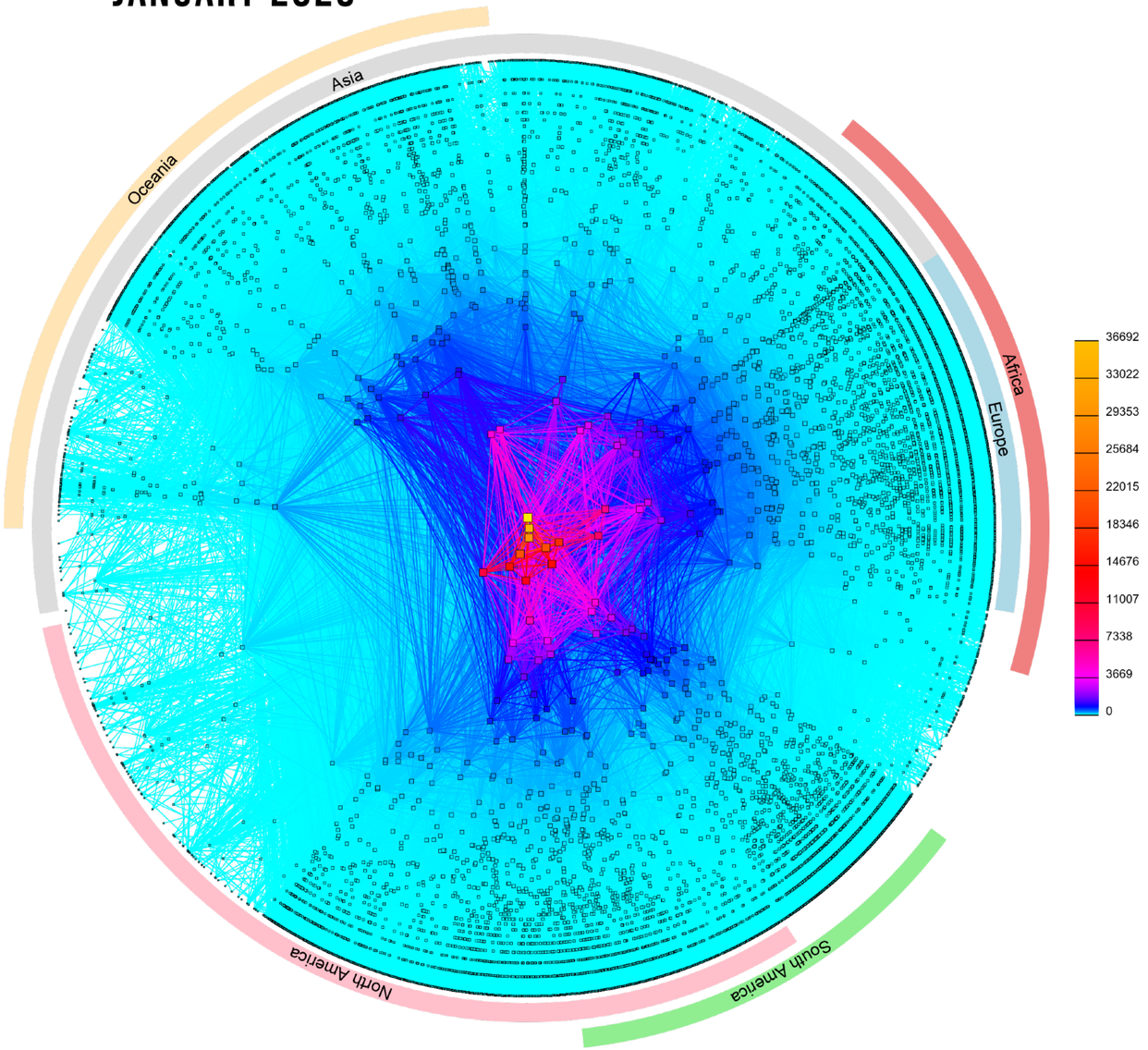
CS249i: The Modern Internet

Stanford Computer Science

Stanford
University

Review

CAIDA'S IPV4 AS CORE GRAPH
JANUARY 2020



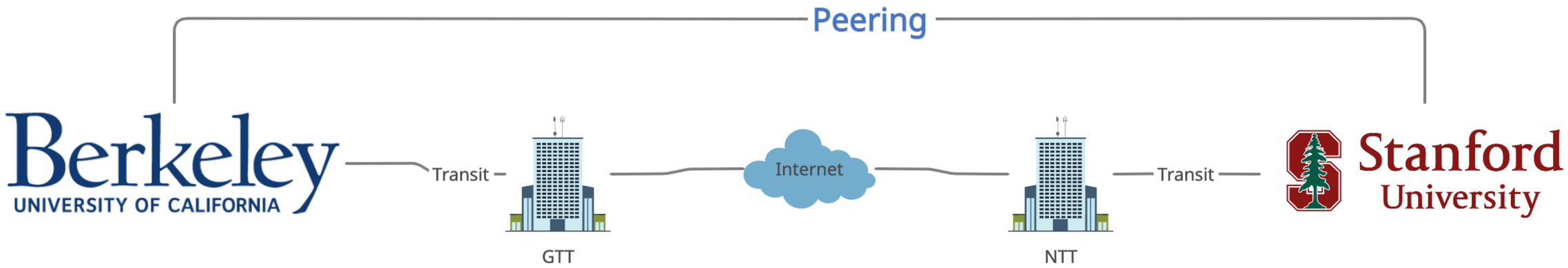
Companies who want Internet access pay IP Transit providers, who in turn pay larger providers

If your Transit Provider didn't pay a larger Transit Provider, then you'd only be able to send traffic to their other customers

Settlement Free Peering

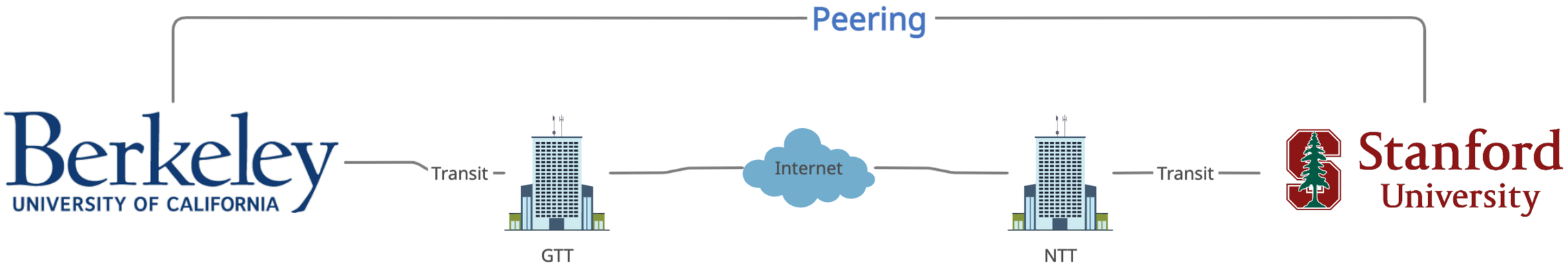
Even if you have IP Transit, two companies may want to exchange traffic without incurring any transit fees — known as "Settlement Free Peering"

Generally the two parties are equals in some way — they are "peers"



Example: Berkeley and Stanford don't want to both pay their ISPs to transit their traffic when they could just peer with one another.

Peering is not Transitive



In this situation, Berkeley and Stanford don't want to be each other's Internet Service Provider — Stanford will only advertise its own IP ranges to Berkeley

Not a substitute for Internet Transit — the peering is only a shortcut for Stanford ↔ Berkeley's traffic to one another

Internet Transit

(Connecting to the Edge)

Choosing an IP Transit Provider

Except for residential access, Internet Transit is a metered service. Different providers have different costs, different assurances

If you're small, you're probably choosing from a limited number of *regional providers* — it's expensive to lease fiber to connect to someone far away from you

Not necessarily a direct correlation between price and quality. ISPs have different reputations. Knowledge is shared through industry groups like NANOG (North American Network Operators' Group)

Pre-Commit + Burst Rate + Port Size

Generally, you negotiate two costs:

Commit ("Flat Rate") — how much you pay each month regardless of whether you use less than the committed amount (at a discounted rate)

Burst Rate — how much you pay per megabit between commit and 95th percentile. Typically more expensive than or tied to flat rate.

For example, you could commit to 3 Gbit on a 10 Gbit port.

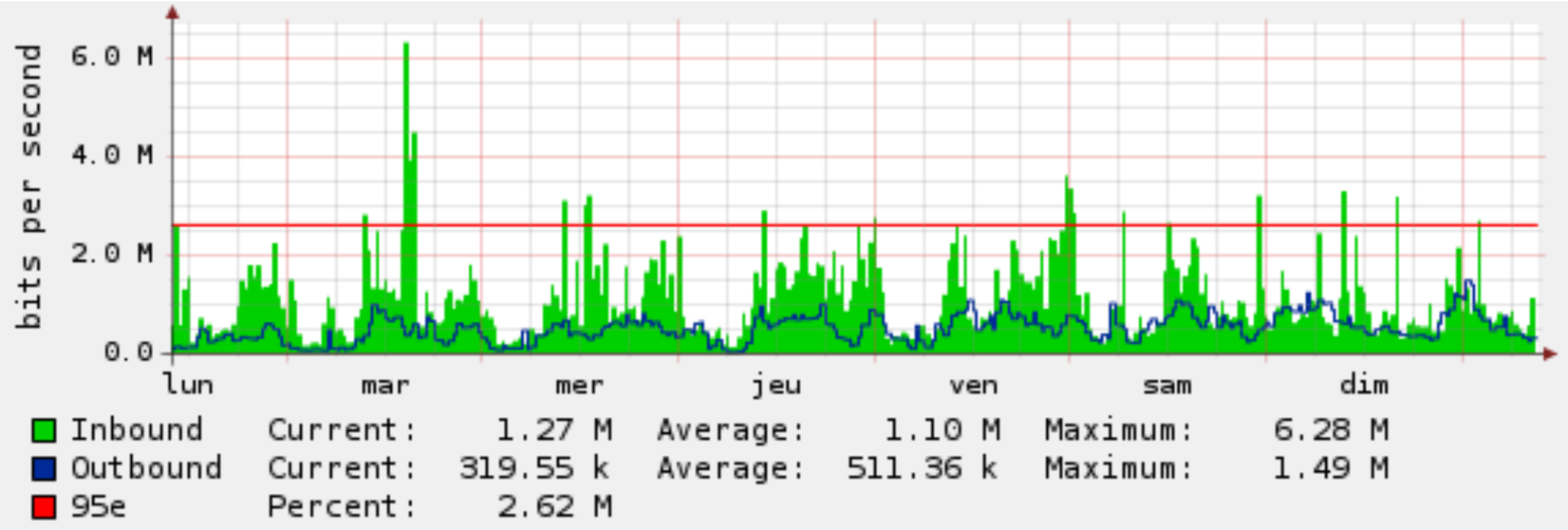
As part of this negotiation, you also define an SLA (Service Level Agreement). This quantifies acceptable downtime, max packet drop rate, transit times within the ISP.

