“UK Tier-2 Storage Evolution”

(A presentation on behalf of the GridPP Storage Group)

UK Tier-2s (Now)

SITE	Capacity(%TOT)	Solution
Manchester	18%	DPM
QMUL	17%	StoRM-Lustre
Glasgow	13%	DPM + Xrootd-Ceph
Imperial	11%	dCache
RAL PPD	9.8%	dCache
Lancaster	9.2%	DPM
Brunel	4.8%	DPM
Birmingham	4.6%	EOS + XCache
Liverpool	3.8%	DPM

Edinburgh	2.8%	DPM
RHUL	2.5%	DPM
Oxford	1.8%	DPM
Durham	0.5%	DPM
Bristol	0.4%	DPM-HDFS
Cambridge	0	XCache
Sheffield	0	storageless
Sussex	0	storageless
UCL	0	storageless

UK Context

GridPP

“Flat Cash” Staff Funding from STFC grants; allocation of site funding and consolidation.

“Funding situations are different in different jurisdictions, and strongly influence which models can work in a given jurisdiction.”

Need to also support other communities, with their own storage requirements:

IRIS UK : [DUNE, LSST, LZ, SKA, ...]
Some sites are tightly entangled with specific communities.

Many DPM sites, but wide size distribution

UK Tier-2s (Future)

SITE	Current	Future
Manchester	DPM	DPM ?
QMUL	StoRM	StoRM
Glasgow	DPM+Xrootd-Ceph	Xrootd-Ceph
Imperial	dCache	dCache
RAL PPD	dCache	dCache
Lancaster	DPM	DPM ?
Brunel	DPM	DPM
Birmingham	EOS + XCache	EOS + XCache
Liverpool	DPM	DPM

Changes highlighted

Edinburgh	DPM	DPM ?
RHUL	DPM	DPM / XCache?
Oxford	DPM	XCache
Durham	DPM	DPM + XCache
Bristol	DPM-HDFS	Xrootd-HDFS + XCache
Cambridge	XCache	XCache
Sheffield	storageless	XCache ?
Sussex	storageless	storageless
UCL	storageless	storageless

General Comments

Storage is inherently more conservative than Compute, as it encodes (important) State.

1. Even “scrapping storage” is hard [if users need to migrate off data]
2. Migrating infrastructure is extremely hard [esp. if some users can’t migrate off data]

non-Core sites will certainly move to “Storageless” [Cachey] solutions (1)

before core sites migrate to any “new/different” solutions (2)

We have several sites in case 1, and only one and a half sites in case 2.

UK Tier-2s - Concerns

Community support model for core software applications

- Requires more expertise of Tier-2 sysadmins, who are already heavily loaded.
- Much of this expertise is WLCG proprietary / not transferable
- Expertise retention in core developers and sys admins.

“Small” sites feedback loop [small workforce ⟳ remove services]

“Provider lock-in”/”High activation energy”: moving from 1 complex system to another, whilst in production, requires more effort + workforce than either the starting or end states.

(And hardware lock-in: buy hardware suited to particular implementations, limits movement to other solutions with different requirements.)

Job mix versus limited site functionality [cacheless or storageless sites might require radically different job types - this also places more pressure on the sites with storage, which will proportionately take the jobs not suitable for the cacheless/storageless ones]

Increased dependence on network for “storageless” solutions

Need solutions accessible outside of WLCG “bubble” for funding and other reasons.

Case 2: Glasgow

Began moving from DPM to Xrootd-on-CEPH ~2019 - complete ~ now, 2020

Triggers:

- Existing proof of concept & expertise - ECHO @ RAL
- Decline in central resource allocated to DPM development
- Significantly advanced resilience (RAIS, HA) features in Ceph wrt DPM
- Significantly advanced data placement (striping, auto optimise) features in Ceph wrt DPM

Why not DPM on Ceph/POSIX?:

- Overcomplicated [most of DPM features redundant wrt Ceph features]
- Lacking transparency [DPM namespace is decoupled from underlying namespace - “dark data” possible; cf transparent Xrootd namespace]

Why could we move?

- Already needed to move to new datacentre with different infrastructure on same timescale - much of the “disruption” was already going to happen.

