

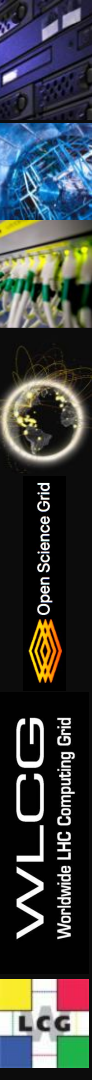
WLCG/OSG Network Activities, Status and Plans

Marian Babik (CERN), Shawn McKee (UM)
on behalf of SAND, IRIS-HEP & OSG/WLCG Network Throughput WG



Outline

- News
 - OSG/WLCG activities and WLCG Network Throughput WG
 - perfSONAR community updates
- LHCOPN/LHCONE perfSONAR status
- 100Gbps Testing
- Network Platform Updates
 - OSG Network Monitoring Platform
- Plans
- Summary

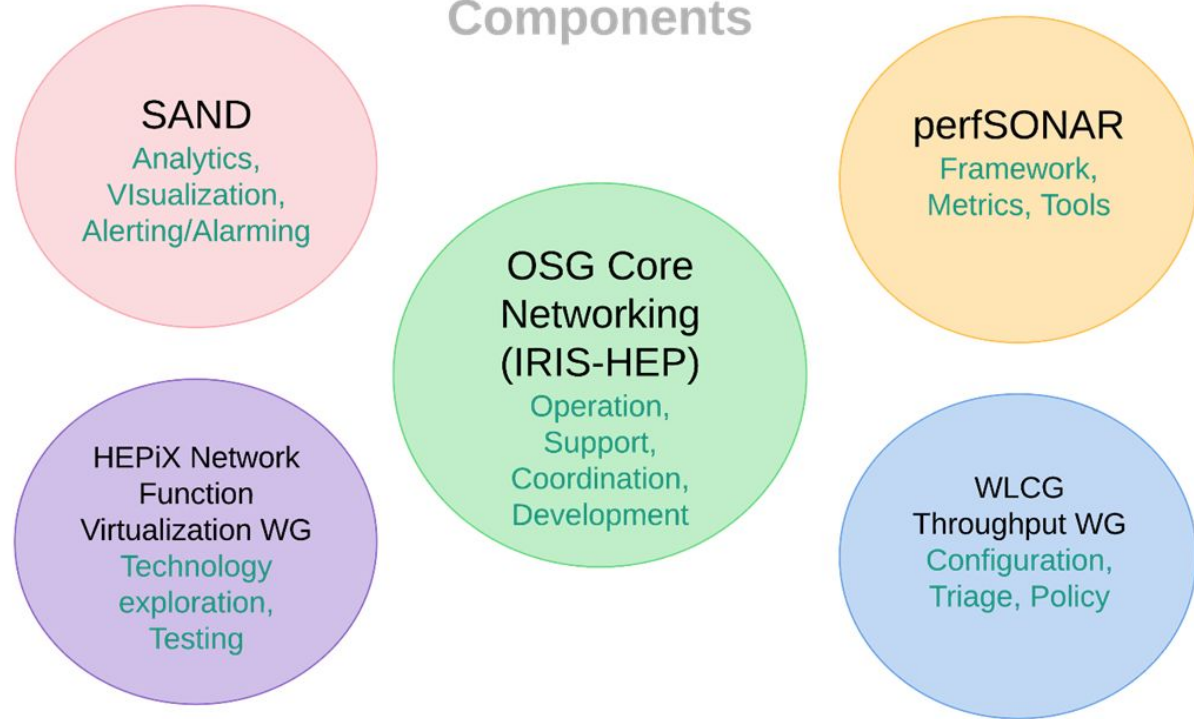


OSG/WLCG networking projects

There are 4 coupled projects around the core **OSG Net Area**

1. **SAND** (NSF) project for analytics
2. **HEPiX** NFV WG
3. **perfSONAR** project
4. **WLCG** Network Throughput WG

