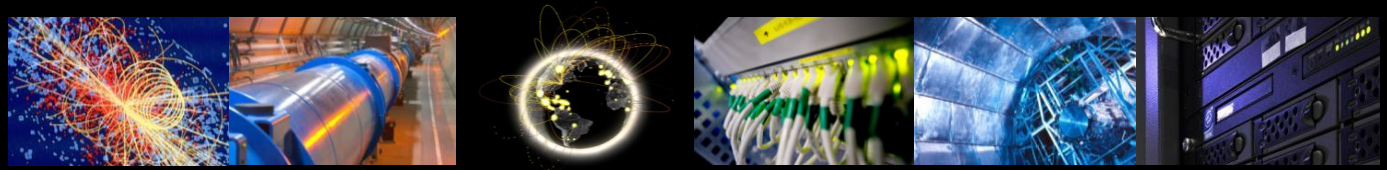
The SAND Project

Shawn McKee / University of Michigan
on behalf of the SAND and OSG Collaborations



WLCG Worldwide LHC Computing Grid

At the GEANT 2nd Performance Management Workshop - March 8, 2022

Open Science Grid

Outline

- Overview of WLCG/OSG Networking; The Motivation for SAND
- The SAND project, goals and outcomes
- Ongoing activities and continuing the SAND work

OSG/WLCG Networking Activities

- OSG is in its 9th year of supporting WLCG/OSG networking focused on:
  - Assisting its users and affiliates in identifying and fixing network bottlenecks
  - **Developing and operating a comprehensive Network Monitoring Platform**
  - Improving our ability to manage and use network topology and network metrics for analytics
- WLCG Network Throughput Working Group was established to ensure sites and experiments can better understand and fix networking issues:
  - Oversees the **OSG/WLCG perfSONAR infrastructure**
    - Core infrastructure for taking network measurements and performing low-level debugging activities
  - **Coordinates WLCG network performance incidents** - runs a dedicated support unit which involves sites, network experts, R&Es and perfSONAR developers
    - Many issues are potentially resolvable within the working group

The NSF SAND Project

SAND: Service Analysis and Network Diagnosis

This was a NSF funded project (award #1827116) focusing on combining, visualizing, and analyzing disparate network monitoring and service logging data. (**GOAL**: capitalize on our rich network dataset!!)

Website
https://sand-ci.org/
(Project started in **September 2018** and finished in **August 2021**)

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

Brian Bockelmann
Associate Scientist
Morgridge Institute for Research, University of Wisconsin
bbockelman@morgridge.org

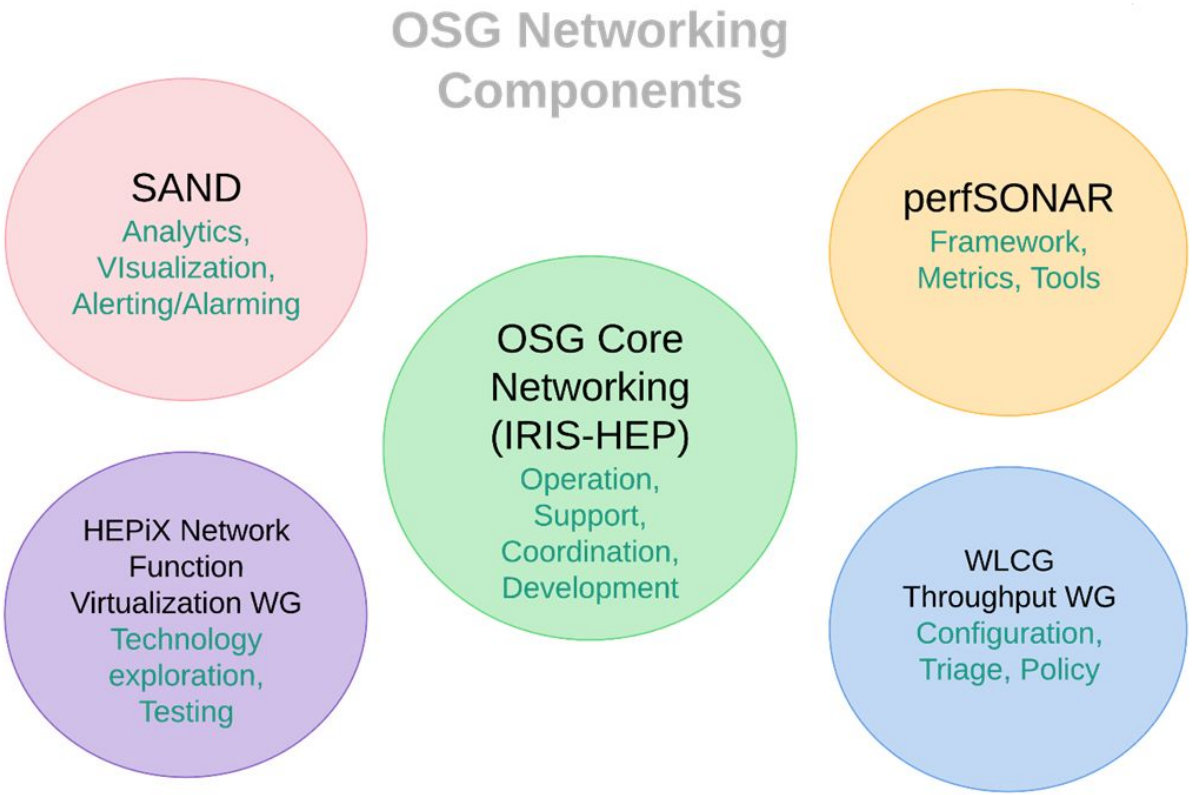
Shawn McKee
Research Scientist
University of Michigan
Physics
smckee@umich.edu

