BGP Best Current Practices

ISP Workshops

Last updated 5th September 2019

Acknowledgements

- This material originated from the Cisco ISP/IXP Workshop Programme developed by Philip Smith & Barry Greene
- Use of these materials is encouraged as long as the source is fully acknowledged and this notice remains in place
- Bug fixes and improvements are welcomed
  - Please email *workshop (at) bgp4all.com*

Philip Smith

Configuring BGP

Where do we start?

Cisco IOS Good Practices

- ISPs should start off with the following BGP commands as a basic template:

```
router bgp 64511
 bgp deterministic-med
 no bgp default ipv4-unicast
 distance bgp 200 200 200
 no synchronization
 no auto-summary
```

Replace with public ASN

Turn off IOS assumption that all neighbours will exchange IPv4 prefixes

Make ebgp and ibgp distance the same & more than any IGP

EBGP Configuration Best Practices

- Industry standard is described in RFC8212
  - https://tools.ietf.org/html/rfc8212
  - External BGP (EBGP) Route Propagation Behaviour without Policies
- NB: BGP in Cisco IOS is permissive by default
  - This is contrary to industry standard and RFC8212
- Configuring BGP peering without using filters means:
  - All best paths on the local router are passed to the neighbour
  - All routes announced by the neighbour are received by the local router
  - Can have disastrous consequences (see RFC8212)

EBGP Configuration Best Practices

- Best practice is to ensure that each eBGP neighbour has inbound and outbound filter applied:

```
router bgp 64511
 address-family ipv4
  neighbor 100.64.0.1 remote-as 64510
  neighbor 100.64.0.1 prefix-list as64510-in in
  neighbor 100.64.0.1 prefix-list as64510-out out
  neighbor 100.64.0.1 activate
```

What is BGP for??

What is an IGP not for?

BGP versus OSPF/ISIS

- Internal Routing Protocols (IGPs)
  - Examples are IS-IS and OSPF
  - Used for carrying **infrastructure** addresses
  - NOT used for carrying Internet prefixes or customer prefixes
  - Design goal is to **minimise** number of prefixes in IGP to aid **scalability** and **rapid convergence**

BGP versus OSPF/IS-IS

- BGP is used
  - Internally (iBGP)
  - Externally (eBGP)
- iBGP is used to carry:
  - Some/all Internet prefixes across backbone
  - Customer prefixes
- eBGP is used to:
  - Exchange prefixes with other ASes
  - Implement routing policy

BGP versus OSPF/IS-IS

- DO NOT:
  - Distribute BGP prefixes into an IGP
  - Distribute IGP routes into BGP
  - Use an IGP to carry customer prefixes
- **YOUR NETWORK WILL NOT SCALE**

Aggregation

Aggregation

- Aggregation means announcing the address block received from the RIR to the other ASes connected to your network
- Subprefixes of this aggregate may be:
  - Used internally in the ISP network
  - Announced to other ASes to aid with multihoming
- Too many operators are still thinking about class Cs, resulting in a proliferation of /24s in the Internet routing table
  - July 2019: 436208 /24s in IPv4 table of 762552 prefixes
- The same is happening for /48s with IPv6
  - July 2019: 34203 /48s in IPv6 table of 71862 prefixes

Configuring Aggregation – Cisco IOS

- ISP has 100.66.0.0/19 address block
- To put into BGP as an aggregate:

```
router bgp 64511
 address-family ipv4
  network 100.66.0.0 mask 255.255.224.0
ip route 100.66.0.0 255.255.224.0 null0
```

- The static route is a “pull up” route
  - More specific prefixes within this address block ensure connectivity to ISP’s customers
  - “Longest match” lookup

Aggregation

- Address block should be announced to the Internet as an aggregate
- Subprefixes of address block should NOT be announced to Internet unless for traffic engineering
  - See BGP Multihoming presentations
- Aggregate should be generated internally
  - Not on the network borders!

Announcing Aggregate – Cisco IOS

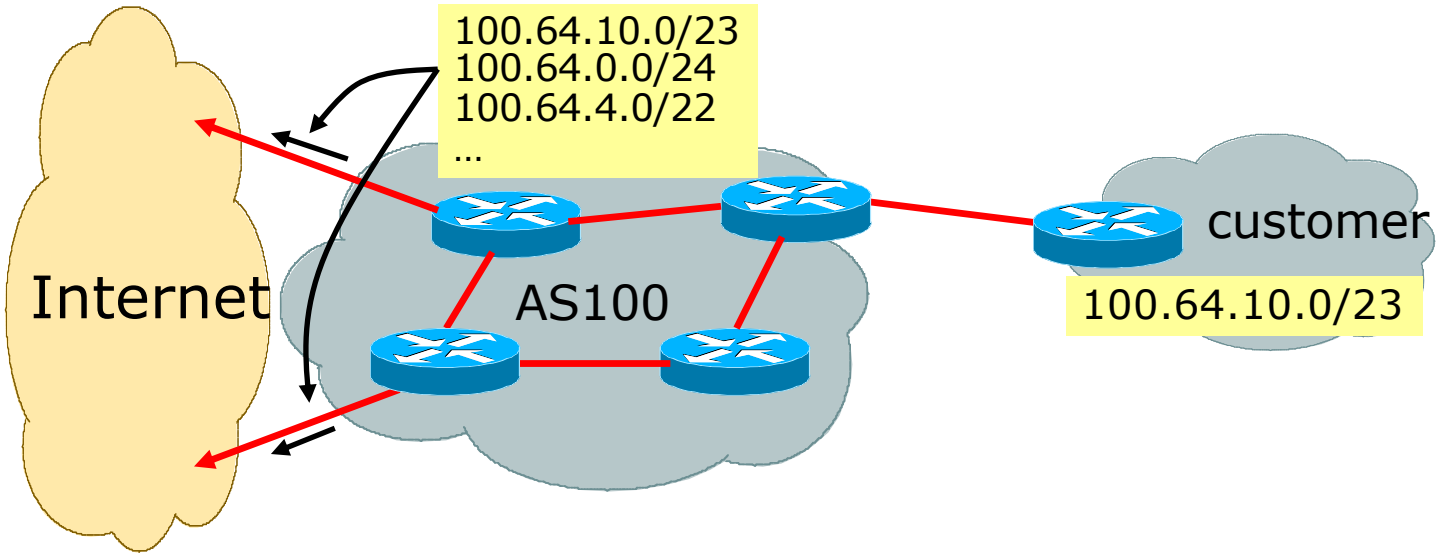
- Configuration Example

```
router bgp 64511
 address-family ipv4
  network 100.66.0.0 mask 255.255.224.0
  neighbor 100.67.10.1 remote-as 101
  neighbor 100.67.10.1 prefix-list out-filter out
  neighbor 100.67.10.1 prefix-list default in
  neighbor 100.67.10.1 activate
!
ip route 100.66.0.0 255.255.224.0 null0
!
ip prefix-list out-filter permit 100.66.0.0/19
ip prefix-list out-filter deny 0.0.0.0/0 le 32
!
ip prefix-list default permit 0.0.0.0/0
```

Announcing an Aggregate

- ISPs who don't and won't aggregate are held in poor regard by community
- Registries publish their minimum allocation size
  - For IPv4:
    - /24
  - For IPv6:
    - /48 for assignment, /32 for allocation
- Until 2010, there was no real reason to see anything longer than a /22 IPv4 prefix in the Internet. But now?
  - IPv4 run-out is having an impact

Aggregation – Example

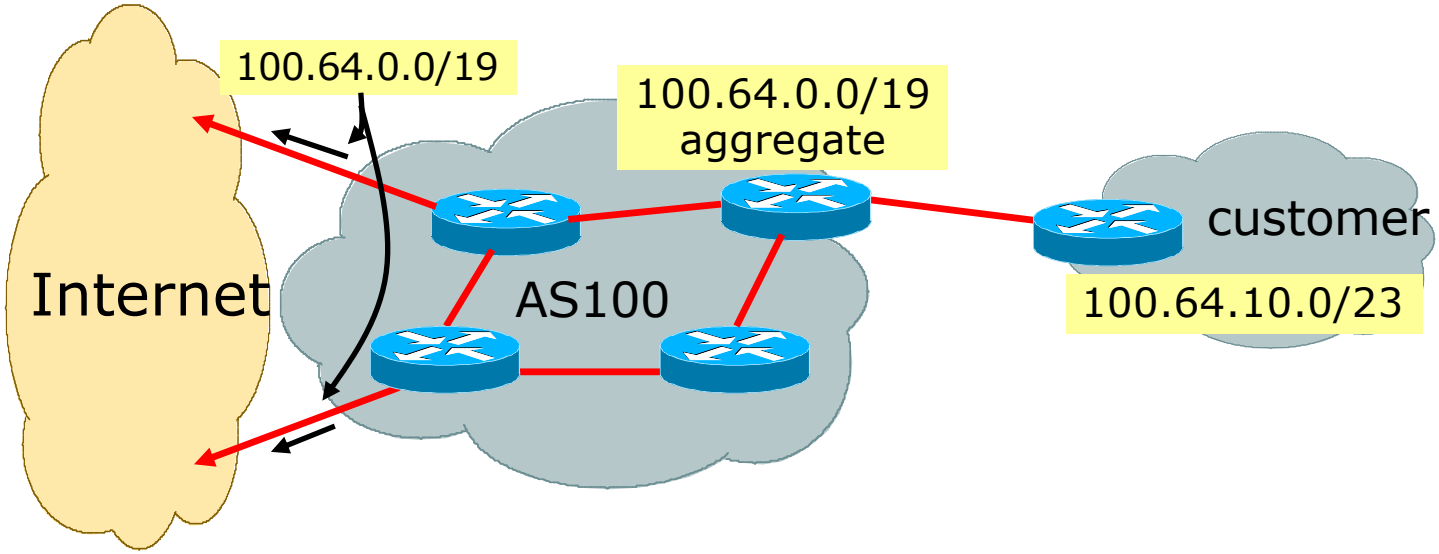


- Customer has /23 network assigned from AS100's /19 address block
- AS100 announces customers' individual networks to the Internet

Aggregation – Bad Example

- Customer link goes down
  - Their /23 network becomes unreachable
  - /23 is withdrawn from AS100's iBGP
- Their ISP doesn't aggregate its /19 network block
  - /23 network withdrawal announced to peers
  - Starts rippling through the Internet
  - Added load on all Internet backbone routers as network is removed from routing table
- Customer link returns
  - Their /23 network is now visible to their ISP
  - Their /23 network is re-advertised to peers
  - Starts rippling through Internet
  - Load on Internet backbone routers as network is reinserted into routing table
  - Some ISP's suppress the flaps
  - Internet may take 10-20 min or longer to be visible
  - Where is the Quality of Service???

Aggregation – Example



- Customer has /23 network assigned from AS100's /19 address block
- AS100 announced /19 aggregate to the Internet

Aggregation – Good Example

- Customer link goes down
  - Their /23 network becomes unreachable
  - /23 is withdrawn from AS100's iBGP
- /19 aggregate is still being announced
  - No BGP hold down problems
  - No BGP propagation delays
  - No damping by other ISPs

- Customer link returns
- Their /23 network is visible again
  - The /23 is re-injected into AS100's iBGP
- The whole Internet becomes visible immediately
- Customer has Quality of Service perception

Aggregation – Summary

- Good example is what everyone should do!
  - Adds to Internet stability
  - Reduces size of routing table
  - Reduces routing churn
  - Improves Internet QoS for **everyone**
- Bad example is what too many still do!
  - Why? Lack of knowledge?
  - Laziness?

