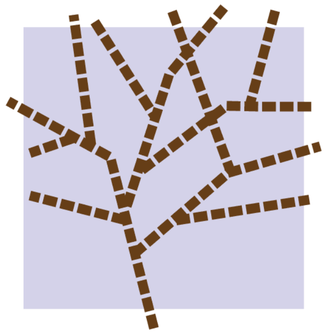
OARC Data Collections

Matthew Pounsett
OARC 43, Prague
October, 2024



DNS-OARC

Domain Name System Operations Analysis and Research Center

Warning:
Contains Aspirations

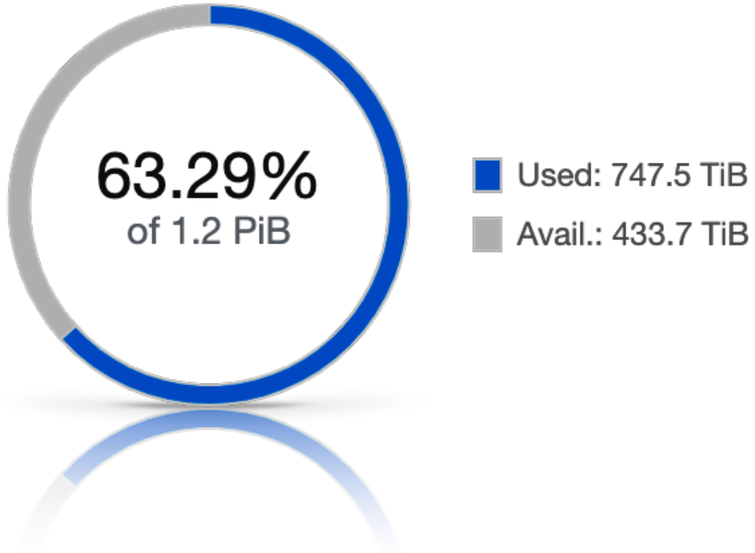
This presentation may mention plans, hopes, dreams, fears, aims, wishes, ambitions, inclinations, pursuits, urges, goals, daydreams, or descriptions of utopias which should not be taken as firm commitments or weak promises. Do not take internally. Keep out of the reach of children.

DNS-OARC

Domain Name System Operations Analysis and Research Center

The Infrastructure

- 1.2PiB Ceph Cluster, 747.5 TiB used
- 354TiB volume with 247.8TiB of actual data
- 3x block duplication, no backups
- 3 Linux analysis servers



Historical Datasets

- Member-contributed DSC
- DNS Testers (DNS Entropy, Port Entropy, Reply Size)
- DNS Testbeds (ODVR, DNS Privacy)
- Derived Data (e.g. 2013 Collisions Study)

Ad-hoc Datasets

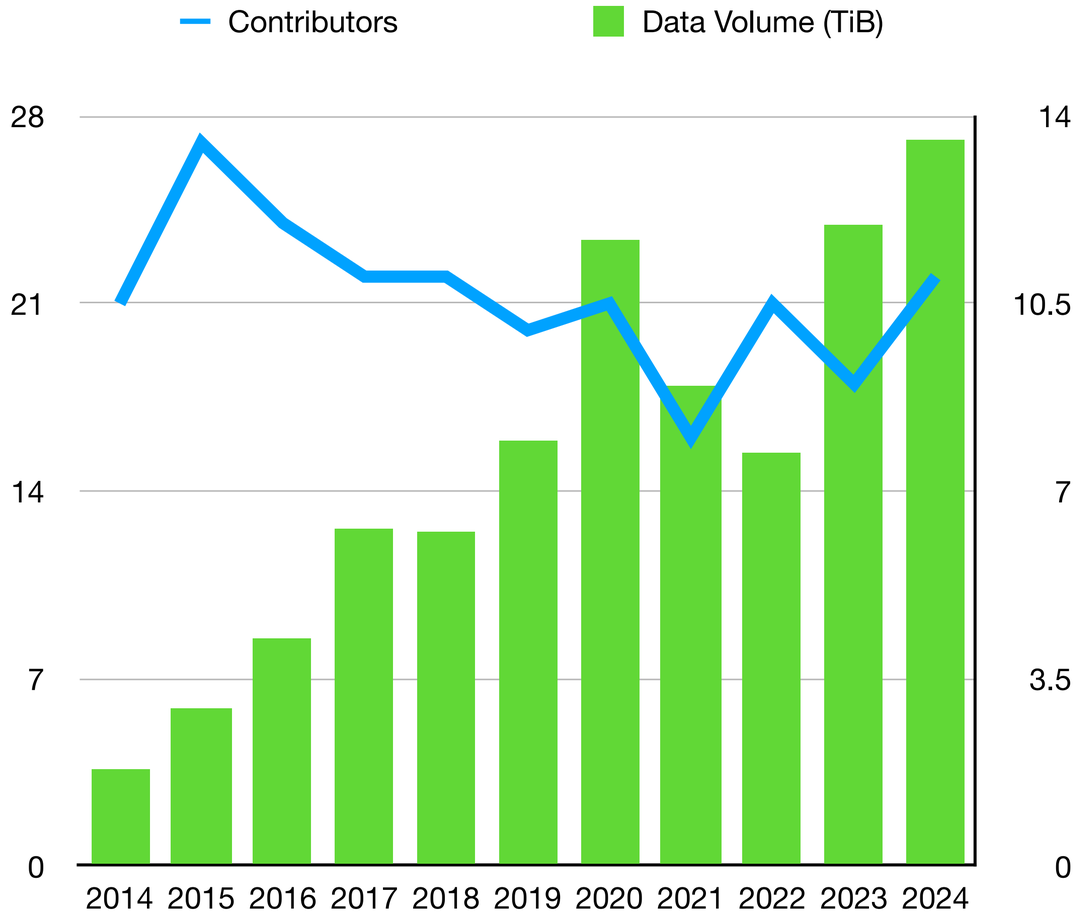
- Non-DITL PCAPs
- Query Logs (e.g. Case Connection Zone Logs)

