Internet Topology

Andrea Weisskopf
Seminar of Distributed Computing
WS 03/04

Outline

- BGP Introduction
- Traceroute and BGP AS Path Incongruities
- Internet Topology: Connectivity of IP Graphs
- Conclusion
- Questions and Discussion

Papers

- Traceroute and BGP AS Path Incongruities.
  Y. Hyun, A. Broido, K. Claffy; 2003
  CAIDA (Cooperative Association for Internet Data Analysis)
- Internet topology: Connectivity of IP Graphs.
  A. Broido, K. Claffy; ITCom 2001
  CAIDA (Cooperative Association for Internet Data Analysis)

BGP (1)

- Internet consists of Autonomous Systems (AS) interconnected with each other.
- AS are numbered like
  - AS 559: SWITCH
  - AS 8803: Migros-Genossenschafts-Bund
- Two kinds of routing
  - Intra-AS (RIP, OSPF, IGRP etc.)
  - Inter-AS (BGP as de-facto standard)

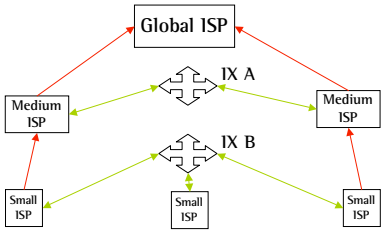
  ... as taught in network classes

BGP (2)

- Routing often done at Internet eXchange points (IXes) like TIX, LINX, DECIX
- 1:n switching
- Most of them neutral (not owned by ISP, special IX organisation)

BGP (3)

- Routing policies
  not always shortest path, contracts between ISPs for peering or transit

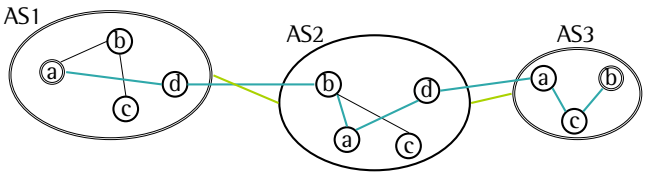


IX = internet exchange

Peering, cost natural
Transit, with costs

1st Paper: Traceroute and BGP AS Path Incongruities

- Goal: Try to explain the differences between IP traceroute path and AS paths



Overview 1st paper

- Collect IP traceroute & BGP AS paths
- Convert IP traceroute to AS traceroute paths
- Pair AS traceroute and BGP AS paths
- Try to explain 2 types of incongruities

Collecting IP traceroute paths

- Locations

		type	# Hosts
`sjc`	San Jose	IP based	301752
`k-peer`	Amsterdam	DNS based	143193
`m-root`	Tokyo		

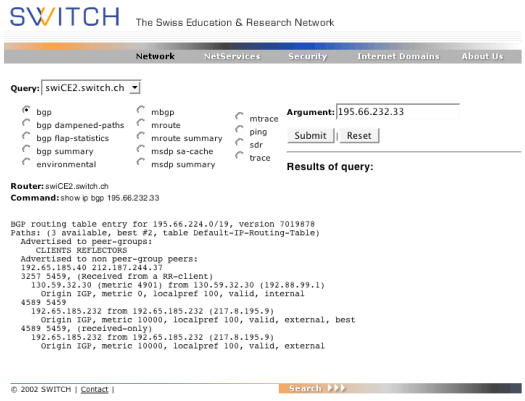
IP based & DNS based have 23903 hosts in common

- Using a modified version of `traceroute` called skitter, to avoid probing the same host more than once
- Probing done between 01:00 and 13:00 on April 1, 2002

Collecting BGP AS paths

- No need to query hosts, needed AS paths are stored in the routing table of a BGP router
- RouteView snapshot taken of closest backbone router to `sjc`, `k-peer` & `m-root`
- Snapshot taken at 06:00 on April 1, 2002 was in the middle of the period used for the skitter probing

Looking glass at lg.lan.switch.ch



Converting IP paths to AS paths (1)

