perfSONAR update

Christopher Walker, Raul Lopes, Duncan Rand, Tim Chown

GridPP52, Ambleside, 30 August 2024

Contact: *netperf@jisc.ac.uk* - reaches all of the above

What is perfSONAR ?

- A tool for monitoring end-to-end network characteristics
  - Open-source project with R&E developers - https://perfsonar.net
  - Assists assessment and debugging of network performance issues
  - Separate network performance from storage performance
  - Adopted by GridPP/WLCG, SKA, others

- Main monitoring tests:
  - Bandwidth (iperf or ethr), by default every 6 hours
  - Latency (one way, using OWAMP) and packet loss, persistently
  - Traceroute, noting path changes
  - Pluggable architecture allows additional tests to be written

Jisc

perfSONAR 5.1

Changes

- perfSONAR 5.0.0 (April 2023)
  - Opensearch – new database backend
- perfSONAR 5.1.0 (June 2024)
  - Grafana – new host and mesh views
  - Threaded *iperf3*
  - *psarchive troubleshoot; pscheduler troubleshoot*
  - Alma/Rocky 9, Debian 11,12, Ubuntu 20,22
- Future
  - LXC containers
  - Development in the area of exploring time series data
  - Use AI to help monitor nodes (identify trends)

Jisc and perfSONAR

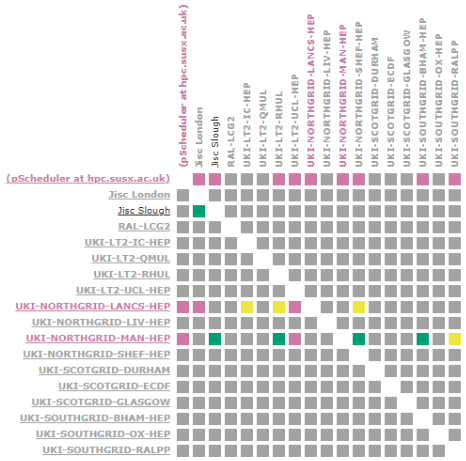
Jisc uses and recommends perfSONAR for member use

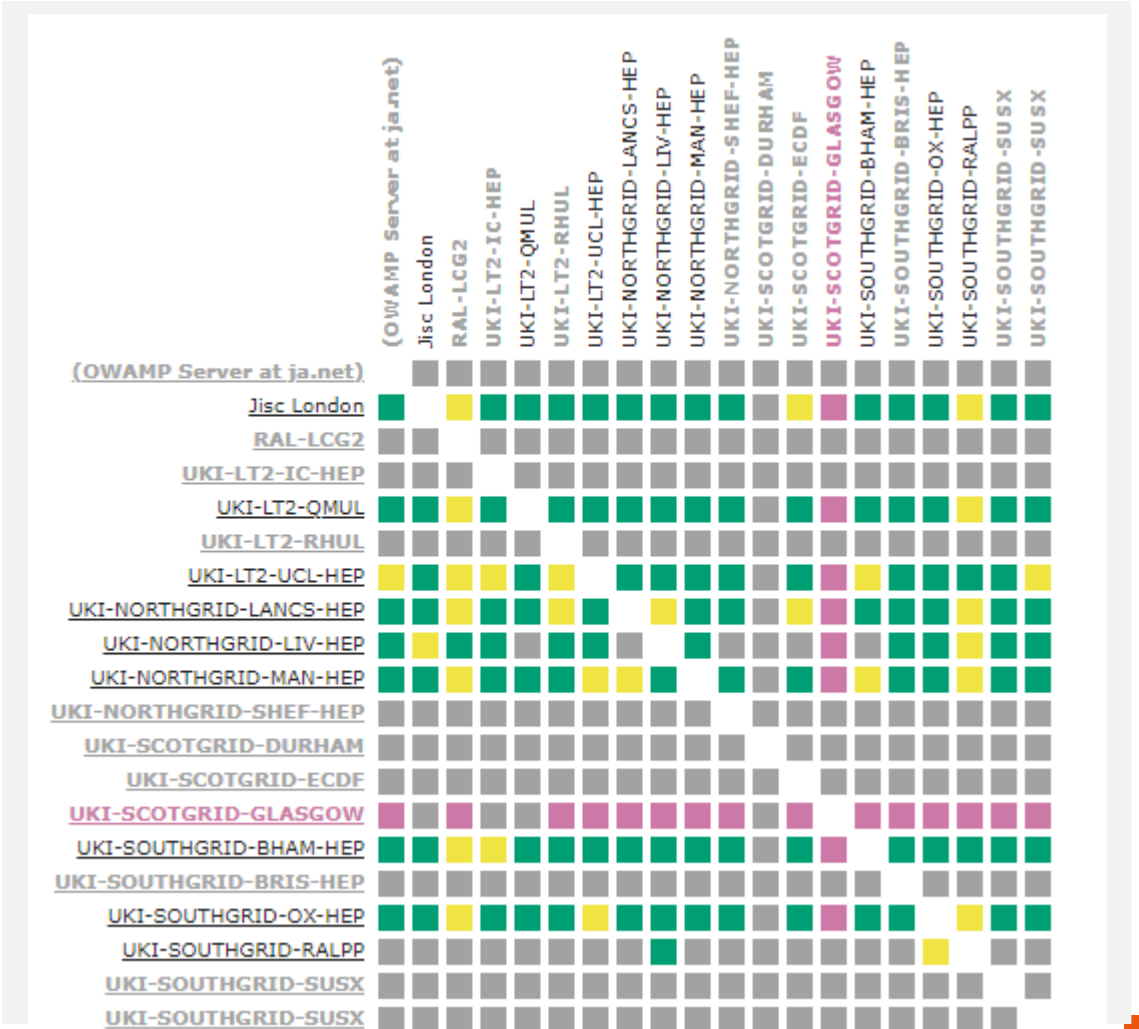
- Jisc (Raul) helps test and debug perfSONAR with the dev team
  - As part of the GÉANT GN5-1 project, and committed into GN5-2 from Jan 2025
- We run (host) and participate in a number of community meshes
  - WLCG/GridPP mesh
  - HEAnet
  - SKA
- Very useful for helping with network issues as part of our toolset:
  - Janet Network performance test facilities
- Jisc happy to assist in solving network performance problems between sites
  - Just email netperf@jisc.ac.uk

perfSONAR network monitoring

Matrix of tests

- Bandwidth, typically run every 6 hours for 30 secs
- One way latency / packet loss
- Traceroute
- GridPP mesh
  - OSG maddash (currently not working due to move to 5.1 and new Grafana-based tooling)
  - Jisc UK mesh