Separation of iBGP and eBGP

- Many ISPs do not understand the importance of separating iBGP and eBGP
  - iBGP is where all customer prefixes are carried
  - eBGP is used for announcing aggregate to Internet and for Traffic Engineering
- Do NOT do traffic engineering with customer originated iBGP prefixes
  - Leads to instability similar to that mentioned in the earlier bad example
  - Even though aggregate is announced, a flapping subprefix will lead to instability for the customer concerned
- Generate traffic engineering prefixes on the Border Router

The Internet Today (July 2019)

- Current IPv4 Internet Routing Table Statistics

BGP Routing Table Entries	762552
Prefixes after maximum aggregation	294687
Unique prefixes in Internet	368695
/24s announced	436208
ASNs in use	65001

  - (maximum aggregation is calculated by Origin AS)
  - (unique prefixes > max aggregation means that operators are announcing aggregates from their blocks without a covering aggregate)

Efforts to improve aggregation

- The CIDR Report
  - Initiated and operated for many years by Tony Bates
  - Now combined with Geoff Huston's routing analysis
    - www.cidr-report.org
    - (covers both IPv4 and IPv6 BGP tables)
  - Results e-mailed on a weekly basis to most operations lists around the world
  - Lists the top 30 service providers who could do better at aggregating
- RIPE Routing WG aggregation recommendations
  - IPv4: RIPE-399 — www.ripe.net/ripe/docs/ripe-399.html
  - IPv6: RIPE-532 — www.ripe.net/ripe/docs/ripe-532.html

Efforts to Improve Aggregation
The CIDR Report

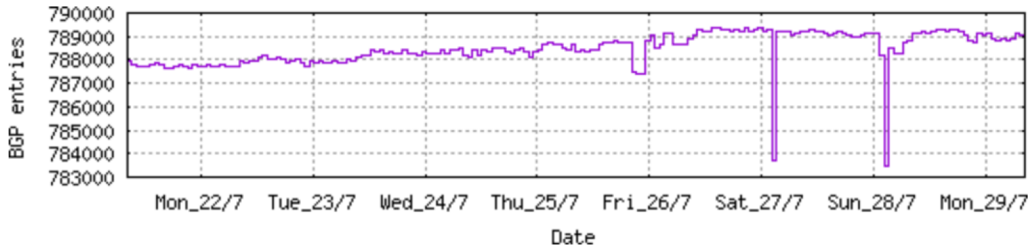
- Also computes the size of the routing table assuming ISPs performed optimal aggregation
- Website allows searches and computations of aggregation to be made on a per AS basis
  - Flexible and powerful tool to aid ISPs
  - Intended to show how greater efficiency in terms of BGP table size can be obtained without loss of routing and policy information
  - Shows what forms of origin AS aggregation could be performed and the potential benefit of such actions to the total table size
  - Very effectively challenges the traffic engineering excuse

Status Summary

Table History

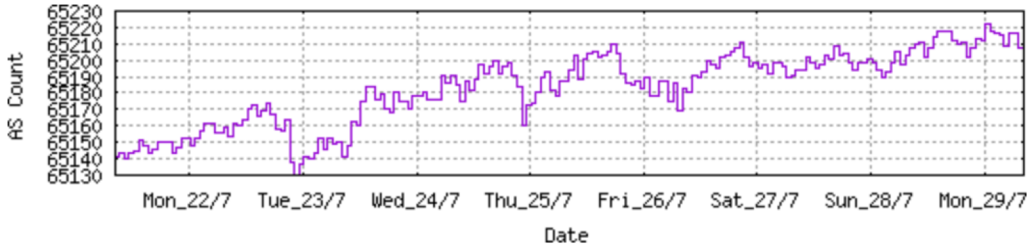
Date	Prefixes	CIDR Aggregated
22-07-19	787721	431337
23-07-19	787946	431503
24-07-19	788320	431918
25-07-19	788351	432129
26-07-19	788783	432386
27-07-19	789374	432301
28-07-19	789131	432251
29-07-19	789064	432141

Plot: BGP Table Size



AS Summary

65220	Number of ASes in routing system
23744	Number of ASes announcing only one prefix
6265	Largest number of prefixes announced by an AS AS8151: Uninet S.A. de C.V., MX
115794432	Largest address span announced by an AS (/32s) AS4134: CHINANET-BACKBONE No.31,Jin-rong Street, CN



Plot: AS count
Plot: Average announcements per origin AS
Report: ASes ordered by originating address span
Report: ASes ordered by transit address span
Report: Autonomous System number-to-name mapping (from Registry WHOIS data)

Announced Prefixes

```
Rank  AS       Type    Originate Addr Space  (pfx)   Transit Addr space  (pfx)  Description
41    AS6389         ORG+TRN Originate:   16211456 /8.05   Transit:      93696 /15.48 BELLSOUTH-NET-BLK - AT&T Corp., US
```

Aggregation Suggestions

Filter: Aggregates, Specifics

This report does not take into account conditions local to each origin AS in terms of policy or traffic engineering requirements, so this is an approximate guideline as to aggregation possibilities.

Rank	AS	AS Name	Current	Wthdw	Aggte	Annce	Redctn	%
23	AS6389	BELLSOUTH-NET-BLK - AT&T Corp., US	1560	1495	8	73	1487	95.32%

Prefix	AS Path	Aggregation Suggestion
12.81.90.0/23	4777 2516 3356 7018 6389	
12.81.120.0/24	4777 2516 3356 7018 6389	
65.4.0.0/14	4777 2516 3356 7018 6389	
65.5.1.0/24	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.12.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.20.0/23	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.21.0/24	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.24.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.32.0/20	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.34.0/24	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.46.0/24	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.57.0/24	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.64.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.68.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.80.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.88.0/21	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.118.0/23	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.120.0/21	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.133.0/24	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.136.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.140.0/23	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.141.0/24	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.148.0/23	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.150.0/23	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.152.0/21	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.156.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.160.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.164.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.172.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.200.0/21	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.228.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.232.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.236.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.240.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.244.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.248.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.5.252.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.6.192.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389
65.6.196.0/22	4777 2516 3356 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4777 2516 3356 7018 6389

Announced Prefixes

```
Rank  AS       Type    Originate Addr Space  (pfx)   Transit Addr space  (pfx)  Description
195   AS18566        ORG+TRN Originate:    3190528 /10.39  Transit:      11008 /18.57 MEGAPATH5-US - MegaPath Corporation, US
```

Aggregation Suggestions

Filter: Aggregates, Specifics

This report does not take into account conditions local to each origin AS in terms of policy or traffic engineering requirements, so this is an approximate guideline as to aggregation possibilities.

Rank	AS	AS Name	Current	Wthdw	Aggte	Annce	Redctn	%
17	AS18566	MEGAPATH5-US - MegaPath Corporation, US	2093	1809	65	349	1744	83.33%

Prefix	AS Path	Aggregation Suggestion
64.6.160.0/23	4777 2516 3257 18566	
64.6.164.0/22	4777 2516 3257 18566	+ Announce - aggregate of 64.6.164.0/23 (4777 2516 3257 18566) and 64.6.166.0/23 (4777 2516 3257 18566)
64.6.164.0/23	4777 2516 3257 18566	- Withdrawn - aggregated with 64.6.166.0/23 (4777 2516 3257 18566)
64.6.166.0/24	4777 2516 3257 18566	- Withdrawn - aggregated with 64.6.167.0/24 (4777 2516 3257 18566)
64.6.167.0/24	4777 2516 3257 18566	- Withdrawn - aggregated with 64.6.166.0/24 (4777 2516 3257 18566)
64.50.206.0/23	4777 2516 3257 18566	
64.51.126.0/23	4777 2516 3257 18566	
64.81.0.0/16	4777 2516 3257 18566	
64.81.16.0/22	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.20.0/22	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.22.0/24	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.24.0/22	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.28.0/22	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.32.0/20	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.32.0/24	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.33.0/24	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.34.0/24	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.35.0/24	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.36.0/24	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.37.0/24	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.38.0/24	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.39.0/24	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.40.0/24	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.44.0/24	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.48.0/20	4777 2516 3257 18566	- Withdrawn - matching aggregate 64.81.0.0/16 4777 2516 3257 18566
64.81.48.0/23	4777 2516 3356 18566	+ Announce - aggregate of 64.81.48.0/24 (4777 2516 3356 18566) and 64.81.49.0/24 (4777 2516 3356 18566)
64.81.48.0/24	4777 2516 3356 18566	- Withdrawn - aggregated with 64.81.49.0/24 (4777 2516 3356 18566)
64.81.49.0/24	4777 2516 3356 18566	- Withdrawn - aggregated with 64.81.48.0/24 (4777 2516 3356 18566)
64.81.50.0/24	4777 2516 3356 18566	
64.81.52.0/22	4777 2516 3356 18566	+ Announce - aggregate of 64.81.52.0/23 (4777 2516 3356 18566) and 64.81.54.0/23 (4777 2516 3356 18566)
64.81.52.0/24	4777 2516 3356 18566	- Withdrawn - aggregated with 64.81.53.0/24 (4777 2516 3356 18566)
64.81.53.0/24	4777 2516 3356 18566	- Withdrawn - aggregated with 64.81.52.0/24 (4777 2516 3356 18566)
64.81.54.0/24	4777 2516 3356 18566	- Withdrawn - aggregated with 64.81.55.0/24 (4777 2516 3356 18566)
64.81.55.0/24	4777 2516 3356 18566	- Withdrawn - aggregated with 64.81.54.0/24 (4777 2516 3356 18566)
64.81.56.0/22	4777 2516 3356 18566	+ Announce - aggregate of 64.81.56.0/23 (4777 2516 3356 18566) and 64.81.58.0/23 (4777 2516 3356 18566)
64.81.56.0/24	4777 2516 3356 18566	- Withdrawn - aggregated with 64.81.57.0/24 (4777 2516 3356 18566)
64.81.57.0/24	4777 2516 3356 18566	- Withdrawn - aggregated with 64.81.56.0/24 (4777 2516 3356 18566)
64.81.58.0/24	4777 2516 3356 18566	- Withdrawn - aggregated with 64.81.59.0/24 (4777 2516 3356 18566)
64.81.59.0/24	4777 2516 3356 18566	- Withdrawn - aggregated with 64.81.58.0/24 (4777 2516 3356 18566)

Announced Prefixes

```
Rank  AS       Type    Originate Addr Space  (pfx)   Transit Addr space  (pfx)  Description
170   AS7545         ORG+TRN Originate:    3920156 /10.10  Transit:   81247232 /5.72  TPG-INTERNET-AP TPG Telecom Limited, AU
```

