

WLCG Site Network Monitoring Campaign

Shawn McKee / University of Michigan

WLCG Ops Coordination <https://indico.cern.ch/event/1272843/>

April 6, 2023

on behalf of WLCG Monitoring Task Force



Overview

- After the WLCG Network Data Challenge in Fall 2021, we identified site network monitoring as one of the key missing components
- **In today's meeting we want to announce a campaign getting larger sites to provide the needed information**
 - Network information targeted for humans
 - Total network traffic IN/OUT targeted for central monitoring
- **Goal:** our Tier-1s and bigger Tier-2s provide URLs to both types of information and register those URLs into [WLCG CRIC](#)
- This presentation provides the details and we can discuss feasibility, site concerns and possible alternatives.

Targeted Site List

Our goal is to have increased visibility into our sites ability to use their networks, both for the upcoming WLCG Data Challenge 2 and for general operations.

The campaign will be targeting (Note: ALL sites are welcome to participate):

- **All Tier-1 sites**
- Larger **Tier-2** sites(**52**) using [WLCG WCW](#) (HEPscore23 Hours > 250 Mil for last 90 days):

MWT2,TOKYO-LCG2,DESY-HH,GRIF,UFlorida-HPC,GLOW,MIT_CMS,praguelcg2,AGLT2,UKI-LT2-IC-HEP,CSCS-LCG2, Nebraska,Purdue,CIT_CMS_T2,UKI-SCOTGRID-GLASGOW,UKI-NORTHGRID-MAN-HEP,GoeGrid,SWT2_CPB, BU_ATLAS_Tier2,CA-SFU-T2,INFN-LNL-2,UCSDT2,UKI-NORTHGRID-LANCS-HEP,UNIBE-LHEP,JINR-LCG2,LBL_HP CS,UKI-LT2-QMUL,RWTH-Aachen,INFN-BARI,DESY-ZN,CA-VICTORIA-WESTGRID-T2,UKI-SCOTGRID-DURHAM, IFIC-LCG2,BEgrid-ULB-VUB,CA-WATERLOO-T2,NCBJ-CIS,RO-07-NIPNE,IN2P3-LAPP,IN2P3-CPPM, UKI-SOUTHGRID-RALPP,OU_OSCER_ATLAS,BUDAPEST,INFN-MILANO-ATLASC,IN2P3-LPC, UKI-SOUTHGRID-OX-HEP,CIEMAT-LCG2,ORNL,UKI-LT2-RHUL,IN2P3-IRES,CYFRONET-LCG2,MPPMU, TECHNION-HEP,IN2P3-LPSC

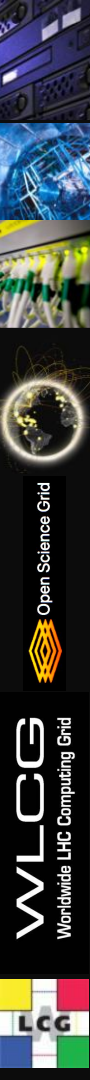
Site Network Monitoring Motivation

During DC1 in October 2021, while we had certain kinds of network monitoring data, we had almost no information about actual network traffic associated with each of our sites.

Without seeing the total traffic to/from our sites it is very difficult to understand **bottlenecks** and **other network issues** that could be impacting site's performance.

In general, we don't have sufficient information or understanding about our site's network configuration, architecture or performance.

We want to find a way to gather the minimal amount of information that will help us understand and improve how our sites work across the WAN.



Providing WLCG Site Network Information

GOAL: Provide human readable information about a sites network (description, link information, peering information, equipment, diagrams, etc.)

We have created a template [markdown file](#) which sites can clone and fill out following the instructions here:

<https://gitlab.cern.ch/wlwg-doma/site-network-information/-/tree/master/>

The template has both mandatory and optional sections.

Examples are in the [SitePages](#) area in Gitlab

The completed site specific markdown file should be converted to HTML, downloaded to a site's web server and the URL registered in [WLCG CRIC](#).

