

Internet Evolution

ISP/IXP Workshops



These materials are licensed under the Creative Commons Attribution-NonCommercial 4.0 International license (<http://creativecommons.org/licenses/by-nc/4.0/>)

Last updated 18th October 2018

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

Terminology



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
 - Non-commercial, educational use only

Definitions

□ Transit

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

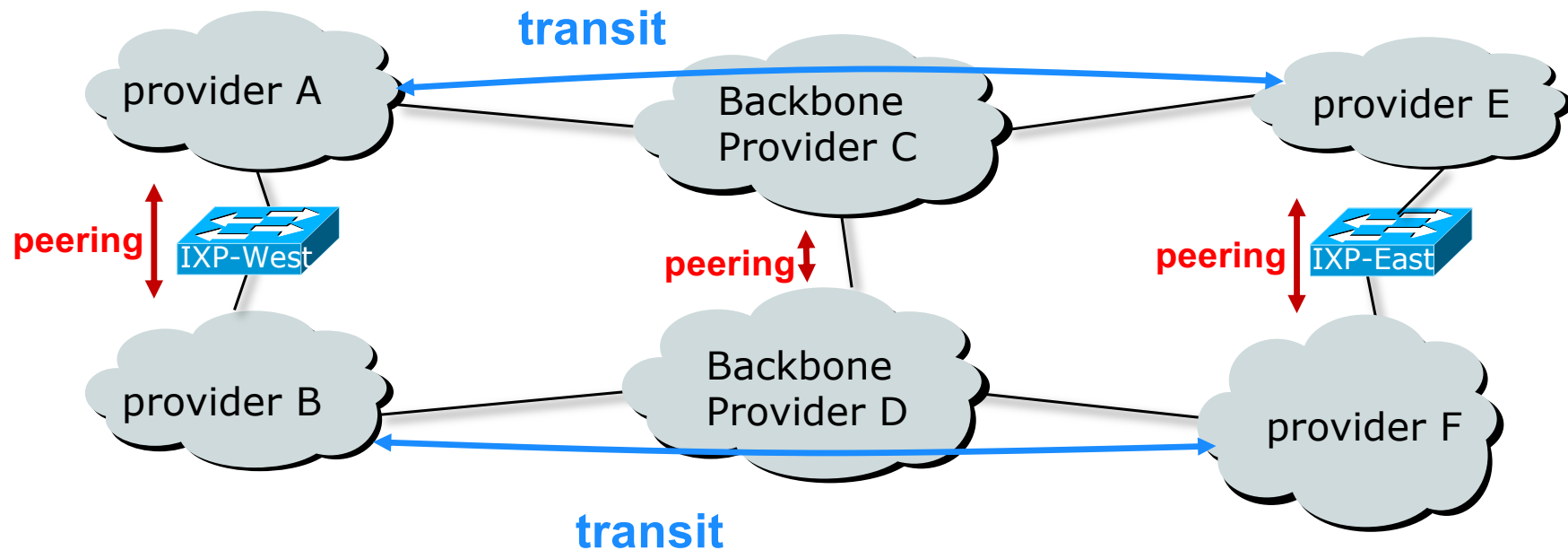
□ Peering

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

□ Default

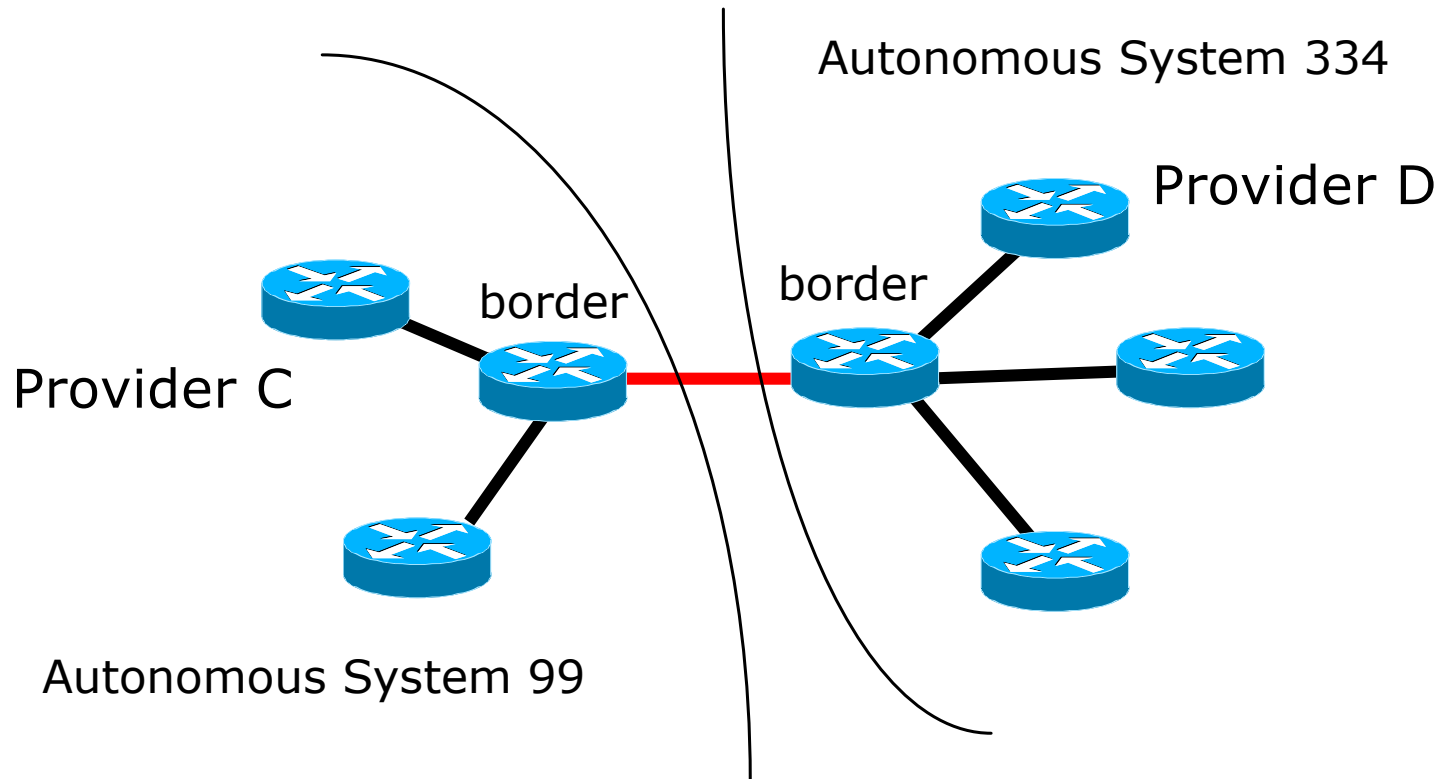
- Where to send traffic when there is no explicit match in the routing table

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



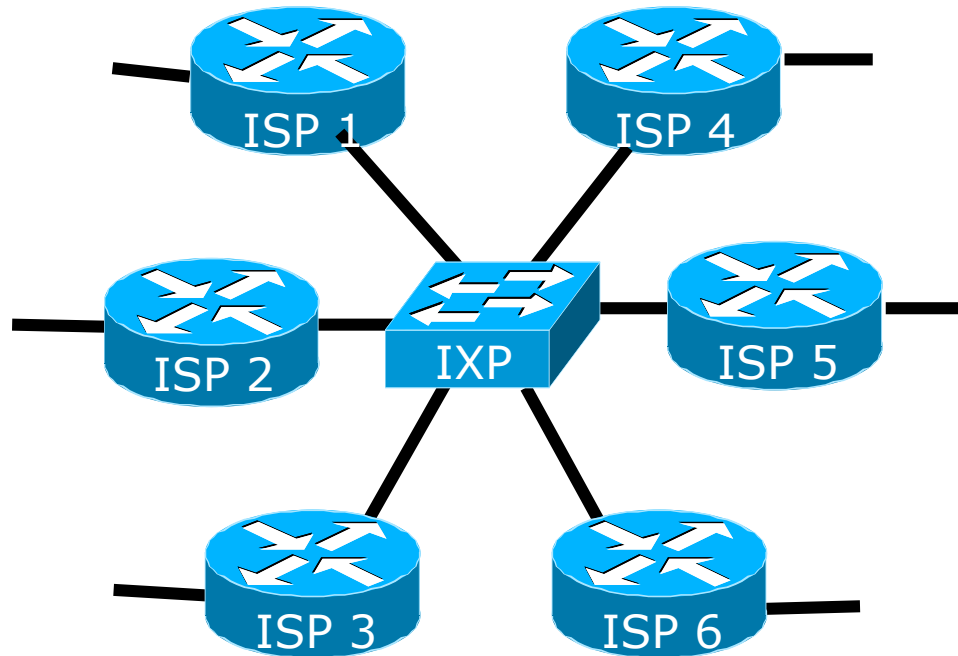
Public Interconnect

- An open and neutral location or facility where several network operators are present and connect to each other over a common shared media
- Why?
 - To save money
 - To reduce latency
 - To improve performance
- IXP – Internet eXchange Point
- NAP – Network Access Point

Public Interconnect

- Centralised (in one facility)
- Larger Interconnects are Distributed (connected via fibre optics) over the local area
- Switched interconnect
 - Ethernet (Layer 2)
 - Technologies such as SRP, FDDI, ATM, Frame Relay, SMDS and even routers have been used in the past
- Each provider establishes **peering** relationship with other providers at the IXP

