Routing in the Future Internet

Marcelo Yannuzzi

Graduate Course (Slideset 5)
Institute of Computer Science
University of the Republic (UdelaR)

August 22nd 2012, Montevideo, Uruguay

Department of Computer Architecture
Technical University of Catalonia (UPC), Spain

Institute of Computer Science
University of the Republic (UdelaR), Uruguay

Outline

1. **Interdomain aspects: Truths and Myths**
   - AS interconnection
   - Peering policies among ASs and valley-free routes
   - Topological properties
   - Evolution and invariant metrics
2. Demystify me!

The Peering Myth

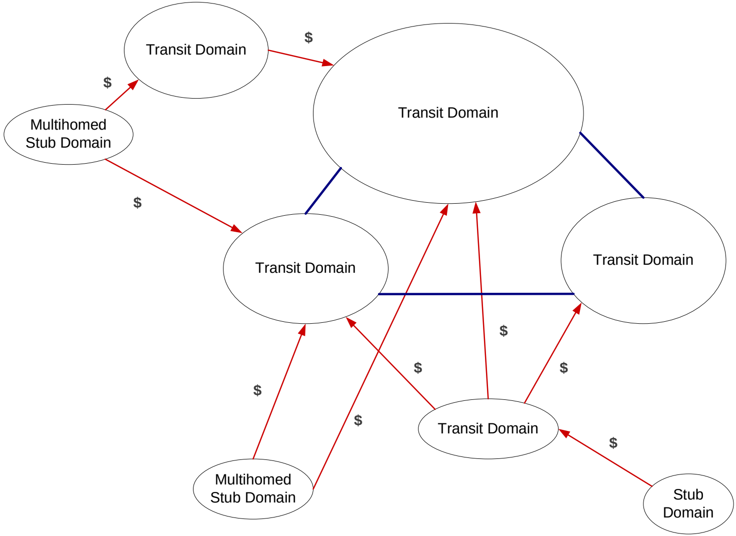
Relationships between ASs **(Myth)**

There are three types of relationships ...

- ... which correspond to the different traffic exchange agreements between neighboring domains
- ***customer-provider***: applies when a domain buys Internet connectivity from a provider.
- ***peer-peer***: applies when two providers that exchange a significant amount of traffic, agree to connect directly to each other to avoid transiting through, and thus pay, a third-party provider. Peers share the costs of the connection between them, so there is no customer-provider relationship in this case.
- ***sibling-sibling***: this relationship is quite infrequent, and are sometimes used between merging companies. According to data from CAIDA's AS Relationships Dataset, less than 0.3% of the total number of relationships between Internet domains were siblings in March of 2010.

The Tiered Structure Myth

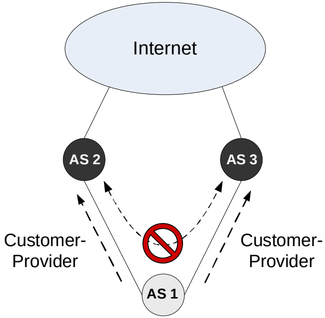
Tiered Hierarchy of Autonomous Systems **(Myth)**



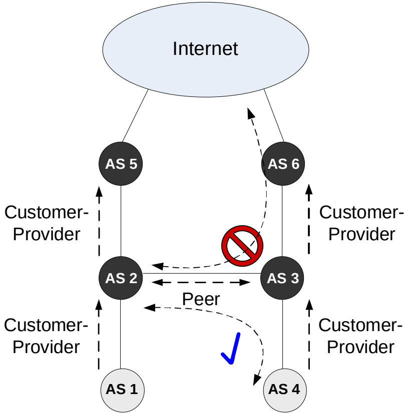
The Valley-free Routes Myth

No Valley Policies **(Myth)**

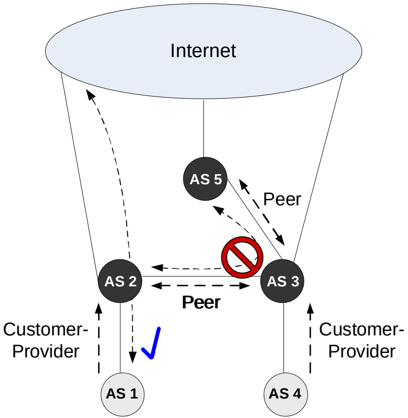
- The commercial relationships between domains impose constraints on the forwarding policies of domains.



