

Measuring Recursive Resolver Centrality

Geoff Huston, Joao Damas
APNIC Labs

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 DNS then to all practical purposes they would control not just the DNS, but the entire Internet!

This Presentation

- What's the problem with centrality anyway?
- What does centrality in the DNS mean?
- How to measure DNS centrality
- What we measure
- What are we see

Centrality

- Many aspects of the Internet's infrastructure are operated by fewer and fewer entities over time

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 every DNS recursive resolver and authoritative server ran Bind software. This has broadened out to a number of software platforms and is less of a concern today
- Where else might we find consolidation in today's DNS?
 - Name Registration services
 - Name Hosting service providers
 - Name Resolution providers

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

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
- Is there consolidation in the DNS recursive resolution function over and above this access market consolidation?
- Where might we see such consolidation?

The Rise of 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
- 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
 - Open DNS 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

What's the Centrality Question here?

- One way to measure centrality is by “market share”
- So the 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 through to our authoritative servers

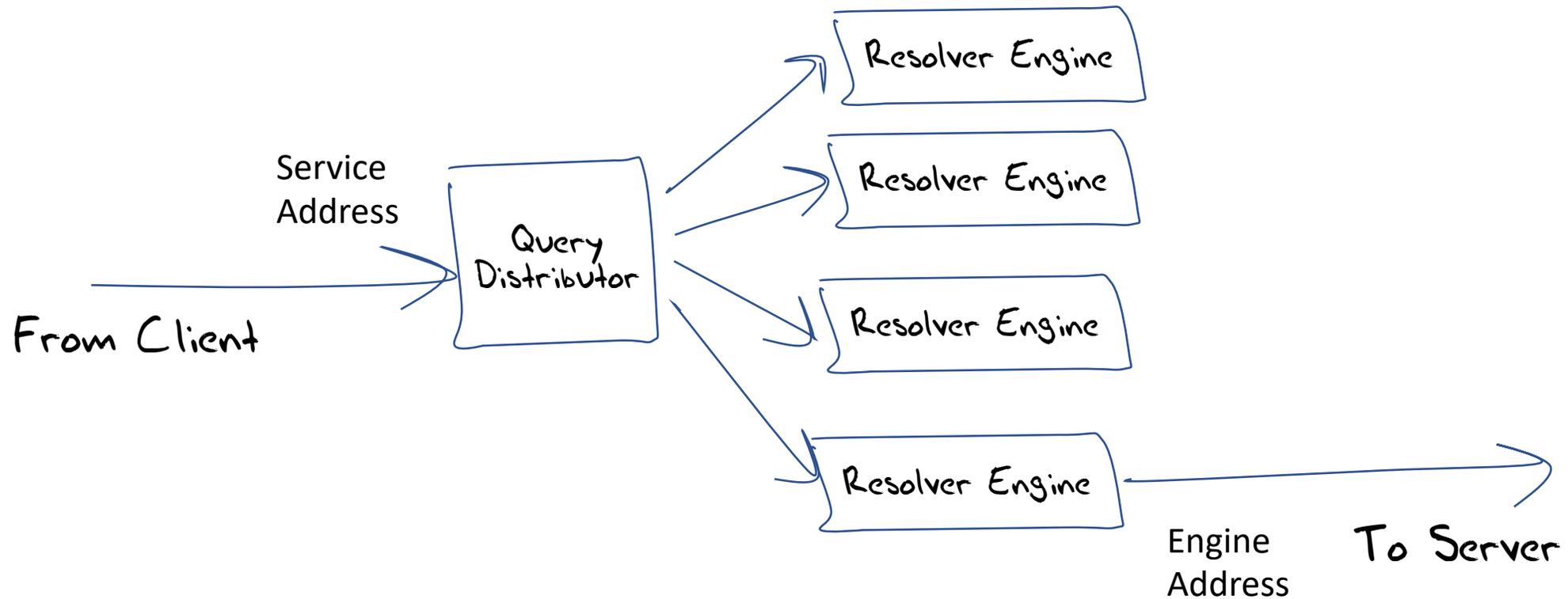
Recursive Resolver Behaviours

- 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

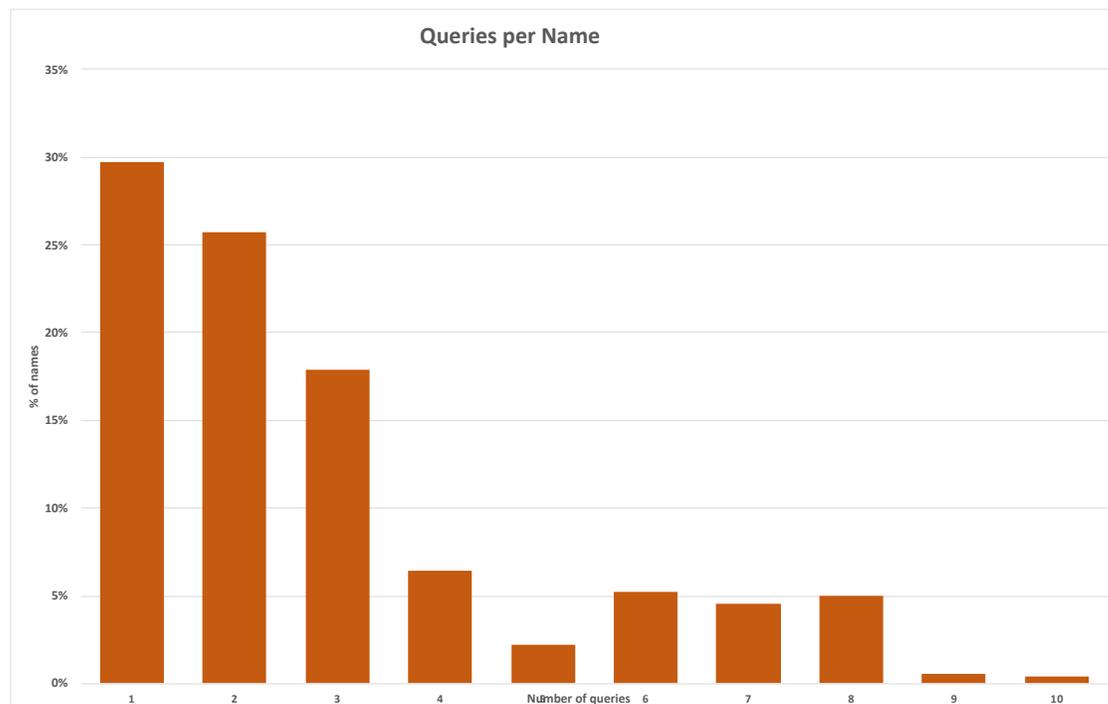
- We use periodic sweeps with RIPE Atlas to reveal the engine addresses used by popular Open DNS resolvers, and load this into an identification database

