A new readout control system for the LHCb upgrade at CERN

# A new readout control system for the LHCb upgrade at CERN

**F. Alessio**[1] **and R. Jacobsson**

*CERN,*
*CH-1211, Switzerland*

*E-mail:* Federico.Alessio@cern.ch

ABSTRACT: The LHCb experiment has proposed an upgrade towards a full 40 MHz readout system in order to run between five and ten times its initial design luminosity. The entire readout architecture will be upgraded in order to cope with higher sub-detector occupancies, higher rate and higher network load. In this paper, we describe the architecture, functionalities and a first hardware implementation of a new fast Readout Control system for the LHCb upgrade, which will be entirely based on FPGAs and bi-directional links. We also outline the real-time implementations of the new Readout Control system, together with solutions on how to handle the synchronous distribution of timing and synchronous information to the complex upgraded LHCb readout architecture. One section will also be dedicated to the control and usage of the newly developed CERN GBT chipset to transmit fast and slow control commands to the upgraded LHCb Front-End electronics. At the end, we outline the plans for the deployment of the system in the global LHCb upgrade readout architecture.

KEYWORDS: Control and monitor systems online; Digital signal processing (DSP); Data acquisition concepts; Trigger concepts and systems (hardware and software)

---

[1]Corresponding author.

## Contents

## 1   The upgrade of the LHCb readout architecture

The LHCb experiment at CERN has submitted a Letter of Intent for an LHCb Upgrade [1] which would allow operating the experiment at luminosity between five and ten times the current design after 2018. It will allow improving the trigger efficiencies in order to collect more than ten times the statistics foreseen in the first phase. Improving the trigger efficiencies requires in practice reading out the entire detector at the full 40 MHz LHC bunch crossing frequency with the consequence that practically all readout electronics have to be replaced [2]. The only exception is the current first-level trigger electronics [3] which already operates at the full frequency. In the upgraded scenario the first-level trigger is referred to as the Low Level Trigger (LLT). It compiles a physics decision per LHC bunch crossing based on the hadron, electron and muon candidates and passes it to the new LHCb Readout Control system (Timing and Fast Control system or, shortly, S-TFC).

Figure 1 shows schematically the upgraded LHCb readout architecture. All Front-End electronics (FE) record and transmit data continuously at 40 MHz to the Readout Boards, while the muon and calorimeter detectors also transmit information about the event to the LLT in parallel. The expected non-zero suppressed event size would result in a very large number of links between the FE and the new Readout Boards. In this regard, it has been shown [2] that almost a factor of ten could be gained by sending zero-suppressed data already at the FE. The zero-suppression will thus be performed in radiation-hard FE chips. The consequence of the varying zero-suppression time is that data are transmitted asynchronously to the Readout Boards and each data frame includes an event identifier in order to realign the event fragments in the Readout Boards.

The new Readout Boards are referred to as TELL40 in the upgraded architecture. A proposal [4] for a new TELL40 board based on ATCA technology has been accepted as the baseline by the LHCb Collaboration. About 15000 optical links will be used for the readout between the
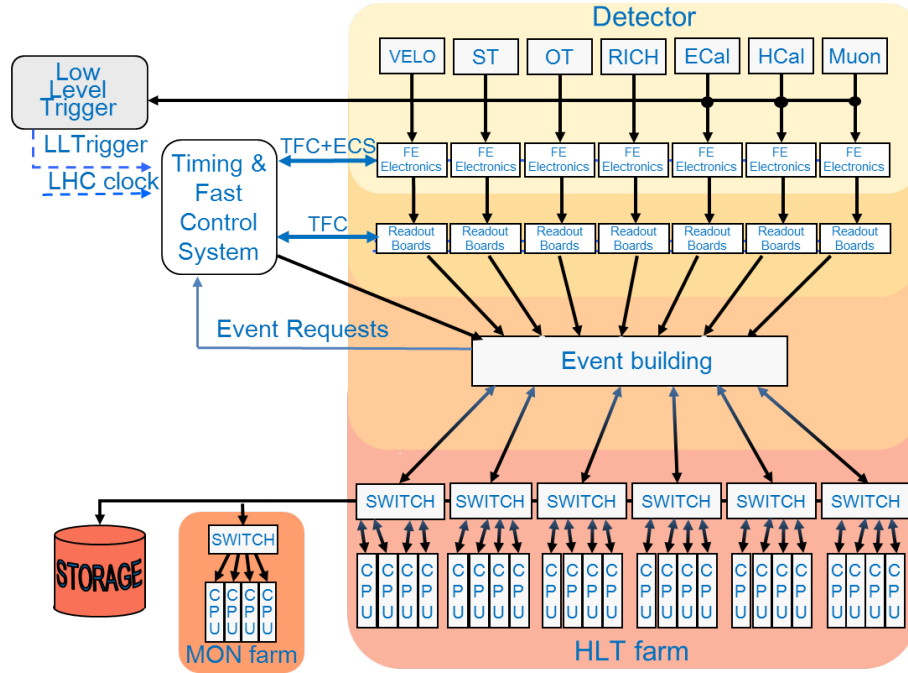
**Figure 1**. The upgraded LHCb readout architecture.

