Peering Only? Analyzing the Reachability Benefits of Joining Large IXPs Today

Lars Prehn[1(✉)], Franziska Lichtblau[1], Christoph Dietzel[1,2], and Anja Feldmann[1]

[1] MPII, Saarbrücken, Germany
{lprehn,franziska.lichtblau,anja}@mpi-inf.mpg.de
[2] DE-CIX, Cologne, Germany
christoph.dietzel@de-cix.net

Abstract. Internet Exchange Points (IXPs) became a fundamental building block of inter-domain routing throughout the last decade. Today, they offer their members access to hundreds—if not thousands—of possible networks to peer.

In this paper, we pose the question: How far can peering at those large IXPs get us in terms of reachable prefixes and services? To approach this question, we first analyze and compare Route Server snapshots obtained from eight of the world's largest IXPs. Afterwards, we perform an in-depth analysis of bi-lateral and private peering at a single IXP based on its peering LAN traffic and queries to carefully selected, nearby looking glasses. To assess the relevance of the prefixes available via each peering type, we utilize two orthogonal metrics: the number of domains served from the prefix and the traffic volume that a large eyeball network egress towards it.

Our results show that multi-lateral peering can cover ˜20% and ˜40% of the routed IPv4 and IPv6 address space, respectively. We observe that many of those routes lead to out-of-continent locations reachable only via three or more AS hops. Yet, most IXP members only utilize "local" (i.e., single hop) routes. We further infer that IXP members can reach more than **half of all routed IPv4** and more than one-third of all routed IPv6 address space via bi-lateral peering. These routes contain almost all of the top 10K egress prefixes of our eyeball network, and hence they would satisfy the reachability requirements of most end users. Still, they miss up to 20% of the top 10K prefixes that serve the most domains. We observe that these missing prefixes often belong to large transit and Tier 1 providers.

1 Introduction

Traditionally, the Internet follows a hierarchical structure. At the top of this hierarchy resides a set of large transit providers—also called Tier 1 networks—that exchange traffic with each other at no monetary compensation. The literature commonly refers to this type of interconnection (and business relation) between two ASes as "peering".

O. Hohlfeld et al. (Eds.): PAM 2022, LNCS 13210, pp. 338–366, 2022.
https://doi.org/10.1007/978-3-030-98785-5_15

When logically descending from the top, higher-tier networks deliver traffic for their lower-tier customers, i.e., they provide transit. Since the early 2000s, the "topology flattening" phenomenon gradually superseded this hierarchical structure. Lower-tier networks started to shift more of their transit traffic to newly established peering connections. The continuous acquisition of new peering partners is often incentivised by cost reduction and potential latency improvements [2].

The fast and widespread deployment of Internet eXchange Points (IXPs) has further accelerated the establishment of new peering connections.

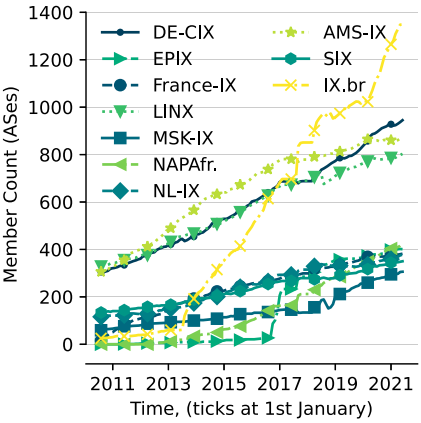

Fig. 1. Number of members over time based on PeeringDB

Traditionally, IXPs allow physically-close networks to exchange traffic via a shared layer-2 switching fabric; thus, they eliminate unnecessary routing detours, which reduces the overall latency and helps to "keep local traffic local". Today, the largest IXPs have grown to multiple hundreds—sometimes even thousands—of members (see Fig. 1) and handle peak traffic volumes of more than 10 Tb/s [4,27,38].

As different networks have different negotiation positions, various forms of peering have emerged. The simplest form, bi-lateral peering, refers to a direct connection between two ASes via the IXP's switching fabric.

To ease the life of their customers, most IXPs also offer Route Servers that redistribute all routes they received from one IXP member to all others via a single BGP session per member. As this form of peering involves more than two networks, the community refers to it as multi-lateral peering. As a third option, networks can establish private peering sessions amongst each other. Instead of using the IXP's layer-2 fabric, ASes establish these peering sessions via a dedicated cross-connect in the same colocation facility (or via layer-2 transport for different colocation facilities).

While peering itself is a well-established concept that has been broadly discussed in the research literature (e.g., [1,6,10,11,20,22,45,50]), we still lack fundamental insights into the actual extent and importance of the routes available at large IXPs. In this paper, we take a closer look at how the different forms of peering translate into transit-free prefix reachability. We characterize and compare the multi-lateral peering routes available at the Route Servers of the world's largest IXPs and further estimate the bi-lateral and private peering routes available at one large IXP in Europe that we refer to as L-IXP. We contrast our reachability analysis using two dimensions of importance: the number of top domains that a route serves and the traffic volume that one of the largest European eyeball networks egresses towards it.

In particular, our contributions can be summarized as follows:

- **Characterization of Multilateral Peering:** We analyze and compare Route Server snapshots from eight of the ten largest IXP peering LANs worldwide (see, Sect. 4). We find that all Route Servers show consistent insights: (1) only 10% of Route Server peers provide more than 100 routes while 30% provide less than ten routes, (2) approximately half of the Route Server routes have a minimum path length of three ASes (announced by close and distance peers alike) and about two-thirds of all routes lead to out-of-continent destinations, and (3) most large Route Servers have a prefix overlap of ˜50% while the actually reachable IPs overlap by ˜60–70%.
- **Characterization of Bi-lateral & Private Peering:** For one of Europe's largest IXPs, we infer routes available via bi-lateral and private peering (see Sect. 5.1). Similar to Ager et al. [1], we observe that most ASes use the switching fabric to establish additional transit sessions. As such connections can drastically influence our inferences of available routes, we developed a methodology to increase the coverage of relationship inference algorithms at IXPs, and we use the resulting relationships to isolate transit connections during the inference process. Similarly, we introduce a methodology to infer routes available via private peering based on the careful selection and querying of looking glass utilities.
- **Route Importance:** We compare the IPv4 and IPv6 routes available via multi-lateral, bi-lateral, and private peering against two top-10K prefix lists: one based on the number of served domains and one based on the traffic volume of a large European eyeball network (see Sect. 6). We find that nearly all top-10k IPv4 prefixes are available via bi-lateral peering. For IPv6, we observe that prefixes serving many domains are often unavailable (up to 15%) or can only be obtained via private peering.

2 Background

In this section, we provide an introduction to the different interconnection models and highlight important observations from related work. We refer to Fig. 2 as a visualization of the individual components explained throughout this section.

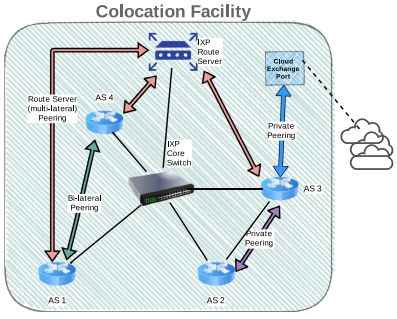

Fig. 2. Illustration of different peering types at an IXP. (Color figure online)

While interconnection agreements can be rather complex in practice, the scientific literature abstracts mainly into two categories: transit and peering.

In a transit agreement, a customer pays a transit provider for delivering its traffic from its egress router to any IP. In a (settlement-free) peering agreement, two ASes—usually of similar size and with roughly equal traffic volume towards each other—forward each other's traffic without substantial

amounts of money flowing in either direction. As neither of the peering partners is a provider for the other, both ASes have to negotiate where to physically interconnect and who is bearing the infrastructure costs. Over time and with the spread of Internet Exchange Points (IXPs) across the globe the peering ecosystem itself became rather complex and different peering practices emerged. In the following, we give an overview of the fundamentals of current peering models.

Internet Exchange Points. As establishing a single BGP peering session for every interconnection partner separately is rather wasteful, operators started building common switching infrastructure that could be shared (w.r.t. usage and cost) among ASes. These switching infrastructures—envisioned to keep local traffic local— belong to so-called Internet eXchange Points (IXPs) located in well-connected colocation facilities. Those colocation facilities provide dedicated infrastructure (e.g., rack space, electricity, and cooling) for the housing of peering equipment. Figure 2 gives an abstract example for a layer-2 peering fabric. While IXPs may attract very diverse sets of members, previous work reported that they observe traffic for 40% or more of all theoretically possible peering connections [13]. As some large IXPs observe traffic originated by or destinated towards tens of thousands of ASes and millions of servers [22] and could theoretically reach 70% of all routed addresses [10], it nowadays is also common that networks pay remote-peering providers to get access to remote IXPs [20]. A recent study by Nomikos et al. [57] revealed that around 90% of 30 tested IXPs had more than 10% of their members connecting via remote peering. They further reported that for certain large IXPs up to 40% of members can be connected via remote-peering.

Bi-lateral Peering. This practice describes a BGP peering session between two member ASes at an IXP via the shared peering fabric as depicted in Fig. 2 (green arrows). While legal processes and concerns of peering policy leakage slow down the acquisition of bi-lateral peering partners [49], Marcos et al. proposed a framework that allows IXP members to quickly provision peering sessions based on an intent abstraction and digitally handled legal contracts [50]. Interestingly, Ager et al. showed in 2012 that also Tier1 providers peer at IXPs and that they use their IXP peerings not only as backup routes. They further showed that these Tier1 providers also abuse the peering LAN for transit connections to their customers [1].

