Imperial College London

GridPP
UK Computing for Particle Physics

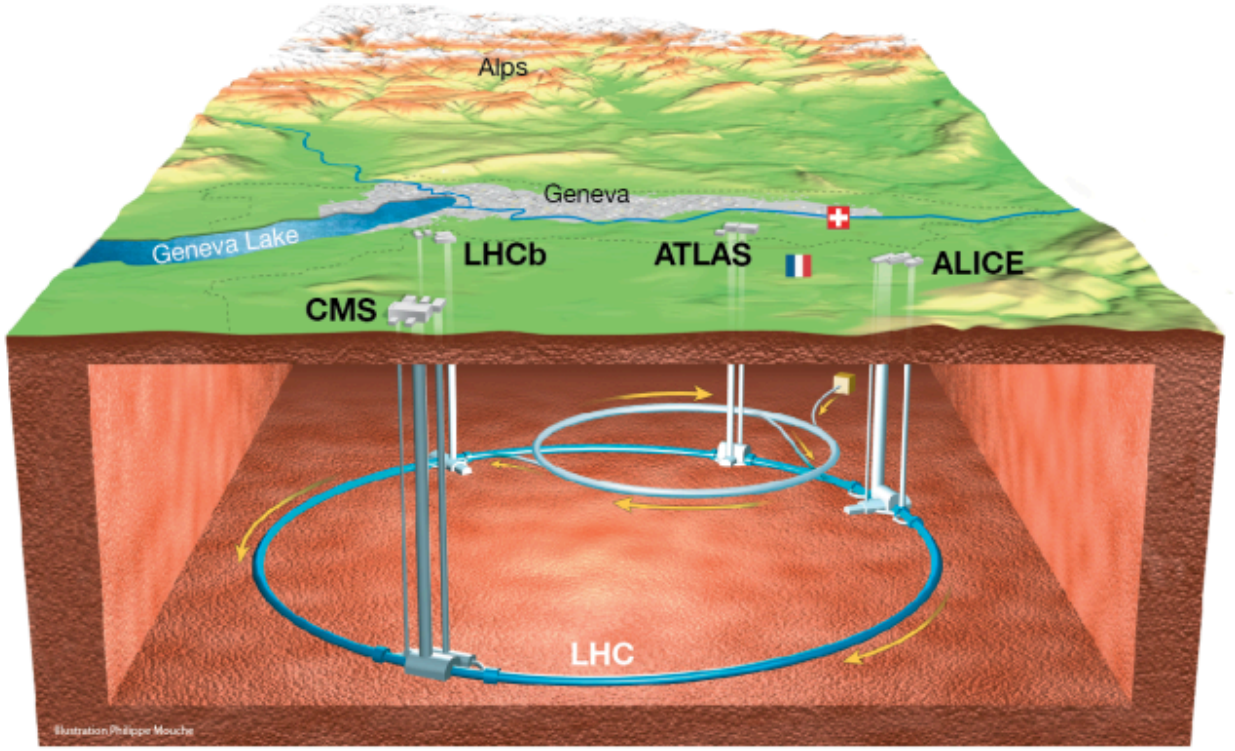
perfSONAR in the WLCG

Duncan Rand
Imperial College London

First European perfSONAR meeting, London, June 2019

The Large Hadron Collider (LHC)

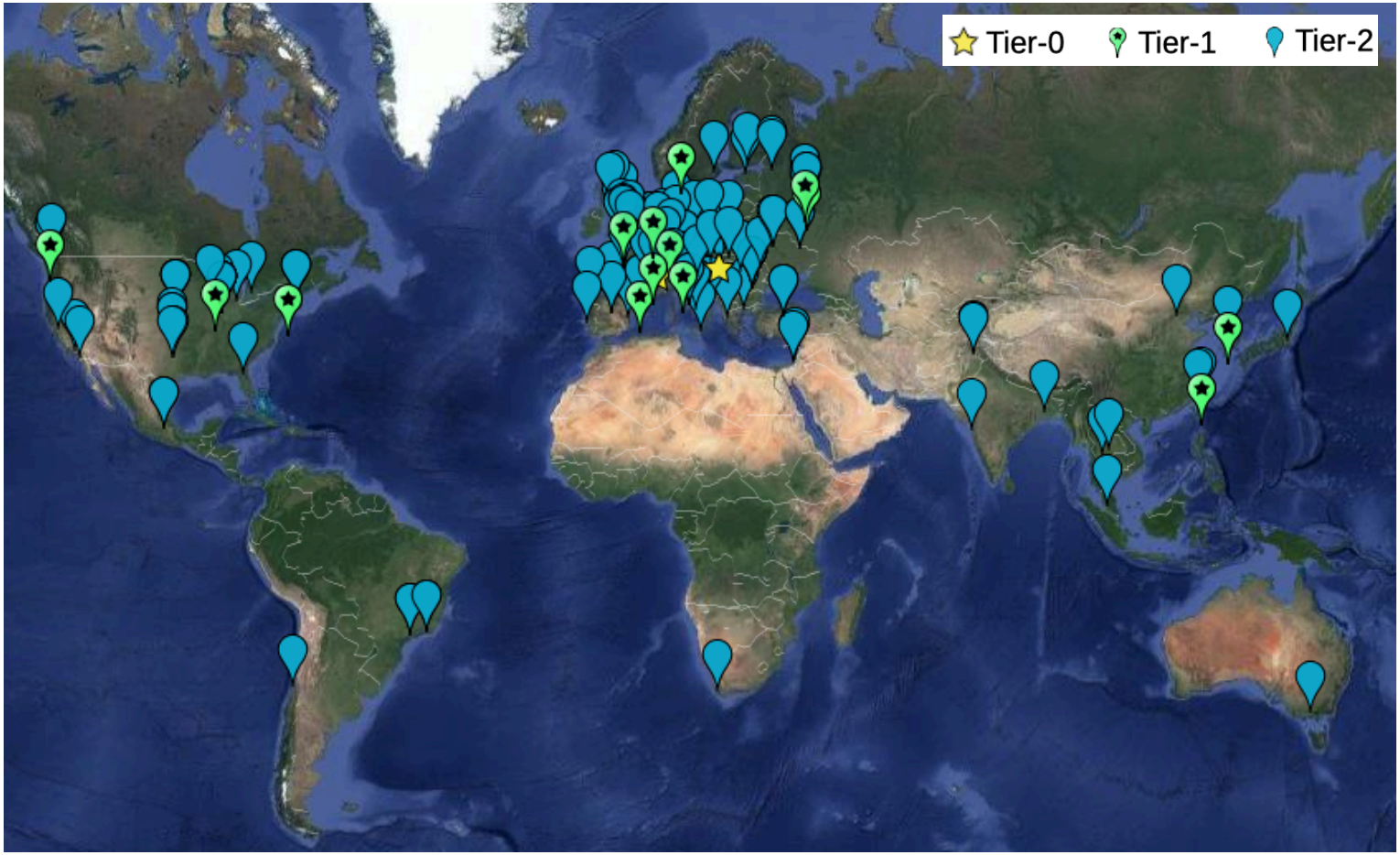
- The LHC is located at CERN on the Franco-Swiss border
- Proton proton and heavy ion collider with four main experiments
- Two general purpose: ATLAS and CMS
- Two specialist: LHCb and ALICE (heavy ions)
- During Run 1 at 8 TeV: found the Higgs particle in 2012
- Started Run 2 in 2015 at 13 TeV, just finished it on Monday 3rd December
- Computing for LHC experiments carried out by the Worldwide LHC Computing Grid (WLCG or 'the Grid')



