Internet Exchange Points (IXPs)

Scalable Infrastructure Workshop

AfNOG 2008

Objectives

- To be able to explain what an Internet Exchange Point (IXP) is
- To be able to explain why ISPs participate in IXPs
- To understand why IXPs are important
- To review some current IXP designs used today
- To think about how to set up an IXP in your environment

Introduction to Internet Exchange Points

- A bit of history
- What are they?
- Why use them?
- Design Considerations

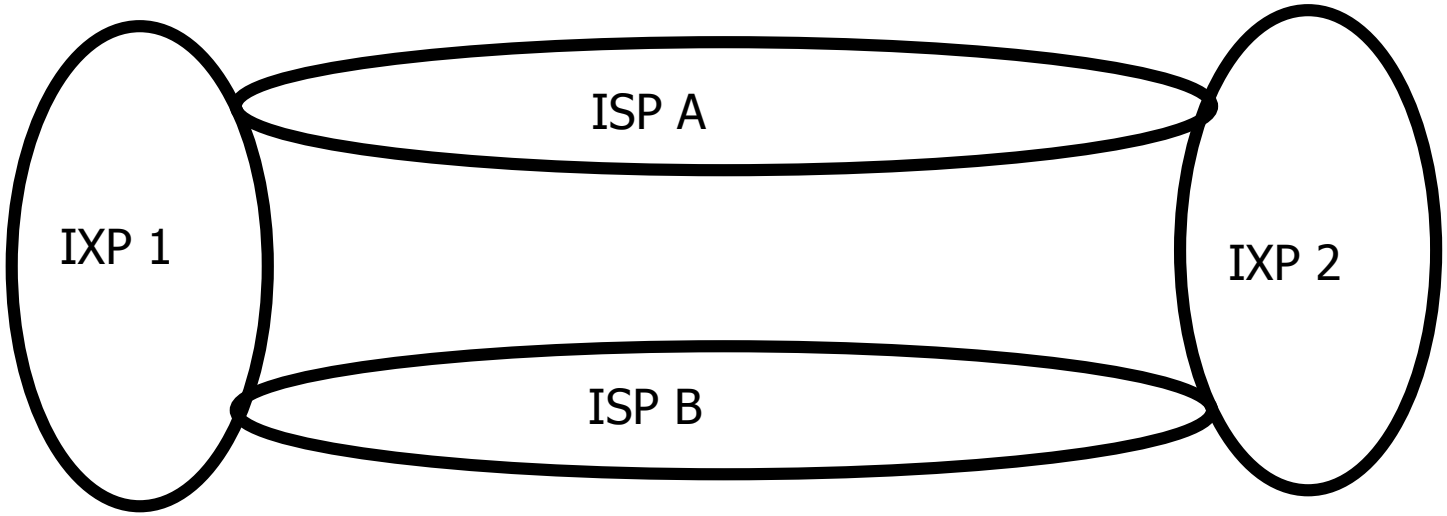
A Bit of History…

- 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

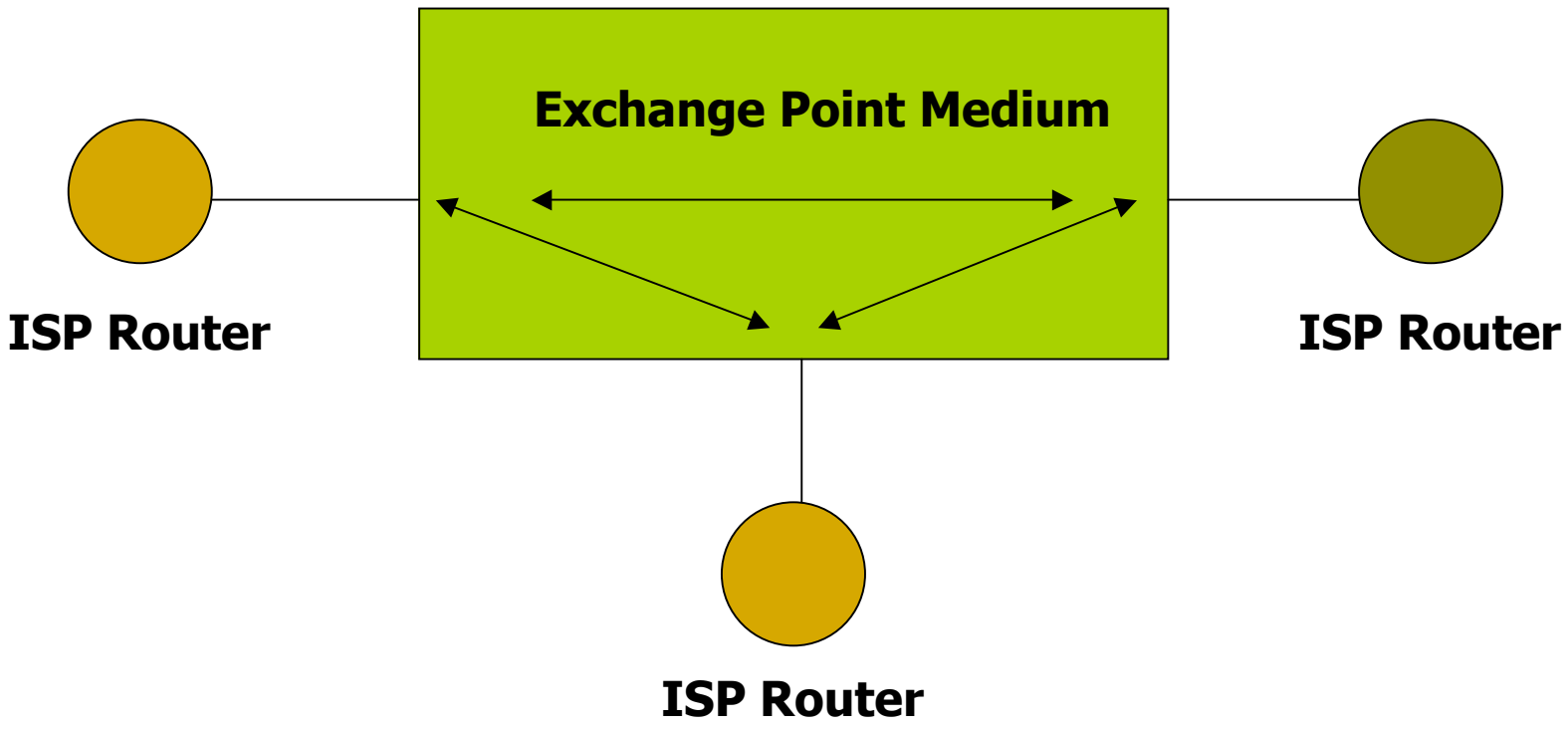
- Network Access Points (NAPs) established at end of NSFnet
  - The original "exchange points"
- Major providers connect their networks and exchange traffic
- High-speed network or ethernet switch
- Simple concept – any place where providers come together to exchange traffic

