Cloud Provider Connectivity in the Flat Internet

Todd Arnold[†], Jia He[†], Weifan Jiang[†], Matt Calder[‡†], Italo Cunha[♮†], Vasileios Giotsas[#], Ethan Katz-Bassett[†]

[†] Columbia University in the City of New York

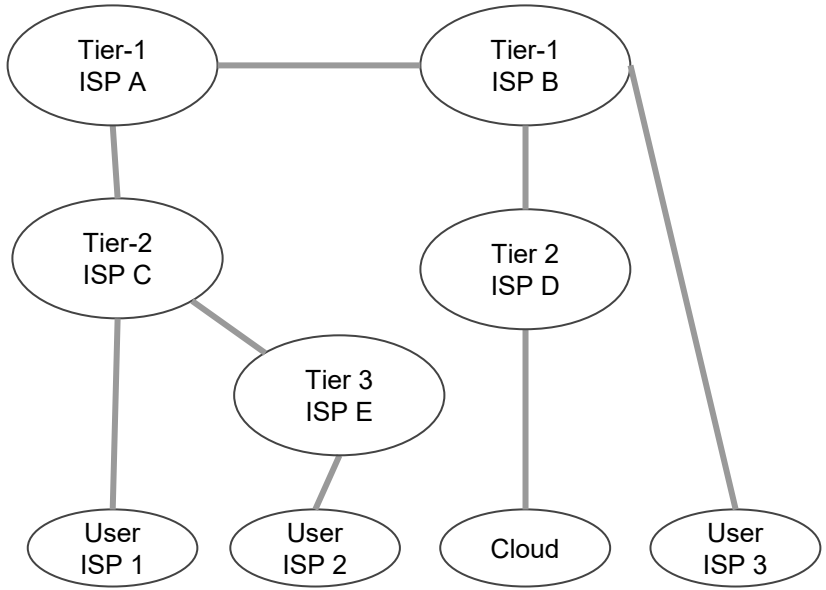
[‡] Microsoft

[♮] UFMG Universidade Federal de Minas Gerais

[#] Lancaster University

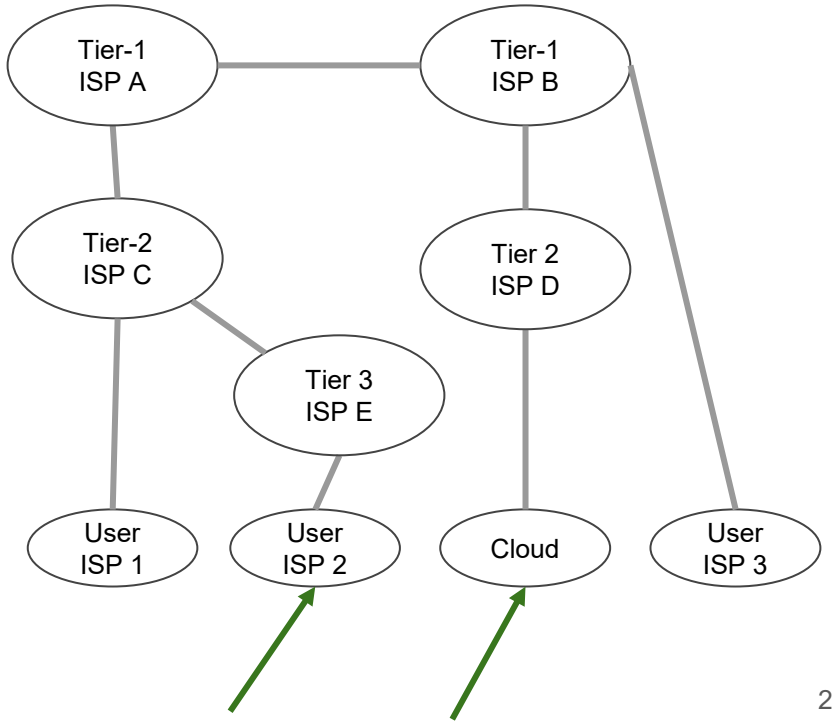
Evolving Internet Topology

- Traditional: Hierarchical
  - Tier-1 ISPs are global networks, and all other networks fall under at least one
  - Tier-2 ISPs are larger, regional networks
  - Tier-3 ISPs interconnect edge networks
  - Edge networks at the bottom
  - Networks pay higher tiers to transport their data (a.k.a. *transit*)



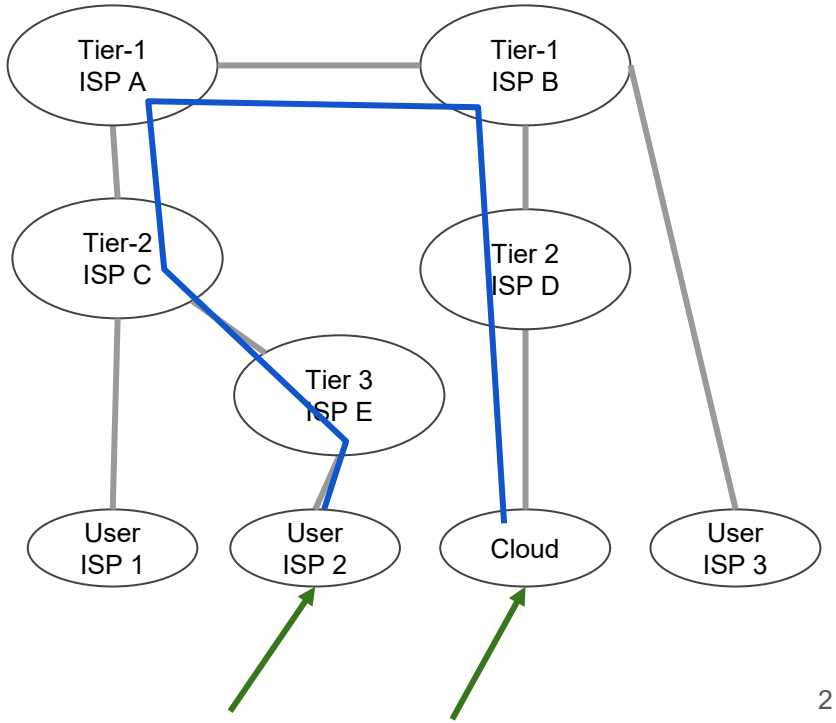
Evolving Internet Topology

- Traditional: Hierarchical
  - Tier-1 ISPs are global networks, and all other networks fall under at least one
  - Tier-2 ISPs are larger, regional networks
  - Tier-3 ISPs interconnect edge networks
  - Edge networks at the bottom
  - Networks pay higher tiers to transport their data (a.k.a. *transit*)