No Valley Policies (cont.) **(Myth)**



No Valley Policies (cont.) **(Myth)**



Summary of Forwarding policies **(Myth)**

Filtering Policies (FPs)

FP1) Traffic coming from a provider will not be forwarded to a peer or another provider.

FP2) Traffic coming from a peer will not be forwarded to another peer or provider.

FP3) Traffic coming from or directed to a customer can always be forwarded by a domain.

More formally ... (Myth)

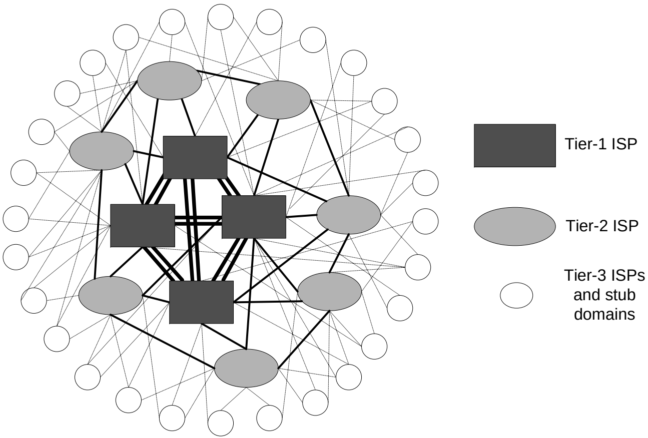
Table: Valley-free policies applied by domain AS_j for the transit from domain AS_i to domain AS_k through AS_j ($AS_i \to AS_j \to AS_k$).

Commercial relationship	AS_j is a customer of AS_k	AS_j is a provider of AS_k	AS_j is a peer of AS_k
AS_i is a provider of AS_j	$\times$	$\checkmark$	$\times$
AS_i is a customer of AS_j	$\checkmark$	$\checkmark$	$\checkmark$
AS_i is a peer of AS_j	$\times$	$\checkmark$	$\times$

The Power-laws Myth

The AS graph shape...an invariant?? **(Myth)**

- Power-laws and scale free shape ...



Internet's Topological Properties **(Myth)**

Invariants after 10 years of evolution

- A hierarchical structure is clearly observable **(Myth)**
- Power-law degree distribution **(Myth)**
- Strong clustering **(Myth)**
- Almost constant average path length **(?)**

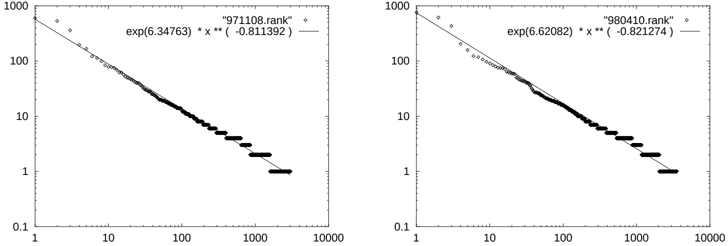
- Sources:

A. Dhamdhere, and C. Dovrolis, "Ten Years in the Evolution of the Internet Ecosystem," in Proc. of ACM Sigcomm/USENIX Internet Measurement Conference (IMC) 2008, Vouliagmeni, Greece, October 2008.

A. Elmokashfi, A. Kvalbein, and C. Dovrolis, "On the Scalability of BGP: the roles of topology growth and update rate-limiting," ACM CoNEXT 2008, Madrid, Spain, December 2008.

Faloutsos 1999: The rank power-law **(Myth)**

The rank power-law: $o_i \propto r_i^{\gamma}$



(a) Int-11-97 (b) Int-04-98

Figure 3: The rank plots. Log-log plot of the outdegree d_v versus the rank r_v in the sequence of decreasing outdegree.

- Source: M. Faloutsos, P. Faloutsos, and C. Faloutsos, “On power-law relationships of the Internet topology,” ACM/SIGCOMM, Cambridge MA, USA, August 1999.

Faloutsos 1999: The outdegree power-law **(Myth)**

The outdegree power-law: $f_o \propto o^{\beta}$

- Chou (2000) proved that the rank and outdegree power laws derived by Faloutsos et al. are equivalent

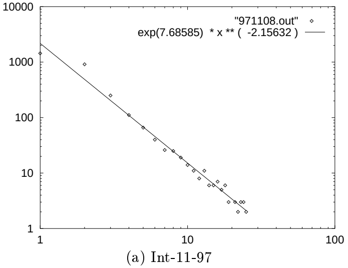


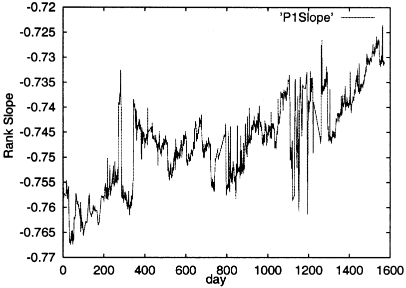
Figure 5: The outdegree plots: Log-log plot of frequency f_d versus the outdegree d.

