Overview

• CAIDA:
  • conducting research
  • building infrastructure
  • data collection and curation
  • tool development
  • informing policy
  • workshops
Research

• Macroscopic Topology Project
  • IPv4 and IPv6 topology discovery
  • hostname collection
  • Alias Resolution
  • Router to AS assignment
  • dual graph
  • AS relationships
Research (cont)

• IPv4 and IPv6 topology discovery
  • daily collected
  • available to researchers

• hostnames
  • collected for every IP address in topology
  • released per cycle
Research (cont)

• Alias Resolution
  • Collapse interfaces to produce router-level graph
  • Analysis run across two months of topology data
  • Additional measurement collected
  • iffinder
  • MIDAR

• Publication
Research (cont)

• Router to AS mapping (presented later)

• Developing new AS Relationships algorithm
  • improved running time (hours vs days)
  • prevent creation of acyclic components
  • validation by providers (work in progress)
Research
Research (cont)

- Internet Topology Data Kit (ITDK) Process
Modeling Complex Networks
Research (cont)

• Modeling Complex Networks
  • Internet AS evolution modeled with multiclass preferential attachment
  • scaling dual graph (router+AS)
  • Internet is made up of ultra small worlds (tiny path relative to number of nodes)

• Routing Complex Networks
  • hyperbolic metric space allows for shortest path routing without global knowledge
Research (cont)

publications

• Curvature and Temperature

• Greedy forwarding

• Navigability
Research (cont)

• Domain Name System (DNS)

• S. Castro, M. Zhang, W. John, D. Wessels, kc claffy, “Understanding and preparing for DNS evolution”, TMA 2010
Research (cont)

• Internet Traffic Analysis
  • Traffic classification overview
    http://www.caida.org/research/traffic-analysis/classification-overview/
    • taxonomy of traffic classification papers and data sets
    • provides filtering/search/sort functionality to taxonomy
    • many media rich entertainment application emerging

• Routing Asymmetry Study
  http://www.caida.org/research/traffic-analysis/asymmetry
  • traffic symmetry does not hold for network location beyond intranet and access links
Collaborative Research

• Using Ark to examine source address spoofing
  • how many networks allow packets with spoofed IP addresses to leave their network
  • working on adding IPv6

• R. Beverly, A. Berger, Y. Hung, kc claffy “Understanding the Efficacy of Deployed Internet Source Address Validation Filtering”, IMC November 2009.

• Improving the efficiency of topology probing
  • implemented Doubletree using Marinda (tuple space)
  • Matthew Luckie and Alistair King
Infrastructure

- Archipelago
  - CAIDA's active measurement infrastructure
  - 43 monitors – growing 1 or 2 per month
  - 11 w/ IPv6 connectivity
  - currently used for
    - Team-probing experiment to collect IPv4 and IPv6 topology
    - alias resolution measurements
    - Spoofer experiment
Infrastructure (cont)

• Passive Trace Capture
  • Tier 1 OC192 backbone link packet header captures

• UCSD Network Telescope
  • 2 days of telescope dataset
    http://www.caida.org/data/passive/telescope-2days-2008_dataset.xml
  • 3 days of Conficker dataset
    http://www.caida.org/data/passive/telescope-3days-conficker_dataset.xml
Data

- **OC192 backbone**: 8.5 TB (3.6 anonymized; 4.9 unanonymized) – curation to quarterlies will reduce
- **UCSD telescope**: 3.4 TB on disk (30 day window) 4.8 T on samqfs
- **topology**: 12.3 TB (skitter+ark uncompressed)
  - routed ipv4: 2.3TB since Sep 2007
  - routed ipv6: 275MB since Dec 2008

**Total**: ~30TB (as of 15 Feb 2010)
## how many total requests for the data?

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Requests</th>
<th>Approved</th>
<th>Accessed</th>
<th>Since</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backscatter</td>
<td>451</td>
<td>241</td>
<td>207</td>
<td>Feb 2003</td>
</tr>
<tr>
<td>Passive</td>
<td>799</td>
<td>585</td>
<td>483</td>
<td>Feb 2004</td>
</tr>
<tr>
<td>Topology</td>
<td>614</td>
<td>372</td>
<td>290</td>
<td>Jul 2004</td>
</tr>
<tr>
<td>Witty</td>
<td>58</td>
<td>38</td>
<td>32</td>
<td>Mar 2008</td>
</tr>
<tr>
<td>Telescope</td>
<td>36</td>
<td>20</td>
<td>16</td>
<td>Jul 2009</td>
</tr>
<tr>
<td>DNS-RTT</td>
<td>40</td>
<td>23</td>
<td>18</td>
<td>Aug 2006</td>
</tr>
<tr>
<td></td>
<td>1998</td>
<td>1279</td>
<td>1046</td>
<td></td>
</tr>
</tbody>
</table>
Data request stats

- All requests (cumulative)

All requests (as of 2010/02/01)

Total received: 1998
Total approved: 1279
Tools

• Coral Reef: software for traffic analysis
  • traffic report generator
• Geocompare: survey of geolocation tools
• topostats: topology statistics web interface
• APAR: software for analytical alias resolution
• RadarGun: software for active measurement alias resolution
• MAARS: software package for alias resolution regularly updated AS-level, router-level and dual graph – ITDKs
Tools (cont)

- Example: Report Generator - Chicago OC192 monitor
Policy

• Policy to support Empirical Internet Research

• Privacy-Sensitive Data Sharing

• Developing Ethical Guidelines for Internet Research
Policy (cont)

- Emperical Internet research to Support Policy
- Advising Regulators on “Network Neutrallity” regulation
  - Historical and Architectural context for Internet Traffic Management (relied heavily on .JP and .CA examples)
Workshops

- Active Internet Measurement Systems (AIMS)
- Workshop on Internet Economics
- Joint workshop with WIDE/CASFI

http://www.caida.org/workshops/
slides

Please email your slides to CAIDA.

- webmaster@caida.org
- talk title:
- author name:
- workshop: CAIDA-WIDE-CASFI
- topic:

active data, bandwidth estimation, data, dns, education, measurement methodology, overview, peer-to-peer, routing, security, software/tools, topology, trends, visualization, workshop report