DNS Centrality

Geoff Huston AM
APNIC

This Presentation

- What's the problem with centrality anyway?
- What does centrality in the DNS mean?
- How to measure DNS centrality
- What we measured
- What I think it means

All in 20 minutes!

Why pick on the DNS?

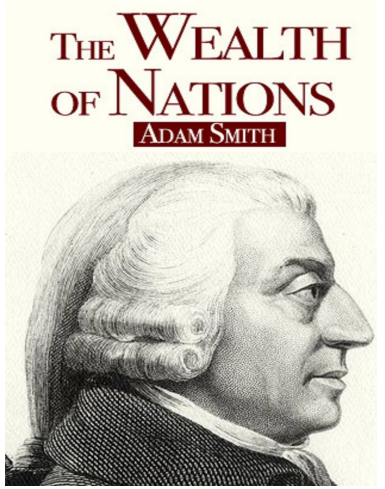
The DNS is **used by everyone and everything**

- Because pretty much everything you do on the net starts with a call to the DNS
- If a single entity “controlled” the entire DNS then to all practical purposes that entity would control not just the DNS, but the entire Internet!

Centrality is today's Internet

- Many aspects of the Internet's infrastructure are operated by fewer and fewer entities over time
  - Shift from entrepreneurial ventures to established business practices have largely driven these broad changes that have resulted in amalgamation and market concentration in many aspects of the Internet's service provision
- Competition has been replaced by Oligarchies

What's the problem?



- Economics A01 (or Adam Smith's Invisible Hand)
  - Competition rewards efficient producers
  - Innovation that increases production efficiency is rewarded
  - Consumers benefit from increased production efficiency and innovation
- Consolidation in the market
  - Distorts the functions of an open competitive market
  - Decreases competition pressure
  - Creates barriers to entry in the market
  - Reduces pressure for increased production efficiency and innovation
  - Consumers end up paying a premium

Consolidation in the DNS

It’s not a new topic:

- For many years BIND was a defacto monopoly provider for DNS software. At the time almost every DNS recursive resolver and authoritative server ran BIND software
- Due to a deliberate effort to broaden the DNS resolver space from a monoculture to a richer space, this picture has broadened out to a number of DNS software platforms and is less of a concern these days

- But there are many other places where the DNS is aggregating
  - Name Registration services
  - Name Hosting service providers
  - Name Resolution providers

Let's Focus!

- Here we are going to concentrate on just one of these areas
- We will look at the recursive resolver market and try to understand the extent to which we are seeing consolidation of the recursive name resolution function
- And then assess to what extent this represents a source of concern in the DNS

Recursive Resolvers

- This function is generally bundled with an ISP's access service for public network services
  - Which means that there is already some level of consolidation in this space as the concentration of these DNS services follows the concentration of ISPs in the retail market

Visible ASNs: Customer Populations (Est.)

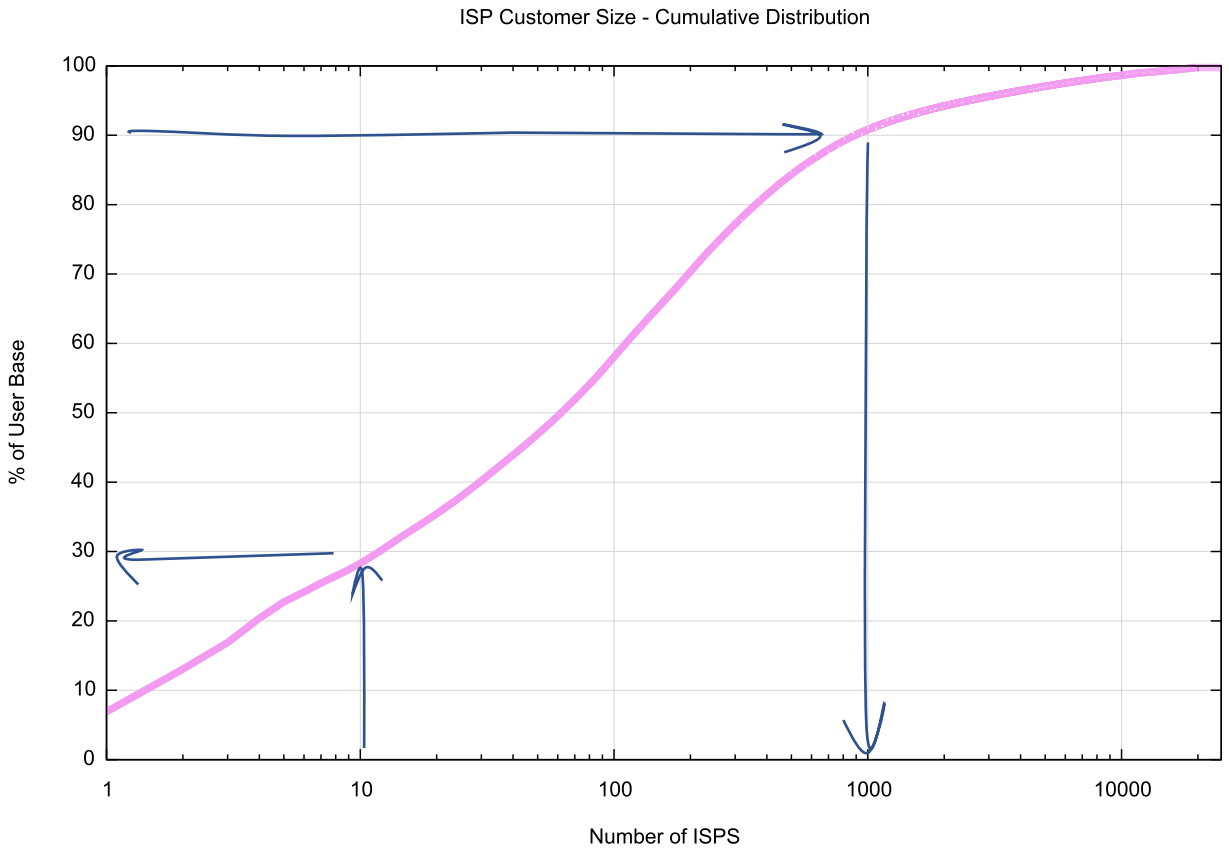
Rank	ASN	AS Name	CC	Users (est.)	% of country	% of Internet
1	AS55836	RELIANCEJIO-IN Reliance Jio Infocomm Limited	IN	288,557,917	48.69	6.902
2	AS4134	CHINANET-BACKBONE No.31,Jin-rong Street	CN	257,212,036	31.08	6.153
3	AS45609	BHARTI-MOBILITY-AS-AP Bharti Airtel Ltd. AS for GPRS Service	IN	159,869,634	26.97	3.824
4	AS4837	CHINA169-BACKBONE CHINA UNICOM China169 Backbone	CN	146,378,038	17.69	3.501
5	AS9808	CMNET-GD Guangdong Mobile Communication Co.Ltd.	CN	98,636,119	11.92	2.359
6	AS197207	MCCI-AS	IR	60,005,138	96.08	1.435
7	AS4812	CHINANET-SH-AP China Telecom (Group)	CN	53,344,384	6.45	1.276
8	AS7922	COMCAST-7922	US	42,464,383	16.98	1.016
9	AS38266	VODAFONE-IN Vodafone India Ltd.	IN	38,481,730	6.49	0.921
10	AS7713	TELKOMNET-AS-AP PT Telekomunikasi Indonesia	ID	38,308,027	32.57	0.916
11	AS4808	CHINA169-BJ China Unicom Beijing Province Network	CN	38,151,882	4.61	0.913
12	AS8151	Uninet S.A. de C.V.	MX	36,536,795	40.07	0.874
13	AS8452	TE-AS TE-AS	EG	35,732,468	67.3	0.855
14	AS29465	VCG-AS	NG	32,678,848	51.21	0.782
15	AS45271	ICLNET-AS-AP Idea Cellular Limited	IN	31,116,521	5.25	0.744
16	AS24757	EthioNet-AS	ET	27,065,478	100	0.647
17	AS56040	CMNET-GUANGDONG-AP China Mobile communications corporation	CN	25,845,793	3.12	0.618
18	AS9299	IPG-AS-AP Philippine Long Distance Telephone Company	PH	24,900,902	33.92	0.596
19	AS28573	CLARO S.A.	BR	23,357,492	14.2	0.559
20	AS2516	KDDI KDDI CORPORATION	JP	22,891,043	20.26	0.548
21	AS23693	TELKOMSEL-ASN-ID PT. Telekomunikasi Selular	ID	22,538,829	19.16	0.539

https://stats.labs.apnic.net/aspop

Aside: Concentration in the retail ISP market

The ISP retail access market is already heavily concentrated/centralised:

- 10 ISPs serve some 30% of the Internet's user base
- 90% of users are served by 1,000 ISPs



DNS Recursive Resolvers

- This function is generally bundled with an ISP's access service for public network services
  - So we would expect to see a level of concentration in recursive resolvers in line with the concentration in the ISP access market
- The question is: Is there consolidation in the DNS recursive resolution function **over and above** the existing access market consolidation?
- Where might we see such consolidation?

Open DNS Resolvers

- There are some 6M open DNS resolvers in operation today*
- Most of these appear to be inadvertently open due to errant CPE equipment
  - Where the resolver implementation does not correctly distinguish between “inside” and “outside” and provides a resolution service on all interfaces
- That may sound like a large number, but it has got a whole lot better over time!
  - 33M open resolvers were seen in 2013 **

* https://scan.shadowserver.org/dns/

** https://indico.dns-oarc.net/event/0/contributions/1/attachments/19/125/201305-dnsoarc-mauch-openresolver.pdf

Open DNS Resolvers as a Service

- Others are explicitly configured to offer DNS resolution services as a open service
  - Hard to say where all this started, but an early example was the the 4.2.2.2 open resolver project offered by BBN Planet in the mid-90's, though there were many others even then
    - At that time many ISPs used recursive resolvers as a service and some operated these platforms as a open service as a least cost / lowest admin overhead option
  - The use of anycast in the DNS made it possible to operate a single service with a distributed footprint
  - OpenDNS was one of the early offerings of a dedicated recursive resolution service with a scaled up infrastructure
  - Google Public DNS entered the picture with a service that took scaling to the next level

* https://scan.shadowserver.org/dns/

What's the Centrality Question here?

