Simultaneous usage of the LHCb HLT farm for Online and Offline processing workflows

Joel Closier-1,*, Clara Gaspar1, Luis Granado Cardoso1,**, Christophe Haen1,***, Beat Jost1 and Niko Neufeld1 on behalf of the LHCb Collaboration

1CERN, European Organization for Nuclear Research, Geneva, Switzerland

Abstract. LHCb is one of the 4 LHC experiments and continues to revolutionise data acquisition and analysis techniques. Already two years ago the concepts of “online” and “offline” analysis were unified: the calibration and alignment processes take place automatically in real time and are used in the triggering process such that Online data are immediately available offline for physics analysis (Turbo analysis), the computing capacity of the HLT farm has been used simultaneously for different workflows: synchronous first level trigger, asynchronous second level trigger, and Monte-Carlo simulation. Thanks to the same software environment, LHCb can switch seamlessly and fast from Online to Offline workflows, run them simultaneously and thus maximize the usage of the HLT Farm computing resources.

1 Introduction

LHCb is one of the 4 LHC experiment at CERN and which is in activity for 10 years. The data taken by the detector are processed by a set of computers in real time to decrease the amount of information that will be stored on tape for analysis purpose afterwards by physics community.

The computing capacity is quite important and LHCb has decided to use this capacity during idle cycle to process some Monte Carlo generations. In this paper, we will describe the environment of the High Level Trigger (HLT) Farm, the various workflows that we use and how we are able to have this concept of “online” and “offline” analysis unified.

2 HLT Farm environment

The HLT Farm is composed by around 1500 PCs, distributed over 60 subfarms. The subfarms are logically divided in the Control System and each subfarm row is composed with 24, 28 or 32 PCs each. Each of the subfarms is controlled by a controller PC with WinCC OA installed on it. This application manages the HLT tasks on the HLT nodes. These controller

* Joel Closier: joel.closier@cern.ch
** Luis Granado Cardoso: luis.granado@cern.ch
*** Christophe Haen: christophe.haen@cern.ch
nodes are also connected to a top level HLT control node, which manages the availability and allocation of the subfarms for the global Experiment Control System (ECS). This number of computers represents roughly 12,000 cores that are idle when there is no data-taking.

2.1 Control System ECS

Experiment Control System (ECS)[2] in LHCb is based on the SCADA WinCC OA (see next section) with custom LHCb developed components.

ECS controls the whole experiment as shown in Fig. 1 in particular:
- Front End Electronics
- High Level Trigger (HLT)
- Data Acquisition System (DAQ)

Fig. 1 ECS overview

ECS is able to configure the whole experiment based on the different states of the LHC accelerator. In the future, we want to integrate also in ECS, the configuration of the production tools for Offline activities (DIRAC tasks)

2.2 WINCC OA

The SCADA WinCC OA is a commercial product from Siemens. It is used throughout CERN and provides, amongst other things, the User Interfaces, the archiving interface and the general alarms for the ECS.

A Control System based on WinCC OA was developed to manage the processing of the offline data on the HLT Farm. It is able to monitor and control all the worker nodes on the HLT Farm which are organized topographically with farms and subfarms as explained previously.
Each worker node can adopt independent configuration and some of the settings that can be used are:

- set the exact number of jobs on each machine
- Possibility to set automatically the number of jobs depending on the machine CPU
- In case of automatic configuration, possibility to set the number of cores to be left unused (for DIRAC)

With this configuration, there is no need to change the settings of each node to switch between task and we can easily use just a part of the farm in case some nodes are needed for data taking or for test purposes.

### 2.3 DIRAC

DIRAC [1] (Distributed Infrastructure with Remote Agent Control) INTERWARE (Fig. 2) is a software framework for distributed computing providing a complete solution to one (or more) user community requiring access to distributed resources. DIRAC builds a layer between the users and the resources offering a common interface to a number of heterogeneous providers, integrating them in a seamless manner, providing interoperability, at the same time as an optimized, transparent and reliable usage of the resources.

![DIRAC Overview](image)

**Fig. 2 DIRAC Overview**

A DIRAC script is a task started on each worker node which perform some actions to run the workflow shown in Fig. 3:

- Sets the proper computing environment
- Launches the agent which will also do several actions:
  - Query the DIRAC workload Management System to check if there some task to be executed
  - If the agent gets a job, it will:
    - Execute in the local disk where the input data, if any, will be downloaded and where the output will be written
    - At the end of the task, upload the output to the storage located in the computer centre
- During the execution of the task, information is sent to DIRAC monitoring to follow the progress of the job

**Fig. 3 Standard DIRAC job submission cycle**

### 2.4 LHCb Software Environment

LHCb is using CernVM File System (CVMFS) to distribute all the LHCb applications. CVMFS is mounted on all the computer centres that are providing computing resources for LHCb: Tier0, Tier1, Tier2 and the HLT Farm.

To run the LHCb applications, we set an environment which is also based on CernVM File System. The facility to have the same environment ease the initialisation phase of any activities on our computing resources.

### 3 LHCb workflows

In the HLT Farm, the resources are used for three main type of activities (Fig. 4) that we will describe in the following sections.

**Fig. 4 LHCb 2015 Trigger Diagram**

#### 3.1 High Level Trigger 1

The First workflow which is used, is the High Level Trigger 1 (HLT1) which reduces the data taking rate from 1 MHz to 100 KHz. This workflow is run synchronously in real time.
3.2 High Level Trigger 2

The High Level Trigger 2 (HLT2) is run asynchronously on the HLT1 output which has been buffered on the local disk of the HLT Farm worker nodes and reduces the rate to 12 KHz. This workflow is completely software based and runs on a dedicated computing farm with around 1500 PCs, which represents, roughly 50 000 Hyper Core. The HLT 2 software is installed on CVMFS on all the worker nodes.

3.3 Monte Carlo Simulation

The HLT Farm is also used during idle cycle to run DIRAC jobs which will simulate events. The simulation software is installed on CVMFS and the environment is set in the same way as the HLT2 software. This eases the usage of the farm for these two different workflows.

4 HLT Farm usage

The configuration that we have described above, allows LHCb to use the HLT Farm resources for various types of workflow and the switch between the two is very easy with the Control System that has been developed for this purpose, which controls the configuration of the farm. In the Fig. 5, you can see the activities of the Farm during data taking start-up.

![HLT Farm activities during data taking start-up](image)

As the Simulation application, based on Gaudi, has also been instrumented to catch an interrupt signal: we can see in Fig. 6, that when we need to free the HLT Farm for data taking, the Control System can send an interrupt signal to all the tasks running in the worker nodes and in few minutes, the tasks terminate gracefully without losing any processed events.
The facility to run Simulation jobs on the HLT Farm provides a big amount of resources to the Simulation production. In fact, the simulation production on the HLT Farm is bigger than other Tier1 that LHCb used in the Grid as shown in Fig. 7. In addition, we can see in Fig. 8 that the number of running jobs during the period without any data taking is put in evidence.
5 Conclusion

The fine grain configuration of the High Level Trigger farm with the Control System developed for the usage of the farm for offline tasks, and the usage of CVMFS for the distribution of any LHCb applications, allow the usage of these resources in an optimal way. The usage of the HLT Farm is maximized for 2 years and 20% of the 10 billion events of Monte Carlo that have been produced over the last year, have been done so on the HLT Farm. The HLT computing resources are now idle only for maintenance purpose.

References