LHCb 2001 – 116 DAQ
September 3, 2001

Partitioning in LHCb

LHCb Technical Note

Issue: 1
Revision: 1

Reference: LHCb DAQ 2001- 116
Created: 7th April 2001
Last modified: 06/09/01

Prepared By: C. Gaspar, R. Jacobsson, B. Jost
CERN
LHCb Computing Group

ABSTRACT

A partition is in LHCb a generic term defining a configurable ensemble of parts of the LHCb online system that can be run concurrently, independently, and with a different configuration than any other partition. The type and degree of sub-division and the meaning of *run* depends on the context. Seen from the Experiment Control System (ECS), running a partition could be controlling a single power-supply. Seen from the Timing and Fast Control (TFC) system running a partition could be timing, triggering and controlling the Front-End electronics of a sub-detector.

This note describes how the partitioning concept has been implemented in the Timing and Fast Control system, the Data AcQuisition (DAQ) system and the Experiment Control System. It also discusses the consequences and proposes two different degrees of partitioning for the TFC system: one in which the LHCb detector is divided into 16 TFC partition elements and another with 32 TFC partition elements.

Table of Contents

1 Introduction

Partitioning is an extremely important concept of the LHCb online system. It denotes the possibility of running parts of the LHCb online system independently and concurrently with other parts of the system, e.g. one partition of the system can perform a calibration activity, while another part can execute electronics tests. The requirement to do this has more or less severe consequences on the overall design of the system. In order to achieve a functional partitioning system, the concepts must be taken into account from the beginning and the system architecture must integrate the appropriate level of individual control of components.

The information in this note is a collection and an elaboration of information scattered around in other documents. Event though most of the information already exists elsewhere, it was felt that a document concentrating on the partitioning concepts and mechanisms was useful. It would also serve to unify and concert the vocabulary.

1.1 Control Domains in the LHCb online system

There are three domains or sub-system in the LHCb online system that each have their specific controls aspects:

- Experiment Controls System (ECS) [1] controls and monitors the operational state of the experiment
- Timing and Fast Control (TFC) distribution [3] distributes timing and trigger information and thus controls the operation of the front-end electronics
- Data AcQuisition (DAQ) system [1] controls the flow of the physics data through the system.

In general these domains exercise controls using a hierarchical structure. Even though these sub-systems are in principle independent, they are of course correlated and connected, either physically or logically, e.g. a detector channel is connected to the TFC system and to the DAQ system and is fed by a HV supply. Each of these sub-systems has to support the concept of partitioning, and the partitioning schemes of all three have to be coherent. This means that the definition of a partition in one domain will influence the definition of the corresponding partition in all the other hierarchies. However, in order to facilitate factorising out problems, the type and degree of finest partitioning is different for the different sub-systems. This is achieved by fully exploiting the ECS control possibilities. For instance, an ECS partition consisting of a single device, such as a detector HV power supply, and its control unit, could be tested while at the same time a sub-detector partition is performing calibration and while at the same time a Readout Unit is being replaced.

A few remarks here on the example above that justify the chapters to follow. Although the ECS partition does not directly involve resources from the other sub-systems, testing a HV power supply certainly prevents some types of activities on the detector to which the power supply belongs. Thus there is need for resource management in the system such that it is clear which resources are affected. The same is true for the intervention on the Readout Unit. However, although intervening on these two components does affect other resources, it does not prevent all activities involving the systems to which the components belong. Thus, particularly for the ECS system, it is interesting to have different *partition modes* that define different levels of individual usage as discussed in Section 2.1 (For more details see Ref. [2]).

The sub-detector partition performing a calibration makes use of resources from all three systems.

Thus, in order to set up a functional partition for this purpose, it is necessary to apply *partitioning rules*. When selecting a part of a detector to readout, the partitioning rules specify the components in the ECS system, the TFC system and the DAQ system that have to be reserved and included.

It is also clear from the examples above that the purpose of the particular partition dictate how the partition and its resources are configured.

Taking ownership of individual components and setting up partitions will be done through a user interface to a *partition database*, which contains all the partition information, inter-coherency and rules. The concept of *partition ownership* and reservation by request becomes imperative to carry out many different tests/calibrations etc by many people at the same time

1.2 Definitions

2.1.1 Partition Element

A *partition element* is the smallest component of a sub-system that can be controlled independently. A partition element determines the maximum partitioning granularity. The identification of a partition element can be different from sub-system to sub-system, e.g. the partition elements are different in the ECS and the TFC system.

