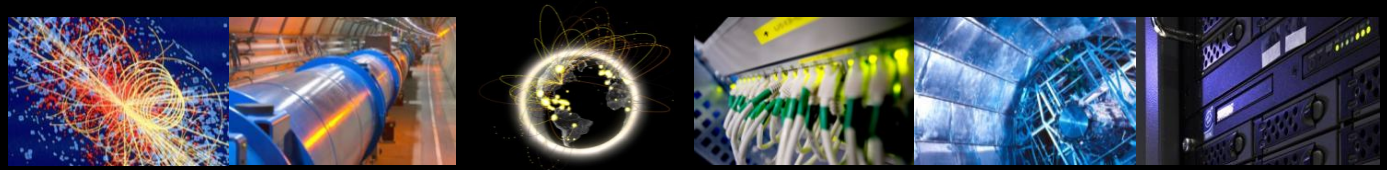
Networking: perfSONAR, SciTags and WLCG Site Network Monitoring

Shawn McKee / University of Michigan Physics

HTC25 USATLAS-USCMS Session 1, June 5, 2025

on behalf of WLCG Network Throughput WG & RNTWG



WLCG Worldwide LHC Computing Grid

OSG

Overview

I want to provide status and updates for WLCG Networking relevant for WLCG HEP, to both inform you and discuss where relevant:

- perfSONAR status, deployment and near-future updates
- SciTags news and deployment
- WLCG Site Network Monitoring update

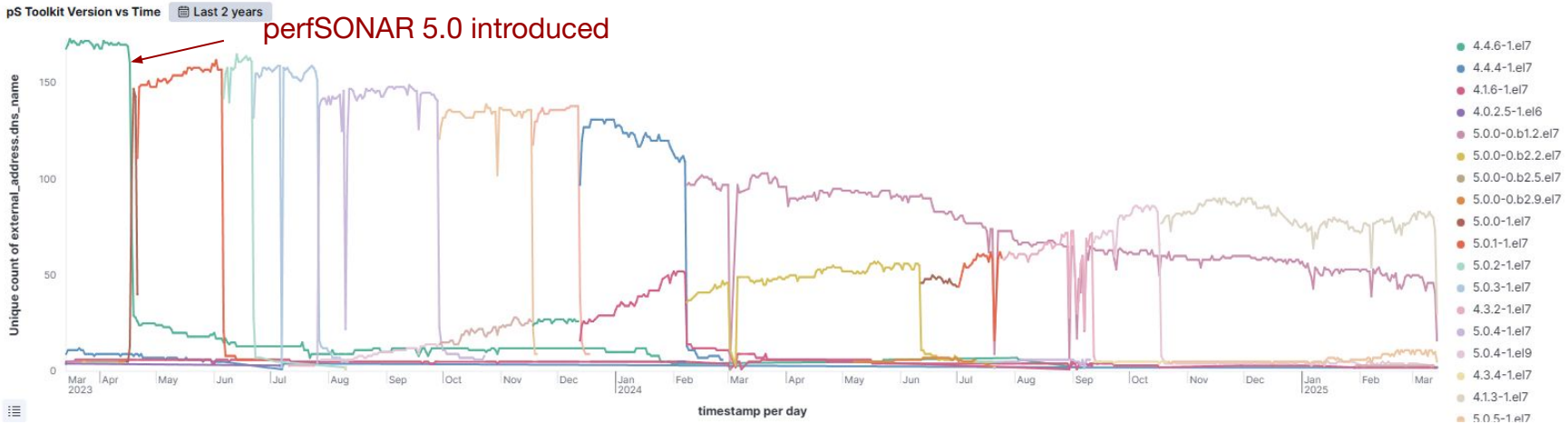
We can cover additional items based upon interest during our discussion block.

