DNS Maze: Testing server selection

Petr Špaček
Štěpán Balážik

petr.spacek@nic.cz
stepan.balazik@nic.cz

(speaking)
(working)
Server selection problem

- Client
- Resolver
- Cache
- Auth 1
- Auth 2
- Auth 3

Cache miss
Server selection problem

- “Which server should I ask next?”
- Optimizing for
  - # of packets
  - Non-predictability
  - RTT for clients
- Messy code
  - Hard to understand (sometimes), arbitrary (?) choice of constants
Server selection testing

- Two sub-problems
- Phase 1 – reproducible test environment
- Phase 2 – test scenarios & data analysis
Part 1

Reproducible test environment
DNS Maze: new tool!

- Open-source and **Work-in-Progress**
- Mock (smallish) Internet!
  - Network environment (IP addr, delay, loss)
  - Authoritative servers
  - Resolver
  - Client
  - Time – if necessary for DNSSEC tests
DNS Maze: design principles

- Minimalistic
- Fast
  - 30 servers in 5 seconds on a laptop
- Flexible
  - Simulating broken behavior
- Mix-and-match existing SW
  - Systemd unit generator
  - Flexible config templating
DNS Maze: scenarios

<table>
<thead>
<tr>
<th>auth</th>
<th>resolver</th>
<th>client</th>
</tr>
</thead>
<tbody>
<tr>
<td>name: &quot;e-root-server&quot;</td>
<td>name: &quot;resolver&quot;</td>
<td>name: &quot;dnsperf&quot;</td>
</tr>
<tr>
<td>service: &quot;knot&quot;</td>
<td>service: &quot;named&quot;</td>
<td>service: &quot;dnsperf&quot;</td>
</tr>
<tr>
<td>config: &quot;knotd.conf&quot;</td>
<td>address: [&quot;1.0.0.50&quot;]</td>
<td>address: [&quot;2.0.0.1&quot;]</td>
</tr>
<tr>
<td>address: [&quot;1.0.0.50&quot;]</td>
<td>config: &quot;named.conf&quot;</td>
<td>dnsperf: -d querylist</td>
</tr>
<tr>
<td>netem: &quot;delay 50ms&quot;</td>
<td>roothints: &quot;root.hints&quot;</td>
<td></td>
</tr>
<tr>
<td>zonefile: &quot;/root.zone&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>domain: &quot;/.&quot;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Config file templates
- Systemd services
- Traffic => PCAP
DNS Maze: Maze of config files :-)

```
git
  ├── configs
  │   ├── knotd.conf.j2
  │   ├── kresd.conf.j2
  │   └── ...
  └── systemd
      ├── dnsperf.service.j2
      │   ├── knot.service.j2
      │   └── named.service.j2
      ├── namespace.service.j2
      └── net.service.j2
  └── process_generator.py
  └── reload_services

scenarios
  └── scenario_name
      └── auth
          ├── root
          │   └── a
          │       └── config.yaml
          └── root.zone
      └── client
          └── config.yaml
          └── querylist
      └── resolver
          └── config.yaml
          └── root.hints
```
IP network simulation

- Linux namespaces – Systemd orchestration
  - [Service]
    PrivateNetwork=yes
    ExecStartPre=/sbin/ip link add dummy0 type dummy
  - [Unit]
    JoinsNamespaceOf=MAZE_NAMESPACE_SERVICE.service
- NetEm – Network Emulator
  - $ tc qdisc ... dev dummy0 add netem OPTIONS
- Gory implementation details
Auth server simulation

- Knot DNS authoritative server
  - Config templates
  - User-provided zones
- Special "test server"
  - \((qname, qtype) \Rightarrow \text{an answer packet}\)
  - Latency/jitter/loss \Rightarrow \text{NetEm rule}
    - delay 68ms 4ms loss random 1.48%
DNS client simulation

- dnsperf
  - Config template
  - User-provided query list
    - 0.aaa A
    - 1.aaa A
    - 2.aaa A
    - 3.aaa A
    - ...

…
Resolve = software-under-test

- Pick your victim :-)
  - Config template
  - libfaketime for DNSSEC tests
- Tested with
  - BIND, Knot Resolver, PowerDNS Recursor,
  - Teachable DNS Resolver ***, Unbound
Running through a maze

- # generate systemd unit files
  
  $ ./process_generator.py /tmp/scenario-bp-simple
  
  • => output /tmp/bp-simple-xe7mwcxf

- # start test (inside systemd machinery)
  
  $ sudo ./reload_services /tmp/tmpname bp-dnsperf
  
  • wait for /tmp/dump.pcap
Part 2

Test scenarios & data analysis
Server selection: theory
Server selection: reality

Source: https://commons.wikimedia.org/wiki/File:Labyrinth.svg CC BY SA 2.5, modified
Test scenarios

- Laboratory
  - Hand-crafted
  - Simulating various errors

- Real-life
  - Generated from RIPE Atlas DNS measurements
  - "Resolver running in networks around the world"
Part 2a
Laboratory
10 x 10 ms steps: # of queries to auth

- Unbound: 100%
- PowerDNS: 100%
- BIND: 101%
- KNOT: 123%
10 x 10 ms steps: BIND packets to auth

95% attempt to contact 10 unique IP addresses

# of queries to auth [%]

10 ms: 95%
20 ms: 0%
30 ms: 1%
40 ms: 0%
50 ms: 0%
60 ms: 0%
70 ms: 1%
80 ms: 1%
90 ms: 1%
100 ms: 1%

latency to auth server
10 x 10 ms steps: KNOT packets to auth

attempt to contact 10 unique IP addresses
10 x 10 ms steps: PowerDNS packets to auth

