FAMoS - an information service on the usage of data files in AliEn

A Abramyan[1,3], L Betev[2], P Buncic[2], C Grigoras[2], A Grigoryan[1], N Manukyan[1,3], M M Pedreira[2] and P Saiz[2]

[1]A.I. Alikhanyan National Science Laboratory Foundation, 2 Brothers Alikhanyan str., 0036 Yerevan Armenia
[2] CERN, CH-1211, Geneve 23 Switzerland

E-mail: aabramya@yerphi.am nmanukya@yerphi.am

Abstract. The File Access Monitoring Service (FAMoS) leverages the information stored in the central AliEn file catalogue, which describes every file in a Unix-like directory structure, as well as metadata on file location and its replicas. In addition, it uses the access information provided by a set of API servers, used by all Grid clients to access the catalogue. The main functions of FAMoS are to sort the file accesses by logical groups, access time, user and storage element. The collected data identifies rarely used groups of files, as well as those with high popularity over different time periods. This information can be further used to optimize file distribution and replication factors, thus increasing the data processing efficiency. The paper describes the FAMoS structure and user interface and presents the results obtained in one year of service operation.

1. Introduction

The aim of ALICE [1], one of 4 Large Hadron Collider (LHC) experiments at CERN, is to explore the primordial state of matter that existed in the first instants of our Universe, immediately after the initial hot Big Bang. The ALICE detector has been built by a collaboration of over 1500 physicists and engineers from 151 Institutes in 37 countries. During data taking, the experiment produces tens of Petabytes of data annually. This data is accessed and analyzed by the members of ALICE Virtual Organization (VO) around the world, using the WLCG Grid infrastructure [2]. The ALICE collaboration deploys ALICE Environment (AliEn) [3] middleware across all Grid sites to facilitate the computing resources access.

The files created in the process of the data processing are stored in Storage Elements (SEs) distributed across the Grid sites. At present there are 82 SEs. Every recorded file and its replicas are annotated in the AliEn File Catalogue (section 2). Since the storage capacities of the sites are limited, the recorded data must be fully monitored for access and relevance, and some of it removed to free space for newly produced files. The FAMoS toolkit (section 3) is build to evaluate the data access and facilitate the removal of sparsely accessed data and replication of popular data sets.

The service functionality can be summarized as follows:
1. Accumulation of file access attributes;
2. Definition of popularity by implementing optimization algorithms, determination of files to be removed, replicated and/or migrated from one site to another;
3. Definition of thresholds for various manipulations of files based on the analysis of information gathered in steps 1 and 2.

This paper describes our work on the 1st step, as well as presents the results obtained in one year of operation of the service (section 4).

2. File System and Data file and file set types of ALICE VO

2.1. File System of ALICE VO

ALICE distributed data storage uses multiple types of storage systems and data transport protocols into one virtual file system called AliEn File Catalogue [4]. Unlike real file systems, the catalogue does not own the files; it only keeps an association between the Logical File Name (LFN) and (possibly more than one) Physical File Names (PFN) on disk or mass storage (tape) systems.

Below we give example of the LFN and PFN of a file named "testfile":

LFN =   */alice/cern.ch/user/n/nmanukya/testfile*

PFN =   *root://pcaliense04.cern.ch:1095//14/22310/e4b050e2…*

protocol   *hostname*   *port*   *path of the file*

Figure1. Structure of LFN and PFN

In this example, the LFN is the address of the file in the catalogue. The PFN string contains the following fields:
- *protocol* to access the file remotely (e.g. gridFTP, root, rfio);
- *hostname* of the SE master node where the file is physically stored, the SE may consist of multiple file servers;
- *port* which is used by the protocol for the communication with SE;
- *path* shows the physical location of the file on the SE file system.

2.2. Data file and file set types in ALICE VO

The following data containers are created and replicated in the processes of the data processing and Monte Carlo simulations:
- ESD (Event Summary Data) contain detailed information about characteristics of the reconstructed collisions. The ESDs are produced during event reconstruction from the raw data.
- For data analysis, the ESDs are filtered to smaller, AOD (Analysis Object Data) data containers.

In ALICE VO each ESD and AOD is stored two or more times on different SEs. The data are grouped in logical periods (categories), each corresponding to a specific set of data taking conditions or MC generation. The data replication is performed by logical periods.

3. FAMoS architecture

FAMoS accumulates file access information from log files of AliEn API and Authen services [5, 6]. The information is sorted by the attributes, shown in table 1.

Table 1. List of file access attributes

Attribute	Description
File name	File LFN
SE name	SE from where the file was accessed
User name	User by whom the file was accessed
Access Time	Time and date of file access
Operation result	Successful or failure

The monitoring is performed not only on the level of the files themselves, but also on the level of their combinations (categories), described in section 2.2.