Multi-lateral Peering. As briefly discussed in Sect. 1, IXPs provide a Route Server for their members to establish multi-lateral peerings. In addition to reducing the number of needed interconnections to reach most IXP members[1], Route Servers can also implement additional functionality (e.g., the frequently used per-peer blackholing [28]) to make them more attractive to IXP members. Those services are often realized by attaching a specifically formatted BGP Community onto Route Server announcements. As a route server has to store such information to act properly based on it, some IXP members do not establish a session with the route server as they expect that it might expose their peering

[1] A Route Server reduces the number of totally needed BGP sessions for a fully-meshed topology from $n * (n-1)/2$ to n, where n is the number of BGP speakers.

policies [23]. As a notable example of such exposition, Giotsas et al. showed that it is possible to uncover 200k multi-lateral peering agreements by analyzing the BGP community values visible at few Route Servers [34].

Private Peering. When present at the same colocation facility, e.g. because they are members of the same IXP, two networks can establish a private peering session via direct cross-connect avoiding the IXP's peering fabric. Especially large ASes prefer this peering practice as it is provides a very fine-grained control over their peering sessions. Hence, networks that, e.g., need to egress a high traffic volume often require direct peering sessions on dedicated physical infrastructure with guaranteed capacity. This form of interconnection usually comes with monetary compensation for certain Service-Level Agreements (SLAs). Even though private peering keeps the peering policies of an AS hidden and often provides dedicated capacity, even private peering sessions can suffer from outages when, e.g., the entire colocation facility goes down—a not so uncommon scenario as Giotsas et al. reported (160 outages in 5 years) [32].

Cloud and Content Provider Connectivity. Many Cloud and Content providers peer at hundreds of physically distinct locations [11] to thousands of different networks [6]. While they often require private peering connections, they sometimes also rely on bi-lateral peering to ensure that they directly connect with as many eyeball ASes as possible [24] or to gain tens of milliseconds of latency improvements over their transit providers [69]. Hence, it is unsurprising that those providers also dominate the peering LAN traffic (as shown for two medium-sized IXPs by Cardona et al. [19]). Yet, as most networks try to establish private peering connections with them directly in the colocation facilities, those facilities have established so-called cloud exchanges—specific ports which directly provide connectivity (called virtual private interconnection (VPI)) to any number of cloud service providers within the colocation facility [79].

Identifying Peering Partners. Many network operators rely on a network policy database called PeeringDB to identify potential peering partners [62]. In particular, PeeringDB differentiates between four peering policy types: (1) open: A network with an open peering policy that peers with any other network, (2) selective: A network that will peer under certain conditions, e.g., minimum traffic volume or location, (3) restrictive: A network that already has an existing set of peers and needs strong, convincing arguments to establish a peering connection, and (4) no peering: These networks do not peer at all and rely entirely upon transit [58]. Notably, the vast majority of peering policies in PeeringDB are of the 'open' type. Yet, PeeringDB is known to have certain inaccurate entries [45,74]. Further, many small networks—especially in developing regions—do simply not register in PeeringDB [45].

3 Preface: Data Sets

While we introduce each data set separately when using it, this section summarizes the used data sets to provide a better overview of time coherence and caveats.

3.1 Main Data Sets

PeeringDB Snapshots (2010/08/01–2021/06/01, Monthly). PeeringDB is a community-effort database containing information about the infrastructure and policies for IXPs, colocation facilities, peering LANs, and networks [62]. PeeringDB is known to have a small set of inaccurate entries [45,74]. Similarly, Lodhi et al. reported that PeeringDB underrepresents small—especially developing country—networks [45]. The Center for Applied Internet Data Analysis (CAIDA) produces monthly snapshots of this database [17].

Route Server Snapshots (2021/06/06–21, Once). WWe compiled a set of Route Server snapshots for the largest (in terms of members) peering LAN for eight of the world's largest IXPs. We received these snapshots via multiple personal contacts throughout 15 days.

IXP Traffic Data (2021/05/01–2021/06/07). We obtain IPFIX traffic captures from one of the largest European IXPs. The traffic is sampled at a rate of 1 out of 10K (1:10k) flows. The captures encompass all traffic exchanged via the peering LAN; hence, it contains traffic exchanged via multi-lateral and bi-lateral peering sessions but misses private peering traffic. In particular, we utilize the data from May 2021 to analyze how our observation period influences our results and subsequently report most of our results based on the first week in June 2021.

ISP Traffic Data (2021/06/10). We obtain a single workday of egress traffic captured from all border routers from a large European eyeball network. The data was sampled at a rate of 1:1K packets.

Domain-Based Prefix Top List (2021/04/30). We obtain a recently recomputed domain-based prefix top list from Naab et al. [55]. Their methodology relies on a domain top list as input, then resolves those domains to IP addresses from a single physical location, and finally aggregated the number of Fully Qualified Domain Names that is served by every norm-prefix (i.e., a /24 prefix in IPv4 and a /48 prefix in IPv6). We use the prefix top list that relied on Umbrella's domain top list [25] as input, as it was the only one that could provide us with 10K IPv6 prefixes. Notably, this domain-based prefix top list is biased towards the European service region as DNS load-balancing [71] and caching [67] may lead to strongly regionalized address resolutions.

Please note that we handled our traffic data sets in compliance with **measurement ethics** and best practices. We performed all data analyses on servers located at the respective premises of our vantage points using data collected as a part of their routine network analysis. We analyzed flow data summaries based on packet headers that did not reveal any payload information. We further anonymized all flow attributes not explicitly needed for the results presented in this paper. This is in line with Ethical Committee policies. For the remaining data sets, we rely on publicly available sources only.

3.2 Orthogonal Data Sets

Maxmind GeoLite2 Snapshot (2021/06/01). We utilize a snapshot of Maxmind's GeoLite2 database [53] to geolocate Route Server prefixes. While they can have significant inaccuracies on a city or country-level [21], even freely available databases achieve near-perfect continent-level predictions [52].

CAIDA's AS Relationships Snapshot (2021/06/01). CAIDA produces monthly snapshots of the business relationships inferred by ASRank [47] based on routing information collected by RouteViews [61] and RIPE/RIS [56] from the first five days within the month [14]. While it misses many peering links, this data is reasonably complete for transit links [35,59,60]. Further, the inference algorithm is known to near-perfectly infer transit relationships but often misinfers peering relationships as transit [30,39,40], i.e., it overestimates the number of transit relationships.

CAIDA's IP-to-AS Mapping Snapshot (2021/06/10). CAIDA generates daily IP to AS mappings based on routing information from selected Route Views [61] collectors [18].

CAIDA's AS-to-Org Mapping Snapshot (2021/04/01). CAIDA produces quarterly snapshots of AS-to-Organization mappings generated based on the WHOIS databases of all Regional and some National Internet Registries [16]. Notably, WHOIS data is known to contain malformatted and hard-to-parse entries [44], leading to potential inaccuracies in the inferred AS-to-Organization mapping. The April snapshot is the latest available snapshot before our measurement period.

4 Multilateral Peering

We start our analysis with the lowest-hanging fruit: multi-lateral peering. While some IXPs have explicit APIs that could be used to re-build the current routing table of their route servers, we explicitly request Route Server snapshots for the largest peering LAN of different IXPs. Out of the ten IXPs shown in Fig. 1, only Nl-IX and EPIX did not fulfil our request. Our eight Route Server snapshots are from differents days between 6th and 21st June, 2021[2] and contain the entire routing information base for each session, i.e., they contain all paths from all neighbours (rather than just one best path) for a given prefix. Using those snapshots, we look at what routes an AS may expect from the Route Server and how consistent those findings are across different IXP Route Server. In particular, we arrive at the following takeaways:

– Large Route Servers across the world are very similar: They not only have the same distribution of routes per peer but also share the majority of reachable

[2] As we obtained similar results for all Route Server related plots for a set of inital snapshots that we obtained throughout January and February, we do not expect any major inconsistencies due to a two week offset.

prefixes and IPs, i.e., joining a second, third, etc. Route Server only negligibly improves reachability.
– Due to the growing trend of remote peering, Route Servers provide only a limited amount of in-continent routes.
– We observe that most routes (at all analyzed Route Servers) contain at least three hops. While both close and distant peers announce those lengthy, unattractive routes, we find that members often only use one-hop Route Server routes.

How Consistent are the Distribution of Routes to Peers Across Route Servers? Our snapshots show that connecting to the Route Server immediately provides routes from up to 650 IXP members. Yet, Richter et al. already reported that not all IXP members announce the same number of prefixes [68]. As a first look at how similar Route Servers are, we analyze whether this distribution is consistent across them. Figure 3 shows the number of prefixes (y-axis, logarithmic) announced by every peer (x-axis) per Route Server. Indeed, we observe strong consistency across different IXP Route Servers regardless of the protocol. For the AMS-IX Route Server (top curve), the top ˜1.5, 10, 30, and 70% of Route Server peers announce routes for more than 10K, 1K, 100, and 10 IPv4 (1K, 100, 20, and 5 IPv6) prefixes. While most Route Servers are close to AMS-IX, peers at NAPAfrica (bottom curve) announce around an order of magnitude fewer prefixes.e fewer prefixes, most other IXPs are closer to AMS-IX.

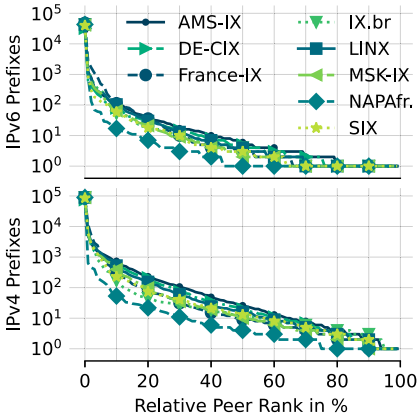

Fig. 3. Number of prefixes announced per peer

Notably, not all prefixes are necessarily exported to all peers by the Route Server. To estimate how many prefixes can only be received conditionally, we inspect the Route Server snapshots for BGP communities that control its redistribution rules. For, e.g., DE-CIX, we inspect routes with the `0:6695` Community that is used to exclude all peers; this community is usually combined with other BGP Communities of the form `6695:X` which instruct the Route Server to explicitly redistribute a route to peer `X`. Overall, we find that 31.3% of IPv4 and 11.2% of IPv6 Route Server prefixes are not globally exported.

