Journal of Physics: Conference Series **513** (2014) 042049 doi:10.1088/1742-6596/513/4/042049

Data federation strategies for ATLAS using XRootD

Robert Gardner[1], Simone Campana[4], Guenter Duckeck[2], Johannes Elmsheuser[2], Andrew Hanushevsky[3], Friedrich G. Hönig[2], Jan Iven[4], Federica Legger[2], Ilija Vukotic[1], Wei Yang[3], on behalf of the ATLAS Collaboration

[1] University of Chicago, Chicago IL, USA
[2] Ludwig-Maximilians-Universität München, Germany
[3] Stanford Linear Accelerator Center, Stanford CA, USA
[4] CERN, Geneva, Switzerland

E-mail: rwg@hep.uchicago.edu

Abstract. In the past year the ATLAS Collaboration accelerated its program to federate data storage resources using an architecture based on XRootD with its attendant redirection and storage integration services. The main goal of the federation is an improvement in the data access experience for the end user while allowing more efficient and intelligent use of computing resources. Along with these advances come integration with existing ATLAS production services (PanDA and its pilot services) and data management services (DQ2, and in the next generation, Rucio). Functional testing of the federation has been integrated into the standard ATLAS and WLCG monitoring frameworks and a dedicated set of tools provides high granularity information on its current and historical usage. We use a federation topology designed to search from the site's local storage outward to its region and to globally distributed storage resources. We describe programmatic testing of various federation access modes including direct access over the wide area network and staging of remote data files to local disk. To support job-brokering decisions, a time-dependent cost-of-data-access matrix is made taking into account network performance and key site performance factors. The system's response to production-scale physics analysis workloads, either from individual end-users or ATLAS analysis services, is discussed.

1. Introduction

The Large Hadron Collider (LHC) concluded its first three-year running period during which the experiments' software and the Worldwide LHC Computing Grid (WLCG) successfully handled the required computing workloads. The next data collection period, due to start in 2015, is expected to bring a significant increase in the amount of data produced with an increase in both luminosity and trigger rate. This, combined with a slow down in the increase of available CPU and disk space, will pose a significant challenge to computing infrastructure. However, a steady increase in the bandwidth capacity of links connecting WLCG sites opens the possibility to modify workloads such that more of the input data are accessed remotely. Thus federating storage resources over wide area networks can mitigate resource shortfalls by reducing the number of copies of data sets required on disk and by enabling access to additional CPU resources. As this approach breaks the long standing computing paradigm - “jobs go to data” – i.e. a tight coupling between CPU and disk, a number of systems had to be made federation-aware. A direct result has been that new applications were made possible and end-user data usage patterns changed.

The ATLAS project to federate data storage – FAX (Federating ATLAS storage using XRootD) has been steadily running for more than one year. The XrootD protocol was chosen as it offers flexible federation topologies, efficiency for wide area direct file accesses, and client-side caching features available in ROOT, which is used in roughly half of all grid-based analysis jobs in ATLAS. Implementation details (how the disparate storage technologies join the federation, authorization and authentication handling, caching possibilities, etc.) have been described in [1]. Here we start with a description of the federation’s current configuration and monitoring systems, followed by a description of changes needed to evolve with the ATLAS data management service. Next we describe a few main use cases as well as the changes in ATLAS production systems that enabled them. The results of two HammerCloud-based [2] tests, one aimed at understanding the system’s reliability and the other at finding the performance envelope are reviewed. The conclusion will present our vision of further expansion, reorganization, and wider integrations.

2. Federation topology

During the last year, the federation went through a significant expansion and a number of changes in its components. As can be seen from Table 1, the federation now covers 42 sites. The increase enlarged the percentage of data accessible through the federation, improved system redundancy, increased the total available bandwidth and made the average distance from an end-user to the closest FAX endpoint shorter (in network round-trip-time, RTT).

Table 1. Number of sites and amount of disk storage currently federated. “T2D” sites are typically larger, have more restrictive reliability requirements, and participate in multi-region processing campaigns.

