



EPOC

Engagement and Performance
Operations Center

perfSONAR and the Engagement Performance Operations Center

Scott Chevalier

Indiana University International Networks



ESnet

ENERGY SCIENCES NETWORK



INDIANA UNIVERSITY

Engagement and Performance Operations Center

- Joint project between Indiana University and ESnet
 - co-PI Jent (IU GlobalNOC) and Zurawski (ESnet)
- Part of CC* program for domestic science support
 - Program Officer: Kevin Thompson
- Award #1826994, \$3.5M over 3 years
- Partnerships with regional, infrastructure, and science communities that span the NSF and DOE continuum of funding

Why an Engagement Operations Center?

- Today's science is collaborative science
- Collaborative science
 - Multiple partners
 - Multiple data sets
 - Many points of connection
 - Cross agency cooperation
- With better access to data we ask harder questions
- Interactive data sources change the science we do

Why should I care who's using my network?

- Furthering research is part of our mission
 - We're not building infrastructure for infrastructure's sake
- Researchers struggle to use advanced services
 - Don't know what to expect
 - Often don't understand the value of CI services
- Knowing user base is critical to helping improve End to End performance
 - Proactive effort is needed

Return on Investment Means More Investment

- If you can't measure it you can't improve it
- If you want to grow, you need to justify it
 - Understanding workflows leads to increasing utilization
 - Utilization = Justification.... = Funding!
- Happy researchers means building critical relationships with stakeholders

- Measure before- Do the work- Measure After

In the Past

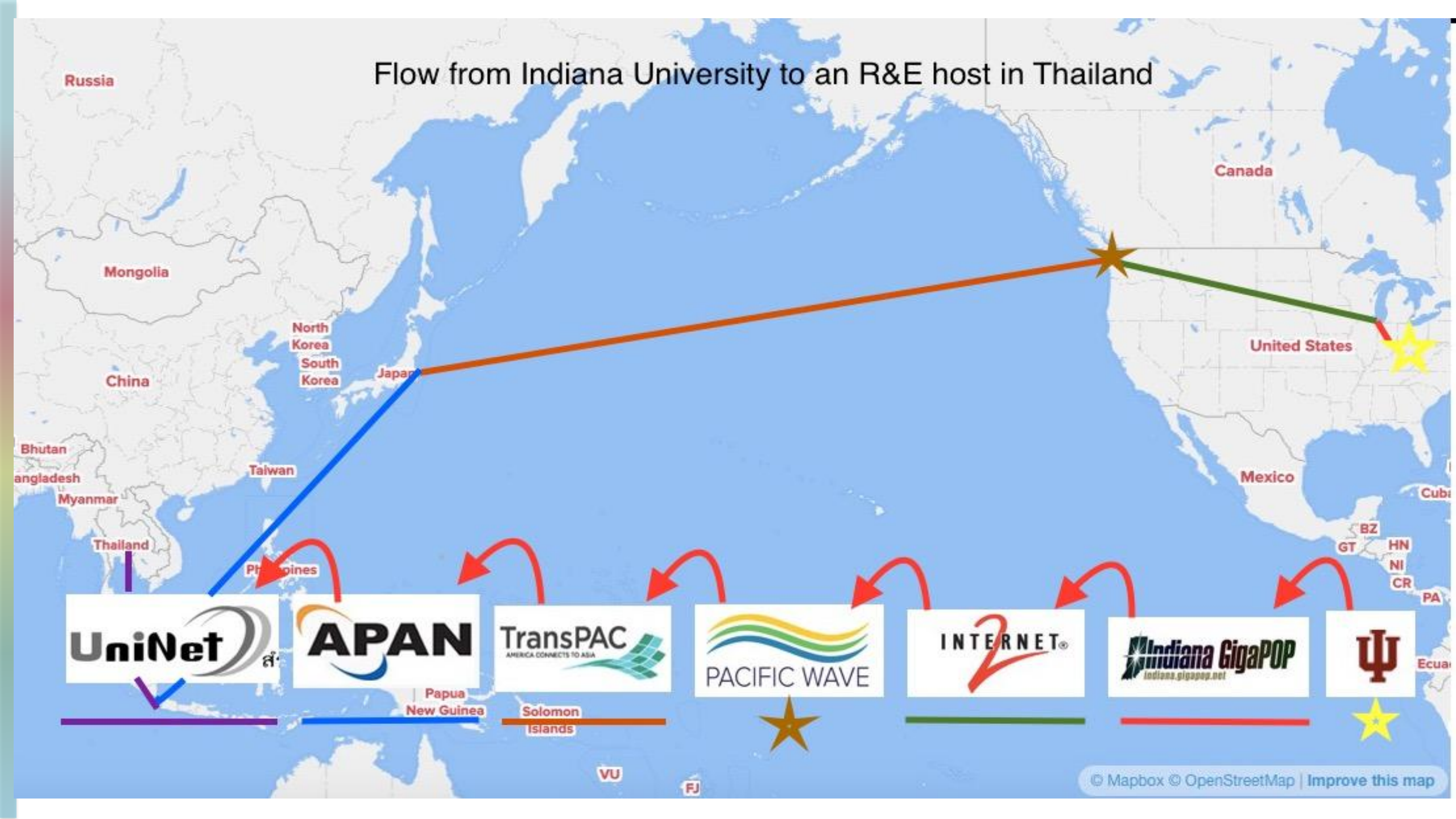
- Infrastructure often built without consulting researchers
 - Capability mismatch
 - Distrust between researchers and IT
 - Research groups build their own solutions
- Which leads to IT eventually supporting many approaches, many of them VERY poorly thought out!

