APNIC 54

SINGAPORE

8 - 15 September 2022

On the Latency Impact of Remote Peering

Fabricio Mazzola, Pedro Marcos, Ignacio Castro, Matthew Luckie, and Marinho Barcellos

UFRGS UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

FURG UNIVERSIDADE FEDERAL DO RIO GRANDE

Queen Mary University of London

THE UNIVERSITY OF WAIKATO Te Whare Wānanga o Waikato

Traffic delivery is an increasingly complex aspect of the Internet today…

Large volumes of traffic and strict service requirements

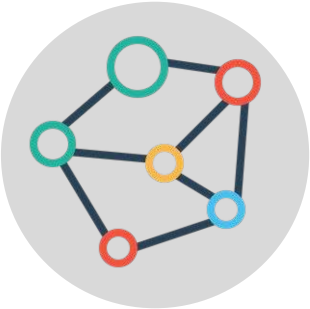
Traffic delivery is an increasingly complex aspect of the Internet today…

Large volumes of traffic and strict service requirements

Need for enhanced interconnection capacities and expanded footprint

Traffic delivery is an increasingly complex aspect of the Internet today…

Large volumes of traffic and strict service requirements



Need for enhanced interconnection capacities and expanded footprint

IXPs are key elements to help shorten paths and reduce interconnection cost

Originally, IXPs were designed to keep local traffic local by directly connecting geographically close ASes, but…

Remote Peering (RP) can simplify and lower the associated cost of peering...

Originally, IXPs were designed to keep local traffic local by directly connecting geographically close ASes, but…

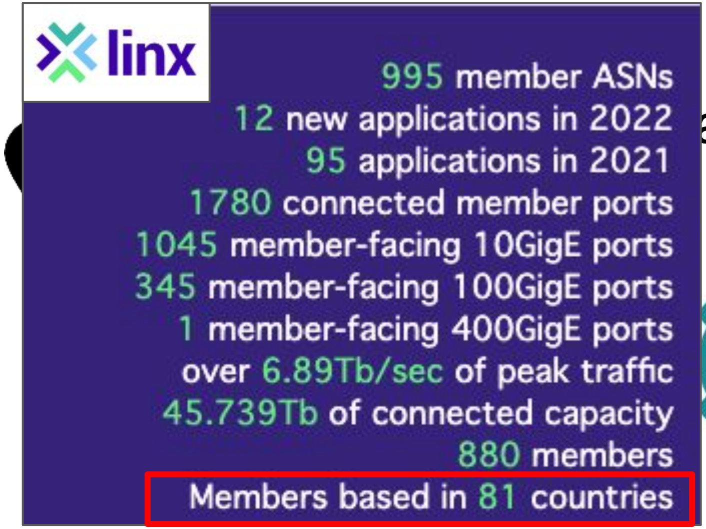
Remote Peering (RP) can simplify and lower the associated cost of peering...

Quicker and easier setup

Lower installation costs

No need for additional hardware

Originally, IXPs were designed to keep local traffic local by directly connecting geographically close ASes, but…



995 member ASNs
12 new applications in 2022
95 applications in 2021
1780 connected member ports
1045 member-facing 10GigE ports
345 member-facing 100GigE ports
1 member-facing 400GigE ports
over 6.89Tb/sec of peak traffic
45.739Tb of connected capacity
880 members
Members based in 81 countries

