Client-IP EDNS Option Concerns

RIPE 67, Athens

Florian Streibelt

<florian@inet.tu-berlin.de>

TU-Berlin, Germany - FG INET

www.inet.tu-berlin.de

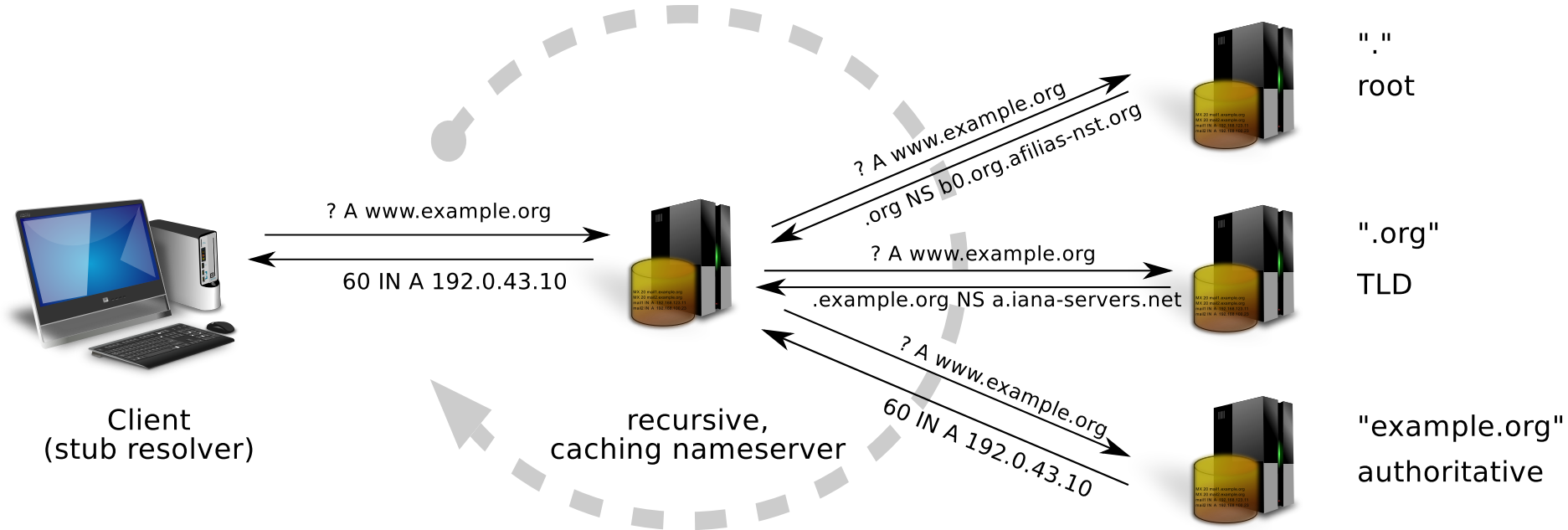
October 16th 2013

Preliminary results, full results at IMC 2013 (see last slide)

