Internet Exchange Points

APNIC
2023-05-30

Pacific Islands Telecommunications Association
Annual General Meeting

Primary Author: Che-Hoo Cheng, APNIC Director of Infrastructure & Development

Agenda

- Objectives of Transit and Peering and Benefits of IXP
- IXP Network Design and Operations
- Operation/Business/Governance Models for IXPs
- IXP Development Work of APNIC

Terry Sweetser

(2023-1989)=34 years of industry experience.

At times: APNIC Community Trainer, CTO, Consultant, Founder, Engineer, Coach, Mentor, Geek, Nerd, etc

APNIC Training Delivery Manager for South Asia and Oceania

Nationality: Australian

Languages: English

Quals: BEng MBA MIEAust

about.me/terry.sweetser

Disclaimer

There is no “One Size Fits All”

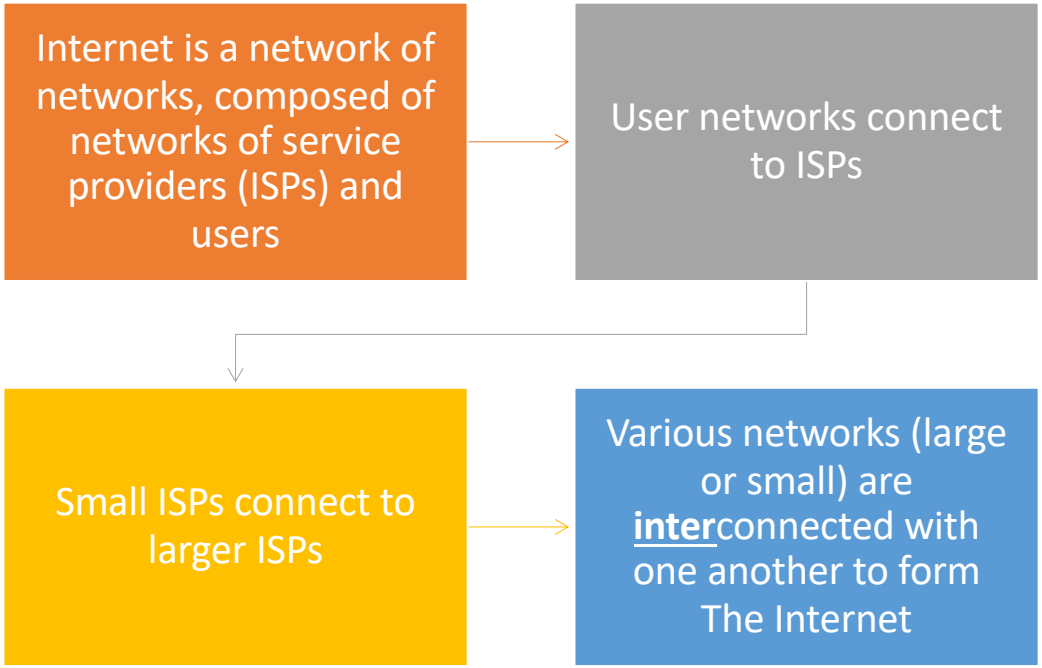
There are many different “recipes” for running an IXP

Just to provide hints, not answers

Cannot cover all scenarios here because of limited time

Objectives of Transit and Peering and Benefits of IXP

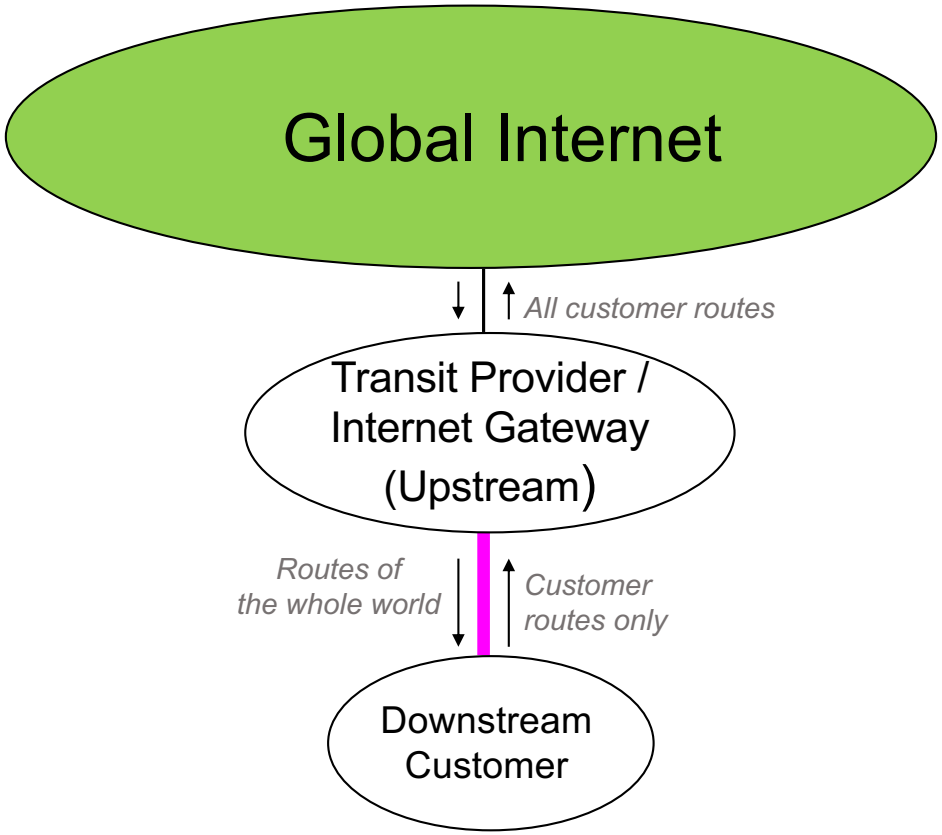
How Does Internet Operate?



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 APAC region (56 economies)
  - 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…

Ordinary Transit Model – Internet Gateway



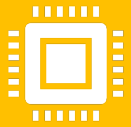
Transit in General

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

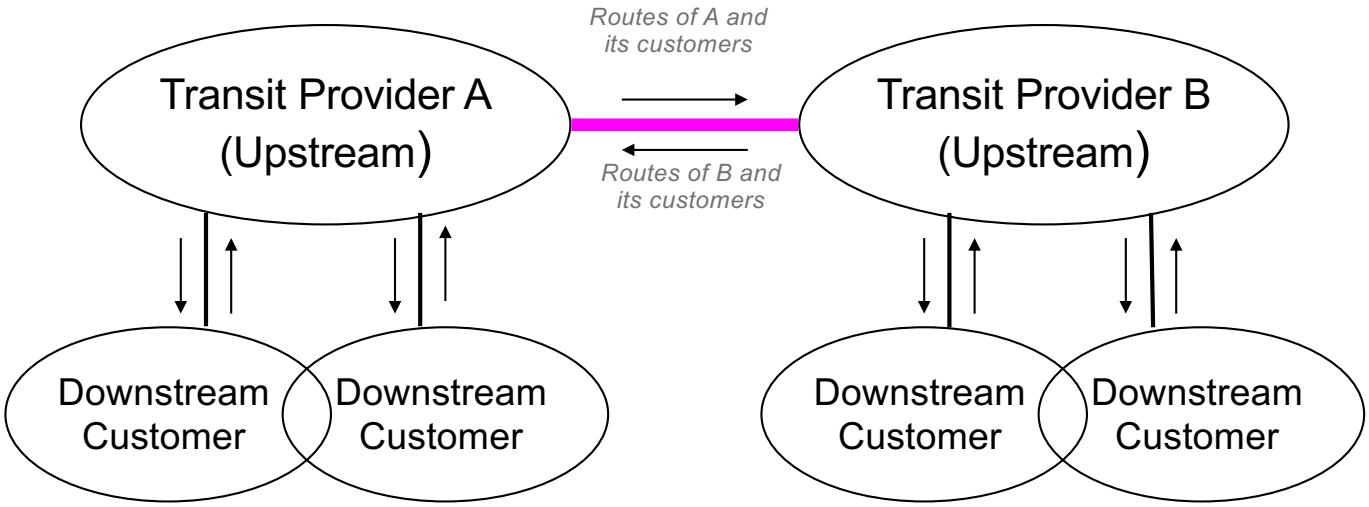
Other ISPs must be transit customers of those tier-1 networks directly or indirectly in order to gain full connectivity