Aggregation Suggestions

Filter: Aggregates, Specifics

This report does not take into account conditions local to each origin AS in terms of policy or traffic engineering requirements, so this is an approximate guideline as to aggregation possibilities.

```
 Rank AS           AS Name                                    Current  Wthdw  Aggte  Annce Redctn       %
   13 AS7545       TPG-INTERNET-AP TPG Telecom Limited, AU       4863   2903    877   2837   2026  41.66%


  Prefix             AS Path                          Aggregation Suggestion
  14.2.0.0/19        4608 4739 7545
  14.2.64.0/18       4608 4739 7545       + Announce - aggregate of 14.2.64.0/19 (4608 4739 7545) and 14.2.96.0/19 (4608 4739 7545)
  14.2.64.0/19       4608 4739 7545       - Withdrawn - aggregated with 14.2.96.0/19 (4608 4739 7545)
  14.2.96.0/19       4608 4739 7545       - Withdrawn - aggregated with 14.2.64.0/19 (4608 4739 7545)
  14.2.192.0/20      4608 4739 7545
  14.200.0.0/14      4608 1221 2764 7545
  14.200.0.0/21      4608 7575 7545 7545  + Announce - aggregate of 14.200.0.0/22 (4608 7575 7545 7545) and 14.200.4.0/22 (4608 7575 7545 7545)
  14.200.0.0/24      4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.1.0/24 (4608 7575 7545 7545)
  14.200.1.0/24      4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.0.0/24 (4608 7575 7545 7545)
  14.200.2.0/24      4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.3.0/24 (4608 7575 7545 7545)
  14.200.3.0/24      4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.2.0/24 (4608 7575 7545 7545)
  14.200.4.0/24      4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.5.0/24 (4608 7575 7545 7545)
  14.200.5.0/24      4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.4.0/24 (4608 7575 7545 7545)
  14.200.6.0/24      4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.7.0/24 (4608 7575 7545 7545)
  14.200.7.0/24      4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.6.0/24 (4608 7575 7545 7545)
  14.200.8.0/24      4608 7575 7545
  14.200.9.0/24      4608 7575 7545 7545
  14.200.10.0/23     4608 7575 7545 7545  + Announce - aggregate of 14.200.10.0/24 (4608 7575 7545 7545) and 14.200.11.0/24 (4608 7575 7545 7545)
  14.200.10.0/24     4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.11.0/24 (4608 7575 7545 7545)
  14.200.11.0/24     4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.10.0/24 (4608 7575 7545 7545)
  14.200.12.0/24     4608 7575 7545 7545
  14.200.13.0/24     4608 9722 7545
  14.200.14.0/23     4608 7575 7545 7545  + Announce - aggregate of 14.200.14.0/24 (4608 7575 7545 7545) and 14.200.15.0/24 (4608 7575 7545 7545)
  14.200.14.0/24     4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.15.0/24 (4608 7575 7545 7545)
  14.200.15.0/24     4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.14.0/24 (4608 7575 7545 7545)
  14.200.16.0/21     4608 7575 7545 7545  + Announce - aggregate of 14.200.16.0/22 (4608 7575 7545 7545) and 14.200.20.0/22 (4608 7575 7545 7545)
  14.200.16.0/24     4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.17.0/24 (4608 7575 7545 7545)
  14.200.17.0/24     4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.16.0/24 (4608 7575 7545 7545)
  14.200.18.0/24     4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.19.0/24 (4608 7575 7545 7545)
  14.200.19.0/24     4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.18.0/24 (4608 7575 7545 7545)
  14.200.20.0/24     4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.21.0/24 (4608 7575 7545 7545)
  14.200.21.0/24     4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.20.0/24 (4608 7575 7545 7545)
  14.200.22.0/24     4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.23.0/24 (4608 7575 7545 7545)
  14.200.23.0/24     4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.22.0/24 (4608 7575 7545 7545)
  14.200.24.0/23     4608 7575 7545 7545  + Announce - aggregate of 14.200.24.0/24 (4608 7575 7545 7545) and 14.200.25.0/24 (4608 7575 7545 7545)
  14.200.24.0/24     4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.25.0/24 (4608 7575 7545 7545)
  14.200.25.0/24     4608 7575 7545 7545  - Withdrawn - aggregated with 14.200.24.0/24 (4608 7575 7545 7545)
  14.200.26.0/24     4608 7575 7545
  14.200.27.0/24     4608 7575 7545 7545
```

Announced Prefixes

```
Rank  AS       Type    Originate Addr Space  (pfx)   Transit Addr space  (pfx)  Description
60    AS12479        ORG+TRN Originate:   12483328 /8.43   Transit:     303872 /13.79 UNI2-AS, ES
```

Aggregation Suggestions

Filter: Aggregates, Specifics

This report does not take into account conditions local to each origin AS in terms of policy or traffic engineering requirements, so this is an approximate guideline as to aggregation possibilities.

```
 Rank AS            AS Name                                  Current  Wthdw  Aggte  Annce Redctn       %
    9 AS12479       UNI2-AS, ES                                 5485   3572    780   2693   2792  50.90%
```

```
 Prefix              AS Path                              Aggregation Suggestion
 37.11.0.0/16        4777 2516 3356 12715 12479
 37.11.0.0/19        4777 2516 3356 5511 12479 + Announce - aggregate of 37.11.0.0/20 (4777 2516 3356 5511 12479) and 37.11.16.0/20 (4777 2516 3356 5511 1247
 37.11.0.0/22        4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.4.0/22 (4777 2516 3356 5511 12479)
 37.11.4.0/22        4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.0.0/22 (4777 2516 3356 5511 12479)
 37.11.8.0/22        4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.12.0/22 (4777 2516 3356 5511 12479)
 37.11.12.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.8.0/22 (4777 2516 3356 5511 12479)
 37.11.16.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.20.0/22 (4777 2516 3356 5511 12479)
 37.11.20.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.16.0/22 (4777 2516 3356 5511 12479)
 37.11.24.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.28.0/22 (4777 2516 3356 5511 12479)
 37.11.28.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.24.0/22 (4777 2516 3356 5511 12479)
 37.11.32.0/22       4608 7575 2914 5511 12479
 37.11.36.0/22       4777 2516 3356 5511 12479
 37.11.40.0/21       4777 2516 3356 5511 12479 + Announce - aggregate of 37.11.40.0/22 (4777 2516 3356 5511 12479) and 37.11.44.0/22 (4777 2516 3356 5511 124
 37.11.40.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.44.0/22 (4777 2516 3356 5511 12479)
 37.11.44.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.40.0/22 (4777 2516 3356 5511 12479)
 37.11.48.0/20       4777 2516 3356 5511 12479 + Announce - aggregate of 37.11.48.0/21 (4777 2516 3356 5511 12479) and 37.11.56.0/21 (4777 2516 3356 5511 124
 37.11.48.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.52.0/22 (4777 2516 3356 5511 12479)
 37.11.52.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.48.0/22 (4777 2516 3356 5511 12479)
 37.11.56.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.60.0/22 (4777 2516 3356 5511 12479)
 37.11.60.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.56.0/22 (4777 2516 3356 5511 12479)
 37.11.64.0/18       4777 2516 3356 5511 12479 + Announce - aggregate of 37.11.64.0/19 (4777 2516 3356 5511 12479) and 37.11.96.0/19 (4777 2516 3356 5511 124
 37.11.64.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.68.0/22 (4777 2516 3356 5511 12479)
 37.11.68.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.64.0/22 (4777 2516 3356 5511 12479)
 37.11.72.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.76.0/22 (4777 2516 3356 5511 12479)
 37.11.76.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.72.0/22 (4777 2516 3356 5511 12479)
 37.11.80.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.84.0/22 (4777 2516 3356 5511 12479)
 37.11.84.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.80.0/22 (4777 2516 3356 5511 12479)
 37.11.88.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.92.0/22 (4777 2516 3356 5511 12479)
 37.11.92.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.88.0/22 (4777 2516 3356 5511 12479)
 37.11.96.0/22       4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.100.0/22 (4777 2516 3356 5511 12479)
 37.11.100.0/22      4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.96.0/22 (4777 2516 3356 5511 12479)
 37.11.104.0/22      4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.108.0/22 (4777 2516 3356 5511 12479)
 37.11.108.0/22      4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.104.0/22 (4777 2516 3356 5511 12479)
 37.11.112.0/22      4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.116.0/22 (4777 2516 3356 5511 12479)
 37.11.116.0/22      4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.112.0/22 (4777 2516 3356 5511 12479)
 37.11.120.0/22      4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.124.0/22 (4777 2516 3356 5511 12479)
 37.11.124.0/22      4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.120.0/22 (4777 2516 3356 5511 12479)
 37.11.128.0/20      4777 2516 3356 5511 12479 + Announce - aggregate of 37.11.128.0/21 (4777 2516 3356 5511 12479) and 37.11.136.0/21 (4777 2516 3356 5511 1
 37.11.128.0/22      4777 2516 3356 5511 12479 - Withdrawn - aggregated with 37.11.132.0/22 (4777 2516 3356 5511 12479)
```

Importance of Aggregation

- Size of routing table
  - Router Memory is not so much of a problem as it was in the 1990s
  - Routers routinely carry over 2 million prefixes
- Convergence of the Routing System
  - This is a problem
  - Bigger table takes longer for CPU to process
  - BGP updates take longer to deal with
  - BGP Instability Report tracks routing system update activity
  - bgpupdates.potaroo.net/instability/bgpupd.html

The BGP Instability Report

The BGP Instability Report is updated daily. This report was generated on 29 July 2019 06:44 (UTC+1000)

50 Most active ASes for the past 14 days

RANK	ASN	UPDs	%	Prefixes	UPDs/Prefix	AS NAME
1	6762	355299	5.38%	363	978.79	SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
2	9829	293312	4.44%	2697	108.75	BSNL-NIB National Internet Backbone, IN
3	13904	93253	1.41%	168	555.08	COSLINK - Cherryland Services Inc, US
4	27738	86788	1.31%	847	102.47	Ecuadortelecom S.A., EC
5	5972	76057	1.15%	1845	41.22	DNIC-ASBLK-05800-06055 - DoD Network Information Center, US
6	47331	72811	1.10%	5663	12.86	TTNET, TR
7	262691	69753	1.06%	136	512.89	CONECTA LTDA., BR
8	2561	69058	1.05%	77	896.86	EUN, EG
9	16509	60610	0.92%	3087	19.63	AMAZON-02 - Amazon.com, Inc., US
10	8151	58509	0.89%	6366	9.19	Uninet S.A. de C.V., MX
11	7939	48591	0.74%	19	2557.42	UNIVCENTFLA - University of Central Florida, US
12	39028	46773	0.71%	27	1732.33	ULSK-AS, RU
13	36903	43331	0.66%	843	51.40	MT-MPLS, MA
14	20852	41202	0.62%	137	300.74	ATLANT-TELECOM-AS AtlantTelecom Autonomus System, BY
15	531	36220	0.55%	126	287.46	DNIC-AS-00531 - Headquarters, USAISC, US
16	46562	33629	0.51%	351	95.81	TOTAL-SERVER-SOLUTIONS - Total Server Solutions L.L.C., US
17	7579	33624	0.51%	108	311.33	INTERNEX-AS-AP InterNex Australia Pty Ltd, AU
18	11492	32828	0.50%	3694	8.89	CABLEONE - CABLE ONE, INC., US
19	138659	30949	0.47%	5	6189.80	CYBERLINK-AS-AP Cyberlink Online, BD
20	10620	30809	0.47%	3420	9.01	Telmex Colombia S.A., CO
21	35487	29933	0.45%	36	831.47	MISAKA, EU
22	15135	28577	0.43%	28	1020.61	DYN-HC - Oracle Corporation, US
23	5800	28329	0.43%	96	295.09	DNIC-ASBLK-05800-06055 - DoD Network Information Center, US
24	58224	26896	0.41%	819	32.84	TCI, IR
25	50710	25790	0.39%	519	49.69	EARTHLINK-AS, IQ
26	21859	24654	0.37%	428	57.60	ZNET - Zenlayer Inc, US
27	3832	22667	0.34%	20	1133.35	CINE-NET - Cinenet Communications, US
28	19058	21300	0.32%	37	575.68	IRTC-NET - Illinois Rural Telecommunication Co., US

50 Most active Prefixes for the past 14 days

RANK	PREFIX	UPDs	%	Origin AS -- AS NAME
1	201.183.255.0/24	51272	0.74%	27738 -- Ecuadortelecom S.A., EC
2	132.170.30.0/23	48591	0.70%	7939 -- UNIVCENTFLA - University of Central Florida, US
3	172.83.45.0/24	29768	0.43%	46562 -- TOTAL-SERVER-SOLUTIONS - Total Server Solutions L.L.C., US
4	216.238.254.0/23	29038	0.42%	13904 -- COSLINK - Cherryland Services Inc, US
5	23.35.212.0/22	27855	0.40%	6762 -- SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
6	23.50.160.0/20	27407	0.39%	6762 -- SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
7	23.50.184.0/22	25845	0.37%	6762 -- SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
8	64.68.236.0/22	25763	0.37%	13904 -- COSLINK - Cherryland Services Inc, US
9	23.50.176.0/20	25749	0.37%	6762 -- SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
10	23.50.188.0/22	25536	0.37%	6762 -- SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
11	2.16.70.0/23	25183	0.36%	6762 -- SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
12	92.123.208.0/22	24972	0.36%	6762 -- SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
13	88.221.100.0/22	24954	0.36%	6762 -- SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
14	92.122.68.0/22	24944	0.36%	6762 -- SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
15	95.101.156.0/22	24821	0.36%	6762 -- SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
16	88.221.28.0/22	24733	0.35%	6762 -- SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
17	2.19.16.0/20	24555	0.35%	6762 -- SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
18	2.16.146.0/23	24498	0.35%	6762 -- SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
19	2.20.4.0/22	24221	0.35%	6762 -- SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
20	193.118.40.0/24	22313	0.32%	21859 -- ZNET - Zenlayer Inc, US
21	162.212.150.0/23	17980	0.26%	30321 -- BURNINGMAN - Burning Man, US
22	177.136.12.0/24	16310	0.23%	52871 -- TASCOM TELECOMUNICAÇÕES LTDA, BR
23	68.69.37.0/24	15469	0.22%	14858 -- HISNET - Hanson Information Systems, Inc. / Family Net, US 19058 -- IRTC-NET - Illinois Rural Telecommunication Co., US
24	109.195.207.0/24	15166	0.22%	39028 -- ULSK-AS, RU
25	209.177.171.0/24	14339	0.21%	18465 -- WORKDAY-01 - Workday, Inc., US
26	68.69.45.0/24	14132	0.20%	14858 -- HISNET - Hanson Information Systems, Inc. / Family Net, US 19058 -- IRTC-NET - Illinois Rural Telecommunication Co., US
27	162.88.46.0/24	14055	0.20%	15135 -- DYN-HC - Oracle Corporation, US
28	204.208.170.0/24	14025	0.20%	5972 -- DNIC-ASBLK-05800-06055 - DoD Network Information Center, US
29	68.70.218.0/24	13193	0.19%	13904 -- COSLINK - Cherryland Services Inc, US
30	68.70.217.0/24	13190	0.19%	13904 -- COSLINK - Cherryland Services Inc, US
31	82.196.154.0/24	12424	0.18%	29651 -- CTCS, RU
32	159.138.67.0/24	11832	0.17%	136907 -- HWCLOUDS-AS-AP HUAWEI CLOUDS, HK