Disclaimer

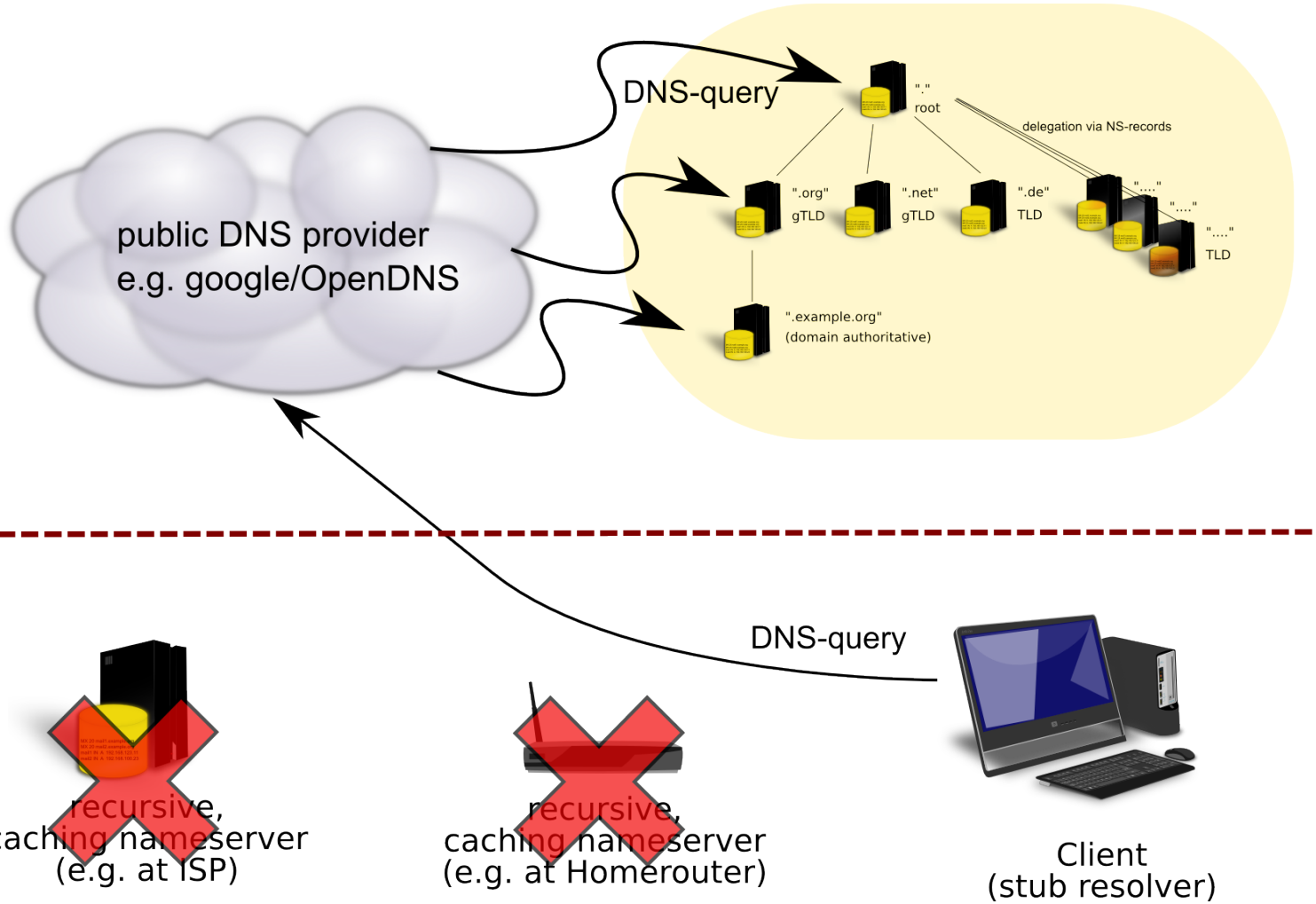
- This talk is not about repeating all the arguments from the IETF dns-ext WG.
- I don't know if this extension is a good solution or not, but it seems to solve a problem for some people and my hope is that the measurements we did may help in understanding some additional side effects.

Textbook DNS-Lookup



- Stub resolver on the client asks a recurser (e.g., at the ISP)
- Recurser follows the delegation

