TSGWG
Internet Draft
Intended status: Experimental
Expires: January 2016

J. Touch
USC/ISI

July 21, 2015

Transport Options for UDP
draft-touch-tsvwg-udp-options-01.txt

Status of this Memo

Copyright Notice

Abstract

Transport protocols are extended through the use of transport header options. This document experimentally extends UDP to provide a location, syntax, and semantics for transport layer options.

Table of Contents

1. Introduction

Transport protocols use options as a way to extend their capabilities. TCP [RFC793], SCTP [RFC4960], and DCCP [RFC4340] include space for these options but UDP [RFC768] currently does not. This document defines an experimental extension to UDP that provides space for transport options including their generic syntax and semantics for their use in UDP's stateless, unreliable message protocol.

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].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lowercase uses of these words are not to be interpreted as carrying significance described in RFC 2119.

In this document, the characters ">>" preceding an indented line(s) indicates a statement using the key words listed above. This convention aids reviewers in quickly identifying or finding the portions of this RFC covered by these key words.

3. Background

Many protocols include a default header and an area for header options. These options enable the protocol to be extended for use in particular environments or in ways unforeseen by the original designers. Examples include TCP's Maximum Segment Size, Window Scale, Timestamp, and Authentication Options [RFC793][RFC5925][RFC7323].

These options are used both in stateful (connection-oriented, e.g., TCP [RFC793], SCTP [RFC4960], DCCP [RFC4340]) and stateless (connectionless, e.g., IPv4 [RFC791], IPv6 [RFC2460] protocols. In stateful protocols they can help extend the way in which state is managed. In stateless protocols their effect is often limited to individual packets, but they can have an aggregate effect on a sequence as well. One example of such uses is Substrate Protocol for User Datagrams (SPUD) [Tr15], and this document is intended to provide an out-of-band option area as an alternative to the in-band mechanism currently proposed [Hi15].

UDP is one of the most popular protocols that lacks space for options [RFC768]. The UDP header was intended to be a minimal addition to IP, providing only ports and a data checksum for protection. This document experimentally extends UDP to provide a trailer area for options located after the UDP data payload.

4. The UDP Option Area

The UDP transport header includes demultiplexing and service identification (port numbers), a checksum, and a field that indicates the payload length. This length field is typically redundant with total IP datagram length and header length.

For IPv4, the total datagram length (including IP header) is the "Total Length" field and the header and its options are 4*IHL ("Internet Header Length") [RFC791]. For IPv6, the last IP option with "Next Header" = UDP (i.e., 17) indicates the size of the transport payload as its "Payload Length" directly [RFC2460]. In both cases, the space available for the UDP transport protocol data unit is indicated by IP

As a result of this redundancy, the UDP length field can be used in other ways. UDP-Lite uses this field to indicate UDP checksum coverage. This document uses this field to create a place for UDP transport options.

The UDP option area is defined as the location between the end of the UDP payload (as indicated by UDP length) and the end of the IP datagram (as indicated by the IP length and IP header length), i.e., as a trailing options area. This area can occur at any valid byte offset, i.e., it need not be 16-bit or 32-bit aligned. In effect, this document redefines the UDP "Length" field as a "trailer offset".

UDP options are defined using a syntax similar to that of TCP [RFC793]. They are typically a minimum of two bytes in length as shown in Figure 1, excepting only the one byte options "No Operation" (NOP) and "End of Options List" (EOL) described below.

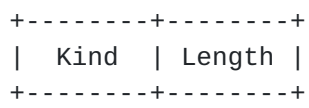

Figure 1 UDP option default format

>> UDP options MAY occur at any UDP length offset.

>> The UDP length MUST be at least as large as the UDP header (8) and no larger than the payload of the IP datagram (IPlen - IPhdrlen). Values outside this range MUST be silently discarded as invalid and logged where rate-limiting permits.

>> UDP options MUST be interpreted in the order in which they occur in the UDP option area.

The following UDP options are currently defined:

Kind	Length	Meaning
0	-	End of Options List
1	-	No operation
128-253		RESERVED
254	N(>=4)	RFC 3692-style experiments
255		RESERVED

>> NOP options SHOULD be used at the beginning of the UDP options area to achieve 32-bit alignment for active (i.e., non-NOP) options.

>> When the UDP options do not consume the entire option area, the last non-NOP option SHOULD be EOL.

>> All bytes after EOL MUST be ignored by UDP option processing.

