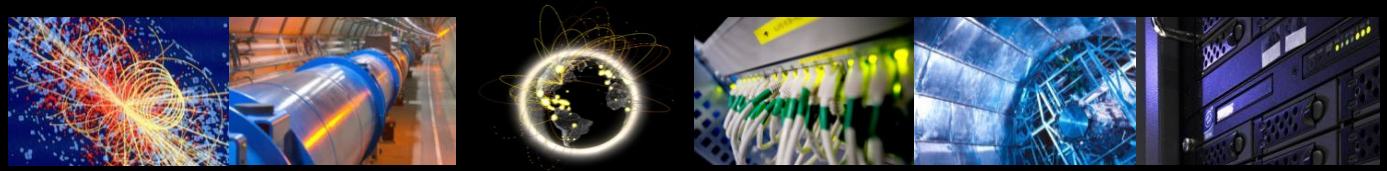
perfSONAR Monitoring Update

Shawn McKee / U Michigan, Marian Babik / CERN
on behalf of WLCG Network Throughput WG

2024

At the #52 LHCOPN/LHCONE Meeting, Beijing, China

https://indico.cern.ch/event/1410638/



WLCG Worldwide LHC Computing Grid

Open Science Grid

Outline

- News and Updates
- Network Measurement Platform
- New Infrastructure Monitoring (psetf)
- New dashboards (maddash)
- Analytics

perfSONAR News

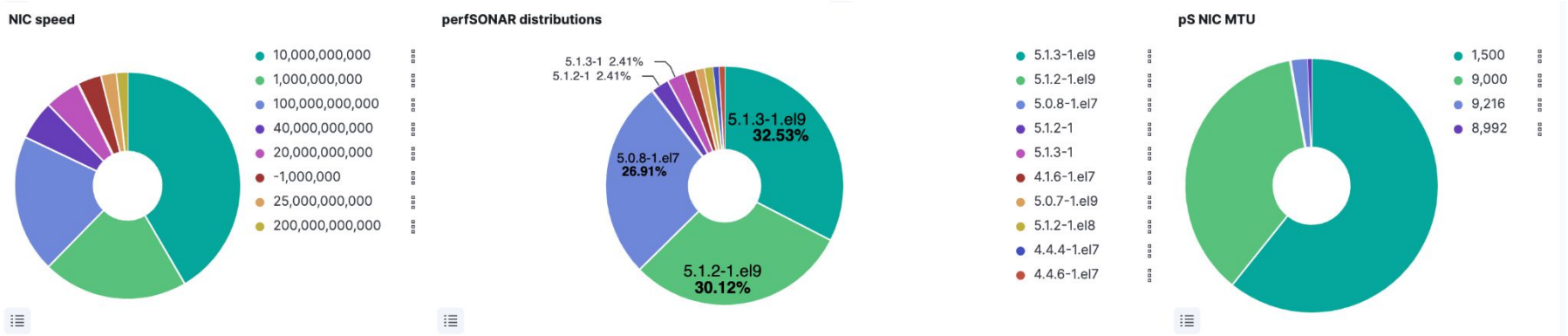
- **perfSONAR 5.1.3** is the latest release
  - Number of bug-fixes since 5.0; bi-weekly meetings with the developers
  - Update campaign in WLCG
    - Various issues, mostly archiving, but also e.g. legacy limits configuration (fix)
  - Toolkit support for latest Alma/Rocky 8 and 9, Debian 11/12, Ubuntu 20/22
    - **CentOS7 is no longer supported, sites should update ASAP**
- **perfSONAR session at LHCOPN/LHCONE tomorrow**

3:45 PM	**perfSONAR News and Updates (TBC)** **Speaker**: Szymon Trocha (Poznan Supercomputing and Networking Centre (PCSS))	20m
4:05 PM	**Experiences with perfSONAR in Jisc/UK** **Speaker**: Tim Chown	20m
4:25 PM	**perfSONAR Network Analytics through Machine Learning** **Speaker**: Petya Vasileva (University of Michigan (US))	20m

perfSONAR Infrastructure

181 Active perfSONARs

68 Communities in Use

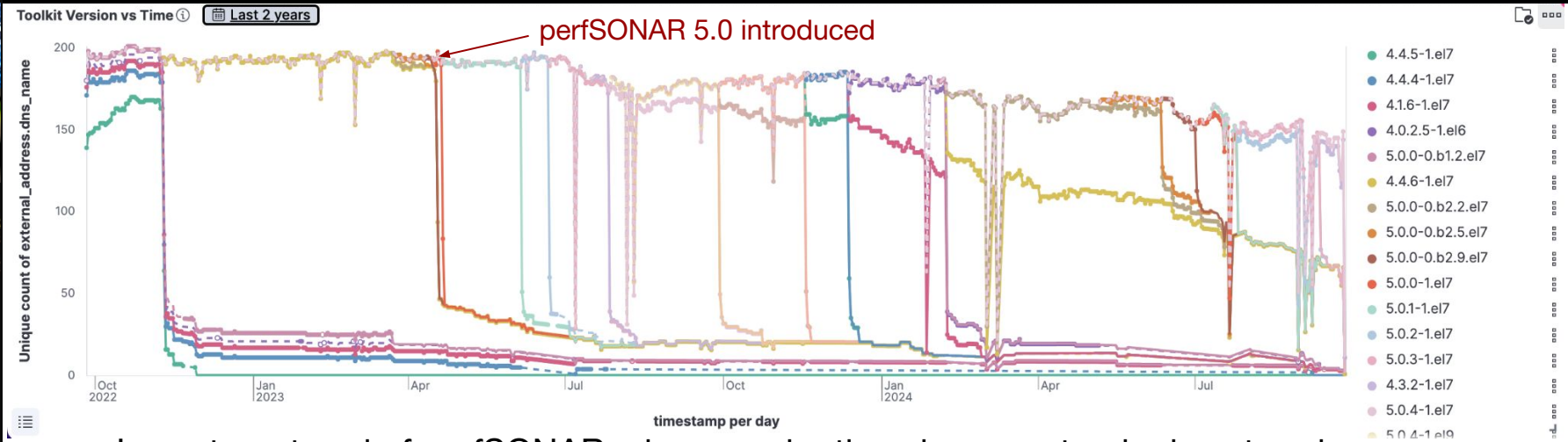


- 64% toolkits now on 5.1 and EL9
  - 25% still on EL7 !
- Core deployments are still on 10Gbps, but we have about 20% with 100Gbps
  - For WLCG/OSG testing purposes 10Gbps is still sufficient
  - Important to refresh HW along with the update to EL9
- MTU - around 36% on jumbo frames (9000), rest is on standard frames (1500)
- We have small testbed with about 10 perfSONARs with BBRv3 enabled

perfSONAR Infrastructure Evolution

181 Active perfSONARs

68 Communities in Use

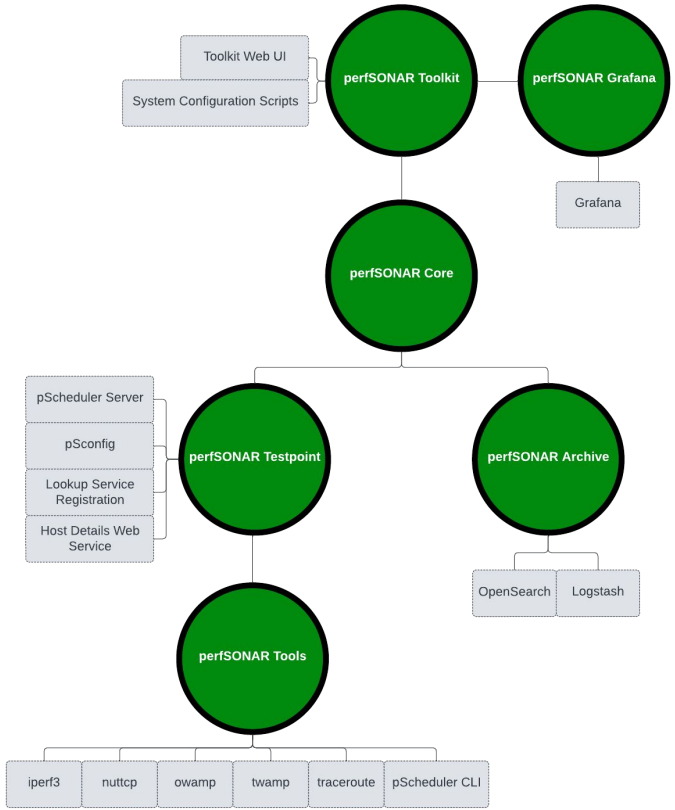


- Long-term trend of perfSONAR releases adoption shows a steady downtrend
  - Only reachable “toolkit version” installations are shown
  - Stability and reliability of the releases clearly playing a role

- New strategy and deployment models will need to be considered
  - Exploring simple deployments (without complex components such as ES)
  - 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)

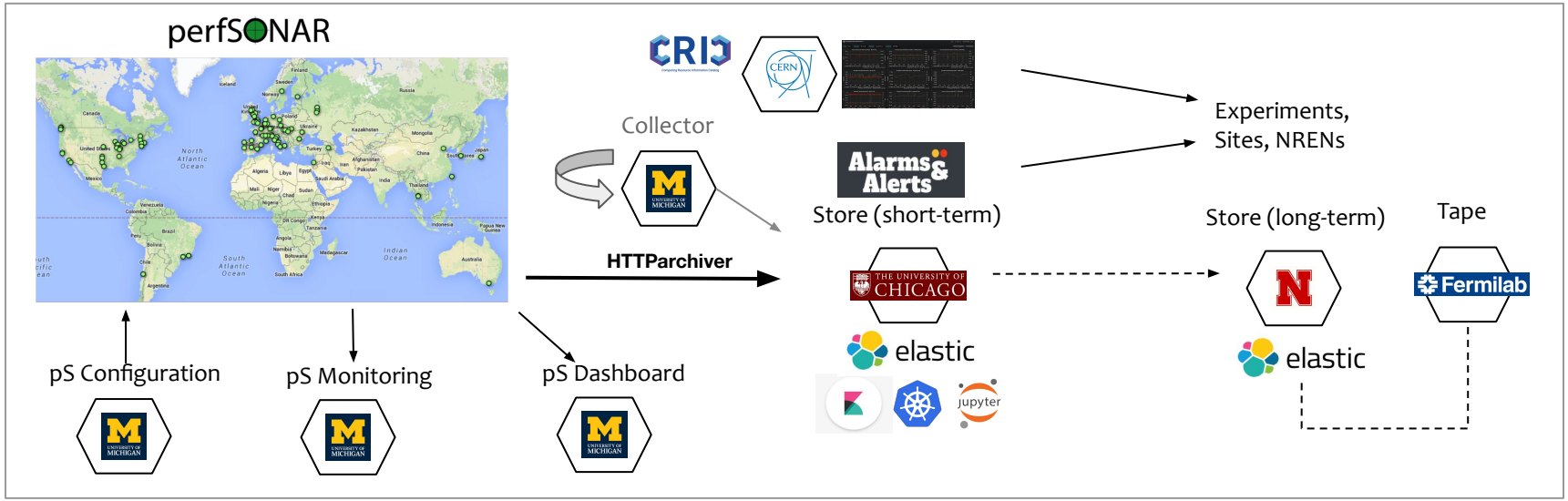
perfSONAR Testpoint

- perfSONAR Testpoint:
  - Automatically runs tests on a regular schedule
  - Participates in a centrally managed set of tests
  - Publishes the existence of a measurement node
- Can run in a container, i.e. starting testpoint as easy as
  - *docker run -d --name perfsonar-testpoint --net=host perfsonar/testpoint*
- Main concerns
  - No local cache that would sustain an outage of central services (48-72 hours)
  - No way to remotely monitor - needs extra packages/containers to run node_exporter and perfsonar exporter
  - Missing auto-updates and persistent configuration
  - Performance and integration with OS needs testing
- Looking for volunteers willing to deploy, test and develop the missing components