FAMoS consists of the following components:

- *Collector* Bash script runs on daily basis on the API and Authen servers and uploads the log file of the last day accesses to the FAMoS server.
- *Parser* Perl module running on the FAMoS server and processing the log files of last day; aggregates the file attributes by categories, access time, user, etc. The Parser inserts the processed data into a database, called *Accesses*.
- *Accesses* MySQL database (section 3.1) contains access information with the time granularity of one day.
- *Web interface* (section 3.2) for statistical representation of categories access information, volumes, access time, user and SE.

3.1. The FAMoS database
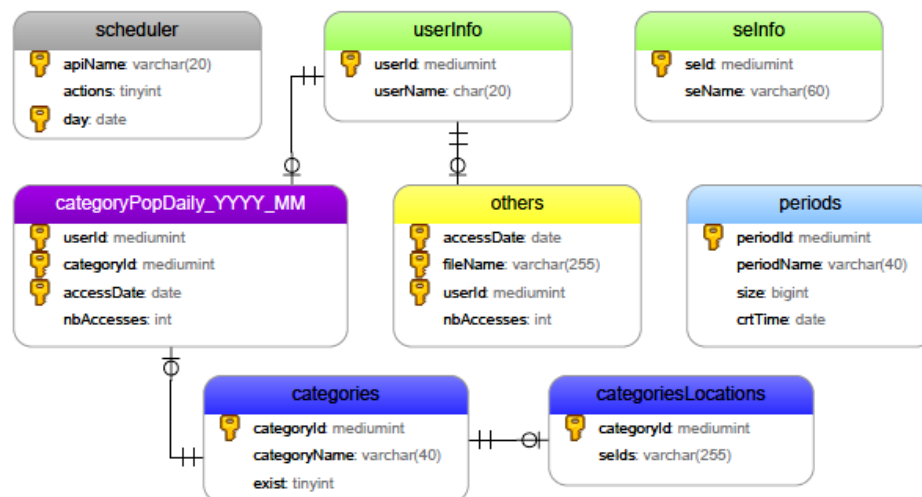The structure of the database is presented in figure 2.

Figure 2. Accesses database structure

The database contains the following tables:

- *scheduler* - scheduling information (start time and state) for the operation of the FAMoS *Parser* module.
- *userInfo* and *seInfo* - values of user name and SE name attributes.
- *categories* and *categoriesLocations* - names of categories (extracted from the LFN), and the list of SE names from where a file of that category was accessed.
- *periods* - names of LHC periods, their volume and creation date. This information is gathered from the AliEn File Catalogue.

- *categoryPopDaily_YYYY_MM* - accesses to the categories. In order to minimize the reading time these tables are created for each month of accesses.
- *others* - access information for those files which cannot be grouped into a category.

3.2. FAMoS Web interface

FAMoS Web interface is an environment for graphical representation of the categories usage statistics as interactive charts and tables. To provide fast load of the web pages the interface design is simplified and page access from any device (from mobile phones to PCs) is done on the base of pure CSS. Currently it consists of the following components:

- *PHP pages* for the dynamic creation of database queries and collecting of their results, mostly in JSON format [7];
- *Highcharts API* [8] for visualization of the FAMoS data in the form of interactive charts;
- Google Chart API [9] for data representation in the form of interactive tables.

4. Statistics obtained in one year of operation of the FAMoS

FAMoS has been deployed and is in operation since August 2013. In this section we present several types of statistics provided by this service.

4.1. Volumes of accessed or not accessed LHC periods

Information on the volumes of used and/or unused periods/categories is critical for optimization of used space on the SEs. For example the plot in figure 3 presents the volumes of the periods which were not accessed during last year and plot in figure 4 presents their volumes grouped by creation year. The unused categories are deleted or archived, freeing space for new data.
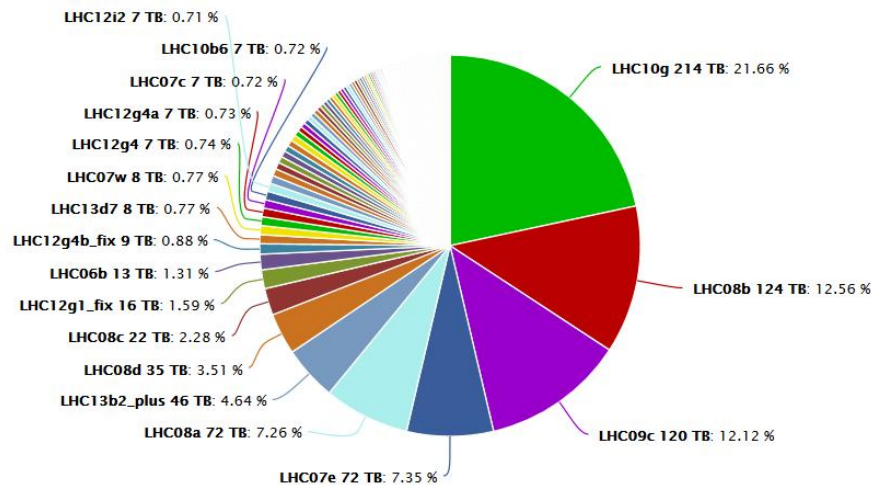

