Internet Engineering Task Force (IETF)
Request for Comments: 6263
Category: Standards Track
ISSN: 2070-1721
June 2011

X. Marjou
A. Sollaud
France Telecom Orange

Application Mechanism for Keeping Alive the NAT Mappings Associated with RTP / RTP Control Protocol (RTCP) Flows

Abstract

This document lists the different mechanisms that enable applications using the Real-time Transport Protocol (RTP) and the RTP Control Protocol (RTCP) to keep their RTP Network Address Translator (NAT) mappings alive. It also makes a recommendation for a preferred mechanism. This document is not applicable to Interactive Connectivity Establishment (ICE) agents.

Status of This Memo

This is an Internet Standards Track document.

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

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc6263.

Copyright Notice

Table of Contents

1. Introduction

[RFC4787] and [RFC5382] describe Network Address Translator (NAT) behaviors and point out that two key aspects of NAT are mappings (a.k.a. bindings) and keeping them refreshed. This introduces a derived requirement for applications engaged in a multimedia session involving NAT traversal: they need to generate a minimum of flow activity in order to create NAT mappings and maintain them.

When applied to applications using the Real-time Transport Protocol (RTP) [RFC3550], the RTP media stream packets themselves normally fulfill this requirement. However, there exist some cases where RTP does not generate the minimum required flow activity.

The examples are:

- In some RTP usages, such as the Session Initiation Protocol (SIP) [RFC3261], agents can negotiate a unidirectional media stream by using the Session Description Protocol (SDP) [RFC4566] "recvonly" attribute on one agent and "sendonly" on the peer, as defined in [RFC3264]. [RFC3264] directs implementations not to transmit media on the receiving agent. If the agent receiving the media is located on the private side of a NAT, it will never receive RTP packets from the public peer if the NAT mapping has not been created.

- Similarly, a bidirectional media stream can be "put on hold". This is accomplished by using the SDP "sendonly" or "inactive" attributes. Again, [RFC3264] directs implementations to cease transmission of media in these cases. However, doing so may cause NAT bindings to time out, and media won’t be able to come off hold.

- Some RTP payload formats, such as the payload format for text conversation [RFC4103], may send packets so infrequently that the interval exceeds the NAT binding timeouts.

To solve these problems, an agent therefore needs to periodically send keepalive data within the outgoing RTP session of an RTP media stream regardless of whether the media stream is currently inactive, sendonly, recvonly, or sendrecv, and regardless of the presence or value of the bandwidth attribute.

It is important to note that NAT traversal constraints also usually require that the agents use Symmetric RTP / RTP Control Protocol (RTCP) [RFC4961] in addition to RTP keepalive.

This document first states the requirements that must be supported to perform RTP keepalives (Section 3). In a second step, the document reports the different mechanisms to overcome this problem (Section 4). Section 5 finally states the recommended solution for RTP keepalive. Section 6 discusses some media format exceptions. Section 7 adds details about timing and transport considerations. Section 8 documents how to maintain NAT bindings for RTCP.

This document is not applicable to Interactive Connectivity Establishment (ICE) [RFC5245] agents. Indeed, the ICE protocol, together with Session Traversal Utilities for NAT (STUN) [RFC5389] and Traversal Using Relays around NAT (TURN) [RFC5766], solves the overall Network Address Translator (NAT) traversal mechanism of media streams. In the context of RTP media streams, some agents may not require all ICE functionalities and may only need a keepalive mechanism. This document thus applies to such agents, and does not apply to agents implementing ICE.

Note that if a given media uses a codec that already integrates a keepalive mechanism, no additional keepalive mechanism is required at the RTP level.

As mentioned in Section 3.5 of [RFC5405], "It is important to note that keepalive messages are NOT RECOMMENDED for general use -- they are unnecessary for many applications and can consume significant amounts of system and network resources".

2. Terminology

In this document, the key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" are to be interpreted as described in RFC 2119 [RFC2119].

3. Requirements

This section outlines the key requirements that need to be satisfied in order to provide RTP media keepalive.

REQ-1 Some data is sent periodically within the outgoing RTP session for the whole duration of the RTP media stream.

REQ-2 Any type of transport (e.g., UDP, TCP) MUST be supported.

REQ-3 Any media type (e.g., audio, video, text) MUST be supported.

REQ-4 Any media format (e.g., G.711, H.263) MUST be supported.

REQ-5 Session signaling protocols SHOULD NOT be impacted.

REQ-6 Impacts on existing software SHOULD be minimized.

REQ-7 The remote peer SHOULD NOT be impacted.

REQ-8 The support for RTP keepalive SHOULD be described in the SDP.

REQ-9 The solution SHOULD cover the integration with RTCP.

4. List of Alternatives for Performing RTP Keepalive

This section lists, in no particular order, some alternatives that can be used to perform a keepalive message within RTP media streams.

4.1. Empty (0-Byte) Transport Packet

The application sends an empty transport packet (e.g., UDP packet, Datagram Congestion Control Protocol (DCCP) packet).

Con:

- This alternative is specific to each transport protocol.