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**

Ordinary Peering Model



Peering in General

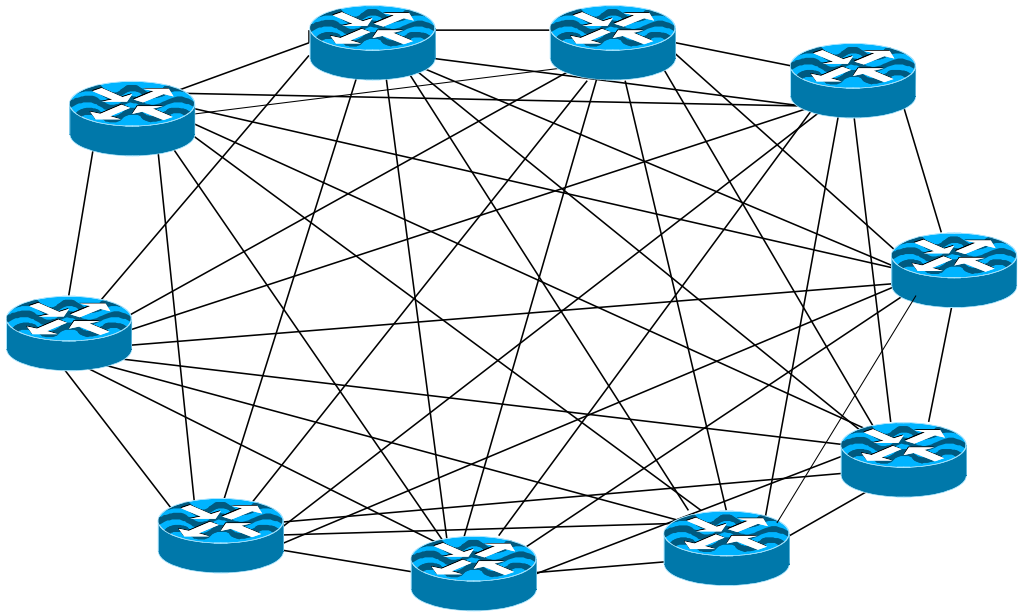
- AS's are interconnected/peered at Internet exchanges points (IXPs) or privately
- Interconnection/peering is among ISPs / data centres / content providers / content distribution network (CDN) providers / cloud services providers which have different ASNs using BGP protocol

For mutual benefits	• **For higher performance, lower latency and lower cost** • **Usually no settlement between peers and 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)

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
  - Spare bandwidth cannot be used for other traffic
  - Unless the traffic volume is really high
- Not very scalable
  - ${}_nC_2$ physical connections for n ASes to peer fully with one another

