



Peering in Hong Kong

Che-Hoo CHENG

CUHK/HKIX

www.hkix.net



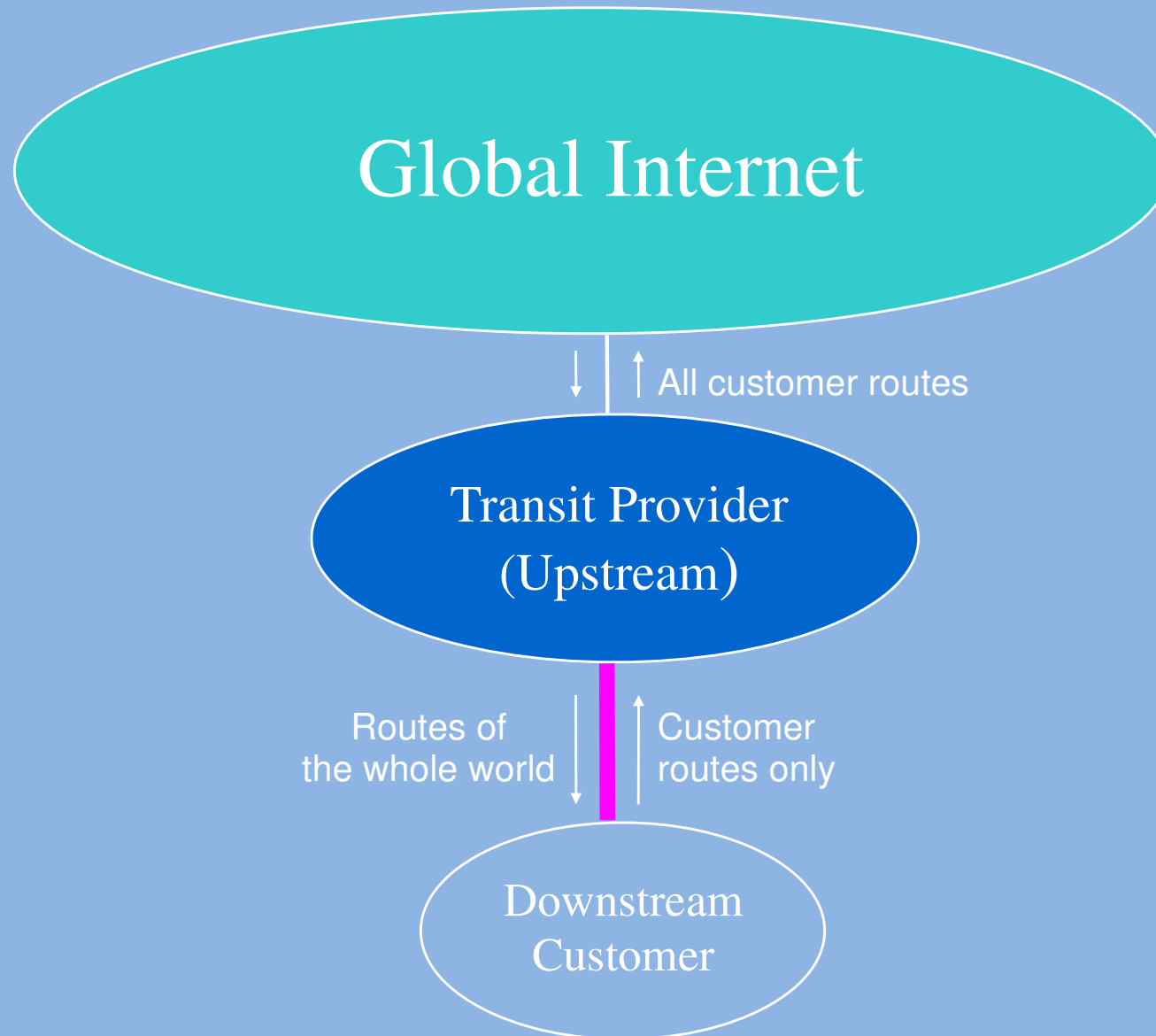
How Internet Operates – in simple terms



- Internet is composed of networks of ISPs and users
- User networks connect to ISPs
- Small ISPs connect to large ISPs
- ISPs (large or small) are interconnected or peered at Internet exchanges points (IXPs) or privately
- A few very large ISPs act as major transit providers (the so-called tier-1 backbones) for the whole world which rely on only peering arrangements to have full connectivity
- Other ISPs must be transit customers of those backbone ISPs directly or indirectly in order to gain full connectivity
- *A network on Internet is called Autonomous System (AS) which is represented by AS Number (ASN)*
 - *ASN is unique around the world*

