DNSOP Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 22, 2017

R. Bellis
ISC
S. Cheshire
Apple Inc.
J. Dickinson
S. Dickinson
Sinodun
A. Mankin
Salesforce
T. Pusateri
Unaffiliated
July 21, 2016

DNS Session Signaling
draft-bellis-dnsop-session-signal-01

Abstract

The Extension Mechanisms for DNS (EDNS(0)) [RFC6891] is explicitly defined to only have "per-message" semantics. This document defines a new Session Signaling OpCode used to carry persistent "per-session" type-length-values (TLVs), and defines an initial set of TLVs used to manage session timeouts and termination.

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 http://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 22, 2017.

Copyright Notice

Table of Contents

1. Introduction 2
2. Terminology 3
3. Protocol Details 3
3.1. Message Format 4
3.2. Message Handling 4
3.3. TLV Format 5
4. Mandatory TLVs 6
4.1. Session Management Support TLVs 6
4.1.1. "Not Implemented" 6
4.2. Session Management TLVs 6
4.2.1. Start Session 6
4.2.2. Terminate Session 6
4.2.3. Idle Timeout 7
5. IANA Considerations 7
5.1. DNS Session Signaling Opcode Registration 8
5.2. DNS Session Signaling Type Codes Registry 8
6. Security Considerations 8
7. Acknowledgements 9
8. References 9
8.1. Normative References 9
8.2. Informative References 9
Authors' Addresses 9

1. Introduction

The Extension Mechanisms for DNS (EDNS(0)) [RFC6891] is explicitly defined to only have "per-message" semantics. This document defines a new Session Signaling OpCode used to carry persistent "per-session" type-length-values (TLVs), and defines an initial set of TLVs used to manage session timeouts and termination.

A further issue with EDNS(0) is that there is no standard mechanism for a client to be able to tell whether a server has processed or otherwise acted upon the individual options contained with an OPT RR. The Session Signaling OpCode therefore requires an explicit response to each request message.

It should be noted that the message format (see Section 3.1) does not conform to the standard DNS packet format.

2. Terminology

The terms "initiator" and "responder" correspond respectively to the initial sender and subsequent receiver of a Session Signaling TLV, regardless of which was the "client" and "server" in the usual DNS sense. The term "sender" may apply to either an initiator or responder.

The term "session" in the context of this document means the exchange of DNS messages over a single connection using an end-to-end transport protocol where:

- connections can be long-lived
- either end of the connection may initiate requests
- message delivery order is guaranteed
- it is guaranteed that the same two endpoints are in communication for the entire lifetime of the session.

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. Protocol Details

Session Signaling messages MUST only be carried in protocols and in environments where a session may be established according to the definition above. Standard DNS over TCP [RFC1035], and DNS over TLS [RFC7858] are appropriate protocols. DNS over plain UDP is not appropriate since it fails on both the bi-directional initiation requirement and the message order delivery requirement.

Session Signaling messages relate only to the specific session in which they are being carried. Where a middle box (e.g. a DNS proxy, forwarder, or session multiplexer) is in the path the message MUST NOT be blindly forwarded in either direction by that middle box. This does not preclude the use of these messages in the presence of a NAT box that rewrites Layer 3 or Layer 4 headers but otherwise maintains the effect of a single session.

A server MUST NOT initiate Session Signaling messages until a client-initiated Session Signaling message is received first. This

requirement is to ensure that the client does not observe unsolicited inbound messages until it has indicated its ability to handle them.

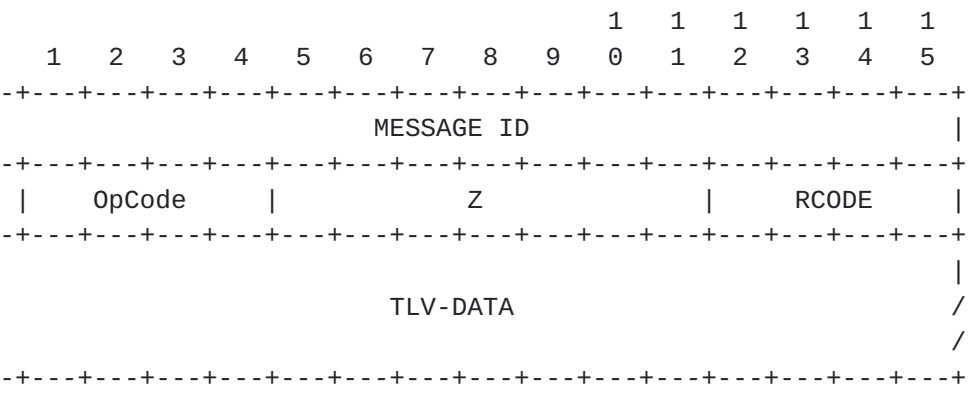
Session Signaling support is therefore said to be confirmed from the client's point of view after the first session signaling TLV has been sent by that client and subsequently successfully acknowledged by the server.

Use of Session Signaling by a client should be taken as an implicit request for a long-lived session.

