Seven Years in the Life of Hypergiants' Off-nets

Petros Gigis[1]
Matt Calder[2,3], Lefteris Manassakis[4], George Nomikos[4,5], Vasileios Kotronis[4],
Xenofontas Dimitropoulos[4,6], Ethan Katz-Bassett[3], Georgios Smaragdakis[7]

[1] University College London, [2] Microsoft, [3] Columbia University,
[4] FORTH-ICS, [5] Lancaster University, [6] University of Crete, [7] TU Delft

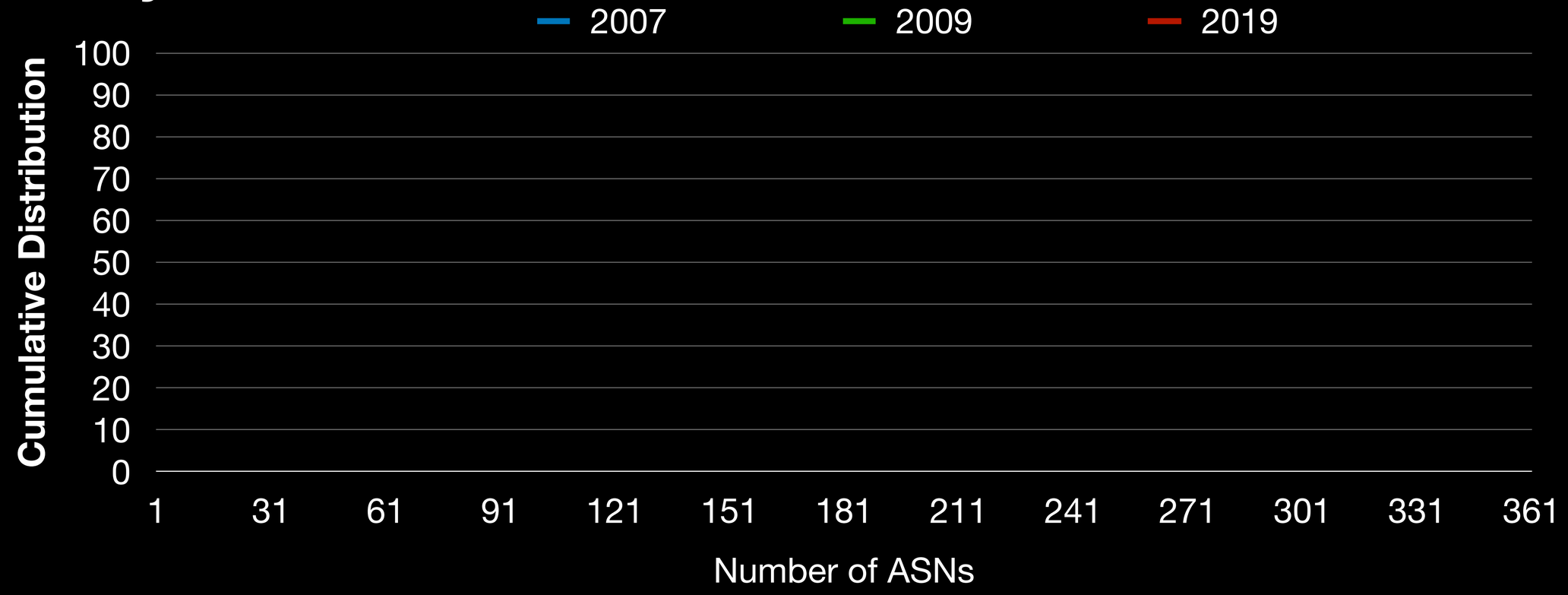
ACM SIGCOMM, August 2021

Hypergiants and Traffic Consolidation

January 2009 - 2019

Hypergiants and Traffic Consolidation

January 2009 - 2019

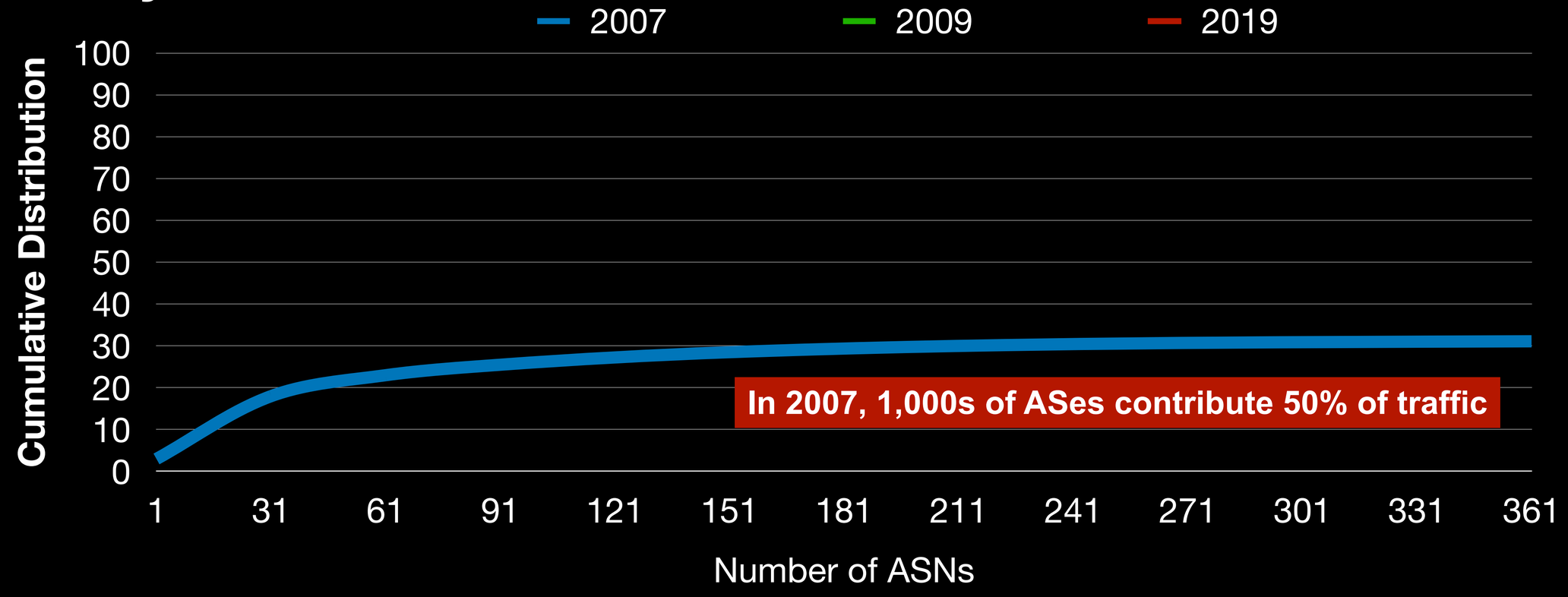


2000 ASN == 50%	150 ASN == 50%
2007	**2009**

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

Hypergiants and Traffic Consolidation

January 2009 - 2019

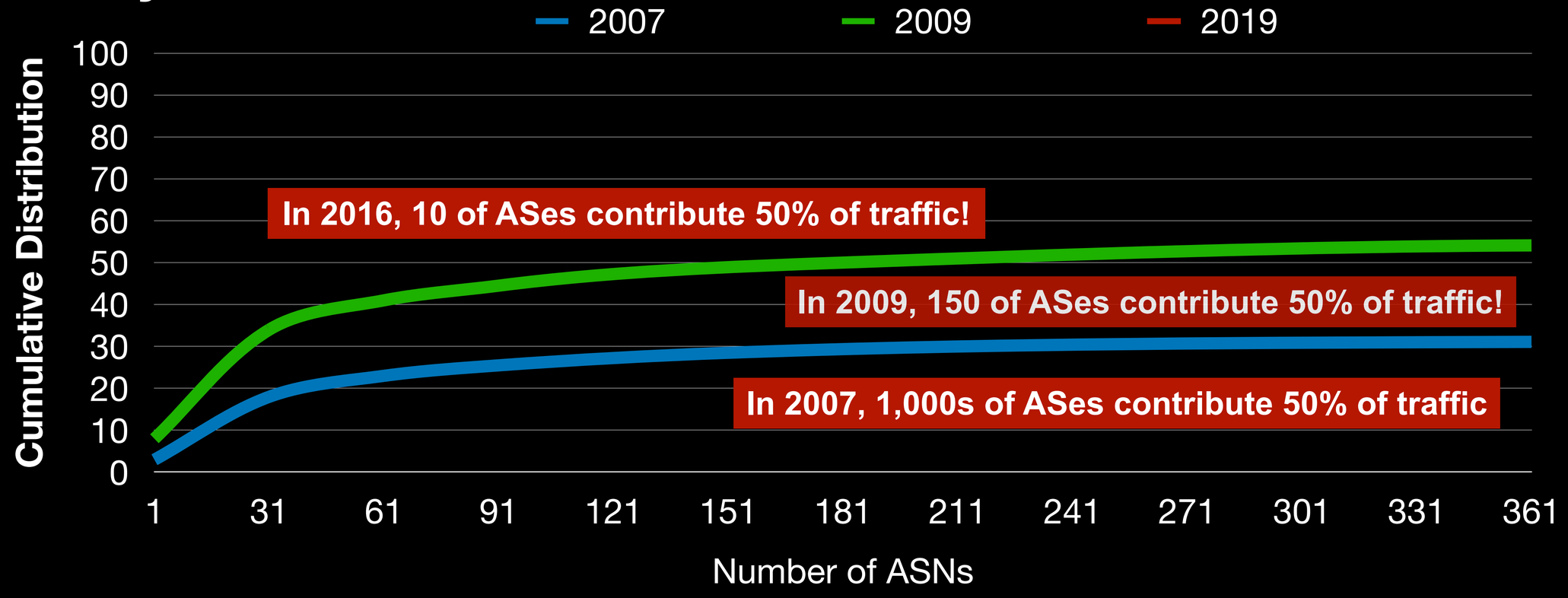


2000 ASN == 50%	150 ASN == 50%
2007	**2009**

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

Hypergiants and Traffic Consolidation

January 2009 - 2019

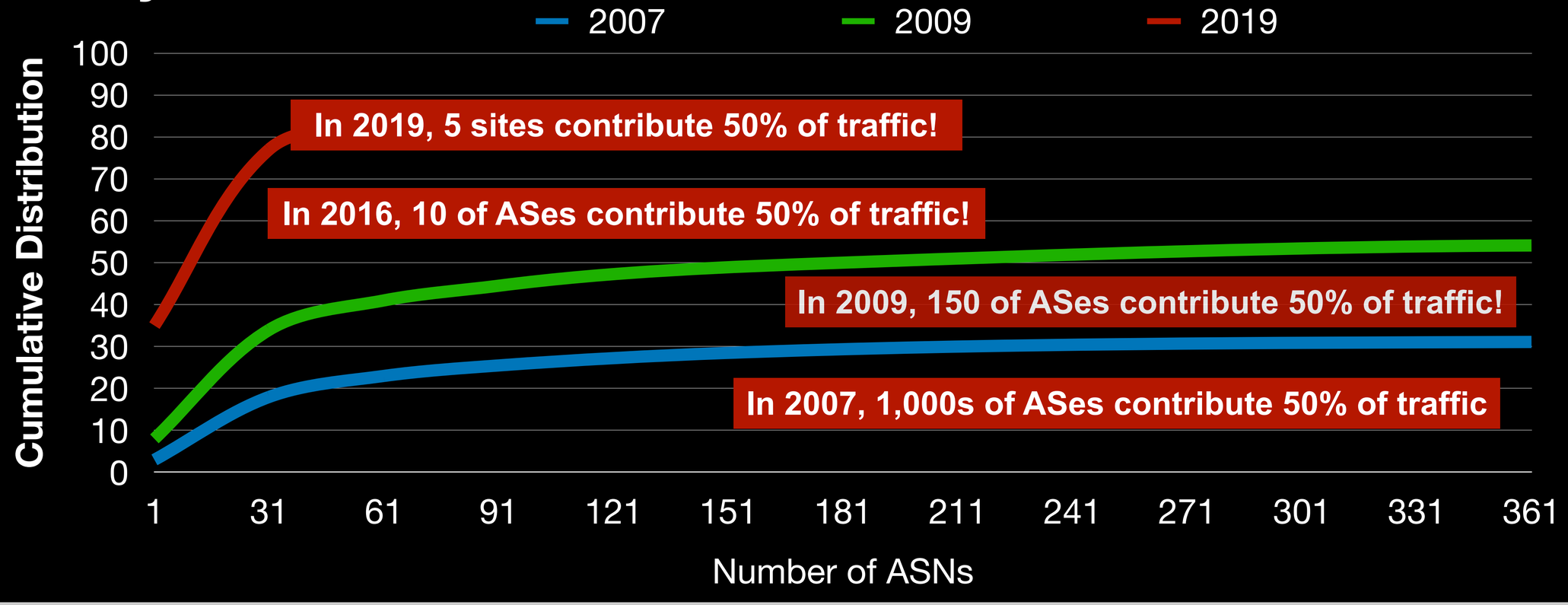


2000 ASN == 50%	150 ASN == 50%
2007	**2009**

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

Hypergiants and Traffic Consolidation

January 2009 - 2019



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

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

Microsoft’s Azure Global Network

Source: Microsoft (January 2020)

Microsoft’s Azure Global Network

Construct datacenters

Source: Microsoft (January 2020)

Microsoft’s Azure Global Network

Construct datacenters

Roll out fiber to built their backbone

Source: Microsoft (January 2020)

Microsoft's Azure Global Network

Construct datacenters

Peer at IXPs and collocation facilities

Roll out fiber to built their backbone

Source: Microsoft (January 2020)

Microsoft’s Azure Global Network

Construct datacenters

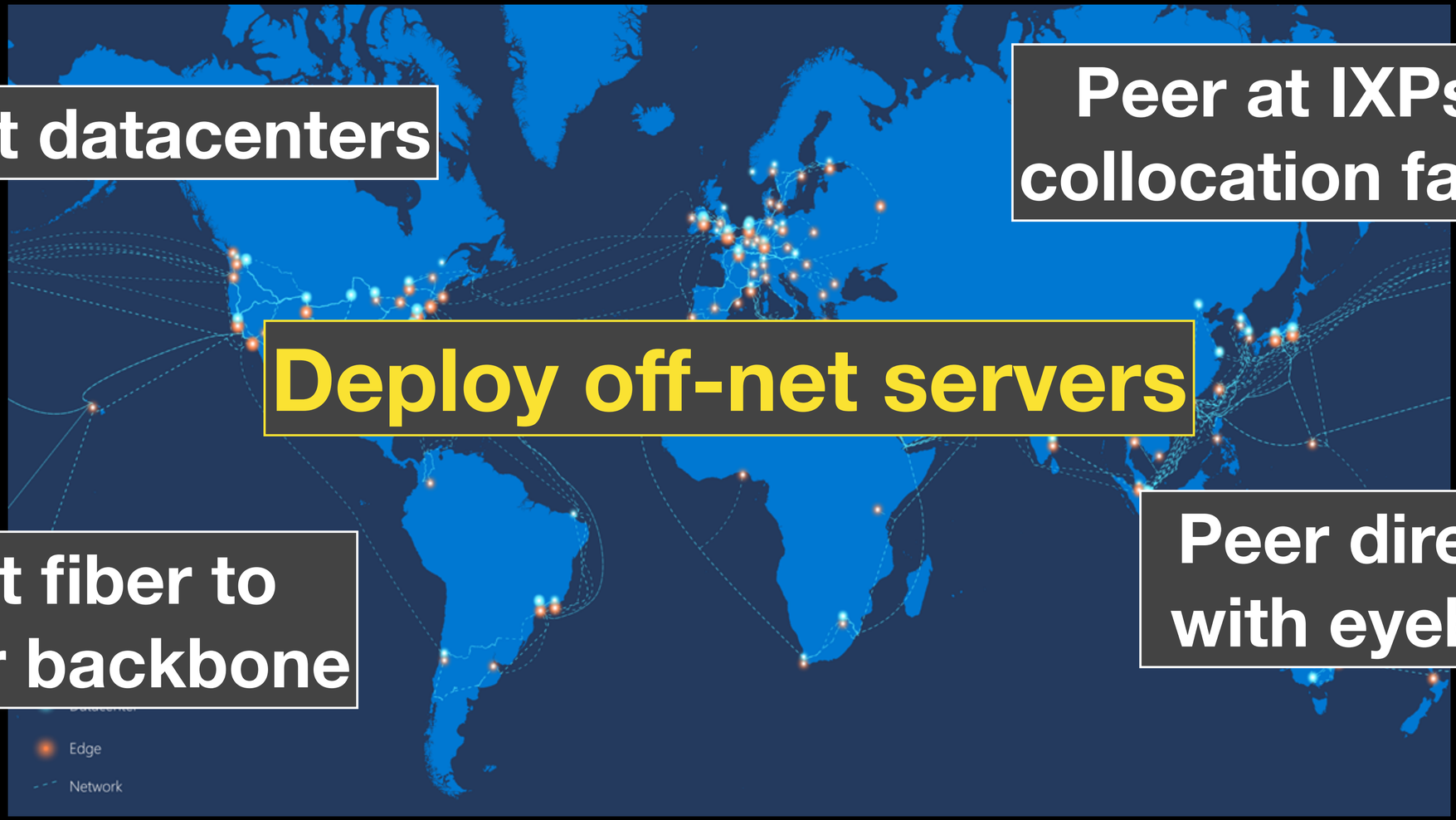
Peer at IXPs and collocation facilities

Roll out fiber to built their backbone

Peer directly with eyeballs

Source: Microsoft (January 2020)

Microsoft's Azure Global Network



Construct datacenters

Peer at IXPs and collocation facilities

Deploy off-net servers

Roll out fiber to built their backbone

Peer directly with eyeballs

Source: Microsoft (January 2020)

Hypergiant’s off-net footprint

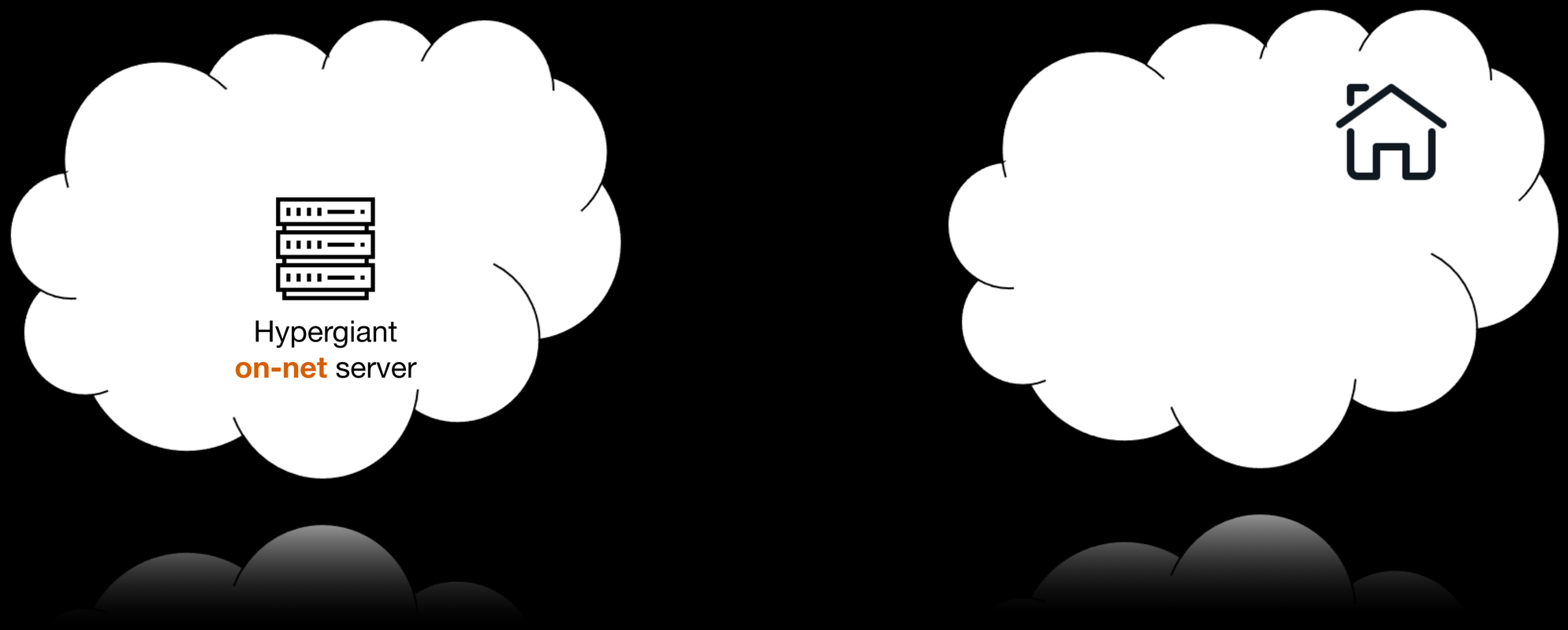
Hypergiant’s off-net footprint

On-net: A server deployment **inside** the HG own network.

Hypergiant’s off-net footprint

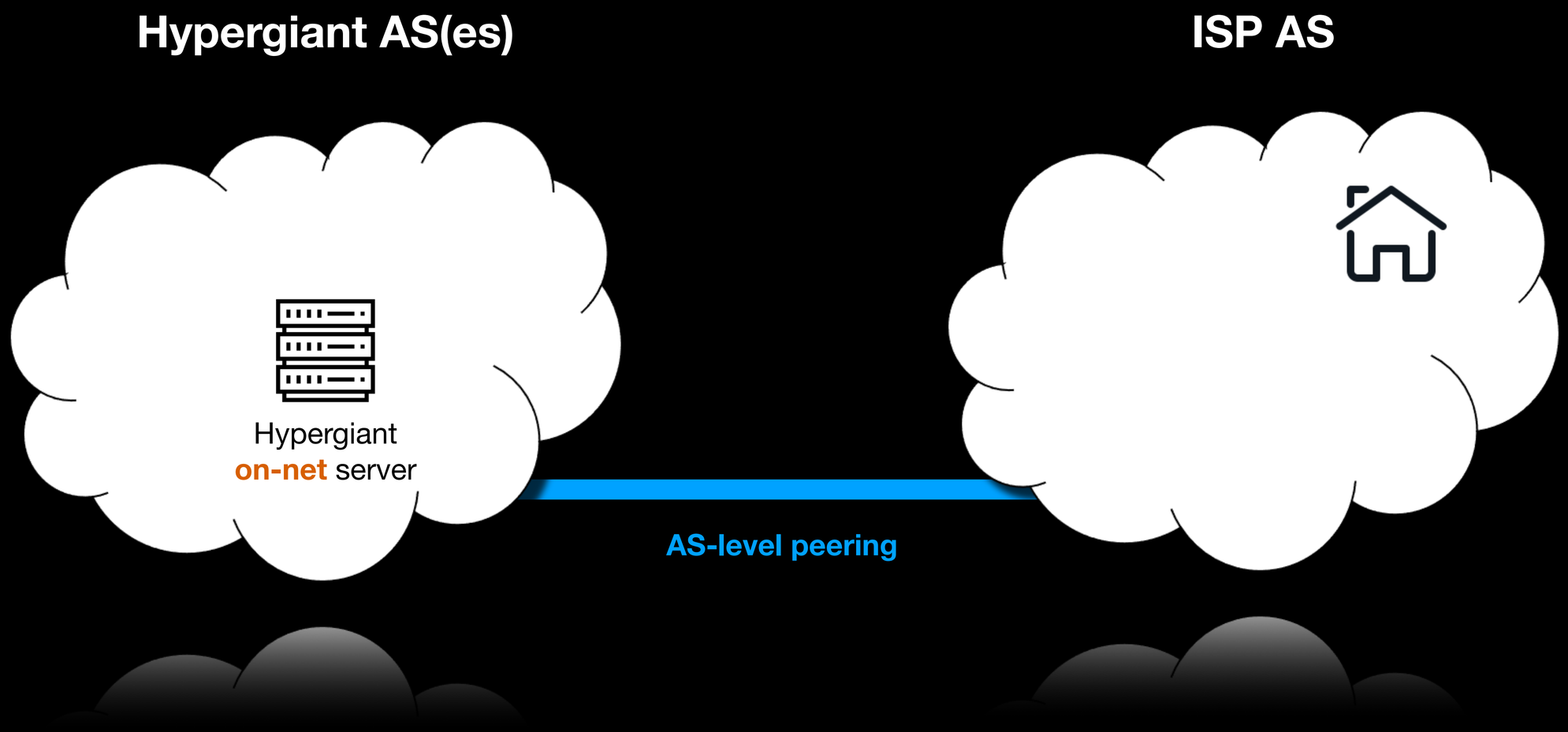
Hypergiant AS(es)

ISP AS



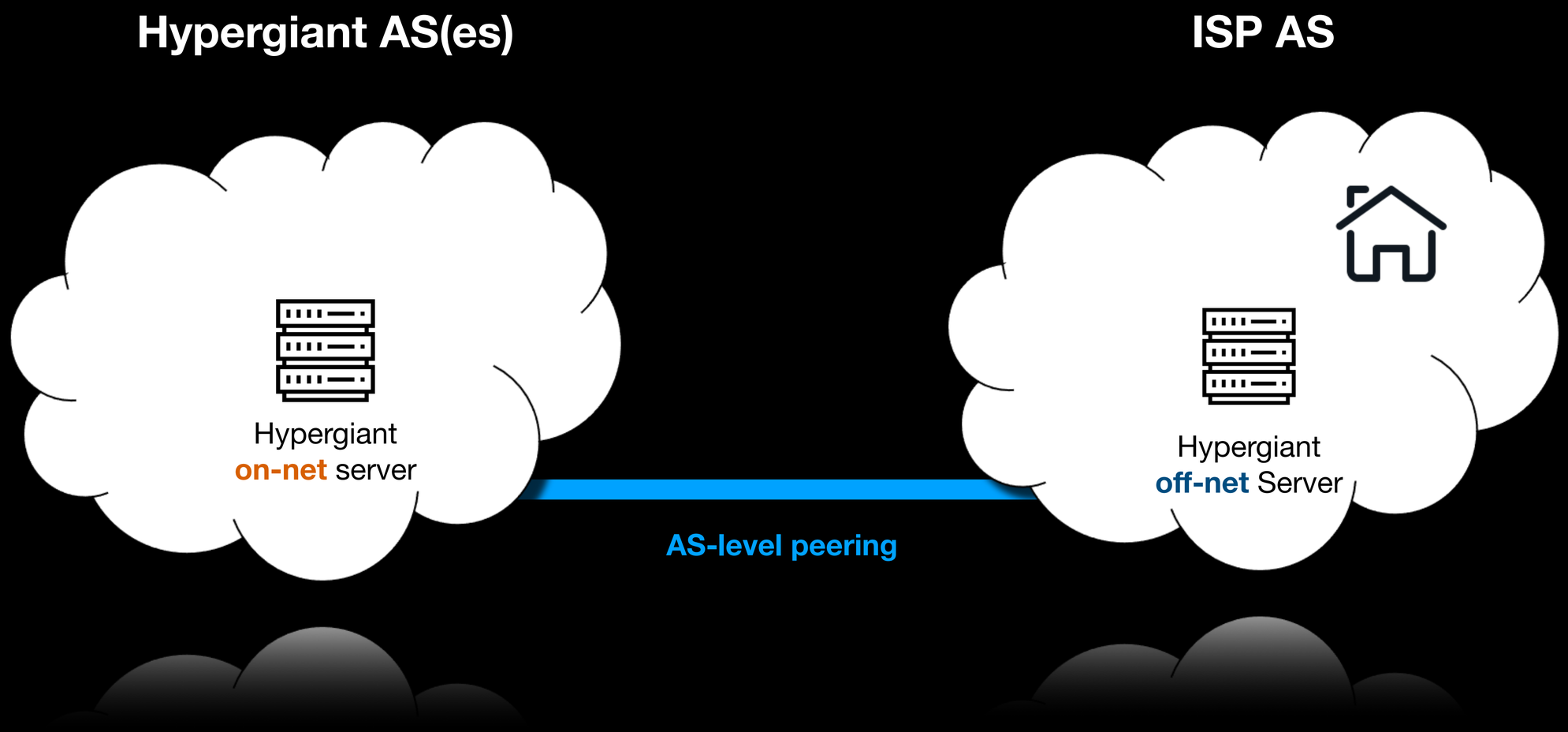
On-net: A server deployment **inside** the HG own network.