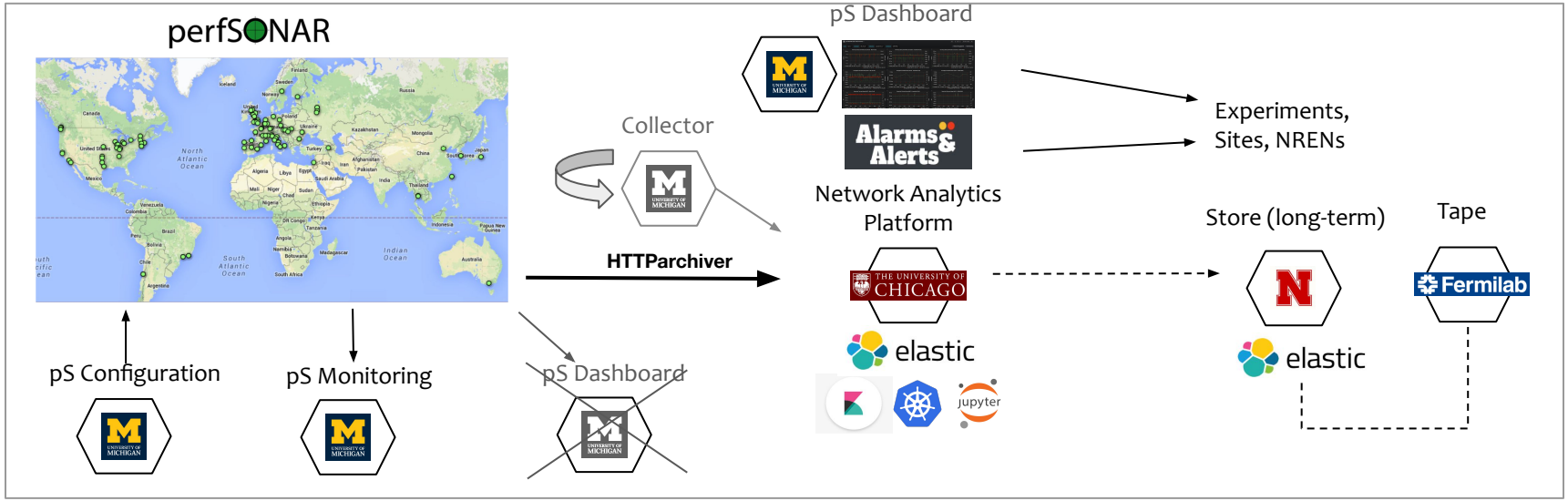
Network Measurement Platform Status

- Our platform collects, stores, configures and transports all network metrics
- Evolution based on the perfSONAR 5 already partially implemented.
  - Now directly publishing results from perfSONARs to ES@UC
  - Collector used only as a fallback;
  - WLCG CRIC now used for topology



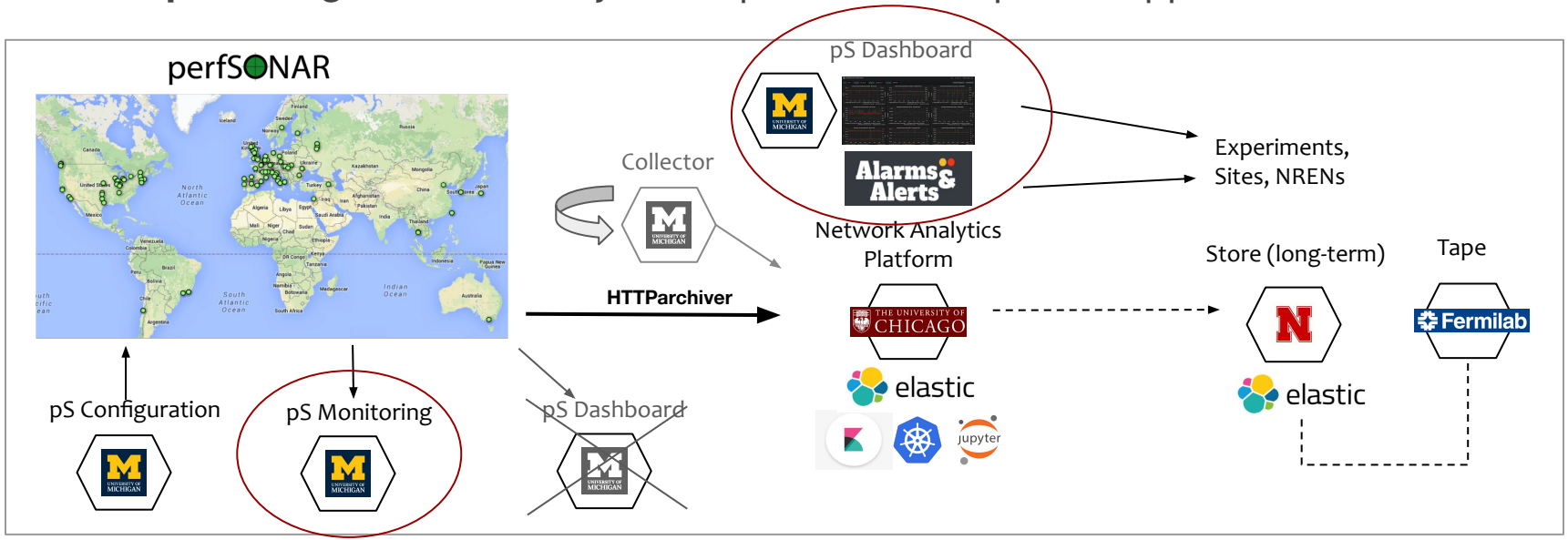
Network Measurement Platform Plans

- Evolution based on the perfSONAR 5 already partially implemented.
  - Forwarding to UNL and backup to FNAL still to be implemented
  - **pS Monitoring** - update to latest Checkmk and enable SSO authentication
  - **ps Dashboard** - integrate with Analytics Platform/Grafana (retire maddash)
  - **ps Configuration** - new tool has been announced



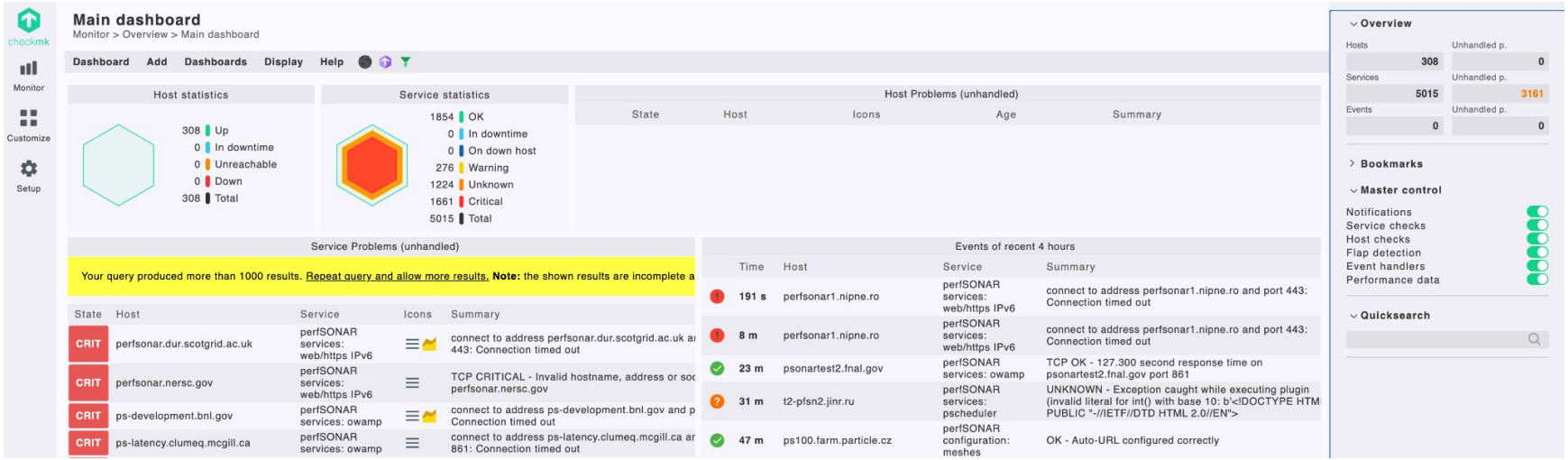
Network Measurement Platform Plans

- Evolution based on the perfSONAR 5 already partially implemented.
  - Forwarding to UNL and backup to FNAL still to be implemented
  - **pS Monitoring** - update to latest Checkmk and enable SSO authentication
  - **ps Dashboard** - integrate with Analytics Platform/Grafana (retire maddash)
  - **ps Configuration** - clarify development roadmap and support

