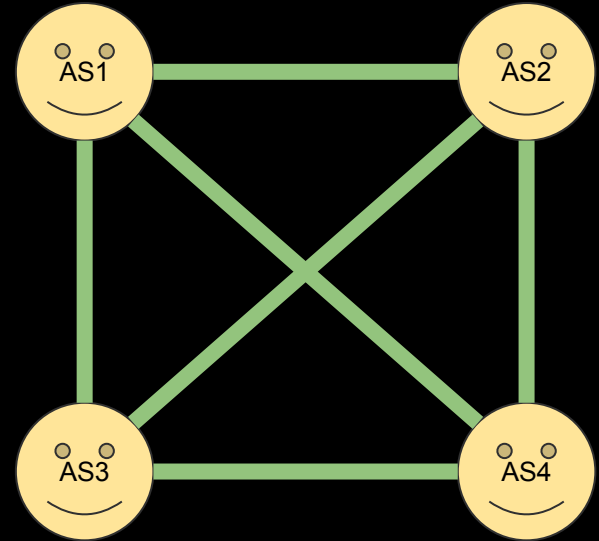


How far can you go with IX Route Servers only?

Ben Cartwright-Cox
BGP.Tools / Port 179 LTD

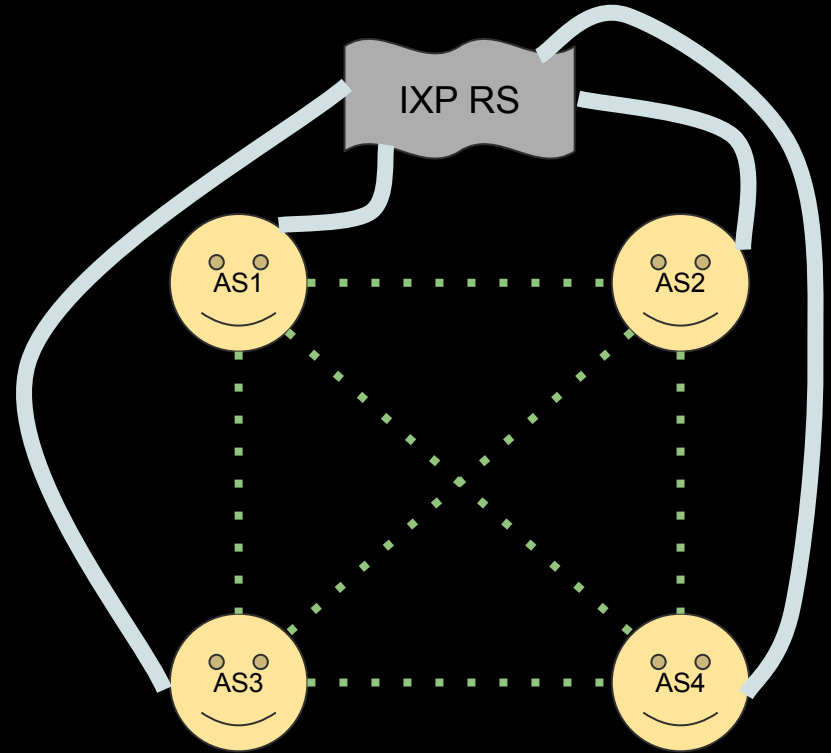
IX Route servers

- Attempts to solve the problem of peers needing to create BGP sessions with nearly every member of a exchange they want to exchange traffic with
- This is bad because networks are lazy/busy, and may not setup sessions when asked



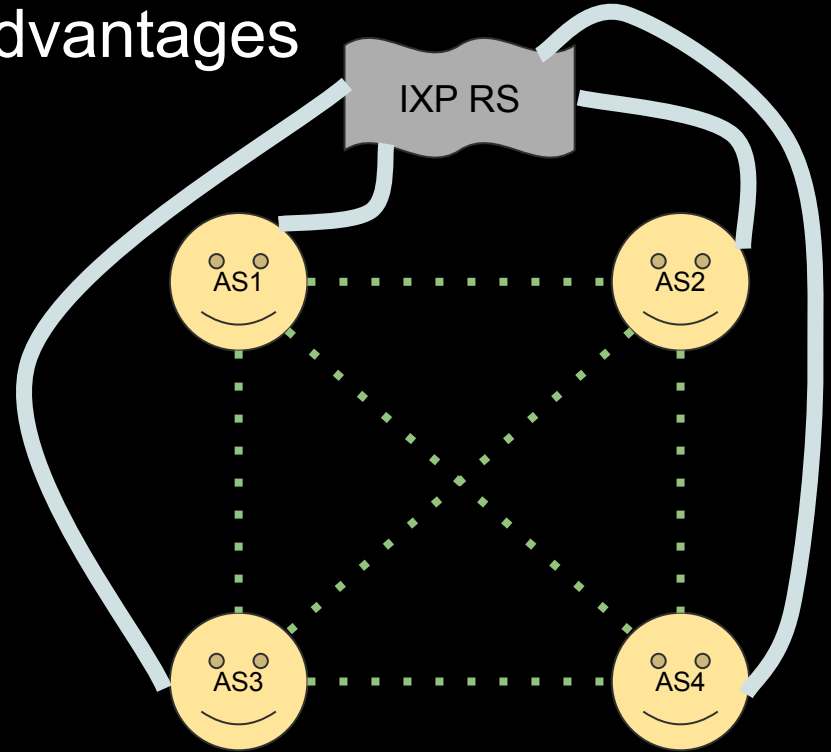
IX Route Servers (RS)

- Instead, everyone peers with the RS, and it distributes the routes sent to it by other members to everyone.
- Since everyone is on the same subnet/layer 2. The BGP next hop is not changed.
- The RS can control terabits of traffic with a 100mbps port.