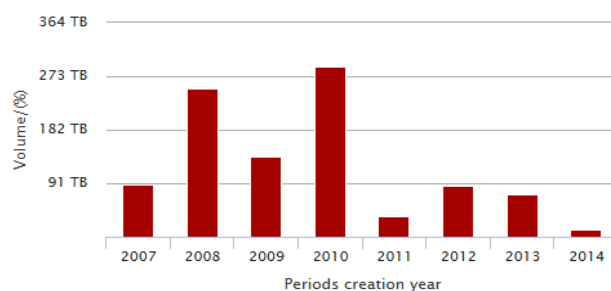Figure 3. Volumes of the not accessed periods

Figure 4. Volumes of not accessed periods grouped by their creation year

4.2. Most and least accessed categories
FAMoS identifies the frequently and rarely used categories over different periods of time. For instance figure 5 illustrates the most and least accessed categories during one year.

Currently we work on the analyses of this data for the precise definition of Popularity Index, by taking into account category volume, access time, number of users, which will be used for the optimization of distribution of the files/categories over SEs.
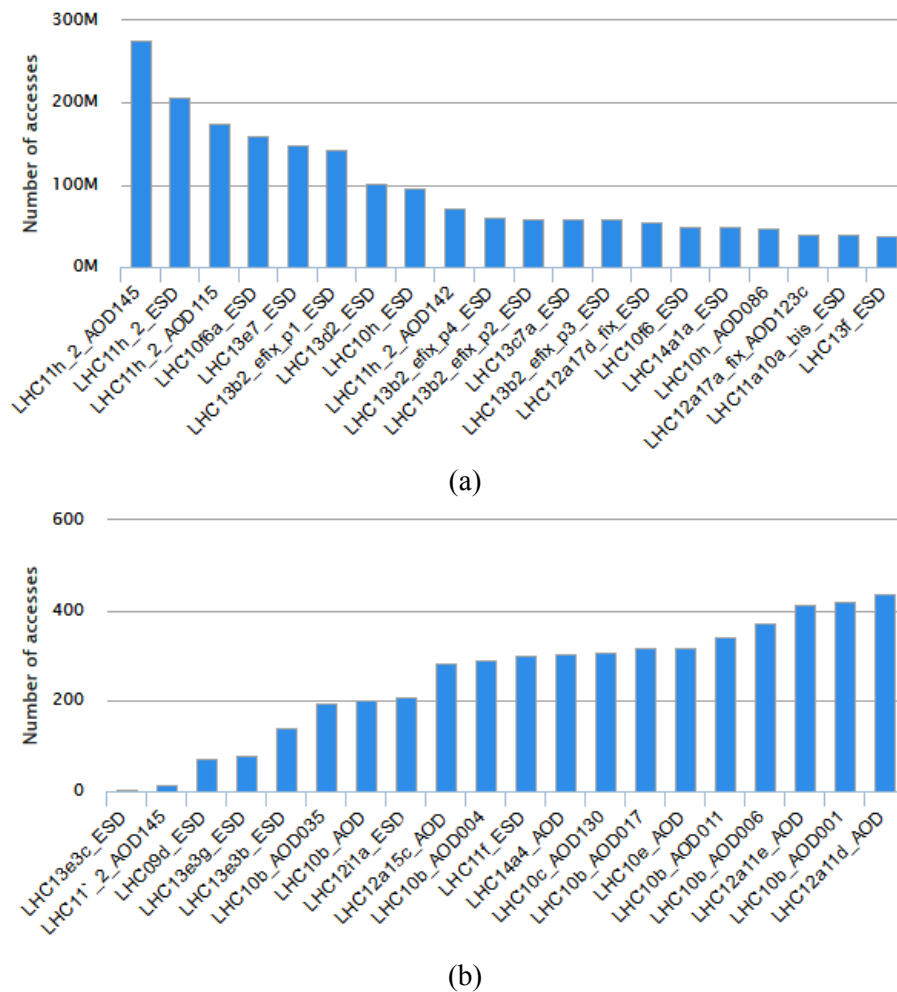

