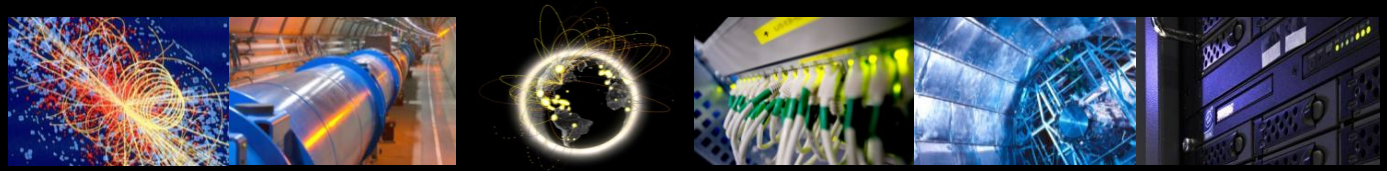
perfSONAR Strategy in Support of DC24 Preparation

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

April 19, 2023

At the #50 LHCOPN/LHCONE Meeting, Prague, Czech Republic

https://indico.cern.ch/event/1234127/



WLCG Worldwide LHC Computing Grid

Open Science Grid

Outline

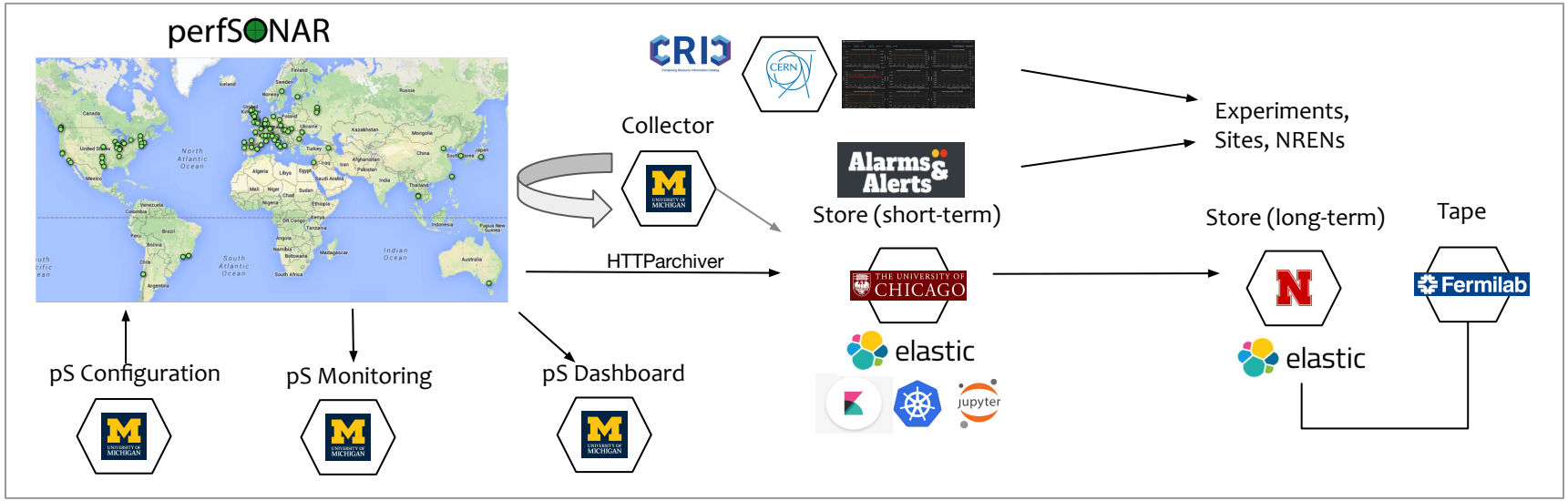
- Infrastructure Reminders
- WLCG Data Challenges: History and Context
- Plans for perfSONAR and associated tools
- Details on tools and problem identification

perfSONAR News

- perfSONAR 5 beta has been out since last summer
  - ElasticSearch as local archive (replacing esmond/Cassandra) + Logstash.
  - Grafana visualisations (dashboards).
  - Toolkit support for latest Debian, RHEL8 compatible systems (Alma).
    - Will require full reinstall (backup **not** needed).
  - **Release of 5.0 was two days ago: April 17, 2023**
  - **OS support:** CentOS 7, Debian 10, Ubuntu 18 and Ubuntu 20
- **Important release information**
  - ESmond (postgresql+cassandra) replaced by Elasticsearch
  - No longer supporting ISO deployments
  - Numerous bug fixes for WLCG, including pscheduler and psconfig issues
  - Within the next two months we expect OS support to include Alma 8, Alma 9, Debian 11 and Ubuntu 22

Network Measurement Platform Evolution

- Our platform collects, stores, configures and transports all network metrics
  - Distributed deployment - operated in collaboration
- Planned evolution based on the perfSONAR 5 already partially implemented.
  - Directly publishing results from perfSONARs to ES@UC (have test_* indices)
  - High-level services provided to the experiments/users (Alerting, Dashboards)



WLCG Data Challenges

The **WLCG Data Challenges** are a ~biennial series of increasingly-complex exercises which started in 2021 and are aimed at demonstrating readiness at the HL-LHC scale.

Such exercises are only meaningful if there is worldwide, multi-experiment participation; they are planned and managed by the WLCG DOMA (Data Organization, Management and Access) group (See DC24 update from Feb MB mtg).