Do Route Servers Help to Keep Local Traffic Local? As briefly discussed in Sect. 2, IXPs initially were established as a solution to interconnect geographically close ASes following the idea to "keep local traffic local". Yet, given that many peers announce tens of thousands of prefixes to hundreds of millions of hosts, we now want to take a look at how strictly this idea is followed through by today's Route Servers. We first use a naïve approach to answering this question:

We look at the AS path length (after removing AS Path Prepending). Figure 4 shows the Route Server prefixes of different IXPs separated by the number of ASes in their shortest route. We observe that for around half of all prefixes the shortest path contains three or more ASes. This result goes against the "keep local traffic local" idea, as local routes would likely either directly lead to an access/eyeball network or indirectly via a national service provider. However, given that the AS path length is often not a good proxy for geographic distance, we now switch to a more insightful perspective.

Fig. 4. Length of shortest AS path per prefix

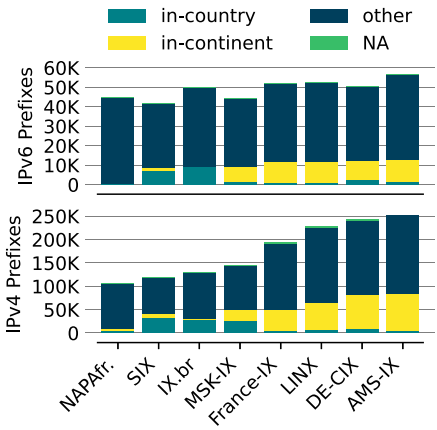

Fig. 5. Geolocation of prefixes relative to Route Server

Rather than looking at the AS path, we now directly map the visible prefixes to countries and continents using a snapshot of Maxmind's GeoLite2 database [53] from 1st June 2021. While perfect IP-to-geolocation mapping is a long-standing research problem, previous work showed that for various public geolocation databases 99% of predictions stay within 600 km of the actual location [21]. Similarly, Maxmind claims that for many countries 0% of predictions are off by more than 250 km [52]. While this large radius might influence the accuracy of country-level predictions, it provides us with near-perfect accuracy for continental predictions as most of our Route Servers have even more distance between their location and the closest continental border. Figure 5 shows the Route Server prefixes of different IXPs separated by whether they lead to in-country, in-continent, or out-of-continent ("other") hosts. Notably, there is a small number of prefixes for which the database did not include a mapping ("NA"). Interestingly, looking at host locations provides an even more drastic result than looking at AS paths: Regardless of the actual Route Server, around two-thirds of all prefixes lead to out-of-continent hosts.

While the growing trend of remote-peering [57] can easily lead to many out-of-continent routes, it is unclear whether it also contributes to the high number of

lengthy routes. To better understand whether this correlation exists, we want to compare the path length of each route with the RTT (as a proxy for distance) to its next-hop interface. Hence, we run ping measurements from a server directly connected to the switching fabric of L-IXP towards each member interface.[3] To account for latency inflations due to, e.g., congestion, we repeated those measurements 100 times and collected the minimum RTT towards each interface throughout all runs. Finally, we associate the shortest path of each prefix with the minimum RTT we measured for its respective next-hop interface. Notably, if there was more than one possible shortest path, we picked the one for which the next-hop RTT was the lowest. Figure 7 shows for each prefix of a given minimum path length the minimum latency to its next-hop.[4] We observe that there is no strict correlation between the distance of a peer and the length of the routes it provides (Fig. 6).

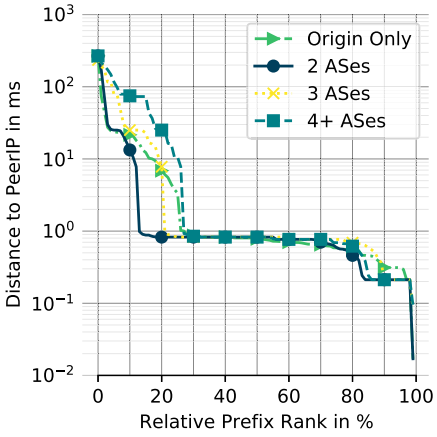

Fig. 6. Distance to next-hop per prefix, separated by length of shortest AS path

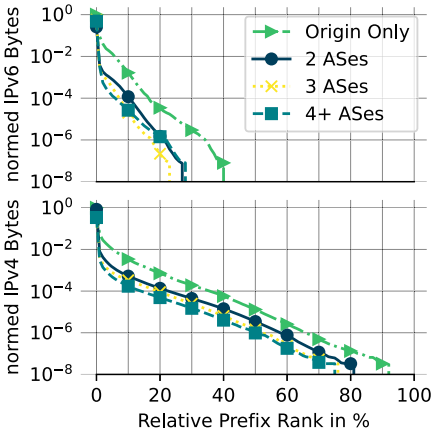

Fig. 7. peering LAN bytes per prefix, separated by length of shortest AS path

Now that we know that even local peers forward lengthy routes to the route server, the question becomes whether those routes see any traffic. For one of our observed IXPs, we obtained IPFIX captures sampling 1 out of every 10K packets traversing its peering LAN. While we can observe multilateral and bilateral peering traffic in this data set, we have no insights into traffic exchanged via private peering established via direct interconnects as it does not traverse the public peering infrastructure. Based on the captured flows between the 1st of June and the 7th of June[5], we calculate the aggregated number of Bytes destined

[3] We neither had probing devices at other peering LANs, nor was our probing device at L-IXP IPv6-enabled at the time of our study.

[4] We explicitly avoid the classification into remote and local peers based on RTT estimates alone given the caveats presented in [57].

[5] We provide details on how we choose this time window in the next section.

towards each prefix. Figure 7 groups Route Server prefix by their shortest path and shows for each prefix (x-axis) the number of bytes (y-axis, logarithmic) relative to the prefix with the most bytes (i.e., we show bytes normalized by the prefix with the maximum byte count, ρ). We observe that 6% of prefixes reachable via one hop carry at least 1% of ρ's bytes while only less than 0.5% of 2 or more hop prefixes carry that much traffic. Apart from the top 6%, prefixes reachable via two or more hops carry around an order of magnitude less traffic—with only minor differences between two, three, and four or more hops. Finally, we observe that 8, 19, 24, and 25% of IPv4 (60, 72, 73, and 77% of IPv6) prefixes with a shortest path of 1, 2, 3, and 4+ hops carry no traffic at all, respectively.

Those observations are likely tied to how long-established IXP members engage with a Route Server: In contrast to new members, long-established members already acquired many bi-lateral peering sessions. It is common that members attribute higher local preference values to such bi-lateral sessions as they often come with Service Level Agreements (SLAs). Hence, long-established members often peer with the Route Server to get an idea of which routes are available at all but only hand-pick routes they actually use based on, e.g., how consistently they are available or how much performance benefit they may introduce. As local preference values only de-prioritize (rather than filtering them) multi-lateral peering routes, Route Servers are also used as automatic fall-back in case a bi-lateral peering session suffers from, e.g., an outage [32,68].

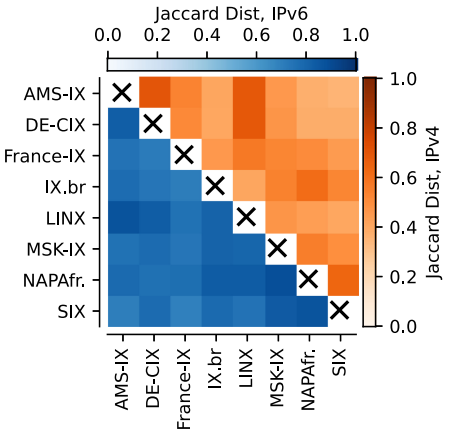

Fig. 8. Similarity of prefixes between Route Servers

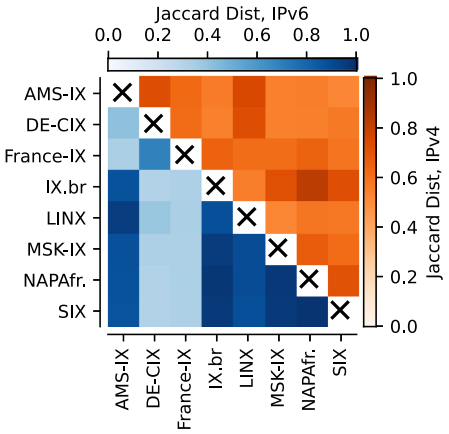

Fig. 9. Similarity of addresses between Route Servers

How Route Server Specific are Multi-lateral Peering Routes? Until now, we saw that most Route Servers have very similar characteristics; hence, we now try to understand where the actual difference lies. As a similarity metric, we use the Jaccard distance. The Jaccard distance between two sets of elements,

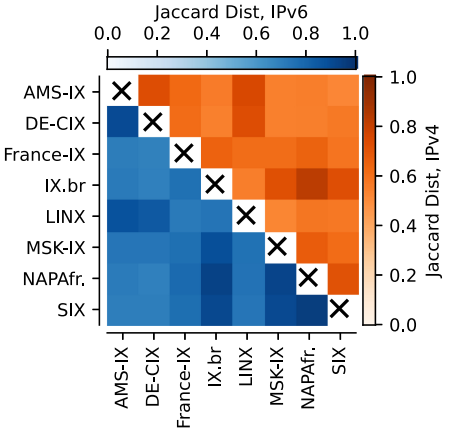

Fig. 10. Similarity of addresses between Route Servers without HE's 2002::/16 route