OSG Networking Components



perfSONAR deployment



288 Active perfSONAR instances

- **207 production endpoints**
- T1/T2 coverage
- Continuously testing over 5000 links
- Testing coordinated and managed from central place
- Dedicated latency and bandwidth nodes at each site
- **Open platform** - tests can be scheduled by anyone who participates in our network and runs perfSONAR

perfSONAR News

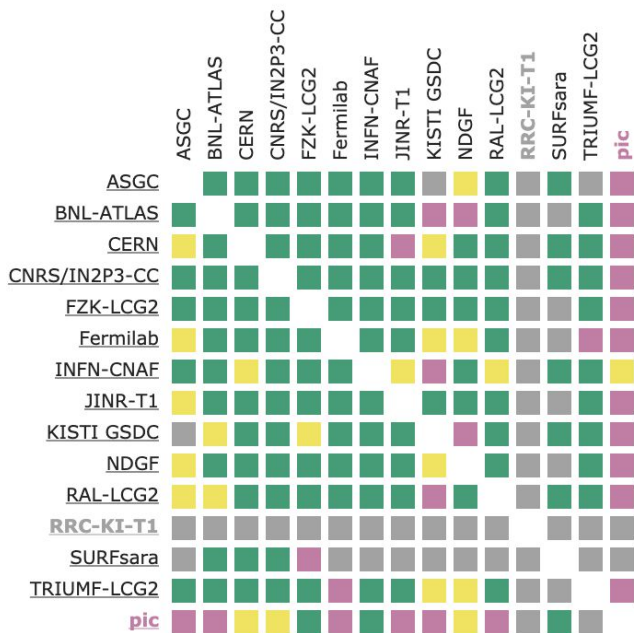
- 4.2.4 is the latest version (last one before 4.3)
 - This release fixes some of the performance issues (seen @CERN and other places)
- 4.3.0 - beta1 was released yesterday ([release notes](#))
 - Python3 support; PWA/psconfig support for multiple archivers
 - Integrated [ethr](#) and S3 benchmark test tools
- Traces were added to all our latency meshes
 - This way we have full coverage, which makes it easier to correlate with other data
- Communities that have joined recently or plan to use our infrastructure in the near-term:
 - WLCG ALICE, EU ESCAPE, EU ARCHIVER, DUNE, SLATE
- perfSONAR F2F developers meeting took place in June
 - Plan to move to ELK stack (ElasticSearch/Grafana)
- perfSONAR session at TechExtra Nov 2nd
 - Will discuss what's in the pipeline for 4.4 release
 - Demo of the ELK stack capabilities, PWA, WIFI and Cloud monitoring

LHCOPN 14th Sept 2020

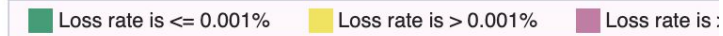
OPN Mesh Config - OPN IPv6 Bandwidth - Throughput



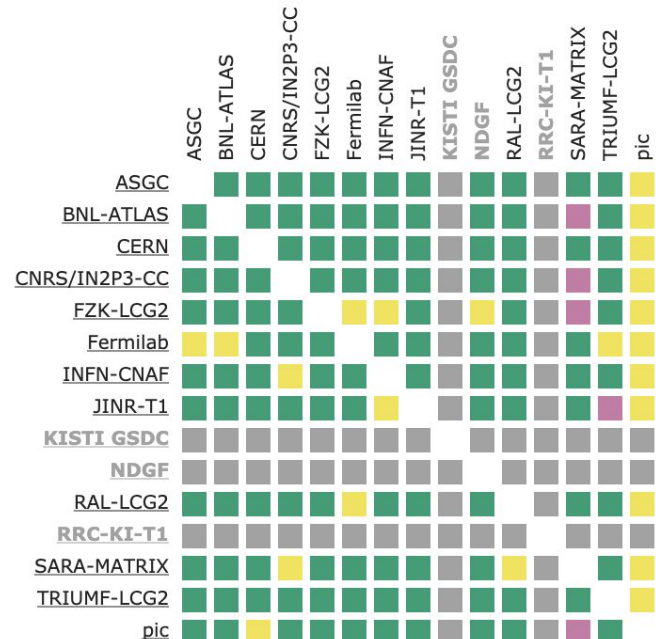
! Found a total of 3 problems involving 2 hosts in the grid