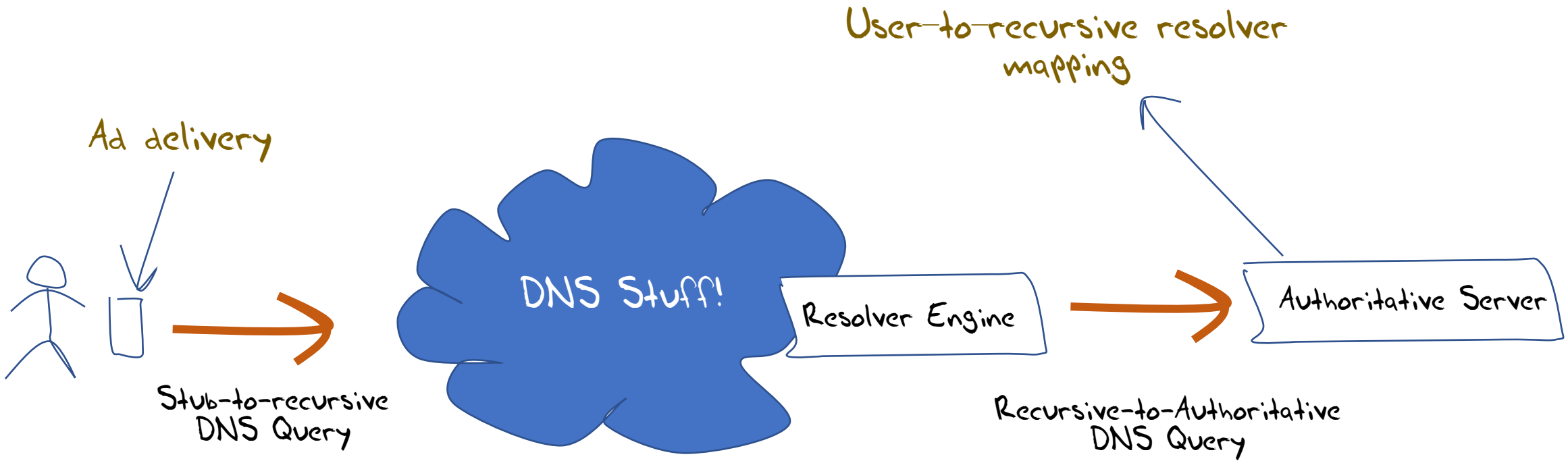
- One way to measure centrality is by “market share”
- So the market share question here would be: What proportion of users of the Internet use <X> as their DNS resolver?
  - We won’t distinguish between end users explicitly adding their own DNS configuration into their platform and ISPs using forwarding structures to pass all DNS queries to an open resolver. Through the lens of “centrality” both paths to using open DNS resolvers look the same!

How we* Measure DNS Centrality

- We use Google Ads as the main element of this measurement
  - The measurement script is an embedded block of HTML5 code in an Ad
  - The Ad runs in campaigns that generate some 10M impressions per day
  - We get to “see” the DNS in operation from the inside of most mid-to-large ISPs and service providers across the entire Internet
- Ads provide very little functionality in the embedded scripts – it’s basically limited to fetching URLs
  - But that’s enough here, as a URL fetch involves the resolution of a domain name
  - So we use unique DNS names in every ad, so the DNS queries will be passed though to our authoritative servers

* by “we” I mean APNIC Labs!

How we Measure DNS Centrality



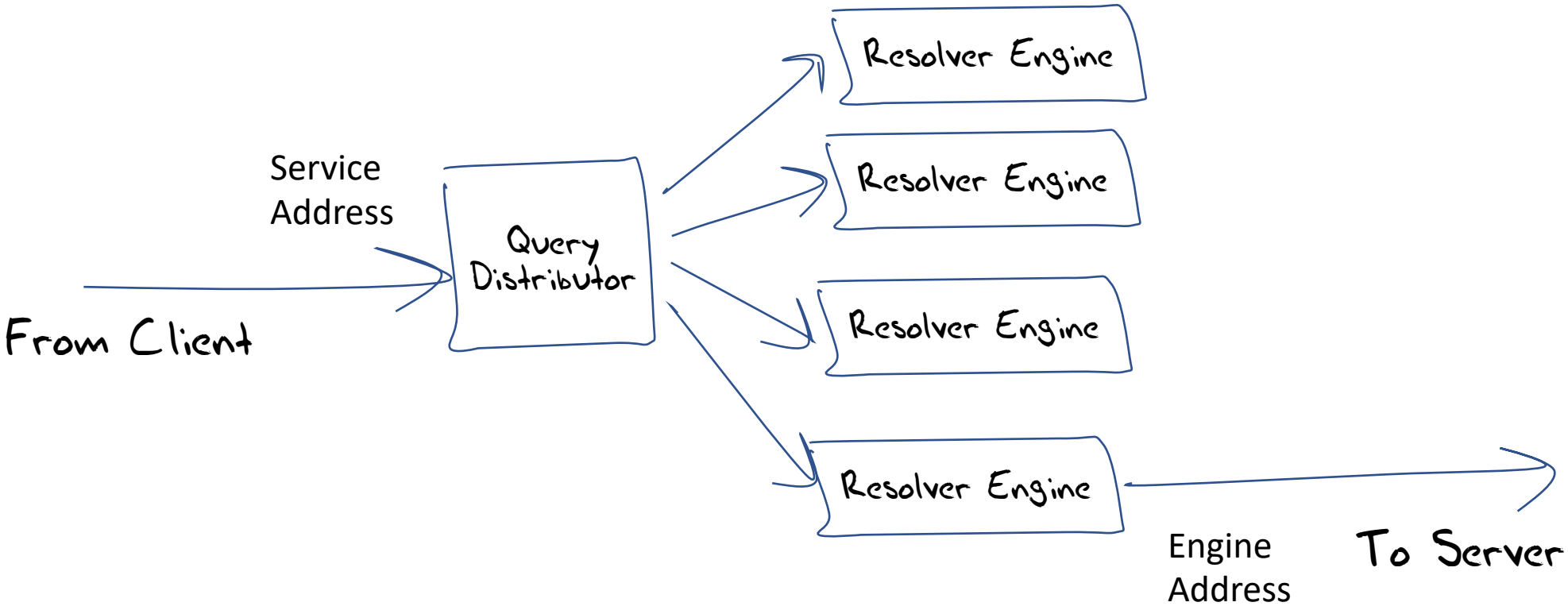
Recursive Resolver Behaviours

- The task is to match the source of a query of a domain name to both a resolver and an end user
- We need to
  - map query IP source addresses to resolvers
  - understand how the DNS “manages” queries
  - how the resolver lists in /etc/resolv.conf are used

Mapping Resolver Addresses

- We use periodic sweeps with RIPE Atlas to reveal the engine addresses used by popular Open DNS resolvers, and load this into an identification database

