Using NetSage to Support ACCESS

Jason Zurawski - zurawski@es.net

Lawrence Berkeley National Laboratory / Energy Sciences Network (ESnet)

2023 Internet2 Technology Exchange

September 20th, 2023

Monitoring using NetSage

- NetSage advanced measurement services for R&E data traffic
  - Better understanding of current traffic patterns across instrumented circuits
  - Better understanding of large flow sources/sinks
  - Performance information for data transfers
- Originally a NSF funded collaboration between Indiana University, LBNL, and University Hawaii Manoa, now primary development is at the Texas Advanced Computing Center
- ACCESS-CI compute resources: https://access.netsage.io

Built around answering questions

- Answers questions asked by network engineers, network owners, and end-users
- Human-readable summaries and patterns
- Big picture overview helps highlight trends and events that can make in-depth analysis of local data more fruitful

Built around answering questions

Interesting pattern.
What is it?

# Flows	Total Volume	Avg Rate
2,791	2.7 TB	4.6 Mb/s

Why so slow?

Common Information

Timestamp	Sensor	5-Tuple Hash	Protocol	Duration
2021-07-15 03:20:53	SingAREN Los Angeles, United States	c9cef62cde44cadab1d152f5e3d37d197f3fa7bf034e74a44b105b1508a3a7f9	tcp	00:00:26

Source Information

Field	Data
Organization	Nanyang Technological University
Country	Singapore
ASN	9419
Subnet	155.69.240.x
Port	443

Destination Information

Field	Data
Organization	Education Bureau, Kaohsiung City Government, Taiwan
Country	Taiwan
ASN	1659
Subnet	140.127.254.x
Port	55125

Singapore to Taiwan
via LA?

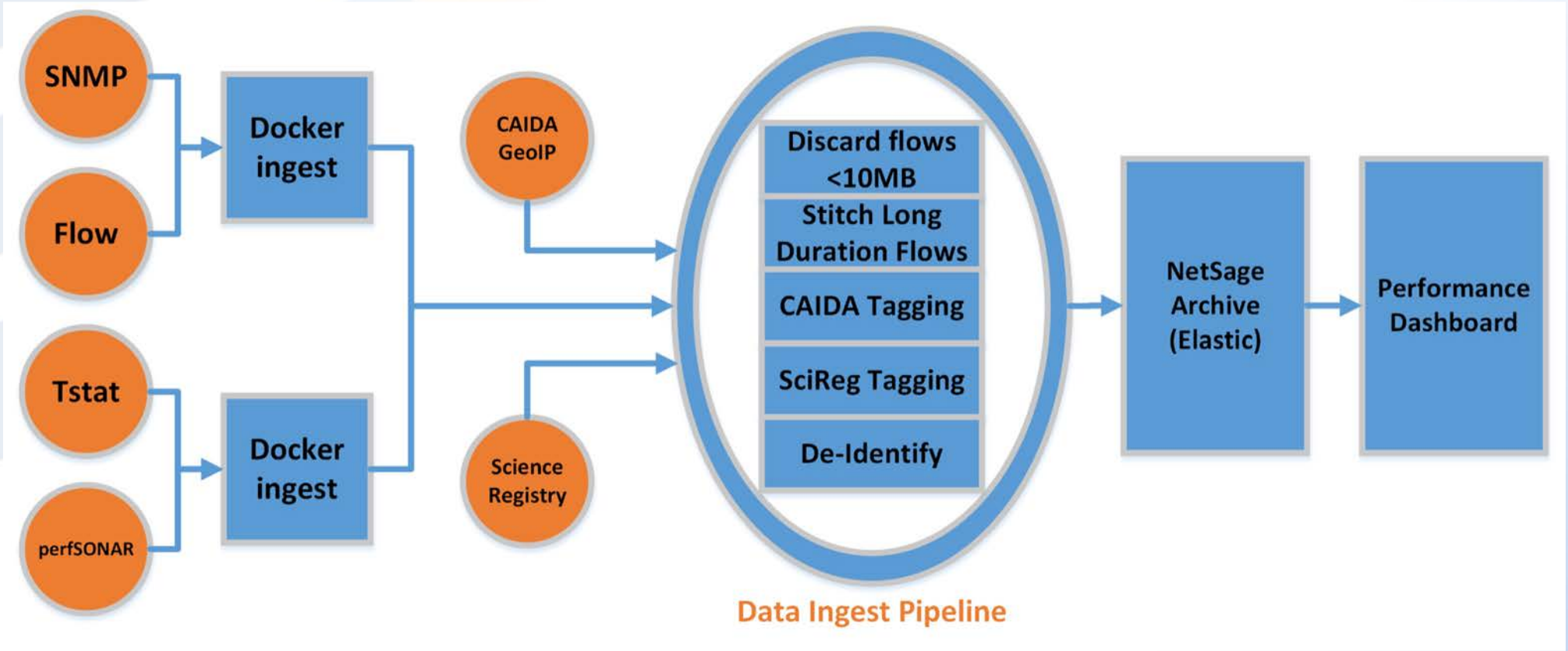
NetSage Focus on Use Cases

- Start with Flow Data Dashboards
  - What are the top sites using my circuits?
  - What are the top sources/destinations for an organization?
  - Who's using my archive?
- Dig deeper with debugging dashboards
  - What are the flows like between these two orgs?
  - There was a performance spike on my circuit – what was it?
  - Who's transferring a lot of data really slowly?

