BGP Best Current Practices

ISP Workshops

Last updated 24th June 2021

Acknowledgements

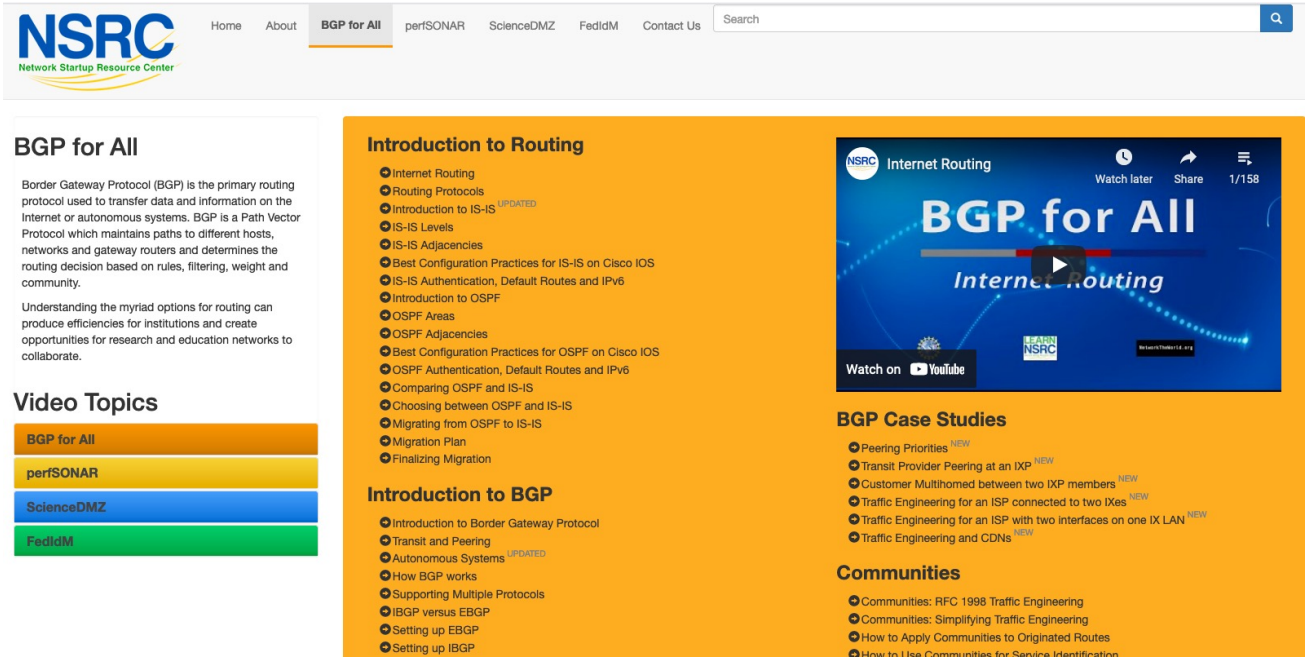
- This material originated from the Cisco ISP/IXP Workshop Programme developed by Philip Smith & Barry Greene

- Bug fixes and improvements are welcomed
  - Please email *workshop (at) bgp4all.com*

Philip Smith

BGP Videos

- NSRC has made a video recording of this presentation, as part of a library of BGP videos for the whole community to use:
  - https://learn.nsrc.org/bgp#bgp_best_practices



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 Default Behaviour

- 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 Default Behaviour

- 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
```

EBGP Default Behaviour

- FRR turns on RFC8212 support by default:
  - https://frrouting.org/

```
frr.pfs.lab(config)# router bgp 64512 view LAB
frr.pfs.lab(config-router)# bgp ?
<snip>
ebgp-requires-policy          Require in and out policy for eBGP peers (RFC8212)
<snip>
```

- No prefixes will be sent or received to external peers in the absence of inbound and outbound policy

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
  - May 2021: 498733 /24s in IPv4 table of 857992 prefixes
- The same is happening for /48s with IPv6
  - May 2021: 54043 /48s in IPv6 table of 115991 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

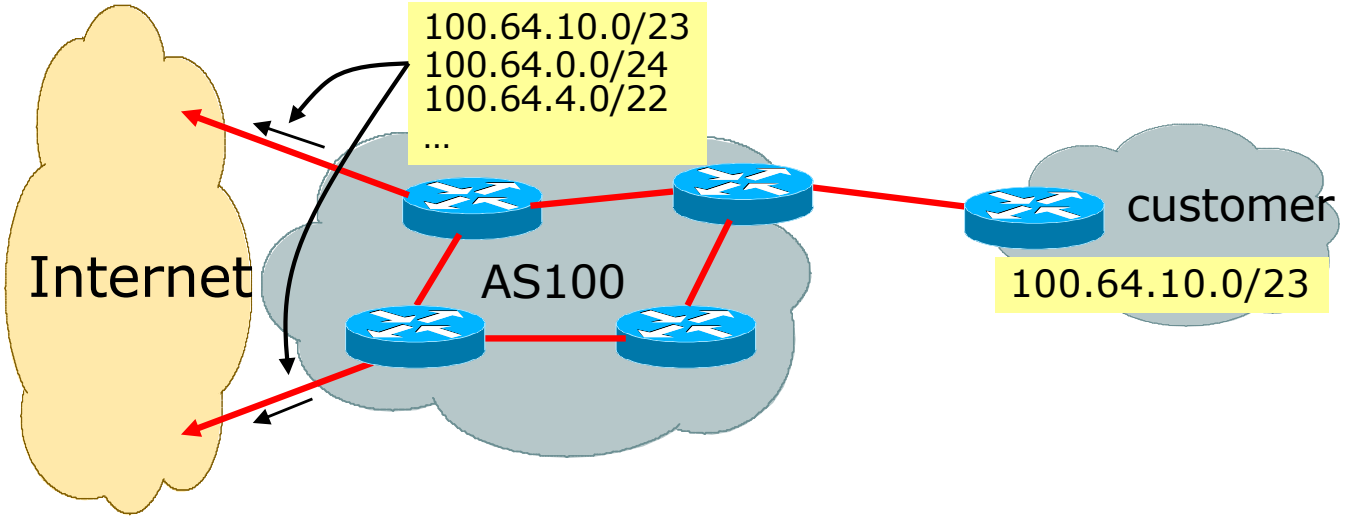
- Configuration Example

```
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 on the Internet. But now?
  - IPv4 run-out is having an impact

