UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

FABRÍCIO MARTINS MAZZOLA

On the Latency and Routing Impacts of Remote Peering to the Internet

Thesis presented in partial fulfillment
of the requirements for the degree of
Doctor of Computer Science

Advisor: Prof. Dr. Luciano Paschoal Gaspary

Porto Alegre
July 2023

CIP — CATALOGING-IN-PUBLICATION

Mazzola, Fabrício Martins

On the Latency and Routing Impacts of Remote Peering to the Internet / Fabrício Martins Mazzola. – Porto Alegre: PPGC da UFRGS, 2023.

106 f.: il.

Thesis (Ph.D.) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR–RS, 2023. Advisor: Luciano Paschoal Gaspary.

1. Internet. 2. Interconnection. 3. Peering. 4. Remote Peering. 5. Internet eXchange Point. I. Paschoal Gaspary, Luciano. II. Título.

"It is more complicated than you think."

— RFC 1925: THE TWELVE NETWORKING TRUTHS

ACKNOWLEDGEMENTS

The thesis is the end of a long journey and represents, directly and indirectly, the work and support of wonderful people. First, I am incredibly thankful to my advisor, Marinho Barcellos, for his guidance over the last eight years. Since day one as a research assistant in my undergraduate degree, Marinho showed me the path to accomplish genuine and impactful research, always looking for significant and real problems and challenges. Thank you for constantly pushing me forward and believing in my potential. I would also like to thank him for the countless hours providing support, patience, constant feedback, and advice. I will always be grateful for the opportunity, our time spent together, and your temporary adoption during my stay in New Zealand. Thank you so much for everything!

I also want to thank my second advisor in this journey, Luciano Gaspary. He has been present in my academic life since my bachelor's degree, sharing all his knowledge about research, critical thinking and computer networks. His support and understanding were crucial, especially in the final stages of my PhD. None of this work would be possible without your contributions.

I am very thankful to Pedro Marcos, which co-advised this thesis. Pedro has always been around collaborating with me since my bachelor's period. His partnership and cooperation at the early stages of my PhD were fundamental in helping set the right direction for our research and making it happen successfully during the whole PhD period. My development as a researcher and the contributions of this work are directly related to his assistance.

I am also beyond grateful to Ignacio Castro (Queen Mary University of London) and Matthew Luckie (University of Waikato, New Zealand), which contributed tremendous help to the development of this thesis. I feel very fortunate for the chance I had to work with them during my PhD. It helped me greatly to become a better researcher and develop a rigorous scientific approach to Internet measurements. My deepest thanks!

I am forever thankful and would like to send my warmest recognition to my family. This journey could not be done without the endless support of my girlfriend Aline, that was there with me every single minute during this experience. Nothing of this would be made without her constant help, love, strength, patience and encouragement. I am also very grateful to my parents Simone and João, and my sister Jessica, for always being there for me as a rock, supporting, cheering, and guaranteeing that I have had everything possible to reach the end of this marathon. Special mention to my cats Tito and Figo, who were adopted in the beginning of this PhD and made this period lighter, funnier, fluffier and most importantly, they were always by my side.

Special thanks go to my close collaborators and people who helped me and encouraged me throughout my research life: Lucas Muller, Miguel C. Neves, Rodrigo Oliveira, Daniel Marcon, Rodrigo Dal Ri, Sergio Gutierrez, Matheus Lehmann, Lucas Leal, and Augusto Cetti. I am grateful to each one of them! I am very thankful to Julio Sirota (IX.br), Antonio M. Moreiras (NIC.br) and Leonardo Almeira (Globenet) for their excellent support, feedback, and discussions that allowed this work to be possible.

Finally, I would like to thank UFRGS and the University of Waikato for all the support during my PhD. I also would like to acknowledge ACM for the travel support to attend conferences. This research was partly supported by CAPES/Brazil - Finance Code 001 and LACNIC FRIDA.

I specially dedicate this thesis to my father João Mário Mazzola, who always dreamt of pursuing a PhD and could not be here to see me achieve my and his dreams.

ABSTRACT

Remote peering (RP) has crucially altered the Internet topology and its economics. Increasingly popular thanks to its lower costs and simplicity, RP has shifted the member base of Internet eXchange Points (IXPs) from strictly local to include ASes located anywhere in the world. While the popularity of RP is well understood, its implications on Internet routing and performance are not. In this thesis, we perform a comprehensive measurement study of RP in the wild, based on a representative set of IXPs (including some of the largest ones in the world, covering the five continents). We first identify the challenges of inferring remote peering and the limitations of the existing methodologies. Next, we perform active measurements to identify the deployment of remote IXP interfaces and announced prefixes in these IXPs, including a longitudinal analysis to observe RP growth over one and a half years. We use the RP inferences on IXPs to investigate whether RP routes announced at IXPs tend to be preferred over local ones and what are their latency and latency variability impacts when using different interconnection methods (remote peering, local peering, and transit) to deliver traffic. Next, we asses the RP latency impact when using a remote connection to international IXPs and reaching prefix destinations announced by their members. We perform measurements leveraging the infrastructure of a large Latin American RP reseller and compare the latency to reach IXP prefixes via RP and four Transit providers. Finally, we glimpse some of the RP implications on Internet routing. We evaluate how RP can considerably affect IXP members' connection stability, potentially introduce routing detours caused by prefix announcement mispractices and be the target of traffic engineering by ASes using BGP communities.

Keywords: Internet. Interconnection. Peering. Remote Peering. Internet eXchange Point.

LIST OF ABBREVIATIONS AND ACRONYMS

API	*Application Program Interface*
AMS-IX	*Amsterdam Internet eXchange*
AS	*Autonomous System*
ASN	*Autonomous System Number*
BGP	*Border Gateway Protocol*
CDN	*Content Delivery Network*
DB	*Data Base*
DE-CIX	*Deutscher Commercial Internet eXchange*
eBGP	*External Border Gateway Protocol*
EGP	*Exterior Gateway Protocol*
Euro-IX	*European Internet Exchange Association*
HE	Hurricane Electric
iBGP	*Internal Border Gateway Protocol*
IGP	*Interior Gateway Protocol*
IP	*Internet Protocol*
ISP	*Internet Service Provider*
IXP	*Internet Exchange Point*
IX.br	*Brazilian Internet eXchange*
LACNIC	*Latin America and Caribbean Network Information Centre*
LG	*Looking Glass*
LINX	*London Internet eXchange*
L2	*Layer-2*
L3	*Layer-3*
MED	*Multi-Exit Discriminator*

MLPE	*Multilateral Peering Exchanges*
NAPAfrica	*NAPAfrica Internet eXchange*
PCH	Packet Clearing House
PDB	PeeringDB
PNI	*Private Network Interconnection*
RIB	*Routing Information Base*
RP	*Remote Peering*
RS	*Route Server*
RTT	*Round Trip Time*
TTL	*Time To Live*
Tbps	*Terabits per second*
VP	*Vantage Points*
VM	*Virtual Machine*

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

Evolution of Internet Structure and Applications. The Internet structure has deeply changed over the last two decades. In the traditional hierarchy, a small set of big Transit providers (a.k.a Tier-1 networks) dominated the Internet structure by providing global connectivity to other networks (BöTTGER et al., 2018; LABOVITZ et al., 2010). In this context, networks were required purchase transit services to obtain access to the global Internet by paying fees proportional to the traffic volume sent/received through the Transit provider infrastructure (NORTON, 2012c).

However, the recent transformations in application characteristics and requirements have led to changes in how Autonomous Systems (ASes) interconnect. The rise of Internet video (e.g., Netflix, YouTube), gaming and social networking (e.g., Meta, Tiktok), along with the tendency of services using shared cloud infrastructure, made a majority of Internet traffic to be generated and concentrated on the hands of a small set of networks (ARNOLD et al., 2020; CHIU et al., 2015). These three types of applications combined represent more than 76% of all the Internet traffic in 2022 (Sandvine, 2023), and only 5 ASes were responsible for originating more than half of Internet traffic in 2019 (TREVISAN et al., 2020; PUJOL et al., 2019).

Internet flattening and rise of peering infrastructures. Since many current applications require strict service requirements (e.g., lower latency) and generate large traffic volumes, ASes must constantly increase their interconnection capacities to stay profitable and provide increased benefits to their customers. One way to improve interconnection is by shortening the distance between their networks and the ASes generating the demanded content, such as Cloud and Content Providers (CP). To achieve this goal, ASes started to directly interconnect through peering relationships with each other. With the increased network interconnectivity, ASes can reduce dependence on Transit providers (i.e. Tier-1 and Tier-2 ISPs) and bypass them through direct peering connections (ARNOLD et al., 2020; BöTTGER et al., 2018).

As many networks established direct connections over Transit providers, the Internet became less hierarchized and "flatter". Peering infrastructures, such as Internet eXchange Points (IXPs) and colocation facilities, became crucial elements of the Internet topology in this process. They concentrated a large number of networks and allowed ASes to directly interconnect in a centralized infrastructure. Compared to Transit providers, they can offer shortened Internet paths, improved performance, and reduced interconnection and operational cost to its members (Internet Society, 2021; CHATZIS et al., 2013b; AGER et al., 2012; AUGUSTIN; KRISHNAMURTHY; WILLINGER, 2009; Cloudflare, 2016; Dr Peering, 2012). As of February 2023, there were more than 800 IXPs deployed worldwide (PeeringDB, 2023; Euro-IX, 2023; Hurricane Electric, 2021), with the largest ones surpassing 1000 members (LINX, 2021c; LINX, 2021a; IX.br, 2023a) and 10 Tbps of peak traffic (CGI.br, 2021; DE-CIX, 2023; LINX, 2021c; AMS-IX, 2023a).

A growing and unexplored scenario. An original motivation of IXPs was to keep local traffic local by having ASes physically present (deploying a router) at an IXP facility. However, IXPs no longer only interconnect members physically present at IXP facilities. *Remote Peering* (RP) – where an AS is not physically present at an IXP facility and reaches the IXP through a layer-2 provider – allows ASes to widen their peering footprint with a quicker setup, no additional hardware, and lower installation costs compared to local peering (DE-CIX, 2021a; AMS-IX, 2021b; CASTRO et al., 2014). For example, ASes from 80 countries connect at LINX remotely (LINX, 2021c) as of February 2023. To cope with the demand for peering, IXPs and remote peering resellers have expanded their offerings (BICS, 2014; Telecomdrive Bureau, 2020; France-IX, 2017) with some IXPs having up to 55 official partners selling remote peering services (IX.br, 2023a; AMS-IX, 2018; LINX, 2021d).

According to the state-of-the-art (GIOTSAS et al., 2021; CASTRO et al., 2014), RP is not limited to a theoretical idea anymore, becoming a significantly common practice on the Internet. For the most relevant IXPs in the world, approximately 40% of their member base is connected via RP (GIOTSAS et al., 2021). However, despite the popularity of this connection approach, little is known about its implications for the Internet and the ability to interconnect with remote members at IXPs adds complexity to traffic engineering choices. Recently, there has been a broad public debate about remote peering performance. Among the concerns are that L2 services can introduce further operational complexity, lead to routing inefficiencies, hinder monitoring and make it harder to evaluate their latency impact on the interconnections (LEVY, 2019; ALMEIDA, 2019; NIPPER et al., 2018; PANEL..., 2016b; PANEL..., 2016a; ALI, 2012; NORTON, 2012a). The very few research studies looking at the RP impacts to the Internet found a link between RP and negative impacts on anycast performance and detecting of peering infrastructure outages (BIAN et al., 2019; GIOTSAS et al., 2017).

Research questions. Investigating the widely unexplored scenario of RP implications to Internet performance, specifically looking at its latency impacts, and Internet routing poses several research questions:

- Are the state-of-the-art methods enough to adequately infer RP on IXPs worldwide? (§5.1)
- Can networks experience latency benefits or penalties when delivering traffic to remotely announced prefixes? If so, how much better/worse is it compared to local peering connections and traditional Transit providers? (§6)
- Considering the high popularity of RP on different IXPs, using RP must provide benefits to the ASes using this approach. Is the obtained advantage associated with latency performance? If yes, how many destinations can benefit from it, and how much better is it compared to Transit providers? (§7)
- Can RP introduce other effects to the interconnection apart from latency performance, such as differences in IXP members' connection stability and a higher prospect of introducing routing detours? (§8)

Openly disclosing which members' networks are connected locally or via RP is not a common practice at most IXPs. The lack of transparency makes it difficult for ASes to know which of its peers are remote and how they can affect the performance of their interconnections. As of today, the RP performance debate is data-poor, and there is no well-defined study which evaluates whether remote connections can positively or negatively impact the interconnection's latency performance. Effectively addressing these questions will enable a series of improvements in the interconnection ecosystem and will highlight the different advantages/drawbacks of RP.

Contributions. In this thesis, we perform an extensive measurement investigation to improve the understanding of deployment and impacts of RP on the Internet. We believe our results are valuable to the interconnection ecosystem and the community as it contributes data to the open performance discussion.

We summarize our contributions as follows.

1. We provide an analysis of the methodological challenges and their implications on inferring RP at IXPs using the state-of-the-art proposal, extended with a comprehensive analysis of previous work (§5.1).

2. We perform active measurements in eight IXPs (which include six of the world's ten largest IXPs by membership) to identify and understand the deployment of remote IXP interfaces and announced prefixes, including a longitudinal analysis to observe RP growth over one and a half years. Partnerships with some of the IXPs and interactions with network operators support our findings with ground truth data and operational insights (§5.3, §5.4, §5.5).

3. We use the RP inferences on IXPs to investigate whether RP routes announced at IXPs tend to be preferred over local ones. Afterwards, we perform an extensive measurement study to understand the latency and latency variability impacts when using different interconnection methods (remote peering, local peering, and transit) to deliver traffic to prefixes announced by remotely connected members (§6).

4. We asses the RP latency impact when using a remote connection to international IXPs and reaching prefix destinations announced by their members. We partnered with a large Latin American RP reseller to leverage their infrastructure and simulate an AS connecting at two IXPs via RP. We then compare the latency to reach IXP prefixes via RP and four Transit providers (§7).

5. We glimpse some of the RP effects on Internet routing. We evaluate how RP can considerably affect IXP members' connection stability, potentially introducing routing detours caused by prefix announcement mispractices and the usage of BGP communities to perform traffic engineering on remotely connected networks at IXPs (§8).

Limitations. Investigating the benefits or drawbacks of a network connection can be a broad and complex problem. While some network operators consider financial advantages as the best metric to evaluate a connection type (e.g., peering, RP, Transit), others mainly rely on connection performance and stability. In terms of performance, the choice of metric can also vary across applications. Some applications, such as video streaming, depend on high network bandwidths to operate correctly. Others, including online gaming, Voice over IP (VoIP) and stock trading, depend highly on low network latencies. This thesis investigated he RP performance impacts compared to other connection types, primarily based on latency metrics. Analyzing performance and routing impacts by other metrics (e.g., bandwidth, economics) is very challenging, considering the lack of reliable information in publicly available datasets, limited access to Vantage Points with full measurements capability, and intrinsic restrictions on the execution of measurement campaigns in-the-wild to avoid damaging impact to the public Internet.

We organize the remainder of this thesis as follows. In §2 we present the fundamentals of the interconnection ecosystem. In §3 we discuss the related work to infer RP on IXPs, investigate the RP implications to the Internet and usage of BGP communities. In §4 we present the measurement architecture and datasets used in our work. §5.1 discusses the current challenges of inferring RP on IXPs, while on §5 we detail the deployment of remote interfaces and announced prefixes on eight IXPs worldwide. In §6 we investigate the latency impacts introduced by RP when looking through an IXP Perspective and in in §7 we evaluate RP impact through an customer AS Perspective. Finally, in §8 we highlight some of the RP implications to Internet routing and §9 describe our findings and discuss the possibilities to continue this research.

2 INTERNET INTERCONNECTION

This chapter presents the Internet interconnection ecosystem foundations. We first describe the basics of the underlying protocol used to interconnect networks via interdomain routing, followed by details on the different interconnection agreement models, the infrastructures used for peering and the distinct peering methods. Readers familiar with these concepts can skip this chapter and continue to Chapter 3.

2.1 Interdomain routing

ASes must have a common language to exchange reachability information to achieve interdomain routing. The Border Gateway Protocol (BGP) is the standard protocol used on the Internet for this purpose. It allows ASes to announce routes to their prefixes, has an algorithm to determine the best paths to a prefix, and provides features for ASes to perform traffic engineering following their desired policies. We further describe these aspects below.

2.1.1 Border Gateway Protocol (BGP)

BGP is the de facto protocol used on the Internet to perform interdomain routing between ASes. It allows networks to exchange reachability data with other BGP systems. It has a protocol specific to configure routing between distinct ASes (*external* BGP - eBGP) and one to configure routing between the routers within one AS (*internal* BGP - iBGP) (REKHTER; HARES; LI, 2006).

When establishing a connection between two BGP routers (or edge routers), the protocol starts a BGP session and exchanges *UPDATE* messages between them. ASes use these messages to announce to their neighbours when a new best route to reach a prefix. Each BGP route has a series of attributes used by networks to reach a prefix. We highlight the most relevant ones below.

- *ORIGIN*: defines the origin or how the route was learned. It can be via the IGP, EGP or *INCOMPLETE*.

- *AS-PATH*: contains the sequence of traversed ASes to reach the advertised prefix. Every time an AS forwards a learned route, it adds its ASN on the *path*;

- *NEXT-HOP*: identifies the router's IP address that packets should use as the next-hop to reach the destination IP;
- *MULTI-EXIT-DISCRIMINATOR* (MED): represents an optional value that indicates a path is preferred when an AS has multiple entry points;
- *LOCAL-PREF*: determines the preferred path to a prefix according to the AS local policy. The default value is 100.

2.1.2 BGP Best Path Selection

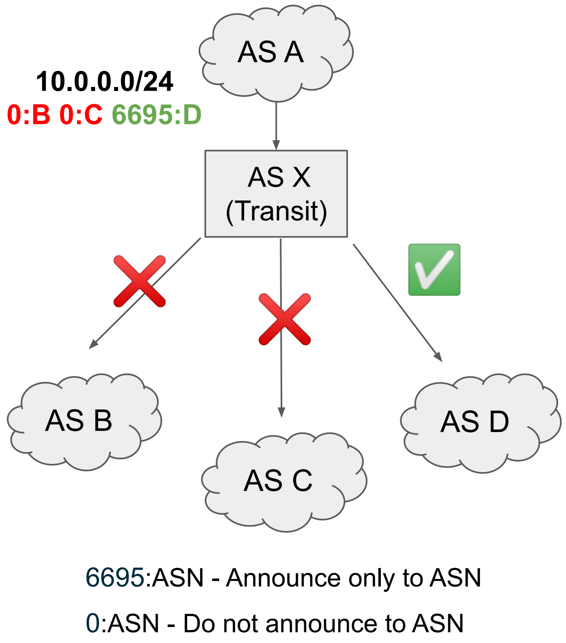
After receiving multiple routes to the same prefix, the router must select a single best path to store in its *Routing Information Base* (RIB). The best path is selected by the BGP protocol, following multiple rules which vary according to different vendors implementation (Cisco, 2023; Juniper, 2023). After the best path is chosen, it is stored and propagated to all the router's neighbours. Below, we highlight the criteria used in most best path selection algorithms, in decreasing order of priority;

1. Highest *LOCAL-PREF*.
2. Routes originated locally by other networks.
3. Paths with the shortest *AS-PATH* (i.e., fewer ASNs in the path).
4. *ORIGIN* field in decreasing priority order: IGP > EGP > *INCOMPLETE*.
5. Lowest Multi-Exit Discriminator (MED). MED is exchanged between ASes and informs other networks which path should be used to enter the AS's networks.

2.1.3 BGP Communities

BGP communities (CHANDRA; TRAINA; LI, 1996) is a variable-length message attribute that can be included in BGP updates to convey some routing information or request. The standard version of these communities is represented by 32 *bits*, written in two groups of 16 *bits* (16:16). To comply with the deployment of 32 bits ASNs, large BGP communities were proposed in 2016, using 96 *bits*, written in three groups of 32 *bits* (32:32:32).

Figure 2.1: Example of the usage of BGP communities.

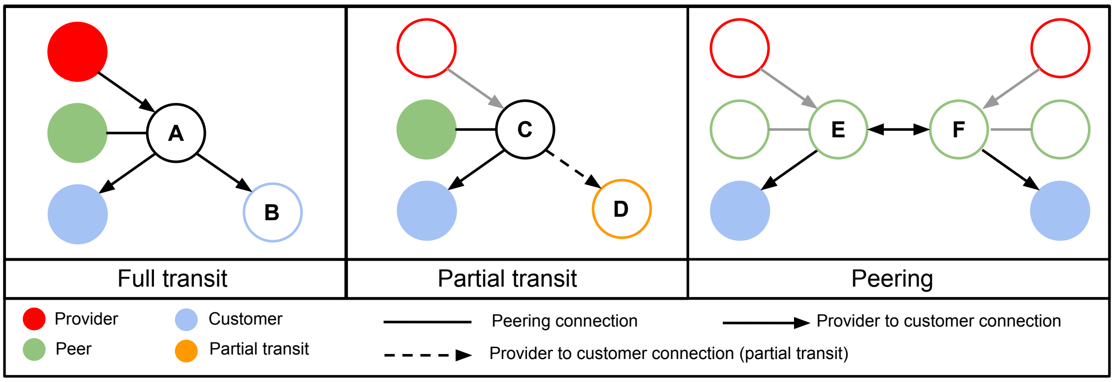


Source: The authors

ASes can tag routes with communities to either ask other ASes to perform some action (i.e., action BGP communities) regarding the route or to add some information about a given route characteristic (i.e., informational BGP communities) which can help improve routing decisions of the AS or its neighbours. Action communities can request that the AS receiving the route treat it in a particular way, such as performing path-prepending, blackholing, or changing the route's propagation characteristics. On the other hand, informational ones act as tags for route features, including the country or continent where an AS received the route and if an AS originated the route or received it externally from other networks.

Figure 2.1 shows an example of how BGP communities work. In this case, ASes A, B, C and D connect to Transit X. Network X allow its clients to use the BGP communities to avoid or specify the route announcement to other ASe. AS A announces prefix 10.0.0.0/24 to X, but it does not want X to forward its route to ASes B and C, only to AS D. To perform this traffic engineering, AS A tags its route with the communities *0:B*, *0:C* and *6695:D*. As AS X receives the route from AS A, it will process the tagged communities and then apply the respective actions.

Figure 2.2: Full Transit, Partial Transit and Peering.



Source: Adapted from (GIOTSAS et al., 2014)

2.2 Interconnection Paradigms

When ASes decide to interconnect, they must choose which interconnection paradigm best reflects their needs in terms of interconnection cost, exchanged traffic volume, and peering and security policies (GIOTSAS et al., 2015). Currently, *transit* and *peering* are the most known interconnection methods, distinguished by aspects such as monetary expenses and Internet reachability (NORTON, 2012c; NORTON, 2012b). These models are non-exclusive, meaning ASes can have both paradigms simultaneously with different networks in multiple locations (GIOTSAS et al., 2014; MüLLER et al., 2019). Figure 2.2 illustrates the different paradigms. We describe each model, along with its characteristics and differences.

2.2.1 Transit

A transit connection paradigm bases itself on ASes already connected to the Internet selling access to the global Internet to other networks. Internet Service Providers (ISPs) are the primary entities selling and operating these Internet Transit services. ASes can have different transit agreements with multiple providers at once to improve connectivity and resilience. According to the literature, there are two main types of transit agreements: *full transit* and *partial transit* (FARATIN et al., 2008; GIOTSAS et al., 2014).