I also wanted to note that this is the work of many people, including Marian Babik/CERN, Ilija Vukotic/UChicago, Petya Vasileva/UM

perfSONAR Infrastructure Evolution

175 Active perfSONARs

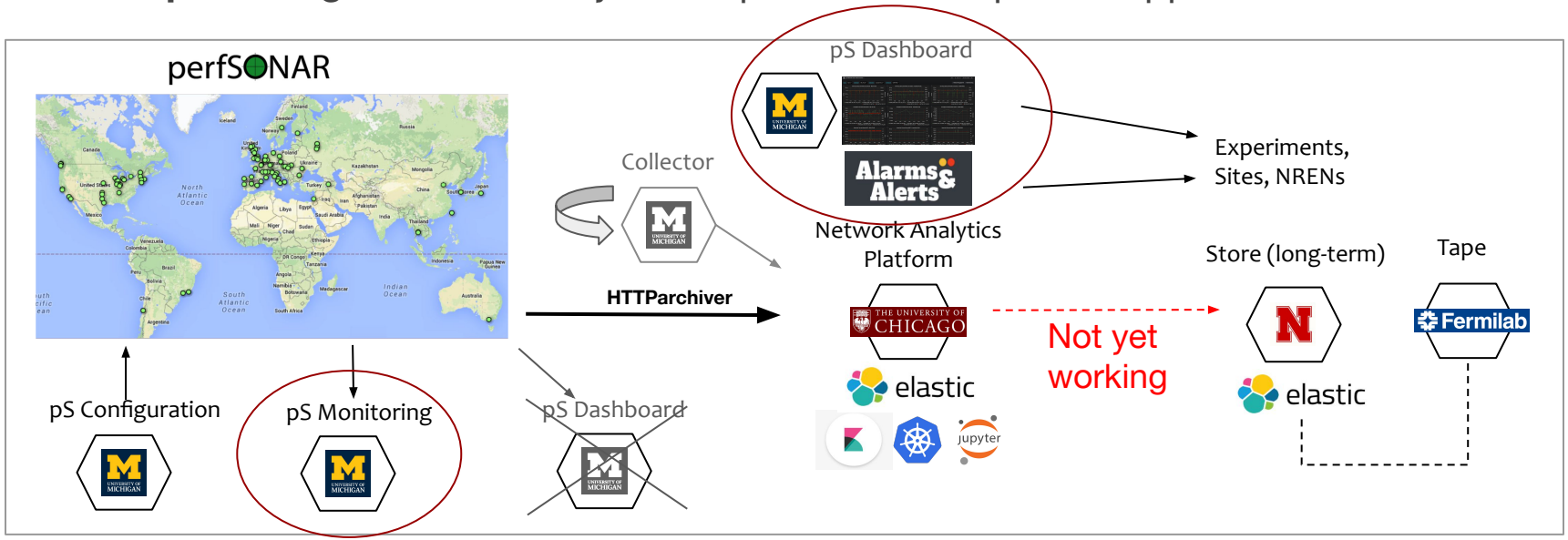
65 Communities in Use



- Long-term trend of perfSONAR releases adoption shows a steady downtrend
  - Only reachable “toolkit version” installations are shown (stacked plot)
  - Stability and reliability of the releases clearly playing a role
- New strategy and deployment models will need to be considered
  - Exploring simple deployments (without complex components such as OS)
  - Providing means to easily reset/restart deployment (to avoid accumulation of issues)
  - Easy to co-locate with storages (or even co-hosted on storages with multiple NICs)

Network Measurement Platform Plans

- Our platform/pipeline is evolving, with some components replaced or retired.
  - Forwarding to UNL and backup to FNAL still to be implemented
  - **pS Monitoring** - Enable changes to get container monitoring (via Prometheus)
  - **ps Dashboard** - integrate with Analytics Platform/Grafana (MaDDash retired)
  - **ps Configuration** - clarify development roadmap and support



perfSONAR News

- **perfSONAR 5.2.0** was just released 3 days ago
  - Number of bug-fixes since 5.1
  - We will need to get a sufficient number of sites running the new version for a while to see if this release helps for robustness and resource limitations we have seen.
  - See Release Notes for more details
- **Our current deployment has issues with resiliency and reliability.**
  - Possible issue: trying to test to many things with some toolkits
  - The transition to deploying OpenSearch(OS) as the MA may also require more resources than our hosts are capable of delivering
  - Now that 5.2 is released, we will be exploring the testpoint container as a new best practice recommendation.
    - Our sites will get instructions on how to deploy and configure in the next few months

Future perfSONAR Deployment Option

The **perfSONAR Testpoint container** allows a simplified deployment model that is less resource intensive and easier to maintain and update.