Aggregation – Example

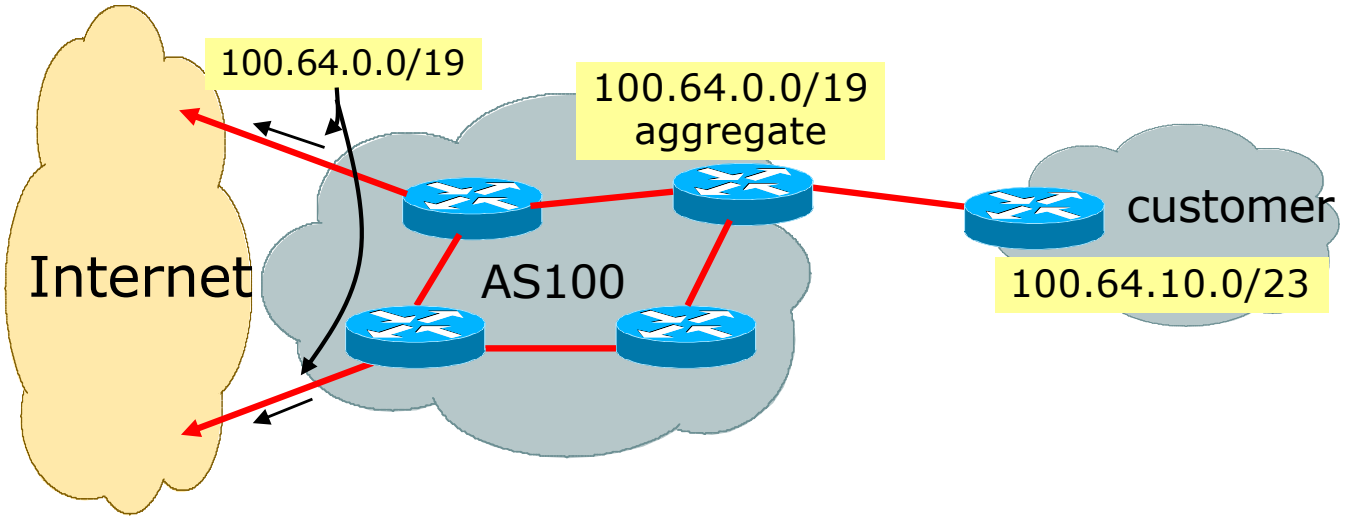


- Customer has /23 network assigned from AS100's /19 address block
- AS100 announces customers' individual networks to the Internet

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 (May 2021)

- Current IPv4 Internet Routing Table Statistics

BGP Routing Table Entries	857992
Prefixes after maximum aggregation	324830
Unique prefixes in Internet	407487
/24s announced	498733
ASNs in use	71226

  - (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

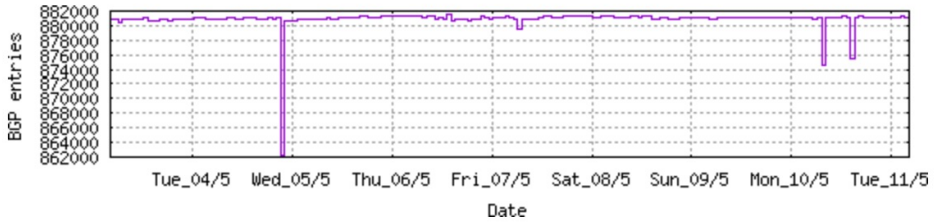
Efforts to Improve Aggregation
The CIDR Report

- 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

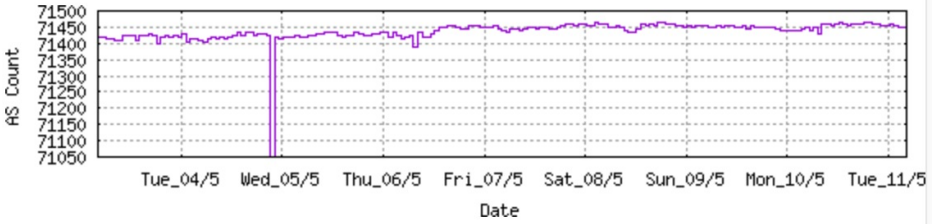
Date	Prefixes	CIDR Aggregated
04-05-21	880976	472773
05-05-21	880679	473912
06-05-21	881316	473846
07-05-21	880934	474664
08-05-21	881215	474805
09-05-21	881090	475032
10-05-21	881101	475222
11-05-21	881101	475332



Plot: BGP Table Size

AS Summary

71443	Number of ASes in routing system
25035	Number of ASes announcing only one prefix
8563	Largest number of prefixes announced by an AS AS8151: Uninet S.A. de C.V., MX
178348544	Largest address span announced by an AS (/32s) AS8003: GRS-DOD, US



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
72	AS6389	ORG+TRN	Originate:	10075904 /8.74	Transit:	60160 /16.12	BELLSOUTH-NET-BLK, 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	%
59	AS6389	BELLSOUTH-NET-BLK, US	902	722	17	197	705	78.16%