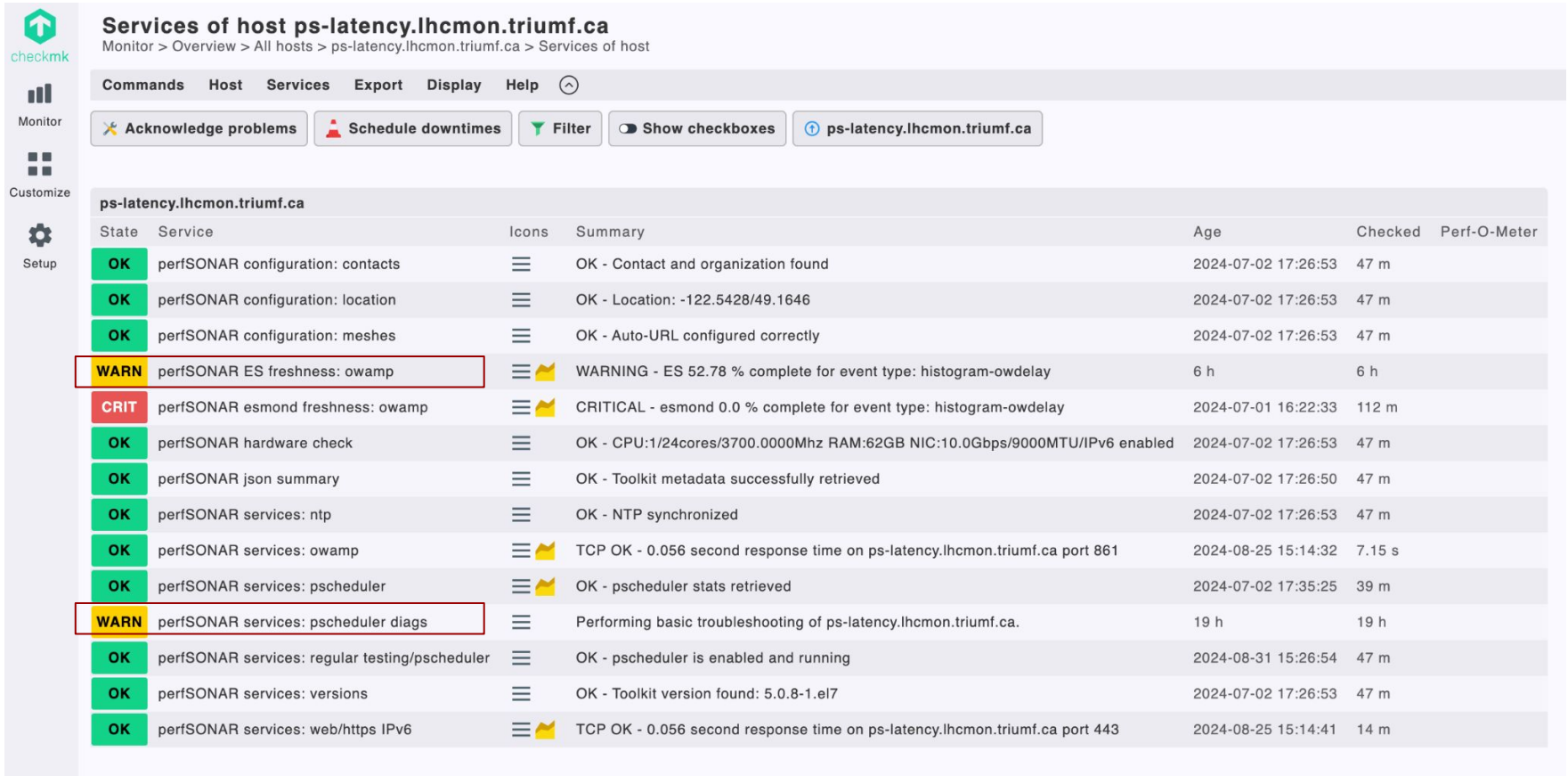


perfSONAR Infrastructure Monitoring

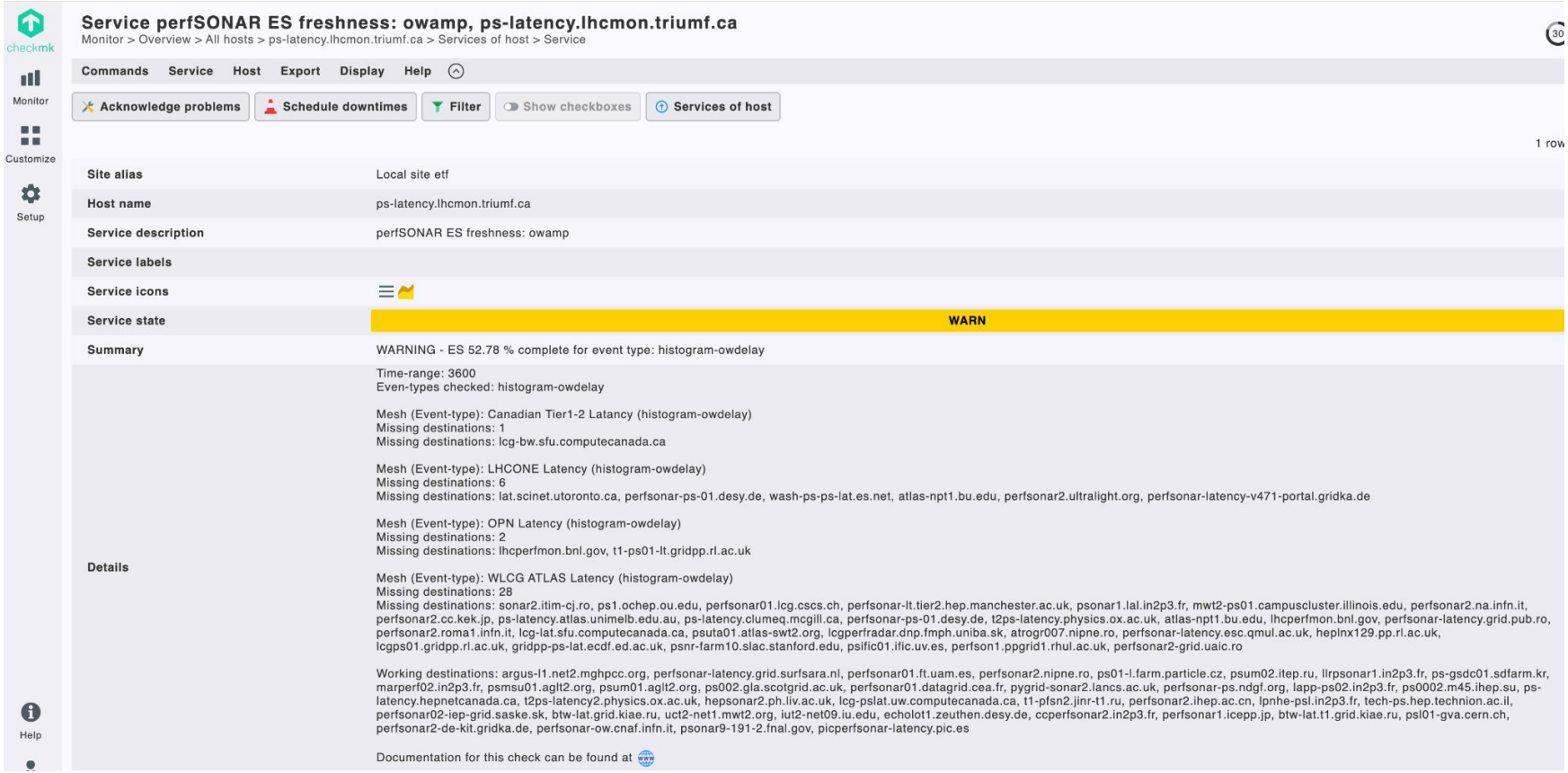
- Updated to Checkmk 2.3.0 (from 1.6.0)
- Integration with CILogon (single-sign on) - moving away from x509 certs
- New tests
  - Node diagnostics based on “pscheduler troubleshoot” command
  - Tracking measurements in central ElasticSearch
- Now in pre-production at psetf-itb.aglt2.org



