Network Working Group G. Hellstrom
Internet-Draft Omnitor
Updates: RFC 4102, RFC 4103 (if approved) February 29, 2020
Intended status: Standards Track
Expires: September 1, 2020

Indicating source of multi-party Real-time text
draft-hellstrom-avtcore-multi-party-rtt-source-01

Abstract

Real-time text mixers need to identify the source of each transmitted text chunk so that it can be presented in suitable grouping with other text from the same source. An enhancement for RFC 4103 real-time text is provided, suitable for a centralized conference model that enables source identification, for use by text mixers and conference-enabled participants. The mechanism builds on use of the CSRC list in the RTP packet. A capability exchange is specified so that it can be verified that a participant can handle the multi-party coded real-time text stream. The capability is indicated by an sdp media attribute "rtt-mix".

Table of Contents

1. Introduction

RFC 4103[RFC4103] specifies use of RFC 3550 RTP[RFC3550] for transmission of real-time text (RTT) and the "text/t140" format. It also specifies a redundancy format "text/red" for increased robustness. RFC 4102 [RFC4102] registers the "text/red" format. The redundancy scheme enables efficient transmission of redundant text in packets together with new text. However the redundant header format has no source indicators for the redundant transmissions. An assumption has had to be made that the redundant parts in a packet are from the same source as the new text. The recommended transmission is one new and two redundant generations of text (T140blocks) in each packet and the recommended transmission interval

is 300 ms. A mixer, selecting between text input from different sources and transmitting it in a common stream need to make sure that the receiver can assign the received text to the proper sources for presentation. Therefore, without any extra rule for source identification, the mixer needs to stop sending new text from that source and then make sure that all text so far has been sent with all intended redundancy levels (usually two) in order to switch source. That causes the very long time of one second to switch between transmission of text from one source to text from another source. Both the total throughput and the switching performance in the mixer is too low for most applications.

A more efficient source identification scheme requires that each redundant T140block has its source individually preserved. The present specification introduces a source indicator by specific rules for populating the CSRC-list in the RTP-packet.

2. Nomenclature

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].

The terms SDES, CNAME, NAME, SSRC, CSRC, CSRC list, CC are explained in [RFC3550]

The term "T140block" is defined in RFC 4103 [RFC4103] to contain one or more T.140 code elements.

3. Intended application

The scheme for identification of source of redundant transmissions is intended for transmission from entities taking the mixer role in centralised mixing configurations for RTT. It is intended for reception by both participants and mixers.

4. Use of fields in the RTP packets

RFC 4103[RFC4103] specifies use of RFC 3550 RTP[RFC3550], and a redundancy scheme "text/red" for increased robustness. This specification updates RFC 4102[RFC4102] and RFC 4103[RFC4103] by introducing a rule for populating and using the CSRC-list in the RTP packet to enhance the performance in multi-party RTT sessions.

The first member in the CSRC-list shall contain the SSRC of the source of the primary T140block in the packet. The second and further members in the CSRC-list shall contain the SSRC of the source of the first, second, etc redundant generation T140blocks included in

the packet. (the recommended level of redundancy is to use one primary and two redundant generations of T140blocks.) In some cases, the primary or redundant T140block is empty, but is still represented by a member in the redundancy header. For such cases, the corresponding CSRC-list member MUST also be included.

The CC field shall show the number of members in the CSRC list.

```
 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=2|P|X| CC=3  |M|  "RED" PT   |   sequence number of primary  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               timestamp of primary encoding "P"               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           synchronization source (SSRC) identifier            |
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|  CSRC list member 1 = SSRC of source of "P"                   |
|  CSRC list member 2 = SSRC of source of "R1"                  |
|  CSRC list member 3 = SSRC of source of "R2"                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|1|   T140 PT   |  timestamp offset of "R2" | "R2" block length |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|1|   T140 PT   |  timestamp offset of "R1" | "R1" block length |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|   T140 PT   | "R2" T.140 encoded redundant data             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+---------------+
|   |  "R1" T.140 encoded redundant data        |               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+         +-+-+-+
|              "P" T.140 encoded primary data           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
Figure 1: text/red packet with sources indicated in the CSRC-list.
```

5. Actions at transmission by a mixer

A text/red transmitter is usually sending packets at a regular transmission interval as long as there is something (new or redundant T140blocks) to transmit. 300 ms is the default transmission interval, but shorter intervals may be considered for specific cases. The transmitter has its own SSRC, and its own RTP sequence number series. At time of transmission, the RTP packet SHALL be populated with next T140block queued for transmission from any of the active sources. This T140block shall be placed in the primary area of the packet. When performing as a mixer, the SSRC of its source shall be placed as the first member in the CSRC-list. The current time is inserted and the timestamp.