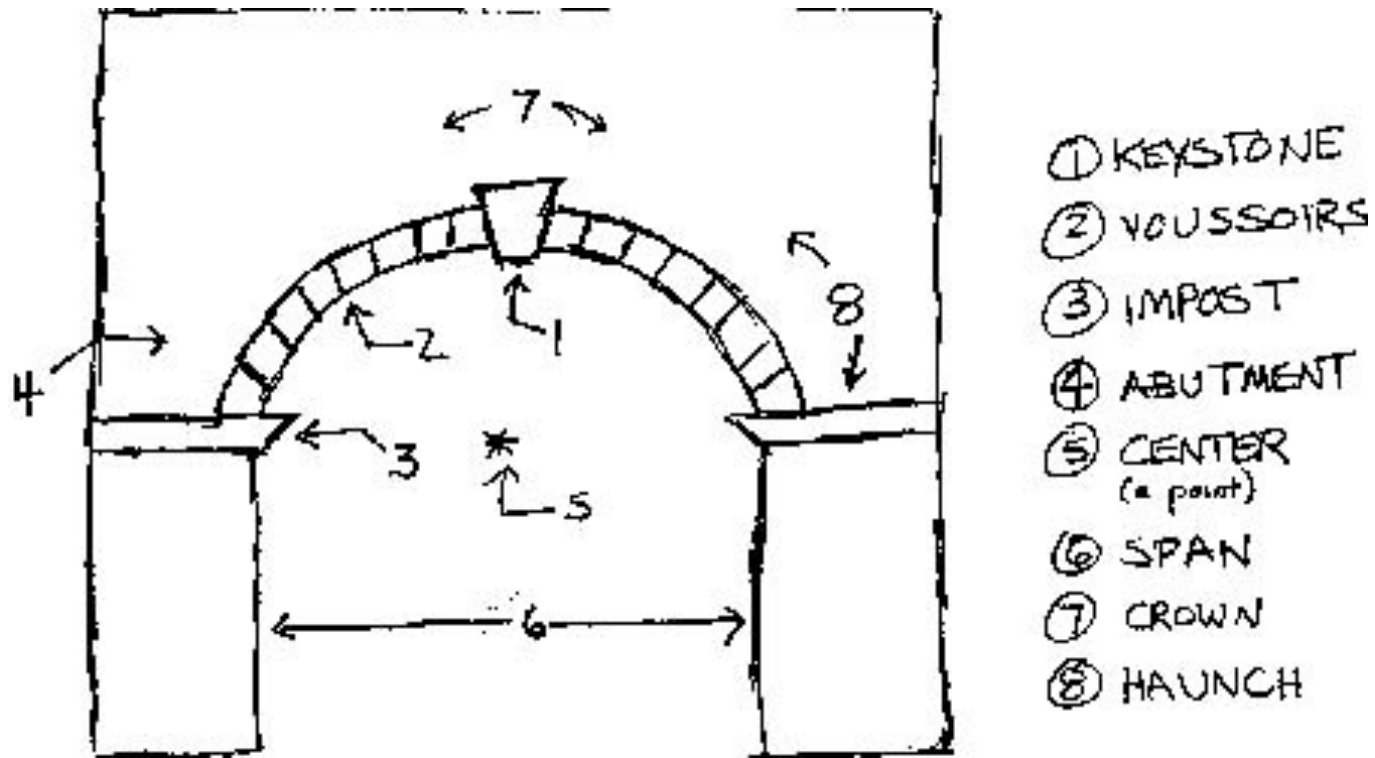
Public Interconnect



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

Public Interconnect

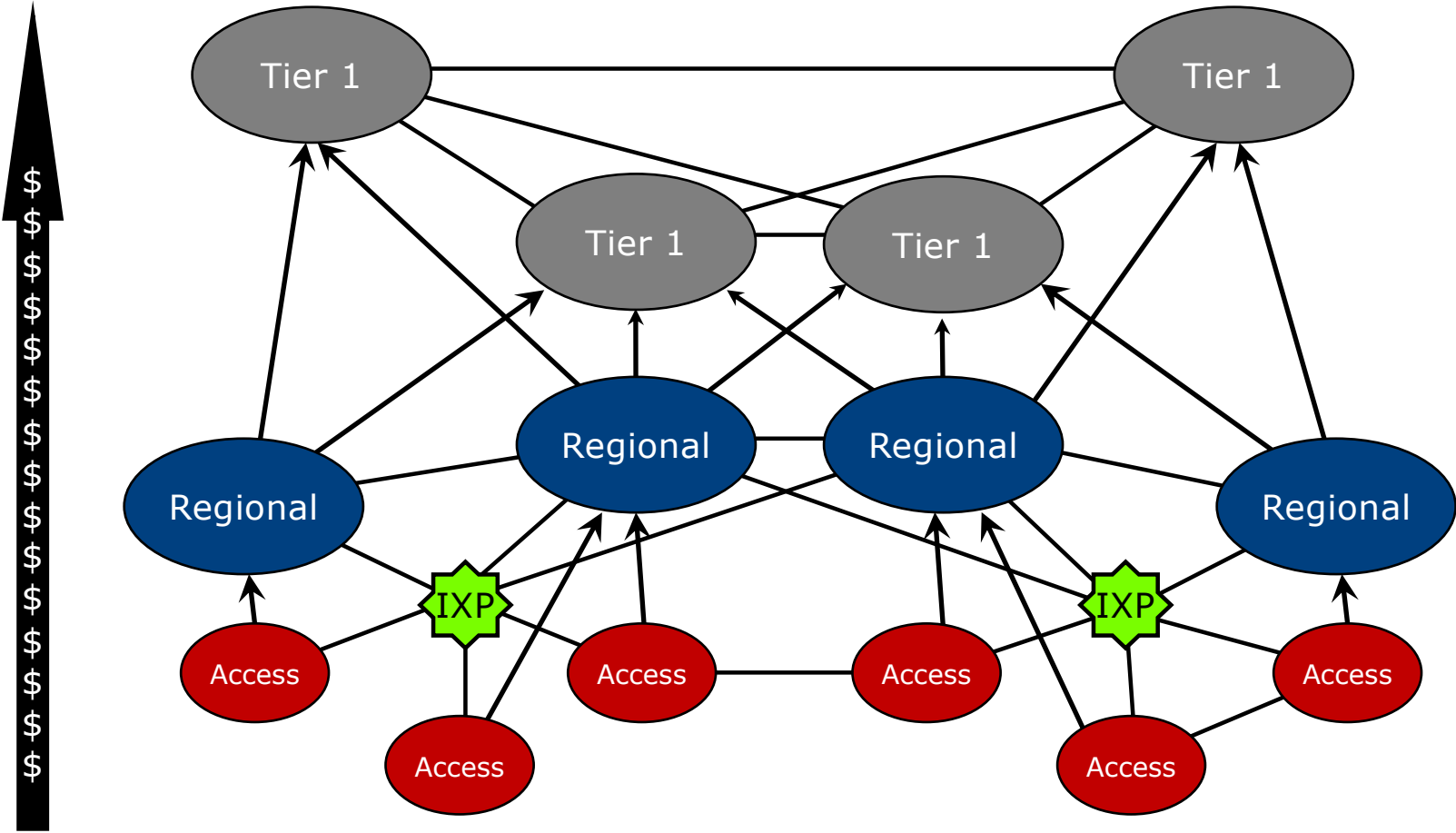
- An IXP is the Keystone of the local Internet Economy



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 61600 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 8400 autonomous networks provide transit

Categorising Network Operators



Categorising Network Operators

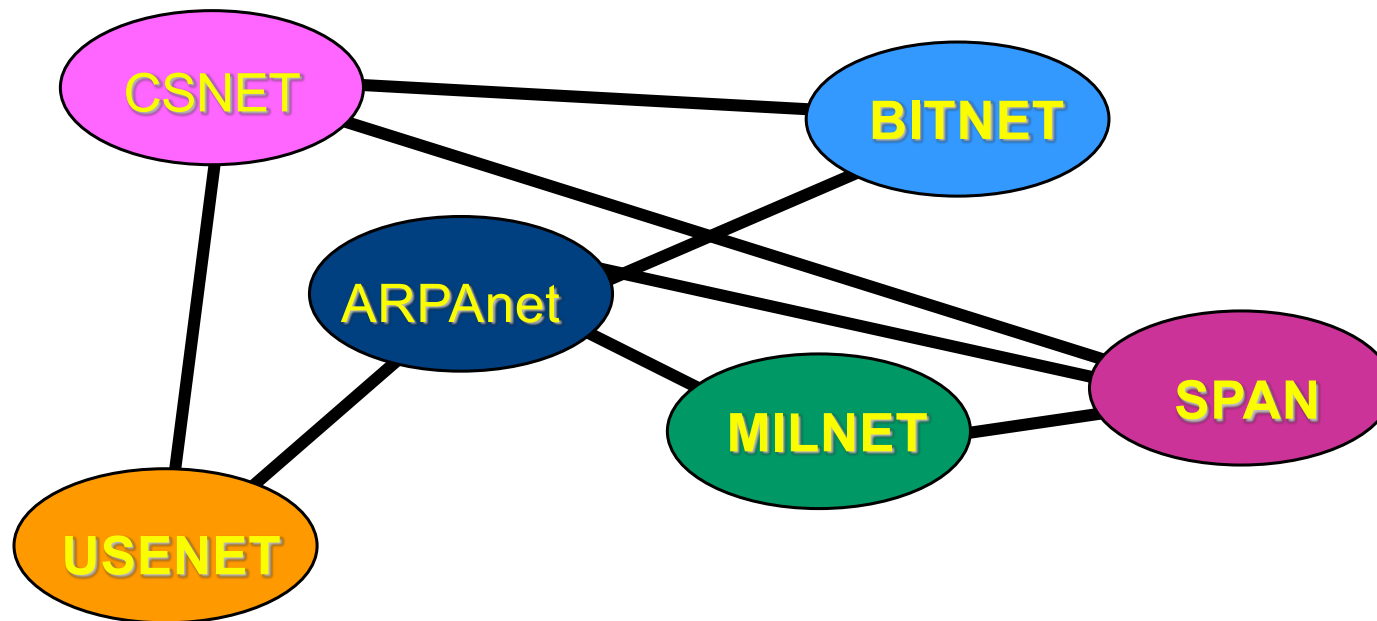
- Tier-1 – definition:
 - A provider which peers with other Tier-1s and does NOT pay for transit
 - Caveat:
 - Many marketing departments call their service provider a Tier-1 – even though that provider may still pay for transit to some parts of the Internet
- Regional providers often have the reach of Tier-1s but still have to rely on maybe one or two Tier-1s to access the whole Internet
 - They often provide access too, via in country domestic access networks
- Access providers work exclusively in their locale

A little bit of History



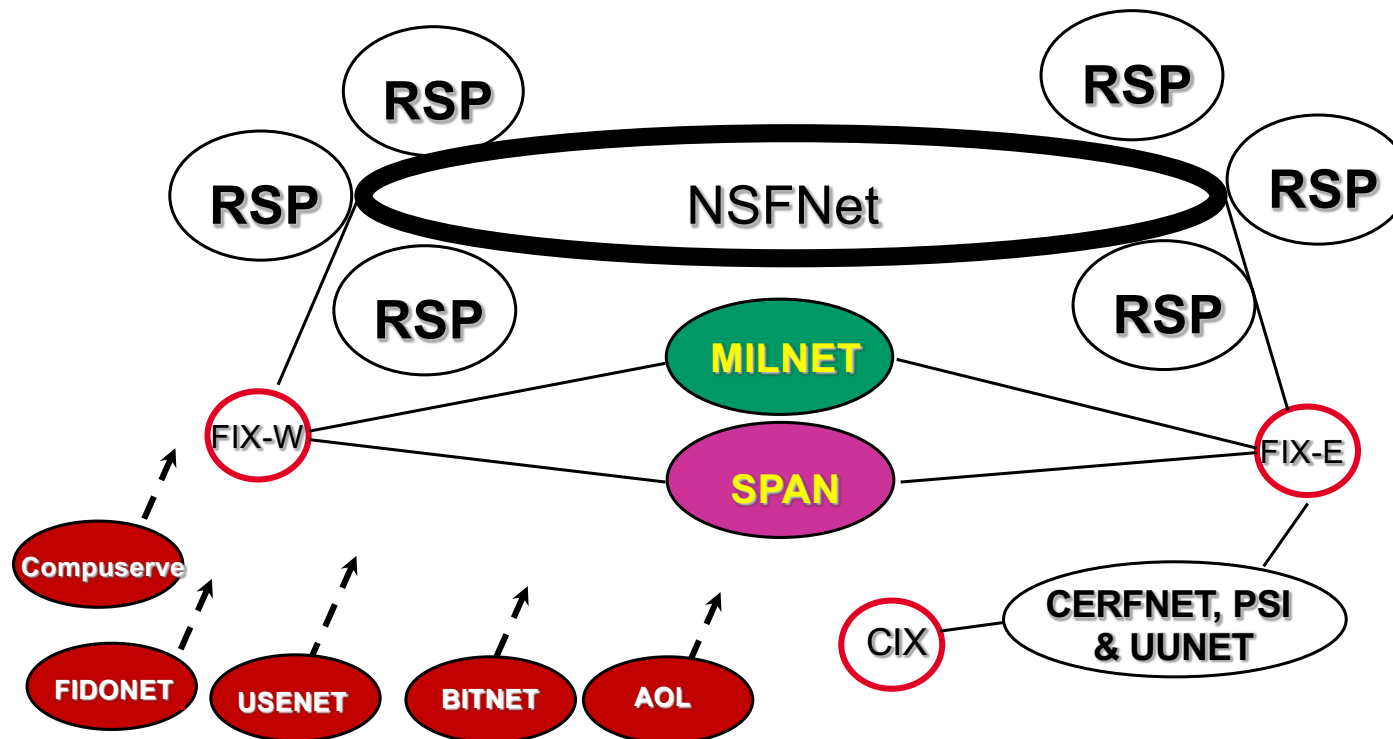
A Bit of History...

