1

Internet Transit, Peering & the Variants

APNIC (AS4608) | 2021-08-26
Che-Hoo Cheng

2

Disclaimers

- My personal viewpoints and observations mostly
  - Not representing the viewpoints of my employer
- A high-level half-technical talk
  - But still assuming audience to have basic understanding of how Internet operates
  - Not covering hands-on router commands
    - Please attend APNIC training instead for hands-on training
- Not able to cover all scenarios here because of limited time
  - Focusing more on networks with majority of their user base in one single country/economy
- Trying not to name names

3

How does the Internet Operate?

- Internet is a network of networks, composed of networks of ISPs and users
- User networks connect to ISPs
- Small ISPs connect to large ISPs
- Various networks (large or small) are **inter**connected with one another to form Internet

4

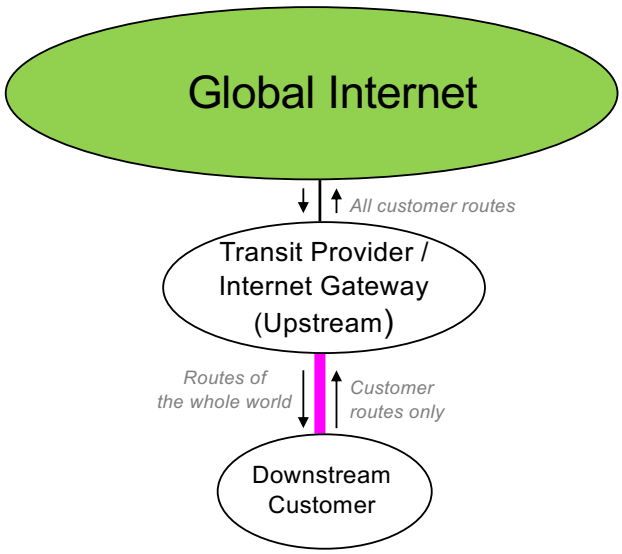
Autonomous Systems

A network on Internet is called Autonomous System (AS) which is represented by AS Number (ASN)

- ASN is unique around the world
  - APNIC is in charge of ASN assignment for AP region
- Used together with **BGP (Border Gateway Protocol)** for interconnections with **multiple** networks (or multi-homing)
- Networks having ASNs can be more independent, or portable
  - Together with portable IP addresses
  - Like what APNIC members are enjoying…

5

Ordinary Internet Transit Model



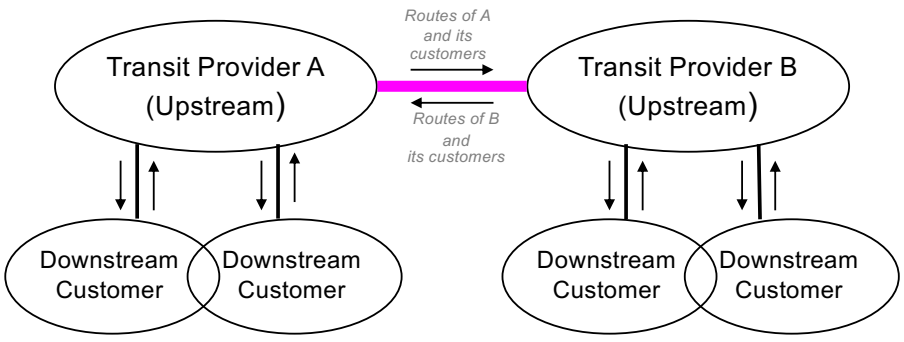
6

Internet Transit in General

- Networks **need to pay transit providers** to get to the whole Internet
  - Can connect to multiple transit providers for resilience and portability
- A few very large ISPs act as transit providers for the whole world (the so-called tier-1 networks) which do not need to pay others to get full Internet connectivity
  - Other ISPs must be transit customers of those tier-1 networks directly or indirectly in order to gain full connectivity
- Networks on Internet are trying to bypass transit providers as much as possible
  - By doing **direct peering** with various networks **for lower cost and higher performance**

7

Ordinary Peering Model



8

Peering in General

- ASes are interconnected/peered at Internet exchanges points (IXPs) or privately
- Interconnection/peering is among ISPs / data centres / content providers / content distribution network (CDN) providers / cloud services providers which have different ASNs using BGP protocol

