On the Asymmetry of Internet eXchange Points - Why Should IXPs and CDNs Care?

Leandro M. Bertholdo*, Sandro L. A. Ferreira§, João M. Ceron†,
Lisandro Zambenedetti Granville‡, Ralph Holz*, Roland van Rijswijk-Deij*¶
*University of Twente, Enschede, The Netherlands - {l.m.bertholdo, r.holz, r.m.vanrijswijk}@utwente.nl
§Federal Institute of Technology, Porto Alegre, Brazil - sandro@ifrs.edu.br
†Botlog.org, Brazil - ceron@botlog.org
‡Federal University of Rio Grande do Sul, Porto Alegre, Brazil - granville@inf.ufrgs.br
¶NLnet Labs, Amsterdam, The Netherlands

***Abstract*—Internet eXchange Points (IXPs) provide an infrastructure where content providers and consumers can freely exchange network traffic. The main incentive for connecting to an IXP is to decrease costs and improve the user experience by having content closer to consumers. Despite these benefits, several small Content Delivery Networks (CDNs) avoid exchanging traffic on IXPs due to the poor routing quality via IXP paths. In this paper, we investigate how traffic asymmetry affects the quality of paths. IXP asymmetry occurs when traffic is sent (or received) via a direct IXP peering but received (or sent) on an alternative path outside the IXP. We employ a new method to quantify a symmetry rate for an IXP, which we evaluate on five IXPs. Our method covers three times more ASes than alternatives, such as using RIPE ATLAS. Our results show that IXPs have 15% asymmetric paths at a distance of one AS hop, *i.e.*, when sending traffic to a given peer on the IXP, 15% of this traffic will be responded via a transit AS that does not use the IXP path. We also identify *deaf neighbors*, *i.e.*, ASes that never return traffic to the IXP. We identify egress-only paths as a major cause of asymmetries and show that this occurs only for a small number of ASes. We also quantify the impact of traffic asymmetry at IXPs in terms of latency and show that traditional traffic engineering on IXP prefixes can actually make route quality worse.**

I. Introduction

The increasing demand for bandwidth from Internet users requires application developers and Internet providers to maximize the use of available bandwidth. To that end, Web applications are in the process of migrating to HTTP/3 [1], using QUIC as a transport protocol [2], and Content Delivery Networks (CDNs) cache to place the content closer to users. Both CDNs and Internet providers use traffic engineering to optimize their networks [3], sometimes with conflicting goals.

At the network level, solutions to maximize bandwidth usage (*e.g.*, load balancing and traffic engineering), however, may lead to the side effect of *routing asymmetry*. Asymmetric routing occurs when packet flows between two endpoints traverse different physical links in the forward and return directions [4]. Routing asymmetry creates problems, for example, in (*a*) estimating one-way latency between hosts, (*b*) troubleshooting routing, (*c*) developing strategies for network optimization [5] [6], (*d*) detecting IP spoofing [7], and (*e*) establishing routing-based DDoS defenses [8]. Asymmetric routing can even be considered a "routing pathology" [9].

Several efforts have addressed diverse aspects of routing asymmetry, including how it degrades TCP performance [10], affects TCP anycast networks [11], impacts security appliances [12], and leads to wrong round-trip-time (RTT) estimates when one assumes that forward and reverse delays are half of the RTT [13]. Still, asymmetry exists and is generally an unwanted feature that negatively impacts the network.

Although asymmetry can happen at any multihomed network, its effect is more damaging when it involves anycast networks in CDNs or cloud providers. Anycast networks use the same IP address in different physical locations, peering with dozen IXPs and transit providers. When one AS prefer transit rather than the direct IXP path for an anycast network, packets can be ended in another anycast site miles away. This situation impacts the user experience and can lead to financial losses. – when a service is hosted on a cloud provider and a user from one continent is mistakenly redirected to another continent – some cloud providers charge up to four times more for intercontinental traffic. Selecting the best path is a challenge for anycast networks.

In a previous work [14], we compared IXPs in terms of coverage, prefix overlapping, and ASes preference to deliver traffic in one or another IXP. We also identify the existence of traffic asymmetry on IXPs and how it can be hurtful during outages. In this paper, we investigate ASes asymmetrical behavior adopting open policy routing on IXPs, and how asymmetry affects IXP customers, especially anycast networks used by CDNs and DNS providers. The identification and quantification of asymmetric paths is a first step to providing a metric to compare IXPs by quality, and better understand what leads small CDNs to prefer leaving the IXP to move behind a transit provider, a movement contrary to research showing that peering outperform transit [15]. The information about traffic asymmetry in IXPs can provide new insights for improvements in the way eyeball networks and CDNs relate to each other, as well as encourage the creation of new services for IXPs.

To observe IXP traffic asymmetry, we developed a methodology that monitors asymmetry over time. This methodology dispenses with private data sources (*e.g.*, flow data from IXPs) and can be deployed by any anycast network. Our contributions are as follows. (i) We present the first evaluation

978-3-903176-51-5

of traffic asymmetry at IXPs in terms of ingress and egress traffic to/from prefixes and ASes, investigating the case of open peering; (ii) we demonstrate the existence of fully asymmetrical ASes (mute and deaf neighbors); (iii) we provide takeaways based on the IXPs' characteristics we identified.

II. Defining route quality

Route quality or *quality of routes* is a term/concept adopted by Content Delivery Network (CDN) providers to refer to quality-related metrics they attribute to routes [16]. The concept only applies to inter-domain routing at CDNs and shapes their *best path* selection process. To the best of our knowledge, the terminology is not formally defined in prior literature.

Route quality adds quality-related metrics to a routing prefix. While some internal routing protocols (*e.g.,* EIGRP) allow metrics such as delay, load, and reliability, external protocols (*e.g.,* BGP) do not. Rather, external protocols have a "political" approach: while they provide means to conduct traffic, actual decisions are operational and not necessarily informed by performance metrics. New software-defined networks gather quality-related information from external sources (*e.g.,* passive measurements from websites or DNS resolvers) and use that data to build a secondary table of prefixes, mixing routing paths and quality metrics associated with each path. The most common metric is delay, frequently obtained from round-trip-time (RTT) measurements. In this case, quality is measured at the application layer, although it reflects changes at the routing layer. From a routing perspective, quality trumps routing attributes when selecting the best path.

CDNs consider a route *poor* or *low quality* when *the best BGP path* received from a transit provider or IXP affects the application in one of these aspects:

(i) Reliability: While routing tables include just "feasible" paths, some paths cannot reach the application at the destination. This situation occurs when a middle element applies filters or lowers the quality of service (QoS) to the destination. Janssen [17] describes several situations where ISP's traffic engineering leads to routes of poor quality for certain types of CDNs.