Rob Gardner
Senior Scientist
University of Chicago
Physics
rwg@hep.uchicago.edu

Context: OSG/WLCG Projects at SAND Start

There were 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

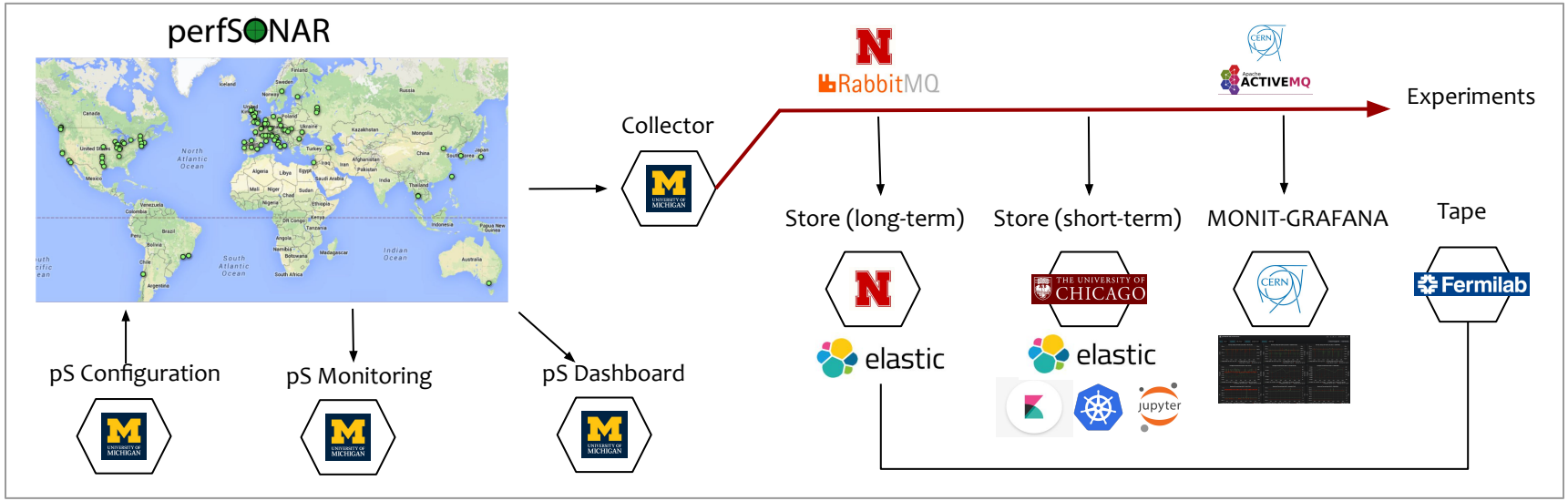


SAND Original Planning and Work Areas

- Main goal of SAND was to create new analytics, visualizations and user-interfaces to extract value from the perfSONAR (and related) network metrics
- **Initial architecture**: Data-pipeline to ELK stack, visualizations via Kibana, Grafana and perhaps other tools, analytics via Jupyter notebooks and creation of "architecture plugins" to leverage this framework.
  - Planned work areas:
    1. Alarming dashboards that show Top-N problem links (SRC-DEST with largest packet loss in last N hours, SRC-DEST with most routes in last N hours, SRC-DEST with largest change in measured throughput in last N hours, SRC with most average packet loss averaged over all DEST, DEST with most average packet loss averaged over all SRC)
    2. Route correlation: Identify SRC-DEST pairs with similar behavior changes at a point in time and analyze common hops in their routes
    3. Alerting system based upon alarming and route work. Users subscribe to various alerts using SRC, DEST, packet-loss, change in BW, etc

Network Measurement 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 ESNet, LHCOPN traffic also via same channels



SAND Outcomes

There are three primary outcomes that **SAND** provided:

1. **Hardening and evolution of the network metric pipeline**
   a. The pipeline was evolved to support use of a central RabbitMQ bus
   b. Data could be pushed from toolkits directly to the bus in addition to being pulled
   c. Added backup of the bus data to tape (and associated playback capabilities)
2. **Creation of an alerting and alarming system based upon metrics**
