Network Working Group
Request for Comments: 3226
Updates: 2874, 2535
Category: Standards Track

O. Gudmundsson

December 2001

DNSSEC and IPv6 A6 aware server/resolver message size requirements

Status of this Memo

Copyright Notice

Abstract

This document mandates support for EDNS0 (Extension Mechanisms for DNS) in DNS entities claiming to support either DNS Security Extensions or A6 records. This requirement is necessary because these new features increase the size of DNS messages. If EDNS0 is not supported fall back to TCP will happen, having a detrimental impact on query latency and DNS server load. This document updates RFC 2535 and RFC 2874, by adding new requirements.

1. Introduction

Familiarity with the DNS [RFC1034, RFC1035], DNS Security Extensions [RFC2535], EDNS0 [RFC2671] and A6 [RFC2874] is helpful.

STD 13, RFC 1035 Section 2.3.4 requires that DNS messages over UDP have a data payload of 512 octets or less. Most DNS software today will not accept larger UDP datagrams. Any answer that requires more than 512 octets, results in a partial and sometimes useless reply with the Truncation Bit set; in most cases the requester will then retry using TCP. Furthermore, server delivery of truncated responses varies widely and resolver handling of these responses also varies, leading to additional inefficiencies in handling truncation.

Compared to UDP, TCP is an expensive protocol to use for a simple transaction like DNS: a TCP connection requires 5 packets for setup and tear down, excluding data packets, thus requiring at least 3 round trips on top of the one for the original UDP query. The DNS

server also needs to keep a state of the connection during this transaction. Many DNS servers answer thousands of queries per second, requiring them to use TCP will cause significant overhead and delays.

1.1. Requirements

The key words "MUST", "REQUIRED", "SHOULD", "RECOMMENDED", and "MAY" in this document are to be interpreted as described in RFC 2119.

2. Motivating factors

2.1. DNSSEC motivations

DNSSEC [RFC2535] secures DNS by adding a Public Key signature on each RR set. These signatures range in size from about 80 octets to 800 octets, most are going to be in the range of 80 to 200 octets. The addition of signatures on each or most RR sets in an answer significantly increases the size of DNS answers from secure zones.

For performance reasons and to reduce load on DNS servers, it is important that security aware servers and resolvers get all the data in Answer and Authority section in one query without truncation. Sending Additional Data in the same query is helpful when the server is authoritative for the data, and this reduces round trips.

DNSSEC OK[OK] specifies how a client can, using EDNS0, indicate that it is interested in receiving DNSSEC records. The OK bit does not eliminate the need for large answers for DNSSEC capable clients.

2.1.1. Message authentication or TSIG motivation

TSIG [RFC2845] allows for the light weight authentication of DNS messages, but increases the size of the messages by at least 70 octets. DNSSEC specifies for computationally expensive message authentication SIG(0) using a standard public key signature. As only one TSIG or SIG(0) can be attached to each DNS answer the size increase of message authentication is not significant, but may still lead to a truncation.

2.2. IPv6 Motivations

IPv6 addresses [RFC2874] are 128 bits and can be represented in the DNS by multiple A6 records, each consisting of a domain name and a bit field. The domain name refers to an address prefix that may require additional A6 RRs to be included in the answer. Answers where the queried name has multiple A6 addresses may overflow a 512-octet UDP packet size.

2.3. Root server and TLD server motivations

The current number of root servers is limited to 13 as that is the maximum number of name servers and their address records that fit in one 512-octet answer for a SOA record. If root servers start advertising A6 or KEY records then the answer for the root NS records will not fit in a single 512-octet DNS message, resulting in a large number of TCP query connections to the root servers. Even if all client resolver query their local name server for information, there are millions of these servers. Each name server must periodically update its information about the high level servers.