"SLAs are widely dismissed by operators as merely insurance policies"
— William Norton (Co-Founder of Equinix)

95th Percentile Pricing (aka 95/5 pricing)

Calculate usage every 5 minutes (max of inbound and outbound). Line up from lowest to highest. Charge based on 95th percentile.

Allows small bursts without affecting overall costs. Big bursts of traffic throw off averages — demonstrated problem in late 90s early 2000s



Bandwidth can be used at a higher rate for up to 72 min/day with no penalty.

Example IP Transit Order

Account Number:	**Account Name:** [redacted]				
Location(s):	*See Special Instructions*				
Port:	☐ Fast Ethernet ☐ Gigabit Ethernet		☒ TenGigabit Ethernet ☐ One Hundred Gigabit Ethernet	**MRC**	**NRC**
	VLAN sub-interface requirements should be detailed in the Special Instructions				
Bandwidth:	☐ IP Fixed ☒ IP Burstable	Price Per Mbit - Committed: [redacted]	Committed Bandwidth: 3000 Mbps	[redacted]	[redacted]
		Price Per Mbit - Burst: [redacted]	Maximum Bandwidth: 10000 Mbps		
		USID: *(Existing Service Only)*	☐ Cumulative Billing *(Specify detail in Special Instructions)*		
Circuit or Cross-Connect:	☐ Customer Provided Circuit or Cross-Connect			$ 0.00	$ 0.00
	☒ NTT Provided Circuit or Cross-Connect *Telco Quote #: Q113588.1*			[redacted]	[redacted]

Real World Pricing

Lst prices are inflated and negotiable (as with any enterprise product). I've seen 30-50% off of advertised. Get multiple quotes.

Prices typically are negotiated under NDA and cannot be publicly shared

Prices are consistently falling — do not sign long contracts — Price != Quality

This year I've signed contracts between
≈ $0.20 - $0.50 / Mbit for IP commit

Year	Internet Transit Prices (in Mbps, min commit)		% Decline
1998	$1200	per Mbps	
1999	$800	per Mbps	33%
2000	$675	per Mbps	16%
2001	$400	per Mbps	40%
2002	$200	per Mbps	50%
2003	$120	per Mbps	40%
2004	$90	per Mbps	25%
2005	$75	per Mbps	17%
2006	$50	per Mbps	33%
2007	$25	per Mbps	50%
2008	$12	per Mbps	52%
2009	$9.00	per Mbps	25%
2010	$5.00	per Mbps	44%
2011	$3.25	per Mbps	35%
2012	$2.34	per Mbps	28%
2013	$1.57	per Mbps	33%
2014	$0.94	per Mbps	40%
2015	$0.63	per Mbps	33%
Source: DrPeering.net			

Hurricane Electric Advertised Special (*Insanely* Cheap)

Specials

Get BGP+IPv6+IPv4 for $0.09/Mbps!

Basic Internet Access

Do you need to BGP peer with anyone for Internet access?

If you're not a large organization, you're likely not going to BGP peer with your ISP

Why No?

- Tremendous room for error in BGP configuration
- Minimum size of routed prefix is /24 = 256 addresses
- Router needs sufficient memory to store Internet routing table ($$)
    - Can sometimes ask for your ISP to provide only default route (0/0)

Why Yes?

- You want to advertise an AS to multiple upstream providers
- You want to control your own AS and IP blocks

Peering Policies

Peering Policies

ISPs generally fall within one of three peering policy categories:

- **Open:** ISP will peer with almost everyone
- **Selective:** ISP has specific criteria for peers
- **Restrictive:** ISP generally does not want to peer with you

Individual Peering Policies can differ dramatically

- Most ISPs are relatively selective (typically, larger = more selective)
- Hurricane Electric will peer with anyone
    - Having the largest number of peers is a marketing win for them
- Content Providers (e.g., Google) have relatively open peering policies
    - They'd prefer not to pay ISPs for traffic

Convergence

Can you access the full Internet through only settlement-only peering (without buying any IP Transit)?

If you peered with enough networks and ISPs, you theoretically might not need to buy IP Transit, but…

- It's difficult to peer with a tremendous number (tens or hundreds of thousands) of networks. How will you connect to all of them?
- Large networks won't peer with smaller players
  (little value for them, big value for you)
- Large ("Tier-1") networks have direct customers, so even if a Tier-2 were to peer with every other Tier-2, the wouldn't have everything

Tier-1 Network Definition

Tier-1 Provider: a network that can access every network on the Internet through settlement-free peering

Tier-1 Networks are so well connected that nobody would charge them for connectivity

Exclusive Club — Tier 1s have little incentive to support new Tier 1 networks

Name	Headquarters	AS number
AT&T[11]	United States	7018
Deutsche Telekom Global Carrier[13]	Germany	3320
GTT Communications	United States	3257
Liberty Global[17][18]	United Kingdom[19]	6830
Lumen Technologies (formerly CenturyLink formerly Level 3)[21][22][23]	United States	3356
Lumen Technologies (formerly CenturyLink formerly Level 3 formerly Global Crossing)[21][22][23]	United States	3549
NTT Communications (formerly Verio)[26]	Japan	2914
Orange[27]	France	5511
PCCW Global	Hong Kong	3491
T-Mobile US (formerly Sprint)[29]	United States	1239
Tata Communications (formerly Teleglobe)[31]	India	6453
Telecom Italia Sparkle (Seabone)[33]	Italy	6762
Telia Carrier[34]	Sweden	1299
Telxius (Subsidiary of Telefónica)[36]	Spain	12956
Verizon Enterprise Solutions (formerly UUNET)[42]	United States	701
Zayo Group (formerly AboveNet)[44]	United States	6461

Telia Carrier (Tier 1)'s Peering Policies

Telia will *consider* peering with you if…

- Peer must be able to exchange traffic in EMEA, AMER, and APAC
- Peer must be able to connect to Telia's backbone network in…
  - At least 40 cities in separate metropolitan areas in Europe,
  - At least 35 cities in separate metropolitan areas in North America,
  - At least 3 cities in separate metropolitan areas in Asia Pacific/Oceania
- Peer must be willing to implement a minimum of 18 points of interconnection
  - Each interconnect must be capable of supporting 100 Gbps of Traffic
- Traffic exchanged shall be balanced (inbound:outbound cannot exceed 3:1)
- Peer must have at least 1,500 distinct BGP customers

Interconnection Relationship Review

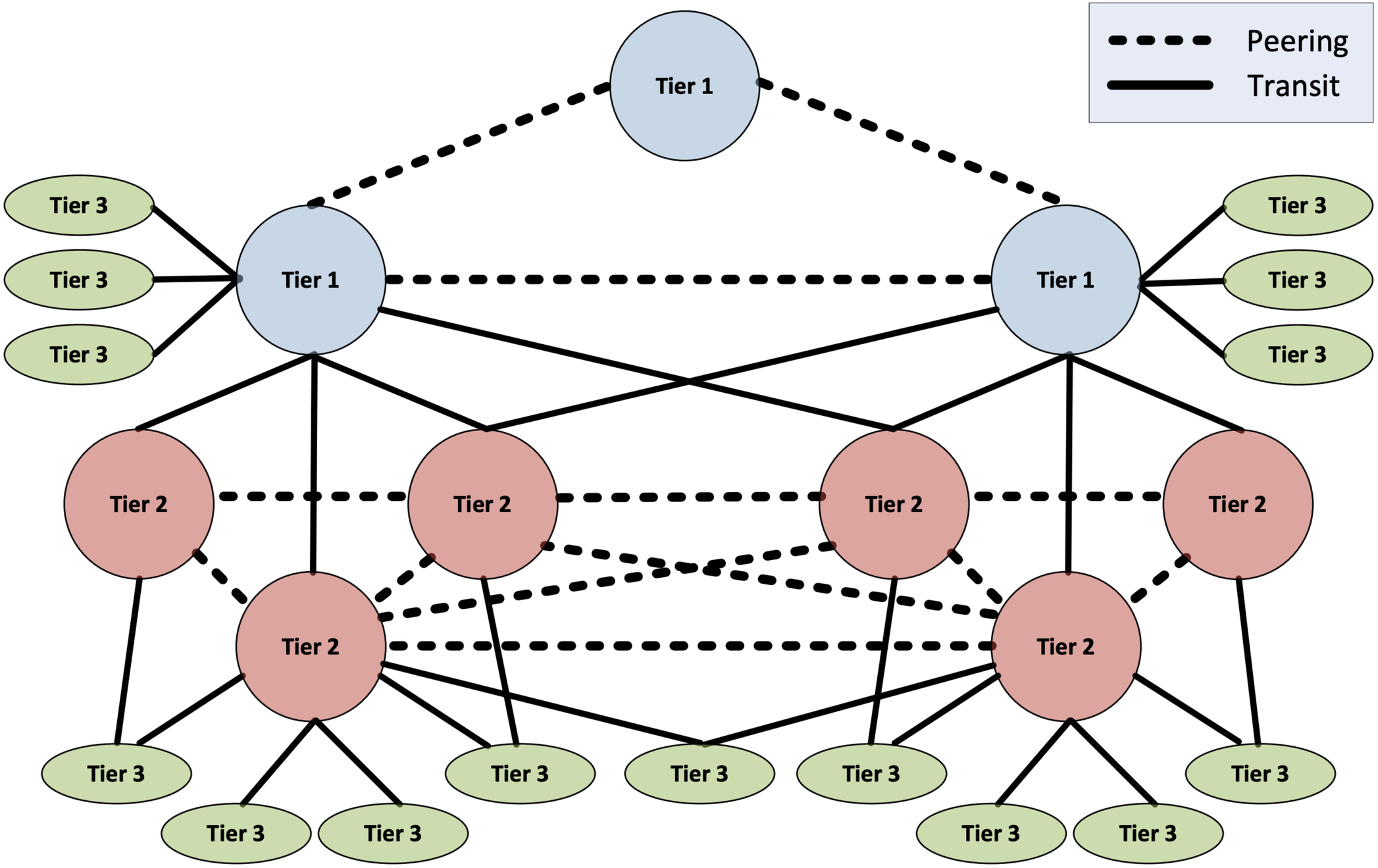
Peering:

- Typically settlement free between two "peers"
- Non-transitive – only gain access to that ISP's customers
- Unless you peer with an unattainable number of ISPs, you're not going to gain access to the full Internet routing table — not possible in practice

IP Transit:

- Pay a larger Transit Provider for access to their full routing table
- Your "default" route — any traffic not to a peer goes to Transit Provider

Most ISPs (unless you're a Tier-1) will have both peers *and* transit providers


Peering
Transit
Tier 1
Tier 1
Tier 1
Tier 2
Tier 2
Tier 2
Tier 2
Tier 2
Tier 2
Tier 3
Tier 3
Tier 3
Tier 3
Tier 3
Tier 3
Tier 3
Tier 3
Tier 3
Tier 3
Tier 3
Tier 3
Tier 3
Tier 3

 Physical Interconnect

Direct Circuit Peering

Historically, ISPs would directly connect their networks with fiber — known as Direct-Circuit Peering

- Typically the two peering ISPs would ask a local provider to provide fiber between two POPs (points of presence) within a local metropolitan area
- Expensive and slow (circuits could easily be delayed by months or years)

Sometimes still happens, but only if a small number of connections *and* a single metropolitan area

- For example, if you just need IP Transit from your regional ISP

Carrier Hotels + Private Peerings

ISPs eventually agreed on specific buildings where ISPs, cloud providers, CDNs, and companies could peer with one another — known as carrier hotels

Originally, many locations were at the landing points of undersea cables. Locations have since increased

Typically owned by a third-party rather than a carrier

Carrier Hotels support **Private Peerings** — two ISPs directly connect their racks together (known as a "cross connect" or "XC")

Downside: You might have to have physical fiber connects with dozens of other people in the same data center. Router ports are expensive!

