



28/11/2017

## perfSONAR – training & dev discussion summary

SIG-PMV Copenhagen, 28 Nov 2017

Tim Chown, Jisc, UK, [tim.chown@jisc.ac.uk](mailto:tim.chown@jisc.ac.uk)

- » Brief introduction to perfSONAR
  
- » Background: current use in Janet context
  - › User communities and Janet NOC
  - › Active vs passive measurements
  - › perfSONAR supporting Janet E2EPI activity
  
- » Outcomes of perfSONAR training held in Jisc's Manchester offices on 22/23 Nov
  - › Attendees
  - › Feature/ wish list
  - › Topics arising

- » An active measurement system:
  - › Open source; Linux platforms; ISO/RPMs; current version 4.0.1
  - › Measures loss, latency, throughput, path
  - › Continuous loss/latency, periodic throughput (default 4x per day)
  - › Scheduled or ad-hoc tests through pScheduler (introduced in 4.0)
  - › Back end database to store historical data
  - › Web or CLI-based management
  - › Web-based visualisation tools; support for measurement meshes
- » Typical deployment:
  - › Alongside data transfer endpoint (DTN) and at campus edge
  - › Allows performance bottlenecks to be identified
- » See <https://www.perfsonar.net/>

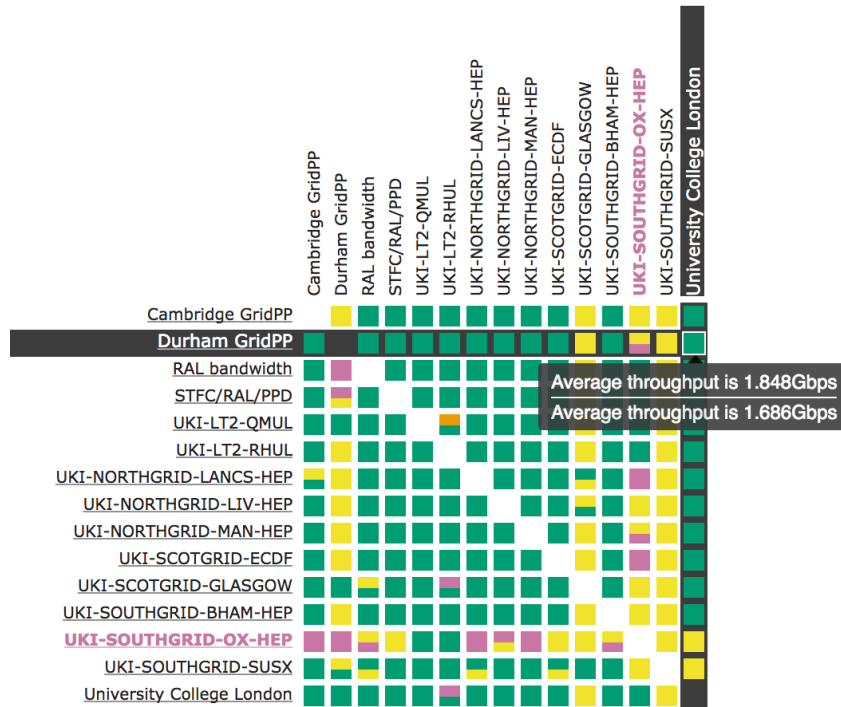
- » Used for some time by the WLCG including UK GridPP community
  - › ~20 UK sites, mesh run between LHC experiment participants
  - › perfSONAR nodes generally installed by campus IT staff
  - › WLCG pushing IPv6, so mesh is dual-stack where possible
  - › Mesh is publically available:
  - › <http://ps-dash.dev.ja.net/maddash-webui/index.cgi?dashboard=UK%20Mesh%20Config>
  
- » Jisc's E2EPI is working with sites to improve application throughput
  - › Identifying 'problem' cases, e.g., data transferred by hard disks / tape
  - › Supporting the work through establishing baseline perfSONAR measurements
  - › Encouraging wider use of perfSONAR; hence the training event

# Example: UK GridPP perfSONAR mesh

## UK Mesh Config - IPv4 Bandwidth Tests



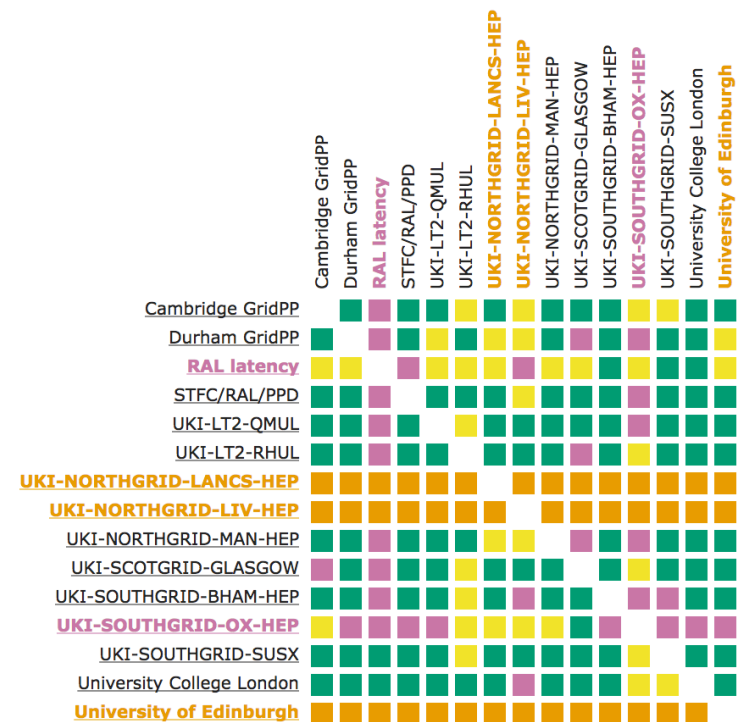
! Found a total of 1 problem involving 1 host in the grid



