding is similar to the TLV encoding/decoding used by the RAMS applications [I-D.ietf-avt-rapid-acquisition-for-rtp].

5.1.1. Vendor-Neutral Extensions

If the goal in defining new TOLV elements is to extend the functionality in a vendor-neutral manner, they MUST be registered with IANA through the guidelines provided in Section 10.2.

The current document defines several vendor-neutral extensions in Section 5.2.

5.1.2. Private Extensions

It is desirable to allow vendors to use private extensions in TOLV format. For interoperability, such extensions MUST NOT collide with each other.

A certain range of TOLV Types is reserved for private extensions (Refer to Section 10.2). IANA management for these extensions is unnecessary and they are the responsibility of individual vendors.

The structure that MUST be used for the private extensions is depicted in Figure 2. Here, the enterprise numbers are used from http://www.iana.org/assignments/enterprise-numbers. This will ensure the uniqueness of the private extensions and avoid any collision.

```
  0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |     Type      |     Order     |             Length            |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                          Ent. Number                          |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                             Value                             /
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: Structure of a private extension

5.2. Vendor-Neutral Extensions

The current document defines several vendor-neutral extensions. In the following extensions, we use 'r' to indicate a reserved bit, which SHALL be set to zero. The fields denoted by 'Reserved' SHALL also be set to zero.

5.2.1. PAT TOLV

Optional TOLV element that is used to encode information associated with a Program Association (PA) section as defined in Section 2.4.4.3

of [MPEG2TS]. It has the following structure:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| PAT_TOLV_Type |     Order     |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            PID          |r r r|        Section Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Section Data                         /
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: Structure of PAT TOLV

The PID is a 13-bit field used to identify the ES in an MPTS or SPTS. This is the PID used in the associated TS to carry the PA Section Data. Its value MUST be 0. The Section Length field is a 16-bit field that specifies the length of the Section Data in octets. The Section Data is a PA section as defined in Section 2.4.4.3 of [MPEG2TS]. Note that the length of the PAT TOLV is variable.

Type: 1

Length: Variable

5.2.2. PMT TOLV

Optional TOLV element that is used to encode information associated with a Program Map (PM) section. It has the following structure:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| PMT_TOLV_Type |     Order     |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            PID          |r r r|        Section Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Section Data                         /
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: Structure of PMT TOLV

The PID is used in the associated TS to carry the PM Section Data. The Section Length field is a 16-bit field that specifies the length of the Section Data in octets. The Section Data is a PM section. Note that the length of the PMT TOLV is variable.

Type: 2

Length: Variable

5.2.3. PCR TOLV

Optional TOLV element that contains a PCR value used to initialize the decoder and system clocks. The PCR value corresponds to the first byte of the first TS packet that will be transmitted as part of the unicast burst. It has the following structure:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| PCR_TOLV_Type |     Order     |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            PID          |r r r|  Reserved   |     PCR_EXT     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           PCR_BASE                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|b|                         Reserved                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: Structure of PCR TOLV

The PID is used in the associated TS to carry the PCR data. PCR_BASE is a 33-bit field that is part of the PCR timestamp. PCR_BASE occupies the entire third 32-bit word along with the first bit of the fourth word. The remainder of the fourth word (31 bits) is reserved. PCR_EXT is a 9-bit field that is part of the PCR timestamp.

Type: 3

Length: 13

5.2.4. PID_LIST TOLV

A PID is a packet identifier that is used to identify ESs of a program in an SPTS or MPTS. The PID_LIST contains a PID Element for each PID referenced in the MPEG2-TS Preamble. Each PID Element contains the PID and associated continuity counter that corresponds to the first packet that will be sent.

PID_LIST TOLV It has the following structure:

```
  0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |PID_L_TOLV_Type|     Order     |            Length             |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                         PID Element 1                         |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 :                                                               :
 :                                                               :
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                         PID Element n                         |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 6: Structure of PID_LIST TOLV

And the PID Element has the following structure:

```
  0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |          PID            |r r r|r r r r|  CC   |   Reserved    |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 7: Structure of PID Element

Here, CC is a 4-bit field that carries the Continuity Counter.

