Observations from the routing diff

Anurag Bhatia

Objective

Comparison of routes in global routing table as seen across various large networks.

Why?

To see if announcements are actually global & find reason for difference in number of routes from various networks.

What we already know?

- Large networks carry relatively similar number of routes

- Route announcements can be limited by use of BGP community & hence may result in difference in number of routes

- Eyeball networks announce more specific routes to steer inbound traffic across various circuits

But we also know...

- Large networks carry relatively similar number of routes <- *But there's a visible difference of few thousand routes!*

- Route announcements can be limited by use of BGP community & hence may result in difference in number of routes <- *Can that alone result in difference of thousands of routes?*

- Eyeball networks announce more specific routes to steer inbound traffic across various circuits <- *But if route reaches default free zone, it should reach everywhere else!*

Networks considered in the study

1. AT&T - AS7018

2. Centurylink / Level3 / Global Crossing - AS3549

3. Cogent - AS174

4. GTT - AS3257

5. Hurricane Electric - AS6939

6. KPN - AS286

7. NTT - AS2914

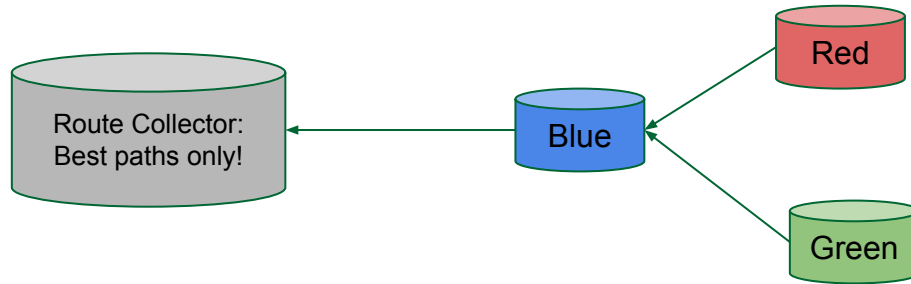8. Sprint - AS1239

9. Verizon - AS701

Why specifically these networks?

1. Large networks by size & reach

2. Full routes are visible at RIPE RIS and Oregon Route Views. For other's it's either no visibility at popular collectors

3. Routes in default free zone without no-export (or similar community) should reach everyone else

Note: Remember downstream give a picture but the picture is seperate from what upstream gets it
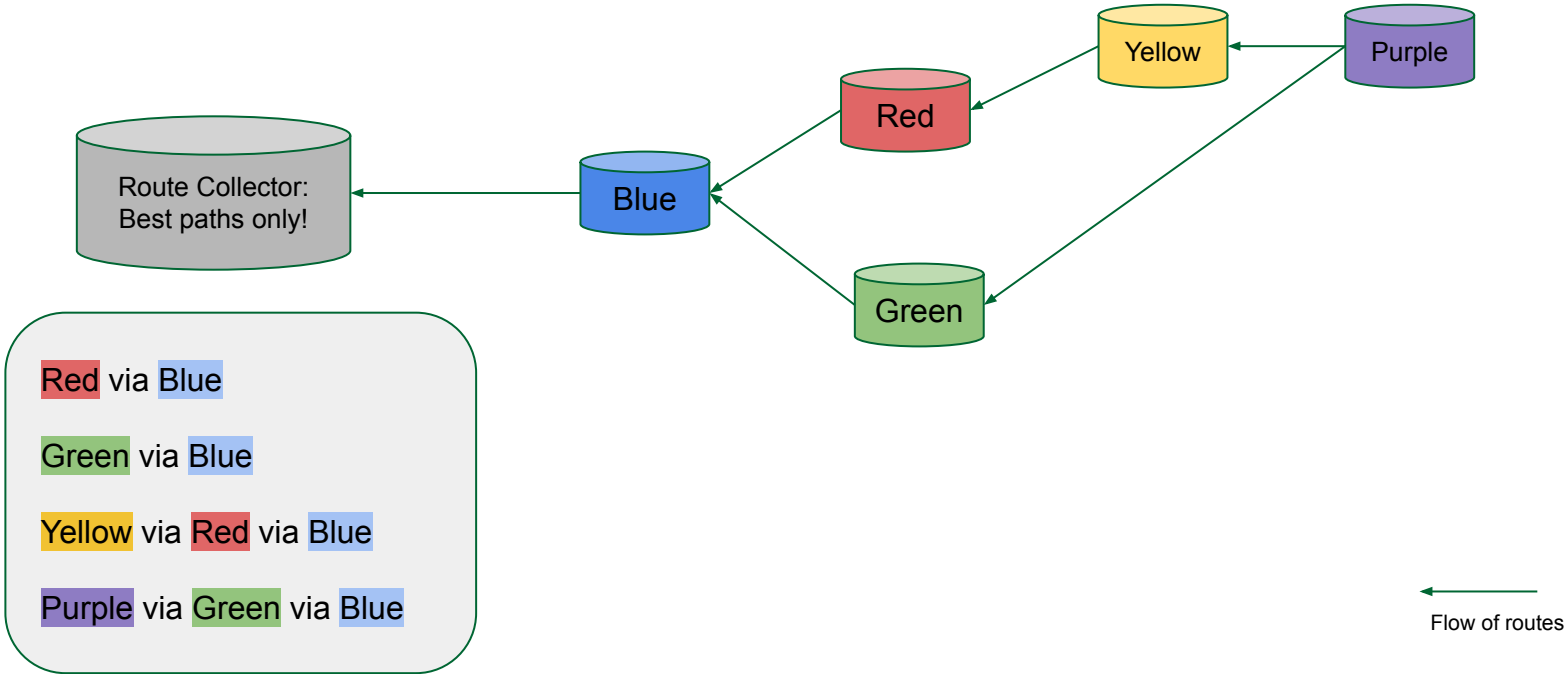
Best path propagation only

