Internet Engineering Task Force (IETF)   Q. Xie
Request for Comments: 6354   August 2011
Updates: 2198, 4102
Category: Standards Track
ISSN: 2070-1721

Forward-Shifted RTP Redundancy Payload Support

Abstract

This document defines a simple enhancement to support RTP sessions with forward-shifted redundant encodings, i.e., redundant data sent before the corresponding primary data. Forward-shifted redundancy can be used to conceal losses of a large number of consecutive media frames (e.g., consecutive loss of seconds or even tens of seconds of media).

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc6354.

Copyright Notice

Table of Contents

1. Introduction 2
   1.1. Sending Redundant Data Inband vs. Out-of-Band 3
2. Conventions 4
3. Allowing Forward-Shifted Redundant Data 4
4. Registration of Media Type "fwdred" 5
5. Mapping Media Type Parameters into SDP 7
6. Usage in Offer/Answer 7
7. IANA Considerations 7
8. Security Considerations 8
9. Normative References 8
Appendix A. Anti-Shadow Loss Concealment Using
            Forward-Shifted Redundancy 9
   A.1. Sender-Side Operations 9
   A.2. Receiver-Side Operations 11
      A.2.1. Normal-Mode Operation 11
      A.2.2. Anti-Shadow-Mode Operation 12

1. Introduction

This document defines a simple enhancement to RFC 2198 [RFC2198] to support RTP sessions with forward-shifted redundant encodings, i.e., redundant data sent before the corresponding primary data.

Forward-shifted redundancy can be used to conceal losses of a large number of consecutive media frames (e.g., consecutive loss of seconds of media). Such capability is highly desirable, especially in wireless mobile communication environments where the radio signal to a mobile wireless media receiver can be temporarily blocked by tall buildings, mountains, tunnels, etc. In other words, the receiver enters into a shadow of the radio coverage. No new data will be received when the receiver is in a shadow.

In some extreme cases, the receiver may have to spend seconds or even tens of seconds in a shadow. The traditional backward-shifted redundant encoding scheme (i.e., redundant data is sent after the primary data), as currently supported by RFC 2198 [RFC2198], does not address such consecutive frame losses.

In contrast, the forward-shifted redundancy scheme allows one to apply effective anti-shadow loss management at the receiver (as illustrated in Appendix A), thus preventing service interruptions when a mobile receiver runs into such a shadow.

Anti-shadow loss concealment as described in this document can be readily applied to the streaming of pre-recorded media. Because of the need of generating the forward-shifted anti-shadow redundant stream, to apply anti-shadow loss concealment to the streaming of live media will require the insertion of a delay equal to or greater than the amount of forward-shifting at the source of media.

1.1. Sending Redundant Data Inband vs. Out-of-Band

Regardless of the direction of time shift (e.g., forward-shifting, or backward-shifting as in RFC 2198) or the encoding scheme (e.g., Forward Error Correction (FEC), or non-FEC), there is always the option of sending the redundant data and the primary data either in the same RTP session (i.e., inband) or in separate RTP sessions (i.e., out-of-band). There are pros and cons for either approach, as outlined below.

Inband Approach:

- o Pro: A single RTP session is faster to set up and easier to manage.
- o Pro: A single RTP session presents a simpler problem for NAT/ firewall traversal.
- o Pro: Less overall overhead -- one source of RTP/UDP/IP overhead.
- o Con: Lack of flexibility -- difficult for middle boxes such as gateways to add/remove the redundant data.
- o Con: Need more specification -- special payload formats need to be defined to carry the redundant data inband.

Out-of-Band Approach:

- o Pro: Flexibility -- redundant data can be more easily added, removed, or replaced by a middle box such as a gateway.
- o Pro: Little or no specification -- no new payload format is needed.
- o Con: Multiple RTP sessions may take longer to set up and may be more complex to manage.
- o Con: Multiple RTP sessions for NAT/firewall traversal are harder to address.
- o Con: Bigger overall overhead -- more than one source of RTP/UDP/IP overhead.

It is noteworthy that the specification of inband payload formats, as described in this document and in RFC 2198, does not preclude a deployment from using the out-of-band approach. Rather, it gives the deployment the choice to use whichever approach is deemed most beneficial under a given circumstance.

2. Conventions

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 [RFC2119].