Internet Exchange Points



ISPs connect at Exchange Points or Network Access Points to exchange traffic

Conceptual Diagram of an IXP



Why use an IXP?

Internet Exchange Point
Why peer?

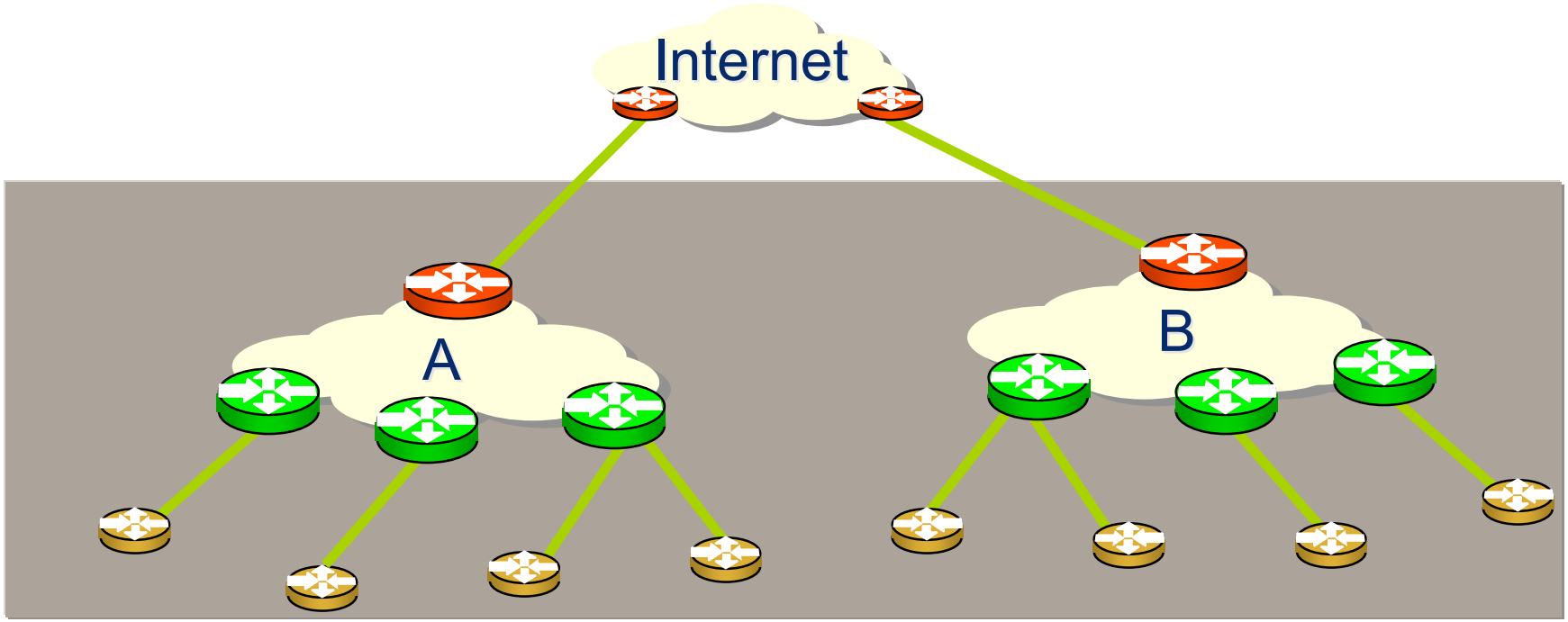
- Consider a region with one ISP
  - They provide internet connectivity to their customers
  - They have one or two international connections
- Internet grows, another ISP sets up in competition
  - They provide internet connectivity to their customers
  - They have one or two international connections
- 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 order of magnitude or two more than domestic bandwidth
  - Becomes congested with local traffic
- Wastes money, harms performance