- In the beginning, there was no Internet Backbone
 - Operators of the early networks just interconnected..



A Bit of History...

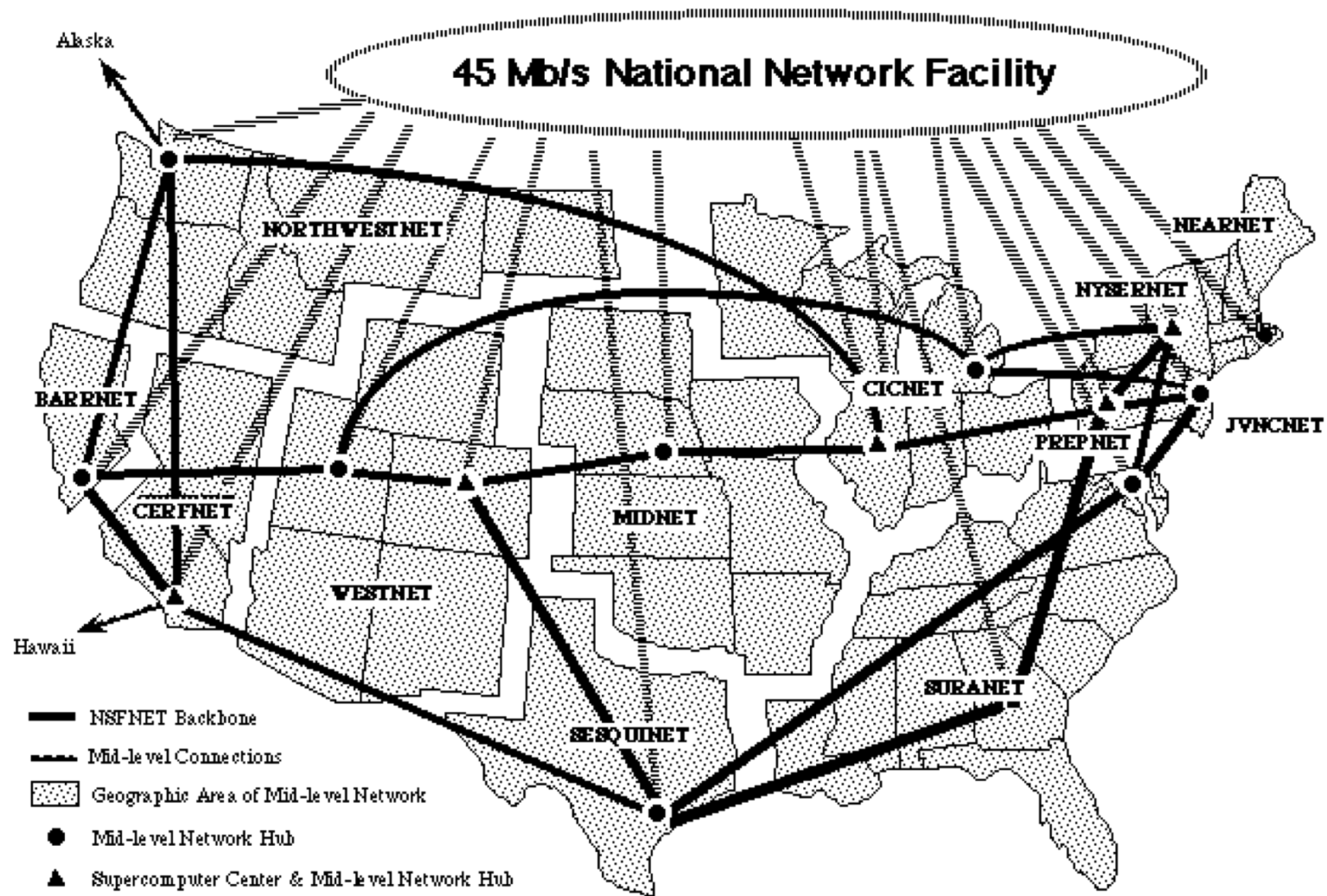
- The NSFNet created the first concept of an Internet Backbone



A Bit of History...

- NSFNet – one major backbone
 - US National Science Foundation funded
 - Connected Universities, Colleges and other educational institutions
 - Connected research laboratories across the US
 - Hosted links to other education and research infrastructure around the world
 - Also connected “private company” networks, under acceptable use policy (AUP), at network access points
 - **AUP: No commercial activity**

The Old NSFNET Backbone



A Bit of History...

- Four Network Access Points (NAPs)
 - Chicago – run by Ameritech
 - New York – run by Sprint
 - San Francisco – run by PacBell
 - Vienna (Virginia) – run by MFS
- These NAPs were the official locations where commercial entities could connect to the NSFNet

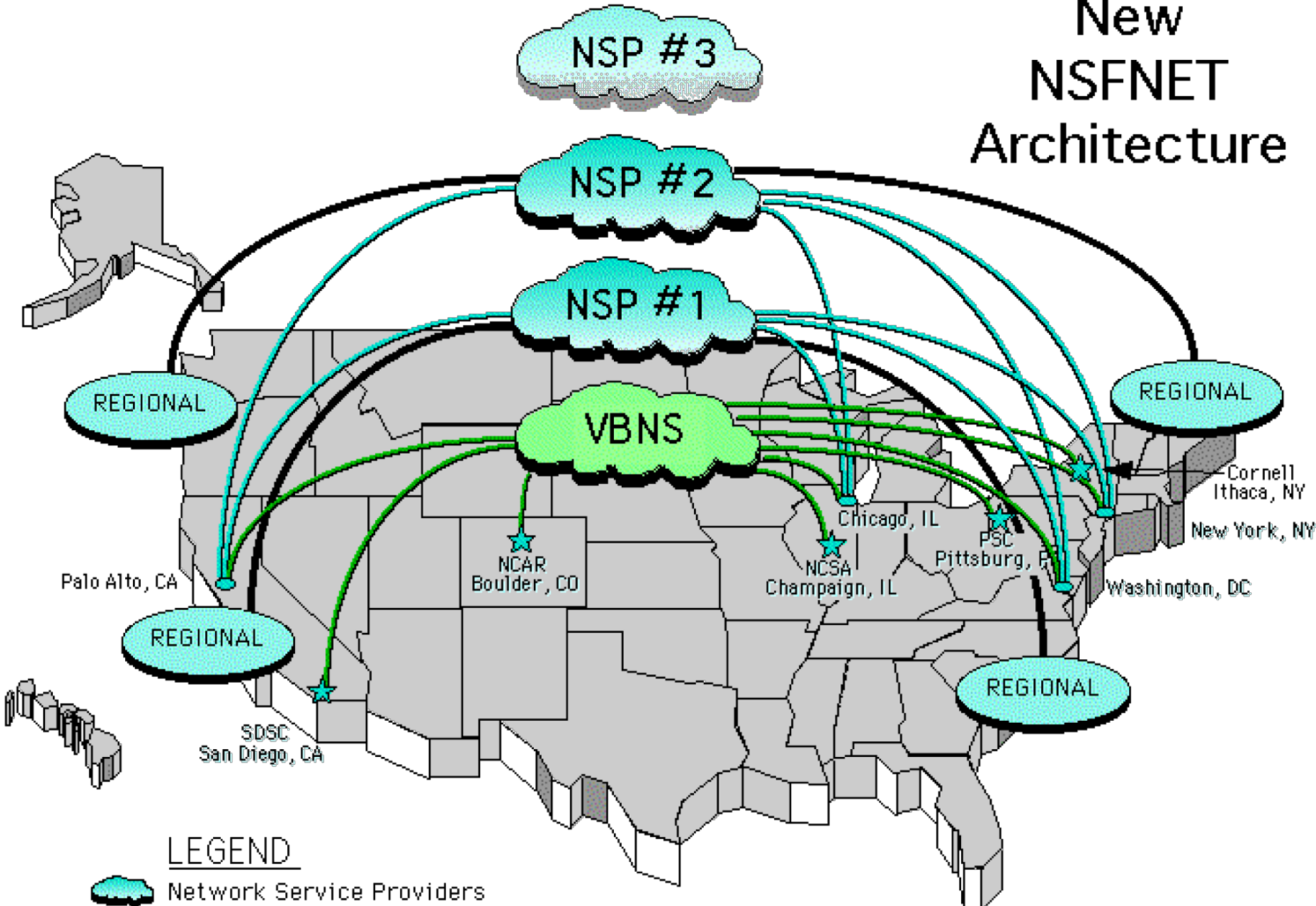
More History...

- Private companies needed to interconnect their networks too
 - Requirement to send “commercial traffic”
 - Could not cross NSFnet due to the AUP
- Resulted in the first “commercial Internet Exchanges” in the early 1990s:
 - CIX-West – west coast USA (San Francisco Bay Area)
 - MAE-East – east coast USA (Falls Church, Virginia)





More History...