(ii) Stability: In traditional external routing, "the oldest route" is considered a better quality route, so longevity is used as a tie-breaker criterion. In the context of route quality we also need to consider jitter and packet loss.

(iii) Latency: Selecting a path based on response time is a target to improve user experience. User experience is (here) a function of all devices' latency, physical distance, and propagation time between client and server. Some internal routing protocols already consider a latency-based path selection [18], but external routing protocols do not. This is the main metric in quality-aware routing.

(iv) Throughput: While a path data rate is a natural metric, we do not know any CDN currently adopting.

The CDNs' wish for better route quality often results in physical changes to their network, such as adding new sites, establishing new peering agreements, or contracting additional transit providers. The main goal is often to increase the number of available paths to be selected. Large CDNs adopt complex mechanisms to compute route quality, or they adopt direct peering strategies. Both are expensive. Part of this decision process is based on *which server has the best path to that user*, where *'the best'* is a function of all aspects identified (reliability, stability, latency and delay).

III. Routing asymmetry on IXPs

IXPs are a natural place through which to deliver content as they are often a central interconnection point in a locality/region. These characteristics make IXPs more likely to provide high-quality routes to a region.

However, the operator community has previously reported the limitations of IXP traffic exchanges, citing the issue of low-quality routes [16]. The difficulty associated with identifying "poor quality routes" has led several CDNs to avoid adopting open peering at IXPs. Larger CDNs have clearly stated their preference for private peering over the open policy option [19]. The *"poor route quality"* issue is frequently caused by saturated links, competition between providers, and also the *"remote peers"* problem. The latter is characterized in the literature as peers adding a considerable interface delay to an AS that is expected to be in the same LAN/Ethernet [20]. To address the problem of remote peers, IXPs have begun to take more explicit steps, such as using BGP communities to encode latency in the announced prefixes. The peer routers' RTT [21] is a first step to improving the route quality on IXPs.

In our reasearch, we identify a number of ASes that avoid receiving traffic on IXPs, as we show in Section VII. We argue that asymmetrical traffic requires attention. Asymmetrical ASes cannot be detected by the same techniques used to identify remote peers. They also do not fit the assumption that a path's delay in one direction is half of the RTT, and invalidate the peer router RTT as a reference for that AS.

In the context of IXPs, symmetry refers to the preference of each IXP customer to deliver traffic towards peers through the IXP, and also receive traffic back from them via the IXP. We can define IXP (a)symmetry as follows and then compute a "(a)symmetry rate" for it. Let AS_A be an AS with an open peering policy on an IXP. If a different AS AS_B peers with AS_A at the IXP, AS_B can deliver traffic for any of the prefixes announced by AS_A. AS_B also announces its prefixes to AS_A. We define AS_A as *symmetric* if it delivers any traffic in response to traffic from AS_B via the peering on the IXP, and as *asymmetric* if the returned traffic is delivered through some other, different path (*e.g.,* using transit provider link).

While we can forecast the IXP egress traffic based on the IXP routing table, inferring ingress traffic is challenging. Thus, to measure IXP symmetry, we use an anycast network. We provide the details in Section IV. One of our goals is to determine the (a)symmetry for each IXP, and for each AS associated to that IXP. We defined *IXP symmetry rate* based on the behavior of all individual /24 networks. The *AS symmetry* is computed considering the volume of (a)symmetric networks we mapped for that AS. This process is detailed in Section VI.

IXP traffic symmetry is affected by diverse factors, *e.g*, hot-potato-routing, configuration mistakes, lack of routing knowledge, temporary IXP connection issues, or commercial interests. Also, some IXPs recommend that their participants employ more specific prefixes announced to IXP route-servers to increase the volume of traffic exchanged. We evaluate the impact of such a policy in subsection VI-B. We also identify "deaf" and "mute" neighbors. These are peers that never return traffic through the IXP or send traffic from networks not announced on the IXP respectively (see subsection VI-E).

IV. Measurement Architecture and Datasets

Figure 1 depicts the methodology we use to determine traffic symmetry. We have three entities: an *ixp* anycast site connected to the IXP infrastructure, a *drain* anycast site connected to a transit provider, and a *pinger*, an application used to generate ICMP requests to hosts on a hitlist. Once an ICMP request is generated on *pinger* using the anycasted IP as source, an ICMP response will go to either the *drain* or *ixp* site, depending on the routing preferences of each network/AS. The *drain site* announces the anycast prefix through a transit provider to the entire Internet. The announcement uses AS-path prepending [22] to work as a last resource route. The *ixp site* announces an equal size, or a more specific prefix than the *drain*. Varying the pinger location enables us to identify other anycast ASes (subsection V-C). Varying the prefix size enables us to identify ingress traffic issues, such as the case of deaf or mute neighbors. Our method is not influenced by anti-spoofing filters as it uses the anycast address as the source.

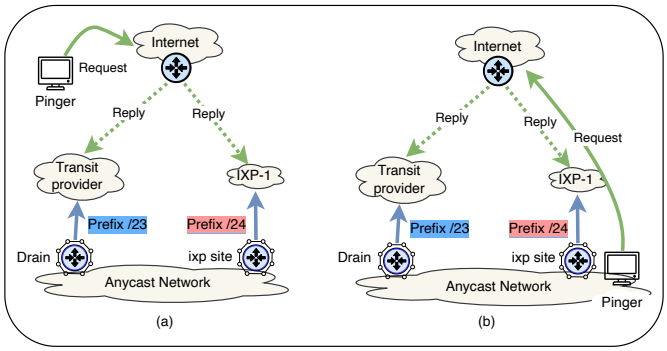


Fig. 1: Anycast setup for asymmetry measurement.

We used the Tangled [23] testbed to implement our *ixp* anycast site. Tangled has multiple sites and provides an interface to set up and manage traffic engineering. It uses Verfploeter [24] for the anycast mapping and information from the Internet address census [25] as a hitlist. The hitlist contains one ICMP responsive IP address to almost all /24 network on the Internet, it provides our individual network sample (subsection VI-A). Table I summarizes the setup used in our experiments. Among the available sides on Tangled, we selected the ones connected to IXPs. In the table, we describe the connection characteristics of these nodes. *rank* is the value that PeeringDB assigns to IXPs *ASes* is the number of ASes connected at the IXP and *Traffic* is the total traffic exchanged there. *Open peering* shows the average number of active open peering ASes during the period of our study.

