Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: October 27, 2019

G. Fairhurst
T. Jones
University of Aberdeen

April 25, 2019

Datagram PLPMTUD for UDP Options
draft-fairhurst-tsvwg-udp-options-dplpmtud-00

Abstract

This document describes how a UDP Options sender may implement Datagram Packetization Layer Path Maximum Transmission Unit Discovery (DPLPMTUD) as a robust method for Path Maximum Transmission Unit Discovery.

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 October 27, 2019.

Table of Contents

1. Introduction 2
2. Terminology 3
3. DPLPMTUD for UDP Options 3
3.1. Sending Packet Probes using Control Information 4
3.2. Sending Packet Probes using Application Data 5
3.3. Validating the Path with UDP Options 5
3.4. Handling of PTB Messages by UDP Options 5
4. Acknowledgements 6
5. IANA Considerations 6
6. Security Considerations 6
7. References 6
7.1. Normative References 6
7.2. Informative References 7
Appendix A. Revision Notes 7
Appendix B. Informative Description of new UDP Options 7
B.1. UDP Probe Request Option 7
B.2. UDP Probe Response Option 8
Authors' Addresses 8

1. Introduction

The User Datagram Protocol [RFC0768] offers a minimal transport service on top of IP and is frequently used as a substrate for other protocols. Applications using UDP frequently have to implement basic transport services such as Path Maximum Transmission Unit Discovery (PMTUD) themselves. Section 3.5 of UDP Guidelines ([RFC8085]) recommends that applications implement some form of Path MTU Discovery to avoid the generation of IP fragments:

"Consequently, an application SHOULD either use the path MTU information provided by the IP layer or implement Path MTU Discovery (PMTUD)".

The UDP API [RFC8304] offers calls for applications to receive ICMP Packet Too Big (PTB) messages and to control the size of messages that are sent, but does not offer any automatic mechanisms for an application to discover the maximum packet size supported by a path. Applications and upper layer protocols are left to implement robust PMTUD mechanisms of their own.

Packetization Layer PMTUD (PLPMTUD) [RFC4821] describes a method for a Packetization Layer (such as UDP with options) to search for the largest MTU supported on a path in the absence of ICMP PTB messages. Datagram PLPMTUD [I-D.ietf-tsvwg-datagram-plpmtud] describes PMTUD probing and search algorithms for datagram transports that does not solely rely on ICMP PTB messages. This allows the Packetization

Layer to offer a probing mechanism which works in the presence of lost probes. However, UDP is unable itself to offer the required probing mechanisms to implement DPLPMTUD without some additional transport services. This document specifies the additional functionality required to perform DPLPMTUD with UDP Options as a service to upper-layer protocols.

The authors solicit comments from the TSV working group. The working group could decide to incorporate the text in the current contribution into subsequent versions of the UDP Options Specification.

The structure of the present document follows the structure used to describe DPLPMTUD for other transports [I-D.ietf-tsvwg-datagram-plpmtud].

2. Terminology

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 BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. DPLPMTUD for UDP Options

UDP Options[I-D.ietf-tsvwg-udp-options] supplies additional functionality that can be used to implement DPLPMTUD within the UDP transport service. Implementing DPLPMTUD using UDP Options avoids the need for each upper layer protocol or application to implement the DPLPMTUD method.

Section 5.6 of[I-D.ietf-tsvwg-udp-options] defines the Maximum Segment Size (MSS) option, which allows the local sender to indicate the EMTU_R to the peer. The value received in this option can be used to initialise MAX_PMTU used by DPLPMTUD.

The DPLPMTUD method [I-D.ietf-tsvwg-datagram-plpmtud] relies upon the sender Packetization Layer to be able to generate probe packets with a specific size. UDP Options enables padding to be added to a UDP datagram that is used as these Probe Packets. Feedback confirming reception of each Probe Packet is provided by two UDP Options described in section 6 of [I-D.ietf-tsvwg-udp-options]:

- The Probe Request Option (Appendix B.1) is set by a sending PL to solicit a response from a remote endpoint. A four-byte token identifies each request.

