Network Working Group
Request for Comments: 4175
Category: Standards Track
L. Gharai
USC/ISI
C. Perkins
University of Glasgow
September 2005

RTP Payload Format for Uncompressed Video

Status of This Memo

Copyright Notice

Abstract

This memo specifies a packetization scheme for encapsulating uncompressed video into a payload format for the Real-time Transport Protocol, RTP. It supports a range of standard- and high-definition video formats, including common television formats such as ITU BT.601, and standards from the Society of Motion Picture and Television Engineers (SMPTE), such as SMPTE 274M and SMPTE 296M. The format is designed to be applicable and extensible to new video formats as they are developed.

1. Introduction

This memo defines a scheme to packetize uncompressed, studio-quality video streams for transport using RTP [RTP]. It supports a range of standard and high-definition video formats, including ITU-R BT.601 [601], SMPTE 274M [274] and SMPTE 296M [296].

Formats for uncompressed standard definition television are defined by ITU Recommendation BT.601 [601] along with bit-serial and parallel interfaces in Recommendation BT.656 [656]. These formats allow both 625-line and 525-line operation, with 720 samples per digital active line, 4:2:2 color sub-sampling, and 8- or 10-bit digital representation.

The representation of uncompressed high-definition television is specified in SMPTE standards 274M [274] and 296M [296]. SMPTE 274M defines a family of scanning systems with an image format of 1920x1080 pixels with progressive and interlaced scanning, while SMPTE 296M defines systems with an image size of 1280x720 pixels and progressive scanning. In progressive scanning, scan lines are displayed in sequence from top to bottom of a full frame. In interlaced scanning, a frame is divided into its odd and even scan lines (called fields) and the two fields are displayed in succession. SMPTE 274M and 296M define images with aspect ratios of 16:9, and define the digital representation for RGB and YCbCr components. In the case of YCbCr components, the Cb and Cr components are horizontally sub-sampled by a factor of two (4:2:2 color encoding).

Although these formats differ in their details, they are structurally very similar. This memo specifies a payload format to encapsulate these and other similar video formats for transport within 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 [2119].

3. Payload Design

Each scan line of digital video is packetized into one or more RTP packets. If the data for a complete scan line exceeds the network MTU, the scan line SHOULD be fragmented into multiple RTP packets, each smaller than the MTU. A single RTP packet MAY contain data for more than one scan line. Only the active samples are included in the RTP payload: inactive samples and the contents of horizontal and vertical blanking SHOULD NOT be transported. In instances where ancillary data is being transmitted, the sender and receiver can disambiguate between ancillary and video data via scan line numbers. That is, the ancillary data will use scan line numbers that are not within the scope of the video frame.

Scan line numbers are included in the RTP payload header, along with a field identifier for interlaced video.

> For SMPTE 296M format video, valid scan line numbers are from 26 through 745, inclusive. For progressive scan SMPTE 274M format video, valid scan lines are from scan line 42 through 1121, inclusive. For interlaced scan SMPTE 274M format video, valid scan line numbers for field one (F=0) are from 21 to 560 and valid scan line numbers for the second field (F=1) are from 584 to 1123. For ITU-R BT.601 format video, the blanking intervals defined in

BT.656 are used: for 625 line video, lines 24 to 310 of field one (F=0) and 337 to 623 of the second field (F=1) are valid; for 525 line video, lines 21 to 263 of the first field, and 284 to 525 of the second field are valid. Other formats (e.g., [372]) may define different ranges of active lines.

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 the payload format to accommodate high data rates without ambiguity. This is necessary as the 16-bit RTP sequence number will roll over very quickly for high data rates. For example, for a 1-Gbps video stream with packet sizes of at least 1000 octets, the standard RTP packet will roll over in 0.5 seconds, which can be a problem for detecting loss and out-of-order packets particularly in instances where the round-trip time is greater than half a second. The extended 32-bit number allows for a longer wrap-around time of approximately nine hours.

