Network Working Group
Request for Comments: 3497
Category: Standards Track

L. Gharai
C. Perkins
USC/ISI
G. Goncher
Tektronix
A. Mankin
Bell Labs, Lucent Corporation

March 2003

RTP Payload Format for Society of Motion Picture and Television Engineers (SMPTE) 292M Video

Status of this Memo

Copyright Notice

Abstract

This memo specifies an RTP payload format for encapsulating uncompressed High Definition Television (HDTV) as defined by the Society of Motion Picture and Television Engineers (SMPTE) standard, SMPTE 292M. SMPTE is the main standardizing body in the motion imaging industry and the SMPTE 292M standard defines a bit-serial digital interface for local area HDTV transport.

1. Introduction

The serial digital interface, SMPTE 292M [1], defines a universal medium of interchange for uncompressed High Definition Television (HDTV) between various types of video equipment (cameras, encoders, VTRs, etc.). SMPTE 292M stipulates that the source data be in 10 bit words and the total data rate be either 1.485 Gbps or 1.485/1.001 Gbps.

The use of a dedicated serial interconnect is appropriate in a studio environment, but it is desirable to leverage the widespread availability of high bandwidth IP connectivity to allow efficient wide area delivery of SMPTE 292M content. Accordingly, this memo defines an RTP payload format for SMPTE 292M format video.

It is to be noted that SMPTE 292M streams have a constant high bit rate and are not congestion controlled. Accordingly, use of this payload format should be tightly controlled and limited to private networks or those networks that provide resource reservation and enhanced quality of service. This is discussed further in section 9.

This memo only addresses the transfer of uncompressed HDTV. Compressed HDTV is a subset of MPEG-2 [9], which is fully described in document A/53 [10] of the Advanced Television Standards Committee. The ATSC has also adopted the MPEG-2 transport system (ISO/IEC 13818-1) [11]. Therefore RFC 2250 [12] sufficiently describes transport for compressed HDTV over RTP.

2. Overview of SMPTE 292M

A SMPTE 292M television line comprises two interleaved streams, one containing the luminance (Y) samples, the other chrominance (CrCb) values. Since chrominance is horizontally sub-sampled (4:2:2 coding) the lengths of the two streams match (see Figure 3 of SMPTE 292M [1]). In addition to being the same length the streams also have identical structures: each stream is divided into four parts, (figure 1): (1) start of active video timing reference (SAV); (2) digital active line; (3) end of active video timing reference (EAV); and (4) digital line blanking. A SMPTE 292M line may also carry horizontal ancillary data (H-ANC) or vertical ancillary data (V-ANC) instead of the blanking level; Likewise, ancillary data may be transported instead of a digital active line.

The EAV and SAV are made up of three 10 bit words, with constant values of 0x3FF 0x000 0x000 and an additional word (designated as XYZ in figure 2), carrying a number of flags. This includes an F flag which designates which field (1 or 2) the line is transporting and also a V flag which indicates field blanking. Table 1, further displays the code values in SAV and EAV. After EAV, are two words, LN0 and LN1 (Table 2), that carry the 11 bit line number for the SMPTE 292M line. The Cyclic Redundancy Check, CRC, is also a two word value, shown as CR0 and CR1 in figure 2.

```
+------------+-----------------------+-----+---------------------+
|            | Digital Line Blanking |     | Digital Active Line |
| EAV+LN+CRC | (Blanking level or    | SAV |  (Active Picture or |
|            |  Ancillary Data)      |     |   Ancillary Data)   |
+------------+-----------------------+-----+---------------------+
```

Figure 1. The SMPTE 292M line format.

```
0       20      40      60     80       0      20      40
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+       +-+-+-+-+-+-+-+-+
|3FF| 0 | 0 |XYZ|LN1|LN2|CR0|CR1|       |3FF| 0 | 0 |XYZ|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+       +-+-+-+-+-+-+-+-+
<---- EAV -----> <- LN-> <- CRC->       <----- SAV ----->
```

Figure 2. Timing reference format.