Understanding Resolver Behaviour

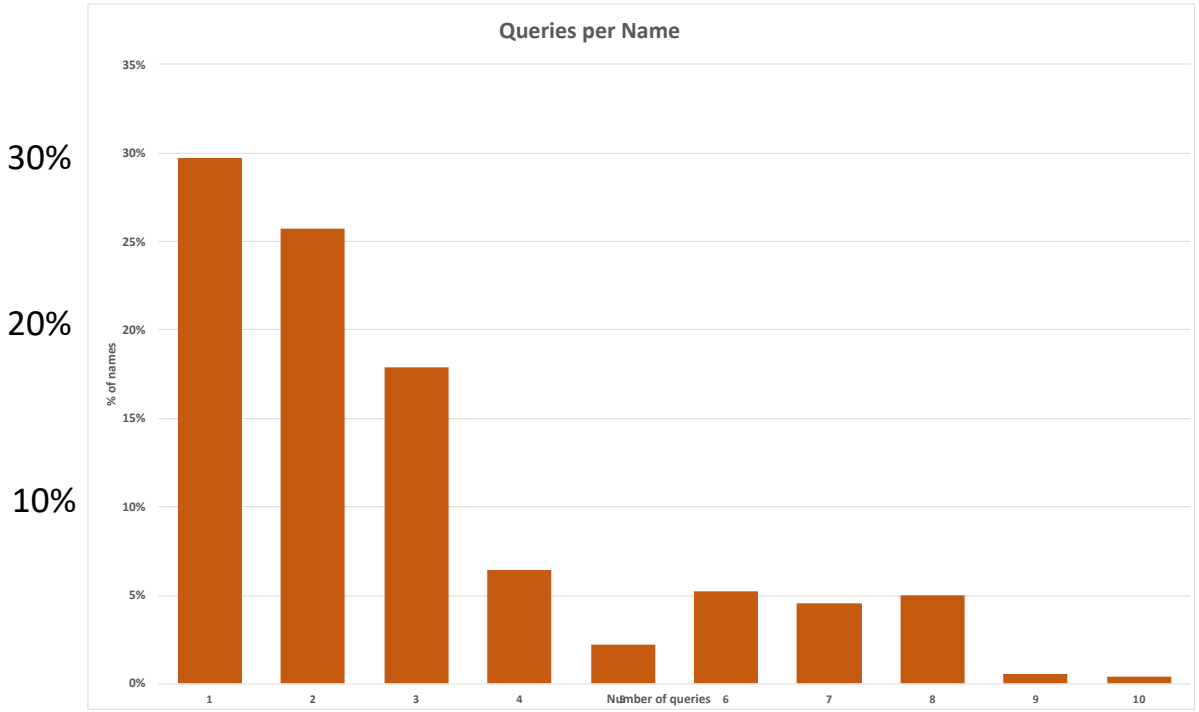


Resolution Metrics

- Average query count per unique name: 3.4

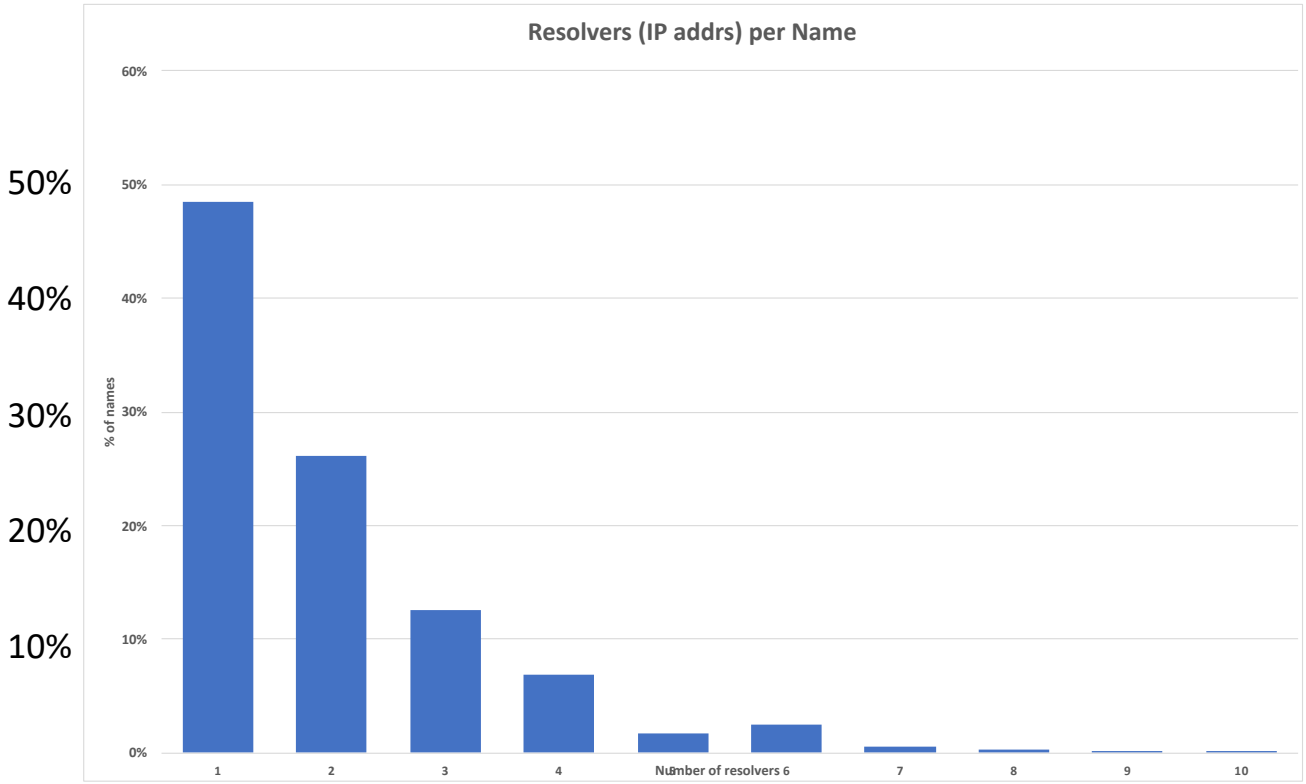
(Dual stack hosts may be a factor here)

  - Max observed query count in 30 seconds is 1,761 queries!



Resolution Metrics

- Average number of resolvers (IP addresses) per unique name: 2.1
- 30 second maximum resolvers seen: 94

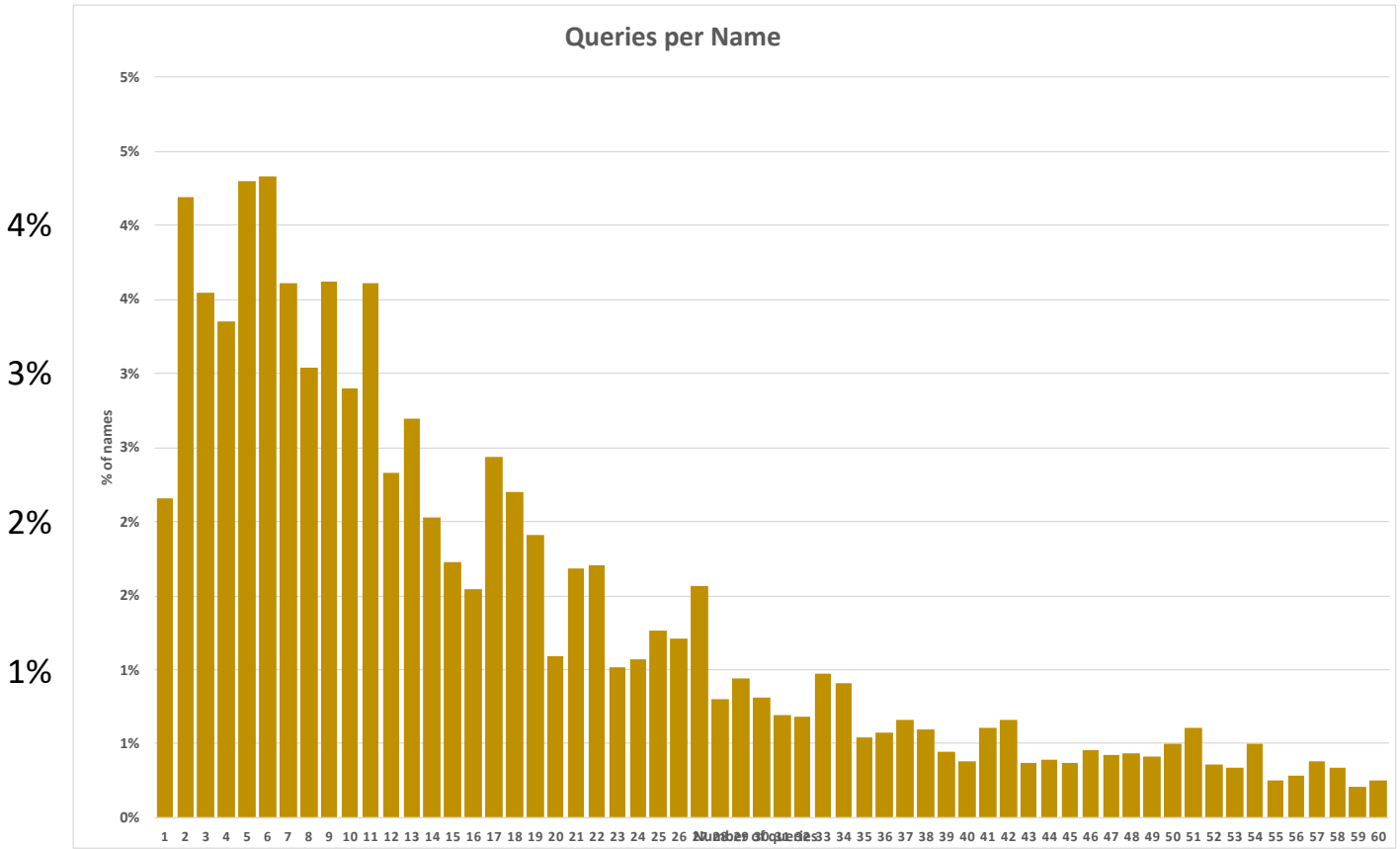


First Resolver vs Full Resolver Set

- What happens if the authoritative server always reports SERVFAIL to all queries?
- We use a server that always returns a SERVFAIL error code to prompt the client to run through its full set of recursive resolvers

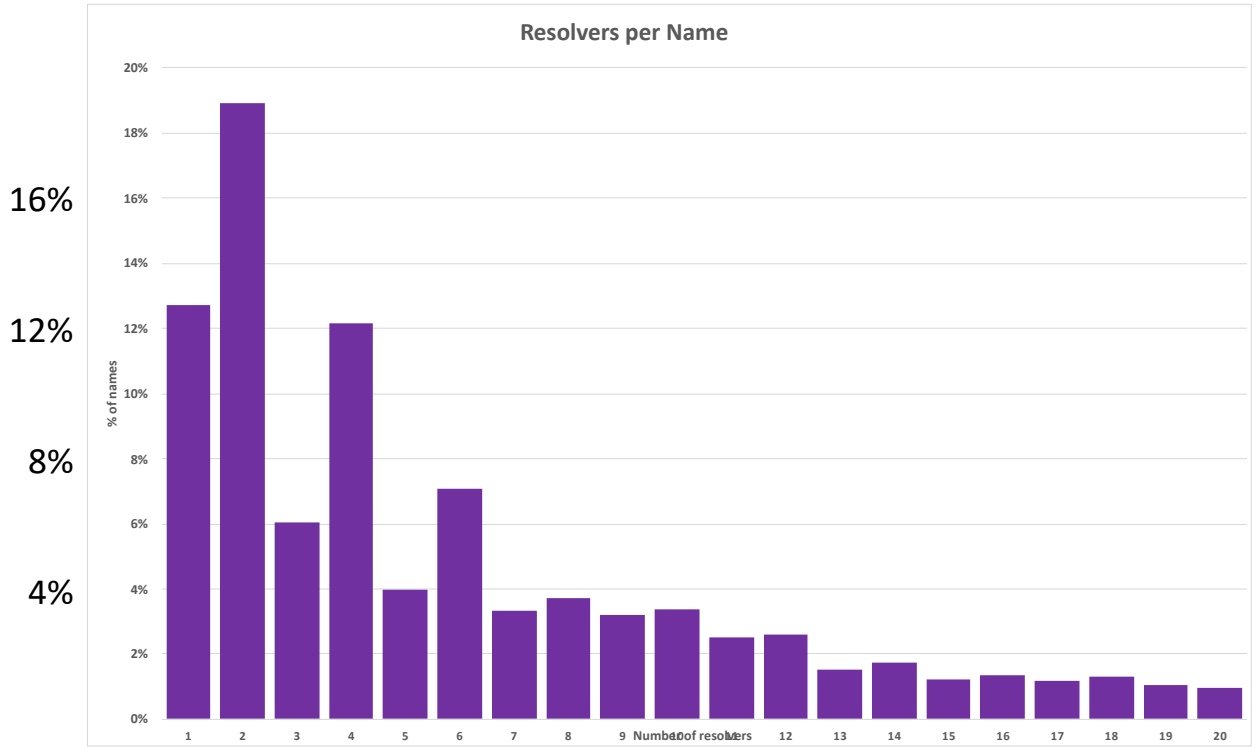
SERVFAIL Resolution Metrics

- Average query count per unique name: 36.5
  - Max observed query count in 30 seconds is 292,942 queries! (yes, really!)

