

EGI Dataset Accounting and the WLCG

Adrian Coveney^{1*}, and Greg Corbett¹

¹Science and Technology Facilities Council, Scientific Computing, Didcot, United Kingdom

Abstract. This paper reviews the prototype dataset accounting developed during the EGI-Engage project and how it could be used to complement the view that the WLCG has of its datasets. This is a new feature of the EGI resource accounting system that will enable storing information on dataset usage such as who has accessed a dataset and how often. The new REST interface used for retrieving usage metrics from the EGI DataHub is described as well as further work that is required.

1 Introduction

While the Worldwide LHC Computing Grid[†] (WLCG) and EGI[‡] have both made significant progress towards solutions for storage space accounting, one area that is still quite exploratory is that of dataset accounting. This type of accounting would enable resource centre and research community administrators to report on dataset usage to the data owners, data providers, and funding agencies. Eventually decisions could be made about the location and storage of datasets to make more efficient use of the infrastructure. By giving insight into data usage, dataset accounting also assists scientists in assessing the impact of their work.

This paper reviews the status of the prototype dataset accounting developed during the EGI-Engage project and how it could be used to complement the view that the WLCG has of its datasets. This is a new feature of the EGI resource accounting system that will enable storing information on dataset usage such as who has accessed a dataset and how often, the transfer volumes, and end points etc. The design of this new feature has been led by the users' requirements collected in the first part of the project, from which a set of dataset accounting metrics were derived. In these trials, the EGI Accounting Repository was integrated with the data provider Onedata[§] (the underlying technology powering the EGI Open Data Platform and EGI DataHub) as an example of a generic data provider.

2 Existing Resource Usage Accounting

APEL [1, 2] is a compute resource usage accounting tool that collects data on the resources, such as CPU time and storage space, used by research communities. It supports two main use cases. The first is that APEL is the Accounting Repository for EGI. This means that

* Corresponding author: apel-admins@stfc.ac.uk

† <http://wlcg.web.cern.ch/>

‡ <https://www.egi.eu/>

§ <https://onedata.org>

every EGI site sends all their usage data to APEL. The second is that some international virtual organisations (VO), such as the LHC experiments, that use multiple e-infrastructures choose to send all of their usage data to APEL so that they have a unified view of all their worldwide usage. Therefore, APEL aggregates data from sites participating in the EGI and WLCG e-infrastructures as well as from sites belonging to other collaborating e-infrastructure organisations, including the Open Science Grid** (OSG) and NorduGrid††. The information is gathered from different collectors in the form of text-based accounting records into a central accounting repository where it is then processed to generate statistical summaries that are made available through the web-based EGI Accounting Portal. As well as the tabular and graphical summaries provided by the Portal, data can be downloaded from the Portal in JSON or CSV format.

APEL currently collects accounting information for grid batch, cloud VM, and storage resources. Typically, a site will deploy some form of accounting collector, which will interact with the underlying resource system and produce an accounting record in a standardised format. The individual records are then accumulated into messages‡‡, usually a thousand at a time, and are sent to the APEL central repository via the EGI message broker network using the Secure STOMP Messenger (SSM) tool [3]. SSM is a component of the APEL software suite that sends these accumulated messages to the EGI Message Brokers so that they can be pulled in by the central APEL accounting server using its own SSM client. The EGI Message Broker Network is a small number of distributed servers (for load balancing and redundancy) which handle asynchronous reliable message passing between EGI operational or user services. Message flows can be one to one, many to one, or one to many, encrypted or not. APEL uses this network to send usage records from sites to the central APEL repository and aggregated summaries from the repository to the EGI Accounting Portal.

APEL provides data collectors but others can be, and have been, written based on a common accounting record, the Open Grid Forum Usage Record [4]. However, APEL is agnostic to the exact source of accounting data, so some regions have set up regional APEL servers that receive the accounting data from national sites before sending a copy of the information on to the central accounting server. A similar model exists for OSG, whereby sites participating in that infrastructure send data to an OSG accounting server, which then sends a copy of the accounting information for selected VOs to the APEL central repository. Figure 1 shows an overview of the current APEL accounting system as described above.

