Network Working Group   P. Luthi
Request for Comments: 3047   PictureTel
Category: Standards Track   January 2001

RTP Payload Format for ITU-T Recommendation G.722.1

Status of this Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Copyright Notice

Abstract

International Telecommunication Union (ITU-T) Recommendation G.722.1 is a wide-band audio codec, which operates at one of two selectable bit rates, 24kbit/s or 32kbit/s. This document describes the payload format for including G.722.1 generated bit streams within an RTP packet. Also included here are the necessary details for the use of G.722.1 with MIME and SDP.

1. 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 [6].

2. Overview of ITU-T Recommendation G.722.1

G.722.1 is a low complexity coder, it compresses 50Hz - 7kHz audio signals into one of two bit rates, 24 kbit/s or 32 kbit/s.

The coder may be used for speech, music and other types of audio.

Some of the applications for which this coder is suitable are:

- Real-time communications such as videoconferencing and telephony.
- Streaming audio
- Archival and messaging

A fixed frame size of 20 ms is used, and for any given bit rate the number of bits in a frame is a constant.

3. RTP payload format for G.722.1

G.722.1 uses 20 ms frames and a sampling rate clock of 16 kHz, so the RTP timestamp MUST be in units of 1/16000 of a second. The RTP payload for G.722.1 has the format shown in Figure 1. No additional header specific to this payload format is required.

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      RTP Header [3]                           |
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|                                                               |
+                 one or more frames of G.722.1                 |
|                             ....                              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: RTP payload for G.722.1

The encoding and decoding algorithm can change the bit rate at any 20ms frame boundary, but no bit rate change notification is provided in-band with the bit stream. Therefore, a separate out-of-band method is REQUIRED to indicate the bit rate (see section 6 for an example of signaling bit rate information using SDP). For the payload format specified here, the bit rate MUST remain constant for a particular payload type value. An application MAY switch bit rates from packet to packet by defining two payload type values and switching between them.

The assignment of an RTP payload type for this new packet format is outside the scope of this document, and will not be specified here. It is expected that the RTP profile for a particular class of applications will assign a payload type for this encoding, or if that is not done then a payload type in the dynamic range shall be chosen.

The number of bits within a frame is fixed, and within this fixed frame G.722.1 uses variable length coding (e.g., Huffman coding) to represent most of the encoded parameters [2]. All variable length codes are transmitted in order from the left most (most significant - MSB) bit to the right most (least significant - LSB) bit, see [2] for more details.

The use of Huffman coding means that it is not possible to identify the various encoded parameters/fields contained within the bit stream without first completely decoding the entire frame.

For the purposes of packetizing the bit stream in RTP, it is only necessary to consider the sequence of bits as output by the G.722.1 encoder, and present the same sequence to the decoder. The payload format described here maintains this sequence.

When operating at 24 kbit/s, 480 bits (60 octets) are produced per frame, and when operating at 32 kbit/s, 640 bits (80 octets) are produced per frame. Thus, both bit rates allow for octet alignment without the need for padding bits.

Figure 2 illustrates how the G.722.1 bit stream MUST be mapped into an octet aligned RTP payload.

An RTP packet SHALL only contain G.722.1 frames of the same bit rate.

```
   first bit                                               last bit
   transmitted                                          transmitted
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                                                               |
   + sequence of bits (480 or 640) generated by the                |
   |             G.722.1 encoder for transmission                  |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

   |            |            |                      |            |
   |            |            |     ...              |            |
   |            |            |                      |            |

   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ ... +-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |MSB...  LSB|MSB...  LSB|                       |MSB...  LSB|
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+ ... +-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     RTP         RTP                                  RTP
     octet 1     octet 2                              octet
                                                      60 or 80
```

Figure 2: The G.722.1 encoder bit stream is split into a sequence of octets (60 or 80 depending on the bit rate), and each octet is in turn mapped into an RTP octet.

The ITU-T standardized bit rates for G.722.1 are 24 kbit/s and 32kbit/s. However, the coding algorithm itself has the capability to run at any user specified bit rate (not just 24 and 32kbit/s) while maintaining an audio bandwidth of 50 Hz to 7 kHz. This rate change is accomplished by a linear scaling of the codec operation, resulting in frames with size in bits equal to 1/50 of the corresponding bit rate.

When operating at non-standard rates the payload format MUST follow the guidelines illustrated in Figure 2. It is RECOMMENDED that values in the range 16000 to 32000 be used, and that any value MUST be a multiple of 400 (this maintains octet alignment and does not then require (undefined) padding bits for each frame if not octet aligned). For example, a bit rate of 16.4 kbit/s will result in a frame of size 328 bits or 41 octets which are mapped into RTP per Figure 2.

