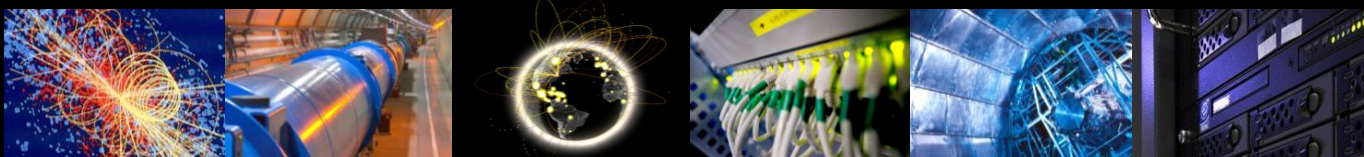# WLCG Network Monitoring and Analytics Updates

Shawn McKee / U Michigan, Marian Babik / CERN
Petya Vasileva / U Michigan, Ilija Vukotic / U Chicago
on behalf of WLCG Network Throughput WG

2024

Spring 2024 HEPiX, Paris, France

https://indico.cern.ch/event/1377701/timetable/#20240417

WLCG
Worldwide LHC Computing Grid

Open Science Grid

Outline

- Infrastructure Status and Updates

- DC24 Review

- Network Analytics
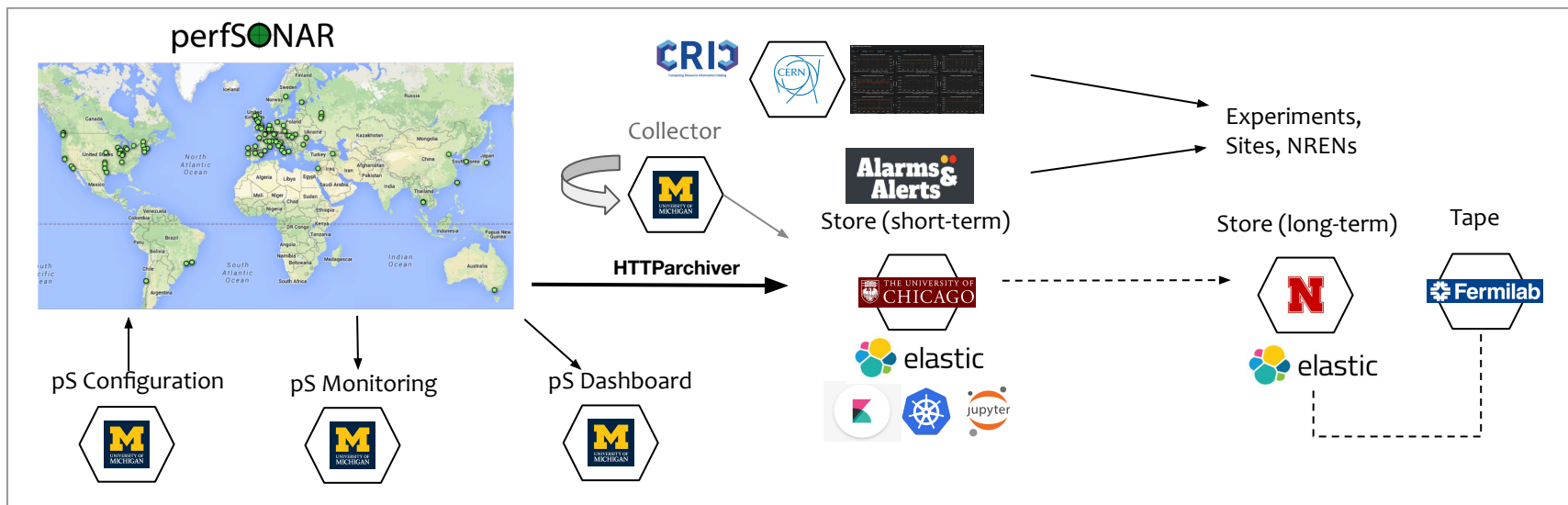
perfSONAR News

- perfSONAR 5.0.8 is the latest release
  - Number of bug-fixes since 5.0
  - Weekly meetings with the developers
  - Update campaign in WLCG
    - Various issues, mostly archiving, but also e.g. legacy limits configuration ([fix](#))
  - Toolkit support for latest CC7 and Alma/Rocky 8 and 9 compatible systems (Alma).
    - **Sites should plan to update by June (end-of-life for CentOS7)**
- **5.1 Beta Release**
  - New Grafana interface - **replacing** toolkit and maddash graphs
  - Threaded iperf3 support
  - Enhanced instrumentation, better troubleshooting of archiving issues
  - **OS support:** Alma/Rocky 9, Debian 11/12, Ubuntu 20/22 (updated docker), No support for CentOS 7, Debian 10, Ubuntu 18