Our goal: Cyberinfrastructure to advance Innovative Research

Understanding End-to-End Performance is Hard

- Lots of pieces - Host system through networks to host system
- No one controls all the pieces
- Unknown expectations for what performance should be
- Soft failures are hard to find
- Many, many points of coordination

Flow from Indiana University to an R&E host in Thailand



EPOC Five Main Focus Areas

1. Roadside Assistance and Consulting
2. Application Deep Dives
3. Network Analysis (NetSage - Andy talked about this Earlier)
4. Services “in a box” (DMZ, testpoint in a box, pS, etc)
5. Training

Roadside Assistance Process

- “This file transfer worked last week, but it doesn’t anymore?”
 - Think of this like a flat tire, crash repair
 - Anyone can submit
- Contact epoc@iu.edu
 - Within 24 hours, gets triaged
 - Some initial investigation to verify the issues
 - A Case Manager and Lead Engineer are assigned
 - Shareable infrastructure set up

Roadside Assistance - Consulting

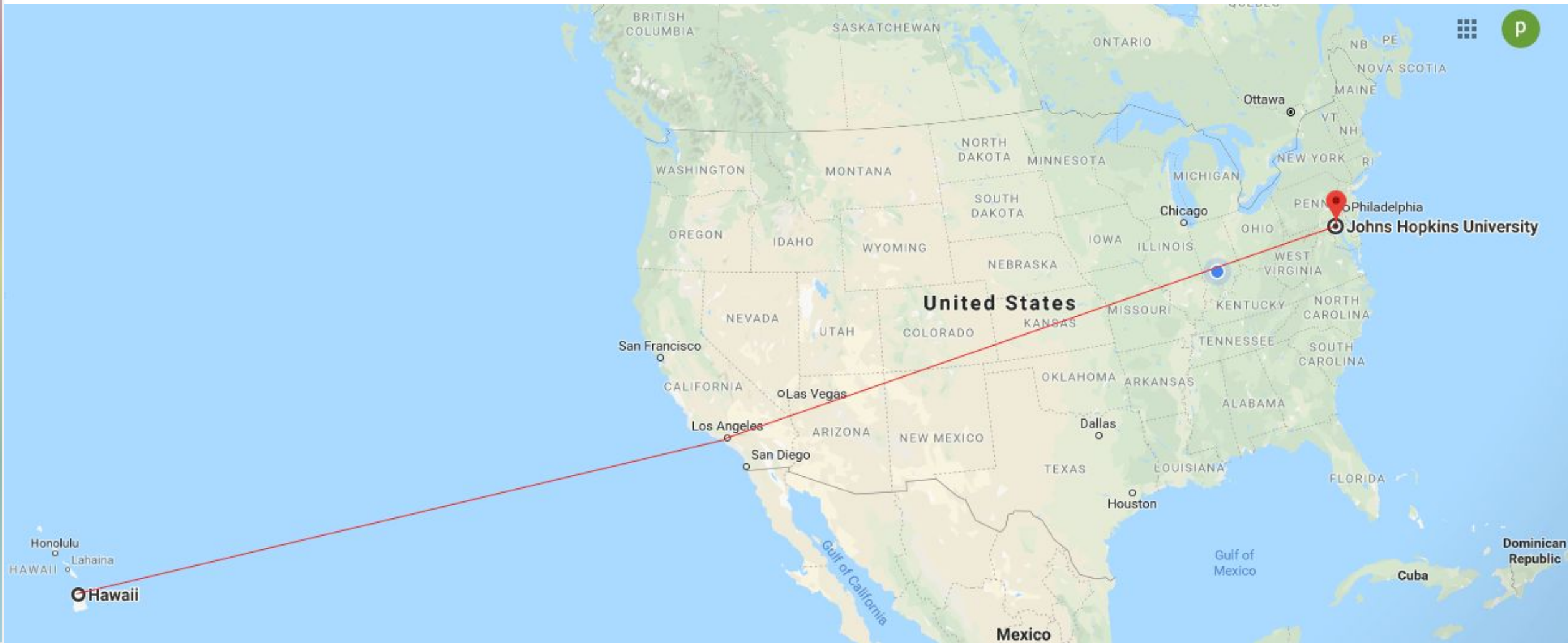
- Lighter weight than a full Roadside Assistance Case
- Submission process same – contact EPOC@iu.edu
 - Suggestions for DTNs, DMZs, firewalls and DMZs,
 - Data projections for science fields
 - Expected (real) performance between two sites
 - Advice on how to conduct a performance assessment of a network and applications
 - Or others!
- Similar operations approach
- Results/suggestions will be added to fasterdata.es.net over time

