

INTERNET-DRAFT
January 7, 2004
Expires: July 7, 2005

John Lazzaro
CS Division
UC Berkeley

Framing RTP and RTCP Packets over Connection-Oriented Transport

<[draft-ietf-avt-rtp-framing-contrans-05.txt](#)>

Status of this Memo

By submitting this Internet-Draft, I certify that any applicable patent or other IPR claims of which I am aware have been disclosed, and any of which I become aware will be disclosed, in accordance with [RFC 3668](#).

By submitting this Internet-Draft, I accept the provisions of [Section 3 of RFC 3667](#) ([BCP 78](#)).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/lid-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on July 7, 2005.

Copyright Notice

Copyright (C) The Internet Society (2004). All Rights Reserved.

Abstract

This memo defines a method for framing Real Time Protocol (RTP) and Real Time Control Protocol (RTCP) packets onto connection-oriented transport (such as TCP). The memo also defines how session descriptions may specify RTP streams that use the framing method.

Table of Contents

1.	Introduction	2
1.1	Terminology	2
2.	The Framing Method	3
3.	Packet Stream Properties	3
4.	Session Descriptions for RTP over TCP	4
5.	Example	6
6.	Congestion Control	7
7.	Acknowledgements	7
8.	Security Considerations	7
9.	IANA Considerations	8
10.	References	8
10.1	Normative References	8
	Authors' Address	9
	Intellectual Property Rights Statement	9
	Full Copyright Statement	10
	Change Log for < draft-ietf-avt-rtp-framing-contrans-05.txt >	11

[1.](#) Introduction

The Audio/Video Profile (AVP, [[1](#)]) for the Real-Time Protocol (RTP, [[2](#)]) does not define a method for framing RTP and Real Time Control Protocol (RTCP) packets onto connection-oriented transport protocols (such as TCP). However, earlier versions of RTP/AVP did define a framing method, and this method is in use in several implementations.

In this memo, we document the method and show how a session description [[4](#)] may specify the use of the method.

[1.1](#) Terminology

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 [BCP 14](#), [RFC 2119](#) [[5](#)].

2. The Framing Method

Figure 1 defines the framing method.



Figure 1 -- The bitfield definition of the framing method.

A 16-bit unsigned integer LENGTH field, coded in network byte order (big-endian), begins the frame. If LENGTH is non-zero, an RTP or RTCP packet follows the LENGTH field. The value coded in the LENGTH field MUST equal the number of octets in the RTP or RTCP packet. Zero is a valid value for LENGTH, and codes the null packet.

This framing method does not use frame markers (i.e. an octet of constant value that would precede the LENGTH field). Frame markers are useful for detecting errors in the LENGTH field. In lieu of a frame marker, receivers SHOULD monitor the RTP and RTCP header fields whose values are predictable (for example, the RTP version number). See [Appendix A.1](#) of [1] for additional guidance.

3. Packet Stream Properties

In most respects, the framing method does not specify properties above the level of a single packet. In particular, [Section 2](#) does not specify:

Bi-directional issues.

[Section 2](#) defines a framing method for use in one direction on a connection. The relationship between framed packets flowing in defined direction and in the reverse direction is not specified.

Packet loss and reordering.

The reliable nature of a connection does not imply that a framed RTP stream has a contiguous sequence number ordering. For example, if the connection is used to tunnel a UDP stream through a network middlebox that only passes TCP, the sequence numbers in the framed stream reflect any packet loss or reordering on the UDP portion of the end-to-end flow.

Out-of-band semantics.

[Section 2](#) does not define the RTP or RTCP semantics for closing a TCP socket, or of any other "out of band" signal for the connection.

Memos that normatively include the framing method MAY specify these properties. For example, [Section 4](#) of this memo specifies these properties for RTP/AVP sessions specified in session descriptions.

In one respect, the framing protocol DOES specify a property above the level of a single packet. If a direction of a connection carries RTP packets, the streams carried in this direction MUST support the use of multiple SSRCs in those RTP packets. If a direction of a connection carries RTCP packets, the streams carried in this direction MUST support the use of multiple SSRCs in those RTCP packets.