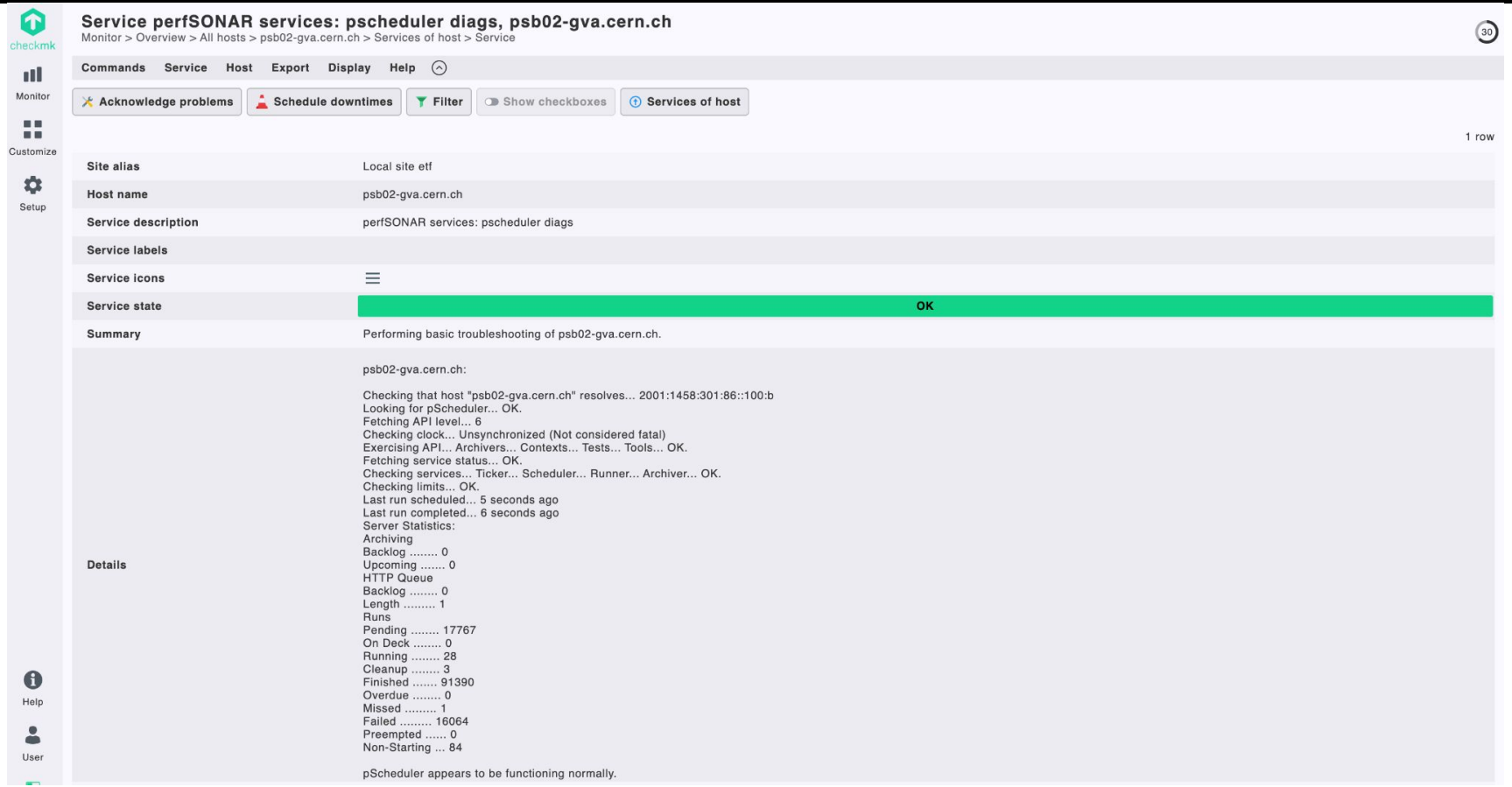
perfSONAR Infrastructure Monitoring



perfSONAR Infrastructure Monitoring: Freshness

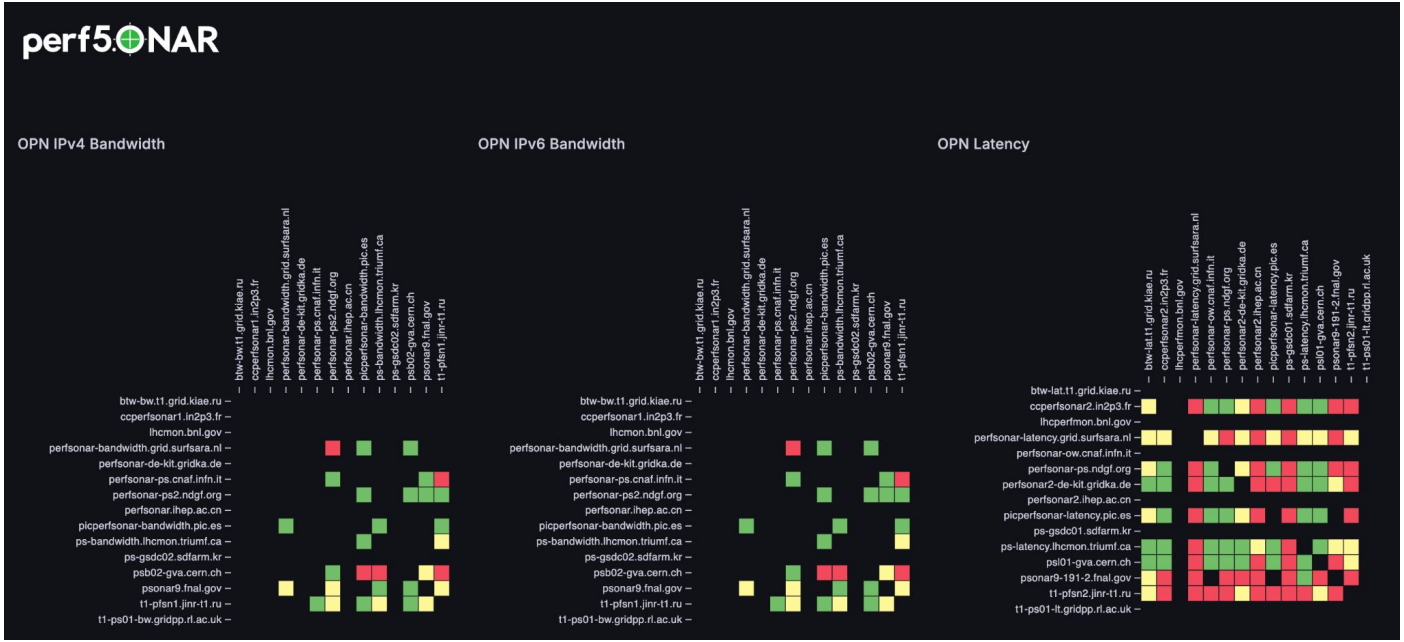


perfSONAR Infrastructure Monitoring: pscheduler

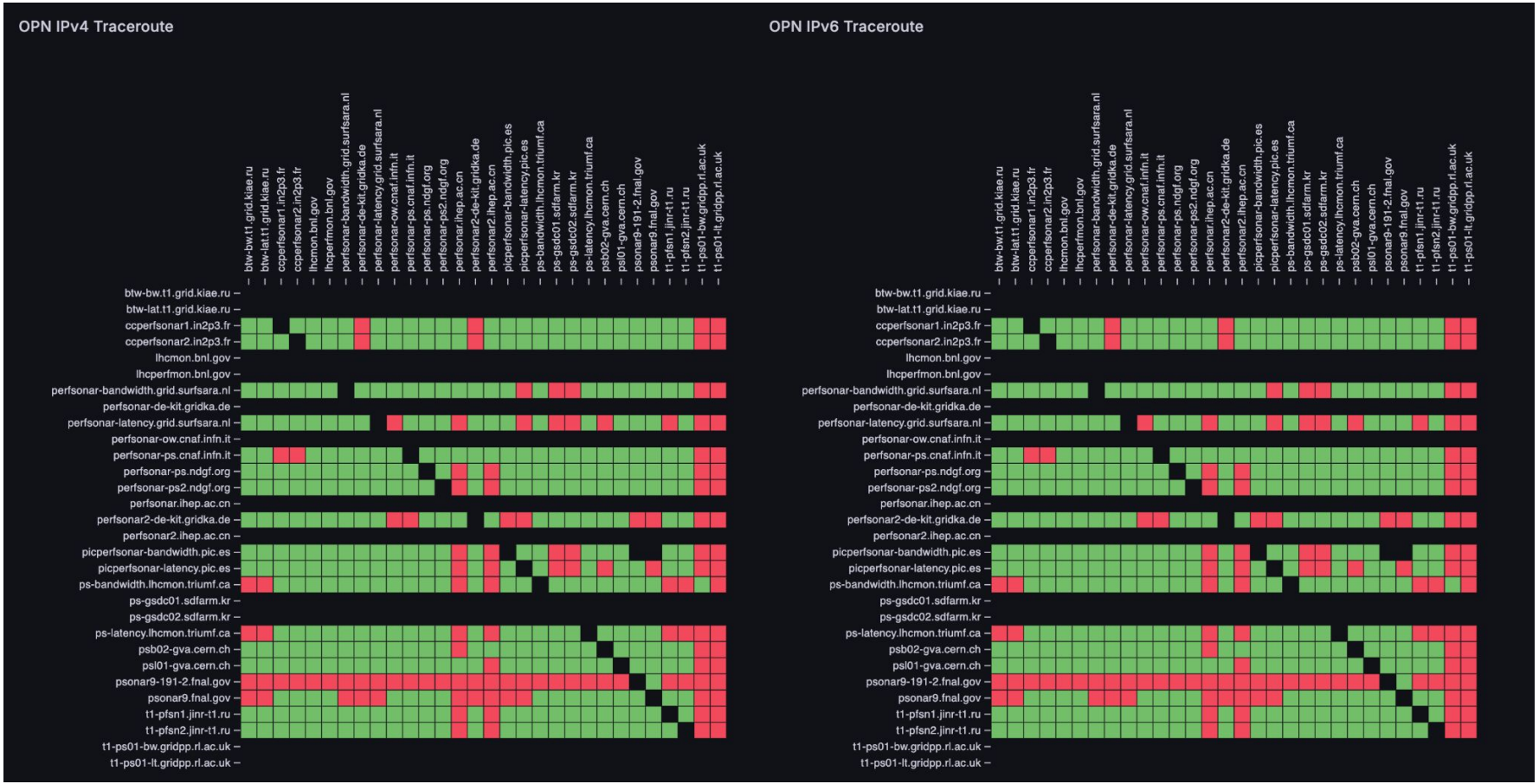


perfSONAR Dashboard

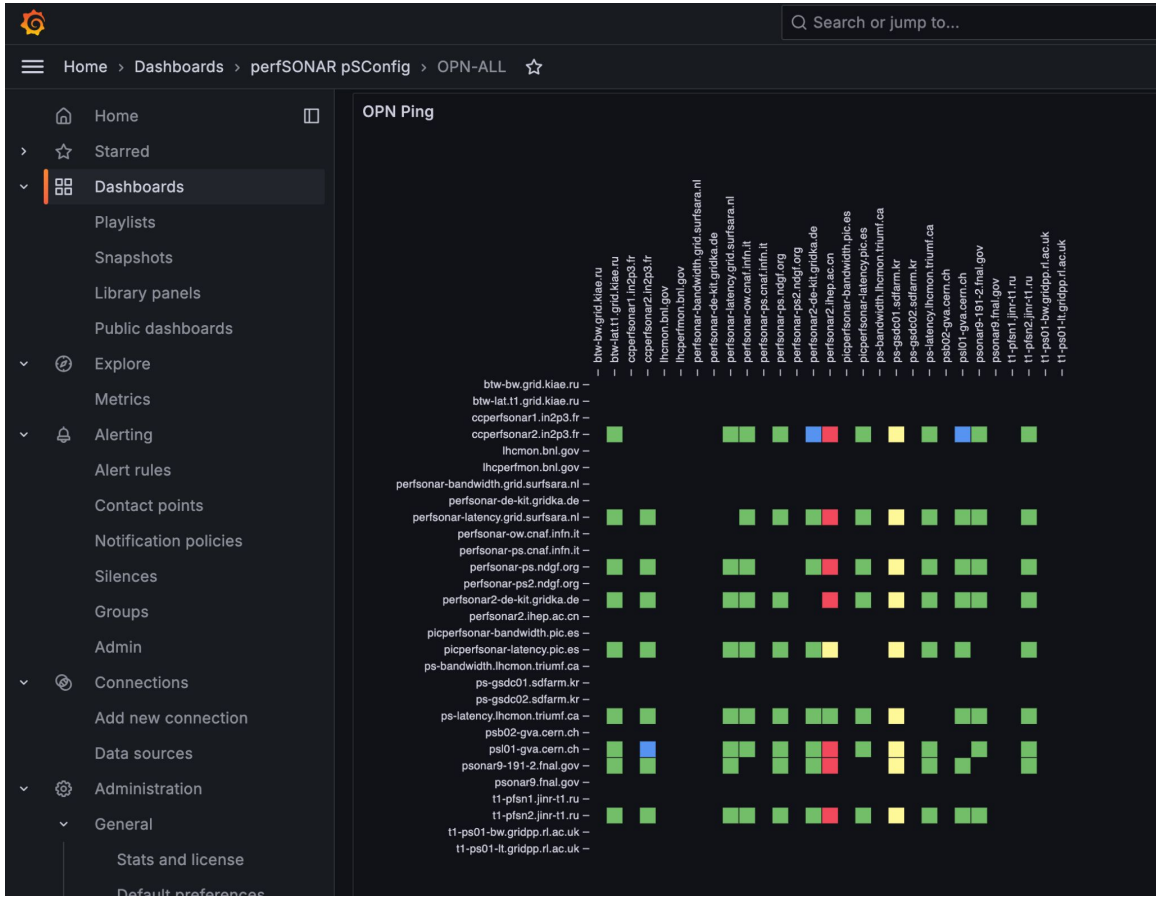
- New dashboard service replacing previous maddash
- Based on Grafana - based on perfSONAR 5 code base
  - Modified to use the central ElasticSearch, which uses different schema
  - Generation of dashboards required some new code which is now upstream



perfSONAR Dashboard



perfSONAR Dashboard



perfSONAR Dashboard

Analytics

Network Analytics R&D

- Investigate ML models/methods to process network measurements

- 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
- Correlate 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.)

Summary

- Updates to perfSONAR and OSG/WLCG network measurement platform
  - perfSONAR 5.2 is coming with new features.
  - New infrastructure monitoring and dashboards will appear in production soon
  - Simplified deployment models are being evaluated
- Ongoing efforts in network analytics and ML methods for our data
  - Focus on pre-processing (gaps, predictive models) and anomaly detection
  - Opportunity to collaborate on models and data sets
- **We are running monthly meetings with site network teams**:
  - Discuss how sites are deploying, managing their network infrastructure and planning for WLCG networking requirements
  - Meetings every 3rd week of a month on Thursday at 10am EST (to join mail wlcg-site-net-requests@umich.edu)

- We have to continue to watch our network monitoring infrastructure as it is a complex system with lots of areas for issues to develop.

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

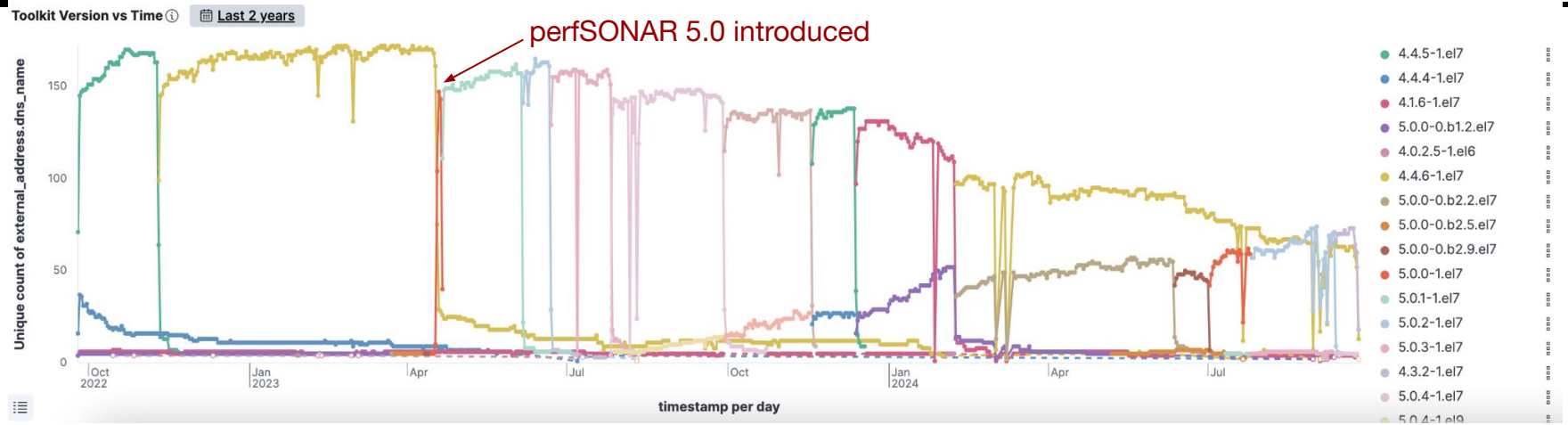
Useful URLs

- OSG/WLCG Networking Documentation
  - https://opensciencegrid.github.io/networking/
- perfSONAR Infrastructure Dashboard
  - https://atlas-kibana.mwt2.org:5601/s/networking/goto/9911c54099b2be47ff9700772c3778b7
- WLCG DOMA DC24 plans
  - https://indico.cern.ch/event/1225415/contributions/5155042/attachments/2593516/4476291/Data%20Challenge%202024.pdf
- perfSONAR Central Configuration
  - https://psconfig.opensciencegrid.org/
- Toolkit information page
  - https://toolkitinfo.opensciencegrid.org/
- Grafana dashboards
  - http://monit-grafana-open.cern.ch/
- ATLAS Alerting and Alarming Service: https://psa.osg-htc.org/
- The perfSONAR Dashboard application: https://ps-dash.uc.ssl-hep.org/
- ESnet WLCG Stardust Dashboard: https://public.stardust.es.net/d/XkxDL5H7z/esnet-public-dashboards?orgId=1

perfSONAR Infrastructure Evolution

181 Active perfSONARs

68 Communities in Use



- Long-term trend of perfSONAR releases adoption shows a steady downtrend
  - Only reachable “toolkit version” installations are shown
  - Stability and reliability of the releases clearly playing a role

- New strategy and deployment models will need to be considered
  - Explore simplified deployments (in containers; without complex components such as ES)
  - Providing means to easily reset/restart deployment (to avoid accumulation of issues)
  - Must be easy to co-locate with storages (or even co-hosted on storages with multiple NICs)

Backup Slides Follow

Alarms & Alerts Interface

Components

Database
Elasticsearch

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

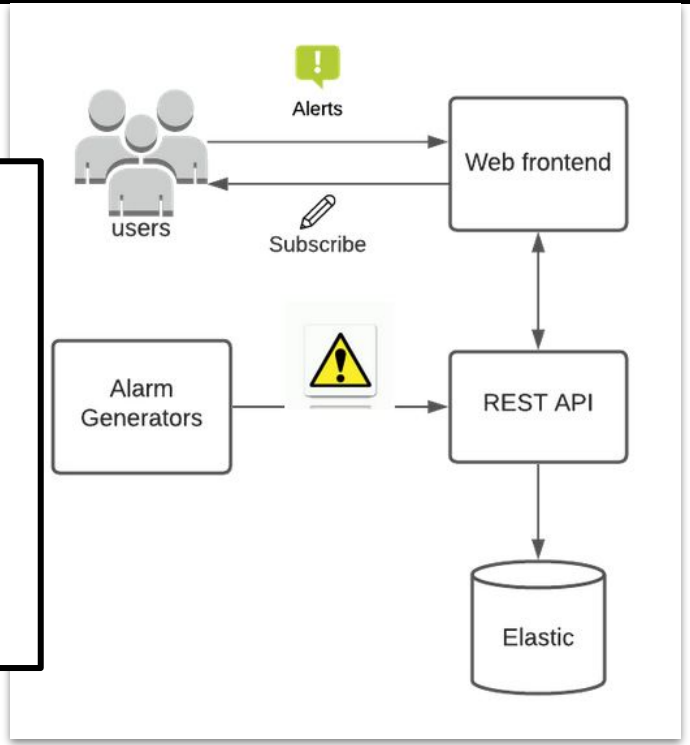
Deployment
Docker, K8s, Helm (soon)

Authentication
Globus InCommon

Authorization
API key

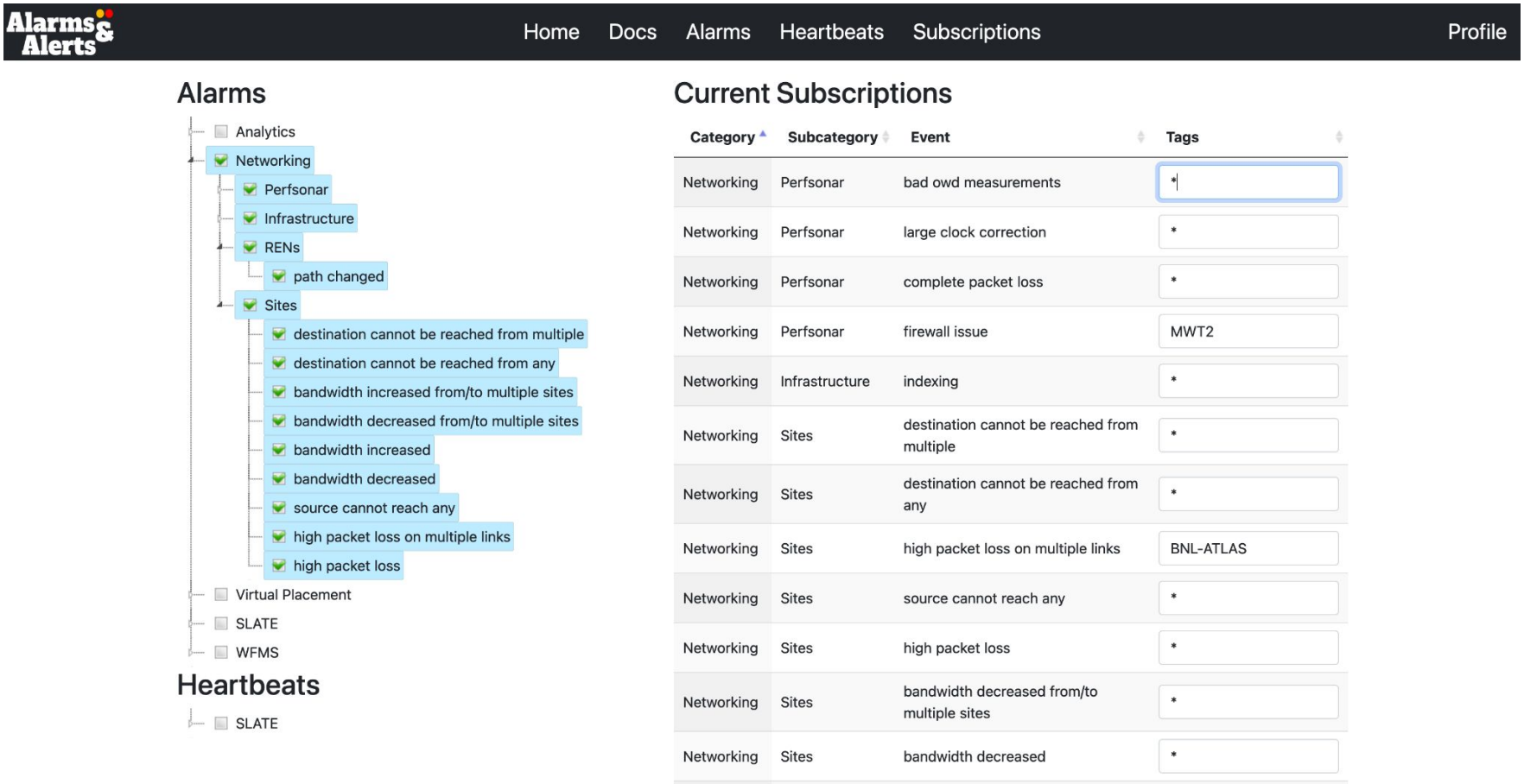
Mail
Mailgun

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



Two main improvements needed: **Acknowledging alerts** that are being worked on and **adding user notification mailing lists**