Full Mesh for Peering



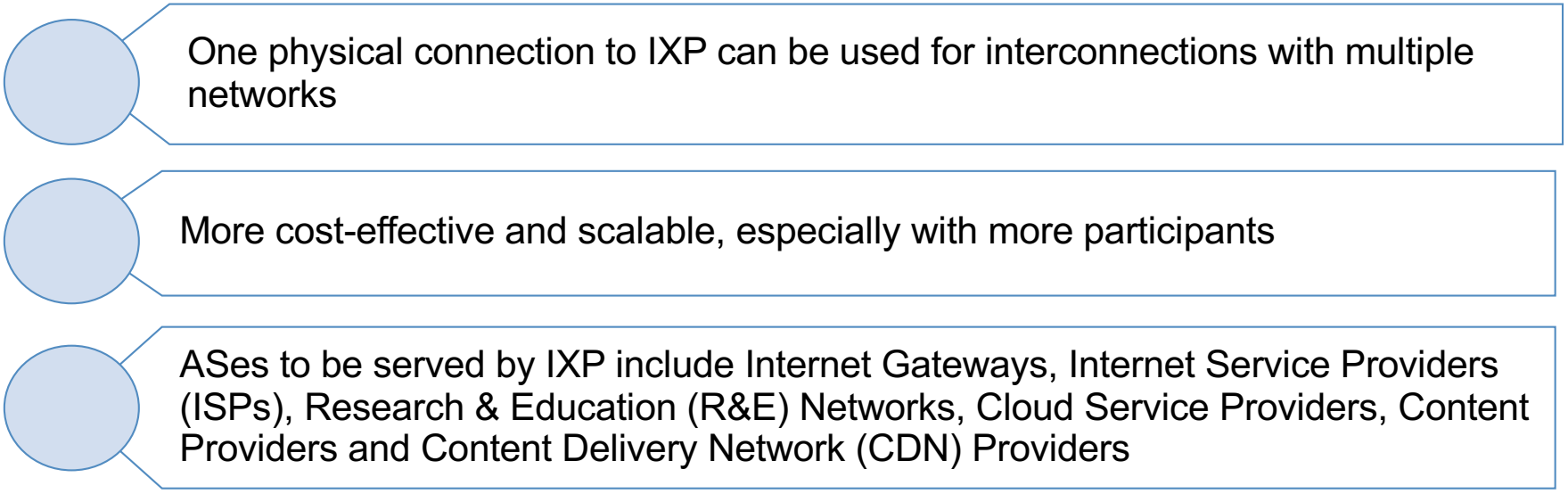
$_{10}C_2$ = 45 circuits

Total Number of Circuits for Full Interconnections

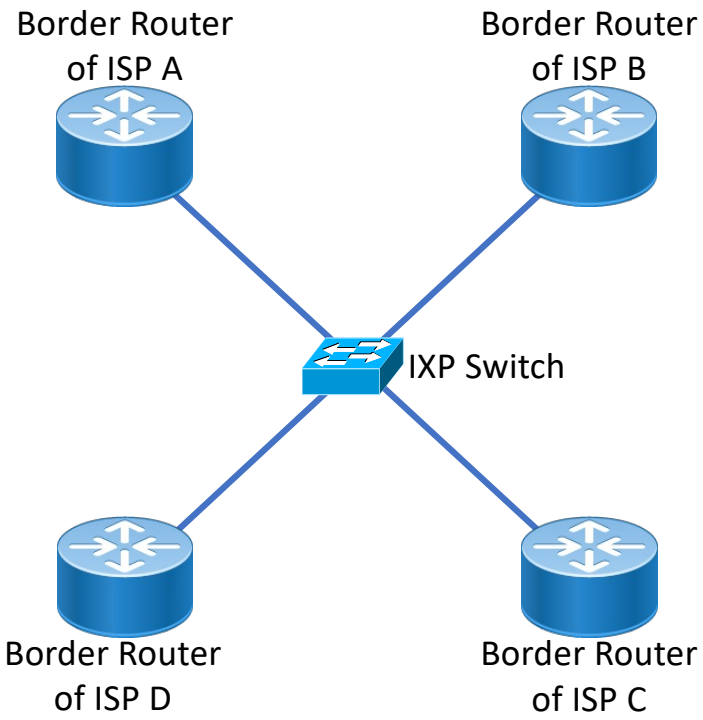
- 3 networks: ${}_3C_2$ = **3**; *plus 2 router ports per network*
- 4 networks: ${}_4C_2$ = **6**; *plus 3 router ports per network*
- 5 networks: ${}_5C_2$ = **10**; *plus 4 router ports per network*
- 6 networks: ${}_6C_2$ = **15**; *plus 5 router ports per network*
- 7 networks: ${}_7C_2$ = **21**; *plus 6 router ports per network*
- 8 networks: ${}_8C_2$ = **28**; *plus 7 router ports per network*
- 9 networks: ${}_9C_2$ = **36**; *plus 8 router ports per network*
- 10 networks: ${}_{10}C_2$ = **45**; *plus 9 router ports per network*

What is an Internet eXchange Point (IXP)?

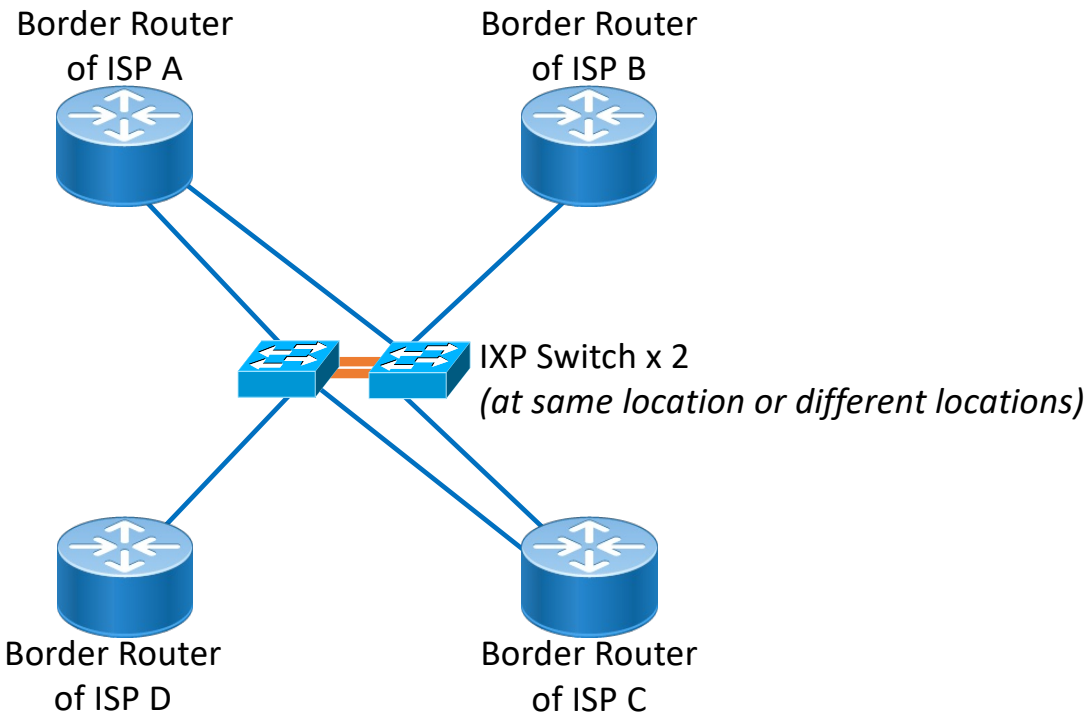
- An IXP is a shared physical network infrastructure over which various Autonomous Systems can do easy peering with one another



Simplest IXP Topology



IXP Topology with Minimal Switch/Site Resilience



Total Number of Circuits for IXP Connections

- 3 networks: **3** *or **6*** (vs **3**); plus 1 or 2* router ports per network* (vs 2)
- 4 networks: **4** *or **8*** (vs **6**); plus 1 or 2* router ports per network* (vs 3)
- 5 networks: **5** *or **10*** (vs **10**); plus 1 or 2* router ports per network* (vs 4)
- 6 networks: **6** *or **12*** (vs **15**); plus 1 or 2* router ports per network* (vs 5)
- 7 networks: **7** *or **14*** (vs **21**); plus 1 or 2* router ports per network* (vs 6)
- 8 networks: **8** *or **16*** (vs **28**); plus 1 or 2* router ports per network* (vs 7)
- 9 networks: **9** *or **18*** (vs **36**); plus 1 or 2* router ports per network* (vs 8)
- 10 networks: **10** *or **20*** (vs **45**); plus 1 or 2* router ports per network* (vs 9)

** 2 circuits and router ports per network for resilience*

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, often low fixed cost
  - Enhanced network performance – faster speed, larger capacities
  - Reduced latency – lower delay, switching at less than a millisecond
- Helps encourage development of more local content and local applications
  - Helps local data centre business and other businesses
- Everybody benefits
  - The gain for each may be different but all will gain
  - In the end, the most important outcome is customer/user experience improves
- Often considered as Critical Internet Infrastructure locally, regionally or globally

Pacific Islands

- Far from any other places
- External connectivity is very expensive
  - More submarine cables are being built for them
- Small markets because of small population
- Usually just a few ISPs but they may not be interconnected locally
- Local traffic across ISPs usually routed through US or Australia
- Local IXP is very much needed for helping Internet development
- Observed immediate benefits on Day 1 of set-up of Fiji-IXP
  - Much improved latency and high volume of traffic
- **Small land-locked economies have more or less similar issues**

Before Fiji-IXP was set up

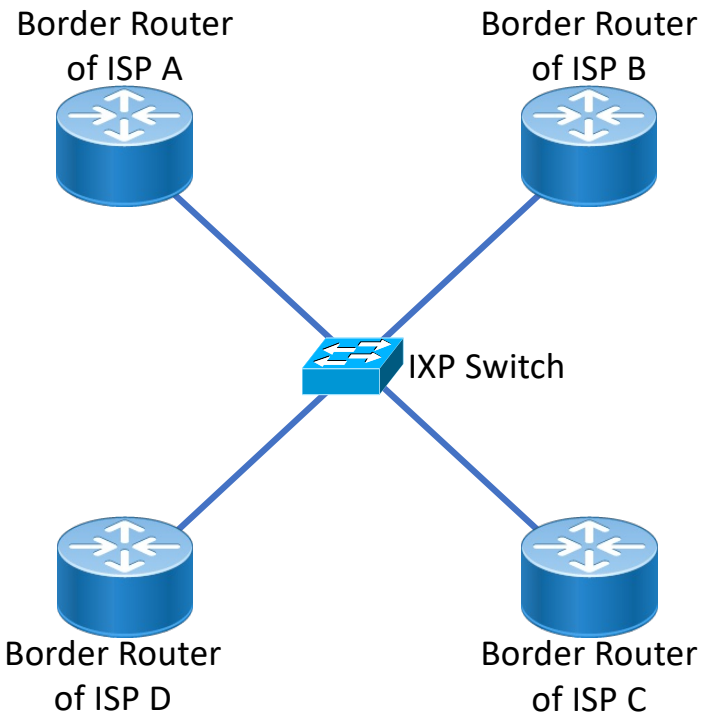
- A Fijian ISP in Suva accessing content at the University of the South Pacific in Suva
- Packet travels > 25,000km
- Physical distance < 10km
- Adding long latency
- Possibly high jitter too
- Using expensive submarine capacity
- Return path had similar issue