- **Key Features and Challenges:**
  - Easy to deploy using Docker, e.g., docker run -d --name perfsonar-testpoint --net=host perfsonar/testpoint.
  - Concerns:
    - Potentially limited local cache to sustain central service outages.
    - Limited remote monitoring capability without additional features or packages.
    - Deployment docs, performance and integration with OS need development & testing (e.g., policy based routing for multi-NIC)
- **Future Development and Operations:**
  - We are developing instructions for OS bootstrapping, Docker-compose files, and OS configuration scripts.
  - Monitoring improvements rely on Prometheus perfSONAR node exporter and host exporter and these need testing and tuning.
  - Focus will be on stability, persistent configuration, caching results, and automated updates via cronjobs.

Target for new deployment: ready by end of summer 2025 (depends on pS developments)

Scitags Reminder: Description and Purpose

- Scientific Network Tags (scitags) is an initiative promoting identification of the science domains and their high-level activities at the network level.

- **Experiments/Collaborations** can better understand how their network flows perform along the path
  - Network performance studies, sites profiling
- **Research and Education Network Providers** gain visibility into sources and purpose of the network traffic
  - Enable tracking and correlation of network flows with REN systems
  - Facilitates debugging and capacity planning

SciTags Goals for the US

We have an IRIS-HEP **metric**, tracking the number of US sites identifying their traffic via SciTags (either Fireflies or packet marking).

For the next WLCG Data Challenge (DC27), we would like to have **at least 80%** of our traffic identified via SciTags.

USCMS has been leading here (6 sites instrumented) because Xrootd has supported sending SciTag fireflies since version 5.0 and USCMS heavily utilizes Xrootd

USATLAS uses dCache which hasn't yet provided a working SciTags firefly capability…but this is fixed in upcoming v 11.0.

- AGLT2 is testing a SciTags enabled version backported 10.2.13

SciTags Support: DDM & Storages

DDM:

- **Rucio** (from 32.4.0) and **FTS/gfal2** from 3.2.10/2.21.0
- Works for both **HTTP and XRoot** protocols (cf. technical specification)

Storages:

- **XRootD** provides Scitags implementation (from 5.0+)
  - In production at several sites
- **EOS** provides Scitags support from 5.2.19+
  - In production at CERN EOS ATLAS, CMS and ALICE
- **dCache** firefly support released in 11.0 replacing PoC introduced in 10.0
  - AGLT2 testing backported dCache with SciTags for version 10.2.13
  - for readout from XRoot and HTTP protocols underway
- **StoRM** provides Scitags support from 1.4.3+
  - In production at **CNAF** for all supported experiments

Collectors:

- Production deployments at ESnet, Jisc, in discussion with GEANT
- Site-collector running at CERN

SciTags Monitoring and Next Steps

ESnet has provided a dashboard for their global receiver:
https://dashboard.stardust.es.net/goto/HVw6rRLHg?orgId=2
As we increase coverage, we should see additional information show up there
We are working on “extending” the information we provide in the SciTag UDP “end” firefly

- We are adding information about the flow (total bytes, duration, etc) as well as harvesting and sending information from the network stack.
- We should discuss the kind of information we may want to collect
- As we determine what should be sent, we will work on improving our dashboards.

