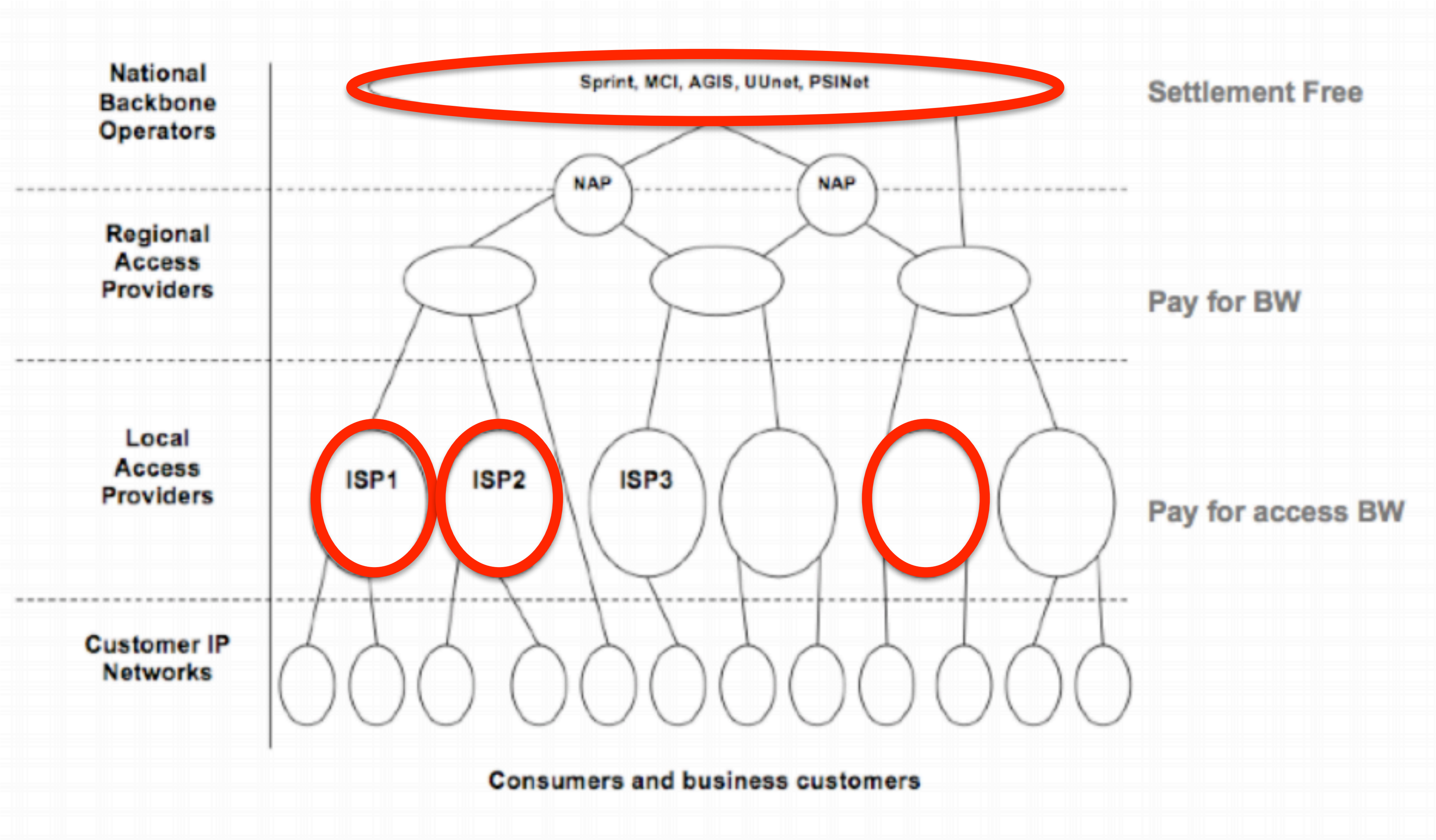
Internet Measurement
& Analysis (IMA)

Measuring the Expanding Internet Topology

ISP and IXP Topologies

Prof. Georgios Smaragdakis, Ph.D.

The Textbook Internet

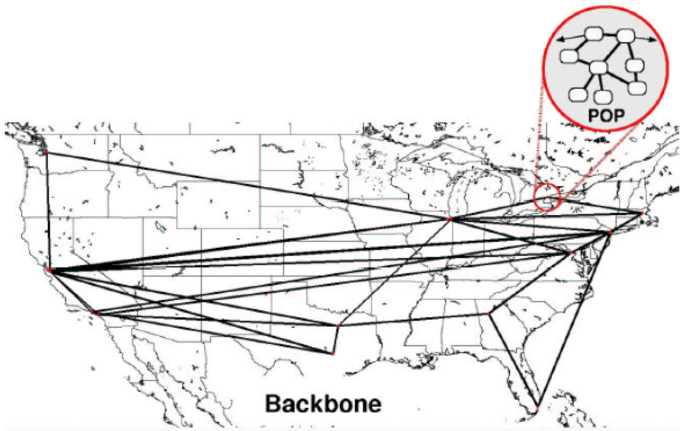


Rocketfuel: Measuring ISP Topologies

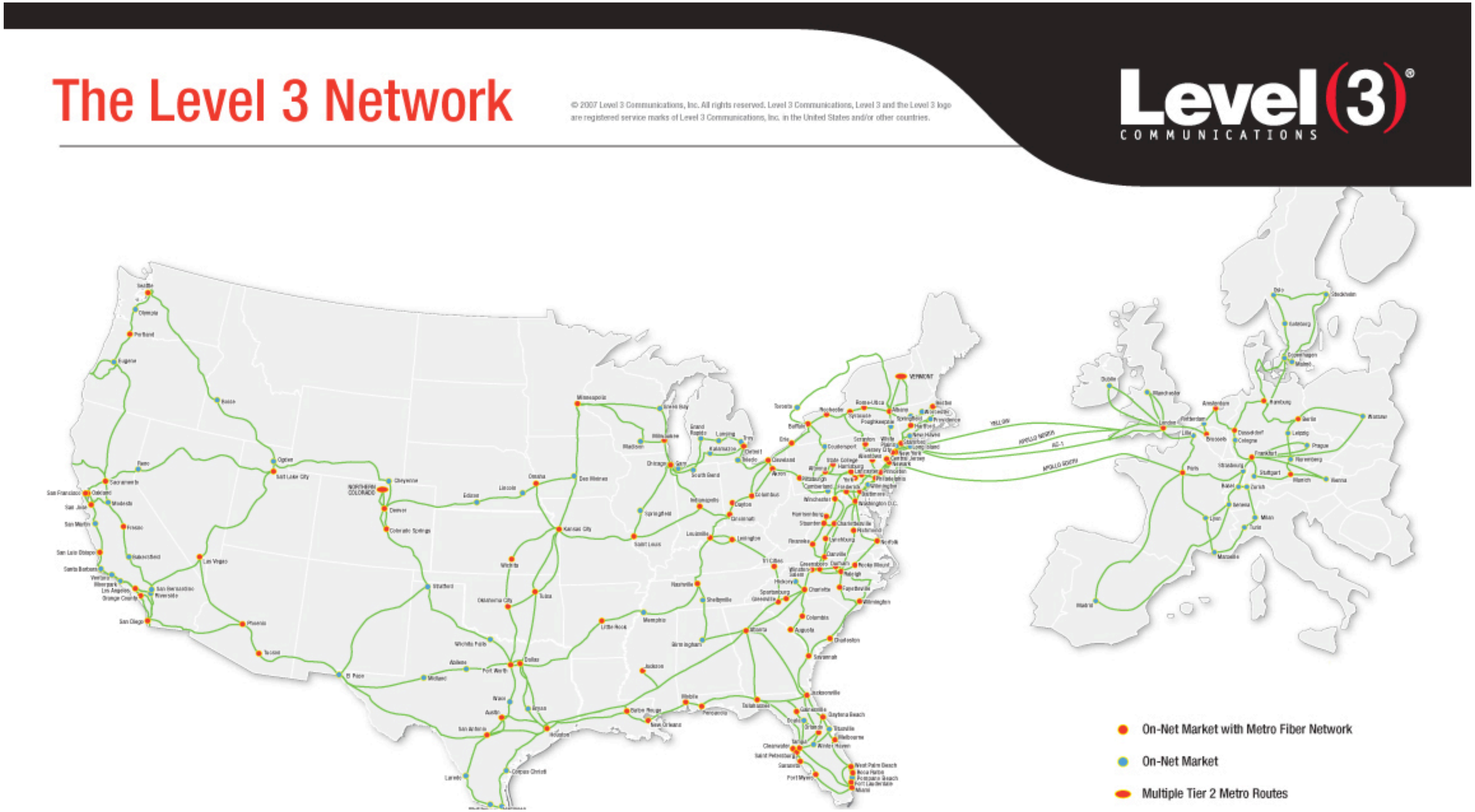
- To understand Internet structure and design
  - How ISP router-level topologies are designed?
  - Are there common best practices?
  - Assessment of security and resilience
- Real maps is a business secret
  - Backbone maps often available in "marketing" form
  - Severely lacking in router-level detail (abstract topologies)

Terminology

- Each POP is a physical location where the ISP houses a collection of routers.
- The ISP *backbone* connects these POPs, and the routers attached to inter-POP links are called *backbone* or *core* routers.
- Within every POP, *access* routers provide an intermediate layer between the ISP backbone and routers in neighboring networks.



Example: Level3 ISP



Rocketfuel Methodology

- ISPs release "helpful" information:
  - BGP – list of prefixes that are served
  - Traceroute – list of paths
  - DNS - where routers are and what they do
- Build detailed maps:
  - Backbone
  - POPs
  - Peering links

Selecting Measurements

- Publicly available traceroute servers (see at end of this presentation)
- Challenge: To build accurate ISP maps using few measurements
- Brute Force Method (2002)
  - 784 vantage points to 120K allocated prefixed in BGP table (700K in 2019)
  - Queried every 1.5 minutes: 125 days to complete a map (around 700 days)

Optimization 1: Directed probing

- Use vantage point inside the network and are customers
- Capitalize on routing information
  - Identify traceroutes which transit the ISP network

