Lawrence Berkeley National Laboratory
LBL Publications

Title

The Engagement and Performance Operations Center: EPOC

Permalink

Authors

Moynihan, Edward
Schopf, Jennifer
Zurawski, Jason

Publication Date

DOI

The Engagement and Performance Operations Center: EPOC

Edward Moynihan, Indiana University
Jennifer Schopf, Indiana University
Jason Zurawaski, ESnet

https://epoc.global/

EPOC — Engagement and Performance Operations Center

INDIANA UNIVERSITY

ESnet — ENERGY SCIENCES NETWORK

OVERVIEW

When researchers have better access to their data, their time to science is reduced, they can ask more complicated questions, and can work in larger teams to solve more complicated problems. The Engagement and Performance Operations Center (EPOC) helps researchers routinely and reliably transfer large datasets faster, thereby improving data access and collaboration. Through our targeted partnerships, EPOC has the potential to benefit nearly all of U.S. science, research and education by working across organizational boundaries to address performance issues that can hinder cooperative research.

EPOC works with researchers, engineers, and cyberinfrastructure planners to help make sense of performance slowdowns that occur during the end-to-end data transfer process, both in the moment and when long term planning. We help under-resourced sites take advantage of modern cyberinfrastructure approaches so that all researchers can better collaborate.

