Overview of Routing & Interconnection

Internet Society Policymaker Programme

IETF 119

Brisbane

Last updated 17th March 2024

Acknowledgements

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

- Use of these materials is encouraged as long as the source is fully acknowledged and this notice remains in place

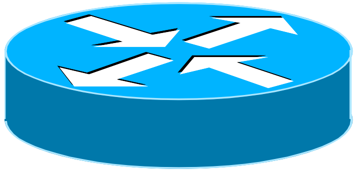
- Bug fixes and improvements are welcomed
  - Please email *workshop (at) bgp4all.com*

Philip Smith

Routing & Interconnection

- Routers
- Routing Protocols
- Peering & Transit
- Internet Hierarchy
- Interconnection Goals

What is a Router?

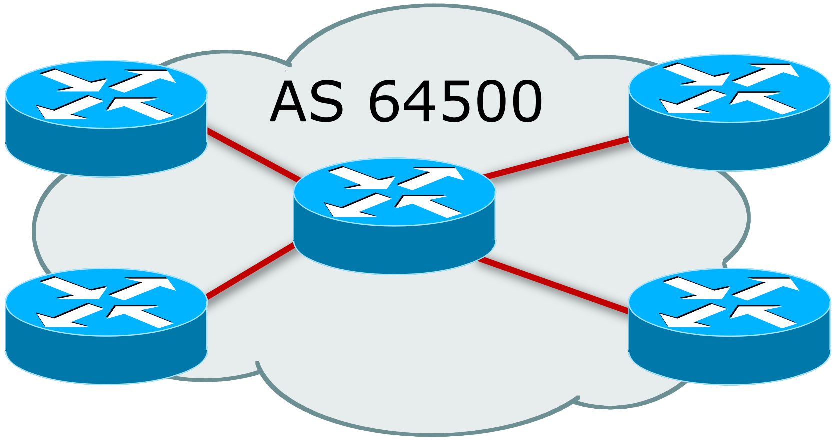


- A router is a device used for connecting different networks together
- A router generally has at least two interfaces
- A router looks at the destination address in the IP packet, and decides how to forward it

The Routing Table

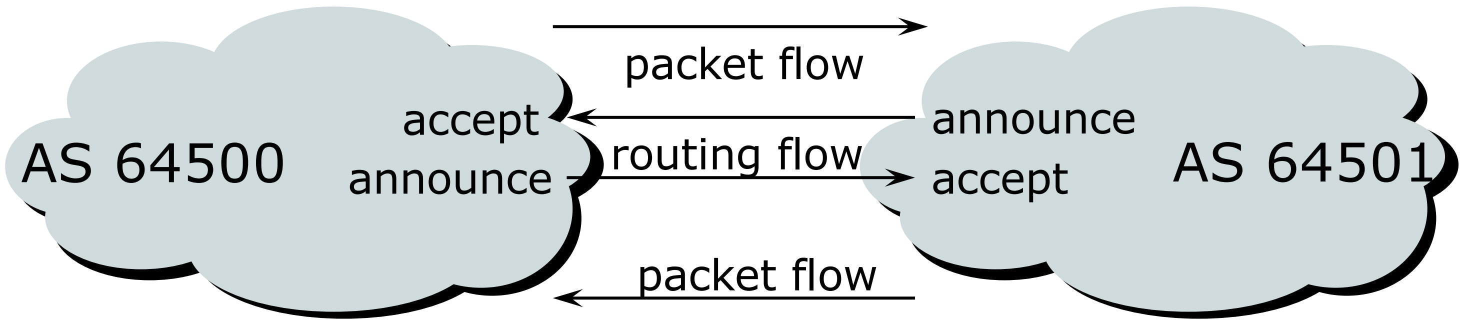
- Each router/host has a *routing table*, indicating the path or the next hop for a given destination host or a network
- The router/host tries to match the destination address of a packet against entries in the routing table
- If there is a match, the router forwards it to the corresponding gateway router or directly to the destination host
- Default route is taken if no other entry matches the destination address
  - Default route is a pointer to a neighbouring router/host that can be used as a last resort

Autonomous System (AS)



- A group of routers with same routing policy
- Single routing protocol
- Single ownership, trust and administrative control

Routing flow and packet flow



For networks in AS645400 and AS64501 to communicate:

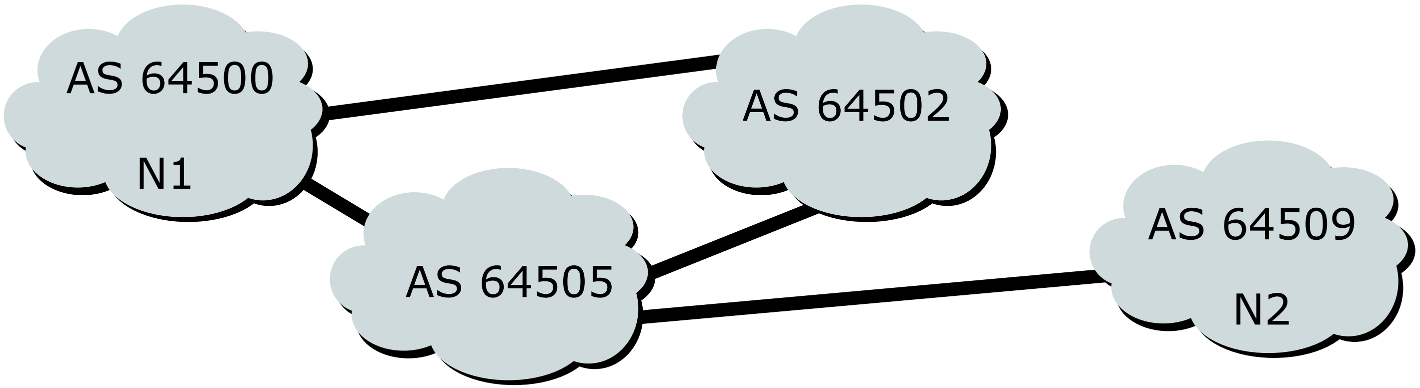
AS64500 must announce to AS64501

AS64501 must accept from AS64500

AS64501 must announce to AS64500

AS64500 must accept from AS64501

Routing Flow/Packet Flow: With multiple ASes



- For net N1 in AS64500 to send traffic to net N2 in AS64509:
  - AS64509 must originate and announce N2 to AS64505.
  - AS64505 must accept N2 from AS64509.
  - AS64505 must announce N2 to AS64500 or AS64502.
  - AS64500 must accept N2 from AS64505 or AS64502.
- For two-way packet flow, similar policies must exist for N1

Routing & Interconnection

- Routers
- Routing Protocols
- Peering & Transit
- Internet Hierarchy
- Interconnection Goals

1: How Does Routing Work?

- Internet is made up of the Network Operators who connect to each other's networks
- How does an operator in Kenya tell an operator in Japan what end-site customers they have?
- And how does that operator send data packets to the customers of the Japanese operator, and get responses back
  - After all, as on a local ethernet, two-way packet flow is needed for communication between two devices

2: How Does Routing Work?

- The operator in Kenya could buy a direct connection to the operator in Japan
  - But this doesn’t scale – there are thousands of distinct networks, would need thousands of connections, and cost would be astronomical
- Instead, the operator in Kenya tells his neighbouring operators what end-sites they have
  - And the neighbouring operators pass this information on to their neighbours, and so on
  - This process repeats until the information reaches the operator in Japan

3: How Does Routing Work?

- This process is called "Routing"
- The mechanisms used are called "Routing Protocols"
- Routing and Routing Protocols ensures that
  - The Internet can scale
  - Thousands of network operators can provide connectivity to each other
  - We have the Internet we see today

4: How Does Routing Work?

- The Network Operator in Kenya doesn’t actually tell its neighbouring operators the names of the end-sites
  - (network equipment does not understand names)
- Instead, it has received an IP address block as a member of the Regional Internet Registry serving Kenya
  - Its customers have received address space from this address block as part of their “Internet service”
  - And it announces this address block to its neighbouring operators – this is called announcing a “route”

Routing Protocols

- Routers use "routing protocols" to exchange routing information with each other
  - IGP is used to refer to the process running on routers inside a provider's network
    - The two commonly used IGPs are OSPF and IS-IS
  - EGP is used to refer to the process running between routers bordering directly connected provider networks
    - There is only one EGP: BGP

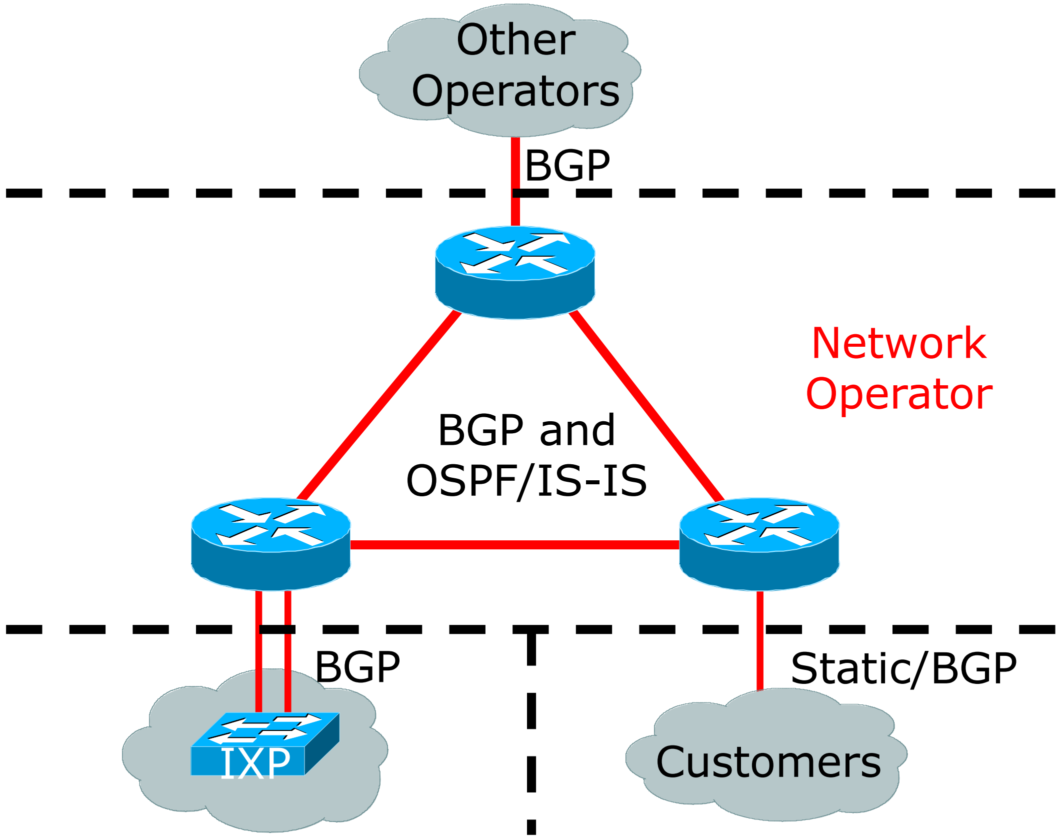
Why an IGP?

- Runs within an Autonomous System
- Carries information about internal infrastructure prefixes
- Computes the optimum path within a Network Operator's infrastructure
- Network Operator backbone scaling
  - Hierarchy within the operator network
  - Limiting scope of failure

Why an EGP?