Each scan line comprises an integer number of pixels. Each pixel is represented by a number of samples. Samples may be coded as 8-, 10-, 12-, or 16-bit values. A sample may represent a color component or a luminance component of the video. Color samples may be shared between adjacent pixels. The sharing of color samples between adjacent pixels is known as color sub-sampling. This is typically done in the YCbCr color space for the purpose of reducing the size of the image data.

Pixels that share sample values MUST be transported together as a "pixel group". If 10-bit or 12-bit samples are used, each pixel may also comprise a non-integer number of octets. In this case, several pixels MUST be combined into an octet-aligned pixel group for transmission. These restrictions simplify the operation of receivers by ensuring that the complete payload is octet aligned, and that samples relating to a single pixel are not fragmented across multiple packets [ALF].

For example, in YCbCr video with 4:1:1 color sub-sampling, each group of 4 adjacent pixels comprises 6 samples, Y1 Y2 Y3 Y4 Cr Cb, with the Cr and Cb values being shared between all 4 pixels. If samples are 8-bit values, the result is a group of 4 pixels comprising 6 octets. If, however, samples are 10-bit values, the resulting 60-bit group is not octet aligned. To be both octet aligned and appropriately framed, two groups of 4 adjacent pixels must be collected, thereby becoming octet aligned on a 15-octet boundary. This length is referred to as the pixel group size ("pgroup").

Formally, the "pgroup" parameter is the size in octets of the smallest grouping of pixels such that 1) the grouping comprises an integer number of octets; and 2) if color sub-sampling is used, samples are only shared within the grouping. When packetizing digital active line content, video data MUST NOT be fragmented within a pgroup.

Video content is almost always associated with additional information such as audio tracks, time code, etc. In professional digital video applications, this data is commonly embedded in non-active portions of the video stream (horizontal and vertical blanking periods) so that precise and robust synchronization is maintained. This payload format requires that applications using such synchronized ancillary data SHOULD deliver it in separate RTP sessions that operate concurrently with the video session. The normal RTP mechanisms SHOULD be used to synchronize the media.

4. RTP Packetization

The standard RTP header is followed by a 2-octet payload header that extends the RTP Sequence Number, and by a 6-octet payload header for each line (or partial line) of video included. One or more lines, or partial lines, of video data follow. This format makes the payload header 32-bit aligned in the common case, where one scan line (or fragment) of video is included in each RTP packet.

For example, if two lines of video are encapsulated, the payload format will be as shown in Figure 1.

```
 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 Number         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Time Stamp                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             SSRC                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Extended Sequence Number    |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|F|          Line No            |C|           Offset            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Length             |F|          Line No            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|C|           Offset            |                               .
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+                               .
.                                                               .
.                 Two (partial) lines of video data             .
.                                                               .
+---------------------------------------------------------------+
```

Figure 1: RTP Payload Format showing two (partial) lines of video

4.1. The RTP Header

The fields of the fixed RTP header have their usual meaning, with the following additional notes:

Payload Type (PT): 7 bits

A dynamically allocated payload type field that designates the payload as uncompressed video.

Timestamp: 32 bits

For progressive scan video, the timestamp denotes the sampling instant of the frame to which the RTP packet belongs. Packets MUST NOT include data from multiple frames, and all packets belonging to the same frame MUST have the same timestamp.

For interlaced video, the timestamp denotes the sampling instant of the field to which the RTP packet belongs. Packets MUST NOT include data from multiple fields, and all packets belonging to the same field MUST have the same timestamp. Use of field timestamps, rather than a frame timestamp and field indicator bit, is needed to support reverse 3-2 pulldown.

A 90-kHz timestamp SHOULD be used in both cases. If the sampling instant does not correspond to an integer value of the clock (as may be the case when interleaving), the value SHALL be truncated to the next lowest integer, with no ambiguity.

