

Internet Exchange Point Design



Philip Smith

<philip@apnic.net>

APNIC 42

Colombo, Sri Lanka

28th September – 5th October 2016

Presentation Slides

- Will be available on
 - <http://bgp4all.com/dokuwiki/conferences/>
 - And on the APNIC42 website
- Feel free to ask questions any time



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, with acceptable use policy
 - **AUP: No commercial activity**
 - Three Network Access Points (NAPs): Chicago, New York, San Francisco
- Private companies needed to interconnect their networks
 - Requirement to send “commercial traffic”
 - Could not cross NSFnet
 - Resulted in the early “commercial Internet Exchanges”

More History...

- ❑ Early Internet Exchanges created in early 90s
 - CIX-West – west coast USA
 - MAE-East – east coast USA
 - D-GIX – Stockholm
- ❑ End of the NSFnet in 1995:
 - Meant move towards commercial Internet
 - Private companies selling their bandwidth
 - ANS (operator of the late NSFnet) had to join IXes
- ❑ Routing Arbiter project helped with coordination of routing exchange between providers
 - Traffic from ISP A needs to get to ISP B
 - The RADB is the remnant of that project

More History still...

- ❑ The NAPs established late in NSFnet life were some of the original “exchange points”
 - NAP operators supported commercial activities as well
 - (Sprint: NY, PacBell: SF, Ameritech: Chicago, MFS: Vienna/VA)
- ❑ The NAPs replaced by IXPs:
 - NAPs didn’t succeed (operated by ISPs), replaced by more neutral IXPs
 - E.g. Virginia NAP replaced by MAE-East (by MFS)
- ❑ Mid 90s saw rapid Internet growth, with major providers connecting...

Even more History

- ❑ D-GIX formed in Stockholm in 1992
 - Three major ISPs interconnected
 - Latency reduction, performance gains
 - Local traffic stays local
- ❑ LINX formed in London in 1994
 - Five UK operators interconnected
 - Latency reduction, performance gains
 - Local traffic stays local
- ❑ HKIX formed in Hong Kong in 1995
 - Vibrant Internet community, many small operators
 - Latency, performance, and local traffic benefits
- ❑ Also AMS-IX in Amsterdam in 1994
 - Same reasons as others

Internet Exchange Point

- What:
 - **A 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 Operator provides the switch and rack space
 - Network Operators bring routers, and interconnect them via the IX fabric
- Very simple concept – any place where providers meet to exchange traffic

Internet Exchange Point

- Layer 2 exchange point
 - Ethernet (100Gbps/10Gbps/1Gbps/100Mbps)
 - Older technologies used in the past included ATM, Frame Relay, SRP, FDDI and SMDS
- Layer 3 exchange point
 - *Has historical status now*
 - Router based
 - Best known example was CIX-West
 - Router very quickly overwhelmed by the rapid growth of the 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...
 - If satellite, RTT is around 550ms per hop
 - So local traffic takes over 1s round trip
- ❑ International bandwidth
 - Costs significantly more than domestic bandwidth
 - Congested with local traffic
 - Wastes money, harms performance

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
- They agree to peer with the two other ISPs
 - To save money
 - To keep local traffic local
 - To improve network performance, QoS,...

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 buy just one whole circuit
 - From their premises to the IXP
- ❑ Rather than $N-1$ half circuits to connect to the $N-1$ other ISPs
 - 5 ISPs have to buy 4 half circuits = 2 whole circuits → already twice the cost of the IXP connection

Internet Exchange Point

□ Solution

- Every ISP participates in the IXP
- Cost is minimal – one local circuit covers all domestic traffic
- International circuits are used for just international traffic – and backing up domestic links in case the IXP fails

□ Result:

- Local traffic stays local
- QoS considerations for local traffic is not an issue
- RTTs are typically sub 10ms
- 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
 - 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 *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

