NSF · CYBERSECURITY INNOVATION FOR CYBERINFRASTRUCTURE (CICI) · IPAAI · AWARD # 2531134

CANIS - Curated AI-ready Network telescope datasets for Internet Security

Safeguarding the data integrity of the UCSD-NT and transforming the applicability of UCSD-NT in ML/AI contexts

PI Ka Pui (Ricky) Mok · **co-PI** kc claffy (CAIDA, UC San Diego)

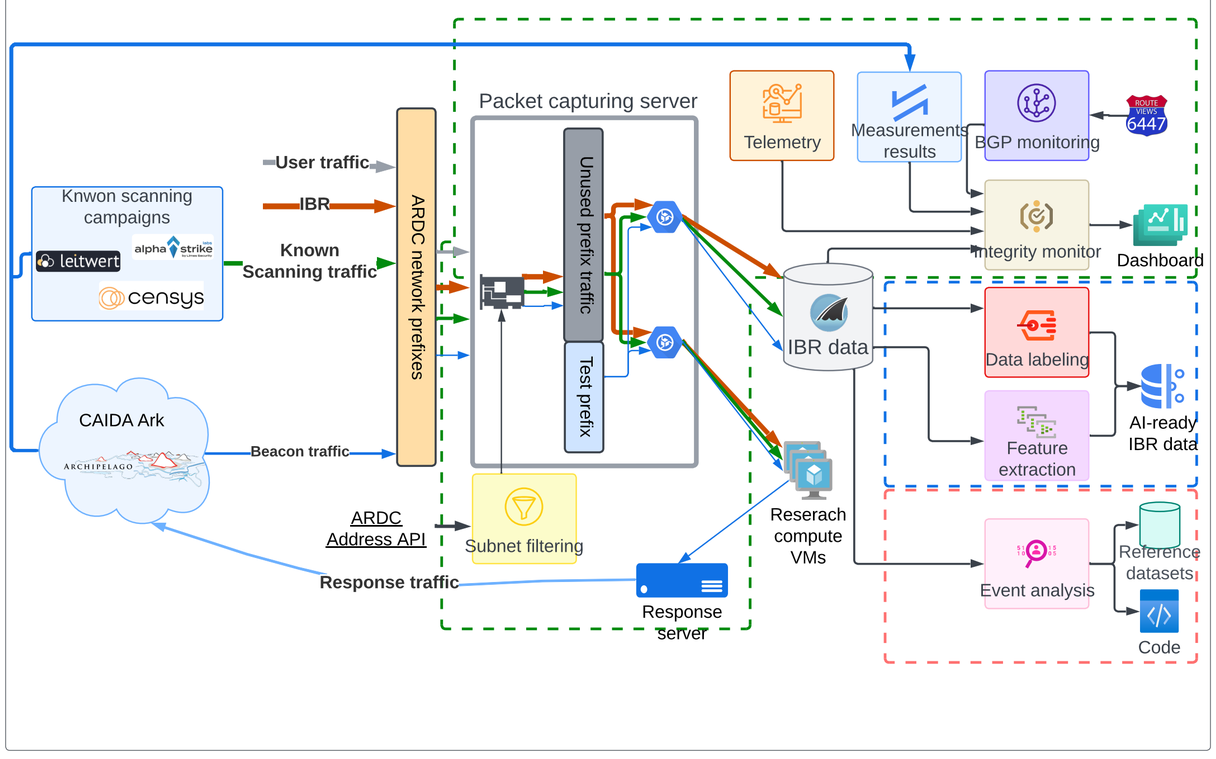
Motivation

- UCSD Network Telescope (UCSD-NT) has been a long-standing NSF-funded scientific cyberinfrastructure (CI), supporting the collection of unsolicited Internet (IPv4) traffic (Internet background radiation or IBR)
- The UCSD-NT faces mounting operational hurdles, harming data integrity:
  - Increasing usage dynamics of underlying address space
  - Rapid growth in IBR traffic volume
  - Disruptions in Internet routing
- Modernize data dissemination pipeline to support ML/AI

Progress

- Deployed Grafana dashboards to monitor the status of UCSD-NT
- Shared data across the country using Open Science Data Federation (OSDF), reducing fraction of using UCSD-NT traffic data
- Prototyping new AI-friendly data query mechanisms using ClickHouse databases
- Enhancing software tools to improve the traffic labeling of IBR, facilitating AI training and evaluation

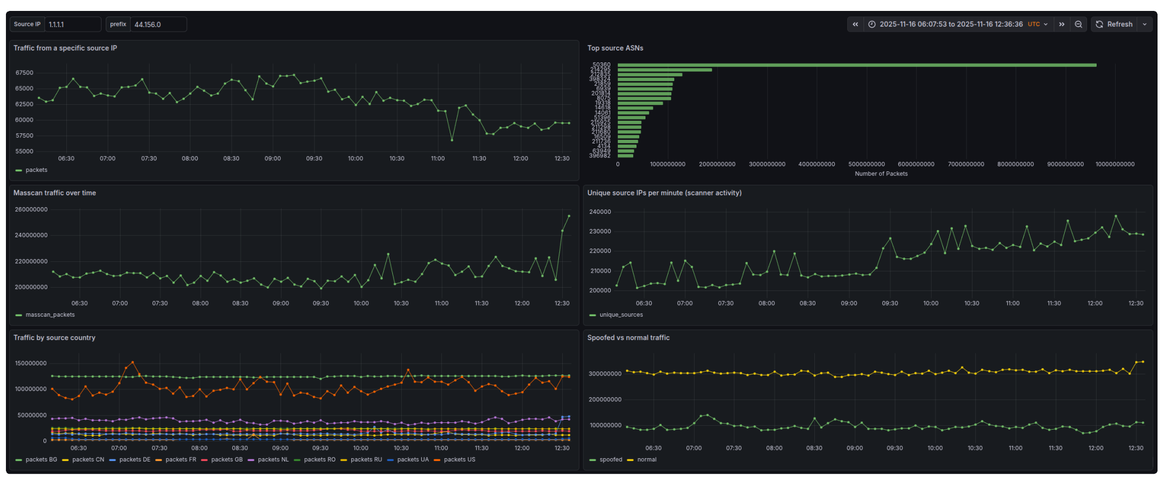
Overview of CANIS



UCSD-NT Montioring dashboards

A screenshot of UCSD-NT monitor dashboard with alerting functions to reduce our reaction time to system issues

AI-ready data exploration



We built a prototype using ClickHouse database which reduces data query time and outputs query results in Apache Parquet format

Publications

Degen et al., "Through a Smaller Lens: Revisiting Opportunistic Analysis using Network Telescopes", Proc. Passive and Active Meaasurement Conference, 2026.

Kempen et al., "A Multi-Cloud View of Internet Background Radiation", To appear in IFIP TMA, 2026.

Gao et al., "DarkBench: Towards Benchmarking the Efficacy of Automatic Darknet Event Detection Methods", Under Review.

Resources

Ricky Mok, "Resource for analyzing UCSD-NT data on OSDF with OSPool", https://github.com/CAIDA/OSP_telescope_tutorial

Ricky Mok, "A tutorial for analyzing UCSD-NT data using SDSC Expanse", https://github.com/CAIDA/telescope_expanse_tutorial

Learn more about the project: https://www.caida.org/funding/cici-canis/
PI Email: cskpmok@caida.org

NSF CICI PI Meeting May 11-12, 2026 · Los Angeles, CA

This material is based upon work supported by the National Science Foundation under Grant #2531134. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.