                Transmission Time Offsets in RTP Streams

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.

Abstract

   This document describes a method to inform Real-time Transport
   Protocol (RTP) clients when RTP packets are transmitted at a time
   other than their 'nominal' transmission time. It also provides a
   mechanism to provide improved inter-arrival jitter reports from the
   clients, that take into account the reported transmission times.

Table of Contents

1. Introduction

   In the Real-time Transport Protocol (RTP) specification [RFC3550],
   network jitter calculations are based on the presumption that packets
   are transmitted essentially in accordance with their RTP timestamps.
   This must be true, of course, on average over longer time intervals,
   as the client is playing the packets out according to those
   timestamps. However, for individual packets, this may not be true
   under some circumstances, such as:

   o When the data rate of the stream is bursty, such as with video
      where I-frames may be significantly larger than P or B frames,
      traffic smoothing may need to be applied to maintain an
      appropriate data rate.

   o In video that has forward-decode dependencies, frames may need to
      be transmitted in decoding order (the sequence number order) but
      with, of course, presentation timestamps. Under these
      circumstances, the transmission time of a frame sent early in
      sequence does not correspond to its RTP timestamp.

   o When retransmissions are sent, the retransmitted packet clearly
      has a different actual transmission time from the original, even
      though they share the same timestamp.

   Under some circumstances, it can help the receiver, or intermediate
   network elements, to know the actual transmission time of the packet.
   This RTP header extension element allows the communication of this
   information.

   The RTP specification does not define a transmission timestamp; nor
   does this specification. This specification merely provides
   information on the relationship between the relative transmission
   times and relative RTP timestamps.

This specification allows the transmitter to indicate to the receiver
any known variation between the spacing of transmission times and the
spacing of RTP timestamps; any unreported variation introduced at or
after the point of measurement of the transmission time will be
treated as network jitter by the receiver. The definition of the
point where the transmission time is measured or defined is left to
the transmitter, though it should, of course, be consistent from
packet to packet.

This information can also be of use to report the inter-arrival
jitter caused by the network, excluding that introduced by the
source. A new RTP Control Protocol (RTCP) packet is defined to
enable this reporting.

2. Requirements Notation

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 [RFC2119].

3. Transmission Offset

Classically, a pair of RTP packets with timestamps S2 and S1 are
transmitted with a time interval between them of (S2 - S1). This
specification permits sending an offset value O in each packet, O1
and O2. One characteristic of these offsets is that the original
transmission interval can be deduced to be (S2 + O2) - (S1 + O1).

More precisely, the offset is defined as follows (with the function
RtoN converting from RTP to Network Time Protocol (NTP) times, and
NtoR doing the reverse):

o   Take an RTP stream that has a recent RTCP sender report relating
    RTP timestamp S0 to NTP timestamp N0;

o   Consider a packet sent after that with RTP timestamp S1.
    Nominally, this is sent at N1 = (N0 + RtoN(S1 - S0));

o   If it was actually sent at a different time, Na, then the offset
    value O1 is O1 = NtoR(Na - N1).

The transmission time is signaled to the receiver in-band using the
general mechanism for RTP header extensions [RFC5285]. The payload
of this extension (the transmitted value) is a 24-bit signed integer.
When added to the RTP timestamp of the packet, it represents the
"effective" RTP transmission time of the packet, on the RTP
timescale. The reported transmission time T1 of a packet with

timestamp S1 and an offset of O1, from the above equations, is T1 =
S1+O1 (though of course the transmission time values only have
meaning when two or more are compared).

The form of the transmission offset extension block is as follows:

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |  ID   | len=2 |              transmission offset               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
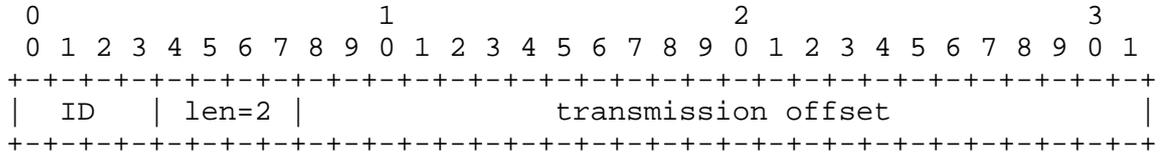
The length field takes the value 2 to indicate that 3 bytes follow.