3.1 Multiple G.722.1 frames in a RTP packet

More than one G.722.1 frame may be included in a single RTP packet by a sender.

Senders have the following additional restrictions:

- o SHOULD NOT include more G.722.1 frames in a single RTP packet than will fit in the MTU of the RTP transport protocol.

- o All frames contained in a single RTP packet MUST be of the same length, that is they MUST have the same bit rate (octets per frame).

- o Frames MUST NOT be split between RTP packets.

It is RECOMMENDED that the number of frames contained within an RTP packet be consistent with the application. For example, in a telephony application where delay is important, then the fewer frames per packet the lower the delay, whereas for a delay insensitive streaming or messaging application, many frames per packet would be acceptable.

3.2 Computing the number of G.722.1 frames

Information describing the number of frames contained in an RTP packet is not transmitted as part of the RTP payload. The only way to determine the number of G.722.1 frames is to count the total number of octets within the RTP packet, and divide the octet count by the number of expected octets per frame (either 60 or 80 per frame, for 24 kbit/s and 32 kbit/s respectively).

4. MIME registration of G.722.1

MIME media type name: audio

MIME subtype: G7221

Required parameters:

bitrate: the data rate for the audio bit stream. This parameter is necessary because the bit rate is not signaled within the G.722.1 bit stream. At the standard G.722.1 bit rates, the value MUST be either 24000 or 32000. If using the non-standard bit rates, then it is RECOMMENDED that values in the range 16000 to 32000 be used, and that any value MUST be a multiple of 400 (this maintains octet alignment and does not then require (undefined) padding bits for each frame if not octet aligned).

Optional parameters:

ptime: RECOMMENDED duration of each packet in milliseconds.

Encoding considerations:
This type is only defined for transfer via RTP as specified in a Work in Progress.

Security Considerations:
See Section 6 of RFC 3047.

Interoperability considerations: none

Published specification:
See ITU-T Recommendation G.722.1 [2] for encoding algorithm details.

Applications which use this media type:
Audio and video streaming and conferencing tools

Additional information: none

Person & email address to contact for further information:
Patrick Luthi
Luthip@pictel.com

Intended usage: COMMON

Author/Change controller:
Author: Patrick Luthi
Change controller: IETF AVT Working Group

5. SDP usage of G.722.1

When conveying information by SDP [5], the encoding name SHALL be "G7221" (the same as the MIME subtype). An example of the media representation in SDP for describing G.722.1 at 24000 bits/sec might be:

```
m=audio 49000 RTP/AVP 121
a=rtpmap:121 G7221/16000
a=fmtp:121 bitrate=24000
```

where "bitrate" is a variable that may take on values of 24000 or 32000 at the standard rates, or values from 16000 to 32000 (and MUST be an integer multiple of 400) at the non-standard rates.

6. Security Considerations

RTP packets 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. This implies that confidentiality of the media streams is achieved by encryption. Because the data compression used with this payload format is applied end-to-end, encryption may be performed after compression so there is no conflict between the two operations.

A potential denial-of-service threat exists for data encodings using compression techniques that have non-uniform receiver-end computational load. The attacker can inject pathological datagrams into the stream which are complex to decode and cause the receiver to be overloaded. However, this encoding does not exhibit any significant non-uniformity.

As with any IP-based protocol, in some circumstances a receiver may be overloaded simply by the receipt of too many packets, either desired or undesired. Network-layer authentication may be used to discard packets from undesired sources, but the processing cost of the authentication itself may be too high. In a multicast environment, pruning of specific sources may be implemented in future versions of IGMP [7] and in multicast routing protocols to allow a receiver to select which sources are allowed to reach it.

7. References

1. Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC 2026, October 1996.

2. ITU-T Recommendation G.722.1, available online from the ITU bookstore at http://www.itu.int.

3. Schulzrinne, H., Casner, S., Frederick, R. and V. Jacobson, "RTP: A Transport Protocol for real-time applications", RFC 1889, January 1996. (Updated by a Work in Progress.)

4. Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies", RFC 2045, November 1996.

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

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

7. Deering, S., "Host Extensions for IP Multicasting", STD 5, RFC 1112, August 1989.

8. Acknowledgments

The author wishes to thank Tony Crossman for starting this work on G.722.1 packetization and for authoring the initial draft. The author also wishes to thank Steve Casner and Colin Perkins for their valuable feedback and helpful comments.

9. Author's Address

Patrick Luthi
PictureTel Corporation
100 Minuteman Road
Andover, MA 01810
USA

Phone: +1 (978) 292 4354
EMail: luthip@pictel.com

10. Full Copyright Statement

Acknowledgement

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