Clouding up the Internet:
how centralized is DNS traffic becoming?

Giovane C. M. Moura (1), Sebastian Castro (2), Wes Hardaker (3), Maarten Wullink (1) and Cristian Hesselman (1,4)

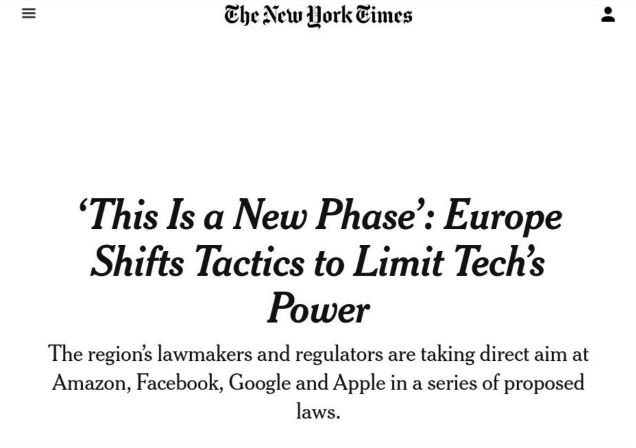
DNS-OARC

2020-09-28

Virtual Conference

1: SIDN Labs 2: InternetNZ 3: USC/ISI 4: University of Twente

Internet centralization concerns: EU



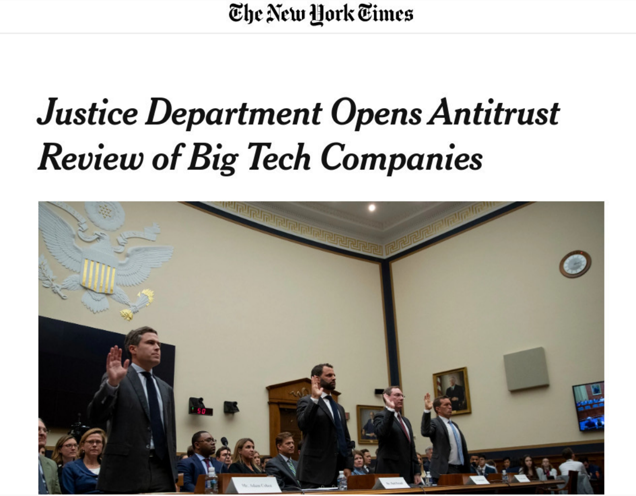
The New York Times

‘This Is a New Phase’: Europe Shifts Tactics to Limit Tech’s Power

The region’s lawmakers and regulators are taking direct aim at Amazon, Facebook, Google and Apple in a series of proposed laws.

source: https://www.nytimes.com/2020/07/30/technology/europe-new-phase-tech-amazon-apple-facebook-google.html

Internet centralization concerns: US DOJ

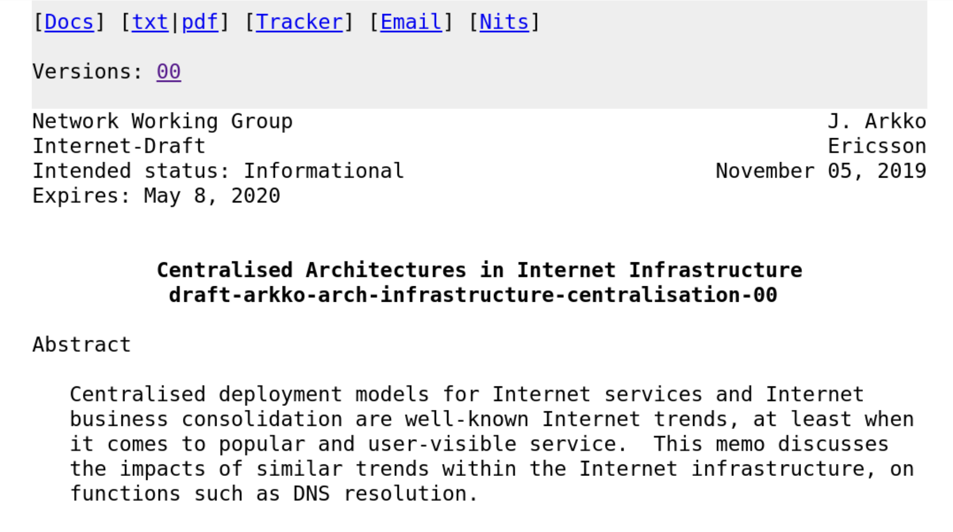


The New York Times

Justice Department Opens Antitrust Review of Big Tech Companies

source: https://www.nytimes.com/2019/07/23/technology/justice-department-tech-antitrust.html

