Center for Applied Internet Data Analysis (CAIDA)

http://www.caida.org/
info@caida.org

UC San Diego

caida

Overview

The Center for Applied Internet Data Analysis (CAIDA) conducts network research and builds research infrastructure to support large-scale data collection, curation, and data distribution to the scientific research community. The group, located at the federally funded San Diego Supercomputer Center located at the University of California, San Diego, designs, deploys and maintains a growing number of computational, data analysis and visualization services. The group also ships and maintains small form factor measurement instrumentation to networks around the world, extending its Archipelago (Ark) Internet measurement platform for use by the network and cybersecurity research community. CAIDA researchers develop novel techniques to collect, analyze, query and visualize the resulting data.

Mission Statement: CAIDA investigates practical and theoretical aspects of the Internet, focusing on activities that:

- ⇒ provide insight into the macroscopic function of Internet infrastructure, behavior, usage, and evolution,
- ⇒ foster a collaborative environment in which data can be acquired, analyzed, and (as appropriate) shared,
- ⇒ improve the integrity of the field of Internet science, and
- ⇒ inform science, technology, and communications public policies.

Research and Analysis

CAIDA's research spans Internet topology, routing, security, economics, future Internet architectures, and public policy. Our infrastructure, software development, and data sharing activities support measurement-based Internet research, both at CAIDA and around the world, with focus on the health and integrity of the global Internet ecosystem.

Mapping the Internet. We pursue Internet cartography to improve our Internet topology mapping capabilities using our expanding and extensible Ark measurement infrastructure. We work to improve the accuracy and sophistication of our topology annotation capabilities, including classification of ISPs and their business relationships. Using our evolving IP address alias resolution measurement system, we collect, curate, and release our flagship data product, the Internet Topology Data Kit (ITDK).

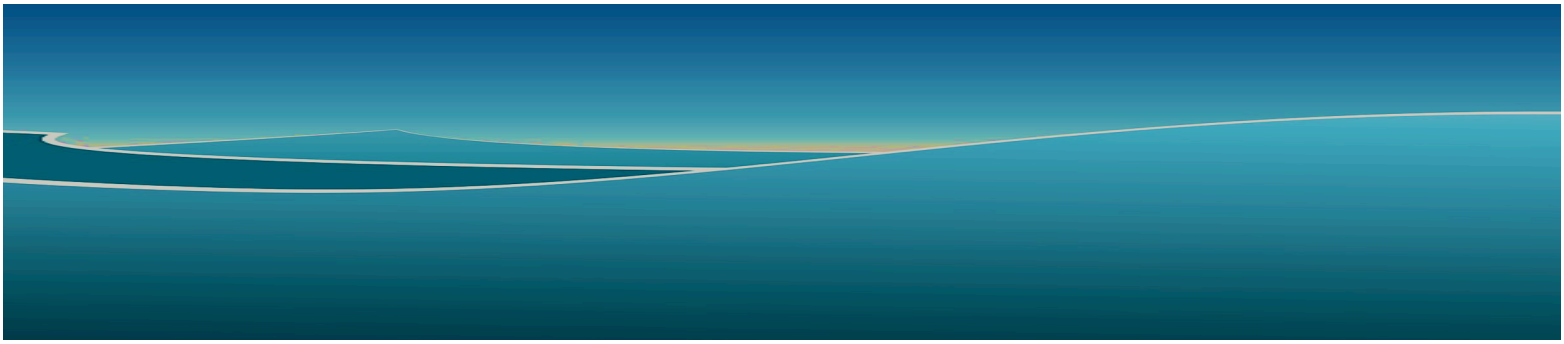
Mapping Interconnection Connectivity and Congestion. We use the Ark infrastructure to support an ambitious collaboration with researchers at Massachusetts Institute of Technology (MIT) to map the rich mesh of interconnection in the Internet, with a focus on congestion induced by evolving peering and traffic management practices of Content Distribution Networks (CDN) and access ISPs, including methods to detect and localize the congestion to specific points in networks. We undertake studies to pursue different dimensions of this challenge: identification of interconnection borders from comprehensive measurements of the global Internet topology; identification of the actual physical location (facility) of an interconnection in specific circumstances; and mapping observed evidence of congestion at points of interconnection. We produce related data collection and analysis to enable evaluation of these measurements in the larger context of the evolving ecosystem: quantifying a given ISP's global routing footprint; classification of autonomous systems (ASes) according to business type; and mapping ASes to