- Network Service Providers started providing transit services coast-to-coast across the US
 - An NSP was the ISP for ISPs
- Small / state level network operators couldn't get to the NAPs or other interconnects
 - They bought transit from the NSPs
 - The first NSP was NSFnet – but had an AUP!
- Other NSPs came to prominence:
 - Sprint, UUNET, PSInet, vBNS, etc

New NSFNET Architecture



LEGEND

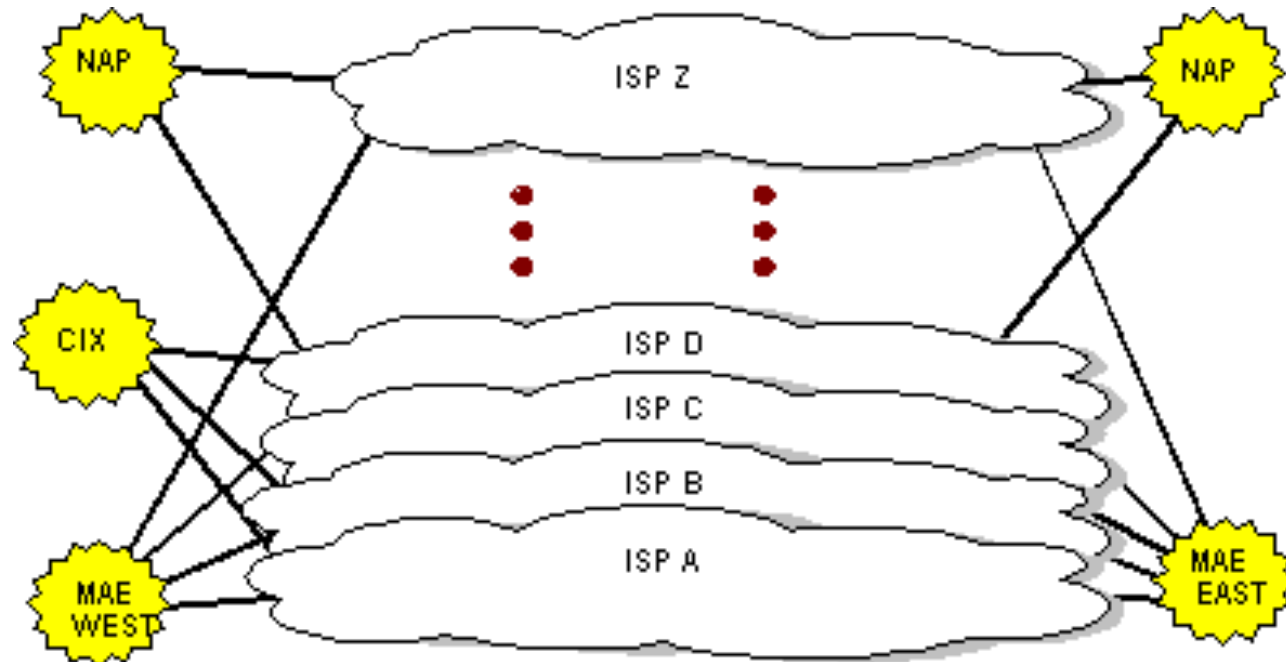
-  Network Service Providers
-  Regional and Midlevel Networks
-  Network Access Points (NAPs)
-  Supercomputer Centers

Note: Supercomputer centers have additional connections for commodity traffic.

Copyright © 1994 General Atomics. NSF Network Plan, NSF Form 1594, 7/94, 1, rev. 6. Reprinted with permission.

More History...

- More interconnects between operators established



Source: WorldCom MAE Services

More History still...

- End of the original NSFnet in 1995:
 - Meant move towards commercial Internet
 - Private companies selling their bandwidth
- The NAPs established late in NSFnet life were some of the original “exchange points”
 - NAP operators were providing commercial Internet access as well
 - Sprint, PacBell and Ameritech NAPs were replaced by neutral/commercial IXPs
 - The MFS hosted MAE-East replaced the Vienna NAP
 - ANS (operator of the late NSFnet) forced to join IXes

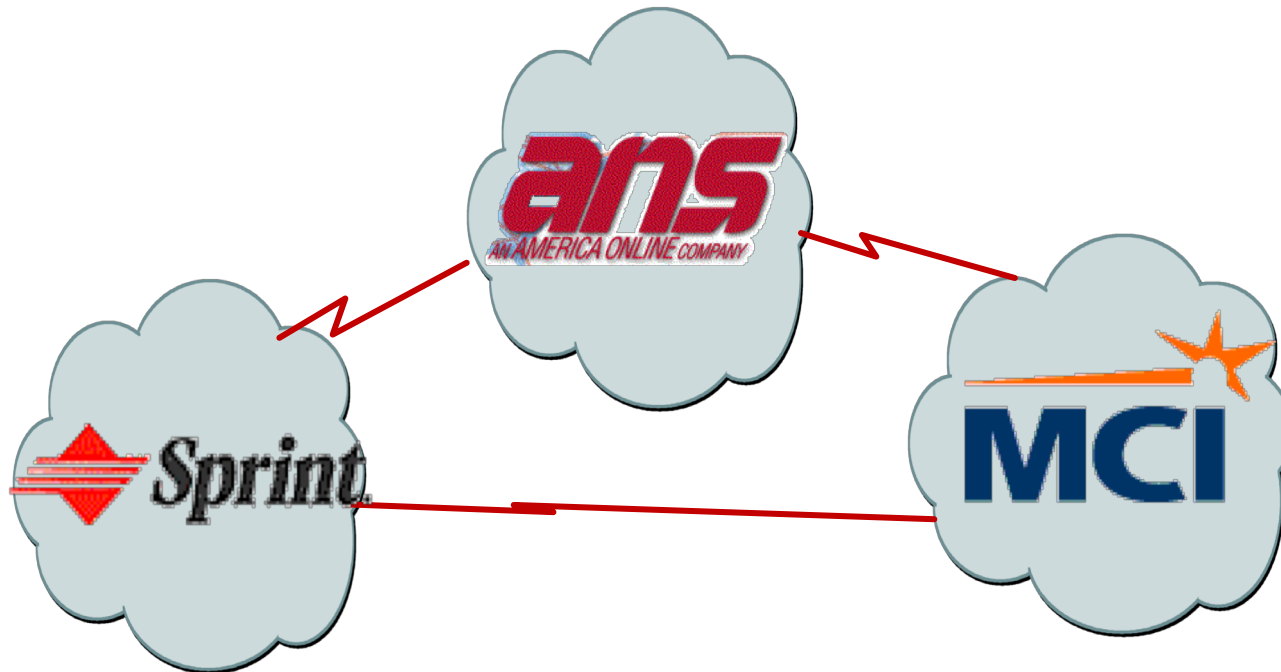
Internet in the 1990s

- By mid-1990s, Internet model looked like this:
 - Very much US centric
 - NSPs provided transit coast-to-coast across the US

- NSPs of the mid-1990s became known as Tier-1s
 - Tier-1 is a network operator who has no need to buy transit from any other operator
 - Interconnect with other Tier-1s by Private Interconnect

Tier-1 Private Interconnects

- "ANS, MCI and Sprint Sign Agreements for Direct Exchange of Internet Traffic" – June 30, 1995



Internet in the 1990s

- For network operators in the 1990s, connecting to the Internet meant:
 - Connecting to one or more US operators for transit
 - Connecting to one of the US IXPs
 - Expensive connections across big oceans (Atlantic, Pacific)

Europe: early 2000s

- European Internet had developed
 - European Commission had removed the trade barriers imposed for cross-border telecommunications between EU member states
 - Prior to 1995, capacity from London to the US was cheaper than the same capacity from London to Paris, or Paris to Frankfurt
 - Allowed growth of early European backbones (Ebone, PIPEX International, EUnet)
 - No longer US hub centric
 - US operators expanded their backbone infrastructure into Europe
 - European infrastructure acquisitions or joint ventures by UUNET, PSInet, Qwest and AT&T

Europe: early 2000s

□ Interconnects!

- Network operators in Europe interconnected at IXPs such as LINX, AMS-IX, DE-CIX etc
- Most countries had at least one IXP

□ Devolution of content distribution

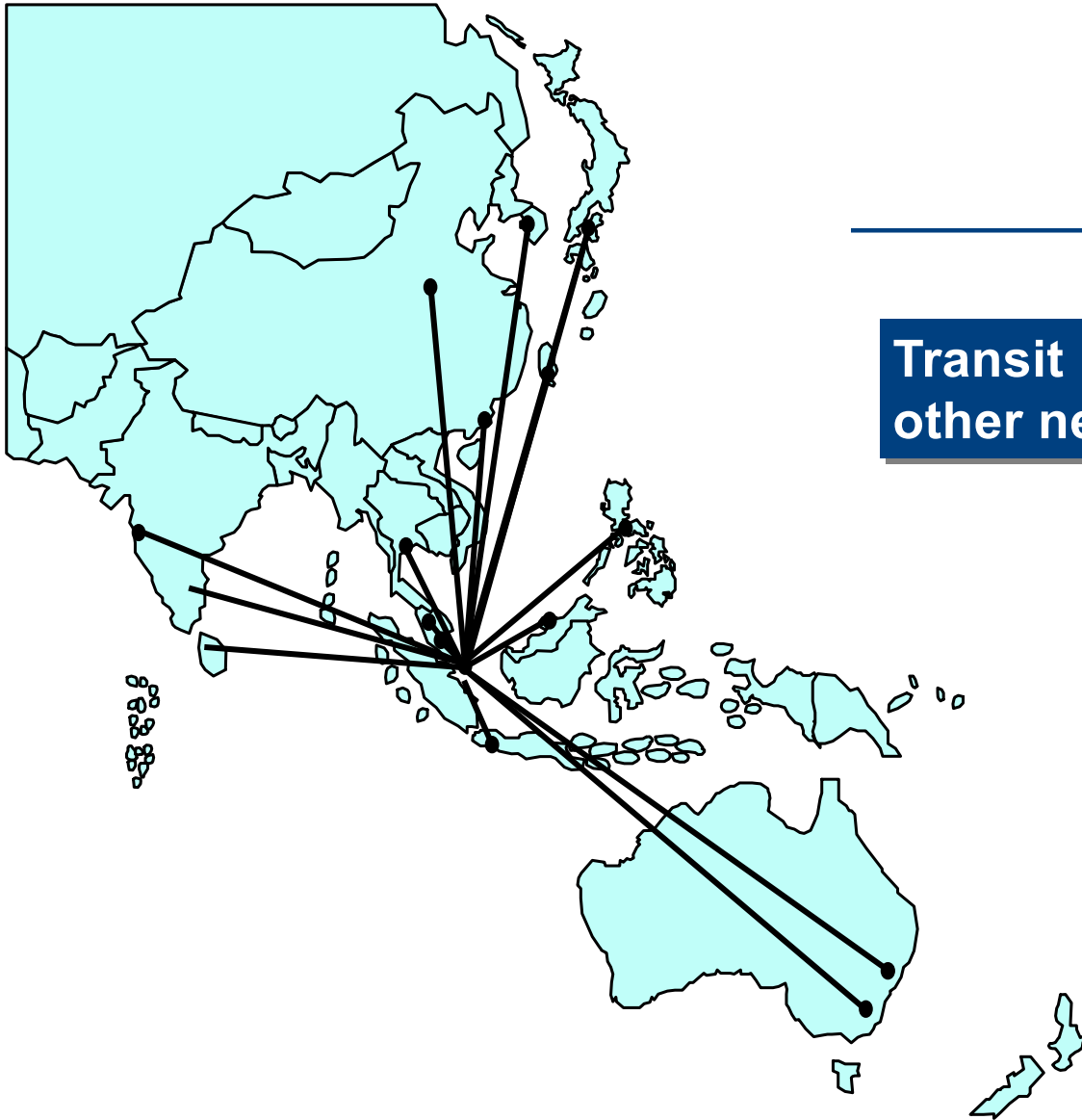
- The news media (eg CNN and BBC) starting to put news and programming onto the Internet
- Microsoft Network (MSN) delivering content from locations other than HQ in Redmond (Seattle), US