The sign of the offset value depends greatly on the choice of the
initial mapping of RTP to NTP times. In general, without scanning a
stream entirely it is not possible to ensure that this mapping would
keep all the offsets positive; therefore, this specification allows
negative values.

Imagine a stream with the following timestamps and sizes (in KB):

```
200     2 KB
300     4 KB
400     2 KB
500    12 KB
600    ...effective end of stream
```

This has 20 KB spread over 400 time units, i.e., on average, 1 KB per
20 time units. We traffic-smooth this, and establish that given a
transmission time of x for the first packet, we would transmit the
following packets at the given intervals later:

```
x + 000    2 KB
x + 040    4 KB
x + 120    2 KB
x + 160   12 KB
x + 400 ...effective end of stream
```

The choice of x is essentially arbitrary: only relative values of
timestamps matter. Now, let's say I claim on the first packet that
it went out *at* its RTP timestamp, i.e., with an offset of 0,
meaning that x is 200. Then the offset values are:

```
   0
 -60
 -80
-140
```

This is because in this case, I traffic-smooth by conceptually
sending the small packets 'early'. But since only the relative
values are significant, it is just as valid to say x is 400,
whereupon the offset values are:

200
140
120
 60

In a stream where this extension is not in effect (i.e., not declared
or negotiated), the actual transmission offset is therefore unknown.
However, when the extension is in effect for the stream, it MAY be
omitted in those packets for which the offset is 0 (zero); that is,
packets sent at their nominal time do not need this to be tagged with
this extension. Therefore, the implied transmission time of an un-
tagged RTP packet depends on whether the extension is in effect for
the stream (and therefore the transmission offset is 0) or not
(whereupon the transmission offset is unknown).

The jitter calculations performed by an RTP client MUST NOT use these
transmission offsets. In general, the sender (or intermediate
network elements doing RTP analysis) cannot always know whether the
offsets have been taken into account or not. Therefore, for
consistency, the jitter calculation should continue to operate on the
'raw' reception times. However, see Section 4 on extended jitter
reports, below.

There are no extensionattributes defined for this extension.

It is structurally possible to have more than one extension of the
same type in a packet. However, this extension is only defined for
the source to report. Intermediate network nodes that are not the
source of the RTP session MUST NOT add this extension (whether or not
it was previously present) and MUST NOT alter the existing
transmission offset value in a packet, if the extension is already
present.

(Of course, it is clear that network elements that terminate an RTP
flow, and are the source for a new RTP flow, can add a transmission
offset extension header to the RTP packets of the new flow, if
desired.)

4. Extended Jitter Reports

The inter-arrival jitter computed as defined in Section 6.4.1 of RFC
3550 provides inter-arrival jitter reports that include any source-
introduced jitter (transmission time offsets). If it is desired to

   indicate the actual network jitter, excluding the source-introduced
   jitter, the new RTCP packet type defined here may be used.

