Experiences of using perfSONAR for network troubleshooting on Janet

Duncan Rand

Jisc

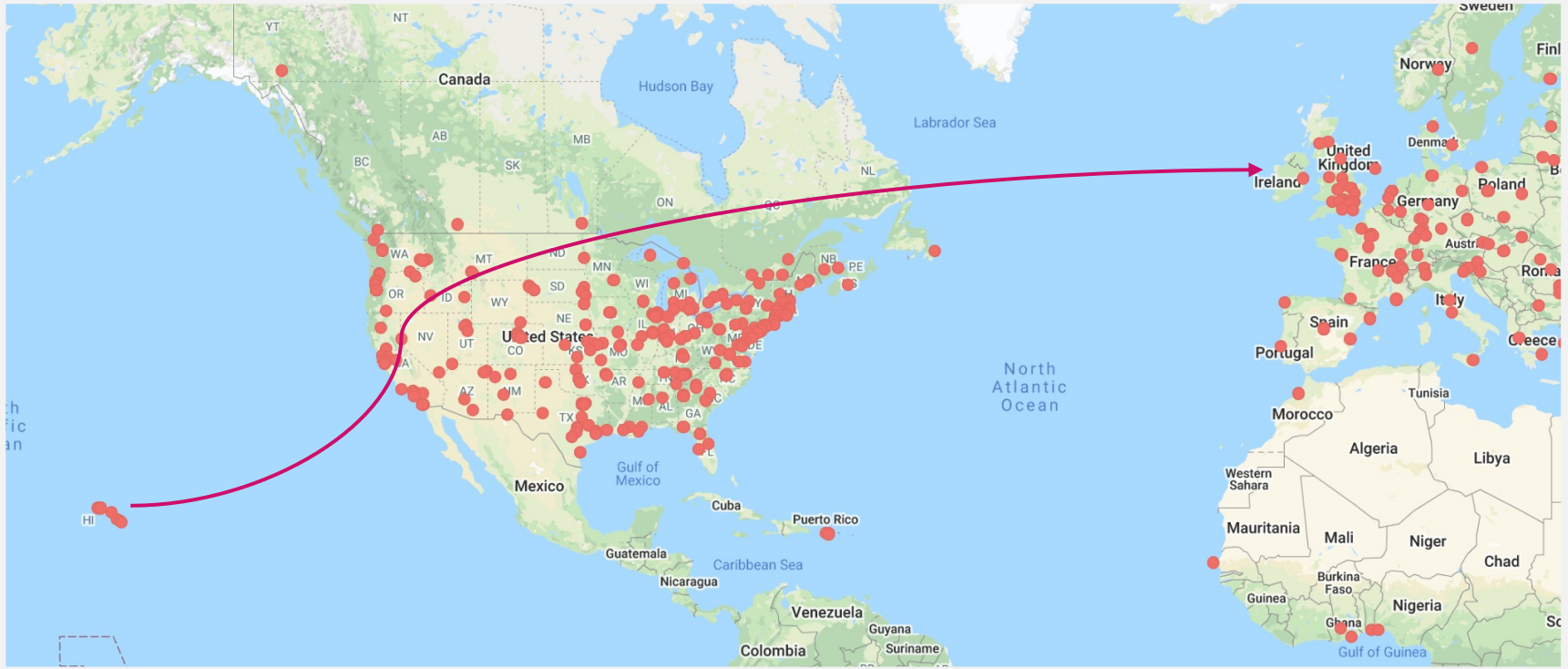netperf@jisc.ac.uk

Jisc

Network Connectivity

- Researchers rely more than ever on good network connectivity to carry out their work

- Datasets are growing in size year on year making the timely transfer of data increasingly more difficult

- But the global research and education network is made up of a disparate array of different networks which are owned, funded and operated independently

- These networks need to operate seamlessly to enable research collaborations to operate successfully

- Where do we start when things go wrong?

Jisc

An example: The ATLAS project

Asteroid Terrestrial-impact Last Alert System (ATLAS, https://fallingstar.com/)

- 'ATLAS will provide one day's warning for a 30-kiloton "town killer," a week for a 5-megaton "city killer," and three weeks for a 100-megaton "county killer".'