3.1. Message Format

A message containing a Session Signaling OpCode does not conform to the usual DNS message format. The 4 octet header format from [RFC1035] is however preserved, since that includes the message ID and OpCode and RCODE fields, and the QR bit that differentiates requests from responses.

Each message MUST contain only a single TLV.



The MESSAGE ID, QR, OpCode and RCODE fields have their usual meaning as defined in [RFC1035].

The Z bits are currently unused, and SHOULD be set to zero (0) in requests and responses unless re-defined in a later specification.

3.2. Message Handling

Both clients and servers may unilaterally send Session Signaling messages at any point in the lifetime of a session and are therefore considered to be the initiator with respect to that message. The initiator MUST set the value of the QR bit in the DNS header to zero (0), and the responder MUST set it to one (1).

Every Session Signaling request message MUST elicit a response (which MUST have the same ID in the DNS message header as in the request).

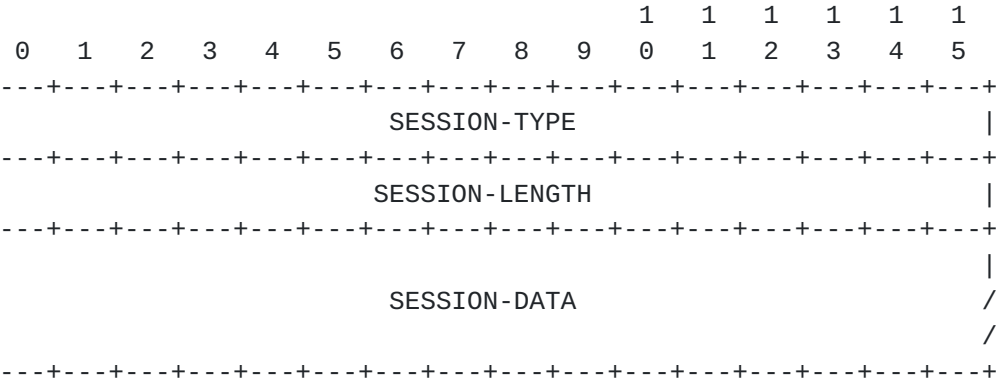
In order to preserve the correct sequence of state, Session Signaling requests MUST NOT be processed out of order.

<< RB: should the presence of a SS message create a "sequencing point", such that all pending responses must be answered? >>

The RCODE value in a response uses a subset of the standard (non-extended) RCODE values from the IANA DNS RCODEs registry, interpreted as follows:

Code	Mnemonic	Description
0	NOERROR	TLV processed successfully
1	FORMERR	TLV format error
4	NOTIMP	Session Signaling not supported
5	REFUSED	TLV declined for policy reasons

3.3. TLV Format



SESSION-TYPE: A 16 bit field in network order giving the type of the current Session Signaling TLV per the IANA DNS Session Signaling Type Codes Registry.

SESSION-LENGTH: A 16 bit field in network order giving the size in octets of SESSION-DATA.

SESSION-DATA: Type-code specific.

4. Mandatory TLVs

4.1. Session Management Support TLVs

4.1.1. "Not Implemented"

Since the "NOTIMP" RCODE is required to indicate lack of support for the Session Signaling OpCode itself, the "Not Implemented" TLV (0) MUST be returned in response to a TLV that is not implemented by the responder.

This TLV has no SESSION-DATA.

4.2. Session Management TLVs

4.2.1. Start Session

The Start Session TLV (1) SHOULD be used by a client to indicate support for Session Signaling. It MUST NOT be initiated by a server.

It is not required that this TLV be used in every session - any valid client-initiated TLV will suffice to initiate Session Signaling support. The intention of this TLV is to provide a suitable "No-Op" TLV to permit Session Signaling support to be negotiated without carrying any other information.

This TLV has no SESSION-DATA.

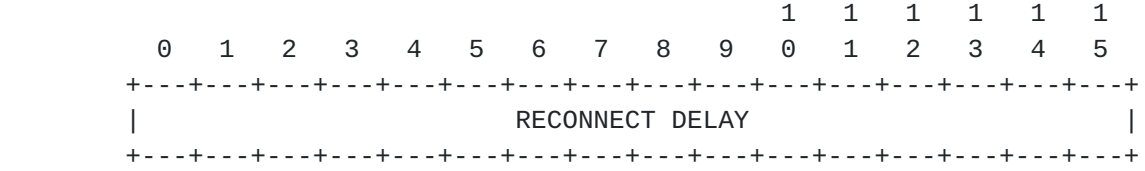
<< RB: this could perhaps also be used as a real "no-op" message to provide application-level keep-alive pings >>

4.2.2. Terminate Session

The Terminate Session TLV (2) MAY be sent by a server to request that the client terminate the session. It MUST NOT be initiated by a client.

The client SHOULD terminate the session as soon as possible, but MAY wait for any inflight queries to be answered. It MUST NOT initiate any new requests over the existing session.

The SESSION-DATA is as follows:



RECONNECT DELAY: a time value, specified as a 16 bit word in network order in units of 100 milliseconds, within which the client MUST NOT establish a new session to the current server.