These data challenges provide many benefits, allowing sites, networks and experiments to evaluate their progress, motivate and validate their developments in hardware and software and show readiness of technologies at suitable scale.

WLCG DC24

The next data challenge (Data Challenge 2024 = DC24) will be ~March 2024 and targeting ~20-25% of HL-LHC

Some of the important **network** related areas we need to work on for DC24:

- 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
  - May need to extend what ESnet has been doing in the US to other regions
- **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!

Plans for perfSONAR and the Analytics Platform

For DC24, we want to, debug and repair our networks using perfSONAR well in advance of March 2024

With the release of perfSONAR 5.0 on April 17, 2023, we are almost ready to initiate a campaign to get perfSONAR upgraded WLCG-wide

- Within a month or two we should have perfSONAR 5.0 available for EL9
- Target is to get sites to upgrade to 5.x using a 5.x kernel ASAP
- Sites should review their hardware to ensure it is still matching storage

By early Fall 2023, we want to review and address issues being identified by our persistent perfSONAR testing, e.g., firewalls, packet loss, non-optimal routing, low throughput, infrastructure issues, etc.

Infrastructure: Network Data Pipeline Transition

- We are updating our network data pipeline which gathers perfSONAR data.
- This is primarily motivated by two things
  - The operational experience with using RabbitMQ (high latency; too many parts)
  - The perfSONAR v5 change to the "default" Measurement Archive (MA) to Elasticsearch
- We are in the process of transitioning away from using a message bus and instead sending data directly to Elasticsearch from each perfSONAR toolkit.
  - **Pros**: Reduces latency, simplifies pipeline with less components required, matches direction the perfSONAR developer have chosen allowing us to utilize and benefit from their work, should be easier to operate and monitor
  - **Cons**: All dependencies on the RabbitMQ bus must be transformed to use the HTTP archiver. To maintain the ability to grab data for those toolkits that cannot "push" also requires updating the perfsonar_collector to send data directly to Elasticsearch. (**SOLVED**)
- **TIMELINE**: Estimate is end of **June 2023** (Approximately 2 months after working pS 5 release)

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

Components

Database
Elasticsearch

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

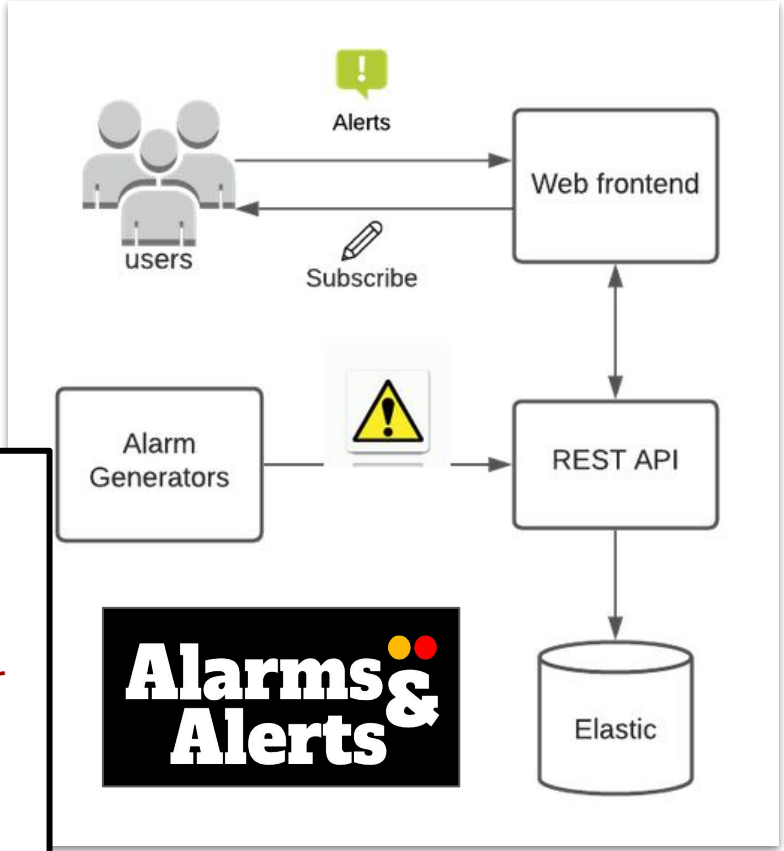
Deployment
Docker, K8s, Helm (soon)