CAIDA's projects made possible by support from:

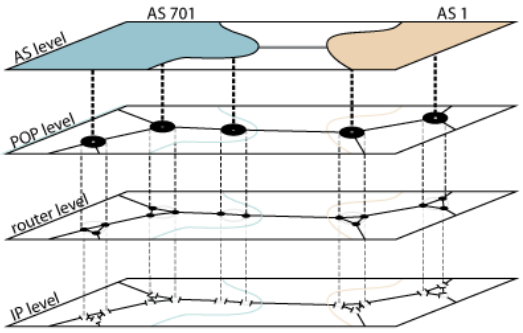
Department of Homeland Security
Science & Technology Division

National Science Foundation

SDSC
SAN DIEGO SUPERCOMPUTER CENTER



their owning organizations. In parallel, we examine the peering ecosystem from an economic perspective, exploring fundamental weaknesses and systemic problems of the current economic framework of Internet interconnection that will continue to cause peering disputes between ASes.



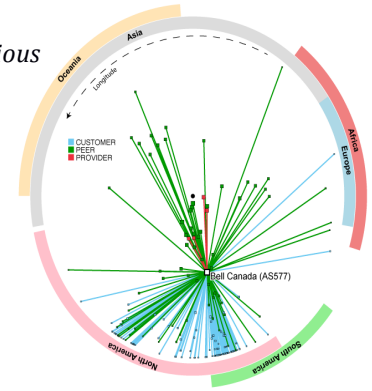
Mapping the Internet at various layers of abstraction

Monitoring Global Internet Security and Stability. Our monitoring projects focus on security and stability aspects of the global Internet: traffic interception events (hijacks), macroscopic outages, and network filtering of spoofed packets. These projects leverages our existing Ark measurement infrastructure, but each has also required the development of new measurement and data aggregation and analysis tools and infrastructure, now at various stages of development. Recently, we released BGPstream, a software framework for processing large amounts of historical and live BGP measurement data. BGPstream serves as one data analysis component of our outage-detection monitoring infrastructure, a prototype of which was operating at the end of 2015. We also undertake studies that leverage the results of internet scanning and other unsolicited traffic to infer macroscopic properties of the Internet.

Future Internet Architectures. The current TCP/IP architecture is showing its age, and the slow uptake and architectural limitations of its proposed upgrade, IPv6, has inspired NSF and other research funding agencies around the world to invest in research on entirely new Internet architectures. We help launch this moonshot from several angles -- routing, security, testbed, management -- while also pursuing and publishing results of six empirical studies of IPv6 deployment and evolution.

Public Policy. Our final research thrust is public policy, an expanding area due to requests from policymakers for empirical research results or guidance to inform industry tussles and communication policies. Most notably, the FCC and AT&T selected CAIDA to be the Independent Measurement Expert to support measurement conditions that accompanied the AT&T/DirecTV merger, which turned out to be as much of a challenge as it was an honor.

We have published several policy analysis papers discussing how to optimize different public policy outcomes in the face of a rapidly evolving information and communication technology landscape. We also contribute to the development of frameworks and supporting tools for ethical assessment of Internet measurement research methods.



The secret keys to Internet science

The most under-appreciated aspect of Internet science is the need for longitudinal support for measurement infrastructure, similar to the need for telescopes and satellites to study outer space. The second most under-appreciated aspect is the incentive misalignments of many players in the ecosystem to supporting rigorous scientific inquiry. Legal, competitive, security, privacy, and cost concerns inhibit development of community measurement instrumentation as well as sharing of data to validate scientific methods and results. Federal support for not only Internet research, but also measurement infrastructure and data sharing efforts, is essential to success of what has become a critical infrastructure science.

CAIDA's projects made possible by support from:

Department of Homeland Security
Science & Technology Division

National Science Foundation

SDSC
SAN DIEGO SUPERCOMPUTER CENTER