

QUIC
Internet-Draft
Intended status: Informational
Expires: February 21, 2019

B. Trammell
ETH Zurich
August 20, 2018

Why do we need passive measurement of round trip time?
draft-trammell-why-measure-rtt-00

Abstract

This document describes the utility of passive two-way latency measurement, both for the generation of latency metrics, as well as for other measurement tasks, when passive latency measurement is the only facility available for measurement. It additionally discusses other metrics derivable from the transport-independent latency spin signal defined in [[TSVWG-SPIN](#)].

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 <https://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 February 21, 2019.

Copyright Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in [Section 4](#).e of

the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	2
1.1.	About This Document	3
2.	Direct Utility of Passive RTT Measurement	3
2.1.	Inter-domain Troubleshooting	3
2.2.	Bufferbloat Mitigation in Cellular Networks	4
2.3.	Locating WiFi Problems in Home Networks	4
2.4.	Internet Measurement Research	5
3.	Indirect Utility of RTT Measurements	5
4.	Additional Metrics Derivable from the Spin Bit	6
4.1.	Derived Loss and Reordering	6
4.2.	Two-Point Intradomain Measurement	7
5.	Contributors	7
6.	Acknowledgments	8
7.	References	8
7.1.	Normative References	8
7.2.	Informative References	8
	Author's Address	9

[1.](#) Introduction

Latency is a key metric to understanding network operation and performance. Passive measurement allows inspection of latency on productive traffic, avoiding problems with different treatment of productive and measurement traffic, and enables opportunistic measurement of latency without active measurement overhead.

Passive measurement of RTT, in particular, has both direct utility (see [Section 2](#)), generating RTT samples for the measurement of RTT for various use cases, as well as indirect utility (see [Section 3](#)), since RTT is correlated with other useful metrics. In addition, the passive latency signal proposed in [\[TSVWG-SPIN\]](#) provides other opportunities for metric generation which are a consequence of its design (see [Section 4](#)).

This document describes these use cases in order to motivate why passive measurability of RTT on a per-flow basis is an interesting and useful feature for a transport protocol to have. In the absence of other directly observable metrics such as loss and retransmission, as is the case with protocols with mostly-encrypted wire images [\[WIRE-IMAGE\]](#) such as QUIC [\[QUIC\]](#).

Trammell

Expires February 21, 2019

[Page 2]

1.1. About This Document

This document is maintained in the GitHub repository <https://github.com/britram/draft-trammell-tsvwg-spin>, and the editor's copy is available online at <https://britram.github.io/draft-trammell-tsvwg-spin>. Current open issues on the document can be seen at <https://github.com/britram/draft-trammell-tsvwg-spin/issues>. Comments and suggestions on this document can be made by filing an issue there, or by contacting the editor.

This document is based in part on [QUIC-SPIN], however, aside from [Section 4](#), it is not specific to the spin bit proposal.

2. Direct Utility of Passive RTT Measurement

RTT measurement generates two-way latency metric samples; these samples are useful in many measurement tasks which directly require latency data. The measurement methodologies using two-way latency measurement samples follow one of a few basic variants:

- o The RTT evolution of a flow or a set of flows can be compared to baseline or expected RTT measurements for flows with the same characteristics in order to detect or localize latency issues in a specific network.
- o The RTT evolution of a single flow can also be examined in detail to diagnose performance issues with that flow.
- o Samples of RTT for a flow aggregate (e.g., all flows between two given networks) can be used without regard to temporal evolution of the RTT, in order to examine the distribution of RTTs for a group of flows that should have similar RTT (e.g., because they should share the same path(s)).

2.1. Inter-domain Troubleshooting

Network access providers are often the first point of contact by their customers when network problems impact the performance of bandwidth-intensive and latency-sensitive applications such as video, regardless of whether the root cause lies within the access provider's network, the service provider's network, on the Internet paths between them, or within the customer's own network.

Points on path can extract spatial delay metric samples [RFC6049] from fields of the transport layer (e.g. TCP) or application layer (e.g. RTP). The information is captured in the upper layer because neither the IP header nor the UDP layer includes fields allowing the measurement of upstream and downstream delay.

Local network performance problems are detected with monitoring tools which observe the variation of upstream latency and downstream latency.

Inter-domain troubleshooting relies on the same metrics but is not a proactive task; instead, it is a recursive process which hones in on the domain and link responsible for the failure. In practice, inter-domain troubleshooting is a communication process between the Network Operations Center (NOC) teams of the networks on the path, because the root cause of a problem is rarely located on a single network, and requires cooperation and exchange of data between the NOCs.

One example is the troubleshooting performance degradation resulting from a change of routing policy on one side of the path which increases queueing on the other side of the path.