Outline of Site Network Info Template

SITE Network Information

Network Overview [**Mandatory**; can be brief]

Network Description [Optional]

Peering Description [Optional]

Network Equipment Details [Optional]

Network Monitoring [**Mandatory**]

Network Monitoring Link Into CRIC [**Mandatory**]

Network Diagrams [Optional]

[Current example for AGLT2](#) and [CERN](#)

WLCG Site Network Monitoring

GOAL: Provide site's **IN** and **OUT** network traffic (total) updated each minute, in a publicly accessible URL in JSON format as shown on the right =>

We have an example python3 script which implements what is required at: <https://gitlab.cern.ch/wlcg-doma/site-network-information/-/tree/master/WLCG-site-sntp>

The example script queries one or more interfaces that represent the **boundary** of the site, adding up IN and OUT traffic of all kinds and producing a JSON output file with the right schema.

JSON Formatted Site Network Stats (AGLT2)

```
{
  Description: "Network statistics for
AGLT2",
  UpdatedLast:
    "2023-04-05T19:59:01.691317+00:00" ,
  InBytesPerSec: 1294612738.7737598 ,
  OutBytesPerSec: 1023097622.4124134 ,
  UpdateInterval: "60 seconds" ,
  MonitoredInterfaces:
    [
      "aglt2-rtr-1.local_Ethernet1/48" ,
      "aglt2-rtr-1.local_Ethernet1/51" ,
      "aglt2-rtr-1.local_Ethernet1/52" ,
      "aglt2-rtr-2.local_Ethernet1/51" ,
      "aglt2-rtr-2.local_Ethernet1/52"
    ]
}
```

Current examples available at

<https://head01.aglt2.org/aglt2-netmon.json>

https://netstat.cern.ch/monitoring/network-statistics/ext/wlcg/total_traffic.json

WLCG Site Network Monitoring

Format of JSON for site network monitoring use

Description: Text

UpdatedLast: ISO 8601 UTC (microsecond resolution)

InBytesPerSec: Total incoming bytes/sec to the site

OutBytesPerSec: Total outgoing bytes/sec from the site

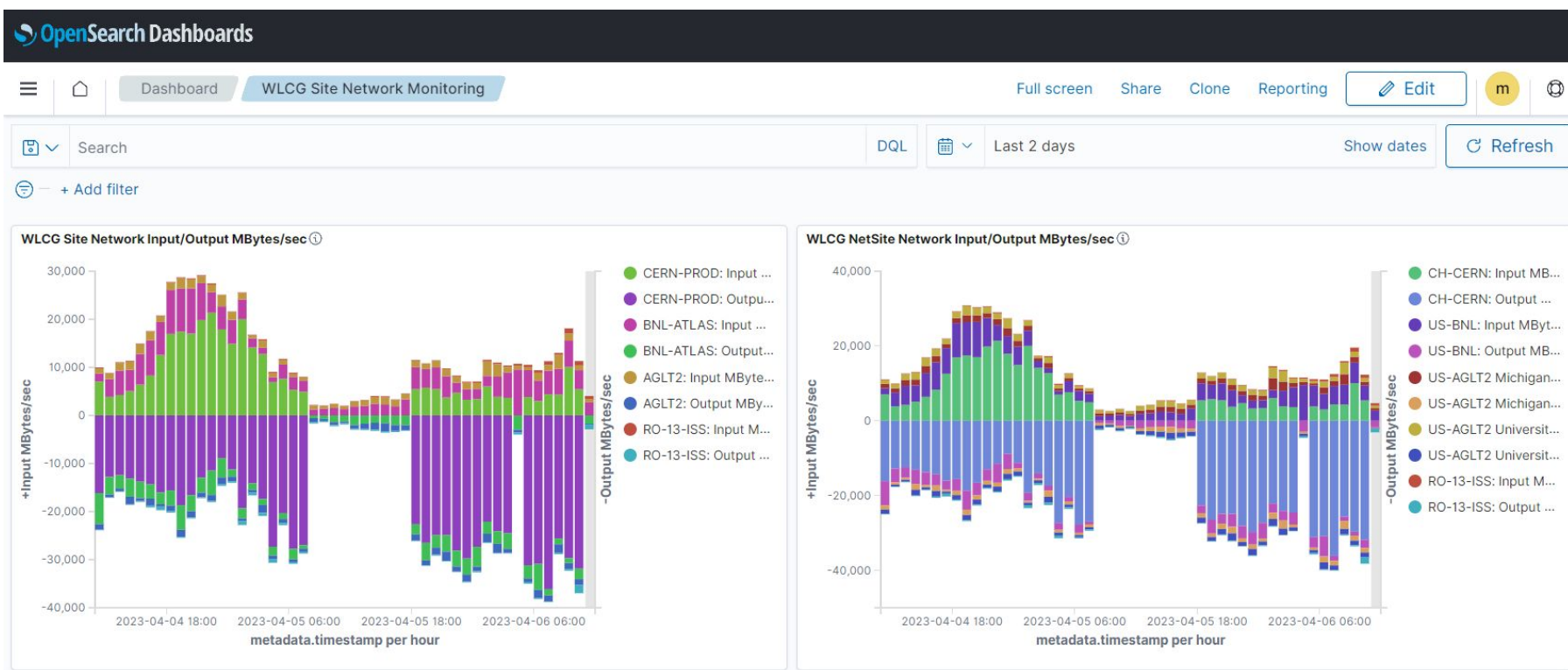
UpdateInterval: Text describing the frequency for updates in seconds: '60 seconds'