Layer 2 Exchange

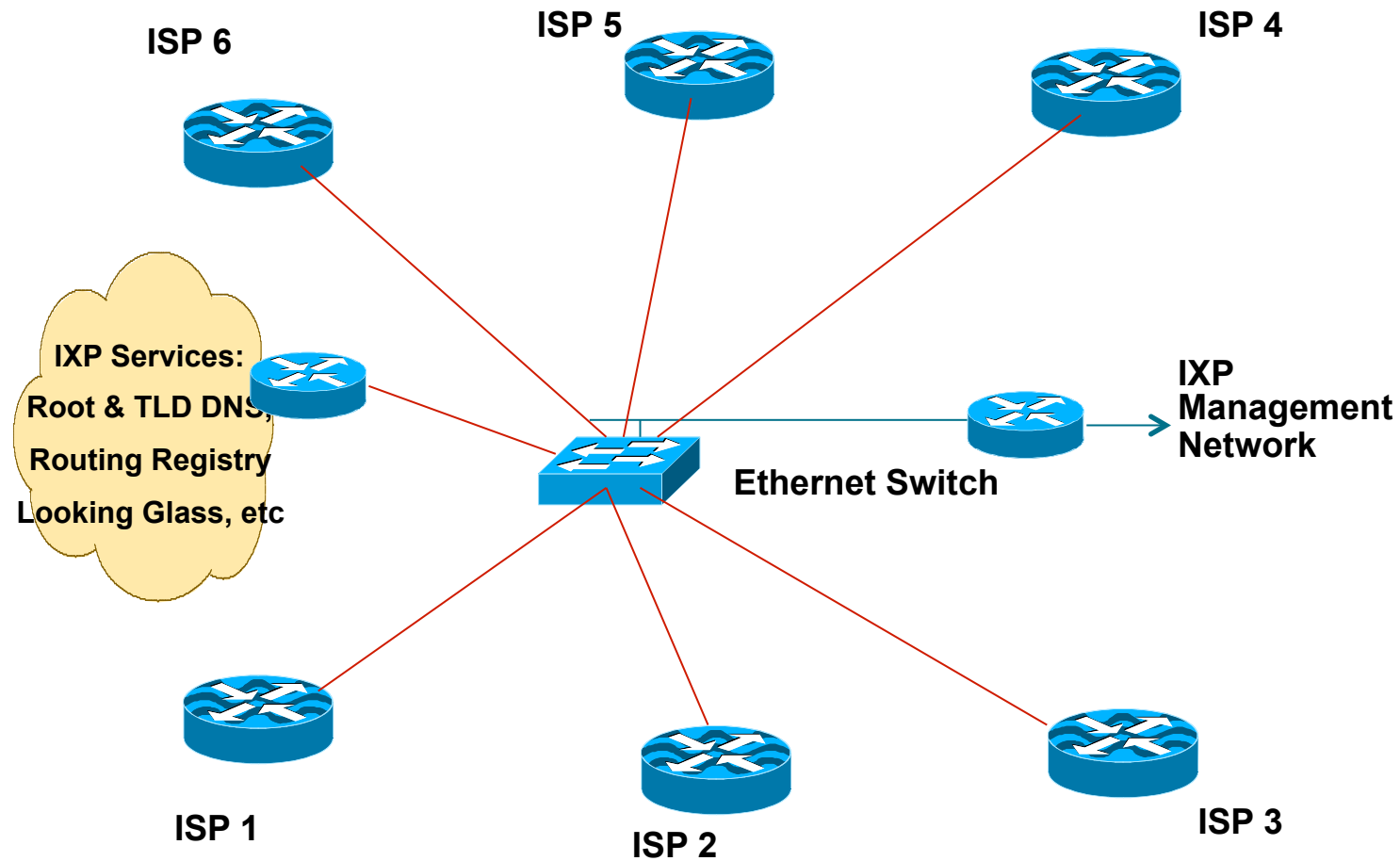


The traditional IXP

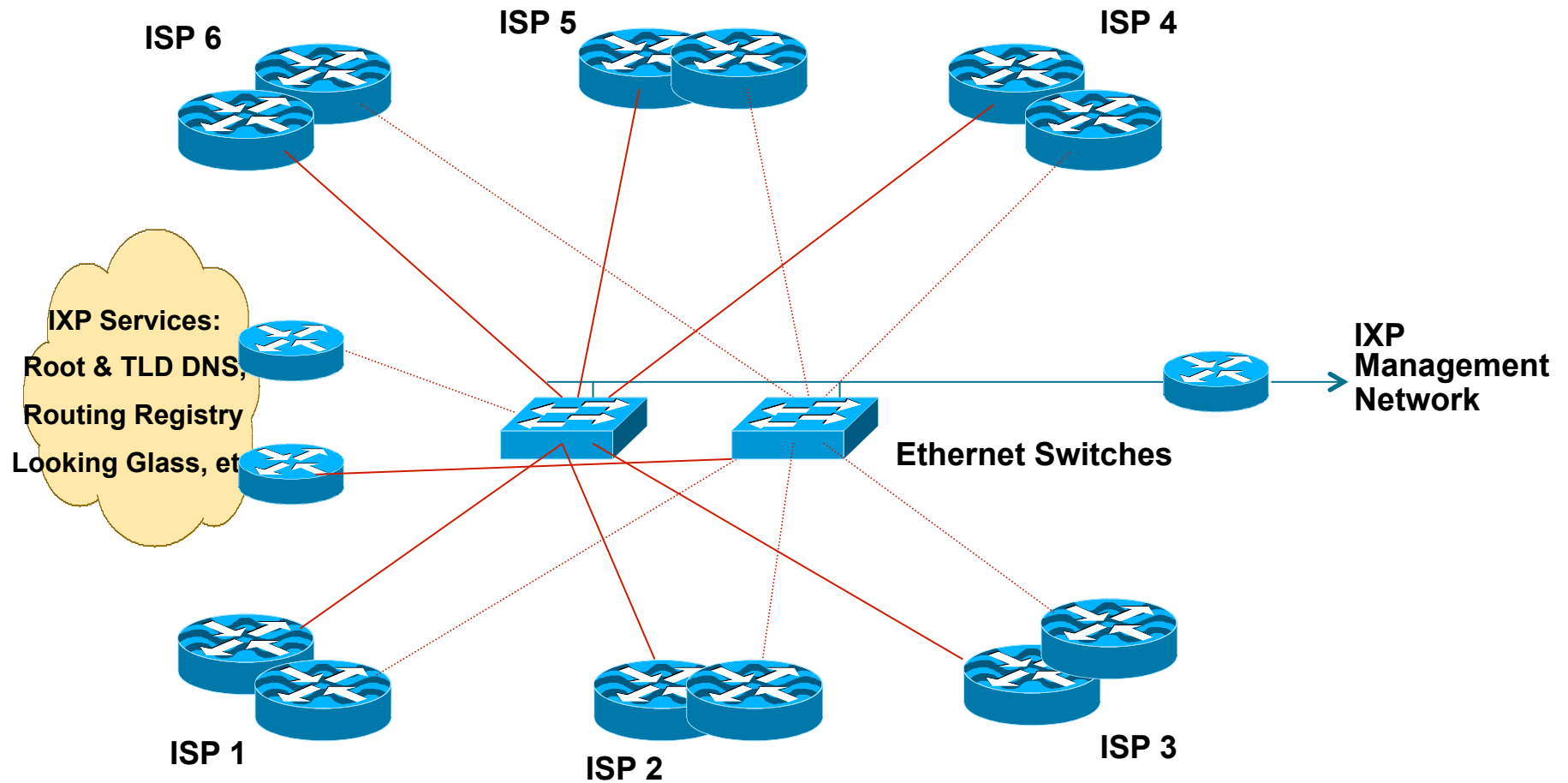
IXP Design

- ❑ Very simple concept:
 - Ethernet switch is the interconnection media
 - ❑ IXP is one LAN
 - Each ISP brings a router, connects it to the ethernet switch provided at the IXP
 - 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



Layer 2 Exchange



Layer 2 Exchange

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

Layer 2 Exchange

- ❑ Neutral location
 - Anyone can install fibre or other connectivity media to access the IXP
 - ❑ Without cost or regulations imposed by location
- ❑ Secure location
 - Thorough security, like any other network data centre
- ❑ Accessible location
 - Easy/convenient for all participants to access
- ❑ Expandable location
 - IXPs result in Internet growth, and increasing space requirements