PanSTARRS Poor Performance



- Panoramic Survey Telescope and Rapid Response System ([Pan-STARRS](#))
 - Collects and shares wide-field astronomy data
 - Used for estimating galaxy redshifts, research in local cosmic expansion and dark energy
- Regular 100TB data transfers
 - Institute for Astronomy at University Hawaii (UH)
 - Space Telescope Science Institute at Johns Hopkins University

PanSTARRS Poor Performance



PanSTARRS Poor Performance



- 100TB transfer - experienced only 320Mbps speed
- End-to-end path believed to be 10 or 100G
 - Expected multi Gbps at least
- Involved engineers from
 - International Networks at Indiana University (IN@IU)
 - IRNC NOC Performance Engagement Team (PET)
 - ESNNet
 - MidAtlantic Crossroads (MAX)
 - Internet2 NOC

PanSTARRS: Problem Identification 1 (The Usual Suspects)



- perfSONAR testing identified JHU did **not** have a 10G connection through MAX to Internet2
 - Campus network upgraded to MAX/Internet2
- perfSONAR testing identified default UH to CONUS route was 10G
 - Updates default route to PIREN 100G Hawaii to LA

PanSTARRS: Problem Identification 2 (The Usual Suspects)



- Maximum Transmission Unit (MTU) setting on several routers was less than 9000 byte size frames (Jumbo Frames)
 - Larger MTU settings make data transmissions more efficient because the CPUs on switches and routers can process a larger payload for each frame
 - Only works if each link in the network path -- including servers and endpoints -- is configured to enable jumbo frames at the same MTU

PanSTARRS: Problem Identification 3 (The Usual Suspects)



- TCP Buffer settings on end hosts were misconfigured
 - ESnet recommended settings, available at:
<http://fasterdata.es.net/host-tuning/background/>

PanSTARRS: Problem Identification 4 (The Less Usual Suspects)



- At UH, underpowered Top of Rack (TOR) switch bottleneck, misconfigured access control lists, and misconfigured firewalls
 - Equipment placement redesigned to remove bottlenecks from path

PanSTARRS Outcome



- Transmission rates went from 320 Mbps to 1Gbps sustained
- Several additional architectural and software issues were identified, which are now part of the project's longer-term upgrade path

<https://epoc.global/wp-content/uploads/2019/04/Roadside-Assistance-case-PanSTARRS-transfers-Hawaii-to-JHU.pdf>

LHC Data Movement Pakistan - UK



- High Energy Physics (Large Hadron Collider) data sharing
 - Tier 1 sites - large, regional sites (all or most of the data)
 - Tier 2 sites - smaller country-level sites (data subset)
 - Local universities and researchers (data for research)

LHC Data Movement Pakistan - UK (2)



- Pakistani Tier 2 site
 - National Center for Physics (NCP), Quaid-i-Azam University Campus in Islamabad, Pakistan
 - 1G connection to Pakistan national network (PERN)
- Regional Tier 1 site
 - Queen Mary University, London, UK

LHC Data Movement Pakistan - UK (3)



- Transfer rates NCP-QM as low as 40 Mbps
 - NCP-Australia Tier 1: 500 Mbps transfers
 - NCP-ESnet Tier 1: 280 Mbps transfers
- Additional intermittent performance problems over previous 2 years

LHC Pakistan-UK Problem ID (1)



