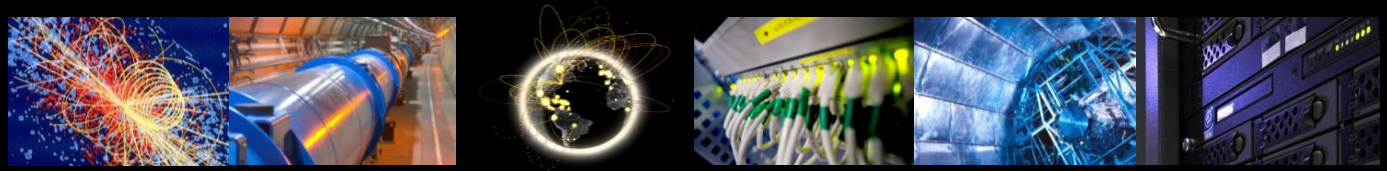
perfSONAR Plans for DC24

Shawn McKee / University of Michigan
on behalf of WLCG Network Throughput WG and DOMA

November 9, 2023

Data Challenge 2024 Workshop

https://indico.cern.ch/event/1307338/



WLCG Worldwide LHC Computing Grid

Open Science Grid

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 can 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.

Plans for perfSONAR and the Analytics Platform

For DC24, we want to, debug and repair our networks using perfSONAR well in advance of February 12, 2024

Original Plan (From Spring 2023)

- In Summer 2023, have a robust, reliable perfSONAR distribution to allow sites to upgrade base OS to EL9 (or similar) AND update hardware if it no longer matches storage.
- In Fall 2023, start utilizing perfSONAR metrics to understand issues being identified by our persistent perfSONAR testing, e.g., firewalls, packet loss, non-optimal routing, low throughput, infrastructure issues, etc.

We registered a project in the WLCG DOMA CERNbox…

perfSONAR Debugging DC24 Project

Description: Utilize the WCLG perfSONAR instances and associated dashboards and analytics to identify severe network issues associated with our sites and follow up with tickets till resolved.

Work: We need to first upgrade and harden our perfSONAR deployment, moving to version 5.0.5+ and updating perfSONAR hardware where it no longer represents storage capability accurately.
- **Timeline**: August 2023-October 2023 Initiate upgrade campaign and expect at least two months to get most sites updated. [**STATUS**: Not officially started]
- **Who**: Shawn McKee, Marian Babik, WLCG Network Throughput WG

Work: Improve our analytics to better identify *network* issues.
- **Timeline**: October 2023 to have basic analytics that reliably identify severe network issues
- **Who**: Petya Vasileva, Jan Perina, IRIS-HEP Student Fellows [**STATUS**: Almost done]

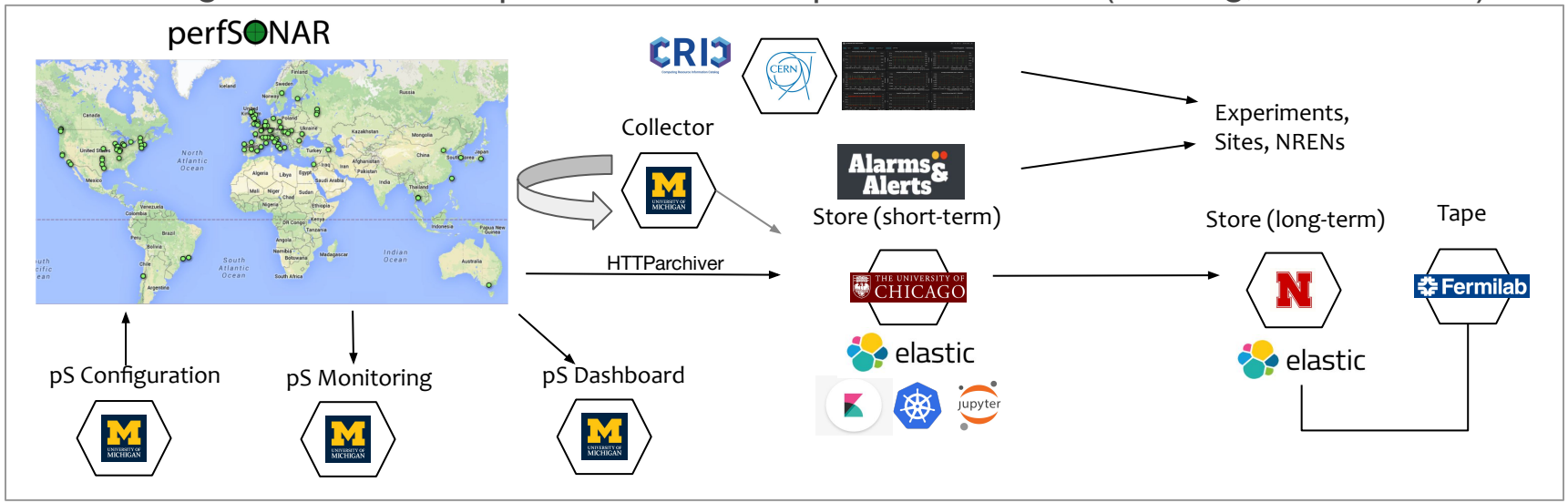
Work: Utilize available tools and perfSONAR results to identify, follow-up and fix network issues
- **Timeline**: November 2023 - January 2024
- **Who**: WLCG Network Throughput WG, WLCG Ops [**STATUS**: Delayed]