For redundancy, latency and load balancing reasons, large numbers of DNS servers are required for some zones. Since the root zone is used by the entire net, it is important to have as many servers as possible. Large TLDs (and many high-visibility SLDs) often have enough servers that either A6 or KEY records would cause the NS response to overflow the 512 byte limit. Note that these zones with large numbers of servers are often exactly those zones that are critical to network operation and that already sustain fairly high loads.

2.4. UDP vs TCP for DNS messages

Given all these factors, it is essential that any implementation that supports DNSSEC and or A6 be able to use larger DNS messages than 512 octets.

The original 512 restriction was put in place to reduce the probability of fragmentation of DNS responses. A fragmented UDP message that suffers a loss of one of the fragments renders the answer useless and the query must be retried. A TCP connection requires a larger number of round trips for establishment, data transfer and tear down, but only the lost data segments are retransmitted.

In the early days a number of IP implementations did not handle fragmentation well, but all modern operating systems have overcome that issue thus sending fragmented messages is fine from that standpoint. The open issue is the effect of losses on fragmented messages. If connection has high loss ratio only TCP will allow reliable transfer of DNS data, most links have low loss ratios thus sending fragmented UDP packet in one round trip is better than establishing a TCP connection to transfer a few thousand octets.

2.5. EDNS0 and large UDP messages

EDNS0 [RFC2671] allows clients to declare the maximum size of UDP message they are willing to handle. Thus, if the expected answer is between 512 octets and the maximum size that the client can accept, the additional overhead of a TCP connection can be avoided.

3. Protocol changes:

This document updates RFC 2535 and RFC 2874, by adding new requirements.

All RFC 2535 compliant servers and resolvers MUST support EDNS0 and advertise message size of at least 1220 octets, but SHOULD advertise message size of 4000. This value might be too low to get full answers for high level servers and successor of this document may require a larger value.

All RFC 2874 compliant servers and resolver MUST support EDNS0 and advertise message size of at least 1024 octets, but SHOULD advertise message size of 2048. The IPv6 datagrams should be 1024 octets, unless the MTU of the path is known. (Note that this is smaller than the minimum IPv6 MTU to allow for some extension headers and/or encapsulation without exceeding the minimum MTU.)

All RFC 2535 and RFC 2874 compliant entities MUST be able to handle fragmented IPv4 and IPv6 UDP packets.

All hosts supporting both RFC 2535 and RFC 2874 MUST use the larger required value in EDNS0 advertisements.

4. Acknowledgments

Harald Alvestrand, Rob Austein, Randy Bush, David Conrad, Andreas Gustafsson, Jun-ichiro itojun Hagino, Bob Halley, Edward Lewis Michael Patton and Kazu Yamamoto were instrumental in motivating and shaping this document.

5. Security Considerations:

There are no additional security considerations other than those in RFC 2671.

6. IANA Considerations:

None

7. References

[RFC1034] Mockapetris, P., "Domain Names - Concepts and Facilities", STD 13, RFC 1034, November 1987.

[RFC1035] Mockapetris, P., "Domain Names - Implementation and Specification", STD 13, RFC 1035, November 1987.

[RFC2535] Eastlake, D. "Domain Name System Security Extensions", RFC 2535, March 1999.

[RFC2671] Vixie, P., "Extension Mechanisms for DNS (EDNS0)", RFC 2671, August 1999.

[RFC2845] Vixie, P., Gudmundsson, O., Eastlake, D. and B. Wellington, "Secret Key Transaction Authentication for DNS (TSIG)", RFC 2845, May 2000.

[RFC2874] Crawford, M. and C. Huitema, "DNS Extensions to Support IPv6 Address Aggregation and Renumbering", RFC 2874, July 2000.

[RFC3225] Conrad, D., "Indicating Resolver Support of DNSSEC", RFC 3225, December 2001.

8. Author Address

Olafur Gudmundsson
3826 Legation Street, NW
Washington, DC 20015
USA

EMail: ogud@ogud.com

9. Full Copyright Statement

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.