



## Internet Exchange Point Design

ISP/IXP Workshops

## IXP Design

- Why set up an IXP?
- Layer 2 Exchange Point
- Layer 3 Exchange Point
- Design Considerations

## Internet Exchange Points

- **Layer 2 exchange point**  
Ethernet (1000/100Mbps)  
Older technologies include ATM, Frame Relay, SRP, FDDI and SMDS
- **Layer 3 exchange point**  
router based  
central or distributed



## 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?

- **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

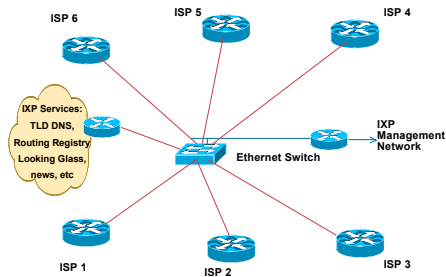
- **Solution**
  - Every ISP participates in the IXP
  - Cost is minimal – one local circuit covers all domestic traffic
  - International circuits are used for just international traffic – and backing up domestic links in case the IXP fails
- **Result:**
  - Local traffic stays local
  - QoS considerations for local traffic is not an issue
  - RTTs are typically sub 10ms
  - Customers enjoy the Internet experience
  - Local Internet economy grows rapidly



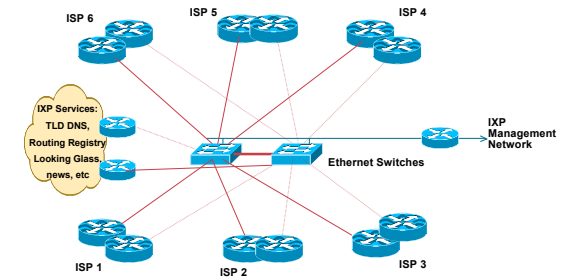
## Layer 2 Exchange

The traditional IXP

## 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 TLD, News, NTP servers
  - Routing Registry and Looking Glass

## Layer 2 Exchange

- Requires neutral IXP management
  - usually funded equally by IXP participants
  - 24x7 cover, support, value add services
- Secure and neutral location
- Configuration
  - private address space if non-transit and no value add services
  - ISPs require AS, basic IXP does not

## Layer 2 Exchange

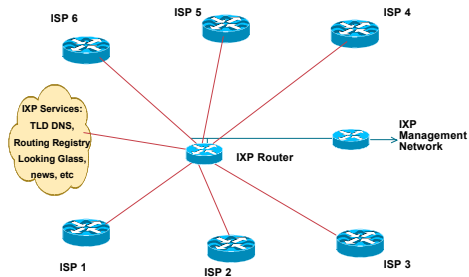
- Network Security Considerations
  - LAN switch needs to be securely configured
  - Management routers require TACACS+ authentication, vty security
  - IXP services must be behind router(s) with strong filters



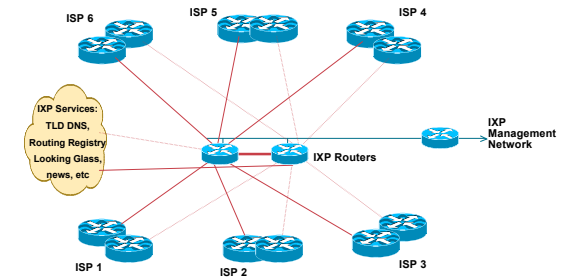
## Layer 3 Exchange

The wholesale transit ISP

## Layer 3 Exchange



## Layer 3 Exchange



## Layer 3 Exchange

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

## Layer 3 Exchange

- Requires neutral IXP management
  - usually funded equally by IXP participants
  - 24x7 cover, support, value add services
  - BGP configuration skills essential
- Secure and neutral location
- Configuration
  - private address space if non-transit and no value add services
  - ISPs and IXP require AS

