Network Working Group
Request for Comments: 4103
Obsoletes: 2793
Category: Standards Track

G. Hellstrom
Omnitor AB
P. Jones
Cisco Systems, Inc.

June 2005

RTP Payload for Text Conversation

Status of This Memo

Copyright Notice

Abstract

This memo obsoletes RFC 2793; it describes how to carry real-time text conversation session contents in RTP packets. Text conversation session contents are specified in ITU-T Recommendation T.140.

One payload format is described for transmitting text on a separate RTP session dedicated for the transmission of text.

This RTP payload description recommends a method to include redundant text from already transmitted packets in order to reduce the risk of text loss caused by packet loss.

Table of Contents

1. Introduction

This document defines a payload type for carrying text conversation session contents in RTP [2] packets. Text conversation session contents are specified in ITU-T Recommendation T.140 [1]. Text conversation is used alone or in connection with other conversational facilities, such as video and voice, to form multimedia conversation services. Text in multimedia conversation sessions is sent character-by-character as soon as it is available, or with a small delay for buffering.

The text is intended to be entered by human users from a keyboard, handwriting recognition, voice recognition or any other input method. The rate of character entry is usually at a level of a few characters per second or less. In general, only one or a few new characters are expected to be transmitted with each packet. Small blocks of text may be prepared by the user and pasted into the user interface for transmission during the conversation, occasionally causing packets to carry more payload.

T.140 specifies that text and other T.140 elements must be transmitted in ISO 10646-1 [5] code with UTF-8 [6] transformation. This makes it easy to implement internationally useful applications and to handle the text in modern information technology environments. The payload of an RTP packet that follows this specification consists of text encoded according to T.140, without any additional framing. A common case will be a single ISO 10646 character, UTF-8 encoded.

T.140 requires the transport channel to provide characters without duplication and in original order. Text conversation users expect that text will be delivered with no, or a low level, of lost information.

Therefore, a mechanism based on RTP is specified here. It gives text arrival in correct order, without duplication, and with detection and indication of loss. It also includes an optional possibility to repeat data for redundancy in order to lower the risk of loss. Because packet overhead is usually much larger than the T.140 contents, the increase in bandwidth, with the use of redundancy, is minimal.

By using RTP for text transmission in a multimedia conversation application, uniform handling of text and other media can be achieved in, for example, conferencing systems, firewalls, and network translation devices. This, in turn, eases the design and increases the possibility for prompt and proper media delivery.

This document obsoletes RFC 2793 [16]. The text clarifies ambiguities in RFC 2793, improves on the specific implementation requirements learned through development experience and gives explicit usage examples.

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 [4].

3. Usage of RTP

The payload format for real-time text transmission with RTP [2] described in this memo is intended for general text conversation use and is called text/t140 after its MIME registration.

3.1. Motivations and Rationale

The text/t140 format is intended to be used for text transmitted on a separate RTP session, dedicated for the transmission of text, and not shared with other media.

The text/t140 format MAY be used for any non-gateway application, as well as in gateways. It MAY be used simultaneously with other media streams, transmitted as a separate RTP session, as required in real time multimedia applications.

The text/t140 format specified in this memo is compatible with its earlier definition in RFC 2793. It has been refined, with the main intention to minimize interoperability problems and encourage good reliability and functionality.

By specifying text transmission as a text medium, many good effects are gained. Routing, device selection, invocation of transcoding, selection of quality of service parameters, and other high and low level functions depend on each medium being explicitly specified.

3.2. Payload Format for Transmission of text/t140 Data

A text/t140 conversation RTP payload format consists of one, and only one, block of T.140 data, referred to as a "T140block" (see Section 3.3). There are no additional headers specific to this payload format. The fields in the RTP header are set as defined in Section 3.5, carried in network byte order (see RFC 791 [12]).

3.3. The "T140block"

T.140 text is UTF-8 coded, as specified in T.140, with no extra framing. The T140block contains one or more T.140 code elements as specified in [1]. Most T.140 code elements are single ISO 10646 [5] characters, but some are multiple character sequences. Each character is UTF-8 encoded [6] into one or more octets. Each block MUST contain an integral number of UTF-8 encoded characters regardless of the number of octets per character. Any composite character sequence (CCS) SHOULD be placed within one block.