IX Route Servers (RS) Bonus Advantages

- Your average person (despite what they say publicly) does not have a good/secure BGP peer configuration
- Modern RS are *far safer* to peer on than bi-lat peering, due to better IRR/RPKI/Sanity/PeerLock automations
- Yes you may have the magic config, but most of the IX is importing almost anything they are sent



IX Route Servers (RS) Bonus Advantages

- Your average person (despite what they say publicly) does not have a good/secure BGP peer configuration
- Modern RS are *far safer* to peer on than bi-lat peering, due to better IRR/RPKI/Sanity/PeerLock automations
- Yes you may have the magic config, but most of the IX is importing almost anything they are sent

AS-DECIX

Database RIPE

Full Name AS-DECIX

Overview

Reverse

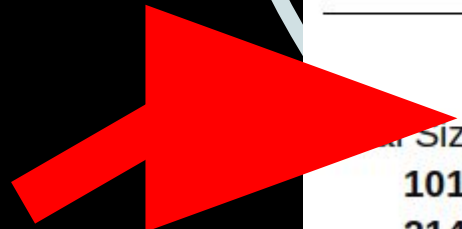
Raw

Size

101625 ASNs

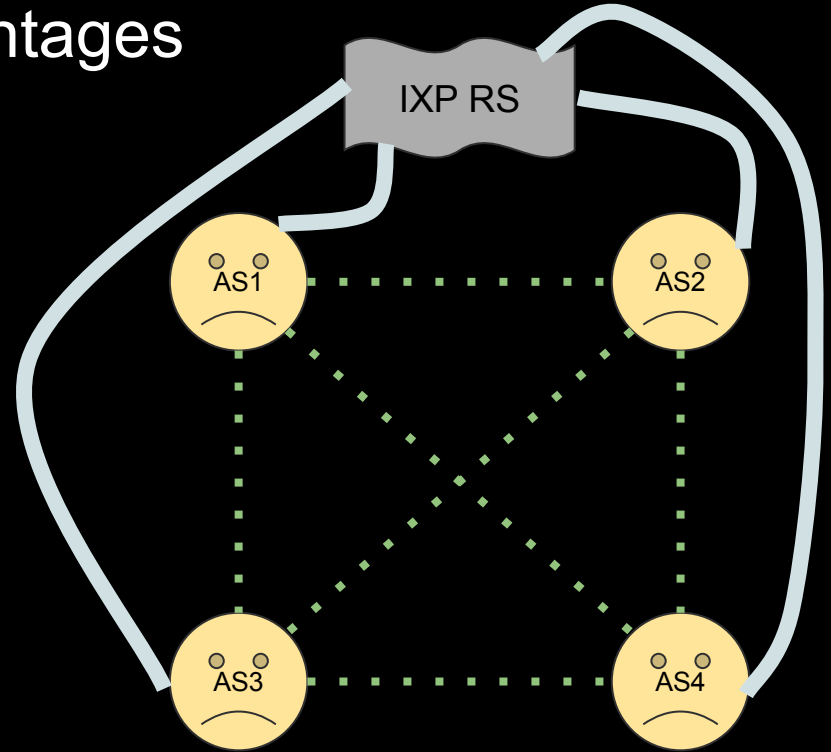
2147500 v4 Prefixes









839717 v6 Prefixes



IX Route Servers (RS) Disadvantages

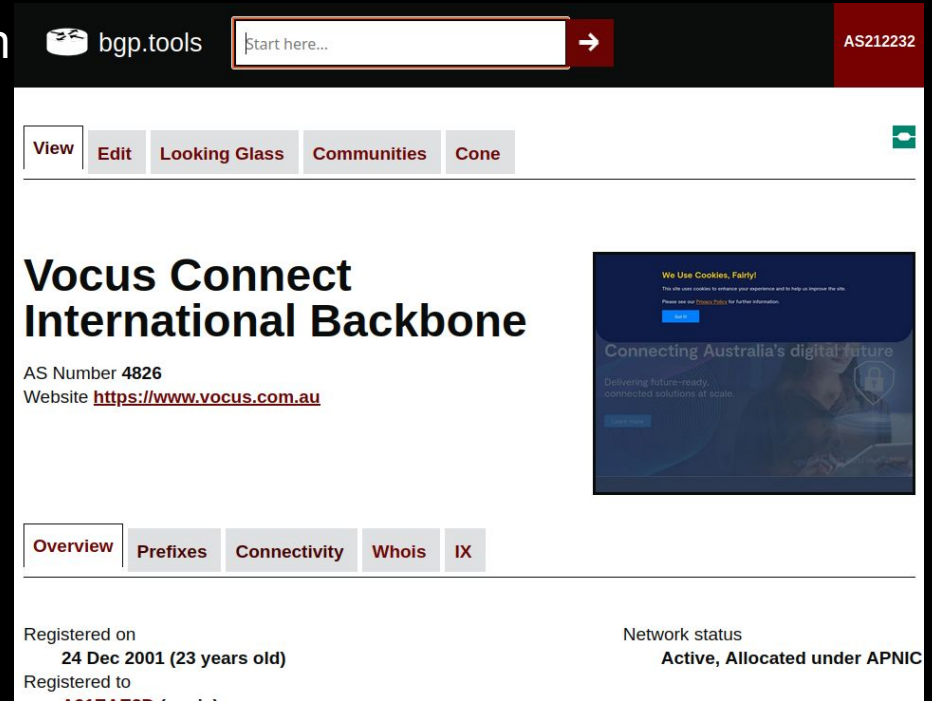
- **Some** networks import/export everything to the Route Server
- **Some** networks just import from the route server, but don't send their routes to them
 - There are some good reasons to do this, some networks are very sensitive poor routing and Route Servers are considered high risk
- **Others** will export/send routes, but not import anything