Word	(MSB) 9	8	7	6	5	4	3	2	1	(LSB) 0
3FF	1	1	1	1	1	1	1	1	1	1
000	0	0	0	0	0	0	0	0	0	0
000	0	0	0	0	0	0	0	0	0	0
XYZ	1	F	V	H	P	P	P	P	P	P

NOTES:
F=0 during field 1; F=1 during field 2.
V=0 elsewhere; V=1 during field blanking.
H=0 in SAV; H=1 in EAV.
MSB=most significant bit; LSB=least significant bit.
P= protected bits defined in Table 2 of SMPTE 292M

Table 1: Timing reference codes.

Word	(MSB) 9	8	7	6	5	4	3	2	1	(LSB) 0
LN0	R	L6	L5	L4	L3	L2	L1	L0	R	R
LN1	R	R	R	R	L10	L9	L8	L7	R	R

NOTES:
LN0 - L10 - line number in binary code.
R = reserved, set to "0".

Table 2: Line number data.

The number of words and the format for active lines and line blanking is defined by source format documents. Currently, source video formats transfered by SMPTE 292M include SMPTE 260M, 295M, 274M and 296M [5-8]. In this memo, we specify how to transfer SMPTE 292M over RTP, irrespective of the source format.

3. Conventions Used in this Document

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 BCP 14, RFC 2119 [2].

4. Payload Design

Each SMPTE 292M data line is packetized into one or more RTP packets. This includes all timing signals, blanking levels, active lines and/or ancillary data. Start of active video (SAV) and end of active video (EAV+LN+CRC) signals MUST NOT be fragmented across packets, as the SMPTE 292M decoder uses them to detect the start of scan lines.

The standard RTP header is followed by a 4 octet payload header. All information in the payload header pertains to the first data sample in the packet. The end of a video frame (the packet containing the last sample before the EAV) is marked by the M bit in the RTP header.

The payload header contains a 16 bit extension to the standard 16 bit RTP sequence number, thereby extending the sequence number to 32 bits and enabling RTP to accommodate HDTV's high data rates. At 1.485 Gbps, with packet sizes of at least one thousand octets, 32 bits allows for an approximate 6 hour period before the sequence number wraps around. Given the same assumptions, the standard 16 bit RTP sequence number wraps around in less than a second (336 milliseconds), which is clearly not sufficient for the purpose of detecting loss and out of order packets.

A 148.5 MHz (or 148.5/1.001 MHz) time-stamp is used as the RTP timestamp. This allows the receiver to reconstruct the timing of the SMPTE 292M stream, without knowledge of the exact type of source format (e.g., SMPTE 274M or SMPTE 296M). With this timestamp, the location of the first sample of each packet can be uniquely identified in the SMPTE 292M stream. At 148.5 MHz, the 32 bit timestamp wraps around in 21 seconds.

The payload header also carries the 11 bit line number from the SMPTE 292M timing signals. This provides more information at the application level and adds a level of resiliency, in case the packet containing the EAV is lost.

The bit length of both timing signals, SAV and EAV+LN+CRC, are multiples of 8 bits, 40 bits and 80 bits, respectively, and therefore are naturally octet aligned.

For the video content, it is desirable for the video to both octet align when packetized and also adhere to the principles of application level framing, also known as ALF [13]. For YCrCb video, the ALF principle translates into not fragmenting related luminance and chrominance values across packets. For example, with the 4:2:0 color subsampling, a 4 pixel group is represented by 6 values, Y1 Y2 Y3 Y4 Cr Cb, and video content should be packetized such that these values are not fragmented across 2 packets. However, with 10 bit words, this is a 60 bit value which is not octet aligned. To be both octet aligned, and adhere to ALF, an ALF unit must represent 2 groups of 4 Pixels, thereby becoming octet aligned on a 15 octet boundary. This length is referred to as the pixel group or pgroup, and it is conveyed in the SDP parameters. Table 3 displays the pgroup value for various color samplings. Typical source formats use 4:2:2 sampling, and require a pgroup of 5 octets, other values are included for completeness.

The contents of the Digital Active Line SHOULD NOT be fragmented within a pgroup. A pgroup of 1 indicates that data may be split at any octet boundary (this is applicable to instances where the source format is not known). The SAV and EAV+LN+CRC fields MUST NOT be fragmented.

Color Subsampling	Pixels	10 bit words	aligned on octet#	pgroup
4:2:0	4	6*10	2*60/8 = 15	15
4:2:2	2	4*10	40/8 = 5	5
4:4:4	1	3*10	4*30/8 = 15	15

Table 3. Color subsampling and pgroups.