Network Measurement Platform Status

- **Our platform collects, stores, configures and transports all network metrics**
- Evolution based on the perfSONAR 5 already partially implemented.
  - Now directly publishing results from perfSONARs to ES@UC
  - Collector used **only as a fallback**;
  - WLCG CRIC now used for topology

Network Measurement Platform Plans

- Evolution based on the perfSONAR 5 already partially implemented.
  - Forwarding to UNL and backup to FNAL still to be implemented
  - **pS Monitoring** - update to latest Checkmk and enable SSO authentication
  - **ps Dashboard** - integrate with Analytics Platform/Grafana (retire maddash)
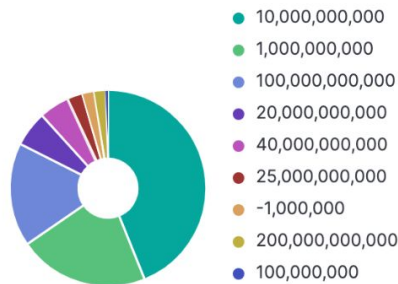 - **ps Configuration** - clarify development roadmap and support
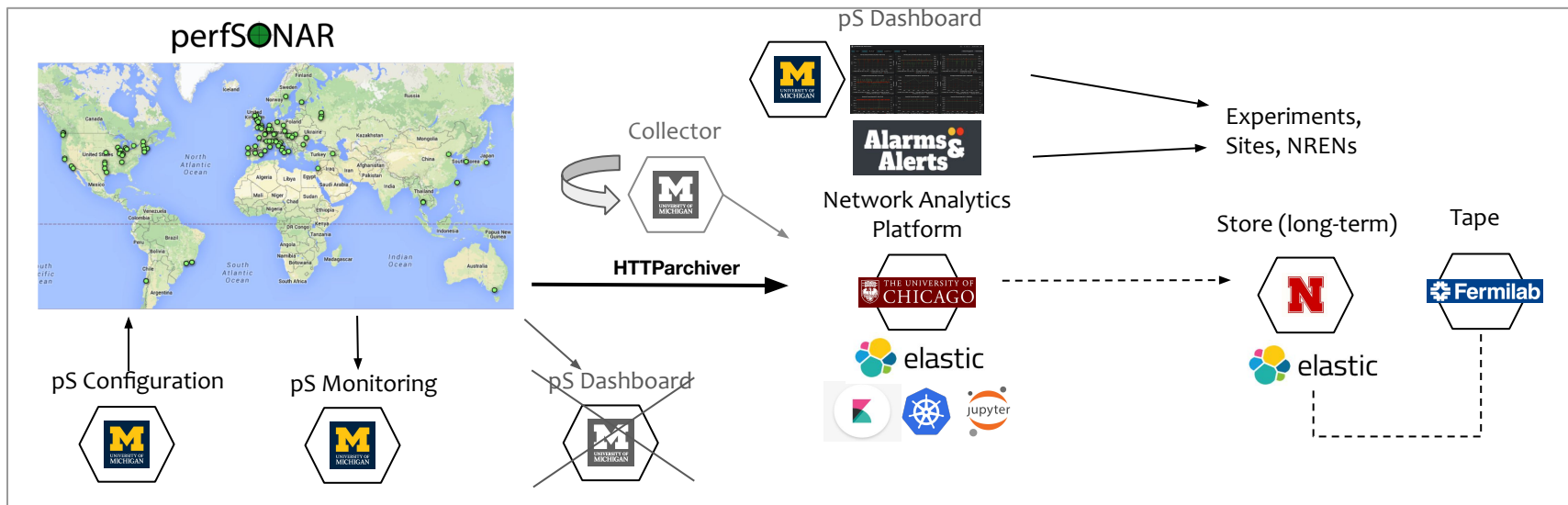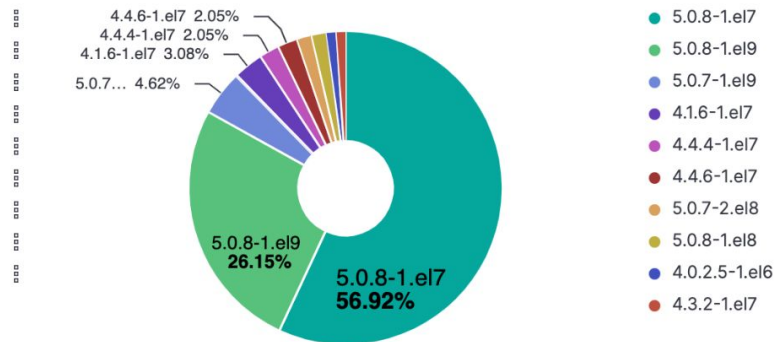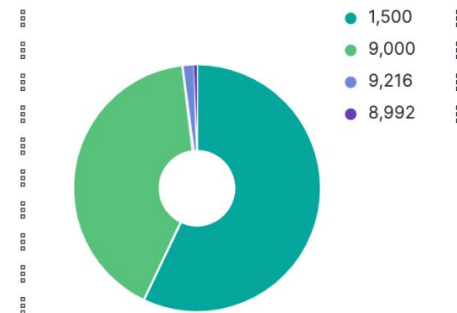
perfSONAR Infrastructure

