Progress and Plans in OSG Networking

July 11, 2023

HTC 23, Madison, Wisconsin

Shawn McKee reporting for the extended team: Marian Babik, Derek Weitzel, John Thiltges, Ilija Vukotic and Petya Vasileva

Why Networking in OSG?

Distributed science collaborations critically depend upon the network to interconnect and enable geographically dispersed compute and storage resources to work together.

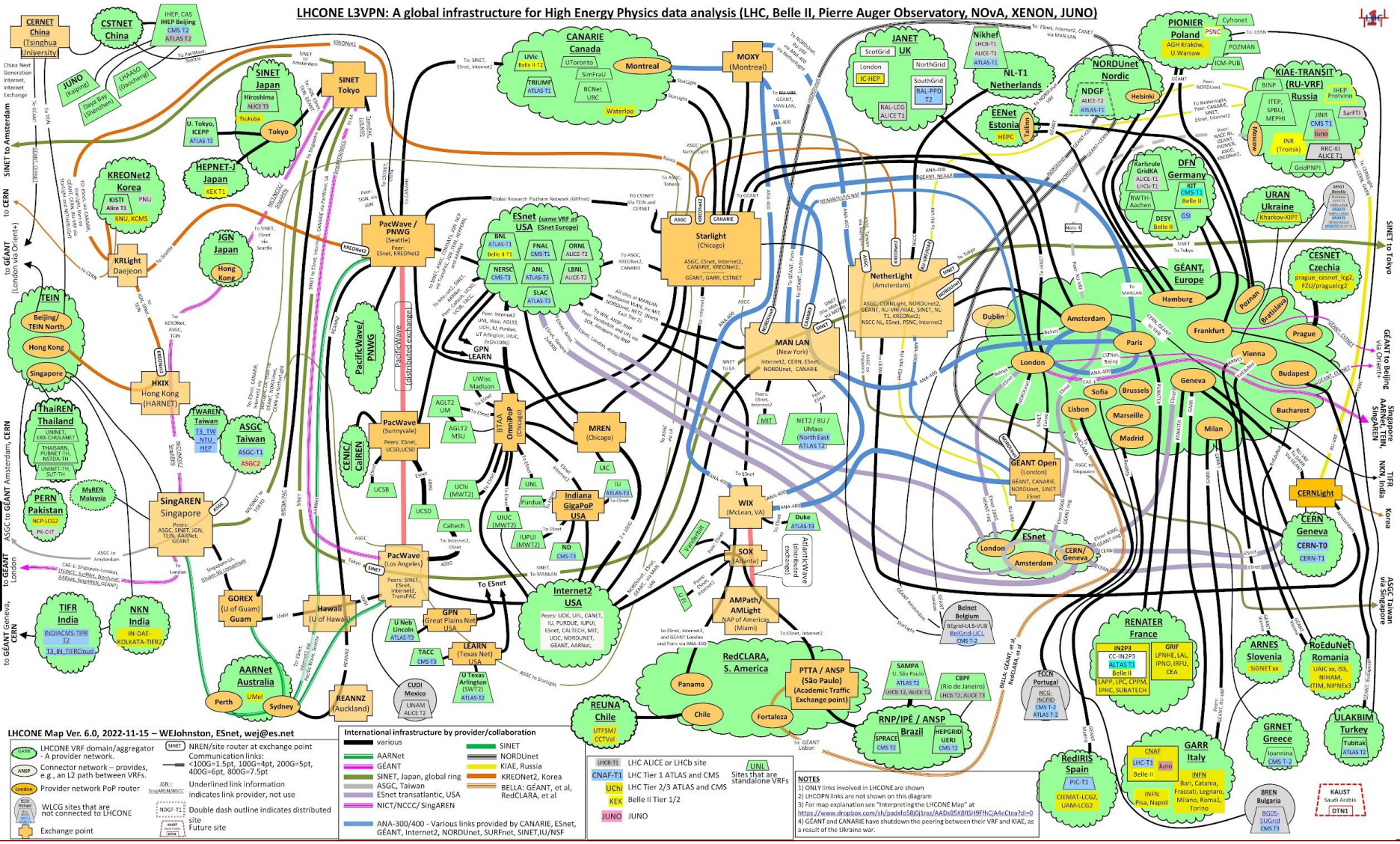
Network problems can significantly impact OSG users, slowing or halting their research.

The “network” is, in almost all cases of interest, actually a set of inter-operating networks that provide end-to-end connectivity for their users; but there is no single entity that owns, configures and maintains any given network path.

- Thus “problems” in the network can be exceedingly difficult to fix
- Sometimes “network” problems are actually end-system or application issues

OSG networking has focused on measuring our networks to identify and, to the extent possible, localize network issues or rule them out.