OPN Mesh Config - OPN Latency - Loss

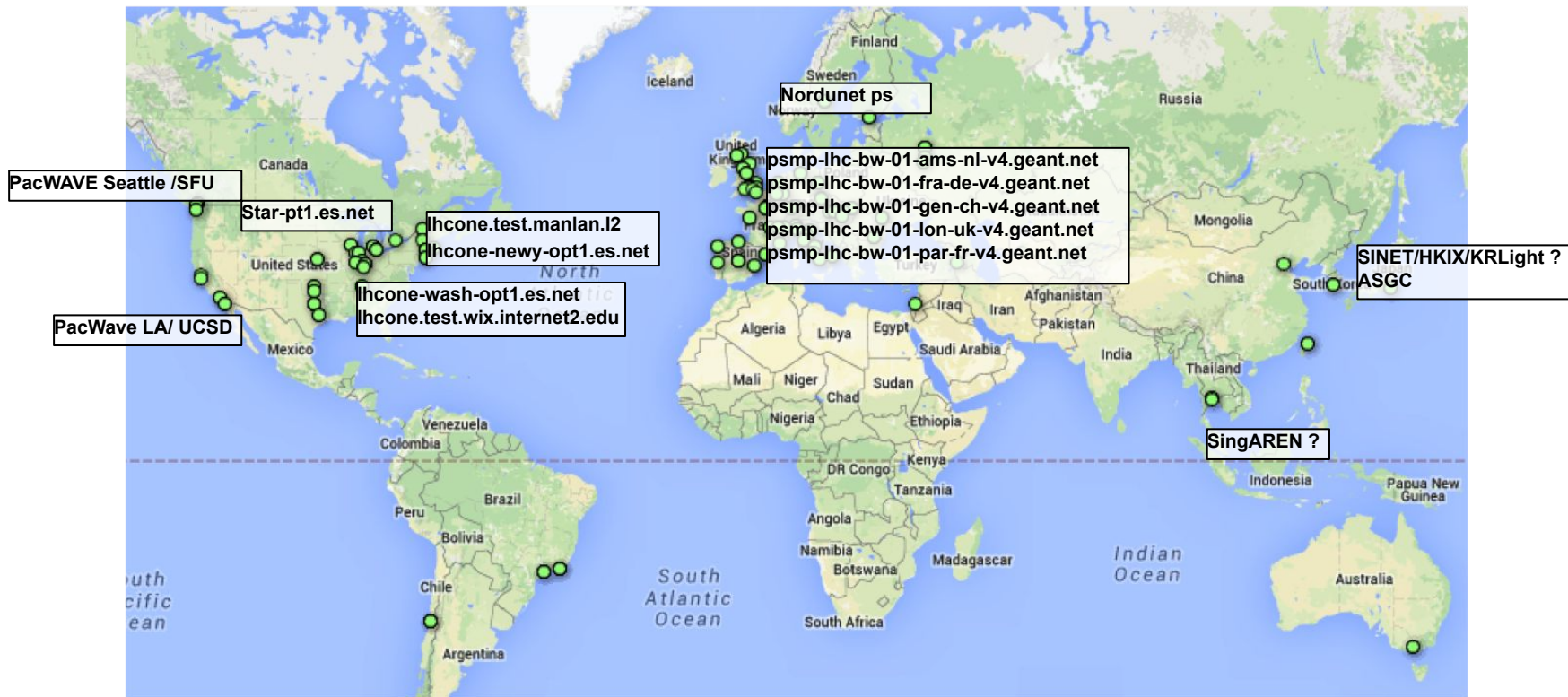


! Found a total of 3 problems involving 3 hosts in the grid



New LHCONE mesh

Testing from sites -> R&E endpoints on LHCONE; shows global connectivity



LHCONE 14th Sept 2020

LHCONE - LHCONE Bandwidth IPv4 - Throughput

■ Throughput \geq 1Gbps
 ■ Throughput $<$ 1Gbps
 ■ Throughput \leq .5Gbps
 ■ Unable to find test data
 ■ Check

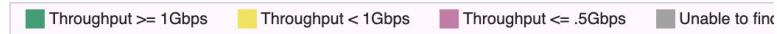
⚠ Found a total of 34 problems involving 27 hosts in the grid