Understanding Resolver Behaviour



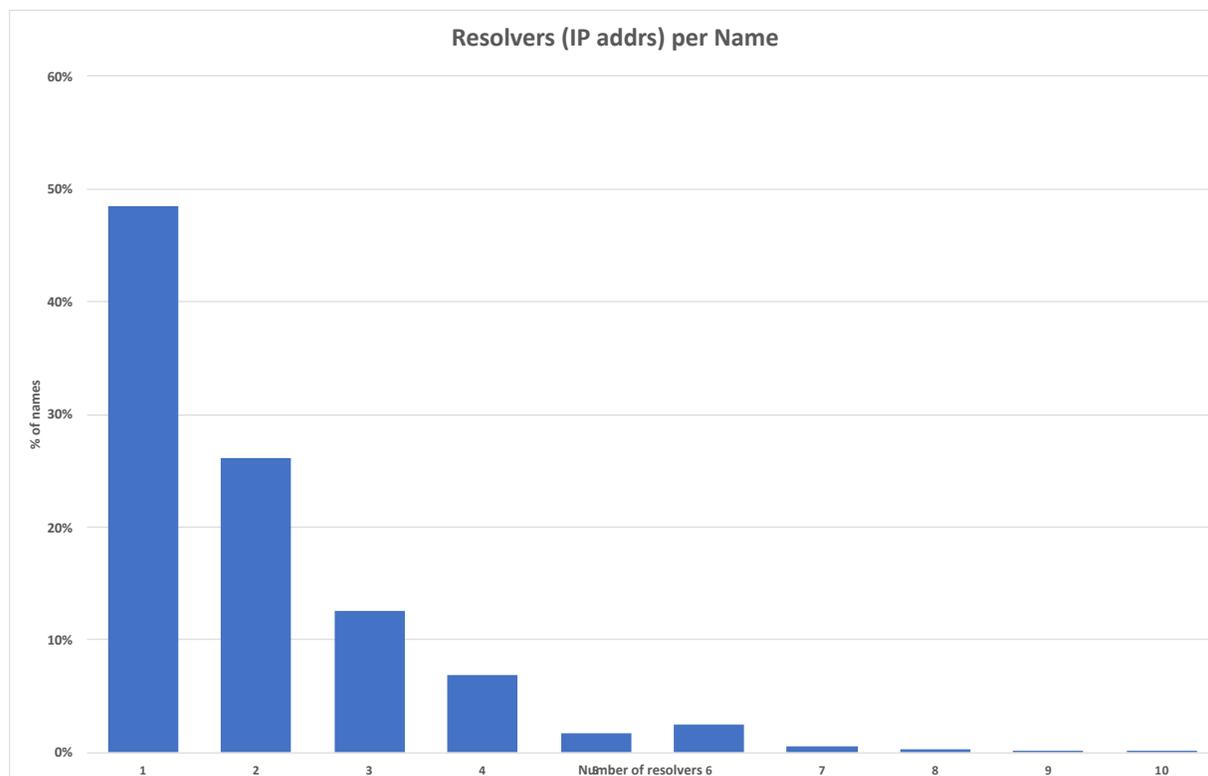
Resolution Metrics

- 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

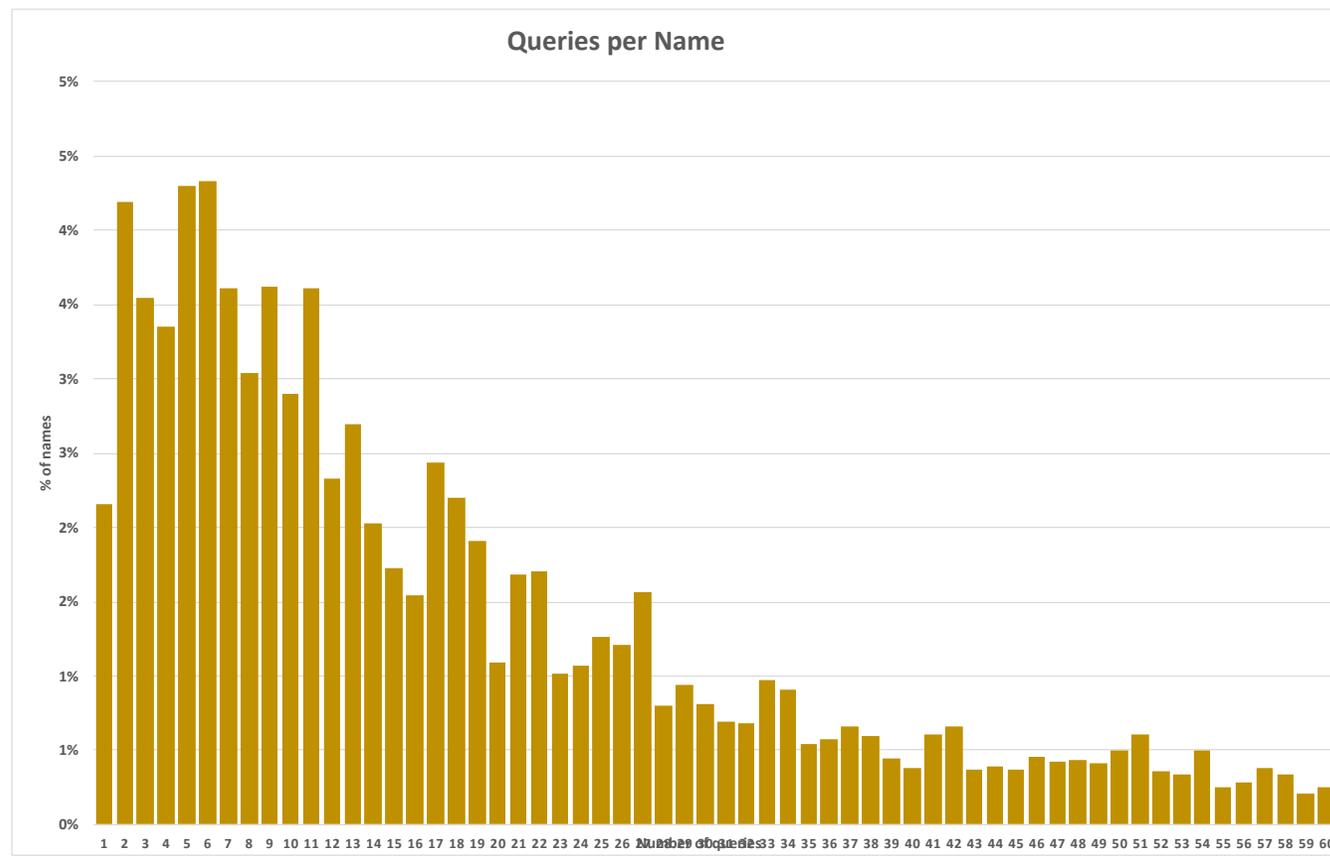


First Resolver vs Full Resolver Set