There is much more information about SciTags from the recent HEPiX presentation

WLCG Site Network Monitoring Update

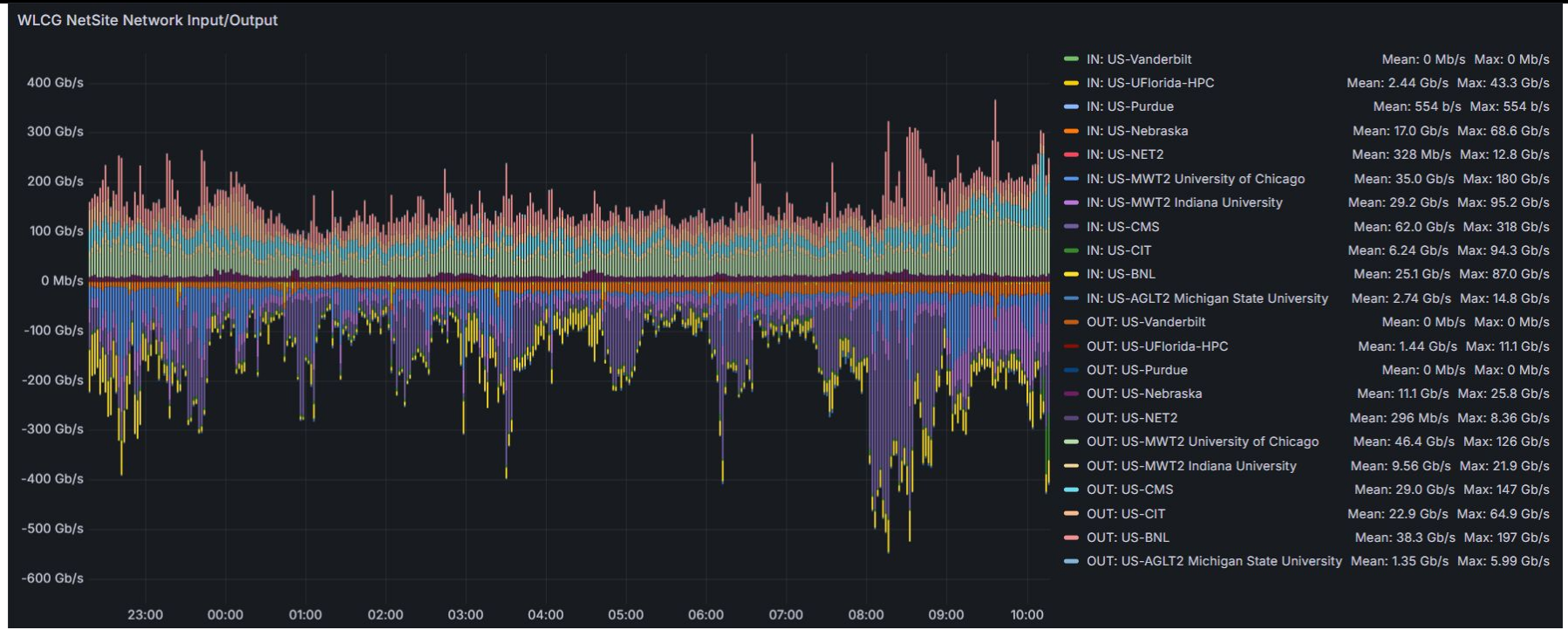
We ran a campaign prior to the WLCG Data Challenge 2024 to instrument as many of our sites as possible, capturing the total traffic IN/OUT

- We have the code and documentation in CERN Gitlab: https://gitlab.cern.ch/wlcg-doma/site-network-information
- Instructions for SNMP monitoring is at https://gitlab.cern.ch/wlcg-doma/site-network-information/-/tree/master/WLCG-site-snmp?ref_type=heads

Many of our US sites already provide monitoring data: http://monit-grafana-open.cern.ch/goto/htIB0gLHg?orgId=16 (see next slide)

Two issues: Not all sites covered and we are missing IPv4 vs IPv6 information

WLCG Site Network Monitoring for US (June 4 2025)



Missing SWT2,OU,MIT,UCSD,Wisconsin (others?)

Can we get these working well in advance of DC27?

WLCG Site Monitoring To-Do

We are planning to improve the WLCG site network monitoring to differentiate **IPv4** vs **IPv6**.

The current MIB we query for the traffic IN/OUT, also includes other OIDs that monitor IPv4 or IPv6 specific traffic.

We will be updating the SNMP tools to also output traffic IN/OUT for both IPv4 and IPv6

We will need to update the central CERN MONIT Grafana instance to capture and display this information.

Timescale: end of summer 2025? Sites will need to update their query script.

Additional Network Monitoring Topics

We have a significant effort underway to support analytics and alerting on the data we are collecting from the network (primarily perfSONAR metrics)

- Requests for particular capabilities ?
- See “backup slides” at end for more details

There is still work to be done in validating all the various monitoring data we are collecting to ensure we clearly understand what is measured.

- We need to make sure units and selection criteria are known for the data

What other monitoring ideas/topics should we discuss?