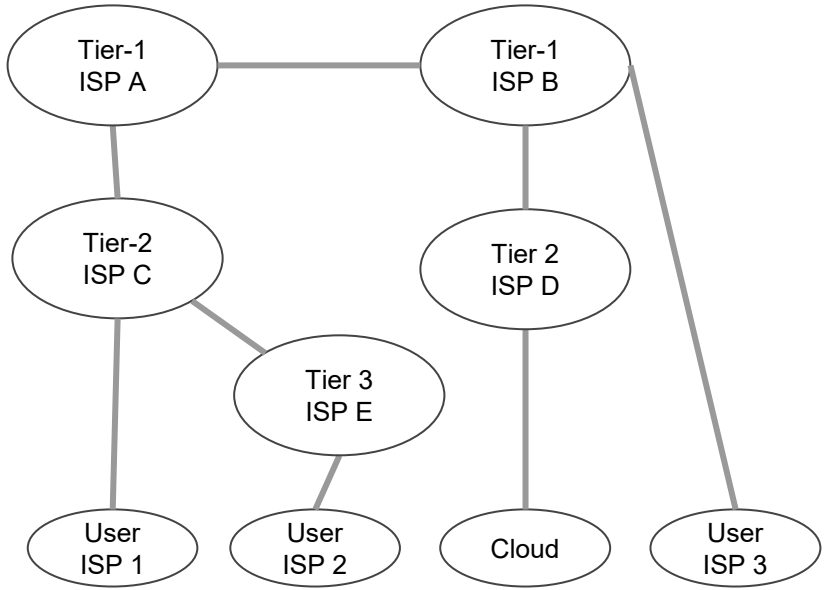
Evolving Internet Topology

- Traditional: Hierarchical
  - Tier-1 ISPs are global networks, and all other networks fall under at least one
  - Tier-2 ISPs are larger, regional networks
  - Tier-3 ISPs interconnect edge networks
  - Edge networks at the bottom
  - Networks pay higher tiers to transport their data (a.k.a. *transit*)



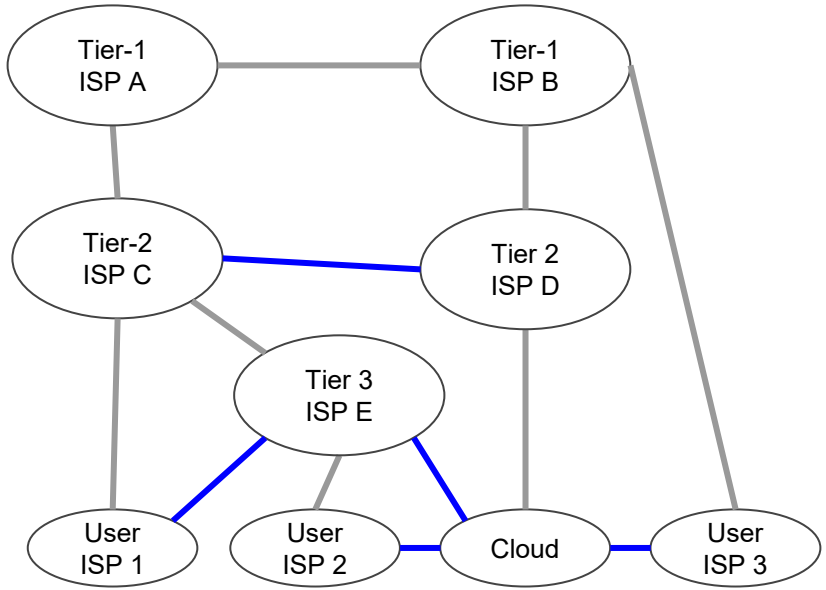
Evolving Internet Topology

- Traditional: Hierarchical
  - Tier-1 ISPs are global networks, and all other networks fall under at least one
  - Tier-2 ISPs are larger, regional networks
  - Tier-3 ISPs interconnect edge networks
  - Edge networks at the bottom
  - Networks pay higher tiers to transport their data (a.k.a. *transit*)
- Recent Years: Internet Flattening
  - Increased direct connectivity between networks at lower tiers



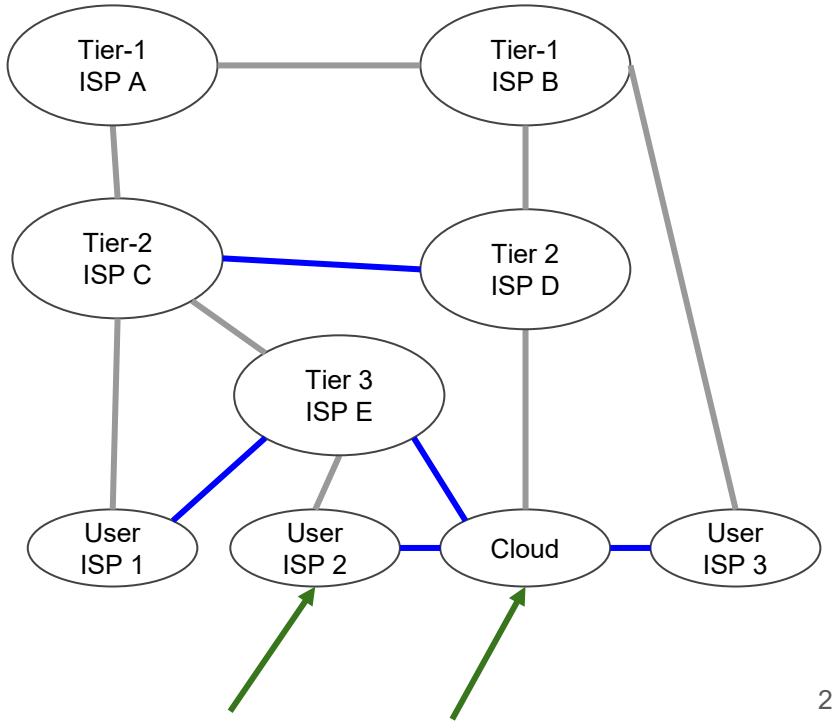
Evolving Internet Topology

- Traditional: Hierarchical
  - Tier-1 ISPs are global networks, and all other networks fall under at least one
  - Tier-2 ISPs are larger, regional networks
  - Tier-3 ISPs interconnect edge networks
  - Edge networks at the bottom
  - Networks pay higher tiers to transport their data (a.k.a. *transit*)
- Recent Years: Internet Flattening
  - Increased direct connectivity between networks at lower tiers



Evolving Internet Topology

- Traditional: Hierarchical
  - Tier-1 ISPs are global networks, and all other networks fall under at least one
  - Tier-2 ISPs are larger, regional networks
  - Tier-3 ISPs interconnect edge networks
  - Edge networks at the bottom
  - Networks pay higher tiers to transport their data (a.k.a. *transit*)
- Recent Years: Internet Flattening
  - Increased direct connectivity between networks at lower tiers



Motivation and Goals

- Majority of Internet traffic now occurs over direct connections
- Impact of flattening is not captured by traditional approaches
  - Invisible to traditional vantage points
  - Existing metrics of importance (*e.g.,* customer cone)
    - Do not reflect the rich peering interconnectivity of the flat Internet
    - Focus on how much transit an AS could provide rather than how much it does provide
- To understand this gap and capture the progress of Internet flattening
  - Uncover the missing links
  - Understand to what degree they enable the major cloud providers (Amazon, Google, IBM, and Microsoft) to bypass the traditional hierarchy