Worldwide LHC Computing Grid (WLCG)

- The WLCG is a global collaboration of more than 170 computing centres in 42 countries.
- Its mission is to provide global computing resources to store, distribute and analyse the ~50-70 petabytes of data generated per year by the LHC experiments

- Sites hierarchically arranged
- Tier-0 at CERN (and Wigner in Hungary)
- 14 Tier-1s (mainly national laboratories)
- 149 Tier-2s (generally university physics laboratories)



WLCG Tiers Hierarchy

- Initial modelling of LHC computing requirements suggested a hierarchical tier-based data management and transfer model
- Data exported from Tier-0 at CERN to each Tier-1 and then on to Tier-2s
- However better than expected network bandwidth means that the LHC experiments have been able to relax this hierarchy
- Now data is transferred in an all-to-all mesh configuration
- Data often transferred across multiple domains
- e.g. a CMS transfer to Imperial College London might come from Fermilab near Chicago

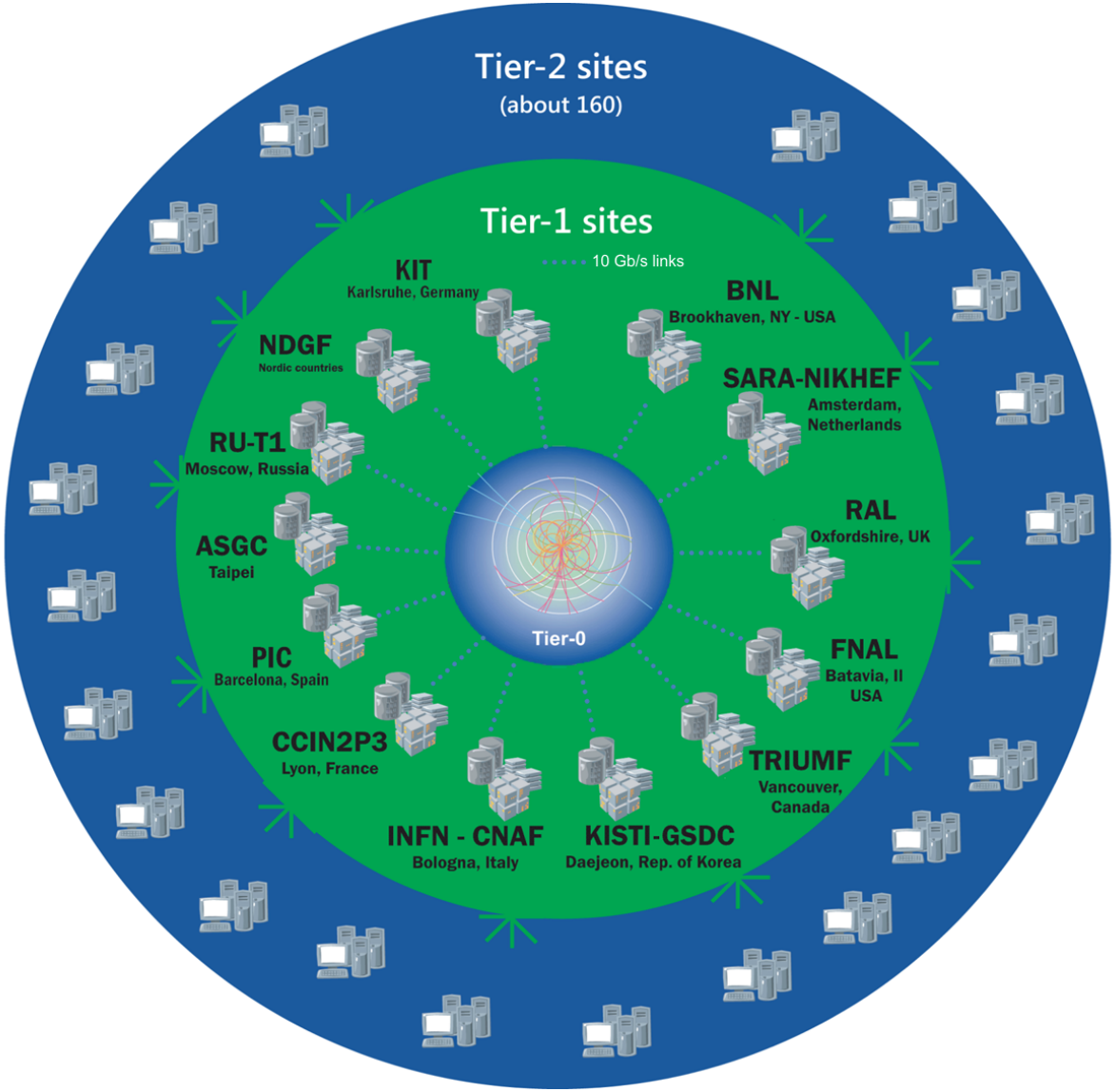


Image from 2014

WLCG Network Throughput WG

- Mandate
  - Ensure sites and experiments can better understand and fix networking issues
- Objectives
  - Oversight of the perfSONAR network infrastructure
  - Coordination of the WLCG network performance incidents
  - Detection and follow up on issues seen by the perfSONAR network
- List of Network Performance Incidents