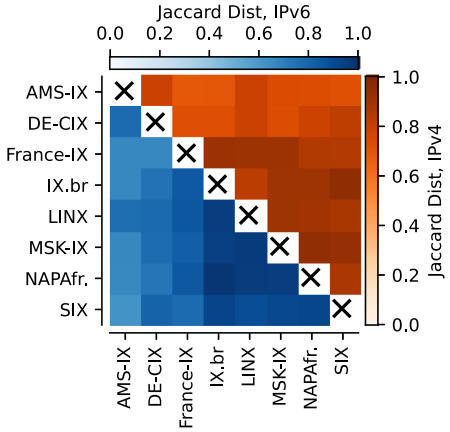

Fig. 11. Similarity of prefixes between Route Servers for common peers

A and B, is calculated as $JD(A,B) = \frac{|A \cap B|}{|A \cup B|}$. In comparison to other common similarity metrics (e.g., the overlap coefficient $OC(A,B) = \frac{|A \cap B|}{min(|A|,|B|)}$), the Jaccard distance also produces small values when A is entirely contained in a significantly larger B, i.e., it not only considers the similarity of elements but also the cardinalities of the sets. For each pair of Route Servers we now compute the Jaccard Distance between prefixes (see, Fig. 8) and reachable IP addresses (see, Fig. 9). As the Jaccard index is symmetric, we show results for IPv4 in the top-right triangle and results for IPv6 in the bottom-left triangle.

While we observe that certain Route Server combinations show more overlap than others (e.g. AMS-IX and DE-CIX), the average similarity for IPv4 lies at around 50% (77% for IPv6). As certain prefixes can be more-specifics of others, it is also unsurprising that the similarity of reachable IP addresses lies roughly 13% higher for IPv4. While we observe similar behaviour for many IPv6 combinations, we observe that France-IX and DE-CIX are different from the others but similar to each other. We observe that this "clustering" is mainly the result of a single route: 2002::/16 announced by AS6939 (Hurrican Electric). When ignoring this route (see Fig. 10), the takeaways for IPv6 are roughly the same as for IPv4.

Finally, we want to know whether ASes with memberships at multiple IXPs share the same routes with the respective Route Servers. Hence, we rerun the same analysis but, this time, focus only on routes announced by the same member ASes at both IXPs (see Fig. 11). While this comparison shows naturally higher overlap compared to Fig. 8, we observe that certain Route Server combinations still show a Jaccard distance of less than 70%; yet those routes barely make a difference for the number of reachable IPs (Figure not shown).

Summary. We observe that the distribution of prefixes across Route Server peers that was presented by Richter et al. [68] is also present in many other

Route Servers across the world. In general, we show that the characteristics of routes at various Route Servers are very similar. We observe that the majority of routes at Route Servers lead to out-of-continent destinations—likely a side-effect of the growing remote-peering trend. Surprisingly, we found that most routes at Route Servers contain three or more ASes and that the distance of the peer is not a factor for this phenomenon, i.e., even local peers provide many unattractive routes to the Route Server. Nevertheless, the peering LAN traffic from one IXP suggests that its members primarily use the routes to direct destinations, and mostly rely on the Route Server for failover or analysis purposes.

5 Inferring Peering Relationships

After we analyzed the routes that are available to newly joined IXP members via multi-lateral peering, we are now interested in the routes that can be obtained by establishing bi-lateral and private peering sessions.

Similar to the work of Richter et al. [68], we infer bi-lateral peerings (and the prefixes that are announced via them) by observing the traffic that flows through the IXP's peering LAN. As shown by Ager et al., some ASes may "abuse" the peering LAN for additional transit connections to their customers. Given that our reachability analysis might be rather sensitive to the presence of transit relationships[6], we substantially extend the method used by Richter et al. to account for them.

As the inference approach for bi-lateral peerings relies on traffic data, we now limit the scope of our analysis to **one** large European IXP, L-IXP. While the IXP's peering LAN may cover most of the bi-lateral peering agreements, it offers no visibility into the private peerings that happen within the co-located data centers; hence, we rely on carefully selected looking glasses within those data centers to uncover routes that are available via private peering. Notably, this approach does not allow us to accurately distinguish between dedicated private peerings and connections to, e.g., cloud exchanges (as discussed in Sect. 2).

5.1 Bilateral Peering

We bootstrap our analysis in a similar way to Richter et al. [68]: Whenever we observe traffic destined towards IP I flowing from A to B, we deduct that the respective covering /24 (or /48 for IPv6) for I must have been announced from B to A. Notably, this approach relies on the assumption that an ASes will *eventually* send traffic to most, if not all, of the prefixes it received from a neighbor. Hence, we first have to understand for how long we need to observe peering LAN traffic before we arrive at a rather static "snapshot".

Picking a Reasonable Window Size. On the one hand, a small window size (e.g., an hour) may underestimate the available routes as not all of them

[6] As customers can potentially send traffic destined for the entire Internet to their transit providers, incorporating such connections would bloat up the set of reachableprefixes.

necessarily continuously see traffic; on the other hand, a large time window (e.g., a year) is more likely to yield an extensive list, yet may provide an overestimate as certain routes are withdrawn in the meantime. To get a better sense of what might be a good window size, we test by how much a certain window size would affect the number of /24s and /48s for which we observe traffic. For various window sizes between 4 hours and 14 days, we calculate the prefix counts and then move the window forward by one hour. Using this method, we generate, e.g., 739, 719, 575, and 407 data points for the window sizes 4 hours, 1 day, 7 days, and 14 days throughout the entire May 2021.

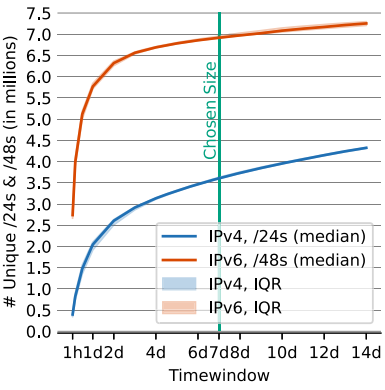

Fig. 12. Influence of window size on visible prefixes

Figure 12 show the median prefixes (y-axis) that we observed for a given window size (x-axis) as well as the Inter Quartile Ranges (IQR) for IPv4 and IPv6. While the knee of the curve (i.e., the point at which further increases of the window size start to yield smaller improvements) lies at around one and a half days, we observe a continuous, almost linear, increase after a window size of six days. We decided to choose a window size of seven days. While this choice might yield a small number of already withdrawn prefixes, it covers workdays as well as weekend days—which are known to exhibit rather different traffic characteristics [29,41,43,72].

Removing Transit Sessions. Now that we have some understanding of the routes that are announced between each member pair, we have to isolate and ignore transit sessions as they might substantially inflate the set of reachable prefixes. Perfectly identifying the business relationships of links has been an academic goal for more than two decades. The current state of the art algorithm, ASRank [47], is well-known for its high accuracy when it comes to identifying transit relationships (even in narrow contexts [64]). CAIDA hosts two versions of monthly-updated business relationship information: serial-1 and serial-2. While serial-1 relies solely on routing information (i.e., AS paths), serial-2 contains serial-1's information but is further extended with topology information inferred via additional sources, e.g., traceroute paths that were mapped to AS Paths. As a result, serial-2 contains more relationships but also inherits inaccuracies from its data extensions (e.g., from IP-to-AS mapping [7,51]). Surprisingly, neither serial-1 nor serial-2 can cover more than 21.2% or 22.3% of the 220k+ IPv4 IXP member pairs that exchanged traffic during that period.

Improving Relationship Coverage via Route Server Paths. Whether the ASRank algorithm produces an inference for a given AS link mostly depends on the set of AS paths that it is executed on. Hence, we can improve our inference coverage by providing additional AS paths that 'cross' (i.e., contain two consec-

utive IXP members) the IXP's peering fabric. To uncover such paths, we revisit the Route Server of our IXP.

Our main idea is as follows: Our Route Server snapshot contains various routes as well as their respective Route Server redistribution communities, i.e., Route Server specific communities to express the instructions: (1) announce to all neighbor, (2) don't announce to any neighbor, (3) announce to a specific neighbor, and (4) do not announce to a specific neighbor. Notably, instruction (1) and (2) are usually paired with instructions of type (3) and (4) but not with one another. By simulating the redistribution, we can deduce the paths that each IXP member received via its Route Server session(s).

More formally, we construct paths as follows: Let AS A announce some route with AS path (A, p') to the Route Server where p' refers to some (potentially empty) sequence of ASes—we ignore the few routes that contain AS_SETs. A also attaches a set of (potentially large) BGP communities that we translate into the previously explained instructions (1)–(4). To retrieve the set RP of Route Server peers to which the route is redistributed, we first sort the set of instructions in the order we introduced them[7]. While we set RP to all Route Server neighbors for instruction (1), we set RP to the empty set for instruction (2); if both instruction (1) and (2) are present we ignored the route. Notably, if n either instruction (1) nor (2) is present, we defaulted to instruction (1). Afterward, we first added and then discarded specific ASes to/from RP according to the instructions of type (3) and (4). respectively. Finally we constructed paths of the form $(B, A, p'), \forall B \in RP$ which 'cross' the IXP at the link (B, A).