Metrics: We need to track identified network issues, resolution of each issue (true network issue, fixed or not, sites impacted). Will track via the WLCG Throughput page

perfSONAR Infrastructure

Network Measurement Platform Overview

- 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 (via HTTParchiver)
  - High-level services provided to the experiments/users (Alerting, Dashboards)



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
- We have developed some higher level tools to create, manage and explore alarms & alerts:
  - User interface to subscribe is **AAAS** (ATLAS Alerting and Alarming Service)
  - Tool to explore alerts is **pS-Dash** (Plotly base perfSONAR dashboard UI tool)

pSDash (perfSONAR Dashboard)

Alarms & Alerts Service

Components

Database
Elasticsearch

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

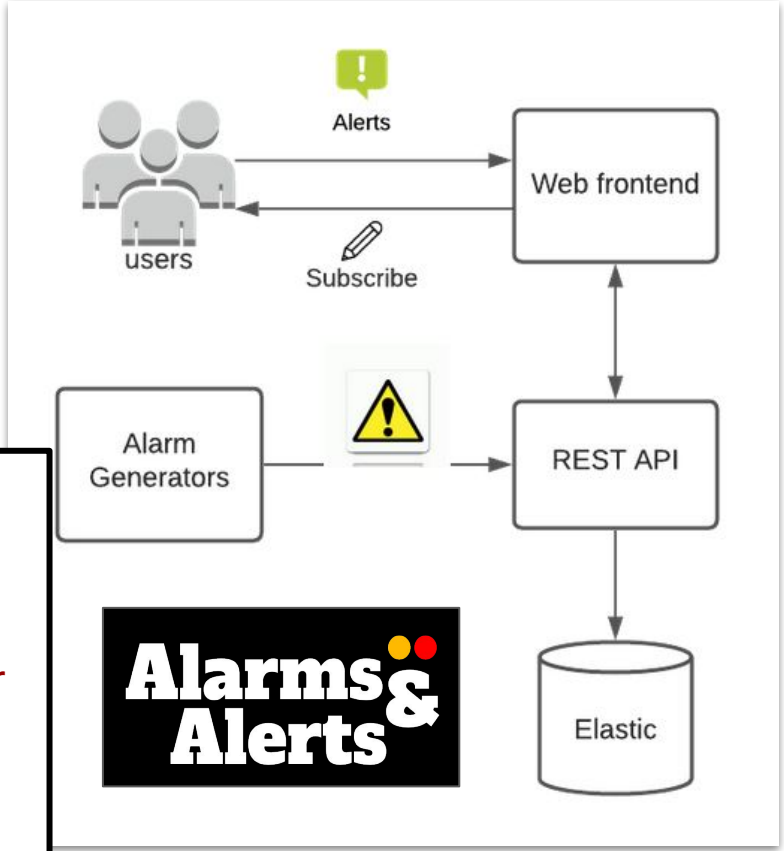
Deployment
Docker, K8s, Helm (soon)

