

Storage SotU

Sam Skipsey (he/they)

GridPP52

Core Tier-2s - CEPH sweeping in from the North

- CEPH/?? Storage

- Glasgow (Ceph/xrdceph, cephfs/xrootd)
 - v ATLAS heavy + LHCb
 - ~10PB
- Lancaster (Cephfs/xrootd)
 - v ATLAS heavy
 - ~10PB
- Manchester
 - v ATLAS heavy (+LHCb?)
 - ~12PB

- Something else [“The London Set”]

- Imperial
 - dCache
 - Most of CMS
 - ~23PB (!)
- QMUL
 - Lustre / StoRM + xrootd (rdonly)
 - ATLAS heavy
 - ~13 PB

Non-Core Tier-2s

- Mostly cache / storageless
- Core site storage providers:
 - Lancs, QMUL [until d/t], RAL
 - A little of MAN + GLA
- VP has been v helpful
 - But we needed to put in work (see Rob C's talk)

- Storage-carrying exceptions
 - Durham – Cephfs/xrootd
 - Brunel – Cephfs/xrootd
 - Bristol – ~~HDFS~~Cephfs/xrootd
 - Birmingham – EOS (for ALICE)
 - RALPPD – dCache
 - Liverpool – dCache [from DPM]

Ceph consolidation

- Lots of Ceph, but not all Ceph are equal
- Newer ceph sites are
 - EL8 or 9 / Reef
- RAL T1 & GLA transitioning
 - EL7/Nautilus -> EL9/Reef
 - Requires multiple steps.



By the end of the year, we should aim to achieve a uniform Ceph/OS combination – and then establish “Golden Releases” for Tier-2s based on this going forwards.

Reef issues:

Reef is now (Aug 2024) acknowledged to have been a very buggy release of Ceph.

Vast number of bugs fixed in the last update

- resolved Durham issues!

This update not available for bare-metal EL8 hosts [support dropped “mid-release”]

- concern in community over implicit contract w/ Ceph devs

Q: what is the best feature / risk tradeoff?

Storage and NetZero

- Work in Progress
- Open questions:
 - Ceph (etc) are active storage solns – they draw more power when doing maintenance work, or changing config.
 - What is the “carbon cost” of rebalancing?
 - RAL + GLA have power monitoring on ceph clusters...
 - Is this better or worse with EC v replication (ignoring cost of media)
 - SSD v HDD embodied carbon issues.
 - Tape Farm considerations
 - Robot costs v cost per tape.
 - What does the v active use of tape farm by T1 mean for the “carbon benefits” of Tape storage v disk?