We combine those paths with routes gathered from five days of the rib snapshots from the route collector projects RIPE RIS and RouteViews (i.e., the same data sources that CAIDA uses to produce serial-1 data). For IPv4-related inferences, we use the publicly available ASRank script that is hosted by CAIDA. For IPv6, we apply the necessary changes described by Giotsas et al. [33] to adjust the inference script to IPv6 routig policies. Both scripts require a list of Route Server ASNs for their inference. To generate this list, we extract all ASNs with the type 'Route Server' from PeeringDB. After these steps, our extended relationship data set covers 69.0% and 63.2% of traffic-carrying IPv4 and IPv6 links.

Improving Relationship Coverage via Manual Search. At this point, we still have various ASes with limited coverage. Hence, we decided to manually search for additional relationship information. We invested three days of manual relationship look-ups for ASes that either (i) are in the top 30 contributors of unclassified links, (ii) have only less than 10% of their links covered, or (iii) have more than 10% of their links inferred to be transit connections.

For our manual search, we mostly relied on entries in PeeringDB (e.g., [63]), RADb/Whois (e.g., [66]), and targeted web searches (e.g., [76]) that clearly described (at least some) relationships of a given ASN—please note that the

[7] This order represents a conservative approach—if both the instruction to add AS X and to delete X are present, x will ultimately not be included in the set of Route Server peers.

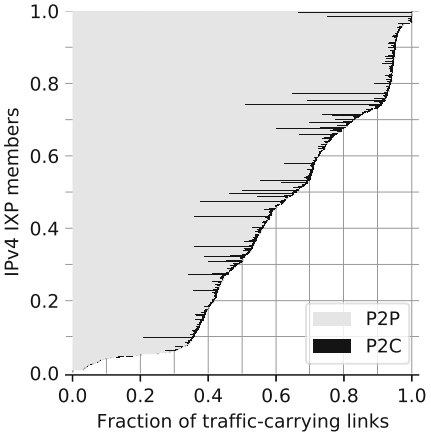

Fig. 13. Coverage of relationships for traffic-carrying links (IPv4).

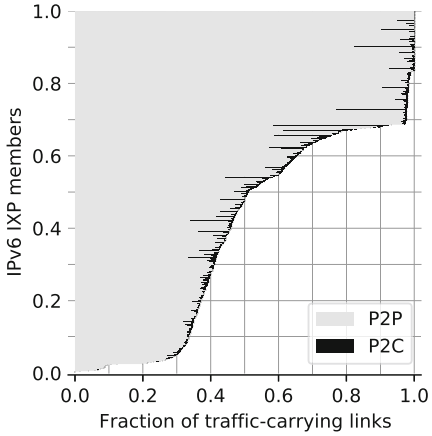

Fig. 14. Coverage of relationships for traffic-carrying links (IPv6).

three given examples are chosen randomly and may or may not belong to members of our studied IXP. For autnum objects in RADb/Whois, we used an approach similar to that described in [47] to infer transit relationships (even though we did not automate the process). We used as-set objects in RADb/Whois with clearly defined names (most commonly, e.g., `AS<XXX>:AS-CUSTOMER(S)`, `AS<XXX>:AS-TRANSIT(S)`, `AS<XXX>:AS-UPSTREAM(S)` or `AS<XXX>:AS-PEER(S)`) to identify relationships. For PeeringDB and the targeted web searches, we searched for exhaustive enumerations of, e.g., providers as part of, e.g., the network infrastructure description. Whenever possible, we differentiated between IPv4 and IPv6 relationships as well as regional relationships (i.e., if a websites described AS X as peer in Europe but as provider in Asia, we noted it as peer giove that our IXP operates in Europe).

While investigating the relationships for the ASes mentioned above, we observed diminishing coverage improvements; hence, we decided to not extend our manual search beyond them. Notably, whenever an AS explicitly specified its providers and customers but not its peers, we assumed that all remaining links are peering relationships.

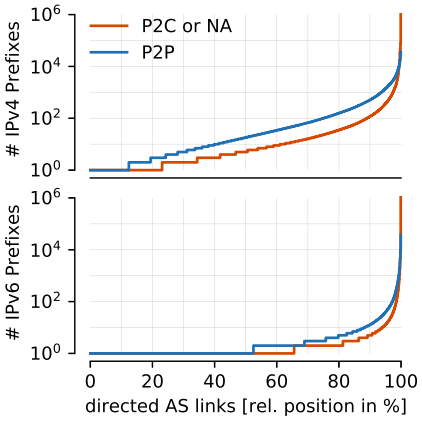

Fig. 15. Norm-prefixes per directed AS link

Our final set of relationships covers 74.2% and 65.9% of traffic-carrying IPv4 and IPv6 links at our IXP. Figure 13 (for IPv4) and Fig. 14 (for IPv6) show the fraction of links for each AS that are inferred to be P2P and P2C relationships. We observe that in both plots our data set covers at least a fourth of all relationships for 93% of ASes. On median, we cover 66% of IPv4 and 51% of IPv6 relationships. While we observe that overall only 1.2% (IPv4) and 1.5% (IPv6) of all inferred links have transit relationships, we also observe that these relationships are distributed across almost all IXP members; hence, it is rather the norm than the exception to establish additional sessions with transit providers via the IXP's peering fabric. Beyond its coverage, we are also interested in the filtering impact of our relationship data set. Figure 15 shows the number of available IPv4 and IPv6 norm-prefixes per traffic-carrying, directed[8] AS link. We observe that certain links carry traffic for more than 10^6 norm-prefixes. Yet, when only considering links that our data set classifies as peering links, we filter out all links that carry traffic for exceptionally many prefixes. Hence, we continue our analysis using only the links explicitly inferred as peering links, i.e., we not only ignore those links explicitly inferred as transit links but also those for which we have no inferred relationship.

5.2 Private Peering

As previously discussed in Sect. 4, our traffic captures do not contain any private peering connections. Therefore, we rely on queries to carefully selected looking glasses (LGs) to infer routes available via private peering. To automatically query looking glass interfaces, we write identification and querying interfaces—similar to those described in [31]—for common looking glass utilities including, e.g., HSDN [73], RESPAWNER [54], and COUGAR [26]. To initially find ASes with looking glasses, we rely on PeeringDB [62] as well as various online lists [8,9,37,42,46,75]. We first narrow down our selection by removing all LGs from ASes that are not members of our IXP. Afterwards, we removed all LGs that our indentification interface could not map to a LG template. Then, we manually went through the looking glass interfaces of the remaining 63 ASes and validated whether they could look at the routing table of a router that is located within one of the IXPs contiguous colocation facilities—we heavily relied on the naming and

[8] If A and B exchange traffic in both directions, we treat the links (A, B) and (B, A) separately.

excluded all entries for which the location was not exactly matching a colocation name. Finally, after removing LGs requiring captchas, exploring rate-limiting, or explicitly stating 'no automation allowed', we are left with LGs from 17 different ASes to trigger.

Triggering Looking Glasses. As looking glasses are usually provided on a voluntary basis from operators to operators, we do not want to abuse them with gazillions of bursty queries. First, we limit the set of norm-prefixes for which we query the LGs to those that are (1) necessary for the analysis in Sect. 6 and (2) not yet covered by multi-lateral or bi-lateral peering. Second, when a looking glass yields a longest-prefix match rather than an exact match and returns a covering prefix that is likely not a default route (i.e., a routes less specific than /8 and /16 for IPv4 and IPv6, respectively), we no longer query for any other norm-prefixes covered by this less-specific. Third, we waited 39.3 seconds[9] on average between two consecutive queries to the same looking glass. With those safeguards in place, we queried looking glasses as follows:

1. **Querying a LG.** We choose a looking glass in round-robin fashion and performed—depending on the LG utility—either an exact match or, preferably, a longest-prefix match query against it.
2. **Ignoring transit routes.** If the LG returned a route for which the first-hop would be a transit provider to the AS the looking glass resides in, we ignore that route. Similarly, if we can't find a relationship and the first hop is a Tier 1 provider, we also ignore the route (given that it likely represents a transit relationship).
3. **Requiring IXP routes.** To ensure that the route is locally available at the IXP, we ensured that the first-hop AS is also an IXP member.

If no route remains after steps 2 and 3, we wait 2 seconds and then query the next looking glass until we have exhausted our LG list. If one LG returned a non-filtered route we marked the norm-prefix as reachable (and queried the next round-robin-order LG for the next norm-prefix), otherwise we mark it as unreachable.

In total, we were able to uncover 2.33M, 6.73M, and 6.77M IPv4 (3.41B, 3.41B, and 3.45B IPv6) norm-prefixes available via multi-lateral, bi-lateral, and private peering covering 19.8, 57.1, and 57.4% (37.3, 37.4, 37.8%) of all routed IPv4 (IPv6) addresses (according to Geoff Houston's Routing Table Analysis Report [36]), respectively. These results provide a real-world calibration for the 70+ % of reachability theoretically calculated by Böttger et al. [10] in 2018.

6 Route Importance

In this section, we present a qualitative analysis of the uncovered peering prefixes with two different measures of importance: (a) How many domains in a top N

[9] A result of multiple small waits between queries to different LGs in combination with the answer time of the other LGs.

ranking are served by transit-free reachable prefixes, and (b) how many of the top destination prefixes of a large eyeball network are reachable without transit. The findings of this section can be summarized as follows:

- For both rankings, around half of the top-100 norm-prefixes can be reached via multi-lateral peering.
- For our traffic-based ranking, nearly all prefixes can be reached via bi-lateral peering with few exceptions that can mostly be reached via private peering.
- For our domain-based ranking, the same holds true for IPv4. For IPv6, we observe that bi-lateral peering has a substantially lower impact. While, in general, more prefixes remain unreachable than for IPv4, most of the top norm-prefixes can be obtained via private peering.
- We observe that the prefixes that remain unreachable even via private peering mostly lead to large Transit and Tier 1 providers.