Authentication
Globus InCommon

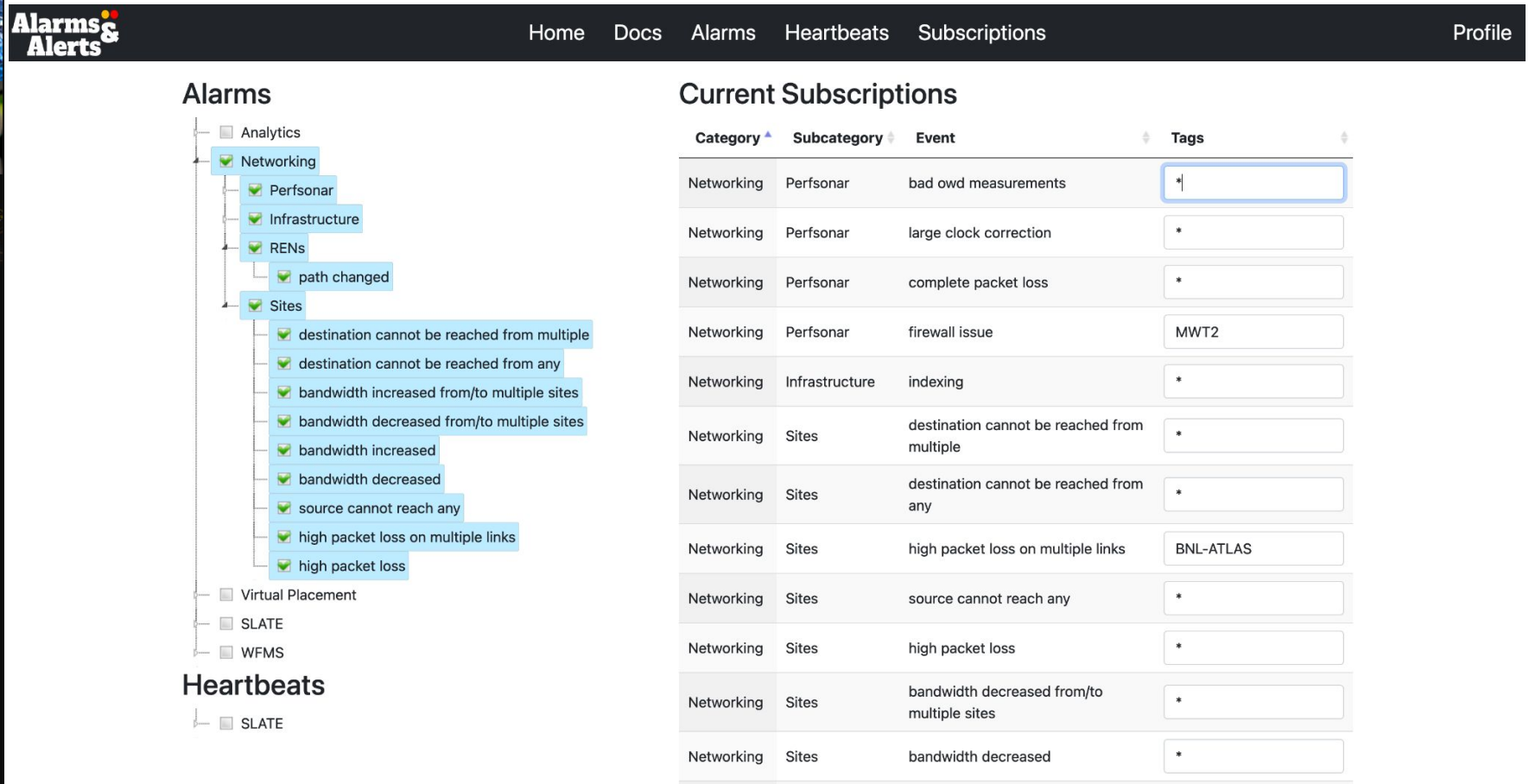
Authorization
API key

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

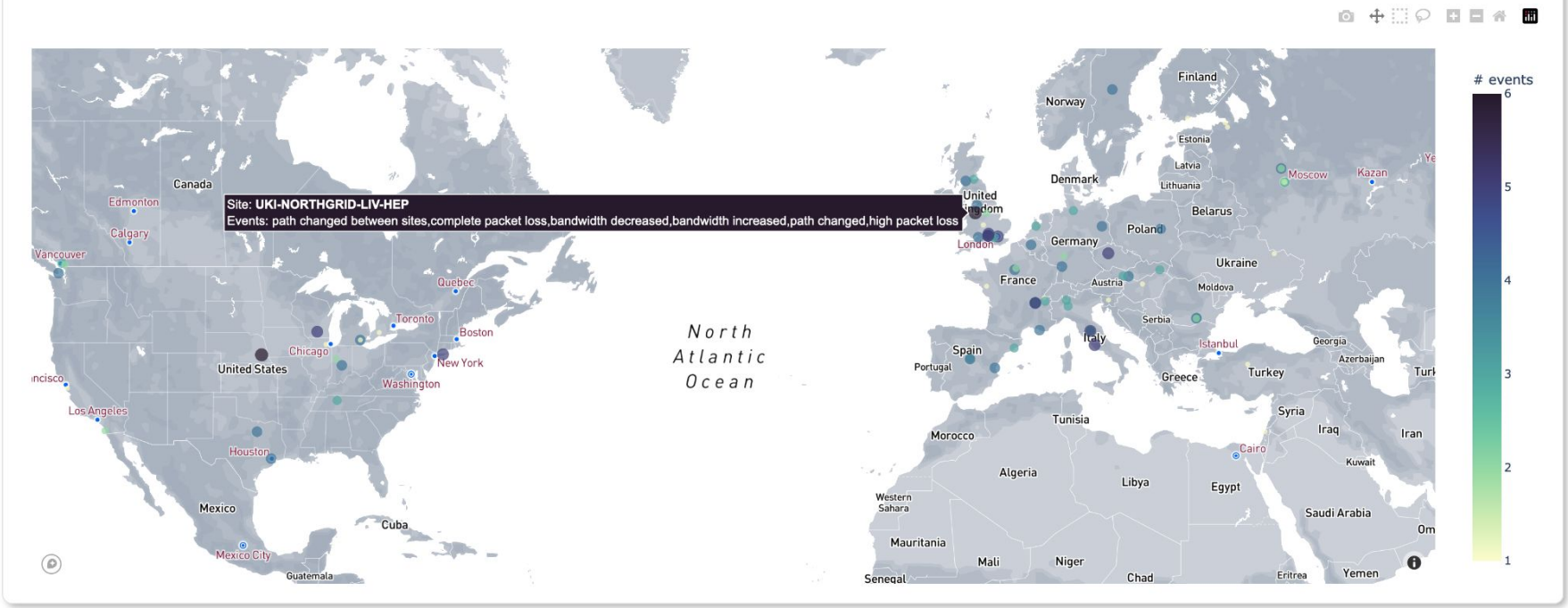


perfSONAR Toolkit Information | Kibana: Packet Loss in OSG/WLCG | Kibana: Packet Loss Tracking | MEPHi Tracer: Traceroute explorer | Alarms description

ps dash | SITES OVERVIEW | SEARCH ALARMS | EXPLORE PATHS

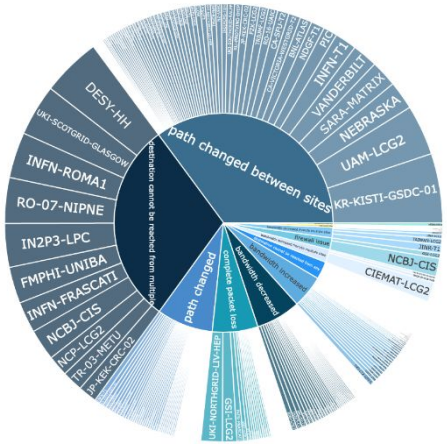
Selected site: UKI-NORTHGRID-LIV-HEP

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



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



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

Project Status

Work Item #1: perfSONAR Upgrade Campaign

STATUS: Not Started. We had originally planned to undertake a perfSONAR upgrade campaign during the summer of 2023 with the following goals:

- Ensure sites upgraded their perfSONAR to the most recent
- Get sites to use an EL9 OS (much newer, more capable kernel)
- Have sites evaluate the perfSONAR hardware, replacing systems that no longer match storage servers in terms of network interface

perfSONAR 5.0.0 was available in May 2023 but was not suitable for our needs

- Initial version didn't support EL9
- Significant bugs prevented reliable deployment and use
- Even with version 5.0.5 (current) we have some major issues with reliability
  - Of ~250 instances only ~50 reliably operate with data appearing in their local MA

Remediation: Worked with perfSONAR devs to debug issues. Possible fix due out next Week (Nov 13-17). Have to start notifying sites about updating…

Work Item #2: Create Analytics Tools to Identify Issues

STATUS: Almost ready. We have had a regular meeting weekly to discuss networking and analytics for the last year.

This group includes various research project members (primarily in the US), members of various WLCG working groups and sites.

The group has been focused on enabling perfSONAR metrics (and other network data) to be used to proactively identify and alert on network issues.

The AAAS (https://psa.osg-htc.org/) system allows anyone to subscribe for alerts. The pS-Dash (https://ps-dash.uc.ssl-hep.org/) application provides a web tool to explore alerts and network data

Plan is to use what we have to start identifying problems and opening tickets.

Work Item #3: Using perfSONAR find/fix network issues

STATUS: Delayed. Because we haven't gotten a suitable perfSONAR version available, we have been delayed in starting to find and ticket for network issues.

Time is short and we basically need to have upgraded and reliable perfSONAR instances running by December. Sites may also need to replace old or unrepresentative hardware to give us the best data.

Plan is to start identifying the issues we can in December and work up until DC24.

This should become a standard part of our toolkit going forward and should help us keep our networks functional and operating well for our use cases.

Estimated Timeline and Activity to DC2

We hope to have a perfSONAR update next week to address the main issues for operational resiliency.

- Once this is ready and verified, we will notify sites about upgrading, including recommendations to match storage and use EL9 (Nov 2024)
- In parallel, we will start gathering a list of major network issues that we can identify (Finish initial list by end of November)
- The WLCG Throughput WG will need to start opening network tickets by site in December
- Analytics will continue to be used to identify new problems and verify resolution of ticketed problems (January 2024)
- Goal is to fix as many of the most important network problems by February 12 (start of DC24)

Summary

- **We are preparing the networks for DC24**: plan to use perfSONAR to proactively debug our main network links before DC24
  - We have issues still to address to ensure we have diagnostic data covering the majority of our networks…
- Developing, improving and hardening 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

- WLCG DOMA project descriptions
  https://cernbox.cern.ch/files/link/public/aCITXJenZxpF5qw?tiles-size=1&items-per-page=100&view-mode=resource-table
- 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

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)

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)

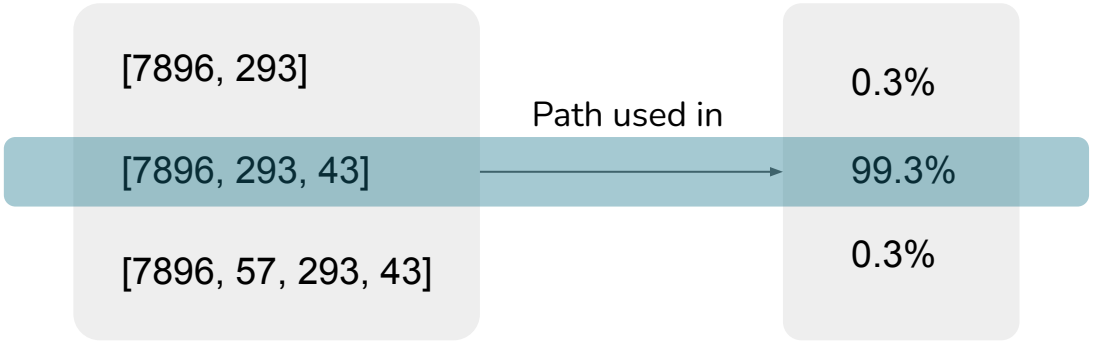
4415 traversed nodes Node=”router”

