# Internet Exchange Point Design

## ISP/IXP Workshops



Last updated 3rd May 2021

# Acknowledgements


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





- I'd like to acknowledge all members of the community who have helped improve this presentation


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

Philip Smith

# IXP Design

- Background
- Why set up an IXP?
- Layer 2 Exchange Point
- Design Considerations
- Route Collectors & Servers
- What can go wrong?

# A bit of history

Where did the IX concept come from?

# A Bit of History…

- NSFnet – one major backbone
  - US "National Science Foundation" funded
  - Connected academic & research institutions
  - Also connected "private company" networks, under acceptable use policy (AUP), at network access points
  - **AUP: No commercial activity**
- Four Network Access Points (NAPs)
  - Chicago (IL) – run by Ameritech
  - New York (NY) – run by Sprint
  - San Francisco (CA) – run by PacBell
  - Vienna (VA) – run by MFS

# More History…

- Private companies needed to interconnect their networks too
  - Requirement to send "commercial traffic"
  - Could not cross NSFnet due to AUP
- Resulted in the first "commercial Internet Exchanges" in the early 1990s:
  - CIX-West – west coast USA (Bay Area)
  - MAE-East – east coast USA (Virginia)
- Routing Arbiter project helped with coordination of routing exchange between providers
  - Traffic from ISP A needs to get to ISP B
  - Now superseded by today's Internet Routing Registries (IRR)
  - The RADB is the remnant of the Routing Arbiter project

# More History still…

- End of the NSFnet in 1995:
  - Meant move towards commercial Internet
  - Private companies selling their bandwidth
- The NAPs established late in NSFnet life were some of the original "exchange points"
  - NAP operators were providing commercial Internet access as well
  - Sprint, PacBell and Ameritech NAPs were replaced by neutral/commercial IXPs
  - The MFS hosted MAE-East replaced the Vienna NAP
  - ANS (operator of the late NSFnet) forced to join IXes
- A global Distributed GIX proposed in mid 1990s
  - But never happened (planned to be CIX-West, MAE-East, SE-GIX and a Paris IX)

# Even more History

- SE-GIX formed in Stockholm in 1992
  - Three major ISPs interconnected
  - Latency reduction, performance gains
  - Local traffic stays local
  - (Proposed to be part of the D-GIX)
- LINX formed in London in 1994
  - Five UK operators interconnected
  - Latency reduction, performance gains
  - Local traffic stays local
  - (Proposed to be part the D-GIX when Paris fell through)
- HKIX formed in Hong Kong in 1995
  - Vibrant Internet community, many small operators
  - Latency, performance, and local traffic benefits

# Internet Exchange Point

- What:
  - **An open & neutral location where network operators freely interconnect their networks to exchange traffic**
- What is the physical IX:
  - An ethernet switch in a neutral location
- How does it work:
  - IX Host provides the switch and rack space
  - Network Operators interconnect via the IX fabric
    - Either bring their own router and install at the IX
    - Or run fibre/wireless link (point to point ethernet) from their location to IX
- Very simple concept – any place where providers meet to exchange traffic

# Internet Exchange Point

- Layer 2 exchange point
  - Ethernet (400Gbps/100Gbps/10Gbps/1Gbps)
  - Older technologies used in the past included ATM, Frame Relay, SRP, FDDI and SMDS
- Layer 3 exchange point
  - *Has had historical status since mid-90s*
  - Router based
    - Best known example was CIX-West
    - Router quickly overwhelmed by the sophisticated requirements of the rapidly growing Internet

# Why an Internet Exchange Point?

Saving money, improving QoS,
Generating a local Internet economy

# Internet Exchange Point
# Why peer?

- Consider a region with one ISP
  - They provide internet connectivity to their customers
  - They have one or two international connections
- Internet grows, another ISP sets up in competition
  - They provide internet connectivity to their customers
  - They have one or two international connections
- How does traffic from customer of one ISP get to customer of the other ISP?
  - Via the international connections

# Internet Exchange Point
# Why peer?

- Yes, International Connections…
  - Major content may be tens if not hundreds of milliseconds away
  - If geostationary satellite, RTT is around 550ms per hop
    - So local traffic between two operators would take over 1s round trip
    - Huge disincentive for a local Internet economy
- International bandwidth
  - Costs significantly more than domestic bandwidth
  - Is congested with local traffic
  - Local traffic on international links wastes money for both operators, harms overall performance for all users

# Internet Exchange Point
# Why peer?

- Solution:
  - Two competing ISPs peer with each other
- Result:
  - Both save money
  - Local traffic stays local
  - Better network performance, better QoS,…
  - More international bandwidth for expensive international traffic
  - Everyone is happy

# Internet Exchange Point
# Why peer?

- A third ISP enters the equation
  - Becomes a significant player in the region
  - Local and international traffic goes over their international connections

- All three ISPs agree to peer with each other to:
  - Save money for all three
  - Keep local traffic local
  - Improve network performance
  - Improve service quality for end users
  - Improve value proposition for local content hosting

# Internet Exchange Point
# Why peer?

- Private peering means that the three ISPs have to buy circuits between each other
  - Works for three ISPs, but adding a fourth or a fifth means this does not scale
- Solution:
  - Internet Exchange Point

# Internet Exchange Point

- Every participant has to deploy just one link
  - From their premises to the IXP
- Rather than N-1 links to connect to the N-1 other ISPs
  - 5 ISPs will have to share the cost of 4 links = 2 whole links → already twice the cost of the IXP connection

# Internet Exchange Point

- Solution
  - Every ISP participates in the IXP
  - Cost is minimal – one local link covers all domestic traffic
  - International links are used for just international traffic – and backing up domestic links in case the IXP suffers any outage
- Result:
  - Local traffic stays local
  - QoS considerations for local traffic is not an issue
  - RTTs between members are typically sub 1ms
  - Customers enjoy the Internet experience
  - Local Internet economy grows rapidly

# Who can join an IXP?

- Requirements are very simple: any organisation which operates their own autonomous network, and has:
  - Their own address space
  - Their own AS number
  - Their own transit arrangements
- This often includes:
  - Commercial ISPs
  - Academic & Research networks
  - Internet infrastructure operators (eg Root/ccTLDs)
  - Content Providers & Content Distribution Services
  - Cloud and Hosting Providers
  - Broadcasters and media
  - Government Information networks

