

Let's talk about routing security

How secure is our routing infrastructure in 2019?

Fundamentals of global routing

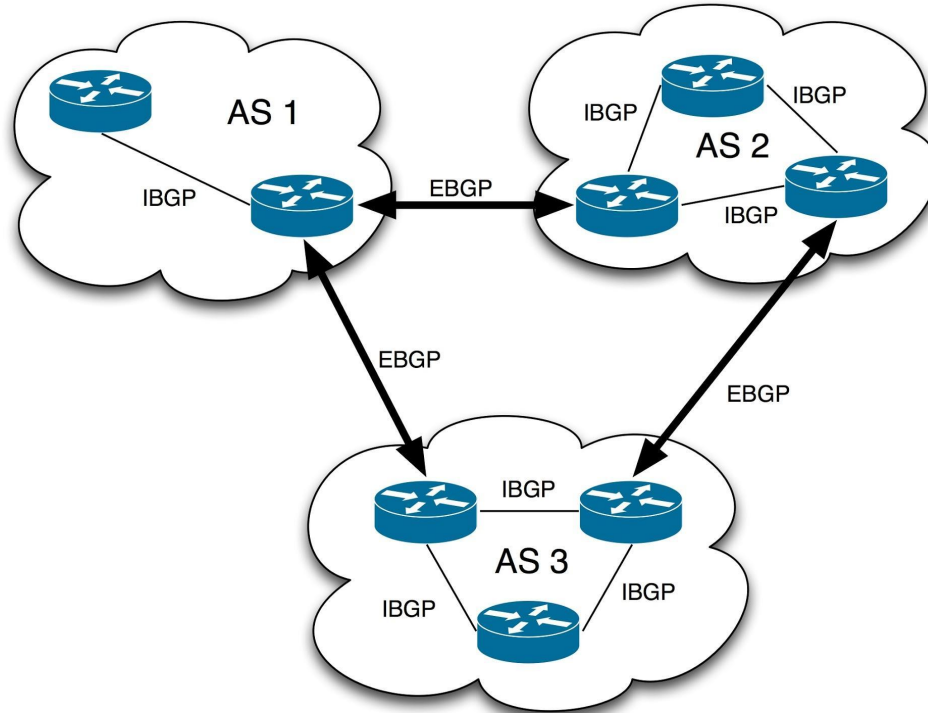
1. Fundamentals of global routing
2. IRR
3. Present statistics
4. Future

Fundamentals of global routing...

Internet - Network of ASNs...

- Internet is simply network of autonomous networks all connected together and speaking “BGP”.
- There are around 65k autonomous networks (known by their number called ASN) in IPv4 routing and 17k ASNs in IPv6 world.
- A set of around 15 networks stitch these ASNs together by forming a “default free / transit free zone” and essentially all ASNs in the world are direct/indirect customer of either of these ASNs.
- A large part of modern traffic flows from a limited set of ASNs (content networks) to eyeball networks via PNI’s and Internet Exchanges

Internet - Network of ASNs...



Internet - Network of ASNs...(+ DNS!)

