Monitoring the DIRAC Distributed System

R. Santinelli, M. Seco, R. Nandakumar on behalf of the LHCb DIRAC team
CERN, University of Santiago and STFC
E-mail : raja.nandakumar@stfc.ac.uk

Abstract. DIRAC, the LHCb community Grid solution, is intended to reliably run large data
mining activities. The DIRAC system consists of various services (which wait to be contacted
to perform actions) and agents (which carry out periodic activities) to direct jobs as required.
An important part of ensuring the reliability of the infrastructure is the monitoring and logging
of these DIRAC distributed systems. The monitoring is done collecting information from two
sources - one is from pinging the services or by keeping track of the regular heartbeats of the
agents, and the other from the analysis of the error messages generated both by agents and
services and collected by a logging system. This allows us to ensure that the components are
running properly and to collect useful information regarding their operations. The process
status monitoring is displayed using the SLS sensor mechanism that also automatically allows
to plot various quantities and keep a history of the system. A dedicated GridMap interface
(ServiceMap) allows production shifters and experts to have an immediate, high-impact view
of all LHCb critical services status while offering the possibility to refer to details of the SLS
and SAM sensors. Error types and statistics provided by the logging service can be accessed
via dedicated web interfaces on the DIRAC portal or programmatically via the python based
API and CLI.

1. Introduction
The LHCb experiment[1] is the Large Hadron Collider beauty experiment at CERN, primarily
intended for precise measurements of CP violation and rare decays in b-physics. Along with the
remaining LHC experiments, LHCb expects to start taking data towards the end of 2009. The data rate
is expected to be about 5 PetaBytes per full year of running. This data will need large computing
resources at a scale in which the only reasonable solution is the worldwide computing grid.

LHCb uses the worldwide computing grid to perform most of the large scale computing tasks. The
DIRAC[2] community grid solution is the interface to perform these tasks. The DIRAC system
consists of various services and agents to direct jobs as required. This paper concerns the different
monitoring systems currently in place to give us a view of the status of the DIRAC systems and ensure
that the needed services and agents are running smoothly and to troubleshoot in case of problems
observed.

Here we describe three different levels of monitoring. The first are the GridMap and SLS[3]
sensors that monitor the different services which make up the overall LHCb grid interface – both
internal (DIRAC) and external (LFC, FTS, MyProxy, etc.). Next is the status information provided by
dashboard which gives the information of the various sites as seen by LHCb. And finally we have the
DIRAC logging system which logs the messages from the DIRAC processes and allows us to understand and target possible problems in the DIRAC system.

2. GridMap and SLS

2.1. GridMap

The GridMap (Figure 1) gives a high level view of all the service statuses in a single interface. It is accessible at the following location - http://servicemap.cern.ch/ccrc08/servicemap.html?vo=LHCb.

The SAM (Service Availability Monitoring described in the next section) and SLS (Service Level Status) sensors feed into this plot and produce data used by the GridMap to compute the health status of any given Grid service.

Moving the cursor to a given service gives us more details of the service status, and a colour gradient indicates the current status of the service. The size of a given service corresponds to the criticality of the service as perceived by LHCb. The GridMap technology is used by LHCb because it offers a top-down view of the Grid service status and allows immediate problem spotting in case some service is not working properly.

2.2. SLS sensors

The SLS sensors were originally only in the CERN domain for service providers are now set up to offer a hierarchical view of the status of different services that the SAM framework cannot cope with. Currently all VOs use these sensors extensively to track their own specific services. These sensors run from CERN and ping the various services (WLCG or otherwise). The information is fed into the GridMap described above and is also made available in a dedicated web page as shown in figure 2.

The sensors can be set to measure any critical metric defined and the SLS server databases maintain historical plots of the system, giving us a fine grained service health status information. The system can also be set up to broadcast alarms in case of need.
3. LHCb dashboard and SAM information

The LHCb dashboard is the central point for monitoring the status of sites and is available at the page [http://dashb-lhcb-sam.cern.ch/](http://dashb-lhcb-sam.cern.ch/). This page collects the information from the LHCb SAM tests and displays it in a convenient view. The SAM graphical user interface of the dashboard introduces a more flexible notion of availability that goes beyond the algorithms to compute site availability used by the WLCG monitoring tools like GridView. It is now left to the VO to decide which tests are really critical and really matters the VO. This translates into high interest levels from the site administrators in using this interface to find out how LHCb perceives their site.

By selecting various options it is possible for the site administrator to monitor for example the historical status of all the sites together (Figure 4), or actual status of various services of a given site (Figure 3). In particular, clicking on any test, service or site in these displays gives more detailed information about the item. The high flexibility and rich set of functionalities of the Dashboard has also made it attractive to LHCb. The dashboard is now used extensively within the LHCb grid operations team for site and service status monitoring during the daily operations activities.