Value and Attractiveness of an IXP

- Proportional to the number of different networks (ASNs) connected and 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
    - Business and Consumer
    - Fixed and Mobile

IXP Network Design and Operations

Simplest IXP Topology



IXPs are Layer-2 Networks

- Switched Ethernet
  - One physical connection for interconnections with multiple networks
  - Only routers are allowed to connect to the switching fabric directly
- IXP participants can do direct Bilateral Peering (BLPA) over the layer 2 infrastructure anytime
- With Route Server added to the layer 2 infrastructure, IXP participants can also do Multilateral Peering (MLPA) for easier interconnections among everybody
  - Traffic exchange is direct and not going through the route server
- Those called themselves “IX” but serving layer-3 services are considered as transit providers
  - Note that IXPs, transit providers and data centres are not the same things

Evolution

IXP development is an evolutionary process done step by step

It can be improved over time, but picking the right initial neutral organisation & governance model and a neutral site at the start is important for future success

Some Local IXPs can evolve into Regional IXPs

Possible Steps for IXP Development

- Can be gradual, step by step (some steps can be skipped or be done at the same time)
  - Layer-2 network is the bare minimal
    - Can use private IP addresses if small amount of participants
  - Public IP addresses next
    - Legal entity issue
  - Site resilience is IMPORTANT while equipment resilience is also included
    - Has to have site resilience sooner or later
  - Route server(s) with ASN follows
    - RPKI/ROV consideration
  - Other value added services
    - DNS: Root / TLDs / Recursive
    - Shared CDN Caches (Transit for Cache Fill)

IP/ASN Resources for an IXP

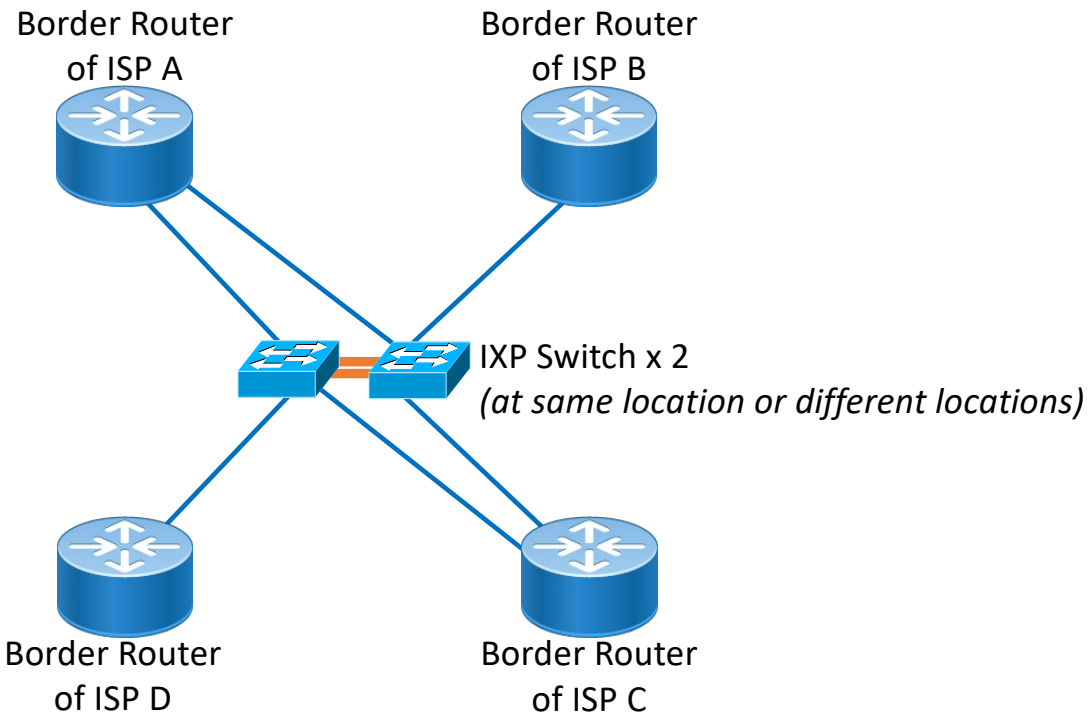
Considered as Critical Infrastructure under APNIC Policy

- Using public IP addresses and ASN is recommended
- IPv4: /24
- IPv6: /48
- ASN: 1 (for route server to facilitate multilateral peering)

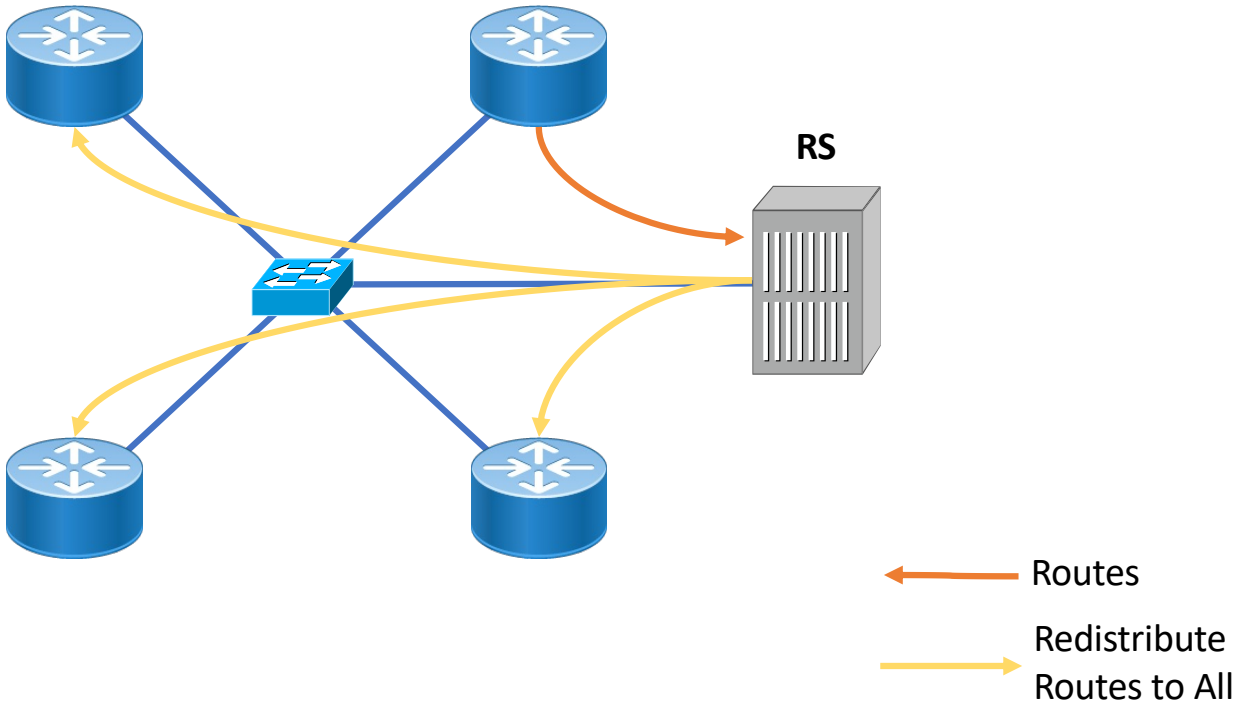
But IXP may need another network to provide transit

