The QUIC Loss Bits
draft-ferrieux-hamchaoui-quic-lossbits-00

Abstract

   This document specifies the addition of loss bits to the QUIC
   transport protocol and describes how to use them to measure and
   locate packet loss.

Note to Readers

   This document specifies an experimental delta to the QUIC transport
   protocol.

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 http://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 October 11, 2019.

Table of Contents

1. Introduction

Packet loss is a hard and pervasive problem of day-to-day network
operation, and locating them is crucial to timely resolution of
crippling end-to-end throughput issues. To this effect, in a TCP-
dominated world, network operators have been heavily relying on
information present in clear in TCP headers: sequence and
acknowledgement numbers, and SACK when enabled. By passive on-path
observation, these allow for quantitative estimation of packet loss.
Additionally, the lossy segment (upstream or downstream from the
observation point) is unambiguous; this is crucial as it gives the
ability to quickly home in on the offending segment, by moving the
passive observer around.

In the QUIC context, the equivalent transport headers being
encrypted, such observation is not possible. To restore network
operators' ability to maintain QUIC clients experience, this document
adds two explicit loss bits to the QUIC short header, named "Q"
(sQuare signal) and "R" (Retransmit). Together, these bits allow the
observer to estimate upstream and downstream loss, enabling the same
dichotomic search as with TCP.

[2](#). Passive Loss measurement

The proposed mechanisms enable loss measurement from observation
points on the network path throughout the lifetime of a connection.
End-to end loss as well as segmental loss (upstream or downstream
from the observation point) are measurable thanks to two dedicated
bits in short packet headers, named loss bits. The loss bits
therefore appear only after version negotiation and connection
establishment are completed.

[2.1](#). Proposed Short Header Format Including Loss Bits

As of the current editor's version of [[QUIC-TRANSPORT](#)], two bits are
"reserved" in the first byte of short headers. This proposal
naturally fits in there, allocating these two bits as Q and R. Of
course, the very purpose of Q and R being to enable on-path
observation, the current restrictions about their encryption and zero
value should be lifted in QUIC versions supporting this proposal.

[2.2](#). Semantics

The semantics of these bits are as follows:

Q: The sQuare bit is toggled every N outgoing packets as explained
below in [Section 2.2.1](#).

R: The Retransmit bit is set to 0 or 1 according to the not-yet-
disclosed-lost-packets counter, as explained below in [Section 2.2.2](#).

[2.2.1](#). Setting the sQuare Bit on Outgoing Packets

Each endpoint independently maintains a sQuare value, 0 or 1, during
a block of N outgoing packets (e.g. N=64), and sets the sQuare bit
in the short header to the currently stored value when a packet with
a short header is sent out. The sQuare value is initiated to 0 at
each endpoint, client and server, at connection start. This
mechanism thus delineates slots of N packets with the same marking.
Observation points can estimate the upstream losses by simply
counting the number of packets during a half period of the square
signal, as described in [Section 3](#).

[2.2.2](#). Setting the Retransmit Bit on Outgoing Packets

Each endpoint, client and server, independently maintains a not-yet-
disclosed-lost-packets counter and sets the Retransmit bit of short
header packets to 0 or 1 accordingly. The not-yet-disclosed-lost-
packets counter is initialized to 0 at each endpoint, client and
server, at connection start, and reflects packets considered lost by

the QUIC machinery, the content of which is pending for
retransmission. When a packet is declared lost by the QUIC
retransmission machinery (see [QUIC-RECOVERY]) the not-yet-disclosed-
lost-packets counter is incremented by 1. When a packet with a short
header is sent out by an end-point, its retransmit bit is set to 0
when the not-yet-disclosed-lost-packets counter is equal to 0.
Otherwise, the packet is sent out with a retransmit bit set to 1 and
the not-yet-disclosed-lost-packets counter is decremented by 1.
Thus, the retransmit bit performs unary encoding of the amount of
loss: observation points can estimate the number of packets
considered lost by the QUIC transmission machinery in a given
direction by counting packets in this direction with a retransmit bit
equal to 1.

2.2.3. Resetting state on CID change

When sending the first packet of a given connection with a new
connection ID, each endpoint resets its sQuare value and not-yet-
disclosed-lost-packets counter to zero. This eliminates the
possibility for transient sQuare or Retransmit bit state to be used
to link flows across connection migration or ID change.