4.2. RTP Packet with Comfort Noise Payload

The application sends an RTP packet with a comfort noise payload [RFC3389].

Cons:

- o This alternative is limited to audio formats only.
- o Comfort noise needs to be supported by the remote peer.
- o Comfort noise needs to be signaled in SDP offer/answer.
- o The peer is likely to render comfort noise at the other side, so the content of the payload (the noise level) needs to be carefully chosen.

4.3. RTCP Packets Multiplexed with RTP Packets

The application sends RTCP packets in the RTP media path itself (i.e., the same tuples for both RTP and RTCP packets) [RFC5761]. RTCP packets therefore keep the NAT mappings open as long as the requirements for parameter selection are fulfilled as discussed in Section 8.

> Note: The "on hold" procedures of [RFC3264] do not impact RTCP transmissions.

Cons:

- o Multiplexing RTP and RTCP must be supported by the remote peer.
- o Some RTCP monitoring tools expect that RTCP packets are not multiplexed.
- o RTCP must be configured so that the Tmin value [RFC3550] is less than or equal to the Tr interval.

4.4. STUN Indication Packet

The application sends a STUN [RFC5389] Binding Indication packet as specified in ICE [RFC5245].

Thanks to the RTP validity check, STUN packets will be ignored by the RTP stack.

Con:

- The sending agent needs to support STUN.

4.5. RTP Packet with Incorrect Version Number

The application sends an RTP packet with a version number set to zero (i.e., an incorrect version number).

Based on the RTP specification [RFC3550], the peer should perform a header validity check and therefore ignore these types of packets.

Cons:

- Only four version numbers are possible. Using one of them for RTP keepalive would be wasteful.

- [RFC4566] and [RFC3264] mandate that media with inactive and recvonly attributes not be sent; however, this is mitigated, as no real media is sent with this mechanism.

4.6. RTP Packet with Unknown Payload Type

The application sends an RTP packet of 0 length with a dynamic payload type that has not been negotiated by the peers (e.g., not negotiated within the SDP offer/answer, and thus not mapped to any media format).

The sequence number is incremented by one for each packet, as it is sent within the same RTP session as the actual media. The timestamp contains the same value that a media packet would have at this time. The marker bit is not significant for the keepalive packets and is thus set to zero.

The synchronization source (SSRC) is the same as for the media for which keepalive is sent.

Normally, the peer will ignore this packet, as RTP [RFC3550] states that "a receiver MUST ignore packets with payload types that it does not understand".

Cons:

- [RFC4566] and [RFC3264] mandate that media with inactive and recvonly attributes not be sent; however, this is mitigated, as no real media is sent with this mechanism.

- [RFC3550] does not preclude examination of received packets by the peer in an attempt to determine if it is under attack.

- The statement "a receiver MUST ignore packets with payload types that it does not understand" of [RFC3550] is not always observed in real life.

- There is no RTCP reporting for the keepalive packets, as [RFC3550] mandates that RTP packets with payload types that the receiver does not understand be ignored.

- Some RTP payload formats do not handle gaps in RTP sequence number well.

5. Recommended Solution for Keepalive Mechanism

The RECOMMENDED mechanism is that discussed in "RTCP Packets Multiplexed with RTP Packets" (Section 4.3). This mechanism is desirable because it reduces the number of ports when RTP and RTCP are used. It also has the advantage of taking into account RTCP aspects, which is not the case with other mechanisms.

Other mechanisms (Sections 4.1, 4.2, 4.4, 4.5, and 4.6) are NOT RECOMMENDED.

6. Media Format Exceptions

When a given media format does not allow the keepalive solution recommended in Section 5, an alternative mechanism SHOULD be defined in the payload format specification for this media format.

7. Timing and Transport Considerations

An application supporting this specification MUST transmit either keepalive packets or media packets at least once every Tr seconds during the whole duration of the media session.

Tr has different value according to the transport protocol.

For UDP, the minimum RECOMMENDED Tr value is 15 seconds, and Tr SHOULD be configurable to larger values.

For TCP, the recommended Tr value is 7200 seconds.

When using the "RTCP packets multiplexed with RTP packets" solution (Section 4.3) for keepalive, Tr MUST comply with the RTCP timing rules of [RFC3550].

Keepalive packets within a particular RTP session MUST use the tuple (source IP address, source TCP/UDP port, target IP address, target TCP/UDP port) of the regular RTP packets.

The agent SHOULD only send RTP keepalive when it does not send regular RTP packets.

8. RTCP Flow Keepalive

RTCP packets are sent periodically and can thus normally keep the NAT mappings open as long as they are sent frequently enough. There are two conditions for that. First, RTCP needs to be used bidirectionally and in a symmetric fashion, as described in [RFC4961]. Secondly, RTCP needs to be sent frequently enough. However, there are certain configurations that can break this latter assumption.