(RP) can simplify and

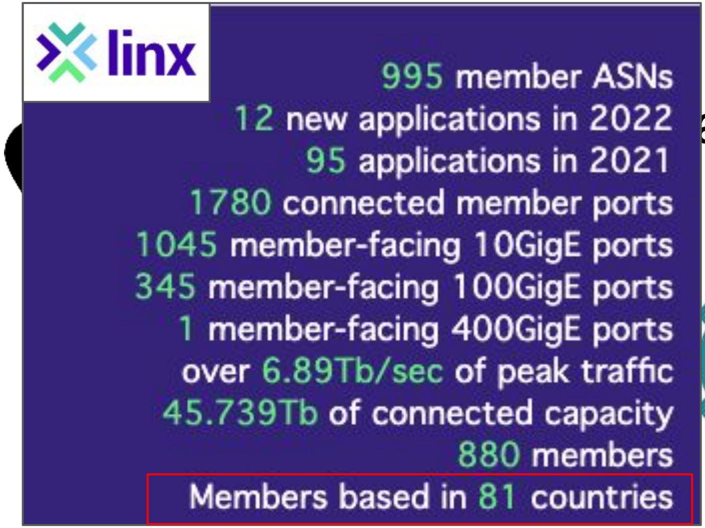
ated cost of peering...

easier setup

Lower installation costs

No need for additional hardware

Originally, IXPs were designed to keep local traffic local by directly connecting geographically close ASes, but…



(R

ate

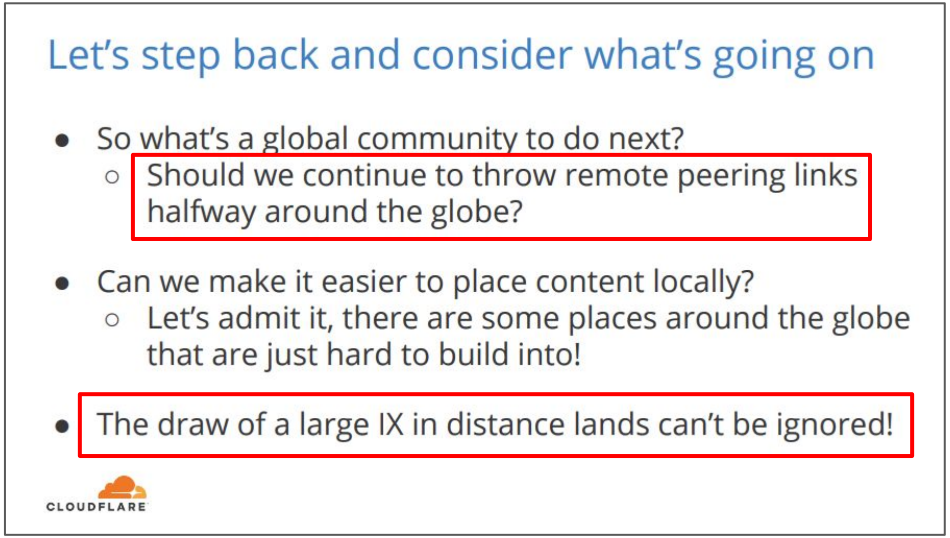
Ability to interconnect with remote members at IXPs **adds complexity to traffic engineering!!**