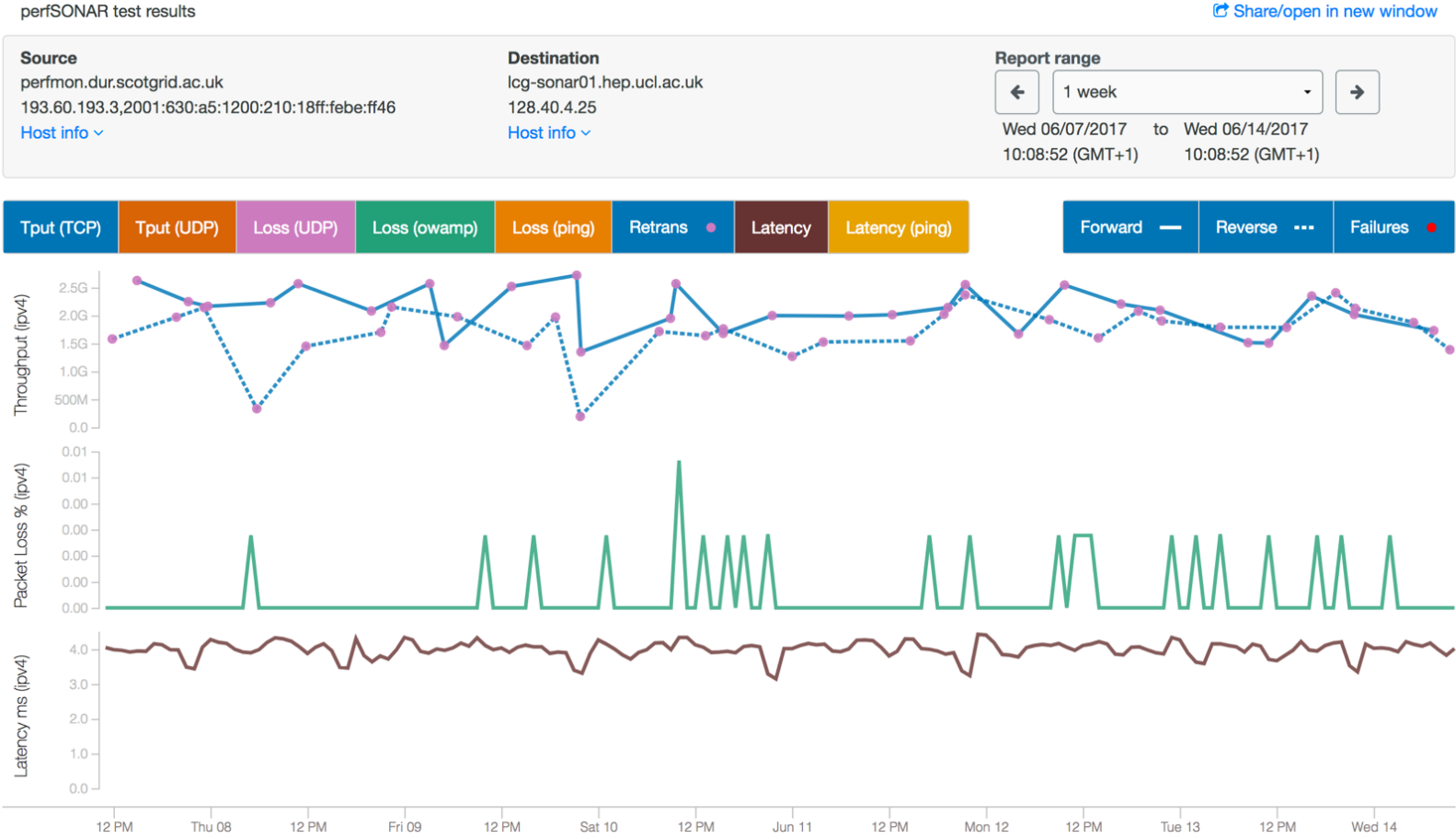
## UK Mesh Config - IPv4 Latency Tests



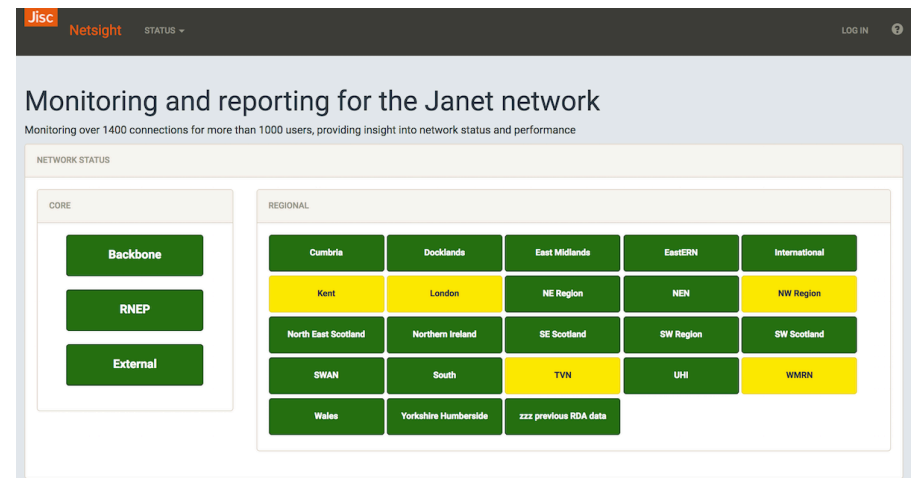
! Found a total of 6 problems involving 5 hosts in the grid



# Drilling down on a specific test pair...



- » We're undergoing a review of our network monitoring and management systems
- » NOC largely uses passive measurement tools
- » Current functions:
  - › Up/down status checks on Janet network elements
  - › Link utilisation data, allowing display of (private) network "weather maps"
  - › Site link utilisation; very helpful for site capacity planning
- » The Janet NOC has an OSS (private)
- » Universities have Netsight view
  - › <https://netsight.ja.net/>
  - › More detail if logged in



