The GBT, a Proposed Architecture for Multi-Gb/s Data Transmission in High Energy Physics

P. Moreira, A. Marchioro, K. Kloukinas
CERN, 1211 Geneva 23, Switzerland
Paulo.Moreira@cern.ch

Abstract

The future upgrade of the LHC accelerator, the SLHC, will increase the beam luminosity by a factor of ten leading to a corresponding growth of the amounts of data to be treated by the data transmission and acquisition systems. The development of the GBT chipset addresses this issue providing a means to increase the bandwidth available to transmit data to and from the counting room. The GigaBit Transceiver (GBT) architecture will provide the support to transmit simultaneously the three types of information required to run an experiment in a hostile radiation environment over a multipurpose link. This paper describes the GBT link architecture and some aspects of its implementation. As this project is still in the specification phase, detailed features might change prior to the final silicon fabrication.

I. Introduction

With the installation of LHC and its associated experiments approaching completion, CERN and its collaborating institutes are now considering an upgrade of the accelerator to achieve higher beam luminosity, the SLHC. Higher luminosity will bring the benefit of improving the statistical accuracy of the measurements but, it will also impose more stringent requirements on the performance of the data acquisition systems as well as on their radiation tolerant characteristics. Some of the systems and their components will have thus to be redesigned to cope with higher data rates and higher radiation levels. In particular, the data transmission links will have to be upgraded to wider bandwidths in order to cope with the larger amounts of physics data being produced by the detectors. To increase the bandwidth without paying a penalty on the detector's material budget, it is necessary to use fewer optical links at higher data rates rather than simply increasing the number of links. The GBT architecture was developed under this perspective. It is targeted to support high speed (~5 Gb/s) data transmission between the detectors and the counting room and it aims at providing simultaneous transmission of physics, trigger and experiment control data over the same link. This last feature represents a significant innovation for systems used in HEP, where these three functions have traditionally been implemented over different and independent physical systems. The large bandwidth available allows the GBT to act simultaneously as a data-link and as a TTC transceiver and therefore to reduce the overall system cost significantly. The GBT architecture will implement point-to-point duplex links allowing bidirectional data transmission between the counting room and the detectors.

In other to simplify the development, embedding in existing systems and maintenance of the links, the GBT interface is proposing to adopt as the high level transport protocol the Ethernet standard. Moreover, to enhance system integration of the off detector electronics, the GBT transceivers in the counting room will be implemented by FPGAs. This requires using a communication protocol between GBT transceivers that can be implemented in standard FPGAs existing today in the market.

Due to the beam luminosity increase, the total dose radiation levels that the on detector electronics will be exposed to are expected to increase in the same proportion reaching the 100 Mrad level for some of the inner detectors. These high levels of radiation will pose long term reliability problems due to total dose effects. These can be minimized by using deep submicron CMOS commercial technologies and by following special layout techniques previously developed for the LHC ASICs. The GBT chipset will be thus fabricated in a commercial 130 nm technology which will ensure the capability to implement large bandwidth transmission circuits (~5 Gb/s). The higher luminosity will also be linked to an increase of the Single Event Upset rates. SEUs are a major impairment to error free data transmission at high data rates. To deal with this situation the GBT line coding adopts a robust error correction scheme that will allow correcting burst of errors caused by SEUs. These features are discussed in more detail in the following sections.

II. Link Architecture Overview

Running high energy physics experiments typically requires the operation of three concurrent systems: Data Acquisition (DAQ), Timing Trigger and Control (TTC) and Slow Control (SC). These systems have very different requirements in what concerns the data transmission bandwidth of their links. In the past specific data transmission links have been custom developed for each of these applications. This was somehow justified by the type of application being targeted and as well by technological limitations. However, with today's maturity of the optoelectronics and CMOS technologies, it is possible to envisage the development of a general purpose optical link which can simultaneously cover those three data transmission applications. Such an approach has a clear advantage of concentrating the development effort on a flexible single system instead of dispersing the development effort on several nearly identical systems. Moreover, economies of scale can be made since a single type of components needs to be produced, qualified, installed and maintained. Additionally, the number of optical fibres required can be drastically reduced due to the use of a higher bandwidth (when compared with the systems currently deployed in the LHC experiments) and because a single optical fibre link will take up the functionally of what was previously requiring three links.