Attempt to contact 10 unique IP addresses

<table>
<thead>
<tr>
<th>Latency to auth server</th>
<th># of queries to auth [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 ms</td>
<td>98%</td>
</tr>
<tr>
<td>20 ms</td>
<td>0%</td>
</tr>
<tr>
<td>30 ms</td>
<td>0%</td>
</tr>
<tr>
<td>40 ms</td>
<td>0%</td>
</tr>
<tr>
<td>50 ms</td>
<td>0%</td>
</tr>
<tr>
<td>60 ms</td>
<td>0%</td>
</tr>
<tr>
<td>70 ms</td>
<td>0%</td>
</tr>
<tr>
<td>80 ms</td>
<td>0%</td>
</tr>
<tr>
<td>90 ms</td>
<td>0%</td>
</tr>
<tr>
<td>100 ms</td>
<td>1%</td>
</tr>
</tbody>
</table>
10 x 10 ms steps: Unbound packets to auth

attempt to contact 10 unique IP addresses

<table>
<thead>
<tr>
<th>Latency to auth server</th>
<th># of queries to auth [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 ms</td>
<td>10 %</td>
</tr>
<tr>
<td>20 ms</td>
<td>10 %</td>
</tr>
<tr>
<td>30 ms</td>
<td>10 %</td>
</tr>
<tr>
<td>40 ms</td>
<td>10 %</td>
</tr>
<tr>
<td>50 ms</td>
<td>10 %</td>
</tr>
<tr>
<td>60 ms</td>
<td>10 %</td>
</tr>
<tr>
<td>70 ms</td>
<td>10 %</td>
</tr>
<tr>
<td>80 ms</td>
<td>10 %</td>
</tr>
<tr>
<td>90 ms</td>
<td>10 %</td>
</tr>
<tr>
<td>100 ms</td>
<td>10 %</td>
</tr>
</tbody>
</table>
10 x 10 ms steps: latency over time

sliding window = 100 queries

- BIND
- PowerDNS
- Unbound
- KNOT

Average response time [ms]

client query no.
10 x 10 ms steps: latency over time (zoom)

sliding window = 10 queries

Average response time [ms]
- 300 ms
- 200 ms
- 100 ms
- 0 s

client query no.

BIND
PowerDNS
Unbound
KNOT
10 x 10 ms steps: latency histogram
Test stability? 10 rounds of PowerDNS
1 NS timeouts: auth servers
1 NS timeouts: # of queries to auth

- PowerDNS: 101%
- BIND: 101%
- Unbound: 102%
- KNOT: 103%
1 NS timeouts: BIND packets to auth

- attempt to contact 6 unique IP addresses

# of queries to auth [%]

- 0 ms: 22%
- 50 ms: 27%
- 50 ms: 26%
- 50 ms: 24%
- dead: 1%
1 NS timeouts: KNOT packets to auth

- attempt to contact 6 unique IP addresses

- 24% at 0 ms
- 24% at 50 ms
- 24% at 50 ms
- 24% at 50 ms
- 3% at dead

# of queries to auth [%]

latency to auth server
1 NS timeouts: PowerDNS packets to auth

75% attempt to contact 6 unique IP addresses

Latency to auth server:
- 0% (0 ms)
- 75% (50 ms)
- 10% (50 ms)
- 7% (50 ms)
- 7% (50 ms)
- 1% (dead)
1 NS timeouts: Unbound packets to auth

Attempt to contact 6 unique IP addresses

<table>
<thead>
<tr>
<th>Latency to Auth Server</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 ms</td>
<td>0%</td>
</tr>
<tr>
<td>50 ms</td>
<td>25%</td>
</tr>
<tr>
<td>50 ms</td>
<td>24%</td>
</tr>
<tr>
<td>50 ms</td>
<td>25%</td>
</tr>
<tr>
<td>50 ms</td>
<td>25%</td>
</tr>
<tr>
<td>dead</td>
<td>2%</td>
</tr>
</tbody>
</table>
1 NS timeouts: latency over time

sliding window = 100 queries

Average response time [ms]

client query no.
1 NS timeouts: latency over time (zoom)
1 NS timeouts: latency histogram
fast DNSSEC expired, slow OK: auth servers
fast DNSSEC expired, slow OK: # of queries to auth

- Unbound: 151%
- BIND: 199%
- PowerDNS: 200%
- KNOT: 203%

queries from resolver to auth
fast DNSSEC expired, slow OK: BIND packets to auth

attempt to contact 3 unique IP addresses

<table>
<thead>
<tr>
<th># of queries to auth [%]</th>
<th>0 ms</th>
<th>10 ms</th>
<th>100 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 %</td>
<td>50 %</td>
<td>50 %</td>
<td></td>
</tr>
</tbody>
</table>
fast DNSSEC expired, slow OK: KNOT packets to auth

attempt to contact 3 unique IP addresses

# of queries to auth [%]

<table>
<thead>
<tr>
<th>Latency to auth server</th>
<th>% of queries</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 ms</td>
<td>0%</td>
</tr>
<tr>
<td>10 ms</td>
<td>93%</td>
</tr>
<tr>
<td>100 ms</td>
<td>7%</td>
</tr>
</tbody>
</table>
fast DNSSEC expired, slow OK: PowerDNS packets to auth

attempt to contact 3 unique IP addresses

# of queries to auth [%]

0 % 100 %

0 ms 10 ms 100 ms

latency to auth server
fast DNSSEC expired, slow OK: Unbound packets to auth

attempt to contact 3 unique IP addresses

<table>
<thead>
<tr>
<th>Latency to Auth Server</th>
<th>% of Queries to Auth</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 ms</td>
<td>0 %</td>
</tr>
<tr>
<td>10 ms</td>
<td>34 %</td>
</tr>
<tr>
<td>100 ms</td>
<td>66 %</td>
</tr>
</tbody>
</table>
fast DNSSEC expired, slow OK: latency over time

sliding window = 100 queries

Average response time [ms]

