

Network Topics Part 1

Shawn McKee

OSG All-Hands Meeting

Joint USATLAS-USCMS Networking Session

March 5, 2021

Online Meeting

Introduction

- In this session on networking we will have two presentations
- This presentation will cover topics related to sites, perfSONAR, RNTWG, packet marking and data challenges (15')
- Harvey will follow with a presentation on ongoing network R&D, network orchestration, GNA-G, data lakes and ESnet activities (15')
- We plan to have at least 10 minutes at the end for discussions

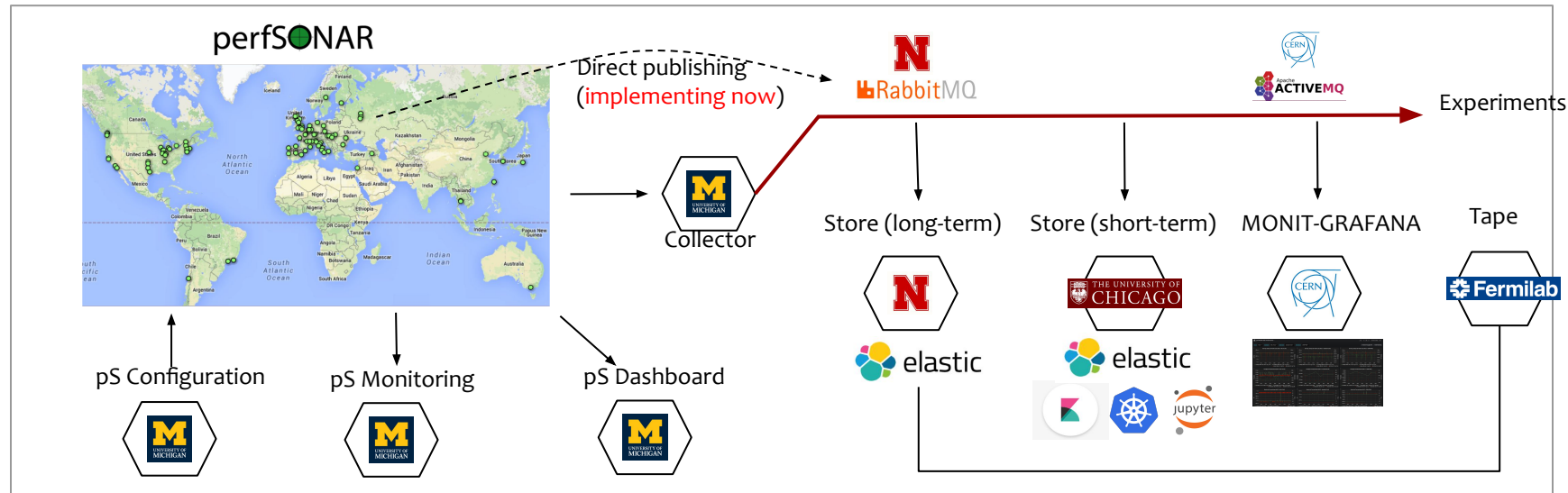
Current Network Status

Regarding our networking efforts, I want to state a few things about our current status up front

- Our networks continue to perform very well for our community
- Most users are happy with the networking we have but see areas for improvement.
- Primary concerns exist around our ability to fully utilize existing networks
- **Visibility** is key to understanding, maintaining and fixing our networks


So there continues to be **near-term** work regarding our networking in **optimizing, monitoring and fixing network problems**, **but** we should also think longer term regarding how the situation may evolve and what that might mean for us.

- 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 ESNet, LHCOPN traffic also via same channels



OSG Networking Near-Term Activities

- We continue working closely with the **SAND** (<https://sand-ci.org/>) project to:
 - Improve the **robustness** and **efficiency** of the data pipeline
 - Adding new analytics capabilities
 - Lots of student project underway the are coming to fruition by summer
- Finishing the pipeline improvements (reducing latency via direct writing to our RabbitMQ bus, improving configuration tools, monitoring and alerting)
- **Implemented a new mesh to track 100G perfSONARs for OSG_LHC**
 - 100Gbps mailing list (please join if you are interested)
 - <http://cern.ch/simba3/SelfSubscription.aspx?groupName=wlcg-perfsonar-100g>
- Working with the data in **Elasticsearch** to **correlate** and **visualize** traceroute paths with their related network metrics (packet-loss, latency, bandwidth)
 - Implementation of the result of student projects to provide new user tools in our production system.
 - Using some of the student developed projects to augment the types of issues that users can **self-subscribe** to for alerting



