Specification and Simulation of ALICE DAQ System

G. Di Marzo Serugendo, P. Jovanovic,
P. Vande Vyvre, O. Villalobos Baillie

for the ALICE collaboration
Cern, 1211 Geneva 23, Switzerland
Giovanna.Di.Marzo@cern.ch

Abstract

The Trigger and Data Acquisition (DAQ) System of the ALICE experiment has been designed to support the high bandwidth expected during the LHC heavy ion run. A model of this system has been developed. The goal of this model is twofold. First, it allows to verify that the system-level design is consistent and behaves according to the requirements. Second, it is used to evaluate the theoretical system performances using the measurements done on sub-systems prototypes. This paper presents the specification and simulation of a model of the ALICE DAQ system using a commercial tool, called Foresight. In addition, this paper describes the performances reached by the model during the simulation.

I. INTRODUCTION

The ALICE Trigger and Data Acquisition System (DAQ) is required to support an aggregate event building bandwidth of up to 4 GBytes/s and a storage capability of up to 1.25 GBytes/s to mass storage. The system must also be able to combine different types of physics events: slow rates of Central and Minbias triggers generating the largest fraction of the total data volume, together with faster rates of Dielectron and Dimuon events.

The ALICE DAQ system has been decomposed in a set of hardware and software components. The detailed system design is going on in parallel with the development of prototypes of these components. We wish to verify this design in order to check that it can reach the expected behaviour and the target performances.

However, such a complex system happens to be difficult to verify manually, since there is no corresponding mathematical description. A tool that enables one to define a model of the system, and to perform its verification is therefore an extremely valuable help.

This paper presents the formal specification and simulation of the DAQ of the ALICE experiment. A modelling and simulation tool, called Foresight, is used to specify the system in an abstract manner (system-level) in order to focus on the functionality.

II. FORESIGHT

A Foresight "specification" [1] is made of hierarchical data flow diagrams, finite state diagrams, and pieces of a procedural modelling language. The specification provides a unambiguous description of the system. The semantics of the specification provides a model of the system whose behaviour is very close to the behaviour of the system. The verification process is performed during the simulation. It demonstrates the functional correctness of the system.

The Foresight simulation consists of the real-time execution of the specification. It offers debugging functions like animation of diagrams, breakpoints, and monitor windows. The simulation is used to evaluate the performances of the specified system. It also makes it possible to perform some analysis such as the system sensitivity to some key parameters. One can also explore other algorithms, and new architectures. This is useful when the final architecture has not yet been defined (as is the case for ALICE), since it helps to compare architectures or choices of implementation.

III. SPECIFICATION

The current ALICE specification describes the *functionality* of the whole experiment and of the major sub-systems: Trigger, Trigger Detectors, Tracking Detectors, DAQ, Permanent Data Storage. The specification follows the description of ALICE DAQ given in [2], using up-to-date parameters values.

A. Overall System

The top-level Foresight specification of the ALICE DAQ system is made of the data flow diagram of Figure 1. Sub-systems are specified with Foresight processes. Links between sub-systems stand for the input/output interfaces among sub-systems.

The Interaction Source element generates events at a rate of 6000 Hz distributed according to a Poisson distribution. Trigger detectors receive the event signal emitted by the Interaction Source and inform the Trigger System. The latter emits L0, L1, L2 signals towards the Tracking Detectors. According to the received signals from the trigger system, the tracking detectors send data to the DAQ Sys (DAQ sub-system), and emit busy signals to the Trigger System. The DAQ Sys performs the sub-event building, then the event building and finally sends the events to the PDS Sys (Permanent Data Storage).

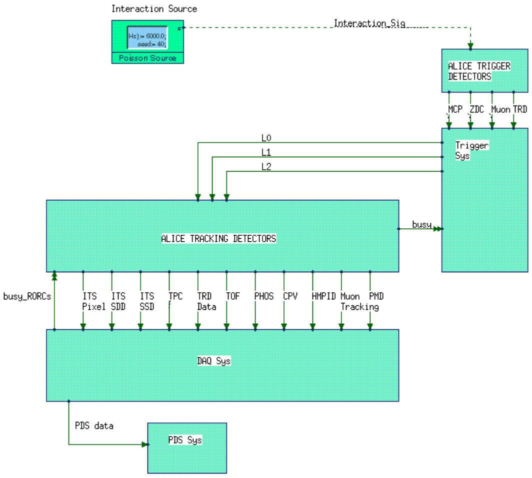

Figure 1: Overall Architecture

B. *Trigger System*

The ALICE trigger system has three levels: level 0 (L0), level 1 (L1), and level 2 (L2).