Marker bit (M): 1 bit

If progressive scan video is being transmitted, the marker bit denotes the end of a video frame. If interlaced video is being transmitted, it denotes the end of the field. The marker bit MUST be set to 1 for the last packet of the video frame/field. It MUST be set to 0 for other packets.

Sequence Number: 16 bits

The low-order bits for RTP sequence number. The standard 16-bit sequence number is augmented with another 16 bits in the payload header in order avoid problems due to wrap-around when operating at high rate rates.

4.2. Payload Header

Extended Sequence Number: 16 bits

The high order bits of the extended 32-bit sequence number, in network byte order.

Length: 16 bits

Number of octets of data included from this scan line, in network byte order. This MUST be a multiple of the pgroup value.

Line No.: 15 bits

Scan line number of encapsulated data, in network byte order. Successive RTP packets MAY contains parts of the same scan line (with an incremented RTP sequence number, but the same timestamp), if it is necessary to fragment a line.

Offset: 15 bits

Offset of the first pixel of the payload data within the scan line. If YCbCr format data is being transported, this is the pixel offset of the luminance sample; if RGB format data is being transported, it is the pixel offset of the red sample; if BGR format data is being transported, it is the pixel offset of the blue sample. The value is in network byte order. The offset has a value of zero if the first sample in the payload corresponds to the start of the line, and increments by one for each pixel.

Field Identification (F): 1 bit

Identifies which field the scan line belongs to, for interlaced data. F=0 identifies the first field and F=1 the second field. For progressive scan data (e.g., SMPTE 296M format video), F MUST always be set to zero.

Continuation (C): 1 bit

Determines if an additional scan line header follows the current scan line header in the RTP packet. Set to 1 if an additional header follows, implying that the RTP packet is carrying data for more than one scan line. Set to 0 otherwise. Several scan lines MAY be included in a single packet, up to the path MTU limit. The only way to determine the number of scan lines included per packet is to parse the payload headers.

4.3. Payload Data

Depending on the video format, each RTP packet can include either a single complete scan line, a single fragment of a scan line, or one (or more) complete scan lines and scan line fragments. The length of each scan line or scan line fragment MUST be an integer multiple of the pgroup size in octets. Scan lines SHOULD be fragmented so that the resulting RTP packet is smaller than the path MTU.

It is possible that the scan line length is not evenly divisible by the number of pixels in a pgroup, so the final pixel data of a scan line does not align to either an octet or a pgroup boundary. Nonetheless, the payload MUST contain a whole number of pgroups; the sender MUST fill the remaining bits of the final pgroup with zero and the receiver MUST ignore the fill data. (In effect, the trailing edge of the image is black-filled to a pgroup boundary.)

For RGB format video, samples are packed in order Red-Green-Blue. For BGR format video, samples are packed in order Blue-Green-Red. For both formats, if 8-bit samples are used, the pgroup is 3 octets. If 10-bit samples are used, samples from 4 adjacent pixels form 15-octet pgroups. If 12-bit samples are used, samples from 2 adjacent pixels form 9-octet pgroups. If 16-bit samples are used, each pixel forms a separate 6-octet pgroup.

For RGBA format video, samples are packed in order Red-Green-Blue-Alpha. For BGRA format video, samples are packed in order Blue-Green-Red-Alpha. For 8-, 10-, 12-, or 16-bit samples, each pixel forms its own pgroup, with octet sizes of 4, 5, 6, and 8, respectively.

If the video is in YCbCr format, the packing of samples into the payload depends on the color sub-sampling used.

For YCbCr 4:4:4 format video, samples are packed in order Cb-Y-Cr for both interlaced and progressive frames. If 8-bit samples are used, the pgroup is 3 octets. If 10-bit samples are used, samples from 4 adjacent pixels form 15-octet pgroups. If 12-bit samples are used, samples from 2 adjacent pixels form 9-octet pgroups. If 16-bit samples are used, each pixel forms a separate 6-octet pgroup.