Internet centralization concerns: IETF members



[Docs] [txt|pdf] [Tracker] [Email] [Nits]

Versions: 00

```
Network Working Group                                          J. Arkko
Internet-Draft                                                 Ericsson
Intended status: Informational                        November 05, 2019
Expires: May 8, 2020
```

Centralised Architectures in Internet Infrastructure
draft-arkko-arch-infrastructure-centralisation-00

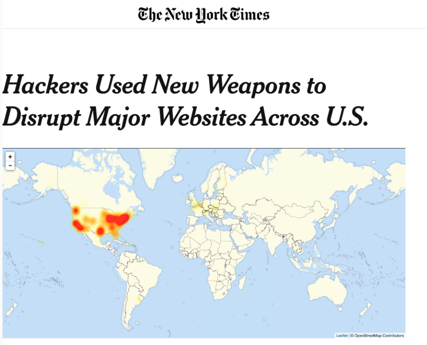
Abstract

Centralised deployment models for Internet services and Internet business consolidation are well-known Internet trends, at least when it comes to popular and user-visible service. This memo discusses the impacts of similar trends within the Internet infrastructure, on functions such as DNS resolution.

source: https://tools.ietf.org/html/draft-arkko-arch-infrastructure-centralisation-00

Centralization poses various risks

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



The New York Times

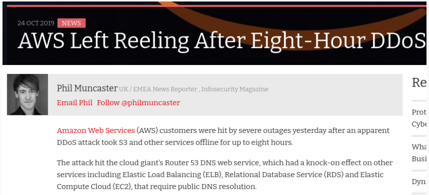
Hackers Used New Weapons to Disrupt Major Websites Across U.S.

DYN **DNS** 2016 Attack

source: https://www.nytimes.com/2016/10/22/business/internet-problems-attack.html

Centralization poses various risks

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



24 OCT 2019 NEWS

AWS Left Reeling After Eight-Hour DDoS

Phil Muncaster UK / EMEA News Reporter , Infosecurity Magazine

Email Phil Follow @philmuncaster

Amazon Web Services (AWS) customers were hit by severe outages yesterday after an apparent DDoS attack took S3 and other services offline for up to eight hours.

The attack hit the cloud giant's Router 53 DNS web service, which had a knock-on effect on other services including Elastic Load Balancing (ELB), Relational Database Service (RDS) and Elastic Compute Cloud (EC2), that require public DNS resolution.

Amazon **Route 53 (DNS)** 2019 Attack

source: https://www.infosecurity-magazine.com/news/aws-customers-hit-by-eighthour-ddos/

Can we measure Internet Centralization?

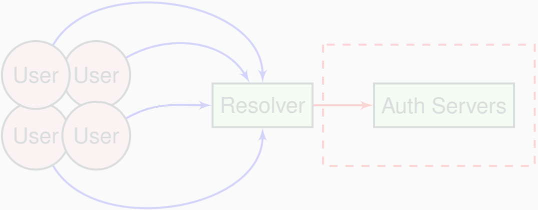
Easier said than done.

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?

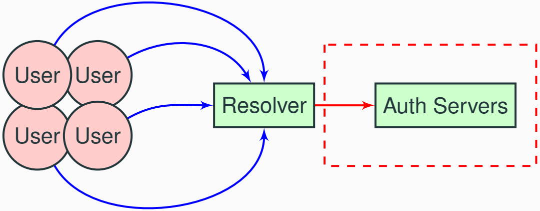
Easier said than done.