- 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

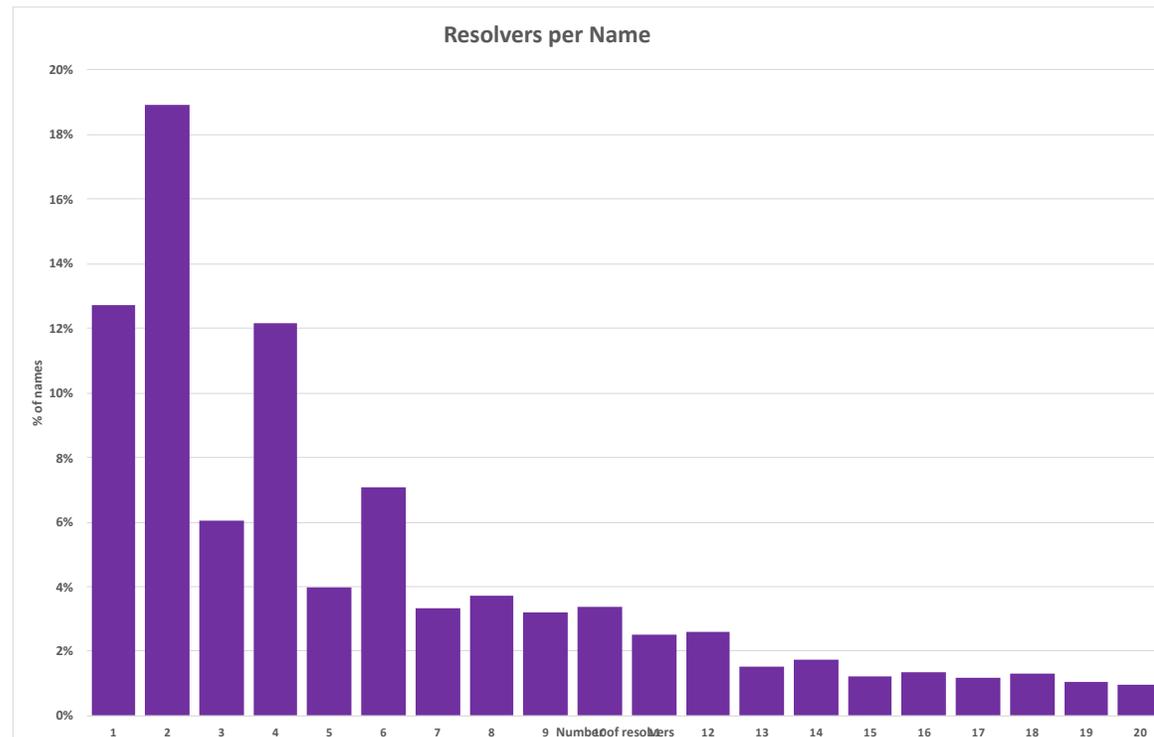
SERVFAIL Resolution Metrics

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

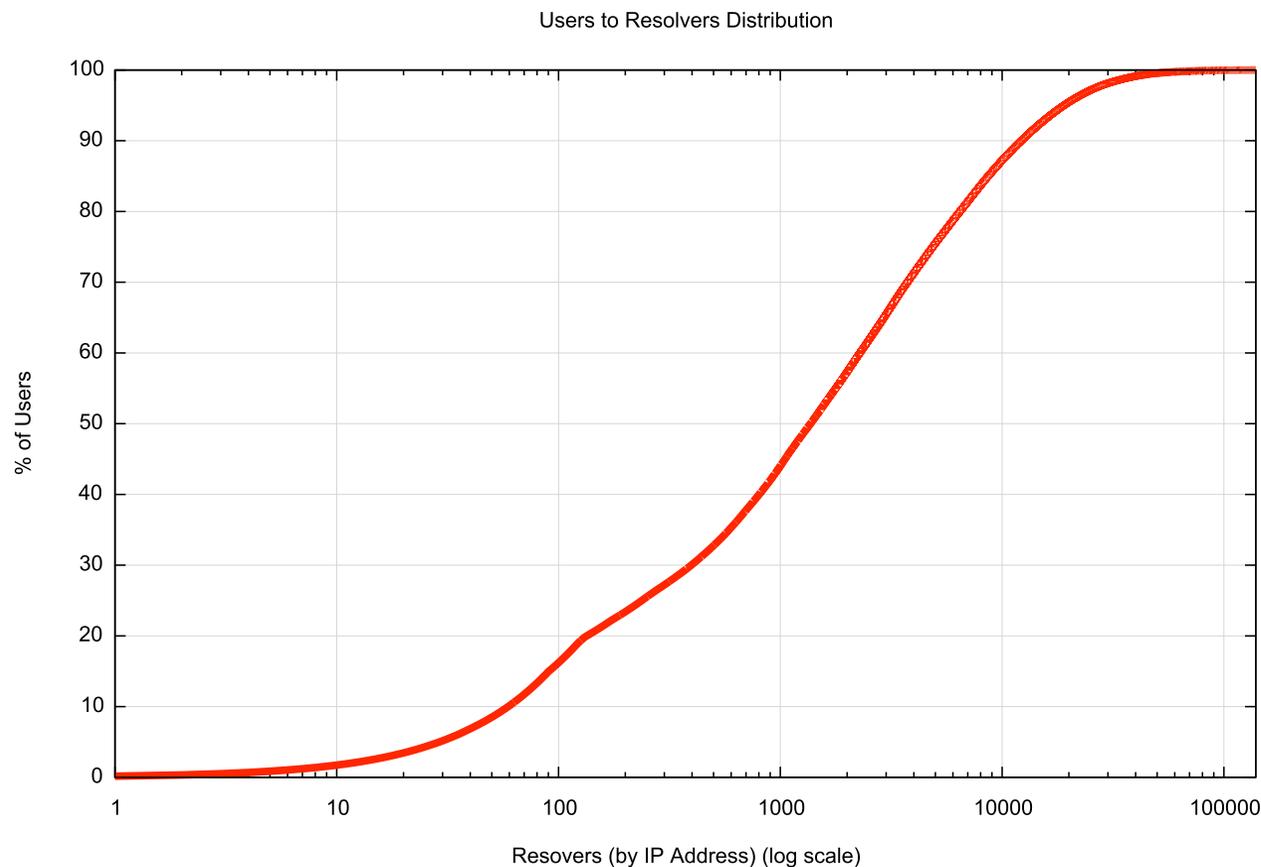


SERVFAIL Resolution Metrics