  	AS140627	OneQode	202.77.88.36	2001:df0:680:5::24	100 gbps
  	AS9280	Servers Australia Pty Ltd	202.77.88.37	2001:df0:680:5::25	40 gbps
  	AS133120	Hosted Network Pty. Ltd.	202.77.88.38	2001:df0:680:5::26	10 gbps
  	AS54113	Fastly, Inc.	202.77.88.39	2001:df0:680:5::2a	200 gbps
  	AS9500	One New Zealand Group Limited	202.77.88.40	2001:df0:680:5::28	100 gbps
  	AS45437	Real World - The Core	202.77.88.41	2001:df0:680:5::29	10 gbps
  	AS54113	Fastly, Inc.	202.77.88.42	2001:df0:680:5::57	200 gbps
  	AS19679	Dropbox, Inc.	202.77.88.43	2001:df0:680:5::2b	100 gbps

Quick tl;dr of what bgp.tools is

























- General internet routing information site
- Runs its own BGP Route Collector that you should feed!
- Has a large IX presence for route collection
- Can also offer rapid BGP (and other BGP related things) network monitoring



The screenshot shows the bgp.tools website interface. At the top, there is a search bar with the text "Start here..." and a red arrow button. To the right of the search bar, the AS number "AS212232" is displayed. Below the search bar, there are navigation tabs: "View", "Edit", "Looking Glass", "Communities", and "Cone". The main content area features the title "Vocus Connect International Backbone" in large, bold black text. Below the title, it lists "AS Number 4826" and "Website <https://www.vocus.com.au>". To the right of the main content, there is a blue banner with the text "We Use Cookies, Fairly!" and "Connecting Australia's digital future". Below the banner, there are more navigation tabs: "Overview", "Prefixes", "Connectivity", "Whois", and "IX". At the bottom of the page, there are two sections: "Registered on 24 Dec 2001 (23 years old)" and "Registered to" (partially obscured), and "Network status Active, Allocated under APNIC".












  	<u>AS140627</u>	OneQode	202.77.88.36	2001:df0:680:5::24	100 gbps
  	<u>AS9280</u>	Servers Australia Pty Ltd	202.77.88.37	2001:df0:680:5::25	40 gbps
  	<u>AS133120</u>	Hosted Network Pty. Ltd.	202.77.88.38	2001:df0:680:5::26	10 gbps
  	<u>AS54113</u>	Fastly, Inc.	202.77.88.39	2001:df0:680:5::2a	200 gbps
  	<u>AS9500</u>	One New Zealand Group Limited	202.77.88.40	2001:df0:680:5::28	100 gbps
  	<u>AS45437</u>	Real World - The Core	202.77.88.41	2001:df0:680:5::29	10 gbps
  	<u>AS54113</u>	Fastly, Inc.	202.77.88.42	2001:df0:680:5::57	200 gbps
  	<u>AS19679</u>	Dropbox, Inc.	202.77.88.43	2001:df0:680:5::2b	100 gbps

[1] <https://bgp.tools/ixp/EdgeIX+-+Sydney>

  	AS140627	OneQode	202.77.88.36	2001:df0:680:5::24	100 gbps
  	AS9280	Servers Australia P		:680:5::25	40 gbps
  	AS133120	Unstated Network Pt		:680:5::26	10 gbps
  	13	Fastly, Inc.	202.77.88.39	2001:df0:680:5::2a	200 gbps
  	AS9500	One New Zealand Group Limited	202.77.88.40	2001:df0:680:5::28	100 gbps
  	AS45437	Real World - The Core	202.77.88.41	2001:df0:680:5::29	10 gbps
  	AS54113	Fastly, Inc.	202.77.88.42	2001:df0:680:5::57	200 gbps
  	AS19679	Dropbox, Inc.	202.77.88.43	2001:df0:680:5::2b	100 gbps

























They are sending routes to the RS



  	AS140627	OneQode	202.77.88.36	2001:df0:680:5::24	100 gbps
  	AS9280	Servers Australia	202.77.88.37	2001:df0:680:5::25	40 gbps
  	AS133120	Hosted Network Pty. Ltd.	202.77.88.38	2001:df0:680:5::26	10 gbps
  	AS54113	Fastly, Inc.	202.77.88.39	2001:df0:680:5::2a	200 gbps
  	AS9500	One New Zealand Group Limited	202.77.88.40	2001:df0:680:5::28	100 gbps
  	AS45437	Real World - The Core	202.77.88.41	2001:df0:680:5::29	10 gbps
  	AS140613	Fastly, Inc.	202.77.88.42	2001:df0:680:5::57	200 gbps
  	AS19679	Dropbox, Inc.	202.77.88.43	2001:df0:680:5::2b	100 gbps

























They are not



  	AS140627	OneQode	202.77.88.36	2001:df0:680:5::24	100 gbps
  	AS9280	Servers Australia P		:680:5::25	40 gbps
  	AS133120	Hosted Network Pt		:680:5::26	10 gbps
  	AS54113	Fastly, Inc.	202.77.88.39	2001:df0:680:5::2a	200 gbps
  	AS9500	The New Zealand Group Limited	202.77.88.40	2001:df0:680:5::28	100 gbps
  	AS45437	Real World - The Core	202.77.88.41	2001:df0:680:5::29	10 gbps
  	AS54113	Fastly, Inc.	202.77.88.42	2001:df0:680:5::57	200 gbps
  	AS19679	Dropbox, Inc.	202.77.88.43	2001:df0:680:5::2b	100 gbps

They *might* be importing the routes from the RS!



			AS140627	OneQode	202.77.88.36	2001:df0:680:5::24	100 gbps
			AS9280	Servers Australia P		:680:5::25	40 gbps
			AS133120	Hosted Network Pt		:680:5::26	10 gbps
			AS13	Fastly, Inc.	202.77.88.39	2001:df0:680:5::2a	200 gbps
			AS9500	One New Zealand Group Limited	202.77.88.40	2001:df0:680:5::28	100 gbps
			AS45437	Real World - The Core	202.77.88.41	2001:df0:680:5::29	10 gbps
			AS54113	Fastly, Inc.	202.77.88.42	2001:df0:680:5::57	200 gbps
			AS19679	Dropbox, Inc.	202.77.88.43	2001:df0:680:5::2b	100 gbps

They *might not* be importing the routes from the RS!



How good is the reach for RS's for outbound heavy nets?