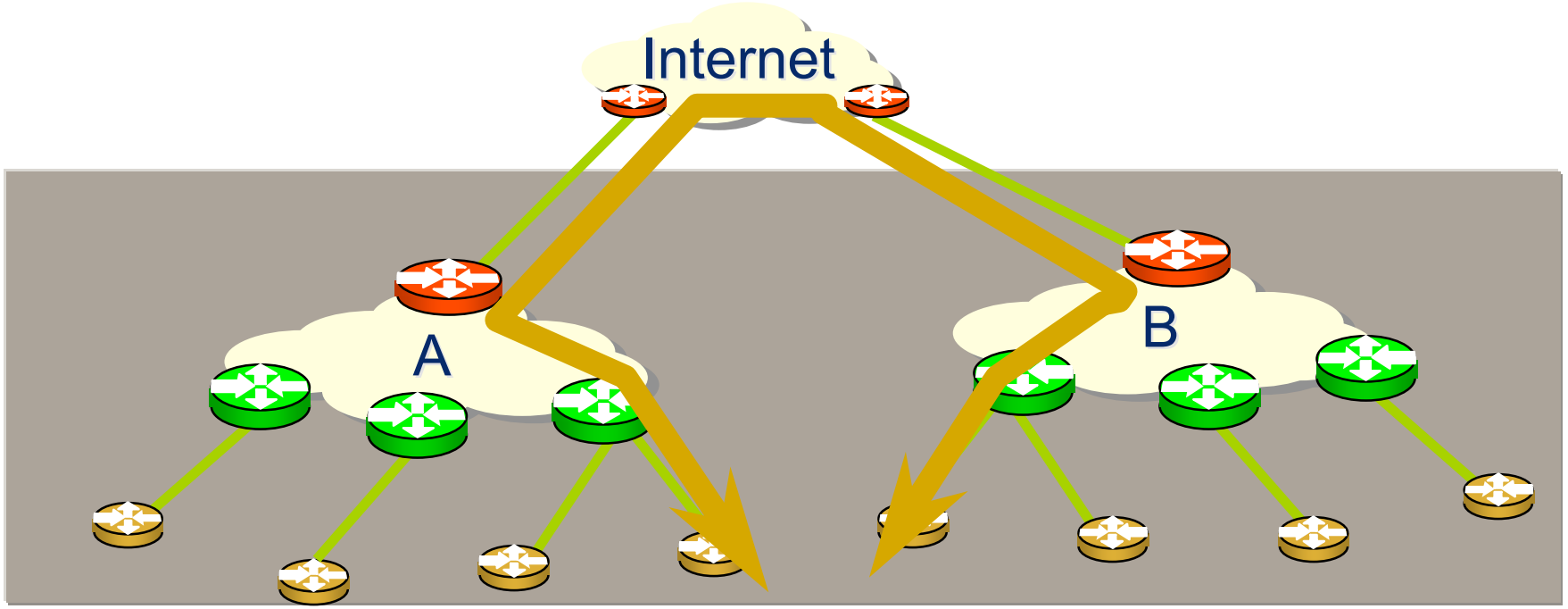
Internet Exchange Point
Why peer?

- Multiple service providers
- Each with Internet connectivity



Why IXPs?

- Is not cost effective
- Backhaul issue causes cost to both parties

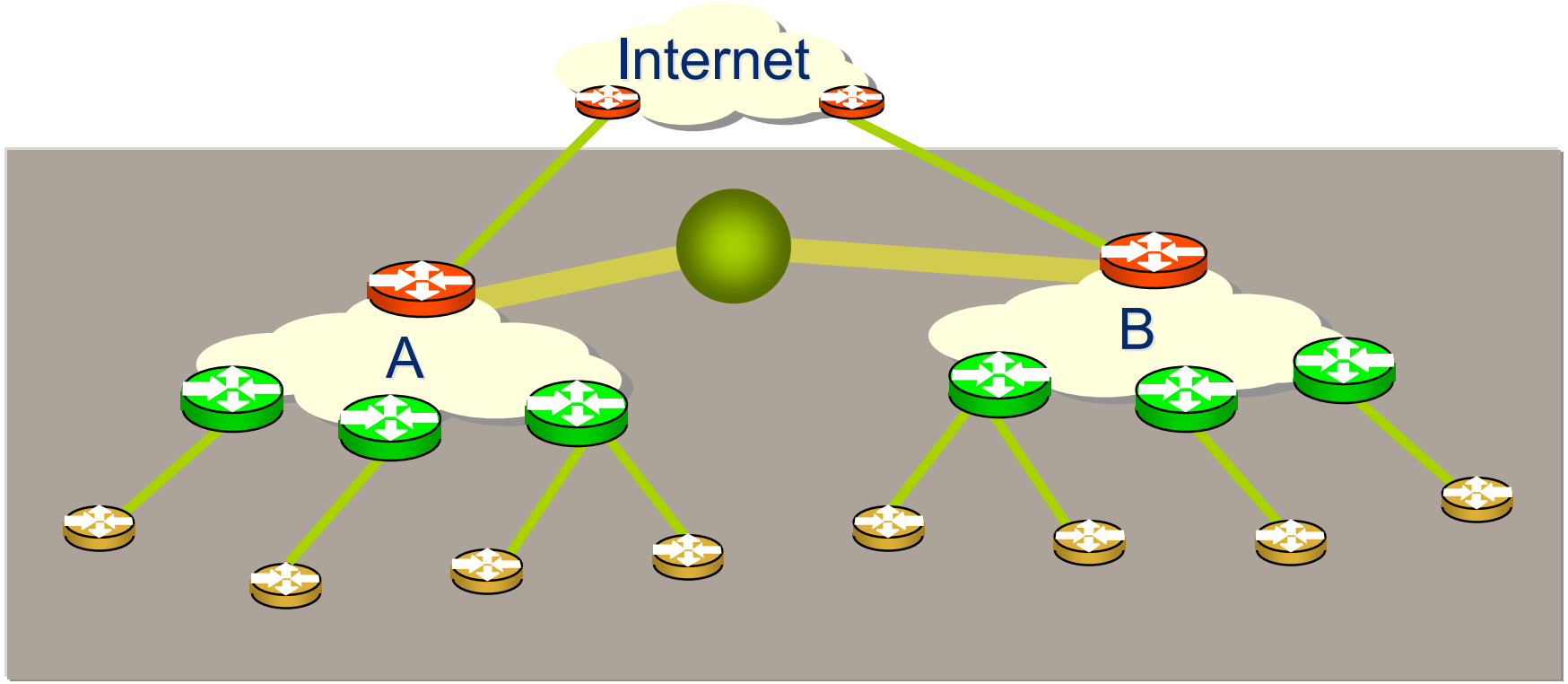


Internet Exchange Point
Why peer?

