The resolvers we use

Geoff Huston
João Damas

APNIC Labs

As far as we are aware nobody actually said this:

> "In the EU the DNS resolvers we use are dominated by a service operated by an American behemoth that is completely unaccountable to EU users and operates entirely outside of our regulatory framework.
>
> And they are operating an essential piece of Internet infrastructure that many EU Internet users rely on.
>
> We feel that this is unacceptable situation for the EU."

But when we look at the possible political motivations behind the DNS4EU initiative it sure feels like this is a credible explanation!

“The resolvers we use”

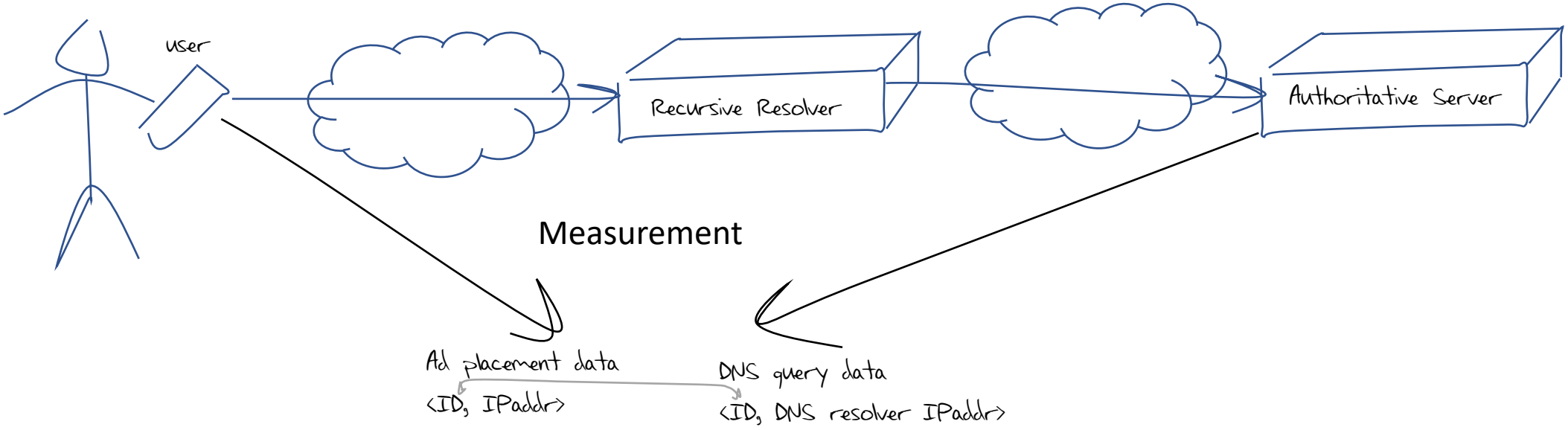
What does this ***mean?***

If you wanted to conduct a measurement experiment to calculate the “user share” of DNS resolvers, then what exactly should be measured?

And once you have that measurement, then what is the data telling us?

Our Initial thoughts

Use Ads to send each user a unique DNS name to resolve. We look at the authoritative server and collect the IP address of the resolvers asking the authoritative server, and use Ad data to match this query to an end user IP address



More Initial Thoughts

- The Authoritative Server always answers queries immediately with the A / AAAA records as requested
- The data is unsigned and the responses fit comfortably within 512 octets of DNS payload
- We try to minimise timeouts and requeries by steering users to a DNS Authoritative server that is (roughly) on the same continent as the user