Authentication
Globus InCommon

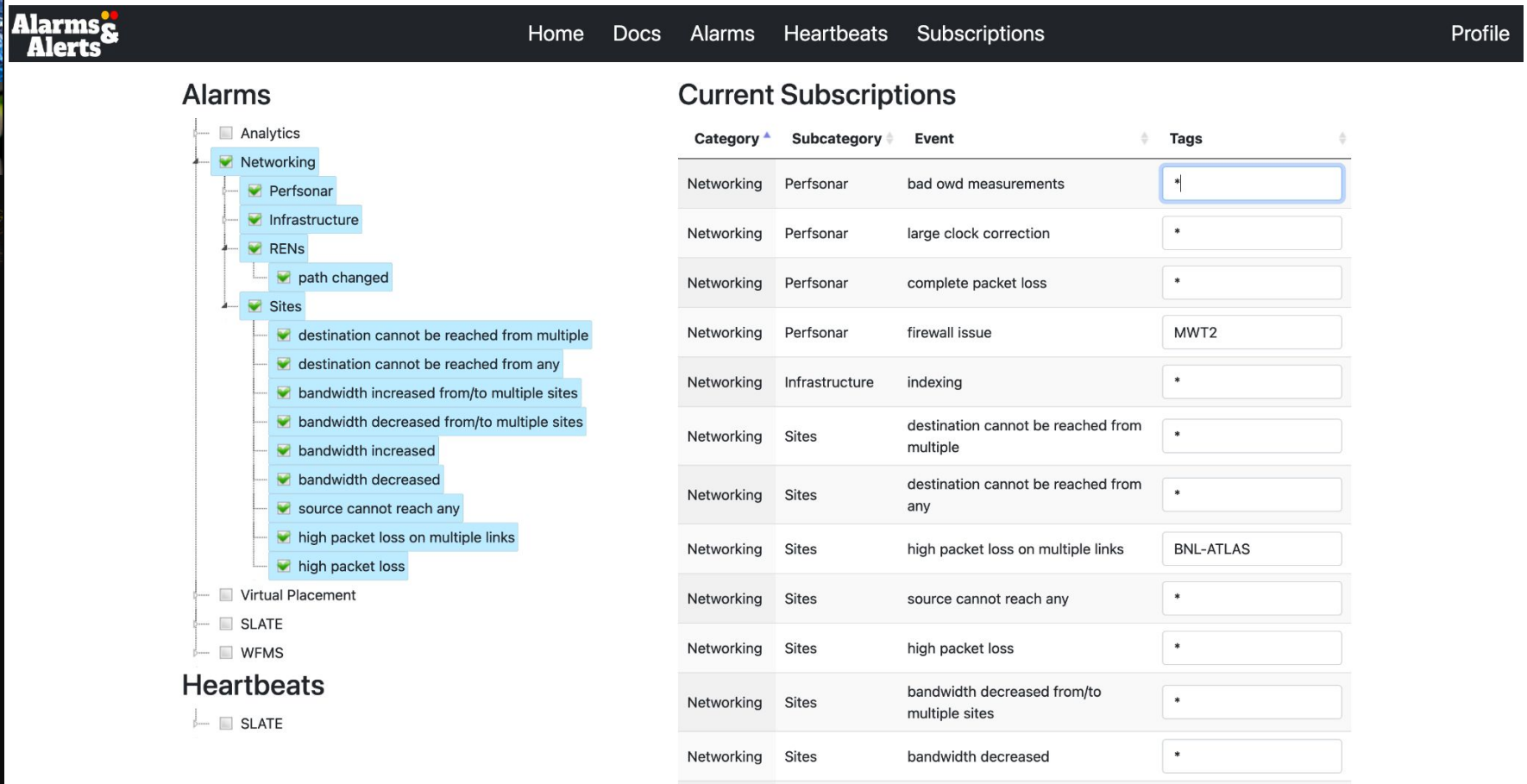
Authorization
API key

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



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)

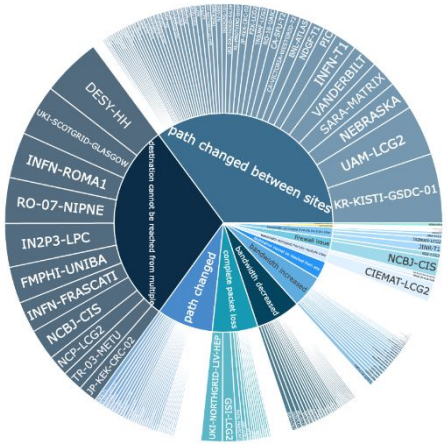
perfSONAR Toolkit Information | Kibana: Packet Loss in OSG/WLCG | Kibana: Packet Loss Tracking | MEPHi Tracer: Traceroute explorer | Alarms description

ps dash | SITES OVERVIEW | SEARCH ALARMS | EXPLORE PATHS

Selected site: UKI-NORTHGRID-LIV-HEP

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

ps dash | SITES OVERVIEW | SEARCH ALARMS | EXPLORE PATHS



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)

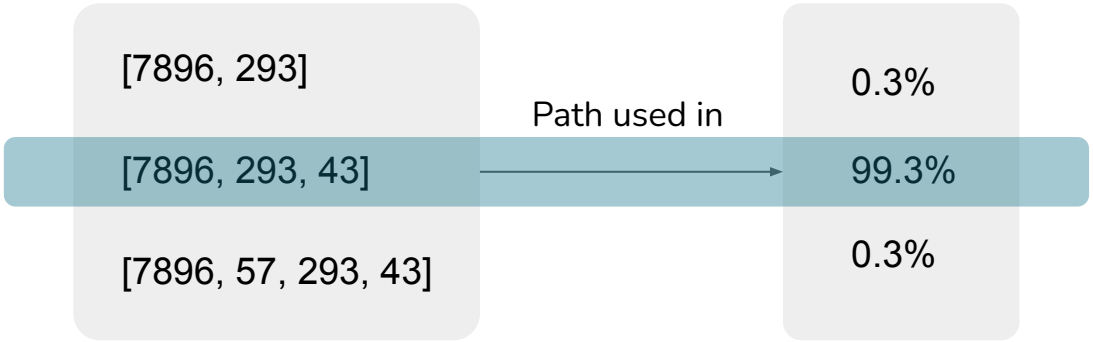
4415 traversed nodes Node=”router”