   a. One of the most requested features from sites that had deployed perfSONAR
   b. Still under very active development but we have a system in production
   c. Additionally we have provided new user interfaces to the underlying data
3. **Development of analytics and associated work in preparing for machine-learning on our data**
   a. Very challenging in that the data has significant noise and we need to work around various measurement tool shortcomings.
   b. Correlations between network path and metrics are central to the work
   c. Development of annotated data suitable for machine learning is the focus

Network Topology Challenges

Whenever we identify a possible network problem, the first question is: **what path is being measured**?

- Knowing the path in place when a problem is identified is critical
- Having many paths continuously monitored is a very powerful tool for both identify network issues and localizing them!
- Gedanken experiment: at approximately the same time, 5 host-pairs show an increase in packet loss. **What is the inference we can make by correlating their paths?**

Fortunately, we have regular "traceroute" tests between our perfSONAR measurement end-points but the data is LARGE and noisy.

A path visualisation tool was developed by our MEPHi SAND collaborators

- Video of first release available at https://yadi.sk/i/tyhiA-e3GGKqDQ
- Note: users need to limit the amount of data by time or by source or the tool becomes unresponsive!

The WLCG and OSG continue to work on identifying anomalous path data for alerting and alarming as well as to prepare data for machine learning

Alerting and Alarming

Much of the work during the SAND project was to create a system to allow users to be alerted when specific problems were identified.

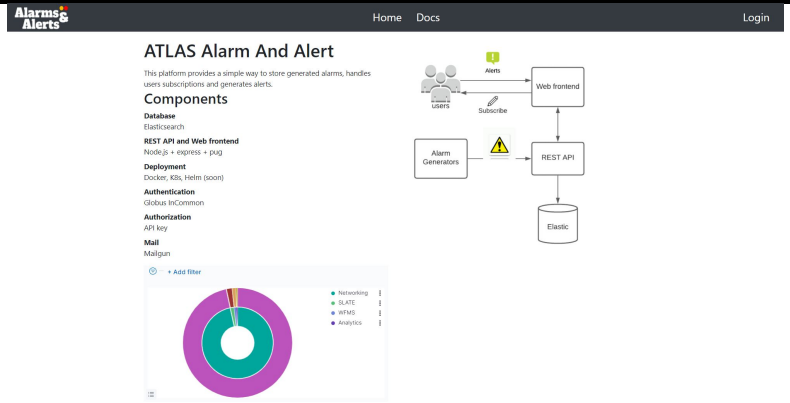
We now have an initial implementation of such a system in production at:

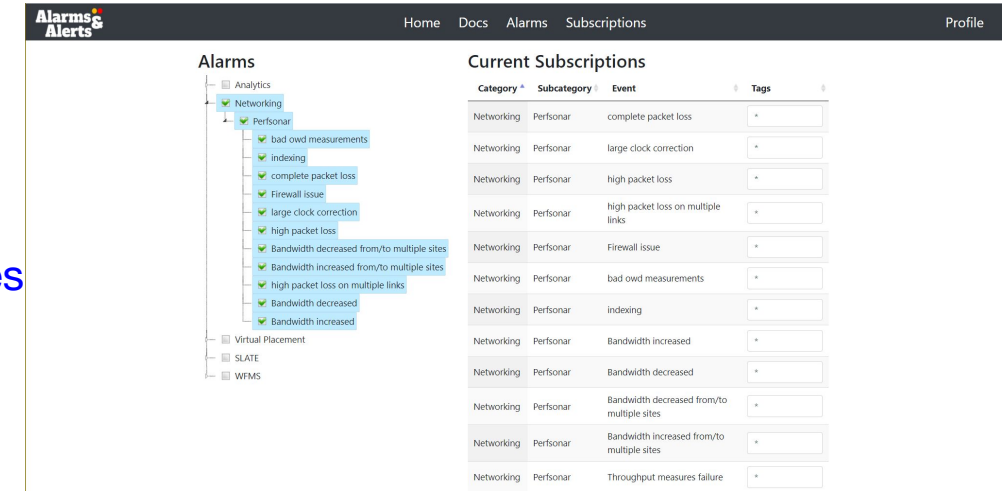
https://aaas.atlas-ml.org/

It supports Federated identity to login

Users can select and filter alerts

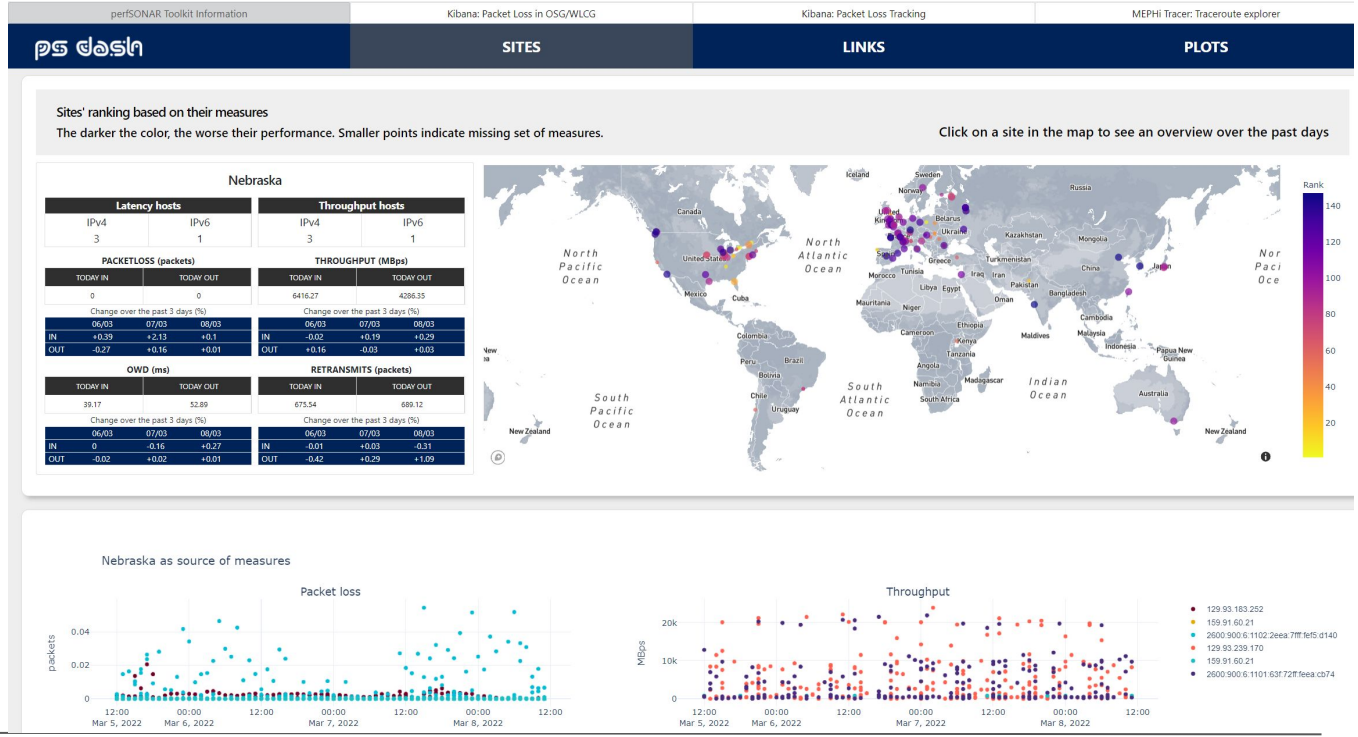
The IRIS-HEP/OSG-LHC team continues to develop and tune this application.





