The LHCb Experience on the Grid from the DIRAC Accounting Data

Adrian Casajús, Ricardo Graciani, Albert Puig, Ricardo Vázquez
Universitat de Barcelona

on behalf of the LHCb Collaboration

Abstract. DIRAC is the software framework developed by LHCb to manage all its computing operations on the Grid. Since 2003 it has been used for large scale Monte Carlo simulation productions and for user analysis of these data. Since the end of 2009, with the start-up of LHC, DIRAC also takes care of the distribution, reconstruction, selection and analysis of the physics data taken by the detector apparatus. During 2009, DIRAC executed almost 5 million jobs for LHCb. In order to execute this workload slightly over 6 million of pilot jobs were submitted, out of which approximately one third were aborted by the Grid infrastructure. In 2010, thanks to their improved efficiency, DIRAC pilots are able, on average, to match and execute between 2 and 3 LHCb jobs during their lifetime, largely reducing the load on the Grid infrastructure.
Given the large amount of submitted jobs and used resources, it becomes essential to store detailed information about their execution to track the behaviour of the system. The DIRAC Accounting system takes care, among other things, to collect and store data concerning the execution of jobs and pilots, making it available to everyone via the public interface of the LHCb DIRAC web portal in the form of time-binned accumulated distributions. The analysis of the raw accounting data stored allow us to improve and debug the system performance, as well as, to give a detailed picture on how LHCb uses its Grid resources. A new tool has been developed to extract the raw records from the DIRAC Accounting database and to transform them into ROOT files for subsequent study. This contribution presents an analysis of such data both for LHCb jobs and the corresponding pilots, including resource usage, number of pilots per job, job efficiency and other relevant variables that will help to further improving the LHCb Grid experience.

LHCb-PROC-2011-026
03/05/2011

1. Introduction

The LHCb detector [1][2][3] is one of LHC four main experiments and has already collected $37\,\mathrm{pb}^{-1}$ of data, which have been reconstructed and are being analyzed on the Grid. The DIRAC (Distributed Infrastructure with Remote Agent Control) project [4] is the Grid solution developed and used by the LHCb community. It forms a layer between the users and its various computing resources to allow optimized, transparent and reliable usage and is used for all LHCb Grid activities.
Given the great number of activities carried out on the Grid by DIRAC, it becomes essential to keep track of actions performed. This task is performed by the DIRAC Accounting System. Not only is this very useful for DIRAC maintenance and monitoring purposes, but also it can be used to extract information about how the LHCb community is using the Grid resources.

It is the purpose of this paper to present the Grid usage patterns that have emerged in LHCb since CHEP'09, as well as the tools used to extract such patterns from the DIRAC Accounting System. In order to do so, the DIRAC Accounting System will be outlined in section 2. General information extracted from time-binned Accounting records will be shown in section 3, while detailed usage patterns extracted directly from the raw Accounting will be shown in section 4. Some conclusions on the LHCb experience on the Grid will be drawn in section 5.

2. The DIRAC Accounting System

DIRAC has been used in LHCb for tasks such as distributed Monte Carlo simulations, data reconstruction and physics analysis. The DIRAC Accounting System takes care of collecting and recording the data regarding LHCb activities on the Grid, from pilot jobs to data transfers. Data stored include properties to help with classification of records, such as site and user, and properties that can be measured, such as consumed CPU and total time. This information is recorded in the Accounting Database in two different formats as can be seen in figure 1:

- *Raw records* The information about the relevant event is stored in the Accounting Database as one new row.
- *Accounting time buckets* Different parameters from the raw records are accumulated in time buckets and classified according suitable properties. This data can be displayed in the DIRAC web portal, where the accumulated parameters are used to generate graphics based on selections of the available properties [5].