Non-ISP (aka 'public') DNS usage increases



Otto et al. [2]: usage at 8.6% in December 2011

Challenge for CDNs/CPs

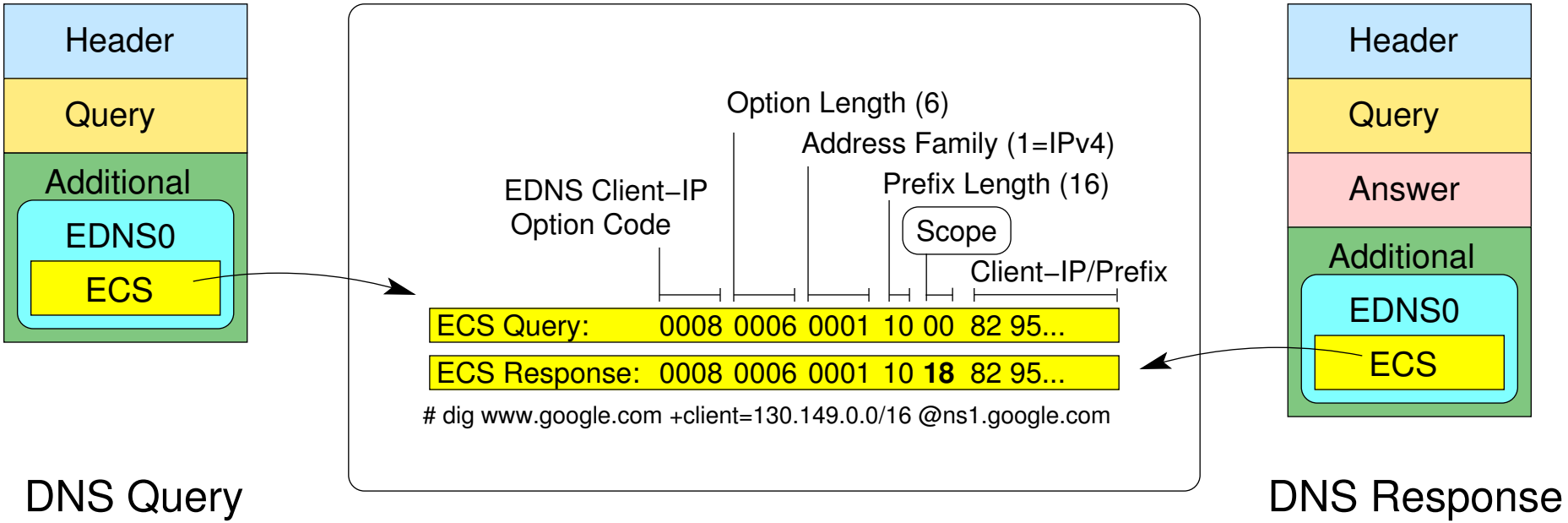
- Non-ISP Resolvers are gaining momentum
- CDNs often rely heavily on 'dns-tricks' for client location
- Using the DNS request origin for client-location now leads to (more) wrong results
- Mis-location of clients gives end-users bad performance
- Some workarounds exist but don't scale well/are inacurate - e.g. check against known list of Google NS IPs and their geolocation[1]

[1] https://developers.google.com/speed/public-dns/faq#locations

Introducing: Client IP information in EDNS (ECS)

- Proposal by Google, OpenDNS and others:
  `http://afasterinternet.com/`
- EDNS0 extension to transport Client Subnet information:
  `http://tools.ietf.org/html/`
  `draft-vandergaast-edns-client-subnet-02`
- Recurser adds client IP-information (network prefix) to the query directed at the authoritative NS
- Performance gain can be observed [2].

Protocol: Client IP information in EDNS (ECS)



- DNS query contains additional section
- EDNS0 is used to transport Client Subnet Information
- Answer differs only in one byte
- The scope returned allows for caching the answer (q-tuple!)

$\Rightarrow$ We can impose every 'location' using arbitrary Client Subnet information

How to enable ECS?

- Primary nameservers must be ECS enabled
  (Supported by PowerDNS: yes, Bind: no)
- If there are e.g., loadbalancers (sic!) in front: these too
- Nameservers need to be whitelisted (manually) by OpenDNS/Google, etc.
- Note: We find that roughly 13% of the top 1 million domains (Alexa) may be already ECS enabled.

Measurements

- Single vantage point[2] is sufficient to use *arbitrary* Client IP/prefix
- We use all network prefixes collected by RIPE RIS (sanity check using Routeviews)
- Measurements done for: Google, YouTube, MySqueezebox, Edgecast, CacheFly, TorrentFreak
- Following is a subset of our experiments, using Google
- In progess: Measurements with traces from an ECS-enabled CDN

[2] we checked from four different locations

Looking at the A-Records of Google

- Resolving www.google.com via `ns1.google.com`
- Using all network prefixes from RIPE RIS as client subnets
- Different synchronized vantage points (plausibility check)

Date (RIPE)	IPs	Sub nets	ASes	Countries
2013-03-26	6340	329	166	47
2013-03-30	6495	332	167	47
2013-04-13	6821	331	167	46
2013-04-21	7162	346	169	46
2013-05-16	9762	485	287	55
2013-05-26	9465	471	281	52
2013-06-18	14418	703	454	91
2013-07-13	21321	1040	714	91
2013-08-08	21862	1083	761	123

see also:

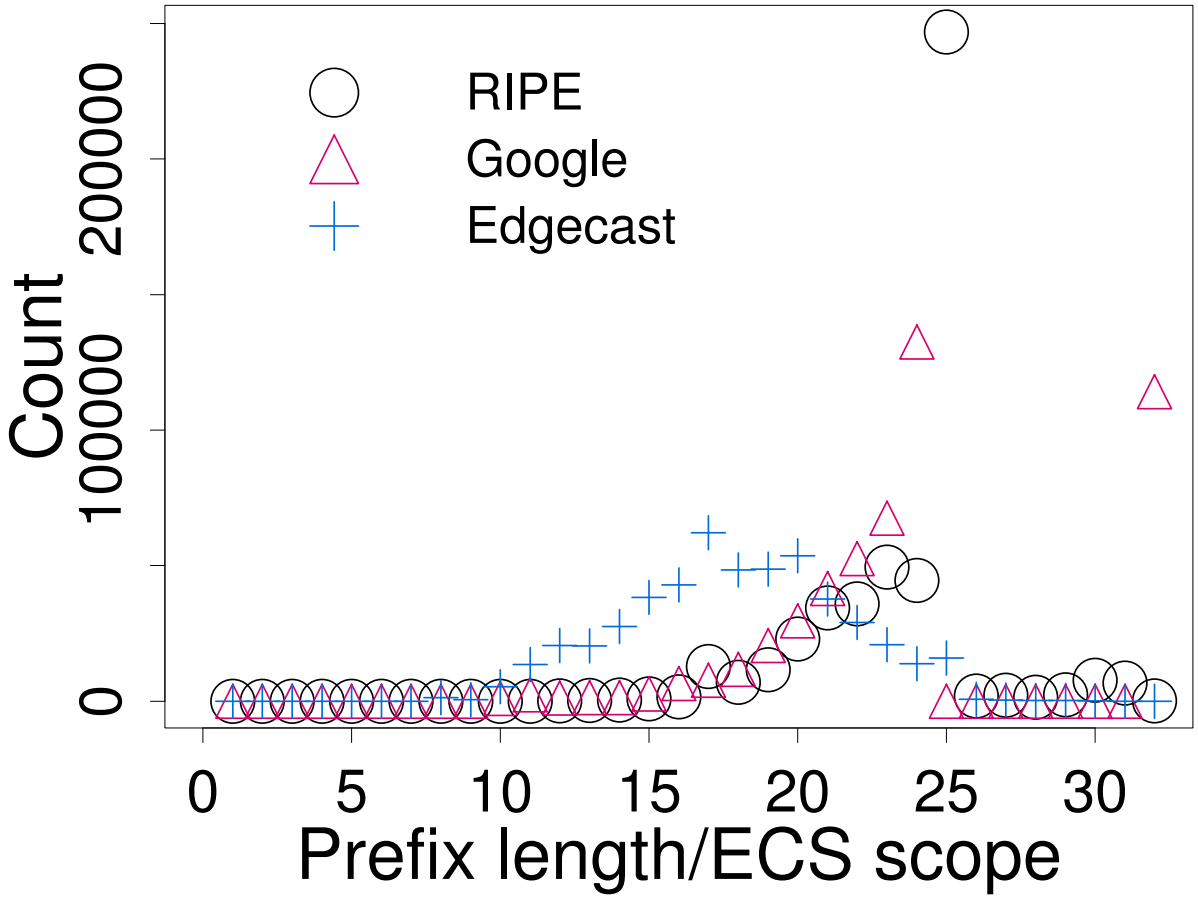
Calder et al.: Mapping the Expansion of Google's Serving Infrastructure [1]

Looking at the A-Records of Google

Preliminary results from combined experiments:

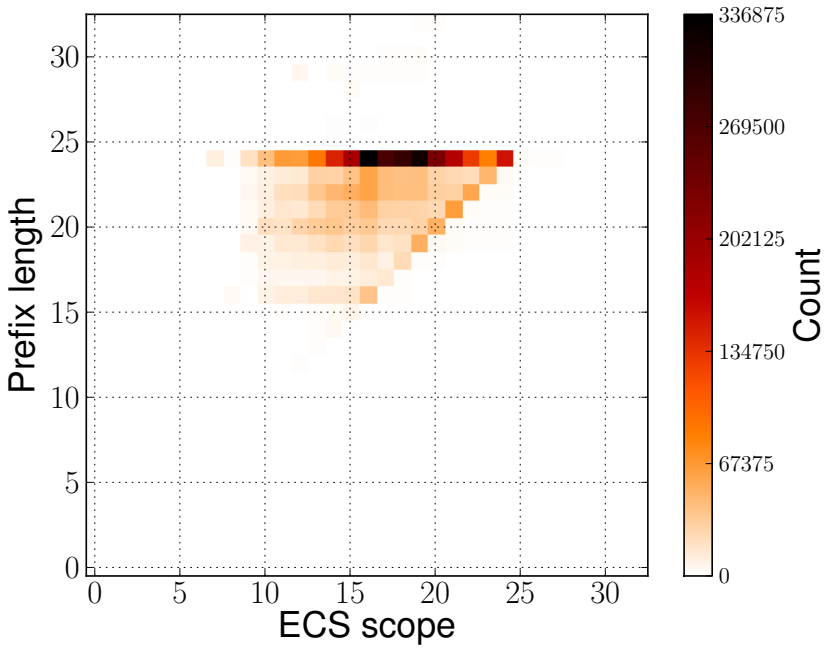
- We see GGC (Google Global Cache edge servers) in various ISP networks
- These ISPs are not allowed to advertise the GGC, but we are
- Huge increase in the footprint can be observed, also for YouTube
- Comparing results from different vantage points we observe redirection of clients and prefixes, probably due to load balancing the GGCs
- We see that most of the time clients indeed are served from caches in their respective AS
- We see large overlap in the returned A records in the results from the different vantage points, both for Google and YouTube

