EPOC Network Analysis with NetSage
Last updated March 22, 2021

ABOUT EPOC

Over the last decade, the scientific community has experienced an unprecedented shift in the way research is performed and how discoveries are made. Highly sophisticated experimental instruments are creating massive datasets for diverse scientific communities and hold the potential for new insights that will have long-lasting impacts on society. However, scientists cannot make effective use of this data if they are unable to move, store, and analyze it.

The Engagement and Performance Operations Center was established in 2018 as a collaborative focal point for operational expertise and analysis and is jointly led by Indiana University (IU) and the Energy Sciences Network (ESnet). EPOC provides researchers with a holistic set of tools and services needed to debug performance issues and enable reliable and robust data transfers. By considering the full end-to-end data movement pipeline, EPOC is uniquely able to support collaborative science, allowing researchers to make the most effective use of shared data, computing, and storage resources to accelerate the discovery process.

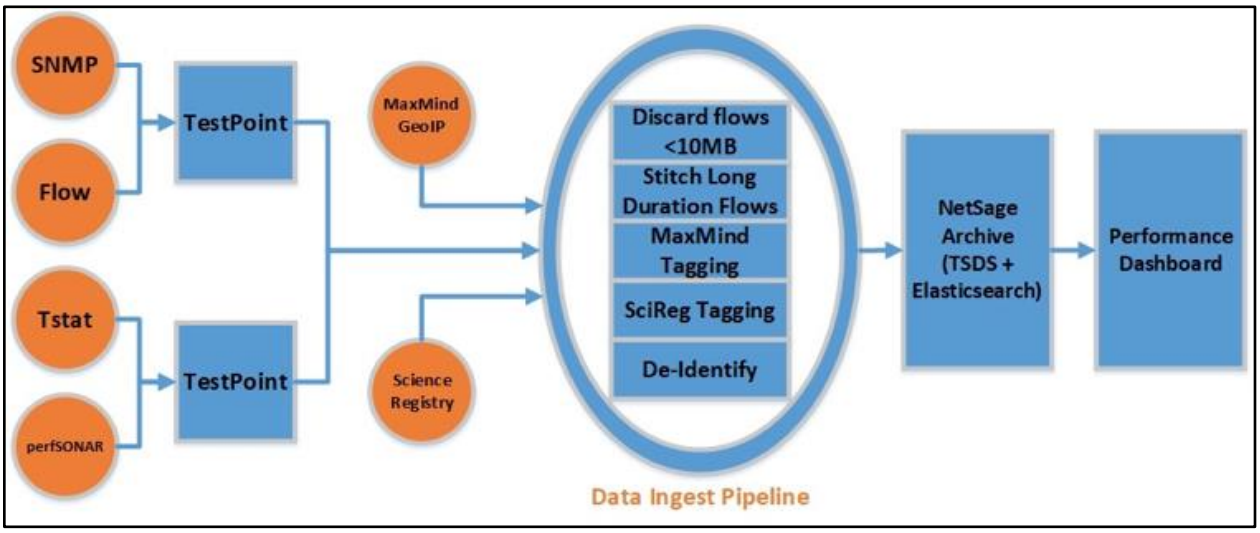
EPOC supports six main activities

- *Roadside Assistance* via a coordinated Operations Center to resolve network performance problems with end-to-end data transfers reactively;
- *Application Deep Dives* to work more closely with application communities to understand full workflows for diverse research teams in order to evaluate bottlenecks and potential capacity issues;
- *Network Analysis* enabled by the NetSage monitoring suite to proactively discover and resolve performance issues;
- The *Data Mobility Exhibition* and associated work with our simplified portal to check transfer times against known “good” end points;
- Provision of *managed services* via support through the IU GlobalNOC and our Regional Network Partners;
- *Coordinated Training* to ensure effective use of network tools and science support.

Network Analysis

Understanding application performance and network measurement are two sides to a single coin - one doesn’t make sense without the other. The EPOC project uses the NetSage tool (http://netsage.global) to collect and evaluate common network measurement data. The initial NetSage software (NSF-1540933) was developed and deployed to understand the use of the NSF-funded international networks. It was meant to work with sparse, international circuits, with circuit owners and operators as the end user. NetSage deployments can collect data from routers, switches, active testing sites, and science data archives, which are common for collaborative research. A NetSage deployment uses a combination of passive and active measurements to provide longitudinal performance visualizations via performance Dashboards. EPOC has expanded the use of this software to more densely defined networks, and supports additional analysis and visualization aspects.

The NetSage Software consists of a set of open source tools that follows a basic monitoring tool architecture. *NetSage TestPoints* are a collection of software and hardware components that gather active and passive data into records that are then sent to the *Data Ingest Pipeline*. The five-step Pipeline filters those records and adds additional tags before de-identifying the data. The records are then stored in the *NetSage Archive*, a centralized storage framework consisting of two different databases, a Time Series Data System (TSDS) archive and an Elasticsearch archive. *Performance Dashboards* built using the open source Grafana analysis and visualization engine, access the records from the NetSage Archive and present visualizations to answer the questions identified by the stakeholders.



Current Deployments

Several EPOC Regional Network Partners are deploying NetSage on their infrastructure. These include:

- The Great Plains Network (GPN) - SNMP and Flow data: http://gpn.netsage.global
- iLight/Indiana GigaPop - Flow data: http://ilight.netsage.global
- Southern Crossroads (SoX) - Flow data: https://sox.netsage.global
- KINBER - Flow data: https://kinber.netsage.global
- CENIC/PacificWave - Flow data: https://pacwave.netsage.global/
- Front Range GigaPop (FRGP) - Flow data: https://frgp.netsage.global
- TACC - Flow data: https://tacc.netsage.global
- Sun Corridor: In progress
- NSF-supported international networks, including TransPAC, NEAAR/NEA^{3}R, AmPath/AmLight, and PIREN
- Plus archive data from NCAR, NERSC, TACC and U. Hawaii – Archive data

NetSage Impact