For mutual benefits	• **For higher performance, lower latency and lower cost** • **Usually no settlement between peers and necessary circuit cost is shared**
Local Peering	• **Local-to-local traffic do NOT need to route through overseas** • **Important to local Internet development**
Between 2 ASes	• BLPA (Bi-Lateral Peering Agreement)
Among > 2 ASes	• MLPA (Multi-Lateral Peering Agreement)

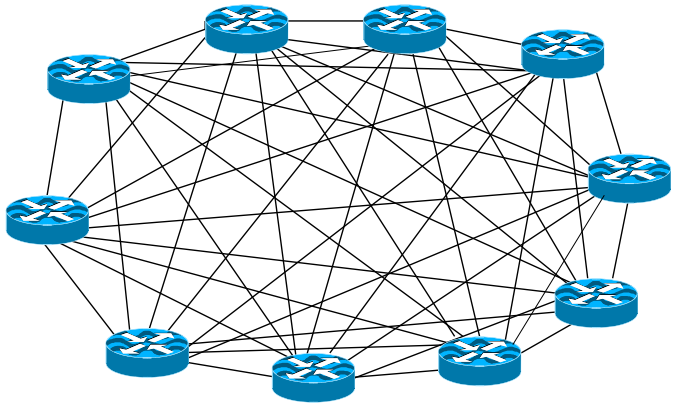
9

Private Peering

- A form of BLPA having dedicated point-to-point connection between 2 ASes
- Using cross-connect or local loop or IPL to interconnect
  - Cost is usually shared between 2 peers
- May have multiple connections between 2 ASes for resiliency
- Not quite cost-effective if traffic volume is not high enough
  - Spare bandwidth cannot be used for other traffic
- Not very scalable, in terms of # of circuits to use
  - ${}_nC_2$ physical connections for n ASes to peer fully with one another

10

Full Mesh for Peering



$_{10}C_2$ = 45 circuits

11

What is an Internet eXchange Point (IXP)?

- An IXP is a shared physical network infrastructure over which various Autonomous Systems can do easy peering with one another

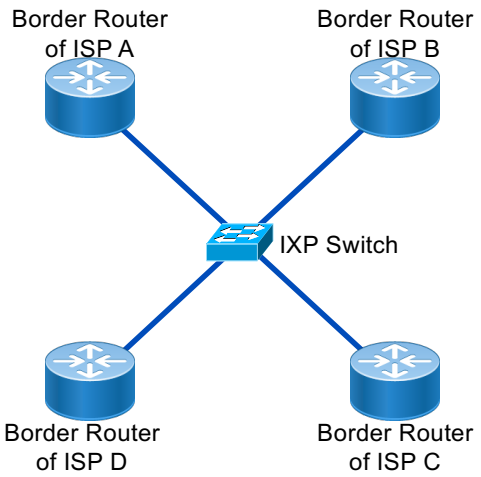
One physical connection to IXP can be used for interconnections with multiple networks

More cost-effective and scalable, especially with more participants

ASes to be served by IXP include Internet Service Providers (ISPs), Research & Education (R&E) Networks, Cloud Service Providers, Content Providers and Content Delivery Network (CDN) Providers

