

Timing and Readout Control in the LHCb Upgraded Readout System

Federico Alessio*, Joao Barbosa, Sophie Baron, Jean-Pierre Cachemiche, Cairo Caplan,
Clara Gaspar, Frederic Hachon, Richard Jacobsson, and Ken Wyllie

Abstract—In 2019, the LHCb experiment at CERN will undergo a major upgrade where its detectors electronics and entire readout system will be changed to read-out events at the full LHC rate of 40 MHz. In this paper, the new timing, trigger and readout control system for such upgrade is reviewed. Particular attention is given to the distribution of the clock, timing and synchronization information across the entire readout system using generic FTTH technology like Passive Optical Networks. Moreover the system will be responsible to generically control the Front-End electronics by transmitting configuration data and receiving monitoring data, offloading the software control system from the heavy task of manipulating complex protocols of thousands of Front-End electronics devices. The way in which this was implemented is here reviewed with a description of results from first implementations of the system, including usages in test-benches, implementation of techniques for timing distribution and latency control.

I. INTRODUCTION

THE LHCb experiment [1] is a high-precision experiment at the LHC devoted to the search for New Physics by precisely measuring its effects in CP violation and rare decays. By applying an indirect approach, LHCb is able to probe effects which are strongly suppressed by the Standard Model, such as those mediated by loop diagrams and involving flavor changing neutral currents. In the proton-proton collision mode, the LHC is to a large extent a heavy flavor factory producing over 100,000 bb-pairs every second at the nominal LHCb design luminosity of $2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$. Given that bb-pairs are predominantly produced in the forward or backward direction, the LHCb detector was designed as a forward spectrometer with the detector elements installed along the main LHC beam line, covering a pseudo-rapidity range of $2 < \eta < 5$ well complementing the other LHC detectors ranges. LHCb proved excellent performance in terms of data taking [2] and detector performance over the period 2010-2015 accumulating about 3.5 fb^{-1} of data and it is foreseen to accumulate about 5 fb^{-1} over the period 2016-2018. Due to the foreseen improved performance of the LHC accelerator, the prospect to augment the physics yield in the LHCb dataset seems very attractive. However, the LHCb detector is limited by design in terms of data bandwidth - 1 MHz instead of

the LHC bunch crossing frequency of 40 MHz - and physics yield for hadronic channels at the hardware trigger. Therefore, a Letter Of Intent [3], a Framework TDR [4] and a Trigger and Online TDR [5] document the plans for an upgraded detector which will enable LHCb to increase its physics yield in the decays with muons by a factor of 10, the yield for hadronic channels by a factor 20 and to collect about 50 fb^{-1} at a leveled constant luminosity of up to $2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$. This corresponds to ten times the current design luminosity and increased complexity (pileup) of a factor 5.

II. THE UPGRADE OF THE LHCb READOUT ARCHITECTURE

In order to remove the main design limitations of the current LHCb detector, the strategy for the upgrade of the LHCb experiment essentially consists of ultimately removing the first-level hardware trigger (L0 trigger) entirely, hence to run the detector fully trigger-less. By removing the L0 trigger, LHC events are recorded and transmitted from the Front-End electronics (FE) to the readout network at the full LHC bunch crossing rate of 40 MHz, resulting in a 40 Tb/s DAQ network. All events will therefore be available at the processing farm where a fully flexible software trigger will perform selection on events, with an overall output of about 20 kHz of events to disk. This will allow maximizing signal efficiencies at high event rates. The direct consequences of this approach are that some of the LHCb sub-detectors will need to be completely redesigned to cope with an average luminosity of $2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ and the whole LHCb detector will be equipped with completely new trigger-less FE electronics. In addition, the entire readout architecture must be redesigned in order to cope with the upgraded multi-Tb/s bandwidth and a full 40 MHz dataflow [6]. Figure 1 illustrates the upgraded LHCb readout architecture. It should be noted that although the final system will ultimately be fully trigger-less, a first-level hardware trigger based on the current L0 trigger will be maintained in software. This is commonly referred to as Software LLT and its main purpose is to allow a staging installation of the DAQ network, gradually increasing the readout rate from the current 1 MHz to the full and ultimate 40 MHz. This however will not change the rate of event recorded at the FE, which will run fully trigger-less regardless of the DAQ output rate.