- Core deployments are still on 10Gbps, we have about 17% with 100Gbps
  - For WLCG/OSG testing purposes 10Gbps is still sufficient
  - Important to **refresh HW** along with the update to 5.1
- Most of the infrastructure is on 5.0.8, but significant fraction still on CentOS 7
- MTU - around 40% on jumbo frames (9000), rest is on standard frames (1500)
- We have small testbed with about 10 perfSONARs with BBRv3 enabled
  - Enabled testing TCP congestion algorithm benefits and jumbo frame trade-offs
  - Open for participation

DC24

WLCG DC24

WLCG Data Challenge 2024 took place in Feb 2024; targeting 25% of HL-LHC

Our DC24 plans included the following:

- Update and utilize **perfSONAR** to clean up links and fix problems before DC24.
- Instrument and document **site networks**, for at least our largest sites.
- **Network planning**: we need to make sure our sites and their local and regional networks are aware of our requirements and timeline and are planning appropriately
- **IPv6** should be enabled everywhere not just because of packet marking, but because it will allow us to get back to a single stack sooner!

psDash Network Status

Network Status dashboard - part of Network Analytics platform - shows network performance based on perfSONAR measurements. Status (ok/warning/critical/unknown) aggregates network and infrastructure metrics.

Site Network Utilisation

WLCG Site Network Input/Output

	Mean	Max
IN: praguelcg2 - CZ-praguelcg	Mean: 27.4 Gb/s	Max: 75.9 Gb/s
IN: pic - ES-PIC	Mean: 41.4 Gb/s	Max: 136 Gb/s
IN: ifae - ES-IFAE	Mean: 41.4 Gb/s	Max: 136 Gb/s
IN: WEIZMANN-LCG2 - IL-WEIZMANN	Mean: 2.02 Gb/s	Max: 9.39 Gb/s
IN: USCMS-FNAL-WC1 - US-CMS	Mean: 84.4 Gb/s	Max: 249 Gb/s
IN: UKI-SOUTHGRID-RALPP - UKI-SOUTHGRID-RALPP	Mean: 14.0 Gb/s	Max: 31.1 Gb/s