easier setup

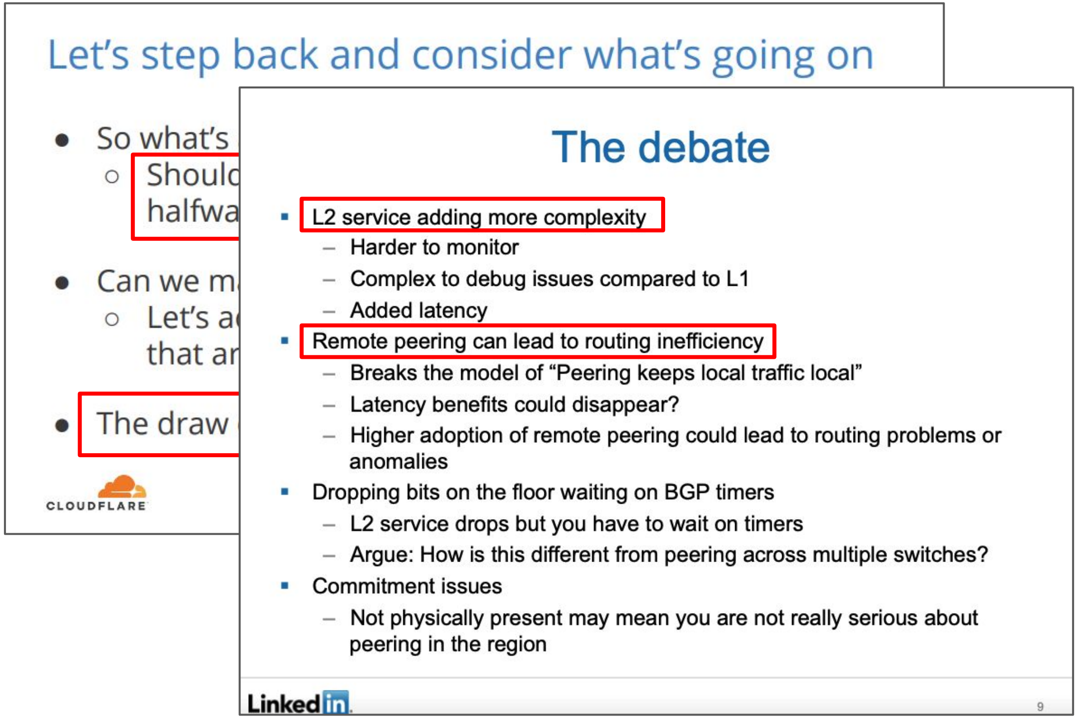
Lower installation costs

No need for additional hardware

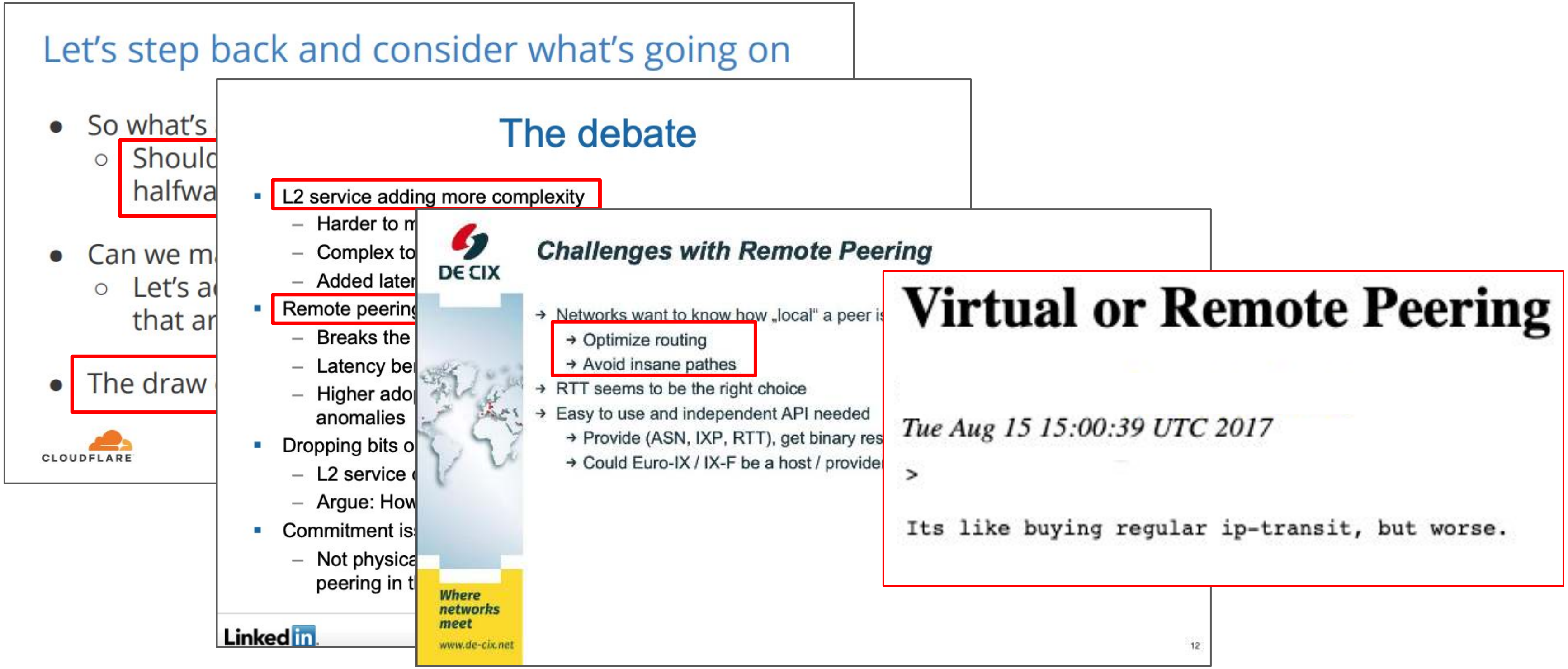
Despite being widely used, there are mixed feelings by the community about RP performance, which is currently data-poor…