Developing Analytics

As part of the SAND work, we also created a Plotly based dashboard to develop and test our analytics and present the results.
This dashboard is available at: https://ps-dash.uc.ssl-hep.org/sites

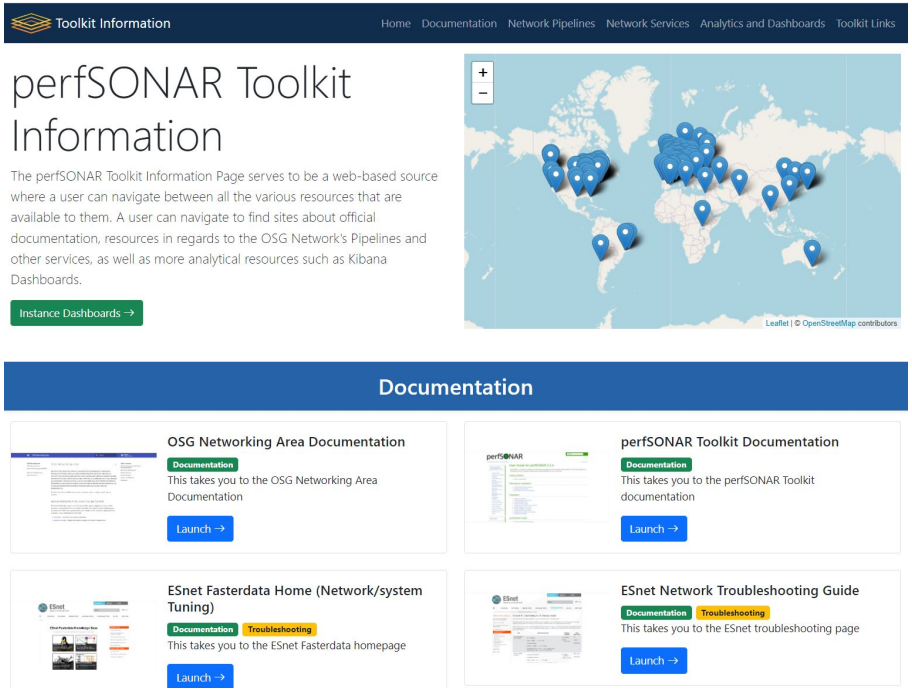


Finding Relevant Information: Toolkitinfo

The SAND project, WLCG and OSG have created a number of resources, making it difficult for end-users to find what they might need. We have set up a web server at: https://toolkitinfo.opensciencegrid.org/

Our goal is to continue to maintain and add-to the various menus available to allow a broad range of users to easily find and access network data and analytics results.

We welcome feedback and suggestions on this resource.



Continuing the Work

As noted, the OSG/WLCG communities are continuing to develop and build upon what SAND created.

A new proposal has been submitted to the NSF IMR program to apply **machine learning** to our data to provide new data sets for network researchers and specific problem identification for network users and administrators.

The Research Networking Technical Working Group is working to augment our network visibility, control and utilization by focusing on:

- Packet marking and flow labeling
- Traffic shaping and packet pacing
- Network orchestration and management

The WLCG Monitoring Task-force is working to get site networking monitoring incorporated into our global monitoring: primarily real-time IN/OUT by site.