Limitations of Available Measurements

- BGP feeds
  - High visibility of transit connections (90+% coverage of Tier-1 and Tier-2 interconnections) [1]
  - Poor coverage of edge networks and peering links (~10% coverage of interconnects) [1]

[1] Oliveira et al. 2010. The (In)Completeness of the Observed Internet AS-level Structure. (ToN)

Limitations of Available Measurements

- BGP feeds
  - High visibility of transit connections (90+% coverage of Tier-1 and Tier-2 interconnections) [1]
  - Poor coverage of edge networks and peering links (~10% coverage of interconnects) [1]
- Traceroutes
  - Can be sourced from inside the cloud providers
  - Can infer false links due to dropped packets or load balancing
  - Networks can interfere with measurements

[1] Oliveira et al. 2010. The (In)Completeness of the Observed Internet AS-level Structure. (ToN)

Methodology

AS topology graph from two sources

- BGP Feeds
  - Use CAIDA’s AS relationship dataset from Sep 2020 as our initial topology

Methodology

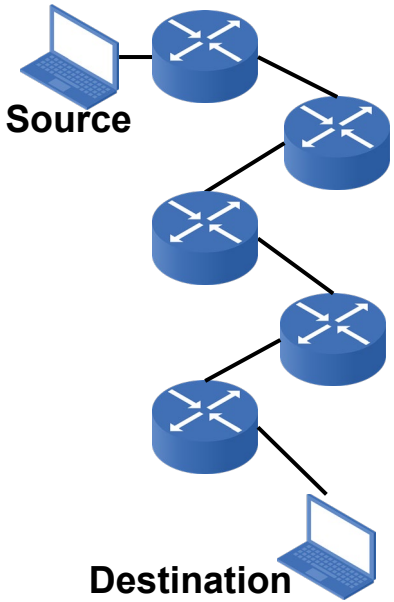
AS topology graph from two sources

- BGP Feeds
  - Use CAIDA's AS relationship dataset from Sep 2020 as our initial topology
- Traceroutes from inside clouds
  - Used to identify directly connected neighbors to add to the topology
  - Target every prefix using ICMP via scamper
  - Map IP-to-ASN using PeeringDB, Cymru IP-to-ASN Mapping Tool, and whois
  - Traceroute must
    - Have at least one cloud provider hop
    - One non-cloud hop
    - No intermediate hop with an unresolved IP address

Methodology

AS topology graph from two sources

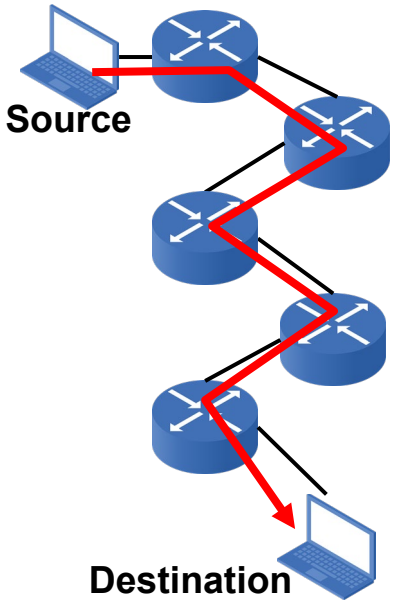
- BGP Feeds
  - Use CAIDA's AS relationship dataset from Sep 2020 as our initial topology
- Traceroutes from inside clouds
  - Used to identify directly connected neighbors to add to the topology
  - Target every prefix using ICMP via scamper
  - Map IP-to-ASN using PeeringDB, Cymru IP-to-ASN Mapping Tool, and whois
  - Traceroute must
    - Have at least one cloud provider hop
    - One non-cloud hop
    - No intermediate hop with an unresolved IP address



Methodology

AS topology graph from two sources

- BGP Feeds
  - Use CAIDA's AS relationship dataset from Sep 2020 as our initial topology
- Traceroutes from inside clouds
  - Used to identify directly connected neighbors to add to the topology
  - Target every prefix using ICMP via scamper
  - Map IP-to-ASN using PeeringDB, Cymru IP-to-ASN Mapping Tool, and whois
  - Traceroute must
    - Have at least one cloud provider hop
    - One non-cloud hop
    - No intermediate hop with an unresolved IP address

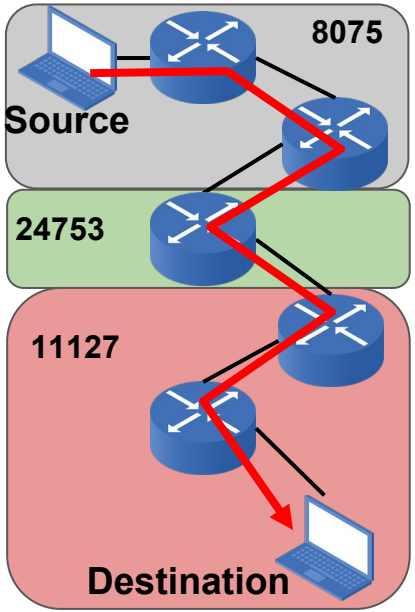


IP Address	ASN
*	
104.44.16.180	
193.238.116.5	
207.241.160.1	
207.241.160.38	
81.85.101.89	

Methodology

AS topology graph from two sources

- BGP Feeds
  - Use CAIDA's AS relationship dataset from Sep 2020 as our initial topology
- Traceroutes from inside clouds
  - Used to identify directly connected neighbors to add to the topology
  - Target every prefix using ICMP via scamper
  - Map IP-to-ASN using PeeringDB, Cymru IP-to-ASN Mapping Tool, and whois
  - Traceroute must
    - Have at least one cloud provider hop
    - One non-cloud hop
    - No intermediate hop with an unresolved IP address

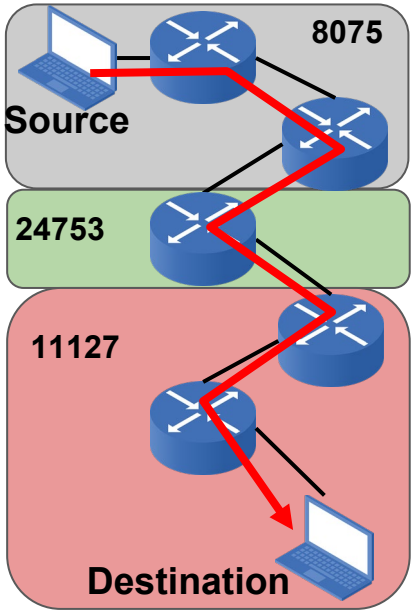


IP Address	ASN
*	*
104.44.16.180	8075
193.238.116.5	24753
207.241.160.1	11127
207.241.160.38	11127
81.85.101.89	11127

Methodology

AS topology graph from two sources

- BGP Feeds
  - Use CAIDA's AS relationship dataset from Sep 2020 as our initial topology
- Traceroutes from inside clouds
  - Used to identify directly connected neighbors to add to the topology
  - Target every prefix using ICMP via scamper
  - Map IP-to-ASN using PeeringDB, Cymru IP-to-ASN Mapping Tool, and whois
  - Traceroute must
    - Have at least one cloud provider hop
    - One non-cloud hop
    - No intermediate hop with an unresolved IP address