1.2.3.0/24	13 4 2 5
	6 9 10 5
	11 7 5
4.5.0.0/16	3 7 8
	7 8

Target ASN: AS 7

Dependent Prefixes: 4.5.0.0/16

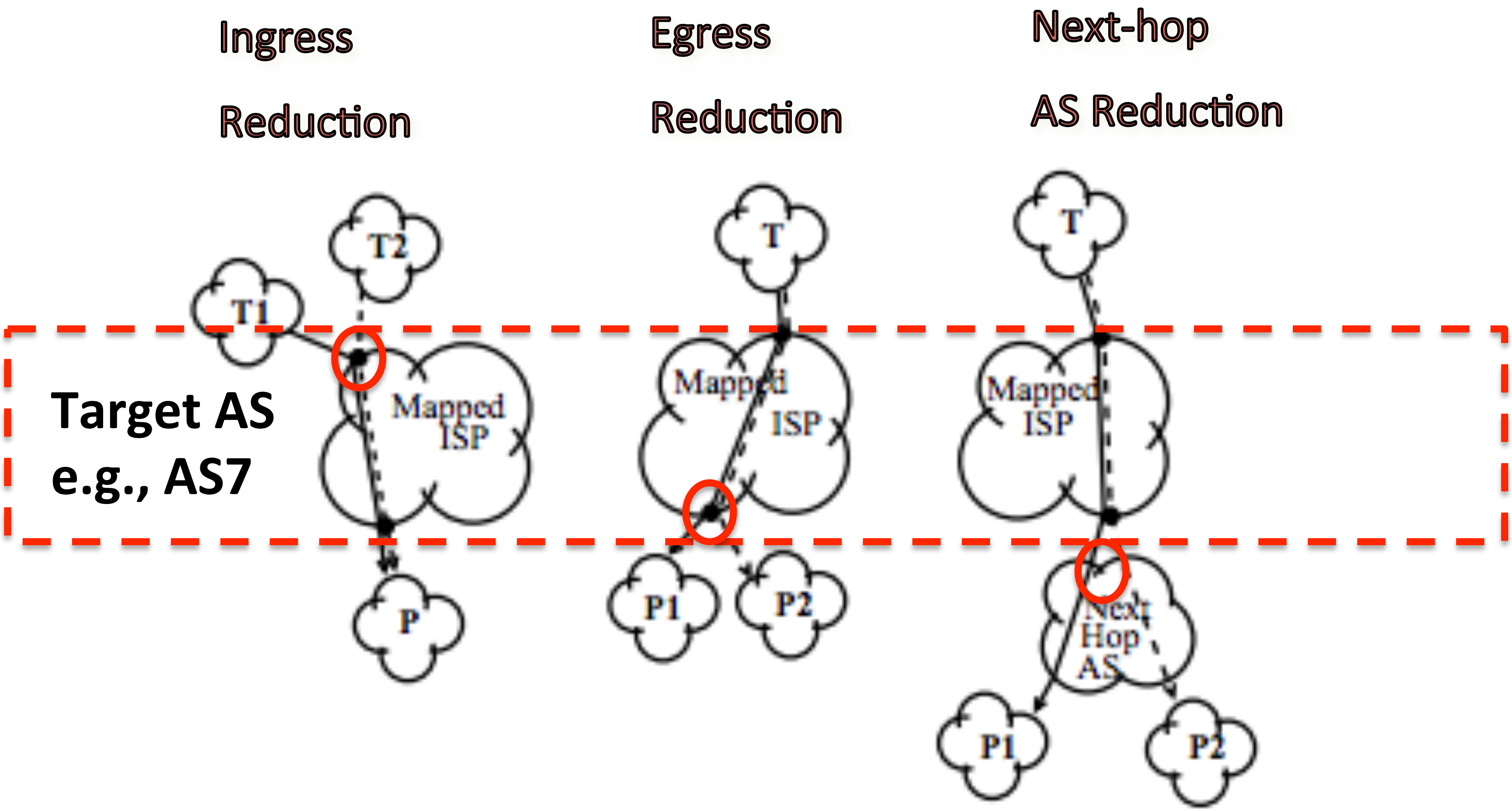
Insiders : 4.5.0.0/16

Up/down traces: AS 11 to 1.2.3.0/24

Optimization 2: Path Reduction

- T1 and T2 enter the ISP at the same point on the way to the same destination

- Paths to P1 and P2 leave the ISP at the same point



Reduction Effectiveness – Across Topologies

- Brute force : 90Million - 150 Million traceroutes required
- BGP directed probes : 200K - 15 Million traceroutes required
- Executed after path reduction : 8K - 300K traceroutes required!!

Location and Role Discovery

- Where is this router located?

  use DNS names

  S1-bb11-nyc-3-0.sprintlink.net is a Sprint router in New York City

  use connectivity information

  if a router connects only to router in Seattle, it is in Seattle

- What role does this router play in the topology?

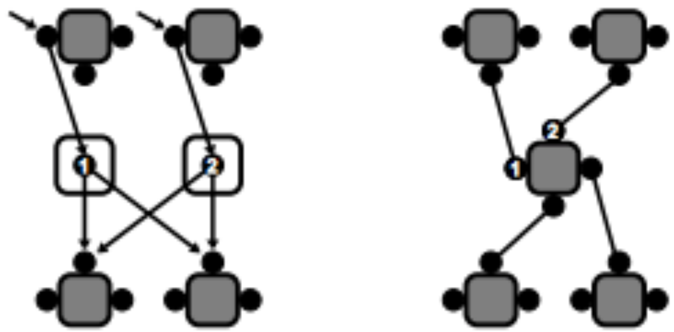
  only backbone routers connect to other cities

  use DNS names

  s1-gw2-sea-3-1.sprintlink.net is a Sprint gateway router

The Alias resolution problem

- The problem of identifying which IP addresses belong to the same router



Interfaces 1 and 2 belong to the same router

Alias resolution solution: Ally

Approach:

Send a packet to each interface (potentially aliased IP) to solicit responses ("UDP port unreachable")

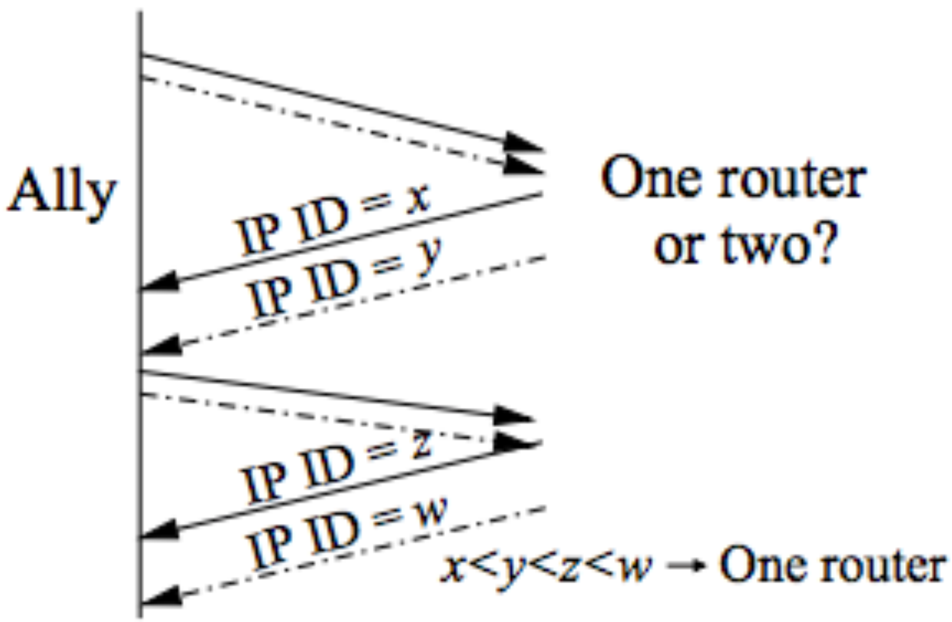
- UDP high port, TTL = 255

- Routers often set source address of the outgoing interface
- Responses have nearby IP identifiers:

IP-ID pair is commonly set from the same counter (incremented after generating a packet)

IP ID method (port unreachable resources)

- x<y<z<w, and w-x small → likely aliases
- If |x-y|>200, then Aliases are disqualified (third and fourth packets may not be sent)



Router Degree Distribution

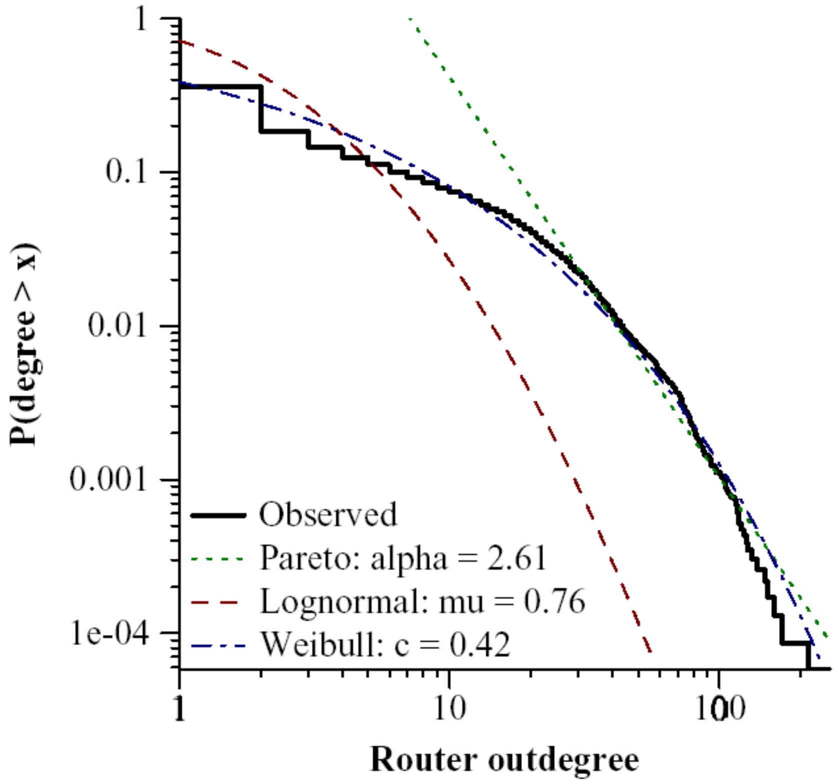
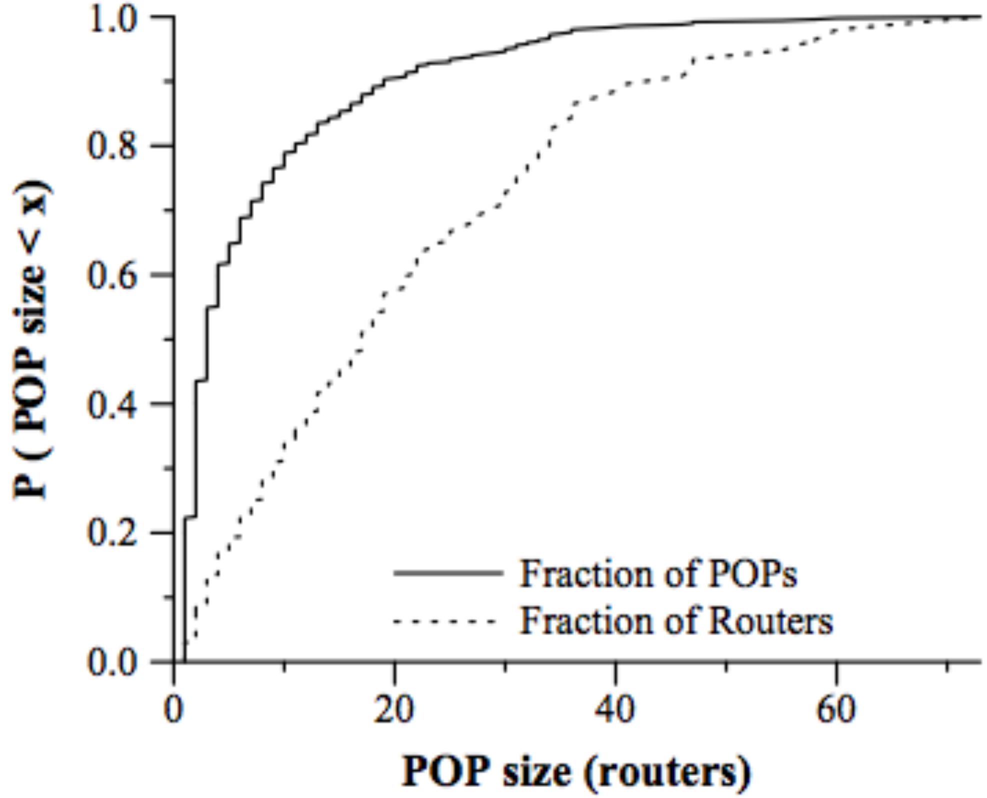


