RTP Encapsulation of MPEG1/MPEG2

Don Hoffman, Gerard Fernando, Vivek Goyal

2 Profiles

1 - Supports maximum interoperability with MPEG2 Systems Transport Stream (MTS). Maintains full functionality of MTS

2 - Supports sufficient functionality for IP-based end-systems.

Sun Microsystems Computer Company

MPEG - Background

- ISO - Motion Picture Experts Group
- MPEG1 and MPEG2 International Standards (IS)
- MPEG standards consist of Video, Audio and Systems specifications
- MPEG Systems supports packetization, timestamping and multiplexing of compressed Video and Audio streams.

MPEG1

- Aimed mainly at CD-ROM apps.
- Video resolutions
  - 352x240 for NTSC countries
  - 352x288 for PAL/SECAM countries
  - 320x240 for square pixel format
  - VHS quality
- Video compression

  Three types of pictures
  - I - Intra coding (similar to JPEG)
  - P - Predictive coding (uses motion vectors)
  - B - Bidirectional coding (uses motion vectors)
  - For the same picture quality, ratio of bits required for I:P:B - 4:2:1

MPEG1

Continued:-

- Audio
    - 2 channel
    - Sample rates of 32K, 44.1K and 48 KHz
    - Compressed rates up to 384kb/s
    - Near CD quality

- Systems
    - Supports packetization, timestamping and multiplexing

MPEG2

- Aimed at a wider range of applications - CD-ROM, broadcast video, cable services, desktop applications, etc.
- Several profiles from SIF resolution up to HDTV
- Bit-rates from 1.5 Mb/s for SIF up to 30 Mb/s for HDTV
- Most common resolutions covered by Main Profile at Main Level (MP@ML)
    - 720(H)x480(V)x60(T) interlace for NTSC countries
    - 720(H)x576(V)x50(T) interlace for PAL/SECAM countries
    - Good NTSC/PAL quality at 4-6 Mb/s

MPEG2

Continued:-

- Audio
  - 5 channel (surround sound)
- Systems
  - Suitable for either storage media, or for distribution over networks. Performs packetization, timestamping and multiplexing of audio and video from one entertainment program, as well as multiplexing multiple entertainment programs.
  - Timestamps used for :

    i. Audio/video synchronization (i.e. for "lip-synch") - presentation timestamp (PTS).

    ii. locking of decoder systems clock to that of the encoder. This prevents buffer overflow (or underflow) - program clock reference (PCR).

MPEG2

- Systems (*continued)*

    - Two Systems streams

    i. Transport Stream - transmission over error-prone channels.

    ii. Program Stream - storage or transmission over error-free channels.

- MTS is likely to be the most widely implemented of the MPEG2 Systems specs. We propose a profile to map MTS over RTP.

Encapsulation of MPEG2 Transport Streams (MTS)

- Utilizes the timestamping and packetization features of MTS
- In MPEG2 Systems spec. PTS (presentation timestamp) need not be in every MTS packet.
- We propose PTS in every RTP fixed header.
- PTS defined to 33 bits with an accuracy of 90KHz.
- Only 32 LSB's. Hence the PTS cycles every 13.26 hours.
- Due to the presence of B pictures PTS may not increase monotonically.
- When present, PCR (program clock reference) and OPCR (original PCR) are carried in optional headers.