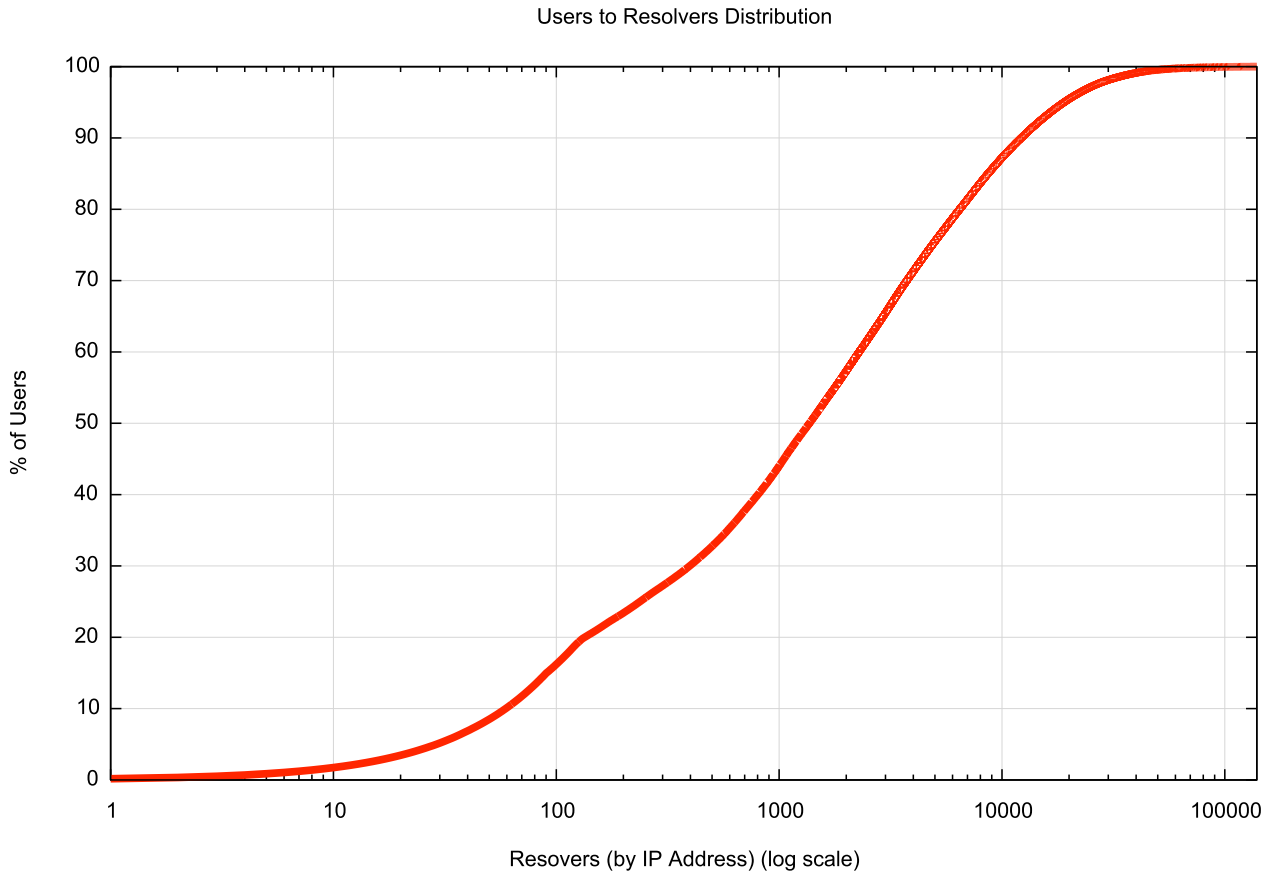


SERVFAIL Resolution Metrics

- Average number of resolvers (IP addresses) per unique name: 8.9
- 30 second maximum resolvers seen: 1,368



Recursive Resolver Stats



Of the 140,000 visible recursive resolvers, just 150 resolvers account for 20% of all users and 1,500 resolvers account for 50% of all users.

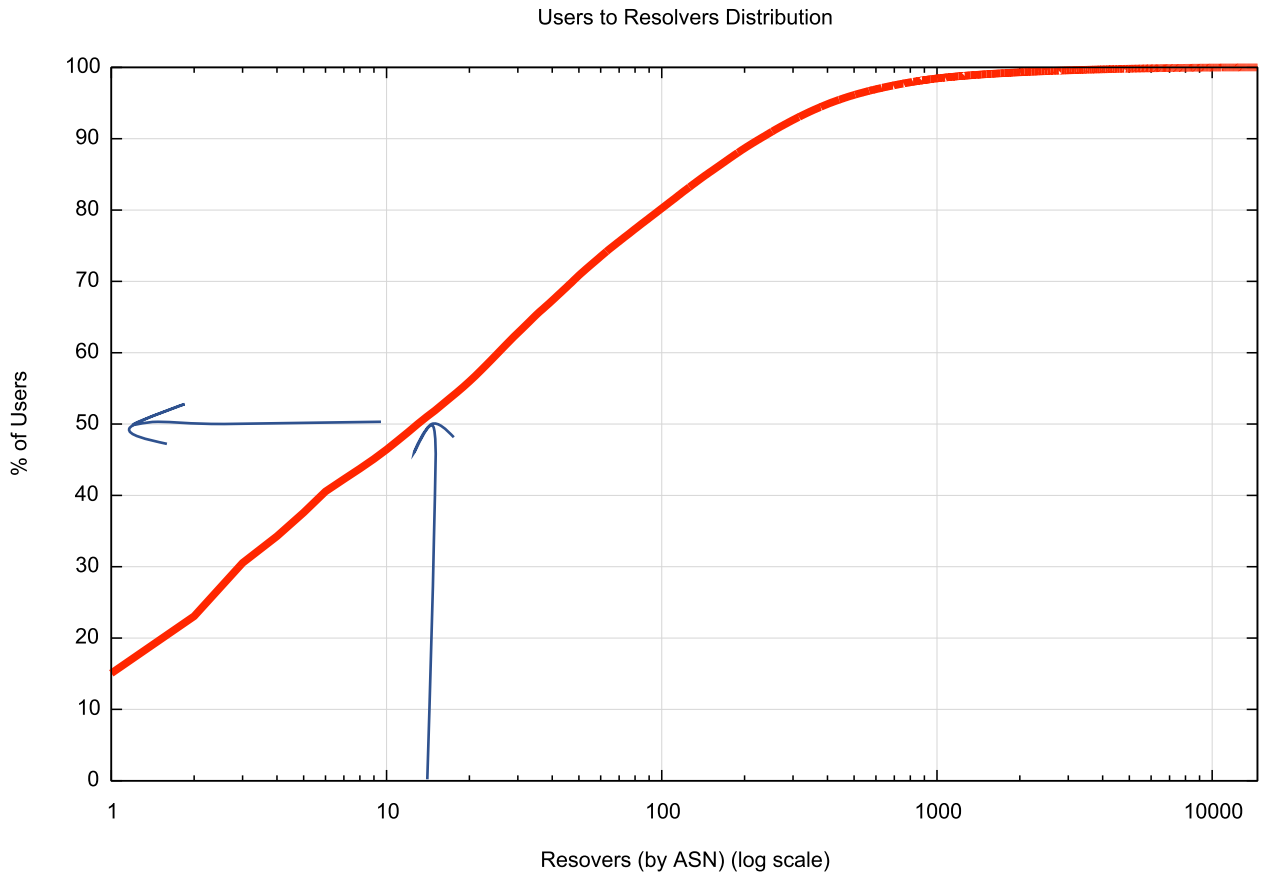
10,000 resolvers account for 90% of all users

However we are looking here at resolver IP addresses, and that's probably misleading.

