QUIC
Internet-Draft
Intended status: Experimental
Expires: April 26, 2019

B. Trammell, Ed.
M. Kuehlewind
ETH Zurich

October 23, 2018

The QUIC Latency Spin Bit
draft-ietf-quic-spin-exp-01

Abstract

This document specifies the addition of a latency spin bit to the QUIC transport protocol and describes how to use it to measure end-to-end latency.

Note to Readers

This document specifies an experimental delta to the QUIC transport protocol. Specifically, this experimentation is intended to determine:

- the impact of the addition of the latency spin bit on implementation and specification complexity; and

- the accuracy and value of the information derived from spin bit measurement on live network traffic.

The information generated by this experiment will be used by the QUIC working group as input to a decision about the standardization of the latency spin bit. Although this is a Working Group document, it is currently NOT a Working Group deliverable.

Discussion of this draft takes place on the QUIC working group mailing list (quic@ietf.org), which is archived at https://mailarchive.ietf.org/arch/search/?email_list=quic [1].

Working Group information can be found at https://github.com/quicwg [2]; source code and issues list for this draft can be found at https://github.com/quicwg/base-drafts/labels/-spin [3].

Status of This Memo

Table of Contents

1. Introduction

The QUIC transport protocol [QUIC-TRANSPORT] uses Transport Layer Security (TLS) [TLS] to encrypt most of its protocol internals. In contrast to TCP where the sequence and acknowledgement numbers and timestamps (if the respective option is in use) can be seen by on-path observers and used to estimate end-to-end latency, QUIC's wire image (see [WIRE-IMAGE]) currently does not expose any information that can be used for passive latency measurement techniques that rely on this information (e.g. [CACM-TCP], [TMA-QOF]).

This document adds an explicit signal for passive latency measurability to the QUIC short header: a "spin bit". Passive observation of the spin bit provides one RTT sample per RTT to passive observers of QUIC traffic. This document describes the mechanism, how it can be added to QUIC, and how it can be used by passive measurement facilities to generate RTT samples.

2. The Spin Bit Mechanism

The latency spin bit enables latency monitoring from observation points on the network path throughout the duration of a connection. Since it is possible to measure handshake RTT without a spin bit, it is sufficient to include the spin bit in the short packet header. The spin bit therefore appears only after version negotiation and connection establishment are completed.

2.1. Proposed Short Header Format Including Spin Bit

As of the current editor's version of [QUIC-TRANSPORT], this proposal specifies using the sixth most significant bit (0x04) of the first octet in the short header for the spin bit.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+
```

```
|0|K|1|1|0|S|R R|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                Destination Connection ID (0..144)           ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Packet Number (8/16/32)                ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Protected Payload (*)                   ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: Short Header Format including proposed Spin Bit

S: The Spin bit is set 0 or 1 depending on the stored spin value that is updated on packet reception as explained in Section 2.2.

R: Two additional bits are reserved for experimentation in the short header.

2.2. Setting the Spin Bit on Outgoing Packets

Each endpoint, client and server, maintains a spin value, 0 or 1, for each QUIC connection, and sets the spin bit in the short header to the currently stored value when a packet with a short header is sent out. The spin value is initialized to 0 at each endpoint, client and server, at connection start. Each endpoint also remembers the highest packet number seen from its peer on the connection.

The spin value is then determined at each endpoint within a single connection as follows:

- When the server receives a packet from the client, if that packet has a short header and if it increments the highest packet number seen by the server from the client, the server sets the spin value to the value observed in the spin bit in the received packet.

- When the client receives a packet from the server, if the packet has a short header and if it increments the highest packet number seen by the client from the server, it sets the spin value to the opposite of the spin bit in the received packet.

This procedure will cause the spin bit to change value in each direction once per round trip. Observation points can estimate the network latency by observing these changes in the latency spin bit, as described in Section 3. See [QUIC-SPIN] for further illustration of this mechanism in action.

2.3. Resetting Spin Value State

Each client and server resets it spin value to zero when sending the first packet of a given connection with a new connection ID. This reduces the risk that transient spin bit state can be used to link flows across connection migration or ID change.

3. Using the Spin Bit for Passive RTT Measurement

When a QUIC flow sends data continuously, the latency spin bit in each direction changes value once per round-trip time (RTT). An on-path observer can observe the time difference between edges (changes from 1 to 0 or 0 to 1) in the spin bit signal in a single direction to measure one sample of end-to-end RTT.

An observer can store the largest observed packet number per flow, and reject edges that do not have a monotonically increasing packet number (greater than the largest observed packet number). This will avoid detecting spurious edges caused by reordering events that include an edge, which would lead to very low RTT estimates if not ignored.