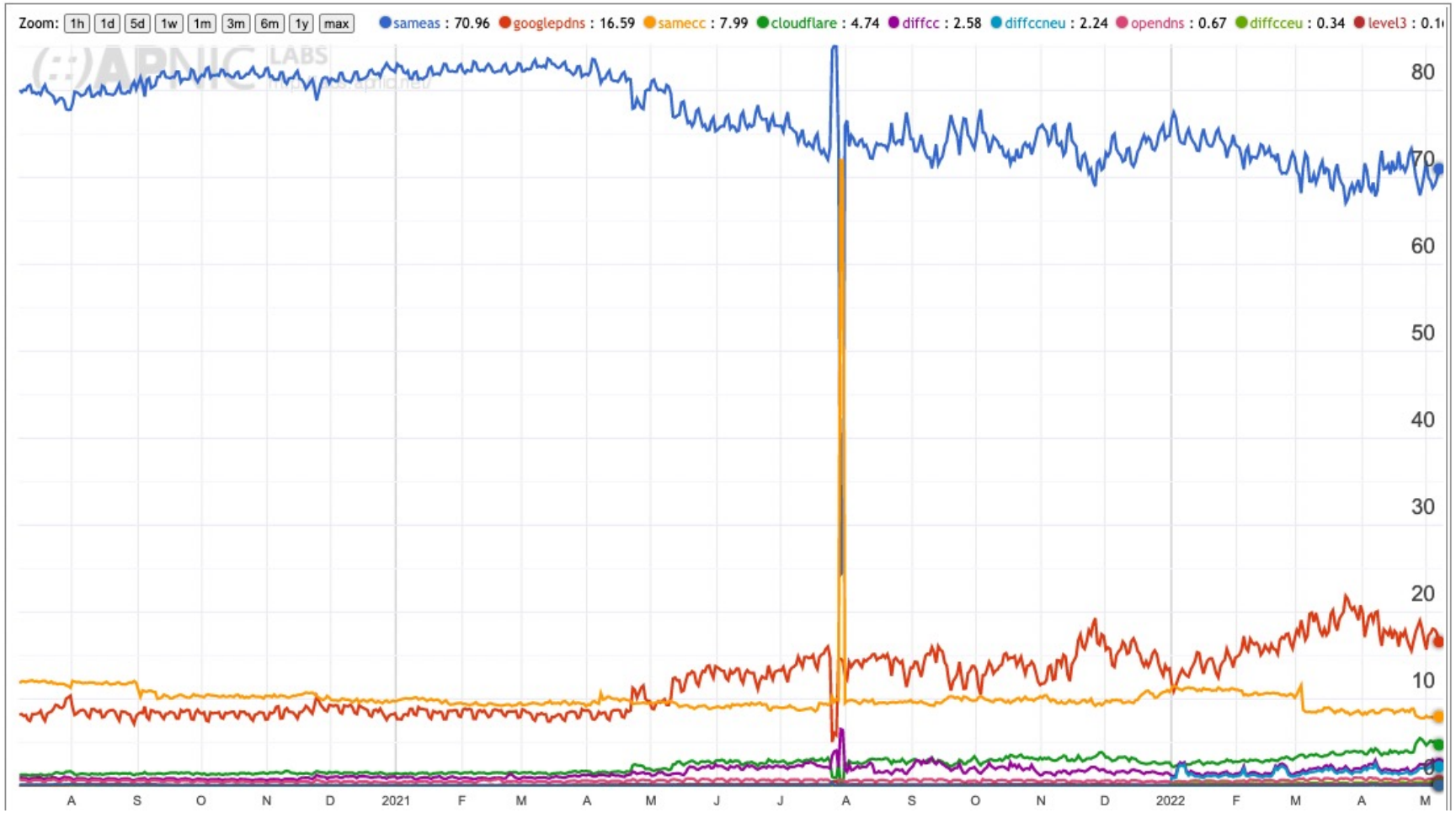
More Initial thoughts

- We need to map the resolver “helper” addresses to a resolver service
  - Which back-end DNS addresses are used by each open resolver?
  - RIPE Atlas helped here for those cases where the open resolver operator does not publish this information
- We map resolvers into a number of categories based on the resolver’s IP address:
  - Resolver is in the same AS as the end user
  - It’s a known Open DNS resolver
  - Resolver is geo-located to the same CC as the end user
  - Resolver is geo-located to a different CC from the end user