** <https://opensciencegrid.org/>

†† <http://www.nordugrid.org/>

‡‡ <https://wiki.egi.eu/wiki/APEL/MessageFormat>

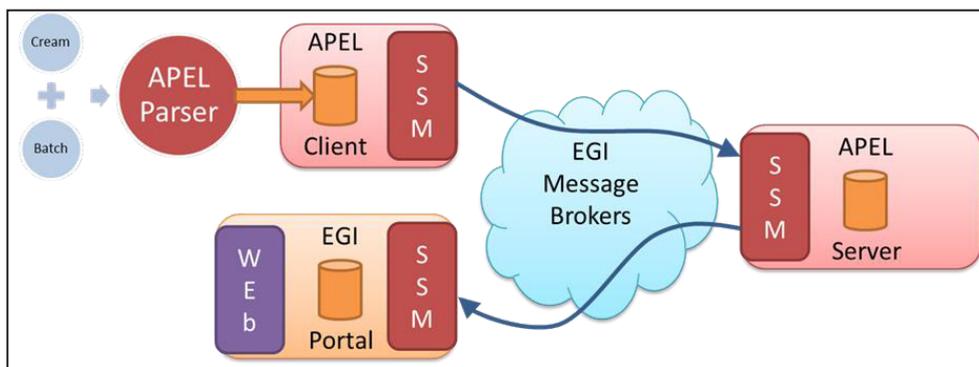


Fig. 1. The current accounting system showing an example data flow from a grid batch system through to the web portal provide by the EGI Accounting Portal. All “APEL” and “SSM” components were developed by the APEL team.

3 Dataset Usage Accounting

3.1 Definitions

The definition of dataset accounting used as part of this prototyping work is a logical set of files which may exist in several places at once but to which it is possible to assign some form of unique persistent identifier (PID). This was used to ensure that the definition was clearly distinct from storage space accounting, which accounts for disk allocation and usage without concern over how it is used. Additionally, for accounting purposes, it was necessary to assume that this unique identifier is available so that multiple records relating to the same dataset could be collated.

There is also a difference between performing file-based accounting and uniquely identified dataset accounting. In the area of data accounting, WLCG is mainly concerned with the former with the intention of optimising the use of storage space by minimising the storage of data that is infrequently read. To do this, WLCG produces reports on a per file basis about how often they have been accessed in periods of particular length.

The details of creating and assigning PIDs were left to the dataset storage system. For the purposes of accounting, it is sufficient that one is available for each dataset that needs to be accounted for and that it is included with the usage statistics when they are collected. It is up to the sites submitting accounting data to ensure that the PIDs they report are correct if they want to have useful accounting information that is correctly collated in the Accounting Repository.

3.2 Objectives

The work done as part of the EGI-Engage project aimed to prototype a system that would enable infrastructure managers to make more efficient use of the infrastructure and enable scientists to assess the impact of their work. In a distributed environment, there can be multiple copies of datasets or files within a dataset so having the right information can allow better replication decisions to be made – i.e. moving the data closer to where it is used and providing sufficient copies.,. One way to assess the impact of a dataset is by measuring how many times datasets are reused and thus how popular they are. This was to be achieved by collecting metrics such as number of access events for each PID.

3.3 Preliminary work

The accounting team published a questionnaire to gather feedback from stakeholders on how best to implement a prototype system. In addition, communities that expressed great interest in this activity were selected for interviews to clarify their needs. Considering the need, identified in our preliminary analysis, for a persistent identifier (PID) management system as a prerequisite to implementing a data accounting feature, special attention was devoted to gathering information about the usage of different methods for identifying datasets including Digital Object Identifiers (DOI) [5], Uniform Resource Identifiers (URI) and persistent Uniform Resource Locators (URL).

The EGI Open Data Platform seemed to be the best candidate in this set and was investigated as a source of dataset accounting data. The EGI Open Data Platform is a collection of technologies that enable the publishing, use and reuse openly accessible data identified by PID. It uses Onedata as the end-user service exposing these technologies, which is a global data management system providing access to distributed storage resources and it handles the creation and registering of PIDs required by users. The planned design of this platform included the integration of current EGI storage services into the platform backend, which should help to hide the complexity caused by the wide range of storage systems adopted.

