



Internet Exchange Point Design

ISP/IXP Workshops

IXP Design

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



A bit of history

In a time long gone...

A Bit of History...

- **End of NSFnet – one major backbone**
- **move towards commercial Internet**
private companies selling their bandwidth
- **need for coordination of routing exchange between providers**
Traffic from ISP A needs to get to ISP B
- **Routing Arbiter project created to facilitate this**

What is an Exchange Point

- **Network Access Points (NAPs) established at end of NSFnet**

The original “exchange points”

- **Major providers connect their networks and exchange traffic**
- **High-speed network or ethernet switch**
- **Simple concept – any place where providers come together to exchange traffic**

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

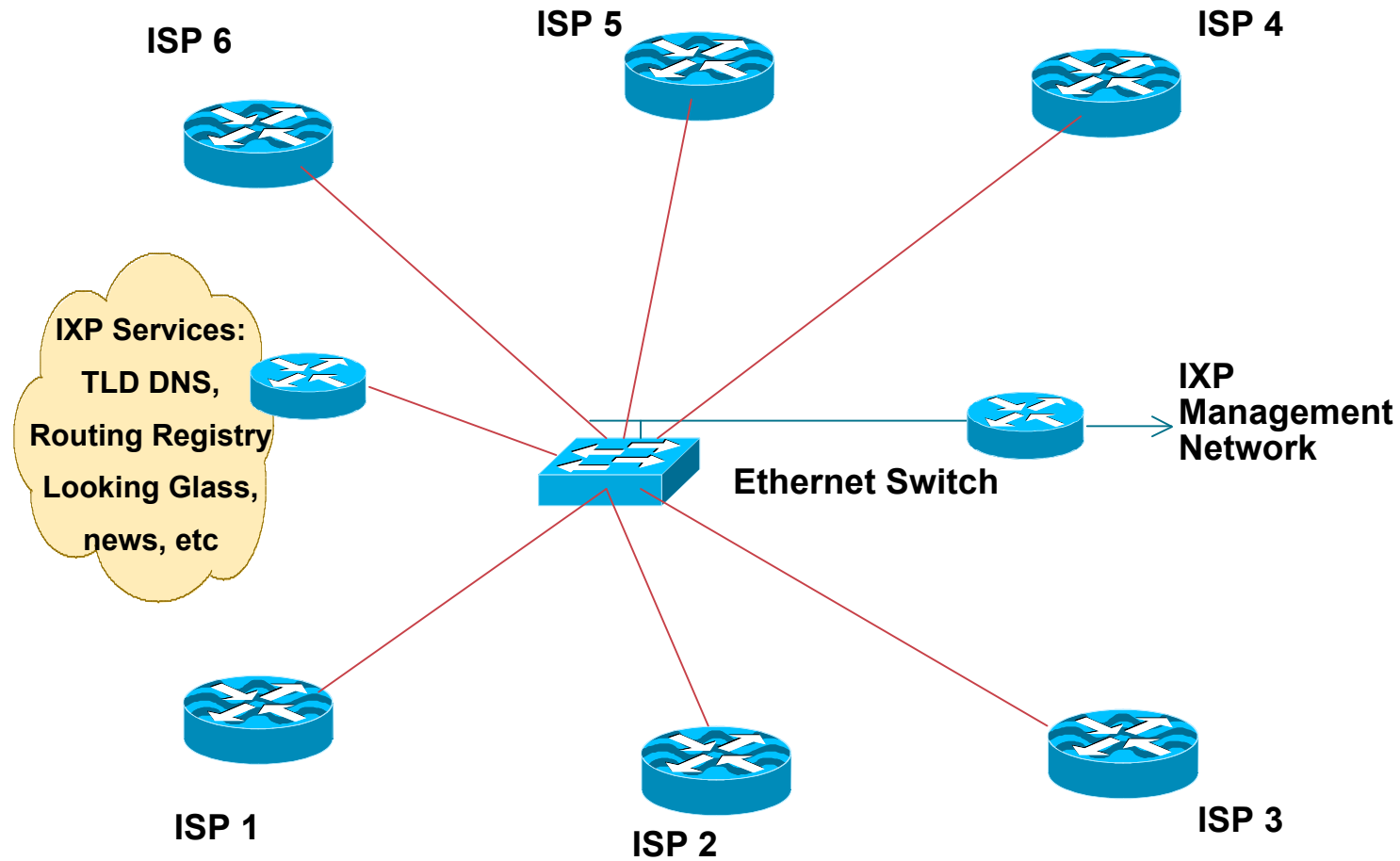
Local Internet economy grows rapidly



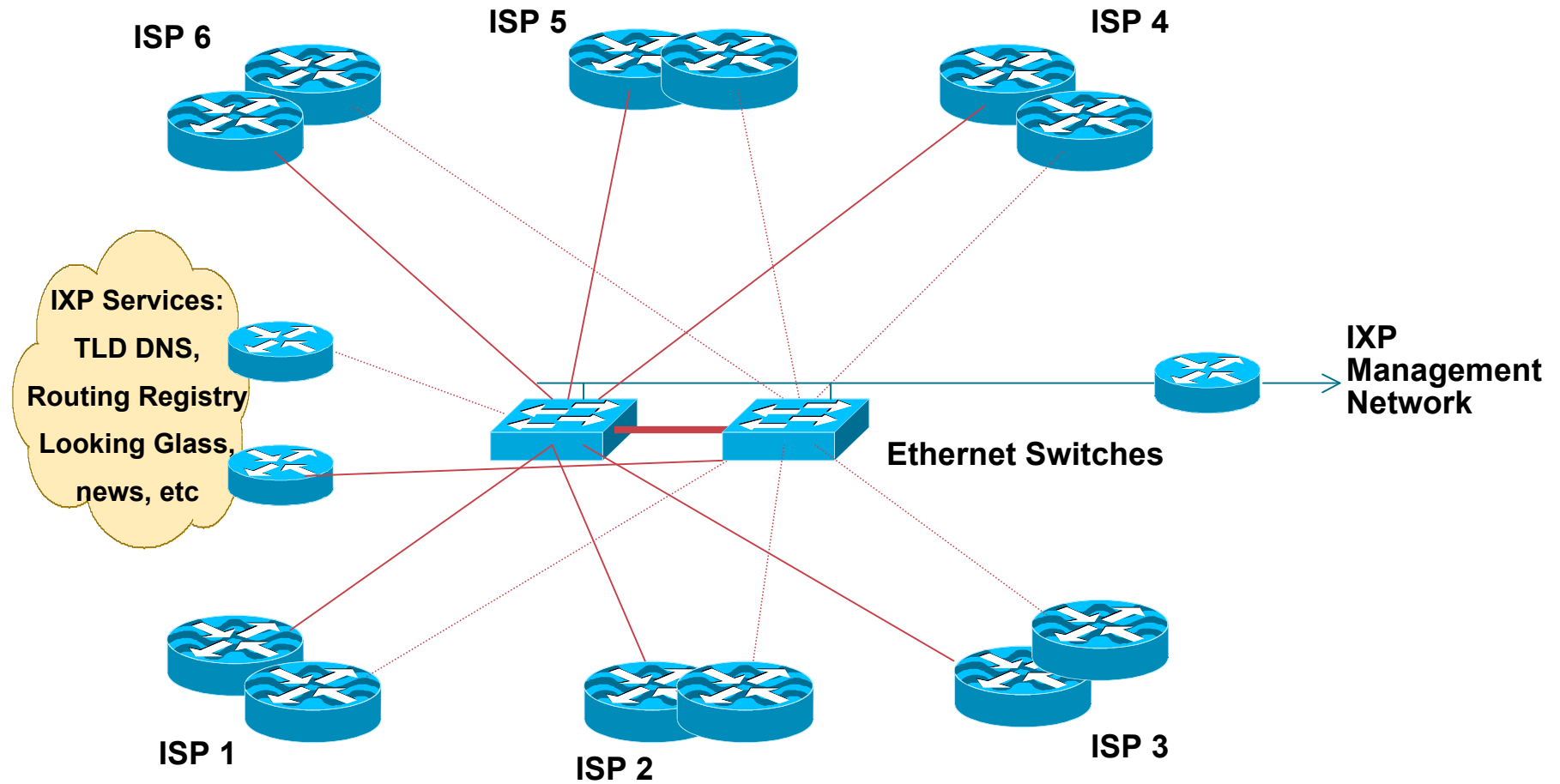
Layer 2 Exchange

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

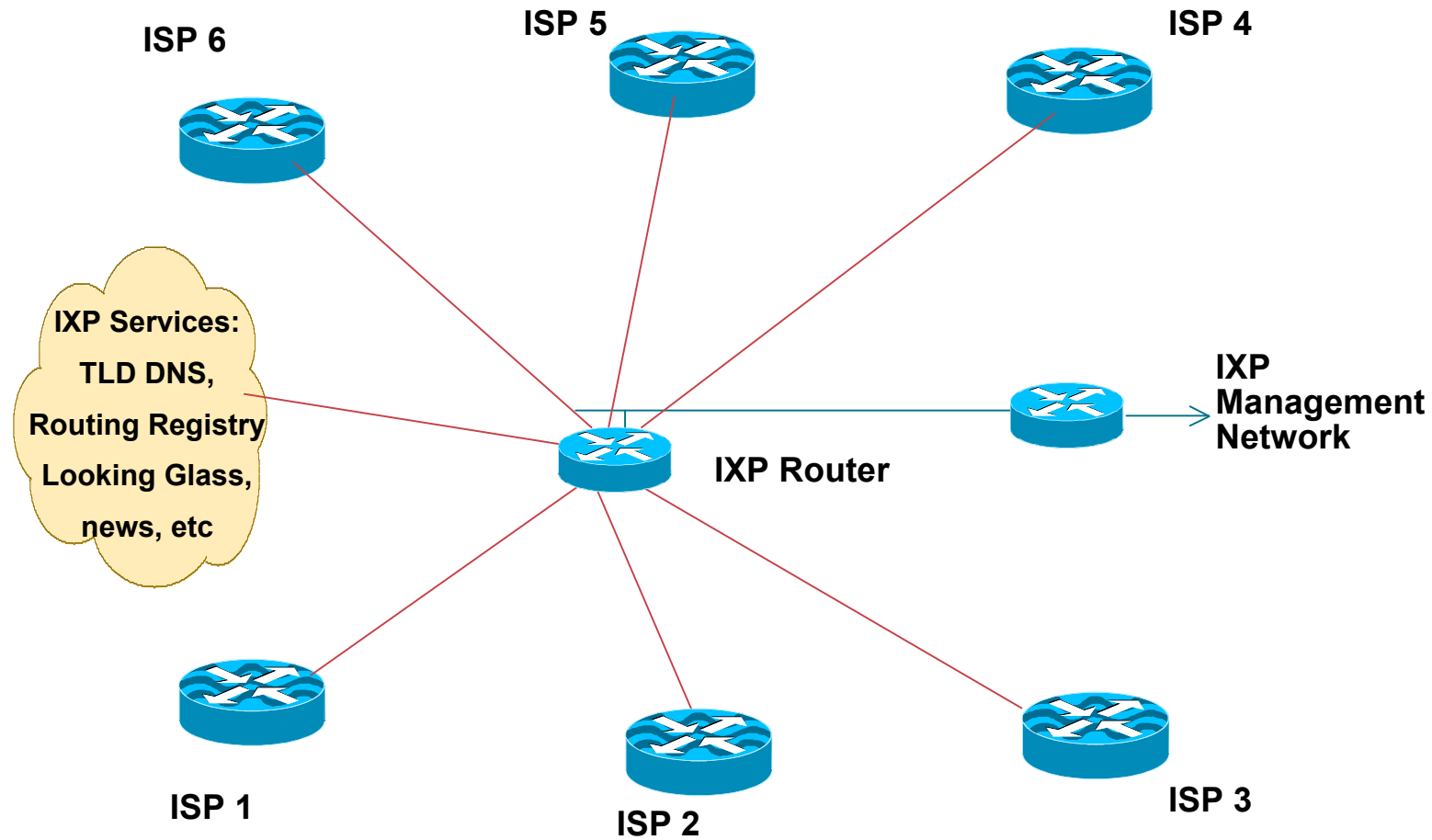
IXP services must be behind router(s) with strong filters