2.2. Bufferbloat Mitigation in Cellular Networks

Cellular networks consist of multiple Radio Access Networks (RAN) where mobile devices are attached to base stations. It is common that base stations from different vendors and different generations are deployed in the same cellular network.

Due to the dynamic nature of RANs, base stations have typically been provisioned with large buffers to maximize throughput despite rapid changes in capacity. As a side effect, bufferbloat has become a common issue in such networks [[WMM-BLOAT](#)].

An effective way of mitigating bufferbloat without sacrificing too much throughput is to deploy Active Queue Management (AQM) in bottleneck routers and base stations. However, due to the variation in deployed base-stations it is not always possible to enable AQM at the bottlenecks, without massive infrastructure investments.

An alternative approach is to deploy AQM as a network function in a more centralized location than the traditional bottleneck nodes. Such an AQM monitors the RTT progression of flows and drops or marks packets when the measured latency is indicative of congestion. Such a function also has the possibility to detect misbehaving flows and reduce the negative impact they have on the network.

2.3. Locating WiFi Problems in Home Networks

Many residential networks use WiFi (802.11) on the last segment, and WiFi signal strength degradation manifests in high first-hop delay, due to the fact that the MAC layer will retransmit packets lost at that layer. Measuring the RTT between endpoints on the customer network and parts of the service provider's own infrastructure (which

have predictable delay characteristics) can be used to isolate this cause of performance problems.

The network provider can measure the RTT at the home gateway, or at an upstream point if there is no access to home gateway. A problem in the WiFi network is identified by seeing high delay and low packet loss.

These measurements are particularly useful for traffic which is latency sensitive, such as interactive video applications. However, since high latency is often correlated with other network-layer issues such as chronic interconnect congestion [[IMC-CONGESTION](#)], it is useful for general troubleshooting of network layer issues in an interdomain setting.

In this case, multiple RTT samples per flow are useful less for observing intraflow behavior, and more for generating sufficient samples for a given aggregate to make a high-quality measurement.

[2.4.](#) Internet Measurement Research

As a large, distributed, engineered system with no centralized control, the Internet has emergent properties of interest to the research community not just for purely scientific curiosity, but also to provide applicable guidance to Internet engineering, Internet protocol design and development, network operations, and policy development. Latency measurements in particular are both an active area of research as well as an important tool for certain measurement studies (see, e.g. [[IMC-TCPSIG](#)], from the most recent Internet Measurement Conference). While much of this work is currently done with active measurements, the ability to generate latency samples passively or using a hybrid measurement approach (i.e., through passive observation of purpose-generated active measurement traffic; see [[RFC7799](#)]) can drastically increase the efficiency and scalability of these studies.

[3.](#) Indirect Utility of RTT Measurements

In addition to the direct generation of RTT metric samples, RTT measurement can also be used for indirect generation of other metrics when more direct means are not available.

A variety of tools are used for detailed troubleshooting of the performance of single flows, both for debugging transport- and application-layer protocol implementations, as well as to determine whether a particular end-to-end performance issue is related to particular network conditions. One common type of visualization used for TCP (implemented, for example, in the TCP Stream Graphs feature

of Wireshark, <https://www.wireshark.org/>) shows the development over time of the sequence and acknowledgment numbers, including retransmissions, and the evolution of the inflight and receiver flow control windows over time. By analyzing the relationship among loss, latency, and throughput, the precise cause of an observed performance on a given flow can be determined.

While RTT measurements on their own are not enough to drive such a visualization, many similar techniques can be built on high-resolution time series RTT data. Here we exploit two properties of transport protocols:

- o The size of the inflight window is equal to the number of bytes/packets sent per RTT, so inflight window evolution can be generated at each RTT sample r at t , and summing the number of bytes/packets sent between $t - r$ and t .
- o Changes in the inflight window can be related to sender reactions to congestion. For common loss- and ECN-based congestion control protocols such as NewReno [[RFC6582](#)] and Cubic [[RFC8312](#)], inflight window reductions are correlated with sender-experienced congestion or loss.

Inflight window evolution over time, together with heuristic assumptions about server behavior, can go a long way toward replacing direct visibility of transport protocol dynamics (sequence and acknowledgment number sequence over time) for encrypted transports; the exact details of this are a subject of present and future research.

4. Additional Metrics Derivable from the Spin Bit

The latency spin signal mechanism itself [[TSVWG-SPIN](#)] has additional measurement utility; these observations do not apply to other methodologies for measuring RTT.

4.1. Derived Loss and Reordering

When used alone (as a one-bit signal), measurement systems using the latency spin bit must use heuristics to reject samples which are potentially-lost, potentially-reordered, or potentially-delayed. When these heuristics are instrumented to note their sample rejection rate, this rate itself is a potentially-useful proxy metric for "difficulty" (vaguely defined) experienced by a flow.

