Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 5, 2012

C. Perkins
University of Glasgow
V. Singh
Aalto University

March 4, 2012

RTP Congestion Control: Circuit Breakers for Unicast Sessions
draft-perkins-avtcore-rtp-circuit-breakers-00

Abstract

The Real-time Transport Protocol (RTP) is widely used for telephony, video conferencing, and telepresence applications. These applications are often used over best-effort UDP/IP networks. If congestion control is not implemented then network congestion will deteriorate the user's multimedia experience. This document does not propose a congestion control algorithm. Instead, it specifies a minimal set of "circuit-breakers". Circuit-breakers are conditions under which an RTP flow should cease to transmit media to protect the network from excessive congestion. It is expected that all RTP applications running on best-effort networks will be able to run without triggering these circuit breakers in normal operation.

Table of Contents

1. Introduction

The Real-time Transport Protocol (RTP) [RFC3550] is widely used in voice-over-IP, video teleconferencing, and telepresence systems. Many of these systems run over best-effort IP networks, and can suffer from packet loss and increased latency due to network congestion. Designing effective RTP congestion control algorithms, to adapt the transmission of RTP-based media to match the available network capacity, while also maintaining the user experience, is a difficult but important problem. Many such congestion control and media adaptation algorithms have been proposed, but to date there is no consensus on the correct approach, or even that a single standard algorithm is desirable.

This memo does not attempt to propose a new RTP congestion control algorithm. Rather, it proposes a minimal set of "circuit breakers"; conditions under which there should be general agreement that an RTP flow is causing serious congestion, and should cease transmission. It is expected that any future standards-track congestion control algorithms for RTP will operate within the envelope defined by this memo.

2. 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 RFC 2119 [RFC2119].

3. Background

We consider congestion control for unicast RTP traffic flows. This is the problem of adapting the transmission of an audio/visual data flow, encapsulated within an RTP transport session, from one sender to one receiver so that it matches the available network bandwidth. Such adaptation must be done in a way that limits the disruption to the user experience caused by both packet loss and excessive rate changes.

Congestion control for unicast RTP traffic can be implemented in one of two places in the protocol stack. One approach is to run the RTP traffic over a congestion controlled transport protocol, for example over TCP, and to adapt the media encoding to match the dictates of the transport-layer congestion control algorithm. This is safe for the network, but may be suboptimal for the media quality unless the transport protocol is designed to support real-time media flows. We do not consider this class of applications further in this memo, as

their network safety is guaranteed by the underlying transport.

Alternatively, RTP flows can be run over a non-congestion controlled transport protocol, for example UDP, performing rate adaptation at the application layer based on RTP Control Protocol (RTCP) feedback. With a well-designed, network-aware, application, this allows highly effective media quality adaptation, but there is potential to disrupt the network operation if the application does not adapt its sending rate in a timely and effective manner. This memo focusses on this class of application.

Congestion control relies on monitoring the delivery of a media flow, and responding to adapt the transmission of that flow when there are signs that the network path is congested. Network congestion may be detected in one of three ways: 1) a receiver may infer the onset of congestion by observing an increase in one-way delay caused by queue build-up within the network; 2) if Explicit Congestion Notification (ECN) [RFC3168] is supported, the network may signal the presence of congestion by marking packets with ECN Congestion Experienced (CE) marks; or 3) in the extreme case, congestion will cause packet loss, which can be detected by observing a gap in the received RTP sequence numbers. Once the onset of congestion is observed, the receiver must send feedback to the sender to indicate that the transmission rate should be reduced. How the sender reduces the transmission rate is highly dependent on the media codec being used, and is outside the scope of this memo.

There are several ways in which a receiver may send feedback to a media sender within the RTP framework:

o The base RTP specification [RFC3550] defines RTCP Reception Report (RR) packets to convey reception quality feedback information, and Sender Report (SR) packets to convey information about the media transmission. RTCP SR packets contain data that can be used to reconstruct media timing at a receiver, along with a count of the total number of octets and packets sent. RTCP RR packets report on the fraction of packets lost in the last reporting interval, the cumulative number of packets lost, the highest sequence number received, and the inter-arrival jitter. The RTCP RR packets also contain timing information that allows the sender to estimate the network round trip time (RTT) to the receivers. RTCP reports are sent periodically, with the reporting interval being determined by the number of participants in the session and a configured session bandwidth estimate. The interval between reports sent from each receiver tends to be on the order of a few seconds on average, and it is randomised to avoid synchronisation of reports from multiple receivers. If a receiver detects problems, the base RTP specification contains no provisions for sending the feedback

report early and must wait until the next scheduled reporting interval.

