Take The Long Way Home - Distant Peering to the Cloud

Esteban Carisimo*, Mia Weaver†, Fabián E. Bustamante*, Paul Barford†
*Northwestern University †University of Wisconsin-Madison

***Abstract*—The emergence of large cloud providers in the last decade has transformed the Internet, resulting in a seemingly ever-growing set of datacenters, points of presence, and network peers. Despite the availability of closer peering locations, some networks continue to peer with cloud providers at distant locations, traveling thousands of kilometers. In this paper, we employ a novel cloud-based traceroute campaign to characterize the distances networks travel to peer with the cloud. This unique approach allows us to gain unprecedented insights into the peering patterns of networks. Our findings reveal that 50% of the networks peer within 300 kilometers of the nearest datacenter. However, our analysis also reveals that over 20% of networks travel at least 6,700 kilometers beyond the proximity of the nearest computing facility, and some as much as 18,791 kilometers! While these networks connect with the cloud worldwide, from South America to Europe and Asia, many come to peer with cloud providers in North America, even from Oceania and Asia. We explore possible motivations for the persistence of distant peering, discussing factors such as cost-effective routes, enhanced peering opportunities, and access to exclusive content.**

***Index Terms*—Cloud Computing, Remote Peering, Long-Haul Links**

I. Introduction

The last decade has radically changed the Internet structure, with large cloud providers emerging as central components of a densely connected topology [1], [2], [3], [4], [5], [6], [7].

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 [8], [9], [10], [11], [12].

This impressive expansion means that most access networks worldwide are now a few hundred kilometers away from cloud-provider datacenters. Figure 3 clearly illustrates this; it shows the distribution of distances to the closest datacenter for 57.5% of the Internet population [13] distributed over 175 countries around the world. Half the networks (hereafter we refer Autonomous Systems (ASes) as networks) are less than ≈800 km (or ≈500 miles) from a cloud datacenter (or ≈1200 km/750 miles for the 75^{th} percentile).

The same footprint 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 [14]. Nevertheless, networks may still opt for remote peering locations due to factors such as cost-effectiveness [15], [16], the prospect of connecting with other networks [17], [3], or simple inertia (e.g., preexisting IRU agreements [18], [19], [20]). This raises the question of 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 decide 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). We launched a network-wide traceroute campaign (§III). We combine the collected data with additional network datasets and apply state-of-the-art tools to identify networks' peering points with the cloud (§IV).

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. Across the networks in our study, we find a median *peering stretch* of 300 kilometers, meaning that most networks travel less than 300 additional kilometers from their nearest to their actual peering point. However, the distribution of the *peering stretch* shows that network on the upper 20% (80^{th} percentile) can travel at least 6,700 km to peer with the cloud.

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 (§V).

We combine our topological findings with additional data sources to explore possible motivations of these peerings (§VI), 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 peering at distant locations despite the growing number of closer peering locations.

In summary, we make the following key contributions:

- We carried out a large-scale analysis of peerings with large cloud providers from networks around the world. Despite the global presence of cloud providers, some networks still choose to travel to different continents, up to ≈19,000 kilometers away from their closest datacenter, to establish peering connections with cloud providers.

- We investigate the characteristics of networks with high peering stretches. We discover variations in preferences for traveling long distances to peer with cloud providers, with almost no adoption in North America but significant adoption in South American and Asian countries, with networks serving 30% and 48% of the Internet populations, respectively. We explore the distances traveled from different continents and the preferred destinations and providers of these peerings.
- Finally, we explore possible explanations behind these preferences. We find that several networks have established presence at distant locations before the rise of cloud computing suggesting an *inertial behavior* as a possible explanation for the selection of peering locations. For example, Telefonica's subsidiary in Colombia has disclosed in its annual reports its subscription to multi-year IRU agreements, some of which extend until 2030 [18], [19].

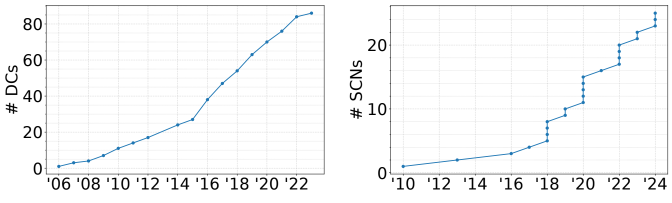
Our findings have clear implications for studies of digital sovereignty and cybersecurity, particularly on the criticality of submarine cable infrastructure [21], [22], [23].

This work does not raise any ethical issues.

II. Peering with an Expanding Cloud

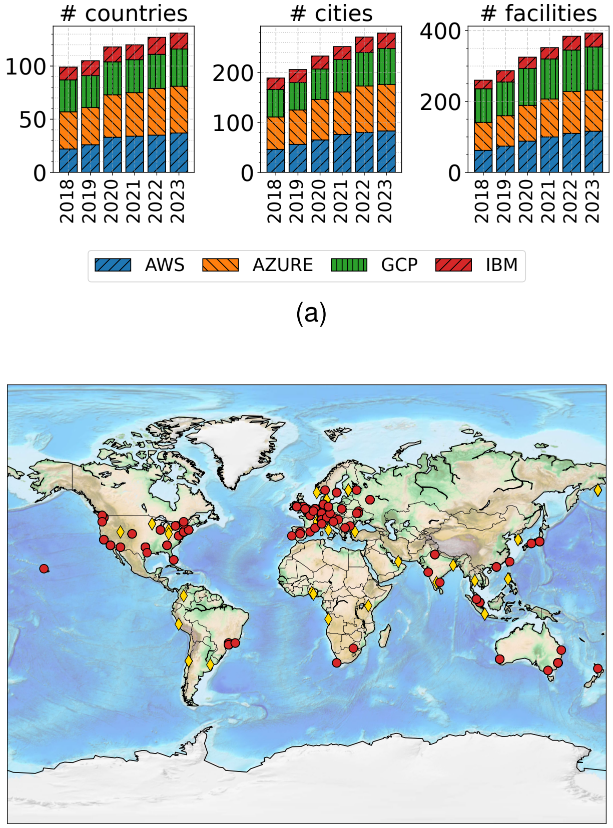
The rise of cloud-based Internet services has been accompanied by a notable expansion of the physical reach of cloud networks.

The expansion has come with a $5x$ increase in the number of datacenters between 2013 and 2023, among the top-3 providers, and the construction of numerous submarine cables from their first cable becoming operational in 2010 to an expected partial or full ownership of 25 submarine cables by 2024. Figure 1 illustrates the combined growth of Amazon Web Services (AWS), Microsoft's Azure and Google Cloud Platform (GCP)[1] in number of datacenters (Fig. 1a) and submarine cable ownership (Fig. 1b).



(a) (b)

Fig. 1. Evolution of AWS, Azure, and GCP infrastructure, shown by the growth in datacenter locations (Fig.1a) and the expansion of submarine cable ownership (Fig.1b) over the years.



(a)

(b)

Fig. 2. Geographical expansion of AWS, Azure, GCP, and IBM's Points of Presence (PoPs) across country, city, and facility levels from 2018 to 2023, as shown in Fig.2a. Fig.2b displays their presence at peering facilities, with gold diamonds representing presence at locations in 2018 and red circles representing their presence across 2018 and 2023. For visual purposes, each marker represents a 100kmx100km cell.

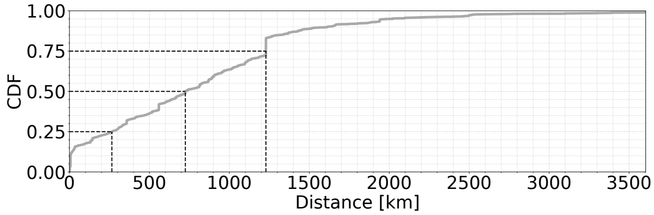


Fig. 3. Distance (km) from the networks found in this work to the closest datacenter.

Combined with leased and terrestrial infrastructure, these submarine deployments multiplied the access points to cloud infrastructure across peering facilities in all regions. Cloud providers can directly peer with networks through these extensive infrastructures, circumventing the public Internet [1], reducing latency, and minimizing network congestion [7]. According to PeeringDB, these cloud providers have collectively expanded their footprint at peering facilities from 150 to 244 in 2018-2023, with PoPs in 17 new cities and 11 new countries (Fig.2).

With cloud providers having computing resources and ingress points virtually everywhere, most networks can find a nearby peering point with the cloud. Figure 3 illustrates this point, plotting the distribution of distances to the closest datacenter for all the 1,928 networks included as part of our

[1] While having a comparatively smaller infrastructure, we were unable to find any information on datacenters or submarine cable growth for IBM.

study – networks around the world are, on average, ≈800km away from the nearest peering point (less than 1200km at the 80th pct). The impressive expansion motivates our work.

III. Measurement campaign

As vantage points for our measurement campaign, we use virtual machine (VM) instances placed in all regions of four major cloud providers – Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft's Azure, and IBM Cloud.