Example: LHCONE Network



Map Credit: [Bill Johnston]

OSG and Networking

OSG has been working in network monitoring and network performance for more than 10 years and we have developed an extensive infrastructure and associated tools for network monitoring and analysis including:

- Over 230 globally deployed **perfSONAR** toolkits at compute/storage sites
- A framework to monitor, manage and configure toolkits to support multiple virtual organization's testing requirements
- A network data pipeline to centrally gather all measurements
- OSG/WLCG specific documentation and guidelines
- An alerting & alarming framework to allow self-subscription to alarms
- An application to visualize and explore network alarms
- An analytics platform to explore our extensive datasets and look for more effective ways to identify & localize network and infrastructure issues

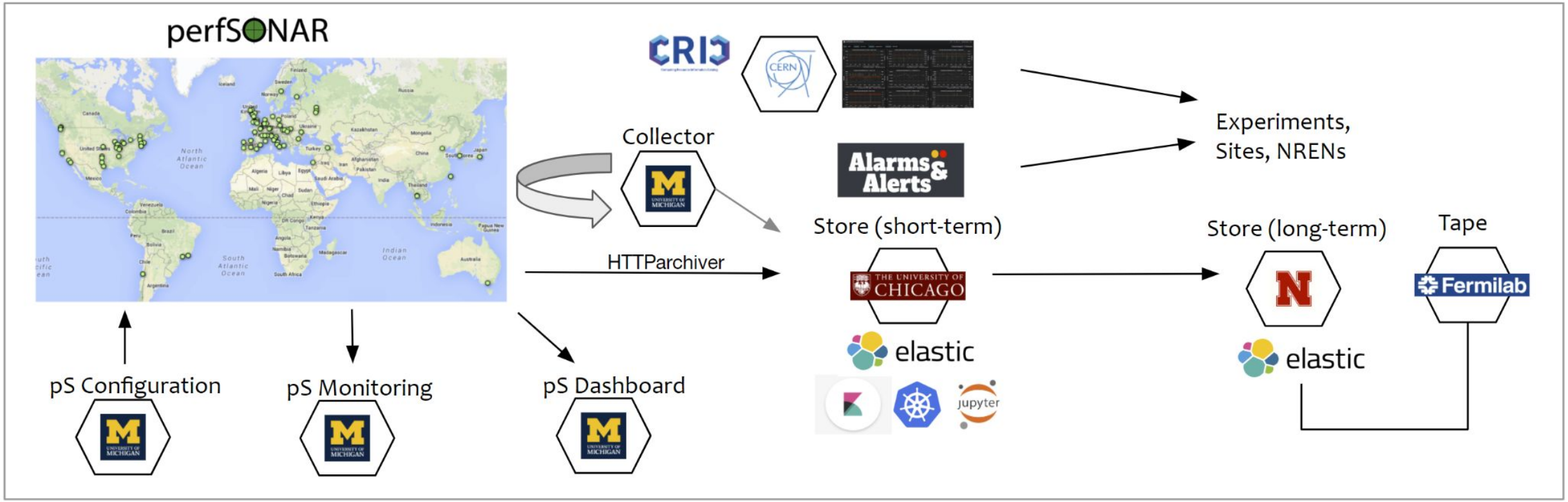
OSG Network Pipeline

- The OSG network pipeline gathers approximately 6 TBs of network metrics and various other data per year.
- The current system has a central collector which gathers data from perfSONAR toolkits and other data sources and sends data to a message bus.
- Another part of the pipeline ingests data from the message bus and augments it with additional meta data before inserting it into Elasticsearch
- With release 5 of the perfSONAR toolkit we are implementing a transition from the current “pull” model for the pipeline to a new “push” model where toolkits send directly to Logstash / Elasticsearch
  - This reduces complexity and components that we have develop and maintain
  - It also significantly reduces latency between measurements and when they are available in Elasticsearch, supporting near real-time use-cases.

OSG/WLCG Network Pipeline Transition

The transitional version of the pipeline is shown in the graphic below and features and updated Collector and use of the HTTParchiver to get data into Elasticsearch.

- The default is still using RabbitMQ and our “Ingestors”
- Broad v5.x deployment will allow us to remove RabbitMQ and retire ingestors