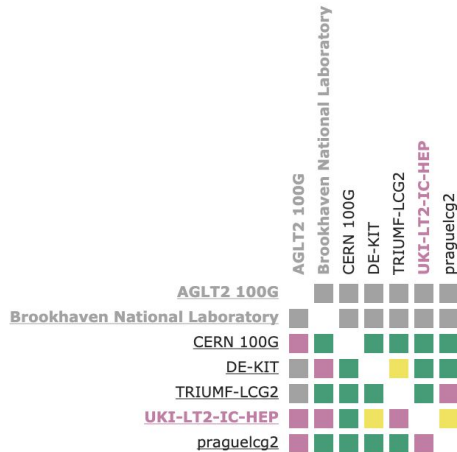
100Gbps Testing

- LHCOPN/LHCONE 100Gbps mesh was created

WLCG 100G Mesh - WLCG 100G IPv4 Bandwidth - Throughput



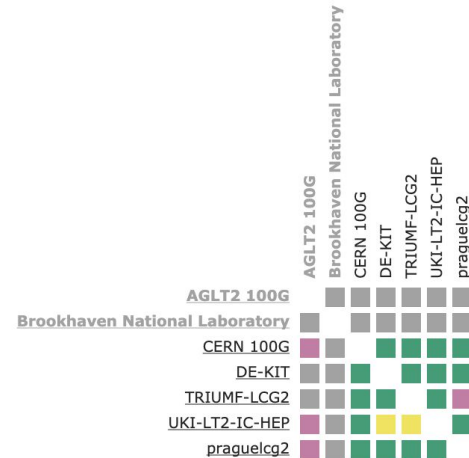
⚠ Found a total of 3 problems involving 3 hosts in the grid



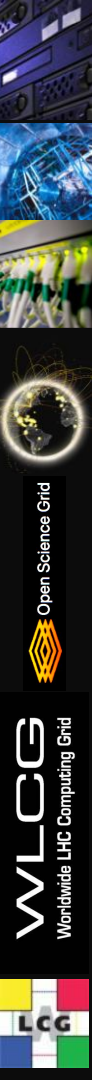
WLCG 100G Mesh - WLCG 100G IPv6 Bandwidth - Throughput



⚠ Found a total of 2 problems involving 2 hosts in the grid



- Plan to establish full US 100Gbps mesh (USATLAS/USCMS)
 - [100Gbps mailing list](#)



Network Operations

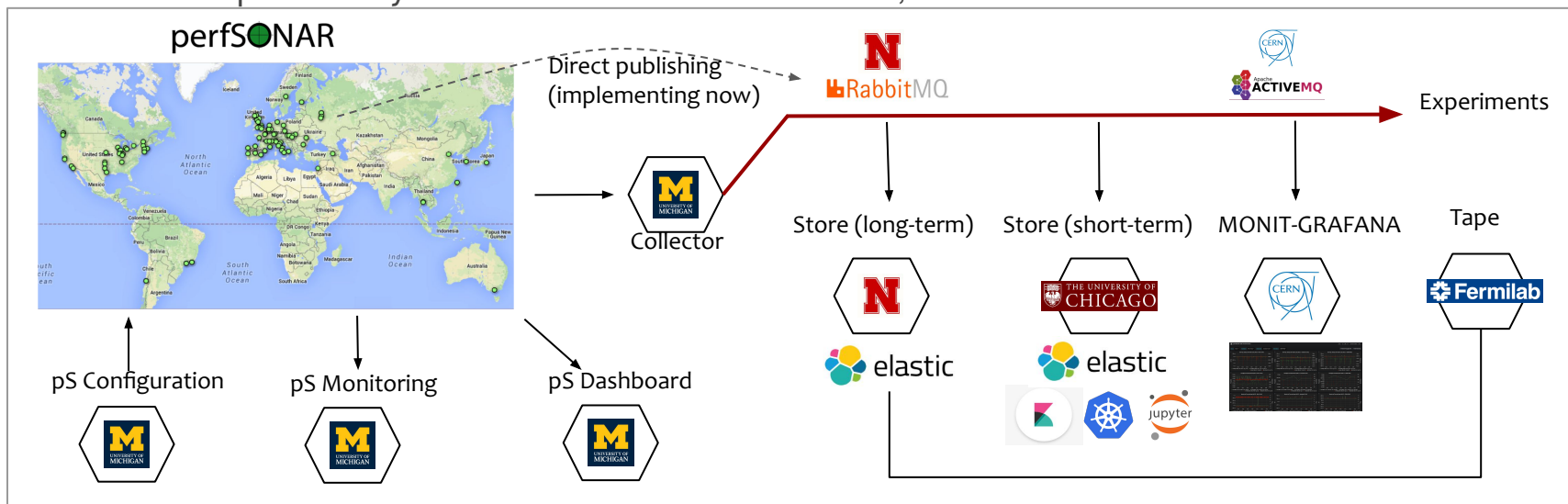
- Sites experiencing network issues should **first** contact their local network team or directly their regional and backbone (R&E) providers
- A group focusing on helping sites and experiments with network performance using perfSONAR - **WLCG Network Throughput**
 - Please include as many details as possible (include any existing tickets with R&Es)
 - For list of existing and recently resolved issues see this [link](#)
- **LHCONE operations** - support for establishing and operating LHCONE infrastructure - regular meetings and support mailing list
- **LHCOPN/LHCONE community** - organised bi-annually - good place to meet R&Es and discuss architecture and plans
 - Last meeting in Sept. (<https://indico.cern.ch/event/932306/>)

