eGee

Enabling Grids for E-sciencE

Data Management
cluster summary

Gavin McCance

JRA1 All Hands meeting, CERN

24 October 2007

www.eu-egee.org

Information Society

Enabling Grids for E-sciencE

- **DPM status**
  - HTTP and xrootd DPM
- **LFC status**
- **FTS status**
- **DM clients: GFAL/lcg_util**
- **Medical Data Management**
- **Bits and pieces**

- **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**

Enabling Grids for E-sciencE

- **Functionality offered in current production release (1.6.5-3):**
  - Control interfaces:
    - socket, SRM v1.0, SRM v2.1, SRM v2.2
  - Data access protocols:
    - secure RFIO, GsiFTP (Globus 2.4)

  - 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 (internally tested)

Enabling Grids for E-sciencE

- **Functionality offered in latest release (1.6.6 / 1.6.7)**
  - Variety of bug fixes
    - GridFTP memory leaks / performance
    - All known SRM 2.2 issues are now addressed
  - Moved to use gsoap 2.7
  - HTTP(s) access to DPM (via apache httpd)
  - Refactored and repackaged xrootd plugin
  - Integration with Yaim 4.0

- **Status: release in preparation**

- **Plans:**
  - SRM copy coming next
  - Quota support

DPM HTTPS & xrootd Protocol

- **DPM HTTPS Protocol**
  - Transparent Access to DPM files & directories with HTTPS/HTTP protocol
  - Full support of X509/proxy & VOMS proxies
  - Full DPM access control - virtual ID mapping
  - High performance for read operation
  - High performance for write operations
    - Write operation via POST from HTTP form or via PUT using e.g. curl or scripting language
  - Default I/O via http - Optional fully encrypted I/O via https protocol
  - Lightweight shell client
  - Configuration via YAIM tools
  - 'dpm-httpd' service based on standard Apache2 daemon

- **DPM xrootd Protocol**
  - Native xrootd daemon serving DPM files
  - Currently no X509/proxy/VOMS authentication
    - All access mapped to single user identity
    - DPM path restricted to directory sub tree
    - Additional authorization plugins possible (e.g. ALICE Authz plugin)

  - Configuration via YAIM tools
  - 'dpm-xrd','dpm-olb','dpm-manager-xrd','dpm-manager-olb' services

DPM Server

https/rfio/xrootd
Clients

TESTSETUP

QuickTime™ and a
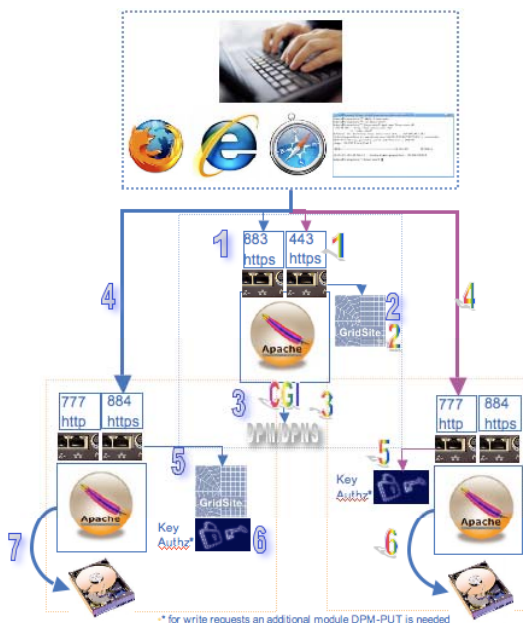TIFF (Uncompressed) decompressor
are needed to see this picture.

- **Parallel Clients reading single identical 1Gb file**
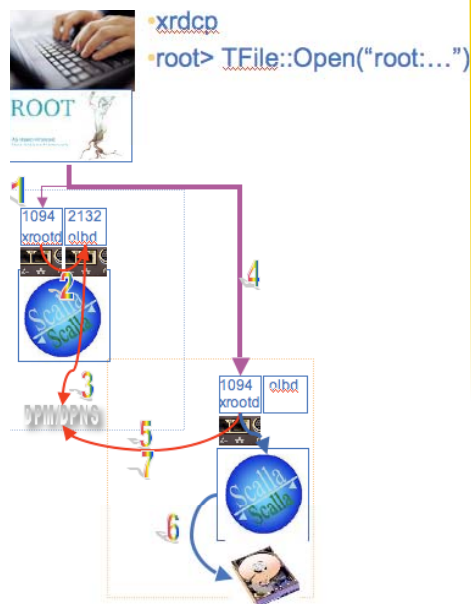 - No disk I/O limitation
  - always sum of I/O for all clients given