2.1.2 Partition

A *partition* is a generic term for a collection of one or more partition elements from one or several sub-systems. It represents a functional subset of the online system for the purpose it is defined for. Hence, once a partition element or several have been selected it is necessary to add other components to the partition in order to construct a functional sub-system. The choice of these other components is governed by the partitioning rules. All components of a partition will be operated in a coherent fashion.

2.1.3 Masking

In some case the size of a partition element may be inconvenient to efficiently test or debug a problem. Particularly, this concerns the TFC system where the partition elements cover a relatively large amount of Front-End electronics. Another case is if a fraction of a partition element is not functioning properly and has to be excluded. To handle this the concept of masking has been introduced. Masking also requires a system-wide awareness, as it may have implications for other parts of the system. For instance, masking a fraction of a Front-End may mean that a Front-End Multiplexer (FEM) does not receive any data. Unless properly configured, the FEM and the sub-sequent DAQ chain would report errors.

It is important to observe that although masking may mean that several components become idle they cannot be used for other independent tests. They are still associated to the partition as in any other case.

2 Partition Elements in LHCb

To identify the partition elements in the LHCb online system it is necessary to differentiate between the TFC system, the DAQ system and the ECS system. This is because there is not a one-to-one relation between the smallest component that can be controlled independently in the three sub-systems and because the mechanism to control the partition elements is different in the three sub-systems. For instance, whereas the controls system basically uses a local area network (LAN) for controlling the partition elements, the TFC sub-system uses the RD-12 TTC system to exercise the controls of its elements.

2.1 ECS Partition Elements

The ECS system has a highly pronounced hierarchical structure with Control Units (CU), and Device Units (DU) as the base layer, in which control commands flow down and status information flows up (cf. Figure 1 and [2]). Due to the fact that the LAN allows access to any node from any other node connected to it, the partition element of the ECS system is defined to be a functional unit, i.e. a device (Device Unit). A device can be a front-end chip, an electronics board, a power supply etc. The controls system will allow controlling a single device independent of the other devices via the Control Units. The high granularity will prove useful when debugging or when making diagnostics of e.g. individual boards is necessary. However, it will most likely not be applied on large scale inter-operating with other parts of the system.

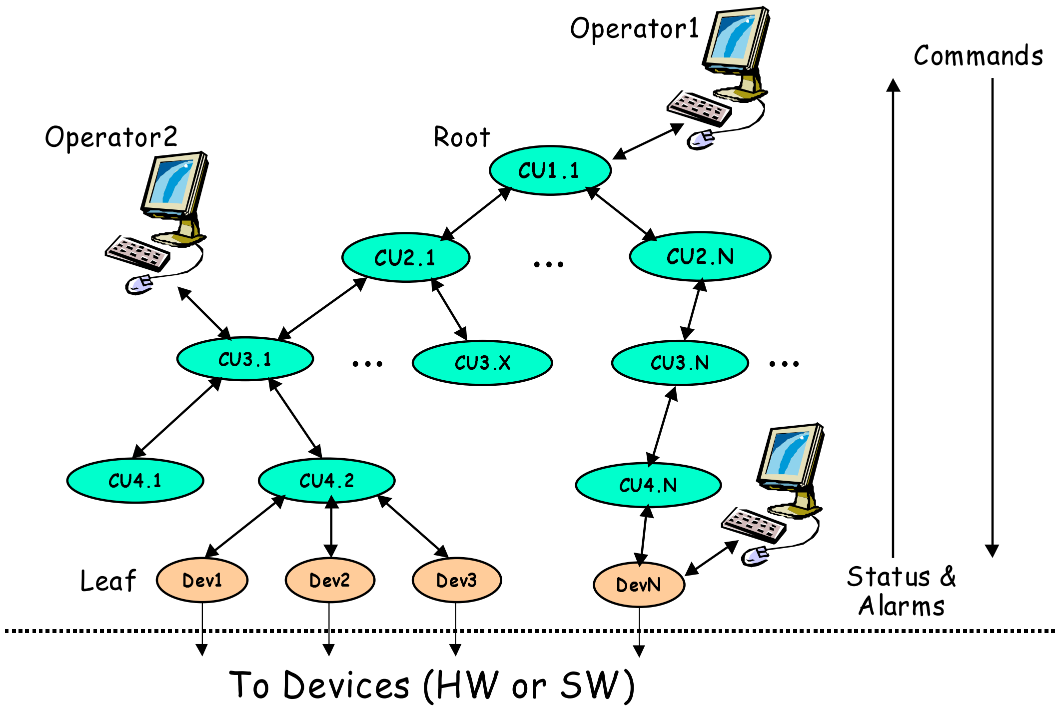

Figure 1: The ECS sub-system hierarchy.