- The RTCP Extended Reports (XR) [RFC3611] allow reporting of more complex and sophisticated reception quality metrics, but do not change the RTCP timing rules. RTCP extended reports of potential interest for congestion control purposes are 1) the extended packet loss, discard or burst/gap metrics for reacting based on loss patterns; and 2) the end-system delay metrics for delay-based congestion control.

- Rapid feedback about the occurrence of congestion events can be achieved using the Extended RTP Profile for RTCP-Based Feedback (RTP/AVPF) [RFC4585]. This modifies the RTCP timing rules to allow RTCP reports to be sent early, in some cases immediately, provided the average RTCP reporting interval remains unchanged. It also defines new transport-layer feedback messages, including negative acknowledgements (NACKs), that can be used to report on specific congestion events. The use of the RTP/AVPF profile is dependent on signalling, but is otherwise generally backwards compatible, as it keeps the same average RTCP reporting interval as the base RTP specification. The Codec control messages [RFC5104] extend the RTP/AVPF profile with additional feedback message that can be used to influence that way in which rate adaptation occurs. The dynamics of how rapidly feedback can be sent are unchanged.

- Finally, the RTP and RTCP extensions for Explicit Congestion Notification (ECN) [I-D.ietf-avtcore-ecn-for-rtp] can be used to provide feedback on the number of packets that received an ECN Congestion Experienced (CE) mark. This extension builds on the RTP/AVPF profile to allow rapid congestion feedback.

In addition to these mechanisms for providing feedback, the sender can include an RTP header extension in each packet to record packet transmission times. There are two methods: [RFC5450] represents the transmission time in terms of a time-offset from the RTP timestamp of the packet, while [RFC6051] includes an explicit NTP-format sending timestamp (potentially more accurate, but a higher header overhead). Accurate sending timestamps can be helpful for estimating queuing delays, to get an early indication of the onset of congestion.

Taken together, these various mechanisms allow receivers to provide feedback on the senders when congestion events occur, with varying degrees of timeliness and accuracy. The key distinction is between systems that use only the basic RTCP mechanisms, without RTP/AVPF rapid feedback, and those that use the RTP/AVPF extensions, and can respond to congestion more rapidly.

4. RTP Circuit Breakers for Systems Using the RTP/AVP Profile

The feedback mechanisms defined in [RFC3550] are the minimum that can be required for a baseline circuit breaker mechanism suitable for all unicast applications of RTP. Accordingly, for an RTP circuit breaker to be useful, it should be able to detect that an RTP flow is causing excessive congestion using only basic RTCP features, without needing RTCP XR feedback or the RTP/AVPF profile for rapid RTCP reports.

Three potential congestion signals are available from the basic RTCP SR/RR packets, and are reported for each SSRC in the RTP session:

1. The sender can estimate the network round-trip time once per RTCP reporting interval, based on the contents and timing of RTCP SR and RR packets.

2. Receivers report estimated jitter (the statistical variance of the RTP data packet inter-arrival time) calculated over the RTCP reporting interval. Due to the nature of the jitter calculation ([RFC3550], section 6.4.4), the jitter is only meaningful for RTP flows that send a single data packet for each RTP timestamp value (i.e., audio flows, or video flows where each frame comprises one RTP packet).

3. Receivers report the fraction of packets lost during the RTCP reporting interval, and the cumulative number of packets lost over the entire RTP session.

These congestion signals limit the possible circuit breakers, since they give only limited visibility into the behaviour of the network.

RTT estimates are widely used in congestion control algorithms, as a proxy for queuing delay measures in delay-based congestion control, or to determine connection timeouts. RTT estimates derived from RTCP SR and RR packets send according to the RTP/AVP timing rules are too infrequent to be useful though, and don't give enough information to distinguish a delay change due to routing updates from queuing delay caused by congestion. Accordingly, we do not use the RTT estimate alone as an RTP circuit breaker.

