# NetSage for LHC update

Doug Southworth

Indiana University

# Today's Discussion

- Current Project Status
- 2.0 Ingest Pipeline
- Science Registry

# NetSage Status

- Original IRNC project is complete on October 28, 2022
- Personnel Changes
  - Jennifer Schopf (PI) → TACC
  - Dan Doyle (Data management) → ESnet
  - Lisa Ensman (Pipeline development) → Retirememement

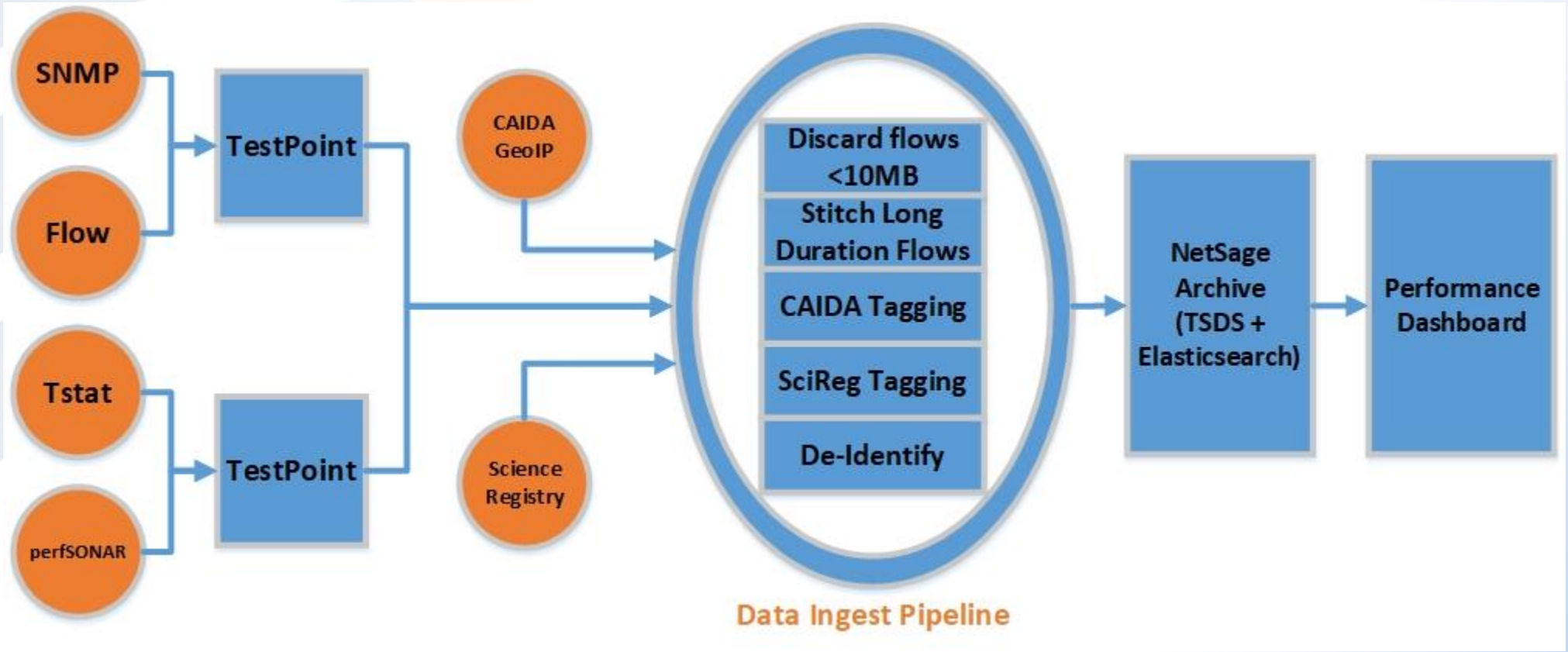
# Is development complete?