- This is the most simple thing to answer, because all you need to do is get all of the RIB's of all of the IXs RS's you are on (bgp.tools is currently on 113~ IXs), and perform a total unique route count vs the full table

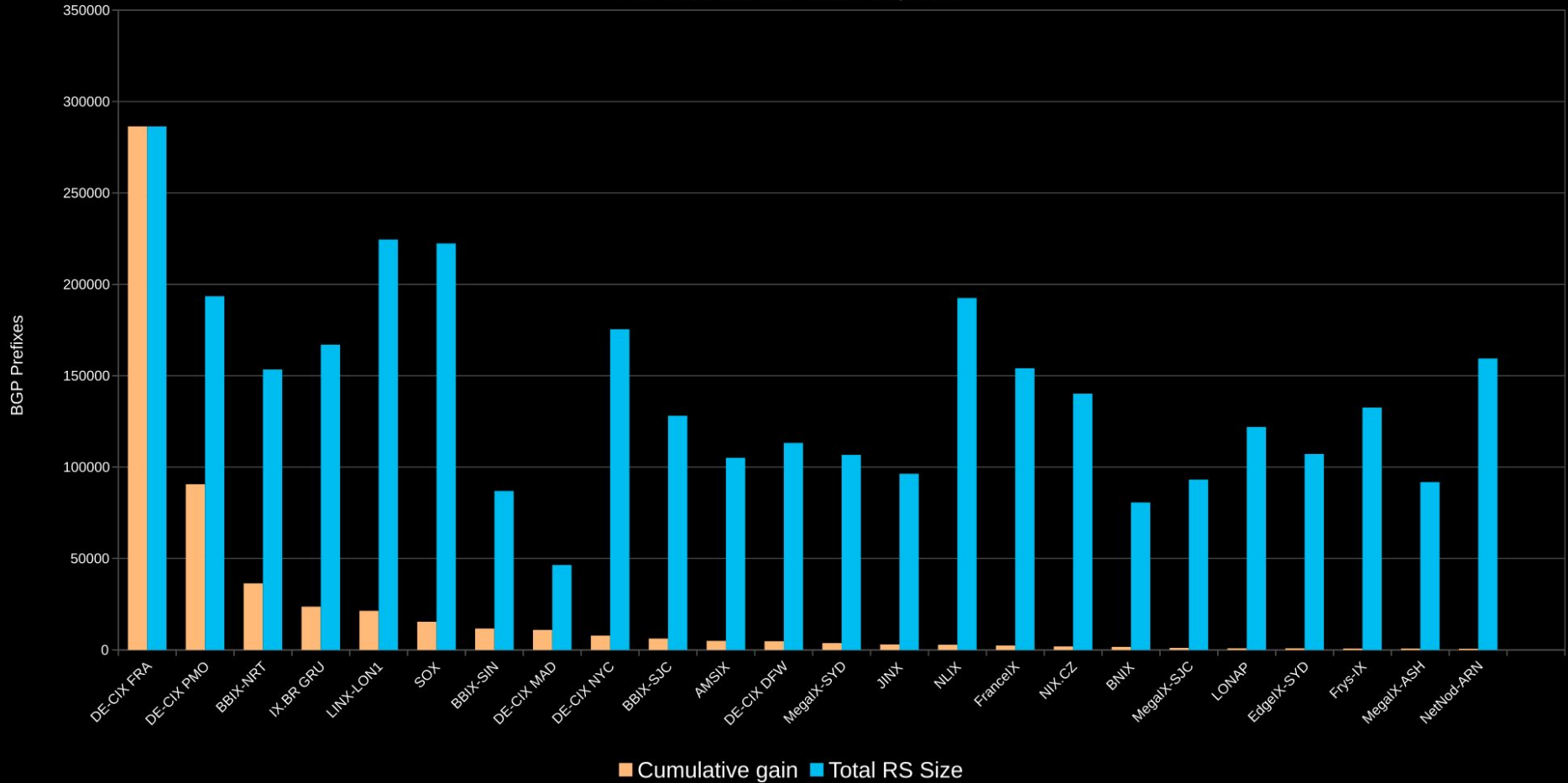
- A full table as of the time of writing is
 - IPv4: 986,502 routes
 - IPv6: 222,168 routes

"Total" Outbound RS Reachability

- IPv4 - 56.6%
 - 557,996 Reachable via RS
 - 986,502 Total Prefixes in DFZ

- IPv6 - 60%
 - 133,401 Reachable via RS
 - 222,168 Total Prefixes in DFZ

(IPv4) IX Route Server Impact
Size of Route Server vs the Gain of adding each IX



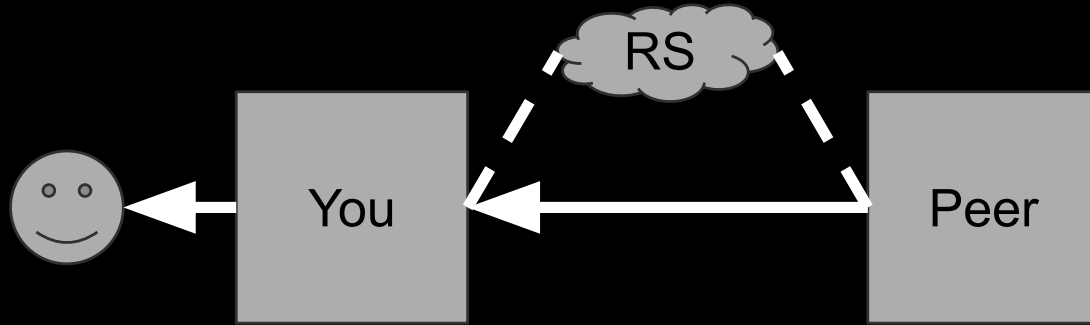
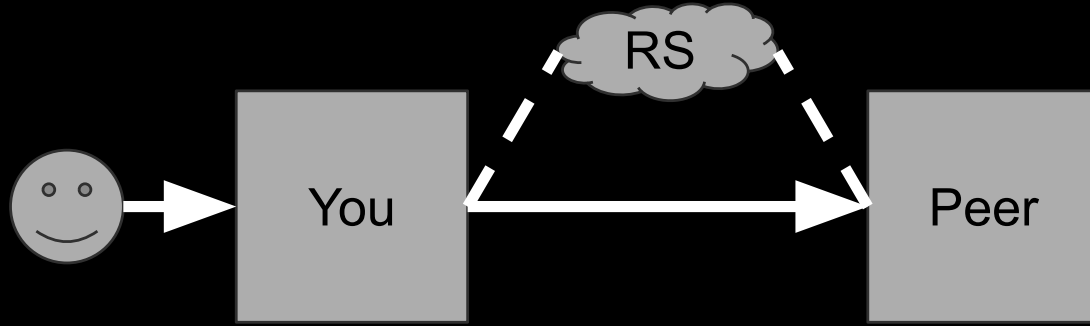
Obvious (to me) caveats to this