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

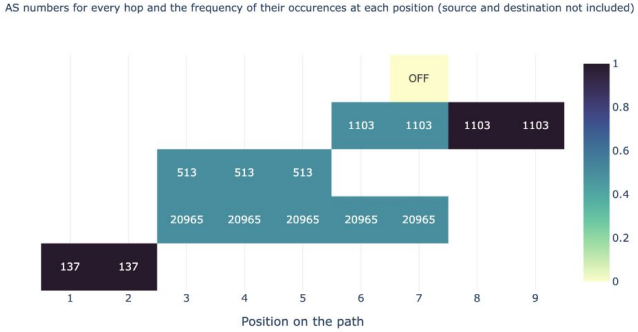
ALTERNATIVE PATHS

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

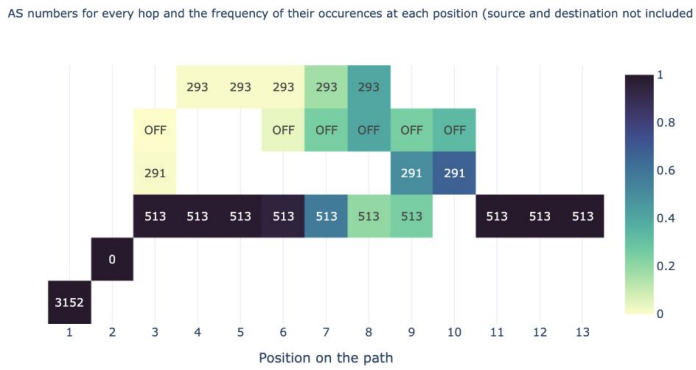
ALTERNATIVE PATHS

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

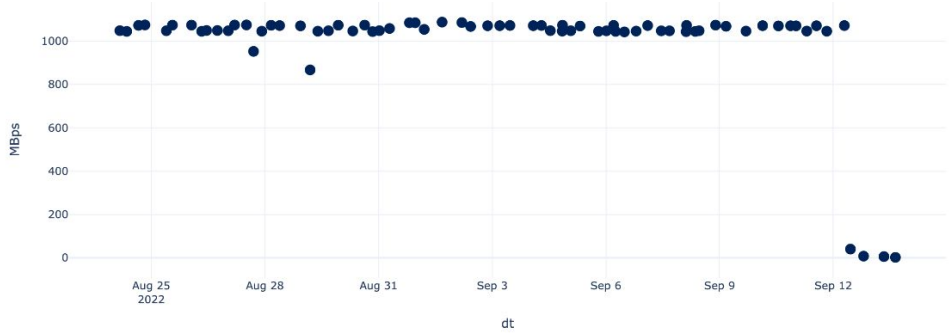
NDGF-T1

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

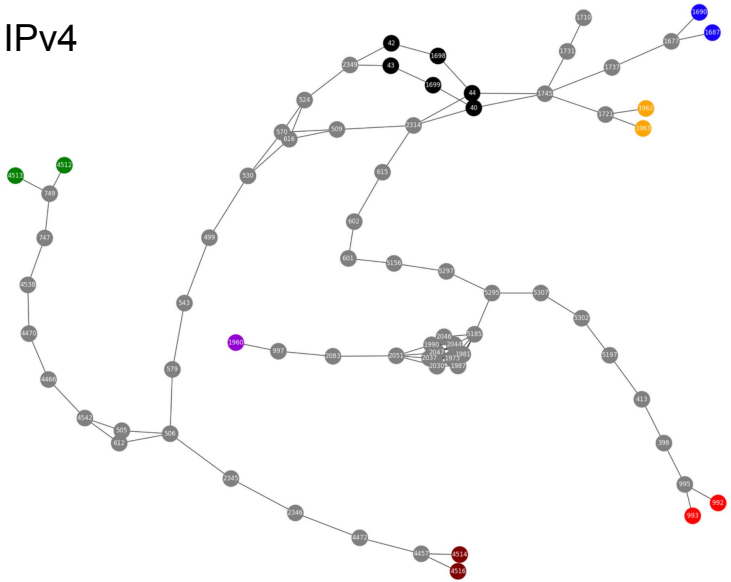
At position	Typically goes through	Changed to
4	**11537** INTERNET2-RESEARCH-EDU, US	**7029** WINDSTREAM, US

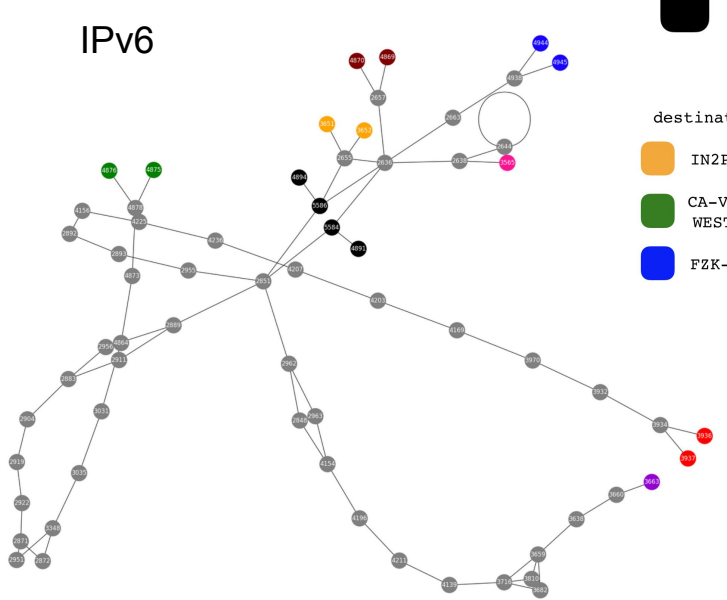
At position	Typically goes through	Changed to
7	**11537** INTERNET2-RESEARCH-EDU, US	**6939** HURRICANE, US

