Internet centralization concerns: US DOJ

Justice Department Opens Antitrust Review of Big Tech Companies

Centralization poses various risks

- Creates a **single point of failure**
- Privacy
- Market consolidation

DYN DNS 2016 Attack

Can we measure Internet Centralization?

Easier said than done.

Measure it in terms of?

- Users?
- Traffic?
- Networking infrastructure?
- Computing infrastructure?
- Market?
- ...

Our approach:

- We focus on DNS traffic
- But NOT on user traffic
- We focus on traffic from resolvers to authoritative servers
Can we measure Internet Centralization?

Easier said than done.

Measure it in terms of?

- Users?
- Traffic?
- Networking infrastructure?
- Computing infrastructure?
- Market?
- ...

Our approach:

- We focus on **DNS traffic**
- But **NOT** on *user* traffic
- We focus on traffic from resolvers to authoritative servers
What we measure: DNS queries to

The Netherlands (.nl)
- 17.1M inhabitants
- 6M domain names (.nl)
- Continent: Europe
- Official language: Dutch

New Zealand (.nz)
- 4.8 M inhabitants
- 700k domain names (.nz)
- Continent: Oceania
- Official languages: English, Maori

B-Root

World
- 7.8 Billion inhabitants
- 1588 TLDs
- Continents: 7
- Language: *
What we measure: DNS queries from

From 5 Cloud/Content Providers

<table>
<thead>
<tr>
<th>Company</th>
<th>ASes</th>
<th>Public DNS?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>15169</td>
<td>Yes</td>
</tr>
<tr>
<td>Amazon</td>
<td>7224, 8987, 9059, 14168, 16509</td>
<td>No</td>
</tr>
<tr>
<td>Microsoft</td>
<td>3598, 6584, 8068–8075, 12076, 23468</td>
<td>No</td>
</tr>
<tr>
<td>Facebook</td>
<td>32934</td>
<td>No</td>
</tr>
<tr>
<td>Cloudflare</td>
<td>13335</td>
<td>Yes</td>
</tr>
</tbody>
</table>
So, what did we find?
Traffic to b.root-servers.net

Year | Queries Ratio
--- | ---
2018 | 0.1
2019 | 0.2
2020 | 0.3

Companies:
- Google
- Amazon
- Microsoft
- Facebook
- Cloudflare
Traffic to .nz
Traffic to .nl

Queries Ratio

Year

Google
Amazon
Microsoft
Facebook
Cloudflare
### IPv4 vs IPv6 Adoption

- **Roughly 50/50%:**
  - Google, Cloudflare

- **More IPv6:**
  - Facebook (2019 onwards)

- **Very little IPv6:**
  - Microsoft, Amazon

#### IPv4 and IPv6 queries proportion

<table>
<thead>
<tr>
<th>Year</th>
<th>IPv4</th>
<th>IPv6</th>
<th>IPv4</th>
<th>IPv6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>0.66</td>
<td>0.34</td>
<td>0.61</td>
<td>0.39</td>
</tr>
<tr>
<td>2019</td>
<td>0.49</td>
<td>0.51</td>
<td>0.54</td>
<td>0.46</td>
</tr>
<tr>
<td>2020</td>
<td>0.52</td>
<td>0.48</td>
<td>0.54</td>
<td>0.46</td>
</tr>
<tr>
<td>Amazon</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>2019</td>
<td>0.98</td>
<td>0.02</td>
<td>0.97</td>
<td>0.03</td>
</tr>
<tr>
<td>2020</td>
<td>0.97</td>
<td>0.03</td>
<td>0.96</td>
<td>0.04</td>
</tr>
<tr>
<td>Microsoft</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>2019</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>2020</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Facebook</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>0.52</td>
<td>0.48</td>
<td>0.51</td>
<td>0.49</td>
</tr>
<tr>
<td>2019</td>
<td>0.24</td>
<td>0.76</td>
<td>0.19</td>
<td>0.81</td>
</tr>
<tr>
<td>2020</td>
<td>0.24</td>
<td>0.76</td>
<td>0.17</td>
<td>0.83</td>
</tr>
<tr>
<td>Cloudflare</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>0.54</td>
<td>0.46</td>
<td>0.54</td>
<td>0.46</td>
</tr>
<tr>
<td>2019</td>
<td>0.57</td>
<td>0.43</td>
<td>0.56</td>
<td>0.44</td>
</tr>
<tr>
<td>2020</td>
<td>0.51</td>
<td>0.49</td>
<td>0.49</td>
<td>0.51</td>
</tr>
</tbody>
</table>
Our measurements revealed:

- Traffic levels differ per cloud provider
  - “Junk” queries vary by provider and year
- Query types sent vary significantly
  - By cloud provider
  - From year to year
- Key technology deployment variations
  - DNSEC
  - IPv4 vs IPv6 usage
  - UDP vs TCP
  - Q-Name minimization
- Pros and Cons of centralization:
  - Rapid upgrades and rapid failures

Real-world cloud types

Paper (IMC2020):
Download it here