Hypergiant's off-net footprint



On-net: A server deployment **inside** the HG own network.

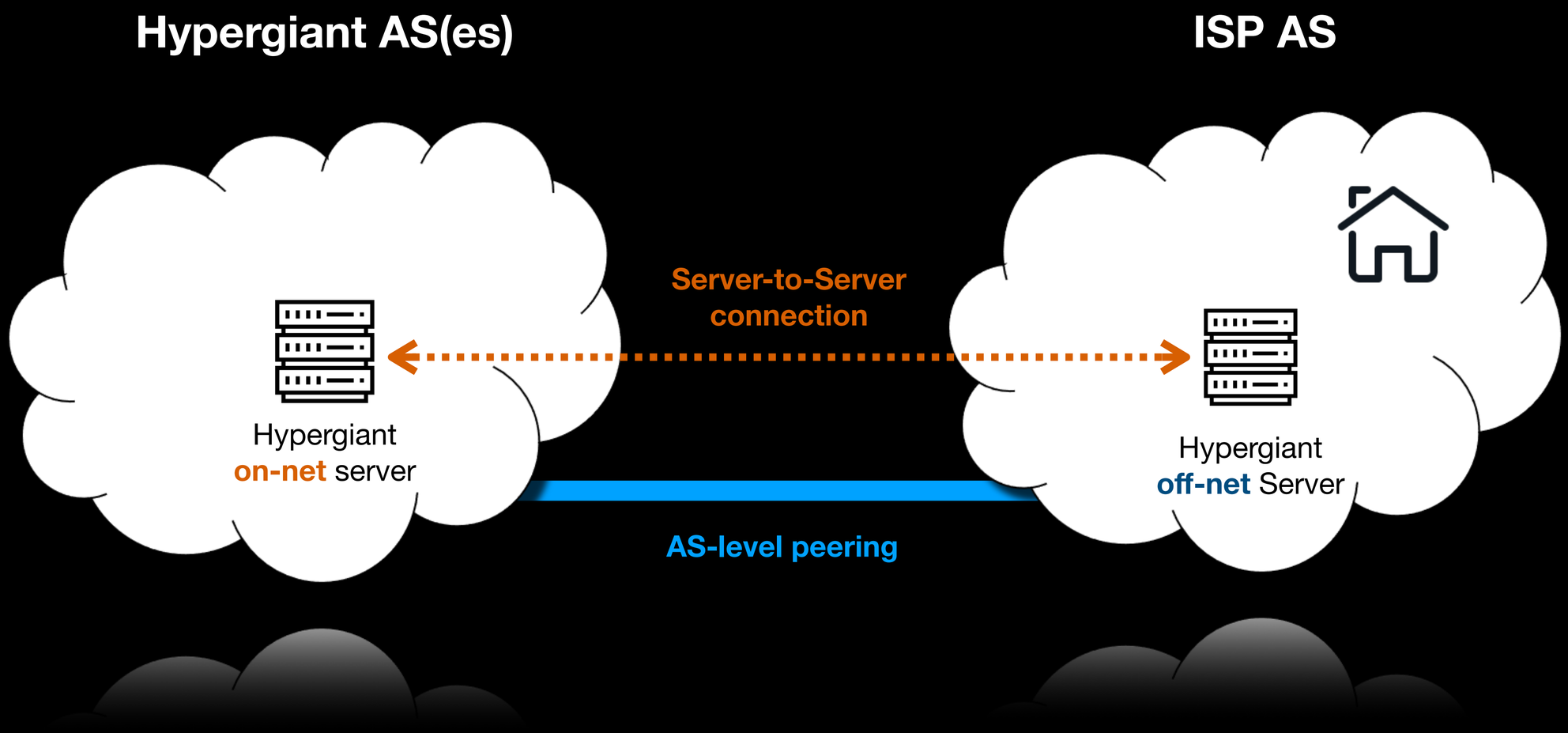
Hypergiant’s off-net footprint



On-net: A server deployment **inside** the HG own network.

Off-net: A server deployment **outside** of the HG own network.

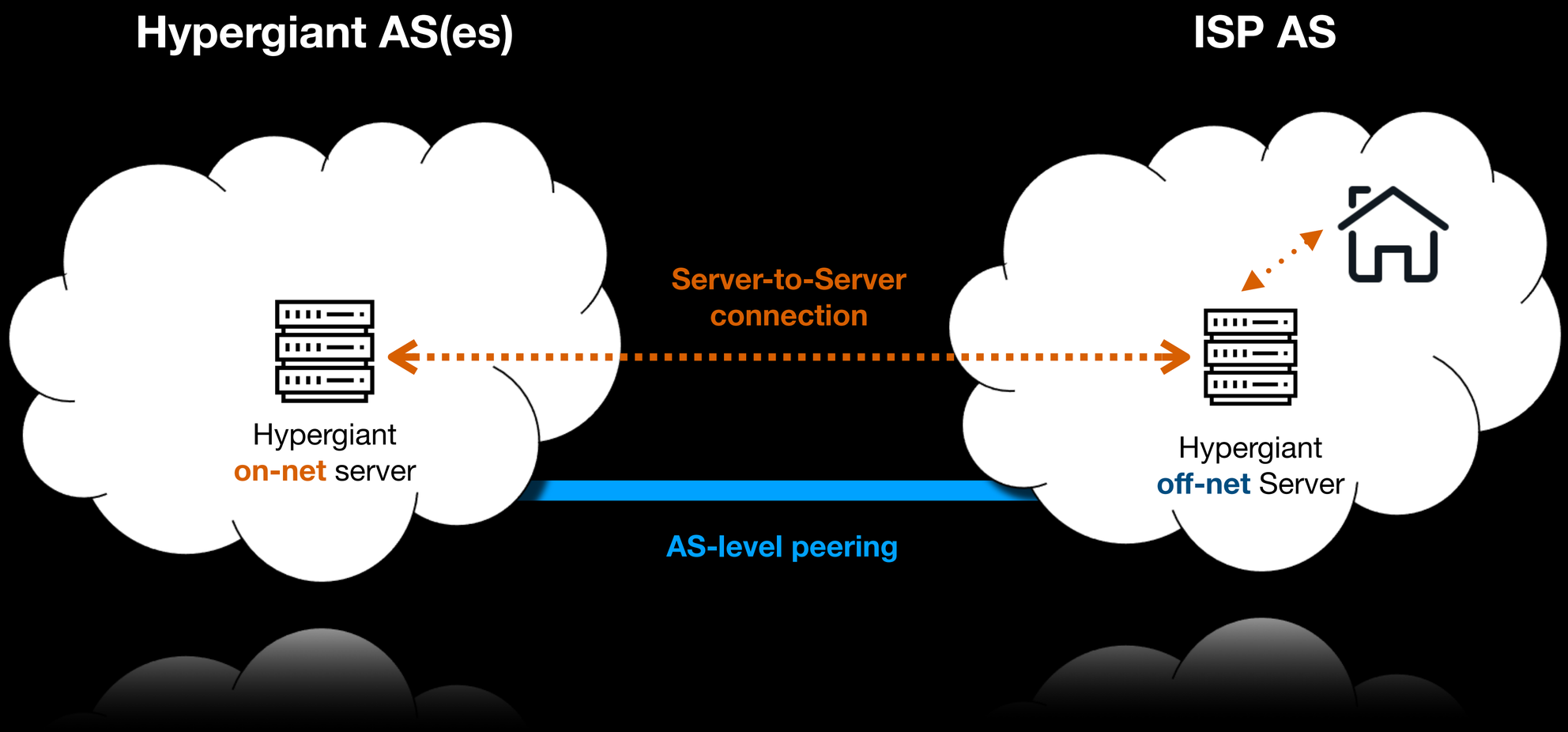
Hypergiant's off-net footprint



On-net: A server deployment **inside** the HG own network.

Off-net: A server deployment **outside** of the HG own network.

Hypergiant's off-net footprint



On-net: A server deployment **inside** the HG own network.

Off-net: A server deployment **outside** of the HG own network.

Why measure Hypergiant’s Off-nets?

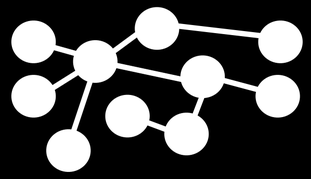
Why measure Hypergiant's Off-nets?

- Revisit the value of peering and how traffic flows in the Internet.

Why measure Hypergiant's Off-nets?

- Revisit the value of peering and how traffic flows in the Internet.

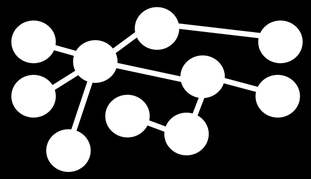
- Understand the evolution of the Internet structure.



Why measure Hypergiant's Off-nets?

- Revisit the value of peering and how traffic flows in the Internet.

- Understand the evolution of the Internet structure.

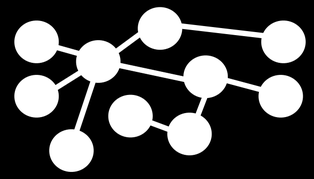


- Localisation of content within an Internet Service Provider (ISP).

Why measure Hypergiant's Off-nets?

- Revisit the value of peering and how traffic flows in the Internet.

- Understand the evolution of the Internet structure.



- Localisation of content within an Internet Service Provider (ISP).

- Performance potential for emerging networks (e.g., 

).

How to measure Hypergiant’s Off-nets?

How to measure Hypergiant's Off-nets?

- Traditional techniques (e.g., IP-to-AS mapping) do not work.

How to measure Hypergiant's Off-nets?

- Traditional techniques (e.g., IP-to-AS mapping) do not work.

- Existing approaches[1,2] to discover off-nets **lack generality** and are **fragile to HG changes**.

[1] "Mapping the Expansion of Google's Serving Infrastructure", Calder et. al., IMC 2013
[2] "Open Connect Everywhere: A Glimpse at the Internet Ecosystem through the Lens of the Netflix CDN", Böttger et. al., CCR 48

How to measure Hypergiant's Off-nets?

- Traditional techniques (e.g., IP-to-AS mapping) do not work.

- Existing approaches[1,2] to discover off-nets **lack generality** and are **fragile to HG changes**.

  [1] "Mapping the Expansion of Google's Serving Infrastructure", Calder et. al., IMC 2013
  [2] "Open Connect Everywhere: A Glimpse at the Internet Ecosystem through the Lens of the Netflix CDN", Böttger et. al., CCR 48

- **So, is there a generic method to uncover the off-nets of all Hypergiants?**

How to measure Hypergiant's Off-nets?

- Traditional techniques (e.g., IP-to-AS mapping) do not work.

- Existing approaches[1,2] to discover off-nets **lack generality** and are **fragile to HG changes**.

  [1] "Mapping the Expansion of Google's Serving Infrastructure", Calder et. al., IMC 2013
  [2] "Open Connect Everywhere: A Glimpse at the Internet Ecosystem through the Lens of the Netflix CDN", Böttger et. al., CCR 48

- **So, is there a generic method to uncover the off-nets of all Hypergiants?**
  - **Surprisingly, yes!**

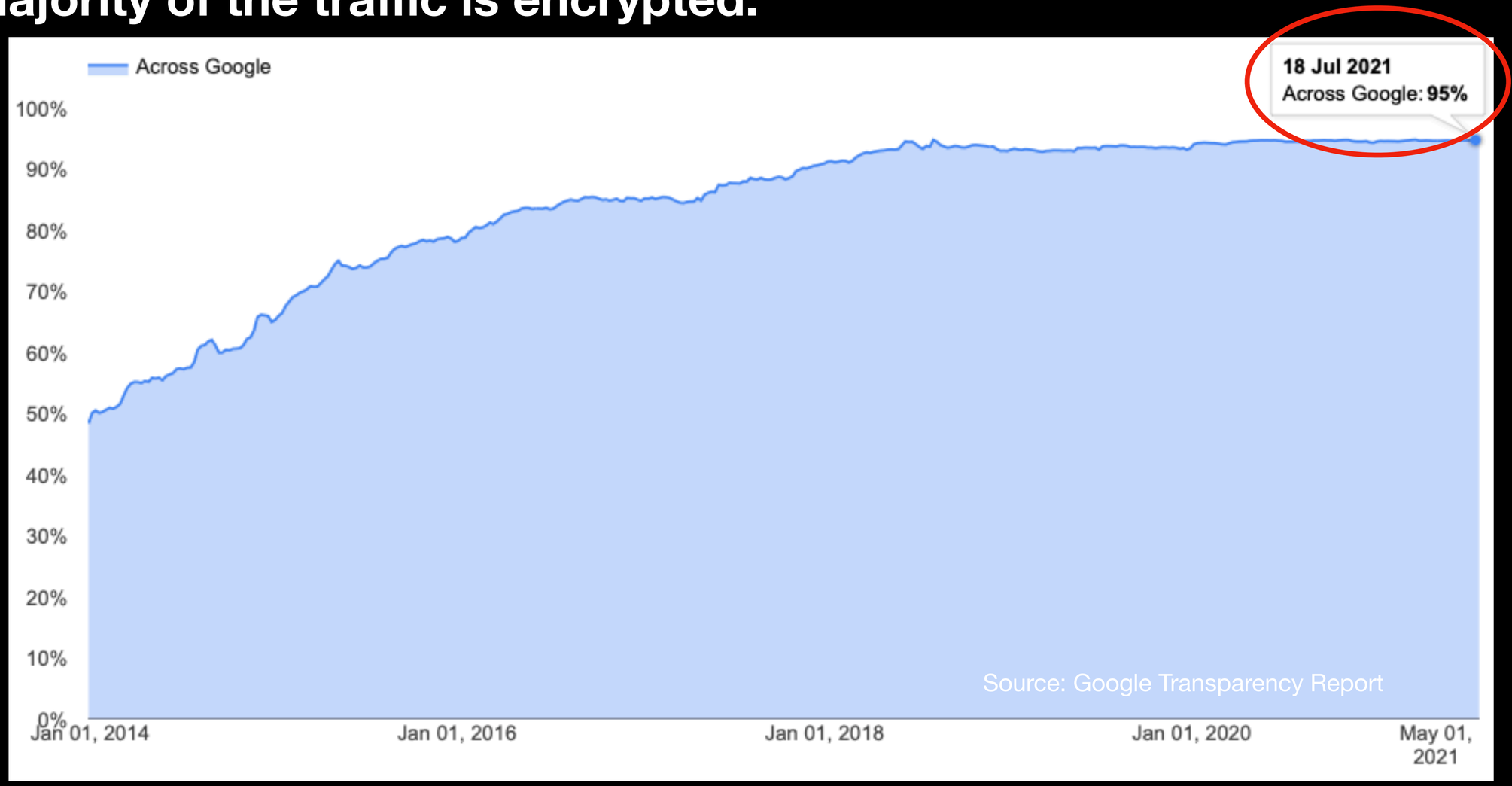
Encrypted traffic is the new standard

Encrypted traffic is the new standard

- **The majority of the traffic is encrypted.**

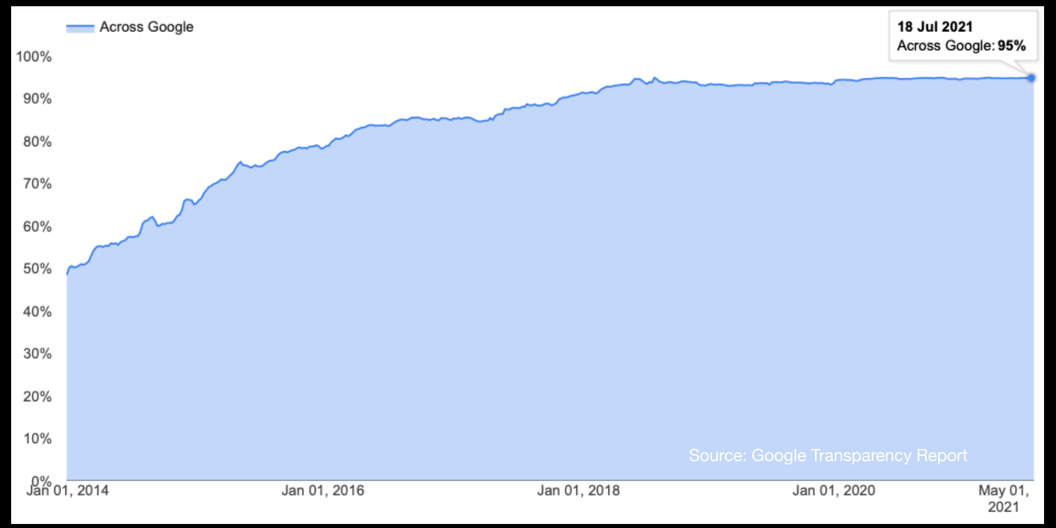
Encrypted traffic is the new standard

- **The majority of the traffic is encrypted.**



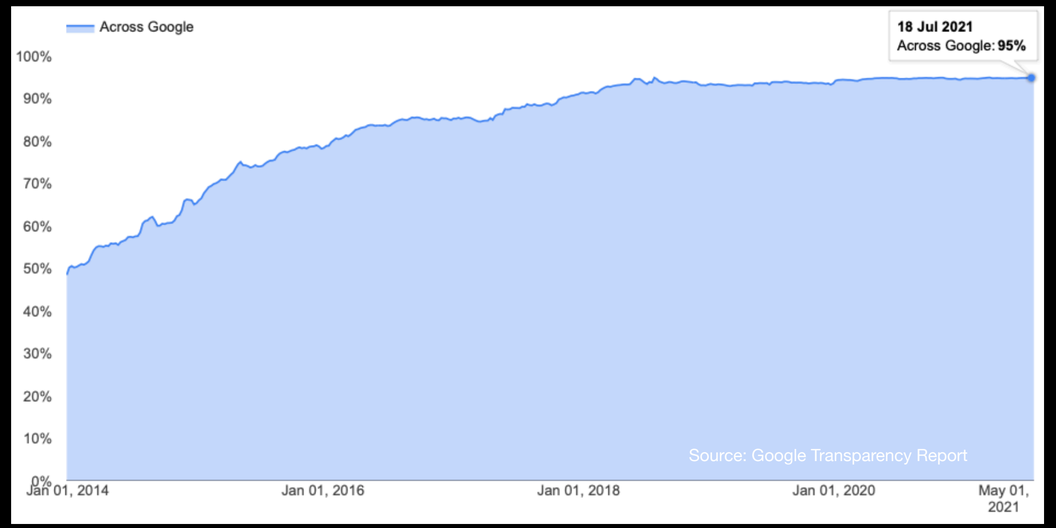
Encrypted traffic is the new standard

- **The majority of the traffic is encrypted.**
- **Enabling change: encrypted traffic is the new standard**



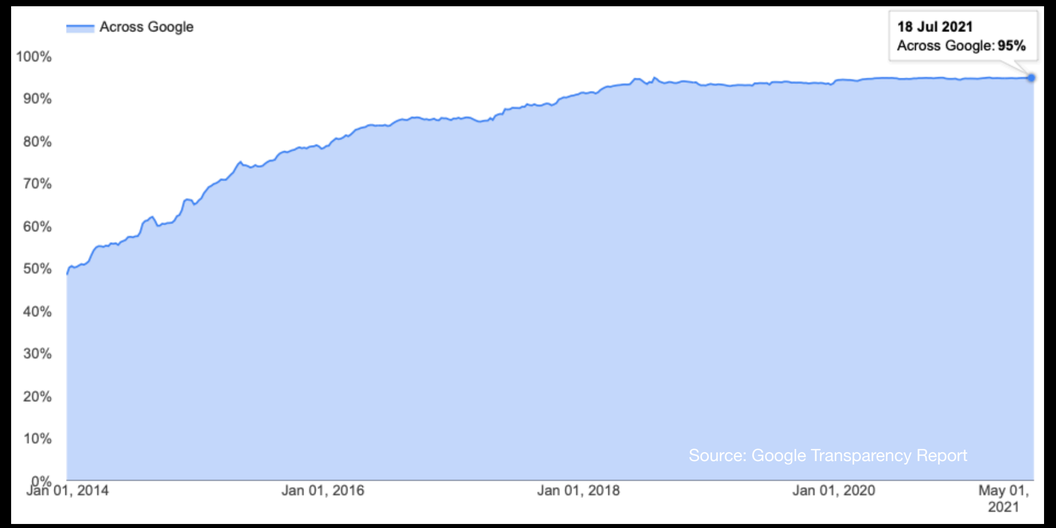
Encrypted traffic is the new standard

- The majority of the traffic is encrypted.
- Enabling change: encrypted traffic is the new standard
  - Using TLS certificates we can find the service owner.



Encrypted traffic is the new standard

- The majority of the traffic is encrypted.
- Enabling change: encrypted traffic is the new standard
  - Using TLS certificates we can find the service owner.
  - Corpuses of TLS data are publicly available.



Contributions

Contributions

- We developed the first technique capable of uncovering the off-net footprint deployment of all Hypergiants.

Contributions

- We developed the first technique capable of uncovering the off-net footprint deployment of all Hypergiants.
- We applied it to map their growth from 2013 to 2021.

Contributions

- We developed the first technique capable of uncovering the off-net footprint deployment of all Hypergiants.
- We applied it to map their growth from 2013 to 2021.
- We found that Google grew by 2,766 ASes, reaching 3,810 in April 2021.

Contributions

- We developed the first technique capable of uncovering the off-net footprint deployment of all Hypergiants.
- We applied it to map their growth from 2013 to 2021.
- We found that Google grew by 2,766 ASes, reaching 3,810 in April 2021.
- facebook and NETFLIX launched their own CDNs and now have presence in at least 2,115 ASes.

What this work is not about:

What this work is not about:

- **Not a head-to-head comparison** of different HGs as we do not know:

What this work is not about:

- **Not a head-to-head comparison** of different HGs as we do not know:
    1. Business strategies.

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.

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.

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.

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**.

Methodology

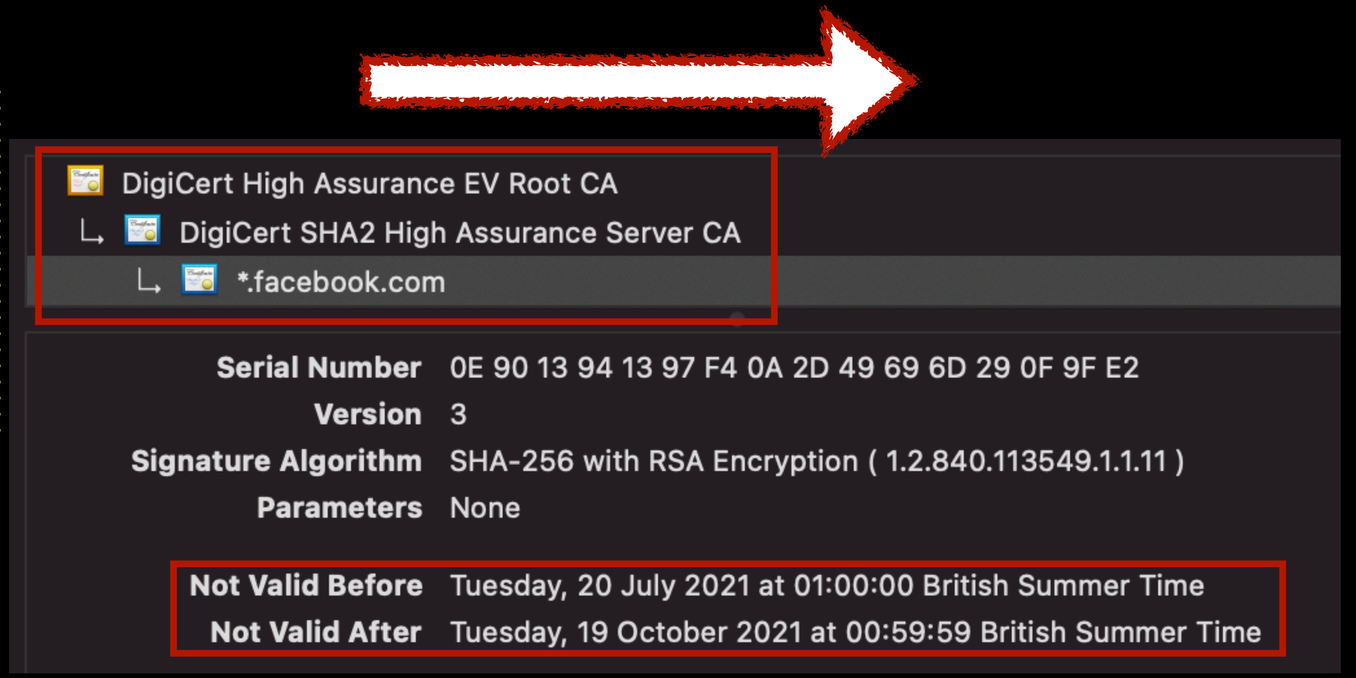
Methodology

- Step 0: **Collect TLS certificates dataset**

Methodology

- Step 0: **Collect TLS certificates dataset**
- Step 1: **Validate Certificates**
  - Exclude **self-signed**, **expired** and certificates with a **non-verified chain**.