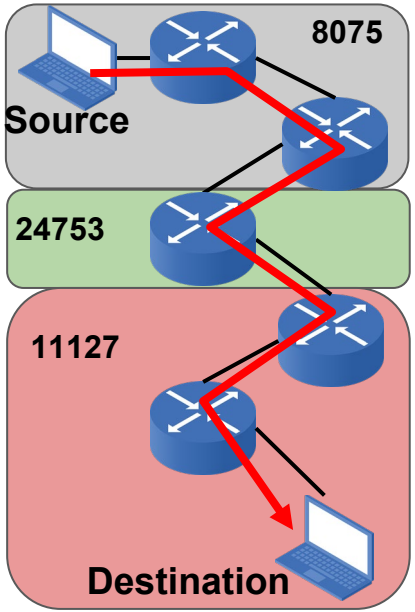


IP Address	ASN
*	*
104.44.16.180	8075
193.238.116.5	24753
207.241.160.1	11127
207.241.160.38	11127
81.85.101.89	11127

Methodology

AS topology graph from two sources

- BGP Feeds
  - Use CAIDA's AS relationship dataset from Sep 2020 as our initial topology
- Traceroutes from inside clouds
  - Used to identify directly connected neighbors to add to the topology
  - Target every prefix using ICMP via scamper
  - Map IP-to-ASN using PeeringDB, Cymru IP-to-ASN Mapping Tool, and whois
  - Traceroute must
    - Have at least one cloud provider hop
    - One non-cloud hop
    - No intermediate hop with an unresolved IP address

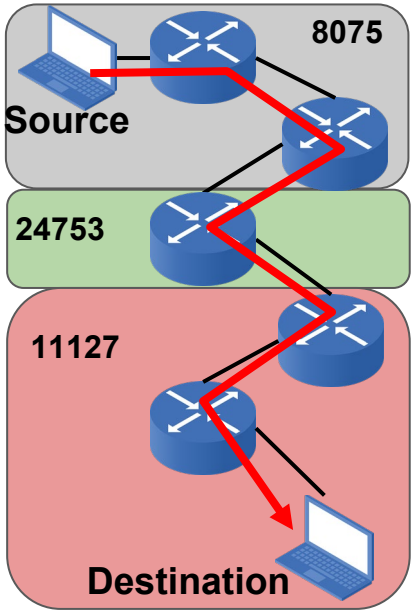


IP Address	ASN
*	*
104.44.16.180	8075
193.238.116.5	24753
207.241.160.1	11127
207.241.160.38	11127
81.85.101.89	11127

Methodology

AS topology graph from two sources

- BGP Feeds
  - Use CAIDA's AS relationship dataset from Sep 2020 as our initial topology
- Traceroutes from inside clouds
  - Used to identify directly connected neighbors to add to the topology
  - Target every prefix using ICMP via scamper
  - Map IP-to-ASN using PeeringDB, Cymru IP-to-ASN Mapping Tool, and whois
  - Traceroute must
    - Have at least one cloud provider hop
    - One non-cloud hop
    - No intermediate hop with an unresolved IP address



IP Address	ASN
*	*
104.44.16.180	8075
*	*
207.241.160.1	11127
207.241.160.38	11127
81.85.101.89	11127

Validation

Iterative process with Microsoft and Google

- True and false positives
  - False Discovery Rate (FDR): FP / (FP +TP)
- True and false negatives
  - False Negative Rate (FNR): FN / (FN +TP)
- Worked with Microsoft while we refined our methodology
- Google’s feedback validated our refinements

Validation (cont.)

- Initial methodology
  - Allow a single unresponsive/unresovled hop between cloud and neighbor; assume the AS immediately following this hop was a neighbor

Validation (cont.)

- Initial methodology
  - Allow a single unresponsive/unresovled hop between cloud and neighbor; assume the AS immediately following this hop was a neighbor

IP Address	ASN
*	*
104.44.16.180	8075
*	*
207.241.160.1	11127
207.241.160.38	11127
81.85.101.89	11127

Validation (cont.)

- Initial methodology
  - Allow a single unresponsive/unresovled hop between cloud and neighbor; assume the AS immediately following this hop was a neighbor

IP Address	ASN
*	*
104.44.16.180	8075
*	*
207.241.160.1	11127
207.241.160.38	11127
81.85.101.89	11127

Validation (cont.)

- Initial methodology
  - Allow a single unresponsive/unresovled hop between cloud and neighbor; assume the AS immediately following this hop was a neighbor
  - Both: ~50% FDR and 23 – 50% FNR. Microsoft: 8,910 neighbors, Google: 13,336 neighbors

IP Address	ASN
*	*
104.44.16.180	8075
*	*
207.241.160.1	11127
207.241.160.38	11127
81.85.101.89	11127

Validation (cont)

- ~~Initial methodology~~
  - ~~Allow a single unresponsive/unresovled hop between cloud and neighbor; assume the AS immediately following this hop was a neighbor~~
  - ~~~50% FDR and 23 – 50% FNR~~
- Evaluated traceroute that caused false positives
  - Some had timeouts for intermediate hop – not possible to resolve

IP Address	ASN
*	*
104.44.16.180	8075
*	*
207.241.160.1	11127
207.241.160.38	11127
81.85.101.89	11127

Validation (cont)

- ~~Initial methodology~~
  - ~~Allow a single unresponsive/unresovled hop between cloud and neighbor; assume the AS immediately following this hop was a neighbor~~
  - ~~~50% FDR and 23 – 50% FNR~~
- Evaluated traceroute that caused false positives
  - Some had timeouts for intermediate hop – not possible to resolve

IP Address	ASN
*	*
104.44.16.180	8075
*	*
207.241.160.1	11127
207.241.160.38	11127
81.85.101.89	11127

Validation (cont)