3.4. Synchronization of Text with Other Media

Usually, each medium in a session utilizes a separate RTP stream. As such, if synchronization of the text and other media packets is important, the streams MUST be associated when the sessions are established and the streams MUST share the same reference clock (refer to the description of the timestamp field as it relates to synchronization in Section 5.1 of RFC 3550 [2]). Association of RTP streams can be done through the CNAME field of RTCP SDES function. It is dependent on the particular application and is outside the scope of this document.

3.5. RTP Packet Header

Each RTP packet starts with a fixed RTP header. The following fields of the RTP fixed header are specified for T.140 text streams:

Payload Type (PT): The assignment of an RTP payload type is specific to the RTP profile under which the payload format is used. For profiles that use dynamic payload type number assignment, this payload format can be identified by the MIME type "text/t140" (see Section 10). If redundancy is used per RFC 2198, another payload type number needs to be provided for the redundancy format. The MIME type for identifying RFC 2198 is available in RFC 4102 [9].

Sequence number: The definition of sequence numbers is available in RFC 3550 [2]. When transmitting text using the payload format for text/t140, it is used for detection of packet loss and out-of-order packets, and can be used in the process of retrieval of redundant text, reordering of text and marking missing text.

Timestamp: The RTP Timestamp encodes the approximate instance of entry of the primary text in the packet. A clock frequency of 1000 Hz MUST be used. Sequential packets MUST NOT use the same timestamp. Because packets do not represent any constant duration, the timestamp cannot be used to directly infer packet loss.

M-bit: The M-bit MUST be included. The first packet in a session, and the first packet after an idle period, SHOULD be distinguished by setting the marker bit in the RTP data header to one. The marker bit in all other packets MUST be set to zero. The reception of the marker bit MAY be used for refined methods for detection of loss.

4. Protection against Loss of Data

Consideration must be devoted to keeping loss of text due to packet loss within acceptable limits. (See ITU-T F.703 [17])

The default method that MUST be used, when no other method is explicitly selected, is redundancy in accordance with RFC 2198 [3]. When this method is used, the original text and two redundant generations SHOULD be transmitted if the application or end-to-end conditions do not call for other levels of redundancy to be used.

Forward Error Correction mechanisms, as per RFC 2733 [8], or any other mechanism with the purpose of increasing the reliability of text transmission, MAY be used as an alternative or complement to redundancy. Text data MAY be sent without additional protection if end-to-end network conditions allow the text quality requirements, specified in ITU-T F.703 [17], to be met in all anticipated load conditions.

4.1. Payload Format When Using Redundancy

When using the payload format with redundant data, the transmitter may select a number of T140block generations to retransmit in each packet. A higher number introduces better protection against loss of text but marginally increases the data rate.

The RTP header is followed by one or more redundant data block headers: one for each redundant data block to be included. Each of these headers provides the timestamp offset and length of the corresponding data block, in addition to a payload type number (indicating the payload format text/t140).

The redundant data block headers are followed by the redundant data fields carrying T140blocks from previous packets. Finally, the new (primary) T140block for this packet follows.

Redundant data that would need a timestamp offset higher than 16383 (due to its age at transmission) MUST NOT be included in transmitted packets.

4.2. Using Redundancy with the text/t140 Format

Because text is transmitted only when there is text to transmit, the timestamp is not used to identify a lost packet. Rather, missing sequence numbers are used to detect lost text packets at reception. Also, because sequence numbers are not provided in the redundant header, some additional rules must be followed to allow redundant data that corresponds to missing primary data to be properly merged into the stream of primary data T140blocks. They are:

- Each redundant data block MUST contain the same data as a T140block previously transmitted as primary data.

- The redundant data MUST be placed in age order, with the most recent redundant T140block last in the redundancy area.

- All T140blocks, from the oldest desired generation up through the generation immediately preceding the new (primary) T140block, MUST be included.

These rules allow the sequence numbers for the redundant T140blocks to be inferred by counting backwards from the sequence number in the RTP header. The result will be that all the text in the payload will be contiguous and in order.

If there is a gap in the received RTP sequence numbers, and redundant T140blocks are available in a subsequent packet, the sequence numbers for the redundant T140blocks should be inferred by counting backwards from the sequence number in the RTP header for that packet. If there are redundant T140blocks with sequence numbers matching those that are missing, the redundant T140blocks may be substituted for the missing T140blocks.