Clients	https/ http-io	rfio	xrootd	https/ https-io
1	100 Mb/s	97.5 Mb/s	96.0 Mb/s	98.5 Mb/s [1 core 100% CPU]
10	99.1 Mb/s	100.7 Mb/s	95.0 Mb/s	
40	95.2 Mb/s	93.4 Mb/s	88.0 Mb/s	
80	100 Mb/s	97.5 Mb/s	95.3 Mb/s	
120	96.6 Mb/s	96.6 Mb/s	96.6 Mb/s	
120	60 %	80 %	65 %	Client Load
10/40/80/120	27 %	27 %	27 %	Server Load

- **Parallel Clients reading randomly 1024kb files (time seen by an individual client)**

#Clients	https/http-io	xrootd	rfio
5	0.27 s / file	0.99 s / file	1.61 s / file
10	0.32 s / file	1.57 s / file	1.64 s / file
20	0.37 s / file	3.21 s / file	1.76 s / file
40	0.71 s / file	6.69 s / file	2.52 s / file

- xrdcp
- root> TFile::Open("root:…")

ROOT

1094 2132
xrootd olbd

1094 olbd
xrootd

883 443
https https

777 884
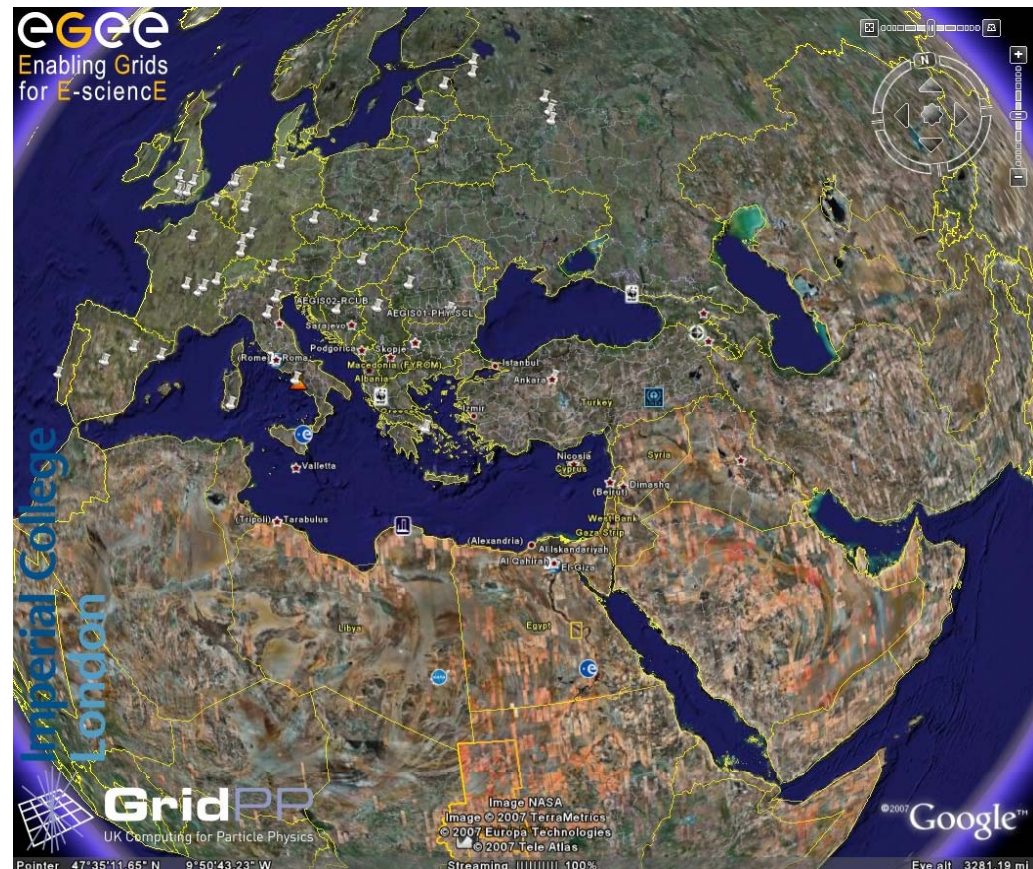http https

777 884
http https

Key Authz*

Key Authz*

* for write requests an additional module DPM-PUT is needed

- **LFC in production at over 110 sites for 158 Vos**

- **Current LFC production version 1.6.5**

- **LFC planning**
  - More bulk operations to improve performance
  - SSL session re-use
  - Operational support