Asia & Pacific early 2000s

- Asia & Pacific Internet started to develop
 - Still dependent on US hub though
 - Australia to SE Asia traffic tended to use low cost path via US
 - Intra-SE Asia connectivity tended to be via US
 - Large geographical region more challenging and costly to cover
 - Satellite dominated in South Asia and the Pacific
 - Public interconnects developed only in Japan and Hong Kong
 - No concept of interconnection between country networks
 - Much talk of Regional Hubs

Regional Hub

**Transit Provider interconnecting
other network operators in the region**

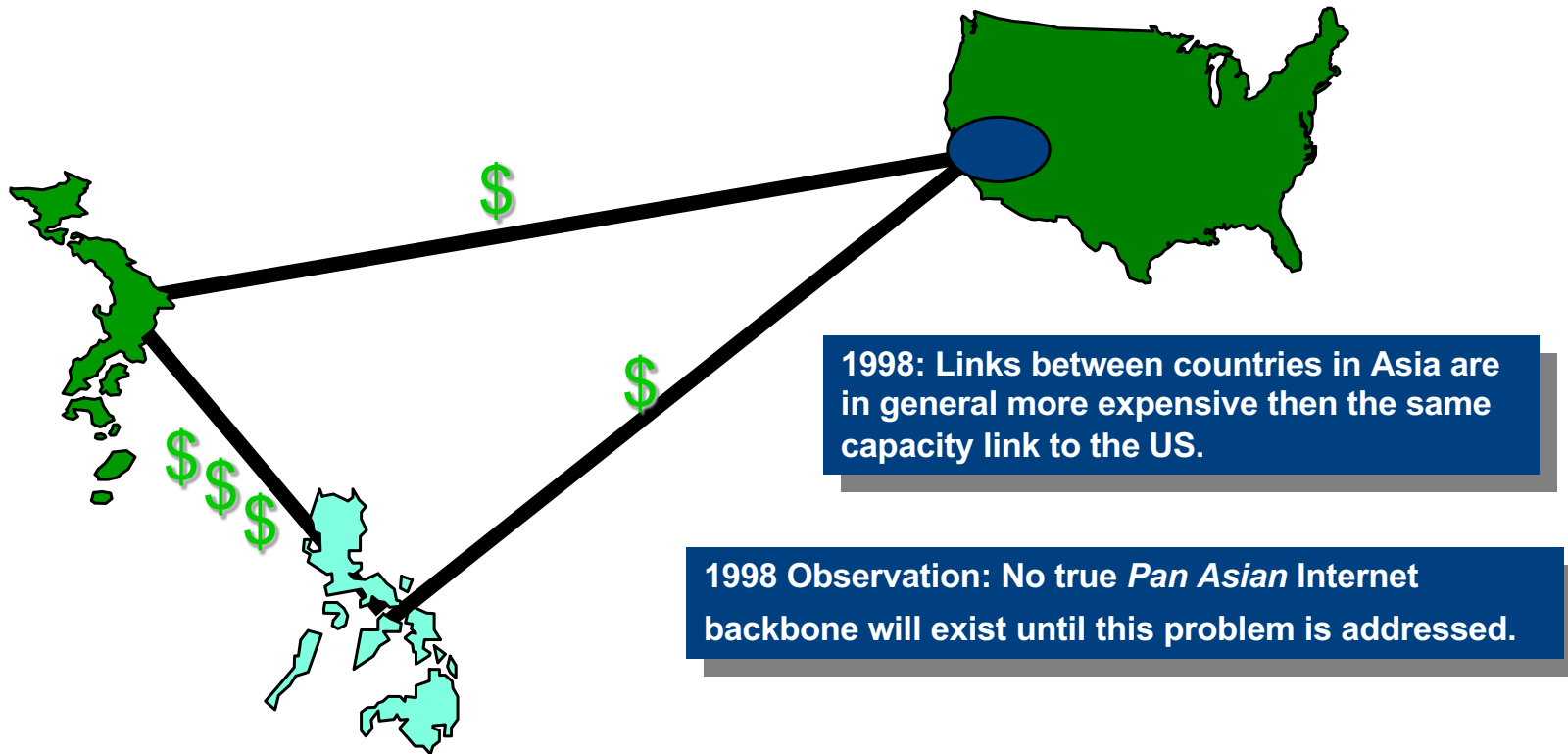


Asia & Pacific early 2000s

- Three factors inhibited growth of Asia & Pacific Internet interconnections during the late 1990s
 - Price:
 - International Private Leased Circuits (IPLC) between Asian and Pacific countries was much higher than the equivalent circuit to the US
 - Regional Rivalries:
 - Everyone wanted to be the hub
 - Multitude of Cultures:
 - Mandarin speaker will not be browsing Hindi content – and vice-versa

Private Interconnects in Asia: early 2000s

- Asian ISPs use the US West Coast as the hub because it is more cost effective – despite the performance impact of crossing the Pacific Ocean twice!

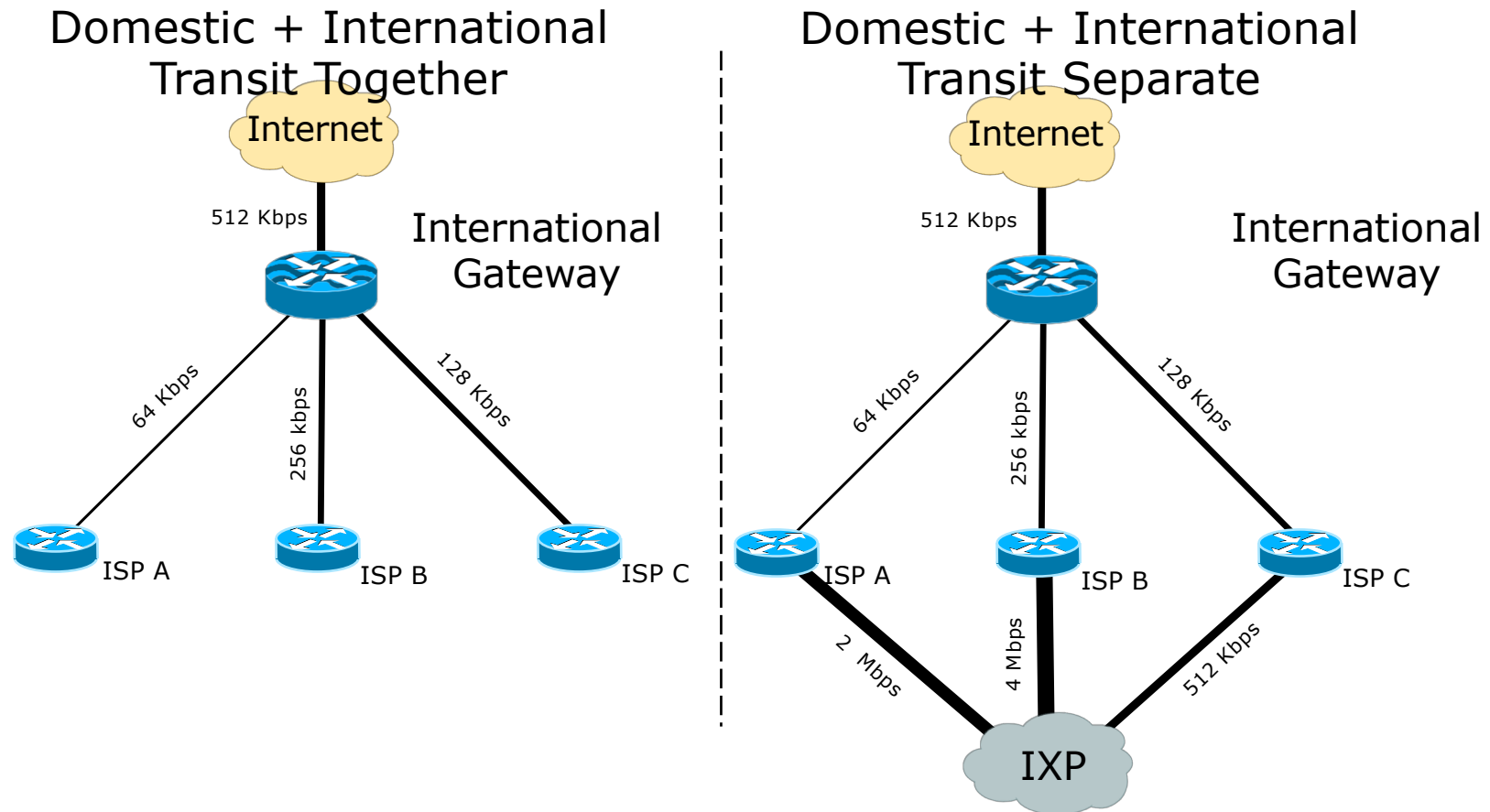


National Internet Gateways

- Unlike in North America & Europe, National Internet Gateways were established in many countries in Asia and in the Pacific
 - Not free neutral interconnects like in Europe or US
 - For profit transit to the Internet

- Many countries mandated that the National Internet Gateway operator also had to operate an “IX”
 - The idea was to keep local traffic local
 - Although this IX was only for the IG’s customers
 - Traffic was charged (as part of the overall service)

National Internet Gateway Models



National Internet Gateways

- Some countries established several National Internet Gateways
 - Regulatory desire to have a Competitive Internet Gateway market
- Advantage:
 - Encouraged several operators to apply for the licence to sell Internet transit to other operators
- Disadvantage:
 - To access all Internet content in one country, operators now had to connect to all National Internet Gateways

National Internet Gateways

- Compared with Europe and North America, this restricted the growth of the Internet in Asia and in the Pacific
- Many issues:
 - Greater expense for traffic exchange
 - Limited interconnect bandwidths
 - Poorer quality of service
 - No incentive to host any content or services locally – content provider had to connect to all IIGs!
- Still a big challenge in many countries today

Content in the 1990s?