Example LINX -> TIX

- IP path

```
0 collector.linx.net (195.66.232.254)
1 195.66.224.110 (195.66.224.110)
2 i68ges-015-pos5-2.bb.ip-plus.net (164.128.33.13)
3 i79zhh-015-pos5-0.bb.ip-plus.net (164.128.33.1)
4 i79tix-005-gig1-0.bb.ip-plus.net (164.128.34.82)
5 cctld.tix.ch (194.42.48.120)
```

- BGP path

```
0 5459
1 3303
2 8235
```

Converting IP paths to AS paths (2)

Example LINX -> TIX

- Involved ASes

```
AS 3303 SWISSCOM (IP-plus)    164.128.0.0/16, …
AS 5459 LINX-AS               195.66.224.0/19, …
AS 8235 TIX-ZH                194.42.48.0/24, …
```

- IP path and corresponding ASes by longest prefix matching

```
0 collector.linx.net (195.66.232.254) [AS 5459]
1 195.66.224.110 (195.66.224.110) [AS 5459]
2 i68ges-015-pos5-2.bb.ip-plus.net (164.128.33.13) [AS 3303]
3 i79zhh-015-pos5-0.bb.ip-plus.net (164.128.33.1) [AS 3303]
4 i79tix-005-gig1-0.bb.ip-plus.net (164.128.34.82) [AS 3303]
5 cctld.tix.ch (194.42.48.120) [AS 8235]
```


Converting IP paths to AS paths (3)

Example LINX -> TIX

Traceroute AS path	BGP AS path
0 … (195.66.232.254) [AS 5459]	
1 … (195.66.224.110) [AS 5459]	
2 … (164.128.33.13) [AS 3303]	0 5459
3 … (164.128.33.1) [AS 3303]	1 3303
4 … (164.128.34.82) [AS 3303]	2 8235
5 … (194.42.48.120) [AS 8235]	

… which are apparently the same

Paring traceroute AS paths and BGP AS paths

- Lookup BGP path in RouteView data for every traceroute IP path by longest prefix matching
- Surjective, not injective mapping between IP traceroute paths and AS paths
- Thus need for eliminating redundant pairs

Eliminating redundant pairs: Example

	Pair A	Pair B
source	collector.linx.net (195.66.225.254)	collector.linx.net (195.66.225.254)
destination	dcg.inf.ethz.ch (129.132.130.158)	www.ethz.ch (129.132.202.79)
IP traceroute	00 collector.linx.net (195.66.225.254) [AS 5459] 01 London-i2.telia.net (195.66.224.48) [AS 5459] 02 ldn-bb1-pos5-2-0.telia.net (213.248.65.97) [AS 1299] 03 pra-bb1-pos1-1-0.telia.net (213.248.64.158) [AS 1299] 04 zch-b1-pos6-1.telia.net (213.248.65.42) [AS 1299] 05 dante-01287-zch-b1.c.telia.net (213.248.79.190) [AS 1299] 06 swiEZ2-G3-2.switch.ch (130.59.36.249) [AS 559] 07 rou-rz-gw-giga-to-switch.ethz.ch (192.33.92.1) [AS 559] 08 rou-ethz-access-intern.ethz.ch (192.33.92.130) [AS 559] 09 rou-ifw-mega-transit-1.ethz.ch (129.132.99.72) [AS 559] 10 dcg.inf.ethz.ch (129.132.130.158) [AS 559]	00 collector.linx.net (195.66.225.254) [AS 5459] 01 London-i2.telia.net (195.66.224.48) [AS 5459] 02 ldn-bb1-pos5-2-0.telia.net (213.248.65.97) [AS 1299] 03 pra-bb1-pos6-0-0.telia.net (213.248.64.9) [AS 1299] 04 zch-b1-pos6-1.telia.net (213.248.65.42) [AS 1299] 05 dante-01287-zch-b1.c.telia.net (213.248.79.190) [AS 1299] 06 swiEZ2-G3-2.switch.ch (130.59.36.249) [AS 559] 07 rou-rz-gw-giga-to-switch.ethz.ch (192.33.92.1) [AS 559] 08 rou-ethz-access-intern.ethz.ch (192.33.92.130) [AS 559] 09 rou-rz-1-mega-transit-2.ethz.ch (129.132.99.195) [AS 559] 10 www.ethz.ethz.ch (129.132.202.79) [AS 559]
IP AS path	00 5459 01 1259 02 559	00 5459 01 1259 02 559
BGP AS path	00 5459 01 1259 02 559	00 5459 01 1259 02 559