- https://twiki.cern.ch/twiki/bin/view/LCG/NetworkTransferMetrics

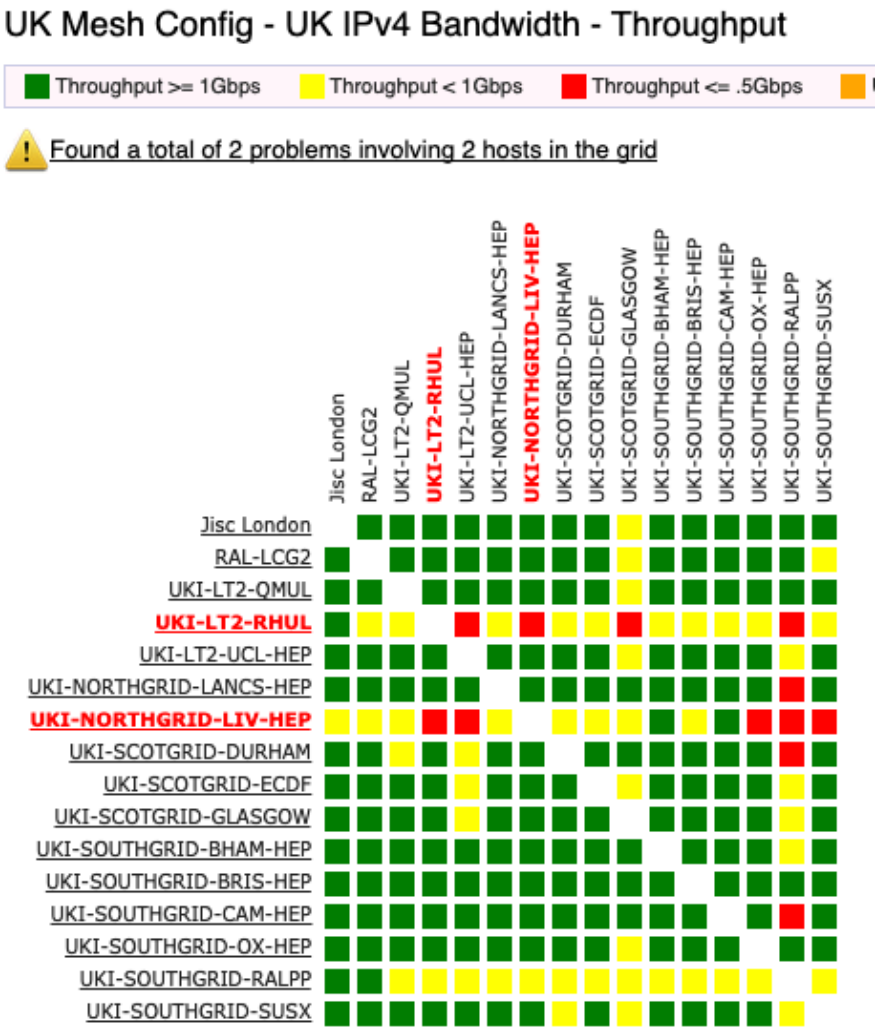
perfSONAR dashboards

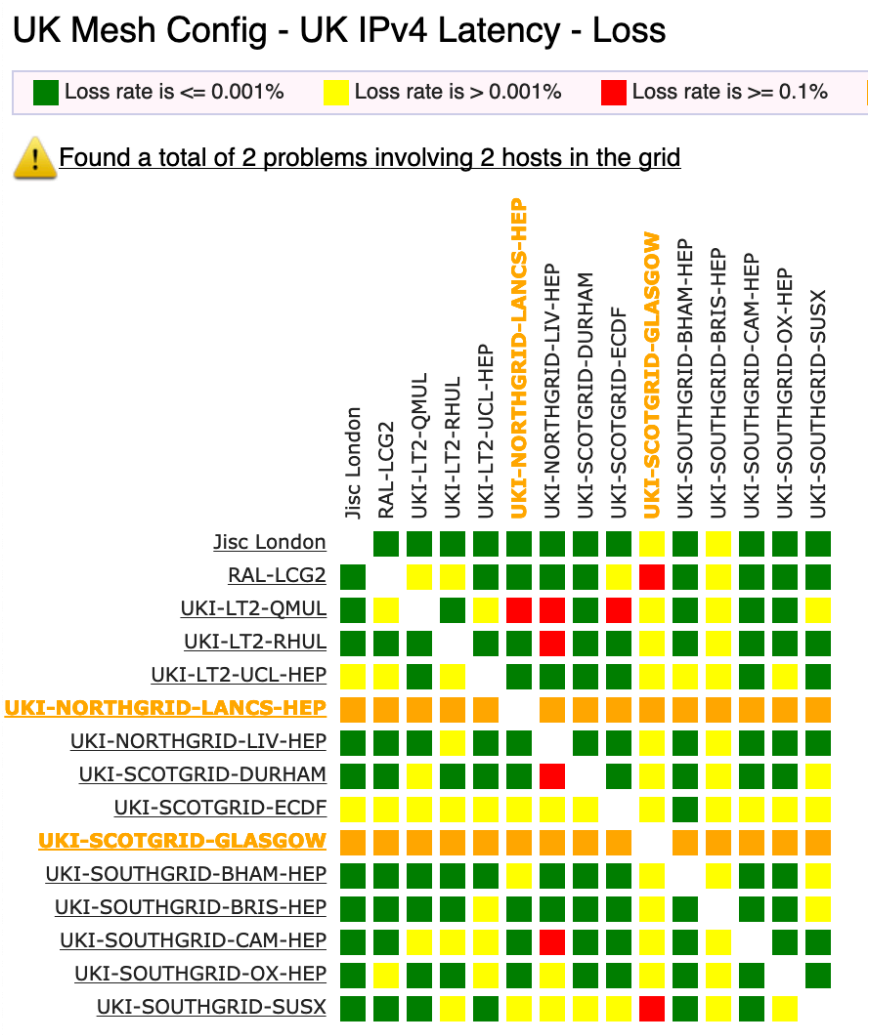
- Each WLCG site requested to deploy perfSONAR
- WLCG has meshes for a variety of groupings e.g. the LHCOPN, LHCONE, USCMS, USATLAS, CMS, ATLAS and LHCb
- UK and France also have their own national meshes
- WLCG IPv6 roll-out requires dual-stack perfSONAR
- Initial meshes were IPv4 only
- Once sites started to make their perfSONAR hosts dual-stack we implemented a dual-stack mesh
- Over time this grew to be too large
- Decision taken to make all other meshes dual-stack
- Traceroute and bandwidth meshes run both IPv4 and IPv6 tests
- Latency and loss tests load pS nodes more heavily and so only tested over either IPv4 or IPv6 depending on pair

