

# The Grid Storage System Deployment Working Group: a status report



7<sup>th</sup> February 2007

Flavia Donno  
IT/GD, CERN

# Grid Storage System Deployment (GSSD)



- <https://twiki.cern.ch/twiki/bin/view/LCG/GSSD>
- Mailing list: [storage-class-wg@cern.ch](mailto:storage-class-wg@cern.ch)
- **Mandate:**
  - Coordinating with sites, experiments, and developers the deployment of the various SRM 2.2 implementations and the corresponding Storage Classes.
  - Establishing a migration plan from SRM v1 to SRM v2 so that the experiments can access the same data from the 2 endpoints transparently.
  - Coordinating the Glue Schema v1.3 deployment for the Storage Element and make sure that the information published are correct.
  - Coordinating the provision of the necessary information by the Storage Providers in order to monitor the status of storage resources, ensuring that all sites provide the experiments with the requested resources and with the correct usage. People involved: developers (SRM and clients), storage providers (CASTOR, dCache, DPM, StoRM), site admins, experiments
  - Ensure transparency of data access and the functionalities required by the experiments (see Baseline Service report).



# Coordination

- Use portions of pre-GDB meetings.
  - Pre-GDB are a good opportunities since the people involved in the work attend usually in person.
  - Pre-GDBs can be extended to last one entire day if other important technical matters need discussion
- Create sub-groups to work on specific issues.
  - Phone conferences whenever needed.
  - The specific deadlines are established depending on the priority of the item under study by the specific subgroup.
- Use mailing list for status reports or problems/issues.
- Use wiki web site for documents.
- Report progress at WLCG MB



# Work in progress

- Coexistence of SRM v1 and v2 endpoints:
  - It is not desirable to have SRM test instances running on production resources.
    - However it is not always possible to replicate the hardware
    - We will target few sites and make tests with them first
    - We will proceed with wider deployment only if no problems found
- Experiments would like to be able to access data transparently from SRM v1 and v2 endpoints for testing purposes.
  - Is this possible ? CASTOR sites use different endpoints for different classes of storage. They run on different machines.
    - Working group ?
  - It could be necessary to modify the entries in the catalogues. Is this acceptable in the long run ? What about experiment private catalogues ?
    - It is a costly operation and time consuming operation. The entire Grid can be down during this time since no data can be accessed. A strategy is needed.
      - Working group ?
    - During the testing phase, how should new data be registered in the catalogue ?
    - What happens in case a roll back is needed during the testing phase?
      - Working group ?
  - We can try the whole experiment using Tier-2 centres with DPM and the test UI setup for SRM 2.2 to verify that we have not overlooked at some details.



# Work in progress

- Detailed analysis of the LHCb requirements in progress:
  - FZK taken as an example Tier-1 for LHCb
  - Nick Brook has provided necessary information.
  - Sites have requested more details:
    - Which data are online at the same time during the processing step ? What can go on the same tape? Are real data separated from MC data ?
    - What is the average file size per space description ?
    - Is there a need for scratch areas ?
    - What is the data lifetime ? What about pin times ?
    - Are user data really managed by SRM or are they accessible also via "backdoors"?
    - Is there an association between Space Description and Path ?
    - How is data staged ? What order ?
    - Data might be stored on different SE at the same site. Is this a problem ?
    - What kind of storage accounting is required ?
  - Three sites are targeted for this exercise: SARA, IN2P3 and FZK
    - Working group ?



# Work in progress

- Reports from Tier-1 sites
  - Report from RAL and CNAF
  - Some issues to consider:
    - “For a VO is it better to segment the storage and separate the flow into multiple pools or run with a single big pool and average out all the I/O”?
    - Should T1D1 be used to be sure that the files remain available on disk for the needed time and with stage requests failing when the space is full? Or is T1D0 sufficient for large-scale production/reconstruction and should T1D1 only be used for data that needs to be online for a long time?



# Work in progress

- **Storage Monitoring:**
  - How can experiments check the existence and setup of the resources dedicated to the VO ?
  - What tools are available to monitor their status ?
  - What information are available with implementation specific tools ?
    - Investigate what information is possible to extract with implementation specific tools: DPM can report on usage per VO group/role/user; usage info per VO group/role is not yet available in dCache.
    - Come up with some general interface to extract this info
    - Make them available through SAM or GridView via a defined schema, as a first approach.
      - Working group ?
  - SRM Developers should not spend time initially on this issue. They should concentrate in providing a working implementation of SRM 2.2