BGP Flexibility and...  
Its Consequences  

Alexander Azimov  
Qrator Labs aa@qrator.net
Rights

• Get address space from RIR;
• Establish BGP sessions;
• Advertise prefixes;
• Receive traffic.
Obligations

• MUST not hijack foreign address space;
• MUST not create route leaks;
• MUST Support anti-spoofing policies;
• MUST Configure Ingress/Egress filters;
• MUST Pay fee to RIR.
Obligations

- MUST not hijack foreign address space;
- MUST not create route leaks;
- MUST support anti-spoofing policies;
- MUST Configure Ingress/Egress filters;
- MUST Pay fee to RIR – that’s all!
BGP Anomalies

BGP Hijacks
Illegitimate advertisement of foreign address space.

BGP Route Leaks
Illegitimate announce of a route received from peer or upstream to another peer or upstream.
It will never happen again!
BGP in the Headlines

Data Centre ▶ Networks

Google routing blunder sent Japan's Internet dark on Friday

Another big BGP blunder

Google linked to internet disruptions in Japan
KYODO, STAFF REPORT

Google hijack made Japan 'land of no internet' for more than 30 minutes

Google accidentally broke the internet throughout Japan

A mistake led to internet outages for about half of the country.
24th April
AS10297, ENET, US

BGP Hijack

+ 

DNS Hijack

= 

Money

26th April
AS267286, DPSTL, Brazil

~268 million unique IPs

https://radar.qrator.net/blog/bgp-hijacks-malicious-or-mistakes
4th May
AS8449, ElCat, KYRGYZSTAN

Route Leak – Connecting People! Oh not?
BGP Exams?

I will never create BGP hijack again.
I will never create BGP hijack again.
I will never create BGP hijack again.
I will never create BGP hijack again.
I will never create BGP hijack again.
I will never create BGP hijack again.
I will never create BGP hijack again.
I will never create BGP hijack again.
I will never create BGP hijack again.
I will never create BGP hijack again.
I will never create BGP hijack again.
BGP Ingress Filtering: AS-SETs

from AS197068 action accept AS-QRATOR

AS-SET to ASNs
200449, 197068, .......

ASNs to route objects
178.248.232.0/21, 185.65.148.0/22, 45.116.91.0/24, 89.218.31.0/24 ...

IRR filters doesn’t perform origin validation!
AS-SETs & AS Cone

In ‘Ideal World’ AS-SETs = AS Cone
AS-SETs & AS Cone

In ‘Ideal World’ AS-SETs = AS Cone

Not in AS-SET? Drop it!
AS-SETs & AS Cone

In an ‘Ideal World’ AS-SETs = AS Cone
But even then it has limitations.
What Do We Know?

• IRR filters doesn’t perform origin validation;
• AS-SET objects are not authorized (even in RIPE);
• Poorly maintained AS-SETs become less affective;
• There are ISPs does not use any IRR filters.
Investigation

How many filters have been already violated?
Methodology

- Route Object Aggregator (RIPE, APNIC, ARIN, AFRINIC, RADB... 27 sources);
- Analyze prefixes with unique asn in route objects;
- Detect c2p links through which route leaks were propagated;
- Check that origin doesn’t belong to customer cone.
Results

IPv4

At least 7% of ISPs have problems with filters

IPv6

At least 1% of ISPs have problems with filters
Results: Explained
Accepts Leaks originated by Tier1

### IPv4
- 4809 - China Telecom
- 4837 - China Unicom
- 6695 - DECIX
- 6939 - HE
- 7363 - YOMURA
- 7552 - Viettel
- 7713 - Telkomnet
- 7843 - TWCABLE
- **8732** - Comcor
- 9583 - Sify
- **12389** - Rostelecom
- 12586 - GHOSTnet

### IPv6
- 13536 - TVC-AS1
- 20485 - TTK
- 20562 - Opentransit
- 22356 - Durand
- 22773 - CXA
- 31025 - GHOSTnet
- 35104 - Kaztranscom
- 40805 - JMFSOLUTIONS
- **48276** - LEVEL-AS
- 50384 - W-IX
- 53211 - ISPRJ

### IPv4
- 6939 - HE
- 16735 - ALGAR
- 23106 - Cemig
- 49697 - Joey-Network
- 199524 - G-Core
Accepts Leaks originated by **Tier1**