On the former, an AS buys connectivity to the rest of the Internet from ISPs. In this case, the ISP will announce to the Internet how to reach its customer prefixes and will send and receive the traffic between the customers and the Internet via its infrastructure. ISP will charge its customers typically on the Internet transit volume exchanged at peak-hour traffic (e.g., 95th Percentile Measurement Method) (NORTON, 2012c). For the latter, the ISP will only sell reachability to a part of the Internet. Usually, this represents the prefixes from the ISP's peers and customers. Since the service is limited, its price is considerably lower than a full transit (FARATIN et al., 2008).

2.2.2 Peering

The peering paradigm allows networks to interconnect with other ASes and provide reachability to each other's customers prefixes. It has gained great popularity in the last decades since it provides economic and performance benefits over relying on transit providers (NORTON, 2012b). Peering agreements can be either *settlement-free* or *paid peering*. For both models, the interconnection between the networks allow them to exchange traffic originated/destined from/to their networks or each other's customer cone (LUCKIE et al., 2013; GIOTSAS et al., 2013). Settlement-free peering does not involve financial compensation between the involved ASes, regardless of the traffic volume exchanged. If the peering benefits are not symetrical between both networks, ASes can agree on a paid peering model. In this case, ASes arrange payment either over traffic over the agreed ratio or traffic volume (FARATIN et al., 2008).

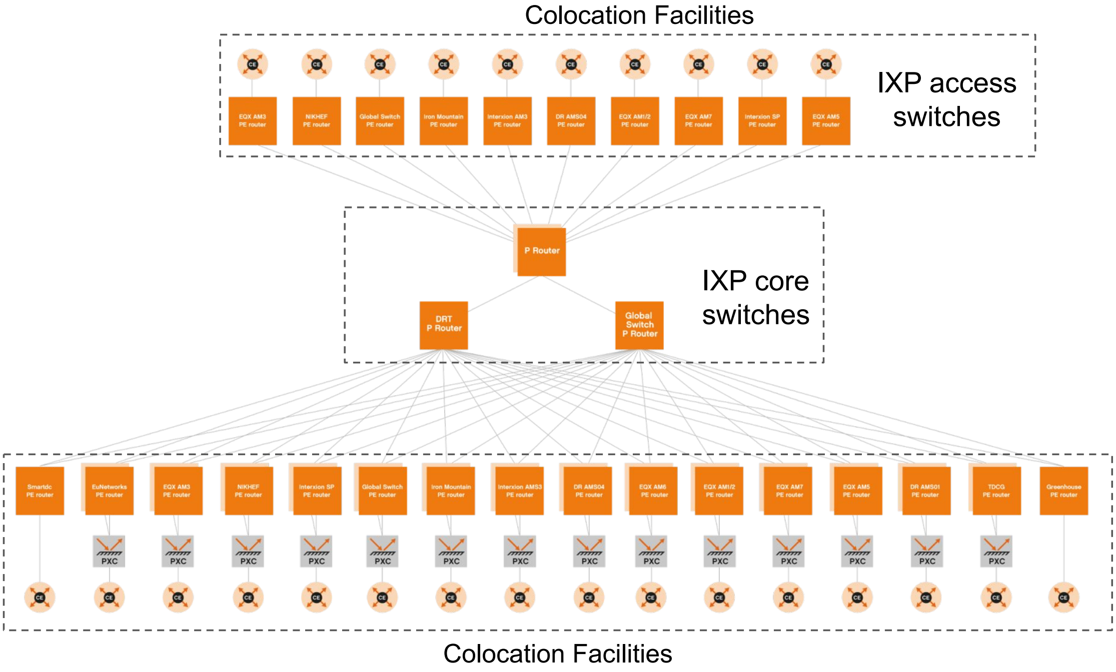
2.3 Where and How to Peer?

After deciding to interconnect via peering, an AS must connect to a peering infrastructure. The connection to IXPs or colocation facilities can be local or remote, depending on the network's primary goal when peering. Next, the AS can decide whether to connect with other networks inside the infrastructure via public or private peering. We describe them below.

2.3.1 Peering Infrastructures

IXPs and colocation facilities represent the primary examples of modern peering infrastructures. They are layer-2 switching fabrics where ASes interconnect to exchange traffic with many other networks also connected in that facility (AGER et al., 2012). We describe them below.

Figure 2.3: AMS-IX Platform Layer 1. Adapted from (AMS-IX, 2023a)



Source: The authors

Internet Exchange Points. IXPs are physical network infrastructures where multiple ASes can interconnect their networks to exchange traffic. Figure 2.3 shows the typical architecture of AMS-IX, one of the largest IXPs in the world. IXPs provide a shared switching fabric where participating networks can interconnect their routers. The switch fabric carries the traffic resulting from public and private peering of all interconnected ASes. Each IXP has one or more *core switches* in the shared fabric for redundancy. They also associate with several Colocation Facilities and install *access switches* to reach city-level interconnection with other networks (GIOTSAS et al., 2015).

Historically, IXPs are considered the successors of Network Access Points (NAPs), responsible for the smooth transition from the monolithic government network to the modern Internet (CHATZIS et al., 2013b). Since 1995, the four existing NAPs have been replaced by more than 850 IXPs in 200+ cities worldwide, interconnecting 50k+ networks (PeeringDB, 2023; Euro-IX, 2023; GIOTSAS et al., 2015). Despite initially deployed mainly in Europe and USA (CHATZIS et al., 2013b; CHATZIS et al., 2015), these peering infrastructures are the leading forces in the Internet development in Latin America and Africa (BRITO et al., 2016; FANOU; VALERA; DHAMDHERE, 2017). In Latin America, the number of IXPs increased by 58.3% from 2016 to 2021, expanding from 60 to 96 IXPs (Rosas, Israel, 2021). Well-established IXPs (e.g., IX.br-SP, DE-CIX Frankfurt) exchange, on average, over 10 Tbps of traffic per day (IX.br, 2023b; DE-CIX, 2023), amounts of data similar to Tier 1 Internet Service Provider (ISPs). The largest IXP in terms of members (IX.br-SP) has over 2450 connected members in 2023 (IX.br, 2023a).

IXPs have been central infrastructures for over ten years in the modern Internet. A study performed in 2012 reported that a single European IXP observed traffic from an extensive share of the Internet, including 42K+ routed ASes, almost all 450K+ routed prefixes and around a quarter billion IP addresses from all around the globe (CHATZIS et al., 2013a). Just five IXPs are sufficient for an AS to reach about 40% of the IP prefixes advertised on the Internet (KOTRONIS et al., 2015). Until 2014, IXPs presented an annual growth of 10-20% on ASes connecting to IXPs and of 50-100% per year in traffic rates (RICHTER et al., 2014).

Colocations Facilities. Colocation facilities (Colos) are physical locations which provide essential infrastructures like power, space, cooling, physical security, and storage to their associated ASes. More specifically, it is a place where operators of multiple networks place their networking equipment for interconnection (INTERCONNECTION..., 2014). These facilities lower the infrastructure costs and help small and medium providers to house their equipment (storage, server, routers) in a centralized and safe location (KOTRONIS et al., 2017; MILOLIDAKIS; FONTUGNE; DIMITROPOULOS, 2019).

The Colos fabric connects the member's network to various IXPs, transit providers, cloud/content providers and ASes also connected in these infrastructures. In large metropolitan areas, a colocation facility operator may install interconnected facilities in the same city to allow access from ASes present at one facility to networks at another facility in the same region (GIOTSAS et al., 2015). Large carrier-neutral companies such as Equinix (EQUINIX..., 2023) and Telehouse (TELEHOUSE..., 2023) are the leading operators of colocation facilities worldwide.

2.3.2 Local vs Remote Peering

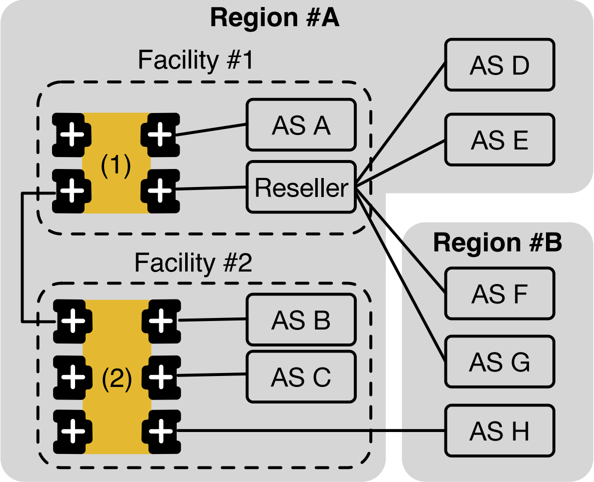
ASes can reach and connect to a peering infrastructure through two alternatives: *local* or *remote*. The first mode is usually used by ASes geographically close to the metro area of a peering infrastructure. Local peering is characterized by ASes physically connecting their access router to a colocation facility or the IXP switching fabric (AGER et al., 2012). RP, on its turn, is when ASes want to connect on an IXP without incurring additional operational and hardware expenses. Studies show that up to 40% of the member base of modern IXPs are using remote peering (CASTRO et al., 2014; NOMIKOS et al., 2018). There are many ways for an AS to connect remotely at one peering infrastructure. We describe them in further detail in the next section (§2.4)

2.3.3 Public vs Private Peering

After connecting to IXPs and colocation facilities, ASes can establish interconnections with other networks via *public* or *private* peering. Public peering is when interconnections use the public infrastructure of the facility. By peering via the IXP's switch fabric, ASes can exchange traffic with a large fraction of IXP members using the same switch port (NORTON, 2012b). One method to connect two ASes is to establish a direct BGP session between their access routers. The *bilateral* BGP peering session allows them to trade routing information between each other's routers and exchange traffic over this peering link at the IXP (GIOTSAS et al., 2015; AGER et al., 2012). Despite providing higher control and security over the ASes' routing policies, creating multiple bilateral sessions yield higher operational complexities due to the crescent number of peering links over peering infrastructures (GIOTSAS et al., 2013). To simplify the interconnection process, IXPs started offering their members the option of establishing multilateral peering to a route server. The benefit of multilateral peering is that an AS only needs to establish one BGP session with the IXP route server to receive all the other IXP members' routes (RICHTER et al., 2014; GIOTSAS et al., 2013).

On the other hand, private peering (a.k.a. private network interconnection - PNI) offers two ASes to physically connect their interfaces at one interconnection facility. They are usually used in cases where networks exchange asymmetric traffic volumes. By using PNIs, ASes can obtain higher interconnection bandwidth capacity, availability, security and monitoring guarantees. Given the performance benefits, PNI is the preferred method used by hypergiants networks (e.g., Google, Netflix) to connect with ISPs providing Internet services to the end-users (i.e., eyeball networks) (WOHLFART et al., 2018; RAPAPORT et al., 2021). Despite the advantages, PNIs have higher operational complexities as operators have to manually set up the environment at the switching fabric of the peering infrastructure.

Figure 2.4: ASes connect to IXPs via local (AS A, AS B, AS C) and RP connections, either via resellers (RES) ports (AS D, AS E, AS F, AS G) or through their own port and buying transport from the remote location to the IXP switch (AS H). Note that RPs can be physically located near the IXP (AS D, AS E) or can also be geographically distant (AS F, AS G, AS H).



Source: The authors

2.4 RP comes in different flavor.

An AS can connect remotely to an IXP through different modes: (i) remote peering resellers, (ii) IP/layer-2 transport, and (iii) interconnected IXPs. Each of these methods presents different characteristics related to the underlying third-party infrastructure being used to reach the IXP remotely. We discuss the three distinct cases associated with RP and use Figure 2.4 to illustrate them. The figure contains one IXP, eight ASes (A-H), one Reseller (RES) and two colocation facilities (1) and (2).

Remote peering resellers. The typical approach to connect from a remote location is via resellers. Many modern IXPs offer official solutions to foster the use of RP on their peering infrastructures, involving agreements with a few selected RP resellers (AMS-IX, 2021a; LINX, 2021e; DE-CIX, 2021a; Netnod, 2021). These services provide access to the IXP, usually connecting the routers of the remote ASes to the IXP switches via layer-2 transport, along with the contractual IXP membership to its customers.

Resellers services can serve two purposes. The first is to help the bridging of large geographical distances by connecting AS located far from the IXP (AS F, AS G). The second one is to facilitate and reduce the costs of peering equipment and its installation, allowing ASes located close to the IXP (AS D, AS E) to easily connect at IXPs.

In Figure 2.4, Reseller RES allows customer ASes D and E to reach other members at the IXP. The client networks can choose to join the IXP using either shared or dedicated resellers ports. In both cases, ASes appear as regular IXP members and have the same benefits as a direct connection to the IXP. In the shared approach, ASes can take capacities smaller than the minimum speed port offered by the IXP (e.g., 100Mbps instead of 1 Gbps port), which can reduce peering costs. Each network using the shared port is assigned with a VLAN, which provides logical isolation to their communications.

IP/layer-2 transport. The second method of connecting remotely consists of obtaining IP or L2-layer transport from the AS's router to an IXP facility. In this case, instead of relying on a reseller, the AS buys the peering ports directly with the IXP and buys transport to it from its remote location. For example, AS H, in Figure 2.4, reaches the IXP by buying transport to CF #2. The remoteness of networks using this approach is undetectable even to the IXPs, as they have no information about the exact location of the member AS's router.

Interconnected IXPs. The last approach to remotely join IXPs is through IXP federations (DE-CIX, 2021b; LINX, 2021b; France-IX, 2021; LU-CIX, 2021; AMS-IX, 2021a). Such IXP Federations can be of two kinds: IXPs in the same organization or Partners IXPs.

The first type represents the peering infrastructures that facilitate the connection to other IXPs belonging to the group. The DE-CIX GlobePEER Remote solution (DE-CIX, 2021b), for example, allows members connected at their IXP in Frankfurt to peer remotely at nine other IXPs across Europe and the USA.

Partners IXPs, on their turn, consist of agreements between IXPs from different organizations that provide interconnections between their infrastructures. Usually, these are connections made via one of their existing IXP reseller partners, who offer IXP members the possibility to exchange traffic with networks of different exchange points without additional cost. If Reseller RES had a connection to other IXPs of the federation besides the one presented in Figure 2.4, ASes D and E could use their already existing connection to RES to interconnect at them. LINX, for example, allows its members to connect to NaMeX (in Rome) (NaMeX, 2021) and JPIX (in Tokyo) (JPIX, 2021) through its IXP Partner Program (LINX, 2021b).

3 RELATED WORK

With the growing deployment of remote peering, there have been several efforts to investigate this interconnection practice. We divide related work into three categories: (1) methods to identify remote peering at IXPs (§3.1.2), (2) studies to explore implications of remote peering on the Internet (§3.1.1), and (3) work investigating the characteristics and usage patterns of BGP communities at IXPs (§3.2).

3.1 Remote Peering

3.1.1 Inferring Remote Peering.

Two main related methodologies have been proposed in the literature. In 2014, Castro *et al.* (CASTRO et al., 2014) introduced a conservative inference method based on measuring propagation delay to IXP interfaces connected to it via pings. Responses to ping probes sent to IXP interfaces that presented latency more than 10ms and whose IP-TTL had not been decremented were classified as remote. The authors reported that 91% of the 22 studied IXPs showed networks connecting via remote peering. Further, using ground-truth traffic from a National Research and Education Network, the paper demonstrated that a network could offload up to 25% of its transit-provider traffic via remote peering.

In 2018, Nomikos *et al.* (NOMIKOS et al., 2018) also proposed a methodology to infer remote peering. Using ground-truth data from seven IXPs, the authors showed that latency alone was not sufficient to make accurate inferences in some cases, such as IXPs with switching fabrics distributed across different countries. The paper proposed combining latency measurements with additional remote peering features, such as port capacity and AS presence at colocation facilities, to obtain a more trustworthy inference methodology. Their method computes the geographical area where an IXP member's router could be located and associates the router with the feasible facilities that a local peering could use. They used this method to infer RP in 30 IXPs worldwide, and reported that 90% of the analyzed IXPs had more than 10% of their members using remote peering, with two of the largest IXPs in terms of members (DE-CIX and AMS-IX) having up to 40% of remote members. In 2021, the authors extended the previous work (GIOTSAS et al., 2021), with changes in the methodology and additional analysis on Wide-Area IXPs.

3.1.2 Implications of Remote Peering.

In 2017, Giotsas *et al.* (GIOTSAS et al., 2017) proposed a methodology for detecting peering infrastructure outages, such as colocation facilities and IXPs. The authors reported that the rise of remote peering made it easier for localized failures in IXP and colocation facilities to become widespread. For two outages observed in London (2016), they showed that more than 45% of the interfaces related to the affected links were from outside England, with more than 20% of them being located outside Europe.

In 2019, Bian *et al.* (BIAN et al., 2019) proposed a methodology to characterize anycast based on archived BGP routing information collected globally. While trying to infer anycast prefixes, the authors found that remote peering caused a significant element of inaccuracy in their method. They reported that RP can cause unintended consequences on anycast performance and potentially affect 19.2% of the anycast prefixes. Active measurements found that 38% of such prefixes were indeed impacted with an average latency increase of 35.1ms.

The work of Bertholdo et. al (BERTHOLDO et al., 2021) analyzed the stability of participants' connections to IXP route-servers. Their results show that unstable interfaces were mainly caused by large regional ASes connected in just one IXP or ASes connected via Remote Peering at IXPs.

3.2 Usage of BGP Communities.

The original BGP communities standard (CHANDRA; TRAINA; LI, 1996) defined values for only three communities, essentially providing a way to limit route propagation. In 2008, Donnet and Bonaventure (DONNET; BONAVENTURE, 2008) proposed a taxonomy, with three classes: inbound, outbound and blackholing. The inbound referred to tagging announcements with information (e.g. where it was learned), while the outbound referred to communities used for traffic engineering, by influencing route propagation. The third class, blackholing (BH), allowed ASes to drop traffic towards some prefix (as a DDoS defence strategy) (KING et al., 2016; DIETZEL; WICHTLHUBER, 2018). BGP communities can also be grouped in *informational* and *action* communities.

Previous work can be roughly divided into studies on the use/se-mantics of communities, and their use for measurement studies. In the first group, Dietzel et al. investigated the usage of blackholing in IXPs (DIETZEL; FELDMANN; KING, 2016), while

Giotsas et al. measured its adoption in the wild (GIOTSAS et al., 2017), and Nawrocki et al. assessed its efficacy against DDoS (NAWROCKI et al., 2019). While communities are useful for AS' operations, they can also lead to problems. Earlier studies examined its use as a vector of routing attacks (STREIBELT et al., 2018; BIRGE-LEE et al., 2019), and how communities can cause overheads (KRENC; BEVERLY; SMARAGDAKIS, 2020). The (lack of) semantics for community values motivated methodologies for semantics inference (SILVA et al., 2022) and best-effort attempts to build community directories (Step, 2021). BGP communities have also been used for inference studies. Examples include finding p2p links at IXP RSes (GIOTSAS et al., 2013), studying RSes in IXPs (RICHTER et al., 2014), inferring complex AS relationships (GIOTSAS et al., 2014), mapping peering interconnections to a facility (GIOTSAS et al., 2015), and detecting outages (GIOTSAS et al., 2017).

Comparing the work in this thesis to prior art, we highlight three studies. In (GIOTSAS et al., 2013), Giotsas et al. collected communities from router server with semantics defined by the IXP in order to infer p2p links. In (RICHTER et al., 2014), Philipp et al. examined the role of route servers in IXPs, using communities for some inferences. Krenc et al. (KRENC; BEVERLY; SMARAGDAKIS, 2021) observed announcements at BGP collectors (e.g. RIPE and RouteViews) aiming to understand better community usage, but limited to when/how ASes *add* communities to announcements and when they *remove*. Mazzola et al (MAZZOLA; MARCOS; BARCELLOS, 2022), on the other hand, evaluated how action communities used for traffic engineering are used by ASes in IXPs and performed a characterization about BGP communities usage patterns.

4 MEASUREMENT ARCHITECTURE

In this section, we explain the measurement architecture used to infer remote interfaces and prefixes for our intended analysis in three parts: we justify the selection of IXPs considered (§4.1), the control plane datasets (with BGP information) we collected for the study (§4.2), how we used VPs to perform necessary dataplane measurements (§4.3). The described measuremente architecture is used on the analysis performed in the two following chapters (§5 and §6)

4.1 Peering Infrastructure Selection

To identify networks connected via remote peering, and prefixes and routes announced via remote peering, we are restricted to peering infrastructures that have (1) publicly available BGP routing data, and (2) an active measurement VP attached to the IXP switching fabric. Table 4.1 presents the eight selected IXPs where we had both BGP routing data and active measurement capability. These IXPs include six of the world's ten largest IXPs by membership (Euro-IX, 2023; Hurricane Electric, 2021) and are deployed in five different countries. The three Brazilian IXPs (i.e., PTT sites) are part of the largest ecosystem of public IXPs in the world (IX.br) and are the leading Latin American IXPs in terms of average traffic volumes (≈12.9, 9.2, and 1.4 Tbps, respectively) (IX. . . , 2020; CARISIMO et al., 2020; BRITO et al., 2016). The eight IXPs together comprise 3466 unique ASes.

4.2 Datasets

Remote Peering Reseller Ground Truth Data. We obtained ground truth information for the ASes remotely connected via resellers for four of the analyzed IXPs: LINX, PTT-SP, PTT-RJ, and PTT-CE. The data set contains information about the ASN and IP interface of remote ASes reaching the IXPs through shared ports or VLANs associated with resellers. For the PTT IXPs, we obtained the ground truth data from their operators on the 20th April 2021. The set of ASes reaching LINX through resellers or locally connected to the IXP is publicly available at their member portal (LINX, 2021d) (collected on 5th May 2021). LINX representatives confirmed that ASes with Port Type labeled as ConneXions

Table 4.1: The eight IXPs analyzed in our study, along with the availability of BGP VPs and ground truth data on remote peering.

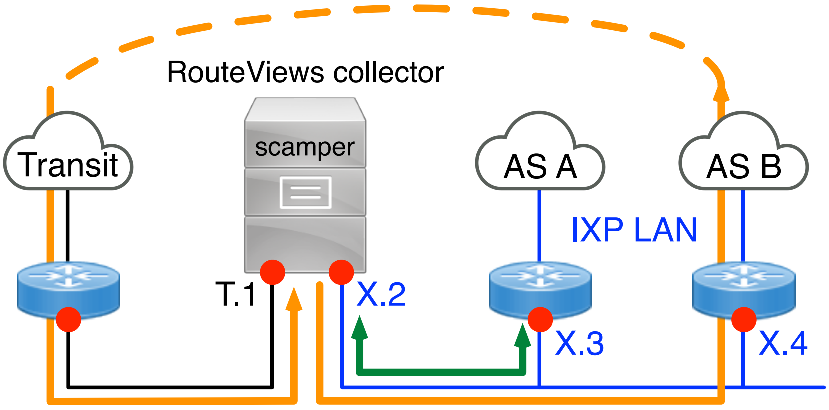
IXP	Location	Observed Interfaces	BGP VPs LG	BGP VPs PCH	Reseller Ground Truth
PTT-SP	Sao Paulo, BR	2,169	✓	✗	✓
LINX	London, UK	911	✓	✓	✓
AMS-IX	Amsterdam, NL	907	✓	✓	✗
NAPAfrica	Johannesburg, ZA	542	✗	✓	✗
PTT-RJ	Rio de Janeiro, BR	462	✓	✗	✓
PTT-CE	Fortaleza, BR	395	✓	✗	✓
Eq-Ash	Ashburn, VA, US	365	✗	✓	✗
Eq-Chi	Chicago, IL, US	259	✗	✓	✗

correspond to ASes using resellers. The ground truth for the four IXPs comprise a list of 1634 unique ASes using remote peering through resellers.