IXP	Rank	ASes	Open Peering	Traffic	Website
IX.br/SP	1	2,324	2,298	15 Tbps	ix.br
AMS-IX	3	847	571	11 Tbps	ams-ix.net
LINX	4	733	554	7 Tbps	linx.net
SIX	9	337	246	2 Tbps	seattleix.net
IX.br/RS	46	302	296	0.5 Tbps	ix.br

TABLE I: Selected IXPs by PeeringDB Ranking (May-2022)

When we *originate and receive* an ICMP ping back at the IXP site (Figure 1-B), we consider that network symmetric. When we *originate the packet at the IXP site* and have a route to that destination, but receive the *answer on our drain site*, we consider that network asymmetric (only egress traffic). To detect reverse asymmetry (ingress asymmetry), we *generate packets using the pinger outside the IXP* site (Figure 1-A). If we receive the ICMP responses *from networks that are not announced in the IXP routing table*, we have detected an asymmetric ingress-only configuration. On Section V we discuss choices, coverage, and limitations of our method.

By manipulating the prefix size on the IXP site—that is, using the same prefix size at the drain and IXP site, or a more specific at the IXP—allows us to quantify how IXP symmetry is affected by the use of more specific prefixes inside the IXP.

Datasets. We have full, daily measurements from 2022-04-18 to 2022-05-02 available [26]. Each daily dataset contains a less specific and more specific experiment for each IXP. We also include pinger configurations, peering-related information, IXP routing tables, and our result tables.

V. Challenges in Inferring Symmetry of IXP Peers

In the following, we discuss the challenges of inferring routing asymmetry and reason about alternative approaches. We also discuss possible variations of our method.

A. Why not traceroute?

Intuitively, the first tool one might consider measuring asymmetry is traceroute. To use traceroute to estimate AS-level symmetry on IXPs, one would need to originate traces from one IXP participant towards all others, and vice versa.

RIPE Atlas has (May-2022) the best coverage in terms of vantage points. It covers over 7,000 Autonomous Systems [27], which would allow us to obtain the routing tables' view from these ASes toward the IXP. In a best case scenario, we would need one RIPE Atlas probe inside each IXP participant. However, we would still need to check whether we are using the IXP path, which is impossible in certain cases [28].

We carried out preliminary experiments using all available Atlas probes to test this approach. Despite the overall large number of probes, a relatively low number of traceroutes reach a prefix we announce through our IXPs, and many of the traceroutes share the last hop-AS toward the IXP. In the best case (AMS-IX), we could cover about 59% of ASes with an open peering policy. In the worst case (IX.br/SP), just 4% of

2,000 ASes could be covered with traceroutes from the Atlas probes. The median for our five IXPs was just 31%.

Contrasting this to our chosen approach, we can analyse the symmetry of 91% of all ASes in the best-case (AMS-IX), and 79% in the worst case (IX.br/SP). The median is 85%.

B. Can we use IXP network flow data?

As we want to compare symmetry differences between multiple IXPs, we need to build metrics that apply to all. IXPs in some countries consider flow data as private data due to legislation. Additionally, many IXPs collect only layer-2 flow data (sFlow or similar) [29]. Since any conversation has one flow in each direction, we cannot be sure the sampling includes both directions, this data would only allow us to give a lower bound on symmetry (the case where we see the flows in both directions). It would not allow us to make claims about asymmetry, as the absence of one of the flows can just as well be caused by sampling rather than asymmetry. Hence, using flow data is not a viable approach.

C. Traffic drain, vantage points, and CDN view

The key point about using anycast to compare and measure other networks on the Internet is the ability to emulate a world-wide backbone. Anycast allows us to send packets originating anywhere in the world without violating Routing Security Standards [30]. Manipulating where and when we generate and receive traffic allows us to use anycast to measurements.

Anycast-based measurements have been used before, for example, to discover and compare entities on the Internet [14], [31]. Here, we improve previous methods by adding a component (routing dynamics) while performing anycast measurements. We do this by announcing the anycast prefix in multiple places and using different prefix sizes. We can also vary where we originate ICMP polling traffic. The response to a packet sent from one site may return to that site or another anycast location, depending on how the polled node will choose the best path to the anycast address. Using anycast measurements gives us new possibilities. So we considered several aspects of anycast routing and topology:

(i) Several anycast sites active (mimicking an anycast CDN): We used this approach to identify other anycast networks (subsection VI-D). But, the same approach does not allow us to identify asymmetry on IXPs. The results are dependent on the number of anycast sites and transit providers we select.

(ii) Just *ixp-site*: This approach allows to estimate the total visibility of the IXP as described in [14], and identify the IXP symmetry. But it does not provide us with information about asymmetrical paths, since only one path is available. We cannot validate cases of *deaf* neighbors—they ignore prefixes from the IXP. To do so, we need to capture messages sent by other paths than IXP.

(iii) Adding a *transit-site*: This approach allows us to capture all answers from IXP peers. Gathering the additional answers to a second site allows us to validate cases of partially symmetric ASes, and *deaf/mute* neighbors. However, still a few IXP peers show a variation on the *symmetry rate* when we change the *transit-site*.

(iv) Prepending the *transit-site*: When we look at the IXP routing tables we identified 95% of the IXP routing table show paths length equal to or less than 10 (Section VII). Taking in account this value, the average Internet as-path size, and option of some providers in filter as-paths longer than 16 we prepended our *transit-site* announcement by 10. Minimizing the influence of the chosen transit provider. Even though, we notice a number of networks preferring the transit path. For example, when testing an IXP in South America, most networks in the US prefer to send replies to a *transit-site* in the US. This "traffic polarization" has also been identified in other studies [32].

The "drain" concept and manipulating prefix sizes. To overcome the polarization problem, we used a more specific prefix on the IXP. The unbalanced prefix size allows us to implement a "packet drain" concept. We turn our drain site (where we announce the less specific prefix) further into a "last resource route". Using this approach, we can identify the worst cases of routing asymmetry within each IXP and compare them with regular ingress asymmetry. In subsection VI-B we compare the use of less and more specific prefixes.

D. Routing Dynamics

During our early tests, we noticed path instabilities for far-away ASes. This resulted in some ASes oscillating between symmetrical and asymmetrical behavior. The number of ASes for which we observe this, however, is very small. For example, in our measurements of AMS-IX, only 0.5% of all AS paths announced to the IXP were affected on the days we analyzed symmetry. These cases were mostly distant ASes in Asia and Eastern Europe. Other studies also report AS instability on long paths [33]. IXP neighbors are stable.