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

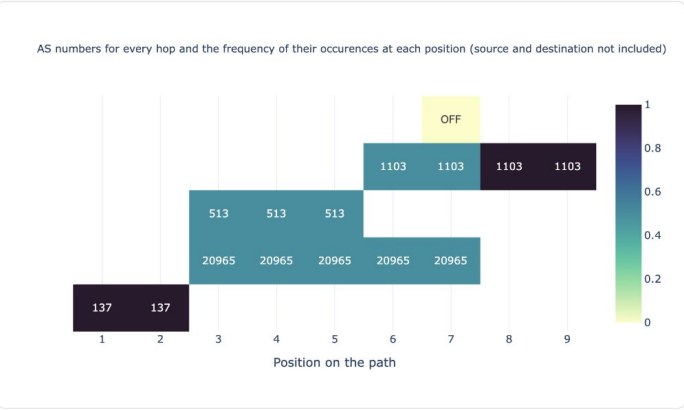
ALTERNATIVE PATHS

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

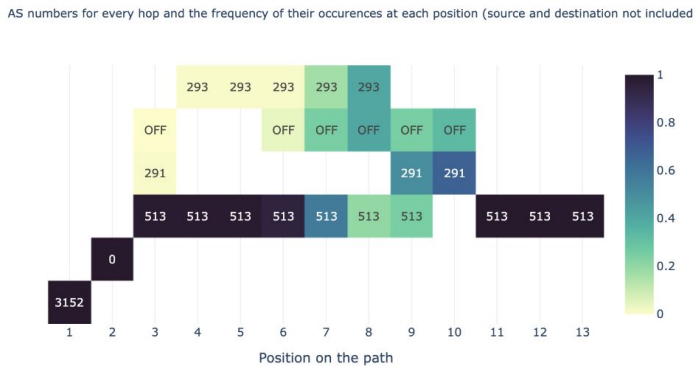
ALTERNATIVE PATHS

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

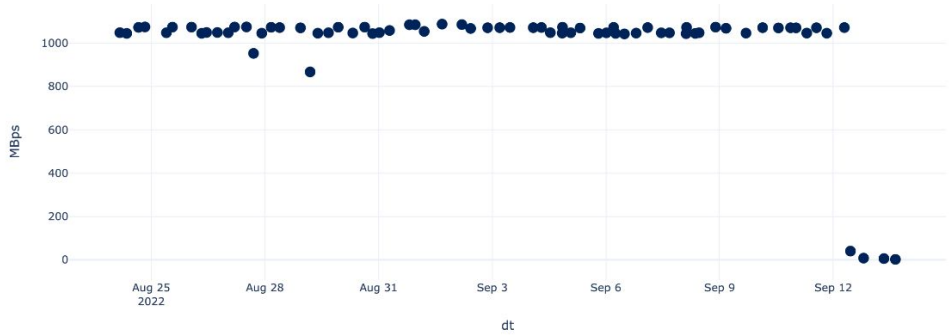
NDGF-T1

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

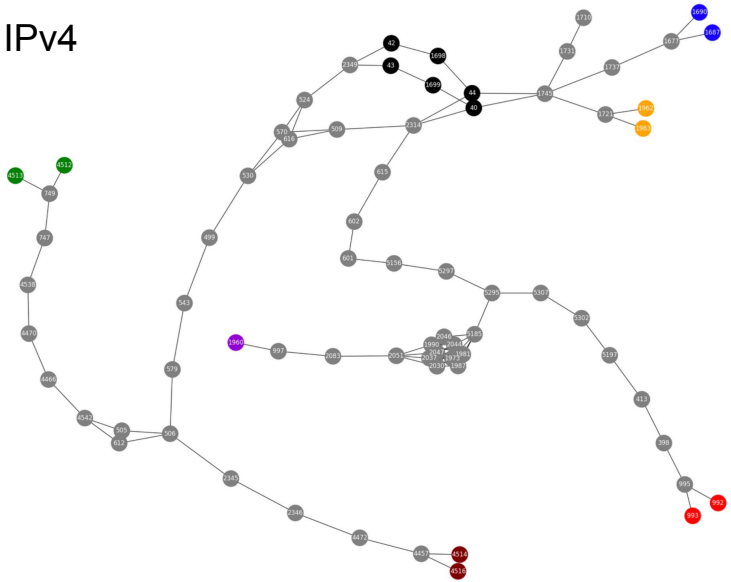
At position	Typically goes through	Changed to
4	**11537** INTERNET2-RESEARCH-EDU, US	**7029** WINDSTREAM, US