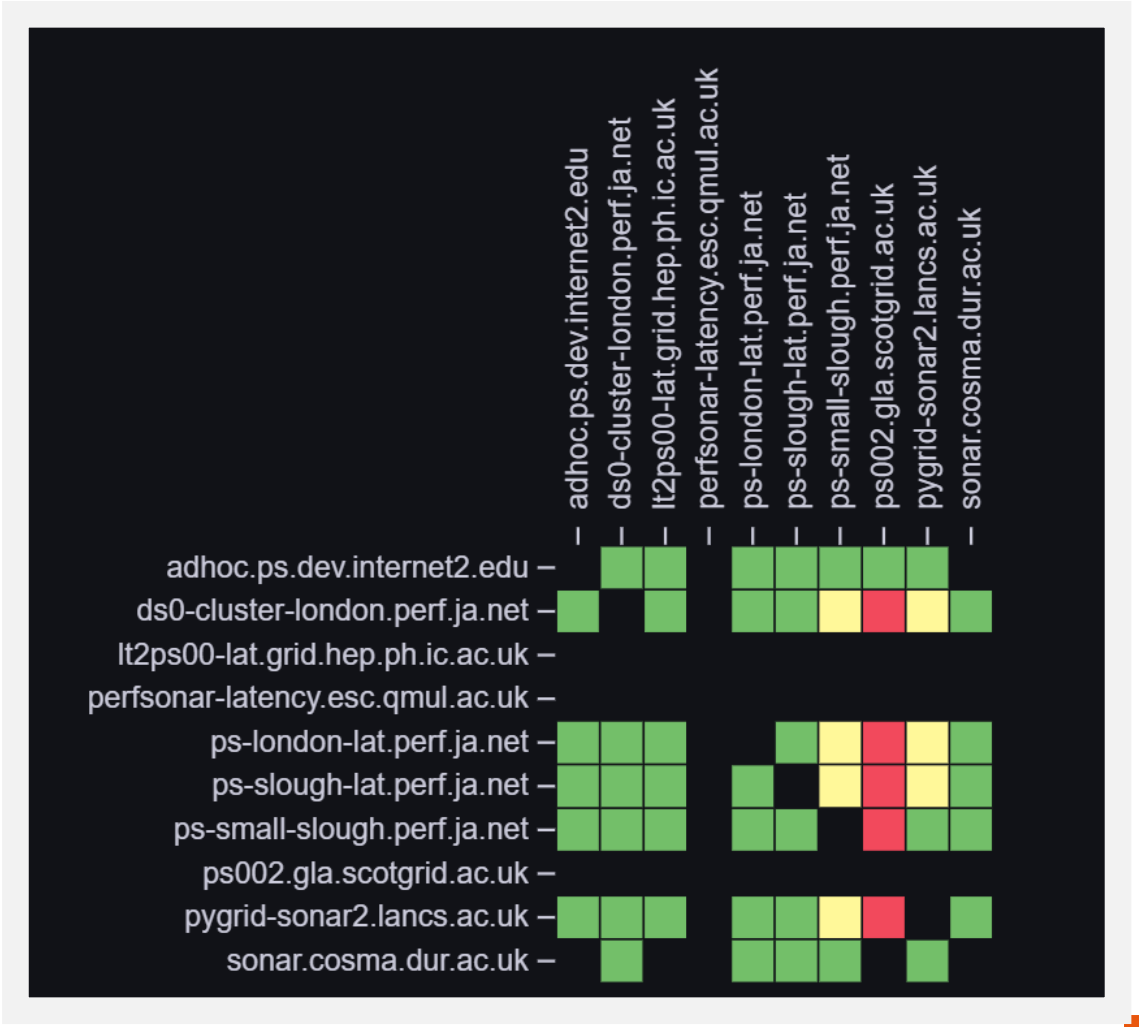


Jisc UK mesh

Piloting 5.1 Grafana tooling

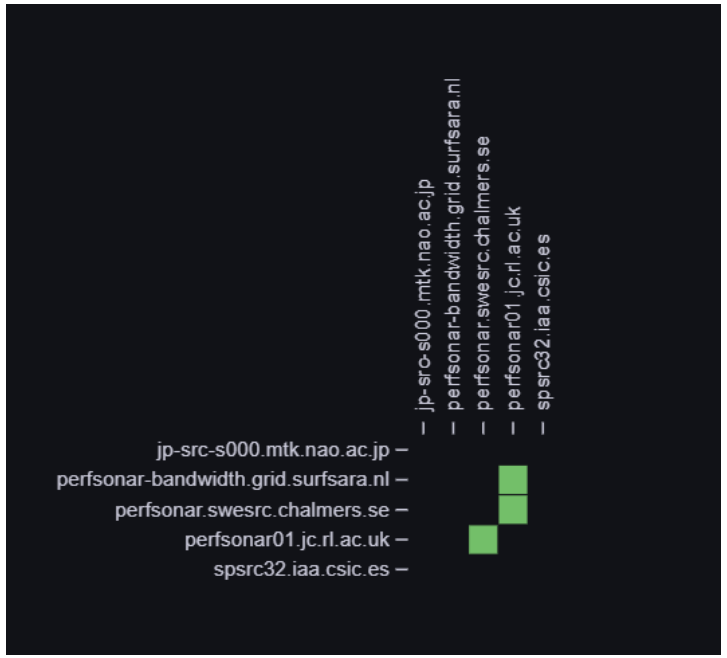
- Interim GridPP/WLCG view
  - psconfig remote add https://ds2-london.perf.ja.net/psconfig/ds2-psconfig.json
  - Email raul.lopes@jisc.ac.uk
- HPC mesh
- HEAnet (latency testing)

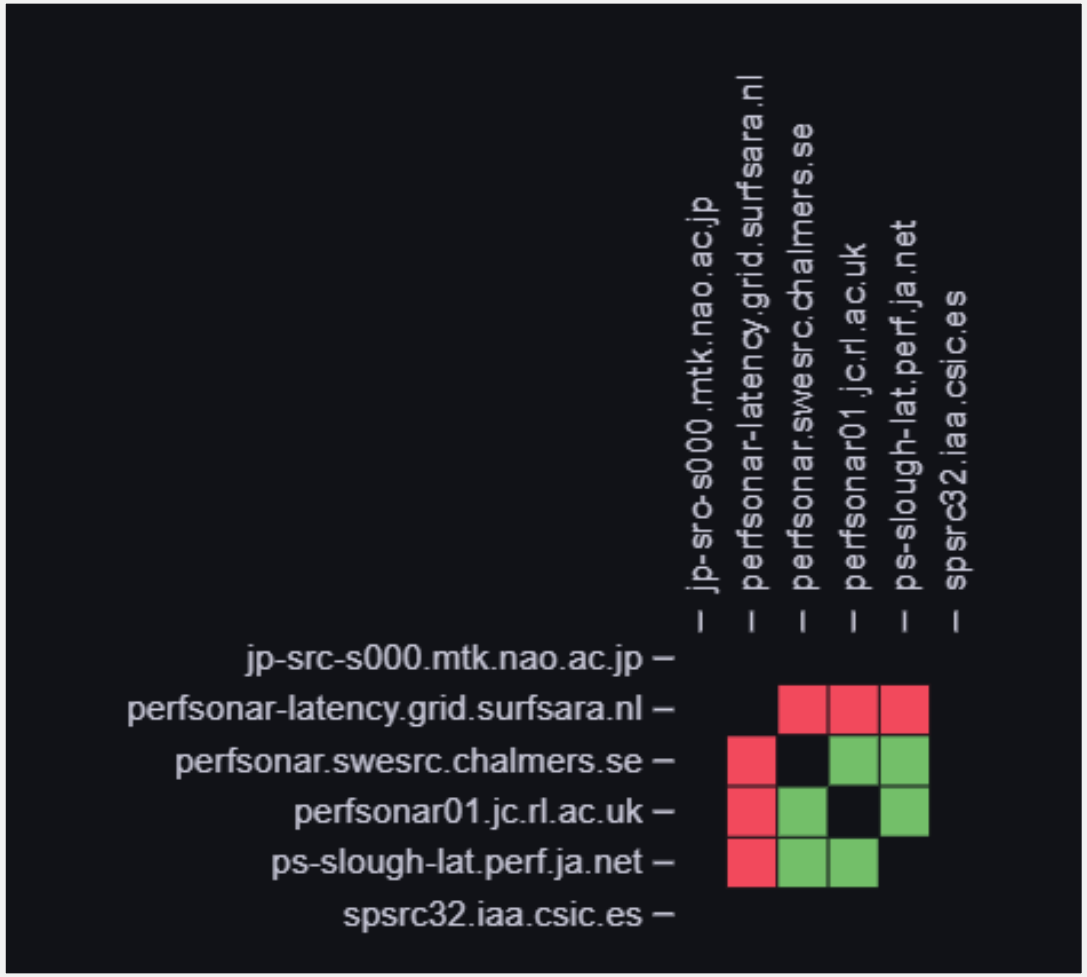
- Jisc currently looking to build a new virtual hosting platform that will further support our ability to host measurement archives and Grafana meshes for members and communities.
  - Likely to use Proxmox