Type: 4

Length: Variable

5.2.5. SEQ TOLV

Optional TOLV element that is used to encode information from the Video Sequence Header of an MPEG2 Video ES. The Video Sequence Header contains information such as frame rate, aspect ratio and picture size. It has the following structure:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| SEQ_TOLV_Type |     Order     |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            PID          |r r r|        Section Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Section Data                        /
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 8: Structure of SEQ TOLV

The PID is used in the associated TS to carry the SEQ Section Data. The Section Length field is a 16-bit field that specifies the length of the Section Data in octets. The Section Data is a SEQ section. See Section 6.2.2 of [ISO13818-2] for a discussion of Video Sequence Header and Sequence Extension.

Note that SEQ TOLV is only applicable to transport streams that carry MPEG2 video.

Type: 5

Length: Variable

5.2.6. SPS TOLV

Optional TOLV element that is used to encode information from the Sequence Parameter Set Network Abstraction Layer (NAL) unit. It has the following structure:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| SPS_TOLV_Type |     Order     |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            PID          |r r r|        Section Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Section Data                        /
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 9: Structure of SPS TOLV

The PID is used in the associated TS to carry the SPS Section Data. The Section Length field is a 16-bit field that specifies the length of the Section Data in octets. The Section Data is a SPS section. See Section 7.4.2.1 of [ISO13818-10] for a discussion of semantics of

the Sequence Parameter Set NAL.

Note that SPS TOLV is only applicable to transport streams that carry AVC (H.264) video.

Type: 6

Length: Variable

5.2.7. PPS TOLV

Optional TOLV element that is used to encode information from the Picture Parameter Set NAL unit. It has the following structure:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| PPS_TOLV_Type |     Order     |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            PID          |r r r|        Section Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Section Data                         /
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 10: Structure of PPS TOLV

The PID is used in the associated TS to carry the PPS Section Data. The Section Length field is a 16-bit field that specifies the length of the Section Data in octets. The Section Data is a PPS section. See Section 7.4.2.2 of [ISO13818-10] for a discussion of semantics of the Picture Parameter Set NAL.

Note that PPS TOLV is only applicable to transport streams that carry AVC (H.264) video.

Type: 7

Length: Variable

5.2.8. SEI TOLV

Optional TOLV element that is used to encode information from the Supplemental Enhanced Information NAL unit. It has the following structure:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| SEI_TOLV_Type |     Order     |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            PID          |r r r|       Section Length          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Section Data                         /
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 11: Structure of SEI TOLV

The PID is used in the associated TS to carry the SEI Section Data. The Section Length field is a 16-bit field that specifies the length of the Section Data in octets. The Section Data is a SEI section. See Annex D of [ISO13818-10] for details.

Note that SEI TOLV is only applicable to transport streams that carry AVC (H.264) video.

Type: 8

Length: Variable

5.2.9. ECM TOLV

Optional TOLV element that contains the ECM from the MPEG2-TS. This applies only to transport streams that are encrypted. It has the following structure:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| ECM_TOLV_Type |     Order     |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            PID          |r r r|       Section Length          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Section Data                         /
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 12: Structure of ECM TOLV

Section Data depends on type of encryption used. It contains a Private Section as defined in Section 2.4.4.10 of [MPEG2TS]. The MPEG2-TS PMT provides a list of PIDs in conditional access descriptors. Private Sections from these PIDs are placed into ECM TOLVs. There may be multiple ECM TOLV elements corresponding to

different PID or table_id values. The contents and use of such a Private Section are vendor specific. It is treated here as opaque data. Conditional access vendors will often use table_id = 0x80 and table_id = 0x81 in accordance with [DVB-ETR-289].

Type: 9

Length: Variable

5.2.10. EMM TOLV

Optional TOLV element that contains the EMM from the MPEG2-TS. This applies only to transport streams that are encrypted. It has the following structure:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| EMM_TOLV_Type |     Order     |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            PID          |r r r|        Section Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Section Data                         /
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 13: Structure of EMM TOLV

Section Data depends on type of encryption used. It contains a Private Section as defined in Section 2.4.4.10 of [MPEG2TS]. The MPEG2-TS CAT provides a list of PIDs in conditional access descriptors. Private Sections from these PIDs are placed into EMM TOLVs. There may be multiple EMM TOLV elements corresponding to different PID or table_id values. The contents and use of such a Private Section are vendor specific. It is treated here as opaque data.