- Own servers such as network management & monitoring
- DNS anycast servers: Authoritative or Cache/Resolving/Recursive
- Shared CDN Caches for Participants (Capacity)
- Small network

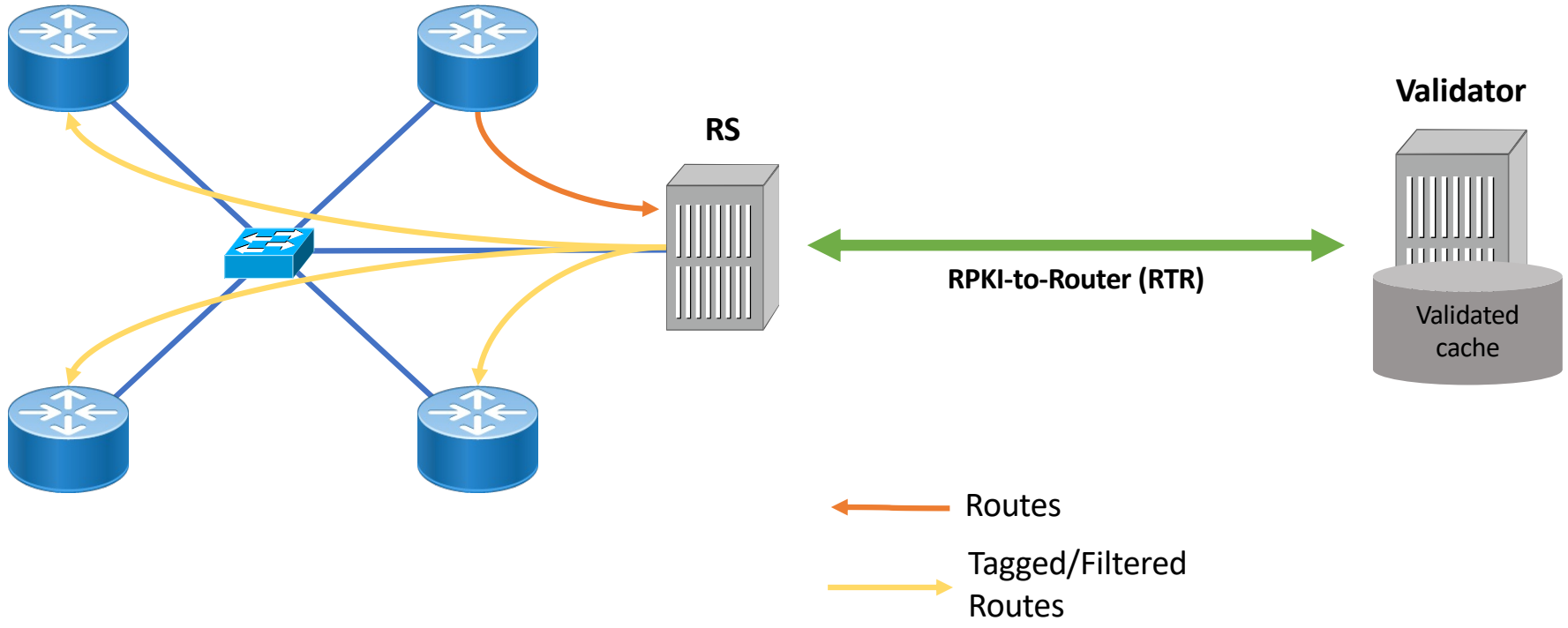
IXP Topology with Minimal Switch/Site Resilience



Adding Route Server for Multilateral Peering



Route Origin Validation (ROV) at IXP – via Route Server for Improved Routing Security



Scalability and Resilience Issues

IXPs were supposed to have no packet loss in its infrastructure

And with very low latency and very good resilience too

Become an issue when IXP grow beyond one switch

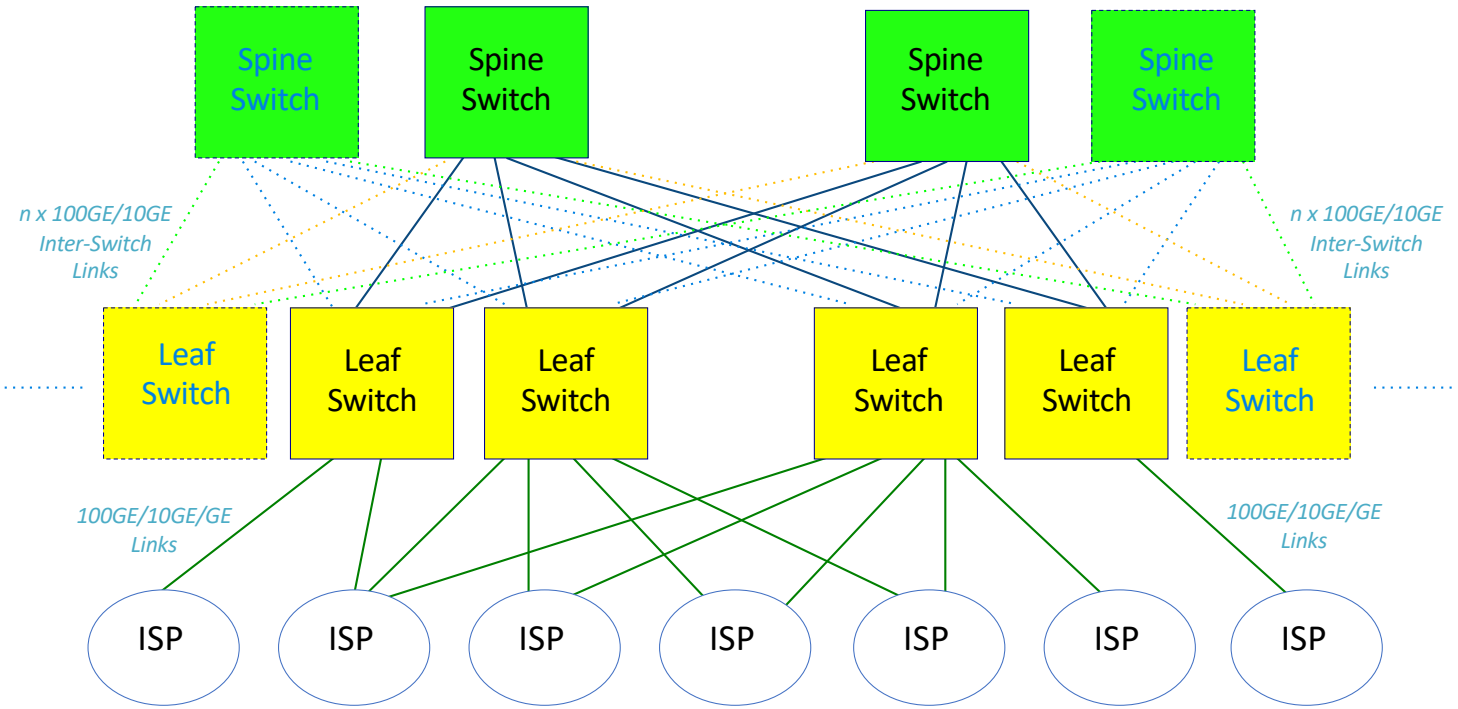
Due to not enough ports or expanding to multiple sites