Lets try and group resolver IP addresses into resolver **services***

* Use techniques of common origin AS, direct probing with RIPE Atlas and (occasionally) operator provided lists

Recursive Resolver Stats



Of the 14,600 visible recursive resolvers **services**, just 15 resolver **services** serve 50% of users

250 resolver **services** serve 90% of users

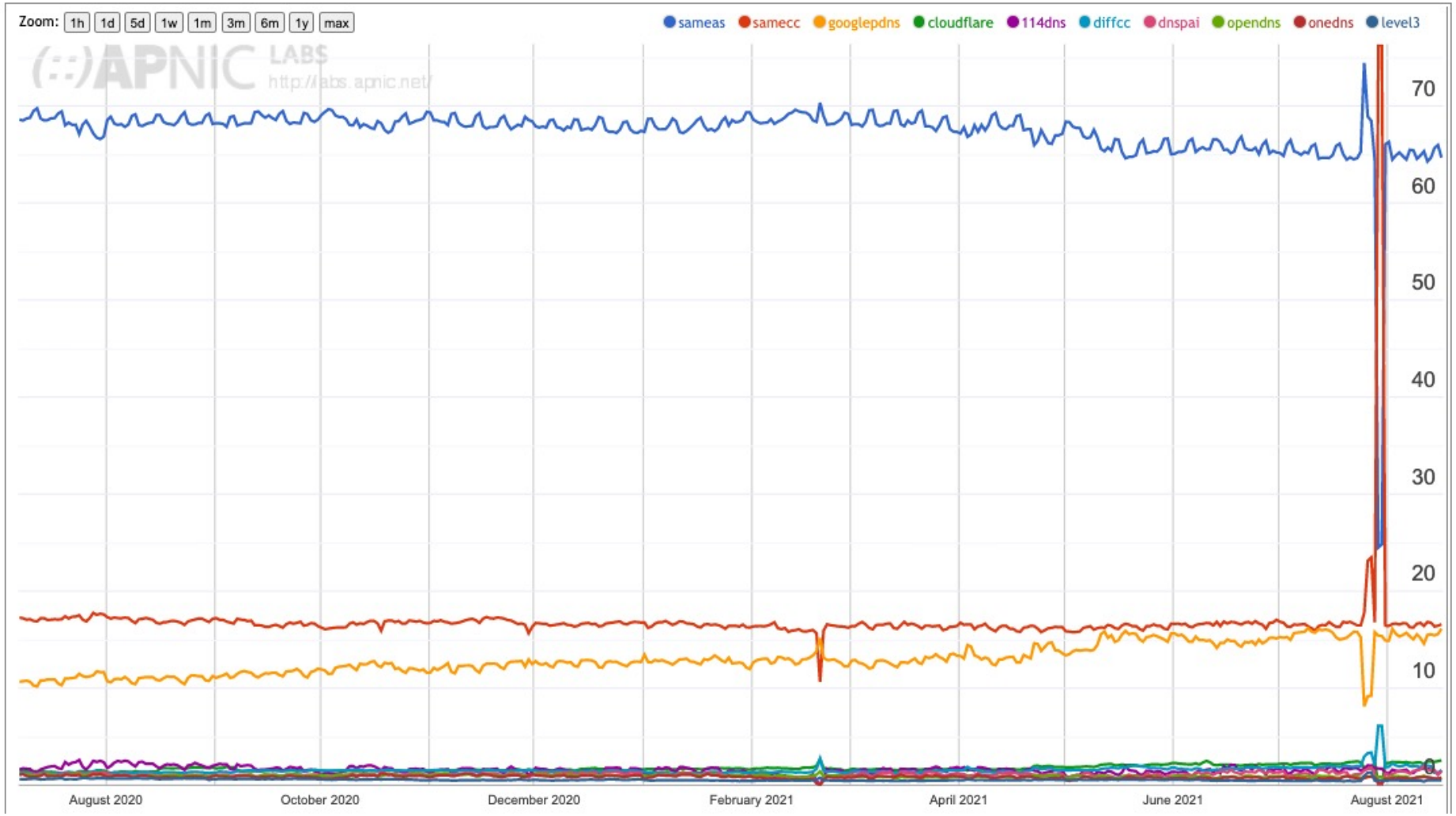
Is this what we mean by “centralisation”?

Details

Lets break this data down into:

- Using a “known” open DNS resolver
- Using a resolver in the same AS as the user
- Using a resolver in the same country as the user
- Others

"First" Resolver Use

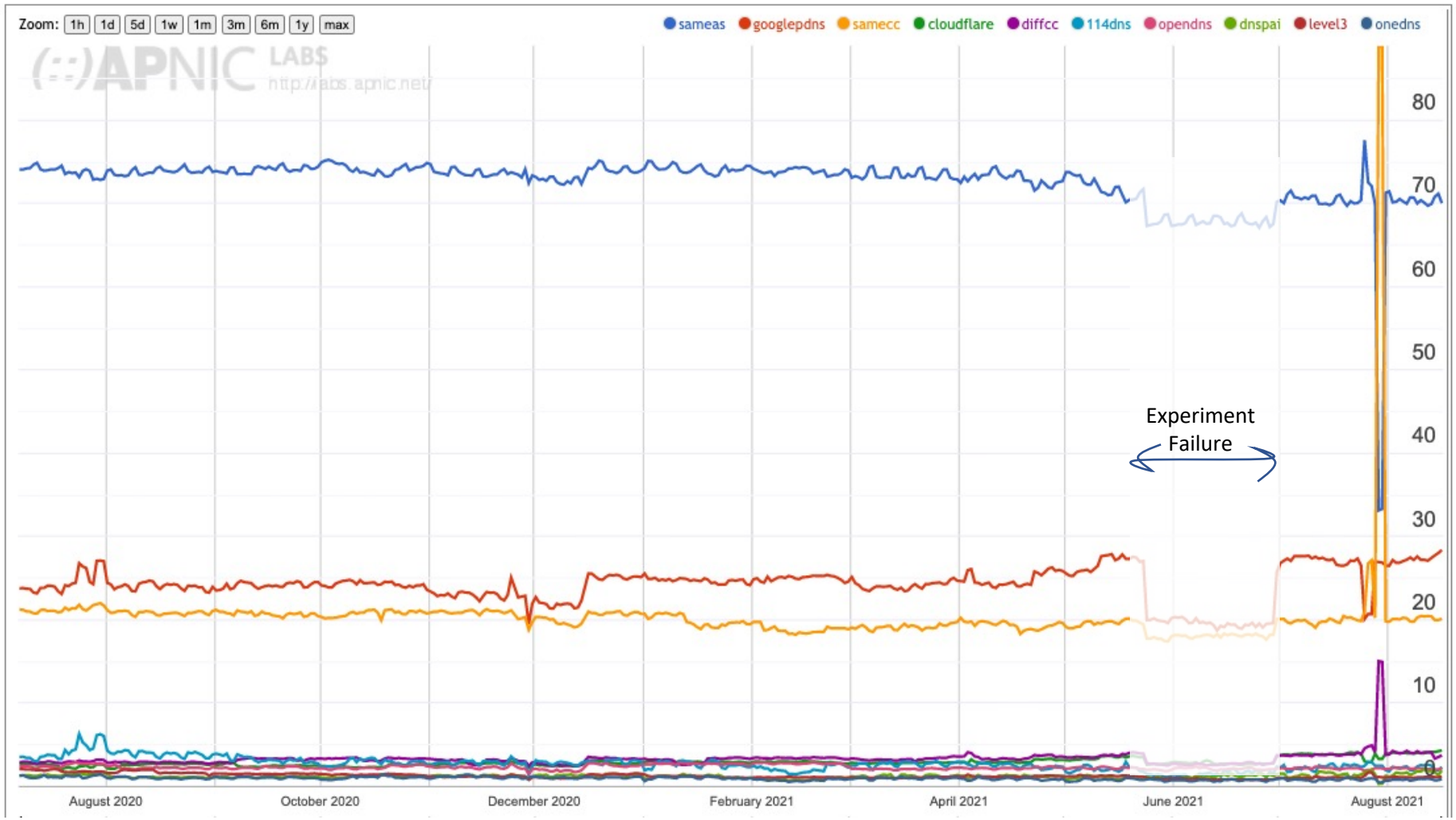


65÷ of users use a resolver located in the same AS as the user (ISP resolver)

17÷ of users use a resolver located in the same CC as the user (ISP resolver?)

17÷ of users use the Google open resolver (8.8.8.8)

All Resolver Use (SERVFAIL)



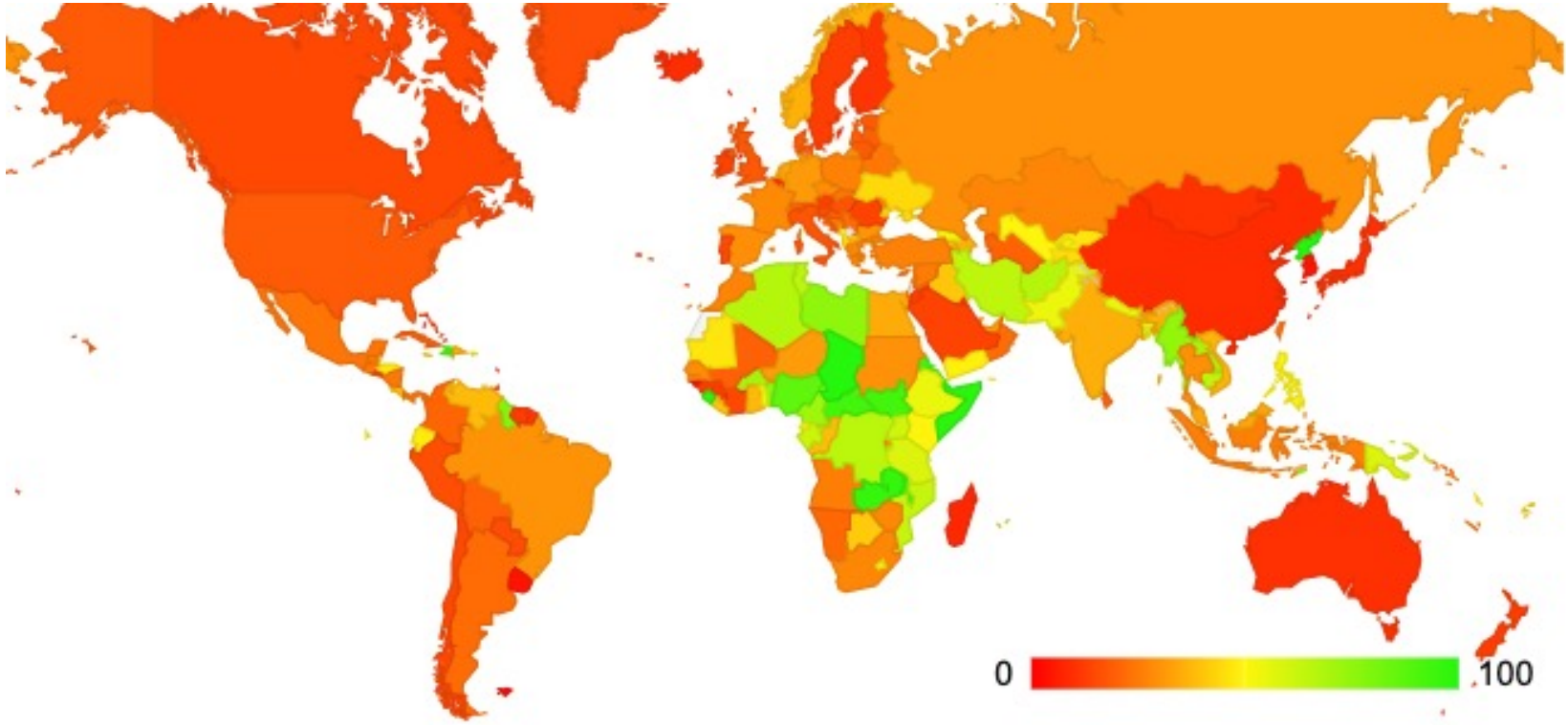
65÷ -> 69÷ for same ISP

15÷ -> 29÷ for Google use

(yes, the plotting software performed a colour change – sorry!)