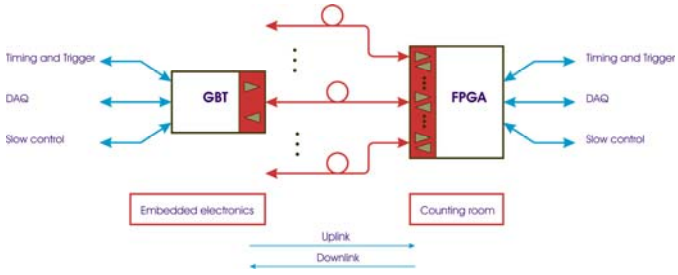

Figure 1: GBT based link architecture

The development being proposed will act thus as a DAQ, TTC and SC link, living still some flexibility to implement custom configurations. At the heart of such a link is the GigaBit Transceiver ASIC (GBT13). The global architecture of an optical link based on the GBT is shown in Figure 1. It consists, essentially, on bidirectional point-to-point optical fibre links between the counting room and the detector embedded electronics. As mentioned before, each of these fibres carries simultaneously DAQ, TTC and SC data. In the counting room, the receiver and transmitters are implemented using Commercial-Of-The-Self (COTS) components and FPGAs while embedded on the detectors the receivers and transmitters are implemented by the GBT chipset which consists on the GBT itself plus the GigaBit Laser Driver (GBLD) and the GigaBit Transimpedance Amplifier (GBTIA). This architecture clearly distinguishes between the counting room and embedded electronics specificities: that is, on-detector electronics works in a hostile radiation environment requiring custom made components while the counting room components operate on a radiation free environment and can be implemented by COTS components. Moreover, the availability of FPGAs with up to 32 serializer macros [1, 2, 3 and 4] will allow concentrating data from several front-end sources into a single module in the counting room facilitating data merging and leading to compact systems.

A. *Link bandwidth*

As seen from its electrical interface, by the embedded electronics (and the counting room FPGA), the GBT link can be though as providing three independent interfaces with one dedicated to the DAQ, a second to the TTC, and a third to SC. All these interfaces provide a fixed bandwidth but in practice fractions of them can be assigned to different channels as will be discussed latter. To understand how the bandwidth is allocated, it is necessary to consider in detail how the data is encapsulated when transmitted between the counting room and the detectors in what is called the GBT frame.

As represented schematically in Figure 2, the GBT frame is composed of 120 bits which are transmitted during a single bunch crossing interval (≅ 25 ns) resulting in a line data rate of 4.8 Gb/s. Of these, 4 bits are used for the frame Header (H) and 32 used for Forward Error Correction (FEC). This leaves a total of 84 bits free for data transmission corresponding to a user bandwidth 3.36 Gb/s. Of the 84-bits four are reserved for the SC field, the 'TTC' field is 16-bits wide and the 'D' field is 64-bits wide resulting in the following bandwidths: SC – 160 Mb/s, TTC – 640 Mb/s and DAQ – 2.56 Gb/s.

The high levels of radiation (up to 10^{16} n/cm^2 and 100 Mrd) foreseen for SLHC environments require strong error correction algorithms to be used in order to achieve error free data transmission. A significant fraction of the channel bandwidth is thus assigned to the transmission of Forward Error Correction (FEC) code. The 32-bit Error Correction (FEC) field carries an interleaved Reed-Solomon double error correction code which is discussed further in section III A.