TLS Certificates Dataset



Valid TLS Certificates

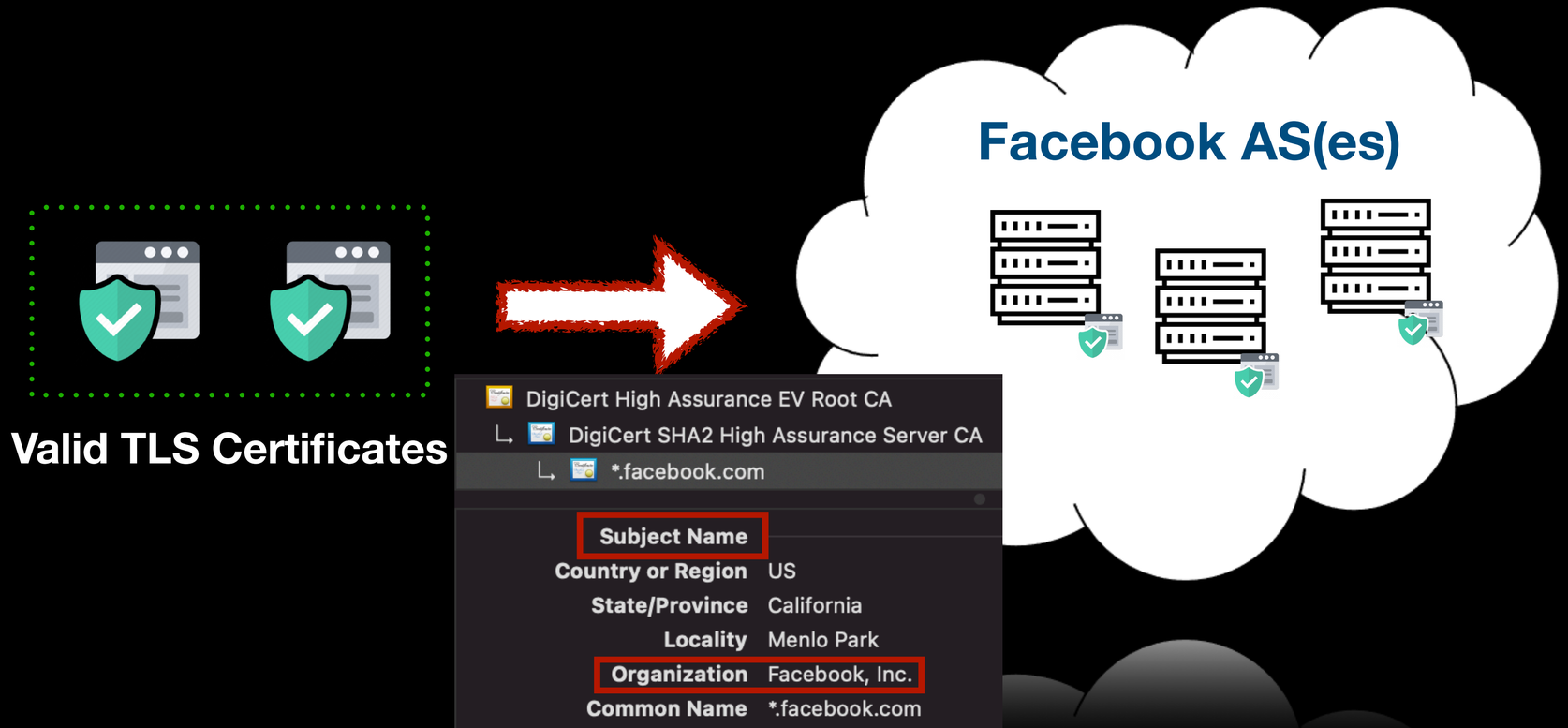
Methodology

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

Valid TLS Certificates

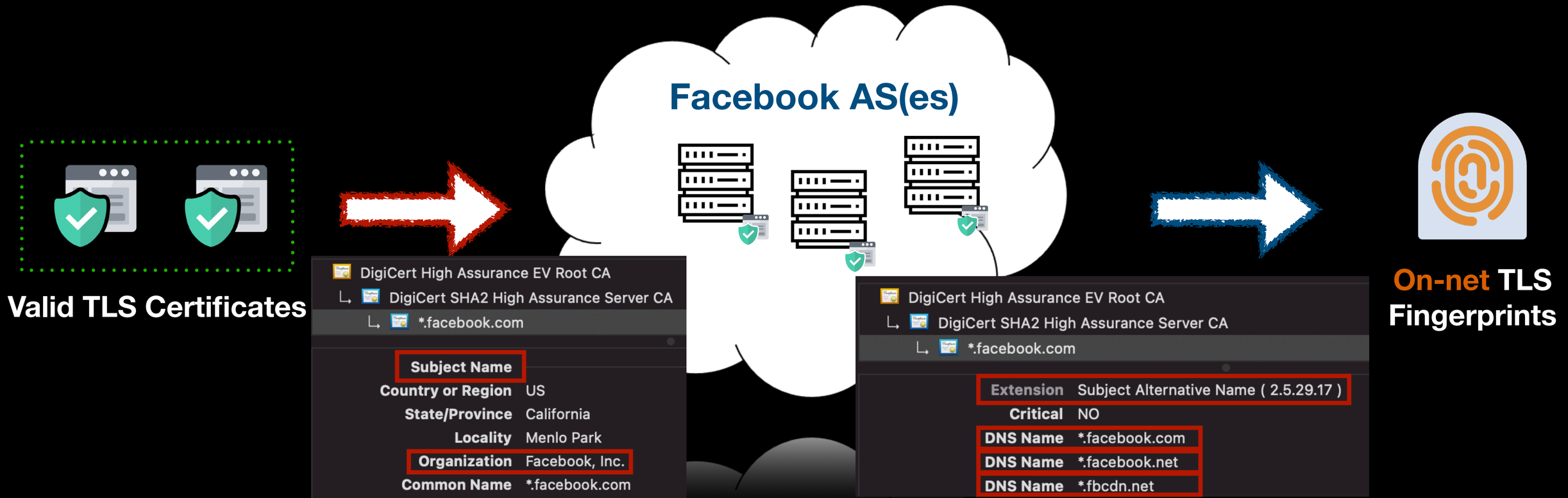
Methodology

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



Methodology

- Step 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.

On-net TLS Fingerprints

Methodology

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

Subject Name	
Country or Region	US
State/Province	California
Locality	Menlo Park
Organization	Facebook, Inc.
Common Name	*.facebook.com

On-net TLS Fingerprints

DNS Name	*.facebook.com
DNS Name	*.facebook.net
DNS Name	*.fbcdn.net
DNS Name	*.fbsbx.com

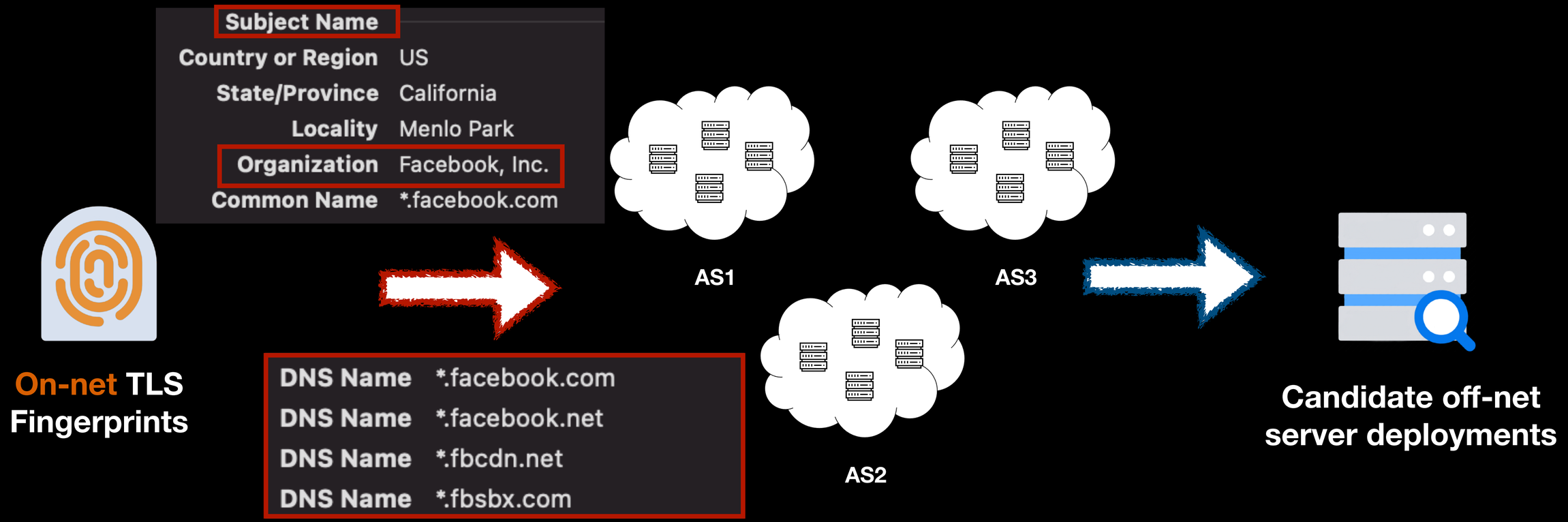
AS1

AS3

AS2

Methodology

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



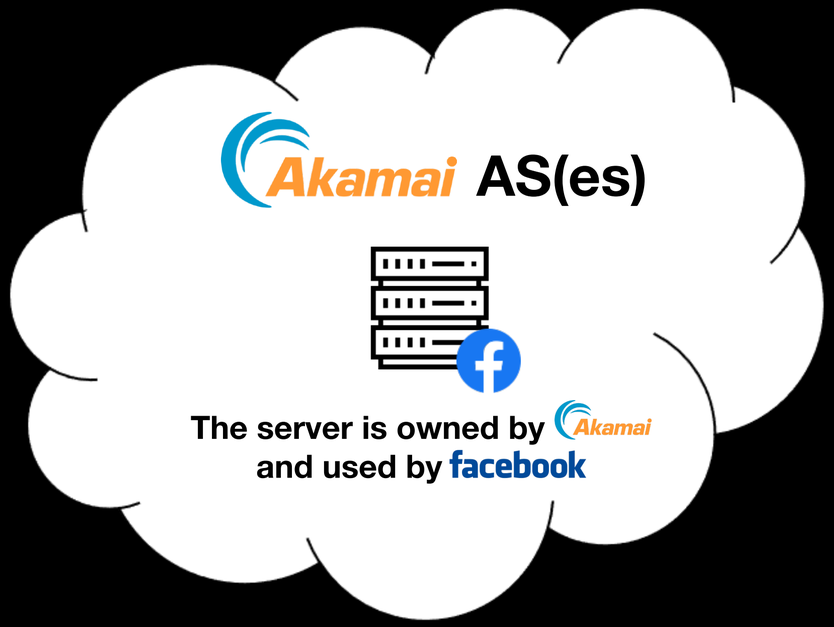
Challenges

Challenges

- Does presence of a Hypergiant certificate on a server outside their own network guarantee that the server is an off-net server deployment?
- **No, it does not!**

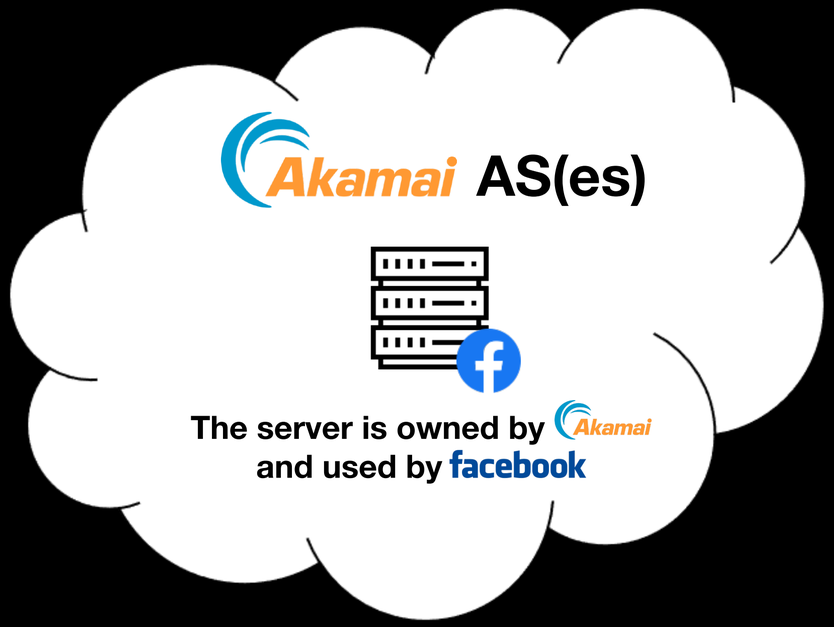
Challenges

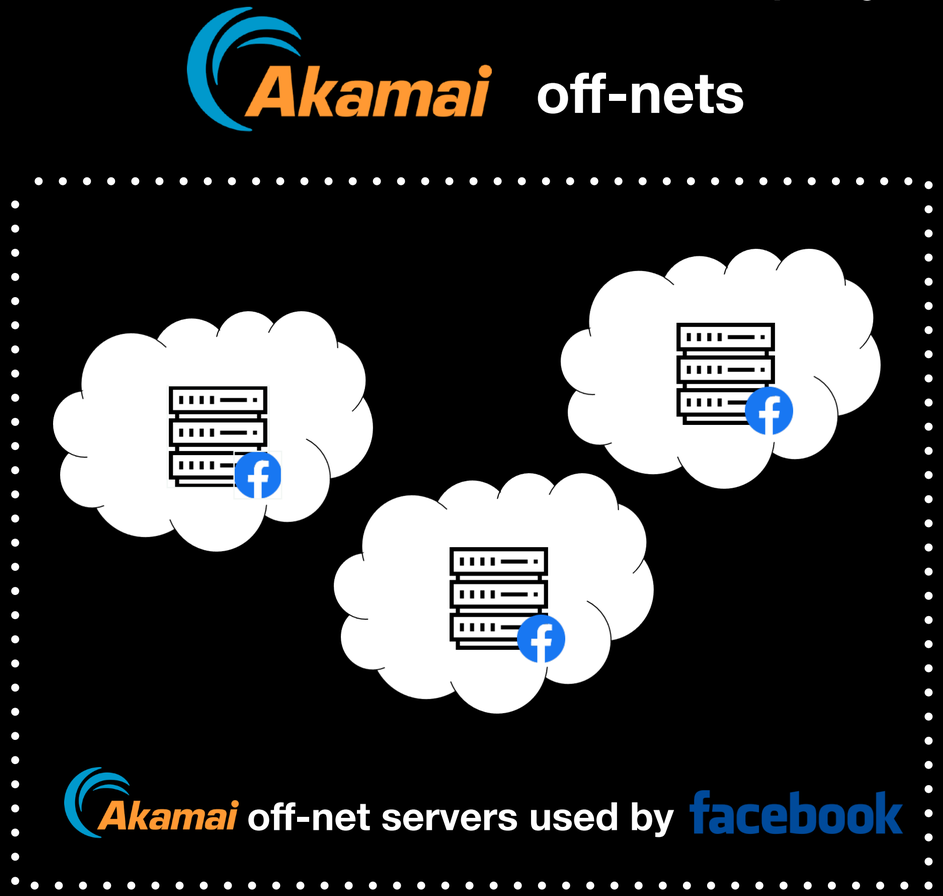
- Does presence of a Hypergiant certificate on a server outside their own network guarantee that the server is an off-net server deployment?
- **No, it does not!**



Challenges

- Does presence of a Hypergiant certificate on a server outside their own network guarantee that the server is an off-net server deployment?
- **No, it does not!**



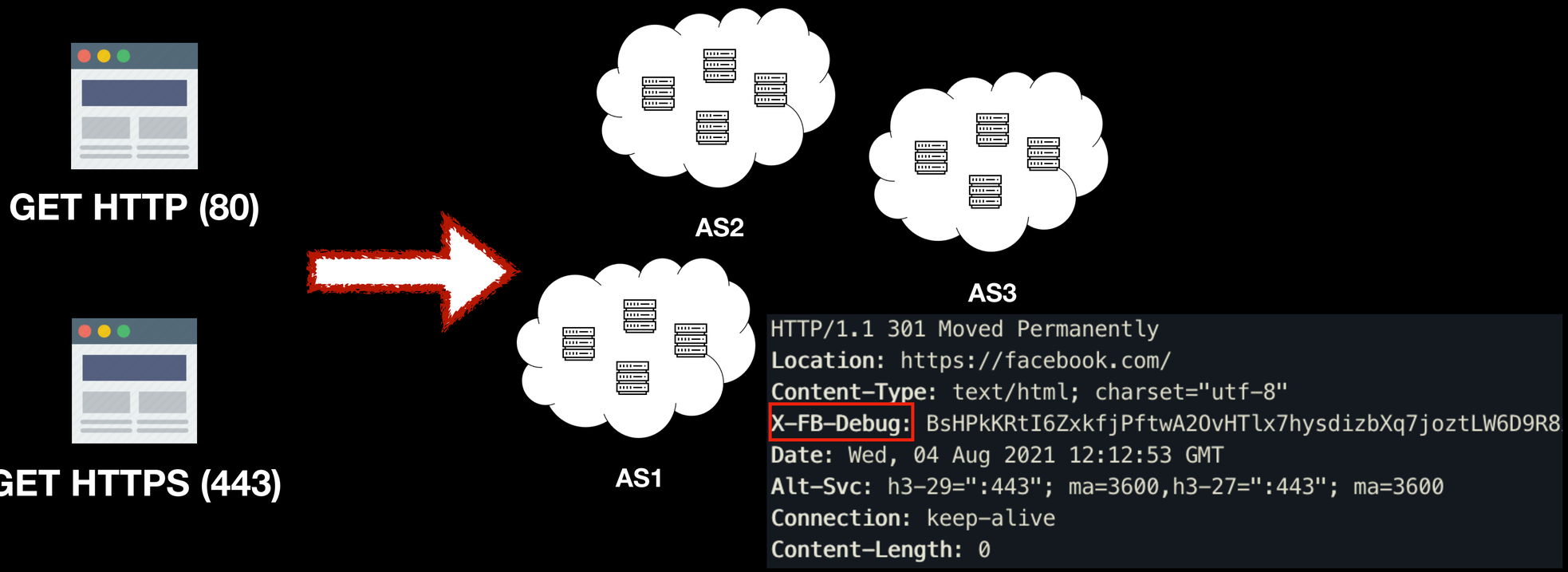


Methodology

- Step 4: **Learn Hypergiant HTTP(S) Fingerprints using headers**

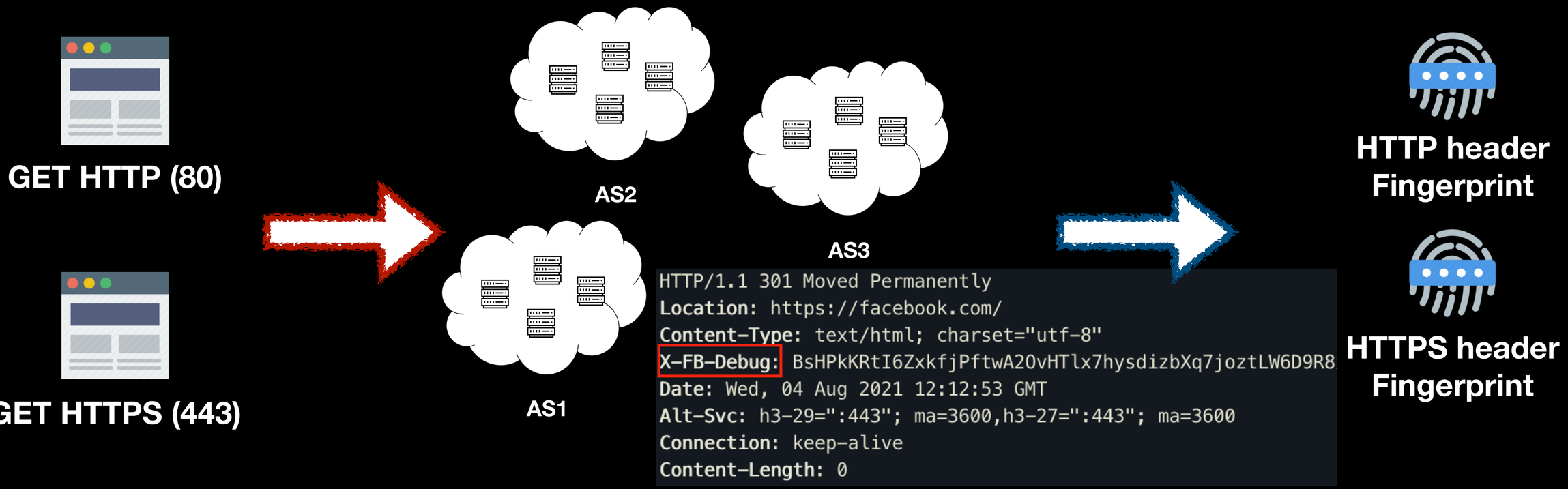
Methodology

- Step 4: **Learn Hypergiant HTTP(S) Fingerprints using headers**



Methodology

- Step 4: **Learn Hypergiant HTTP(S) Fingerprints using headers**



Methodology

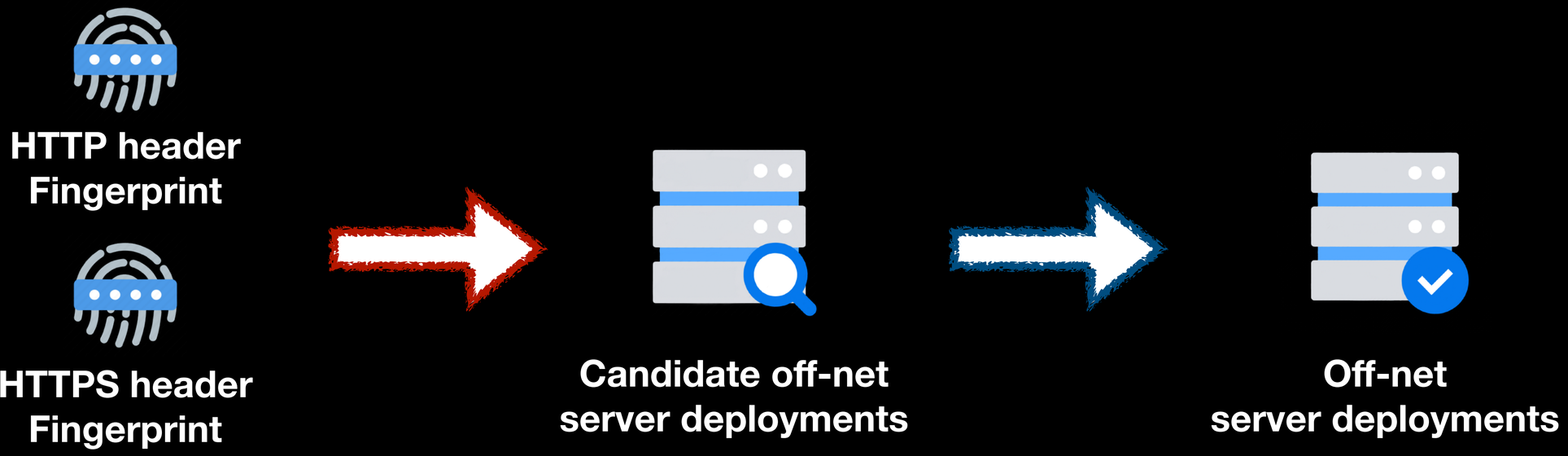
- Step 5: **Confirm Candidates Using HTTP(S)**
  - Apply the HTTP(S) fingerprints to the off-net candidates and classify as off-nets any that match the HG fingerprints.

HTTP header Fingerprint

HTTPS header Fingerprint

Methodology

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



Technique Outline

Technique Outline

- The **TLS certificate** reveals if an IP hosts a service for the Hypergiant.
- **HTTP(S) header** reveals who operates the server.
- The **IP address** reveals if it is an on-net or off-net server.

Technique Outline

- The **TLS certificate** reveals if an IP hosts a service for the Hypergiant.
- **HTTP(S) header** reveals who operates the server.
- The **IP address** reveals if it is an on-net or off-net server.

- To consider a server as a Hypergiant off-net deployment:

Technique Outline

- The **TLS certificate** reveals if an IP hosts a service for the Hypergiant.
- **HTTP(S) header** reveals who operates the server.
- The **IP address** reveals if it is an on-net or off-net server.

- To consider a server as a Hypergiant off-net deployment:
  - **TLS certificate** and **HTTP(S) headers** must map to the Hypergiant.

Technique Outline

- The **TLS certificate** reveals if an IP hosts a service for the Hypergiant.
- **HTTP(S) header** reveals who operates the server.
- The **IP address** reveals if it is an on-net or off-net server.

- To consider a server as a Hypergiant off-net deployment:
  - **TLS certificate** and **HTTP(S) headers** must map to the Hypergiant.
  - The IP address is not part of the Hypergiant own network.

Datasets

Datasets

- TLS certificate scans:
  -  collects certificates in IPv4-wide scans on port 443.

Datasets

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

Datasets

- TLS certificate scans:
  - 

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

+ Custom active scan.

Datasets

- TLS certificate scans:
  - 

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

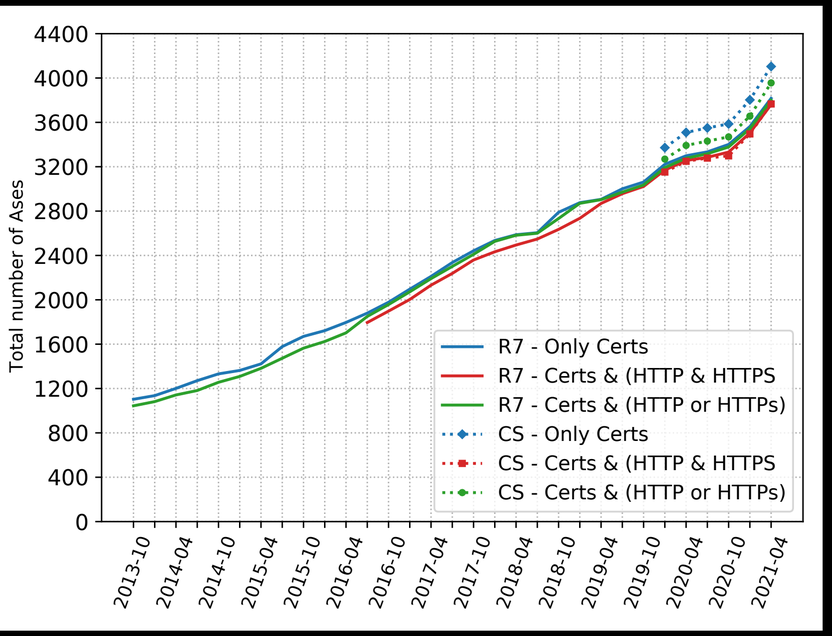
+ Custom active scan.
- HTTP(S) headers (Validation):
  - We used corpuses of available HTTP(S) headers from 

from Oct. 2013 to Apr. 2021.