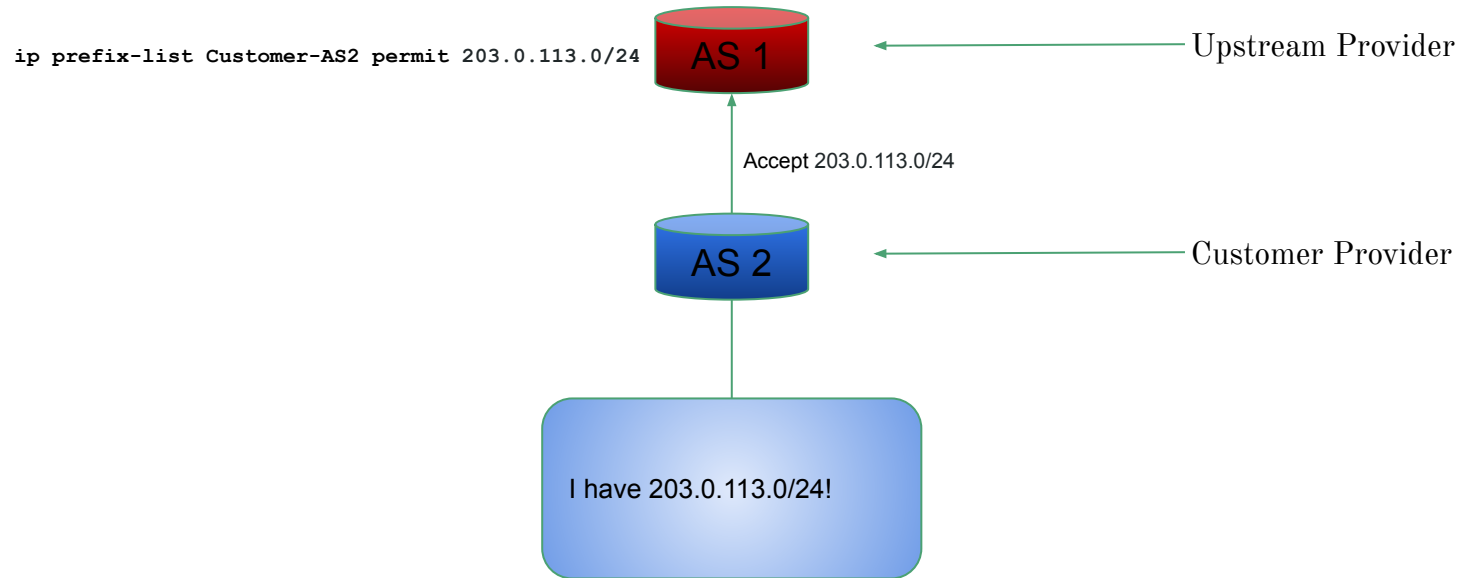
- BGP ensures interconnection of networks and DNS ensures domain to IP mapping.
- DNS relies on set of 13 logical root DNS servers and practically as many as 980 instances across the world via anycast.
- These 13 root DNS addresses are hardcoded in DNS resolver software (like BIND, powerdns etc) and hence security of these 13 IPs is important.
- DNS resolver contacts either of 13 based on reply time and other factors in the resolution algorithm.

So how “trust” in the BGP works?

Trust in the BGP...

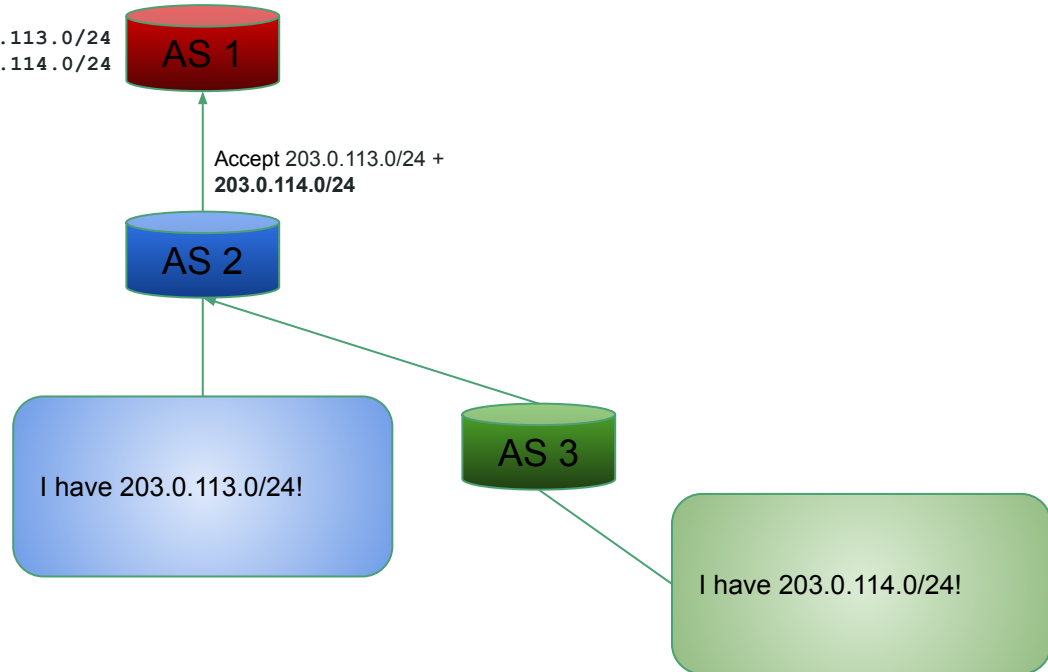
- BGP supports filtering and networks can define in filter what they can accept or reject and the default action (accept/reject).
- Filters can be based on IP prefix, ASN or AS Path or other factors like BGP community.
- Edge filtering - Filter the networks which connect to you based on static filter based on prefix and some other basic rules and full stop.
- Filtering beyond the edge - Allow prefixes of your downstream customer + their downstream + further their downstream and so on...

