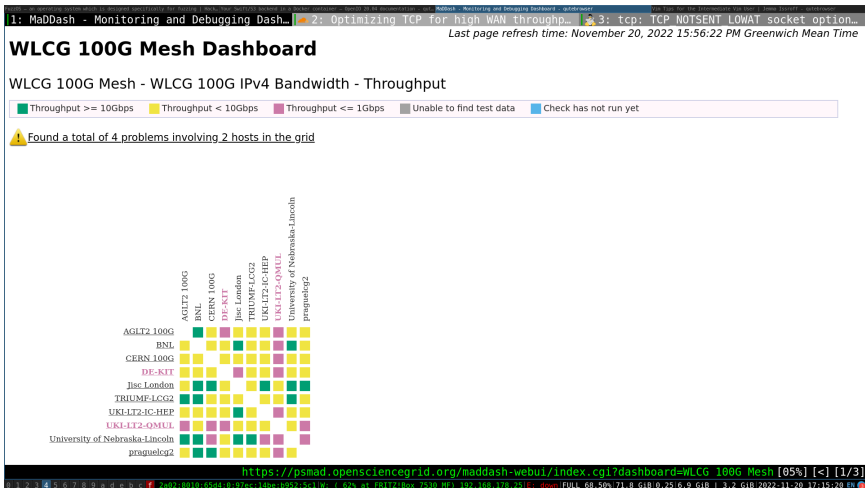


Odd and nice things in perfSONAR land

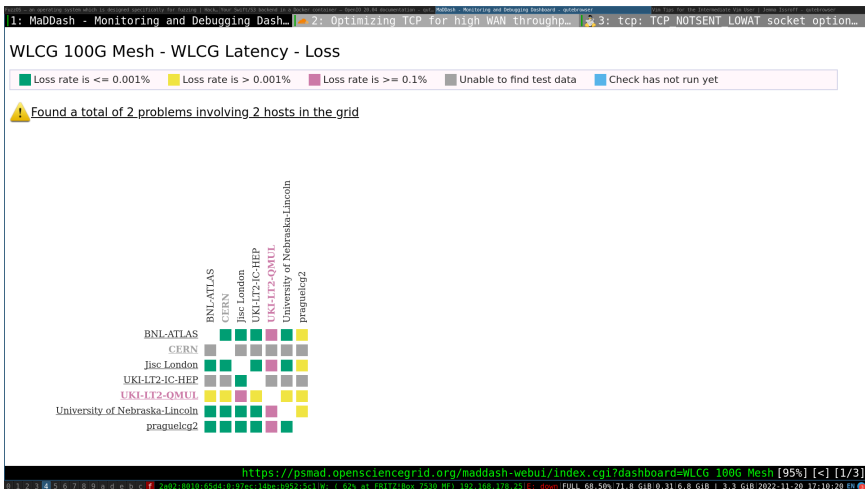
Raul H C Lopes (Jisc)
Raul.Lopes@jisc.ac.uk

- Jisc London was added to the WLCG dashboard in July.
- It shows green with 5 out of 9 peers.
- We've seen at least one serious glitch.
- We see several permanent issues in the dashboard related to
 - latency
 - plotting
 - throughput
- I would like to talk about issues and further use of perfSONAR. I have lots of questions. Sorry!

