Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: April 4, 2020

G. Fairhurst
T. Jones
University of Aberdeen

October 2, 2019

Datagram PLPMTUD for UDP Options
draft-fairhurst-tsvwg-udp-options-dplpmtud-01

Abstract

This document specifies how a UDP Options sender implements 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 April 4, 2020.

Table of Contents

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 do not

solely rely on ICMP PTB messages. This allows the Packetization Layer (PL) 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 [I-D.ietf-tsvwg-udp-options]as a service to upper-layer protocols. 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.

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.

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

3. DPLPMTUD for UDP Options

The DPLPMTUD PL endpoint implements the method specified in [I-D.ietf-tsvwg-datagram-plpmtud].

3.1. Confirmation of Connectivity across a Path

The DPLPMTUD method requires that the PL is able to confirm connectivity on the path (Section 5.2.1 of [I-D.ietf-tsvwg-datagram-plpmtud]).

The UDP API does not offer a mechanism for a sender to confirm connectivity, a UDP Options sender SHOULD use reception of an option that elicits a positive confirmation (i.e. Timestamps, ECHO Request/ Response) to confirm connectivity of the path.

3.2. Sending UDP-Options Probe Packets

This method 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.

A PL needs to determine that the current path continues to support the size of datagram that the application is currently sending when in the DPLPMTUD search_done state i.e., to detect black-holing of data (Section 4.2 of [I-D.ietf-tsvwg-datagram-plpmtud]). UDP Options can achieve this by sending probe packets padded to the size of the PLPMTU.

A PL also needs, from time to time, to determine whether the current path can support a larger size of datagram that the current PLPMTU. UDP Options can achieve this by sending probe packets padded to the size required for the Packet Probe.

A UDP options sender MUST be able to send probes up to the maximum for the size the local interface supports, and not constrained by the maximum PMTU set by network layer mechanisms (such as PMTUD [RFC1063][RFC8201]). There are various options described in DPLPMTUD to send a Packet Probe to test the size of packet supported by a path (see Section 4.1 of [I-D.ietf-tsvwg-datagram-plpmtud]). This secification recommends "Probing using padding data".

3.2.1. Sending Packet Probes using the Echo Request Option Request Option

The RECOMMENDED method sends a Packet Probe with the Echo Request Option (RES) together with any padding needed to be inflated to the size required for the Packet Probe. The reception of this option generates an Echo Response Option that confirms reception of each received Packet Probe.

Implementations ought to only send a Packet Probe with a Request Probe Option when required by their local DPLPMTUD state machine, i.e., when probing to grow the PLPMTU or to confirm the current PLPMTU.

Packet Probes consume network capacity and incur endpoint processing (see Section 4.1 of [I-D.ietf-tsvwg-datagram-plpmtud]).

Implementations MAY track multiple requests and respond to them with a single packet.

The UDP Options used in this method are described in section 6 of [I-D.ietf-tsvwg-udp-options]:

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

- The Echo Response Option (REQ) is generated by the UDP Options receiver in response to reception of a previously received Echo Request Option. Each Echo Response Option echoes a previously received four-byte token.

The token value allows implementations to distinguish between acknowledgements for initial Packet Probes and acknowledgements confirming receipt of subsequent Packet Probes (e.g., travelling along alternate paths with a larger round trip time). Each Packet Probe 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 to enhance protection from off-path attacks, as described in section 5.1 of [RFC8085]).

The procedure to handle the loss of a datagram is the responsibility of the sender of the request. Implementations MAY track multiple requests and respond to them with a single packet carrying the Echo Response Option (REQ).

3.2.2. Sending Packet Probes that include Application Data

The RECOMMENDED approach to generate Packet Probes is to send a probe that contains only control information.

A sender could include application data in Packet Probes (Section 4.1 of [I-D.ietf-tsvwg-datagram-plpmtud] discusses the merits and demerits of this approach). A Packet Probe 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 include an Echo Request Option or Timestamp Option (Section 5.9 [I-D.ietf-tsvwg-udp-options]).

If the application/transport needs protection from the loss of data in the Packet Probe payload, 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].

3.3. Validating the Path with UDP Options

A PL needs to validate that the path continues to support the PLPMTU discovered in a previous search for a suitable PLPMTU value (see Section 6.1.4 of [I-D.ietf-tsvwg-datagram-plpmtud]). This could be provided by an upper layer protocol confirming correct reception of data by the remote PL, but there is no generic mechanism to provide this in UDP Options, and therefore requires generation of a Packet Probe of size PLPMTU to confirm the path. This Packet Probe could use either the ECHO Response Option or the TimeStamp option to elicit a response from the remote PL.

3.3.1. Sending Packet Probes using Timestamps

Reception of a valid Timestamps Option echoed by the remote endpoint can be used to infer connectivity and that packets of the current size are being received by the remote PL. 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 (see Section 2.1 of [RFC2923]).

When sending a probe to increase the PLPMTU, a Timestamp might be unable to unambiguously identify that a specific Packet Probe has been received [KP87]. Timestamp mechanisms therefore cannot be used to confirm the reception of individual probe messages and cannot be used to stimulate a response from the remote peer. Packet Probes used to search for a larger PLPMTU MUST include the Echo Request Option.

3.4. PTB Message Handling for this Method

A UDP Options sender MAY ignore received ICMP PTB messages.

A UDP Options sender that utilises ICMP PTB messages received to a Packet Probe MUST use the quoted packet to validate the UDP port information in combination with the token and/or timestamp value contained in the UDP Option, before processing the packet using the DPLPMTUD method (Section 4.4.1 of [I-D.ietf-tsvwg-datagram-plpmtud]). An implementation unable to support this validation needs to ignore received ICMP PTB messages.

As in other implementations of DPLPMTUD, a PL implementing this specification MUST suspend processing of ICMP PTB by the network layer (as specified in PMTUD [RFC1191] [RFC8201]) for each flow utilising DPLPMTUD.

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 using UDP Options 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.

The security considerations for using DPLPMTUD are described in [I-D.ietf-tsvwg-datagram-plpmtud]. The proposed new method does not change the ICMP PTB message validation method described DPLPMTUD.

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>.

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

[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>.

[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

[KP87] Karn, P. and C. Partridge, "Improving Round-Trip Time Estimates in Reliable Transport Protocols", 1987.

[RFC1063] Mogul, J., Kent, C., Partridge, C., and K. McCloghrie, "IP MTU discovery options", RFC 1063, DOI 10.17487/RFC1063, July 1988, <https://www.rfc-editor.org/info/rfc1063>.

[RFC2923] Lahey, K., "TCP Problems with Path MTU Discovery", RFC 2923, DOI 10.17487/RFC2923, September 2000, <https://www.rfc-editor.org/info/rfc2923>.

[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>.

[RFC8201] McCann, J., Deering, S., Mogul, J., and R. Hinden, Ed., "Path MTU Discovery for IP version 6", STD 87, RFC 8201, DOI 10.17487/RFC8201, July 2017, <https://www.rfc-editor.org/info/rfc8201>.

[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.

Individual draft-01.

- Address Nits

- Change Probe Request and Probe Reponse options to Echo to align names with draft-ietf-tsvwg-udp-options

- Remove Appendix B, Informative Description of new UDP Options

- Add additional sections around Packet Probe generation

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