Despite being widely used, there are mixed feelings by the community about RP performance, which is currently data-poor…



Despite being widely used, there are mixed feelings by the community about RP performance, which is currently data-poor…



Our goal is to investigate the latency impact of RP by contributing data to the performance discussion

Challenges of inferring RP

Our goal is to investigate the latency impact of RP by contributing data to the performance discussion

Challenges of inferring RP

RP's numbers on IXP routing data

Our goal is to investigate the latency impact of RP by contributing data to the performance discussion

Challenges of inferring RP

RP's numbers on IXP routing data

RP's latency impact on the data plane

Which IXPs and datasets do we analyze?

- **IXPs**: 8 IXPs (6 of the 10 largest IXPs by membership in the world)

IXP	Location	Observed Interfaces	BGP VPs LG	BGP VPs PCH
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	✗	✓

Which IXPs and datasets do we analyze?

- **IXPs**: 8 IXPs (6 of the 10 largest IXPs by membership in the world)

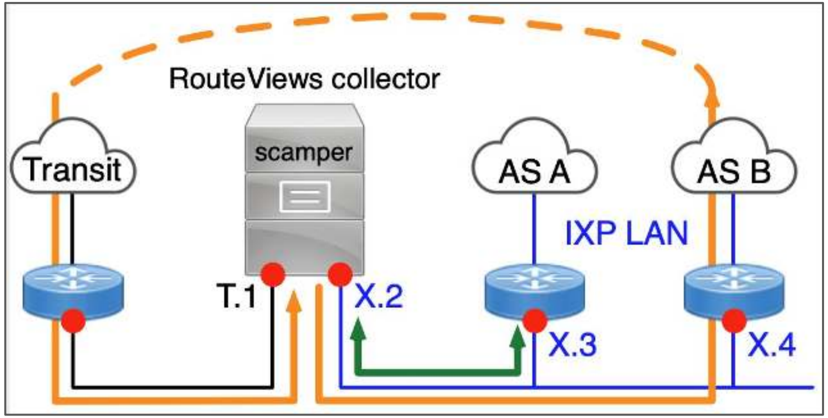
IXP	Location	Observed Interfaces	BGP VPs LG	BGP VPs PCH
PTT-SP	Sao Paulo, BR	2,169	✓	✗
LINX	London, UK	911	✓	✓
AMS-IX	Amsterdam, NL	907	✓	✓
NAPAfrica	Johannesburg, ZA	542	✗	✓
PTT-RJ	Rio de Janeiro, BR	462	✓	✗
PTT-CE	Fortaleza, BR	395	✓	✗
Eq-Ash	Ashburn, VA, US	365	✗	✓
Eq-Chi	Chicago, IL, US	259	✗	✓

- **Datasets and tools**:

PCH Packet Clearing House   LG   ROUTE VIEWS 6447   Scamper

How and from where do we perform our measurements?