DNS resolver use in EU region

Resolvers seen from a single initial query

Same AS (ISP) – 70%
Google – 16%
Same CC – 8%
Cloudflare – 5%

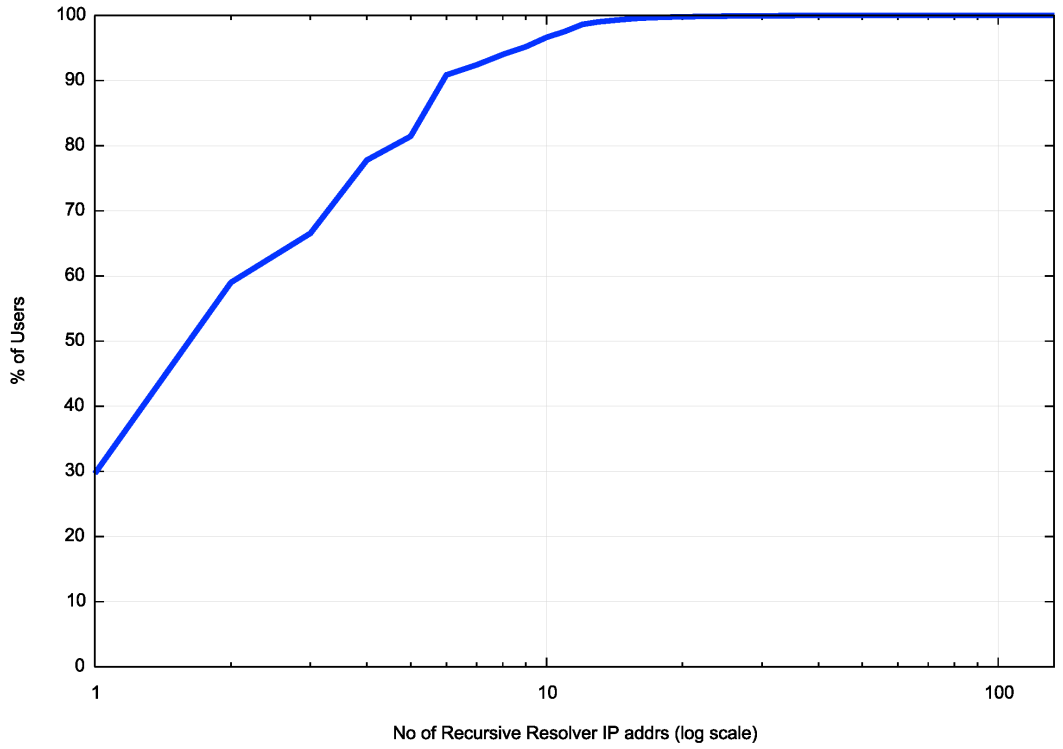


https://stats.labs.apnic.net/rvrs/XE?hc=XE&hl=1&hs=0&ht=10&w=1&t=10&s=2

However, its not as simple as it may suggest

- We observe that this single initial query generates 1 or more queries from a single recursive resolver IP address just 30% of the time
- 2 or more different resolvers are queried in 60% of cases
- Most of the time (90% of cases) these multiple resolver IP addresses are all in the same AS

Cumulative Distribution of number of resolver IP addresses seen to query for a unique DNS name



Multiple resolvers “see” individual stub queries

- We see an average of 3.23 distinct resolver IP addresses at the authoritative server for each queried domain name within the first 15 seconds
- What should we do with these “extra” DNS queries?
- In this case we just add them to the count
- Could we do better?

What are we measuring here?

- Seems that this experiment is not clear about what is being measured
- So we thought that maybe we really wanted to know ***all*** the resolvers who ***might*** see your query
- But to flush out all of these resolvers we need to adjust this experiment