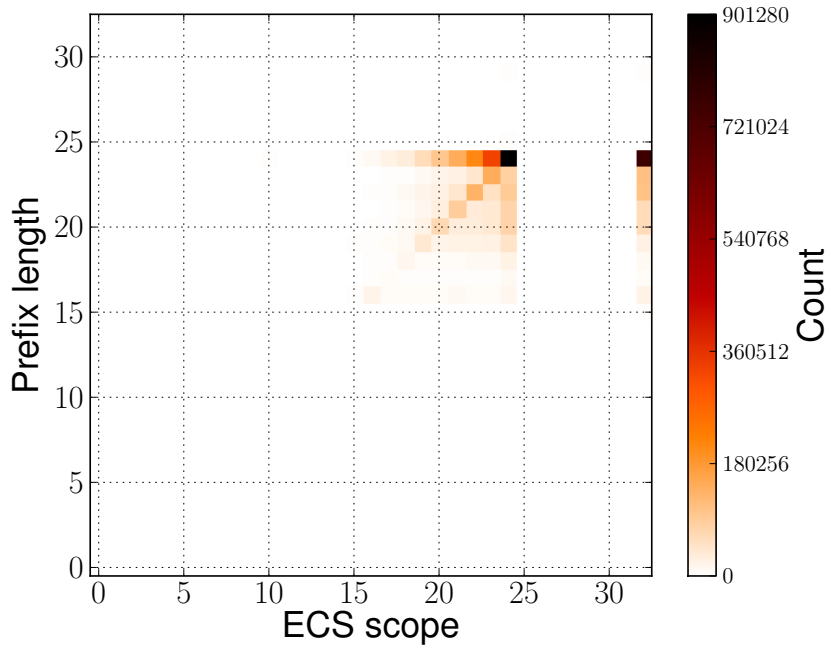
RIPE RIS prefix length vs. ECS-scopes



Prefix length and scope distribution do not match and differ between adopters, also note the /32s!