Long term deaggregator – BellSouth in the US

Prefix	AS Path	Aggregation Suggestion
12.81.120.0/24	4608 7575 6461 7018 6389	
12.130.209.0/24	4608 7575 2914 7018 6389 6389 6389 6389	
65.4.0.0/14	4608 7575 6461 7018 6389	
65.4.0.0/19	4608 7575 2914 7018 6389	
65.5.1.0/24	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.12.0/22	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.64.0/22	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.88.0/21	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.118.0/23	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.140.0/23	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.141.0/24	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.160.0/22	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.164.0/22	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.172.0/22	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.200.0/21	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.228.0/22	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.232.0/22	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.236.0/22	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.240.0/22	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.244.0/22	4608 7575 2914 7018 6389	
65.5.248.0/22	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.5.252.0/22	4608 7575 6461 7018 6389	- Withdrawn - matching aggregate 65.4.0.0/14 4608 7575 6461 7018 6389
65.6.192.0/21	4608 7575 2914 7018 6389	+ Announce - aggregate of 65.6.192.0/22 (4608 7575 2914 7018 6389) and 65.6.196.0/22 (4608 7575 2914 7018 6389)
65.6.192.0/22	4608 7575 2914 7018 6389	- Withdrawn - aggregated with 65.6.196.0/22 (4608 7575 2914 7018 6389)
65.6.196.0/22	4608 7575 2914 7018 6389	- Withdrawn - aggregated with 65.6.192.0/22 (4608 7575 2914 7018 6389)
65.7.64.0/18	4608 7575 2914 7018 6389	
65.7.116.0/22	4608 4826 3257 7018 6389	+ Announce - aggregate of 65.7.116.0/23 (4608 4826 3257 7018 6389) and 65.7.118.0/23 (4608 4826 3257 7018 6389)
65.7.116.0/24	4608 4826 3257 7018 6389	- Withdrawn - aggregated with 65.7.117.0/24 (4608 4826 3257 7018 6389)
65.7.117.0/24	4608 4826 3257 7018 6389	- Withdrawn - aggregated with 65.7.116.0/24 (4608 4826 3257 7018 6389)
65.7.118.0/24	4608 4826 3257 7018 6389	- Withdrawn - aggregated with 65.7.119.0/24 (4608 4826 3257 7018 6389)
65.7.119.0/24	4608 4826 3257 7018 6389	- Withdrawn - aggregated with 65.7.118.0/24 (4608 4826 3257 7018 6389)
65.7.120.0/22	4608 4826 3257 7018 6389	+ Announce - aggregate of 65.7.120.0/23 (4608 4826 3257 7018 6389) and 65.7.122.0/23 (4608 4826 3257 7018 6389)
65.7.120.0/24	4608 4826 3257 7018 6389	- Withdrawn - aggregated with 65.7.121.0/24 (4608 4826 3257 7018 6389)

Announced Prefixes

Rank	AS	Type	Originate Addr Space	(pfx)	Transit Addr space	(pfx)	Description
212	AS18566	ORG+TRN	Originate:	2836480 /10.56	Transit:	6912 /19.25	MEGAPATH5-, 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	%
27	AS18566	MEGAPATH5-, US	1990	1646	119	463	1527	76.73%

Prefix	AS Path	Aggregation Suggestion
64.6.160.0/23	4608 4826 3257 18566	
64.6.164.0/22	4608 4826 3257 18566	+ Announce - aggregate of 64.6.164.0/23 (4608 4826 3257 18566) and 64.6.166.0/23 (4608 4826 3257 18566)
64.6.164.0/23	4608 4826 3257 18566	- Withdrawn - aggregated with 64.6.166.0/23 (4608 4826 3257 18566)
64.6.166.0/24	4608 4826 3257 18566	- Withdrawn - aggregated with 64.6.167.0/24 (4608 4826 3257 18566)
64.6.167.0/24	4608 4826 3257 18566	- Withdrawn - aggregated with 64.6.166.0/24 (4608 4826 3257 18566)
64.50.206.0/23	4608 4826 3257 18566	
64.51.126.0/23	4608 4826 3257 18566	
64.81.0.0/16	4608 4826 3356 18566	
64.81.4.0/24	4608 4826 3257 18566	
64.81.16.0/20	4608 4826 3257 18566	+ Announce - aggregate of 64.81.16.0/21 (4608 4826 3257 18566) and 64.81.24.0/21 (4608 4826 3257 18566)
64.81.16.0/22	4608 4826 3257 18566	- Withdrawn - aggregated with 64.81.20.0/22 (4608 4826 3257 18566)
64.81.20.0/22	4608 4826 3257 18566	- Withdrawn - aggregated with 64.81.16.0/22 (4608 4826 3257 18566)
64.81.22.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.20.0/22 4608 4826 3257 18566
64.81.24.0/22	4608 4826 3257 18566	- Withdrawn - aggregated with 64.81.28.0/22 (4608 4826 3257 18566)
64.81.28.0/22	4608 4826 3257 18566	- Withdrawn - aggregated with 64.81.24.0/22 (4608 4826 3257 18566)
64.81.32.0/19	4608 4826 3257 18566	+ Announce - aggregate of 64.81.32.0/20 (4608 4826 3257 18566) and 64.81.48.0/20 (4608 4826 3257 18566)
64.81.32.0/20	4608 4826 3257 18566	- Withdrawn - aggregated with 64.81.48.0/20 (4608 4826 3257 18566)
64.81.32.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.32.0/20 4608 4826 3257 18566
64.81.33.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.32.0/20 4608 4826 3257 18566
64.81.34.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.32.0/20 4608 4826 3257 18566
64.81.35.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.32.0/20 4608 4826 3257 18566
64.81.36.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.32.0/20 4608 4826 3257 18566
64.81.37.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.32.0/20 4608 4826 3257 18566
64.81.38.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.32.0/20 4608 4826 3257 18566
64.81.39.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.32.0/20 4608 4826 3257 18566
64.81.40.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.32.0/20 4608 4826 3257 18566
64.81.44.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.32.0/20 4608 4826 3257 18566
64.81.48.0/20	4608 4826 3257 18566	- Withdrawn - aggregated with 64.81.32.0/20 (4608 4826 3257 18566)
64.81.48.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.48.0/20 4608 4826 3257 18566
64.81.50.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.48.0/20 4608 4826 3257 18566
64.81.53.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.48.0/20 4608 4826 3257 18566
64.81.54.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.48.0/20 4608 4826 3257 18566
64.81.57.0/24	4608 4826 3257 18566	- Withdrawn - matching aggregate 64.81.48.0/20 4608 4826 3257 18566

Long term deaggregator – Megapath in the US

Announced Prefixes

Rank	AS	Type	Originate Addr Space (pfx)	Transit Addr space (pfx)	Description
143	AS7545	ORG+TRN	Originate: 5090560 /9.72	Transit: 2886144 /10.54	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	%
5	AS7545	TPG-INTERNET-AP TPG Telecom Limited, AU	5758	5170	139	727	5031	87.37%

Prefix	AS Path	Aggregation Suggestion
14.2.0.0/19	4608 4739 7545	
14.2.32.0/19	4608 7575 7545	
14.2.32.0/21	4608 7575 7545	- Withdrawn - matching aggregate 14.2.32.0/19 4608 7575 7545
14.2.40.0/21	4608 7575 7545	- Withdrawn - matching aggregate 14.2.32.0/19 4608 7575 7545
14.2.48.0/21	4608 7575 7545	- Withdrawn - matching aggregate 14.2.32.0/19 4608 7575 7545
14.2.56.0/21	4608 7575 7545	- Withdrawn - matching aggregate 14.2.32.0/19 4608 7575 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)
14.2.128.0/18	4608 7575 7545	
14.2.192.0/20	4608 4739 7545	
14.200.0.0/14	4608 7575 7545	
14.200.0.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.1.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.2.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.3.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.4.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.5.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.6.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.7.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.8.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.9.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.10.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.11.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.12.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.13.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.14.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.15.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.16.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.17.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.18.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.19.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.20.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545
14.200.21.0/24	4608 7575 7545	- Withdrawn - matching aggregate 14.200.0.0/14 4608 7575 7545

