draft-fujiwara-dnsop-avoid-fragmentation-00

Kazunori Fujiwara @ IETF 105
Path MTU discovery is vulnerable

• DNS is said to be the biggest user of IP fragmentation.
  • EDNS0 (and DNSSEC) is widely deployed
• Research papers described effective cache poisoning attacks using IP fragmentation and path MTU discovery
  • Fragmentation Considered Poisonous, 2013
  • IP fragmentation attack on DNS, 2013
  • Domain Validation++ For MitM-Resilient PKI, 2018
• As a result, we cannot trust fragmented UDP packets and path MTU discovery
We can avoid large UDP responses

• EDNS0 has requestor's UDP payload size field
  • We can choose smaller value (smaller than path MTU)
  • Note that path MTU, with or without fragmentation, could be smaller than this. (Quoted from Section 6.2.3, RFC 6891)

• Truncation works well
  • When responses exceed specified EDNS0 size, servers return truncated responses, and clients retry by TCP.

• TCP is considered resistant against IP fragmentation attacks
  • RFC 7766 states that all general-purpose DNS implementations MUST support both UDP and TCP
New recommendations

• 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 set EDNS0 responder's maximum payload size to 1220
  • And more, authoritative servers MAY send DNS responses with IP_DONTFRAG / IPV6_DONTFRAG options.

• Full-service resolvers MAY drop fragmented UDP responses derived from DNS before IP reassembly.
  • It is a countermeasure against DNS cache poisoning attacks using IP fragmentation.
Special consideration in small MTU network

• When DNS servers 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/IPv6 header size and UDP header size.

• Or (maybe) another recommendation: DNS servers SHOULD be located at networks where MTU value to the major part of the Internet is larger than or equal to 1280
Deployment

• The proposed method supports incremental deployment.

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

• When an authoritative server implements the proposal, the authoritative server becomes to avoid IP fragmentation in DNS.

• DNSSEC, or TSIG with shared-key require both requestor’s and responder’s support.
Concerns about dropping fragments (not yet written in draft)

• Drop fragmented responses and DNS responses with IP_DONTFRAG / IPV6_DONTFRAG options may cause DNS communication error (timeout)
  • To recover the situation, full-service resolvers need to retry the query by TCP transport
  • It increases complexity of full-service resolvers
How do you consider?

- Do you support this recommendation?
- Do you like fragmentation?