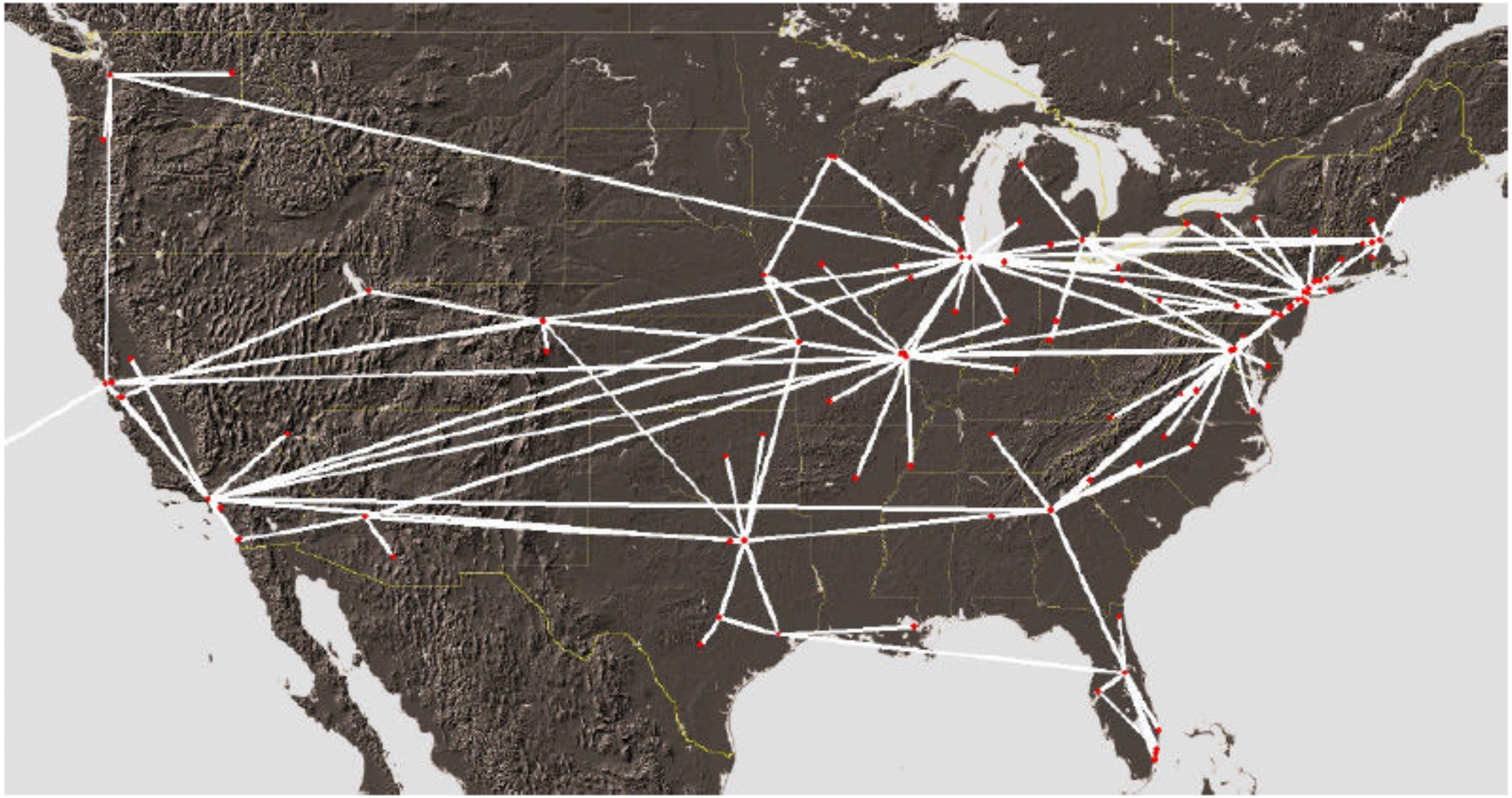
Figure 16: Router outdegree ccdf. The Pareto fit is only applied to the tail.

POP size CDF



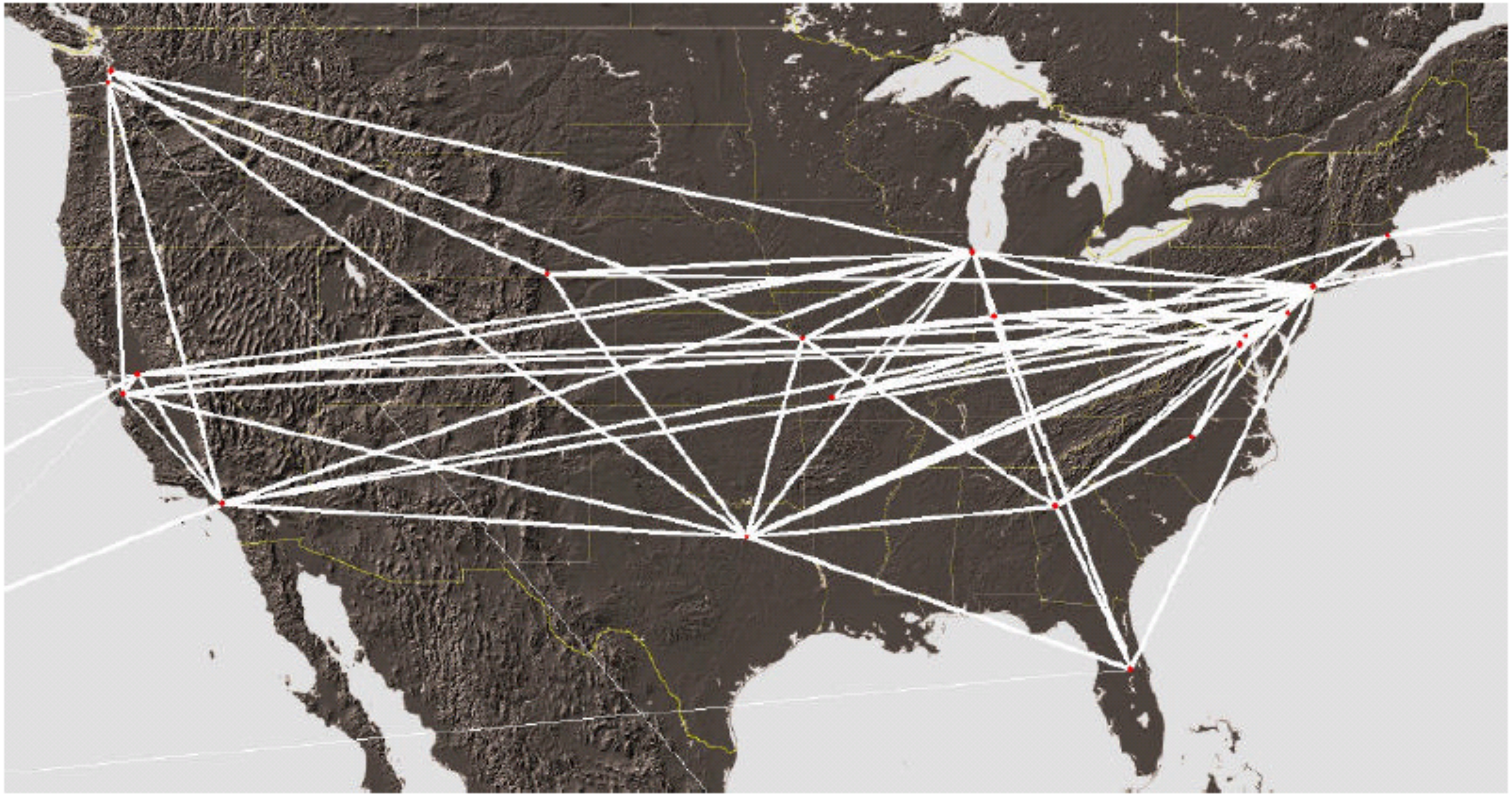
1.0
0.8
0.6
0.4
0.2
0.0
P (POP size < x)
0
20
40
60
POP size (routers)
Fraction of POPs
Fraction of Routers

Examples – AT&T



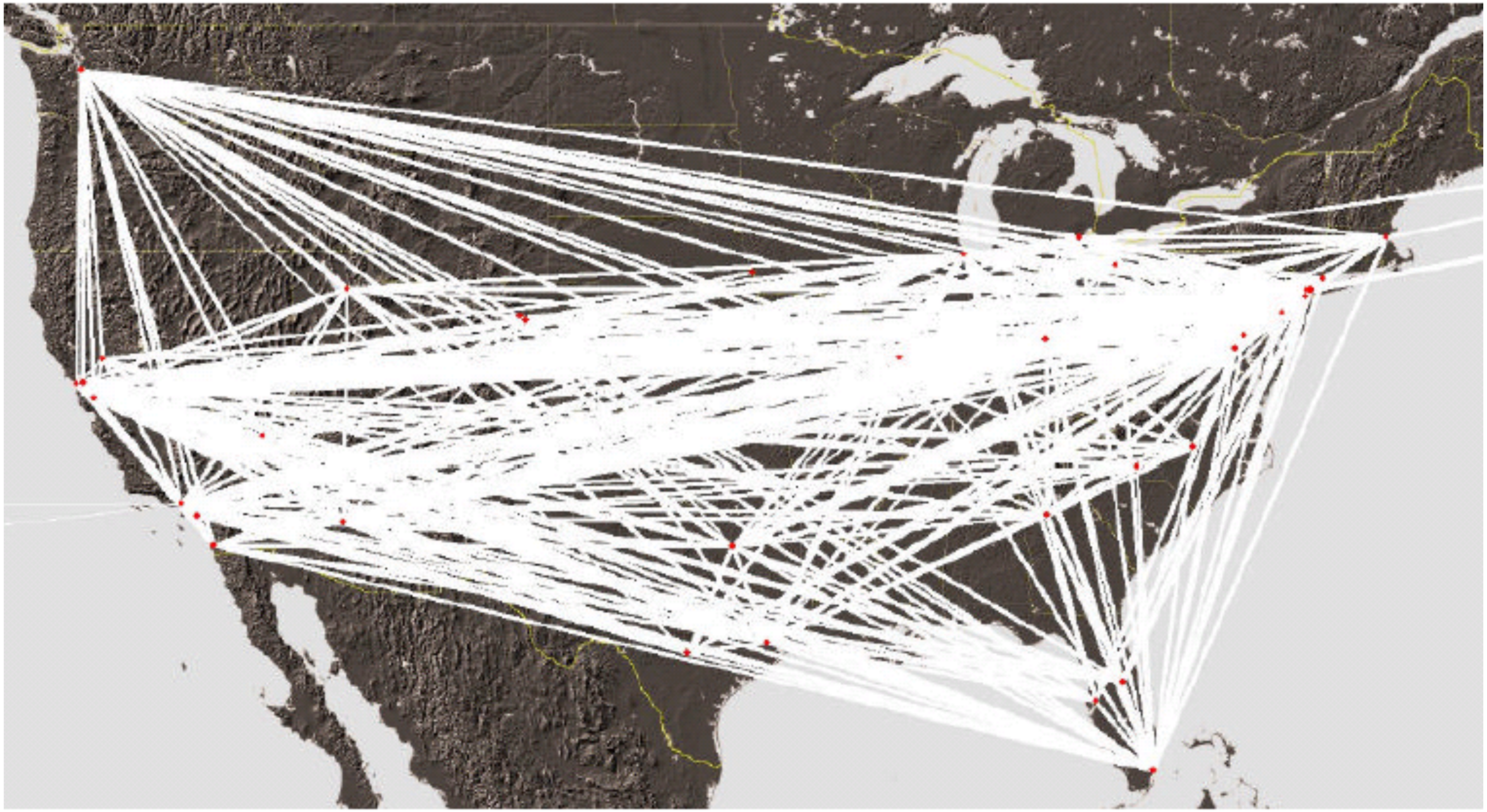
Background image courtesy JHU, applied physics labs