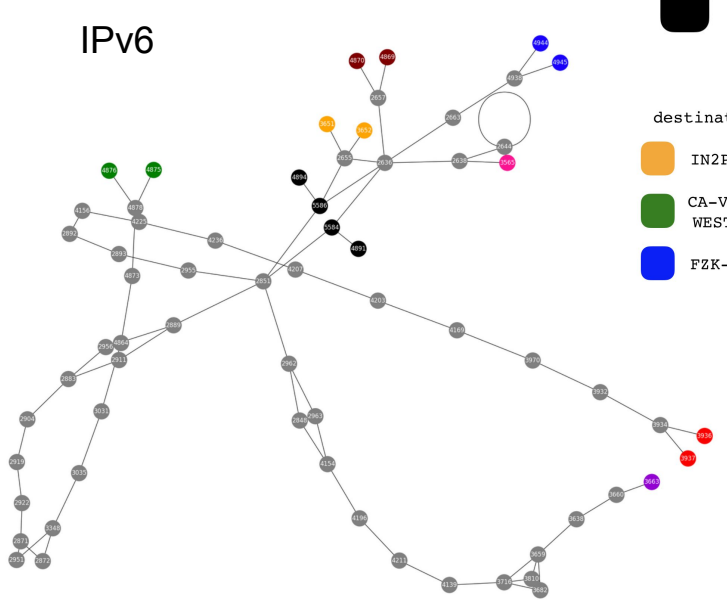
At position	Typically goes through	Changed to
7	**11537** INTERNET2-RESEARCH-EDU, US	**6939** HURRICANE, US