IN: UKI-SCOTGRID-GLASGOW - UKI-SCOTGRID-GLASGOW	Mean: 6.58 Gb/s	Max: 14.8 Gb/s
IN: UKI-NORTHGRID-MAN-HEP - UKI-NORTHGRID-MAN-HEP	Mean: 13.7 Gb/s	Max: 33.2 Gb/s
IN: UKI-NORTHGRID-LANCS-HEP - UKI-NORTHGRID-LANCS-HEP	Mean: 12.3 Gb/s	Max: 24.5 Gb/s
IN: UKI-LT2-RHUL - UKI-LT2-RHUL	Mean: 1.78 Gb/s	Max: 6.04 Gb/s
IN: UFlorida-HPC - US-UFlorida-HPC	Mean: 3.10 Gb/s	Max: 23.8 Gb/s
IN: UCSDT2 - UCSD	Mean: 9.84 Gb/s	Max: 28.2 Gb/s
IN: UAM-LCG2 - ES-UAM	Mean: 4.06 Gb/s	Max: 9.50 Gb/s
IN: TRIUMF-LCG2 - CA-TRIUMF	Mean: 25.4 Gb/s	Max: 67.6 Gb/s
IN: TOKYO-LCG2 - JP-University-Tokyo	Mean: 32.6 Gb/s	Max: 72.4 Gb/s
IN: TECHNION-HEP - TECHNION-HEP	Mean: 3.99 Gb/s	Max: 14.0 Gb/s
IN: SPRACE - BR-SPRACE	Mean: 5.34 Gb/s	Max: 26.7 Gb/s
IN: RWTH-Aachen - DE-RWTH-Aachen	Mean: 21.2 Gb/s	Max: 105 Gb/s
IN: RO-13-ISS - RO-13-ISS	Mean: 7.57 Gb/s	Max: 18.8 Gb/s
IN: RO-07-NIPNE - RO-NIPNE_1	Mean: 13.7 Gb/s	Max: 28.5 Gb/s
IN: RAL-LCG2 - UK-RAL	Mean: 87.8 Gb/s	Max: 201 Gb/s
IN: Purdue - US-Purdue	Mean: 17.4 Gb/s	Max: 59.1 Gb/s
IN: Nebraska - US-Nebraska	Mean: 19.2 Gb/s	Max: 69.9 Gb/s
IN: NET2 - US-NET2	Mean: 8.43 Gb/s	Max: 9.43 Gb/s

Site Network Utilisation - computed from aggregated utilisation (SNMP counters) provided by sites via simple API. Screenshot shows network utilisation during DC24 as seen by the sites.

Analytics

Alarms & Alerts Interface

Components

Database
Elasticsearch

REST API and Web frontend
Node.js + express + pug

Deployment
Docker, K8s, Helm (soon)

Authentication
Globus InCommon

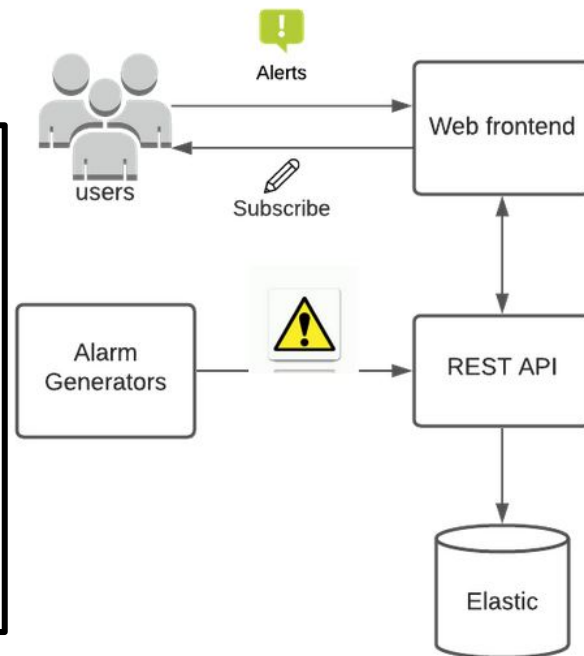**Authorization**
API key

Mail
Mailgun
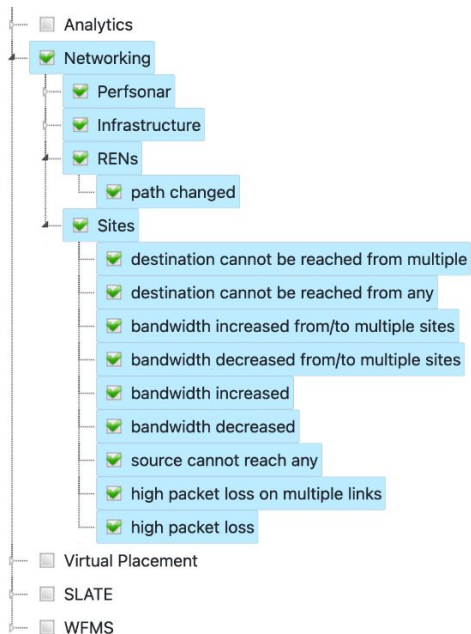
https://psa.osg-htc.org
(Uses EDUGain/InCommon)
Purpose: provides user-subscribable alerting for specific types of network issues found by analyzing perfSONAR data

