Seven Years in the Life of Hypergiants' Off-nets

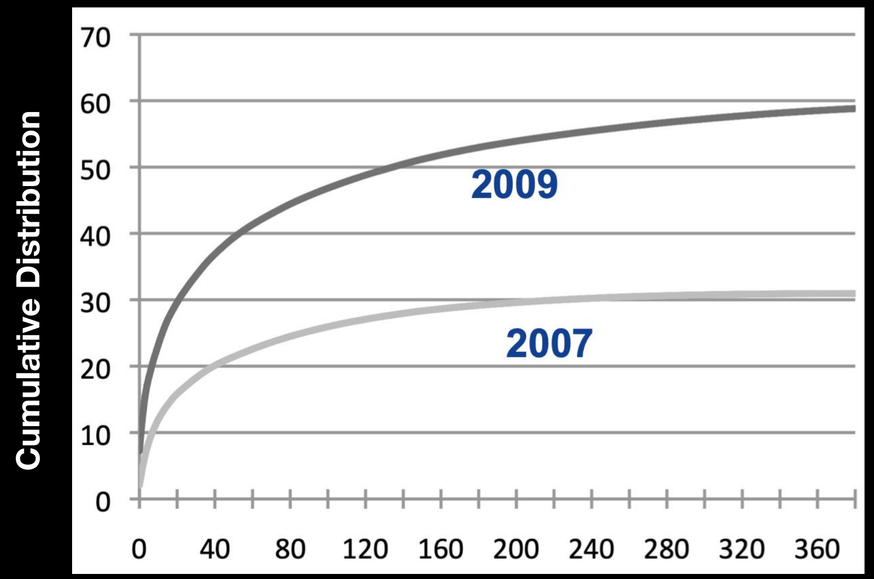
Coseners 2021

Petros Gigis, Matt Calder, Lefteris Manassakis, George Nomikos,
Vasileios Kotronis, Xenofontas Dimitropoulos, Ethan Katz-Bassett, Georgios Smaragdakis

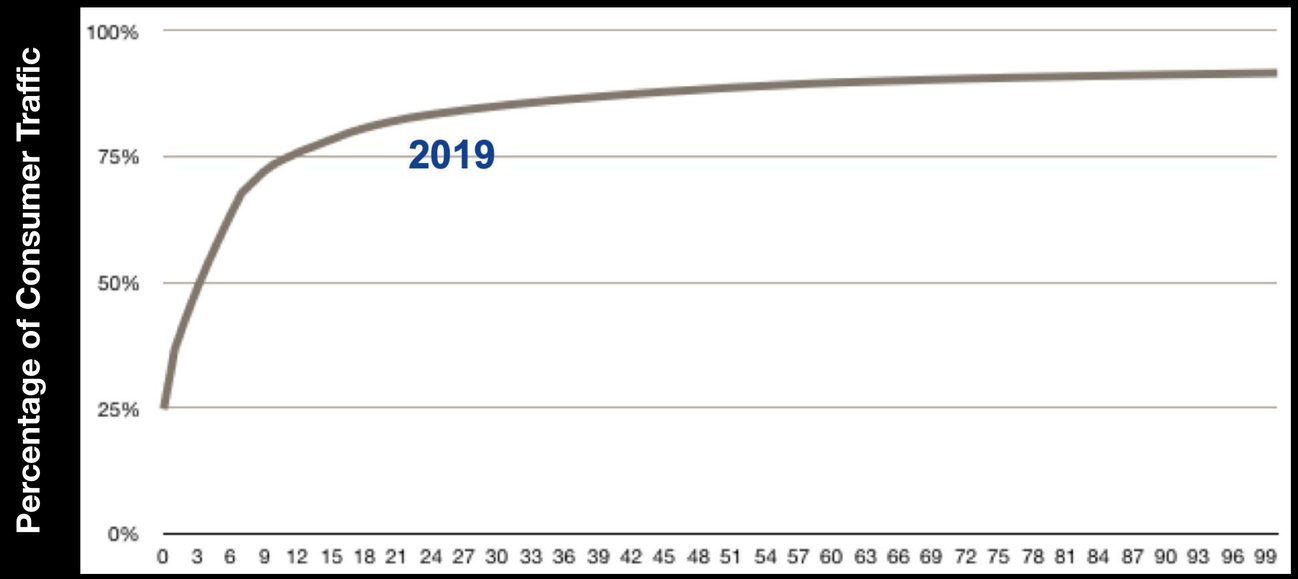
UCL · Microsoft · FORTH Institute of Computer Science · Lancaster University · Columbia University in the City of New York · TU Delft