SKA Mesh

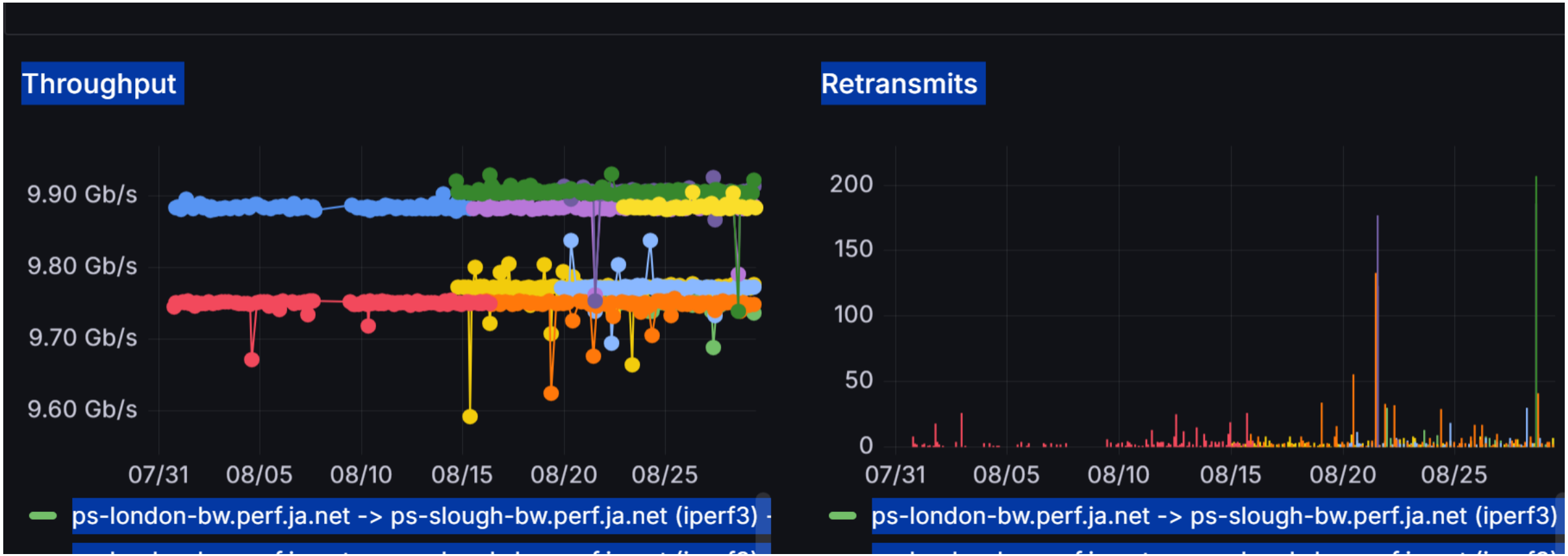
Bandwidth and Latency/packet loss





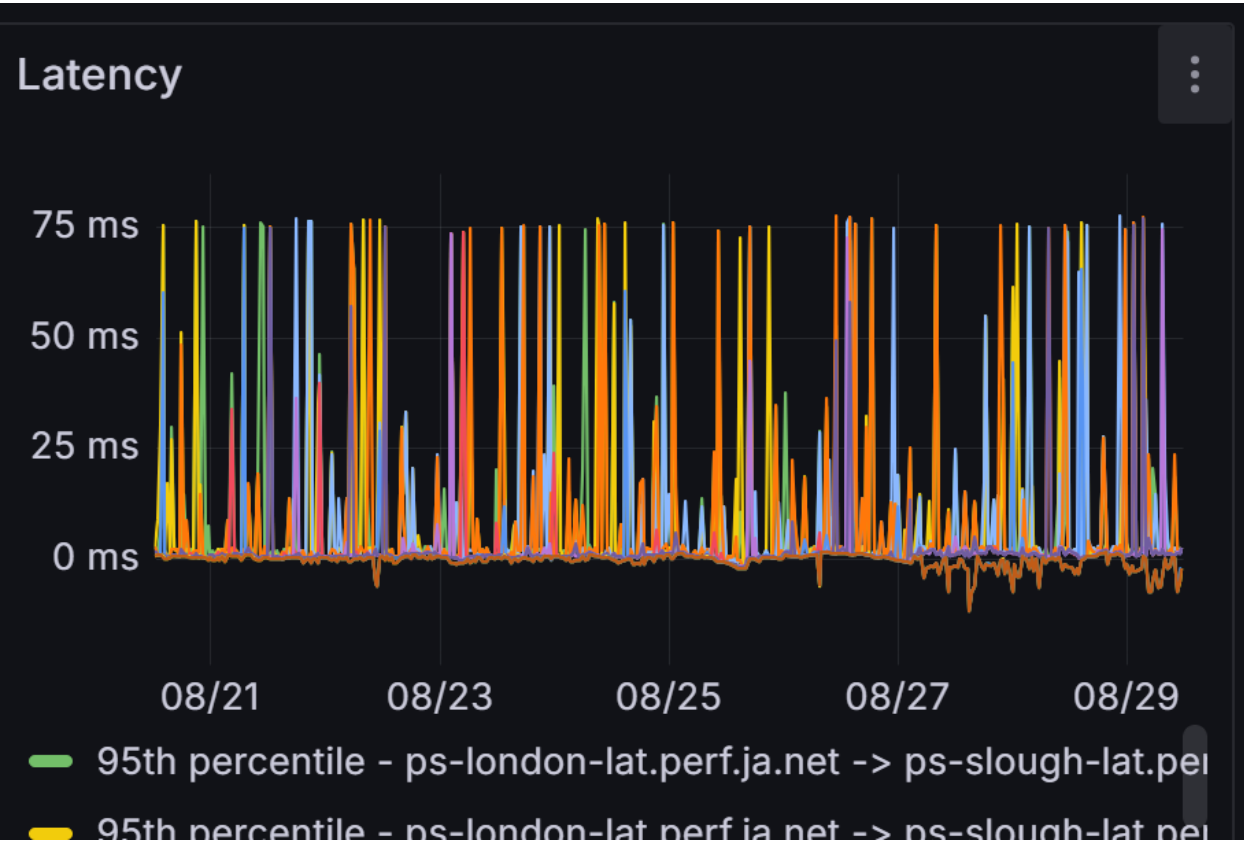
Bandwidth measurements

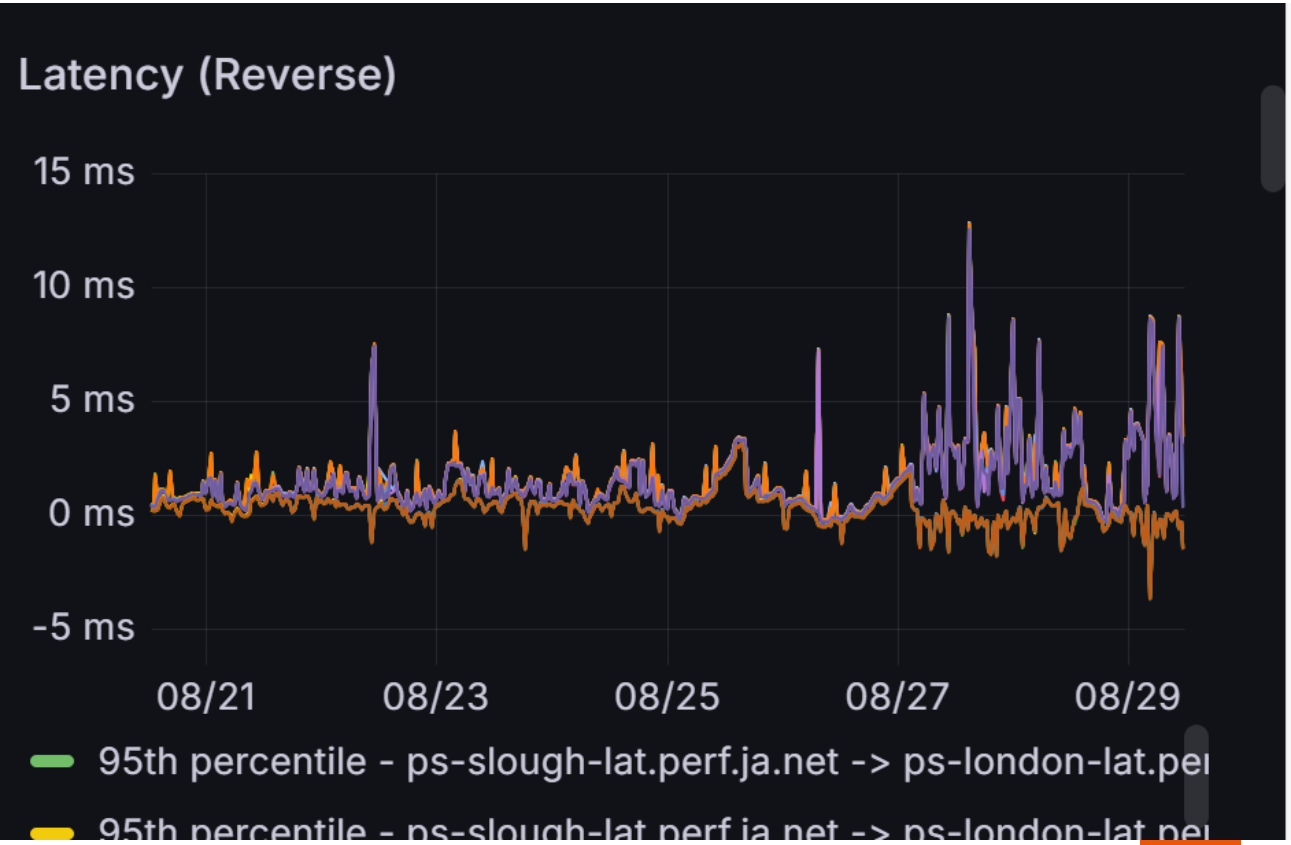
Jisc London -> Jisc Slough



Latency

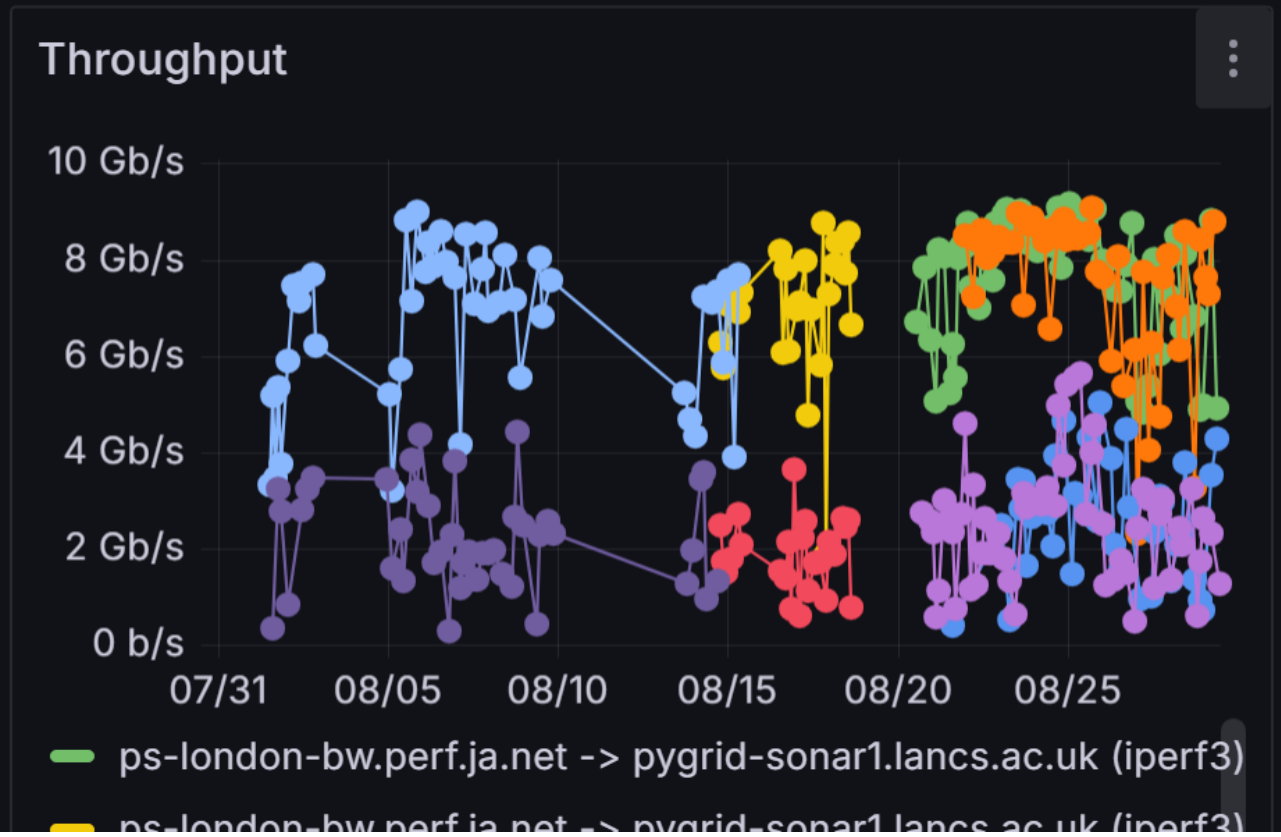
Jisc London -> Slough