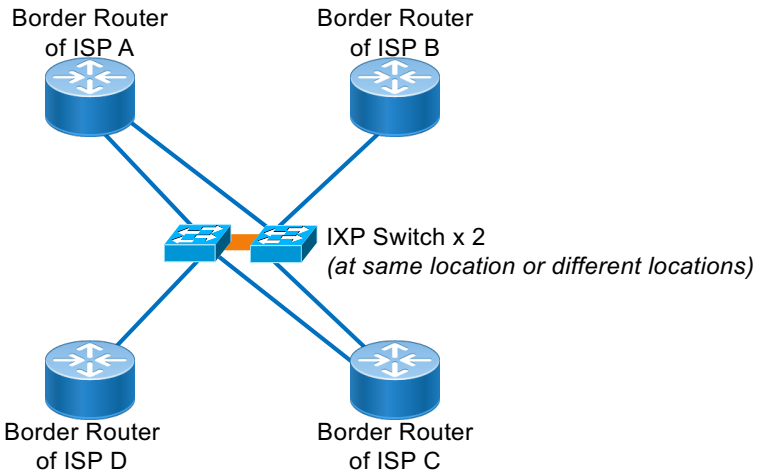
12

Simplest IXP Topology



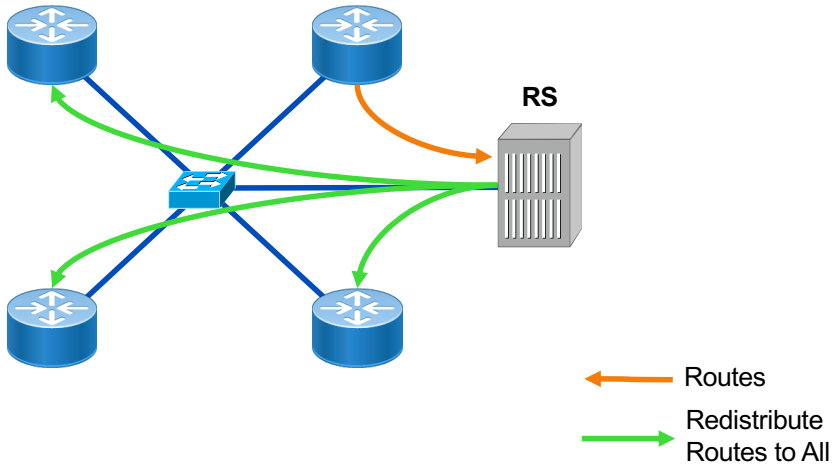
13

IXP Topology with Minimal Switch/Site Resilience



14

Adding Route Server for Multilateral Peering



15

Main Benefits of IXP

- One main objective of an IXP is to keep local traffic local
  - Important to local Internet development
- Helps bypass 3rd-party network infrastructure for easy interconnection and **direct** traffic exchange among participating networks
  - Reduced cost – cheaper connectivity
  - Enhanced network performance – faster speed
  - Reduced latency – lower delay
- Helps encourage development of more local content and local applications
  - Helps local data centre business and other businesses
- Everybody is benefited
  - The gain for each may be different but all will gain
  - At the end, it is the most important that end users or consumers are benefited
- Often considered as Critical Internet Infrastructure locally, regionally or globally

16

Value and Attractiveness of an IXP

- Proportional to the number of different networks (ASNs) connected and also the amount of traffic volume
- Snowball effect after reaching critical mass
  - The initial period usually is the hardest
    - Most will take wait-and-see approach
  - Gradually will have good mix of networks of different types
    - E.g. Eyeballs vs Content

17

IXPs and Data Centres

- They are natural partners
- Common situation in advanced/large metro cities
  - Multiple IXPs (commercial and/or not-for-profit) in one Data Centre
    - Usually in carrier-neutral Data Centres
    - A lot of data centres have their IXPs
  - One IXP in multiple Data Centres
    - The same layer-2 broadcast domain
    - Circuit cost of the backbone links is a burden to the IXP
  - A lot of telco's have their IXPs too
  - Healthy competition would be good
    - Customers have choices
    - Also for better resilience
    - As long as Keeping Local Traffic Local can be achieved, all good

18

Multi-Homing

- Having multiple transit connections plus doing peering
- Most service providers are doing it for improved connectivity and lower cost
  - They are doing peering extensively
- More and more common for large enterprise networks which have high traffic volume and/or have needs for better performance
  - Government networks
  - Universities and research institutes
  - Banks and financial institutions
  - Hospitals

19

Typical Set-up for Multi-Homing

- Connect to 2 to 3 Transit Providers with your own portable IP addresses and ASN
  - You may switch provider as you wish if service provided is not satisfactory
- Analyse and know your traffic with proper tools
  - But traffic is dynamic though so you need to do it continuously
- Do peering with as many close partners and target peers as possible
  - For higher performance, lower latency and possibly lower cost as well
  - With Private Peering and/or via IXPs – Local and/or Remote
  - Bilateral peering (BLPA) may help mitigate asymmetrical routing because of shorter AS Path
- Announce all your routes (including your downstream routes) to all BGP neighbours
  - But maybe with different preferences (using relevant LocalPref and BGP communities) and prefix-lengths for traffic engineering
- *Request for CDN Caches if meeting the traffic requirements – out of scope…*
- **NOTE:** Should set up different ASNs for your different services with different peering policies

20

Selecting IP Transit Provider(s)