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



Recursive Resolver Stats



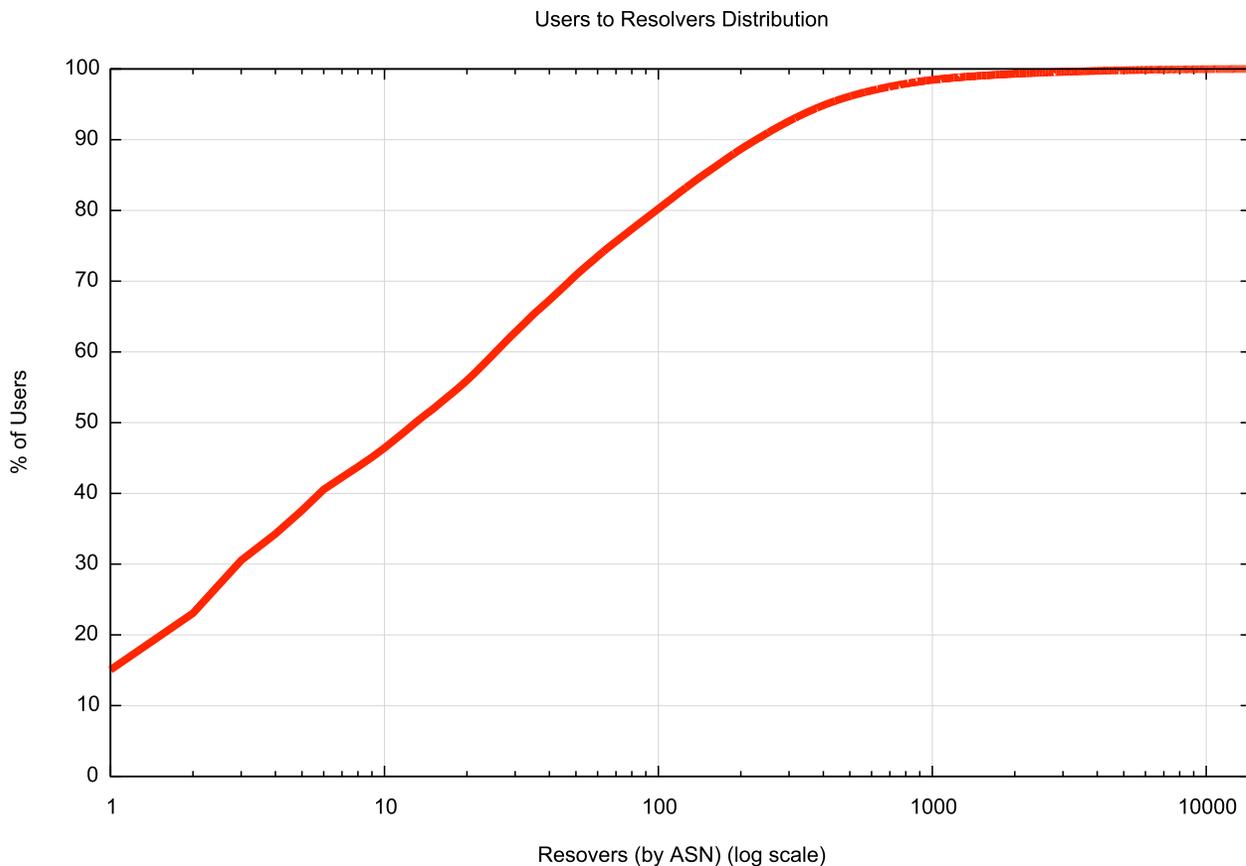
Of the 140,000 visible recursive resolvers, just 150 resolvers account for 20% of all users and 1,500 resolvers account for 50% of all users.