Google DNS

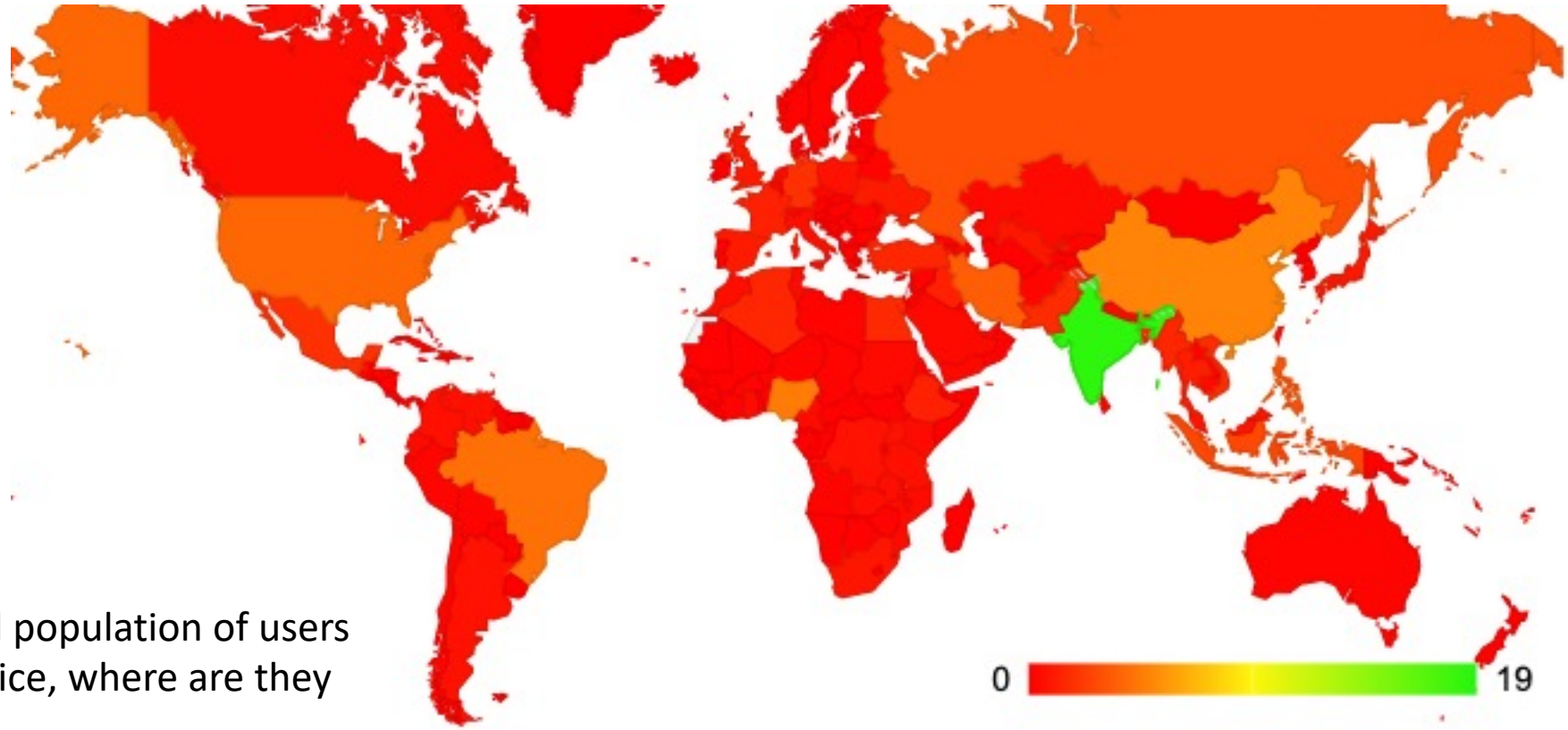
Use of Google Service per CC



Within each country how many users (%)
in that country use Google's resolver service?

Google DNS

Use of Google Service by User Count



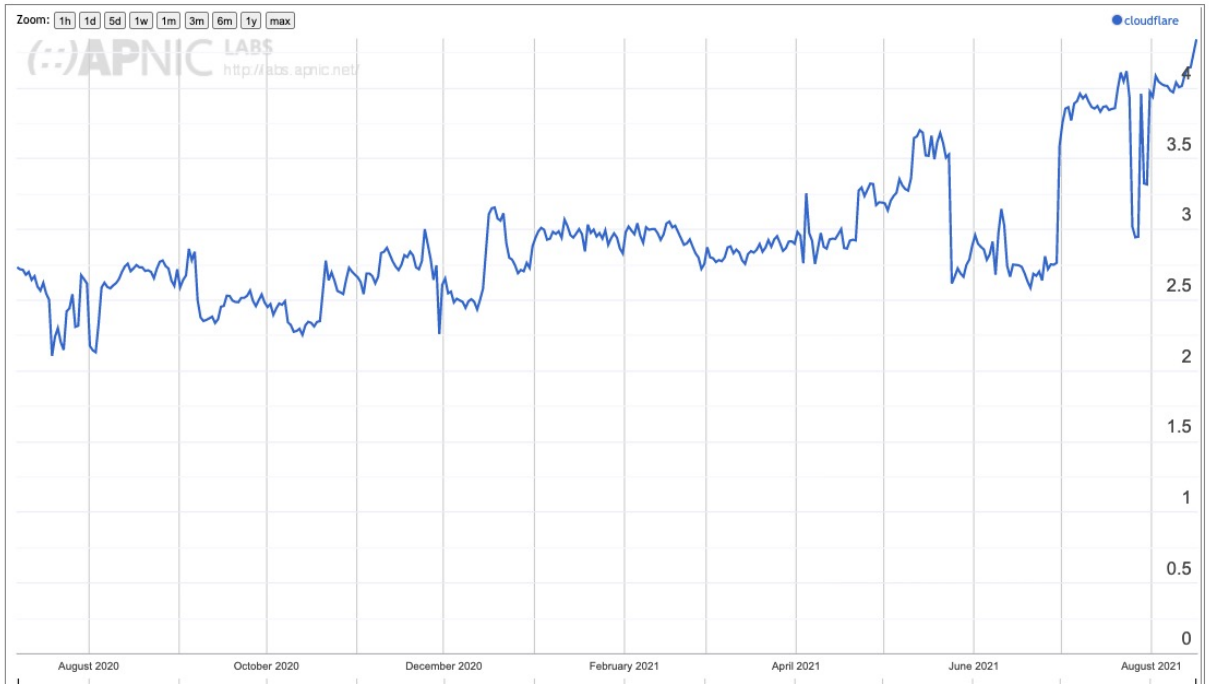
Looking at the total population of users using Google's service, where are they located?

Google DNS

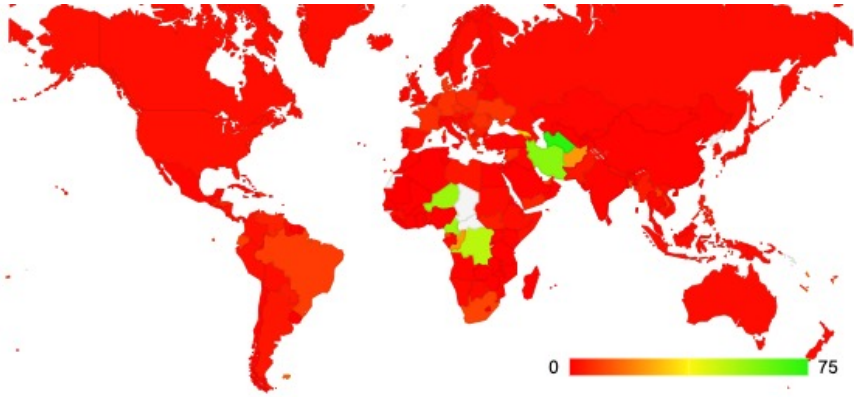
- Google DNS use appears to be equally split between first use (15% of users) and backup resolvers (a further 14% of users)
- Within each economy Google DNS is heavily used in some African economies, and central and southern Asian economies
- The largest pool of Google DNS users are located in India (19% of Google DNS users)
- Significant pools Google users are also seen in the US, China, Nigeria, Brazil and Iran (each CC has some 4% - 6% of Google's DNS users)

Cloudflare's 1.1.1.1 service

Cloudflare market share

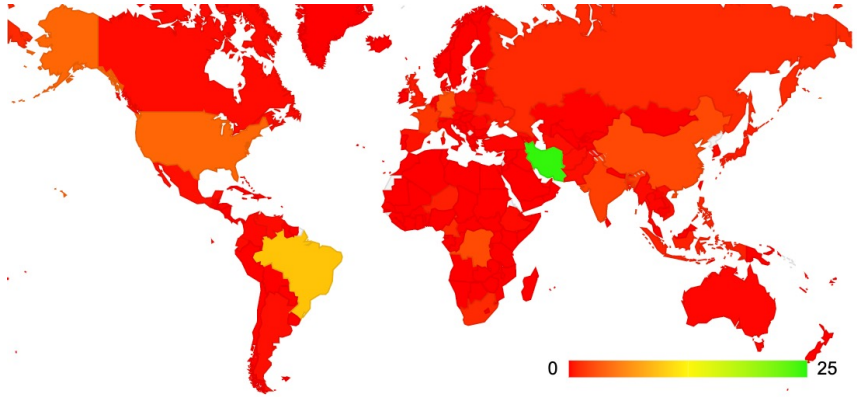


Where is Cloudflare used?