Inter-switch links are the risk

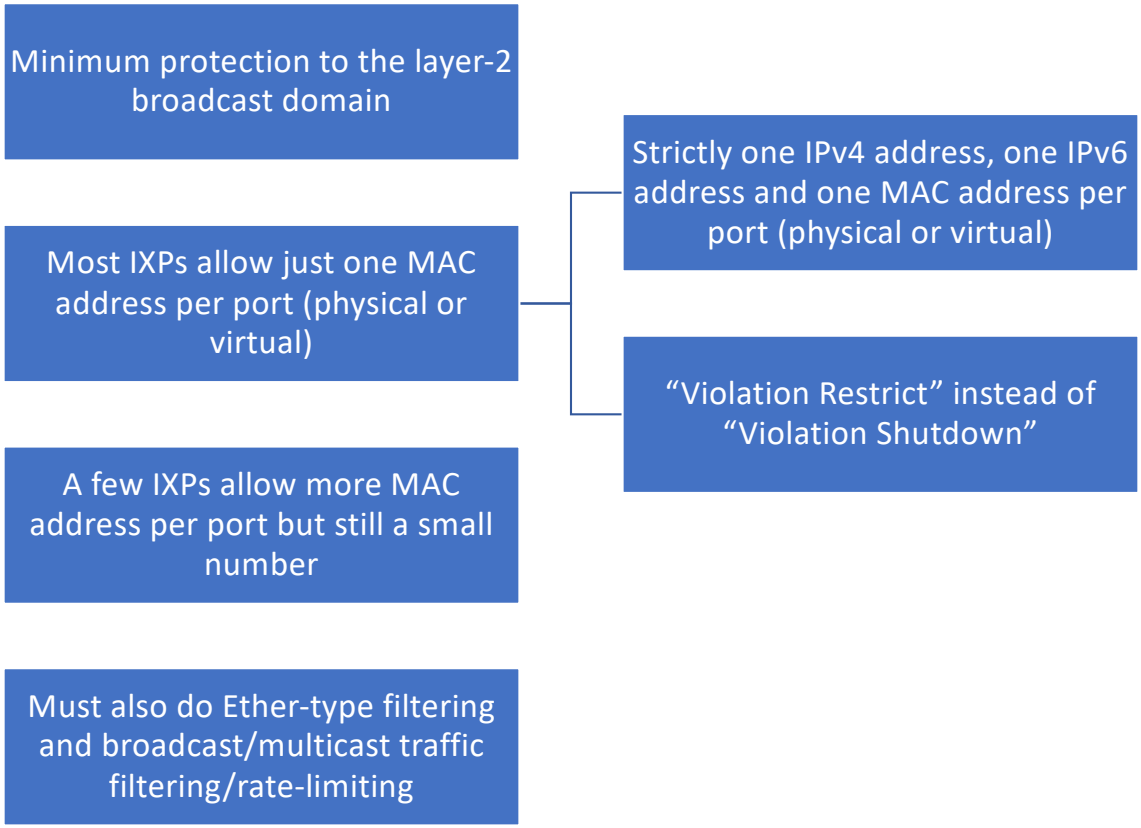
Over-subscription has to be minimised

Also need to minimise single point of failure

Spine-and-Leaf Architecture for Beyond 2 Switches



Port Security Is Important



Vulnerabilities of IXPs

Proxy ARP

- Why can't all router vendors have Proxy ARP disabled as default?
- Cannot stop it totally because of possible human errors
- Can only do regular monitoring by checking the ARP table
- EVPN over VxLAN technology may help but it is not a simple technology

Unknown Unicast Flooding

- May happen when there is asymmetric routing seen from an IXP
- Can be mitigated by sending proactive ARP check to all active addresses every hour or so
- EVPN over VxLAN technology may help but it is not a simple technology

Shared Buffer over Multiple Switch Ports

- Can cause trouble to multiple connections when there is big congestion on one port
  - Unknown to innocent participants which do not have any congestion
- Just be careful when choosing switch models
  - Also avoid switch models with small buffer

Operation/Business/Governance Models for IXPs

IXP Models

Commercial vs Non-profit

Developed economies vs Developing economies

Subsidized vs Self-financed

Government-led vs Industry-led

- No one single model which can suit all situations
- Neutrality is very important, but not always achieved

Commercial vs Non-Profit

Commercial set-up is free to do anything

No need to care about neutrality too much

IXP may be a service to help other business

Non-profit set-up tends to be more cautious

Neutrality is more important, at least to the target participants

Tend to be more independent from individual participants

Tend to offer fewer services

Developed Economies

IXPs are business	• Even for not-for-profit set-up • Less government involvement
Multiple IXPs	• Keen competition
But if they cannot keep intra-economy traffic local, someone needs to step up	• Government? Industry group? Customer pressure?

IXPs and Data Centres in Developed Economies

- They are natural partners
- Common situation in advanced metro cities
  - Multiple IXPs in one Data Centre
    - A lot of data centres have their IXPs
  - One IXP in multiple Data Centres
    - The same layer-2 broadcast domain
    - Circuit cost is a burden to the IXP
    - A lot of telco's have their IXPs
  - Healthy competition would be good
    - Customers have choices
    - Also for better resilience

Developing Economies

- Some do not have any IXPs yet
- Local traffic does not stay local
  - A lose-lose situation for everybody
- IXPs can help Internet development a lot
  - Better to be non-for-profit set-up
  - May need to start with subsidized model
  - May not be a business at all
  - Help from government is mostly needed
  - Active participation of the biggest players is also very important

Governance for Non-Profit IXPs

- Multi-stakeholder bottom-up approach is an approach with good acceptance by the community
  - E.g. membership-based model
- Government support is also important
  - At the very least, should get the proper license
- Be as inclusive as possible in order to provide maximum benefits to the whole community which it serves
- Should be fair and consistent to every participant or member
- Should be open and transparent as much as possible

Membership-Based Model

- Networks connected can join as members
  - Open membership vs closed membership
  - Mandatory membership vs optional membership
- Full Members with voting rights vs Associate Members without voting rights
  - Licensed ISPs vs others
  - Local legal entities vs overseas legal entities
- Governance by the Board elected
  - Policy decisions
  - Financial responsibilities
- Government's role
  - A facilitator at the very least