The survey identified the most important attributes needed for meaningful dataset accounting as

- how often a dataset is accessed,
- who accessed them, and
- what transfers of the dataset occurred.

Other high priority data fields that should be included are as follows: the different forms of user identification, such as the Distinguished Name (DN) from an x.509 certificate, widely used within grid computing, or an eduPersonPrincipleName (ePPN) attribute from a federated identity provider; user groupings such as VO, or home institution; number of store and retrieve operations; number of files transferred; success or failure of the transfer; and the dataset identifier.

Other, medium priority data fields which should probably be accounted for include: storage system implementation, i.e. the type of storage system this data was extracted from; transfer start time and end time or duration; the source and destination IP address; and the volume of data transferred.

4 EGI-Engage Prototype

The prototyping work focused on getting basic metrics out of Onedata and integrating that with the APEL accounting system. A new record format was designed as an extension to the XML-based Open Grid Forum Usage Record version 2 to enable dataset usage to be collected.

4.1 Transferring the data

Under the current model, a provider-side parser creates APEL accounting records in a message format suitable for transmission. These records are then sent by a provider-side instance of the APEL Secure Stomp Messenger (SSM) to the EGI Message Broker Network. A server-side SSM then pulls down the data from the Message Broker Network, where it is loaded into the APEL server, summarised and finally these summaries are sent to the EGI Accounting Portal, again via SSM and the Message Broker.

For dataset usage accounting, the SSM component was updated to add support for retrieving records from a REpresentational State Transfer (REST) HTTP interface. This means that the SSM instance on the APEL server can pull dataset accounting records directly from storage systems that expose these records via a REST interface. This new architecture, which runs alongside the existing accounting infrastructure, is shown in Figure 2. This allowed the APEL software to integrate with the EGI DataHub, which only exposes a REST interface for obtaining accounting metrics and so does not require the use of a message broker. Transfers are performed over HTTPS to ensure the integrity and security of the data.

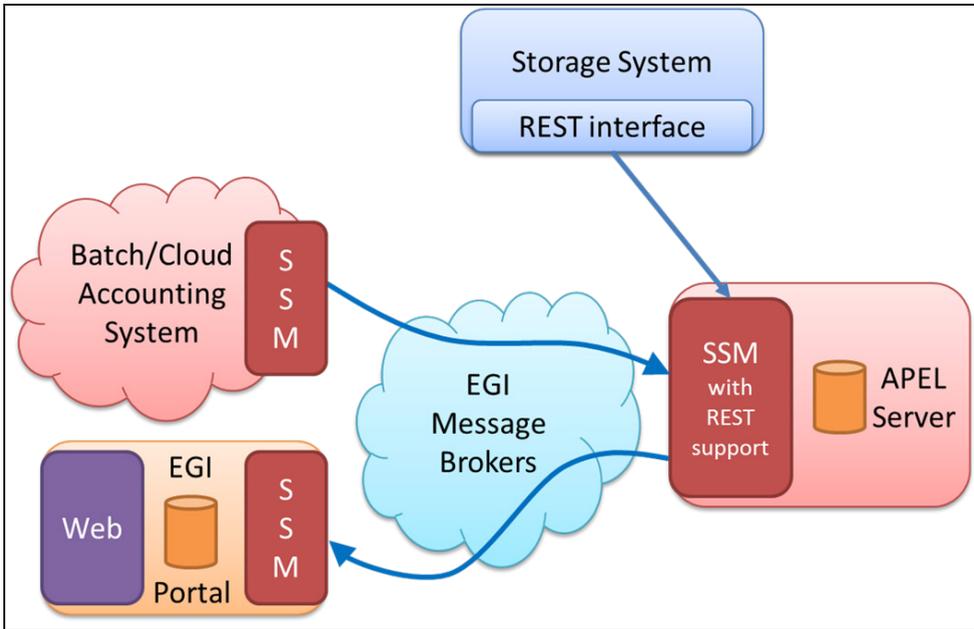


Fig. 2. The updated accounting system. In this case, the Storage System is Onedata. The SSM on the APEL server has been modified to add REST support. Existing accounting batch and cloud accounting sources shown in a simplified manner.