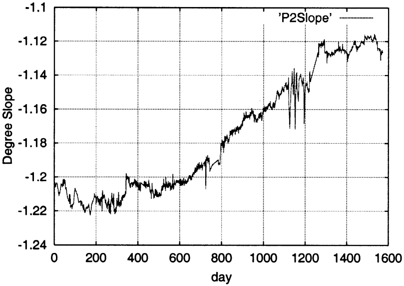
- Source: M. Faloutsos, P. Faloutsos, and C. Faloutsos, “On power-law relationships of the Internet topology,” ACM/SIGCOMM, Cambridge MA, USA, August 1999.

Siganos 03: Rank and outdegree power-laws **(Myth)**

- Five-year span from November 1997 to February 2002 (*)

  (*) Reverse cumulative distribution function of power-law 2





- Source: G. Siganos, M. Faloutsos, P. Faloutsos, and C. Faloutsos, “Power-laws and the AS-level Internet Topology,” IEEE/ACM Trans. on Networking, Vol. 11, no. 4, pp. 514-524, August 2003.

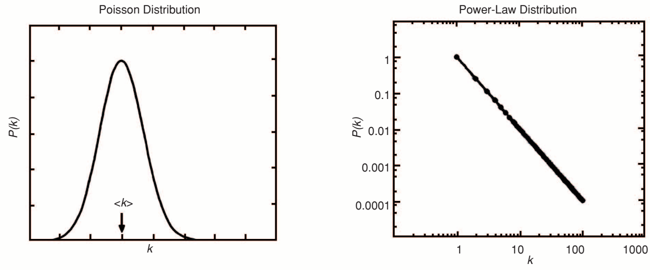
Wang 2003: The rank and outdegree power-laws

Table 1.
Small-world pattern and scale-free property of several real networks. Each network has the number of nodes N, the clustering coefficient C, the average path length L and the degree exponent γ of the power-law degree distribution. The WWW and metabolic network are described by directed graphs.

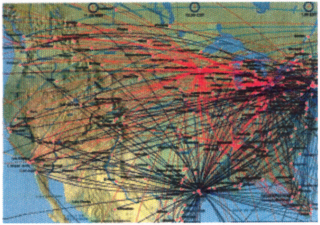
Network	Size	Clustering coefficient	Average path length	Degree exponent
Internet, domain level [13]	32711	0.24	3.56	2.1
Internet, router level [13]	228298	0.03	9.51	2.1
WWW [14]	153127	0.11	3.1	$\gamma_{\text{in}} = 2.1 \quad \gamma_{\text{out}} = 2.45$
E-mail [15]	56969	0.03	4.95	1.81
Software [16]	1376	0.06	6.39	2.5
Electronic circuits [17]	329	0.34	3.17	2.5
Language [18]	460902	0.437	2.67	2.7
Movie actors [5, 7]	225226	0.79	3.65	2.3
Math. co-authorship [19]	70975	0.59	9.50	2.5
Food web [20, 21]	154	0.15	3.40	1.13
Metabolic system [22]	778	–	3.2	$\gamma_{\text{in}} = \gamma_{\text{out}} = 2.2$

- Source: X. F. Wang, G. Chen, “Complex networks: small-world, scale-free and beyond,” Circuits and Systems Magazine, IEEE, Vol. 3, No. 1., pp. 6-20, 2003.

Wang 2003: Small-World and Scale-Free



Exponential Network



Scale-Free Network

- Source: X. F. Wang, G. Chen, “Complex networks: small-world, scale-free and beyond,” Circuits and Systems Magazine, IEEE, Vol. 3, No. 1., pp. 6-20, 2003.

Small-World and Scale-Free

Small World

- In the small-world model, the connectivity distribution of a network peaks at an average value and decays exponentially. Such networks are called “exponential networks" or “homogeneous networks", because each node has about the same number of link connections.