3.1. Service Availability Monitoring (SAM)

The LHCb SAM framework [4] is currently one of the main sources of information for the LHCb operational team to monitor the overall site status for all the sites in the production mask. The LHCb SAM suite has been deliberately split into two parts:

3.1.1. A DIRAC based suite that is meant to run LHCb specific application software (and to install software on
the sites in case it is required). These tests are run by submitting simple standard jobs through DIRAC to the sites. The DIRAC system is interfaced with SAM framework and the tests are guaranteed to be run quickly by means of the prioritization mechanism in place in DIRAC. This suite is targeted to check that the software is properly installed on the sites (and eventually to repair/re-install it) by running specific jobs using the whole LHCb framework. The insight offered by this suite is very significant for LHCb team and gives a perception of the site usability.

3.1.2. A Grid middleware suite: These tests are executed by running regular probes on the various sites and are completely disentangled from DIRAC. The results of these tests are exposed to sites as an LHCb perception of the site availability. Tests falling in this suite are meant to provide to site administrators important and basic infrastructure information like shared area access, certificates validity, job submission but also other site service components – for example, checking that the SRM interfaces to the SE are working, space tokens are properly defined, the LFC up and running, conditions DB are accessible, etc.

Although site managers would appreciate to have insights about the LHCb site usability offered by the DIRAC-based set of tests, the site availibity is computed only from the test in this suite. The tests on the LHCb dashboard shown in figure 3 group the two suites in a unique omni-comprehensive view. These tests are run on a regular basis on all the sites publishing to be available for LHCb to run jobs on them. The actual tests are tailored to the type of site that they are (Tier-0/1/2) and the services that they provide.

The information available in the SAM, SLS, GridMap and dashboard monitoring relates to the information on grid services relevant to the typical Tier-1 administrator and also to the LHCb operations crew. The information mostly public and is either forwarded to the site administrators or to alert LHCb experts as needed.

4. DIRAC logging system
The DIRAC system itself consists of many different services and agents distributed over many machines at various sites. A dedicated logging system has been developed for handling and storing at a central facility if requested, the log records generated by the different DIRAC processes and to provide status information for both, other DIRAC components and to the administrators.

![Figure 5: Workflow of the DIRAC logging system](image)

The architecture of this subsystem follows the generic architecture designed for DIRAC and consists of services, agents and a client. The client allows DIRAC systems to generate the log records and send them to the standard output, a file and / or a central repository. The Services allow interested parties to store log records in a central database and retrieve information from the records stored in that database.
The information that each log record contains is the standard for a logging system – that is a timestamp, the system and subsystem that generated the message and the log message. The log message itself has been broken in two parts – a fixed part that describes the generic status and another transient that contains mutable information like filenames, connections, etc. If the message is sent to the central repository extra information is added: the username, groupname, the machine name and IP and the site name. This extra information is extracted from the secure channel that is established between the client and the server.

The life cycle of a log message is represented in the figure 5 above. A DIRAC process generates a log record and, depending on the configuration, the message will be sent to the standard output, a file, the central server or a combination of the above. More output handlers could be added in the future, if needed. If the log record is sent to the server, a receiver service will store the message in the database along with the extra information. When the message becomes old a cleaner agent will remove it.

Users will have several means of accessing the information stored in the database. Reporter services will collect information about the records stored in the logging database and submit this information to relevant users or administrators through the DIRAC notification system. This way the Reporters can be considered to be extra sensors probing the system. A web service (figure 6) and an API are also provided to query the database. These two methods contact an

Finally there will be a set of analyzers that will analyze the errors stored on the database and feed that information to other DIRAC components so they could react to changes automatically. As an example one such agent will analyze the transfer failures between sites to extract information on whether the source or the destination is to blame for the errors and send this information to the Data transfer agent; this agent could then send less requests to the failing site or even stop them altogether until the service is restored.

5. Acknowledgements
The authors would like to thank the LHCb collaboration and in particular, the members of the LHCb DIRAC production and development team for their help and support.

References
[1] The LHCb Collaboration, A Augusto Alves Jr et al 2008 JINST 3 S08005
[2] A Tsaregorodtsev et al 2008 J. Phys.: Conf. Ser. 119 062048

[3] SLS documentation: https://twiki.cern.ch/twiki/bin/view/FIOgroup/ServiceLevelStatusProject

[4] R. Santinelli et al. Ensuring GRID resource availability with the SAM framework in LHCb J. Phys. Conf. Ser. 119:062025, 2008.