- ~~Initial methodology~~
  - ~~Allow a single unresponsive/unresovled hop between cloud and neighbor; assume the AS immediately following this hop was a neighbor~~
  - ~~~50% FDR and 23 – 50% FNR~~
- Evaluated traceroute that caused false positives
  - Some had timeouts for intermediate hop – not possible to resolve
  - Most had intermediate hop IP
    - AS not resolved by Cymru using BGP
    - Registered in whois to an IXP – examined PeeringDB

IP Address	ASN
*	*
104.44.16.180	8075
193.238.116.5	*
207.241.160.1	11127
207.241.160.38	11127
81.85.101.89	11127

Validation (cont)

PeeringDB

Search here for a network, IX, or facility.

Advanced Search

Register or Login

NL-ix Silver Sponsor

Organization	NL-ix
Long Name	Neutral Internet Exchange
City	Amsterdam
Country	NL
Continental Region	Europe
Media Type	Ethernet
Protocols Supported	✓ Unicast IPv4 ○ Multicast ✓ IPv6
Last Updated	2020-08-04T14:01:37Z
Notes	

Contact Information

Company Website	https://www.nl-ix.net/
Traffic Stats Website	https://www.nl-ix.net/traffic.php
Technical Email	support@nl-ix.net
Technical Phone	+31703120710
Policy Email	info@nl-ix.net
Policy Phone	

LAN

DOT1Q	○
MTU	1500
IX-F Member Export URL Visibility	Private

Peers at this Exchange Point

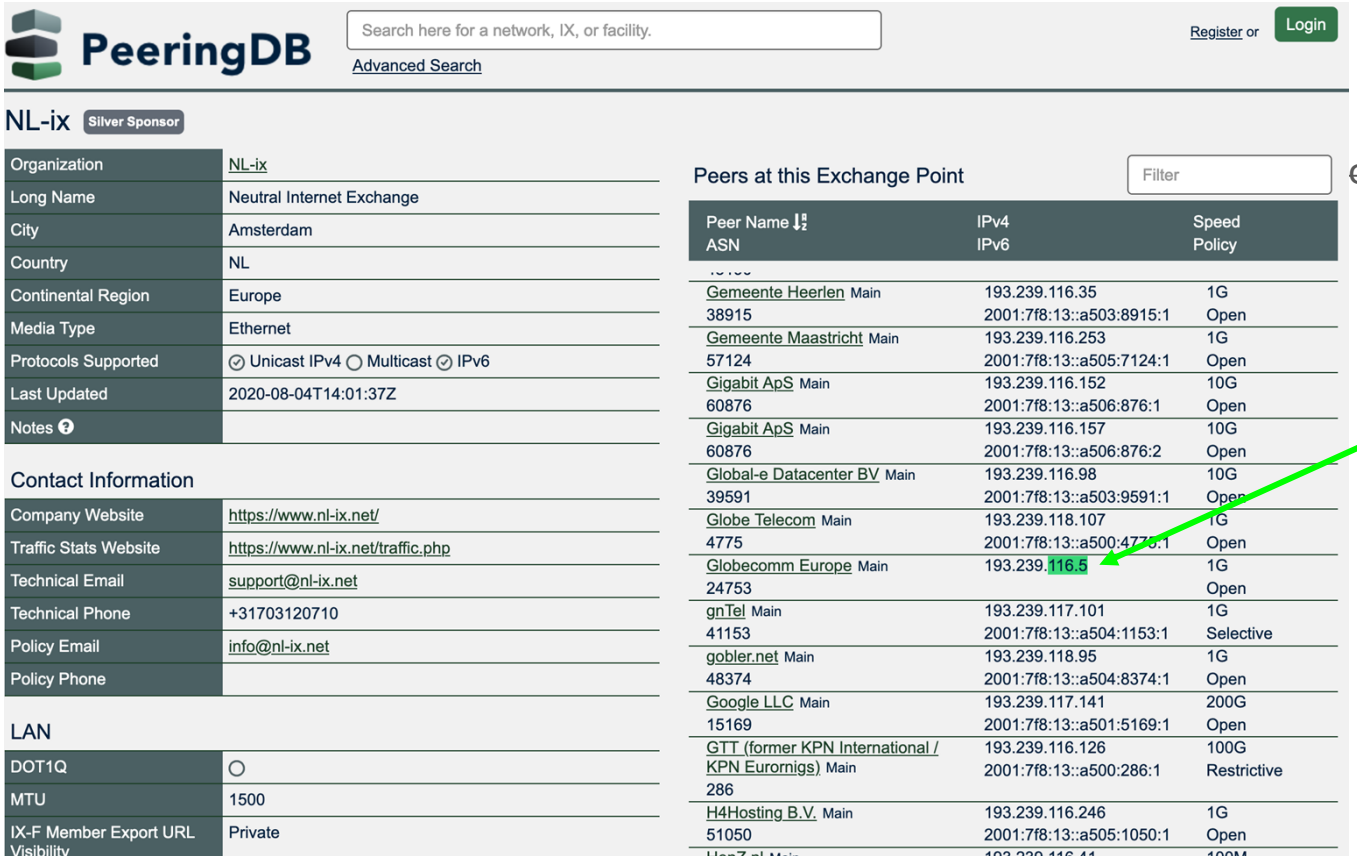
Filter

Peer Name / ASN	IPv4 / IPv6	Speed / Policy
Gemeente Heerlen Main 38915	193.239.116.35 2001:7f8:13::a503:8915:1	1G Open
Gemeente Maastricht Main 57124	193.239.116.253 2001:7f8:13::a505:7124:1	1G Open
Gigabit ApS Main 60876	193.239.116.152 2001:7f8:13::a506:876:1	10G Open
Gigabit ApS Main 60876	193.239.116.157 2001:7f8:13::a506:876:2	10G Open
Global-e Datacenter BV Main 39591	193.239.116.98 2001:7f8:13::a503:9591:1	10G Open
Globe Telecom Main 4775	193.239.118.107 2001:7f8:13::a500:4775:1	1G Open
Globecomm Europe Main 24753	193.239.116.5	1G Open
gnTel Main 41153	193.239.117.101 2001:7f8:13::a504:1153:1	1G Selective
gobler.net Main 48374	193.239.118.95 2001:7f8:13::a504:8374:1	1G Open
Google LLC Main 15169	193.239.117.141 2001:7f8:13::a501:5169:1	200G Open
GTT (former KPN International / KPN Eurornigs) Main 286	193.239.116.126 2001:7f8:13::a500:286:1	100G Restrictive
H4Hosting B.V. Main 51050	193.239.116.246 2001:7f8:13::a505:1050:1	1G Open

~~or~~

IP Address	ASN
*	*
104.44.16.180	8075
193.238.116.5	*
207.241.160.1	11127
207.241.160.38	11127
81.85.101.89	11127

Validation (cont)



or

IP Address	ASN
*	*
104.44.16.180	8075
193.238.116.5	24753
207.241.160.1	11127
207.241.160.38	11127
81.85.101.89	11127

Validation (cont)

- ~~Initial methodology~~
  - ~~Allow a single unresponsive/unresovled hop between cloud and neighbor; assume the AS immediately following this hop was a neighbor~~
  - ~~~50% FDR and 23 – 50% FNR~~
- Evaluated traceroute that caused false positives
  - Some had timeouts for intermediate hop – not possible to resolve
  - Most had intermediate hop IP
    - AS not resolved by Cymru using BGP
    - Registered in whois to an IXP – examined PeeringDB
  - Microsoft: FDR 8%, FNR 34%, 2,982 neighbors

IP Address	ASN
*	*
104.44.16.180	8075
193.238.116.5	24753
207.241.160.1	11127
207.241.160.38	11127
81.85.101.89	11127

Validation (cont)