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

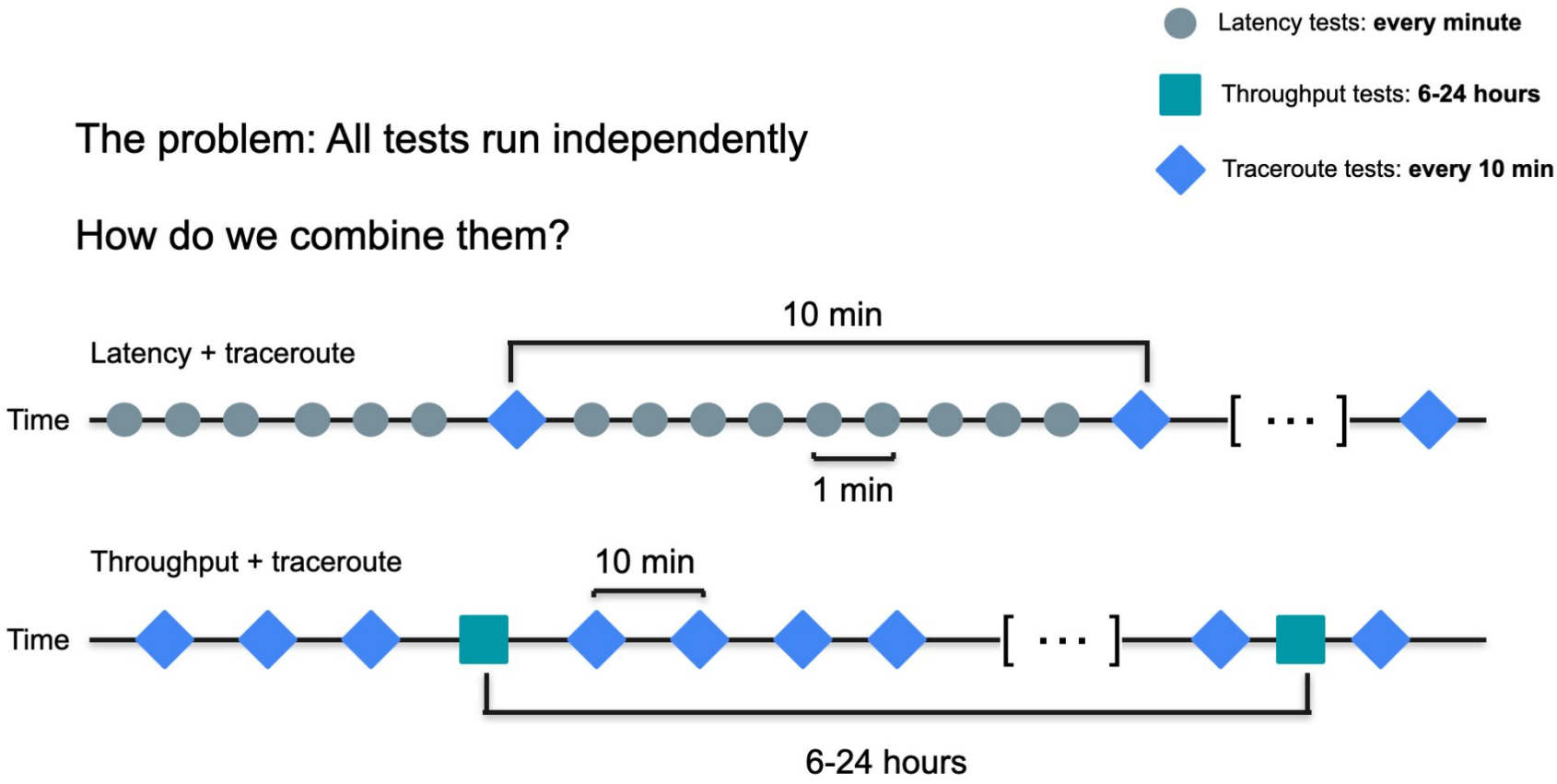
Challenges and Ongoing Work

Paths differ significantly



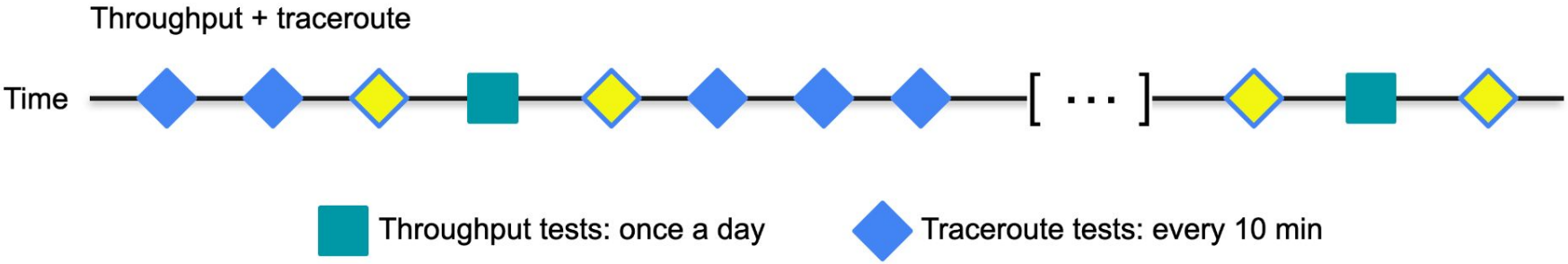


Correlating Tests with Paths: Two Timescales



Connecting Throughput to Traceroute

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

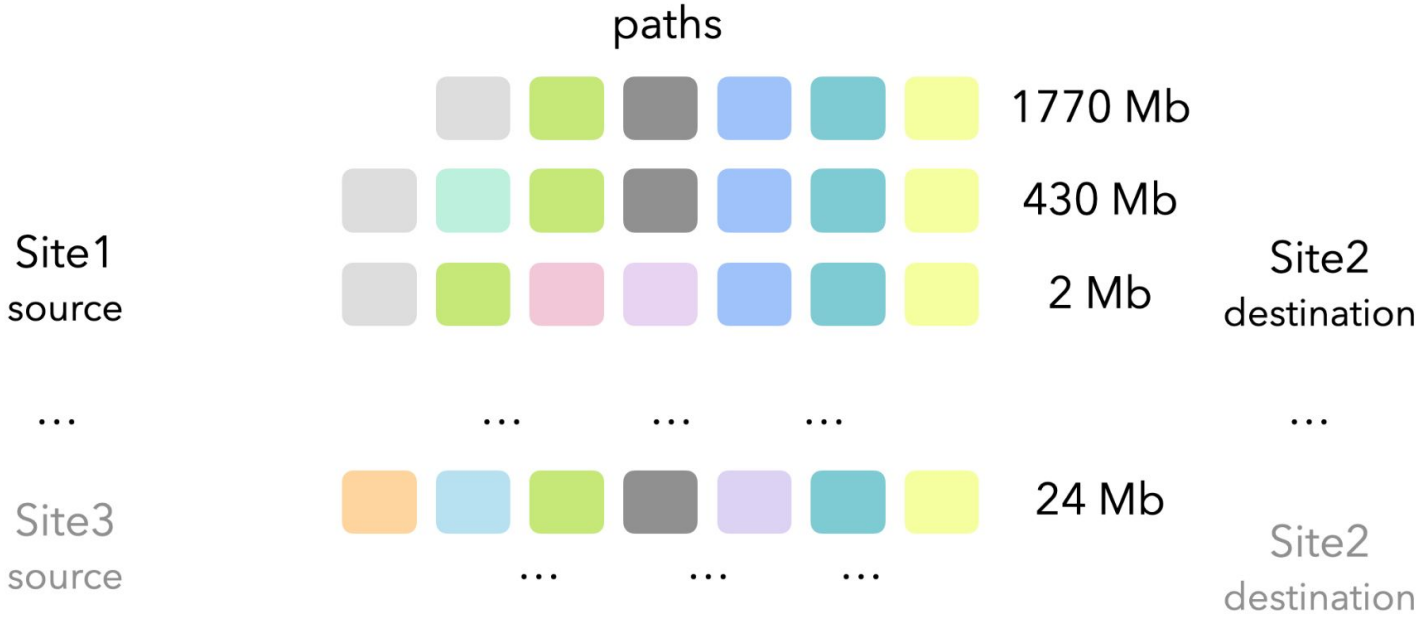


Our starting choice: Use **both** tracepaths (just before; just after) as valid paths and attribute BW to both.

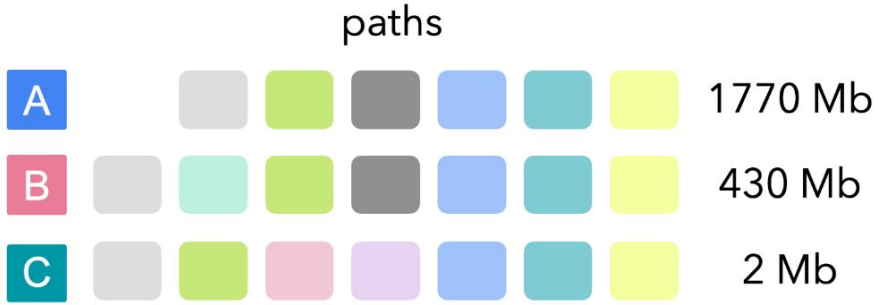
Have to see if this is superior to just using the last measured route before the measurement…