The screenshot shows the 'The perfSONAR Toolkit Information Page'. At the top left is the Open Science Grid logo. The title 'The perfSONAR Toolkit Information Page' is centered. On the top right is the WLCG logo. Below the title is a navigation bar with six buttons: 'Main Page', 'Documentation', 'OSG Network Pipelines', 'OSG Network Services', 'Analytics and Dashboards', and 'Toolkit Specific Page'. The main content area has the heading 'Please Select A Category Based on The Question You Have'. Below this are six buttons arranged in two columns, each containing a question: 'Do You Suspect General Networking Problems with Your Site?', 'Are you Interested in your Site's Overall Network Performance?', 'Do You Suspect a lot of Packetloss when Transmitting Data to Another Site?', 'Are you Interested in the Path that Data Takes from your Site to Another?', 'Do You Suspect Slow Connections between Your Site and Another?', and 'Are you Seeing a lot of Throughput Issues or Retransmitted Data for your Site?'. At the bottom of the page are three logos: SAND, iris hep (Institute for Research & Innovation in Software for High Energy Physics), and NSF. A link 'Contact us about this webpage' is centered below the logos.

Open Science Grid

The perfSONAR Toolkit Information Page

WLCG
Worldwide LHC Computing Grid

Main Page Documentation OSG Network Pipelines OSG Network Services Analytics and Dashboards Toolkit Specific Page

Please Select A Category Based on The Question You Have

[Do You Suspect General Networking Problems with Your Site?](#)

[Are you Interested in your Site's Overall Network Performance?](#)

[Do You Suspect a lot of Packetloss when Transmitting Data to Another Site?](#)

[Are you Interested in the Path that Data Takes from your Site to Another?](#)

[Do You Suspect Slow Connections between Your Site and Another?](#)

[Are you Seeing a lot of Throughput Issues or Retransmitted Data for your Site?](#)

SAND

iris hep
Institute for Research & Innovation
in Software for High Energy Physics

NSF

[Contact us about this webpage](#)

We have been evolving the perfSONAR toolkit “front-end” web page to better guide end-users to the information they need. As noted in Tommy Shearers’ slide above, we now have an updated interface running that we plan to improve based upon your feedback: <https://toolkitinfo.opensciencegrid.org/> Please give us feedback via email to shearert@umich.edu

Our community has identified a few critical networking capabilities they want to have:

1. Making our network use visible (marking)
2. Shaping WAN data flows (pacing)
3. Orchestrating the network to enable multi-site infrastructures (orchestrating)

To move forward we organized a new **Research Networking Technical Working Group (RNTWG)** in spring 2020 along with three sub-groups (one per topic above)

Charter for the group is at

<https://docs.google.com/document/d/1I4U5dpH556kCnolHzyRpBI74IPc0gpgAG3VPUp98lo0/edit?usp=sharing>

The current focus is on “packet marking” and initial testing work used our OSG/LHC perfSONAR’s to verify markings could make it across our networks.

OSG could be a natural place to test some developments as they are ready

Network Packet Marking

- Full stack packet marking for **visibility, accounting** and **future network optimizations**
 - Work being organized by the RNTWG Packet Marking subgroup
 - Charter link
 - Presentation to HEPiX IPv6 WG (January 19th 2021)
 - Last meeting was earlier this week (March 3, 2021)
 - Would like to see XRootD able to mark packets ASAP.
 - Need to discuss how best to communicate what flow-label to set.
 - Proposal from Wei Yang and Andy Hanushevsky
 - **Contact:** Shawn McKee / Marian Babik

This work could benefit not only the collaborations using our global Research and Education (R&E) networks but also the R&E networks themselves.

The visibility of what traffic is flowing, ***at any point in the network***, is critical for understanding and optimizing how we utilize the network.