If no unsent T140blocks were available, at this time, but T140blocks available which have not been yet been sent the full intended number of redundant transmissions, then The primary T140block is created by an empty T140block, and populated in a packet for transmission. The SSRC of the transmitter is included in the first place of the CSRC-list.

The primary T140block, in the latest transmission is used to populate the first redundant T140block, and its source is placed as the second member of the CSRC-list. The first redundant T140block from the latest transmission is now placed as the second redundant T140block, and the corresponding CSRC placed in its place in the CSRC-list. Usually this is the level of redundancy used. If a higher number of redundancy is used, then the procedure is maintained until all available redundant levels of T140blocks and their sources are placed in the packet. If a receiver has negotiated a lower number of text/red generations, then that level shall be the maximum used by the transmitter.

The timer offset values are inserted in the redundancy header, with the time offset from when the corresponding T140block was sent as original. (usually one or two times the transmission interval).

The number of members in the CSRC list shall be placed in the "CC" header field.

When there is no new T140block to transmit, and no redundant T140block that has not been retransmitted the intended number of times, the transmission process can stop until either new T140blocks arrive, or if a keep-alive method calls for transmission of keep-alive packets.

6. Actions at reception

The enhanced "text/red" receiver shall receive RTP packets in the single stream from the transmitter, and distribute the T140blocks for presentation in presentation areas for each source. Other tasks for receivers, such as gateways or chained mixers are also feasible.

The CC field of the received packets indicate the used level of redundancy for the current packet.

The RTP sequence numbers of the received packets are monitored for gaps or packets out of order.

As long as the sequence is correct, each packet is unpacked in order. The T140blocks are extracted from the primary area, and the corresponding SSRC is extracted from the first position in the CSRC

list and used for assigning the new T140block to the correct input area (or correspondingly).

If a sequence number gap appears and is still there after some short defined time for jitter resolution, T140data needs to be recovered from redundant data. If the gap is wider than the number of generations of redundant T140blocks in the packet, then a t140block is created with a marker for text loss [T140ad1] and assigned to the SSRC of the transmitter as a general input from the mixer.

Then, the T140blocks in the received packet are retrieved beginning with the highest redundant generation, grouping them with the corresponding SSRC from the CSRC-list and assigning them to further processing per source. Finally the primary T140block is retrieved from the packet and similarly its source retrieved from the first position in the CSRC-list, and then assigned to the corresponding input handling for that source.

If the sequence number gap was equal or less than the number of redundancy generations in the received packet, no missing text marker shall be inserted, and instead the T140blocks and their SSRCs recovered from the redundancy information and the CSRC-list in the way indicated above.

Unicode character BOM is sometimes used as a filler or keep alive by transmission implementations. These should be deleted on reception.

Note that empty T140blocks are sometimes included in the packets. They just do not provide any contents.

7. RTCP considerations

A mixer should send RTCP reports with SDES, CNAME and NAME information about the sources in the conference. This makes it possible for participants to compose a suitable label for text from each source.

8. Chained operation

By strictly applying the rules for CSRC-list population by all conforming devices, mixers can be arranged in chains.

9. Use with SIP centralized conferencing framework

The SIP conferencing framework, mainly specified in RFC 4353[RFC4353], RFC 4579[RFC4579] and RFC 4575[RFC4575] is suitable for coordinating sessions including multi-party RTT. The RTT stream is one and the same during the conference. Participants get

announced by notifications when participants are joining or leaving, and further user information may be provided. The SSRC of the text to expect from joined users can be included in a notification. This can be used both for security purposes and for preparation of an SSRC to label translation for presentation to other users.