- 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

Why IXPs?

- Domestic Interconnection



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

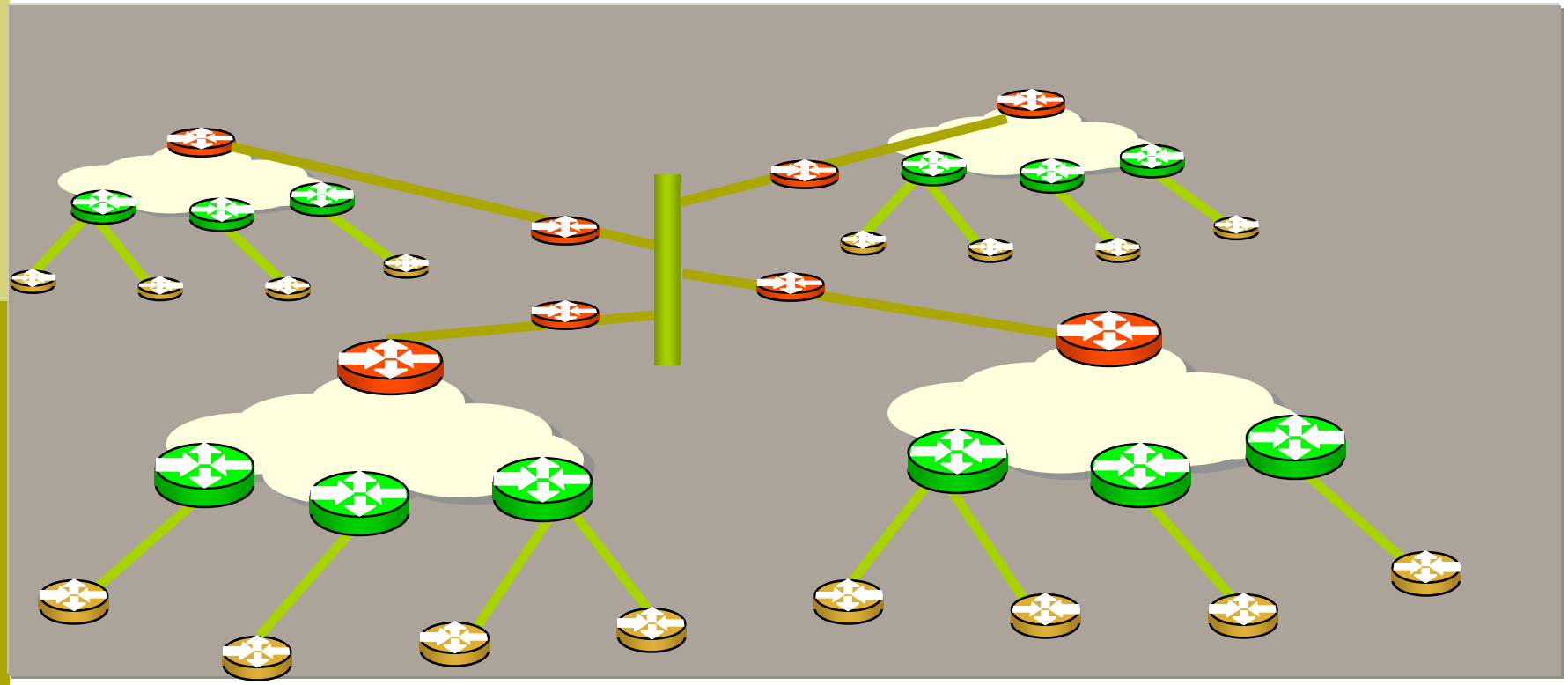
- Every participant has to buy just one whole circuit
  - From their premises to the IXP
- Rather than N-1 half circuits to connect to the N-1 other ISPs
  - 5 ISPs have to buy 4 half circuits = 2 whole circuits → already twice the cost of the IXP connection

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

Internet Exchange Point

- Ethernet switch in the middle



Why use an IXP?

- PEERING
  - Shared medium vs. point-to-point
  - Shared
    - can exchange traffic with multiple peers at one location via one interface
  - Point-to-Point
    - for high volumes of traffic

Why use an IXP?

- KEEP LOCAL TRAFFIC LOCAL!!!
  - ISPs within a region peer with each other at the local exchange
  - No need to have traffic go overseas only to come back
  - Much reduced latency and increased performance

Why use an IXP?

- SAVES MONEY!!!
  - Traffic going overseas means transit charges paid to your upstream ISP
  - Money stays in local economy
    - Used to provide better local infrastructure and services for customers
  - Customers pay less for Internet access
    - Therefore more customers sign up
    - ISP has more customers, better business

Why use an IXP?

- VASTLY IMPROVES PERFORMANCE!!!
  - Network RTTs between organisations in the local economy is measured in milliseconds, not seconds
  - Packet loss becomes virtually non-existent
  - Customers use the Internet for more products, services, and activities

Why use an IXP?

