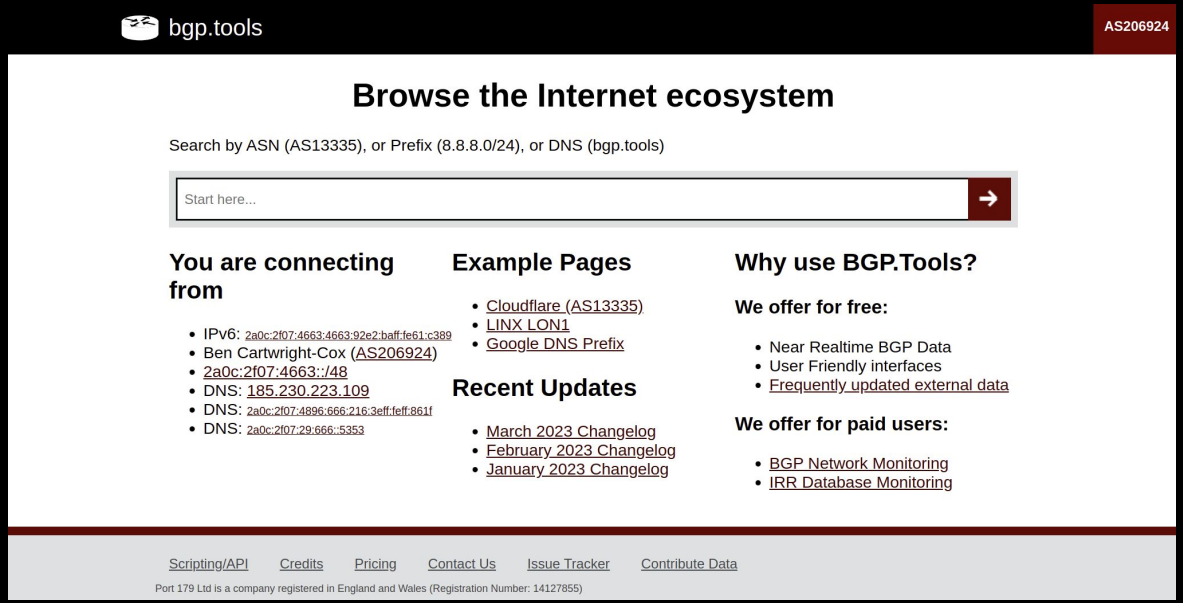
Getting creative with IXP route collection

