A/V Transport Payloads Workgroup
Internet-Draft
Intended status: Informational
Expires: August 26, 2019

J. Sandford
British Broadcasting Corporation

February 22, 2019

RTP Payload for TTML Timed Text
draft-sandford-payload-rtp-ttml-03

Abstract

This memo describes a Real-time Transport Protocol (RTP) payload format for TTML, an XML based timed text format for live and file based workflows from W3C. This payload format is specifically targeted at live workflows using TTML.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on August 26, 2019.

Table of Contents

1. Introduction

TTML (Timed Text Markup Language)[TTML2] is a media type for describing timed text such as closed captions (also known as subtitles) in television workflows or broadcasts as XML. This document specifies how TTML should be mapped into an RTP stream in live workflows including, but not restricted to, those described in the television broadcast oriented EBU-TT Part 3[TECH3370] specification. This document does not define a media type for TTML but makes use of the existing application/ttml+xml media type [TTML-MTPR].

2. Conventions, Definitions, and Abbreviations

Unless otherwise stated, the term "document" is used in this draft to refer to the TTML document being transmitted in the payload of the RTP packet(s).

Where the term "word" is used in this draft, it is to refer to byte aligned or 32-bit aligned words of data in a computing sense and not to refer to linguistic words that might appear in the transported text.

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 BCP14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Media Format Description

3.1. Relation to Other Text Payload Types

Prior payload types for text are not suited to the carriage of closed captions in Television Workflows. RFC 4103 for Text Conversation [RFC4103] is intended for low data rate conversation with its own session management and minimal formatting capabilities. RFC 4734 Events for Modem, Fax, and Text Telephony Signals [RFC4734] deals in large parts with the control signalling of facsimile and other systems. RFC 4396 for 3rd Generation Partnership Project (3GPP) Timed Text [RFC4396] describes the carriage of a timed text format with much more restricted formatting capabilities than TTML. The lack of an existing format for TTML or generic XML has necessitated the creation of this payload format.

4. Payload Format

In addition to the required RTP headers, the payload contains a section for the TTML document being transmitted (User Data Words), and a field for the Length of that data. Each RTP payload contains one or part of one TTML document.

A representation of the payload format for TTML is Figure 1.

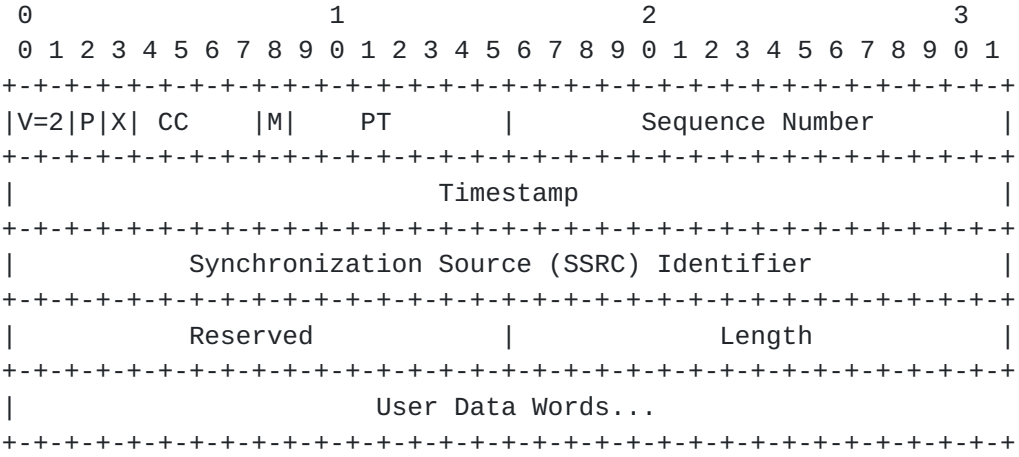

Figure 1: RTP Payload Format for TTML

4.1. RTP Header Usage

RTP packet header fields SHALL be interpreted as per RFC 3550 [RFC3550], with the following specifics:

Marker Bit (M): 1 bit
The Marker Bit is set to "1" to indicate the last packet of a document. Otherwise set to "0". Note: The first packet might also be the last.