MonitoredInterfaces: List of interfaces that are used to calculate the In/OutBytesPerSec in <host>_<interface> format

Summary: Instructions To Sites

- **Describe** your network in Gitlab (has Mandatory and Optional parts)
<https://gitlab.cern.ch/wlcg-doma/site-network-information/>
 - Copy the Template/ [SitePageTemplate.md](#) to the SitePages/<Sitename>.md
 - Edit <Sitename>.md to describe your site's network
 - Following the README on the URL above, convert the .md to HTML and serve it at a URL
 - Register that URL in WLCG-CRIC's [NetSite](#) Info URL
- **Identify** and **gather** your IN/OUT network statistics
 - For each `NetSite` associated with your RC Site (site), determine all the network links which connect to that `NetSite`. **NOTE:** Each [NetworkRoute](#) can also provide a Monitoring URL
 - For the identified connections, gather the input and output counters for those interfaces following instructions at <https://gitlab.cern.ch/wlcg-doma/site-network-information/-/tree/master/WLCG-site-snmp>
 - “**Publish**” the statistics every 60 seconds at a URL reachable by CERN Monit and record that URL in the appropriate Monitoring URL for the `NetSite` or `NetworkRoute` in [WCLG CRIC](#)
- **Maintain** the information as things change

Example of Monitoring in CERN MONIT



Dashboard URL (still working out some data acquisition bugs)

https://monit-opensearch.cern.ch/dashboards/goto/f0607fb8528ce6b7c9a336aef74be40b?security_tenant=global

Site Expectation Summary

Sites will be expected to:

- **Document** their network at a high-level, with options to provide helpful details
 - Information about site topology, peering, hardware and capacity can allow WLCG network experts to better support, diagnose and fix network problems
- **Gather** (via snmp or other data source) the IN/OUT traffic of their whole site
 - This requires identifying the interface(s) that represent the “border” of the site and a mechanism to gather the interface(s) traffic
- **Provide** URLs to access the network information and monitoring files
- **Maintain** the documentation and update monitoring as networking at the site evolves

Notification for Sites: UDP Fireflies Being Sent

One quick heads-up for sites and network providers: we are beginning to send **UDP fireflies** from some of our sites along data transfer paths. UDP fireflies (by default) are sent to the same destination as the data transfer flow. This means UDP packets arriving at storage servers on port 10514.

A site can choose to [ignore](#), [block](#) or [capture](#) these packets

We are working on an [informational RFC](#) (target to publish Fall 2023)

One implication: if packets hit iptables, it may generate noise in the logging that may be a concern (fill /var/log?)

Recommendation is to open port 10514 for incoming UDP packets or explicitly 'drop' them.

