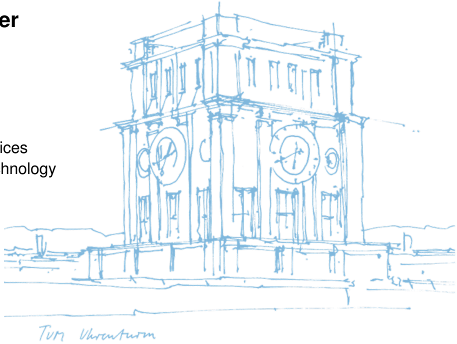
Distributed ECS Measurement with Ark

Patrick Sattler, Mattijs Jonker

Wednesday 12th February, 2025

Chair of Network Architectures and Services
School of Computation, Information, and Technology
Technical University of Munich



Motivation

- Why look at ECS at authoritative nameservers?
  - Uncovers infrastructure
  - Evaluate load balancing properties
  - Distributed platforms do not cover enough client networks to account for all possible responses
  - Last big ECS studies on authoritative nameservers from 2013

Motivation

- Why look at ECS at authoritative nameservers?
  - Uncovers infrastructure
  - Evaluate load balancing properties
  - Distributed platforms do not cover enough client networks to account for all possible responses
  - Last big ECS studies on authoritative nameservers from 2013
- Our contributions:
  - We developed a response-aware ECS scanner (ECSplorer)
  - Analyzed the current ECS landscape
  - Available on arXiv and submitted to *CoNEXT*
  - Hackaton to implement it using Ark for distributed scanning

EDNS Client Subnet (ECS)

- Defined in RFC7871 with EDNS OPTION-CODE 8
- Resolver forwards the client IP address to the authoritative name server
- Sends:
  - IP address family
  - IP address
  - Source prefix length (number of relevant bits in the IP address)
  - Scope prefix length (number of bits the response covers)

Our Approach

- Send DNS queries including ECS information directly to authoritative nameservers
- Use routed address space to seed *client subnets*
    - Unrouted and special prefixes get only a limited number of queries

Our Approach

- Send DNS queries including ECS information directly to authoritative nameservers
- Use routed address space to seed *client subnets*
    - Unrouted and special prefixes get only a limited number of queries
- Skip address space covered by responses (scope prefix length $<$ source prefix length)
    - Reduces the queries up to 97 % depending on the deployed ECS response strategy
- Use patricia trie and prefix length based query limits

Our Approach

- Send DNS queries including ECS information directly to authoritative nameservers
- Use routed address space to seed *client subnets*
  - Unrouted and special prefixes get only a limited number of queries
- Skip address space covered by responses (scope prefix length $<$ source prefix length)
  - Reduces the queries up to 97 % depending on the deployed ECS response strategy
- Use patricia trie and prefix length based query limits
- The first scanner to support IPv6 probing
- Code is public `github.com/tumi8/ecsplorer`

Results

- Scanned a collection of top list domains (3.2 M) with a limited number of queries
  - 53 % of nameservers have ECS enabled (authoritative for 1.2M domains on top lists)
  - Only 15 % also return multiple RRsets for four different ECS subnet queries.

Results

- Scanned a collection of top list domains (3.2 M) with a limited number of queries
  - 53 % of nameservers have ECS enabled (authoritative for 1.2M domains on top lists)
  - Only 15 % also return multiple RRsets for four different ECS subnet queries.
- Full address space scan for selected domains
  - Meta uses 137-140 IPv4 and IPv6 addresses (Facebook, Instagram, Whatsapp)
  - Google uses different deployments for Google.com (~2.1k) and YouTube (~1.8k)
  - AWS Route 53 always returns 24 scope prefix lengths
    - Customer can apply their custom mapping

Results

- Scanned a collection of top list domains (3.2 M) with a limited number of queries
  - 53 % of nameservers have ECS enabled (authoritative for 1.2M domains on top lists)
  - Only 15 % also return multiple RRsets for four different ECS subnet queries.
- Full address space scan for selected domains
  - Meta uses 137-140 IPv4 and IPv6 addresses (Facebook, Instagram, Whatsapp)
  - Google uses different deployments for Google.com (~2.1k) and YouTube (~1.8k)
  - AWS Route 53 always returns 24 scope prefix lengths
    - Customer can apply their custom mapping
  - Cloudflare is the largest provider with such domains (99,7 % of all probed domains witch a Cloudflare authoritative nameserver)
  - It seems to always return the same RRset using an ECS scope length of 24

→ Perform distributed measurements

Initial Distributed Results

- We used four different vantage points in two ASes
- Cloudflare is the provider with the largest number of domains with differing answers between VPs
- NSID values suggest we hit different anycast deployments

Initial Distributed Results

- We used four different vantage points in two ASes
- Cloudflare is the provider with the largest number of domains with differing answers between VPs
- NSID values suggest we hit different anycast deployments
- Cloudflare has a peculiar behavior:
  - From a single vantage point we always receive the same address independent of the ECS subnet
  - Address seems to be bound to vantage point
  - Why do Cloudflare nameservers behave like that?

