Poster: Beyond Proximity: Exploring Remote Cloud Peering

Esteban Carisimo
esteban.carisimo@northwestern.edu
Northwestern University
Evanston, Illinois, USA

Mia Weaver
mweaver6@wisc.edu
University of Wisconsin-Madison
Madison, Wisconsin, USA

Fabián E. Bustamante
fabianb@cs.northwestern.edu
Northwestern University
Evanston, Illinois, USA

Paul Barford
pb@cs.wisc.edu
University of Wisconsin-Madison
Madison, Wisconsin, USA

ABSTRACT

We investigate network peering location choices, focusing on whether networks opt for distant peering sites even when nearby options are available. We conduct a network-wide cloud-based traceroute campaign using virtual machine instances from four major cloud providers to identify peering locations and calculate the "peering stretch": the extra distance networks travel beyond the nearest data center to their actual peering points. Our results reveal a median peering stretch of 300 kilometers, with some networks traveling as much as 6,700 kilometers. We explore the characteristics of networks that prefer distant peering points and the potential motivations behind these choices, providing insights into digital sovereignty and cybersecurity implications.

CCS CONCEPTS

• **Networks → Network measurement**.

KEYWORDS

Cloud Computing, Remote Peering, Long-Haul Links (LHLs)

ACM Reference Format:

Esteban Carisimo, Mia Weaver, Fabián E. Bustamante, and Paul Barford. 2024. Poster: Beyond Proximity: Exploring Remote Cloud Peering. In *ACM SIGCOMM 2024 Conference (ACM SIGCOMM Posters and Demos '24), August 4–8, 2024, Sydney, NSW, Australia.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3672202.3673715

1 INTRODUCTION

The last decade has radically changed the Internet structure, with large cloud providers emerging as central components of a densely connected topology [1, 3, 6, 10, 12, 17, 18].

The change has come with, and as a result of, the global expansion of cloud providers' footprints. Large providers, such as Amazon, Google, IBM, and Microsoft, have deployed data centers and Points of Presence (PoPs) in virtually every region in the world, nearly doubling their geographic footprint in just five years as they become the source and destination of the majority of today's Internet traffic [2, 13, 15, 22, 26].

This impressive expansion means that most access networks around the world are now a few hundred kilometers away from cloud-provider datacenters – our preliminary analysis shows that half the networks [14] are less than ≈800 km (or ≈500 miles) from a cloud datacenter!

The expansion should prompt a shift in the places where networks peer with cloud providers, from early, faraway locations to proximate ones. This shift could reduce transit costs, enhance control over routing, and enable latency-sensitive applications [7]. Nevertheless, networks may still opt for remote peering locations due to factors such as cost-effectiveness [5, 25], the prospect of connecting with other networks [1, 4], or simple inertia (e.g., preexisting IRU agreements [8, 9, 24]).

Our work explores whether the availability of closer peering options leads to a preference for closer peering. Specifically, we are interested in understanding if networks choose to travel to a distant peering location to peer with cloud providers despite the availability of nearby options and which networks choose to do so.

We conduct a cloud-based traceroute campaign to identify the networks peering with the cloud and their peering locations. We set up virtual machine instances in all regions available from four large cloud providers (Amazon Web Services, Microsoft's Azure, Google Cloud Platform, and IBM Cloud Services) and launched a network-wide traceroute campaign. We combine the collected data with additional network datasets and apply state-of-the-art tools to identify networks' peering points with the cloud.

To measure the additional distance covered by a network from its nearest datacenter to its current peering location, we introduce a new metric: *peering stretch*. This metric, constructed based on a simple model of a network's peering point options, captures the difference between the geographic distances from the network to its potential nearest peering point and its actual peering point. We explore the characteristics of networks that establish peering connections with cloud providers at faraway locations, the popularity of these options across continents and countries, and the preferred destinations and providers for these peerings.

We combine our topological findings with additional data sources to explore possible motivations of these peerings, including a preference for locations with more cost-effective routes, richer peering opportunities, and access to specific content. Our analysis offers additional insights that could explain the persistent preference for

ACM SIGCOMM Posters and Demos '24, August 4–8, 2024, Sydney, NSW, Australia

ACM ISBN 979-8-4007-0717-9/24/08
https://doi.org/10.1145/3672202.3673715

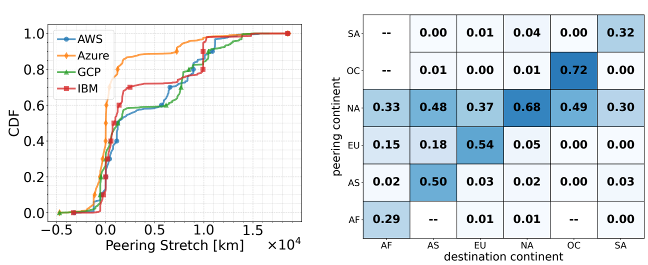
peering at distant locations despite the growing number of closer peering locations.