Type: 10

Length: Variable

5.2.11. CAT TOLV

Optional TOLV element that is used to encode information associated with a Conditional Access (CA) section as defined in Section 2.4.4.6 of [MPEG2TS]. It has the following structure:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| CAT_TOLV_Type |     Order     |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            PID          |r r r|        Section Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Section Data                         /
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 14: Structure of CAT TOLV

The PID is a 13-bit field used to identify the ES in an MPTS or SPTS. This is the PID used in the associated TS to carry the CA Section Data. The Section Length field is a 16-bit field that specifies the length of the Section Data in octets. The Section Data is a CA section as defined in Section 2.4.4.6 of [MPEG2TS].

Type: 11

Length: Variable

5.2.12. PTS TOLV

Optional TOLV element that is used to encode the PTS timestamp corresponding to the first picture in the unicast burst. It has the following structure:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| PTS_TOLV_Type |     Order     |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            PID          |r r r|           Reserved            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            PTS_BASE                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|b|                          Reserved                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 15: Structure of PTS TOLV

The PID is used to identify the associated TS to carry the PTS data. PTS_BASE is a 33-bit field that is taken from the PES header of the MPEG2-TS that will be sent. Note that this TOLV is purely informational.

Type: 12

Length: 13

6. Payload Format Parameters

This section provides the media subtype registration for the MPEG2-TS Preamble.

6.1. Media Type Registration

This registration is done using the template defined in [RFC4288] and following the guidance provided in [RFC4855].

6.1.1. Registration of audio/mpeg2-ts-preamble

Type name: audio

Subtype name: mpeg2-ts-preamble

Required parameters:

- o rate: The RTP timestamp (clock) rate. The rate SHALL be larger than 1000 Hz to provide sufficient resolution to RTCP operations.

Optional parameters: None.

Encoding considerations: This media type is framed (See Section 4.8 in the template document [RFC4288]) and contains binary data.

Security considerations: See Section 9 of this document.

Interoperability considerations: None.

Published specification: This document.

Applications that use this media type: Applications that transmit MPEG2-TS content and want to provide the Preamble information for the transport stream they are transmitting to the receiver(s).

Additional information: None.

Person & email address to contact for further information: Ali Begen <abegen@cisco.com> and IETF Audio/Video Transport Working Group.

Intended usage: COMMON.

Restriction on usage: This media type depends on RTP framing, and hence, is only defined for transport via RTP [RFC3550].

Author: Ali Begen <abegen@cisco.com>.

Change controller: IETF Audio/Video Transport Working Group delegated from the IESG.

6.1.2. Registration of video/mpeg2-ts-preamble

Type name: video

Subtype name: mpeg2-ts-preamble

Required parameters:

- o rate: The RTP timestamp (clock) rate. The rate SHALL be larger than 1000 Hz to provide sufficient resolution to RTCP operations.

Optional parameters: None.

Encoding considerations: This media type is framed (See Section 4.8 in the template document [RFC4288]) and contains binary data.

Security considerations: See Section 9 of this document.

Interoperability considerations: None.

Published specification: This document.

Applications that use this media type: Applications that transmit MPEG2-TS content and want to provide the Preamble information for the transport stream they are transmitting to the receiver(s).

Additional information: None.

Person & email address to contact for further information: Ali Begen <abegen@cisco.com> and IETF Audio/Video Transport Working Group.

Intended usage: COMMON.

Restriction on usage: This media type depends on RTP framing, and hence, is only defined for transport via RTP [RFC3550].

Author: Ali Begen <abegen@cisco.com>.

Change controller: IETF Audio/Video Transport Working Group delegated from the IESG.