60 Hudson Street, New York City

Internet Exchange Points (IXPs) + Public Peering

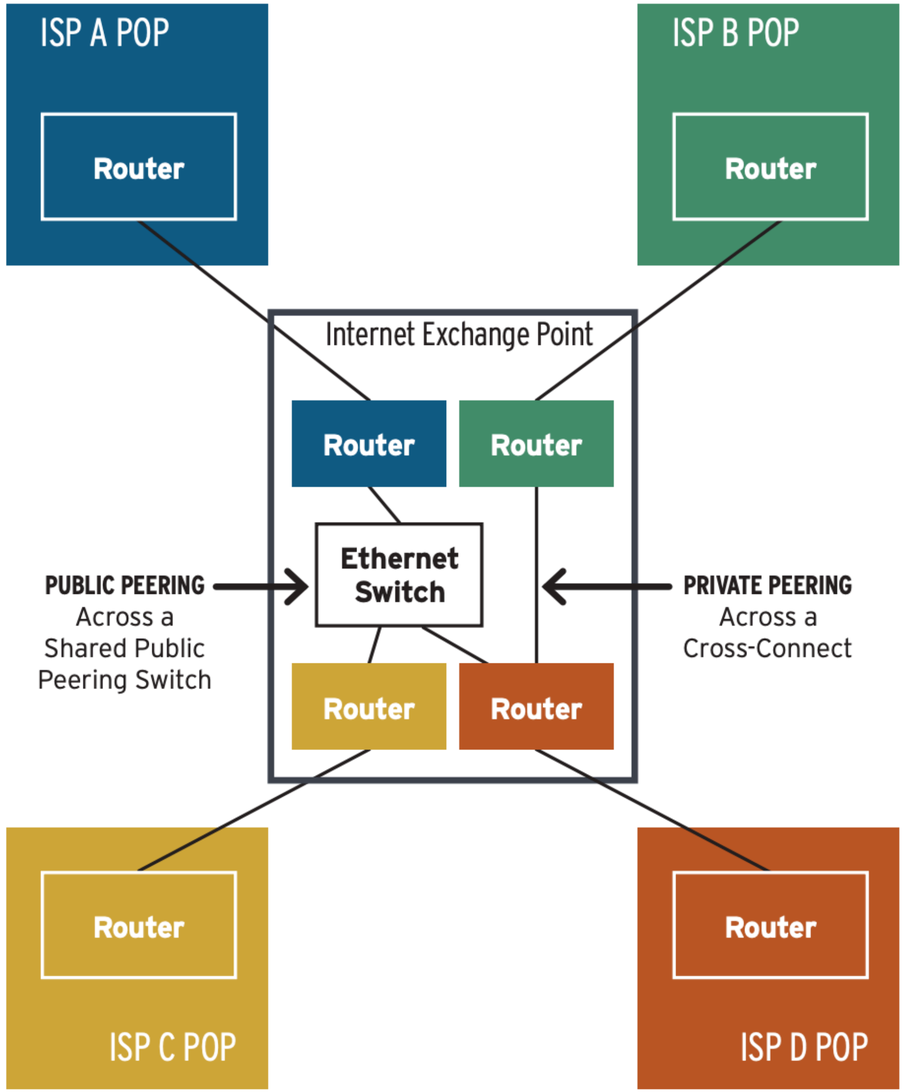
Carrier Hotels provide shared interconnect (switching fabric) between ISPs

Allow ISPs to BGP peer with a large number of organizations through a single link

Peerings that use shared fabric ("public interconnect") known as a "**Public Peering**"

You still have to negotiate the BGP peering with others on the exchange

In the U.S., private peering more common.
In Europe, public peering more common.



Many ISPs Will be Your Friend at an IXP

Most content providers will peer with you over public exchanges

So will cloud providers

Little downside for them not to if they're already on the exchange

Most ISPs (even Tier 1s) will sell you transit at an IXP (via private peering)

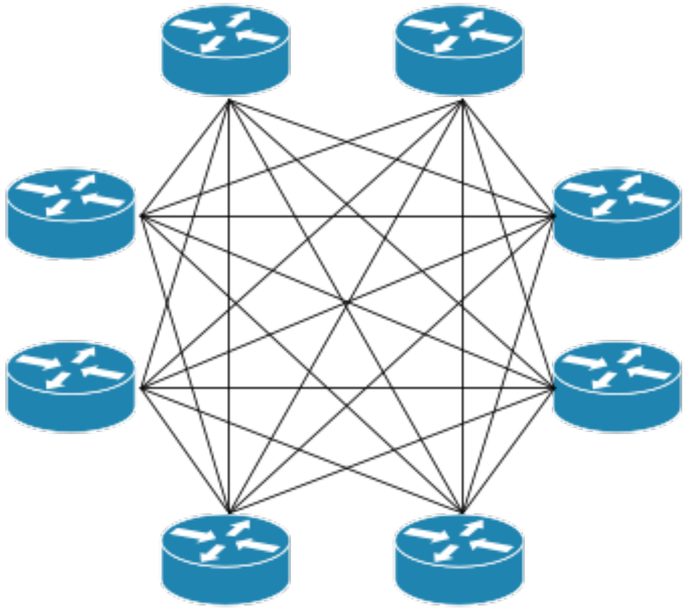
Multilateral Peering Exchanges

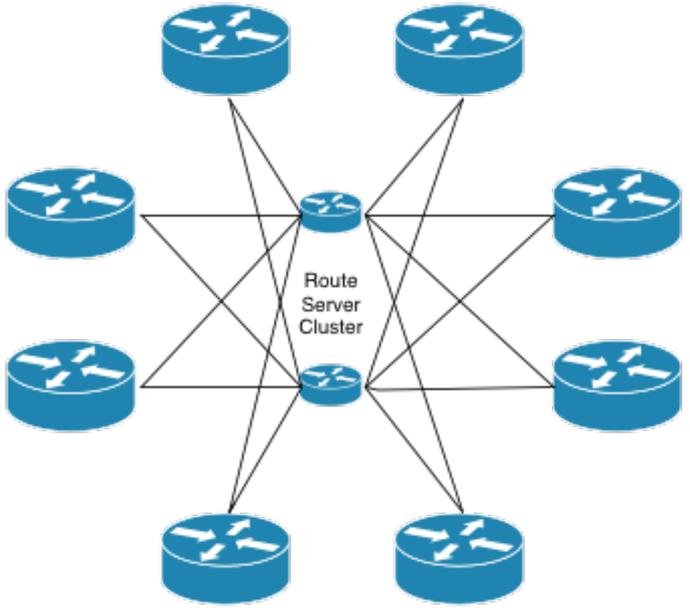
Typically, you'd establish peering relationships with others at an exchange

This can become cumbersome if there are a large number of players (could be 100s or 1000s at a single IXP)

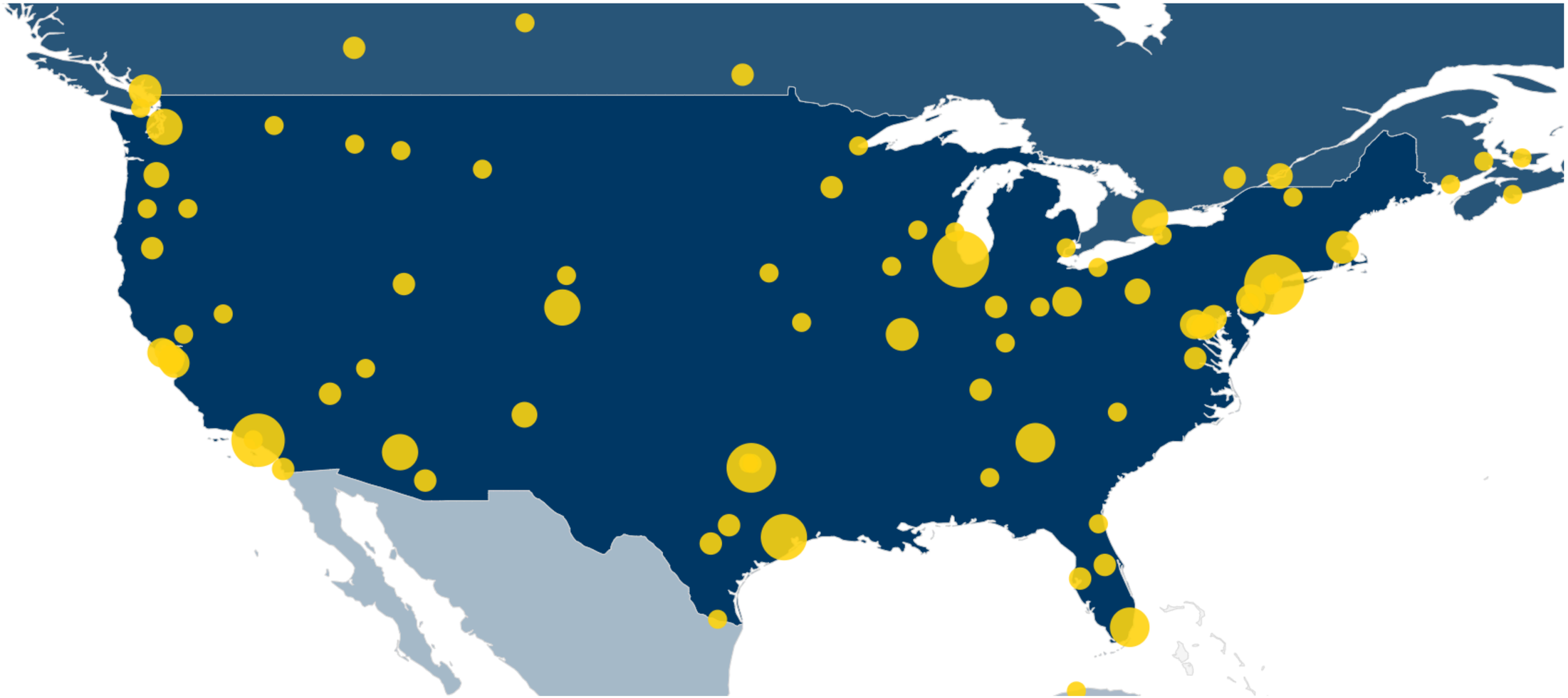
Route server allows those with open policies to all BGP peer with only a single entity to both advertise its routes and collect routes from others on the exchange