- ~~Initial methodology~~
  - ~~Allow a single unresponsive/unresovled hop between cloud and neighbor; assume the AS immediately following this hop was a neighbor~~
  - ~~~50% FP and 23–50% FN~~
- Evaluated traceroute that caused false positives
  - Some had timeouts for intermediate hop – not possible to resolve
  - Most had intermediate hop IP
    - Not resolved in BGP
    - Registered in whois to an IXP -- examined PeeringDB
  - Microsoft: FDR 8%, FNR 34%, 2,982 neighbors
- Additional VM locations
  - Cloud providers have neighbors only available at specific locations
  - Microsoft: 16% FDR, 24% FNR, 3,619 neighbors

Validation (cont)

- ~~Initial methodology~~
  - ~~Allow a single unresponsive/unresovled hop between cloud and neighbor; assume the AS immediately following this hop was a neighbor~~
  - ~~~50% FDR and 23-50% FNR~~
- Evaluated traceroute that caused false positives
  - Some had timeouts for intermediate hop – not possible to resolve
  - Most had intermediate hop IP
    - Not resolved in BGP
    - Registered in whois to an IXP -- examined PeeringDB
  - Microsoft: FDR 8%, FNR 34%, 2,982 neighbors
- Additional VM locations
  - Cloud providers have neighbors only available at specific locations
  - Microsoft: 16% FDR, 24% FNR, 3,619 neighbors
- Prefer PeeringDB
  - Other IPs may have resolved using Cymru, but have more specific entry in PeeringDB
  - Microsoft: 11% FDR, 21% FNR, 3,565 neighbors
  - Google: 15% FDR, similar FNR, 7,554 neighbors

Validation (cont)

- ~~Initial methodology~~
  - ~~Allow a single unresponsive/unresovled hop between cloud and neighbor; assume the AS immediately following this hop was a neighbor~~
  - ~~~50% FDR and 23-50% FNR~~
- Evaluated traceroute that caused false positives
  - Some had timeouts for intermediate hop – not possible to resolve
  - Most had intermediate hop IP
    - Not resolved in BGP
    - Registered in whois to an IXP -- examined PeeringDB
  - Microsoft: FDR 8%, FNR 34%, 2,982 neighbors
- Additional VM locations
  - Cloud providers have neighbors only available at specific locations
  - Microsoft: 16% FDR, 24% FNR, 3,619 neighbors
- Prefer PeeringDB
  - Other IPs may have resolved using Cymru, but have more specific entry in PeeringDB
  - Microsoft: 11% FDR, 21% FNR, 3,565 neighbors
  - Google: 15% FDR, similar FNR, 7,554 neighbors

IP Address	ASN
*	*
104.44.16.180	8075
193.238.116.5	34307
207.241.160.1	11127
207.241.160.38	11127
81.85.101.89	11127

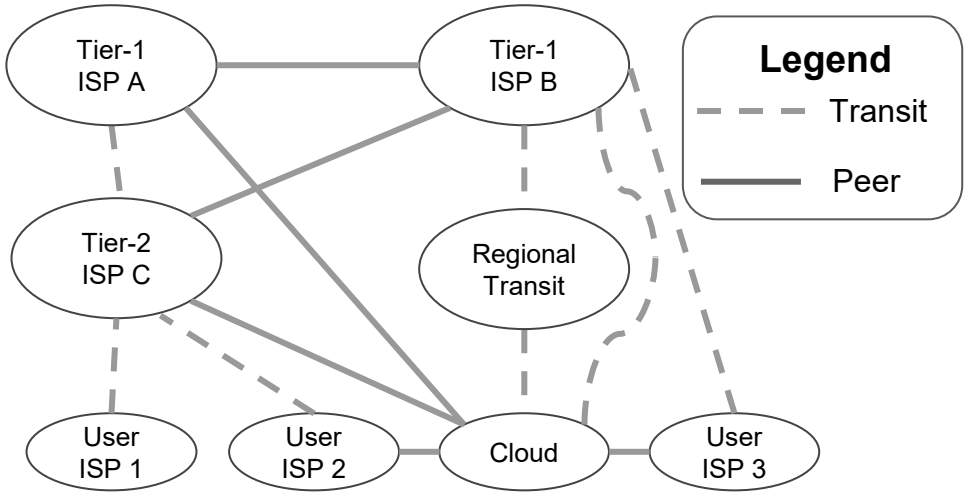
Validation (cont)

- ~~Initial methodology~~
  - ~~Allow a single unresponsive/unresovled hop between cloud and neighbor; assume the AS immediately following this hop was a neighbor~~
  - ~~~50% FDR and 23-50% FNR~~
- Evaluated traceroute that caused false positives
  - Some had timeouts for intermediate hop – not possible to resolve
  - Most had intermediate hop IP
    - Not resolved in BGP
    - Registered in whois to an IXP -- examined PeeringDB
  - Microsoft: FDR 8%, FNR 34%, 2,982 neighbors
- Additional VM locations
  - Cloud providers have neighbors only available at specific locations
  - Microsoft: 16% FDR, 24% FNR, 3,619 neighbors
- Prefer PeeringDB
  - Other IPs may have resolved using Cymru, but have more specific entry in PeeringDB
  - Microsoft: 11% FDR, 21% FNR, 3,565 neighbors
  - Google: 15% FDR, similar FNR, 7,554 neighbors

IP Address	ASN
*	*
104.44.16.180	8075
193.238.116.5	24753
207.241.160.1	11127
207.241.160.38	11127
81.85.101.89	11127

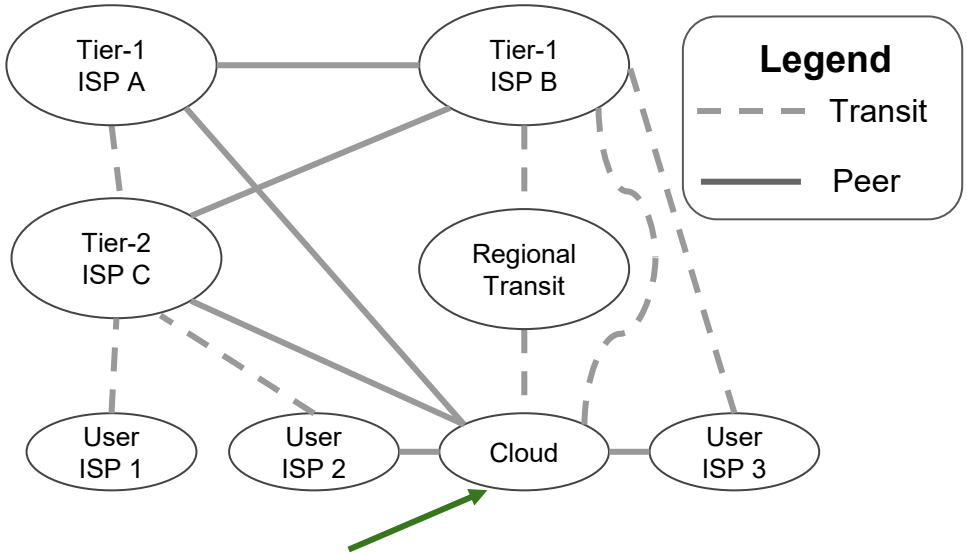
Calculating Hierarchy-free Reachability

- Calculate reachability propagating announcements through customers and peers, but not
  - Tier-1 ISPs
  - Tier-2 ISPs
- Reachability
  - If AS B receives a prefix announcement from the cloud, AS B is reachable by the cloud
  - Announcing AS called the *origin*
  - Uses augmented topology
  - Enforces common routing policies
- ***Hierarchy-free Reachability***
  - Count of reachable ASes when using peer links and not the hierarchical Internet