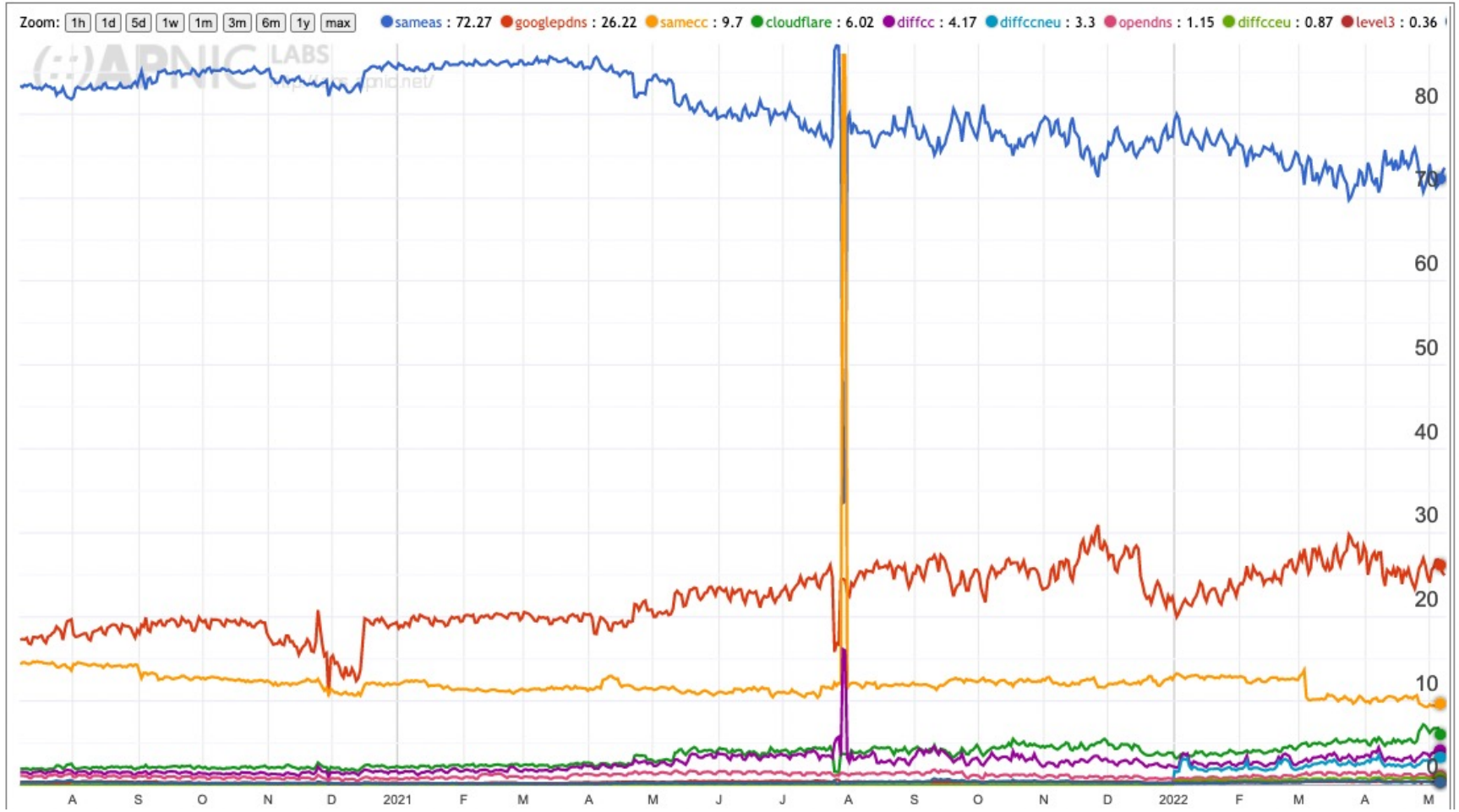
Second Thoughts

- Get the authoritative server to return SERVFAIL all the time
- This way the stub resolver is likely to cycle through all the locally configured recursive resolvers to find a non-SERVFAIL DNS response