E. Comparing with Internet asymmetry studies

Our solution aims to infer IXP symmetry just based on directly connected ASes. Previous studies on global Internet symmetry, using different techniques, reports low symmetry values of 10-35% [34], [35]. So, we applied our approach to all prefixes received on the IXPs to measure the asymmetry between the IXP and transit paths. We found similar values *e.g.,* LINX is just 30% symmetric considering all IXP-Cone.

F. Limitations and accuracy

We are limited by ICMP responsive ASes (89% of all active ASes). Also, we do not use a fully independent prefix. Our /23 prefix is part of a bigger /16 prefix advertised by SURFnet.

VI. IXP SYMMETRY RESULTS

Our goal is to measure the lower bound of asymmetry in IXPs for ASes that adopt open peering, regardless of the transit provider that each AS has. This information can help small CDNs and eyeball networks to identify cases where traffic engineering is ineffective or even harmful, considering that path asymmetry often has a negative effect on route quality.

We present our results in the following. We begin computing the *symmetry rate* for each IXP. Then, we check how individual ASes affect the overall rate and characterize those ASes in terms of symmetry. Finally, we analyze the impact on symmetry that using more or less specific announcements on the IXPs has, as well as the impact on RTT.

A. Network symmetry rate on IXPs

We compute the *IXP symmetry rate* by quantifying the behavior of all individual /24 networks we mapped (ICMP responsive addresses) from all peers at that IXP. The highest *symmetry rate* we could establish across IXPs was just about 88%. In total, we mapped more than 230,000 /24 prefixes in 2,800 ASes. This represents a median of 85% of all ASes connected to each IXP (subsection V-A).

In Table II, we summarize our results for the experiment where we advertised a more specific prefix. Here, the mean IXP *symmetry rate* is 83%. We observe, across all IXPs, an asymmetry of 11-21% of networks. This is highly influenced by networks that do not return traffic to the IXP (only egress). The egress traffic depends on whether we use the IXP paths or not, or whether we prefer the IXP path over our transit provider paths. The table shows the case of optimizing for both: we use a more specific prefix and always prefer the IXP path, which is a common IXP recommendation. This approach also yields the best symmetry values for all IXPs, which confirms the IXPs' recommendation in the context of traffic symmetry.

ixp	Neig. Net.	Symmetric	Only Ingress	Only Egress
AMS-IX	90,064	79.4%	6.6%	13.8%
LINX	66,040	88.5%	7.0%	4.2%
IX.br/RS	7,917	78.9%	1.1%	20.0%
SIX	31,286	88.1%	3.7%	8.1%
IX.br/SP	35,327	85.3%	1.7%	12.2%

TABLE II: Network symmetry using more specific prefix

Table III shows symmetry results when we use a prefix of same length for the *ixp* and *drain* sites (it is a /23, but drain is prepended as described before). We observe higher asymmetry rates in this experiment. The mean *symmetry rate* falls to 76%. Curiously, while symmetry on AMS-IX and LINX falls by 15%, the other IXPs show a smaller variation of under 3%. This suggests differences in the operational setups of ASes at the respective IXPs. We also receive traffic from fewer networks, which leads to a reduction in ingress asymmetry.

No IXP we measure comes even close to 100% symmetry, even though we follow the IXPs' recommendations for traffic engineering in all cases. We also already filter out all asymmetrical cases arising from other anycast prefixes (see subsection VI-D). Even so, we are unable to ever reach a 100% rate for ingress traffic. We contacted several operators of networks that we identified as not returning traffic to the IXP. They confirmed to us that they announce some prefixes to the IXP but filter the IXP announcements in some parts of their network. Two operators stated that this was a desired behavior, but did not disclose the reason - once a common reason is because traffic/contractual agreements. Three others identified a possible mistake in their configuration, and agreed that fixing the asymmetry would be an improvement for their networks. We discuss the more extreme cases of 100% asymmetric ASes in *deaf neighbors* in subsection VI-E.

ixp	Neig. Net.	Symmetric	Only Ingress	Only Egress
AMS-IX	85,967	63.2% ↓↓	2.1% ↓↓	34.4% ↑↑
LINX	65,258	74.3% ↓↓	6.0% ↓	19.3% ↑↑
IX.br/RS	7,903	76.7% ↓	0.9% →	22.2% ↑
SIX	31,310	86.3% ↓	3.4% →	10.1% ↑
IX.br/SP	34,984	85.4% →	0.8% →	13.6% ↑

TABLE III: Network symmetry using same size prefix

Takeaway: Some operators intentionally generate asymmetry, but more than half of those we contacted acknowledged misconfigurations.

B. Impact of more specific prefixes on traffic symmetry

A common technique used in traffic engineering is to break prefixes into smaller subprefixes to attract more traffic. Several IXPs incentivize this technique, although there is no consensus if this is a good strategy [36]. We obtained our best IXP symmetry rate using this strategy and analyze the impact of following it in more detail in this section.

When we use this technique and announce a more specific prefix in the IXP, we perceive a small difference in new /24 networks sending the answer back to the *ixp-site* (1%). The exception is AMS-IX, where we attracted traffic from 5,000 additional networks (an additional 4.7% - column Neig.Net on Table III and Table II). We believe that this outlier might be due to us operating a SurfNet subprefix (subsection V-F).

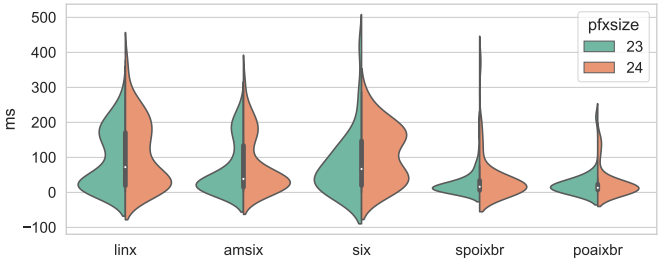


Fig. 2: More specific prefix impact on RTT

As expected, the more specific prefix attracts more traffic, reducing egress-only traffic. We evaluate the quality of these "new" paths. In Figure 2, we compare the RTT of the less-specific case (green) with the new answers received by our more-specific experiment (orange). The negative values are a graphical tweak to show the volume of measurements close to zero. All outlier values over 500ms are omitted for legibility; outliers are more common when using a more-specific prefix.

The mean RTT using both experiments does not show a significant difference. However, the standard deviation is high for new paths. In some cases, it is four times higher than in the less-specific experiment. In the case of IX.br/SP, the more-specific experiment attracted some low-latency paths, which

would be beneficial for CDNs. For SIX, there is a second peak of RTTs around 180ms. Here, the new paths are clearly *poor quality routes*. On other IXPs, we notice similar behavior to SIX and IX.br/SP, but in a less pronounced way.