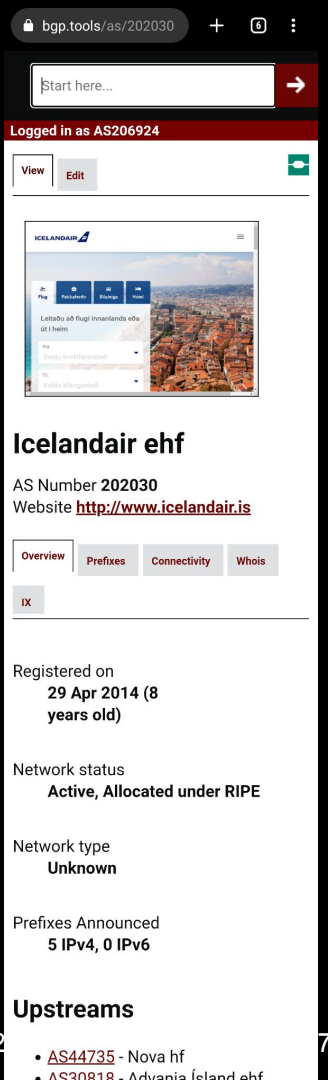
Ben Cartwright-Cox - PeeringDays 2023

bgp.tools

Quick overview of bgp.tools



bgp.tools



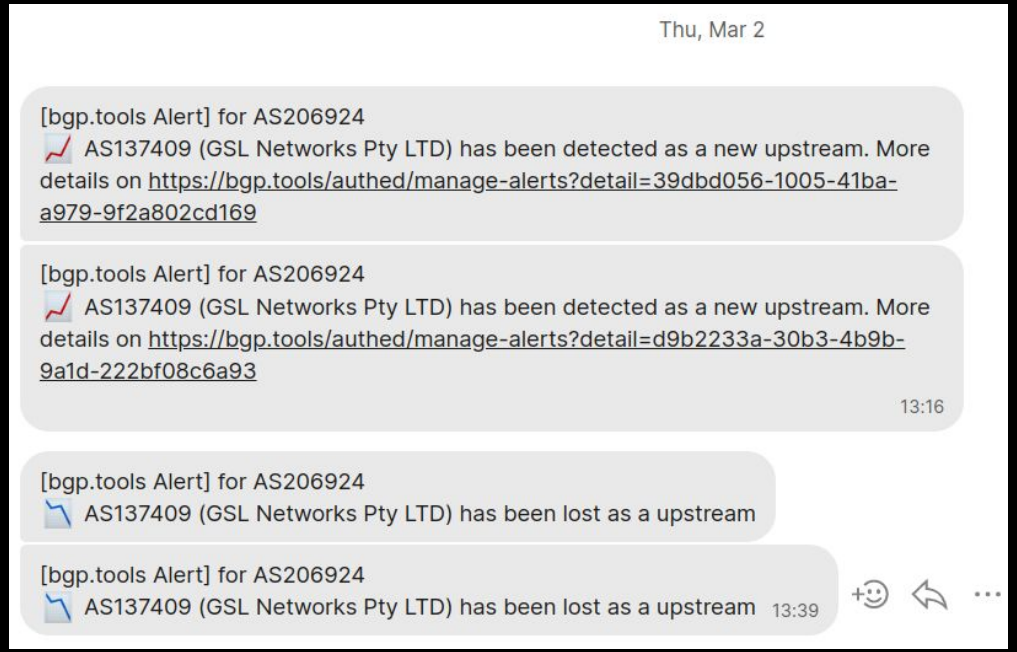
Core points

- Built out of the frustration I had with other tools
- 910~ BGP sessions established
- Practically realtime BGP peer updates
- WHOIS is handled with more care, and in some cases is updated in near realtime
- Frequently updated (~14 days):
  - ICMP Ping data scans of IPv4 /0
  - IPv4 and IPv6 RDNS data
  - Forward DNS data (Looking what A records point to a prefix)
- Peering IXP data is provided (MAC address vendors, ping data etc)

bgp.tools

Monitoring services

- Goal is to fund the site (and myself) by proving nice paid tools like rapid BGP/RPKI/IRRdb monitoring, managed looking glasses, BGP Session data recording



bgp.tools

ANYWAY

Sorry didn't mean to do a sales pitch

bgp.tools

162634855247785eb40804512729e88e

One of the core bits of today's adventure

Upstreams

	ASN	Description	IPv4	IPv6
	AS20473	The Constant Company, LLC	✘	✔
	AS5511	Orange S.A.	✔	✔
	AS6939	Hurricane Electric LLC	✘	✔
	AS3170	VeloxServ Communications Ltd	✔	✔
	AS44684	Mythic Beasts Ltd	✔	✔

bgp.tools

162634855247785eb40804512729e88e

Determining the upstreams of an ASN is a non trivial task

- You have to have loads of vantage points
- You have to detect and scrub route optimisers who lie to you

bgp.tools

Determining the upstreams of an ASN is a non trivial task

- You have to have loads of vantage points
- You have to detect and scrub route optimisers who lie to you