- Not at all! NetSage is still in use in several projects, including LHC, EPOC, and upcoming uses at TACC
- Development is currently frozen during the transition, however in the near future active development will resume.

# Pipeline 2.0

- Represents a significant upgrade to the existing ingest pipeline
- Significant development work has been done, along with initial testing, however there is still work to be done

# NetSage Ingest



# Pipeline 2.0 changes

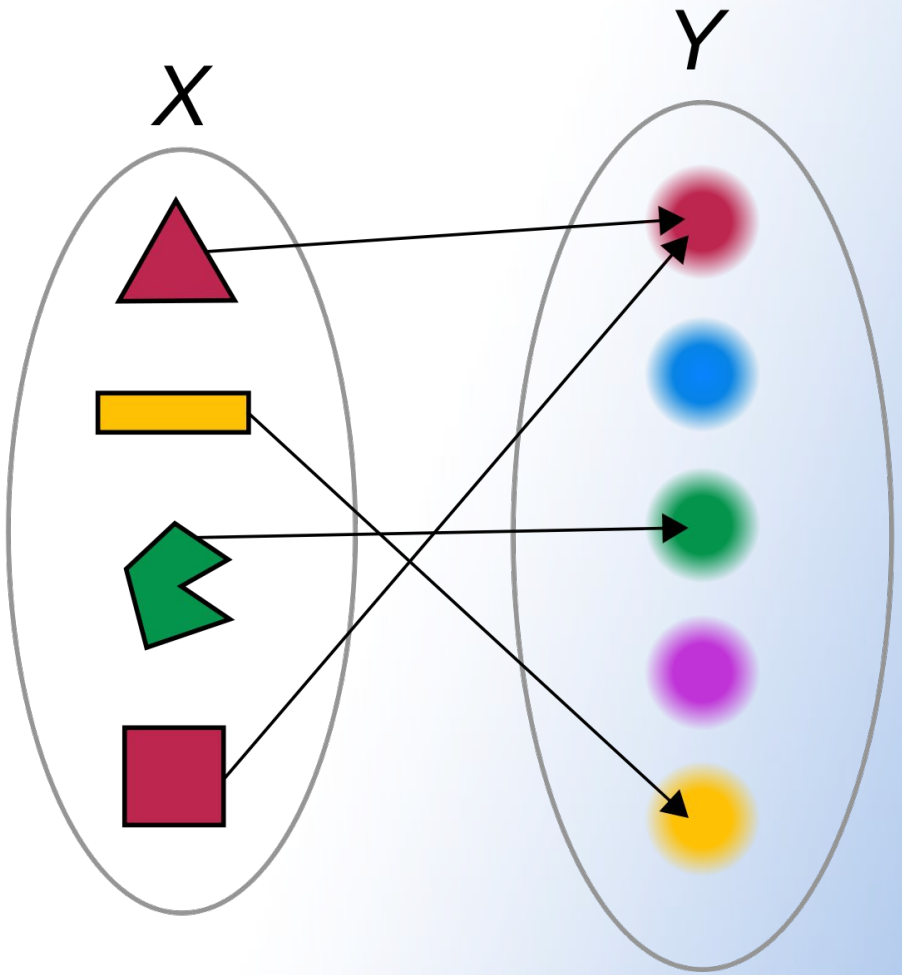
- Switch from nfdump to pmaccount for flow aggregation (mostly complete)
- Separate tagging for IPv4 and IPv6
- Single alias for multiple ifindexes
- Drop IP addresses starting with 0 (system network management, or “this network”)
- Drop ICMP pre-logstash

# Pipeline 2.0 changes

- Labeling of 6to4 relay flows
- RabbitMQ to Kafka
- Several other internal process efficiency and security changes

# 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



# Science Registry

- Product of scope creep
- Now that the actual use case is known, a wholesale rewrite is in order
  - API for external contributions/updates
  - Elimination of unused fields
  - Build for collaboration with major projects

# NetSage for LHC update

Doug Southworth

Indiana University