LHCb distributed computing operations

F Stagni, R Santinelli, M Cattaneo, S Roiser
PH Department, CH-1211 Geneva 23 Switzerland
E-mail: federico.stagni@cern.ch, roberto.santinelli@cern.ch, marco.cattaneo@cern.ch, stefan.roiser@cern.ch

Abstract. In the Grid world, there are many tools for monitoring both activities and infrastructure. The huge amount of information available needs to be well organized, especially considering the pressing need for prompt reaction in case of problems impacting the activities of a large Virtual Organization. Such activities include data taking, data reconstruction, data reprocessing and user analysis. The monitoring system for the LHCb Grid Computing relies on many heterogeneous and independent sources of information. These offers different views for a better understanding of problems, while an operations team follow defined procedures that have been put in place to handle them. This work summarizes the state-of-the-art of LHCb Grid operations, emphasizing the reasons that brought to various choices, and what are the tools currently in use to run our daily activities. We highlight the most common problems experienced across years of activities on the WLCG infrastructure, the services with their criticality, the procedures in place, the relevant metrics, the tools available and the ones still missing.

1. Introduction

This paper presents how LHCb organized its distributed computing operations, in light of the resources, the tools and the activities. The reader is expected to be familiar with the Tier levels of the MONARC model [1]. This paper is organized as follows: section 2 presents the LHCb distributed computing activities. Section 3 presents the organization given to the operations team. Section 4 introduces the tools used for monitoring, management, reporting and testing. Conclusions are given in section 5.

2. Distributed Computing Activities

The LHCb Distributed Computing Activities are based on the experiment Computing Model [2]. Even if the computing model has been recently subject to some relaxation, in its fundamentals it is still in use. The original computing model was to use Tier 2 for simulations activities, and Tier 0 and Tier 1 for real data manipulations: today, some Tier 2 are actively used for part of our real data reprocessing activity. Apart from few exceptions, computing resources are mainly provided by WLCG [3].

The distributed computing activities can be split between data manipulation, handled in the following section, and data management, explained in section 2.2.

2.1. Data Manipulation

We split Data Manipulation activities into production and non-production activities. Production activities include:

- **Prompt Data Reconstruction**: data coming directly from the experiment is distributed, as *RAW* files, to the six Tier 1 centers participating to the LHCb Grid, and reconstructed. The model dictates that a copy of each RAW file is stored at CERN, which is the Tier 0, and a replica is at a Tier 1. The distribution of the RAW files and their subsequent reconstruction is driven by shares, agreed between the experiment and the resource providers.
- **Data Reprocessing**: this activity is done from time to time, usually at the end of a data talking period. From a workflow point of view, it is identical to the data reconstruction process. Its objective is to provide a second reconstruction of already reconstructed RAW files. Data Reprocessing is usually a highly-demanding activity, in terms of CPU and storage usage. For this reason, the used resources can be different from the ones used in normal data reconstruction. For the reprocessing of 2011 data, some Tier 2 have been used in addition to Tier 0 and Tier 1.
- **Data Stripping**: stripping is the name LHCb gave to the selection of interesting events. Data Stripping comes just after Data Reconstruction or Data Reprocessing, and it is usually a much lighter activity with respect to computing needs. It's anyway stressing for the storage resources, due to the high I/O involved. When reconstructed data is stripped, a number of *streams* are created, each containing a selection of physics events.
- **Calibration**: from the operational point of view, calibration is a lightweight activity that follows the same workflow of Data Stripping, thus starts from already reconstructed data. Sometimes, though, calibration activities take a very high priority and should be done as fast as possible.
- **Monte Carlo (MC) simulations**: the simulation of real collisions started way before real data recording, and it is still a dominant activity for what regards purely the number of jobs handled by the system. Even if simulated events are continuosly produced, large MC campaigns are commonly scheduled during longer shut-down periods, when LHCb is not recording collisions. Since Tier 0 and Tier 1 centers are for most of the time busy with real data processing, Tier 2 centers represent the main resource when coming to production of simulated events. This does not exclude using Tiers 0 and 1 resources.