Case 2: Birmingham

UNIVERSITY OF BIRMINGHAM

Storage at Birmingham - XCache

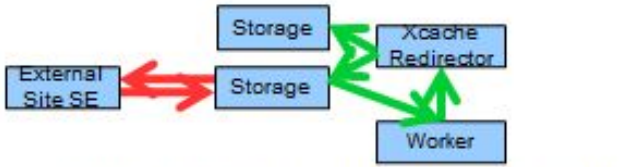
The Bham site originally provided 1/3 of storage it's 800TB (at the time, now 1.4PB) to ATLAS via DPM and 2/3 of storage to ALICE through plain xrootd

There was little prospect of getting enough storage to be really useful to ATLAS and available manpower was becoming an issue. To this end, XCache was setup to provide data access and Bham went diskless for ATLAS

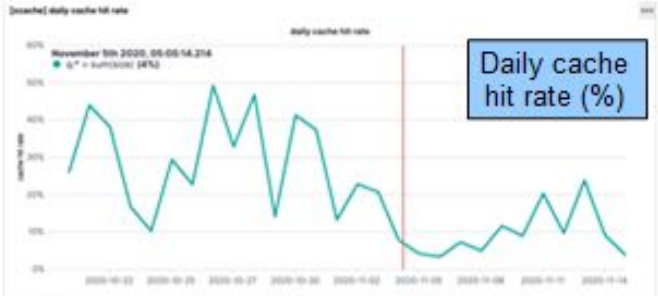
The current setup includes:

- Basic server (16C, 24GB) to act as redirector
- Two 40TB pool nodes
- ~400 cores for Atlas → 50 simultaneous 8-core jobs
- Pool nodes are connected at 10Gb/s, workers at 1Gb/s.

To ensure compliance with university IT policy, setup was done 'manually' by the site admin which was not too difficult. Some changes were required on the ATLAS side to change the path that jobs used to access data



root://xcache.bham.ac.uk/root://dpm.manc.ac.uk/..



Case 2: Birmingham

UNIVERSITY OF BIRMINGHAM

Storage at Birmingham - EOS

In order to continue to provide the vast majority of our disk space to ALICE, it was decided to switch from plain xrootd to EOS. This would give more (and easier) admin control and at the time, was considered to be used by ATLAS as well.

EOS boasts many desirable features:

- Low latency for operations due to in-memory namespace
- Unix user, GSI or Kerberos access
- User, Group and Project Quotas
- SAMBA and WebDAV access available

It also has a lot of administrative tools available:

- Comprehensive CLI tools to control all aspects of the system
- Automatic balancing between and within groups
- Draining of FSs and auto-repair from broken disks
- Highly configurable redundancy on a per file/directory basis

Initial setup wasn't too hard though at the time, documentation was a little lacking.

It was initially hoped to use the erasure encoding between servers to provide resilience to individual server failure. However, this could not be got to work and in the end ZFS was used for redundancy

After this however, the system has run smoothly (from Autumn 2018) with very few significant interventions or issues.

Other examples / Shorter term changes

Bristol: HDFS behind DPM site, low staff effort -> xrootd-on-HDFS

DOME DPM does not support HDFS; [xrootd-hdfs OSG-supported plugin]

(DPM namespace replicated in underlying HDFS so no data migration required.)

HDFS storage used by other parts of Group, so can be relied on.

Oxford: DPM DOME -> (test Xrootd proxy cache / XCache)