- L0 serves to strobe the detectors at $t_0 + 1.2\mu s$, where t_0 is the event time. L0 performs past/future (P/F) protection at t_0, i.e., no other interaction must have taken place during a given time interval before and after the occurrence of the current event. The time interval depends on the type of the event. At L0, we have simulated the case where an event can be either Central, Dimuon, Dielectron, Minbias, or a combination of two or more of these types. In the case of a combined event, two or more classes of detectors may be strobed. A class of detectors is the set of detectors involved in an event of a given type. L0 checks the busy flag of the detectors concerned by the current event. If one of them is busy, the classes of detectors to which this detector belongs are not strobed. If no class of detector can be strobed, then L0 does not send any signal to the detectors. Otherwise, the remaining classes of detectors receive the L0 signal.

- L1 performs a new P/F protection for the period up to $t_0 + 4.3\mu s$. If the event can be accepted as it is, the corresponding classes of detectors are informed at $t_0 + 5.5\mu s$. The trigger numbers are distributed by L1. In the case of a combined event, the P/F protection may act as a filter, so that only some classes of detectors are allowed to continue with the event. In this case, only some classes of detectors receive an L1 accept. If no L1 is generated, the event is rejected;

- Following the L1, level L2 makes the final trigger decision (no more combinations). It performs a P/F protection for the period up to $t_0 + 88\mu s$, and detectors are informed at $t_0 + 89.2\mu s$. The P/F protection may remove some classes of detectors. If after P/F protection, the final trigger outcome still includes more than one class, one of the participating classes is chosen proportionally to the required DAQ rates. The remaining class of detectors receives an L2accept signal, the others receive an L2reject signal.

- L0, L1, L2 signals always arrive in this order. It may happen that L0', L1', for a consecutive event, arrive before L2 (i.e., the sequence L0 L1 L0' L1' L2 is allowed). However L2 arrives always before L2', since level 2 signals reach the detectors at $t_0 + 89.2\mu s$.

1) *P/F protection Intervals*

For all trigger types, no other interaction can take place during the P/F protection interval $[t_0 - \Delta, t_0 + \Delta]$, where Δ is the P/F protection time. The parameters used for the P/F protection intervals are the following:

Dimuon: $\Delta = 3\mu s$
Dielectron: $\Delta = 7\mu s$
Combination(Dimuon,Dielectron): $\Delta = 7\mu s$
Central, Minbias, Other combinations: $\Delta = 88\mu s$.

2) *Event Rates at L0 Input*

The event rates produced by the collisions and received at the L0 level have been estimated as follows. The total number of interactions that have to be taken into account for P/F protection is set to 6000 Hz. Only 4000 Hz are interesting at the physics level (physics events). All interactions participate in the P/F protection process, i.e., their occurrence may spoil another event.

Table 1 shows the distribution of the 4000 Hz of interesting physics events among the different classes of triggers. For instance, the rate of Central events has been estimated to be 1000 Hz, distributed in the following manner: 403 Hz are "pure" Central events, 272 Hz are a combination of Central and Dimuon events, 272 Hz a combination of Central and Dielectron events, and 53 Hz are given by the combination of Central, Dimuon and Dielectron events. Minbias events do not combine with other events.

Table 1: Event Rates at L0 Input

	C	DM	DIEL	All	Total
C	**403**	272	272	53	1000 Hz
DM	**272**	**272**	53	53	650 Hz
DIEL	**272**	**53**	**272**	**53**	650 Hz
MB	**2403**				2403 Hz
Sub-Total					**4000** Hz
Others					**2000** Hz
Total					**6000** Hz

C. Tracking Detectors

The functional behaviour of the tracking detectors has been specified in the following manner:

- As soon as it receives a L0 signal, a detector becomes "busy". The trigger system does not send another L0 signal to a busy detector;
- The detectors wait for a L1 signal that has to arrive at $t_0 + 5.5\mu s$. If there is no L1 signal, then the detectors reset the collected data, and become available (not busy). If the L1 signal arrives, the detectors remain busy until the end of the reading period, unless the multi-event buffer is full (all event slots occupied);
- The data are stored in a multi-event buffer (one buffer for each DDL). In this simulation, every detector has a multi-event buffer of 4 positions;
- Following the L1, the detectors wait for the L2 decision. If it is an L2reject, the data are discarded and a slot is freed in the multi-event buffer (this operation takes $1.0\mu s$). If it is an L2accept, the data are transferred to the DAQ over the corresponding DDLs. Multi-event buffers are FIFO queues: data are sent or discarded in the order they were received - no "event overtaking" is allowed.