6.1.3. Registration of text/mpeg2-ts-preamble

Type name: text

Subtype name: mpeg2-ts-preamble

Required parameters:

- o rate: The RTP timestamp (clock) rate. The rate SHALL be larger than 1000 Hz to provide sufficient resolution to RTCP operations.

Optional parameters: None.

Encoding considerations: This media type is framed (See Section 4.8 in the template document [RFC4288]) and contains binary data.

Security considerations: See Section 9 of this document.

Interoperability considerations: None.

Published specification: This document.

Applications that use this media type: Applications that transmit MPEG2-TS content and want to provide the Preamble information for the transport stream they are transmitting to the receiver(s).

Additional information: None.

Person & email address to contact for further information: Ali Begen <abegen@cisco.com> and IETF Audio/Video Transport Working Group.

Intended usage: COMMON.

Restriction on usage: This media type depends on RTP framing, and hence, is only defined for transport via RTP [RFC3550].

Author: Ali Begen <abegen@cisco.com>.

Change controller: IETF Audio/Video Transport Working Group delegated from the IESG.

6.1.4. Registration of application/mpeg2-ts-preamble

Type name: application

Subtype name: mpeg2-ts-preamble

Required parameters:

- o rate: The RTP timestamp (clock) rate. The rate SHALL be larger than 1000 Hz to provide sufficient resolution to RTCP operations.

Optional parameters: None.

Encoding considerations: This media type is framed (See Section 4.8 in the template document [RFC4288]) and contains binary data.

Security considerations: See Section 9 of this document.

Interoperability considerations: None.

Published specification: This document.

Applications that use this media type: Applications that transmit MPEG2-TS content and want to provide the Preamble information for the transport stream they are transmitting to the receiver(s).

Additional information: None.

Person & email address to contact for further information: Ali Begen <abegen@cisco.com> and IETF Audio/Video Transport Working Group.

Intended usage: COMMON.

Restriction on usage: This media type depends on RTP framing, and hence, is only defined for transport via RTP [RFC3550].

Author: Ali Begen <abegen@cisco.com>.

Change controller: IETF Audio/Video Transport Working Group delegated from the IESG.

6.2. Mapping to SDP Parameters

Applications that are using RTP transport commonly use Session Description Protocol (SDP) [RFC4566] to describe their RTP sessions. The information that is used to specify the media types in an RTP session has specific mappings to the fields in an SDP description. In this section, we provide these mappings for the media subtype registered by this document ("mpeg2-ts-preamble"). Note that if an application does not use SDP to describe the RTP sessions, an appropriate mapping must be defined and used to specify the media types and their parameters for the control/description protocol employed by the application.

The mapping of the media type specification for "mpeg2-ts-preamble" and its parameters in SDP is as follows:

- The media type (e.g., "application") goes into the "m=" line as the media name.

- The media subtype ("mpeg2-ts-preamble") goes into the "a=rtpmap" line as the encoding name. The RTP clock rate parameter ("rate") also goes into the "a=rtpmap" line as the clock rate.

An SDP example is provided in Section 8.

6.2.1. Offer-Answer Model and Declarative Considerations

There no configuration parameters or optional format parameters for the MPEG2-TS Preamble payload format. Thus, when offering MPEG2-TS Preamble over RTP using SDP in an Offer/Answer model [RFC3264] or in a declarative manner (e.g., SDP in the Real-time Streaming Protocol (RTSP) [RFC2326] or the Session Announcement Protocol (SAP) [RFC2974]), there are no specific considerations.

7. Post-Processing of the MPEG2-TS Preamble

Once received, the RTP packet(s) carrying the MPEG2-TS Preamble cannot be fed directly to the MPEG transport demux/decoder since the demux/decoder would not be able to recognize the TOLV elements within the Preamble packets. Thus, the Preamble information MUST be first processed, and a demux/decoder-friendly stream MUST be formed. The demux/decoder-friendly stream MUST conform to [MPEG2TS]. In this section, we briefly explain this process.

The Preamble TOLV elements contain several different types of Transport Stream data. The first type is Program Specific Information (PSI) Section Data. The second type is PCR Data, and the third is Elementary Stream Data.