Last updated: 10th March 2017

## South

- RNEP PoP
- Regional PoP
- Janet6 Router

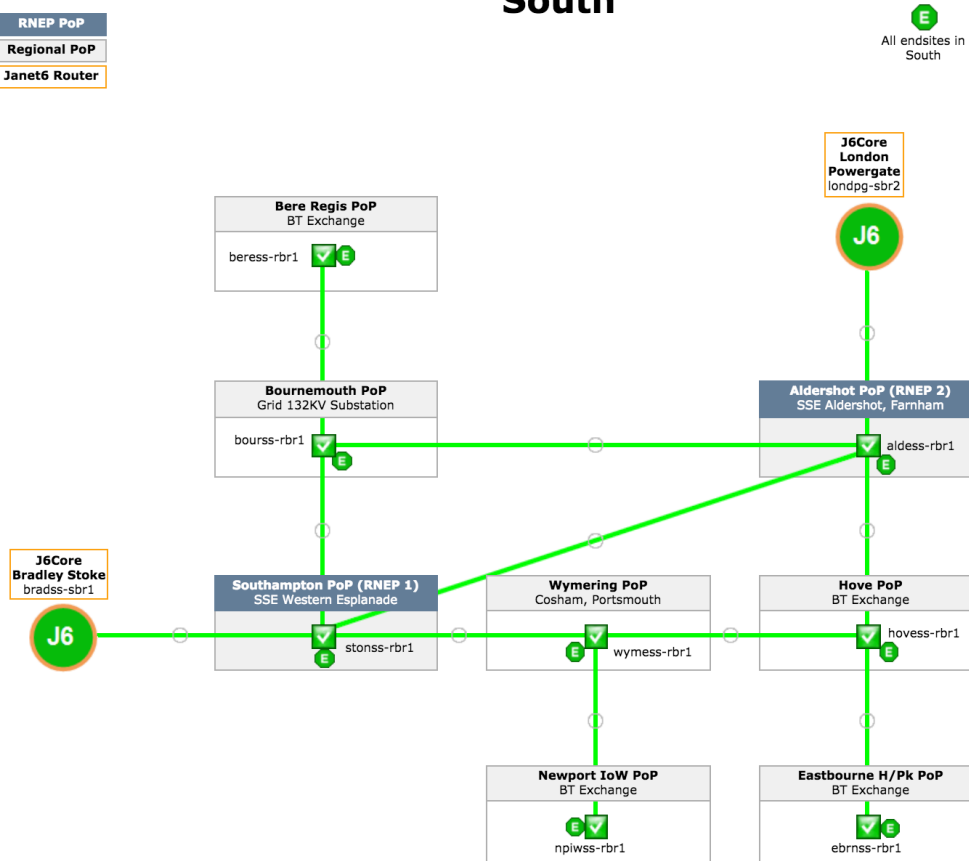
**Current Status**

**Tactical Overview**

- Maps**
- Overview
  - IP Core
  - Optical(Geo)
  - Optical(Topo)
  - Cumbria
  - East
  - East Midlands
  - Highlands & Islands
  - Kent / KPSN
  - London
  - London Docklands
  - North East
  - North East Scotland
  - North West
  - Northern Ireland
  - South
  - South East Scotland
  - South West
  - South West Scotland
  - Stirling
  - SWAN
  - Thames Valley
  - Transpennine
  - Wales / PSBA
  - West Midlands
  - Yorks & Humber
- Alarms (Live)**
- Optical
  - Optical (2)
  - Optical (jsd)
  - IP
  - IP (jsd)
- Alarms History (24hrs)**
- Optical
  - IP
- Interface tables**
- Management Routers**
- MRS Routers**
- Trouble Ticket Info**

**Utilities**

Quick Search:



**Navigation Panel**

**Regions**

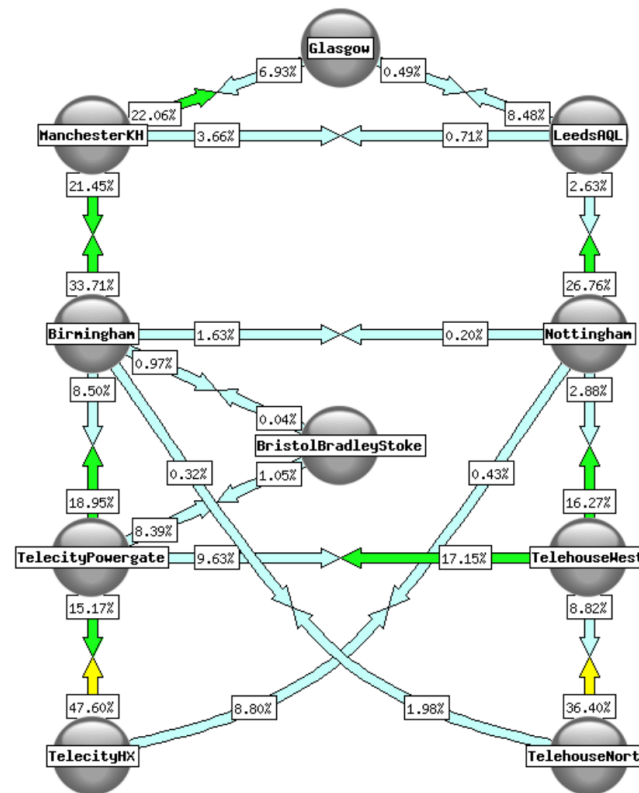
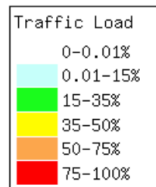
<input checked="" type="checkbox"/> Cumbria & N. Lancs	<input checked="" type="checkbox"/> South
<input checked="" type="checkbox"/> East England	<input checked="" type="checkbox"/> South East Scotland
<input checked="" type="checkbox"/> East Midlands	<input checked="" type="checkbox"/> South West
<input checked="" type="checkbox"/> Highlands & Islands	<input checked="" type="checkbox"/> South West Scotland
<input checked="" type="checkbox"/> Kent	<input checked="" type="checkbox"/> Stirling
<input checked="" type="checkbox"/> London	<input checked="" type="checkbox"/> SWAN
<input checked="" type="checkbox"/> London Docklands	<input checked="" type="checkbox"/> Thames Valley
<input checked="" type="checkbox"/> North East	<input checked="" type="checkbox"/> Trans Pennine
<input checked="" type="checkbox"/> North East Scotland	<input checked="" type="checkbox"/> Wales
<input checked="" type="checkbox"/> North West	<input checked="" type="checkbox"/> West Midlands
<input checked="" type="checkbox"/> Northern Ireland	<input checked="" type="checkbox"/> Yorkshire & Humberside

**Core**

<input checked="" type="checkbox"/> IP	<input checked="" type="checkbox"/> Optical
--	---