Layer 2 Exchange

❑ 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

□ 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

Layer 2 Exchange

- ❑ Industry Standards documented by Euro-IX, the European IXP Association
 - Contributed to by the Euro-IX members
 - <https://www.euro-ix.net/starting-an-ixp>
- ❑ IXP BCP
 - General overview of the infrastructure, operations, policies and management of the IXP
 - <https://www.euro-ix.net/euro-ix-bcp>
- ❑ IXP Website BCP
 - <https://www.euro-ix.net/website-bcp>

“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 gets 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 10/100 switch
 - To the largest switches now handling high densities of 10GE, 40GE and 100GE interfaces

Exchange Point Design

- ❑ Each ISP participating in the IXP brings a router to the IXP location
 - Note that with increased availability of fibre access, ISPs 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 participant ISPs located in neighbouring/adjacent rack(s)
- ❑ Copper (UTP) connections made for 10Mbps, 100Mbps or 1Gbps connections
- ❑ Fibre used for 1Gbps, 10Gbps, 40Gbps or 100Gbps connections

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
 - 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

❑ Closed 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
 - <http://peeringdb.com>
- ❑ All operators of ASNs should register in the peeringdb
 - All operators who are considering peering or are peering must be in the peeringdb to enhance their peering opportunities
- ❑ Participation in peering fora is encouraged too
 - Global Peering Forum (GPF)
 - Regional Peering Fora (European, Middle Eastern, Asian, Caribbean, Latin American)
 - Many countries now have their own Peering Fora

Routing

- ❑ ISP 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 ISP 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 ISPs offer transit across IX fabrics
 - They do so at their own risk – see above

Routing (more)

- ❑ ISP border routers at the IXP should not be configured to carry the IXP LAN network within the IGP or iBGP
 - Use next-hop-self BGP concept
- ❑ Don't generate ISP prefix 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
 - Public address space means IXP network could be leaked to Internet which may be undesirable
 - Because most ISPs filter RFC1918 address space, 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 ISP member backbone
- ❑ IXPs provide both IPv6 and IPv4 support on IX LANs

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

Hardware

- ❑ Ethernet switch needs to be managed
 - Including CLI access rather than only SNMP
 - 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 ISP connection is responsibility of the ISP, not the IXP

Charging

- ❑ IXPs need to be run at minimal cost to its member participants
- ❑ Common examples:
 - Datacentre hosts IX for free
 - IX operates cost recovery
 - Different pricing for different ports
- ❑ 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
 - 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

Services Offered

- ❑ Services offered should not compete with member ISPs (basic IXP)
 - e.g. web hosting at an IXP is a bad idea unless all members agree to it
- ❑ IXP operations should 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

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)

Introduction to Route Collectors



What routes are available at the
IXP?