Hypergiants and Traffic Consolidation

January 2009 - 2019



2009 Cumulative ASN Traffic



2019 Cumulative Hypergiant Traffic

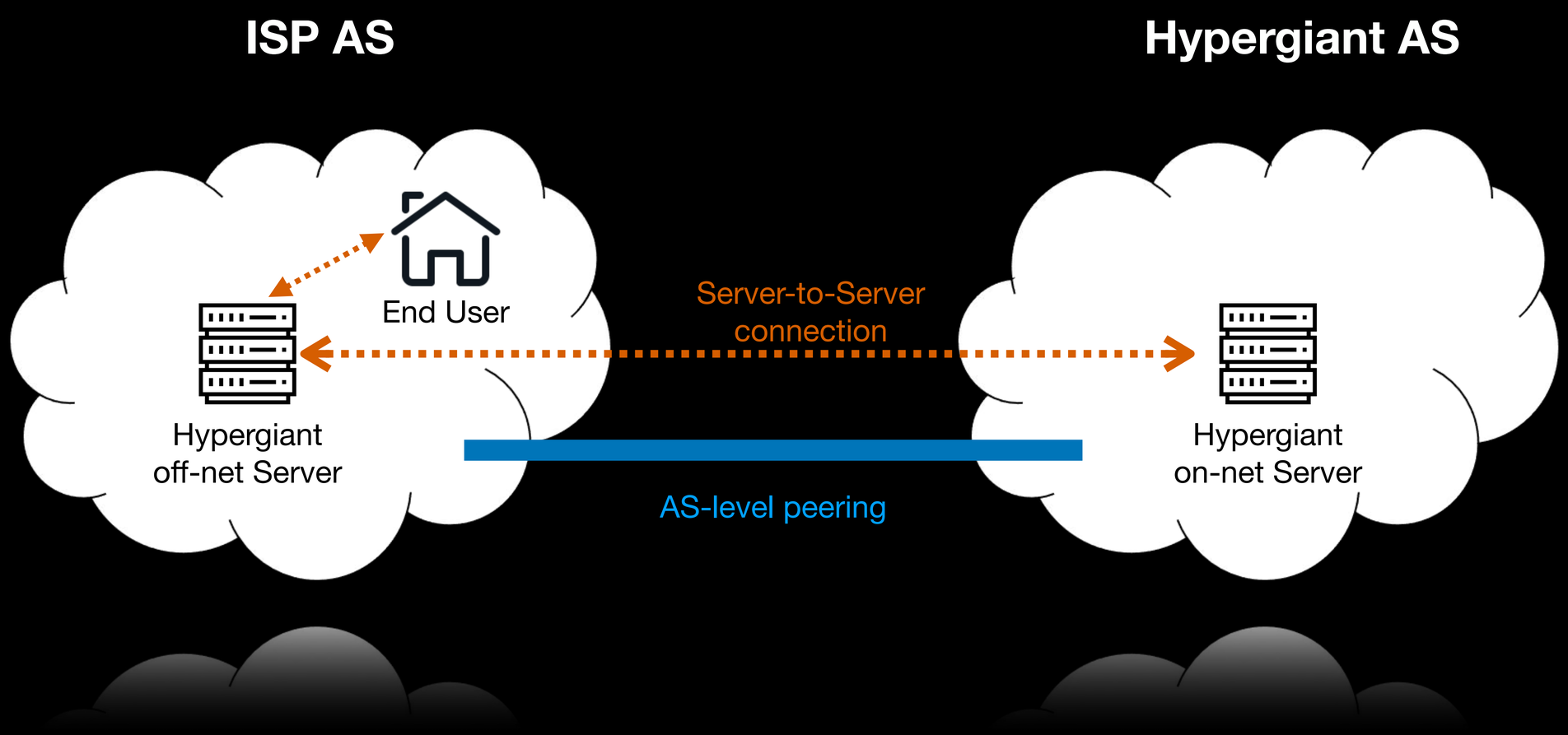
2000 ASN == 50%	150 ASN == 50%	5 HG == 50%
2007	**2009**	**2019**

Source: Internet Traffic 2009-2019, C. Labovitz.

Microsoft's Azure

Source: Microsoft

Hypergiant's *off-net* footprint



Why measuring Hypergiant's Off-nets?