□ Popular Content & Activities:

- FTP sites
- Usenet News groups
- Education archives (usually University or National Libraries)
- Bulletin Boards
- Internet Relay Chat (IRC)

□ Search engines:

- Gopher was popular before ubiquitous web browsing in 1996
- Altavista
- Google became the go-to search engine by 2000

Content in the 1990s

- Early content was hosted at the site that created it.
- Examples:
 - BBC News website hosted by the BBC in London
 - Users browsing the website connected to the server in the UK
 - CNN.com hosted in the US by CNN
 - Google search engine hosted in the US by Google
 - Etc
- Content distribution was centralised

Content delivery scaling

- Operators in late 1990s and early 2000s wishing to scale their network infrastructure
 - User experience starting to matter
 - Bigger pipes and faster speeds was fine for the operator network
 - But with content not hosted locally, many dependencies for delivering quality for the end user
 - “Internet Broken” is the operator’s problem, regardless of where the problem really is

Content delivery scaling

- Recognition in the late 1990s that content delivery had to move to the access edge
 - Not entirely clear how to do this at that time
 - Huge growth of Google, Facebook, YouTube etc had not yet taken off
- Usenet News still had relatively useful content
 - Large volumes of content every day
 - Network operators had deployed Usenet News distribution infrastructure in their access networks
 - The precursor to the Content Distribution Networks we see today

Today



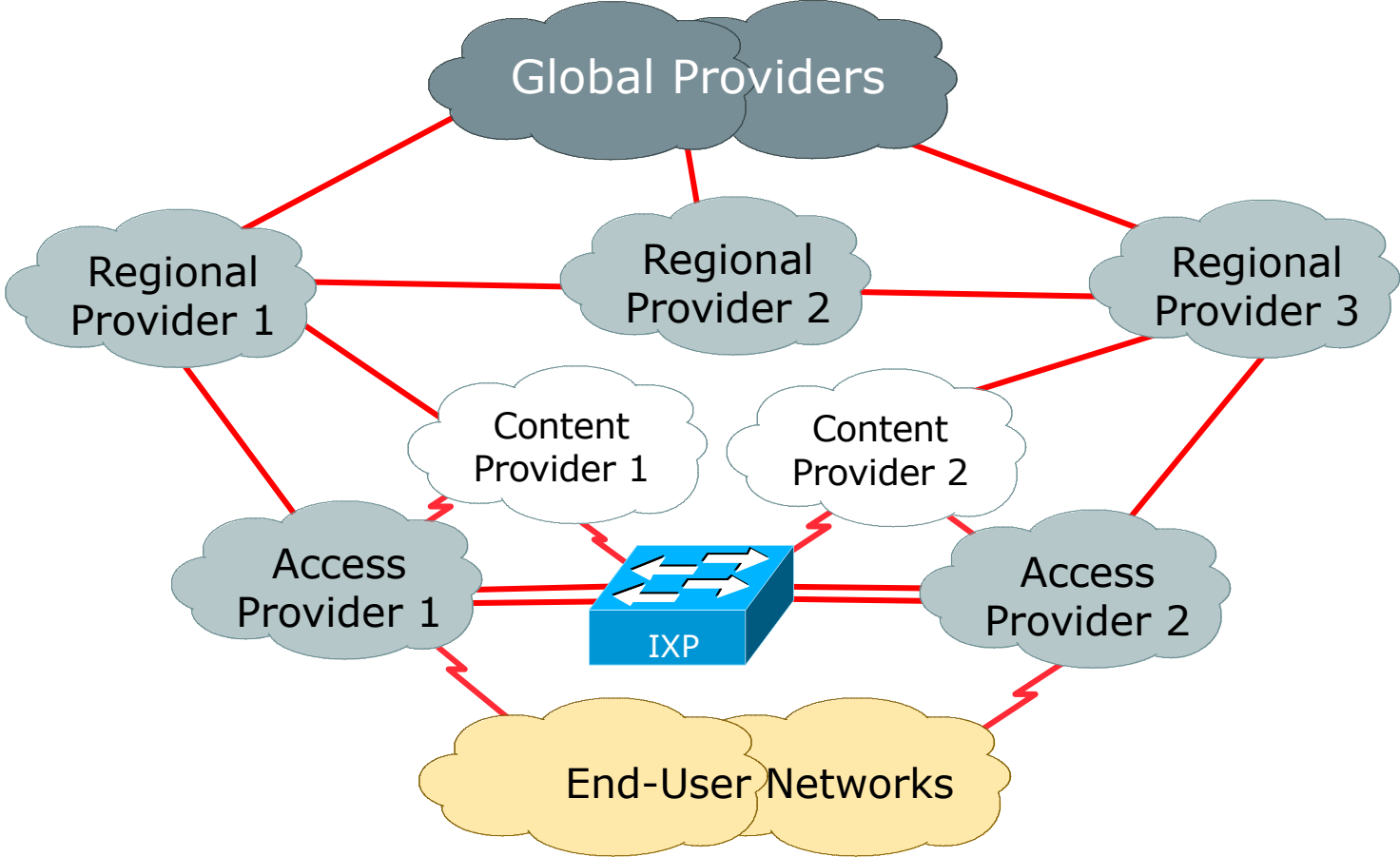
The Internet Today

- “Content is King”
- The typical end-user traffic profile shows:
 - 50% of all Internet traffic is Google/YouTube
 - 25% of all Internet traffic is Facebook
 - 10% of all Internet traffic is Content hosted by Akamai, Cloudflare, Netflix, Microsoft, and other content operators
 - (“typical” in this author’s experience)
- This is a significant change over the traffic profile from the late 1990s and early 2000s

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

Global Internet: High Level View

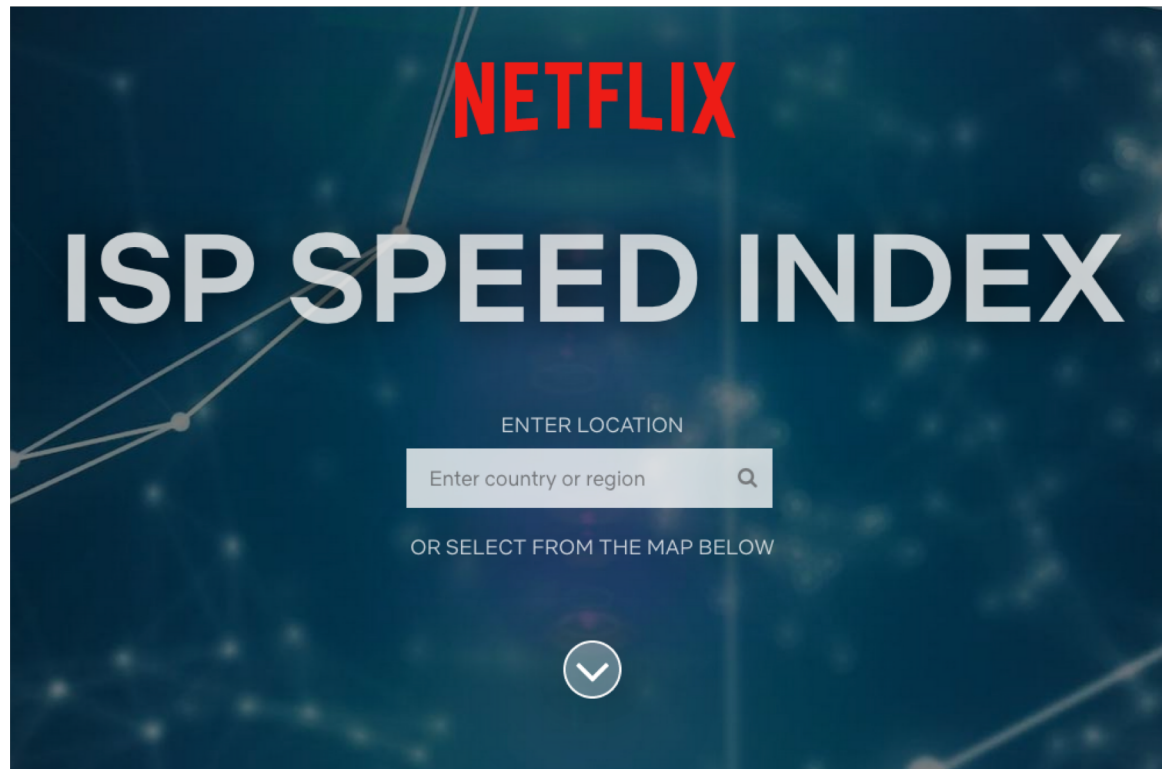


Internet Provider Profile

- 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 directly connect with Access Providers
 - PNI – Private Network Interconnect, or
 - Across IXPs, and
 - Provide a local cache for most frequently used content

Content delivery is competitive!

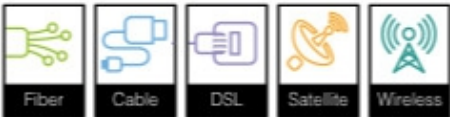
- Competition in local marketplace is all about speed and quality of content delivery
 - e.g.



AUSTRALIA

ISP LEADERBOARD - JULY 2018

RANK	ISP	SPEED Mbps	PREVIOUS Mbps	RANK CHANGE	TYPE				
					Fiber	Cable	DSL	Satellite	Wireless
1	Telstra	3.92	3.86		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	Optus	3.76	3.71		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	Exetel	3.74	3.70		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
4	iiNet	3.68	3.64		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	Dodo/iPrimus	3.65	3.61		<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
6	TPG	3.63	3.59		<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



What happened?

□ In the late 1990s:

- US was hub of global Internet
- Europe was becoming a hub of the European Internet
- Asia, Pacific, Latin America still mostly connected to the US, rather than interconnected within region
- Africa mostly connected to Europe, rather than interconnected within region

- Internet access was by desktop or, more rarely, laptop computer
 - Content by static web pages, UseNet, some news media
- No smartphones or tablets or 3G or LTE

What happened?

- Apple iPhone launch in January 2007
 - Availability of 3G networks
 - Smartphones took off
 - Google's Android quick to follow
- Dominance of Google as search engine
- Dominance of Facebook for social networking

- By 2010, users could be online 24x7 through their increasingly smarter and more data-hungry devices