# When an IXP is not beneficial

- Legislation: When there is one legislated monopoly transit provider
  - With all other network operators are legislated to be customers of this monopoly provider
- Geography: When the local economy is so small that it cannot sustain more than one network operator
  - Very small nations (maybe less than 10000 population?)
  - Sparsely populated / remote areas

# When an IXP is not permitted

- This is still the situation in several countries around the world
- Usually it is a Government operated "national telco"
  - ISP licence mandates connecting to "national telco" for Internet services
- Implications:
  - Expensive domestic connectivity
  - Expensive international connectivity
  - Restricted and poor service offerings
  - No domestic Internet economy
  - Everyone loses, especially the "national telco"

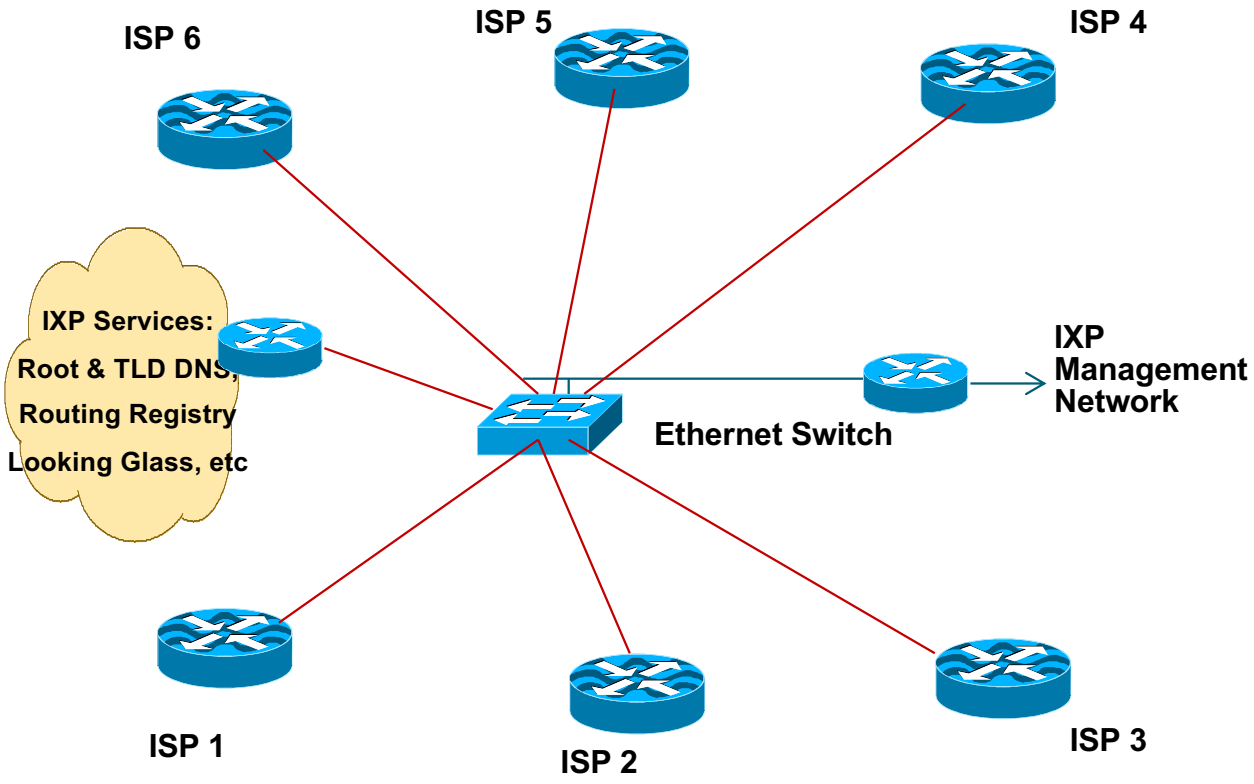
# Layer 2 Exchange

The global industry standard IXP

# IXP Design

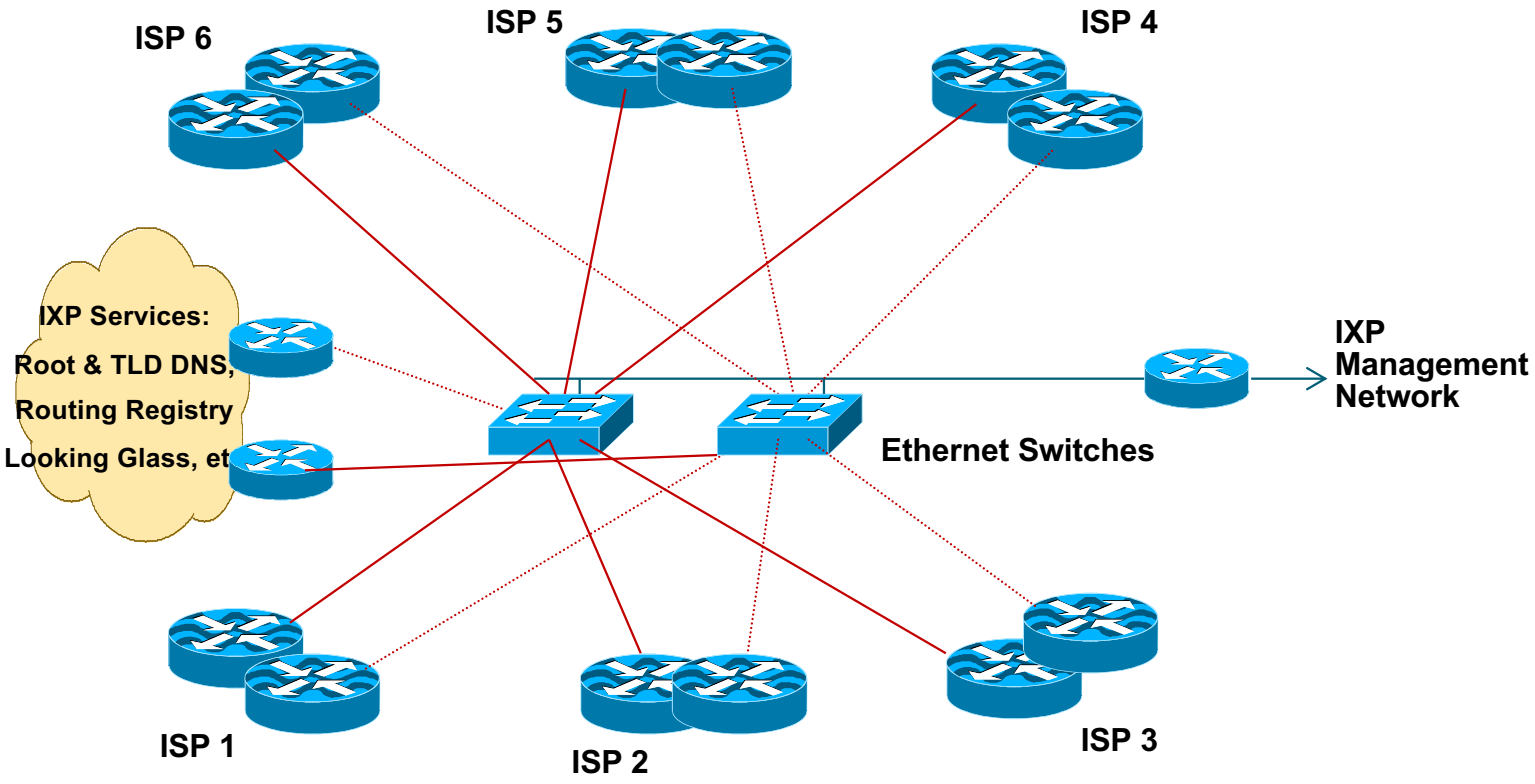
- Very simple concept:
  - Ethernet switch is the interconnection media
    - IXP is one LAN
  - Each member brings a router, connects it to the ethernet switch provided at the IXP
    - Alternatively, the member runs fibre/wireless from their datacentre to the IXP location, and connects to the IXP switch (point-to-point ethernet)
  - Each ISP peers with other participants at the IXP using BGP
- Scaling this simple concept is the challenge for the larger IXPs