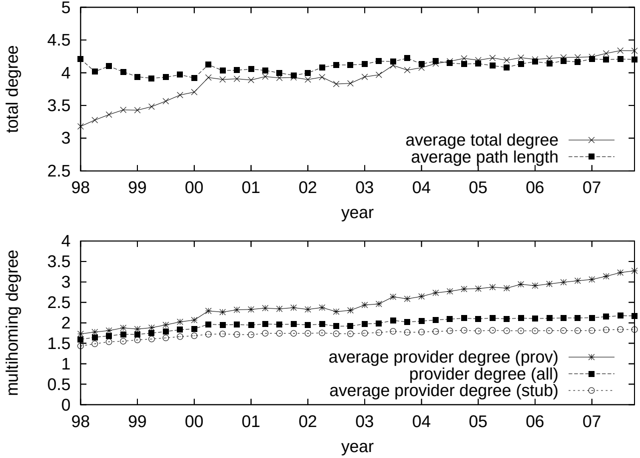
Scale-free

- Preferential attachment rule (“rich get richer").

- Source: X. F. Wang, G. Chen, “Complex networks: small-world, scale-free and beyond,” Circuits and Systems Magazine, IEEE, Vol. 3, No. 1., pp. 6-20, 2003.

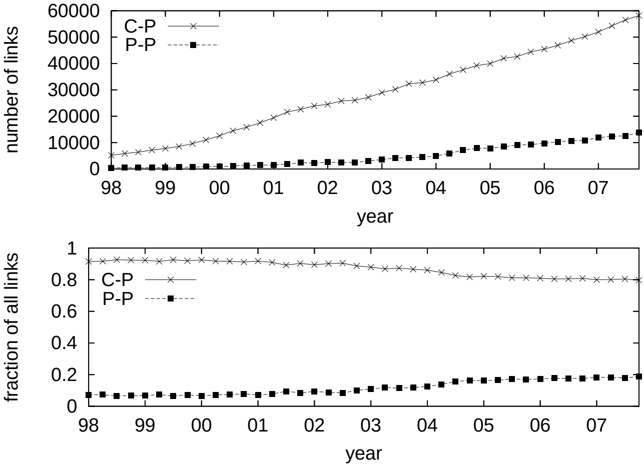
The Evolution and Invariant Metrics Myth

Evolution of the Internet **(Myth)**



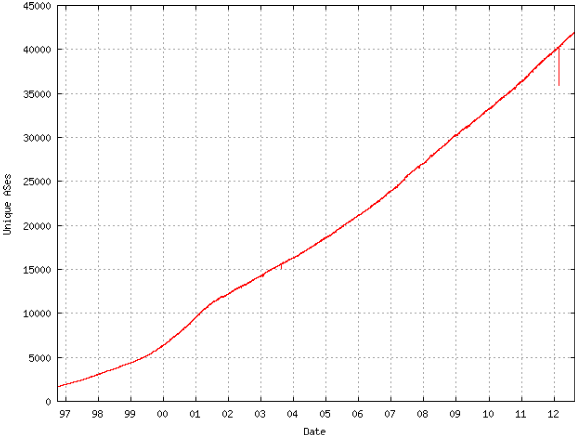
- Source: A. Dhamdhere and C. Dovrolis, “Ten Years in the Evolution of the Internet Ecosystem,” in Proc. of ACM SIGCOMM/USENIX Internet Measurement Conference (IMC), Vouliagmeni, Greece, Oct 2008.

Evolution of the Internet (cont.) **(Myth)**



- Source: A. Dhamdhere and C. Dovrolis, “Ten Years in the Evolution of the Internet Ecosystem,” in Proc. of ACM SIGCOMM/USENIX Internet Measurement Conference (IMC), Vouliagmeni, Greece, Oct 2008.

Evolution of the number of ASs (as of August 2012)



- Source: CIDR Report

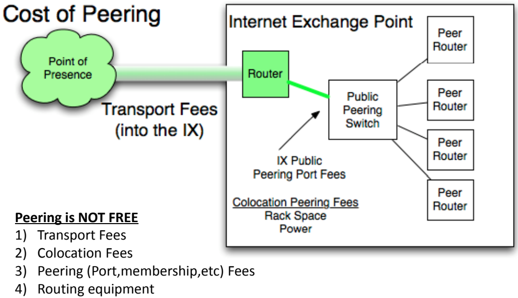
Outline

1. Interdomain aspects: Truths and Myths
   - AS interconnection
   - Peering policies among ASs and valley-free routes
   - Topological properties
   - Evolution and invariant metrics
2. **Demystify me!**

What do IXPs reveal?

Peering: Internet Exchange Points (IXPs)

- Support easy interconnection between member ASes: physical space, caches, cabling, power, A/C supply, secure access, etc.)