Increased jitter can be a signal of transient network congestion, but in the highly aggregated form reported in RTCP RR packets, it offers insufficient information to estimate the extent or persistence of congestion. Jitter reports are a useful early warning of potential network congestion, but provide an insufficiently strong signal to be used as a circuit breaker.

The remaining congestion signals are the packet loss fraction and the

cumulative number of packets lost. These are robust indicators of congestion in networks where packet loss is primarily due to queue overflows, although less accurate in networks where losses can be caused by non-congestive packet corruption. TCP also uses packet loss as a congestion signal.

Two packet loss regimes can be observed: 1) RTCP RR packets show a non-zero packet loss fraction, while the extended highest sequence number received continues to increment; and 2) RR packets show a loss fraction of zero, but the extended highest sequence number received does not increment even though the sender has been transmitting RTP data packets. The former corresponds to the TCP congestion avoidance state, and indicates a congested path that is still delivering data; the latter corresponds to a TCP timeout, and is most likely due to a path failure. We derive circuit breaker conditions for these two loss regimes.

4.1. RTP/AVP Circuit Breaker #1: Timeout

If RTP data packets are being sent while the corresponding RTCP RR packets report a non-increasing extended highest sequence number received, this is an indication that those RTP data packets are not reaching the receiver. This could be a short-term issue affecting only a few packets, perhaps caused by a slow-to-open firewall or a transient connectivity problem, but if the issue persists, it is a sign of a more ongoing and significant problem. Accordingly, if a sender of RTP data packets receives two or more consecutive RTCP RR packets from the same receiver that correspond to its transmission and have a non-increasing extended highest sequence number received field, then that sender SHOULD cease transmission.

Systems that usually send at a high data rate, but which can reduce their data rate significantly (i.e., by an order of magnitude), MAY first reduce their sending rate to this lower value, but MUST then cease transmission if the problem does not resolve itself within a further two RTCP reporting intervals. An example of this might be a video conferencing system that backs off to sending audio only, before completely dropping the call.

The choice of two RTCP reporting intervals is to give enough time for transient problems to resolve themselves, but to stop problem flows quickly enough to avoid causing serious problems. A single RTCP report showing no reception could be caused by numerous transient faults, and so should not stop transmission. More than two RTCP reports could avoid false positives, but would lead to problematic flows running for a long time before being cut off.

4.2. RTP/AVP Circuit Breaker #2: Congestion

If RTP data packets are being sent, and the corresponding RTCP RR packets show non-zero packet loss fraction and increasing extended highest sequence number received, then the RTP data packets are arriving at the receiver, but some degree of congestion is occurring. The RTP/AVP profile [RFC3551] states that:

> If best-effort service is being used, RTP receivers SHOULD monitor packet loss to ensure that the packet loss rate is within acceptable parameters. Packet loss is considered acceptable if a TCP flow across the same network path and experiencing the same network conditions would achieve an average throughput, measured on a reasonable timescale, that is not less than the RTP flow is achieving. This condition can be satisfied by implementing congestion control mechanisms to adapt the transmission rate (or the number of layers subscribed for a layered multicast session), or by arranging for a receiver to leave the session if the loss rate is unacceptably high.
>
> The comparison to TCP cannot be specified exactly, but is intended as an "order-of-magnitude" comparison in timescale and throughput. The timescale on which TCP throughput is measured is the round-trip time of the connection. In essence, this requirement states that it is not acceptable to deploy an application (using RTP or any other transport protocol) on the best-effort Internet which consumes bandwidth arbitrarily and does not compete fairly with TCP within an order of magnitude.

The throughput of a long-lived TCP connection can be estimated using the TCP throughput equation:

$$X = \frac{s}{R*\mathrm{sqrt}(2*b*p/3) + (t\_RTO * (3*\mathrm{sqrt}(3*b*p/8) * p * (1+32*p^2)))}$$

Where:

X is the transmit rate in bytes/second.

s is the packet size in bytes. If the RTP data packets vary in size, then the average size should be used.

R is the round trip time in seconds.

p is the loss event rate, between 0 and 1.0, of the number of loss events as a fraction of the number of packets transmitted.

t_RTO is the TCP retransmission timeout value in seconds, approximated by setting t_RTO = 4*R.

b is the number of packets acknowledged by a single TCP acknowledgement ([RFC3448] recommends the use of b=1 since many TCP implementations do not use delayed acknowledgements).