	Total	Federated	Coverage [%]
Sites (Tier0/Tier1/Tier2D)	1/12/44	1/6/35	100/50/78
Space tokens	523	305	58.3
Total disk space [PB]	277.5	176.5	63.6
Used disk space [PB]	242.7	153.8	63.4

At present, users of ATLAS data can expect an average of 2.4 dataset replicas to be accessible through FAX. From a common data sample used by the Standard Model physics working group consisting of 694 individual datasets, only 24 or 3.5% could not be reached in this way. As shown in figure 1, the numbers are similar across subsets and MC samples.

The federation topology is organized in such a way as to group sites with good network connections (high bandwidth, low RTT latency) under the same redirector and is shown in figure 2. Two top-level redirectors are peers, a regional redirector (GLRD) hosted at Brookhaven Lab and an EU redirector hosted at CERN. A redirector “ASIA”, hosted at Academia Sinica in Taiwan, is a level below and redirects upstream to the GLRD redirector since their best network connection is to the U.S. All of the “national cloud” redirectors are hosted at CERN in order to ease installation and support. The maximum number of redirections possible is 6. To ease integration into other frameworks, the federation topology (endpoint and redirector names and their relations, addresses, and status) is encoded in AGIS [3] from where it can be obtained via a RESTful API.

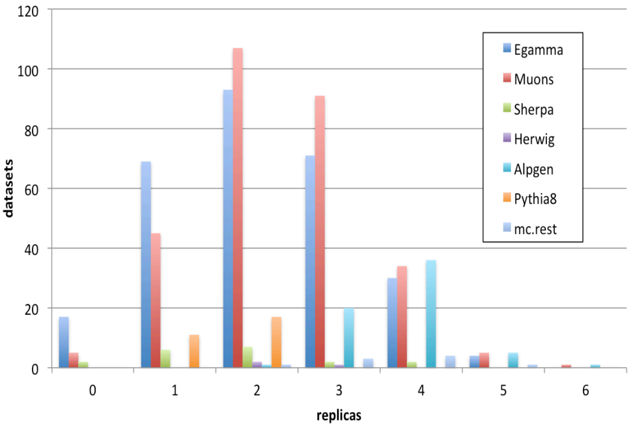

Figure 1. Number of dataset replicas accessible through the federation for real data and MC samples.

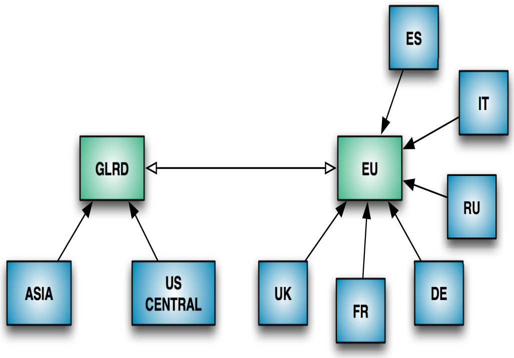

Figure 2. Topology of the regional redirection network in the U.S., Asia, and European regions.

The ATLAS global name space relies on the LCG File Catalog (LFC) [4] to translate global logical file names (gLFNs) to local physical file names . Depending on the storage technology deployed, sites use one of three name-to-name (N2N) plugins that perform queries and locally cache results in order to reduce load on the LFC catalog. There are two LFC catalog instances for ATLAS data: one at BNL covering ATLAS sites in the U.S. and the other at CERN covering all the rest. Due to query result caching by an Oracle database (on which the LFC runs), individual translations were always in the sub-second time range.

ATLAS data on the grid is managed by the Distributed Data Management (DDM) system called Don Quijote 2 (DQ2) [5]. During the long shutdown DDM will transition from DQ2 to a new service Rucio [6] – a system designed to ensure system scalability, reduce operational overhead and support new use cases. This will present a major simplification and performance improvement in FAX operations as LFC access will no longer be needed. With Rucio it becomes possible for a N2N plugin for a given gLFN to obtain all the possible corresponding physical file names without a remote catalog lookup. The translation is not necessarily one-to-one as the file can exist in multiple locations ("space token" directories) at the site. Although there are only a few of these per site, the number of file system calls made by the N2N plugin is reduced by ordering them according to their probability to be successful. The N2N plugins obtain path prefixes for all of the space tokens from AGIS. Tools to obtain gLFNs in Rucio format have been made, new N2N plugins supporting both versions of gLFNs have been rolled out, and all of the files at a number of sites have already been renamed. While still in a testing phase, we expect a smooth transition and an improvement in scalability and reliability of the federation when complete.