3. Allowing Forward-Shifted Redundant Data

In RFC 2198, the timestamp offset in the additional header corresponding to a redundant block is defined as a 14-bit unsigned offset of the timestamp relative to the timestamp given in the RTP header. As stated in RFC 2198:

> The use of an unsigned offset implies that redundant data must be sent after the primary data, and is hence a time to be subtracted from the current timestamp to determine the timestamp of the data for which this block is the redundancy.

This effectively prevents RFC 2198 from being used to support forward-shifted redundant blocks.

In order to support the use of forward-shifted redundant blocks, the media type "fwdred", which allows a parameter called "forwardshift", is introduced to indicate the capability of and willingness to use forward-shifted redundancy and the base value of timestamp forward-shifting. The base value of "forwardshift" is an integer equal to or greater than '0' in RTP timestamp units.

In an RTP session that uses forward-shifted redundant encodings, the timestamp of a redundant block in a received RTP packet is determined as follows:

```
timestamp of redundant block = timestamp in RTP header
                    - timestamp offset in additional header
                    + forward-shift base value
```

Note that generally, in a forward-shifted session, the timestamp offset in the additional header is set to '0'.

The sender MUST NOT change the contents of a packet that appears in a forward-shifted stream when it is time to send it in the main stream.

4. Registration of Media Type "fwdred"

The definition is based on media type "red" defined in RFC 2198 [RFC2198] and RFC 4102 [RFC4102], with the addition of the "forwardshift" parameter.

Type names: audio, text

Subtype names: fwdred

Required parameters:

rate: as defined in [RFC4102].

pt: as defined in [RFC4102].

forwardshift: An unsigned integer can be specified as the value.

If this parameter has a value greater than '0', it indicates that the sender of this parameter will use forward-shifting with a base value as specified when sending out redundant data. This value is in RTP timestamp units.

If this parameter has a value of '0', it indicates that the sender of this parameter will not use forward-shifting when sending its redundant data; i.e., the sender will have the same behaviors as defined in RFC 2198.

Optional parameters:

ptime: as defined in [RFC4102] [RFC4566].

maxptime: as defined in [RFC4102] [RFC4867].

Encoding considerations:

This media type is framed binary data (see RFC 4288, Section 4.8) and is only defined for transfer of RTP redundant data frames specified in RFC 2198.

Security considerations: See Section 6 of RFC 2198.

Interoperability considerations: none.

Published specification:

RTP redundant data frame format is specified in RFC 2198.

Applications that use this media type:

It is expected that real-time audio/video, text streaming, and conferencing tools/applications that want protection against losses of a large number of consecutive frames will be interested in using this type.

Additional information: none.

Person & email address to contact for further information:

Qiaobing Xie <Qiaobing.Xie@gmail.com>

Intended usage: COMMON

Restrictions on usage:

This media type depends on RTP framing, and hence is only defined for transfer via RTP (RFC 3550 [RFC3550]). Transfer within other framing protocols is not defined at this time.

Author:

Qiaobing Xie

Change controller:

IETF Audio/Video Transport working group delegated from the IESG.

5. Mapping Media Type Parameters into SDP

The information carried in the media type specification has a specific mapping to fields in the Session Description Protocol (SDP) [RFC4566], which is commonly used to describe RTP sessions. When SDP is used to specify sessions employing the forward-shifted redundant format, the mapping is as follows:

- o The media type ("audio") goes in SDP "m=" as the media name.
- o The media subtype ("fwdred") goes in SDP "a=rtpmap" as the encoding name.
- o The required parameter "forwardshift" goes in the SDP "a=fmtp" attribute by copying it directly from the media type string as "forwardshift=value".

The following is an example of usage that indicates forward-shifted (by 5.1 sec) redundancy:

```
m=audio 12345 RTP/AVP 121 0 5
a=rtpmap:121 fwdred/8000/1
a=fmtp:121 0/5 forwardshift=40800
```

The following is an example of usage that indicates sending redundancy without forward-shifting (equivalent to RFC 2198):

```
m=audio 12345 RTP/AVP 121 0 5
a=rtpmap:121 fwdred/8000/1
a=fmtp:121 0/5 forwardshift=0
```

6. Usage in Offer/Answer

The "forwardshift" SDP parameter specified in this document is declarative, and all reasonable values are expected to be supported.