## China

<table>
<thead>
<tr>
<th>IPv4</th>
<th>IPv6</th>
</tr>
</thead>
<tbody>
<tr>
<td>4809 - China Telecom</td>
<td>13536 - TVC-AS1</td>
</tr>
<tr>
<td>4837 - China Unicom</td>
<td>20485 - TTK</td>
</tr>
<tr>
<td>6695 - DECIX</td>
<td>20562 - Opentransit</td>
</tr>
<tr>
<td>6939 - HE</td>
<td>22356 - Durand</td>
</tr>
<tr>
<td>7363 - YOMURA</td>
<td>22773 - CXA</td>
</tr>
<tr>
<td>7552 - Viettel</td>
<td>31025 - GHOSTnet</td>
</tr>
<tr>
<td>7713 - Telkomnet</td>
<td>35104 - Kaztranscom</td>
</tr>
<tr>
<td>7843 - TWCABLE</td>
<td>40805 - JMFSOLUTIONS</td>
</tr>
<tr>
<td>8732 - Comcor</td>
<td>48276 - LEVEL-AS</td>
</tr>
<tr>
<td>9583 - Sify</td>
<td>50384 - W-IX</td>
</tr>
<tr>
<td>12389 - Rostelecom</td>
<td>53211 - ISPRJ</td>
</tr>
<tr>
<td>12586 - GHOSTnet</td>
<td>199524 - G-Core</td>
</tr>
<tr>
<td>16735 - ALGAR</td>
<td>49697 - Joey-Network</td>
</tr>
<tr>
<td>23106 - Cemig</td>
<td>6939 - HE</td>
</tr>
<tr>
<td>40805 - JMFSOLUTIONS</td>
<td>16735 - ALGAR</td>
</tr>
<tr>
<td>49697 - Joey-Network</td>
<td>23106 - Cemig</td>
</tr>
<tr>
<td>199524 - G-Core</td>
<td>49697 - Joey-Network</td>
</tr>
</tbody>
</table>

- IPv4
- IPv6
Accepts Leaks originated by **Tier 1**

**IXes**

<table>
<thead>
<tr>
<th>IPv4</th>
<th>IPv6</th>
</tr>
</thead>
<tbody>
<tr>
<td>4809 - China Telecom</td>
<td>13536 - TVC-AS1</td>
</tr>
<tr>
<td>4837 - China Unicom</td>
<td>20485 - TTK</td>
</tr>
<tr>
<td><strong>6695 - DECIX</strong></td>
<td><strong>20562 - Opentransit</strong></td>
</tr>
<tr>
<td>6939 - HE</td>
<td>6939 - HE</td>
</tr>
<tr>
<td>7363 - YOMURA</td>
<td>16735 - ALGAR</td>
</tr>
<tr>
<td>7552 - Viettel</td>
<td>23106 - Cemig</td>
</tr>
<tr>
<td>7713 - Telkomnet</td>
<td>49697 - Joey-Network</td>
</tr>
<tr>
<td>7843 - TWCABLE</td>
<td>199524 - G-Core</td>
</tr>
<tr>
<td>8732 - Comcor</td>
<td></td>
</tr>
<tr>
<td>9583 - Sify</td>
<td></td>
</tr>
<tr>
<td>12389 - Rostelecom</td>
<td>50384 - W-IX</td>
</tr>
<tr>
<td>12586 - GHOSTnet</td>
<td>53211 - ISPRJ</td>
</tr>
</tbody>
</table>

Additionally, IPv6 addresses:**

- 6939 - HE
- 16735 - ALGAR
- 23106 - Cemig
- 49697 - Joey-Network
- 199524 - G-Core
AcceptsLeaks originated by **Tier1**

**Hurricane Electric!**

<table>
<thead>
<tr>
<th>IPv4</th>
<th>IPv6</th>
</tr>
</thead>
<tbody>
<tr>
<td>4809 - China Telecom</td>
<td>13536 - TVC-AS1</td>
</tr>
<tr>
<td>4837 - China Unicom</td>
<td>20485 - TTK</td>
</tr>
<tr>
<td>6695 - DECIX</td>
<td>20562 - Opentransit</td>
</tr>
<tr>
<td><strong>6939 - HE</strong></td>
<td>22356 - Durand</td>
</tr>
<tr>
<td>7363 - YOMURA</td>
<td>22773 - CXA</td>
</tr>
<tr>
<td>7552 - Viettel</td>
<td>31025 - GHOSTnet</td>
</tr>
<tr>
<td>7713 - Telkomnet</td>
<td>35104 - Kaztranscom</td>
</tr>
<tr>
<td>7843 - TWCABLE</td>
<td>40805 - JMFSOLUTIONS</td>
</tr>
<tr>
<td>8732 - Comcor</td>
<td>48276 - LEVEL-AS</td>
</tr>
<tr>
<td>9583 - Sify</td>
<td>50384 - W-IX</td>
</tr>
<tr>
<td>12389 - Rostelecom</td>
<td>53211 - ISPRJ</td>
</tr>
<tr>
<td>12586 - GHOSTnet</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>6939 - HE</strong></td>
</tr>
<tr>
<td></td>
<td>16735 - ALGAR</td>
</tr>
<tr>
<td></td>
<td>23106 - Cemig</td>
</tr>
<tr>
<td></td>
<td>49697 - Joey-Network</td>
</tr>
<tr>
<td></td>
<td>199524 - G-Core</td>
</tr>
</tbody>
</table>
Key Findings: AS-SETs

AS-SET Can be Used to:

• Filter some hijacks;
• Filter some route leaks.