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

Company	ASes	Public DNS?
Google	15169	Yes
Amazon	7224, 8987, 9059, 14168, 16509	No
Microsoft	3598,6584, 8068–8075, 12076, 23468	No
Facebook	32934	No
Cloudflare	13335	Yes

Datasets: 55 Billion Queries, 1week/year, 3 years

$.nl$

Week	Queries(total)	Queries (valid)	Resolvers	ASes
w2018	7.29B	6.53B	2.09M	41276
w2019	10.16B	9.05B	2.18M	42727
w2020	13.75B	11.88B	1.99M	41716

$.nz$

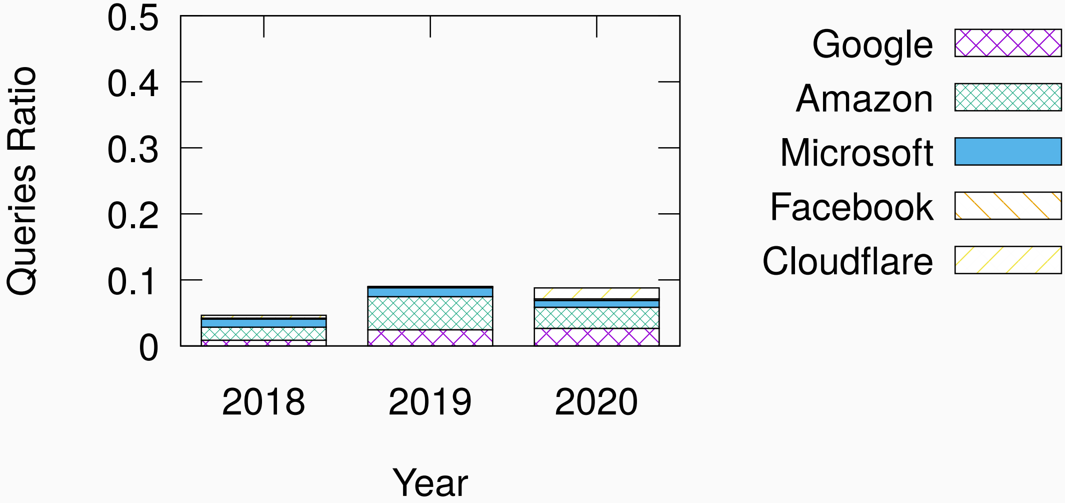
Week	Queries(total)	Queries (valid)	Resolvers	ASes
w2018	2.95B	2.00B	1.28M	37623
w2019	3.48B	2.81B	1.42M	39601
w2020	4.57B	3.03B	1.31M	38505

b.root-servers.net

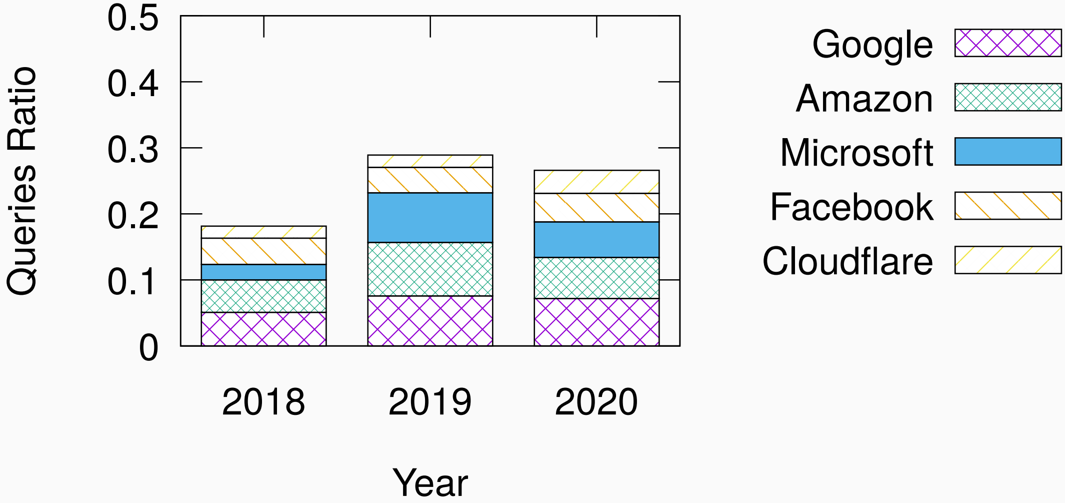
Date	Queries(total)	Queries (valid)	Resolvers	ASes
2018/04/10	2.68B	0.93B	4.23M	45210
2019/04/09	4.13B	1.43B	4.13M	48154
2020/05/06	6.70B	1.34B	6.01M	51820

So, what did we find?