5. Recommended Procedure

This section contains RECOMMENDED procedures for usage of the payload format. Based on the information in the received packets, the receiver can:

- reorder text received out of order.
- mark where text is missing because of packet loss.
- compensate for lost packets by using redundant data.

5.1. Recommended Basic Procedure

Packets are transmitted when there is valid T.140 data to transmit.

T.140 specifies that T.140 data MAY be buffered for transmission with a maximum buffering time of 500 ms. A buffering time of 300 ms is RECOMMENDED when the application or end-to-end network conditions are not known to require another value.

If no new data is available for a longer period than the buffering time, the transmission process is in an idle period.

When new text is available for transmission after an idle period, it is RECOMMENDED to send it as soon as possible. After this transmission, it is RECOMMENDED to buffer T.140 data in buffering time intervals, until the next idle period. This is done in order to keep the maximum bit rate usage for text at a reasonable level. The buffering time MUST be selected so that text users will perceive a real-time text flow.

5.2. Transmission before and after "Idle Periods"

When valid T.140 data has been sent and no new T.140 data is available for transmission after the selected buffering time, an empty T140block SHOULD be transmitted. This situation is regarded as the beginning of an idle period. The procedure is recommended in order to more rapidly detect potentially missing text before an idle period.

An empty T140block contains no data.

When redundancy is used, transmission continues with a packet at every transmission timer expiration and insertion of an empty T.140block as primary, until the last non-empty T140block has been transmitted, as primary and as redundant data, with all intended generations of redundancy. The last packet before an idle period will contain only one non-empty T140block as redundant data, while the remainder of the redundancy packet will contain empty T140blocks.

Any empty T140block sent as primary data MUST be included as redundant T140blocks in subsequent packets, just as normal text T140blocks would be, unless the empty T140block is too old to be transmitted. This is done so that sequence number inference for the redundant T140blocks will be correct, as explained in Section 4.2.

After an idle period, the transmitter SHOULD set the M-bit to one in the first packet with new text.

5.3. Detection of Lost Text Packets

Packet loss for text/t140 packets MAY be detected by observing gaps in the sequence numbers of RTP packets received by the receiver.

With text/t140, the loss of packets is usually detected by comparison of the sequence of RTP packets as they arrive. Any discrepancy MAY be used to indicate loss. The highest RTP sequence number received may also be compared with that in RTCP reports, as an additional check for loss of the last packet before an idle period.

Missing data SHOULD be marked by insertion of a missing text marker in the received stream for each missing T140block, as specified in ITU-T T.140 Addendum 1 [1].

Because empty T140blocks are transmitted in the beginning of an idle period, there is a slight risk of falsely marking loss of text, when only an empty T140block was lost. Procedures based on detection of the packet with the M-bit set to one MAY be used to reduce the risk of introducing false markers of loss.

If redundancy is used with the text/t140 format, and a packet is received with fewer redundancy levels than normally in the session, it SHOULD be treated as if one empty T140block has been received for each excluded level in the received packet. This is because the only occasion when a T140block is excluded from transmission is when it is an empty T140block that has become too old to be transmitted.

If two successive packets have the same number of redundant generations, it SHOULD be treated as the general redundancy level for the session. Change of the general redundancy level SHOULD only be done after an idle period.

The text/t140 format relies on use of the sequence number in the RTP packet header for detection of loss and, therefore, is not suitable for applications where it needs to be alternating with other payloads in the same RTP stream. It would be complicated and unreliable to

try to detect loss of data at the edges of the shifts between t140 text and other stream contents. Therefore, text/t140 is RECOMMENDED to be the only payload type in the RTP stream.

5.4. Compensation for Packets Out of Order

For protection against packets arriving out of order, the following procedure MAY be implemented in the receiver. If analysis of a received packet reveals a gap in the sequence and no redundant data is available to fill that gap, the received packet SHOULD be kept in a buffer to allow time for the missing packet(s) to arrive. It is RECOMMENDED that the waiting time be limited to 1 second.

If a packet with a T140block belonging to the gap arrives before the waiting time expires, this T140block is inserted into the gap and then consecutive T140blocks from the leading edge of the gap may be consumed. Any T140block that does not arrive before the time limit expires should be treated as lost and a missing text marker should be inserted (see Section 5.3).