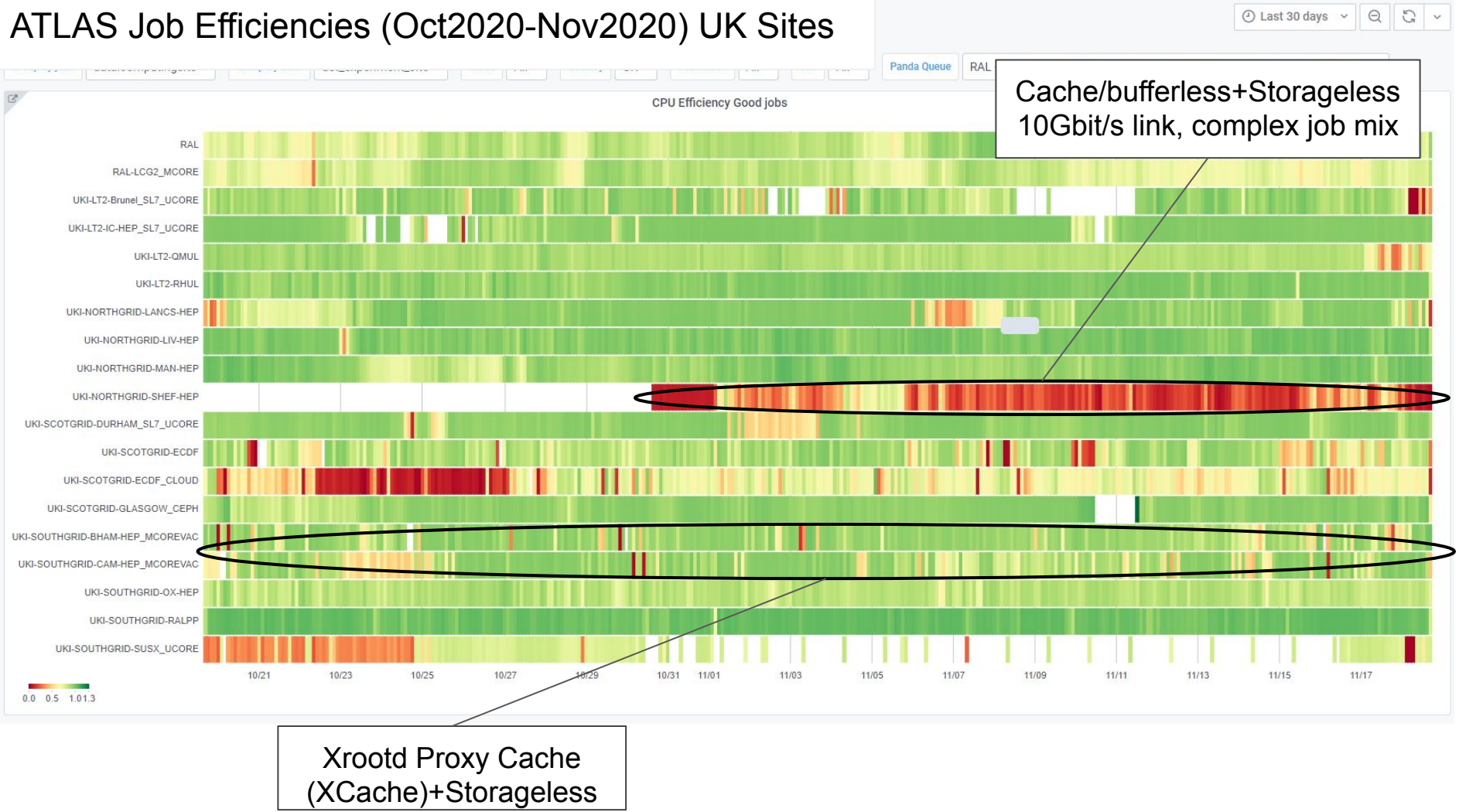
Staff effort at site, funding, expertise changes

Useful test instance for future specific advice to other “medium” sites.

Job mix from ATLAS workloads versus cache effect/efficiency.

(XCache monitoring hosted by Edinburgh, running for Birmingham atm)

ATLAS Job Efficiencies (Oct2020-Nov2020) UK Sites



Testing space of config for “storageless” sites

Efficiency of storageless sites is a multidimensional problem, with non-orthogonal axes.

Job mix: Simulation (almost no network requirement) -> Skimming / Derivation

- Job mix constraints for many sites reduces VO flexibility
- Can also result in “hard” job concentration.

Access model: staged versus streamed [or both]

Cache configuration / buffering: “caches” most useful for data read more than once ; but buffering via a cache can remove latency issues.

Plan at Oxford is for extensive, structured plan to explore interdependencies.

Scalability

CPU/Disk ratios are not a constant across UK sites, and the two are only somewhat correlated.

Caching/buffering models for sites with large CPU capacity are a particular concern for the testing work in the previous slide.

(If you assume as much as 2MB/s per job slot for IO heavy work, then that implies significant network requirements for an (unbuffered/cached) high CPU site.)
This also affects storage-holding sites which provide the sources for these sites [by adding to their total network load].
Esp. for ATLAS sites, where we need to pair [storage site] with [storageless site] this requires care.

Summary

Storage planning and evolution is inherently conservative, esp in production.

