

Data integrity check on LCG sites

DDM workshop

Cédric Serfon

Ludwig-Maximilians Universität, München

26 January 2007

Reminders

- Two kind of integrity check :
 - "Direct integrity check"
 - Check Files in LFC/not on SE + check discrepancies between file size in LFC and in SE.
 - Script already in CVS (offline/Production/swing/scripts/ddm)
 - "Reverse integrity check"
 - Check Files on SE/not on LFC
 - First release available in CVS.
- Scripts need to be run locally. For dCache SE, need to be run on a machine where /pnfs is mounted.
- During last SW week, we agreed on unregistering files on LFC when not on the storage element.
 - Script won't do this deletion before being validated. Needs to be done manually.
 - It only drops the list of files on the SE and their status (OK/ERROR).
- If file is the only replica, no deletion on DQ2 catalog.

Some results

- Integrity check performed on some of the GridKa cloud's Tier2s and on FZK as well on a few other sites. Results below.

Site	SE	In LFC not in SE	In SE not in LFC	Different size	0 length in LFC	0 length on SE
FZK	dCache	12 766	31	56	627	18
CYF	dpm	1	?	0	0	2
DESY-ZN	dCache	599	?	0	0	0
WUP	dCache	12	?	0	0	0
ASGC*	Castor	?	20	?	?	1559
CERNCAF*	Castor	?	3 095	?	?	5
CERNCAF	Castor	2 386	?	14	307	154
CPPM	dpm	67	?	1	0	1

* Results provided by Zhijun's scripts.

- All these numbers are not up-to-date :
 - Run on FZK (each day) and CERNCAF (each week).
 - For other sites, don't have access and need to be run be people on these sites.
- For FZK, a large amount of files which doesn't exist on the SE (12 656) are due to the automatic deletion of sc4 data.

Running script remotely

- As said before, for the time being, scripts can only be run locally. It will be good to run them remotely using grid jobs.
- Many problems to run scripts remotely on a WN :
 - For dCache, /pnfs is not mounted on all machines. Can't access the data.
 - Even when /pnfs is mounted, need to know the path where atlas data are put. New entry in ToA ?
 - Not all sites have afs access → Need to install dq2 client locally. Only available on SLC3/4.

Speed

- Although much faster than using lcg commands (lcg-lr), speed will be quickly a problem.
- Number for test in GridKa ($\sim 190\,000$ files) :
 - Dumping all LFC content for a given site : 47 seconds !
 - Dumping all LFC content for a given site on the screen : 13 minutes...
 - Integrity check : $\sim 20\,000$ files per hour.
- The check itself is really the most time consuming.
- A solution to speed up could be to use parallel processes on the LFC dump.
 - First attempt made in CERNCAF using LAM/MPI.
 - Tried to run several processes (up to 20/CPU) on one or more computer.

Issues

- Test on Ixplus machines :
 - Biproc Intel(R) Xeon(TM) CPU 2.80GHz
 - 2 GB RAM
- Results in following tables :

	Number of processes per computer					
	1	3	5	10	15	20
Time to process 10 000 files	2 567 s	1 156 s	836 s	775 s	627 s	694 s

Time to process 10 000 files on one computer, with different number of processes.

		Number of computer used			
		1	2	4	6
Number of process per computer	5	6 430 s	-	-	3 832 s
	10	-	-	3 421 s	3 162 s
	20	-	-	-	3 063 s

Time to process all the files in CERNCAF (~96 000), with different number of processes.

- First results show we can improve a little bit the speed (~ 100 000 files/hour).
- But not easy to setup and gain in speed not so high.

Conclusion

- Lots of errors due to files registered in LFC but not on SE : e.g. in FZK :

```
Failed on SRM get: Failed SRM get on httpg://srm-durable-atlas.cern.ch:8443/srm/managerv1 ; id=560727043 call
Error is specified file(s) does not exist
Failed SRM get on httpg://dcache.gridpp.rl.ac.uk:8443/srm/managerv1 ; id=-2147472787 call. Error isRequestFil
failed with error:[ file not found : can't get pnfsId (not a pnfsfile)]
```

- By applying integrity check on all Tier1s, we can remove most of these errors.
- As already said, can't do it remotely from the time being (need someone to run it locally). But I can provide some help.