The BGP IPv6 Instability Report

This report is updated daily. The current report
was generated on 29 July 2019 01:22 (UTC+1000)

50 Most active ASes for the past 14 days

RANK	ASN	UPDs	%	Prefixes	UPDs/Prefix	AS NAME
1	23650	328690	16.51%	76	4324.87	CHINANET-JS-AS-AP AS Number for CHINANET jiangsu province backbone, CN
2	6718	135267	6.79%	8	16908.38	NAV NAV Communications, RO
3	26615	122768	6.17%	107	1147.36	Tim Celular S.A., BR
4	9829	89551	4.50%	453	197.68	BSNL-NIB National Internet Backbone, IN
5	38457	89496	4.50%	5	17899.20	HNS-AS-AP Honesty Net Solution (I) Pvt Ltd, IN
6	28573	83535	4.20%	1123	74.39	CLARO S.A., BR
7	28176	81751	4.11%	12	6812.58	Quick Soft tecnologia da Informacao Ltda, BR
8	38082	59790	3.00%	9	6643.33	IIT-TIG-AS-AP True International Gateway Co., Ltd., TH
9	30036	58836	2.96%	927	63.47	MEDIACOM-ENTERPRISE-BUSINESS - Mediacom Communications Corp, US
10	50937	45342	2.28%	2	22671.00	PAGINIEUROPENE-AS, RO
11	5588	33135	1.66%	20	1656.75	GTSCE GTS Central Europe / Antel Germany, CZ
12	203271	28032	1.41%	1	28032.00	BNTPRO, TR
13	12222	27426	1.38%	258	106.30	AKAMAI - Akamai Technologies, Inc., US
14	8376	25785	1.30%	1330	19.39	, JO
15	53692	23054	1.16%	3	7684.67	ISN-4 - Interop Show Network, US
16	32629	20033	1.01%	2	10016.50	CITY-OF-CHARLOTTE-ASN - City of Charlotte, US
17	49762	19577	0.98%	1	19577.00	SMN-AS, FI
18	33047	18887	0.95%	23	821.17	INSTART - Instart Logic, Inc, US
19	56485	18811	0.94%	3	6270.33	THEHOST-AS, UA
20	30361	18242	0.92%	2	9121.00	SWIFTWILL2 - Swiftwill, Inc., US
21	12654	17891	0.90%	37	483.54	RIPE-NCC-RIS-AS Reseaux IP Europeens Network Coordination Centre (RIPE NCC), NL
22	3573	14723	0.74%	143	102.96	ACCENTURE - Accenture LLP, US
23	263885	12959	0.65%	9	1439.89	Central NET, BR
24	9890	12865	0.65%	1	12865.00	ATOSINFOTECH-SG-AP ATOS Information Technology (Singapore) Pte Ltd, SG
25	36351	12684	0.64%	111	114.27	SOFTLAYER - SoftLayer Technologies Inc., US
26	2571	11254	0.57%	7	1607.71	DHLNET - DHL Information Services (Europe) s.r.o, CZ
27	16331	11095	0.56%	1	11095.00	TELE-ENTRE-FI-AS, FI
28	53184	8162	0.41%	19	429.58	INB Telecom EIRELI - ME, BR
29	22394	8086	0.41%	680	11.89	CELLCO - Cellco Partnership DBA Verizon Wireless, US

50 Most active Prefixes for the past 14 days

RANK	PREFIX	UPDs	%	Origin AS -- AS NAME
1	2a06:9087:ffff::/48	28032	1.30%	203271 -- BNTPRO, TR
2	2620:144:a00::/40	23054	1.07%	53692 -- ISN-4 - Interop Show Network, US
3	2a0a:8880::/48	22671	1.05%	50937 -- PAGINIEUROPENE-AS, RO
4	2a0a:8880:1::/48	22671	1.05%	50937 -- PAGINIEUROPENE-AS, RO
5	2a05:1c01::/32	22528	1.04%	6718 -- NAV NAV Communications, RO
6	2a00:ece0::/32	22528	1.04%	6718 -- NAV NAV Communications, RO
7	2a05:1c02::/32	22528	1.04%	6718 -- NAV NAV Communications, RO
8	2a05:1c03::/32	22528	1.04%	6718 -- NAV NAV Communications, RO
9	2a05:1c04::/32	22528	1.04%	6718 -- NAV NAV Communications, RO
10	2a05:1c00::/32	22528	1.04%	6718 -- NAV NAV Communications, RO
12	2620:0:2f0::/48	20025	0.93%	32629 -- CITY-OF-CHARLOTTE-ASN - City of Charlotte, US
15	2001:67c:245c::/48	19577	0.91%	49762 -- SMN-AS, FI
16	2a03:8160:14::/48	18837	0.87%	33047 -- INSTART - Instart Logic, Inc, US
17	2401:4800:2::/48	18639	0.86%	38457 -- HNS-AS-AP Honesty Net Solution (I) Pvt Ltd, IN
18	2401:4800::/48	18624	0.86%	38457 -- HNS-AS-AP Honesty Net Solution (I) Pvt Ltd, IN
19	2401:4800:feed::/48	18597	0.86%	38457 -- HNS-AS-AP Honesty Net Solution (I) Pvt Ltd, IN
20	2a01:9dc0::/32	18241	0.85%	30361 -- SWIFTWILL2 - Swiftwill, Inc., US
21	2401:4800:3021::/48	16821	0.78%	38457 -- HNS-AS-AP Honesty Net Solution (I) Pvt Ltd, IN
22	2401:4800:2011::/48	16815	0.78%	38457 -- HNS-AS-AP Honesty Net Solution (I) Pvt Ltd, IN
23	2804:214:85cb::/48	14333	0.66%	26615 -- Tim Celular S.A., BR
24	2804:214:85ca::/48	14333	0.66%	26615 -- Tim Celular S.A., BR
25	2804:214:85c9::/48	14331	0.66%	26615 -- Tim Celular S.A., BR
26	2804:214:85cc::/48	14330	0.66%	26615 -- Tim Celular S.A., BR
27	2804:214:85c8::/48	14330	0.66%	26615 -- Tim Celular S.A., BR
28	2804:214:85cd::/48	14315	0.66%	26615 -- Tim Celular S.A., BR
29	2804:214:85cf::/48	14312	0.66%	26615 -- Tim Celular S.A., BR
30	2804:214:85ce::/48	14308	0.66%	26615 -- Tim Celular S.A., BR
31	2a00:ad87:4600::/48	12865	0.60%	9890 -- ATOSINFOTECH-SG-AP ATOS Information Technology (Singapore) Pte Ltd, SG
32	2804:14c:8586::/48	12619	0.59%	28573 -- CLARO S.A., BR
33	2604:2e89:8014::/48	11954	0.55%	30036 -- MEDIACOM-ENTERPRISE-BUSINESS - Mediacom Communications Corp, US
34	2804:14c:8584::/48	11767	0.55%	28573 -- CLARO S.A., BR
35	2804:14c:85a4::/48	11738	0.54%	28573 -- CLARO S.A., BR
36	2804:14c:8585::/48	11129	0.52%	28573 -- CLARO S.A., BR

Receiving Prefixes

Receiving Prefixes

- There are three scenarios for receiving prefixes from other ASNs
  - Customer talking BGP
  - Peer talking BGP
  - Upstream/Transit talking BGP
- Each has different filtering requirements and need to be considered separately

Receiving Prefixes: From Customers

- ISPs should only accept prefixes which have been assigned or allocated to their downstream customer
- If ISP has assigned address space to its customer, then the customer IS entitled to announce it back to his ISP
- If the ISP has NOT assigned address space to its customer, then:
  - Check in the five RIR databases to see if this address space really has been assigned to the customer
  - The tool: whois –h jwhois.apnic.net x.x.x.0/24
    - (jwhois is "joint whois" and queries all RIR databases)

Receiving Prefixes: From Customers

- Example use of whois to check if customer is entitled to announce address space:

```
$ whois -h jwhois.apnic.net 202.12.29.0

inetnum:        202.12.29.0 - 202.12.29.255
netname:        APNIC-SERVICES-AU
descr:          Asia Pacific Network Information Centre
descr:          Regional Internet Registry for the Asia-Pacific Region
descr:          6 Cordelia Street
descr:          South Brisbane
geoloc:         27.4731138 153.0141194
country:        AU
admin-c:        AIC1-AP
tech-c:         AIC1-AP
mnt-by:         APNIC-HM
mnt-irt:        IRT-APNIC-IS-AP
status:         ASSIGNED PORTABLE
changed:        hm-changed@apnic.net 20170327
changed:        hm-changed@apnic.net 20170331
source:         APNIC
```

inetnum – means it is an address delegation to an entity

Portable – means its an assignment to the customer, the customer can announce it to you

Receiving Prefixes: From Customers

- Example use of whois to check if customer is entitled to announce address space:

```
$ whois -h jwhois.apnic.net 193.128.0.0/16

inetnum:        193.128.0.0 - 193.133.255.255
netname:        UK-PIPEX-193-128-133
country:        GB
org:            ORG-UA24-RIPE
admin-c:        WERT1-RIPE
tech-c:         UPHM1-RIPE
status:         ALLOCATED UNSPECIFIED
remarks:        Please send abuse notification to abuse@uk.uu.net
mnt-by:         RIPE-NCC-HM-MNT
mnt-by:         AS1849-MNT
mnt-routes:     AS1849-MNT
mnt-routes:     WCOM-EMEA-RICE-MNT
mnt-irt:        IRT-MCI-GB
created:        2002-06-25T15:05:40Z
last-modified:  2016-10-31T12:20:01Z
source:         RIPE
```

ALLOCATED – means that this is Provider Aggregatable address space and can only be announced by the ISP holding the allocation (in this case Verizon UK)

Receiving Prefixes from customer: Cisco IOS

- For Example:
  - Downstream has 100.69.0.0/20 block
  - Should only announce this to upstreams
  - Upstreams should only accept this from them
- Configuration on upstream

```
router bgp 100
 address-family ipv4
  neighbor 100.67.10.1 remote-as 101
  neighbor 100.67.10.1 prefix-list customer in
  neighbor 100.67.10.1 prefix-list default out
  neighbor 100.67.10.1 activate
!
ip prefix-list customer permit 100.69.0.0/20
!
ip prefix-list default permit 0.0.0.0/0
```

Receiving Prefixes: From Peers

- A peer is an ISP with whom you agree to exchange prefixes you originate into the Internet routing table
  - Prefixes you accept from a peer are only those they have indicated they will announce
  - Prefixes you announce to your peer are only those you have indicated you will announce

Receiving Prefixes: From Peers

- Agreeing what each will announce to the other:
  - Exchange of e-mail documentation as part of the peering agreement, and then ongoing updates