Timestamp: 32 bits
The RTP Timestamp encodes the time of the text in the packet. The clock frequency used is dependent on the application and is specified in the media type rate parameter as per Section 7.1. Documents spread across multiple packets MUST use the same timestamp but different consecutive Sequence Numbers. Sequential documents MUST NOT use the same timestamp. Because packets do not represent any constant duration, the timestamp cannot be used to directly infer packet loss.

Reserved: 16 bits
These bits are reserved for future use and MUST be set to 0x0.

Length: 16 bits
The length of User Data Words in bytes.

User Data Words: integer number of words whose length is defined by the character encoding
User Data Words contains the text of the whole document being transmitted or a part of the document being transmitted. Documents using character encodings where characters are not represented by a single byte MUST be serialized in big endian order, a.k.a. network byte order. When the document spans more than one RTP packet, the entire document is obtained by concatenating User Data Words from each contributing packet in ascending order of Sequence Number.

4.2. Payload Data

Documents carried in User Data Words are encoded in accordance with one of the defined TTML profiles specified in its registry [TTML-MTPR]. These profiles specify the document structure used, systems models, timing, and other considerations.

Additionally, documents carried over RTP MUST conform to the following profile.

4.2.1. TTML Profile for RTP Carriage

This section defines constraints on the content and processing of the TTML payload for RTP carriage.

4.2.1.1. Payload content restrictions

Multiple TTML subtitle streams MUST NOT be interleaved in a single RTP stream.

The TTML document instance MUST use the "media" value of the "ttp:timeBase" parameter attribute on the root element.

This is equivalent to the following TTML2 content profile definition document:

```
<?xml version="1.0" encoding="UTF-8"?>
<profile xmlns="http://www.w3.org/ns/ttml#parameter"
    xmlns:ttm="http://www.w3.org/ns/ttml#metadata"
    xmlns:tt="http://www.w3.org/ns/ttml"
    type="content"
    designator="urn:ietf:rfc:XXXX#content"
    combine="mostRestrictive">
    <features xml:base="http://www.w3.org/ns/ttml/feature/">
        <tt:metadata>
            <ttm:desc>
                This document is a minimal TTML2 content profile
                definition document intended to express the minimal
                requirements to apply when carrying TTML over RTP.
            </ttm:desc>
        </tt:metadata>
        <feature value="required">#timeBase-media</feature>
        <feature value="prohibited">#timeBase-smpte</feature>
        <feature value="prohibited">#timeBase-clock</feature>
    </features>
</profile>
```

4.2.1.2. Payload processing requirements

If the TTML document payload is assessed to be invalid then it MUST be discarded. When processing a valid document, the following requirements apply.

The epoch E relative to which computed TTML media times are offset MUST be set to the RTP Timestamp in the header of the RTP packet in which the TTML document instance is carried.

When processing a sequence of TTML documents each delivered in the same RTP stream, exactly zero or one document SHALL be considered active at each moment in the RTP time line.

Each TTML document becomes active at E. In the event that a document D_(n-1) with E_(n-1) is active, and document D_(n) is delivered with E_(n) where E_(n-1) < E_(n), processing of D_(n-1) MUST be stopped at E_(n) and processing of D_(n) MUST begin.

When all defined content within a document has ended, i.e. the active intermediate synchronic document contains no content, then processing of the document MAY be stopped.

4.2.1.2.1. TTML Processor profile

4.2.1.2.1.1. Feature extension designation

This specification defines the following TTML feature extension designation:

o urn:ietf:rfc:XXXX#rtp-relative-media-time

The namespace "urn:ietf:rfc:XXXX" is as defined by [RFC2648].

A TTML content processor supports the "#rtp-relative-media-time" feature extension if it processes media times in accordance with the payload processing requirements specified in this document, i.e. that the epoch E is set to the time equivalent to the RTP Timestamp as detailed above in Section 4.2.1.2.