FE and a set of about 500 new TELL40s, by profiting from the CERN development of the Giga-Bit Transceiver (GBT [5]). The GBT is a common CERN Ser-Des rad-hard chipset for data, fast and slow control distribution to and from the FE electronics for the upgraded LHC experiments. The TELL40s will also act as interfaces to the event-building multi-Terabit/s network which connects the detector readout to the EFF. The EFF is to be based on COTS multi-cores and it will be responsible to process the events in order to select at least ∼20 kHz of them to be written to storage.

The new S-TFC system will play a central role in the upgraded readout architecture by distributing timing and fast commands to the TELL40s and to the FE electronics and by managing the dispatching of events to the EFF. It will also be responsible to handle back-pressure (*trigger throttle*) due to readout load and to interface the LHCb readout architecture with the LHC clock and LHC interfaces for synchronicity with the LHC accelerator operation. It will ultimately decide whether to accept an event basing its decision on the LLT decision, on the network load and on processing nodes availability. Lastly, the new S-TFC system will also profit from the bidirectional capability of the CERN GBT by interfacing the FE electronics to the global LHCb Experimental Control System (ECS). This will be done by transmitting slow control configuration data to the FE electronics and receiving slow control monitoring data from the FE electronics and relaying it back to the global LHCb ECS.

The experience with the current Timing and Fast Control system [6] allows a critical examination and inheriting features which are viable in the LHCb upgrade and which have evolved and matured over already ten years and are outlined throughout this paper, with particular attention to the novel approach of transmitting slow control information via the GBT chip to the FE electronics.

## 2 Readout control system functional requirements

Based on the role that the S-TFC system will play in the upgraded readout architecture, a list of the global functions which the new S-TFC system must support is given below.

Since the system must be ready before the whole readout electronics in order to be used in the development of the sub-detector electronics and detector test beams, the ultimate requirements are obviously flexibility and versatility for any changes in the readout strategy which may be decided on later.

**Bidirectional communication network.**  The new TFC network must allow distributing synchronous information to all parts of the readout electronics and must allow collecting trigger and buffer status information to be used for rate control.

**Low clock jitter, and deterministic phase and latency control.**  The synchronous distribution system must allow transmitting a clock to the readout electronics with a known and stable phase. The transmitted clock must have a jitter at each destination well within the specifications of what is needed for the high-speed data links from FE electronics. It must also allow fully controlling and maintaining stable the latency of the distributed information.

**Partitioning.**  The S-TFC architecture must allow partitioning, that is the possibility of running autonomously one or any ensemble of sub-detectors in a special running mode independently of all the others. In practice, this means that the new S-TFC system should contain a set of independent S-TFC masters, each of which may be invoked for local sub-detector activities or used to run the whole LHCb detector in a global data taking. A configurable switch fabric in the TFC communication network should associate the selected master to the ensemble.

**LHC interface.**  The system must be able to receive and operate directly with the LHC clock and revolution frequency, and must allow full control of the exact phase of the received clock.