- Should select those transit providers which have good connectivity with your target networks or good customer base in your target economies
  - After knowing more about your own traffic
- Best with interconnections with more of other transit providers
- Best with good resilience and good customer satisfaction
- Have POP closer to your network
  - Lower latency and/or lower cost
- Reasonable pricing
  - Not too expensive and not too cheap
- Do not sign contract of longer than 1 year
- Tier-1 providers or not not really matter

21

Direct Internet Access (DIA) Service

- An ambiguous term
- May mean dedicated Internet access service with supplied fixed non-portable IP addresses without need to run BGP
  - You cannot switch provider easily without renumbering
- May also mean IP transit service with BGP using your own portable IP addresses and ASN
  - You can switch provider easily without renumbering
  - This is the focus of this talk
- Physical ports usually Ethernet now
  - GE/10GE/100GE
  - Rate-limiting possible

22

Paying IP Transit Service

- Need circuit cost
  - Circuit can be cross-connect or local loop or IPL or any Layer-2 circuit
- You get full routes supposedly
  - Minus INVALID routes probably, if the provider does RPKI Invalid Routes Filtering
  - Having multiple transit providers plus peering gives you higher chance of getting the original "valid" routes to tackle route-hijacking cases
- Charging Model
  - Flat-Rate Charging – Rate-limiting possible
    - Lowest /Mbps price; Common; Good for stable traffic volume but not allowing traffic burst
  - Usage-Based Charging – With committed volume; Rate-limiting possible
    - 95th Percentile
      - Higher /Mbps price; Most common; Allowing traffic burst; some free traffic
    - Average
      - Highest /Mbps price; Not common; Allowing traffic burst; no free traffic
    - Should support usage aggregation for multiple ports, via Link Aggregation at one routing/location, or across multiple routers/locations

23

95th Percentile Usage-Based Charging

- Take traffic measurement every 5 mins
  - 8,640 samples every month (for 30-day months)
  - Inbound and outbound measurements to be checked separately
  - All measurements are ranked in descending order
  - The top 432 (5%) samples are discarded
  - **Then take the highest number of inbound and outbound measurements as the usage to charge**
  - *For the case of multiple ports, should add up the measurements of all samples for the same 5-min time period first before being ranked in descending order*
- You can have up to 1.5 days of traffic burst up to the port speed (or rate limit) every month without extra charge
  - Allowing you to take appropriate action if abnormal traffic volume is observed, such as attack traffic
- You need to select a committed volume which is usually your normal traffic volume
  - The higher the committed volume, the lower the /Mbps rate
- Should keep enough headroom so as to enjoy the free traffic burst
- 90th Percentile option was available long time ago…

24

Selecting IXP(s)

- Should select those IXPs with more of your target peers
  - After knowing more about your own traffic
- Best with more other networks connected for more peering opportunities
- Best with good resilience and good customer satisfaction
- Have POP closer to your network, to save circuit cost
- Reasonable pricing
  - Not too expensive and not too cheap
- Do not sign contract of longer than 1 year

25

Paying IXP Connections

- Port charge, plus circuit cost
  - Ethernet only – GE/10GE/40GE/100GE/400GE
- Port charge is usually fixed monthly charge according to port speed
  - Usually not usage based
    - Unless for value-added services or overseas connections
- Circuit can be cross-connect or local loop or IPL or any Layer-2 circuit
- Link aggregation should be common
  - To allow for gradual upgrade and added resilience
  - Should look at the total cost including the circuit cost