Alarms & Alerts Service

Components

Database
Elasticsearch

REST API and Web frontend
Node.js + express + pug

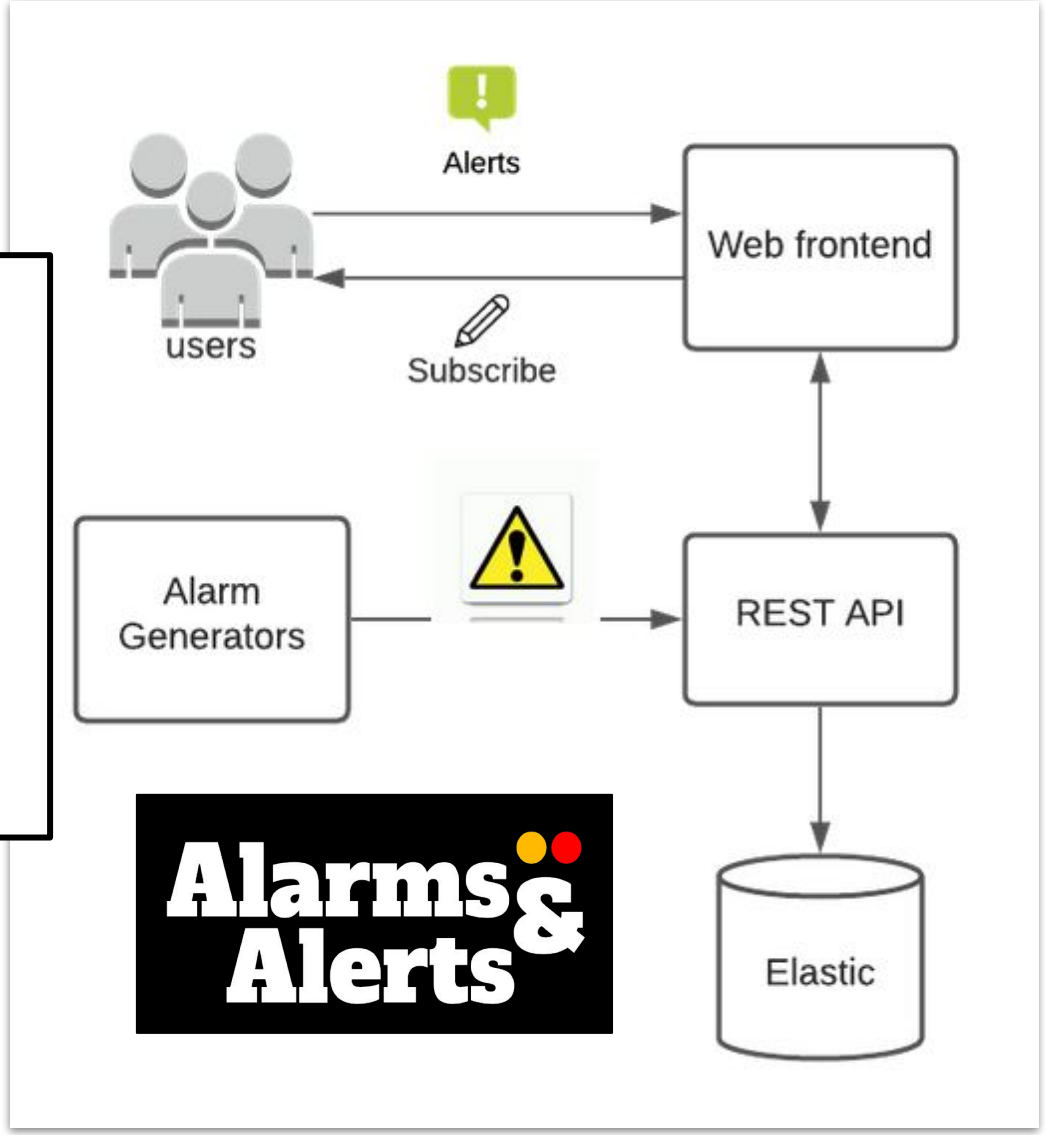
Deployment
Docker, K8s, Helm (soon)