We can wait to cover this till the Discussion at the end of this session.

Summary

- Updates to perfSONAR and OSG/WLCG network measurement platform
  - perfSONAR 5.2 is out with new features required for our evolution.
  - New infrastructure monitoring and dashboards are (or will appear) in production.
  - Simplified deployment models are being evaluated and documented.
- The SciTags project is making progress
  - Goal is to instrument at least 80% of our network transfers via SciTags before DC27
- WLCG Site Network monitoring will update soon to provide IPv4/IPv6 visibility
- We have to continue to watch our network monitoring infrastructure as it is a complex system with lots of areas for issues to develop.

Question or Comments?

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

Useful URLs

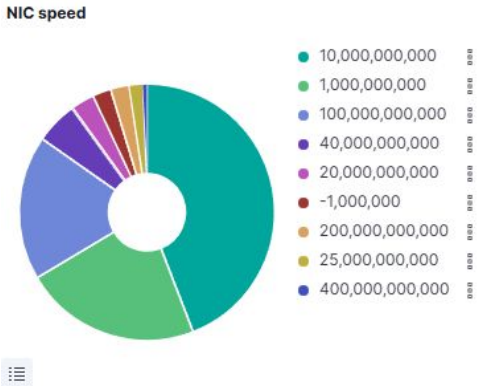
- OSG/WLCG Networking Documentation
  - https://opensciencegrid.github.io/networking/ (old, being updated soon)
- perfSONAR Infrastructure Dashboard
  - https://atlas-kibana.mwt2.org:5601/s/networking/goto/9911c54099b2be47ff9700772c3778b7
- perfSONAR Dashboard and Monitoring
  - http://maddash.aglt2.org/maddash-webui
  - (old) https://psetf.opensciencegrid.org/etf/check_mk
  - (new, beta) https://psetf-itb.aglt2.org/etf/check_mk
- Scitags Update from HEPiX https://docs.google.com/presentation/d/1Vdhs1Lc5asMKSccrFkiaBA0IluS9kl5PBoBQf6FdIRU/edit?slide=id.g8036819354_0_7#slide=id.g8036819354_0_7
- Toolkit information page
  - https://toolkitinfo.opensciencegrid.org/
- ATLAS Alerting and Alarming Service: https://aaas.atlas-ml.org/
- The pS Dash application: https://ps-dash.uc.ssl-hep.org/
- ESnet WLCG DC Dashboard: https://public.stardust.es.net/d/IkFCB5Hnk/lhc-data-challenge-overview?orgId=1