Filtering chain...

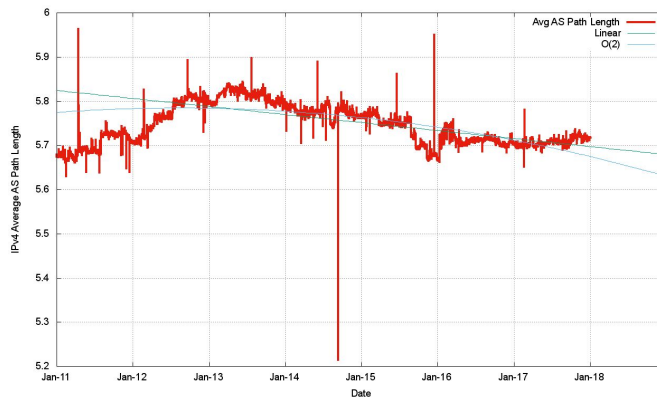


Filtering chain...

```
ip prefix-list Customer-AS2 permit 203.0.113.0/24  
ip prefix-list Customer-AS2 permit 203.0.114.0/24
```



Trust in the BGP...



As per research data by Mr Geoff Huston (Scientist at APNIC) average AS path length in IPv4 world is around 5.7 and hence for a case like $AS\ 1 \leftarrow AS2 \leftarrow AS3 \leftarrow AS4 \leftarrow AS5$ it's very hard for AS1 to what to allow for AS4 (learnt via AS2).

How does filtering works at the “Internet scale” ?

IRR

1. Fundamentals of global routing
2. IRR
3. Present statistics
4. Future

IRR - Internet Routing Registries

- IRRs are the public “registers” where one can log what they want to do and then just do it.
- IRRs use RPSL (Routing Policy Specific Language) to define “route object” where one defines prefix, origin AS, description etc and upstream can generate filters based on that.
- IRRs use “AS SETs” which define ASNs in a set (for instance a set of customer ASNs) and that is used to define customers ASNs.
- AS SETs can further have AS Sets of customer and that helps to generate downstream’s downstream’s downstream filter.

IRR - Route Object Example

```
whois -h whois.radb.net 216.218.128.0/17
```

```
route:      216.218.128.0/17
descr:      Hurricane Electric
            55 South Market St
            San Jose, CA
origin:     AS6939
notify:     noc@he.net
changed:    noc@he.net 20170407
mnt-by:     HE-NOC
source:     RADB
```

IRR - Route Object Example

```
whois -h whois.radb.net 216.218.128.0/17
```

```
route:      216.218.128.0/17    <- Prefix
descr:      Hurricane Electric
             55 South Market St
             San Jose, CA
origin:     AS6939      <- Origin AS
notify:     noc@he.net
changed:    noc@he.net 20170407
mnt-by:     HE-NOC
source:     RADB
```


IRR - AS SET Example

```
whois -h whois.radb.net AS-ACT
```

```
as-set:          AS-ACT
descr:           ACT-AS
country:         IN
members:         AS55577, AS131269, AS131219,
AS18209,AS45196,AS24309,AS13335,AS138318
tech-c:          TB103-AP
admin-c:         AB208-AP
mnt-by:          MAINT-IN-BEAMTELECOM
mnt-lower:       MAINT-IN-BEAMTELECOM
last-modified:  2019-01-23T18:04:59Z
source:         APNIC
```

More on Internet Routing Registries

- There are as many as 25 IRRs and were created for different reasons historically.
- Non-for profit RADB used mostly by larger organisations, free option ALTDB (for general Internet).
- One can define which IRR one is using at the peeringdb e.g RADB::AS-HURRICANE for Hurricane Electric or APNIC::AS9498:AS-BHARTI-IN for Airtel.
- RADB mirrors all major IRRs and thus a query to RADB includes it's own database as well as data of other mirrors IRRs.