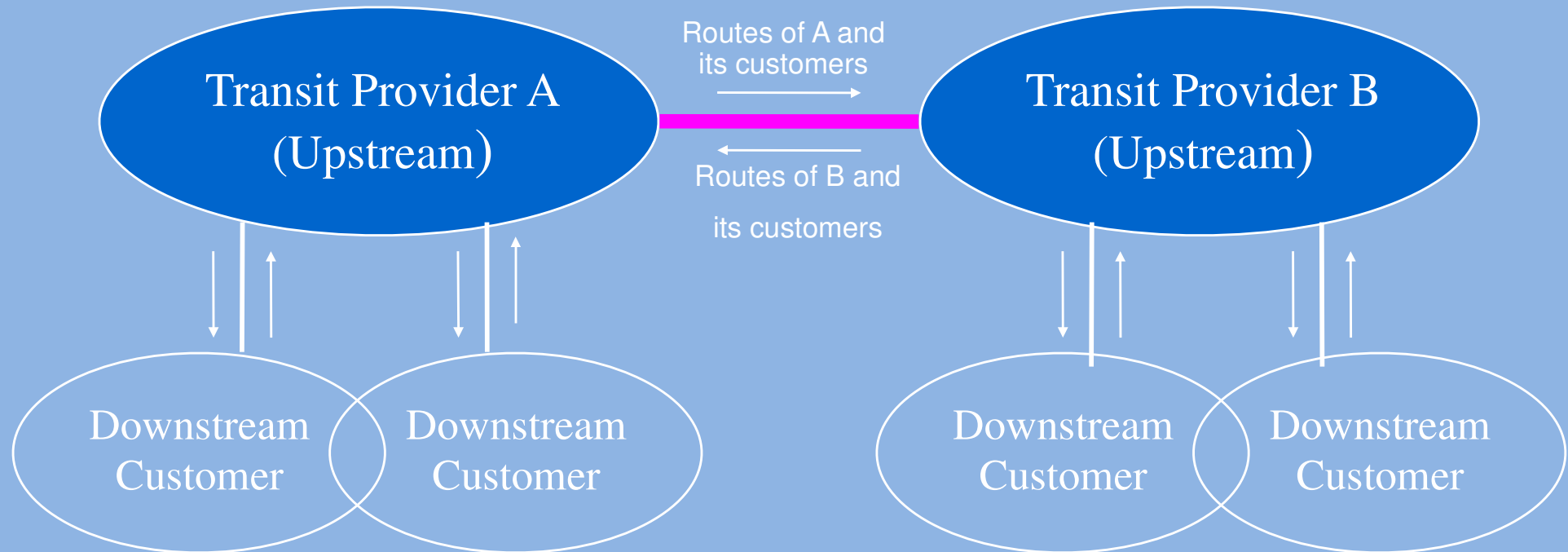


Transit Relationship





Ordinary Peering Model





Peering in General

- Interconnection among ISPs / data centres / content providers / cloud services providers which have different ASNs using BGP protocol
- For mutual benefits
 - For higher performance and lower cost
- Local Peering
 - Local-to-local traffic do NOT need to route through overseas
 - Important to Internet development
- Between 2 AS'es
 - BLPA (Bi-Lateral Peering Agreement)
- Among > 2 AS'es
 - MLPA (Multi-Lateral Peering Agreement)
- May have formal agreement / contract



Private Peering

- A form of BLPAs having dedicated point-to-point connection between 2 AS'es
- Using cross-connect or local loop or IPL (or dedicated VLAN over IXP) to interconnect
 - Cost is usually shared between 2 peers
- May have multiple connections between 2 AS'es for resiliency
- Not quite cost effective
 - Spare bandwidth cannot be used
- Not very scalable
 - ${}_n C_2$ physical connections for n AS'es to peer fully