Examples – SPRINT

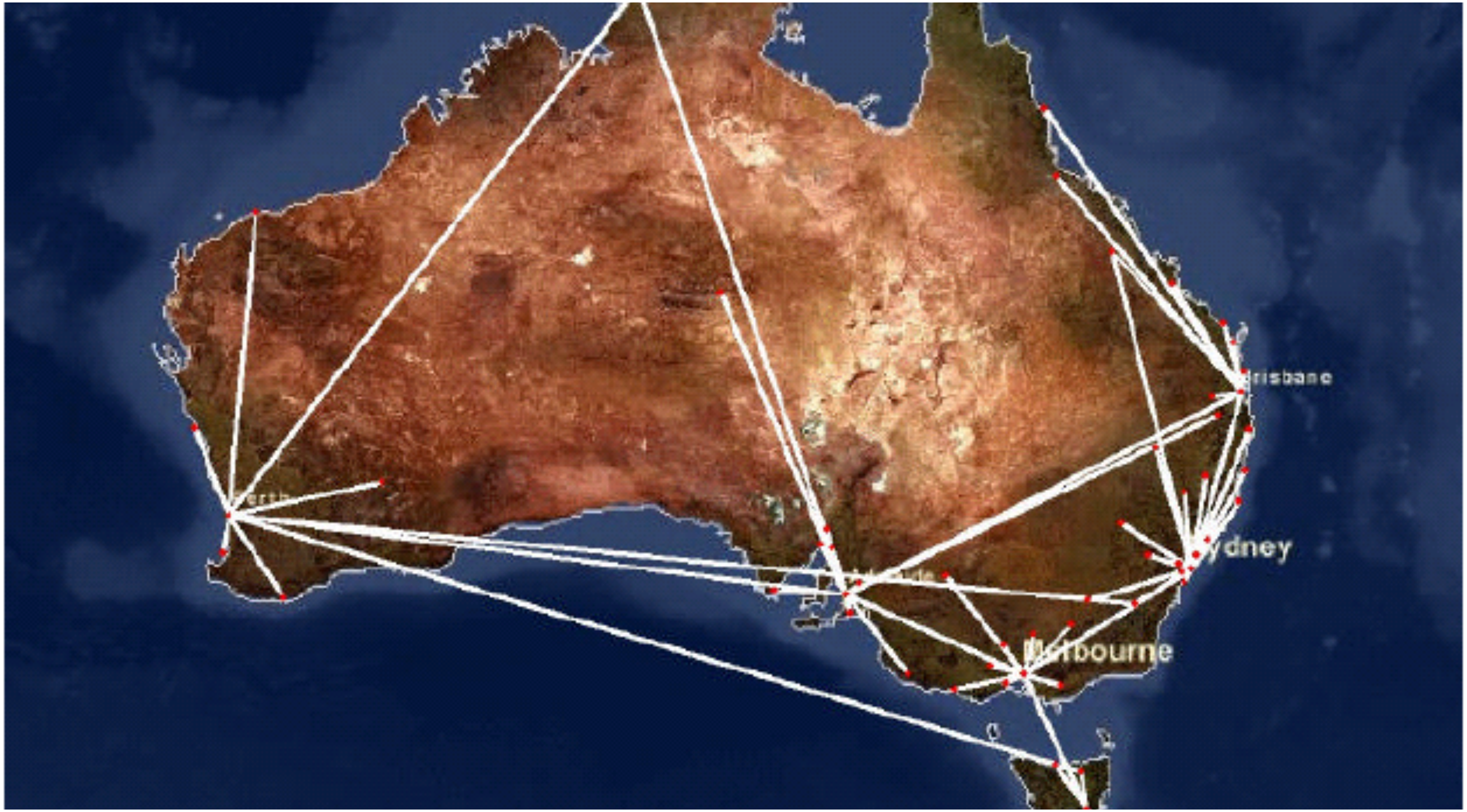


Background image courtesy JHU, applied physics labs

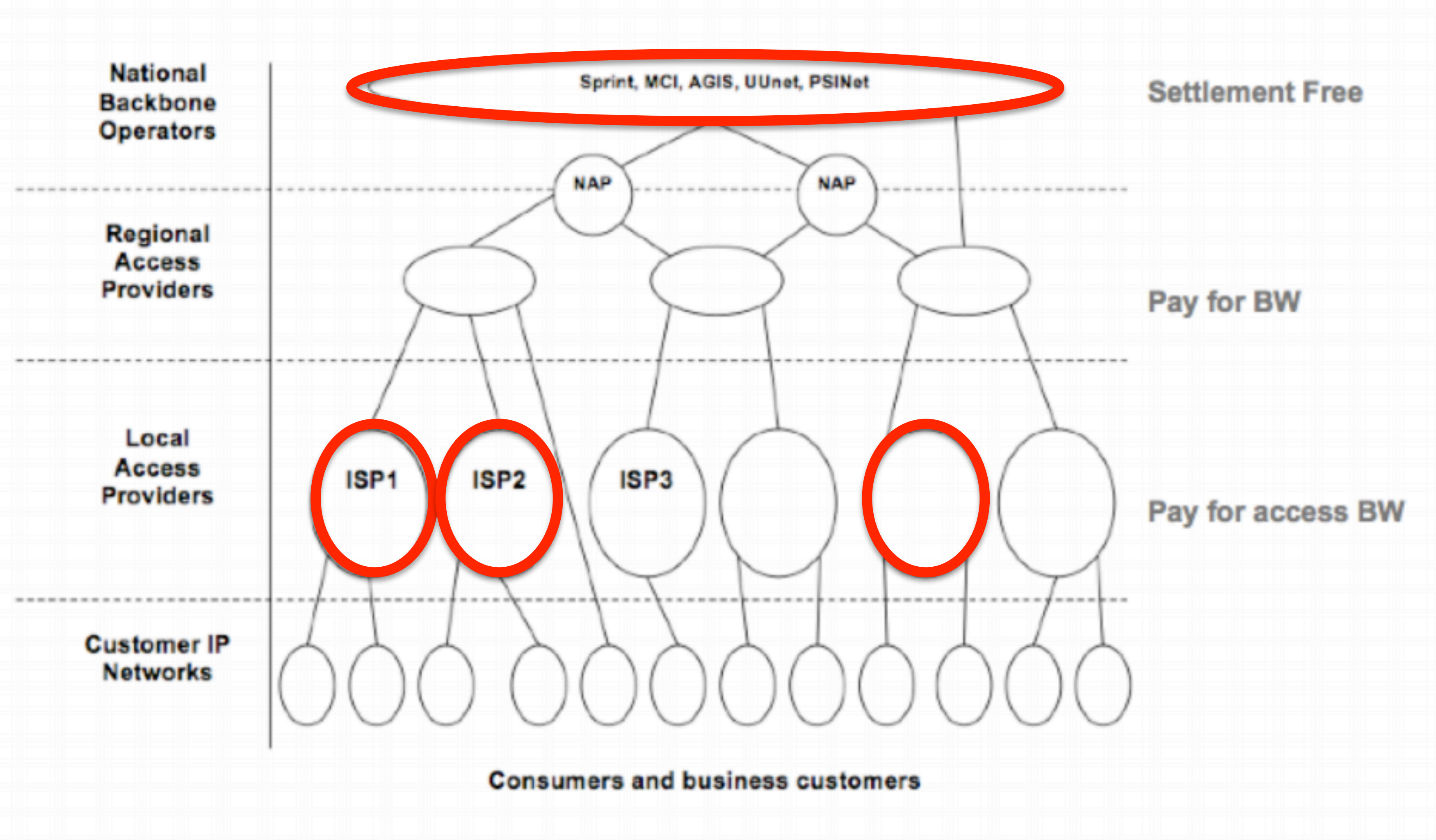
Examples – Level3



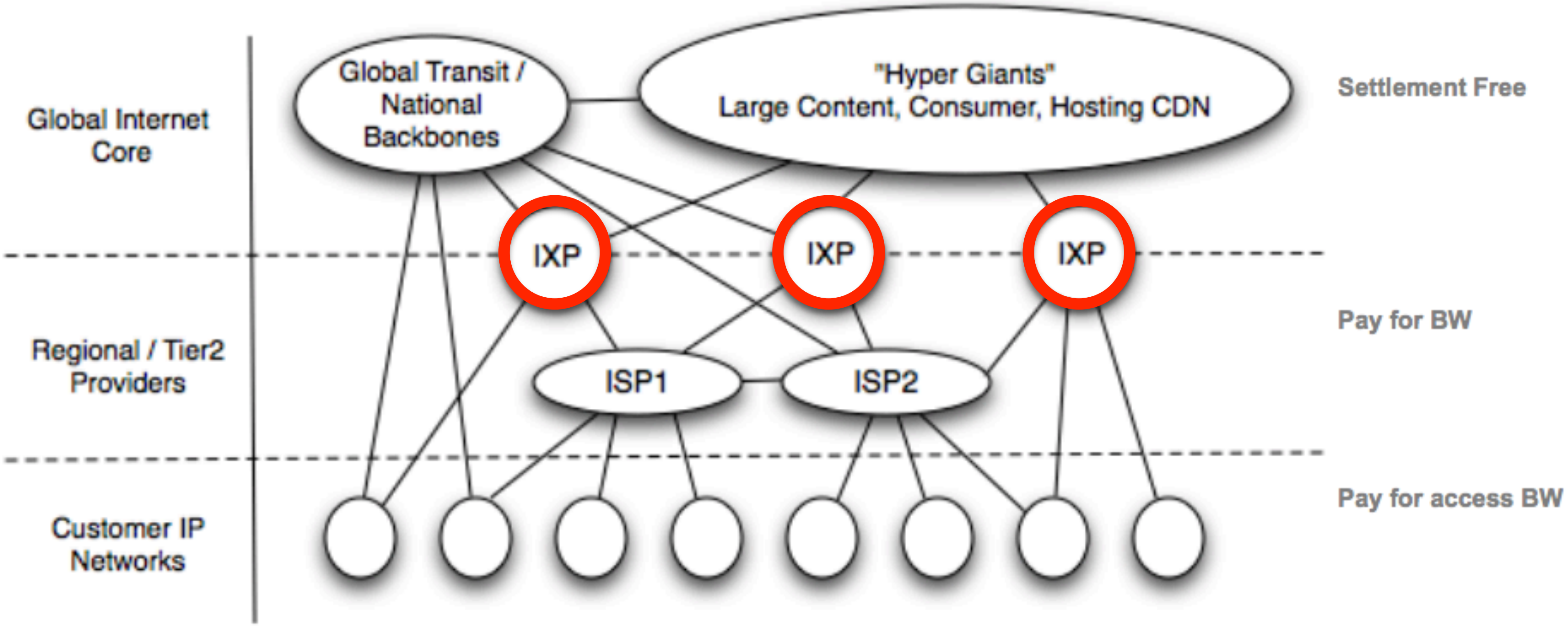
Examples – Telstra



The Textbook Internet



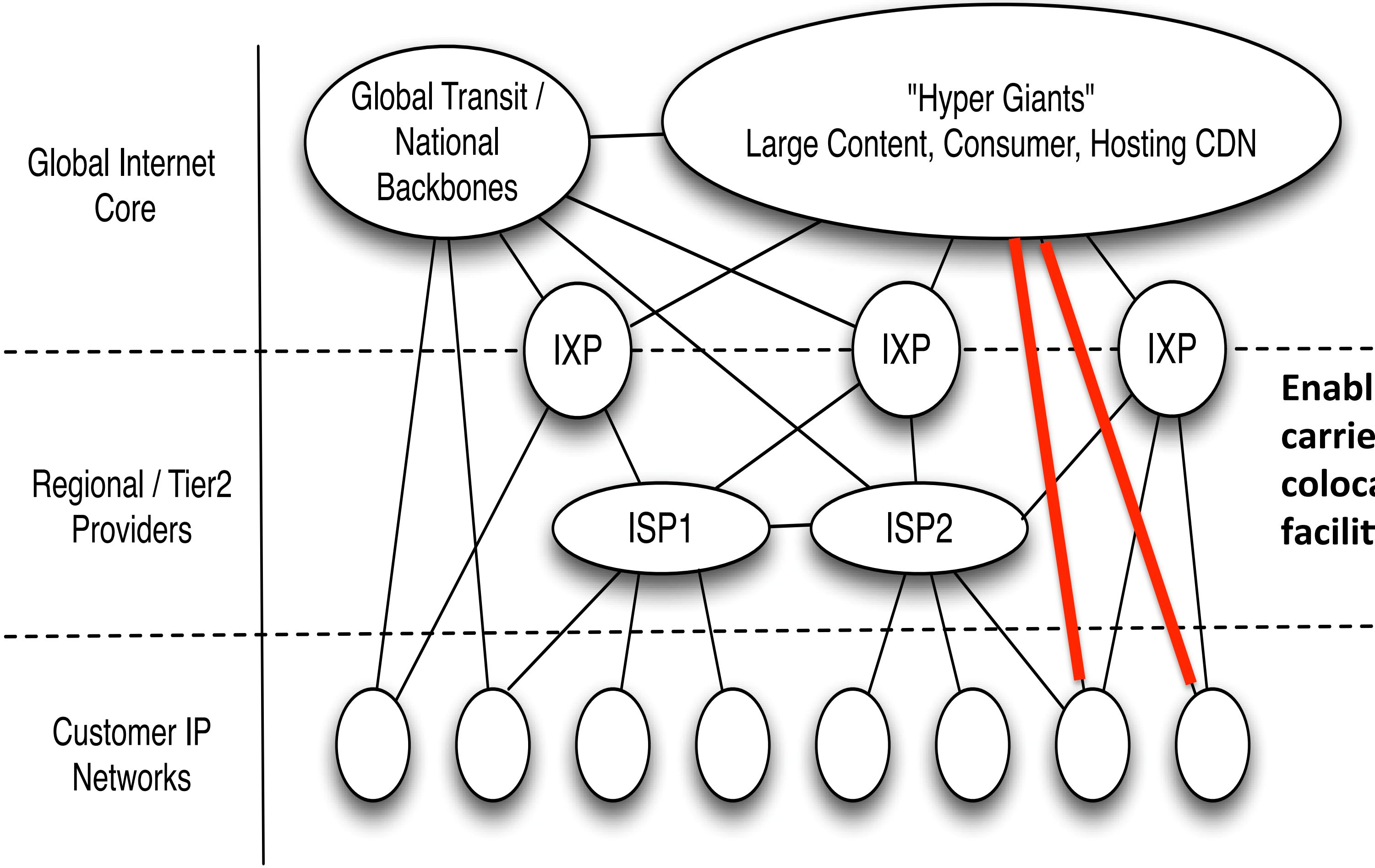
The New Internet



- Flatter and much more densely interconnected Internet
- Disintermediation between content and “eyeball” networks
- New commercial models between content, consumer and transit