Flow Data collection

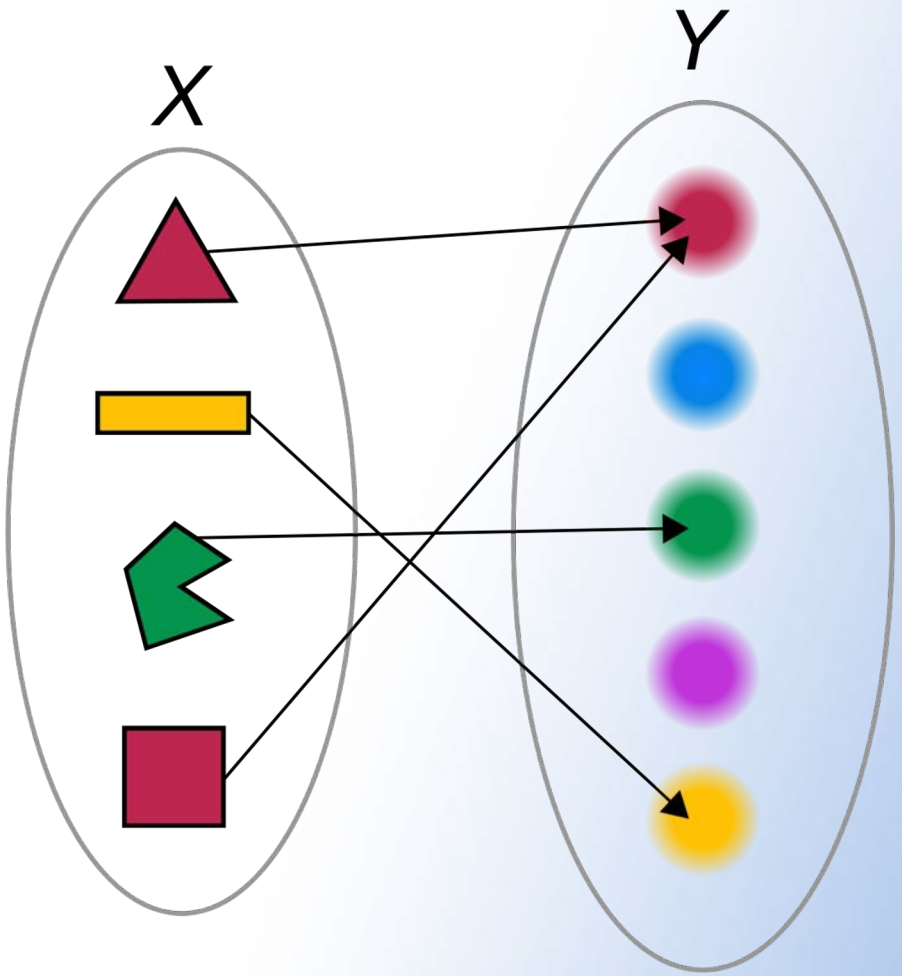
- Flow data is sent to a NetSage Ingest Pipeline, de-identified, and then sent to NetSage archive
- Runs as a set of Docker containers on existing hardware/VM infrastructure
  - Linux distro of your choice underneath
  - Can be anywhere on your network as long as your router can send raw flow data to it
  - Support from TACC if you run into any issues, but you handle general updates

NetSage Ingest



Global Science Registry

- Maps flows to specific resources
  - DTN
  - Instrument
  - Compute
- Contributions come from resource owners
- Currently matched by IP, however packet marking is in the roadmap for development



NetSage Privacy

- NetSage is committed to privacy, and preemptively addressing any security or data sharing concerns
  - No PII collected
  - Remove the last octet from IP address
  - Only keep data on flows over 10MB for circuits
    - 1MB for archives
- Data Privacy Policy
  - https://tinyurl.com/netsage-privacy
- Prototypes are behind a password until we're told to make it public

NetSage Domestic:

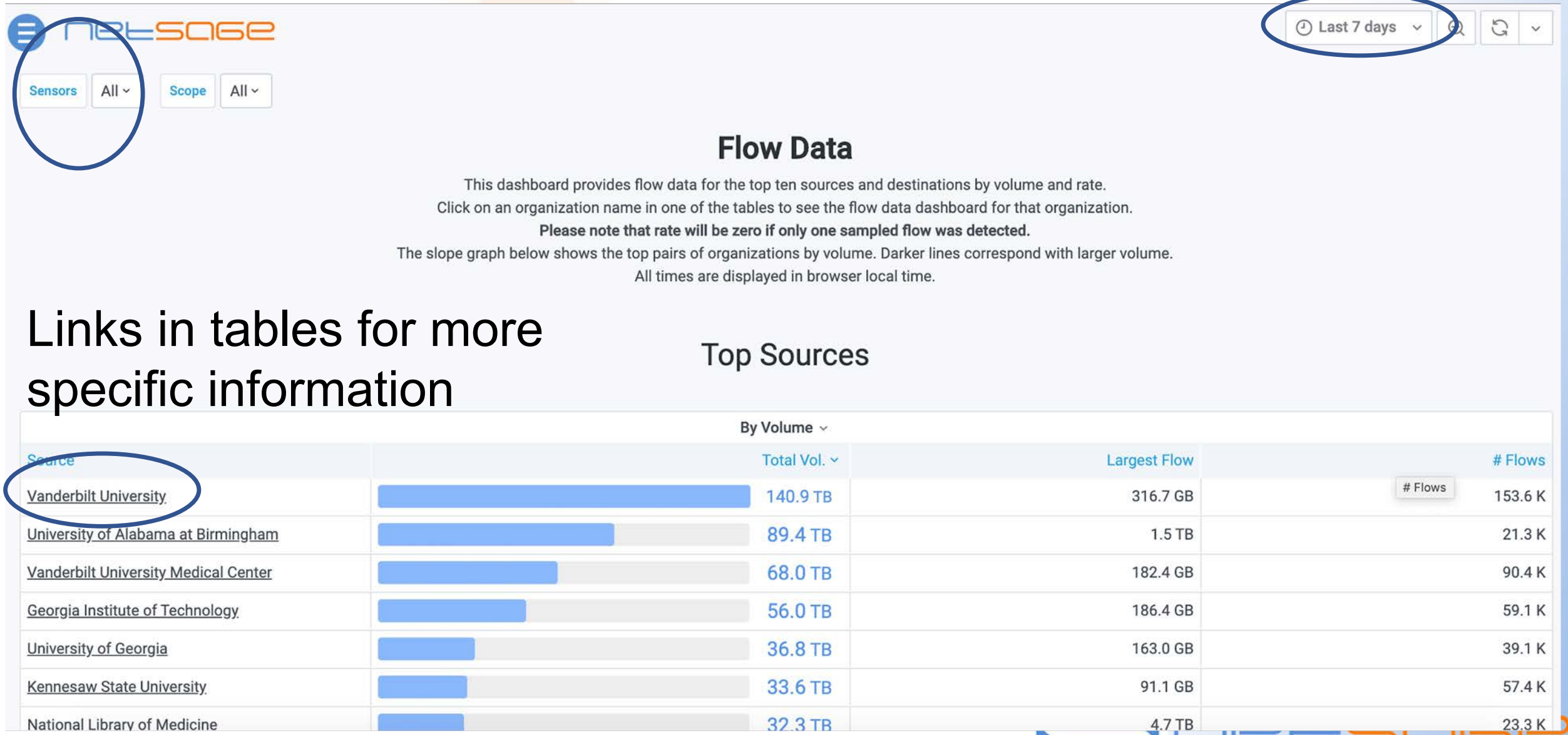
- FRGP
- GPN
- PACWAVE
- SoX
- Sun Corridor
- TACC
- LEARN
- …and now ACCESS-CI!

Netsage for ACCESS-CI

- One set of dashboards for all ACCESS-CI compute resources
  - Currently, TACC and PSC are contributing flow data
  - SDSC, NCSA, and others coming soon
  - Portal available at https://access.netsage.io
- Currently just flow data, however there are plans for future development, including:
  - In-band telemetry
  - Tie in flow data with job scheduling/allocation to see the big picture of flow data in, computation, and flow data out

Pick a sensor

Pick a timeframe



Links in tables for more specific information

What NetSage Does Best

- Answers questions asked by network engineers and network owners
- Human-readable summaries and patterns
- Gives people the higher level pattern so they can narrow down a time frame and then use local tools that have more detail
- Questions? Contact Jen Schopf & Doug Southworth (jms@tacc.utexas.edu & dsouthworth@tacc.utexas.edu)

Using NetSage to Support ACCESS

Jason Zurawski - zurawski@es.net

Lawrence Berkeley National Laboratory / Energy Sciences Network (ESnet)

2023 Internet2 Technology Exchange

September 20th, 2023