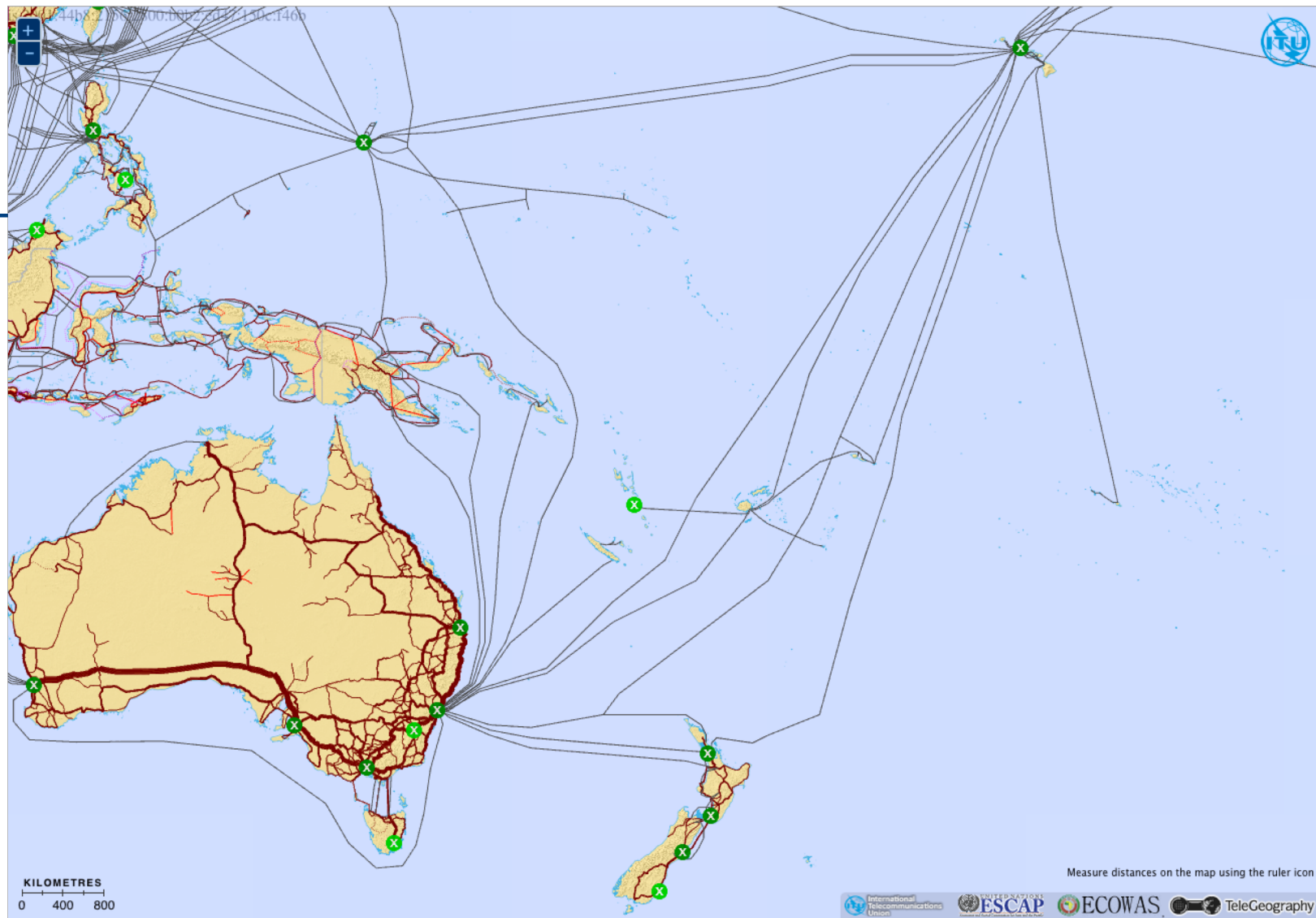
Asia in the 2000s

- Emergence of Singapore as regional hub to complement Hong Kong and Japan
 - Fibre cuts caused by the Taiwanese earthquake of December 2006 forced many Asian network operators to reconsider “US hub / go East” model
 - Singapore is now the interconnect for almost all South East and South Asian network operators
 - (The next regional interconnect heading west is in France!)

Pacific in the 2010s

- Sydney has emerged as the hub for the South Pacific
 - Southern Cross Cable to US via Auckland, Fiji and Hawaii created opportunities
 - Sydney to Guam fibre giving access to Japan and SE Asia
 - Papua New Guinea to Sydney fibre
 - New Caledonia to Sydney fibre
 - Vanuatu to Fiji fibre
 - No break out in Fiji means capacity from Vanuatu direct to Sydney on Southern Cross Cable
 - Tonga to Fiji fibre
 - No break out in Fiji means capacity from Tonga direct to Sydney on Southern Cross Cable

Pacific Fibre Map

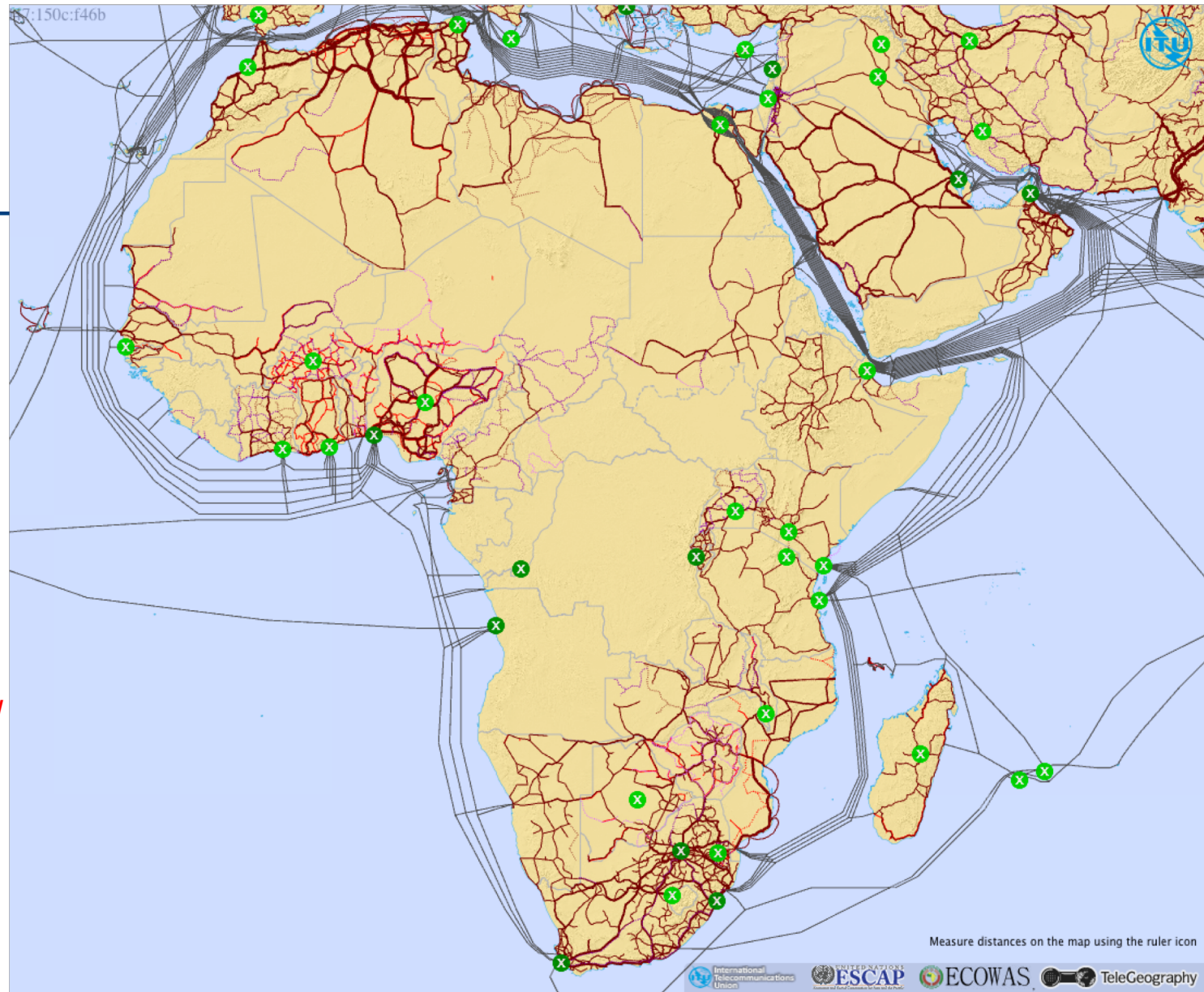


Source: <https://www.itu.int/itu-d/tnd-map-public/>

Africa in the 2010s

- With the new East Africa cable, operators like SEACOM and Liquid Telecom flourished
 - Before then, Internet was universally expensive and low bandwidth via national telecom operators to France or UK
 - (That's where the fibre went)
 - Regional fibre infrastructure in East Africa has caused rapid development for much of the region
 - Now viable for content distribution networks to look at locating on the continent, rather than feeding from Europe
 - Example:
 - <https://www.internetsociety.org/news/press-releases/2018/internet-society-partners-with-facebook-to-expand-internet-connectivity-in-africa/>

Africa Fibre Map



Source:
<https://www.itu.int/itu-d/tnd-map-public/>

What is a Content Cache: Network Operator

- ❑ CDN provides a device (usually a server or cluster of servers rackmounted) which stores content frequently requested by end users
- ❑ The device is hosted in the core of the network operator's infrastructure
- ❑ The network operator announces to the cache the address space to be served by the cache
 - Often announce the address space of customer operators and even peers too
 - The more address space announced to the cache, the greater the number of “eyeballs”, the more efficient the cache becomes

What is a Content Cache: End-User

- The first request from end user for content is downloaded over international transit link directly from the CDN provider's main infrastructure
 - Served to end user
 - Stored in content cache
- The next request to the CDN provider for the same content is redirected to the local cache
 - Fast response for end user
 - Minimal use of the network operator's international transit link (only for initial request and control traffic)

Content Distribution Today

- ❑ CDNs such as Google, Facebook, Cloudflare and Akamai have built considerable content distribution infrastructure
- ❑ Several have large stake holdings in global submarine fibre
 - Example: <https://www.wired.co.uk/article/google-facebook-plcn-internet-cable>
- ❑ Several have built their own large data centres at strategic locations around the globe
- ❑ This has all supplanted the Tier-1 operator as the content delivery vehicle to the regions around the globe
- ❑ The CDNs encourage operators to connect to their datacentres to maximise performance for content delivery

Content Distribution Today

- ❑ CDNs such as Google, Facebook and Akamai also supply and operate content caches
- ❑ Operators with a few Gbps of content being served from these CDNs usually qualify for a cache
- ❑ Caches are found in most larger operators today

- ❑ Many IXPs have CDNs present
- ❑ Many operators at smaller IXPs will share their content caches with their peers across the fabric

Content Distribution Today

- CDNs at IXPs:
 - Lowest possible latency between the content and the end-user
 - Highest possible bandwidth between the content and the end-user

 - Which means happy end-user!

 - Which means end-user keeps connected to the CDN operator, rather than moving to a competitor

 - Onus on network operator to maintain high capacity at IXP and on to end-user
 - International connectivity is usually much more expensive!

Content Distribution Today

- Not every operator qualifies for a content cache
 - The CDNs usually require a minimum of 5Gbps of traffic to subscribers of the network operator before they will provision a cache
 - This is **not** about being unfair to smaller operators!
 - Content caches, in the experience of the CDN operators, only show effectiveness when end-user traffic volume is around 5Gbps
 - Lower traffic volumes result in poor cache hits and minimal savings for the network operator

