DPM at GRIF

Michel Jouvin
January 9, 2007
GDB Storage Classes Meeting

GRIF at a Glance

- GRIF : distributed T2
  - 1 middleware site
  - 5 geographical sites in Paris Region (join effort by 5 labs)
    - Distributed technical team : 9 FTE
  - Currently 1 CE and 1 SE per site
  - Open to LCG and non LCG VOs
    - Strong increase of non LCG VOs resources

- Current status
  - CPU : ~1 MSI2K
  - Storage : ~150 TB
  - Network : 1 Gb/s inter-site, 1 Gb/s site-external

- Planned evolution in 2007

- 1 DPM SE per site
  - Between 20 and 75 TB per SE
  - Between 1 and 4 disk servers per DPM instance
  - SAN based or DAS storage
  - Currently on SL3, soon evaluation of SL4 64-bit
    - Plan to evaluate Solaris 10 for disk servers (Solaris zfs)

- 1 general purpose pool + some VO dedicated pools
  - Only permanent type

- Managed by Quattor
  - LAL : main maintainer of QWG templates (M. Jouvin/Cal)

- RFIO client library incompatible with DPM
  - Hot issues for Atlas and CMS
  - Workaround available but risk of conflict with standard part of the middleware
    - Not acceptable for large fabrics with automated deployment
  - New unified library late
    - More or less available but not tested

- Lack of tools for resynchronizing disk servers and database
  - Needed in case of problems with disk servers
    - LAL experienced a problem after a device name change following addition of new SAN based LUNs
  - In particular, need ability to list files missing on disk (but present in db) and file present on disk but not registered in db
    - Required to give feedback to VOs in case of problems

- Lack of flexibility for specifying VO access to pools
  - Currently : open to all VOs or only one
  - Need pools open to a limited list of VO
    - Example : ILC has 2 VOs (ilc and calice) sharing the same resource
  - Need ability to give an exclusion list
    - Example : general purpose pool open to all VOs except atlas
- Problem with users belonging to several VOs
  - Seems to be a bug in srm daemon

Feature Requests

- Improved draining tools
  - Ability to drain a VO from a pool
    - Move VO files from one pool to another

- VO quotas per pool
  - Should allow to define a default quota in a pool to simplify management of shared pools
  - Should allow overbooking (sum of quotas > 100%)

- Long term project : unified SE with distributed disk servers
  - Require 10 Gb/s between GRIF sites
  - Require ability to replicate DPM head node (MySQL)
  - Selective inter-site replication of heavily used files

- Usability of volatile storage for scratch areas
  - Not a DPM issue, rather a use case problem and how to advertise them to VOs