lundi 8 septembre 2008

SRM/dCache issues at IN2P3-CC Tier-1

Jonathan Schaeffer / Lionel Schwarz

dcachemaster@cc.in2p3.fr

dapnia
cea
saclay

CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE

Contents

- Current top issues
  - Prevent VOs accessing data
- Other hot issues
  - Need a fix but workarounds have been found
- Other issues
  - Currently not a showstopper for VOs
- Configuration details
- Stability/performance improvements

Current top issues

- SpaceManager SQL Error (#3454) preventing VOs to write into the system in some cases
- Atlas SrmBringOnline with multiple files fails with 'lifetime expires' even if all files were successfully brought online (#3451)
  - Serious issue but occurred only once /1000

Other hot issues

- Calculation of available space does not work when HoppingManager is used (Atlas T0D1) (#2427)
  - As workaround : UpdateLinkGroupsPeriod set to high value and “available space” changed directly in the database
- Srmls -l gives wrong LOCALITY for T1D0 (#3186)
  - Workaround in PoolManager.conf with SRM IP@
- Missing checksums in PNFS: not understood but seems to have disappeared

Other issues

- gsidcap multi-thread issue, impacting LHCb jobs
  - A bugfix allows the jobs to run but the issue can still be artificially reproduced.
- Inability to write with gsidcap when ReserveSpaceForNonSRMtransfers=true (#2280)
- Deletions are not traced in billing database
- Missing documentation (database schema, parameters usage, changeLogs not detailed...)
- Need to distinguish between patches and releases
  - A patch is just applied, a release is tested before...

Configuration details

- Dedicated dcap doors for Atlas and CMS
- Separate queues on pools for gsiftp, dcap, xrootd
- Separating pools by usage (transfers or jobs IO)
  - spacecostfactor=0.0 for incoming transfer pools
    Avoids to have all transfers on a single pool
- Isolate infrastructure (power, machines...) for T0 import (in progress)
  - Ability to keep minimum SE availability for T0 import

Stability/availability improvements

- Use Java6 instead of Java5 helped avoid memory problems
- double power supply for each core server
- Redundant gsiftp, gsidcap doors (not yet dcap)
- Core services failures detection and auto-restart

Performance improvements

- Network interface aggregation on pool nodes serving the data (2x1Gbs)
- Archiving of old (>1 month) billing DB entries
  - Keeps live billing DB small
  - Helps data mining on archived DB
- PostgreSQL 8.3 seems much better than 8.1
  - Installed for all core services now