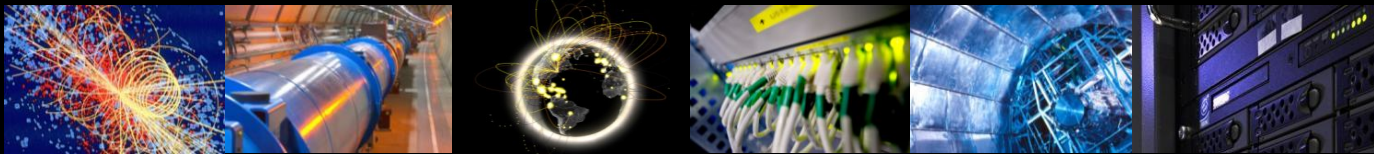


# Network Monitoring R&D (and data challenge needs)

Shawn McKee

WLCG Data Challenge Monitoring Mini-Workshop

April 27, 2021



# Overview

Our upcoming series of **data challenges** will require detailed monitoring, much of which we don't yet have. This meeting starts the discussion about what we need and how we get it.

My bias is that the data challenges should identify shortcomings in our existing infrastructures and fix them before the next challenge, incrementally working our way to the HL-LHC era

In this talk I will focus on the **network aspects** of the data challenges.

Note that the perfSONAR network monitoring infrastructure we have provides **network diagnostics, problem identification and localization** which is only part of what we need.

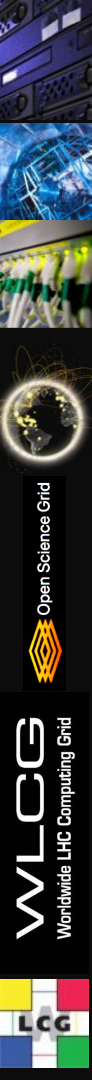
For the data challenges we also need to understand how we are using our networks, what we are able to achieve in terms of sustained throughput and, perhaps most importantly, where we have bottlenecks, whatever their cause.



# Likely Network-related Issues To Expect

I want to start by summarizing the most obvious challenges we should expect to see in the network:

- **Firewalls** interfering with throughput or connectivity
- **Misconfigurations** in MTU, network stack/host tuning, routing and applications which can severely impact achievable performance.
- **Old equipment** (switches, routers and hosts) which are unable to exploit the available network capacity or unable to utilize newer features available in the network
- **Not-old equipment** but in need of upgrades (in OS or firmware) or tuning to to fix bugs, performance or provide improved robustness or features
- **Architectural problems** in the LAN and WAN
  - A critical feature of the data challenges is the **simultaneous** use of our shared infrastructure to be able to identify potential bottlenecks that appear when jointly stressing the networks
  - The existing design of many campus and site networks may not be appropriate for their current and future use-cases



# Site Networking Considerations

We have an **opportunity** to prepare for the HL-LHC era over the next six years and we should use the Data Challenges 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 (soon in [CRIC](#))

# Proposed Site Monitoring Page

We need sites to provide us with a site network information we can use to understand the results of our data challenges. My proposal for discussion is that sites create a web page with three sections:

**First section** should include links to real-time network monitoring that, **at a minimum**, provide the site ingress and egress network rate (MBytes/sec)

- More detailed monitoring is welcome and encouraged
- Adding descriptions of the monitoring technology and characteristics is strongly recommended

**Second section** should provide a summary of the site network: overview of equipment (vendor/model) and a description of the LAN and connectivity to the WAN.

**Third section** (optional?) should provide site network diagrams showing how storage and compute are connected to the LAN and how the WAN is reached.

# Research Networking Technical Collaboration

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

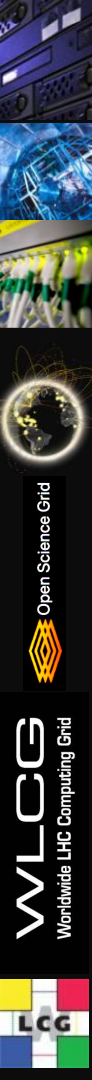
1. Making our network use visible (marking)
2. Shaping WAN data flows (pacing)
3. Orchestrating the network to enable multi-site infrastructures (orchestrating)

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/1I4U5dpH556kCnoIHzyRpBI74IPc0gpgAG3VPUp98lo0/e/dit?usp=sharing>

This group is focused on a few practical and technical issues to improve how HEP (and others) are utilizing their networks.