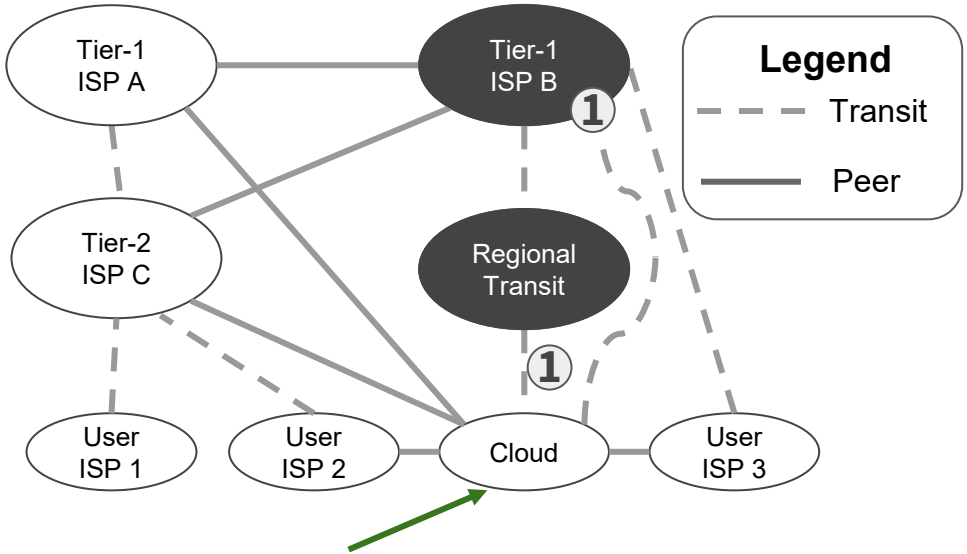
Calculating Hierarchy-free Reachability

- Calculate reachability propagating announcements through customers and peers, but not
  - Tier-1 ISPs
  - Tier-2 ISPs
- Reachability
  - If AS B receives a prefix announcement from the cloud, AS B is reachable by the cloud
  - Announcing AS called the *origin*
  - Uses augmented topology
  - Enforces common routing policies
- ***Hierarchy-free Reachability***
  - Count of reachable ASes when using peer links and not the hierarchical Internet



Calculating Hierarchy-free Reachability

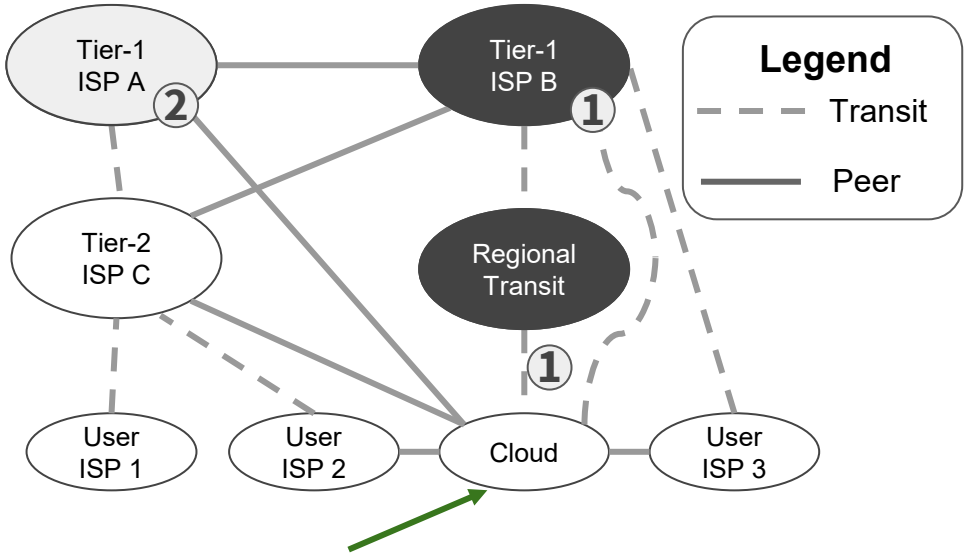
- Calculate reachability propagating announcements through customers and peers, but not
  - Tier-1 ISPs
  - Tier-2 ISPs
- Reachability
  - If AS B receives a prefix announcement from the cloud, AS B is reachable by the cloud
  - Announcing AS called the *origin*
  - Uses augmented topology
  - Enforces common routing policies
- ***Hierarchy-free Reachability***
  - Count of reachable ASes when using peer links and not the hierarchical Internet



Do not propagate routes via
(1) Origin's transit providers

Calculating Hierarchy-free Reachability

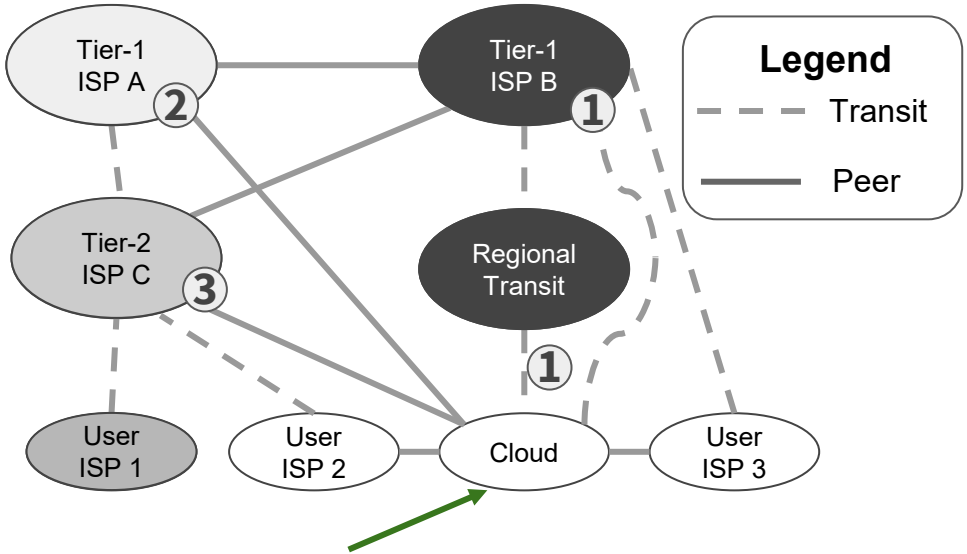
- Calculate reachability propagating announcements through customers and peers, but not
  - Tier-1 ISPs
  - Tier-2 ISPs
- Reachability
  - If AS B receives a prefix announcement from the cloud, AS B is reachable by the cloud
  - Announcing AS called the *origin*
  - Uses augmented topology
  - Enforces common routing policies
- ***Hierarchy-free Reachability***
  - Count of reachable ASes when using peer links and not the hierarchical Internet