Known as Multilaterial Peering

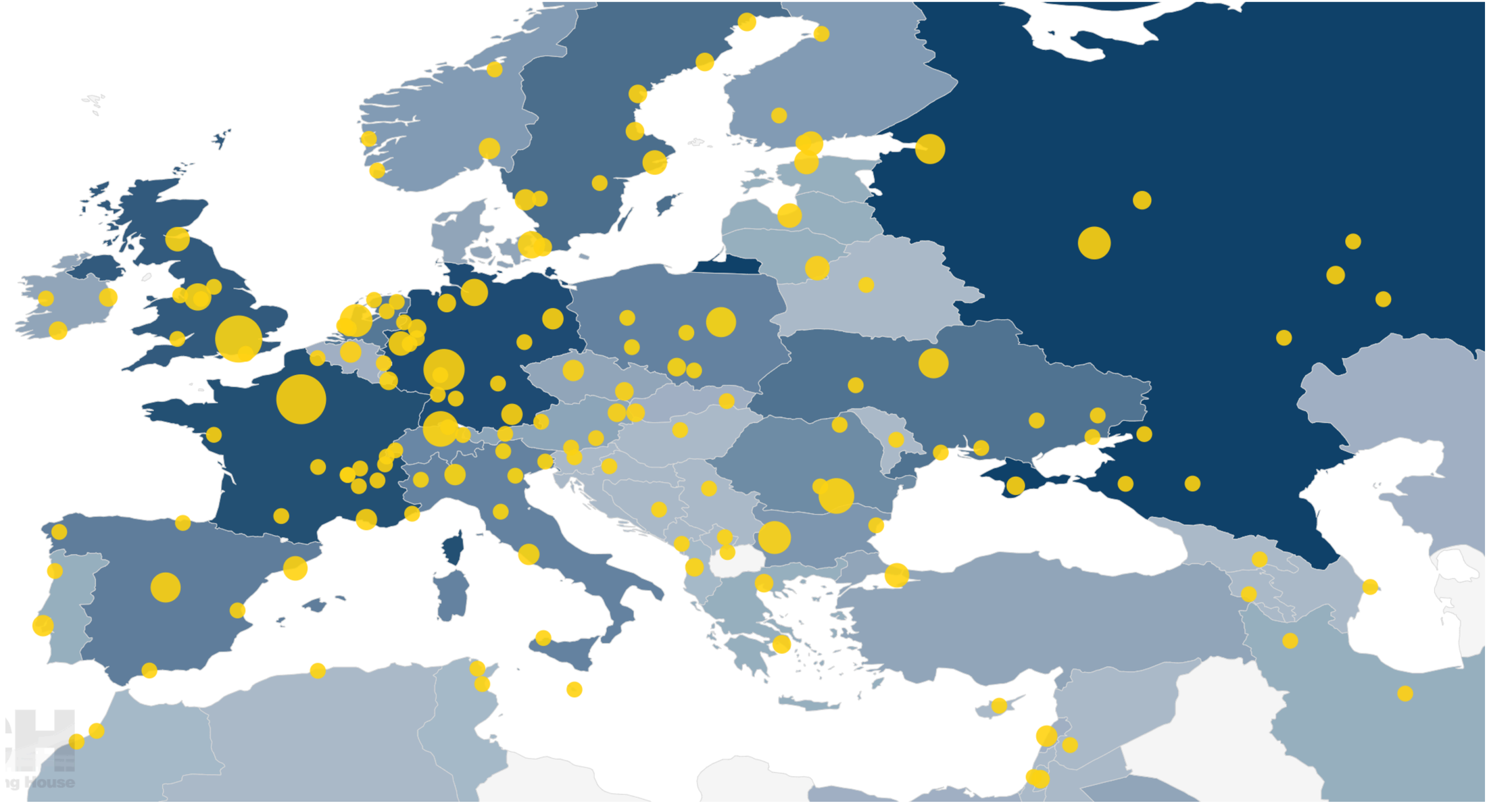




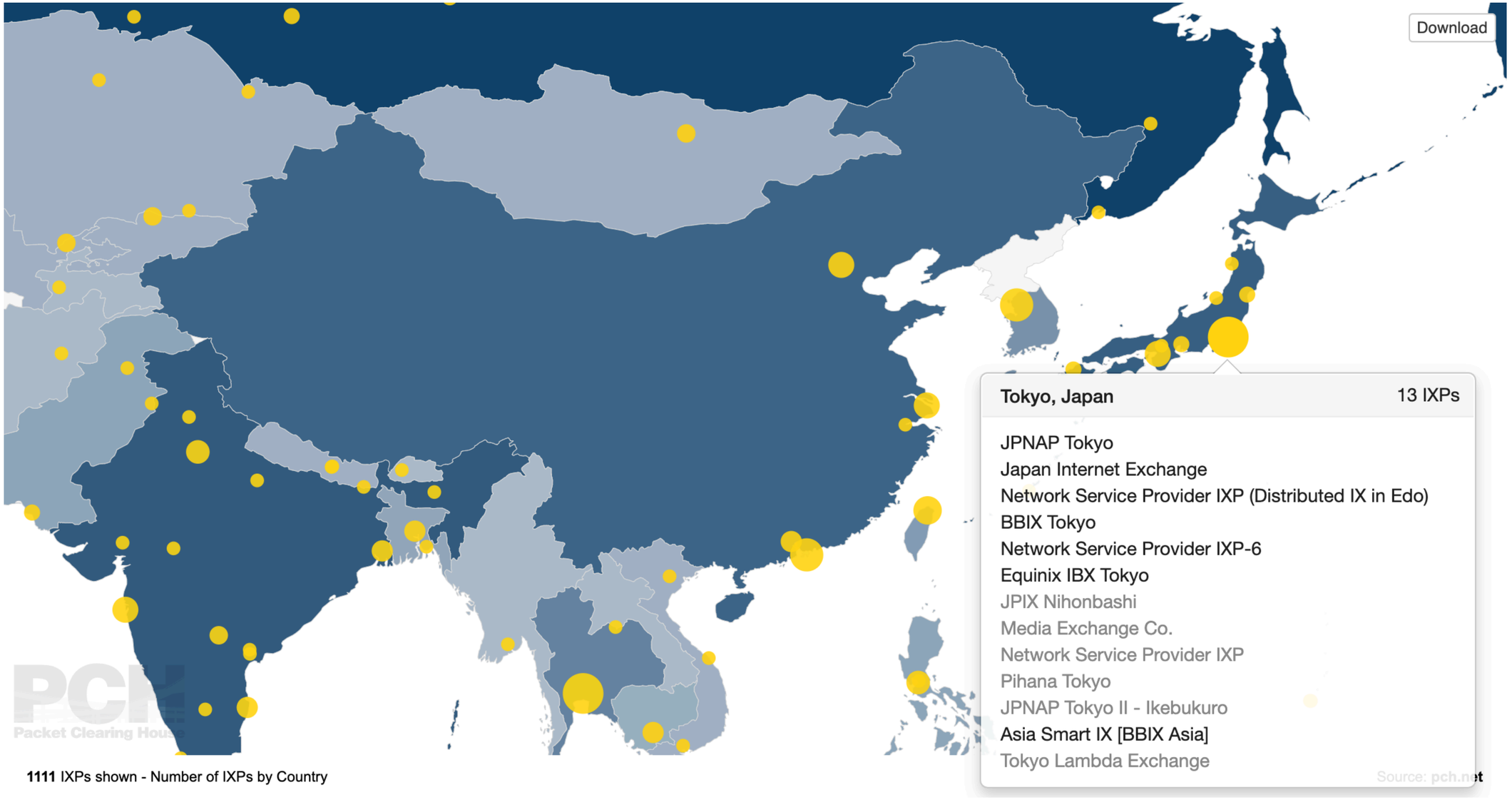
Packet Clearing House — Where are U.S. IXPs?