While the use of a more specific prefix provides us with better symmetry, we also get a slightly lower route quality. Our results suggest it might be possible to improve symmetry by reducing the egress traffic rather than increasing ingress traffic by announcing a more specific prefix. In Section VII, we look at routing tables at IXPs and evaluate whether it is a good idea to give priority to IXP routes (adjust egress traffic).

Takeaway: The use of unbalance prefix between IXP/ISP are prone to attract routes with higher RTT.

C. Symmetry at AS level

Here, we investigate the AS-level asymmetry within an IXP. We do this to understand if asymmetry is a common phenomenon in the IXP or linked to only a few ASes. In the latter case, operators would be able to tackle asymmetry by ceasing to exchange traffic with asymmetric ASes or even cease to be part of the open peering policy in the IXP.

Our results show a slightly greater symmetry at AS level: the mean value for *symmetry by AS* is 86%. Asymmetry, on the other hand, is more concentrated in a few ASes. We observe this behavior across all of the IXPs we analyze. We apply our classification scheme (symmetric, asymmetric ingress, asymmetric egress) on ASes with a granularity of /24 and classify them. We say that a predominant category characterizes an AS. However, in cases where the classification is less compelling—*i.e.*, we find a similar distribution of /24—we call the classification result "hybrid".

In Table IV, we give the median values for open peering ASes, which had an established BGP session at the moment we measured (column ASes). This number is significantly smaller than that reported by the IXP administration (Table I).

ixp	ASes	Unk	Symm	Hybrid	Ingress	Egress
AMS-IX	472	28	86.0 % ↑	12	20	30
LINX	439	32	83.8 % ↓	10	22	35
IX.br/RS	220	18	94.1 % ↑↑	2	3	7
SIX	204	22	84.2 % ↓	9	12	26
IX.br/SP	1,879	261	90.7 % ↑	13	20	116

TABLE IV: ASes symmetry using more specific prefix

Column *Symm* shows the percentage of active symmetrical ASes—with at least one responsive network. Column *Unk* are those ASes with unclear behavior (without any samples) and *Hybrid, Ingress and Egress* refer to our classification already described. Additionally, we added arrows for easy comparison with Table II. Upward arrows mean that symmetrical networks are spread over many ASes within a IXP (asymmetry is concentrated). Downward arrows mean that symmetrical networks are linked to a few ASes (asymmetry is more spread). Double arrows indicate more than ten percentage points difference.

The symmetry value in Table IV reflects how many ASes are symmetrical. For example, in IX.br/RS—the smallest and more regional IXP—the symmetry by AS is 94%. When we compare this to symmetry at /24 level (78%), it is easy to conclude that just a few ASes (6%) are responsible for the asymmetry problem. The value also indicates that, if one was to connect at IX.br/RS, 94% of all peers would return traffic.

When showing our network and AS results to network operators, they were surprised by the small difference between the number of neighbor networks attracted when using less/more-specific prefix-size. One operator asked us to compare globally—without restricting reachability to neighbors. We performed this experiment for three IXPs. We registered between 4-8% of new networks in the /24 experiment compared to the /23 and just 1% new ASes.

Takeaway: In most cases, few ASes use to be responsible for asymmetry on IXPs.

D. Characterizing asymmetrical ASes

In this section, we investigate possible links between asymmetry and business of respective IXP participants. We used *ip2location* data to classify ASes' business type. Figure 3 shows symmetry by business type. The graph is normalized and shows the percentage of ASes for each case. CDNs and Mobile operators have the biggest asymmetric traffic.

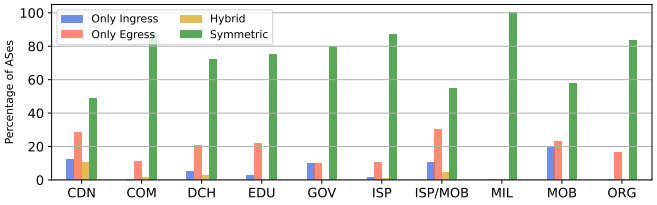


Fig. 3: All asymmetrical ASes by business type

Mobile operators use IXP multilateral peering as a second path, preferring to return traffic by another link, or on a preferred IXP. They also show cases of traffic sent to the IXP without announcing a corresponding route. These inconsistencies are expected since they often run large networks.

CDNs make extensive use of anycast. So much that it influences our results, as we discussed earlier (Section V). We expect anycasted prefixes to contribute significantly to asymmetry. In addition, we see extensive evidence of traffic engineering by CDNs. For example, the same prefix is announced from different ASes, or they use a specific subset of /24 prefixes one each IXP. These cases are hard to identify and we expect this to influence our egress-only metric. We note, however, that the evidence of ingress-only CDNs is rather unexpected. Unfortunately, we were unable to determine whether this is due to configuration mistakes or possibly a backup solution for serving traffic in the IXP's region.

In our characterization process, we filter out anycasted prefixes using the method described in [31]. In Figure 4, we visualize the impact of anycast on asymmetrical traffic on IXPs. The left-hand shows our results of asymmetrical networks including other anycast networks. The right-hand subfigure shows what happens when we remove prefixes that we detect as being anycasted. It is easy to identify

those anycasted networks that represent a significant parcel of asymmetrical networks we detected. So we exclude them from our results.

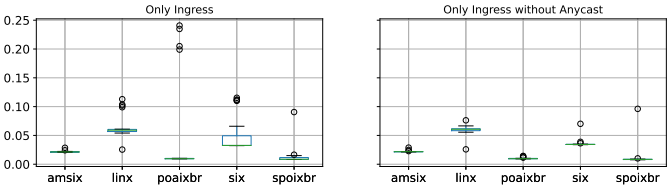


Fig. 4: Identifying anycasted CDN on ingress-only networks.

Takeaway: ISPs are more symmetric than expected. Mobile operators are the most asymmetrical and have good space for improvement on the IXPs we analyzed.

E. Deaf and Mute neighbors

We identified fully asymmetrical neighbors in both directions. Some always ignore our announcements on the IXP (deaf neighbors), and others forward traffic to us but never announce their prefixes (mute neighbors). After identifying this strange behavior, we contacted the IXP administration and several AS owners to validate our findings. We validated more than 50 ASes as mute or deaf from Table IV. The result of this validation is a graph showing almost zero inbound (deaf) or outbound (mute) traffic from the IXP's point of view.