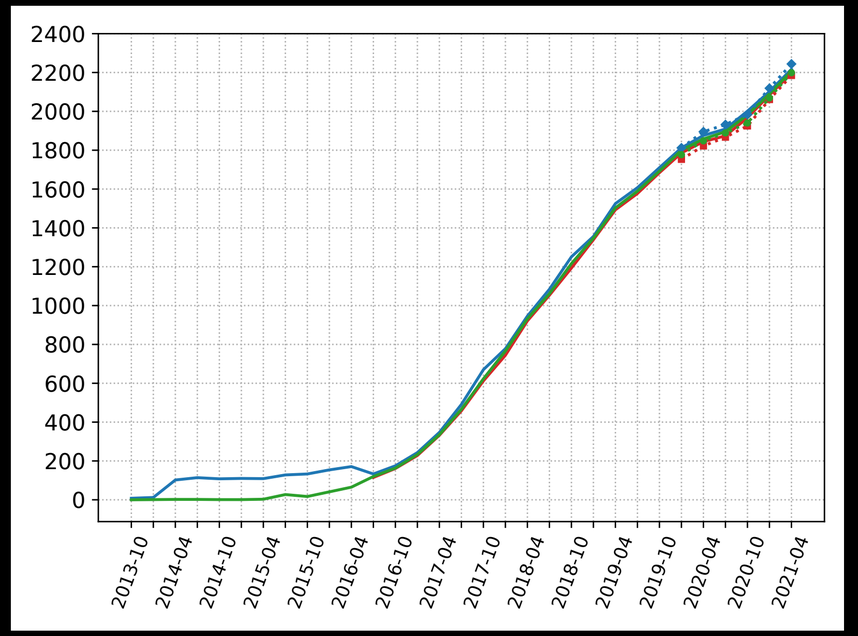
Dataset and HTTP(S) header validation

Dataset and HTTP(S) header validation

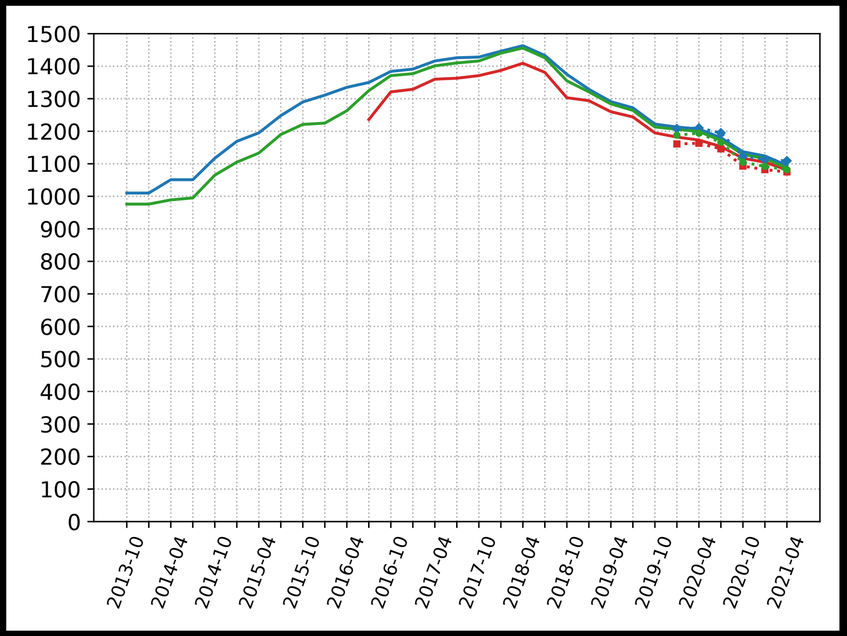
Google



facebook

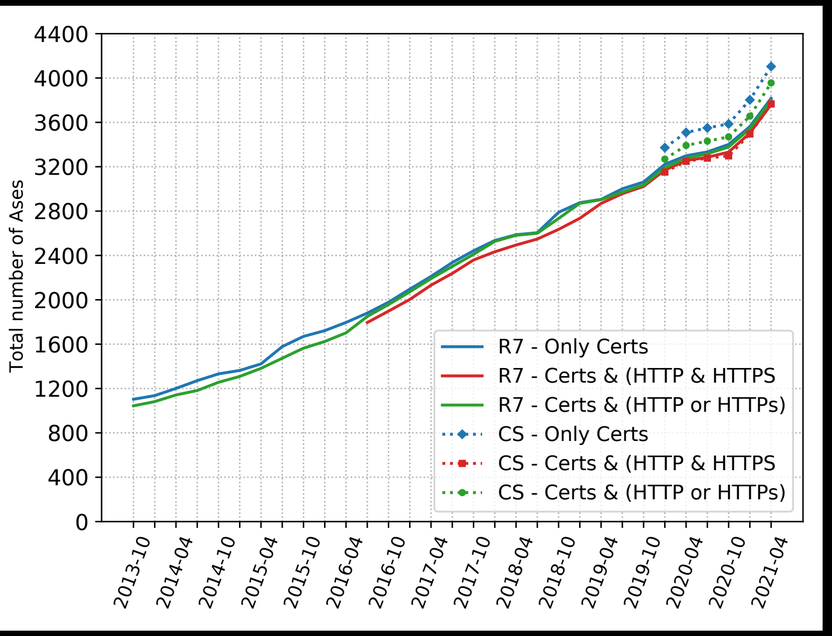


Akamai

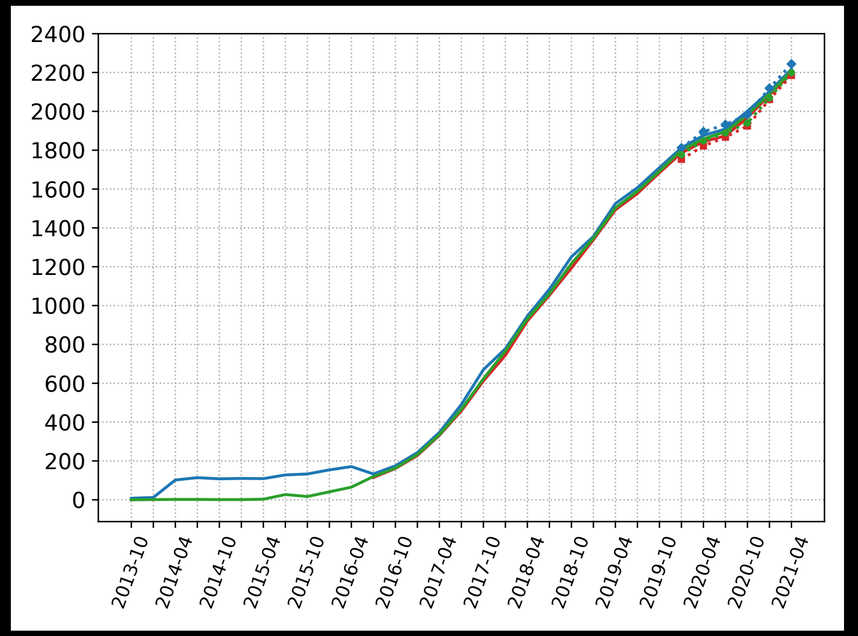


Dataset and HTTP(S) header validation

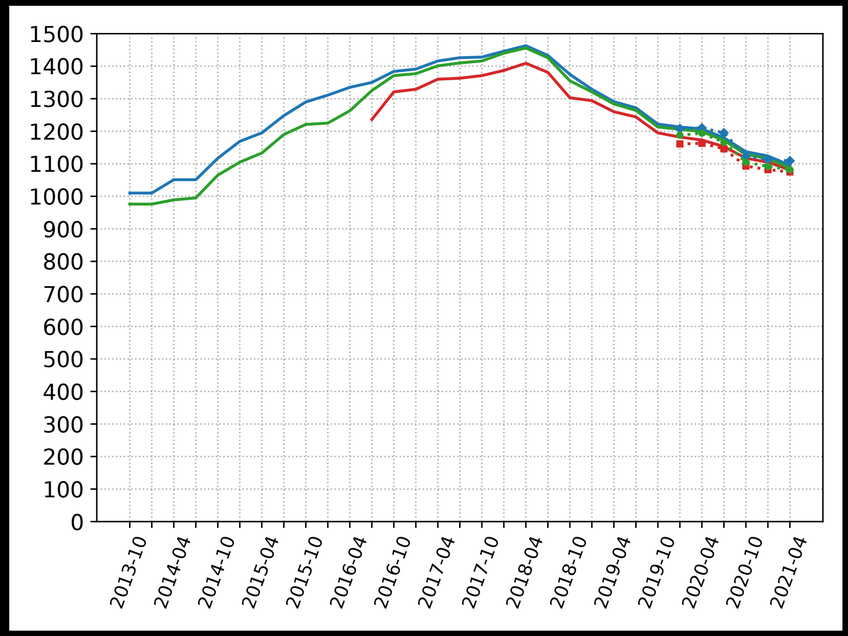
Google



facebook



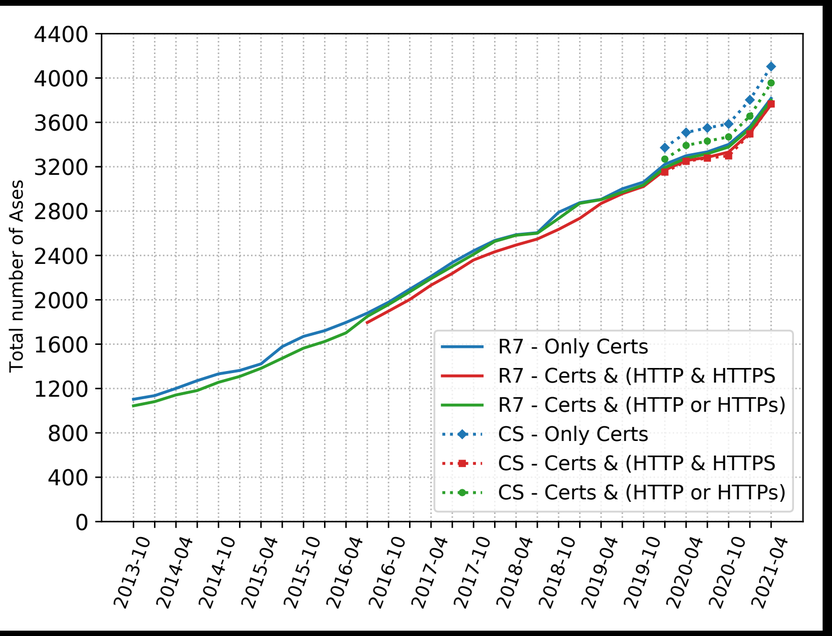
Akamai



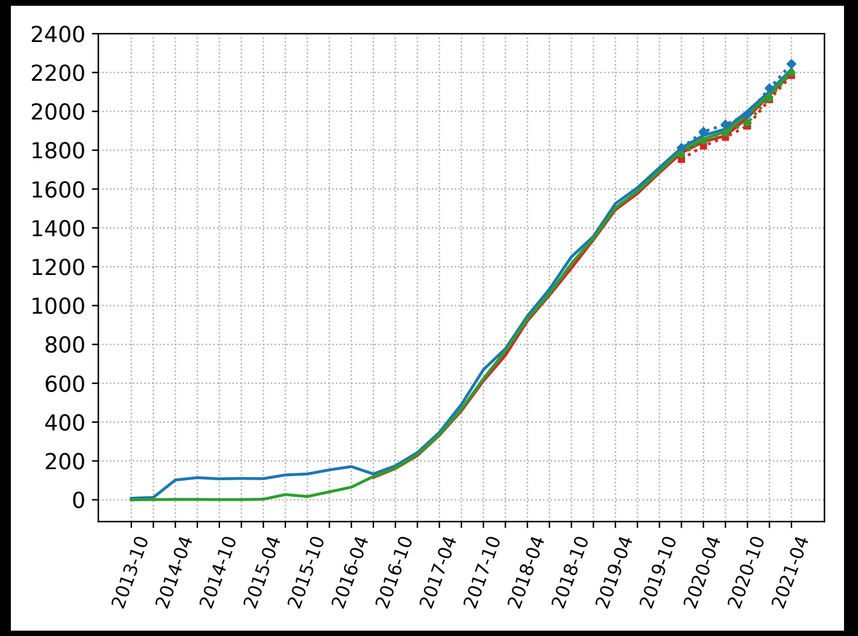
- Differences between **only-certificates** and **&/or HTTP(S)** are minimal.

Dataset and HTTP(S) header validation

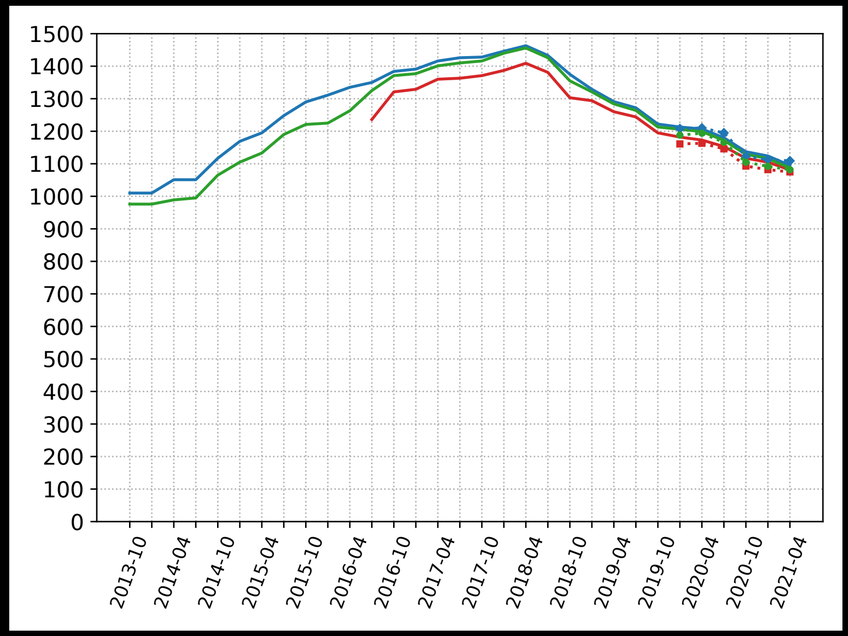
Google



facebook



Akamai



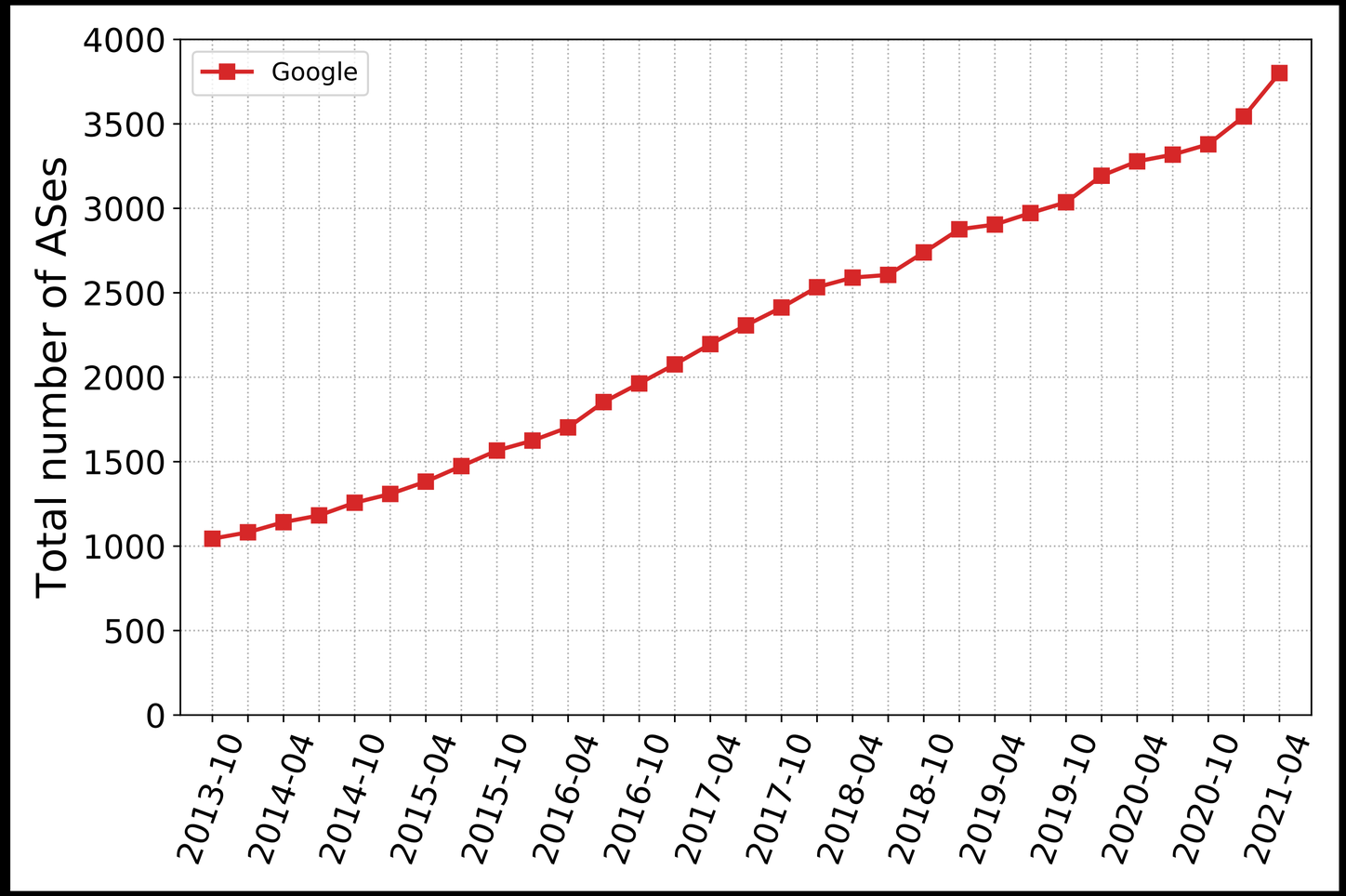
- Differences between **only-certificates** and **&/or HTTP(S)** are minimal.
- For Google using Censys we are able to identify more ASes.

Hypergiants’ off-nets Expansion

Longitudinal Growth (2013-2021)

Hypergiants' off-nets Expansion

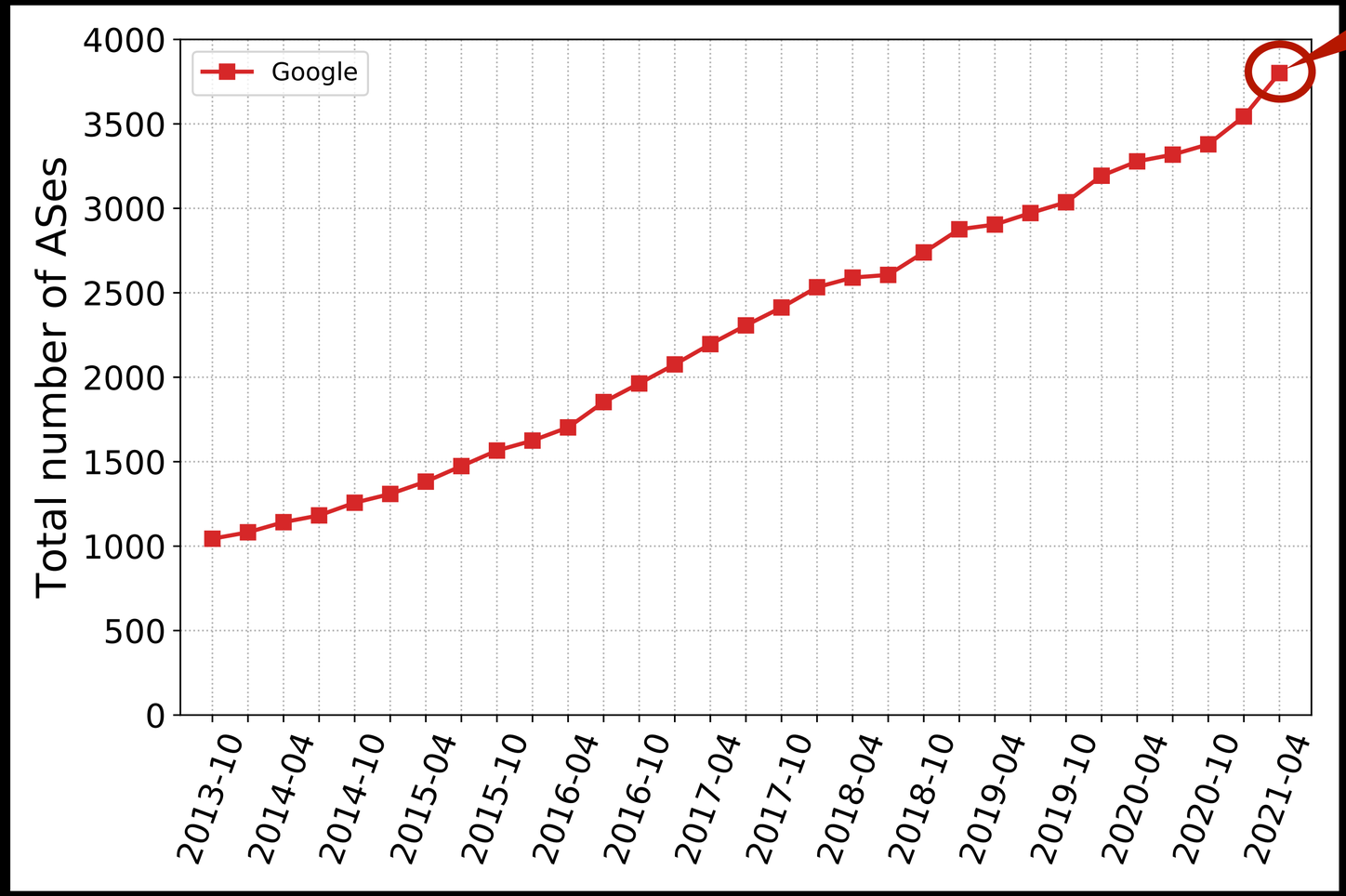
Longitudinal Growth (2013-2021)



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

Hypergiants' off-nets Expansion

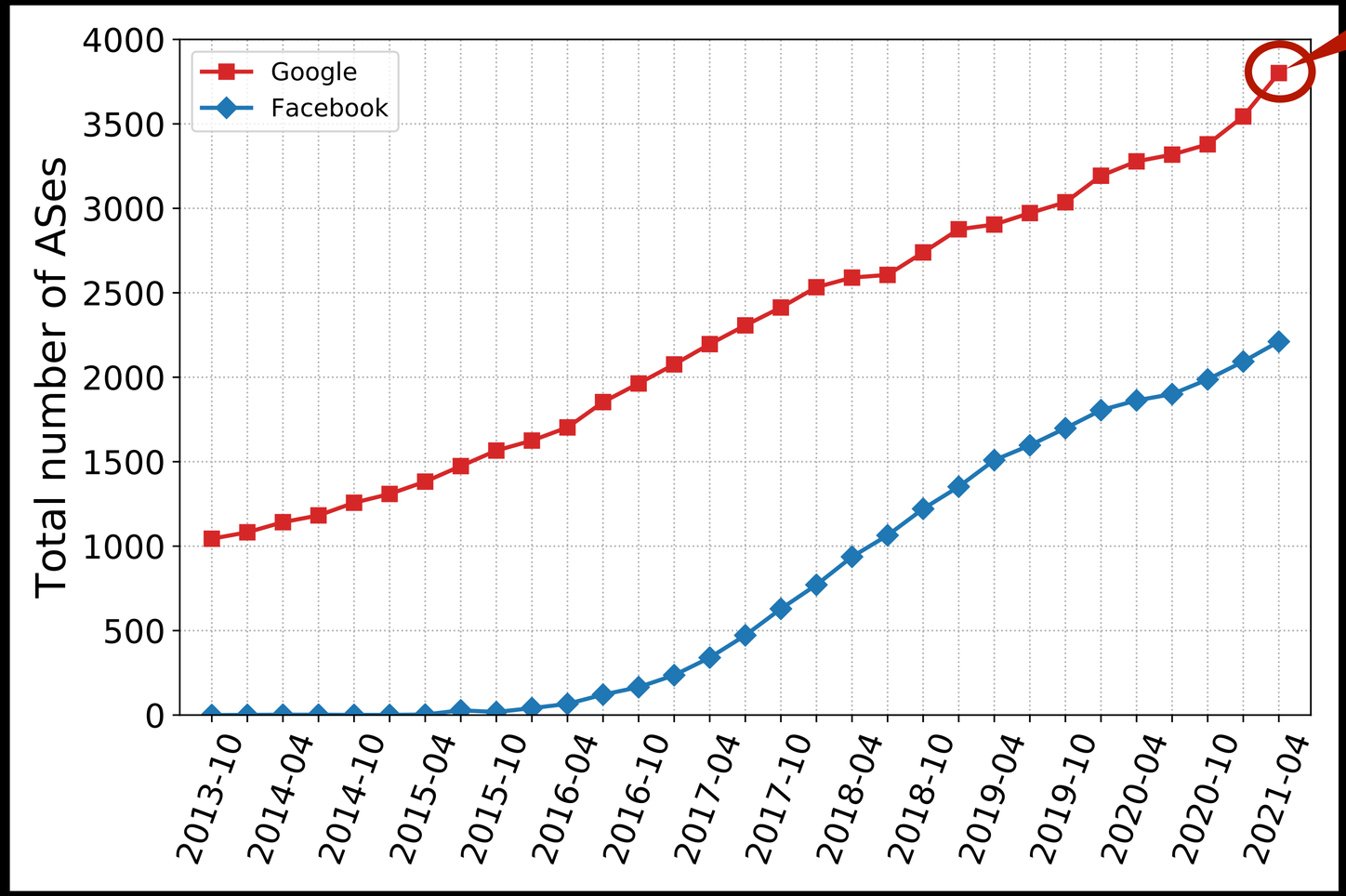
Longitudinal Growth (2013-2021)



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

Hypergiants' off-nets Expansion

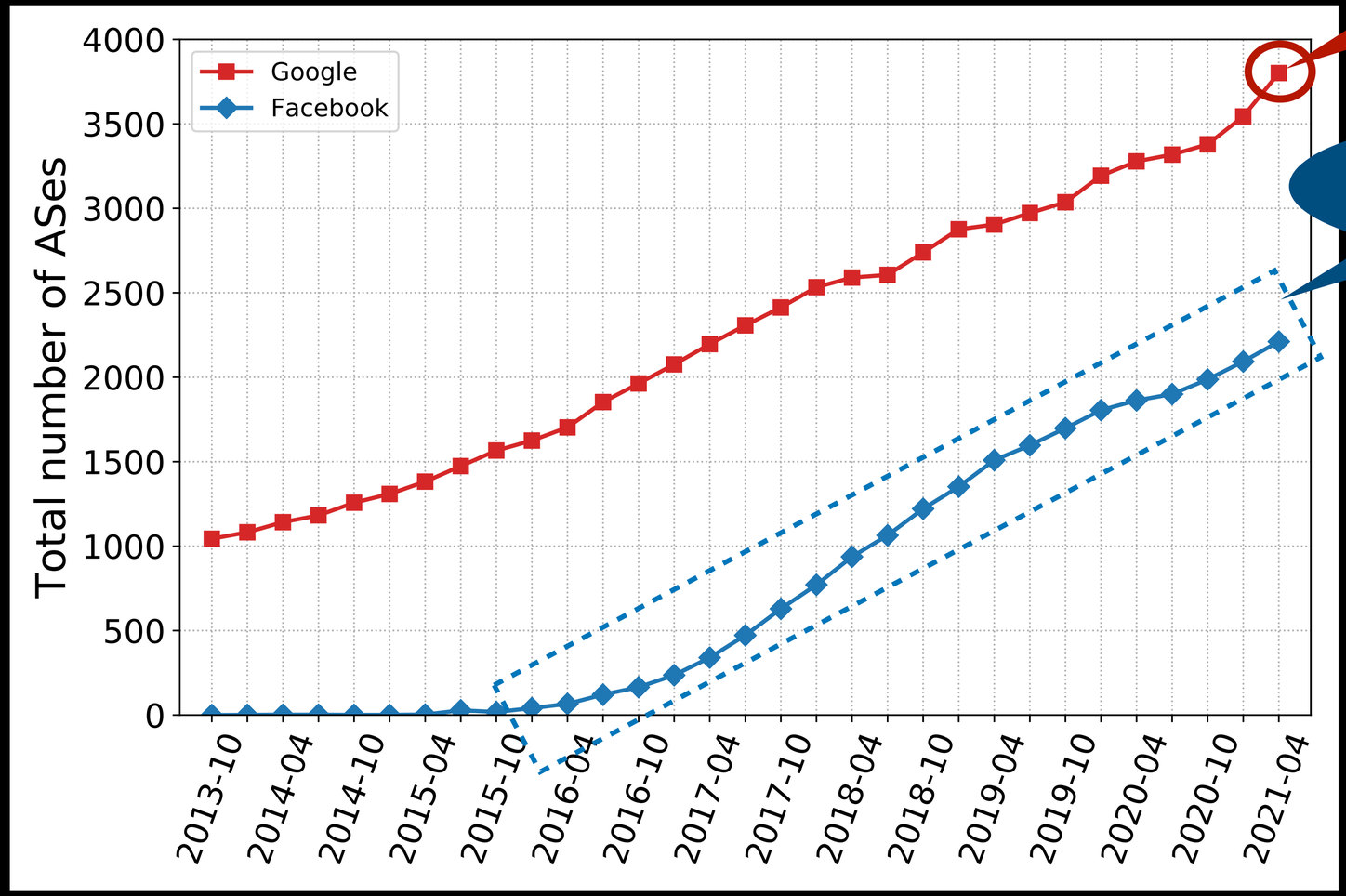
Longitudinal Growth (2013-2021)