Upgrading Issues

- Upgrading to CentOS7 and pS version 4.1.6
- Many hosts and meshes had fallen into disrepair
- Sustained effort to get UK and OPN hosts upgraded
- Submitted WLCG GGUS tickets
- Check_MK instance proved to be extremely useful discovering basic status of hosts (now also offers email notifications)
- Some sites awaiting new hardware etc
- UK and OPN meshes better but still many hosts requiring attention

perfSONAR dashboards

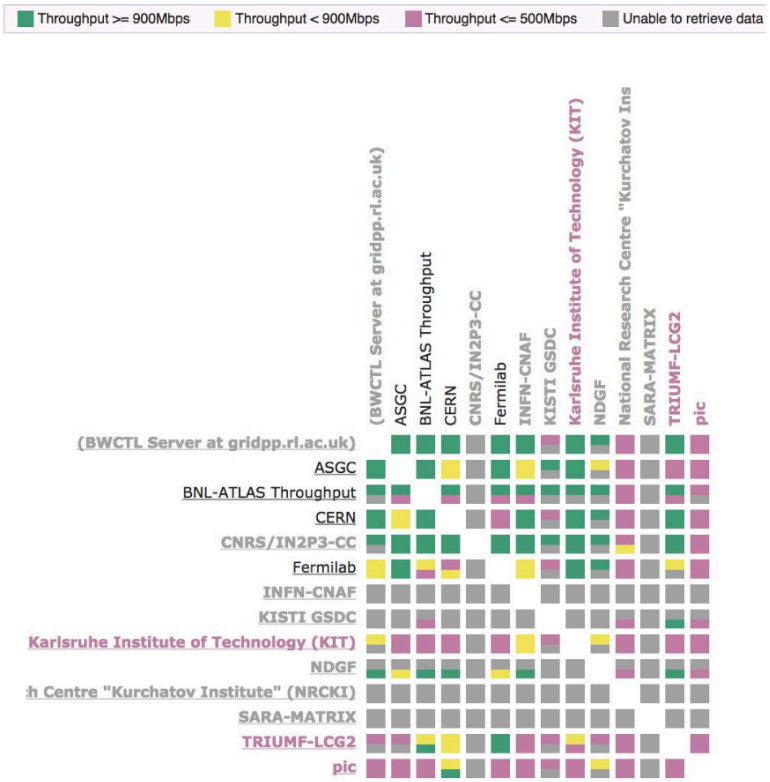




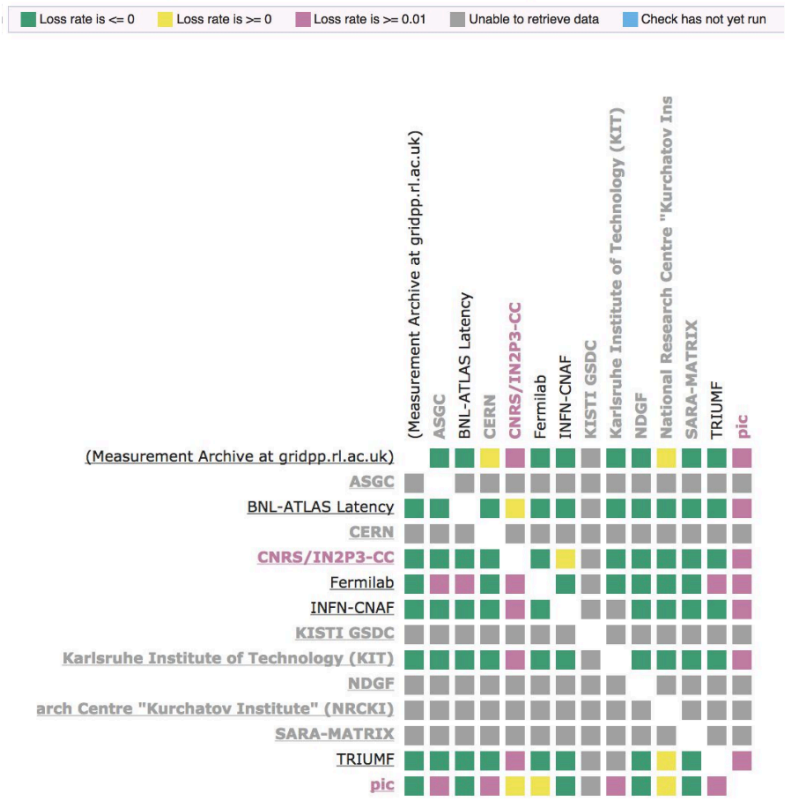
https://psmad.opensciencegrid.org/maddash-webui/index.cgi?dashboard=UK%20Mesh%20Config