- **Current FTS production status**
  - All T1 sites now have FTS 2.0 installed
- **FTS 2.0 has been well tested and is now in production**
  - It's being run hard in current experiment activities
    - Currently CMS' CSA'07, more to follow
    - Support for this is absorbing most of the team's effort
  - Even after all the testing, we still find issues in production!
    - Slow job cancelation
    - Corner cases where things break (e.g. MAXTRANSFERS bug)
    - These are being addressed with urgency
  - Focus is now upon operational procedures
    - Integration with experiment operations

Enabling Grids for E-sciencE
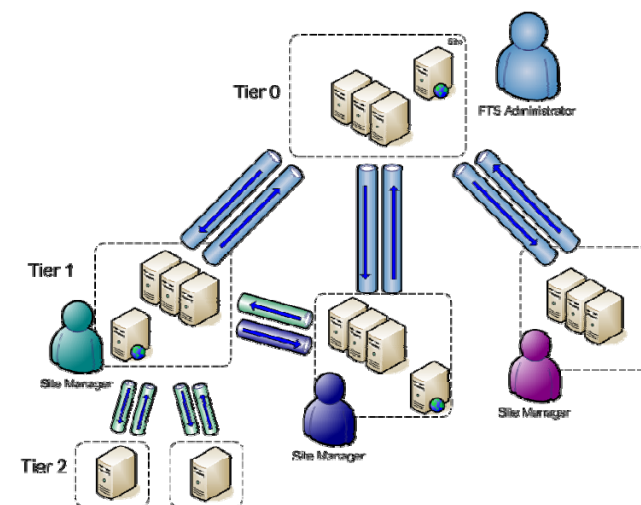
- **FTS development and integration activities**
  - Continue SRM 2.2 integration on production service
  - Incrementally improve service monitoring and admin
    - Better service admin tools being released
    - Monitoring work now prototyped and running on pilot
    - FTM node currently in PPS
    - Incremental approach
  - Site grouping in channel definition ("clouds")
    - To make it easier to implement the computing models of CMS and ALICE, help with ATLAS' operations
    - Being tested at CERN and some T1 (production) sites
  - SRM/gridFTP split
  - Notification of job state changes

Enabling Grids for E-sciencE

- **A number of bug fixes and improvements**
  - More SRM functionality exposed via GFAL / lcg_uitl
  - Cleaned up SRM and GFAL methods
  - Improved error messages
  - Improved python interface
  - All known SRM 2.2 issues addressed

- **Planning**
  - Longer term: thread safe version of GFAL / lcg_util
  - Automatic BDII failover

- **DICOM integration with DPM**
- **Initial design was completed**
  - Implementation proceeding and further design iterations have been made
  - Development of the DPM staging backend
    - Work almost complete
  - Prototype DICOM plug-in developed
  - Key-splitting work done

- **Next step: integration of these into a prototype**
  - Aiming for end of November

eGee

Enabling Grids for E-sciencE

- **Rationalise gSoap dependencies**
  - Done for all components

- **Remove unnecessary CGSI_gSOAP usage**
  - Replace by openssl and GridSite
  - Done for all components it makes sense for

- **Migrate to libxml2**
  - 75% of the work done on the development branches

- **General cleanup of build dependencies**
  - Mostly done

- **32 bit / 64 bit SLC4 now ~completed**
  - LFC/DPM done and in production + FTS underway

- **Test DM code with VDT 1.6 / Tomcat 5.5 / Java 1.5**
  - DPM/LFC done, lcg-util + FTS underway

- **Client / Server split**
  - The runtime packages are cleanly separated
  - On the source side there is much overlap: doesn't make much sense

Enabling Grids for E-sciencE

- **Following minimal model**
  - Don't disturb current production services. The services are critical to the ongoing WLCG production
  - There are lots of scripts / tools using the current logs
  - Close relationship anyway with SA1 / WLCG sys-admins: their comments / suggestions about the logs are usually fed into the next version quite quickly

- **We'll carefully introduce the security related ones**

Enabling Grids for E-sciencE

- **Full service administration guide / FAQ / troubleshooting guide available for all production services (LFC / DPM and FTS)**
  - Mostly (well-edited) Twiki

Enabling Grids for E-sciencE

- **Data management session – Thursday morning**

Enabling Grids for E-sciencE

- **Summary and plans**
  - gLite restructuring ~done
  - FTS 2.0 in production for experiment FDRs
  - New functionality for LFC/DPM
    - HTTPs / xrootd / SRM copy
  - Lcg_util / GFAL improvements and cleanup
  - Full DM stack moving SRM 2.2 into production
  - Medical data management prototype implementation proceeding

- **Focus is on the service delivered:**
  - Significant effort provided to support operational services, user support and bug fixes
  - Significant resources going into supporting current production challenges on WLCG / EGEE / OSG grids