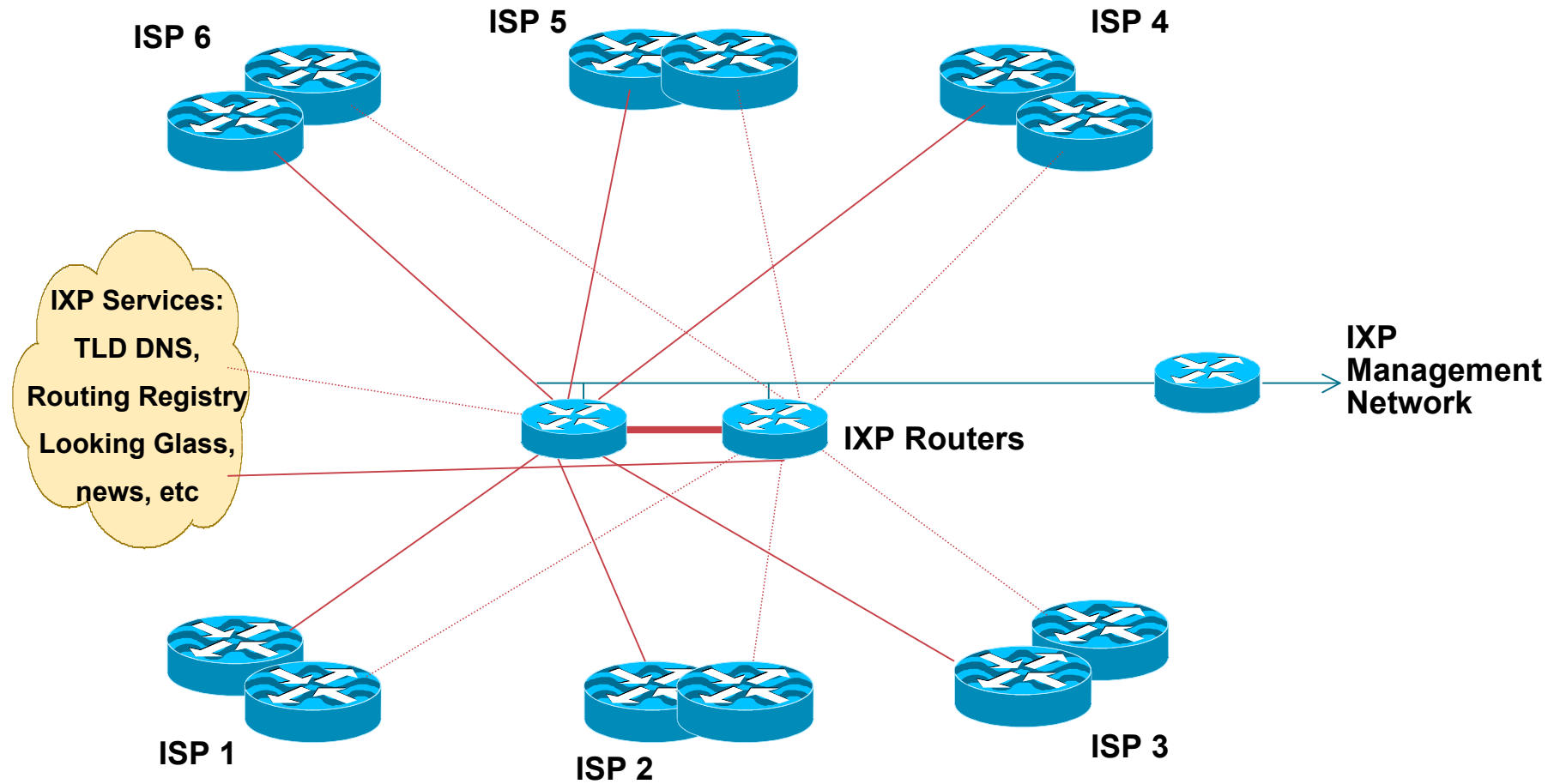
Layer 3 Exchange

The wholesale transit ISP

Layer 3 Exchange/Wholesale Transit ISP



Layer 3 Exchange/Wholesale Transit ISP



Layer 3 Exchange/Wholesale Transit ISP

- **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/Wholesale Transit ISP

- **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/Wholesale Transit ISP

- **Network Security Considerations**

Core IXP router(s) require strong security, preferably with BGP neighbour authentication