Authentication
Globus InCommon

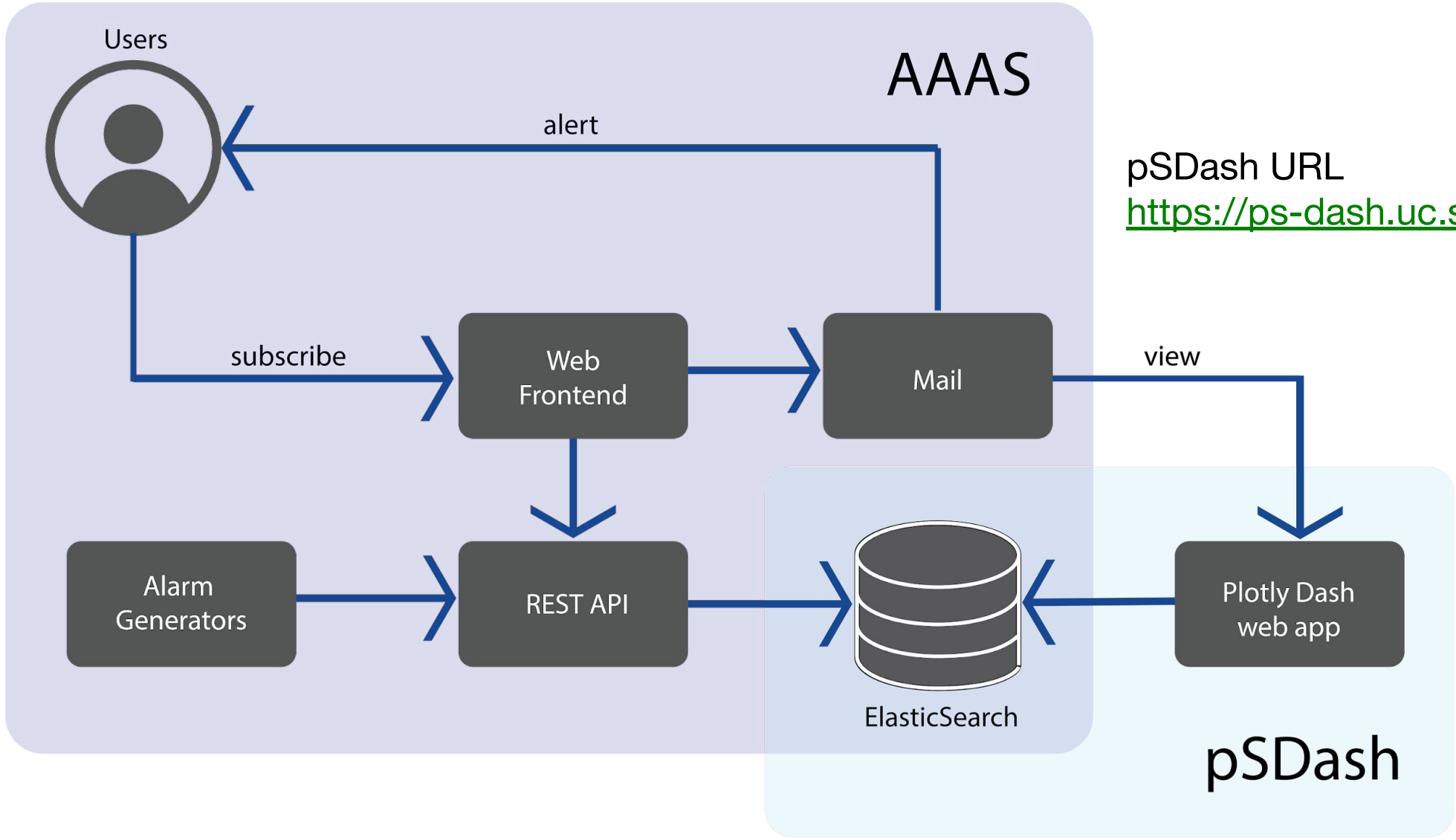
Authorization
API key

Mail
Mailgun

https://psa.osg-htc.org
(Uses EDUGain/InCommon)
Purpose: provides user-subscribable alerting for specific types of network issues found by analyzing perfSONAR data



Components: AAS & pSDash



pSDash URL
https://ps-dash.uc.ssl-hep.org/

OSG Example Data Analysis

Correlating Tests with Paths: Two Timescales

Our perfSONAR infrastructure runs many varied tests

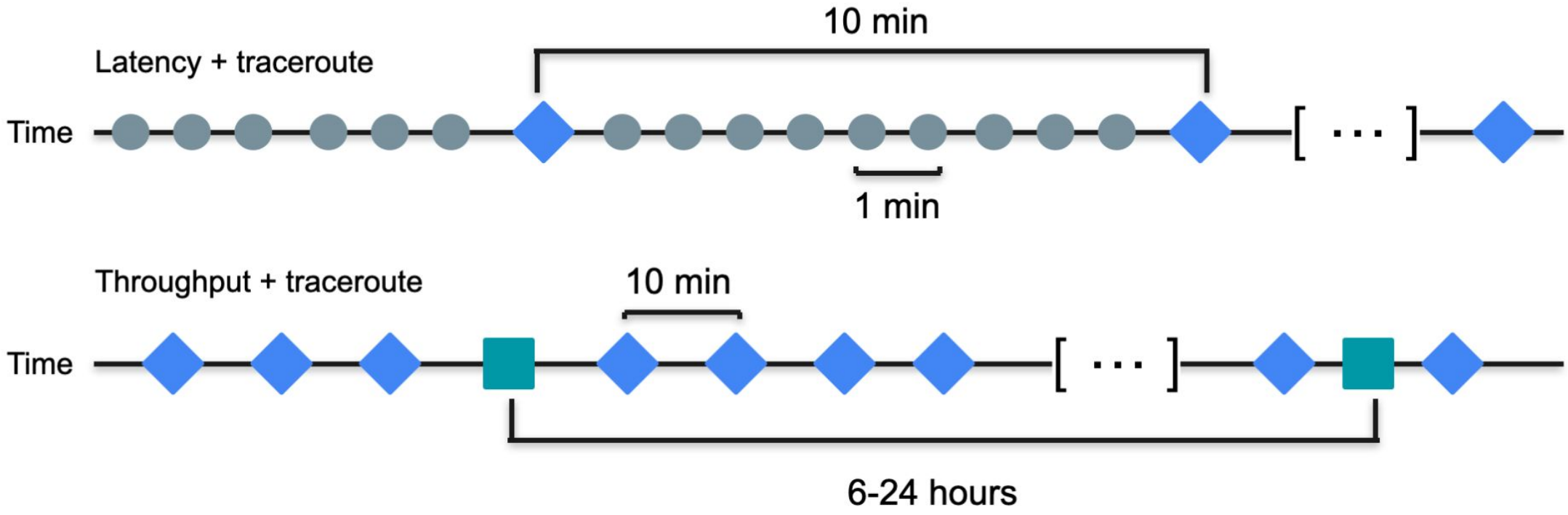
Latency tests: **every minute**

Throughput tests: **6-24 hours**

Traceroute tests: **every 10 min**

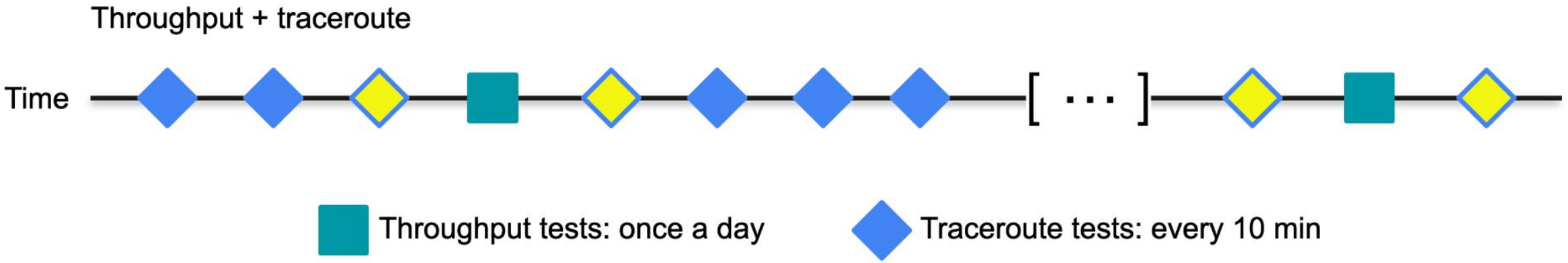
The problem: All tests run independently

How do we combine them?