AS 559 ETHZ 129.132.0.0/16, …

Summary of simplifications

	sjc		k-peer		m-root	
probed hosts	301752		143193		143193	
completed traceroutes	220088	73%	89667	63%	89317	62%
non-redundant IP AS path – BGP AS path pairs	60271	20%	36950	26%	38527	27%
incongruent paths	11279	4%	36888	26%	38460	27%
covered BGP prefixes (of total announced 113563)	58037	51%	36170	31%	37292	32%

Incongruent paths

- Should not occur in theory
- But at k-peer and m-root almost every pair is incongruent
- Which one is the “real” path of an IP packet?

Incongruities from IX ASes (1)

- List of IXes compiled by
  - querying internet registries whois databases (RIPE, APNIC, LACNIC, ARIN) and looking for strings like “internet exchange”, “IX”, etc.
  - consulting some unofficial lists floating around the internet
- Leads to 60 IX ASes

Incongruities from IX ASes (2)

- Just a few IX ASes responsible for majority of occurrences
- k-peer is located near AMS-IX (`AS 1200`), m-root is located near WIDE/NSPIXP (`AS 2500`)

sjc			k-peer			m-root		
IX AS	freq	cum. %	IX AS	freq	cum. %	IX AS	freq	cum. %
6695	2174	48.1%	1200	36908	98.0%	2500	31679	90.9%
5459	1187	74.3%	10764	305	98.8%	7527	1423	95.0%
7527	546	86.4%	6695	252	99.5%	6695	949	97.7%
total	4331		total	36150		total	33782	

Incongruities from ASes under same ownership (1)

- In theory no organization needs more than one AS
- In practice many organization have more than one, due:
  - Simpler routing policies
  - Segregating traffic classes (academic vs. commercial)
  - Business merges and acquisitions

Incongruities from ASes under same ownership (2)

sjc			k-peer			m-root		
group	freq	cum. %	group	freq	cum. %	group	freq	cum. %
MCI[a]	665	16%	Level3	3518	53%	MCI	1719	33%
SBC[b]	571	29%	C&W[d]	465	60%	Telia	444	41%
Qwest[c]	557	42%	Telia	329	65%	Qwest	377	48%
Telia	494	54%	Qwest	322	70%	SBC	369	55%
AT&T	323	62%	XO	300	75%	Sprint	355	62%
total	4232		total	6623		total	5284	

[a]MCI/WorldCom/UUNET/AlterNet/ANS/Bertelsmann

[b]SBC/Pacific Bell/Nevada Bell/Southwestern Bell

[c]Qwest/US West/SuperNet/Touch America

[d]C&W/Exodus/PSI

Remaining incongruities (1)

sjc			k-peer			m-root		
t - b	freq		t - b	freq		t - b	freq	
2	1597	33%	2	1861	35%	2	7795	46%
-1	1203	25%	0	1485	28%	1	6217	37%
1	1151	24%	3	941	18%	3	1220	7%
0	474	10%	1	679	13%	0	1126	7%
*	394	8%	*	295	6%	*	569	3%
+	3125	65%	+	3673	70%	+	15765	93%
-	1220	25%	-	103	2%	-	36	0%
total	4819		total	5261		total	16927	

* Remaining length differences other than those included in the table

t Traceroute AS path length

b BGP AS path length

Remaining incongruities (2)

- Edit distance to transform a BGP AS path to a traceroute AS path

Operation	sjc		k-peer		m-root	
insertions only	2788	58%	2764	53%	13661	81%
deletions only	1132	23%	1	0%	0	0%
substitutions only	813	17%	1813	34%	2648	15%
mixture	86	2%	683	13%	618	4%
total paths	4819		5261		16927	

Remaining incongruities (3)