Query to RADB...

```
whois -h whois.radb.net 59.145.135.0/24

route:          59.145.135.0/24
descr:         BHARTI-IN
descr:         Bharti Airtel Limited
descr:         Class A ISP in INDIA .
descr:         234 , OKHLA PHASE III ,
descr:         NEW DELHI
descr:         INDIA
country:       IN
origin:        AS9498
mnt-by:        MAINT-IN-BBIL
changed:       rar.data@airtel.in 20070814
source:        APNIC
```

Query to RADB...

```
whois -h whois.radb.net 59.145.135.0/24
```

```
route:          59.145.135.0/24
descr:          BHARTI-IN
descr:          Bharti Airtel Limited
descr:          Class A ISP in INDIA .
descr:          234 , OKHLA PHASE III ,
descr:          NEW DELHI
descr:          INDIA
country:        IN
origin:         AS9498
mnt-by:         MAINT-IN-BBIL
changed:        rar.data@airtel.in 20070814
source:        APNIC <- Shows the source database
```

bgpq3 - Tool for generating filters

- Open source tool bgpq3 can be used for generating filters based on IRR.
- It supports syntax of Cisco, JunOS out of the box.
- It also supports generating filter list based on custom syntax (including JSON) of any given hardware.
- Includes supports for AS Path based filters as well as filters for IPv6 world.
- Supports only generation of filters and one needs to have a mechanism to push these filters to the routers.

bgpq3 - in action

```
bgpq3 -l Anurag AS58901 -6
no ipv6 prefix-list Anurag
ipv6 prefix-list Anurag permit 2402:b580::/32
ipv6 prefix-list Anurag permit 2402:b580:1::/48
ipv6 prefix-list Anurag permit 2402:b580:2::/48
ipv6 prefix-list Anurag permit 2402:b580:3::/48
```

```
bgpq3 -J -l Anurag AS58901 -6
policy-options {
  replace:
    prefix-list Anurag {
      2402:b580::/32;
      2402:b580:1::/48;
      2402:b580:2::/48;
      2402:b580:3::/48;
    }
}
```

bgpq3 - in action

```
bgpq3 -l Anurag AS58901 -6
no ipv6 prefix-list Anurag
ipv6 prefix-list Anurag permit 2402:b580::/32
ipv6 prefix-list Anurag permit 2402:b580:1::/48
ipv6 prefix-list Anurag permit 2402:b580:2::/48
ipv6 prefix-list Anurag permit 2402:b580:3::/48
```

<- Cisco iOS style syntax

```
bgpq3 -J -l Anurag AS58901 -6
policy-options {
  replace:
    prefix-list Anurag {
      2402:b580::/32;
      2402:b580:1::/48;      <- JunOS syntax
      2402:b580:2::/48;
      2402:b580:3::/48;
    }
}
```

It's querying RADB and formatting

```
whois -h whois.radb.net '!6as58901'  
A66  
2402:b580:1::/48 2402:b580:3::/48 2402:b580:2::/48 2402:b580::/32  
C
```

More on this on RADB here: <https://www.radb.net/query/help>

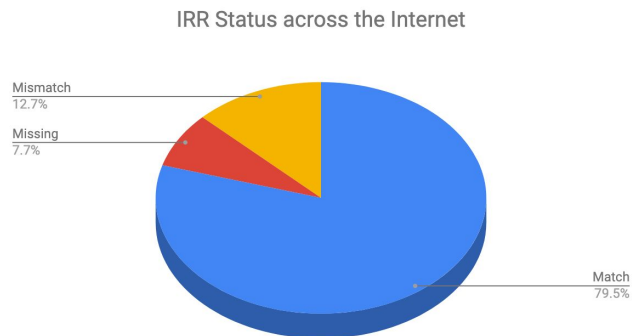
So how well IRR based filtering works?