1. This is a survey of 113 exchanges (a lot of exchanges)
 - a. You are not likely to go out and build a network like this to offload this traffic, you would just buy more transit instead
 - b. Also this 113 exchanges is missing a few notable groups of exchanges, notably, Equinix, (and Internet Association of Australia)
2. Just because it's 60% of your prefixes does not mean it is 60% of your traffic
 - a. As far as I can tell for eyeball traffic profiles, 50%+ of all traffic is concentrated in just 5 ASNs
3. This calculation does not account for people pushing more specifics into the IX RS for traffic engineering, A technique very popular in places like Brazil

Full list of exchanges

AMS-IX	YXEIX	Netnod Helsinki GREEN	MegaIX Los Angeles	DE-CIX Munich
OGIX	SOX Serbia	Netnod Copenhagen GREEN	MegaIX Denver	MegaIX Brisbane
INTERIX	Frys-IX	MegaIX Chicago	BIX.BG	MegaIX New York
NMBINX	EdgeIX - Perth	BBIX Amsterdam	DE-CIX Phoenix	DE-CIX New York
SONIX Stockholm	France-IX Toulouse	BBIX Tokyo	DE-CIX Hamburg	DE-CIX Richmond
PIT-IX	LINX LON1	BBIX Singapore	MegaIX Las Vegas	DE-CIX Chicago
CINX	Stuttgart-IX	BBIX Dallas	MegaIX Ashburn	LU-CIX
QIX Montreal	FIXO	BBIX US-West	DE-CIX Palermo	DE-CIX Dallas
JINX	EdgeIX - Auckland	BBIX Chicago	Douala-IX	MegaIX Frankfurt
DINX	BreizhIX	BBIX London	MegaIX Bay Area	DE-CIX Marseille
GPC Missouri	NIX.CZ	MegaIX Auckland	Netnod Copenhagen BLUE	B-IX
STUIX	EdgeIX - Adelaide	Netnod Helsinki BLUE	DE-CIX Barcelona	MegaIX Adelaide
FD-IX - Indianapolis	LONAP	France-IX Paris	RomandIX	DE-CIX Frankfurt
EdgeIX - Brisbane	EdgeIX - Melbourne	Netnod Sundsvall	MegaIX Dusseldorf	MegaIX Berlin
THINX Warsaw	France-IX Lille	Netnod Gothenburg	MegaIX Sofia	MegaIX Hamburg
InterLAN-IX	France-IX Marseille	MegaIX Perth	NL-ix	MegaIX Sydney
SIX.SK	BCIX	MegaIX Singapore	MegaIX Melbourne	DE-CIX Lisbon
France-IX AURA	GetaFIX Davao	MegaIX Seattle	DE-CIX Madrid	ONIX
EdgeIX - Sydney	GetaFIX Cebu	MegaIX Atlanta	DE-CIX Dusseldorf	MSK-IX Moscow
BNIX	IRAQ-IXP	MegaIX Dallas	DE-CIX Istanbul	MegaIX Munich
IX.br (PTT.br) São Paulo	IXP.mk	MegaIX Miami	Netnod Stockholm BLUE	DE-CIX Leipzig
NIX.SK	Lillix	MegaIX Charlotte	Netnod Stockholm GREEN	
		MegaIX Toronto		

Traffic directions covered



Inbound will need a different strategy

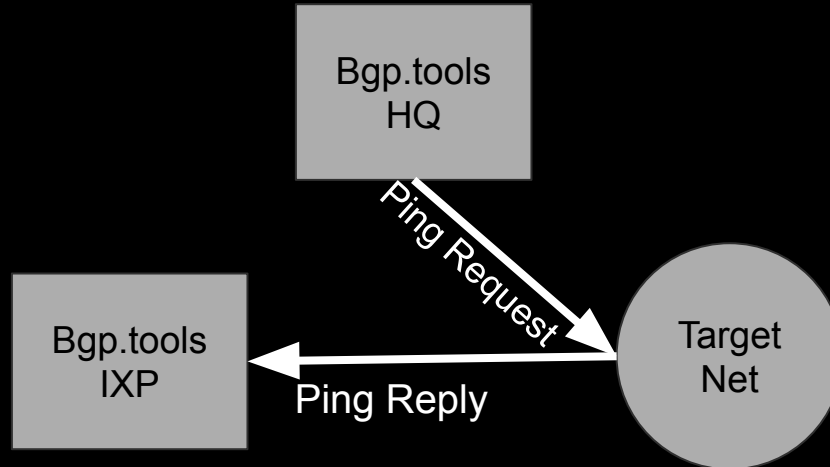
- BGP has no feedback signal to a peer for "yes I accepted a prefix"
- So we need to find another way to see if they have accepted the prefix
 - One way could be to use the bgp.tools live data set, however even though this is big (3k sessions), that is not big enough to cover the "whole" internet
- So we need to figure out a way to do test this on the data plane!