- Transferring astronomy data from Hawaii to Queens University Belfast (QUB), Northern Ireland

- Copying data from Hawaii – 11,000 km, 180 ms RTT

- Encountered a problem with throughput – significant drops

- Unable to transfer data in time to keep up with schedule

- The problem was intermittent - a "soft fault"

Jisc

Hawaii to Belfast (11,000 km, 180 ms RTT)

2nd European perfSONAR User Workshop

http://stats.es.net/ServicesDirectory/

Jisc

Transferring data at high throughput

Over long distances is not that easy…

- TCP-based transfers with high latencies are very sensitive to packet loss

- We need good monitoring to understand what is going on

- iperf3 (https://iperf.fr/) simulates data transfers with memory to memory test, typically 10-20s

- We can also use ping to measure latency and packet loss

- The traceroute command shows routes taken between hosts

- Useful for detecting asymmetric routing which can be a problem

Jisc

Network monitoring with perfSONAR

- perfSONAR (https://www.perfSONAR.net) is a tool for the automated collection of network characteristics

- Open source, free measurement middleware

- Installed on linux hosts which test each other

- Locate pS hosts alongside data stores and transfer nodes

- Used to find network problems e.g. soft failures

- Multi-domain environment

- Can form meshes of perfSONAR hosts
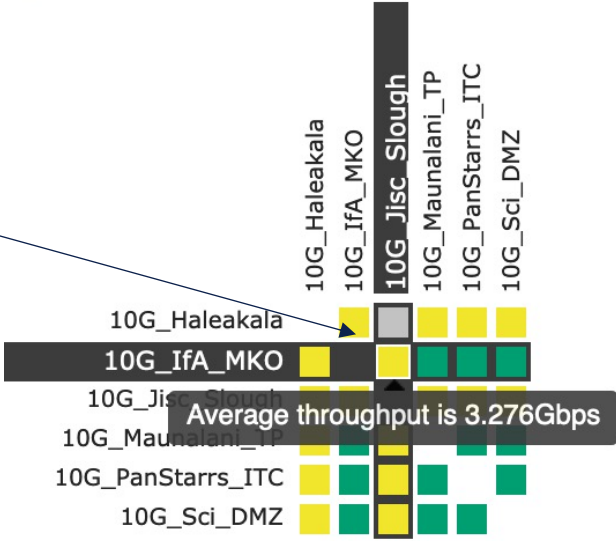
Jisc

perfSONAR hosts at Hawaii and in Belfast

- The usefulness of perfSONAR is maximised when the hosts are installed in good time before problems occur

- Generate a history of chosen metrics and their fluctuations over time

- Easier to troubleshoot when have historic data over time rather than trying to "run iperf" when problems arise

- Often the change as well as the absolute magnitude of a metric, such as throughput or loss, is valuable

- The ATLAS project had previously set up perfSONAR hosts at both endpoints and Baltimore, USA - in ATLAS perfSONAR mesh

Jisc

UH IfA - 10G Bandwidth Tests - Throughput

■ Throughput >= 5Gbps	■ Throughput < 5Gbps	■ Throughput <= 1Gbps

✓ No problems found in grid

Average throughput from Institute for Astronomy (IfA) Hawaii to Jisc host in Slough is 3.3 Gbps

	10G_Haleakala	10G_IfA_MKO	10G_Jisc_Slough	10G_Maunalani_TP	10G_PanStarrs_ITC	10G_Sci_DMZ
10G_Haleakala		■	■ (gray)	■	■	■
10G_IfA_MKO	■		■	■	■	■
10G_Jisc_Slough						
10G_Maunalani_TP					■	■
10G_PanStarrs_ITC	■	■	■			■
10G_Sci_DMZ	■	■	■	■	■	