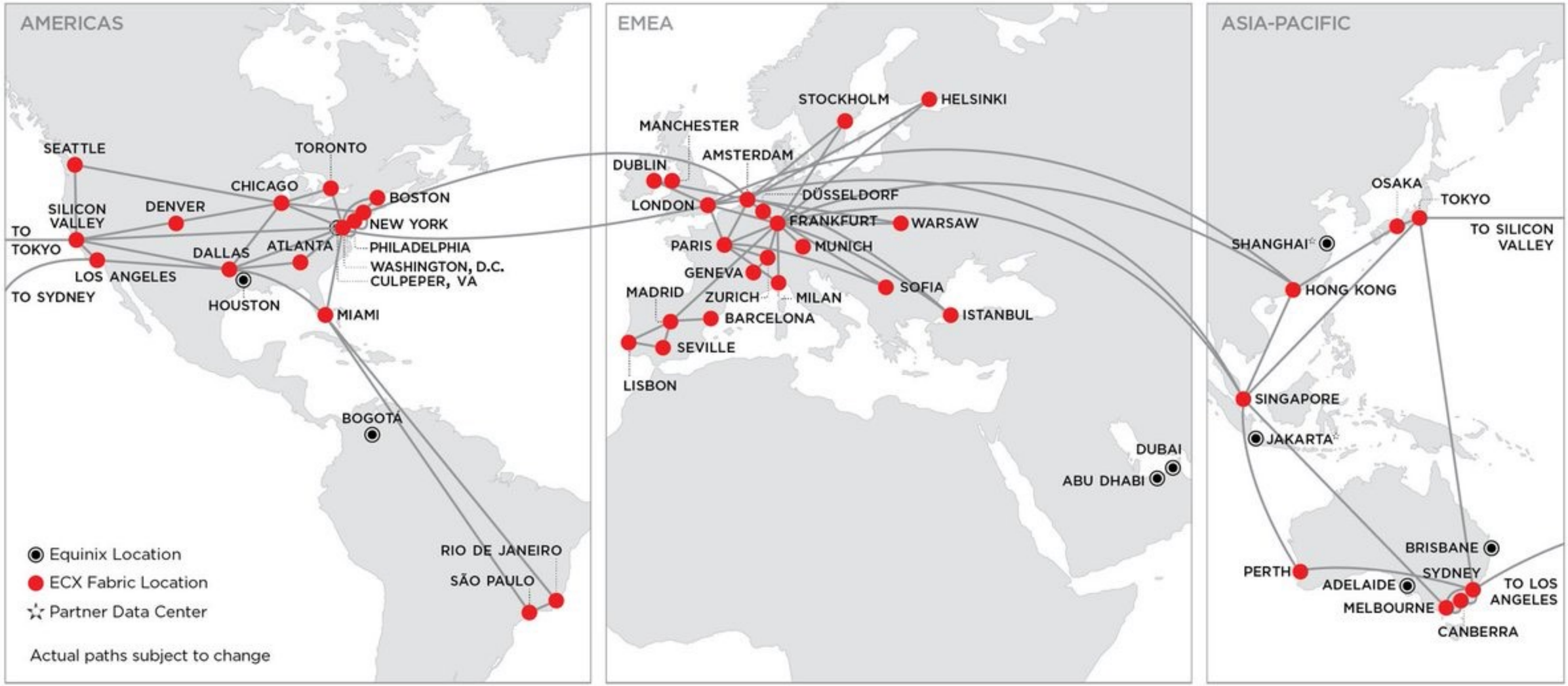
Packet Clearing House — Europe



Packet Clearing House — Asia



Equinix — Largest Commercial IXP Provider



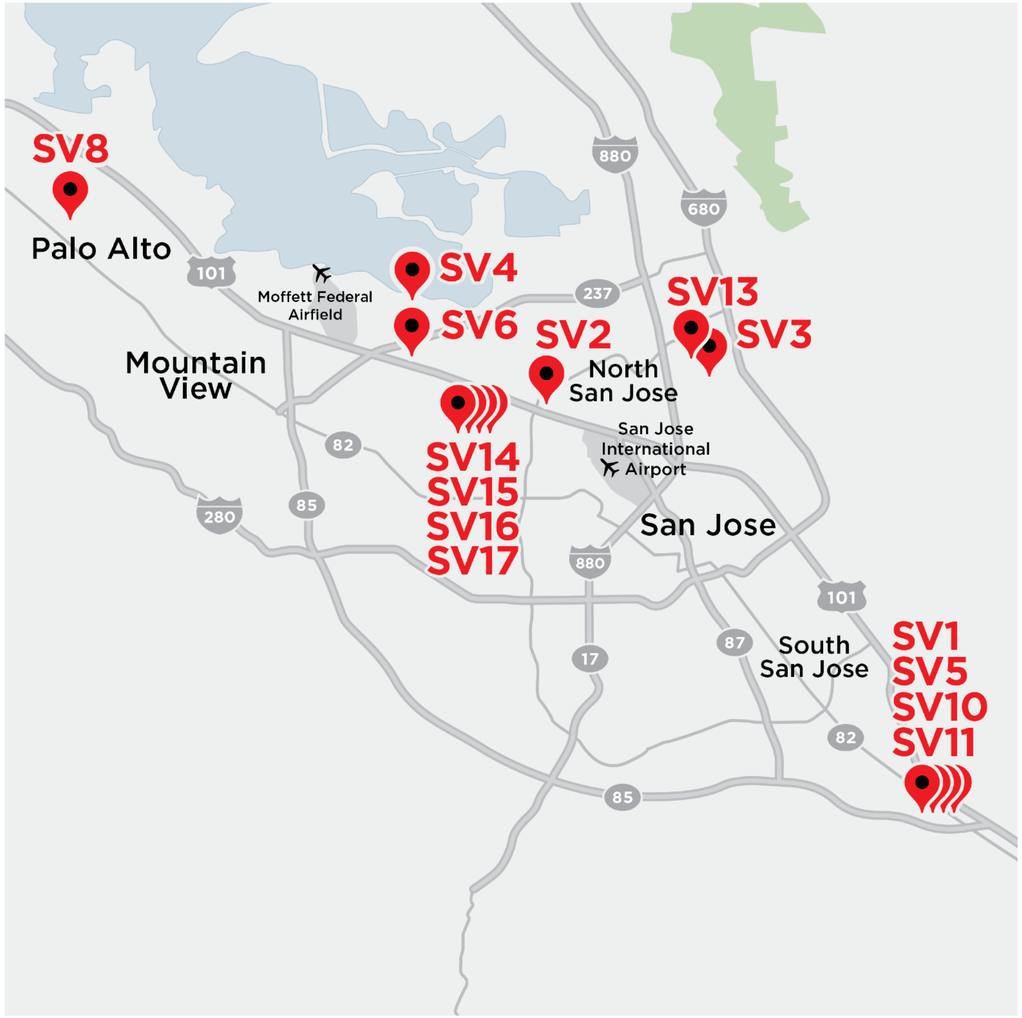
Equinix Internet Exchange San Jose (Bay Area)

Facilities:

- Equinix SV1/SV5/SV10 - San Jose
- Equinix SV2 - Santa Clara
- Equinix SV3 - San Jose
- Equinix SV4 - Sunnyvale
- Equinix SV8 - Palo Alto

~210 (publicly listed) participants on San Jose Exchange. ~101 on Palo Alto Exchange.

~500 organizations listed between those 5 hotels (many more than on public exchange)

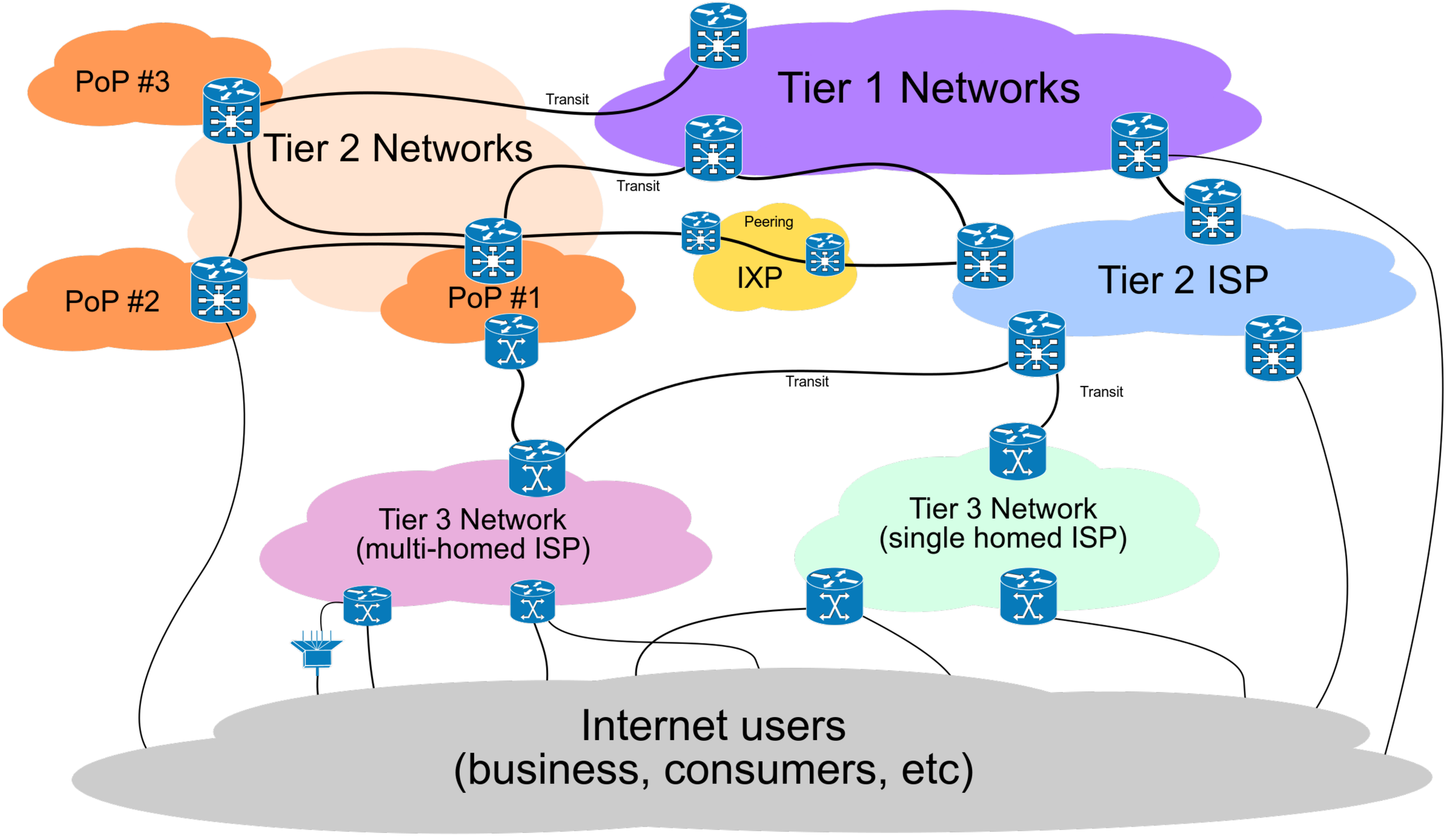


European Example: AMS-IX

AMS-IX Points of Presence

	Data Centre	Phone	Address
1	NorthC Amsterdam (AMS01)	+31 (0)20 486 9773	Kabelweg 48a, Amsterdam
2	Digital Realty AMS17	-	Science Park 120, Amsterdam
3	Digital Realty AMS04	+31 (0)20 480 4415	H.J.E. Wenckebachweg 127, Amsterdam
4	Equinix AM1/2	+31 (0)53 434 0570	Luttenbergweg 4, Amsterdam
5	Equinix AM3	+31 (0)20 808 0015	Science Park 610, Amsterdam
6	Equinix AM5	+31 (0)20 592 8263	Schepenbergweg 42, Amsterdam
7	Equinix AM6	+31 (0)53 436 2666	Duivendrechtsekade 80A, Amsterdam
8	Equinix AM7	+31 (0)53 434 0570	Kuiperbergweg 13, Amsterdam
9	EuNetworks	+31 (0)20 354 8098	Paul van Vlissingenstraat 16, Amsterdam
10	Iron Mountain	+31 (0)20 316 5170	J.W. Lucasweg 35, Haarlem
11	Global Switch	+31 (0)20 666 6300	Henk Sneevlietweg 2-6, Amsterdam
12	Interxion	+31 (0)20 880 7700	Tupolevlaan 101, Schiphol-Rijk
13	Interxion	+31 (0)20 560 6600	Science Park 121, Amsterdam
14	Nikhef	+31 (0)20 592 2037	Science Park 105, Amsterdam