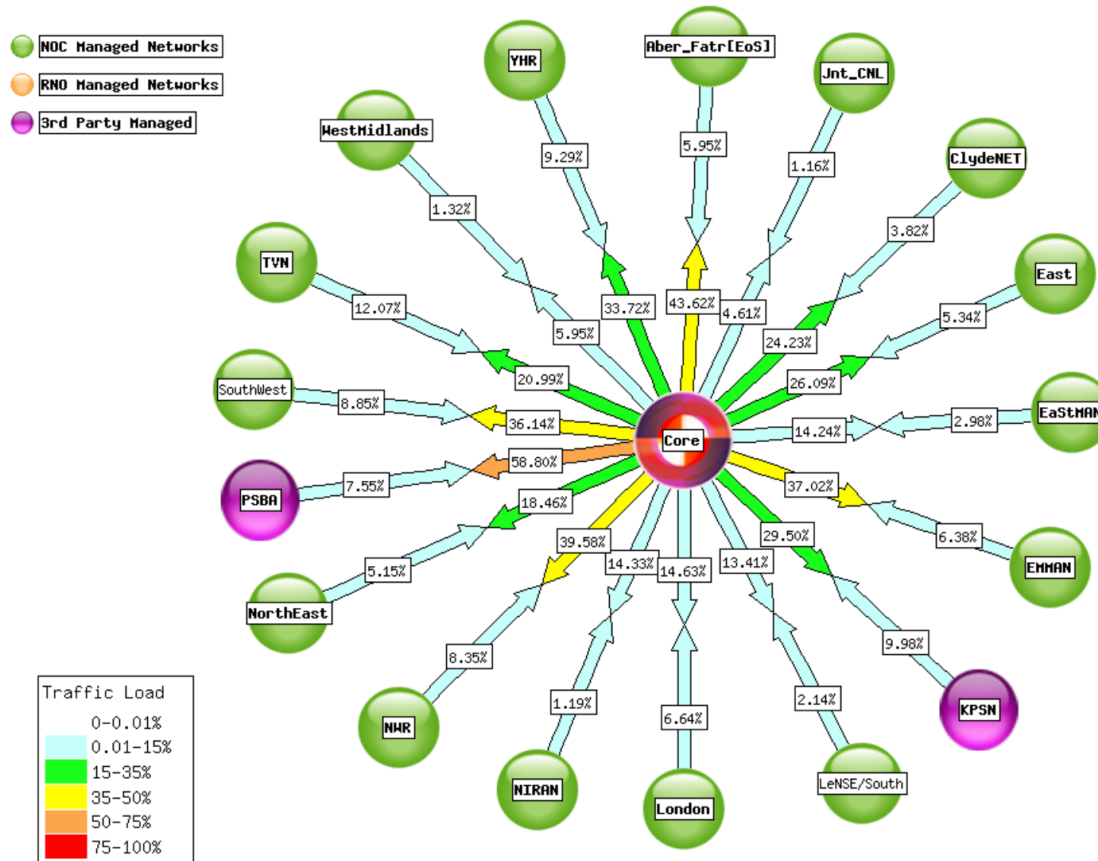


## Core Network J6



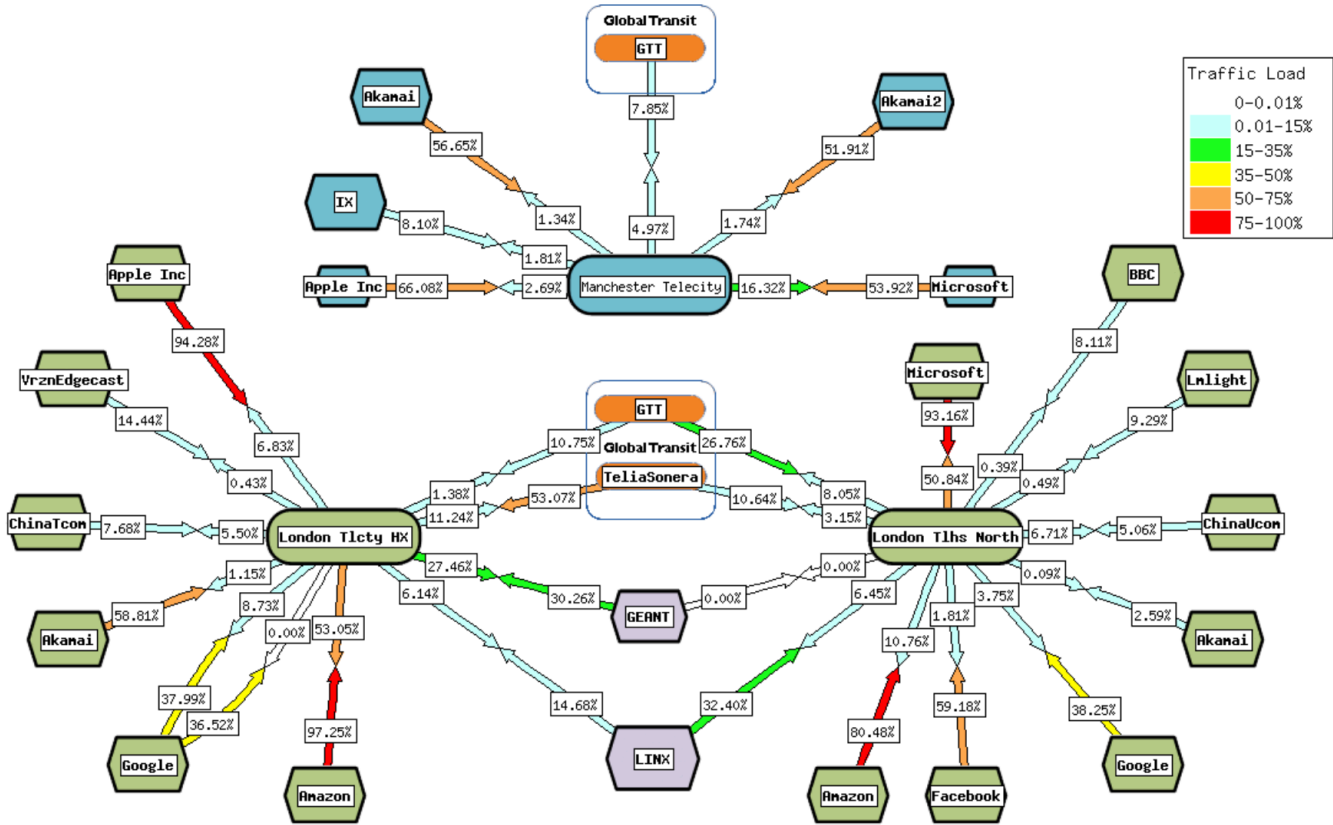
Created: Nov 14 2017 11:25:01

## Regional Network Entry Points



Created: Nov 14 2017 11:20:02

## Major External Links



- » Piloting perfSONAR nodes on Janet network
  - › With a view to a wider rollout; details TBD, hence dev discussion
- » Have deployed 10G perfSONAR node at London in Harbour Exchange
  - › Dell R620; 2 Intel 2.6 GHz processors (32 cores) and 32 GB memory
  - › Intel X520 DP 10Gb DA/SFP+ Server Adapter with 10G and 1 G single-mode SFPs
- » Second 10G perfSONAR node coming soon in our Slough data centre
  - › Alongside a 10G DTN for disk to disk & transfer tool tests
- » Jisc is running a VM that acts as a mesh server
  - › Results currently pulled from measurement points
  - › In future, likely to also store data centrally