This is the same approach to estimated TCP throughput that is used in [RFC3448]. Two parameters must be estimated in order to calculate the throughput: the round trip time, R, and the loss event rate, p. The round trip time can be estimated from RTCP. This is done too infrequently for accurate statistics, but is the best that can be done with the standard RTCP mechanisms.

RTCP RR packets contain the packet loss fraction, rather than the loss event rate, so p cannot be reported (TCP typically treats the loss of multiple packets within a single RTT as one loss event, but RTCP RR packets report the overall fraction of packets lost, not caring about when the losses occurred). Using the loss fraction in place of the loss event rate can overestimate the loss. We believe that this overestimate will not be significant, given that we are only interested in order of magnitude comparison (Floyd et al, "Equation-Based Congestion Control for Unicast Applications", Proc. SIGCOMM 2000, section 3.2.1, show that the difference is small for steady-state conditions and random loss, but using the loss fraction is more conservative in the case of bursty loss).

The congestion circuit breaker is therefore: when RTCP RR packets are received, estimate the TCP throughput using the above equation and the measured R, p (approximated by the loss fraction), and s. Compare this with the actual sending rate. If the actual sending rate has been more than an order of magnitude greater than the throughput equation estimate for two or more RTCP reporting intervals, stop transmitting.

Again, we use two reporting intervals to avoid triggering the circuit breaker on transient failures. This circuit breaker is a worst-case condition, and congestion control should be performed to keep well within this bound. It is expected that the circuit breaker will only be triggered if the usual congestion control fails for some reason.

(tbd -- we need to base the circuit breaker condition on something, so TCP seems a logical choice. Following TCP limits too closely is inappropriate for many applications of RTP, though, since they have different dynamics. Is the above lax enough to not disrupt valid

applications, but tight enough to provide meaningful protection for the network?)

5. RTP Circuit Breakers for Systems Using the RTP/AVPF Profile

More rapid feedback allows more responsiveness. The receiver SHOULD provide feedback more often during, or at onset of, congestion, and provide feedback less often when there is no congestion.

(tbd -- mechanisms may probably need to be designed in conjunction with the different classes of congestion control that can leverage RTP/AVPF; e.g., we might need to specify limits for TFRC-like or delay-based algorithms using RTP/AVPF feedback.)

(tbd -- a high-level question to be answered is whether we need to specify anything different for the circuit breaker for AVPF, or if we leave that unchanged, and focus solely on the dynamics, to ensure the circuit breaker is never triggered.)

6. Impact of RTCP XR

(tbd)

This improves the information, but doesn't change the dynamics of the congestion control loop. Suspect the impact will actually be quite small.

Packets discarded [I-D.ietf-xrblock-rtcp-xr-discard] or bytes discarded [I-D.ietf-xrblock-rtcp-xr-discard-rle-metrics] due to late arrival by the receiver may indicate congestion. Congestion control should consider the discarded packets as if they were lost packets.

The RTCP RR reports the loss fraction over an RTCP interval which is insufficient to distinguish between solitary or bursty losses. To provide rough sense of duration of losses or discards, an endpoint may use burst/gap reporting for loss [I-D.ietf-xrblock-rtcp-xr-burst-gap-loss] and discard [I-D.ietf-xrblock-rtcp-xr-burst-gap-discard]. For more accurate reporting the receiver may use Run-length encoded (RLE) lost [RFC3611] or discarded [I-D.ietf-xrblock-rtcp-xr-discard-rle-metrics] packets.

For precise measurement of network roundtrip delay the receiver can signal its end-system delay [I-D.ietf-xrblock-rtcp-xr-delay] [RFC3611].

A receiver may also indicate onset or end of congestion by reporting the distribution of the inter-packet delay variation [I-D.ietf-xrblock-rtcp-xr-pdv] [RFC3611].

7. Impact of Explicit Congestion Notification (ECN)

ECN-CE marked packets SHOULD be treated as if it were lost for the purposes of congestion control, when determining the optimal rate at which to send. However, it seems unwise to treat the receipt of multiple ECN-CE marked packets as a circuit breaker, since it is likely that ECN-capable and non-ECN-capable paths will exist for a long time to come. Rather, consider packet loss as the circuit breaker condition as for non-ECN flows.