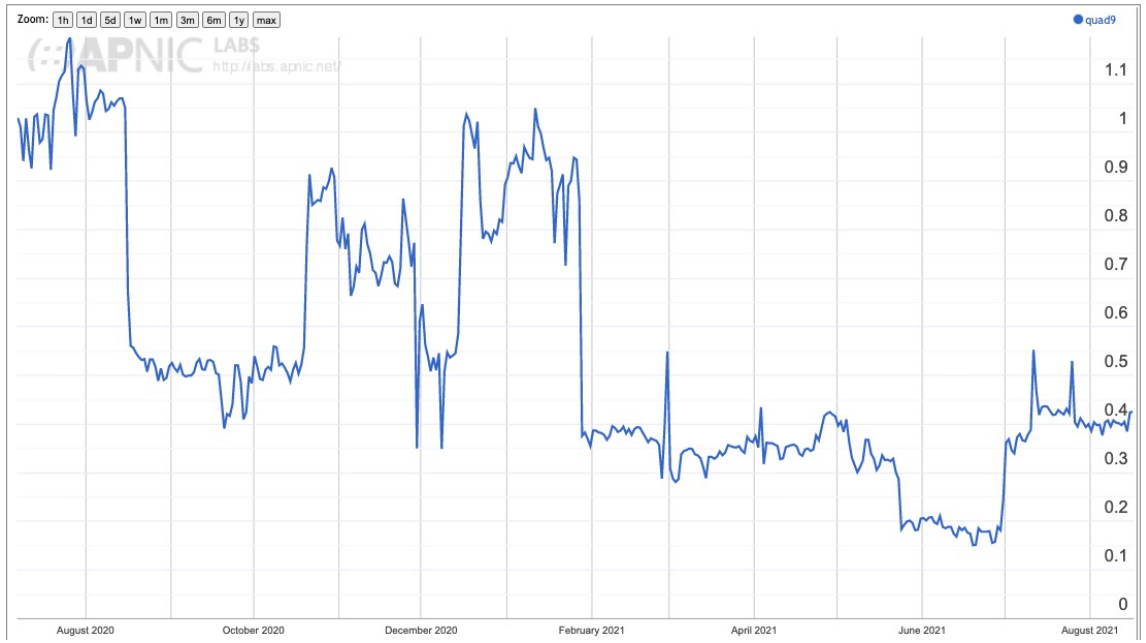
Cloudflare is extensively used in Turkmenistan (80%), Iran (57%), Niger (54%) Cameroon (54%) and the Congo (49%)

Cloudflare User breakdown?

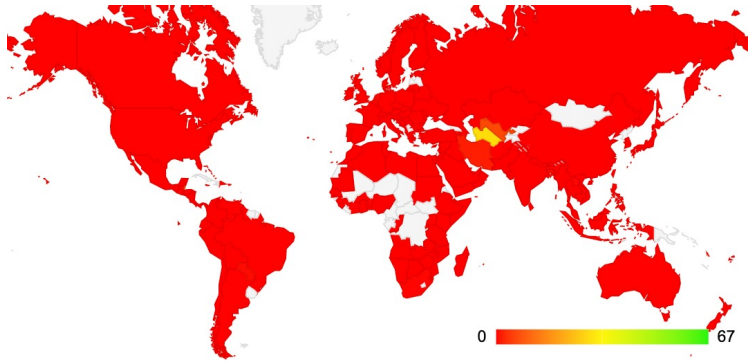


Quad9 service

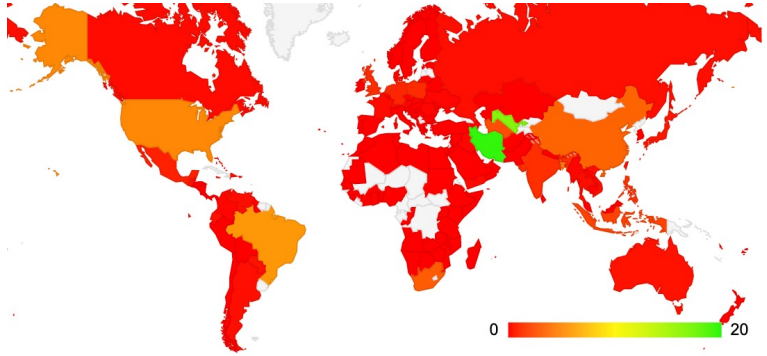
Quad9 market share



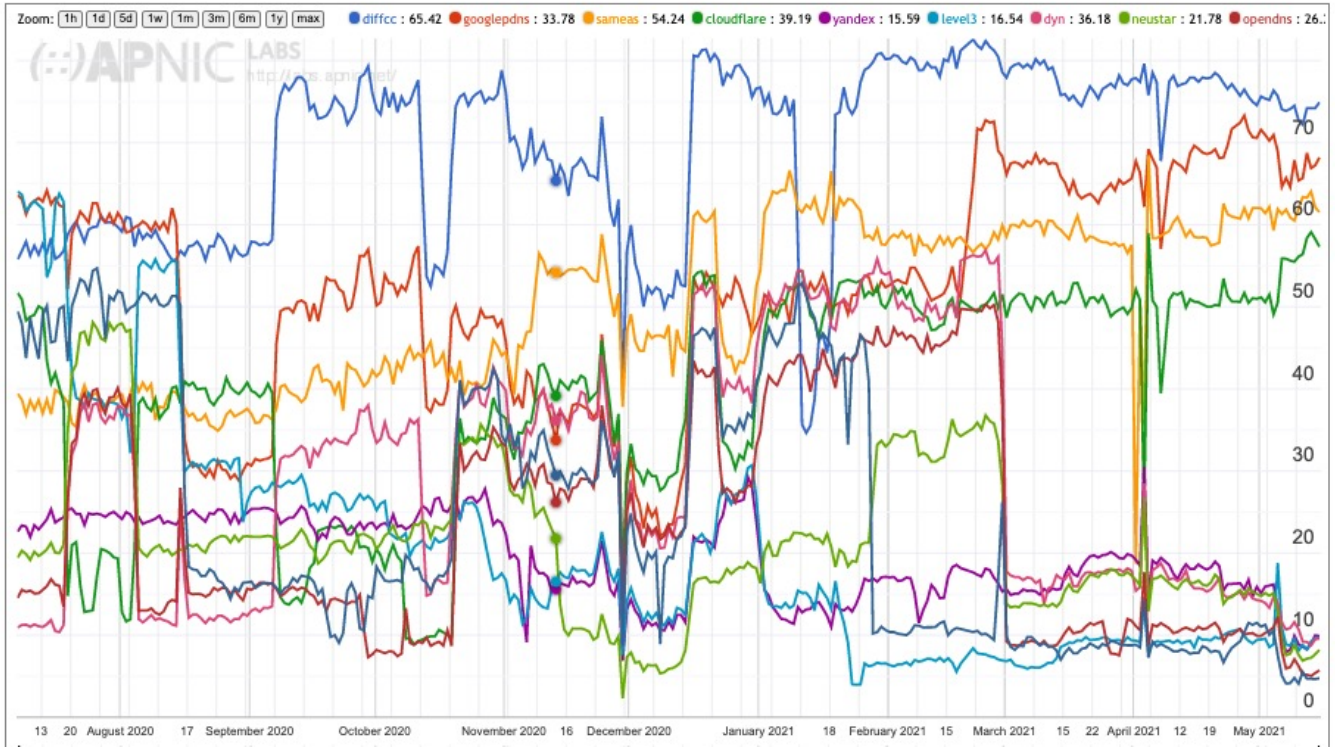
Where is Quad9 used?



Quad9 User breakdown?



Iran



A major ISP in IRAN, MCCI, distributes its queries across Google, Cloudflare, Yandex, Neustar, OpenDNS, Quad9 and others – all at once!

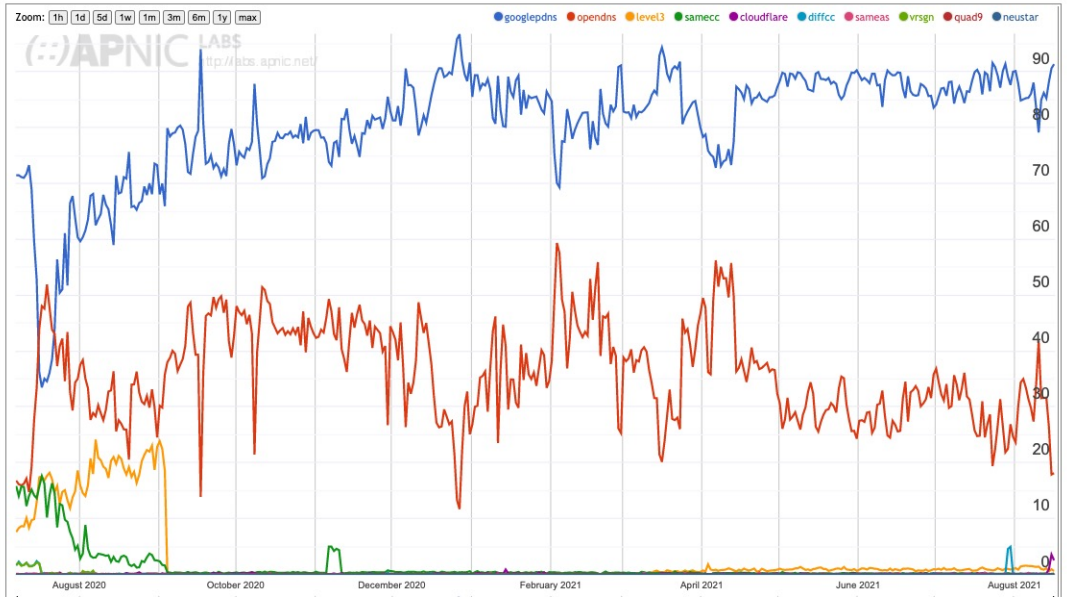
Who makes the choice?