IPv4 throughput on the WLCG 100G dashboard

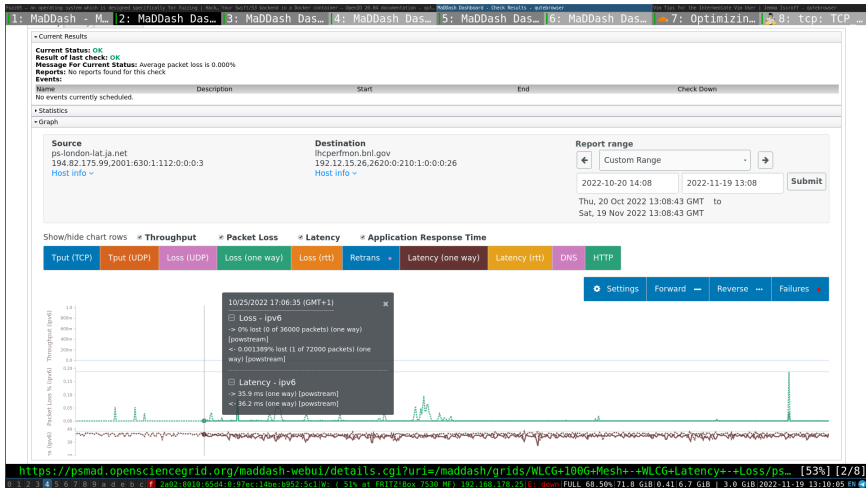


CERN and Imperial do not run latency

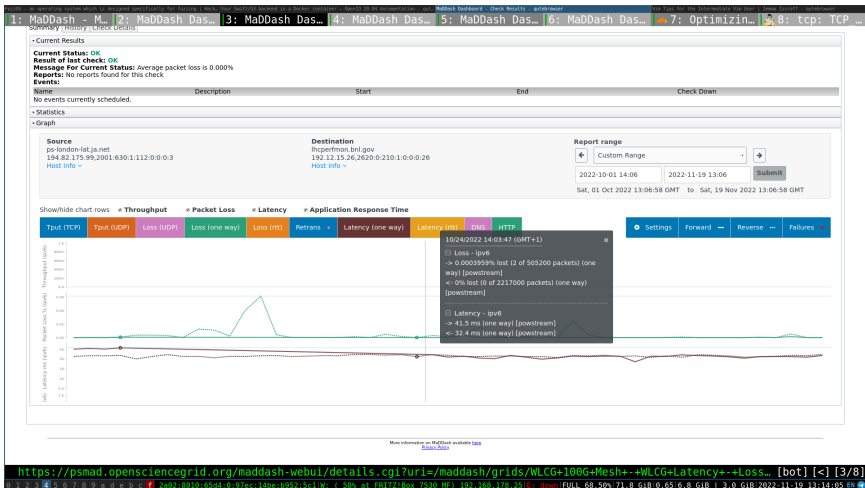


Jisc London went dark from 18 to 25 October

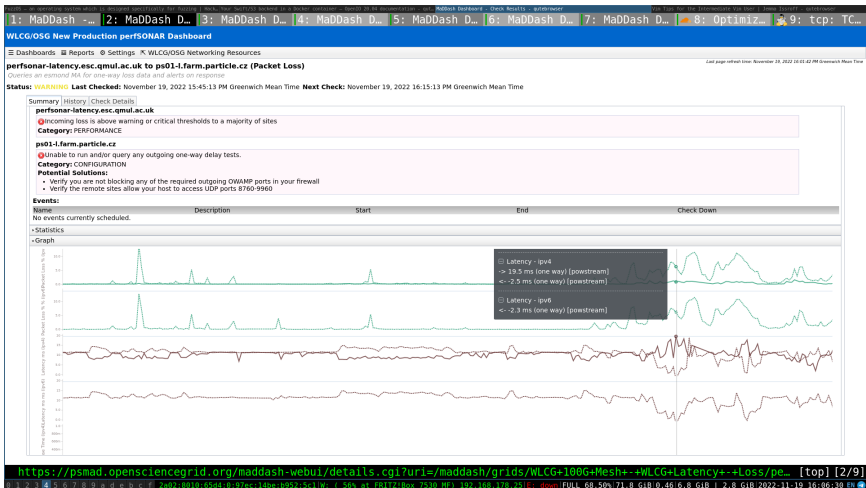
- Failure in reverse IP resolution
- **Only** latency tests were affected. Why only latency?



Surprise when zooming in

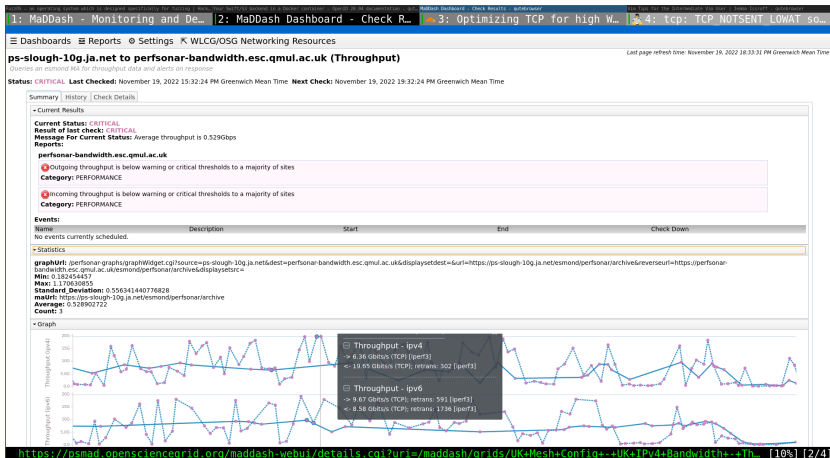


Negative latency for QMUL-Prague



QMUL to Jisc Slough: 20G on a 10G link

- The negative latency (previous slide) coincides with the 15-20G spikes.
- Google on TrueTime <https://cloud.google.com/spanner/docs/true-time-external-consistency>



Statistics associated with plots are frequently incorrect

1: MaDDash - Monitoring and De... 2: MaDDash Dashboard - Check R... 3: Optimizing TCP for high W... 4: tcp: TCP NOTSENT LOWAT so...