Long term deaggregator – TPG in Australia

Announced Prefixes

```
Rank  AS       Type    Originate Addr Space  (pfx)   Transit Addr space  (pfx)  Description
56    AS12479        ORG+TRN Originate:   14214912 /8.24   Transit:     273664 /13.94 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       %
    3 AS12479      UNI2-AS, ES                                 6587   6079     60    568   6019  91.38%
```

Prefix	AS Path	Aggregation Suggestion
1.178.224.0/19	4608 4826 5511 12479	
37.11.0.0/16	4608 4826 5511 12479	
37.11.0.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.4.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.8.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.12.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.16.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.20.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.24.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.28.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.32.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.36.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.40.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.44.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.48.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.52.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.56.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.60.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.64.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.68.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.72.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.76.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.80.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.84.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.88.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.92.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.96.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.100.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.104.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.108.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.112.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.116.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479
37.11.120.0/22	4608 4826 5511 12479	- Withdrawn - matching aggregate 37.11.0.0/16 4608 4826 5511 12479

Long term deaggregator – Orange in Spain

Announced Prefixes

Rank	AS	Type	Originate Addr Space	(pfx)	Transit Addr space	(pfx)	Description
1	AS8003	ORIGIN	Originate: 178348544 /4.59		Transit: 0 /0.00		GRS-DOD, 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	%
55	AS8003	GRS-DOD, US	754	725	3	32	722	95.76%

Prefix	AS Path	Aggregation Suggestion
6.132.0.0/14	4777 6939 8003	
6.136.0.0/13	4777 6939 8003	
6.144.0.0/12	4777 6939 8003	
6.160.0.0/11	4777 6939 8003	
6.192.0.0/10	4777 6939 8003	
6.243.234.0/24	4777 6939 8003	- Withdrawn - matching aggregate 6.192.0.0/10 4777 6939 8003
7.0.0.0/8	4777 6939 8003	
7.0.0.0/22	4777 6939 8003	- Withdrawn - matching aggregate 7.0.0.0/8 4777 6939 8003
7.0.0.0/23	4777 6939 8003	- Withdrawn - matching aggregate 7.0.0.0/8 4777 6939 8003
7.0.0.0/24	4777 6939 8003	- Withdrawn - matching aggregate 7.0.0.0/8 4777 6939 8003
7.6.2.0/24	4777 6939 8003	- Withdrawn - matching aggregate 7.0.0.0/8 4777 6939 8003
7.7.7.0/24	4777 6939 8003	- Withdrawn - matching aggregate 7.0.0.0/8 4777 6939 8003
7.8.0.0/19	4777 6939 8003	- Withdrawn - matching aggregate 7.0.0.0/8 4777 6939 8003
7.8.0.0/22	4777 6939 8003	- Withdrawn - matching aggregate 7.0.0.0/8 4777 6939 8003
7.8.0.0/23	4777 6939 8003	- Withdrawn - matching aggregate 7.0.0.0/8 4777 6939 8003
7.8.0.0/24	4777 6939 8003	- Withdrawn - matching aggregate 7.0.0.0/8 4777 6939 8003
7.64.3.0/24	4777 6939 8003	- Withdrawn - matching aggregate 7.0.0.0/8 4777 6939 8003
11.0.0.0/8	4777 6939 8003	
11.0.0.0/13	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003
11.0.0.0/22	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003
11.0.0.0/24	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003
11.1.1.0/24	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003
11.4.0.0/15	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003
11.4.24.0/21	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003
11.4.144.0/21	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003
11.5.40.0/21	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003
11.5.88.0/21	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003
11.5.120.0/21	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003
11.5.240.0/21	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003
11.5.248.0/21	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003
11.6.148.0/23	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003
11.7.220.0/23	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003
11.8.0.0/15	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003
11.11.11.0/24	4777 6939 8003	- Withdrawn - matching aggregate 11.0.0.0/8 4777 6939 8003

New announcements by US DoD as from late January 2021 – but why the deaggregation??

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 10 May 2021 06:41 (UTC+1000)

50 Most active ASes for the past 14 days

RANK	ASN	UPDs	%	Prefixes	UPDs/Prefix	AS NAME
1	132839	485394	4.07%	293	1656.63	POWERLINE-AS-AP POWER LINE DATACENTER, HK
2	8151	208606	1.75%	8563	24.36	Uninet S.A. de C.V., MX
3	21859	189141	1.59%	844	224.10	ZNET, US
4	19429	188165	1.58%	1220	154.23	ETB - Colombia, CO
5	12696	185302	1.55%	20	9265.10	AXA-TECH Paris, FR
6	16509	156176	1.31%	5217	29.94	AMAZON-02, US
7	40065	118462	0.99%	567	208.93	CNSERVERS, US
8	33548	112983	0.95%	18	6276.83	UNWIRED-NOC, US
9	9829	88042	0.74%	1893	46.51	BSNL-NIB National Internet Backbone, IN
10	47331	83064	0.70%	7831	10.61	TTNET, TR
11	58424	76538	0.64%	49	1562.00	XINWEITELECOM-KH # 3BEo, Sangkat Beoun Prolit, Khan 7Makara, Phnom Penh., KH
12	13904	67563	0.57%	175	386.07	COSLINK, US
13	5972	60645	0.51%	2237	27.11	DNIC-ASBLK-05800-06055, US
14	7713	58881	0.49%	3529	16.68	TELKOMNET-AS-AP PT Telekomunikasi Indonesia, ID
15	58224	56960	0.48%	1598	35.64	TCI, IR
16	45899	53751	0.45%	3383	15.89	VNPT-AS-VN VNPT Corp, VN
17	9009	53322	0.45%	2443	21.83	M247, GB
18	36935	52393	0.44%	528	99.23	Vodafone-, EG
19	47730	47255	0.40%	2	23627.50	JEANDELESTRE, FR
20	12479	45458	0.38%	6694	6.79	UNI2-AS, ES
21	6147	44993	0.38%	1132	39.75	Telefonica del Peru S.A.A., PE
22	3551	41015	0.34%	246	166.73	Universidad Tecnologica de Panama, PA
23	9808	40697	0.34%	8226	4.95	CMNET-GD Guangdong Mobile Communication Co.Ltd., CN
24	39891	39521	0.33%	3266	12.10	ALJAWWALSTC-AS, SA
25	4755	37815	0.32%	2136	17.70	TATACOMM-AS TATA Communications formerly VSNL is Leading ISP, IN