ROADSIDE ASSISTANCE

A key aspect of EPOC is the process pipeline for immediate help, referred to as "Roadside Assistance". A Roadside Assistance case is submitted by a scientist or other end user who is having difficulties transferring or receiving files from a collaborator at another site. A Case Manager and Primary Engineer are assigned, and basic tracking infrastructure is set up. Troubleshooting of these cases involves understanding the end-to-end path of the data transfer and then using public testing and validation services (such as perfSONAR nodes and router proxies) to identify and address potential issues. The Case Manager will discuss these issues with the customer and the relevant support staff and describe possible next steps with help from EPOC. This may include EPOC-led Deep Dives, managed service offerings, or larger scale interactions with other organizations.
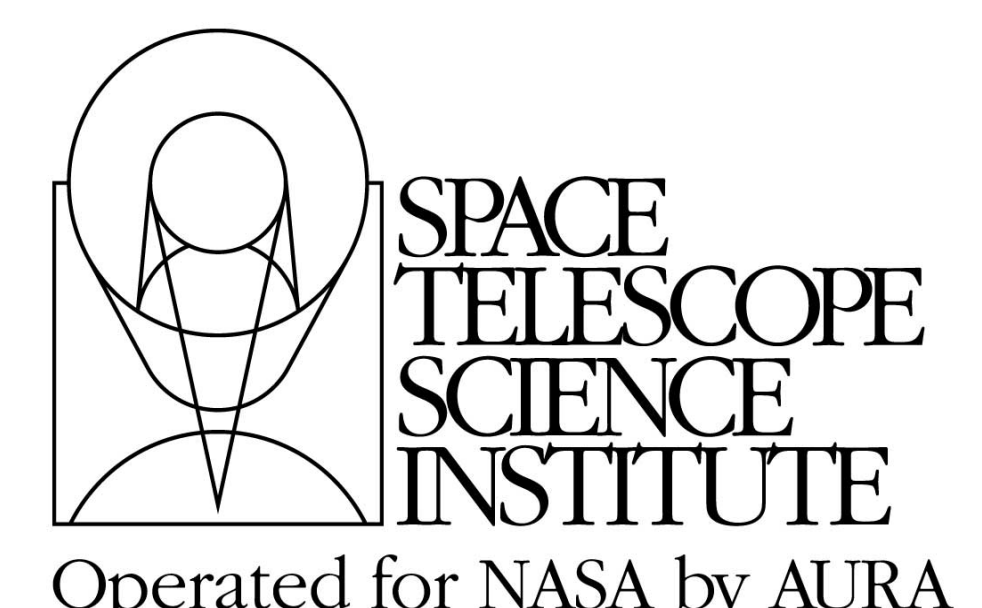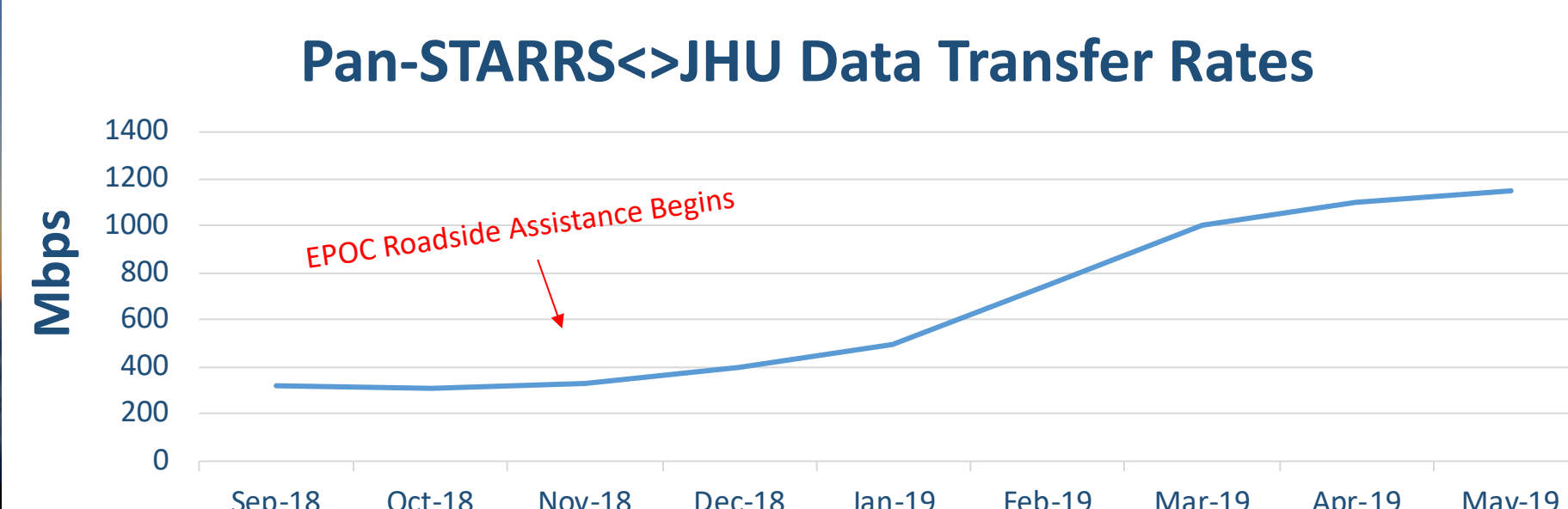
SCIENCE DEEP DIVES

Science Deep Dives aim to understand the full science pipeline for research teams and suggest alternative approaches for the scientists, local IT support, and national networking partners as relevant to achieve the long-term research goals. The Deep Dive approach is based on an almost 10-year practice used by ESnet to understand the growth requirements of DOE facilities . We have adapted their approach for work with individual science groups. At the end of the engagement with the research team, EPOC staff can come away with a good understanding of the research, data movement, who's using what pieces, dependencies, and time frames. This enables us to identify possible bottlenecks or areas that may not scale in the coming years, and to point the research team towards existing resources so that they can reach their goals more effectively.

SCIENCE SUPPORT USE CASE

The Panoramic Survey Telescope and Rapid Response System (Pan-STARRS) is a system for wide-field astronomical imaging, developed and operated by the Institute for Astronomy at the University of Hawaii (UH), that is enabling researchers to more accurately estimate galaxy redshifts and better predict cosmic expansion. Data produced by Pan-STARRS is archived and made accessible to researchers at the Space Telescope Science Institute at John Hopkins University (JHU) in Baltimore, Maryland. Pan-STARRS relies on high-speed research networks to transfer nearly 100TB of data per year between UH and the data archives at JHU. Despite an expected 100Gbps end-to-end network path, networking engineers at UH noticed that data transfers were only reaching maximum speeds of 320Mbps. In November 2018, the UH engineers reached out to EPOC to take advantage of our Roadside Assistance service. Over the next 3 months, the EPOC team worked closely with UH, JHU, and other networking partners to identify and resolve a number of issues that were preventing Pan-STARRS from getting optimal network performance, including:

Identified Networking Issue	EPOC Resolution
• Top of the Rack switches in the UH data center were underpowered	• Critical data servers moved away from this set up
• Misconfigured access control lists and firewalls in the Science DMZ	• Equipment layout redesigned so that data transfer nodes were not behind the firewall
• Default routing between UH and JHU hosts was suboptimal	• UH Peering moved to a 100G link in Los Angeles
• Path between JHU and the Mid-Atlantic Crossroads (MAX) GigaPop was only 1Gbps	• Network was upgraded to 10Gbps from the end data receiver host to the MAX GigaPop
• Maximum Transmission Unit (MTU) setting on several jumbo frame routers was set too low	• Increased MTU setting to 9000-byte size frames
• TCP Buffer settings for the end hosts were misconfigured for large scale data transfers	• Settings updated to the ESnet recommended settings
• Transfers were delayed by a per file DNS lookup piece of the Pan-STARRS file transfer mechanism	• Configuration of the proxy within the Pan-STARRS software framework was updated
• Pan-STARRS data was spread out across 160 discrete logical storage volumes on 32 hosts	• Recommendations for a new, unified storage system

Pan-STARRS<>JHU Data Transfer Rates
(Mbps, Sep-18 through May-19, EPOC Roadside Assistance Begins)

SPACE TELESCOPE SCIENCE INSTITUTE — Operated for NASA by AURA

The EPOC team re-evaluated performance once theses changes had been made and were able to achieve a 3x increase in overall performance; from 320Mpbs to a sustained 1Gbps transfer rate

ARE YOU EXPERIENCING SCIENCE AND RESEARCH DATA TRANSFER PROBLEMS?

IF SO, EPOC CAN HELP:

1. ROADSIDE ASSISTANCE
via a coordinated operations center to resolve network performance problems with end-to-end data transfers.
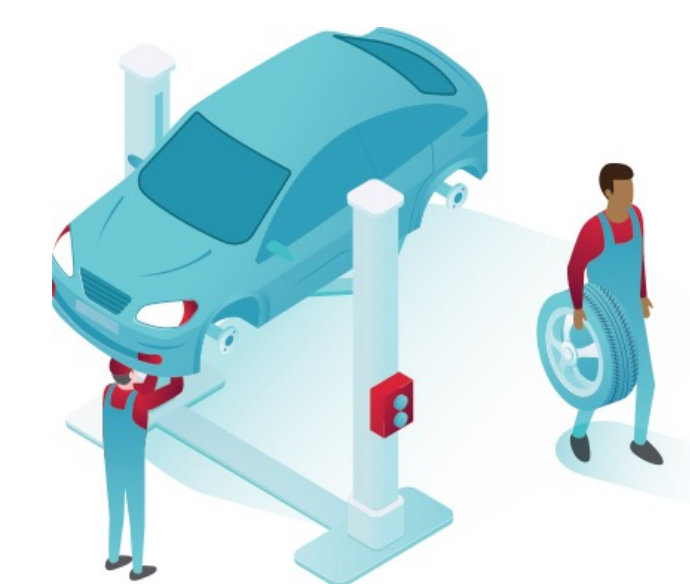
2. APPLICATION DEEP DIVES
to understand full science workflows and to evaluate bottlenecks and potential capacity issues

3. NETWORK ANALYSIS
enabled by the NetSage monitoring suite to proactively discover and resolve performance issues.

4. MANAGED DATA SERVICES
(aka service in a box) enabling the use of advanced technology components supported by IU GlobalNOC and our Regional Network Partners

5. COORDINATED TRAINING
to ensure effective use of network tools and science support.

Contact us: epoc@iu.edu; https://epoc.global