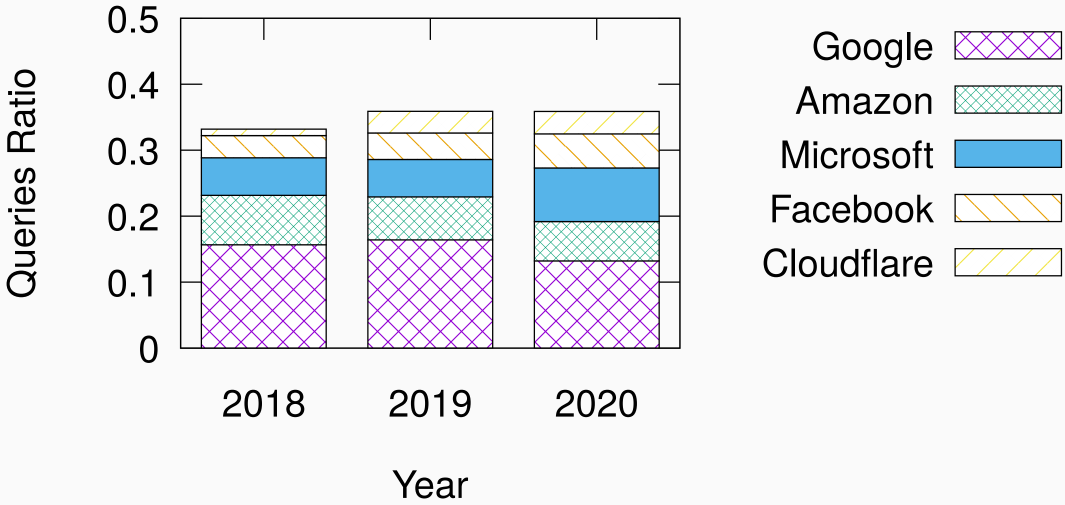
Traffic to b.root-servers.net



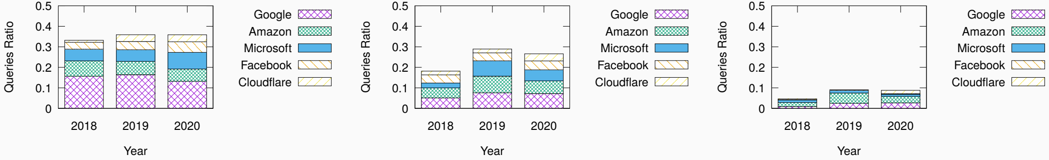
Traffic to .nz



Traffic to .nl



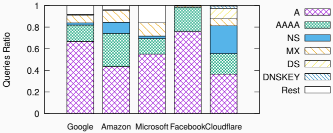
5 clouds → 1/3 of ccTLDs traffic



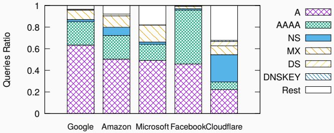
(a) `.nl` **(b)** `.nz` **(c)** b.root-servers.net

- The 5 clouds account for **roughly 1/3 of all queries** to `.nl` and `.nz`
  - `.nl` and `.nz` see 40k+ Autonomous Systems in total
- b.root-servers.net receives less, with than 9% of traffic from clouds
  - likely affected by tons of chromium-based garbage [5, 6]
- Oddity: Google sends more traffic to `.nl` than `.nz`

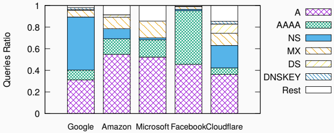
What do clouds dream of when visiting the Netherlands?



(d) 2018 – `.nl`

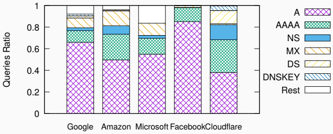


(e) 2019 – `.nl`

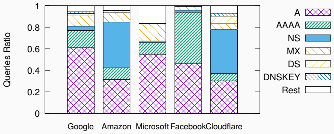


(f) 2020 – `.nl`

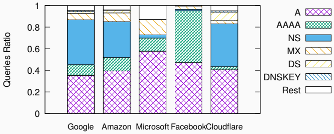
What do clouds dream of when visiting New Zealand?