In order to keep synchronicity across the readout system, to control the FE electronics and to distribute clock and synchronous information to the whole readout system, a centralized Timing and Fast Control system (TFC, highlighted in

Manuscript submitted May 30, 2016

J. Barbosa, S. Baron, C. Gaspar, R. Jacobsson and K. Wyllie are with CERN, Geneva, Switzerland

J.P. Cachemiche, F. Hachon are with CPPM, Marseille, France

C. Caplan is with CBPF, Rio de Janeiro, Brazil

*F. Alessio is the main author of this manuscript. At the time of the manuscript, the author is with CERN, Geneva, Switzerland. E-mail contact: federico.alessio@cern.ch

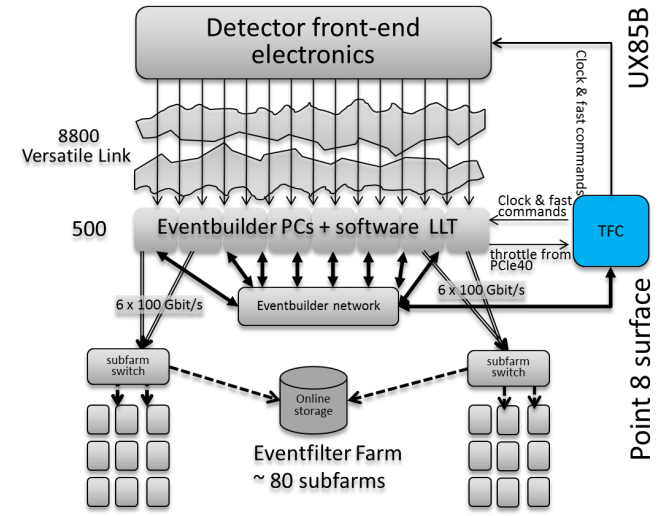


Fig. 1. The upgraded LHCb readout architecture.

Figure 1) has been envisaged, as an upgrade of the current TFC system [7]. The upgraded TFC system will then be interfaced to all elements in the readout architecture by heavily profiting from the bidirectional capability of optical links and FPGA transceivers and a high level of interconnectivity. In particular, the TFC system will heavily profit from the capabilities of the GigaBit Transceiver chipset (GBT) [8] currently being developed at CERN for its communication to the FE electronics. In addition, the TFC system will also be responsible to transmit slow control (ECS) information to the FE, by means of FPGA-based electronics cards interfaced to the global LHCb ECS.

III. THE UPGRADED LHCb TIMING AND READOUT CONTROL SYSTEM

Figure 2 illustrates in detail the logical architecture of the upgraded TFC system. A pool of Readout Supervisors (commonly referred to as S-ODIN) centrally manages the readout of events, by generating synchronous and asynchronous commands, by distributing the LHC clock and by managing the dispatching of events. Each S-ODIN is associated with a sub-detector partition which effectively is a cluster of Readout Boards (TELL40) and Interface Boards (SOL40). While the TELL40s are dedicated to read out fragments of events from the FE and send them to the DAQ for software processing, the SOL40 boards are dedicated to distribute fast and slow control to the FE, by relaying timing information and clock onto the optical link to the FE, and by appending ECS information onto the same data frame. Thanks to the characteristics of the GBT chipset [8], fast commands, clock and slow control are therefore transmitted on the same bidirectional optical link. This is a major novelty with respect to the current LHCb experiment where fast control and slow control are sent over different networks. At the FE, the synchronous fast control information are decoded and fanned out by a GBT Master per FE board, also responsible to recover and distribute the clock in a deterministic way. The slow control information is relayed to the GBT-SCA (which stands for Slow Control Adapter)