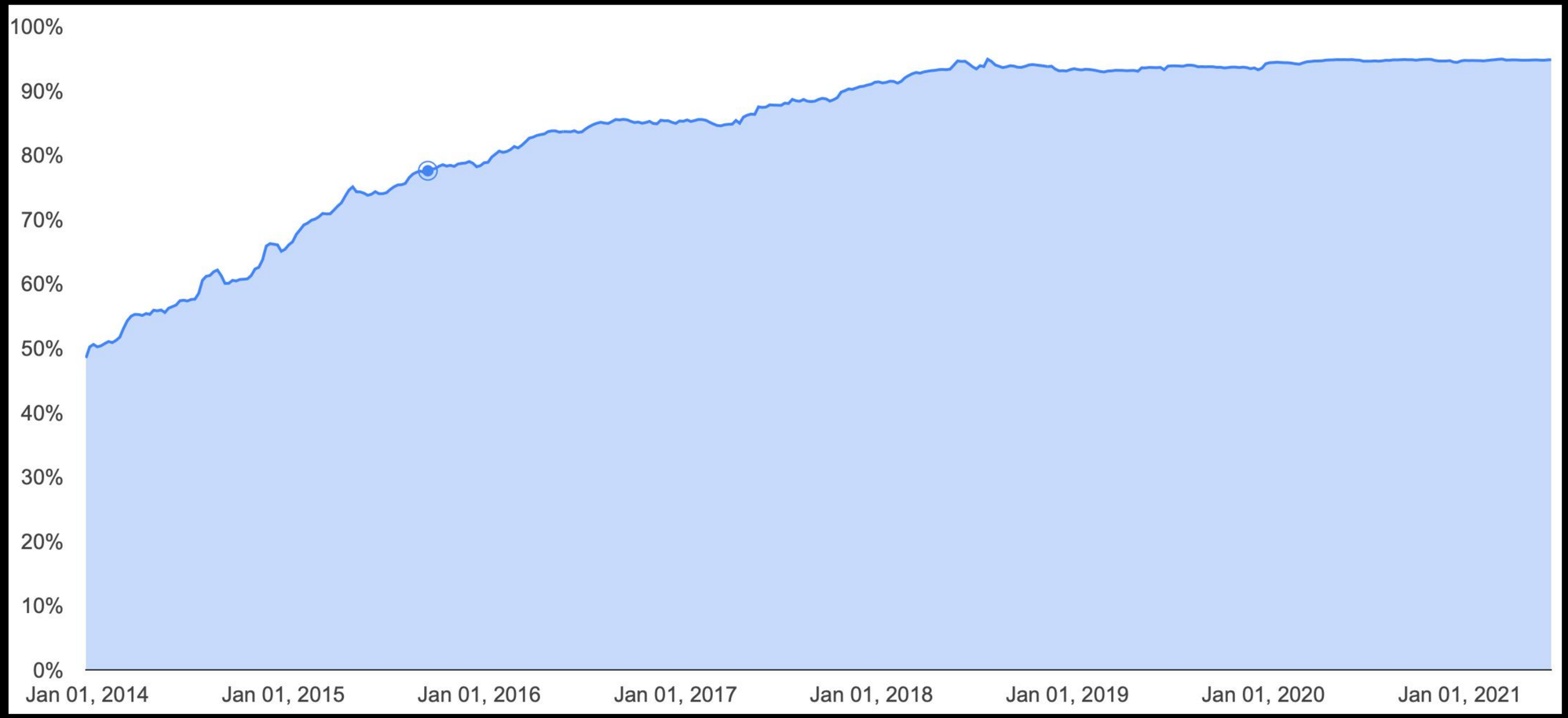
- Understanding the value of peering.
- Understanding how traffic flows in the Internet.
- Localisation of content within an Internet Service Provider (ISP).

Why measuring Hypergiant's Off-nets?

- Is there a generic method to uncover the deployments of Hypergiant off-nets for all Hypergiants?

- Surprisingly, yes! As nowadays traffic is mostly encrypted, HGs include their organisation information in their **TLS certificates**.

Encrypted traffic across Google



Source: Google Transparency Report

What this work is not about:

- Not a head-to-head comparison of different HGs as we do not know:
  1. Business strategies.
  2. Peering agreements.
  3. Performance and cost goals.
- Performance evaluation of different HG off-net footprints is out of the scope of this work.
- In this work, we focus only on uncovering the off-net deployments.