Conclusion

- WLCG would like sites to provide some additional information about their networking, including regularly updated IN/OUT statistics
- Sites should provide two URLs, entered into the WLCG CRIC that provide the required information
 - While we have provided a python3/systemd examples, sites can choose to use whatever method works best for them to provide IN/OUT statistics in JSON format.
- WLCG will programmatically gather the data, making it accessible via MONIT at CERN
- **Deadline: September 2023** (well in advance of WLCG Data Challenge 2)

Examples and detailed instructions are in the [CERN Gitlab](#)

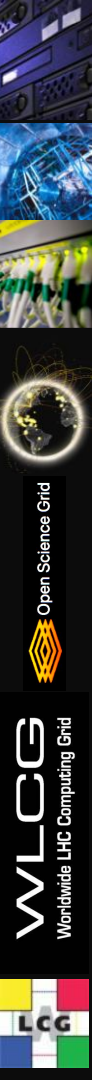
This is the plan but we want feedback from the sites and network admins

Acknowledgements

We would like to thank the **WLCG**, **HEPiX**, **perfSONAR** and **OSG** organizations for their work on the topics presented.

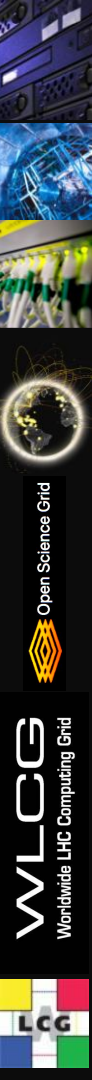
In addition we want to explicitly acknowledge the support of the **National Science Foundation** which supported this work via:

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



Questions / Discussion

Questions, Comments, Suggestions?



Useful Networking URLs

- OSG/WLCG Networking Documentation
 - <https://opensciencegrid.github.io/networking/>
- perfSONAR Infrastructure Dashboard
 - <https://atlas-kibana.mwt2.org:5601/s/networking/goto/9911c54099b2be47ff9700772c3778b7>
- perfSONAR Dashboard and Monitoring
 - <http://maddash.opensciencegrid.org/maddash-webui>
 - https://psetf.opensciencegrid.org/etf/check_mk
- 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://aaas.atlas-ml.org/>
- The pS Dash application: <https://ps-dash.uc.ssl-hep.org/>
- ESnet WLCG DC Dashboard:
<https://public.stardust.es.net/d/lkFCB5Hnk/lhc-data-challenge-overview?orgId=1>

Backup Slides Follow

Review of Existing Network Monitoring

While the site specific network IN/OUT metrics are generally not available we do have a number of network related activities and data:

- **WLCG data challenge monitoring:**

- <https://monit-grafana.cern.ch/d/W2Uj1gDnz/wlcg-transfers-playground?orgId=20>

- **ESnet network monitoring** - ESnet created a monitoring page specifically for our WLCG Network Data Challenge:

- <https://public.stardust.es.net/d/lkFCB5Hnk/lhc-data-challenge-overview?orgId=1>

- **The NetSage project** has LHC specific information

- <https://lhc.netsage.global/grafana/d/xk26lFhmk/flow-data-for-circuits?orgId=2>

- **perfSONAR global deployment** and associated analytics

- Details documented in many presentations; see last [LHCONE meeting](#)

- **Research Networking Technical Working Group ([RNTWG](#))**

- Has working areas in packet marking / flow labeling, traffic shaping / packet pacing and network orchestration.

Example in WLCG CRIC

Home

Site Topology

Pledges

Accounting

Downtime

Admin

Logs

Help

shawn.

Export

Filter

Reload

Columns

12/16

NetSite list

RC Site	NetSite	NOC	monit URL	info URL	AUP	LHCONE active	LHC act
AGLT2	US-AGLT2 Michigan State University	aglt2-noc@umich.edu		info URL	✓	✓	✗
AGLT2	US-AGLT2 University of Michigan	aglt2-noc@umich.edu	monit URL	info URL	✓	✓	✗
ANLASC	US-ANL				✓	✓	✗
ARNES	SL-ARNES-NREN				✓	✓	✗
ARNES	SL-IJS-Ljubljana				✓	✓	✗
ARNES	SL-IJL IM-Marihor				✓	✓	✗