6. Parameter for Character Transmission Rate

In some cases, it is necessary to limit the rate at which characters are transmitted. For example, when a Public Switched Telephone Network (PSTN) gateway is interworking between an IP device and a PSTN textphone, it may be necessary to limit the character rate from the IP device in order to avoid throwing away characters (in case of buffer overflow at the PSTN gateway).

To control the character transmission rate, the MIME parameter "cps" in the "fmtp" attribute [7] is defined (see Section 10). It is used in SDP with the following syntax:

```
a=fmtp:<format> cps=<integer>
```

The <format> field is populated with the payload type that is used for text. The <integer> field contains an integer representing the maximum number of characters that may be received per second. The value shall be used as a mean value over any 10-second interval. The default value is 30.

Examples of use in SDP are found in Section 7.2.

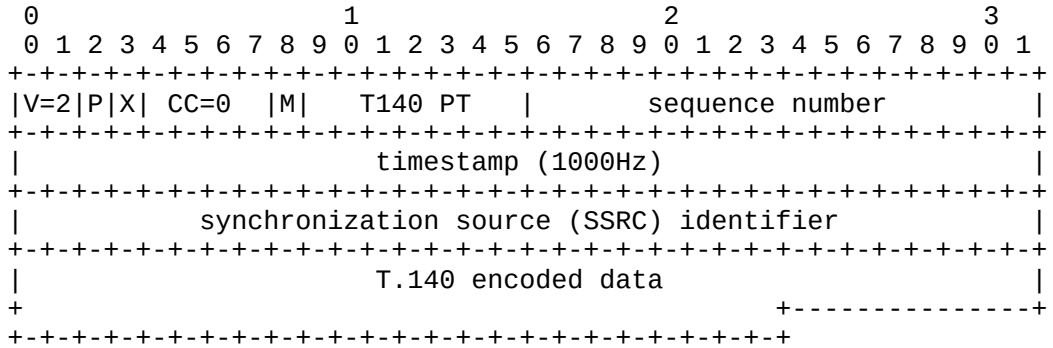
In receipt of this parameter, devices MUST adhere to the request by transmitting characters at a rate at or below the specified <integer> value. Note that this parameter was not defined in RFC 2793 [16]. Therefore implementations of the text/t140 format may be in use that do not recognize and act according to this parameter. Therefore,

receivers of text/t140 MUST be designed so they can handle temporary reception of characters at a higher rate than this parameter specifies. As a result malfunction due to buffer overflow is avoided for text conversation with human input.