Datasets

- TLS certificate scans:
  - 

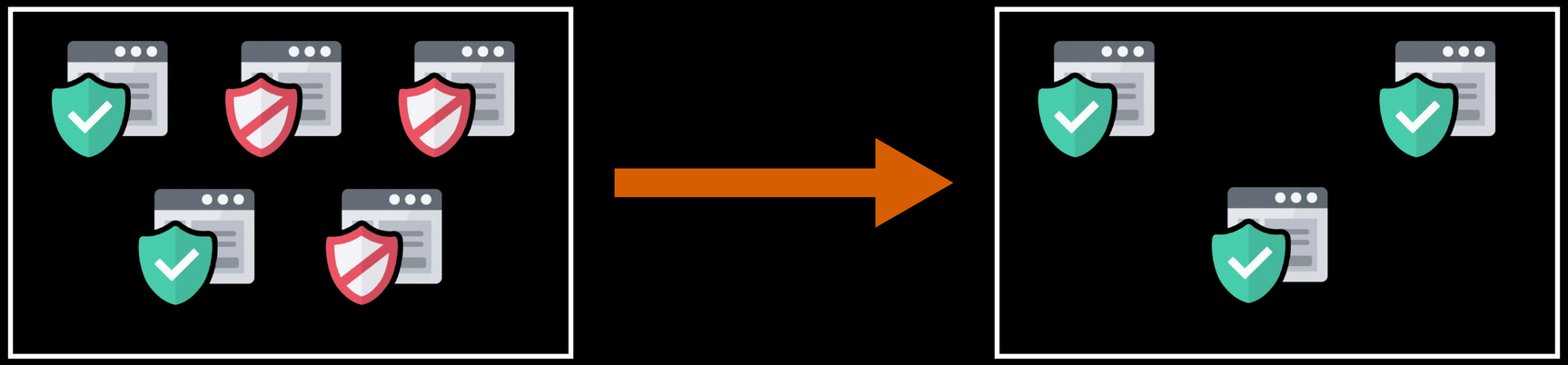
collects certificates in IPv4-wide scans on port 443.
  - Quarterly snapshot from Oct. 2013 to Apr. 2021.
  - 

+ Custom active Scan.
- HTTP(S) headers (Validation):
  - We used corpuses of available HTTP(S) headers from 

from Oct. 2013 to Apr. 2021.

Methodology

- Step 1: **Validate Certificates**
  - Exclude self-signed, expired and certificates with a non-verified chain.