Connecting Throughput to Traceroute

Can we consider the trace routes closest in time to the throughput records?

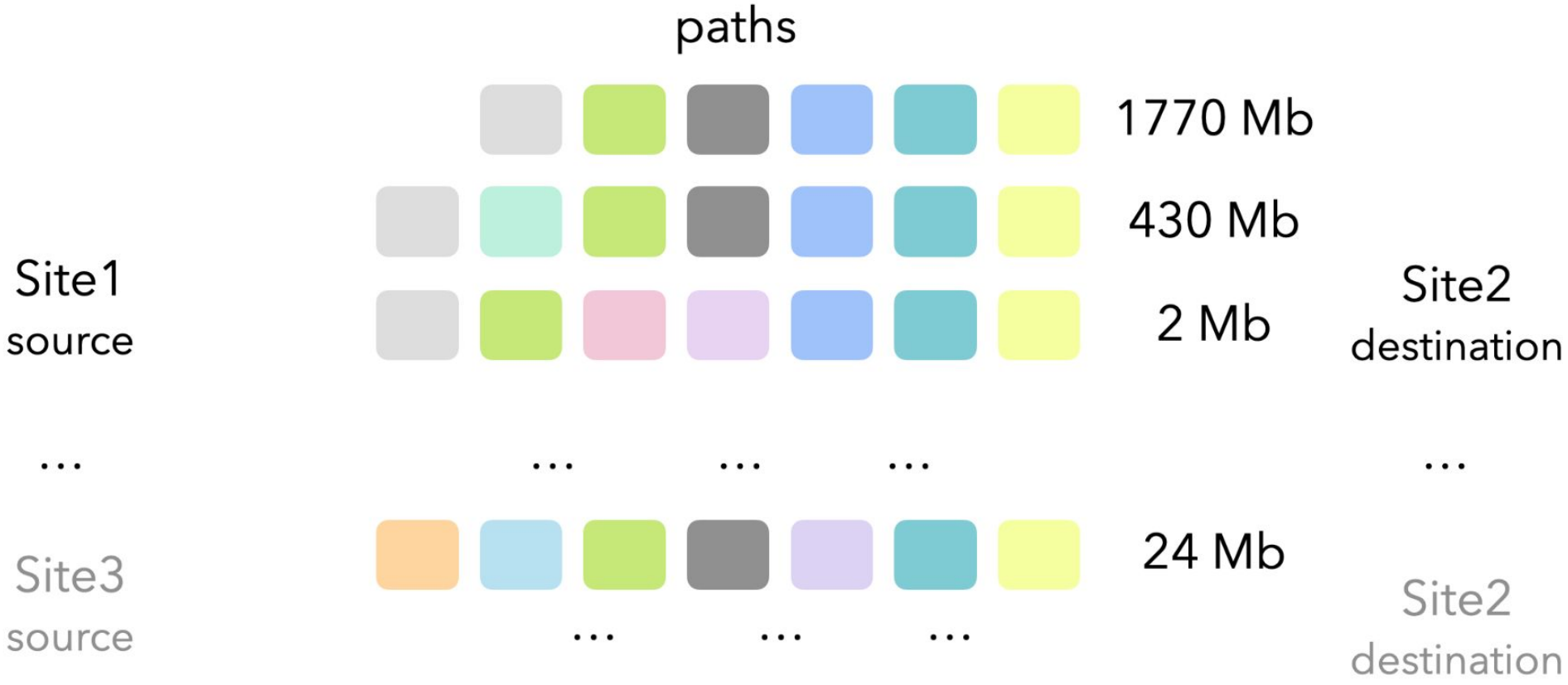


Our starting choice: Use **both** tracepaths (just before; just after) as valid paths and attribute BW to both.

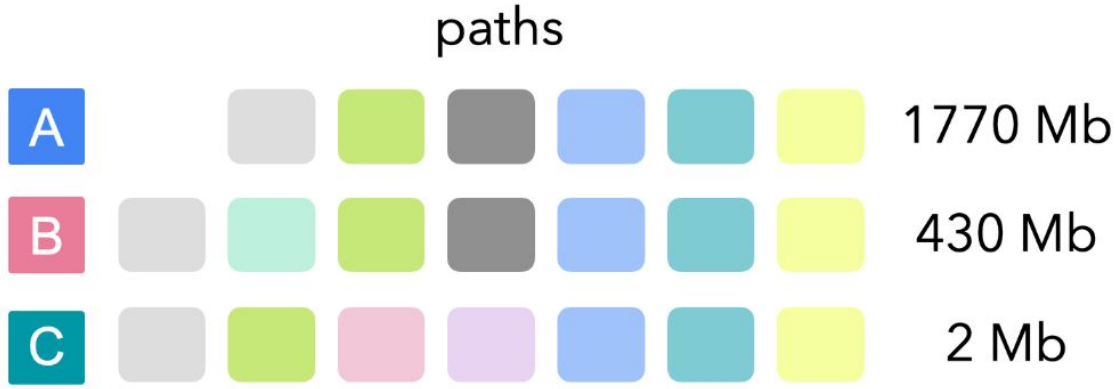
Evaluating to see if this is superior to just using the last measured route before the measurement…