There are two factors that need to be considered to ensure that RTCP is sent frequently enough. First, the RTCP bandwidth needs to be sufficiently large so that transmission will occur more frequently than the longest acceptable packet transmission interval (Tr). The worst-case RTCP interval (Twc) can be calculated using this formula by inserting the max value of the following parameters:

- Maximum RTCP packet size (avg_rtcp_size_max)

- Maximum number of participants (members_max)

- RTCP receiver bandwidth (rtcp_bw)

The RTCP bandwidth value to use here is for a worst case, which will be the receiver proportion when all members except one are not senders. This can be approximated to be all members. Thus, for sessions where RR and RS values [RFC3556] are used, then rtcp_bw shall be set to RR. For sessions where the [RFC3550]-defined proportions of RTCP bandwidth are used (i.e., 1/4 of the bandwidth for senders and 3/4 of the bandwidth for receivers), then rtcp_bw will be 5% of 3/4 of the AS value [RFC4566] in bits per second.

Twc = 1.5 / 1.21828 * members_max * rtcp_bw / avg_rtcp_size_max * 8

The second factor is the minimum RTCP interval Tmin defined in [RFC3550]. Its base value is 5 seconds, but it might also be scaled to 360 divided by the session bandwidth in kbps. The Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF) [RFC4585] also allows for the setting of a trr-int parameter, which is a minimal RTCP interval for regular RTCP packets. It is also used as the Tmin value in the regular Td calculation. An analysis of the algorithm shows that the longest possible regular RTCP interval is:

```
RTCP_int_max = trr-int * 1.5 + Td * 1.5 / 1.21828
```

And as long as there is sufficient bandwidth according to criteria 1 below, then the algorithm can be simplified by setting Td = trr-int, giving

```
RTCP_int_max = trr-int * (1.5 + 1.5 / 1.21828) = 2.73123 * trr-int
```

Thus, the requirements for the RTCP parameters are as follows for functioning keepalive:

1. Ensure that sufficient RTCP bandwidth is provided by calculating Twc, and ensure that the resulting value is less than or equal to Tr.

2. If AVP or SAVP [RFC3711] is used, the Tmin value can’t be greater than Tr divided by 1.5 / (e-3/2).

3. If AVPF or SAVPF [RFC5124] is to be used, trr-min must not be set to a value greater than Tr / 3.

9. Security Considerations

The RTP keepalive packets are sent on the same path as regular RTP media packets and may be perceived as an attack by a peer. However, [RFC3550] mandates that a peer "ignore packets with payload types that it does not understand". A peer that does not understand the keepalive message will thus appropriately drop the received packets.

10. Acknowledgements

Jonathan Rosenberg provided the major inputs for this document via the ICE specification. Magnus Westerlund provided the text for the RTCP flow keepalive section. In addition, thanks to Alfred E. Heggestad, Colin Perkins, Dan Wing, Gunnar Hellstrom, Hadriel Kaplan, Randell Jesup, Remi Denis-Courmont, Robert Sparks, and Steve Casner for their useful inputs and comments.

11. References

11.1. Normative References

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

[RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.

[RFC4961] Wing, D., "Symmetric RTP / RTP Control Protocol (RTCP)", BCP 131, RFC 4961, July 2007.

[RFC5405] Eggert, L. and G. Fairhurst, "Unicast UDP Usage Guidelines for Application Designers", BCP 145, RFC 5405, November 2008.

[RFC5761] Perkins, C. and M. Westerlund, "Multiplexing RTP Data and Control Packets on a Single Port", RFC 5761, April 2010.

11.2. Informative References

[RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, June 2002.

[RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, June 2002.

[RFC3389] Zopf, R., "Real-time Transport Protocol (RTP) Payload for Comfort Noise (CN)", RFC 3389, September 2002.

[RFC3556] Casner, S., "Session Description Protocol (SDP) Bandwidth Modifiers for RTP Control Protocol (RTCP) Bandwidth", RFC 3556, July 2003.

[RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, March 2004.

[RFC4103] Hellstrom, G. and P. Jones, "RTP Payload for Text Conversation", RFC 4103, June 2005.

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

[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, July 2006.

[RFC4787] Audet, F., Ed., and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP", BCP 127, RFC 4787, January 2007.

[RFC5124] Ott, J. and E. Carrara, "Extended Secure RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/SAVPF)", RFC 5124, February 2008.

[RFC5245] Rosenberg, J., "Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols", RFC 5245, April 2010.

[RFC5382] Guha, S., Ed., Biswas, K., Ford, B., Sivakumar, S., and P. Srisuresh, "NAT Behavioral Requirements for TCP", BCP 142, RFC 5382, October 2008.

[RFC5389] Rosenberg, J., Mahy, R., Matthews, P., and D. Wing, "Session Traversal Utilities for NAT (STUN)", RFC 5389, October 2008.

[RFC5766] Mahy, R., Matthews, P., and J. Rosenberg, "Traversal Using Relays around NAT (TURN): Relay Extensions to Session Traversal Utilities for NAT (STUN)", RFC 5766, April 2010.

Authors' Addresses

Xavier Marjou
France Telecom Orange
2, avenue Pierre Marzin
Lannion 22307
France

EMail: xavier.marjou@orange-ftgroup.com

Aurelien Sollaud
France Telecom Orange
2, avenue Pierre Marzin
Lannion 22307
France

EMail: aurelien.sollaud@orange-ftgroup.com