4.2 Record Format Evolution

During the period that the prototype was being developed, the metrics from Onedata API were limited to only those metrics that the resource provider actually recorded but it was still possible to iterate the design to a certain extent.

The record format that was introduced specifically for this project changed over the course of development. It was realised that the count of access events could be split into separate events of how many times the dataset was read and written to. Support for different types of data set ID (non-DOI unique IDs, such as PIDs) was added as well as a way to distinguish which data management system a record had come from. Table 1 shows a summary of the metrics and what block they were included in the XML-based record format.

Table 1 Metrics used in the final prototype.

<i>Block</i>	<i>Key</i>	<i>Type</i>	<i>Description</i>
Record Identity Block	Resource provider	string	Resource provider at which the resource is located
	Infrastructure	string	High level system that the provider is part of (e.g. https://datahub.egi.eu)
Subject Identity Block	GlobalUserID	string	e.g. X.509 certificate DN
	GlobalGroupId	string	e.g. VO
	GlobalGroupAttribute	string	e.g. VO Group and/or Role
	ORCID	string	ORCID iD of the user
Dataset Usage Block	DatasetID	string	Unique identifier such as a PID / DOI
	DataSetIDType	string	“PID”, “DOI”, etc.
	ReadAccessEvents	integer	Number of read operations in period
	WriteAccessEvents	integer	Number of write operations in period
	Source	IP address / other	Source of transfer at resource provider
	Destination	IP address / other	Destination of transfer
	StartTime	ISO 8601 timestamp	Start time of transfer
	Duration	ISO 8601 duration	Duration of transfer
	EndTime	ISO 8601 timestamp	End time of transfer
	TransferSize	integer	Bytes transferred
	HostType	string	Storage system Type
	FileCount	integer	Number of files accessed
	Status	string	Success / failure / partial transfer

5 Conclusion

A basic dataset accounting prototype was developed that introduced a new route (REST APIs) for retrieving metrics to the APEL software system, allowing accounting data to be aggregated from a wider range of types of resource. This is beneficial as not every resource system that may want to be integrated with APEL in the future will be able to push data to a message broker and may already provide REST interfaces that the data can be retrieved from. This experience with REST APIs is being applied in other projects that APEL is involved in.

Further development will be needed as part of EOSC-hub as Onedata matures. Additionally, there is a need to investigate integration with other storage systems that serve as sources of dataset usage metrics, across both EGI and EUDAT^{§§} infrastructures (as part of an overarching EOSC-hub objective). Dataset accounting is also currently being

^{§§} <https://www.eudat.eu/>

explored as part of the SeaDataCloud^{***} project. Finally, there needs to be development work to provide a unified view on datasets across these different infrastructures.

This work is co-funded by the EOSC-hub project (Horizon 2020) under Grant number 777536.

References

1. M. Jiang, C. Del Cano Novales, G. Mathieu, J. Casson, W. Rogers, J. Gordon (2011) *An APEL Tool Based CPU Usage Accounting Infrastructure for Large Scale Computing Grids*. In: Data Driven e-Science. Springer, New York, NY
2. A. Coveney, G. Corbett, W. Rogers, et al (2018) apel/apel: APEL 1.7.1-1. doi: 10.5281/zenodo.1694524
3. A. Coveney, G. Corbett, W. Rogers (2018) apel/ssm: SSM 2.3.0-2. doi: 10.5281/zenodo.1642101
4. A. Cristofori, J. K. Nilsen, J. Gordon, M. Jones, J. A. Kennedy, R. Müller-Pfefferkorn (2013) *Usage Record - Format Recommendation* GFD.204 In: Open Grid Forum Recommendation
5. International Organization for Standardization (2012) *ISO 26324:2012: Information and documentation — Digital object identifier system* Retrieved from: <https://www.iso.org/obp/ui/#iso:std:iso:26324>

^{***} <https://www.seadatanet.org/About-us/SeaDataCloud>