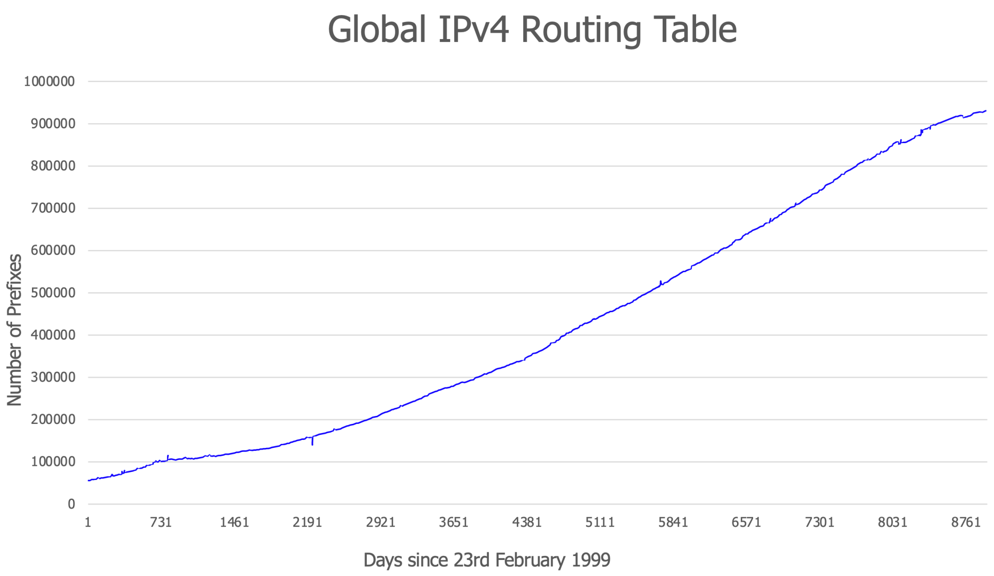
- Used to convey routing information between Autonomous Systems
- Scaling to a large network (today's Internet!)
  - Hierarchy
  - Limit scope of failure
- Define Administrative Boundary (AS relationships)
- BGP speaking routers dynamically compute the best path across the Internet
- Operator Policy
  - Control reachability of prefixes
  - Adjust/optimise paths to destinations

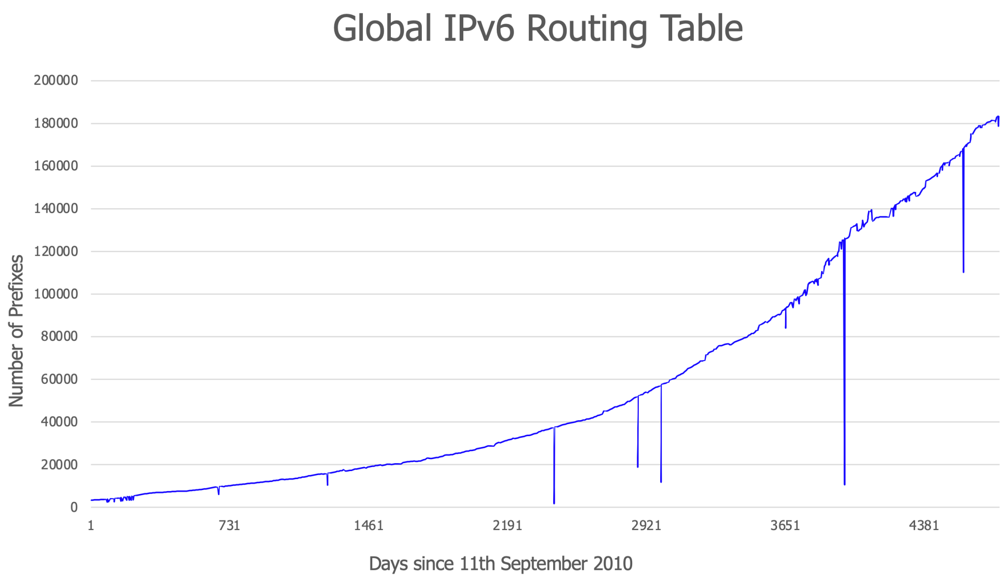
Hierarchy of Routing Protocols



BGP today

- Carries ~940k IPv4 and ~190k IPv6 prefixes
  - The network destinations of the global Internet
- Carries IPv4 routes of ~75k Autonomous Networks
- Carries IPv6 routes of ~32k Autonomous Networks





BGP today

- Adaptable & Self-Healing:
  - If an operator announces/withdraws a prefix this update message is propagated by BGP across the entire Internet
  - If the best path between two destinations is no longer available, an alternative path (if it exists) will be dynamically calculated by BGP speaking routers
- BGP is trusting by design
  - Allows rapid deployment of new infrastructure
  - Policy language implemented by network operators determines:
    - What is accepted/rejected from a neighbouring network
    - What is announced to a neighbouring network
  - Absence of policy means all routes known to that AS are shared with the neighbour
    - RFC8121 (2017) states requirements, but few implementations mandate this by default

Routing & Interconnection

- Routers
- Routing Protocols
- Peering & Transit
- Internet Hierarchy
- Interconnection Goals