LHCOPN - 29th October 2018

OPN Mesh Config - TCP BWCTL Test Between OPN Bandwidth Hosts

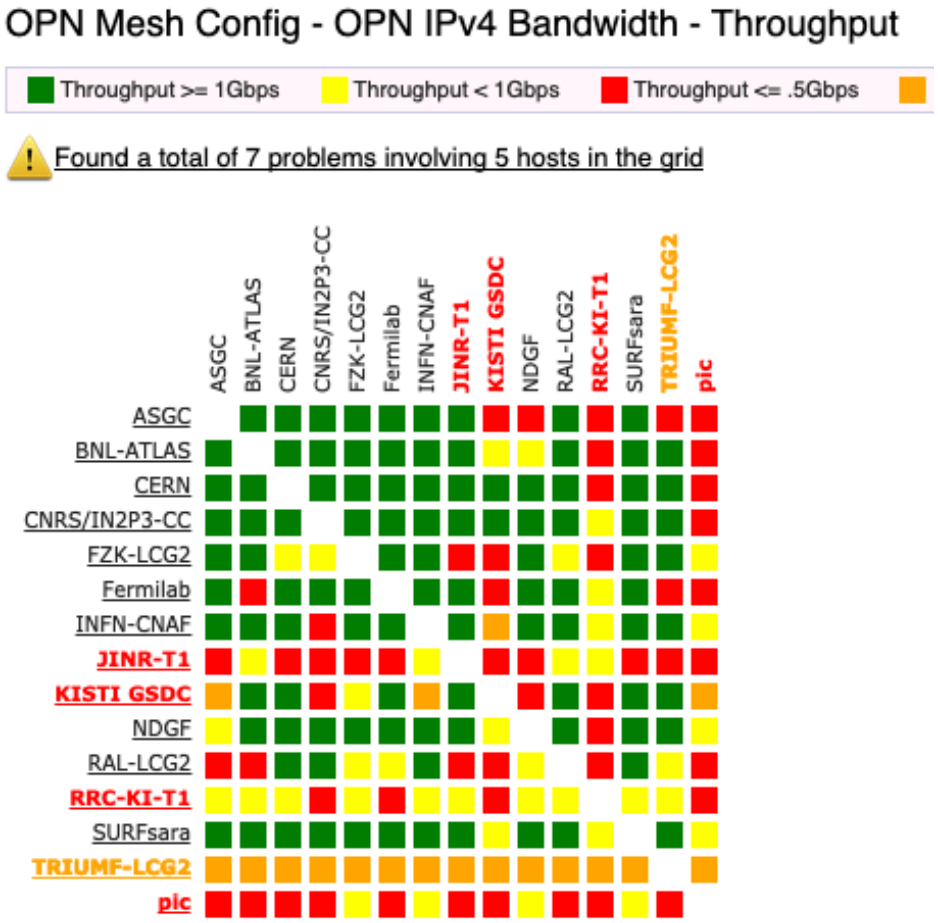


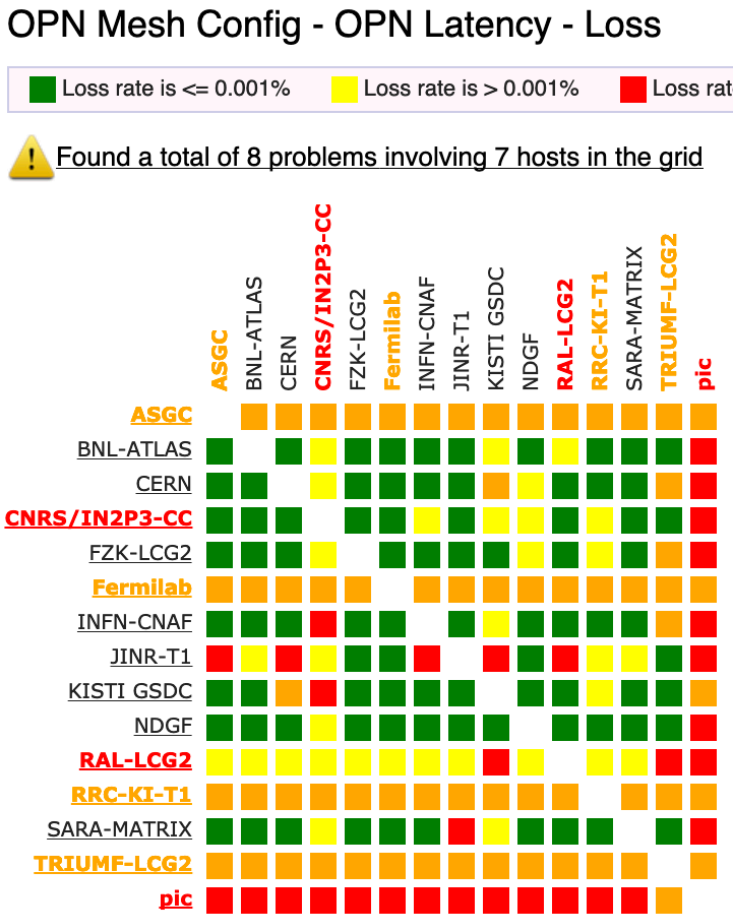
OPN Mesh Config - OWAMP Test Between OPN Latency Hosts



https://indico.cern.ch/event/772031/contributions/3360614/attachments/1855592/3047650/LHCOPN_LHCONE_perfSONAR_Update_2019spring.pdf