# Layer 2 Exchange




Single site internet exchange point

# Layer 2 Exchange




Dual site internet exchange point – not interconnected

# Layer 2 Exchange

- Two switches for redundancy
- ISPs use dual routers for redundancy or loadsharing
- Hosts services for the “common good”
  - Internet portals and search engines
  - DNS Root & TLDs, NTP servers
  - Routing Registry and Looking Glass

# Layer 2 Exchange: Location

- Neutral
  - Anyone can install fibre or other connectivity media to access the IXP
    - Without cost or regulations imposed by location
- Secure
  - Best practice physical and logical security, like any other network data centre
- Accessible
  - Easy/convenient for all participants to access
- Safe
  - Low risk from natural disaster (cyclone, earthquake, tsunami, wildfire)
- Expandable
  - IXPs cause local Internet economy growth, and therefore increasing space requirements within the facility

# Layer 2 Exchange: Location

- Good:
  - Carrier Neutral Data Centre (ideal!)
  - Technology Park
  - University
  - Large Enterprise
  - Check: All need 24x7 power/access/security
- Bad:
  - ISP Data Centre (not neutral / ISP perceived advantage)
  - Government Data Centre (highly sensitive / security risk)
  - Cable Landing Station (national security / physical access / not neutral / tsunami risk)

# Layer 2 Exchange: Operations

- Operation:
  - Requires neutral IXP management
  - "Consortium"
    - Representing all participants
    - "Management Board" etc
- Funding:
  - All costs agreed and covered equally by IXP participants
  - Hosting location often contributes – the IXP brings them more business
- Availability:
  - 24x7 cover provided by hosting location
    - Managed by the consortium

# Layer 2 Exchange: Technical

- Configuration
  - Recommendation: Public address space for IXP LAN
    - IPv4 (/24) and IPv6 (/64)
  - ISPs require AS, basic IXP does not
- Network Security Considerations
  - LAN switch needs to be securely configured
  - IXP Management & Services router(s) require well protected access
  - IXP services must be behind router(s) with strong filters

# IXP Standards

- Industry Standards documented by Euro-IX, the European IXP Association
  - Contributed to by the Euro-IX members
  - https://www.euro-ix.net/en/forixps/set-ixp/
- IXP BCP
  - General overview of the infrastructure, operations, policies and management of the IXP
  - https://www.euro-ix.net/en/forixps/set-ixp/ixp-bcops/
- IXP Website BCP
  - https://www.euro-ix.net/en/forixps/set-ixp/ixp-bcops/ixp-website/

# “Layer 3 Exchange”

Why this is not an IXP

# “Layer 3 IXP”

- Layer 3 IXP today is marketing concept used by Transit ISPs
  - Some incumbent telcos call their domestic or international transit businesses “Exchanges”
- Real Internet Exchange Points are only Layer 2
  - L2 is the accepted International standard

# “Layer 3 IXP” – what breaks

- One extra AS hop between peers
  - Makes path via IXP suboptimal/less preferred
  - Path between peers usually remains with upstream transit provider
    - Unless both peers actively implement BGP policies to prefer the L3 IXP
- Members cannot peer with whom they please
  - Mandatory multilateral peering
  - Third party (L3 IXP operator) required to configure peering sessions and peering policy

# “Layer 3 IXP” – what breaks

- More complicated troubleshooting
  - Troubleshooting peering problems has to involve IXP operator too
- No policy control
  - BGP attributes shared between members get dropped by IXP router
  - (Examples are BGP communities, MEDs)

# “Layer 3 IXP” – what breaks

- CDNs won’t join
  - They have requirements to peer directly with IXP members
- Redundancy problems
  - L3 IXPs with dual sites appear as two separate transit providers between peers
  - Traffic engineering?
- L3 IXP Operator requires strong BGP skills

# IXP Design Considerations

# Exchange Point Design

- The IXP Core is an Ethernet switch
  - It must be a managed switch
  - It must have reasonable security features
  - https://www.euro-ix.net/ixp-wishlist has more details
- Has superseded all other types of network devices for an IXP
  - From the cheapest and smallest managed 12 or 24 port 100M/1G switch
  - To the largest switches now handling high densities of 10GE, 100GE and 400GE interfaces

# Exchange Point Design

- Each member participating at the IXP brings a router to the IXP location
  - Note that with increased availability of fibre access, members connect directly to the IXP without provisioning a dedicated router at the IXP location
- Router needs:
  - One Ethernet port to connect to IXP switch
  - One WAN port to connect to the WAN media leading back to the ISP backbone
  - To be able to run BGP

# Exchange Point Design

- IXP switch located in one equipment rack dedicated to IXP
  - Also includes other IXP operational equipment
- Routers from participants are located in neighbouring/adjacent rack(s)
- Copper (UTP) connections used for 1Gbps connections
- Fibre used for 1Gbps, 10Gbps, 40Gbps, 100Gbps and 400Gbps connections
  - Fibre is more scalable, allowing IXP member to start with 1Gbps and upgrade to bigger bandwidths simply by swapping the SFP