**Events rate control.**  The new system should allow controlling the rate, either relying on physics decisions from the LLT or on non-biased trigger rejection such as throttling from the TELL40s or the EFF. At the simplest level, the rate control should be based on the actual LHC collisions scheme, i.e. the scheme which describes how the various bunches of protons are distributed along the LHC ring and which allows collisions at the LHCb interaction point.

**Low Level Trigger input.**  There should be means to interface the first-level trigger with the new S-TFC system.

**Support for old TTC-based distribution.**  In order to replace the current readout electronics and commission the new electronics in steps, and make use of the current first-level trigger system, the new S-TFC system must support the old CERN TTC system [7]. This will allow also operating a *hybrid* system.

**Destination control for the events packets.** The system should provide means to synchronously distribute the EFF destination to the Readout Boards for each event packet. In fact, in the current system ($n$) events are packed in a Multi-Event Packet (MEP) to reduce the overhead from Ethernet transmission. This function should also include the request mechanism by which the EFF nodes declare themselves ready to the S-TFC to receive the next events for processing. The event transfer from the Readout Boards is thus a push scheme with a passive pull, akin to a credit-based system. The scheme avoids the risk of sending events to non-functional nodes and produces a level of load balancing as well as a rate control in the intermediate upgrade phase with a staged farm. Ultimately this would rather be the only emergency control of the rate when the system has been fully upgraded to the 40 MHz readout.

The mechanism is already in place in the current system and it is proposed to be kept as in [8].

**Event data bank.** An event data bank containing the information about the identity of an event (Run Number, Orbit Number, Event ID, and UTC Time) and trigger source information is currently produced by the current TFC system and added to each event [9]. A similar block should also be produced in the new S-TFC system.

**Sub-detector calibration triggers.** The system must allow generating sub-detector calibration triggers by transmitting synchronous commands to the FE electronics.

**Test-bench support.** The system and its components must be built in a way that they can be used stand-alone in small test-benches and test-beams, and they have to be made available at an early stage in the development of the FE and TELL40 logic.

## 3   Readout control system architecture

Based on the system requirements, the new S-TFC system has been outlined and specified. The logical scheme of the new S-TFC architecture and the data flow is represented in figure 2, in which the S-TFC components are shaded.

A *Readout Supervisor* (S-ODIN) is responsible for controlling the entire upgraded readout scheme by distributing timing and synchronous commands. The commands maintain synchronicity of the system, provide the mechanism for special monitoring triggers, manage the dispatching of the events to the EFF and regulate the transmission of events through the entire readout chain taking into account buffer occupancies, throttles from the TELL40s, the LHC collisions scheme and the physics decisions from the LLT.

The sub-detector readout electronics is connected to the Readout Supervisor via a set of 2.4Gb/s high-speed bidirectional optical links located on an interface board (called TFC+ECSInterface) positioned in each of the TELL40 crates. This board serves two main purposes:

1. Interface all the TELL40 boards of a crate to the S-ODIN by *fanning-out* the synchronous TFC information to the TELL40 boards and *fanning-in* throttle information.