For YCbCr 4:2:2 format video, the Cb and Cr components are horizontally sub-sampled by a factor of two (each Cb and Cr sample corresponds to two Y components). Samples are packed in order Cb0-Y0-Cr0-Y1 for both interlaced and progressive scan lines. For 8-, 10-, 12-, or 16-bit samples, the pgroup is formed from two adjacent pixels (4, 5, 6, or 8 octets, respectively).

For YCbCr 4:1:1 format video, the Cb and Cr components are horizontally sub-sampled by a factor of four (each Cb and Cr sample corresponds to four Y components). Samples are packed in order Cb0-Y0-Y1-Cr0-Y2-Y3 for both interlaced and progressive scan lines. For 8-, 10-, 12-, or 16-bit samples, the pgroup is formed from four adjacent pixels (6, 15, 9, or 12 octets, respectively).

For YCbCr 4:2:0 video, the Cb and Cr components are sub-sampled by a factor of two both horizontally and vertically. Therefore, chrominance samples are shared between certain adjacent lines. Figure 2 shows the composition of luminance and chrominance samples for a 6x6 pixel grid of 4:2:0 YCbCr video. The pixel group is a group of four pixels arranged in a 2x2 matrix. The octet size of the pgroup for progressive scan 4:2:0 video with samples sizes of 8, 10, 12, and 16 bits is 6, 15, 9, and 12 octets, respectively. For interlaced 4:2:0 video, the corresponding pgroups are 4, 5, 6, and 8 octets.

```
line 0:  Y00   Y01   Y02   Y03   Y04   Y05
         Cb00 Cr00   Cb01 Cr01   Cb02 Cr02
line 1:  Y10   Y11   Y12   Y13   Y14   Y15

line 2:  Y20   Y21   Y22   Y23   Y24   Y25
         Cb10 Cr10   Cb11 Cr11   Cb12 Cr12
line 3:  Y30   Y31   Y32   Y33   Y34   Y35

line 4:  Y40   Y41   Y42   Y43   Y44   Y45
         Cb20 Cr20   Cb21 Cr21   Cb22 Cr22
line 5:  Y50   Y51   Y52   Y53   Y54   Y55
```

Figure 2: Chrominance/luminance composition in 4:2:0 YCbCr video

When packetizing progressive scan 4:2:0 YCbCr video, samples from two consecutive scan lines are included in each packet. The scan line number in the payload header is set to that of the first scan line of the pair:

```
line 0/1:
Y00-Y01-Y10-Y11-Cb00-Cr00 Y02-Y03-Y12-Y13-Cb01-Cr01
                                Y04-Y05-Y14-Y15-Cb02-Cr02

line 2/3:
Y20-Y21-Y30-Y31-Cb10-Cr10 Y22-Y23-Y32-Y33-Cb11-Cr11
                                Y24-Y25-Y34-Y35-Cb12-Cr12

line 4/5:
Y40-Y41-Y50-Y51-Cb20-Cr20 Y42-Y43-Y52-Y53-Cb21-Cr21
                                Y44-Y45-Y54-Y55-Cb22-Cr22
```

Figure 3: Packetization of progressive 4:2:0 YCbCr video

For interlaced transport, chrominance samples are transported with every other line. The first set of chrominance samples may be transported with either the first line of field 0, or the first line of field 1. Figure 4 illustrates the transport of chrominance samples starting with the first line of field 0 (signaled by the "top-field-first" MIME parameter).

```
  field 0:
     line 0: Y00-Y01-Cb00-Cr00 Y02-Y03-Cb01-Cr01 Y04-Y05-Cb02-Cr02
     line 2: Y20-Y21 Y22-Y23 Y24-Y25
     line 4: Y40-Y41-Cb20-Cr20 Y42-Y43-Cb21-Cr21 Y44-Y45-Cb22-Cr22

  field 1:
     line 1: Y10-Y11 Y12-Y13 Y14-Y15
     line 3: Y30-Y31-Cb10-Cr10 Y32-Y33-Cb11 Cr11 Y34-Y35-Cb12-Cr12
     line 5: Y50-Y51 Y52-Y53 Y54-Y55
```

Figure 4: Packetization of interlaced 4:2:0 YCbCr video with top-field-first.