# Peering

- Each participant needs to run BGP
  - They need their own AS number
  - **Public** ASN, **NOT** private ASN
- Each participant configures external BGP directly with the other participants in the IXP
  - Peering with all participants

    *or*
  - Peering with a subset of participants

# Peering (more)

- Mandatory Multi-Lateral Peering (MMLP)
  - Each participant is forced to peer with every other participant as part of their IXP membership
  - **Has no history of success** — the practice is strongly discouraged
- Multi-Lateral Peering (MLP)
  - Each participant peers with the other participants (usually via a Route Server)
- Bi-Lateral Peering (BLP)
  - Participants set up peering with each other according to their own requirements and business relationships
  - This is the most common situation at IXPs today

# Types of Operator Peering Policies

- 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 a "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 the "Route Server" but will only peer bilaterally
- Restrictive Peering
  - Where an ISP decides who its peering partners are, and is generally not approachable to creating peering opportunities

# Operators Peering Activities

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

# Routing Advice

- Member border routers at the IXP must NOT be configured with a default route or carry the full Internet routing table
  - Carrying default or full table means that this router and the member network is open to abuse by non-peering IXP members
  - Correct configuration is only to carry routes offered to IXP peers on the IXP peering router
- Note: Some network operators offer transit across IX fabrics
  - They do so at their own risk – see above
  - (It's not the IXP's concern what the member offers by way of services as long as what they do does not impact the IXP's operation)

# Routing Advice (more)

- Member border routers at the IXP should not be configured to carry the IXP LAN network within their iBGP
  - Use next-hop-self BGP concept (refer to BGP Introduction and BGP Attributes presentations)
  - Keeping IXP LAN address block in IGP ensures that traceroutes do not break
- Member should not generate their aggregates on IXP peering router
  - If connection from backbone to IXP router goes down, normal BGP failover will then be successful

# Address Space

- Some IXPs use private addresses for the IX LAN
  - Their claim is that public address space means IXP network could be leaked to Internet which may be undesirable
  - But most network operators filter RFC1918 address space, so this avoids the problem
- Most IXPs use public addresses for the IX LAN
  - Address space available from the RIRs via specific RIR policies
  - IXP terms of participation often forbid the IX LAN to be carried in the member backbone
- IXPs provide both IPv6 and IPv4 support on IX LANs
  - No need for separate LANs for IPv6 and IPv4

# Autonomous System Numbers

- IXPs by themselves do not require ASNs
  - Ethernet switch is L2 device, and does not run BGP
- Some IXPs have a Route Collector
  - This usually runs in a private ASN
- Some IXPs have a Route Server
  - This usually runs in a public ASN
- Some IXPs have “common good services”
  - These usually require Internet transit
  - Meaning the IXP requires a transit router
    - IXP arranges transit for services with a couple of providers
  - And this transit router requires a Public ASN and Public Address space
    - This ASN/address space is separate from what is used by the IXP fabric itself

# Hardware

- Ethernet switch needs to be managed
  - This means CLI access as well as SNMP (at least SNMPv2)
    - CLI allows direct troubleshooting of configuration and operational problems
  - Unmanaged switches mean an unmanageable IXP
- Insist that IXP participants connect a router (L3) port to the IXP switch
  - Avoid spanning tree and L2 security issues
  - Run port security or MAC filtering to protect the IX
- Insist that IXP participants bring their own router
  - Moves buffering problem off the IXP switch
  - (Fibre access to IX reduces this requirement)
  - Security of the member connection is responsibility of the member, not the IXP

# Charging

- IXPs needs to be run at minimal cost to its member participants
- Common examples:
  - Datacentre hosts IX for free
  - IX members pay a flat annual fee (cost recovery)
  - Differential pricing per port (line card basis)
- IXes do **NOT** charge for traffic crossing the switch fabric
  - They are a peering enabler, encouraging as much traffic as possible between members

# Charging:
# Datacentre hosts IX for free

- Datacentre covers all costs relating to the IX
  - They provide the switch and supporting infrastructure
  - They provide the operator cover
  - They benefit from the business the IX members and their customers bring to the DC
  - They benefit from the "prestige" of hosting the IX and its ancillary services
- The IX does not charge members for anything at all
  - Example: Seattle IX

# Charging:
# IX Members pay flat fee

- Each member pays a flat annual fee towards their IX membership
- How it works:
  - Cost of switch and ports
  - Cost of operator support
  - Datacentre cost: power, air-conditioning, etc
  - Cost of IX membership association
  - Contingency needed for new equipment and upgrades
- Total annual cost shared equally amongst members
  - The more members, potentially the lower the costs for each

# Charging: Differential pricing by port

- IXP Member pays according to the port speed they require (big IXP switches)
  - One linecard may handle 4 100GE ports
  - Or one linecard may handle 24 10GE ports
  - Or one linecard may handle 96 1GE ports
  - 96 port 1GE card is tenth price of 24 port 10GE card
  - Relative port cost is passed on to participants
  - Plus share in the cost of the switch
  - Plus all the costs mentioned in the flat-fee model
- IX members pay according to the cost of provisioning their port speed
  - Example: Netnod IXes in Sweden

# Notes about charging

- Smaller or new IXPs:
  - Free, or flat fee, for members
  - 1RU switch supporting 1G/10G on all ports
  - Members are responsible for providing suitable optics

- Larger or longer established IXPs:
  - Chassis based switches, linecards have different costs
  - Members pay contribution to cost of linecard (hence port charge), often including cost of optics too

# Services Offered

- Services offered should not compete with members (basic IXP)
  - e.g. web hosting by the IXP is a bad idea unless all members agree to it
- IXP operations need to make performance and throughput statistics available to members
  - Use tools such as LibreNMS to produce IX throughput graphs for member (or public) information

# Services to Offer

- Root server
  - Anycast instances of F, I and L root nameservers are present at many IXes
- ccTLD DNS
  - The country IXP could host the country's top level DNS
  - e.g. "SE." TLD is hosted at Netnod IXes in Sweden
  - Offer back up of other country ccTLD DNS
- gTLD DNS
  - .com & .net are provided by Verisign at many IXes

# Services to Offer

- Route Server
  - Helps scale IXes by providing easier BGP configuration & operation for participants with Open Peering policies
  - Technical detail covered later on
- Looking Glass
  - One way of making the Route Server routes available for global view (e.g. www.traceroute.org)
  - Public or members-only access
- RPKI Validator Cache
  - Implements RPKI Origin Validation on prefixes shared on the Route Server
  - Also make cache available for member use

# Services to Offer

- Content Redistribution/Caching
  - Various providers offering content distribution services
  - Broadcast media
- Network Time Protocol
  - Locate a stratum 1 time source (GPS receiver, atomic clock, etc) at IXP
- Routing Registry
  - Used to register the routing policy of the IXP membership (more later)

# Notes on IXP Services

- If IXP is offering services to members:
  - Services need transit access
  - Transit needs to be arranged with one or two IXP members (cost shared amongst all members)

- Consider carefully:
  - Should services be located at the IXP itself?
    - How to arrange and pay for the transit to those services?