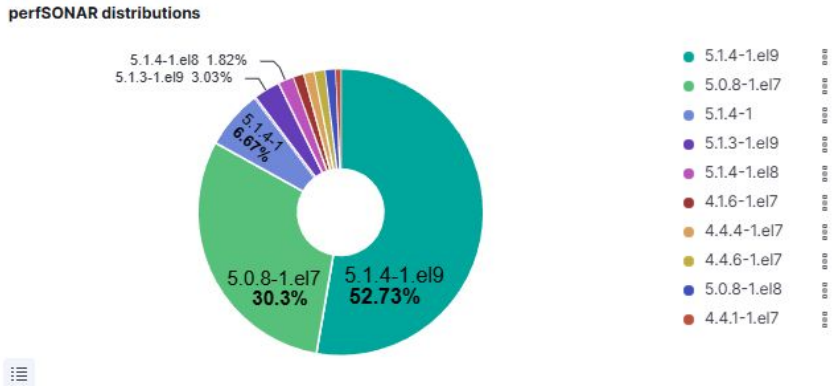
Backup Slides Follow

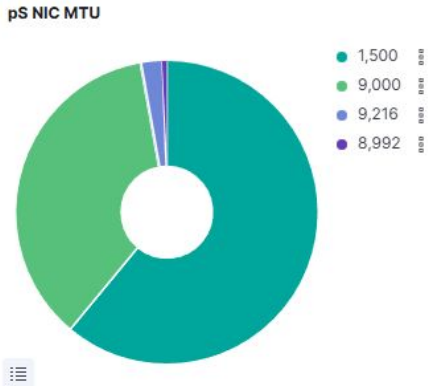
perfSONAR Infrastructure

175 Active perfSONARs

65 Communities in Use



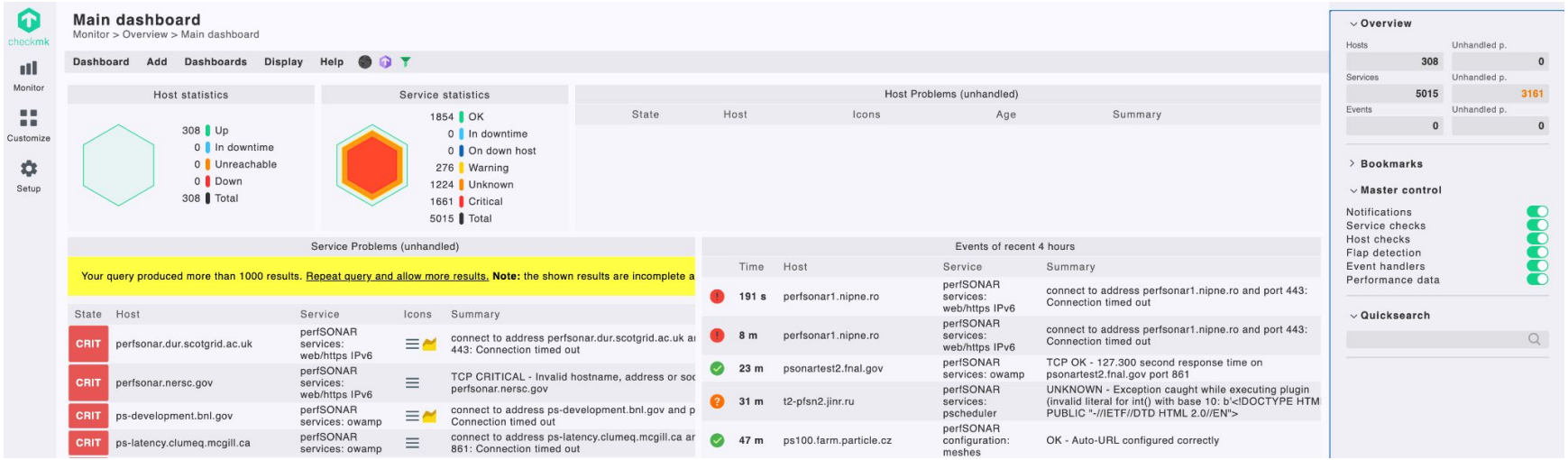




- 95.8% toolkits now on 5.x
  - Still about 4.2% on 4.x :(
  - 35% still on EL7 (all the 4.x and some of the 5.x)
- Core deployments are still on 10Gbps, but we have about 20% with 100Gbps
  - For WLCG/OSG testing purposes 10Gbps is still sufficient
  - We have some 200 Gbps and 400 Gbps hosts
  - **Important to refresh HW along with the update to EL9**
- MTU - around 36% on jumbo frames (9000), rest is on standard frames (1500)

perfSONAR Infrastructure Monitoring

- Updated to CheckMK 2.3.0 (from 1.6.0)
- Integration with CILogon (single-sign on) - moving away from x509 certs
- New test
  - Node diagnostics based on “pscheduler troubleshoot” command
- Now in pre-production at psetf-itb.aglt2.org (uses OSG CILogon registr.)
- Moving into production within the next week on psetf.aglt2.org



Measurements via Grafana

Our **MaDDash** has been replaced with **Grafana + ESnet plugins** that allow us to directly display data already gathered in our Central Measurement Archive.

