Avoid IP fragmentation in DNS
draft-fujiwara-dnsop-avoid-fragmentation-00

Abstract

Path MTU discovery is vulnerable and IP fragmentation may cause protocol weakness. Currently, DNS is said to be the biggest user of IP fragmentation. However, it is possible to avoid IP fragmentation in DNS because TCP transport and truncation work well. This document proposes to avoid IP fragmentation in DNS.

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 January 5, 2020.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.
1. Introduction

DNS has EDNS0 [RFC6891] mechanism. It enables that DNS server can send large size response using UDP. Now EDNS0 is widely deployed, then DNS is said to be the biggest user of IP fragmentation.

However, "Fragmentation Considered Poisonous" [Herzberg2013] proposed effective off-path DNS cache poisoning attack vectors using IP fragmentation. "IP fragmentation attack on DNS" [Hlavacek2013] and "Domain Validation++ For MitM-Resilient PKI" [Brandt2018] proposed that off-path attackers can intervene in path MTU discovery [RFC1191] to perform intentionally fragmented responses from authoritative servers. [RFC7739] stated security implications of predictable fragment identification values.

As a result, we cannot trust fragmented UDP packets and path MTU discovery. By the way, TCP is considered resistant against IP fragmentation attacks because TCP has sequence number and acknowledgement number in each sequence.

This document proposes to avoid IP fragmentation in DNS.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

Many of the specialized terms used in this document are defined in DNS Terminology [RFC8499].
3. Proposal to avoid IP fragmentation in DNS

The methods to avoid IP fragmentation in DNS are described below:

- Full-service resolvers SHOULD set EDNS0 requestor’s UDP payload size to 1220. (defined in [RFC4035] as minimum payload size)
- Authoritative servers and full-service resolvers SHOULD choose EDNS0 responder’s maximum payload size to 1220 (defined in [RFC4035] as minimum payload size)
- Special consideration in small MTU network: If authoritative servers and full-service resolvers are located across the link with the MTU value less than 1280, choose EDNS0 requestor’s and responder’s maximum payload size fit to the smallest link MTU value. (the smallest MTU value minus IPv4 header size and UDP header size.)
- Full-service resolvers MAY drop fragmented UDP responses derived from DNS before IP assembly. (It is a countermeasure against DNS cache poisoning attacks using IP fragmentation.)
- And more, authoritative servers MAY send DNS responses with IP_DONTFRAG / IPV6_DONTFRAG options.

When responses exceed specified EDNS0 size 1220 (or smaller value), servers return truncated responses, and clients retry by TCP. [RFC1035]

4. Incremental deployment

The proposed method supports incremental deployment.

When a full-service resolver implements the proposed method, the full-service resolver becomes to avoid IP fragmentation in DNS.

When an authoritative server implements the proposed method, the authoritative server becomes to avoid IP fragmentation in DNS.

5. Considerations

In past researches ([Fujiwara2018] / dns-operations mailing list discussions), there are some authoritative servers that ignore EDNS0 requestor’s UDP payload size, and return large UDP responses.

And it is known that there are some authoritative servers that do not support TCP transport.
6. IANA Considerations

This document has no IANA actions.

7. Security Considerations

8. References

8.1. Normative References


8.2. Informative References

[Brandt2018]

[Fujiwara2018]

[Herzberg2013]

[Hlavacek2013]

Author’s Address

Kazunori Fujiwara
Japan Registry Services Co., Ltd.
Chiyoda First Bldg. East 13F, 3-8-1 Nishi-Kanda
Chiyoda-ku, Tokyo  101-0065
Japan

Phone: +81 3 5215 8451
Email: fujiwara@jprs.co.jp