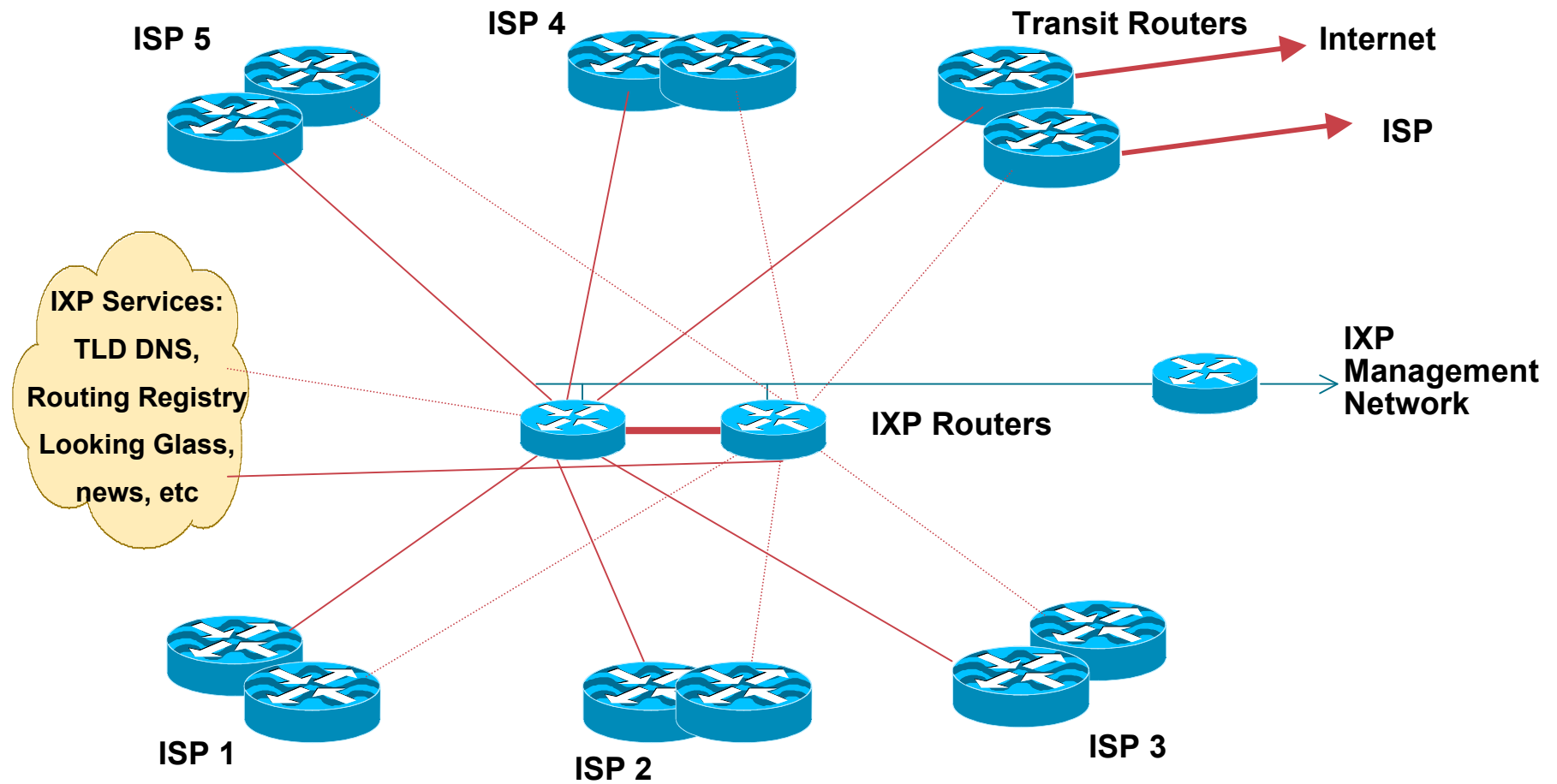
Management routers require TACACS+ authentication, vty security

IXP services must be behind router(s) with strong filters

Transit IXPs/Wholesale Transit ISP

- **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/Transit ISP



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

Usually competes 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

Exchange Point Design

- **The IXP Core is an Ethernet switch**
- **Has superseded all other types of network devices for an IXP**

**From the cheapest and smallest 12 or 24 port 10/100 switch
To the largest 32 port 10GigEthernet switch**

Exchange Point Design

- **Each ISP participating in the IXP brings a router to 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 10Gbps and 40Gbps**

Peering

- **Each participant needs to run BGP**

They need their own AS number

Public ASN, **NOT** private ASN

- **Each participant configures external BGP 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 required to peer with every other participant as part of their IXP membership

Has no history of success — the practice is **discouraged**

- **Multi-Lateral Peering (MLP)**

Each participant peers with every other participant

- **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

Routing

- **ISP border routers at the IXP generally should 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

- **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

- **Try not to 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

- **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

- **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**

Route collector shows the reachability information available at the exchange

Technical detail covered later on

- **Looking Glass**

One way of making the Route Collector routes available for global view (e.g. www.traceroute.org)

Public or members only access

Services to Offer

- **Content Redistribution/Caching**

For example, Akamised update distribution service

- **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)



What can go wrong...