Chrominance values may be sampled with different offsets relative to luminance values. For instance, in Figure 2, chrominance values are sampled at the same distance from neighboring luminance samples. It is also possible for a chrominance sample to be co-sited with a luminance sample, as in Figure 5:

```
    line 0:  Y00-C   Y01   Y02-C   Y03   Y04-C   Y05

    line 1:  Y10     Y11   Y12     Y13   Y14     Y15

    line 2:  Y20-C   Y21   Y22-C   Y23   Y24-C   Y25

    line 3:  Y30     Y31   Y32     Y33   Y34     Y35

    line 4:  Y40-C   Y41   Y42-C   Y43   Y44-C   Y45

    line 5:  Y50     Y51   Y52     Y53   Y54     Y55
```

Figure 5: Co-sited video sampling in 4:2:0 YCbCr video where C designates a CbCr pair

In general, chrominance values may be placed between luminance samples or co-sited. Positions can be designated by an integer numbering system starting from left to right and top to bottom. The position matrices shown in Figures 6, 7, and 8 apply for 4:2:0, 4:2:2, and 4:1:1 video, respectively:

```
    line N:     Y[0] [1] Y[2]   Y[0] [1] Y[2]
                 [3] [4] Y[5]    [3] [4]  [5]
    line N+1:   Y[6] [7] Y[8]   Y[6] [7] Y[8]
```

Figure 6: Chrominance position matrix for 4:2:0 YCbCr video

```
line N:    Y[0] [1] Y[2] [3]  Y[0] [1] Y[2] [3]
line N+1:  Y[0] [1] Y[2] [3]  Y[0] [1] Y[2] [3]
```

Figure 7: Chrominance position matrix for 4:2:2 YCbCr video

```
line N:    Y[0] [1] Y[2] [3] Y[4] [5] Y[6]
line N+1:  Y[0] [1] Y[2] [3] Y[4] [5] Y[6]
```

Figure 8: Chrominance position matrix for 4:1:1 YCbCr video

Although these positions do not affect the packetization order of chrominance and luminance samples, the information is needed for interpolation prior to display and therefore should be signaled to the receiver.

5. RTCP Considerations

RTCP SHOULD be used as specified in RFC 3550 [RTP]. It is to be noted that the sender's octet count in SR packets and the cumulative number of packets lost will wrap around quickly for high data rate streams. This means that these two fields may not accurately represent octet count and number of packets lost since the beginning of transmission, as defined in RFC 3550. Therefore, for network monitoring purposes, other means of keeping track of these variables SHOULD be used.

6. IANA Considerations

The IANA has registered one new MIME subtype along with an associated RTP Payload Format, and has created two sub-parameter registries, as described in the following.

6.1. MIME type registration

MIME media type name: video

MIME subtype name: raw

Required parameters:

rate: The RTP timestamp clock rate. Applications using this payload format SHOULD use a value of 90000.

sampling: Determines the color (sub-)sampling mode of the video stream. Currently defined values are RGB, RGBA, BGR, BGRA, YCbCr-4:4:4, YCbCr-4:2:2, YCbCr-4:2:0, and YCbCr-4:1:1. New values may be registered as described in section 6.2 of RFC 4175.

width: Determines the number of pixels per line. This is an integer between 1 and 32767.

height: Determines the number of lines per frame. This is an integer between 1 and 32767.

depth: Determines the number of bits per sample. This is an integer with typical values including 8, 10, 12, and 16.

colorimetry: This parameter defines the set of colorimetric specifications and other transfer characteristics for the video source, by reference to an external specification. Valid values and their specification are:

BT601-5	ITU Recommendation BT.601-5 [601]
BT709-2	ITU Recommendation BT.709-2 [709]
SMPTE240M	SMPTE standard 240M [240]

New values may be registered as described in section 6.2 of RFC 4175.