Home / Core / NetSite / US-AGLT2 University of Michigan

NetSite: US-AGLT2 University of Michigan

RCSite [AGLT2](#)

Network site US-AGLT2 University of Michigan

NOC contacts [aglt2-noc@umich.edu](#)

WAN bandwidth 80

Monitoring URL <https://head01.aglt2.org/aglt2-netmon.json>

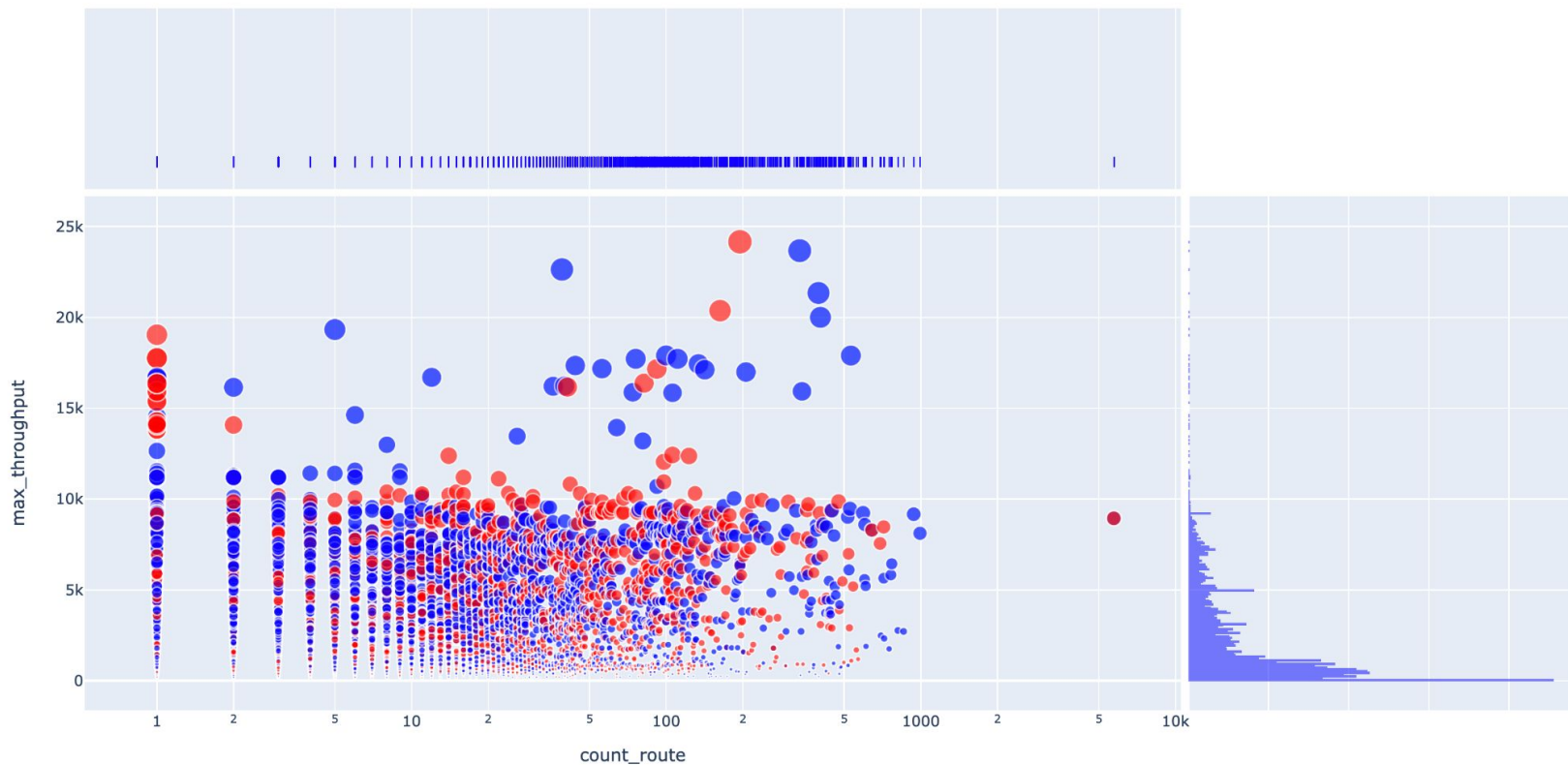
Info URL <https://head01.aglt2.org/AGLT2.html>