Neutral Locations

- May choose one of the followings as starting point:
  - University
  - Landing Station
  - Technology Park
  - **Carrier Neutral Commercial Data Center**
  - Government Data Center
- Having multiple carrier options with easy access is important
- Should maintain neutrality continuously
- Expansion to multiple sites within the same metro area can be done on Day 1 or be done gradually coupled with growth
  - This also helps improve neutrality as more options are provided

Geography

- IXP usually is NOT expanded beyond a metro area so as to avoid competing with IXP participants and to maintain neutrality
  - And simple port charging model can be used
- Usually start with the biggest city first and gradually set up separate infrastructure in other bigger cities one by one

To Add Value

- **Domain Name Infrastructure:** DNS infrastructure is very important to Internet operations so Root/TLD DNS server instance(s) should be connected directly to IXP for direct peering in order to benefit all participants for better DNS performance and resilience
- **Shared CDN Caches:** Connecting cache servers of popular content to the IXP will help everyone save bandwidth, but the cost of the transit bandwidth needed for cache-fill has to be properly shared by the ISPs benefited
  - Different CDN providers have different supported models
  - Need to think about long-term sustainability
- ***NOTE:*** *Transit for the above should NOT be used for providing usual transit service to IXP participants so as to maintain neutrality*

Financial Model for Non-Profit IXPs

- Seed money?
  - Perhaps from subsidies and/or sponsorship/donation
- IXPs need money to operate continuously
  - Need to have a long-term sustainable finance model for full cost recovery of CapEx and OpEx
    - All cost should be covered
  - Should not forget about the limited lifetime of the equipment used so must save money for future major equipment upgrade say once every 5 years
    - By setting up a reserve fund
- Subsidies, sponsorship or volunteers support may not be stable
- Those networks which are benefited should all contribute in a fair way
  - A charging model should be devised to help achieve that
- A good financial model will help sustain the IXP operations in long term

Charging Model

- Simple port charging model is the most common model in the industry
  - Monthly Recurrent Charge (MRC) provides stable income
  - 100GE port MRC / 10GE port MRC = 3 to 6
  - 10GE port MRC / GE port MRC = 1 to 4
  - Volume discount may be applied to encourage more connections for various purposes
    - With or without Link Aggregation
    - For better resilience and/or more bandwidth buffer (headroom)
  - NRC (Non-Recurrent Charge) charged with no contract or no NRC for a fixed-term contract
- Charge by usage for shared CDN cache service
  - Accurate usage accounting by trusted party is crucial

Which Models Can Sustain?

Pure Business Model	• IXP alone cannot make big money • IXP may just be a value-added service • Little issue if it is with good financial model
Subsidized Model	• Funding may or may not be long-term • Little issue if funding is long-term
Model relying on sponsorship and/or volunteers	• May be risky as sponsorship or support of volunteers is not guaranteed unless it is small enough and without growth
Membership-based Model	• Most neutral • Proper governance is important • Need to have good financial model for long-term sustainability

Politics Involved in Early IXP Development

- Usually larger ISPs like IXP less than smaller ISPs
  - Smaller ISPs are target customers of larger ISPs so larger ISPs have fear of losing market share
  - Dominant ISPs having >60% of market share think they are the IXPs
- Larger ISPs refuse to connect to IXP making the value of IXP lower
  - But IXPs do help provide a level-playing field for smaller ISPs
- There are multiple possible mitigation options for that but in any case, larger ISPs need to collaborate
  - E.g. separating access networks from Internet gateway / transit network
- If hurting the goal of “Keeping Local Traffic Local”, then it is lose-lose to everybody
- Government involvement may help or may hurt the case
  - It depends on the relationship between the industry and the government
  - Forcing large ISPs to do peering may not achieve the expected outcomes
- Having an IXP is NOT a magic wand to solve all the issues but collaborative spirit is ☺

Independent Legal Entity for IXP?

- Not critical but highly desirable
- Allow for demonstration of independence and/or neutrality
  - Such as jointly owned by members
  - Or a separate company from the mother company
- Possible use of the legal entity
  - License
  - Agreements with participants / members
    - Bank accounts for collecting incomes and spending
  - APNIC membership
    - Do allow transitional arrangement

Government Funding for IXPs?

- More needed during infancy stage of IXP development
  - Government usually can only provide one-off funding support
- For long-term, IXPs need to have a long-term sustainable financial model
  - Better be together with bottom-up industry-led governance for IXP
  - Align with bottom-up multi-stakeholder approach

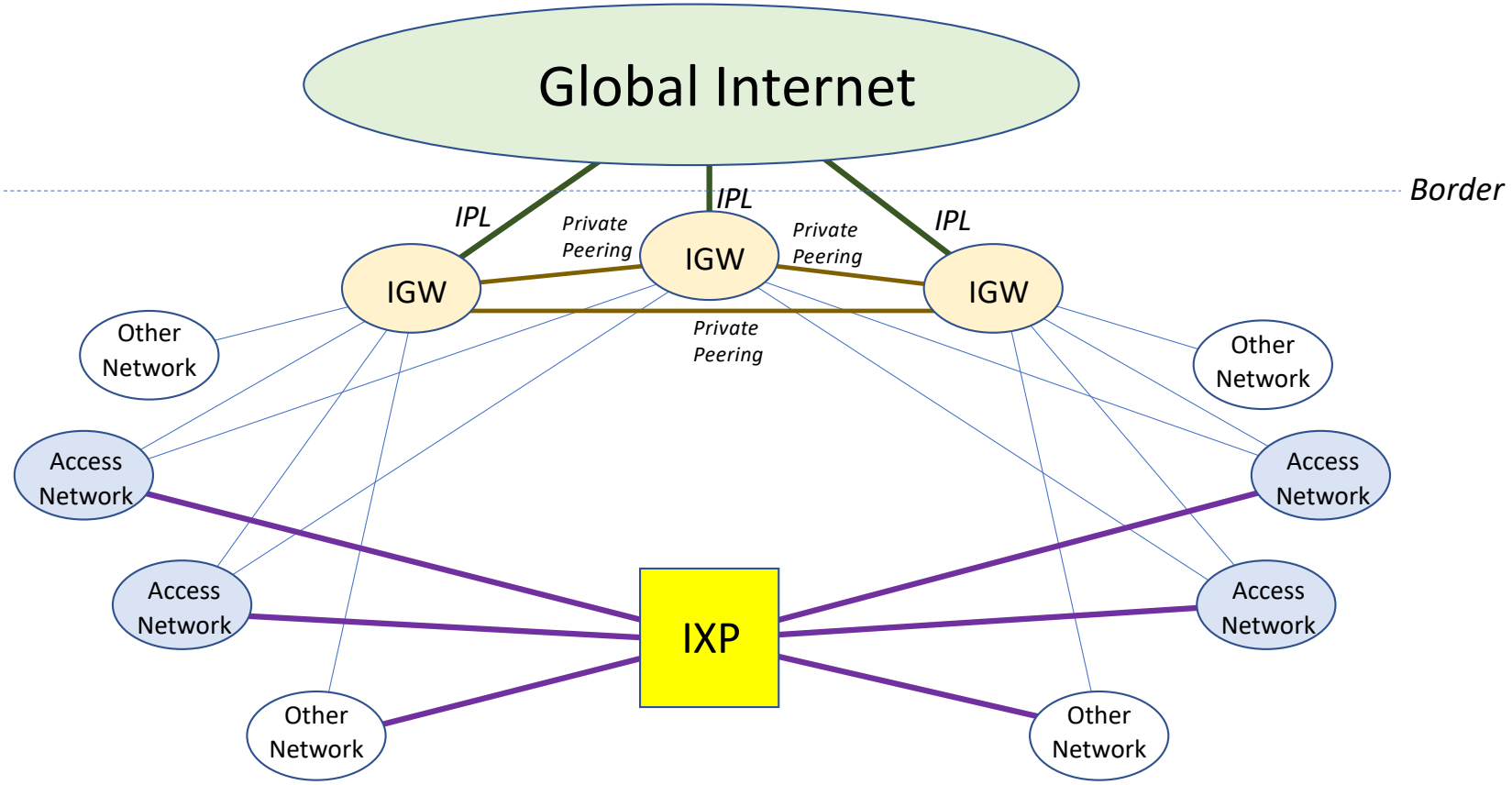
IXP Serving Licensed ISPs Only?