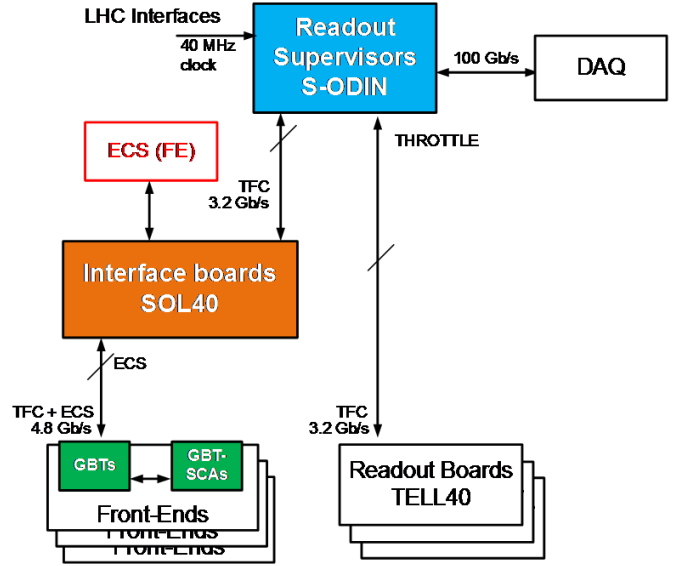


Fig. 2. The upgraded LHCb timing and readout control system architecture.

chip via the GBT Master. The GBT-SCA chip is capable of efficiently distributing ECS configuration data to the FE chips by means of a complete set of buses and interfaces, in a generic way [9]. Monitoring data is sent back on the uplink of the same optical link by following the return path, from the GBT-SCA to the Master GBT to the corresponding SOL40.

The hardware backbone of the entire readout architecture is a PCIe Gen3 electronics card hosted in a commercial PC. The same hardware is used for the TELL40, the SOL40 and the S-ODIN boards, only the different firmware changes the flavor of the board. The board will be equipped with up to 48 bidirectional optical link, an Altera Arria 10 FPGA (GX 1150) and a 16x PCIe Gen3 bus interfaced to a multi-core PC. The card is being currently being developed at CPPM in Marseille [10].

Figure 3 shows schematically the implementation of the merging of fast and slow control information on the same optical link to the FE electronics [11] in the firmware at the SOL40 board. A TFC Relay and Alignment block extracts at maximum 24 bits out of the full TFC word which was transmitted by S-ODIN encoding the fast commands, timing information and various resets. These 24 bits are then relayed onto the GBT link to be transmitted to the FE. The word is generated at 40 MHz and transmitted with constant latency. The TFC word from S-ODIN is used to reconstruct the clock locally in the FPGA to then be used to drive the logic in the firmware.

A. Timing and clock distribution in the LHC upgraded system

Due to the centralized nature of the LHCb Readout Supervisory system (TFC), its implementation within the upgrade of the LHCb experiment poses some challenges in terms of timing distribution, clock recovery, jitter, readout synchronization and ultimately robustness/reliability and control. The TFC system is in fact a single point of entry for synchronizing the LHCb experiment and its entire electronics - from the FE to

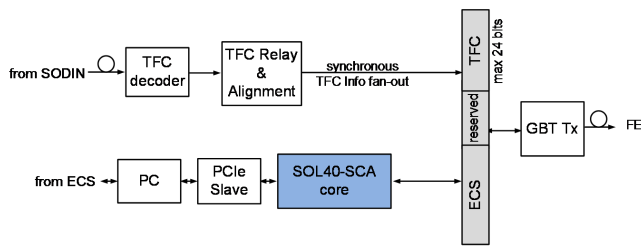


Fig. 3. Schematic view of the algorithm to merge TFC and ECS information on the GBT link towards the FE electronics in the SOL40 firmware.

