QUIC WG
Internet Draft
Intended status: Informational
Expires: November 2019

E. Stephan
M. Cayla
A. Braud
F. Fieau
A. Ferrieux
Orange
M. Ihlar
Ericsson

May 20, 2019

QUIC Interdomain Troubleshooting
draft-stephan-quic-interdomain-troubleshooting-02.txt

Status of this Memo

Copyright Notice

Abstract

On-path network performance measurements methods currently deployed contribute to the ossification of the Internet because they are expensive to deploy and to maintain. This draft motivates the exposure of QUIC header fields for on-path network measurements and their specification in the QUIC core protocol as a solution to avoid on-path network performance measurements to ossify the IP stack in the future.

Table of Contents

1. Introduction

The IP layer does not include the material for measuring the delay and packet losses of segments of a path. The network performance is currently measured by points of presence of the path [SPATIAL], [COMPO] using transport fields of the upper layers: TCP transport layer, RTP application layer...

The evolution of the Internet stack toward end-to-end integrity protection is unavoidable [IABSEC]. This document presents the benefits of preserving the same on-path network performance

measurement capabilities in the evolution of TCP (TCP/TLS, TCPinc...) and UDP (QUIC/UDP...) currently specified at the IETF.

On-path network performance measurement methods currently deployed contribute to the ossification of the Internet because they are expensive to deploy and complex to maintain. This is due to the use of protocol fields not primarily designed for this purpose. This draft motivates the exposure of the fields for on-path network performance measurements of the delay [SPINBIT] and of the packet losses [SQUAREBIT] in the QUIC core protocol to avoid network performance measurements to ossify the IP stack in the future.

The memo recalls the UX interdomain troubleshooting complexity [FALLBACK] introduced by QUIC deployment. Then it describes operational concerns QUIC fallbacks and discusses the potential impacts on the security. Finally it discusses the benefits of exposing durably the fields needed for measuring packet delay and losses.

2. Conventions used in this document

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. Interdomain UX troubleshooting

Fast troubleshooting of network performance is mandatory to maintain end-users high Quality of Experience.

Troubleshooting is the act of identifying the origin of a problem. A major case is the localization of troubles impacting a large number of customers. This case becomes critical when it appears suddenly and represents a noticeable part of the traffic between the entity that connects customers (ISP) and the entity that provides the data (APP).

It becomes critical because network operation center (NOC) teams of the two entities are expected to immediately identify the causes in order to restore UX as quickly as possible. Each team checks that the point of failure is either in their domain or outside. When they located the point of failure in their entity they investigate their own chains of components (network, routers, reverse proxies,...) and quickly fix the issue.

It becomes extremely critical when an entity locates the point of failure outside of their entity. In this case the time needed to fix the problem is much longer and unpredictable because it expects other entities on the path to perform the same actions on their segments.

There are many cases of troubleshooting. A typical example starts by signaling to the ISP that its end-users are experiencing a significant decrease of QoE when using an Internet application. Typical point of failures can be line card memory errors or overloaded routers located somewhere on the path, either in the ISP domain or outside.

The ISP NOC has to localize asap the point of failure. Currently it proceeds by dichotomy (is the point of failure inside ISP domain or outside?) using passive monitoring of packet loss and congestion.

Following is the description of the parameters in use:

Packet lost downstream (vice versa for upstream):

- Measure of packet loss before the point of measure needs TCP sequence number;

- Measure of loss located after the point of measure needs TCP ACK + SACK.

Congestion:

- Congestion also manifests as increased delays in queues, located by measuring half-RTTs upstream and downstream from the point of measure;

- The analysis is based on TCP SEQ/ACK/Timestamp correlation.

4. Reference of Network Performance

A reference of the real performance of the network is not always provided by counters of network equipment. Counters may not be implemented, values are not always stable, their values could be compromised in case of software bugs or equipment congestion... In addition, to face the increasing network architecture complexity involved by the evolution of access networks, security and hosting infrastructures, NOCs need reliable network performance measurement in near real time.

In practice, these measurement tools shall be able to monitor numerous points and interfaces within the network to provide near real time network performances indicators taking into account the global network state.

These measurements systems can also allow detecting issues or unexplained behaviors on equipment, links, peers: for instance, in mobile networks, operators shall be able to identify in real time bottleneck links responsible for customer experience degradation and take necessary actions to avoid further snowball effect.

Charging of data usage is another key feature for Mobile Network Operators where per flow accurate information is collected. It is important to reconcile values between amount of charged data and amount of data seen by the network (cf discussion on goodput in section 6)). This could allow detecting fraud attempts or dysfunctions within the network. In case of significant gap, operators must be able to react quickly to isolate this traffic. Additionally, charging may require the differentiation of the goodput from the throughput.

Continuous network performance monitoring requires packet losses and delay measurements to allow operators to manage properly their networks and to provides them with a reference of performance of their network for interdomain troubleshooting.

5. Explicit measurement signals

An alternative to exposing raw transport protocol data to the path is to have explicitly designed signals with the purpose of facilitating on-path measurements. To facilitate troubleshooting, such a signal should enable passive measurement of RTT, packet-loss and other congestion indicators such as ECN. An example of how a transport protocol could expose measurement information to the path would be three flag bits available in a public header. The first bit would be used for passive RTT measurement, bit 2 would indicate whether the packet contains retransmitted data and bit 3 would indicate whether the packet contains an ECN echo. An explicit signal is unambiguous and simpler for a middlebox to interpret, than parsed transport headers. Furthermore, it could be invariant between revisions of the transport protocol that exposes it, which minimizes the risk of network ossification.

At this step the WG specified a signal to measure the round trip delay delay [SPINBIT].