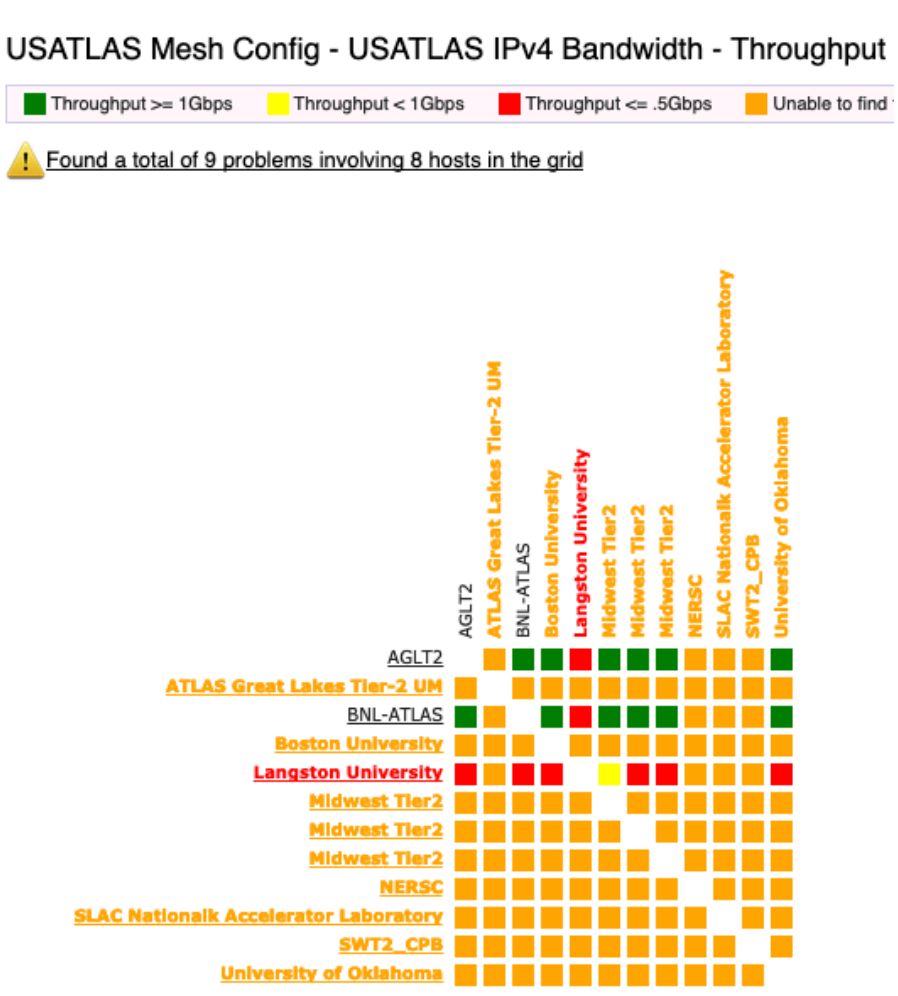
LHCOPN mesh in June 2019 looking better

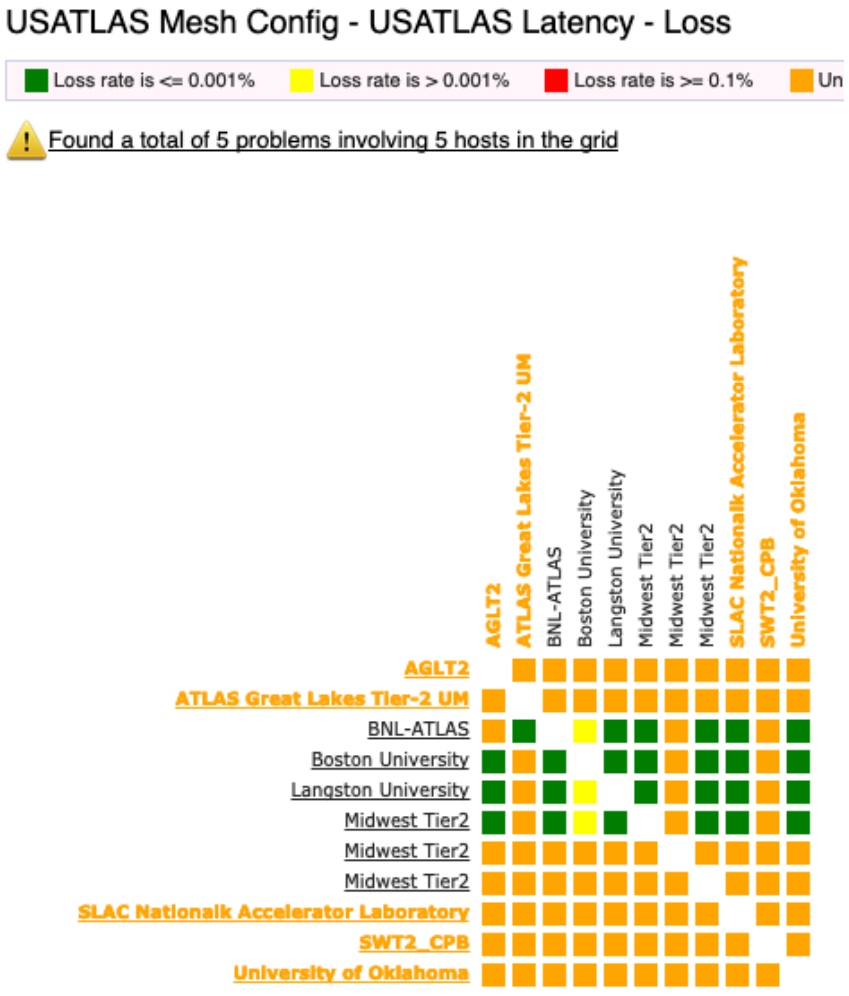




https://psmad.opensciencegrid.org/maddash-webui/index.cgi?dashboard=OPN%20Mesh%20Config