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

Hypergiants' off-nets Expansion

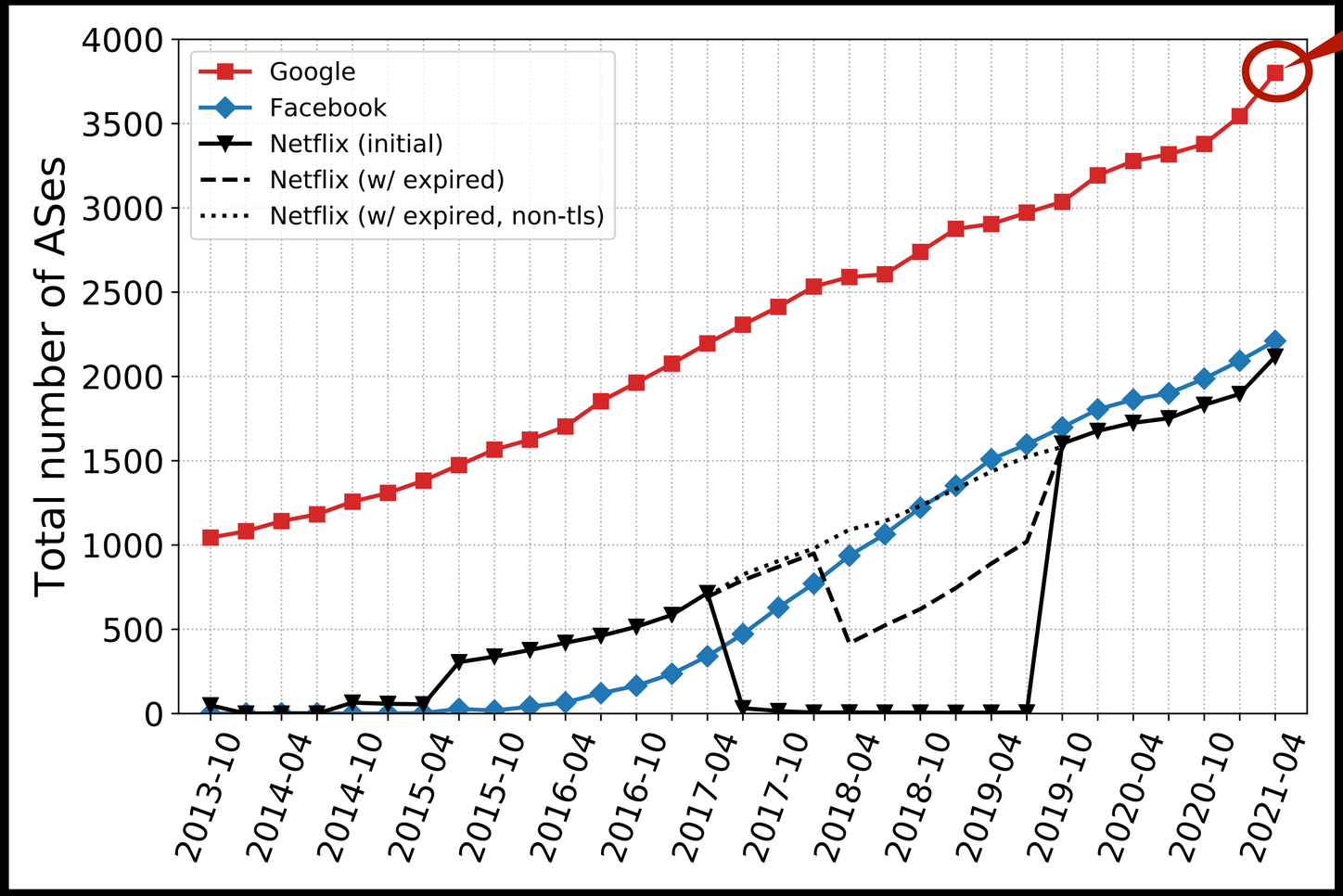
Longitudinal Growth (2013-2021)



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

Hypergiants' off-nets Expansion

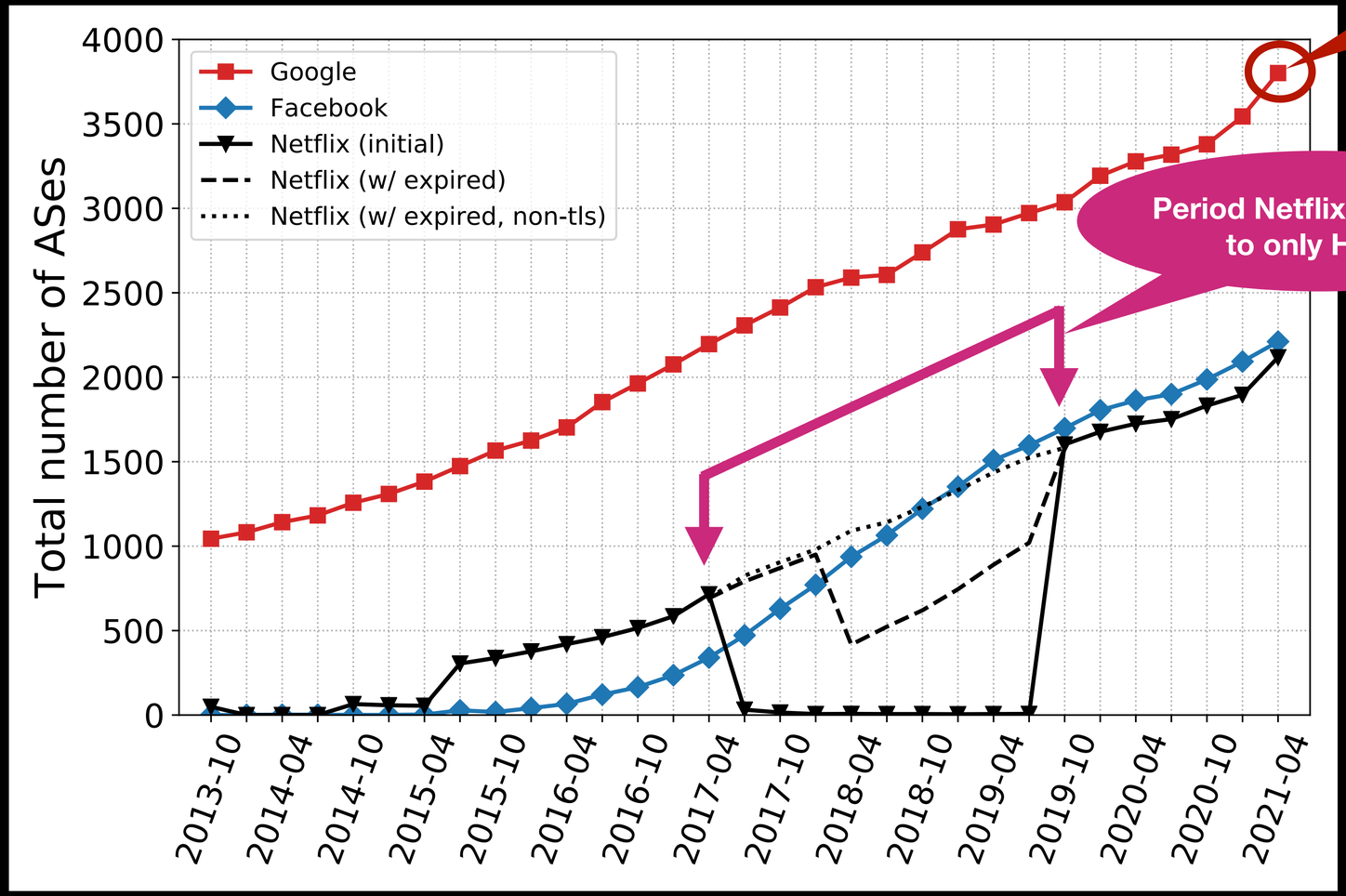
Longitudinal Growth (2013-2021)



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

Hypergiants' off-nets Expansion

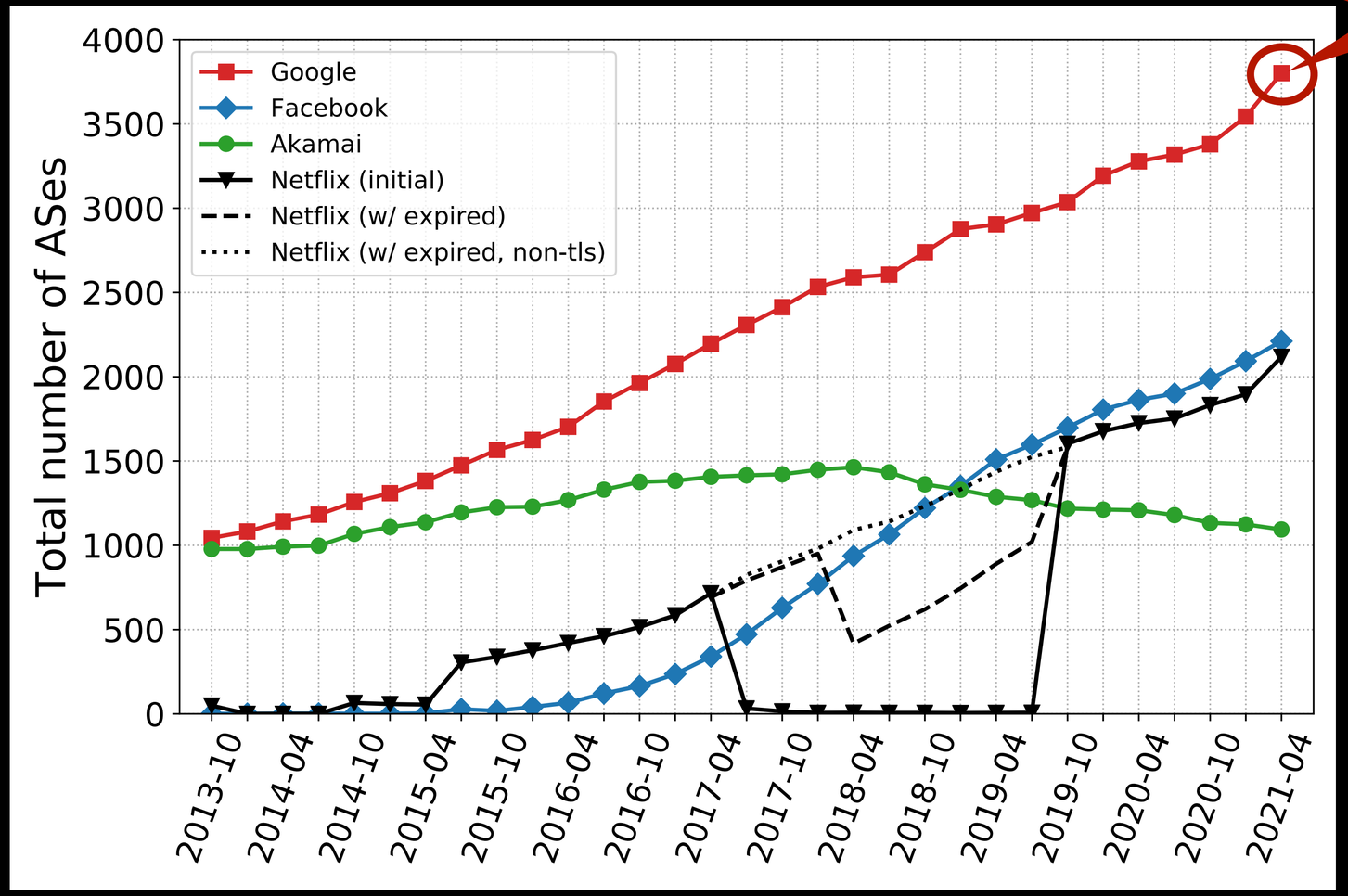
Longitudinal Growth (2013-2021)



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

Hypergiants' off-nets Expansion

Longitudinal Growth (2013-2021)



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

Growth by Network Type

Understanding the “demographics” (Mode I)

Growth by Network Type

Understanding the “demographics” (Mode I)

- We label the ASes hosting off-nets based on their customer cone size*.

* Extracted from the CAIDA AS Relationships Dataset.

Growth by Network Type

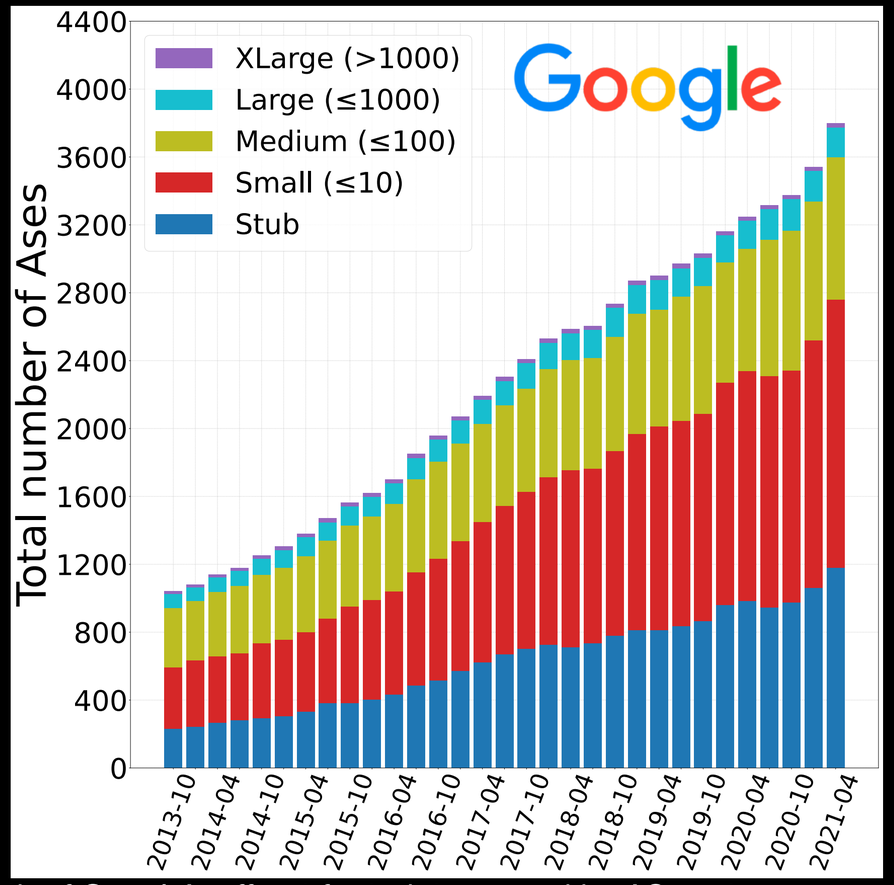
Understanding the “demographics” (Mode I)

- We label the ASes hosting off-nets based on their customer cone size*.
- We consider 5 categories of ASes:

* Extracted from the CAIDA AS Relationships Dataset.

Growth by Network Type

Understanding the “demographics” (Mode I)

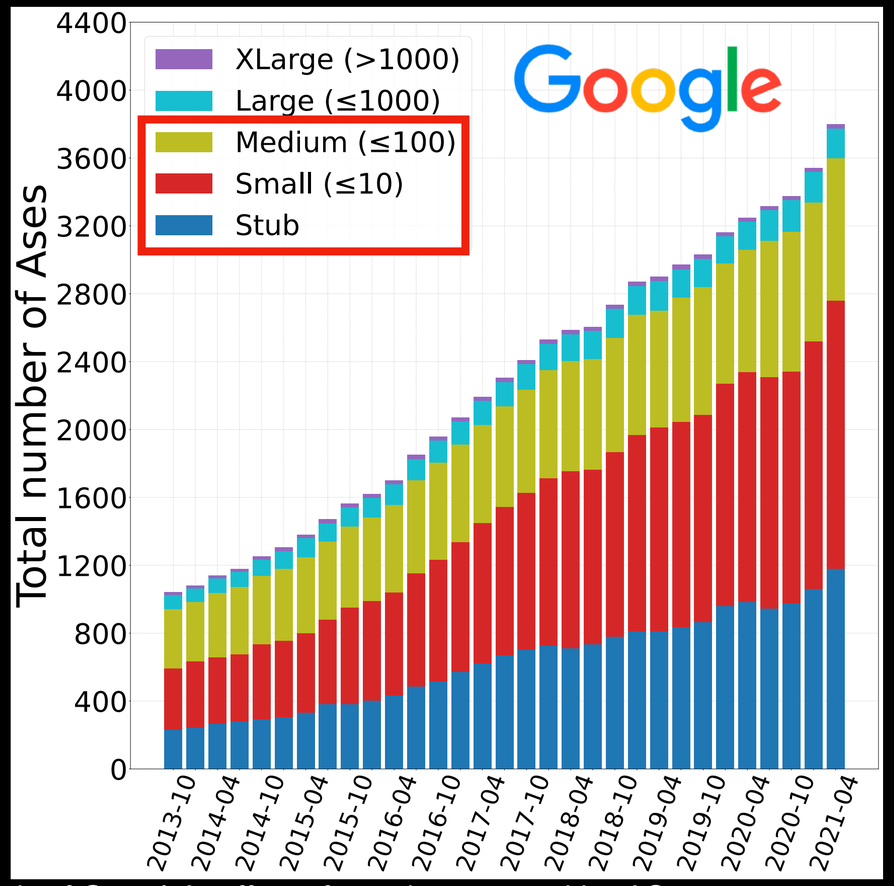


Growth of Google’s off-net footprint grouped by AS customer cone size.

* Extracted from the CAIDA AS Relationships Dataset.

Growth by Network Type

Understanding the “demographics” (Mode I)



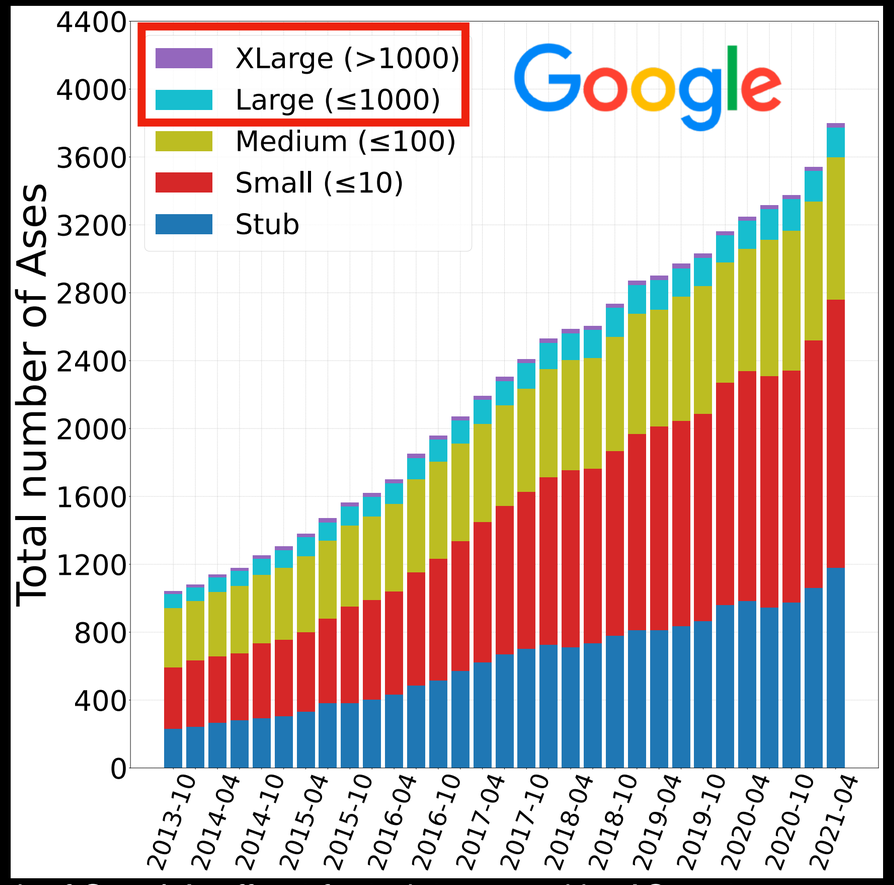
Growth of Google’s off-net footprint grouped by AS customer cone size.

- **4x** increase in Stub, Small and, Medium ASes.

* Extracted from the CAIDA AS Relationships Dataset.

Growth by Network Type

Understanding the “demographics” (Mode I)



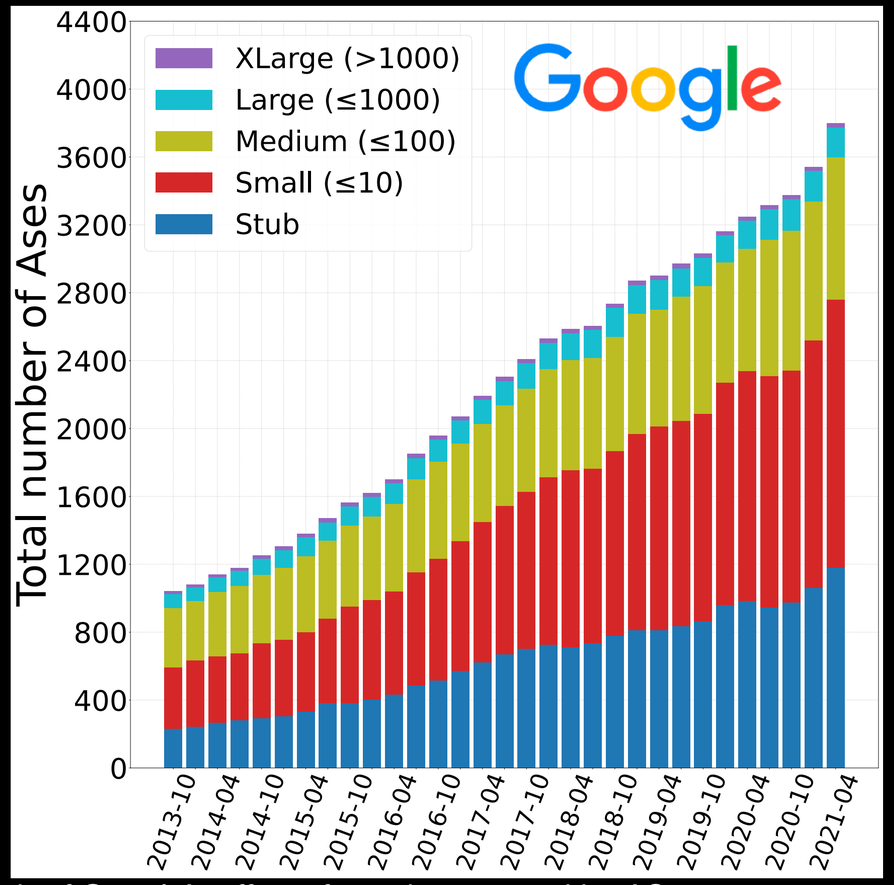
Growth of Google’s off-net footprint grouped by AS customer cone size.

- **4x** increase in Stub, Small and, Medium ASes.
- **2x** increase in Large and XLarge ASes.

* Extracted from the CAIDA AS Relationships Dataset.

Growth by Network Type

Understanding the “demographics” (Mode I)



Growth of Google’s off-net footprint grouped by AS customer cone size.

- **4x** increase in Stub, Small and, Medium ASes.
- **2x** increase in Large and XLarge ASes.
- Growth significantly increase after the open of the economy.

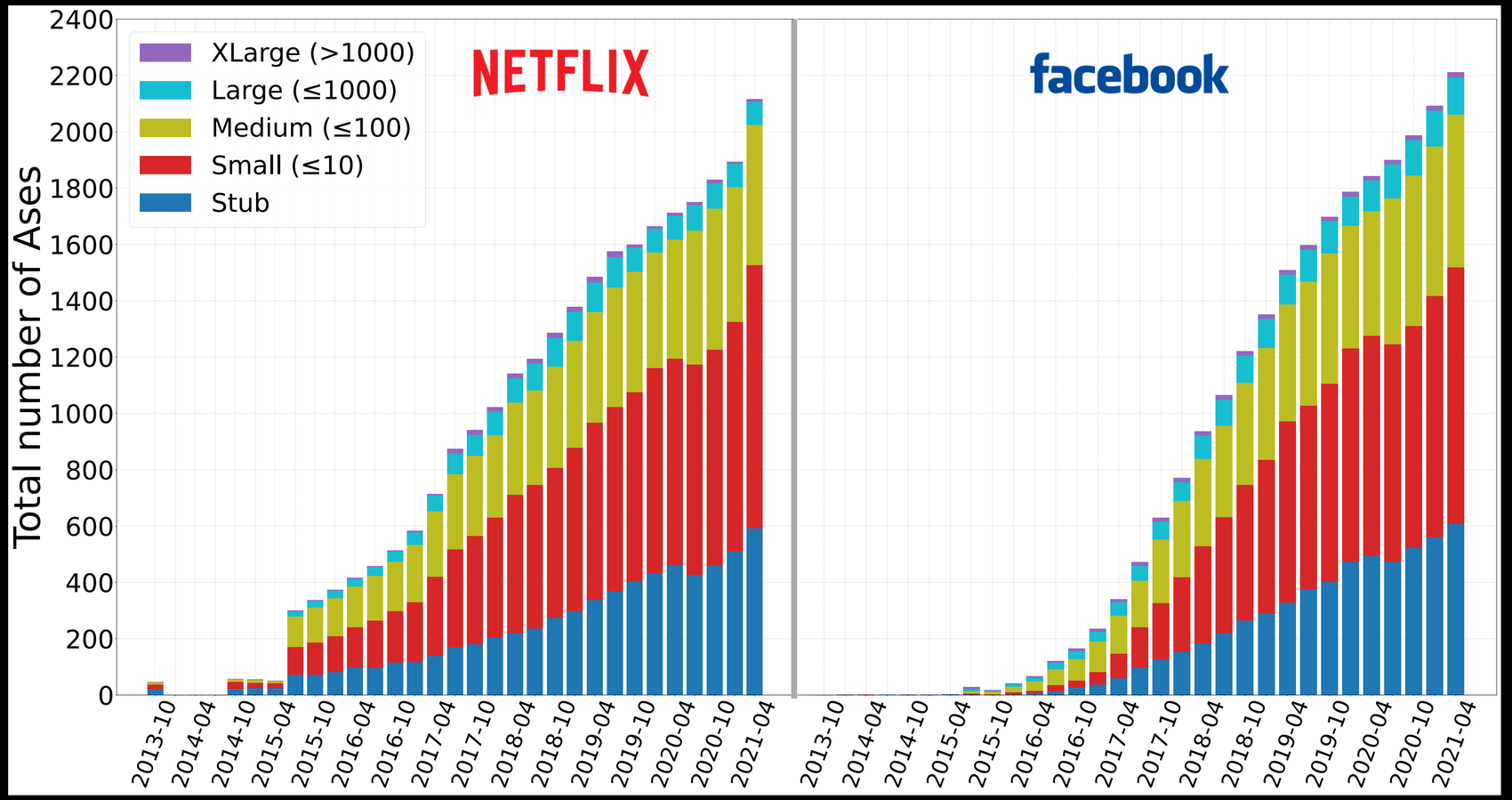
* Extracted from the CAIDA AS Relationships Dataset.