WLCG/OSG New Production perfSONAR Dashboard

Dashboards Reports Settings WLCG/OSG Networking Resources

ps-london-bw.ja.net to lhcmn3.bnl.gov (Throughput)

Queries an endpoint for throughput data and alerts on response

Status: **OK** Last Checked: November 19, 2022 15:30:22 PM Greenwich Mean Time Next Check: November 19, 2022 19:30:22 PM Greenwich Mean Time

Summary History Check Details

Result of last check: OK
Message For Current Status: Average throughput is 14.508Gbps
Reports: No reports found for this check

Events:

Name	Description	Start	End	Check Down
No events currently scheduled.				

Statistics

graphURL: /perfsonar-graph/widget.cgi?source=ps-london-bw.ja.net&dest=lhcmn3.bnl.gov&displaysetdest=+url=https://ps-london-bw.ja.net/esmond/perfsonar/archive&reverseurl=https://lhcmn3.bnl.gov/esmond/perfsonar/archive&displaysetrc=
Min: 14.330350512
Max: 14.733236492
Standard Deviation: 0.094984638138057
maURL: https://ps-london-bw.ja.net/esmond/perfsonar/archive
Average: 14.5082340169319
Count: 27

Graph

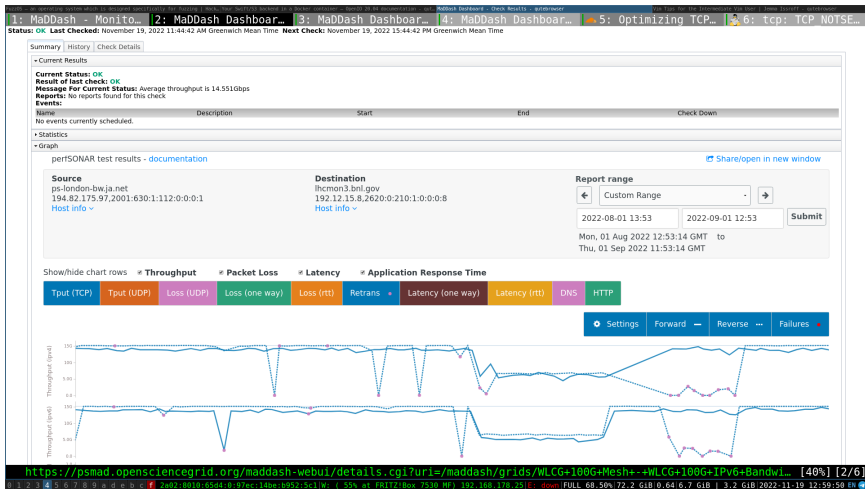
Tput (TCP) Tput (UDP) Loss (UDP) Loss (one way) Loss (rtt) Retrans Latency (one way) Latency (rtt) DNS HTTP

Settings Forward Reverse Failures

https://psmad.opensciencegrid.org/maddash-webui/details.cgi?url=/maddash/grids/WLCG+100G+Mesh+-+WLCG+100G+IPv4+Bandwid_ [top] [2/4]

2402:8010:65d4:0:97ec:14be:b952:5c1W: (58v at FRITZ!Box 7530 MF) 192.168.178.25E: down FULL 68.50% 71.8 G1B 0.14 6.6 G1B | 3.2 G1B 2022-11-19 18:07:45 EN

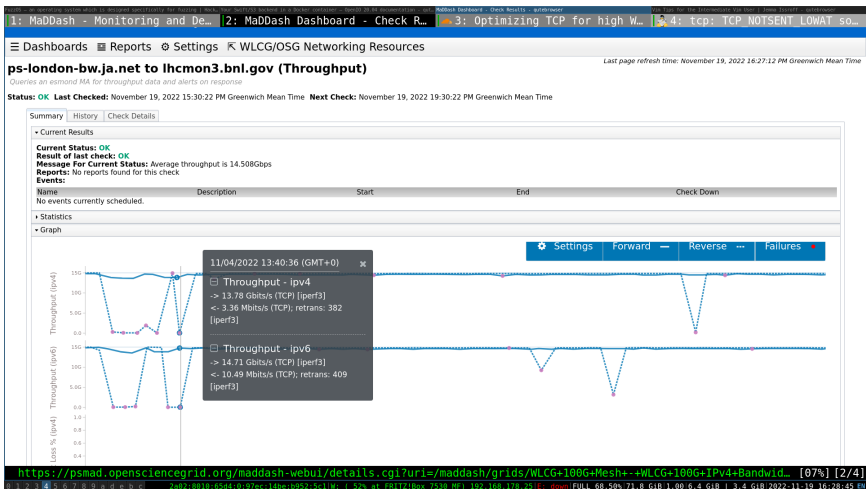
MTU: Tim's request for a comparison of MTU at 9000 vs 1500