Operations

Architecture,
Infrastructure

Network Platform Overview

- Collects, stores, configures and transports all network metrics
 - Distributed deployment - operated in collaboration
- All perfSONAR metrics are available via **API, live stream or directly on the analytical platforms**
 - Complementary network metrics such as ESNNet, LHCOPN traffic also via same channels



Network Platform News

- Updated set of dashboards created in UC ES/Kibana
 - <https://atlas-kibana.mwt2.org/s/networking/goto/20dd25907d61df98a0b85b1dfaed54e1>
 - Provides high-level overview of latency, loss, traces and throughput
 - Implementation of the result of student projects to provide new user tools in our production system
 - Path analytics - Identify common hops between multiple src and dst pairs; detect path symmetry; find all unique paths across a period (Manjari Trivedi/UM)
 - [Site summaries](#) and reports of problematic node pairs (Petya Vasileva/UM) - blacklisting the problematic nodes (outliers) from the analytics
 - Use machine-learning to lay the foundations for a performance anomaly detection service (Edris Qarghah/UC)
- TRACer - [Path visualisation tool](#) developed by MEPHi
 - Still in [Beta](#), but already provides very interesting views into perfSONAR traces
 - Video of next release available at <https://yadi.sk/i/tyhiA-e3GGKqDQ>
- Updated toolkit info web page
 - <https://toolkitinfo.opensciencegrid.org/toolkitinfo> (Tommy Shearer/IRIS-HEP)