the DAQ - to the LHC accelerator, by being interfaced to the main LHC 40 MHz clock and its timing information. The clock must be received, recovered, cleaned and fanned out to all elements in the readout architecture, down to the very last FE chip with a deterministic phase and constant latency. In addition, it must be monitored and controlled in a reliable way: once each detectors partition itself is locally time aligned, the timing of the experiment is adjusted globally to match the LHC beam structure and the fine phase of collisions, ideally with a final tolerance of below 100 ps. As a consequence, every clock source and clock reception device must move accordingly in a deterministic way, in order to have the full system globally aligned. This is simplified schematically in Figure 4 in the context of the LHCb upgrade. Particular care is given to clock crossing domain paths. In the context of the TFC system, this is particularly important in the SOL40 cards where the clock recovered from the TFC optical stream must also be used to drive the FPGA transceivers which fan out the timing information towards the FE chips. The FE chips ultimately must recover the clock and use it to sample detector data and drive their internal DSPs and analog electronics. At the SOL40 boards, only one optical link is used to connect the cards to the S-ODIN cards to receive timing information to be fanned out, while between 24 to 48 links are used to drive the timing information to the FE depending on the configuration of each subdetector's partition. Each transceiver must be configured, monitored and controlled so that it maintains the clock quality transmission, its phase and its latency. It is estimated that the whole TFC system will have to drive up to 2500 destinations at the FE in this way.

In order to fulfill the previous requirements, the architecture of the TFC system has been finalized following currently available technology and the developments from generic projects in our community. The architecture is illustrated in Figure 5, where only a detectors partition is highlighted: each partition will be controlled via a hierarchy composed of a single S-ODIN and a set of SOL40 boards, enough to cover the entire FE electronics Master GBTs. Each S-ODIN will act as a central Readout Supervisor for that partition, distributing the global clock, fast commands and reset, and other synchronous and asynchronous commands all based on configurable recipes. This will allow each partition to work and run independently from one another, while maintaining scalability. Synchronicity across all units in the partition is ensured by the usage of Passive Optical Network technology (PON), a technology used in the Fibers-to-The-Home (FTTH)

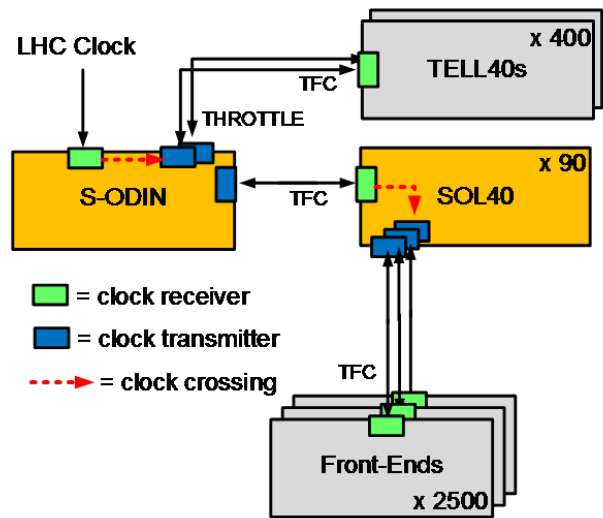


Fig. 4. Illustration of the clock distribution paths in the upgraded LHCb readout architecture.

where a passive optical splitter allows to reach many destinations from a single point of start without the need to re-drive the initial stream. Since the clock, timing and readout control signals are centrally generated and they are the same for all destinations in a partition, they can be distributed passively. A centralized effort at CERN [12] started looking into the possibility of using 10G-PON technology to replace the CERN wide Timing-Trigger and Control (TTC) legacy system at the experiments. The main advantages of having this technology in comparison to the current TTC system are bi-directionality, high number of destinations from a single point of start, high bandwidth, software partitioning and possibilities to use FPGA technology directly to recover the stream of data. This allows the possibility to tune the technology to each experiments needs. In practice, in the LHCb case, each partitions timing and readout control information are generated centrally in S-ODIN. These are then synchronized with the LHC main clock which is used to drive the transmitters at the FPGA. The stream of data is split across all destinations in the partitions, where clock and data are recovered with fixed latency and deterministic phase. At the TELL40, the clock and timing information are used to decode the stream of data coming from the detectors FE. At the SOL40, the clock and timing information - merged together with slow control information - are then relayed onto other FPGA transmitters, to finally reach the FE Master GBTs.