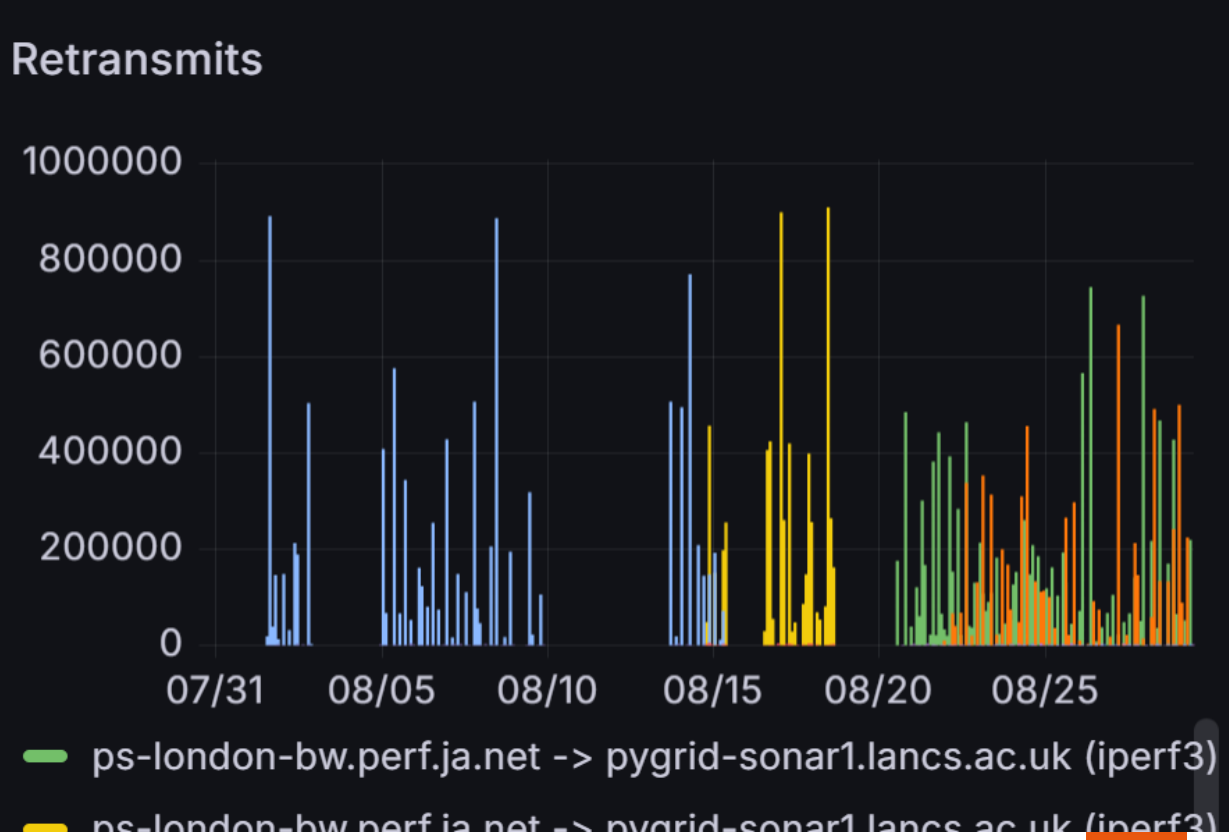




Bandwidth Measurement

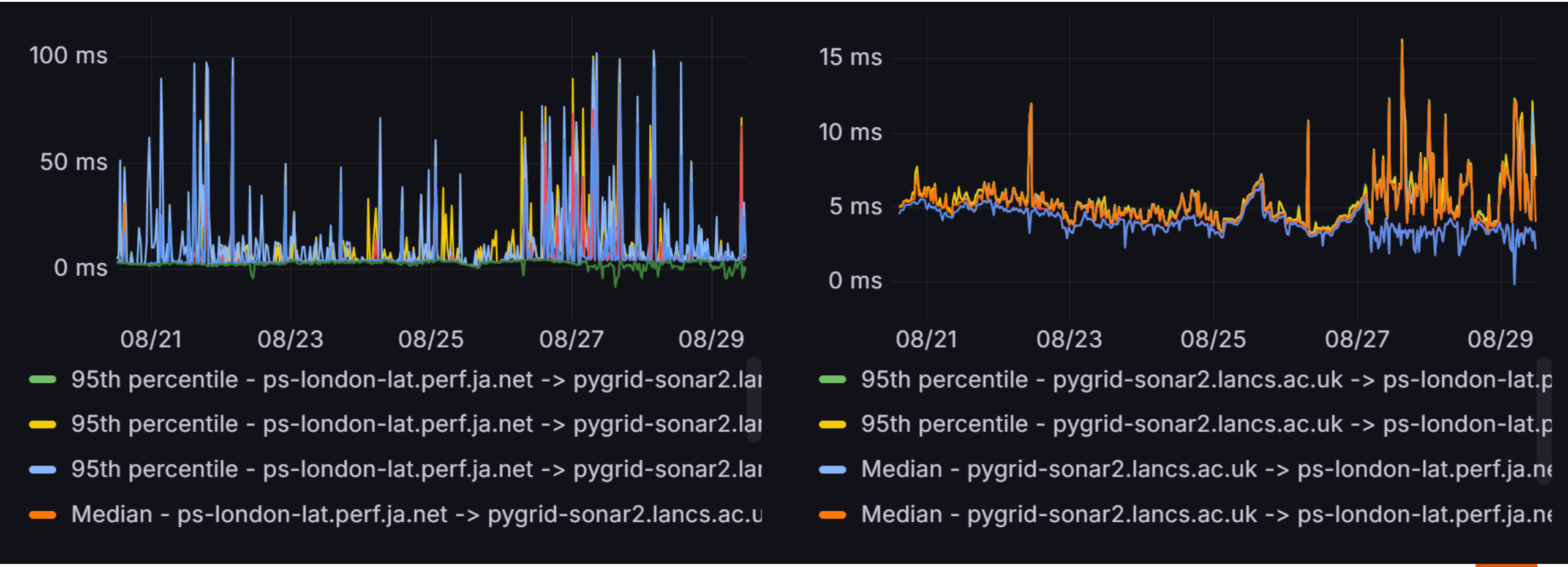
Jisc London -> Lancaster





Latency

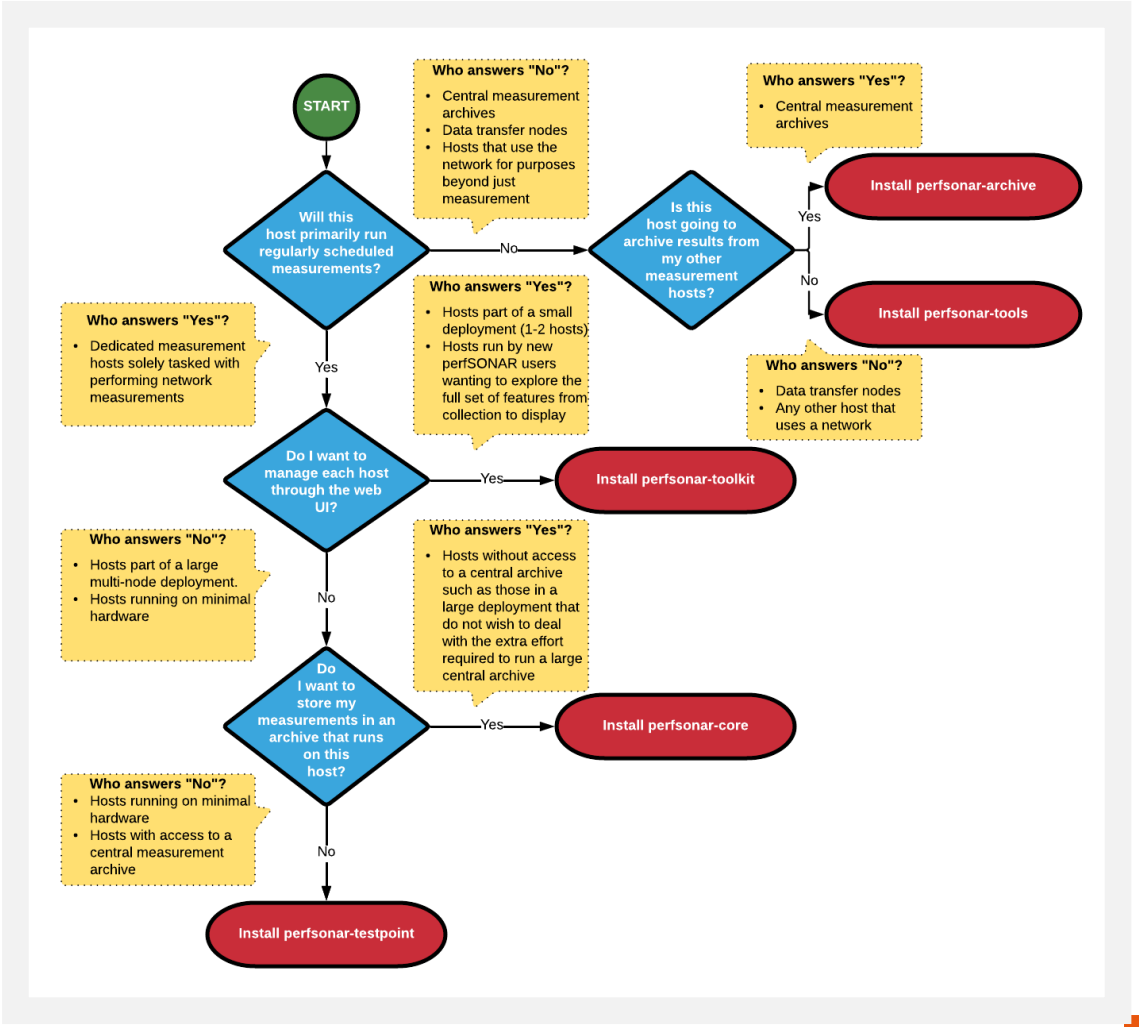
London -> Lancaster



Different perfSONAR Installation Options

These include:

- Toolkit
  - Fully featured