The DIRAC Accounting system, with the plots from the web portal and information extracted from the raw records, provides an insight of the usage and performance of LHCb on the Grid. Even though the Accounting system provides information on many different types of Grid operations, such as jobs, pilot jobs, data transfers and so on, in this document only Job records will be analyzed, since they most easily provide an insight on how the LHCb community has been using the Grid.

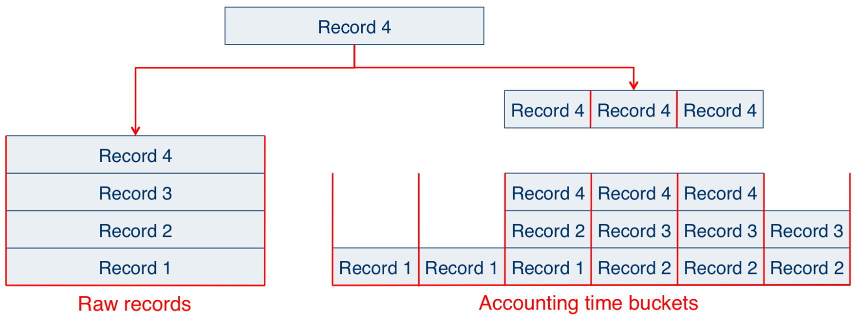


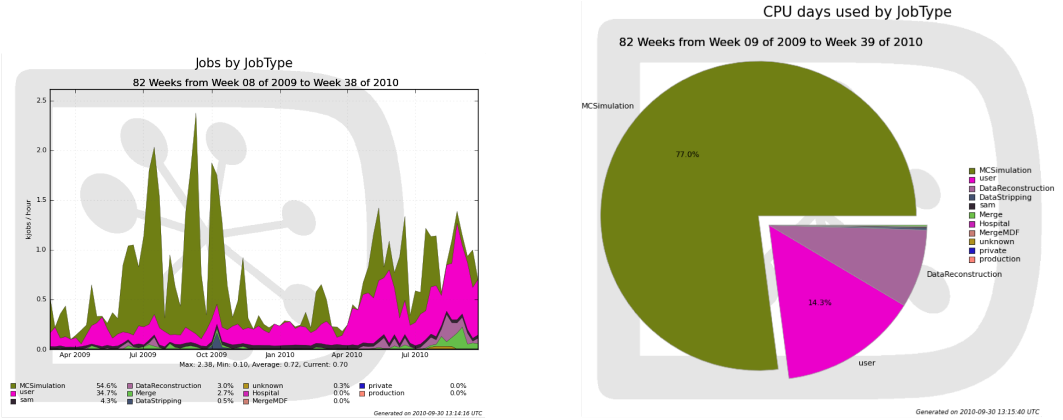
Figure 1. Schema of how a record is stored in the DIRAC Accounting System

3. Accounting with time-binned records

Time-binned accounting plots, which can be displayed in the DIRAC web portal [5], are very useful to grasp the bigger picture on how LHCb has been using the Grid. Since CHEP'09 (March 2009), the usage of the Grid by the collaboration has changed a great deal as a consequence of the arrival of first data from the LHC.

3.1. LHCb Jobs

Plots extracted directly from the DIRAC web page, such as the ones in figure 2, show that, from CHEP'09 until CHEP 2010, LHCb has executed over 8M jobs which represent almost 4500 CPU years.



(a) Jobs by Job Type

(b) Percentage share of CPU time by job type

Figure 2. Usage of the grid by the LHCb collaboration mainly divided in MC simulation (green), User jobs (pink) and Data Reconstruction (purple).

The fact that plots can be easily split using different criteria allows us to extract deeper information regarding the composition of the jobs that have been executed on the Grid.
Monte Carlo Simulation is the dominant contribution both in number of jobs, with over 50% of jobs (figure 2(a)), and CPU days (figure 2(b)), with over 75% of CPU usage. However, these simulation activities are far from constant in time, mainly due to software cycles and physics requirements. For example, the simulation requests made by the various physics working groups and collaboration-wide simulation campaigns.
Nonetheless, figure 2(a) shows an increasing contribution from user activities to the point of becoming dominant in recent months. Moreover, following the LHC start-up, data processing activities have become significant up to the point of representing 30% of used CPU in the last two months.