Initial Distributed Results

- We used four different vantage points in two ASes
- Cloudflare is the provider with the largest number of domains with differing answers between VPs
- NSID values suggest we hit different anycast deployments
- Cloudflare has a peculiar behavior:
  - From a single vantage point we always receive the same address independent of the ECS subnet
  - Address seems to be bound to vantage point
  - Why do Cloudflare nameservers behave like that?
    - *Better data on clients with ECS resolvers?*

Initial Distributed Results

- We used four different vantage points in two ASes
- Cloudflare is the provider with the largest number of domains with differing answers between VPs
- NSID values suggest we hit different anycast deployments
- Cloudflare has a peculiar behavior:
  - From a single vantage point we always receive the same address independent of the ECS subnet
  - Address seems to be bound to vantage point
  - Why do Cloudflare nameservers behave like that?
    - *Better data on clients with ECS resolvers?*

→ Hackaton topic on distributed ECS measurements with Ark

Hackaton Results

We use 30 *distributed* subnets and send queries from 130 Ark nodes

- In total 11 RRsets for Cloudflare
  - → Cloudflare performs DNS-based load balancing
- Each VP only observes a single RRset for all 30 client subnets
  - → No ECS-based load balancing

Hackaton Results

We use 30 *distributed* subnets and send queries from 130 Ark nodes

- In total 11 RRsets for Cloudflare
  - → Cloudflare performs DNS-based load balancing
- Each VP only observes a single RRset for all 30 client subnets
  - → No ECS-based load balancing
- 58 RRsets for Amazon
- VPs observes 19 to 23 RRsets
  - → Indicates localized ECS-based load balancing

Hackaton Results

We use 30 *distributed* subnets and send queries from 130 Ark nodes

- In total 11 RRsets for Cloudflare
  - → Cloudflare performs DNS-based load balancing
- Each VP only observes a single RRset for all 30 client subnets
  - → No ECS-based load balancing
- 58 RRsets for Amazon
- VPs observes 19 to 23 RRsets
  - → Indicates localized ECS-based load balancing
- Google and Wikipedia provide consistent ECS-based responses within these 30 queries across all VPs

Conclusion

- ECS scanning helps to better cover ECS-enabled services and their DNS load balancing
- Provide an efficient ECS scanning approach
- Increases usefulness of single VP measurements

Conclusion

- ECS scanning helps to better cover ECS-enabled services and their DNS load balancing
- Provide an efficient ECS scanning approach
- Increases usefulness of single VP measurements
- Distributed measurements are still necessary
- Distributed ECS scans are the next step to high quality data

Conclusion

- ECS scanning helps to better cover ECS-enabled services and their DNS load balancing
- Provide an efficient ECS scanning approach
- Increases usefulness of single VP measurements
- Distributed measurements are still necessary
- Distributed ECS scans are the next step to high quality data
- We have indicators that ECS is used to collect fine-grained data on the nameserver side
- More analysis load balancing algorithms needed

Hackaton results

Domain	Total RRsets	Per VP RRsets	# VPs	NSIDs
Domain on Cloudflare 1	11	1	130	130
Domain on Cloudflare 2	11	1	130	130
Domain on Cloudflare 3	2	1	130	125

- Using 30 *distributed* subnets on 130 Ark nodes
- Cloudflare returns a single address to a single VP for all 30 queries
- AWS does uses several answer patterns depending on the VP

Hackaton results

Domain	Total RRsets	Per VP RRsets	# VPs	NSIDs
Domain on Cloudflare 1	11	1	130	130
Domain on Cloudflare 2	11	1	130	130
Domain on Cloudflare 3	2	1	130	125
www.amazon.com	58	19	1	37
	58	20	9	37
	58	21	42	37
	58	22	77	37
	58	23	1	37

- Using 30 *distributed* subnets on 130 Ark nodes
- Cloudflare returns a single address to a single VP for all 30 queries
- AWS does uses several answer patterns depending on the VP

Hackaton results

Domain	Total RRsets	Per VP RRsets	# VPs	NSIDs
Domain on Cloudflare 1	11	1	130	130
Domain on Cloudflare 2	11	1	130	130
Domain on Cloudflare 3	2	1	130	125
www.amazon.com	58	19	1	37
	58	20	9	37
	58	21	42	37
	58	22	77	37
	58	23	1	37
www.facebook.com	22	21	1	3899
	22	22	129	3899
www.wikipedia.org	6	6	130	3
www.google.co.jp	28	28	130	0
www.google.com	29	29	130	0

- Using 30 *distributed* subnets on 130 Ark nodes
- Cloudflare returns a single address to a single VP for all 30 queries
- AWS does uses several answer patterns depending on the VP