# Network Packet Marking

The current focus of the RNTWG is on developing the ability to mark our packets to provide **visibility, accounting and future network optimizations.**

- Work is 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

The visibility of what traffic by owner and purpose, ***at any point in the network***, is critical for understanding and optimizing how we utilize the network and would be **especially useful for understanding WAN flows during our data challenges!**

# Identifying Expected Network-related Issues

Let's revisit the expected network issues and how we might identify them:

- **Firewalls** interfering with throughput or connectivity perfSONAR Monitoring
- **Misconfigurations** in MTU, network stack/host tuning, routing and applications which can severely impact achievable performance. perfSONAR Monitoring
- **Old equipment** (switches, routers and hosts) which are unable to exploit the available network capacity or unable to utilize newer features available in the network. SiteWeb Monitoring
- **Not-old equipment** but in need of upgrades (in OS or firmware) or tuning to to fix bugs, performance or provide improved robustness or features Monitoring
- **Architectural problems** in the LAN and WAN
  - A critical feature of the data challenges is the **simultaneous** use of our shared infrastructure to be able to identify potential bottlenecks that appear when jointly stressing the networks
  - The existing design of many campus and site networks may not be appropriate for their current and future use-cases SiteWeb Monitoring



# Data Challenge Preparation: Possible Areas of Work

Here is a discussion list of possible areas of work to pursue for the DCs:

- Create complete **example** network monitoring page for a site
- Aggressively pursue problems identified by perfSONAR monitoring
  - Firewalls, packet-loss, low throughput, flapping routing, etc
- Develop and deploy script to check and recommend host network tuning
- Work with sites to get appropriate network monitoring page in place
  - Initial focus should be on the biggest sites
- Work with R&E networks to identify suitable monitoring resources
- Centralize data from site and R&E network monitoring
- Prototype and deploy packet marking as possible for this Fall's Challenge
- Begin traffic shaping testing and deployment (likely a focus for 2023).
- Implement and test Network Orchestration capabilities (GNA-DIS)
- Create site best practices documentation regarding network monitoring, tuning and architecture

**Others?**

# Summary

- Our **Data Challenges** will help expedite getting the right monitoring in place to understand and identify how we are using our infrastructure and where bottlenecks exist.
- We have a large amount of work in various areas that need attention, even just within networking, and we need to plan for incremental improvements over time.
  - We need to target specific components we deem most important
  - Identifying longer term items is important to ensure they will be usable before HL-LHC
  - Finding the effort will be challenging but we also need to consider that effort is often not shareable or redirectable; if we don't push ahead on multiple fronts we may lose time or effort.

# 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:

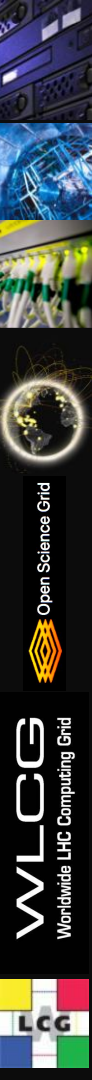
- [OSG: NSF MPS-1148698](#)
- [IRIS-HEP: NSF OAC-1836650](#)



**Questions or Comments?**

# References

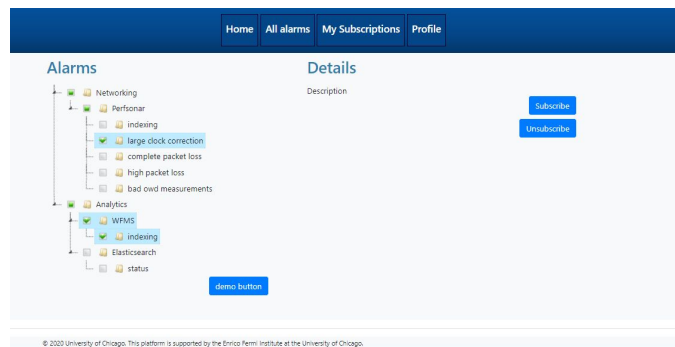
- OSG/WLCG Networking Documentation
  - <https://opensciencegrid.github.io/networking/>
- perfSONAR Infrastructure Dashboard
  - <https://atlas-kibana.mwt2.org:5601/s/networking/goto/9911c54099b2be47ff9700772c3778b7>
- perfSONAR Dashboard and Monitoring
  - <http://maddash.opensciencegrid.org/maddash-webui>
  - [https://psetf.opensciencegrid.org/etf/check\\_mk](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/>



