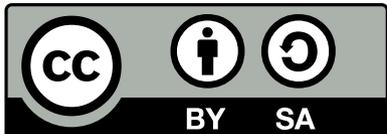


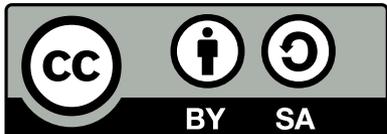
DNS message fragments

Shane Kerr, Davey Song / BII Lab
Mukund Sivaraman / ISC
2015-11-05 / Yokohama, Japan



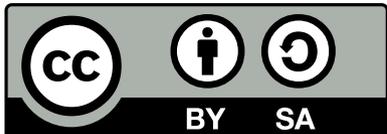
Problem Statement

- IP fragments have issues
 - UDP checksum failures
 - Middleboxes dropping IP fragments
 - Any PMTU discovery failure
 - Timeouts are very costly
- TCP relatively expensive
 - 2 packets worth of data become 3 round trips



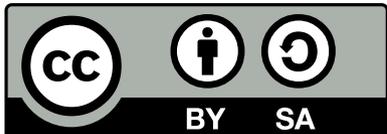
Overview of Protocol

- Resolver puts EDNS option on query
- Authority server processes this
 - If answer > defined size, fragment
 - Each fragment is a DNS message
 - Original DNS message split on RR boundaries
 - Each fragment's DNS header is identical, except counts
 - Including per-fragment name compression
 - EDNS option with fragment count and ID
- Resolver reconstructs original answer
 - Uses normal timeout for answer



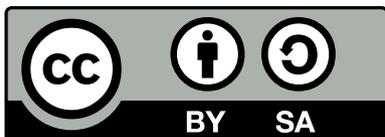
In the Details (1 of 2)

- DNSSEC validation on assembled answer
- Amplification
 - Some small increase in data (5%? 10%?)
 - Cookies? RRL?
- Limit on number of packets
 - Reliability
 - Avoid network disruption



In the Details (2 of 2)

- Increasing fragment sizes
 - IPv4: 512 → 1460 → 1480
 - IPv6: 1280 → 1420 → 1460
 - Allows resolver to infer PMTU



Some Open Issues

- More defined resolver behavior
- TSIG behavior
- RRset splitting
- When to NOT fragment

