Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: 6 January 2023

P. S. Kim
TU Korea

7 July 2022

Passive Probing for Path MTU Discovery with QUIC
draft-pskim-passive-probing-pmtud-00

Abstract

This draft consider an alternative PMTUD for QUIC. To discover the best PMTU, the passive probing approach is adopted. The process of discovering the best PMTU is not carried out separately, but is carried out simultaneously in the actual application data communication. A probe packet is defined newly using 1-RTT packet which includes actual application data as well as a short packet header and a PING_EXT frame. The PING_EXT frame is also defined newly. Until the best PMTU is discovered, the size of the probe packet is changed according to the size of the PMTU candidate. A simple discovery algorithm using only the PMTU candidate sequence with linear upward is described in this draft. Other rather complex discovery algorithms that consider various PMTU candidate sequences will be dealt with in the future.

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 6 January 2023.

Table of Contents

1. Introduction

The maximum transmission unit (MTU) is the largest size frame or packet - in bytes or octets - that can be transmitted across a data link. It is most used in reference to packet size on an Ethernet network using the Internet Protocol (IP). The Path MTU (PMTU) is the smallest MTU of all involved network interfaces for a network path and limits the size of IP packets.

A PMTU Discovery (PMTUD) is a standardized technique in computer networking for determining the PMTU size on the network path between two IP hosts, usually with the goal of avoiding IP fragmentation for IPv4[RFC1191] and for IPv6[RFC8201]. When a packet too large for the path was sent, the PMTUD expects to receive a Packet Too Big (PTB) message. However, there are multiple reasons why a PTB message might not arrive at the sender.

Therefore, the PMTUD for the Packetization Layer (PL) that selects the size of IP packets is specified recently in [RFC8899]. RFC8899 works without a signal from the network and covers generic PL protocols such as QUIC of [RFC9000]. However, RFC8899 does not contain details about how to discovery for the best PMTU.

Recently, therefore, [Q-PMTUD] complements RFC8899 by presenting a discovery algorithm with QUIC. Using the discovery algorithm with a set of possible PMTU candidates and their possible probing sequences, the best PMTU is obtained. However, to discover the best PMTU, some probe packets which have no semantic value might be injecting into network, which is called active probing or active measurement. The active probing approach can increase a network load and perturb the network. In addition, [UDP-PMTUD] also complements RFC8899 by specifying how a UDP Options sender implements Datagram PL PMTUD

(DPLPMTUD). It allows a datagram application to discover the largest size of datagram that can be sent across a specific network path.

Based on [Q-PMTUD] and [UDP-PMTUD], this draft consider an alternative PMTUD for QUIC. To discover the best PMTU, the passive probing approach is adopted. The process of discovering the best PMTU is not carried out separately, but is carried out simultaneously in the actual application data communication. A probe packet is defined newly using 1-RTT packet which includes actual application data as well as a short packet header and a PING_EXT frame. The PING_EXT frame is also defined newly. Until the best PMTU is discovered, the size of the probe packet is changed according to the size of the PMTU candidate. A simple discovery algorithm using only the PMTU candidate sequence with linear upward is described in this draft. Other rather complex discovery algorithms that consider various PMTU candidate sequences will be dealt with in the future.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. Active Probing for PMTUD with QUIC[Q-PMTUD]

The specification of QUIC in RFC9000 recommends to use the PMTUD framework of RFC8899. However, RFC8899 does not contain details about how to discovery for the best PMTU.

Therefore, [Q-PMTUD] complements the specification, RFC8899, by presenting a discovery algorithm with QUIC. From a practical point of view, it might be a good choice to consider only a set of common PMTU values. However, the PMTU value may usually change over time. Thus, [Q-PMTUD] considers a set of possible PMTU candidates. Then, a discovery algorithm is proposed, which probes one PMTU candidate after the other. This means, it starts the probe for the next candidate not before the probe for the current candidate either succeeded or failed. Then endpoint uses this discovery algorithm that repeatedly chooses PMTU candidates to probe.

The candidate sequence is required to specify the order in which the discovery algorithm probes PMTU candidates. The endpoint must choose a PMTU candidate larger than the largest successfully probed candidate and smaller than any other probed candidate with a lost probe packet. Seven candidate sequences are considered, evaluated, and compared in [Q-PMTUD].

To probe one PMTU candidate, according to RFC9000, the endpoint builds a probe packet with a short packet header, a PING frame and PADDING frames. The endpoint controls the size of the probe packet by the number of PADDING frames, whose size is one byte each. The PING frame makes the packet ack-eliciting.

However, to discover the best PMTU, some probe packets which have no semantic value might be injecting into network, which is thus called active measurement or active probing. This active probing approach can increase a network load and perturb the network.

3. Passive Probing for PMTUD with QUIC

There are three possible ways to create a PMTU probe packet as follows[RFC8899]:

- Probing using padding data
- Probing using application data and padding data
- Probing using application data

[UDP-PMTUD] describes "Probe Packets that include Application Data" to implement "Probing using application data" of [RFC8899].

3.1. A new PMTU probe packet (1-RTT packet format)

(1) Probe packet format for active probing [Q-PMTUD]

IP header + UDP header + Short header(QUIC header) + PING frame + PADDING frames

The size of the probe packet is controlled by the number of PADDING frames.