The network pipeline is being evolved to leverage HTTP archivers -> ELK

SAND Summary

- The SAND project worked to:
  - Maintain an effective, efficient metrics pipeline
  - Provide an infrastructure to monitor our infrastructure and analyze various metrics
  - Extract new insights from measurements of our existing, complex global infrastructure.
- **The primary goal for SAND was to better extract "value" for our Scientists, Site and Network Administrators from the extensive network metrics OSG/WLCG is gathering.**
- The OSG and WLCG communities are continuing to work together to evolve what SAND started. We welcome your feedback and participation!

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:

- SAND: NSF OAC-1827116
- IRIS-HEP: NSF OAC-1836650

SAND and Network Links

- SAND webpage
  - http://sand-ci.org
- OSG/WLCG Networking Documentation
  - https://opensciencegrid.github.io/networking/
- perfSONAR Toolkit and network resources link
  - https://toolkitinfo.opensciencegrid.org
- perfSONAR Dashboard and Monitoring
  - http://maddash.opensciencegrid.org/maddash-webui
  - https://psetf.opensciencegrid.org/etf/check_mk
- perfSONAR Central Configuration
  - https://psconfig.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/

SAND Backup Slides

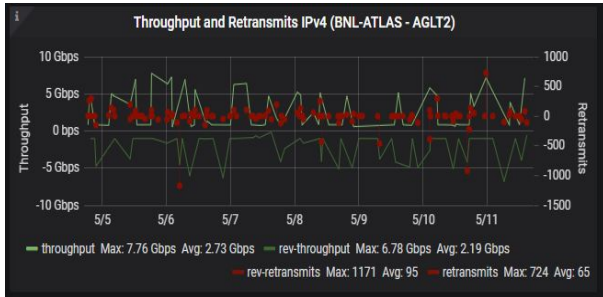
Grafana - perfSONAR dashboard

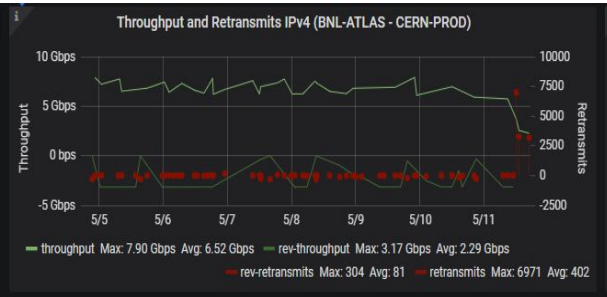
- Now has all WLCG sites that run perfSONAR
- Updated dashboards to support latest Grafana version

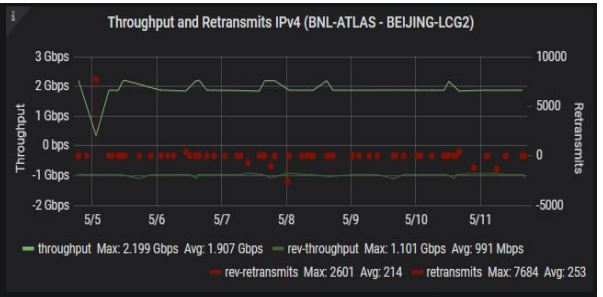
- Now includes all WLCG sites that run perfSONAR
- Added row that tracks RTT and number of hops as reported by traceroute/tracepath