Growth by Network Type

Understanding the “demographics” (Mode II)

Growth by Network Type

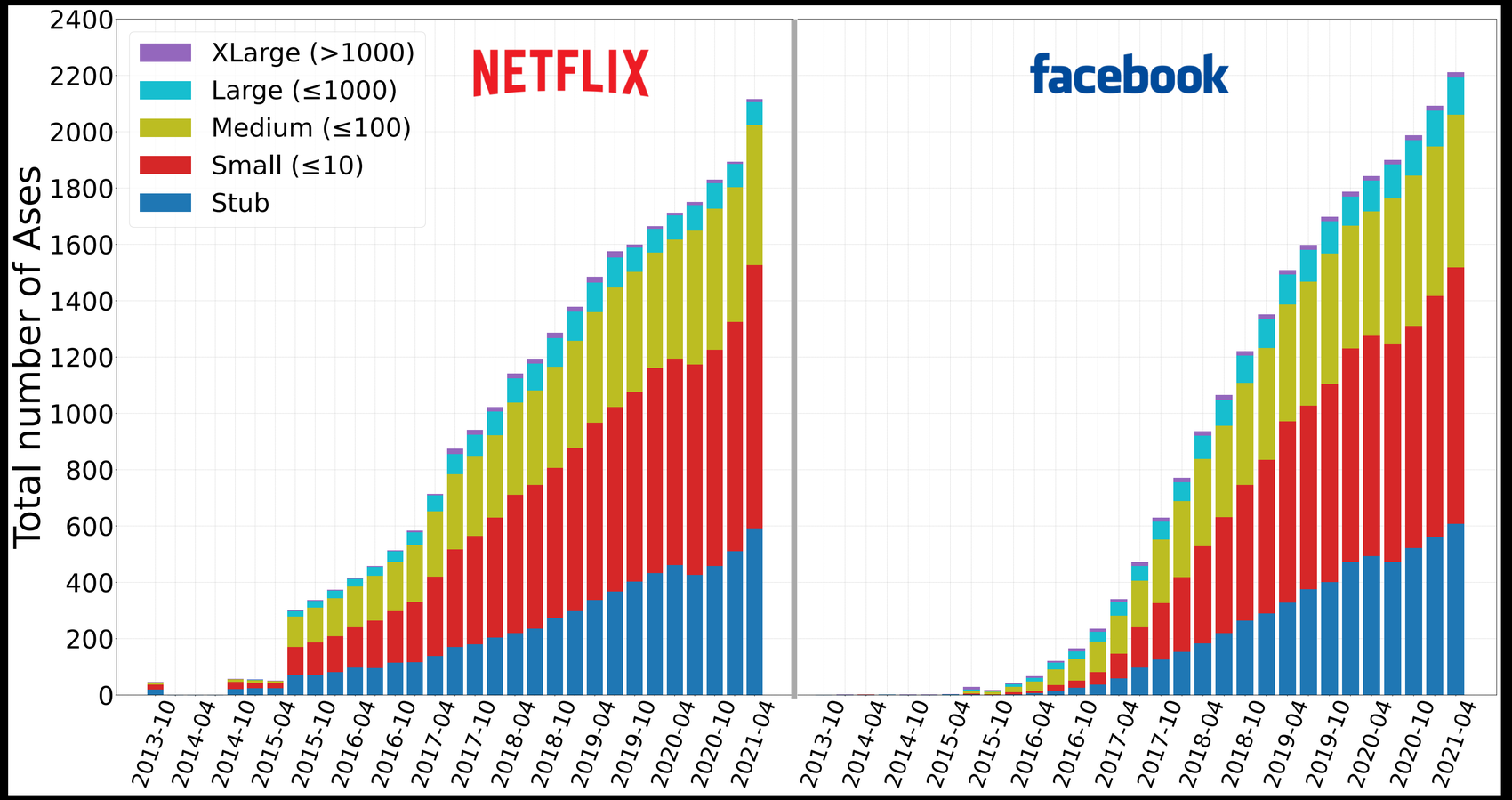
Understanding the “demographics” (Mode II)



Growth of Netflix and Facebook off-net footprints grouped by AS customer cone size.

Growth by Network Type

Understanding the “demographics” (Mode II)

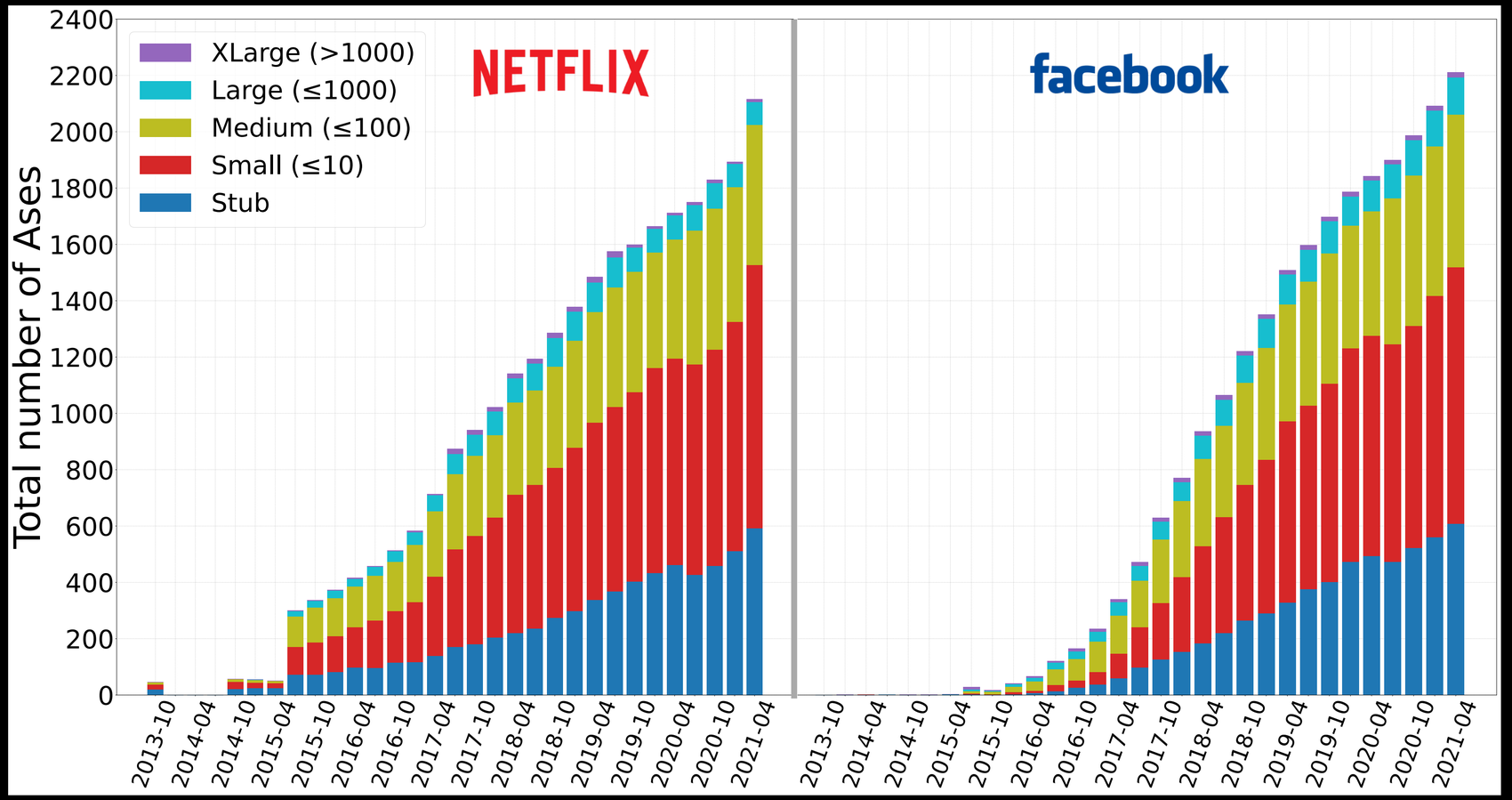


Growth of Netflix and Facebook off-net footprints grouped by AS customer cone size.

- Birth of HG CDNs.

Growth by Network Type

Understanding the “demographics” (Mode II)

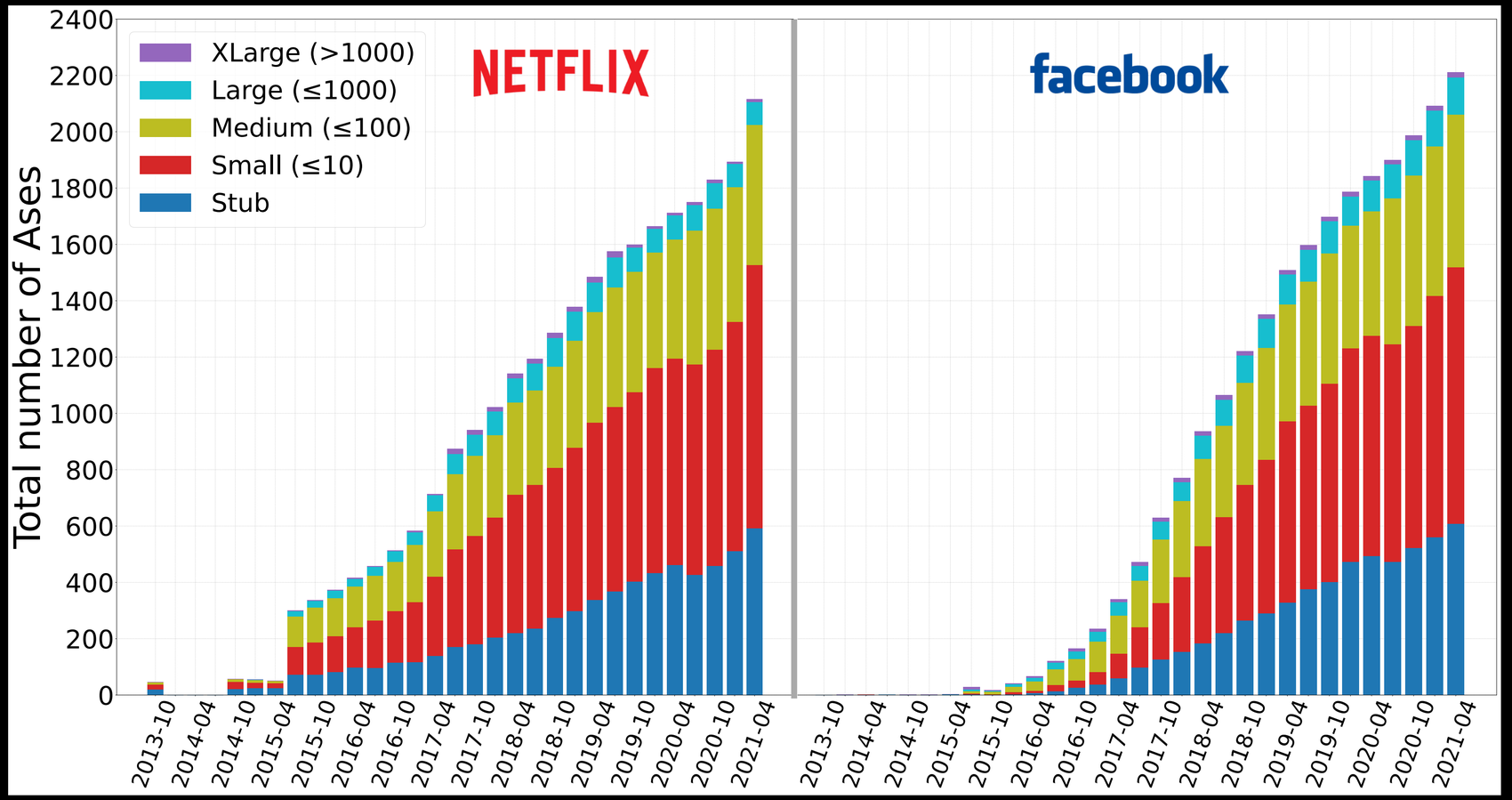


Growth of Netflix and Facebook off-net footprints grouped by AS customer cone size.

- Birth of HG CDNs.
- More aggressive increase 10x.

Growth by Network Type

Understanding the “demographics” (Mode II)

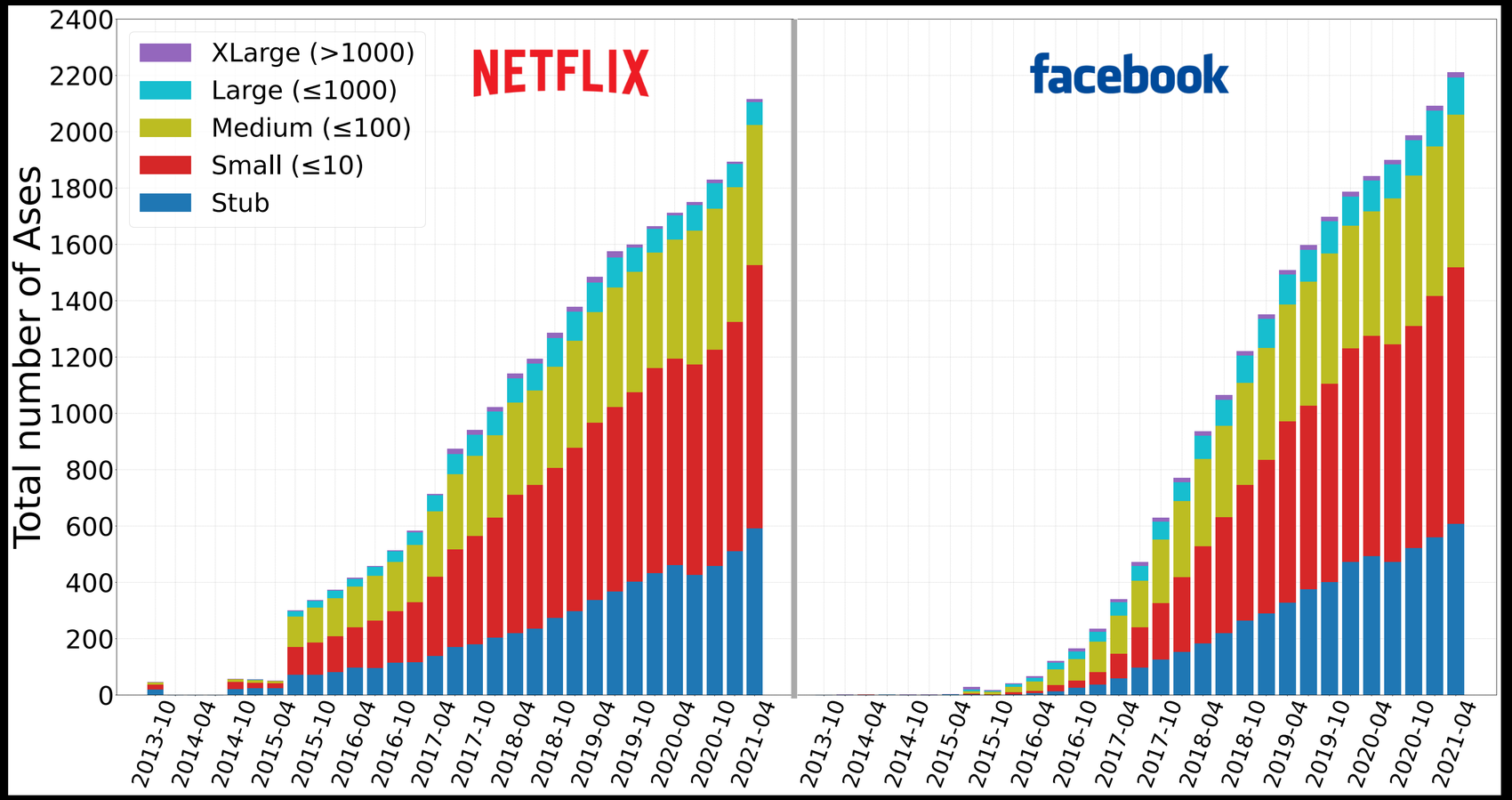


Growth of Netflix and Facebook off-net footprints grouped by AS customer cone size.

- Birth of HG CDNs.
- More aggressive increase 10x.
- Similar contributions of different types of networks.

Growth by Network Type

Understanding the “demographics” (Mode II)



Growth of Netflix and Facebook off-net footprints grouped by AS customer cone size.

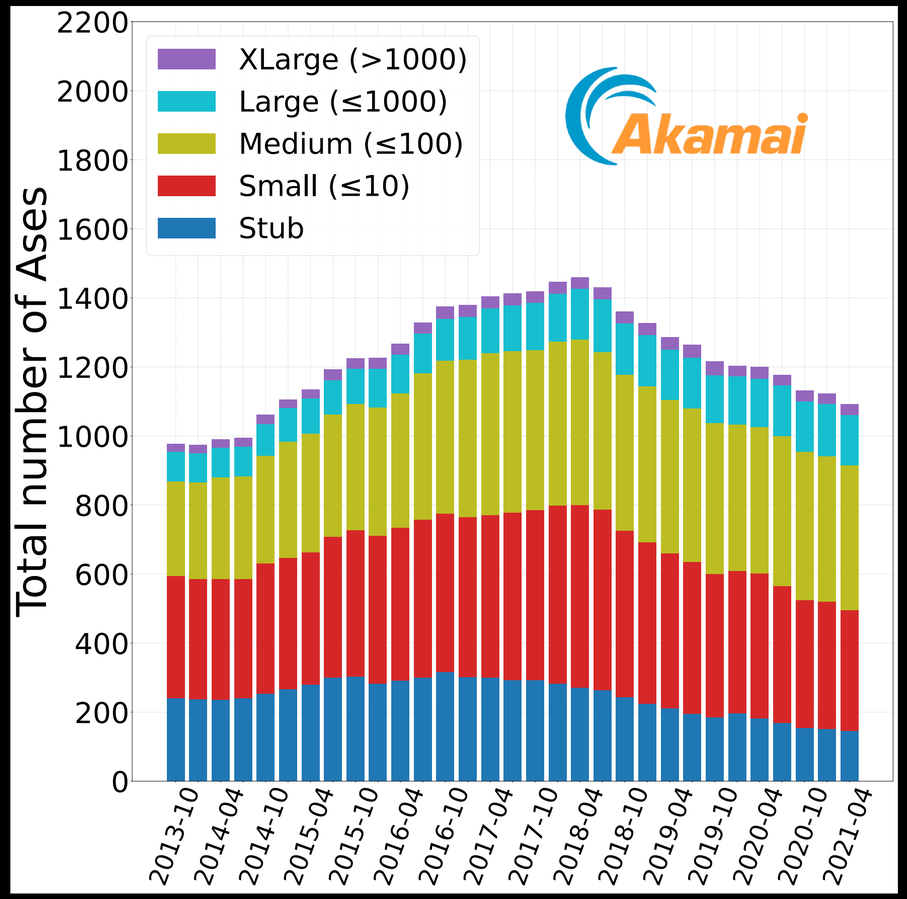
- Birth of HG CDNs.
- More aggressive increase 10x.
- Similar contributions of different types of networks.
- Significant increase after the lockdown.

Growth by Network Type

Understanding the “demographics” (Mode III)

Growth by Network Type

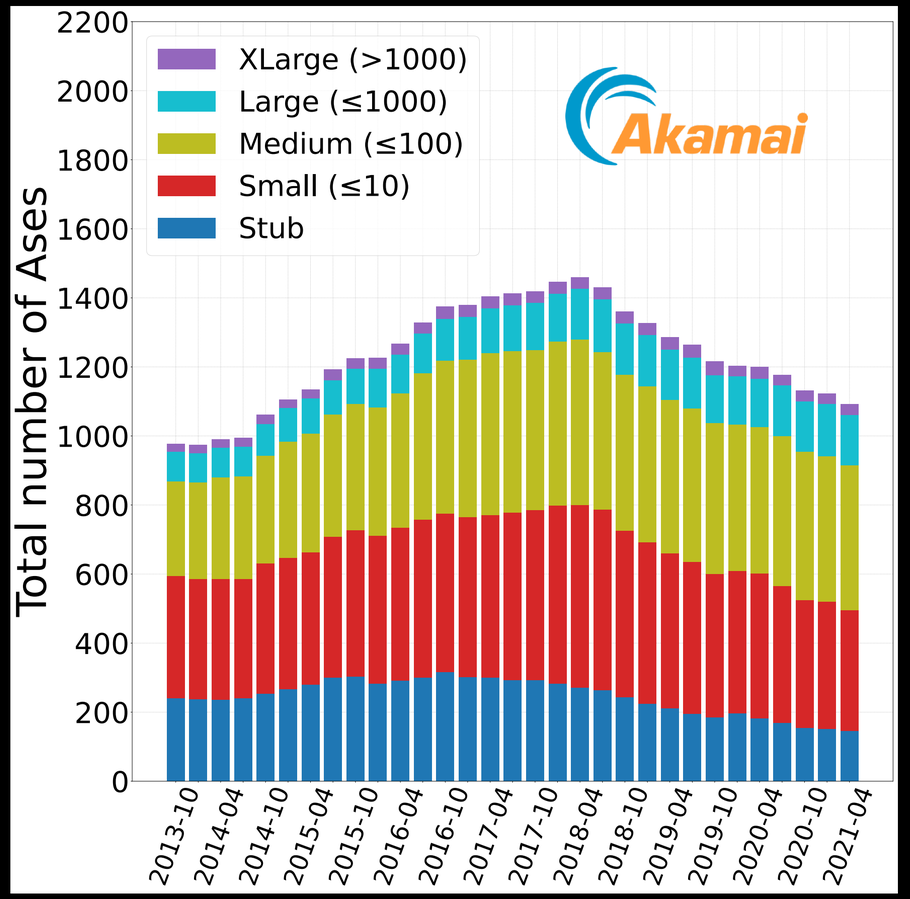
Understanding the “demographics” (Mode III)



Growth of Akamai’s off-net footprint grouped by AS customer cone size.

Growth by Network Type

Understanding the “demographics” (Mode III)

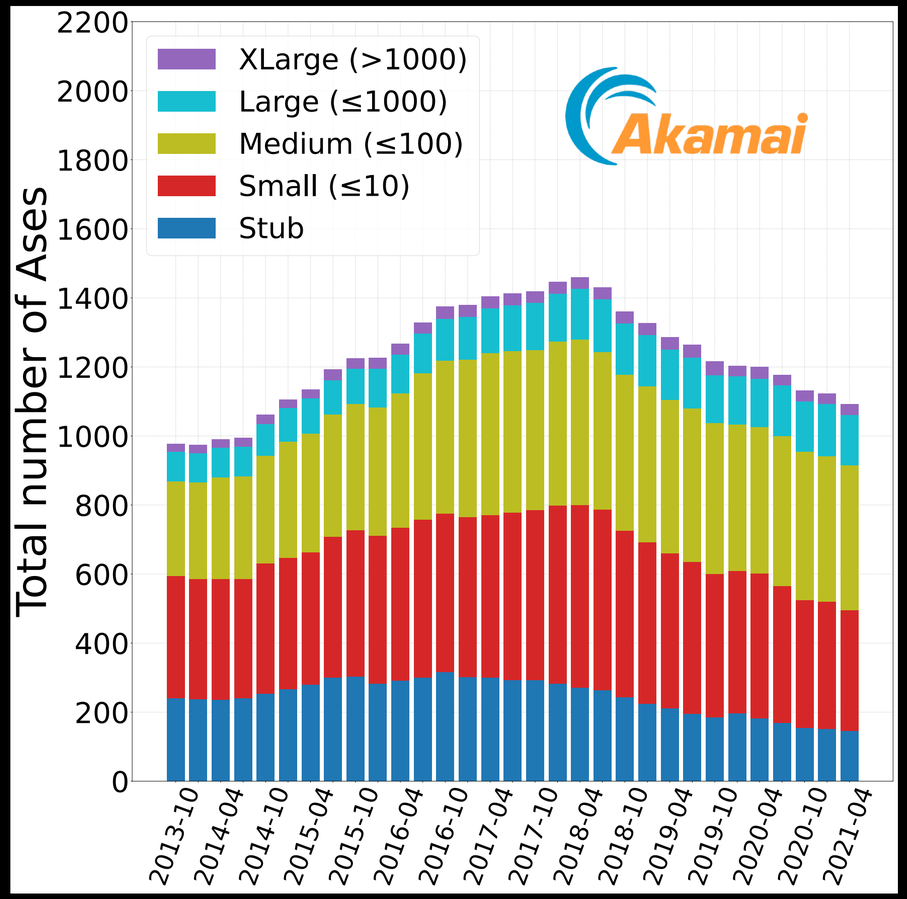


- Contribution of **stub** ASes since 2018 decline.

Growth of Akamai’s off-net footprint grouped by AS customer cone size.

Growth by Network Type

Understanding the “demographics” (Mode III)

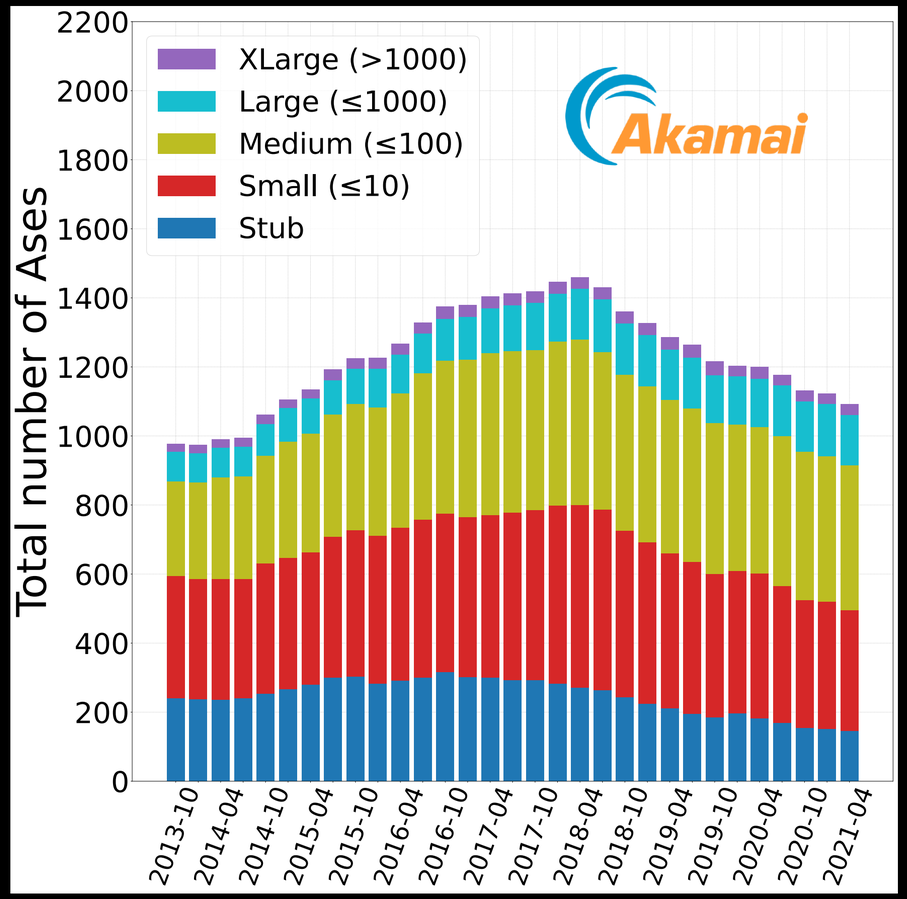


- Contribution of **stub** ASes since 2018 decline.
- Contribution of **small** & **medium** remain stable.

Growth of Akamai’s off-net footprint grouped by AS customer cone size.

