Cartographic Capabilities for Critical Cyber Infrstructure

Center for Applied Internet Data Analysis (CAIDA)

KC Claffy

UC San Diego

Overview

Researchers at the Center for Applied Internet Data Analysis (CAIDA) build and maintain Internet measurement infrastructure, and develop novel techniques to collect, analyze, query and visualize the resulting data. As part of this effort, CAIDA creates and supports two services t for users to interact with data:

- *AS Rank* - integrates multiple diverse types and sources of data (e.g., BGP routing data from the Route Views and RIPE NCC RIS projects, WHOIS data, PeeringDB, CAIDA's inferred AS relationships and AS classification datasets, geolocation annotations) into an updated coherent representation of macroscopic Internet topology at multiple granularities.
- *Vela* – enables on-demand measurements and queries of historical topology data, and provides analysis and visualization that increases accessibility of data and insights to a variety of disciplines and domains.


Customer Need

The "cyber threat" is one of the most serious challenges we face as a nation. America's economic prosperity in the 21st century crucially depends upon a secure and trustworthy communications fabric. Yet, we lack a thorough understanding of the vulnerabilities of the global Internet. Versatile, secure measurement infrastructures; reliable, representative, high-quality Internet data sets; and advanced measurement and analysis tools are rarely available to researchers and developers. CAIDA researchers bridge this gap, advancing situational awareness of the Internet by providing tools and data that support topological analysis and predictive modeling of the structure, dynamics, performance, and vulnerabilities of the global Internet.

Our Approach

We use the *scamper* and MIDAR tools for systematic data collections on the Archipelago (Ark) measurement platform designed, deployed, and maintained by CAIDA. Distributed all over the world, Ark monitors (162 and growing, see map) are tailored to support active network measurements. The monitors use *scamper* to send probes and collect traceroute data to all routed /24 networks in the IPv4 address space and to all announced IPv6 prefixes (/48 or shorter) every 2-3 days. Researchers and analysts use this data is to derive maps of the Internet at various granularity levels: IP, router, PoP, and Autonomous Systems (AS) (shown below). The Ark monitors run the MIDAR tool for alias resolution measurements to obtain the input data necessary for constructing router-level Internet topologies. These topologies form the basis of CAIDA's Internet Topology Data Kits (ITDK), periodically produced and distributed to network and security researchers and analysts.
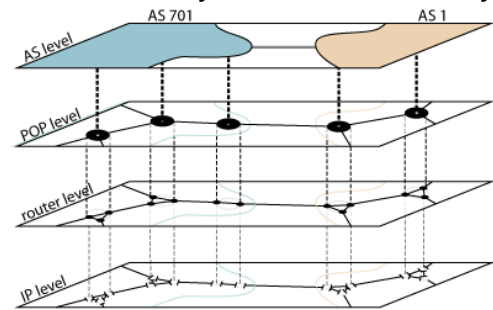
Benefits

The integration of strategic measurement and analysis capabilities allows us to deliver comprehensive annotated Internet topology maps and a platform capable of critical Internet infrastructure security assessments.

ITDK, CAIDA's flagship product, contains richly annotated topology maps of the observable Internet at multiple granularities, providing a more detailed topological view than ever achieved before. These data sets enable empirical research in the network and security fields, advance our ability to identify, monitor, and model critical cyber infrastructure, and deepen our insight into the structure, behavior, and evolution of the global Internet.

CAIDA researchers continue to increase the number of Ark vantage points, refine measurement methods, develop new tools, and improve analysis and inference algorithms. The DHS S&T-funded Ark active measurement platform supports cybersecurity-related situational awareness through macroscopic active measurements, including projects that estimate the geographic diversity of an Autonomous System (AS) presence and connectivity (depicted for one AS and its neighbors in the diagram on right). Cybersecurity-related uses of the platform include measurement of security best practice compliance, IPv4 and IPv6 stability, TCP security vulnerabilities, middle-box behavior, and detection of large-scale outages and BGP hijacks.

Competitive Advantage

The _AS Rank_ service delivers an ordered list of Autonomous Systems (AS) derived from Internet routing data (BGP), registry data (WHOIS), address ownership and transit metrics associated with each AS and CAIDA's inferred AS relationships. AS Rank allows the user to view the data at various levels of granularity and drill down on a single AS to see its customers, peers, siblings and providers as well as geographical information about its neighbors. It enables vetted operators to enter ground truth to identify and remove false links; to correct wrongly inferred relationships; and to increase the accuracy of geographical annotations. AS Rank represents the state-of-the-art with regard to the completeness, accuracy, and richness of the macroscopic Internet maps we provide to the research community.

The _Vela_ service builds on two underlying command-line tools: one that enables users to execute ad hoc, on-demand measurement experiments on the Ark platform. and a second tool that enables the user to query, process, analyze and visualize historical data (over 43 billion traces, 20 TB, since 2007).
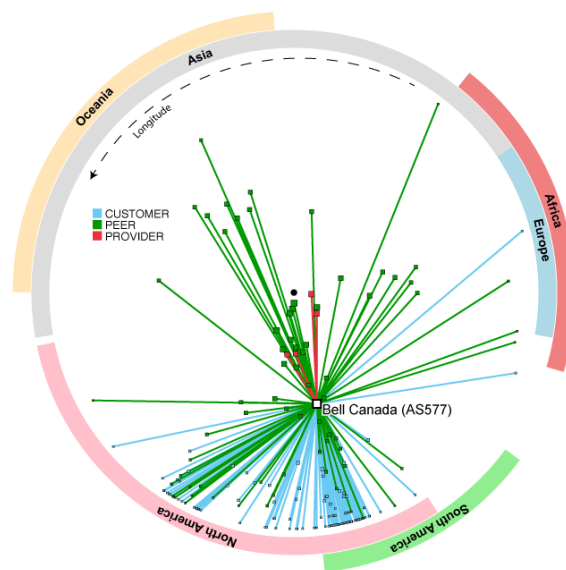
Next Steps

We continue to extend and improve the _AS Rank_ service through deployment of new nodes to fill gaps in our global footprint, improved algorithms for inferring router ownership, location, and mapping interconnection borders. Through _Vela_ we plan to widen the audience for our trace data with pre-made queries, analyses, and visualizations that cater to operators and researchers. For other uses of the Ark platform, contact CAIDA, see our CAIDA's Tools page or the Archipelago (Ark) Measurement Infrastructure page for more details.