Note that Kind=254 is reserved for experiments [RFC3692]. Only one such value is reserved because it experiments are expected to already apply the shared use approach developed for TCP experimental options [RFC6994].

>> The length of the experimental option MUST be at least 4 to account for the Kind, Length, and the minimum 16-bit UDP ExID identifier (similar to TCP ExIDs [RFC6994]).

5. UDP options vs. UDP-Lite

UDP Lite provides partial checksum coverage, so that packets with errors in some locations can be delivered to the user [RFC3828]. It uses a different transport protocol number (136) than UDP (17) to interpret the UDP length field as the prefix covered by the UDP checksum.

UDP already defines the UDP length field as the limit of the UDP checksum but that would also limit the data provided to the user (application). A goal of UDP-Lite is to deliver data beyond that length offset, which is why a separate transport protocol number was required.

UDP options do not need a separate transport protocol number because the data beyond the UDP length offset is never provided to the user. It is interpreted exclusively within the UDP transport layer.

6. Options in a Stateless, Unreliable Transport Protocol

There are two ways to interpret options for a stateless, unreliable protocol -- an option is either local to the message or intended to affect a stream of messages in a soft-state manner. Either interpretation is valid for defined UDP options.

It is impossible to know in advance whether an endpoint supports a UDP option.

>> Options MUST allow for silent failure on first receipt.

>> Options that rely on soft-state exchange MUST allow for message reordering and loss.

>> A UDP option MUST be silently optional until confirmed by exchange with an endpoint.

(I'm sure there will be more here)

7. Security Considerations

(to be addressed)

8. IANA Considerations

Upon publication, IANA is hereby requested to create a new registry for UDP Option Kind numbers, similar to that for TCP Option Kinds. Values in this registry are to be assigned by IESG Approval or Standards Action [RFC5226].

Upon publication, IANA is hereby requested to create a new registry for UDP Experimental Option Experiment Identifiers (UDP ExIDs) for use in the same manner as [RFC6994]. Values in this registry are to be assigned using first-come, first-served (FCFS) rules [RFC5226].

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

9.2. Informative References

[Hi15] Hildebrand, J., B. Trammel, "Substrate Protocol for User Datagrams (SPUD) Prototype," draft-hildebrand-spud-prototype-03, Mar. 2015.

[RFC768] Postel, J., "User Datagram Protocol", RFC 768, August 1980.

[RFC791] Postel, J., "Internet Protocol," RFC 791, Sept. 1981.

[RFC793] Postel, J., "Transmission Control Protocol" RFC 793, September 1981

[RFC2460] Deering, S., R. Hinden, "Internet Protocol Version 6 (IPv6) Specification," RFC 2460, Dec. 1998.

[RFC4340] Kohler, E., M. Handley, and S. Floyd, "Datagram Congestion Control Protocol (DCCP)", RFC 4340, March 2006.

[RFC4960] Stewart, R. (Ed.), "Stream Control Transmission Protocol", RFC 4960, September 2007.

[RFC3692] Narten, T., "Assigning Experimental and Testing Numbers Considered Useful," RFC 3692, Jan. 2004.

[RFC3828] Larzon, L-A., M. Degermark, S. Pink, L-E. Jonsson (Ed.), G. Fairhurst (Ed.), "The Lightweight User Datagram Protocol (UDP-Lite)," RFC 3828, July 2004.

[RFC5226] Narten, T., H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs," RFC 5226, May 2008.

[RFC5925] Touch, J., A. Mankin, R. Bonica, "The TCP Authentication Option," RFC 5925, June 2010.

[RFC6994] Touch, J., "Shared Use of Experimental TCP Options," RFC 6994, Aug. 2013.

[RFC7323] Borman, D., R. Braden, V. Jacobson, R. Scheffenegger (Ed.), "TCP Extensions for High Performance," RFC 7323, Sep. 2014.

[Tr15] Trammel, B. (Ed.), M. Kuelewind (Ed.), "Requirements for the design of a Substrate Protocol for User Datagrams (SPUD)," draft-trammell-spud-req-00, July 2015.

10. Acknowledgments

This work benefitted from feedback from Ken Calvert, Ted Faber, and Gorry Fairhurst, as well as discussions on the IETF SPUD email list.

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

Authors' Addresses

Joe Touch
USC/ISI
4676 Admiralty Way
Marina del Rey, CA 90292 USA

Phone: +1 (310) 448-9151
Email: touch@isi.edu