    OR
  - Use of the Internet Routing Registry and configuration tools such as the IRRToolSet

    https://github.com/irrtoolset/irrtoolset

Receiving Prefixes from peer: Cisco IOS

- For Example:
  - Peer has 220.50.0.0/16, 61.237.64.0/18 and 81.250.128.0/17 address blocks
- Configuration on local router

```
router bgp 100
 address-family ipv4
  neighbor 100.67.10.1 remote-as 101
  neighbor 100.67.10.1 prefix-list my-peer in
  neighbor 100.67.10.1 prefix-list my-prefix out
  neighbor 100.67.10.1 activate
!
ip prefix-list my-peer permit 220.50.0.0/16
ip prefix-list my-peer permit 61.237.64.0/18
ip prefix-list my-peer permit 81.250.128.0/17
ip prefix-list my-peer deny 0.0.0.0/0 le 32
!
ip prefix-list my-prefix permit 100.67.16.0/20
```

Receiving Prefixes: From Upstream/Transit Provider

- Upstream/Transit Provider is an ISP who you pay to give you transit to the WHOLE Internet
- Receiving prefixes from them is not desirable unless really necessary
  - Traffic Engineering – see BGP Multihoming presentations
- Ask upstream/transit provider to either:
  - originate a default-route

    OR
  - announce one prefix you can use as default

Receiving Prefixes: From Upstream/Transit Provider

- Downstream Router Configuration

```
router bgp 100
 address-family ipv4
  network 100.66.0.0 mask 255.255.224.0
  neighbor 100.65.7.1 remote-as 101
  neighbor 100.65.7.1 prefix-list infilter in
  neighbor 100.65.7.1 prefix-list outfilter out
  neighbor 100.65.7.1 activate
!
ip prefix-list infilter permit 0.0.0.0/0
!
ip prefix-list outfilter permit 100.66.0.0/19
```

Receiving Prefixes: From Upstream/Transit Provider

- Upstream Router Configuration

```
router bgp 101
 address-family ipv4
  neighbor 100.65.7.2 remote-as 100
  neighbor 100.65.7.2 default-originate
  neighbor 100.65.7.2 prefix-list cust-in in
  neighbor 100.65.7.2 prefix-list cust-out out
  neighbor 100.65.7.2 activate
!
ip prefix-list cust-in permit 100.66.0.0/19
!
ip prefix-list cust-out permit 0.0.0.0/0
```

Receiving Prefixes: From Upstream/Transit Provider

- If necessary to receive prefixes from any provider, care is required.
  - Don't accept default (unless you need it)
  - Don't accept your own prefixes
- Special use prefixes for IPv4 and IPv6:
  - http://www.rfc-editor.org/rfc/rfc6890.txt
- For IPv4:
  - Don't accept prefixes longer than /24 (?)
    - /24 was the historical class C
- For IPv6:
  - Don't accept prefixes longer than /48 (?)
    - /48 is the design minimum delegated to a site

Receiving Prefixes: From Upstream/Transit Provider

- Check Team Cymru's list of "bogons"
  - http://www.team-cymru.com/bogon-reference.html
- For IPv4 also consult:
  - https://www.rfc-editor.org/rfc/rfc6441.txt (BCP171)
- For IPv6 also consult:
  - http://www.space.net/~gert/RIPE/ipv6-filters.html
- Bogon Route Server:
  - https://www.team-cymru.com/bogon-reference-bgp.html
  - Supplies a BGP feed (IPv4 and/or IPv6) of address blocks which should not appear in the BGP table

Receiving IPv4 Prefixes

```
router bgp 100
 network 101.10.0.0 mask 255.255.224.0
 neighbor 100.65.7.1 remote-as 101
 neighbor 100.65.7.1 prefix-list in-filter in
!
ip prefix-list in-filter deny 0.0.0.0/0              ! Default
ip prefix-list in-filter deny 0.0.0.0/8 le 32        ! RFC1122 local host
ip prefix-list in-filter deny 10.0.0.0/8 le 32       ! RFC1918
ip prefix-list in-filter deny 100.64.0.0/10 le 32    ! RFC6598 shared address
ip prefix-list in-filter deny 101.10.0.0/19 le 32    ! Local prefix
ip prefix-list in-filter deny 127.0.0.0/8 le 32      ! Loopback
ip prefix-list in-filter deny 169.254.0.0/16 le 32   ! Auto-config
ip prefix-list in-filter deny 172.16.0.0/12 le 32    ! RFC1918
ip prefix-list in-filter deny 192.0.0.0/24 le 32     ! RFC6598 IETF protocol
ip prefix-list in-filter deny 192.0.2.0/24 le 32     ! TEST1
ip prefix-list in-filter deny 192.168.0.0/16 le 32   ! RFC1918
ip prefix-list in-filter deny 198.18.0.0/15 le 32    ! Benchmarking
ip prefix-list in-filter deny 198.51.100.0/24 le 32  ! TEST2
ip prefix-list in-filter deny 203.0.113.0/24 le 32   ! TEST3
ip prefix-list in-filter deny 224.0.0.0/3 le 32      ! Multicast & Experimental
ip prefix-list in-filter deny 0.0.0.0/0 ge 25        ! Prefixes >/24
ip prefix-list in-filter permit 0.0.0.0/0 le 32
```

Receiving IPv6 Prefixes

```
router bgp 100
 network 2020:3030::/32
 neighbor 2020:3030::1 remote-as 101
 neighbor 2020:3030::1 prefix-list v6in-filter in
!
ipv6 prefix-list v6in-filter permit 64:ff9b::/96          ! RFC6052 v4v6trans
ipv6 prefix-list v6in-filter deny 2001::/23 le 128        ! RFC2928 IETF prot
ipv6 prefix-list v6in-filter deny 2001:2::/48 le 128      ! Benchmarking
ipv6 prefix-list v6in-filter deny 2001:10::/28 le 128     ! ORCHID
ipv6 prefix-list v6in-filter deny 2001:db8::/32 le 128    ! Documentation
ipv6 prefix-list v6in-filter deny 2002::/16 le 128        ! Deny all 6to4
ipv6 prefix-list v6in-filter deny 2020:3030::/32 le 128   ! Local Prefix
ipv6 prefix-list v6in-filter deny 3ffe::/16 le 128        ! Formerly 6bone
ipv6 prefix-list v6in-filter permit 2000::/3 le 48        ! Global Unicast
ipv6 prefix-list v6in-filter deny ::/0 le 128
```

Note: These filters block Teredo (serious security risk) and 6to4 (deprecated by RFC7526)

Receiving Prefixes

- Paying attention to prefixes received from customers, peers and transit providers assists with:
  - The integrity of the local network
  - The integrity of the Internet
- Responsibility of all ISPs to be good Internet citizens

Prefixes into iBGP

Injecting prefixes into iBGP

- Use iBGP to carry customer prefixes
  - Don’t use IGP
- Point static route to customer interface
- Use BGP network statement
- As long as static route exists (interface active), prefix will be in BGP

Router Configuration: network statement

- Example:

```
interface loopback 0
 ip address 100.64.3.1 255.255.255.255
!
interface Serial 5/0
 ip unnumbered loopback 0
 ip verify unicast reverse-path
!
ip route 100.71.10.0 255.255.252.0 Serial 5/0
!
router bgp 100
 address-family ipv4
  network 100.71.10.0 mask 255.255.252.0
!
```

Injecting prefixes into iBGP

- Interface flap will result in prefix withdraw and reannounce
  - use "**`ip route . . . permanent`**"
- Many ISPs redistribute static routes into BGP rather than using the network statement
  - Only do this if you understand why

Router Configuration: redistribute static

- Example:

```
ip route 100.71.10.0 255.255.252.0 Serial 5/0
!
router bgp 100
 address-family ipv4
  redistribute static route-map static-to-bgp
<snip>
!
route-map static-to-bgp permit 10
 match ip address prefix-list ISP-block
 set origin igp
 set community 100:1000
<snip>
!
ip prefix-list ISP-block permit 100.71.10.0/22 le 30
```

Injecting prefixes into iBGP

- Route-map **`static-to-bgp`** can be used for many things:
  - Setting communities and other attributes
  - Setting origin code to IGP, etc
- Be careful with prefix-lists and route-maps
  - Absence of either/both means all statically routed prefixes go into iBGP

Summary

- Best Practices Covered:
  - When to use BGP
  - When to use ISIS/OSPF
  - Aggregation
  - Receiving Prefixes
  - Prefixes into BGP

Interconnection Best Practices

PeeringDB and the Internet Routing Registry

Types of Peering (1)

- Private Peering
  - Where two network operators agree to interconnect their networks, and exchange their respective routes, for the purpose of ensuring their customers can reach each other directly over the peering link
- Settlement Free Peering
  - No traffic charges
  - **The most common form of peering**
- Paid Peering
  - Where two operators agree to exchange traffic charges for a peering relationship

Types of Peering (2)

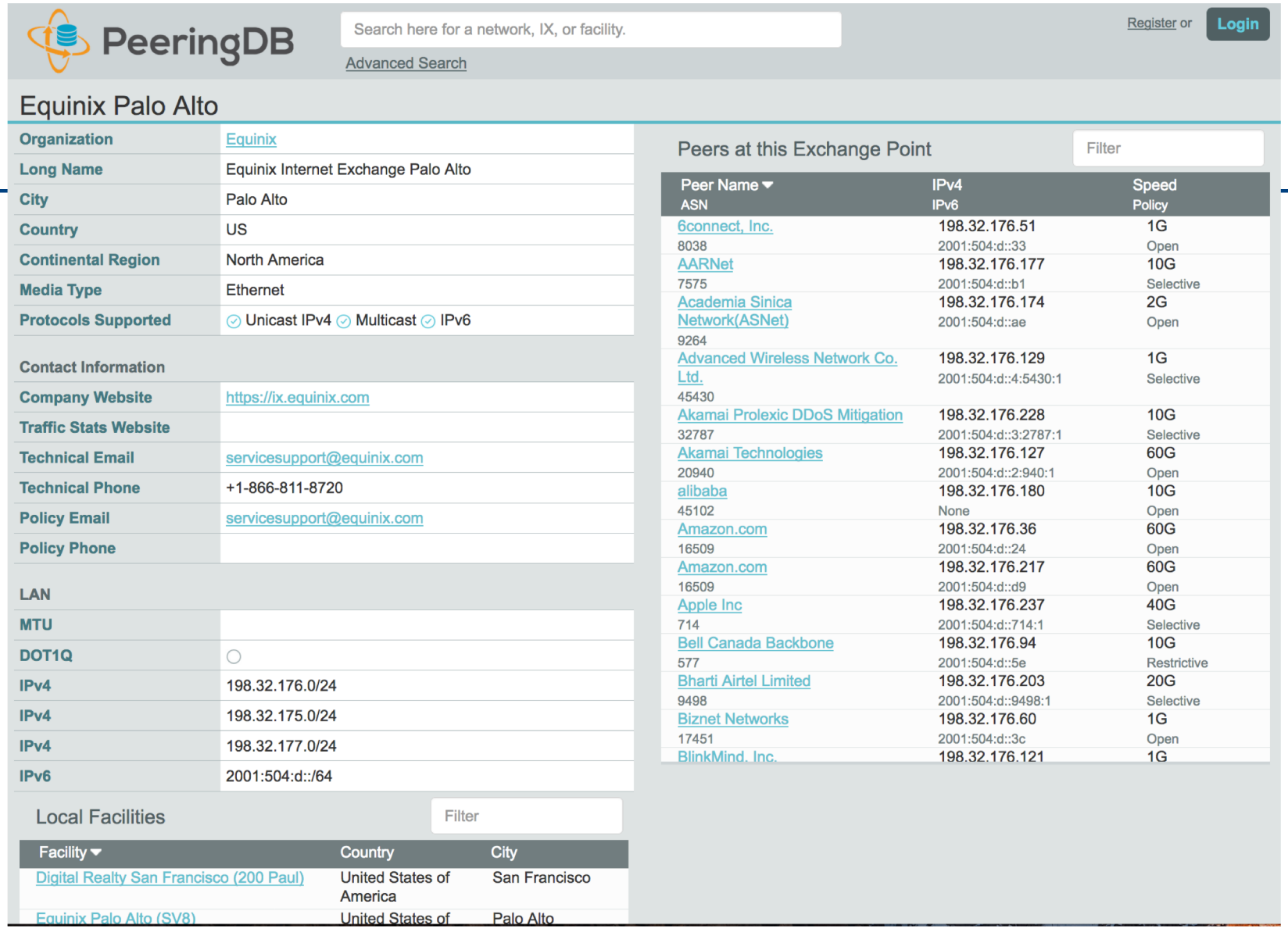
- Bi-lateral Peering
  - Very similar to Private Peering, but usually takes place at a public peering point (IXP)
- Multilateral Peering
  - Takes place at Internet Exchange Points, where operators all peer with each other via a Route Server
- Mandatory Multilateral Peering
  - Where operators are forced to peer with each other as condition of IXP membership
  - **Strongly discouraged: Has no record of success**