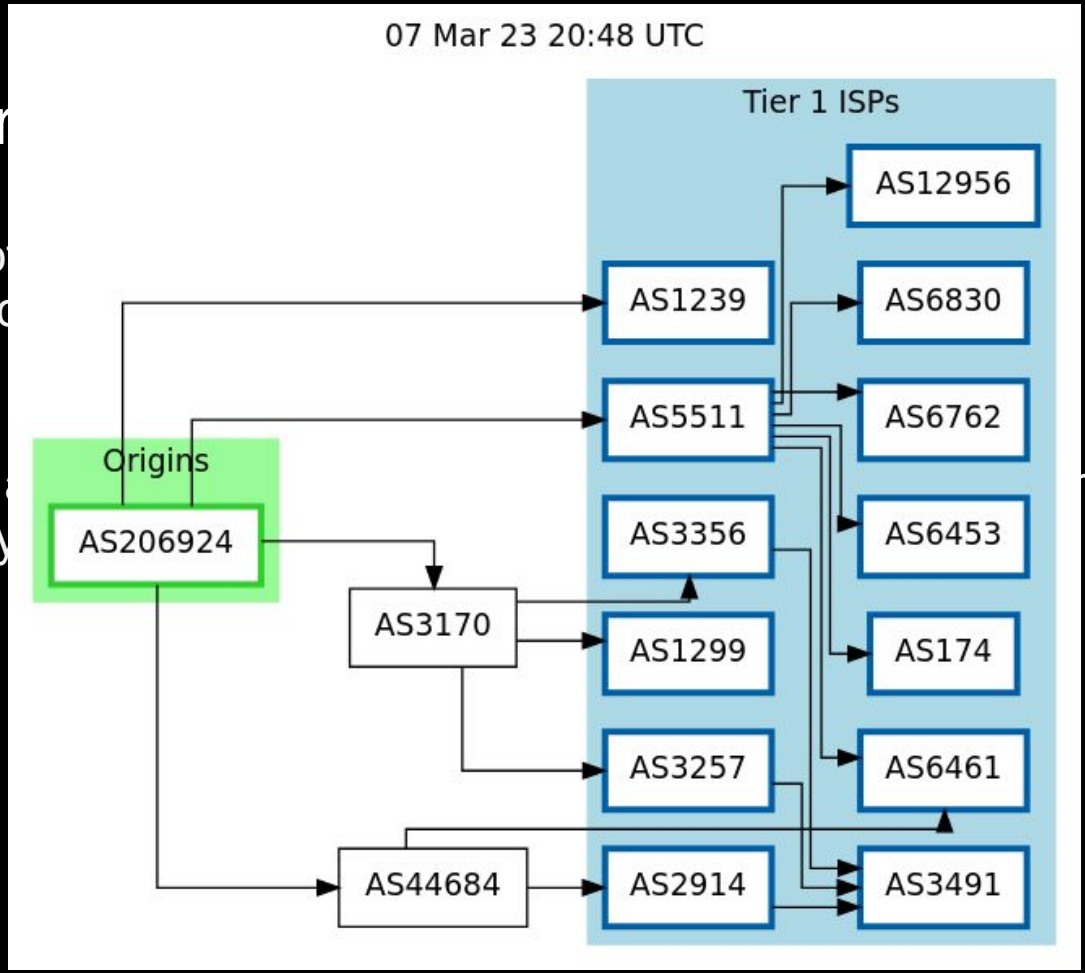
- "Upstreams" is not really a truly well defined concept, but I view it as someone who sends you eventually up to a T1 network