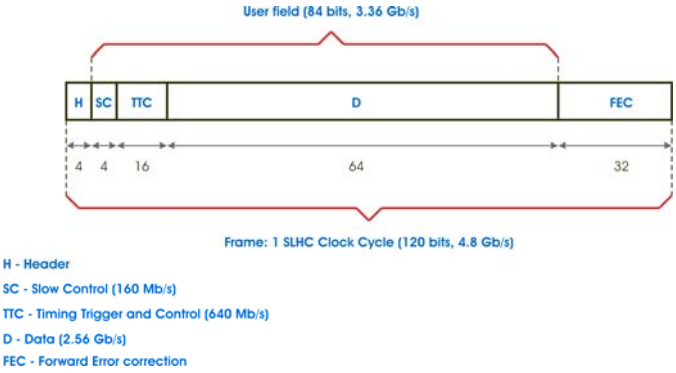

Figure 2: GBT frame

B. *E-Links*

The full bandwidth available for data transmission might not be well matched to the requirements of a single front-end device. It is thus conceived that the GBT13 chip will connect with the Front-End ASICs through a set of 32 bi-directional serial links which are called E-Links. Typically the available bandwidth of a GBT will exceed the throughput of single front-end device being possible for a single GBT13 to serve several front-end ASICs simultaneously. The general front-end link topology using the e-links is schematized in Figure 3. Each e-link is physically implemented by three pairs of differential lines (6 wires) being them the clock (Clk+ and Clk-), Data In (Din+ and Din-) and Data Out (Dout+ and Dout-). The electrical levels follow the LVDS standard [5] and the data in both directions are simply clock synchronous to the 80 MHz clock, allowing to transmit data up to 1-2 m distance. An important feature of the GBT is that several e-links can be easily grouped together to serve a single front-end device achieving bandwidths that are multiples of 80 Mb/s. This flexibility allows the GBT architecture to be scalable and adaptable to different front-end systems, and thus to be used in a number of different applications. In addition, even for given detectors, there might well be regions where detectors provide more data, and the GBT can be adapted to such requirements by grouping a different number of e-links to the demands of the specific front-end modules.

To save in wiring, the data link to the front-end modules is instead always limited to a single pair of signals (Din+ and Din-, for a total of 80 Mb/s) broadcasting the same information to the modules connected to the same e-link.

1) *E-link protocol*

Users will connect to an e-link through a protocol that defines from the very source the packet format. It is suggested that packets of data, representing parts of physics events or multiple events, be created using the well defined Ethernet protocol. The addition of a source address and a destination

address to the user data allows the construction of simple routing devices on the receiving boards in the experiment counting room and supports easily the utilization of high level software now widely in use in computers and industrial control systems. For instance, event building at the counting room level can be achieved by addressing all pieces of data of one event to a same destination, thus greatly simplifying the job of constructing the back-end of the data acquisition system.

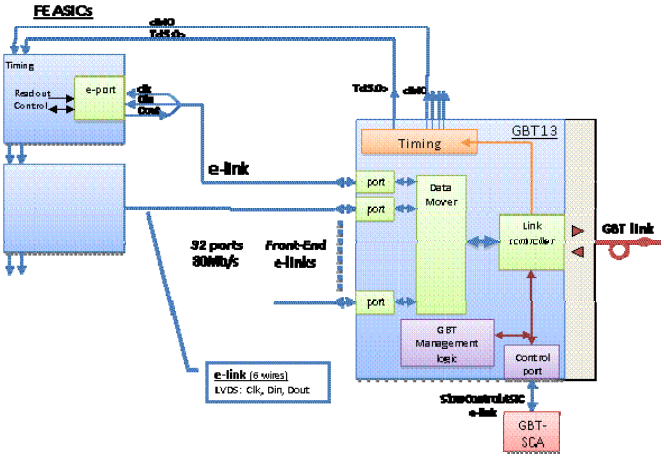

Figure 3: General Front End link topology

In turn, control of front end modules will be possible by generating Ethernet formatted packets for specific Ethernet addresses, each of the front-end modules being assigned such a unique address.

An e-link compatible macro block (called E-link Port Adapter (EPA)) will be made available to ASIC designers in the form of synthesizable HDL code by the GBT project to be integrated in the back-end of their front-end chips. Protocol construction and compliance, line coding and synchronization will be taken care by this macro which is intended to interface to users through a very simple FIFO-like interface. This macro will also provide a method for dispatching the information coming from the counting room to a standard set of user registers defined in the e-link port adapter.

2) Slow control channel

A pair of E-links will be reserved to implement a slow control channel with a total available bandwidth of 160 Mb/s. The GBT13 itself will not act as a slow control management unit but reserves a fraction of the slow control bandwidth to implement functions that have to do with the management of the GBT link it self. All the other slow control functions will be done by a dedicated ASIC called the GBT – Slow Control ASIC (GBT – SCA) (see Figure 3). This chip will interpret the Ethernet frames carried by the SC field of the GBT link and will serve as node controller routing them to the appropriate channels waiting for actions and feeding back to the counting room the requested responses.

III. The GBT chipset

The block diagram of the GBT chipset (which is shown in Figure 4) is composed of the following parts: The GBTIA whose function is to convert the weak photocurrent generated by the photodiode into a digital differential voltage capable of driving the Clock and Data Recovery (CDR) circuit of the GBT13. The GBLD which uses the serial data produced by the GBT Serializer (SER) to modulate the laser diode optical power. The GBT13, which acts as a receiver and transmitter for the optical link, as a communications controller for the e-links and as a TTC receiver.