RTP Fixed Header

```
 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.
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|T=2|P|X|  CC   |M|     PT      |       sequence number         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           timestamp                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           synchronization source identifier (SSRC)            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              content source identifiers(CSRCs)                |
|                             ....                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

M bit: Set to 1 at end of each PES packet

PT: (value TBD) Fixed value for all packet streams of this type.

sequence number: As per RTP spec. Incremented on each packet.

timestamp: 32 LSB's of MPEG Presentation Time Stamp (PTS) (32 bits).

Note - May not be monotonically increasing if pictures present in stream.

SSRC/CSRC: As per RTP spec.

MPEG Transport Header

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|A|B|C|                 PID         |1|2|SC |3|4|5|6| splice counter|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

A: transport_error_indicator (1 bit)

B: payload_unit_start_indicator (normally start of PES packet) (1 bit)

C: transport_priority (1 bit)

PID (Program ID): 13 bits

1: PCR present

2: OPCR present

SC: transport scrambling control (2 bits) (default = 00)

3: discontinuity indicator (1 bit) (default = 0)

4: random access indicator (1 bit) (default = 0)

5: elementary stream priority indicator (1 bit) (default = 0)

6: splice point flag (1 bit) (default = 0)

splice counter: 8 bits (default = 0)

PCR (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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      PCR_base (32 LSB's)                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Unused   |W| PCR_clock_ext       |   Unused                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

W: MSB of program_clock_reference_base (1 bit)

PCR_clock_ext: program_clock_reference_extension (9 bits)
Modulo 300 27MHz clock (default = 0 if not represented in MTS)

PCR_base: 32 LSB's of program_clock_reference_base (32 bits)
90kHz clock

OPCR (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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  OPCR_base (32 LSB's)                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Unused     |W| OPCR_clock_ext  |     Unused                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

W: MSB of program_clock_reference_base (1 bit)

OPCR_clock_ext: program_clock_reference_extension (9 bits)
Modulo 300 27MHz clock (default = 0 if not represented in MTS)

OPCR_base: 32 LSB's of program_clock_reference_base (32 bits)
90kHz clock

Encapsulation of MPEG Elementary Streams

- Following elementary streams (ES) would be encapsulated:
    - i. MPEG1 Video
    - ii. MPEG2 Video
    - iii. MPEG1 Audio
    - iv. MPEG2 Audio

- Only two separate payload types needed as MPEG1 ES can be distinguished from MPEG2 ES from the Video or Audio headers.

RTP Fixed Header for MPEG ES Encapsulation

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|T=2|P|X|   CC    |M|        PT     |   sequencenumber            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          timestamp                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                synchronization source identifier (SSRC)         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         content source identifiers (CSRCs)                      |
|                    ....                                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

M bit: Set to 1 on packet containing MPEG frame end code.

PT: (values TBD) MPEG video or audio stream ID.

sequence number: As per RTP spec. Incremented on each packet.

timestamp: 32 LSB's of MPEG Presentation Time Stamp (PTS) (32 bits).

Note - May not be monotonically increasing if pictures present in stream.

SSRC/CSRC: As per RTP spec.

MPEG Video Specific Headers

Slice/Macro-block fragmentation information:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|S|U|             SC            |A|B|           AP              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

S: Sequence_header_state (1 bit). Initially 0 and toggled at the occurrence of each MPEG sequence header. Used to detect loss of sequence header

U: Unused. Must be set to zero in current profile.

SC: Slice counter (14 bits). Reset to 0 at beginning of each picture and incremented on each slice. Fourteen bits allows pictures with up to 16384 slices (one macro-block each) to be covered by this profile. This should be sufficient to cover the High level of MPEG2 (1920 pixels X 1152 lines), as well as the highest resolution for pictures from the US HDTV system from the "Grand Alliance".

A: Beginning-of-slice (BS) (1 bit).

B: End-of-slice (ES) (1 bit).

AP: Absolute position (expressed as absolute macro-block number) of 1st macro-block in a slice (14 bits). This value shall be constant for each RTP packet for a given slice. Fourteen bits allows slices with up to 16384 macro-blocks to be covered by this profile. This should be sufficient to cover the High level of MPEG2 (1920 pixels X 1152 lines), as well as the highest resolution for pictures from the US HDTV system from the "Grand Alliance". Used to recover synchronization at the macro-block

MPEG Audio Specific Headers

This header shall be attached to each RTP packet at the start of the payload and after any RTP headers for an MPEG1/2 Audio payload type (i.e. PT=TBD).

Audio fragmentation information:

```
 0                   1
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|0|      Frag_offset          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Frag_offset: Byte offset into the frame for the data
in this packet.