In addition, all SOL40 and TELL40 boards FPGAs are configured with a centralized GBT-FPGA core [13], whose aim is to generically be able to drive and receive data from any GBT chip, while maintaining constant phase and minimizing latency. The GBT-FPGA core includes critical features to be able to correctly drive the transmitters at the FE so that the stream of data at the GBT keeps information regarding the clock phase. Globally, a centralized shift at the LHC clock reception location will propagate correspondingly to the entire partition. This is then followed by a resynchronization mechanism (not described here as it specific to the LHCb

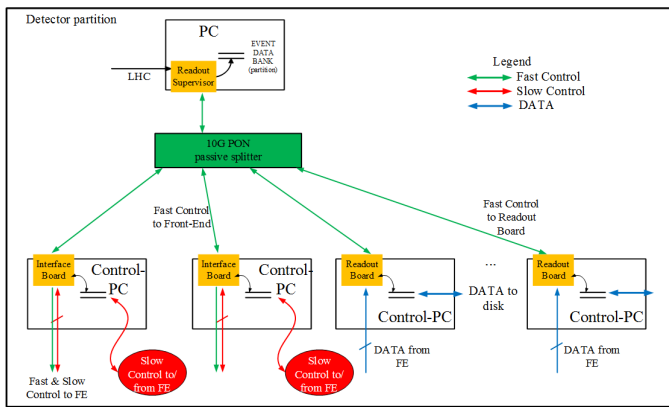


Fig. 5. Architecture of the TFC system for a sub-detector partition.

upgraded system) whose aim is to get all receivers locked again.

B. Critical aspects in the upgraded LHCb clock and timing distribution system

The most important aspect of choosing PON technology for the upgrade of the LHCb clock and timing distribution system is the possibility of reaching many destinations without the need of active fan-out and fan-in boards. This reduces drastically the complexity of the architecture as the clock needs to be recovered from an optical serial stream only once: at the SOL40/TELL40 boards. At the FE, the GBT chipset has robust clock recovery capabilities as the chip was designed to contain PLLs and CDR blocks for this specific purpose. In the case of the need of an active fan-out/fan-in, the clock would have had to be recovered twice before even being sent out to the FE, augmenting the risk of introducing jitter and noise. On the other hand, ad-hoc solutions must be envisaged [12] in order to fulfill the requirements in terms of clock recovery and latency. Moreover, the possibility of having bi-directionality available across the TTC network allows for a high level of interconnectivity. PON technology uses Time Division Multiple Access (TDMA) for the upstream signals. This implies that each sender is allocated a time slot and that the central receiver (S-ODIN in this particular context) needs to wait for each sender to be done transmitting its information. Again, an advanced study on this was done and the maximum round-trip time was limited to less than $5 \mu s$ for up to 128 destinations. This is perfectly compatible with the LHCb upgraded readout architecture as the upstream path is only used to transmit asynchronous busy information (throttle) back to S-ODIN for monitoring purposes. In addition, various techniques can be adopted to reduce such delay: not all senders must send at all times, thus reducing the effective round-trip delay or throttling could be centrally extended for more consecutive clock cycles rather than for just one single clock cycle. Global synchronicity during global running is ensured by a Master S-ODIN who is dedicated to centrally generate and distribute commands, reset and synchronous/asynchronous commands. The concept is illustrated in Figure 5. In this case, the fine and deterministic phase of the clock is maintained

as each partitions S-ODIN is interfaced to the LHC clock thus using it to drive its serializers towards the SOL40 and TELL40 boards. Constant latency is ensured by buffering timing and readout commands at the partitions S-ODINs in order to compensate for the delay in adding one level in the hierarchy.