We selected datacenters locations that maximize the geographic diversity of the vantage points, prioritizing metro areas or regions with multiple cloud providers to enable comparative analysis. Besides these areas, our deployment includes vantage points in South America, the Middle East and Africa, regions typically underrepresented in network measurement studies. Figure 4 shows, on a world map, all the metro areas we selected to place our vantage points (the number of VMs per region): North America (13), South America (5), Europe (10), Africa (1), Middle East (2), Asia (19) and Oceania (4).

We opted for the entry-level VM available for each provider and region, since our measurement campaign is neither computing nor storage intensive. Appendix A details the VM specifications used in each cloud provider.

We ran a network-wide traceroute campaign using a /24 prefix granularity to probe all prefixes visible from RouteViews [24] in the snapshot of March 27, 2022. This granularity aligns with previous works that ran traceroute campaigns from both the cloud [25], [26] and the edge [27], [28]. Prior work assumes that (cloud) networks will not receive prefixes more specific than /24 or their import policies would filter them out. Indeed, Google [29], Amazon [30] and IBM [31] peering policies explicitly state that they will not accept longer prefixes.

Following Ark's measurement design [28], we randomly selected hosts in each /24. We distributed the measurement load by slicing (after random shuffling) the number of probed destinations across all the vantage points of each cloud provider. We use Scamper [32] to collect our traceroute measurement using ICMP packet probes at a maximum rate of 1000 packets per second following the setup of previous measurement studies [7].

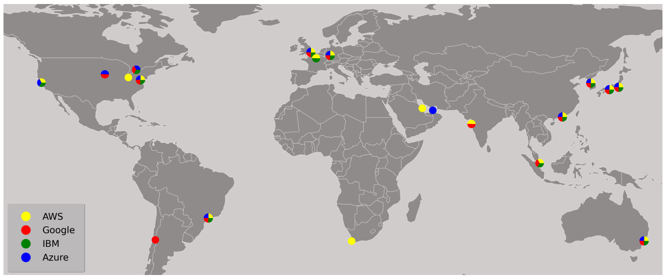


Fig. 4. Datacenter locations of the cloud providers in the study. Pie chart markers indicate regions with overlapping presence of providers.

In total, we collected 42.5M traceroutes, 12M from AWS, 11.7M from Azure, and 9.4M from each GCP and IBM Cloud. *We will make our dataset available to the research community.*

IV. Detecting Peerings to the Cloud

We now describe our methodology for identifying networks peering with the cloud and their peering locations. We process our traceroute dataset to extract two relevant features: (1) interdomain router interfaces (§IV-A), and (2) geographic information of both routers and destinations (§IV-B). To gain additional confidence in our findings, we consult external data sources that provide information on the presence of cloud providers and peers at different locations (§IV-D). We encounter challenges and limitations in inference methods and data sources, some of which are intrinsic to the cloud environment, and describe how we mitigate these issues.

A. IP-to-AS traceroute mappings

We apply bdrmapIT [33] to identify interdomain router interfaces and ASes along the traceroute paths.[2] The tool combines multiple topological data sources to infer ASes in the near and far side of each IP address of a traceroute sequence. Applying bdrmapIT to our dataset, we identify 394,211 unique interfaces interconnecting different pairs of networks. Our analysis focuses on border router interfaces that interconnect cloud providers with their network peers. These interfaces are found when the near side corresponds to a cloud AS and the far side to a non-cloud AS, after removing intercloud connectivity from our dataset.[3]

The existence of sibling ASes that are part of organizations operated by cloud providers, such as AS15619 (Google's primary ASN) and AS396982 (Google Cloud Platform ASN), challenges our methodology. We use $as2org+$ [34] to address this, a tool combining WHOIS and PeeringDB to identify sibling ASes associated with a given organization.

We generate custom decision rules to extend bdrmapIT inferences in cases where it is unable to produce inferences for some public IP address. We investigated such cases and discovered that some traceroute sequences contain IP addresses (e.g., 15.230.129.41) allocated to AWS, but not visible from public route collectors. Since bdrmapIT inferences rely on data derived from these public collectors, the near side of these interfaces is not inferred to belong to a network within Amazon's organization. In such cases, we override this inference and consider it as part of AWS as well.

We rely on APNIC's eyeball population estimates [13] at the AS level to weigh the relevance of networks peering with cloud providers. Network data is insufficient for this purpose given the mismatch between number of users and address spaces due to the widespread use of NAT [35], and the fact that some networks, such as academics, announce large legacy address blocks [36].

a) Challenges and limitations: The effectiveness of traceroute data for discovering AS interdomain relationships depends on ICMP responsiveness of border routers along the path. While the state-of-the-art bdrmapIT has made significant progress in inferring router ownership, it still suffers

[2] We use a containerized version of bdrmapIT that simplifies the setup and execution process. This container can be found: https://github.com/dioptra-io/docker-images/tree/main/bdrmapit

[3] Marder et al. [25] finds that cloud providers peer between them in the same city.

from some limitations that could affect our work. A recent study [37] has shown that bdrmapIT inferences are sensitive to the presence of "off-path" or third-party addresses in traceroute sequences. These inaccuracies could lead to erroneous inferences when discovering networks that engage in peering relationships with cloud providers. In the context of cloud networks, we have also encountered unique challenges and limitations. For instance, our measurements show AWS extensively uses a Carrier-Grade-NAT (CGN) reserved address pool (164.10.0.0/10, RFC6598 [38]), challenging the correct inference of router ownership. Additionally, as previous works have pointed out [25], Google Cloud Platform manipulates packet TTLs, which reduces traceroute visibility and likely compromises our ability to identify Google's peers.

Despite the opaque nature of cloud networks, our methodology successfully identified peering locations in every AWS and GCP region. Our resulting dataset provides a baseline of cloud-to-public-Internet peering.

B. IP-to-country mappings

We combine multiple databases and geolocation heuristics to identify the country- and continent-level locations of routers and destinations. Given that geolocation databases are known for containing some inaccuracies [39], [40], we take steps to improve confidence on the inferred locations.

For starters, we prioritize different sources differently as follows. We give priority to HOIHO geolocation inferences [41] when locating IP addresses of routers along the traceroute path, considering that HOIHO uses an extensively validated ruled-based inference technique based on (operator's assigned) geolocation hints embedded in DNS PTR records of router interfaces [41]. While offering high confidence, HOIHO coverage is limited to routers with DNS PTR records and for which HOIHO has extraction rules. For the remaining routers and destinations, we rely on MaxMind GeoLite2 [42] for geolocation, when corroborated by ipInfo [43].

We focus on a country-level granularity, shown to be more reliable [44]. If a border router, the interface of a cloud provider is within a 10 ms range of the datacenter from which the traceroute was launched, we override the router's assigned country as that of the datacenter.

As a last step to mitigate potential geolocation errors, we use speed-of-light constraints as a lower-bound to discard inferences containing latencies of distant peerings (between the datacenter and the border router) that are inconsistent with the minimum distance between both end-points.

Overall, from the collection of router and destinations in our traceroute collection, we geolocated 65.28% using MaxMind, 0.55% using IPInfo, and 34.17% using HOIHO.

a) Challenges and limitations: IP geolocation is a known challenge for infrastructure-based studies, and ours is no exception. To begin with, we are unable to geolocate IP addresses that belong to reserved address pools. Opting for an active-probing geolocation method, similar to the RIPE IPmap single-radium mechanism [45], to enhance geolocation accuracy faced several limitations, including allocated but non-announced prefixes that remain unreachable from probes outside cloud networks, and hops within the cloud that does not respond to ICMP Echo requests (pings)[4], essential for active probing.

Our study aims to provide a coarse-grained indicator of when networks peer with cloud providers in a different continent from where the prefix is geolocated. Despite the known accuracy issues of geolocation, they have been found to be generally reliable at the country level [46], [40]. This is sufficient given the goal of our study. In addition, we assume that inaccuracies within the same country are minimal, compared to the intercontinental distances involved, and do not change our main findings.

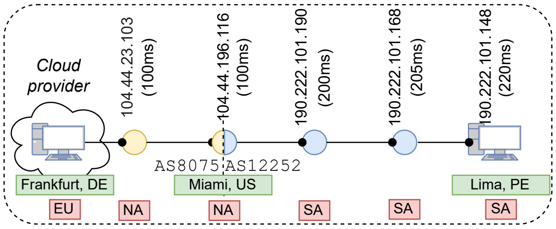


Fig. 5. Illustrative traceroute

C. Final selection

TABLE I
EXAMPLE OF THE FINAL DATA PRODUCT

Source	*Cloud:*	Azure
	Datacenter:	Frankfurt
	Country:	Germany
	Continent:	Europe
Destination	*ASname:*	America Movil Peru S.A.C.
	AS:	12252
	Country:	Peru
	Continent:	South America
Peering Point	*Far-side AS:*	12252
	Near-side AS:	8075
	Hop country:	United States
	Hop continent:	North America
	Hop rDNS:	`ae61-0.ier01.mia.ntwk.msn.net`

As the last step in our data processing, we identify cloud peers and the locations where they peer with their cloud providers. Our analysis only focuses on the peer-originated address space, as these networks control both the geographic regions where those prefixes are physically deployed and the peering point where that prefix is annouced to the network's peers, in this case the cloud.

