



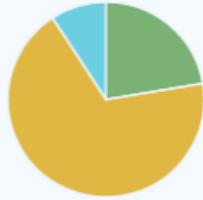
# DPM Upgrade TF

# DPM Upgrade TF

- <https://twiki.cern.ch/twiki/bin/view/LCG/DPMupgrade>
- WLCG TF Triggered by Storage Resource Reporting
  - Reason for assembling the TF
    - Introduced at the WLCG ops meeting Sept 2018
  - Still a good reason to upgrade
- Now the legacy stack is end-of-life
  - Upgrades are a major operational concern
- Objective:
  - By summer 2019 80% of DPM storage (in terms of capacity) to be upgraded and reconfigured
  - By the end of 2019 80% of DPM sites to be upgraded and reconfigured

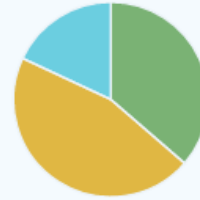
## ▼ Pie Charts

### Used



		used	percentage
ALICE	66.0 PB	22%	
ATLAS	202 PB	68%	
LHCb	27.1 PB	9%	

### Free



		free	percentage
ALICE	17.6 PB	36%	
ATLAS	22.1 PB	46%	
LHCb	8.74 PB	18%	

WLCG Storage Accountin... ☆ ↻ 🖨 🕒 Last 30 days 🔍 ↺ ▼

VO ALICE + ATLAS + LHCb ▼ Tier All ▼ Country All ▼ Federation All ▼ Medium Disk ▼  
Site All ▼ Service All ▼ Area All ▼ Group By vo ▼ Binning 1d ▼

<http://cern.ch/go/6Tvf>

## ▼ Table View

### Summary

vo ▲	Used	Free	Total
ALICE	66.0 PB	17.6 PB	83.6 PB
ATLAS	202.4 PB	22.1 PB	224.5 PB
LHCb	27.1 PB	8.7 PB	35.9 PB

### Usage

	usage	percentage
Used	296 PB	86%
Free	48.4 PB	14%



# TF Progress

- Phase I – 9 sites- completed (mostly)
- Now in phase II – 8 sites
- Along the way, we've...
  - Found and fixed issues with the new stack and the docs
    - (see later)
  - Explored DPM parameter space...
    - DPM installations dating back years
  - Seen some sites upgrade with very few (reported) problems
    - ...but they are therefore less visible
  - Begun to synchronise with EGI to cover all user communities

## Detected problems

- PRAGUELCG2
  - forgot to [tune open file handles limit on dedicated DB machine](#) (during DPM upgrade we moved DB from SLC6 to CC7)
  - **DONE** 1.12 - dmlite-mysql-dirspaces.py doesn't fix data assigned to wrong (empty) spacetoken which can/will cause problems for SRM transfers
  - **⚠** Diverging directory space usage for concurrent [GridFTP](#) uploads [LCGDM-2730](#)
  - Slowly diverging space counters (size(parentdir) != sum(size(childdir))) e.g. for ATLASDATADISK using [pydmlite API](#) or [direct DB query](#)
  - **⚠** [GridFTP](#) transfers confirmed OK to client before processed by DPM (e.g. DB update for filesize, replica status, ...)
    - [GridFTP](#) dsi plugin bug and ugly workaround in DPM doesn't always work
    - lead to lost files at our site when replica status was not updated by DPM (reason not fully understood), but for FTS/Rucio this was successful transfer
  - **DONE** 1.11 + xrootd 4.9.x - xrootd checksums not implemented [LCGDM-2726](#)
  - **DONE** 1.12 - Unable to correctly disable gridftp redirection with [puppet](#)
  - **DONE** 1.12 - Problems with default non-optimal configuration options [LCGDM-2743](#), [LCGDM-2745](#), [GGUS:139803](#)
  - **CLOSED** - SRM upload/deletion doesn't immediately update directory size [LCGDM-2731](#) - db updated => space accounting should work find
  - **CLOSED** - Removing directory with SRM not seen by DAVS [LCGDM-2732](#) - can't be fixed (don't mix SRM with DOME protocols)
  - **DONE** 1.12 - Directory size not correctly updated with gfal-rename [LCGDM-2733](#)
  - **DONE** 1.12 - Unable to disable bad filesystem with DOME [LCGDM-2740](#)
  - SSR script robustness [LCGDM-2714](#), [LCGDM-2744](#) (can be mitigated by separate quotatoken for storagesummary.json)
  - **DONE** 1.12 - DPM hammer itself with checksum requests [LCGDM-2747](#)
  - **DONE** 1.13 - Checksum issues (deadlock, popen3 blocking read) [LCGDM-2791](#)
  - **CLOSED** - I'm not really happy with given answer - Confusing usage of the head.db.poolasz configuration option [LCGDM-2749](#)
  - **DONE** 1.12 - Database connections are not reused - individual connection for each query [LCGDM-2754](#)
  - **DONE** 1.13 Disabled DPM disknodes can be selected for [GridFTP](#) transfers [LCGDM-2748](#)
  - Complain when setting quotatokens at a level deeper than the one where space is calculated [LCGDM-2727](#)
  - **DONE** 1.13 Problematic graceful restart of httpd [LCGDM-2699](#), [LCGDM-2707](#) + "DPM http graceful restart" discussion in dpm-upgrade mailing list
    - problems caused by sub-optimal default Apache configuration ([ServerStart](#) == [ServerLimit](#) can by design cause troubles during graceful restart)
      - processes not terminated till at least one thread deals with existing transfer
      - some transfers can hang forever [LCGDB-2787](#) (rare but happens, probably fixed in lcgdm-dav 0.23)