50 Most active Prefixes for the past 14 days

RANK	PREFIX	UPDs	%	Origin AS -- AS NAME
1	118.127.80.0/22	37291	0.30%	18117 -- HARBOURMSP-AU-AP NTT Communications ICT Solutions, AU
2	95.173.129.0/24	36252	0.29%	8291 -- RSNET-AS RSNET, RU
3	85.208.229.0/24	36145	0.29%	209348 -- NPISP, IQ
4	216.238.254.0/23	35589	0.29%	13904 -- COSLINK, US
5	67.211.53.0/24	33713	0.27%	26405 -- HDCS, US
6	64.68.236.0/22	29685	0.24%	13904 -- COSLINK, US
7	168.243.93.0/24	29351	0.24%	262179 -- Telenetwork de El Salvador, SV
8	196.46.27.0/24	24987	0.20%	37207 -- NNPC, NG 37248 -- PHASE3TEL, NG
9	64.192.5.0/24	24223	0.20%	33548 -- UNWIRED-NOC, US
10	45.148.160.0/24	23635	0.19%	47730 -- JEANDELESTRE, FR
11	45.148.160.0/22	23620	0.19%	47730 -- JEANDELESTRE, FR
12	128.14.218.0/23	23113	0.19%	21859 -- ZNET, US
13	128.14.246.0/24	23111	0.19%	21859 -- ZNET, US
14	107.155.14.0/24	23110	0.19%	21859 -- ZNET, US
15	107.150.113.0/24	23108	0.19%	21859 -- ZNET, US
16	23.251.106.0/23	22932	0.19%	21859 -- ZNET, US
17	23.248.179.0/24	22930	0.19%	21859 -- ZNET, US
18	23.236.116.0/24	22836	0.18%	21859 -- ZNET, US
19	209.177.95.0/24	22617	0.18%	21859 -- ZNET, US
20	64.192.1.0/24	20347	0.16%	33548 -- UNWIRED-NOC, US
21	64.192.0.0/24	20284	0.16%	33548 -- UNWIRED-NOC, US
22	64.192.2.0/24	18900	0.15%	33548 -- UNWIRED-NOC, US
23	64.192.3.0/24	18550	0.15%	33548 -- UNWIRED-NOC, US
24	88.135.160.0/19	17598	0.14%	50188 -- KOLNET, PL
25	82.163.205.0/24	17164	0.14%	25160 -- VORBOSS_AS, GB
26	213.5.248.0/21	17080	0.14%	50188 -- KOLNET, PL
27	200.52.142.0/24	15323	0.12%	16531 -- TOPNET SA de CV, MX
28	99.194.200.0/22	14626	0.12%	22561 -- CENTURYLINK-LEGACY-LIGHTCORE, US
29	199.192.180.0/22	14590	0.12%	20101 -- MICHWAVE, US

The BGP IPv6 Instability Report

This report is updated daily. The current report
was generated on 11 May 2021 01:21 (UTC+1000)

50 Most active ASes for the past 14 days

RANK	ASN	UPDs	%	Prefixes	UPDs/Prefix	AS NAME
1	18403	925592	15.34%	878	1054.21	FPT-AS-AP The Corporation for Financing & Promoting Technology, VN
2	270590	302790	5.02%	33	9175.45	M Andrade dos Santos, BR
3	20473	215176	3.57%	2606	82.57	AS-CHOOPA, US
4	131429	168391	2.79%	403	417.84	MOBIFONE-AS-VN MOBIFONE Corporation, VN
5	267560	137252	2.27%	7	19607.43	PC CONNECT TELECOMUNICACOES LTDA, BR
6	27951	131285	2.18%	50	2625.70	Media Commerce Partners S.A, CO
7	28573	117432	1.95%	1241	94.63	CLARO S.A., BR
8	140485	104448	1.73%	454	230.06	CHINATELECOM-ZHEJIANG-HANGZHOU-5GC-NETWORK CHINATELECOM ZHEJIANG province HANGZHOU 5GC network, CN
9	14840	96252	1.59%	20	4812.60	BR.Digital Provider, BR
10	24429	95641	1.58%	9	10626.78	TAOBAO Zhejiang Taobao Network Co.,Ltd, CN
11	140313	91555	1.52%	304	301.17	CHINATELECOM-GUANGDONG-ZHAOQING-5G-NETWORK CHINATELECOM Guangdong province Zhaoqing 5G network, CN
12	140486	79902	1.32%	453	176.38	CHINATELECOM-ZHEJIANG-JINHUA-5GC-NETWORK CHINATELECOM ZHEJIANG province JINHUA 5GC network, CN
13	140345	71651	1.19%	262	273.48	CHINATELECOM-YUNNAN-SHENGJI-5G-NETWORK CHINATELECOM Yunnan province Shengji 5G network, CN
14	205908	71546	1.19%	6	11924.33	PETERCXY-NETWORKS PeterCxy Networks, DE
15	5588	65862	1.09%	16	4116.38	GTSCE GTS Central Europe / Antel Germany, CZ
16	9829	52634	0.87%	416	126.52	BSNL-NIB National Internet Backbone, IN
17	136440	50249	0.83%	1	50249.00	SASPL-AS-AP Sungard Availability Services (India) Private Limited, IN
18	12208	50197	0.83%	18	2788.72	TRUVISTA, US
19	60626	47015	0.78%	4	11753.75	LEASEWEBCDN, NL
20	52861	44895	0.74%	7	6413.57	SN Internet Navegantes Ltda ME, BR
21	38082	39431	0.65%	8	4928.88	IIT-TIG-AS-AP True International Gateway Co., Ltd., TH
22	31514	34053	0.56%	2	17026.50	INF-NET-AS, RU
23	266390	33949	0.56%	1	33949.00	Tajo Tecnologia Ltda, BR