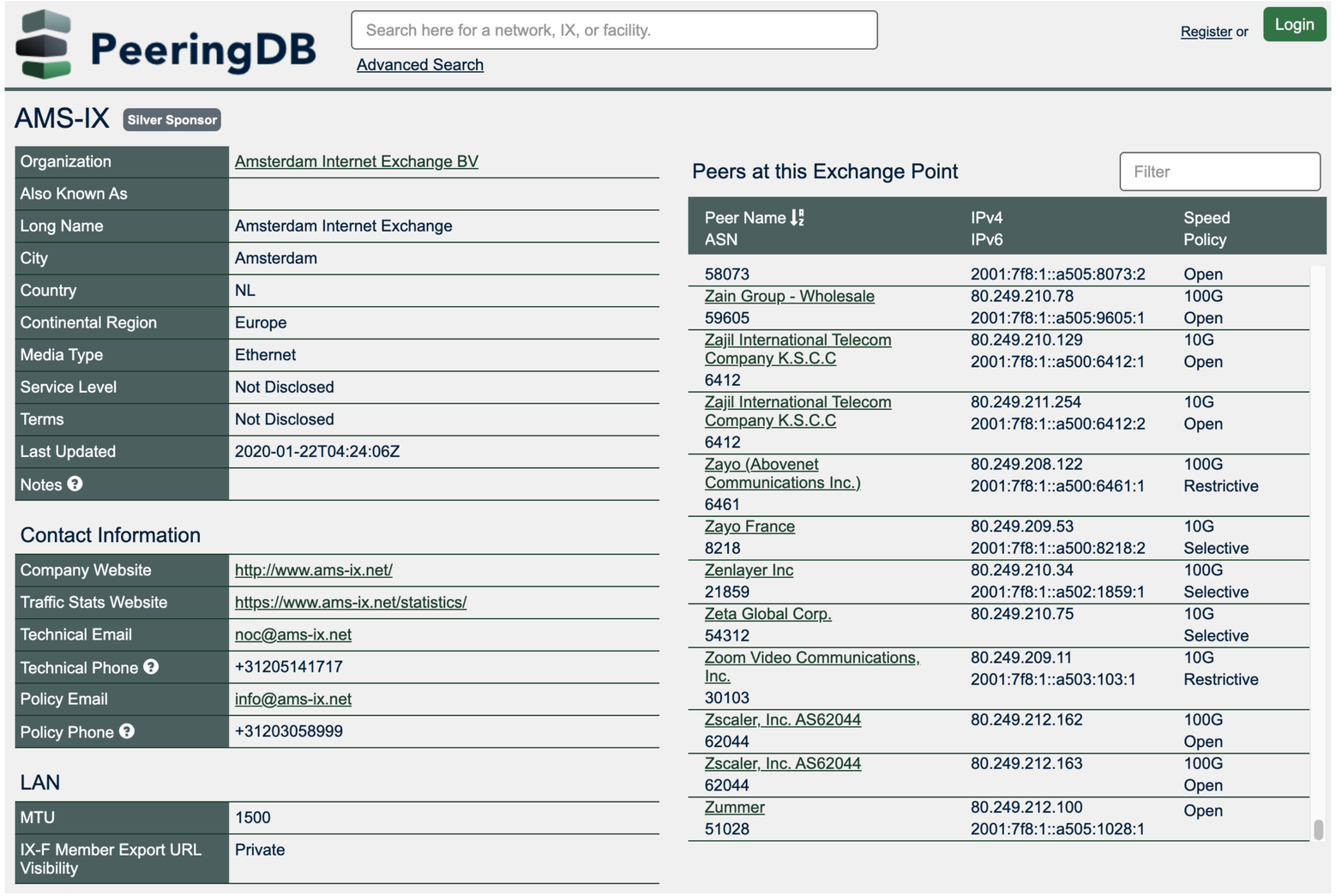
Members	882[1]
Ports	1,438 [1]
Peers	1,316 [1]
Peak in	9.022 Tb/s[2]
Peak out	10.287 Tb/s[2]
Daily in (avg.)	6.42[8] Tb/s[2]
Daily out (avg.)	6.43[8] Tb/s[2]


PoP #3
Tier 1 Networks
Transit
Tier 2 Networks
Transit
Peering
IXP
Tier 2 ISP
PoP #2
PoP #1
Transit
Transit
Tier 3 Network
(multi-homed ISP)
Tier 3 Network
(single homed ISP)
Internet users
(business, consumers, etc)