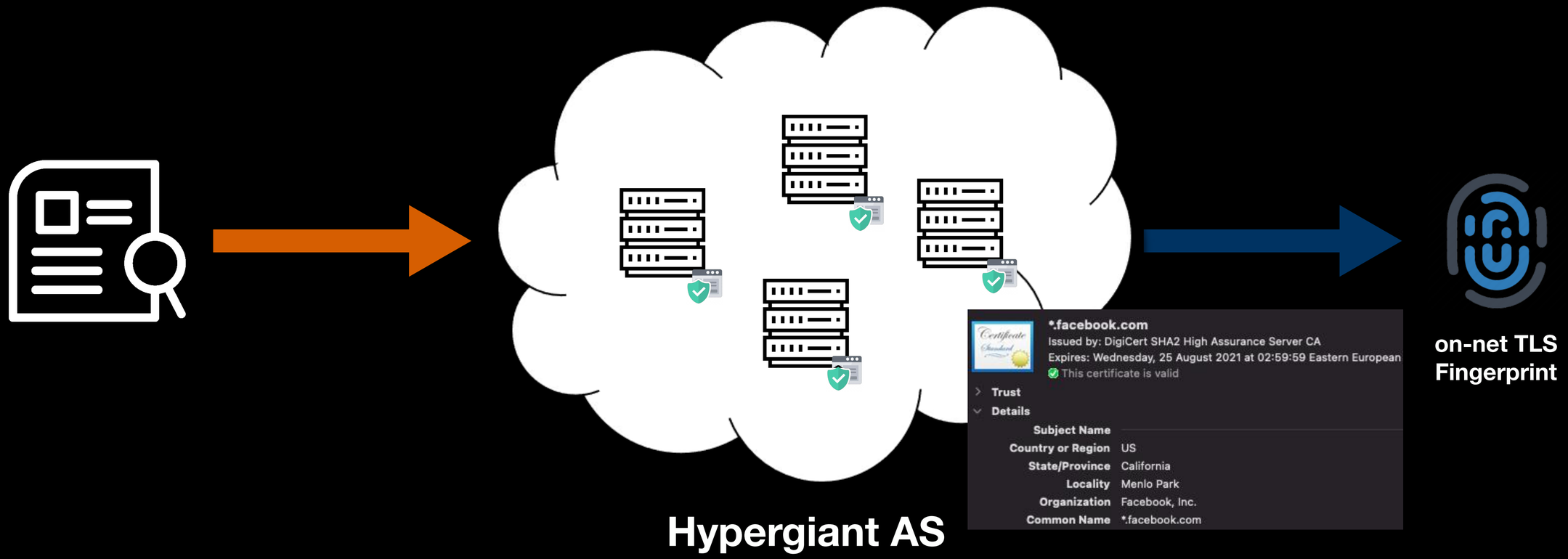


TLS certificate dataset

Valid certificates

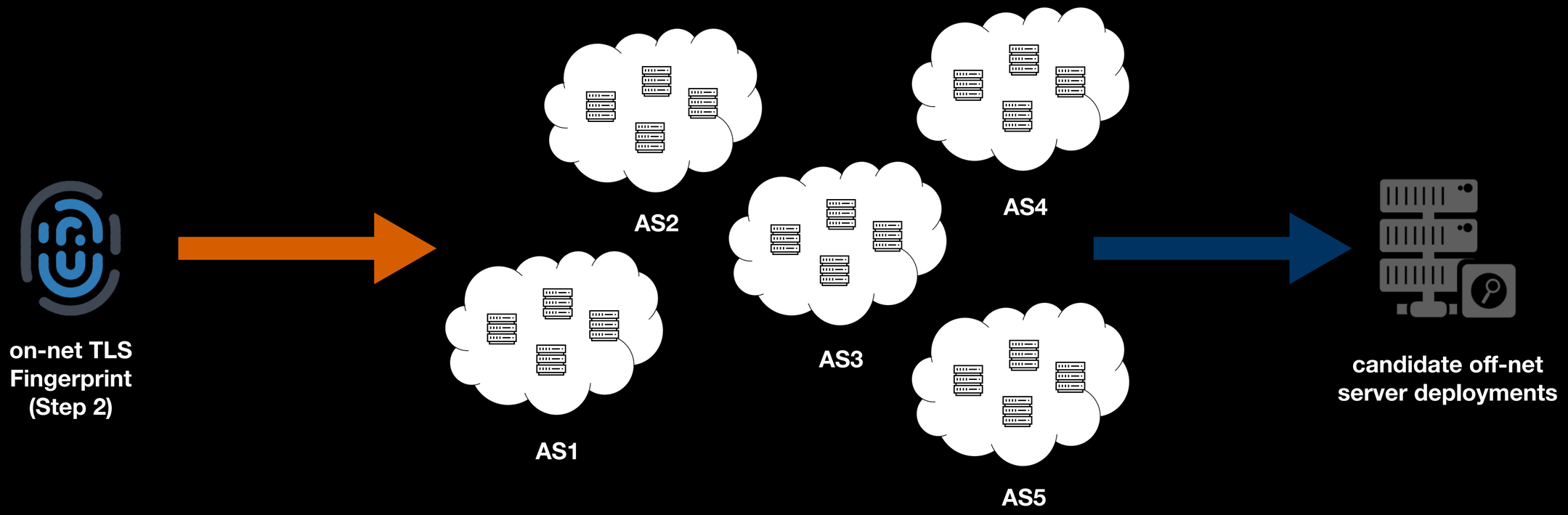
Methodology

- Step 2: **Learn Hypergiant TLS Fingerprints**
  - Input the HG keyword e.g., “facebook” and the TLS scans for all on-net IPs.