Figure 5 uses an illustrative example traceroute to show the described analysis and its output. Figure 5 shows a traceroute launched from the Frankfurt datacenter of Azure, in Germany towards a /24 within AS12253 of America Movil Peru SAC. The analysis reveals a peering between AS8075 (Azure) and AS12252 (America Movil Peru) in Miami, US. Table I lists the associated output of our analysis, including AS name (e.g., 'America Movil Peru S.A.C.'), cloud provider (e.g., Azure), AS numbers at both ends of the border router (e.g., 8075

[4] Despite generating ICMP Time Exceeded in Transit messages that reveals the traceroute path.

and 12252 in the near and far sides, respectively), and the DNS PTR record (including the substring `mia`) of the border interface, among others.

D. Validating peering locations

We further confirm our inferences by consulting PeeringDB. While our methodology includes several steps to build confidence in geolocation inferences (§IV-B), external data sources offer a complementary perspective that increases confidence in our findings.

We compare our results with publicly available information on the presence of peers and cloud providers from PeeringDB. We argue that PeeringDB is a reliable source for validating public peering with cloud providers, as cloud providers and other large content providers require their peers to be listed on PeeringDB in order to establish peering relationships [47], [48], [29] at public peering locations.

Given the focus of our work– characterizing networks with high peering stretch – we validate those networks traveling to another continent to establish cloud peering relationships. We use data from peering facilities and Internet Exchange Points (IXPs) to determine whether both peers and cloud providers are present at the same facility in a country, as indicated by our findings. This analysis shows that only 3.17% do not match that information. Nearly 57% (56.92%) of the evaluated cases align with the information from PeeringDB. The remaining cases include 15.85% which are not registered in PeeringDB, while 24.06% are registered but do not disclose any presence.

V. Distant Peering to the Cloud

The methodology described in the preceding section results in a set of networks along with their respective peering locations with cloud providers. In this section, we focus on networks traveling far to peer with the cloud, their cloud-peering locations, and alternative closer locations where they might have potentially established peering.

We introduce a new metric, *peering stretch*, to quantify the extra distance traveled by a network from the alternative to its actual peering location (§V-A). We show a significant fraction of high-peering stretch across cloud providers and find that networks, on the 80^{th} percentile, travel as much as 6,700 kilometers beyond the distance to the nearest computing facility.

We then examine networks with high peering stretch, focusing on variations observed across different regions, and the preferred destinations and providers of these peerings. We conclude our analysis exploring the role of transit providers in delivering content to other continents (§V-C).

A. 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.[5]

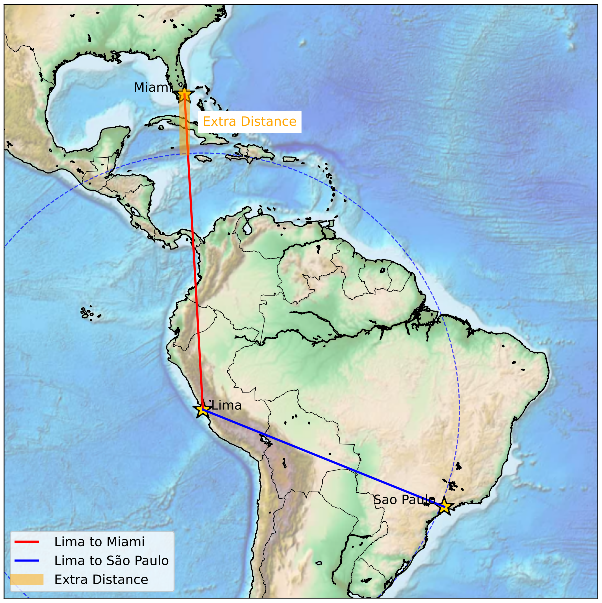


Fig. 6. Illustrative example of *Peering Stretch*: A prefix in Lima identifies Sao Paulo (blue line) as the closest peering location, but it actually peers in Miami (red line). The orange line represents the additional distance traveled when the closest available location is not selected.

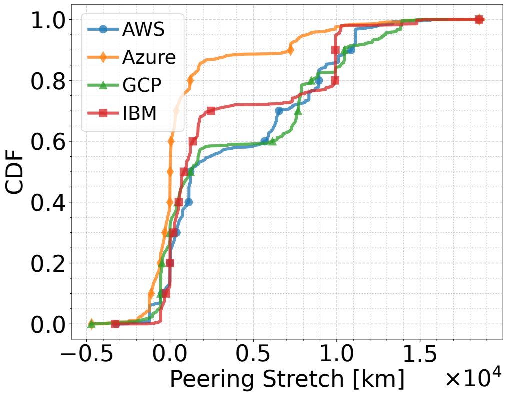


Fig. 7. *Peering stretch* for networks connecting to different large cloud providers.

Figure 6 illustrates the concept of peering stretch with an example of a network with a prefix in Lima, Peru. Although the closest peering location for this network is Sao Paulo, Brazil, the network instead peers with the same cloud provider in Miami. The peering stretch calculates the distance between the prefix and both points: the closest available location (blue) and the actual peering point (red). It then computes the extra

[5] Peering stretch, as defined here, represents a conservative estimate of the true peering stretch. when considering that peering occurs at a PoP geographically closer than the closest datacenter. Given its conservative and approximate nature, this metric can be slighly negative in case of close proximity between network locations, peering points and datacenters.

distance traveled to exchange traffic as a result of the peering policy decision (orange).

Figure 7 shows the cumulative distribution (CDF) of peering stretch for each destination prefix within a cloud peer. Note that these peerings are associated with a subset of the prefixes advertised by these networks. The figure includes curves for peerings related to each of the four cloud providers in our study – Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft's Azure and IBM Cloud.

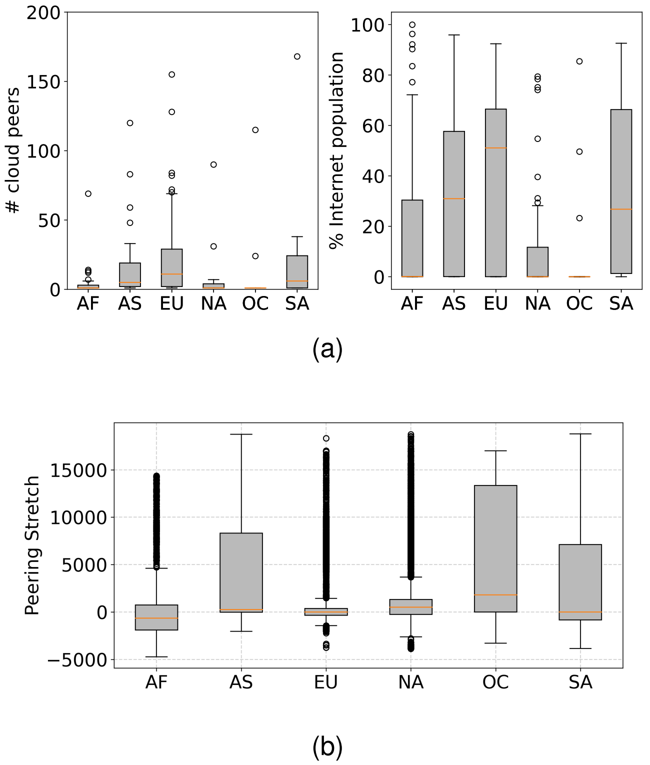
The analysis of the peering stretch reveals significant variation in the peering locations used to exchange traffic with cloud providers. A large percentage of prefixes in the cloud peers are accessible from close locations, with peering stretches below 500 km for 33%, 38%, 41%, and 72% of the fraction of all prefixes reachable from AWS, IBM, Google, and Azure, respectively. However, a notable large fraction of networks travels more than 5,000 km to peer with the cloud. This is the case for those peering with Google and AWS, where 41 and 42% of networks travel over 5,000 km to peer with them. It is important to note that the cloud providers footprint does not solely determine these high-peering stretches. Other factors, such as the specific networks that choose to peer with them and the existence of Indefeasible Rights of Use (IRU) – long-term contracts granting access to infrastructure, typically over 10 years – also influence peering decisions, along with other considerations.

We repeated this analysis for a traceroute campaign collected in October 2023 and found similar results that are detailed in Appendix A.

B. *Peers meeting the cloud*

Figure 7 shows a wide range of peering stretch with a standard deviation across all providers of ≈4,000 km and a range of ≈18,000 kms. In the following paragraphs, we explore the networks incurring high peering stretch and the user population they capture using an estimate of Internet populations offered by APNIC eyeball [13], the distances traveled within and across continents, and the preferred destinations and providers of these peerings.