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

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

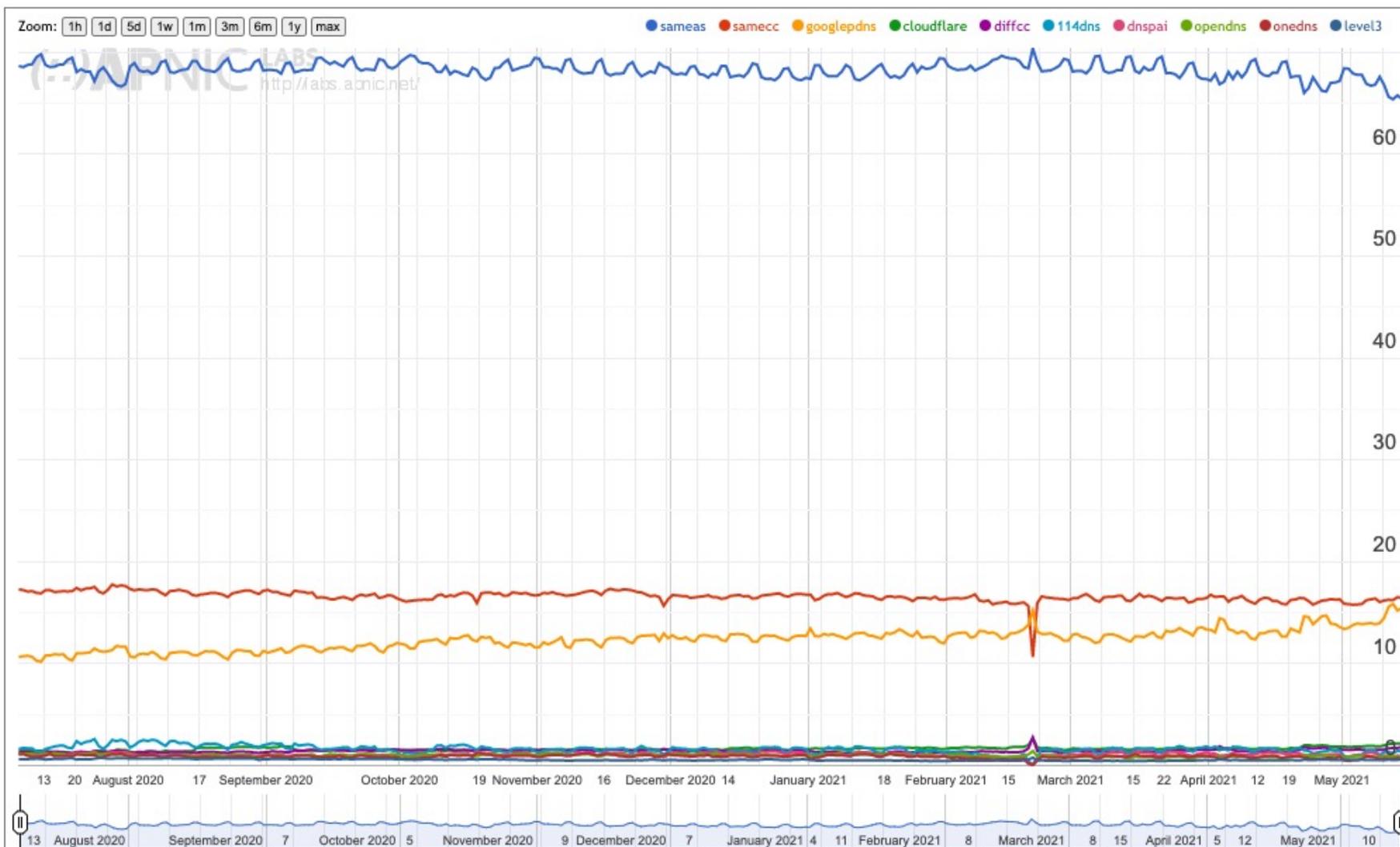
Is this what we mean by “centralisation”?