The different ways IXP operators harm their IXP...

What can go wrong?

Concept

- **Some ISPs attempt to cash on the reputation of IXPs**
- **Market 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 — really Internet Transit Providers**
 - Router used rather than a Switch**
 - Most famous example: SingTelIX**

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**
- **An IXP:**
 - is **NOT** a competition
 - is **NOT** a profit making business

What can go wrong?

Rules and Restrictions

- **IXPs try to compete with their membership**
Offering services that ISPs would/do offer their customers
- **IXPs run as a closed privileged club e.g.:**
Restrictive membership criteria (closed shop)
- **IXPs providing access to end users rather than just Service Providers**
- **IXPs interfering with ISP business decisions e.g. 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 (e.g. LINX)

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

- **iBGP Route Reflector used to distribute prefixes between IXP participants**
- **Claimed Advantage (1):**
 - Participants don't need to know about or run BGP**
- **Actually a Disadvantage**
 - IXP Operator has to know BGP**
 - ISP not knowing BGP is big commercial disadvantage**
 - ISPs who would like to have a growing successful business need to be able to multi-home, peer with other ISPs, etc — these activities require BGP**

What can go wrong?

Routing Design Errors (cont)

- **Route Reflector Claimed Advantage (2):**

Allows an IXP to be started very quickly

- **Fact:**

IXP is only an Ethernet switch — setting up an iBGP mesh with participants is no quicker than setting up an eBGP mesh

What can go wrong?

Routing Design Errors (cont)

- **Route Reflector Claimed Advantage (3):**

IXP operator has full control over IXP activities

- **Actually a Disadvantage**

ISP participants surrender control of:

Their border router; it is located in IXP's AS

Their routing and peering policy

IXP operator is single point of failure

If they aren't available 24x7, then neither is the IXP

BGP configuration errors by IXP operator have real impacts on ISP operations

What can go wrong?

Routing Design Errors (cont)

- **Route Reflector Disadvantage (4):**

Migration from Route Reflector to “correct” routing configuration is highly non-trivial

ISP router is in IXP’s ASN

Need to move ISP router from IXP’s ASN to the ISP’s ASN

Need to reconfigure BGP on ISP router, add to ISP’s IGP and iBGP mesh, and set up eBGP with IXP participants and/or the IXP Route Server



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**
- **Filter! Filter! Filter!**

And do reverse path check

Exchange Point examples

- **LINX in London, UK**
Ethernet switches
- **AMS-IX in Amsterdam, NL**
Ethernet switches
- **SIX in Seattle, US**
Ethernet switches
- **JPNAP in Tokyo, Japan**
Ethernet switches

Features of IXPs

- **Redundancy & Reliability**
Multiple switches, UPS
- **Support**
NOC to provide 24x7 support for problems at the exchange
- **DNS, Route Collector, Content & NTP servers**
ccTLD & root servers
Content redistribution systems such as Akamai
Route Collector – Routing Table view

Features of IXPs

- **Location**
neutral co-location facilities
- **Address space**
Peering LAN
- **AS**
If using Route Collector/Server
- **Route servers (optional)**
- **Statistics**
Traffic data – for membership

More info about IXPs

- **<http://www.ep.net/ep-main.html>**

Excellent resource for ip address allocation for exchanges, locations of XPs in the world, AUPs and other policies

- **<http://www.pch.net/documents>**

Another excellent resource of IXP locations, papers, IXP statistics, etc



IXP Design

Summary

Summary

- **L2 IXP – most commonly deployed**

The core is an ethernet switch

ATM and other old technologies are obsolete

- **L3 IXP – nowadays is 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

Avoid!



Internet Exchange Point Design

ISP/IXP Workshops