Content Distribution Today

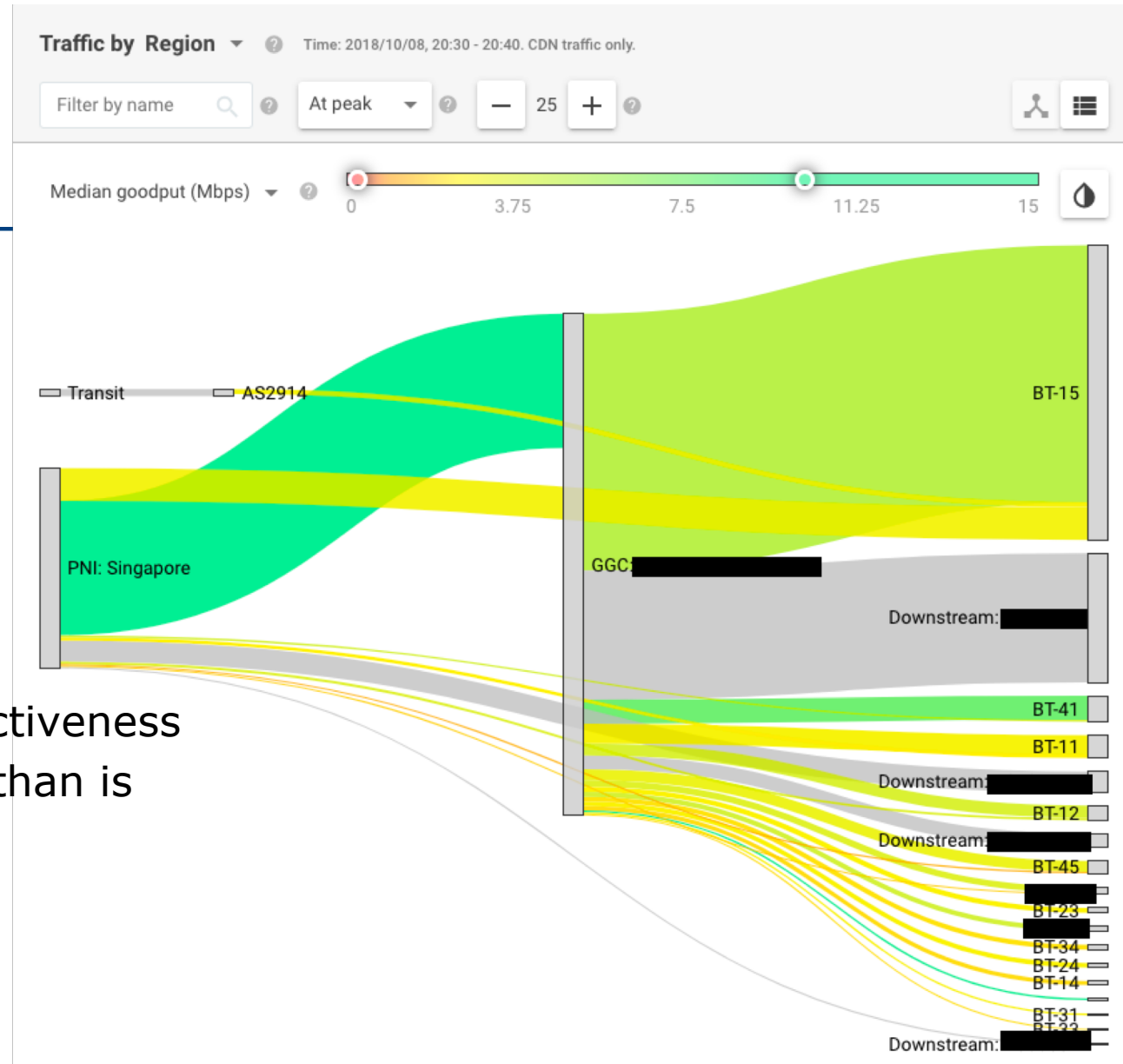
- Many countries do not have content caches
 - Individual operators are not large enough to qualify
 - And therefore are burdened with expensive transit costs
- Solution:
 - Cooperation!
 - Network Operators work together
 - Agree to interconnect their networks
 - Private peering, or more usually, via an Internet Exchange Point
 - And share their hosted content cache across the peerings
 - A significant value proposition for founding any IXP
 - Not only keeping local traffic local, but sharing commonly accessed content

Content Distribution Today

- Well known cooperation examples:
 - Nepal, Bhutan, Vanuatu, Fiji,...
- How does it work?
 - The network operators each share the content caches they host across the IX
 - Operator hosting the cache improves the cache effectiveness for their hosted cache, benefiting their users
 - The transit traffic for cache fill is usually unchanged when adding other operator access to it
 - Their customers are usually looking at the same content!
 - All operators benefit, and the country qualifies for content caches it would otherwise not get

CDN

- ❑ Example of CDN Cache effectiveness
- ❑ Feeding over 3 times more than is arriving via transit link
- ❑ Peers benefiting



Content Distribution Summary

□ Key requirements:

- Low latency to end-user
- High bandwidth to end-user

□ Achievable by:

- Deployment of local caches
- High bandwidth Interconnects between network operators in smaller markets

Evolution Summary

- 20 years ago:
 - Centralised Internet (in US & Europe)
 - Very diverse content, and hosted at origin
 - Clear hierarchy of Tier-1s, Regional providers, and Access providers
 - Access provider goal was to provide international connectivity to that content

Evolution Summary

□ Today:

- Model of centralised Internet is no more
- “Content is King”
 - >80% of traffic volume is from the major content providers
 - Network operator focus today is on delivering content from the major content providers more efficiently than their competitors
 - CDN “performance meters” and Speed Tests now are customer measures of Internet Quality of Service
- Geoff Huston opinion piece:
 - <https://blog.apnic.net/2016/10/28/the-death-of-transit/>

IP Addressing



Where to get address space and who from

IP Addressing Basics

- Internet uses two types of addressing:
 - IPv6 – the new IP protocol
 - IPv4 – legacy IP protocol
- Internet uses classless routing
 - Routers must be CIDR capable
 - Classless InterDomain Routing
 - No routing assumptions made based on the address block
 - Engineers talk in terms of prefix length
 - For example: 158.43/16 and 2001:DB8::/32

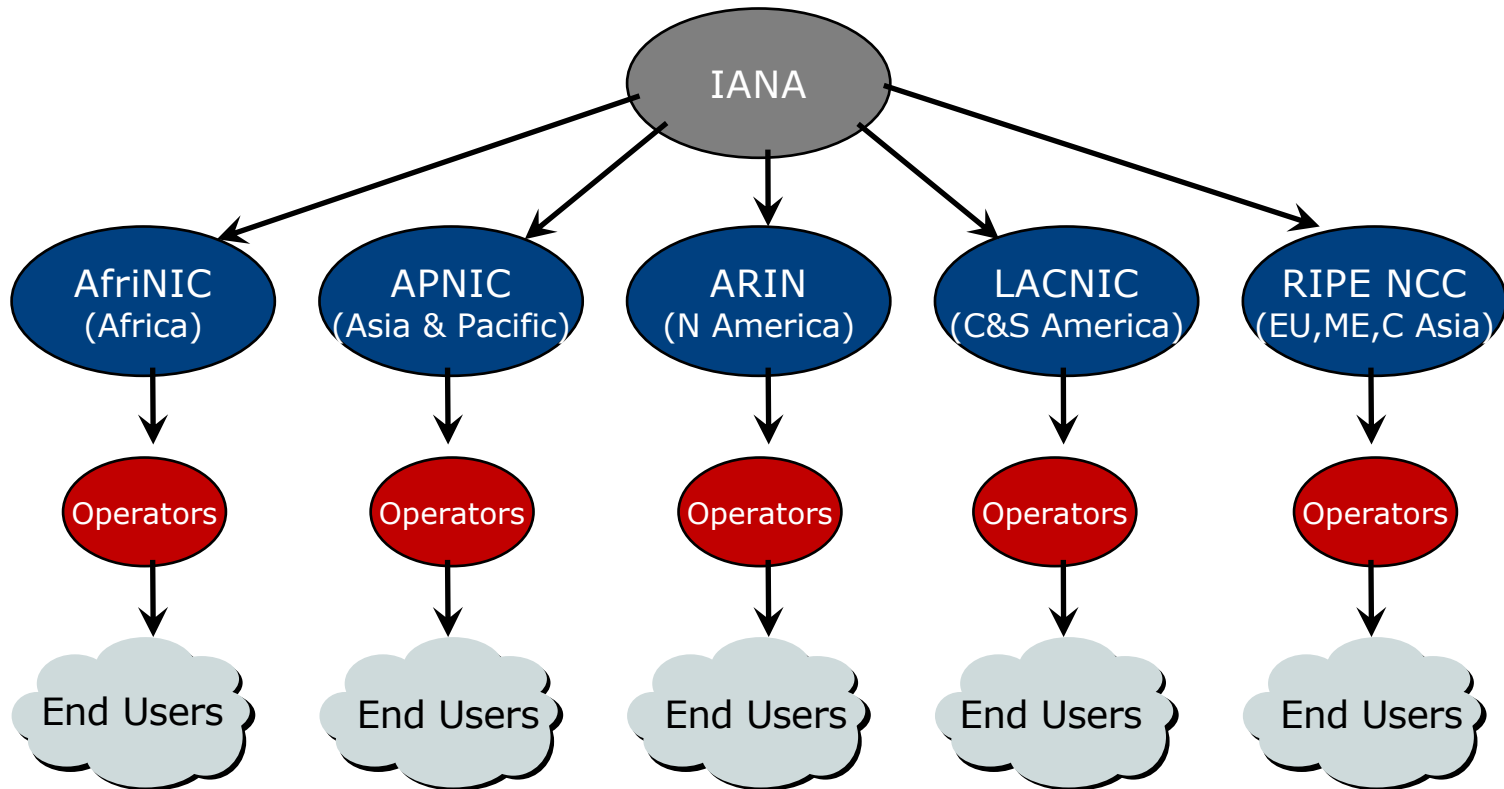
History of IP Addressing

- Pre-CIDR (before 1994)
 - Big networks got a class A
 - Medium networks got a class B
 - Small networks got a class C
- The CIDR IPv4 years (1994 to 2010)
 - Sizes of IPv4 allocations/assignments made according to demonstrated need
 - **CLASSLESS**
- **IPv6 adoption (from 2011)**
 - Network Operators get at least one /32
 - End Sites get /48
 - IANA's free pool is depleted (February 2011) – the size of IPv4 address allocations and assignments is now very limited

IP Addressing

- IP Address space is a resource shared amongst all Internet users
 - Regional Internet Registries delegated allocation responsibility by the Internet Assigned Numbers Authority (IANA)
 - AfriNIC, APNIC, ARIN, LACNIC & RIPE NCC are the five RIRs
 - RIRs **allocate** address space to Network Operators/Local Internet Registries
 - Operators/LIRs **assign** address space to end customers or other Operators
- RIRs address distribution:
 - IPv6 is plentiful
 - IPv4 is very limited

Address delegation hierarchy



Gluing it together



Gluing it together

- Who runs the Internet?
 - No one
 - (Definitely not ICANN, nor the RIRs, nor the US,...)
- How does it keep working?
 - Inter-provider business relationships and the need for customer reachability ensures that the Internet by and large functions for the common good
- Any facilities to help keep it working?
 - Not really. But...
 - Technical staff at Network Operators keep working together!

Operators keep talking to each other...

□ North America

- NANOG (North American Network Operators Group)
- NANOG meetings and mailing list
- www.nanog.org

□ Latin America

- Foro de Redes
- NAPLA
- LACNOG – www.lacnog.org

□ Middle East

- MENO (Middle East Network Operators Group)
- www.menog.org

Operators keep talking to each other...

□ Asia & Pacific

- APRICOT annual conference
 - www.apricot.net
- APOPS mailing list
 - mailman.apnic.net/mailman/listinfo/apops
- PacNOG (Pacific NOG)
 - mailman.apnic.net/mailman/listinfo/pacnog
- SANOG (South Asia NOG)
 - lists.sanog.org/mailman/listinfo/sanog

Operators keep talking to each other...

- Europe
 - RIPE meetings, working groups and mailing lists
 - e.g. Routing WG: www.ripe.net/mailman/listinfo/routing-wg
- Africa
 - AfNOG meetings and mailing list
 - SAFNOG – Southern Africa NOG – www.safnog.org
- Caribbean
 - CaribNOG meetings and mailing list
- And many country NOGs

Operators keep talking to each other...

□ Participation in Peering Fora

- Meetings of the Peering Coordinators of many network operators
 - Planning interconnects between operators, content providers, etc
- Global Peering Forum (GPF)
- Regional Peering Fora (European, Middle Eastern, Asian, Caribbean, Latin American)
- Many NOGs host their own Peering Fora
- Many countries now have their own Peering Fora

□ IETF meetings and mailing lists

- www.ietf.org

Internet Evolution



ISP/IXP Workshops