BIND
PowerDNS
Unbound
KNOT

0 s
2 k
4 k
6 k
8 k
10 k

client query no.
fast DNSSEC expired, slow OK: latency over time (zoom)
fast DNSSEC expired, slow OK: latency histogram

![Graph showing response time latency histogram for different DNS resolvers: BIND, KNOT, PowerDNS, and Unbound. The graph indicates latency distribution over different percentile thresholds from 1 packet to 10000 packets.]
Part 2b

Real world
Prague: auth servers

RTT to auth server [ms]

1 ± 0 ms
1 ± 0 ms
1 ± 0 ms
1 ± 0 ms
6 ± 5 ms
7 ± 0 ms
8 ± 7 ms
14 ± 0 ms
14 ± 0 ms
21 ± 4 ms
25 ± 45 ms
100 ± 2 ms
122 ± 6 ms
Prague: # of queries to auth

- PowerDNS: 100%
- Unbound: 100%
- BIND: 112%
- KNOT: 113%
Prague: BIND packets to auth

61% attempt to contact 13 unique IP addresses

Latency to auth server:
- 1 ± 0 ms: 33%
- 1 ± 2 ms: 1%
- 6 ± 5 ms: 1%
- 7 ± 0 ms: 0%
- 8 ± 7 ms: 0%
- 14 ± 0 ms: 1%
- 21 ± 4 ms: 0%
- 25 ± 45 ms: 0%
- 100 ± 2 ms: 0%
- 122 ± 6 ms: 0%

# of queries to auth [%]
Prague: KNOT packets to auth

Attempt to contact 26 unique IP addresses
Prague: PowerDNS packets to auth

attempt to contact 26 unique IP addresses

### Latency to Auth Server

- **1 ms**: 14 queries (8.0%)
- **2 ms**: 12 queries (7.2%)
- **3 ms**: 23 queries (13.8%)
- **5 ms**: 27 queries (16.2%)
- **7 ms**: 13 queries (8.0%)
- **11 ms**: 1 query (0.6%)
- **14 ms**: 1 query (0.6%)
- **17 ms**: 1 query (0.6%)
- **21 ms**: 1 query (0.6%)
- **24 ms**: 1 query (0.6%)
- **25 ms**: 1 query (0.6%)
- **33 ms**: 1 query (0.6%)
- **100 ms**: 1 query (0.6%)
- **107 ms**: 1 query (0.6%)
- **122 ms**: 1 query (0.6%)
- **157 ms**: 1 query (0.6%)
Prague: Unbound packets to auth

attempt to contact 26 unique IP addresses

<table>
<thead>
<tr>
<th>Latency to auth server</th>
<th># of queries to auth [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>1 ± 0 ms</td>
<td>4</td>
</tr>
<tr>
<td>1 ± 0 ms</td>
<td>4</td>
</tr>
<tr>
<td>1 ± 0 ms</td>
<td>4</td>
</tr>
<tr>
<td>1 ± 0 ms</td>
<td>4</td>
</tr>
<tr>
<td>1 ± 0 ms</td>
<td>4</td>
</tr>
<tr>
<td>1 ± 0 ms</td>
<td>4</td>
</tr>
<tr>
<td>1 ± 2 ms</td>
<td>4</td>
</tr>
<tr>
<td>2 ± 2 ms</td>
<td>4</td>
</tr>
<tr>
<td>6 ± 5 ms</td>
<td>4</td>
</tr>
<tr>
<td>7 ± 0 ms</td>
<td>4</td>
</tr>
<tr>
<td>7 ± 0 ms</td>
<td>4</td>
</tr>
<tr>
<td>7 ± 7 ms</td>
<td>4</td>
</tr>
<tr>
<td>8 ± 7 ms</td>
<td>4</td>
</tr>
<tr>
<td>14 ± 1 ms</td>
<td>4</td>
</tr>
<tr>
<td>14 ± 0 ms</td>
<td>4</td>
</tr>
<tr>
<td>14 ± 0 ms</td>
<td>4</td>
</tr>
<tr>
<td>21 ± 4 ms</td>
<td>4</td>
</tr>
<tr>
<td>22 ± 7 ms 18.0%</td>
<td>4</td>
</tr>
<tr>
<td>24 ± 5 ms</td>
<td>4</td>
</tr>
<tr>
<td>25 ± 45 ms</td>
<td>4</td>
</tr>
<tr>
<td>33 ± 3 ms</td>
<td>4</td>
</tr>
<tr>
<td>100 ± 2 ms</td>
<td>4</td>
</tr>
<tr>
<td>107 ± 4 ms</td>
<td>4</td>
</tr>
<tr>
<td>122 ± 6 ms</td>
<td>4</td>
</tr>
<tr>
<td>157 ± 2 ms 10.6%</td>
<td>4</td>
</tr>
</tbody>
</table>
Prague: latency over time

sliding window = 100 queries
Prague: latency histogram

![Prague latency histogram graph](image_url)
New Caledonia: auth servers

RTT to auth server [ms]

- 3 ± 1 ms
- 3 ± 6 ms
- 24 ± 1 ms
- 25 ± 7 ms
- 69 ± 7 ms
- 86 ± 119 ms
- 138 ± 4 ms
- 138 ± 4 ms
- 172 ± 6 ms
- 178 ± 2 ms
- 220 ± 92 ms
- 232 ± 9 ms
- 238 ± 5 ms
New Caledonia: # of queries to auth

- PowerDNS: 100%
- Unbound: 100%
- KNOT: 155%
- BIND: 198%
New Caledonia: BIND packets to auth

- Attempt to contact 13 unique IP addresses

- 87% of queries to auth:
  - 3 ± 1 ms
  - 3 ± 6 ms

- 13% of queries to auth:
  - 24 ± 1 ms
  - 25 ± 7 ms
  - 69 ± 7 ms
  - 86 ± 119 ms
  - 138 ± 4 ms
  - 138 ± 4 ms
  - 172 ± 6 ms
  - 178 ± 2 ms
  - 220 ± 92 ms
  - 232 ± 9 ms
  - 238 ± 5 ms

- Latency to auth server
New Caledonia: KNOT packets to auth

Attempt to contact 26 unique IP addresses.