3. Using the loss bits for Passive Loss Measurement

3.1. End-to-end loss

The Retransmit bit mechanism merely reflects the number of packets
considered lost by the sender QUIC stack with a slight delay. In
case of fast retransmit due to repeted acknowlegments of a packet,
this delay is at least equal to the one way delay in the reverse
direction. It is larger otherwise (eg RTO). The retransmit
mechanism alone suffices to estimate the end-to-end losses; similar
to TCP passive loss measurement, its accuracy depends on the loss
affecting the retransmit-bit-marked packets, which are in themselves
proof of previous loss.

3.2. Upstream loss

During a QUIC connection lifetime, the sQuare bit mechanism
delineates slots of N packets with the same marking. When focusing
on the sQuare bit of consecutive packets in a direction, this
mechanism sketches a periodic square signal which toggles every N
packets. On-path observers can then estimate the upstream losses by
simply counting the number of packets during a half period (level 0
or level 1) of the square signal. Packets with a long header are not
marked, but yet taken into account by the sender when counting the N
outgoing packets before its next toggle. Observers should assign
long header packets to the pending slot if possible (i.e. up to N

packets counted in this slot), to the next one otherwise. Thus,
slots with less than N packets, whatever their header length,
generally denote upstream loss. As with TCP passive detection based
on missing sequence numbers, this estimation may become inaccurate in
case of packet reordering which blurs the edges of the square signal
; heuristics may be proposed to filter out this noise in the
observation points.

The slot size N should be carefully chosen : too short, it becomes
very sensitive to packet reordering and loss. Too large, short
connections may end before completion of the first square slot,
ruining any loss estimation. Slots of 64 packets are suggested as a
reasonable trade-off.

3.3. Downstream loss

The Retransmit bit mechanism can be coupled with the sQuare bit
mechanism to estimate downstream losses. Indeed, passive observers
can infer downstream losses by difference between end-to-end and
upstream losses.

The sQuare bit mechanism allows for observers to compute loss
measurement at the end of every half square signal period (level 0 or
level 1).

The Retransmit bit mechanism provides for the end-to-end loss after
reaction of the sender stack.

On-path observers can estimate upstream and downstream loss at
various scales, from the square slot level to the connection lifetime
level.

Note that observers should perform a loose synchronisation between
the sQuare and the Retransmit measurements when accurate evolution of
segmental loss over connection lifetime is sought, so as to compare
the same portion of the packet stream.

3.4. Bidirectional flows

The Q and R bits sent by one endpoint cover loss of packets sent by
the same endpoint, allowing a midpoint observer to estimate loss in
that direction; no specific cooperation is needed between the
endpoints beyond negotiating a QUIC version that supports this
proposal. Hence, the server will be enabling troubleshooting of the
download path, and the client will work for the upload path. This
allows to be confident about getting a useful signal in asymmetric
situations: clients may for example implement Q and R improperly, the

download path will still be debuggable as long as servers do it
right.

It should also be noted that the method does not suffer from the
natural asymmetry in packet rate of a typical download or upload
scenario. Indeed, although there are often fewer acknowledgements
than payload-bearing packets, the unary encoding by R of payload loss
is borne by the payload stream itself. This allows to report loss in
the important direction in both a timely and accurate fashion without
sampling or quantization.

4. Security and Privacy Considerations

The loss bits are intended to expose loss to observers along the
path, so the privacy considerations for the loss bits are essentially
the same as those for passive loss measurement in general. Loss
gives no hint on customer geolocalisation; moreover, reset of loss
accounting state on CID changes prevents linkability.

5. IANA Considerations

An IANA registry has been suggested for QUIC versions. In support of
the fully negotiated status of the proposed extension, a natural way
of deploying this feature would be through such a registered version.

6. Acknowledgments

The sQuare Bit was originally specified by Kazuho Oku in early
proposals for loss measurement.

7. Normative References

[QUIC-RECOVERY]
          Iyengar, J., Ed. and I. Swett, Ed., "QUIC Loss Detection
          and Congestion Control".

[QUIC-TRANSPORT]
          Iyengar, J., Ed. and M. Thomson, Ed., "QUIC: A UDP-Based
          Multiplexed and Secure Transport", draft-ietf-quic-
          transport-latest (work in progress).

Authors' Addresses

Alexandre Ferrieux (editor)
Orange Labs

Email: alexandre.ferrieux@orange.com

      Isabelle Hamchaoui (editor)
      Orange Labs

      Email: isabelle.hamchaoui@orange.com