- The Probe Response Option (Appendix B.2 is generated by the UDP Options receiver in response to reception of a previously received Probe Request Option. Each Probe Response Option echoes a previously received four-byte token.

The token value allows implementations to be distinguish between acknowledgements for initial probe packets and acknowledgements confirming receipt of subsequent probe packets (e.g., travelling along alternate paths with a larger RTT). Each probe packet needs to be uniquely identifiable by the UDP Options sender within the Maximum Segment Lifetime (MSL). The UDP Options sender therefore needs to not recycle token values until they have expired or have been acknowledged. A four byte value for the token field provides sufficient space for multiple unique probes to be made within the MSL.

The initial value of the four byte token field SHOULD be assigned to a randomised value, as described in section 5.1 of [RFC8085]) to enhance protection from off-path attacks.

Implementations ought to only send a probe packet with a Request Probe Option when required by their local state machine, i.e., when probing to grow the PLPMTU or to confirm the current PLPMTU. The procedure to handle the loss of a response packet is the responsibility of the sender of the request. Implementations are allowed to track multiple requests and respond to them with a single packet.

A PL needs to determine that the current path can still support the size of datagram that the application is currently sending in the DPLPMTUD search_done state (i.e., to detect black-holing of data). One way to achieve this is to send probe packets of size PLPMTU or to utilise a higher-layer method that provides explicit feedback indicating any packet loss.

3.1. Sending Packet Probes using Control Information

The recommended approach to implementing DPLPMTUD sends a Probe Packet that contains only control information together with any padding that is needed to be inflated to the size required for the probe packet. These probe packets do not carry an application-supplied data block and therefore they do not typically require retransmission, although they do still consume network capacity and incur endpoint processing (see Section 4.1 of [I-D.ietf-tsvwg-datagram-plpmtud]).

3.2. Sending Packet Probes using Application Data

Another possibility is to utilise data packets for Probe Packets (Section 4.1 of [I-D.ietf-tsvwg-datagram-plpmtud] discusses the merits and demerits of this approach). A probe packet then contains a data block supplied by an application that is combined with padding to inflate the length of the datagram to the size required for the probe and additionally contains a Timestamp Option.

Reception of a valid Timestamp Option that was echoed by the remote endpoint can also be used to infer connectivity. This can provide useful feedback even over paths with asymmetric capacity and/or that carry UDP Option flows that have very asymmetric datagram rates, because an echo of the most recent Timestamp still indicates reception of at least one packet of the transmitted size. This is sufficient to confirm there is no black hole.

If the application/transport needs protection from the loss of this Probe Packet, the application/ transport could perform transport-layer retransmission/repair of the data block (e.g., by retransmission after loss is detected or by duplicating the data block in a datagram without the padding) [RFC8085].

When sending a probe to increase the PLPMTU, a Timestamp might be unable to unambiguously identify that a specific probe packet has been received. Timestamp mechanisms cannot be used to confirm the reception of individual probe messages and cannot be used to stimulate a response from the remote peer.

3.3. Validating the Path with UDP Options

The UDP sender validates the responses to a Packet Probe [I-D.ietf-tsvwg-datagram-plpmtud] using the UDP port information in combination with the token and/or Timestamp value contained in the UDP Option.

3.4. Handling of PTB Messages by UDP Options

The UDP sender validates any received ICMP PTB message that is received in response to a Packet Probe [I-D.ietf-tsvwg-datagram-plpmtud] using the quoted packet to validate the UDP port information in combination with the token and/or timestamp value contained in the UDP Option.

4. Acknowledgements

Gorry Fairhurst and Tom Jones are supported by funding provided by the University of Aberdeen.

5. IANA Considerations

This memo includes no requests to IANA.

6. Security Considerations

The security considerations for are described in [I-D.ietf-tsvwg-udp-options]. The proposed new method does not change the integrity protection offered by the UDP options method.