5. RTP Packetization

The standard RTP header is followed by a 4 octet payload header, and the payload data, as shown in Figure 3.

```
     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
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    | V |P|X|   CC  |M|     PT      |     sequence# (low bits)      |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                          time stamp                           |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                             ssrc                              |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |    sequence# (high bits)      |F|V| Z |        line no        |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                                                               |
    :                          SMPTE 292M data                      :
    :                                                               :
    |                                                               |
    +---------------------------------------------------------------+
```

Figure 3: RTP Packet showing SMPTE 292M headers and payload

5.1. The RTP Header

The following fields of the RTP fixed header are used for SMPTE 292M encapsulation (the other fields in the RTP header are used in their usual manner):

Payload Type (PT): 7 bits
   A dynamically allocated payload type field that designates the payload as SMPTE 292M.

Timestamp: 32 bits
   For a SMPTE 292M transport stream at 1.485 Gbps (or 1.485/1.001 Gbps), the timestamp field contains a 148.5 MHz (or 148.5/1.001 MHz) timestamp, respectively. This allows for a unique timestamp for each 10 bit word.

Marker bit (M): 1 bit
   The Marker bit denotes the end of a video frame, and is set to 1 for the last packet of the video frame and is otherwise set to 0 for all other packets.

Sequence Number (low bits): 16 bits
The low order bits for RTP sequence counter. The standard 16 bit RTP sequence number is augmented with another 16 bits in the payload header in order to accommodate the 1.485 Gbps data rate of SMPTE 292M.

5.2. Payload Header

Sequence Number (high bits): 16 bits
The high order bits for the 32 bit RTP sequence counter, in network byte order.

F: 1 bit
The F bit as defined in the SMPTE 292M timing signals (see Table 1). F=1 identifies field 2 and F=0 identifies field 1.

V: 1 bit
The V bit as defined in the SMPTE 292M timing signals (see Table 1). V=1 during field blanking, and V=0 else where.

Z: 2 bits
SHOULD be set to zero by the sender and MUST be ignored by receivers.

Line No: 11 bits
The line number of the source data format, extracted from the SMPTE 292M stream (see Table 2). The line number MUST correspond to the line number of the first 10 bit word in the packet.

6. RTCP Considerations

RFC 1889 should be used as specified in RFC 1889 [3], which specifies two limits on the RTCP packet rate: RTCP bandwidth should be limited to 5% of the data rate, and the minimum for the average of the randomized intervals between RTCP packets should be 5 seconds. Considering the high data rate of this payload format, the minimum interval is the governing factor in this case.

It should be noted that the sender's octet count in SR packets wraps around in 23 seconds, and that the cumulative number of packets lost wraps around in 93 seconds. This means these two fields cannot accurately represent the octet count and number of packets lost since the beginning of transmission, as defined in RFC 1889. Therefore, for network monitoring purposes or any other application that requires the sender's octet count and the cumulative number of packets lost since the beginning of transmission, the application itself must keep track of the number of rollovers of these fields via a counter.

7. IANA Considerations

This document defines a new RTP payload format and associated MIME type, SMPTE292M. The MIME registration form for SMPTE 292M video is enclosed below:

MIME media type name: video

MIME subtype name: SMPTE292M

Required parameters: rate
The RTP timestamp clock rate. The clock runs at either 148500000 Hz or 148500000/1.001 Hz. If the latter rate is used a timestamp of 148351648 MUST be used, and receivers MUST interpret this as 148500000/1.001 Hz.

Optional parameters: pgroup
The RECOMMENDED grouping for aligning 10 bit words and octets. Defaults to 1 octet, if not present.

Encoding considerations: SMPTE292M video can be transmitted with RTP as specified in RFC 3497.

Security considerations: see RFC 3497 section 9.

Interoperability considerations: NONE

Published specification: SMPTE292M
RFC 3497

Applications which use this media type:
Video communication.

Additional information: None

Magic number(s): None

File extension(s): None

Macintosh File Type Code(s): None

Person & email address to contact for further information:
Ladan Gharai <ladan@isi.edu>
IETF AVT working group.

