

Lost/corrupted files : Correction procedure and status report

ATLAS Software & Computing Workshop

Cédric Serfon

Ludwig-Maximilians Universität, München

November 29th 2010

Consistency Service

- Already presented in last Software week in DDM session.
- This is a set of tools composed by :
 - Client tools to declare suspicious files (files that might be bad)/bad files (files that are confirmed to be lost or corrupted).
 - A database backend where all the bad files declared and the history of the recovery of these files is recorded.
 - An agent that treats the reported error running on one machine at CERN.
 - A monitoring to get the history of the losses, follow the evolution of the recovery : <http://bourricot.cern.ch/dq2/consistency/>

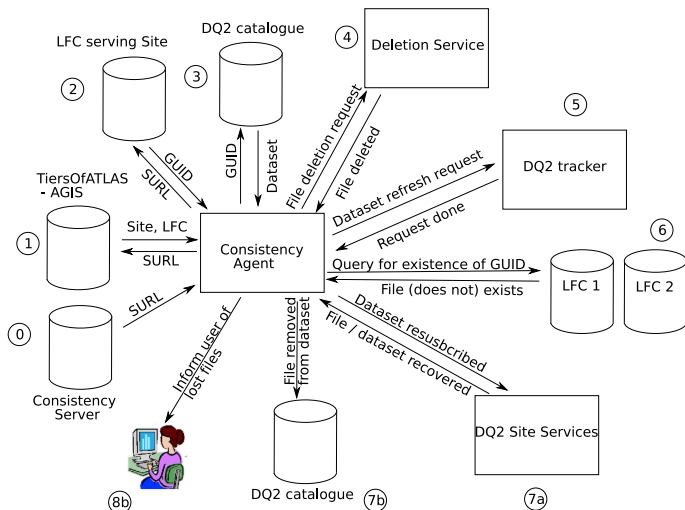
Consistency Service

- Since last time, many improvements in :
 - Monitoring :
 - Search engine/summary for "incidents".
 - File recovery history.
 - Improved speed.
 - Agent :
 - More rare errors handled.
 - Improved speed.

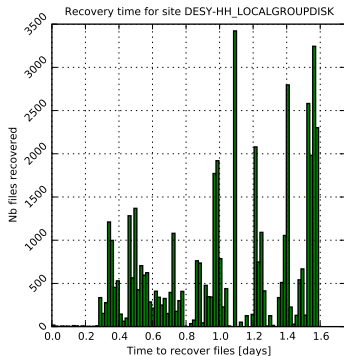
How does it works ?

- Bad files can be reported by every people with Production Role and will trigger action (catalog clean-up, recovery : see next slide). You just need to provide :
 - List of SURLs.
 - A "reason" (max 64 characters) describing the problem (e.g. pool crash, tape damaged, SE database crashed...)
- Suspicious files can be reported by every ATLAS users, DA tools, Sites Services... It doesn't trigger any recovery. People with Prod Role can confirm that a suspicious file is actually bad and the recovery will be done.

Agent description



Recovery performances

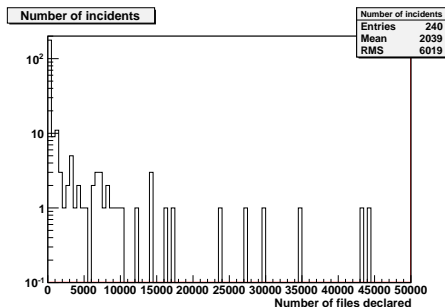


- In general full recovery fast : For loss up-to 20 000 the recovery generally takes less than a day.
- Recovery can be longer if problem with deletion service (e.g. long deletion queue), high transfer rate from other activities, datasets open...

Recovery time for a loss of 44 000 files

Statistics

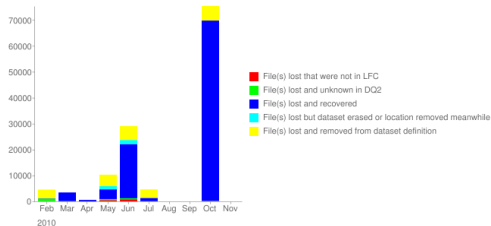
- Service running since beginning of February.
- 240 "incidents"¹ affecting 122 DQ2 sites (45 physical sites) and 490k files.