1. Section Data

   This includes the PAT TOLV, PMT TOLV and CAT TOLV. Section Data also include ECMs or EMMs encapsulated in private sections as per [DVB-ETR-289]. These packets are all processed similarly.

   Each TS packet begins with a 4-byte TS header containing the PID value given in the Preamble TOLV element. There is no adaptation field present. The continuity counter value can be retrieved from the PIDLIST TOLV element for the current PID. If the TS packet contains the beginning of the Section, the Payload Unit Start Indicator (PUSI) bit should be set in the TS header. Also, a 1-byte pointer field follows the TS header, which should be set to 0. The Section Data immediately follow the pointer field. If the Section Data span multiple TS packets, they are recreated in a similar manner, except without the PUSI bit set and the pointer field.

   The continuity counter value of a packet must be one greater than the previous packet on that PID. This applies to packets reconstructed from the Preamble TOLV elements and packets from the MPEG2-TS stream. The continuity counter value of the final packet of Preamble data should be one less than the first packet on the MPEG2-TS stream with that same PID.

   The final TS packet may contain stuffing bytes of 0xFF as necessary to pad the packet to 188 bytes.

   The Section Data may optionally be programmed automatically into decoder hardware registers. This allows the decoder to rapidly learn the PSI data without having to parse the MPEG2-TS preamble output. This technique decreases presentation times by decreasing the amount of time needed to acquire the relevant PSI

data in the decoder. It also provides an opportunity to learn when the demux/decoder is ready to accept the unicast (RAMS) burst [I-D.ietf-avt-rapid-acquisition-for-rtp]. If the unicast burst is given to the decoder before it is fully programmed with the PSI data, the decoder may discard part of the unicast burst, which delays the presentation time significantly.

2. PCR Data

A single TS packet is constructed from the PCR TOLV. This contains a 4-byte TS header with correct PID and continuity counter as retrieved from the PIDLIST. The Adaptation Field control bits should indicate that only an adaptation_field and no payload is present. The adaptation field should contain the PCR value from the PCR TOLV followed by padding to fill the TS packet. The discontinuity_indicator in the Adaptation Field should be set to one.

The PCR value MUST be constructed to be continuous with the MPEG2-TS timebase. The PCR TOLV packet contains the PCR value of the first packet in the unicast burst and this value MUST be adjusted by an amount equal to the time needed to the write the preamble into the decoder.

3. Elementary Stream Data

The final type of Preamble data is that contained in an elementary stream. This includes data from the SEQ, SPS, PPS and SEI TOLVs. Each TS packet begins with a 4-byte TS header containing the PID value given in the Preamble TOLV element. The adaptation field padding can be used to pad the last TS packet to 188 bytes. The continuity counter value can be retrieved from the PIDLIST TOLV element for the current PID. If the TS packet contains the beginning of the TOLV element, the Payload Unit Start Indicator (PUSI) bit should be set in the TS header. A PES header must immediately follow the start of the first TS packet. The stream_type of the PES packet should match the contents of elementary stream data.

The TS packets constructed as above should be passed to the demux/ decoder in the following order: PAT, PMT, PCR, EMM, ECM, {Elementary Stream Data}. The PAT and PMT MUST be the first two packets because they contain required information to program the demux. The EMM and ECM MUST occur before Elementary stream data because they contain conditional access data that will be needed to descramble and decode the elementary stream.

For all PIDs, the continuity counter values MUST be reconstructed to

the continuous with the RAMS session continuity counter values. The PIDLIST TOLV contains the continuity counter value of the first TS packet for each PID present in the preamble. For every TS packet generated in the preamble output, the continuity counter value SHOULD be decremented by 1. For example, the PIDLIST TOLV indicates the continuity counter of the first burst packet on PID 0x0024 is 0xe and the preamble contains two packets on PID 0x0024. The first packet on PID 0x0024 from the preamble sent into the decoder should have continuity counter 0xc and the second preamble packet should have continuity counter 0xd.

8. Session Description Protocol (SDP) Signaling

This section shows how the MPEG2-TS Preamble packets can be used in a RAMS session [I-D.ietf-avt-rapid-acquisition-for-rtp]. The example SDP [RFC4566] description is taken from [I-D.ietf-avt-rapid-acquisition-for-rtp], where additional information about this description is available.