Forming a partition means to reserve one or a set of Device Units and get control of the partition via the appropriate Control Unit covering the devices of interest. There may be several layers of Control Units in between. Generally, the type of activity on the ECS partition prevents the components to be used in partitions in the other sub-systems. However, in some case it may still be possible to use it at the same time in other types of partitions or this may even be necessary in order to trace a problem. For this purpose, the concept of different partition modes has been introduced. The modes are shown in Figure 2. (For more details, see Ref [2]). For instance, "*InManual*" means that an expert has reserved the ECS partition for controlling but it can still be used in a partition with the purpose of data taking.

Masking will be useful to debug or exclude for instance individual power supply channels or sensors.

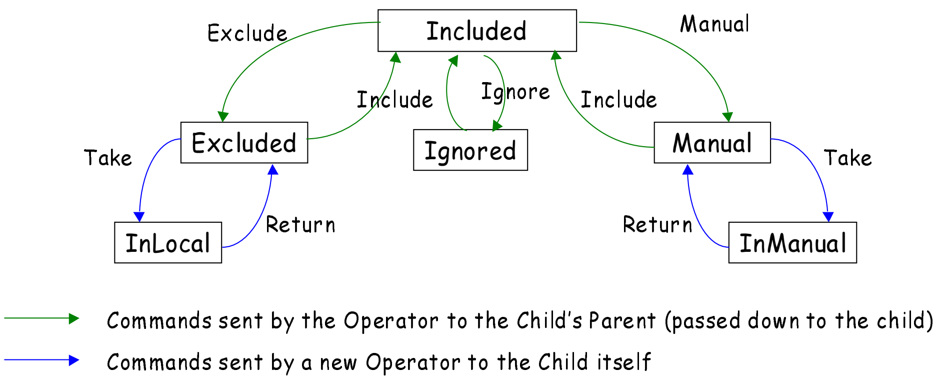

Figure 2: The figure shows the different partition modes in which an ECS partition can be controlled.