- 13 incidents with more than 10 000 files. Mainly hardware failures.
- Most of the incidents reported by UK and DE squads that systematically report all bad files. Not sure to which extent other clouds also systematically report.

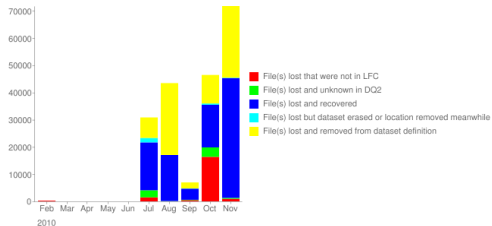
¹An incident is a triplet (site,date,reason)

Statistics : DE cloud



- 7 sites reported losses. ~130k files.
- In more than 95% of the cases : pool crash, controller problem, file system (xfs) failure.
- About 20% of the files were unique and couldn't be recovered.

Statistics : UK cloud



- 6 sites reported losses. ~200k files.
- Big contribution from RAL (~130k files) due to faulty batch of pools.
- In other cases very similar to German sites.
- About 35% of the files were unique and couldn't be recovered.
- Significant quantity of Dark Data (in red and green on the plot) in the files reported lost.

Message to the squads

- Statistics for other clouds can be obtained on <http://bourricot.cern.ch/dq2/consistency/history/search/>
- The system should be systematically used by squad. Many advantages :
 - Easy to use : Just have to run one command and nothing else.
 - Standard recovery procedure.
 - Notification of users in case of definitively lost files.
 - Long term book-keeping.

- Instructions :

https://twiki.cern.ch/twiki/bin/view/Atlas/DDMOperationProcedures#In_case_some_files_were_confirme

Things still to be done

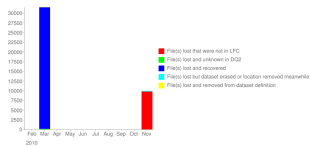
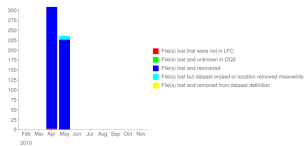
- Integration in pilot.
- Automatic notification to squads about suspicious files.
- A problem with user notification needs to be fixed : Notification of lost files sent to replicas owners. In LOCALGROUPDISK sites, many replicas have Kors, Graeme, Mikhail... as owners. → Sometimes real owners are not notified. Would need to put the correct owner. Probably needs to be done by squads.
- Possibility to flag the datasets replicas with suspicious files so that they won't be used by DA tools.
- Currently only API available. CLI needed ?

Conclusions

- File losses/corruption are frequent (at least $\sim 500k$ files in last 10 months).
- The consistency service allows to correct these bad files with a well defined procedure.
- It aims also at making life easier for sites/squad in case of big loss since cleaning/recovery is fully automated.

Backup

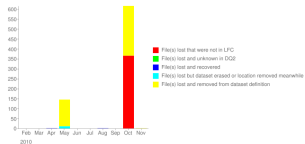
Statistics



CA cloud



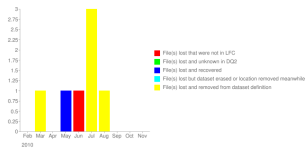
CERN



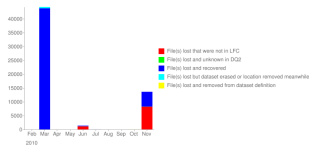
ES cloud

FR cloud

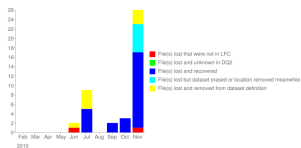
Statistics



IT cloud

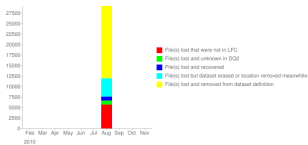


ND cloud



NL cloud

TW cloud



US cloud