8. Session Timeout

From a usability perspective, if there is no audio or video response from the other peer, it is likely that the user may terminate the session.

According to RFC 3550 [RFC3550], any participant that has not sent an RTP packet within the last two RTCP interval is removed from the sender list. To avoid timing out the specific flow, the endpoint MUST send corresponding RTCP reports. Interactive Connectivity Establishment (ICE) [RFC5245] recommends that the timeout MUST NOT be less than 15 seconds.

If no RTCP RR arrives for two complete SR intervals, the sender SHOULD cease transmission. However, if the endpoint can reduce the media rate then it MAY first reduce the rate to the lower value, but terminate the transmission if still no RTCP RR is received in the next two SR intervals.

9. References

9.1. Normative References

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

[RFC3448] Handley, M., Floyd, S., Padhye, J., and J. Widmer, "TCP Friendly Rate Control (TFRC): Protocol Specification", RFC 3448, January 2003.

[RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V.

Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.

[RFC3551] Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video Conferences with Minimal Control", STD 65, RFC 3551, July 2003.

[RFC3611] Friedman, T., Caceres, R., and A. Clark, "RTP Control Protocol Extended Reports (RTCP XR)", RFC 3611, November 2003.

[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.

9.2. Informative References

[I-D.ietf-avtcore-ecn-for-rtp]
Westerlund, M., Johansson, I., Perkins, C., and K. Carlberg, "Explicit Congestion Notification (ECN) for RTP over UDP", draft-ietf-avtcore-ecn-for-rtp-06 (work in progress), February 2012.

[I-D.ietf-xrblock-rtcp-xr-burst-gap-discard]
Clark, A., Hunt, G., Wu, W., and R. Huang, "RTCP XR Report Block for Burst/Gap Discard metric Reporting", draft-ietf-xrblock-rtcp-xr-burst-gap-discard-02 (work in progress), January 2012.

[I-D.ietf-xrblock-rtcp-xr-burst-gap-loss]
Clark, A., Hunt, G., Zhao, J., Wu, W., and S. Zhang, "RTCP XR Report Block for Burst/Gap Loss metric Reporting", draft-ietf-xrblock-rtcp-xr-burst-gap-loss-01 (work in progress), January 2012.

[I-D.ietf-xrblock-rtcp-xr-delay]
Hunt, G., Gross, K., and A. Clark, "RTCP XR Report Block for Delay metric Reporting", draft-ietf-xrblock-rtcp-xr-delay-01 (work in progress), December 2011.

[I-D.ietf-xrblock-rtcp-xr-discard]
Hunt, G., Clark, A., Zorn, G., and W. Wu, "RTCP XR Report Block for Discard metric Reporting", draft-ietf-xrblock-rtcp-xr-discard-01 (work in progress), December 2011.

[I-D.ietf-xrblock-rtcp-xr-discard-rle-metrics]
Ott, J., Singh, V., and I. Curcio, "Real-time Transport Control Protocol (RTCP) Extension Report (XR) for Run Length Encoding of Discarded Packets", draft-ietf-xrblock-rtcp-xr-discard-rle-metrics-03 (work in progress), February 2012.

[I-D.ietf-xrblock-rtcp-xr-pdv]
Hunt, G. and A. Clark, "RTCP XR Report Block for Packet Delay Variation Metric Reporting", draft-ietf-xrblock-rtcp-xr-pdv-02 (work in progress), December 2011.

[RFC3168] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of Explicit Congestion Notification (ECN) to IP", RFC 3168, September 2001.

[RFC5104] Wenger, S., Chandra, U., Westerlund, M., and B. Burman, "Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF)", RFC 5104, 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.

[RFC5450] Singer, D. and H. Desineni, "Transmission Time Offsets in RTP Streams", RFC 5450, March 2009.

[RFC6051] Perkins, C. and T. Schierl, "Rapid Synchronisation of RTP Flows", RFC 6051, November 2010.

Authors' Addresses

Colin Perkins
University of Glasgow
School of Computing Science
Glasgow G12 8QQ
United Kingdom

Email: csp@csperkins.org

Varun Singh
Aalto University
School of Science and Technology
Otakaari 5 A
Espoo, FIN 02150
Finland

Email: varun@comnet.tkk.fi
URI: http://www.netlab.tkk.fi/~varun/