So how well IRR based
filtering works? <- ***Not so well!***

Present statistics

1. Fundamentals of global routing
2. IRR
3. Present statistics
4. Future

Filtering Statistics across the Internet



- There are as many as 758313 prefixes visible in global routing table (IPv4 + IPv6)
- Out of total routes: 603185 (79.54%) have valid route objects, 58587 (7.73%) have no valid route objects and 96514 (12.73%) have mismatching route object.
- Thus IRR based filtering can filter/blackhole 155101 routes or 20.45% of the total routes in the global table.

Is there a real world problem here?

Last “few” route leaks/hijacks

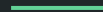
1. YouTube route hijack by PTCL, Pakistan in 2008 -
<https://www.cnet.com/news/how-pakistan-knocked-youtube-offline-and-how-to-make-sure-it-never-happens-again/>
2. Google’s route leake by Airtel in 2015 -
<https://arstechnica.com/information-technology/2015/03/indian-isps-routing-hiccup-briefly-takes-google-down-worldwide/>
3. BGP/DNS Hijacks Target Payment Systems in 2018 -
<https://hub.dyn.com/dyn-research/bgp-dns-hijacks-target-payment-systems-2>
4. Large European Routing Leak Sends Traffic Through China Telecom -
<https://blogs.oracle.com/internetintelligence/large-european-routing-leak-sends-traffic-through-china-telecom>
5. And many more!

Challenges with IRR based filtering

- IRR is old and not very easy to integrate with the routers.
- IRRs by design are log books and whatever goes in there, usually stays in there. In other words they are full of old outdated route objects.
- There's no direct incentive for smaller networks to maintain their entries in IRR and once they go online without IRR entries, they join the list of networks without valid route objects.
- The “software speaking to the routers & pushing config” isn't very common across smaller networks.

Future

1. Fundamentals of global routing
2. IRR
3. Present statistics
4. Future



Some of developments in routing security...

- Folks who are filtering: Hurricane Electric (IRR), BharatIX (IRR + RPKI), Google (based on IRR starting in Sept 2019), AT&T (based on RPKI since Nov 2018), DECIX, INEX, Equinix Singapore IX etc.
- RPKI is being pushed for to use cryptography to validate prefix origin and is supported in latest version of various vendors. For supported hardware, it's very easy to implement in a route-map / routing-policy.
- RPKI is being integrated in next version of IRR (IRR 4) to ensure route objects cannot be created where ROA mismatch happens.
- More tools coming up to track RPKI deployment. E.g <https://sg-pub.ripe.net/jasper/rpki-web-test/>

RPKI in action...

```
router bgp 58901

  address-family ipv4 unicast
  neighbor 1.2.3.4 route-map Customer-IN in
  bgp bestpath prefix-validate allow-invalid
!
route-map Customer-IN permit 10
  match rpki invalid
  set local-preference 50
!
route-map Customer-IN permit 20
  match rpki not-found
  set local-preference 100
!
route-map Customer-IN permit 30
  match rpki valid
  set local-preference 200
!
route-map Customer-IN permit 40
```






RPKI in action...

```
router bgp 58901

  address-family ipv4 unicast
  neighbor 1.2.3.4 route-map Customer-IN in
  bgp bestpath prefix-validate allow-invalid
  !
  route-map Customer-IN permit 10
  match rpki invalid
  set local-preference 50  <- Low localpref on route if RPKI check is invalid (Remember: High localpref wins)
  !
  route-map Customer-IN permit 20
  match rpki not-found
  set local-preference 100 <- Mid level localpref on route is no ROA is present
  !
  route-map Customer-IN permit 30
  match rpki valid
  set local-preference 200 <- High localpref when RPKI check is valid and route is preferred
  !
  route-map Customer-IN permit 40
```

Easy way to check IRR as well as RPKI for prefixes...



















Check for IRR / RPKI ROA validation

- Hurricane Electric's BGP toolkit (free web tool!) supports both IRR as well as RPKI checks.
- Simply go to **bgp.he.net** and search with network name or AS number or prefix and you will see the status of prefixes.
-  Reflects when correct matching route object exists.
-  Reflects when parent route object exists (say for /17 or /21 etc when announcement is for /22).
-  Reflects when there is a mismatch of route objects.
-  Reflects when RPKI check is valid.
-  Reflects when RPKI check is invalid.

