Ladan Gharai
USC/ISI
Colin Perkins
USC/ISI
Gary Goncher
Tektronix
Allison Mankin
USC/ISI
May 30, 2002

RTP Payload Format for SMPTE 292M Video
<draft-ietf-avt-smpte292-video-05.txt>

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

Abstract

This memo specifies a RTP payload format for encapsulating uncompressed High Definition Television (HDTV) as defined by the Society of Motion Picture and Television Engineers 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 10bit words and the total data rate be 1.485 Gbps or 1.485/1.001 Gbps.

A SMPTE 292M television line 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, and likewise, ancillary data may transported instead of a digital active line.

The EAV and SAV are made up of three 10 bit words, with constant values of 0x000 0x000 0x3FF and an additional word (designated as XYZ in figure 2), carrying a number of flags. This includes an F flag which designate 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 LN1 and LN2 (Table 2), which carry the 11bit line number for the SMPTE 292M line, immediately following. The Cyclic Redundancy Check, CRC, is a two word value, shown as CR0 and CR1 in figure 2.

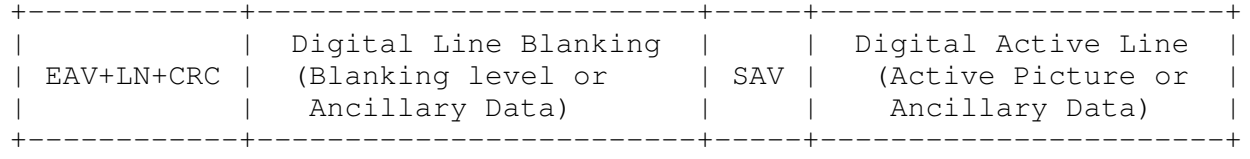


Figure 1. The SMPTE 292M line format.

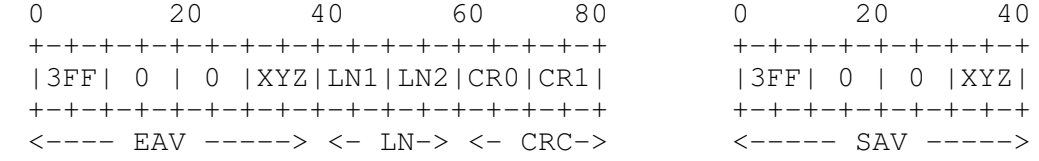


Figure 2. Timing reference format.

Table 1: Timing reference codes.

	(MSB)								(LSB)
Word	9	8	7	6	5	4	3	2	1	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 2: Line number data.

	(MSB)								(LSB)
Word	9	8	7	6	5	4	3	2	1	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".

The number of words and format for active lines and line blanking is defined by source format documents. Currently, source video formats transfered by SMPTE 292M includes SMPTE 260M, 295M, 274M and 296M[2-5]. In this memo we specify how to transfer SMPTE 292M over RTP, irrespective of the source format.

This memo only addresses the transfer of uncompressed HDTV. Compressed HDTV is a subset of MPEG-2 [6], which is fully described in document A/53 [7] of the Advanced Television Standards Committee. The ATSC has also adopted the MPEG-2 transport system (ISO/IEC 13818-1)[8]. Therefore RFC 2250 [9] sufficiently describes transport for compressed HDTV over RTP.

2. 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 RFC 2119[10].

3. 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) signals MUST NOT be fragmented across packets, as the SMPTE 292M decoder uses the sync info in the scan lines 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 is marked by the M bit in the RTP header.

The payload header contains a 16bit 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, 32bits allows for an approximate 6 hour period before the sequence number wraps around.

The payload header also carries the 11bit 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.

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).

The bit length of both timing signals, SAV and EAV+LN+CRC, are multiples of 8bits, 40bits and 80bits, 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 [11]. 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