But funding and effort require some moves regardless within UK

“non-Core” Tier-2s -> (cache-only) supporting Tier-3 accessible storage

“Core” Tier-2s -> [most conservative, longest-term changes, HL-LHC?]
Some sites *considering* moves to new technologies.

Very long timescales: current solutions *need to work* for several years

Ongoing work for Tier-2 site optimisation for Cache config and topology.

Backup Slides

Case 2: Glasgow - Issues

Initial issues:

RAL deployment of Ceph is conservative; tracking Ceph releases versus community versions caused some desync

Xrootd-ceph builds are not automatic: needed to build our own xrootd releases.

Longer-term issues:

Xrootd-ceph plugin had almost no development support, and was several years behind xrootd mainline api functionality.

Xrootd documentation frequently assumes you have expert knowledge of source code, or, for some components, is written for OSG users [needs translation for other cases]

Case 2: Glasgow - Successes

Successes, As of (today):

Xrootd5/Ceph SE is primary production SE for ATLAS @ Glasgow

Ceph metrics, monitoring, automatic recovery, features significant improvement on DPM.

HTTP-TPC enabled @ Glasgow and passing tests [in production]

Xrootd-ceph plugin dev effort now healthy [effort from RAL, Glasgow - see Tom’s talk on ECHO later in this conference]

UK Tier-2s - Concerns [extra detail]

Community support model for core software applications

Requires more expertise of Tier-2 sysadmins, who are already heavily loaded.

Effort at many sites [see slide 3] is contended.

Much of this expertise is WLCG proprietary / not transferable

Current employees do not always stay within our community: learning systems which are not widely used outside of WLCG hinders their “employability”.
(Even within their current jobs, it is useful if a sysadmin can need to master a smaller number of solutions - they will often also be maintaining other Departmental IT systems - and if their experience can be transferable across their work, rather than only being narrowly applicable to a part of it.)

Expertise retention in core developers and sys admins.

Any suitable storage solution is a complex piece of software; development expertise takes time to build for such a product. Developers are not a fungible resource in these roles!

To an extent, this also applies to sys administration expertise.

UK Tier-2s - Concerns [extra detail]

“Small” sites feedback loop [small workforce ⟳ remove services]

Some sites worry that removing services also makes it harder to keep engaged effort at a high level [as those staff have less “contact points” with as many meetings etc]. This is ameliorated by increasing engagement in other areas, but we need to do that...

“Provider lock-in”/”High activation energy”: moving from 1 complex system to another, whilst in production, requires more effort + workforce than either the starting or end states.
(And hardware lock-in: buy hardware suited to particular implementations, limits movement to other solutions with different requirements.)

Most existing Grid Storage solutions conflate “access protocols” and “metadata + namespace” functionality.
(this is partly a consequence of the existence of SRM as a dominant negotiation protocol)
Moving to a different storage solution, without data loss, would therefore require migrating across the entire namespace to the new solution [and keeping the two synchronised during movement]; or maintaining two separate systems and thus running twice as much hardware.

“Dumb disk servers” bought for “classical” file-distribution based solutions are often underpowered in CPU terms for solutions like Ceph (which distributes more effort across its storage nodes). [Conversely, some solutions prefer smaller, “smart disk” solutions.] Since hardware lasts, ideally, for many years, planning architectural moves needs planning on the 3+ year scale.

UK Tier-2s - Concerns [extra detail]

Increased dependence on network for “storageless” solutions

Many GridPP sites are already the dominant users of network traffic to/from their host University.
Moving to storageless solutions increases network use for those sites - it is not clear if this is net saving; as University networking teams need to be on side [and network use competes with other legitimate users]

Additionally, moving to storageless solutions also increases network use for the remaining sites *with* storage: the storageless sites need to get their data from *somewhere*! This, again, needs to be understood as a thing that University networking teams need to be on side for.

[In 2020, with increased remote working for University employees, this has become more “visible” to many Universities.]

Need solutions accessible outside of WLCG “bubble” for funding and other reasons.

As the DOMA Access and TPC groups understand already [see Desirable traits for TPC protocols on https://twiki.cern.ch/twiki/bin/view/LCG/ThirdPartyCopy], many other user communities desire “standard” solutions for storage to work with us. (S3, Swift, non-X509 auth, etc etc)
Providing Tier-3 resources ; and making use of shared resources within Depts or Universities; also easier if we use as much “non-Grid proprietary” technology as possible. (Distributed filesystems, object stores, etc)