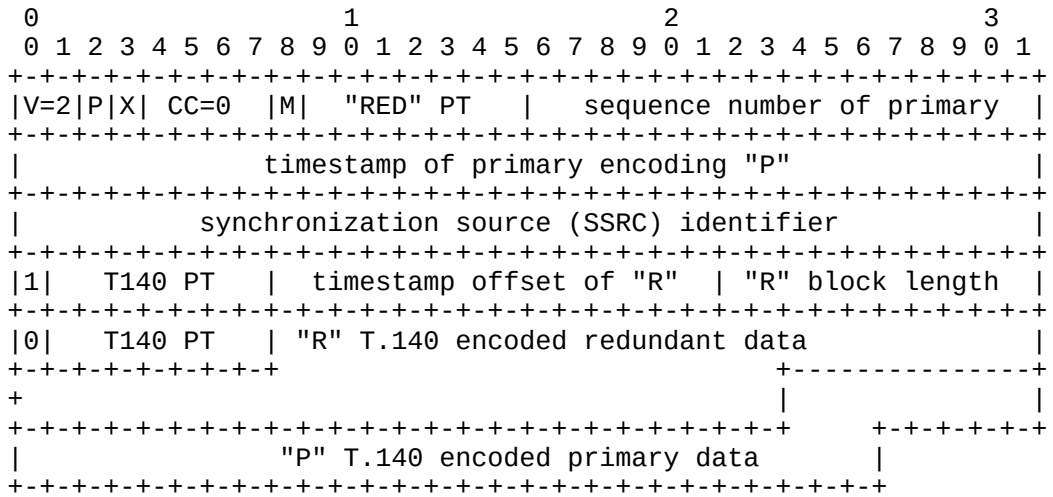
7. Examples

7.1. RTP Packetization Examples for the text/t140 Format

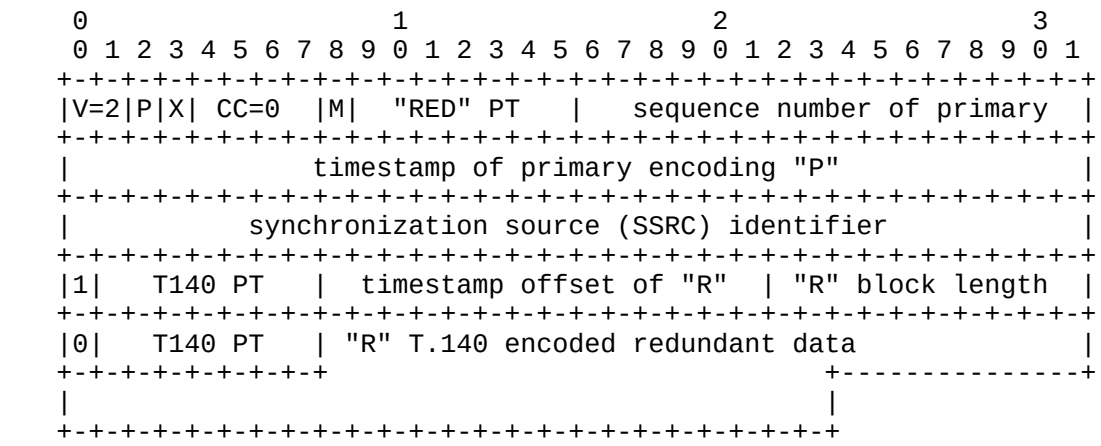
Below is an example of a text/t140 RTP packet without redundancy.



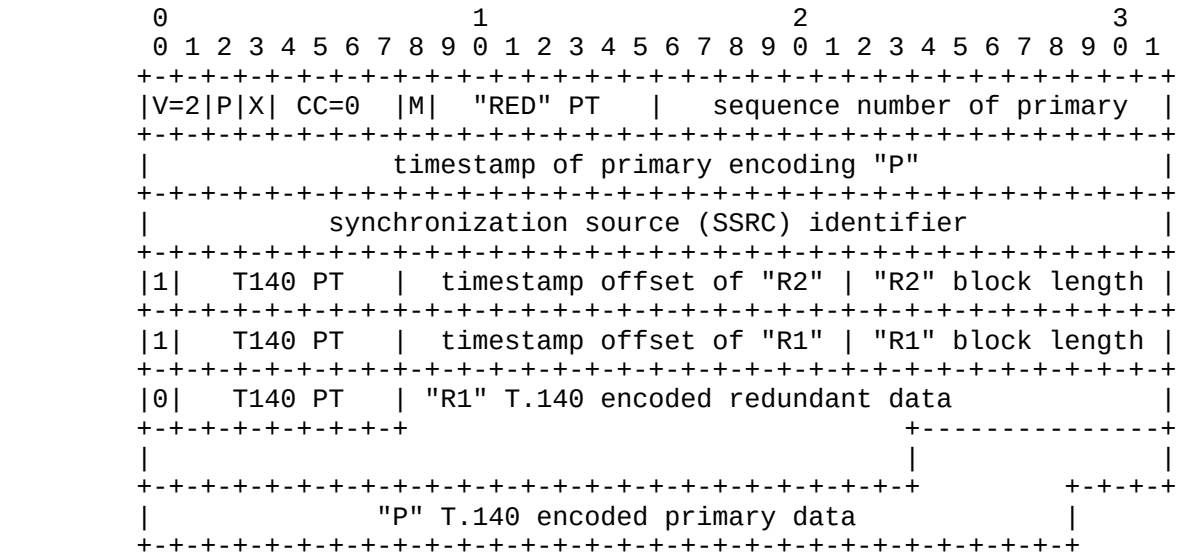
Below is an example of a text/t140 RTP packet with one redundant T140block.



Below is an example of an RTP packet with one redundant T140block using text/t140 payload format. The primary data block is empty, which is the case when transmitting a packet for the sole purpose of forcing the redundant data to be transmitted in the absence of any new data.



As a follow-on to the previous example, the example below shows the next RTP packet in the sequence, which does contain a real T140block when using the text/t140 payload format. Note that the empty block is present in the redundant transmissions of the text/t140 payload format. This example shows two levels of redundancy and one primary data block. The value of the "R2 block length" would be set to zero in order to represent the empty T140block.