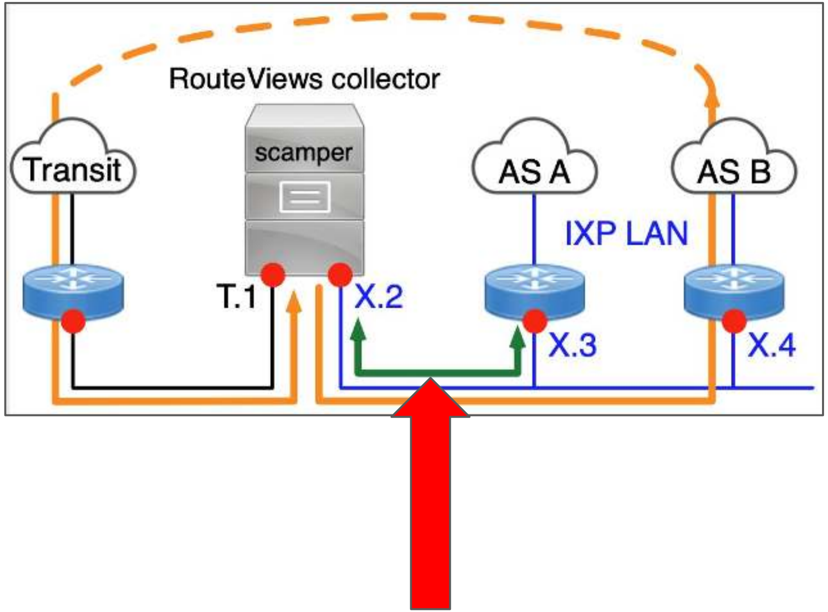
- Most RouteViews collectors are directly connected to an IXP LAN



How and from where do we perform our measurements?

- Most RouteViews collectors are directly connected to an IXP LAN

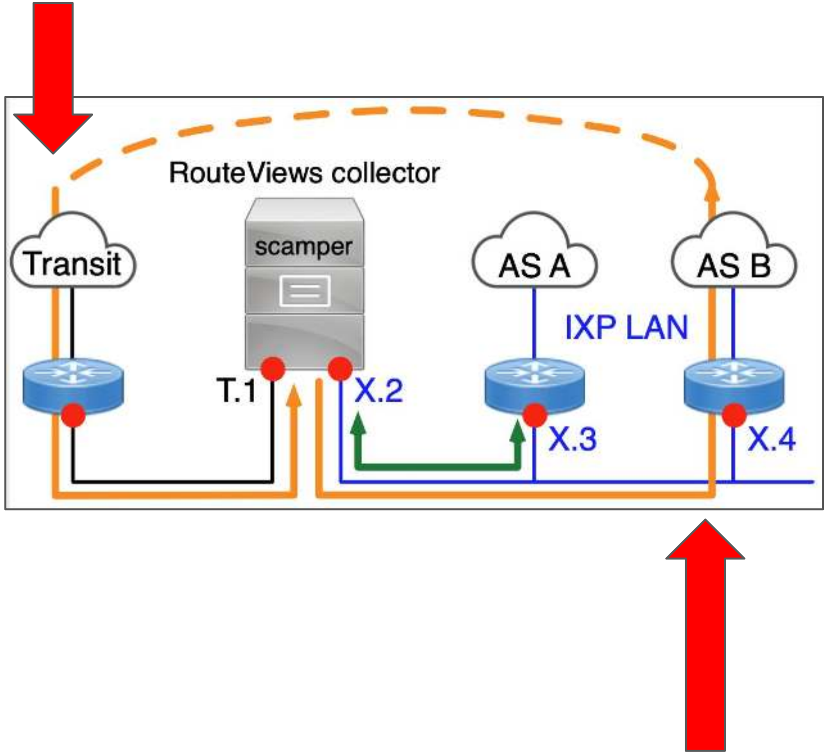
- Data plane measurements
    1. Latency to each IXP member's interface



How and from where do we perform our measurements?

- Most RouteViews collectors are directly connected to an IXP LAN

- Data plane measurements
    1. Latency to each IXP member's interface
    2. Path and latency to prefixes announced by remote IXP members using remote, local peering and transit connections



How do we infer RP at the IXPs?

Using the state-of-the-art method?

How do we infer RP at the IXPs?

Using the state-of-the-art method?

- Unfortunately, it provided insufficient inferences for some IXPs
- Low classification caused by lack of AS peering information on public data sources (PeeringDB)

How do we infer RP at the IXPs?