bgp.tools

Determining the upstr

- You have to have loads o
- You have to detect and s

- "Upstreams" is not really
  who sends you eventually

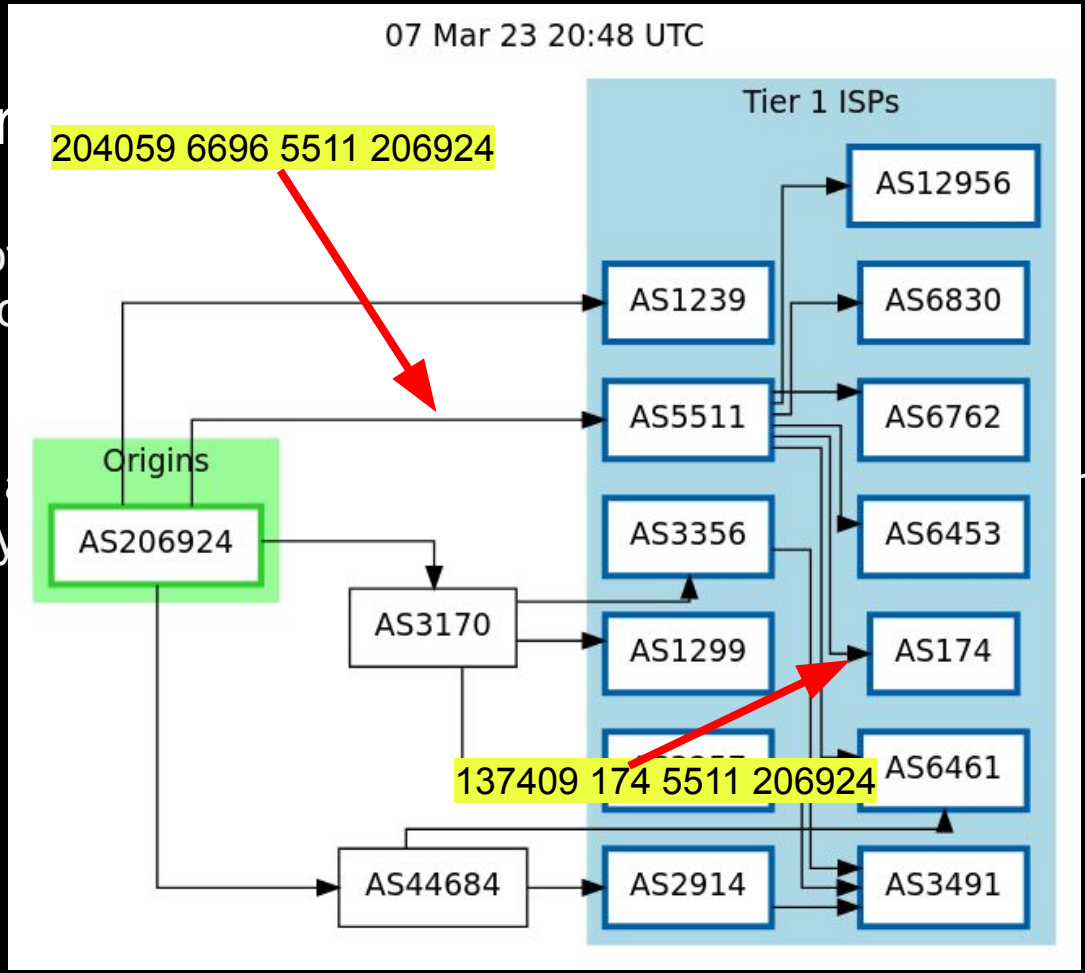


bgp.tools

Determining the upstr

- You have to have loads o
- You have to detect and s

- "Upstreams" is not really
  who sends you eventually



bgp.tools

Determining the upstreams of an ASN is a non trivial task

- You have to have loads of vantage points
- You have to detect and scrub route optimisers who lie to you

- "Upstreams" is not really a truly well defined concept, but I view it as someone who sends you eventually up to a T1 network

bgp.tools

Determining the upstreams of an ASN is a non trivial task

- You have to have loads of vantage points
- You have to detect and scrub route optimisers who lie to you

- "Upstreams" is not really a truly well defined concept, but I view it as someone who sends you eventually up to a T1 network

- It is not clear that this is a good definition in some regions
- ISPs like Zayo and Sprint peer a *lot* on Route Servers, so you have to apply logic to them to avoid everyone on a IXP having Zayo and Sprint as a upstream

bgp.tools

The two problems this talk will explore

You have to have loads of vantage points

It is not clear that this is a good definition in some regions

bgp.tools

16263485524785eb40804512729e88e

What is upstreaming?

bgp.tools

Current status quo

- Bgp.tools considers any path that crosses a T1 a path where upstreams can be learned.
- Current calculation goes:
  - 917 137409 174 5511 206924 112
  - 917 -> 137409 -> 174 is ignored, since the feed could easily be lying about this to juice their upstream stats (or incompetence)
  - 174->5511 are considered peers, since they are T1s
  - 5511->206924 results in 206924 having 5511 as a upstream (and 206924 being a downstream of 5511)
  - 206924->112 results in the same effect, 112 as a upstream of 206924
  - All of this with a single BGP ASPATH!

bgp.tools

Current status quo (problems #1)