Subscription Interface



Alarm Types and Relation to perfSONAR Data

All based on perfSonar data

One-Way Delay

- bad owd measurements
- large clock correction

Traceroute

- path changed
- destination cannot be reached
- source cannot reach any

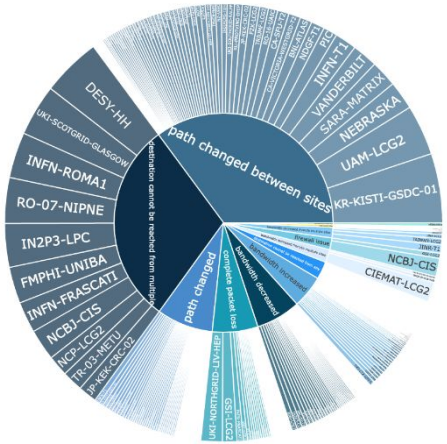
Packet loss

- complete packet loss
- firewall issue
- high packet loss (on multiple links)

Throughput

- bandwidth decreased (from/to multiple sites)
- bandwidth increased (from/to multiple sites)

psDash Alarms Dashboard



Search & Explore the Networking Alarms

03/04/2023 → 03/05/2023

Rounded to the day

Search for a site

Search for an event type

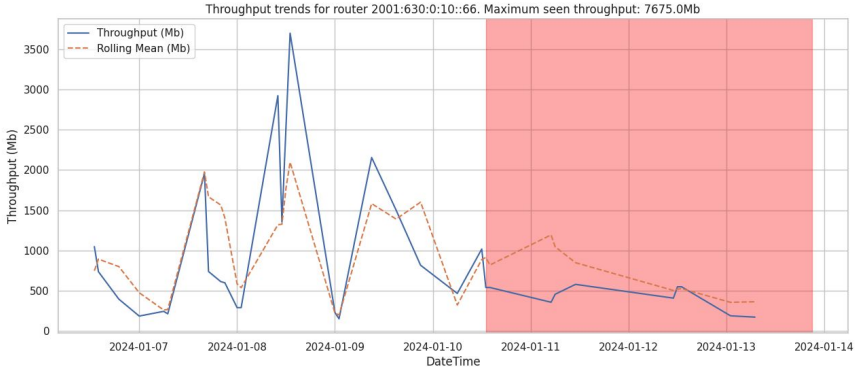
List of alarms

BANDWIDTH DECREASED

from	to	src_site	dest_site	ipv	ipv6	last3days_avg	%change	alarm_link
filter data...								
2023-02-12 04:08	2023-03-05 04:08	AGLT2	RRC-KI-T1	ipv4	false	96	-64	VIEW
2023-02-12 04:08	2023-03-05 04:08	IN2P3-CC	FZK-LCG2	ipv4	false	1008	-74	VIEW

Plans for the Analytics Platform

- Production of the anomaly detection based on Bayesian inference
  - Uses RTT, traceroutes, TTLs as input and detects anomalies
- Continue working on the neural network models that correlate throughputs and traceroutes
  - Generating real-world model of our entire network (all routers)
  - Not only detecting anomalies, but also trying to pinpoint the location of the issue



- Improve infrastructure alarming to the point where we can reliably differentiate infrastructure and network issues
- Network availability dashboard in production

DC24

WLCG DC24

WLCG Data Challenge 2024 took place in Feb 2024; targeting 25% of HL-LHC

Our DC24 plans included the following:

- Update and utilize **perfSONAR** to clean up links and fix problems before DC24.
- Instrument and document **site networks**, for at least our largest sites.
- **Network planning**: we need to make sure our sites and their local and regional networks are aware of our requirements and timeline and are planning appropriately
- **IPv6** should be enabled everywhere not just because of packet marking, but because it will allow us to get back to a single stack sooner!

psDash Network Status

ps dash | SITES OVERVIEW | SEARCH ALARMS | EXPLORE PATHS | MAJOR ALARMS

Status of all sites in the past 24 hours

Summary

0	28	104	3

Highest number of alarms from site

KR-KISTI-GSDC-1-LHCOPNE (KR): 75

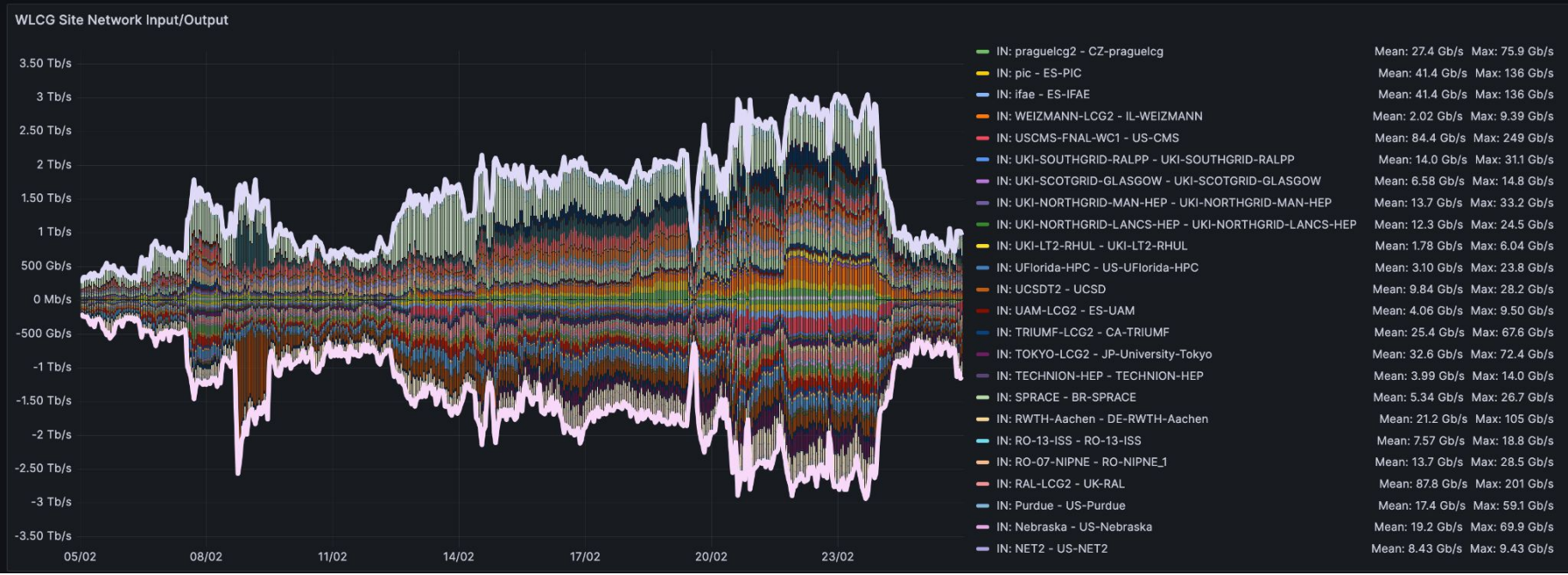
Highest number of alarms from country

United Kingdom: 108

SITE	STATUS	NETWORK	INFRASTRUCTURE	OTHER	URL
filter data...					
KR-KISTI-GSDC-1-LHCOPNE		75	0	0	See latest alarms
TR-ULAKBIM-LHCONE		1	1	0	See latest alarms
INFN-LNL-2-LHCONE		1	3	0	See latest alarms
FMPHI-UNIBA		2	1	0	See latest alarms
BUDAPEST		2	1	0	See latest alarms
BEGRID-ULB-VUB		2	1	0	See latest alarms
TECHNION-HEP		2	1	1	See latest alarms
UKI-LT2-RHUL		2	2	0	See latest alarms
NCBJ-LHCOPN		2	2	0	See latest alarms

Network Status dashboard - part of Network Analytics platform - shows network performance based on perfSONAR measurements. Status (ok/warning/critical/unknown) aggregates network and infrastructure metrics.

Site Network Utilisation



Site Network Utilisation - computed from aggregated utilisation (SNMP counters) provided by sites via simple API. Screenshot shows network utilisation during DC24 as seen by the sites.

Tools and Applications for Network Data

- To organize access to all the various resources we recommend using our Toolkitinfo page: https://toolkitinfo.opensciencegrid.org/
- Reminder: we already have Kibana dashboards looking at
  - Bandwidth
  - Traceroute
  - Packetloss / Latency
  - Infrastructure
- For this meeting we want to update our recent work towards a user subscribable alerting and alarming service
  - User interface to subscribe is **AAAS** (ATLAS Alerting and Alarming Service)
  - Tool to explore alerts is **pS-Dash** (Plotly base perfSONAR dashboard UI tool)

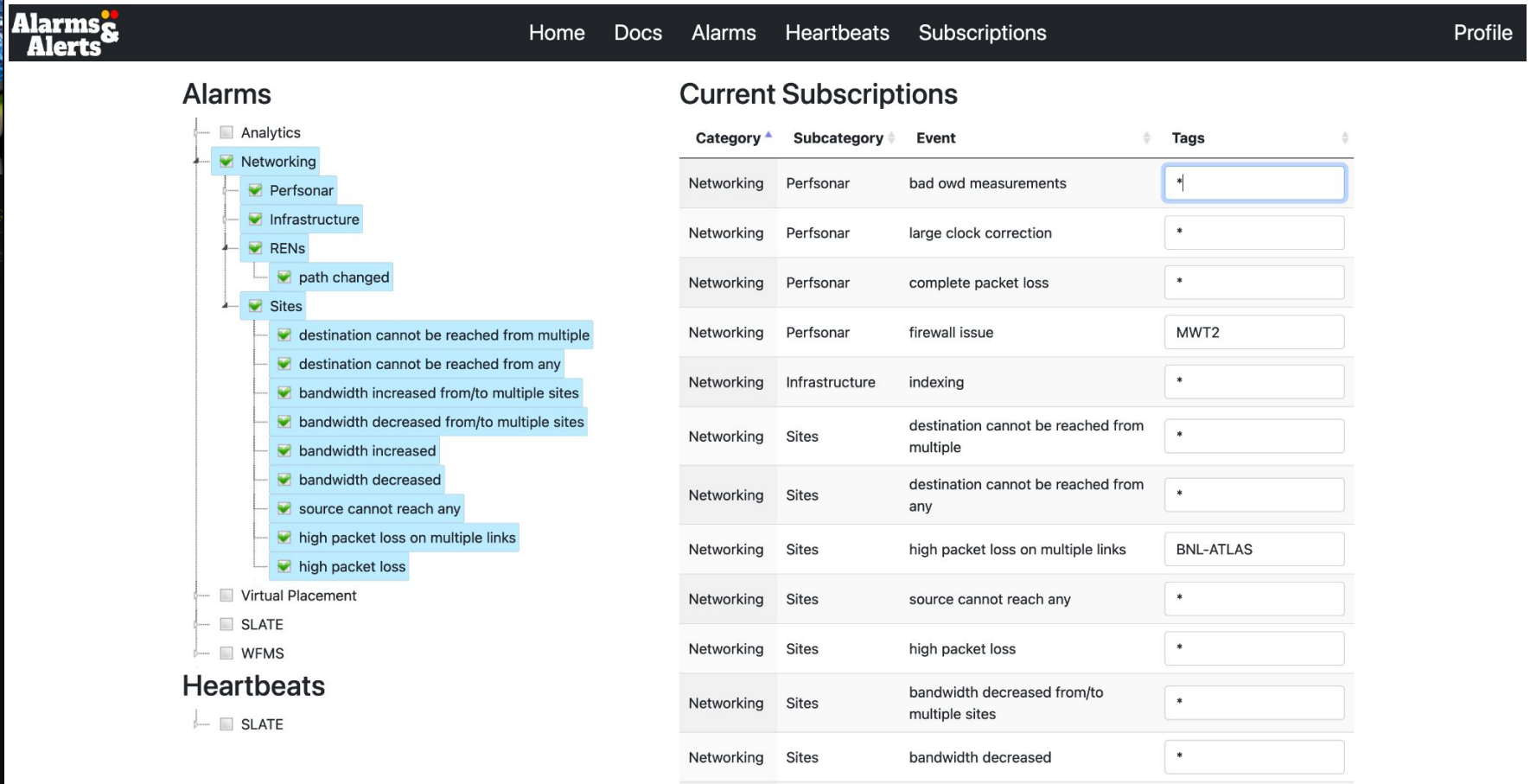
Alarms & Alerts Service

https://psa.osg-htc.org

(Uses EDUGain/InCommon)

Purpose: provides user-subscribable alerting for specific types of network issues found by analyzing perfSONAR data