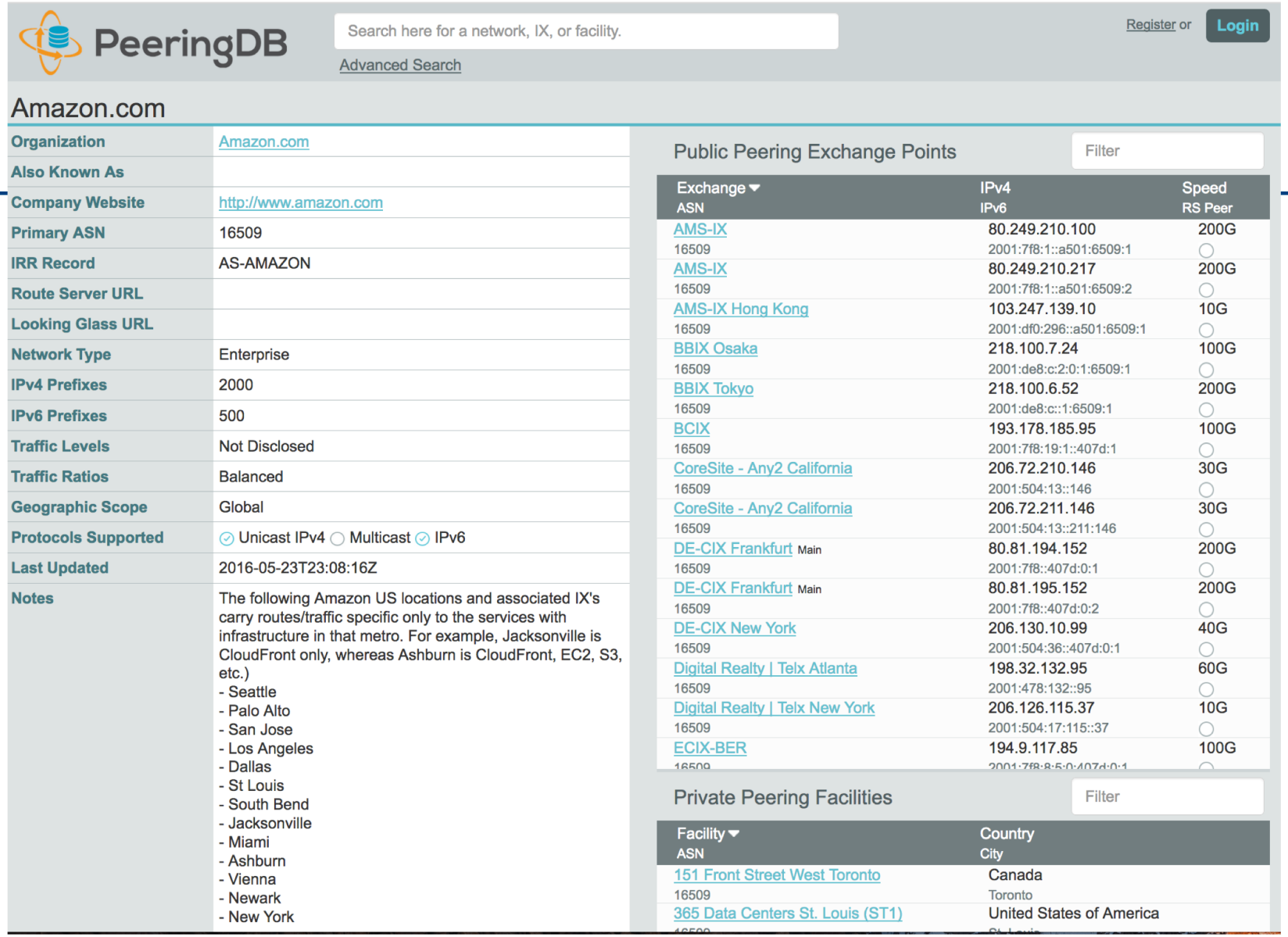
Types of Peering (3)

- Open Peering
  - Where an ISP publicly states that they will peer with all parties who approach them for peering
  - Commonly found at IXPs where ISP participates via the Route Server
- Selective Peering
  - Where an ISP's peering policy depends on the nature of the operator who requests peering with them
  - At IXPs, operator will not peer with RS but will only peer bilaterally
- Restrictive Peering
  - Where an ISP decides who its peering partners are, and is generally not approachable to considering peering opportunities

Types of Peering (4)

- The Peering Database documents ISPs peering policies
  - https://www.peeringdb.com
- All operators of ASNs should register in the PeeringDB
  - All operators who are considering peering or are peering must be in the PeeringDB to enhance their peering opportunities
- Participation in peering fora is encouraged too
  - Global Peering Forum (GPF)
  - Regional Peering Fora (European, Middle Eastern, Asian, Caribbean, Latin American)
  - Many countries now have their own Peering Fora


PeeringDB
Search here for a network, IX, or facility.
Advanced Search
Register or
Login
Equinix Palo Alto
Organization
Equinix
Long Name
Equinix Internet Exchange Palo Alto
City
Palo Alto
Country
US
Continental Region
North America
Media Type
Ethernet
Protocols Supported
Unicast IPv4 Multicast IPv6
Contact Information
Company Website
https://ix.equinix.com
Traffic Stats Website
Technical Email
servicesupport@equinix.com
Technical Phone
+1-866-811-8720
Policy Email
servicesupport@equinix.com
Policy Phone
LAN
MTU
DOT1Q
IPv4
198.32.176.0/24
IPv4
198.32.175.0/24
IPv4
198.32.177.0/24
IPv6
2001:504:d::/64
Local Facilities
Filter
Facility
Country
City
Digital Realty San Francisco (200 Paul)
United States of America
San Francisco
Equinix Palo Alto (SV8)
United States of
Palo Alto
Peers at this Exchange Point
Filter
Peer Name
ASN
IPv4
IPv6
Speed
Policy
6connect, Inc.
8038
198.32.176.51
2001:504:d::33
1G
Open
AARNet
7575
198.32.176.177
2001:504:d::b1
10G
Selective
Academia Sinica Network(ASNet)
9264
198.32.176.174
2001:504:d::ae
2G
Open
Advanced Wireless Network Co. Ltd.
45430
198.32.176.129
2001:504:d::4:5430:1
1G
Selective
Akamai Prolexic DDoS Mitigation
32787
198.32.176.228
2001:504:d::3:2787:1
10G
Selective
Akamai Technologies
20940
198.32.176.127
2001:504:d::2:940:1
60G
Open
alibaba
45102
198.32.176.180
None
10G
Open
Amazon.com
16509
198.32.176.36
2001:504:d::24
60G
Open
Amazon.com
16509
198.32.176.217
2001:504:d::d9
60G
Open
Apple Inc
714
198.32.176.237
2001:504:d::714:1
40G
Selective
Bell Canada Backbone
577
198.32.176.94
2001:504:d::5e
10G
Restrictive
Bharti Airtel Limited
9498
198.32.176.203
2001:504:d::9498:1
20G
Selective
Biznet Networks
17451
198.32.176.60
2001:504:d::3c
1G
Open
BlinkMind, Inc.
198.32.176.121
1G


PeeringDB
Search here for a network, IX, or facility.
Advanced Search
Register or
Login
Amazon.com
Organization
Amazon.com
Also Known As
Company Website
http://www.amazon.com
Primary ASN
16509
IRR Record
AS-AMAZON
Route Server URL
Looking Glass URL
Network Type
Enterprise
IPv4 Prefixes
2000
IPv6 Prefixes
500
Traffic Levels
Not Disclosed
Traffic Ratios
Balanced
Geographic Scope
Global
Protocols Supported
Unicast IPv4 Multicast IPv6
Last Updated
2016-05-23T23:08:16Z
Notes
The following Amazon US locations and associated IX's carry routes/traffic specific only to the services with infrastructure in that metro. For example, Jacksonville is CloudFront only, whereas Ashburn is CloudFront, EC2, S3, etc.)
- Seattle
- Palo Alto
- San Jose
- Los Angeles
- Dallas
- St Louis
- South Bend
- Jacksonville
- Miami
- Ashburn
- Vienna
- Newark
- New York
Public Peering Exchange Points
Filter
Exchange
ASN
IPv4
IPv6
Speed
RS Peer
AMS-IX
16509
80.249.210.100
2001:7f8:1::a501:6509:1
200G
AMS-IX
16509
80.249.210.217
2001:7f8:1::a501:6509:2
200G
AMS-IX Hong Kong
16509
103.247.139.10
2001:df0:296::a501:6509:1
10G
BBIX Osaka
16509
218.100.7.24
2001:de8:c:2:0:1:6509:1
100G
BBIX Tokyo
16509
218.100.6.52
2001:de8:c::1:6509:1
200G
BCIX
16509
193.178.185.95
2001:7f8:19:1::407d:1
100G
CoreSite - Any2 California
16509
206.72.210.146
2001:504:13::146
30G
CoreSite - Any2 California
16509
206.72.211.146
2001:504:13::211:146
30G
DE-CIX Frankfurt Main
16509
80.81.194.152
2001:7f8::407d:0:1
200G
DE-CIX Frankfurt Main
16509
80.81.195.152
2001:7f8::407d:0:2
200G
DE-CIX New York
16509
206.130.10.99
2001:504:36::407d:0:1
40G
Digital Realty | Telx Atlanta
16509
198.32.132.95
2001:478:132::95
60G
Digital Realty | Telx New York
16509
206.126.115.37
2001:504:17:115::37
10G
ECIX-BER
194.9.117.85
100G
Private Peering Facilities
Filter
Facility
ASN
Country
City
151 Front Street West Toronto
16509
Canada
Toronto
365 Data Centers St. Louis (ST1)
United States of America

Internet Routing Registry

- Many major transit providers and several content providers pay attention to what is contained in the Internet Routing Registry
  - There are many IRRs operating, the most commonly used being those hosted by the Regional Internet Registries, RADB, and some transit providers
- Best practice for any AS holder is to document their routing policy in the IRR
  - A route-object is the absolute minimum requirement

Internet Routing Registry

- IRR objects can be created via the database web-interfaces or submitted via email
- Policy language used known as RPSL
- Problems:
  - IRR contains a lot of outdated information
  - Network operators not following best practices
- Some network operators now using RPKI and ROAs to securely indicate the origin AS of their routes
  - Takes priority over IRR entries
  - RPKI and ROAs covered later in the presentation

Route Object: Purpose

- Documents which Autonomous System number is originating the route listed
- Required by many major transit providers
  - They build their customer and peer filter based on the route-objects listed in the IRR
  - Referring to at least the 5 RIR routing registries and the RADB
  - Some operators run their own instance of the IRR as well
    - May require their customers to place a Route Object there (if not using the 5 RIR or RADB versions of the IRR)

Route Object: Examples

```
route:          202.144.128.0/20
descr:          DRUKNET-BLOCK-A1
country:        BT
notify:         ioc@bt.bt
mnt-by:         MAINT-BT-DRUKNET
origin:         AS18024
last-modified:  2018-09-18T09:37:40Z
source:         APNIC
```

This declares that AS18024 is the origin of 202.144.128.0/20

```
route6:         2405:D000::/32
descr:          DRUKNET-IPV6-BLOCK
origin:         AS17660
notify:         netops@bt.bt
mnt-by:         MAINT-BT-DRUKNET
last-modified:  2010-07-21T03:46:02Z
source:         APNIC
```