Definitions

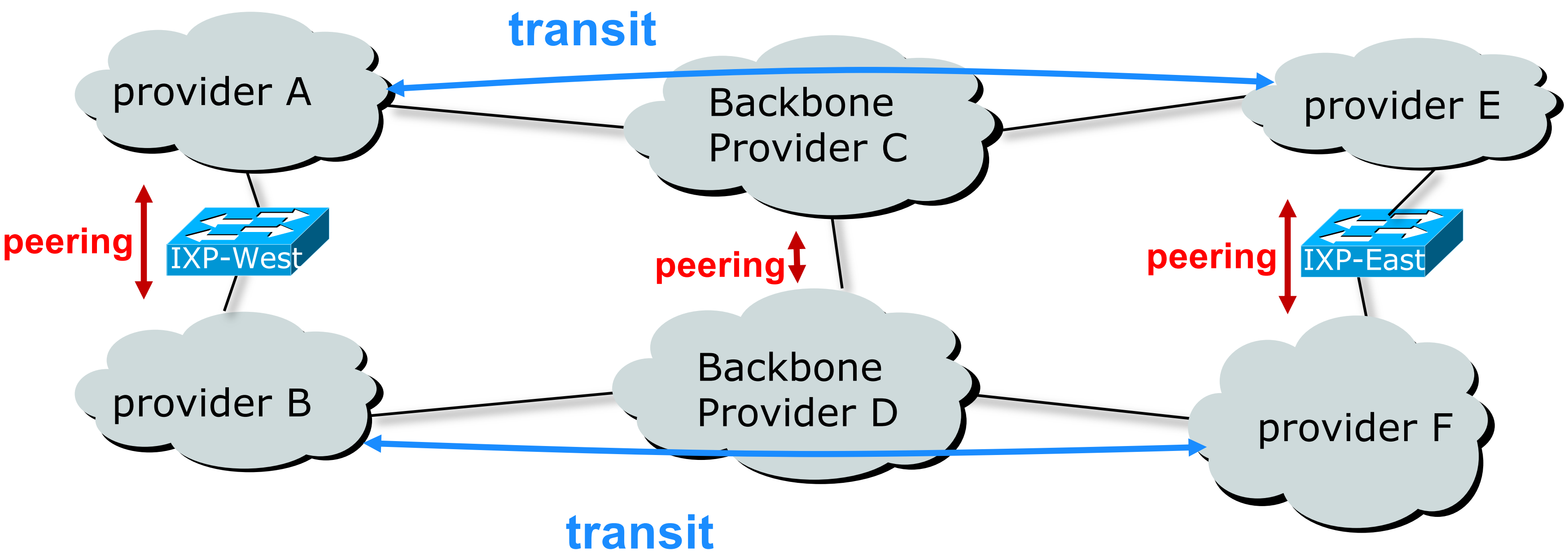
- Network Operator
  - An organisation running an IP backbone
  - Provides access to end users or other network operators
  - Sometimes called a **Service Provider** or a **Network Provider**
- ISP
  - Internet Service Provider
  - Usually commercial, for profit
- REN
  - Research & Education Network
  - Providing access for Universities, Colleges & Research Institutions
  - Non-commercial, research & educational use only
  - Acceptable Use Policies apply

Definitions

- Transit
  - Carrying traffic across a network
  - Usually **for a fee**

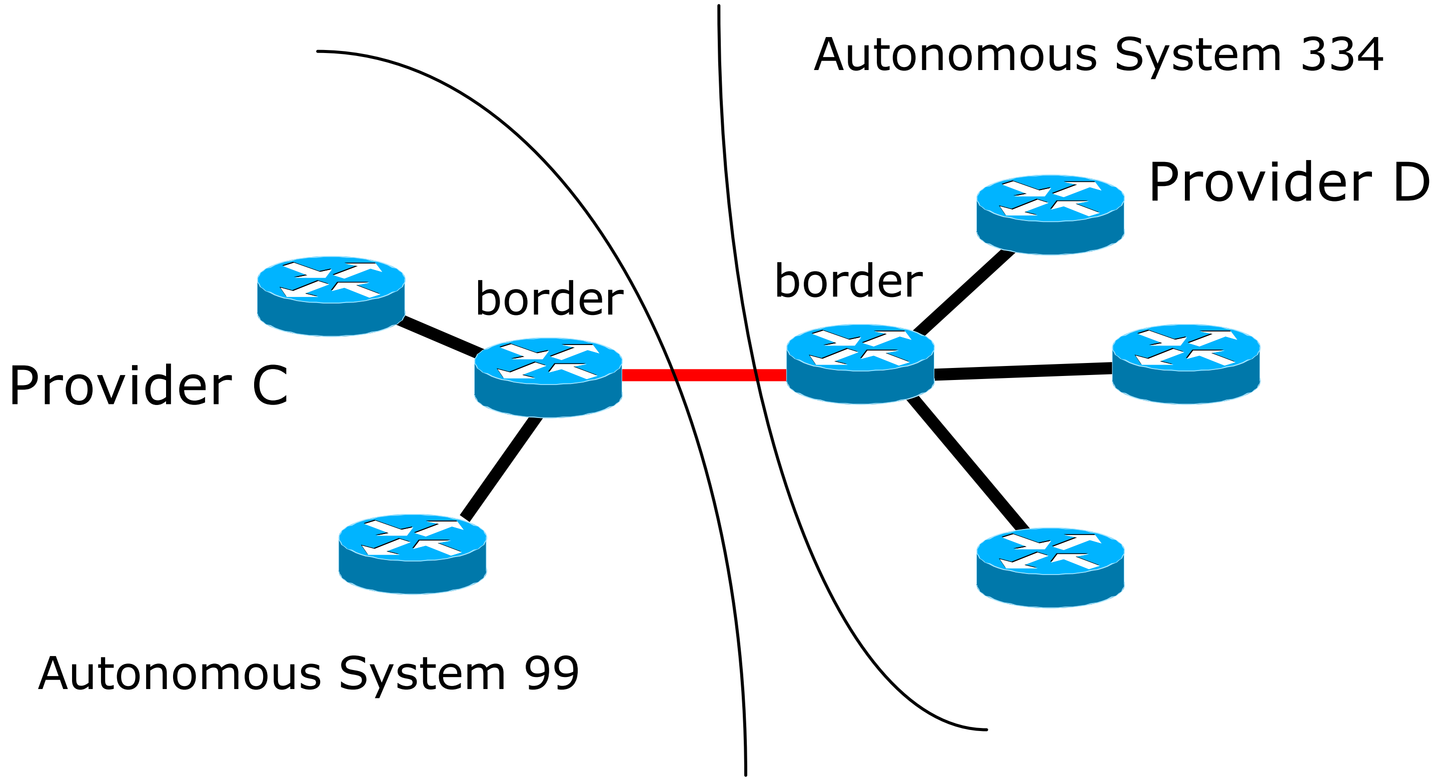
- Peering
  - Exchanging routing information and traffic
  - Usually **for no fee**
  - Sometimes called **settlement free peering**