Active Collections

- Zone File Repository
- Check My DNS Logs
- RSSAC-002 Stats Collection
- Annual “Day in the Life” (DITL) Collection

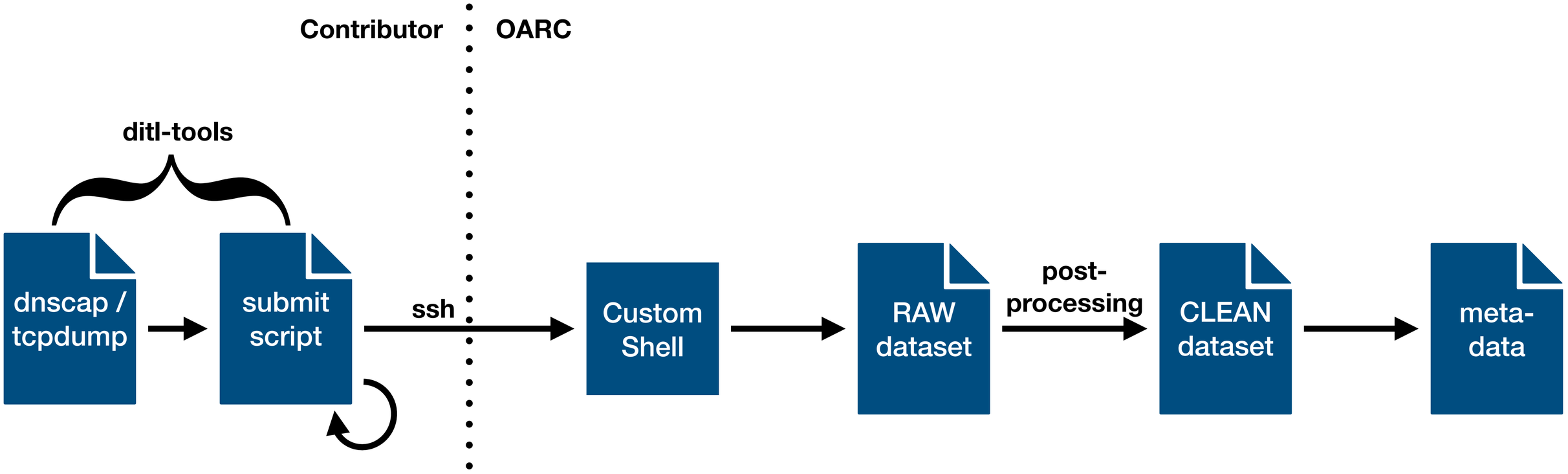
DITL: What is it?

- A Day in the Life of the DNS
- Annual PCAP collection
- 50 straight hours of data
- Multiple operators contributing



DNS-OARC
Domain Name System Operations Analysis and Research Center

DITL: How does it?



DNS-OARC

Domain Name System Operations Analysis and Research Center

How to Get Access

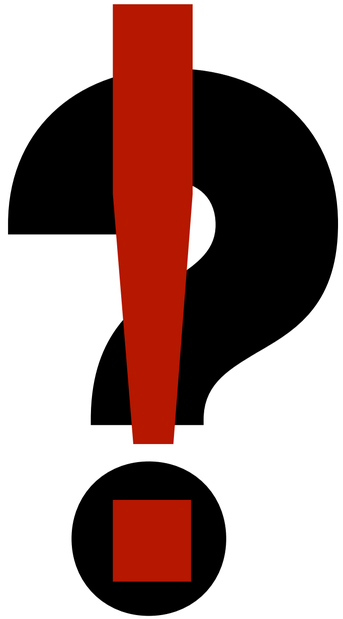
- Requires a Data Sharing Agreement
- Requires a contact in the OARC portal
- Upload an SSH key, and ask

How to Use the Data

- Highly dependent on:
  - Which data
  - What you want to know
- Minimal documentation or help
- Limited processing resources
- But we’re working on that!

DNS-OARC

Domain Name System Operations Analysis and Research Center



DNS-OARC

Domain Name System Operations Analysis and Research Center