    -or-
  - Should services be hosted by members and shared with the others?

# Introduction to Route Collectors

What routes are available at the IXP?

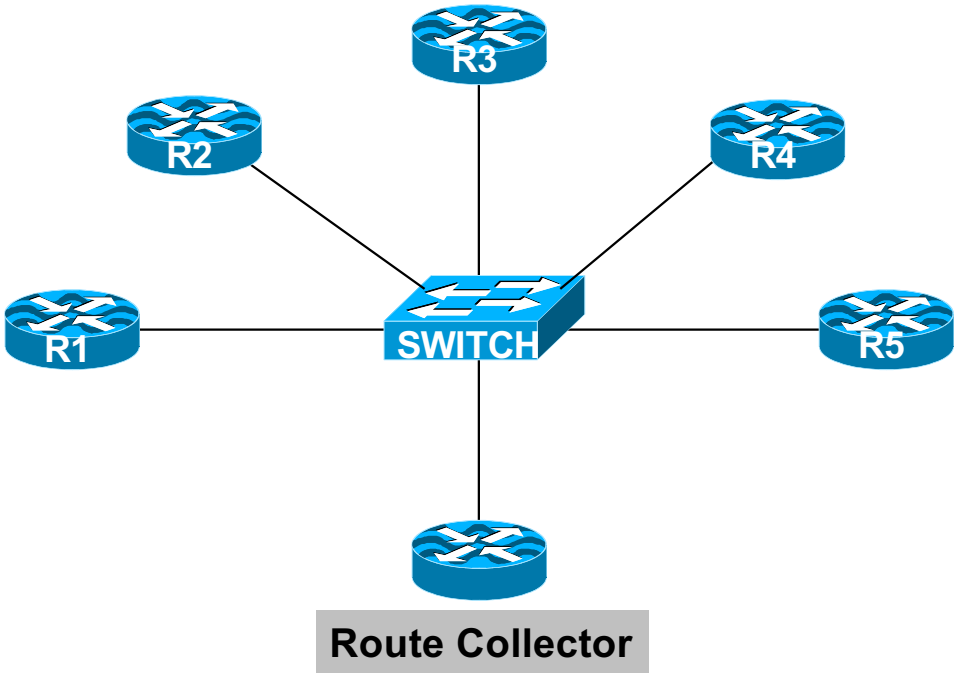
# What is a Route Collector?

- Usually a router or Unix system running BGP software
- Gathers routing information from member routers at the IXP
  - Peers with each member using BGP
- Does not forward packets
- Does not announce any prefixes to members

# Purpose of a Route Collector

- To provide a public view of the Routing Information available at the IXP
  - Useful for existing members to check functionality of BGP filters
  - Useful for prospective members to check value of joining the IXP
  - Useful for the Internet Operations community for troubleshooting purposes
    - E.g. **www.traceroute.org**

# Route Collector at an IXP



# Route Collector Requirements

- Router or Unix system running BGP
  - Minimal memory requirements – only holds IXP routes
  - Minimal packet forwarding requirements – doesn't forward any packets
- Peers eBGP with every IXP member
  - Accepts everything; Gives nothing
  - Uses a private ASN
  - Connects to IXP Transit LAN
- "Back end" connection
  - Second Ethernet globally routed
  - Connection to IXP Website for public access

# Route Collector Implementation

- Most IXPs now implement some form of Route Collector
  - Usually as a Route Server (see next section)
- Benefits already mentioned
- Great public relations tool
- Unsophisticated requirements
  - Just runs BGP

# Introduction to Route Servers

How to scale IXPs

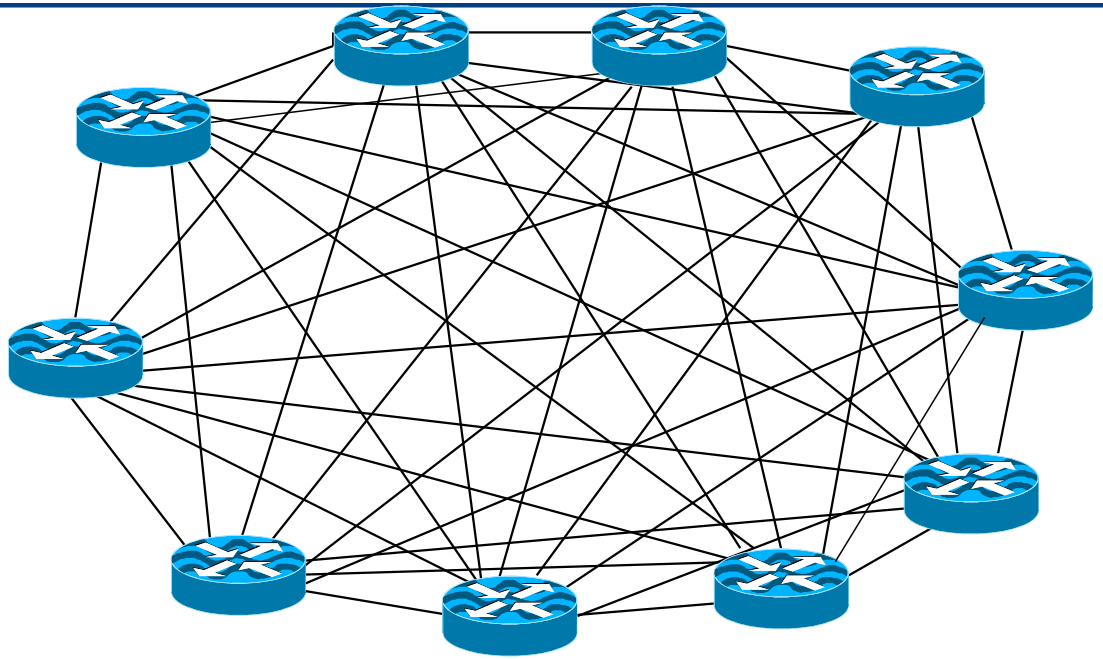
# What is a Route Server?