  - Local Opensearch, local Grafana UI, etc.
- **Testpoint (reduced resources)**
  - Sends results to a central archive
    - No local Opensearch
- Container
  - Docker
  - LXC containers (not yet released)
  - Recommend binding to real network card



Jisc perfSONAR node specs

Jisc public servers

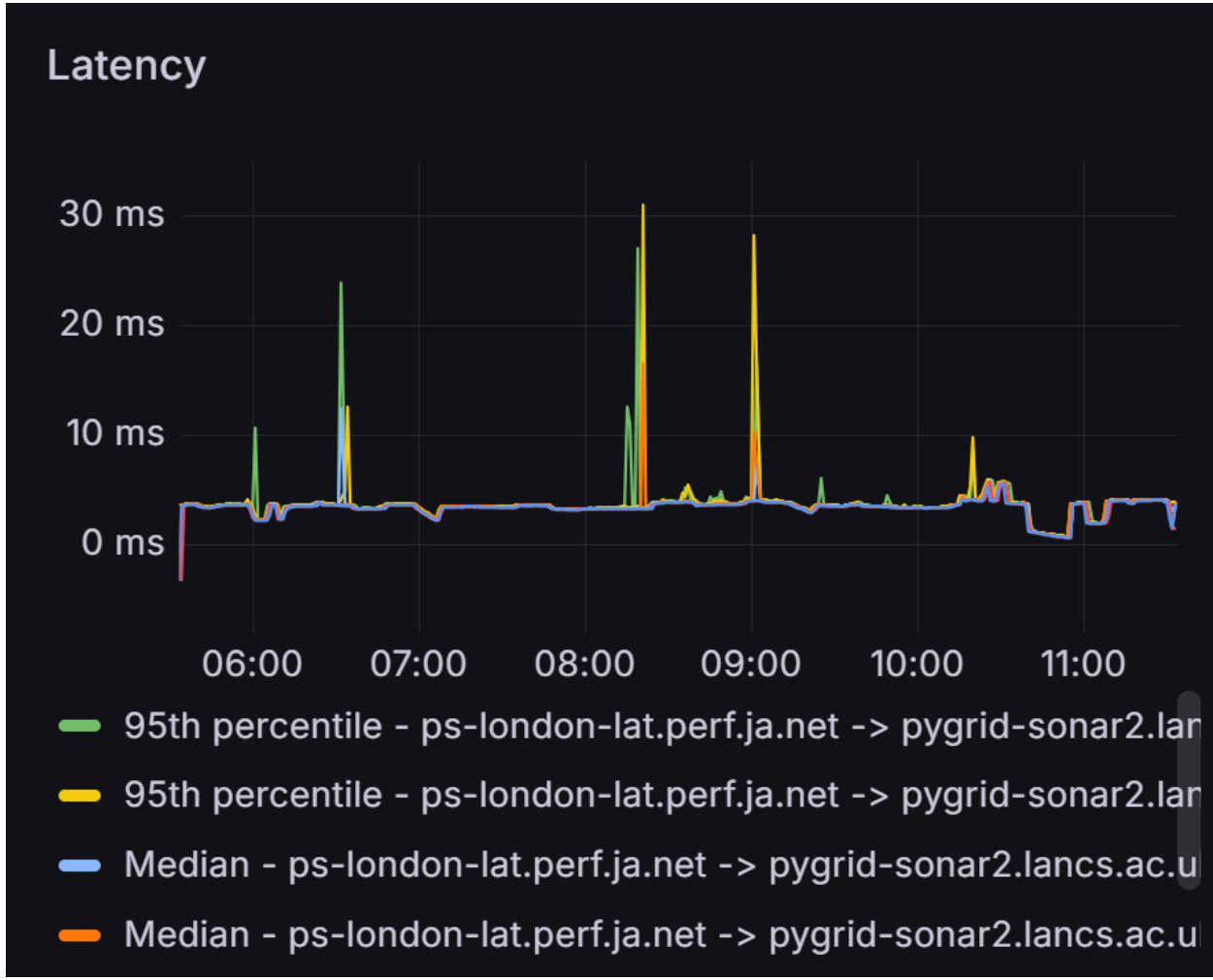
	London	Slough	small
Version	Toolkit	Toolkit	Testpoint
Network card	ConnectX-4	BCM5720	Intel X710
Bandwidth	100 Gb/s (MTU=9000)	10 Gb/s (MTU=9000)	1 Gb/s
CPU	Xeon 6130 (2.1 ->3.7 GHz)	Xeon 5122 (3.6 -> 3.7 GHz)	Celeron J1900
	2018	2017	~2016
RAM	128 GB	32 GB	8 GB
Swap	4 GB	14 GB	1 GB

