Securing Internet Routing

Tashi Phuntsho (tashi@apnic.net)
Senior Network Analyst/Technical Trainer
Why should we bother?

• As a Manager
  – I don’t want to be front page news of a IT paper, or an actual newspaper for routing errors
How Verizon and a BGP Optimizer Knocked Large Parts of the Internet Offline Today

24 Jun 2019 by Tom Strickx.

Quick dumps through the data, showing about 2400 ASNs (networks) affected. Cloudflare being hit the hardest. Top 20 of affected ASNs below:

- sourceAS=13335
- sourceAS=4323
- sourceAS=7618
- sourceAS=63949
- sourceAS=2828
- sourceAS=26769
- sourceAS=209
- sourceAS=6428
- sourceAS=16509
- sourceAS=45899
- sourceAS=852
- sourceAS=12576
- sourceAS=20473
- sourceAS=54113
- sourceAS=55081
- sourceAS=2914

https://twitter.com/atoonk/status/1143143943531454464/photo/1
looking into BGP leak incident involving @google prefixes, AS37282 out of Niger and China Telecom.

3:40 AM - 13 Nov 2018

54 Retweets 48 Likes

MainOne @Mainoneservice

We have investigated the advertisement of @Google prefixes through one of our upstream partners. This was an error during a planned network upgrade due to a misconfiguration on our BGP filters. The error was corrected within 74mins & processes put in place to avoid reoccurrence

6:29 PM - 13 Nov 2018

36 Retweets 50 Likes

BREAKING: Potential hijack underway. ThousandEyes detected intermittent availability issues to Google services from some locations. Traffic to certain Google destinations appears to be routed through an ISP in Russia & blackholed at a China Telecom gateway router.

2:57 AM - 13 Nov 2018

609 Retweets 525 Likes

https://blog.thousandeyes.com/internet-vulnerability-takes-down-google/
Headlines

BORDER GATEWAY PROTOCOL ATTACK —

Suspicious event hijacks Amazon traffic for 2 hours, steals cryptocurrency

Almost 1,300 addresses for Amazon Route 53 rerouted for two hours.

DAN GOODIN - 4/25/2018, 1:30 AM

Kevin Beaumont         @GooseTheDog · Apr 24, 2018
MyEtherWallet subject to a DNS hijack. DNS was redirected via AWS DNS to a server in Russia, Ether stolen. Server is https only so users clicked through certificate errors.

Doug Madory            @DougMadory
Maybe related to this: twitter.com/InternetIntel/

InternetIntelligence   @InternetIntel
BGP hijack this morning affected Amazon DNS. eNet (AS10297) of Columbus, OH announced the following more-specifics of Amazon routes from 11:05 to 13:03 UTC today:
205.251.192.0/24
205.251.193.0/24
205.251.195.0/24
205.251.197.0/24
205.251.199.0/24

7:52 AM · 24 Apr 2018

https://blog.cloudflare.com/bgp-leaks-and-crypto-currencies
Headlines

Large BGP Leak by Google Disrupts Internet in Japan

Research // Aug 28, 2017 // Doug Madory

YouTube blames Pakistan network for 2-hour outage

Company appears to confirm reports that Pakistan Telecom was responsible for routing traffic according to erroneous Internet Protocols.

BY GREG SANDOVAL  |  FEBRUARY 24, 2008 10:15 PM PST

Pakistan hijacks YouTube

Research // Feb 24, 2008 // Dyn Guest Blogs
Why do we keep seeing these?

• Because NO ONE is in charge?
  – No single authority model for the Internet
    • No reference point for what’s right in routing
Why do we keep seeing these?

• Routing works by RUMOUR
  – Tell what you know to your neighbors, and Learn what your neighbors know
  – Assume everyone is correct (and honest)
    • Is the originating network the rightful owner?
Why do we keep seeing these?

- Routing is VARIABLE
  - The view of the network depends on where you are
    - Different routing outcomes at different locations
  - ~ no reference view to compare the local view 😞
Why do we keep seeing these?

