

Introduction

Here some more results from LFC-bulk method rests will be presented

Module description

- **delfilesbyguid** gets two parameters: *list of GUIDs* and a *flag* ("force"), if replicas should be deleted (1) or not (0)
- **delfilesbyname** gets two parameters: *list of LFNs* and the *flag*
- **delfilesbypattern** gets three parameters: *list of directories*, a *pattern* (eg "file%") and the *flag*
- **delreplicas** gets two parameters: *list of GUIDs* and the *storage element* from which the files should be deleted

Test description

Kind of tests

- Functional tests of:
 - ◆ lfc_deletefilesbyguid
 - ◆ lfc_deletefilesbyname
 - ◆ lfc_deletefilesbypattern
 - ◆ lfc_readdirxp
(tobedone)
 - ◆ lfc_readdirxr (tobedone)
- Performance test for deleting files without replicas and with replicas (force =0 or 1)
- Performance test deleting files from multiple machines
- Performance test deleting files from remote sites
- Performance tests for files with multiple replicas
- Test for gfal modules in respect to bulk methods
- for comparison the deletion via lfc_delete within a lfc_session was used for files without replicas

Test setup

```
. /afs/cern.ch/project/gd/egee/glite/ui_PPS/etc/profile.d/grid_env.sh
. /afs/cern.ch/project/gd/egee/glite/ui_PPS_testing/etc/pps_test_env.sh
export PYTHONPATH=$PWD:$PYTHONPATH
export LD_LIBRARY_PATH=$PWD:$LD_LIBRARY_PATH
export LFC_HOST=lx0717.cern.ch
export LCG_CATALOG_TYPE=lfc
```

the nodes got their old names back. Looking at the dashboard I cannot find any obvious anomalies during the last hours. So it seems we got away with a blue eye

- This LFC is based on a Oracle Database. The modules `_lfc.so`, `lfclib.so`, `lfc.py` should be in the current directory
- For each test several thousand entries are created, then these are deleted by the different modules.
- For comparison the entries have been deleted with the entry by entry method currently used

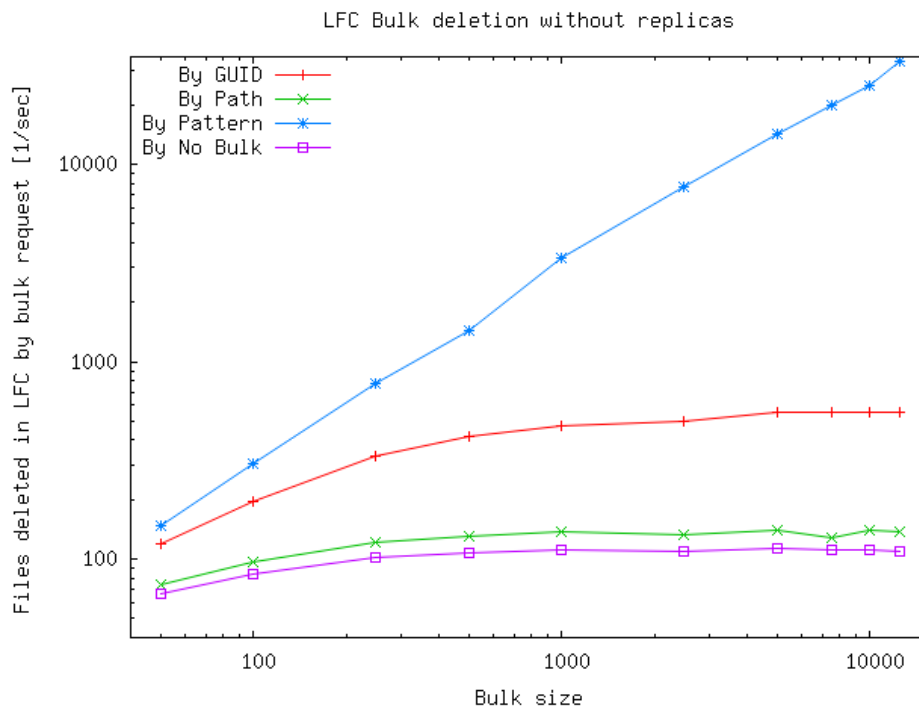
Performance Tests

Just one entry in the LFC, no replica deletion necessary (or possible)

The time is given in seconds, the last test is a comparison with my current solution for deleting entries from the LFC