If the spin bit edge occurs after a long packet number gap, it should be ignored: this filters out high RTT estimates due to loss of an actual edge in a burst of lost packets.

Note that this measurement, as with passive RTT measurement for TCP, includes any transport protocol delay (e.g., delayed sending of acknowledgements) and/or application layer delay (e.g., waiting for a request to complete). It therefore provides devices on path a good instantaneous estimate of the RTT as experienced by the application. A simple linear smoothing or moving minimum filter can be applied to the stream of RTT information to get a more stable estimate.

However, application-limited and flow-control-limited senders can have application and transport layer delay, respectively, that are much greater than network RTT. When the sender is application-limited and e.g. only sends small amount of periodic application traffic, where that period is longer than the RTT, measuring the spin bit provides information about the application period, not the network RTT.

Simple heuristics based on the observed data rate per flow or changes in the RTT series can be used to reject bad RTT samples due to application or flow control limitation; for example, QoF [TMA-QOF] rejects component RTTs significantly higher than RTTs over the history of the flow. These heuristics may use the handshake RTT as an initial RTT estimate for a given flow.

An on-path observer that can see traffic in both directions (from client to server and from server to client) can also use the spin bit to measure "upstream" and "downstream" component RTT; i.e, the component of the end-to-end RTT attributable to the paths between the observer and the server and the observer and the client, respectively. It does this by measuring the delay between a spin edge observed in the upstream direction and that observed in the downstream direction, and vice versa.

4. IANA Considerations

This document has no actions for IANA.

5. Security and Privacy Considerations

The spin bit is intended to expose end-to-end RTT to observers along the path, so the privacy considerations for the latency spin bit are essentially the same as those for passive RTT measurement in general. It has been shown [PAM-RTT] that RTT measurements do not provide more information for geolocation than is available in the most basic, freely-available IP address based location databases. The risk of exposure of per-flow network RTT to on-path devices is therefore negligible.

6. Change Log

RFC Editor's Note: Please remove this section prior to publication of a final version of this document.

6.1. Since draft-ietf-spin-exp-00

Nothing yet.

Acknowledgments

This document is derived from [QUIC-SPIN], which was the work of the following authors in addition to the editor of this document:

- o Piet De Vaere, ETH Zurich
- o Roni Even, Huawei
- o Giuseppe Fioccola, Telecom Italia
- o Thomas Fossati, Nokia
- o Marcus Ihlar, Ericsson
- o Al Morton, AT&T Labs
- o Emile Stephan, Orange

The QUIC Spin Bit was originally specified in a slightly different form by Christian Huitema.

This work is partially supported by the European Commission under Horizon 2020 grant agreement no. 688421 Measurement and Architecture for a Middleboxed Internet (MAMI), and by the Swiss State Secretariat for Education, Research, and Innovation under contract no. 15.0268. This support does not imply endorsement.

8. References

8.1. Normative References

[QUIC-TRANSPORT]
Iyengar, J., Ed. and M. Thomson, Ed., "QUIC: A UDP-Based Multiplexed and Secure Transport", draft-ietf-quic-transport-15 (work in progress), October 2018.

8.2. Informative References

[CACM-TCP]
Strowes, S., "Passively Measuring TCP Round-Trip Times (in Communications of the ACM)", October 2013.

[PAM-RTT] Trammell, B. and M. Kuehlewind, "Revisiting the Privacy Implications of Two-Way Internet Latency Data (in Proc. PAM 2018)", March 2018.

[QUIC-SPIN]
Trammell, B., Vaere, P., Even, R., Fioccola, G., Fossati, T., Ihlar, M., Morton, A., and S. Emile, "Adding Explicit Passive Measurability of Two-Way Latency to the QUIC Transport Protocol", draft-trammell-quic-spin-03 (work in progress), May 2018.

[TLS] Rescorla, E., "The Transport Layer Security (TLS) Protocol Version 1.3", RFC 8446, DOI 10.17487/RFC8446, August 2018, <https://www.rfc-editor.org/info/rfc8446>.

[TMA-QOF] Trammell, B., Gugelmann, D., and N. Brownlee, "Inline Data Integrity Signals for Passive Measurement (in Proc. TMA 2014)", April 2014.

[WIRE-IMAGE]
Trammell, B. and M. Kuehlewind, "The Wire Image of a Network Protocol", draft-trammell-wire-image-04 (work in progress), April 2018.

8.3. URIs

[1] https://mailarchive.ietf.org/arch/search/?email_list=quic

[2] https://github.com/quicwg

[3] https://github.com/quicwg/base-drafts/labels/-spin

Authors' Addresses

Brian Trammell (editor)
ETH Zurich

Email: ietf@trammell.ch

Mirja Kuehlewind
ETH Zurich

Email: mirja.kuehlewind@tik.ee.ethz.ch