Average throughput is 3.276Gbps

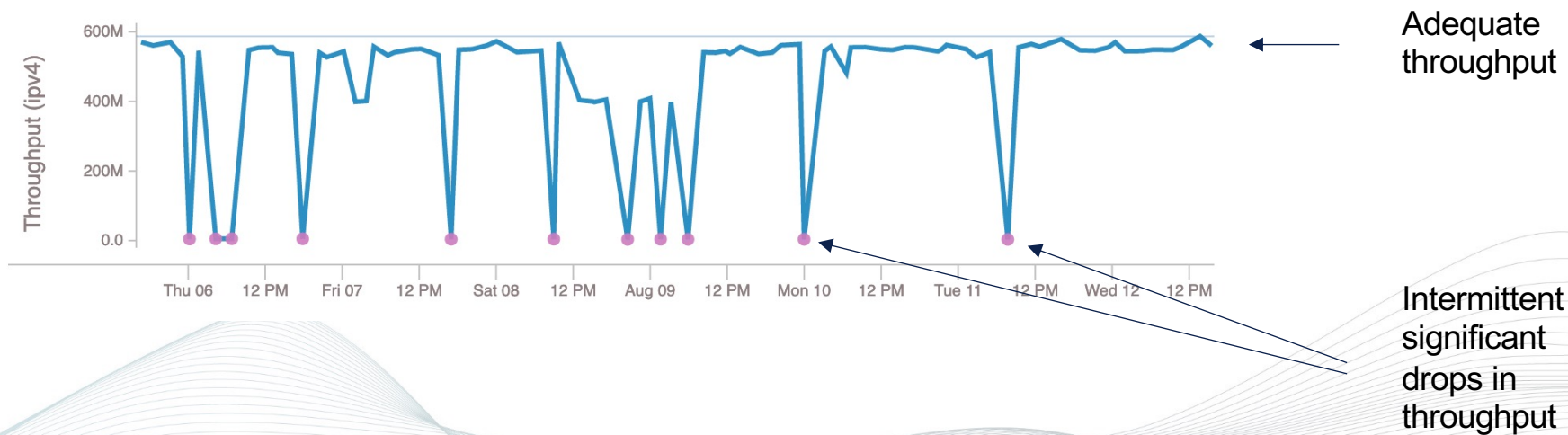2nd European perfSONAR User Workshop

Jisc

Intermittent or Soft faults

Potentially a significant problem: can be difficult to track down

- Hard faults (e.g. cut fibre or power failure) are easy to detect, but soft faults can remain undetected for years

- Can be caused by congested or faulty links – locally, intra-campus, or between domains

- Cause packet loss and degraded TCP performance

- Sometimes the problem is local but only appears when interacting with more distant hosts because TCP is more sensitive to higher latencies

(thanks to ESnet, see https://fasterdata.es.net/ for more details)

Jisc

ATLAS (Hawaii) to QUB (Belfast): intermittent fault

Adequate throughput

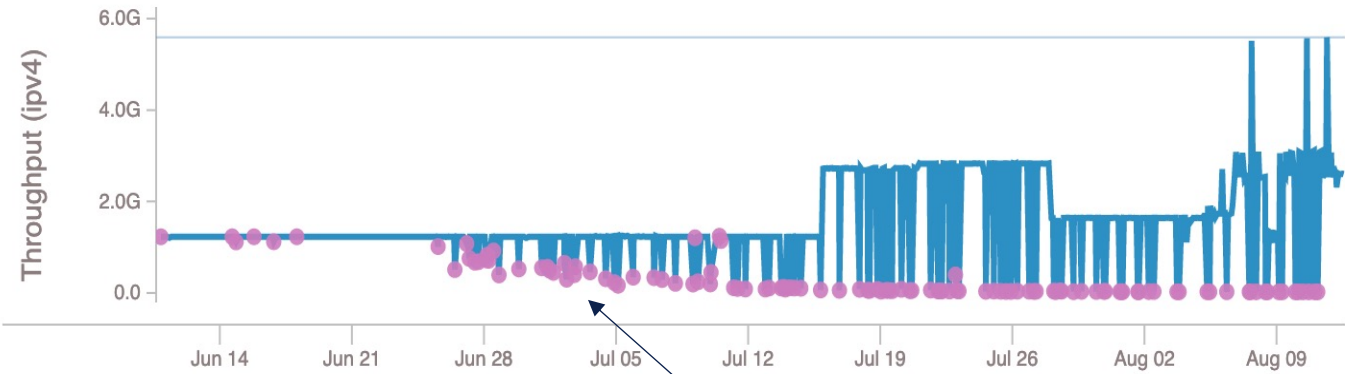Intermittent significant drops in throughput

2nd European perfSONAR User Workshop

Max (Baltimore) to Jisc (Slough)

Not just Hawaii to Belfast, also present between completely different perfSONAR hosts