- Causes for incongruities
  - ISPs participating at IXes erroneously announce IX prefixes
  - Customer ASes at tail of traceroute IP path hidden by prefix aggregation in BGP path
  - IP stacks not conform to RFC1812 in setting source ICMP reply addresses
  - Asymmetric routing with multihomed nodes
  - Misconfiguration of BGP routers (common excuse for unexplainable things)

Summary of incongruities

	sjc		k-peer		m-root	
non-redundant IP AS path – BGP AS path pairs	60271	534%	36950	100%	38257	100%
incongruent paths	11279	100%	36888	100%	38460	100%
IX ASes	3749	33%	30163	82%	20601	54%
ASes same ownership	2711	24%	1464	4%	932	2%
Remaining	4819	43%	5261	14%	16927	44%

- Still a lot of unexplainable incongruities
- Probing host location really matters
- Real world routing policies (and business relationships) not in BGP data

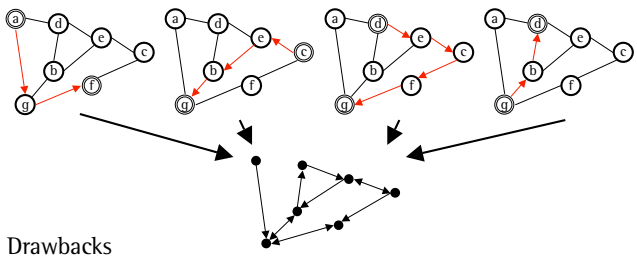
2nd paper: Internet topology: Connectivity of IP Graphs

- Goal: measure a lot of IP graph properties, confuse the reader and hardly provide an explanation...
- In other words: what would you do with an IP graph of 655k nodes?

Research ideas for 655k internet node graph

- Scalability of the internets core
- Simulate new routing algorithms with a realistic connectivity model
- Discover business relationships, decision help for peering strategies
- Be an artist and draw nice pictures of it

Graph construction from IP traceroute paths



Drawbacks

- only ICMP forward data paths
- probing from only 17 skitter monitors 220M nodes

Introduced concepts (1)

- Cones: all nodes reachable from node A via the acyclic sub graph
- (Stub) trees: sub graph connected to the rest of the graph only through its root A
- Stripping: reducing graph G to its core
- Placeholder graph: replace non-responding nodes in IP graph with arcs or placeholders

Introduced concepts (2)

- Use ccdf (complementary cumulative distribution function) rather than frequencies of object sizes

$$F(x) = P(X > x)$$

- Different measures of internet's objects follow Weibull distribution

$$N\{X > x\} = a \exp(-(x/b)^c)$$

Stripping (1)

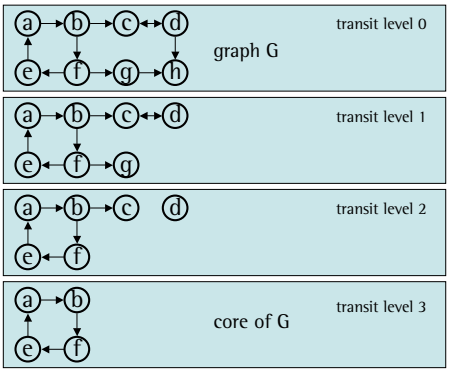
Transit sub graph of G: tsg(G)
remove all nodes with out degree 0
remove all edges of terminal 2 loops

Transit level n sub graph of G: tlsg(G, n)
tlsg(G, 0) = G
tlsg(G, n) = tsg(tlsg(G, n - 1))

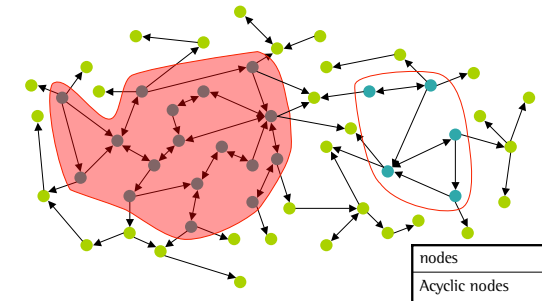
Stripping (2)

- Core of a graph: lowest stable transit level n sub graph
- Nodes not in the core belong to the acyclic sub graph
- Giant Component: largest connected component of core