6.1 Prefix Rankings

Traffic-Based Ranking. To provide a traffic-based importance ranking from an independent source, we use traffic statistics from one of the largest European ISPs. In particular, we collect egress traffic from all the ISP's eyeball source addresses at all edge routers over one day (10th June 2021) at a sample rate of 1:1000 packets. For each destination IP, we sum the number of egress bytes throughout the day, aggregate these values to norm-prefixes, and cluster the top 10k norm-prefixes for IPv4 and IPv6.

Domain-Based Ranking. To quantify the importance of IPs with another metric, we obtain a domain-based importance ranking. Thus, we rely on re-computed results from a previous work by Naab et al. [55]. The domain-based norm-prefix top list is generated by picking a common domain top list (e.g., from Alexa [3], Majestic [48], or Umbrella [25]), resolving these domains to as many IPs as possible, and then ranking each norm-prefix by the number of Fully Qualified Domain Names (FQDNs) that can be resolved to an IP. We requested an updated snapshot of the top list from the authors of [55] and promptly received a re-computation from 30th April 2021. We decide to use the Umbrella-based norm-prefix top list because it is the only one from which we can derive 10K IPv4 as well as 10k IPv6 prefixes.

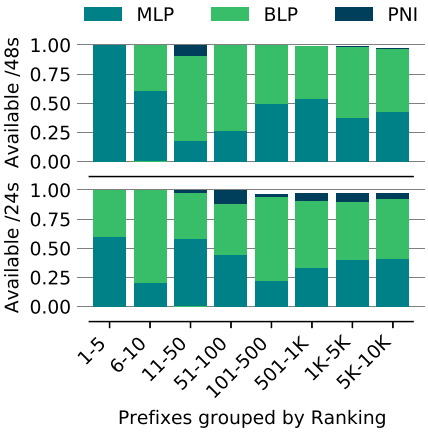

Fig. 16. Coverage of eyeball-based top-10K prefix ranking

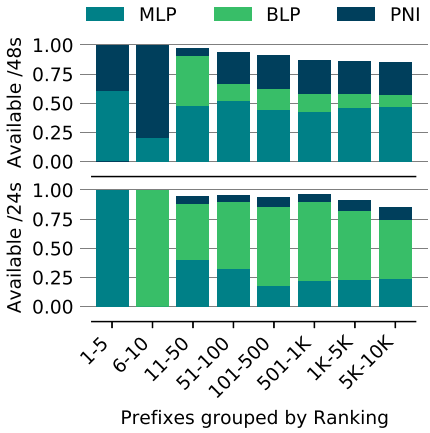

Fig. 17. Coverage of domain-based top-10K prefix ranking

6.2 Reachability of the Top-10K

Now that we got the domain- and traffic-based top 10 IPv4 and IPv6 norm-prefixes, we can analyze how many of those prefixes are reachable via different peering types.

Traffic-Based Ranking. Figure 16 separates the top 10k prefixes into different classes based on their respective ranking (x-axis) and shows for each class the fraction of reachable prefixes (y-axis) for IPv4 at the bottom and IPv6 at the top. In addition, each prefix is colored by the lowest-requirement peering type (requirement and economical costs for PNI > BLP > MLP) that it can be reached by (if any). We observe that the top 100 prefixes for both protocols can be fully covered using all peering types. In general, we observe that only very few prefixes can not be reached. Notably, the vast majority of top-10k prefixes can solely be reached via bi-lateral peering agreements. This result benefits aspiring IXP members who, if they carefully select a few private peering partners, can keep their operational costs minimal.

Domain-Based Ranking. Figure 17 shows our results for the domain-based top 10k prefixes in the same style as the previous figure. First, we observe that significantly more—especially lower rank—prefixes are unreachable (e.g., approx. 15% of the lowest 5k IPv4 prefixes are not reachable). Second, we see a drastic shift in patterns for IPv6: The difference between routes available via multi-lateral and bi-lateral peering is almost negligible compared to IPv4. Consequently, IXP members have to rely substantially more on private peering to reach the prefixes with the highest domain counts. Yet, for approx. 15% of 500-or-lower prefix class prefixes IXP members still have to rely on their transit as they are unreachable via peering.

To reduce their operational costs, members of large IXPs may egress most—if not all—of their high-volume destination traffic via (mostly bi-lateral) peering connections while using their transit to egress low-volume yet domain-heavy prefixes. Notably, between 25 and 50% of both top-10k prefix lists can be reached via multi-lateral peering—a finding that further highlights the importance of Route Server connections especially for new IXP members.

6.3 Missing Routes

To get some idea of which routes were not available, we mapped norm-prefixes to ASes via a longest-prefix match on the previously mentioned IP-to-AS data set from CAIDA. We further map each origin AS to a class using CAIDA's AS Classification data set [15].

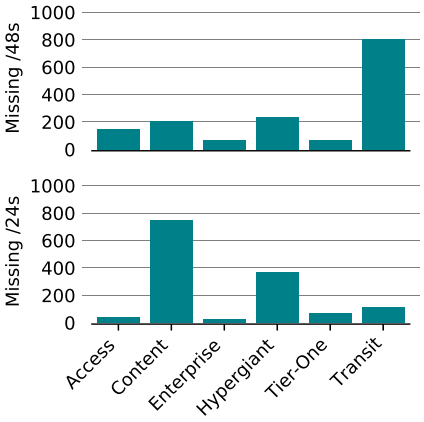

Fig. 18. Unavailable prefixes by origin AS type.

We further refine the classification using lists of Tier 1 Networks [77] and Hypergiants [12]. Figure 18 shows the number of missing norm-prefixes (y-axis) that are originated by the ASes of different classes (x-axis) for IPv4 (bottom) and IPv6 (top). For IPv4, we observe that most of the missing/24 prefixes belong to content providers/hypergiants. In particular, we observe that more than half of the prefixes in both of those classes can be attributed to Amazon's AS14618 and AS16509. Notably, most of the missing prefixes for Amazon do not see any peering LAN traffic (regardless of the business relationship) throughout our measurement period. As most of these prefixes are unique to the traffic-based prefix ranking, we suspect that our eyeball vantage point has access to routes that are only announced via private peering on dedicated connections, and, hence, remain hidden from the peering LAN. Taking Amazon out of the picture, the most prominent class would be the same as for IPv6: Transit ASes. Notably, the individual contributions made by single ASes are much more uniformly distributed; out of the 61 and 231 total ASes contributing to the IPv4 and IPv6 Transit AS class, the top ASes contribute no more than 21 and 29 prefixes respectively. Further, we observe that the vast majority of the prefixes that belong to Transit ASes are only present in the domain-based top list but not in the traffic-based top list. In summary, our observations suggest that ASes can indeed offload high-volume prefixes to peering links by joining an IXP but they still require transit to reach the heavy tail of (potentially low-traffic) domains.

6.4 Limitations

Next, we discuss limitations and specifically elaborate on the generalization of our findings. **Multi-lateral Peering:** We analyzed the Route Servers of different IXPs based on separate snapshots generated throughout seven days. Hence, our observations may be biased by sequences of high-frequency updates (as described by Ariemma et al. [5]). Yet, we discussed our results with some of the IXP operators that provided Route Server snapshots, and they told us that they did not observe unusual behavior during the days from which the snapshot was taken. Yet, as many prefixes can only be seen when aggregating updates over some amount of time, a single snapshot might miss unstable routing information. **Bi-lateral Peering:** Our analysis of bi-lateral peering reachability relied on sampled peering LAN traffic data and inferred business relationships. While we used an entire week of traffic data to partially overcome the problem of missing traffic for existing routes, we likely still missed a few routes as (1) they genuinely did not receive any traffic during our observation period or (2) they small amounts of traffic yet the sampling algorithm did not incorporate any of their packets. While we did our best to improve the coverage of inferred business relationships, we can not guarantee for the correctness of the business inference algorithm. While both algorithms were shown to provide high-quality inferences on public data [33,47], we utilize them in a rather different context which could potentially lead to iimpairments in their performance [64]. **Private Peering:** For the inference of private peering routes, we used a very small set of looking glasses and queried them in a restrictive manner. Especially for our findings regarding the summed reachability, our observations can only be seen as a lower bound. If our number of vantage points would have been significantly higher and we could have triggered queries at a high rate, the amount of private peering prefixes would have certainly increased leading to overall higher estimates for the total achievable reachability. **Regional Importance Bias:** The utilized data sets to infer peering relations and qualify the importance of IPs and prefixes (see Sect. 5 and Sect. 6) are biased towards the European service region. While it is for the conducted analysis required to compare reachability at IXPs and relevance (ISP data set and DNS) in the very same region, it may not necessarily apply to others. As different cultures may have unique eyeball behaviors, a traffic-based ranking for other large eyeball networks around the world may lead to different prefixes especially in the lower part of the top-10k ranking. As address resolution is often location-skewed (e.g., due to DNS load balancing) our domain-based ranking is likely biased towards norm-prefixes primarily used in the European region. While we expect unmatching biases (e.g., comparing American top lists to European IXP) to lower the overall top list coverage based on, e.g., routing policy differences [34], we do not expect that such a comparison would yield considerable differences.

7 Discussion

Our results suggest that networks that peer at one of the larger IXPs can indeed move most traffic to bi-lateral peerings, yet (especially for IPv6) not all prefixes that serve a high number of domains are reachable via peering. While an assessment of the quality of those available peering relationships (i.e., the capacity and latency guarantees they provide) goes beyond the scope of this work, previous works already hinted at certain obtainable benefits [2], e.g., Schlinker et al. [69] showed that the latencies for 10% of Facebook's traffic can be decreased by up to 10ms when switching from transit to peering routes.

