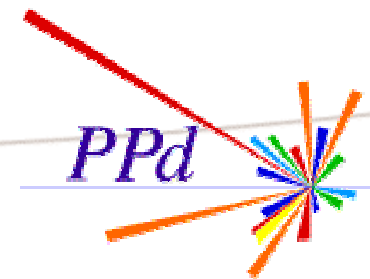# Overview of applications view of the data management middleware

Stephen Burke

Software areas discussed

- BrokerInfo et al

- Replica Manager

- Storage Element

Based on testing on the EDG dev TB and LCG C&T TB, plus preliminary testing on the EDG app TB

Brokerinfo etc

◆ Issues of file matching and use of BrokerInfo fall between WPs, not always obvious who is responsible

- WP1: matching in JDL, .BrokerInfo file

- WP2: general file management, edg-brokerinfo (???)

- WP3 + GLUE: information schema

- WP4: CESEBind info provider

- WP5: TURLs and SE info provider

- WP6: testbed configuration

◆ *This area is very important to users!*

File matching

- Basic matching works OK

- Problems with more than one close SE per CE

- Matching algorithm has some quirks
  - With multiple files will match on 1 even if the others are not available

- Automatic uploading of output files was broken
  - Supposed to be fixed but not deployed

- Gang-matching (matching on SE attributes) apparently does not work

- Optimisation (access cost ranking) not yet tested

- Can only match/use files from one VO

File access in jobs

- Functionality of BrokerInfo getSelectedFile in TB 1 is missing

  - Should return a usable TURL given an LFN and protocol

- Configuration of NFS mount point is less flexible than TB 1

  - Schema does not properly describe mount points, interim solution relies on having a close CE for every SE with identical mount points

- No file locking, job may have to call getBestFile again to recover a deleted file

- Still no local disk space management on the WN

- Slashgrid?

Replica Manager - General

◆ Extensively tested on EDG dev TB

◆ Some testing on LCG

  ▪ No ROS, no WP5 SE, MDS instead of R-GMA

◆ 134 bugs in bugzilla, 33 still open

◆ Biggest problem is speed: various reasons, but much too slow

◆ LFN/GUID/SURL/SFN/StFN/TURL system is very complex!

  ▪ Hidden inside POOL?

RM - Outstanding bugs

- RM does not make full use of Service/ServiceStatus tables

- getBestFile does not deal properly with multiple close SEs

- Some bugs are marked as fixed but not yet deployed

RM - Command line

- edg-rm is a general interface to all DM functions

  - Metadata and wildcard operations excepted

  - Interface is fairly intuitive

- Some errors generate a java backtrace or misleading error messages, but problems are being fixed as they are found

- Exception recovery is not perfect

  - e.g. failed write to catalogue may leave file on disk

  - Different operations deal with failures in different ways

- No equivalent of GDMP (everything is done by the client)

  - No bulk file transfers, no subscriptions, no server-based actions

RM - Catalogues

- ◆ Currently have a single LRC and RMC for each VO
  - ROS is distributed

- ◆ Distributed system desirable for fault tolerance, but may be complex to configure?

- ◆ Nothing to keep catalogues and SEs consistent
  - Are the catalogue data backed up?

- ◆ Still needs testing with large numbers of files

- ◆ Not clear if experiments need metadata in RMC – or even LFNS? (POOL)

RM - Security

◆ Not clear what experiments want, or what they will get!

- Namespace control on LFNs?

- ACLs on files?

- ACLs on SEs?

◆ Will we get the secure web service?

- Denial of service etc.

WP8 view of replica management

◆ **Summary:**

- New replica manager works well in general

- File registration is **_very_** slow (> 20 secs for 1 12-byte file!)

◆ **Issues:**

- Moving to the distributed RLI/LRC system is a big change, can we get it working in time?

- BrokerInfo interaction / getSelectedFile functionality

- VOMS integration – what do we want and can we get it?

- Schema changes – lots of interested parties

Summary of Priorities for WP2

Existing system:

- Registration time

- getSelectedFile

- Exception handling

- Metadata

- Optimisation

New features:

- Secure web service

- Distributed RLI/LRC

- VOMS integration

Vital

Desirable

Low priority

SE - General

- Tested on dev TB only, not installed by LCG
  - Basic functionality works
  - Castor and ADS interfaces seem to be (mostly) working
- Configuration seems fragile, not much guidance for sysadmins (e.g. use of partitions)
  - Validation tests?
- 220 bugs in bugzilla, 51 still open
- Delete still does not work
- Very limited user documentation
  - But largely hidden behind the RM

SE - Command line

- ele* commands removed due to ssh problems, but were more intuitive than edg-se-webservice

- edg-se-webservice command is clumsy (especially in insecure mode), but users will normally only use edg-rm

- Error reporting is poor, most errors give an XML dump, error messages are often meaningless

SE - Architecture

- Separation between permanent storage and cache area makes sense, but only if the cache gets cleaned, otherwise all files are stored twice

- Create/write/commit and cache/getTURL/read cycles are reasonably intuitive

- No space reservation yet

- Storage under hashed names – what is the hash algorithm? Is the mapping from SFNs stored securely?

- Should it be possible to overwrite an existing file? (General model is that files are read-only)

- Each VO has a separate name space, no way to identify a file uniquely

SE - Usage

- create, getTURL, commit, cache, ls, mkdir, getMetaData all OK

  - Easy to forget to call getTURL!

- Can also register an existing file

- No way to get back the TURL for a file which has been created but not committed

- Metadata is limited (creator CN, size, and create, modify and access timestamps) and is only available when the file is cached

- getSECost returns 1 if the file is on disk and –1 if not

SE - Security

- Secure web service badly needed

- No way to specify VO or user with insecure interface, hence everyone is John Gordon

- ACLs?

SE - Information providers

- ◆ Schema is part of GLUE, supposed to be generic for all storage systems
  - ▪ SE providers come from WP5

- ◆ Free disk space reporting is problematic
  - ▪ Does it report the space in the permanent file area or the cache?
  - ▪ Still doesn't deal properly with partitions (but may not be such a problem)
  - ▪ Can cope with different areas per VO

- ◆ Technical problem in the way the DN for the SA (Storage Area) object is constructed
  - ▪ First test of GLUE change procedure?

- ◆ Policy values are published `(MaxFileSize,MinFileSize,MaxData,MaxNumFiles,MaxPinDuration)` but they are not enforced, hence they are meaningless

WP8 view of SE

- ◆ Situation:
  - Functionality is mostly there
  - Lots of minor bugs and problems
  - Configuration seems delicate, SEs often fail to work for no obvious reason, diagnosing problems is hard
  - Error reporting is poor

- ◆ Issues:
  - Space management
  - ACLs
  - GLUE schema

Summary of Priorities for WP5

Existing system:

- Stability/reliability

- Delete

- Secure web service

- Interaction with WP2

- Error reporting

- ele* command line

New features:

- Space management

- VOMS integration (ACLs)