- Has all the features of a Route Collector
- But also:
  - Announces routes to participating IXP members according to their routing policy definitions
- Implemented using the same specification as for a Route Collector

# Features of a Route Server

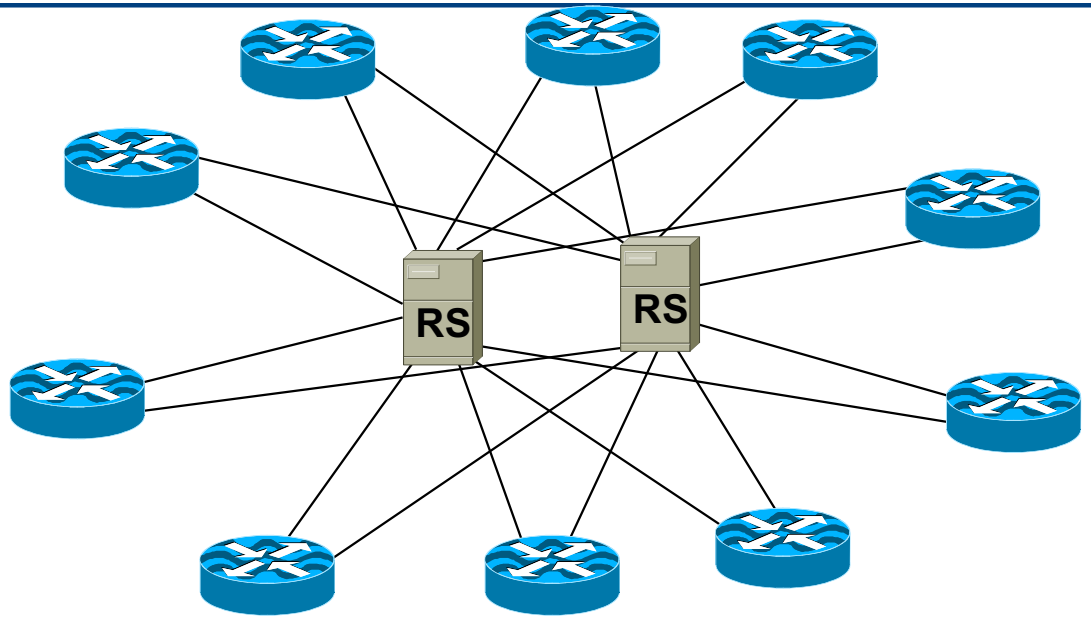
- Helps scale route distribution for IXPs
  - Forwarding of packets is unaffected
  - Makes use of BGP functionality known as “third party next-hop”
- Simplifies Routing Processes on ISP Routers
- Optional participation
  - Provided as service, is NOT mandatory
- If traditional router used, will result in insertion of RS Autonomous System Number in the AS Path
  - To be avoided
- Optionally could use Policy registered in the Internet Routing Registry

# Diagram of N-squared Peering Mesh



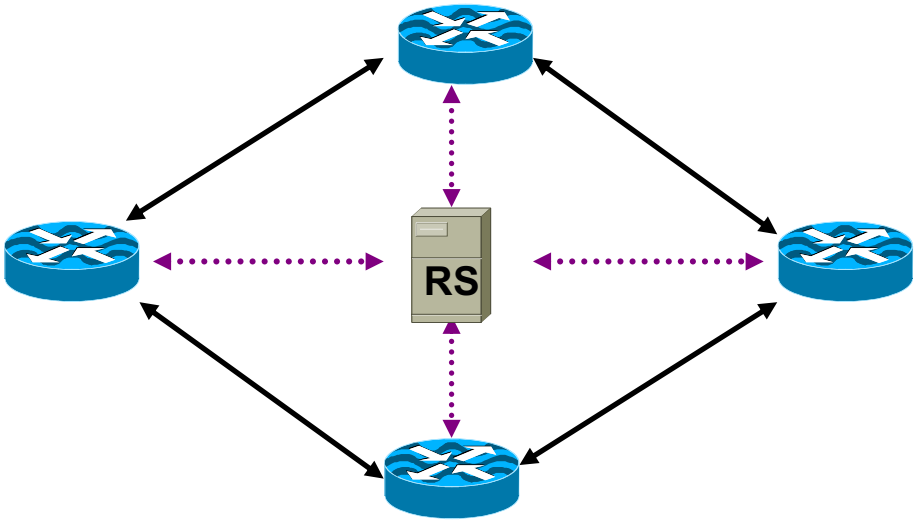
- For large IXPs (dozens of participants) maintaining a larger peering mesh becomes cumbersome and often too hard

# Peering Mesh with Route Servers




- ISP routers peer with the Route Servers
  - Only need to have two eBGP sessions rather than N

# Route Server based Exchange Point Routing Flow




**TRAFFIC FLOW**

**ROUTING INFORMATION FLOW**

# Using a Route Server: Advantages

- Advantageous for large IXPs
  - Helps scale eBGP mesh
  - Helps scale prefix distribution
- Separation of Routing and Forwarding
- Simplifies BGP Configuration Management on ISP routers
  - Don’t need to maintain a large number of eBGP peers
  - eBGP peering only with the Route Server

# Using a Route Server: Disadvantages

- ISPs can lose direct policy control
  - If RS is the only peer, ISPs have no control over who their prefixes are distributed to
    - (Okay if ISP has Open Peering Policy though)
- Completely dependent on 3rd party
  - Configuration, troubleshooting, reliability, etc…
- Possible insertion of RS ASN into routing path
  - (If using a router rather than a dedicated route-server BGP implementation)
  - Traffic engineering/multihoming needs more care

# Typical usage of a Route Server

- Route Servers are provided as an **OPTIONAL** service
  - Most IXPs now offer a Route Server as a service to members
- ISPs peer:
  - Directly with significant peers

    -and-
  - With Route Server for the rest
- ISPs with an Open Peering Policy usually prefer to peer with a Route Server