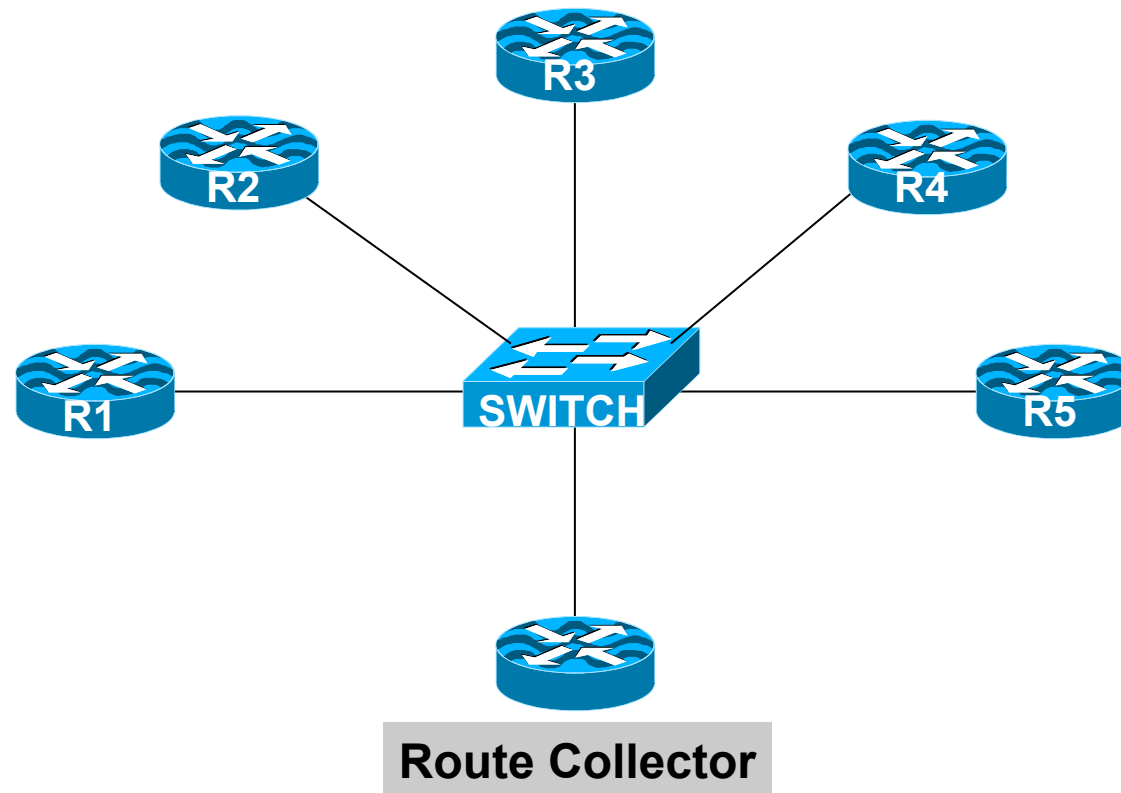
What is a Route Collector?

- ❑ Usually a router or Unix system running BGP software
- ❑ Gathers routing information from service provider routers at an IXP
 - Peers with each ISP using BGP
- ❑ Does **not** forward packets
- ❑ Does **not** announce any prefixes to ISPs

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

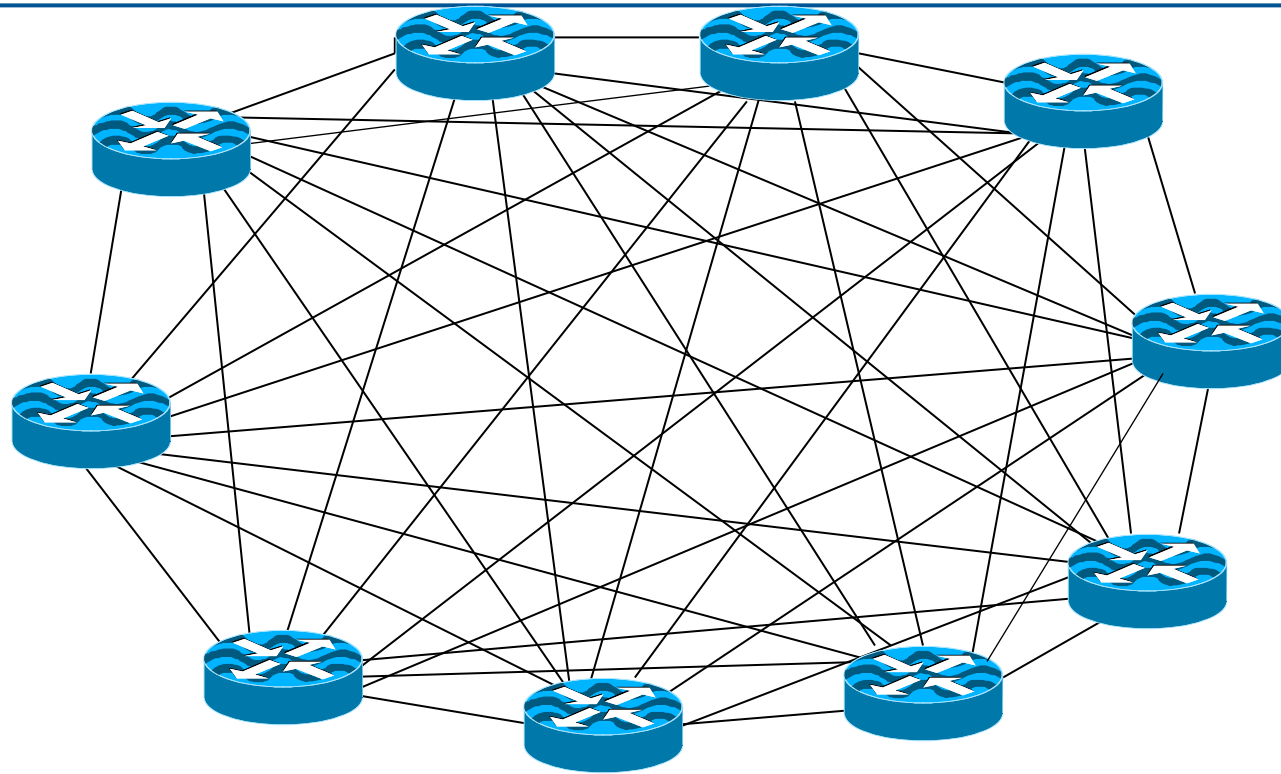
What is a Route Server?