```
v=0
o=ali 1122334455 1122334466 IN IP4 rams.example.com
s=Rapid Acquisition Example with MPEG2-TS Preamble Data
t=0 0
a=group:FID 1 2
a=rtcp-unicast:rsi
m=video 41000 RTP/AVPF 98
i=Primary Multicast Stream
c=IN IP4 233.252.0.2/255
a=source-filter:incl IN IP4 233.252.0.2 192.0.2.2
a=rtpmap:98 MP2T/90000
a=rtcp:42000 IN IP4 192.0.2.1
a=rtcp-fb:98 nack
a=rtcp-fb:98 nack rai
a=ssrc:123321 cname:iptv-ch32@rams.example.com
a=mid:1
m=video 51000 RTP/AVPF 99 100
i=Unicast Retransmission Stream + Preamble Data
c=IN IP4 192.0.2.1
a=sendonly
a=rtpmap:99 rtx/90000
a=rtcp-mux
a=fmtp:99 apt=98;rtx-time=5000
a=rtpmap:100 mpeg2-ts-preamble/90000
a=mid:2
```

Figure 16: SDP description showing the payload-type multiplexed Preamble and retransmission packets in a RAMS session

9. Security Considerations

RTP packets using the payload format defined in this specification are subject to the security considerations discussed in the RTP specification [RFC3550] and in any applicable RTP profile. The main security considerations for the RTP packet carrying the RTP payload format defined within this memo are confidentiality, integrity and source authenticity. Confidentiality is achieved by encrypting the RTP payload. Integrity of the RTP packets is achieved through a suitable cryptographic integrity protection mechanism. Such a cryptographic system may also allow the authentication of the source of the payload. A suitable security mechanism for this RTP payload format should provide confidentiality, integrity protection, and at least source authentication capable of determining if an RTP packet is from a member of the RTP session.

Note that the appropriate mechanism to provide security to RTP and payloads following this memo may vary. It is dependent on the application, transport and signaling protocol employed. Therefore, a single mechanism is not sufficient, although if suitable, using the Secure Real-time Transport Protocol (SRTP) [RFC3711] is recommended. Other mechanisms that may be used are IPsec [RFC4301] and Transport Layer Security (TLS) [RFC5246] (RTP over TCP); other alternatives may exist.

The primary application area for the Preamble packets is to provide accelerated acquisition and presentation of multicast sessions carrying MPEG2-TS content [I-D.ietf-avt-rapid-acquisition-for-rtp]. During the creation of the Preamble packets, no new stream data are created. That is, the only data present in the Preamble are gathered from the recent packets sent in the corresponding primary multicast stream. Thus, there are no additional security risks for an attacker which may capture both the Preamble packets and the primary multicast stream. The Preamble data can contain EMM or ECM packets, but they are also drawn from the multicast stream. No personalized data is included in the Preamble packets, so the Preamble data may not be used to decrypt the stream data. Modifying the Preamble data or preventing the Preamble data from reaching the RTP receiver could lead to increased acquisition delays or presentation artifacts.

10. IANA Considerations

10.1. Payload Format

New media subtypes are subject to IANA registration. For the registration of the payload format and its parameters introduced in this document, refer to Section 6.

10.2. MPEG2-TS Preamble TOLV Space Registry

This document creates a new IANA TOLV space registry for the extensions. The registry is called the MPEG2-TS Preamble TOLV Space Registry. This registry is to be managed by the IANA according to the Specification Required policy of [RFC5226].

The length of the Type field in the TOLV elements is a single octet, allowing 256 values. The registry is initialized with the following entries:

Type	Description	Reference
1	PAT (Program Association Table) Data	This document
2	PMT (Program Map Table) Data	This document
3	PCR (Program Clock Reference) Data	This document
4	PID List	This document
5	SEQ (Video Sequence Header) Data	This document
6	SPS (Sequence Parameter Set) Data	This document
7	PPS (Picture Parameter Set) Data	This document
8	SEI (Supplemental Enhanced Information) Data	This document
9	ECM (Entitlement Control Message) Data	This document
10	EMM (Entitlement Management Message) Data	This document
11	CAT (Conditional Access Table) Data	This document
12	PTS (Presentation Timestamp)	This document

The Type values 0 and 255 are reserved for future use. The Type values between (and including) 128 and 254 are reserved for private extensions.