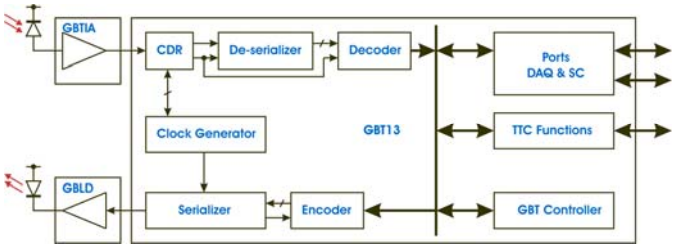

Figure 4: CBT chip set block diagram

The optical link receiver section of the GBT contains the CDR circuit, the De-Serializer (DES) and a decoder (DEC). The CDR circuit recovers the serial clock (4.8 GHz) from the incoming data while the DES circuit converts the 120-bit serial frame into a parallel word. The decoder detects and corrects any errors (which are within the error correction capabilities of the code) that have been introduced during the transmission. After error detection and correction the decoder de-scrambles the data which is then feed to the “Ports” the “TTC functions” or the “GBT controller”. On the optical link transmitter side data is scrambled encoded and serialized. The GBT13 contains as well a clock generator that can be used as a clock source for “standalone” operation or as a phase noise reduction circuit for an external reference clock.

A. *Forward error correction*

The high radiations levels expected for SLHC will result in relatively high rates of SEUs for the embedded electronics. A typical approach to overcome this problem is to use triple-modular redundancy [6] or Hamming encoding [7] to obtain some error correction capability. In addition to errors in the internal ASIC logic, particles can further generate spurious events in the photodiodes, thus simulating the arrival of a data bit, and therefore generating data and phase errors in the receiver. The use of triple-redundancy techniques would impose a speed penalty on the serializer-deserializer (SERDES) circuit preventing the implementation of logic operating at 4.8 Gb/s. The approach adopted was thus to not apply any redundancy techniques on the data path between the input of the serializer and the output of the de-serializer but instead to transmit the data with a Forward Error Correction code (FEC) generated at the input of the former block and verified at the output of the latter. Any transmission errors occurring on the SER, GBLD, PIN-diode, GBTIA, CDR and DES will be corrected at the receiver during the decoding operation. The code to be used has however, to provide a high level of protection, since errors occurring during transmission tend to occur as burst errors [8, 9] and not as isolated events. Because of this, a double interleaved Reed-Solomon two-errors correcting code was chosen. The code is built by interleaving two Reed-Solomon [10] encoded words with 4-bit symbols width, each capable of correcting a double symbol error. This in practice means that a sequence of up to 16 consecutive wrong bits can be corrected. This correction

technique requires an extra field of 32 bits in the frame of 120-bits, resulting in a code efficiency of 73%. In the 130 nm CMOS technology used both the encoding and decoding operations can be done well within a single machine clock cycle (<25 ns) thus having a minimal impact on the transmission latency. In a particular logic implementation, that is currently under evaluation and that has been implemented in a test ASIC currently under test, the entire encoding (and decoding) operation is performed in one single and fairly deep level of logic, using no internal register and/or state machines. This implementation minimizes the impact of SEU on registers, as no register or latch is used throughout the entire block. Taking into account that a complete Reed-Solomon encoder (decoder) needs between one and two hundred lines of code to be written in a high level language (such as C or C++), we have unrolled the entire algorithm in a single sequential logic block and synthesized the whole block without intermediate registers.

The encoding scheme used combines also the capability of DC-balancing the transmitted data [11, 12]. This is achieved by using a self-synchronizing scrambler that "randomizes" the data pattern and guarantees a proper distribution of 0's and 1's in the data stream. In our implementation the scrambling function is implemented before the Reed-Solomon encoding (and the de-scrambling after the decoding).

One should stress that the proposed code guaranties DC balance of the transmitted data while simultaneously providing a strong error detection and correction capability. When compared with a standard 8B/10B line code, which provides no error correction functionality and has only very limited error detection ability, the proposed code introduces only an additional 7% loss of coding efficiency.

B. TTC functions

The first LHC TTC system was designed to distribute uni-directionally a single bit trigger along with some control information in addition to a precisely timed clock to the front-ends of experiments. With the advent of the high bandwidth available today, the GBT can now distribute 16 bits of trigger information, and a very high precision accelerator synchronous clock. It also has the new capacity to transmit the trigger information from the experiment to the counting room, allowing the design of systems where local pre-trigger information is computed promptly on the detector and sent from there to the counting room with a minimum of latency.

The trigger information is transmitted and received with fixed and small latency. Of the 120 GBT frame bits, 16 (the TTC field) are allocated for purposes of trigger broadcast. These bits are distributed to the front-end electronics in parallel (see Figure 3).