3. Monitoring the federation

An essential part of the federation is its monitoring system. Based on the XRootD monitoring protocol [7], it provides two streams of information of different granularity levels. A summary stream consists of input/output data rates, number of connections, authentications, redirection, etc., thus providing a way to centrally monitor federation functionality. A detailed stream consists of several sub-streams and gives information on files, file operations (open, close, read), users, applications, etc. Data mining of the collected detailed monitoring data provides insight into patterns of usage of the storage resources and provides feedback needed for optimization of applications, data placement strategies, and data storage.

The schema shown in figure 3 illustrates the flow of monitoring information. There are essentially three ways storage systems are connected to the federation: a native XRootD storage service with or without a proxy server, an XRootD server providing a proxy service to a backend filesystem (such as a POSIX, DPM [8], or dCache [9] filesystem), or storage systems which internally implement the XRootD protocol. In the first two cases, monitoring comes directly from the XRootD server, while the third case is more complex as each storage technology collects the information differently. For example, a number of dCache sites join FAX by exposing XRootD doors to the backend system via a special “xrootd4j” plugin. In order to collect monitoring information from these sites, we developed a dCache plugin that intercepts all messages passed between the client and the XRootD door. The plugin collects monitoring information and sends it out according to the XRootD monitoring specification.

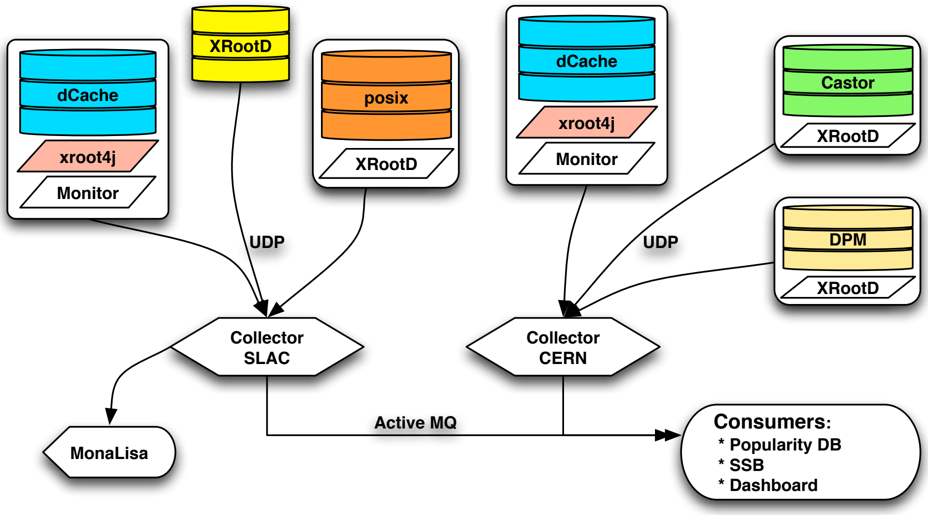

Figure 3. Schema of the FAX monitoring system.

To reduce overhead on the servers, both summary and detailed monitoring streams are sent using the non-blocking stateless protocol UDP. To improve collection reliability, the monitoring system has two collectors: one at Stanford (SLAC) and the other at CERN. The collectors (developed for CMS federation project [10]) provide visualizations of both detailed (GLED [11]) and summary (based on MonALISA [12]) streams. The integrated information is sent to an Oracle DB at CERN serving as a backend to several web based visualization interfaces. Figure 4 was produced using one of them (WLCG dashboard) and shows that the total throughput of the federation, while relatively low in volume, sometimes exceeds 10% of the standard FTS [13] traffic, the WLCG file transfer service used by DQ2 and Rucio.

4. Functional testing

To ensure high availability of the federation’s endpoints and redirectors, a suite of tests are run at one-hour intervals. The main tests are: functioning of direct access to site unique files; that a site properly redirects upstream on request for a file not existing at that site; that a file can be obtained on request from an upstream redirector; and that the grid (X509) authentication works. Test results are sent to the Site Status Board (SSB [14]) which provides visualization and an archive. Automatic notifications are sent to regional cloud support mailing lists when one of the tests fails for more than a certain predefined time interval.