Intended usage: COMMON

Author/Change controller:
Ladan Gharai <ladan@isi.edu>

8. Mapping to SDP Parameters

Parameters are mapped to SDP [14] as follows:

```
m=video 30000 RTP/AVP 111
a=rtpmap:111 SMPTE292M/148500000
a=fmtp:111  pgroup=5
```

In this example, a dynamic payload type 111 is used for SMPTE292M. The RTP timestamp is 148500000 Hz and the SDP parameter pgroup indicates that for video data after the SAV signal, it must be packetized in multiples of 5 octets.

9. Security Considerations

RTP sessions using the payload format defined in this specification are subject to the security considerations discussed in the RTP specification [3] and any appropriate RTP profile (e.g., [4]).

This payload format does not exhibit any significant non-uniformity in the receiver side computational complexity for packet processing to cause a potential denial-of-service threat for intended receivers.

The bandwidth of this payload format is high enough (1.485 Gbps without the RTP overhead) to cause potential for denial-of-service if transmitted onto most currently available Internet paths. Since congestion control is not possible for SMPTE 292M over RTP flows, use of the payload SHOULD be narrowly limited to suitably connected network endpoints, or to networks where QoS guarantees are available.

If QoS enhanced service is used, RTP receivers SHOULD monitor packet loss to ensure that the service that was requested is actually being delivered. If it is not, then they SHOULD assume that they are receiving best-effort service and behave accordingly.

If best-effort service is being used, RTP receivers MUST monitor packet loss to ensure that the packet loss rate is within acceptable parameters and MUST leave the session if the loss rate is too high. The loss rate is considered acceptable if a TCP flow across the same network path, experiencing the same network conditions, would achieve an average throughput, measured on a reasonable timescale, that is not less than the RTP flow is achieving. Since congestion control is not possible for SMPTE 292M flows, this condition can only be satisfied if receivers leave the session if the loss rate is unacceptably high.

10. Acknowledgments

We would like to thank David Richardson for his insightful comments and contributions to the document. We would also like to thank Chuck Harrison for his input and for explaining the intricacies of SMPTE 292M.

11. Normative References

[1] Society of Motion Picture and Television Engineers, Bit-Serial Digital Interface for High-Definition Television Systems, SMPTE 292M-1998.

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

[3] Schulzrinne, H., Casner, S., Frederick, R. and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", RFC 1889, January 1996.

[4] Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video Conferences with Minimal Control", RFC 1890, January 1996.

12. Informative References

[5] Society of Motion Picture and Television Engineers, Digital Representation and Bit-Parallel Interface - 1125/60 High-Definition Production System, SMPTE 260M-1999.

[6] Society of Motion Picture and Television Engineers, 1920x1080 50Hz, Scanning and Interface, SMPTE 295M-1997.

[7] Society of Motion Picture and Television Engineers, 1920x1080 Scanning and Analog and Parallel Digital Interfaces for Multiple Picture Rates, SMPTE 274M-1998.

[8] Society of Motion Picture and Television Engineers, 1280x720 Scanning, Analog and Digital Representation and Analog Interfaces, SMPTE 296M-1998.

[9] ISO/IEC International Standard 13818-2; "Generic coding of moving pictures and associated audio information: Video", 1996.

[10] ATSC Digital Television Standard Document A/53, September 1995, http://www.atsc.org

[11] ISO/IEC International Standard 13818-1; "Generic coding of moving pictures and associated audio information: Systems",1996.

[12] Hoffman, D., Fernando, G., Goyal, V. and M. Civanlar, "RTP Payload Format for MPEG1/MPEG2 Video", RFC 2250, January 1998.

[13] Clark, D. D., and Tennenhouse, D. L., "Architectural Considerations for a New Generation of Protocols", In Proceedings of SIGCOMM '90 (Philadelphia, PA, Sept. 1990), ACM.

[14] Handley, H. and V. Jacobson, "SDP: Session Description Protocol", RFC 2327, April 1998.

13. Authors' Addresses

Ladan Gharai
USC/ISI
3811 Fairfax Dr.
Arlington VA 22203

EMail: ladan@isi.edu

Colin Perkins
USC/ISI
3811 Fairfax Dr.
Arlington VA 22203

EMail: csp@csperkins.org

Allison Mankin
Bell Labs, Lucent Corporation

EMail: mankin@psg.com

Gary Goncher
Tektronix, Inc.
P.O. Box 500, M/S 50-480
Beaverton, OR 97077

EMail: Gary.Goncher@tek.com

14. Full Copyright Statement

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.