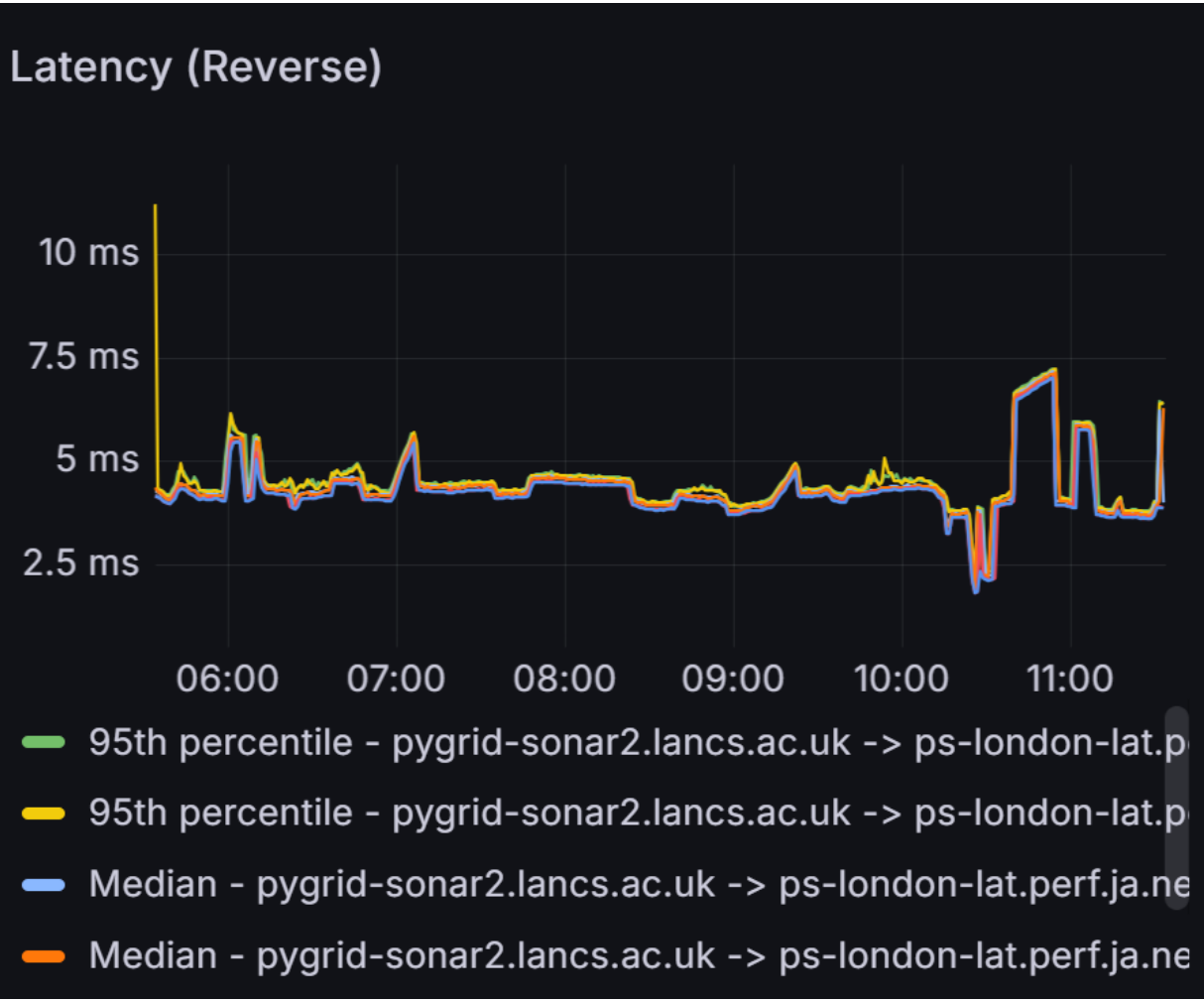
A little aside into hardware timestamping...

Improving NTP-based synchronisation

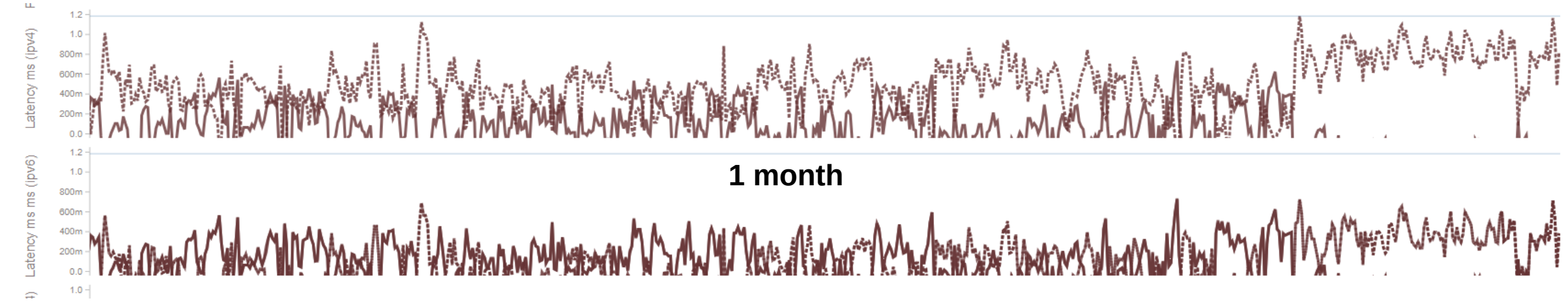
And measuring impact with perfSONAR latency tests

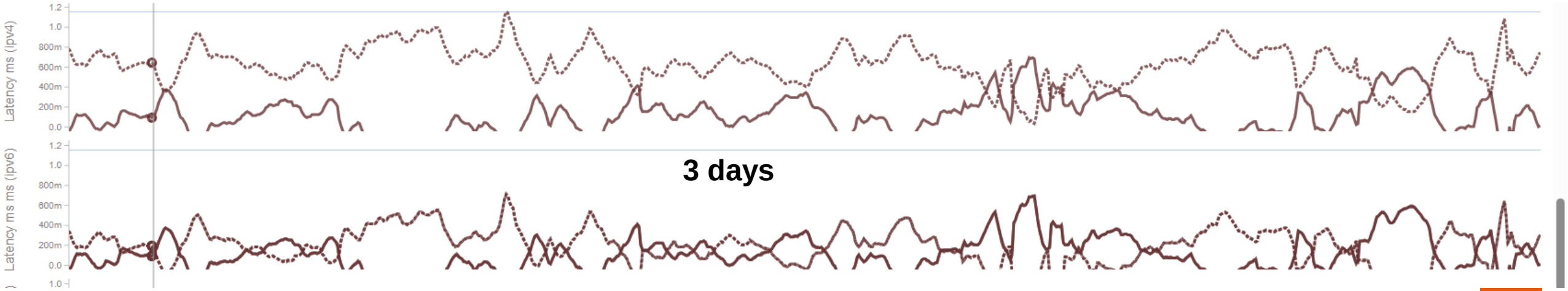
Latency Measurement





Perfsonar one way latency measurements
Jisc (London) to QMUL





Network Card Hardware timestamping

RHEL 9 “Basic System Settings - Chrony”