When the latency signal is used with the Valid Edge Counter (VEC), additional information is available in the wire image to reject samples due to loss, delay, or reordering. Analysis of the VEC

together with the series of spin bit values can be used to recognize single loss and reordering events, which can be used to generate loss and reordering metrics at the resolution of the flow's round trip time. Optimal use of the VEC signal to generate loss and reordering metric signals is a subject of ongoing research.

4.2. Two-Point Intradomain Measurement

The spin bit is also useful as a basic signal for instantaneous measurement of the treatment of traffic carrying the latency spin signal within a single network. Though the primary design goal of the spin bit signal is to enable single-observer on-path measurement of end-to-end RTT, the spin bit can also be used by two cooperating observers with access to traffic flowing in the same direction as an alternate marking signal, as described in [\[ALT-MARK\]](#). The only difference from alternate marking with a generated signal is that the size of the alternation will change with the flight size each RTT. However, these changes do not affect the applicability of the method that works for each marking batch separately applied between two measurement points on the same direction. This two point measurement is an additional feature enabled "for free" by the spin bit signal.

So, with more than one observer on the same direction, it can be useful to segment the RTT and deduce the contribution to the RTT of the portion of the network between two on-path observers. This can be easily performed by calculating the delay between two or more measurement points on a single direction by applying [\[ALT-MARK\]](#). In this way, packet loss, delay and delay variation can be measured for each segment of the network depending on the number and distribution of the available on-path observation points. When these observation points are applied at network borders, the alternate-marking signal can be used to measure the performance of QUIC traffic within a network operator's own domain of responsibility. own portion of the network.

5. Contributors

This document contains text from [\[QUIC-SPIN\]](#), which is 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

6. Acknowledgments

Thanks to Mark Nottingham for suggesting that this document should exist.

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.

7. References

7.1. Normative References

[TSVWG-SPIN]

Trammell, B., "A Transport-Independent Explicit Signal for Hybrid RTT Measurement", [draft-trammell-tsvwg-spin-00](#) (work in progress), July 2018.

7.2. Informative References

[ALT-MARK]

Fioccola, G., Capello, A., Cociglio, M., Castaldelli, L., Chen, M., Zheng, L., Mirsky, G., and T. Mizrahi, "Alternate Marking method for passive and hybrid performance monitoring", [draft-ietf-ippm-alt-mark-14](#) (work in progress), December 2017.

[IMC-CONGESTION]

Luckie, M., Dhamdhere, A., Clark, D., Huffaker, B., and k. claffy, "Challenges in Inferring Internet Interdomain Congestion (in Proc. ACM IMC 2014)", November 2014.

[IMC-TCPSIG]

Sundaresan, S., Dhamdhere, A., Allman, M., and . k claffy, "TCP Congestion Signatures (in Proc. ACM IMC 2017)", n.d..

[QUIC]

Iyengar, J. and M. Thomson, "QUIC: A UDP-Based Multiplexed and Secure Transport", [draft-ietf-quic-transport-14](#) (work in progress), August 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.

[RFC6049] Morton, A. and E. Stephan, "Spatial Composition of Metrics", [RFC 6049](#), DOI 10.17487/RFC6049, January 2011, <<https://www.rfc-editor.org/info/rfc6049>>.

[RFC6582] Henderson, T., Floyd, S., Gurtov, A., and Y. Nishida, "The NewReno Modification to TCP's Fast Recovery Algorithm", [RFC 6582](#), DOI 10.17487/RFC6582, April 2012, <<https://www.rfc-editor.org/info/rfc6582>>.

[RFC7799] Morton, A., "Active and Passive Metrics and Methods (with Hybrid Types In-Between)", [RFC 7799](#), DOI 10.17487/RFC7799, May 2016, <<https://www.rfc-editor.org/info/rfc7799>>.

[RFC8312] Rhee, I., Xu, L., Ha, S., Zimmermann, A., Eggert, L., and R. Scheffenegger, "CUBIC for Fast Long-Distance Networks", [RFC 8312](#), DOI 10.17487/RFC8312, February 2018, <<https://www.rfc-editor.org/info/rfc8312>>.

[WIRE-IMAGE]

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

[WMM-BLOAT]

Alfredsson, S., Giudice, G., Garcia, J., Brunstrom, A., Cicco, L., and S. Mascolo, "Impact of TCP Congestion Control on Bufferbloat in Cellular Networks (in Proc. IEEE WoWMoM 2013)", June 2013.

Author's Address

Brian Trammell
ETH Zurich

Email: ietf@trammell.ch