2 METHODOLOGY

We launched a traceroute campaign from the cloud to obtain a cloud-side global perspective of *who* and *where* with the cloud. To that end, we deployed virtual machine (VM) instances across all regions of four major cloud providers – Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft's Azure, and IBM Cloud – to serve as vantage points. We selected datacenter locations to maximize geographic diversity, focusing on metro areas with multiple cloud providers for comparative analysis and including often underrepresented regions such as South America, the Middle East, and Africa. Our selection criteria opted for entry-level VMs, given that our traceroute campaign was neither computing nor storage intensive. We probed all prefixes visible from RouteViews as of March 27, 2022, and September 29, 2023, using a /24 prefix granularity, a common practice in network measurement that aligns with the peering policies of Google, Amazon, and IBM, which do not accept more specific prefixes. Our traceroute measurements were conducted using Scamper to send ICMP packet probes at a rate of 1000 packets per second. Our first campaign collected 42.5 million traceroutes from these diverse global points, while our second run discovered a growing opacity, especially in Microsoft's Azure network, limiting our ability to contrast both snapshots.

We process traceroute data to extract interdomain router interfaces and geographical information using bdrmapIT [20] and geolocation databases to detect networks peering with the cloud and their locations. Considering the known limitations of geolocation methods, we utilize multiple datasets and heuristics, including HOIHO [19], MaxMind [21], and IPinfo [16] and we also verify the validity of these locations by ensuring they comply with speed-of-light constraints. We utilize multiple topological and geographical data sources to increase the accuracy of gelocation and also consult PeeringDB to validate the presence of cloud providers and peers partially. We finally utilize these results to compute the peering stretch of all networks peering with the cloud.

3 RESULTS



(a) (b)

Figure 1: Fig. 1a shows the *Peering stretch* for networks connecting to large cloud providers. Fig. 1b presents the peering matrix that shows the fraction of continent eyeballs (columns) that peer with the cloud in a given continent (rows).

We present a preliminary analysis of the extra distance traveled by a network from the alternative to its actual peering location. To this end, we introduce a new metric, peering stretch. We define *peering stretch* as the difference between great-circle distances from traceroute destinations to peering points and to the nearest data centers. While this simple model obviates geographic barriers (e.g., deserts, mountains), diplomatic tensions, and other factors that may prevent the use of closer locations, it nevertheless provides a first approximation of the overhead opted by a network peering at a distant location.

Figure 1a presents the cumulative distribution function (CDF) of peering stretch for specific prefixes within cloud peers, focusing on a subset advertised by these networks. The figure displays data for four major cloud providers: Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft Azure, and IBM Cloud.

Our analysis shows considerable variance in the peering distances with these providers. Notably, up to 72% of prefixes at the nearest cloud peers are within 500 km, with Azure showing the highest proximity. In contrast, over 40% of the network paths to Google and AWS extend beyond 5,000 km. These longer distances reflect the cloud providers' global reach and the strategic choices of networking routes, existing remote peering facilities or Indefensible Rights of Use (IRUs).

Where Do Peers Originate and Where Do They Connect? Figure 1b shows a heatmap of peering points by continent (rows) against traceroute destinations (columns), showing the fraction of the Internet population for each intersection.

The heatmap indicates that while most Internet populations primarily peer within their own continents, many also connect remotely, especially outside North America. North America is the predominant remote peering destination, attracting significant traffic from key networks like Bharti Airtel and China Telecom. Angola Cables is a notable example of the interest in peering in the US, as the company invested in deploying a transoceanic cable to Brazil, aiming to reduce latency and improve connectivity from Angola to Miami [11, 23]. Europe is another significant destination, especially for networks from Africa and Asia, supported by connectivity through cables like ACE and SeaMeWe-4. Outside these regions, remote peering is minimal.

4 ACKNOWLEDGEMENTS

This work is supported by the National Science Foundation grants CNS-2107392, CNS-2312709, CNS-2319367, and CNS-2106517. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] Bernhard Ager, Nikolaos Chatzis, Anja Feldmann, Nadi Sarrar, Steve Uhlig, and Walter Willinger. 2012. Anatomy of a large European IXP. In *Proc. of ACM SIGCOMM.*
[2] Amazon Web Services. [n. d.]. AWS Global Infrastructure. https://aws.amazon.com/about-aws/global-infrastructure/.
[3] Todd Arnold, Jia He, Weifan Jiang, Matt Calder, Italo Cunha, Vasileios Giotsas, and Ethan Katz-Bassett. 2020. Cloud Provider Connectivity in the Flat Internet. In *Proc. of IMC.*
[4] Ignacio Castro, Juan Camilo Cardona, Sergey Gorinsky, and Pierre Francois. 2014. Remote Peering: More Peering without Internet Flattening. In *Proc. of CoNEXT.*

[5] Ignacio Castro, Rade Stanojevic, and Sergey Gorinsky. 2014. Using Tuangou to Reduce IP Transit Costs. *IEEE/ACM Transactions on Networking* (2014).