- Network card support?

```
# ethtool -T enp175s0f0np0
```

- Output

```
Time stamping parameters for enp175s0f0np0:
Capabilities:
        hardware-transmit
        hardware-receive
        hardware-raw-clock
PTP Hardware Clock: 0
Hardware Transmit Timestamp Modes:
        off
        on
```

Network Card timestamping (2)

- Enabling in **/etc/chrony.conf**

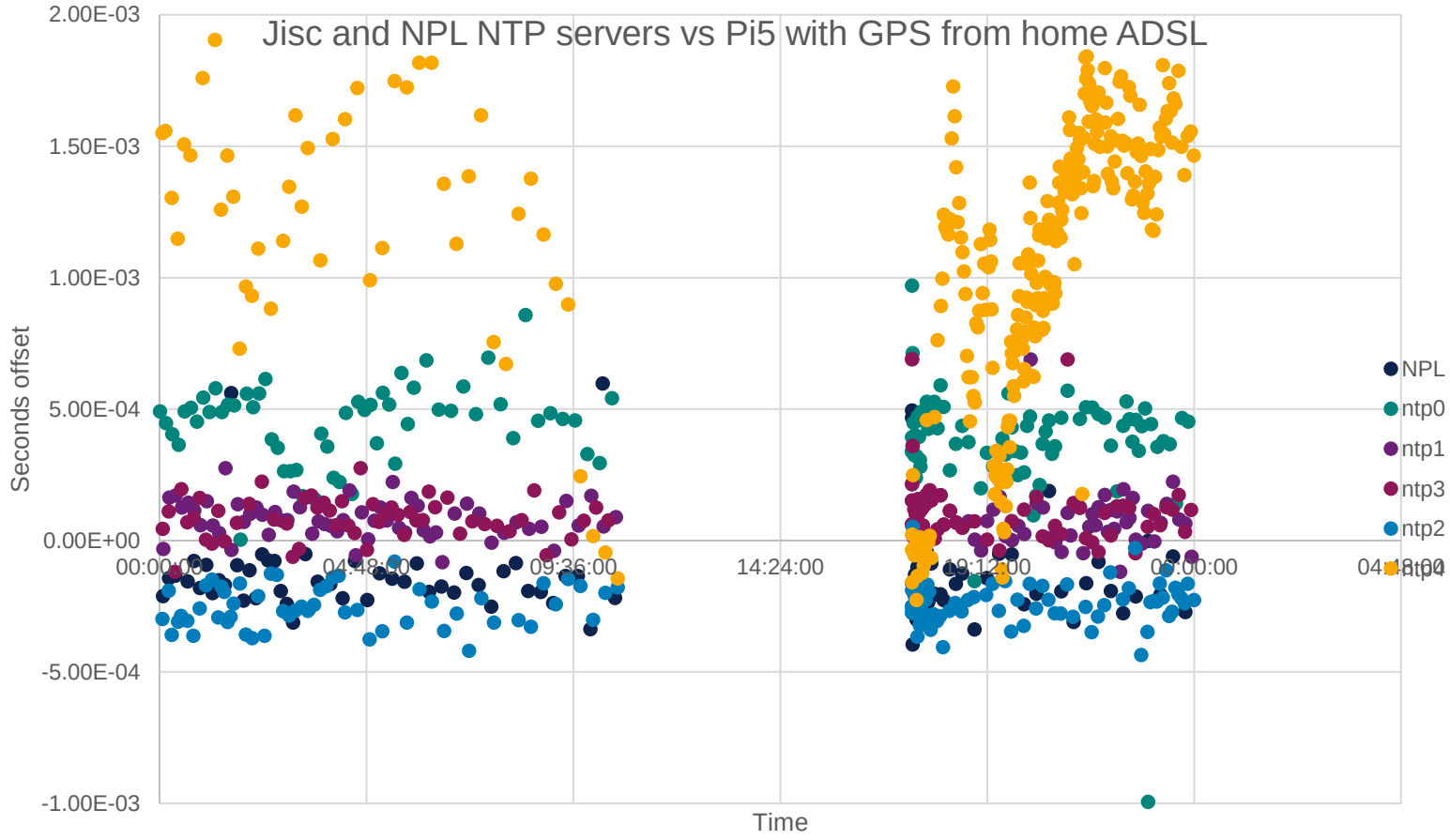
```
# Enable hardware timestamping on all interfaces that support it.

hwtimestamp *
```

- Checking

```
# chronyc ntpdata
…
Offset        : -0.000132777 seconds
Peer delay    : 0.005870986 seconds
Peer dispersion : 0.000003835 seconds
Response time   : 0.000374075 seconds
…
TX timestamping : Hardware
RX timestamping : Hardware
```

Improving clock accuracy/stability

- Accurate time servers?
  - How accurate/stable are yours?
    - RIPE Atlas Measurements?
    - Jisc plans to take a feed from NPL
- More accurate time transfer
  - NTP (draft-ietf-ntp-over-ptp-02)
  - PTP
    - See Simple PTP at Meta
  - White Rabbit



Related Jisc resources

Some related reading...

- Research Network Engineering (RNE) community talks, 2pm last Friday of each month
  - https://www.jisc.ac.uk/get-involved/research-network-engineering-rne-community-group
  - What talks would you like to see, or might you offer?
- Tim presented on IPv6 adoption on the WLCG at IETF 120 last month:
  - https://docs.google.com/presentation/d/1riTdi7zgoJ3ig31Hp-gy4Z089PeKaoRJHALOcvjGn5E/edit#slide=id.g279a4c0f603_0_12
- Expanding / enhancing our Jisc network performance test facilities
  - We aim to soon have a DTN with 100Gbit/s write/read
  - Beefing up our VM platform to support hosting perfSONAR archives & Grafana meshes
- **What's missing**?
  - What test (or other) facilities could help you? Let us know!

perfSONAR TLC

Some things to watch for

- If your IPv6 connectivity is broken, then various services fail
  - Caused by communications not falling back to IPv4
  - Only one or two sites have run into this of late
- The 5.1 testpoint build seems to have a memory bug – something eats memory over time then the "out of memory" handler can kill necessary processes
  - Devs actively working on this, expect a patch soon
- We believe that long-term running GridPP testpoint builds with a central archive (hosted by Shawn/Marian or Jisc) is the simplest and least moving parts model

- Worth remembering that all perfSONAR development is open
  - See https://github.com/perfsonar
  - Includes issue tracking; Raul submits and follows up via these

Jisc

Summary

We're here to help – talk to us about anything...

- perfSONAR:
  - Meshes (for GridPP or any other communities you're involved in)
  - Bugs or desired features – we have regular dialogue with the devs
  - Testpoint installations (rather than toolkit)
- Network time services
  - Hardware timestamping to improve NTP accuracy and stability
  - PTP – in the future

- Jisc will continue to support you
- Happy to support other communities – just ask

Jisc

Questions

Christopher.Walker@jisc.ac.uk

netperf@jisc.ac.uk

– Netperf team.

4 Portwall Lane, Bristol,
BS1 6NB

help@jisc.ac.uk

jisc.ac.uk