Avoid Fragmentation in DNS

Kazunori Fujiwara @ IETF 106 dnsop WG
RFC 8085: UDP Usage Guidelines
= BCP 145 (March 2017)

• Section 3.2. Message Size Guidelines
  • an application SHOULD NOT send UDP datagrams that result in IP packets that exceed the Maximum Transmission Unit (MTU) along the path to the destination.
  • An application SHOULD either use the path MTU information provided by the IP layer or implement Path MTU Discovery (PMTUD) itself [RFC1191] [RFC1981] [RFC4821] to determine whether the path to a destination will support its desired message size without fragmentation.

• Previous BCP 145: RFC 5405 (Nov 2008) Unicast UDP Usage Guidelines for Application Designers
  • already have same text
  • Then, avoid fragmentation in UDP is effective after Nov. 2008.
“Fragmentation” in DNS RFCs

- RFC 2671, RFC 6891: EDNS0
  - Discuss issues of fragmentation
  - “Note that path MTU, with or without fragmentation, may be smaller than this.”

- RFC 3226: DNSSEC and IPv6 A6 aware server/resolver message size requirements
  - “All RFC 2535 (DNSSEC) and RFC 2874 (IPv6 DNS) compliant entities MUST be able to handle fragmented IPv4 and IPv6 UDP packets.”
  - “MUST support EDNS0 and advertise message size of at least 1220 octets, but SHOULD advertise message size of 4000.”

- RFC 4035: DNSSEC
  - “MUST support a message size of at least 1220 octets, and SHOULD support a message size of 4000 octets”
  - “A security-aware resolver's IP layer MUST handle fragmented UDP packets correctly regardless of whether any such fragmented packets were received via IPv4 or IPv6.”
Maximum UDP message size

- RFC 8085: use the path MTU information provided by the IP layer or implement Path MTU Discovery (PMTUD) itself

- Then,
  - Retrieve Path MTU value to each destination
    - Or use interface MTU value
    - Or minimal MTU 1280 (IPv6), 576 (IPv4)
  - Minus (IP header size + UDP header size)

- Possible magic number (maximum DNS/UDP payload size)
  - $\geq 1220$: RFC 4035 defines at least 1220 octets
  - $\leq 1400$: Most of Internet support MTU 1500 (minus some headers)
  - $\geq 1232$: DNS Flag Day 2020 proposed
Recommendations

• UDP requestors and responders SHOULD send DNS responses with IP_DONTFRAG / IPV6_DONTFRAG
  • Upon a timeout, UDP requestors may retry using TCP or UDP, per local policy

• Responders SHOULD compose UDP responses that result in IP packets that do not exceed the path MTU to the requestor.
  • The estimated maximum DNS/UDP payload size SHOULD be the actual or the default maximum DNS/UDP payload size
  • $1220 \leq$ default maximum DNS/UDP size $\leq 1400$
    • May be 1232

• Zone operator SHOULD consider small response size configurations

• Fragmented DNS/UDP messages may be dropped before IP assembly.

• How to retrieve path MTU value to a destination
  • getsockoptIP_MTU, IPV6_MTU on Linux
Changes from 00 to 01

• New Co-author: Paul Vixie
• Refer RFC 8085 UDP Usage Guidelines
• Updated: Responders SHOULD compose UDP responses that result in IP packets that do not exceed the path MTU to the requestor
• Added: the estimated maximum DNS/UDP payload size SHOULD be the actual or the default maximum DNS/UDP payload size
  • $1220 \leq$ default maximum DNS/UDP size $\leq 1400$
• Added: Zone operator SHOULD consider small response size configurations
• Added: How to retrieve path MTU value to a destination
• Is the draft useful?
• Adopt?