50 Most active Prefixes for the past 14 days

RANK	PREFIX	UPDs	%	Origin AS -- AS NAME
1	2400:dc40::/32	50249	0.77%	136440 -- SASPL-AS-AP Sungard Availability Services (India) Private Limited, IN
2	2404:2280:147::/48	45431	0.69%	24429 -- TAOBAO Zhejiang Taobao Network Co.,Ltd, CN
3	2405:4000:800:8::/64	39396	0.60%	38082 -- IIT-TIG-AS-AP True International Gateway Co., Ltd., TH
4	2404:2280:106::/48	38682	0.59%	24429 -- TAOBAO Zhejiang Taobao Network Co.,Ltd, CN
5	2804:371c:8000::/48	33949	0.52%	266390 -- Tajo Tecnologia Ltda, BR
6	2402:53c0:2000::/36	28339	0.43%	137144 -- PASSLIR-AS Passit Media And Communication Pvt. Ltd., IN
7	2804:448c::/32	26546	0.41%	267638 -- Wind Telecomunicacao do Brasil Ltda - ME, BR
8	2a0e:b107:2f0::/44	25417	0.39%	205908 -- PETERCXY-NETWORKS PeterCxy Networks, DE
9	2a06:e881:2502::/48	24143	0.37%	205908 -- PETERCXY-NETWORKS PeterCxy Networks, DE
10	2801:17:4800::/48	23673	0.36%	27951 -- Media Commerce Partners S.A, CO
11	2a00:9d20:5353::/48	23548	0.36%	60626 -- LEASEWEBCDN, NL
12	2a00:9d20:53::/48	23461	0.36%	60626 -- LEASEWEBCDN, NL
13	2801:174:3::/48	23218	0.35%	27951 -- Media Commerce Partners S.A, CO
14	2801:1b6::/44	23215	0.35%	27951 -- Media Commerce Partners S.A, CO
15	2801:172:2::/48	23121	0.35%	27951 -- Media Commerce Partners S.A, CO
16	2804:4348::/40	22898	0.35%	267560 -- PC CONNECT TELECOMUNICACOES LTDA, BR
17	2804:4348:100::/40	22896	0.35%	267560 -- PC CONNECT TELECOMUNICACOES LTDA, BR
18	2804:4348:200::/40	22872	0.35%	267560 -- PC CONNECT TELECOMUNICACOES LTDA, BR
19	2804:4348:500::/40	22861	0.35%	267560 -- PC CONNECT TELECOMUNICACOES LTDA, BR
20	2804:4348:300::/40	22857	0.35%	267560 -- PC CONNECT TELECOMUNICACOES LTDA, BR
21	2804:4348:400::/40	22854	0.35%	267560 -- PC CONNECT TELECOMUNICACOES LTDA, BR
22	2a0e:b107:2f0::/48	21968	0.34%	205908 -- PETERCXY-NETWORKS PeterCxy Networks, DE
23	2620:13a:f001::/48	21836	0.33%	33110 -- UHSC-MBC, CA
24	2804:4ea8::/32	18999	0.29%	268275 -- GOOD TELECOM PROVEDOR DE INTERNET LTDA, BR
25	2a00:5240::/32	18953	0.29%	16331 -- TELE-ENTRE-FI-AS, FI
26	2804:20fc:1b00::/48	17891	0.27%	264525 -- Coelho Tecnologia, BR
27	2801:186::/44	17842	0.27%	27951 -- Media Commerce Partners S.A, CO
28	2a05:3181:ffff::/48	17027	0.26%	31514 -- INF-NET-AS, RU
29	2a05:3181::/32	17026	0.26%	31514 -- INF-NET-AS, RU
30	2001:4190::/32	16497	0.25%	5588 -- GTSCE GTS Central Europe / Antel Germany, CZ

Receiving Prefixes

Receiving Prefixes

- There are three scenarios for receiving prefixes from other ASes
  - 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 194.15.141.0

inetnum:        194.15.141.0 - 194.15.141.255
netname:        INETTECH
country:        SE
org:            ORG-ITAS2-RIPE
admin-c:        KEL5-RIPE
tech-c:         KEL5-RIPE
status:         ASSIGNED PI
mnt-by:         RIPE-NCC-END-MNT
mnt-by:         KURTIS-PP-MNT
mnt-routes:     KURTIS-PP-MNT
mnt-domains:    KURTIS-PP-MNT
created:        2003-12-04T09:33:09Z
last-modified:  2016-04-14T08:21:55Z
source:         RIPE
sponsoring-org: ORG-NIE1-RIPE
```

inetnum – means it is an address delegation to an entity

Assigned PI – 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 193.128.0.0/22

inetnum:        193.128.0.0 - 193.128.6.255
netname:        UK-PIPEX-19931014
country:        GB
org:            ORG-UA24-RIPE
admin-c:        WERT1-RIPE
tech-c:         UPHM1-RIPE
status:         ALLOCATED PA
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:        2018-07-30T09:42:04Z
last-modified:  2018-07-30T09:42:04Z
source:         RIPE # Filtered
```

inetnum – means it is an address delegation to an entity

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:
    - IRRToolSet:
      https://github.com/irrtoolset/irrtoolset
    - bgpq3:
      https://github.com/snar/bgpq3

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 it is 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

Interconnection Best Practices

- Types of Peering
- Using the PeeringDB and IXPDB
- Using 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)

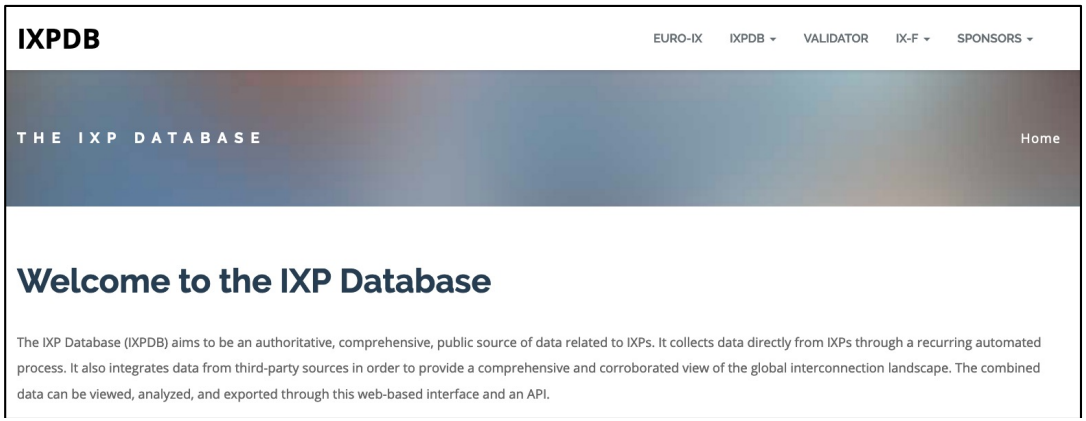
- Open Peering
  - Where an ISP publicly states that they will peer with all parties who approach them for peering
  - Commonly found at IXPs where ISP participates via the Route Server
- Selective Peering
  - Where an ISP's peering policy depends on the nature of the operator who requests peering with 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 AS operators 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) – (for North American peering)
  - Regional Peering Fora (European, Middle Eastern, Asian, Caribbean, Latin American)
  - Many countries now have their own Peering Fora