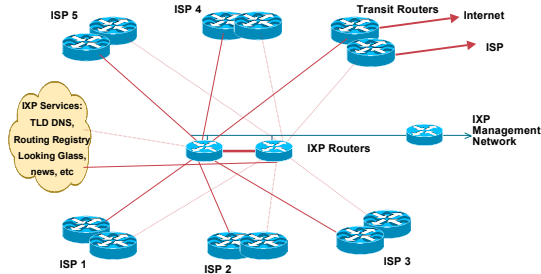
## Layer 3 Exchange

- Network Security Considerations
  - Core IXP router(s) require strong security, preferably with BGP neighbour authentication
  - Management routers require TACACS+ authentication, vty security
  - IXP services must be behind router(s) with strong filters

## Transit IXPs

- Provides local Internet exchange facility to members
- Also provides transit to Internet or upstream ISP
- Usually operated as a commercial service
- Usually layer 3 design

## Layer 3 Transit Exchange



## Layer 2 versus Layer 3

- **Layer 3**
  - IXP team requires good BGP knowledge
  - Rely on 3rd party for BGP configuration
  - Less freedom on who peers with whom
  - Could potentially compete with IXP membership
  - Easier to distribute over wide area

## Layer 2 versus Layer 3

- **Layer 2**
  - IXP team does not need routing knowledge
  - Easy to get started
  - More complicated to distribute over wide area
  - ISPs free to set up peering agreements with each other as they wish

## Layer 2 versus Layer 3 Summary

- **Layer 2 is a REAL internet exchange point**
- **Layer 3 is marketing concept used by Transit ISPs**



## IXP Design Considerations

## Routing

- **ISP border routers at the IXP generally should NOT be configured with default route or carry 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
- **Some IXPs use public addresses for the IX LAN**
  - Address space available from the RIRs
  - IXP terms of participation often forbid the IX LAN to be carried in the ISP member backbone

## Hardware

- **Don't mix port speeds**
  - if 10Mbps and 100Mbps connections available, terminate on different switches (L2 IXP)
- **Don't mix transports**
  - if terminating ATM PVCs and G/F/Ethernet, terminate on different devices
- **Insist that IXP participants bring their own router**
  - moves buffering problem off the IXP
  - security is responsibility of the ISP, not the IXP

## 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 MRTG to produce IX throughput graphs for member (or public) information

## Services to Offer

- **TLD 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 TLD DNS
- **Root server**
  - Anycast instances of I.root-servers.net, F.root-servers.net etc are present at many IXes
- **Usenet News**
  - Usenet News is high volume
  - could save bandwidth to all IXP members

## Services to Offer

- **Route Collector**
  - All IXP members peer with the route collector
  - Route collector shows the reachability information available at the exchange
  - Requires a simple router with large memory
- **Looking Glass**
  - one way of making the Route Collector routes available for global view
  - public or members only access

### Services to Offer Route Server

- **Reduces admin burden on IXP member routers**
  - only BGP session is with Route Server
  - Route Server supplies all paths it knows to the IXP member routers – no best path selection
- **Can use private AS**
  - Route Server software does not prepend its AS to the AS path
- **RSd (from Merit Network) has been used**

### Services to Offer

- **Network Time Protocol**
  - Locate a stratum 1 time source (GPS receiver, atomic clock, etc) at IXP
- **Multicast**
  - Provide MBONE and other multicast services for the common good

### Services to Offer Routing Registry

- **Routing Registry is used to register the routing policy of the IXP membership**
  - documenting peering relationships
  - auto-configuring of peer routers
- **Alternative is to use the public Internet Routing Registry (IRR)**



### IXP Design

Summary

### Summary

- **L2 IXP – most commonly deployed**
  - Based around ethernet switches
  - ATM and other old technologies are obsolete
- **L3 IXP – nowadays generally a marketing concept used by wholesale ISPs**
  - Does not offer the same flexibility as L2
  - Not recommended unless there are overriding regulatory or political reasons to do so



### Internet Exchange Point Design

ISP/IXP Workshops