Actual strategy

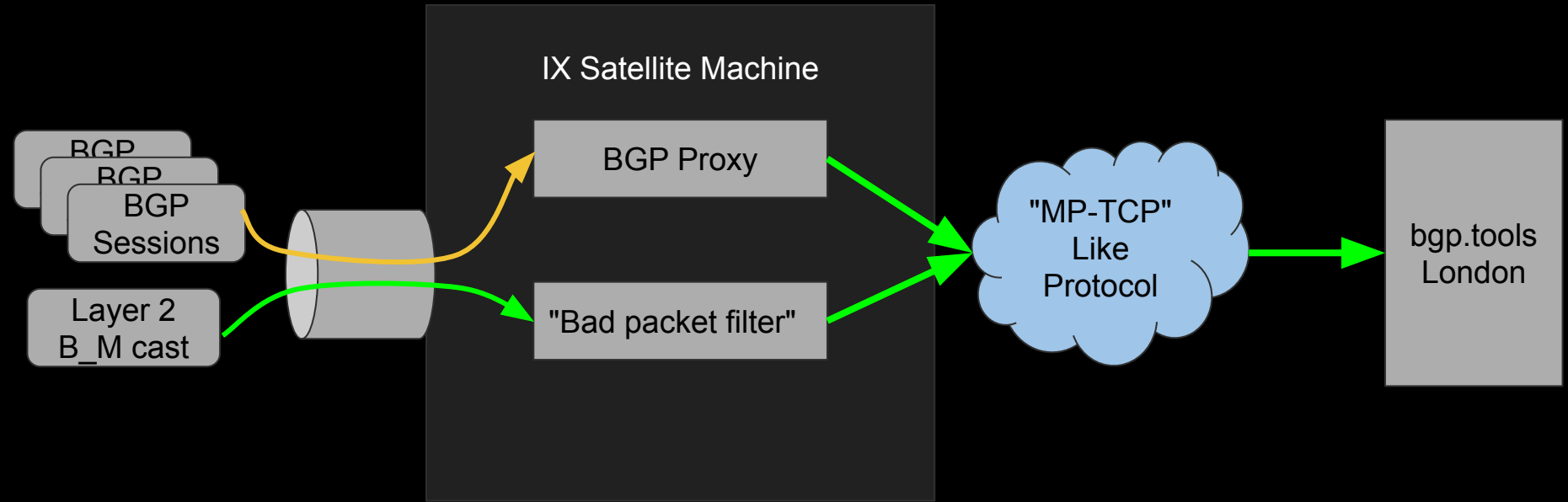
- Announce a prefix to all route servers
- From somewhere else, send ICMP pings to internet address, with the source IP being that RS only prefix
- If they are accepting a route, it will go back to the IX node, if they don't the reply as nowhere to go

Actual strategy

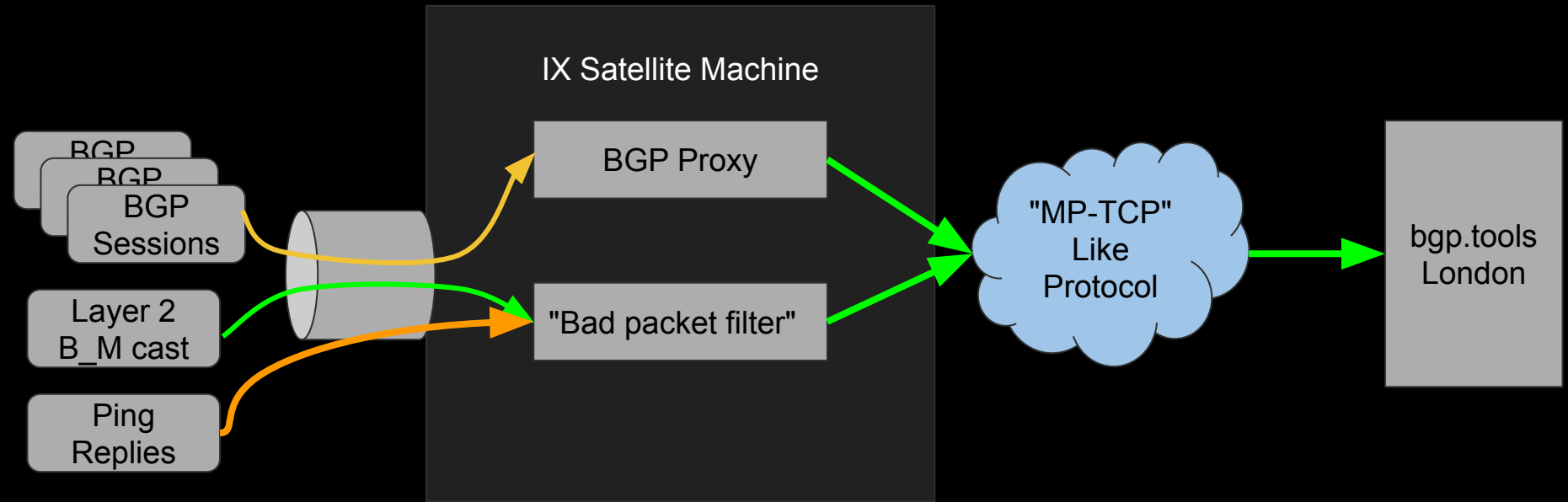
- Announce a prefix to all route servers
- From somewhere else, send ICMP pings to internet address, with the source IP being that RS only prefix
- If they are accepting a route, it will go back to the IX node, if they don't the reply as nowhere to go



How does that work?



How does that work?



Actual strategy

- Announce a prefix to all route servers
- From somewhere else, ping the route server with the source IP being that RS only
- If they are accepting the prefix, ping the node, if they don't the reply as nowhere



203.10.63.0/24

Thanks to Geoff Huston/APNIC!



Optimisations

- To speed up things, we will only be pinging 1 IP out of each /24 (one that is known to reply)
- Bgp.tools already knows what responds to pings because there is map.bgp.tools !