Source: “Interdomain Internet Traffic” SIGCOMM 2010

The New Internet



Enabled by carrier-neutral colocation facilities

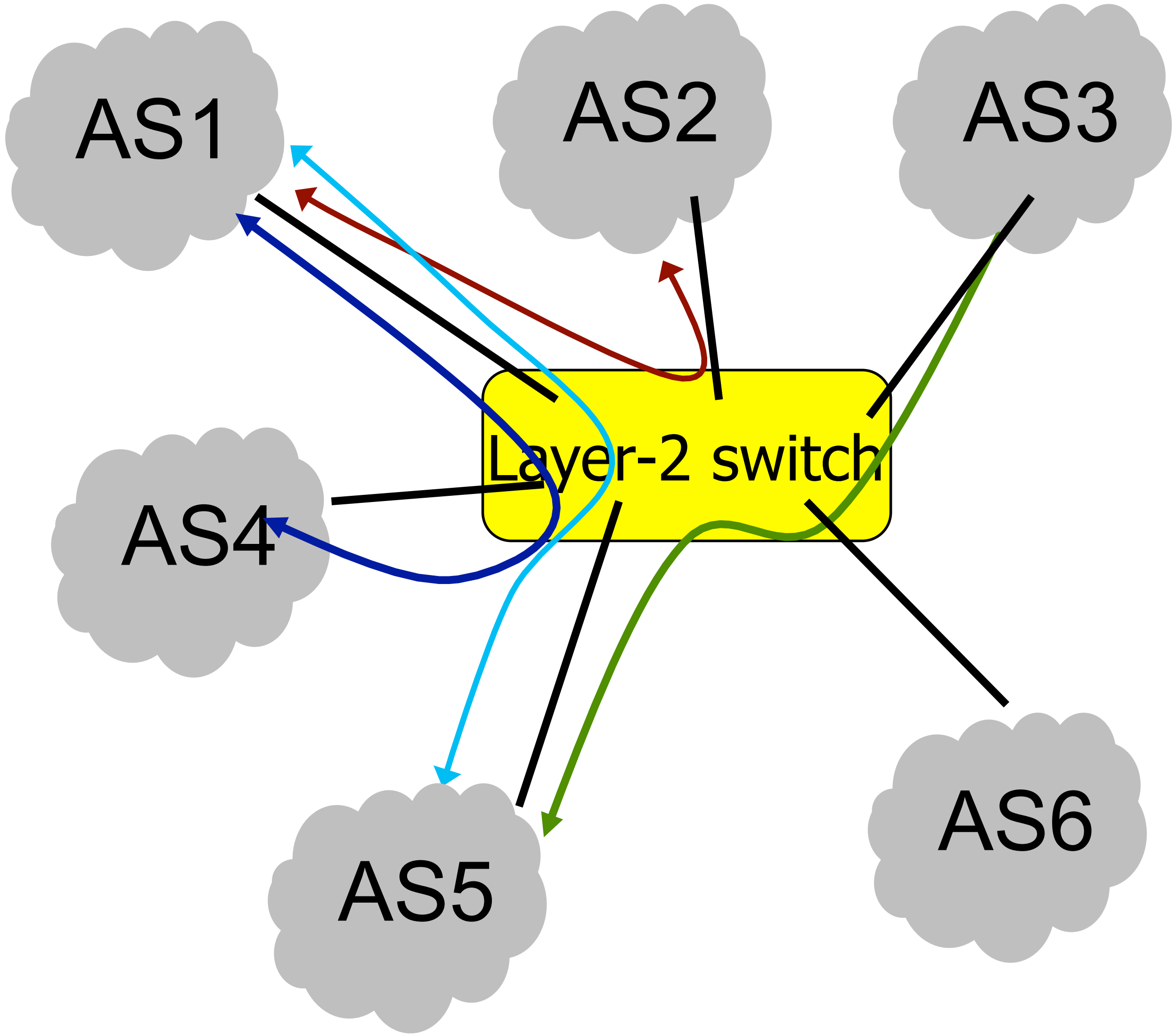
source: "Internet Interdomain Traffic", Labovicz et al. SIGCOMM 2010

Internet Exchange Point (IXP)

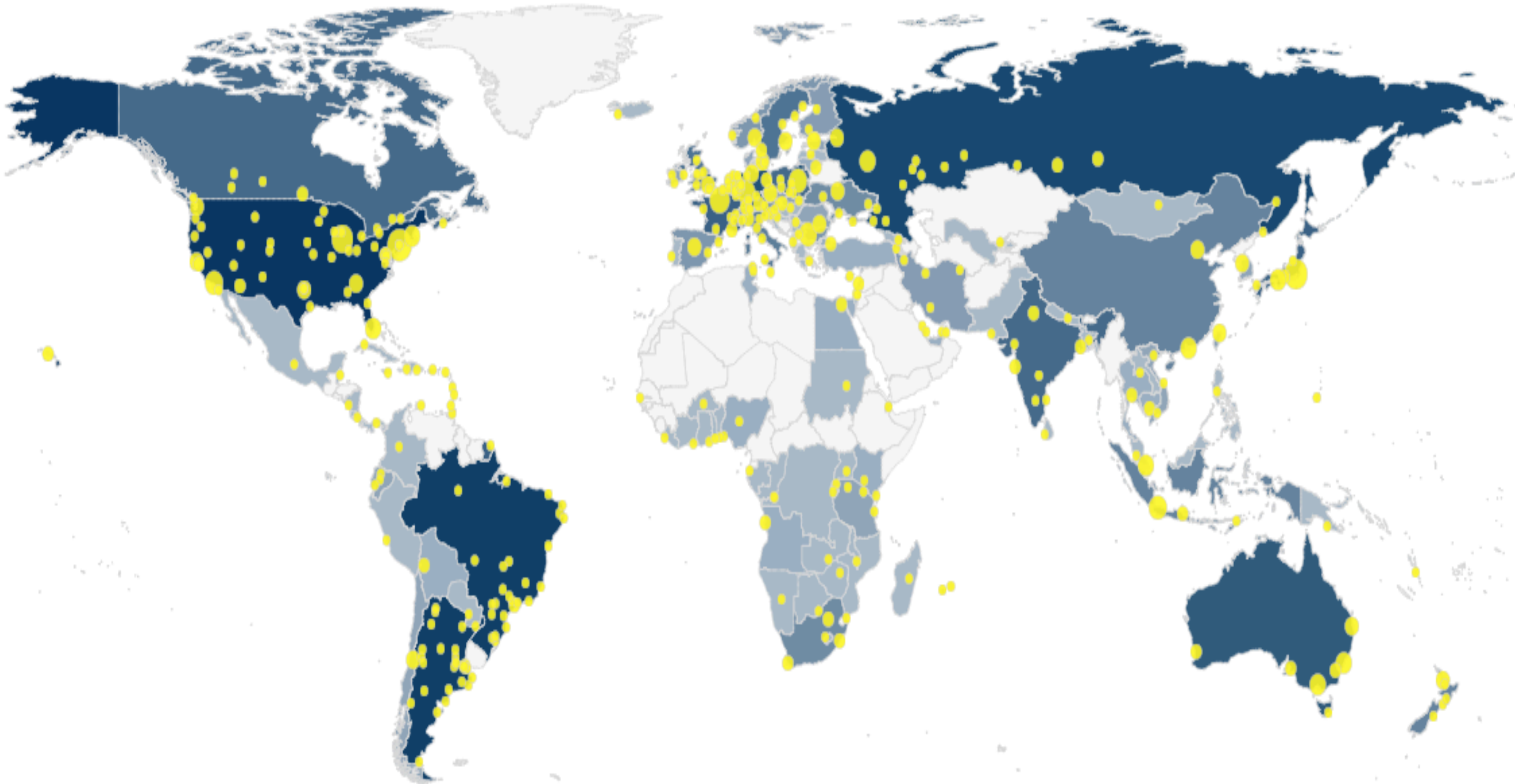
(IXPs) provide a shared switching fabric for layer-2 “Public” peering.

- Bilateral Public Peering: between two ASNs

- Multilateral Public Peering: using Route Server (see slides 35 and 36)



IXPs around the Globe



>**300** active IXPs

The Rise of the IXPs

Large IXPs (AMS-IX, DE-CIX, LINX) have over **800** members, support over **100K** peerings and carry up to **5 Tb/s** peak traffic (comparable to large Transit/ISPs) ["There is more to IXPs than Meets the eye", CCR 2014]

Estimated Peak Total traffic > **125 Tb/s** [Euro-IX 2016]

Estimation of more than **2 million peerings** are established using IXPs ["2016 Survey of Internet Carrier Interconnections", Woodcock et al. 2017]

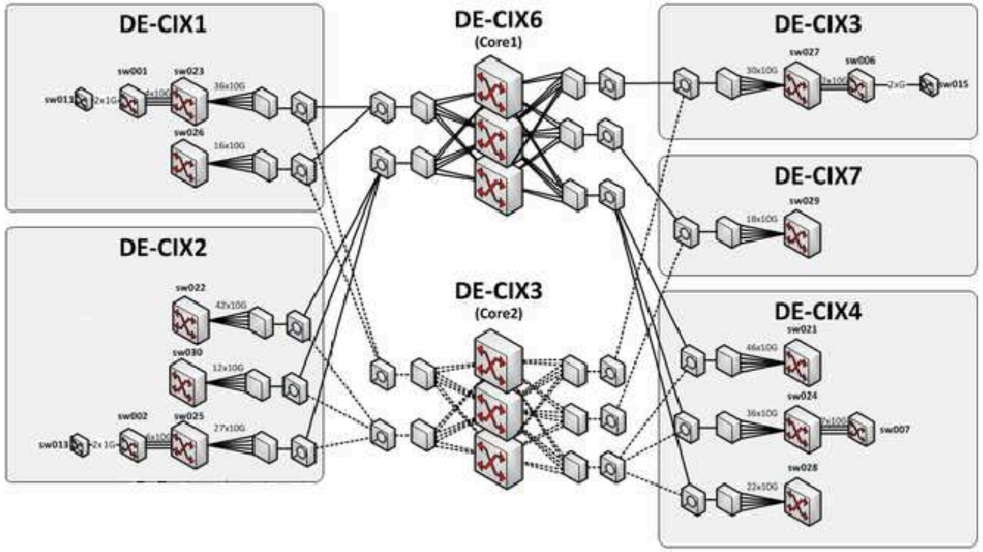
Large CDNs are present in tens of IXPs, e.g., Google and Akamai are in more than **100 IXPs** [PeeringDB]