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

Challenges and Ongoing Work

Paths differ significantly





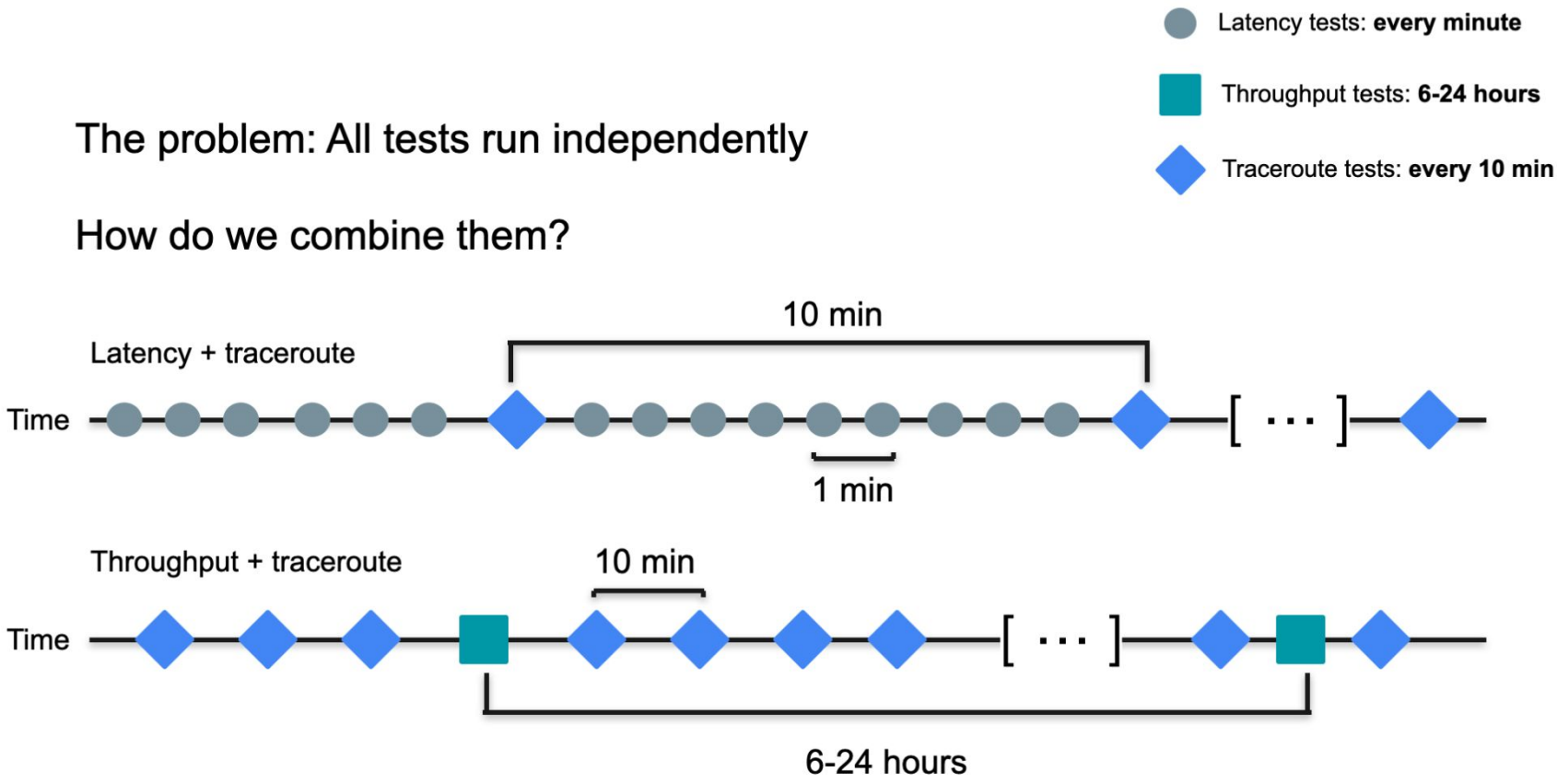
source

BNL-ATLAS
closest 2 hops

destinations

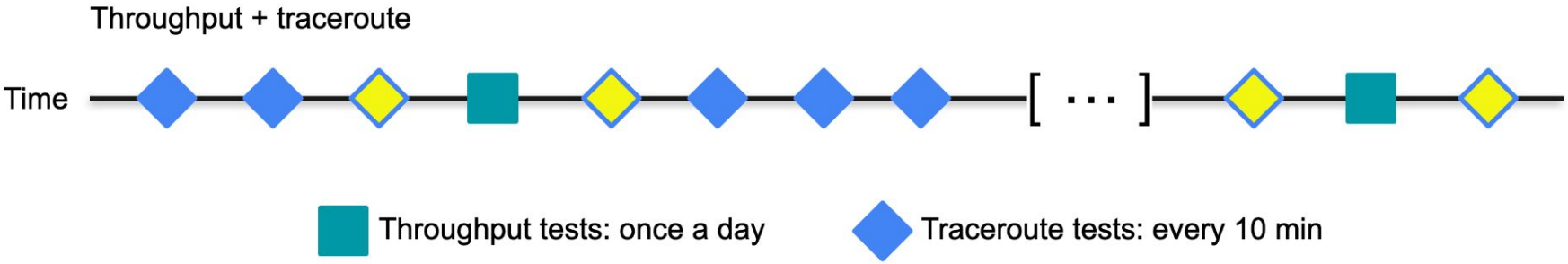
IN2P3-CC
IFIC-LCG2
CA-VICTORIA-WESTGRID-T2
TRIUMF-LCG2
FZK-LCG2
IN2P3-LPSC
RAL-LCG2