7. References

7.1. Normative References

[I-D.ietf-tsvwg-datagram-plpmtud]
Fairhurst, G., Jones, T., Tuexen, M., Ruengeler, I., and T. Voelker, "Packetization Layer Path MTU Discovery for Datagram Transports", draft-ietf-tsvwg-datagram-plpmtud-07 (work in progress), February 2019.

[I-D.ietf-tsvwg-udp-options]
Touch, J., "Transport Options for UDP", draft-ietf-tsvwg-udp-options-07 (work in progress), March 2019.

[RFC0768] Postel, J., "User Datagram Protocol", STD 6, RFC 768, DOI 10.17487/RFC0768, August 1980, <https://www.rfc-editor.org/info/rfc768>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC4821] Mathis, M. and J. Heffner, "Packetization Layer Path MTU Discovery", RFC 4821, DOI 10.17487/RFC4821, March 2007, <https://www.rfc-editor.org/info/rfc4821>.

[RFC8085] Eggert, L., Fairhurst, G., and G. Shepherd, "UDP Usage Guidelines", BCP 145, RFC 8085, DOI 10.17487/RFC8085, March 2017, <https://www.rfc-editor.org/info/rfc8085>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

7.2. Informative References

[RFC8304] Fairhurst, G. and T. Jones, "Transport Features of the User Datagram Protocol (UDP) and Lightweight UDP (UDP-Lite)", RFC 8304, DOI 10.17487/RFC8304, February 2018, <https://www.rfc-editor.org/info/rfc8304>.

Appendix A. Revision Notes

XXX Note to RFC-Editor: please remove this entire section prior to publication. XXX

Individual draft-00.

- This version contains a description for consideration and comment by the TSVWG.

Appendix B. Informative Description of new UDP Options

XXX Note to RFC-Editor: please remove this entire section prior to publication (including subsections) before publication. XXX

This annexe contains a provisional description of the UDP Options that will be specified in [I-D.ietf-tsvwg-udp-options]. The information is provided for information only, to enable understanding of the algorithm and will be superseded by text in the UDP Options specification.

B.1. UDP Probe Request Option

The Probe Request Option allows a sending endpoint to solicit a response from a destination endpoint.

The Probe Request Option carries a four byte token set by the sender. This token can be set to a value that is likely to be known only to the sender (and is sent along the end-to-end path). The initial value of the token SHOULD be assigned to a randomised value, as described in section 5.1 of [RFC8085]) to enhance protection from off-path attacks.

The sender needs to then check the value returned in the UDP Probe Response Option. The value of the Token field, uniquely identifies a probe within the maximum segment lifetime.

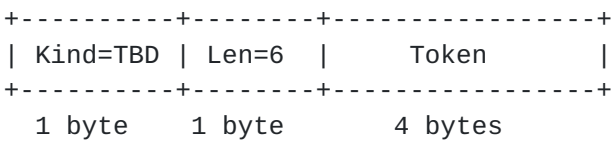

Figure 1: UDP Probe REQ Option Format

B.2. UDP Probe Response Option

The Probe Response Option is generated in response to reception of a previously received Probe Request Option. This response is generated by the UDP Option processing.

The Probe Response Option carries a four byte Token field. The Token field associates the response with the token value carried in the most recently-received Echo Request. The rate of generation of UDP packets carrying a Probe Response Option is expected to be less than once per RTT and SHOULD be rate-limited (see Section 6).

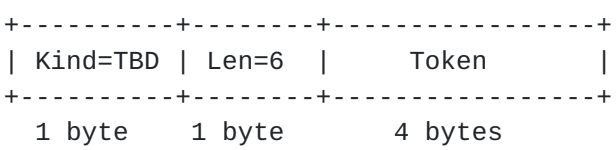

Figure 2: UDP Probe RES Option Format

Authors' Addresses

Godred Fairhurst
University of Aberdeen
School of Engineering
Fraser Noble Building
Aberdeen AB24 3UE
UK

Email: gorry@erg.abdn.ac.uk

Tom Jones
University of Aberdeen
School of Engineering
Fraser Noble Building
Aberdeen AB24 3UE
UK

Email: tom@erg.abdn.ac.uk