IXP is more than a Big Switch

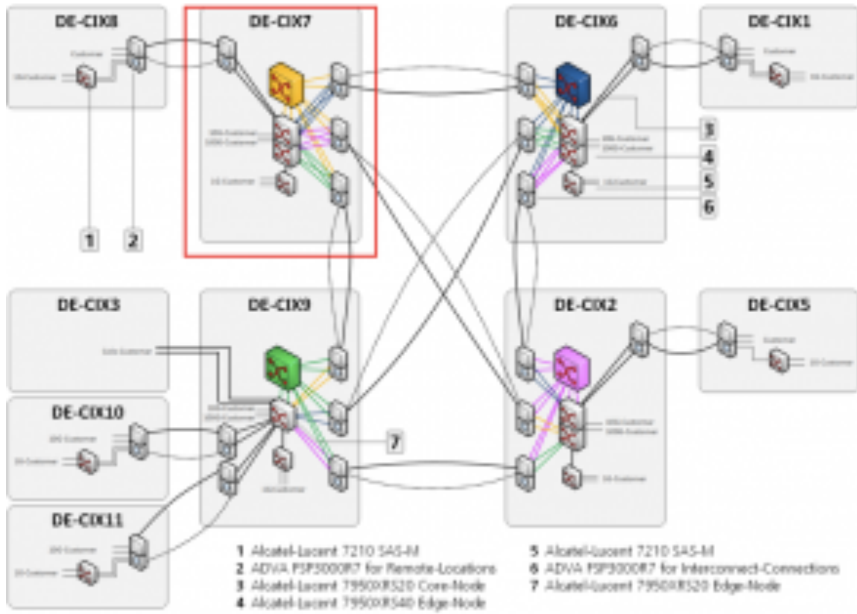
LINX (London Internet Exchange) in Telehouse Colocation Facility (Telehouse North at Docklands)

IXPs are affiliated with peering facilities and anable “private” peering via dedicated lines – **Private Peering Interconnect** (PNIs)

Topology of IXP: DE-CIX in Frankfurt

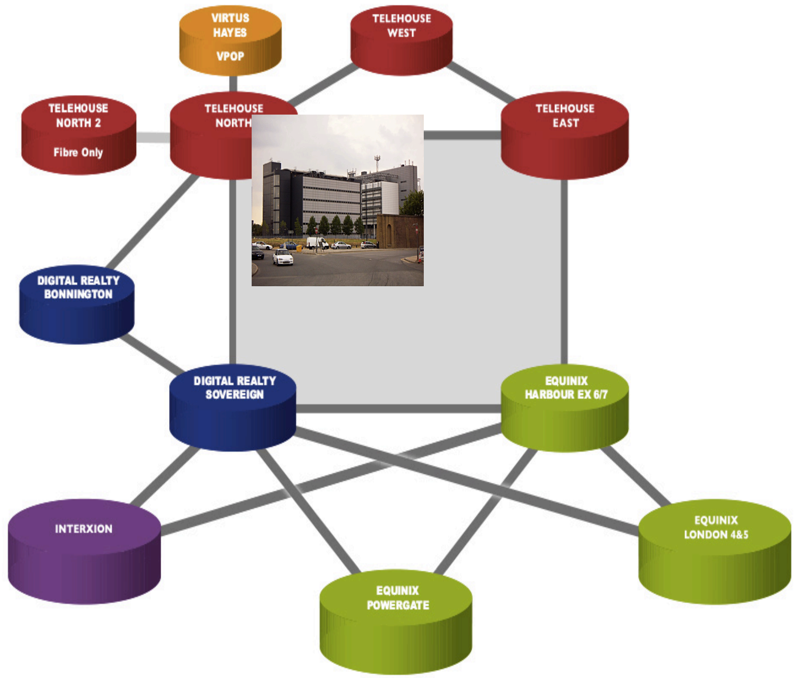


ca. 2012

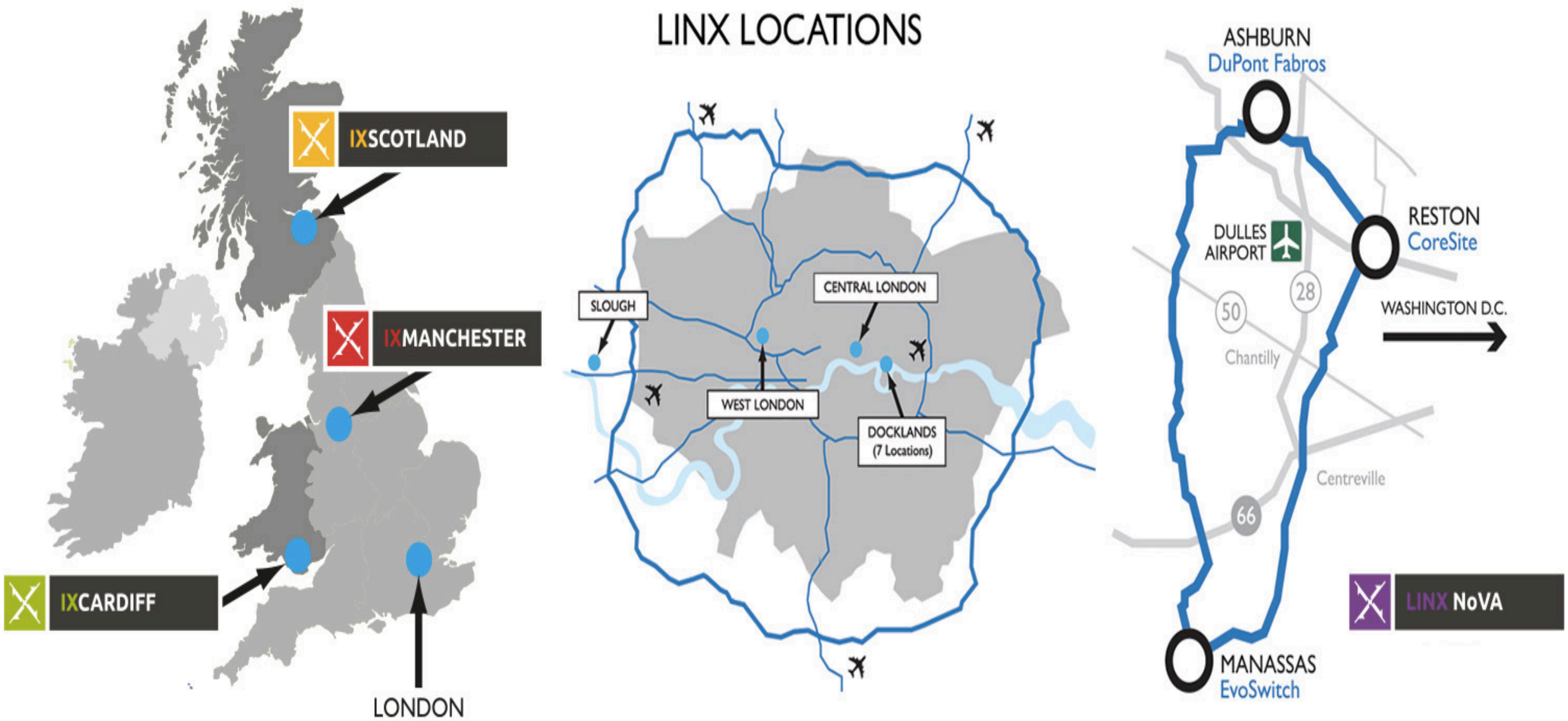


ca. 2018

IXPs in Metropolitan Areas: LINX and Colocations Facilities in London



IXPs' Global Footprint: LINX around



Colocation Facilities (CFs) around the Globe

Colocation facilities (carrier-neutral) provide infrastructure for physical co-location, peering, and cross-connect interconnections.

Today there are about **2,000** peering facilities; around **400** facilities with more than 10 members [PeeringDB]

Largest facility operators, e.g., Equinix, support > **170K** of cross-connects and interconnections

Peering Infrastructures are Critical Infrastructures

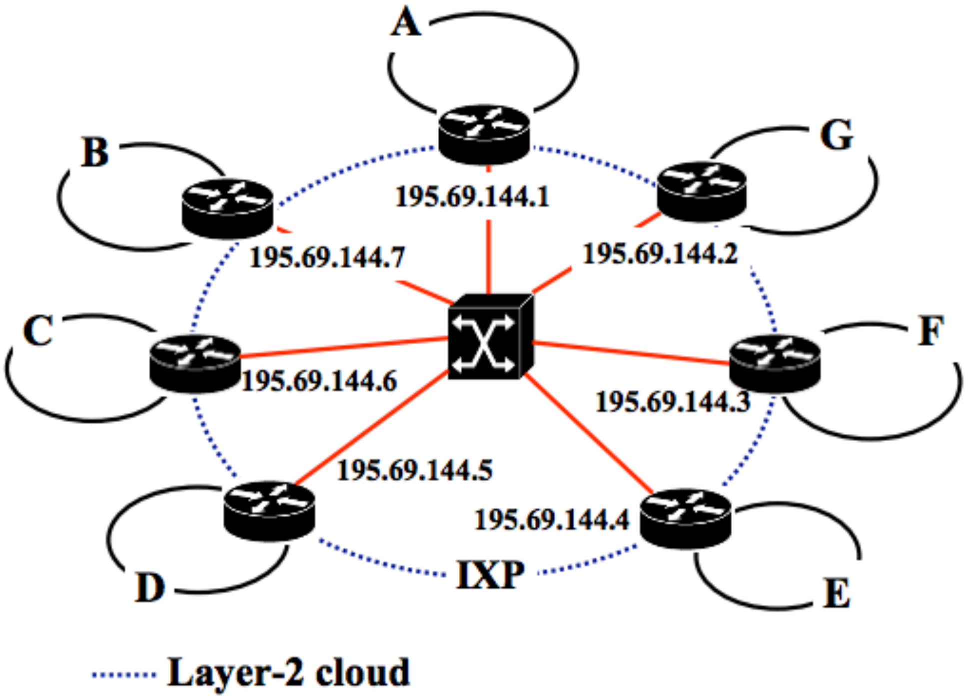
DHS and **ENISA** have characterized peering infrastructures as critical infrastructures – in the same category as nuclear reactors and power powerhouses. [An Annex to the National Infrastructure Protection Plan, 2010, 2015; Critical Infrastructures and Services, Internet Infrastructure: Internet Interconnections, 2010]

Internet Exchange Points: Typical SLA 99.99% (~52 min. downtime/year)[1]

Colocation facilities: Typical SLA 99.999% (~5 min. downtime/year)[2]

