



Enabling Grids for E-science

Data Management cluster summary

David Smith

EGEE'06

25 -29 September 2006

www.eu-egee.org



Information Society



- **Will summarise**
 - Recently completed and ongoing work
 - Next tasks
 - Longer term issues
- **For each of the components**
 - FTS, DPM, LFC, EDS & Hydra, GFAL, lcg_utils

- **Support: 60% of effort**
 - Support of CERN FTS instance and expert support for existing ROC/T1 instances / PPS instances
 - Patches and fixes as required
 - Inter-site expert support to debug service problems

- **Next user-visible features**
 - **Delegation in FTS**
 - Will also allow for better use of MyProxy for proxy renewal
 - Planned for certification end September
 - **SRM v2.2 support**
 - Both SRM v2.2 and SRM v1.1 supported
 - Cleaner error classes and messages for monitoring
 - Cleaner SRM interactions to ease service debugging
 - Planned for certification end October

- **Better service monitoring and debugging tools is the next highest development priority**
 - This is required to provide the necessary service level for EGEE VOs

- **Other feature requests are pending but service management and stability is the real focus for next year**

- **Provide a scalable solution to replace the Classic Storage Elements at Tier2s**
- **The DPM has become very popular (59 sites are using it for 82 VOs)**
- **This implies:**
 - Users/sites support (5-10 % FTE at CERN)
 - More features are requested by
 - Site admins
 - Large variety of users (biomed needs encryption, different data access protocols)
- **But manpower very limited**

- **SRM v2.2 interface**
 - Defined at FNAL workshop in May
 - A test version has already been made available in July and is being tested by Jiri Mencak (RAL) and Alex Sim (LBNL) for stability and compatibility with other implementations
- **CASTOR-DPM common RFIO library**
 - CASTOR RFIO does not integrate grid security
 - Different interface to the backend but same I/O protocol
 - Requested by CMS for the Data Challenge in October
 - Development done by Jiri Kosina in collaboration with the CASTOR team
- **Xrootd plugin**
 - Requested by ALICE
 - Could also be used by ATLAS and CMS later
- **Perl interface (Tomas Kouba)**

- **Short Term (September - October 2006)**
 - CASTOR-DPM common RFIO library (now)
 - Support for XROOTD transfer protocol (October 2006)
 - SRM v2.2 interface (October 2006)

- **Medium Term (end of 2006)**
 - Port to SLC4 and VDT 1.3
 - MySQL backups
 - srmCopy
 - DPM DSI plugin for Globus 4 gridFTP server
 - Integration with fabric monitoring
 - Lemon sensors?

- **Longer Term (first half of 2007)**
 - ACLs propagation during replication between SEs (RRS/RMS ?)
 - Quotas
 - Streaming mode (SRM v3)
 - Limit number of streams per disk server (may be useful for some applications like bulk replication, so pool dependent)
 - Implementation of a migrator/recaller to either recall/migrate files automatically between Tier1 and Tier2 or interface to a tape/DVD backend

- **LFC: recent changes (v1.5.7 to 1.5.9)**
 - Addition of client side retry timeout
 - New calls to set/get some replica attributes
 - Some bug fixes

- **LFC longer term**
 - Some changes relating to performance
 - Ssl session reuse
 - Buffer size tuning
 - Possibly introduce some new server side functions
 - encryption of network traffic between clients and LFC

- **LFC is also being used**
 - By CMS as a Dataset catalog
 - By the RCAT project to hold database endpoints and passwords

- **EDS & Hydra keystore: to do**
 - Integrate use of GFAL into EDS tools
 - Splitting of keys between hydra stores

- **GFAL and lcg_utils current activity**
 - SRM v2.2 support is ongoing (for October)
 - A number of bug fixes are to be introduced and tested
- **Longer term**
 - Cache information system queries
 - Retry with different replica in case of failure
 - Thread safety