Correlating Tests with Paths: Two Timescales



Connecting Throughput to Traceroute

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

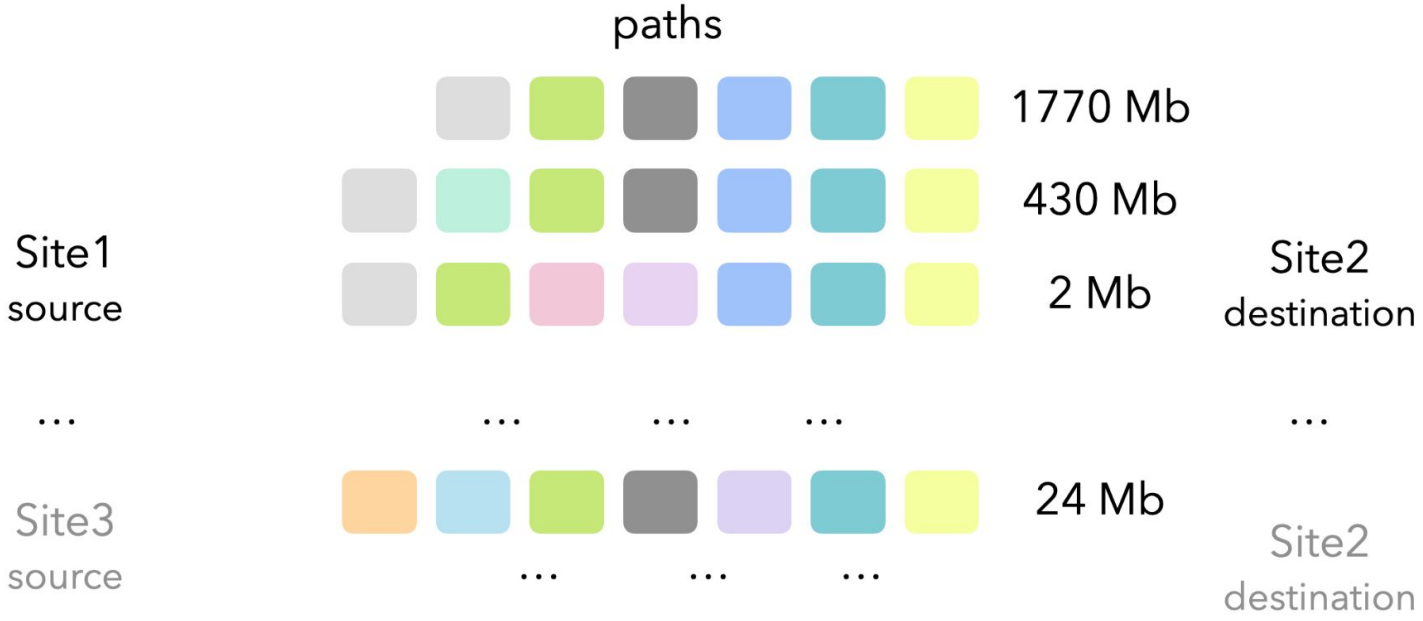


Our starting choice: Use **both** tracepaths (just before; just after) as valid paths and attribute BW to both.

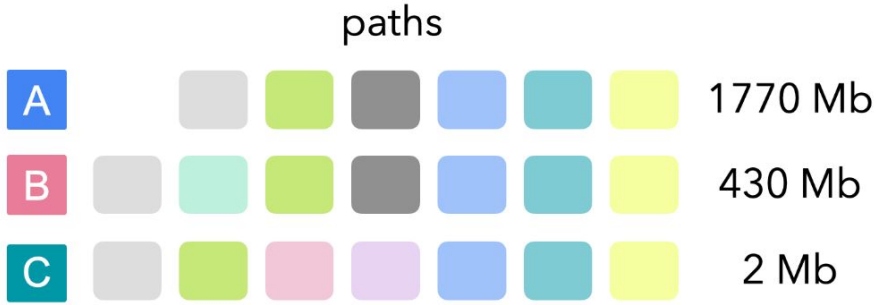
Have to see if this is superior to just using the last measured route before the measurement…