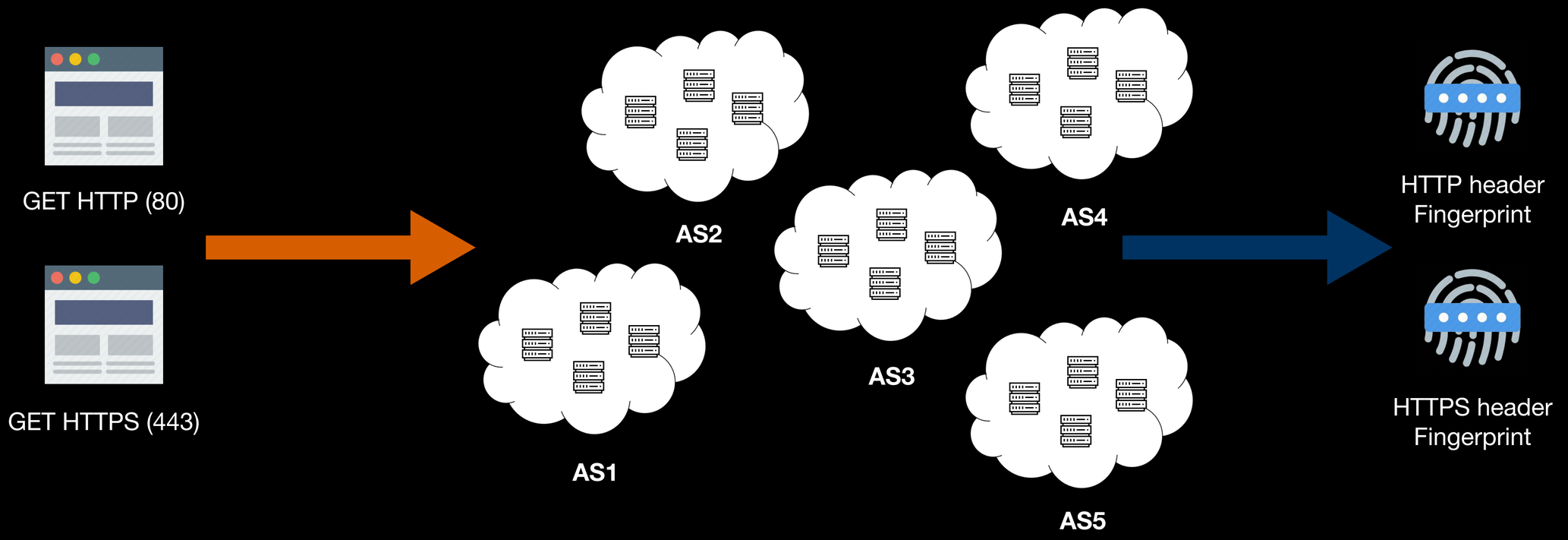
Methodology

- Step 3: **Use Fingerprints to Identify candidate off-nets**
  - Search for certificates matching the on-net fingerprints.



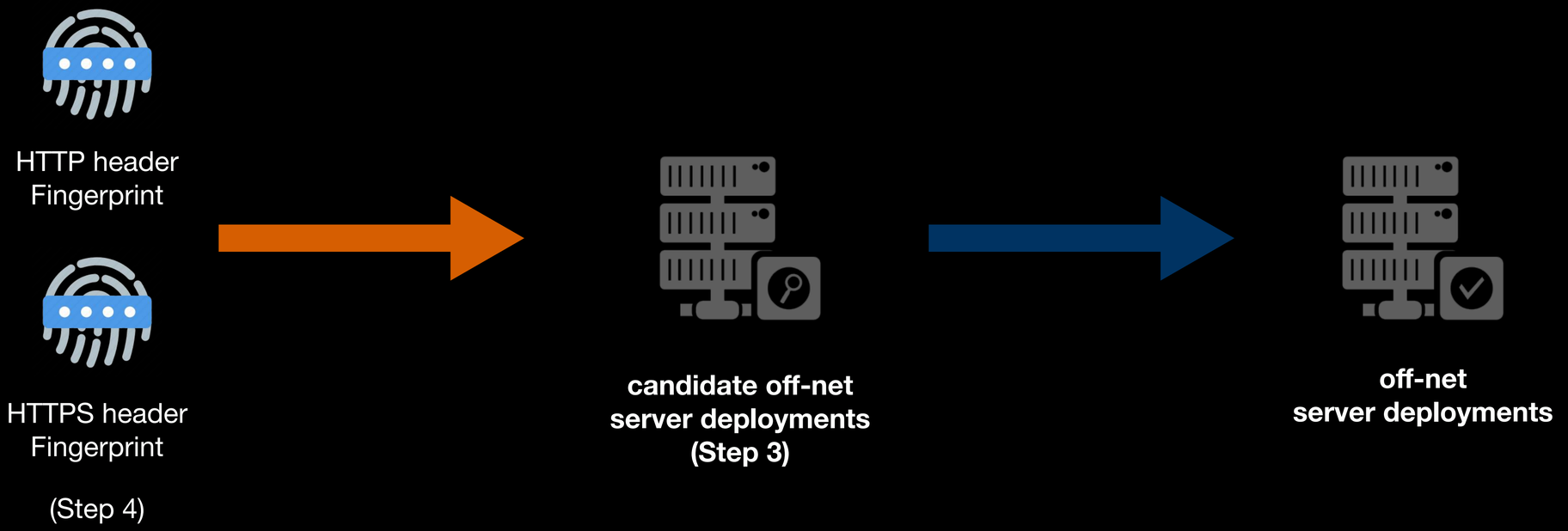
Methodology

- Step 4: **Learn Hypergiant HTTP(S) Fingerprints**
  - Identify fingerprints in Hypergiant HTTP(S) headers.