(g) 2018 – `.nz`

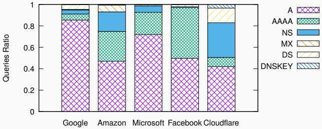


(h) 2019 – `.nz`

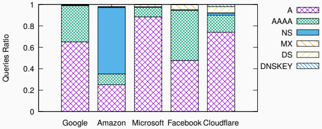


(i) 2020 – `.nz`

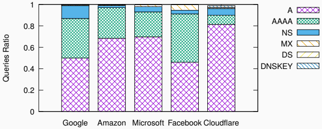
What do clouds dream of when visiting the Root?



(j) 2018 – B

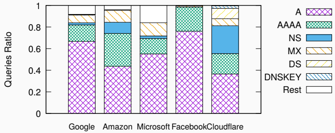


(k) 2019 – B

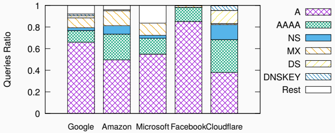


(l) 2020 – B

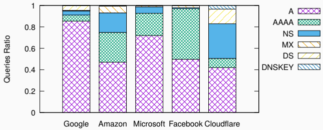
What did clouds dream of in 2018?



(m) 2018 – .nl

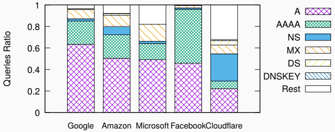


(n) 2018 – .nz

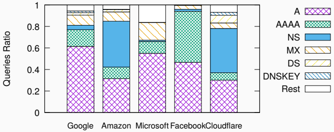


(o) 2018 – B

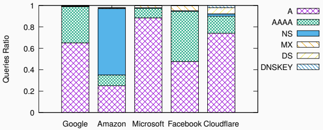
What do clouds dream of in 2019?



(p) 2019 – `.nl`

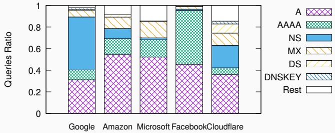


(q) 2019 – `.nz`

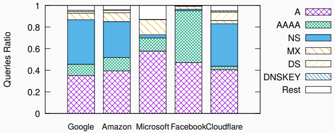


(r) 2019 – B

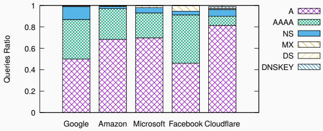
What do clouds dream of in 2020?



(s) 2020 – `.nl`

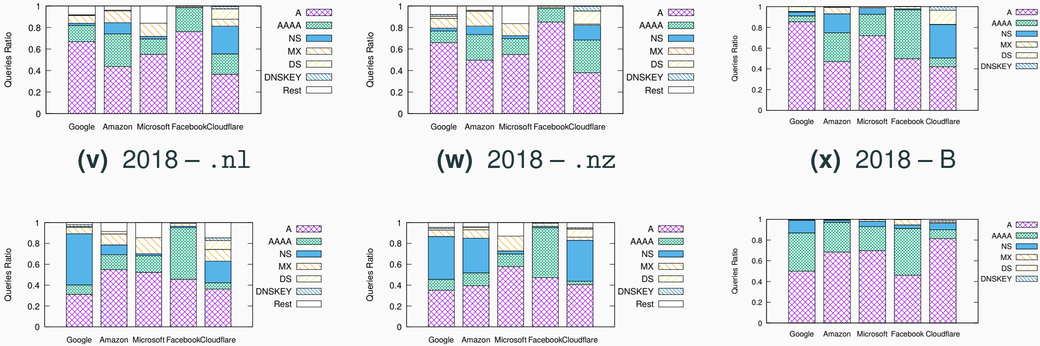


(t) 2020 – `.nz`



(u) 2020 – B

What do clouds dream of?



(v) 2018 – `.nl`

(w) 2018 – `.nz`

(x) 2018 – B

(y) 2020 – `.nl`

(z) 2020 – `.nz`

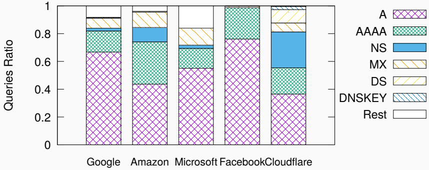
(aa) 2020 – B

Resource Records per Cloud provider

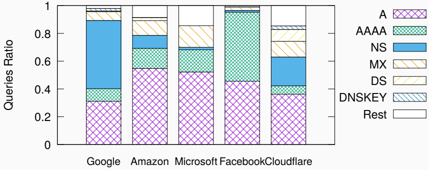
Mostly A records, but...