It is also a role of the TTC functions to allow for compensation of the cable/fiber lengths, the time-of-flight of particles and delays in the electronic circuits. To do so, the GBT13 provides four clocks at 40 or 80 MHz whose phases can be programmed independently in steps of 50 ps.

As part of the TTC functions the GBT13 ASIC contains also a trigger emulator which serves to mimic trigger sequences and the accelerator beam structure cycle. Its main purpose is to aid in the development phase of data acquisition systems, for instance, repetitive events can be generated at fixed positions in the beam structure by programming this emulator. It basically consists of a 4-bit memory with a depth as long as the (S)LHC machine beam, that is freely user programmable and is read in a cyclic manner. The memory is addressed sequentially and its contents are made available on 4 trigger emulation bits. The contents of the trigger emulator memory are programmable through the SC channel.

IV. Summary

In this paper a link architecture is proposed for high speed data transmission in the future upgrade of the LHC experiments. This work suggests the use of point-to-point bidirectional links to simultaneously transmit data, timing, trigger and slow control information between the counting room and the embedded detector electronics. The proposed scheme addresses the problems of data transmission in a radiation hard environment where single event upsets will present a serious impairment for error free data transmission and operation. The hart of such a system will be a chipset, called the GBT chipset, that will transmit data at 4.8Gb/s via a bidirectional optical links between the counting room and the detectors. The GBT chip set will further dispatch and collect data from the front-end electronics using local electrical links whose bandwidth can be adapted to the front-end electronics requirements. It is estimated that such an approach can address the needs of the future LHC upgrade in terms of data bandwidth and that it can substantial lower the development, installation and maintenance cost when compared with solutions that consider and treat data, timing and slow control links as separated entities.

V. Acknowledgements

This work would not have been possible without the collaboration of many people currently involved in several activities at CERN and in particular the GBT project. In particular we would like to thank the many discussions we had with the following people:

CERN – Geneva Switzerland: P. Aspell, S. Baron, J. Christiansen, F. Faccio, J. Troska, F. Vasey and K. Wyllie. CPPM – Marseille, France: M. Menouni and J. Cachemiche. INFN – Bari, Italy: F. Loddo, A. Ranieri and G. De Robertis. INFN – Bologna, Italy: A. Gabrielli. INFN – Torino, Italy: O. Cobanoglu, G. Mazza and A. Rivetti. SMU – Dallas, USA: P. Gui.

VI. References

[1] http://www.latticesemi.com/

[2] http://www.xilinx.com/

[3] http://www.altera.com/index.jsp

[4] http://www.actel.com/

[5] 'LVDS owner's manual low-voltage differential signaling', Spring 2004, 3rd Edition, http://www.national.com/appinfo/lvds/files/ownersmanual.pdf

[6] R. D. Schrimpf and D. M. Fleetwood, Editors, “Radiation effects and soft errors in integrated circuits and electronic devices,” World scientific publishing Co. 2004

[7] H. Nussbaumer, “Computer Communication Systems: Data Circuits Error Detection Data Links,” Vol. 1, John Wiley & Sons, 1990

[8] C.J.Marshall, P.W.Marshall, M.A.Carts, R.Reed, S.Baier, K.LaBel “Characterization of transient error cross sections in high speed commercial fiber optic data links”, IEEE Radiation Effects Data Workshop, 2001, pp142-145.

[9] P.W.Marshall, P.T.Wiley, R.N.Prusia, G.D. Rash, H.Kim, K.A.LaBel “Proton induced bit error studies in a 10Gb/s Fiber Optic Link” IEEE Transactions on Nuclear Science, vol.51, no.5, October 2004, pp.2736-2739.

[10] S. Lin, D. J. Costello, “Error Control Coding”, 2nd ed., Prentice Hall, 2004, ch.2-7.

[11] G. Papotti, “An Error-Correcting Line Coding ASIC for a HEP Rad-Hard Multi-GigaBit Optical Link”, Proc. 2nd Conference on Ph.D. Research in Microelectronics and Electronics (PRIME 2006), Otranto (Lecce), Italy, 12-15 June 2006, pp.225-8.

[12] Giulia Papotti, ‘Architectural studies of a radiation-hard transceiver ASIC im 0.13 mm CMOS for digital optical links in high energy physics applications,” PhD thesis, University of Parma, Italy, January 2007, http://papotti.web.cern.ch/papotti/tesi.pdf