DNS in EU

All Resolvers seen from SERVFAIL

Same AS (ISP) – 72%
Google – 26%
Same CC – 9%
Cloudflare – 6%



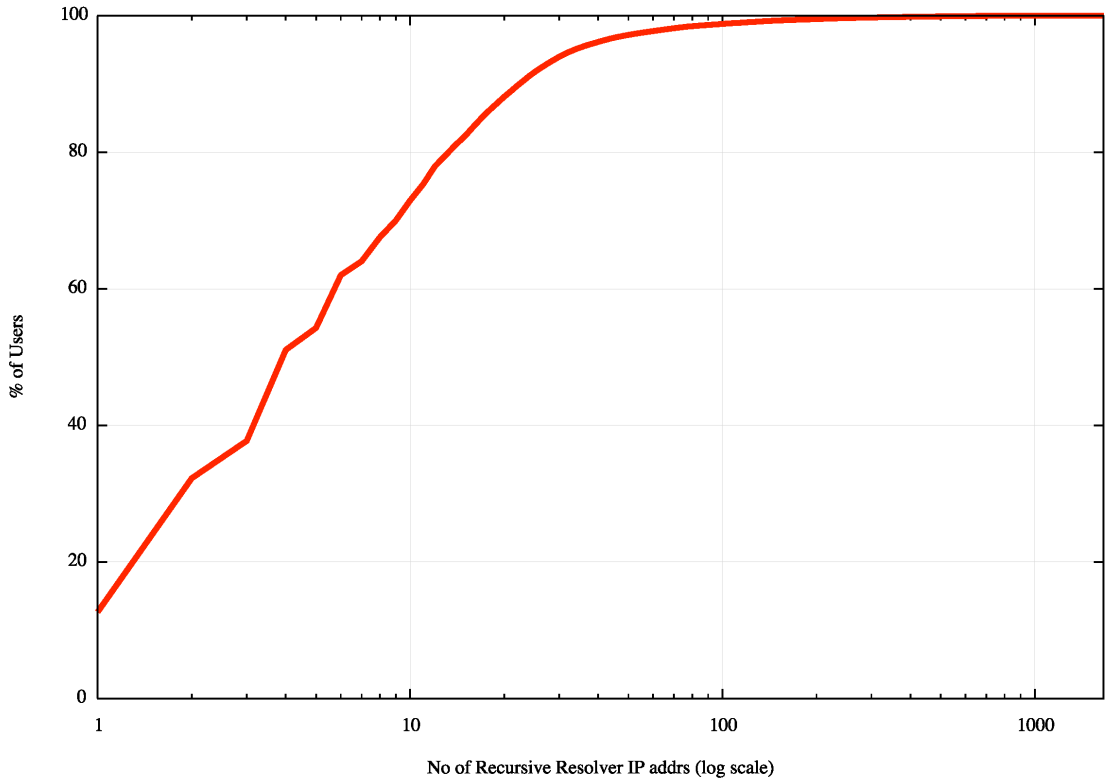
2021 2022

https://stats.labs.apnic.net/rvrs/XE?hc=XE&hl=1&hs=2&ht=10&w=1&t=10&s=1

How many resolvers see the query now?

- We observe that this single initial query generates 1 or more queries from a single recursive resolver IP address just 12% of the time
- 2 or more different resolvers are queried in 30% of cases
- Most of the time (75% of cases) these multiple resolver IP addresses are all in the same AS

Cumulative Distribution of number of resolver IP addresses seen to query for a unique DNS name when the response is SERVFAIL



Are we there yet?

- No, not really

- Perhaps it is also useful to understand ***which resolver provides the response that the user will use***

Third Pass

- Single query – same as Pass 1
- But only record the first query at the auth server for each unique ID
  - We assume that the first recursive resolver to ask the auth server is the first to provide a response to the stub resolver

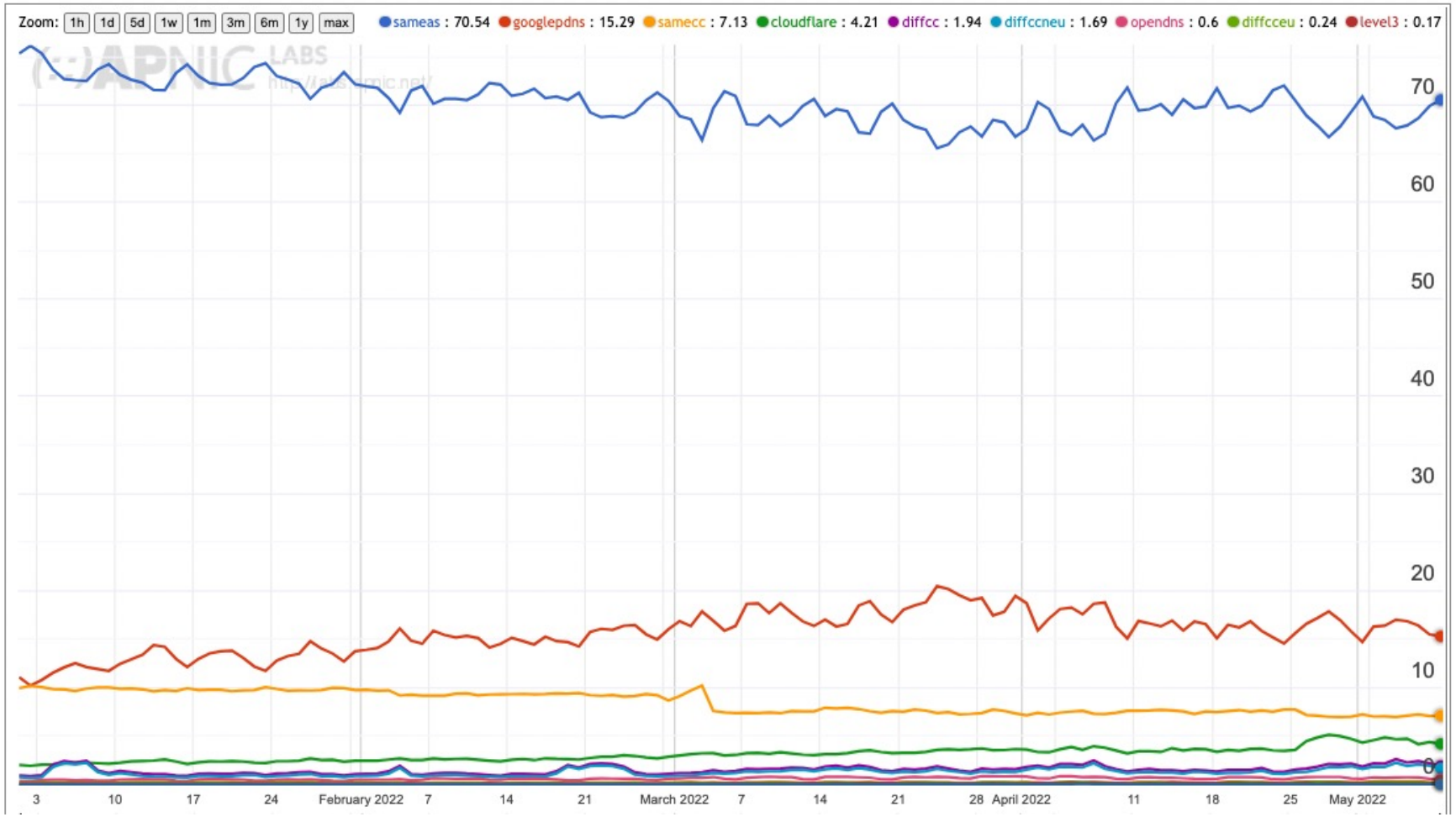
- How does this change the measurements?