This declares that AS17660 is the origin of 2405:D000::/32

AS Object: Purpose

- Documents peering policy with other Autonomous Systems
  - Lists network information
  - Lists contact information
  - Lists routes announced to neighbouring autonomous systems
  - Lists routes accepted from neighbouring autonomous systems
- Some operators pay close attention to what is contained in the AS Object
  - Some configure their border router BGP policy based on what is listed in the AS Object

AS Object: Example

```
aut-num:        AS17660
as-name:        DRUKNET-AS
descr:          DrukNet ISP, Bhutan Telecom, Thimphu
country:        BT
org:            ORG-BTL2-AP
import:         from AS6461   action pref=100;     accept ANY
export:         to AS6461     announce AS-DRUKNET-TRANSIT
import:         from AS2914   action pref=150;     accept ANY
export:         to AS2914     announce AS-DRUKNET-TRANSIT
<snip>
import:         from AS135666 action pref=250;   accept AS135666
export:         to AS135666   announce {0.0.0.0/0} AS-DRUKNET-TRANSIT
admin-c:        DNO1-AP
tech-c:         DNO1-AP
notify:         netops@bt.bt
mnt-irt:        IRT-BTTELECOM-BT
mnt-by:         APNIC-HM
mnt-lower:      MAINT-BT-DRUKNET
mnt-routes:     MAINT-BT-DRUKNET
last-modified:  2019-06-09T22:40:10Z
source:         APNIC
```

Examples of inbound and outbound policies – RPSL

AS-Set: Purpose

- The AS-Set is used by network operators to group AS numbers they provide transit for in an easier to manage form
  - Convenient for more complicated policy declarations
  - Used mostly by network operators who build their EBGP filters from their IRR entries
  - Commonly used at Internet Exchange Points to handle large numbers of peers

AS-Set: Example

```
as-set:          AS-DRUKNET-TRANSIT
descr:           DrukNet transit networks
members:         AS17660
members:         AS38004
members:         AS132232
members:         AS134715
members:         AS135666
members:         AS137925
members:         AS59219
members:         AS18024
members:         AS18025
members:         AS137994
admin-c:         DNO1-AP
tech-c:          DNO1-AP
notify:          netops@bt.bt
mnt-by:          MAINT-BT-DRUKNET
last-modified:   2019-01-15T08:51:21Z
source:          APNIC
```

Lists all the autonomous systems within the AS-DRUKNET-TRANSIT group

Summary

- PeeringDB
  - An industry Best Practice so that:
    - Network operators can promote the interconnects they participate in and attract more peering partners
    - Internet Exchange Points can show their participants and help make the interconnect more attractive for potential participants
- IRR
  - An industry Best Practice:
    - So that network operators can document which autonomous system is originating their prefixes
    - Used by network operators to filter prefixes received from their customers and peers

Route Origin Authorisation

Steps to securing the Routing System

What is RPKI?

- Resource Public Key Infrastructure (RPKI)
  - A security framework for verifying the association between resource holder and their Internet resources
  - Created to address the issues discussed in RFC 4593 "Generic Threats to Routing Protocols" (Oct 2006)
- Helps to secure Internet routing by validating routes
  - Proof that prefix announcements are coming from the legitimate holder of the resource
  - RFC 6480 – An Infrastructure to Support Secure Internet Routing (Feb 2012)

Benefits of RPKI - Routing

- Prevents route hijacking
  - A prefix originated by an AS without authorisation
  - Reason: malicious intent
- Prevents mis-origination
  - A prefix that is mistakenly originated by an AS which does not own it
  - Also route leakage
  - Reason: configuration mistake / fat finger

BGP Security (BGPsec)

- Extension to BGP that provides improved security for BGP routing
- Being worked on by the SIDR Working Group at IETF
- Implemented via a new optional non-transitive BGP attribute that contains a digital signature
- Two components:
  - BGP Prefix Origin Validation (using RPKI)
  - BGP Path Validation

Route Origin Authorisation (ROA)

- A digital object that contains a list of address prefixes and one AS number
- It is an authority created by a prefix holder to authorise an AS Number to originate one or more specific route advertisements
- Publish a ROA using your RIR member portal
  - Consult your RIR for how to use their member portal to publish your ROAs

Route Origin Validation

- Router must support RPKI
- Checks an RP cache / validator
- Validation returns 3 states:

State	Description
Valid	When authorisation is found for prefix X coming from ASN Y
Invalid	When authorisation is found for prefix X but **not** from ASN Y
Unknown	When no authorisation data is found for prefix X

Route Origin Validation

- RFC6483 also describes "Disavowal of Routing Origination"
  - AS 0 has been reserved for network operators and other entities to identify non-routed networks
  - Which means:
    - "A ROA with a subject of AS0 (AS0 ROA) is an attestation by the holder of a prefix that the prefix described in the ROA, and any more specific prefix, should not be used in a routing context"
- Any prefixes with ROA indicating AS 0 as the origin AS need to be dropped
  - If these prefixes appear with any other origin, their ROAs will be invalid, achieving this goal

Route Origin Validation – AS0

- Possible use cases:
  - Internal use of a prefix that should not appear in the global BGP table
  - Internet Exchange Point LAN must never appear in the global BGP table
  - Private Address space (IPv4) and non-Global Unicast space (IPv6)
  - Unassigned address space
  - IPv4 and IPv6 resources which should not appear in the global BGP table
    - For example, the special use address space described in RFC6890

Route Origin Validation

- Vendor support:
  - Cisco IOS – available from release 15.2
  - Cisco IOS/XR – available from release 4.3.2
  - Juniper – available from release 12.2
  - Nokia – available from release R12.0R4
  - Huawei – available from release V800R009C10
  - Brocade – available from release TBA
  - FRR – available from release 4.0
  - BIRD – available from release 1.6

RPKI Validator Caches

- NLnet Labs Routinator
  - https://www.nlnetlabs.nl/projects/rpki/routinator/
  - https://github.com/NLnetLabs/routinator
- Dragon Research validator
  - https://rpki.net
  - https://github.com/dragonresearch/rpki.net/
- RIPE NCC validator
  - https://github.com/RIPE-NCC/rpki-validator-3/wiki
- Cloudflare validator (OctoRPKI)
  - https://github.com/cloudflare/cfrpki

Configure Router to Use Cache: Cisco IOS

- Point router to the local RPKI cache
  - Server listens on port 43779
  - Example:

```
router bgp 64512
 bgp rpki server tcp 10.0.0.3 port 43779 refresh 60
```

  - Once the router's RPKI table is populated, router indicates validation state in the BGP table

BGP Table (IPv4)

```
RPKI validation codes: V valid, I invalid, N Not found

Network          Metric LocPrf Path
N*>  1.0.4.0/24       0         37100 6939 4637 1221 38803 56203 i
N*>  1.0.5.0/24       0         37100 6939 4637 1221 38803 56203 i
...
V*>  1.9.0.0/16       0         37100 4788 i
N*>  1.10.8.0/24      0         37100 10026 18046 17408 58730 i
N*>  1.10.64.0/24     0         37100 6453 3491 133741 i
...
V*>  1.37.0.0/16      0         37100 4766 4775 i
N*>  1.38.0.0/23      0         37100 6453 1273 55410 38266 i
N*>  1.38.0.0/17      0         37100 6453 1273 55410 38266 {38266} i
...
I*   5.8.240.0/23     0         37100 44217 3178 i
I*   5.8.241.0/24     0         37100 44217 3178 i
I*   5.8.242.0/23     0         37100 44217 3178 i
I*   5.8.244.0/23     0         37100 44217 3178 i
...
```

Courtesy of SEACOM: http://as37100.net

BGP Table (IPv6)

```
RPKI validation codes: V valid, I invalid, N Not found

Network                Metric LocPrf Path
N*>  2001::/32              0        37100 6939 i
N*   2001:4:112::/48        0        37100 112 i
...
V*>  2001:240::/32          0         37100 2497 i
N*>  2001:250::/48          0         37100 6939 23911 45
N*>  2001:250::/32          0         37100 6939 23911 23910 i
...
V*>  2001:348::/32          0         37100 2497 7679 i
N*>  2001:350::/32          0         37100 2497 7671 i
N*>  2001:358::/32          0         37100 2497 4680 i
...
I*   2001:1218:101::/48     0         37100 6453 8151 278 i
I*   2001:1218:104::/48     0         37100 6453 8151 278 i
N*   2001:1221::/48         0         37100 2914 8151 28496 i
N*>  2001:1228::/32         0         37100 174 18592 i
...
```

Courtesy of SEACOM: http://as37100.net

RPKI BGP State: Valid

```
BGP routing table entry for 2001:240::/32, version 109576927
Paths: (2 available, best #2, table default)
  Not advertised to any peer
  Refresh Epoch 1
  37100 2497
    2C0F:FEB0:11:2::1 (FE80::2A8A:1C00:1560:5BC0) from
                                   2C0F:FEB0:11:2::1 (105.16.0.131)
      Origin IGP, metric 0, localpref 100, valid, external, best
      Community: 37100:2 37100:22000 37100:22004 37100:22060
      path 0828B828 RPKI State valid
      rx pathid: 0, tx pathid: 0x0
```

Courtesy of SEACOM: http://as37100.net

RPKI BGP State: Invalid

```
BGP routing table entry for 2001:1218:101::/48, version 149538323
Paths: (2 available, no best path)
  Not advertised to any peer
  Refresh Epoch 1
  37100 6453 8151 278
    2C0F:FEB0:B:3::1 (FE80::86B5:9C00:15F5:7C00) from
                                    2C0F:FEB0:B:3::1 (105.16.0.162)
      Origin IGP, metric 0, localpref 100, valid, external
      Community: 37100:1 37100:12
      path 0DA7D4FC RPKI State invalid
      rx pathid: 0, tx pathid: 0
```

Courtesy of SEACOM: http://as37100.net

RPKI BGP State: Not Found

```
BGP routing table entry for 2001:200::/32, version 124240929
Paths: (2 available, best #2, table default)
  Not advertised to any peer
  Refresh Epoch 1
  37100 2914 2500
    2C0F:FEB0:11:2::1 (FE80::2A8A:1C00:1560:5BC0) from
                                  2C0F:FEB0:11:2::1 (105.16.0.131)
      Origin IGP, metric 0, localpref 100, valid, external, best
      Community: 37100:1 37100:13
      path 19D90E68 RPKI State not found
      rx pathid: 0, tx pathid: 0x0
```

Courtesy of SEACOM: http://as37100.net

Configure Router to Use Cache: JunOS

1. Connect to validation cache:

```
routing-options {
  validation {
    group ISP {
      session 10.0.0.3;
      port 43779;
      refresh-time 600;
      hold-time 1800;
    }
  }
}
```

- (using same parameters as for the Cisco IOS example)

Configure Router to Use Cache: JunOS

2. Configure validation policies:

```
policy-options {
  policy-statement RPKI-validation {
    term VALID {
      from {
        protocol bgp;
        validation-database valid;
      }
      then {
        validation-state valid;
        next policy;
      }
    }
    term INVALID {
      from {
        protocol bgp;
        validation-database invalid;
      }
      then {
        validation-state invalid;
        next policy;
      }
    }
```

```
(continued)...

    term UNKNOWN {
      from {
        protocol bgp;
        validation-database unknown;
      }
      then {
        validation-state unknown;
        next policy;
      }
    }
  }
}
```

Configure Router to Use Cache: JunOS

3. Apply policy to eBGP session:

```
protocols {
  bgp {
    group EBGP {
      type external;
      local-address 10.0.1.1;
      neighbor 10.1.15.1 {
        description ”ISP Upstream";
        import [ RPKI-validation Upstream-in ];
        export LocalAS-out;
        peer-as 64511;
      }
    }
  }
}
```

- Note that policy options *Upstream-in* and *LocalAS-out* are the typical inbound and outbound filters needed for an eBGP session

Using RPKI for Route Origin Validation

- Network operators can make decisions based on RPKI state:
  - Invalid – discard the prefix – several do this now!
  - Not found – let it through (maybe low local preference)
  - Valid – let it through (high local preference)
- Some operators even considering making "not found" a discard event
  - But then Internet IPv4 BGP table would shrink to about 55000 prefixes and the IPv6 BGP table would shrink to about 9600 prefixes!