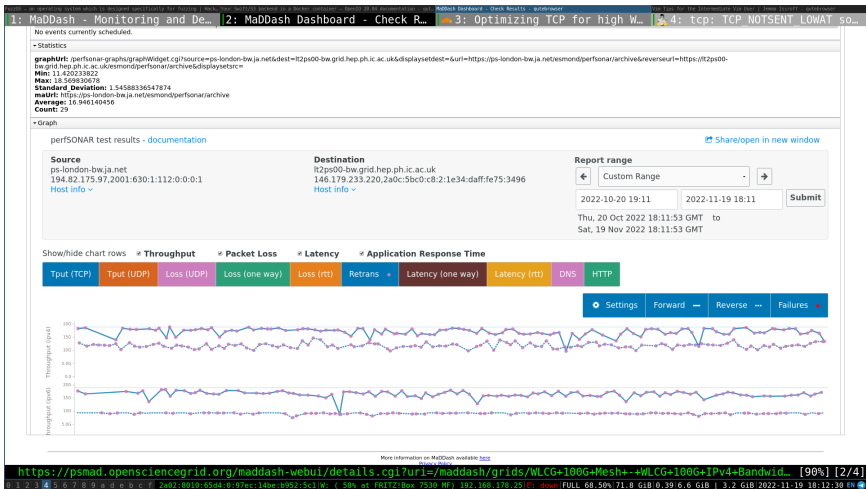
Is Brian showing a link possibly maxed out?



