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

[Diagram of an Internet topology with IX exchanges labeled as IX A and IX B, showing Global ISP and Medium ISP connections with arrows indicating peering and transit, costs are denoted with different line styles.]
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

<table>
<thead>
<tr>
<th>Location</th>
<th>Type</th>
<th># Hosts</th>
</tr>
</thead>
<tbody>
<tr>
<td>sjc</td>
<td>San Jose</td>
<td>301752</td>
</tr>
<tr>
<td>k-peer</td>
<td>Amsterdam</td>
<td>143193</td>
</tr>
<tr>
<td>m-root</td>
<td>Tokyo</td>
<td></td>
</tr>
</tbody>
</table>

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 160gw-015-poa5-2.bb.ip-plus.net (164.128.33.13)
  3 l79zhb-015-poa5-0.bb.ip-plus.net (164.128.33.1)
  4 l79tix-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
  1. AS 3303 SWISSCOM (IP-plus)
  2. AS 5459 LINX-AS
  3. AS 8235 TIX-ZH

- IP path and corresponding ASes by longest prefix matching
  0 collector.linx.net (195.66.225.254) [AS 5459]
  1 195.66.224.110 (195.66.224.110) [AS 5459]
  2 68pea-015-005-0-bb.ip-plus.net (164.128.33.13) [AS 3303]
  3 195.66.224.110 (195.66.224.110) [AS 5459]
  4 195.66.224.110 (195.66.224.110) [AS 5459]
  5 cctld.tix.ch (194.42.48.120) [AS 8235]

Converting IP paths to AS paths (3)

Example LINX -> TIX

<table>
<thead>
<tr>
<th>Traceroute AS path</th>
<th>BGP AS path</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 ... (195.66.224.254) [AS 5459]</td>
<td>... (195.66.224.254) [AS 5459]</td>
</tr>
<tr>
<td>1 ... (195.66.224.110) [AS 5459]</td>
<td>... (195.66.224.110) [AS 5459]</td>
</tr>
<tr>
<td>2 ... (164.128.33.13) [AS 3303]</td>
<td>... (164.128.33.13) [AS 3303]</td>
</tr>
<tr>
<td>3 ... (164.128.33.1) [AS 3303]</td>
<td>... (164.128.33.1) [AS 3303]</td>
</tr>
<tr>
<td>4 ... (164.128.34.82) [AS 8235]</td>
<td>... (164.128.34.82) [AS 8235]</td>
</tr>
<tr>
<td>5 ... (194.42.48.120) [AS 8235]</td>
<td>... (194.42.48.120) [AS 8235]</td>
</tr>
</tbody>
</table>

... 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

<table>
<thead>
<tr>
<th>Pair A</th>
<th>Pair B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 collector.linx.net (195.66.225.254) [AS 5459]</td>
<td>0 collector.linx.net (195.66.225.254) [AS 5459]</td>
</tr>
<tr>
<td>1 London-i2.telia.net (195.66.224.48) [AS 5459]</td>
<td>1 London-i2.telia.net (195.66.224.48) [AS 5459]</td>
</tr>
<tr>
<td>3 rou-ifw-mega-transit-1.ethz.ch (129.132.99.72) [AS 559]</td>
<td>3 rou-ifw-mega-transit-1.ethz.ch (129.132.99.72) [AS 559]</td>
</tr>
<tr>
<td>4 dcg.inf.ethz.ch (129.132.130.158) [AS 559]</td>
<td>4 dcg.inf.ethz.ch (129.132.130.158) [AS 559]</td>
</tr>
</tbody>
</table>

AS 559 ETHZ 129.132.0.0/16...

Summary of simplifications

<table>
<thead>
<tr>
<th></th>
<th>sjc</th>
<th>k-peer</th>
<th>m-root</th>
</tr>
</thead>
<tbody>
<tr>
<td>probed hosts</td>
<td>301752</td>
<td>143193</td>
<td>143193</td>
</tr>
<tr>
<td>completed traceroutes</td>
<td>220888 73%</td>
<td>89667 63%</td>
<td>89317 62%</td>
</tr>
<tr>
<td>non-redundant IP AS path – BGP AS path pairs</td>
<td>60271 20%</td>
<td>36950 26%</td>
<td>38527 27%</td>
</tr>
<tr>
<td>incongruent paths</td>
<td>11279 4%</td>
<td>36888 26%</td>
<td>38460 27%</td>
</tr>
<tr>
<td>covered BGP prefixes (of total announced 113563)</td>
<td>58037 51%</td>
<td>36170 31%</td>
<td>37292 32%</td>
</tr>
</tbody>
</table>

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)