10. Usage without redundancy

The CSRC list member should be used as source indicator also for cases when the text/t140 format is used. That may be the case when robustness in transmission is provided by some other means than by redundancy and the text/red format. All aspects of this memo applies except the redundant generations in transmission.

For the use case without redundancy, the CC field in the RTP packet shall have the value 1, and the CSRC list contain one member.

11. SDP Capability negotiation

There are RTT implementations which implement RFC 4103 [RFC4103] but not the present specification. Sending mixed text according to the present specification to a device implementing only RFC 4103 [RFC4103] would lead to unreadable output. Therefore, in order to negotiate RTT Mixing capability according to the present specification, all devices supporting the present specification shall include an sdp media attribute "rtt-mix" indicating this capability in offers and answers. Multi-party streams using the coding of this specification must not be sent to devices who have not indicated the "rtt-mix" capability.

Implementations not understanding this parameter MUST ignore it according to common SDP rules.

An sdp media attribute is defined here, named "rtt-mix", without any parameter. It is intended to be used in "text" media descriptions with "text/red" and "text/t140" formats. It indicates capability to use source indications in the CSRC list according to this specification. Syntax:

```
a=rtt-mix
```

The attribute is used in offer/answer procedures in a declarative way. Both parties express their capability to use sources in the CSRC list as specified in this specification.

A party who has expressed the "rtt-mix" capability MUST populate the CSRC-list according to this specification if it acts as an rtp-mixer and sends to a party who has expressed the "rtt-mix" capability.

A party who has expressed the "rtt-mix" capability MUST interpret the contents of the CSRC-list according to this specification in received rtp packets from parties who have expressed "rtt-mix" capability .

A party MUST NOT transmit packets with redundancy format according to this specification to a party who has not expressed "rtt-mix" capability.

12. Examples

This example shows a symbolic flow of packets from a mixer with loss and recovery. A, B and C are sources of RTT. M is the mixer. P indicates primary data. R1 is first redundant generation data and R2 is second redundant generation data. A1, B1, A2 etc are text chunks (T140blocks) received from the respective sources. X indicates dropped packet between the mixer and a receiver.

```
----------------
Seq no 1
CSRC list A,M,B
R2 B99
R1: Empty
P: A1
----------------
```

Assuming that earlier packets were received in sequence, text A1 is received and assigned to reception area A.

```
----------------
Seq no 2
CC=3
CSRC list C,A,M
R2 Empty
R1: A1
P: C1
----------------
```

Text C1 is received and assigned to reception area C.

```
----------------
X Seq no 3
X CC=3
X CSRC list A,C,A
X R2: A1
X R1: C1
X P: A2
----------------
```

Assumed to be dropped in network problems

```
----------------
X Seq no 4
X CC=3
X CSRC list B,A,C
X R2: C1
X R1: A2
X P: B1
----------------
Assumed to be dropped in network problems

----------------
X Seq no 5
X CC=3
X CSRC list A,B,A
X R2: A2
X R1: B1
X P: A3
----------------
Assumed to be dropped in network problems

----------------
Seq no 6
CC=3
CSRC list C,A,B
R2: B1
R1: A3
P: C2
----------------
The latest received sequence number before 6 was 2.
Recovery is therefore tried for 3,4,5.
But there is no coverage for seq no 3. A missing text mark (U'FFFD)
is created and assigned to the mixer reception area.
For seqno 4, text B1 is recovered and assigned to reception area B.
For seqno 5, text A3 is recovered and assigned to reception area A.
Primary text C2 is received and assigned to reception area C.
```

With only one or two packets lost, there would not be any need to create a missing text marker, and all text would be recovered.

It will be a design decision how to present the missing text markers assigned to the mixer as a source.

13. Performance considerations

This specification allows new text from one source per packet. Packets are transmitted with timed intervals. The default transmission interval is 300 ms for RFC 4103[RFC4103], and is suitable for transmission from single sources. However when more

sources contribute to the flow, a shorter transmission interval may be applicable. The transmission interval is therefore recommended to be 100 ms for mixers. This interval provides for smooth flow of text from 5 sources simultaneously.

