Packet Loss Signaling for Encrypted Protocols
draft-ferrieuxhamchaoui-tsvwg-lossbits-01

Abstract

This document describes a protocol-independent method that employs two bits to allow endpoints to signal packet loss in a way that can be used by network devices to measure and locate the source of the loss. The signaling method applies to all protocols with a protocol-specific way to identify packet loss. The method is especially valuable when applied to protocols that encrypt transport header and do not allow an alternative method for loss detection.

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 January 21, 2020.

Copyright Notice

Copyright (c) 2019 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
1. Introduction

Packet loss is a pervasive problem of day-to-day network operation, and proactively detecting, measuring, and locating it is crucial to maintaining high QoS and 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 the clear in TCP headers: sequence and acknowledgment numbers and SACKs when enabled (see \[RFC8517\]). These allow for quantitative estimation of packet loss by passive on-path observation, and the lossy segment (upstream or downstream from the observation point) can be quickly identified by moving the passive observer around.

With encrypted protocols, the equivalent transport headers are encrypted and passive packet loss observation is not possible, as described in \[TRANSPORT-ENCRIPT\].
Since encrypted protocols could be routed by the network differently, and the fraction of Internet traffic delivered using encrypted protocols is increasing every year, it is imperative to measure packet loss experienced by encrypted protocol users directly instead of relying on measuring TCP loss between similar endpoints.

Following the recommendation in [RFC8558] of making path signals explicit, this document proposes adding two explicit loss bits to the clear portion of the protocol headers to restore network operators’ ability to maintain high QoS for users of encrypted protocols. These bits can be added to an unencrypted portion of a header belonging to any protocol layer, e.g. IP (see [IP]) and IPv6 (see [IPv6]) headers or extensions, UDP surplus space (see [UDP-OPTIONS] and [UDP-SURPLUS]), reserved bits in a QUIC v1 header (see [QUIC-TRANSPORT]).

2. Notational Conventions

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 [RFC2119].

3. Loss Bits

The proposal introduces two bits that are to be present in every packet capable of loss reporting. These are packets that include protocol headers with the loss bits. Only loss of packets capable of loss reporting is reported using loss bits.

Whenever this specification refers to packets, it is referring only to packets capable of loss reporting.

- Q: The "sQuare signal" bit is toggled every N outgoing packets as explained below in Section 3.1.

- L: The "Loss event" bit is set to 0 or 1 according to the Unreported Loss counter, as explained below in Section 3.2.

Each endpoint maintains appropriate counters independently and separately for each connection (each subflow for multipath connections).

3.1. Setting the sQuare Bit on Outgoing Packets

The sQuare Value is initialized to the Initial Q Value (0 or 1) and is reflected in the Q bit of every outgoing packet. The sQuare value is inverted after sending every N packets (Q Period is 2^N). The Q bit represents "packet color" as defined by [RFC8321].
The choice of the Initial Q Value and Q Period is determined by the protocol containing Q and L bits. For example, the values can be protocol constants (e.g. "Initial Q Value" is 0, and "Q Period" is 128), or they can be set explicitly for each connection (e.g. "Initial Q Value" is whatever value the initial packet has, and "Q Period" is set per a dedicated TCP option on SYN and SYN/ACK).

Observation points can estimate the upstream losses by counting the number of packets during a half period of the square signal, as described in Section 4.

### 3.2. Setting the Loss Event Bit on Outgoing Packets

The Unreported Loss counter is initialized to 0, and L bit of every outgoing packet indicates whether the Unreported Loss counter is positive (L=1 if the counter is positive, and L=0 otherwise).

The value of the Unreported Loss counter is decremented every time a packet with L=1 is sent.

The value of the Unreported Loss counter is incremented for every packet that the protocol declares lost, using whatever loss detection machinery the protocol employs. If the protocol is able to rescind the loss determination later, the Unreported Loss counter SHOULD NOT be decremented due to the rescission.

Observation points can estimate the end-to-end loss, as determined by the upstream endpoint’s loss detection machinery, by counting packets in this direction with a L bit equal to 1, as described in Section 4.

### 4. Using the Loss Bits for Passive Loss Measurement

There are three sources of observable loss:

- **_upstream loss_** - loss between the sender and the observation point (Section 4.2)

- **_downstream loss_** - loss between the observation point and the destination (Section 4.4)

- **_observer loss_** - loss by the observer itself that does not cause downstream loss (Section 4.5)

The upstream and downstream loss together constitute _end-to-end loss_ (Section 4.1).

The Q and L bits allow detection and measurement of the types of loss listed above.
4.1. End-To-End Loss

The Loss Event bit allows an observer to calculate the end-to-end loss rate by counting packets with L bit value of 0 and 1 for a given connection. The end-to-end loss rate is the fraction of packets with L=1.

The simplifying assumption here is that upstream loss affects packets with L=0 and L=1 equally. This may be a simplification, if some loss is caused by tail-drop in a network device. If the sender congestion controller reduces the packet send rate after loss, there may be a sufficient delay before sending packets with L=1 that they have a greater chance of arriving at the observer.