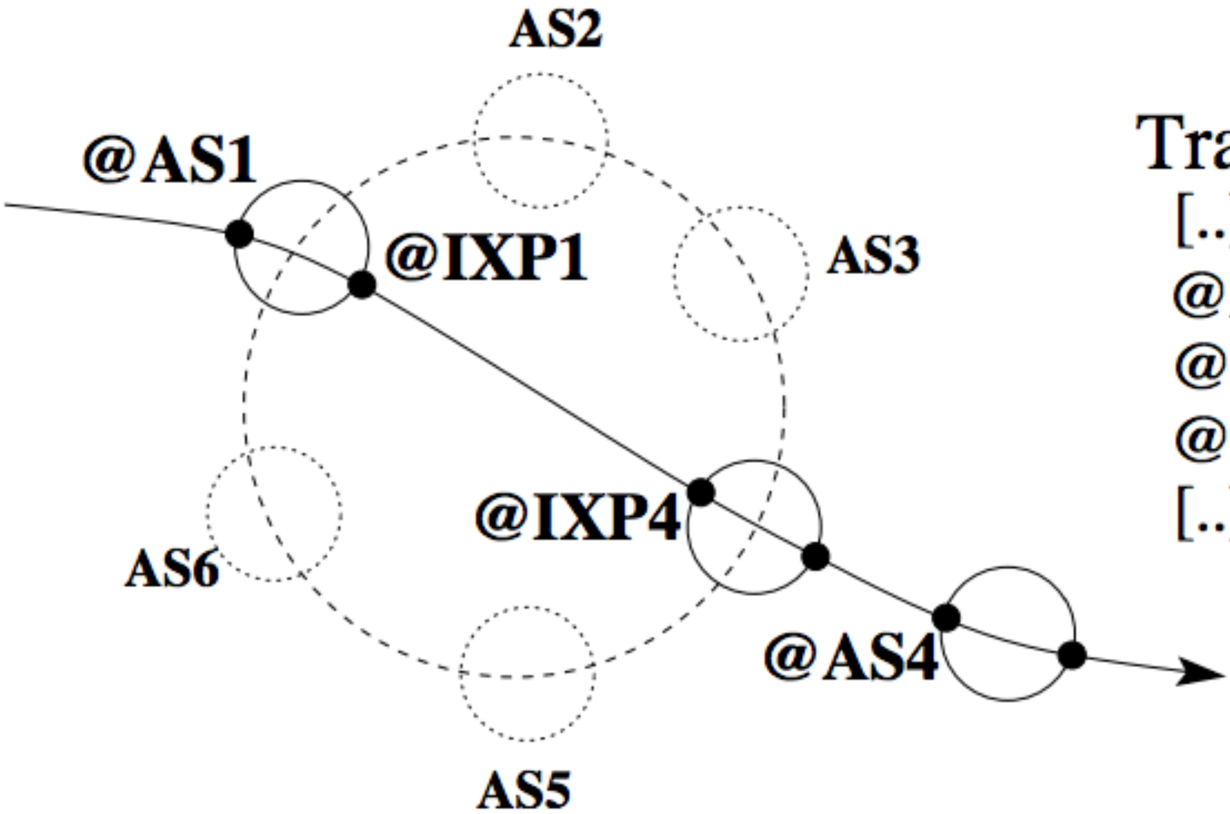
[1] https://ams-ix.net/services-pricing/service-level-agreement [2] http://www.telehouse.net/london-colocation/

Internet Exchange Points (IXPs)



Source: “The (in)Completeness of the Observed Internet AS-level Structure” IEEE/ACM ToN 2010

IXP Mapped?



Traceroute output:
[..]
@AS1
@IXP4
@AS4
[..]

Internet Exchange Points (IXPs)

Table 2: Overview of routing and looking glass datasets for November. The numbers show P-P links.

Dataset	Unique LGs / ASN	Visible links	only in this dataset
RV	78	5,336	1,084
RIPE	319	10,913	5,460
NP	723	3,419	684
RV+RIPE+NP	997	13,051	10,472
LG	821 / 148	4,892	2,313
RV+RIPE+NP+LG	1,070	15,364	15,364

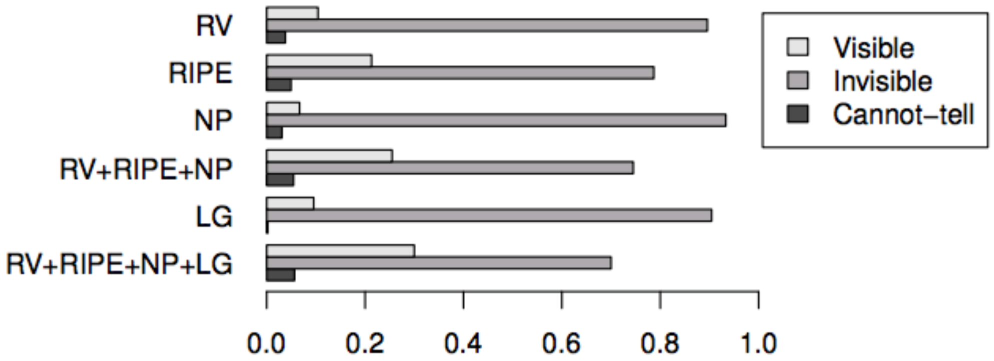
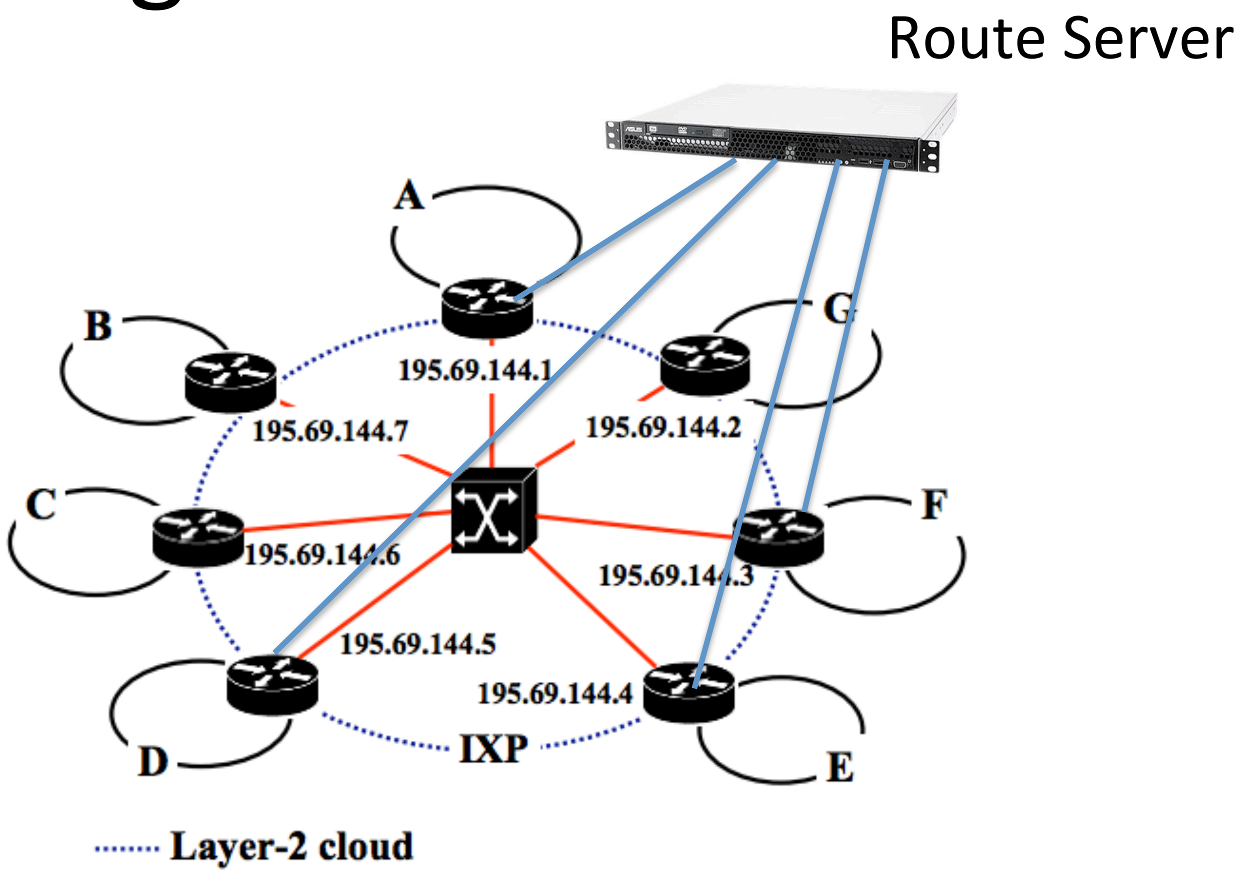


Figure 2: Peering links and visibility in control/data plane (normalized by number of detected P-P links).

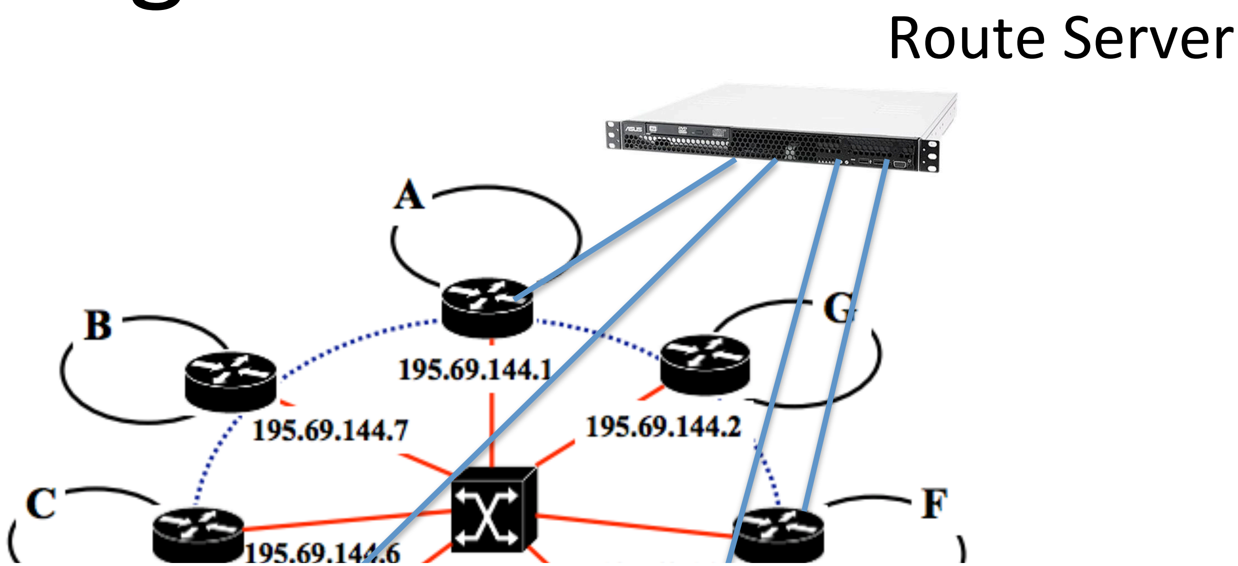
Source: "Anatomy of a Large European IXP" SIGCOMM 2012

Multilateral Public Peering at IXPs using the Route Server



Source: "The (in)Completeness of the Observed Internet AS-level Structure" IEEE/ACM ToN 2010

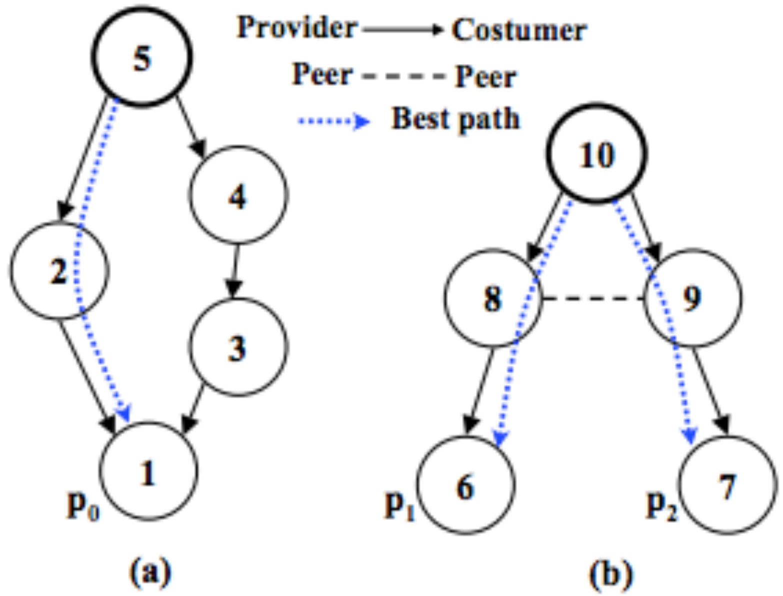
Multilateral Public Peering at IXPs using the Route Server



	DE-CIX	MSK-IX
RS-ASN	6695	8631
ALL	6695:6695	8631:8631
EXCLUDE	0:peer-asn	0:peer-asn
NONE	0:6695	0:8631
INCLUDE	6695:peer-asn	8631:peer-asn

Source: "Inferring Multilateral Peering" CoNEXT 2013

Missing AS-level Links



- Hidden: A link that has not yet been observed but could possible revealed at later time (a).