# perfSONAR toolkit home page

Added a Jisc certificate

Dual-stack

Possible to set up tests manually, but better to set up a mesh...

perfSONAR Toolkit on ps-londhx1-mgmt.ja.net

ps-londhx1-mgmt.ja.net at 194.83.97.214, 2001:630:3c:f801::6

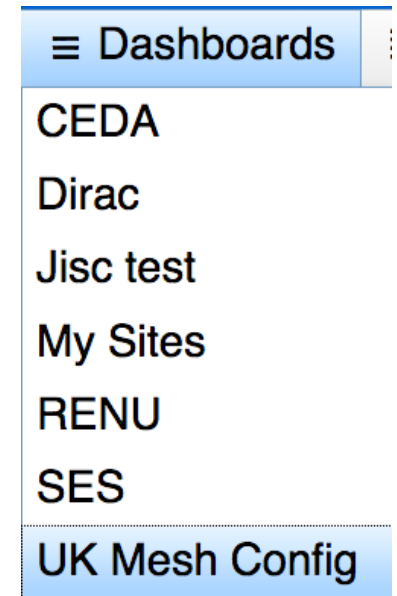
Organization: Jisc  
Address: London GB (map)  
Administrator: Duncan Rand (duncan.rand@jisc.ac.uk)

SERVICE	STATUS	VERSION	PORTS	SERVICE LOGS
bwctl	Running	1.6.4-1.el6	4823	<a href="#">View</a>
esmond	Running	2.1.1-1.el6		<a href="#">View</a>
lsregistration	Running	4.0.1-1.el6		<a href="#">View</a>
meshconfig-agent	Running	4.0.1-1.el6		<a href="#">View</a>
owamp	Running	3.5.4-1.el6	861	<a href="#">View</a>
pscheduler	Running	1.0.1.2-1.el6		<a href="#">View</a>

Test Results (18 Results) [Configure tests](#)

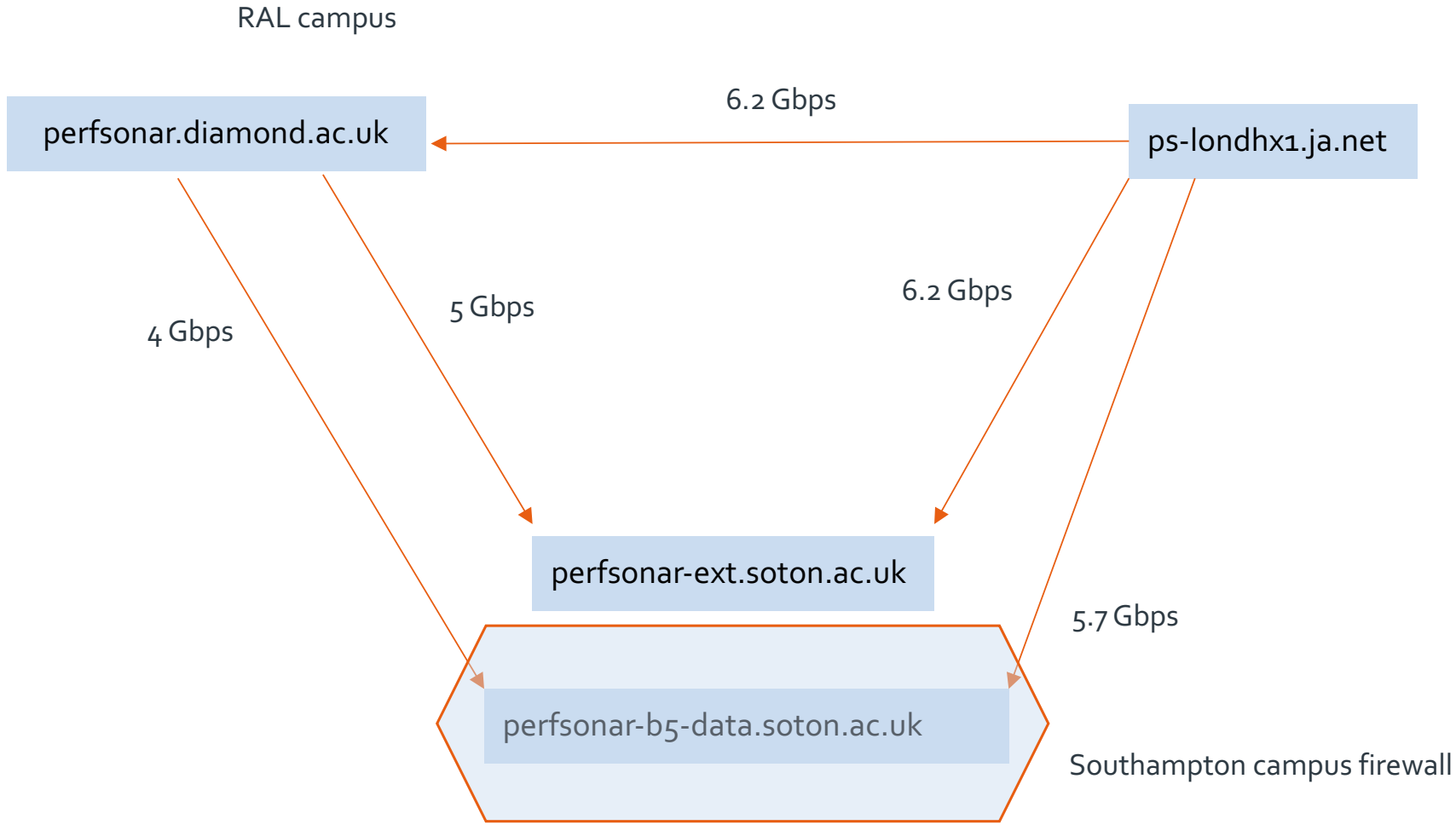
Search:  Results for the last...