Latency to auth server:
- 2 ± 1 ms: 5 queries
- 3 ± 1 ms: 6 queries
- 3 ± 6 ms: 5 queries
- 3 ± 5 ms: 3 queries
- 24 ± 1 ms: 6 queries
- 25 ± 7 ms: 6 queries
- 27 ± 41 ms: 2 queries
- 65 ± 69 ms: 2 queries
- 69 ± 7 ms: 3 queries
- 137 ± 7 ms: 3 queries
- 138 ± 4 ms: 4 queries
- 138 ± 4 ms: 4 queries
- 152 ± 14 ms: 2 queries
- 153 ± 24 ms: 1 query
- 171 ± 9 ms: 9 queries
- 172 ± 6 ms: 4 queries
- 173 ± 2 ms: 2 queries
- 173 ± 11 ms: 1 query
- 178 ± 2 ms: 3 queries
- 180 ± 6 ms: 5 queries
- 203 ± 3 ms: 4 queries
- 220 ± 92 ms: 9 queries
- 232 ± 9 ms: 5 queries
- 238 ± 5 ms: 2 queries
New Caledonia: PowerDNS packets to auth

attempt to contact 26 unique IP addresses

# of queries to auth [%]

2 ± 1 ms | 0
3 ± 1 ms | 0
3 ± 6 ms | 0
3 ± 5 ms | 0
24 ± 1 ms | 0
24 ± 1 ms | 0
25 ± 7 ms | 0
65 ± 69 ms | 0
69 ± 7 ms | 0
86 ± 119 ms | 0
137 ± 7 ms | 0
138 ± 4 ms | 0
138 ± 4 ms | 0
152 ± 14 ms | 0
153 ± 24 ms | 1.2%
172 ± 6 ms | 0
173 ± 2 ms | 0
173 ± 11 ms | 0
178 ± 2 ms | 0
180 ± 6 ms | 0
203 ± 3 ms | 0
220 ± 92 ms | 0
232 ± 9 ms | 0
238 ± 5 ms | 0

latency to auth server
New Caledonia: Unbound packets to auth

attempt to contact 26 unique IP addresses

# of queries to auth [%]

latency to auth server:

- 2 ± 1 ms: 9
- 3 ± 1 ms: 8
- 3 ± 6 ms: 8
- 3 ± 5 ms: 8
- 24 ± 1 ms: 8
- 24 ± 1 ms: 8
- 25 ± 7 ms: 8
- 65 ± 69 ms: 8
- 69 ± 7 ms: 8
- 86 ± 119 ms: 8
- 137 ± 7 ms: 8
- 138 ± 4 ms: 8
- 152 ± 14 ms: 8
- 153 ± 24 ms: 1.2%
- 157 ± 9 ms: 0.3%
- 171 ± 6 ms: 0
- 172 ± 6 ms: 0
- 173 ± 2 ms: 0
- 173 ± 11 ms: 0
- 178 ± 2 ms: 0
- 180 ± 6 ms: 0
- 203 ± 3 ms: 0
- 220 ± 92 ms: 0
- 223 ± 9 ms: 0
- 238 ± 5 ms: 0
New Caledonia: latency over time

sliding window = 100 queries

Average response time [ms]

- BIND
- PowerDNS
- Unbound
- KNOT

client query no.
New Caledonia: latency over time (zoom)

sliding window = 10 queries

Average response time [ms]

- BIND
- PowerDNS
- Unbound
- KNOT

client query no.
New Caledonia: latency histogram
Brazil: auth servers

- 34 ± 23 ms 1.3%
- 36 ± 33 ms 0.6%
- 43 ± 33 ms 1.0%
- 150 ± 31 ms 0.3%
- 153 ± 35 ms 0.6%
- 159 ± 23 ms 0.3%
- 167 ± 16 ms 1.0%
- 168 ± 59 ms 0.3%
- 184 ± 19 ms 0.3%
- 186 ± 45 ms 0.3%
- 203 ± 19 ms 0.3%
- 204 ± 17 ms 0.3%

Note: The values indicate the round-trip time (RTT) to the auth servers with error margins and percentages.
Brazil: # of queries to auth

- PowerDNS: 101%
- Unbound: 101%
- BIND: 124%
- KNOT: 144%
Brazil: BIND packets to auth

79 attempts to contact 26 unique IP addresses

Number of queries to auth [%]

34 ± 23 ms 0.3%
36 ± 33 ms 0.3%
43 ± 33 ms 0.6%
150 ± 31 ms 0.3%
153 ± 35 ms 0.3%
159 ± 23 ms 0.3%
165 ± 84 ms 0.6%
167 ± 16 ms 0.3%
168 ± 59 ms 0.3%
184 ± 19 ms 0.3%
186 ± 45 ms 0.3%
203 ± 19 ms 0.3%
204 ± 17 ms 0.3%