IV. SLOW CONTROL TO THE FRONT-END ELECTRONICS IN THE LHCb UPGRADE

For what concern the slow control part to the FE electronics, LHCb has developed a firmware core, commonly referred to as SOL40-SCA, in order to generically drive each GBT-SCA chip at the FE, covering all of its functionality and protocols. Its location within the SOL40 firmware is highlighted in Figure 3. This is achieved by developing the firmware in a completely configurable way, i.e. the chosen SCA protocol can be selected in real-time via commands issued by the LHCb ECS system [14] together with the configuration data. The destination of such data can be selected via a configurable mask. The core is designed to cover a full GBT link with up to 16 GBT-SCAs connected to it. It can then be replicated as many times as needed to cover all GBT links connected to a SOL40 board. In total, the same firmware will allow driving generically the entire LHCb upgraded FE electronics over a total of about 2500 duplex optical links and about 90 SOL40 10 boards. The firmware core is technology independent, developed in HDL language, it does not make use of any technology specific element and it is completely agnostic of the content of the data field.

The core provides a way to control with high parallelism and flexibility many FE chips via the GBT-SCA interfaces through GBT links. Its main functionalities can be listed as follows:

- Provide a generic hardware interface (FPGA) between the ECS system and the FE electronics.
- Build and encode/decode GBT-SCA compliant packets.
- Serialize and de-serialize command packets in the command word sent to the FE electronics according to GBT-SCA specifications.
- Support for all GBT-SCA protocols (SPI, IC, JTAG, GPIO and ADC+DAC).
- Support for all GBT-SCA commands and channels.
- Support for many GBT-SCAs per GBT link and many GBT links per FPGA.
- Possibility of re-transmission of packets and transmission monitoring.
- Modularity, i.e. components can be removed if not needed.
- Robustness, reliability, programmability, flexibility.

The core is essentially composed of a series of layers as it is illustrated in Figure 6. Their main roles are to:

- store the ECS configuration packets and decode them as commands and viceversa in the ECS Interface and ECS Packets Buffers Layers.
- build the corresponding GBT-SCA packets with the selected protocol in the Protocol Layer.

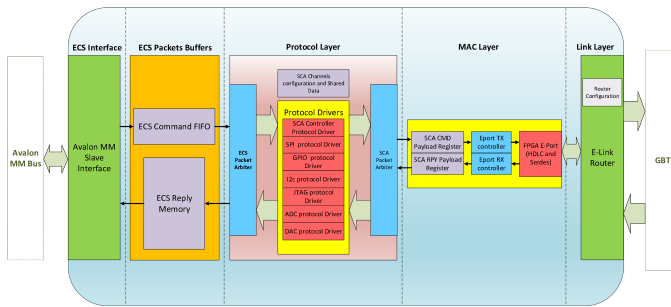


Fig. 6. Architecture of the SOL40-SCA firmware core.

- encode it in the specified communication protocol (HDLC [15]) in the MAC Layer.
- serialize and route the packets to the selected GBT-SCA connected to a GBT at the FE in the Link Layer.

In practice, the ECS generates a command which is transmitted to the FPGA via the PCIe bus of the generic LHCb hardware readout board. This command contains an extended addressing scheme to tell the core where and how to route the configuration packet and a command code scheme which tells the core what actions to perform (i.e. read/write or wait for response/do not wait). In addition, it may contain the configuration data to be sent to the FE in case of a write operation. In the FPGA, the command is stored in a buffer in order to be picked up by the Protocol Layer when not busy. The ECS command is then decoded and the SCA specific protocol packets are built accordingly. The information about which protocol to be built is in the ECS command and it is completely generic, that is the core is able to build run-time any SCA packet simply based on the content of the command. Finally, the packet is encapsulated in the HDLC protocol to then be routed to the corresponding bit field in the GBT word to be sent through the optical link to the corresponding Master GBT at the FE. The corresponding bit field is selected based on the connections at the FE. In order to be as generic as possible, this is also a configurable parameter so that the core can be used with any FE configuration. The core also features the possibility of packet retransmission in case a particular transaction failed. Yet another feature the core implements is the agglomeration of several GBT-SCA commands into one ECS command, denominated a *composite command*. The ECS only has to send one command to the core for it to send several commands to the GBT-SCA.