Membership and Interface Addresses. To identify the peering router's IP and ASN of all members at each IXP, we combine multiple public data sources for all IXPs except for LINX, which publishes this information through their member portal (LINX, 2021d). We collected membership data and subnet information from Euro-IX (Euro-IX, 2023) and the publicly available databases of Hurricane Electric (HE) (Hurricane Electric, 2021), PeeringDB (PDB) (PeeringDB, 2023), and Packet Clearing House (PCH) IXP Directory (PCH, 2020). In cases of conflicts, we followed the preference ordering described in (NOMIKOS et al., 2018): *Euro-IX* > *HE* > *PDB* > *PCH*.

BGP Datasets and Sanitization. We used two sources of routing data: (i) Looking Glass (LG) of the IXP which observes routes from the IXP's Route Server and (ii) routes from the archive collected by PCH (PCH, 2021). For IXPs with both PCH and LG views, we used data archived by PCH because it has greater visibility of routes advertised by IXP members. For example, when comparing both datasets for AMS-IX and LINX, we observed 3.4–3.9x more routes and 1.9–2.0x more prefixes from PCH than from LG views. For our analysis of RP latency impact on the IXP perspective (§6), we prefer the dataset from PCH whenever it is available, as it provides us with better visibility of the IXP routes (*PCH* > *LG*). On IXPs with only LG views (PTT sites), we have observed that, in 2021, LGs were configured to output only the best routes, lowering the number of cases with multiple routes for the same prefix at that time. We have observed that this was no longer happening in 2022. Additionally, we collected BGP data from RouteViews collectors at each IXP to understand the types of routes that RouteViews peers actually chose. We discarded: (i) routes with artifacts, such as reserved/unassigned ASes (IANA, 2020) and

Figure 4.1: Architecture of our data plane measurements. We used RouteViews collectors with an interface connected to a transit provider and an interface in the IXP LAN as VPs for data plane measurements. Delay measurements to the peering router of each IXP member (e.g., X.3) used the collector's IP address in the IXP LAN (X.2), so the probes and responses crossed the IXP LAN. Other measurements used the Transit IP address T.1 as the source address, and were delivered to each IXP member using the layer-2 address corresponding to their IXP LAN IP address (e.g., X.4).



loops; (ii) prefixes shorter than /8 or longer than /24.

4.3 Data plane measurements

Vantage Points. At each IXP listed in Table 4.1, we used RouteViews collectors which were directly connected to the IXP LAN to conduct active measurements using scamper (LUCKIE, 2010). Figure 4.1 illustrates the measurement architecture of each RouteViews collector and how we used them to conduct active measurements.

Measurement Types. We conducted two types of measurements. In the first, we measured the latency to each IXP member's peering router. These measurements use the IP address that the collector has in the IXP LAN (X.2), so that probes and responses cross the IXP LAN, as in when we probe X.3 in Figure 4.1. In the second, we measured the path and latency to IP addresses within prefixes announced by each IXP member. Note that these prefixes are peering routes, and not transit routes. These measurements go out via a selected IXP member (e.g. AS B, using the layer-2 address of X.4 in Figure 4.1) but used the collector's Transit IP address T.1 as the source address, so that we could receive a response. This strategy allowed us to maintain the same return path from the probed address back to the RouteViews collector, while varying the forward path as we selected different IXP members. We provide further details about the measurement methodology in the sections describing our results (§5,§6.1, and §6.2).

5 REMOTE PEERING DEPLOYMENT AT IXPS

This chapter first discusses the main challenges identified in inferring RP using state-of-the-art methodologies (§5.1). Next, we describe the method used in this thesis to identify RP at IXPs (§5.2) and present numbers on both the remotely connected members (§5.3) and remotely announced prefixes and routes (§5.4) obtained in 2021. Finally, we present a longitudinal analysis of the deployment of RP over three months in 2022 and compare the results with 2021 data (§5.5).

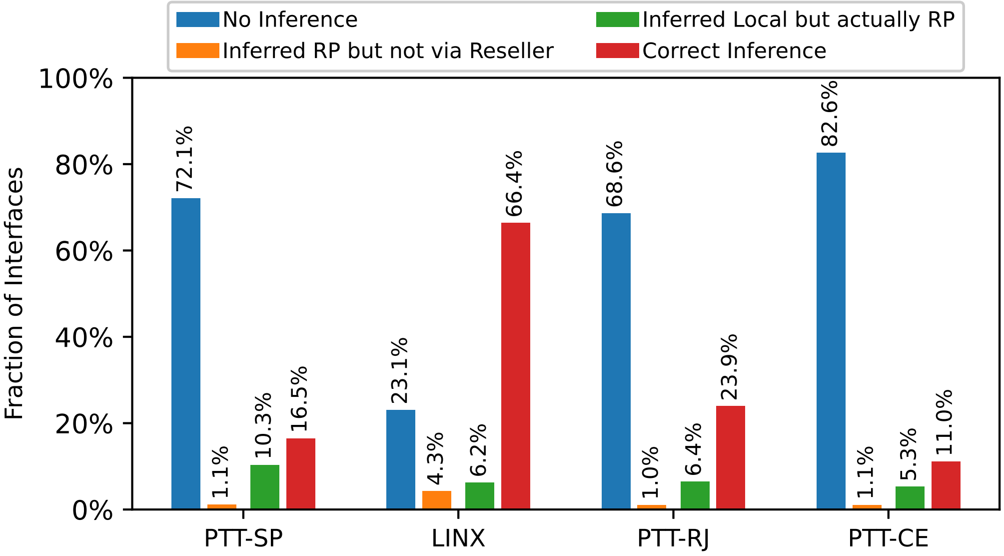
5.1 Challenges in Inferring Remote Peering

Available Data Limits Accuracy of Remote Peering Inferences. The current state-of-the-art methodology for inferring remote peering proposed by Giotsas *et al.* (GIOTSAS et al., 2021) infers remote peering (1) through a reseller and/or (2) geographically distant from the IXP. The method combines delay measurements with additional features, such as port capacity and AS presence at colocation facilities; if an AS is not present in one of the feasible IXP facilities, their method infers the AS is remotely connected. We used available ground truth (§4.2) for four IXPs (LINX, PTT-SP, PTT-RJ, and PTT-CE) and applied their method to all interfaces connected to these IXPs.

We implemented the four steps from the Giotsas *et al.* (GIOTSAS et al., 2021) method. The first step (*ping measurement campaign*) measures the latency to IXP member interfaces from a VP within the IXP. Using the scamper probers on the RouteViews collectors (§4.3), we performed delay measurements to the peering interfaces of IXP members every two hours for two days, and discarded measurements where the replies might have come from outside the peering infrastructure because they had an IP-TTL value that appeared to have been decremented (i.e., the received IP-TTL was not 64 or 255). The second step (*colocation-informed RTT interpretation*) computes a geographical area where the IXP member router could be located using an AS to colocation facility mapping obtained from PeeringDB and IXP websites. Then, we obtained publicly available RIPE Atlas IPv4 traceroute measurements collected on the same days as our ping campaign and applied step 3 (*multi-IXP router inference*) and step 4 (*finding remote peers via port capacities and lack of private connectivity*) to complete the methodology.

Figure 5.1 presents the results we obtained. In (GIOTSAS et al., 2021), public

Figure 5.1: Classification of interfaces we obtained when we applied our implementation of the current state-of-the-art methodology for inferring remote peering (GIOTSAS et al., 2021). The high percentage of no inference for the three Brazilian IXPs was a consequence of the method's high reliance on public information (PeeringDB) which was not widely available for members of Brazilian IXPs.



information about AS presence at colocation facilities was missing for ≈25% of remote peers and ≈18% of local peers. When we reproduced the study, the number of unknown inferences for LINX was low and the fraction of remote and local interfaces inferred was similar with the published work (GIOTSAS et al., 2021). The case for Brazilian IXPs was different. For PTT-SP and PTT-CE, only 27.0% of the members had PeeringDB entries that reported both the IXP and facilities where they were present, leading the current state-of-the-art method to only classify 17.1%, on average, of the interfaces at the Brazilian IXPs. This low classification was because only a few ASes connected to the Brazilian IXPs shared their information in PeeringDB. Openly publishing peering data has only recently been encouraged by IXP operators in Brazil as best practice (MENDES, 2020).

In addition, 5.3-10.3% of the interfaces inferred as local peerings were actually remote, according to ground truth. We believe the misclassification was related to incorrect information about the presence of ASes in colocation facilities. In many cases, an AS using a reseller recorded the facility their reseller connected to in their PeeringDB record, leading the method (GIOTSAS et al., 2021) to infer the AS was locally connected. The other 1.0-4.3% of interfaces inferred as remote were correct, but they did *not* observably connect to the IXP via a reseller. In summary, the methodology of (GIOTSAS et al., 2021) may not be suitable for accurately inferring remote peering for IXPs that have incomplete or inaccurate publicly available data.

5.2 Inferring RP on IXPs

Identifying remote peering (RP) based solely on reseller connections is imprecise, as it ignores geographically distant ASes not using reseller ports which also incur a latency penalty. However, examining only remote peers that are geographically distant overlooks RP through resellers. This diversity in the notion of RP led us to evaluate RP both by (1) connection type (*Reseller RP*), and (2) geographical distance to the IXP (*Geographical RP*).

To identify members using Reseller RP, we used ground truth that identified members connected to an IXP using a reseller for four IXPs (§4.2). To infer members using Geographical RP at all eight IXPs, we used the method in (CASTRO et al., 2014), which uses latency measurements and empirically obtained thresholds as a proxy of physical distance, with the following approach. For each IXP, we associated IXP member ASes and their assigned IXP IP addresses using the datasets mentioned in §4.2. We performed latency measurements to these addresses on 5-6 May 2021. From each RouteViews scamper instance, we probed each interface every two hours for two days, and used the minimum latency for each address to account for cases of transient congestion. To ensure that the ping replies returned directly over the peering infrastructure, we discarded measurements where the replies had an IP-TTL value that appeared to have been decremented (i.e., not 64 or 255). If the minimum latency from a given interface was 10ms or higher, we classified the member's router as remotely connected to the IXP; a latency of 10ms would roughly correspond to a distance of up to 1000km from the IXP (KATZ-BASSETT et al., 2006; TRAMMELL; KÜHLEWIND, 2018). We adopted (CASTRO et al., 2014)'s method because its latency threshold alone yielded accurate results for single metropolitan area peering infrastructures (GIOTSAS et al., 2021), which is the case of the analyzed IXPs in our work (see §4.1).

To further assess the correctness of our inferences – and similar to step 2 in (GIOTSAS et al., 2021) (*colocation-informed RTT interpretation*) – we obtained the colocation facilities of each of the eight analyzed IXPs in public data sources (IXP websites and PeeringDB) and computed the distance between them. We observed that Equinix Ashburn has the largest distance between facilities (i.e., 80km), which corresponds to a latency of

Table 5.1: Number and percentage of routes and prefixes announced by members using a shared port via resellers. Members connecting to an IXP via a reseller announced fewer routes than members connecting locally. LINX had a considerable percentage (78.7%) of the same prefixes being announced by both remote and local peers.

	Reseller Remote Peering			
IXP	Interfaces (I)	Routes (R)	Prefixes (P)	P also Local
PTT-SP	1,265 of 2,169 (58.3%)	28,385 of 154,509 (18.4%)	27,148 of 158,880 (17.1%)	577 of 27,148 (2.1%)
LINX	189 of 911 (20.7%)	107,533 of 1,018,593 (10.6%)	90,633 of 486,171 (18.6%)	71,357 of 90,633 (78.7%)
PTT-RJ	172 of 462 (37.2%)	5,525 of 128,961 (4.3%)	5,502 of 128,478 (4.3%)	25 of 5,502 (0.5%)
PTT-CE	214 of 395 (54.2%)	7,098 of 26,025 (27.3%)	7,095 of 26,012 (27.3%)	10 of 7,095 (0.1%)

≈1ms. Therefore, any IXP peer interface with latency consistently higher than 10ms is unlikely to be a local peer at the IXPs we examined.

5.3 Remotely Connected Members

Tables 5.1 and 5.2 summarize the number and percentage of interfaces connected via remote peering at each IXP.

Reseller RP. We observed a large percentage of Reseller RP at the three Brazilian IXPs, representing more than 37.2% of their member base (Table 5.1). According to network operators at these IXPs, the IXPs' members are spread across Brazil, which has a large land mass, and members connect to the IXP to reach large content and cloud providers. We encountered a substantially smaller fraction of Reseller RP at LINX (20.7%).

Geographical RP. We inferred that at least a quarter of the ASes connected to PTT-CE, AMS-IX, and PTT-SP were Geographical RP (Table 5.2). The remaining IXPs had less than 13.3% Geographical RP members inferred. This indicates that even though remote peering is widely used at IXPs (as shown by (NOMIKOS et al., 2018; GIOTSAS et al., 2021)), the majority of the member ASes are physically connected to the IXPs or closely located to them.

Table 5.2: Number and percentage of routes and prefixes announced by inferred geographically remote members. Members we infer to connect to an IXP from some geographical distance announced fewer routes than members connecting locally. LINX, AMS-IX, Eq-Ash, and Eq-Chi all had a considerable percentage (71.4%) of the same prefixes announced by both remote and local peers.

	Geographical Remote Peering			
IXP	Interfaces (I)	Routes (R)	Prefixes (P)	P also Local
PTT-SP	681 of 2,169 (31.4%)	20,289 of 158,932 (12.8%)	19,612 of 154,561 (12.7%)	1,118 of 19,612 (5.7%)
LINX	121 of 911 (13.3%)	92,975 of 1,015,040 (9.2%)	71,452 of 482,643 (14.8%)	65,060 of 71,452 (91.1%)
AMS-IX	238 of 907 (26.2%)	67,397 of 978,225 (6.9%)	63,323 of 485,933 (13.0%)	56,503 of 63,323 (89.2%)
NAPAfrica	40 of 542 (7.4%)	7,256 of 159,100 (4.6%)	7,252 of 144,513 (5.0%)	88 of 7,252 (1.2%)
PTT-RJ	61 of 462 (13.2%)	3,861 of 129,135 (3.0%)	3,850 of 128,652 (3.0%)	355 of 3,850 (9.2%)
PTT-CE	139 of 395 (35.2%)	6,870 of 26,610 (25.8%)	6,869 of 26,597 (25.8%)	8 of 6,869 (0.1%)
Eq-Ash	35 of 365 (9.6%)	49,157 of 967,133 (5.1%)	46,752 of 525,688 (8.9%)	43,455 of 46,752 (92.9%)
Eq-Chi	17 of 259 (6.6%)	8,382 of 347,788 (2.4%)	8,120 of 271,855 (3.0%)	5,795 of 8,120 (71.4%)

5.4 Remotely Announced Prefixes and Routes

For each IXP, we examined the proportion of BGP routes in the IXP routing data, and the percentage of prefixes that could be reached via both local and remote peers (i.e. local and remote routes). To identify whether routes were local or remote, we compared routes observed in the BGP data with inferred remote networks. We labeled routes as remote when the next-hop IP interface belonged to the IXP subnet and belonged to the list of networks we classified as remote.

We show the percentage of remote interfaces, routes, and prefixes we inferred at each IXP, along with absolute numbers, in Table 5.1 for Reseller RP and in Table 5.2 for Geographical RP. In all IXPs, remote peers announced proportionally fewer routes than local peers, both for Reseller RP (Table 5.1) and Geographical RP (Table 5.2). For example, in PTT-SP and PTT-RJ, the fraction of peers using Reseller RP was 3.2x and 8.7x higher than the fraction of routes they announced, respectively. For LINX, the 189 remote peers (20.7% of all interfaces) announced just 10.6% of the routes (107k/1M). For the Geographical RP inferences, PTT-RJ shows the highest difference between the fraction of remote interfaces and remote routes (4.4x), with 61 (13.2%) remote interfaces announcing just 3.0% of all routes (67k/981k). The results suggest that remotely connected ASes tend to announce fewer prefixes than local networks into the IXP. Conversations with IXP network operators revealed that remote peers mainly use their connections to obtain specific content not available at their local IXPs.

Interestingly, we observed a sizeable percentage of prefixes announced by both remote and local peers in some IXPs. At LINX, AMS-IX, Eq-Ash, and Eq-Chi, at least 71.4% of remotely announced prefixes also had a route announced by a local peer in May 2021. These cases can be a problem for traffic engineering since remote peering is invisible to Layer-3 protocols, and there is no guarantee that BGP will choose the lowest latency route.

5.5 Longitudinal Analysis

According to (GIOTSAS et al., 2021), as of mid-2018, the deployment of new remote peering connections had been a significant factor in the recent IXP growth. To further investigate these findings, we pose two questions: (i) are RP connections still a major contributor to the growth of IXPs? (ii) How does the prevalence of remote peering changes over time in IXPs?

Methodology. We answer these questions in three steps, as follows. To infer RP at IXPs during three months in 2022, we continuously collect IXP membership data, announced routes at IXPs, and latency measurements from a VP inside the IXP to each member interface (§4.3). Then, we compare our results regarding remotely inferred interfaces (§5.5.1) and prefixes (§5.5.2) to the ones previously described above and obtained in 2021.

5.5.1 How does remote membership vary with time?

First, we investigate how the prevalence of remote peering (in terms of member interfaces) changes with time, as shown in Table 5.3. We select two days from our 2022 data collection (one in the first month and one in the third) and compare them with the numbers obtained in May 2021. We selected the two days with the highest numbers of prefixes and connected interfaces at IXPs for each month, to increase the visibility on the

Table 5.3: Number of interfaces connected via RP at IXPs and the percentage they represent on the total IXP members. The prevalence of RP grew slightly in most IXPs (6/7) between 05-2021 and 10-2022 (matched by an increase in the absolute number of remote interfaces). Eq-Ash was the only IXP with a decrease in remote interfaces, both absolutely and in prevalence.

	Remote Interfaces		
	05-05-2021	16-08-2022	14-10-2022 (17-10-2022 for PTT-RJ)
PTT-SP	681 of 2,169 (31.4%)	720 of 2,156 (33.4%)	728 of 2175 (33.5%)
LINX	121 of 911 (13.3%)	129 of 878 (14.7%)	130 of 888 (14.6%)
NAPAfrica	40 of 542 (7.4%)	52 of 612 (8.5%)	49 of 619 (7.9%)
PTT-RJ	61 of 462 (13.2%)	67 of 427 (15.7%)	67 of 444 (15.1%)
PTT-CE	139 of 395 (35.2%)	169 of 423 (39.9%)	161 of 421 (38.2%)
Eq-Ash	35 of 365 (9.6%)	31 of 375 (8.3%)	30 of 378 (7.9%)
Eq-Chi	17 of 259 (6.6%)	18 of 265 (6.8%)	19 of 265 (7.2%)

IXP routing data.

We found that remote member interfaces' deployment has grown since 2021 in almost all IXPs (6/7) and that the amount of (aggregated) growth varied according to the characteristics of the IXP. The numbers indicate that well-established IXPs, such as LINX and PTT-SP, have grown less, which conforms intuition: these massive IXPs already have a widespread member base and less space to grow in local and remote networks compared to emergent infrastructures, such as NAPAfrica and PTT-CE. In three months, the growth ranged between 6.9% (PPT-SP) and 22.5% (NAPAfrica). The odd case, Equinix Ashburn, actually had a decrease in the number of remote member interfaces (-14.29%), showing that the RP growth cannot be simply assumed.

We contrast our findings with earlier work (GIOTSAS et al., 2021), whose authors looked at the RP evolution at five IXPs between 2017 and 2018. They concluded that remote member interfaces on the five IXPs grew 20% over one year, but this analysis was for all IXPs combined *aggregated*. We look in more depth and find that this growth is not equally distributed, being influenced by basic properties of the IXP (e.g., size, traffic, location).

We can also observe (from Table 5.3) that the fraction of remote interfaces remained relatively stable, with changes under 3% (in PTT-CE, it went from 35.2% to 38.2%). This is so because the number of local interfaces fell in some IXPs *despite* the growth in RP. For PTT-SP and PTT-RJ, the number of local interfaces has decreased -2.76% and -5.99%, respectively. In virtually all IXPs (6/7), the number of remote interfaces increased, reaching up to 10× when compared than local interfaces. NAPAfrica, the younger/smaller IXP in our set, had a subtle difference (1.6×), hinting that for emergent IXPs the accelerated grow may happen similarly for both remote and local interfaces.

5.5.2 How do remotely announced prefixes vary with time?

To answer the question, first we combine the information about remote interfaces with IXP BGP routing data to identify the prefixes announced by these members. To classify a *prefix* as local or remote, we tag it according to the `nextHop` interface of the BGP route, which refers to the IXP member announcing it at the IXP. A prefix that has routes being announced by *both* remote *and* local members is denoted as *hybrid*.

Table 5.4 shows the results obtained for the three-month data collection compared with data from 2021. For most IXPs (5/7), the number of remote prefixes decreased from

Table 5.4: Number and percentages of remote and prefixes at IXPs. Remote prefixes decreased in four IXPs even though they showed a growth in remote members over time.

Remote Prefixes			
	05-05-2021	16-08-2022	14-10-2022 (17-10-2022 - PTT-RJ)
PTT-SP	19,612 of 154,561 (12.7%)	13,764 of 122,237 (11.3%)	12,453 of 130,707 (9.5%)
LINX	71,452 of 482,643 (14.8%)	67,794 of 472,746 (14.3%)	70,163 of 463,036 (15.2%)
NAPAfrica	7,252 of 144,513 (5.0%)	8,301 of 174,449 (4.8%)	8,435 of 175,797 (4.8%)
PTT-RJ	3,850 of 128,652 (3.0%)	2,654 of 98,176 (2.7%)	1,640 of 93,589 (1.7%)
PTT-CE	6,869 of 26,597 (25.8%)	5,356 of 28,017 (19.1%)	5,367 of 27,137 (19.8%)
Eq-Ash	46,752 of 525,688 (8.9%)	21,698 of 600,632 (3.6%)	21,679 of 597,586 (3.6%)
Eq-Chi	8,120 of 271,855 (3.0%)	8,686 of 344,233 (2.5%)	8,745 of 344,312 (2.5%)

2021 to 2022. The reasons for this, according to our private talks with network operators, is that networks use RP more often to fetch/download content (faster) from remote locations (via more established IXPs) than to deliver/upload content to distant networks. We did not observe a change pattern in terms of remote prefix prevalence, ranging broadly between -57.4% (2,210 fewer) and 16.3% (1,183 more). One might expect that an increase in the number of remote members would lead to an increase in remote prefixes, but we found no correlation: the four IXPs with a reduction in remote prefixes *actually increased* the number of remote members in the same period (as shown in §5.5.1). In nearly half of the IXPs (4/7), we saw a shift in the prevalence of prefixes, from remote to local ones.

5.6 Summary

Our results show that inferring RP in the wild using state-of-the-art methodologies is still challenging since they do not adequately embrace the different characteristics of IXPs worldwide and the ASes' adoption of publicly sharing their interconnection data. Considering this, we use a more conservative method to identify remote connections on IXPs and observe that RP deployment is indeed prevalent nowadays, accounting for a substantial fraction of the IXP membership. To observe whether the expansion of RP was consistent, we used a three-month data collection from 2022 and compared it with our findings from 2021 to find that remote interfaces have grown over the one-and-a-half-year period in almost all IXPs evaluated. We have observed that the growth was directly related to how developed and prevalent the peering infrastructure was. While the growth was lower in well-established IXPs, RP development was predominant in more emergent peering infrastructures. The financial benefits of RP, along with its lower bureaucracy and full advantages of regular peering, make it an appealing alternative for ASes to obtain direct access to the entire IXP routes and offload traffic that would be sent via a Transit provider's infrastructure.