In reality:

• Many AS-SETs are poorly maintained;
• No filters at some huge Tier-2 networks;
• Even some Tier1 networks fail to configure filters.
Resource Public Key Infrastructure
check_roa(asn, prefix)

Valid – prefix matches ROA with corresponding asn and within maximum length;

Invalid – ROA(s) record exists, but without corresponding asn or exceeds maximum length;

Unknown – ROA that covers prefix does not exist.

But you can’t retrieve all valid prefixes for an asn!
Integration with AS-SET? Ooops...
ROA Validation: AS_PATH Violation

ROA (178.248.232.0/21, 197068, 32)

ASYYY

178.248.232.0/21
AS_PATH: ASXXX AS197068

ROA check OK!
ROA Records

IPv4

IPv6

Valid  Invalid  No object

Valid  Invalid  No object
No ROA validation – no ROAs?
Key Findings: ROA validation

ROA Validation Can be Used to:
• filter mistake hijacks;

ROA Validation Can’t:
• filter route leaks;
• filter malicious hijacks.

In reality:
• Only 10% of prefixes are signed;
• Transit ISPs doesn’t perform origin validations.
### BGP Quadrant

<table>
<thead>
<tr>
<th>Mistake</th>
<th>Hijacks</th>
<th>Route Leaks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malicious</td>
<td><strong>BGPSec</strong></td>
<td><strong>BGPSec</strong></td>
</tr>
</tbody>
</table>
28.09.2017

BGPsec Protocol Specification
RFC 8205

<table>
<thead>
<tr>
<th>Status</th>
<th>IESG evaluation record</th>
<th>IESG writeups</th>
<th>Email expansions</th>
<th>History</th>
</tr>
</thead>
<tbody>
<tr>
<td>Versions: 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

draft-lepinski-bgpsec-protocol 0 00
draft-bgpsec-spec 00 01 02 03 04 05 06 07 08 09 10 11 13 14 15 17 18 22 22 23
rfc8205

RFC 8205: BGPsec Protocol Specification
RFC 8206: BGPsec Considerations for Autonomous System (AS) Migration
RFC 8207: BGPsec Operational Considerations
RFC 8208: BGPsec Algorithms, Key Formats, and Signature Formats
RFC 8209: A Profile for BGPsec Router Certificates, Certificate Revocation Lists, and Certification Requests
RFC 8210: The Resource Public Key Infrastructure (RPKI) to Router
RFC 8211: Adverse Actions by a Certification Authority (CA) or Repository Manager in the Resource Public Key Infrastructure (RPKI)
AS_PATH Validation

ASN1
- pCount
- Flags
- Target ASN2
- Signature

ASN2
- pCount
- Flags
- Target ASN3
- Signature

ASN3
- pCount
- Flags
- Target ASN4
- Signature

Everything is signed and validated, no more hijacks! Right?
BGPSec: Backward Compatibility

ROA (178.248.232.0/21, 197068, 32)

ASXXX

I don’t know BGPSec

178.248.232.0/21
AS_PATH: ASXXX AS197068

ASYYY

Ok, plain BGP.

ROA check OK!
Key Findings: BGPSec

BGPSec Can be Used to:
• make things more complicated.

In reality:
It seems that authors forgot the goal.
# BGP Quadrant

<table>
<thead>
<tr>
<th></th>
<th>Hijacks</th>
<th>Route Leaks</th>
</tr>
</thead>
</table>
| Mistake | IRR filters; ROA validation.         | IRR filters; [draft-ietf-idr-bgp-open-policy](https://datatracker.ietf.org/doc/draft-ietf-idr-bgp-open-policy/)  
                  [draft-ymbk-idr-bgp-eotr-policy](https://datatracker.ietf.org/doc/draft-ymbk-idr-bgp-eotr-policy/)  
| Malicious | ![warning](https://i.imgur.com/2Q5Q5Q5.png) | ![warning](https://i.imgur.com/2Q5Q5Q5.png) |
What can YOU do?
What Can Transit Do?

- IRR filters at your customer links, no exceptions!
- Work with customers, that corrupt AS-SETs;
- Consider using IRR filters with your private peers;
- Ad-hoc filtering ([NTT Peering Lock](http://www.ntt.com/peering-lock));
- Perform constant BGP monitoring.
What Can IX Do?

- IRR filters at all links, **no exceptions**!
- **Work with customers**, that corrupt AS-SETs.
- Remove legacy filters (LOA, route-sets);
- Consider using ROA validation at RS;
What Can Multihomed Do?

- Keep Route Objects up to date;
- Keep AS-SETs up to date;
- Create ROA records ([RIPE Wizard](#));
- Perform constant BGP monitoring.