4. Session Descriptions for RTP/AVP over TCP

Session management protocols that use the Session Description Protocol [\[4\]](#) in conjunction with the Offer/Answer Protocol [\[6\]](#) MUST use the methods described in [\[3\]](#) to set up RTP/AVP streams over TCP. In this case, the use of Offer/Answer is REQUIRED, as the setup methods described in [\[3\]](#) rely on Offer/Answer.

In principle, [\[3\]](#) is capable of setting up RTP sessions for any RTP profile. In practice, each profile has unique issues that must be considered when applying [\[3\]](#) to set up streams for the profile.

In this memo, we restrict our focus to the Audio/Video Profile (AVP, [\[2\]](#)). Below, we define a token value ("TCP/RTP/AVP") that signals the use of RTP/AVP in a TCP session. We also define the operational procedures that a TCP/RTP/AVP stream MUST follow.

We expect that other standards-track memos will appear to support the use of the framing method with other RTP profiles. The support memo for a new profile MUST define a token value for the profile, using the style we used for AVP. Thus, for profile xyz, the token value MUST be

"TCP/RTP/xyz". The memo SHOULD adopt the operational procedures we define below for AVP, unless these procedures are in some way incompatible with the profile.

The remainder of this section describes how to setup and use an AVP stream in a TCP session. Figure 2 shows the syntax of a media (m=) line [4] of a session description:

```
"m=" media SP port ["/" integer] SP proto 1*(SP fmt) CRLF
```

Figure 2 -- Syntax for an SDP media (m=) line (from [4]).

The <proto> token value "TCP/RTP/AVP" specifies an RTP/AVP [1] [2] stream that uses the framing method over TCP.

The <fmt> tokens that follow <proto> MUST be unique unsigned integers in the range 0 to 127. The <fmt> tokens specify an RTP payload type associated with the stream.

In all other respects, the session description syntax for the framing method is identical to [3].

The TCP <port> on the media line carries RTP packets. If a media stream uses RTCP, a second connection carries RTCP packets. The port for the RTCP connection is chosen using the algorithms defined in [4] or by the mechanism defined in [7].

The TCP connections MAY carry bi-directional traffic, following the semantics defined in [3]. Both directions of a connection MUST carry the same type of packets (RTP or RTCP). The packets MUST exclusively code the RTP or RTCP streams specified on the media line(s) associated with the connection.

As noted in [1], the use of RTP without RTCP is strongly discouraged. However, if a sender does not wish to send RTCP packets in a media session, the sender MUST add the lines "b=RS:0" AND "b=RR:0" to the media description (from [8]).

If the session descriptions of the offer AND the answer both contain the "b=RS:0" AND "b=RR:0" lines, a TCP flow for the media session MUST NOT be created by either endpoint in the session. In all other cases, endpoints MUST establish two TCP connections for an RTP AVP stream, one for RTP and one for RTCP.

As described in [6], the use of the "sendonly" or "sendrecv" attribute

in an offer (or answer) indicates that the offerer (or answerer) intends to send RTP packets on the RTP TCP connection. The use of the "recvonly" or "sendrecv" attributes in an offer (or answer) indicates that the offerer (or answerer) wishes to receive RTP packets on the RTP TCP connection.

5. Example

The session descriptions in Figure 3-4 define a TCP RTP/AVT session.

```
v=0
o=first 2520644554 2838152170 IN IP4 first.example.net
s=Example
t=0 0
c=IN IP4 192.0.2.105
m=audio 9 TCP/RTP/AVP 11
a=setup:active
a=connection:new
```

Figure 3 -- TCP session description for first participant.

```
v=0
o=second 2520644554 2838152170 IN IP4 second.example.net
s=Example
t=0 0
c=IN IP4 192.0.2.94
m=audio 16112 TCP/RTP/AVP 10 11
a=setup:passive
a=connection:new
```

Figure 4 -- TCP session description for second participant.

The session descriptions define two parties that participate in a connection-oriented RTP/AVP session. The first party (Figure 3) is capable of receiving stereo L16 streams (static payload type 11). The second party (Figure 4) is capable of receiving mono (static payload type 10) or stereo L16 streams.

The "setup" attribute in Figure 3 specifies that the first party is "active" and initiates connections, and the "setup" attribute in Figure 4 specifies that the second party is "passive" and accepts connections [3].

The first party connects to the network address (192.0.2.94) and port

(16112) of the second party. Once the connection is established, it is used bi-directionally: the first party sends framed RTP packets to the second party on one direction of the connection, and the second party sends framed RTP packets to the first party in the other direction of the connection.