Optional parameters:

Interlace: If this OPTIONAL parameter is present, it indicates that the video stream is interlaced. If absent, progressive scan is implied.

Top-field-first: If this OPTIONAL parameter is present, it indicates that chrominance samples are packetized starting with the first line of field 0. Its absence implies that chrominance samples are packetized starting with the first line of field 1.

chroma-position: This OPTIONAL parameter defines the position of chrominance samples relative to luminance samples. It is either a single integer or a comma separated pair of integers. Integer values range from 0 to 8, as specified in Figures 6-8 of RFC 4175. A single integer implies that Cb and Cr are co-sited. A comma separated pair of integers designates the locations of Cb and Cr samples, respectively. In its absence, a single value of zero is assumed for color-subsampled video (chroma-position=0).

gamma: An OPTIONAL floating point gamma correction value.

Encoding considerations:

Uncompressed video is only transmitted over RTP as specified in RFC 4175. No file format media type has been defined to go with this transmission media type at this time.

Security considerations: See section 9 of RFC 4175.

Interoperability considerations: NONE.

Published specification: RFC 4175.

Applications which use this media type: Video communication.

Additional information: None

Person & email address to contact for further information:

Ladan Gharai <ladan@isi.edu>
IETF Audio/Video Transport working group.

Intended usage: COMMON

Author: Ladan Gharai <ladan@isi.edu>
Change controller: IETF AVT Working Group
delegated from the IESG

6.2. Parameter Registration

New values of the "sampling" parameter MAY be registered with the IANA provided they reference an RFC or other permanent and readily available specification (the Specification Required policy of RFC 2434 [2434]). A new registration MUST define the packing order of samples and a valid combinations of color and sub-sampling modes.

New values of the "colorimetry" parameter MAY be registered with the IANA provided they reference an RFC or other permanent and readily available specification if colorimetric parameters and other applicable transfer characteristics (the Specification Required policy of RFC 2434 [2434]).

7. Mapping MIME Parameters into SDP

The information carried in the MIME media type specification has a specific mapping to fields in the Session Description Protocol (SDP) [SDP], which is commonly used to describe RTP sessions. When SDP is used to specify sessions transporting uncompressed video, the mapping is as follows:

- The MIME type ("video") goes in SDP "m=" as the media name.

- The MIME subtype (payload format name) goes in SDP "a=rtpmap" as the encoding name.

- Remaining parameters go in the SDP "a=fmtp" attribute by copying them directly from the MIME media type string as a semicolon-separated list of parameter=value pairs.

A sample SDP mapping for uncompressed video is as follows:

```
m=video 30000 RTP/AVP 112
a=rtpmap:112 raw/90000
a=fmtp:112 sampling=YCbCr-4:2:2; width=1280; height=720; depth=10;
                         colorimetry=BT.709-2; chroma-position=1
```

In this example, a dynamic payload type 112 is used for uncompressed video. The RTP sampling clock is 90 kHz. Note that the "a=fmtp:" line has been wrapped to fit this page, and will be a single long line in the SDP file.

8. Security Considerations

RTP packets using the payload format defined in this specification are subject to the security considerations discussed in the RTP specification [RTP] 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 important to note that uncompressed video can have immense bandwidth requirements (up to 270 Mbps for standard-definition video, and approximately 1 Gbps for high-definition video). This is sufficient to cause potential for denial-of-service if transmitted onto most currently available Internet paths.