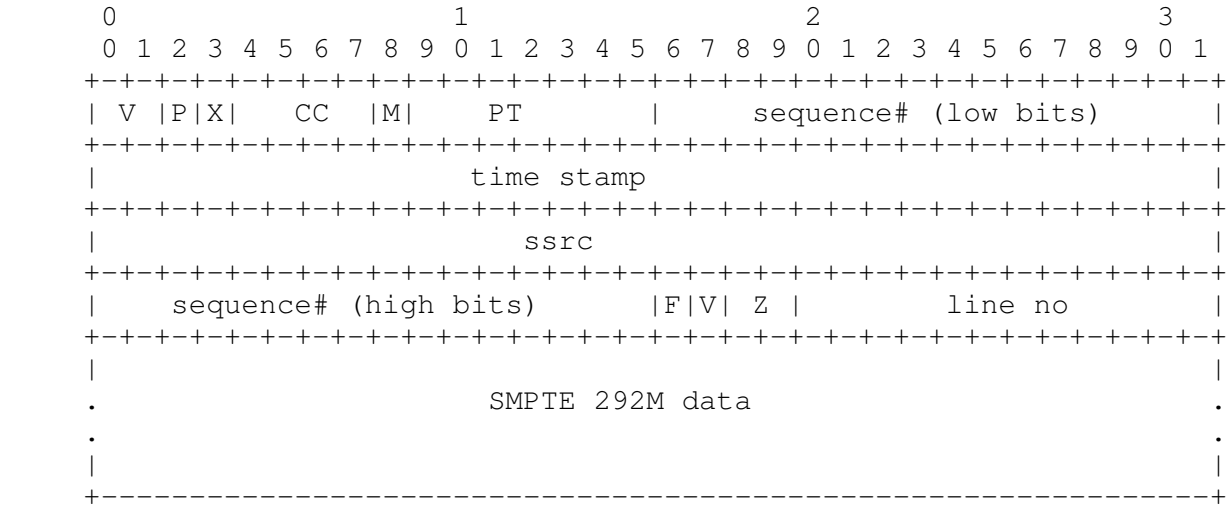
should be packetized such that these values are not fragmented across 2 packets. However, with 10bit words this is a 60bit value which is not octet aligned. To be both octet aligned, and adhere to ALF, an ALF unit must represent 2 groups of 4 Pixels, thereby becoming octet aligned on a 15 octet boundary. This length is referred to as the pixel group or pgroup, and it is conveyed in the SDP parameters. The table below, displays the pgroup value for 4:2:2 and 4:4:4 color samplings.

When packetizing digital active line content, video data MUST NOT be fragmented within a pgroup. This restraint only applies to active video data between the SAV and EAV timing signals. 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.

Color Subsampling	Pixels	10bit 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

4. RTP Packetization

The standard RTP header is followed by a 4 octet payload header, and the payload data.



4.1. The RTP Header

The following fields of the RTP fixed header are used for SMPTE 292M encapsulation:

Payload Type (PT): 7bits
A dynamically allocated payload type field which 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 10bit word.

Marker bit (M): 1bit
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.

4.2. Payload Header

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

F: 1bit
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: 1bit
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: 2bits
SHOULD be set to zero by the sender and MUST be ignored by receivers.

Line No: 11bits
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 10bit word in the packet.

5. RTCP Considerations

RFC1889 recommends transmission of RTCP packets every 5 seconds or at a reduced minimum in seconds of 360 divided by the session bandwidth in kilobits/seconds. At 1.485 Gbps the reduced minimum interval computes to 0.2ms or 4028 packets per second.

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 octet count and number of packets lost since the beginning of transmission, as defined in RFC1889. Therefore for network monitoring purposes other means of keeping track of these variables should be used.

6. 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 148351000
  MUST be used, and receivers MUST interpret this as 148500000/1.001 Hz.

Optional parameters: pgroup
  The RECOMMENDED grouping for aligning 10bit words and octets.

Encoding considerations: SMPTE292M video can be transmitted with
  RTP as specified in "draft-ietf-avt-smpte292-video-05".

Security considerations: see draft "draft-ietf-avt-smpte292-video-05"
  section 8.

Interoperability considerations: NONE

Published specification: SMPTE292M
  draft-ietf-avt-smpte292-video-05

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>

7. Mapping to SDP Parameters

Parameters are mapped to SDP [12] 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, must be packetized in multiples of 5 octets.

8. Security Considerations

RTP packets using the payload format defined in this specification are subject to the security considerations discussed in the RTP specification, and any appropriate RTP profile. This implies that confidentiality of the media streams is achieved by encryption.

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

It is perhaps to be noted that the bandwidth of this payload 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. In the absence from the standards track of a suitable congestion control mechanism for flows of this sort, use of the payload should be narrowly limited to suitably connected network endpoints, or to networks where QoS guarantees are available, and great care taken with the scope of multicast transmissions. This potential threat is common to all high bit rate applications without congestion control.

9. Full Copyright Statement

Copyright (C) The Internet Society (2002). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works.

However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE."

10. Authors' Addresses

Ladan Gharai
ladan@isi.edu
USC/ISI
3811 Fairfax Dr
Arlington, VA 22203-1695

Gary Goncher
ggoncher@tek.com

Colin Perkins
csp@isi.edu
USC/ISI
3811 Fairfax Dr
Arlington, VA 22203-1695

Allison Mankin
mankin@isi.edu
USC/ISI
3811 Fairfax Dr
Arlington, VA 22203-1695

11. Acknowledgment

We would like to thank David Richardson for his insightful comments and contributions to the draft.

12. Bibliography

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

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

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

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

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

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

[7] ATSC Digital Television Standard Document A/53, September 1995,

http://www.atsc.org

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

[9] Hoffman, Fernando, Goyal, Civanlar, "RTP Payload Format for MPEG1/MPEG2 Video", RFC 2250, IETF, January 1998.

[10] S. Bradner, "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119.

[11] 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.

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