ILANDS: Integrated Library for Advancing Network Data Science

Enabling scientific research to understand, assess, model, and improve the evolving security and resilience of the Internet

Kc Claffy (PI), B. Huffaker
CAIDA, UC San Diego

W.H. Allen, S. Huter
NSRC, U Oregon

D. Clark (co-PI)
MIT CSAIL

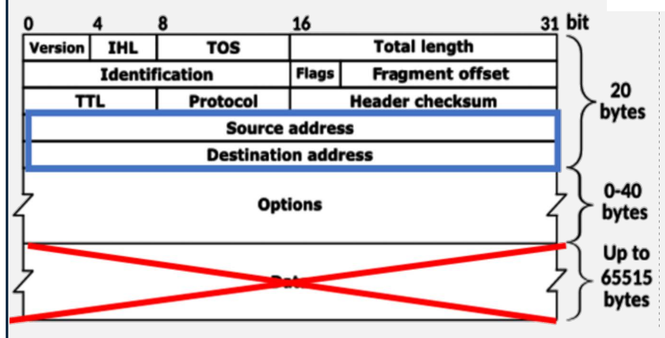
Motivation: Understand the Internet's changing robustness, resilience, security, and stability characteristics requires realistic and representative datasets and infrastructure that can support longitudinal measurements and new measurement experiments.

Goal: Upgrade two measurement infrastructures: **2-way passive traffic capture at 100 Gbps link** and **BGP routing data collection**

Challenges: (1) Growth in traffic and routing data that overwhelms current measurement architectures and infrastructures. (2) Navigating privacy concerns. (3) Sustaining long-term sharing of strategic Internet data sets.

Two-Way Passive Data Collection

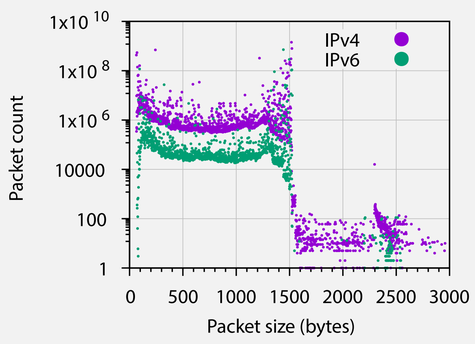
Deployed 100GB monitor at a backbone link in LA (Napatech NIC).



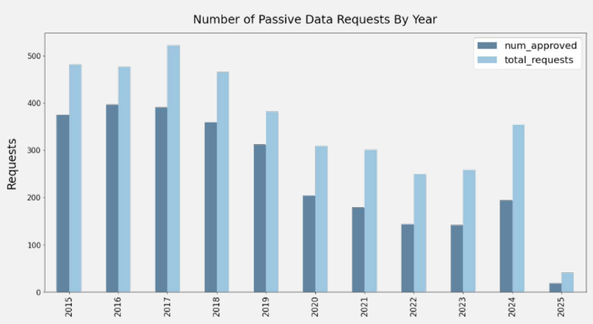
Sharing monthly 1-hour (~2TB) traces since April 2024

CryptptoPan prefix-preserving anonymization.
Discard payload after layer 4 headers .

Reported Uses: anomaly detection; traffic and performance analysis, modeling, and simulations, caching algorithm evaluation; protocol research (high-speed packet processing, adaptive load balancing).



www.caida.org/catalog/datasets/100g_trace_stats



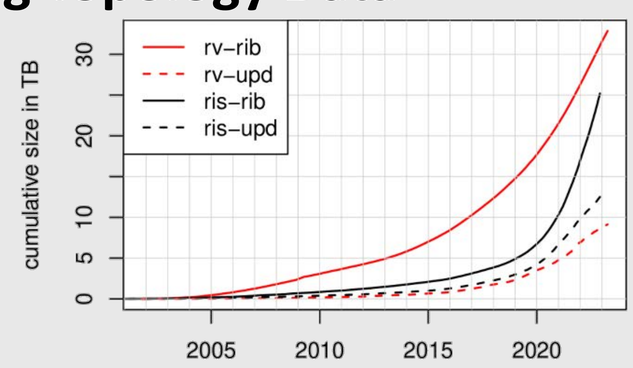
Loss of monitor in 2019 led to fewer requests as traces became stale. Jump in 2024 after deployment of new monitors.

Smaller, 5-second sample (Nov 24) for test/evaluation/lightweight use: catalog.caida.org/datasets/passive_100g_sampler

Acceptable Use Agreement (AUA) prohibits attempts to de-anonymize or o/w misuse data

Internet Routing Topology Data

Growth in BGP data archives challenges data management

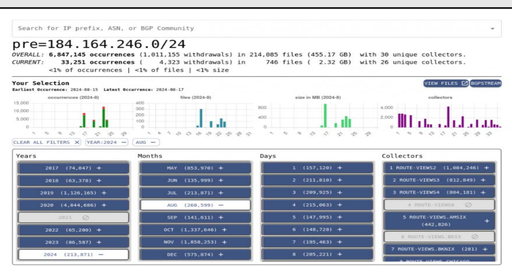


BGP2GO: Lowered barriers to Find, Access, Interoperate, Reproduce (FAIR) data

Indexed: 13 years (12.7TB) of MRT BGP update files into 316+ SQL databases. 12TB, 38.5M IP prefixes (14.9M IPv4 (7.6M /32s), 23.5M IPv6), 420K BGP communities, 117K ASes.