- Can be a starting point for non-profit local IXP if so desired
- But sooner or later, the IXP should be opened up for all kinds of networks including CDN networks, authoritative DNS servers, large enterprise networks (e.g. government departments, universities, banks and hospitals) and overseas networks so as to further enhance the importance and the status of the IXP
  - Can drive down the pricing of Internet connectivity further
- With proper environment, the local IXP may become an regional IXP

Only a Few International Gateway Licensees?

- 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
- 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

Possible Scenario for Improving Local Peering



IXP across Multiple Cities / Economies?

- Affect neutrality?
  - Generally considered as competing with participants which provide services across the same set of locations
- Bad for non-profit IXPs targeting all kinds of networks or providers?
  - Those that see competition may not join and then it may affect the goal of “keeping local traffic local”
  - Commercial IXPs can take this business risk especially if this may help their other business

Shared CDN Caches Offered by IXP?

- More and more local IXPs in developing economies want to provide shared caches for their participants to increase their value
  - Cost recovery and cost sharing / accounting are major issues to them though
  - Can do charging by usage which should be fairer
- CDN providers are starting to accept such model
  - They still mostly look at cache efficiency (cache hit ratio) and traffic volume for justifications
- It should be good to consider it
  - At smaller developing economies
  - Especially for gathering small ISPs together to meet the requirements of CDN providers

IXP Development Work of APNIC

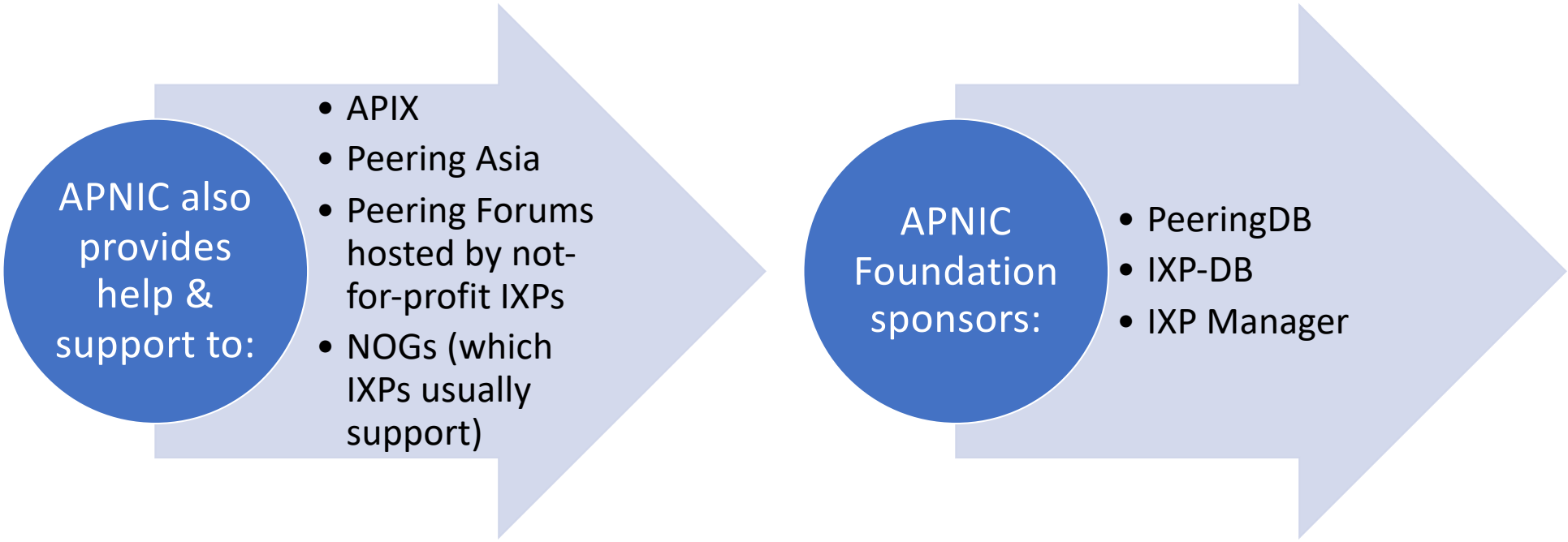
IXP Development Work of APNIC

- APNIC strongly believes IXPs help Internet development
  - After all, IXPs serve and benefit APNIC members
  - In fact, IXPs need IP addresses and ASNs and so are APNIC members themselves
- Do more on helping those developing economies
  - Especially those which do not have any IXP yet
  - Or those which their only IXP is not functioning well
- Training and Technical Assistance work primarily
  - Not just for IXP operators but also for IXP participants
  - Also help talk to major stakeholders to convince them of the benefits of having a local IXP while maintaining neutrality
  - May need help of Community Trainers and Consultants from time to time
- Having been supporting IXP development in Fiji, PNG, Vanuatu, Mongolia, Bhutan, Myanmar, Pakistan, India and others

IXP Development Package of APNIC

- Training & Technical Assistance primarily
  - Technical & non-technical
- Other possible support items (on case-by-case basis according to individual needs):
  - Ethernet switch
  - Route Server
  - ROV & IPv6 deployment support
  - IXP Manager
  - Root Server anycast instance
  - RIPE Atlas Anchor
- Collaboration with APIX & ISOC along with APNIC Foundation
  - MOU signed: https://www.apnic.net/wp-content/uploads/2022/03/Memorandum-of-Understanding-IXP-Support-APIX-ISOC-APNIC-APNIC-Foundation-FINAL-SIGNED.pdf
- APIX & MANRS Memberships are recommended to all IXPs

Other Help & Support by APNIC



Final Remarks

Final Remarks

- IXPs will continue to play a key role for easier interconnections among networks
  - Especially for developing economies
  - But IXP is NOT a magic wand to solve all the issues
    - Collaborative spirit is
- Need to find a suitable model for long-term sustainability
- Relative neutrality is important
  - So have to maintain it as much as possible
- After all, “Keeping Local Traffic Local” is the most important thing for the whole country/economy

Thank You!