7. IANA Considerations

IANA made the assignments described below per this document.

- o IANA added the following to the "Audio Media Types" registry:

  fwdred [RFC6354]

- o IANA added the following to the "Text Media Types" registry:

  fwdred [RFC6354]

8. Security Considerations

Security considerations discussed in Section 6 of [RFC2198], Section 4 of [RFC4856], and Sections 9 and 14 of [RFC3550] apply to this specification. In addition, to prevent denial-of-service attacks, a receiver SHOULD be prepared to ignore a 'forwardshift' parameter declaration if it considers the offset value in the declaration excessive. In such a case, the receiver SHOULD also ignore the redundant stream in the resultant RTP session.

9. Normative References

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

[RFC2198] Perkins, C., Kouvelas, I., Hodson, O., Hardman, V., Handley, M., Bolot, J., Vega-Garcia, A., and S. Fosse-Parisis, "RTP Payload for Redundant Audio Data", RFC 2198, September 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.

[RFC4102] Jones, P., "Registration of the text/red MIME Sub-Type", RFC 4102, June 2005.

[RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.

[RFC4856] Casner, S., "Media Type Registration of Payload Formats in the RTP Profile for Audio and Video Conferences", RFC 4856, February 2007.

[RFC4867] Sjoberg, J., Westerlund, M., Lakaniemi, A., and Q. Xie, "RTP Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs", RFC 4867, April 2007.

Appendix A. Anti-Shadow Loss Concealment Using Forward-Shifted Redundancy

(This Appendix is included for Informational purposes.)

It is not unusual in a wireless mobile communication environment that the radio signal to a mobile wireless media receiver can be temporarily blocked by tall buildings, mountains, tunnels, etc. for a period of time. In other words, the receiver enters into a shadow of the radio coverage. When the receiver is in such a shadow, no new data will be received. In some extreme cases, the receiver may have to spend seconds or even tens of seconds in such a shadow.

Without special design considerations to compensate for the loss of data due to shadowing, a mobile user may experience an unacceptable level of service interruptions. Traditional redundant encoding schemes (including RFC 2198 and most FEC schemes) are known to be ineffective in dealing with such losses of consecutive frames.

However, the employment of forward-shifted redundancy, in combination with the anti-shadow loss concealment at the receiver, as described here, can effectively prevent service interruptions due to the effect of shadowing.

A.1. Sender-Side Operations

For anti-shadow loss management, the RTP sender simply adds a forward-shifted redundant stream (called anti-shadow or AS stream) while transmitting the primary media stream. The amount of forward-shifting, which should remain constant for the duration of the session, will determine the maximal length of shadows that can be completely concealed at the receiver, as explained below.

Except for the fact that the redundant stream is forward-shifted relative to the primary stream (i.e., the redundant data is sent ahead of the corresponding primary data), the design decision and trade-offs on the quality, encoding, bandwidth overhead, etc. of the redundant stream are not different from the traditional RTP payload redundant scheme.

The following diagram illustrates a segment of the transmission sequence of a forward-shifted redundant RTP session, in which the AS stream is forward-shifted by 155 frames. If, for simplicity here, we assume that the value of the timestamp is incremented by 1 between two consecutive frames, this forward-shifted redundancy can then be indicated with:

```
forwardshift=155
```

and the setting of the timestamp offset to 0 in all the additional headers. This can mean 3.1 seconds of forward-shifting if each frame represents 20 ms of original media.

```
                        Primary stream    AS stream

            ...               |                |
                              v                v
            Pkt k+8       [ 111 ]          [ 266 ]
                              |                |
                              v                v
            Pkt k+7       [ 110 ]          [ 265 ]
                              |                |
                              v                v
        ^   Pkt k+6       [ 109 ]          [ 264 ]
        |                     |                |
        |                     v                v
            Pkt k+5       [ 108 ]          [ 263 ]
        T                     |                |
        I                     v                v
        M   Pkt k+4       [ 107 ]          [ 262 ]
        E                     |                |
                              v                v
            Pkt k+3       [ 106 ]          [ 261 ]
                              |                |
                              v                v
            Pkt k+2       [ 105 ]          [ 260 ]
                              |                |
                              v                v
            Pkt k+1       [ 104 ]          [ 259 ]
                              |                |
                              v                v
            Pkt k         [ 103 ]          [ 258 ]
                              |                |
                              v                v

                                 (Transmitted first)
```