4.2.1.2.1.2. Processor profile document

The required syntax and semantics declared in the following minimal TTML2 processor profile MUST be supported by the receiver, as signified by those "feature" or "extension" elements whose "value" attribute is set to "required":

```
<?xml version="1.0" encoding="UTF-8"?>
<profile xmlns="http://www.w3.org/ns/ttml#parameter"
    xmlns:ttm="http://www.w3.org/ns/ttml#metadata"
    xmlns:tt="http://www.w3.org/ns/ttml"
    type="processor"
    designator="urn:ietf:rfc:XXXX#processor"
    combine="mostRestrictive">
    <features xml:base="http://www.w3.org/ns/ttml/feature/">
        <tt:metadata>
            <ttm:desc>
                This document is a minimal TTML2 processor profile
                definition document intended to express the minimal
                requirements of a TTML processor able to process TTML
                delivered over RTP according to RFC XXXX.
            </ttm:desc>
        </tt:metadata>
        <feature value="required">#timeBase-media</feature>
        <feature value="optional">#profile-full-version-2</feature>
    </features>
    <extensions xml:base="urn:ietf:rfc:XXXX">
        <extension restricts="#timeBase-media" value="required">
            #rtp-relative-media-time
        </extension>
    </extensions>
</profile>
```

Note that this requirement does not imply that the receiver needs to support either TTML1 or TTML2 profile processing, i.e. the TTML2 "#profile-full-version-2" feature or any of its dependent features.

4.2.1.2.1.3. Processor profile signalling

The "codecs" media type parameter MUST specify at least one processor profile. The processor profiles specified in "codecs" MUST be compatible with the processor profile specified in this document. Where multiple options exist in "codecs" for possible processor profile combinations (i.e. separated by "|" operator), every permitted option MUST be compatible with the processor profile specified in this document. Where processor profiles other than the one specified in this document are advertised in the "codecs" parameter, the requirements of the processor profile specified in this document MAY be signalled additionally using the "+" operator with its registered short code.

A processor profile (X) is compatible with the processor profile in this document (P) if X includes all the features and extensions in P, identified by their character content, and the "value" attribute of each is at least as restrictive as the "value" attribute of the

feature or extension in P that has the same character content. The term "restrictive" here is as defined in [TTML2] Section 6.

Note that short codes for TTML profiles are registered at [TTML-MTPR].

4.2.1.2.2. EBU-TT Live considerations

EBU-TT Live is a profile of TTML intended to support live contribution of TTML documents as a stream independently of the carriage mechanism. When EBU-TT Live documents are carried in an RTP stream, or when the TTML documents being transferred over RTP use EBU-TT Live semantics, the following considerations apply:

E is considered to be the Availability Time as defined by EBU-TT Live. It is an error if two documents are delivered such that E_(n-1) < E_(n) and the "ebuttp:sequenceNumber" of E_(n-1) is greater than the "ebuttp:sequenceNumber" of E_(n). Every EBU-TT Live document in a single RTP stream MUST have a "ebuttp:sequenceIdentifier" with the same value.

5. Payload Examples

The following is an example of a valid TTML document that may be carried using the payload format described in this document:

```
<?xml version="1.0" encoding="UTF-8"?>
<tt xml:lang="en"
 xmlns="http://www.w3.org/ns/ttml"
 xmlns:ttm="http://www.w3.org/ns/ttml#metadata"
 xmlns:ttp="http://www.w3.org/ns/ttml#parameter"
 xmlns:tts="http://www.w3.org/ns/ttml#styling"
 ttp:timeBase="media"
 >
  <head>
    <metadata>
      <ttm:title>Timed Text TTML Example</ttm:title>
      <ttm:copyright>The Authors (c) 2006</ttm:copyright>
    </metadata>
    <styling>
      <!-- s1 specifies default color, font, and text alignment -->
      <style xml:id="s1"
        tts:color="white"
        tts:fontFamily="proportionalSansSerif"
        tts:fontSize="100%"
        tts:textAlign="center"
      />
    </styling>
    <layout>
      <region xml:id="subtitleArea"
        style="s1"
        tts:extent="78% 11%"
        tts:padding="1% 5%"
        tts:backgroundColor="black"
        tts:displayAlign="after"
      />
    </layout>
  </head>
  <body region="subtitleArea">
    <div>
      <p xml:id="subtitle1" dur="5.0s" style="s1">
        How truly delightful!
      </p>
    </div>
  </body>
</tt>
```

6. Congestion Control Considerations

Congestion control for RTP SHALL be used in accordance with RFC 3550 [RFC3550], and with any applicable RTP profile: e.g., RFC 3551 [RFC3551]. An additional requirement if best-effort service is being used is users of this payload format MUST monitor packet loss to ensure that the packet loss rate is within acceptable parameters.