Some of these production activities also include a merging phase: for example, streamed data that is produced by stripping jobs is usually stored into small files. Thus, it is convenient to merge them into larger files that are then stored in the Grid Storage Elements. While merging is computationally a light activity, it can very well represent a stress in terms of usage of storage resources, especially when the number of files to merge is large.

Non-production activities include:

- **User data analysis**: user analysis jobs represent a good fraction of the total amount of jobs handled by the system. User jobs, when compared to production jobs, are usually less demanding in terms of computing and storage. User jobs have a priority over production jobs, and when a large number of them are queued, they may represent an operational burden, slowing down if not completely blocking part of the production activities. For this reason, pressure has recently put on analysis working groups to integrate as much as possible their activities in the production framework. This would allow the operations team to better control the stress put on computing and storage resources, depending on the relative priorities of the various scheduled production activities. Working Groups will also gain advantages when coming to data provenance, recording and distribution.
- **Monitoring and testing**: continuous testing is done using test jobs. Some of these test jobs are sent by WLCG, and are VO-agnostic. LHCb complements the tests with VO-specific ones.

2.2. Data Management

Data Management activities are usually a direct consequence of the production activities of Data Manipulation. Tier 0 and Tier 1 centers provide LHCb with storage resources, that are split in three *space tokens*: *user*, *tape*, and *disk*. Staging data from tape to disk is an important part of daily activities, which requires a careful monitoring. Data management activities include:

- **Replication**: data is replicated, on disks, as soon as it is produced. The number of replicas produced follows the computing model. There are usually more replicas of real data than replicas of simulated data.
- **Archival**: at least one copy of each data that is produced is archived, on a tape system. This custodial copy is not meant to be used by users or productions. Archival is a safe measure introduced not long ago in the computing model.
- **Removal**: data is from time to time selectively removed, to free disk space for newer data. Sometimes, the number of replicas of older datasets is reduced, based on a simple popularity algorithm.
- **Checks**: these checks include consistency, for example between the file catalogs in use, or with the real content of the storage elements, and user quota checks.

Even if a large part of these Data Managements operations are automatically executed, data managers are in charge of performing part of these activities, especially for what regards reducing the number of replicas, or removing datasets.

3. The Operations team

The distributed computing operations team comprises a varying number of shifters, while few roles are covered continuosly by a small set of experts. Most of the members of the team meet each other at the daily LHCb offline operations meeting to discuss progress, problems encountered, and to take decisions on short term operational issues. Figure 1 shows the connections between the roles of the operations team.

Within this section, we explain each role's duties, its connection with the other members, a bit of history behind it, and who is covering each of them.

- **GEOC (Grid Expert On Call)**: this role is covered on a week-long shift base by a small set of experts (not more than 5). The GEOC is a central role in the daily operations schedule. A GEOC acts first of all as an information collector, is in charge of running the daily operations meeting, and to represent LHCb at the daily WLCG operations meeting. When issues have to be discussed with resource owners, GEOC deals with site managers and with LHCb contacts at the Tier 1. She receives input from the computing shifters, but it is also their main support when in need of an expert opinion. Other shifters, like distributed analysis support shifters, are referring to the GEOC when a grid-releated problem can't be simply solved. A GEOC is first of all an expert shifter, and since LHCb main resource provider is WLCG, she has to have a good knowledge of how LHCb is using WLCG resources. A GEOC is not an all-round expert, is in fact a connection role inside the operations team.
- **Production Managers**: the role of production manager is covered on a six months base. In today's team, there is a production manager in charge of running Monte Carlo simulations, and a second has been introduced since when real data taking started, and runs and controls all the other production activities. All real data activities are highly stressing in terms of computing and storage, and also very delicate: thus, the role of the real data production manager is fundamental for the collaboration. Production managers are in close contact not only with the computing shifters, but also with application managers, and translates in production terms the decisions taken by Physics and Operations Planning Groups.