That many high-volume prefixes can be served via bi-lateral peering at IXPs is strongly correlated with the observation that Hypergiants—large content providers such as Google, Facebook, or Amazon [12]— interconnect at tens (if not hundreds) of IXPs (see PeeringDB). According to Pujol et al. [65], these relatively few Hypergiants can be responsible for up to 80% of all ingress traffic of large eyeball networks.

Similar to hypergiants, the routes of many lower-tier networks are also available via peering. To them, broadly announcing their routes allows them to reduce the volume of ingress traffic delivered via some of their transit providers. Over time, such an approach may transform an asymmetric traffic ratio into a symmetric one, and allows these networks to re-negotiate their previous transit providers into a peering relationship.

In contrast, we observe that many of the domain-based top prefixes belong to large transit providers and Tier-1s. To reach those prefixes, IXP members often still have to rely on transit.

But how do those findings relate to different types of networks? **Large networks and hypergiants** already established thousands of peering connections [6] and use sophisticated traffic engineering strategies [70,78] among those connections. Their egress traffic mapping is already automated to a degree where adding new peers does not pose a challenge anymore which leads to constant growth of their peering edges and continuous dwindling of dependence on their transit connections.

In contrast, **small (access) networks** may rely on a few border routers operated mostly manually by a small group of network engineers. Adding new bi-lateral peers for these networks often poses a challenge in terms of resources and network complexity (operational costs). Hence, despite our findings, many of such networks may only peer with a Route Server and a few carefully selected bi-lateral peers on purpose. To them, the reduced supplier cost that comes with sophisticated peering is often not worth the increasing added operational complexity.

Medium-Sized Networks. (e.g., smaller national service providers) sit in between those two extremes. While many of them have neither automated their egress traffic mapping nor their peer acquisition yet, they are typically run by competent IT staff capable of anticipating how much their network would benefit from a particular peer. The earlier those networks transition from a few

expensive yet feature-rich routers to a distributed fleet of cheaper routers (with potentially partial visibility), the sooner they can quickly scale their peering edge allowing them to take full advantage of the opportunities provided by large IXPs.

8 Conclusion

Throughout this paper, we analyzed the routes available via multi-lateral, bi-lateral, and private peering. For multi-lateral peering, we analyzed Route Server snapshots from eight of the world's largest peering LANs and showed that most of their routes lead to out-of-continent locations via three or more AS hops. While remote peering might be a major contributor to the geographic distance of Route Server destinations, we observe that close and distant IXP members alike provide lengthy, unattractive routes to the Route Server. When comparing those findings to peering LAN traffic, obtained through a collaboration with one large IXP, we saw that mostly one-hop routes saw substantial traffic. In fact, we observed that 25% and 77% of IPv4 and IPv6 Route Server prefixes with at least four hop long paths see no traffic at all. This indicates that even though Route Servers provide many routes, most IXP members only make use of local routes. Afterwards, we used two heuristic-based methodologies to infer bi-lateral and private peering routes from the IXP's peering LAN traffic. During our inferences, we carefully isolated transit connections that were established over the peering LAN—a phenomenon previously reported by Ager et al. [1]. Based on our inference, we observe that at least 19.8, 57.1, and 57.4% (37.3, 37.4, 37.8%) of all routed IPv4 (IPv6) address space can be reached at our IXP via multi-lateral, bi-lateral, and private peering, respectively. Those results provide practical contrast to the 70+ % reachability theoretically calculated by Böttger et al. [10]. Finally, we show that almost all of the top 10k egress prefixes of a large European eyeball network can be reached via bi-lateral peerings. In contrast, we also find that up to 15% of top 10k domain-serving prefixes can not be reached via any type of peering at our IXP. Notably, we observe that most of these prefixes belong to large transit and Tier 1 providers.

References

1. Ager, B., Chatzis, N., Feldmann, A., Sarrar, N., Uhlig, S., Willinger, W.: Anatomy of a large European IXP. In: Proceedings of the ACM SIGCOMM 2012 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, pp. 163–174 (2012)
2. Ahmed, A., Shafiq, Z., Bedi, H., Khakpour, A.: Peering vs. transit: performance comparison of peering and transit interconnections. In: 2017 IEEE 25th International Conference on Network Protocols (ICNP), pp. 1–10. IEEE (2017)
3. Alexa: The top 500 sites on the web (2021). https://www.alexa.com/topsites. Accessed 21 June 2021

4. AMS-IX: Total traffic statistics (2021). https://stats.ams-ix.net/index.html. Accessed 27 June 2021. Archived version. https://web.archive.org/web/20210627072325/stats.ams-ix.net/index.html
5. Ariemma, L., Liotta, S., Candela, M., Di Battista, G.: Long-lasting sequences of BGP updates. In: Hohlfeld, O., Lutu, A., Levin, D. (eds.) PAM 2021. LNCS, vol. 12671, pp. 213–229. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-72582-2_13
6. Arnold, T., et al.: Cloud provider connectivity in the flat internet. In: Proceedings of the ACM Internet Measurement Conference, pp. 230–246 (2020)
7. Augustin, B., et al.: Avoiding traceroute anomalies with Paris traceroute. In: Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement, pp. 153–158 (2006)
8. BGP4.as: BGP looking glasses for IPv4/IPv6, traceroute & BGP route servers (2021). https://www.bgp4.as/looking-glasses. Accessed 21 June 2021
9. bgplookingglass.com: BGP looking glass database (2021). http://www.bgplookingglass.com/. Accessed 21 June 2021
10. Böttger, T., et al.: The elusive internet flattening: 10 years of IXP growth. arXiv e-prints (2018)
11. Böttger, T., Cuadrado, F., Tyson, G., Castro, I., Uhlig, S.: Open connect everywhere: a glimpse at the internet ecosystem through the lens of the Netflix CDN. ACM SIGCOMM Comput. Commun. Rev. **48**(1), 28–34 (2018)
12. Böttger, T., Cuadrado, F., Uhlig, S.: Looking for hypergiants in peeringDB. ACM SIGCOMM Comput. Commun. Rev. **48**(3), 13–19 (2018)
13. Brito, S.H.B., Santos, M.A.S., Fontes, R.R., Perez, D.A.L., Rothenberg, C.E.: Dissecting the largest national ecosystem of public internet eXchange points in Brazil. In: Karagiannis, T., Dimitropoulos, X. (eds.) PAM 2016. LNCS, vol. 9631, pp. 333–345. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-30505-9_25
14. CAIDA: The CAIDA as relationships dataset, 2021/06 (2021). https://publicdata.caida.org/datasets/as-relationships/. Accessed 21 June 2021
15. CAIDA: The CAIDA UCSD as classification dataset, 2021–04-01 (2021). https://www.caida.org/catalog/datasets/as-classification
16. CAIDA: The CAIDA UCSD as to organization mapping dataset, 2021/06 (2021). https://www.caida.org/data/as_organizations. Accessed 21 June 2021
17. CAIDA: The CAIDA UCSD peeringDB dataset, 2010/08-2021/06 (2021). https://www.caida.org/catalog/datasets/peeringdb. Accessed 21 June 2021
18. CAIDA: Routeviews prefix to as mappings dataset for IPv4 and IPv6, 2021/06. https://www.caida.org/datasets/routeviews-prefix2as/. Accessed 21 June. 2021
19. Cardona Restrepo, J.C., Stanojevic, R.: IXP traffic: a macroscopic view. In: Proceedings of the 7th Latin American Networking Conference, pp. 1–8 (2012)
20. Castro, I., Cardona, J.C., Gorinsky, S., Francois, P.: Remote peering: more peering without internet flattening. In: Proceedings of the 10th ACM International on Conference on Emerging Networking Experiments and Technologies, pp. 185–198 (2014)
21. Chandrasekaran, B., et al.: Alidade: IP geolocation without active probing. Department of Computer Science, Duke University, Technical report CS-TR-2015.001 (2015)
22. Chatzis, N., Smaragdakis, G., Böttger, J., Krenc, T., Feldmann, A.: On the benefits of using a large IXP as an internet vantage point. In: Proceedings of the 2013 Conference on Internet Measurement Conference, pp. 333–346 (2013)