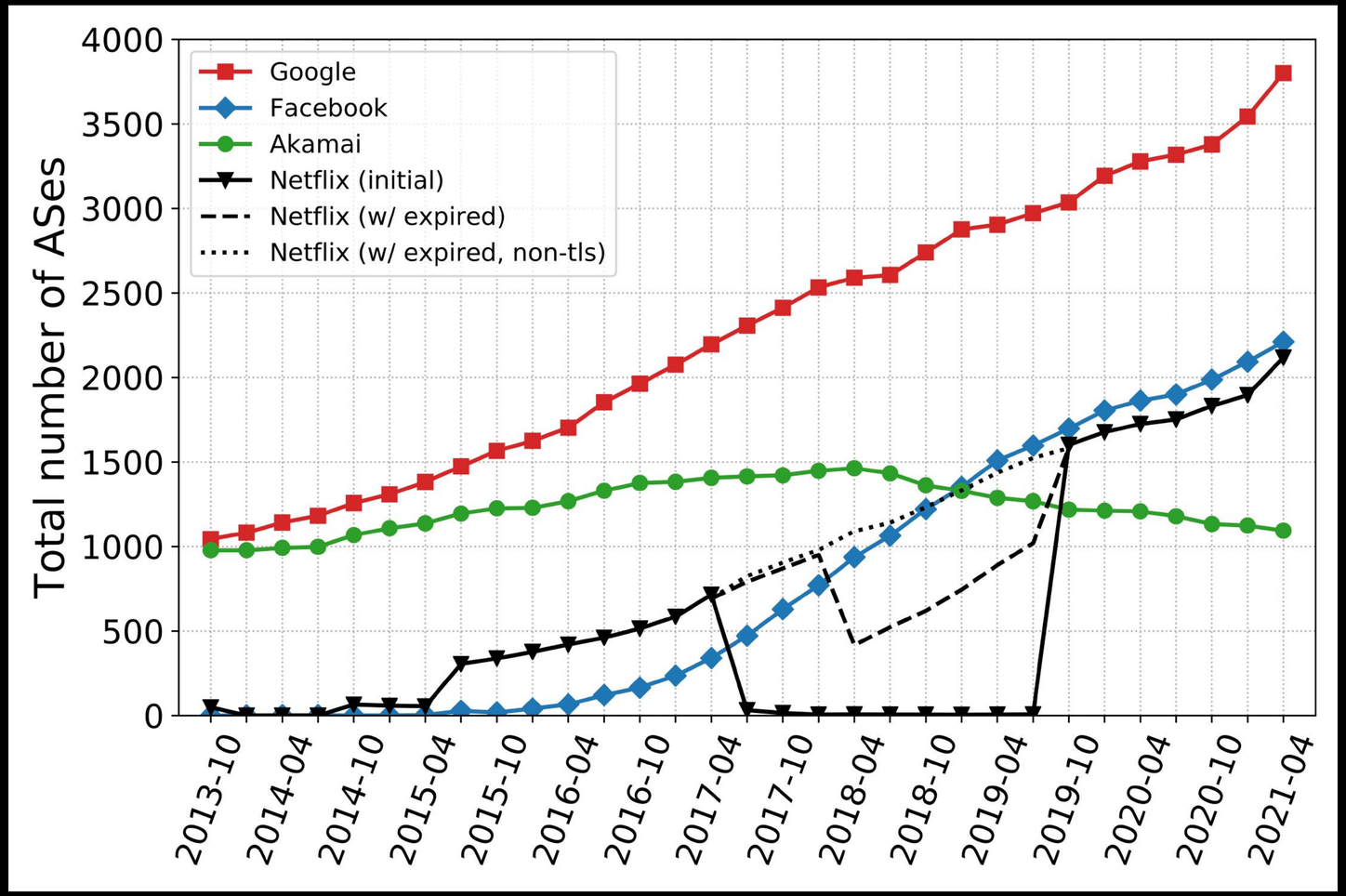
Methodology

- Step 5: **<u>Confirm Candidates Using HTTP(S)</u>**
  - Apply HTTP(S) fingerprints (Step 4) to the off-net candidates (Step 3) and classify as off-nets any that match the HG fingerprints.