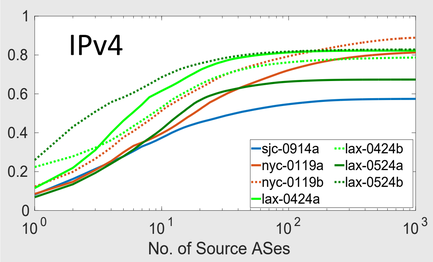
Science Gateway: Select files by prefix, ASN, community; minimize resource needs

Tutorial: Analyzing Prefix Propagation using PEERING Testbed, BGP2GO, and BGPStream—A guide for network operators to advertise a prefix and track its propagation via RouteViews.
blog.caida.org/best_available_data/

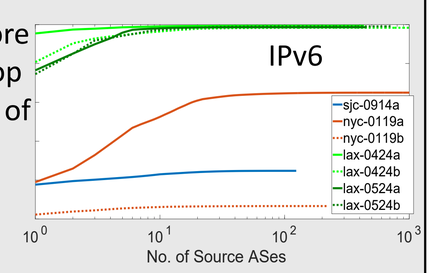


Created new tool for data assurance:
interprets MRT files, identifies data corruption
catalog.caida.org/software/bgp_explain

Integrating traffic and topology data to analyze infrastructure trends



IPv6 traffic significantly more concentrated than IPv4: top 10 IPv6 ASes sent > 96.5% of IPv6 traffic. (IP->AS mapping needs operator approval.)



Community Outreach:

2025 AIMS Workshop – Feb. 8-14

82 Participants; 2-day Hackathon; 60+ presentations on: Active Measurements, BGP, DNS, Traffic

Publications (Oct 2021-Feb 2025)

350 using (10Gb) traces
320 using BGP data

Data Findability, Accessibility, and Interoperability

catalog.caida.org

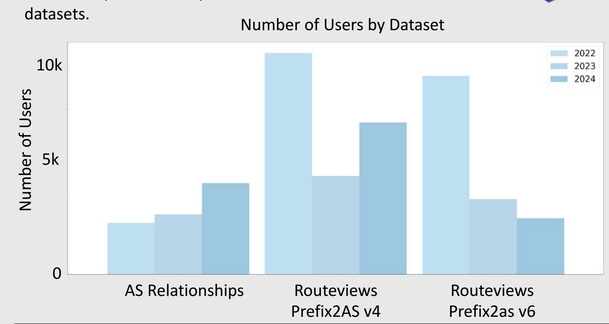
CAIDA's Internet Science Catalog (catalog.caida.org) connects datasets, analysis recipes, research papers, and media, enabling discovery of relationships between them.

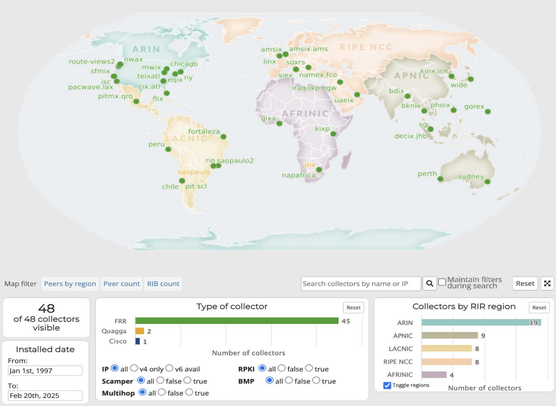
RouteViews Expanded, Created new API

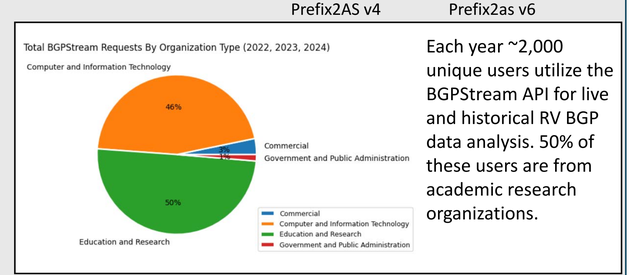
api.routeviews.org

28-year data set of Border Gateway Protocol (BGP) data archive helps operators and researchers investigate Internet routing stability, security and performance. New API supports most common queries of the data.

Ten CAIDA datasets and three APIs are powered by RV BGP data. Each year 20K unique users download CAIDA BGP-related datasets.







Each year ~2,000 unique users utilize the BGPStream API for live and historical RV BGP data analysis. 50% of these users are from academic research organizations.

13 Presentations since Oct 2021

- https://www.routeviews.org/routeviews/routeview-presentations-over-the-years/

Detecting hijacks and tracking global RPKI deployment

- https://www.kentik.com/blog/rpki-rov-deployment-reaches-major-milestone/

"RouteViews data is relied upon in NIST routing security research and the NIST RPKI Monitor, and the National Cybersecurity Strategy Initiative " -- R. Cannon, Policy Analysis/Development, NTIA, U.S. Dept of Commerce

Supported by NSF CIRC grant CNS-2120399 Oct 21 – Sep 26.