- ❑ 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

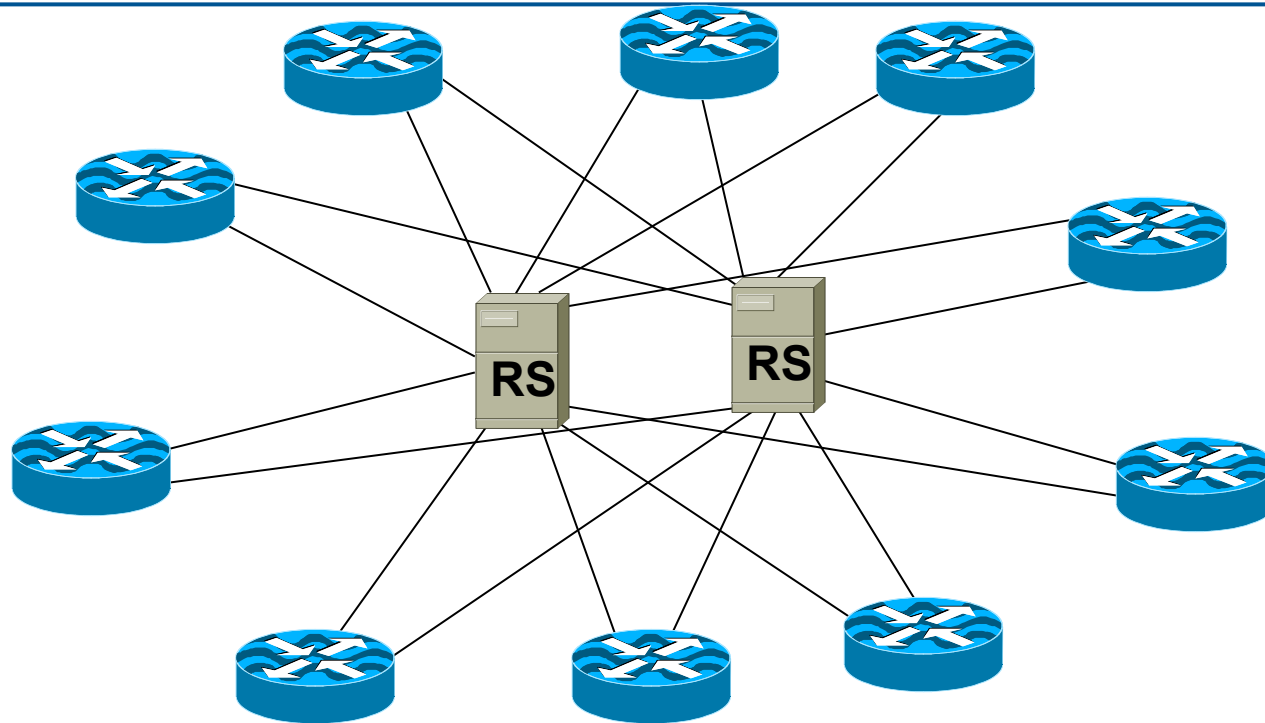
- ❑ Helps scale route distribution for IXPs
 - Forwarding of packets is unaffected
 - Makes use of BGP functionality known as “third party next-hop”
- ❑ Simplifies Routing Processes on ISP Routers
- ❑ Optional participation
 - Provided as service, is **NOT** mandatory
- ❑ If traditional router used,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