DNS in EU

First Responder

Resolver

Same AS (ISP) – 71%
Google – 15%
Same CC – 7%
Cloudflare – 4%

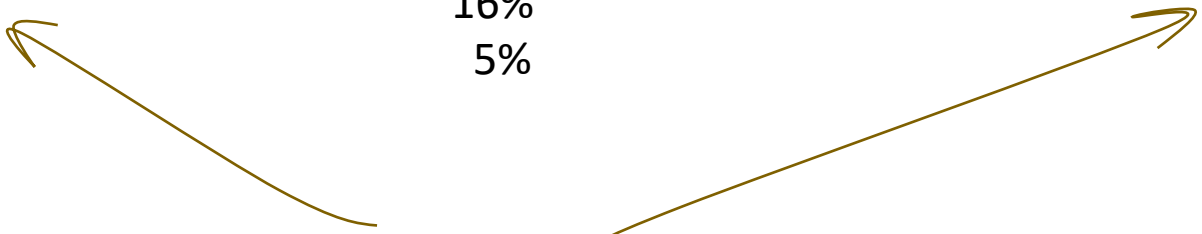


2022

https://stats.labs.apnic.net/rvrs/XE?hc=XE&hl=1&hs=1&ht=10&w=1&t=10&s=0

Where are we now?

	All configured Resolvers	Resolvers in the initial query set	First Responder
Same AS	72%	70%	71%
Different AS	9%	8%	7%
Google	26%	16%	15%
Cloudflare	6%	5%	4%



What's Google's market share of DNS resolution in EU? 26%? Or 15%

Who is “*we*”?