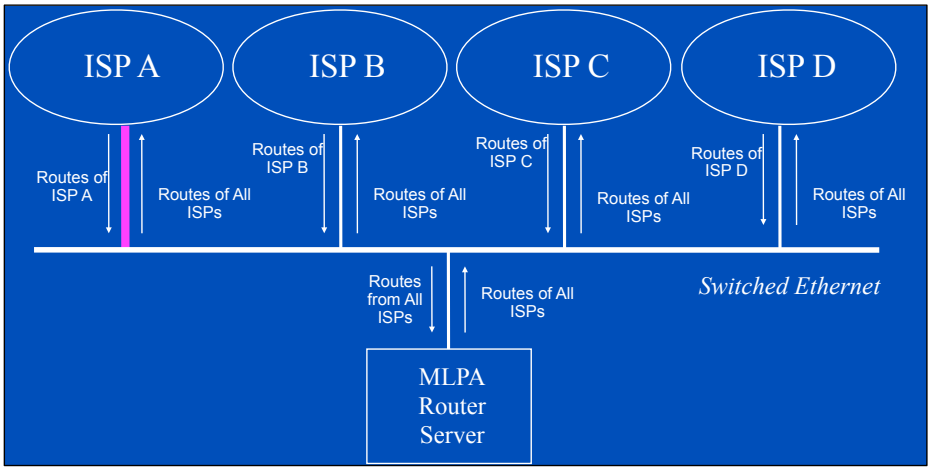
26

Private Peering vs IXP Connections

- Private Peering over dedicated circuit (not shared circuit) is best for your BLPA peers which have high volume of traffic
  - Better control and visibility, and best performance
  - May set up multiple circuits to the same peers for resiliency
    - Easier for circuit cost sharing as well
- IXP connections are good for most other peers
  - Cost-effective as traffic is aggregated
  - But you do not know easily whether your peers have congestion problem at their IXP connections

27

Multilateral Peering (MPLA) over IXP



28

Points to Note for MLPA

- Convenient for connecting to multiple networks
  - Just one BGP session
  - Facilitated by MLPA route server
- You have less control of your routing under MLPA though
  - Because MLPA route server select the best routes for you
- With BLPA (over IXP or privately), you should have better routes and connectivity
  - Possibly one AS hop less than MLPA
  - May get more routes from your BLPA peers than MLPA
  - Have direct control of your routing
- Do NOT blindly prefer all routes learnt from MLPA route server using higher LocalPref
  - **Doing BLPA more in addition to MLPA over your IXP connection is highly recommended**

29

Paid Peering

- Usually with incumbents in each country/economy which have large market share
  - One mostly, up to three in large countries/economies
- Cannot have settlement-free peering with them locally
  - They are more willing to doing peering when they are farther away from their home-base
  - They will only buy transit far away, such as in US
- If you want to get the best connectivity to them with the lowest latency and highest performance, you have to pay them for peering locally
  - Or, you may do free peering with them far away from their home-base but network latency may be much higher
- You are NOT recommended to use paid peering connection to one provider to reach the other provider in the same country/economy which also enforces paid peering
- It is the reality although not encouraged or preferred

30

Variant of Full Transit – Country Routes

- Transit service without full routes (offered by some providers)
  - Routes of identified country/economy only
    - Especially country/economy with multiple enormous incumbents
  - Mostly for paid peering
    - Save your trouble of dealing with multiple providers for paid peering
    - May cover multiple countries/economies with one single connection if supported
- May use different ASN from the normal transit service of the transit provider
- For better connectivity, not lower cost

31

Variant of Full Transit – Direct Routes

- Transit service without full routes (offered by a few providers)
  - Just peering routes of the transit provider
  - May not include routes which involve paid peering
- Lower /Mbps price
  - So as to enjoy lower cost but better performance and more direct connectivity
  - The main selling point

32

Variant of Usual Transit – Transit Sharing

- Sharing transit to other networks between two networks
  - More popular among R&E networks
    - As they are more willing to share connectivity with other R&E networks
- However, the downside is there will be many paths to reach one single destination which engineers may not be able to select the optimal paths easily
  - And the connectivity is being changed continuously
- Asymmetric routing is even worse than commercial networks
  - Suboptimal routing is very common
  - Should use direct bilateral peering over remote layer-2 VLAN more to alleviate the issue

33

Bandwidth Price Drop

- Submarine cable capacity price drop is on-going, especially hot routes
- IP Transit price drop is the fastest
- IXP Port price drop is not as fast
- Local Loop price drop is slow
  - Unless there is good local competition
- Cross-connect price drop is the slowest
  - Or may go reverse in certain cases
  - Still cheaper than IXP port though so it should be used for private peering if feasible
- Cost of Transit vs Cost of Peering
  - Become closer and closer
  - Should also look at quality of connectivity, i.e. cost-benefit

34

Reselling of Transit, Peering and IXP Services

- More and more common offerings by various service providers
- Regional/global layer-2 service providers with good automation via web portal offering these services are becoming popular
  - Supporting connections to various cloud providers as well
- Some IXPs even help you to connect to oversea IXPs with one physical port
- One physical Ethernet connection to multiple networks
  - Just use different VLANs for connections to different networks
  - It saves you a lot of time and effort with good automation
- The only question remains is whether you can get the performance you need with those connections
  - Ask for how you can monitor the performance continuously