(2) Probe packet format for passive probing

In this drfat, a probe packet is defined newly using 1-RTT packet including actual application data as well as a PING_EXT frame as follows:

IP header + UDP header + Short header(QUIC Header) + PING_EXT frame + Actual application data

- PING_EXT frame (defined newly)
  . Frame Type Name : PING_EXT
  . Type Value : 0x20
  . The PING_EXT frame makes the packet ack-eliciting. In addition, the PING_EXT frame indicates that the current 1-RTT packet is now discovering the best PMTU as well as transmitting actual application data.

- Application data
  . Actual application data controls the size of the probe packet by a multiple of four bytes.

The size of probe packet is changed according to PMTU candidates (=1280 + incremental where, for example, incremental can be a multiple of four as shown in [Q-PMTUD]).

3.2. Passive probing to both discover best PMTU and transmit actual application data

Through the new probe packet, it is possible not only to discovery the best PMTU, but also to transmit actual application data. That is, to discover the best PMTU size and carry actual application data, the endpoint expand the payload of all UDP datagrams.

(1) A simple algorithm for discovering the best PMTU

As specified in RFC9000, QUIC must send QUIC packets with the smallest allowed maximum datagram size when validating a path during connection initiation or migration. Thus, the endpoint sets the probe packet initially to the smallest allowed maximum datagram size of 1280 bytes including actual application data as well as a short packet header, a PING_EXT frame.

As mentioned, until the best PMTU is discovered, the size of the probe packet is changed successively according to the size of the PMTU candidate. The size of the probe packet is controlled with the size of actual application data. The size of actual application data is a multiple of four.

In the active probing approach [Q-PMTUD], the endpoint uses a simple discovery algorithm that repeatedly chooses PMTU candidates to probe. Thus, seven PMTU candidate sequences are considered and each candidate sequence specifies the order in which the discovery algorithm probes PMTU candidates. In addition, four metrics such as number of probed PMTU candidates, time to discover the best PMTU, network load, average PMTU estimation are defined for performance evaluations of seven sequences.

However, because the process of discovering the best PMTU is carried out simultaneously in the actual application data communication, only the PMTU candidate sequence with linear upward is adopted first in this draft. The linear upward sequence selects one candidate after the other from a list of candidates in ascending order, starting with the second one (the first one was probed with the smallest allowed maximum datagram size of 1280 bytes). Other rather complex discovery algorithms that consider various PMTU candidate sequence will be dealt with in the future.

Until the best PMTU is discovered, the endpoint repeats a series of probing steps. In absence of a PTB message, the discovery algorithm considers a probe for a PMTU candidate as failed, only if the probe packet of the size of the candidate were detected as lost. A probe for a PMTU candidate that fails, lets all other probes for larger candidates fail as well. Therefore, the best PMTU is the PMTU candidate that succeeded just before the failure.

(2) Discovery complete and PMTU cache

When the algorithm determines that it has discovered the best PMTU, the endpoint terminates the probing. Then, the endpoint sets the 1-RTT packet finally to the best datagram size using the best PMTU discovered. From now on, the 1-RTT packet does not include a PING_EXT frame. QUIC can cache the best PMTU discovered and use it for future connections to the same endpoint.

(3) Other rather complex discovery algorithms

Other rather complex discovery algorithms that consider various PMTU candidate sequences will be dealt with in the future.

4. IANA Considerations

This memo includes no request to IANA.

5. Security Considerations

The same security considerations as those described in RFC7880 will apply to this document.

6. References

[RFC1191] Mogul, J. and S. Deering, "Path MTU discovery", RFC 1191, DOI 10.17487/RFC1191, November 1990, <https://www.rfc-editor.org/info/rfc1191>.

[RFC8201] McCann, J., S. Deering, J. Mogul, R. Hinden, Ed. "Path MTU Discovery for IP version 6", RFC 8201, DOI 10.17487/RFC8201, July 2017, <https://www.rfc-editor.org/info/rfc8201>.

[RFC8899] Fairhurst, G., T. Jones, M. Tuxen, I. Rungeler, T. Volker, "Packetization Layer Path MTU Discovery for Datagram Transports", RFC 8899, DOI 10.17487/RFC8899, September 2020, <https://www.rfc-editor.org/info/rfc8899>.

[RFC9000] J. Iyengar, Ed., M. Thomson, Ed., "QUIC: A UDP-Based Multiplexed and Secure Transport", RFC 9000, DOI 10.17487/RFC9000, May 2021, <https://www.rfc-editor.org/info/rfc9000>.

[Q-PMTUD] Timo Volker, Michael Tuxen, "The search of the path MTU with QUIC", EPIQ '21: Proceedings of the 2021 Workshop on Evolution, Performance and Interoperability of QUIC, December 2021

[UDP-PMTUD] Work in Progress, Internet-Draft, draft-ietf-tsvwg-udp-options-dplpmtud-03, 25 February 2022, <https://www.ietf.org/archive/id/draft-ietf-tsvwg-udp-options-dplpmtud-03.txt>.

Authors' Addresses

Pyung Soo Kim
Tech University of Korea
Siheung, Gyeonggi
Korea
Email: pskim@tukorea.ac.kr