PeeringDB

PeeringDB is a free database of networks, IXPs, and facilities

https:/peeringdb.com/



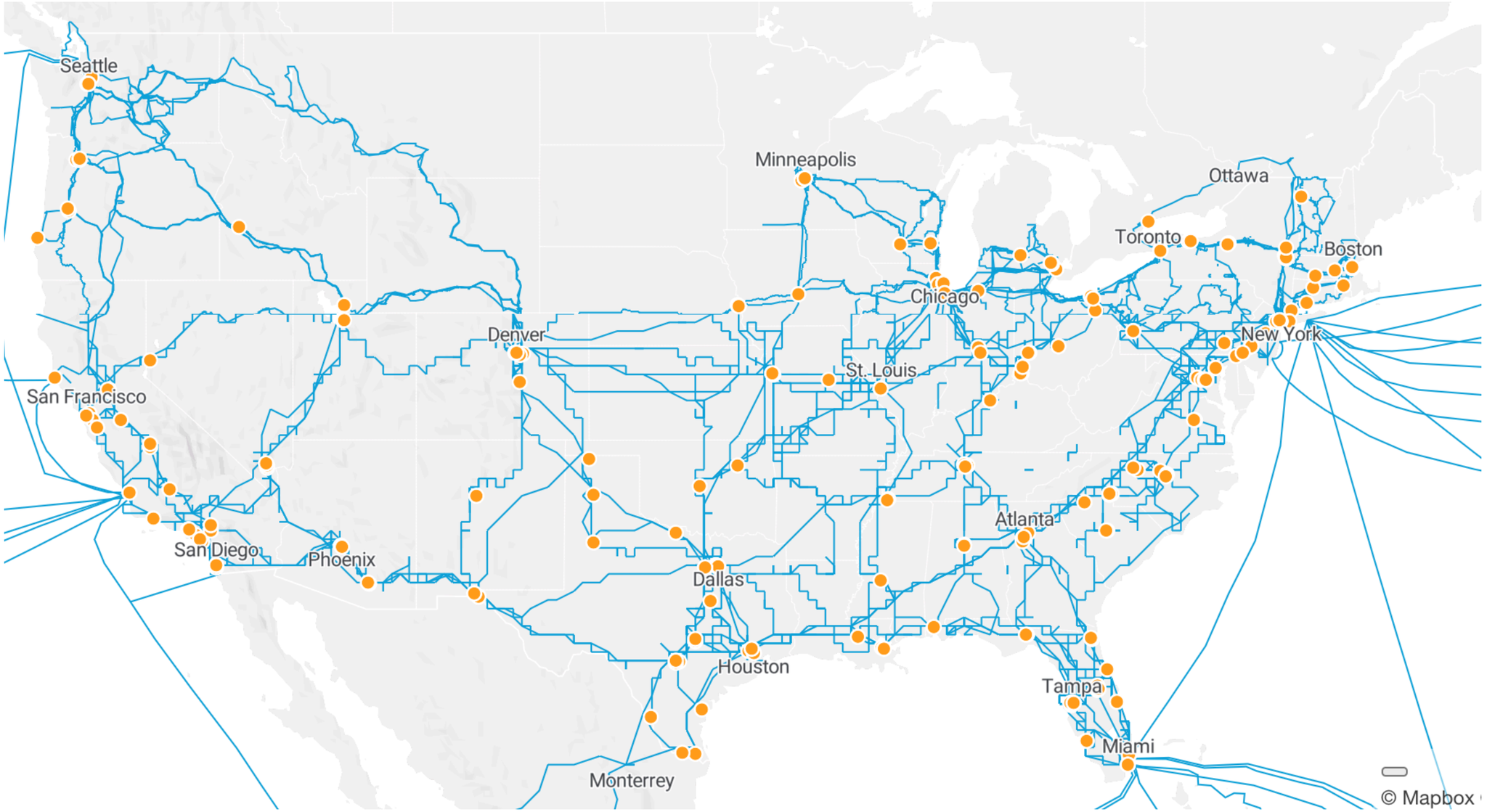
Tier 1 Network Backbone

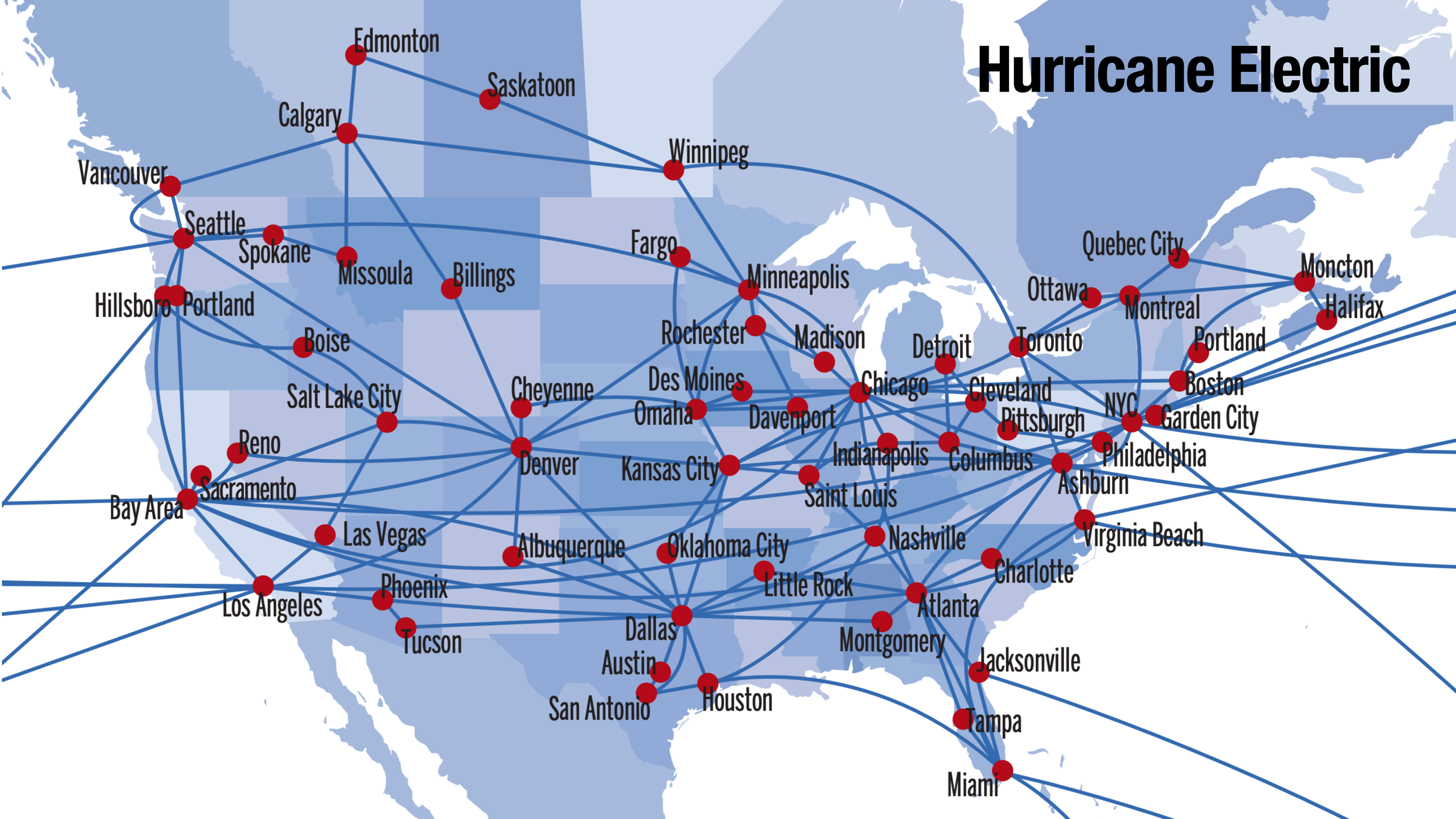
AT&T

gtt

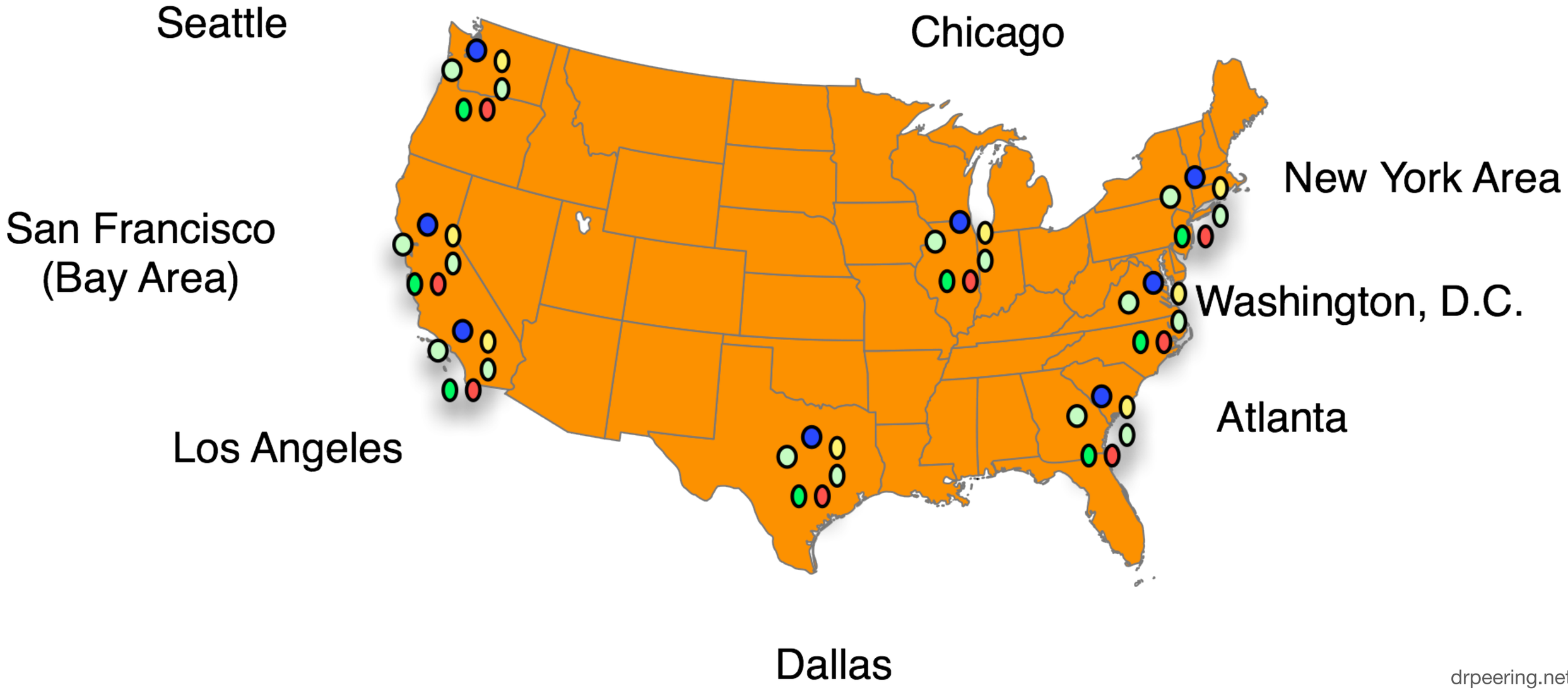

Vancouver
Seattle
Minneapolis
Montreal
Moncton
Halifax
Toronto
Albany
Osaka
Tokyo
Chicago
Detroit
Buffalo
Boston
Cleveland
New York
Pittsburgh
Philadelphia
Salt Lake City
Columbus
Kansas City
Ashburn
San Francisco
St. Louis
Denver
Washington, D.C.
San Jose
Las Vegas
Hong Kong
Charlotte
Dallas
Los Angeles
Phoenix
San Diego
Atlanta
Baton Rouge
Jacksonville
Houston
Miami
Singapore
Mexico City
to Sydney, to Melbourne
to Sydney
to Sydney
to São Paulo
to São Paulo
to Johannesburg
to Johannesburg
Copenhagen
Dublin
Cork
Manchester
Amsterdam
Hamburg
London
Berlin
Warsaw
Reading
Brussels
Frankfurt
Prague
Paris
Munich
Vienna
Zürich
Budapest
Geneva
Milan
Bordeaux
Bucharest
Sofia
Marseille
Istanbul
Barcelona
Rome
Madrid

Lumen


Seattle
Minneapolis
Ottawa
Toronto
Boston
Chicago
Denver
New York
St. Louis
San Francisco
Atlanta
San Diego
Phoenix
Dallas
Houston
Tampa
Miami
Monterrey
© Mapbox


Hurricane Electric
Edmonton
Saskatoon
Calgary
Winnipeg
Vancouver
Seattle
Spokane
Missoula
Billings
Fargo
Minneapolis
Quebec City
Moncton
Ottawa
Montreal
Halifax
Hillsboro
Portland
Rochester
Madison
Detroit
Toronto
Portland
Boise
Des Moines
Chicago
Cleveland
Boston
Salt Lake City
Cheyenne
Omaha
Davenport
Pittsburgh
NYC
Garden City
Reno
Denver
Indianapolis
Columbus
Philadelphia
Sacramento
Kansas City
Saint Louis
Ashburn
Bay Area
Las Vegas
Albuquerque
Oklahoma City
Nashville
Virginia Beach
Charlotte
Phoenix
Little Rock
Los Angeles
Atlanta
Dallas
Montgomery
Tucson
Jacksonville
Austin
San Antonio
Houston
Tampa
Miami

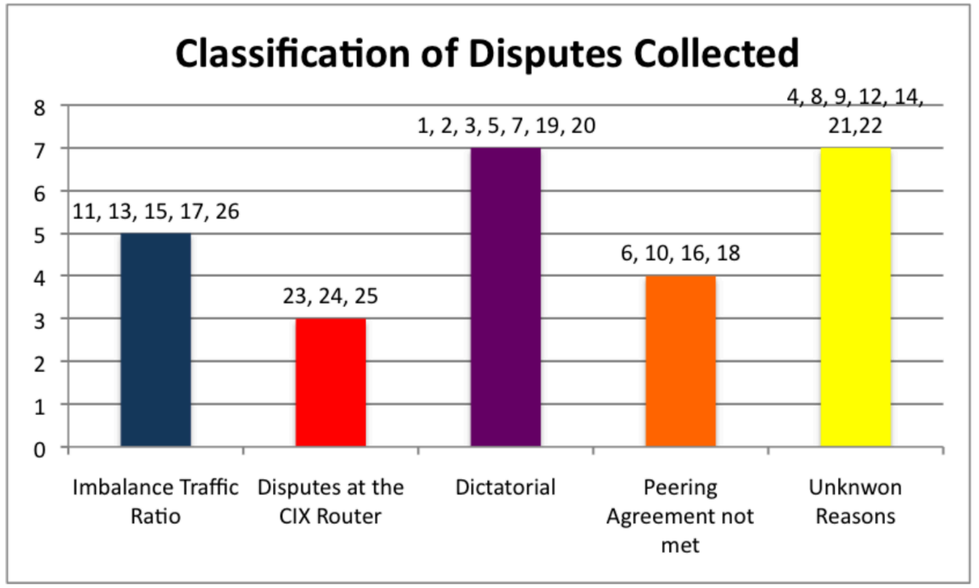
The 8 U.S. Interconnection Regions



Peering Disputes

What happens if Tier-1s De-Peer?

"Cogent has decided not to exchange traffic directly with TeliaSonera's AS 1299 or indirectly with AS 1299 through a third-party provider," Telia told its customers. "As a result, Cogent has partitioned the Internet and disrupted the flow of traffic between Cogent and TeliaSonera customers."



Source: Anatomy of the Internet Peering Disputes

Flattening

The Internet is Flat: Modeling the Transition from a Transit Hierarchy to a Peering Mesh

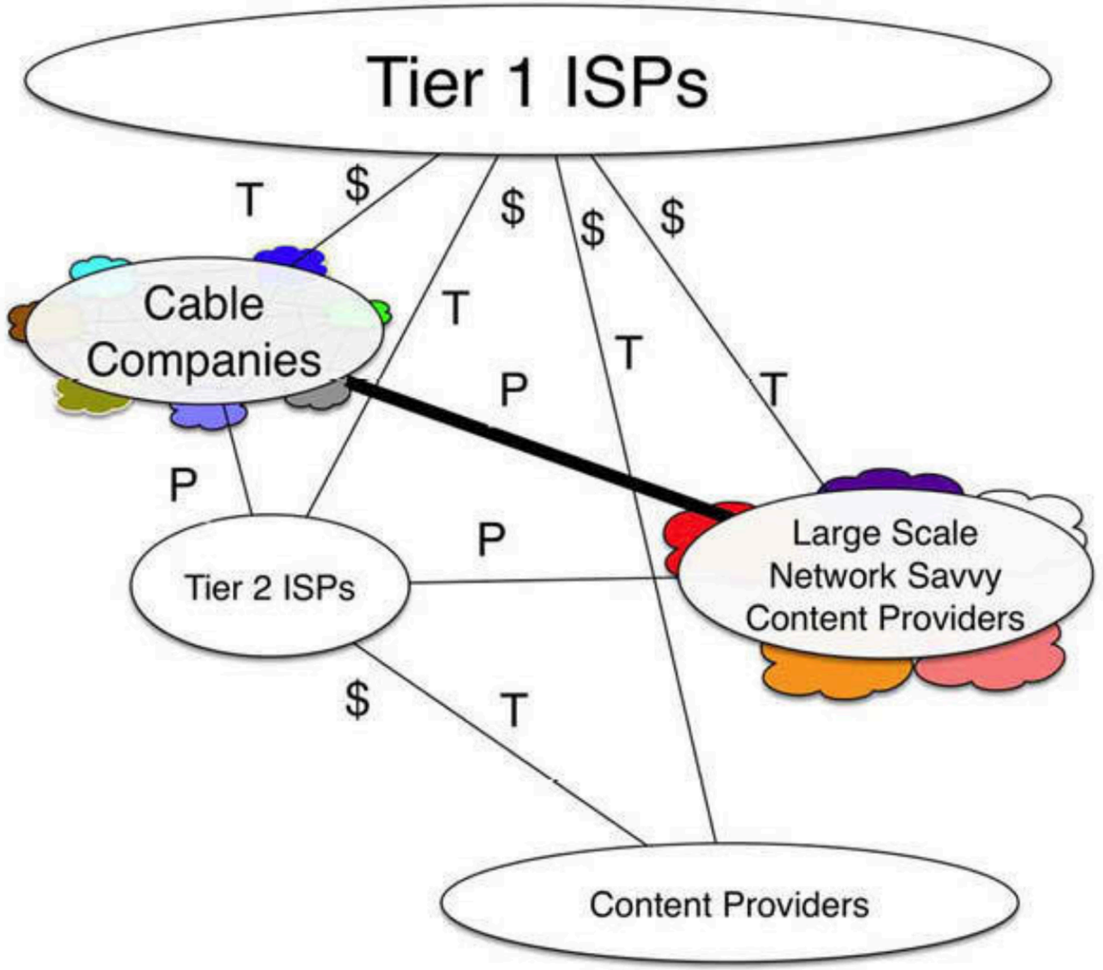
Three Changes in the World:

- An increasing fraction of Internet traffic originates from a few CPs or CDNs (e.g., Google, YouTube, Akamai, Cloudflare). This shift is due to the large penetration of video streaming.

- The major CDNs and CPs have expanded to almost every region of the developed world, so that they can be co-located with many ASes at Internet Exchange Points (IXPs).

- IXPs have increased rapidly in number, making it easy and cheap for an AS to establish peering links with other ASes co-located at the same IXP.