   It has the following form:

```
        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
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   hdr |V=2|P|    RC   |    PT=IJ=195   |             length            |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       |                     inter-arrival jitter                      |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
       .                                                               .
       .                                                               .
       .                                                               .
       |                     inter-arrival jitter                      |
       +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
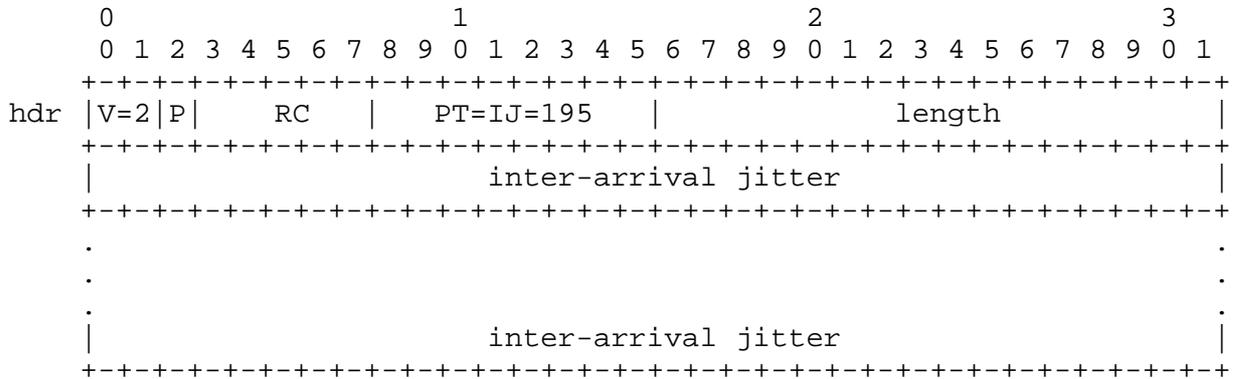
   If present, this RTCP packet must be placed after a receiver report
   (inside a compound RTCP packet), and MUST have the same value for RC
   (reception report count) as the receiver report. The content is
   exactly that number of inter-arrival jitter calculations, calculated
   using the same formula as for sender and receiver reports, but taking
   into account the transmission offsets for the streams (if any). That
   is, the formula uses the values T1=S1+O1, T2, etc., as defined above,
   instead of S1, S2, etc. (If no transmission offset information is
   given for a stream, then the value of inter-arrival jitter in this
   packet and in the receiver report will be identical).

   Precisely, the replacement equation for the equation in the RTP
   specification is as follows, where Rj is the most recent arrival
   time:

$$D(i,j) = (R_j - R_i) - ((S_j + O_j) - (S_i + O_i))$$
$$\qquad = (R_j - (S_j + O_j)) - (R_i - (S_i + O_i))$$

5. Signaling (Setup) Information

   The URI for declaring this header extension in an extmap attribute is
   "urn:ietf:params:rtp-hdrext:toffset". There is no additional setup
   information needed for this extension (no extensionattributes).

6. Security Considerations

   The given transmission offsets are only informative, and it is hard
   to see security considerations from associating them with media
   streams.

The underlying security considerations of [RFC3550] should be taken
into account.

It is possible that malicious senders (or systems tampering with
packets in transit) could send offsets that are implausible, could
confuse the receiver, or result in calculated jitter values that
might mislead the sender. Both the sender and receiver of the
transmission offsets and jitter values should take care that such
behavior does not result in denial of service or other problems.

7. IANA Considerations

The RTCP packet type used for the adjusted inter-arrival jitter has
been registered, in accordance with Section 15 of [RFC3550]. IANA
has added a new value to the RTCP Control Packet types subregistry of
the Real-Time Transport Protocol (RTP) Parameters registry, according
to the following data:

abbrev.	name	value	Reference
IJ	Extended inter-arrival jitter report	195	RFC 5450

Additionally, IANA has registered a new extension URI to the RTP
Compact Header Extensions subregistry of the Real-Time Transport
Protocol (RTP) Parameters registry, according to the following data:

```
   Extension URI: urn:ietf:params:rtp-hdrext:toffset
   Description:   Transmission Time offsets
   Contact:       singer@apple.com
   Reference:     RFC 5450
```

8. Acknowledgments

Ron Frederick, Colin Perkins, and Steve Casner all contributed
substantially to this document, and their help and contributions
helped turn an idea into a specification.

9. Normative References

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", BCP 14, RFC 2119, March 1997.

   [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V.
              Jacobson, "RTP: A Transport Protocol for Real-Time
              Applications", STD 64, RFC 3550, July 2003.

   [RFC5285] Singer, D. and H. Desineni, "A General Mechanism for RTP
              Header Extensions", RFC 5285, July 2008.

Authors' Addresses

   David Singer
   Apple Computer Inc.
   1 Infinite Loop
   Cupertino, CA 95014
   US

   Phone: +1 408 996 1010
   EMail: singer@apple.com

   Harikishan Desineni
   Qualcomm
   5775 Morehouse Drive
   San Diego, CA 92121
   US

   Phone: +1 858 845 8996
   EMail: hd@qualcomm.com