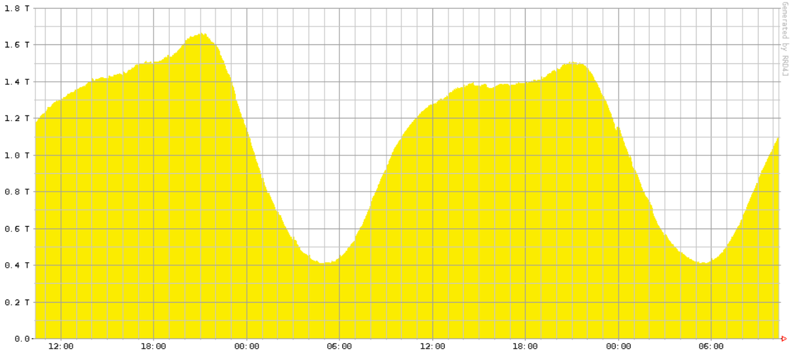


- Source: William B. Norton, “A Business Case for Peering in 2010,” in the 15 YEAR ANNIVERSARY of the 1st DE-CIX CUSTOMER SUMMIT Frankfurt, Germany, August 2010.

DE-CIX: the largest IXP in the world

- Centrally located in Frankfurt right in the heart of Europe.
- Largest Internet Exchange in the world
- Leading Internet Exchange for Central and Eastern Europe
- State of the art switching platform
- 465+ participants (56+ new in 2011)
- 7 Tbps of connected capacity (Public Peering)
- 500+ Private Interconnects in service
- 200+ Gigabit-Ethernet ports
- 600+ 10Gigabit-Ethernet ports
- 100% uptime since 2007

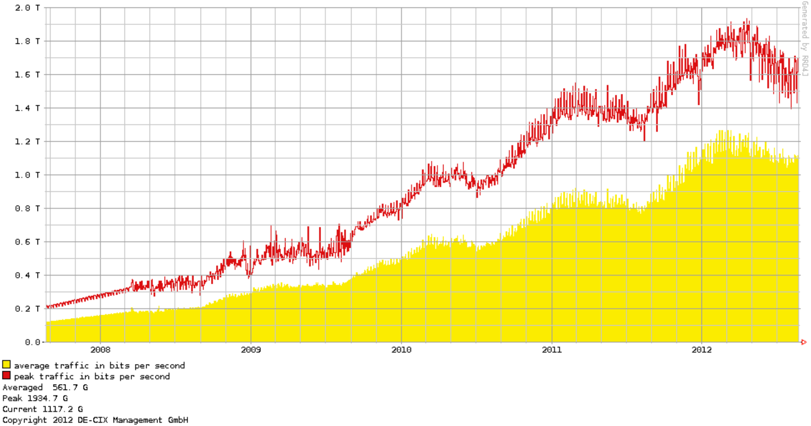
- IPv4 traffic:



average traffic in bits per second
Averaged 1088.4 G
Peak 1670.0 G
Current 1090.3 G

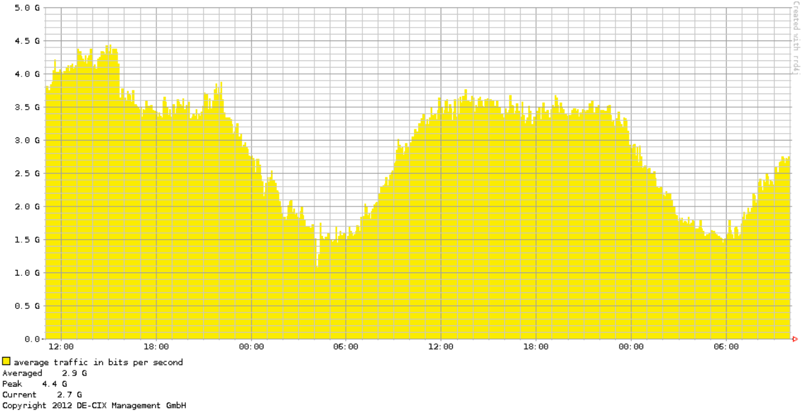
DE-CIX: the largest IXP in the world (cont.)

- IPv4 traffic:



DE-CIX: the largest IXP in the world (cont.)