Another testing chain (shown in figure 5) measures transfer rates grid jobs might expect from all of the FAX endpoints. The chain starts with jobs submitted from the HammerCloud (HC) system to the analysis queues of the 40 largest ATLAS computing sites. Each job copies a 100 MB file from each of the FAX endpoints in parallel, sends the resulting transfer rate to the SSB, and then sleeps a predefined time interval before starting a new round of tests. The SSB collected information is then filtered, averaged, and stored in a different table together with measurements obtained from all of the FTS transfers and perfSonar [15] throughput measurements. Finally all of this information is used to calculate a so called “Cost Matrix” that is delivered to the PanDA workload management system [16].

From one thousand links currently monitored, average rates for these particular tests of <1MB/s, 1-10 MB/s and >10 MB/s are observed at 13%, 61%, and 25% of links, respectively.

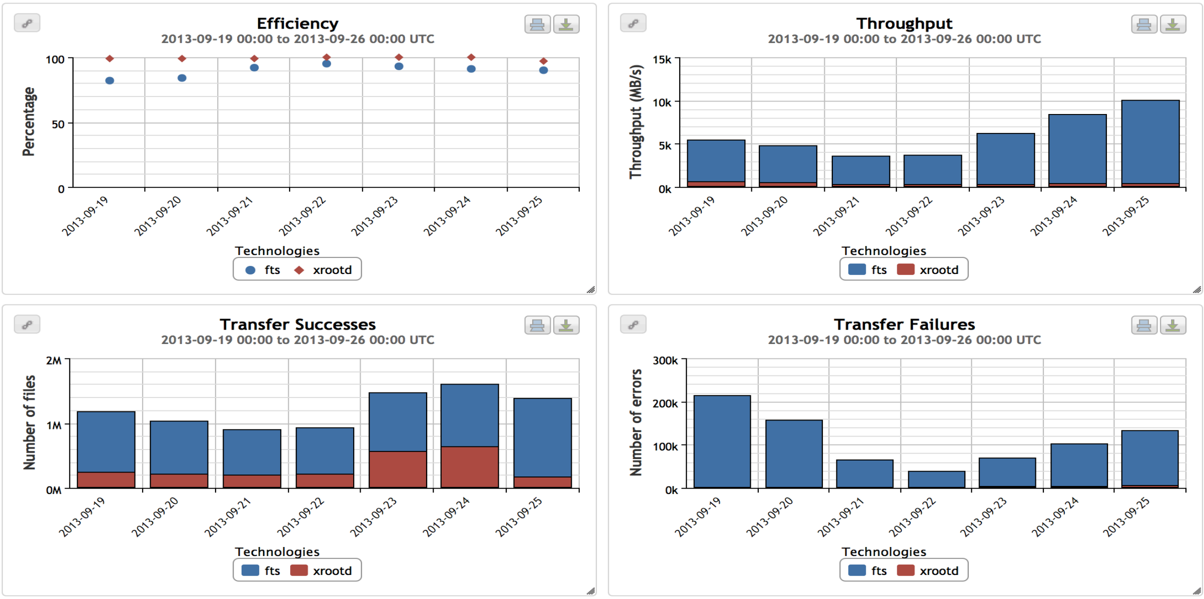

Figure 4. WLCG dashboard view of traffic through FTS and FAX. Only remote traffic is shown. While still smaller (in average 1.5GB/s), federation traffic is not negligible and shows high efficiency.

5. Using the federation

Direct access via FAX to most ATLAS datasets is readily available to end-users, while several other FAX use cases require substantial development in both integration with existing services and in some cases developing new ones.

5.1. Interactive and batch processing by end-users

Since its inception, users have been able to use FAX to directly access data (both interactively and in batch processing) from a variety of computing resources (PROOF clusters, Tier 3 clusters, opportunistic grid computing facilities). A significant use of FAX has come from co-located or nearby Tier 3 users. The data transfer request interface (DaTRI [17]) is used to request dataset replication to nearby FAX-enabled Tier 2 or Tier 1 centers, and Tier 3 disk space is used only for output. Typically a few hundred Tier 3 jobs results in a few GB/s of throughput, underscoring the importance and utility of high capacity wide area network links.