<table>
<thead>
<tr>
<th>IX AS</th>
<th>freq</th>
<th>cum. %</th>
<th>IX AS</th>
<th>freq</th>
<th>cum. %</th>
<th>IX AS</th>
<th>freq</th>
<th>cum. %</th>
</tr>
</thead>
<tbody>
<tr>
<td>sjc</td>
<td>6695</td>
<td>4.8%</td>
<td>k-peer</td>
<td>1200</td>
<td>36008</td>
<td>98.0%</td>
<td>m-root</td>
<td>2500</td>
</tr>
<tr>
<td>5450</td>
<td>1187</td>
<td>74.3%</td>
<td>10764</td>
<td>305</td>
<td>98.8%</td>
<td>7517</td>
<td>1423</td>
<td>95.0%</td>
</tr>
<tr>
<td>7527</td>
<td>546</td>
<td>86.4%</td>
<td>6695</td>
<td>252</td>
<td>99.5%</td>
<td>6695</td>
<td>948</td>
<td>97.7%</td>
</tr>
<tr>
<td>total</td>
<td>4311</td>
<td>total</td>
<td>total</td>
<td>total</td>
<td>total</td>
<td>total</td>
<td>total</td>
<td>total</td>
</tr>
</tbody>
</table>

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)

<table>
<thead>
<tr>
<th>group</th>
<th>freq</th>
<th>cum. %</th>
<th>group</th>
<th>freq</th>
<th>cum. %</th>
<th>group</th>
<th>freq</th>
<th>cum. %</th>
</tr>
</thead>
<tbody>
<tr>
<td>sjc</td>
<td>MCI</td>
<td>565</td>
<td>16%</td>
<td>Level3</td>
<td>3518</td>
<td>5.3%</td>
<td>MCI</td>
<td>1719</td>
</tr>
<tr>
<td>k-peer</td>
<td>SBC</td>
<td>371</td>
<td>29%</td>
<td>CitW</td>
<td>2465</td>
<td>60%</td>
<td>Sprint</td>
<td>444</td>
</tr>
<tr>
<td>m-root</td>
<td>Quest</td>
<td>557</td>
<td>42%</td>
<td>Telia</td>
<td>1329</td>
<td>65%</td>
<td>Quest</td>
<td>377</td>
</tr>
<tr>
<td></td>
<td></td>
<td>494</td>
<td>54%</td>
<td>Quest</td>
<td>1322</td>
<td>70%</td>
<td>SBC</td>
<td>369</td>
</tr>
<tr>
<td></td>
<td></td>
<td>323</td>
<td>62%</td>
<td>AT&amp;T</td>
<td>1001</td>
<td>7%</td>
<td>Sprint</td>
<td>355</td>
</tr>
<tr>
<td>total</td>
<td>4331</td>
<td>total</td>
<td>total</td>
<td>total</td>
<td>total</td>
<td>total</td>
<td>total</td>
<td>total</td>
</tr>
</tbody>
</table>

Remaining incongruities (1)

<table>
<thead>
<tr>
<th>sjc</th>
<th>freq</th>
<th>t - b</th>
<th>freq</th>
<th>t - b</th>
<th>freq</th>
<th>t - b</th>
<th>freq</th>
</tr>
</thead>
<tbody>
<tr>
<td>1599</td>
<td>2</td>
<td>-1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1151</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>474</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>394</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3126</td>
<td>65%</td>
<td>+</td>
<td>373</td>
<td>70%</td>
<td>+</td>
<td>15765</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>4819</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Remaining incongruities (2)

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

<table>
<thead>
<tr>
<th>Operation</th>
<th>sjc</th>
<th>k-peer</th>
<th>m-root</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insertions</td>
<td>2788</td>
<td>58%</td>
<td>2764</td>
</tr>
<tr>
<td>Deletions</td>
<td>1132</td>
<td>23%</td>
<td>1</td>
</tr>
<tr>
<td>Substitutions</td>
<td>813</td>
<td>17%</td>
<td>1813</td>
</tr>
<tr>
<td>Mixture</td>
<td>86</td>
<td>2%</td>
<td>683</td>
</tr>
<tr>
<td>total paths</td>
<td>4819</td>
<td>5261</td>
<td>16927</td>
</tr>
</tbody>
</table>

Note: Remaining length differences other than those included in the table

1 TraceRoute AS path length
2 BGP AS path length

Operation
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

<table>
<thead>
<tr>
<th></th>
<th>ejc</th>
<th>k-peer</th>
<th>m-root</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-redundant IP AS path – BGP AS path pairs</td>
<td>60271 (64%)</td>
<td>36950 (100%)</td>
<td>38257 (100%)</td>
</tr>
<tr>
<td>Incongruent paths</td>
<td>11279 (100%)</td>
<td>36888 (100%)</td>
<td>38460 (100%)</td>
</tr>
</tbody>
</table>

| IX ASes             | 3749 (13%) | 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 internet’s 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
  \[ N\{X > x\} = a \exp\left(-\frac{x}{b}\right)^c \]

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

Topological resilience of giant component

<table>
<thead>
<tr>
<th>nodes</th>
<th>629647</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acyclic nodes</td>
<td>60008</td>
<td>90.5%</td>
</tr>
<tr>
<td>Core nodes</td>
<td>52505</td>
<td>8.3%</td>
</tr>
<tr>
<td>Giant component</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Nodes removed from giant component by outdegree

- 0
- 500
- 1000
- 1500
- 2000
- 2500
- 3000
- 3500
- 4000
- 4500
- 5000
- 5500

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?