The first party also initiates an RTCP TCP connection to port 16113 (16112 + 1, as defined in [4]) of the second party. Once the connection is established, the first party sends framed RTCP packets to the second party on one direction of the connection, and the second party sends framed RTCP packets to the first party in the other direction of the connection.

6. Congestion Control

The RTP congestion control requirements are defined in [1]. As noted in [1], all transport protocols used on the Internet need to address congestion control in some way, and RTP is not an exception.

In addition, the congestion control requirements for the Audio/Video Profile are defined in [2]. The basic congestion control requirement defined in [2] is that RTP sessions should compete fairly with TCP flows that share the network. As the framing method uses TCP, it competes fairly with other TCP flows by definition.

7. Acknowledgements

This memo, in part, documents discussions on the AVT mailing list about TCP and RTP. Thanks to all of the participants in these discussions.

8. Security Considerations

Implementors should carefully read the Security Considerations sections of the RTP [1] and RTP/AVP [2] documents, as most of the issues discussed in these sections directly apply to RTP streams framed over TCP.

Session descriptions that specify connection-oriented media sessions (such as the example session shown in Figures 3-4 of [Section 5](#)) raise unique security concerns for streaming media. The Security Considerations section of [3] describes these issues in detail.

Below, we discuss security issues that are unique to the framing method defined in [Section 2](#).

Attackers may send framed packets with large LENGTH values, to exploit security holes in applications. For example, a C implementation may declare a 1500-byte array as a stack variable, and use LENGTH as the bound on the loop that reads the framed packet into the array. This code would work fine for friendly applications that use Etherframe-sized RTP packets, but may be open to exploit by an attacker. Thus, an implementation needs to handle packets of any length, from a NULL packet (LENGTH == 0) to the maximum-length packet holding 64K octets (LENGTH = 0xFFFF).

9. IANA Considerations

[4] defines the syntax of session description media lines. We reproduce this definition in Figure 2 of [Section 4](#) of this memo. In [Section 4](#), we define a new token value for the <proto> field of media lines: "TCP/RTP/AVP". [Section 4](#) specifies the semantics associated with the <proto> field token, and [Section 5](#) shows an example of its use in a session description.

10. References

10.1 Normative References

- [1] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson. "RTP: A transport protocol for real-time applications", [RFC 3550](#), July 2003.
- [2] Schulzrinne, H., and S. Casner. "RTP Profile for Audio and Video Conferences with Minimal Control", [RFC 3551](#), July 2003.
- [3] Yon, D. and G. Camarillo. Connection-Oriented Media Transport in the Session Description Protocol (SDP), [draft-ietf-mmusic-sdp-comedia-10.txt](#).
- [4] Handley, M., Jacobson, V., and C. Perkins. "SDP: Session Description Protocol", [draft-ietf-mmusic-sdp-new-22.txt](#).
- [5] Bradner, S. "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [6] Rosenberg, J. and H. Schulzrinne. "An Offer/Answer Model with SDP", [RFC 3264](#), June 2002.
- [7] C. Huitema. "Real Time Control Protocol (RTCP) attribute in Session Description Protocol (SDP)", [RFC 3605](#), October 2003.

[8] S. Casner. "Session Description Protocol (SDP) Bandwidth Modifiers for RTP Control Protocol (RTCP) Bandwidth", [RFC 3556](#), July 2003.

Authors' Address

John Lazzaro
UC Berkeley
CS Division
[315 Soda Hall](#)
Berkeley CA 94720-1776
Email: lazzaro@cs.berkeley.edu

Intellectual Property Rights Statement

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Full Copyright Statement

Copyright (C) The Internet Society (2004). This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Acknowledgement

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

Change Log for <[draft-ietf-avt-rtp-framing-contrans-05.txt](#)>

[Note to RFC Editors: this Appendix, and its Table of Contents listing, should be removed from the final version of the memo]

Changes were made in response to Magnus's comments on AVT.

[Issue 1] I am concerned that the draft is written in a bit to AVP centric. I know that the draft only registers, and apparently there is not enough consensus and interest to define any other profile for the moment. However the format and its basic signalling properties would be the same independent of the profile in use.

[Response 1] [Section 4](#) has been rewritten to be less AVP centric. See the first four paragraphs of [Section 4](#).