Collaboration with MEPhI on Network Visualization

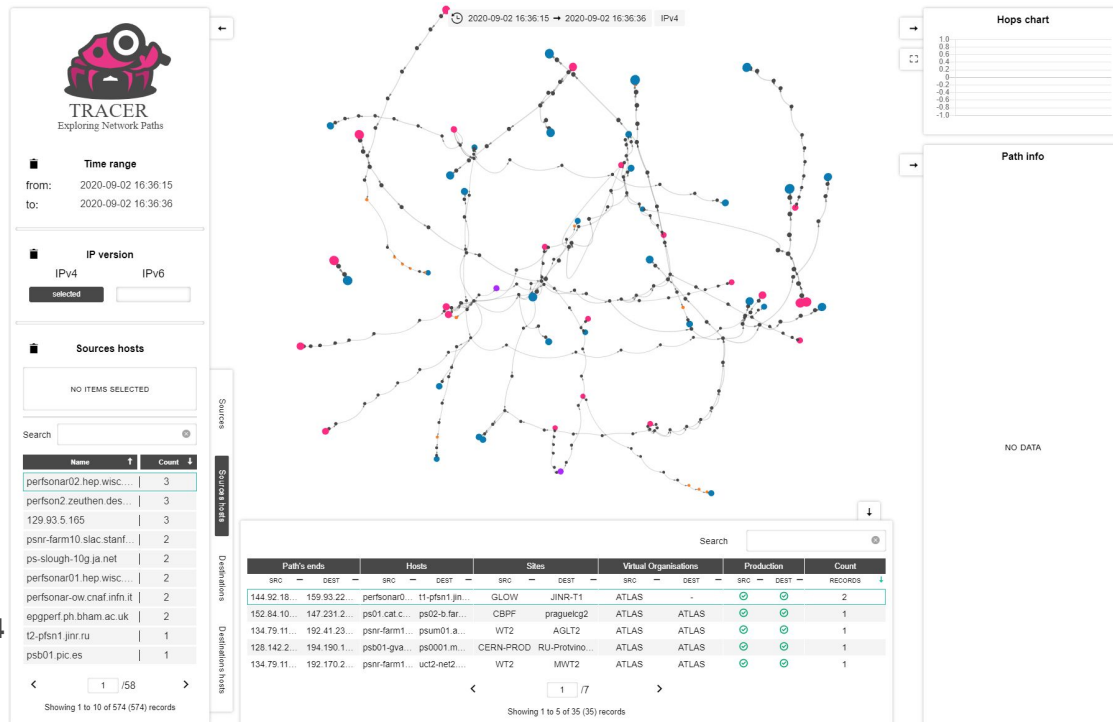
Containerized Version running at UC

We have been collaborating with [MEPhI](#) (Moscow Engineering Physics Institute) on network path visualization and have developed **TRACer** to visualize our perfSONAR path measurement data

TRACer Allows us to explore traceroute measurements from our OSG/WLCG perfSONAR measurements

NOTE: We have LOTS of traceroute measurements and TRACer can overwhelm your browser. When TRACer starts, it will load a very small interval of recent data, allowing the use to apply additional filters of interest (selectings sources, destinations and time windows). If you try to load too much data it will be sloooow. We are working on improving this behavior.

For more information, we have a preprint of a paper on TRACer available at <https://doi.org/10.6084/m9.figshare.12724865.v1>



Network Platform News

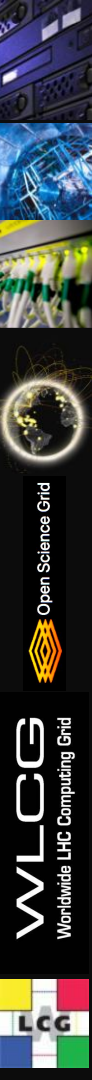
- Added **new data sources** to the messaging
 - Re-established SNMP from ESNet and added HTCondor transfer metrics
- **Directly publishing (push-mode)** from perfSONAR toolkit
 - Most of the developments are ready, testing has started (US region for now)
 - Developed in collaboration with perfSONAR developers
- **Infrastructure monitoring** - updates and new capabilities coming this fall:
 - Updates to **CheckMK version 1.6.x**, including plugin updates
 - Improvements to the configuration system to make it useful for other communities (dependent only on the perfSONAR ecosystem)
 - **Self-subscription for alerts** to toolkit contact, based on joining a particular community (using perfSONAR lookup service/information system)
 - **Migration to CRIC for topology** as well as working with CRIC team on potential improvements related to mapping perfSONARs to WLCG experiment site names
 - LHCONE Grafana dashboards based on newly established LHCONE mesh
 - New metric to track overall toolkit utilisation (for throughput tests)

Summary

- OSG in collaboration with WLCG are operating a comprehensive network monitoring platform
- Platform has been used in a wide range of activities from core OSG/WLCG operations to Cloud testing and IPv6 deployment
- Number of communities are now using our infrastructure:
 - ATLAS, CMS, LHCb, StashCache, Belle II, CCSTAR, Jisc (UK), GRIF/IN2P3 (FR), LHCOPN/LHCONE, ALICE, EU ESCAPE, EU ARCHIVER, DUNE and SLATE
- IRIS-HEP and SAND contributing to the R&D in the network area
- Current focus is on analytical studies that would lead to a production level network alerting service
- We also expect 100 Gbps mesh to grow with potential to contribute to the network data challenges and network performance studies