35

Remote Peering Getting Popular

- Via regional/global layer-2 service providers or using dedicated IPLs
  - Connect remotely to overseas IXPs
  - Or do direct remote bilateral peering
- No need to set up POPs overseas
- Just need to set up proper cross-connect or local loop both ends
  - One-off work
  - You can select the physical paths more easily by looking at the latency info
  - If using remote layer-2 service, Just need one single physical circuit to support multiple connections with different VLANs
    - Do changes via web portals easily with virtual cross-connect service
    - Some service providers are resellers of IXPs which make the remote IXP connections even easier

36

General Guidelines for Peering

- Look for shorter AS Path's and physical paths to more networks
- Seek for more direct bi-lateral peering over more direct paths
- Asian networks should not peer with Asian networks outside of Asia
- Not peering with your downstream customers
  - Messy for routing
- Not pointing default to your peers
- Upgrade your peering circuit when it is almost full
- Never leak the routes that you learn from your transit providers to your peers
  - Set-up proper filters with good automation
    - Tag the ingress routes that you learn properly to facilitate good automation

37

Peering Policies in General

- Tier-1 Transit Providers – the strictest: traffic volume, ratio and locations
- Tier-2 / Tier-3 Transit Providers – strict; do not like MLPA of IXPs
- Eyeball Access Networks – less strict but still selective, except those with dominant market share which would be very strict
- Content / CDN / OTT / Cloud Networks – usually open
  - They prefer private peering more than peering over IXPs
  - They prefer to peer with those ISP which directly-owned eyeballs, not via transit providers
- Enterprises – usually selective
- *NOTE: Should set up different ASNs for your different services with different peering policies*

38

Default Route(s)

- For backup and resilience
  - You can have multiple default routes with different preferences
- Point to your most trusted transit provider(s)
  - But not point to your peers
- If you do RPKI/ROV invalid routes filtering in your network, then you must NOT have default routes
  - Otherwise, it is not effective

39

Routing Preferences for Outbound Traffic

1. Routes from downstream customers
   - Including routes from paid peering customers
2. Routes from peers
   - Private Peering > Bilateral Peering over IXPs > Multilateral Peering over IXPs
3. Routes from transit providers
4. Default routes (if you need it)
   - Last resort

- For Inbound Traffic, use BGP communities provided by your providers and peers to do better control

40

Internet Routing Registry (IRR)

- A lot of transit providers and some IXPs (for their MLPA) require you to register your route/route6 objects on IRR-DB with the correct Origin AS so as to automate the generation of their inbound prefix filters
  - If you are an APNIC member and all your routes to be announced are registered with APNIC, you can just use APNIC RR-DB
  - If you serve multiple AS'es, you need to set up as-set object
  - as-set object can help generate the full prefix list for inbound prefix filtering purpose by doing reverse look-up on IRR-DB
- With MyAPNIC, it is easy to also register ROAs for your prefixes (along with proper prefix lengths) for better routing security with RPKI
- Must match with what you announce
  - Otherwise you may lose connectivity or at the very least, you are showing your incompetence