What do they ask for?

- Google sends more NS queries in 2020 than in 2018
- Why?
  - QNAME-minimization [4]
  - Q-min first query for the NS records
- We confirmed with Google that they deployed QNAME-minimization in Dec. 2019



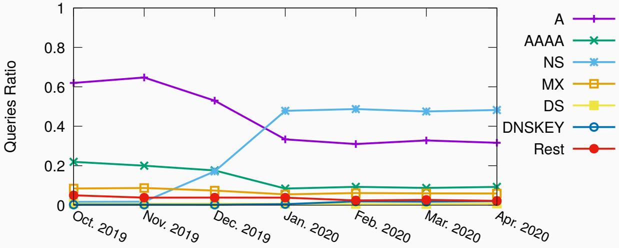
(ab) 2018 – `.nl`



(ac) 2020 – `.nl`

Identifying when Google deployed RFC7816

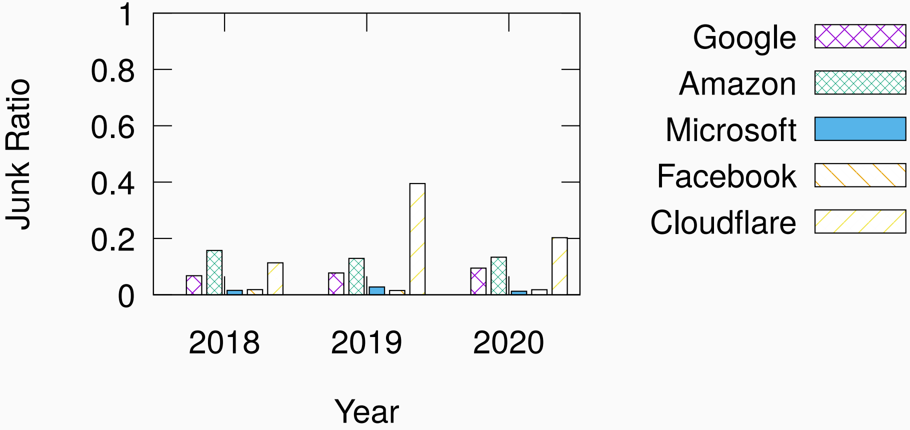
- As Google deployed QNAME-minimization it created a visible shift in query types
- **Centralization Pro**: new security feature deployments benefits many users all at once
    - DNSSEC validation
    - QNAME-minimization



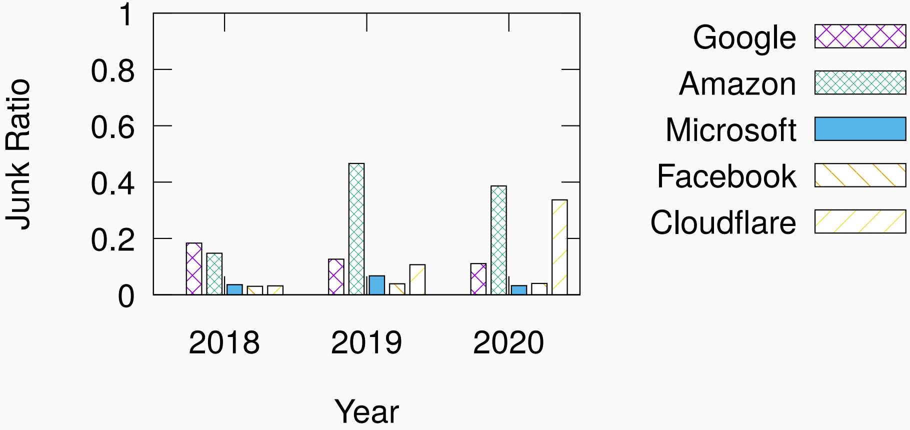
(ad) Google – `.nl`

Queries distribution per month for Google.