D. DAQ Sub-system

Figure 2 depicts the schematic view of the DAQ sub-system, showing the case of a detector having 20 DDLs. The elements in this figure are described below:

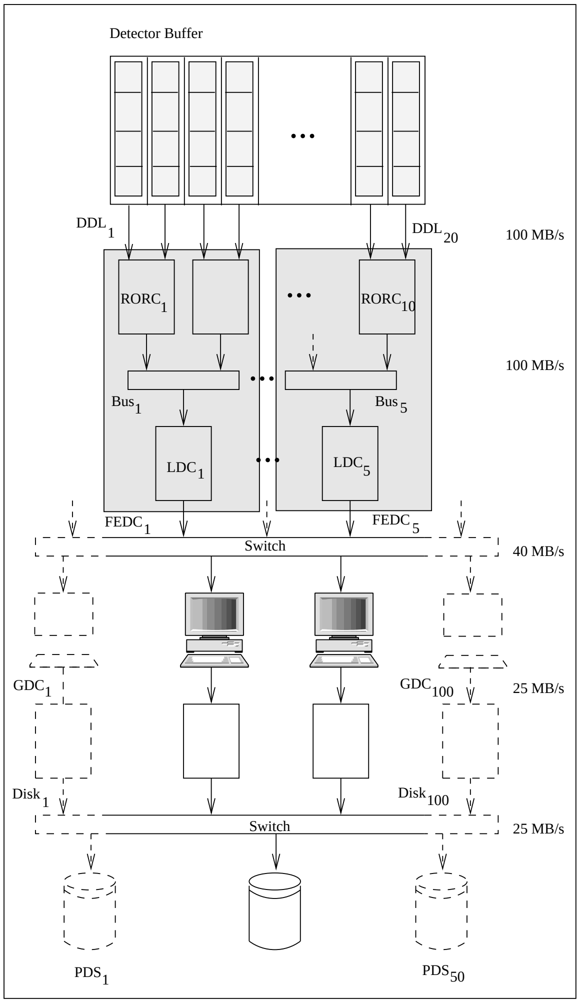

Figure 2: DAQ Sub-System

- A *Detector Data Link (DDL)* transfers the data from a buffer of the detector to the DAQ, at the rate of 100 MBytes/s. Data arrive in a Read-out Receiver Card (RORC). Several DDLs may be connected to the same RORC. In the example, two DDLs are connected to each RORC. Data are sent independently and in parallel along every DDL. If the corresponding RORC is full, the buffer slot cannot be released;
- A *Read-Out Receiver Card (RORC)* stores the data received from the DDL. A bus of 100 MBytes/s is used to transfer the data from a RORC to a Local Data Concentrator (LDC). Several RORCs can be plugged in the same LDC (two RORCs, in the example). Data are transferred to the LDC in the received order. The capacity of a RORC is set to 12 MBytes. If a LDC is full, the corresponding RORCs are blocked;

- The *bus* transports the data from RORCs to an LDC. The bus is connected at most to one LDC. It sends data sequentially to the LDC. Therefore, a RORC, which wants to send data, has to wait upon the completion of the sending of the data of the preceding RORC;

- A *Local Data Concentrator (LDC)* is responsible for sub-event building. As soon as a LDC receives all the data corresponding to an event, it builds the sub-event and sends it to a given Global Data Collector (GDC). In the example, every LDC receives four pieces of data (two from each connected RORC) before sending it to the GDC. Sub-events are sent in chronological order to the GDC. The capacity of a LDC is set to 128 MBytes. Every LDC sends the data of a given event to the same GDC. The choice of the GDC for a trigger event number n is: n MOD 100 (where 100 is the number of GDCs). Sub-events are sent to a Switch that forwards the sub-event to the GDC at a rate of 40 MBytes/s. If the corresponding GDC is full, the LDC is blocked. All LDCs and GDCs are linked to the same Switch;

- The *Front-End Digital Crate (FEDC)* is the set made of one LDC, the bus and the connected RORCs;

- A *Global Data Collector (GDC)* is responsible for the event building. It waits for all LDCs to send the sub-events. Once the event is complete, the whole event is stored on a disk. Events are sent in chronological order. Each GDC has its own disk. The link between the GDC and the disk has a rate of 25 MBytes/s. The GDC capacity is 512 MBytes;

- There is one *Disk* for each GDC, used for storing events. As soon as a file of 1 GBytes is built on the disk, the whole file is sent to an available Permanent Data Storage (PDS). Files are sent to a Switch that forwards the data to the chosen PDS, at a rate of 25 MBytes/s. All disks and PDSs are linked to the same Switch;

- A *Permanent Data Storage (PDS)* receives files of 1 GBytes from the DAQ. A PDS is considered to be an infinite buffer. There are 50 PDSs.