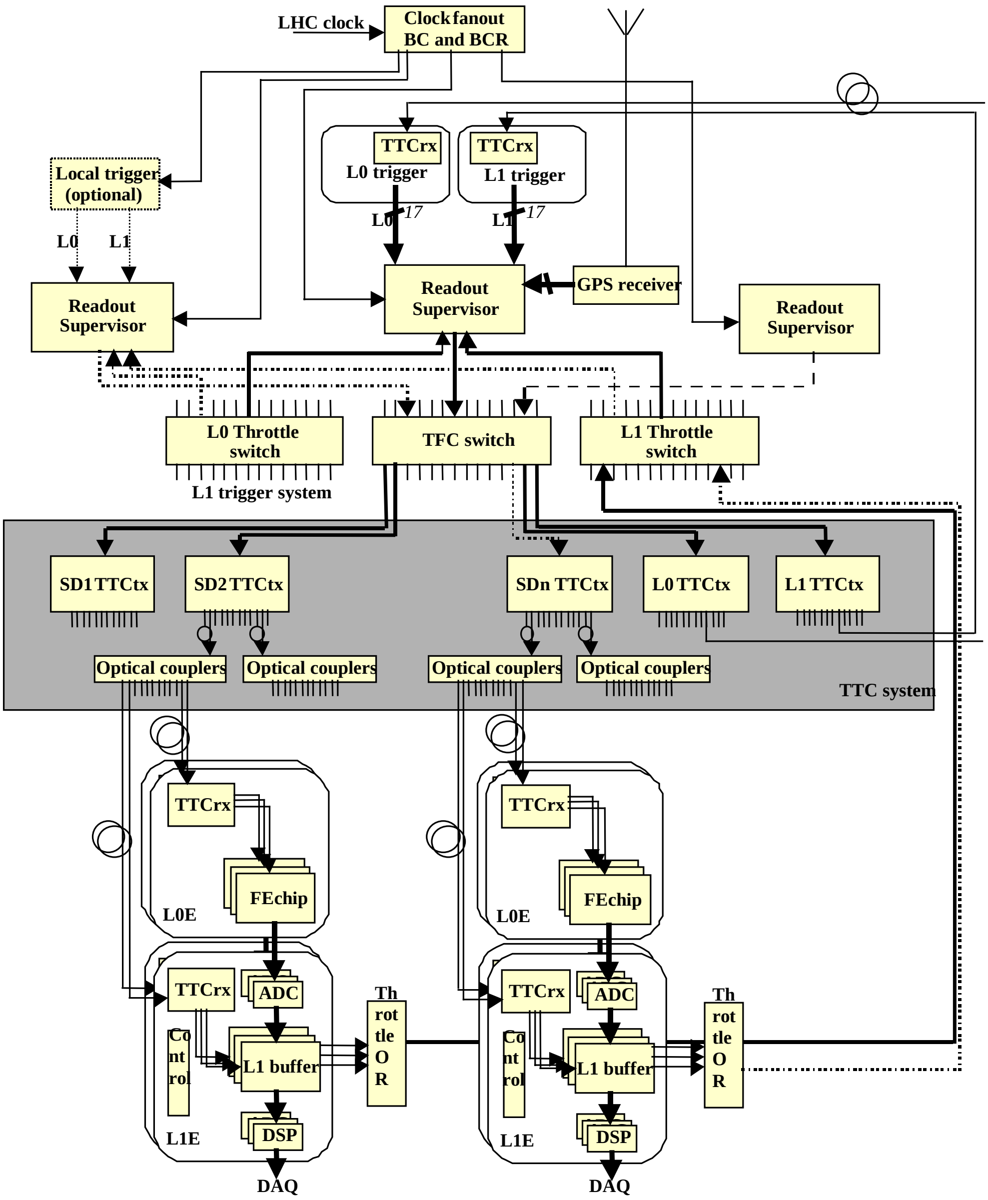

Figure 3: The TFC sub-system hierarchy.

2.2 TFC Partition Elements

The "size" of a TFC partition element is dictated by the TFC distribution system (RD-12 TTC)[5]. Since the TTC system is a passive optical distribution system, the smallest item that can be controlled individually is a TTC Transmitter (TTCtx). All devices connected to one transmitter will receive the same timing, trigger and control signals and will have to obey the same protocol. Hence the partition elements of the TFC system are defined by the sub-systems associated to each TTCtx. All other components that will be part of the partition will be added by the rules implemented in the partition database. In particular, it implies reserving a Readout Supervisor (RS) to drive the partition and configuring the Switches with the appropriate paths (Figure 3).

Observe that the Switches define the TFC partition granularity as they allow distributing the TFC signals on independent paths between sources (RS') and destinations (FEs). The TFC Switch has been designed to have 16 inputs and 16 outputs. Reserving one output for the TTCtx transmitting to the L0 trigger partition element and one output for the TTCtx transmitting to the L1 trigger partition element, the detector Front-End could be divided into 14 partition elements (Table 1). If this turns out to be limiting a second option would be to organise four TFC switches according to Figure 4, and have 32 partition elements of which two are dedicated to the L0 and the L1 trigger systems. However, note that additional cost is associated with increased granularity. Three additional TFC Switches and two additional Throttle Switches are needed plus more spares. It will also be necessary to add more TTC transmitters. However, it may not be necessary with a TTC transmitter per partition. A single transmitter can support up to 448 TTCrx's but it can also be divided into two independent transmitters and thereby support two partitions with up to 224 TTCrx's.

Table 1: The table presents a possible configuration with the LHCb detector sub-divided into 16 TFC (DAQ) partition elements.

Detector	16 elements
VELO	2
IT	2
RICH 1	2
RICH 2	1
OT	2
PRS	1
ECAL	1
HCAL	1
MUON	2
L0 trigger	1
L1 trigger	1

With the switch configuration in Figure 4, some care has to be taken when associating the TFC outputs to the different part of the detector as the arrangement does not allow all combinations of outputs.