2. Interface all the FE electronics to the S-ODIN by relaying the TFC information onto fibres towards the FE electronics via the GBT chipset. In addition, the flexibility of the optical link and the hardware architecture also allow accommodating the function of relaying ECS
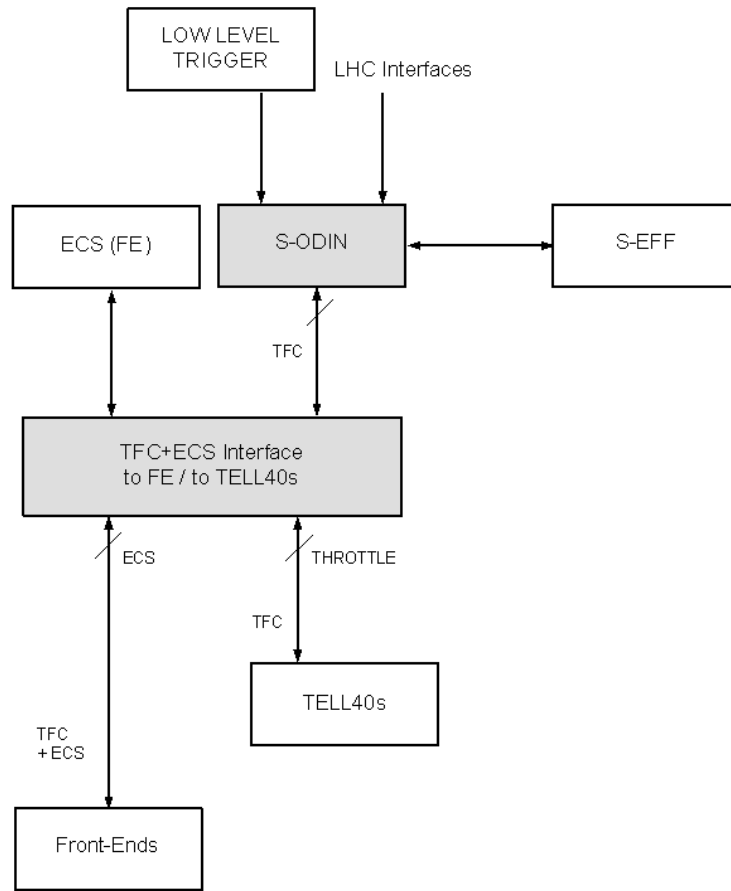
**Figure 2**. Logical architecture of the new S-TFC system with the TFC information dataflow.

configuration and control data to the FE, and use the return path for read-back and receiving monitoring data back.

The TFC+ECSInterface boards may be cascaded and configured differently to support different requirements in terms of number of links and bandwidth and accommodate the granularity of the system supporting partitioning.

A proposal for the use of the ATCA technology as the main backbone for the upgraded readout system has become the baseline solution for the LHCb upgrade. A common flexible ATCA electronics board [4] has been envisaged to be the main component of the upgraded readout architecture in order to minimize the development of custom-made electronics for the upgrade. The board's architecture essentially consists of an ATCA motherboard with slots for four AMC cards. The motherboard provides the powering, the clock fan-out and the interface to the control system. It incorporates a dense fabric of interconnectivity between the four AMC slots, in addition to an access to the ATCA backplane bus via a crossbar. The proposed dual-star topology backplane allows distributing a clock and provide two sets of point-to-point serial links from the hub board slots in the ATCA crate. Each of the four AMC cards is composed of a large FPGA - ALTERA Stratix V or later in the final stage, while Stratix IV is used in the prototyping stage - with on one side
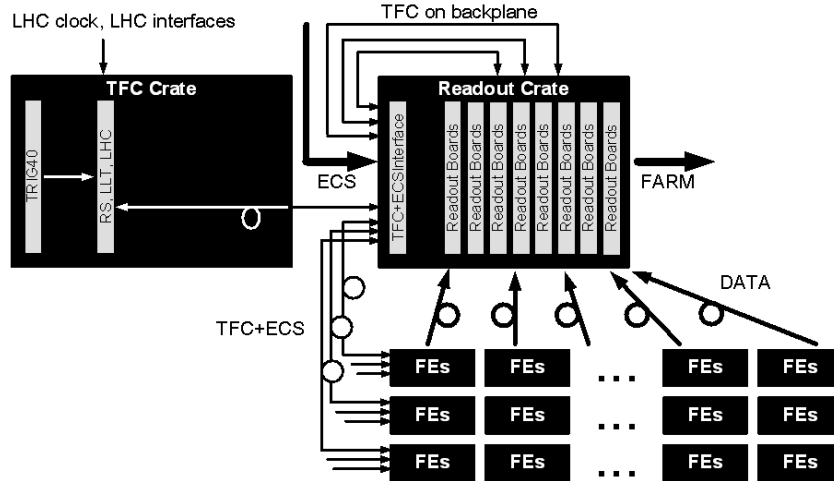
**Figure 3**. Physical architecture of a partition of the upgraded readout system using ATCA technologies.

the bus to the motherboard interconnectivity and switch fabric and on the other end 3x12 optical bi-directional transceivers with MPO connectors. Each of the transceivers may be implemented to operate the custom-made protocol for the CERN GBT transceiver as the transmission of data from the FE electronics to the readout board will be done using the GBT chipset. Such a board is flexible enough to acquire different purpose by simply reprogramming its FPGAs with special-purpose logic. Therefore, each of the AMC cards can be customized to a particular task.