References

- OSG/WLCG Networking Documentation
 - <https://opensciencegrid.github.io/networking/>
- perfSONAR Stream Structure
 - http://software.es.net/esmond/perfsonar_client_rest.html
- perfSONAR Dashboard and Monitoring
 - <http://maddash.opensciencegrid.org/maddash-webui>
 - https://psetf.opensciencegrid.org/etf/check_mk
- perfSONAR Central Configuration
 - <https://psconfig.opensciencegrid.org/>
- Toolkit information page
 - <https://toolkitinfo.opensciencegrid.org/>
- Grafana dashboards
 - <http://monit-grafana-open.cern.ch/>
- ATLAS Analytics Platform
 - <https://indico.cern.ch/event/587955/contributions/2937506/>
 - <https://indico.cern.ch/event/587955/contributions/2937891/>



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:

- **OSG: NSF MPS-1148698**
- **IRIS-HEP: NSF OAC-1836650**
- **SAND: NSF CC* OAC-1827116**

Backup Slides Follow

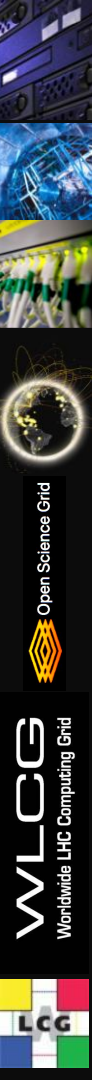
WLCG Network Throughput Support Unit

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. (See coming [talk](#) :))



IRIS-HEP (<http://iris-hep.org>)

The Institute for Research and Innovation in Software in High Energy Physics (**IRIS-HEP**) project has been funded by National Science Foundation in the US as grant OAC-1836650 as of 1 September, 2018.

The institute focuses on preparing for **High Luminosity (HL) LHC** and is funded at **\$5M** / year for 5 years. There are three primary development areas:

- Innovative algorithms for data reconstruction and triggering;
- Highly performant analysis systems that reduce 'time-to-insight' and maximize the HL-LHC physics potential;
- Data organization, management and access systems for the community's upcoming Exabyte era.

The institute also funds the **LHC part of Open Science Grid, including the networking area** and will create a new integration path (the **Scalable Systems Laboratory**) to deliver its R&D activities into the distributed and scientific production infrastructures. **Website for more info:** <http://iris-hep.org/>



The NSF funded SAND Project

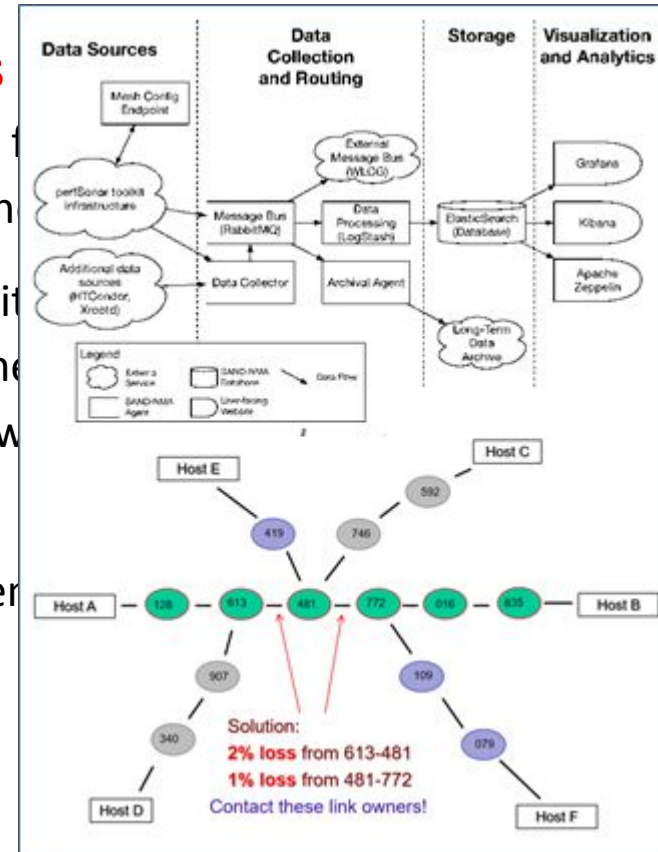
SAND: Service Analysis and Network Diagnosis

This is a newly funded NSF project (award #1827116) focused on visualizing, and analyzing disparate network monitoring and