Peering and Transit example



A and B peer for free, but need transit arrangements with C and D to get packets to/from E and F

Private Interconnect



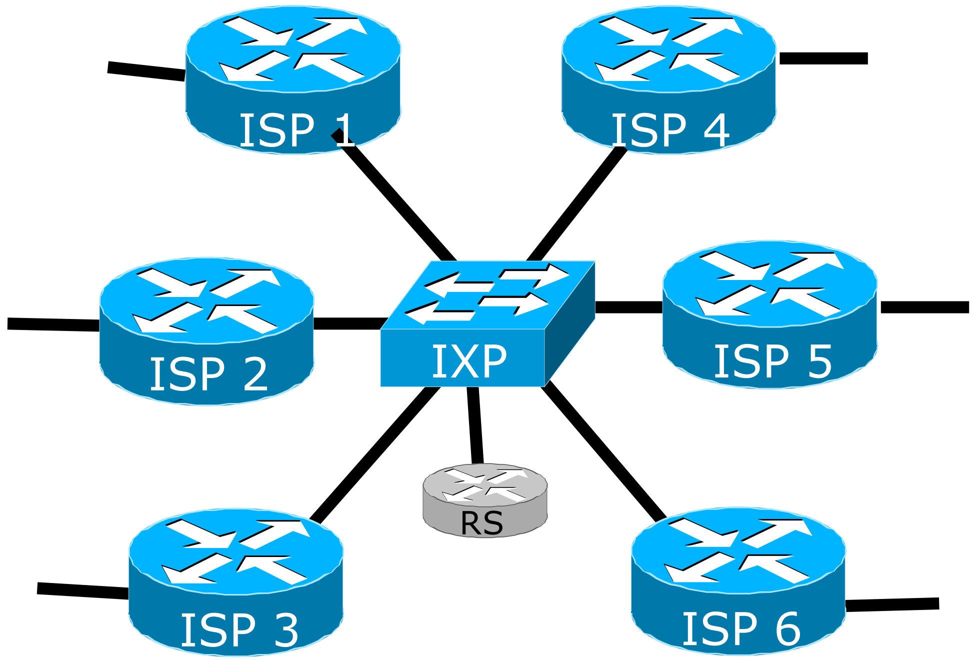
Public Interconnect

- A location or facility where several network operators are present and connect to each other over a common shared media
- Why?
  - Reduce latency
  - Increase bandwidth
  - Improve performance
  - Save money
- Called an Internet eXchange Point (IXP)
  - Some locations use the name “Network Access Point” (NAP)

Public Interconnect (IXP)

- Centralised (in one facility)
- Larger Interconnects are Distributed (connected via fibre optics) over the local area
- Switched interconnect
  - Global standard: Ethernet (Layer 2)
  - Several older technologies have been used in the past
- Each operator establishes **peering** relationship with the other operators at the IXP
- The IXP's Route Server helps scale peering

Public Interconnect (IXP)



Each of these represents a peering router in a different autonomous system

IXPs today

- Cooperative/home-grown:
  - Formed by the network operators taking part
    - Run as non-profit, membership organisation, equal say, shared costs
    - Neutral location

- Commercial
  - Run as a business by a third party, *or*
  - Offered by a datacentre facility as a service to customers

Types of Peering

- Private Peering
- Public Peering (at IXPs)
  - Bi-Lateral
  - Multi-Lateral
  - Mandatory Multi-Lateral

Private Peering

- Where two network operators agree to interconnect their networks via a private interconnect
- Once the operators interconnect:
  - Settlement Free Peering
    - No traffic charges
    - **The most common form of peering**
  - Paid Peering
    - Where two operators agree to exchange traffic charges for a peering relationship
    - (Marketing name for Local Transit? 🤨)

Public Peering

- *Bi-lateral Peering*
  - Very similar to Private Peering, but usually takes place at a public peering point (IXP)
- *Multilateral Peering*
  - Takes place at IXPs, where operators all peer with each other via the IXP Route Servers
- *Mandatory Multilateral Peering*
  - Where operators are forced to peer with each other as condition of IXP membership
  - **Strongly discouraged: Has no record of success**
  - (But some are still determined to prove 30 years of industry experience wrong 😬)