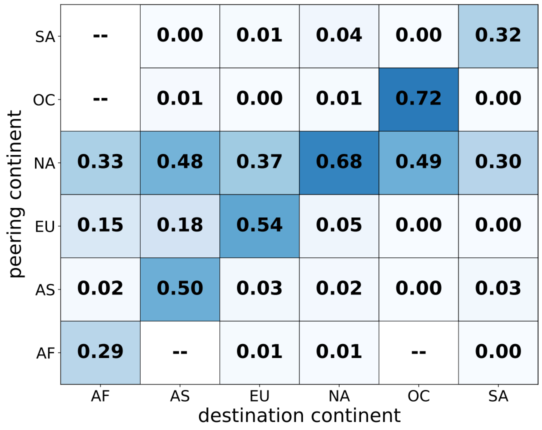
a) Who is peering with the cloud? A regional view: Figure 8 presents, per continent, the number of eyeball networks peering with the cloud, the percentage of users they capture, and their associated peering stretch. Figure 8a shows a boxplot of the aggregated number of eyeball networks that peers with the cloud (left) and the estimated Internet population they capture, aggregated per continent. The analysis shows significant variations across continents. Considering their peering stretch (Fig. 8b), we observe that certain regions have a large fraction of prefixes traveling significantly farther from the nearest location. While 90% of North American prefixes are served from the closest location, the corresponding fraction is much lower in other regions. For instance, 25% of prefixes in South America and Oceania travel more than 7,000 km and 13,000 km, respectively. Appendix A includes a CDF of peering stretch per continent (Fig. 8b), allowing a more detailed analysis of the long-tail distributions.



(a)

(b)

Fig. 8. On the top, Fig. 8a shows the eyeball networks peering with cloud services across continents, aggregating the Internet users they represent. On the bottom, Fig. 8b presents the 'peering stretch' for these networks, offering a continent-wise comparison.



	AF	AS	EU	NA	OC	SA
SA	--	0.00	0.01	0.04	0.00	0.32
OC	--	0.01	0.00	0.01	0.72	0.00
NA	0.33	0.48	0.37	0.68	0.49	0.30
EU	0.15	0.18	0.54	0.05	0.00	0.00
AS	0.02	0.50	0.03	0.02	0.00	0.03
AF	0.29	--	0.01	0.01	--	0.00

Fig. 9. Peering matrix that shows the fraction of continent eyeballs (columns) that peers with the cloud in a given continent (rows).

b) Where do peers come from and where do they peer?: Figure 9 illustrates this as a heatmap with the continents of peering points (rows) and traceroute destinations (columns), with each cell showing the fraction of the Internet population from networks that meet the row-column pairs.

The colors and numbers in the heatmap show a clear pattern. As expected, the majority of Internet populations peer with the cloud within their home continents, shown in the matrix diagonal. Nevertheless, the analysis reveals a substantial portion of these Internet populations – with the exception of North America – also engage in cloud peering at distant locations, either as a secondary peering point or, in some cases, as the primary one.

The most common destination for these remote peerings is North America, where a significant number of networks and their users – ranging from 0.29 to 0.47, go to peer with larger cloud providers. Major networks such as Bharti Airtel-9498 and China Telecom-4134, both significant providers in the APNIC region, provide additional examples of networks opting for remote peering as they travel to the US for peering with cloud providers (possibly not exclusively). Besides North America, Europe serves as a popular alternative destination for eyeball networks, particularly those in Africa (0.15) and Asia (0.18), likely due to geographical proximity and the availability of submarine cable (e.g., ACE, SeaMeWe-4) connecting the continents. We find negligible remote eyeball networks outside North America and Europe in other continents. An interesting case is Angola Cables-37468, a state-funded transit company [49] aimed at establishing low-latency connectivity between Africa and the Americas. The company invested in deploying a transoceanic cable in the South Atlantic [50] to connect to Brazil, the ultimate goal is optimizing routes from Angola to Miami [51], [52].

c) With whom do they peer?: We examine the prevalence of country-level Internet populations that peer with the cloud among cloud providers. Figure 10 shows the cumulative distribution of the fraction of countries' Internet populations that peer with each cloud service. The plot shows two clear clusters, with IBM separated from the second cluster, which includes GCP, Azure, and AWS. Their respective mean values of the size of the countries' Internet population hosted by these peers show a similar grouping with IBM (11.33%), significantly different than GCP (27.27%), Azure (32.19%) and AWS (32.28%). Looking at the specific countries and networks contributing to these populations shows networks in countries as diverse as Uruguay, Vietnam, Qatar, and New Zealand.

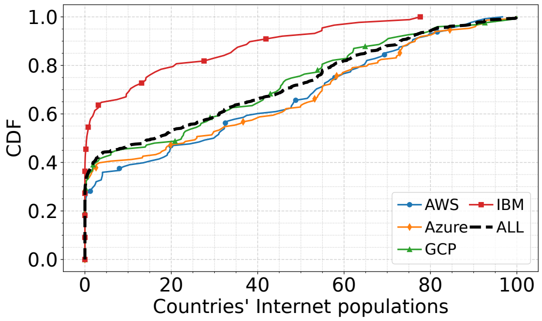


Fig. 10. Cumulative distribution of the country-level eyeballs peering with the cloud in a different continent.

As a side note, we found several organizations using different networks for peering and serving eyeballs and accounted for this in our analysis. For instance, our estimation of the Internet population of networks peering with the cloud refers to all eyeball networks within a cloud-peering organization. We include a detailed analysis of these organizations in Appendix A.

C. *Bypassing* transit-free-clique *transits*

In recent work, Arnold et al. [7] shows the extent to which networks bypass TIER-1 transit providers to peer with the cloud. Considering this work and the long tail distribution of peering stretch, we focus on the subset of peers that travel to a different continent to peer with the cloud and investigate the role of these peers as transits, their peering and destination points.

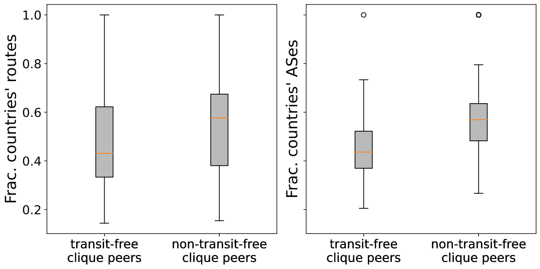


Fig. 11. Boxplot of the fraction of routes (left) and networks (right) in each country that is reachable via *tfc* and *non-tfc* networks peering with the cloud.

We download the entire CAIDA's PPDC dataset[6] to compile a list of all networks that have been at some point identified as members of the *transit-free clique* [53]. Using this membership information, we classify networks peering with cloud providers into two categories: *(i)* transit-free-clique (*tfc*) peers, and *(ii)* non-transit-free-clique (*non-tfc*) peers.

To compare *tfc* and *non-tfc* peers in routes to intercontinental destinations, Figure 11 shows a boxplot of the fraction of routes and networks to each country that use either type of network. Our measurements show that *non-tfc* peers are more prevalent in the path to intercontinental destinations in both routes and networks. This *"second tier"* is composed of diverse networks that established peering with the cloud at intercontinental locations to serve smaller networks in their regions. For instance, a notable example in Brazil involves large regional transit networks such as Vtal-7738 and Algar-16735, which peer with AWS in the Ashburn area.

VI. Why going further?

In this section, we discuss some of the possible reasons that could motivate operators to establish distant peering to the cloud. While this discussion is not intended to be exhaustive, it explores some of the explanations brought up in discussions with network operators and other researchers, including the role of content availability and clouds' pricing strategies, the availability of peering options as an attractor, the influence of physical infrastructure's availability, and decreasing cost, especially for those networks with no proprietary long-haul infrastructure, and whether presence at remote locations predates the growth of cloud computing.

A. *Availability of Long-Haul Infrastructure*

In the last section, our findings reveal that 20% of the prefixes are accessible through peering locations situated over

[6] CAIDA's PPDC files can be found at: https://publicdata.caida.org/datasets/as-relationships/serial-1/

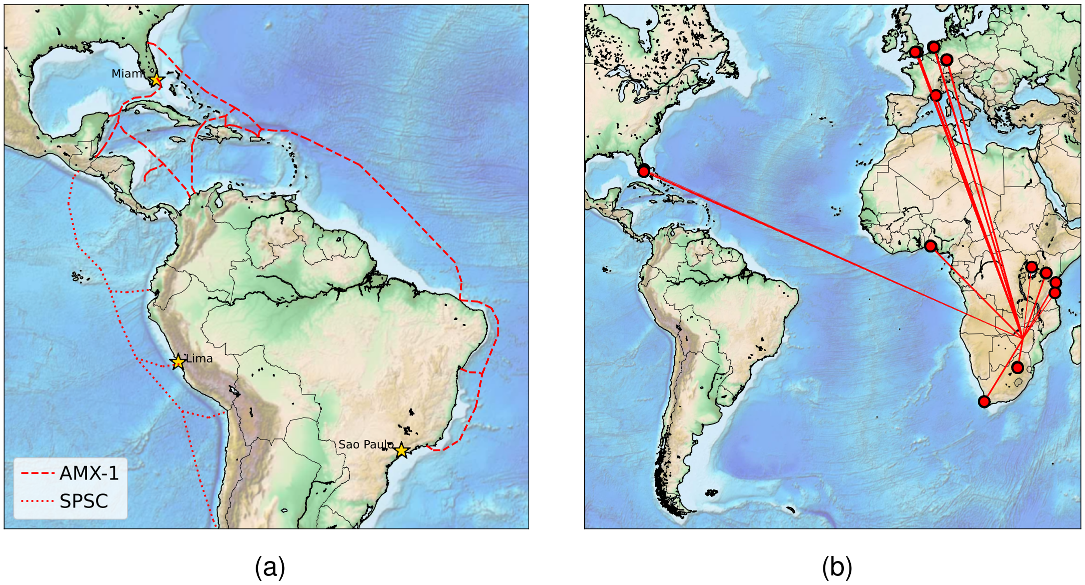


Fig. 12. Comparative analysis of the network expansion strategies through submarine and terrestrial cables. On the left, the network of the Latin American mobile carrier Claro extends across Central and South America, highlighting its strategic connectivity to the US, particularly Miami (Fig. 12a). On the right, Liquid Telecom's extensive overseas presence showcases an example of networks peering abroad to serve the domestic markets, such as Zimbabwe in this example (Fig. 12b).