Number of files	delfilebypattern			delfilebyguid			delfilebyname			unlink file			
	best	worst	per file	best	worst	per file	best	worst	per file	best	worst	per file	
50	0.34		0.0068	0.42		0.0084	0.68		0.0136	0.75		0.80	0.015
50 (3 dirs)	1.65		0.0333	0.42	0.82	0.0084	0.67		0.0134	0.77			0.0154
100	0.33		0.0033	0.51	0.88	0.0051	1.03	1.10	0.0103	1.20			0.0120
250	0.33		0.0013	0.76	1.77	0.0030	2.06		0.0082	2.48	2.63		0.0099
500	0.33		0.0007	1.2	2.5	0.0024	3.79		0.0076	4.68			0.0093
1000	0.33		0.0003	2.07	4.51	0.0021	7.30		0.0073	8.99			0.0090
10000	129	235	0.0129	148	225	0.0148	96	132	0.0096	160	238		0.016
12000	162	253	0.0135	180	211	0.015	102	126	0.0085	179	233		0.0149

- Updated results for LFC Bulk deletion without replicas:

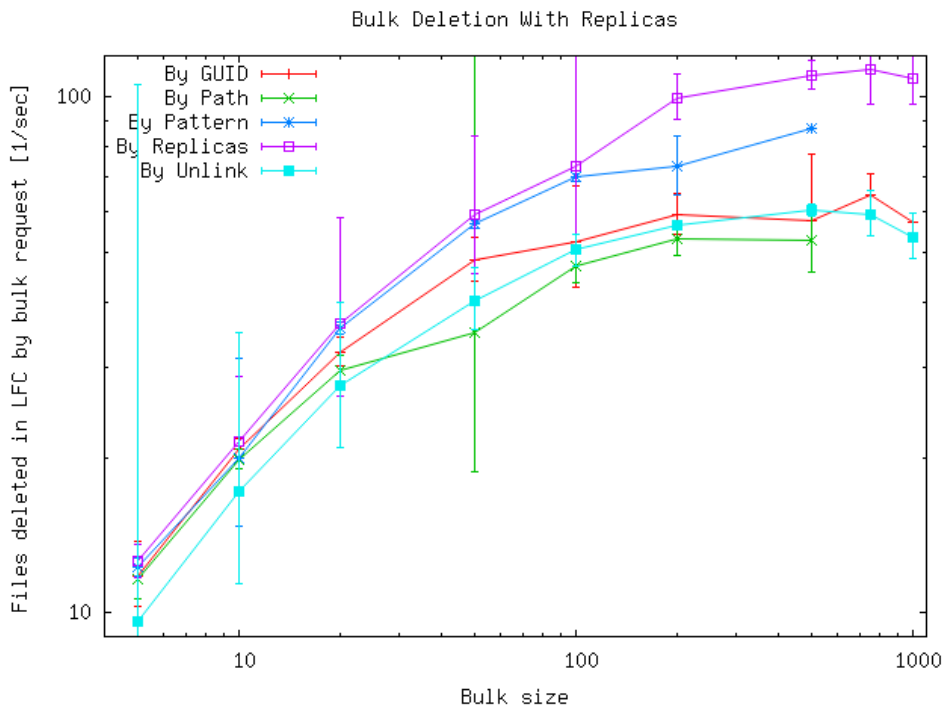


- Data are just reliable for there relative relationship to each other. For absolute numbers are depending on the load of te database server and the numbers of concurrent queries.

Deletion with replicas

- These tests have been done in connection with a srm2 d-cache. (Thanks to Birger Koblitz for the help with the srm2 scripts)
- The data has been taken during a longer period of time with several clients querying the LFC, the errorbars are the standard variation.
- Again these tests just have an impact on the relation between different modules, for absolute numbers a dedicated machine would be necessary

- LFC bulk deletion with replicas:



Experiences from the tests

- Strong performance variation, but this could be due to different load of the testmachine (which wasn't dedicated to the tests), couldn't check
- Oracle DB crashed once while multiple user access (which can be due to another user)
- While uploading files to multiple directories, the LFC gave back errors (which is not clear now)
- Remote access (Desy instead of Cern) did not change the overall deletion time
- Strong performance variation * Machine wasn't dedicated to these tests * Results are not to be taken as absolute numbers...
 - ◆ ... but as a relation between the modules.
- While uploading files to multiple directories, the LFC gave back errors
- Remote access (Desy instead of Cern) was not detectable
- Upload of files to srm2 fails quite often (has to be discussed with Birger)

Preparation of scripts for automation of bulk deletion

- Skripts are functional for bulk operations on the production system. Some more work still necessary
- Deployment of the new Version has to be done.

-- KaiLefhalm - 19 Dec 2007

This topic: LCG > LFCTestForBulk

Topic revision: r9 - 2008-01-23 - KaiLefhalm



Copyright &© 2008-2020 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Send feedback