Route Collector:
Best paths only!

Blue

Red

Green

Flow of routes

Best path propagation only

Red via Blue

Green via Blue

Yellow via Red via Blue

Purple via Green via Blue

Flow of routes

For vision of the route collector

Route Collector: Best paths only!

Blue

Red

Yellow

Purple

Green

Red via Blue

Green via Blue

Yellow via Red via Blue

Purple via Green via Blue

Flow of routes
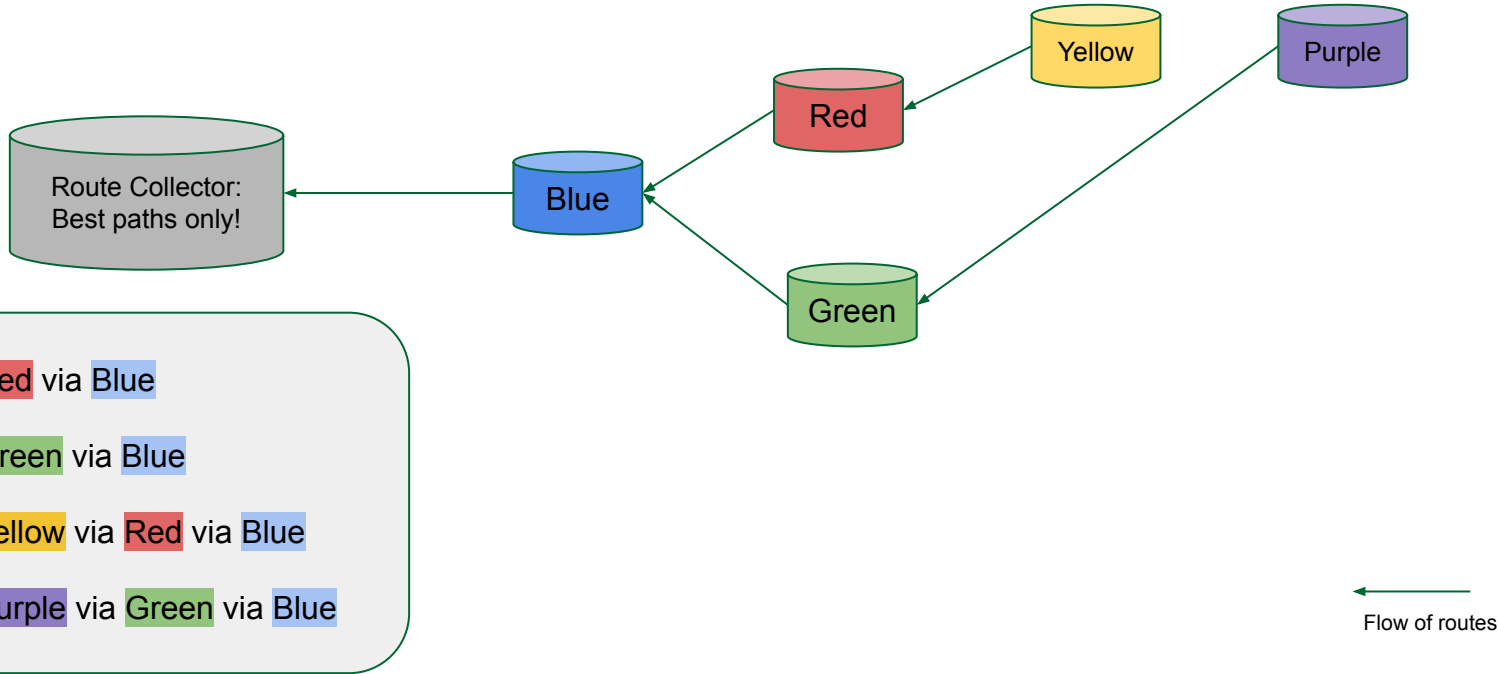
Warning! Slides with lot of numbers ahead!

