Data Management cluster summary

Krzysztof Nienartowicz

JRA1 All Hands meeting, Helsinki

18 June 2007

Outlook

- Disk Pool Manager status
- LHC File Catalog status
- File Transfer Service status
- Grid File Access Library and lcg_util status
- Biomed
- Glite restructuring

DPM current status

- **DPM running in production at more than 120 sites**
  - Austria, Belgium, Canada, France, Greece, Hungary, India, Italy, Japan, Holland, Poland, Russia, Spain, Switzerland, Taiwan, UK
- **DPM serving 148 Virtual Organizations**
  - HEP experiments, biomed, esr, fusion …
- **Collaboration with NAREGI on interoperability**
- **Collaboration with Globus**

DPM I

- **Functionality offered in current production release (1.6.4):**
  - Management of disk space on geographically distributed disk servers
  - Management of name space (including ACLs)
  - Control interfaces:
    - socket, SRM v1.0, SRM v2.1, SRM v2.2 (no srmCopy)
  - Data access protocols:
    - secure RFIO, GsiFTP (Globus 2.4)

DPM II

- **Functionality offered in latest release (1.6.5)**
  - Full support for permanent and volatile spaces
  - ACLs on disk pools
  - Improved version of dpm-qryconf
  - Recursive srmLs and srmRmdir
  - SL4 32 bits release including gridFTP 2 plugin
  - SL4 on 64 bits architecture (being tested at a few sites)

  - Status: certified – moving onto PPS today (17 June)

DPM III

- **SRM v2.2 interface**
  - Basic tests and use case tests are ok
  - Current stress tests show that the system is stable
  - Being deployed at selected sites in UK and France
  - srmCopy and srmChangeSpaceForFiles still missing
- **Xrootd plugin**
  - Requested by ALICE
  - Prototype delivered in November 2006
  - Currently being tested by ALICE
  - Could also be used by ATLAS and CMS later (ongoing discussions)

LFC

- **LFC in production at over 70 sites for 158 VOs**
- **Current LFC production version 1.6.4:**
  - Support for secondary groups
  - Bulk queries
- **LFC planning**
  - SSL enabled CSec version exists
  - SSL session re-use via Csec or openssl, Csec lib prepared
  - More bulk operations

FTS current status

- **Current FTS production status**
  - CERN and all WLCG T1 sites currently FTS v1.5
  - > 10 petabytes exported from CERN since SC4

- **FTS 2.0 has been well tested on pilot service**
  - > 500 terabytes transferred
  - All VOs have agreed to upgrade to FTS 2.0
    - Validation took a while but it was vital to ensure the continuity of the production service
  - Scheduled intervention this week to upgrade CERN-PROD
  - Release will be made available to T1 sites after 3-4 weeks

FTS II

- **FTS 2.0 new features**
  - Delegation of proxy from the client to the FTS service
  - Improved monitoring capabilities
    - Critical to the ‘overall transfer service’ operational stability
    - Much more data retained in the database, some new methods to access them in the admin API
  - *Beta* SRM 2.2 support
    - This is now being tested on the PPS
  - Better administration tools
    - Make it easier to run the service
  - Better database model
    - Improve the performance and scalability
  - Placeholders for future functionality
    - Minimise the impact of future upgrade interventions

FTS III

- **FTS planning**
  - Evolve the SRM 2.2 code as we understand the SRM 2.2 implementations (based on feedback from PPS)
  - Incrementally improve service monitoring
    - FTS will have the capacity to give very detailed measurements about the current service level and problems currently being observed with sites
    - Integration with experiment and operations dashboards
    - Design work ongoing
  - Site grouping in channel definition (“clouds”)
    - To make it easier to implement the computing models of CMS and ALICE
    - Code exists: to be tested on pilot service
  - Incrementally improve service administration tools
  - SRM/gridFTP split
  - Notification of job state changes
- **Not planned**
  - Not planning to produce a non-Oracle version
    - Sites with no Oracle license can use CPU/memory restricted Oracle XE
  - Not planning to produce a more complex data movement orchestration service (Data Scheduler)
    - No real driving requirements for this

GFAL and lcg_util

- **A number of bug fixes and improvements**
- **Fixes on the way…**
  - Improved error messages
  - Improved python interface
  - Improvements to SRM 2.2 functionality: list support, bulk requests (srmRm, etc)
  - VO argument now not needed if not necessary
  - VO name max length is now larger
- **Planning**
  - Shorter term: more requested fixes, cleaner SRM GFAL functions
  - Longer term: thread safe version of GFAL / lcg_util

BIOMED (background)

- **BIOMED**
  - BIOMED VO would like certain data to be ‘visible on the grid’, with the ability to replicate it to ordinary sites with ordinary SEs
  - Have their own storage and wire protocol called DICOM
  - It is required that data be encrypted (and anonymised) after leaving the DICOM server and before going elsewhere on the site or the grid.

- **Currently BIOMED are**
  - Using a modified version of dCache with gLite I/O and Fireman to provide access the DICOM service (all together called the *DICOM SE*)
  - Hydra is used as a keystore for encryption keys, with the hydra and EDS clients (which in turn use the gLite I/O client) to access the data, retrieve the key and decrypt or encrypt the data.

BIOMED requirements

- **Requirement summary**
  - Files/pictures are in the DICOM server
  - LFN can be study/series/ID, which can be used with DICOM (no need for hierarchical directories and custom file names)
  - SRM interface
  - Efficient access, e.g. gridFTP or https - either are OK
  - Mandatory encryption
  - ACL on each file or picture (synchronized between SEs)
  - Use cases are still not fully understood
    - Need some more iteration with NA4

DICOM integration with DPM

- **Initial design complete, prototyping underway**
  - Need to develop the DPM staging backend
    - Design understood, work pending
  - Need to write the DICOM staging backend plug-in, with a callout to Hydra, to copy between DICOM server and the DPM
    - Designed, prototyping underway

- **Hydra client using GFAL rather than gLite I/O prepared**

ACL synchronization

- **ACL synchronization**
  - To synchronize ACL updates on replicas at arbitrary sites
  - We believe we will need to provide explicit synchronization between LFC, Hydra, (DICOM SE) and other SEs.
  - Design and architecture discussions. Work about to start.
    - Likely based on queuing toolsets (JMS)

gLite restructuring

- **Rationalise gSoap dependencies**
  - Currently upgrading all to 2.7.6b
- **Remove unnecessary CGSI_gSOAP usage**
  - Replace by openssl and GridSite
  - Done for all components it makes sense for
- **Migrate to libxml2**
  - Work started
- **Test DM code with VDT 1.6**
  - DPM/LFC done, lcg-utils + FTS to do
- **General cleanup of build dependencies**
  - Much work done.. ongoing.
- **Significant effort spent on ETICS integration**
  - ETICS build is now working for us
- **32 bit / 64 bit SLC4 builds now ~completed**
  - 32 bit SLC4 build OK. Final fixes for 64 bit build being done.
  - LFC/DPM tested on 64 bit. Other components to test.

Summary

- **Summary and plans**
  - Build work completed: gLite restructuring underway
  - Various improvements
    - FTS 2.0 released
    - New functionality for LFC/DPM
    - Lcg_util / GFAL bug fixes and new methods
    - SRM 2.2 introduced and now in active testing
  - BIOMED design work done, prototyping underway

- **As always…**
  - Significant effort provided to support operational services, user support and bug fixes