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

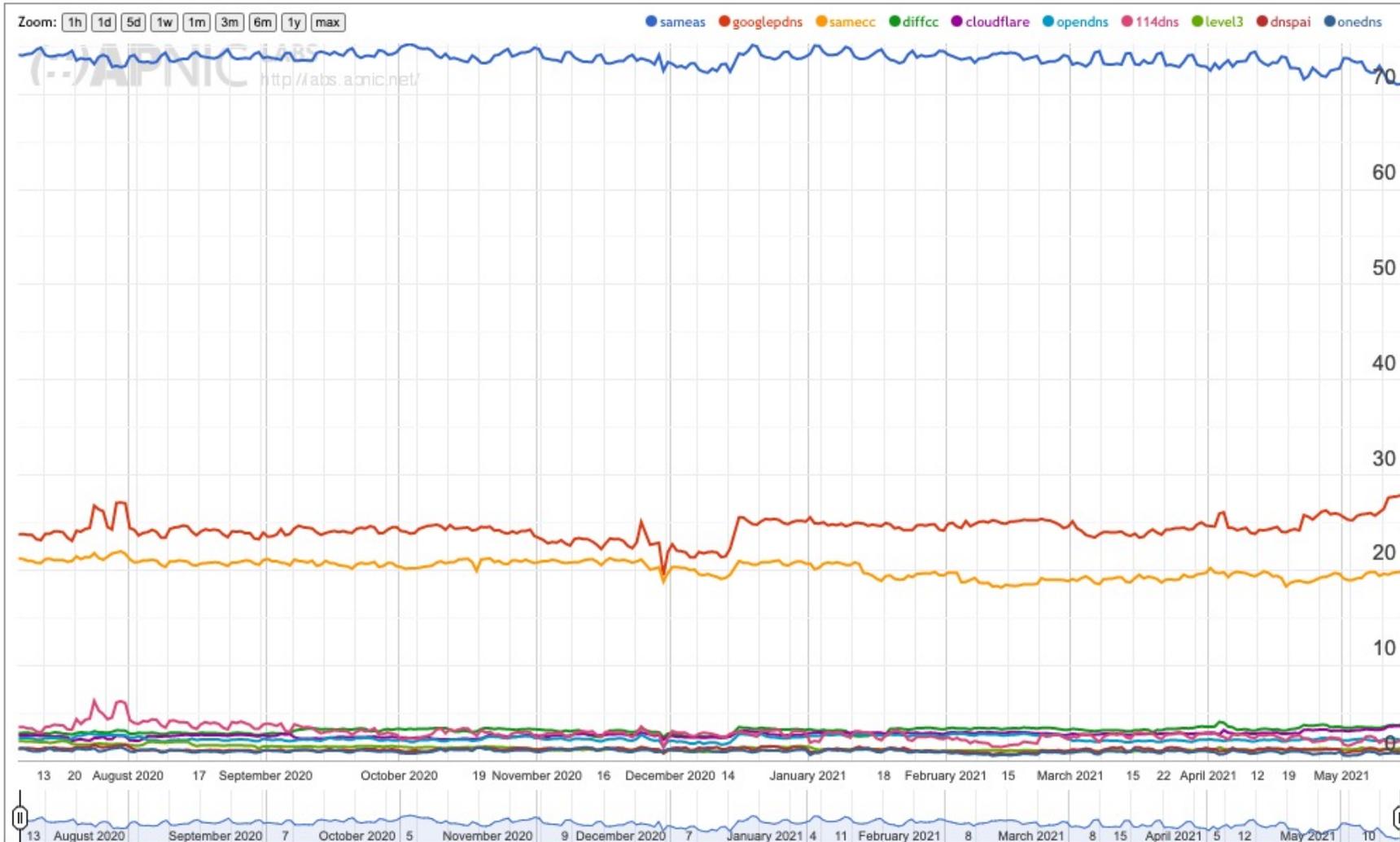


70% of users use a resolver located in the same AS as the user (ISP resolver)

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

15% of users use the Google open resolver (8.8.8.8)

All Resolver Use (SERVFAIL)