Circuit Breakers [RFC8083] is an update to RTP [RFC3550] that defines criteria for when one is required to stop sending RTP Packet Streams and applications implementing this standard MUST comply with it. RFC 8085 [RFC8083] provides additional information on the best practices for applying congestion control to UDP streams.

7. Payload Format Parameters

This RTP payload format is identified using the existing application/ ttml+xml media type as registered with IANA [IANA] and defined in [TTML-MTPR].

7.1. Clock Rate

The default clock rate for TTML over RTP is 1000Hz. The clock rate SHOULD be included in any advertisements of the RTP stream where possible. This parameter has not been added to the media type definition as it is not applicable to TTML usage other than within RTP streams. In other contexts, timing is defined within the TTML document.

When choosing a clock rate, implementers should consider what other media their TTML streams may be used in conjunction with (e.g. video or audio). It may be appropriate to use the same Synchronization Source and Clock Rate as the related media. As TTML streams may be aperiodic, implementers should also consider the frequency range over which they expect packets to be sent and the temporal resolution required.

7.2. Mapping to SDP

The mapping of the application/ttml+xml media type and its parameters [TTML-MTPR] SHALL be done according to Section 3 of RFC 4855 [RFC4855].

- The type name "application" goes in SDP "m=" as the media name.
- The media subtype "ttml+xml" goes in SDP "a=rtpmap" as the encoding name,
- The clock rate also goes in "a=rtpmap" as the clock rate.

Additional format specific parameters as described in the media type specification SHALL be included in the SDP file in "a=fmtp" as a semicolon separated list of "parameter=value" pairs as described in [RFC4855]. The "codecs" parameter MUST be included in the SDP file. Specific requirements for the "codecs" parameter are included in Section 4.2.1.2.1.3.

7.2.1. Examples

A sample SDP mapping is as follows:

```
m=application 30000 RTP/AVP 112
a=rtpmap:112 ttml+xml/90000
a=fmtp:112 charset=utf-8;codecs=im1t
```

In this example, a dynamic payload type 112 is used. The 90 kHz RTP timestamp rate is specified in the "a=rtpmap" line after the subtype. The codecs parameter defined in the "a=fmtp" line indicates that the TTML data conforms to IMSC 1 Text profile.

8. IANA Considerations

No IANA action.

9. Security Considerations

RTP packets using the payload format defined in this specification are subject to the security considerations discussed in the RTP specification [RFC3550] , and in any applicable RTP profile such as RTP/AVP [RFC3551], RTP/AVPF [RFC4585], RTP/SAVP [RFC3711], or RTP/SAVPF [RFC5124]. However, as "Securing the RTP Protocol Framework: Why RTP Does Not Mandate a Single Media Security Solution" [RFC7202] discusses, it is not an RTP payload format's responsibility to discuss or mandate what solutions are used to meet the basic security goals like confidentiality, integrity, and source authenticity for RTP in general. This responsibility lays on anyone using RTP in an application. They can find guidance on available security mechanisms and important considerations in "Options for Securing RTP Sessions" [RFC7201]. Applications SHOULD use one or more appropriate strong security mechanisms. The rest of this Security Considerations section discusses the security impacting properties of the payload format itself.

To avoid potential buffer overflow attacks, receivers should take care to validate that the User Data Words in the RTP payload are of the appropriate length (using the Length field).

This payload format places no specific restrictions on the size of TTML documents that may be transmitted. As such, malicious implementations could be used to perform denial-of-service (DoS) attacks. RFC 4732 [RFC4732] provides more information on DoS attacks and describes some mitigation strategies. Implementers should take into consideration that the size and frequency of documents transmitted using this format may vary over time. As such, sender implementations should avoid producing streams that exhibit DoS-like

behaviour and receivers should avoid false identification of a legitimate stream as malicious.

As with other XML types and as noted in RFC 7303 [RFC7303], XML Media Types, Section 10, repeated expansion of maliciously constructed XML entities can be used to consume large amounts of memory, which may cause XML processors in constrained environments to fail.

In addition, because of the extensibility features for TTML and of XML in general, it is possible that "application/ttml+xml" may describe content that has security implications beyond those described here. However, TTML does not provide for any sort of active or executable content, and if the processor follows only the normative semantics of the published specification, this content will be outside TTML namespaces and may be ignored. Only in the case where the processor recognizes and processes the additional content, or where further processing of that content is dispatched to other processors, would security issues potentially arise. And in that case, they would fall outside the domain of this RTP payload format and the application/ttml+xml registration document.