(a)

(b)

Figure 5. 20 most (a) and least (b) accessed categories during one year

4.3. Number of ESD and AOD accesses
FAMoS provides statistics on the usage of the ESD and AOD files over time. Figure 6 shows the number of accesses to the ESD and AOD files, as well as the number of users who have accessed these files during the last year. This information will be used in the definition of Popularity Index for further location and number of replicas optimization.
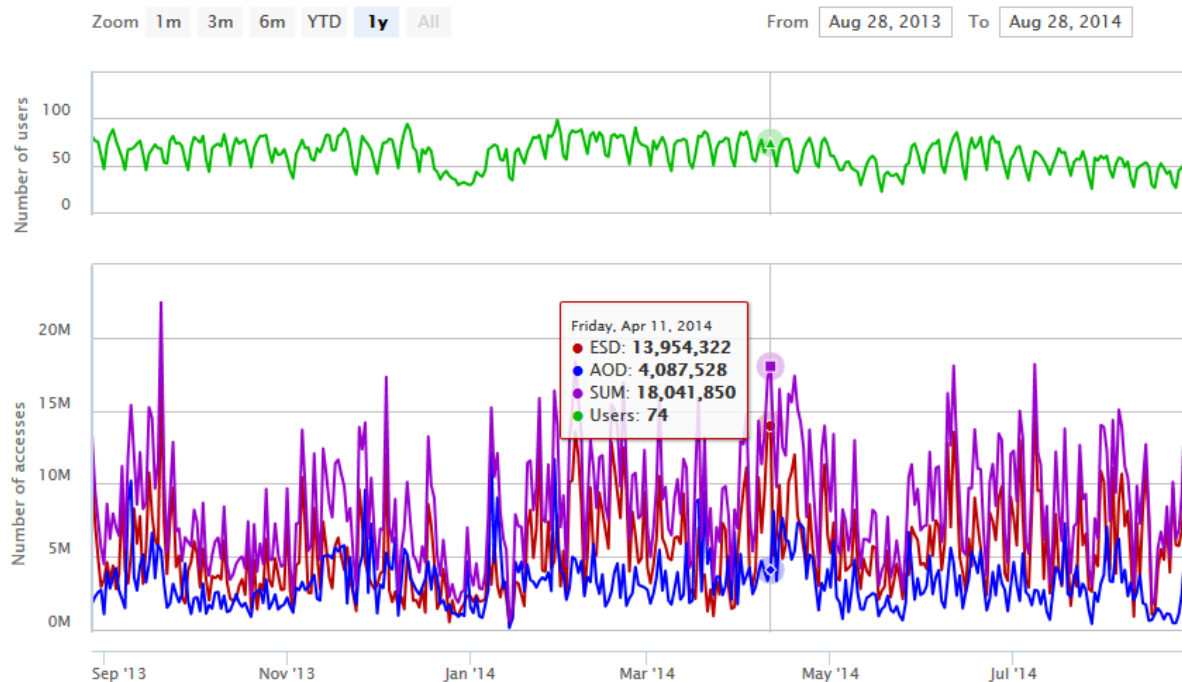
Figure 6. Evaluation of the number of ESD and AOD accesses during last year

5. Conclusions and future work

The File Access Monitoring Service has been deployed and is in operation since August 2013. It provides the information, necessary to optimize the file distribution over the Grid storage elements. The file Popularity Index calculated from the collected access parameters (described above) will determine which files should be removed, replicated and/or migrated from one site to another. The work on the distribution optimization algorithms and file Popularity Index is in progress.

References
[1]   ALICE Collaboration 'The ALICE Experiment at the CERN LHC' J. Instrum. 3, S08002 (2008); Web site URL http://aliweb.cern.ch
[2]   Web site of WLCG infrastructure - http://wlcg-public.web.cern.ch/
[3]   Web site of ALICE Grid Environment (AliEn) - http://alien.cern.ch
[4]   P. Buncic, A. Peters, P.Saiz, The AliEn system, status and perspectives. Computing in High Energy and Nuclear Physics, 24-28 March 2003, La Jolla, California.
[5]   Peters, Andreas-J. et al. – 'AliEnFS: A Linux file system for the AliEn grid services'. eConf C0303241 (2003) THAT005 cs/0306071 [cs-dc] CHEP-2003-THAT005
[6]   J. Zhu et al, Enhancing the AliEn Web Service Authentication. 2011 J. Phys.: Conf. Ser. 331 062048;
[7]   JavaScript Object Notation (JSON) - http://www.json.org/
[8]   Highcharts charting API web site - http://www.highcharts.com/
[9]   Google Chart API - https://developers.google.com/chart/interactive/docs/index