Latency to auth server

dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
Brazil: KNOT packets to auth

attempt to contact 26 unique IP addresses

<table>
<thead>
<tr>
<th>Latency (ms)</th>
<th>Percentage</th>
<th>Number of Queries</th>
</tr>
</thead>
<tbody>
<tr>
<td>34 ± 23</td>
<td>0.3%</td>
<td>11</td>
</tr>
<tr>
<td>36 ± 33</td>
<td>0.6%</td>
<td>10</td>
</tr>
<tr>
<td>43 ± 33</td>
<td>0.6%</td>
<td>10</td>
</tr>
<tr>
<td>150 ± 31</td>
<td>0.3%</td>
<td>4</td>
</tr>
<tr>
<td>153 ± 35</td>
<td>0.3%</td>
<td>4</td>
</tr>
<tr>
<td>159 ± 23</td>
<td>0.3%</td>
<td>4</td>
</tr>
<tr>
<td>165 ± 84</td>
<td>0.6%</td>
<td>4</td>
</tr>
<tr>
<td>167 ± 16</td>
<td>0.6%</td>
<td>5</td>
</tr>
<tr>
<td>168 ± 59</td>
<td>0.6%</td>
<td>5</td>
</tr>
<tr>
<td>184 ± 19</td>
<td>0.6%</td>
<td>5</td>
</tr>
<tr>
<td>186 ± 45</td>
<td>0.6%</td>
<td>5</td>
</tr>
<tr>
<td>203 ± 19</td>
<td>0.3%</td>
<td>5</td>
</tr>
<tr>
<td>204 ± 17</td>
<td>0.3%</td>
<td>5</td>
</tr>
<tr>
<td>Dead</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Dead</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Dead</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Dead</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Dead</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Dead</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Dead</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Dead</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Dead</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Dead</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Dead</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Dead</td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

Latency to auth server
Brazil: PowerDNS packets to auth

Attempt to contact 26 unique IP addresses

<table>
<thead>
<tr>
<th># of queries to auth [%]</th>
<th>latency to auth server</th>
</tr>
</thead>
<tbody>
<tr>
<td>34 ± 23 ms 0.3%</td>
<td>36 ± 33 ms 0.3%</td>
</tr>
<tr>
<td>43 ± 33 ms 0.6%</td>
<td>150 ± 31 ms 0.3%</td>
</tr>
<tr>
<td>153 ± 35 ms 0.3%</td>
<td>159 ± 23 ms 0.3%</td>
</tr>
<tr>
<td>165 ± 84 ms 0.6%</td>
<td>167 ± 16 ms 0.3%</td>
</tr>
<tr>
<td>168 ± 59 ms 0.3%</td>
<td>184 ± 19 ms 0.3%</td>
</tr>
<tr>
<td>186 ± 45 ms 0.3%</td>
<td>203 ± 19 ms 0.3%</td>
</tr>
<tr>
<td>204 ± 17 ms 0.3%</td>
<td>0</td>
</tr>
</tbody>
</table>
Brazil: Unbound packets to auth

- Attempt to contact 26 unique IP addresses

<table>
<thead>
<tr>
<th># of queries to auth (%)</th>
<th>Latency to auth server</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>34 ± 23 ms 0.3%</td>
</tr>
<tr>
<td>11</td>
<td>36 ± 33 ms 0.3%</td>
</tr>
<tr>
<td>11</td>
<td>43 ± 33 ms 0.6%</td>
</tr>
<tr>
<td>10</td>
<td>153 ± 35 ms</td>
</tr>
<tr>
<td>11</td>
<td>159 ± 23 ms 0.3%</td>
</tr>
<tr>
<td>11</td>
<td>165 ± 84 ms 0.6%</td>
</tr>
<tr>
<td>11</td>
<td>167 ± 16 ms</td>
</tr>
<tr>
<td>11</td>
<td>168 ± 59 ms</td>
</tr>
<tr>
<td>0</td>
<td>184 ± 19 ms</td>
</tr>
<tr>
<td>0</td>
<td>186 ± 45 ms</td>
</tr>
<tr>
<td>0</td>
<td>203 ± 19 ms 0.3%</td>
</tr>
<tr>
<td>0</td>
<td>204 ± 17 ms</td>
</tr>
<tr>
<td>0</td>
<td>dead</td>
</tr>
<tr>
<td>0</td>
<td>dead</td>
</tr>
<tr>
<td>0</td>
<td>dead</td>
</tr>
<tr>
<td>0</td>
<td>dead</td>
</tr>
<tr>
<td>0</td>
<td>dead</td>
</tr>
<tr>
<td>0</td>
<td>dead</td>
</tr>
<tr>
<td>0</td>
<td>dead</td>
</tr>
<tr>
<td>0</td>
<td>dead</td>
</tr>
<tr>
<td>0</td>
<td>dead</td>
</tr>
<tr>
<td>0</td>
<td>dead</td>
</tr>
<tr>
<td>0</td>
<td>dead</td>
</tr>
</tbody>
</table>
Brazil: latency over time

sliding window = 100 queries

Average response time [ms]

client query no.
Brazil: latency over time (zoom)
Brazil: latency histogram

- BIND
- 1.00% SERVFAIL
- KNOT
- PowerDNS
- Unbound

Response time [ms]

Slowest percentile

1 packet | 10 packets | 100 packets | 1000 packets | 10000 packets
Hawai: auth servers

RTT to auth server [ms]

- 17 ± 2 ms
- 63 ± 43 ms
- 72 ± 46 ms
- 73 ± 47 ms
- 76 ± 37 ms
- 76 ± 47 ms
- 76 ± 59 ms
- 93 ± 37 ms
- 106 ± 55 ms
- 132 ± 53 ms
- 234 ± 5 ms
- 299 ± 45 ms

dead, dead, dead, dead, dead, dead, dead, dead, dead, dead
Hawai: # of queries to auth