Diagram of N-squared Peering Mesh



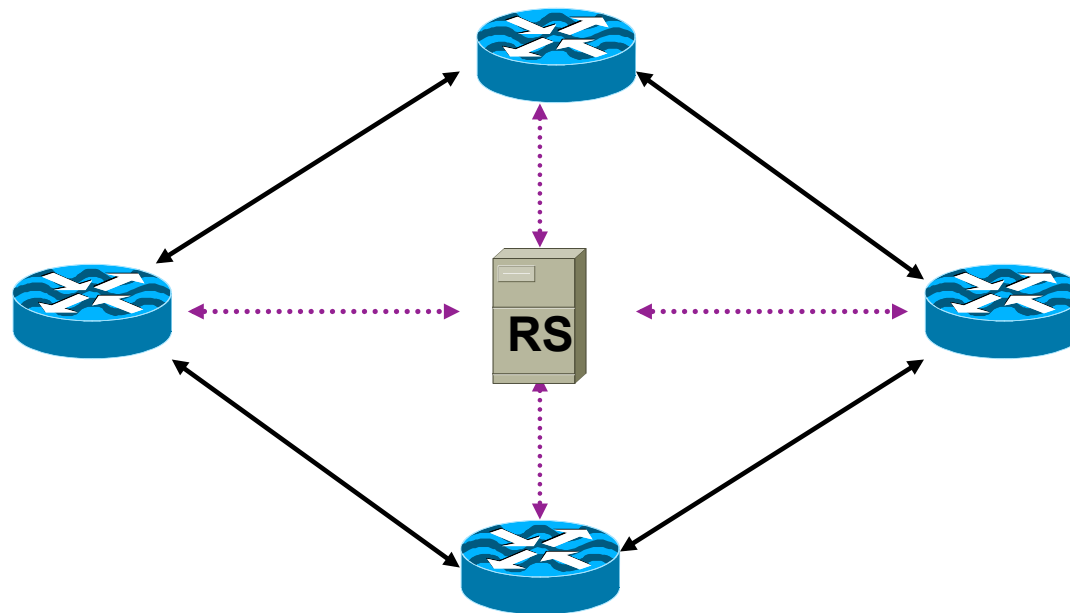
- ❑ For large IXPs (dozens for participants) maintaining a larger peering mesh becomes cumbersome and often too hard

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 may be provided as an **OPTIONAL** service
 - Most IXPs now offer a Router 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>
 - Quagga
 - <http://www.nongnu.org/quagga/>
- 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 Point!”
 - 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 100Mbps, 1Gbps, 10Gbps etc ports as they all have different hardware costs on the switch.)

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 and their own independent address space**
- ❑ 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

- ❑ ISPs 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 ISP 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 ISP 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 ISPs, 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

- ❑ LINX in London, UK
- ❑ TorIX in Toronto, Canada
- ❑ AMS-IX in Amsterdam, Netherlands
- ❑ SIX in Seattle, Washington, US
- ❑ PA-IX in Palo Alto, California, US
- ❑ JPNAP in Tokyo, Japan
- ❑ DE-CIX in Frankfurt, Germany
- ❑ HK-IX in Hong Kong
- ...
- ❑ All use Ethernet Switches

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/Server
 - Public ASN needed for IXP Services
- ❑ Route servers (for larger IXPs)
- ❑ 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!

Industry Associations

□ Euro-IX

- 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>

□ APIX

- Asia Pacific Internet Exchange association
- Members from Asia Pacific region
- Meets twice a year, during APRICOT and APNIC conferences
- <http://apix.asia>

More info about interconnects

□ Telegeography

- <http://www.telegeography.com/telecom-resources/internet-exchange-map/>
- A collection of ISP interconnect points
- **Beware!! Not all Telegeography listings are IXPs!**

□ Internet Society

- <http://www.internetsociety.org/what-we-do/issues/internet-exchange-points-ixps>
- <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

