

# Storage Space Accounting

*...trying to make a summary of pre-GDB discussion...  
and bootstrap today's discussion.*

*N.b.: similar discussions ongoing since many years, don't shoot the messenger!*

*Ale Di Girolamo CERN IT/LCG  
on behalf of WLCG Ops Coord*

# Status

All OK! Experiments work! WLCG is effective! Everyone Happy! Hurrah!!

- ALICE:
  - not too aware of the internal details, basically they ask their site to run some plugin to calculate the space
- ATLAS:
  - frequent (30mins/1h) check on each endpoint (each spacetoken) to get Total/Available. Asking/getting (monthly) full storage dumps
- CMS:
  - space Phedex aware was already under control, the rest of the space handled by local contacts at sites. Now developing and deploying this new “StorageSpace monitoring” relying on full storage dumps (each file, size, etc, monthly frequency)
- LHCb:
  - Similar to ATLAS: frequent check on Total/Available.

# ... is it really all OK?

..... yes but:

- Experiments are really (almost) all full
- Very difficult (as of today almost impossible) to have a WLCG storage overview of the storage capacity total/used.
- In the future we will be (more) limitation on the storage resource wrt what Physics folks would like:
  - Fighting for fundings
  - Our resource usage reports need to go deeper and deeper (e.g. why 5% there and 7% there?)
- New SRM-less “world” in front of us, it is approaching! (rapidly?)
  - How do we handle these storage resources?
- How do we account storage provided by e.g. Cloud Providers?
  - We can't really impose anything here, flexibility is the key!

# Do we need a revolution?

- Most probably not. But we need r(eal) evolution!
  - Experiments are doing the same thing several times, asking sites and storage developers similar but not exactly identical things.
    - This can (and should) be improved
- What I think we need (to start with):
  - Experiments could agree on the minimal set of info needed
    - is Total/Available per quota node enough? I think so...
  - Experiments could agree on formats in which they would like to get and store the storage accounting (monitoring) information
  - Experiments could agree on common tools that they use to get this storage information
    - As it was successfully done for lcg-stmd now evolved into gfal-xattr
  - WLCG Storage providers could agree in common ways to get the total/available numbers in the SRM-less (no spacetoken) era.
- Setting up an overview portal is the last (but important) step
  - not without nasty difficulties: e.g. Naming Convention

# More use cases: not really for Storage accounting

- Storage dump
  - Full details of each file, size, access time
  - Useful for both ATLAS and CMS
- Format can be agreed
  - Simplification for the sites and the storage providers
  - Can be an extension of the “simplified” basic format that would be nice to have to get the “total/available” info

# Practical examples of basic info needed: SRM-less and gfal2

```
-bash-4.1$ gfal-xattr srm://atlassrm-fzk.gridka.de/srm/managerv2?SFN=/ "spacetime.description?ATLASSCRATCHDISK"
[{"spacetime":"25353616","owner":"VoGroup=/atlas VoRole=*","totalsize":1000000000000000,"guaranteedsize":1000000000000000,"unuse
dsize":40101043548662,"lifetimeassigned":-1,"lifetimeleft":-1,"retentionpolicy":"REPLICA","accesslatency":"ONLINE"}]
-bash-4.1$
```

- On top the “standard” gfal2 query
  - This is OK!
  - (should or plan to have it) work for SRM, gridftp, xrootd and http (but w spacetimes)
- On the right
  - the proposed structure that an SRM-less site (w/o spacetimes) could expose.
  - Up to the Storage Developers to extend/modify etc... experiments can adapt, but we need them!

```
{
  "ATLASDATADISK":
  {"status":"online/offline",
   "status_message": "The report can not be created because ...",
   "list_of_paths":
   ["/castor/ads.rl.ac.uk/prod/atlas/stripInput/atlasdatadisk/",
    "/castor/ads.rl.ac.uk/prod/atlas/stripInput/atlasgroupdisk/],
   "total_space": 5000000000,
   "used_space": 2000000000,
   "num_files": 123456,
   "time_stamp": 1447936989},
  "ATLASSCRATCHDISK":
  {same}
  "ATLASLOCALGROUPDISK":
  {same}
}
```

Remember: frequency not KHz (<<1Hz!), precisions +/- few GB

# Conclusion

- ... not really any conclusion:
  - there could be real work to do!
- Experiments do have working systems
  - Whenever sites do not publish correct numbers (e.g. through gfal/lcg) they get GGUS tickets
- Storage providers have their own ways/ideas.
- Changing already working systems could be not effective:
  - But would be very useful if the evolution is done together