Figure 1: An Example of Forward-Shifted Redundant RTP Packet Transmission

A.2. Receiver-Side Operations

The anti-shadow receiver is illustrated in the following diagram.

```
                                                 +---------+
                               normal mode   sw1 | media   |     media
    Primary stream ======================o___o==>| decoder |===> output
    AS stream     ----                           +---------+     device
                     |             AS mode o
                     |       +---------+   |
                     |       | anti-   |   |
                     ------->| shadow  |----
                             | buffer  |
                             +---------+
                                  |
                                  V
                             expired frames
                             discarded
```

Figure 2: Anti-Shadow RTP Receiver

The anti-shadow receiver operates between two modes -- "normal mode" and "AS mode". When the receiver is not in a shadow (i.e., when it still receives new data), it operates in the normal mode. Otherwise, it operates in the AS mode.

A.2.1. Normal-Mode Operation

In the normal mode, after receiving a new RTP packet that contains the primary data and forward-shifted AS data, the receiver passes the primary data directly to the appropriate media decoder for play-out (a de-jittering buffer may be used before the play-out, but for simplicity we assume that none is used here), while the received AS data is stored in an anti-shadow buffer.

Moreover, data stored in the anti-shadow buffer will be continuously checked to determine whether it has expired. If any redundant data in the anti-shadow buffer is found to have a timestamp older (i.e., smaller) than that of the last primary frame passed to the media decoder, it will be considered expired and be purged from the anti-shadow buffer.

The following example illustrates the operation of the anti-shadow buffer in normal mode. We use the same assumption as in Figure 1, and assume that the initial timestamp value is 103 when the session starts.

Time (in ms)	Timestamp being played out	Timestamp of media in AS buffer	Note
t < 0		--	(buffer empty)
...			
t=0	103	258	(hold 1 AS frame)
t=20	104	258-259	(hold 2 AS frames)
t=40	105	258-260	(hold 3 AS frames)
...			
t=3080	257	258-412	(full, hold 155 AS frames)
t=3100	258	259-413	(full, frame 258 purged)
t=3120	259	260-414	(full, frame 259 purged)
...			
t=6240	415	416-570	(always holds 3.1 sec worth of redundant data)
...			

Figure 3: Example of Anti-Shadow Buffer Operation in Normal Mode

Before the beginning of the session (t<0), the anti-shadow buffer will be empty. When the first primary frame is received, the play-out will start immediately, and the first received AS frame is stored in the anti-shadow buffer. With the arrival of more forward-shifted redundant frames, the anti-shadow buffer will gradually be filled up.

For the example shown in Figure 3, after 3.08 seconds (the amount of the forward-shifting minus one frame) from the start of the session, the anti-shadow buffer will be full, holding exactly 3.1 seconds worth of redundant data, with the oldest frame corresponding to t=3.1 sec and the youngest frame corresponding to t=6.18 sec.

It is not difficult to see that in normal mode, because of the continuous purge of expired frames and the addition of new frames, the anti-shadow buffer will always be full, holding the next 'forwardshift' amount of redundant frames.

A.2.2. Anti-Shadow-Mode Operation

When the receiver enters a shadow (or any other conditions that prevent the receiver from getting new media data), the receiver switches to the anti-shadow mode, in which it simply continues the play-out from the forward-shifted redundant data stored in the anti-shadow buffer.

For the example in Figure 3, if the receiver enters a shadow at t=3120, it can continue the play-out by using the forward-shifted redundant frames (ts=260-414) from the anti-shadow buffer. As long as the receiver can move out of the shadow by t=6240, there will be no service interruption.

When the shadow condition ends (meaning new data starts to arrive again), the receiver immediately switches back to the normal mode of operation, playing out from newly arrived primary frames. At the same time, the arrival of new AS frames will start to re-fill the anti-shadow buffer.

However, if the duration of the shadow is longer than the amount of forward-shifting, the receiver will run out of media frames from its anti-shadow buffer. At that point, service interruption will occur.

Author's Address

Qiaobing Xie
15901 Wetherburn Rd.
Chesterfield, MO 63017
US

Phone: +1-847-893-0222
EMail: Qiaobing.Xie@gmail.com