Attaching Throughput Results to Sets of Routers/Links

Each colored box represents a specific router along the path

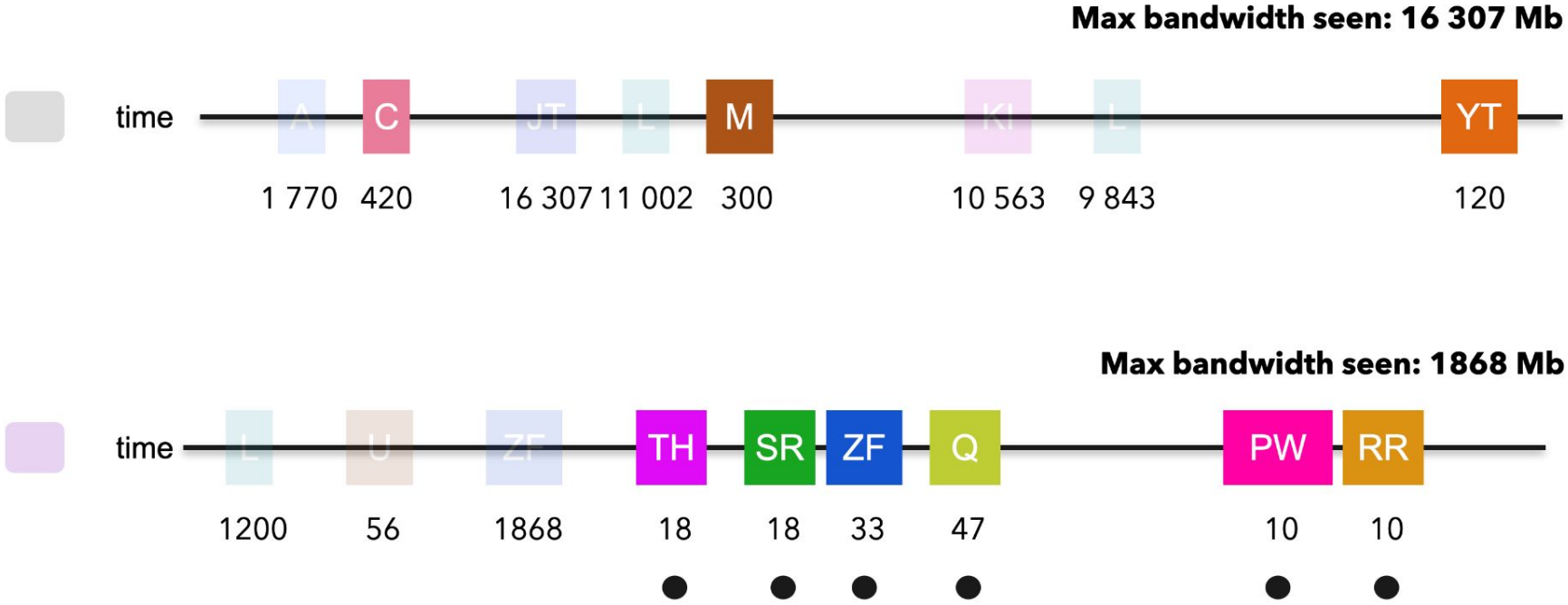


Example Throughput Attribution by Router



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

Checking Router Results vs Time

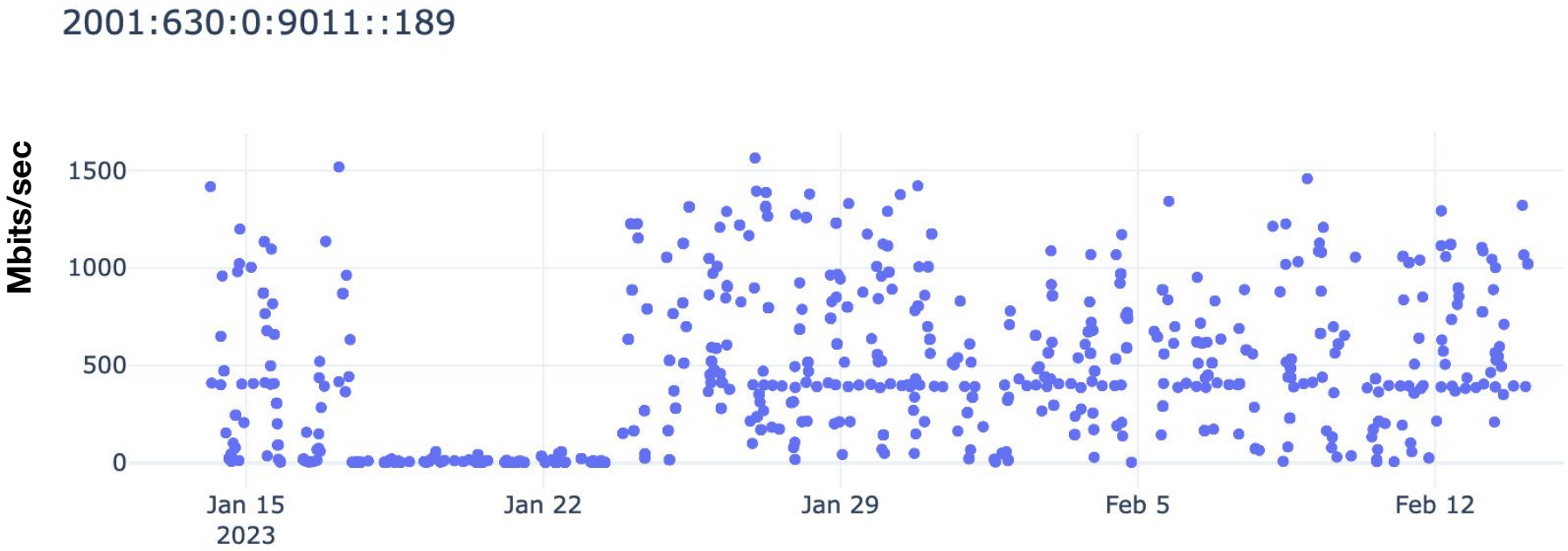


Look for a down trend

(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)