6 THE IMPACT OF RP TO LATENCY: THE IXP PERSPECTIVE

When an AS is at an IXP, it does not have the information about which other members connect via local or remote connections. The invisibility of RP to layer-3 protocols makes it harder for networks to understand the latency implications of sending traffic to ASes connecting via RP. Besides, it makes it complex for them to evaluate whether delivering traffic via remote or local peering connections at the IXP or a Transit provider yields the best latency performance.

This chapter first describes the results of active measurements to remotely announced prefixes comparing the latency to reach the same addresses using routes from remote peers and local peers (§6.1). Next, we perform measurements to remotely announced prefixes that only have the remote routes at the IXP and examine the latency differences between using remote peers and transit providers to reach them (§6.2).

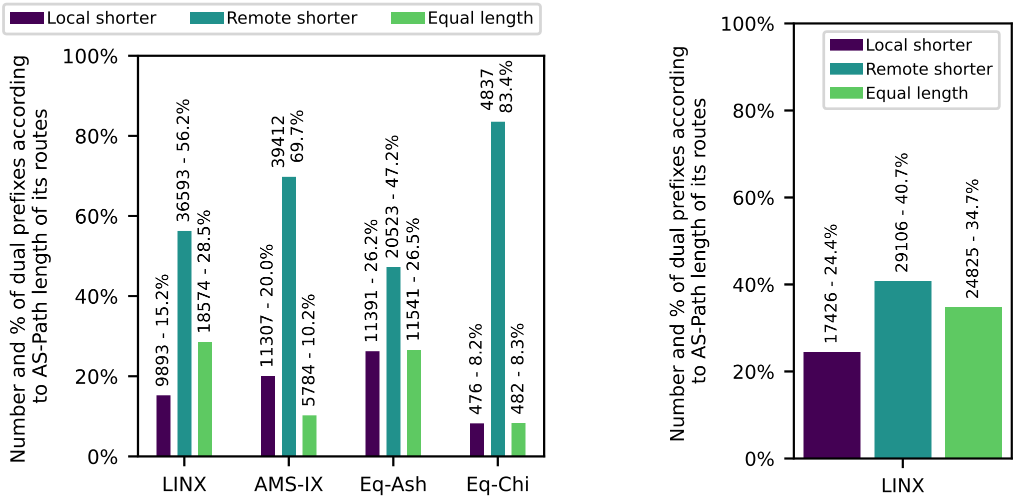
6.1 Choosing Between Remote and Local Peering

Sending traffic via an IXP rather than a transit provider can potentially offer lower latency by keeping local traffic local. However, it is currently unknown whether remote peering might hinder that benefit. The geographical distance of an AS or its connection type can introduce undesired latency implications to peering. In this section, we first investigate whether remote routes have shorter AS paths than local routes (§6.1.1). Next, we analyze routing data from RouteViews collectors at each IXP and find that remote routes are chosen by BGP in the majority of cases (§6.1.2). Then, we measure latency, and compare the latency of remote routes with the latency of local routes (§6.1.3). Finally, we measure the latency variation of each route and evaluate if remote peering introduces higher latency variability compared to the local route (§6.1.4).

6.1.1 Which Route had the Shortest AS Path?

Prefixes with both local and remote routes can be problematic for traffic engineering because an AS might choose a higher-latency route with a shorter AS path, since AS path length is the BGP second tie-breaker (after local preference) (REKHTER; HARES; LI, 2006). To examine whether this was the case, we compared the AS path length of

Figure 6.1: AS path lengths of prefixes reachable via both remote and local peers. Regardless of the method to infer RP, the majority of prefixes with both local and remote routes had remote routes with an AS path length shorter or the same length as the local route, and therefore likely chosen by BGP, a hypothesis we have confirmed using data from RouteViews peers (§6.1.2).



(a) Geographical RP.

(b) Reseller RP.

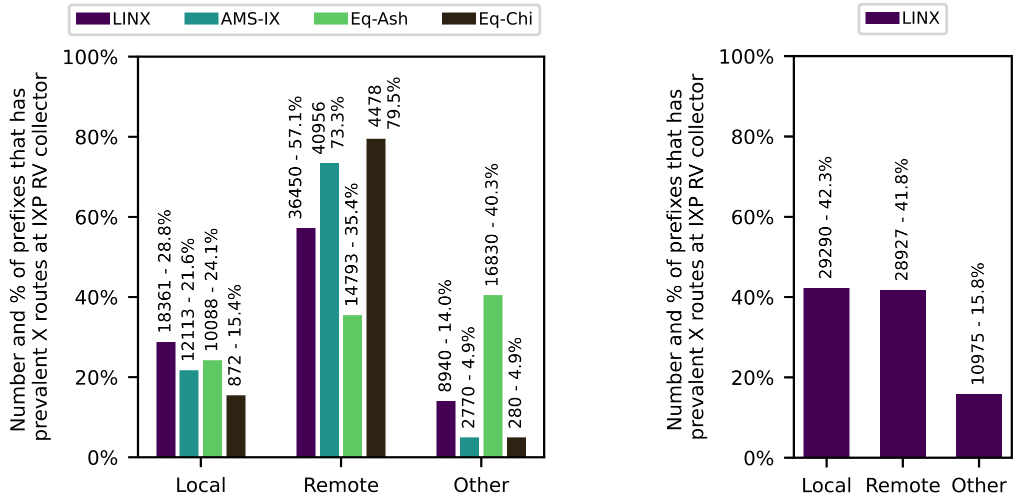
Source: The authors

routes for every prefix announced via remote and local peerings seen in IXP routing data, reporting the analysis for the IXPs that had a considerable number of these cases, namely LINX, AMS-IX, Eq-Ash, and Eq-Chi (§5.4). To compare routes, we selected the shortest AS path route of each type, local and remote. In order to observe the path lengths as they appear in the routing data, we do not reduce paths with AS path prepending.

Remote Routes had Shorter AS Paths than Local Routes. Figure 6.1 shows the percentage of prefixes with a shorter AS path length per peering type. In Figure 6.1a, most Geographical RP routes (an average of 82.5%) had shorter (or equal) AS path lengths, with the remaining 17.4% having a shorter AS path for the local route. Thus, BGP may choose a remote route over a local route if BGP uses AS path length as a tie breaker. The difference in AS path lengths for most prefixes with different length routes was a single ASN (82.1%, 79.0%, 73.9%, 89.9% for LINX, AMS-IX, Eq-Ash, and Eq-Chi). This happened because the local route was usually announced by large transit providers connected to the IXPs, which include the transit provider's ASN in the path.

Figure 6.1b shows the distribution when looking at the Reseller RP inferences for LINX. We only show LINX because the PTT-SP, PTT-RJ, and PTT-CE results are similar but from a much smaller number of prefixes associated with resellers (fewer than 600

Figure 6.2: The type of selected route by peers of RouteViews collectors at each IXP for prefixes with both local and remote routes. The remote route was more likely to be selected for Geographical RP. For Reseller RP, preference between remote and local routes was the same – ≈42%.



(a) Geographical RP routes.

(b) Reseller RP routes.

Source: The authors

prefixes each). Again, we find that the remote routes tend to have shorter AS paths – 40.7% of remote prefixes had the shortest AS path, whereas only 24.4% of local prefixes had the shortest AS path. The difference in path length for most prefixes with different length routes was also a single ASN – 62.5% of the prefixes with different AS path lengths for LINX.

6.1.2 Are Shorter AS Path Remote Routes Chosen?

Next, we want to understand the extent to which remote routes are preferred over local routes. We analyze how frequently the remote routes appear in routes shared by RouteViews peers in the IXPs (§4.2). For each prefix with both local and remote routes announced, we find all the routes the RouteViews peers see and compare them with the routes in the dataset used in the previous section. A remote (or local) route is prevalent among RouteViews peers when most peers see the route. It was also possible that most peers reported a different route, neither local nor remote, which we did not observe in the IXP routing data that we used.

Figure 6.2 shows how often each kind of route was preferred according to Route-

Table 6.1: Number of prefixes that had lower latency via remote or local peers. Generally, a route from a local peer had lower latency than a route from a remote peer to reach addresses in the same prefix.

IXP	Reseller RP		Geographical RP	
	Remote lower	Local lower	Remote lower	Local lower
PTT-SP	131 (51.1%)	125 (48.9%)	112 (20.9%)	423 (79.1%)
LINX	21,001 (45.5%)	25,155 (54.5%)	13,721 (33.0%)	27,903 (67.0%)
AMS-IX	-	-	6,644 (38.8%)	10,477 (61.2%)
NAPAfrica	-	-	14 (28.0%)	36 (72.0%)
PTT-RJ	10 (76.9%)	3 (23.1%)	53 (26.1%)	150(73.9%)
PTT-CE	4 (57.1%)	3 (42.9%)	4 (66.7%)	2 (33.3%)
Eq-Ash	-	-	2,230 (9.4%)	21,561 (90.6%)
Eq-Chi	-	-	830 (25.0%)	2,486 (75.0%)

Views peers: the local, the remote, or a different route which was not in our data set (*other* in Fig 6.2). We find that the remote route was more commonly chosen. For Geographical RP routes at LINX, AMS-IX, and Eq-Chi, these remote routes were chosen for at least 57.1% of the prefixes, compared to 28.8% or fewer local routes, and 14.0% or fewer other routes. When a remote route was prevalent among RouteViews peers, the remote route had the shortest AS path among the routes (local, other) for most prefixes (83.5%, 90.0%, 81.3%, and 98.5% of these prefixes, respectively, for LINX, AMS-IX, Eq-Ash, and Eq-Chi). When local routes were prevalent, they were not always the shortest AS path routes available, and the IXP had a remote route with shorter or equal AS path length (64.5%, 39.7%, 76.6%, and 61.0%, respectively, for LINX, AMS-IX, Eq-Ash, and Eq-Chi). This suggests that operators might have been using local policy to prefer local routes so that the remote routes with shorter AS paths were not selected by BGP.

For Reseller RP routes (Figure 6.2b) the situation was different: preference between remote and local routes was similar (≈42%), with other paths accounting for the remaining 15.8%. For 75.2% of the prefixes with remote routes prevalent, the remote paths had shorter AS paths. When local routes were prevalent, 58.4% of prefixes had a remote alternative with shorter or equal AS path length available at the IXP.

6.1.3 Is There a Latency Penalty Using a Remote Route?

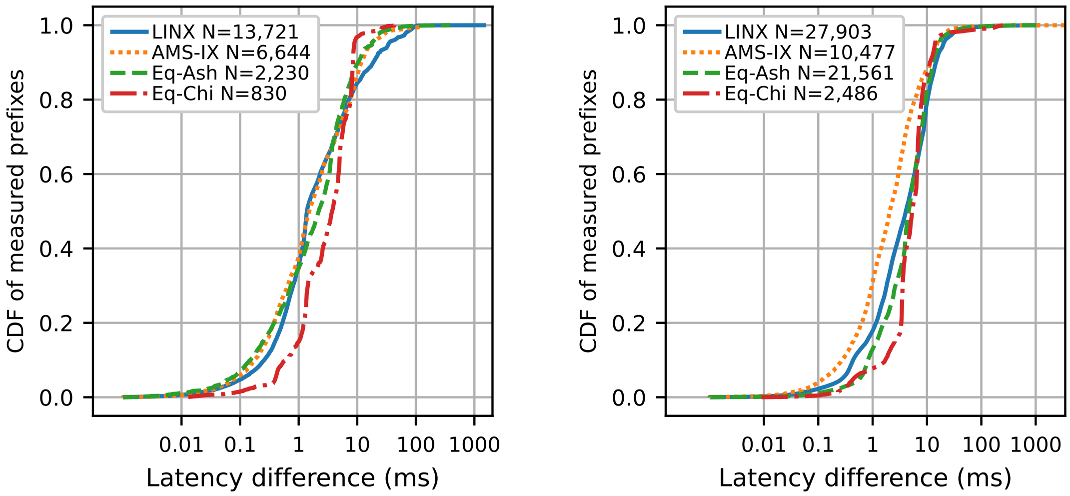
Considering the current preference for peers to select remote routes, we wanted to understand whether they were also the best route latency-wise. We performed active measurements, using traceroutes toward IP addresses within the prefixes set seen in IXP routing data. Since we did not have a pre-selected list of responding servers, we initially probed the first ten addresses in the IP block of the prefix, followed by thirty IP addresses randomly selected, from a system external to the IXP. Because not every prefix had a responsive address, the set of measured prefixes is smaller than the original set of prefixes. We then ran ICMP-Paris traceroute measurements to these IP addresses from RouteViews VPs in the IXPs over two days and compared the latency of the remote and local routes, provided we had obtained at least five responses from addresses in each type of route. Because a prefix can have multiple remote or local routes, we used the lowest latency measured when comparing each route type – i.e., we compared the lowest latency local and remote routes.

Local Routes had Predominantly Lower Latency than Remote Routes. Table 6.1 shows the number (percentage) of prefixes where a remote route had lower latency than the local routes. Looking at Geographical RP first, local routes had lower latency than remote routes for nearly all analyzed IXPs. When focusing on the IXPs with a higher prevalence of prefixes with both local and remote routes (e.g., LINX, AMS-IX, Eq-Ash, and Eq-Chi), up to 90.6% of the measured prefixes had lower latency using a local route. Similarly, for the Reseller RP inferences in LINX, the majority of prefixes also had a lower latency local route.

The previous analysis was binary – which route had the lowest latency. We now analyze the differences in latency. Figure 6.3 shows the latency difference between remote and local routes. The figures have a different number of points, as the number of prefixes with lower latency for remote or local routes shown in Table 6.1 are different. Figure 6.3a shows that when a Geographical RP provided a route with lower latency than the local route, the advantage was small: for at least 72.9% of the prefixes, the latency benefit of the remote route was restricted to 5ms or less for three IXPs. In contrast, when the local route was faster, as shown in Figure 6.3b, the latency advantage was more pronounced. For at least 44.7% of prefixes in three IXPs, the latency benefit for the local route was more than 5ms when compared to the corresponding remote route. When looking at Reseller RP for LINX in Figure 6.3c, we observe that the distribution of latency differences was similar for both remote and local routes, with nearly 20% of the prefixes having a latency difference above 10ms.

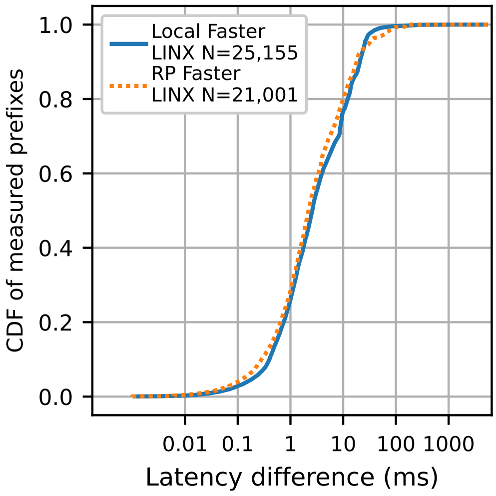
Figure 6.4 shows a CCDF of the *relative* latency difference between remote and local routes when the latency differed by more than 5ms. The left side of the figure shows the prefixes where the local route had lower end-to-end latency than the remote route,

Figure 6.3: Latency difference between remote and local routes measured by end-to-end latency to reach an address in a remote prefix. For Geographical RP, when local routes had lower latency, the advantage compared to the remote route was more than 5ms for at least 44.7% of prefixes in three IXPs



(a) Geographical RP lower latency

(b) Local peering lower latency

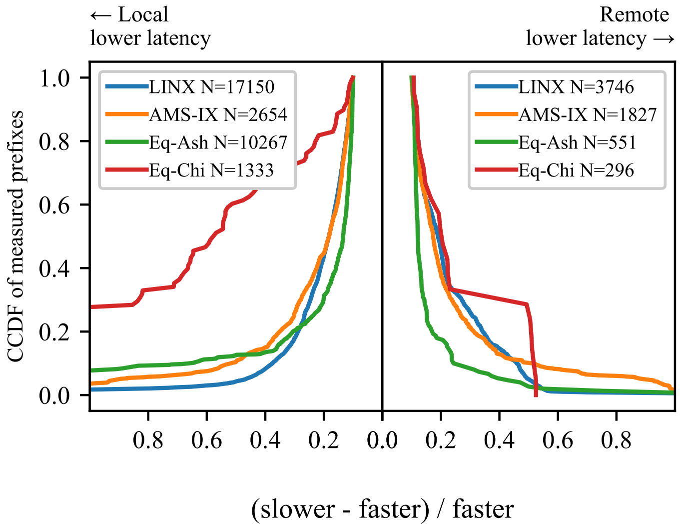


(c) Reseller RP vs. local peering

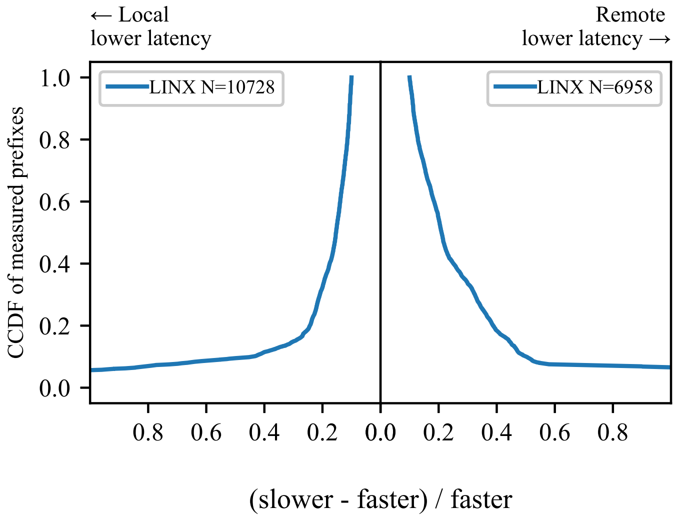
Source: The authors

while the right side shows when the remote route had a lower latency than the local route. The x-axis represents how much faster one route was when compared with the other. For example, an x equals 0.2 shows that for some fraction of prefixes (in the y-axis), one type of route was 20% faster than the other type of route. We see on the left side of Figure 6.4a that local routes are up to 30% faster (better) for 75.1% of prefixes observed in three IXPs. For Eq-Chi, 50% of prefixes are at least 57.8% faster (better) via a local route than using

Figure 6.4: Relative comparison of end-to-end latencies. For Geographical RP, when either the local or remote route had lower latency, the route had up to 30.7% lower latency than when compared with the other route type for 75.1% of prefixes in three IXPs. For Reseller RP, when a remote route had lower latency, its advantage over the local route tended to be higher than vice-versa.



(a) Geographical RP perspective.



(b) Reseller RP perspective.

Source: The authors

the remote one. On the right side, we see a similar pattern, where remote routes have RTTs less than 30.7% lower (better) for 75.1% of prefixes observed in three IXPs. The situation was different for Reseller RP inferences for LINX. As shown in Figure 6.4b, when the remote routes via reseller had lower latency, they were at least 20% faster for 54.6% of prefixes, while when the local route had lower latency, they were at least 20% faster for only 32.5% of measured prefixes. In summary, the results suggest that with

Table 6.2: Breakdown per IXP when comparing remote and local routes for each prefix in terms of latency and AS path length – Geographical RP only. A large number of local routes had lower latency but had a longer AS path than the remote route.

IXP	Total prefixes	Remote lower latency, *longer* AS path length	Remote lower latency, *equal* AS path length	Local lower latency, *longer* AS path length	Local lower latency, *equal* AS path length
LINX	41,624	1,177 (2.8%)	2,185 (5.2%)	12,950 (31.1%)	9,636 (23.2%)
AMS-IX	17,121	1,397 (8.2%)	657 (3.8%)	4,798 (28.0%)	1,828 (10.7%)
Eq-Ash	23,791	270 (1.1%)	674 (2.8%)	9,547 (40.1%)	5,579 (23.5%)
Eq-Chi	3,316	57 (1.7%)	161 (4.9%)	2,149 (64.8%)	111 (3.3%)

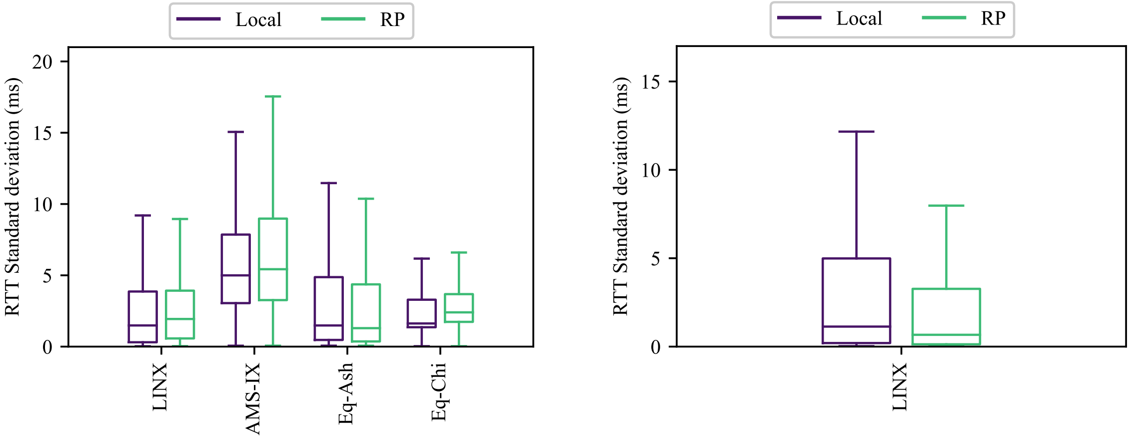
proper configuration and knowledge about these cases, ASes can decide which route to select and steer their traffic, potentially enabling better performance according to their specific goals.

The Path with Lowest Latency was Not Always Preferred by BGP. Table 6.2 shows the percentage of prefixes where the route with lowest latency would not match the route specified in a BGP tie-breaker. We observed a small percentage of prefixes where the remote route had lower latency but also had a longer AS path when compared to the local route (no more than 8.2%.). In contrast, there were proportionally more cases of prefixes for which the local route had lower latency but a longer AS path than the remote route, varying from 28% (AMS-IX) up to 64.8% (Eq-Ash). When both the remote and local routes had the same path length, the local peering predominantly had a latency advantage over the remote routes despite the latency benefit not being higher than 5ms for most routes. The results for Reseller RP, obtained from LINX, follow a similar pattern (as in Table 6.2) and are omitted. In summary, the results indicate that the shortest AS path route may often not match the route with the lowest latency.

6.1.4 Do Remote Routes Have More Latency Variability than Local Routes?

In discussion with network operators, there was a concern about potential latency variability that could be introduced by a layer-2 connection or the geographic distance separating the AS's router to the IXP. To compare the relative latency variability of remote routes over local routes, we performed active measurements by sending at least 120 ping packets from the scamper prober at the IXP RouteViews node to an address in each of the prefixes with both local and remote routes seen in Table 6.1 over ≈4 days (depending on the size of the IXP): at least 60 packets via the local route and at least 60 via the remote

Figure 6.5: Latency variability to remotely announced prefixes via remote and local routes. The latency variability to reach remote destinations was similar for both local and remote routes, suggesting that reseller connections and geographical distance had limited impact on latency variability.



(a) Geographical RP vs. Local.

(b) Reseller RP vs. Local.

Source: The authors

route. We computed the latency standard deviation for the best remote and local routes for the prefixes we used in the latency comparison in the previous section.