- IPv6 traffic:



Traffic Volumes on a daily basis

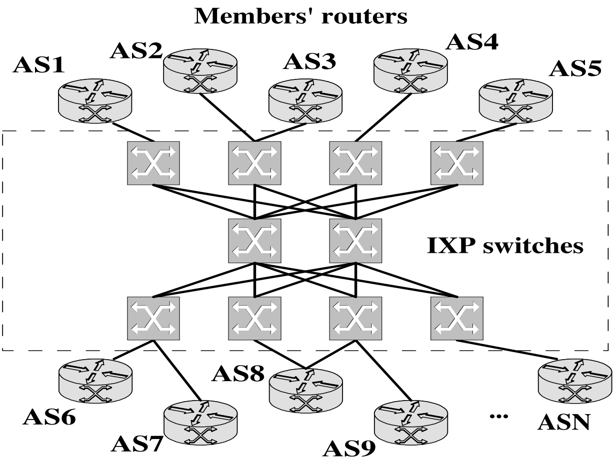
Comparison

- AT&T around 30 petabytes of traffic per day on average.
- Deutsche Telekom around 15 petabytes of traffic per day on average.
- An IXP....similar traffic volumes...due to the traffic exchanged by a few hundreds of member ASs (300 – 500 ASs) among each other.
- There are > 300 IXPs worldwide.

1 peta = $2^{50} \approx 10^{15}$

- Source: B. Ager et al. "Anatomy of a Large European IXP," ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012.

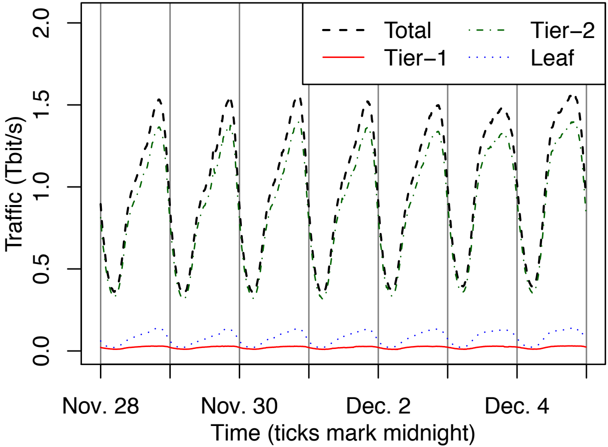
The internals of an IXP



(a) A typical IXP architecture.

- Source: B. Ager et al. “Anatomy of a Large European IXP,” ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012.

The internals of an IXP (cont.)



(b) Traffic volume distribution across time.

- Source: B. Ager et al. “Anatomy of a Large European IXP,” ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012.

The internals of an IXP (cont.)

Table 1: Overview of IXPs sFlow dataset.

	Apr 25 May 1	Aug 22 Aug 28	Oct 10 Oct 16	Nov 28 Dec 4
Identified member ASes	358	375	383	396
Router IPs	426	445	455	474
MAC addresses	428	448	458	474
Tier-1	13	13	13	13
Tier-2	281	292	297	306
Leaf	64	70	73	77
Countries of member ASes	43	44	45	47
Continents of member ASes	3	3	3	3
Average packet rate (Mpps)	142	150	166	174
Average bandwidth (Gbps)	838	863	954	992
Daily avg volume (PB)	9.0	9.3	10.3	10.7

- Source: B. Ager et al. “Anatomy of a Large European IXP,” ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012.

Resetting our mental picture of the Internet...

Study performed by Ager et al. **(public peerings)**:

9 months' worth of sFlow records collected at an IXP in 2011.

- Main findings:
  - *"... this IXP has close to 400 members which have established some 67% (or* ***more than 50,000****) of all possible such peerings and use them for exchanging some 10 PB of IP traffic daily."*

    "To put this number in perspective, note that as of 2010, the number of inferred AS links of the peer-peer type in the Internet was reported to be around 40,000 – ***less than what we observe at this particular IXP alone!****"*

- Source: B. Ager et al. "Anatomy of a Large European IXP," ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012.

Resetting our mental picture of the Internet...(cont.)

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

- Source: B. Ager et al. "Anatomy of a Large European IXP," ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012.

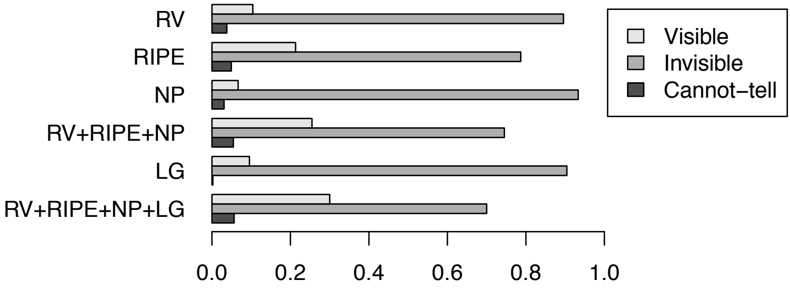


Figure 2: Peering links and visibility in control/data plane (normalized by number of detected P-P links).