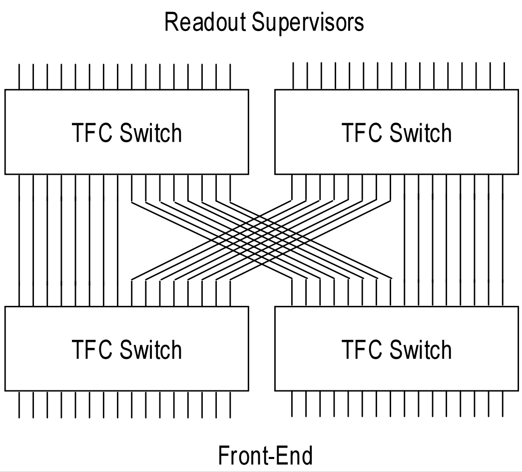

Figure 4: TFC Switch configuration for 32 partition elements.

It is important to remember the interdependence of the three sub-systems. Several of the TFC components, the Readout Supervisors and the Switches, are devices in the ECS sub-system, that is, ECS partition elements. Therefore, setting up a TFC partition also implies reserving and including the ECS devices for an appropriate control. This is defined by the partitioning rules.

For what concerns masking, the possibility will exist to program TTC receiver chips and possibly even individual front-end chips to ignore the TFC signals.

2.3 DAQ Partition Elements

The DAQ sub-system (Figure 5) is slightly special with respect to partitioning in that no individual parts can run its DAQ task stand-alone and form a proper DAQ partition analogously to those of the ECS and the TFC sub-systems. The components in the DAQ path are driven only by the data flowing down the chain from the detector Front-Ends. Of course the DAQ components are or contain devices in the context of the ECS sub-system and are therefore ECS partition elements. Hence, a “DAQ partition”, carrying out its DAQ task, only makes sense if also a part of the Front-End is included, and consequently the TFC sub-system and the ECS sub-system. The set of DAQ components (or ECS devices) that is associated to a “DAQ partition” is thus mainly determined by the partitioning rules, that is, through the connections that exist between the components in the three sub-systems. For example, if a detector channel is associated to a partition, automatically the higher layers of the readout system (Level-1 Electronics, FEMs etc) that are connected to this particular channel also have to be associated to the partition. On the other hand there are components in the DAQ system that are shareable among partitions. One example of a shareable component is the readout network. This particularity also has to be included in the partition database.

The consequence of the facts above is that the granularity and the partition element boundaries of the DAQ sub-system from the DAQ point of view are the same as those of the TFC sub-system. An exception is the shareable components, which span over several DAQ partition elements.

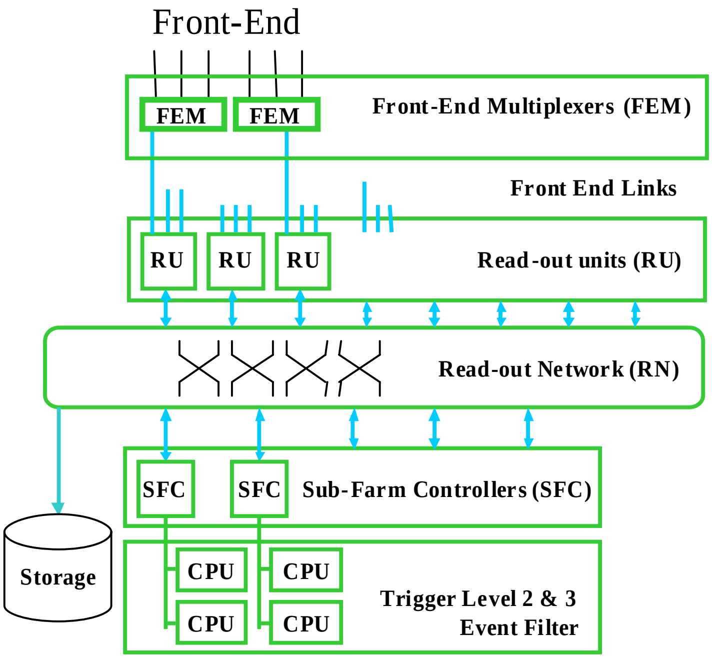

Figure 5: The DAQ sub-system..

The term DAQ partition will be used when we refer to a system, which includes the DAQ components and where the principal aim is data collection after the DAQ system. Under other circumstances the DAQ components are viewed as ECS partition elements.

3 Partition Definition and Partitioning Rules

Following the definition of a partition in section 2.1.2 it is necessary to define how partition elements of the different controls domains are associated to a partition. In the following sections we will outline the mechanisms by which partitions are defined and configured.

