

GridPP DPM Experience

Alessandra Forti
DPM Community Workshop
27 February 2012

UK Status (1)

(repetita iuvant)

- DPM is the dominant storage in the UK
 - 1 T1 on Castor
 - 2 T2 on dcache
 - 3 T2 on Storm
 - 14 T2 on DPM**
- Everybody playing cool but most sites are extremely worried at the idea of changing storage
 - Even during the gap year it will not be easy if this happens
 - Lots of accumulated experience will be lost

EMI-*

[\[edit\]](#)

Sites running Atlas

Sites	CE status	WN status	Storage status	WN Tested	WN In Production	Tier	Problems
RAL-LCG2	Done	In progress	n/a	yes	yes	T1	
UKI-LT2-Brunel	Done	Done	Done	yes	yes	T3	
UKI-LT2-IC-HEP	Done	date	n/a	no	no	T3	Waiting for WN tarball
UKI-LT2-QMUL	Done	date	Done	no	no	T2	Waiting for WN tarball, and investigating move to rpms.
UKI-LT2-RHUL	Done	Done	Headnode done	yes	yes	T2	Will do rolling upgrade of SL6 poolnodes
UKI-LT2-UCL-HEP	In progress	In progress	30/11/12	no	no	T2	Manpower problems
UKI-NORTHGRID-LANCS-HEP	Done	In progress	Done	yes	yes	T2	Helping out with WN tarball. Testing on 1 cluster nodes are in production so far so good.
UKI-NORTHGRID-LIV-HEP	Done	Done	Done (EMI2/SL6)	yes	yes	T2	
UKI-NORTHGRID-MAN-HEP	Done	Done	Head node done (EMI2/SL6)	yes	yes	T2	
UKI-NORTHGRID-SHEF-HEP	Done (EMI1)	In progress	Done (EMI2,SL5)	no	no	T2	
UKI-SCOTGRID-DURHAM	date	date	Done	no	no	T3	Manpower problems
UKI-SCOTGRID-ECDF	Done	date	Done	no	no	T2	Waiting for SL6 WN tarball
UKI-SCOTGRID-GLASGOW	Done (EMI1)	In progress	Done (EMI1)	In progress (test node)	no	T2	
UKI-SOUTHGRID-BHAM-HEP	Done (UMD1)	Done (UMD1)	PNs: Done	6 months ago	no	T2	EMI2 head node didn't work, EMI2 WN needs retesting. Still affected by Bham power loss 2 weeks ago.
UKI-SOUTHGRID-CAM-HEP	Done (EMI2)	Done (EMI2)	Head node UMD1. Pool nodes done (EMI2)	yes	yes	T2	
UKI-SOUTHGRID-OX-HEP	Done	Done	Done (EMI1)	yes	yes	T2	
UKI-SOUTHGRID-RALPP	Done	date	n/a	no	no	T2	
UKI-SOUTHGRID-SUSX	Done	Done	Done	In progress	no	T2	Working on atlas integration

[\[edit\]](#)

Sites not running Atlas

Sites	CE date	WN date	Storage date	WN Tested	WN In Production	Tier	Problems
UKI-SOUTHGRID-BRIS-HEP	Done(EMI2/SL5)	Done(EMI2/SL5)	Done(EMI2/SL6)	yes	yes	not an atlas site	Only 10% LCG site support
EFDA-JET	Done	Done	Done	n/a	n/a	not an atlas site	

Upgrade from glite3.2 to EMI-2

- Upgrade generally smoothly

Most sites have upgraded in the last couple of months

- See State of the Nation page for DPM version numbers

As seen from previous page a mixture of
glite3.2/UMD/EMI-1/EMI-2 and SL5/SL6

- To DPM credit all working fine

- Couple of tools to make life easier during the upgrade

SQL script to drop request tables accelerates dramatically
the reload of the database

Yaim capability to create more than one pool (solution
suggested too)

- <https://svnweb.cern.ch/trac/lcgdm/ticket/442>

Problems reported

(Raul Lopes and fellow DPMers)

- Drain

If anything it's worst than it used to in 1.8.3. Multi-threaded just doesn't work. Still leaving lots of files in an "undecided" state.

- https://ggus.eu/ws/ticket_info.php?ticket=88633

Draining process stops many times and needs to be restarted

There is always a 1% left undrained needs to be cleaned up

- Checksum

Transfers of big CMS files to Brunel have become a serious issue. I've requested computation simultaneous with the transfer

- https://ggus.eu/ws/ticket_info.php?ticket=88431

- Missing files

More than once I have notice the case of files that the SE reports as available, I can see them in the pool node, but both lcg-cp and dpm-replicate fail to copy them.

- https://ggus.eu/ws/ticket_info.php?ticket=88634

Problems & Features (2)

- DPM service still dying if a pool is momentarily unavailable.
- xroot still segfaulting but not at all sites

It's a known issue on SL6

- Number of rfio threads on a pool node can grow out of hand

Connected to the round robin way to fill the nodes

- Weight eased handling of this problem but didn't solve it
- Common space reporting

Each VO reports the same space

- Not possible to separate by VO

When more space unavailable space reported is still negative

- NFS4 interface would be useful (there alternatives though)
- User quotas would be useful especially if SRM goes.
- DMLite plus Ceph/Hadoop/.... (Brunel offered testing Ceph)

Community Effort



GridPP Storage Group

- Very active community

DPM biggest slice of effort

- Weekly meetings
- Community effort when sites have a problem
- Testing new features a number of sites are joining the testing

Glasgow, Edinburgh, Liverpool, Brunel

GridPP DPM tools supported by Sam (See his talk)

Other tools still hosted in the repository requested long time ago by WLCG/Hepix

- <http://www.sysadmin.hep.ac.uk/svn/fabric-management/dpm/>
- Resurrected few weeks ago for UK sites sharing

Still mostly on best effort

- If a site needs it they'll do it

Conclusions

- DPM is important in the UK

Moving 14 sites to another storage system is no minor task

Sites move to EMI-* (mostly EMI-2) should grant some time to decide what to do (1 year? 2 years? Forever?)

Community effort is fine but is on best effort and cannot solve all the problems.

- As Ronald Starink pointed out when queried about a new version of the `dpm_log_parser`

I am always in favour of sharing scripts. [...] I can have a look at that, but I won't make any promises about a time scale. But if you're in a hurry or like to do it yourself, feel free.

- And others are in his same predicament: it's on the TODO list but....

Dedicated effort is important to keep everything together!