IXP –



Internet eXchange Point

- For multiple AS'es to do peering easily over a well-managed dedicated network infrastructure
- Layer 2 infrastructure in most of the cases:
 - Switched Ethernet
 - Co-location of Routers?
 - Not necessarily the case now because of advancement of Ethernet technologies
 - Only routers are allowed to connect to the switching fabric directly
 - Support BLPA
 - Most IXPs facilitate MLPA
- Having multiple sites within one metro area (extended layer 2) is common
- Often considered as Critical Internet Infrastructure locally, regionally or globally



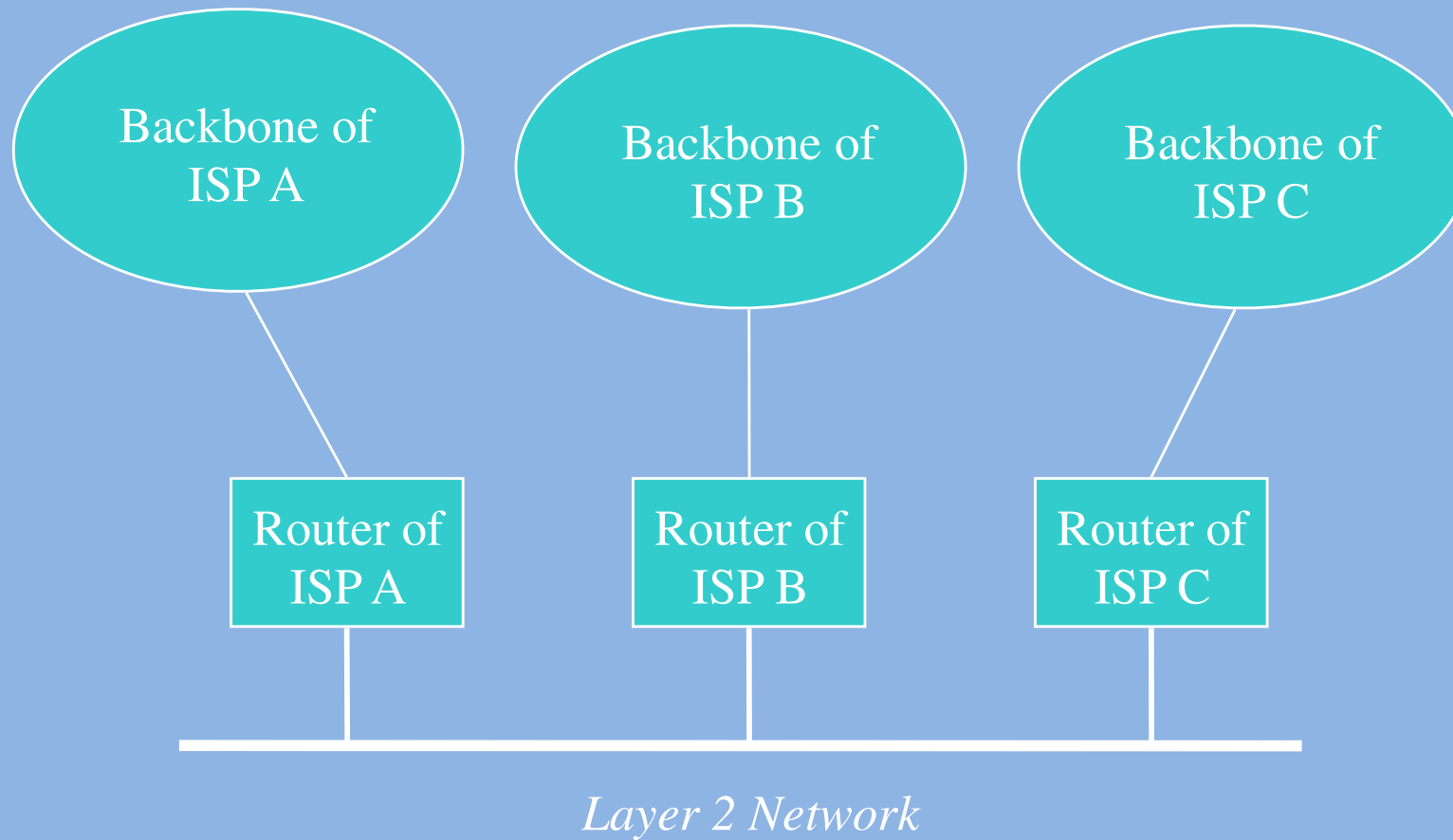
IXPs in HK

- HKIX
- Equinix HK
- AMS-IX HK
- iAdvantage IX

- Different IXPs have difference focuses
 - They may or may not serve you



Layer 2 IXP





BLPA over Layer 2 IXP

- Can set up BLPA with multiple peers using BGP over the same layer 2 connection
- Convenient and cost effective
- GE connection probably is the minimum for IXP connection
 - 10GE connection is getting more and more popular
 - Link Aggregation (LACP) can be used for incremental upgrade
 - Best to have 2, 4 or 8 links for more balanced traffic volume
 - 40GE/100GE support is coming
- The only disadvantage properly is you do not know whether your peers have congestion problem at their IXP connections
- And of course, you need to pay the IXP port charge