Attaching Throughput Results to Sets of Routers/Links

Each colored box represents a specific router along the path

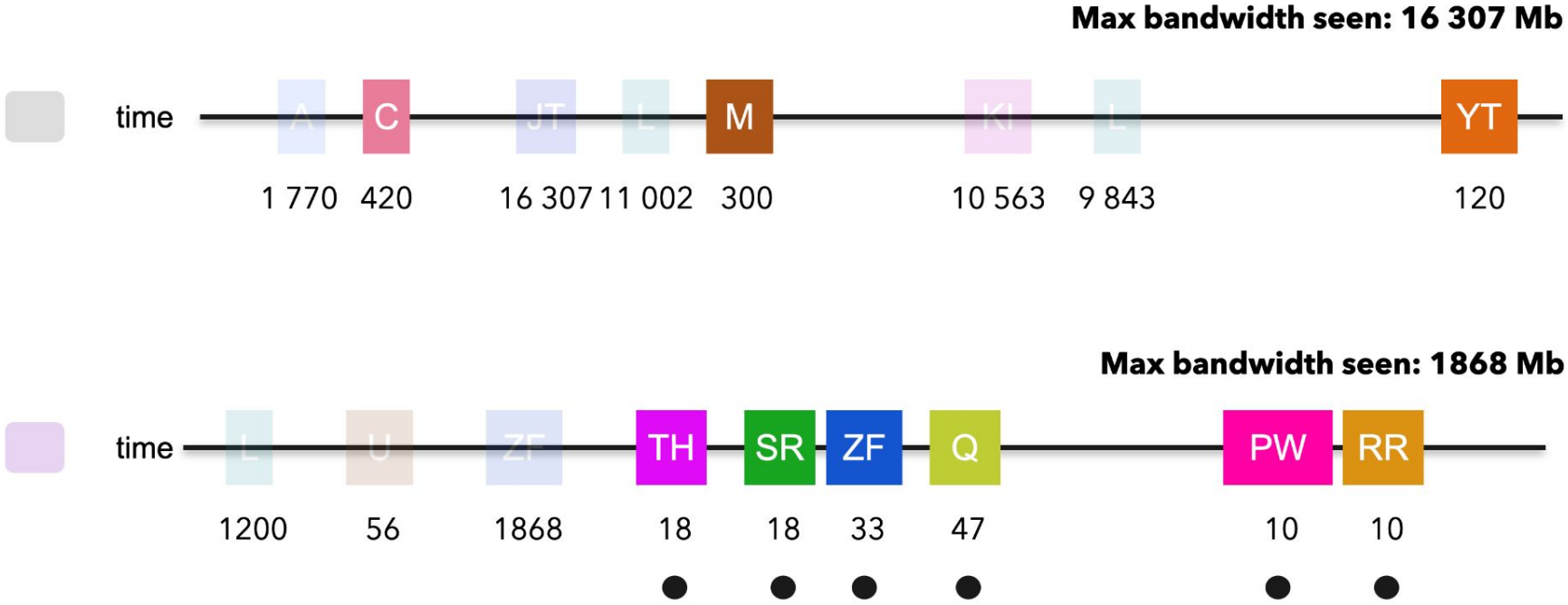


Example Throughput Attribution by Router



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

Checking Router Results vs Time

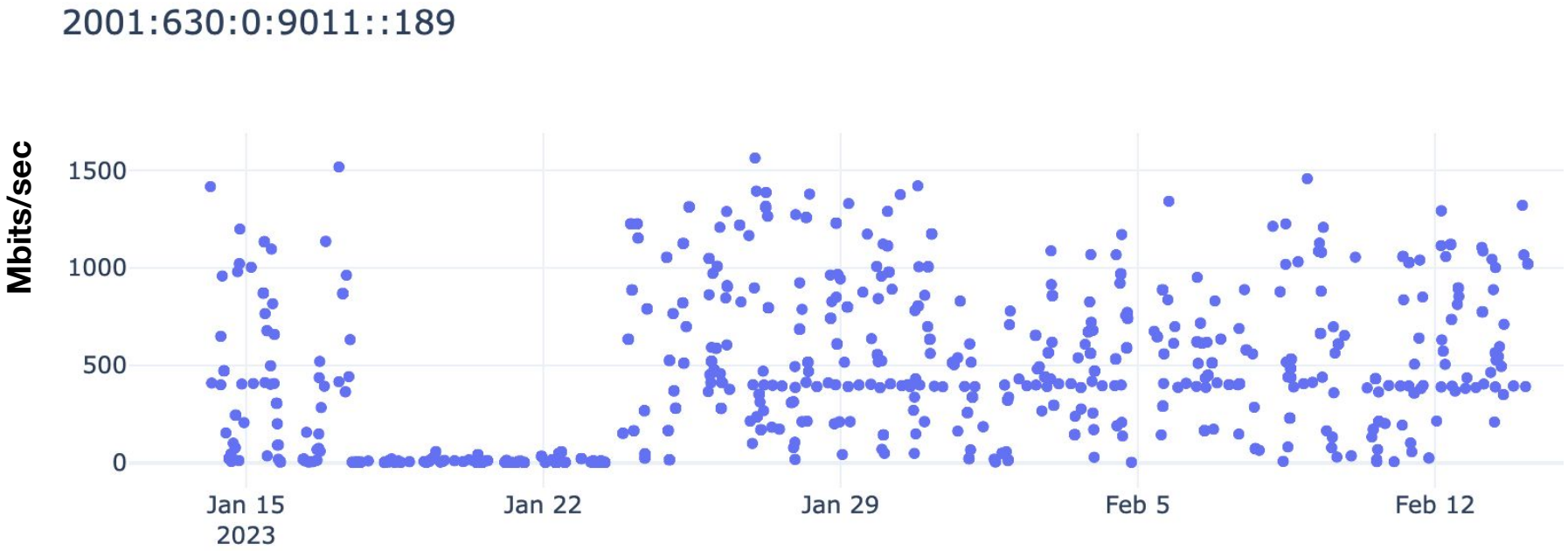


Look for a down trend

(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

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**