14. Presentation level considerations

ITU-T T.140 [T140] provides the presentation level requirements for the RFC 4103 [RFC4103] transport. T.140 [T140] has functions for erasure and other formatting and has this general statement for the presentation:

"The display of text from the members of the conversation should be arranged so that the text from each participant is clearly readable, and its source and the relative timing of entered text is visualized in the display. Mechanisms for looking back in the contents from the current session should be provided. The text should be displayed as soon as it is received."

There is no strict "message" concept in real-time text. Line separator is used as a separator allowing a part of received text to be grouped in presentation. The receiving party may separate presentation of parts of text from a source in readable groups based on other criteria (as comma, full stop, or other phrase delimiters, or a long pause) when it benefits the user to most easily find new text or correlated text from different parties.

Further presentation level considerations are out of scope for this document.

15. Congestion considerations

The congestion considerations and recommended actions from RFC 4103 [RFC4103] are valid also in multi-party situations.

16. Acknowledgements

17. IANA Considerations

[RFC EDITOR NOTE: Please replace all instances of RFCXXXX with the RFC number of this document.]

IANA is asked to register the new sdp attribute "rtt-mix".

Contact name:	IESG
Contact email:	iesg@ietf.org
Attribute name:	rtt-mix
Attribute syntax	a=rtt-mix
Attribute semantics	See RFCXXXX section <xref target="nego"/>
Attribute value	-
Usage level:	media
Purpose:	Indicate support for the rtp-mixer format for real-time text transmission
O/A procedure	Declarative
Mux Category	normal
Reference:	RFCXXXX

18. Security Considerations

The RTP-mixer model requires the mixer to be allowed to decrypt, pack and encrypt secured text from the conference participants. Therefore the mixer needs to be trusted. This is similar to the situation for central mixers of audio and video.

The requirement to transfer information about the user in RTCP reports in SDES, CNAME and NAME fields for creation of labels may have privacy concerns as already stated in RFC 3550 [RFC3550], and may be restricted of privacy reasons. The receiving user will then get a more symbolic label for the source.

19. Change history

19.1. Changes from version -00 to -01

Editorial cleanup.

Changed capability indication from fmtp-parameter to sdp attribute "rtt-mix".

Swapped order of redundancy elements in the example to match reality.

Increased the SDP negotiation section

20. References

20.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, DOI 10.17487/RFC3550, July 2003, <https://www.rfc-editor.org/info/rfc3550>.

[RFC4102] Jones, P., "Registration of the text/red MIME Sub-Type", RFC 4102, DOI 10.17487/RFC4102, June 2005, <https://www.rfc-editor.org/info/rfc4102>.

[RFC4103] Hellstrom, G. and P. Jones, "RTP Payload for Text Conversation", RFC 4103, DOI 10.17487/RFC4103, June 2005, <https://www.rfc-editor.org/info/rfc4103>.

[T140] ITU-T, "Recommendation ITU-T T.140 (02/1998), Protocol for multimedia application text conversation", February 1998.

[T140ad1] ITU-T, "Recommendation ITU-T.140 Addendum 1 - (02/2000), Protocol for multimedia application text conversation", February 2000.

20.2. Informative References

[RFC4353] Rosenberg, J., "A Framework for Conferencing with the Session Initiation Protocol (SIP)", RFC 4353, DOI 10.17487/RFC4353, February 2006, <https://www.rfc-editor.org/info/rfc4353>.

[RFC4575] Rosenberg, J., Schulzrinne, H., and O. Levin, Ed., "A Session Initiation Protocol (SIP) Event Package for Conference State", RFC 4575, DOI 10.17487/RFC4575, August 2006, <https://www.rfc-editor.org/info/rfc4575>.

[RFC4579] Johnston, A. and O. Levin, "Session Initiation Protocol (SIP) Call Control - Conferencing for User Agents", BCP 119, RFC 4579, DOI 10.17487/RFC4579, August 2006, <https://www.rfc-editor.org/info/rfc4579>.

Author's Address

Gunnar Hellstrom
Omnitor
Esplanaden 30
Vendelso 13670
Sweden

Phone: +46708204288
Email: gunnar.hellstrom@omnitor.se