(see beta version at https://maddash.aglt2.org/dashboards)

This is a much better situation than what we had with **MaDDash** because:

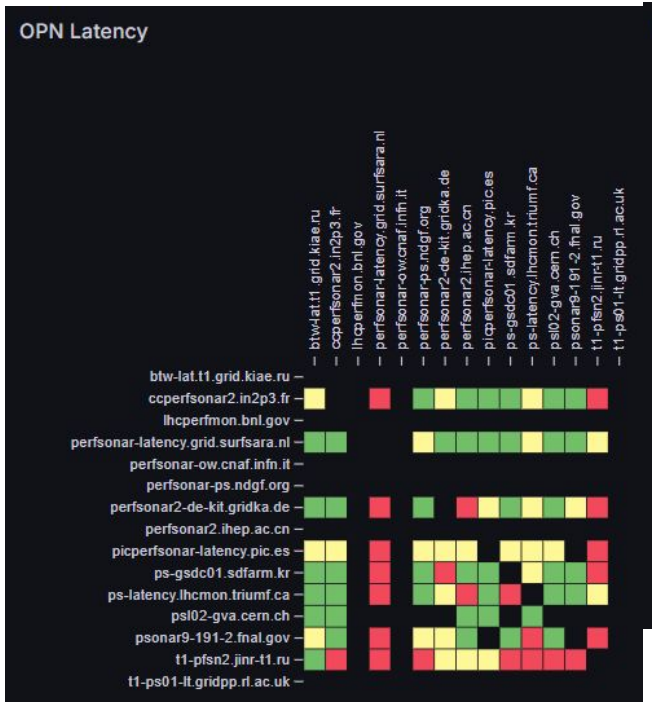
- No need to “re”gather data from each perfSONAR instance.
- No latency issues for the displayed data (previously up to tens of hours).
- Provides direct visibility for the centrally gathered data.

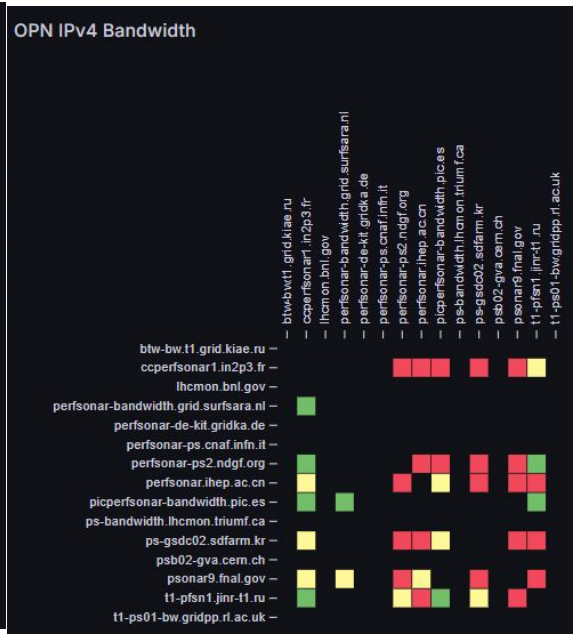
We do have some challenges to address

- The drill-down capabilities (clicking on a mesh box) needs work because our central data format differs from perfSONAR devs version
- We also need to change what data is gathered since we are missing some required metadata for some measurement types.

Examples for LHCOPN from MaDDash Replacement

Below are some examples of current LHCOPN meshes from the MaDDash replacement built upon the new perfSONAR developers Grafana+ESnet plugin





Analytics Summary

Network Analytics has been a long-term effort for our small group.

We want to better exploit the large complex network information we gather.

There are four general areas of work here:

- **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, Bayesian/probabilistic approaches
  - Detect anomalies in network paths and bandwidth measurements
  - Compare with the existing heuristic algorithms that we have developed
- **Correlation** 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.)

Analytics Core Team

We have a number of people who meet weekly to discuss network analytics and our measurement and analysis platforms.

The core group is comprised of Marian Babik, Shawn McKee, **Petya Vasileva** and Ilija Vukotic.

We often have others join our meetings from ATLAS, CMS, R&E networks, the perfSONAR team and site or network admins. (Notes available here)

Through **IRIS-HEP** (NSF funded institute) we have also managed to bring in a few fellows who have worked on specific projects identifying network issues or creating alerts and alarms. This year we have a returning IRIS-HEP fellow, **Yana Holoborodko** working with us at CERN till October.

Analytics Tools

We have developed tools to analyze and alert upon alarms identified.

Petya has given recent detailed presentation on this work at the February 2025 ATLAS S&C meeting at CERN.

The main user interface is pS-Dash (shown here =>)

This is a web GUI that serves as a front-end to our gathered data and generated alarms.

Explore it and let us know your comments/suggestions