Growth by Network Type

Understanding the “demographics” (Mode III)

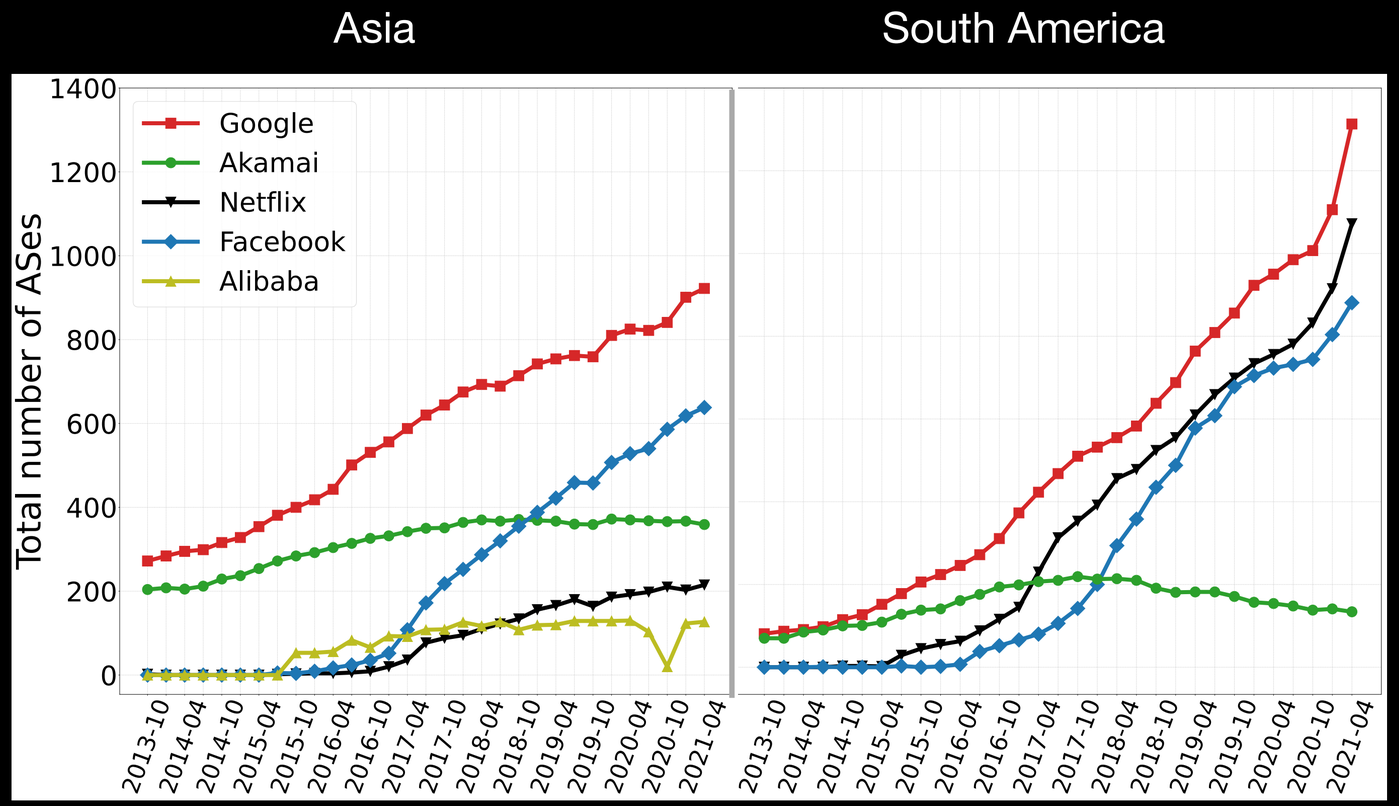


- Contribution of **stub** ASes since 2018 decline.
- Contribution of **small** & **medium** remain stable.
- Sum of **stub**, **small** and **medium** remains 84%.

Growth of Akamai’s off-net footprint grouped by AS customer cone size.

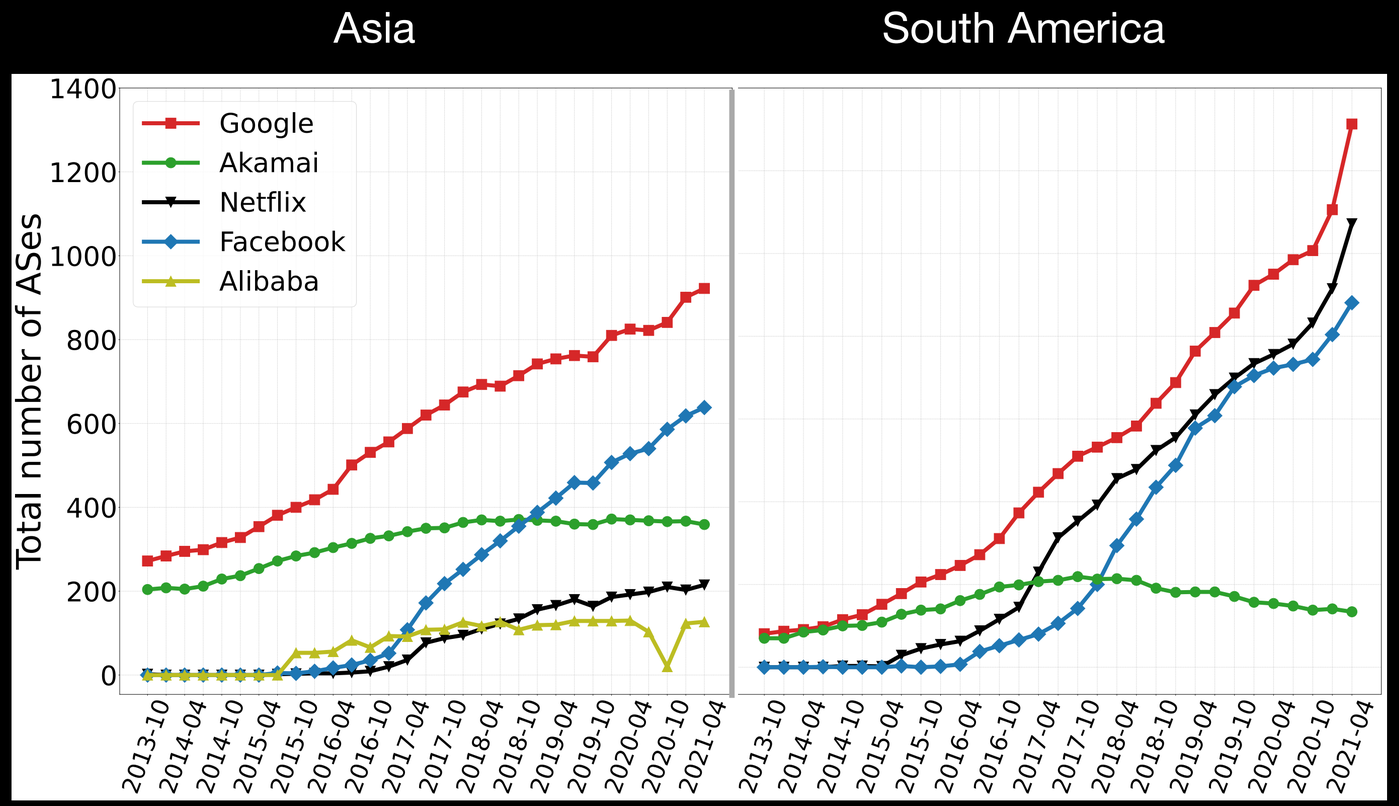
Regional Growth

Regional Growth



Growth of top-4 HGs (plus Alibaba) in Asia and South America continent over time.

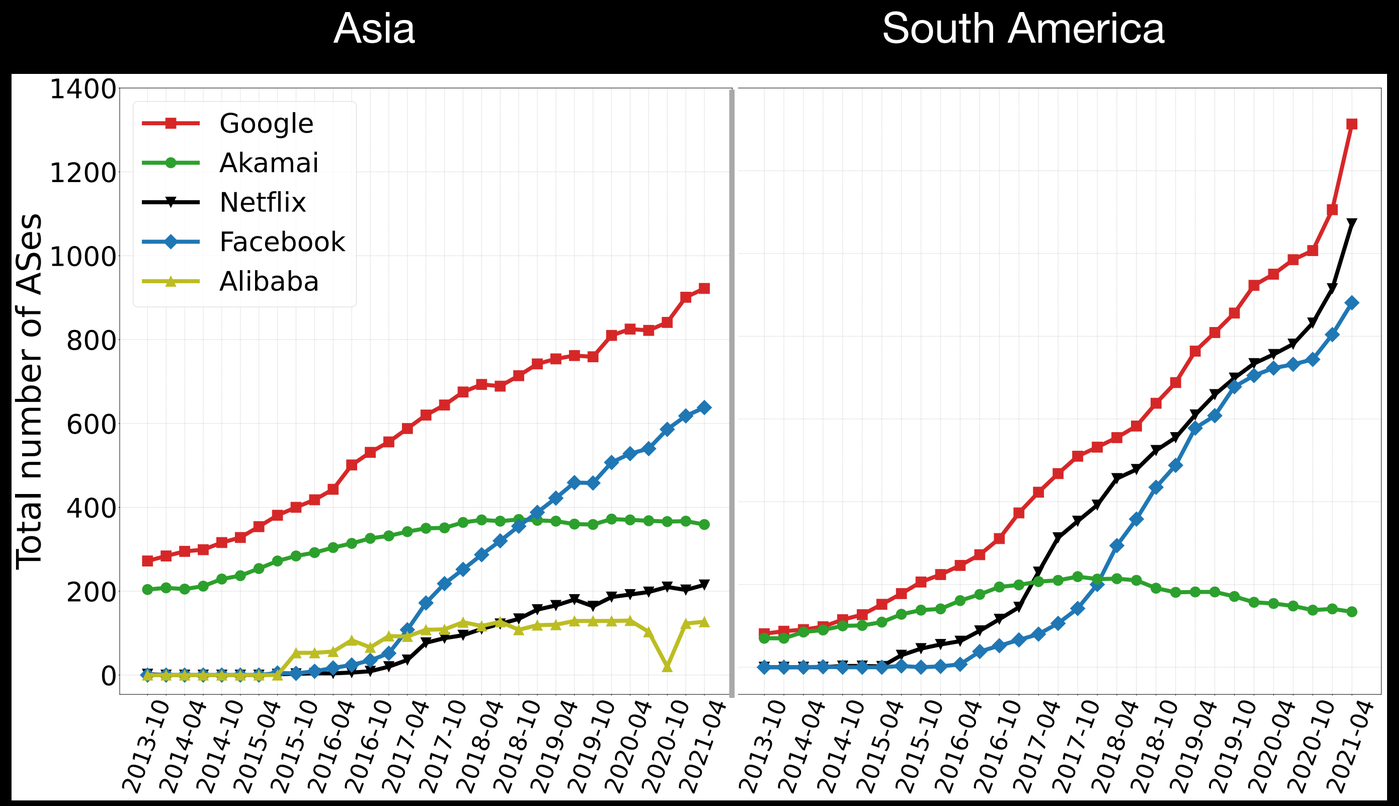
Regional Growth



Growth of top-4 HGs (plus Alibaba) in Asia and South America continent over time.

- Significant increase of all HGs (except Akamai) in all regions.

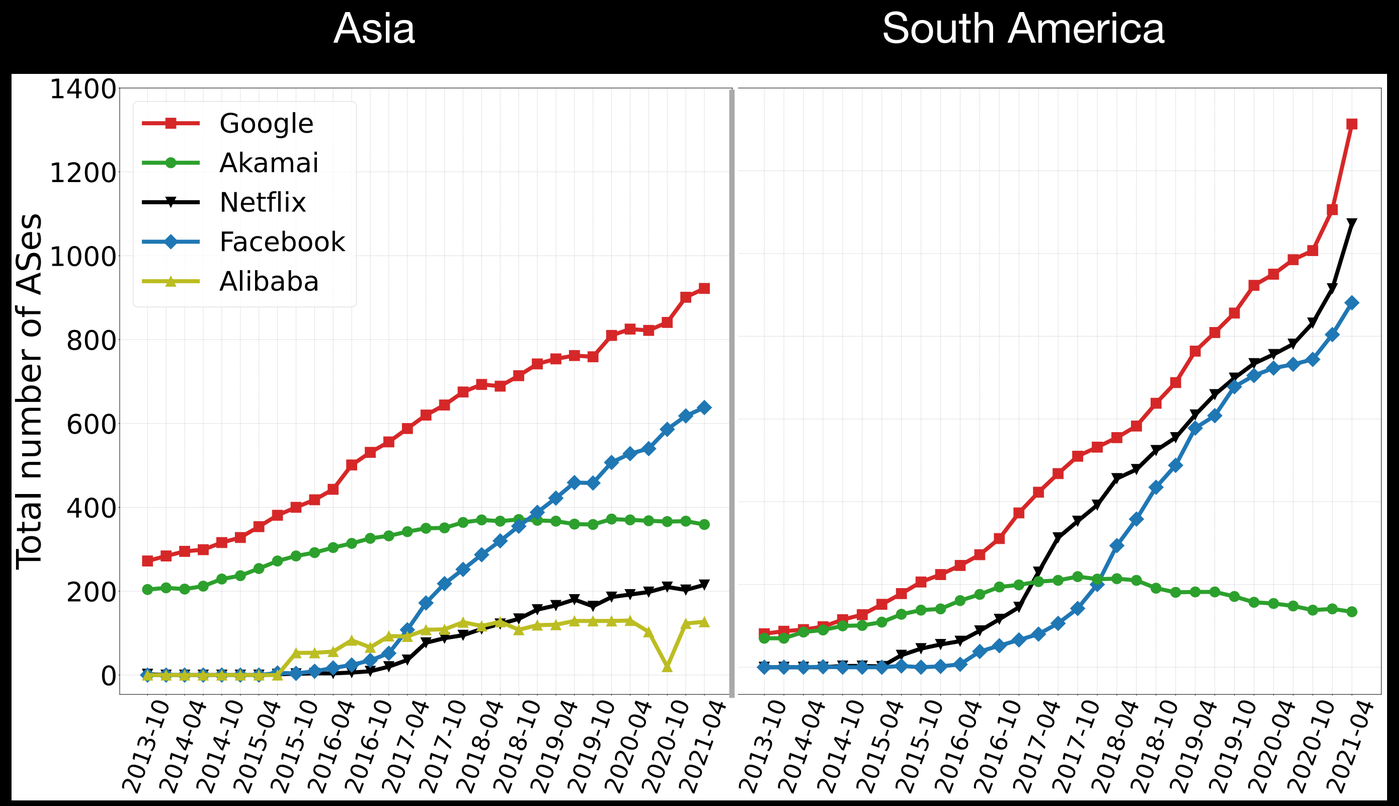
Regional Growth



Growth of top-4 HGs (plus Alibaba) in Asia and South America continent over time.

- Significant increase of all HGs (except Akamai) in all regions.
- Exponential growth in Asia and South America.

Regional Growth



Growth of top-4 HGs (plus Alibaba) in Asia and South America continent over time.

- Significant increase of all HGs (except Akamai) in all regions.
- Exponential growth in Asia and South America.
- Regional growth of some HGs (e.g., Alibaba in Asia).

APNIC Internet User Population Estimates

Source: How Big is that Network? | labs.apnic.net

APNIC Internet User Population Estimates

- APNIC conducts measurement campaigns ()
  to estimate the user population per AS.

Source: How Big is that Network? | labs.apnic.net

APNIC Internet User Population Estimates

- APNIC conducts measurement campaigns (Google Ads) to estimate the user population per AS.
- Uses ITU data to normalise findings.

Source: How Big is that Network? | labs.apnic.net

APNIC Internet User Population Estimates

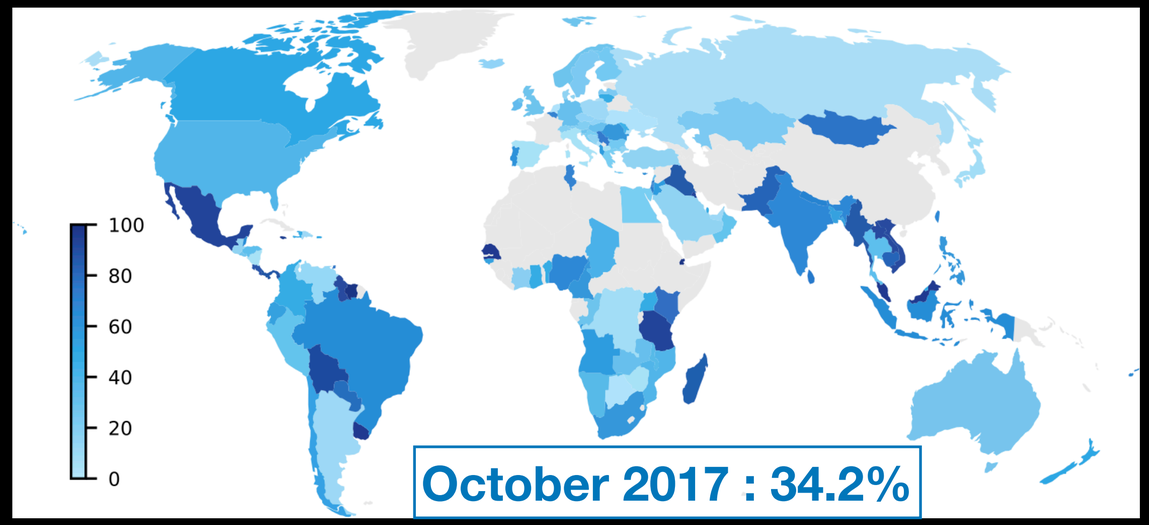
- APNIC conducts measurement campaigns (Google Ads) to estimate the user population per AS.
- Uses ITU data to normalise findings.
- The only available dataset that provides this type of information.

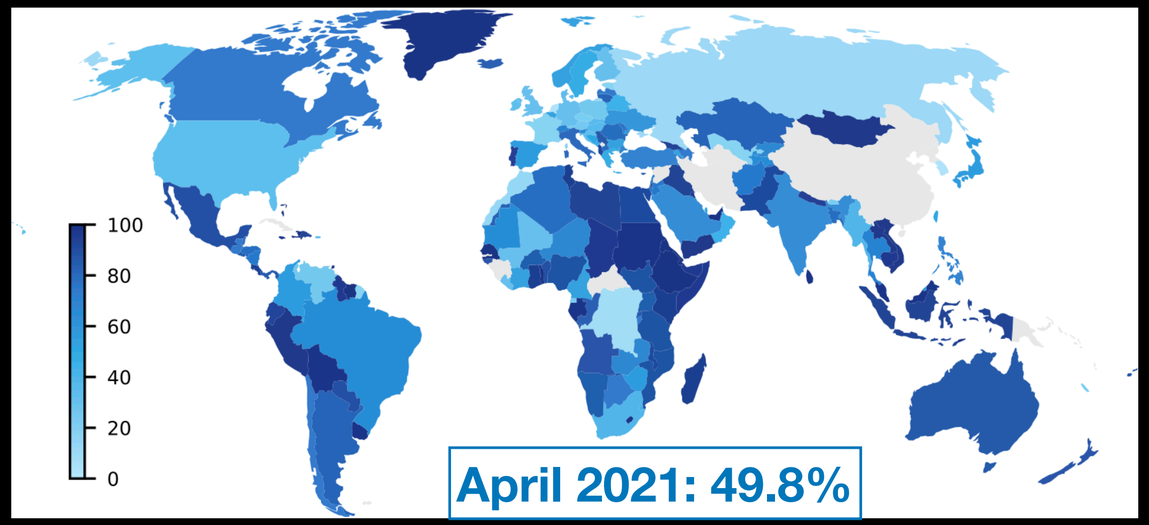
APNIC Internet User Population Estimates

- APNIC conducts measurement campaigns (Google Ads) to estimate the user population per AS.
- Uses ITU data to normalise findings.
- The only available dataset that provides this type of information.
- Daily snapshots from October 2017 to date.

Hypergiants’ off-nets Expansion

Internet User Population Coverage ***over time***





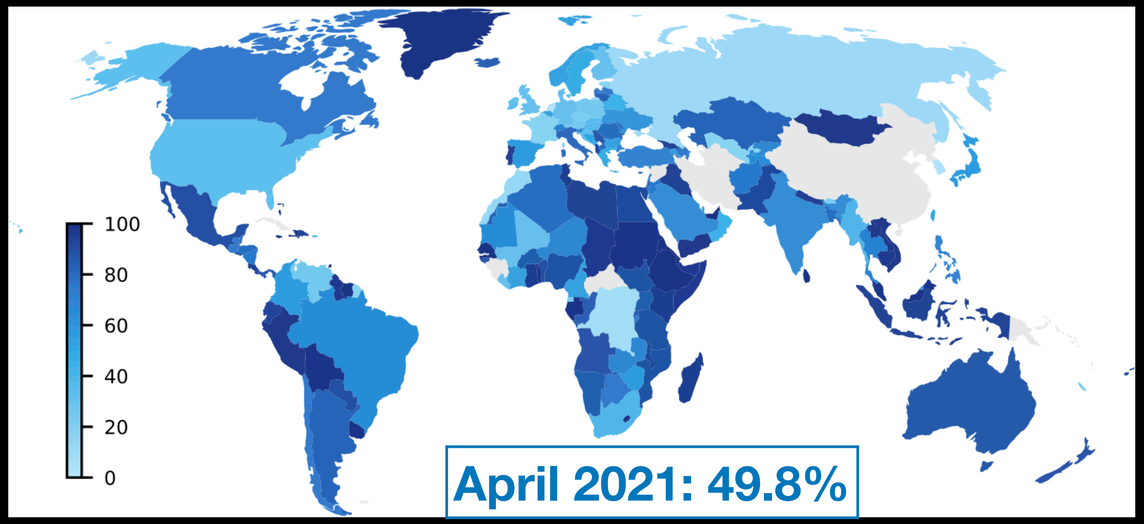
Facebook’s off-net footprint user coverage (%).

Hypergiants’ off-nets Expansion

Internet User Population Coverage ***over time***

Mexico: 90%

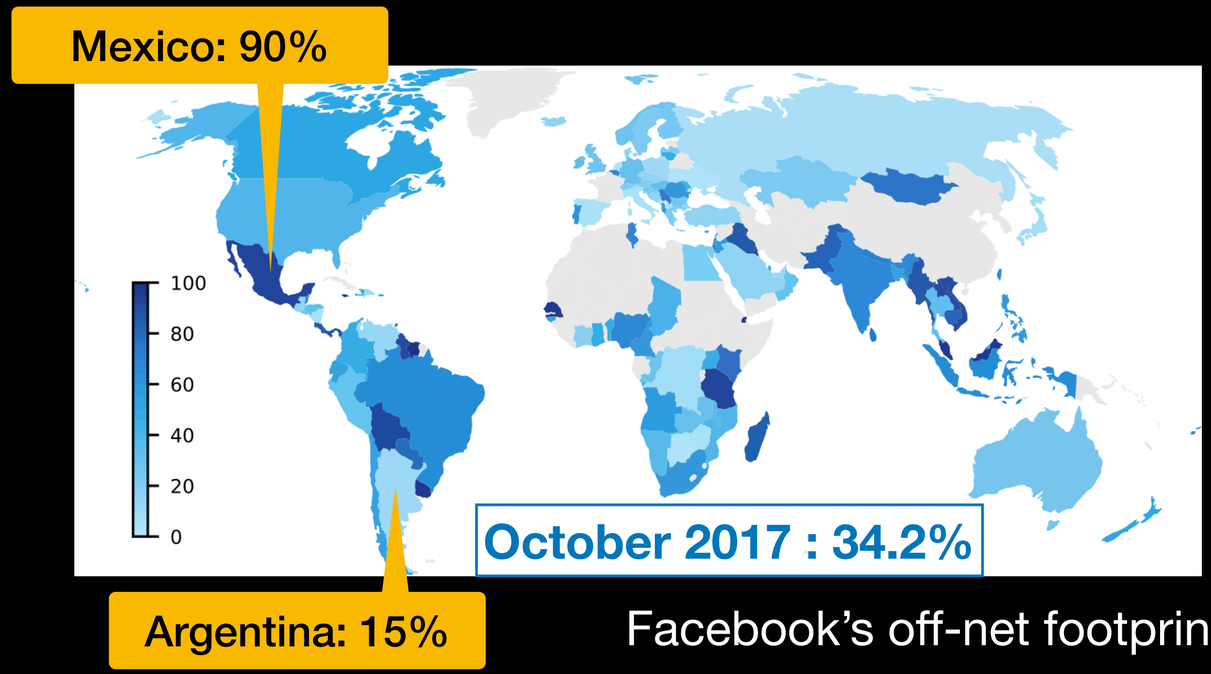
Argentina: 15%

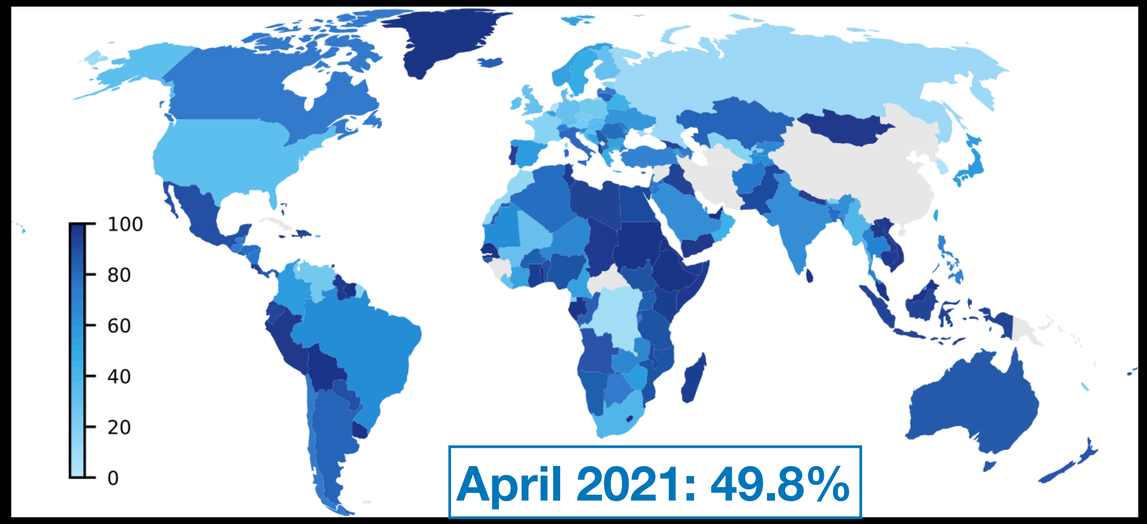


Facebook’s off-net footprint user coverage (%).

Hypergiants' off-nets Expansion

*Internet User Population Coverage **over time***



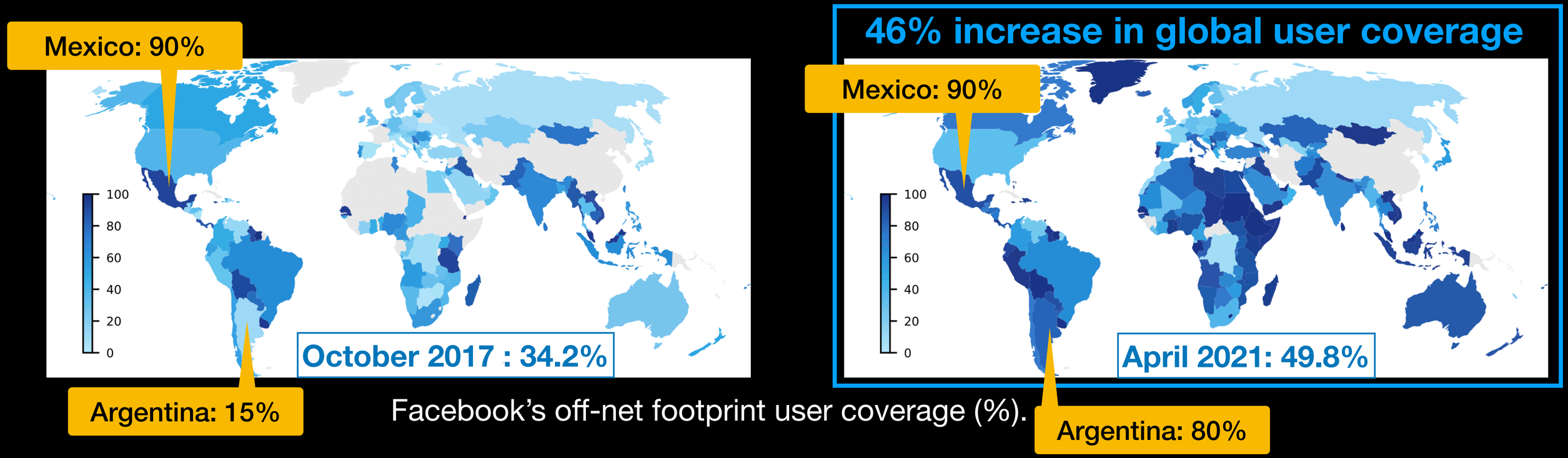


Facebook's off-net footprint user coverage (%).