Remote and Local Routes had Similar Latency Variability. Figures 6.5a and 6.5b show the latency variability was similar between remote and local routes. Regardless of peering type or remote peering perspective, 75% of the prefixes had less than 10ms of latency variability. More specifically, for three of the four analyzed IXPs, the same fraction of prefixes had latency variability below 5ms. The results indicate that variability was *not* a distinguishing feature at least for the IXPs we considered.

6.2 Does Remote Peering have Lower Latency than Transit?

When prefixes announced via RP do not also have routes from a local peer at the IXP, ASes must decide between delivering their traffic via the remote peer at the IXP or using a transit provider. Which connection type presents the lower latency to reach these prefixes? Discussions in the network operator community concern whether remote peering is an inferior alternative to transit in both latency and connection stability (NANOG, 2017; LEVY, 2019).

To assess whether remote peering or transit had lower latency to reach addresses in prefixes exclusively announced at an IXP via remote peers, we performed traceroute

Table 6.3: Latency comparison between remote peering or transit, showing the number of prefixes with lower latency. For Reseller RP, in four IXPs, at least 64.9% of the prefixes had lower latency via Reseller RP routes than via transit. For Geographical RP, seven of eight IXPs had at least 57.6% of prefixes with lower latency via remote peering routes than via transit.

IXP	Reseller RP latency		Geographical RP latency	
	Remote lower	Transit lower	Remote lower	Transit lower
PTT-SP	8,886 (74.2%)	3,085 (25.8%)	5,657 (72.0%)	2,205 (28.0%)
LINX	10,342 (77.7%)	2,973 (22.3%)	2,724 (71.0%)	1,108 (29.0%)
AMS-IX	-	-	2,651 (57.6%)	1,950 (42.4%)
NAPAfrica	-	-	1,787 (98.1%)	35 (1.9%)
PTT-RJ	1,929 (64.9%)	1,045 (35.1%)	1,113 (59.6%)	754 (40.4%)
PTT-CE	3,014 (71.7%)	1,190 (28.3%)	2,648 (71.3%)	1,065 (28.7%)
Eq-Ash	-	-	708 (28.9%)	1,740 (71.1%)
Eq-Chi	-	-	1,204 (94.6%)	69 (5.4%)

measurements through the remote peers at eight IXPs, as well as a transit provider from the same location (§6.2.1). We compared the latency variability of both RP and transit to reach these remotely announced prefixes (§6.2.2).

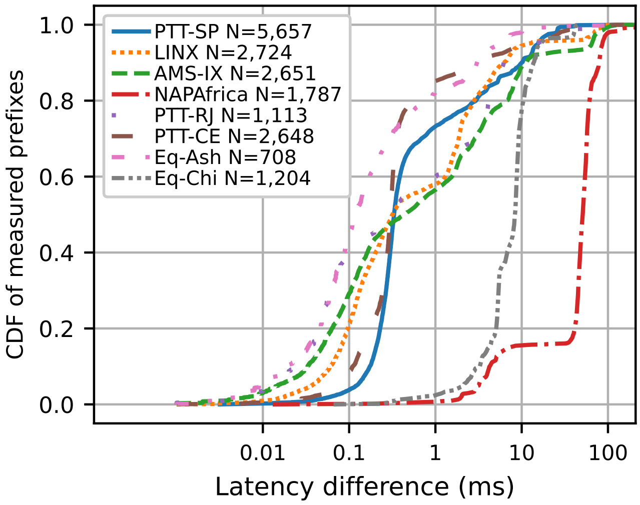
6.2.1 Does Transit Offer Lower Latency than Remote Peering?

We collected latency measurements to addresses in prefixes announced by remote peers both using the remote routes and a transit route using a similar approach to §6.1.3 – we first identified remote prefixes without a local route and responsive IP addresses in each prefix. We collected at least five latency samples for each remote prefix using a remote peer and the transit provider.

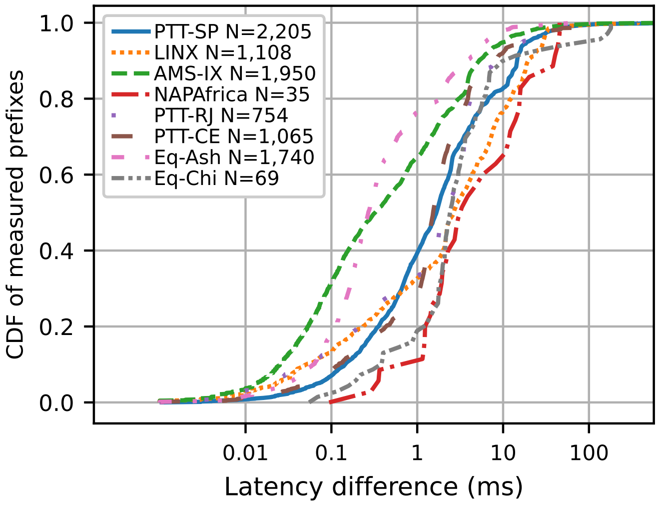
Table 6.3 shows the number of probed prefixes per IXP, along with the connection type (remote or transit) with lowest latency. Note that the number of prefixes with a measurement is lower than the number of prefixes observed in the routing table (§6.1.3), as in some cases we failed to identify a responsive address for the prefix. The remote route had lower latency for most prefixes: 57.6% of the prefixes had lower latency with Geographical RP routes for seven out of eight IXPs, and 64.9% for Reseller RP.

Remote Routes can have a Substantial Latency Advantage. Figure 6.6a and 6.6b show the absolute latency difference for Geographical RP. Figure 6.6a shows that some remote routes had latencies substantially lower than the the transit alternative in some IXPs. In

Figure 6.6: Latency difference between Geographical RP and transit provider routes measured by latency to addresses in remote prefixes. Remote peering had a substantial advantage for a few IXPs (NAPAfrica, Eq-Chi), but not as a substantial advantage for others (less than 5ms for 78.1% of measured prefixes).



(a) RP with lower latency.

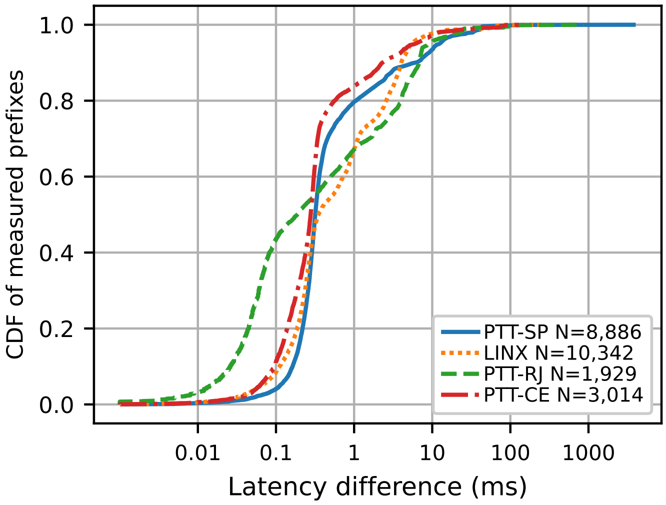


(b) Transit with lower latency.

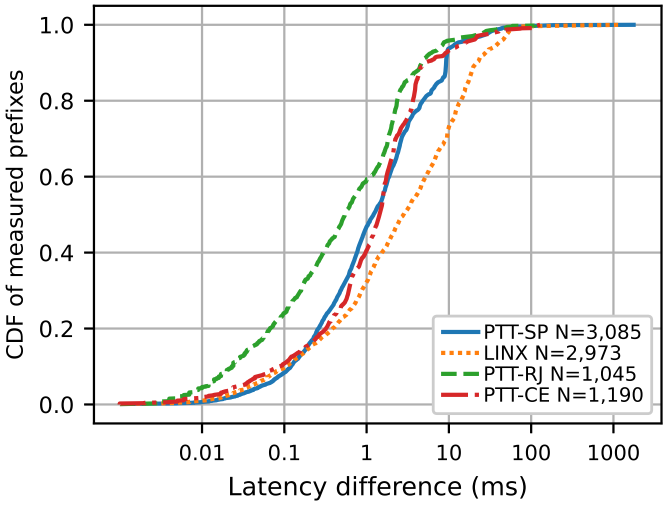
Source: The authors

NAPAfrica, 81.4% of remote routes with lower latency than transit had at least 40ms lower latency. When we discussed our results with resellers, they suggested that high IP transit prices, along with poor ISP interconnectivity and performance in Africa, made remote peering a lower latency and cheaper option, in line with the published literature (GUPTA et al., 2014; FANOU; VALERA; DHAMDHERE, 2017; FORMOSO et al., 2018). For the remaining IXPs, the latency difference between remote routes and transit was not

Figure 6.7: Latency difference between Reseller RP and transit measured by the latency to reach remote prefixes. When Reseller RP had lower latency, the latency advantage was not substantial (below 1ms for over 67.2% of the measured prefixes). When transit routes had lower latency, the latency advantage was a bit higher (more than 1ms for 53.1% of the measured prefixes in three IXPs).



(a) RP with lower latency.



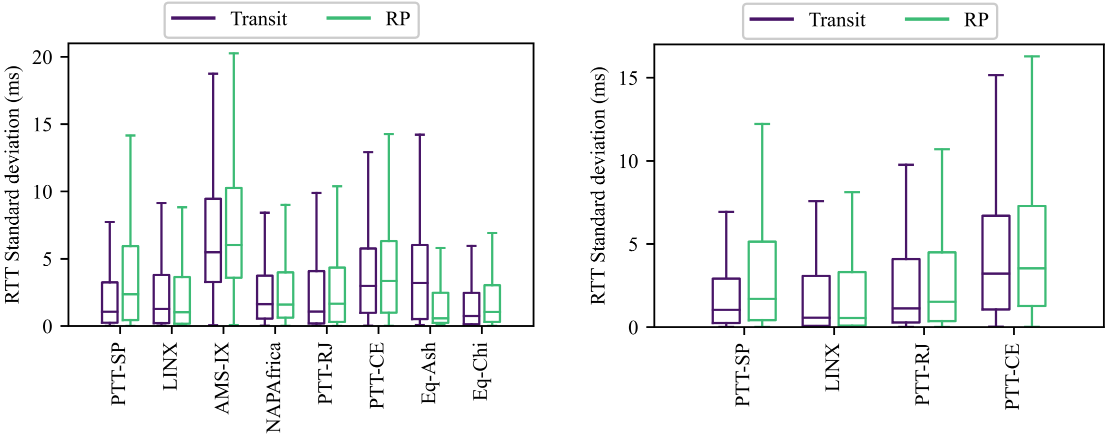
(b) Transit with lower latency.

Source: The authors

substantial. Regardless of which route had lower latency, in six IXPs, we observed that the latency difference was below 5ms for at least 78.1% of the measured prefixes.

Figure 6.7a and 6.7b show the results for Reseller RP. Figure 6.7a suggests that any latency advantage of remote peering was not substantial. For more than 67.2% of remote routes with lower latency, the latency advantage was within 1ms. In comparison,

Figure 6.8: Latency variability to remotely announced prefixes via remote peers and transit providers. The latency variability to reach addresses in remote prefixes was similar between transit and remote peering in all IXPs (latency standard deviation less than 10ms for 75% of measured prefixes), suggesting that neither transit or the remote peering had a substantial effect on latency variability.



(a) Geographical RP perspective (b) Reseller RP perspective

Source: The authors

Figure 6.7b suggests that when transit was faster for three out of four IXPs, the latency advantage was a bit higher: in at least 53.1% of transit routes with lower latency, the advantage was more than 1ms.

6.2.2 RTT Variability of Remote Prefixes

In §6.1.4, the latency variability to reach addresses using either remote or local routes was similar. To understand if using a transit provider introduces more latency variability, we performed ping measurements to exclusively announced prefixes seen at Table 6.3. Similar to the previous measurements, we sent at least 120 ping probes from to each prefix over up to 4 days (depending on the size of the IXP): 60 (at least) via the transit provider and 60 (at least) via the remote route. We then computed the latency standard deviation among the ping probes for the measurements via remote peering and transit.

Transit and Remote Peering had Similar Latency Variability. Figures 6.8a and 6.8b show the latency variability for remote peering. The latency variability to reach prefixes exclusively announced at an IXP via a remote peer was equivalent for both remote and transit. PTT-SP and Eq-Ash were the only IXPs where a fraction of the prefixes had higher latency variability (see Figure 6.8a).

Still, for all the IXPs, the standard deviation for 75% of the prefixes was below 10ms. We observed a similar trend for Reseller RP inferences, where resellers and transit had comparable latency variability.

6.3 Summary

This chapter investigates the latency implications of sending traffic to ASes connecting via RP at IXPs. First, we examine whether remote routes tend to be preferred over local ones in the IXP BGP routing data. Our results show that most remote routes for prefixes with both local and remote routes had a shorter or equal AS path length than the available local routes and tended to be preferred by the peers of RouteViews collectors.

Next, we performed latency measurements to all prefixes with remote and local routes at IXPs, using each connection alternative as path to reach the destination. Our results show that despite being shorter and indeed preferred, remote routes were not necessarily the lowest latency route. For at least 61.2% of these prefixes in seven IXPs, the local route had lower latency compared to the geographically distant remote peering routes.

Finally, we analyse if RP could be a reliable alternative to Transit regarding latency for prefixes with only remote routes at IXPs. Our measurements suggest that relying on remote routes can be an advantageous option for end-to-end latencies. In some scenarios (NAPAfrica and Eq-Chi), remote routes at the IXPs had considerably better latency results when compared to Transit, showing latency improvements of at least 40ms for 81.4% of the measured prefixes when the remote route was faster than Transit. For the other six IXPs, we observed that the latency difference of using the remote route or the Transit was no higher than 5ms for 78.1% of the measured prefixes.

7 THE IMPACT OF RP TO LATENCY: THE REMOTE AS PERSPECTIVE

After evaluating the latency impacts of sending traffic to prefixes announced remotely, we investigate how RP compares to Transit providers performance-wise when used by ASes to reach prefixes announced in international IXPs. In this chapter, we leverage the infrastructure of a large RP reseller and compare the latency benefits of reaching prefixes announced at two IXPs via RP or relying on Transit providers. First, we detail our measurement methodology (§7.1), followed by the obtained results from our analyses (§7.2).

7.1 How Do We Measure It?

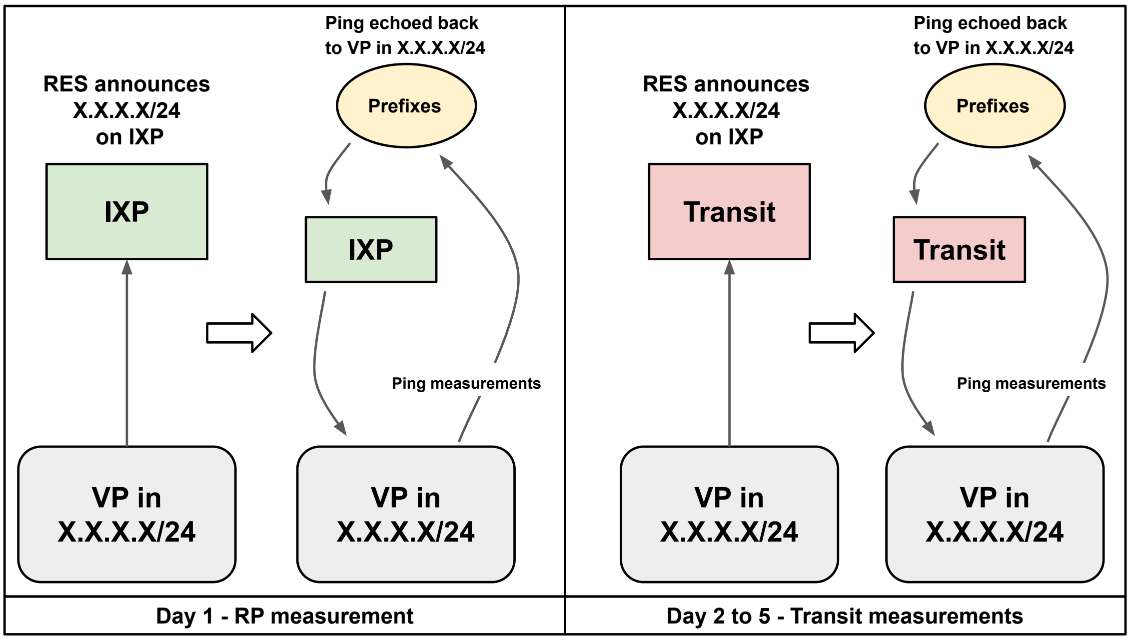
7.1.1 IXP BGP Data.

We used prefixes announced at two IXPs (AMS-IX and LINX) as measurement destinations to evaluate the performance differences between RP and Transit providers. We collected these routes using the Looking Glass ALICE API of each IXP (AMS-IX, 2023b; LINX, 2023). The Looking Glasses receive all the routes being announced to the IXP's Route Server via multilateral peering. Our data collection was performed on 2023-02-01 for AMS-IX and 2023-02-29 for LINX, resulting in 297,217 and 288,291 prefixes, respectively. After collecting the announced prefixes for both IXPs, we matched them with the CAIDA IPv4 Prefix-Probing Traceroute Dataset (ARK. . . , 2023) to obtain responsive addresses inside each prefix and use as the destination for our active measurements. After this step, we obtained 124,838 pingable addresses for AMS-IX and 118,591 for LINX.

7.1.2 Measurement Infrastructure.

As our current goal is to perform measurements originating from the AS connected via RP, we now need a measurement infrastructure different from the one previously that used VPs located inside the IXP LAN (§4.3). We take advantage of the infrastructure of a large Latin American RP reseller (we will call it RES for the remainder of the chapter) as VP for our measurements. RES has remote connections, offering its customers access

Figure 7.1: Illustrated example of the measurement experiment considering the RP customer AS perspective. The prefixes in the yellow elispsis represent all the announced prefixes at the measured IXP and are not related to the /24 prefix being announced by RES.



to multiple IXPs in Europe and North America. Besides the remote services, RES also maintains connections to multiple Transit providers and sells IP Transit for its customers. Our VP consists of a Linux Virtual Machine (VM) directly connected to one of RES's core routers deployed in Brazil.

7.1.3 Experiment Design.

We perform active latency measurements to prefixes announced in the IXPs and compare the remote connection of each IXP with four different Transit providers. According to RES, when ASes connect via RP at international IXPs, the return path and traffic (IXP to eyeball networks) are critical for ASes' application performance. To guarantee the return path will be the IXP via RP or any of the four Transits, we announce the VP X.X.X.X/24 prefix on the desirable provider for the measurements period. We exemplify our experiment design in Figure 7.1.

First, on day 1, we announce the VP prefix on the IXP Route Server. In this way, all IXP members receiving traffic originating in this prefix will send back the responses via the IXP. After waiting 15 minutes to allow BGP to converge, we perform 30 pings for each prefix with identified pingable addresses announced at the IXP RS using the *nping* tool (NMAP.org, 2023). Our measurements last approximately 10 hours.

For days 2 to 5, we conduct a similar process and announce the VP prefix to one of each target Transit provider. In this case, ASes receiving traffic originating in this prefix will send back the responses via the targetted Transit. We then perform measurements to the same prefixes measured on day 1.

We apply a few steps to validate our results and guarantee they are accurate. First, we remove any prefix also announced at Brazilian IXPs to avoid the possibility of receiving the measurement responses over a local peering interconnection. Next, after announcing the prefix, we validate that the return path is going via the targetted provider or IXP by performing traceroute measurements from at least 50 RIPE atlas probes to the VP IP Address. Lastly, we discard prefixes where the results for the IXP measurement were lower than 150ms (representing less than 7% of the cases). RTTs lower than 150ms are a great indication that the response for that prefix is not arriving over the RES RP connection, considering the average distance between the VP and the IXPs (i.e., Amsterdam and London) and speed of light on fibre constraint (WonderNetwork, 2023).

7.2 Deciding Between RP and Transit Providers

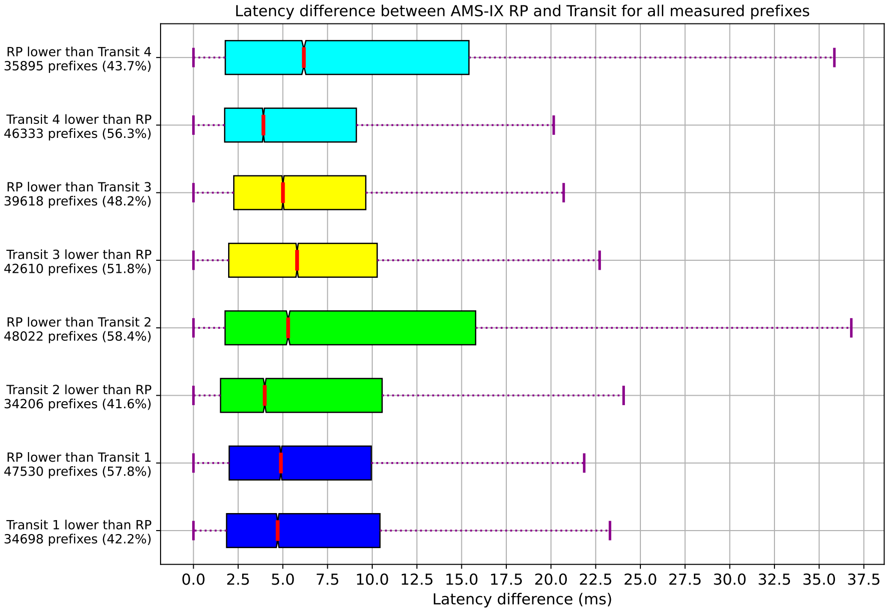
We compare the latency performance of all connections using two primary metrics: minimum RTT and RTT variability. Studies frequently employ these metrics to assess VoIP, Video, and Web application performance and estimate QoS/QoE for delay-sensitive services (IORIO; RISSO; CASETTI, 2021; LIOTOU; TSOLKAS; PASSAS, 2016; DIMOPOULOS et al., 2016; HOHLFELD et al., 2014; KUSCHNIG; KOFLER; HELLWAGNER, 2011)

7.2.1 Does RP Really Improve Latency Performance?

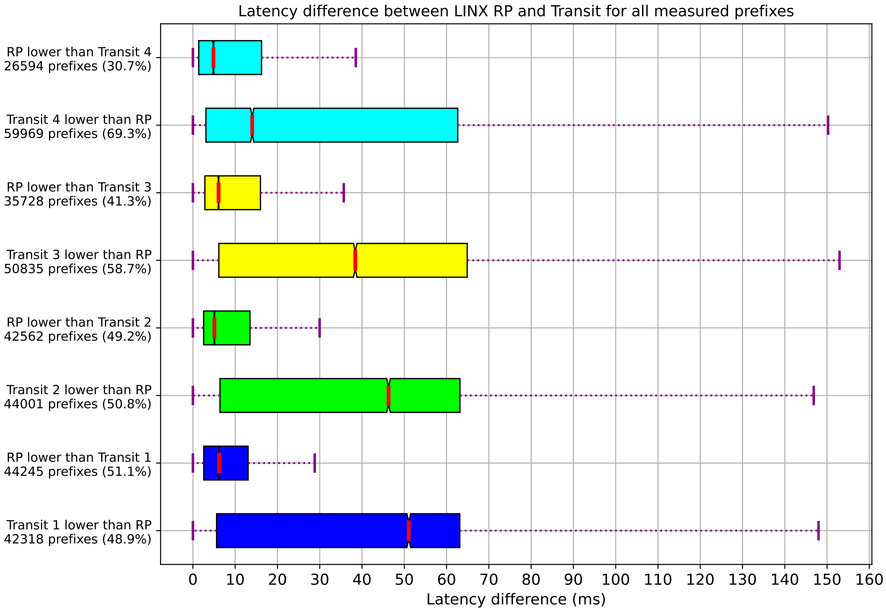
To define the connection with the best latency performance to a prefix, we compare the minimum RTT observed among the 30 ping measurements between the RES RP connection and each of the N Transit.