- » Supporting several communities that we have been working with in the E2EPI
- » *UK Mesh Config*: GridPP mesh, part of the WLCG
- » *SES* - Science and Engineering South
  - › Pilot with Southampton University and Diamond Light Source
- » <https://ps-dash.dev.ja.net/maddash-webui/>



≡ Dashboards
CEDA
Dirac
Jisc test
My Sites
RENU
SES
UK Mesh Config

- » Materials science researcher at Southampton, getting a few 10's of Mbit/s for data transfers from Diamond to local lab filestore
  - › Moving 10-40TB, six times a year, on physical disks
- » Initial work involved adopting Globus Connect to transfer files from DLS
  - › Achieved a significant improvement; able to fill 1 Gbit/s local link
- » Also installed a perfSONAR host (*perfsonar-b5-data.soton.ac.uk*) on campus next to data storage
- » Network to storage upgraded to 10 Gbps
  - › Then achieving a few Gbit/s
- » Later a perfSONAR host (*perfsonar-ext*) was installed at the Soton border, outside the firewall
  - › perfSONAR very useful for understanding effects of changes

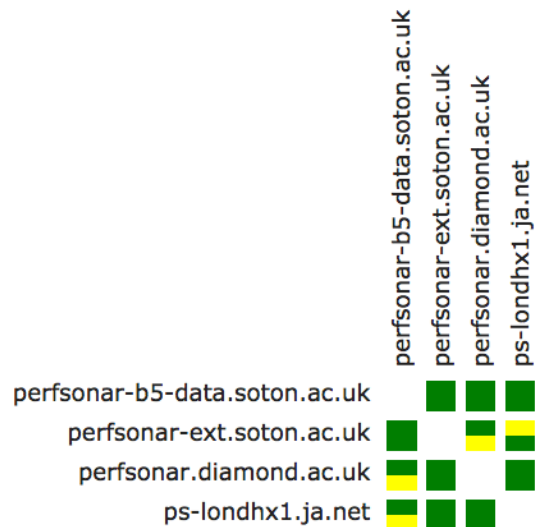




## SES - Traceroute

■ Number of Paths is <= 1
 ■ Number of Paths is >= 2

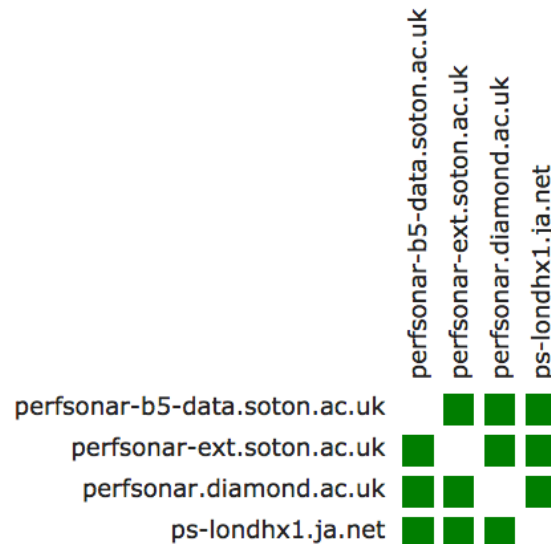
✔ No problems found in grid



## SES - Throughput Testing

■ Throughput >= 900Mbps
 ■ Throughput < 900Mbps

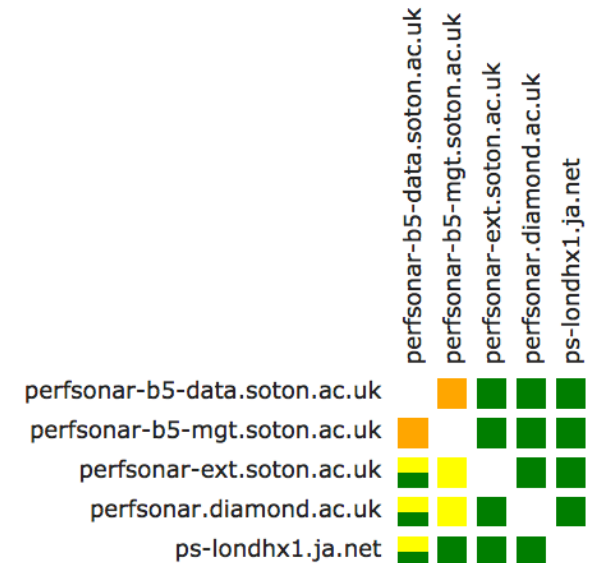
✔ No problems found in grid