6,700 km away from their nearest datacenter. With these findings in mind, we now focus on network structure characteristics, such as submarine connectivity, leasing contracts, and capacity costs, that could explain these preferences for peering at distant locations.

a) Availability of Submarine Infrastructure: We use Claro Peru-AS12252 as an example to illustrate a broader trend in Latin America, where there is a consistent inclination to establish a presence in the US. The submarine cable ring of Claro, depicted in Figure 12a[7], encircles South America and extends to Florida, showing an interest in connecting Claro's subsidiaries in South America with the US. Although IX.br Sao Paulo [54] currently serves as the preferred interconnection point for cloud providers and regional networks, historical preferences of Latin American providers for Miami persist. This trend extends beyond Claro Peru-AS12252, as recent submarine cable deployments, including Mistral (2021) and Monet (2017), show a persistent interest in establishing connections between the US and Latin America. Despite the growing local hosting presence, PeeringDB reveals the continued presence of numerous large and medium-sized Latin American networks in various facilities across Southern Florida and the US.

b) Indefeasible Right of Use: A common practice in the network industry involves subscribing long-term agreements known as *Indefeasible Right of Use* (IRU) contracts in which companies secure access to submarine routes for 10 to 25 years [55], [56], [57], [58], [59], [60], [61], [62]. For instance, Telecom Argentina S.A. (AS7303) disclosed in a 2013 filing to the US Securities and Exchange Commission (SEC) that it had acquired IRUs for over 15 years from Latin America Nautilus (LAN), a subsidiary of Telecom Italia to establish connectivity with Miami [20]. These extended agreements enable networks to reach overseas destinations to obtain access to content, pricing structures, and peer connections unavailable in their respective regions. Liquid Telecom-AS30844, a prominent provider in South Africa and Zimbabwe, is an example of purchase capacities (probably over IRU contracts) to establish a presence in peering facilities across multiple continents, as shown in Figure 12b. The period of these contracts suggests that many could have been established before the expansion of cloud services that could explain the *inertia* observed in peering locations, resulting in high peering stretch values.

c) Declining costs of International capacity: Despite the historical decline in capacity costs, we search for public records that could explain the persistent interest of networks in establishing remote presence. While typically confidential, Table II aggregates reported price reductions from Telegeography's public articles [63], [64], [65], [66], highlighting significant drops, especially in the Miami-Sao Paulo route. Despite suggestions in some reports, leasing submarine capacity may not be the predominant factor in pricing structures (eg., backhaul and rack costs) [67], [55], networks may establish a presence in global hubs to secure competitive transit agreements. Although transit costs have seen consistent overall reductions [67], transit providers in the US and Europe offer more economical transit agreements compared to other regions [16], [15].

B. Content availability and pricing strategies

Two other reasons that could help explain networks peering at remote locations are the availability of particular content and pricing considerations. To obtain a view of cloud regions used to serve content, we examine top-ranked websites in countries

[7] Cable layout sourced from Telegeography's GitHub repository, retrieved during its public availability.

TABLE II
VARIATION IN 10/100 GBPS WAVELENGTH MEDIAN PRICES. EMPTY CELLS ARE FOR UNAVAILABLE DATA.

route	’14-’17	’17-20’	19’-22’
Miami-Sao Paulo	-79.46%	- 47.45%	-56.10%
London-Singapore			- 16.94%
Los Angeles-Tokyo	-64.21%	-27.10%	
Hong Kong-Singapore		-62.68%	
London-New York	-46.86%	-34.15%	

across different continents. We also consider pricing of cloud services in different regions.

The web is a complex ecosystem involving multiple resources and services that third-party providers often serve. In the current service model, websites commonly rely on third-party providers for – at least – content delivery, DNS management, and domains’ certificates. While prior efforts have explored third-party dependency in The Web [68], [69], our analysis focuses on its potential relation with peering decisions.

We run a non-exhaustive analysis to obtain a qualitative understanding of using remote cloud regions to serve cloud resources. To collect data for this analysis, we focus on Alexa’s list of TOP500 regional websites for different countries and use a VPN service to obtain all resources of these websites including CNAMES, A and PTR DNS records [69]. We limit our analysis to those servers that embed their geolocation information in PTR DNS records. A complete geolocation of the serving infrastructure of cloud-based services is a complex task beyond this work’s scope. We collect this data from vantage points in six different countries, one per continent: Australia (Oceania), Argentina (South America), United States (North America), Germany (Europe), India (Asia) and South Africa (Africa).

Table III shows the percentage of websites that use cloud-based services and the fraction that has at least one resource hosted in another continent. While we see limited adoption of IBM and Azure in the surveyed countries, AWS and GCP are, on average 44% and 47%, respectively, of their most popular websites. These results reveal a significant adoption of content served from remote locations (likely non-cacheable content), particularly in South Africa, Argentina and Australia. For instance, Google resources are being served from overseas – primarily Singapore – for 61% of Australia’s top-ranked websites. Websites in the United States and Germany, on the other hand, are primarily relying on local cloud resources. While the line of causation is unclear, this dependency on overseas-hosted content may help explain cloud peering at distant locations.

To explore pricing consideration as another explanation for remote peerings, we investigate the per-region price strategies implemented by cloud providers. Figure 13 shows the percentage difference of VM prices in different regions when compared with the base price (US-based resources). The price spreads are remarkable, reaching 61% (Sao Paulo) and 115% (Rio de Janeiro) for AWS and Azure, respectively. The price gaps across computing facilities may incentivize service providers to host resources overseas, especially when these resources are latency-insensitive, encouraging some networks to establish remote peerings to enable direct access to this content.

C. Peering opportunities

TABLE III
PERCENTAGE (WITH TOTAL NUMBERS IN PARENTHESES) OF TOP500 WEBSITES IN COUNTRY-LEVEL RANKINGS THAT HAVE RESOURCES HOSTED AT CLOUD PROVIDERS (ALL) AND INTERCONTINENTAL CLOUD FACILITIES (ICT).

	AWS		GCP	
Country	ALL	ICT	ALL	ICT
Argentina	46 (231)	17 (84)	46 (350)	7 (33)
Australia	55 (273)	15 (76)	55 (317)	61 (304)
Germany	40 (201)	4 (19)	40 (261)	1 (4)
India	38 (192)	10 (48)	38 (27)	3 (13)
United States	63 (316)	1 (6)	63 (289)	0 (2)
South Africa	23 (115)	23 (115)	23 (180)	36 (180)

We examine the preferences of networks when establishing a presence abroad, aiming to understand how these choices influence the composition of IXPs.

Internet Exchange Points (IXPs) have proliferated across all regions [4], [3], [70], [71], [51], [72], [73], serving as local peering hubs that facilitate settlement-free traffic exchange among members, bypassing transit providers and reducing latency. This model has thrived globally, fostering dense peering ecosystems in Europe and Latin America, with North America following to a lesser extent [4], [3], [70], [71]. Since IXPs are designed to *keep local traffic local*, it is important to explore possible motivations behind *networks traveling significant distances to participate in an IXP*.