For example, in the case of the ATCA board which will host the main Readout Supervisor logic, only one AMC card will be configured to carry the Readout Supervisor firmware. Thanks to the flexibility of ATCA board, another AMC card can be programmed differently to assume the role of the upgraded hardware trigger LLT, working in tight connection with the Readout Supervisor via the fabric interconnections on the ATCA motherboards. An AMC slot will be reserved for the LHC interfaces but it is likely to require dedicated hardware design. Figure 3 shows the physical architecture of a partition of the upgraded readout system with ATCA technologies highlighting the various connections between the main components of the system.

Moreover, the new S-TFC should support the old TTC protocol currently in place in order to allow hybrid operation with a slice of the upgraded readout architecture in parallel to the current readout architecture, and in order to maintain the old calorimeter and muon trigger logic for the LLT. The support for this hybrid functionality can be implemented via a PICMG3.8 compatible Rear Transition Module (RTM) which would allow connection to the old Readout Supervisor, fitting on the rear part of the ATCA crate.

The tasks of the TFC+ECSInterface board to relay and fan-out the TFC protocol and fan-in the throttle protocol can be implemented in one of the AMC cards together with a single stand-alone Readout Supervisor instantiation for local tests. The communication with the TELL40 boards and the distribution of the TFC clock are ensured by the access to the bus on the backplane instead of individual optical links. For this reason, the TFC+ECSInterface should be located in the ATCA hub position of the TELL40 crates so that it can have access to all the TELL40 sitting in the corresponding Readout Crate.

The TFC+ECSInterface boards are also used for the second purpose of transmitting the timing, clock and commands to the FE electronics. Since the estimated TFC bandwidth requirements to the FE are low, the rest of the bandwidth may be used for ECS configuration and control of the FE electronics. The card therefore also incorporates the logic to transmit ECS information to the FE electronics and to receive back ECS data and monitoring information. As a matter of design, the other three AMC cards on the same ATCA board may contain the same logic as the first and according to the need of a particular sub-detector they may provide more TFC+ECS links to the FE to cover the complete set of FE electronics connected to the TELL40s in the crate. Thus, a single TFC+ECSInterface could potentially drive as many as 132 bidirectional FE links with TFC and ECS for a total potential of just above 10Gb/s of user data.

## 4    Timing and fast commands distribution to the entire readout architecture

In the upgraded scenario, the S-ODIN receives directly the LHC clocks via an LHC Interface card on the ATCA motherboard. These clocks make up the global master clocks of the entire LHCb readout system. The presence and stability of the clocks are monitored locally. A local quartz-based PLL circuit allows providing a stable continuous bunch clock and a digital PLL allows recovering the turn signal, which is used to define the length of a full LHC turn, in case of a temporary transmission problem. In case the external clocks are absent, these circuits provide local clocks to allow LHCb to operate the readout system for calibrations and tests.

The local timing distribution consists firstly of distributing the LHC bunch clock to all electronics modules operating in the readout. The distribution must satisfy the strict requirements of sufficiently low jitter ( $O$(50ps) peak-to-peak), and a stable, reproducible, and controllable fine phase across the entire distribution chain up to each destination. The low jitter is required for the high-speed data links to function and the phase allows adjusting the detector signal sampling to the optimal point. The distribution is implemented by means of clock and data recovery (CDR) on two types of optical serial links, using the GBT protocol. The first type uses commercial FPGA-based transceivers for both transmission and reception, and the second uses commercial FPGA-based transceivers for transmission while the reception is handled by the CERN GBT. In addition, the clock distribution between the TFC+ECSInterface and the TELL40 boards is done by means of the ATCA backplane.