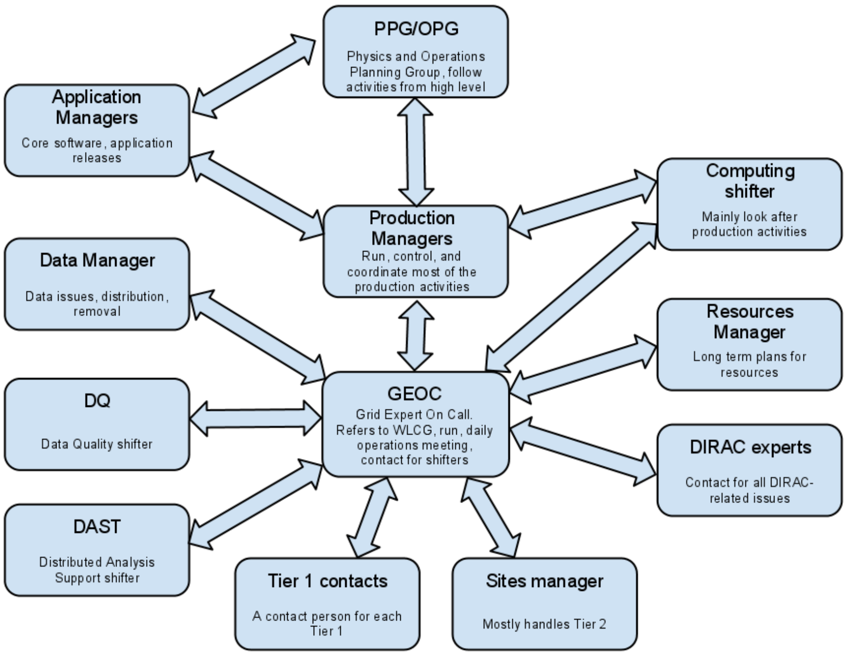

Figure 1. Organization of the Operations team

- **Computing shifters**: acting as helpers of the GEOC and production managers, computing shifters are selected from a pool that range from 20 to 30 people. A shift lasts 4 to 7 days. Shifters' reports are produced daily, and are publicly available on the LHCb computing pages.
- **Data Quality (DQ) shifters**: as well as computing shifters, DQ shifters are selected from a pool that range from 20 to 30 users with the same 4 to 7 days turnover. DQ shifters constitute one of the link between Online and Offline activities. Their role is most of all the flagging of real data based on quality parameters, making the data available for its processing.
- **Distributed Analysis Support (DAST)**: users' issues are answered in first place by the DAST shifter. DAST shifters are selected from a pool of less than 10 experienced users, and rotate each 7 days.
- **Data Manager**: this role, as well as the ones that follow, is not subject to any turnover. The Data Manager became operationally an important role since when real data taking started. Tools for bulk dataset handling have been recently developed, as well as consistency checks and automated monitoring.
- **Sites manager**: mostly dedicated to the handling of Tier 2 resources, the sites manager adds and removes sites and resources from the production mask, depending on their availability.
- **Tier 1 contact persons**: at each Tier 1, LHCb has a contact person. Even if a contact person doesn't usually dedicate 100% of her time to LHCb computing issues, they become useful when issues with the site arise.
- **Application Managers**: each LHCb data processing application has a manager, who is in charge of formally specifying the application versions to be used by Data Manipulation productions at any time.

- **Resources Mananager**: defines long term planning of resources.
- **LHCbDIRAC experts**: being LHCbDIRAC the tool LHCb developed for handling its distributed computing activities, some DIRAC developers are also part of the operations team.

4. Used Tools

DIRAC (Distributed Infrastructure with Remote Agent Control) [4] is the software used by LHCb for its distributed computing activities. It has been initially developed as a LHCb-specific project, but many efforts have been made to re-engineering it into a generic framework, capable to serve the distributed computing needs of a number of Virtual Organizations. Now DIRAC, which is developed in python, is a community Grid solution, offering powerful job submission functionalities, and a developer-friendly way to create services and agents. After this complex reorganization, the LHCb-specific code resides in the LHCbDIRAC extension while a core, VO-agnostic, DIRAC project has been disentangled. Even if DIRAC and LHCbDIRAC are now following distinct development cycles, for simplicity within this paper we refer to DIRAC and LHCbDIRAC as synonyms. DIRAC includes many components, some of which are used daily in the operations.