Two main improvements needed: **Acknowledging alerts** that are being worked on and **adding user notification mailing lists**

Subscription Interface

Alarm Types and Relation to perfSONAR Data

All based on perfSonar data

One-Way Delay
▶ bad owd measurements
▶ large clock correction

Traceroute
▶ path changed
▶ destination cannot be reached
▶ source cannot reach any

Packet loss
▶ complete packet loss
▶ firewall issue
▶ high packet loss
(on multiple links)

Throughput
▶ bandwidth decreased
(from/to multiple sites)
▶ bandwidth increased
(from/to multiple sites)

psDash Alarms Dashboard

# Network Analytics R&D

- Investigate ML models/methods to process network measurements
- Data-preprocessing, e.g.
    - Train neural networks to predict network paths,
      e.g. help us fill the gaps in traceroute(s)
- Build model(s) that represents
  our network(s)
    - Network measurements are inherently
      noisy and therefore require robust
      models
- Use ML models for anomaly detection (for alerts & alarms)
    - Neural networks (which ones ?), Bayesian/probabilistic approaches,
    - Detect anomalies in network paths and bandwidth measurements
    - Compare with the existing heuristic algorithms that we have developed
- Correlate with other data
    - Traceroutes with throughput for example, but also outside of perfSONAR, e.g. FTS
    - New types of data appearing (high-touch, scitags, in-band telemetry, etc.)

Uncertainty in traceroute data

Repairing the path
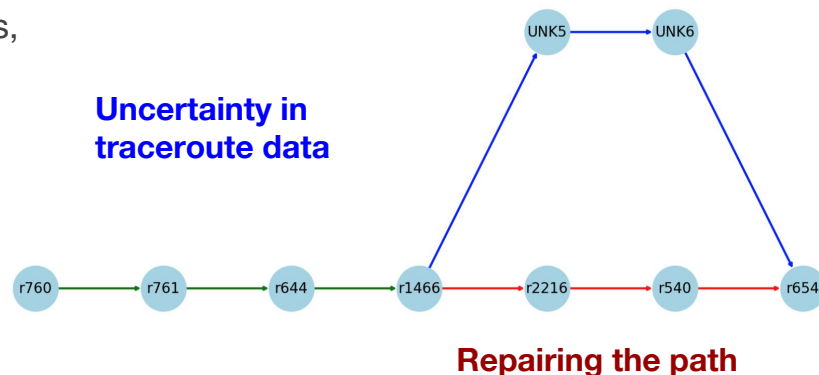
Plans for the Analytics Platform

- Production of the anomaly detection based on Bayesian inference
  - Uses RTT, traceroutes, TTLs as input and detects anomalies
- Continue working on the neural network models that correlate throughputs and traceroutes

  - Generating real-world model of our entire network (all routers)
  - Not only detecting anomalies, but also trying to pinpoint the location of the issue ⟹

- Improve infrastructure alarming to the point where we can **reliably differentiate infrastructure and network issues**
- Network availability dashboard in production
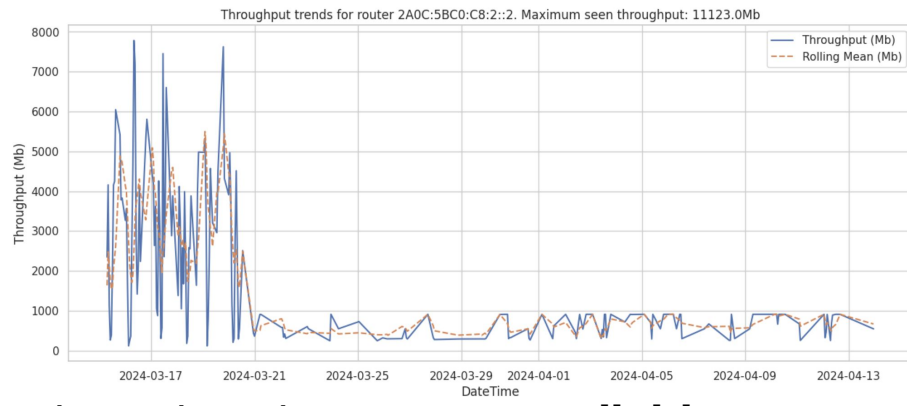
Summary