We analyze potential regional correlations between the number of IXP members and the diversity of their nationalities, using the latter as a proxy for whether larger memberships promote distant peering relationships. To explore these dynamics, we used a snapshot of PeeringDB (PDB) from April 2022. Our analysis draws on the reported presence at IXPs to assess membership and utilizes the nationality of the organizations managing these networks, as recorded in PeeringDB, to evaluate the diversity of their origins. Notably, despite hosting the largest number of members, IX.br Sao Paulo shows less diversity than its European counterparts, DE-CIX, LINX, and AMS-IX, which have evolved into global hubs attracting networks from various regions. We further examine PeeringDB records over six years (2018-2023) to investigate a potential growth

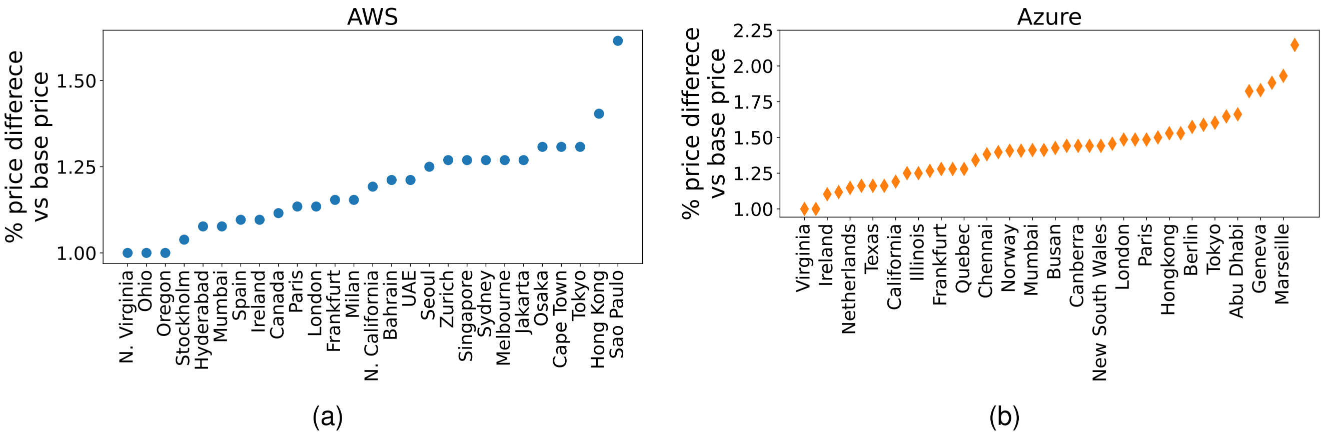


Fig. 13. AWS's (Fig. 13a) and Azure's (Fig. 13b) per-region price differentiation strategies. Regional prices are compared as a percentual increment concerning a base price.

in the international presence and find a consistent fraction (0.17) of networks present at a different continent from where they were registered. The analysis suggests networks may be present abroad when IXPs offer a diverse network ecosystem. Interestingly, this preference has remained consistent over the past six years.

VII. Related Work

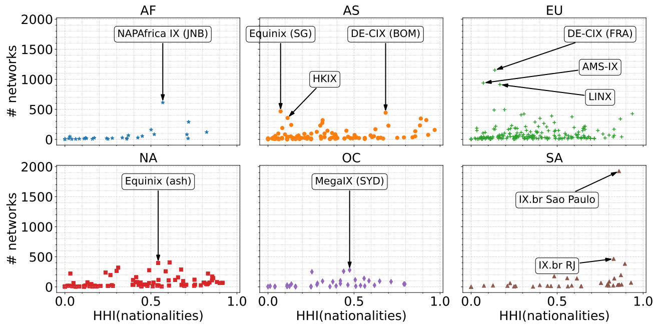


Fig. 14. Herfindahl-Hirschman Index (HHI) (x-axis) measures the diversity of the nationality of IXP members and the number of members (y-axis) for IXPs registered in PDB. Marker shapes distinguish the continents where these IXPs are located.

The rise of cloud providers in recent years has drawn the attention of the research community. Many of these efforts reviewed changes in the topological characteristics of the network brought about by the expansion of cloud networks. Arnold et al. [7] build on a cloud-centric measurement campaign to show that cloud providers deliver content through a dense non-hierarchical interdomain network. Applying a similar methodology to ours, Marder et al. [25] identified and geolocated clouds' border routers but focused on the interdomain connectivity between cloud providers. Yeganeh et al. [26] also relied on traceroutes to the entire /24 address space to reveal Amazon's peering fabric and classified these links into public, private and *virtual private interconnections*. Salamatian et al. [74] addressed the rising concern of opacity in cloud-based measurements by proposing Ricci's curvature to identify clouds' private links. In contrast, our work extends beyond network topology by looking into the locations where networks peer with the cloud. We compare these peering locations with the respective locations of network prefixes and evaluate the proximity of available cloud alternatives.

Other studies have gone beyond the examination of topological characteristics. Dang et al. [75] explored latency characteristics to major cloud providers and identified distance and last-mile bottlenecks as significant factors. Mok et al. [76] ran speed test clients from Google Cloud Platform to detect network congestion during COVID-19 lockdowns in the US. Kashaf et al. [68] and Kumar et al. [69] quantified the reliance of most visited websites on cloud providers and discussed the risks of Internet centralization. Similarly, Moura et al. [77] examined queries to root and ccTLD DNS servers to quantify the level of Internet centralization of traffic to large cloud and content providers.

Similar to our findings, the existing literature on IXPs has also examined the role of these interconnection points in offering content and peering opportunities. The growth of IXPs transformed the network by attracting both participants and content providers to the same place, creating a *virtuous* cycle of growth. Ager et al. [3] documented this transformation, showing that by 2011, the traffic exchanged at a prominent European IXP had exceeded the volume carried by one of the largest Internet Service Providers. Considering the close relationship between IXPs and content providers, Bottger et al. [78] proposed a ranking mechanism for large content providers based on reported capacity and presence at public exchanges. Castro et al. [17] further explored the phenomenon of *network flattening* finding that networks often travel great distances to join IXPs.

VIII. Conclusions and Future Work

The emergence of large cloud providers in the last decade has transformed the Internet, resulting in a seemingly evergrowing set of datacenters, points of presence, and consequently network peers. Yet, despite the availability of closer peering locations, some networks continue to peer with cloud providers at distant locations, traveling thousands of kilometers, beyond the nearest computing facility. This paper presented the first examination of the distances network travel to peer with the cloud. We analyzed different characteristics of those networks opting for distant peering locations, revealing that networks serving large Internet populations still travel to North America and, to a lesser extent, to Europe to peer with the cloud. We also discuss potential explanations for these extended distances and argue about multiple factors such as preexisting infrastructures and contractual agreements such as Indefeasible Rights of Use (IRU). These findings contribute to a deeper understanding of peering decisions that highlight the inertia of some characteristics of the network topology and the persistent use of submarine cables to peer at distant locations.

This work suggests promising directions for future work extending the analysis of distant peering. Our analysis focuses on four large cloud providers – Amazon, Google, IBM and Microsoft. It may be interesting to expand this analysis to other cloud providers, particularly Alibaba and Tencent. Another research direction could examine how distant peering translates into end-to-end latency for real web services, assess whether failover events risk SLA violations, and extend our measurements longitudinally as new data centers and submarine cables come online to observe how the stretch evolves over time.

IX. Acknowledgements

This work is supported by the National Science Foundation grants CNS-2107392, CNS-1703592, CNS-2039146, and CNS-2106517. Any opinions, findings, and 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. The authors would like to thank the reviewers for their valuable feedback and for their comments on early versions of this manuscript.

References

[1] A. Dhamdhere and C. Dovrolis, "The Internet is flat: modeling the transition from a transit hierarchy to a peering mesh," in *Proc. of CoNEXT*, 2010.

[2] C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian, "Internet inter-domain traffic," in *Proc. of ACM SIGCOMM*, 2010.

[3] B. Ager, N. Chatzis, A. Feldmann, N. Sarrar, S. Uhlig, and W. Willinger, "Anatomy of a large european ixp," in *Proc. of ACM SIGCOMM*, 2012.

[4] N. Chatzis, G. Smaragdakis, A. Feldmann, and W. Willinger, "There is more to ixps than meets the eye," *ACM SIGCOMM Computer Communication Review*, 2013.

[5] V. Giotsas, S. Zhou, M. Luckie, and k. claffy, "Inferring multilateral peering," in *Proc. of CoNEXT*, 2013.

[6] A. Lodhi, N. Larson, A. Dhamdhere, C. Dovrolis, and K. Claffy, "Using peeringdb to understand the peering ecosystem," *ACM SIGCOMM Computer Communication Review*, 2014.

[7] T. Arnold, J. He, W. Jiang, M. Calder, I. Cunha, V. Giotsas, and E. Katz-Bassett, "Cloud provider connectivity in the flat internet," in *Proc. of IMC*, 2020.

[8] A. Weissberger, "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/, 2022.

[9] Amazon Web Services, "AWS Global Infrastructure," 2024, https://aws.amazon.com/about-aws/global-infrastructure/.

[10] Microsoft Azure, "Azure geographies," https://azure.microsoft.com/en-us/explore/global-infrastructure/geographies/.

[11] Google Cloud, "Cloud locations," https://cloud.google.com/about/locations.

[12] IBM Cloud, "Regions and locations," https://cloud.ibm.com/docs/hs-crypto?topic=hs-crypto-regions.

[13] G. Huston, "How Big is that Network?" https://labs.apnic.net/?p=526, 2014.

[14] L. Corneo, M. Eder, N. Mohan, A. Zavodovski, S. Bayhan, W. Wong, P. Gunningberg, J. Kangasharju, and J. Ott, "Surrounded by the clouds: A comprehensive cloud reachability study," in *Proc. of the WWW*, 2021.

[15] R. Stanojevic, I. Castro, and S. Gorinsky, "Cipt: Using tuangou to reduce ip transit costs," in *Proc. of CoNEXT*, 2011.

[16] I. Castro, R. Stanojevic, and S. Gorinsky, "Using tuangou to reduce ip transit costs," *IEEE/ACM Transactions on Networking*, 2014.

[17] I. Castro, J. C. Cardona, S. Gorinsky, and P. Francois, "Remote peering: More peering without internet flattening," in *Proc. of CoNEXT*, 2014.