# Route Server implementations

- Linux/FreeBSD server:
  - BIRD – the standard & works best
    - http://bird.network.cz
  - GoBGP
    - https://osrg.github.io/gobgp/
  - Quagga (LINX fork)
    - https://github.com/bbonev/quagga.euro-ix/
  - FRR (origins in Quagga)
    - https://www.frrouting.org/
- Router:
  - Any router (but has RS AS in the AS-path)
  - Cisco IOS 15.2 and IOS XE 3.7 onwards has route-server-client configuration:

```
neighbor 172.16.1.1 route-server-client
```

# Things to think about…

- Would using a route server benefit you?
  - Helpful when BGP knowledge is limited (but is NOT an excuse not to learn BGP)
  - Avoids having to maintain a large number of eBGP peers

# What can go wrong…

The different ways IXP operators harm their IXP…

# What can go wrong? Concept

- Some Service Providers attempt to cash in on the reputation of IXPs
- Market their Internet transit services as "Internet Exchange Point"
  - "We are exchanging packets with other ISPs, so we are an Internet Exchange!"
  - So-called Layer-3 Exchanges – they really are Internet Transit Providers
  - Router(s) used rather than a Switch
  - Most famous example: SingTelIX

# What can go wrong? Financial

- Some IXPs price the IX out of the means of most providers
  - IXP is intended to encourage local peering
  - Acceptable charging model is minimally cost-recovery only
- Some IXPs charge for port traffic
  - IXPs are not a transit service, charging for traffic puts the IX in competition with members
  - (There is nothing wrong with charging different flat fees for 1Gbps, 10Gbps, 100Gbps etc ports as they all have different hardware costs on large chassis switches)

# What can go wrong? Competition

- Too many exchange points in one locale
  - Competing exchanges defeats the purpose
- Becomes expensive for ISPs to connect to all of them
  - So they don't, or won't, and local traffic suffers, defeating the viability of the IXPs
- An IXP:
  - is **NOT** a competition
  - is **NOT** a profit making business

# What can go wrong?
# Rules and Restrictions

- IXP tries to compete with their membership
  - Offering services that ISPs would/do offer their customers
  - **In reality, IXPs are operated by the members for the members**
- IXP is run as a closed privileged club e.g.:
  - Restrictive membership criteria
  - **In reality, a participant needs to have an ASN, their own independent address space, and their own transit arrangements**
- IXP located in a data centre with restricted physical/transmission access
  - **IXP must be a neutral interconnect in a neutral location**

# What can go wrong? Rules and Restrictions

- IXP charges for traffic
  - So do transit providers – **charging for traffic is a sure way of ending the viability of the IXP**
- IXPs providing access to end users rather than just Network Operators & Service Providers
  - **A participant at an IXP needs to have their own address space, their own ASN, and their own transit arrangements**
- IXPs interfering with member business decisions
  - **The most common error: Mandatory Multi-Lateral Peering**

# What can go wrong? Technical Design Errors

- Interconnected IXPs
  - IXP in one location believes it should connect directly to the IXP in another location
  - Who pays for the interconnect?
  - How is traffic metered?
  - Competes with the ISPs who already provide transit between the two locations (who then refuse to join IX, harming the viability of the IX)
  - Metro interconnections work ok

# What can go wrong?
# Technical Design Errors

- Members bridge the IXP LAN back to their offices
  - "We are poor, we can't afford a router"
  - Financial benefits of connecting to an IXP far outweigh the cost of a router
  - In reality it allows the members to connect any devices to the IXP LAN — with disastrous consequences for the security, integrity and reliability of the IXP

# What can go wrong? Routing Design Errors

- Route Server mandated
  - Mandatory peering has no history of success
  - ISPs have no incentive to learn BGP
  - Therefore have no incentive to understand peering relationships, peering policies, &c
  - Entirely dependent on operator of RS for troubleshooting, configuration, reliability
    - RS can't be run by committee!
- Route Server is designed to assist with scaling peering at IXPs

# What can go wrong?
# Routing Design Errors (cont)

- iBGP Route Reflector used to distribute prefixes between IXP participants
- Claimed advantages:
  - Participants don't need to know about or run BGP
  - Allows an IXP to be started very quickly
  - IXP operator has full control over member activities
  - ISP participants routers sit inside IXP's ASN
- All are disadvantages!
  - Participants never learn BGP
  - Participants have no policy control, IXP policies could impact the participants networks
  - IXP is an ethernet switch, not an Internet operator
  - IXP operator is single point of failure
  - Migration to true IXP with RS is very difficult

# What can go wrong: Summary

- Not a transit business, just an L2 switch
- If charging, fair cost recovery only
- Not a competitive service
- No oppressive rules & restrictions
- No Mandatory Peering
- No bureaucratic management
- No interconnection with other IXPs
- No bridging of IX LAN back to members
- No Route Reflector, use a Route Server to scale

# More Information

# Exchange Point Policies & Politics

- AUPs
  - Acceptable Use Policy
  - Minimal rules for connection
- Fees?
  - Some IXPs charge no fee
  - Other IXPs charge cost recovery
  - A few IXPs are commercial
- Nobody is obliged to peer
  - Agreements left to individual members, not mandated by IXP