A recent focus for the WLCG experiments has been on defining and executing a series of data challenges that can bridge where we are now with where we need to be for HL-LHC

- We need to test and improve our ability to exploit networks, both for Run-3 and for HL-LHC
 - Initial discussions held (ATLAS, CMS, WLCG); the consensus, so far, is to define a series of incremental tests (every ~ 2 years) to build to HL-LHC scale
- Short draft [document is here](#) which started the discussions
 - Need to understand our workflow for T2s, analysis, production (impact on LHCONE)
- Challenges should be relevant for each experiment (not one-off “demos”) so the effort put into each challenge can incrementally improve each experiment’s ability to use the network. **Goal** is to identify deficiencies and incorporate improvements.

We need to start this year and most discussions to date point to a fall data challenge where all experiments can stress their production infrastructure to identify bottlenecks and components in need to tuning or upgrades.

A major challenge will be getting the needed monitoring in place to understand what the challenge showed us!

We have an **opportunity** to prepare for the HL-LHC era over the next six years and we should use this period to prepare our sites

To be successful, we will need a large fraction of our sites participating in developing the next generation of infrastructure and associated networking.

Suggestions for site networking:

- Update/replace your perfSONAR hardware; match site storage interfaces
- Renew/refresh/replace your LAN and WAN equipment, enabling your site to test and prototype the next generation of capabilities.
- Don't forget security; participation in WLCG SOC is a good starting point.
 - This involves some hardware and services as part of any implementation.
- Plan for automating your network provisioning to allow agile, consistent site configuration.
- Identify and publish your site network monitoring links (in CRIC?)

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

Special thanks to Marian Babik, Derek Weitzel, Ilija Vukotic, John Thiltges and many others!

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* INTEGRATION 1827116

OSG has a working network monitoring infrastructure measuring our sites and research and education networks.

The near-real-time network data we are gathering is a unique resource we need to exploit to proactively identify network problems and provide network capacity information to users and applications.

Sites should be aware of network technology evolution and be planning how best to take advantage of them as they evolve their infrastructure

Let's move on to Harvey and discuss after...

Draft document on HL-LHC Network Challenges

https://docs.google.com/document/d/1sVnfkUS_7uh892eTtHUnPPcbpEaYwyCxcUEWAZtmAQA/edit?usp=sharing

Research Networking Technical Working Group

https://docs.google.com/document/d/1l4U5dpH556kCnoIHzyRpBI74IPc0gpgAG3VPU_p98lo0/edit#heading=h.nxusujd3hvwy

Packet Marking WG

<https://docs.google.com/document/d/1aAnsujpZnxn3oIUl9JZxcw0ZpoJNVXkHp-Yo5oj-B8U/edit#heading=h.kjs85ae6lo7a>

Xrootd Flow-Label Passing

<https://docs.google.com/document/d/1HTaNwv7huRqdNUvgHJTjlow8MivJgoknRUKgADNlvgY/edit#heading=h.c84ryvst43hq>

DOMA Challenges and Prototypes

<https://docs.google.com/document/d/1-eY8nKyAFZBnMZKvE0d5SB8UuwEKgY5rzJ4dxbS8e8g/edit>

- OSG/WLCG Networking Documentation
 - <https://opensciencegrid.org/networking/>
- 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/toolkitinfo>
- 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/>