2nd European perfSONAR User Workshop
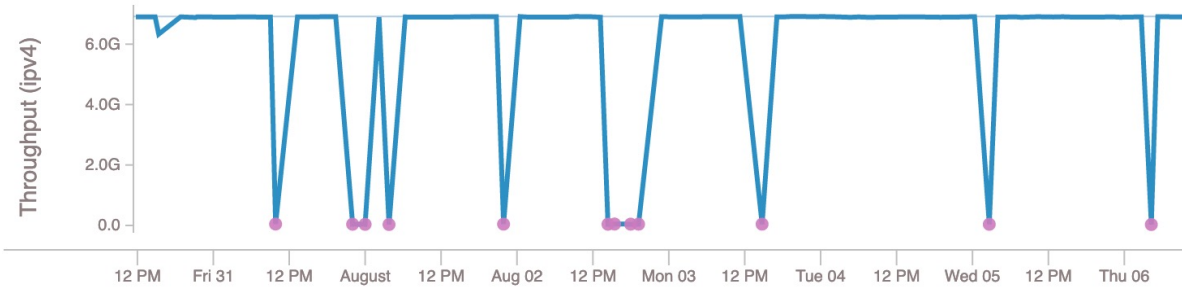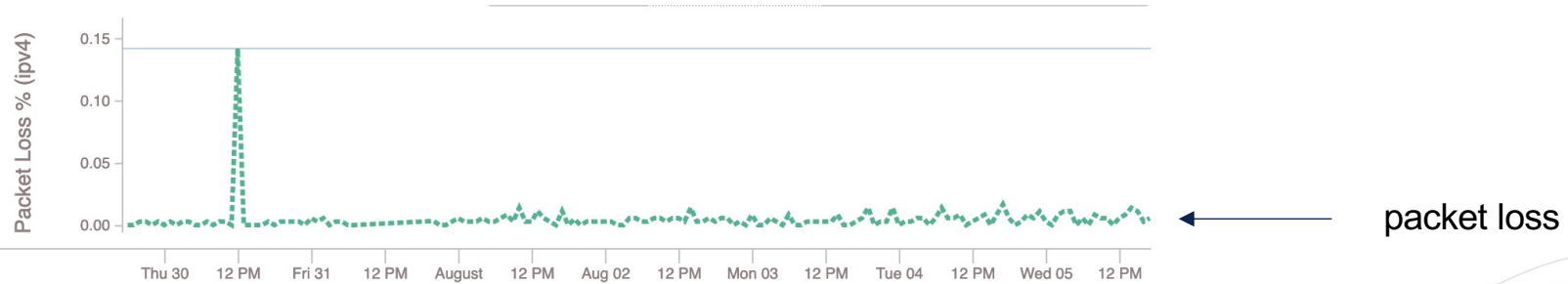
Jisc

ITC (Hawaii) to Jisc (Slough)

Gradual worsening of drops in throughput

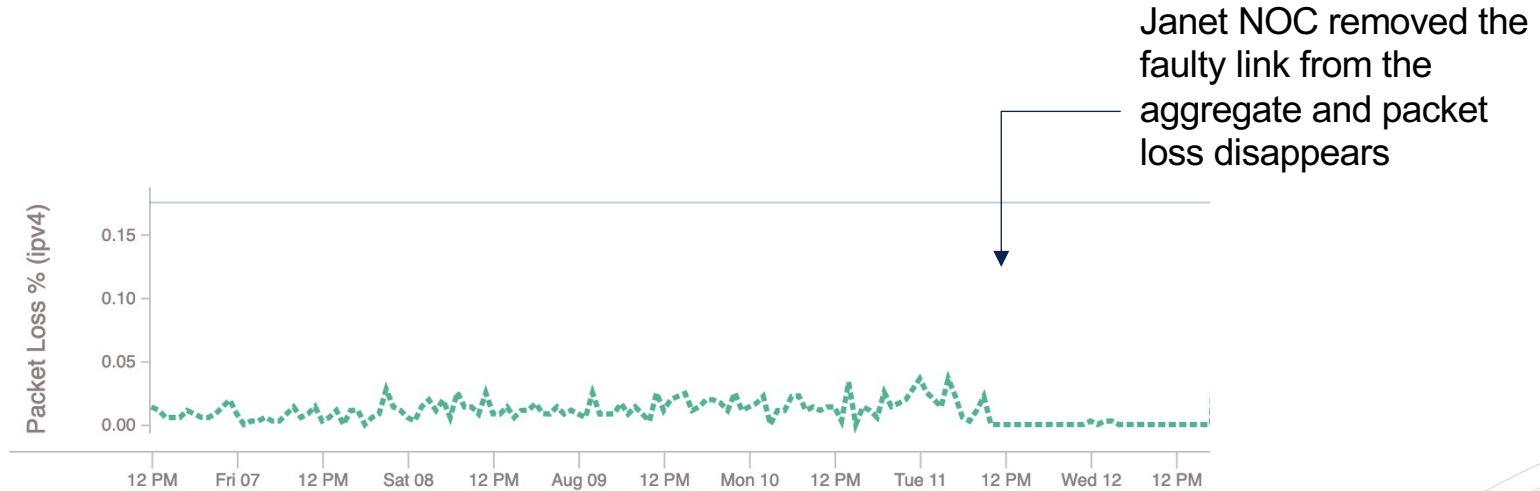2nd European perfSONAR User Workshop

Jisc

GÉANT (London) to Jisc (Didcot)