Using the state-of-the-art method?

- Unfortunately, it provided insufficient inferences for some IXPs
- Low classification caused by lack of AS peering information on public data sources (PeeringDB)

We analyse two RP perspectives:

- **Geographical RP** - Interfaces with min RTT higher than 10ms threshold
- **Reseller RP** - Interfaces shown in the Reseller ground truth (PTT-SP, LINX, PTT-RJ, and PTT-CE)

How widely deployed is RP? How does it reflect on routing data?

RP is widely deployed at IXPs, but that does not reflect on the announced remote routes/prefixes…

RP interfaces

Geographical RP usage varies in different IXPs:

Max of 35.2% of interfaces at PTT-CE (BR)

Less than 13.3% for other 5 IXPs

RP is widely deployed at IXPs, but that does not reflect on the announced remote routes/prefixes…

RP interfaces

Geographical RP usage varies in different IXPs:

Max of 35.2% of interfaces at PTT-CE (BR)

Less than 13.3% for other 5 IXPs

RP prefixes/routes

Remote peers announced proportionally fewer routes/prefixes than local peers

Highest difference for PTT-RJ: 13.2% IXP interfaces using Geographical RP announced just 3.0% of all routes

RP is widely deployed at IXPs, but that does not reflect on the announced remote routes/prefixes…

RP interfaces

Geographical RP usage varies in different IXPs:

Max of 35.2% of interfaces at PTT-CE (BR)

Less than 13.3% for other 5 IXPs

RP prefixes/routes

Remote peers announced proportionally fewer routes/prefixes than local peers

Highest difference for PTT-RJ: 13.2% IXP interfaces using Geographical RP announced just 3.0% of all routes

Multiple options

More than 71.4% of prefixes announced via Geographical RP at LINX, AMS-IX, Eq-Ash, and Eq-Chi also had a route announced by a local peer

RP is widely deployed at IXPs, but that does not reflect on the announced remote routes/prefixes…

RP interfaces

Geographical RP usage varies in different IXPs:

Max of 35.2% of interfaces at PTT-CE (BR)

Less than 13.3% for other 5 IXPs

RP prefixes/routes

Remote peers announced proportionally fewer routes/prefixes than local peers

Highest difference for PTT-RJ: 13.2% IXP interfaces using Geographical RP announced just 3.0% (67k/981k) of all routes

Multiple options

More than 71.4% of prefixes announced via Geographical RP at LINX, AMS-IX, Eq-Ash, and Eq-Chi also had a route announced by a local peer

Complexity to traffic engineering!

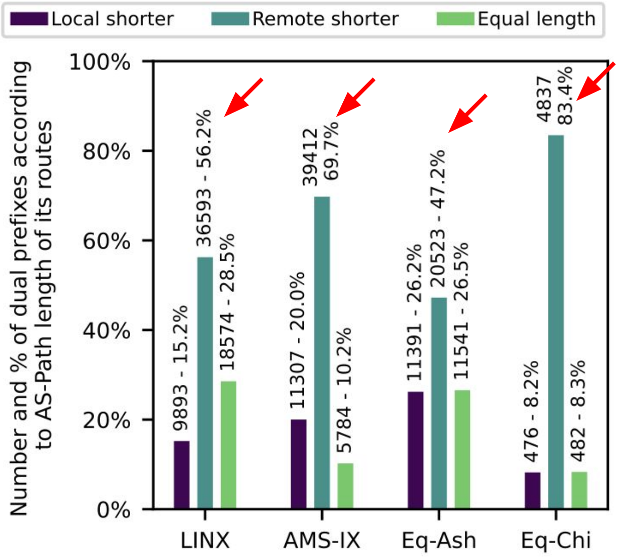
Which route has the best performance latency-wise?

For prefixes with multiple route alternatives at IXPs,
which routes are shorter and preferred?
Is there a latency penalty using a remote route?

