DNSEXT Working Group
INTERNET-DRAFT
<draft-vixie-dnsext-edns2-00.txt>

Paul Vixie
ISC
August, 2002

Extensions to DNS (EDNS2)

Status of this Memo

Abstract

This document specifies a number of extensions within the Extended DNS framework defined by [RFC2671] and [EDNS1], including several new extended label types.

1 - Rationale and Scope

1.1. EDNS (see [RFC2671]) specifies an extension mechanism to DNS (see [RFC1035]) which provides for larger message sizes, additional label types, and new message flags.

1.2. This document makes use of the EDNS extension mechanisms to add several new extended label types and message options.

2 - Affected Protocol Elements

2.1. Compression pointers are 14 bits in size and are relative to the start of the DNS Message, which can be 64KB in length. 14 bits restrict pointers to the first 16KB of the message, which makes labels introduced in the last 48KB of the message unreachable by compression pointers. A longer pointer format is needed.

2.2. DNS Messages are limited to 65535 octets in size when sent over TCP. This acts as an effective maximum on RRset size, since multiple TCP messages are only possible in the case of zone transfers. Some mechanism must be created to allow normal DNS responses (other than zone transfers) to span multiple DNS Messages when TCP is used.

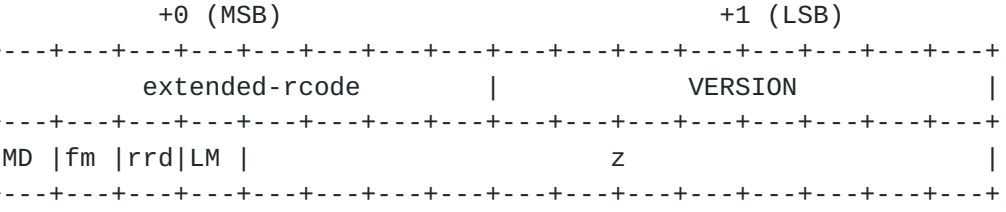
3 - Extended Label Types

**3.1. In [EDNS0], the ``0 1'' label type was specified to denote an** extended label type, whose value is encoded in the lower six bits of the first octet of a label, and an extended label type of ``1 1 1 1 1 1'' was further reserved for use in future multibyte extended label types.

3.2. The ``0 0 0 0 0 0'' extended label type will indicate an extended compression pointer, such that the following two octets comprise a 16-bit compression pointer in network byte order. Like the normal compression pointer, this pointer is relative to the start of the DNS Message.

4 - OPT pseudo-RR Flags and Options

4.1. The extended RCODE and flags are structured as follows:



VERSION Indicates the implementation level of whoever sets it. Full conformance with the draft standard version of this specification is version ``2.'' Note that both requestors and responders should set this to the highest level they implement, that responders should send back RCODE=BADVERS and that requestors should be prepared to probe using lower version

numbers if they receive an RCODE=BADVERS.

MD ``More data'' flag. Valid only in TCP streams where message ordering and reliability are guaranteed. This flag indicates that the current message is not the complete request or response, and should be aggregated with the following message(s) before being considered complete. Such messages are called ``segmented.'' It is an error for the RCODE (including the EXTENDED-RCODE), AA flag, or DNS Message ID to differ among segments of a segmented message. It is an error for TC to be set on any message of a segmented message. Any given RR must fit completely within a message, and all messages will both begin and end on RR boundaries. Each section in a multipart message must appear in normal message order, and each section must be complete before later sections are added. All segments of a message must be transmitted contiguously, without interleaving of other messages.

LM ``Longest match'' flag. If this flag is present in a query message, then for any question whose QNAME is not fully matched by zone or cache data, the longest trailing label-bounded suffix of the QNAME for which zone or cache data is present will be eligible for use as an answer. Note that an intervening wildcard name shall supercede this behaviour and the rules described in [RFC1034 4.3.2, 4.3.3] shall apply, except that the owner name of the answer will be the wildcard name rather than the QNAME. Any of: QTYPE=ANY, or QCLASS=ANY, or QCOUNT>1, shall be considered an error if the LM flag is set.

If LM is set in a request, then LM has meaning in the response as follows: If the content of the response would have been different without the LM flag being set on the request, then the response LM will be set; If the content of the response was not determined or affected by the request LM, then the response LM will be cleared. If the request LM was not set, then the response LM is not meaningful and should be set to zero by responders and ignored by requestors.

Z Set to zero by senders and ignored by receivers, unless modified in a subsequent specification.

5 - References

[RFC1035] P. Mockapetris, ``Domain Names - Implementation and Specification,'' RFC 1035, USC/Information Sciences Institute, November 1987.

[RFC2671] P. Vixie, ``Extension mechanisms for DNS (EDNS0),'' RFC 2671, IETF DNSIND, September 1998

6 - Author's Address

Paul Vixie
Internet Software Consortium
950 Charter Street
Redwood City, CA 94063
+1.650.779.7001
<vixie@isc.org>