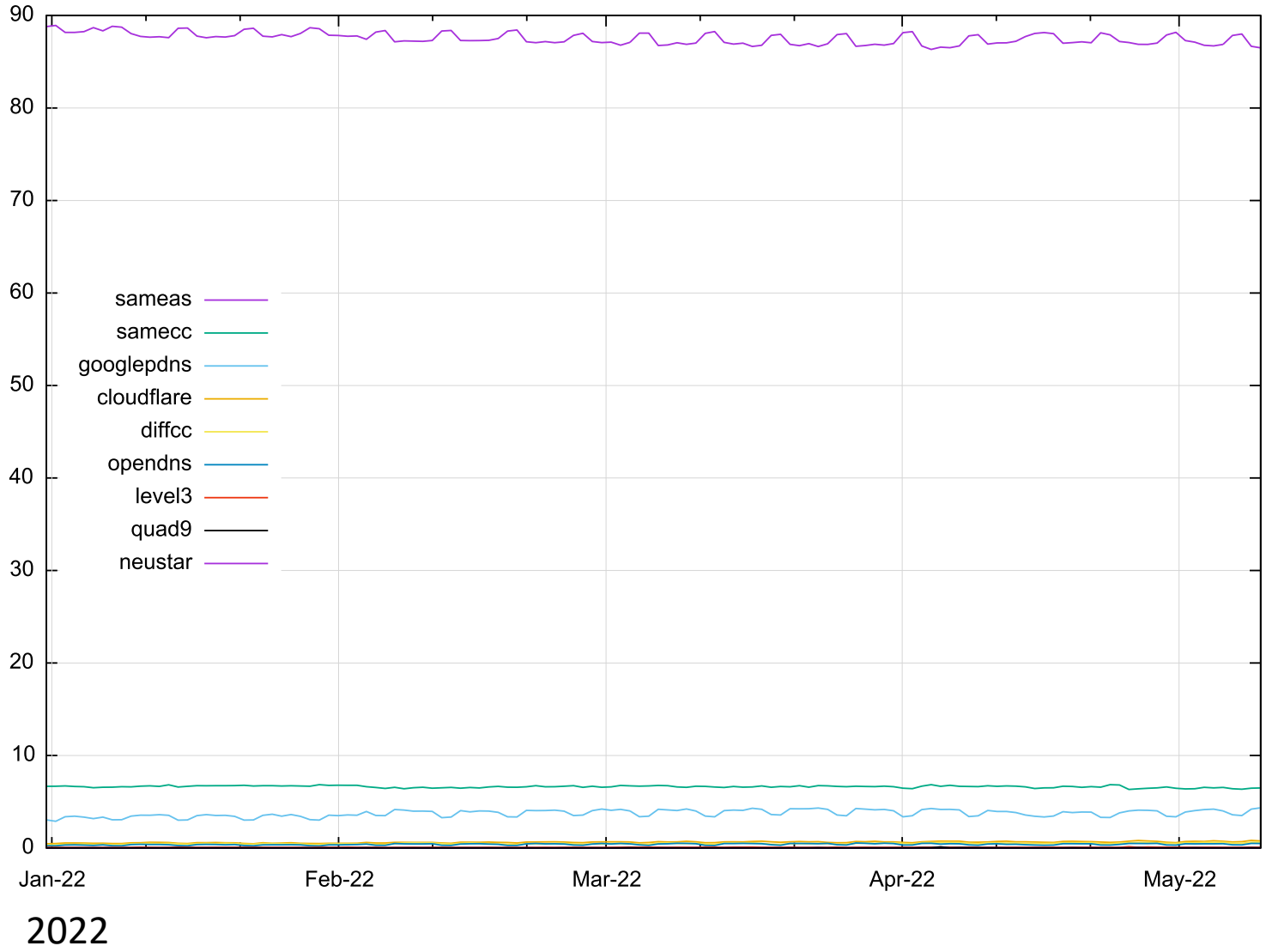
- Ads are directed to all kinds of end points
  - There are various forms of Enterprise and B2B networks that are measured as well a consumer networks. These enterprise endpoints have a different DNS profile as compared to consumer retail services, as they tend to be more intensive users of open DNS resolver services than we see in consumer networks.
- What if the “we” we are referring to in this measurement are individual consumers in the EU region?
- What happens if we filter the ad data for Europe to look only at what we believe are consumer retail ISPs?

DNS in EU

First Responder,

Consumer ISPs

Same AS (ISP) – 87%
Same CC – 6%
Google – 4%
Cloudflare – 1%



Where are we now?

	All Resolvers	Initial query set	First Responder	Consumer ISPs
Same AS	72%	70%	71%	87%
Different AS	9%	8%	7%	6%
Google	26%	16%	15%	4%
Cloudflare	6%	5%	4%	1%

What does all this mean?

- Most users of retail ISPs use their ISP-provided resolver (same AS)
- Some additional users in those ISPs use resolvers in different networks (ASN) but same country
  - This can be due to the ISP having services across different ASNs
- The picture for enterprise networks is slightly different.

What we didn’t measure

- How many different DNS resolver service operators process DNS queries?
  - Many ISPs are contracting their DNS resolution functions to a third party

    For example, Nominum is a significant player in this space
  - This is subtly different question to just a count of the open DNS resolvers, as we would need to look inside each ISP to identify how they provision their DNS resolution service
- Who on-sells query log data or just passes the query logs onto others?
  - Again, this is not something that we can measure directly
  - Previously we’ve looked at the persistence of DNS queries over time

More hard-to-measure questions

- Who retains DNS queries?
  - Obviously this is not easy to tell
  - But we can look at the distribution of the time gap between the time original query and the recycled query