Types of Peering (5)

- The IXPDB documents IXPs and their participants around the world
  - https://ixpdb.euro-ix.net/en/
- All Internet Exchange Point operators should register their IXP in the database
  - IXPs using IXP Manager will have this happen as part of the IXP Manager set up
  - Provides the LAN IP addresses of each member to facilitate automation



PeeringDB

Search here for a network, IX, or facility.

Advanced Search

pfsinoz

HKIX

Organization	Hong Kong Internet eXchange Limited
Long Name	Hong Kong Internet Exchange
City	Hong Kong
Country	HK
Continental Region	Asia Pacific
Media Type	Ethernet
Protocols Supported	✓ Unicast IPv4 ○ Multicast ✓ IPv6
Notes	

Contact Information

Company Website	https://www.hkix.net/
Traffic Stats Website	https://www.hkix.net/hkix/stat/aggt/hkix-aggregate.html
Technical Email	noc@hkix.net
Technical Phone	+85239439900
Policy Email	info@hkix.net
Policy Phone	+85239438800

LAN

MTU	1500
DOT1Q	○
IPv6	2001:7fa:0:1::/64
IPv4	123.255.88.0/21

Local Facilities

Filter

Facility	Country	City
CUHK	Hong Kong	Hong Kong
MEGA Two (iAdvantage Hong Kong)	Hong Kong	Hong Kong
MEGA-i (iAdvantage Hong Kong)	Hong Kong	Hong Kong

Peers at this Exchange Point

Filter

Peer Name ASN	IPv4 IPv6	Speed Policy
ASGCNET HKIX Peering LAN 24167	123.255.91.53 2001:7fa:0:1::ca28:a135	10G Open
Asia Pacific Telecom HKIX Peering LAN 17709	123.255.91.86 2001:7fa:0:1::ca28:a156	10G Open
ASLINE HKIX Peering LAN 18013	123.255.92.13 2001:7fa:0:1::ca28:a20d	10G Open
AT&T AP - AS2687 HKIX Peering LAN 2687	123.255.91.46 2001:7fa:0:1::ca28:a12e	10G Selective
Automattic HKIX Peering LAN 2635	123.255.90.71 2001:7fa:0:1::ca28:a047	10G Open
Badoo Ltd HKIX Peering LAN 12678	123.255.90.220 None	2G Open
Baidu HKIX Peering LAN 55967	123.255.90.131 2001:7fa:0:1::ca28:a083	10G Open
Baidu HKIX Peering LAN 55967	123.255.91.61 2001:7fa:0:1::ca28:a13d	10G Open
Bayan Telecommunications Inc. HKIX Peering LAN 6648	123.255.91.45 2001:7fa:0:1::ca28:a12d	3G Open
BGP Network Limited HKIX Peering LAN 64050	123.255.91.177 2001:7fa:0:1::ca28:a1b1	100G Open
BIGHUB-ISP HKIX Peering LAN 137989	123.255.90.207 2001:7fa:0:1::ca28:a0cf	1G Open
BIGHUB-ISP HKIX Peering LAN	123.255.91.98	10G

PeeringDB

Search here for a network, IX, or facility.

Advanced Search

pfsinoz

Amazon.com Diamond Sponsor

Organization	Amazon.com
Also Known As	Amazon Web Services
Company Website	http://www.amazon.com
Primary ASN	16509
IRR as-set/route-set	AS-AMAZON
Route Server URL	
Looking Glass URL	
Network Type	Enterprise
IPv4 Prefixes	5000
IPv6 Prefixes	2000
Traffic Levels	Not Disclosed
Traffic Ratios	Balanced
Geographic Scope	Global
Protocols Supported	⊘ Unicast IPv4 ◯ Multicast ⊘ IPv6 ◯ Never via route servers
Last Updated	2019-12-29T14:56:38Z
Notes	If you have a connectivity issue to Amazon then please visit: • IPv4: http://ec2-reachability.amazonaws.com/ • IPv6: http://ipv6.ec2-reachability.amazonaws.com/ And include detail on prefixes you think you have a problem with if you contact our Ops alias. This will reduce time with troubleshooting. 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

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	400G ◯
AMS-IX 16509	80.249.210.217 2001:7f8:1::a501:6509:2	400G ◯
AMS-IX Chicago 16509	206.108.115.36 2001:504:38:1:0:a501:6509:1	100G ◯
AMS-IX Hong Kong 16509	103.247.139.10 2001:df0:296::a501:6509:1	100G ◯
AMS-IX India 16509	223.31.200.29 2001:e48:44:100b:0:a501:6509:2	10G ◯
AMS-IX India 16509	223.31.200.30 2001:e48:44:100b:0:a501:6509:1	10G ◯
BBIX Osaka 16509	218.100.9.24 2001:de8:c:2:0:1:6509:1	40G ◯
BBIX Tokyo 16509	218.100.6.52 2001:de8:c::1:6509:1	200G ◯
BBIX Tokyo 16509	218.100.6.207 2001:de8:c::1:6509:2	200G ◯
BCIX BCIX Peering LAN 16509	193.178.185.95 2001:7f8:19:1::407d:1	200G ◯
BIX.BG Main 16509	193.169.198.87 2001:7f8:58::407d:0:1	100G ◯
BNIX	194.53.172.122	100G

Private Peering Facilities

Filter

Facility ASN	Country City
151 Front Street West Toronto 16509	Canada Toronto

PeeringDB

Search here for a network, IX, or facility.