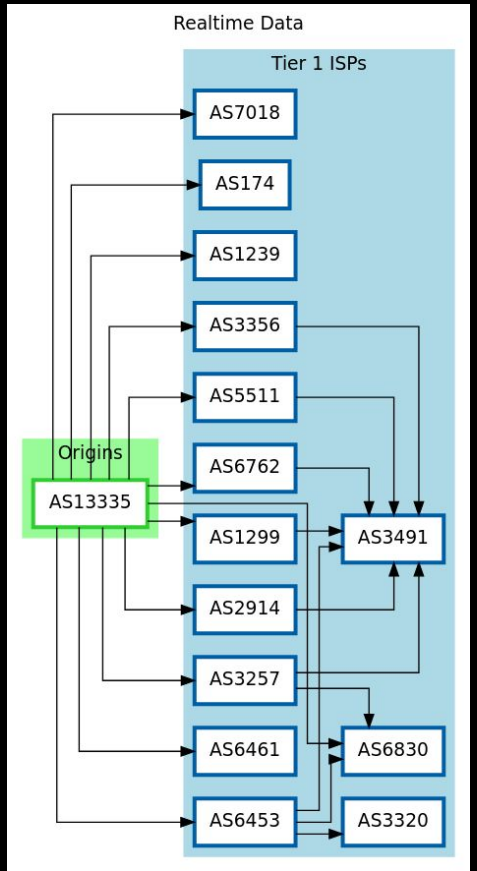
- This does not consider that you may be a peer of the T1.
  - If I see 137409 174 206924 It could (but very unlikely) be that AS206924 is a peer of cogent
  - This is more likely with (A) Zayo, (B) Sprint since (A) peers a lot, (B) imports peers on the route servers of IXPs
    - Though sprint won't be a problem for too much longer I guess…
  - Zayo and Sprint have a requirement of a "double hop", meaning we **must** observe two T1 ASN hops before considering it a upstream route
    - Since that would imply that route is in their customer cone
  - This trick works well, but it does not scale to all T1's because…

bgp.tools

Current status quo (problems #2)

My code

AT&T, Liberty Global, ✨, 🍊, TATA, GTT, NTT, Lumen, Cogent and 🦁 are upstreams



My (own brain) interpretation

AT&T might be peering???
Cogent is almost defo transit
Sprint, Zayo are likely peering

🍊, Lumen, ✨ ,🦁,NTT, GTT, TATA are all double hops. Assumed upstreams

AT&T, Liberty are likely not upstreams

bgp.tools

New idea

- A upstream is someone who (*with consent)* carries your routes to their peers/upstreams
- Basically, if it's in your cone, you are a upstream for them

- The basic way of implementing this is to keep doing the T1 logic, but also import in IXP route servers, since they are a great insight into the cone of networks in a pure-ish way

- Downside: You must get feeds from loads of different IXP route servers

bgp.tools

The two problems this talk will explore

You have to have loads of vantage points

It is not clear that this is a good definition in some regions

bgp.tools

16263485524785eb40804512729e88e

RIS and Routeviews

bgp.tools

Peering LAN first

- Most collectors live on IXPs
- Kind of a mess of IXP peers just sending their customer cone, not their full table
- Previously (Maybe still currently?) RIS has been limiting new sessions due to capacity concerns?

Name	Physical Location	Type	Scope	Raw Data
RRC00	Amsterdam, NL	multihop	global	data
RRC01	London, GB	IXP	LINX, LONAP	data
RRC03	Amsterdam, NL	IXP	AMS-IX, NL-IX	data
RRC04	Geneva, CH	IXP	CIXP	data
RRC05	Vienna, AT	IXP	VIXP	data
RRC06	Otemachi, JP	IXP	DIX-IE	data
RRC07	Stockholm, SE	IXP	Netnod	data
RRC10	Milan, IT	IXP	MIX	data
RRC11	New York, NY, US	IXP	NYIIX	data
RRC12	Frankfurt, DE	IXP	DE-CIX	data
RRC13	Moscow, RU	IXP	MSK-IX	data
RRC14	Palo Alto, CA, US	IXP	PAIX	data
RRC15	Sao Paolo, BR	IXP	PTTMetro-SP	data
RRC16	Miami, FL, US	IXP	Equinix Miami	data
RRC18	Barcelona, ES	IXP	CATNIX	data
RRC19	Johannesburg, ZA	IXP	NAP Africa JB	data

bgp.tools

But I can't deny IXP collection does not work

- RIPE has ~1500 BGP sessions online,
  - 374 / 403 Full IPv4/IPv6 tables
  - (by their own calculations)
  - Some of these sessions have issues about to be mentioned, **some are immensely useful views of the internet though!**

- bgp.tools is eBGP Multihop only
  - 910~ Sessions online
  - 522 / 800 Full IPv4/IPv6 tables (!!)