Others still need attention

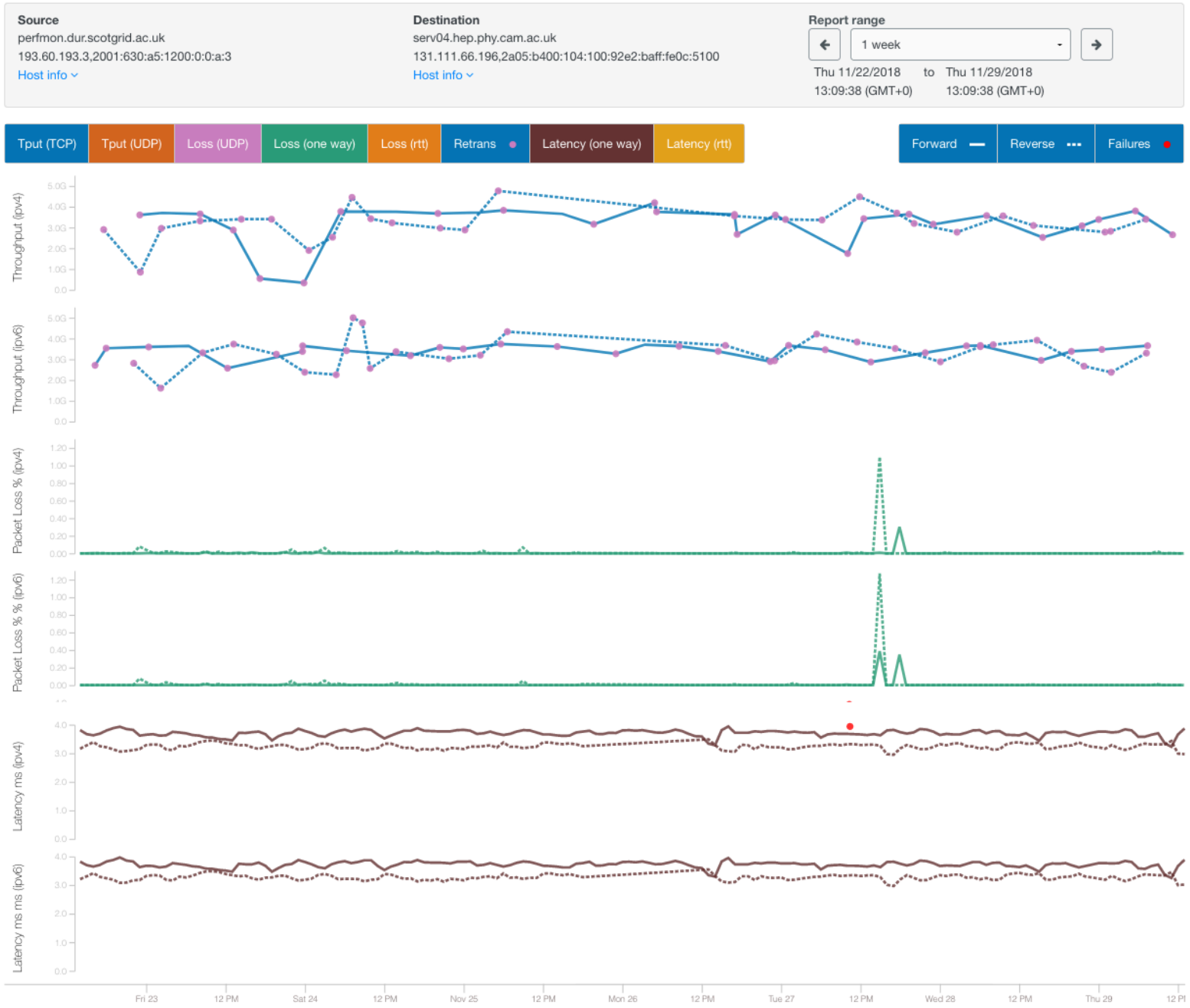




Roll out of IPv6

- WLCG IPv6 roll-out requires dual-stack perfSONAR along with dual-stack storage
- Each site received a GGUS ticket requesting status report
  - Check host status with Check_MK site
  - Request upgrade to CentOS7 if necessary
  - Some sites new to IPv6 unaware that ports need to be open over both IPv6 and IPv4
  - New Check_MK test added to check port 443 over IPv6
- Currently 163 of 293 (~55%) WLCG perfSONAR hosts are now reporting 'IPv6-enabled'
- More details available at https://twiki.cern.ch/twiki/bin/view/LCG/WlcgIpv6

Example perfSONAR results: Durham to Cambridge



IPv4 throughput

IPv6 throughput

IPv4 packet loss

IPv6 packet loss

IPv4 latency

IPv6 latency

UK RAL Tier-1 outbound packet loss over IPv6

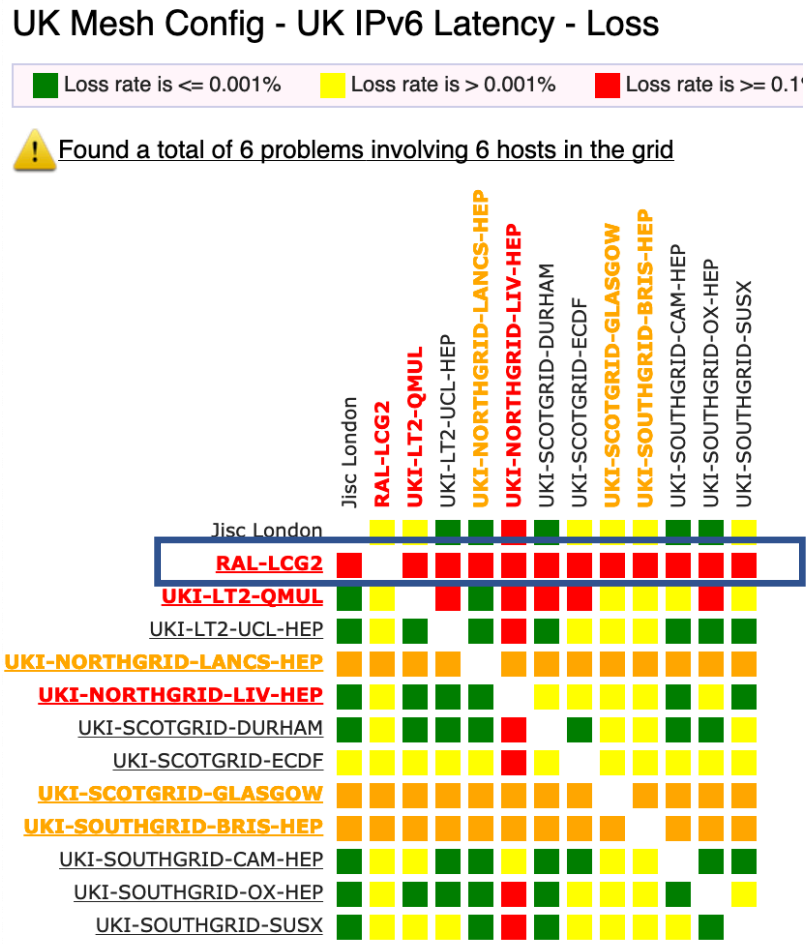
Tier-1 is still seeing unnacceptably high outbound packet loss over IPV6