The Alerting and Alarming Tools Subscription Interface



Alarm Types and Relation to perfSONAR Data

All based on perfSonar data

One-Way Delay

- bad owd measurements
- large clock correction

Traceroute

- path changed
- destination cannot be reached
- source cannot reach any

Packet loss

- complete packet loss
- firewall issue
- high packet loss (on multiple links)

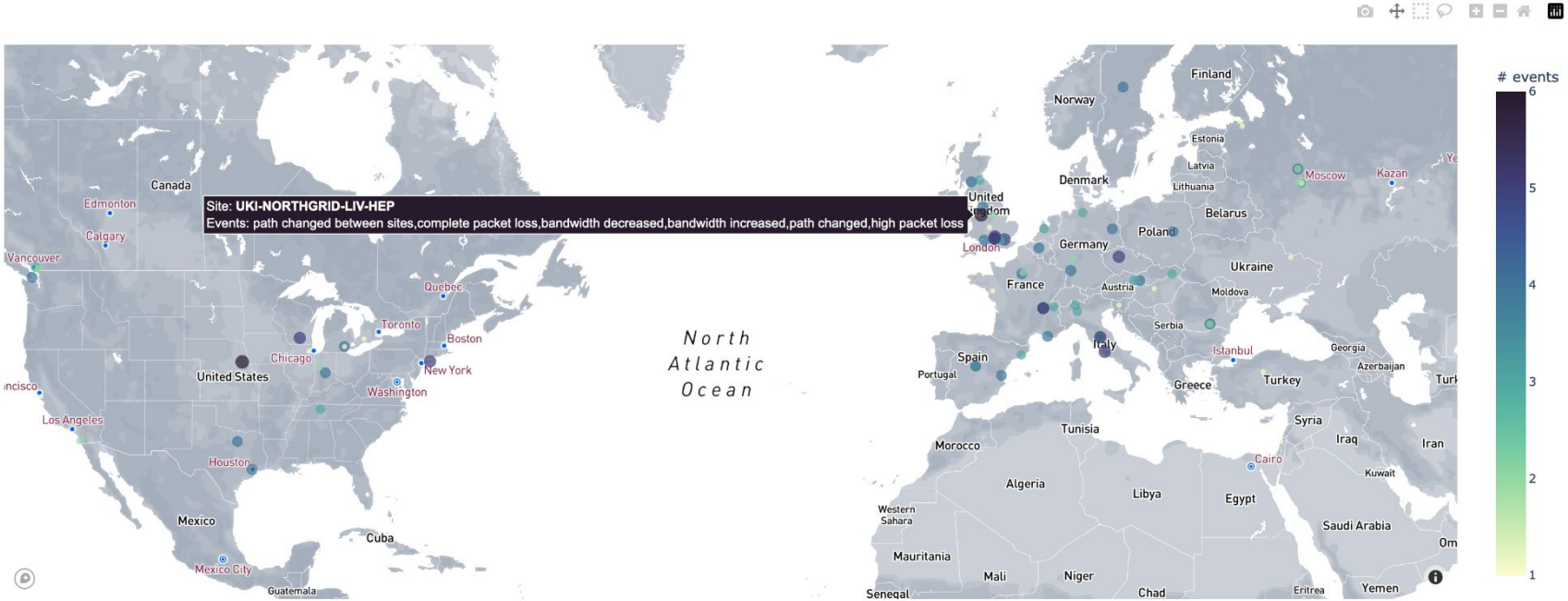
Throughput

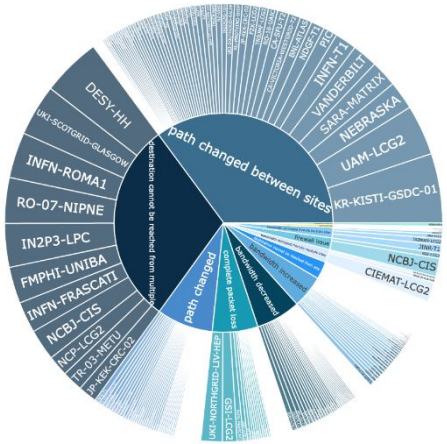
- bandwidth decreased (from/to multiple sites)
- bandwidth increased (from/to multiple sites)

pSDash (perfSONAR Dashboard)

Selected site: UKI-NORTHGRID-LIV-HEP

Alarms reported in the past 24 hours (2023-03-05 19:00 UTC)





Search & Explore the Networking Alarms

03/04/2023 → 03/05/2023

Rounded to the day

Search for a site

Search for an event type

List of alarms

BANDWIDTH DECREASED

from	to	src_site	dest_site	ipv	ipv6	last3days_avg	%change	alarm_link
filter data...								
2023-02-12 04:08	2023-03-05 04:08	AGLT2	RRC-KI-T1	ipv4	false	96	-64	VIEW
2023-02-12 04:08	2023-03-05 04:08	IN2P3-CC	FZK-LCG2	ipv4	false	1008	-74	VIEW

WLCG perfSONAR Path Statistics

We uniquely identify each traceroute (route **IP** path) with a SHA1 hash.

	route-sha1
count	19995.000000
mean	19.911678
std	43.373343
min	1.000000
25%	2.000000
50%	4.000000
75%	12.000000
max	377.000000

5264 links tested Link=”hop” (IP-to-IP)

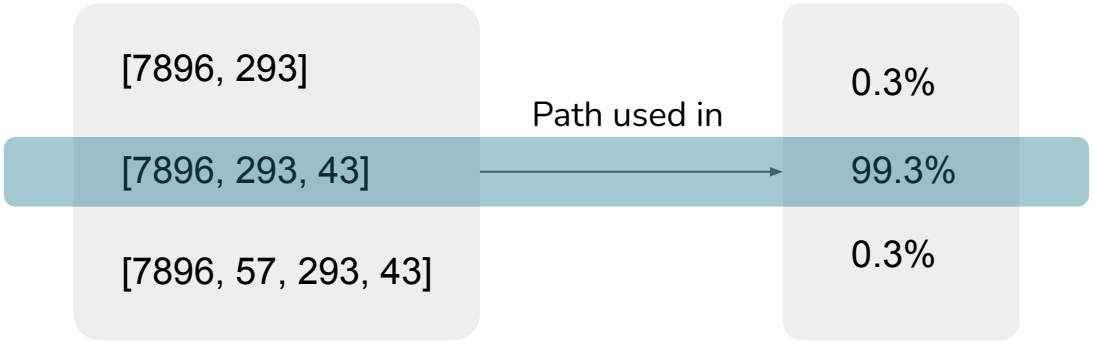
4415 traversed nodes Node=”router”

Statistics on the left concern all the “paths” we are tracking with about 20K unique paths found
About 50% of src-dest pairs have 4 or less paths.

AS (Autonomous System) Path Changed

NOTE: Paths denoted by route IP are too noisy; instead use AS number

ASN sequence	Reduced ASNs
[7896, 7896, 293, 293, 293, 293, 293, 293]	[7896, 293]
[7896, 7896, 293, 293, 293, 293, 293, 293, 43]	[7896, 293, 43]
[7896, 7896, 7896, 7896, 57, 57, 57, 293, 293, 293, 293, 293, 293, 43]	[7896, 57, 293, 43]



Baseline

[7896, 293, 43]

57

Example: LHCOPN/LHCONE Load Balancing

Source

INFN-T1

2001:760:4205:254::11

Destination

SARA-MATRIX

2001:610:108:203a::31

Total number of traceroute measures: 280

Other networking alarms: None found

BASELINE PATH

Taken in 50% of time — Always reaches destination: NO

137 513 1103

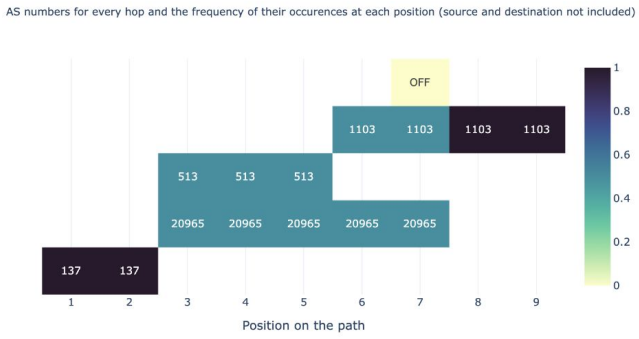
ALTERNATIVE PATHS

Taken in 50.0% of time — Always reaches destination: NO

137 20965 1103

Taken in 0.0% of time — Always reaches destination: NO

137 20965



At position	Typically goes through	Changed to
3	**513** CERN, CH	**20965** GEANT The GEANT IP Service, NL
4	**513** CERN, CH	**20965** GEANT The GEANT IP Service, NL
5	**513** CERN, CH	**20965** GEANT The GEANT IP Service, NL
6	**1103** SURFNET-NL SURFnet, The Netherlands, NL	**20965** GEANT The GEANT IP Service, NL

Example: LHCOPN Alternate via ESnet

USCMS-FNAL-WC1->CERN-PROD

Source

USCMS-FNAL-WC1

131.225.205.23

Destination

CERN-PROD

128.142.208.134

Total number of traceroute measures: 248

Other networking alarms: None found

BASELINE PATH

Taken in 99% of time | Always reaches destination: NO

3152 513

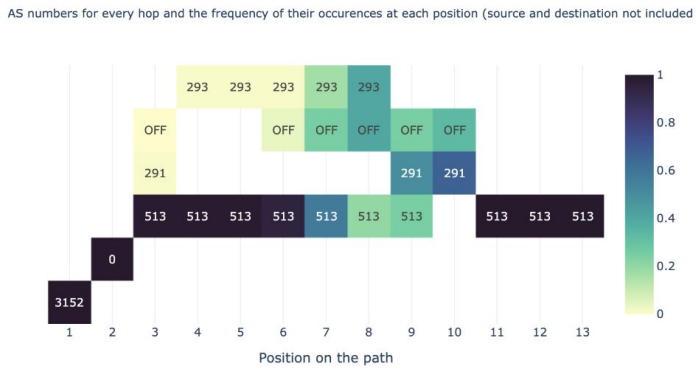
ALTERNATIVE PATHS

Taken in 0.0% of time | Always reaches destination: NO

3152

Taken in 1.0% of time | Always reaches destination: NO

3152 291 293 513



At position	Typically goes through	Changed to
3	**513** CERN, CH	**291** ESNET-EAST, US

At position	Typically goes through	Changed to
4	**513** CERN, CH	**293** ESNET, US

At position	Typically goes through	Changed to
5	**513** CERN, CH	**293** ESNET, US

Example: FNAL Incident (BW drop)