[18] T. de Colombia, "Informe anual," https://descubre.movistar.co/informe-de-gestion-responsable-2018/estados_financieros/Telefonica_Informe2018_.pdf, 2018.

[19] ——, "Informe anual," https://www.telefonica.co/wp-content/uploads/sites/4/2023/03/Informe-Gestion-Responsable-2022-Reporte-Gestion-BIC.pdf, 2022.

[20] T. A. S.A., "United states securities and exchange commission: Form 20-f," https://www.sec.gov/Archives/edgar/data/932470/000119312514142506/d707745d20f.htm, 2013.

[21] D. Levin, Y. Lee, L. Valenta, Z. Li, V. Lai, C. Lumezanu, N. Spring, and B. Bhattacharjee, "Alibi routing," in *Proc. of ACM SIGCOMM*, 2015.

[22] J. A. Obar and A. Clement, "Internet surveillance and boomerang routing: A call for canadian network sovereignty," in *TEM 2013: Proceedings of the Technology & Emerging Media Track-Annual Conference of the Canadian Communication Association (Victoria*, 2012.

[23] M. E. Hathaway, "Connected choices: how the internet is challenging sovereign decisions," *American Foreign Policy Interests*, vol. 36, no. 5, pp. 300–313, 2014.

[24] "RouteViews," http://www.routeviews.org/routeviews/, 2020.

[25] A. Marder, K. C. Claffy, and A. C. Snoeren, "Inferring cloud interconnections: Validation, geolocation, and routing behavior," in *Proc. of PAM*, 2021.

[26] B. Yeganeh, R. Durairajan, R. Rejaie, and W. Willinger, "How cloud traffic goes hiding: A study of amazon's peering fabric," in *Proc. of IMC*, 2019.

[27] X. Fan and J. Heidemann, "Selecting representative ip addresses for internet topology studies," in *Proc. of IMC*, 2010.

[28] "Archipelago measurement infrastructure updates," https://catalog.caida.org/details/media/2011_archipelago, 2023, accessed: 2021-9-30.

[29] Google, "Prerequisites to peer with Google," https://peering.google.com/\#/options/peering, 2022.

[30] "Amazon Web Services," https://aws.amazon.com/peering/policy/, 2020.

[31] IBM, "Ibm cloud overview: Public peering," https://cloud.ibm.com/docs/overview?topic=overview-public-peering, 2022.

[32] M. Luckie, "Scamper: a scalable and extensible packet prober for active measurement of the Internet," in *Proc. of IMC*, 2010.

[33] A. Marder, M. Luckie, A. Dhamdhere, B. Huffaker, K. Claffy, and J. M. Smith, "Pushing the boundaries with bdrmapit: Mapping router ownership at internet scale," in *Proc. of IMC*, 2018.

[34] A. Arturi, E. Carisimo, and F. E. Bustamante, "as2org+: Enriching as-to-organization mappings with peeringdb," in *Proc. of PAM*, 2023.

[35] P. Richter, F. Wohlfart, N. Vallina-Rodriguez, M. Allman, R. Bush, A. Feldmann, C. Kreibich, N. Weaver, and V. Paxson, "A multi-perspective analysis of carrier-grade nat deployment," in *Proc. of IMC*, 2016.

[36] A. Dainotti, K. Benson, A. King, B. Huffaker, E. Glatz, X. Dimitropoulos, P. Richter, A. Finamore, and A. C. Snoeren, "Lost in space: improving inference of ipv4 address space utilization," *IEEE Journal on Selected Areas in Communications*, 2016.

[37] A. Marder, M. Luckie, B. Huffaker, and K. Claffy, "Vrfinder: Finding outbound addresses in traceroute," *Proc. ACM Meas. Anal. Comput. Syst.*, 2020.

[38] J. Weil, V. Kuarsingh, C. Donley, C. Liljenstolpe, and M. Azinger, "IANA-Reserved IPv4 Prefix for Shared Address Space," IETF, RFC 6598, Apr. 2012. [Online]. Available: http://tools.ietf.org/rfc/rfc6598.txt

[39] I. Poese, S. Uhlig, M. A. Kaafar, B. Donnet, and B. Gueye, "Ip geolocation databases: Unreliable?" *ACM SIGCOMM Computer Communication Review*, 2011.

[40] I. Livadariu, T. Dreibholz, A. S. Al-Selwi, H. Bryhni, O. Lysne, S. Bjørnstad, and A. Elmokashfi, "On the accuracy of country-level ip geolocation," in *Proceedings of the Applied Networking Research Workshop*, 2020, pp. 67–73.

[41] M. Luckie, B. Huffaker, A. Marder, Z. Bischof, M. Fletcher, and k. claffy, “Learning to extract geographic information from internet router hostnames,” in *Proc. of CoNEXT*, 2021.

[42] “Maxmind geolocation data,” https://www.maxmind.com/en/geoip2-services-and-databases, 2023.

[43] ipInfo, “ipinfo,” 2023. [Online]. Available: https://ipinfo.io

[44] A. Edmundson, R. Ensafi, N. Feamster, and J. Rexford, “Nation-state hegemony in internet routing,” in *Proc. of the ACM SIGCAS Conference on Computing and Sustainable Societies*, 2018.

[45] RIPE NCC, “IPmap,” https://ipmap.ripe.net/.

[46] B. Huffaker, M. Fomenkov, and K. Claffy, “Geocompare: a comparison of public and commercial geolocation databases,” *Proc. NMMC*, pp. 1–12, 2011.

[47] Microsoft, “Prerequisites to set up peering with Microsoft,” https://docs.microsoft.com/en-us/azure/internet-peering/prerequisites, 2022.

[48] Facebook, “Peering: Technical requirements,” https://www.facebook.com/peering, 2022.

[49] E. Carisimo, A. Gamero-Garrido, A. C. Snoeren, and A. Dainotti, “Identifying ases of state-owned internet operators,” in *Proc. of IMC*, 2021.

[50] W. Bank, “COUNTRY PRIVATE SECTOR DIAGNOSTIC: CREATING MARKETS IN ANGOLA,” https://documents1.worldbank.org/curated/en/606291556800753914/pdf/Angola-Country-Private-Sector-Diagnostic-Creating-Markets-in-Angola-Opportunities-for-Development-Through-the-Private-Sector.pdf, 2019.

[51] R. Fanou, P. Francois, and E. Aben, “On the diversity of interdomain routing in africa,” in *Proc. of PAM*, 2015.

[52] B. Prior, “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, 2018.

[53] M. Luckie, B. Huffaker, A. Dhamdhere, V. Giotsas, and K. Claffy, “As relationships, customer cones, and validation,” in *Proc. of IMC*, 2013, pp. 243–256.

[54] IX.br, “IX.br - Participants at the Sao Paulo IXP,” https://ix.br/particip/sp, 2023.

[55] “International Cables, Gateways, Backhaul and International Exchange Points,” https://www.oecd-ilibrary.org/science-and-technology/international-cables-gateways-backhaul-and-international-exchange-points_5jz8m9jf3wkl-en, 2014.

[56] R. Durairajan, P. Barford, J. Sommers, and W. Willinger, “InterTubes: A Study of the US Long-haul Fiber-optic Infrastructure,” in *Proc. of ACM SIGCOMM*, 2015.

[57] ——, “Greyfiber: A system for providing flexible access to wide-area connectivity,” *CoRR*, 2018.

[58] CommsUpdates, “Cable compendium: a guide to the week’s submarine and terrestrial developments,” https://www.commsupdate.com/articles/2014/01/24/cable-compendium-a-guide-to-the-weeks-submarine-and-terrestrial-developments/, 2014.

[59] ——, “Cable compendium: a guide to the week’s submarine and terrestrial developments,” https://www.commsupdate.com/articles/2019/01/25/cable-compendium-a-guide-to-the-weeks-submarine-and-terrestrial-developments/, 2019.

[60] ——, “Cable compendium: a guide to the week’s submarine and terrestrial developments,” https://www.commsupdate.com/articles/2021/10/01/cable-compendium-a-guide-to-the-weeks-submarine-and-terrestrial-developments/, 2021.

[61] ——, “Cable compendium: a guide to the week’s submarine and terrestrial developments,” https://www.commsupdate.com/articles/2018/08/03/cable-compendium-a-guide-to-the-weeks-submarine-and-terrestrial-developments/, 2018.

[62] ——, “Cable compendium: a guide to the week’s submarine and terrestrial developments,” https://www.commsupdate.com/articles/2016/01/15/cable-compendium-a-guide-to-the-weeks-submarine-and-terrestrial-developments/, 2016.

[63] B. Boudreau, “Global Pricing in Uncharted Territory,” https://blog.telegeography.com/global-pricing-in-uncharted-territory, 2023.

[64] ——, “Trends Stay on Track: A Massive Global Bandwidth Pricing Update for 2021,” https://blog.telegeography.com/2021-global-bandwidth-pricing-update-trends, 2021.

[65] ——, “Bandwidth Prices: How Low Can They Go?” https://blog.telegeography.com/bandwidth-prices-price-erosion-median-monthly-lease-prices, 2018.