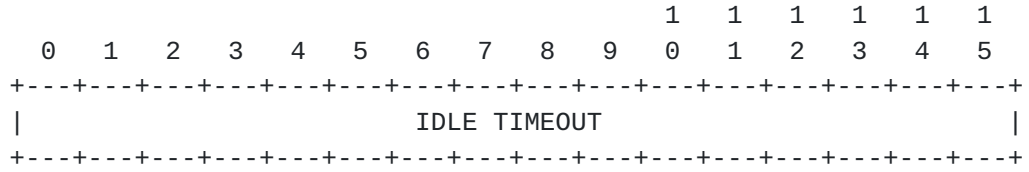
The RECOMMENDED value is 10 seconds. << RB: text required here about default values for load balancers, etc >>

4.2.3. Idle Timeout

The Idle Timeout TLV (3) has similar intent to the EDNS TCP Keepalive Option [RFC7828]. It is used by a server to tell the client how long it may leave the current session idle for. a client. The definition of an idle session is as specified in [RFC7766].

Messages generate by the client have no SESSION-DATA (whether in requests or responses). A client-initiated Idle Timeout TLV allows the client to request the current timeout value, whereas a server-initiated request allows the server to unilaterally update the current timeout value.

Messages generated by the server contain SESSION-DATA as follows:



IDLE TIMEOUT: the idle timeout for the current session, specified as a 16 bit word in network order in units of 100 milliseconds.

The client SHOULD terminate the current session if it remains idle for longer than the specified timeout (and MAY of course terminate the session earlier). The server MAY unilaterally terminate the connection at any time, but SHOULD allow the client to keep the connection open if further messages are received before the idle timeout expires.

A client / server pair that supports Session Signaling MUST NOT use the EDNS TCP KeepAlive option within any message once bi-directional Session Signaling support has been confirmed.

5. IANA Considerations

5.1. DNS Session Signaling Opcode Registration

IANA are directed to assign the value TBD for the Session Signaling OpCode in the DNS OpCodes Registry.

5.2. DNS Session Signaling Type Codes Registry

IANA are directed to create the DNS Session Signaling Type Codes Registry, with initial values as follows:

Type	Name	Status	Reference
0	Not implemented		RFC-TBD1
1	Start Session	Standard	RFC-TBD1
2	Terminate Session	Standard	RFC-TBD1
3	Idle Timeout	Standard	RFC-TBD1
4 - 63	Unassigned, reserved for session management TLVs		
64 - 63487	Unassigned		
63488 - 64511	Reserved for local / experimental use		
64512 - 65535	Reserved for future expansion		

Registration of additional Session Signaling Type Codes requires Expert Review. << RB: definition of process required? >>

6. Security Considerations

If this mechanism is to be used with DNS over TLS, then these messages are subject to the same constraints as any other DNS over TLS messages and MUST NOT be sent in the clear before the TLS session is established.

7. Acknowledgements

TBW

8. References

8.1. Normative References

[RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, DOI 10.17487/RFC1035, November 1987, <http://www.rfc-editor.org/info/rfc1035>.

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

[RFC6891] Damas, J., Graff, M., and P. Vixie, "Extension Mechanisms for DNS (EDNS(0))", STD 75, RFC 6891, DOI 10.17487/RFC6891, April 2013, <http://www.rfc-editor.org/info/rfc6891>.

[RFC7766] Dickinson, J., Dickinson, S., Bellis, R., Mankin, A., and D. Wessels, "DNS Transport over TCP - Implementation Requirements", RFC 7766, DOI 10.17487/RFC7766, March 2016, <http://www.rfc-editor.org/info/rfc7766>.

[RFC7828] Wouters, P., Abley, J., Dickinson, S., and R. Bellis, "The edns-tcp-keepalive EDNS0 Option", RFC 7828, DOI 10.17487/RFC7828, April 2016, <http://www.rfc-editor.org/info/rfc7828>.

8.2. Informative References

[RFC7858] Hu, Z., Zhu, L., Heidemann, J., Mankin, A., Wessels, D., and P. Hoffman, "Specification for DNS over Transport Layer Security (TLS)", RFC 7858, DOI 10.17487/RFC7858, May 2016, <http://www.rfc-editor.org/info/rfc7858>.

Authors' Addresses

Ray Bellis
Internet Systems Consortium, Inc.
950 Charter Street
Redwood City CA 94063
USA

Phone: +1 650 423 1200
Email: ray@isc.org

Stuart Cheshire
Apple Inc.
1 Infinite Loop
Cupertino CA 95014
USA

Phone: +1 408 974 3207
Email: cheshire@apple.com

John Dickinson
Sinodun Internet Technologies
Magadalen Centre
Oxford Science Park
Oxford OX4 4GA
United Kingdom

Email: jad@sinodun.com

Sara Dickinson
Sinodun Internet Technologies
Magadalen Centre
Oxford Science Park
Oxford OX4 4GA
United Kingdom

Email: sara@sinodun.com

Allison Mankin
Salesforce

Email: allison.mankin@gmail.com

Tom Pusateri
Unaffiliated

Phone: +1 843 473 7394
Email: pusateri@bangj.com