Overview

CAIDA activities

- conducting research
- building infrastructure
- data collection and curation
- software tool development
- informing policy
- workshops
Research

• Macroscopic Topology Project
  • IPv4 and IPv6 topology discovery
  • Hostname collection
  • Alias resolution
  • Router-to-AS assignment
  • AS relationships
Research (cont)

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

- hostnames
  - collected for every observed IP address
  - released per cycle
Research (cont)

• Scalable IPv4 alias resolution
  • aggregate interfaces to form router-level graph
  • uses two-month window of topology data
  • integrates additional measurements:
    • iffinder
    • MIDAR

• tech report (revised version submitted to ToN)
Research (cont)

• Measuring the impact of per-flow load balancing on link inference methods using classic traceroute

• Examined effects of false inferences on accuracy of topology map, both macroscopic and per-ISP


[Diagram showing interface graphs inferred with multipath detection algorithm (MDA) and classic traceroute.]

A→E is a valid router-level link if A and B are aliases
A→E is a valid router-level link if C and E are aliases
A→E is a false link in this router-level topology
Research (cont)

• Geocompare: a comparison of public and commercial geolocation databases
    - Providers agree with the majority 98%~99.7% with the majority
    - RIPE-NCC’s address were the source of many disagreements
    - MaxMind Geo and Netacuity are within 33 km of each other? for 79% and 80% respectively
    - Netacuity had shortest median distance to the Tier 1 and Planet Lab ground truth data sets.
    - MaxMind Geo had shortest median distance to French Home Networks ground truth data set. why?
Research: router infoviz prototype)
Research (cont.)

• Internet Topology Data Kit (ITDK) Process
Research (cont)

- Modeling Complex Networks
  - hidden variables influence structure of network
  - (which nodes are connected)
  - variables form a hidden metric space that can be used to enable shortest path routing without global knowledge of topology
  - suggests potential direction toward infinitely scalable Internet routing architecture
  - rich research results with broad applications beyond CS/networking
Research (cont)

Publications


Research (cont)

• Internet Peering Analysis

• Economics of Transit & Peering Interconnections
  http://www.caida.org/funding/netse-econ/
  • create empirically parameterized computational tools for modeling
  • enable broader validation
  • several papers published already (on caida web site)

• Workshop on Internet Economics 2012
  • last week at UCSD (1-2 December)
  • peering, pricing, industry structure, IPv6
  • report forthcoming
Research (cont)

• Dark Space Analysis
• CRI-Telescope: A Real-time Lens into Dark Address Space
  http://www.caida.org/funding/cri-telescope/
  • enable near realtime sharing of traffic data in a way that maximizes data utility for research while protecting user privacy,
  • improve classification of traffic using more modern taxonomy (DOS attacks, vulnerability scans, malware spread),
  • build infrastructure to allow vetted researchers to run analysis
  • workshop for darknet researchers and operators planned for May 2012
Research (cont)

- Internet background radiation (darknet traffic) as an indicator of impacts of natural disaster on nearby communications infrastructure.
  - Example: ratio of unique IP addresses reaching UCSD darknet in two successive 24-hour periods from radius around Tohoku earthquake epicenter
Research (cont)

New project: Empirically exploring IPv6 evolution

Topology, performance, traffic
http://www.caida.org/funding/nets-ipv6/

• Three proposed tasks:
1. Correlating rate of IPv6 deployment with socioeconomic parameters
2. Capture a comprehensive view of IPv6 topology from core to edge
3. Conducting quantitative assessment of IPv6 performance, including the impact of transition technologies and traffic characteristics

[Matthew Luckie starts postdoc on this project in January]
Infrastructure

- Archipelago
  - CAIDA's active measurement infrastructure
  - 57 monitors – growing 1 or 2 per month
  - 28 w/ IPv6 connectivity
  - currently used for
    - Team-probing experiment to collect IPv4 and IPv6 topology
    - alias resolution measurements
    - researcher experiments, e.g., spoofer
Infrastructure (cont)

- Passive Trace Capture
  - Tier 1 10GE backbone link packet headers
- 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
- Near Realtime Telescope Data
  (Invitation only)
Data

- **10GE backbone:** 16.9 TB (3.6 anonymized; 4.9 unanonymized) – curation to quarterlies will reduce
- **UCSD telescope:** 5 TB on disk (30 day window) 85 TB on tape (soon to migrate to cloud disk)
- **topology:** 5.5 TB (skitter+ark compressed)
  - routed ipv4: 5.4TB since Sep 2007
  - routed ipv6: 8GB since Dec 2008

**Total:** ~105TB (as of 30 Nov 2011)
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>530</td>
<td>311</td>
<td>256</td>
<td>Feb 2003</td>
</tr>
<tr>
<td>Passive</td>
<td>1140</td>
<td>878</td>
<td>744</td>
<td>Feb 2004</td>
</tr>
<tr>
<td>Topology</td>
<td>1201</td>
<td>736</td>
<td>559</td>
<td>Jul 2004</td>
</tr>
<tr>
<td>Witty</td>
<td>134</td>
<td>100</td>
<td>88</td>
<td>Mar 2008</td>
</tr>
<tr>
<td>Telescope</td>
<td>95</td>
<td>63</td>
<td>52</td>
<td>Jul 2009</td>
</tr>
<tr>
<td>DNS-RTT</td>
<td>52</td>
<td>30</td>
<td>24</td>
<td>Aug 2006</td>
</tr>
<tr>
<td></td>
<td>3152</td>
<td>2118</td>
<td>1723</td>
<td></td>
</tr>
</tbody>
</table>
Data request stats

- All requests (cumulative)

3335 requests received
2251 requests approved
1835 accounts accessed
Data download stats

- All downloads (cumulative)

*47 TiB downloaded*
Tools

- Coral Reef: software for traffic analysis
  - traffic report generator
- Geocompare: survey of geolocation tools
- topostats: topology statistics web interface
- kapar: software for analytical alias resolution
- MIDAR: software for active measurement alias resolution (partly released)
- Motu: a tool for dealiasing pairs of IPv4 addresses.
Tools (cont)

• Example: Report Generator - Chicago 10GE monitor
Empirical Internet research to Support Policy

Advising Researchers on Guidelines for Ethical Research and Data Disclosure


Workshops

• Active Internet Measurement Systems (AIMS)
  • Feb ’09, Feb ’10, Feb ’11, coming Feb ‘12
• Workshop on Internet Economics
  • Sep ’09, Dec ‘11
• Joint workshop with WIDE/CASFI
  • Aug ’08, Apr ’09, Apr ’10, Dec ‘11

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