Summary

- OSG in collaboration with WLCG operates a comprehensive network monitoring platform
  - Provides data and feedback to LHCOPN/LHCONE, HEPiX, WLCG and OSG communities
- **We are preparing for DC24**: plan to use perfSONAR to proactively debug our main network links before DC24
- Developing high-level services based on perfSONAR measurements that will help sites, experiments and R&Es receive targeted alarms/alerts on existing issues in the infrastructure
- We have to continue to watch our network monitoring infrastructure as it is a complex system with lots of areas for issues to develop.

Questions / Discussion?

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:

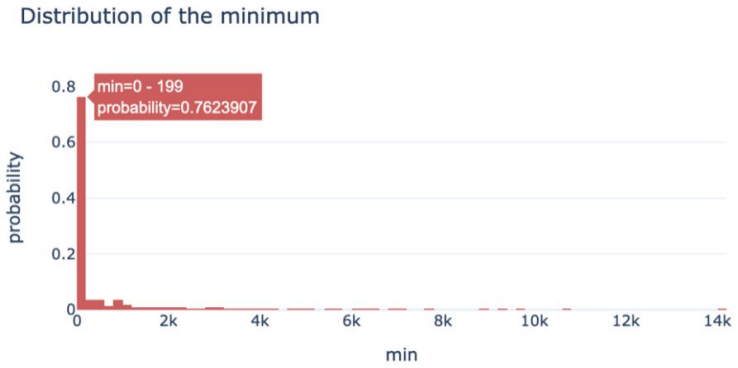
- OSG: NSF MPS-1148698
- IRIS-HEP: NSF OAC-1836650

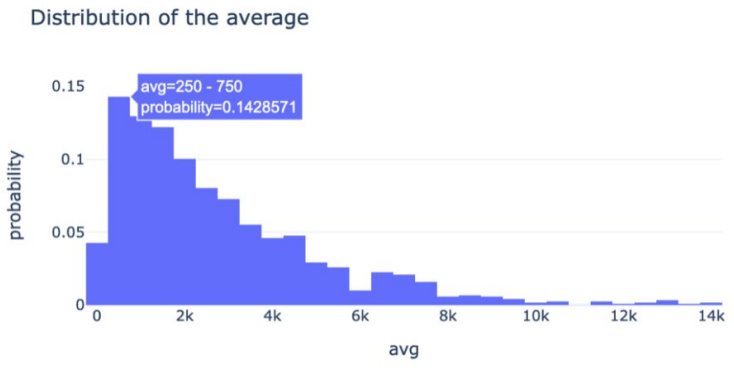
Useful URLs

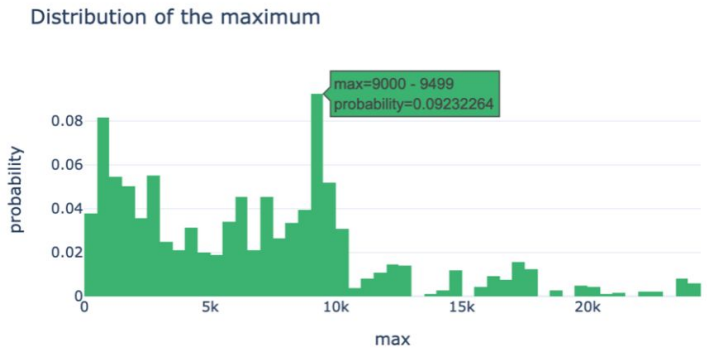
- OSG/WLCG Networking Documentation
  - https://opensciencegrid.github.io/networking/
- perfSONAR Infrastructure Dashboard
  - https://atlas-kibana.mwt2.org:5601/s/networking/goto/9911c54099b2be47ff9700772c3778b7
- WLCG DOMA DC24 plans
  - https://indico.cern.ch/event/1225415/contributions/5155042/attachments/2593516/4476291/Data%20Challenge%202024.pdf
- 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

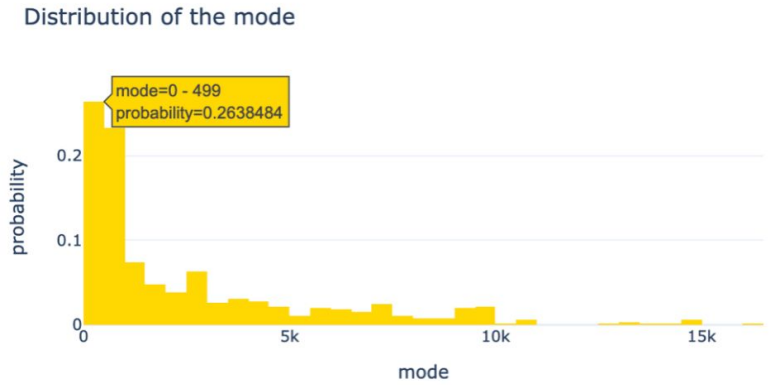
Backup Slides Follow

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**