ATLAS DQ2 Deletion Service

D.A. Oleynik, A.S. Petrosyan

Laboratory of Information Technologies, JINR, Dubna

Introduction

The ATLAS Distributed Data Management project DQ2 [1] is responsible for the replication, access and bookkeeping of ATLAS data across more than 130 distributed grid sites (700 storage endpoints, 250 millions of files, 75 PB). It also enforces data management policies decided on by the collaboration and defined in the ATLAS computing model.

The deletion service is one of the most important DDM services. This distributed service interacts with a 3rd party grid middleware and DQ2 catalogs to serve data deletion requests on the grid. Furthermore, it also takes care of retrial strategies, check-pointing transactions, and performance throttling, ensuring DQ2ʹs scalability and fault tolerance.

Main development and deployment for production of Deletion service was done in 2010 [2]. The aims of works at 2011 was improvement in productivity of the service (this was done by tuning algorithms and optimization of DB backend usage) and support in operation issues.

Initial clarification

The logical unit of data in ATLAS computing model is a dataset. The dataset consists of files: form one to tens of thousands. One file may belong to more than one dataset (overlapping). All operations with data (collecting, replication, deletion) are provided with datasets.

According to this declaration, the Deletion service should produce the following operations:

- Defining list of non overlapped files for deletion for dataset on storage element
- Remove associated LFC entry for this list of files
- Physically remove files on storage
- Mark dataset as deleted on site

Deletion service should provide a reasonable deletion rate and take care of loading each component (LFC servers, Storages, DB backend).

Service architecture

Deletion service is a distributed information system which consists of the following components:

- Server: collects deletion requests, stores detailed information about datasets, their contents (files) and the ongoing state of deletion requests.
- Client: provides communication between server and other components of the service.
- Deletion Agent: realizes deletion process on storages.
- Monitoring: display the deletion process with the current and errors.

Deletion Client/Server has been implemented with web services technologies: Apache, mod_python and gridsite. They are also designed to encapsulate the interaction with the Oracle DBMS. Using web service technology increases scalability of service, simplify deployment model, and provide easy integration with other DQ2 services.

Deletion Agent

Deletion Agent provides a logic of the deletion process: resolve list of files for deletion on site, deletion from LFC catalog, deletion from mass storage systems. Deletion agent works as three independent (parallel) process: resolver, catalog cleaner, storage cleaner. For increasing the productivity of service and reducing the loading of the server part, most operations are processed with chunk of files (bulk operations implemented).

Resolver: resolves content of dataset (list of files) for deletion, store this content in database for future processing.

Catalog cleaner: deletes files on file catalogs(LFC). It depends upon results of the operation, 'catalog cleaner': mark files as deleted or deleted with error, or mark files for repeating deletion. In case all files of dataset were deleted from LFC, this dataset is marked as deleted from LFC.

Storage cleaner: deletes files in storage systems. It depends on results of deletion, files can be set in states: deleted or to repeat deletion again. For balancing the loading of storage systems provides delays between deletion operations. In case all files of dataset was deleted on storage - this dataset is marked as deleted on storage.

If dataset marked as deleted on storage and deleted from LFC, this dataset is marked as successfully removed.

A lot of parameters like chunk size for bulk operation, delays between operations are configurable for each site. It gives possibility to determinate different deletion strategies for different sites.

Deletion agent is a multithread application. One instance of 'resolver', 'catalog cleaner', 'storage cleaner' serve for each DDM site. So, for the moment, Deletion agent creates about 400 parallel process.

'Storages plug-ins'

Interfaces with different mass storage systems can be done via different protocols. Using of storage specific protocol can significant increases performance of operations. The idea of 'storage plug in' is to be able to implement the same method for any protocol type.

Integration with other DQ2 services

Deletion service is a part of Distributed Data Management system DQ2, so it intercommunicates with some other subsystem: Blacklisting service provide possibility of automatic disabling of deletion operations for some endpoints. Special service for definition and declaration for deletion of unneeded information on sites – Victior, has a close intercommunication with Deletion service: Victor collects some information about the deletion process and provides a lot of deletion requests (list of datastes for deletion).

Deletion service monitoring

Deletion monitoring is a web application based on the Django framework, which provides live graphical and stats reports about deletion process at ATLAS sites, involved.

The info is available at the cloud/site/endpoint levels. It allows one to select statistics at different periods: from last hour/last 4 hours/last 24 hours/last week. Also, there is a future plan report which gives the amount of datasets to be deleted in next 24 hours.

In addition to graphical reports, monitoring generates table with info about waiting/resolved/queued/deleted datasets, amount of files deleted, GBs deleted and amount of errors. Table is expandable. There are datasets and errors browsers. The information is generated via jQuery AJAX calls and uses BBQ plug-in to maintain history and bookmarks.

Current status and future works

Initial productivity of the deletion service was around 500k of deleted files per day (autumn 2010). This productivity is not enough in 2011 (data distribution policy was changed) [3].

After optimization in 2011, the deletion rate estimates 6M of files per day (deletion campaigns – removing large scale of old data), with ordinary deletion rate 2-2,5M of files per day.

Some improvements and extension of functionality should be implemented in the nearest future:

- Improving Deletion monitoring:

  - Aggregation operations took much time - summarizing needed
  - Some plots – not useful now, and some data missed (LFC deletion)
  - Publishing of additional workflow metrics

- Resolving of Jumbo datasets (more 30k of files.. We know one with 5 000 000 of files):

  - High loading for DB backend operation – algorithm should be reviewed (issue with version)
  - 11g validation

- Directory deletion:

  - Current realization does not care about empty directory deletion.
  - Hard to do a recursive delete. At least the latest level. Issue with the naming conventions at site

- Priority in deletion:

  - Some data (test results etc.) should be deleted at first steps.

- SRMless intercommunication with storages

  - Code for storage intercommunication through plugin already in place and used for SRM implementation. Solutions for other interfaces should be adapted and tested.

ATLAS DQ2 Deletion Service was presented at the ATLAS computing & software workshop 12-16 July 2010 [2], International Conference on Computing in High Energy and Nuclear Physics (CHEP) Taipei 18-22 October 2010 [4], ATLAS computing & software workshop 17-21 October 2011 [3] and conference AYSS-2011 [5].

References

[1] Distributed Data Management (DDM). https://twiki.cern.ch /twiki/bin/view/Atlas /DistributedDataManagement

[2] Oleynik D., Petrosyan A. "ATLAS Deletion Service". ATLAS computing&software workshop 12-16 July 2010.

[3] Oleynik D. "Deletion service: Recent progress and improvements"ATLAS computing & software workshop 17-21 October 2011.

[4] Oleynik D., Petrosyan A., Garonne V. "Grid Data Deletion Service (Poster)"International Conference on Computing in High Energy and Nuclear Physics (CHEP) Taipei 18-22 October 2010.

[5] Олейник Д.А. "Централизованный сервис удаления данных для эксперимента ATLAS. ATLAS DQ2 Deletion Service."XV Конференция молодых ученых и специалистов ОМУС 2011. ОИЯИ. Дубна.