We received answers from four IXP participants when we asked them to explain the reason for their behavior. Two recognized the mute/deaf effect as a router configuration mistake, and two reports they used the IXP as a secondary/backup path—they had announced only a prepended less specific prefix and use the IXP routes in case of a failure in their transit provider. The fraction of mute and deaf neighbors appears to be stable. The only fluctuation we observe are anycast prefixes changing category during our period of observation.

Takeaway: Deaf and mute neighbors may be linked to configuration mistakes or the use of the IXP as a backup path.

VII. Do IXP customers deprefer IXPs?

So far, we showed that egress-only traffic is a primary cause of asymmetry and that using more specific prefixes in announcements comes with drawbacks. Here, we investigate how IXP routing tables are linked to egress-only traffic. Our analysis applies to all prefixes announced by IXP peers.

A. General observations

Poor routing paths—when the AS does not return traffic—increase our egress-only asymmetry. On the other hand, deprefered paths can increase ingress-only asymmetry.

In our symmetry experiment, we maximize the ingress path (using a more specific prefix in the IXP) and the egress path (preferring the IXP paths). We took this decision as a way to isolate IXP paths from the influence of different transit providers. However, assuming that IXP paths should be always preferred is a recurrent subject of discussions [36].

Intuitively, one would expect the IXP path to always be the shortest between IXPs' customers. By analyzing the IXP routing table, however, we notice several cases of ASes depreferring the IXP path using path prepending; that is a clear preference to receive traffic from a path other than the IXP. Recent research [37] observed that origin ASes today prepend more than 25% of their prefixes in the global routing table, but seeing this done at IXPs warrants further investigation.

In Table V, we summarize one year's worth of routing tables from IXPs. We found that 26-31% (median) of all prefixes received at the IXPs are prepended—in total, not just at the origin AS. The IXPs show distinct behavior related to prepends. Here we can see three clusters, AMS-IX and LINX, IX.br/RS and IX.br/SP, and SIX.

ixp	any_prep	nei_prep	org_prep	nei_client	nei_org
AMS-IX	26.52	20.52	17.86	16.57	8.52
LINX	28.91	23.43	17.94	19.74	9.21
IX.br/RS	31.00	9.56	14.39	8.79	3.25
SIX	30.19	11.96	22.10	5.22	24.29
IX.br/SP	26.15	8.72	18.43	5.84	6.41

TABLE V: Percentual analysis of IXP's prepended paths.

Column *any_prepend* in Table V represents the percentage of prepended prefixes considering all prefixes on that IXP; *nei_prep* shows how many IXP paths are directly prepended by the IXP customer; *org_prep* is the percentage of paths prepended at the origin by the AS owner. Both are relative to the total announced prefixes at the IXP. When a prefix is prepended in multiple ways (origin and neighbor), we count these separately. *nei_org* shows prepends added by IXP customers to their own prefixes. Here, we consider all neighbor prefixes relative to the total number of prefixes the AS originates. *nei_client* is the percentage of prepended paths added by IXP customers to prefixes of other ASes they announce on the IXP (neighbor clients). We compute this as the fraction of prepends for other ASes over the total of prepended paths. Figure 5 shows the values along the year at LINX (others IXPs were omitted due to lack of space). The IXP customer is directly responsible for adding prepends for up to 23% of prepended prefixes. They normally prepend their client prefixes (downstream), but not their prefixes. The strong variations are typically caused by just one large AS. For example, AS6939 was responsible for more than 50% of all prefixes announced at two IXPs. We sampled a number of long AS paths announced by this AS. They were mostly *low quality paths*. In subsection VII-B, we further analyze the impact of such a big player on an IXP's open peering policy.

Figure 6 compares prepended paths in a global route table and the LINX routing table. While LINX and other IXPs have around 30% paths prepended, the AS3333 global routing table view[1] shows less than 10% prepended paths over 12 years. Unfortunately, one year of IXP data is insufficient to show trends or to reveal if there is a point in time when prepending on IXPs became common practice[2].

[1]Table extracted from RIS [38] collector `rrc01-as3333` (Ripe-AS)

[2]It is not possible to extract IXPs open peering routing table from RIS, Routeviews, or PCH. They do not peer directly with the IXP route-server, or they include private peers, or they log the resulting RIB–normally full routing.

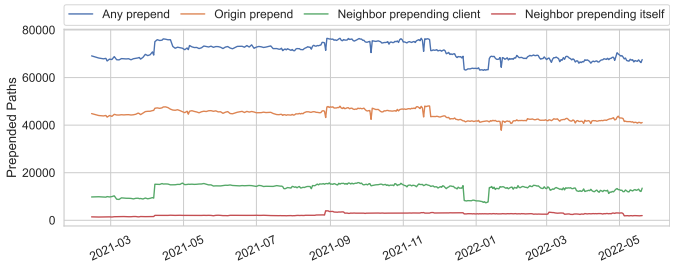
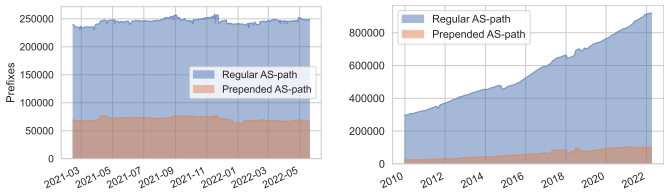


Fig. 5: Who is prepending on LINX



(a) LINX prepended paths. (b) AS3333 view from RIS

Fig. 6: Prepended prefixes at IXP vs. Global prepending

B. *Hurricane Electric's (AS6939) impact on open peering*

Hurricane Electric (HE, AS6939) is a global player who peers openly on several IXPs, announcing a large number of prefixes. HE is one of the ASes we find to exhibit asymmetry. They are not classified as a remote-peer [20].

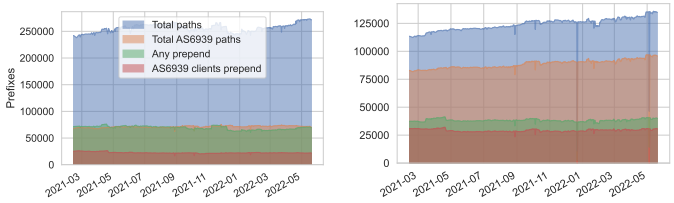
In Figure 7, we show the impact of AS6939's choices at two IXPs. In both cases, they announce about 80,000 prefixes. This number represents less than 30% of all routes on AMS-IX and more than 65% of all prefixes announced on SIX. The announcements include intercontinental routes with high RTT. The accumulated AS paths are long, with a mean of 4, and 25% of them have a length between 5-38. HE never prepends any client path. However, their customers do prepend the path through AS6939 in many cases (look at SIX numbers on Table V). When one AS prepend five times its path, this is a strong indication that is a bad path, normally congested or with high latency (*e.g.,* remote peer).