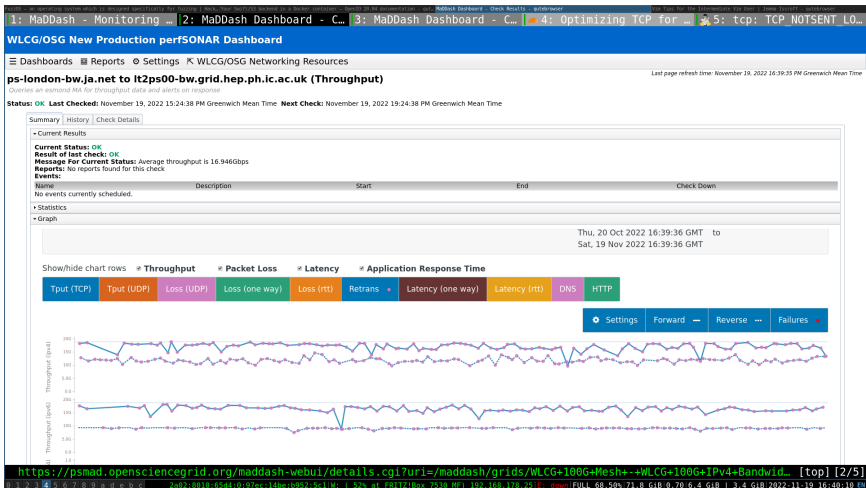
Jisc London view — BNL path



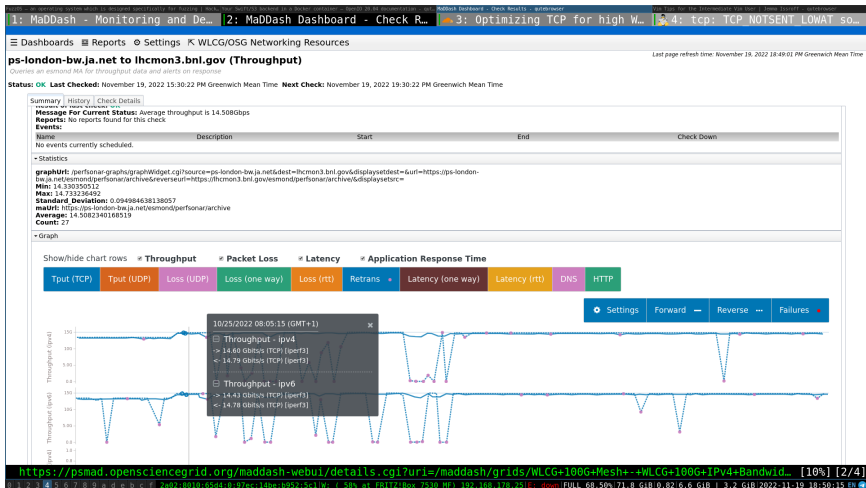
Test in the UK went unaffected



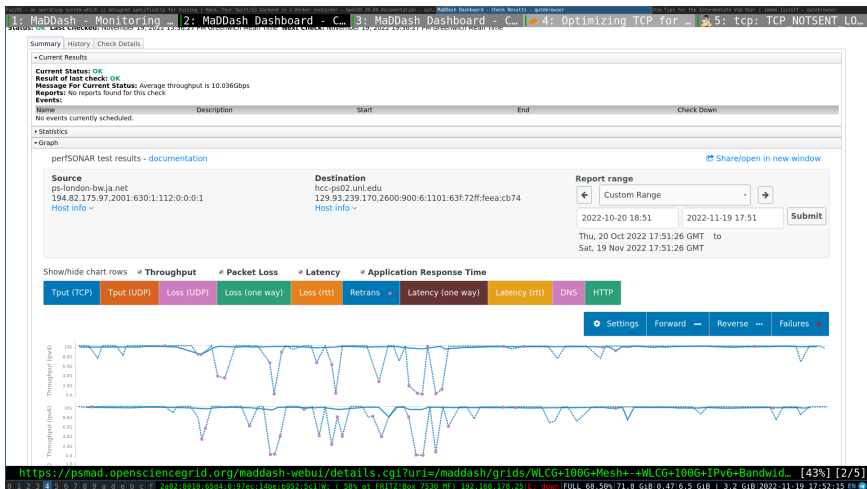
Tim: "Is Imperial capping outbound IPv6 at 10G?"



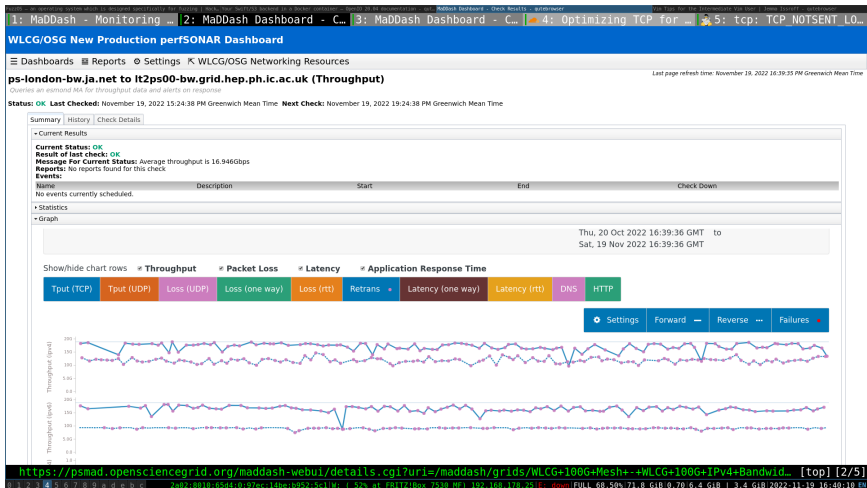
Jisc London to BNL: feel good of MTU=9000



Jisc London to UNL: feel good of MTU=9000



Jisc London to Imperial: MTU=9000 vs MTU=1500



Jisc London configuration

- Kernel: tested 5.14, 5.17, 6.0
- CWA: tested reno, htcp, cubic, bbrv1; bbrv1 being used.
- tcp_rmem = max test 256MB and 512MB; 512MB in use.
 - Google and Cloudflare have reported that small tcp_rmem on high latency limits throughput.
- tcp_adv_win_scale = -2; it limits the TCP window to 3/4 of available receiving memory.
- tcp_wmem = max test 256MB and 512MB; 512MB in use.
- net.ipv4.tcp_notsent_lowat = 131072
 - limit number of unsent bytes in the writing queue
 - reduce buffer bloat <https://lwn.net/Articles/560082/>
- cpu power set to performance.
 - Important for smooth receiving.
 - Evidence when testing with RNP: a 4 cores server was capable of sending at 97 Gbps.
 - Same tests, same server: it never achieved more than 70Gbps while receiving.

- Tim Chown (task leader)
- Duncan Rand (Speaker of House and WLCG Ambassador)
- Chris Walker (University Ambassador)
- David Richardson (Network Engineer)
- Raul Lopes (some semblance of a sysadmin)
- Supermicro servers behind a Freertr.