- PowerDNS: 101%
- Unbound: 102%
- BIND: 107%
- KNOT: 137%
Hawai: BIND packets to auth

attempt to contact 26 unique IP addresses

<table>
<thead>
<tr>
<th># of queries to auth [%]</th>
<th>latency to auth server</th>
</tr>
</thead>
<tbody>
<tr>
<td>17 ± 2 ms 1.0% 1.3%</td>
<td>61 ± 32 ms 0.3% 0.6%</td>
</tr>
<tr>
<td>63 ± 43 ms 1.0% 0.3%</td>
<td>72 ± 46 ms 0.3% 0.6%</td>
</tr>
<tr>
<td>76 ± 47 ms 0.3% 0.6%</td>
<td>76 ± 47 ms 0.3% 0.6%</td>
</tr>
<tr>
<td>76 ± 47 ms 0.3% 0.6%</td>
<td>76 ± 47 ms 0.3% 0.6%</td>
</tr>
<tr>
<td>83 ± 37 ms 1.0% 0.6%</td>
<td>93 ± 37 ms 0.3% 0.6%</td>
</tr>
<tr>
<td>106 ± 55 ms 0.3% 0.6%</td>
<td>234 ± 45 ms 0.3% 0.6%</td>
</tr>
<tr>
<td>299 ± 45 ms 0.3% 0.6%</td>
<td>dead</td>
</tr>
</tbody>
</table>

dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
Hawai: KNOT packets to auth

Attempt to contact 26 unique IP addresses

<table>
<thead>
<tr>
<th>Latency to auth server</th>
<th># of queries to auth (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>17 ± 2 ms 1.0.3%</td>
<td>14</td>
</tr>
<tr>
<td>63 ± 43 ms 1.0.3%</td>
<td>6</td>
</tr>
<tr>
<td>72 ± 46 ms</td>
<td>5</td>
</tr>
<tr>
<td>76 ± 37 ms</td>
<td>5</td>
</tr>
<tr>
<td>76 ± 47 ms</td>
<td>5</td>
</tr>
<tr>
<td>76 ± 59 ms</td>
<td>5</td>
</tr>
<tr>
<td>93 ± 37 ms</td>
<td>5</td>
</tr>
<tr>
<td>106 ± 55 ms</td>
<td>5</td>
</tr>
<tr>
<td>234 ± 5 ms</td>
<td>5</td>
</tr>
<tr>
<td>299 ± 45 ms</td>
<td>5</td>
</tr>
<tr>
<td>132 ± 53 ms 1.0.6%</td>
<td>1</td>
</tr>
<tr>
<td>61 ± 32 ms</td>
<td>1</td>
</tr>
</tbody>
</table>

Latency categories: dead, 1, 2, 5
Hawai: PowerDNS packets to auth

attempt to contact 26 unique IP addresses

# of queries to auth [%]

latency to auth server

17 ± 2 ms 0.3%
Hawai: Unbound packets to auth

attempt to contact 26 unique IP addresses

<table>
<thead>
<tr>
<th># of queries to auth [%]</th>
<th>latency to auth server</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>17 ± 2 ms 10.3%</td>
</tr>
<tr>
<td>10</td>
<td>61 ± 32 ms 10.3%</td>
</tr>
<tr>
<td>10</td>
<td>63 ± 43 ms 10.3%</td>
</tr>
<tr>
<td>9</td>
<td>72 ± 46 ms 10.3%</td>
</tr>
<tr>
<td>10</td>
<td>76 ± 37 ms 10.3%</td>
</tr>
<tr>
<td>10</td>
<td>76 ± 47 ms 10.3%</td>
</tr>
<tr>
<td>10</td>
<td>76 ± 59 ms 10.3%</td>
</tr>
<tr>
<td>9</td>
<td>93 ± 37 ms 10.3%</td>
</tr>
<tr>
<td>10</td>
<td>106 ± 55 ms 10.3%</td>
</tr>
<tr>
<td>10</td>
<td>132 ± 53 ms 10.6%</td>
</tr>
<tr>
<td>234 ± 5 ms</td>
<td>dead</td>
</tr>
<tr>
<td>299 ± 45 ms</td>
<td>dead</td>
</tr>
<tr>
<td>dead</td>
<td>dead</td>
</tr>
<tr>
<td>dead</td>
<td>dead</td>
</tr>
<tr>
<td>dead</td>
<td>dead</td>
</tr>
<tr>
<td>dead</td>
<td>dead</td>
</tr>
<tr>
<td>dead</td>
<td>dead</td>
</tr>
<tr>
<td>dead</td>
<td>dead</td>
</tr>
<tr>
<td>dead</td>
<td>dead</td>
</tr>
<tr>
<td>dead</td>
<td>dead</td>
</tr>
<tr>
<td>dead</td>
<td>dead</td>
</tr>
<tr>
<td>dead</td>
<td>dead</td>
</tr>
<tr>
<td>dead</td>
<td>dead</td>
</tr>
</tbody>
</table>
Hawai: latency over time

sliding window = 100 queries

Average response time [ms]

client query no.
Hawai: latency over time (zoom)
Tasmania: auth servers

RTT to auth server [ms]

- 27 ± 218 ms
- 36 ± 16 ms
- 37 ± 214 ms
- 48 ± 221 ms
- 50 ± 6 ms
- 70 ± 2 ms
- 164 ± 2 ms
- 167 ± 3 ms
- 172 ± 70 ms 4.2%
- 177 ± 13 ms
- 177 ± 11 ms
- 179 ± 8 ms
- 194 ± 4 ms 3.3%

dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
dead
Tasmania: # of queries to auth