• Routing works in REVERSE
  – Outbound advertisement affects inbound traffic
  – Inbound (*Accepted*) advertisement influence outbound traffic
Why do we keep seeing these?

• And as always, there is no E-bit
  – a bad routing update does not identify itself as BAD
    • RFC 3514 😊

• So tools/techniques try to identify GOOD updates
Why should we worry?

• Because it’s just so easy to do bad in routing!

Why should we bother?

• As a Engineer
  – I don’t want to be told at 3AM my routing is broken
Current Practice

Peering/Transit Request → LOA Check → Filters (in/out)
Tools & Techniques

• Look up **whois**
  – verify holder of a resource

```bash
tashi@tashi -> whois -h whois.apnic.net 202.125.96.0
% [whois.apnic.net]
% Whois data copyright terms http://www.apnic.net/db/dbcopyright.html
% Information related to '202.125.96.0 - 202.125.96.255'
% Abuse contact for '202.125.96.0 - 202.125.96.255' is 'training@apnic.net'

inetnum: 202.125.96.0 - 202.125.96.255
netname: APNICTRAINING-AP
descr: Prefix for APNICTRAINING LAB DC
country: AU
admin-c: AT480-AP
tech-c: AT480-AP
status: ALLOCATED NON-PORABLE
mnt-by: MAINT-AU-APNICTRAINING
mnt-irt: IRT-APNICTRAINING-AU
last-modified: 2016-06-17T00:17:28Z
source: APNIC

irt: IRT-APNICTRAINING-AU
address: 6 Cordelia Street
address: South Brisbane
address: QLD 4101
e-mail: training@apnic.net
abuse-mailbox: training@apnic.net
admin-c: AT480-AP
tech-c: AT480-AP
auth: # Filtered
mnt-by: MAINT-AU-APNICTRAINING
last-modified: 2013-10-31T11:01:10Z
source: APNIC
```

```bash
tashi@tashi -> whois -h whois.apnic.net 202.125.96.0
% [whois.apnic.net]
% Whois data copyright terms http://www.apnic.net/db/dbcopyright.html
% Information related to '202.125.96.0/24AS131107'
% Abuse contact for '202.125.96.0/24AS131107' is 'training@apnic.net'

inetnum: 202.125.96.0/24
netname: APNICTRAINING-AP
descr: Prefix for APNICTRAINING LAB DC
origin: AS131107
mnt-by: MAINT-AU-APNICTRAINING
country: AU
last-modified: 2016-06-16T23:23:00Z
source: APNIC
```
Tools & Techniques

• Ask for a **Letter of Authority**
  – Absolve from any liabilities
Tools & Techniques

• Look up/ask to enter details in **IRR**
  – describes route origination and inter-AS routing policies
Tools & Techniques

• **IRR**
  - Helps generate network (prefix & as-path) filters using RPSL tools
  - Filter out route advertisements not described in the registry
Tools & Techniques

• Problem(s) with IRR
  – No single authority model
    • How do I know if a RR entry is genuine and correct?
    • How do I differentiate between a current and a lapsed entry?
  – Many RRs
    • If two RRs contain conflicting data, which one do I trust and use?
  – Incomplete data - Not all resources are registered in an IRR
    • If a route is not in a RR, is the route invalid or is the RR just missing data?
  – Scaling
    • How do I apply IRR filters to upstream(s)?
Back to basics – identify GOOD

• Using digital signatures to convey the “authority to use”?
  – A private key to sign the authority, and
  – the public key to validate that authority
How about trust?