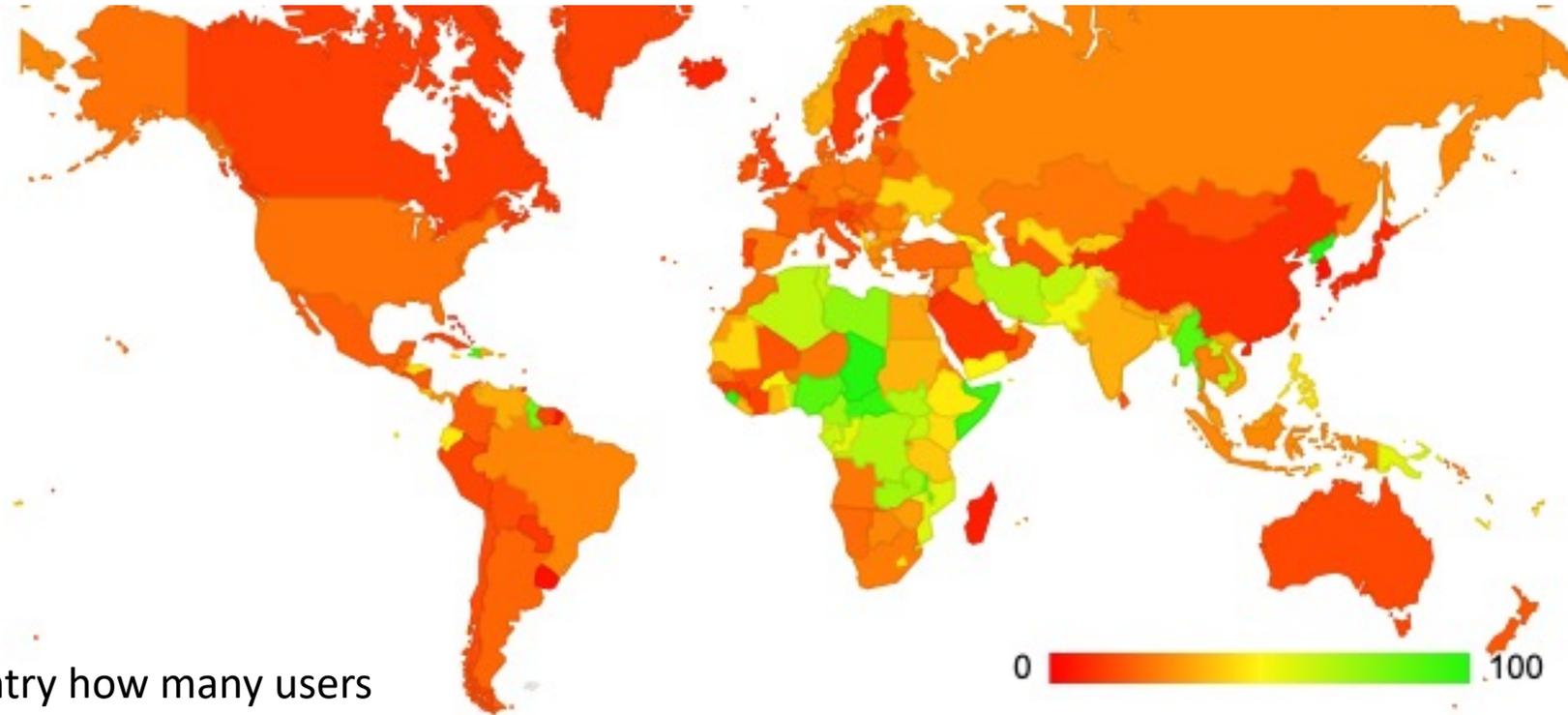
70% → 72% for same ISP

15% → 29% for Google use

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

Google DNS

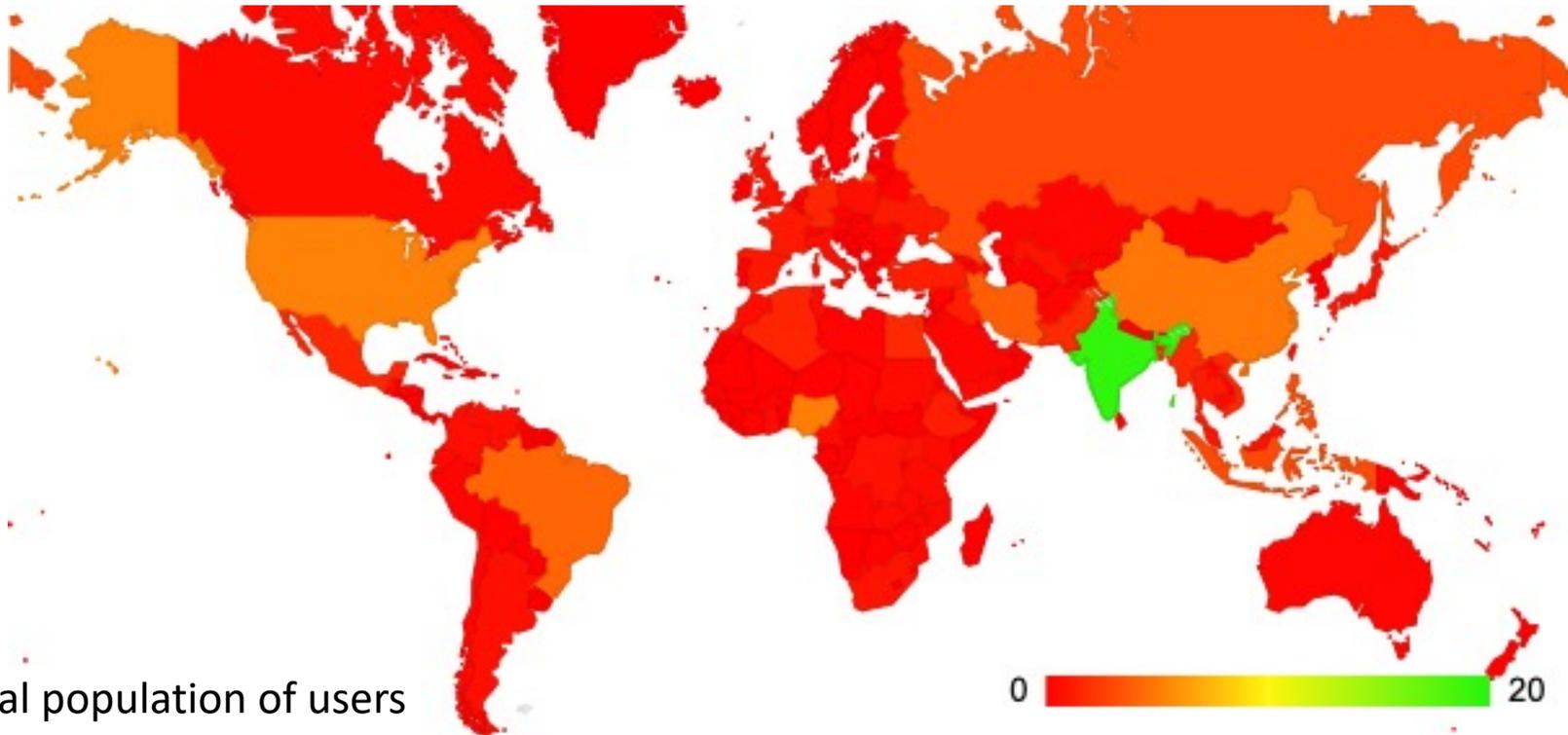
Use of Google Service per CC



Within each country how many users
In that country use Google's resolver?

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

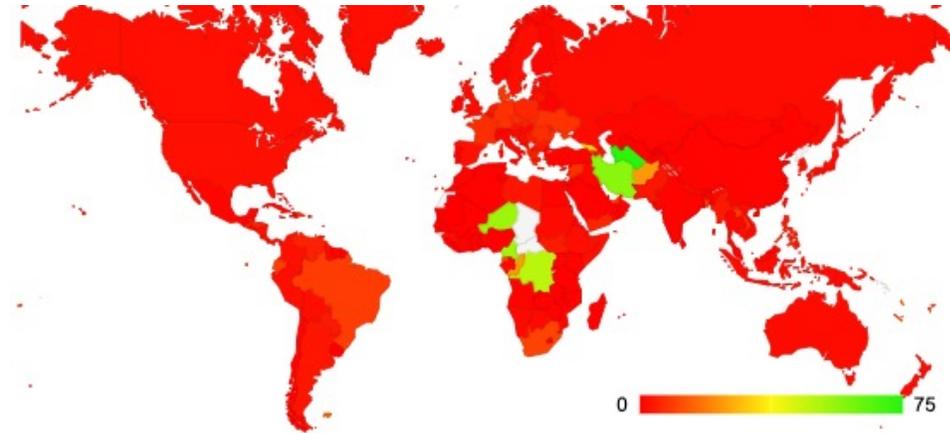
- Google DNS use appears to be equally split between first use (15% of users) and backup resolvers (a further 14% of users)
- Within each economy Google DNS is heavily used in some African economies, and central and southern Asian economies
- The largest pool of Google DNS users are located in India (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