## SES - Latency Testing

■ Loss rate is <= 0
 ■ Loss rate is >= 0
 ■ Loss rate is >= 0

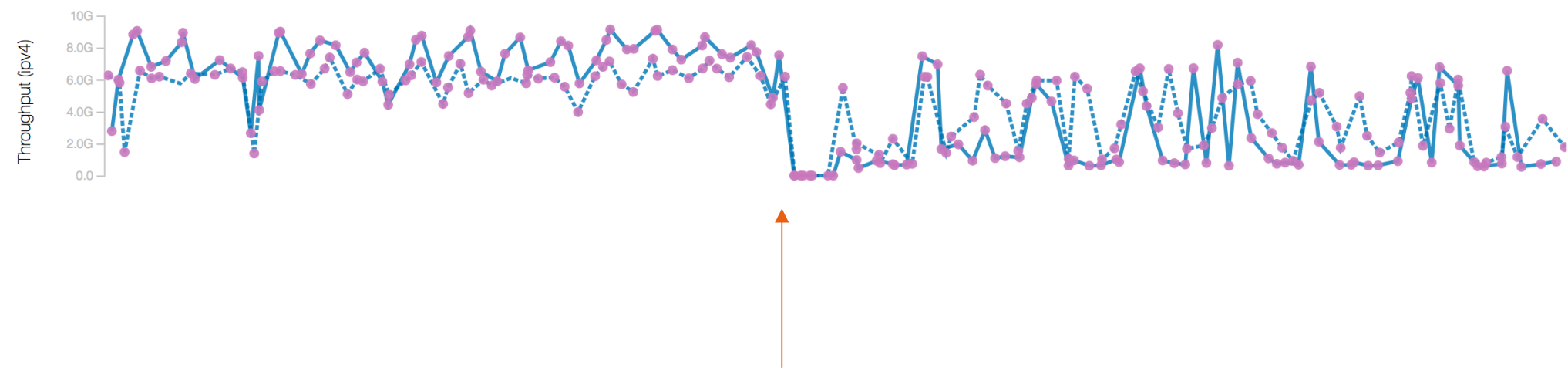
✔ No problems found in grid



# Inbound throughput to Soton

<b>Source</b> ps-londhx1.ja.net 194.83.97.209,2001:630:3c:f800:0:0:0:209 <a href="#">Host info</a> ▾	<b>Destination</b> perfonar-b5-data.soton.ac.uk 152.78.176.16 <a href="#">Host info</a> ▾	<b>Report range</b> ← 1 month → Tue 08/15/2017 12:08:30 (GMT+1) to Fri 09/15/2017 12:08:30 (GMT+1)
---	--	--

Tput (TCP) Tput (UDP) Loss (UDP) Loss (owamp) Loss (ping) Retrans Latency Latency (ping) Forward Reverse Failures

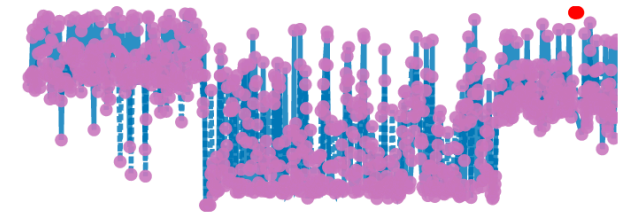
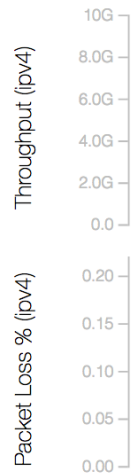


Firewall update  
(30th Aug)

# Inbound throughput to Soton

<b>Source</b> ps-londhx1.ja.net 194.83.97.209,2001:630:3c:f800:0:0:0:209 <a href="#">Host info</a> ▾	<b>Destination</b> perfsonar-b5-data.soton.ac.uk 152.78.176.16 <a href="#">Host info</a> ▾	<b>Report range</b> ← 1 year → Wed 11/16/2016 11:12:04 (GMT+0) to Thu 11/16/2017 11:12:04 (GMT+0)
---	---	---

Tput (TCP)	Tput (UDP)	Loss (UDP)	Loss (owamp)	Loss (ping)	Retrans ●	Latency	Latency (ping)	Forward —	Reverse ---	Failures ●
------------	------------	------------	--------------	-------------	-----------	---------	----------------	-----------	-------------	------------

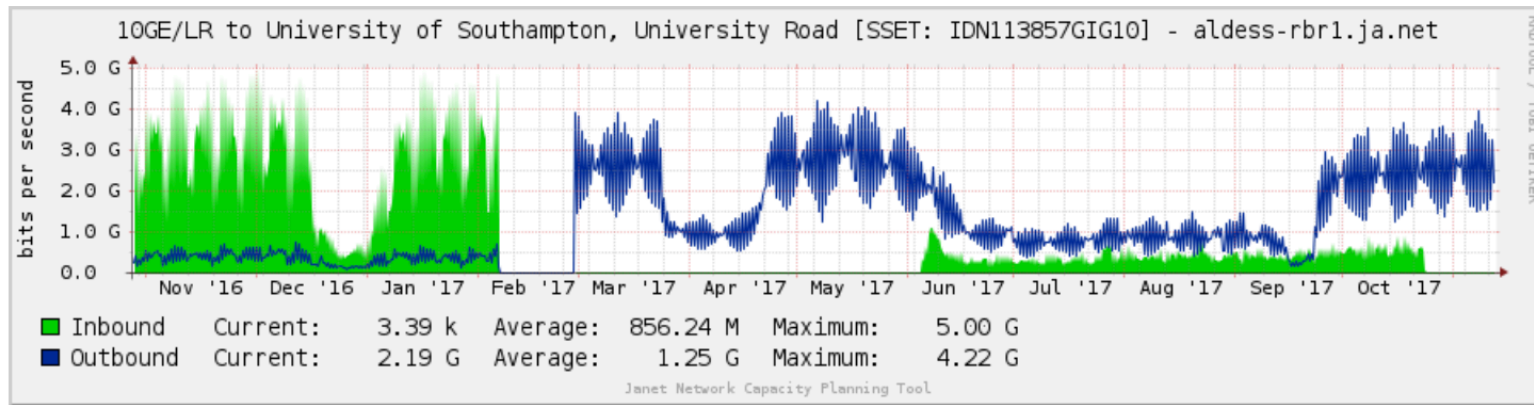


Firewall update  
(30th Aug)