For prefixes with multiple route alternatives at IXPs,
RP routes are shorter/preferred but tend to have higher latency than local routes…

For prefixes with multiple route alternatives at IXPs, RP routes are shorter/preferred but tend to have higher latency than local routes…

Routes according to their AS-Path length

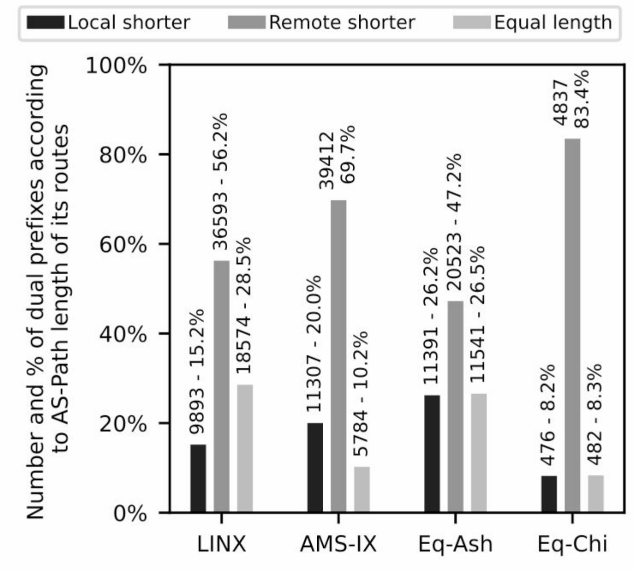


(a) Geographical RP.

Remote routes tend to be shorter and preferred than local ones

For prefixes with multiple route alternatives at IXPs, RP routes are shorter/preferred but tend to have higher latency than local routes…

Routes according to their AS-Path length



(a) Geographical RP.

Remote routes tend to be shorter and preferred than local ones

Latency measurements to prefixes with remote and local routes

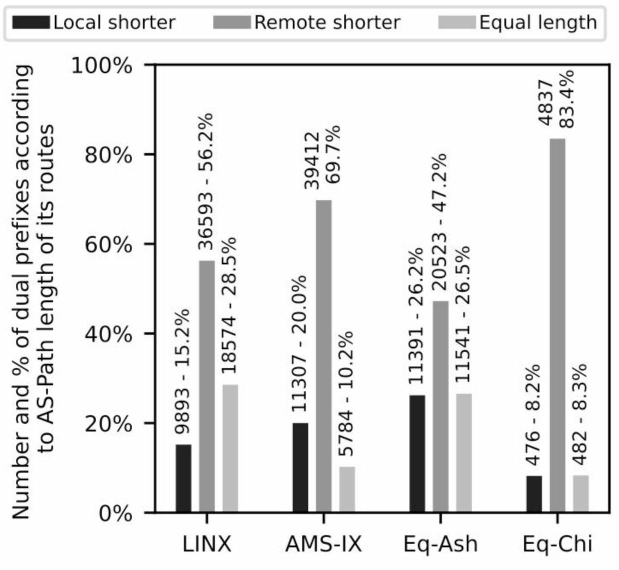
IXP	Geographical RP	
	Remote lower	Local lower
LINX	13,721 (33.0%)	27,903 (67.0%)
AMS-IX	6,644 (38.8%)	10,477 (61.2%)
Eq-Ash	2,230 (9.4%)	21,561 (90.6%)
Eq-Chi	830 (25.0%)	2,486 (75.0%)

For most prefixes, local routes were indeed better:

Latency benefit higher than 5ms for 44.7% or more prefixes

For prefixes with multiple route alternatives at IXPs, RP routes are preferred but tend to have higher latency than local routes…

Routes according to their AS-Path length



(a) Geographical RP.

Remote routes tend to be shorter and preferred than local ones

Latency measurements to prefixes with remote and local routes

IXP	Geographical RP	
	Remote lower	Local lower
LINX	13,721 (33.0%)	27,903 (67.0%)
AMS-IX	6,644 (38.8%)	10,477 (61.2%)
Eq-Ash	2,230 (9.4%)	21,561 (90.6%)
Eq-Chi	830 (25.0%)	2,486 (75.0%)

For most prefixes, local routes were indeed better:

Latency benefit higher than 5ms for 44.7% or more prefixes

Proper traffic engineering to decide which route to steer the traffic can lead to alternatives with better latencies!