Secondly, the readout synchronization is achieved by distributing the LHC turn signal in the form of a Bunch Counter Reset command. All the readout control logic and the synchronous TFC readout control commands are aligned to an LHC turn signal across the entire distribution chain up to each destination. Since the commands are encoded across the optical links, the distribution must satisfy the strict requirement of stable, reproducible, and controllable transmission latency, where latency is defined in terms of LHC bunch clock cycles. The transmission between the TFC+ECSInterface and the TELL40 boards is ensured by serial busses on the ATCA backplane and is also subject to these requirements.

The GBT chipset is meant to satisfy all of these requirements, solving the problem of a constant latency and fine phase at the FE, but special implementations and validation tests are required for the commercial FPGA-based transceivers, as well as the electrical transmission across the ATCA backplane.

Based on these considerations, the main critical area has been identified in the link between S-ODIN and the TFC+ECSInterface boards. This is the most critical link in the system as COTS FPGA-to-FPGA transceivers will be used for this link. Moreover, the S-ODIN information must be received by all the TFC+ECSInterface boards with a specific and deterministic latency which takes into account cable length and processing. For this reason, each link has adjustable delays.

An extensive system validation phase is currently ongoing to qualify each area of the timing distribution for the final system. In this regard, a custom-made S-TFC protocol has been specified in order to transmit fast commands and reset across the links and a validation phase of FPGA-to-FPGA transceivers using ALTERA FPGAs is ongoing. The protocol employs a 60bits GBT-like encoding stage, by profiting from scrambling, interleaving and forward error correction to make the transmission of such critical TFC commands as robust as possible, even if it is not done in a radiation critical environment. However, special configuration of the ALTERA FPGAs must be envisaged to obtain a deterministic phase and latency after the FPGA's CDR. These configurations have been identified in ALTERA Stratix IV FPGAs and a future validation test campaign will be performed on ALTERA Stratix V FPGAs.

A simple 8b/10b encoder can also be used over the same links profiting from the embedded encoder/decoder of the ALTERA FPGAs. It was already shown how the 8b/10b decoder in the ALTERA FPGAs is able to recover a deterministic and constant phase of the clock from the serial data stream [4].

## 5 Encoding of fast and slow control information to the FE electronics

Each TFC+ECSInterface board performs the function of distributing the timing and readout control commands together with the ECS information needed to configure and control the operation of the FE electronics which serves the TELL40 crate in which the TFC+ECSInterface is located.

Figure 4 shows schematically the implementation of such functionality in the TFC+ECSInterface board. A *TFC Relay and Alignment* logical block extracts 24 bits out of the TFC word which was transmitted by S-ODIN encoding the fast commands and resets. This word is used to reconstruct the clock locally in the FPGA to be then used to drive the logic inside the TFC+ECSInterface.

In parallel, a Credit-Card sized PC [10] which is interfaced to the global LHCb ECS [11] receives configuration and slow control information to be sent to the FE electronics. The received data is accompanied with an *extended addressing scheme* which allows routing the data to the correct GBT-link, and further to the correct FE chip, by means of the special GBT-SCA chip. This chip is part of the main GBT chipset and it is meant to be able to interface the GBT chipset with FE chips by providing different types of busses and slow control functionalities. The extended addressing scheme also includes a bus type to handle different types of data sequencing. A *PCIe Memory map* provides intermediate storage for the addresses and data while the *e-link Protocol Driver* actually drives the appropriate bit-sets in the GBT frame. This may also include special sequencing for different bus types.

In addition to the write commands for configuration and control, the ECS link also provides read commands. The return path of the FE link is reserved for receiving ECS data and monitoring
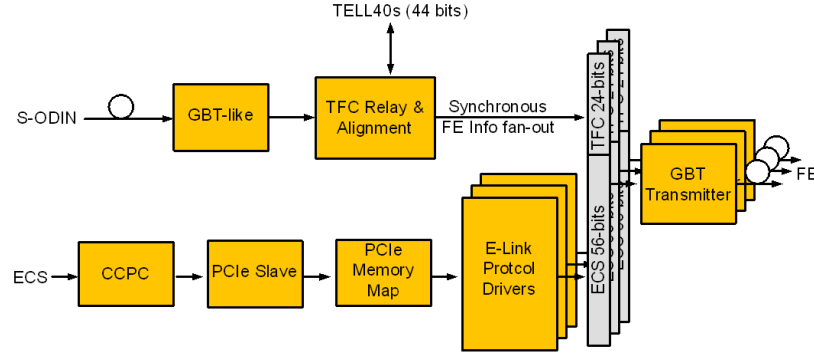
**Figure 4**. Schematic view of the packing algorithm to merge TFC and ECS information on the GBT link towards the FE electronics.