Types of Operator Peering Policies

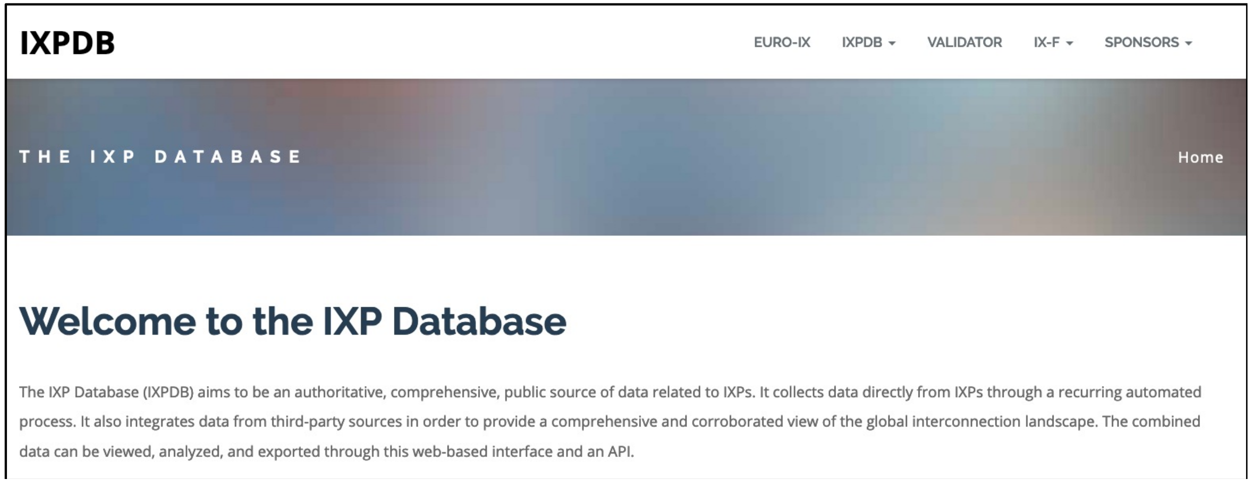
- *Open Peering*
  - Where a network operator publicly states that they will peer with all parties who approach them for peering
  - Commonly found at IXPs where the member participates via the Route Server
- *Selective Peering*
  - Where a network operator's peering policy depends on the nature of the operator who requests peering with them
  - At IXPs, the operator will not peer with the Route Server but will only peer bilaterally
- *Restrictive Peering*
  - Where a network operator decides who its peering partners are, and is generally not approachable to considering peering opportunities

The Peering Database

- The Peering Database documents network operator peering policies
  - https://www.peeringdb.com
- All operators with AS numbers are recommended to register in the PeeringDB
  - All operators who are considering peering or are peering must be in the PeeringDB to enhance their peering opportunities
  - Most major network operators will not peer with you unless you have a PeeringDB entry
- Participation in peering fora is encouraged too
  - Global Peering Forum (GPF) – for North American operators
  - Many Regional Peering Fora (including Europe, Middle East, Africa, Asia, Caribbean, Latin America)
  - Many countries now have their own Peering Fora

The IXP Database

- The IXPDB documents IXPs and their participants around the world
  - https://ixpdb.euro-ix.net/en/
- All Internet Exchange Point operators are recommended to register their IXP in the database
  - IXPs using IXP Manager (https://www.ixpmanager.org) will have this happen as part of the IXP Manager set up
  - Provides the LAN IP addresses of each member to facilitate automation