If multiple connections are available, RP provides lower RTT for a fifth of the prefixes. First, we investigate how the RP connections at IXPs perform in a scenario where RP and the four Transit connections are available for reaching the measured prefixes. We

Figure 7.2: Results considering the Min RTT latency difference between RP and the N Transit for the measured prefixes.



(a) AMS-IX vs. four Transit providers

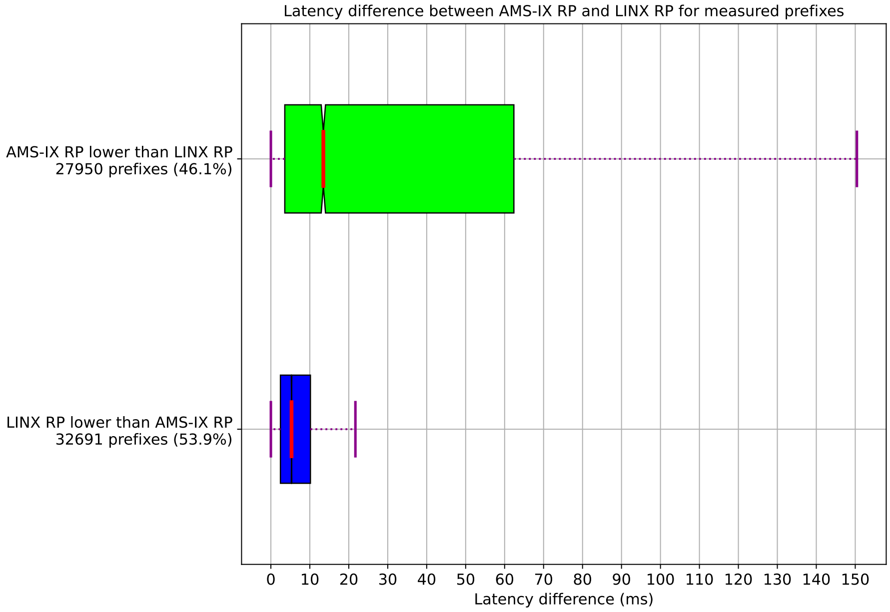


(b) LINX vs. four Transit providers

Source: The authors

compute it by looking at each connection's ranking position among the five alternatives on the end-to-end latency to reach a prefix. For AMS-IX, we observe that the RP connection is the best option for 21.7% of the measured prefixes. For the remaining 78,3% of the 82,228 measured prefixes, RP is the second best option for 22.6%, 3rd for 18.1%, 4th for 17.1% and 5th for 20.4%. When looking at LINX RP, using the remote connection

Figure 7.3: Results considering the Min RTT latency difference between the RP connection at AMS-IX and LINX for prefixes announced at both IXPs.

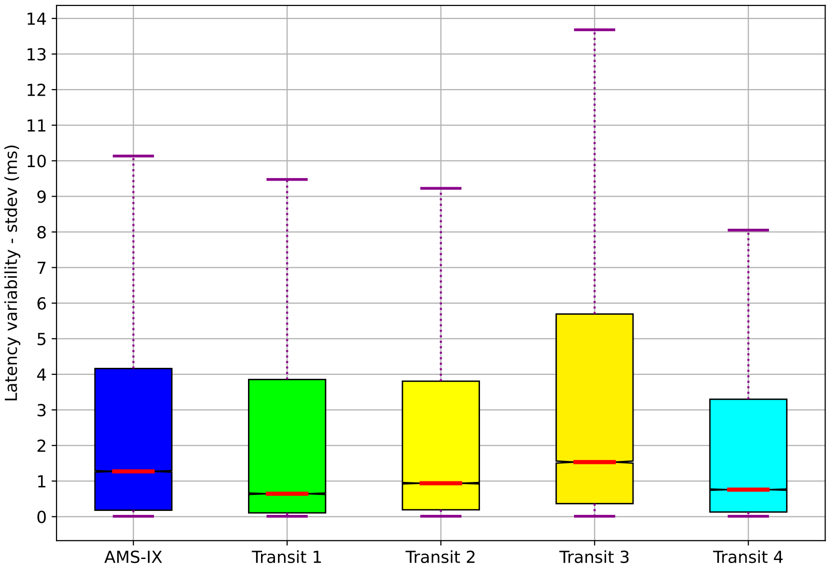


Source: The authors

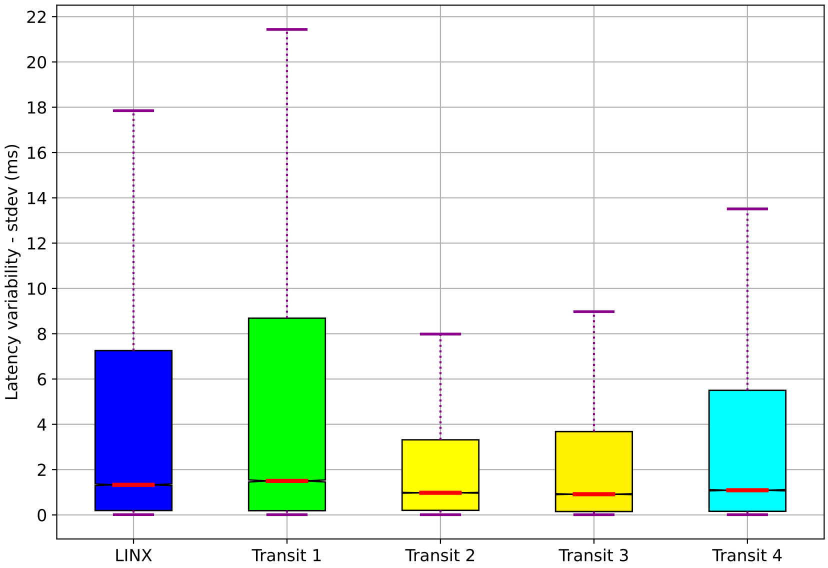
is the best option for 19.2% of the measured prefixes, 2nd for 19.7%, 3rd for 12.4%, 4th for 11.4% and, surprisingly, the worst option for almost 40% of the destinations (37.2%). The results show that when multiple connection alternatives are available, relying on RP can provide better latency to a small fraction of the measured prefixes.

The IXP where the RP connection is can considerably influence the peering latency performance. After comparing the RP performance with all the available alternatives, we investigate how it compares with each Transit individually for all the 84255 measured prefixes. Such analysis allows us to isolate the benefits/drawbacks and the absolute differences in latency for each set of connection alternatives. Figure 7.2 shows boxplots with the latency difference results for the minimum RTT analysis comparing AMS-IX and LINX with the four Transit providers. Looking at the AMS-IX case (Figure 7.2a), we can observe that, on average, RP and Transit providers share the advantage of having a lower min RTT for half of the prefixes. Regarding how much difference they deliver, when RP have lower min RTT, it tends to have comparable or better latency difference to Transits. In two cases, for 25% of these prefixes, RP can have a difference to the Transit connection reaching up to 36.8ms. On the other hand, we can observe that the results for LINX (Figure 7.2b) are reasonably different. At LINX, RP has a lower min RTT than Transit for half of the prefixes or less (30.7% to 51.1%, depending on the Transit). It is also possible to notice a considerably higher latency difference when the Transit alterna-

Figure 7.4: Results considering the latency variability (stdev) between providers for a prefix.



(a) AMS-IX vs. four Transit providers



(b) LINX vs. four Transit providers

Source: The authors

tive is better for a prefix. For at least 50% of the destinations with lower min RTT, Transit has an advantage varying from at least 14ms to at least 51ms, giving the provider.

To understand whether the meaningful difference is solely caused by the IXP where the RP connection is deployed, we select the prefixes being announced at both AMS-IX and LINX and compare the min RTT measurements performed to them. Figure 7.3 displays boxplots with the latency difference results for each IXP RP with lower RTT. Even though most prefixes have lower RTT via LINX, the performance difference to AMS-IX is relatively irrelevant. When AMS-IX have better performance, the latency difference is higher than 13.5ms for more than 50% of the prefixes, reaching a difference of up to 150ms.

7.2.2 Does RP Deliver Less Latency Variability than Transit?

Now, we analyse the RTT variability among the five different connection options to reach the measured prefixes. We compute the standard deviation for each connection using the 30 ping measurements to a prefix to obtain the results.

RP has similar RTT variability to Transit connections. Figure 7.4 shows the RTT variability in comparing RP via AMS-IX, LINX, and the four Transit. We can observe that the latency variability is comparable for the RP and Transit connections. On AMS-IX, RP and all providers have variability lower than 6ms to 75% of the prefixes. On the other hand, data for the LINX scenario is more volatile. Despite the median variability being similar to all, RP shows more variability than most Transit, with a standard deviation reaching up to 18ms.

7.3 Summary.

In this chapter, we investigate the RP latency impact to ASes using it as a way to reach destinations in international IXPs. Our measurements indicate that RP has lower RTT for, on average, half of the measured prefixes while maintaining a comparable latency variability to Transit providers. As RP primarily intends to reach prefixes unavailable locally with better performance or connectivity and exchange traffic in bulk on an international IXP, results show that RP can be a reliable alternative to Transits. If performance is not the primary goal for one AS, RP can still be an attractive interconnection option, as it allows direct access to the entire IXP routes while vastly decreasing the financial payments to rely on the Transit provider's infrastructure. Besides, our results show that when a network is considering alternatives to obtain connectivity to a set of prefixes unavailable locally, it must evaluate not only the type of connection (RP or Transit) but also the location of the connection in the RP case. Although located geographically close, a shift in the selected IXP can considerably affect the connection performance for reaching a set of destinations.

8 ROUTING IMPLICATIONS OF RP ON THE INTERNET

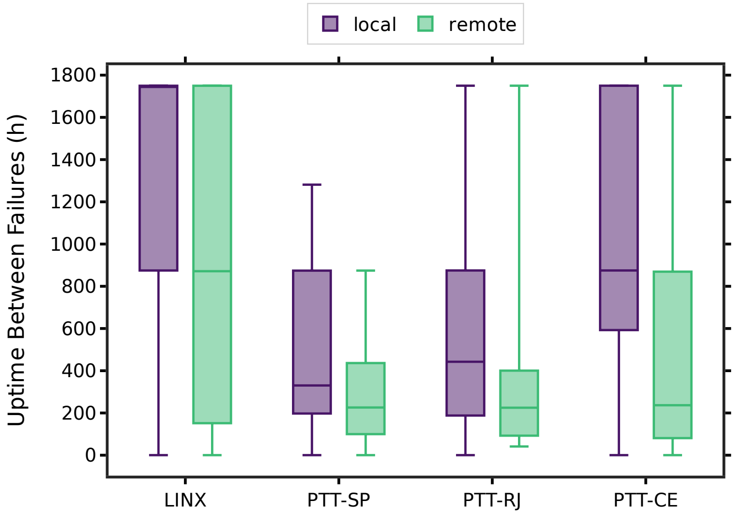
This chapter aims to better understand how RP can affect parts of the Internet, primarily focusing on routing features. We note that all of the following analysis were done considering a Geographical RP perspective (§5). First, we investigate whether RP can affect the stability of IXP members' interface connections (§8.1). Next, we analyze the usage of BGP communities at IXPs explicitly created to perform traffic engineering on remote networks (§8.2). Finally, we investigate cases of routing mispractices caused by RP which may introduce trombone paths (§8.3). Our goal is to provide a glimpse into the popularity of these BGP communities and their usage patterns by IXP members.

8.1 Connection Stability of IXP Members

In this section, we investigate to which extent RP can affect the stability of IXP members' interface connections to the IXP route server. Understanding how much a peering connection type is less stable than others can be helpful since downtime can cause significant financial, performance and reputational loss for ASes (HOLTERBACH et al., 2017). Based on earlier work (BERTHOLDO et al., 2021), we expect to see remote interfaces being less stable than local ones because they are not physically connected at the IXPs and rely on a third-party infrastructure in which they do not have control to reach the peering facilities. Besides, since many RP resellers usually connect multiple remote networks through shared logical ports, in which higher instability to one physical interface can affect many clients.

Methodology. To analyze the connection stability of remote and local interfaces at IXPs, we collected membership data from IXP LGs every 15 minutes over two months (2022-11-19 to 2023-01-23). This analysis is only feasible in IXPs with LGs allowing each of the IXP interfaces to be classified according to their remoteness: LINX, PTT-SP, PTT-RJ, and PTT-RJ. In AMS-IX, the LG presents all the IXP member interfaces under the Route Server interface address, while in NAPAfrica, Equinix-Ash, and Equinix-Chi, no LGs offer open access to BGP data (see §4.1). We exclude from the analysis a small fraction of IXP interfaces (1.6%) which either changed from remote to local and vice-versa or were not present during the entire collection period. We collected 7008 samples of the connected interfaces to RS on four IXPs (LINX, PTT-SP, PTT-RJ e PTT-CE) between

Figure 8.1: Median uptime between failures per IXP interface. Remote interfaces stay less time continuously without failures.



2022-11-19 and 2023-01-23 (see § 4.2).

Interfaces switch between two states, up or down, which we use to compute the following metrics:

- **Median Uptime Between Failures.** It represents the stability of a connection and is computed by the median of all the hours one interface stayed in the up state uninterruptedly during the collection period. If the interface remains up between data collection X and $X + 1$, we add 15 minutes of uptime to the interface.

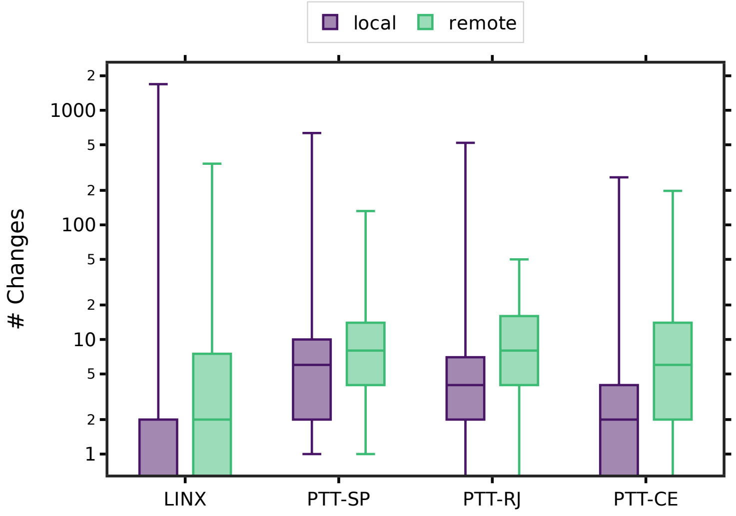
- **State changes.** The metric indicates how common errors are on IXP's interfaces. We compute it as the sum of state changes (up to down, and vice-versa), between all data collections.

- **Mean Time Between Failures.** MTBF is (a well-known metric) which we use to indicate the level of availability of interfaces. We compute it by dividing the total uptime by the number of failures.

- **Reliability.** It measures the ability of a system or component to perform its intended function for a specified period. The system's reliability can be calculated using various methods, including reliability models, failure data analysis, and reliability testing. We computed it using the MTBF metric and two different intervals, day and month: $Reliability = e^{-\frac{1}{MTBF} \times Time}$

RP interfaces stay less time continuously functioning than local ones. We first evaluate whether the connection type affects the interfaces' stability, by comparing the median

Figure 8.2: Number of state changes per IXP interface; the *y* axis is in log scale. Remote interfaces are subject to many more state changes.

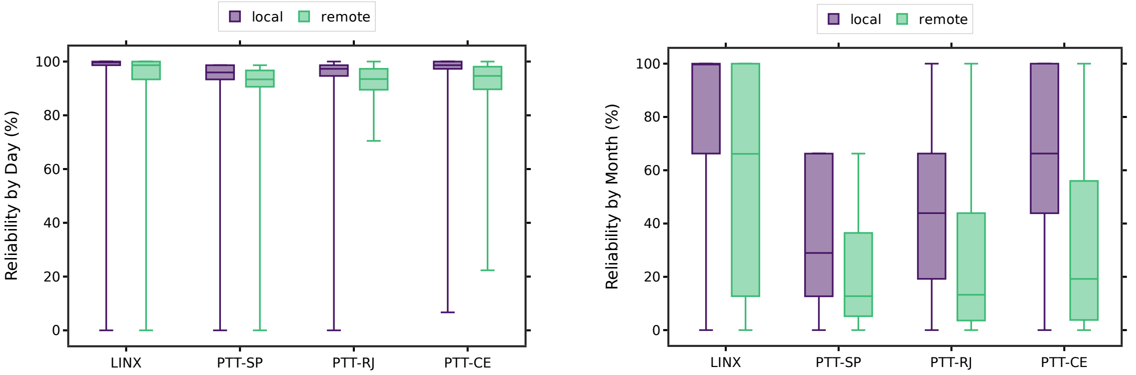


uptime between failures for remote and local interfaces. Figure 8.1 shows that the uptime for local interfaces is considerably higher than remote ones. In other words, remote interfaces tend to stay less time continuously functioning before a change of state (e.g., failure, BGP session issue). This difference is striking in the case of LINX and PTT-CE. For LINX, the disparity reaches two times higher uptime for local members (1749 hours to local vs 871 hours to remotes) when considering at least 50% of the interfaces. For PPT-CE, at least 50% of the local interfaces have uptime higher than 874 hours, while the uptime of at least 50% of the remote ones does not reach 236 hours (3.7 times lower).

Local interfaces suffer fewer state changes. Next, we look at the state change number for each connection type. Too many state changes indicate lower stability, since the BGP session of the members is constantly shutdown and re-established. Figure 8.2 shows the results using a log-scale on the Y-axis. We observe that remote interfaces consistently present more state transitions than local ones, reaching up to 3.5× more changes. While 75% of the local members in each IXP have no more than between 10 and 2 changes, depending on the IXP, the corresponding values for remote connections were between 16 and 5. Combined with the uptime hours previously presented, these numbers strongly indicate less connection stability for RP.

Local peers are greatly more reliable than RP. Finally, we compare the metric Reliability for the IXP interfaces, for periods of different durations, day and month. Members with higher reliability also present higher availability and less time between intermittent failures. Figures 8.3a and 8.3b show the results 1-day and 1-month, respectively. Even though reliability is directly related to the time interval, remote interfaces have lower re-

Figure 8.3: Level of reliability using the Reliability metric. Results show that remote interfaces are less reliable than local ones, reaching a difference up to 47.05%.



(a) Interface's reliability considering 1-day period.

(b) Interface's reliability considering 1-month period.

liability than local ones in both scenarios for all IXPs. However, the results vary considerably depending on the duration of the period considered. While in the 1-day analysis, there was little difference between local and remote (the median differences varied between 1.37% and 3.98%), in the 1-month, there are massive differences (reaching 47.05% and 33.85% for PTT-CE and LINX, respectively).

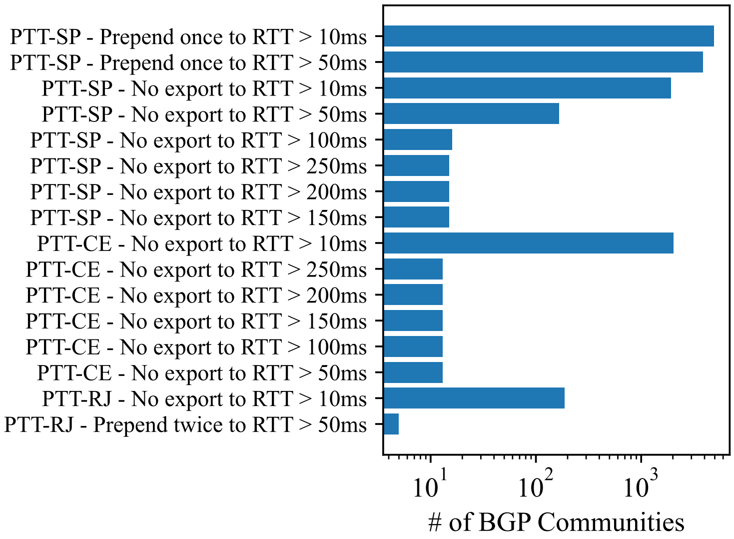
All results combined showed that, in addition to having less uptime and more state changes, remote interfaces also show less reliability over time and are less stable when compared to local ones in the analyzed IXPs. This confirms our expectations and shows there can be a substantial difference considering in some large IXPs.

8.2 Usage of RP BGP Communities for TE

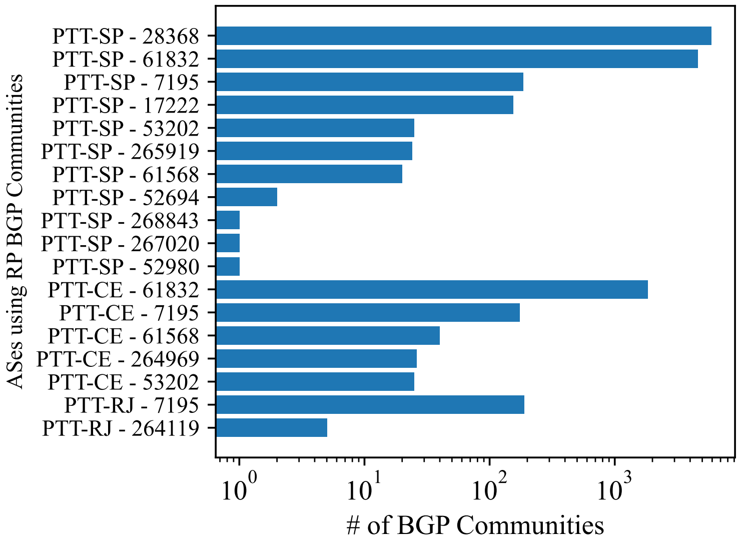
BGP communities have been used in IXP to tag routes and help with traffic engineering (MAZZOLA; MARCOS; BARCELLOS, 2022). Communities can also be used to perform TE in remote peering connections, such as tagging connections whose latency exceeds some threshold. IX.br started supporting BGP communities targetting remote networks in March 2022. This was accomplished with two changes: the RS in IXPs started tagging remote routes with *informational communities* to assist ASes in making traffic engineering decisions; and IXPs started supporting *action communities* added by the IXP members to impose their routing policies on their routes. We look at which communities networks add when sharing routes with the IXP.

Methodology. We analyze the usage of RP-related BGP communities for the three IXPs among our set that have support for such communities (PTT-SP, PTT-RJ, and PTT-CE).

Figure 8.4: Results for number of occurrences and ASes using RP BGP Communities. We observe that despite being used, these communities represente a negligible number in comparison with the other action communities seen on IXP routes.



(a) Number of occurrences for RP BGP communities.



(b) ASes that most use RP BGP communities at the three IXPs.

RP-related communities can be used to avoid route exports and to add different levels of prepending for ASes when matching certain conditions, such as RTT thresholds, packet loss thresholds, or announcement within some RIR region. The IX.br is responsible for measuring both latency and packet loss and marking up the routes with such characteristics.

Using the LG API in each IXP, we collected a routing data snapshot on 2023-01-17 from their primary IPv4 RS. The information captured for every route includes a prefix, next-hop address, AS-Path, and lists of BGP communities. We also fetched the RS configuration file containing the semantics of available informational and action BGP communities. To identify the communities defined by the IXPs, we build a *dictionary* using the LG API. We identified 96 communities related to RP on IX.br, out of the total 649 available. Their actions allow ASes to avoid route export and path-prepend once, twice or thrice to peers with RTT higher than 10, 50, 100, 150, 200 and 250 ms.

We process every route observed in the IXP BGP data for our analyses and compare the BGP communities in it with our community dictionary. We then counted the number of RP communities used in each route by each AS (i.e., if a route have two RP communities, we add two to the AS count).