- Countries or regions with a successful IXP have a successful Internet economy
- Local traffic stays local
- Money spent on local ‘net infrastructure
- Service Quality not an issue

- All this attracts businesses, customers, and content

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

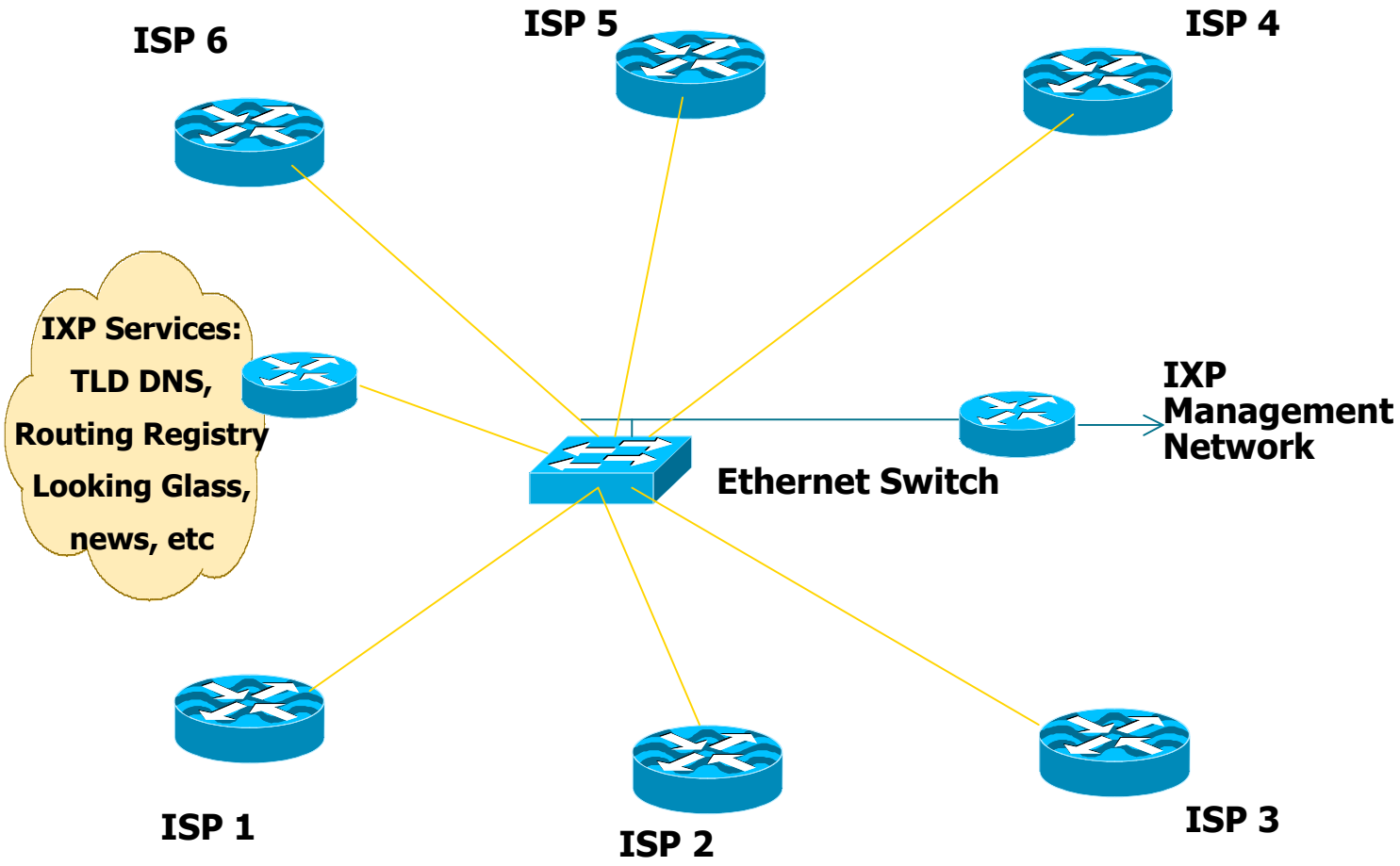
Exchange Point Design

- 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

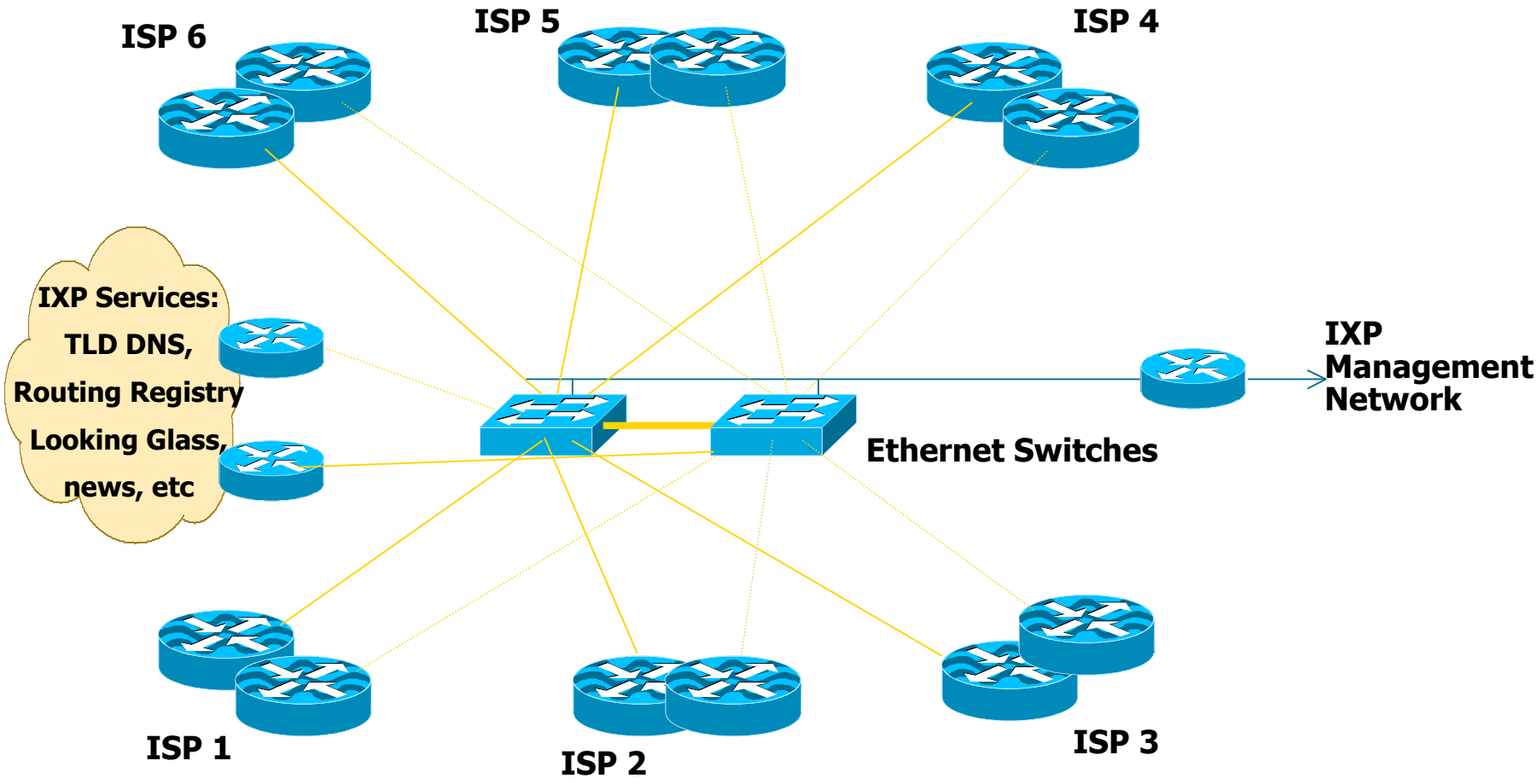
Exchange Point Design

- 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 higher speeds

Exchange Point Design



Exchange Point Design



Peering at an IXP

- 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 forced to peer with every other participant as part of their IXP membership
  - Has no history of success — strongly 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
  - Set BGP next-hop to local router (Cisco IOS next-hop-self)
- 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

IP Address Space

- Some IXPs use private addresses for the IXP LAN
  - Public address space means the IXP network can be leaked to the Internet, which could be undesirable
  - Filtering RFC1918 address space by ISPs is Best Practice; this avoids leakage
- Some IXPs use public addresses for the IXP LAN
  - Address space is available from the RIRs for IXPs
  - IXP terms of participation usually forbid carrying the IXP LAN addressing in the ISP backbone

Hardware

- Try not to mix port speeds
  - if 10Mbps and 100Mbps connections available, terminate on different switches
- Insist that IXP participants bring their own router
  - Moves buffering problem off the IXP
  - Ensures integrity of the IXP
  - Security is responsibility of the ISP, not the IXP

Services to Locate at an IXP

- 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 F, I, etc root nameservers are present at many IXes
- Usenet News
  - Usenet News is high volume
  - Could save bandwidth to all IXP members

Services to Locate at an IXP

- 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 Locate at an IXP

- 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…

What can go wrong? Concept

- Some ISPs attempt to cash on the reputation of IXPs
- Market Internet transit services as "Internet Exchanges"
  - "We are exchanging packets with other ISPs, so we are an Internet Exchange!"
  - 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 or exclusive membership criteria
- 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 are ok (e.g. LINX, AMSIX)

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 at a 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 impact 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