Comparing Google and Edgecast Scopes





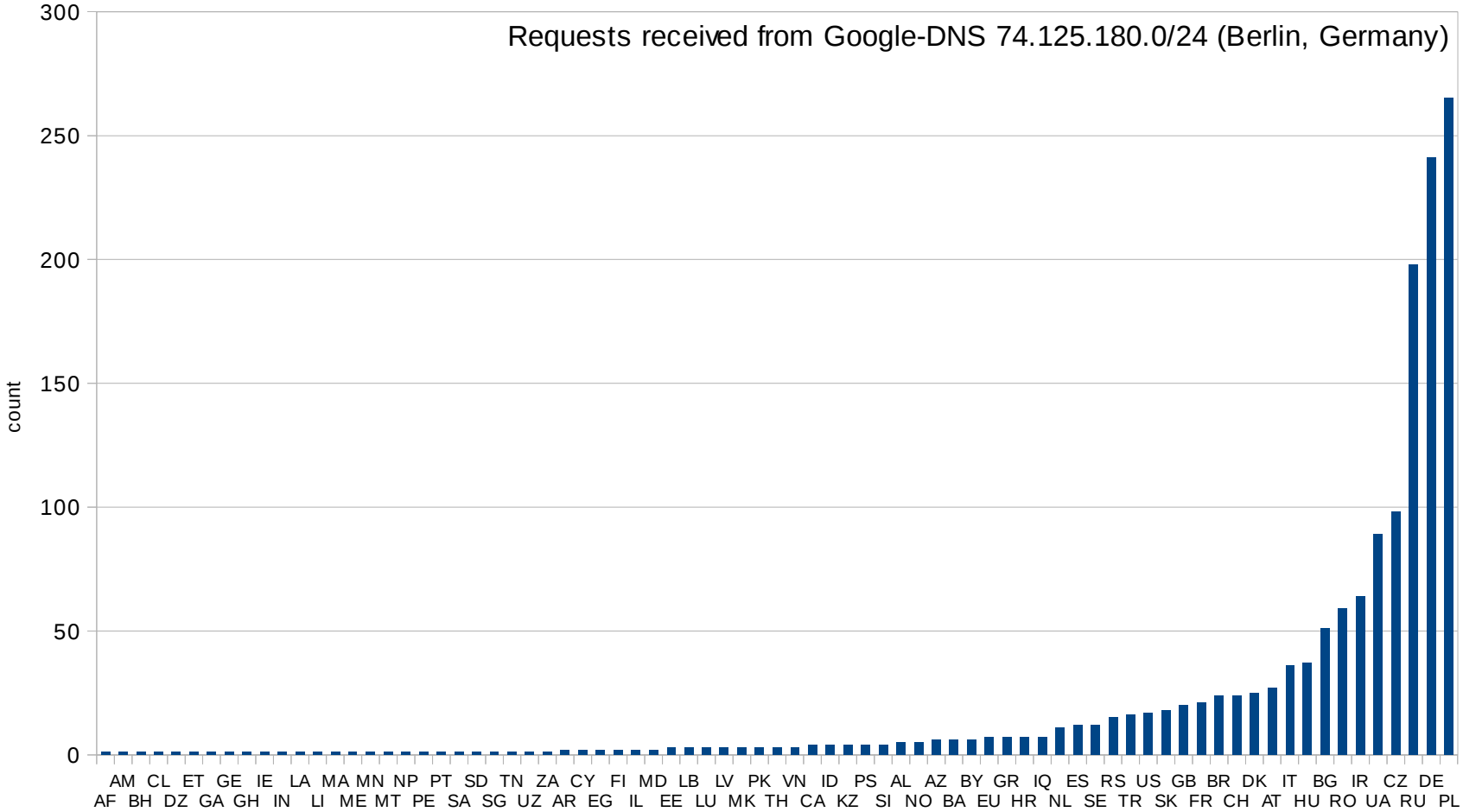
Edgecast (left) aggregates while Google (right) returns more specific scopes.

Looking at the CDN side

- We have access to all dns-requests sent to all authoritative nameservers of a CDN
- For Google we receive queries from the known backend subnets[3]
- We can map the client prefixes to these locations and infer data from the Google location DB
- There will be future work with this dataset...

[3] again see:
`https://developers.google.com/speed/public-dns/faq#locations`

Client subnet: country to DNS-Server mapping



Conclusion

- Enabling ECS gives better performance for clients
- This comes with a tradeoff for DNS providers and CDNs: it also reveals internal information
- It enables researchers (and competitors) to investigate e.g. global footprint, growth-ate, user-to-server mapping
- Thus it reveals more information than desired (server and service distribution)
- This is in fact an experiment running on the public Internet and might not be as 'harmless' as it seemed
- Future Adopters and the community should be aware

Bibliography I

[1] Matt Calder, Xun Fan, Zi Hu, Ethan Katz-Bassett, John Heidemann, and Ramesh Govindan.
Mapping the expansion of Google's serving infrastructure.
Technical Report TR 13-935, University of Southern California Computer Science Department, June 2013.

[2] John S. Otto, Mario A. Sánchez, John P. Rula, and Fabián E. Bustamante.
Content delivery and the natural evolution of dns: remote dns trends, performance issues and alternative solutions.
In *Proceedings of the 2012 ACM conference on Internet measurement conference*, IMC '12, pages 523–536, New York, NY, USA, 2012. ACM.

Contact:
Florian Streibelt <florian@inet.tu-berlin.de>

Related publication:
Unintended Consequences: Exploring EDNS-Client-Subnet Adopters in your Free Time
Internet Measurement Conference, October 2013
http://conferences.sigcomm.org/imc/2013/
Authors:
Florian Streibelt, Jan Böttger, Nikolaos Chatzis, Georgios Smaragdakis, Anja Feldmann

The paper, software and raw data will be published in November 2013.

Image sources:
own work and http://openclipart.org/