3.1 Defining Partitions

As already mentioned, it is necessary to distinguish between cases when a partition is being defined, that is, whether the purpose of the partition is data-taking or it is purely for controls. Whatever the purpose, the corner stone of the partitioning software and the mechanisms for partition definition is a configuration database. This database will describe the entire LHCb online system's connectivity, which in some places is physical and in some others logical[1]. Based on this description, the partitioning software will choose the elements that have to be included in the partition through the partitioning rules (see section 3.2).

1.3.1 Partitions for Controls Purposes only

Here the case is relatively straightforward. One chooses the device(s) (ECS partition elements) to be controlled and the partitioning rules will ensure that the correct controls hierarchy is established. The mechanism for selecting the devices needs to be defined. Clearly it will be impossible to

[1] The configuration database will contain more information than just the connectivity of the components of the system. It will also contain the threshold values to download into front-end modules, and (references to) the software that has to run on embedded processors, etc.

specify all the thousands of devices individually. Instead, in order to simplify, the devices will be grouped in several ways into sub-systems that have logical meanings such as sub-detectors, geographical closeness etc. It still has to be possible, however, to identify and address in the partition definition individual devices[2].

By means of masking it should also be possible to control a part of an ECS partition element.

1.3.2 Partitions for DAQ Purposes

A partition for the purpose of data taking is generally defined by firstly selecting one or a set of TFC partition elements, i.e. sub-systems governed by different TTCtx’s, as data taking is primarily conducted by a Readout Supervisor via the TTC system. However, again it should be possible to select individual components of a TFC partition element, such as for instance individual TTCrx's or front-end chips. Obviously, selecting individual components involves reserving the entire TFC partition element(s) and the selection is by means of masking as described above. The partitioning rules will ensure that the correct controls hierarchy is established. In addition, the partitioning rules will automatically add the necessary TFC components (Readout Supervisor, TFC switch paths, (local) trigger sources, etc.) and DAQ components downstream of the Front-End Chips (e.g. Level-1 Electronics boards, Readout Units, etc) to form a fully functional DAQ partition. Optionally, the corresponding ECS components could be added to control e.g. the high voltages of the sub-detectors involved or the crates that hold the readout electronics etc. The inclusion of these components (or parts thereof) will also necessitate including an integration component between the ECS and the DAQ part to allow controlling both activities through one supervising component.

3.2 Partitioning Rules

The partitioning rules will be implemented in software and will make use of a very elaborate configuration database. The configuration database has to contain all the connectivity between the different components, like for instance the association of front-end chips to TTCrx's, to Level-1 electronics boards and to macroscopic sub-detector elements. It has to contain the connections between Level-1 electronics and Readout Units, and the connections between high-voltage channels and sub-detector elements, etc. Once a partition is defined, its contents or definition can be saved in the Partition Database to ease its retrieval at a later time for reuse. The partition definition software, starting from selected partition elements and the purpose, will have to navigate through the database to find the additional components necessary to obtain a fully functional partition and to produce configuration data wherever needed.

4 Partition Awareness

Given the flexibility in defining partitions it is clear that many (if not all) of the components have to be aware of the partition contents. This does not need to be the case globally, but rather locally. For example, suppose that a phi-station of the VELO detector is not read out. This fact has to be taken into account by the connected Level-1 electronics board and potentially by the corresponding Front-End Multiplexer. Later on, the Level-2 and the Level-3 trigger software must also be aware of the fact that there will never be data from this particular phi-station. The awareness in this case is local to the, more or less, immediate neighbourhood of the phi-station, however a component of the system belonging to the calorimeter will not need to know about the absence of this phi-station. It

[2] This is basically a matter of user interface to the partition definition software, i.e. how the choices are presented to the user. Internally, a partition will most likely be described by the individual devices that make it up.

will be a task of the initialisation software to configure the components belonging to the partition according to the partition contents. From the above it is evident, that all components in the system have to be configurable in such a way, that they can be made aware of the partition contents.

5 References

[1] LHCb Collaboration, LHCb Technical Proposal, LHCC-98-04

[2] Report of the JCOP Architecture Working Group, To be released.

[3] R. Jacobsson and B. Jost, "The LHCb Timing and Fast Control", LHCb Technical Note LHCb 2001-016.

[4] LHCb Collaboration, LHCb Technical Proposal, LHCC 98-04.
M. Frank et al "DAQ architecture and read-out protocols" LHCb Technical Note, LHCb 98-028.

[5] RD-12 Documentation on WWW (http://www.cern.ch/TTC/intro.html) and references therein