Data Management tools and operational procedures in ATLAS : Example of the German cloud

Cédric Serfon (Ludwig-Maximilians-Universität)

1. Introduction

Distributed Data Management

Distributed Data Management (DDM) is one of the key component in ATLAS that is used by many other domains (Production system, user analysis...). In ATLAS, DDM is performed through the use of a software called DQ2 that interacts with different services and catalogs. There are 2 kinds of catalogs :

- LCG File Catalog (or LFC) : Maps GUIDs to LFNs and also store for each GUID a list of Physical File Names. There are 17 different LFCs (1 at CERN, 10 on the Tier1s, and 6 at US Tier2s).
- DQ2 Central catalogs : Located at CERN. There are many different catalogs, one of them called location catalog, lists the datasets names, their replicas...

German cloud

In ATLAS, each Tier2 (T2) is associated with a Tier1 (T1). The association means that the T1 hosts some of the services that are used by the T2s (LFC/FTS). The structure formed form the T1 and the list of its associated T2s is called cloud. One of the 10 clouds in ATLAS is the German cloud[1] (that also contains sites from A, CH, CZ, PL) whose T1 is GridKa (Karlsruhe).

2. Tools developed

Introduction

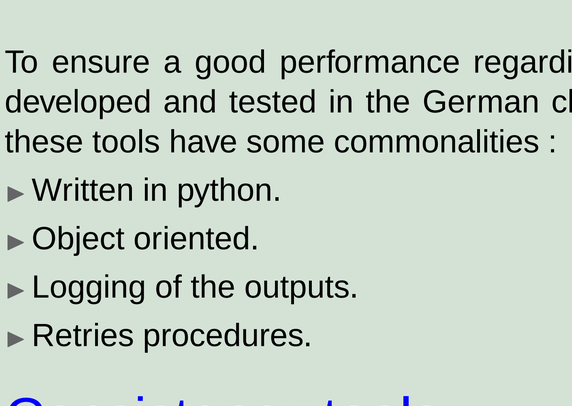
To ensure a good performance regarding DDM issues, a set of tools has been developed and tested in the German cloud. These tools are detailed below. All these tools have some commonalities :

- Written in python.
- Object oriented.
- Logging of the outputs.
- Retries procedures.

Consistency tools

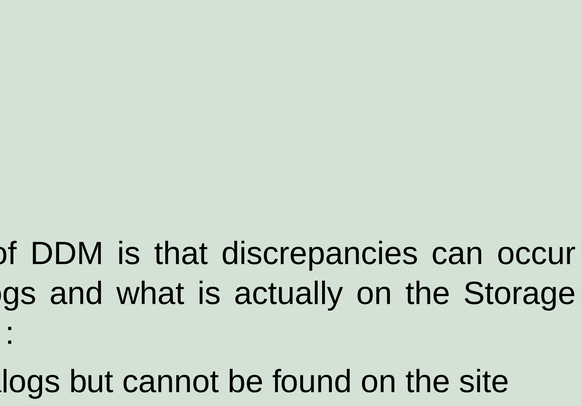
A major issue linked to the complexity of DDM is that discrepancies can occur between what is registered in the catalogs and what is actually on the Storage Element. There are 2 differents problems :

- Files are registered correctly in the catalogs but cannot be found on the site
  → Bad for users, make their jobs crash.
- Files are on the Storage Element but not registered in catalogs (aka Dark Data)
  → Bad for the sites since these data are unusable and waste disk space.



There are 2 types of Dark Data : File not registered in DQ2 (a), that are mainly leftovers from the Production System ; files not registered in the LFC (b) coming from failed FTS attempts or problems with DQ2 (recreation of DQ2 database that let unregistered files).

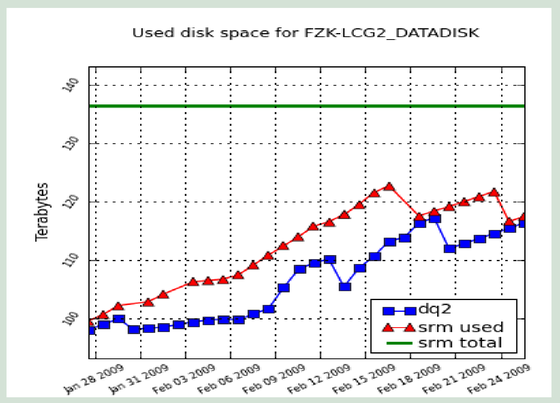
The lost files (c) are files that disappeared from the Storage Element but that have not been cleaned from catalogs.



Number of unregistered files of type (b) versus time

To correct these inconsistencies, a set of tools is available.

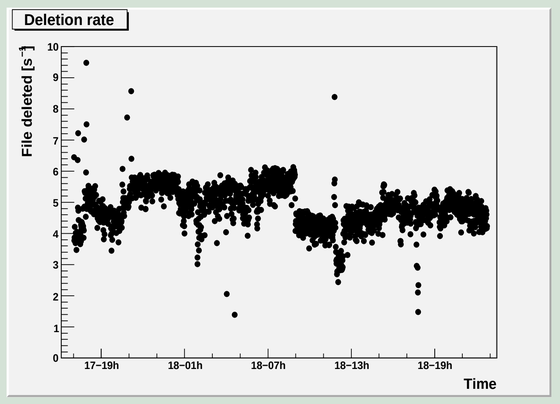
- The first of these tools ensures the consistency between LFC and DQ2 (it checks that every files in the LFC is registered in DQ2) : it loops over all files registered in the LFC for a given site and checks if these files are declared in the location catalog. If the files are known from DQ2 (the GUID belongs to a dataset) the location is added, otherwise, the file is cleaned-up from LFC and from the Storage Element.
- The second of these tools ensures the consistency between LFC and SE in both directions. This consistency check cannot be performed remotely using SRM commands (too slow), therefore it uses a dump of the Storage Element provided by sites. This dump is uploaded into a dedicated space and an agent retrieves it and compares it to the LFC content.



Comparison between information from the DQ2 catalogs (blue curve) and SRM information (red curve). Effect of running regular consistency checks (18th February, 1st March) can be seen. Plot extrated from DQ2 accounting service[2].

Local deletion

An other development performed consists in implementing a local deletion method to the deletion service which is part of DQ2. This service is used to clean consistently the catalogs and the Storage Elements. The default method is to use SRM commands to perform the clean up of the Storage Element. Implementation of local deletion is now available for dCache (deletion on Castor still under testing) and is regularly used at GridKa (plot bellow). It allows faster deletion 6-7 Hz (1 Hz for SRM access) without putting high load on SRM.



Performance of the deletion service at GridKa with local access.

3. Operational procedures

In order to keep the consistency, regular checks are performed on all sites of the German cloud :

- The tool that detects lost files and Dark Data of type (b) needs as input a fresh dump of the Storage Element and therefore is run once a month (all sites perform and upload their dumps each 7th).
- The tool to detect Dark Data of type (a) can be run at any time since it doesn't need any dump. Run also once a month.

References

[1] J. KENNEDY *et al.*, *DE cloud operations, CHEP 2009*

[2] V. GARONNE *et al.*, *The ATLAS DQ2 Accounting and Storage Usage service, CHEP 2009*