# Phase I

Site	Total Size PB (BDII)
Brunel	1.8
Napoli	2.3
Recas-Napoli	0.4
Manchester	6.8
Prague	5.2
LAL	1.5
LLR	1.7
IRFU	4.6
Lancaster	4.3
<b>Total</b>	<b>28.6</b>

# Phase II

Site	Total Size PB (BDII)
CPPM	2.6
Auvergrid	0.02
Cosenza	0.5
Budapest	1.1
Glasgow	3.8
LPC	1.6
LHEP	1.1
Oxford	1
<b>Total</b>	<b>11.7</b>



# Phase II

Site	Total Size PB (BDII)
CPPM	2.6
Auvergrid	0.02
Cosenza	0.5
Budapest	1.1
Glasgow	3.8
LPC	1.6
LHEP	1.1
Oxford	1
<b>Tokyo</b>	<b>10.6</b>
<b>Total</b>	<b>22.3</b>

# Progress & objectives

	Capacity	Number
Phase I or II	50PB	18
Close collaborators	70PB	45
All installations	97PB	102

- Reminder of Objective:
  - By summer 2019 80% of DPM storage (in terms of capacity) to be upgraded and reconfigured
    - This means 77PB
  - By the end of 2019 80% of DPM sites to be upgraded and reconfigured
    - This means 80 sites

# DOME/upgrade related issues

- **gridftp redirection**
  - Old client tunnelling
  - Issue with stale gridftp processes?
  - **Success before put-done**
- **Disabled filesystems**
  - deletions and drains – fix 1.12
  - gridftp redirection – fix 1.13
- **Checksums**
  - Performance - fix 1.11
  - deadlocks etc - fix1.13
- **Divergences between dir counters and their contents**
  - LCGDM-2730 concurrent gsiftp uploads
  - (size(parentdir) != sum(size(childdir)))
- **dmlite-mysql-dirspaces issues**
- **SRR publisher could be improved**
- **xrootd/dome crashes**
  - Stability/performance things (brunel) - fix 1.12/xrd4.9.0
  - Davix – fix 0.7.3
- **Various certificate problems**
  - Dome has better internal authentication
- **Puppet special password chars**

# Other issues

- Existing issues
  - Apache graceful restart
  - GridFTP
    - ipv4 client + ipv6 server
      - fix under validation
- Old/New interactions
  - Divergences between SRM and QT
  - Divergence SRM write and davs delete divergence – fix 1.13
  - xroot stat / SRM write / xroot read + small file

# What's the destination?

- Gridftp redirection
  - In principle this is not coupled to the upgrade
    - It's coupled to SRM retirement
  - In practise...
    - “Mixed mode” numbers divergence could cause sites move experiments off SRM
    - *“never go to sea with two chronometers, take one or three”*
  - What tools are available for detecting older client populations?
    - Ideally before upgrading...
- “Mixed mode” is the worst of all worlds
  - You are at maximum complexity
  - We need to clear the way to SRM-free operation for WLCG
- Role of the experiments?
  - And their tests



# Discussion - how to progress

- What are the current impediments?
  - Particular bugs?
  - General complexity?
  - Scheduling?
    - Synced with hardware or OS upgrades?
- Ticket everyone in the “third wave”
  - Atlas tickets
  - Natural triggers through other tickets
- What role for EGI?
  - And their tests
- What should we do about
  - Old applications requiring SRM
  - Old gridftp clients
- We need to take care scheduling this over the summer
  - DPM team are happy to discuss dates