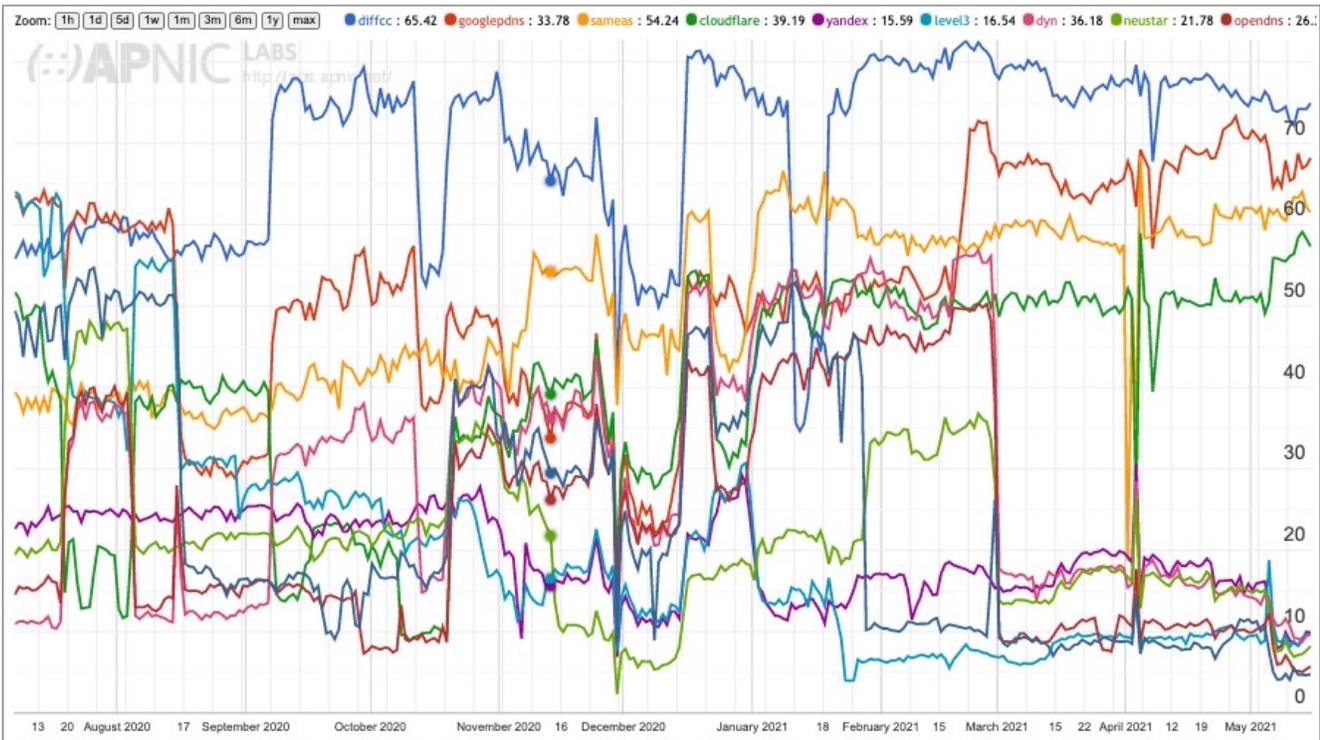


Where is Cloudflare used?



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

Iran

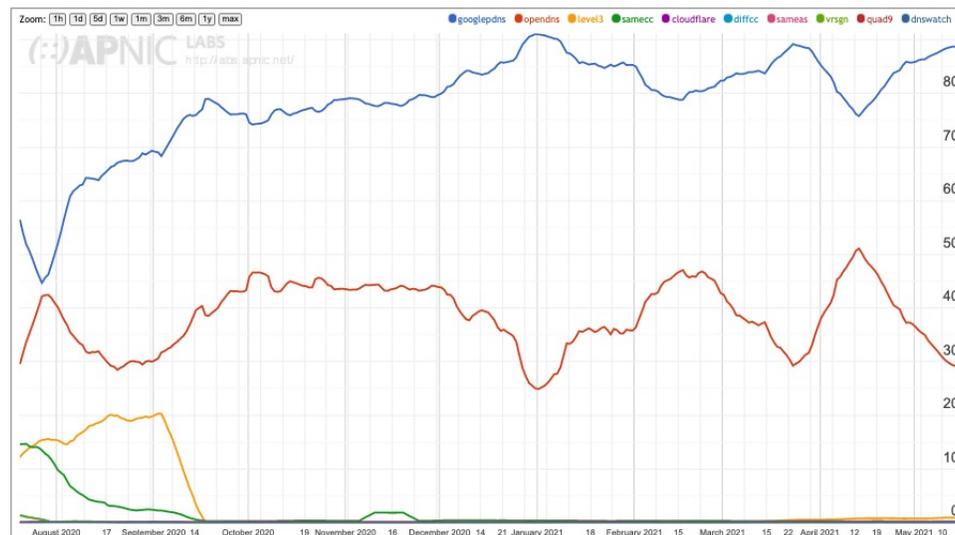


A major ISP in IRAN, MCCI, distributes its queries across Google, Cloudflare, Yandex, Neustar, OpenDNS, Quad9 and others

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

DNS Resolver Use Measurements for AS131429: MOBIFONE-AS-VN MOBIFONE (Vietnam (VN))



Google DNS at 86%

Open DNS at 27%

Resolver Centrality?

- Its not a “small number” of open resolvers
 - It’s 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

Is this a "problem"?

- It this an emerging distortion of the market that puts excessive market control in the hands of a small set of providers?
 - No, not so far
- 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

Thanks!