Which type of BGP communities is more popular? First, we identify which specific RP-related communities are prevalent in the IXP route server member's routes (Figure 8.4a). On PTT-SP, the most common actions are *adding a single prepend* to members with RTT higher than 10ms and 50ms (4911 and 3860 instances, respectively). In contrast, the most common action in PTT-CE and PTT-RJ was to *avoid route export* for ASes with RTT higher than 10ms (2024 and 188 occurrences, respectively).

Are RP-related BGP Communities used? These fine-grained traffic engineering BGP communities for RP were introduced in March 2022 to all the 35 Brazilian IXPs. According to network operators from IX.br, this was prompted by a broad request from the network operation community. The general understanding has been that the communities would be widely used, and to confirm whether this was true or not, and to which extent, we examined the usage of RP BGP communities. Surprisingly, we discover that they are still not so widespread, at least when compared to the rest of BGP communities. For example, in PTT-SP, the fraction of RP-related action communities was under 0.27%.

We believe the low usage is because of two main reasons: the lack of knowledge about the potential impacts that remote peers may cause on performance; and the lack familiarity with the communities available at the IXPs. According to network operators, measurements to evaluate the peering performance are not proactively and regularly done to every individual peer. Instead, networks generally perform more detailed traffic engineering analysis when clients complain of reduced connectivity or performance. Unfortunately, the lack of periodic analysis and the opportunity to use RP-related BGP communities may hinder optimal peering decisions overall. As previously shown, using remote connections to deliver traffic can impose latency penalties compared to using a local route alternative for many cases. With respect to the lack of familiarity, given that different IXPs provide different services and deploy distinct BGP communities, network operators may still need to learn about the possibility of fine-grained traffic engineering.

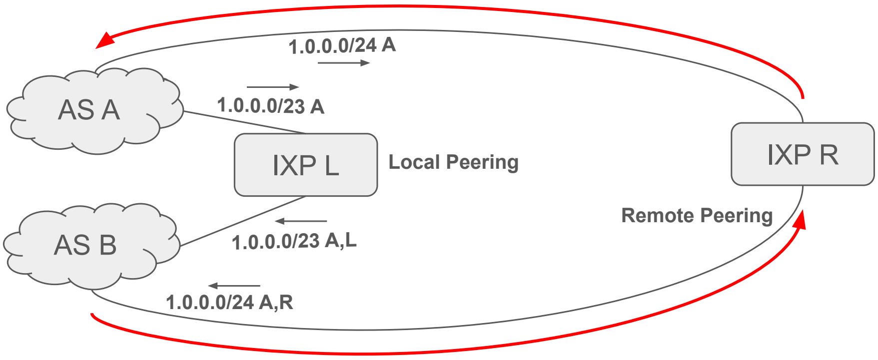
We discussed this with the operational community, and learned that often they deploy a single configuration model: a configuration is performed considering communities that are available at every IXP (e.g., avoid export to ASN, export only to ASN) and then replicated to all other IXPs .

Do ASes apply different RP communities in distinct IXPs to steer traffic? We now examine the numbers and characteristics of the ASes that are using RP BGP communities to their routes (Figure 8.4b). We found that only a tiny number of ASes use these communities, representing 0.5% (11 ASes' interfaces out of 2175) in PTT-SP and 1.1% (5 ASes' interfaces out of 471) in PTT-CE. When looking further at these ASes, we can highlight that one AS (AS7195) uses RP communities for its routes on the three IXPs examined, and three ASes (AS61832, AS53202, and AS61568) use them in two IXPs (PTT-SP and PTT-CE). AS7195 (i.e., EdgeUno) is a continental company focusing on IP transit, remote peering, and cloud hosting services. It adds the community to avoid route export to ASes with RTT $>$ 10 to, on average, 2.5% of their announced prefixes. They could represent prefixes needing better performance, which would not benefit from having long-distance networks redistributing them to their customers. The other three ASes, AS61832, AS53202, and AS61568, are ISPs connected to multiple IXPs worldwide. AS53202 and AS61568 add communities to avoid route export for ASes with RTT $>$ 50ms to 7.4% and 0.4%, respectively, of their announced prefixes in both IXPs. AS61832, in contrast, makes a distinguished use of these actions against remote peers: in PTT-CE, AS61832 uses the "No export to RTT $>$ 10ms" action on 74.4% of its prefixes, but on PTT-SP, it applies the community "Add one prepend to RTT $>$ 10ms" to all its announced routes. This approach inflates the AS-Path of the routes distributed in PTT-SP and forces traffic for some of its routes to go preferably via its connection in PTT-CE.

8.3 Routing Mispractices Caused by RP

If remote peering is not correctly configured by operators, ASes connected to multiple IXPs with different types of peering (i.e., a combination of local and remote connections) may introduce *trombone paths*. In this section, we investigate likely routing mispractice cases associated with RP misconfiguration. Trombone paths happen when two networks connected in a geographically-close location (e.g., London) exchange traffic via a distant peering facility (e.g., Chicago), causing poor routing performance and poor

Figure 8.5: Routing mispractice behavior when announcing at remote and local IXPs. AS A announces a more specific prefix to the IXP it connects via RP and a less specific prefix to the IXP it connects via Local Peering. When AS B receives both routes, it will give preference to the one via IXP R (red path) considering Internet routing best practices.

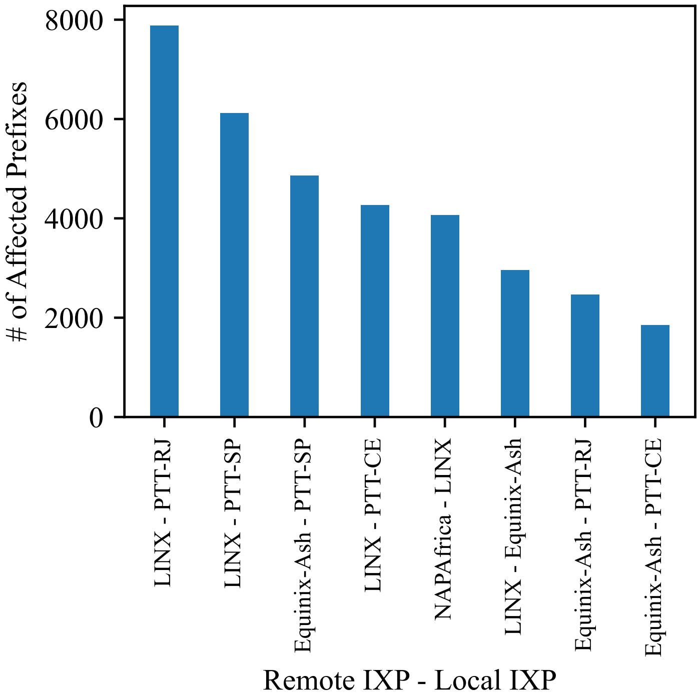


connectivity.

The IP Longest Prefix Match Routing indicate that more specific prefixes have a preference (as the intended destination for packets) over less specific prefixes. For example, a prefix 1.0.0.0/24 will be preferred over 1.0.0.0/16. In a peering scenario, the more specific prefixes (routes) announced by one AS in one IXP will be preferred for traffic exchange over less specific prefixes that can be announced at one transit provider. As reported by (ALMEIDA, 2019), an AS (in São Paulo) that has peering connections to a local IXP and a remote IXP can damage its peering connection performance if the operator does not correctly configure its prefix announcements in both IXPs. We exemplify this scenario in Figure 8.5, where an AS A connects to an IXP nearby (i.e. local peering on IXP X) and to a distant IXP (i.e. RP on IXP Y). By wrongly announcing more specific prefixes to the remote IXP instead of the local IXP, it can inadvertently steer its traffic exchange thousands of kilometres away (green path) from the optimal peering infrastructure (IXP X). For example, if an AS B also connected at IXP X via local peering and IXP Y via RP, the more specific prefix wrongly announced would lead the traffic exchange between them to go via the distant IXP Y, even though the better option (IXP X) is where both ASes are locally connected. In the rest of this section, we denote these likely problematic cases as *trombone prefixes*.

Methodology. We use the BGP data collected between 2022-08-01 and 2022-10-31 along with ping measurements performed from VPs in IXPs to corresponding IXP member interfaces (§ 5). The data allows us to identify the ASes that are present in multiple IXPs (in our set), by definition with one local connection and two or more,

Figure 8.6: IXPs involved in the identified prefixes with route announcements mispractices.



remote. For these ASes, we check the set of prefixes they announce in more than one IXP. The last step is to verify the of trombone prefix cases, that is, where a more specific prefix is announced at the remote IXP while a less specific prefix is announced at the local IXP.

We quantify the daily number of trombone prefixes and identify the ASes that announce them. We observed more than 480 trombone prefixes/day, with a peak of 1069/day. The daily number of ASes varied between 4 and 31, with an average of 17 ASes daily over the entire period; the 31 ASes were responsible for 1408 prefixes. We discover that both the number of trombone prefixes and ASes responsible for them can vary over time, indicating that they usually represent a *transient* routing problem that ASes fix in a matter of days. Looking further at the ASes that own such prefixes, we see just two ASes being accountable for 56.8% of all the overall trombone prefixes on average in the two-month period (AS52320 and AS262589, with 12616 and 8417 occurrences each). Not surprisingly, perhaps, we find that these two ASes were also the ones responsible for the peak occurrences (representing 78.6% of the prefixes on this day). Both ASes are continental networks offering IP transit, cloud and remote peering services for Latin American customers. We believe these prefixes represent routes from these ASes' customers that were improperly announced by the transit provider and subsequentially fixed.

IX.br and LINX are involved in most cases identified. Here we shift the analysis to the *IXPs* most commonly involved in the trombone prefixes. Figure 8.6 shows the most relevant combinations of IXPs for two months; we only display the IXP pairs with more than 1000 occurrences, which represent 93.0% of the total. The order is relevant, such that the first and second IXPs of the pair are, respectively, announcing the most specific prefix and the less specific prefix. We can observe in the figure that 50% of the cases are due to less specific prefixes being announced locally by networks connected at one of the IX.br IXPs (i.e., PTT-SP, PTT-RJ, and PTT-CE) and more specific versions of it announced remotely at LINX. Five ASes were responsible for more than 86.8% of the cases on the three IXPs: the previously mentioned AS52320 and AS262589, and AS28329, AS14840, and AS267613. The latter three are Brazilian networks with continental coverage, also offering IP transit and cloud services to Latin American clients.

Even though looking only at control plane data cannot guarantee that these cases represent actual routing problems, anedoctally, they often do. Our study reveals that there is a considerable number of occurrences happening over time. If these cases are not dealt with properly, they may affect the performance and connectivity of these networks and their customers and impose financial losses.

8.4 Summary

This chapter examines implications that RP can introduce to Internet Routing beyond latency. First, we evaluate an existing concern in the community about a higher instability in RP connection to IXP RSes. Our measurements show that this is indeed the case in all analyzed IXPs. Remote interfaces seemed less reliable, presenting differences to local peers, reaching up to 47.05% on a monthly analysis. Besides, the remote interfaces remain less time in a UP state between failures, with local interfaces staying up to 3.7× more active. Finally, remote interfaces showed up to 3.5× more state changes (from up to down) than local peers. The remote interfaces' stability and reliability differences to local ones are directly connected to the usage of shared ports and the additional underlying infrastructure combined with the geographic distance to IXPs.

Next, we collect routing data from RSes of IXPs with BGP communities explicitly created to perform TE on remote members. Our results show that while many modern IXPs started to offer specific BGP communities to filter route export or perform some action (e.g., prepend) to networks connected remotely, their usage is still not widespread.

Our analysis of the usage of these communities at three IXPs (PTT-SP, PTT-RJ, and PTT-CE) revealed that they are still negligible, representing less than 0.27% of all the action communities seen at the IXP routes.

Lastly, using BGP data collected and IXP membership data, we identify networks connected to multiple IXPs simultaneously and examine their route announcement patterns on different peering facilities. Our results show that undesirable trombone prefixes may be common due to route announcements mispractices. More than 60 distinct ASes announced most specific prefixes on the remote IXP instead of prioritizing their local IXP connection, which tends to prioritize traffic preference to go via the distant IXP. Among these ASes, we found over 37k prefixes with highly likely trombone paths, which could impair peering performance.

9 CONCLUSION

In the following sections, we summarize our findings and the main takeaways of this thesis and discuss the possibilities to continue this research.

9.1 Takeaways

IXPs are critical infrastructures that support ever-increasing data volumes and service requirements of modern Internet services. However, the recent growth of remote peering introduces new challenges for traffic engineering because peering may no longer keep local traffic local. This thesis shed light on the RP impacts on Internet performance and routing, and had the following key findings.

Inferring remote peering is still challenging. Using IXP ground truth and delay measurements, we showed that current state-of-the-art methodologies have limitations. We show that relying on public network data can result in a sizable fraction of unknown inferences for some IXPs, caused by public data being unavailable for some classes of networks. Compared to the European, American, and Asian IXPs evaluated in (NOMIKOS et al., 2018), reduced data availability in some regions, such as Latin America, limits the accuracy of remote peering inferences.

The route preferred by BGP is not always the lowest latency route. When investigating the use of remote routes in the BGP routing, we detected a high prevalence of prefixes announced both by remote and local peerings in four IXPs (LINX, AMS-IX, Eq-Ash, and Eq-Chi). We found that most remote routes for these prefixes had a shorter or equal AS path length compared to the available local routes and tended to be preferred by the peers of RouteViews collectors. Despite being shorter and indeed preferred, they were not necessarily the lowest latency route. For at least 61.2% of these prefixes in seven IXPs, the local route had lower latency compared to the geographically distant remote peering routes.

Remote routes are a reliable option to deliver traffic at IXPs. Some prefixes have only remote routes at IXPs, and ASes must choose between delivering their traffic via remote peering or a transit provider. Our measurements suggest that relying on remote routes can be an advantageous option for end-to-end latencies. In some scenarios (NAPAfrica and Eq-Chi), remote routes at the IXPs had considerably better latency results when compared to transit, showing latency improvements of at least 40ms for 81.4% of the measured prefixes, when the remote route was faster than transit. For the other six IXPs, we observed that the latency difference of using the remote route or the transit was no higher than 5ms for 78.1% of the measured prefixes.

The connection type or geographical distance does not directly impact latency variability for remote routes. A concern about remote peering growth at IXPs is that networks using a reseller or being geographically distant limits the original performance benefits of peering. Our measurements suggest that remote peering does not introduce additional latency variability to reach addresses in these prefixes. For 75% of the remote prefixes, we observed less than 10ms of latency variability for remote connections.

ASes using RP services to reach international IXPs can have considerable latency benefits. Our measurements suggest that using RP services and relying on them to reach traffic available on international IXPs can offer advantages to end-to-end latencies when compared to four Transit providers. For a fifth of the measured prefixes, the RP connection reached latency differences of up to 17 ms to the best Transit provider, keeping a comparable latency variability. However, the remaining 75% of the prefixes can suffer a latency penalty when using RP, indicating that operators should be conscient about the benefits and drawbacks of each connection alternative to reach their desired prefixes.

9.2 Future Research Directions

Despite the results reported in the thesis, promising opportunities for future research remain. In the following, we discuss the most prominent ones.

- **RP Inference Methodology.** Improving current methodologies is also crucial to promote further research on RP implications to performance and security. Our methodology used a 10ms latency threshold to infer geographical remote peering. While the threshold is conservative, it was adequate to identify networks connected far from IXPs. However, a deeper analysis of the impact of using different latency thresholds (e.g., 2ms and 5ms) is needed.

- **Route Selection.** Route selection is a complex problem faced by network operators, as there are many metrics that could affect traffic delivery performance. In this thesis we focused on investigating AS-Path length and latency (§6.1 and §6.2).

Analyzing routing by other metrics is challenging, because of the lack of reliable information in publicly available datasets, but can give interesting insights route selection regarding transit costs, economic decisions, and local preference.

- **Path Relevance.** Despite analyzing a considerable number of remote routes, one question that stands is the relevance of such paths, both in terms of destination popularity and traffic carried. Investigating this problem requires data protected by confidentiality terms and not publicly available (e.g., IXP traffic data) for all IXPs. Additionally, many IXPs do not have an automated way to measure traffic flowing through each announced route, and are able to only share aggregated traffic per AS.

- **Distributed IXPs.** Our analysis considered only IXP facilities within a single metropolitan area, avoiding wide-area peering infrastructures. In distributed IXPs, local members connected at facilities far from the IXP region could present very high latencies. An investigation on the performance implications of exchanging traffic in these peering infrastructures, considering RP, high and low latencies local peers can bring better understanding on the end-to-end latency to reach these different peering variations.

- **Additional IXPs.** Beyond the analysis we performed, we believe that considering additional IXPs would improve the community's understanding of remote peering in different parts of the world. With the ever-growing IXP ecosystem (Rosas, Israel, 2021), RP also tends to be a widely employed approach to connect to these peering infrastructures. Studying and analysing how RP is deployed in locations not widely explored, such as Africa, Asia and Oceania, can provide interesting insights into these peering ecosystems.

- **IPv6.** In this thesis, we focus on IPv4 exclusively, as the employed VPs did not have compatibility with IPv6. However, with the increasing adoption of IPv6 (NRO, 2023), we can expect more networks connecting to IXPs using IPv6. A further study to evaluate the RP deployment in IPv6 and whether they imply performance differences to ASes's interconnections is needed.

The previous lines of work will expand the relevance of our work in the measurement research community, as well as the technical community, and further our insights in new scenarios.

REFERENCES

AGER, B. et al. Anatomy of a large European IXP. In: **SIGCOMM**. [S.l.: s.n.], 2012.

ALI, Z. **Panel: Remote Peering, RIPE 65, Amsterdam, Netherlands**. 2012. <https://ripe65.ripe.net/presentations/196-RIPE65_EIXWG_ZA.pdf>.

ALMEIDA, L. **Good Practices for ISP connection to International IXPs**. 2019. In IX Forum 13 Brazil. Available from Internet: <https://youtu.be/bIfwuFVm0iI>.

AMS-IX. **The AMS-IX Partner Program**. 2018. <https://www.ams-ix.net/ams/partners>.

AMS-IX. **EasyAccess**. 2021. Available from Internet: <https://www.ams-ix.net/ams/service/easyaccess/>.

AMS-IX. **Pricing**. 2021. <https://www.ams-ix.net/ams/pricing>.

AMS-IX. **AMS-IX Infrastructure**. 2023. <https://www.ams-ix.net/ams/documentation/ams-ix-topology>.

AMS-IX. **Official AMS-IX Looking Glass**. 2023. <https://lg.ams-ix.net>.

ARK IPv4 prefix-probing data. 2023. <https://catalog.caida.org/dataset/ark_ipv4_prefix_probing>. Accessed: 2023-2-20.

ARNOLD, T. et al. Cloud provider connectivity in the flat internet. In: **Proceedings of the ACM Internet Measurement Conference**. New York, NY, USA: Association for Computing Machinery, 2020. (IMC '20), p. 230–246. ISBN 9781450381383. Available from Internet: <https://doi.org/10.1145/3419394.3423613>.

AUGUSTIN, B.; KRISHNAMURTHY, B.; WILLINGER, W. IXPs: Mapped? In: **IMC**. [S.l.: s.n.], 2009. p. 336–349. ISBN 978-1-60558-771-4.

BERTHOLDO, L. M. et al. Forecasting the impact of ixp outages using anycast. In: **TMA**. [S.l.: s.n.], 2021.

BIAN, R. et al. Towards passive analysis of anycast in global routing: Unintended impact of remote peering. **SIGCOMM CCR**, v. 49, n. 3, p. 18–25, nov. 2019. ISSN 0146-4833.

BICS. **BICS Launches Remote Peering to Offer Communications Providers an Affordable Solution for Network Expansion**. 2014. <https://prn.to/3lMetTQ>.

BIRGE-LEE, H. et al. Sico: Surgical interception attacks by manipulating bgp communities. In: **ACM CCS**. [S.l.: s.n.], 2019. p. 431–448. ISBN 9781450367479.

BRITO, S. H. B. et al. Dissecting the largest national ecosystem of public Internet eXchange points in Brazil. In: **PAM**. [S.l.: s.n.], 2016.

BöTTGER, T. et al. The elusive internet flattening: 10 years of ixp growth. 10 2018.

CARISIMO, E. et al. A first look at the Latin American IXPs. **SIGCOMM CCR**, 2020.

CASTRO, I. et al. Remote peering: More peering without Internet flattening. In: **CoNEXT**. [S.l.: s.n.], 2014. p. 185–198.

CGI.br. **IX.br reaches mark of 10 Tb/s of peak Internet traffic**. 2021. Available from Internet: <https://cgi.br/noticia/releases/ix-br-reaches-mark-of-10-tb-s-of-peak-internet-traffic/>.

CHANDRA, R.; TRAINA, P.; LI, T. **RFC 1997: BGP Communities Attribute**. 1996. RFC 1997. (Request for Comments, 1997).

CHATZIS, N. et al. On the benefits of using a large IXP as an Internet vantage point. In: **IMC**. [S.l.: s.n.], 2013.

CHATZIS, N. et al. There is more to IXPs than meets the eye. **SIGCOMM CCR**, 2013.

CHATZIS, N. et al. Quo vadis Open-IX? **SIGCOMM CCR**, v. 45, n. 1, p. 12–18, jan. 2015. ISSN 0146-4833.

CHIU, Y.-C. et al. Are we one hop away from a better internet? In: **Proceedings of the 2015 Internet Measurement Conference**. New York, NY, USA: Association for Computing Machinery, 2015. (IMC ’15), p. 523–529. ISBN 9781450338486. Available from Internet: <https://doi.org/10.1145/2815675.2815719>.

Cisco. **BGP Best Path Selection Algorithm**. 2023. <https://www.cisco.com/c/en/us/support/docs/ip/border-gateway-protocol-bgp/13753-25.html>.

Cloudflare. **Bandwidth Costs Around the World**. 2016. The Cloudflare Blog. Available from Internet: <https://blog.cloudflare.com/bandwidth-costs-around-the-world/>.

DE-CIX. **Connect to DE-CIX via a reseller**. 2021. Available from Internet: <https://www.de-cix.net/en/access/how-to-connect/connect-via-reseller>.

DE-CIX. **DE-CIX GlobePEER Remote Service.** 2021. Available from Internet: <https://www.de-cix.net/en/de-cix-service-world/globepeer-remote>.

DE-CIX. **DE-CIX Frankfurt statistics**. 2023. <https://www.de-cix.net/en/locations/germany/frankfurt/statistics>.

DIETZEL, C.; FELDMANN, A.; KING, T. Blackholing at IXPs: On the Effectiveness of DDoS Mitigation in the Wild. In: **PAM Conference**. [S.l.: s.n.], 2016.

DIETZEL, C.; WICHTLHUBER, M. Stellar: Network Attack Mitigation using Advanced Blackholing. In: **ACM CoNEXT**. [S.l.: s.n.], 2018.

DIMOPOULOS, G. et al. Measuring video qoe from encrypted traffic. In: **Proceedings of the 2016 Internet Measurement Conference**. New York, NY, USA: Association for Computing Machinery, 2016. (IMC ’16), p. 513–526. ISBN 9781450345262. Available from Internet: <https://doi.org/10.1145/2987443.2987459>.

DONNET, B.; BONAVENTURE, O. On bgp communities. **ACM CCR**, v. 38, n. 2, p. 55–59, mar 2008.

Dr Peering. **What are the economics of Internet Peering?** 2012. Available from Internet: <https://drpeering.net/FAQ/What-are-the-economics-of-peering.php>.