bgp.tools

Problems with IXP Route Collection

- Really expensive if you don't have friends
  - IXP Membership fees + XC fees + colo fees
  - IXP membership alone can be more than the last two
  - https://peering.exposed
- ***Huge*** bias to AS6939
  - They are on almost all of the large IXPs, and provide you 180k+ of peered v4 routes that will likely be preferred over transit, hiding transit paths from the collector

bgp.tools

Solving for XC Fees / Colo

- What is the cheapest, smallest, most insane thing we could ship to a willing IXP?

bgp.tools https://blog.benjojo.co.uk/post/smart-sfp-linux-inside

Solving for XC Fees / Colo

- What is the cheapest, smallest, most insane thing we could ship to a willing IXP?

bgp.tools https://blog.benjojo.co.uk/post/smart-sfp-linux-inside

162634855247785eb40804512729e88e

Solving for XC Fees / Colo

- What is the cheapest, smallest, most insane thing we could ship to a willing IXP?

- No XC, The switch is the power supply, you can hitch backhaul either via someone friendly on the IXP, or relaying via a VPS or something
- Cheap, Around 150 USD all in
- Single core ARMv7, with 512M of RAM running Debian Jessie
- ***Completely deranged***. Everyone is going to look at you like you lost the plot!
- Made by a Russian company who is totally now a Dubai company since the Russian Invasion of Ukraine

Looking in the right places

bgp.tools

Looking in the right places

- 80~ USD from FS.com
- Runs a 400Mhz 32bit MIPS core, 64MB of RAM
- Literally running OpenWRT out of the packaging
- Also, here are the SSH credentials to all of them
  - user `ONTUSER`
  - password `7sp!lwUBz1`
- The constrained RAM and MIPS CPU µArch makes this a challenge to program for
- Thankfully Zig lang has a mostly working MIPS target!
- Can change the vendor to w/e with just `` `sfp_i2c -i 0 -s "CISCO"` ``

bgp.tools

Dialing down the insanity

- A lot of IXPs have reseller programs, and with friends you can pick up very cheap ports and maybe a tagged vlan on a virtual machine.

- Downside is that even with this, a lot of the IXPs still require you to be a full member to be present on the LAN

bgp.tools

All roads lead back to London

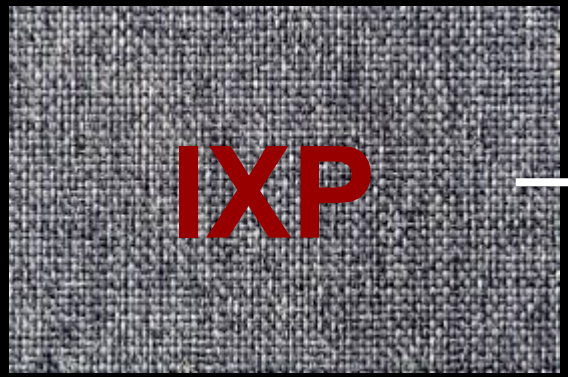
- You have have noticed it isn't really possible to store a *modern* full internet table on 64MB of RAM.

- Instead of storing sessions locally, the local collector will "rehost" the BGP session back in London where all of the infra is.

- This is because with how bgp.tools is designed, all BGP data has to be within 1ms~ of the web server to ensure a enjoyable experience

bgp.tools

Overall



Relay 1

Relay 2

Relay 3

bgp.tools London

bgp.tools

162634855247785eb40804512729e88e

Current expanding plans

- NL-IX is coming up via a reseller port
- African IXP relays are slowly being figured out via two different carriers
- Got contacts at AMS-IX / DE-CIX / FranceIX / MIX? I'd love to know them
- IX.BR is home to a lot of interesting networks that don't wish to eBGP multihop peer, I am slowly working out a reseller port setup there!
    - Same with some African networks

- Some IXPs are setting up eBGP multihop sessions from their route servers! Very nice of them!

bgp.tools

Questions?

Want to feed bgp.tools?

go to bgp.tools and go to to bottom link "Contribute Data"

More complex queries:
IRC: Benjojo-bgptools (terahertz) / benjojo (everything else)
Or email: **admin@bgp.tools**

bgp.tools