Attaching Throughput Results to Sets of Routers/Links

Each colored box represents a specific router along the path



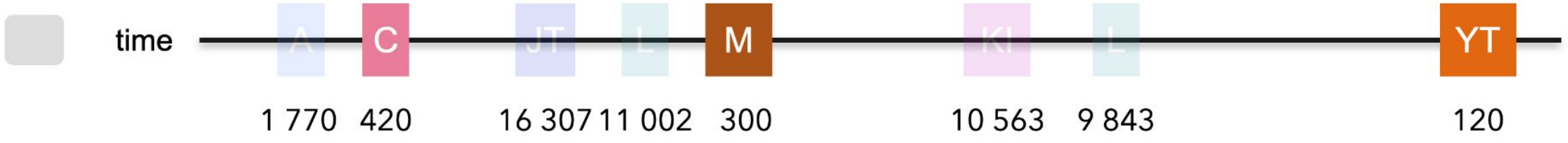
Example Throughput Attribution by Router



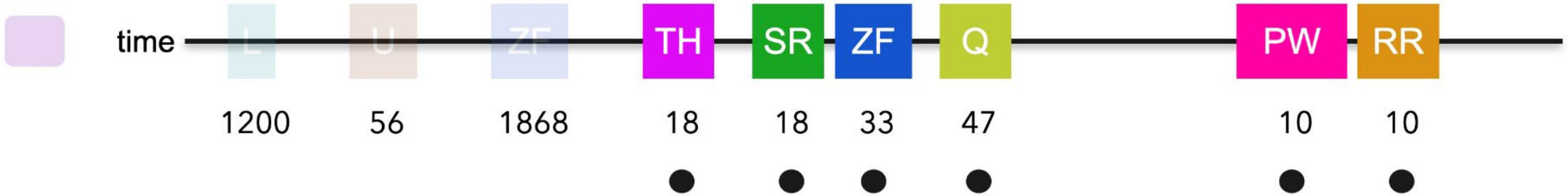
Each router on the path gets the closest (in time) throughput values

Checking Router Results vs Time

Max bandwidth seen: 16 307 Mb



Max bandwidth seen: 1868 Mb



Look for a down trend

(threshold below 10% of the max throughput)

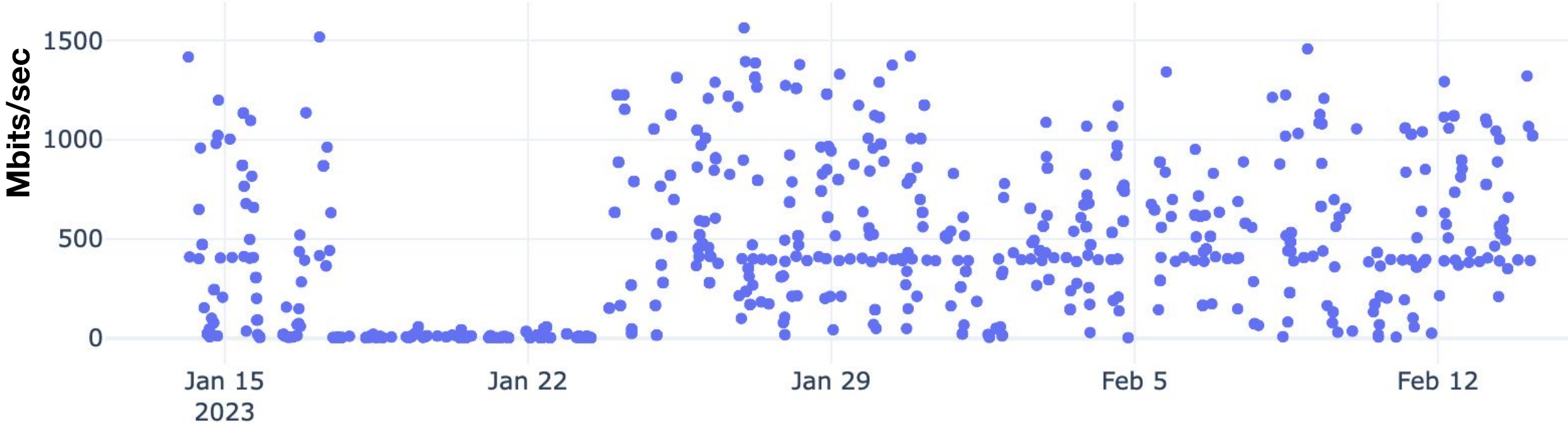
Is everything OK with/around that router?

Example Result: One Router; Throughput vs Time

Router IPv6 Address

2001:630:0:9011::189

Each **point** represents the throughput values collected when the router was on the path



This is a very powerful diagnostic tool because it localizes a problem to a specific link in the path!