- Is this the ISP‘s resolver performing forwarding of the query to an open resolver, or the users themselves opting out of the ISP service?
  - The numbers vary, but it is quite common to see 60% - 80% of users in an AS having their queries sent to an open resolver when open resolvers are used

Who makes the choice?

- Is this the ISP‘s resolver performing forwarding of the query to an open resolver, or the users themselves opting out of the ISP service?
  - The numbers vary, but it is quite common to see 60% - 80% of users in an AS having their queries sent to an open resolver when open resolvers are used

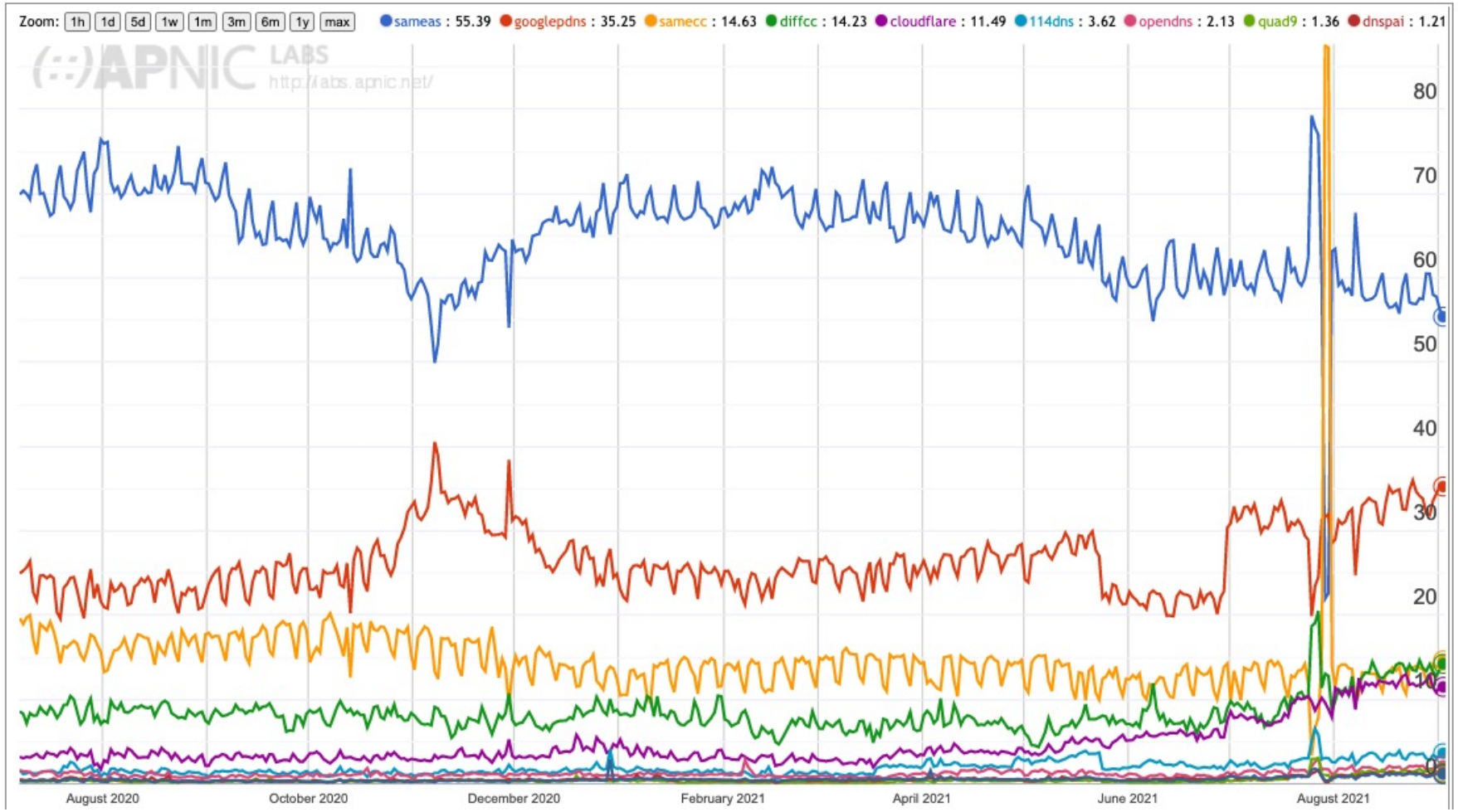
Resolver Use – AS131429, Mobifone, Vietnam



Google DNS at 90%

OpenDNS at 20%

DNS Use in Hong Kong



ISP Resolver

Google

Other HK
Other Country
Cloudflare

Resolver Centrality?

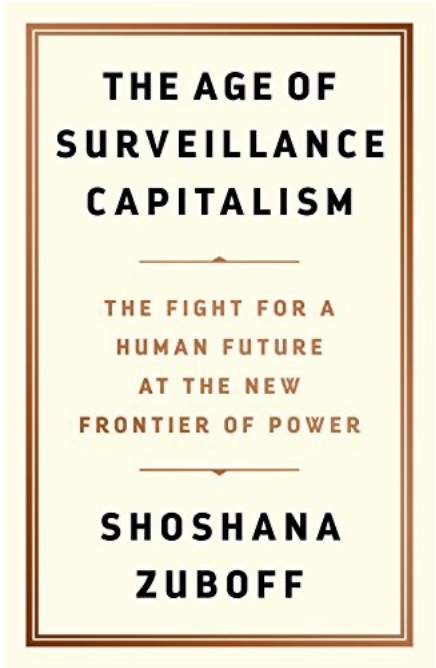
- Its not a “small number” of open resolvers
  - It’s actually just 1 – Google’s Public DNS
- Its not end users reconfiguring their devices
  - It’s the ISP
  - And where its not the ISP it’s mainly enterprise customers of ISPs
- Is this changing?
  - Yes, but quite slowly

Commentary and Opinions

What follows are my opinions!

Is this a centrality "problem"?

- Is this an emerging distortion of the market that puts excessive market control in the hands of a small set of providers?
  - A lot of users have the DNS queries passed on to authoritative servers via Google's open resolver service
  - But does this present us with issues?
    - 8.8.8.8 is fast, supports DNSSEC validation and does not filter or alter DNS responses (as far as I am aware)
    - Its cheap, its fast, its well managed, and it works reliably
    - So what's the issue here?



https://xkcd.com/1361/

What's the problem here?

- It's a sensitive issue these days
- There are many privacy undertakings in our space, but the undeniable fact is that many "free" services are indirectly funded through advertising revenue, and advertising is based on individual tracking and profiling
- Open DNS providers typically provide undertakings that they do not use their query traffic for profiling - and I have **no** evidence that these undertakings are not being adhered to

  But I still have some questions as a consumer of their services:
  - How are these undertakings audited and/or enforced? By whom?
  - Are there penalties for breaches of these undertakings?
  - Considering the size of these actors are any of these penalties even meaningful?

Barriers to Entry

- Why is there only 1 very large open DNS provider in today's Internet?
- Is it because the incumbent is raising the barriers of entry to all potential competitors?
  - Unlikely, as there is no evidence that this is the case
- Or are there "natural" barriers to entry?
  - Possibly – lets look

"Natural" Barriers to DNS Entry

- The DNS economy is such a financial wasteland that very few have a natural incentive to enter this market
  - **No one pays for queries!**
  - Selling query logs can very damaging in terms of reputation and liability – particularly when you cannot get the users' informed consent to do so
  - Selling NXDOMAIN substitution is also very damaging in terms of reputation
- It can be argued* that only someone with a massive presence is search has a commercial case for deploying a DNS resolver that is "honest" about the DNS (including NXDOMAIN)

* And some have from time to time

But…

Is all this a distraction?

- It's more likely that the shift of DNS functions into application realms using DoH services as an application function is a far greater threat to the current model of the DNS as a common single infrastructure
- Maybe the convergence of:
    - increased autonomy of applications in today's Internet
    - the dominant position of Android
    - the dominant position of Chrome

  poses a greater *potential* threat to the integrity of the name infrastructure of the Internet than the issue of recursive resolver use

Is this a problem?

- If applications develop their own application-centric name universes then what is the issue here?
- While we've used a single name space for the past 40 years or so, that does not mean that we have to keep on doing so
- And just because many others use a single name space as a means of surveillance and imposed control does not mean that we have to keep on using this single name space
  - Particularly when applications can be far more effective in cloaking the user's interactions in the application space than we've been able to do in a single DNS so far

Maybe this is a Good Thing

- Again, it could be argued that this deliberate fragmentation of a single cohesive name space for the Internet into a set of application-specific spaces is a shift that lies in protecting the interests of privacy and service integrity for users
- Its early days, and right now we just don’t know!

Watch This Space

The Internet remains one where the stasis of incumbent services has to compete against the pressure of continued innovation

While the pressures of aggregation and centralisation are evident in the space of common infrastructure services such as bandwidth, the DNS and CDN services, innovation in the Internet is moving up into the application space

We are inexorably moving to application-centric networking where both the network and the platforms are stripped back and the functionality and value is loaded into application-environments

Thanks!

Report on Resolver Use: https://stats.labs.apnic.net/rvrs