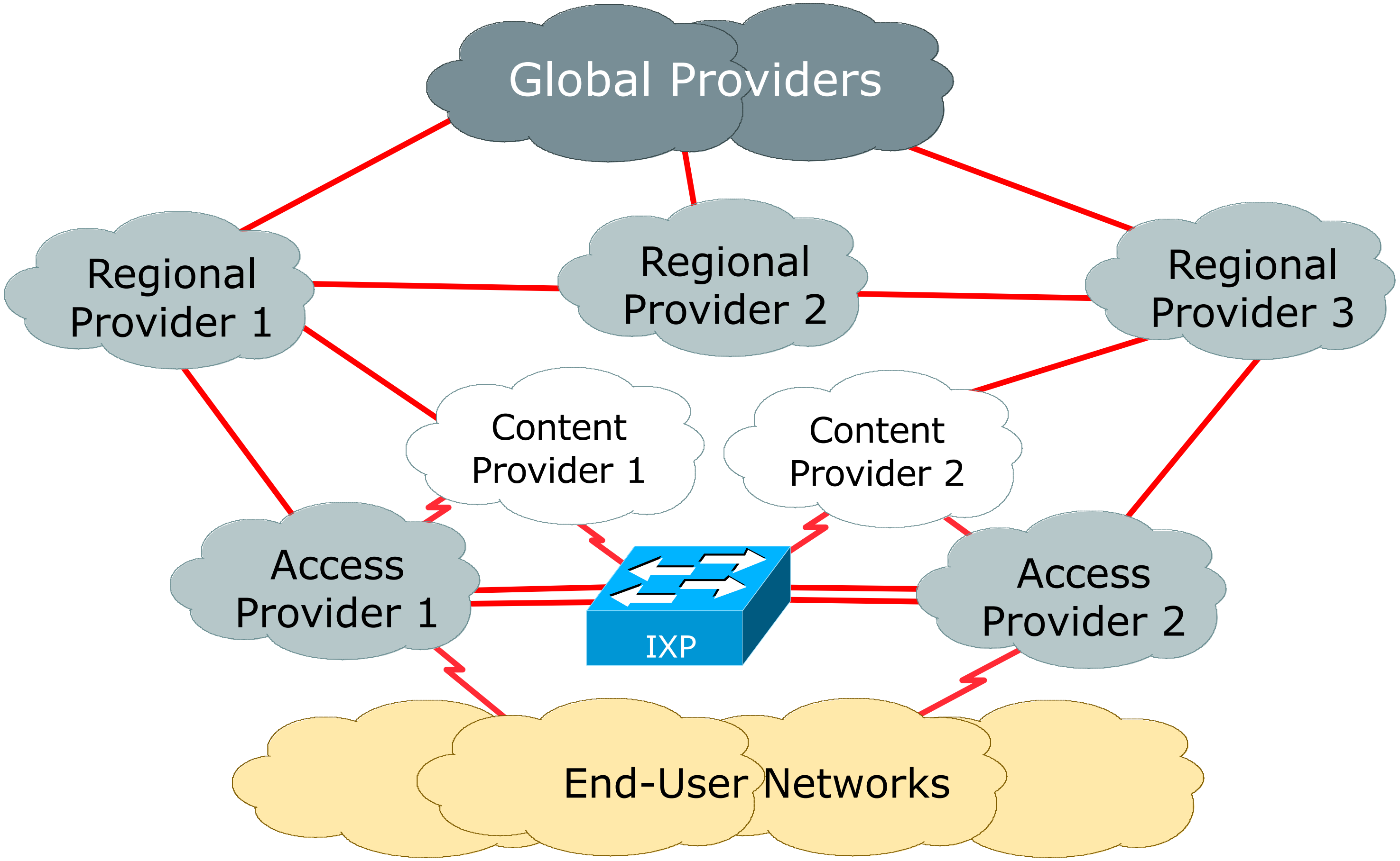
Routing & Interconnection

- Routers
- Routing Protocols
- Peering & Transit
- Internet Hierarchy
- Interconnection Goals

The Internet Today

- Internet is made up of Network Operators of all shapes and sizes
  - Some have local coverage (access providers)
  - Others can provide regional or per country coverage
  - And others are global in scale
- These Operators interconnect their businesses
  - They don’t interconnect with every other Operator (over 75500 distinct autonomous networks) – won’t scale
  - They interconnect according to practical and business needs
- Some Operators provide transit to others
  - They interconnect other Operator networks
  - Just over 10700 autonomous networks provide transit

Global Internet: High Level View



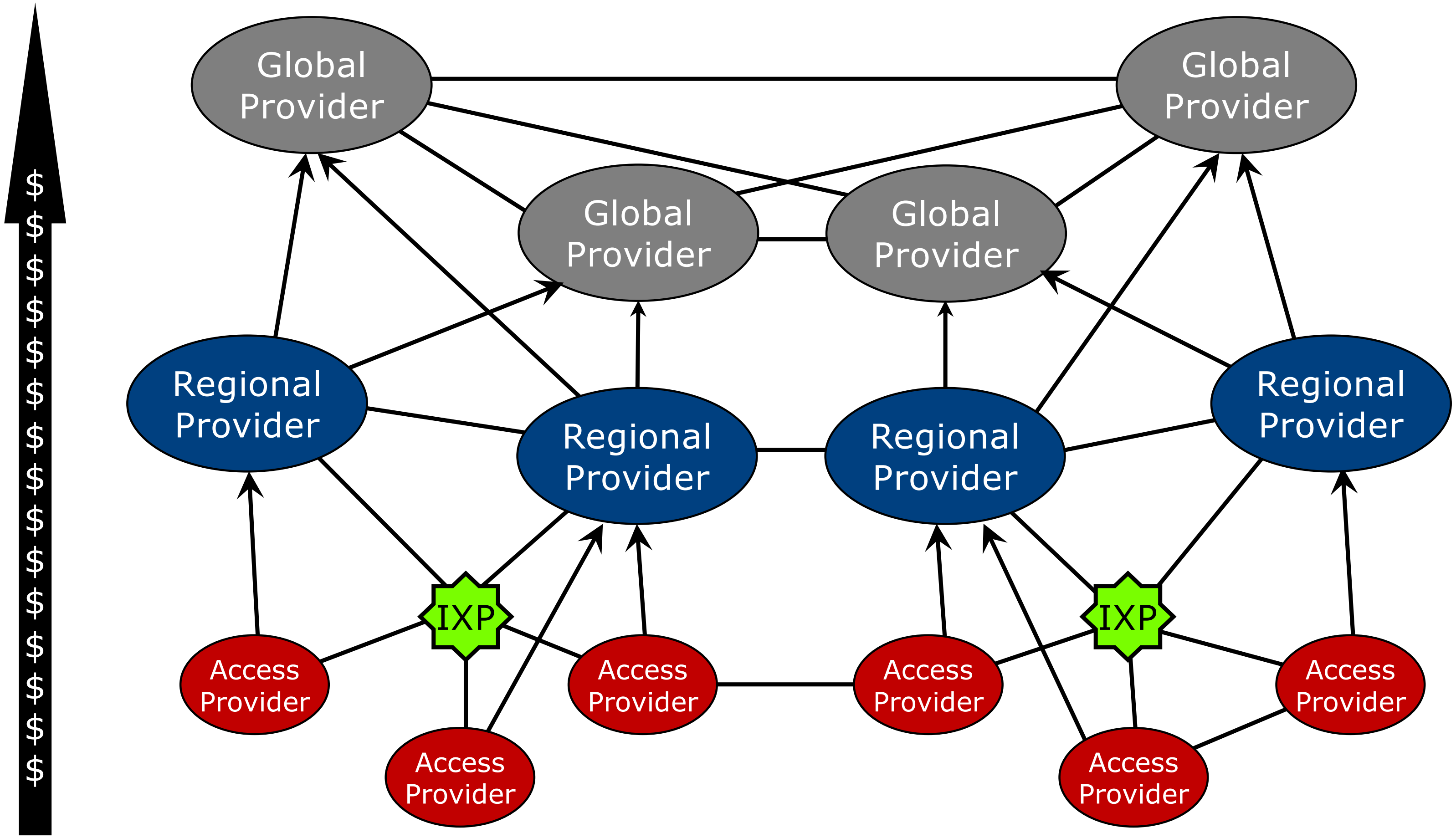
Internet Provider Profile

- Content Providers have moved close to the Access Providers and to Public Interconnects
- Access Providers are simply a vehicle to deliver content as fast as possible to end-user
- Content Providers connect directly with Access Providers
  - PNI – Private Network Interconnect, *or*
  - Across IXPs, *and*
  - Provide a local cache for most frequently used content, *and*
  - Nowadays are building their own global backbones

The Internet Today

- Major content distribution networks no longer have “one big server”
- They each operate a substantial distributed network of content delivery caches from multiple regional datacentres
- Goal:
  - Content as close to the “eyeballs” (the end users) as possible
  - Lowest latency possible
  - Highest bandwidth possible
- The average consumer’s tolerance of non-working websites or delays is only a few seconds

Categorising Network Operators



Routing & Interconnection

- Routers
- Routing Protocols
- Peering & Transit
- Internet Hierarchy
- Interconnection Goals

Network Operator Goals?