Although not prohibited, there are no expectations that XML signatures or encryption would normally be employed.

Further information related to privacy and security at a document level can be found in TTML 2 Appendix P [TTML2].

10. References

10.1. Normative References

[IANA] "IANA - Media Types - Application", February 2019, <https://www.iana.org/assignments/media-types/media-types.xhtml#application>.

[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>.

[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>.

[RFC4732] Handley, M., Ed., Rescorla, E., Ed., and IAB, "Internet Denial-of-Service Considerations", RFC 4732, DOI 10.17487/RFC4732, December 2006, <https://www.rfc-editor.org/info/rfc4732>.

[RFC4855] Casner, S., "Media Type Registration of RTP Payload Formats", RFC 4855, DOI 10.17487/RFC4855, February 2007, <https://www.rfc-editor.org/info/rfc4855>.

[RFC7201] Westerlund, M. and C. Perkins, "Options for Securing RTP Sessions", RFC 7201, DOI 10.17487/RFC7201, April 2014, <https://www.rfc-editor.org/info/rfc7201>.

[RFC7303] Thompson, H. and C. Lilley, "XML Media Types", RFC 7303, DOI 10.17487/RFC7303, July 2014, <https://www.rfc-editor.org/info/rfc7303>.

[RFC8083] Perkins, C. and V. Singh, "Multimedia Congestion Control: Circuit Breakers for Unicast RTP Sessions", RFC 8083, DOI 10.17487/RFC8083, March 2017, <https://www.rfc-editor.org/info/rfc8083>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[TECH3370] "TECH 3370 - EBU-TT PART 3: LIVE CONTRIBUTION", May 2017, <https://tech.ebu.ch/publications/tech3370>.

[TTML-MTPR] "TTML Media Type Definition and Profile Registry", January 2017, <https://www.w3.org/TR/ttml-profile-registry/>.

[TTML2] "Timed Text Markup Language 2 (TTML2)", November 2018, <https://www.w3.org/TR/ttml2/>.

10.2. Informative References

[RFC2648] Moats, R., "A URN Namespace for IETF Documents", RFC 2648, DOI 10.17487/RFC2648, August 1999, <https://www.rfc-editor.org/info/rfc2648>.

[RFC3551] Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video Conferences with Minimal Control", STD 65, RFC 3551, DOI 10.17487/RFC3551, July 2003, <https://www.rfc-editor.org/info/rfc3551>.

[RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, DOI 10.17487/RFC3711, March 2004, <https://www.rfc-editor.org/info/rfc3711>.

[RFC4396] Rey, J. and Y. Matsui, "RTP Payload Format for 3rd Generation Partnership Project (3GPP) Timed Text", RFC 4396, DOI 10.17487/RFC4396, February 2006, <https://www.rfc-editor.org/info/rfc4396>.

[RFC4585] Ott, J., Wenger, S., Sato, N., Burmeister, C., and J. Rey, "Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF)", RFC 4585, DOI 10.17487/RFC4585, July 2006, <https://www.rfc-editor.org/info/rfc4585>.

[RFC4734] Schulzrinne, H. and T. Taylor, "Definition of Events for Modem, Fax, and Text Telephony Signals", RFC 4734, DOI 10.17487/RFC4734, December 2006, <https://www.rfc-editor.org/info/rfc4734>.

[RFC5124] Ott, J. and E. Carrara, "Extended Secure RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/SAVPF)", RFC 5124, DOI 10.17487/RFC5124, February 2008, <https://www.rfc-editor.org/info/rfc5124>.

[RFC7202] Perkins, C. and M. Westerlund, "Securing the RTP Framework: Why RTP Does Not Mandate a Single Media Security Solution", RFC 7202, DOI 10.17487/RFC7202, April 2014, <https://www.rfc-editor.org/info/rfc7202>.

Appendix A. RFC Editor Considerations

TODO To be filled

Appendix B. Acknowledgements

TODO

Author's Address

James Sandford
British Broadcasting Corporation
Dock House, MediaCityUK
Salford
United Kingdom

Phone: +44 30304 09549
Email: james.sandford@bbc.co.uk