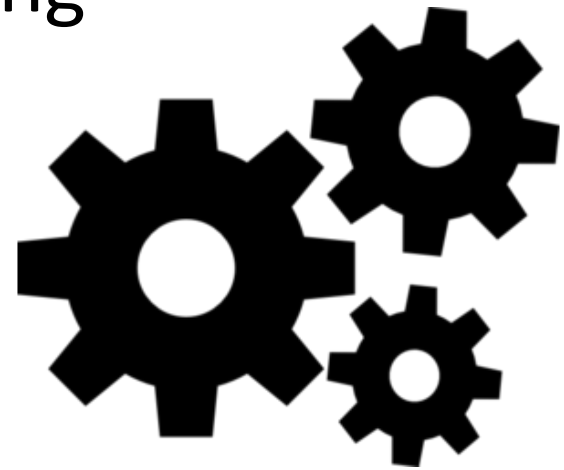
-

- Invisible: A link that it is impossible to be observed with the given set of monitors (b).

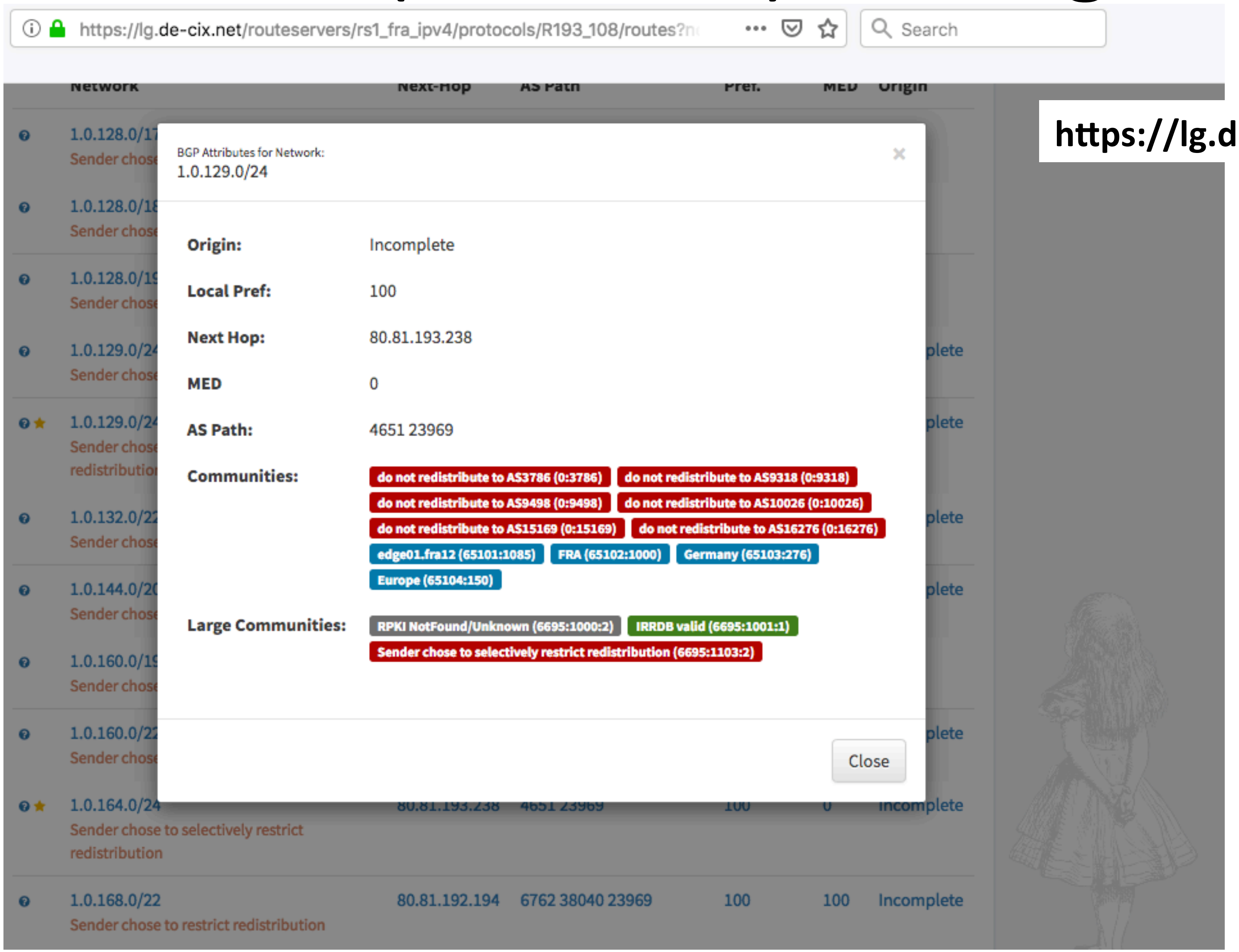
Source: "The (in)Completeness of the Observed Internet AS-level Structure" IEEE/ACM ToN 2010

Hands-on Exercise

Inferring Invisible links in IXPs
(Multilateral Public Peering
via Route Server)



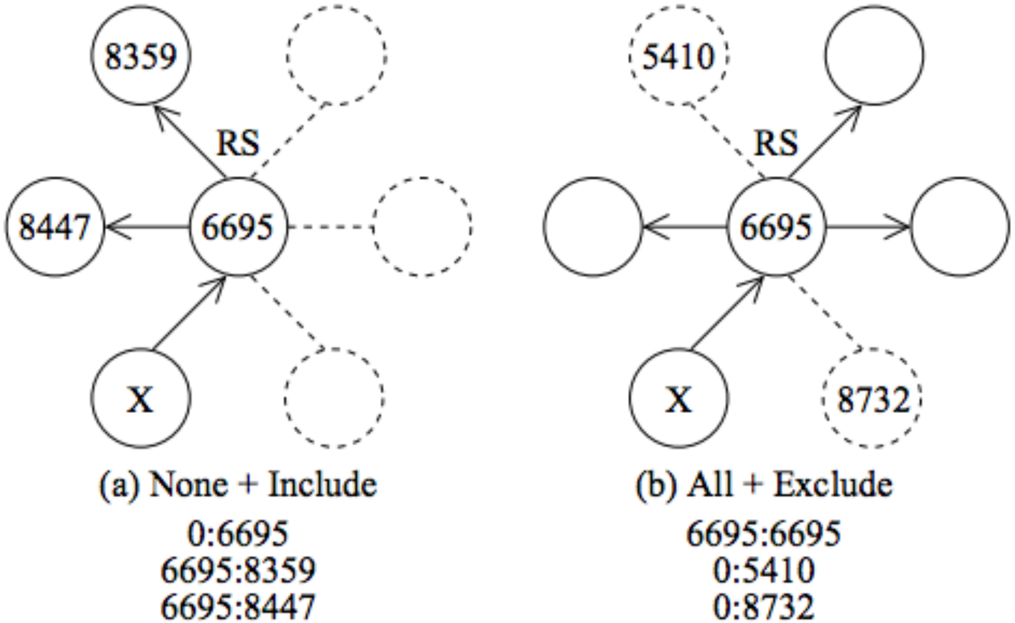
DE-CIX (Frankfurt) Looking Glass



https://lg.de-cix.net/

Multilateral Peering Links in IXPs

	DE-CIX	MSK-IX
RS-ASN	6695	8631
ALL	6695:6695	8631:8631
EXCLUDE	0:peer-asn	0:peer-asn
NONE	0:6695	0:8631
INCLUDE	6695:peer-asn	8631:peer-asn



Source: "Inferring Multilateral Peering" CoNEXT 2013

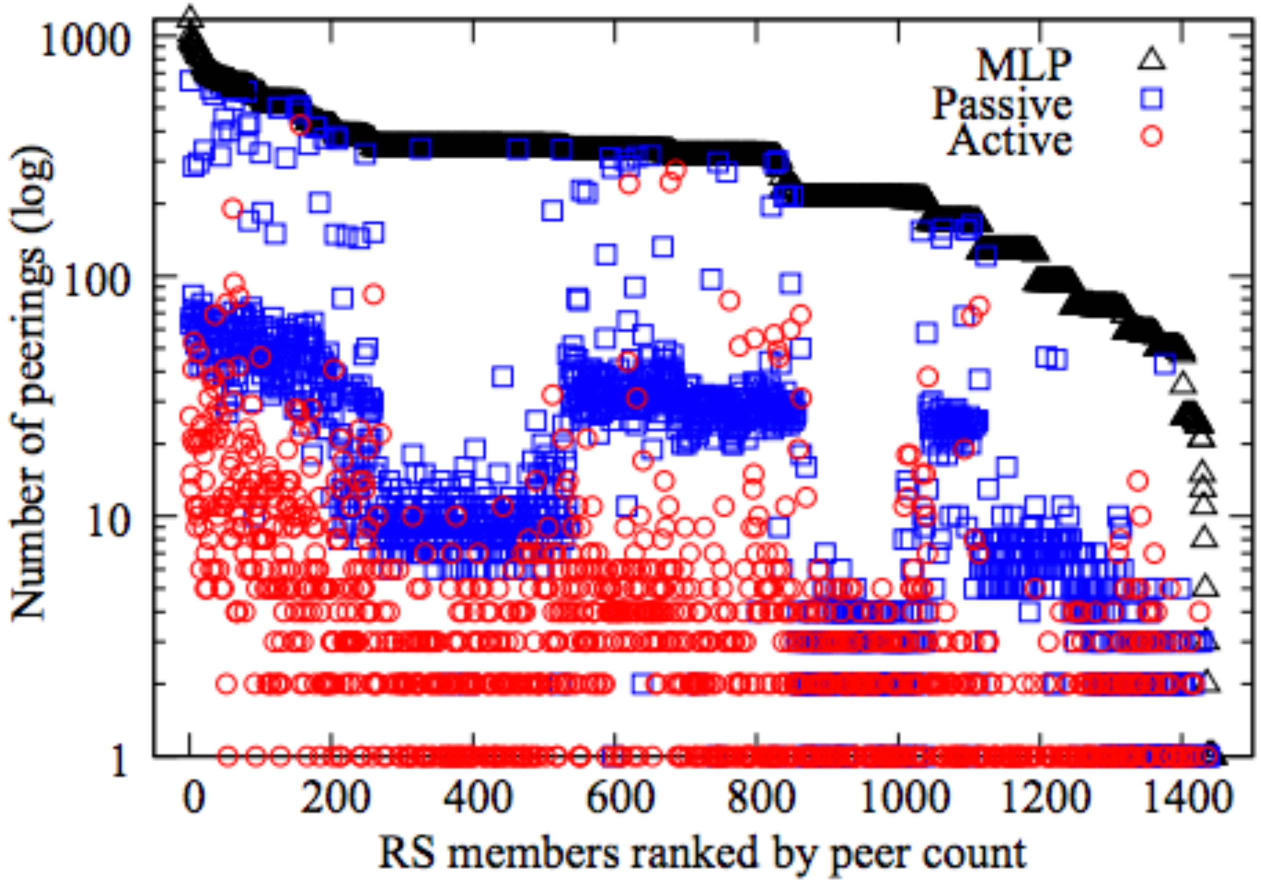
Multilateral Peering Links in IXPs

IXP	LG	ASes	RS	Pasv	Active	Links
AMS-IX	N	574	444	296	55	49249
DE-CIX	Y	483	369	113	256	54082
LINX	N	457	177*	137	39	14759
MSK-IX	Y	374	348	23	325	58501
PLIX	Y	222	211	37	174	21911
France-IX	Y	193	169	103	25	8117
LONAP	N	120	109	30	65	4458
ECIX	Y	102	83	33	50	2751
SPB-IX	Y	89	78	0	78	2828
DTEL-IX	Y	74	71	0	66	1725
TOP-IX	Y	71	52	19	33	1272
STHIX	N	69	42	4	23	340
BIX.BG	Y	53	52	0	52	950

206K p-p links from these 13 IXPs, 4x more than seen in BGP raw data

Source: “Inferring Multilateral Peering” CoNEXT 2013

Multilateral Peering Links in IXPs



Source: “Inferring Multilateral Peering” CoNEXT 2013

Readings

"Measuring ISP Topologies with Rocketfuel". Neil Spring, Ratul Mahajan, and David Wetherall. SIGCOMM 2002

"IXPs: Mapped?", Brice Augustin, Balachander Krishnamurthy , and Walter Willinger. IMC 2009