Description The AGLT2 network for the University of Michigan location

Last modified 2022-05-13 19:07:26.867507

Network Analytics

Number of times path taken vs. maximum throughput on path



ESnet Monitoring for WLCG Data Challenge

ESnet created a very nice [monitoring dashboard](#)

LHC Data Challenge / LHC Data Challenge Overview

Last 6 hours

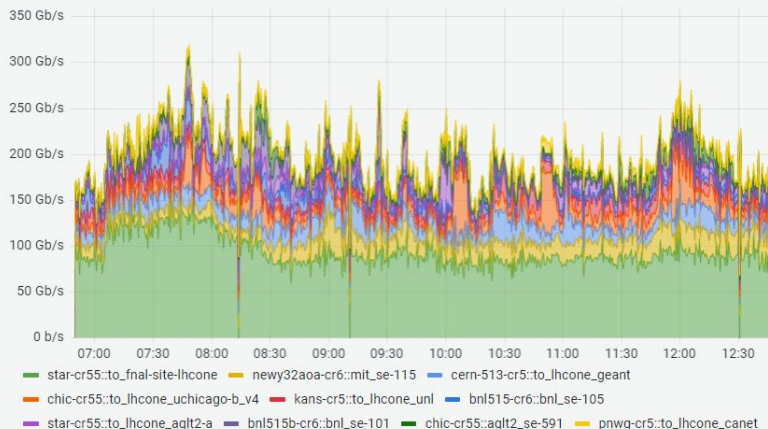
LHC Data Challenge Overview

Menu: Overview | Interfaces | Sites | Transatlantic | LHCOPN

This dashboard shows an overview of statistics relevant to the LHC data challenge. It contains a combination of SNMP and flow statistics from ESnet's Stardust measurement system. Use the navigation menu above this text or links in the data below to move to other dashboards that provide different views of the data.

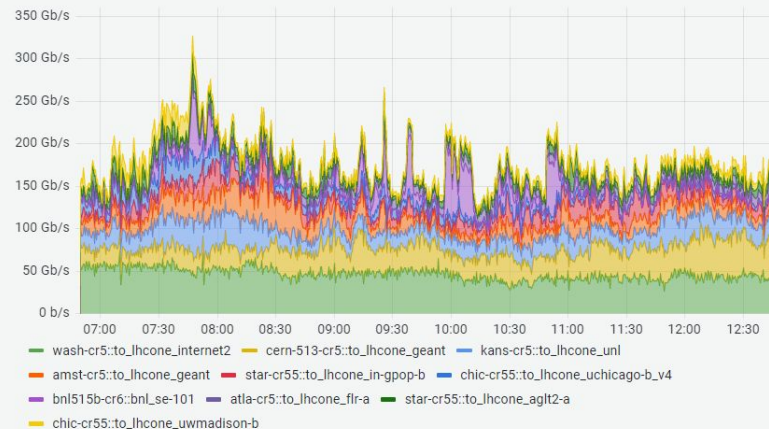
SNMP Statistics

Top 10 Interfaces by Incoming Rate (SNMP)



Top Interfaces by Incoming Volume (SNMP)

Top 10 Interfaces by Outgoing Rate (SNMP)



Top Interfaces by Outgoing Volume (SNMP)