- PowerDNS: 101%
- Unbound: 104%
- BIND: 127%
- KNOT: 131%
Tasmania: BIND packets to auth

attempt to contact 26 unique IP addresses
Tasmania: KNOT packets to auth

attempt to contact 26 unique IP addresses

# of queries to auth [%]

latency to auth server
Tasmania: PowerDNS packets to auth

96 attempts to contact 26 unique IP addresses

Latency to auth server:
- 27 ± 18 ms (0 %)
- 36 ± 16 ms (0 %)
- 37 ± 214 ms (0 %)
- 48 ± 221 ms (0 %)
- 50 ± 6 ms (0 %)
- 70 ± 2 ms (0 %)
- 164 ± 2 ms (0 %)
- 167 ± 3 ms (0 %)
- 172 ± 70 ms (14.2 %)
- 177 ± 13 ms (4.2 %)
- 177 ± 11 ms (4.2 %)
- 179 ± 8 ms (0.3 %)
- 194 ± 4 ms (0.3 %)

All other latencies are marked as 'dead'.
Tasmania: Unbound packets to auth

- Attempt to contact 26 unique IP addresses

- Latency to auth server:
  - 27 ± 218 ms: 2 queries
  - 36 ± 16 ms: 9 queries
  - 37 ± 214 ms: 0 queries
  - 48 ± 221 ms: 1 query
  - 50 ± 6 ms: 10 queries
  - 70 ± 2 ms: 9 queries
  - 164 ± 2 ms: 10 queries
  - 167 ± 3 ms: 9 queries
  - 177 ± 13 ms: 10 queries
  - 177 ± 11 ms: 9 queries
  - 179 ± 8 ms: 9 queries
  - 194 ± 4 ms: 10 queries

- Latency range: 27 ± 218 ms to 194 ± 4 ms
Tasmania: latency over time

sliding window = 100 queries

Average response time [ms]

client query no.
Tasmania: latency over time (zoom)

sliding window = 10 queries

Average response time [ms]

client query no.
Tasmania: latency histogram

- BIND
- 1.00% SERVFAIL
- KNOT
- PowerDNS
- Unbound
China Atlas #34903: auth servers

RTT to auth server [ms]

26 ± 13 ms | 0.3 %
71 ± 33 ms | 19.5 %
97 ± 39 ms | 32.7 %
151 ± 55 ms | 14.9 %
153 ± 50 ms | 18.6 %
156 ± 88 ms | 19.6 %
190 ± 62 ms | 18.3 %
191 ± 67 ms | 12.5 %
193 ± 68 ms | 12.2 %
208 ± 69 ms | 10.1 %
252 ± 38 ms | 12.5 %
256 ± 33 ms | 12.5 %
267 ± 33 ms | 17.6 %
272 ± 32 ms | 19.8 %
292 ± 30 ms | 21.1 %
313 ± 35 ms | 16.8 %
315 ± 38 ms | 13.7 %
317 ± 32 ms | 19.0 %
324 ± 35 ms | 11.6 %
353 ± 62 ms | 28.9 %
508 ± 70 ms | 21.1 %
China Atlas #34903: # of queries to auth

- PowerDNS: 101%
- BIND: 107%
- Unbound: 120%
- KNOT: 211%
China Atlas #34903: BIND packets to auth

attempt to contact 26 unique IP addresses

<table>
<thead>
<tr>
<th># of queries to auth [%]</th>
<th>latency to auth server</th>
</tr>
</thead>
<tbody>
<tr>
<td>72</td>
<td>26 ± 13 ms</td>
</tr>
<tr>
<td>3</td>
<td>71 ± 33 ms</td>
</tr>
<tr>
<td>3</td>
<td>97 ± 39 ms</td>
</tr>
<tr>
<td>0</td>
<td>151 ± 55 ms</td>
</tr>
<tr>
<td>0</td>
<td>153 ± 50 ms</td>
</tr>
<tr>
<td>2</td>
<td>208 ± 69 ms</td>
</tr>
<tr>
<td>0</td>
<td>191 ± 67 ms</td>
</tr>
<tr>
<td>1</td>
<td>156 ± 88 ms</td>
</tr>
<tr>
<td>1</td>
<td>190 ± 62 ms</td>
</tr>
<tr>
<td>2</td>
<td>252 ± 38 ms</td>
</tr>
<tr>
<td>2</td>
<td>267 ± 35 ms</td>
</tr>
<tr>
<td>1</td>
<td>271 ± 35 ms</td>
</tr>
<tr>
<td>2</td>
<td>276 ± 33 ms</td>
</tr>
<tr>
<td>1</td>
<td>292 ± 30 ms</td>
</tr>
<tr>
<td>0</td>
<td>313 ± 35 ms</td>
</tr>
<tr>
<td>1</td>
<td>317 ± 32 ms</td>
</tr>
<tr>
<td>1</td>
<td>324 ± 35 ms</td>
</tr>
<tr>
<td>1</td>
<td>353 ± 62 ms</td>
</tr>
<tr>
<td>1</td>
<td>508 ± 70 ms</td>
</tr>
</tbody>
</table>
China Atlas #34903: KNOT packets to auth