The core has been extensively tested with real FE chips and with ECS software that is being developed in parallel. The I2C protocol is the most tested so far and it has been used to program several kinds of FE chips in both 7b and 10b addressing modes. A 16-byte transaction over I2C is performed in about 3 ms, when issued from software. Fully configuring a chip with 366 registers (i.e. the GBT) of 1 byte takes about 100 ms currently - work is ongoing to reduce this time by applying some optimization both in the core and in ECS software. The SPI protocol was tested both in prototype chips and FPGA emulated chips showing good robustness. The GPIO protocol was fully tested to scan all of the lines in both direction (writing and reading). ADC and DAC channels were

tested by connecting them to each other, showing consistent results in the measurements. JTAG tests are done through the emulation of a slave controller inside a small FPGA and will later be used to remotely configure FPGA devices. Each core demonstrated being able to operate several GBT-SCA chips in parallel robustly and in a generic way.

V. CONCLUSIONS

Within its upgrade, the LHCb experiment has finalized the specifications of its sub-system. The timing and readout control system TFC is a crucial system in the upgrade of the LHCb detector as it is responsible to centrally manage the readout of event, the distribution of synchronous and asynchronous commands and the distribution of the global clock, received from the LHC. In this paper, the current implementations for the architecture of the TFC system have been presented. PON technology seems to be the ideal solution for such a system, together with optical links, FPGAs and high-level of interconnectivity. As commercial PON components will become more available in the near future, an extensive testing campaign within CERN is about to be performed in order to validate the ideas and deploy the system in its entire scale. Moreover, the LHCb experiment has developed a generic firmware core to drive any GBT-SCA at the FE electronics. This is achieved by implementing an HDLC based code, capable of driving any protocol of any GBT-SCA over any GBT link, programmable at run-time. The core is so generic that can be used in any FE environment featuring the presence of the GBT chipset. The firmware core is currently in use by sub-detectors in test-benches, test-beams and it will be extensively used in order to commission the FE electronics for the upgrade of the LHCb. A heavy testing campaign together with the very first GBT-SCA chips has been performed in order to test robustness, reliability and compatibility issues and showed excellent results.

REFERENCES

- [1] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005
- [2] R. Jacobsson (for the LHCb Collaboration), *Performance of the LHCb Detector during the LHCb Proton Runs 2010-2012*, Proceedings of 2012 IEEE/NSS, pp. **1479-1486**
- [3] LHCb Collaboration, *Letter of Intent for the LHCb Upgrade*, CERN-LHCC-2011-001
- [4] LHCb Collaboration, *Framework TDR for the LHCb Upgrade*, CERN-LHCC-2012-007
- [5] LHCb Collaboration, *LHCb Trigger and Online Upgrade Technical Design Report*, CERN-LHCC-2014-016
- [6] F. Alessio (for the LHCb Collaboration), *Trigger-less readout architecture for the upgrade of the LHCb experiment at CERN*, JINST **8** (2013) C12019
- [7] F. Alessio and R. Jacobsson, *Timing and Fast Control for the upgraded readout architecture of the LHCb experiment at CERN*, IEEE TNS vol. **60**, issue 5
- [8] P. Moreira et al., *The GBT Ser-Des ASIC prototype*, JINST **5** (2010) C11022
- [9] A. Caratelli et al., *The GBT-SCA, a radiation tolerant ASIC for detector control and monitoring applications in HEP experiments*, JINST **10** (2015) C03034
- [10] J.P. Cachemiche et al., *The readout system upgrade for the LHCb experiment*, proceedings of 2016 IEEE/Real-Time, this conference.
- [11] F. Alessio and R. Jacobsson, *A new readout control system for the LHCb upgrade at CERN*, JINST **7** (2012) C11010

- [12] D. Kolotouros et al., *A TTC upgrade proposal usign bidirectional 10G-PON FTTH technology*, JINST **10** (2015) C04001
- [13] M. Barros Marin et al., *The GBT-FPGA core: features and challenges*, JINST **10** (2015) C03021
- [14] C. Gaspar et al., *The LHCb Experiment Control System: on the path to full automation*, Proceedings of 2011 ICALEPCS, **pp. 20-23**
- [15] International Standards Organization, *Telecommunications and information exchange between systems - HDLC procedures*, ISO/IEC 13239:2002