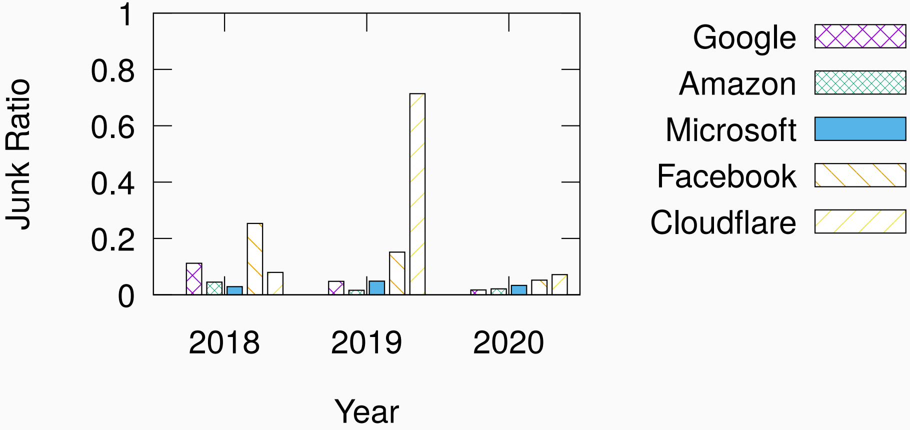
Junk queries sent to .NL from clouds



Junk queries sent to .NZ from clouds

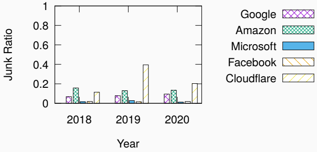


Junk queries sent to b.root-servers.net from clouds

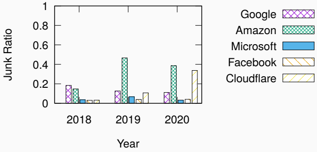


(ag) b.root-servers.net

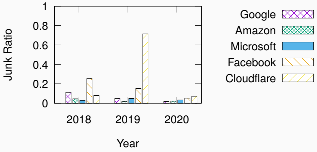
Junk queries raining from the clouds



(ah) `.nl`



(ai) `.nz`

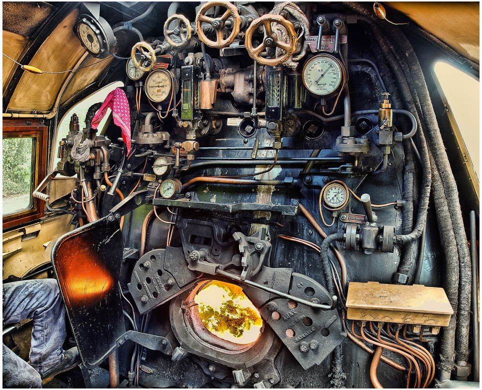


(aj) b.root-servers.net

- Junk := queries received for non-authoritative domains
- Distribution varies widely per zone
- ccTLDs: clouds send junk as all ASes do
- reduction in junk in junk levels to b.root-servers.net in 2020:
  - Proportionally, less junk from clouds
  - NSEC aggressive caching?
  - Chromium deployments now dominates root junk

Measuring Cloud Technology Adoption

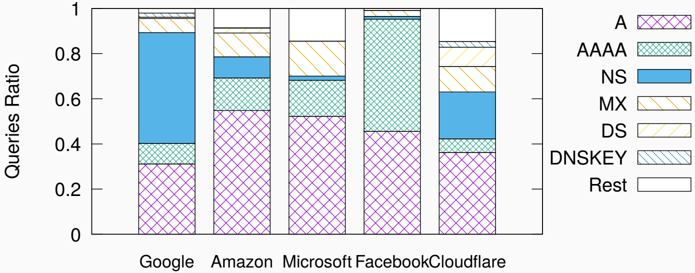
- DNSSEC
- IPv4 vs IPv6
- UDP vs TCP



source: https://www.flickr.com/photos/anguskirk/4817305157

DNSSEC

- DNSSEC provides authenticity and integrity [1, 3, 2].
- Do clouds use it equally?
  - They need DS and DNSKEY records



w2020: `.nl`

- Adoption measured by DNSKEY queries:
  - Microsoft: 0.02M / 1.1B
  - Cloudflare: 11M / 460M

IPv4 vs IPv6 Adoption

- Roughly 50/50%:
  Google, Cloudflare
- More IPv6:
  Facebook (2019 onwards)
- Very little IPv6:
  Microsoft, Amazon