4.2. Upstream Loss

Blocks of N (half of Q Period) consecutive packets are sent with the same value of the Q bit, followed by another block of N packets with inverted value of the Q bit. Hence, knowing the value of N, an on-path observer can estimate the amount of upstream loss after observing at least N packets. If "p" is the average number of packets in a block of packets with the same Q value, then the upstream loss is "1-p/N".

The observer needs to be able to tolerate packet reordering that can blur the edges of the square signal.

The Q Period needs to be chosen carefully, since the observation could become too unreliable in case of packet reordering and loss if Q Period is too small. However, when Q Period is too large, connections that send fewer than half Q Period packets do not yield a useful upstream loss measurement.

The observer needs to differentiate packets as belonging to different connections, since they use independent counters.

4.3. Correlating End-to-End and Upstream Loss

Upstream loss is calculated by observing the actual packets that did not suffer the upstream loss. End-to-end loss, however, is calculated by observing subsequent packets after the sender’s protocol detected the loss. Hence, end-to-end loss is generally observed with a delay of between 1 RTT (loss declared due to multiple duplicate acknowledgments) and 1 RTO (loss declared due to a timeout) relative to the upstream loss.

The connection RTT can sometimes be estimated by timing protocol handshake messages. This RTT estimate can be greatly improved by
observing a dedicated protocol mechanism for conveying RTT information, such as the Latency Spin bit of [QUIC-TRANSPORT].

Whenever the observer needs to perform a computation that uses both upstream and end-to-end loss rate measurements, it SHOULD use upstream loss rate leading the end-to-end loss rate by approximately 1 RTT. If the observer is unable to estimate RTT of the connection, it should accumulate loss measurements over time periods of at least 4 times the typical RTT for the observed connections.

If the calculated upstream loss rate exceeds the end-to-end loss rate calculated in Section 4.1, then either the Q Period is too short for the amount of packet reordering or there is observer loss, described in Section 4.5. If this happens, the observer SHOULD adjust the calculated upstream loss rate to match end-to-end loss rate.

### 4.4. Downstream Loss

Because downstream loss affects only those packets that did not suffer upstream loss, the end-to-end loss rate ("e") relates to the upstream loss rate ("u") and downstream loss rate ("d") as 

\[(1-u)(1-d)=1-e\].

Hence, 

\[d=(e-u)/(1-u)\].

### 4.5. Observer Loss

A typical deployment of a passive observation system includes a network tap device that mirrors network packets of interest to a device that performs analysis and measurement on the mirrored packets. The observer loss is the loss that occurs on the mirror path.

Observer loss affects upstream loss rate measurement since it causes the observer to account for fewer packets in a block of identical Q bit values (see {{upstreamloss}}). The end-to-end loss rate measurement, however, is unaffected by the observer loss, since it is a measurement of the fraction of packets with the set L bit value, and the observer loss would affect all packets equally (see Section 4.1).

The need to adjust the upstream loss rate down to match end-to-end loss rate as described in Section 4.3 is a strong indication of the observer loss, whose magnitude is between the amount of such adjustment and the entirety of the upstream loss measured in Section 4.2.
5. Ossification Considerations

Accurate loss information is not critical to the operation of any protocol, though its presence for a sufficient number of connections is important for the operation of the networks.

The loss bits are amenable to "greasing" described in [GREASE], if the protocol designers are not ready to dedicate (and ossify) bits used for loss reporting to this function. The greasing could be accomplished similarly to the Latency Spin bit greasing in [QUIC-TRANSPORT]. Namely, implementations could decide that a fraction of connections should not encode loss information in the loss bits and, instead, the bits would be set to arbitrary values. The observers would need to be ready to ignore connections with loss information more resembling noise than the expected signal.

6. Security Considerations

Passive loss observation has been a part of the network operations for a long time, so exposing loss information to the network does not add new security concerns.

7. Privacy Considerations

Guarding user’s privacy is an important goal for modern protocols and protocol extensions per [RFC7285]. While an explicit loss signal—a preferred way to share loss information per [RFC8558]—helps to minimize unintentional exposure of additional information, implementations of loss reporting must ensure that loss information does not compromise protocol’s privacy goals.

For example, [QUIC-TRANSPORT] allows changing Connection IDs in the middle of a connection to reduce the likelihood of a passive observer linking old and new subflows to the same device. A QUIC implementation would need to reset all counters when it changes the destination (IP address or UDP port) or the Connection ID used for outgoing packets. It would also need to avoid incrementing Unreported Loss counter for loss of packets sent to a different destination or with a different Connection ID.

8. IANA Considerations

This document makes no request of IANA.
9. Change Log

9.1. Since version 00
   - Addressed review comments
   - Improved guidelines for privacy protections for QIUC

10. Acknowledgments

The sQuare bit was originally suggested by Kazuho Oku in early proposals for loss measurement.

11. References

11.1. Normative References


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


11.2. Informative References


Authors’ Addresses

Alexandre Ferrieux
Orange Labs
EMail: alexandre.ferrieux@orange.com

Isabelle Hamchaoui
Orange Labs
EMail: isabelle.hamchaoui@orange.com