- A traffic shaping misconfiguration on the NCP connection to PERN
 - Limited R&E traffic to 50Mbps
 - PERN removed traffic shaping for R&E traffic
- Top of rack switch bottleneck between NCP's file transfer node and edge router
 - Moved file transfer node to the edge router, performance increased from 40Mbps to 100Mbps or better

LHC Pakistan-UK Problem ID (2)



- Small amounts of ongoing, intermittent packet loss within the Quaid-i-Azam University campus network
 - Identified by perfSONAR, cause unclear
 - Moving data node closer to the edge of their network to alleviated the issue
 - Work continues to identify source of loss



LHC Pakistan-UK Problem ID (3)

- Packet loss identified inside the PERN regional network
 - Specific cause of the loss still unclear, work ongoing
- Additional bottlenecks between PERN and TEIN (Asian) networks
 - 1Gbps between national and regional network
 - Congestion is common, therefore so is packet loss
 - Upgrade to 10Gbps being explored
 - Temporary use of commercial path being explored

LHC Pakistan-UK Outcomes



- Original Data transfer NCP to Queen Mary: 40 Mbps
- After engagement transfer speed: ~480 Mbps
- Additional areas for performance improvements identified
 - Larger scale and longer term changes to infrastructure needed
 - Discussions ongoing
- <https://epoc.global/wp-content/uploads/2019/04/Roadside-Assistance-Case-LHC-Transfers-Pakistan-to-London.pdf>

Some Other Common Problems

- Social, not technical
 - Coordination between teams
 - Lack of response somewhere on the (people) path
 - Very few want to know the problem is theirs
- Problems are often not network issues
 - Most of the issues are in systems and software
 - End-to-end path **INCLUDES** wherever the scientist is sitting

EPOC Deep Dives

- Think of this as regular maintenance, oil change, or planning to buy a new car
- Based on seminal work by ESnet to develop Scientific Case Studies
 - Walk through science workflow with the actual scientists
 - Way to understand needs and planning
- Often identifies issues that have **nothing** to do with networks, and everything to do with sociology

We Walk Through Scientific Components...

1. Background information
 - Brief overview of the facility, nature of the science being performed
2. Collaborators
 - Identify people and institutions that a science group interacts with
3. Instrumentation
 - Local and remote scientific instruments and facilities.
4. Process of Science
 - Explain 'a day in the life' of the science group
 - Should tie together the instruments, people, and resources

And Also More Technical Aspects...

5. Software Infrastructure

6. Network and Data Architecture

7. Cloud Services

8. Outstanding Issues and Pain Points

- Local and regional IT staff are critical to these parts, and help form valuable partnerships that may not exist, or could use strengthening

Need for Network Instrumentation

- Performance and measurement are 2 sides of a coin
- Common basic measurement data is the first step to understanding performance issues
 - E.g. Global perfSONAR Deployment, <http://my.es.net>

What is a “Service-in-a-Box”?

- Basic idea:
 - Only large facilities with dedicated funding can afford the time/effort to design/install/operate/maintain a dedicated science infrastructure
 - Ameliorate the costs of design/install at a higher level (e.g. regional network).
 - Create infrastructure that can be delivered as a service
 - Operation can be local or regional (offer flexibility based on the environment and resources available)
 - ***Develop a business model that facilitates cost recovery and upgrade schedules***

Anticipated Offerings

- perfSONAR
- Science DMZ
 - Deployment of regional hardware to support campus high-performance needs
- Data Transfer Hardware/Software
 - Rental or co-location of capable hardware and storage
- Network Capacity Testing
 - Use of 10G/40G/100G/(400G?) hardware to prove out new circuits, or debug old ones
- Intrusion Detection System

Training

- Follow on to OIN (<http://oinworkshop.com>) series that reached over 750 people in the NSF/DOE funding space during the 3 year operational period
- Hands on perfSONAR sessions
 - Especially for small nodes, includes file transfer tests
- How to do an Application Deep Dive
 - Also known as “How to talk to Scientists”
- DMZ/DTN Set Up
- To request – send mail to epoc@iu.edu
 - include “Training Request: in the subject line

EPOC

- Everything thing we do has the goal of helping researchers do their research easier, faster more
 - Faster/better access to data changes the research questions you ask
 - Faster/better access to data speeds up time to results, lets researchers try more things
- We spend a LOT of time talking to researchers and thinking about what they might need from us

Take Aways

- perfSONAR is foundational to what we do
- EPOC provides a single point of contact to help with end-to-end performance issues: epoc@iu.edu
- More about EPOC:
 - <http://epoc.global>