- Updates to perfSONAR and OSG/WLCG network measurement platform
  - perfSONAR 5.1 is coming with new features and will require all sites to update OS.
  - Plan to adapt the network measurement platform to benefit from changes in 5.1
- Ongoing efforts in network analytics and ML methods for our data
  - Focus on pre-processing (gaps, predictive models) and anomaly detection
  - Opportunity to collaborate on models and data sets
- **We are preparing monthly meetings with site network teams**:
  - Discuss how sites are deploying, managing and planning for WLCG networking requirements
  - Next meeting April 18th 10am EST (to join mail wlcg-site-net-requests@umich.edu)

- We have to continue to watch our network monitoring infrastructure as it is a complex system with lots of areas for issues to develop.

Questions / Discussion?
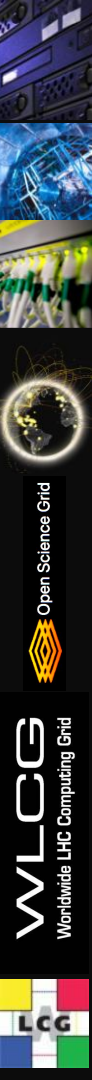
Acknowledgements

We would like to thank the **WLCG**, **HEPiX**, **perfSONAR** and **OSG** organizations for their work on the topics presented.

In addition we want to explicitly acknowledge the support of the **National Science Foundation** which supported this work via:

Useful URLs

- OSG/WLCG Networking Documentation
  - https://opensciencegrid.github.io/networking/
- perfSONAR Infrastructure Dashboard
  - https://atlas-kibana.mwt2.org:5601/s/networking/goto/9911c54099b2be47ff9700772c3778b7
- WLCG DOMA DC24 plans
  - https://indico.cern.ch/event/1225415/contributions/5155042/attachments/2593516/4476291/Data%20Challenge%202024.pdf
- perfSONAR Central Configuration
  - https://psconfig.opensciencegrid.org/
- Toolkit information page
  - https://toolkitinfo.opensciencegrid.org/
- Grafana dashboards
  - http://monit-grafana-open.cern.ch/
- ATLAS Alerting and Alarming Service: https://psa.osg-htc.org/
- The perfSONAR Dashboard application: https://ps-dash.uc.ssl-hep.org/
- ESnet WLCG Stardust Dashboard:
  https://public.stardust.es.net/d/XkxDL5H7z/esnet-public-dashboards?orgId=1

Backup Slides Follow

WLCG Network Throughput Support Unit

Support channel where sites and experiments can report potential network performance incidents:

- Relevant sites, (N)RENs are notified and perfSONAR infrastructure is used to narrow down the problem to particular link(s) and segment. Also [tracking past incidents](#).
- Feedback to WLCG operations and LHCOPN/LHCONE community

Most common issues: MTU, MTU+Load Balancing, routing (mainly remote sites), site equipment/design, firewall, workloads causing high network usage

As there is no consensus on the MTU to be recommended on the segments connecting servers and clients, LHCOPN/LHCONE working group was established to investigate and produce a recommendation.
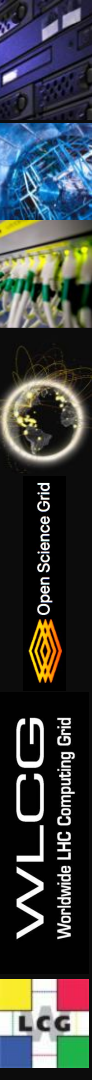
Importance of Measuring Our Networks

- **End-to-end network issues are difficult to spot and localize**
  - Network problems are multi-domain, complicating the process
  - Performance issues involving the network are complicated by the number of components involved end-to-end
  - Standardizing on specific tools and methods focuses resources more effectively and provides better self-support.
- **Network problems can severely impact experiments workflows and have taken weeks, months and even years to get addressed!**
- **perfSONAR provides a number of standard metrics we can use**
  - Latency, Bandwidth and Traceroute
  - These measurements are critical for network visibility
- **Without measuring our complex, global networks we wouldn't be able to reliably use those network to do science**