5.2. PanDA failover

In the event of two unsuccessful attempts to open a file from local storage, a PanDA pilot job (a generic script submitted to the site which receives the actual job payload) will try to copy it from remote sites in the federation instead. Enabling this failover mode is done per PanDA queue and is controlled by the AGIS information service. Currently, approximately 30 production and analysis queues are enabled to use this failover functionality. Each job that tries to failover to FAX reports the number of files accessed, file sizes, and data transfer rates to a monitoring service. The PanDA job queues are already highly efficient and thus the failover rates are quite low.

5.3. Scheduling jobs against FAX - PanDA overflow

It is often the case, especially with user analysis jobs, that an input dataset exists at only one site. This can result in very long job queues at a particular site. PanDA can solve the problem by submitting the job to a different site where data access would be provided by the federation. Overflow decisions will be based on the following algorithm: for each job that would normally end up waiting in a queue, PanDA will estimate the expected time a job would start. Transfer times are estimated for each queue with idle job slots using data from the Cost Matrix. An overflow scheduling decision is then based on a comparison of the queue wait time and the file transfer times. Most of infrastructure to support this use case is ready and we expect the first overflow jobs to be submitted in the near future.

5.4. FAX enabled applications

Having access to the majority of ATLAS data at once raises the possibility for new types of services. As a proof of principle, prototypes for an *Event Picking Service* and a *Skim Slim Service* have been developed. These services are designed to relieve end-users from optimizing their code for PB-scale data, and to provide a faster turn-around experience as resources can be scheduled more efficiently by the service while using expert versions of the skimming code optimized for WAN access.

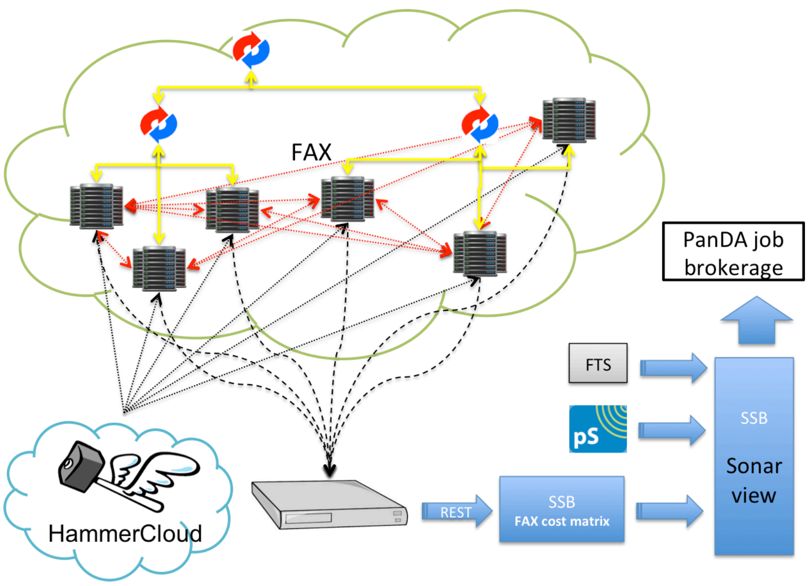

Figure 5. Measuring rates between job queues and FAX endpoints. Jobs submitted by HammerCloud tasks measure the interlinks and report values to the SSB. Information is filtered and averaged before sending to PanDA (standalone or in combination with FTS and perfSonar measurements).

6. Performance measurements

The federation performance depends on a number of factors: the available bandwidth of WAN links between a given CPU and data source; the load on disk servers; traffic on the local area network (LAN); the load on individual FAX endpoints; and the load on the LFC databases. All of these naturally have a strong time dependence, changing dynamically with workloads in the system. The best measures of how much these factors influence system usability are:

- *Probability of finding a file known to exist in the system* – when a query for a known, unique file takes longer than a pre-defined time, the user will be given a "file not found" error. This is observed to occur in less than 1% of all file accesses under normal running conditions.
- *Probability of file transfer failure* – recent HammerCloud stress tests show this to be less than 1%.

To measure rates expected from a set of endpoints under high load, we needed a large number of clients simultaneously accessing each. We used HammerCloud functional tests to submit up to 50 concurrent jobs reading a specific set of test files or randomly selected ntuples. Rates observed in individual transfers were often not dependent on the number of concurrent transfers (a subset of measurements is shown in figure 6). By comparing rates obtained from FTS transfers, and FAX Cost Matrix and perfSonar measurements, we find that the WAN is often not the rate limiting factor but more likely the LAN congestion at the source sites.