DIRAC is not the only tool used, but has a central role in all LHCb activities. In what follows, some of its components are highlighted.

We split operations actions into 4 areas: *monitoring*, *management*, *reporting*, and *testing*.

4.1. Monitoring

Monitoring is the source of the majority of operation activities. The proliferation of monitoring tools brought the VO administrators to interact with a variety of dispersed information somehow difficult to handle successfully. This is one of the main reasons to promote initiatives for aggregating various infrastructure monitoring data, which is a commonly recognized requirement expressed by many Virtual Organizations. DIRAC philosophy has always been of creating a single system as a central point of control, and also the development of the monitoring infrastructure falls into this paradigm. For completeness, the tools used are:

- **DIRAC portal**: highly integrated, and redundant, the DIRAC web portal provides the majority of the information monitored daily by the team. Users and productions jobs, data replications and removals, as well as sites monitoring, historical accounting for jobs and data, are all available via the portal. As an of example, figure 2 shows the *productions monitoring page*, as well as some accounting plots.
- **DIRAC notifications**: DIRAC is also used for sending notifications, usually in the form of e-mails. Notifications come from monitoring agents. Examples of such notifications are, for example, sent to users who are over-using their storage quota.
- **SLS**: Service Level Status (SLS) is web-based tool that dynamically shows availability of IT services. LHCb uses SLS mostly for infrastructure monitoring.
- **GOCDB**: The Grid Operations Center Database (GOCDB) contains general information about the sites participating to the production Grid that are part of WLCG. A visualisation portal is provided for storing and presenting Grid topology. Downtime announcement are sent regularly as mails.
- **Dashboard**[5]: extensively used by different VOs, for different purposes, dashboard offers a toolkit for monitoring both community activities and status of resources aggregating them in useful views. LHCb uses it only for a very specific monitoring, namely the Nagios [6] testing infrastructure, which provides VO agnostic testing of grid services.

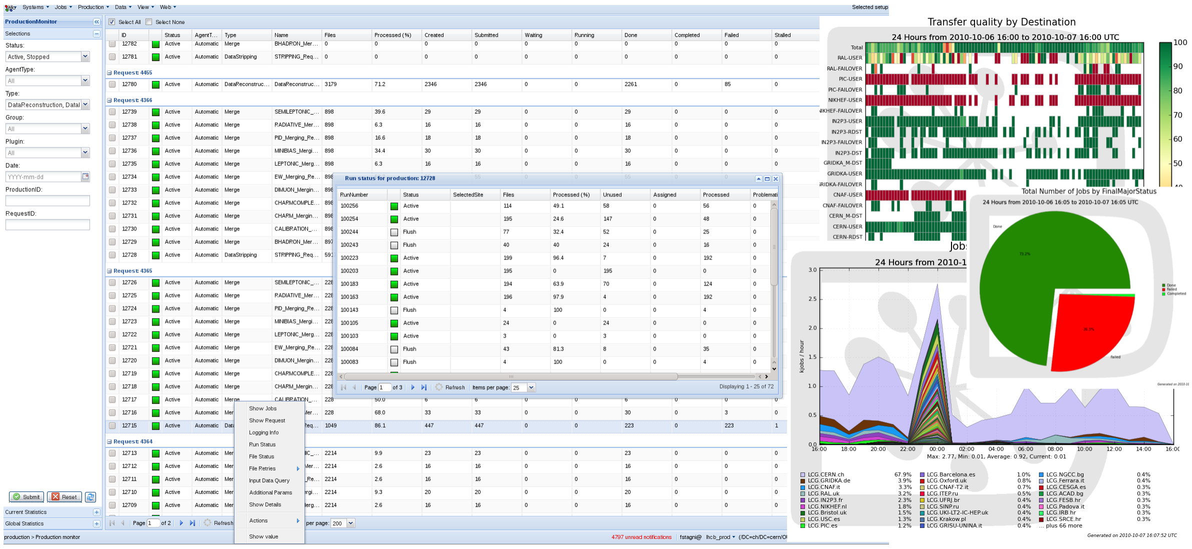

Figure 2. DIRAC Production monitoring page

4.2. Management of operational parameters

There are few places where operational parameters are set. A change of some parameters can have a drastic effect on how the resources are used: this highlights once more how vital is the experience of the members of the operations team.