Re-Query distribution

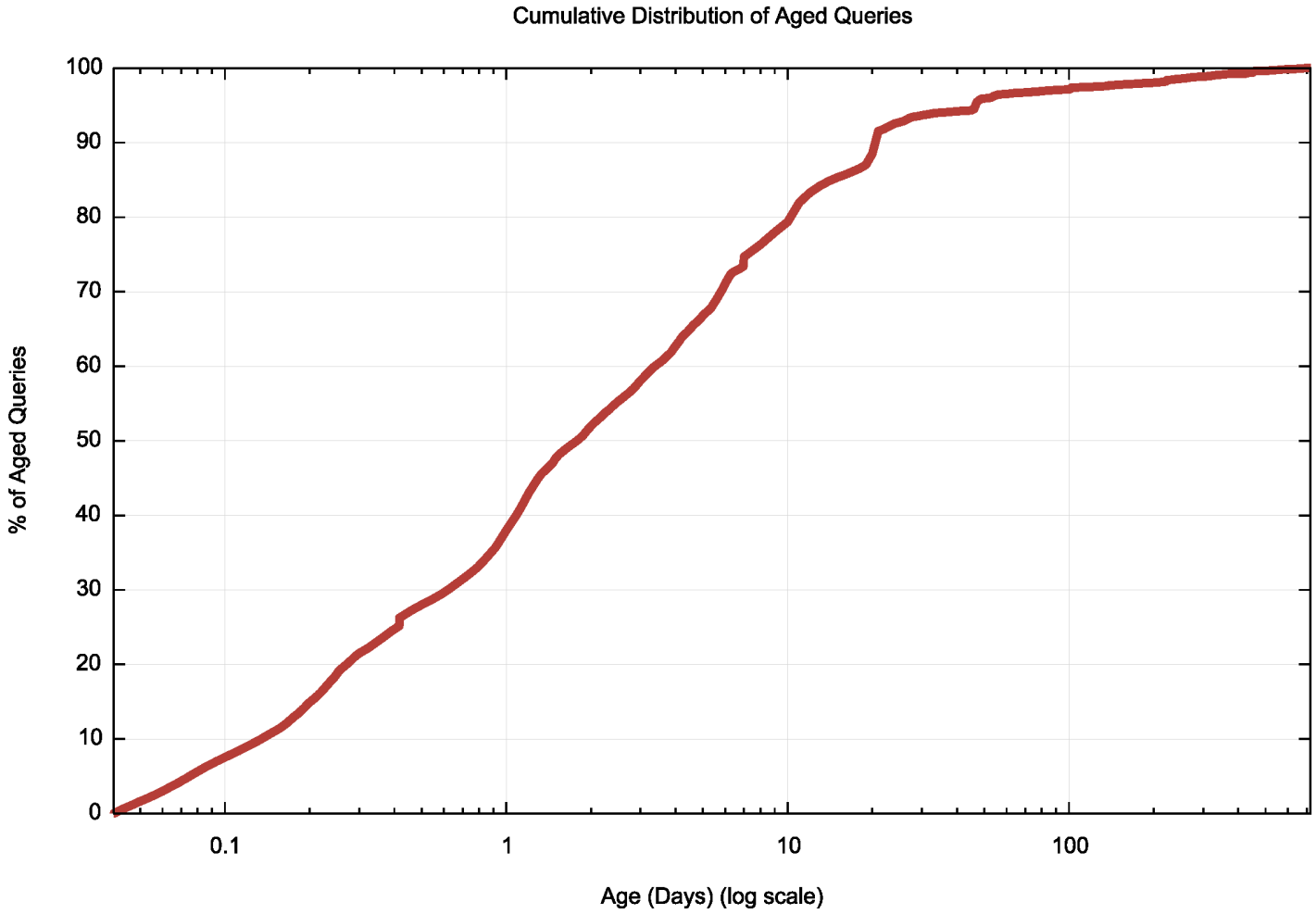
In a 24 hour period we received 612M DNS queries, or which 23M were for names which were “older” than 1 hour

50% of queries for aged names were between 1 and 12 days old

The DNS names are single use names

The DNS responses have a TTL of 60 seconds

Why do we get a 4% query retention and re-query rate?



So, is there really a problem for the EU?

“Yes and No” or perhaps “No and Yes”

- The majority of end users use the ISP-provided default in the EU region
- However, there is an undeniable issue about the emergence of aspects of centrality in DNS resolution that should be a concern for us all

Questions?