Status and plans on PhEDEx-LFC integration

A. Sciabà (CERN), N. De Filippis (Bari)

CMS/LCG/EGEE Task Force
Integration meeting
Aug 31, 2005

Status

- POOL LFCCatalog plugin released (2.1.0)
- Functionality tests performed
  - POOL CLI thoroughly tested
  - some bugs found, fixed in later releases (now 2.1.2)
- Performance tests performed
  - test suite provided by CMS, same used to validate any POOL catalog
  - no changes needed to use it with LFC
  - performances look reasonable, but no clear indication on whether they are up to production requirements

Test setup

Site	LFC client	LFC server	Backend
CERN	PIII 1 GHz	dual Xeon 2.4 GHz	Oracle
Bari	PIII 1.266 GHz	PIV 2.8 GHz	MySQL

- CERN is the LFC pilot installation
  - database shared among all VOs
- 10000 fake entries inserted
  - LFN-PFN pairs, mimicking typical CMS names

Performance summary of POOL catalogs

	LFC				Oracle(*)		MySQL(*)	
No. of clients	**Oracle**		**MySQL**					
	Query by GUID (ms)	Query by PFN (ms)	Query by GUID (ms)	Query by PFN (ms)	Query by GUID (ms)	Query by PFN (ms)	Query by GUID (ms)	Query by PFN (ms)
1	384	87	264	102	13	7	4	4
5	111	18	79	20	5	6	12	13
20	159	6	120	6	5	5	15	14
50		6		8	7	7	16	14

(*) Results from Jens Rehn

Remarks on LFC performances

- Substantial overhead compared to database access (performances improve a lot with parallel queries)
- Queries by GUID are very slow due to authentication being repeated for each file
- Queries by PFN as fast as with other catalogs for parallel queries
- Time to publish an XML fragment of 19 files to DB:
  - ~ 0.5 s Oracle
  - ~ 3 s LFC Oracle
  - ~ 2.2 s LFC MySQL

Plans on integration with PhEDEx

- Short-term (2-3 weeks)
  - Configure PhEDEx with POOL-LFC interface
  - Run realistic transfers from remote sites
  - Test migration from existing POOL catalogs to LFC
  - More extensive scalability tests?
- Medium term (1-2 months)
  - Test LFC as a POOL file catalog for CMS analysis jobs
- Long term (3-4 months)
  - Stronger integration with LCG SRM SEs (DPM, dCache, CASTOR)?