Stripping: Example



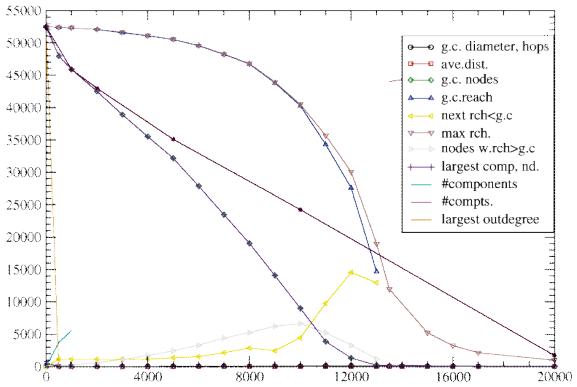
Internet IP graph



Acyclic sub graph
Core
Connected components
Giant component

nodes	629647	100%
Acyclic nodes		90.5%
Core nodes	60008	9.6%
Giant component	52505	8.3%

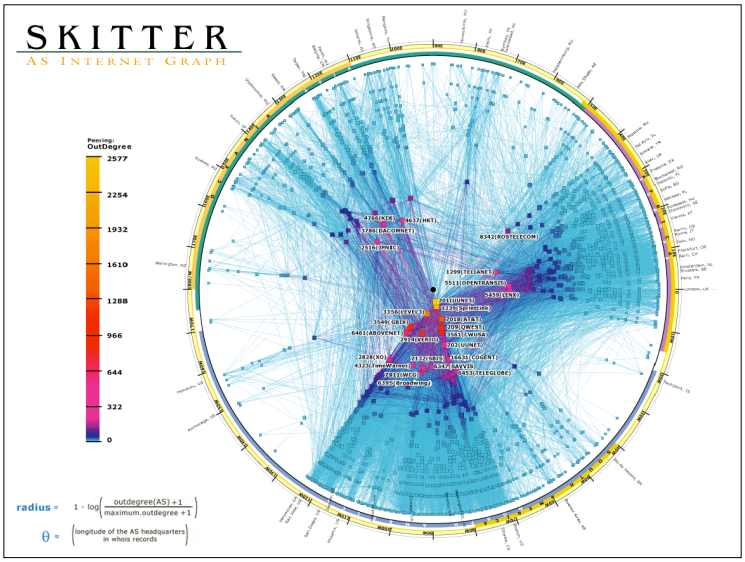
Topological resilience of giant component



Nodes removed from giant component by outdegree

Conclusion 2nd paper

- No real results, nor a motivation for measuring all this properties
- One paper belong many others covering this topic
- Weibull distribution for modelling internet characteristics, brute-force attack for best fit
- Most of the results are not comprehensible
  - unintroduced or undefined terms, not even defined otherwise
  - plots with missing y-axis label or overlapping, unlabeled curves
  - reference to unavailable papers
    - "The extended version of this paper includes..." but there is no extended version



Questions?

Discussion

- What should be first, collected data or research idea?
- What is better suited for topology analysis, BGP AS paths or IP traceroute paths?

Converting IP paths to AS paths (4)

Example TIX -> LINX

- IP path and corresponding ASes

```
0  cctld.tix.ch (194.42.48.120) [AS 8235]
1  194.42.48.125 (194.42.48.125) [AS 8235]
2  tix-1.ip-plus.net (194.42.48.12) [AS 8235]
3  i79zhh-015-gig8-2.bb.ip-plus.net (164.128.34.81) [AS 3303]
4  i68ges-015-pos5-0.bb.ip-plus.net (164.128.33.2) [AS 3303]
5  i00lon-005-pos2-0.bb.ip-plus.net (164.128.33.14) [AS 3303]
6  collector.linx.net (195.66.225.254) [AS 5459]
```

- BGP path

```
0 8235
1 4589
2 5459
```

- Involved ASes

```
AS 3303 SWISSCOM (IP-plus)  164.128.0.0/16, …
AS 4589 EASYNET
AS 5459 LINX-AS             195.66.224.0/19, …
AS 8235 TIX-ZH              194.42.48.0/24, …
```