EQUINIX - Internet services company. 2023. <https://www.equinix.com/>.

Euro-IX. **Euro-IX. The IXP Database.** 2023. <https://ixpdb.euro-ix.net/en/>.

FANOU, R.; VALERA, F.; DHAMDHERE, A. Investigating the causes of congestion on the African IXP substrate. In: **IMC**. [S.l.: s.n.], 2017. p. 57–63. ISBN 978-1-4503-5118-8.

FARATIN, P. et al. The growing complexity of internet interconnection. **Communications & strategies**, n. 72, p. 51, 2008.

FORMOSO, A. et al. Deep diving into Africa's inter-country latencies. In: **INFOCOM**. [S.l.: s.n.], 2018. p. 2231–2239.

France-IX. **Angola Cables launches remote peering service at France-IX Marseille**. 2017. <https://www.franceix.net/en/events-and-news/presse/angola-cables-lance-un-service-de-peering-distance-france-ix-mar/>.

France-IX. **Interconnection with other IXPs**. 2021. Available from Internet: <https://www.franceix.net/en/solutions/interconnection/>.

GIOTSAS, V. et al. Detecting peering infrastructure outages in the wild. In: **SIGCOMM**. [S.l.: s.n.], 2017. p. 446–459. ISBN 978-1-4503-4653-5.

GIOTSAS, V. et al. Inferring complex AS relationships. In: **IMC**. [S.l.: s.n.], 2014. p. 23–30. ISBN 978-1-4503-3213-2.

GIOTSAS, V. et al. O peer, where art thou? uncovering remote peering interconnections at IXPs. **IEEE/ACM Transactions on Networking**, v. 29, n. 1, p. 1–16, 2021.

GIOTSAS, V. et al. Inferring BGP blackholing activity in the internet. In: **ACM IMC**. [S.l.: s.n.], 2017. p. 1–14.

GIOTSAS, V. et al. Mapping peering interconnections to a facility. In: **CoNEXT**. [S.l.: s.n.], 2015. p. 37:1–37:13. ISBN 978-1-4503-3412-9.

GIOTSAS, V. et al. Inferring multilateral peering. In: **CoNEXT**. [S.l.: s.n.], 2013. p. 247–258. ISBN 978-1-4503-2101-3.

GUPTA, A. et al. Peering at the Internet's frontier: A first look at ISP interconnectivity in Africa. In: **PAM**. [S.l.: s.n.], 2014. p. 204–213. ISBN 978-3-319-04918-2.

HOHLFELD, O. et al. A qoe perspective on sizing network buffers. In: **Proceedings of the 2014 Conference on Internet Measurement Conference**. New York, NY, USA: Association for Computing Machinery, 2014. (IMC '14), p. 333–346. ISBN 9781450332132. Available from Internet: <https://doi.org/10.1145/2663716.2663730>.

HOLTERBACH, T. et al. Swift: Predictive fast reroute. In: **Proceedings of the Conference of the ACM Special Interest Group on Data Communication**. New York, NY, USA: Association for Computing Machinery, 2017. (SIGCOMM '17), p. 460–473. ISBN 9781450346535. Available from Internet: <https://doi.org/10.1145/3098822.3098856>.

Hurricane Electric. **Internet Exchange Report**. 2021. <https://bgp.he.net/report/exchanges>.

IANA. **IANA Autonomous System (AS) Numbers.** 2020. <https://www.iana.org/assignments/as-numbers/as-numbers.xml>.

INTERCONNECTION and Traffic Exchange on the Internet: A Broadband Internet Technical Advisory Group Technical Working Group Report. 2014. <https://www.bitag.org/documents/Interconnection-and-Traffic-Exchange-on-the-Internet.pdf>.

Internet Society. **IXPs make the Internet faster and more affordable**. 2021. Available from Internet: <https://www.internetsociety.org/issues/ixps/>.

IORIO, M.; RISSO, F.; CASETTI, C. When latency matters: Measurements and lessons learned. **SIGCOMM Comput. Commun. Rev.**, Association for Computing Machinery, New York, NY, USA, v. 51, n. 4, p. 2–13, dec 2021. ISSN 0146-4833. Available from Internet: <https://doi.org/10.1145/3503954.3503956>.

IX.br. **IX.br members**. 2023. <http://ix.br/particip/sp>.

IX.br. **IX.br São Paulo statistics**. 2023. <https://ix.br/trafego/agregado/sp>.

IX.BR - Brazilian Public IXP Project. 2020. <https://ix.br/intro>.

JPIX. **JPIX: Japan Internet Exchange Co.** 2021. Available from Internet: <https://www.jpix.ad.jp/en/index.php>.

Juniper. **Understanding BGP Path Selection**. 2023. <https://www.juniper.net/documentation/us/en/software/junos/vpn-l2/bgp/topics/concept/routing-protocols-address-representation.html>.

KATZ-BASSETT, E. et al. Towards IP geolocation using delay and topology measurements. In: **IMC**. [S.l.: s.n.], 2006. p. 71–84. ISBN 1-59593-561-4.

KING, T. et al. **RFC 7999: BLACKHOLE Community**. 2016. RFC 7999. (Request for Comments, 7999).

KOTRONIS, V. et al. Investigating the potential of the Inter-IXP multigraph for the provisioning of guaranteed end-to-end services. In: **SIGMETRICS**. [S.l.: s.n.], 2015. p. 429–430. ISBN 978-1-4503-3486-0.

KOTRONIS, V. et al. Shortcuts through colocation facilities. In: **IMC**. [S.l.: s.n.], 2017. p. 470–476. ISBN 978-1-4503-5118-8.

KRENC, T.; BEVERLY, R.; SMARAGDAKIS, G. Keep your communities clean: Exploring the routing message impact of bgp communities. In: ____. **Proceedings of the 16th International Conference on Emerging Networking EXperiments and Technologies**. [S.l.: s.n.], 2020. p. 443–450. ISBN 9781450379489.

KRENC, T.; BEVERLY, R.; SMARAGDAKIS, G. As-level bgp community usage classification. In: **ACM IMC**. [S.l.: s.n.], 2021. p. 577–592.

KUSCHNIG, R.; KOFLER, I.; HELLWAGNER, H. Evaluation of http-based request-response streams for internet video streaming. In: **Proceedings of the Second Annual ACM Conference on Multimedia Systems**. New York, NY, USA: Association for Computing Machinery, 2011. (MMSys '11), p. 245–256. ISBN 9781450305181. Available from Internet: <https://doi.org/10.1145/1943552.1943585>.

LABOVITZ, C. et al. Internet inter-domain traffic. In: **Proceedings of the ACM SIGCOMM 2010 Conference**. New York, NY, USA: Association for Computing Machinery, 2010. (SIGCOMM '10), p. 75–86. ISBN 9781450302012. Available from Internet: <https://doi.org/10.1145/1851182.1851194>.

LEVY, M. J. **Peer locally, with local network on local IXs or not! (Alternate title: Remote peering is bad for content).** 2019. <https://www.globalpeeringforum.org/pastEvents/gpf14/presentations/Wed_2_MartinLevy_remote_peering_is_bad_for.pdf>.

LINX. **100GE Member Facing Port Milestone Reached**. 2021. Available from Internet: <https://www.linx.net/newsletter/linx-newsletter-august-2021/>.

LINX. **LINX Partner Exchanges**. 2021. Available from Internet: <https://www.linx.net/join-linx/connecting-remotely/ixp-partners/>.

LINX. **LINX Quick Facts**. 2021. Available from Internet: <https://quickfacts.linx.net/>.

LINX. **Members by IP/ASN**. 2021. Available from Internet: <https://portal.linx.net/members/members-ip-asn>.

LINX. **What is ConneXions**. 2021. Available from Internet: <https://www.linx.net/join-linx/connexions/>.

LINX. **LINX Route Server Looking Glass**. 2023. <http://alice-rs.linx.net>.

LIOTOU, E.; TSOLKAS, D.; PASSAS, N. A roadmap on qoe metrics and models. In: **2016 23rd International Conference on Telecommunications (ICT)**. [S.l.: s.n.], 2016. p. 1–5.

LU-CIX. **IX Partners - LU-CIX special partners**. 2021. Available from Internet: <https://www.lu-cix.lu/solutions/ix-partners/>.

LUCKIE, M. Scamper: A scalable and extensible packet prober for active measurement of the Internet. In: **IMC**. [S.l.: s.n.], 2010. p. 239–245. ISBN 9781450304832.

LUCKIE, M. et al. As relationships, customer cones, and validation. In: **IMC**. [S.l.: s.n.], 2013. p. 243–256. ISBN 978-1-4503-1953-9.

MAZZOLA, F.; MARCOS, P.; BARCELLOS, M. Light, camera, actions: Characterizing the usage of ixps' action bgp communities. In: **Proceedings of the 18th International Conference on Emerging Networking EXperiments and Technologies**. New York, NY, USA: Association for Computing Machinery, 2022. (CoNEXT '22), p. 196–203. ISBN 9781450395083. Available from Internet: <https://doi.org/10.1145/3555050.3569143>.

MENDES, J. L. **IX Forum 14: PeeringDB**. 2020. <https://forum.ix.br/files/apresentacao/arquivo/1020/IX_Forum_14_Sao_Paulo-peeringdb-20201204-V3.pdf>.

MILOLIDAKIS, A.; FONTUGNE, R.; DIMITROPOULOS, X. Detecting network disruptions at colocation facilities. In: **IEEE INFOCOM 2019 - IEEE Conference on Computer Communications**. [S.l.: s.n.], 2019. p. 2161–2169.

MüLLER, L. et al. Challenges in inferring spoofed traffic at ixps. In: **Proceedings of the 15th International Conference on Emerging Networking Experiments And Technologies**. New York, NY, USA: Association for Computing Machinery, 2019. (CoNEXT '19), p. 96–109. ISBN 9781450369985. Available from Internet: <https://doi.org/10.1145/3359989.3365422>.

NaMeX. **NaMeX: The Roman (Em)Peer**. 2021. Available from Internet: <https://www.namex.it/>.

NANOG. **Virtual or Remote Peering**. 2017. <https://mailman.nanog.org/pipermail/nanog/2017-August/091963.html>.

NAWROCKI, M. et al. Down the black hole: Dismantling operational practices of BGP blackholing at IXPS. In: **ACM IMC**. [S.l.: s.n.], 2019. p. 435–448.

Netnod. **Netnod Reach - Remote peering that grows with your network**. 2021. Available from Internet: <https://www.netnod.se/ix/reach>.

NIPPER, A. et al. **Remote Peering Panel - Providers**. 2018. In IX Forum 12 Brazil. Available from Internet: <https://www.youtube.com/watch?v=HQZY-DJNlxU&ab_channel=NICbrvideos>.

NMAP.org. **Nping: Measuring the Network**. 2023. <https://nmap.org/nping/>.

NOMIKOS, G. et al. O peer, where art thou? uncovering remote peering interconnections at IXPs. In: **IMC**. New York, NY, USA: [s.n.], 2018. p. 265–278. ISBN 9781450356190.

NORTON, W. B. **The Great Remote Peering Debate**. 2012. <http://drpeering.net/AskDrPeering/blog/articles/Ask_DrPeering/Entries/2012/9/18_The_Great_Remote_Peering_Debate.html>.

NORTON, W. B. **Internet Peering**. 2012. <http://drpeering.net/tools/HTML_IPP/chapters/ch04-0-Internet-Peering/ch04-0-Internet-Peering.html>.

NORTON, W. B. **Internet Transit**. 2012. <http://drpeering.net/tools/HTML_IPP/chapters/ch02-Internet-Transit/ch02-Internet-Transit.html>.

NRO. **The Number Resource Organization Extended Allocation and Assignment Reports**. 2023. <https://www.nro.net/about/rirs/statistics/>.

PANEL: Remote Peering, With a look at Resellers as well, 29th Euro-IX Forum, Krakow, Poland. 2016. <https://www.euro-ix.net/media/filer_public/ba/61/ba61bc32-f506-4ac7-b037-4dab8549c8a2/e-an-20161108-remote_peering.pdf>.

PANEL: What are the challenges of remote peering?, European Peering Forum 2016, Sofia, Bulgaria. 2016. <https://www.peering-forum.eu/agenda?year=2016>.

PCH. **Packet Clearing House, Internet Exchange Directory**. 2020. <https://www.pch.net/ixp/dir>.

PCH. **PCH. Daily Routing Snapshots**. 2021. <https://www.pch.net/resources/>.

PeeringDB. **PeeringDB. IXPs and colocation database.** 2023. <https://www.peeringdb.com>.

PUJOL, E. et al. Steering hyper-giants' traffic at scale. In: **Proceedings of the 15th International Conference on Emerging Networking Experiments And Technologies**. New York, NY, USA: Association for Computing Machinery, 2019. (CoNEXT '19), p. 82–95. ISBN 9781450369985. Available from Internet: <https://doi.org/10.1145/3359989.3365430>.

RAPAPORT, E. et al. Spillover today? predicting traffic overflows on private peering of major content providers. **IEEE Transactions on Network and Service Management**, v. 18, n. 4, p. 4169–4182, 2021.

REKHTER, Y.; HARES, S.; LI, T. **A Border Gateway Protocol 4 (BGP-4)**. 2006. RFC 4271. (Request for Comments, 4271).

RICHTER, P. et al. Peering at peerings: On the role of IXP route servers. In: **IMC**. [S.l.: s.n.], 2014. p. 31–44. ISBN 978-1-4503-3213-2.

Rosas, Israel. **Four New IXPs Take off in Central America and the Caribbean**. 2021. <https://www.internetsociety.org/blog/2021/06/four-new-ixps-take-off-in-central-america-and-the-caribbean/>.

Sandvine. **The Global Internet Phenomena Report 2023**. 2023. <https://www.sandvine.com/phenomena>.

SILVA, B. A. et al. Automatic Inference of BGP Location Communities. **ACM SIGMETRICS**, v. 6, n. 1, p. 3:1–3:23, feb. 2022.

Step, O. **BGP Community Guides**. 2021. Available from Internet: <https://onestep.net/communities/>.

STREIBELT, F. et al. Bgp communities: Even more worms in the routing can. In: **ACM IMC**. [S.l.: s.n.], 2018. p. 279–292.

Telecomdrive Bureau. **Epsilon Grows Remote Peering Presence at LINX**. 2020. <https://telecomdrive.com/epsilon-grows-remote-peering-presence-at-linx/>.

TELEHOUSE - Data center company. 2023. <https://www.telehouse.net/>.

TRAMMELL, B.; KÜHLEWIND, M. Revisiting the privacy implications of two-way Internet latency data. In: **PAM**. [S.l.: s.n.], 2018. p. 73–84. ISBN 978-3-319-76481-8.

TREVISAN, M. et al. Five years at the edge: Watching internet from the isp network. **IEEE/ACM Transactions on Networking**, v. 28, n. 2, p. 561–574, 2020.

WOHLFART, F. et al. Leveraging interconnections for performance: the serving infrastructure of a large cdn. In: **Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication**. [S.l.: s.n.], 2018. (ACM SIGCOMM '18), p. 206–220.

WonderNetwork. **Global Ping Statistics**. 2023. <https://wondernetwork.com/pings>.

APPENDIX A — RESUMO EXPANDIDO

Uma Análise Sobre os Impactos de Latência e Roteamento de Peering Remoto na Internet.

As infraestruturas de peering, como os Pontos de Troca de Tráfego da Internet (IXPs) e as instalações de *colocation*, são elementos cruciais da topologia da Internet. Eles concentram um grande número de redes e permitem a interconexão direta de ASes em uma infraestrutura centralizada. Em comparação com os provedores de trânsito, eles podem oferecer caminhos de internet mais curtos, melhor desempenho e custos de interconexão e operacional reduzidos para seus membros (Internet Society, 2021; CHATZIS et al., 2013b; AGER et al., 2012; AUGUSTIN; KRISHNAMURTHY; WILLINGER, 2009; Cloudflare, 2016; Dr Peering, 2012). Em fevereiro de 2023, havia mais de 800 IXPs implantados em todo o mundo (PeeringDB, 2023; Euro-IX, 2023; Hurricane Electric, 2021), com os maiores ultrapassando 1000 membros (LINX, 2021c; LINX, 2021a; IX.br, 2023a) e 10 Tbps de tráfego máximo (CGI.br, 2021; DE-CIX, 2023; LINX, 2021c; AMS-IX, 2023a).

A motivação original dos IXPs era manter o tráfego local ao ter ASes fisicamente presentes em uma instalação IXP. No entanto, os IXPs não interconectam apenas membros fisicamente presentes nas instalações do IXP. Peering Remoto (RP) - onde um AS não está fisicamente presente em uma instalação IXP e chega ao IXP por meio de um provedor de camada 2 - permite que os ASes ampliem sua oportunidades de peering com uma configuração mais rápida, sem hardware adicional e custos de instalação mais baixos em comparação ao peering local (DE-CIX, 2021a; AMS-IX, 2021b; CASTRO et al., 2014). Por exemplo, ASes de 80 países se conectam ao LINX remotamente (LINX, 2021c) em fevereiro de 2023. Para lidar com a demanda de peering, IXPs e revendedores de peering remoto expandiram suas ofertas (BICS, 2014; Telecomdrive Bureau, 2020; France-IX, 2017), com alguns IXPs tendo até 55 parceiros oficiais vendendo serviços de RP (IX.br, 2023a; AMS-IX, 2018; LINX, 2021d).

De acordo com o estado-da-arte, o RP não é mais apenas uma ideia, tornando-se uma prática significativamente comum na Internet. Para os IXPs mais relevantes do mundo, aproximadamente 40% de sua base de membros está conectada via RP (GIOTSAS et al., 2021). No entanto, apesar da popularidade desse método de conexão, pouco se sabe sobre suas implicações para a Internet. Além disso, a capacidade de conectar com membros remotos em IXPs adiciona complexidade às escolhas de engenharia de tráfego.

Recentemente, há um amplo debate público sobre o desempenho de RP. Entre as preocupações estão o fato de que serviços L2 podem introduzir ainda mais complexidade operacional, levar a ineficiências de roteamento, dificultar o monitoramento e tornar mais difícil avaliar o impacto de latência nas interconexões (LEVY, 2019; ALMEIDA, 2019; NIPPER et al., 2018; PANEL..., 2016b; PANEL..., 2016a; ALI, 2012; NORTON, 2012a). Os poucos estudos de pesquisa que examinam os impactos do RP na Internet encontraram uma ligação entre o RP e impactos negativos no desempenho anycast e na detecção de quedas na infraestrutura de peering (BIAN et al., 2019; GIOTSAS et al., 2017).

Nesta tese, foi realizada uma investigação extensiva de medição para melhorar a compreensão da implantação e dos impactos do remote peering (RP) na Internet. Acredita-se que os resultados obtidos são valiosos para o ecossistema de interconexão e para a comunidade, uma vez que contribuem com dados para a contínua discussão de desempenho de RP.

As contribuições são resumidas da seguinte forma:

1. Fornecemos uma análise dos desafios metodológicos e suas implicações na inferência de RP em IXPs usando a proposta mais avançada, estendida com uma análise abrangente de trabalhos anteriores.

2. Foram realizadas medições ativas em oito IXPs (incluindo seis dos dez maiores IXPs do mundo em termos de membros) para identificar e compreender a implantação de interfaces remotas e prefixos anunciados via conexões remotas, incluindo uma análise longitudinal para observar o crescimento do RP ao longo de um ano e meio. Parcerias com alguns dos IXPs e interações com operadores de rede ajudaram a validar nossas descobertas com dados de *ground-truth* e observações operacionais.

3. A partir das inferências de RP em IXPs, foi investigado se rotas anunciadas de forma remota em IXPs tendem a ser preferidas às rotas locais. Em seguida, foi realizado um estudo extensivo de medição para compreender os impactos de latência e variabilidade de latência ao usar diferentes métodos de interconexão (RP, peering local e Trânsito) para entregar tráfego a prefixos anunciados por membros conectados remotamente.

4. Foi avaliado o impacto de latência do RP ao usar uma conexão remota com IXPs internacionais e alcançar destinos de prefixos anunciados nessas infrastruturas de peering. Através de uma parceria com um grande revendedor de RP latino-americano,

foi utilizada a infraestrutura para simular um AS conectando-se a dois IXPs via RP. Em seguida, foram comparadas as latências para alcançar prefixos anunciados nos IXPs via RP e quatro provedores de trânsito.

5. Foram observados alguns dos efeitos do RP no roteamento da Internet. Pode-se notar que RP pode afetar consideravelmente a estabilidade da conexão dos membros do IXP, potencialmente introduzir desvios de roteamento causados por práticas inadequadas de anúncio de prefixo e como diferentes ASes usam comunidades BGP para realizar engenharia de tráfego em redes conectadas remotamente em IXPs.

APPENDIX B — ACHIEVEMENTS

B.1 Peer-reviewed publications

The development of this thesis has led to the publication of the following peer-reviewed/journal papers:

- **Journal**: ACM/IEEE Transaction on Networking (TON)
  - Title: Analyzing Remote Peering Deployment and its Implications for Internet Routing
  - Authors: **MAZZOLA F.**, CETTI A., MARCOS P., BARCELLOS M.
  - Qualis: A2
  - ISSN: 1558-2566
  - Status: Submitted
- **Conference**: 23rd International Conference Passive and Active Measurement (PAM 2022)
  - Title: On the Latency Impacts of Remote Peering
  - Main track (full-paper)
  - Authors: **MAZZOLA F.**, MARCOS P., CASTRO I., LUCKIE M., BARCELLOS M.
  - Qualis: A2
  - Date: March 28-30, 2022
  - Location: Virtual Conference
  - Digital Object Identifier (DOI): <https://doi.org/10.1007/978-3-030-98785-5_16>
- **Conference**: 19th ACM Internet Measurement Conference (IMC 2019)
  - Title: Are You Really There? Analyzing the Deployment of Remote Peering in the Brazilian IXP Ecosystem
  - Posters: Early Work, Tools, and Datasets Track

  - Authors: **MAZZOLA F.**, BARCELLOS M.
  - Qualis: A1
  - Date: March 21-23, 2019
  - Location: Amsterdam, Netherlands

Collaboration projects

In addition to the aforementioned main outcomes of this thesis, we further authored/coauthored some others studies on correlated large network measurement studies, Internet eXchange Points, and routing problems. These publications are listed next.

- **Conference**: The 18th International Conference on emerging Networking EXperiments and Technologies (CoNEXT '22)
  - Title: Light, Camera, Actions: characterizing the usage of IXPs' action BGP communities
  - Main track (short-paper)
  - Authors: **MAZZOLA F.**, MARCOS P., BARCELLOS M.
  - Qualis: A2
  - Date: December 6-9, 2022
  - Location: Rome, Italy
  - Digital Object Identifier (DOI): <https://doi.org/10.1145/3555050.3569143>

B.2 Invited Talks

- **MAZZOLA F.**, MARCOS P., CASTRO I., LUCKIE M., BARCELLOS M. On the Latency Impact of Remote Peering. In: 37th Euro-IX Forum, Edinburgh, United Kingdom, October 2022.
- **MAZZOLA F.**, MARCOS P., CASTRO I., LUCKIE M., BARCELLOS M. On the Latency Impact of Remote Peering. In: APNIC 54, Singapore, September 2022.
- MARCOS P., **MAZZOLA F.**, BARCELLOS M. On the Impacts of Remote Peering on Internet Routing. In: LACNIC 39, Merida, México, May 2023.
- BARCELLOS M. A Close Look At Remote Peering. In: BKNIX Peering Forum, Bangkok, Thailand, May 2023.

B.3 Research Projects

- FRIDA LACNIC: O impacto do peering remoto sobre o roteamento da Internet. 2022. Funding source: LACNIC.