- **DIRAC CS**: The DIRAC Configuration Service is the source of a very large part of configuration parameters. Reliable, easily accessible, integrated and redundant, its parameters are used at the client and server side, while each and every job runs with a dump of its content. A selected number of experts are capable to modify the parameters.
- **VO ID Card**: expresses the minimal requirements each site has to fullfill to be part of the LHCb Grid, like memory or CPU requirements.
- **DIRAC Production Parameters**: every time a new production is created, a number of parameters are set, influencing how the jobs are created, the input and output files treated, and the CPU requested.

Once the Grid operations within a team are established with clear procedures in place, running the system becomes a routine job for an experienced Grid administrator. Taking a decision or applying a procedure should be as much as possible an automatic operation. We refer to such operations as policies. As of today, a tool to evaluate and enforce such policies is missing. The lack of automation in the management of such policies always requires a Grid operations administrator to be available in order to react manually when changes happen in the system. LHCb is integrating within LHCbDIRAC an autonomic system that will help the operation team automating certain operations.

4.3. Reporting

Reporting of progress, issues, or simply as knowledge spreading, is an often undervalued activity. It's instead an important part of the daily duties of an operations team. The tools used by LHCb comprise:

- **eLog**: an electronic Log application is used for maintaining an operational logbook.
- **GGUS**: the Global Grid User Support (GGUS), part of EGI, the European Grid Infrastructure, is the official tool for VO support and issues reporting.

- **Twiki**: a wiki implementation, mostly used at CERN.
- **Savannah**: bugs reporting, discussions with developers.

4.4. Testing
Testing of resources, and application stacks, are performed regurarly:

- **Nagios**: Nagios tests are sent continuously at each Grid site as part of the monitoring and testing activity.
- **Production Validation**: when a new LHCb software stack is available, a validation production is performed. A validation production is an integration test which proved to be very useful for both the distributed computing team and the application developers.

5. Conclusions
In this paper we presented the computing activities performed by the LHCb offline operations team. We presented the team organization, and the monitoring system for the LHCb Grid Computing, which relies on many heterogeneous and independent sources of information. This work summarized the state-of-the-art of LHCb Grid operations. We showed the reasons that brought to various choices and what are the tools currently in use to run our daily activities.

References

[1] Collaboration M Models of networked analysis at regional centres for lhc experiments URL `http://monarc.web.cern.ch/MONARC/`

[2] Brook N 2004 Lhcb computing model Tech. Rep. LHCb-2004-119. CERN-LHCb-2004-119 CERN Geneva

[3] Burke S, Campana S, Lanciotti E, Mendez Lorenzo P, Miccio V, Nater C, Santinelli R and Sciaba A 2009 *gLite 3.2 User Guide* URL `https://edms.cern.ch/file/722398//gLite-3-UserGuide.pdf`

[4] Tsaregorodtsev A, Bargiotti M, Brook N, Ramo A C, Castellani G, Charpentier P, Cioffi C, Closier J, Diaz R G, Kuznetsov G, Li Y Y, Nandakumar R, Paterson S, Santinelli R, Smith A C, Miguelez M S and Jimenez S G 2008 *Journal of Physics: Conference Series* **119** 062048 URL `http://stacks.iop.org/1742-6596/119/i=6/a=062048`

[5] Andreeva J, Gaidioz B, Herrala J, Maier G, Rocha R, Saiz P and Sidorova I 2009 131–139 10.1007/978-0-387-78417-5_12 URL `http://dx.doi.org/10.1007/978-0-387-78417-5_12`

[6] Barth W 2006 *Nagios. System and Network Monitoring* u.s. ed ed (No Starch Press) ISBN 1593270704 URL `http://www.amazon.de/gp/redirect.html=wsbscriptionId=13CT5CVB80YFWJEPWS02`