		.nl		.nz	
	Year	**IPv4**	**IPv6**	**IPv4**	**IPv6**
Google	2018	0.66	0.34	0.61	0.39
	2019	0.49	0.51	0.54	0.46
	2020	0.52	0.48	054	0.46
Amazon	2018	1	0	1	0
	2019	0.98	0.02	0.97	0.03
	2020	0.97	0.03	0.96	0.04
Microsoft	2018	1	0	1	0
	2019	1	0	1	0
	2020	1	0	1	0
Facebook	2018	0.52	0.48	0.51	0.49
	2019	0.24	0.76	0.19	0.81
	2020	0.24	0.76	0.17	0.83
Cloudflare	2018	0.54	0.46	0.54	0.46
	2019	0.57	0.43	0.56	0.44
	2020	0.51	0.49	0.49	0.51

IPv4 and IPv6 queries proportion

UDP vs TCP

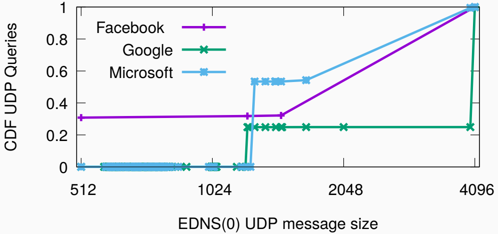
- UDP dominates
- TCP for large queries
- Facebook does more TCP (from 2019 onwards). Why?

		.nl		.nz	
	Year	**UDP**	**TCP**	**UDP**	**TCP**
Google	2018	1	0	1	0
	2019	1	0	1	0
	2020	1	0	1	0
Amazon	2018	1	0	0.98	0.02
	2019	0.98	0.02	0.96	0.04
	2020	0.95	0.05	0.95	0.05
Microsoft	2018	1	0	1	0
	2019	1	0	1	0
	2020	1	0	1	0
Facebook	2018	0.79	0.21	0.52	0.48
	2019	0.85	0.15	0.83	0.17
	2020	0.86	0.14	0.85	0.15
Cloudflare	2018	1	0	1	0
	2019	0.99	0.01	1	0
	2020	0.98	0.02	0.99	0.01

UDP and TCP queries proportion

Why Facebook queries more TCP than others

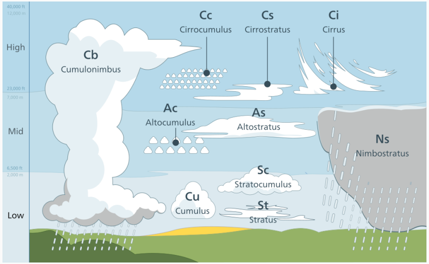
- 1/3 of Facebook queries: EDNS(0) UDP size < 1024
- Sometimes caused truncated answers
- TCP required afterward



CDF of EDNS(0) UDP message size for `.nl` (w2020).

Conclusion: Clouds ain’t all the same

- DNS concentration:
  5 Clouds, 1/3 of ccTLD queries
- Technology adoption varies significantly
  - DNSSEC
  - Transport
  - Routing
- Centralization
  - Pro: new security feature deployments
    benefits many users all at once
  - Con: if it breaks, it can
    affect many users all at once
- **Questions?**



***real-world* cloud types**

Paper (IMC2020):
Download it here

References i

[1] Arends, R., Austein, R., Larson, M., Massey, D., and Rose, S.
DNS Security Introduction and Requirements.
RFC 4033, IETF, Mar. 2005.

[2] Arends, R., Austein, R., Larson, M., Massey, D., and Rose, S.
Protocol Modifications for the DNS Security Extensions.
RFC 4035, IETF, Mar. 2005.

References ii

[3] Arends, R., Austein, R., Larson, M., Massey, D., and Rose, S.

Resource Records for the DNS Security Extensions.

RFC 4034, IETF, Mar. 2005.

[4] Bortzmeyer, S.

DNS Query Name Minimisation to Improve Privacy.

RFC 7816, IETF, Mar. 2016.

References iii

[5] Hardaker, W.

What's in a name?

https://blog.apnic.net/2020/04/13/whats-in-a-name/.

[6] Thomas, M.

Chromium's impact on root dns traffic.

https://blog.apnic.net/2020/08/21/chromiums-impact-on-root-dns-traffic/.