EPOC Network Analysis
Last updated September 19, 2019

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 five 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;
- 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://portal.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. EPOC has expanded the use of this software to more densely defined networks, and supports additional analysis and visualization aspects.

The NetSage testpoint, which gathers the data used in the system, is a collection of software and hardware elements that will be expanded to include additional passive and active collection

techniques. The testpoints communicate with an archiving framework that makes use of an analysis and visualization engine based on Grafana.

The test points collect data from Simple Network Management Protocol (SNMP), which provides basic information about circuits and is collected by almost all network operators. NetSage uses this data in a simple map display as well as to supply basic bandwidth and uptime information.

The second dataset collected by the testpoint makes use of perfSONAR to collect active measurements for throughput, latency, and loss. This data is displayed as heatmaps, but more importantly, these measurement points are a fundamental tool for the EPOC Roadside Assistance process to be able to evaluate user-level performance between sites.

The third data set includes passive TCP flow statistics that are de-identified before archiving. We will also collect this data on some domain science archives to gather data about specific transfers directly from the source. This allows us to compare real file transfer throughput to perfSONAR throughput and assist in identifying network versus disk performance issues.

The testpoints communicate with an archiving framework that consists of a Time Series Data System archive and the ELK Elastic Stack archive. Datasets can be accessed from these archives directly or through a variety of analytical tools and visualizations that uses the open source tool Grafana, which provides a wide variety of basic analysis and visualization widgets, can interact with multiple backend data sources, and can also provide authenticated access as needed.

Current Deployments
Several EPOC Regional Network Partners are deploying NetSage on their infrastructure. These include:
- The Great Plains Network SNMP Dashboard: http://gpn.netsage.global
- iLight/Indiana GigaPop Flow Data Dashboards: http://ilight.netsage.global