QUIC packets numbering) are available in QUIC but encrypted an additional signal like described in [SQUAREBIT] is needed to measure and locate packet losses.

6. QUIC Fallback

Fallback is necessary to address cases where a QUIC connection establishment fails [QUICAPP] (A device of the path blocks UDP, the stack blocks 0-RTT...).

Fallback may occur additionally when an active QUIC connection drops and tries to reconnect. As an example, the steps of the fallback could be:

- The QUIC connection drops accidently;
- The UA fallback and connects in TCP/TLS to the origin server;
- The UA receives from the origin server an indication for an alternate service [ALTSVC] supporting QUIC;
- The UA ends gracefully the TCP connection;
- The UA tries to establish a QUIC connection to the server and port described in the alternate service indication;
- The QUIC 1-RTT connection is established;

6.1. Flapping

A fallback may suddenly occur when one or more elements (links, nodes, reverse proxy, switch, server ...) of the path fail or are reconfigured.

There are cases where the fallback loops and triggers flapping between the origin server and the alternate server. As an example, this might happen when an alternate service indication is outdated and points to a server which does not support QUIC anymore.

This becomes critical for UX when numerous fallbacks occur suddenly on the same path between a set of customers of an ISP and another entity which provides the application data. The time to troubleshoot can be very long. The origin server and the alternate servers can be hosted by different entities.

This should be specified either in QUIC core specifications.

7. Versioning and Implementations

Versioning is an important part of the QUIC protocol framework [QUICCORE]. Multiple versions of the protocol are expected to be deployed and used concurrently. In order to encourage networks to rapidly support the QUIC protocol and to support any versions of QUIC in the future, the exposure of the fields for on-path network performance measurement must not depend on the version.

There might be numerous implementations of the QUIC protocol in the future. An important part of them will implement the congestion control at application level. There will be unfair behaviors like abnormal retransmission rate which will impact the fairness of the repartition of the bandwidth amongst the customers of the network. By consequence the network needs to be able to detect connections which have abnormal throughput/goodput.

8. Security Considerations

The integrity of the parameters exposed for measuring on-path delay and losses can be end-to-end protected to increase the security of the connection.

Flapping from QUIC to a fallback protocol might overload on-path devices and end-points and by consequence affect the stability of the connections and introduces weaknesses.

The fallback from encrypted headers to clear headers transport protocols might open the door to new types of active attacks.

It is not clear yet whether a network can distinguish numerous QUIC fallback flappings from an active attack:

o What is the expected behavior from the network?

o Will networks detect QUIC flapping as an active attack?

9. IANA Considerations

This draft does not request any IANA actions.

10. Discussions

10.1. Fallback

Troubleshooting QUIC traffic and its fallbacks requires measuring similar metrics. One suggestion is to use the integrity mechanism of the TCPinc WG [TCPINC] to protect and keep visible the fields used for on-path measurement.

Fallback must be precisely specified in the core specification of QUIC [QUICCORE].

To avoid unnecessary flapping [QUICCORE] might clarify the usage of the advertisement of QUIC support in HTTP protocols [ALTSVC].

[QUICMAN] should propose guidance for the management of QUIC fallback in a way to avoid flapping situations.

10.2. On-path Measurement

QUIC is designed to carry other traffic than HTTP such as DNS and Web. End-to-end encryption of the transport headers prevents the use of models [E-MODEL] and heuristics to estimate UX on a path segment. To maintain a high level of UX, QUIC capabilities should support the measurement of the delay and the losses of a segment of a source to destination path.

On-path measurement techniques are currently ad hoc. Adding the exposure of the fields for on-path packet delay and losses in the core specification of the QUIC protocol creates a stable network performance measurement framework. It will be a real incentive for networks to support QUIC rapidly and to support the numerous QUIC versions in the future. This will reduce network impacts on the ossification of the IP stack in the future.

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.

[QUICCORE] https://tools.ietf.org/html/draft-ietf-quic-transport

[FALLBACK] https://github.com/quicwg/base-drafts/issues/166

[IABSEC] https://www.iab.org/2014/11/14/iab-statement-on-internet-confidentiality/

[SPINBIT] https://tools.ietf.org/html/draft-ietf-quic-transport-20#section-17.3.1

[SQUAREBIT] https://tools.ietf.org/html/draft-ferrieux-hamchaoui-quic-lossbits-00

11.2. Informative References

[E-MODEL] https://www.itu.int/rec/T-REC-G.107-199812-S/en

[SPATIAL] https://tools.ietf.org/html/rfc5644

[COMPO] https://tools.ietf.org/html/rfc6049

[QUICAPP] https://tools.ietf.org/wg/quic/draft-ietf-quic-applicability/

[QUICMAN] https://tools.ietf.org/wg/quic/draft-ietf-quic-manageability/

[TCPINC] https://tools.ietf.org/wg/tcpinc/

[ALTSVC] https://tools.ietf.org/html/rfc7838

12. Acknowledgments

This document was prepared using 2-Word-v2.0.template.dot.

Authors' Addresses

Emile Stephan
Orange
2, avenue Pierre Marzin
Lannion 22300
France

Email: emile.stephan@orange.com

Mathilde Cayla
Orange
6, avenue Albert Durand
Blagnac 31700
France

Email: mathilde.cayla@orange.com

Arnaud Braud
Orange
2, avenue Pierre Marzin
Lannion 22300
France

Email: arnaud.braud@orange.com

Fred Fieau
Orange
40-48, avenue de la Republique
Chatillon 92320
France

Email: frederic.fieau@orange.com

Alex Ferrieux
Orange
2, avenue Pierre Marzin
Lannion 22300
France

Email: alexandre.ferrieux@orange.com

Marcus Ihlar
Ericsson

Email: marcus.ihlar@ericsson.com