7.2. SDP Examples

Below is an example of SDP, which describes RTP text transport on port 11000:

```
m=text 11000 RTP/AVP 98
a=rtpmap:98 t140/1000
```

Below is an example of SDP that is similar to the above example, but also utilizes RFC 2198 to provide the recommended two levels of redundancy for the text packets:

```
m=text 11000 RTP/AVP 98 100
a=rtpmap:98 t140/1000
a=rtpmap:100 red/1000
a=fmtp:100 98/98/98
```

Note: Although these examples utilize the RTP/AVP profile, it is not intended to limit the scope of this memo. Any appropriate profile may be used in conjunction with this memo.

8. Security Considerations

All of the security considerations from Section 14 of RFC 3550 [2] apply.

8.1. Confidentiality

Because the intention of the described payload format is to carry text in a text conversation, security measures in the form of encryption are of importance. The amount of data in a text conversation session is low. Therefore, any encryption method MAY be selected and applied to T.140 session contents or to whole RTP packets. Secure Real-time Transport Protocol (SRTP) [14] provides a suitable method for ensuring confidentiality.

8.2. Integrity

It may be desirable to protect the text contents of an RTP stream against manipulation. SRTP [14] provides methods for providing integrity that MAY be applied.

8.3. Source Authentication

There are several methods of making sure the source of the text is the intended one.

Text streams are usually used in a multimedia control environment. Security measures for authentication are available and SHOULD be applied in the registration and session establishment procedures, so that the identity of the sender of the text stream is reliably associated with the person or device setting up the session. Once established, SRTP [14] mechanisms MAY be applied to ascertain that the source is maintained the same during the session.

9. Congestion Considerations

The congestion considerations from Section 10 of RFC 3550 [2], Section 6 of RFC 2198 [3], and any used profile (e.g., the section about congestion in chapter 2 of RFC 3551 [11]) apply with the following application-specific considerations.

Automated systems MUST NOT use this format to send large amounts of text at rates significantly above those a human user could enter.

Even if the network load from users of text conversation is usually very low, for best-effort networks an application MUST monitor the packet loss rate and take appropriate actions to reduce its sending rate (if this application sends at higher rate than what TCP would

achieve over the same path). The reason for this is that this application, due to its recommended usage of two or more redundancy levels, is very robust against packet loss. At the same time, due to the low bit-rate of text conversations, if one considers the discussion in RFC 3714 [13], this application will experience very high packet loss rates before it needs to perform any reduction in the sending rate.

If the application needs to reduce its sending rate, it SHOULD NOT reduce the number of redundancy levels below the default amount specified in Section 4. Instead, the following actions are RECOMMENDED in order of priority:

- Increase the shortest time between transmissions (described in Section 5.1) from the recommended 300 ms to 500 ms, which is the highest value allowed according to T.140.

- Limit the maximum rate of characters transmitted.

- Increase the shortest time between transmissions to a higher value, not higher than 5 seconds. This will cause unpleasant delays in transmission, beyond what is allowed according to T.140, but text will still be conveyed in the session with some usability.

- Exclude participants from the session.

Please note that if the reduction in bit-rate achieved through the above measures is not sufficient, the only remaining action is to terminate the session.

As guidance, some load figures are provided here as examples based on use of IPv4, including the load from IP, UDP, and RTP headers without compression .

- Experience tells that a common mean character transmission rate, during a complete PSTN text telephony session, is around two characters per second.

- A maximum performance of 20 characters per second is enough even for voice-to-text applications.

- With the (unusually high) load of 20 characters per second, in a language that makes use of three octets per UTF-8 character, two redundant levels, and 300 ms between transmissions, the maximum load of this application is 3300 bits/s.

- When the restrictions mentioned above are applied, limiting transmission to 10 characters per second, using 5 s between transmissions, the maximum load of this application, in a language that uses one octet per UTF-8 character, is 300 bits/s.

Note that this payload can be used in a congested situation as a last resort to maintain some contact when audio and video media need to be stopped. The availability of one low bit-rate stream for text in such adverse situations may be crucial for maintaining some communication in a critical situation.

10. IANA Considerations

This document updates the RTP payload format named "t140" and the associated MIME type "text/t140", in the IANA RTP and Media Type registries.