USCMS-FNAL-WC1 to NDGF-T1

Source

USCMS-FNAL-WC1

Destination

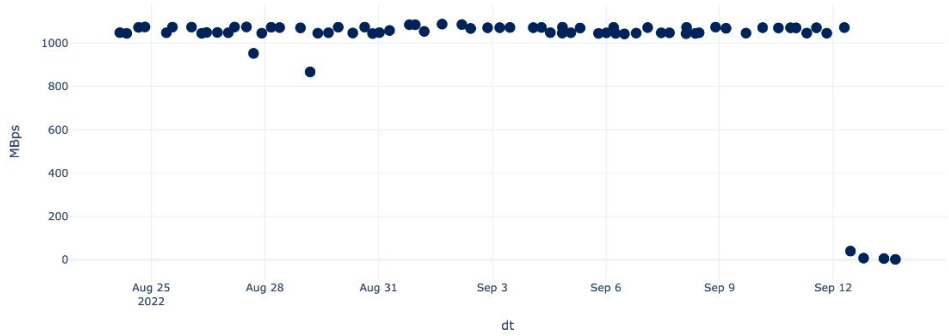
NDGF-T1

Change: -100%

Total number of throughput measures: 71

Other networking alarms

| High packet loss: 2 | High packet loss on multiple links: 1 |



push	MA	src	dest	src_host	dest_host	ipv6	src_site	src_VO	dest_site	dest_VO	src_production	dest_production	timestamp	throughput	pair	dt	MBps
filter da																	
false	131.225.205.23	131.225.205.23	109.105.124.88	psonar3.fnal.gov	perfsonar-ps2.ndgf.org	false	USCMS-FNAL-WC1	UNKNOWN	NDGF-T1	ATLAS	true	true	1661561263000	1048161726	131.225.205.23->109.105.124.88	2022-08-27T00:47:43	1048.16
false	131.225.205.23	131.225.205.23	109.105.124.88	psonar3.fnal.gov	perfsonar-ps2.ndgf.org	false	USCMS-FNAL-WC1	UNKNOWN	NDGF-T1	ATLAS	true	true	1661618925000	952946516	131.225.205.23->109.105.124.88	2022-08-27T16:48:45	952.95
false	131.225.205.23	131.225.205.23	109.105.124.88	psonar3.fnal.gov	perfsonar-ps2.ndgf.org	false	USCMS-FNAL-WC1	UNKNOWN	NDGF-T1	ATLAS	true	true	1662626239000	1045220096	131.225.205.23->109.105.124.88	2022-09-08T08:37:19	1045.22
false	131.225.205.23	131.225.205.23	109.105.124.88	psonar3.fnal.gov	perfsonar-ps2.ndgf.org	false	USCMS-FNAL-WC1	UNKNOWN	NDGF-T1	ATLAS	true	true	1661678463000	1072068304	131.225.205.23->109.105.124.88	2022-08-28T09:21:03	1072.07
false	131.225.205.23	131.225.205.23	109.105.124.88	psonar3.fnal.gov	perfsonar-ps2.ndgf.org	false	USCMS-FNAL-WC1	UNKNOWN	NDGF-T1	ATLAS	true	true	1662439905000	1072905581	131.225.205.23->109.105.124.88	2022-09-06T04:51:45	1072.91
false	131.225.205.23	131.225.205.23	109.105.124.88	psonar3.fnal.gov	perfsonar-ps2.ndgf.org	false	USCMS-FNAL-WC1	UNKNOWN	NDGF-T1	ATLAS	true	true	1661659004000	1073324325	131.225.205.23->109.105.124.88	2022-08-28T03:56:44	1073.32
false	131.225.205.23	131.225.205.23	109.105.124.88	psonar3.fnal.gov	perfsonar-ps2.ndgf.org	false	USCMS-FNAL-WC1	UNKNOWN	NDGF-T1	ATLAS	true	true	1662672411000	1074163359	131.225.205.23->109.105.124.88	2022-09-08T21:26:51	1074.16
false	131.225.205.23	131.225.205.23	109.105.124.88	psonar3.fnal.gov	perfsonar-ps2.ndgf.org	false	USCMS-FNAL-WC1	UNKNOWN	NDGF-T1	ATLAS	true	true	1662902418000	1071231326	131.225.205.23->109.105.124.88	2022-09-11T13:20:18	1071.23
false	131.225.205.23	131.225.205.23	109.105.124.88	psonar3.fnal.gov	perfsonar-ps2.ndgf.org	false	USCMS-FNAL-WC1	UNKNOWN	NDGF-T1	ATLAS	true	true	1662093921000	1085912472	131.225.205.23->109.105.124.88	2022-09-02T04:45:21	1085.91
false	131.225.205.23	131.225.205.23	109.105.124.88	psonar3.fnal.gov	perfsonar-ps2.ndgf.org	false	USCMS-FNAL-WC1	UNKNOWN	NDGF-T1	ATLAS	true	true	1662696230000	1068710540	131.225.205.23->109.105.124.88	2022-09-09T04:03:50	1068.71

Example: Fail-over to Commodity Network

Nebraska -> RAL-LCG2

Source

Nebraska

2600:900:6:1102:2eea:7fff:fef5:d140

Destination

RAL-LCG2

2001:630:58:1820::82f6:b06d

Total number of traceroute measures: 280
Other networking alarms: None found

BASELINE PATH

Taken in 99% of time — Always reaches destination: NO

7896 11537 20965 786

ALTERNATIVE PATHS

Taken in 1.0% of time — Always reaches destination: NO

7896 6939 786

Taken in 0.0% of time — Always reaches destination: YES

7896 7029 6461 3257 1299 786

At position	Typically goes through	Changed to
4	**11537** INTERNET2-RESEARCH-EDU, US	**7029** WINDSTREAM, US

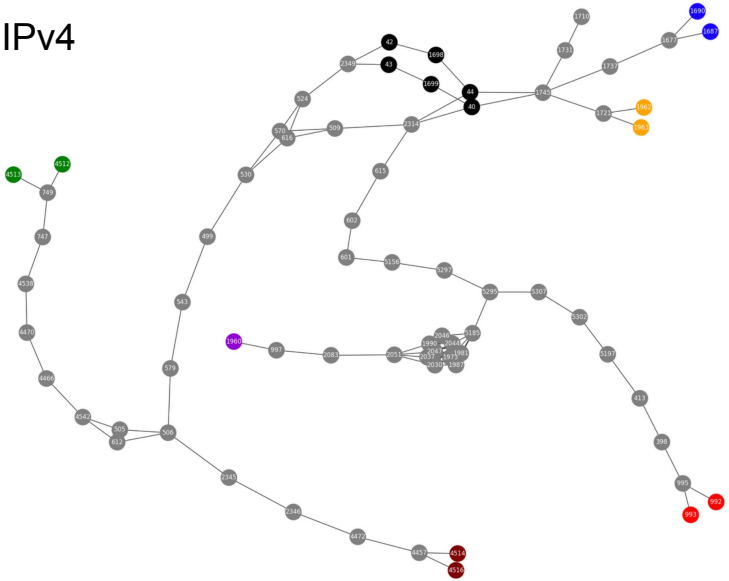
At position	Typically goes through	Changed to
7	**11537** INTERNET2-RESEARCH-EDU, US	**6939** HURRICANE, US

At position	Typically goes through	Changed to
7	**6461** ZAYO-6461, US	**6939** HURRICANE, US

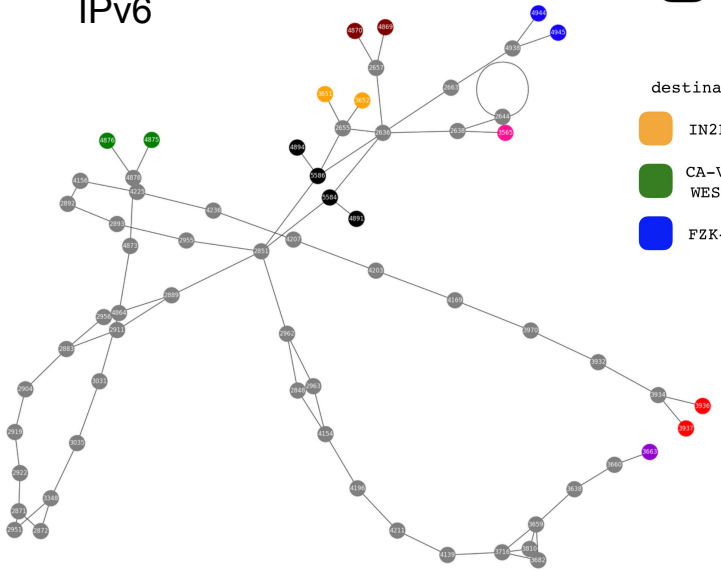
Challenges and Ongoing Work

Paths differ significantly

IPv4



IPv6



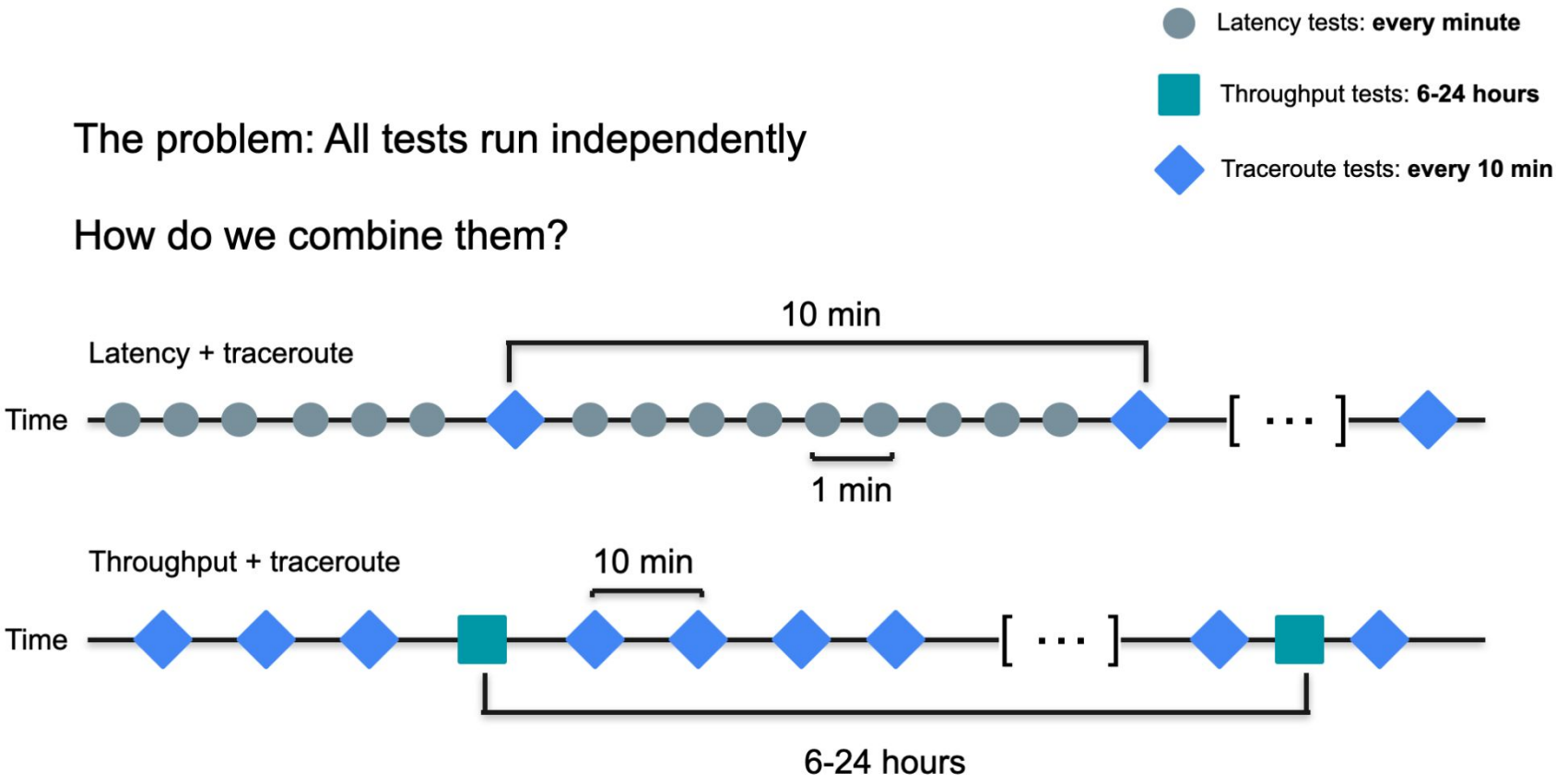
source

BNL-ATLAS
closest 2 hops

destinations

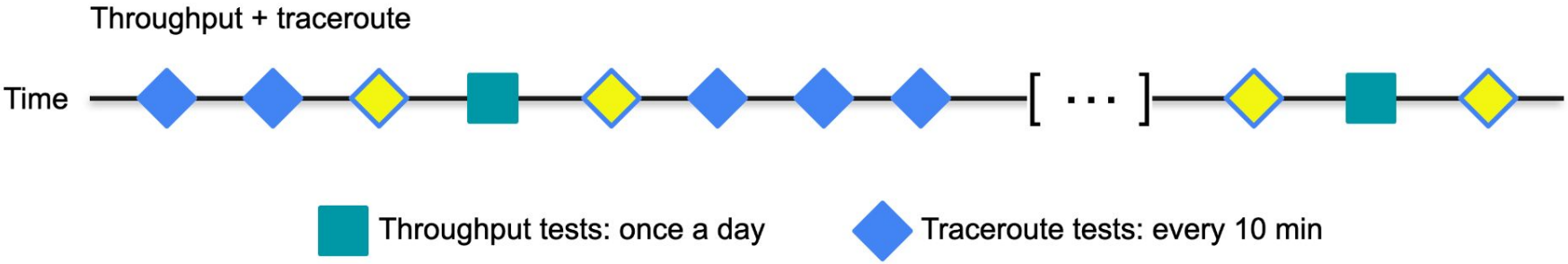
- IN2P3-CC
- IFIC-LCG2
- CA-VICTORIA-WESTGRID-T2
- TRIUMF-LCG2
- FZK-LCG2
- IN2P3-LPSC
- RAL-LCG2

Correlating Tests with Paths: Two Timescales



Connecting Throughput to Traceroute

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

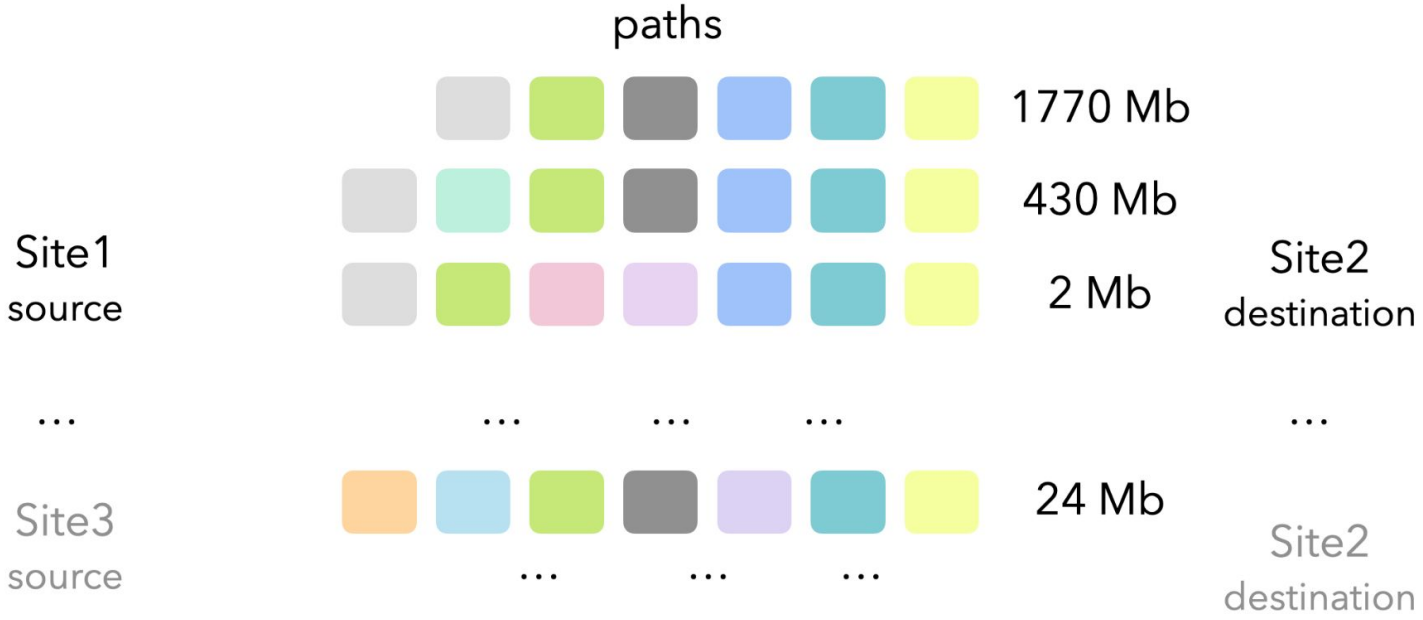


Our starting choice: Use **both** tracepaths (just before; just after) as valid paths and attribute BW to both.