Tier 1 Peerings Mattered Less



ars TECHNICA BIZ & IT TECH SCIENCE POLICY CARS GAMING & CULTURE

POLICY —

How Comcast became a toll-collecting, nuke-wielding hydra

Comcast wants cash to deliver cached Netflix traffic to its subscribers. Has ...

NATE ANDERSON - 11/30/2010, 1:35 PM

A Small Transit Provider Case Study

AS19653 – Small Transit Provider in Climax, Michigan
Founded in 1911 as an Independent Telephone Company.
Started as a CLEC in 1996.
Independent ILEC-CLEC-ISP. CLLI = CLMXMIXI

2011 – Joined NANOG

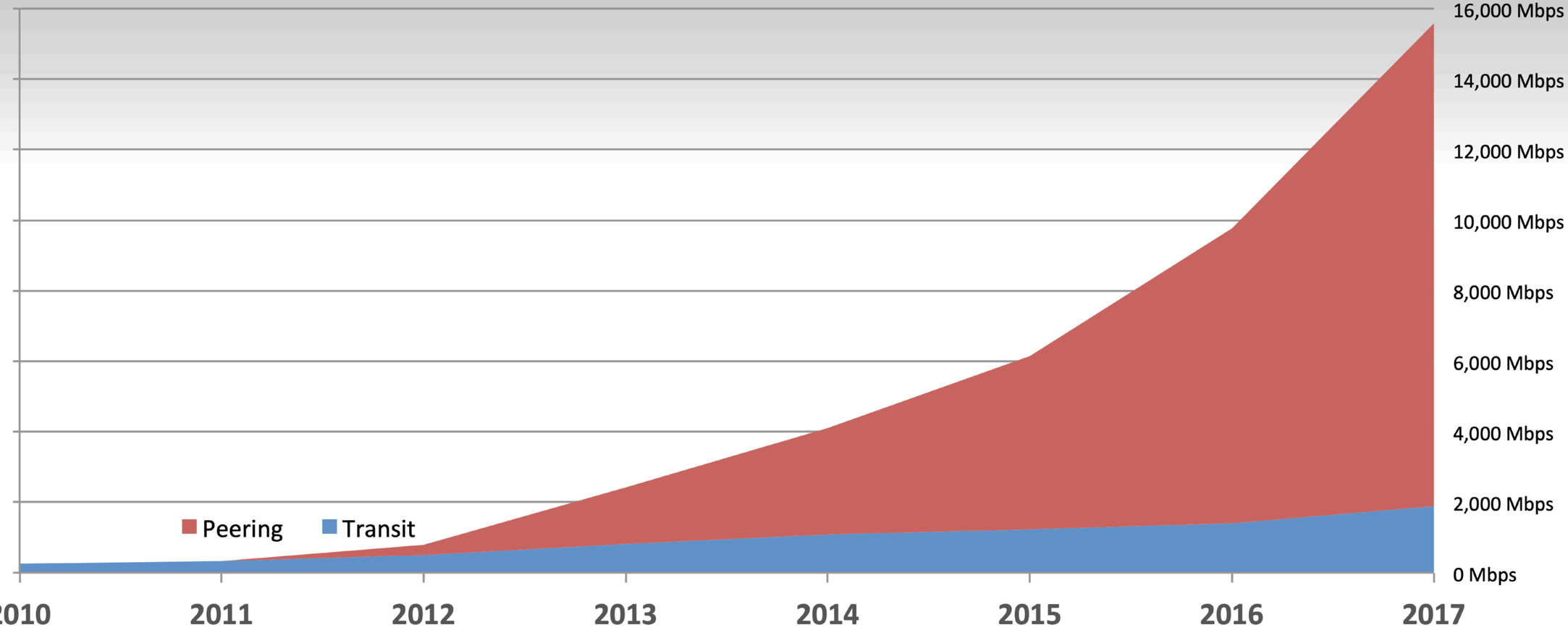
Telephone Company (ILEC-CLEC)
Tier 3 ISP
100% transit (two OC-12s)

2017 – (after 18 NANOGs)

Packet Optical Service Provider
Tier 2 ISP
88% Peering
12% transit
More than 100G in upstream ports

Source: The Internet is Flat: Revisited

AS19653 Evolution of Transit to Peering

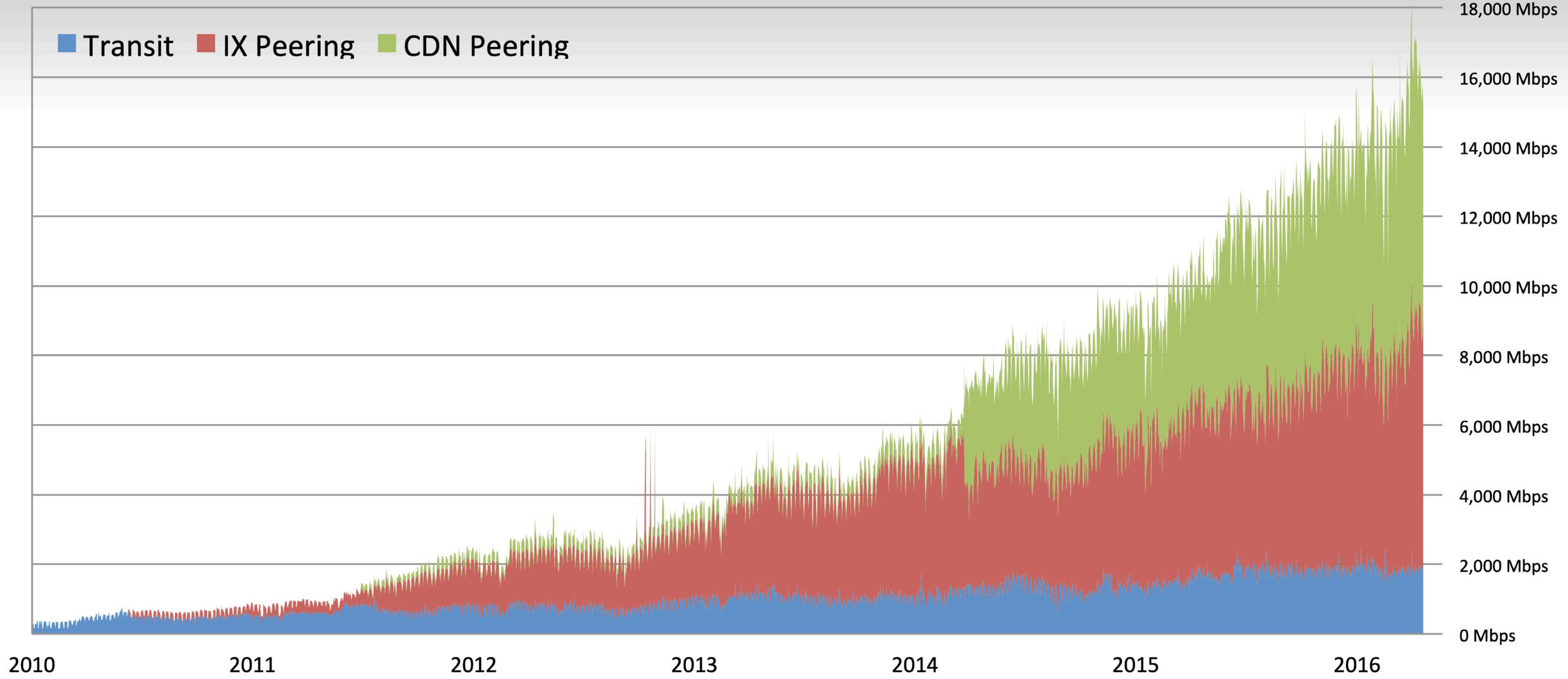


	2010	2011	2012	2013	2014	2015	2016	2017
Peering	0%	0%	36%	68%	74%	80%	86%	88%
Transit	100%	100%	64%	32%	26%	20%	14%	12%

A Small Transit Provider Case Study

SOURCE: AS19653

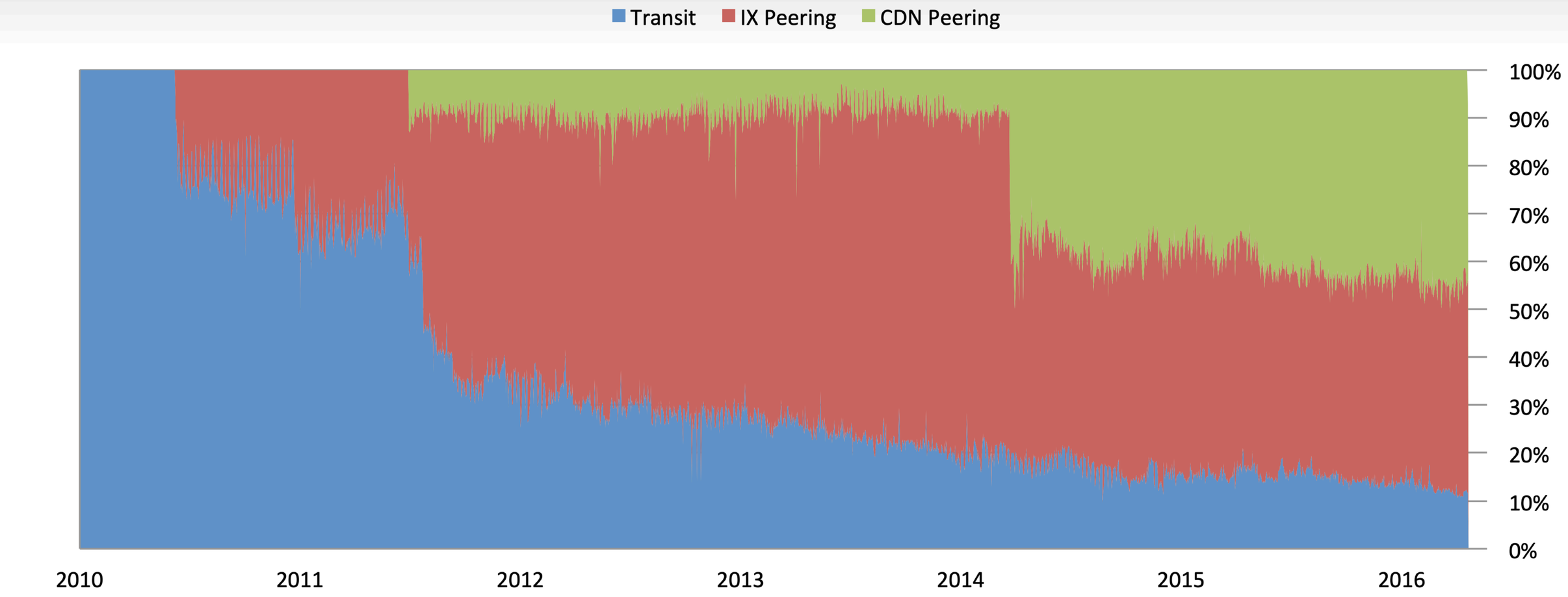
AS19653 From Transit to IXP Peering to CDN



A Small Transit Provider Case Study

SOURCE: AS19653

Percentage of Total Traffic AS19653 Transit/Peering/CDN



A Small Transit Provider Case Study

SOURCE: AS19653