Ongoing Network-related Collaborations

RNTWG (Research Networking Technical Working Group)

Packet marking / flow labeling, packet pacing, orchestration

Network user-facing tools (pSDash, AAS, Toolkitinfo, Kibana, Grafana,TRACer)

Network analytics+IETF ALTO WG is exploring our data to identify & localize issues

perfSONAR v5 the next version of perfSONAR is out; upgrade campaign now

WLCG SOC exploring best practices for network monitoring of attacks/malware

WLCG DOMA (Data Organization, Management and Access) includes networking

- **WLCG DOMA BDT** Bulk Data Transfer working group has network component

WLCG Monitoring Task Force has activities to make site network more visible.

Driver: Data Challenges preparing for HL-LHC; network data challenges ~2 years

- Focus is on DC24 scheduled for Feb 26-Mar 8, 2024 at 25% of HL-LHC scale

Near Term Work Areas

- Evolve pSDash, improving problem identification and localization
  - Interest from WLCG Operations in Tier-1 alarms related to network.
- Continue network topology cleaning, analyzing and alerting.
- Exploring the value of network flow labeling and packet marking (with RNTWG)
  - Developing new perfSONAR tests for flow-labels and/or packet marking
- Increase deployment of **site specific monitoring** (from CRIC URLs)
  - Campaign to get largest sites deployed starting in summer 2023
- Finishing our pipeline by transition by mid summer
- Maintaining the engagement with the Global research community
  - Access to R&E monitoring and integration with our tools and datastores
- Creating new dashboards with site data, augmenting existing dashboards.
- Participating in the planning and execution of mini-data challenges to prepare for second WLCG Network Data Challenge
  - New table of sites, dates and targets being developed

Summary and Conclusion

The **collaboration of OSG, WLCG and various research projects** have created an extensive, reliable infrastructure to monitor our networks via **perfSONAR** and provide associated analytics and visualization.

- We are engaged with various groups, working on improving networking for scientific research.
- While we continue to monitor and maintain our infrastructure, we need to also develop and tune our tools and applications.
- There are a number of challenges remaining, including the difficult continuing meta-challenge of improving our tools for our end users.

Questions 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

Other Presentations for More Detail

perfSONAR Global Monitoring and Analytics Framework Update at Spring 2023 HEPiX https://indico.cern.ch/event/1222948/contributions/5321010/

Status and Plans for the Research Networking Technical WG at Spring 2023 HEPiX https://indico.cern.ch/event/1222948/contributions/5321009/

Identifying and Understanding Scientific Network Flows at CHEP 2023 https://indico.jlab.org/event/459/

Tools and methods for tracking network issues based on perfSONAR datasets at ESnet CI Lunch & Learn YouTube https://youtu.be/tOv8EMwmKb0

Useful URLs to Reference

- Current OSG network documentation
  - https://osg-htc.org/networking/index.html
- IRIS-HEP network page https://iris-hep.org/projects/osg-networking.html
- pS Toolkit Info Page: https://toolkitinfo.opensciencegrid.org/
  - You can find lots of pointers to information and tools here
- The psetf CheckMK pS monitoring page: https://psetf.opensciencegrid.org/etf/check_mk/index.py?start_url=%2Fetf%2Fcheck_mk%2Fdashboard.py
- The RNTWG has information at https://docs.google.com/document/d/1I4U5dpH556kCnoIHzyRpBI74IPc0gpgAG3VPUp98lo0/edit#heading=h.nxusujd3hvwy

Backup Slides

Research Networking Technical WG (RNTWG)

This working group is focused on some specific, practical network efforts:

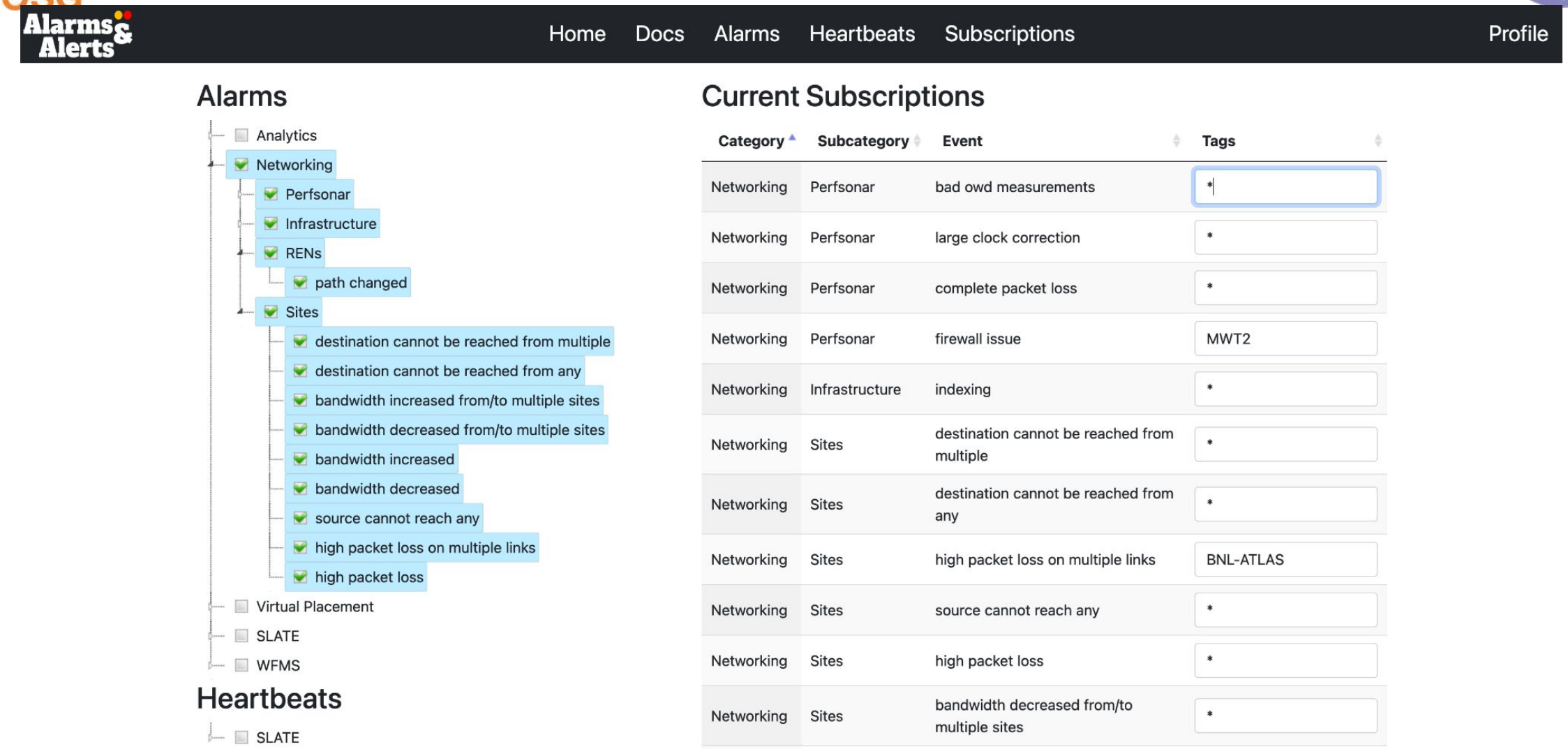
1. **Network visibility** via Packet Marking / Flow Labeling
2. **Network usage optimization** via Packet Pacing / Traffic Shaping
3. **Network management** via Network Orchestration / GNA-G DIS / SENSE / NOTED

Charter for the main group is at https://zenodo.org/record/6470973#.YmamPNrMJD8

Are meetings are available in Indico: https://indico.cern.ch/category/10031/

To undertake the above efforts we have created three subgroups looking into each of the areas above.

The Alerting and Alarming Tools Subscription Interface



Network Visibility and Scitags

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

- Enable tracking and correlation of our transfers with Research and Education Network Providers (R&Es) network flow monitoring
- **Experiments** can better understand how their network flows perform along the path
  - Improve visibility into how network flows perform (per activity) within R&E segments
  - Get insights into how experiment is using the networks, get additional data from R&Es on behaviour of our transfers (traffic, paths, etc.)
- Sites can get visibility into how different network flows perform
  - Network monitoring per flow (with experiment/activity information)
    - E.g. RTT, retransmits, segment size, congestion window, etc. all per flow

Finding More Information: https://scitags.org

Code

Technical Spec

Mailing List

Presentations

scitags.org

Network Flow and Packet Marking for Global Scientific Computing

View On GitHub | Download Tech. Spec | Join scitags.org

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

It provides an open system using open source technologies that helps *Research and Education (R&E) providers* in understanding how their networks are being utilised while at the same time providing feedback to the *scientific community* on what network flows and patterns are critical for their computing.

Our approach is based on a network tagging mechanism that marks network packets and/or network flows using the science domain and activity fields. These tags can then be captured by the *R&E providers* and correlated with their existing netflow data to better understand existing network patterns, estimate network usage and track activities.

The initiative offers an **open collaboration on the research and development of the packet and flow marking prototypes** and works in close collaboration with the scientific storage and transfer providers to enable the marking capability. The project is currently in the prototyping phase and is open for participation from any science domain that require or anticipate to require high throughput computing as well as any interested *R&E providers*.

Participants

ESnet GÉANT INTERNET2 RNP Jisc

XRootD dCache OFTS RUCIO

NORDUnet STARLIGHT OSG

Upcoming and Past Events

- March 2022: LHCOPN/LHCONE workshop
- November 2021: GridPP Technical Seminar (slides)
- November 2021: ATLAS ADC Technical Coordination Board
- October 2021: LHCOPN/LHCONE workshop (slides)
- September 2021: 2nd Global Research Platform Workshop (slides)

Hosted on GitHub Pages — Theme by orderedlist

Collaborative Work with WLCG Monitoring TF

The WLCG Monitoring Task Force is about to undertake a campaign to instrument our larger WLCG site networks with simple monitoring

The network related work within the task force is to get more information about WLCG site's, their network details and traffic flows IN and OUT.

We are currently finalizing our site instructions and plan to start the campaign in May

Code/instructions on GitLab: https://gitlab.cern.ch/wlcg-doma/site-network-information

Monitoring (Need CERN/Grafana access)