3.2. LHCb User Jobs
As stated above, the starting of LHC operations and the possibility to access a significant amount of data taken by the experiment has drawn many users to using Grid facilities. Figure 3 shows a threefold increase on the number of unique users that have used the Grid on a given month since CHEP'09 with a heavy jump in the number of monthly unique users around March 2010.

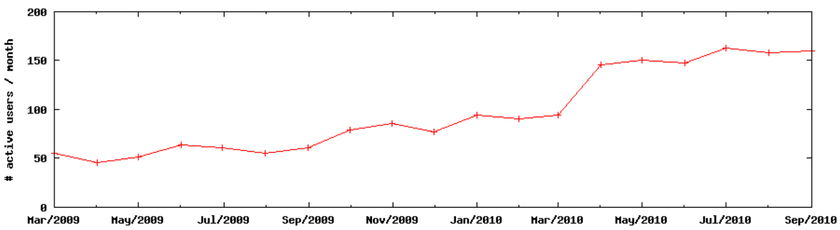
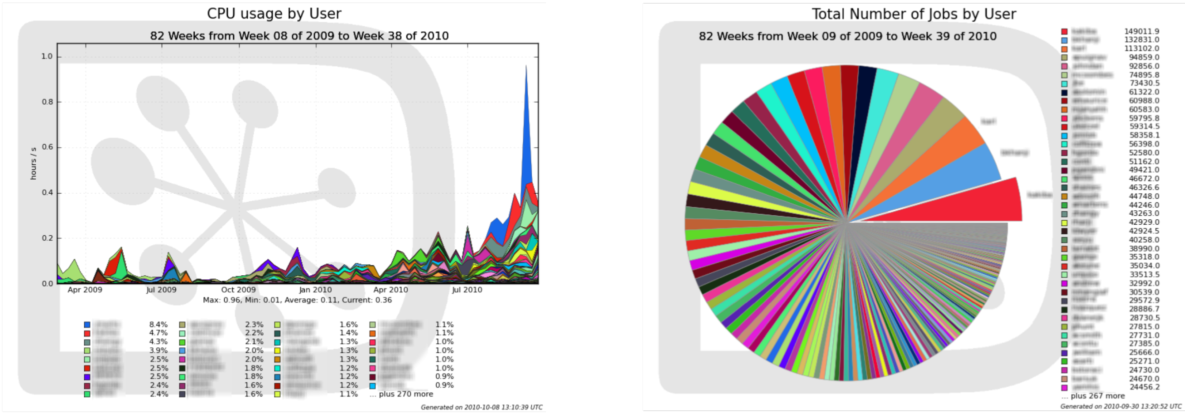


Figure 3. Number of different LHCb Grid users on a given month

As for how this increasing number of users have used Grid resources, figure 4(a) shows almost 800 years of CPU consumed, amounting to over 2.5M jobs (figure 4(b)). It is also clear from



(a) CPU usage by user

(b) Total number of jobs by user

Figure 4. LHCb user jobs.

figure 4 the aforementioned increase in user jobs with the arrival of LHCb data. Not only has the number of users increased, but also their usage is fairly alike (similar number of jobs among the basis of users), as can be seen in figure 4(b).

4. Accounting with raw records

In the process of time-binning and aggregating data information of individual events accounted is lost. That means that the time-binned accounting plots present an average view of the real performance, being very useful when detecting trends but lacking the necessary granularity for more detailed analyses. Job-by-job distributions provide the suitable detail for studying performance of LHCb jobs on the Grid.
To perform this detailed analysis a new set of tools has been developed. These tools are used to go from the raw records to the final plots, performing the following tasks:

- *Convert raw Accounting records to ROOT NTuples.* Dumping the Accounting Database allows to perform intensive selection and plotting tasks without increasing the load on the Database.
- *Produce plots* from the NTuples according to multiple selection and display criteria.