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

Hypergiants' off-nets Expansion

Internet User Population Coverage ***over time***



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

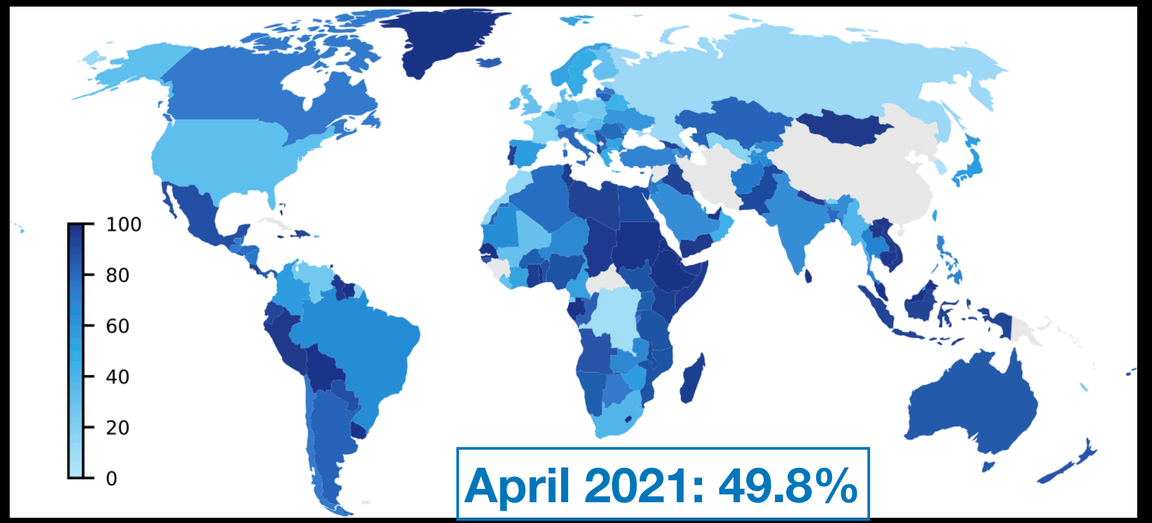
- Example: Facebook in 2017 announced that it had plans to expand in Africa and other developing regions.
- Our analysis reflect that they achieved this goal.

Hypergiants’ off-nets Expansion

Internet User Population Coverage ***based on customer cone***

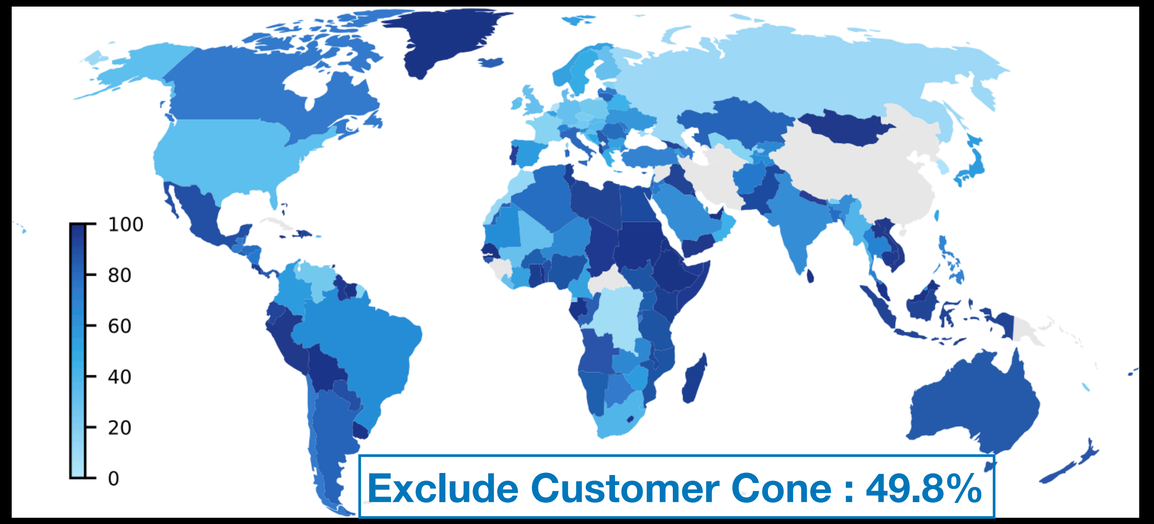
Hypergiants’ off-nets Expansion

Internet User Population Coverage ***based on customer cone***



Hypergiants' off-nets Expansion

Internet User Population Coverage ***based on customer cone***

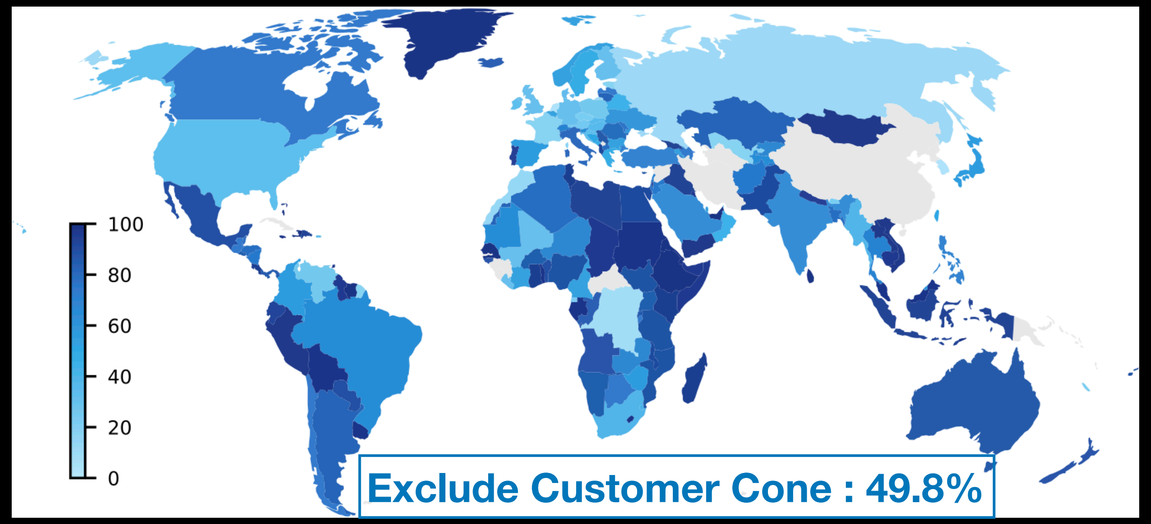


% of a country's Internet users **including** and **excluding** the customer cones of ASes hosting Facebook off-nets (April 2021).

Hypergiants' off-nets Expansion

Internet User Population Coverage ***based on customer cone***

27% increase in global user coverage



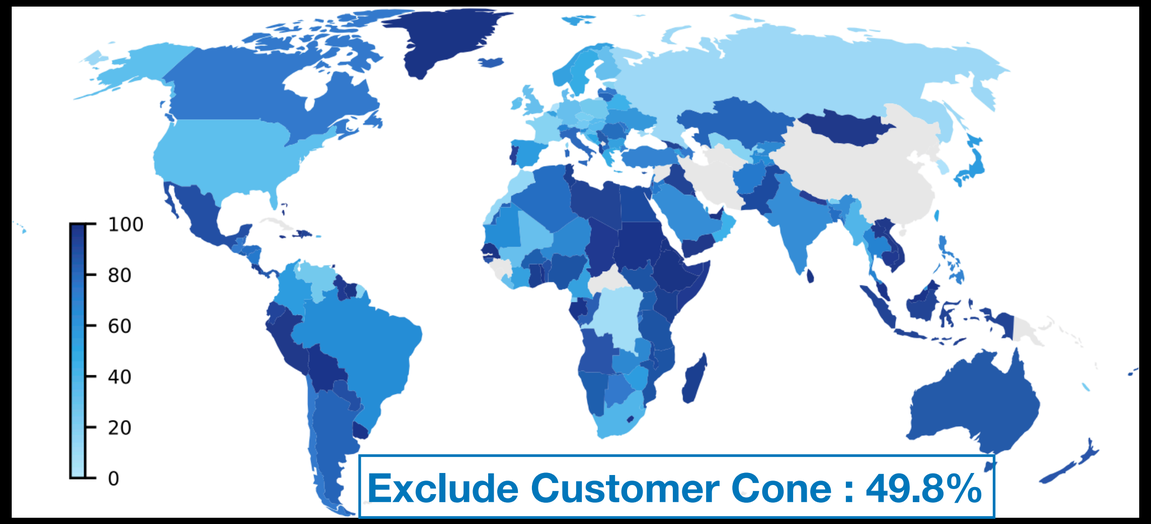
% of a country's Internet users **including** and **excluding** the customer cones of ASes hosting Facebook off-nets (April 2021).

- What-If #1: Serving into the customer cone noticeably expands coverage in parts of Africa, Asia, Europe and South America.

Hypergiants' off-nets Expansion

Internet User Population Coverage ***based on customer cone***

27% increase in global user coverage

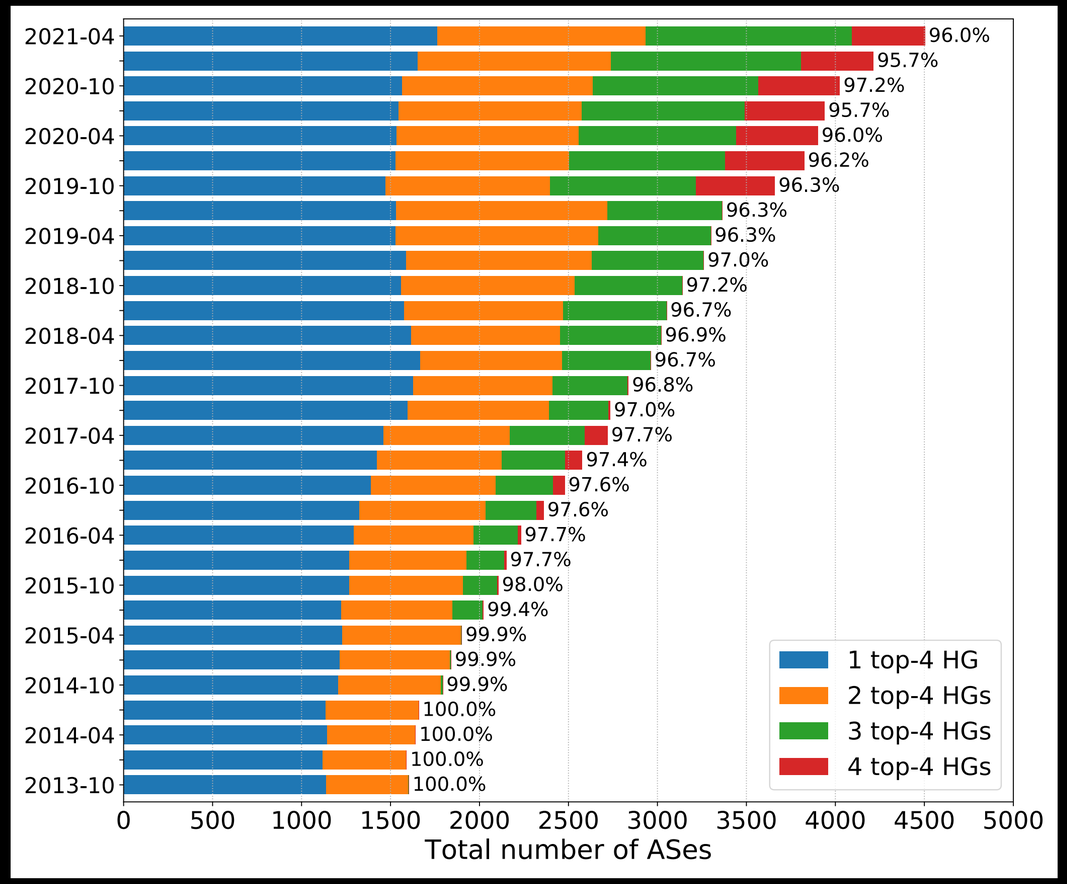


% of a country's Internet users **including** and **excluding** the customer cones of ASes hosting Facebook off-nets (April 2021).

- What-If #1: Serving into the customer cone noticeably expands coverage in parts of Africa, Asia, Europe and South America.
- What-If #2: Facebook coverage could significantly increase in the US from **33.9%** to **61.8%** by deploying off-nets in only 5 ASes.

Network Providers’ Hosting Strategies

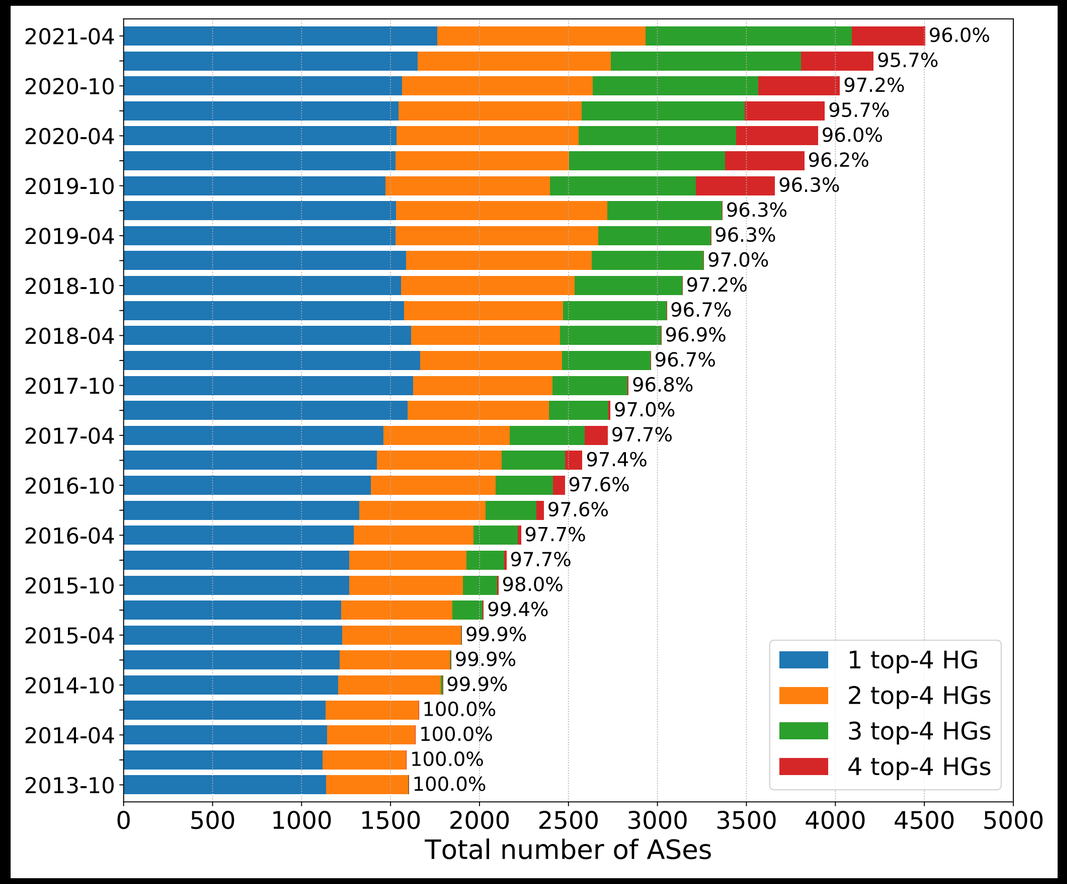
Network Providers' Hosting Strategies



ASes that host at least one top-4 HGs.

- More than 97% of ASes hosting off-nets, host at least one of the top-4 HGs.

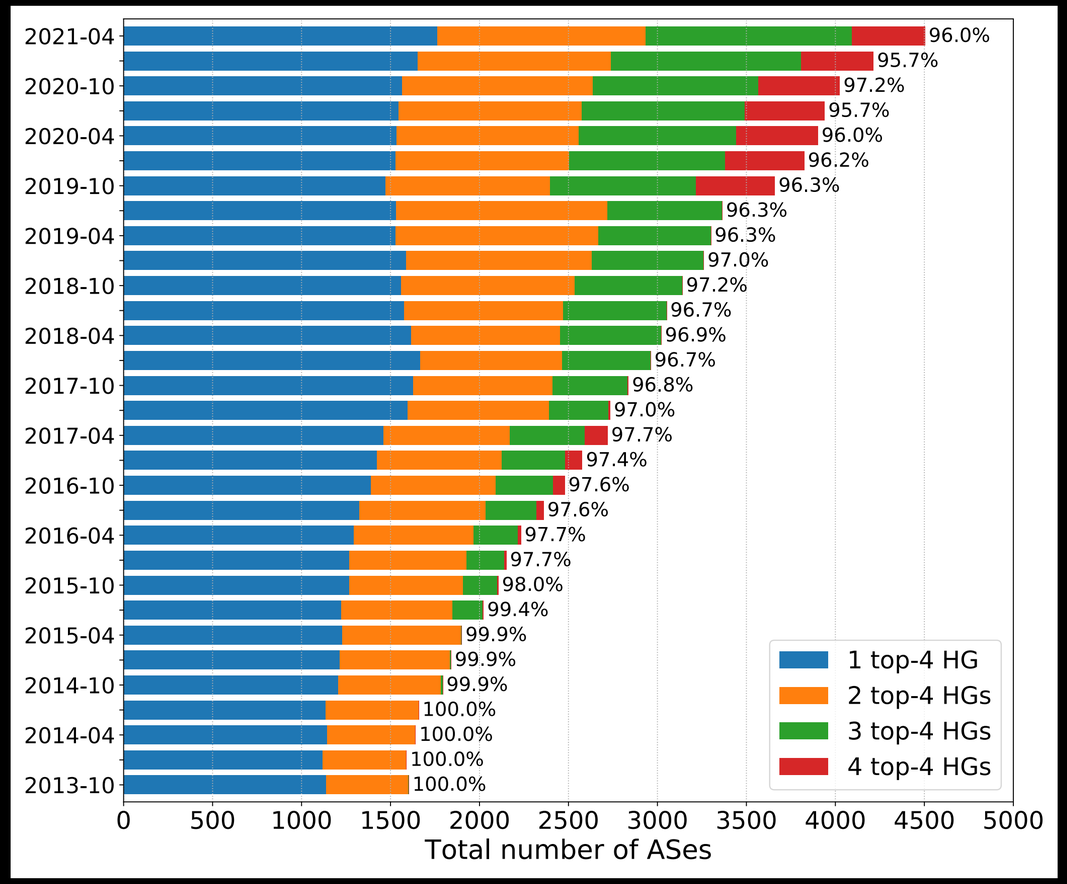
Network Providers' Hosting Strategies



ASes that host at least one top-4 HGs.

- More than 97% of ASes hosting off-nets, host at least one of the top-4 HGs.
- Top-4 HGs have increasingly similar footprints.

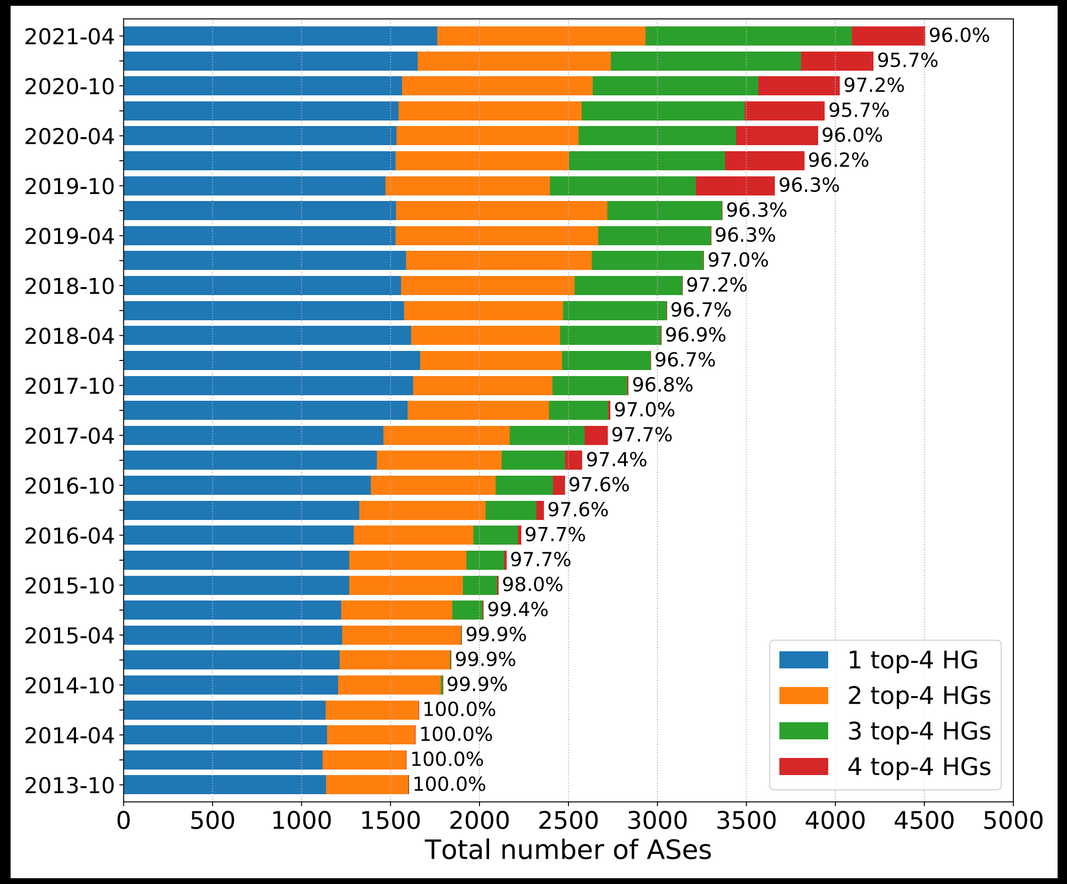
Network Providers' Hosting Strategies



ASes that host at least one top-4 HGs.

- More than 97% of ASes hosting off-nets, host at least one of the top-4 HGs.
- Top-4 HGs have increasingly similar footprints.
  - In 2021, more than 70% of ASes with off-nets host 2-4 top-4 HGs.
  - In 2013, it was less than 30%.

Network Providers' Hosting Strategies



ASes that host at least one top-4 HGs.

- More than 97% of ASes hosting off-nets, host at least one of the top-4 HGs.
- Top-4 HGs have increasingly similar footprints.
  - In 2021, more than 70% of ASes with off-nets host 2-4 top-4 HGs.
  - In 2013, it was less than 30%.
- A networks that already hosts one of the top-4 HGs is likely to later host more.

Validation

Validation

- **Validation from Hypergiants.**

Validation

- **Validation from Hypergiants.**
  - Four replied to our survey, all of them indicated that we **correctly uncovered 89-95%** of ASes hosting their off-nets.

Validation

- **Validation from Hypergiants.**
  - Four replied to our survey, all of them indicated that we **correctly uncovered 89-95%** of ASes hosting their off-nets.
- **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.

Limitations

Limitations

- **TLS-SNI.**
  - Multiple TLS certificates in a single IP address.

Limitations

- **TLS-SNI.**
  - Multiple TLS certificates in a single IP address.
- **Missing Headers.**
  - Default HTTP(S) headers (e.g., NGINX).

Limitations

- **TLS-SNI.**
  - Multiple TLS certificates in a single IP address.
- **Missing Headers.**
  - Default HTTP(S) headers (e.g., NGINX).
- **Special Architectures by HGs.**
  - HGs acting as middleware proxies (e.g., CLOUDFLARE).

Discussion

Discussion

- **An Increasingly Private Internet.**
  - Regulatory implications.
  - ISPs lose negotiation power in peering agreements with HGs.
  - Small networks reduce cost as upstream/downstream traffic is reduced.

Discussion

- **An Increasingly Private Internet.**
  - Regulatory implications.
  - ISPs lose negotiation power in peering agreements with HGs.
  - Small networks reduce cost as upstream/downstream traffic is reduced.
- **Unintended Consequences.**
  - Knowing HGs off-net servers makes it easier for attackers to be effective.
  - Business intelligence by competitors.

Discussion

- **An Increasingly Private Internet.**
  - Regulatory implications.
  - ISPs lose negotiation power in peering agreements with HGs.
  - Small networks reduce cost as upstream/downstream traffic is reduced.
- **Unintended Consequences.**
  - Knowing HGs off-net servers makes it easier for attackers to be effective.
  - Business intelligence by competitors.
- **Hide-and-Seek.**
  - Increase the bar for server identification by implementing TLS-SNI.

Takeaways

Takeaways

- Large content providers serve most of the traffic on today's Internet. To send this traffic, some of them host *off-net* servers in user ASes **around the world**.

Takeaways

- Large content providers serve most of the traffic on today's Internet. To send this traffic, some of them host *off-net* servers in user ASes **around the world**.
- **Generic methodology** to uncover off-net deployments.

Takeaways

- Large content providers serve most of the traffic on today's Internet. To send this traffic, some of them host *off-net* servers in user ASes **around the world**.
- **Generic methodology** to uncover off-net deployments.
- **Significant growth** of 3 (Google , facebook , NETFLIX) out of 4 top HGs, **hosted in more than 4.5k ASes.**

Takeaways

- Large content providers serve most of the traffic on today's Internet. To send this traffic, some of them host *off-net* servers in user ASes **around the world**.
- **Generic methodology** to uncover off-net deployments.
- **Significant growth** of 3 (Google , facebook , NETFLIX) out of 4 top HGs, **hosted in more than 4.5k ASes.**
- Significant fraction of **user population** can be served by off-nets in their ISP.

Takeaways

- Large content providers serve most of the traffic on today's Internet. To send this traffic, some of them host *off-net* servers in user ASes **around the world**.
- **Generic methodology** to uncover off-net deployments.
- **Significant growth** of 3 (Google , facebook , NETFLIX) out of 4 top HGs, **hosted in more than 4.5k ASes.**
- Significant fraction of **user population** can be served by off-nets in their ISP.
- Study of **ISPs wiliness to host HG off-nets**.

Takeaways

- Large content providers serve most of the traffic on today's Internet. To send this traffic, some of them host *off-net* servers in user ASes **around the world**.
- **Generic methodology** to uncover off-net deployments.
- **Significant growth** of 3 (Google , facebook , NETFLIX) out of 4 top HGs, **hosted in more than 4.5k ASes.**
- Significant fraction of **user population** can be served by off-nets in their ISP.
- Study of **ISPs wiliness to host HG off-nets**.
- **Artifacts**, **datasets,** and **an interactive portal** are available at:

https://pgigis.github.io/hypergiants-offnets/