- Source: B. Ager et al. “Anatomy of a Large European IXP,” ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012.

Resetting our mental picture of the Internet...(cont.)

Study performed by Ager et al. **(public peerings)**:

- *"even when relying on all the available datasets,* ***about 70% of the P-P links at this IXP remain invisible****"*
- The Internet is "flattening" .... through a myriad of shortcuts.
- The tiered picture is still there though the hierarchy as such seems to be an "illusion" ... since "*...the observed rich peering fabric at this IXP enables connectivity among networks of all different types and is essentially agnostic of any tier structure.*"

- Source: B. Ager et al. "Anatomy of a Large European IXP," ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012.

Resetting our mental picture of the Internet...(cont.)

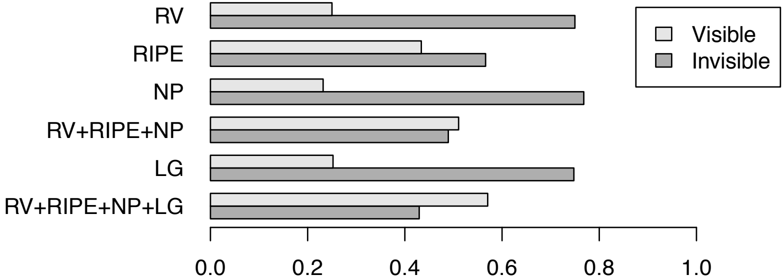
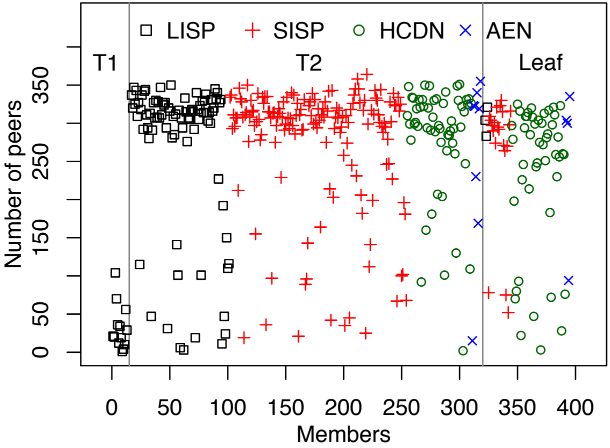


Figure 3: Peering traffic and visibility in control/data plane (normalized by total traffic volume).

- Source: B. Ager et al. "Anatomy of a Large European IXP," ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012.

Resetting our mental picture of the Internet...(cont.)



(b) Scatter-plot of num. of peers per member.

- Source: B. Ager et al. “Anatomy of a Large European IXP,” ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012.

The “Why" of these discrepancies?

Why is the peering fabric seen from inside the IXP so different from the one seen from outside the IXP?

- ...
- ...

- Source: B. Ager et al. “Anatomy of a Large European IXP,” ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012.

The “Why" of these discrepancies?

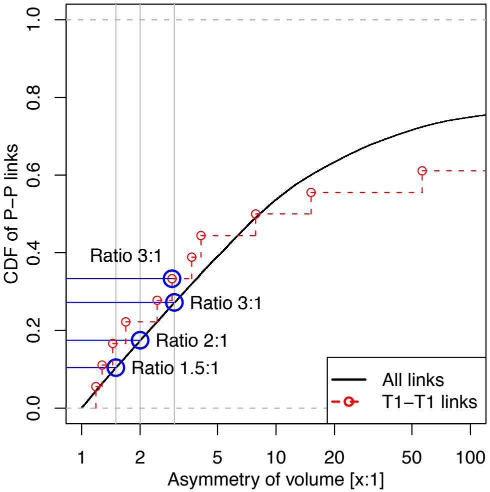
The Wrong Tools ...

- BGP-based solutions...tough BGP is an information-hiding protocol
  - The vantage points problem
  - Routing policies ... “have an intrisically hiding nature"
- Traceroute-based solutions....though traceroute was not devised to the end of topology discovery

- Source: B. Ager et al. “Anatomy of a Large European IXP,” ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012.

Other Myths...