We also verify that AS6936 makes use of route-server BGP communities to *avoid* announcing its prefix to big CDNs (*e.g.,* Netflix, Akamai, Cloudflare, Edgecast, Google, OVHCloud, Amazon, and several others). This strategy avoids peering with the main CDNs in a region. While this strategy solves the "HE effect" for the big CDNs, small CDNs and eyeball networks need to deal with poor quality routes.

Takeaway: We consistently find IXP customers depreferring IXP routes when comparing with transit paths. Longer paths normally indicate poor quality routes. CDNs without *quality-aware routing* should de-peer with global networks in the open peering model.

VIII. Discussion and Conclusion

In this work, we quantified traffic symmetry on five Internet eXchange Points (IXPs). We defined a methodology based on active anycast measurements to map routing (a)symmetry, yielding almost three times better coverage than RIPE ATLAS.



(a) AMSIX. (b) SIX.

Fig. 7: The impact of AS6939 on IXP routing table.

Our measurements identified traffic symmetry values between 79-88% at the level of prefixes and between 66-86% symmetry at the level of ASes. This symmetry value reflects up to 24% of connected ASes avoid exchanging traffic using the IXP infrastructure. On average 28% of all IXP paths are prepended while we registered 10% in a global routing table view. 15% of IXP prepends are added by the AS connected on the IXP over its clients' prefixes. Up to 8% of ASes filter out IXP routes and never return any traffic to the IXP.

Our experiments show that the main cause of IXP traffic asymmetry is egress-only paths, where an AS announces a prefix on the IXP but does not send traffic back to peers through that IXP. We observed up to 34% of egress-only paths. This means that if one forces traffic to the IXP by using traffic engineering, 34% of those networks will not send traffic back. This is a reason not to increase indiscriminately the preference for routes learned on IXP paths. The origin always decides where to deliver traffic. Trying to change this behavior actually increases asymmetry as well as response times, leading to a decrease in the quality of IXP routes.

We also analyzed inbound traffic symmetry. We repeated our experiment with more specific announcements on the IXP, compared to our transit provider. Our results show that using more specific prefixes only attracts additional traffic from between 4-8% of networks and less than 1% of new ASes globally. Regarding route quality, the use of more specific prefixes increases the mean RTT. In most cases, it is a modest increase, but in others, the RTT increases to 180 ms. When discussing our results with network operators, they commented that they intend to re-evaluate how they connect to IXPs.

Some ASes have a usual behavior: some are deaf, *i.e.*, they ignore prefixes announced on the IXP. Others are mute: they do not announce any prefix to the IXP but use IXP prefixes internally. More than half of the ASes we have contacted identified a configuration mistake. These cases show that alerting ASes about IXP asymmetry can improve IXP route quality.

Some possible solutions to help on addressing the asymmetry problem are: (i) Informational: Our technique to assure IXP neighbors symmetry can be used at scale to help CDNs, AS operators, and IXP administration to identify places of improvement, increasing the IXP route quality. (ii) Business model: IXPs can use local flow data to identify symmetrical paths and provide new multi-lateral views. Also, to add symmetry information on open peering view can make

easier for IXP customers to select routes with better quality. (iii) Standardization: Anycast networks frequently demand a special treatment from routing peers. Designate a special AS-range for anycast networks, or at least label it in the IXPs [39] could make easier for any AS operator to recognize networks using anycast routing.

As there is no ground truth about traffic symmetry in IXPs we made public for the community our datasets, code, and results [26], so we can track trends and compare the results.

Acknowledgements

This work was funded by the Netherlands Organisation for Scientific Research projects PAADDoS (628.001.029), and CATRIN (NWA.1215.18.003), and the European Union's Horizon 2020 - CONCORDIA (830927). It was also partially supported by CNPq 423275/2016-0, 316662/2021-6, and FAPESP 2020/05152-7, 2015/24494-8. We Thanks IXPs' teams and Cesar Augusto Haas Loureiro for their support.

References

[1] M. Bishop, "HTTP/3," Internet Engineering Task Force (IETF) Request for Comments - RFC 9114, pp. 1–57, Jun. 2022.

[2] J. Iyengar and M. Thomson, "QUIC: A UDP-Based Multiplexed and Secure Transport," Internet Engineering Task Force (IETF) Request for Comments - RFC 9000, pp. 1–151, May 2021.

[3] "A First Measurement with BGP Egress Peer Engineering," in *International Conference on Passive and Active Network Measurement (PAM)*, Mar. 2022, pp. 199–215.

[4] W. John, M. Dusi, and K. C. Claffy, "Estimating Routing Symmetry on Single Links by Passive Flow Measurements," in *6th International Wireless Communications and Mobile Computing Conference (IWCMC)*, Jun. 2010, pp. 473–478.

[5] L. Wei, M. Flores, H. Bedi, and J. Heidemann, "Bidirectional Anycast/Unicast Probing (BAUP): Optimizing CDN Anycast," in *4th Network Traffic Measurement and Analysis Conference (TMA)*, Jun. 2020.

[6] X. Zhang, T. Sen, Z. Zhang, T. April, B. Chandrasekaran, D. Choffnes, B. M. Maggs, H. Shen, R. K. Sitaraman, and X. Yang, "AnyOpt: Predicting and Optimizing IP Anycast Performance," in *ACM SIGCOMM 2021*, Aug. 2021, pp. 447–462.

[7] E. Muhati and D. B. Rawat, "ASAP: Anti-Spoofing Aphorism Using Path-analysis," in *IEEE/ACM 29th International Symposium on Quality of Service (IWQOS)*, Jun. 2021, pp. 1–10.

[8] A. S. M. Rizvi, L. Bertholdo, J. Ceron, and J. Heidemann, "Anycast agility: Network playbooks to fight DDoS," in *31st USENIX Security Symposium (USENIX Security 22)*. Boston, MA: USENIX Association, Aug. 2022, pp. 4201–4218.

[9] V. Paxson, "End-to-End Routing Behavior in the Internet," *IEEE/ACM transactions on Networking*, vol. 5, no. 5, pp. 601–615, Oct. 1997.

[10] H. Balakrishnan, V. N. Padmanabhan, and R. H. Katz, "The Effects of Asymmetry on TCP Performance," *Mobile Networks and applications*, vol. 4, no. 3, pp. 219–241, 1999.