[6] Nikolaos Chatzis, Georgios Smaragdakis, Anja Feldmann, and Walter Willinger. 2013. There is more to IXPs than meets the eye. *ACM SIGCOMM Computer Communication Review* (2013).

[7] Lorenzo Corneo, Maximilian Eder, Nitinder Mohan, Aleksandr Zavodovski, Suzan Bayhan, Walter Wong, Per Gunningberg, Jussi Kangasharju, and Jörg Ott. 2021. Surrounded by the Clouds: A Comprehensive Cloud Reachability Study. In *Proc. of the WWW*.

[8] Telefonica de Colombia. 2018. Informe Anual. https://descubre.movistar.co/informe-de-gestion-responsable-2018/estados_financieros/Telefonica_Informe2018_.pdf.

[9] Telefonica de Colombia. 2022. Informe Anual. https://www.telefonica.co/wp-content/uploads/sites/4/2023/03/Informe-Gestion-Responsable-2022-Reporte-Gestion-BIC.pdf.

[10] Amogh Dhamdhere and Constantine Dovrolis. 2010. The Internet is flat: modeling the transition from a transit hierarchy to a peering mesh. In *Proc. of CoNEXT*.

[11] Rodérick Fanou, Pierre Francois, and Emile Aben. 2015. On the diversity of interdomain routing in africa. In *Proc. of PAM*.

[12] Vasileios Giotsas, Shi Zhou, Matthew Luckie, and kc claffy. 2013. Inferring Multilateral Peering. In *Proc. of CoNEXT*.

[13] Google Cloud. [n. d.]. Cloud locations. https://cloud.google.com/about/locations.

[14] Geoff Huston. 2014. How Big is that Network? https://labs.apnic.net/?p=526.

[15] IBM Cloud. [n. d.]. Regions and locations. https://cloud.ibm.com/docs/hs-crypto?topic=hs-crypto-regions.

[16] ipInfo. 2023. ipInfo. https://ipinfo.io

[17] Craig Labovitz, Scott Iekel-Johnson, Danny McPherson, Jon Oberheide, and Farnam Jahanian. 2010. Internet inter-domain traffic. In *Proc. of ACM SIGCOMM*.

[18] Aemen Lodhi, Natalie Larson, Amogh Dhamdhere, Constantine Dovrolis, and Kc Claffy. 2014. Using peeringDB to understand the peering ecosystem. *ACM SIGCOMM Computer Communication Review* (2014).

[19] M. Luckie, B. Huffaker, A. Marder, Z. Bischof, M. Fletcher, and k. claffy. 2021. Learning to Extract Geographic Information from Internet Router Hostnames. In *Proc. of CoNEXT*.

[20] Alexander Marder, Matthew Luckie, Amogh Dhamdhere, Bradley Huffaker, KC Claffy, and Jonathan M Smith. 2018. Pushing the boundaries with bdrmapit: Mapping router ownership at Internet scale. In *Proc. of IMC*.

[21] MaxMind. 2023. MaxMind Geolocation Data. https://www.maxmind.com/en/geoip2-services-and-databases.

[22] Microsoft Azure. [n. d.]. Azure geographies. https://azure.microsoft.com/en-us/explore/global-infrastructure/geographies/.

[23] Bradley Prior. 2018. Teraco data centres will benefit from SACS cable. https://mybroadband.co.za/news/cloud-hosting/284682-teraco-data-centres-will-benefit-from-sacs-cable.html.

[24] TELECOM ARGENTINA S.A. 2013. UNITED STATES SECURITIES AND EXCHANGE COMMISSION: FORM 20-F. https://www.sec.gov/Archives/edgar/data/932470/000119312514142506/d707745d20f.htm.

[25] Rade Stanojevic, Ignacio Castro, and Sergey Gorinsky. 2011. CIPT: Using Tuangou to Reduce IP Transit Costs. In *Proc. of CoNEXT*.

[26] Alan Weissberger. 2022. Sandvine: Google, Facebook, Microsoft, Apple, Amazon & Netflix generate almost 57% of Internet traffic. https://techblog.comsoc.org/2022/02/01/sandvine-google-facebook-microsoft-apple-amazon-and-netflix-generate-almost-57-of-internet-traffic/.