[66] ——, “Telxius is Selling Its Cables. Here’s What Potential Buyers Are Assessing.” https://blog.telegeography.com/telxius-is-selling-its-cables-what-potential-buyers-are-assessing, 2020.

[67] A. Rebatta, “A tale of prices Latin America vs. The rest of the world,” https://forum.ix.br/files/apresentacao/arquivo/408/10\%20-\%20Anahi\%20Rebatta-IX\%20Forum-Dec\%2010.pdf, 2020.

[68] A. Kashaf, V. Sekar, and Y. Agarwal, “Analyzing third party service dependencies in modern web services: Have we learned from the mirai-dyn incident?” in *Proc. of IMC*, 2020, pp. 634–647.

[69] R. Kumar, S. Asif, E. Lee, and F. E. Bustamante, “Each at its own pace: Third-party dependency and centralization around the world,” in *Proc. of ACM SIGMETRICS*, 2023.

[70] E. Carisimo, J. M. D. Fiore, D. Dujovne, C. Pelsser, and J. I. Alvarez-Hamelin, “A first look at the latin american ixps,” *ACM SIGCOMM Computer Communication Review*, 2020.

[71] S. H. B. Brito, M. A. S. Santos, R. d. R. Fontes, D. A. L. Perez, and C. E. Rothenberg, “Dissecting the largest national ecosystem of public internet exchange points in brazil,” in *Proc. of PAM*, 2016.

[72] A. Gupta, M. Calder, N. Feamster, M. Chetty, E. Calandro, and E. Katz-Bassett, “Peering at the internet’s frontier: A first look at isp interconnectivity in africa,” in *Proc. of PAM*, 2014.

[73] J. C. Cardona Restrepo and R. Stanojevic, “A history of an internet exchange point,” *ACM SIGCOMM Computer Communication Review*, 2012.

[74] L. Salamatian, S. Anderson, J. Matthews, P. Barford, W. Willinger, and M. Crovella, “Curvature-based analysis of network connectivity in private backbone infrastructures,” *Proc. of ACM SIGMETRICS*, 2022.

[75] T. K. Dang, N. Mohan, L. Corneo, A. Zavodovski, J. Ott, and J. Kangasharju, “Cloudy with a chance of short rtts: Analyzing cloud connectivity in the internet,” in *Proc. of IMC*, 2021.

[76] R. K. P. Mok, H. Zou, R. Yang, T. Koch, E. Katz-Bassett, and K. C. Claffy, “Measuring the network performance of google cloud platform,” in *Proc. of IMC*, 2021.

[77] G. C. M. Moura, S. Castro, W. Hardaker, M. Wullink, and C. Hesselman, “Clouding up the internet: How centralized is dns traffic becoming?” in *Proc. of IMC*, 2020.

[78] T. Böttger, F. Cuadrado, and S. Uhlig, “Looking for hypergiants in peeringdb,” *ACM SIGCOMM Computer Communication Review*, 2018.

APPENDIX

This work does not raise any ethical issues.

Table IV summarizes the locations, number and type of VMs we instantiated in each of the cloud providers. We opted for the entry-level VM available in each regions assuming that traceroute measurements are neither computing nor storage intensive tasks. We discovered that Azure’s entry-level VMs were unable to collect any ICMP response and upgraded those instances to the immediate upper tier obtaining successful traceroute measurements.

CP	VMs	Locations	VM Types
AWS	2	Cape Town, Hong Kong,	t3.micro
	14	Tokyo, Seoul, Osaka, Mumbai, Singapore, Sydney, Frankfurt, Paris, London, Bahrein, Sao Paulo, N. Virginia, Ohio, N. California	t2.micro
GCP	14	Iowa, Toronto, Sao Paulo, Santiago, London, Frankfurt, Mumbai, Singapore, Hong Kong, Tokyo, Osaka, Seoul, Sydney, N. Virginia	e2-micro
Azure	8	Blue Ridge, San Francisco, Sydney, Hong Kong, Tokyo, Seoul, Osaka, London, Frankfurt, Toronto, Dubai, Sao Paulo	Std. B2s
	4	London, Frankfurt, Sao Paulo, Osaka	Std. D2as v4
	1	Des Moines	Std. DS1
IBM	13	San Jose, Washington, Toronto, Sao Paulo, Frankfurt, London, Paris, Hong Kong, Osaka, Seoul, Singapore, Tokyo, Sydney	Bal. B1.2x4

TABLE IV
LOCATIONS AND INSTANCE SPECIFICATIONS OF THE CLOUD RESOURCES USED IN THE MEASUREMENT CAMPAIGNS.

Figure 15 shows the peering stretch for networks in each continent peering with the cloud. These curves show that North America has the most compact distribution with 90% of the

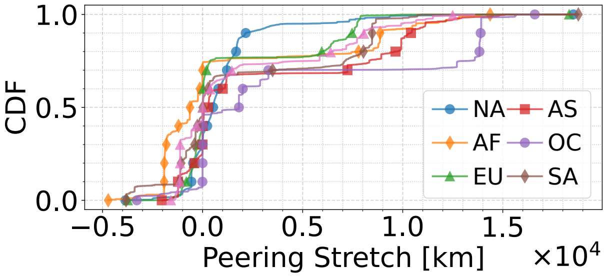


Fig. 15. ”peering stretch” to provide context to the distances traveled to peer with the cloud.

connections having detours of less than 1000 km. However, Africa, South America and Oceania are notable examples of heavy tails since networks from these continents often peer with the cloud in North America.

We now turn our attention to the organizational structure of these entities that facilitate access to the cloud for their respective countries.

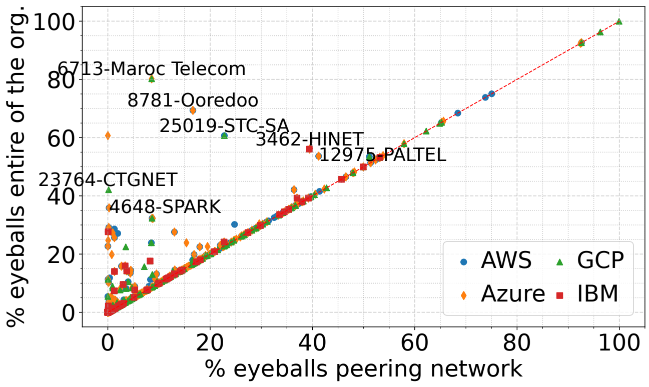


Fig. 16. In many cases, the organization that peers with the cloud uses different ASes to peer with the cloud and to provide access to the eyeballs.

To identify these organizations, we run $as2org+$ [34] to obtain all sibling networks associated with those networks that engage in cloud peering. In Figure 16, we examine the difference between the percentage of eyeballs in the network peering with the cloud and the percentage of eyeballs within the entire organization. While the majority of these peers (85.5%) show no difference with their respective organizations, we do observe several notable cases with substantial disparities. These differences indicate that certain networks within the organization engage in cloud peering and provide access to a large number of eyeballs within their organizational structure. State-owned providers of the Middle East [49], such as Ooredoo-8781 in Qatar and STC-39386 in Saudi Arabia, are prominent examples of separating both Internet population and peering networks as they do it when they peer with the cloud in Europe.

We validate the consistency of our findings over time, by replicating our experiments on September 29, 2023. We repeated the process of downloading announced prefixes, partitioning them at /24 granularity, and slicing them into non-overlapping segments across all VMs of each cloud provider. After executing traceroute measurements from the same locations with identical configurations, we discovered that Azure has shifted towards complete opacity with no hops visible in the traceroute data. Despite testing various configurations, we were unable to collect data from any hop along the path, suggesting that Azure is currently filtering TTL-limited packets.

Despite the absence of Azure data, our analyses with both snapshots revealed similar general trends. For example, Figure 17 shows the peering stretch distribution for traceroutes launched from AWS in both years, a cloud provider with consistent visibility in both snapshots. The comparable structure observed in the distribution indicates that the overall configuration of peering networks has remained consistent between the two campaigns. This finding indicates an overlap among the peers of the cloud providers. Repeating the analysis in §V-B and comparing both snapshots show temporal stability in the results.

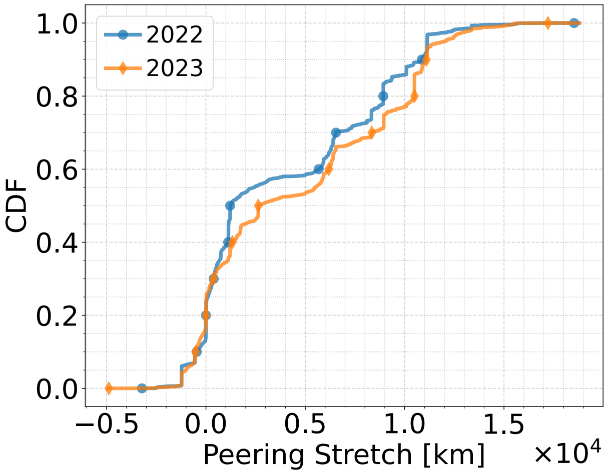


Fig. 17. Comparison peering stretch distributions for traceroutes launched from AWS in 2022 and 2023.