Accordingly, if best-effort service is being used, users of this payload format MUST monitor packet loss to ensure that the packet loss rate is within acceptable parameters. Packet loss is considered acceptable if a TCP flow across the same network path, and 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. This condition can be satisfied by implementing congestion control mechanisms to adapt the transmission rate (or the number of layers subscribed for a layered multicast session), or by arranging for a receiver to leave the session if the loss rate is unacceptably high.

This payload format may also be used in networks that provide quality-of-service guarantees. If enhanced service is being used, 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.

9. Relation to RFC 2431

In comparison with RFC 2431, this memo specifies support for a wider variety of uncompressed video, in terms of frame size, color sub-sampling and sample sizes. Although [BT656] can transport up to 4096 scan lines and 2048 pixels per line, our payload type can support up to 32768 scan lines and pixels per line. Also, RFC 2431 only address 4:2:2 YCbCr data, while this memo covers YCbCr, RGB, RGBA, BGR, BGRA, and most common color sub-sampling schemes. Given the variety of video types that we cover, this memo also assumes out-of-band signaling for sample size and data types (RFC 2431 uses in band signaling).

10. Relation to RFC 3497

RFC 3497 [292RTP] specifies a RTP payload format for encapsulating SMPTE 292M video. The SMPTE 292M standard defines a bit-serial digital interface for local area High-Definition Television (HDTV) transport. As a transport medium, SMPTE 292M utilizes 10-bit words and a fixed 1.485 Gbps (and 1.485/1.001 Gbps) data rate. SMPTE 292M is typically used in the broadcast industry for the transport of other video formats such as SMPTE 260M, SMPTE 295M, SMPTE 274M, and SMPTE 296M.

RFC 3497 defines a circuit emulation for the transport of SMPTE 292M over RTP. It is very specific to SMPTE 292 and has been designed to be interoperable with existing broadcast equipment with a constant rate of 1.485 Gbps.

This memo defines a flexible native packetization scheme that can packetize any uncompressed video, at varying data rates. In addition, unlike RFC 3497, this memo only transports active video pixels (i.e., horizontal and vertical blanking are not transported).

11. Acknowledgements

The authors are grateful to Philippe Gentric, Chuck Harrison, Stephan Wenger, and Dave Singer for their feedback.

This memo is based upon work supported by the U.S. National Science Foundation (NSF) under Grant No. 0230738. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF.

Normative References

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

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

[2434] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 2434, October 1998.

[601] International Telecommunication Union, "Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios", Recommendation BT.601, October 1995.

[709] International Telecommunication Union, "Parameter Values for HDTV Standards for Production and International Programme Exchange", Recommendation BT.709-2

[240] Society of Motion Picture and Television Engineers, "Television - Signal Parameters - 1125-Line High-Definition Production", SMPTE 240M-1999.

Informative References

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

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

[372] Society of Motion Picture and Television Engineers, "Dual Link 292M Interface for 1920 x 1080 Picture Raster", SMPTE 372M-2002.

[ALF] 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.

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

[BT656] Tynan, D., "RTP Payload Format for BT.656 Video Encoding", RFC 2431, October 1998.

[292RTP] Gharai, L., Perkins, C., Goncher, G., and A. Mankin, "RTP Payload Format for Society of Motion Picture and Television Engineers (SMPTE) 292M Video", RFC 3497, March 2003.

[656] International Telecommunication Union, "Interfaces for Digital Component Video Signals in 525-line and 625-line Television Systems Operating at the 4:2:2 Level of Recommendation ITU-R BT.601 (Part A)", Recommendation BT.656, April 1998.

Authors' Addresses

Ladan Gharai
USC Information Sciences Institute
3811 N. Fairfax Drive, #200
Arlington, VA 22203
USA

EMail: ladan@isi.edu

Colin Perkins
University of Glasgow
Department of Computing Science
17 Lilybank Gardens
Glasgow G12 8QQ
United Kingdom

EMail: csp@csperkins.org

Acknowledgement

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