- Today, the majority of content and resources consumed by end-users is available by peering:
  - The multi-national content providers (Google, Meta, Amazon etc)
  - The multi-national "cloud" providers (Cloudflare, AWS, etc)
  - Private cross connects
  - Internet Exchange Points
- A network operator's goal is to obtain as much peering as possible
- Transit is the last resort, for any content not available by peering

Network Operator Goals?

- Peering
  - Locally by direct cross-connect with other providers
  - Locally at an Internet Exchange Point
  - Getting to the nearest IXP or other interconnect
- Transit
  - Relying on another network operator to get the rest of the Internet
    - And, in some cases, to the closest/best IXP(es)

Other concerns for Network Operators

- More economies and regions worried about "data sovereignty"
- Data held at multi-national cloud operators could be anywhere on Earth
- Local peering and local interconnects mean:
  - Less "domestic" traffic going crossing national boundaries
  - Greater opportunity for domestic cloud/hosting providers
  - More responsive hosting and "cloud" services
  - More assurance about "data sovereignty"
  - Greater opportunity for creating a vibrant local Internet economy

Network Operator Goals

- Minimise the cost of operating the business
- Transit
  - Have to pay for circuit (international or domestic)
  - Have to pay for data (at $ cost per Mbps)
  - Applicable to each transit service purchased
  - Significant cost of being a service provider
- Peering
  - Shares circuit cost with peer (private) or runs circuit to public peering point (IXP) (one off cost)
  - No need to pay for data
  - Reduces transit data volume, therefore reducing cost

The IXP's role

- Private peering makes sense when there are very few equivalent players
  - Connecting to one other operator costs X
  - Connecting to two other operators costs 2 times X
  - Connecting to three other operators costs 3 times X
  - Etc… (where X is half the circuit cost plus a port cost)
- The more private peers, the greater the cost
- IXP is a more scalable solution to this problem

The IXP's role

- Connecting to an IXP
  - Operator costs: one router port, one circuit, and (maybe) one router to locate at the IXP
  - Some IXPs charge annual maintenance fees
    - The maintenance fee has potential to significantly influence the cost balance for an operator
    - Commercial IXPs charge service fees depending on number of ports consumed and bandwidth connected
- In general, connecting to an IXP and peering there becomes cost effective when there are at least three other peers
  - The real $ amount varies from region to region, IXP to IXP

Peering or Transit?

- How to choose?
- Or do both?
- It comes down to cost of going to an IXP
  - Free peering
  - Paying for transit from an operator co-located in same facility, or perhaps close by
- Or not going to an IXP and paying for the cost of transit directly to an upstream provider and/or IXP
  - Operator has to determine what makes financial & operational sense

Private or Public Peering

- Private peering
  - Scaling issue, with costs, number of providers, and infrastructure provisioning
- Public peering
  - Makes sense the more potential peers there are (more is usually greater than "two")
- Which public peering point?
  - Local Internet Exchange Point: great for local traffic and local peers
  - Regional Internet Exchange Point: great for meeting peers outside the locality, might be cheaper than paying transit to reach the same consumer base

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 operators at a common location than it is to pay transit to another provider to reach the same consumer base
  - "Local" can mean different things in different regions!

Regional Internet Exchange Point

- Regional Internet Exchanges are NOT built from scratch
  - Even today, there are too many "well meaning" attempts to build so-called Regional IXPs
  - There have been several attempts since the 1990s, all have failed
  - Yet there are still entities determined to repeat the same mistakes thinking they can get a better outcome

- Definition: A Regional IXP is a Local IXP that has become so successful that it attracts members from outside its normal service area

Regional Internet Exchange Point

- Regional IXPs are also local IXPs:
  - Regional operators join and peer with each other
  - And show up at several of these Regional IXPs

- Local operators peer with operators from outside the locality
  - They don't compete in each other's markets
  - Local operators don't have to pay transit costs
  - Operators from outside the locality don't have to pay transit costs
  - Quite often operators of disparate sizes and influences will agree to peer – to defray transit costs

Which IXP?

- How many routes are available?
  - What is traffic to & from these destinations, and by how much will it reduce cost of transit?
- What is the cost of co-lo space?
  - If prohibitive or space not available, pointless choosing this IXP
- What is the cost of running a circuit to the location?
  - If prohibitive or competitive with transit costs, pointless choosing this IXP
- What is the cost of remote hands/assistance?
  - If no remote hands, doing maintenance is challenging and potentially costly with a serious outage

Value propositions

- Peering at a local IXP
  - Reduces latency & transit costs for local traffic
  - Improves Internet quality perception
  - Encourages local Internet economy (content, hosting, "cloud" services)
- Participating at a Regional IXP
  - A means of offsetting transit costs
- Managing connection back to home network
- Improving Internet Quality perception for end users

Routing & Interconnection

- Routers
- Routing Protocols
- Peering & Transit
- Internet Hierarchy
- Interconnection Goals

Overview of Routing & Interconnection