4.1. LHCb Jobs

Job-by-job information allows to study different patterns in the usage of the Grid. It is illustrative to study the distribution of unnormalized CPU Time split by job type (figure 5) from May 2010 to CHEP 2010, where different patterns emerge:

- *Data reconstruction* consists of jobs of varying length depending on input file size and version of reconstruction software. This is still evolving, so the pattern is still changing.
- *MC Simulation* presents well-defined peaks, corresponding to given simulation conditions and to different number of events per job.
- *User jobs* are usually very short (< 1h), and only a few extend beyond the 2-hour mark.

4.2. LHCb User Jobs

As a complement of figure 3, figure 6 shows that between May 2010 and CHEP 2010 over 200 LHCb members have been running analysis jobs on the Grid. The average number of successful jobs per user is 5.4k, while failed jobs per user are, on average, 1.7k.

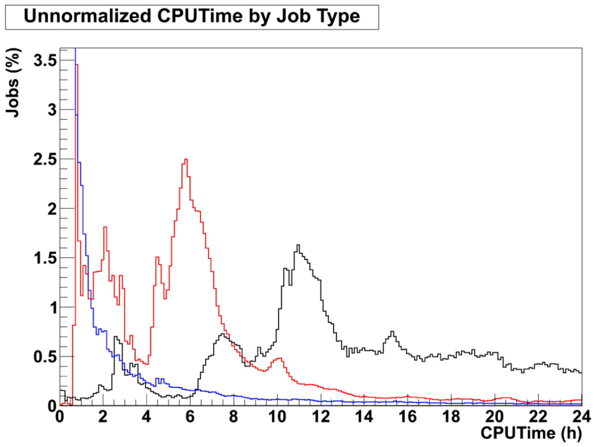


Figure 5. Unnormalized CPU Time for user (blue), reconstruction (black) and MC simulation (red) jobs.

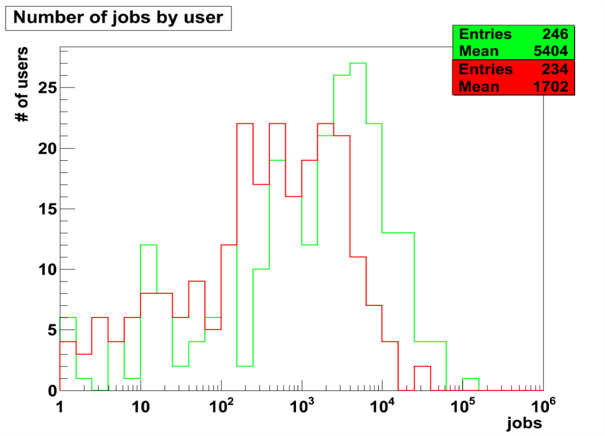


Figure 6. Number of successful jobs (green) and failed jobs (red) per user.

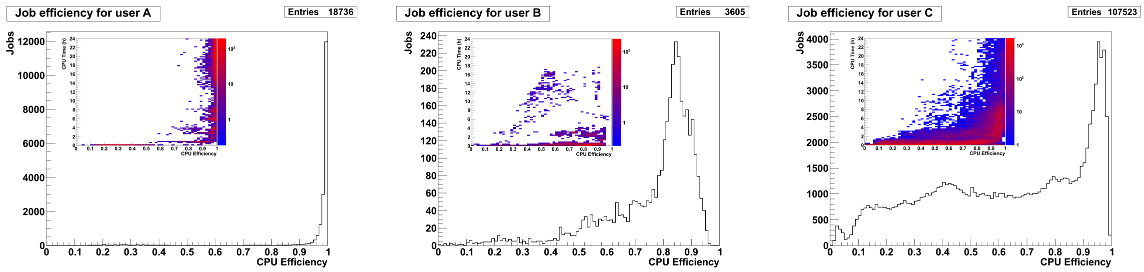
Moreover, the study of the raw records allows to monitor the performance of the individual users' jobs. Figure 7 shows different patterns that emerge from studying individual users:

- Long ($> 12h$), very efficient jobs (figure 7(a)), similar to MC Simulation jobs (figure 7(e)).
- Medium ($> 2h$, $< 12h$), efficient jobs (figures 7(a), 7(c)).
- Short ($< 2h$), efficient jobs (figure 7(b)).
- Short jobs, with irregular efficiency (figure 7(d)).
- Short, inefficient jobs (present in all plots in figure 7).

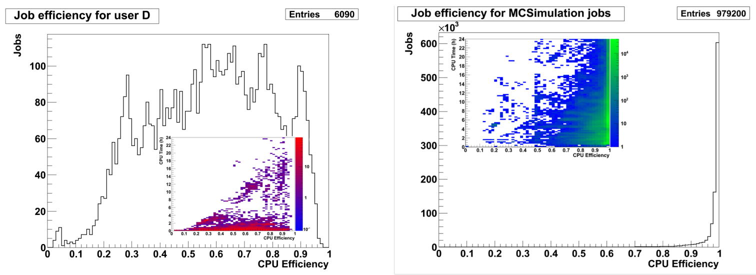
As a general conclusion, we can see that users value response time over CPU efficiency. That means sending many short jobs instead of a few long ones, even though the longer ones are more efficient.

5. Conclusions

The DIRAC Accounting system allows to monitor and analyze the usage of the Grid by the LHCb collaboration. On one hand, it allows to visualize (from the DIRAC website) time-binned averages and aggregates that allow to study trends in the behavior. On the other hand, the raw



(a) Job efficiency for user A (b) Job efficiency for user B (c) Job efficiency for user C



(d) Job efficiency for user D (e) Job efficiency for MC Simulation

Figure 7. Job efficiency distribution from example users (a-d) and MC simulation (e). The small plots represent the CPU time used versus job efficiency.

accounting records can be used to perform more detailed studies on the usage patterns. At the moment, these latter studies have to be performed manually, but integration with the DIRAC website is foreseen.

Time-binned plots show that, in the period from CHEP'09 to CHEP 2010, over 8M successful jobs have been run, representing almost 4500 CPU years. Even though MC simulation usage is still dominant, data-related activities (such as user analysis and data reconstruction) have greatly increased since March 2010, along with the arrival of LHCb data. In fact, the number of LHCb users using the Grid has increased threefold, going up to 200 active users.

User profiles extracted from the raw accounting records can be used to help users improve their analysis jobs, as well as understanding their needs. Several patterns have been extracted from user jobs efficiency distributions, as well as CPU time distributions. These are usually different from patterns that emerge from MC simulation jobs, for example, since user needs and priorities are different.

Overall, the DIRAC Accounting system has proven to be a very useful tool to analyze the LHCb experience on the Grid and has allowed us to see the shift in usage pattern since the arrival of detector data.

6. Bibliography

[1] Amato S *et al.* 1998 LHCb technical proposal. Tech. rep. LHCb CERN-LHCC-98-04. LHCC-P-4

[2] Antunes-Nobrega R *et al.* (LHCb) 2003 LHCb technical design report: Reoptimized detector design and performance. Tech. rep. LHCb CERN-LHCC-2003-030

[3] Alves A A *et al.* (LHCb) 2008 *JINST* **3** S08005

[4] Tsaregorodtsev A *et al.* 2003 *Computing in High Energy and Nuclear Physics, 24-28 March 2003, La Jolla, California* (*Preprint* `cs/0306060v1`)

[5] URL `https://lhcbweb.pic.es/DIRAC/`