New DPM/Xrootd Experience Tests with ALICE at GRIF

DPM workshop
LAL, 3 décembre 2012

Michel Jouvin
jouvin@lal.in2p3.fr

Outline

DPM/Xroot at GRIF

- ◉ 1 ALICE SE at GRIF/Irfu: never worked properly
  - › Many unexplained timeouts
  - › Slow
  - › Huge demotion factor in ALICE: never used for writing
    - • 100 TB free… for many years!
  - › Not a lot of motivation in ALICE and DPM to fix the problem

- ◉ 1 ATLAS SE at GRIF/LAL for a local PROOF cluster: similar situation
  - › Missing Xrootd method for PROOF support: fixed at some point but a regression problem later
  - › When we managed to get it working, we were killed by the 1s minimum time to open a file due to asynchronous DPNS Get
  - › An interface never really used to the standard ATLAS DPM
    - • Heavily used through SRM + gsiftp/rfio

Motivations for Retrying

- ⦿ ALICE : wanting to avoid running a separate storage infrastructure for ALICE
  - › ALICE DPM at Irfu was a subset of a multi-VO large DPM instance
    - • Different from IPNO dedicated to ALICE

- ⦿ ATLAS and CMS started Xrootd federation pilots
  - › Gaining momentum in each VO
  - › GRIF a major site for these Vos: they wanted to have it part of the pilot

- ⦿ PROOF use case is still there at GRIF/LAL
  - › Would like to avoid duplicating storage: money + data management
  - › Demand may increase with the local cloud available to end users
    - • Easy to instantiate a private PROOF cluster

Configuration Used

- New tests based on EMI-2 DPM 1.8.3 + DPM/Xrootd 3.0
  - › Complete rewrite of the DPM backend for Xrootd: gsi security support, synchronous interaction with DPM…

- Deployment integrated with our local configuration management system (Quattor)
  - › Part of QWG templates, available to the community
  - › Includes the Xrootd part but some polishing still required

- Deployed mid-October at GRIF/IRFU and last week at GRIF/LLR
  - › GRIF/LAL about to deploy it after the EMI-2 migration delayed by the need to integrate the new http interface to our configuration

Achievements

- GRIF/Irfu DPM is now an operational SE for ALICE
  - Passing with success ALICE functional tests since mid-October
    - For some reason, production jobs requiring a SE restarted only mid-November…
  - Null demotion factor now: means that the SE is used by production job if it is close to the CE where the job runs
    - Probably mainly jobs also running at GRIF/Irfu
  - Space used increasing continuously: means that jobs write successfully on the SE
    - No negative impact on job success rate…

- Performances: no figure yet… but no complaint!
  - Not clear how to assess I/O performances with ALICE jobs: still trying to get the information…

Problems

- 1 Xrootd crash right after the SE was reenabled for production jobs
  - › Reason still unclear: crash dump were disabled, not really possible to do any troubleshooting
  - › Xrootd crash: not necessarily due to DPM backend
  - › David Smith reported a known problem with gsi authentication support recenly fixed
    - • Waiting for the new Xrootd version

- Performances data: ALICE relies on ApMon
  - › Component specific to ALICE: not part of standard Xrootd
    - • It seems it is not the mainstream approach to connect data from Xrootd
  - › Contact established between ALICE experts (Costin) and DPM developers (David) to understand how ApMon could be integrated with an Xrootd server + DPM backend

Future Work

- Following up the crash issue

- ApMon deployment to get complete statistics in MonALISA

- Integration in ATLAS and CMS federations
  - Done last week for CMS

- Understanding performance limitations
  - Rerun the PROOF test