attempt to contact 26 unique IP addresses

<table>
<thead>
<tr>
<th>Latency to auth server</th>
<th>Queries to auth [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>26 ± 13 ms</td>
<td>9</td>
</tr>
<tr>
<td>71 ± 33 ms</td>
<td>6</td>
</tr>
<tr>
<td>97 ± 39 ms</td>
<td>4</td>
</tr>
<tr>
<td>99 ± 39 ms</td>
<td>5</td>
</tr>
<tr>
<td>151 ± 55 ms</td>
<td>5</td>
</tr>
<tr>
<td>153 ± 50 ms</td>
<td>3</td>
</tr>
<tr>
<td>153 ± 52 ms</td>
<td>4</td>
</tr>
<tr>
<td>182 ± 349 ms</td>
<td>5</td>
</tr>
<tr>
<td>190 ± 62 ms</td>
<td>4</td>
</tr>
<tr>
<td>191 ± 67 ms</td>
<td>5</td>
</tr>
<tr>
<td>193 ± 68 ms</td>
<td>4</td>
</tr>
<tr>
<td>208 ± 69 ms</td>
<td>5</td>
</tr>
<tr>
<td>252 ± 38 ms</td>
<td>3</td>
</tr>
<tr>
<td>256 ± 38 ms</td>
<td>3</td>
</tr>
<tr>
<td>267 ± 33 ms</td>
<td>3</td>
</tr>
<tr>
<td>271 ± 35 ms</td>
<td>4</td>
</tr>
<tr>
<td>272 ± 32 ms</td>
<td>3</td>
</tr>
<tr>
<td>276 ± 39 ms</td>
<td>4</td>
</tr>
<tr>
<td>292 ± 30 ms</td>
<td>3</td>
</tr>
<tr>
<td>313 ± 35 ms</td>
<td>3</td>
</tr>
<tr>
<td>315 ± 38 ms</td>
<td>3</td>
</tr>
<tr>
<td>324 ± 35 ms</td>
<td>3</td>
</tr>
<tr>
<td>353 ± 62 ms</td>
<td>3</td>
</tr>
<tr>
<td>508 ± 70 ms</td>
<td>2</td>
</tr>
</tbody>
</table>
China Atlas #34903: PowerDNS packets to auth

attempt to contact 26 unique IP addresses

<table>
<thead>
<tr>
<th>Latency to auth server</th>
<th>% of queries to auth</th>
</tr>
</thead>
<tbody>
<tr>
<td>26 ± 13 ms</td>
<td>0.3 %</td>
</tr>
<tr>
<td>71 ± 33 ms</td>
<td>1.9 %</td>
</tr>
<tr>
<td>97 ± 39 ms</td>
<td>2.6 %</td>
</tr>
<tr>
<td>99 ± 55 ms</td>
<td>3.7 %</td>
</tr>
<tr>
<td>153 ± 50 ms</td>
<td>8.6 %</td>
</tr>
<tr>
<td>156 ± 88 ms</td>
<td>9.6 %</td>
</tr>
<tr>
<td>182 ± 349 ms</td>
<td>14.6 %</td>
</tr>
<tr>
<td>191 ± 67 ms</td>
<td>18.3 %</td>
</tr>
<tr>
<td>193 ± 68 ms</td>
<td>18.5 %</td>
</tr>
<tr>
<td>208 ± 69 ms</td>
<td>12.5 %</td>
</tr>
<tr>
<td>252 ± 38 ms</td>
<td>14.6 %</td>
</tr>
<tr>
<td>267 ± 33 ms</td>
<td>14.6 %</td>
</tr>
<tr>
<td>271 ± 35 ms</td>
<td>17.6 %</td>
</tr>
<tr>
<td>272 ± 33 ms</td>
<td>14.6 %</td>
</tr>
<tr>
<td>276 ± 39 ms</td>
<td>17.6 %</td>
</tr>
<tr>
<td>292 ± 30 ms</td>
<td>21.1 %</td>
</tr>
<tr>
<td>313 ± 35 ms</td>
<td>16.8 %</td>
</tr>
<tr>
<td>317 ± 32 ms</td>
<td>19.0 %</td>
</tr>
<tr>
<td>324 ± 35 ms</td>
<td>11.1 %</td>
</tr>
<tr>
<td>353 ± 62 ms</td>
<td>28.9 %</td>
</tr>
<tr>
<td>508 ± 70 ms</td>
<td>21.1 %</td>
</tr>
</tbody>
</table>
China Atlas #34903: Unbound packets to auth

The graph shows the latency to the auth server and the percentage of queries to auth that attempt to contact 26 unique IP addresses. The latencies range from 26 ms to 508 ms with corresponding percentages ranging from 0.3% to 1%.
China Atlas #34903: latency over time

sliding window = 100 queries

Average response time [ms]

- BIND
- PowerDNS
- Unbound
- KNOT

client query no.
China Atlas #34903: latency over time (zoom)

sliding window = 10 queries

Average response time [ms]

- BIND
- PowerDNS
- Unbound
- KNOT

client query no.
China Atlas #34903: latency histogram

The graph shows the response time in milliseconds for different DNS implementations across various packet counts. The x-axis represents the slowest percentile of packet counts, ranging from 1 packet to 100,000 packets. The y-axis represents the response time in milliseconds, ranging from 0.01 ms to 10,000 ms.

Four different DNS implementations are compared:
- **BIND**: Black line
- **KNOT**: Blue line
- **PowerDNS**: Orange line
- **Unbound**: Green line

The graph illustrates how each implementation handles latency under different load conditions.
DNS Maze: limitations

- Black-box approach
  - No insight "why"
  - Client-visible latency
    - Not only server selection ...
Further work

- Automate evaluation
- Re-implement algorithms in Teachable DNS resolver
  - Isolate server selection algorithm
- Simulate dynamic changes
- Rewrite server selection in Knot Resolver :-(
Conclusion: Latency != throughput

- QPS throughput practically **NOT** limited by server selection
- Server selection quality
  - this talk
- Throughput
  - talk "Benchmarking and Optimizing DNS Resolvers on the ISP level"
  - http://ripe79.ripe.net/programme/meeting-plan/dns-wg/
- Test your code! :-)