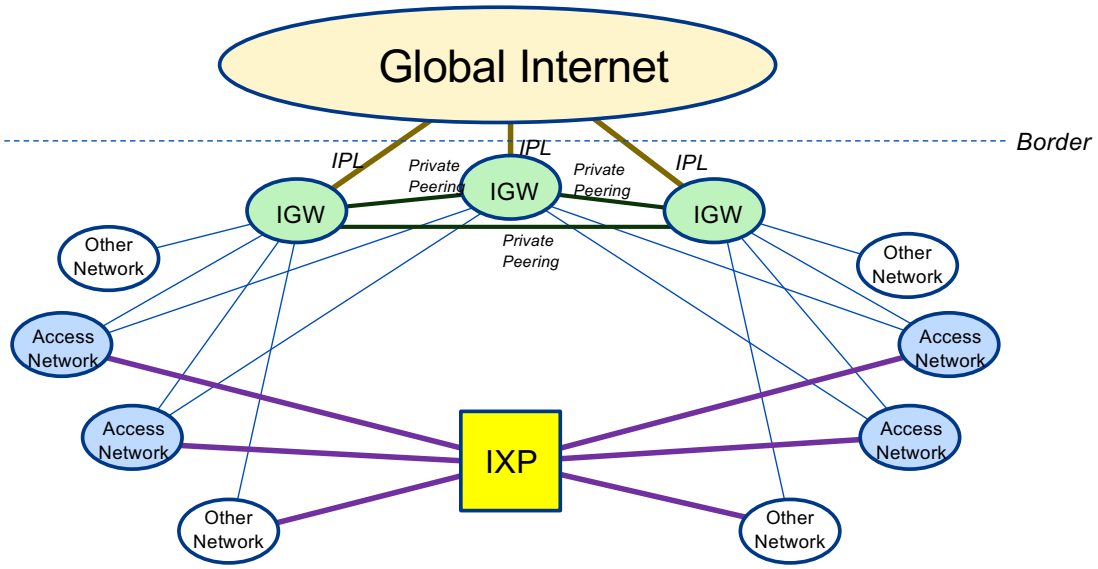
41

Only a Few International Gateways Allowed?

- International Gateways may just do private peering with full mesh among themselves for keeping local traffic local
  - But this may not help lower the cost of local traffic as local traffic is mixed with international traffic and this may affect the network performance as well
- To improve the situation, a local IXP can be set up just for local traffic
  - To separate local traffic from international traffic
  - Access networks and other networks can all connect to it for exchanging local traffic
  - But access networks should be separated from international gateways under the same groups with different ASNs in order to take full advantage of this set-up
- Same can be used to rely on dominant incumbents less

42

Possible Scenario for Improving Local Peering



43

Some Other Observations

- Connecting just to Tier-1 providers for IP transit may not give you the best connectivity to the whole world
  - Must complement with other methods if you are willing to invest for premium connectivity
- ISPs setting up remote POPs less and less common
  - Maybe just in key hubs/markets
  - Doing remote peering more
- Peering cannot replace transit totally
- Link aggregation can be used as intermediate upgrade path
- Do utilize peering with root / TLD DNS servers on IXPs better to enjoy faster and more stable DNS queries
- A lot of information can be found on PeeringDB for find peers for peering

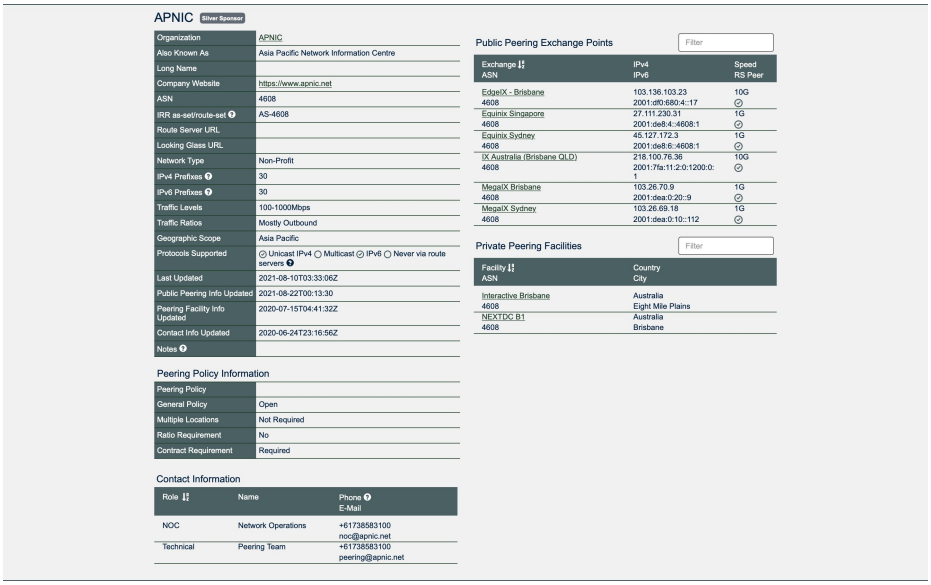
44

PeedingDB

- Set up a record of your ASN on www.peeringdb.com and tell everyone where you are (at which IXPs and/or data centres) and that you are willing to do BLPA
  - With contact info
- Also use it to find your potential BLPA peers
- Some bigger players do require you to have accurate and up-to-date record of your ASN on peeringDB before they peer with you