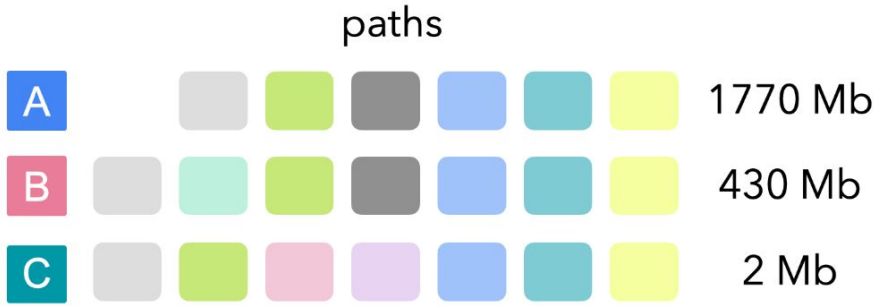
Have to see if this is superior to just using the last measured route before the measurement…

Attaching Throughput Results to Sets of Routers/Links

Each colored box represents a specific router along the path

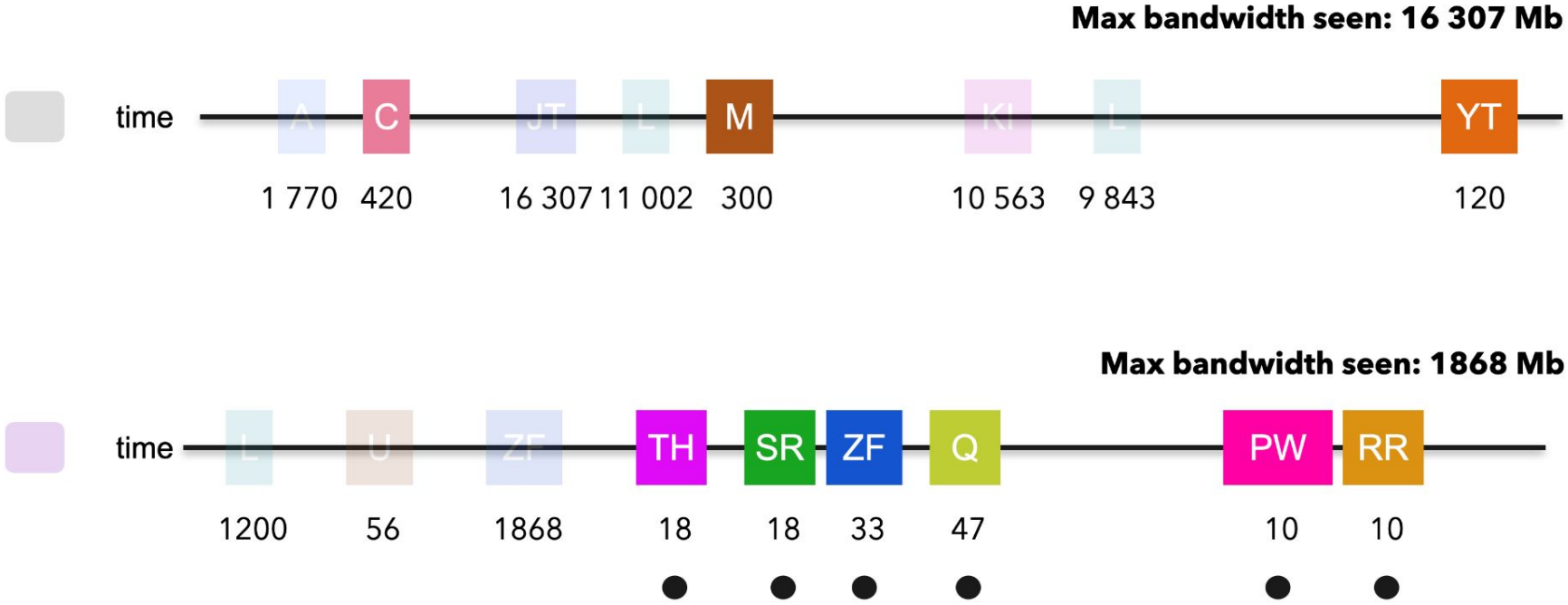


Example Throughput Attribution by Router



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

Checking Router Results vs Time

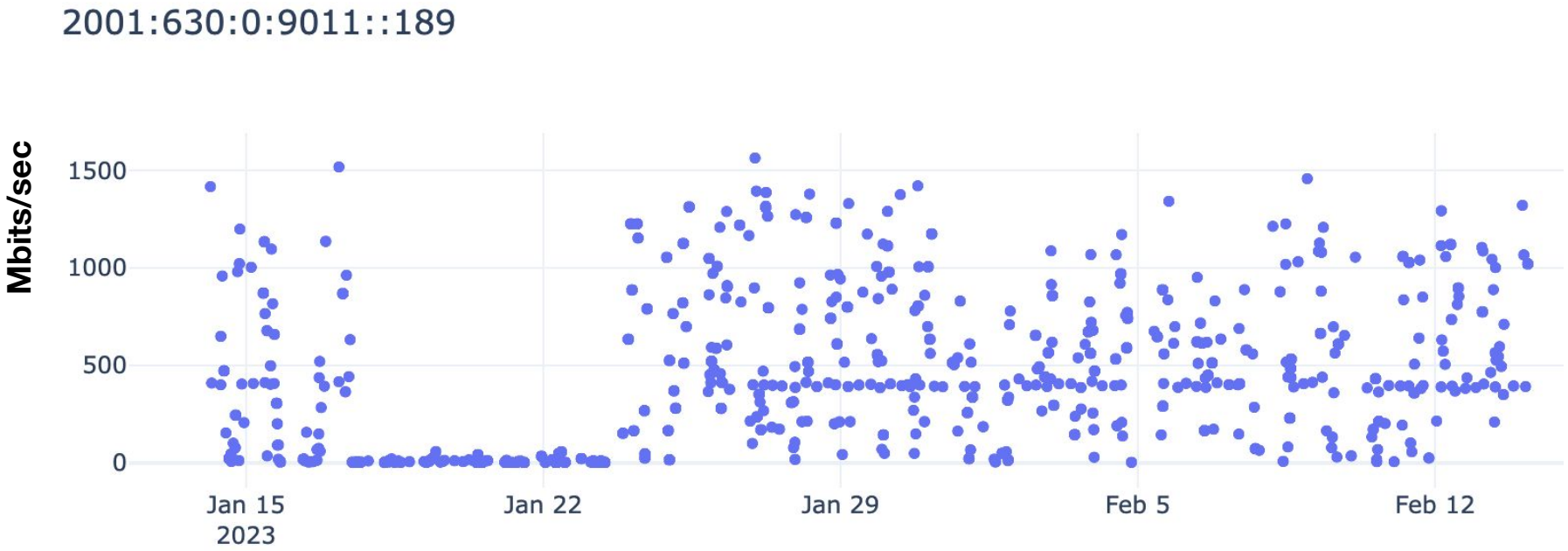


Look for a down trend

(threshold below 10% of the max throughput)

Is everything OK with/around that router?

Initial Example Result: One Router; Throughput vs Time



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

Other Activities / Plans

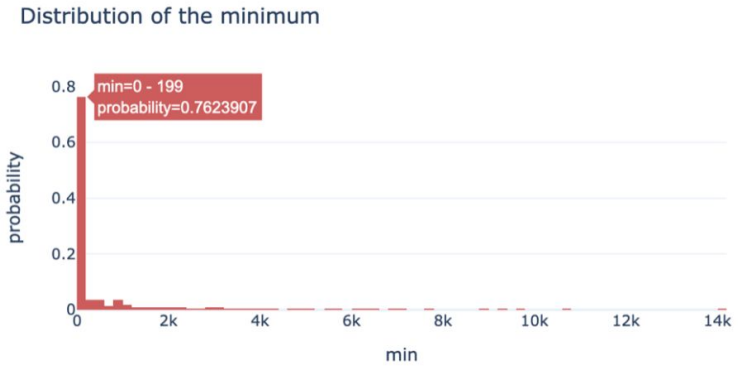
Working to organize and annotate our data for ML/AI work (Petya Vasileva)

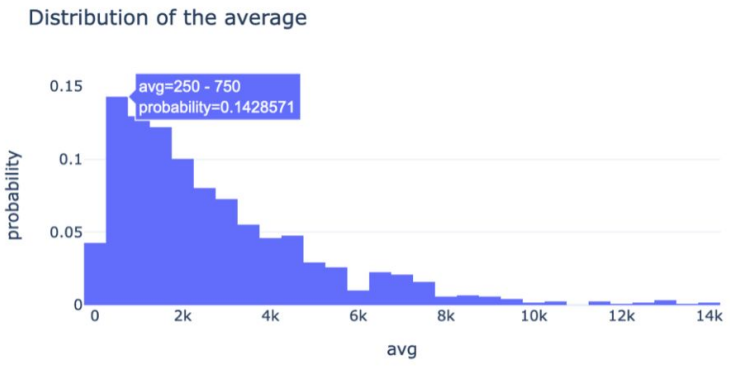
Working with the RNTWG (see previous RNTWG update talk) on identifying and monitoring network traffic details via the SciTags initiative.

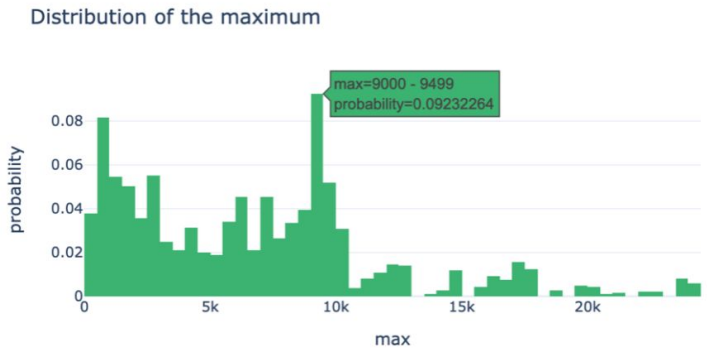
Exploring other network monitoring activities in the perfSONAR space including ARGUS

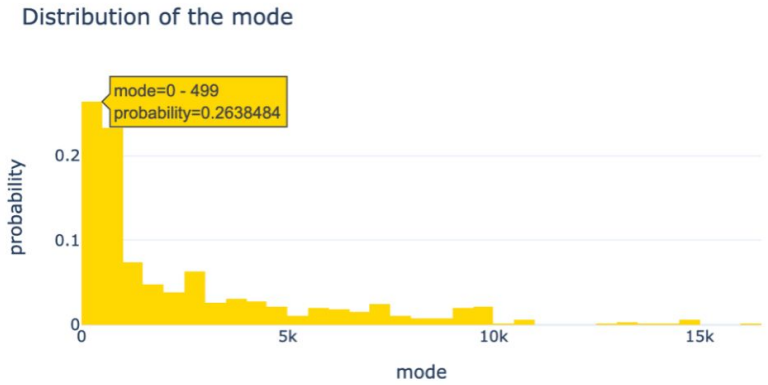
Planning to augment WLCG-CRIC (yesterday's discussion) network meta data (which paths/networks are LHCOPN / LHCONE / Research&Education / Commercial)

Distributions of Throughput









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 :))

Importance of Measuring Our Networks

- End-to-end network issues are difficult to spot and localize
  - Network problems are multi-domain, complicating the process
  - Performance issues involving the network are complicated by the number of components involved end-to-end
  - Standardizing on specific tools and methods focuses resources more effectively and provides better self-support.
- **Network problems can severely impact experiments workflows and have taken weeks, months and even years to get addressed!**
- perfSONAR provides a number of standard metrics we can use
  - Latency, Bandwidth and Traceroute
  - These measurements are critical for network visibility
- **Without measuring our complex, global networks we wouldn't be able to reliably use those network to do science**