Additional Slides



Open Science Grid

Collaboration with MEPhI on Network Visualization



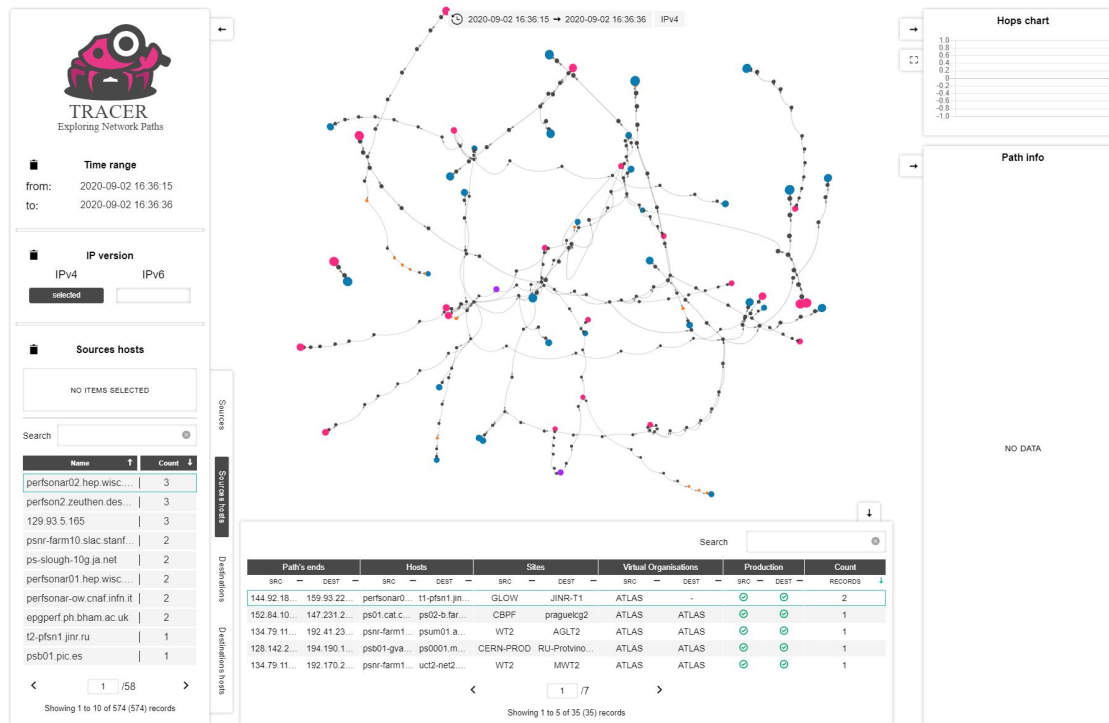
Containerized Version running at UC <https://perfsonar.uc.ssl-hep.org/> (Try it!)

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 Security Considerations

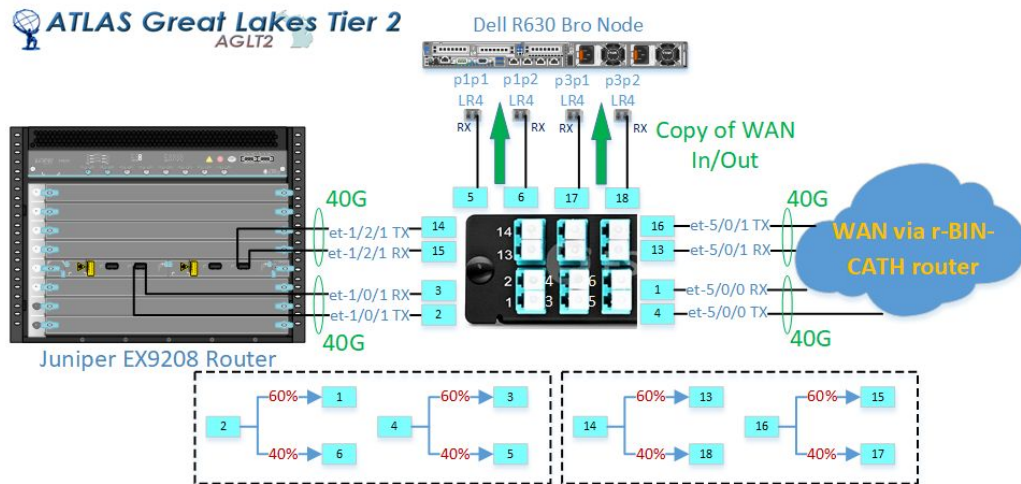
Enabling some additional security via traffic monitoring can help your site identify and defend against attacks AND help other collaborating sites to benefit.

The **WLCG Security Operations Center** working group has a documented solution involving **Zeek** (formerly “Bro”), **MISP** and **Elastiflow**

Info at <http://wlcg-soc-wg-doc.web.cern.ch/wlcg-soc-wg-doc/>

As part of site infrastructure evolution, OSG networking would like to recommend that sites deploy security monitoring and collaborate between institutions.