MLPA –

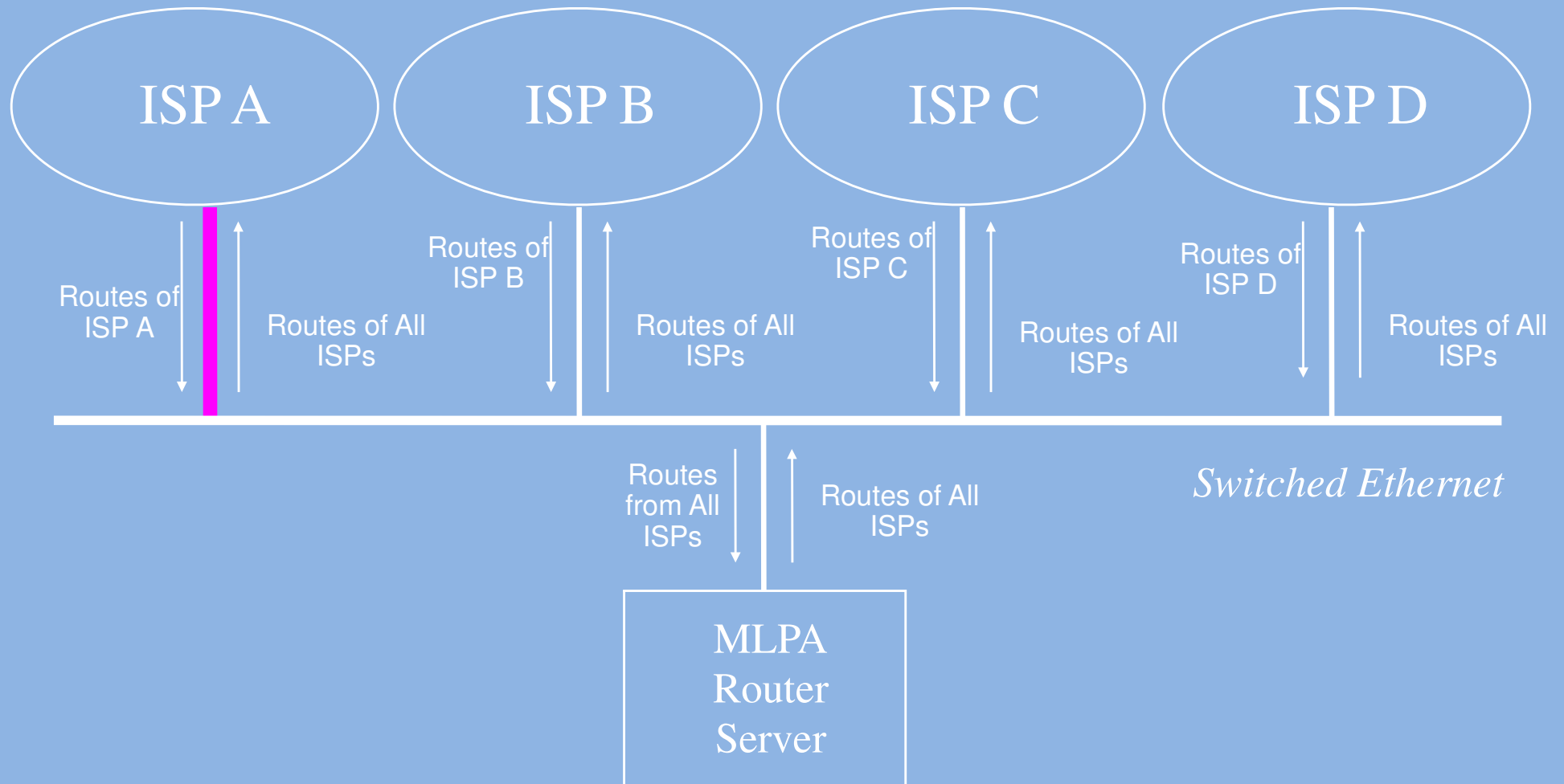


Multi-Lateral Peering Agreement

- Convenient for connecting to multiple ISPs
 - Just one BGP session
 - Facilitated by MLPA route servers
- Can be provided over layer 2 IXP
 - BLPA is supported at the same time



MLPA at Layer 2 IXP





Points to Note for MLPA

- You have less control of your routing under MLPA
 - Because MLPA route servers select the best routes for you
- With BLPA, 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
- Do NOT blindly prefer all routes learn from MLPA route servers using higher LocalPref
 - **Doing BLPA more in addition to MLPA over your IXP connection is highly recommended**



Peeringdb.com

- 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
- Also use it to find your potential BLPA peers
- Most content providers are willing to do bilateral peering
- Do set up BLPA with root / TLD DNS servers on IXPs to enjoy faster DNS queries



HK is an Internet hub

- A lot of overseas AS'es from all over the world have presence in HK
- They may be willing to set up direct peering with anyone for mutual benefit
 - After all, they pay for the circuits to come to HK so they want higher ROI
- A lot of intra-Asia traffic is being exchanged in HK now
- HK is indeed a telecom / Internet hub



ASN & IP Addresses for Peering

- ASN
 - For BGP connections, must be unique
 - Get one from APNIC for each of your networks which has different routing policy / arrangement
 - If you get one from your upstream, you may have it transferred to you later
- IP Addresses
 - **Do NOT use those assigned to you by your upstream which are NOT supposed to be portable**
 - Get your own portable addresses from APNIC
 - IPv4
 - Get your last /22 block (probably + another /22 block later) with demonstrated need
 - If you still do NOT have enough, you may need to buy from others
 - IPv6:
 - Get /48 or /32 block for each of your own AS