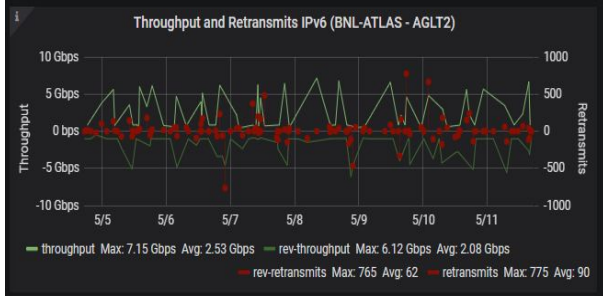
Grafana - IPv6 dashboard

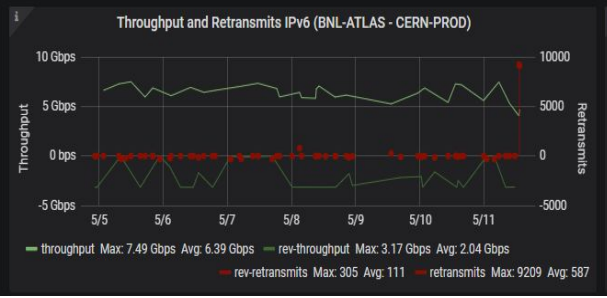
- Added IPv6 dashboard
  - Side-by-side comparison btw. IPv4 and IPv6 performance
- Due to performance limitations it was agreed that won't configure IPv6 latency tests

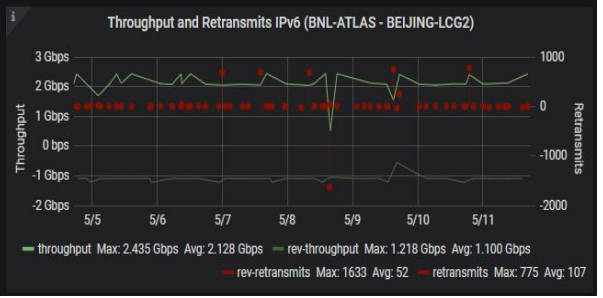












See more Grafana dashboards at http://monit-grafana-open.cern.ch/

Available Data Overview

SAND and OSG/WLCG are gathering a number of potentially very useful metrics

- **perfSONAR** data from over 260 instances all over the world
- **ESnet** network traffic (snmp counters)
- **WLCG** data transfers (FTS)
- **LHCOPN** data (from CERN networking)

This data is being transferred using message bus technologies (RabbitMQ (OSG) and ActiveMQ (CERN)) and ends up in two different Elasticsearch instances (University of Chicago analytics platform and University of Nebraska)

This data could provide powerful insights into our R&E network infrastructure by using the **temporal** and **spatial** information we have available.

Some Context: IRIS-HEP

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 starting 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 created 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/

perfSONAR Data Details

We are collecting a number of different types of data from perfSONAR which are sent to different “topics” on the RabbitMQ bus and put into their own index in Elasticsearch:

- **ps_alarms** : These are generated alarms based on other ps indices
- **ps_meta :** Tracks toolkit version, host info, various metadata
- **ps_owd :** One-way Delay measurements from perfSONAR (latency)
- **ps_packet_loss :** The percentage of packets lost in latency testing (10 Hz)
- **ps_retransmits :** During throughput testing, tracks retransmits
- **ps_status :** Tracks status of measurements (coverage, efficiency)
- **ps_throughput :** Measures throughput via iperf
- **ps_trace :** Measures the layer-3 network path via traceroute

You can explore the details via Kibana:

https://atlas-kibana.mwt2.org/s/networking/app/kibana#/discover?_g=()

SAND Collaboration Meeting Details

Our first in person collaboration meeting was June 17-18, 2019 at U Chicago

Main topic areas discussed day 1

- Network pipeline
- Monitoring tools
- Containerizing perfSONAR
- Engaging with and enabling a broader community
- Topology and data cleaning

The second day was a “hackathon” were we worked on items from day 1.

The “**Team**”
Picture credit: **Rob Gardner** (that’s why he’s missing)

Issues with Traceroute and Network Paths

While we regularly try to measure the network paths between our hosts (and by proxy, between our sites), the traceroute tools has some limitations

- It sometimes doesn't reach the destination
- Hops along the way can fail to respond in time, leaving "holes" in the path
- The trivial variations in traceroutes can lead to 10's of thousands of routes
- The "route" it delivers can be false
  https://www.cellstream.com/reference-reading/tipsandtricks/403-ecmp-linux-paristr

For all these reasons, we have **challenges** in trying to use our traceroute results to understand the network topology

The SAND project is planning to work on cleaning things up

- We are trying to identify logical paths to contain trivially varying physical paths to simplify things
- We need to identify when multiple links might exist at L2
- We have added "AS" number to the traceroute data to simplify understand when a major route change happens.
- We are working on ways to visualize, compare and understand our network paths