Do not propagate routes via
(1) Origin's transit providers
(2) Tier 1 ISPs

Calculating Hierarchy-free Reachability

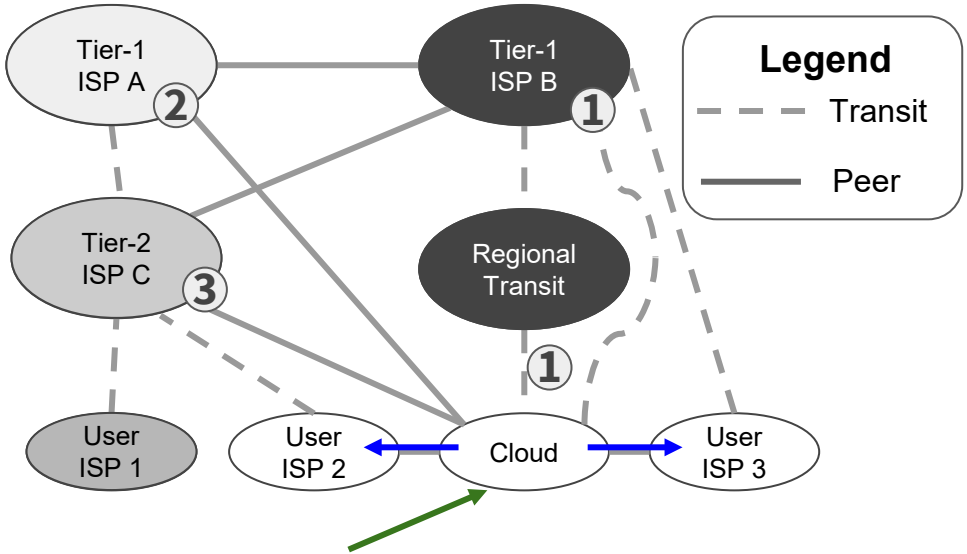
- Calculate reachability propagating announcements through customers and peers, but not
  - Tier-1 ISPs
  - Tier-2 ISPs
- Reachability
  - If AS B receives a prefix announcement from the cloud, AS B is reachable by the cloud
  - Announcing AS called the *origin*
  - Uses augmented topology
  - Enforces common routing policies
- ***Hierarchy-free Reachability***
  - Count of reachable ASes when using peer links and not the hierarchical Internet



Do not propagate routes via
(1) Origin's transit providers
(2) Tier 1 ISPs
(3) Tier 2 ISPs

Calculating Hierarchy-free Reachability

- Calculate reachability propagating announcements through customers and peers, but not
  - Tier-1 ISPs
  - Tier-2 ISPs
- Reachability
  - If AS B receives a prefix announcement from the cloud, AS B is reachable by the cloud
  - Announcing AS called the *origin*
  - Uses augmented topology
  - Enforces common routing policies
- ***Hierarchy-free Reachability***
  - Count of reachable ASes when using peer links and not the hierarchical Internet

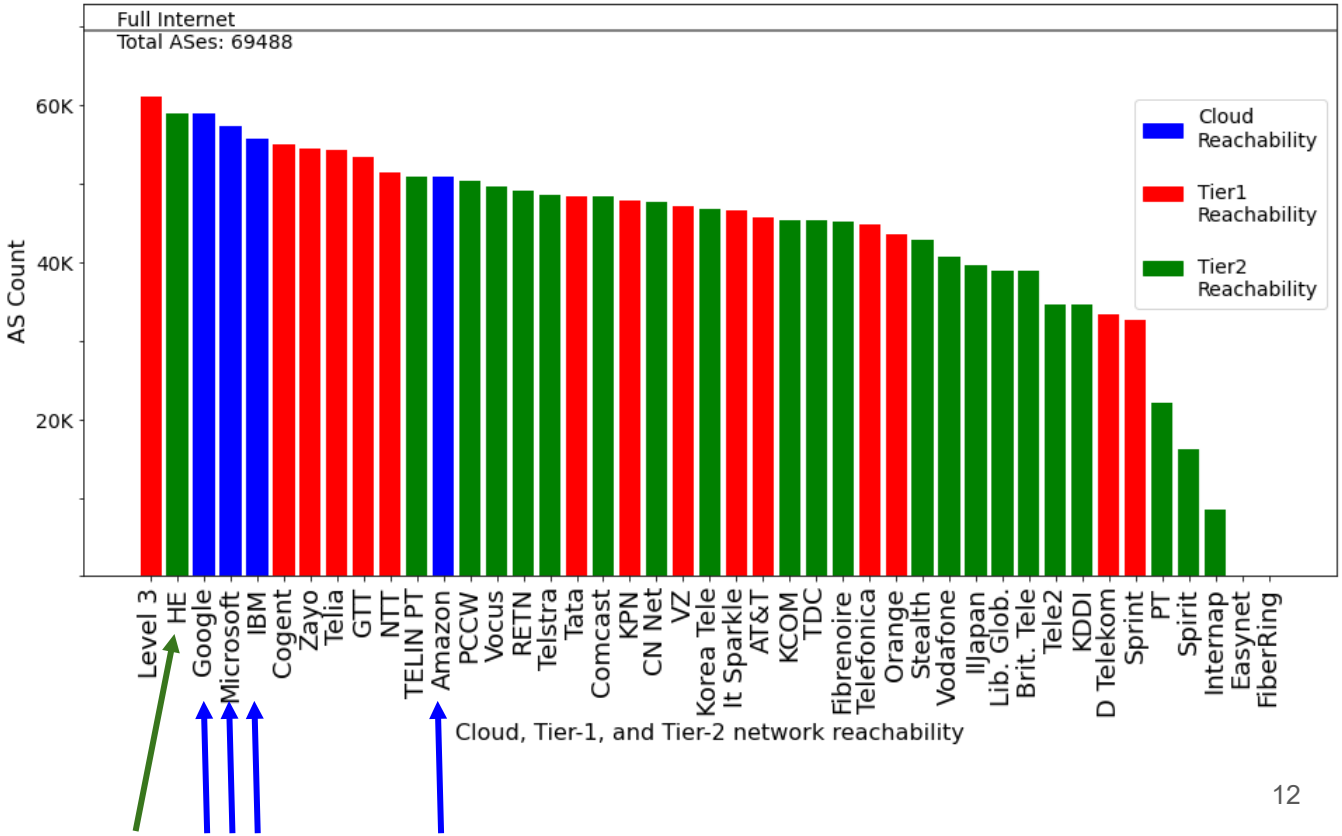


Do not propagate routes via
(1) Origin's transit providers
(2) Tier 1 ISPs
(3) Tier 2 ISPs

Hierarchy-free Reachability Results

Takeaway

- Cloud providers have higher reachability than most Tier 1 and Tier 2 ISPs
- They are able to reach the majority of networks even when bypassing their transit providers, Tier 1 ISPs, and Tier 2 ISPs.



Conclusions

- Emulated connectivity using an AS-level topology graph constructed from
  - BGP data
  - Traceroutes
  - Validated cloud neighbor lists
- Hierarchy-free Reachability quantifies the extent of Internet flattening and a network's potential to bypass the Internet hierarchy
- Show that thousands of networks benefit from flattening
  - Insights that are not captured by other metrics of measurements of a network's importance
  - The cloud providers are better able to bypass the hierarchy than most other networks, including the Tier 1 and Tier 2 ISPs