Traffic Engineering

- Load balancing
- Hot potato routing
- Make better use of BGP community



Route / Prefix Aggregation

- Do route / prefix aggregation as much as possible
- Using longer prefix to do traffic engineering is NOT a good practice
 - Use BGP community instead
- You should NOT allow your downstream customers to announce to other networks the portable addresses that you assign to them



Route Leakage

- Do NOT distribute (leak) to your peers (directly via BLPA or indirectly via MLPA) the routes which do not belong to yourself and your transit customers
- It will waste your bandwidth serving those which do not pay you
- May also affect the overall performance and people may not appreciate you at all
- You should do this ONLY to your transit customers (as part of full routes provisioning)



Dual-Stack

- Should do IPv6 in addition to IPv4 for all peering connections
 - Encourage your peers to enable and support IPv6
- Almost all IXPs support dual-stack
 - If they do not, you should disconnect



Transparency of Ethernet Local Loops

- Metro Ethernet services provided by some local loop providers in HK may not be transparent enough
 - May observe some unrelated traffic
 - May cause problems to your IXP connection
 - May not support LACP
 - GE is usually worse
 - 10GE is usually better (but not 100%)
 - Trial or guarantee is a must before you pay for them
- Same applies to IEPL



Other Tips

- Make sure your routers support 4-byte ASNs
- Do monitor the growth of number of routes learnt from your peers (or MLPA route servers) and adjust your max prefix settings accordingly
- Do monitor the utilization of your links closely and do upgrade before they are full
- Set up your own route / route6 / as-set objects on IRRDB and keep them up-to-date
 - APNIC RRDB is free if you are an APNIC member
 - Do not let your upstream providers to do it for you
 - They will mess things up



Thank You!