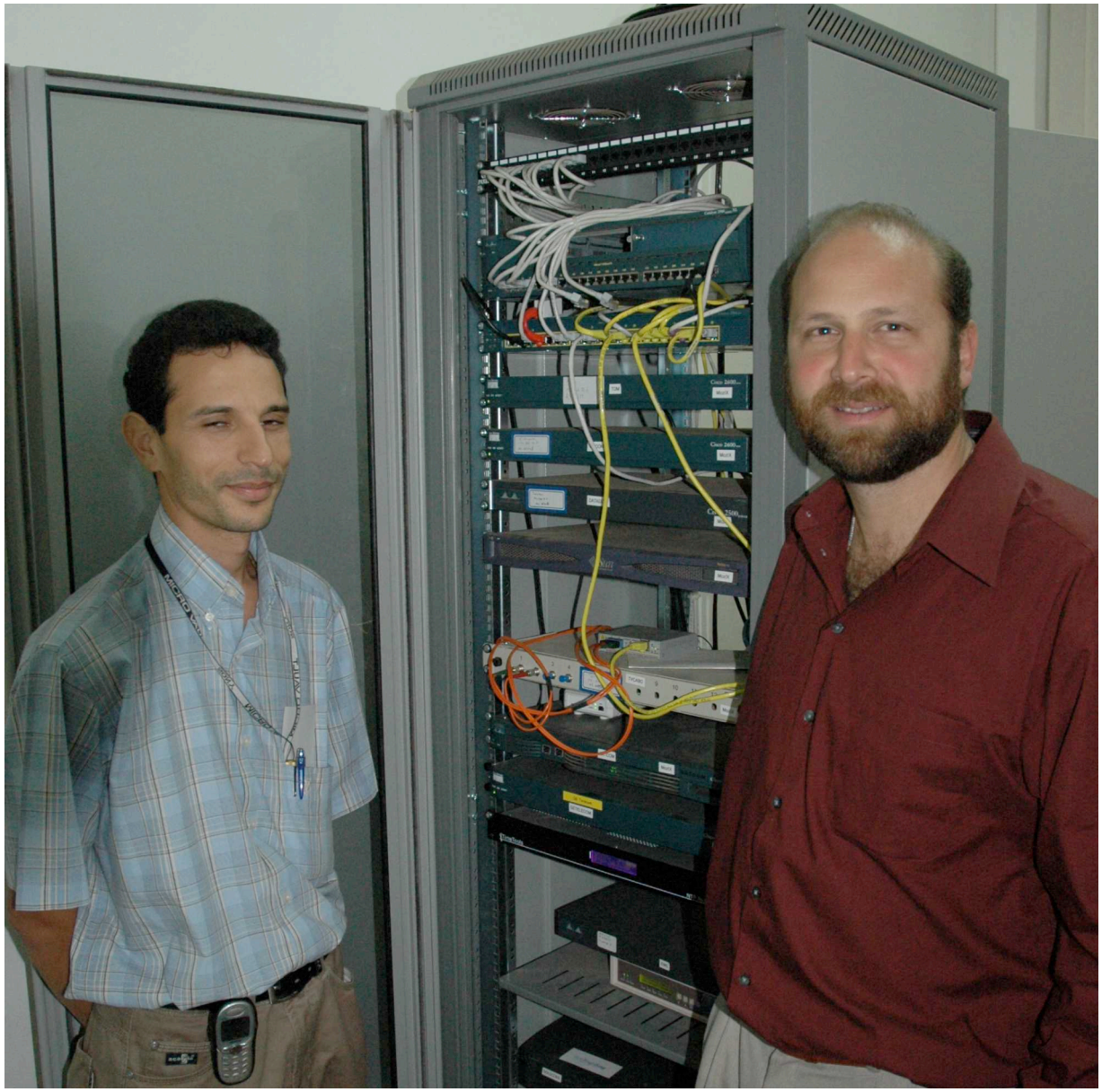
Exchange Point examples

- 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

Exchange Points in Africa

- BINX – Botswana
- CR-IX – Cairo
- CI-XP – Abidjan
- GIXP – Accra
- iBiX – Ibadan
- IXP-Ang – Luanda
- JINX – Johannesburg
- KINIX – Kinshasa
- KIXP – Nairobi
- MOZIX – Maputo
- RINEX – Kigali
- SZIXP – Mbabane
- TIX – Dar es Salaam
- UiXP – Kampala
- ZINX – Harare

Source: http://www.nsrc.org/AFRICA/afr_ix.html



Mozambique Internet Exchange, Maputo

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

Things to think about...

- Do you need to be at an Exchange Point?
- Would you want to start an Exchange Point?
- Would keeping local traffic local benefit your ISP?
- Would your environment (politically, etc.) support an Exchange Point?

Discussion

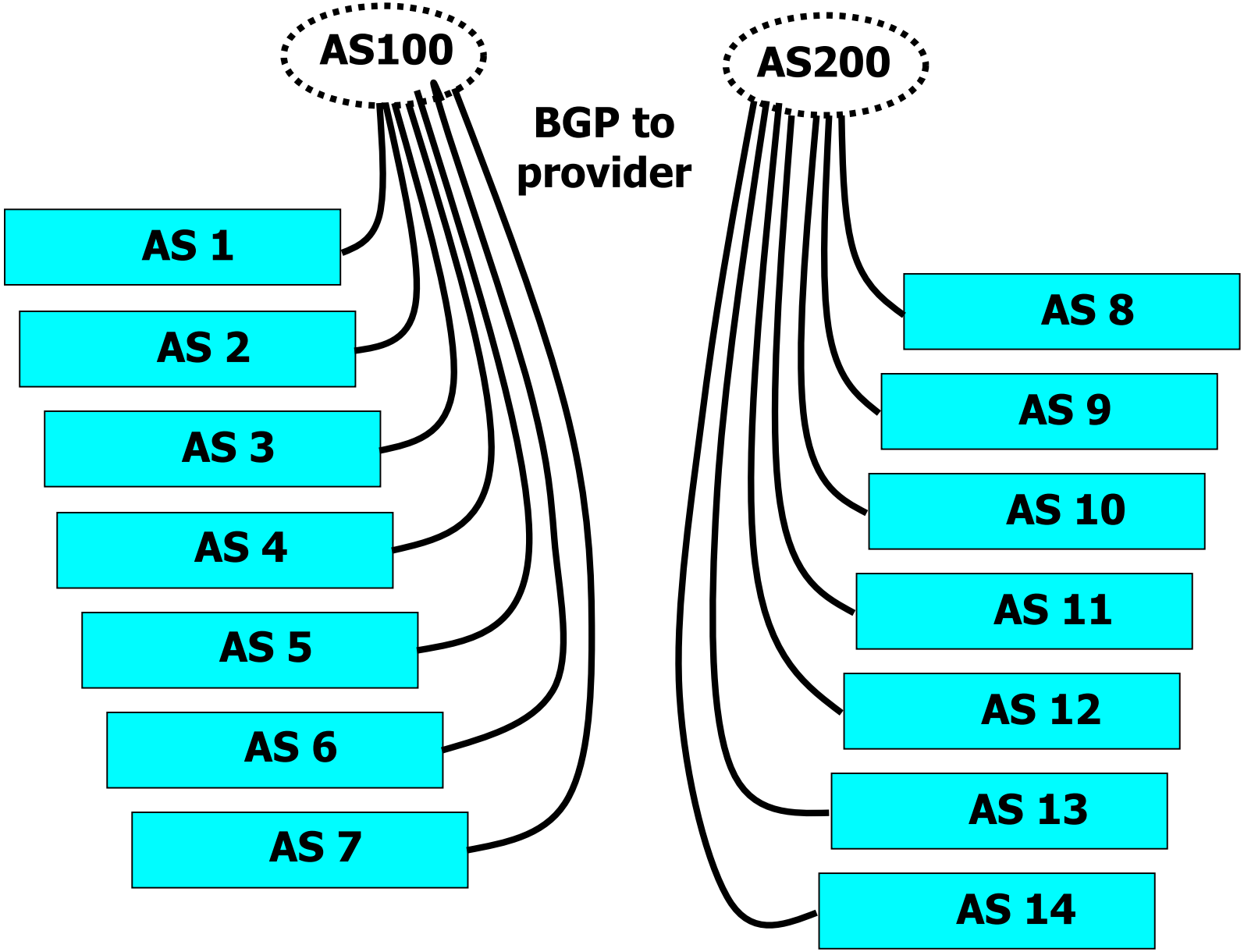
- How would you build an exchange point in your environment?
- Who would connect?
- What services would you provide?
- What policies would you enforce?
- What does your environment look like?
  - Is it feasible to set up an IXP?