Any registration for an unassigned Type value needs to contain the following information:

- o Contact information of the one doing the registration, including at least name, address, and email.

- o A detailed description of what the new TOLV element represents and how it shall be interpreted.

11. Open-Source Implementation

An open-source RTP receiver code that implements the functionalities introduced in [I-D.ietf-avt-rapid-acquisition-for-rtp] and this document is available. For documentation, visit the following URL:

http://www.cisco.com/en/US/docs/video/cds/cda/vqe/3_5/user/guide/vqe_guide3_5.html

The code is available at:

ftp://ftpeng.cisco.com/ftp/vqec/

Note that this code is under development and may be based on an earlier versions of [I-D.ietf-avt-rapid-acquisition-for-rtp] and this document. As progress is made in the specifications, the source code will be updated to reflect the changes. Also note that the current version of the source code packages the Preamble information in RTCP APP packet(s) rather than RTP packet(s).

12. Acknowledgments

The authors would like to thank Dave Oran, Bill Ver Steeg, Art Allison, Sandy MacInnis, Roni Even and Ross Finlayson for their reviews and feedback.

13. References

13.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.

[MPEG2TS] ITU-T H.222.0, "Generic Coding of Moving Pictures and Associated Audio Information: Systems", May 2006.

[ISO13818-2] ITU-T H.262, "Generic Coding of Moving Pictures and Associated Audio Information: Video", February 2000.

[ISO13818-10] ITU-T H.264, "Advanced Video Coding for Generic Audiovisual Services", March 2005.

[RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.

[RFC4288] Freed, N. and J. Klensin, "Media Type Specifications and Registration Procedures", BCP 13, RFC 4288, December 2005.

[RFC4855] Casner, S., "Media Type Registration of RTP Payload Formats", RFC 4855, February 2007.

[RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, June 2002.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

13.2. Informative References

[I-D.ietf-avt-rapid-acquisition-for-rtp] Steeg, B., Begen, A., Caenegem, T., and Z. Vax, "Unicast-Based Rapid Acquisition of Multicast RTP Sessions", draft-ietf-avt-rapid-acquisition-for-rtp-15 (work in progress), September 2010.

[RFC2250] Hoffman, D., Fernando, G., Goyal, V., and M. Civanlar,

"RTP Payload Format for MPEG1/MPEG2 Video", RFC 2250, January 1998.

[RFC5760] Ott, J., Chesterfield, J., and E. Schooler, "RTP Control Protocol (RTCP) Extensions for Single-Source Multicast Sessions with Unicast Feedback", RFC 5760, February 2010.

[RFC0768] Postel, J., "User Datagram Protocol", STD 6, RFC 768, August 1980.

[RFC5740] Adamson, B., Bormann, C., Handley, M., and J. Macker, "NACK-Oriented Reliable Multicast (NORM) Transport Protocol", RFC 5740, November 2009.

[I-D.ietf-fecframe-framework] Watson, M., "Forward Error Correction (FEC) Framework", draft-ietf-fecframe-framework-10 (work in progress), September 2010.

[RFC2326] Schulzrinne, H., Rao, A., and R. Lanphier, "Real Time Streaming Protocol (RTSP)", RFC 2326, April 1998.

[RFC2974] Handley, M., Perkins, C., and E. Whelan, "Session Announcement Protocol", RFC 2974, October 2000.

[DVB-ETR-289] DVB ETR-289, "Digital Video Broadcasting (DVB); Support for use of scrambling and Conditional Access (CA) within digital broadcasting systems", October 1996.

[RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, March 2004.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, December 2005.

[RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, August 2008.

Authors' Addresses

Ali Begen
Cisco
181 Bay Street
Toronto, ON M5J 2T3
Canada

Email: abegen@cisco.com

Eric Friedrich
Cisco
1414 Massachusetts Ave.
Boxborough, MA 01719
USA

Email: efriedri@cisco.com