Advanced Search

pfsinoz

Telia Carrier

Organization	Telia Group
Also Known As	TeliaSonera, Telia, TSIC
Company Website	http://www.teliacarrier.com/
Primary ASN	1299
IRR as-set/route-set	RIPE::AS-TELIANET RIPE::AS-TELIANET-V6
Route Server URL	
Looking Glass URL	https://lg.telia.net/
Network Type	NSP
IPv4 Prefixes	426000
IPv6 Prefixes	40000
Traffic Levels	1 Tbps+
Traffic Ratios	Balanced
Geographic Scope	Global
Protocols Supported	⊘ Unicast IPv4 ○ Multicast ⊘ IPv6 ○ Never via route servers
Last Updated	2020-02-05T11:43:25Z
Notes	IPv4 + IPv6 Prefixes above would be actuals, not proposed max- prefix values. AS1299 is matching RPKI validation state and reject invalid prefixes from peers and customers. Our looking-glass marks validation state for all prefixes. Please review your registered ROAs to reduce number of invalid prefixes. All trouble ticket requests or support related emails should be sent to carrier-csc@teliacompany.com.

Peering Policy Information

Peering Policy	https://www.teliacarrier.com/dam/jcr:d1e83942-3db1-4334-a5f8-431578633d26/Telia_Carrier_Global_Peering_Policy.pdf
General Policy	Restrictive

Public Peering Exchange Points

Filter

Exchange ▾ / ASN	IPv4 / IPv6	Speed / RS Peer

No filter matches.
You may filter by **Exchange**, **ASN** or **Speed.**

Private Peering Facilities

Filter

Facility ▾ / ASN	Country / City
365 Data Centers Buffalo (BU1)	United States of America
1299	Buffalo
365 Data Centers Detroit (DT1)	United States of America
1299	Southfield
365 Data Centers Nashville (NA1)	United States of America
1299	Nashville
365 Data Centers Tampa (TA1)	United States of America
1299	Tampa
3U Rechenzentrum Berlin	Germany
1299	Berlin
Altus IT	Croatia
1299	Zagreb
Borovaya 57	Russia
1299	St. Petersburg
CE Colo Prague	Czechia
1299	Prague
CINECA - DC NaMeX	Italy
1299	Roma
COD BM-18	Russia
1299	St.Petersburg
Caldera21	Italy
1299	Milan
CarrierColo Berlin Luetzow (I/P/B/ site B)	Germany
1299	Berlin
Cologix MTL3	Canada
1299	Montreal

Screenshot

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 to be 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 in other presentations

Internet Routing Registry

- Which IRR database to use?
  - Members of a Regional Internet Registry are recommended to use their RIR's Internet Routing Registry instance
    - Usually managed via the RIR's member portal giving easy access for creation and update of objects
    - Provided as part of the RIR's services to its members
  - Operators who do not belong to any RIR generally use:
    - Their upstream transit provider's Routing Registry (if provided)
    - The RADB
      - https://www.radb.net
      - Note: Placing objects in the RADB requires an annual subscription fee

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 Routing Registry
    - 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
```

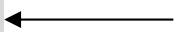
Examples of inbound and outbound policies – RPSL

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
- IXPDB
  - An industry Best Practice so that:
    - 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

Route Origin Authorisation

- Essential first step to secure the global routing system
- Covered in detail in separate presentation slide deck:
  - http://www.bgp4all.com.au/pfs/_media/workshops/02-rpki.pdf

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 ASNs in a typical AS_PATH in the Internet Routing Table 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

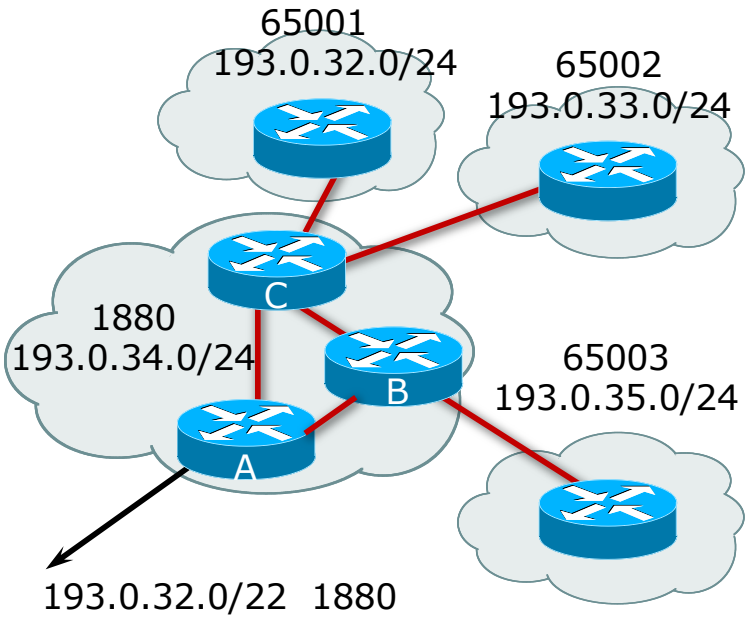
- A network operator with end-sites multihomed on their backbone (RFC2270)

  or

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

  or

- Within a BGP Confederation



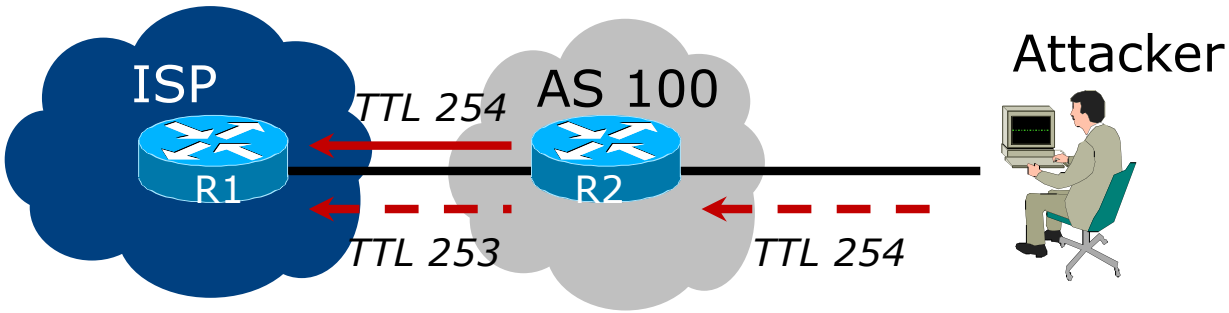
Private-AS – Removal

- 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 ASNs 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

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
   - Up-to-date details in Route and AS Objects, and PeeringDB
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
  - Up to date info in Route and AS Objects
  - Up to date AS info in PeeringDB
- 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 for 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 for CDN & Cloud Providers

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