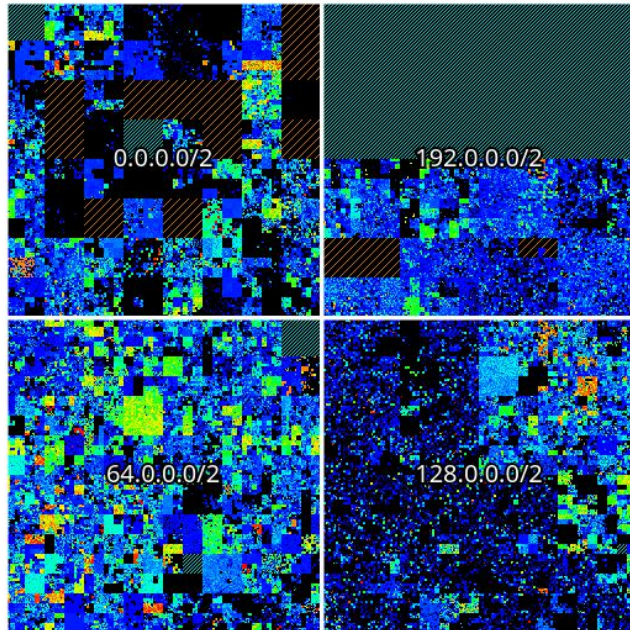
Hover: 197.141.0.0/18



2025-08-01

+

-



Getting left behind

- Because not everything responds to ICMP, some prefixes are just left behind in this test because there is no easy way to test them
- This is sadly common for CGNAT blocks

58.238.112.0/24

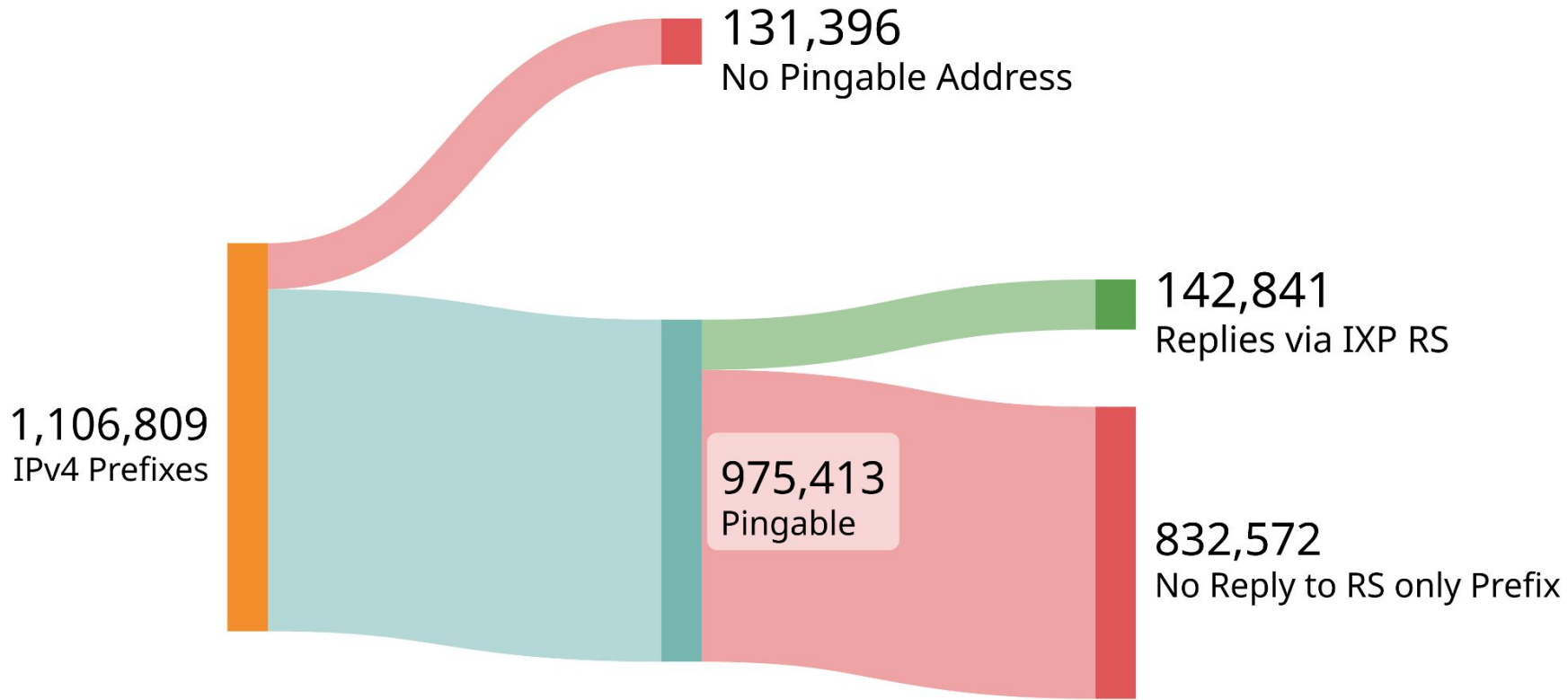
58.238.115.0/24

58.238.113.0/24

58.238.114.0/24

If you use all* RS-es, you get

- 88.1 % {975413} of prefixes can be tested
- 11.8 % {131396} of prefixes can't be tested
- 14.3 % {139523} of prefixes can reach you if you are IX RS only
- 85 % {693049} of prefixes can't reach you
- Using 113 internet exchanges
- Most packets go to the following exchanges:
 - Most packets go to the following exchanges:
 - DE-CIX FRA - 25 %
 - BBIX SIN - 13 %
 - BBIX NRT - 11 %
 - AMS-IX - 9 %
 - Frys-IX - 4 %