For now, email expressions of interest to smckee@umich.edu

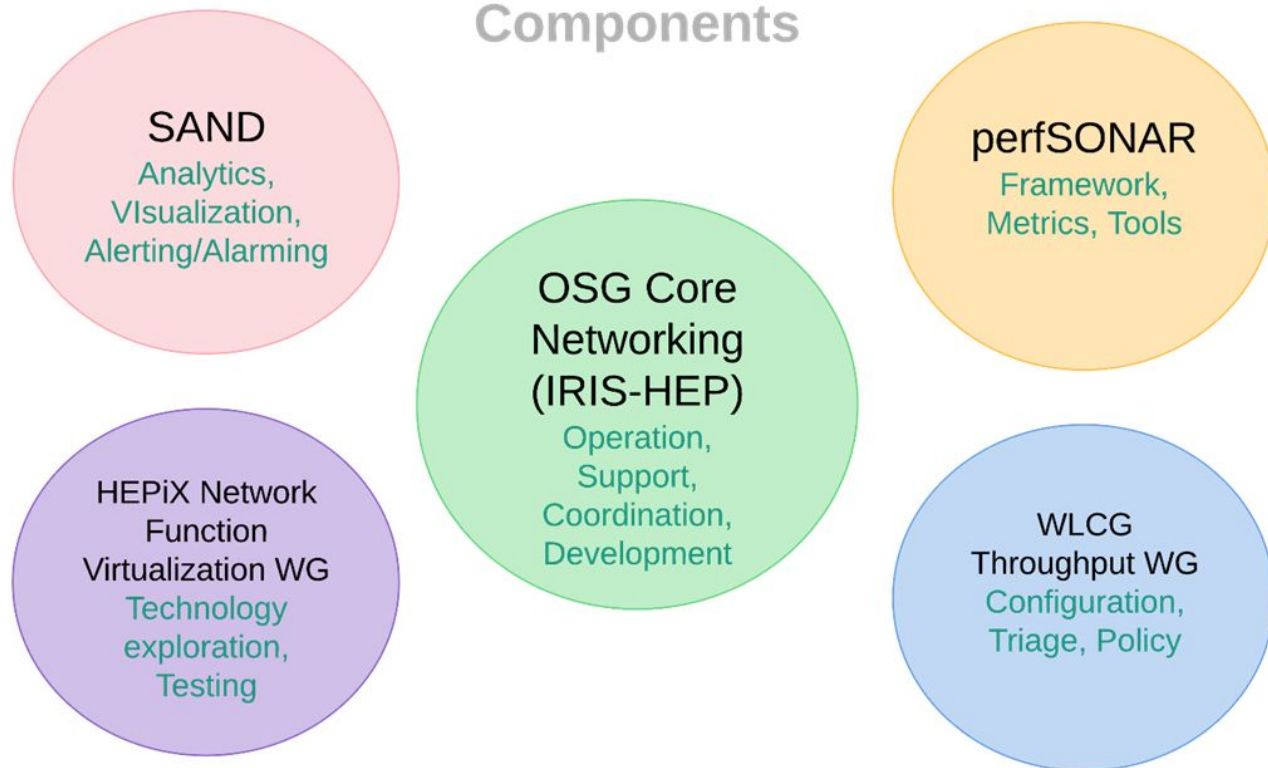


- o Deployment of perfSONARs at most OSG/WLCG sites made it possible for us to see and debug end-to-end network problems
 - o OSG gathers global perfSONAR data and making it available to collaborators
- o We have a group focusing on helping sites and experiments with network performance using perfSONAR - WLCG Network Throughput
 - o Reports of non-performing links are actually quite common
- o Sites with assumed network problems can open a ticket with OSG to allow us to help diagnose the issue
- o Sites experiencing **known** network issues should first contact their local network team, who can pursue the issues with the regional and backbone providers on their behalf

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

OSG Networking Components

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



Longer Term Outlook

- Long term outlook (5-10 years) will likely involve:
 - **Capacity sharing** - other big research domains coming online
 - **Better Monitoring** - Packet Marking at the ends, flow monitoring
 - **Traffic Shaping** - Optimizing flows to match resources
 - **(Re)organisation** - evolution of LHCONE (MULTIONE?), possibly some form of SD-WAN (dynamic circuits/L3 VPNs on demand)
 - **Cloud networking** - network federations spanning multiple data centres (inc. commercial clouds)
 - **Network orchestration**: ability to develop and operate services across large number of heterogenous sites via managed networks
- Are we going to be ready ?