Check for IRR / RPKI ROA validation





































<u>104.20.128.0/20</u>	 	Cloudflare, Inc.	
<u>104.20.144.0/20</u>	 	Cloudflare, Inc.	
<u>104.20.160.0/20</u>	 	Cloudflare, Inc.	
<u>104.20.176.0/20</u>	 	Cloudflare, Inc.	

https://bgp.he.net/AS13335#_prefixes

<u>113.21.244.0/24</u>	 	True Internet Corporation Co. Ltd.	
<u>113.21.245.0/24</u>	 	True Internet Corporation Co. Ltd.	
<u>113.21.247.0/24</u>	 	True Internet Corporation Co. Ltd.	
<u>122.144.24.0/22</u>	 	True Internet Corporation Co. Ltd.	
<u>122.144.24.0/24</u>	 	True Internet Gateway Co.,Ltd.	
<u>122.144.25.0/24</u>	 	DIX	

https://bgp.he.net/AS38082#_prefixes

Check for IRR / RPKI ROA validation

49.228.0.0/14	 	Asia Pacific Network Information Centre	
49.228.0.0/16	 	408/60 PHP Bld. 15th Fl Phaholyothin Rd Samsen Nai Phayathai	
49.229.0.0/16	 	408/60 PHP Bld. 15th Fl Phaholyothin Rd Samsen Nai Phayathai	
49.230.0.0/16	 	408/60 PHP Bld. 15th Fl Phaholyothin Rd Samsen Nai Phayathai	
49.230.40.0/24	 	Assign for AIS_Internet Customers	
49.230.41.0/24	 		
49.230.42.0/24	 		
49.230.43.0/24	 		
49.231.0.0/16	 	Asia Pacific Network Information Centre	
49.231.44.0/24	 	408/60 PHP Bld. 15th Fl Phaholyothin Rd Samsen Nai Phayathai	
49.231.45.0/24	 	408/60 PHP Bld. 15th Fl Phaholyothin Rd Samsen Nai Phayathai	
49.231.46.0/24	 	408/60 PHP Bld. 15th Fl Phaholyothin Rd Samsen Nai Phayathai	

https://bgp.he.net/AS45430#_prefixes

Contribute in the cleanup!

How can you contribute?

- If you maintain resources (IPv4, IPv6 or AS numbers) then ensure to register route objects for them in either of databases - database of your RIR (in India? Contact IRINN, in Asia - do via My APNIC portal, in US - use ARIN portal).
- Create ROAs with your origin ASN and prefix length you intend to announce.
- Report all incorrect IRR entries you encounter to those registries to help them in removing old junk.
- DO NOT register route object on behalf of someone as a proxy entry as that has been a bad practice.
- If you have downstream ASNs behind you, register a AS SET.
- Register yourself on peeringdb.com portal and remember to mention your AS SET in the IRR section.

References

1. Tier 1 Networks Wikipedia Page - https://en.wikipedia.org/wiki/Tier_1_network#List_of_Tier_1_networks
2. BGP Version 4 RFC - <https://tools.ietf.org/html/rfc4271>
3. Root DNS servers list/locations - <https://root-servers.org/>
4. BGP in 2017 (APNIC Blog) - <https://blog.apnic.net/2018/01/10/bgp-in-2017/>
5. RSPL - <http://www.irr.net/docs/rpsl.html>
6. bgpq3 - <https://github.com/snar/bgpq3>
7. Hurricane Electric's route filtering algorithm - <http://routing.he.net/algorithm.html>
8. Google route filtering announcement NANOG75 - https://pc.nanog.org/static/published/meetings/NANOG75/1959/20190220_Morrow_Lightning_Talk_Prefix_v1.pdf
9. IRR (present one) - <https://github.com/irrdnet/irrd>
10. IRR v4 (in development) - <https://github.com/irrdnet/irrd4>
11. AT&T drops RPKI invalid for peers - <https://mailman.nanog.org/pipermail/nanog/2019-February/099501.html>

Questions/Comments?

Anurag Bhatia,
Hurricane Electric (AS6939)
anurag@he.net
Twitter: @anurag_bhatia
Web: <https://he.net>