IV. RESULTS

The Foresight specification described above has been simulated with the DAQ sub-system being able to absorb the full rate of the DDLs. A second simulation, where the DAQ sub-system is modelled in more details and offers a restricted bandwidth to the tracking detectors is currently being performed, but results are not yet available.

In the first case, the DAQ sub-system actually works as an infinite buffer capable of absorbing all data coming from the tracking detectors. Detectors send data to the DAQ at a rate which is equal to the number of DDLs at their disposal times the DDL rate (100 MBytes/s). For instance, in the case of the TPC detector, data are sent at a rate of 18000 MBytes/s, since TPC has 180 DDLs.

Table 2 summarises the results of the trigger outputs observed during this simulation.

L0 rate depends upon the rate of the L0 trigger inputs (Table 1), P/F protection, and the busy status of the detectors. The high number of Minbias events is the result of the high rate of Minbias L0 trigger inputs. These Minbias events feed into the final Dimuon event rate. Indeed, while TPC and TRD are busy with central and Minbias events, Dimuon events can still be accepted by other detectors (TPC, and TRD are not required for the Dimuon class).

In this simulation, L1 outputs only take into account P/F protection. L0 and L1 rates are similar since it has been assumed that dielectron events are identified at L0. In practice this information is available at L1. The final L2 rates are not affected much by this simplification.

In Table 2, L2 rates for Dimuon and Dielectron events appear higher than L1 rates. As a result of re-classification, some of the combined events, in the row "Misc" of Table 2, become Dimuon or Dielectron events.

If we compare the column of L2 outputs obtained with the expected L2 rates for ALICE, given by Table 3, we notice that Dimuon events are very well represented. It is important to notice that the event rate at L0 input for Dimuon is 650 Hz (combined with other events). Therefore, with such an input rate, it is impossible (due to P/F protection) that we obtain the expected rate of 650 Hz at L2 output.

Table 2: Event Rates with Full DDL Bandwidth

	L0	L1	L2	L0%	L1%	L2%
Central	79	78	58	1.3%	1.3%	1.0%
Dimuon	523	517	602	8.7%	8.5%	9.9%
Dielectron	108	104	143	1.8%	1.7%	2.4%
Minbias	535	516	314	8.7%	8.5%	5.2%
Misc	146	142		2.4%	2.3%	
Sub-total	1391			23.0%		
Others	1997			33.0%		
Total	6038			100%		

Table 3: Expected Rates at L2

	C	MB	DM	DIEL	Total
L2	20	20	650	200	890 Hz

The simulation assumes each detector uses the full DDL bandwidth. Therefore, the obtained rates are the *maximum* rates that we can expect from the experience, if we consider the input event rates of Table 1, and the chosen detectors and DAQ sub-system parameters.

V. Conclusion

The next step is related to the establishment of the event rates with restricted bandwidth. Future work will focus on more detailed evaluation of the DAQ performances, checking of key parameters, and the evaluation of different architectural choices.

A. Restricted Bandwidth

The DAQ sub-system described in Section III. has been fully specified with Foresight. The simulation of this model will provide the event rates for the *restricted* bandwidth. The precise establishment of these rates is currently under work. In addition, this simulatin will allow to verify that the DAQ is able to sustain the maximum storage bandwidth of up to 1.25 GBytes/s.

B. DAQ Performances

Statistics on the buffer occupancy for the given parameters can be obtained. For instance, we already notice that, with the given parameters, TPC and TRD detectors frequently have all buffers full. This should be more accurately verified.

C. Architecture Alternatives

In the current model, the L2 trigger outcome reaches the detectors at $t_0 + 89.2\mu s$. It should be interesting to see, if sending the L2 outcome as soon as it is available, changes significantly the Dimuon rate.

The current algorithm for the GDC choice takes into account the trigger number, but not the availability of the GDCs. This algorithm is currently used for the DAQ sub-system in ALICE. With such an algorithm, it may happen that LDCs are blocked, waiting for a specific GDC, which is busy, while there is another GDC completely free. A more interesting algorithm would be to take into account the availability of the GDCs.

D. More Detailed Model

Till now, we have focused on the functionality of the different sub-systems of the ALICE DAQ and trigger systems. The model will be enriched with a more detailed specification of some existing DAQ components such as the ALICE Detector Data Link (DDL) or the DAQ software framework (DATE).

References

[1] Foresight-Systems, inc., Austin, TX 78759. *Foresight User's Guide*. http://www.nuthena.com.

[2] O. Villalobos Baillie, D. Swoboda, and P. Vande Vyvre. Data Acquisition, Control and Trigger: Common Report for the preparation of the ALICE Technical Design Reports. Internal Note DAQ, DCS, Trigger, ALICE/98-23, CERN, 1999.