Summary of experiment open issues with data management and dCache

Flavia Donno

CERN/IT-GS

dCache T1 data management workshop,
Forshungszentrum Karlsruhe (FZK)
14-15 January 2009

Introduction

Introduction

- ***The Pre-GDB held at CERN in November 2008* was dedicated to Grid Storage Services. The goals were:**
    - Bring to the attention of site administrators and developers the most urgent issues and missing features needed by the experiments in order get ready for LHC operations
    - Make release plans well known for each storage implementation
    - Share experiences and best practices
    - Get configuration and management advice
    - *http://indico.cern.ch/conferenceDisplay.py?confId=20247*

- ***WLCG daily operation meetings* reports on incidents seen by the experiments. Very often such incidents are sign of real problems within the storage services.**

Most urgent issues: experiment view

Most urgent issues: experiment view

- **Stability and reliability of storage services for basic experiment activities (import, export, processing)**
  - The response of the service should be good for requests that are being processed even if the server is loaded. Refused connections are better than making a user job wait for days.
    - This applies to disk movers, recall from tape operations, etc. In general, all dCache components should not just make clients wait too long if busy but the connection should be not allowed.
  - The services should not crash under any condition.
    - Would it be possible to tune the internal storage services so that they are correctly configured for the available hardware resources (memory, number of threads, etc.) ?
    - Protection against DoS on SRM doors
  - *Documentation about how to tune the various dCache components*
    - Some parameters need to be well exposed to system administrators
    - Clear directions from the developers

dCache T1 data management workshop, FZK - 14-15 January 2009

Most urgent issues: experiment view

Most urgent issues: experiment view

- **Stability and reliability of storage services for basic experiment activities (import, export, processing)**
  - The activity of one VO should not penalize the concurrent activity of another VO
    - Splitting of service instances whenever possible. SRM ? xxxManagers, etc.
    - Anything else ?
  - It should be possible to privilege certain activities over others
    - Limiting the number of requests of a given type per DN or FQAN. This includes srmLs, pinning, releasing pins, mkdirs, put, get, etc.
    - Pinning should be allowed only by given FQANs
  - *User analysis activities* are somehow difficult: multiple threads/connections, file access, etc.

Most urgent issues: experiment view

Most urgent issues: experiment view

- ## Stable and performing pre-staging service
    - Attempts to pre-stage datasets from tapes from VOs have not been successful at an acceptable level. Furthermore, concurrent multi-VOs exercises have not yet been systematically performed.
    - Therefore, a coordinated effort to test multi-VO prestaging is highly desirable. The tests should exercize all levels of interactions: performance of the tape system, interaction between the dCache systems and the tape system, improving dCache internal tuning for tapes, SRM implementation in dCache for stage activities support, interaction between dCache and WLCG high level interfaces/GFAL/lcg-utils (a few bugs have been found: is this all?), interaction between dCache and experiment applications. Furthermore, things such as access control, priorities, etc. need to be looked at.
        - CERN can take the lead to coordinate such an activity. A set of standard tests must be put in place. Sites can test more if needed.
        - However, contributions from site administrators, dCache developers and experiments are needed.
        - Each site would need to document and publish the tests executed, its status and performance numbers, together with configuration practice and short-cut taken.

Most urgent issues: experiment view

Most urgent issues: experiment view

■ **Performing name service: Chimera ?**

- ■ PNFS has shown criticalities when performing operations on the name space. Moving to FastPNFS and to Postgres 8.3 has improved the situation but sites have sometimes touched new limitations. Is Chimera really solving the problems ?

- ■ A coordinated effort to test Chimera and its promised increased performance should be promoted by some Tier-1. Any volunteer ? Experiments are really interested in the outcome of this test.

- ■ Experiments would like to have *directions on the usage of the system*: for instance, requests such as srmLs can create bottlenecks in the systems. This should be one of the outcome of the Chimera testing activity.

    - ■ How many files can be listed with srmLs ?
    - ■ Is directory listing discouraged or not supported ?
    - ■ Can the directory listing be limited by an internal count ?
    - ■ Should a request lifetime be always specified ? What is a reasonable lifetime ?
    - ■ Which other operations are critical ?

Most urgent issues: experiment view

Most urgent issues: experiment view

- **Access Control at the level of resources and namespace**
  - With the increasing number of users accessing the system, the problem of controlling resources usage and operations on the name space becomes more and more important and pressing
    - Tape protection, access to some disk resources (spaces), selection of staging pools based on some criteria, ACLs on directories (and files ?).
  - In general it would be nice to have _**quota**_ based on DNs or FQANs.

Most urgent issues: experiment view

Most urgent issues: experiment view

- ## Better monitoring of resource usage
  - It is necessary to easily gather the information concerning resource usage from a client
  - Both CMS and ATLAS have requested the possibility to check the consistency between the name service and the files available on the pools.
  - LHCb interested in having a list of LOST files if available. The clients should report as LOST the files that cannot longer be read.
    - Some general inconsistence between various databases (SRM, pin, PNFS) are observed and already reported to developers

giovedì 15 gennaio 2009

Most urgent issues: experiment view

Most urgent issues: experiment view

■ **Support for user analysis**

- ■ Directories with large number of entries
- ■ Support for small files
- ■ Reliable file access via xroot or gsidcap
- ■ Support for a high number of concurrent requests

Thank You

dCache T1 data management workshop,
Forshungszentrum Karlsruhe (FZK)
14-15 January 2009