Also noticed continuous packet loss

packet loss

Narrowed it down to interface between GÉANT and Janet
Asked Janet NOC to investigate

MAX (Baltimore) to Jisc (Slough)

Diagnosis: faulty optic in 6 x 100G aggregate between
Janet London Harbour Exchange and London Powergate

Janet NOC removed the
faulty link from the
aggregate and drops in
throughput disappear

GÉANT (London) to Jisc (Didcot)

Janet NOC removed the faulty link from the aggregate and packet loss disappears
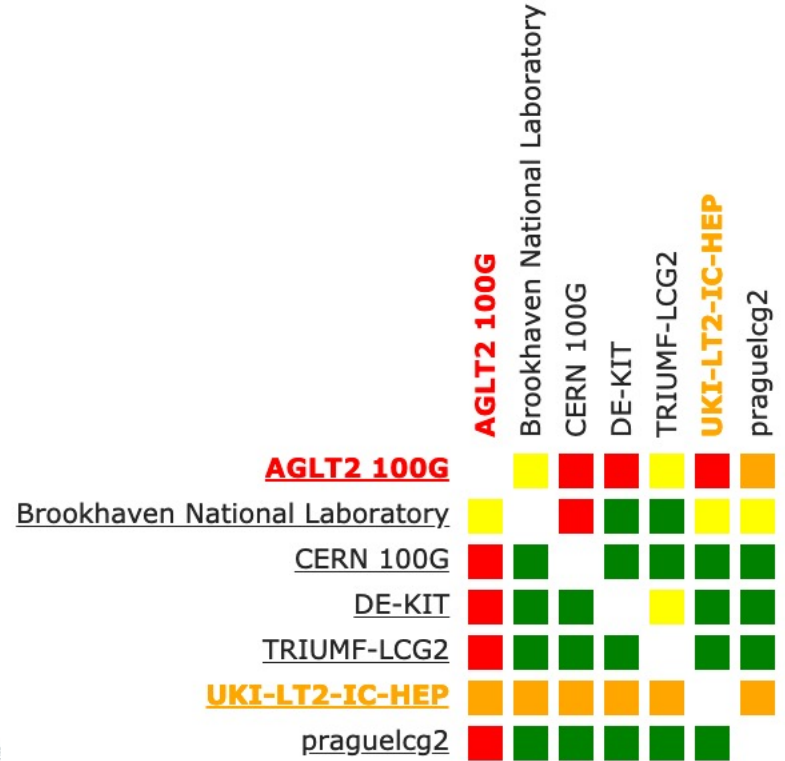

2nd European perfSONAR User Workshop

Jisc

In Conclusion

- Diagnosis: faulty optic in 6 x 100G aggregate between London Harbour Exchange and London Powergate

- Monitoring that simply measures traffic levels on interfaces often won't pick up these issues in the way perfSONAR can

- A number of perfSONAR hosts around the world enabled us to diagnose the problem

- Easily resolved once diagnosed – faulty link was removed from the aggregate

- Customer was able to resume timely copying of data from Hawaii to Belfast

- More examples of soft failures on fasterdata.net here https://fasterdata.es.net/performance-testing/perfsonar/perfsonar-success-stories/

Jisc

100 Gbps perfSONAR testing

- A number of sites are having their connections to Janet upgraded to 100 Gbps

- Imperial College and Rutherford Appleton Laboratory (RAL) have been upgraded

- Now a requirement to start testing with perfSONAR at rates greater than 10 Gbps

- Installed a 100Gbps testbed to gain experience

- The WLCG have set up a 100Gbps perfSONAR mesh:

  https://psmad.opensciencegrid.org/maddash-webui/index.cgi?dashboard=WLCG%20100G%20Mesh

Jisc

Services available for Janet members

- The Janet End to End Performance Initiative have the following available for testing connectivity

  - a 10 Gbps perfSONAR host in Slough: https://ps-slough-10g.ja.net/toolkit/ (part of GridPP mesh)

  - a 10 Gbps iperf server for ad hoc tests:
    e.g. $ iperf3 -c iperf-10g.ja.net

  - a 10 Gbps DTN running Globus Connect and a gridftp server

  - a RIPE ATLAS probe

- Feel free to contact us: netperf@jisc.ac.uk

Jisc