Hypergiants' *off-nets* Expansion

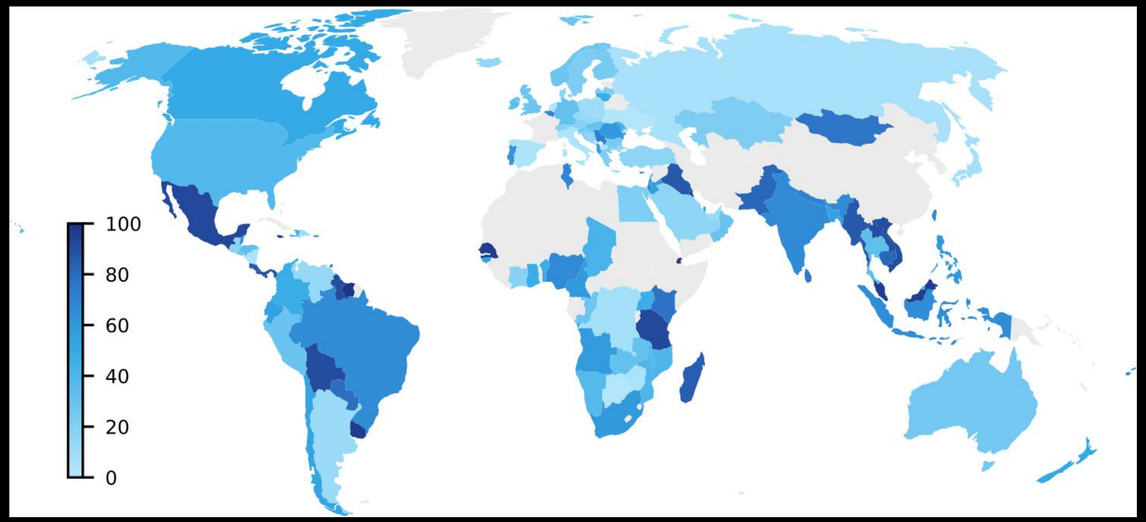
Longitudinal Growth (2013-2021)



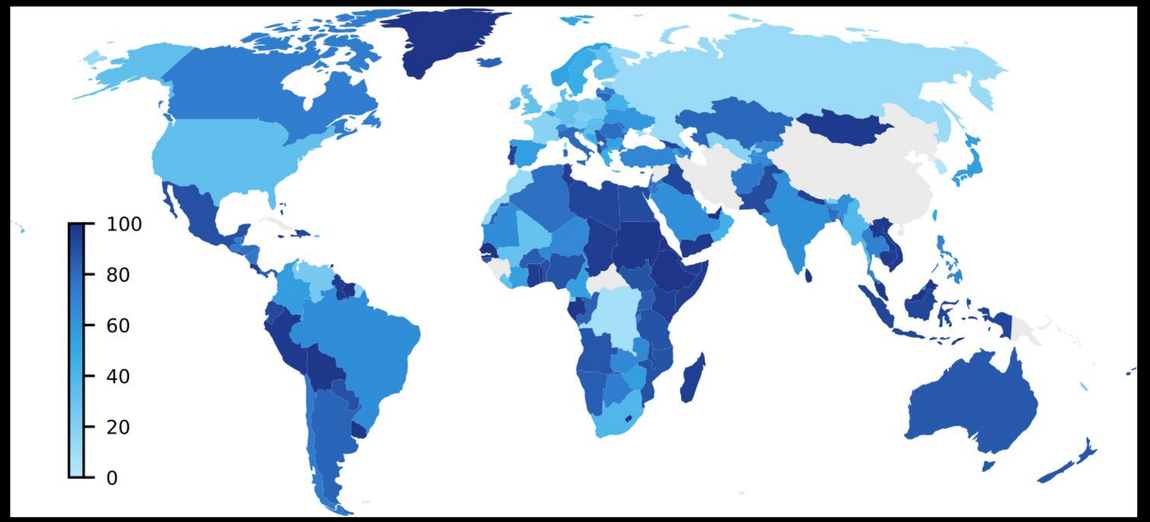
Off-net footprint growth for top-4 HGs (Google, Facebook, Netflix and Akamai) over time.

Hypergiants' Off-net Expansion

Internet User Population Coverage



Oct. 2017



Apr. 2021

Facebook's off-net footprint Internet user coverage (percentage).

- Facebook in 2017 announced that it had plans to expand in Africa and other developing regions.

Validation

- **Validation from Hypergiants.**
  - Three replied to our survey. All reported estimation with less than 10% error.
- **Comparison to Earlier Results.**
  - Google: Previous study in April 2016 reported 1445 ASes.
    We identified 98% of them, plus 283 additional ASes.
  - *Facebook: Comparison with three studies:*
    We identified 96% (2018), 94% (2019) and 95% (2021) of the ASes.
  - *Netflix: Previous study in May 2017 reported 743 ASes, we report 769 ASes.*

Summary

- Generic methodology to uncover off-net deployments.
- Exponential growth of 3 out of 4 top HGs.
- More than 4.5k ASes host any of the top-4 HGs.
- All results and many more are available on the upcoming SIGCOMM paper.
- Artifacts including our software and active scan will be available.
- We will launch an interactive portal (in the following weeks).

Thank you!

Questions?