

## EPOC Data Transfer Testing/Data Mobility Exhibition (DME)

Contact Point: Ken Miller, Dr. Jennifer Schopf

Audience: General

Last updated: August 17, 2022



### 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 and Consultations** via a coordinated Operations Center to resolve network performance problems with end-to-end data transfers;
- **Application Deep Dives** to work more closely with application communities and 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;
- **Data Transfer Testing/ Data Mobility Exhibition** to check transfer times against known good end points;
- **Provision of managed services** via support through the IU GlobalNOC and our Network Partners;
- **Coordinated Training** to ensure effective use of network tools and science support.

### DATA TRANSFER TESTING/DATA MOBILITY EXHIBITION

Data mobility is a critical component of the process of science. Being able to predictably and efficiently move scientific data between experimental sources, processing facilities, long term storage, and collaborators are common use cases that transcend the boundaries of research disciplines.

There are several common barriers to high performance transfers of R&E data:

- Configuration: lack of BGP tuning for large data transfers over R&E networks (BGP best path), MTU mismatches, PMTUD black holes
- Expertise gap: Knowledgeable staff that are capable of designing and operating performance-based infrastructure are often in short supply, and the demands of day-to-day operations consume most, if not all, of their time
- Other priorities: The typical designs of campus networks are sufficient for 99% of users, thus the addition of new technology for a single use case may not be a high priority
- Advocacy: Researchers often don't know what to expect regarding data transfer speeds or what resources are available to them, and therefore don't speak up about network issues that hamper scientific research
- Baselineing: Routine data transfer testing is often not scheduled or validated

The Data Mobility Exhibition (DME) is designed to address many of these concerns by giving institutions a way to model the performance of their network infrastructure using real-world data transfer scenarios. The DME testing infrastructure consists of 5 remote test sites. These well tuned sites host datasets ranging from a single file to over 100,000 files and vary in size from 100MB to 5TB. After a test dataset is successfully transferred, upload and download rates are displayed in MB/s. A 1TB/hr (277 MB/s) data transfer rate is the initial baseline goal for two 10Gb/s capable institutions.

The Data Mobility Exhibition uses the following steps for testing data transfer performance:

- Create a brief (1-2 page) description of the network and data architecture for your campus environment using this template.
- Prepare a local data transfer machine by installing and configuring the Globus Connect software. Help can be requested at [epoc@iu.edu](mailto:epoc@iu.edu).
- Request access to the DME group in order to transfer the reference datasets and record the results. Information on the DME group and datasets is available here: <https://www.globusworld.org/tour/data-mobility-exhibition>
- Share test setup information and results on a standardized spreadsheet with the organizers. These results are recorded as MB/s rates.

For those interested in improving data transfer performance, personalized assistance from the Engagement and Performance Operations Center (EPOC) is available: [epoc@iu.edu](mailto:epoc@iu.edu)