And for prefixes with only remote routes at IXPs, is it better to rely on them or use a transit provider to deliver traffic?

Prefixes with only remote routes at IXPs are better latency-wise than transit, but not by much…

Prefixes with only remote routes at IXPs are better latency-wise than transit, but not by much…

Most RP routes had lower latency than Transit (>57.6% of prefixes)!

But latency diff was below 5ms for more than 78.1% prefixes at 6 IXPs

IXP	Geographical RP latency	
	Remote lower	Transit lower
PTT-SP	5,657 (72.0%)	2,205 (28.0%)
LINX	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	1,113 (59.6%)	754 (40.4%)
PTT-CE	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%)

Prefixes with only remote routes at IXPs are better latency-wise than transit, but not by much…

Most RP routes had lower latency than Transit (>57.6% of prefixes)!

But latency diff was below 5ms for more than 78.1% prefixes at 6 IXPs

IXP	Geographical RP latency	
	Remote lower	Transit lower
PTT-SP	5,657 (72.0%)	2,205 (28.0%)
LINX	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	1,113 (59.6%)	754 (40.4%)
PTT-CE	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%)

For some IXPs, RP can have substantial advantage!

In NAPAfrica, 81.4% of remote routes were at least 40ms better than Transit

Prefixes with only remote routes at IXPs are better latency-wise than transit, but not by much…

Most RP routes had lower latency than Transit (>57.6% of prefixes)!

But latency diff was below 5ms for more than 78.1% prefixes at 6 IXPs

IXP	Geographical RP latency	
	Remote lower	Transit lower
PTT-SP	5,657 (72.0%)	2,205 (28.0%)
LINX	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	1,113 (59.6%)	754 (40.4%)
PTT-CE	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%)

For some IXPs, RP can have substantial advantage!

In NAPAfrica, 81.4% of remote routes were at least 40ms better than Transit

Relying on remote routes at IXPs can be a beneficial option for end-to-end latencies!

What did we learn in our investigations?

Inferring RP is still challenging and state-of-the-art methods have limitations

What did we learn in our investigations?

Inferring RP is still challenging and state-of-the-art methods have limitations

Despite being shorter and preferred, RP routes had higher latencies for a considerable % of prefixes

What did we learn in our investigations?

Inferring RP is still challenging and state-of-the-art methods have limitations

Despite being shorter and preferred, RP routes had higher latencies for a considerable % of prefixes

In most cases, remote routes are better latency-wise than transit, but not by much

What are we looking forward to next?

Extend our analysis to more IXPs, especially in areas we have not investigated yet (i.e., APNIC)

What are we looking forward to next?

Extend our analysis to more IXPs, especially in areas we have not investigated yet (i.e., APNIC)

Establish partnership with ASes using RP to investigate what is the RP impact on their networks and their customers

What are we looking forward to next?

Extend our analysis to more IXPs, especially in areas we have not investigated yet (i.e., APNIC)

Establish partnership with ASes using RP to investigate what is the RP impact on their networks and their customers

Understand if the findings for IPv4 reflect on IPv6 prefixes

APNIC 54
SINGAPORE
8 - 15 September 2022

Thanks for the attention!

If you are interested in contributing to our work
or have any questions, please send us an email!

Fabricio Mazzola
fmmazzola@inf.ufrgs.br

Joint work with:

Pedro Marcos

Ignacio Castro

Matthew Luckie

Marinho Barcellos

For further details, have a look on our PAM 2022 paper!
https://link.springer.com/chapter/10.1007/978-3-030-98785-5_16

UFRGS UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FURG UNIVERSIDADE FEDERAL DO RIO GRANDE
Queen Mary University of London
THE UNIVERSITY OF WAIKATO Te Whare Wānanga o Waikato