information. The continuous monitoring of counters and status registers is performed by polling, that is, explicit read commands (not drawn in figure 4).

## 6    Deployment of the system in the upgraded readout architecture

Due to the complexity of the S-TFC system and the strict timing requirements, a clock-level simulation and verification framework of the new readout architecture and VHDL code is being developed. This includes a detailed synthesizable simulation of all the new S-TFC components and emulation of the surrounding components such as the GBT-to-FPGA links, the TELL40 boards, the FE electronics, the ECS and the EFF for verification purposes.

Moreover, first validation tests are being performed using a first version of the AMC card developed at CPPM in Marseille. The AMC card is powered by a 12V power supply and it is controlled via an ALTERA USB-JTAG Blaster, connected to a PC interfacing the embedded processor on FPGA (NIOS II-based) to control the board hardware resources. The setup allowed demonstrating the feasibility of the system and allowed finding the right configuration for the ALTERA FPGA transceivers to obtain a deterministic and fixed latency for the serial data stream as described in section 4.

Lastly, the possibility of interfacing the new S-TFC system to the old-TTC based TFC system allows for *hybrid* operations of the readout architecture. The new electronics can be controlled via the new S-TFC system, while the old TTC-based electronics can be controlled via the current TFC system. This would allow installing an active test-bench during the Long Shutdown 1 to even improve the LHCb physics potential in the running period between 2015 and 2018 by installing an upgraded detector which can be read out via an upgraded readout slice.

## 7    Conclusions

In this paper, a new Fast Readout Control system for the upgrade of the LHCb readout architecture has been presented. The new functionalities have been listed and the new architecture presented in detail. As the system plays a central role in the upgraded readout architecture, it must be ready before the development of the new electronics for the upgraded LHCb detector. Therefore, a first hardware implementation is already ready and it is being fully validated.

The first version of the system is meant to be ready by Q1 of 2013 in order to allow for a first validation test run in parallel to the current LHCb detector and readout system during the course of 2013.

## References

[1] LHCB collaboration, *letter of intent for the LHCb upgrade*, CERN-LHCC-2011-001 (2011)

[2] K. Wyillie et al., *Electronics architecture for the LHCb upgrade*, LHCb-PUB-2011-001 (2011).

[3] R. Cornat et al., *Level-0 decision unit for LHCb*, LHCb-PUB-2003-065 (2003).

[4] J.P. Cachemiche et al., *Study for the LHCb upgrade readout board*, 2010 *JINST* **5** C12036.

[5] P. Moreira et al., *The GBT Ser-Des ASIC prototype*, 2010 *JINST* **5** C11022.

[6] R. Jacobsson, Z. Guzik and B. Jost, *Driving the LHCb front-end readout*, *IEEE Tr. Nucl. Sci.* **51** (2004) 508.

[7] S. Baron et al., *CERN TTC webpage*, http://ttc.web.cern.ch/TTC.

[8] R. Jacobsson, *Central FPGA-based destination and load control in the LHCb* 1 *MHz readout*, *Nucl. Instrum. Meth.* **668** (2012) 41.

[9] R. Jacobsson, *ODIN Raw Data Format v5.0*, CERN EDMS 704084 (2006).

[10] C. Gaspar, B. Jost and S. Schmeling, *The use of credit card-size PC for interfacing electronics boards in the LHCb ECS*, LHCb-PUB-2001-147 (2001).

[11] C. Gaspar et al., *An integrated experiment control system, architecture and benefits: the LHCb approach*, *IEEE Tr. Nucl. Sci* **51** (2004) 513.