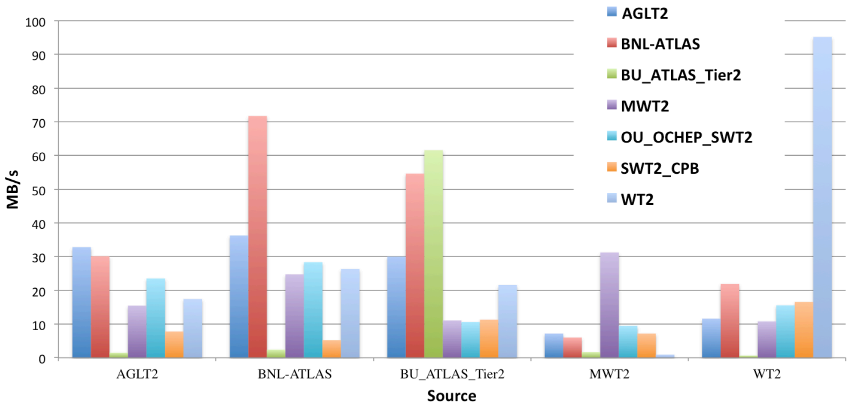

Figure 6. Rates observed over sub-sample of US links under loads of up to 50 concurrent transfers. The small differences between internal and external rates point to high load on disk servers or the LAN.

7. Conclusions and future work

In the past year, the FAX project was extended in the number of sites and now covers a large fraction of all available ATLAS storage resources. Further expansion will improve on all its performance indicators. The federation provides a powerful tool to be used for physics analysis while reducing failure rates and improving turn-around time for both analysis and production grid jobs. Replacing LFC catalog lookups with a local lookup algorithm using the Rucio hash-based naming convention removes the last potential bottleneck. Finalizing integration of Cost Matrix measurements in PanDA will enable the PanDA overflow use case. Two applications relying on FAX for fast access to most of the ATLAS data are in development. Future developments include testing and deployment of file caching services, further integration of monitoring and reporting into the usual production environment, and fine-tuning of operational parameters.

8. Acknowledgements

The federated storage project involves the work of many projects and computing infrastructure providers of the WLCG. We gratefully acknowledge the support of the National Science Foundation (grant PHY-1119200), and the technical staffs of the participating institutes.

References

[1] L Bauerdick *et al* 2012 *J. Phys.: Conf. Ser.* **396** 042009
[2] Daniel C van der Ster *et al* 2011 *J. Phys.: Conf. Ser.* **331** 072036
[3] A Anisenkov *et al* 2012 *J. Phys.: Conf. Ser.* **396** 032006
[4] J P Baud, S Lemaitre 2005 CERN http://hepix.fzk.de/upload/ lectures/LCG-File-Catalog-HEPIX-2005-1.pdf
[5] M Branco *et al* 2008 *J. Phys.: Conf. Ser.* **119** 062017
[6] V Garonne *et al* 2012 *J. Phys.: Conf. Ser.* **396** 032045
[7] A Hanushevsky 2013 XRootD System Monitoring Reference, retrieved from http://XRootD.slac.stanford.edu/doc/prod/xrd_monitoring.pdf
[8] DPM – Disk Pool Manager https://svnweb.cern.ch/trac/lcgdm/wiki/Dpm
[9] dCache http://www.dcache.org
[10] CMS Xrootd Architecture, https://twiki.cern.ch/twiki/bin/view/Main/CmsXrootdArchitecture
[11] M Tadel 2004 *Applied Parallel and Distributed Computing* ISBN 1-59454-174-4
[12] H B Newman, I C Legrand, P Galvez, R Voicu, C Cirstoiu 2003 **cs/0306096**
[13] Frohner A et al. 2010 Data management in EGEE *J. Phys.: Conf. Ser.* 219 062012
[14] J Andreeva *et al* 2012 *J. Phys.: Conf. Ser.* **396** 032072
[15] A Hanemann *et al* 2005 Springer Verlag, LNCS 3826, pp. 241–254, ACM
[16] T Maeno 2008 *J. Phys.: Conf. Ser.* **119** 062036
[17] The Data Transfer Request Interface (DaTRI) http://panda.cern.ch/?mode=ddm_req