Comparing routes

Total unique IPv4 Prefixes: 8,02,156

Prefixes with limited visibility: 34,872

That's 4.34% routes!

Removing announcements smaller than /24…

Comparing routes

Total unique IPv4 Prefixes: 801834

Prefixes with limited visibility: 34565
ASNs with limited visibility: 3866

Prefixes with limited visibility + invalid ROAs: 4477
ASNs with limited visibility + invalid ROAs: 987

That's 4.31% routes!

Comparing routes (cont.)

ASN	IPv4 Prefixes	Diff from total	Missing %
AS7018	769143	32691	4.08
AS174	769917	31917	3.98
AS3257	772781	29053	3.62
AS6939	796930	4904	0.61
AS286	772654	29180	3.64
AS2914	773027	28807	3.59
AS1239	771004	30830	3.84
AS701	771563	30271	3.78

Top 10 ASNs with limited visibility

ASN	Prefix Count	Prefix %	AS Name
47331	5453	15.78	TTNET, TR
31334	3203	9.27	KABELDEUTSCHLAND-AS, DE
18403	983	2.84	FPT-AS-AP The Corporation for Financing & Promoting Technology, VN
45090	481	1.39	CNNIC-TENCENT-NET-AP Shenzhen Tencent Computer Systems Company Limited, CN
4787	447	1.29	ASN-CBN PT Cyberindo Aditama, ID
7545	427	1.24	TPG-INTERNET-AP TPG Telecom Limited, AU
24955	359	1.04	UBN-AS, RU
24560	298	0.86	AIRTELBROADBAND-AS-AP Bharti Airtel Ltd., Telemedia Services, IN
38333	275	0.80	SYMBIO-AS-AU-AP Symbio Networks, AU
9808	225	0.65	CMNET-GD Guangdong Mobile Communication Co.Ltd., CN

Top 10 countries with limited visibility

Country	Prefix Count	Percentage
TR	5770	16.69
BR	3856	11.16
DE	3411	9.87
US	3002	8.69
CN	2281	6.60
RU	1290	3.73
AU	1218	3.52
ID	1207	3.49
IN	1203	3.48
VN	1050	3.04

Possible reasons for the difference

1. RPKI based filtering by network like AS7018 etc and dropping 4k+ invalids in diff list

2. IRR based filtering resulting in drop due to bad route object by certain networks

3. IRR based filtering resulting in drop due to missing AS-SET in the chain

4. Route announcement by anycast sensitive networks with no-export to these specific networks

5. Delays due to slow BGP convergence

6. Misconfiguration - Announcing route only to a "peer" and not all default free zone peers or transit

7. BGP filtering: Some networks filtering downstream & found with unopened filters for few days while some accepted without filtering.

Some misc things...

1. Large operators like AS18101 signed prefixes and missed to create ROAs for smaller chunks assigned & announced by downstream resulting in RPKI invalids

2. Bad operational practice like removing a prefix without covering prefix from "link 1" and announcing on "link 2"

3. Around 454 routes of AS45090 were missing specifically from AS701 and AS1239

4. In total 1178 routes belonging to China not found from AS701 & AS1239 where ROAs were OK or not present (but not invalid)

What operators can do to ensure global reachability?

1. Ensure proper IRR hygiene. Remember to create IRR route objects, mantain AS SETs (if you have downstream), ensure your ASN is in upstream's AS SET

2. Ensure RPKI hygiene. Ensure ROAs are present with correct origin AS and max length

3. If load balancing traffic with more specific, ensure covering prefix is always present to avoid convergence issues

4. Know the difference between your "Peers" and "Transit". Don't assume peers to be transit! :)

Questions?

Anurag Bhatia
anurag@he.net