[Issue 2] [Section 2](#), second paragraph: I think the last sentence could benefit a informative reference to RTP section A.1 for checks that can be used to verify correct alignment.

[Response 2] New text similar to that recommended by Magnus has been added to the final paragraph of [Section 2](#).

[Issue 3] [Section 3](#), the first Undefined property: "The framing method is commonly used for sending a single RTP or RTCP stream over a connection. However, [Section 2](#) does not define this common use as normative, so that (for example) a memo that defines an RTP SSRC multiplexing protocol may use the framing method."

The expected property must be that any contrans supports usage of multiple SSRCS. The behavior to expect needs to be the same for RTP over UDP and RTP over TCP. What comes in on the TCP connection, can be the same as what can come in over UDP port in unicast mode from a single source. The difference between TCP and UDP is really only that you can't receive from multiple sources to the same port as I understand the signalling.

I would like to rephrase and move the paragraph. It should define the expected properties in this case for clarity.

[Response 3] [Section 3](#) has been renamed as "Packet Stream Properties". It begins with the list of unspecified properties, which no longer includes the property discussed in Issue 3. Following this list of unspecified properties is the following text:

In one respect, the framing protocol does specify a property above the level of a single packet. If a direction of a connection carries RTP packets, the streams carried in this direction MUST support the use of multiple SSRCs in those RTP packets. If a direction of a connection carries RTCP packets, the streams carried in this direction MUST support the use of multiple SSRCs in those RTCP packets.

[Issue 4] [Section 4](#). I think the signaling section should clearly define that the basic procedure for establishing the TCP connection that the RTP framing is sent over is using COMEDIA. This should be in the first paragraph.

For example a sentence like: The transport of RTP/AVP over TCP when signaled using SDP and the offer/answer method [[RFC3264](#)] SHALL establish its TCP connection as defined by comedia [xx]. The RTP/AVP over TCP is identified in SDP using the "proto" identifier "TCP/RTP/AVP".

For this SDP "proto" identifier the fmt list ...

I would also like to point out that due to comedia it doesn't seem that this framing method can be used in any non Offer/Answer usage. Or have I missed something in the comedia draft?

[Response 4] Done, see the first four paragraphs of [Section 4](#).

[Issue 5] [Section 4](#), second last paragraph: "The RTP stream MUST have an unbroken sequence number order. RTCP stream packets MUST appear as defined in [[2](#)], with no lost or re-ordered packets. IETF standards-track documents MAY loosen these restrictions on packet loss and packet ordering."

This paragraph is in contradiction with a statement in [section 3](#). I also think that it is wrong to make this requirement on the packets entering the TCP connection.

[Response 5] I deleted the offending paragraph. So, the statement in [Section 3](#) on the topic holds for RTP/AVP by default.

[Issue 6] What is meant with the following sentence in [section 4](#): "The out-of-band semantics for the connection MUST comply with [[3](#)]." I don't think it is clear what is meant with out-of-band semantics.

[Response 6] I deleted the offending sentence. Note that the new text at the start of [Section 4](#) (described in "Response 4" earlier) makes the point I was trying to make in this sentence.

[Issue 7] [Section 4](#). Does this section also need to define the usage of the "a=rtcp" SDP attribute under this profile. Because I think there are advantages of being able to define another TCP port for RTCP than the RTP port + 1.

[Response 7] Normative reference to [RFC 3605](#) was added to the document, which is referenced in the second-to-last paragraph of [Section 4](#).

[Issue 8] [Section 4](#), which method are used to indicate the non-presence of RTCP when using this transport?

[Response 8] A mechanism was added, described in the final 3 paragraphs of [Section 4](#). If a different mechanism is desired, please submit replacement paragraphs that describe the candidate mechanism, so that upon WG approval, we can quickly insert it into the document.

[Issue 9] [Section 8](#). The Length field consideration. I think one can be a bit more direct in the recommendation. I would like to add this following sentence to the end of the paragraph: "Thus, a implementation needs to handle packets of any length from the NULL packet (Length=0) to max length 64K packet (Length=0xFFFF).

[Response 9] Done.

[Issue 10] The lack of recommendations on how to register more identifiers for other profiles and what they would need to consider.

[Response 10] I believe this is now covered in the early part of [Section 4](#).