• Follows the resource allocation/delegation hierarchy
Chain of Trust - RPKI

Certificate chain mirrors the allocation hierarchy
Resource Certificates

• When an address holder A (*IRs) allocates resources \((IP\ \text{address}/ASN)\) to B (end holders)
  
  – A issues a resource certificate that binds the allocated address with B’s public key, all signed by A’s (CA) private key
  
  – proves the holder of the private key (B) is the legitimate holder of the resource!
Route Origin Authority

• B can now sign *authorities* using its private key, – which can be validated by any third party against the TA

• For routing, the address holder can *authorize* a network (ASN) to *originate* a route, and *sign* this permission with its private key (ROA)

<table>
<thead>
<tr>
<th>Prefix</th>
<th>202.144.128.0/19</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max-length</td>
<td>/24</td>
</tr>
<tr>
<td>Origin ASN</td>
<td>AS17660</td>
</tr>
</tbody>
</table>
Filtering with ROAs – Route Origin Validation

Global (RPKI) Repo

ROA
2406:6400::/32-48
17821

RPKI-to-Router (RtR)

RPKI Validator/
RPKI Cache server

AS17821

2406:6400::/48

17821

ASXXXX
Are ROAs enough?

• What if I forge the origin AS in the AS path?
  – Would be accepted as “good” – pass origin validation!

• Which means, we need to secure the AS path as well
  – need AS path validation (per-prefix)
A BGPsec speaker validates the received update by checking:

- If there is a ROA that describes the prefix and origin AS, and
- If the received AS path can be validated as a chain of signatures (for each AS in the AS path) using the AS keys
AS-PATH validation issues...

- More resources
  - CPU - high crypto overhead to validate signatures, and
  - Memory
    - Updates in BGPsec would be per prefix – update packing??
    - New attributes carrying signatures and certs/key-id for every AS in the AS path

- How do we distribute the certificates required?

- Can we have partial adoption?

- Given so much overhead, can it do more - Route leaks?
So, what can we do?

• Basic BGP OpSec hygiene – RFC7454/RFC8212
  – *RFC 8212* – BGP default reject or something similar
  – Filters with your *customers* and *peers*
    • Prefix filters, Prefix limit
    • AS-PATH filters, AS-PATH limit
    • Use IRR objects (source option) or ROA-to-IRR
  – Filter what you receive from your *upstream(s)*
  – Create ROAs for your resources
  – Filter inbound routes based on ROAs ~ ROV

• Join industry initiatives like MANRS
  • [https://www.manrs.org/](https://www.manrs.org/)
Industry Trends

AT&T/as7018 now drops invalid prefixes from peers

Jay Borckenagen jayb at braeburn.org
Mon Feb 11 14:53:45 UTC 2019

• Previous message (by thread): BGP topological vs
• Next message (by thread): AT&T/as7018 now drop
• Messages sorted by: [ date ] [ thread ] [ subject ]

Hello all,

In November 2018 during the ZAPF (South Africa Peering Forum) meeting in Cape Town, 3 major ISPs in Africa announced that they would enable BGP’s “Route Origin” and the dropping of invalid routes as part of an effort to clean up the BGP Internet, on the 1st April, 2019.

On the 1st of April, Workline Communications (AS37271) enabled ROV and the dropping of Invalid routes. This applies to all eBGP sessions for IPv4 and IPv6.

On the 5th of April, SEACOM (AS65790) enabled ROV and the dropping of Invalid routes. This applies to all eBGP sessions with public peers, private peers and transit providers...

IPv4 and IPv6 eBGP sessions toward downstream customers will follow in 3 months from now.

We are still standing by for the 3rd ISP to complete their implementation, and we are certain they will communicate with the community accordingly.

Please note that for the legal reasons previously discussed on various fora, neither Workline Communications nor SEACOM are utilising the ARIN TAL. As a result, any valid or invalid prefixes on our networks will be resolved. We will each re-evaluate this decision if and when ARIN’s policy changes. We are hopeful that this will happen soon.

If you interconnect with either of us and may be experiencing any routing issues potentially related to this new policy, please feel free to reach out to:

• noc@workline.africa
• peering@seacom.ru

Workline Communications and SEACOM hope that this move encourages the rest of the ISP community around the world to ramp up their deployment of BGP ROV and Invalid routes, as we appreciate the work that AT&T have carried out in the same vein.

In the mean time, we are happy to answer any questions you may have about our deployments. Thanks.

Mark Tinka (SEACOM) & Ben Maddison (Workline Communications).
Acknowledgement

• Geoff Huston, APNIC
• Randy Bush, IIJ Labs/Arrcus
THANK YOU