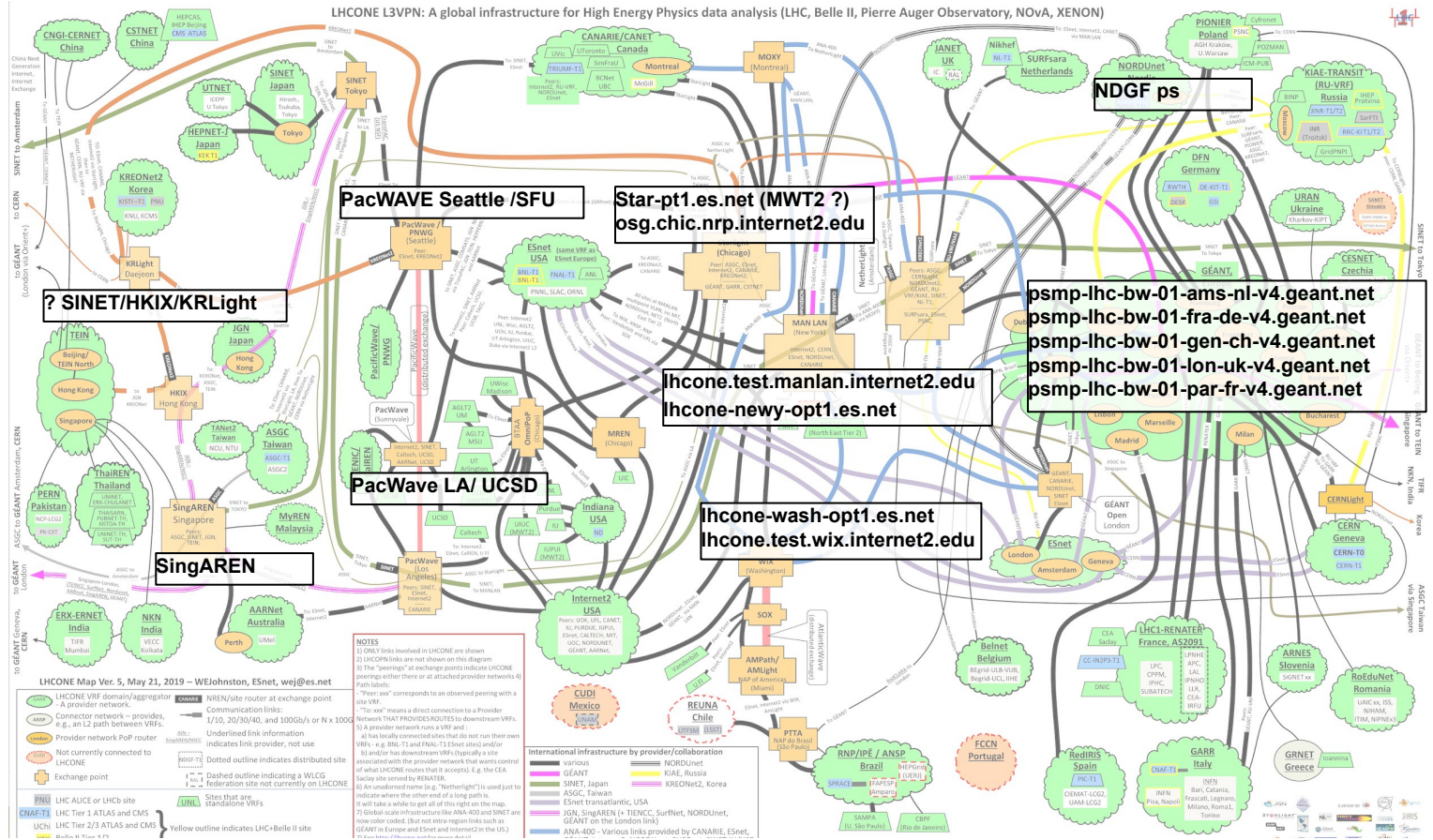
**Backup Slides Follow**

# perfSONAR Plans and Near-term Activities

- The focus is on analytics and getting issue identification into the production pipeline
  - Students from IRIS-HEP and SAND are working on this
  - Prototype user-alarm-subscription developing ->
- We already have Kibana dashboards looking at
  - [Bandwidth](#)
  - [Traceroute](#)
  - [Packetloss](#) / [Latency](#)
  - [Infrastructure](#)
- To organize access to all the various resources we have developed we created <https://toolkitinfo.opensciencegrid.org> (Try it; give feedback!)



# Reminder: New LHCONE mesh





# 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](#) :) )