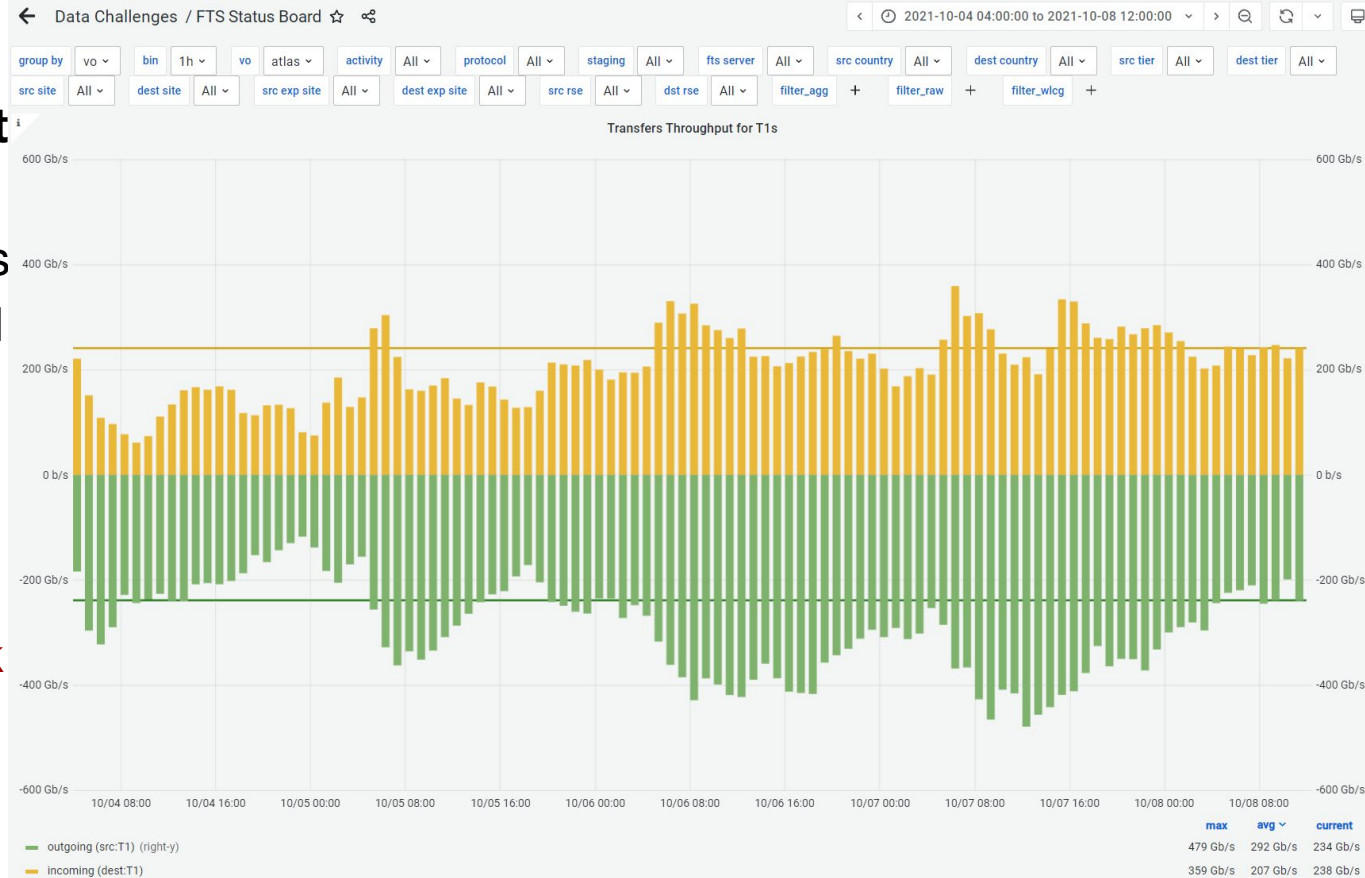
Update on WLCG Network Data Challenge (1 / 2)

WLCG data challenge was Oct 04-08

Goal was 240 Gbps from T0 to T1's and from T1's to T2's involving primarily ATLAS and CMS

The network was NOT the bottleneck in general!

This week is the Tape challenge!



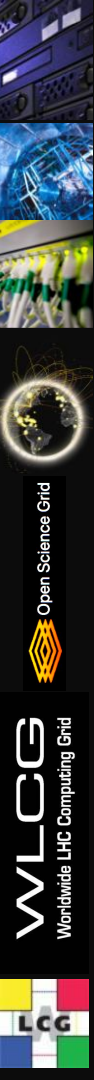
WLCG Network Throughput Support Unit

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**