RPKI Summary

- All AS operators must consider deploying:
  - **Signing ROAs**
  - **Dropping Invalids** (ROV)
- An important step to securing the routing system
- Doesn't secure the path, but that's the next hurdle to cross
- With origin validation, the opportunities for malicious or accidental mis-origination disappear
- FAQ:
  - https://nlnetlabs.nl/projects/rpki/faq/

Configuration Tips

Of passwords, tricks and templates

iBGP and IGPs Reminder!

- Make sure loopback is configured on router
  - iBGP between loopbacks, NOT real interfaces
- Make sure IGP carries loopback IPv4 /32 and IPv6 /128 address
- Consider the DMZ nets:
  - Use unnumbered interfaces?
  - Use next-hop-self on iBGP neighbours
  - Or carry the DMZ IPv4 /30s and IPv6 /127s in the iBGP
  - Basically keep the DMZ nets out of the IGP!

iBGP: Next-hop-self

- BGP speaker announces external network to iBGP peers using router's local address (loopback) as next-hop
- Used by many ISPs on edge routers
  - Preferable to carrying DMZ point-to-point link addresses in the IGP
  - Reduces size of IGP to just core infrastructure
  - Alternative to using unnumbered interfaces
  - Helps scale network
  - Many ISPs consider this "best practice"

Limiting AS Path Length

- Some BGP implementations have problems with long AS_PATHS
  - Memory corruption
  - Memory fragmentation
- Even using AS_PATH prepends, it is not normal to see more than 20 ASes in a typical AS_PATH in the Internet today
  - The Internet is around 5 ASes deep on average
  - Largest AS_PATH is usually 16-20 ASNs

```
neighbor x.x.x.x maxas-limit 20
```

Limiting AS Path Length

- Some announcements have ridiculous lengths of AS-paths
  - This example is an error in one IPv6 implementation

```
*> 3FFE:1600::/24       22 11537 145 12199 10318 10566 13193 1930 2200 3425 293 5609 5430
13285 6939 14277 1849 33 15589 25336 6830 8002 2042 7610 i
```

  - This example shows 100 prepends (for no obvious reason)

```
*>i193.105.15.0         2516 3257 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404
50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404
50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404
50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404
50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404
50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404
50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404 50404
50404 i
```

- If your implementation supports it, consider limiting the maximum AS-path length you will accept

BGP Maximum Prefix Tracking

- Allow configuration of the maximum number of prefixes a BGP router will receive from a peer
- Two level control:
  - Warning threshold: log warning message
  - Maximum: tear down the BGP peering, manual intervention required to restart

```
neighbor <x.x.x.x> maximum-prefix <max> [restart N] [<threshold>] [warning-only]
```

- *restart* is an optional keyword which will restart the BGP session N minutes after being torn down
- *threshold* is an optional parameter between 1 to 100
  - Specify the percentage of <max> that will cause a warning message to be generated. Default is 75%.
- *warning-only* is an optional keyword which allows log messages to be generated but peering session will not be torn down

Private-AS – Application

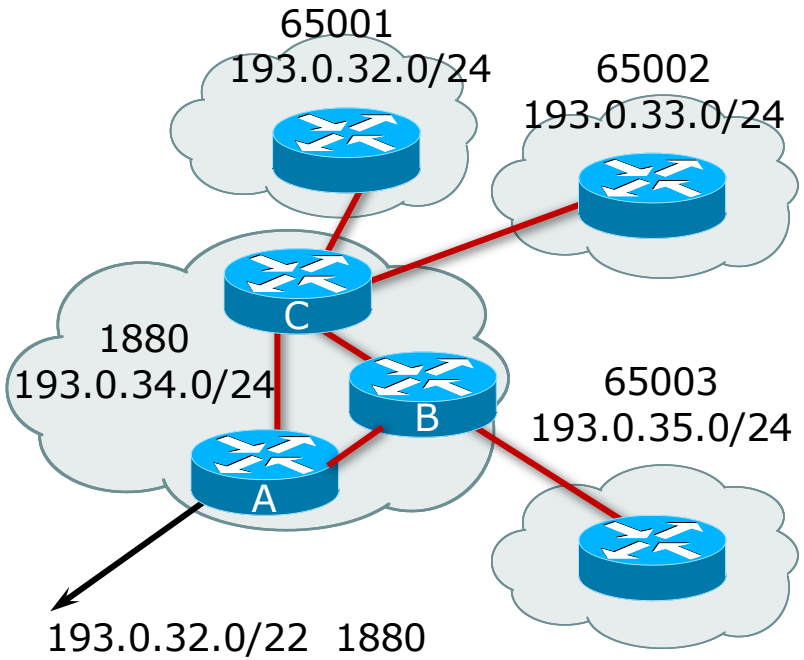
- A network operator with end-sites multihomed on their backbone (RFC2270)

  or

- A corporate network with several regions but connections to the Internet only in the core

  or

- Within a BGP Confederation



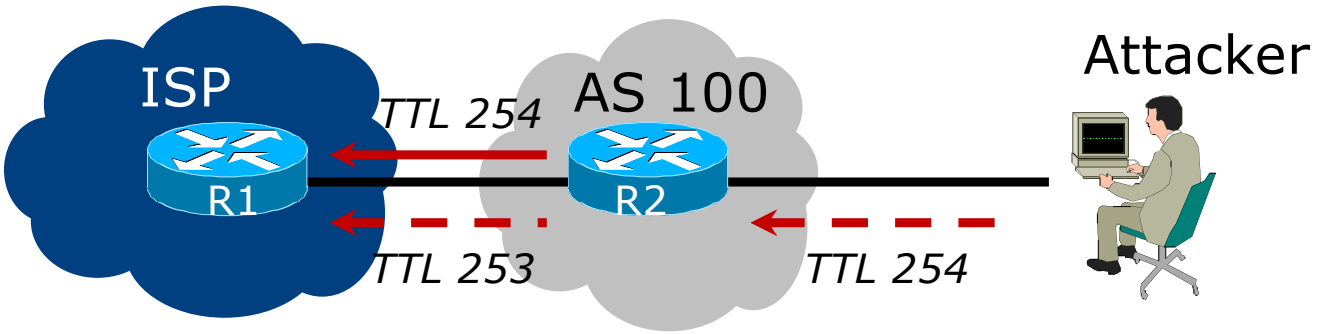
Private-AS – Removal

- Private ASNs MUST be removed from all prefixes announced to the public Internet
  - Include configuration to remove private ASNs in the eBGP template
- As with RFC1918 address space, private ASNs are intended for internal use
  - They must not be leaked to or used on the public Internet
- Cisco IOS

```
neighbor x.x.x.x remove-private-AS
```

BGP TTL "hack"

- Implement RFC5082 on BGP peerings
  - (Generalised TTL Security Mechanism)
  - Neighbour sets TTL to 255
  - Local router expects TTL of incoming BGP packets to be 254
  - No one apart from directly attached devices can send BGP packets which arrive with TTL of 254, so any possible attack by a remote miscreant is dropped due to TTL mismatch



BGP TTL "hack"

- TTL Hack:
  - Both neighbours must agree to use the feature
  - TTL check is much easier to perform than MD5
  - (Called BTSH – BGP TTL Security Hack)
- Provides "security" for BGP sessions
  - In addition to packet filters of course
  - MD5 should still be used for messages which slip through the TTL hack
  - See https://www.nanog.org/meetings/nanog27/presentations/meyer.pdf for more details

BGP TTL "hack"

- Configuration example:

```
neighbor 100.121.0.2 ttl-security hops 1
```

- BGP neighbour status:

```
Router# sh ip bgp neigh 100.121.0.2
...
Mininum incoming TTL 254, Outgoing TTL 255
Local host: 100.121.0.1, Local port: 41103
Foreign host: 100.121.0.2, Foreign port: 179
```

- The neighbour must set the same configuration
  - If they don't, the BGP session will not come up

Templates

- Good practice to configure templates for everything
  - Vendor defaults tend not to be optimal or even very useful for ISPs
  - ISPs create their own defaults by using configuration templates
- eBGP and iBGP examples follow
  - Also see Team Cymru's BGP templates
    - http://www.team-cymru.com/community-services.html

iBGP Template Example

- iBGP between loopbacks!
- Next-hop-self
  - Keep DMZ and external point-to-point out of IGP
- Always send communities in iBGP
  - Otherwise BGP policy accidents will happen
  - (Default on some vendor implementations, optional on others)
- Hardwire BGP to version 4
  - Yes, this is being paranoid!
  - Prevents accidental configuration of BGP version 3 which is still supported in some implementations

iBGP Template Example continued

- Use passwords on iBGP session
  - Not being paranoid, VERY necessary
  - It's a secret shared between you and your peer
  - If arriving packets don't have the correct MD5 hash, they are ignored
  - Helps defeat miscreants who wish to attack BGP sessions
- Powerful preventative tool, especially when combined with filters and the TTL "hack"

eBGP Template Example

- BGP damping
  - Do **NOT** use it unless you understand the impact
  - Do **NOT** use the vendor defaults without thinking
- Cisco's Soft Reconfiguration
  - Do **NOT** use unless troubleshooting – it will consume considerable amounts of extra memory for BGP
- Remove private ASes from announcements
  - Common omission today
- Use extensive filters, with "backup"
  - Use AS-path filters to backup prefix filters
  - Keep policy language for implementing policy, rather than basic filtering

eBGP Template Example continued

- Use password agreed between you and peer on eBGP session
- Use maximum-prefix tracking
  - Router will warn you if there are sudden increases in BGP table size, bringing down eBGP if desired
- Limit maximum as-path length inbound
- Log changes of neighbour state
  - …and monitor those logs!
- Make BGP admin distance higher than that of any IGP
  - Otherwise prefixes heard from outside your network could override your IGP!!

Mutually Agreed Norms for Routing Security

Industry Best Practices to ensure Security of the Routing System

MANRS

Routing Security

- Implement the recommendations in https://www.manrs.org/manrs

1. Prevent propagation of incorrect routing information
   - Filter BGP peers, in & out!
2. Prevent traffic with spoofed source addresses
   - BCP38 – Unicast Reverse Path Forwarding
3. Facilitate communication between network operators
   - NOC to NOC Communication
4. Facilitate validation of routing information
   - Route Origin Authorisation using RPKI

MANRS 1)

- Filtering prefixes inbound and outbound
  - RFC8212 requires all EBGP implementations to reject prefixes received and announced in the absence of any policy
- Advice: ***Never*** set up an EBGP session without inbound and outbound prefix filters
  - If full table required, block at least the bogons (see earlier)

MANRS 2)

- Implementing BCP 38
  - Unicast Reverse Path Forwarding
  - (Deny outbound traffic from customers which has spoofed source addresses)
- Advice: implement uRPF on ***all*** single-homed customer facing interfaces
  - Cheaper (CPU & RAM) than implementing packet filters

MANRS 3)

- Facilitate NOC to NOC communication
  - Know the **direct** NOC contacts for your customer Network Operators, your peer Network Operators, and your upstream Network Operators
  - This is not calling their "customer support line"
  - Make sure NOC contact info is part of any service contract
- Advice: NOC contact info for all connected Autonomous Networks is known to your NOC

MANRS 4)

- Facilitate validation of Routing Information
  - RPKI and Route Origin Authorisation (ROA)
  - All routes originated need to be signed to indicate that your AS is authorised to originate these routes
    - Helps secure the global routing system
- Advice: Sign ROAs for all originated routes using RPKI
  - And make sure all customer originated routes are also signed
  - Validate received routes from all peers
    - High priority to validated routes
    - Discard invalid routes
    - Low priority for unsigned routes

MANRS summary

- If your organisation supports and implements all 4 techniques in your network
  - Then join MANRS
  - https://www.manrs.org/join/
  - MANRS for Operators
  - MANRS for IXPs

MANRS

Summary

- Use configuration templates
- Standardise the configuration
- Be aware of standard “tricks” to avoid compromise of the BGP session
- Anything to make your life easier, network less prone to errors, network more likely to scale
- Implement the four fundamentals of MANRS
- It’s all about scaling – if your network won’t scale, then it won’t be successful

BGP Best Current Practices

ISP Workshops