Dealing with fragmentation in EDNS0
Proposal for a recommendation

Roland van Rijswijk - Deij
roland.vanrijswijk@surfnet.nl

Sean McAfee (AP)
Problem recap

1. Authoritative Name Server
2. Recursive Caching Name Server (resolver)
3. Firewall
4. Firewall
Extent of the problem

• 9% of all internet hosts may have problems receiving fragmented UDP messages [1];

• 2% – 10% of all resolving name servers experience problems receiving fragmented DNS responses [2]


Solutions

• Resolving name servers should advertise a proper max. response size to avoid fragmentation issues [RFC 2671BIS (DRAFT)];

  Not explicitly stated in standards yet, nor widely implemented;

• Until then: set maximum response size at some authoritative name servers
Resolver experiments (1)
Normal operations

Response time (ms.)

Time (ms.)

Windows Server 2012  Unbound  BIND

Response (>ms.)

281  150  109
388  83  105
687  381  105

SURFnet: we make innovation work
Resolver experiments (2)
Blocking fragments

Response time (ms.) [0/5 altered Authoritative Name Servers]

- Windows Server 2012
  - Time x100+ (!!!)
  - x̅ = 17,787

- Unbound
  - Time x2
  - 3.435

- BIND
  - Time x10 (!)
  - 4.663
  - 2.524

- 1.175
  - 760
  - 465

SURFnet: we make innovation work
Resolver experiments (3)
Max. resp. size on 1 authNS

Response time (ms.) [1/5 altered Authoritative Name Servers]

- Windows Server 2012: Max. = 16,162
- Unbound: 109 ms
- BIND: 1,169 ms

(Responses: 1/5 altered authoritative Name Servers)
Resolver experiments (4)  
Max. resp. size on 2 authNS

Response time (ms.) [2/5 altered Authoritative Name Servers]

- Windows Server 2012: Time x10
  - 290 ms. (3.295)

- Unbound: Time x1.5
  - 99 ms. (1.036)

- BIND: Time x2
  - 126 ms. (1.408)

---

SURFnet: we make innovation work

8
## Experiment on live authNS

<table>
<thead>
<tr>
<th>Traffic (IPv4 + IPv6)</th>
<th>Normal Operations</th>
<th>Max. response size 1232 bytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragmented responses</td>
<td>28.9%</td>
<td>0.0%*</td>
</tr>
<tr>
<td>Fragment receiving resolvers</td>
<td>57.3%</td>
<td>0.0%*</td>
</tr>
<tr>
<td>Truncated UDP responses</td>
<td>0.8%</td>
<td>0.9%</td>
</tr>
<tr>
<td>ICMP FRTE messages</td>
<td>5649/h</td>
<td>&lt; 1/h*</td>
</tr>
<tr>
<td>ICMP FRTE sending resolvers</td>
<td>1.3%</td>
<td>0.0%*</td>
</tr>
<tr>
<td>Total retries</td>
<td>25.8%</td>
<td>25.5%</td>
</tr>
</tbody>
</table>

*Statistically significant difference between experiments
Rise in truncated answers

• Experiment:
  - Querying 995 zones in .com, .edu, .mil, .net and .nl
  - All zones are signed and have a www-node
  - Results:

<table>
<thead>
<tr>
<th>Max. response</th>
<th>A for www</th>
<th>AAAA for www</th>
<th>DNSKEY</th>
</tr>
</thead>
<tbody>
<tr>
<td>4096</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>1472</td>
<td>1.8%</td>
<td>1.8%</td>
<td>8.1%</td>
</tr>
<tr>
<td>1232</td>
<td>2.9%</td>
<td>3.5%</td>
<td>40.0%</td>
</tr>
</tbody>
</table>

Proposed recommendation

1. At least 50% of all authoritative name servers for a zone **SHOULD** be set to limit the overall response size to 1472 bytes, but **MAY** be set as low as 1232 bytes;

2. At least 50% of all in-zone authoritative name servers for a zone **SHOULD** be set to limit the overall response size to 1472 bytes, but **MAY** be set as low as 1232 bytes;

3. Authoritative name servers to which the above recommendations are applied **MUST** accept DNS queries over TCP.