45

PeeringDB Entry of AS4608/APNIC



APNIC Silver Sponsor

Organization	APNIC
Also Known As	Asia Pacific Network Information Centre
Long Name	
Company Website	https://www.apnic.net
ASN	4608
IRR as-set/route-set	AS-4608
Route Server URL	
Looking Glass URL	
Network Type	Non-Profit
IPv4 Prefixes	30
IPv6 Prefixes	30
Traffic Levels	100-1000Mbps
Traffic Ratios	Mostly Outbound
Geographic Scope	Asia Pacific
Protocols Supported	Unicast IPv4 Multicast IPv6 Never via route servers
Last Updated	2021-08-10T03:33:06Z
Public Peering Info Updated	2021-08-22T00:13:30
Peering Facility Info Updated	2020-07-15T04:41:32Z
Contact Info Updated	2020-06-24T23:16:56Z
Notes	

Peering Policy Information

Peering Policy	
General Policy	Open
Multiple Locations	Not Required
Ratio Requirement	No
Contract Requirement	Required

Contact Information

Role	Name	Phone / E-Mail
NOC	Network Operations	+61738583100 noc@apnic.net
Technical	Peering Team	+61738583100 peering@apnic.net

Public Peering Exchange Points

Filter

Exchange / ASN	IPv4 / IPv6	Speed / RS Peer
EdgeIX - Brisbane 4608	103.136.103.23 2001:df0:680:4::17	10G
Equinix Singapore 4608	27.111.230.31 2001:de8:4::4608:1	1G
Equinix Sydney 4608	45.127.172.3 2001:de8:6::4608:1	1G
IX Australia (Brisbane QLD) 4608	218.100.76.36 2001:7fa:11:2:0:1200:0:1	10G
MegaIX Brisbane 4608	103.26.70.9 2001:dea:0:20::9	1G
MegaIX Sydney 4608	103.26.69.18 2001:dea:0:10::112	1G

Private Peering Facilities

Filter

Facility / ASN	Country / City
Interactive Brisbane 4608	Australia Eight Mile Plains
NEXTDC B1 4608	Australia Brisbane

46

Peering Managers

- A role within each network
  - Half business half technical role
  - May be with Product, Business Development, Engineering or Operations
  - Some networks may have Peering Committee to do joint decision
  - Only large networks can afford dedicated Peering Managers
- Negotiate peering with other networks
- Act according to Peering Policies set
- Look for mutual benefits with peers
  - And redundancy as well
- May apply discretion if necessary
  - Relationship matters
- May involve in buying IP Transit as well for better strategy alignment
- Attend Peering Forums to meet with other Peering Managers

47

Peering Forums

- Local, regional and global events for Peering Managers to meet one another
  - Conference plus meetings (a lot of meetings)
- Mostly facilitated by IXPs/DCs
- NOGs may also play a role
- One key part is Peering Personals for introducing your networks to your potential peers

48

Peering Personal for AS4608/APNIC

AS Number	4608
Network Type	Non-Profit/Education/Research
IPv4 Prefixes	30
IPv6 Prefixes	30
Traffic Volume	<1Gbps
Peering Policy	Open
Peering Locations	Brisbane: QLD-IX, MegaIX, EdgeIX Sydney: Equinix EIE, MegaIX Singapore: Equinix EIE Hong Kong: HKIX (coming soon)
PeeringDB Entry	as4608.peeringdb.net
Contact Information	peering@apnic.net

49

Network Engineers

- Have to know BGP routing very well and with good understanding of peering
- Need to manage and monitor the network and traffic continuously with good tools
  - Do proper route filtering at all BGP ports
  - Mitigate asymmetric routing wherever possible
  - Do traffic engineering whenever necessary
  - Do proper automation as much as possible
  - Must dedicate enough time to all BGP routing related work
- Better develop your Network Engineers by providing them good training
  - Such as attending APNIC training
  - Your investment will be paid off with a faster and more stable network

50

Final Remark

- Buying Internet transit helps you connect to Internet while investing in peering with good strategies and competent people to implement them helps you get much better Internet connectivity for your network and your users
- **Happy Peering!!!**

51

QUESTIONS?

52

Thank You!

53