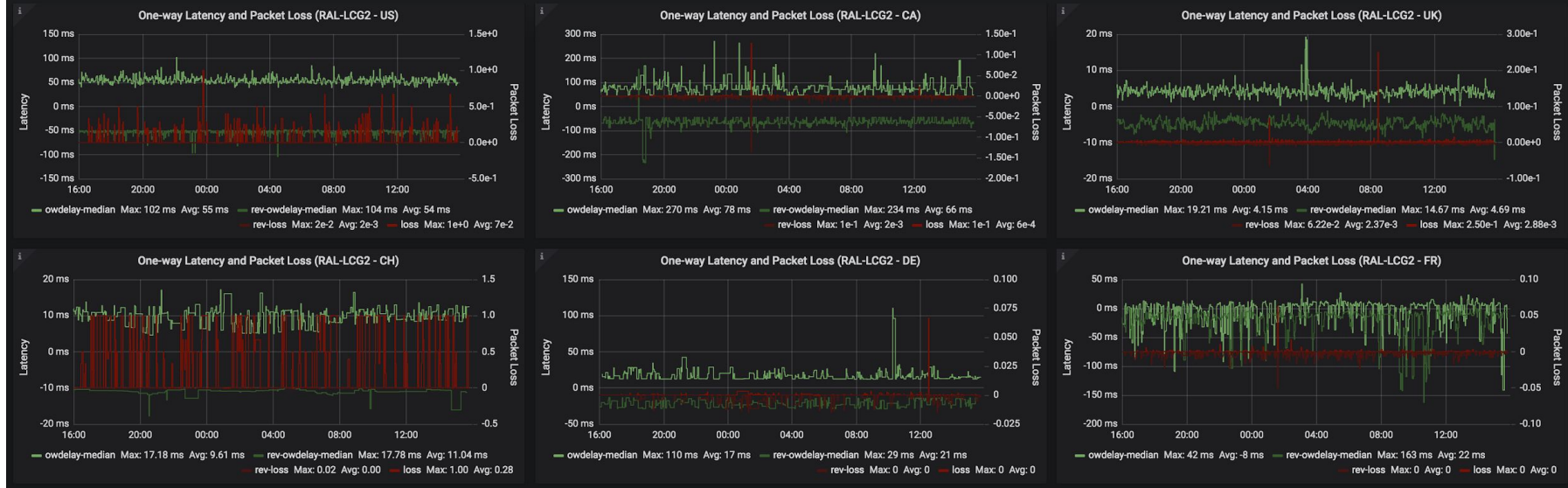
It will **extend** and **augment** the **OSG networking** efforts with extracting useful insights and metrics from the wealth of network data from perfSONAR, FTS, R&E network flows and related network monitoring tools like HTCondor and others.

Website <https://sand-ci.org/> (Project just started in September 2018)

PI: **Brian Bockelman**, Co-PIs: **Shawn McKee**, **Rob Gardner**



Grafana - Inter-Regional Latency Dashboard



Networking Challenges

There are number of challenges in the networking, which will require improved collaboration with other sciences as well as HEP-focused R&D:

- **Capacity/share for data intensive sciences**
 - No issues wrt available technology, however
 - What if N more HEP-scale science domains start competing for the same resources ?
- **Remote data access proliferating in the current DDM design**
 - Promoted as a way to solve challenges within experiment's DDM
 - Different patterns of network usage emerging
 - Moving from large streams to a mix of large and small frequent event streams
- **Integration of Commercial Clouds**
 - Impact on funding, usage policies, security, etc.
- **Technology evolution**
 - Software Defined Networking (SDN)/Network Functions Virtualisation (NFV)



Network Evolution Areas

The following are some of the key areas for HEP Networking R&D:

- Improving efficiency of data transfers
 - TCP BBR - version 2 is in the works with promising improvements
 - Exploring alternative protocols for transfers (UDP)
- Caching
 - Data caches co-located with network hubs in a similar way as on commercial CDNs
- Federations/Clouds
 - Overlay networks spanning multiple domains
 - Multi-clouds - expanding DC networking via L3VPNs
- Technology
 - SDN/NFV approaches - currently looked at by HEPiX NFV WG
 - Compute - Agile service delivery on Cloud Infrastructures (OpenStack, Kubernetes)
 - Data Transfers - Network resource optimisation - dynamically optimising the network based on its load and state (more in Shawn/Ilija)
 - SD-WAN approaches - <https://www.mode.net/>

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**