10.1. Registration of MIME Media Type text/t140

MIME media type name: text

MIME subtype name: t140

Required parameters: rate: The RTP timestamp clock rate, which is equal to the sampling rate. The only valid value is 1000.

Optional parameters: cps: The maximum number of characters that may be received per second. The default value is 30.

Encoding considerations: T.140 text can be transmitted with RTP as specified in RFC 4103.

Security considerations: See Section 8 of RFC 4103.

Interoperability considerations: This format is the same as specified in RFC2793. For RFC2793 the "cps=" parameter was not defined. Therefore, there may be implementations that do not consider this parameter. Receivers need to take that into account.

Published specification: ITU-T T.140 Recommendation. RFC 4103.

Applications which use this media type: Text communication terminals and text conferencing tools.

Additional information: This type is only defined for transfer via RTP.

Magic number(s): None

File extension(s): None
Macintosh File Type Code(s): None

Person & email address to contact for further information:
Gunnar Hellstrom
E-mail: gunnar.hellstrom@omnitor.se

Intended usage: COMMON

Author	/ Change controller:
Gunnar Hellstrom	IETF avt WG
gunnar.hellstrom@omnitor.se	

10.2. SDP Mapping of MIME Parameters

The information carried in the MIME media type specification has a specific mapping to fields in the Session Description Protocol (SDP) [7], which is commonly used to describe RTP sessions. When SDP is used to specify sessions employing the text/t140 format, the mapping is as follows:

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

- The MIME subtype (payload format name) goes in SDP "a=rtpmap" as the encoding name. The RTP clock rate in "a=rtpmap" MUST be 1000 for text/t140.

- The parameter "cps" goes in SDP "a=fmtp" attribute.

- When the payload type is used with redundancy according to RFC 2198, the level of redundancy is shown by the number of elements in the slash-separated payload type list in the "fmtp" parameter of the redundancy declaration as defined in RFC 4102 [9] and RFC 2198 [3].

10.3. Offer/Answer Consideration

In order to achieve interoperability within the framework of the offer/answer model [10], the following consideration should be made:

- The "cps" parameter is declarative. Both sides may provide a value, which is independent of the other side.

11. Acknowledgements

The authors want to thank Stephen Casner, Magnus Westerlund, and Colin Perkins for valuable support with reviews and advice on creation of this document, to Mickey Nasiri at Ericsson Mobile Communication for providing the development environment, Michele Mizarro for verification of the usability of the payload format for its intended purpose, and Andreas Piirimets for editing support and validation.

12. Normative References

[1] ITU-T Recommendation T.140 (1998) - Text conversation protocol for multimedia application, with amendment 1, (2000).

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

[3] 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.

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

[5] ISO/IEC 10646-1: (1993), Universal Multiple Octet Coded Character Set.

[6] Yergeau, F., "UTF-8, a transformation format of ISO 10646", STD 63, RFC 3629, November 2003.

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

[8] Rosenberg, J. and H. Schulzrinne, "An RTP Payload Format for Generic Forward Error Correction", RFC 2733, December 1999.

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

[10] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with the Session Description Protocol (SDP)", RFC 3264, June 2002.

[11] Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video Conference with Minimal Control", STD 65, RFC 3551, July 2003.

[12] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.

13. Informative References

[13] Floyd, S. and J. Kempf, "IAB Concerns Regarding Congestion Control for Voice Traffic in the Internet", RFC 3714, March 2004.

[14] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, March 2004.

[15] Schulzrinne, H. and S. Petrack, "RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals", RFC 2833, May 2000.

[16] Hellstrom, G., "RTP Payload for Text Conversation", RFC 2793, May 2000.

[17] ITU-T Recommendation F.703, Multimedia Conversational Services, November 2000.

Authors' Addresses

Gunnar Hellstrom
Omnitor AB
Renathvagen 2
SE-121 37 Johanneshov
Sweden

Phone: +46 708 204 288 / +46 8 556 002 03
Fax: +46 8 556 002 06
EMail: gunnar.hellstrom@omnitor.se

Paul E. Jones
Cisco Systems, Inc.
7025 Kit Creek Rd.
Research Triangle Park, NC 27709
USA

Phone: +1 919 392 6948
EMail: paulej@packetizer.com

Full Copyright Statement

Intellectual Property

Acknowledgement

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