[11] D. McPherson, D. Oran, D. Thaler, and E. Osterweil, "TCP Performance Implications of Network Path Asymmetry," Internet Engineering Task Force (IETF) Request for Comments - RFC 3449, pp. 1–41, Jan. 2002.

[12] K. Oztoprak and M. A. Yazici, "A Hybrid Asymmetric Traffic Classifier for Deep Packet Inspection Systems with Route Asymmetry," in *IEEE 35th International Performance Computing and Communications Conference (IPCCC)*, Dec. 2016, pp. 1–8.

[13] A. Pathak, H. Pucha, Y. Zhang, Y. C. Hu, and Z. M. Mao, "A Measurement Study of Internet Delay Asymmetry," in *International Conference on Passive and Active Network Measurement (PAM)*, 2008.

[14] L. M. Bertholdo, J. M. Ceron, L. Z. Granville, and R. van Rijswijk-Deij, "Forecasting the Impact of IXP Outages Using Anycast," in *Network Traffic Measurement and Analysis Conference (TMA)*, Sep. 2021.

[15] A. Ahmed, Z. Shafiq, H. Bedi, and A. Khakpour, "Peering vs. transit: Performance comparison of peering and transit interconnections," in *2017 IEEE 25th International Conference on Network Protocols (ICNP)*. IEEE, 2017, pp. 1–10.

[16] L. M. Bertholdo, "Private comunication with cdn provider," 03 2021.

[17] M. Jansen, "Traffic Engineering for CDNs," https://2015.apricot.net/#sessions/peeringforum3, Mar. 2015.

[18] D. Savage, J. Ng, S. Moore, P. Paluch, and R. White, "Cisco's Enhanced Interior Gateway Routing Protocol (EIGRP)," Internet Engineering Task Force (IETF) Request for Comments - RFC 7868, pp. 1–80, May 2016.

[19] K. Yeung, "Akamai - The Path to 100+ IXPs," https://conference.apnic.net/data/41/ix_100-akamai-apricot2016-23feb2016_1456157526.pdf, Feb. 2016.

[20] G. Nomikos, V. Kotronis, P. Sermpezis, P. Gigis, L. Manassakis, C. Dietzel, S. Konstantaras, X. Dimitropoulos, and V. Giotsas, "O Peer, Where Art Thou? Uncovering Remote Peering Interconnections at IXPs," in *ACM Internet Measurement Conference (IMC)*, Oct. 2018, pp. 265–278.

[21] J. Snijders, J. Heasley, and M. Schmidt, "Use of BGP Large Communities," Internet Engineering Task Force (IETF) Request for Comments - RFC 8195, pp. 1–15, Jun. 2017.

[22] M. McBride, D. Madory, J. Tantsura, R. Raszuk, H. Li, J. Heitz, and G. Mishra, "AS Path Prepending," Internet Engineering Task Force (IETF), Internet-Draft draft-ietf-grow-as-path-prepending-06, Feb. 2022.

[23] L. M. Bertholdo, J. M. Ceron, W. B. d. Vries, R. d. O. Schmidt, L. Z. Granville, R. v. Rijswijk-Deij, and Pras, "Tangled: A Cooperative Anycast Testbed," in *IFIP/IEEE International Symposium on Integrated Network Management (IM)*, May 2021, pp. 766–771.

[24] W. B. de Vries, R. de O. Schmidt, W. Hardaker, J. Heidemann, P.-T. de Boer, and A. Pras, "Verfploeter: Broad and Load-Aware Anycast Mapping," in *ACM Internet Measurement Conference (IMC)*, Nov. 2017.

[25] L. project, "LANDER:internet address history it91w-20200710," https://ant.isi.edu/datasets/readmes/internet_address_census_it91w-20200710.README.txt, Sep. 2020.

[26] L. M. Bertholdo, "IXP Symmetry-rate: Datasets and source code," https://github.com/LMBertholdo/ixp-symmetry-rate, Jun. 2022.

[27] Ripe, "Ripe atlas probes archive," https://ftp.ripe.net/ripe/atlas/probes/archive/2022/05/, Feb. 2022.

[28] G. Nomikos and X. Dimitropoulos, "traIXroute: Detecting IXPs in traceroute paths," in *International Conference on Passive and Active Network Measurement (PAM)*, Mar. 2016, pp. 346–358.

[29] Amsterdam Internet Exchange, "sFlow at AMS-IX," https://www.ams-ix.net/ams/documentation/more, Jun. 2022.

[30] I. Society, "Mutually Agreed Norms for Routing Security (MANRS)." [Online]. Available: https://manrs.org

[31] R. Sommese, L. Bertholdo, G. Akiwate, M. Jonker, R. van Rijswijk-Deij, A. Dainotti, K. C. Claffy, and A. Sperotto, "MAnycast2: Using Anycast to Measure Anycast," in *ACM Internet Measurement Conference (IMC)*, Oct. 2020, p. 456–463.

[32] G. C. M. Moura, J. Heidemann, W. Hardaker, P. Charnsethikul, J. Bulten, J. M. Ceron, and C. Hesselman, "Old but Gold: Prospecting TCP to Engineer and Live Monitor DNS Anycast," in *International Conference on Passive and Active Network Measurement (PAM)*, Mar. 2022.

[33] J. Rexford, J. Wang, Z. Xiao, and Y. Zhang, "BGP Routing Stability of Popular Destinations," in *2nd ACM SIGCOMM Workshop on Internet Measurement (IMW)*, Nov. 2002, p. 197–202.

[34] Y. He, M. Faloutsos, and S. Krishnamurthy, "Quantifying Routing Asymmetry in the Internet at the AS Level," in *IEEE Global Telecommunications Conference (GLOBECOM)*, Nov. 2004, pp. 1474–1479.

[35] Y. He, M. Faloutsos, S. Krishnamurthy, and B. Huffaker, "On Routing Asymmetry in the Internet," in *IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 2, Nov. 2005, pp. 1–6.

[36] J. Ginter, "Higher BGP Local Preference for IX Peering Good Practice or Bad," https://www.bitsinflight.com/higher-local-preference-for/, Sep. 2021.

[37] P. Marcos, L. Prehn, L. Leal, A. Dainotti, A. Feldmann, and M. Barcellos, "AS-Path Prepending: There is No Rose without a Thorn," in *ACM Internet Measurement Conference (IMC)*, Oct. 2020, p. 506–520.

[38] RIPE NCC, "Routing Information Service (RIS)," https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris, 2019.

[39] M. Wilhelm and F. Künzler, "A well-known BGP community to denote prefixes used for Anycast," Internet Engineering Task Force, Internet-Draft draft-wilhelm-grow-anycast-community-01, Jul. 2022.