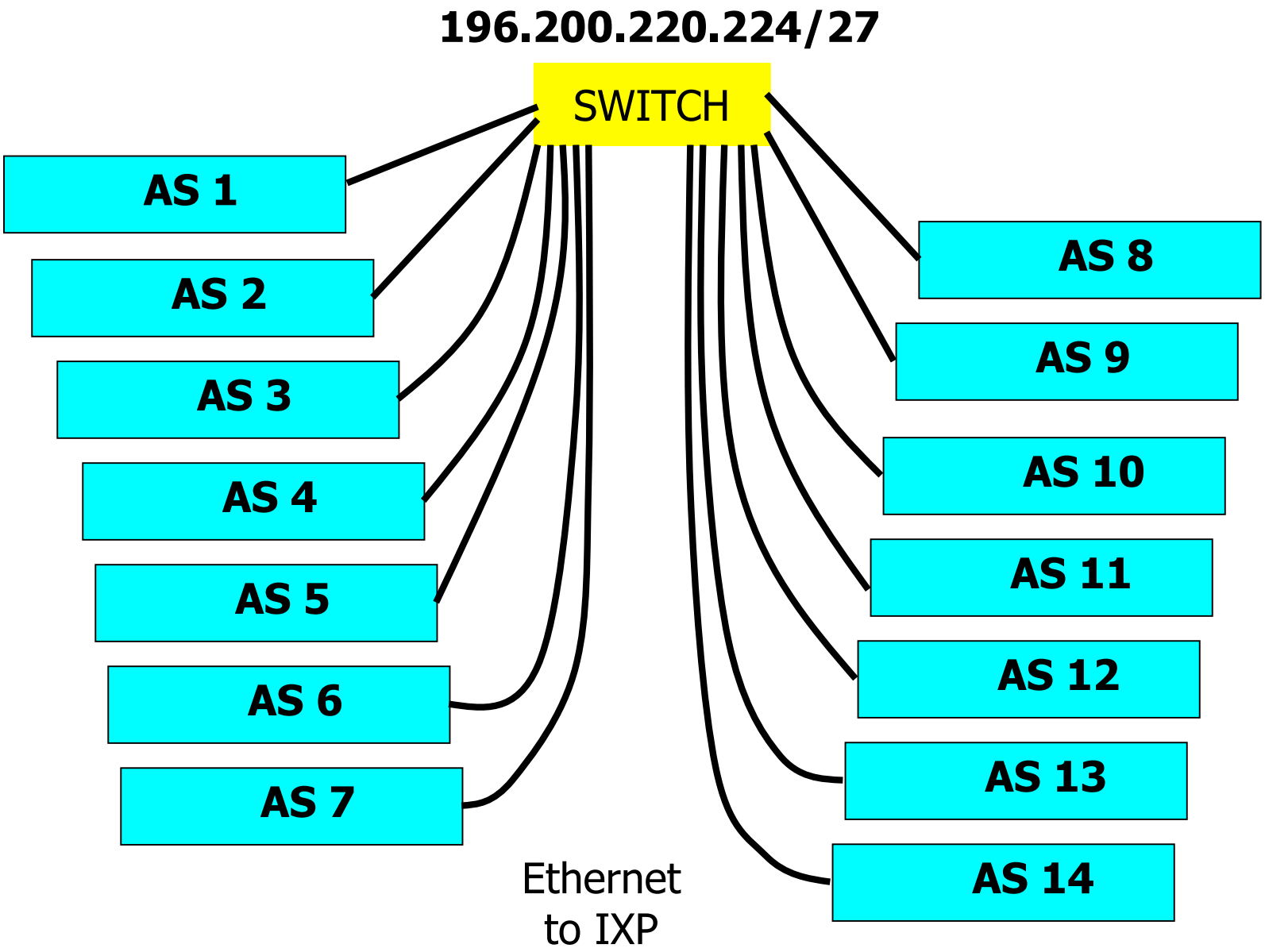
Important to Remember...

- Exchange Points can be as simple as an ethernet HUB!!!!
- Keeping local traffic local
  - improves performance
  - cheaper
  - often simple to do!

Exercise

Building an IXP


AS100
AS200
BGP to
provider
AS 1
AS 2
AS 3
AS 4
AS 5
AS 6
AS 7
AS 8
AS 9
AS 10
AS 11
AS 12
AS 13
AS 14


196.200.220.224/27
SWITCH
AS 1
AS 2
AS 3
AS 4
AS 5
AS 6
AS 7
AS 8
AS 9
AS 10
AS 11
AS 12
AS 13
AS 14
Ethernet
to IXP