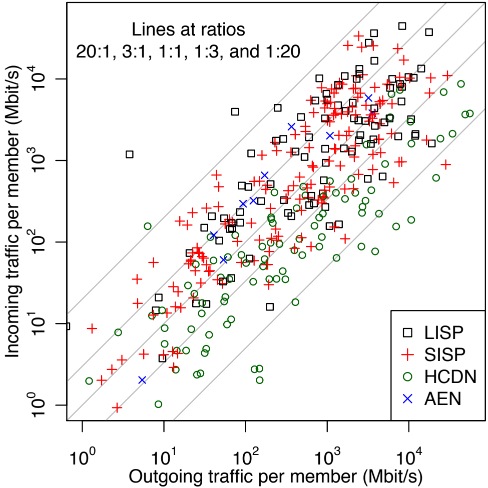
Resetting our mental picture of the Internet...(cont.)



(a) Traffic asymmetry across P-P links.

- Source: B. Ager et al. "Anatomy of a Large European IXP," ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012.

Resetting our mental picture of the Internet...(cont.)



(b) Traffic asymmetry (out/in) per member.

- Source: B. Ager et al. "Anatomy of a Large European IXP," ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012.

Some meat for thought ...

"...considering only the European IXP scene (see [18] for details) and being conservative in using our large IXP as a baseline (i.e., **assuming only a 50% peering rate at IXPs**), counting up the P-P links we expect to encounter at the four largest IXPs (with, say, 400 unique members each) and at the 10 next-largest IXPs (with, say, 100 unique members each), we obtain a realistic lower bound for the estimated number of P-P links for just the European portion of the Internet of some 200,000. **This number is more than 100% larger than the number of all AS links (i.e., customer-provider and peer-peer) in the entire Internet in 2010 as reported in [16]**...."

- Source: B. Ager et al. "Anatomy of a Large European IXP," ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012.

It seems that ...

1. There are easily an order of magnitude more P-P links in today's Internet than previously assumed.
2. Contrary to what we thought, there are many more P-P links in today's Internet than customer-provider type peerings, with twice as many being a conservative estimate.
3. Most of these P-P links are of critical importance as they carry significant traffic.

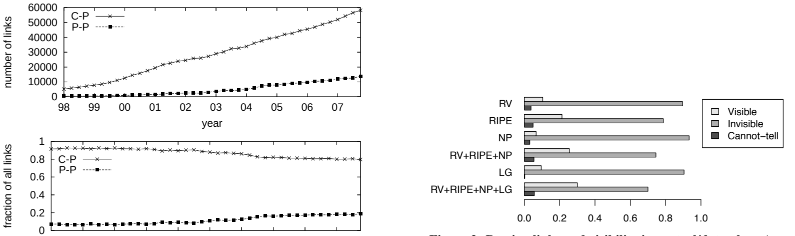


Figure 2: Peering links and visibility in control/data plane (normalized by number of detected P-P links).

- Sources: (bottom-left) B. Ager et al. "Anatomy of a Large European IXP," ACM SIGCOMM 2012, Helsinki, Finland, August 13–17, 2012 and (top and bottom-right) A. Dhamdhere and C. Dovrolis, "Ten Years in the Evolution of the Internet Ecosystem," in Proc. of ACM SIGCOMM/USENIX Internet Measurement Conference (IMC), Vouliagmeni, Greece, Oct 2008.

So now what?

- What about the tiered structure then?
- What about the 3 classical policies (C-P, P-P, S-S)?
  - M. Roughan et al. “10 Lessons from 10 Years of Measuring and Modeling the Internet’s Autonomous Systems,” IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, VOL. 29, NO. 9, OCTOBER 2011.
  - W. B. Norton, “A Study of 28 Peering Policies," http://drpeering.net/white-papers/Peering-Policies/A-Study-of-28-Peering-Policies.html.
- What about valley-free policies?
  - see, e.g., “partial transit policies" in the first paper listed above.
  - S. Y. Qiu, P. D. Mcdaniel, and F. Monrose, “Toward valley-free interdomain routing," in IEEE ICC, 2007.
- ...so....what about the power-laws and the topological properties found by the Faloutsos in 1999?
- ....and what about what we think we know about the Internet’s evolution and its invariant metrics?

Readings for Monday, August 27th

1. P. Brighten Godfrey, Igor Ganichev, Scott Shenker, and Ion Stoica, “Pathlet Routing," ACM SIGCOMM 2009.
2. Stefano Vissicchio, Luca Cittadini, Laurent Vanbever, and Olivier Bonaventure, “iBGP Deceptions: More Sessions, Fewer Routes," IEEE INFOCOM 2012.

Template available from Moodle.

Questions?