23. Chiesa, M., di Lallo, R., Lospoto, G., Mostafaei, H., Rimondini, M., Di Battista, G.: PrIXP: preserving the privacy of routing policies at internet exchange points. In: 2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), pp. 435–441. IEEE (2017)
24. Chiu, Y.C., Schlinker, B., Radhakrishnan, A.B., Katz-Bassett, E., Govindan, R.: Are we one hop away from a better internet? In: Proceedings of the 2015 Internet Measurement Conference, pp. 523–529 (2015)
25. CISCO: Cisco umbrella 1 million (2021). https://umbrella.cisco.com/blog/cisco-umbrella-1-million. Accessed 21 June 2021
26. COUGAR: Cougar looking glass utility (2021). https://github.com/Cougar/lg. Accessed 21 June 2021
27. DE-CIX: DE-CIX Frankfurt statistics (2021). https://www.de-cix.net/en/locations/germany/frankfurt/statistics. Accessed 27 June 2021. Archived version. https://web.archive.org/web/20210620110006/www.de-cix.net/en/locations/germany/frankfurt/statistics
28. Dietzel, C., Feldmann, A., King, T.: Blackholing at IXPs: on the effectiveness of DDoS mitigation in the wild. In: Karagiannis, T., Dimitropoulos, X. (eds.) PAM 2016. LNCS, vol. 9631, pp. 319–332. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-30505-9_24
29. Feldmann, A., et al.: The lockdown effect: implications of the COVID-19 pandemic on internet traffic. In: Proceedings of the ACM Internet Measurement Conference, pp. 1–18 (2020)
30. Feng, G., Seshan, S., Steenkiste, P.: UNARI: an uncertainty-aware approach to as relationships inference. In: Proceedings of the 15th International Conference on Emerging Networking Experiments and Technologies, pp. 272–284 (2019)
31. Giotsas, V., Dhamdhere, A., Claffy, K.C.: Periscope: unifying looking glass querying. In: Karagiannis, T., Dimitropoulos, X. (eds.) PAM 2016. LNCS, vol. 9631, pp. 177–189. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-30505-9_14
32. Giotsas, V., Dietzel, C., Smaragdakis, G., Feldmann, A., Berger, A., Aben, E.: Detecting peering infrastructure outages in the wild. In: Proceedings of the Conference of the ACM Special Interest Group on Data Communication, pp. 446–459 (2017)
33. Giotsas, V., Luckie, M., Huffaker, B., Claffy, K.: IPv6 AS relationships, cliques, and congruence. In: Mirkovic, J., Liu, Y. (eds.) PAM 2015. LNCS, vol. 8995, pp. 111–122. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-15509-8_9
34. Giotsas, V., Zhou, S., Luckie, M., Claffy, K.: Inferring multilateral peering. In: Proceedings of the Ninth ACM Conference on Emerging Networking Experiments and Technologies, pp. 247–258 (2013)
35. He, Y., Siganos, G., Faloutsos, M., Krishnamurthy, S.: Lord of the links: a framework for discovering missing links in the internet topology. IEEE/ACM Trans. Netw. **17**(2), 391–404 (2008)
36. Houston, G.: Address span metrics (2021). https://bgp.potaroo.net/as6447/. Accessed 27 June 2021
37. ipinsight.io: Looking glass (2021). https://whois.ipinsight.io/looking-glass/. Accessed 21 June 2021
38. IX.br: Total traffic (2021). https://ix.br/agregado/. Accessed 27 June 2021. Archived version. https://web.archive.org/web/20210627071318/ix.br/agregado/
39. Jin, Y., Scott, C., Dhamdhere, A., Giotsas, V., Krishnamurthy, A., Shenker, S.: Stable and practical {AS} relationship inference with problink. In: 16th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 2019), pp. 581–598 (2019)

40. Jin, Z., Shi, X., Yang, Y., Yin, X., Wang, Z., Wu, J.: TopoScope: recover as relationships from fragmentary observations. In: Proceedings of the ACM Internet Measurement Conference, pp. 266–280 (2020)
41. Karagiannis, T., Papagiannaki, K., Faloutsos, M.: BLINC: multilevel traffic classification in the dark. In: Proceedings of the 2005 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, pp. 229–240 (2005)
42. Kernen, T.: Looking glass (2021). http://traceroute.org/. Accessed 21 June 2021
43. Lakhina, A., Papagiannaki, K., Crovella, M., Diot, C., Kolaczyk, E.D., Taft, N.: Structural analysis of network traffic flows. In: Proceedings of the Joint International Conference on Measurement and Modeling of Computer Systems, pp. 61–72 (2004)
44. Liu, S., Foster, I., Savage, S., Voelker, G.M., Saul, L.K.: Who is .com? learning to parse WHOIS records. In: Proceedings of the 2015 Internet Measurement Conference, pp. 369–380 (2015)
45. Lodhi, A., Larson, N., Dhamdhere, A., Dovrolis, C., Claffy, K.: Using peeringDB to understand the peering ecosystem. ACM SIGCOMM Comput. Commun. Rev. **44**(2), 20–27 (2014)
46. lookinglass.org: BGP looking glass services (2021). https://lookinglass.org/. Accessed 21 June 2021
47. Luckie, M., Huffaker, B., Dhamdhere, A., Giotsas, V., Claffy, K.: AS relationships, customer cones, and validation. In: Proceedings of the 2013 Conference on Internet Measurement Conference, pp. 243–256 (2013)
48. Majestic: The majestic million (2021). https://majestic.com/reports/majestic-million/. Accessed 21 June 2021
49. Marcos, P., Chiesa, M., Dietzel, C., Canini, M., Barcellos, M.: A survey on the current internet interconnection practices. ACM SIGCOMM Comput. Commun. Rev. **50**(1), 10–17 (2020)
50. Marcos, P., et al.: Dynam-IX: a dynamic interconnection exchange. In: Proceedings of the 14th International Conference on Emerging Networking EXperiments and Technologies, pp. 228–240 (2018)
51. Marder, A., Luckie, M., Dhamdhere, A., Huffaker, B., Claffy, K., Smith, J.M.: Pushing the boundaries with bdrmapIT: mapping router ownership at internet scale. In: Proceedings of the Internet Measurement Conference 2018, pp. 56–69 (2018)
52. MAXMIND: Geoip2 city accuracy (2021). https://www.maxmind.com/en/geoip2-city-accuracy-comparison?country=&resolution=250&cellular=excluding. Accessed 27 June 2021. Archived version. https://web.archive.org/web/20210627075223/www.maxmind.com/en/geoip2-city-accuracy-comparison?country=&resolution=250&cellular=excluding
53. MAXMIND: Geolite2 free geolocation data (2021). https://dev.maxmind.com/geoip/geolite2-free-geolocation-data. Accessed 27 June 2021
54. Mazoyer, G., Schmidt, M., Correa, A.J., et al.: Respawner looking glass utility (2021). https://github.com/gmazoyer/looking-glass. Accessed 21 June 2021
55. Naab, J., Sattler, P., Jelten, J., Gasser, O., Carle, G.: Prefix top lists: gaining insights with prefixes from domain-based top lists on DNS deployment. In: Proceedings of the Internet Measurement Conference, pp. 351–357 (2019)
56. RIPE NCC: Routing information service (RIS) (2021). https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris. Accessed 21 June 2021

57. Nomikos, G., et al.: O peer, where art thou? Uncovering remote peering interconnections at IXPs. In: Proceedings of the Internet Measurement Conference 2018, pp. 265–278 (2018)
58. Norton, W.B.: The Internet Peering Playbook: Connecting to the Core of the Internet. DrPeering Press (2014)
59. Oliveira, R., Pei, D., Willinger, W., Zhang, B., Zhang, L.: The (in) completeness of the observed internet AS-level structure. IEEE/ACM Trans. Netw. **18**(1), 109–122 (2009)
60. Oliveira, R.V., Zhang, B., Zhang, L.: Observing the evolution of internet AS topology. In: Proceedings of the 2007 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, pp. 313–324 (2007)
61. University of Oregon: University of Oregon route views project (2021). http://www.routeviews.org/routeviews/. Accessed 21 June 2021
62. PeeringDB: PeeringDB (2021). https://www.peeringdb.com. Accessed 21 June 2021
63. PeeringDB: Hivelocity Inc (2021). https://web.archive.org/web/20220130161818/www.peeringdb.com/net/2159
64. Prehn, L., Feldmann, A.: How biased is our validation (data) for AS relationships? In: Proceedings of the ACM Internet Measurement Conference, p. TBA (2021)
65. Pujol, E., Poese, I., Zerwas, J., Smaragdakis, G., Feldmann, A.: Steering hypergiants' traffic at scale. In: Proceedings of the 15th International Conference on Emerging Networking Experiments and Technologies, pp. 82–95 (2019)
66. RADb: aut-num: As213045 (2021). https://web.archive.org/web/20220130152317/www.radb.net/query?keywords=AS213045
67. Randall, A., et al.: Trufflehunter: cache snooping rare domains at large public DNS resolvers. In: Proceedings of the ACM Internet Measurement Conference, pp. 50–64 (2020)
68. Richter, P., Smaragdakis, G., Feldmann, A., Chatzis, N., Boettger, J., Willinger, W.: Peering at peerings: on the role of IXP route servers. In: Proceedings of the 2014 Conference on Internet Measurement Conference, pp. 31–44 (2014)
69. Schlinker, B., Cunha, Í., Chiu, Y.C., Sundaresan, S., Katz-Bassett, E.: Internet performance from Facebook's edge. In: Proceedings of the Internet Measurement Conference, pp. 179–194 (2019)
70. Schlinker, B., et al.: Engineering egress with edge fabric: steering oceans of content to the world. In: Proceedings of the Conference of the ACM Special Interest Group on Data Communication, pp. 418–431 (2017)
71. Schomp, K., Bhardwaj, O., Kurdoglu, E., Muhaimen, M., Sitaraman, R.K.: Akamai DNS: Providing authoritative answers to the world's queries. In: Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication, pp. 465–478 (2020)
72. Shafiq, M.Z., Ji, L., Liu, A.X., Wang, J.: Characterizing and modeling internet traffic dynamics of cellular devices. ACM SIGMETRICS Perform. Eval. Rev. **39**(1), 265–276 (2011)
73. Shin, D., Guaitanele, R.G., Vine, B.: HSDN PHP looking glass (2021). https://github.com/hsdn/lg. Accessed 21 June 2021
74. Snijders, J.: PeeringDB accuracy: is blind faith reasonable? (2013). https://archive.nanog.org/sites/default/files/wed.general.peeringdb.accuracy.snijders.14.pdf. Accessed 24 June 2021
75. subnets.ru: Looking glass list (2021). http://subnets.ru/wrapper.php?p=1. Accessed 21 June 2021

76. Virtua.Cloud: Our network (2022). https://web.archive.org/web/20220130154537/www.virtua.cloud/our-infrastructure/our-network
77. Wikipedia: Tier 1 network, May 2021. https://en.wikipedia.org/wiki/Tier_1_network
78. Yap, K.K., et al.: Taking the edge off with espresso: scale, reliability and programmability for global internet peering. In: Proceedings of the Conference of the ACM Special Interest Group on Data Communication, pp. 432–445 (2017)
79. Yeganeh, B., Durairajan, R., Rejaie, R., Willinger, W.: How cloud traffic goes hiding: a study of Amazon's peering fabric. In: Proceedings of the Internet Measurement Conference, pp. 202–216 (2019)