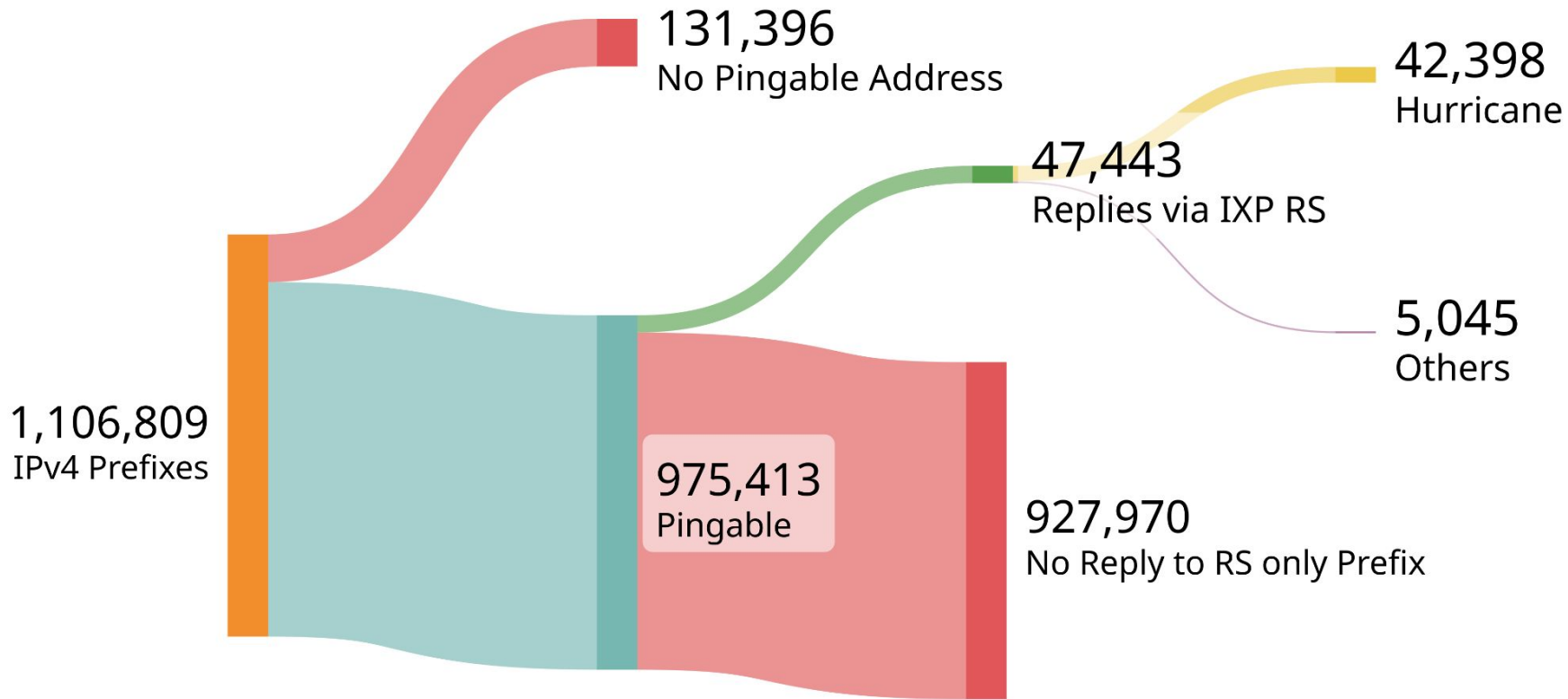
Localised surveys (Just Oceania MegaIX / EdgeIX)

- 4.8% {47443} of testable prefixes can reach you
- Using 12 internet exchanges
- Most packets go to the following exchanges:
 - MegaIX Perth - 70%
 - MegaIX Auckland - 23%
 - EdgeIX Perth - 1.8%
 - MegaIX Sydney - 1.8%
 - MegaIX Melbourne - 0.9%

Localised surveys (Just Oceania MegaIX / EdgeIX)

Without
HE / AS6939

- 0.51 % {5045} of prefixes can reach you
- Using 12 internet exchanges
- Packets go to the following exchanges:
 - MegaIX Perth 20.97%
 - EdgeIX Perth 17.52%
 - MegaIX Sydney 17.44%
 - MegaIX Melbourne 8.80%
 - EdgeIX Melbourne 8.52%
 - EdgeIX Auckland 8.50%
 - MegaIX Auckland 7.51%
 - EdgeIX Sydney 7.20%
 - MegaIX Brisbane 1.96%
 - EdgeIX Brisbane 1.31%
 - EdgeIX Adelaide 0.50%



Localised surveys (Just the big euro ones)

- 12% {125525} of prefixes can reach you
- Using LINX LON1, DE-CIX FRA, AMS-IX
- Most packets go to the following exchanges:
 - DE-CIX Frankfurt - 49.6%
 - LINX LON1 - 29.2%
 - AMS-IX - 21.3%

Thoughts

- The difference between 3 big european IXs and all 113~ exchanges is 5871 routes
 - Now those 5871 routes may matter a lot!
 - But clearly a lot of this is *dominated* by a small number of networks with large downstream customer counts who also import route servers
- Many people export to route servers, few people import from them
 - 557,996 exported prefixes to RS
 - 131,396 prefixes reachable from RS only (aka, networks who import from RS)
- That's a huge difference!
- Some of this is also limited by some prefixes (maybe the "juicy" CGNAT) ones being untestable

Bonus Thoughts

- Because this system uses the same pipeline bgp.tools uses to detect "naughty packets" on exchanges, I also know what networks these route server replies came from
- The results surprised me! On the All IXPs run the top source networks were:
 - BBIX Tokyo AS58807 / China Mobile International Limited
 - BBIX Tokyo AS9498 / Bharti Airtel Ltd.
 - BBIX Singapore AS18403 / FPT Telecom Company
 - DE-CIX Frankfurt AS7713 / PT Telkom Indonesia Tbk
 - BBIX Singapore AS9002 / RETN Limited
 - DE-CIX Frankfurt AS7195 / EdgeUno

Thanks!

Questions? Comments? Stories?
Shy? Email ausnog@benjojo.co.uk
(or fediverse/mastodon [@benjojo@benjojo.co.uk](https://benjojo@benjojo.co.uk))