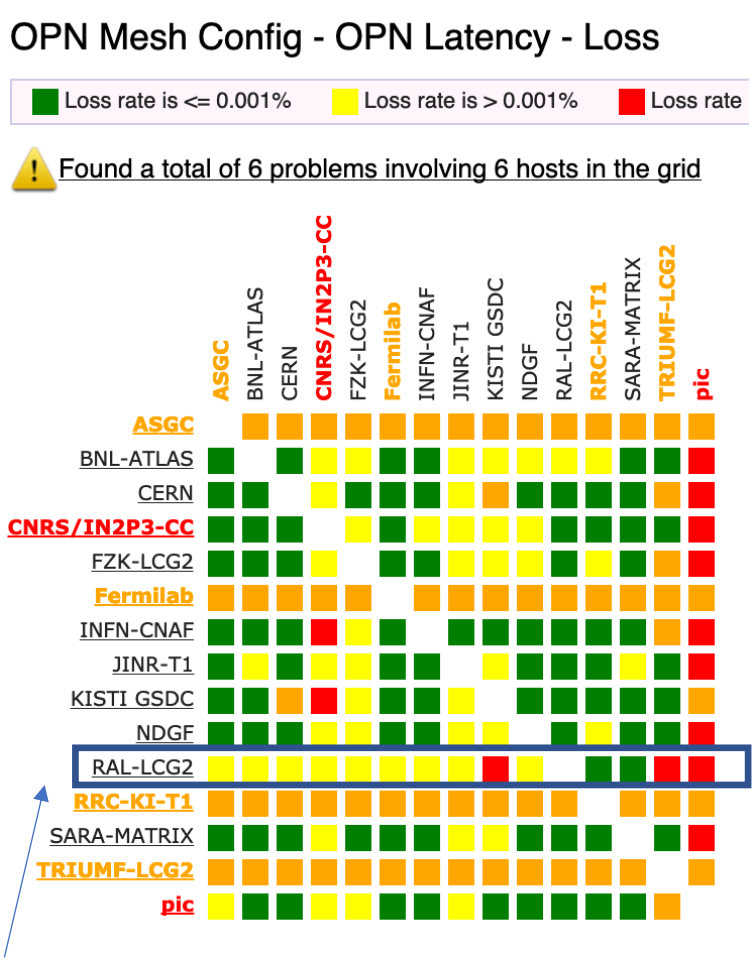
Science & Technology Facilities Council

UK Research and Innovation

https://indico.cern.ch/event/772031/contributions/3449119/attachments/1855535/3047636/STFC-LHCONE-STATUS-IPC-2019-06-04.pdf

UK RAL Tier-1 outbound packet loss over IPv6 (not visible over OPN)





Not visible over OPN

100G perfSONAR hosts

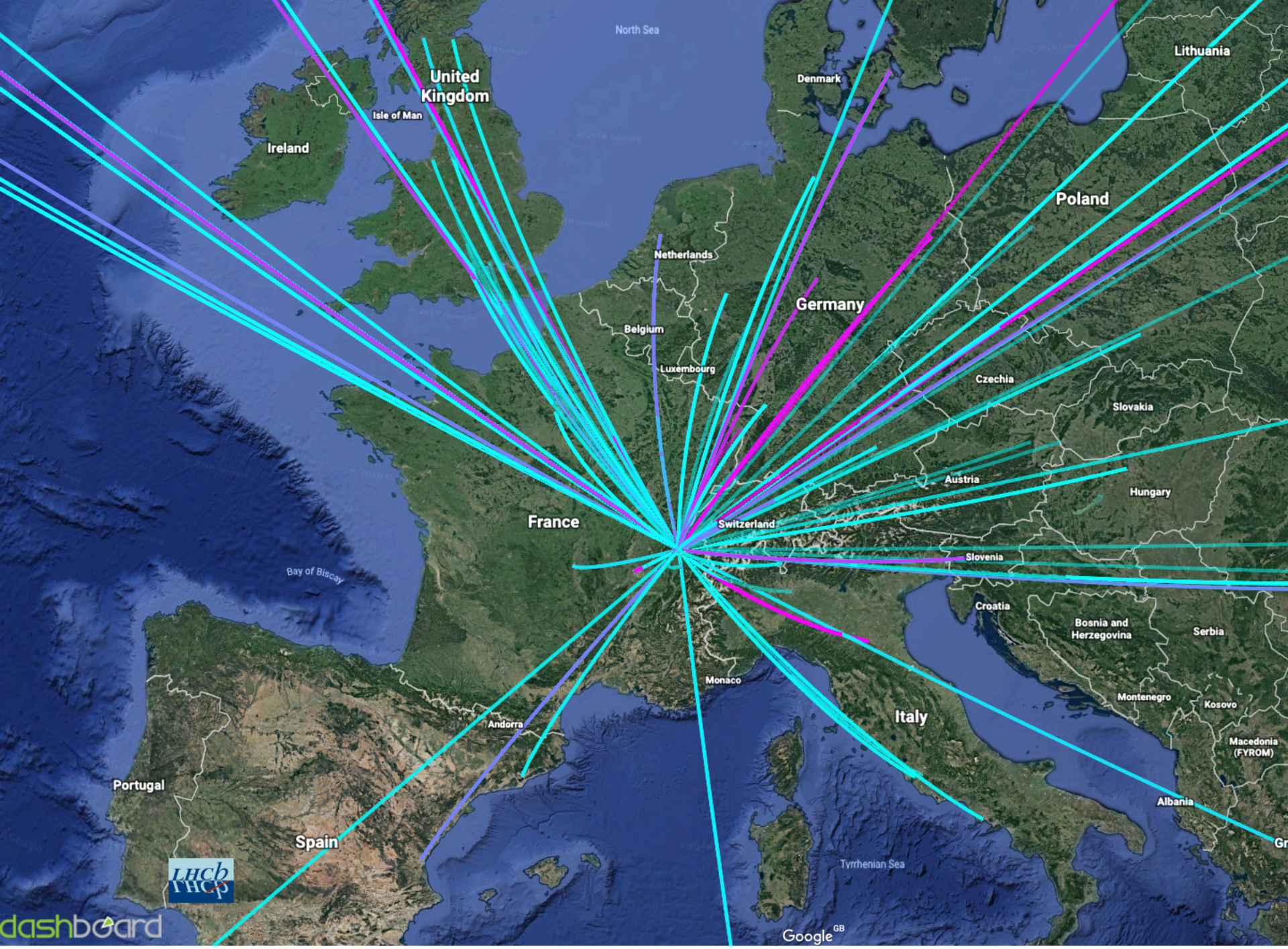
- A number of WLCG sites now have 100G wide area connections
- Some also installing 100G perfSONAR hosts
  - SurfSARA (Amsterdam)
  - CSCS (Lugano)

  - CERN (Geneva) now with 40 Gbps, 100 Gbps soon
  - KIT (Karlsruhr) testing 100 Gbps
  - BNL (80 Gbps)

- Possibility of setting up a 100G mesh
- More on 100G perfSONAR hosts in the next talk

Summary

- The WLCG is a highly distributed computing project
- Data travels often long distances over multiple domains
- The WLCG is using meshes of perfSONAR hosts to monitor network characteristics
- We are in the process of upgrading hosts to CentOS7 and pS version 4.1.6
- About half the WLCG perfSONAR hosts are now reporting ‘IPv6-enabled’
- Sites are starting to install 100G perfSONAR hosts


North Sea
United Kingdom
Isle of Man
Ireland
Denmark
Lithuania
Poland
Netherlands
Germany
Belgium
Luxembourg
Czechia
Slovakia
Austria
Hungary
France
Switzerland
Slovenia
Bay of Biscay
Croatia
Bosnia and Herzegovina
Serbia
Monaco
Montenegro
Kosovo
Italy
Andorra
Macedonia (FYROM)
Portugal
Albania
Spain
Tyrrhenian Sea
LHCb
dashboard
Google GB