# Exchange Point etiquette

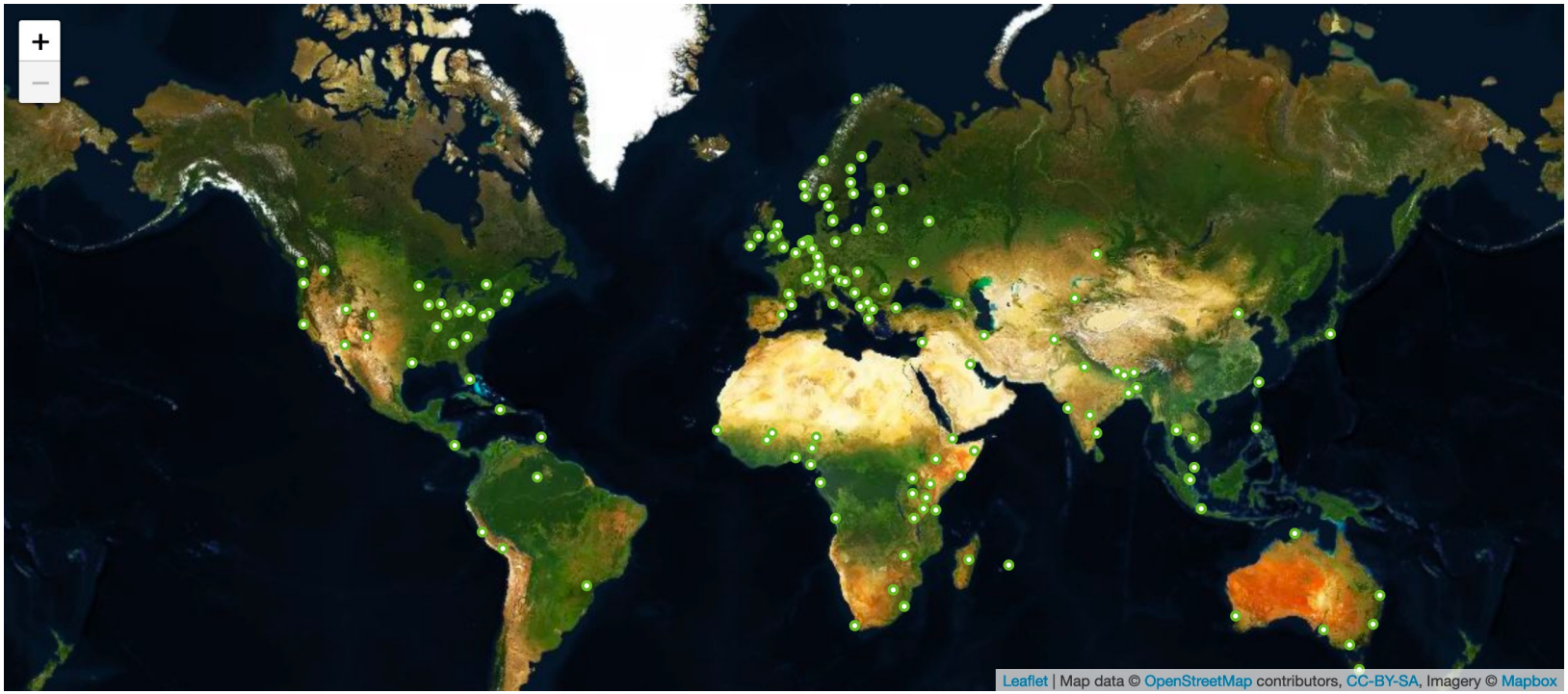
- Don’t point default route at another IXP participant
- Be aware of third-party next-hop
- Only announce your aggregate routes
  - Read RIPE-399 and RIPE-532 first

    www.ripe.net/ripe/docs/ripe-399

    www.ripe.net/ripe/docs/ripe-532
- Filter! Filter! Filter!

# Exchange Point Examples

- Hundreds of Internet Exchange Points globally
- Those using IXP Manager are on this map:
  - https://www.ixpmanager.org/community/world-map



# Features of IXPs (1)

- Redundancy & Reliability
  - Multiple switches, UPS/Generator
- Support
  - NOC to provide 24x7 support for problems at the exchange
- DNS, Route Collector/Server, Content Caches & NTP servers
  - ccTLD & root servers
  - Content caches
  - Content redistribution systems
  - Route Collector – Routing Table view

# Features of IXPs (2)

- Location
  - Neutral, secure & accessible co-location facilities
- Address space
  - Public address for Peering LAN
  - Public address for IXP Services LAN
- AS Number
  - Private ASN needed for Route Collector (if deployed)
  - Public ASN needed for Route Server
  - Public ASN needed for IXP Services
- Route servers
- Statistics
  - Traffic data – for membership

# IXP Creation

- No economy or circumstance is unique or different
  - The first excuse for not creating an IXP is "we don't need one"
  - The second excuse for not creating an IXP is "oh, it is different here"
- Every locality has its differences
  - But every locality wants to
    - Keep local traffic local
    - Improve network performance and QoS
    - Improve local Internet economy
  - The available technology is the same for every network operator everywhere
  - There is no excuse for not improving the local Internet

# Eco System Development

- Create IXP association
  - Formed by members who have a port on the IXP
- IXP members meet regularly
  - IXP Board meetings
  - IXP Operational strategy and direction
- IXP Technical community could also meet too
  - Network operators meeting, involving network and systems operations technicians & engineers
  - Aligned with IXP Association/member meetings
  - Could lead to creation of a Network Operators Group
- IXP could facilitate the creation of a NOG
  - The same technicians & engineers are involved in both!

# Local Internet Exchange Point

- Defined as a public peering point serving the local Internet industry
- Local means where it becomes cheaper to interconnect with other members at a common location than it is to pay transit to another network operator to reach the same consumer base
  - Local can mean different things in different regions!

# Regional Internet Exchange Point

- These are also "local" Internet Exchange Points
- But also attract regional ISPs and ISPs from outside the locality
  - Regional ISPs peer with each other
  - And show up at several of these Regional IXPs
- Local ISPs peer with ISPs from outside the locality
  - They don't compete in each other's markets
  - Local ISPs don't have to pay transit costs
  - ISPs from outside the locality don't have to pay transit costs
  - Quite often ISPs of disparate sizes and influences will happily peer – to defray transit costs

# Industry Associations

- IX-F
  - The Internet Exchange Federation
    - http://www.ix-f.net/
  - The federation of Internet Exchange Associations
- Euro-IX
  - The European Internet Exchange Association
  - Members from Europe, associate members from around the world
  - Website has all the information needed to start an IXP
    - https://www.euro-ix.net/starting-an-ixp
  - IXP Best Practice documentation:
    - https://www.euro-ix.net/euro-ix-bcp

# Industry Associations

- APIX
  - Asia Pacific Internet Exchange association
  - Meets twice a year, during APRICOT and APNIC conferences
    - http://apix.asia
- Af-IX
  - The African IXP Association
  - Meets along with the African Peering Forum
    - http://www.af-ix.net/
- LAC-IX
  - The Latin American & Caribbean IX Association
    - http://www.lac-ix.org/

# More info about interconnects

- Telegeography
  - http://www.telegeography.com/telecom-resources/internet-exchange-map/
  - A collection of ISP interconnect points
  - Beware!! Not all of the Telegeography listings are IXPs!
- Packet Clearing House
  - IXP Directory: https://www.pch.net/ixp/dir
- Internet Society
  - IXP Toolkit: http://www.ixptoolkit.org/

# Summary

- IXP is a Layer 2 infrastructure
- At least three players required (two is okay too)
  - Meeting in an open and neutral location
- Minimal rules
- Minimal bureaucracy
- Cost recovery
- Encourage participation by all autonomous networks
- Develop the local Internet eco-system

# Internet Exchange Point Design

ISP/IXP Workshops