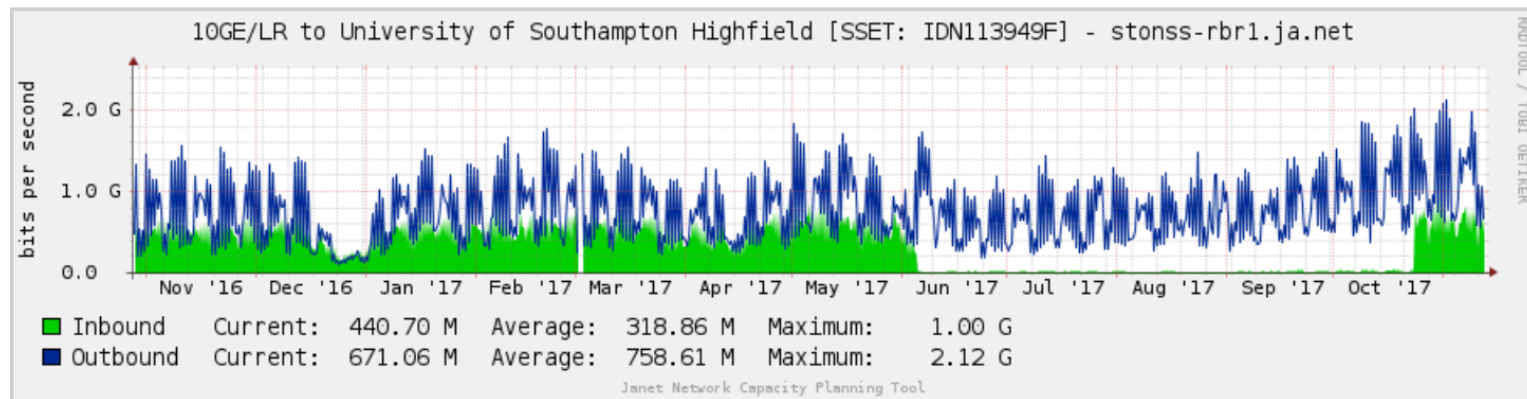
Issue fixed  
(24<sup>th</sup> Oct)

# Univ. Southampton – passive link measurement

## Last 12 Months



## Last 12 Months



- » As we saw earlier, the Janet NOC largely uses passive measurements
  - › Perfectly good for checking device status
  - › Or router interface utilisation; e.g., campus uplinks to Janet
  - › But as the Southampton example shows; it's not so good for detecting performance issues
  - › Note: if users have never experienced good throughput, they're unlikely to complain if they get a bad experience
  
- » Hence the desire/interest in also using active measurements
  - › Juniper RPM probes, and maybe TWAMP, as presented at a previous SIG-PMV
  - › perfSONAR; whether deployed by campuses or the Janet NOC
  - › Ideally we would maximise the value of campus and NOC deployments

- » Organised by Jisc, and held in Jisc's offices in Manchester on 22/23 Nov
- » Trainers were developers on the GEANT project
  - › Ivan Granizov and Antoine Delvaux
- » Plus training VMs provided by Kurt Baumann at SWITCH
- » 1.5 days training, 0.5 days developer discussion
- » 20 attendees; free attendance
  - › From Jisc, STFC (science facilities), ECMWF, universities
- » Logistics and attendees:
  - › <https://eventr.geant.org/events/2785>
- » Agenda and materials:
  - › <https://wiki.geant.org/display/gn42na1/perfSONAR+training+@+Jisc+2017>
- » Many feature requests raised, and ideas discussed in the day 2 discussion







- » The training event raised a few “why can’t I?” type questions, e.g.:
  - › In the web visualisation, can we have specific time periods entered rather than fixed durations?
  - › Displaying results in two directions in the mesh is confusing; is that really needed?
  - › The traceroute ‘errors’ shown in the mesh may be due to local ECMP/load balancing; can we instruct the visualisation tool to ignore certain hops in the path?
  - › Can I cancel a mesh and all future tests associated with it?
  - › How can I easily archive an ad-hoc test result that I run?
  - › Is there an open API to get the historic data out?

- » Could we use perfSONAR as a harness to measure the performance of different transfer tools, e.g., GridFTP, WDT, Aspera, ... perhaps between DTNs?
- » pScheduler has already been extended for DNS tests, tcpdump, ...
- » There are now multiple plugins and categories of plugins, including measurement archive category
- » Can extend pScheduler functionality to support additional archivers, e.g. Elastic, or OpenTSDB
- » In principle can schedule tests and post results to a database
- » Requires development effort on a transfer/GridFTP plugin
- » Recorded webinar on the topic on the perfSONAR YouTube channel

» Can we use perfSONAR to run TWAMP tests against our router infrastructure (largely Cisco and Juniper)?

- » Coming in 4.0.2/4.1
- » Available for testing soon
- » Developers would welcome people offering to test
- » Will be a 'standard' measurement

- » How might we integrate active perfSONAR measurements with other passive measurements, e.g. Netsight (SNMP) data on link utilisation?
- » Better analysis of variation in observed pS measurements
- » Possible work area for GN4-3?
- » Technically possible
- » Need to identify resource
- » Interesting questions over user interface; overlay of utilisation with perfSONAR view?
- » Overlay a trend / baseline?
- » Backend Esmond / Cassandra has a REST API
  - › <http://software.es.net/esmond/>

- » What are the recommendations for automated management of an organisation's perfSONAR infrastructure?
- » ESnet are using Ansible
- » GEANT operations are using puppet in general, and for perfSONAR
- » GEANT small node service will need a solution
- » The perfSONAR packages do not hinder any particular approach

- » Is it possible to use network topology information, and known network locations of perfSONAR nodes, to arrange that tests minimise potential duplication of measurements?
- » Not clear
- » There is an issue if a mesh is 'full' of tests, particularly where the same pairs of hosts appear
- » How to identify duplication?
- » Some smarts in principle possible via meshconfig UI

- » How do we best design perfSONAR tests to effectively show up soft problems in the network?
- » What general principles should we follow?
- » Comes with experience
- » General recommendation is for a perfSONAR node alongside your DTN, and at the campus edge, and build from there
- » Look at the data flows, and place measurement points along the path

- » When setting up pScheduler tests on demand, how should we best ensure appropriate authentication / authorisation for the test to be run?
- » Concern expressed on possible remote 'abuse' of the test infrastructure
- » The system is inherently open
- » Authorisation is IP-based through the limits file; can also limit on duration, throughput, ...
- » Perhaps pScheduler can alert on test requests above a certain threshold?
- » You can monitor traffic volume to/from a perfSONAR node via SNMP
- » Perhaps nodes on a common mesh can be trusted more?



- » What tools should we use to measure the load on a perfSONAR node? At least CPU and memory, perhaps more?
- » SNMP for network link utilisation
- » CPU/memory use
- » Disk space for archives
- » Free / available time in the test schedule
- » Ganglia?

- » What considerations are there for running a perfSONAR instance in a cloud / virtual environment, e.g., to measure performance to/from a commercial cloud provider, or a private OpenStack instance?
- » Docker work is ongoing
- » There is interest



**Tim Chown**

Network Development Manager

Jisc, UK

**tim.chown@jisc.ac.uk**

**jisc.ac.uk**