

TOPICAL WORKSHOP ON ELECTRONICS FOR PARTICLE PHYSICS
BERGEN, NORWAY
19–23 SEPTEMBER 2022

Evaluation of GBT-FPGA for timing and fast control in CBM experiment

V. Sidorenko,^{a,*} W.F.J. Müller,^b W. Zabolotny,^c I. Fröhlich,^b D. Emschermann^b
and J. Becker^a on behalf of CBM collaboration

^aKarlsruhe Institute of Technology, Engesserstraße 5, 76131 Karlsruhe, Germany

^bGSI Helmholtz Centre for Heavy Ion Research, Planckstraße 1, 64291 Darmstadt, Germany

^cWarsaw University of Technology, Pl. Politechniki 1, 00-661 Warsaw, Poland

E-mail: vladimir.sidorenko@kit.edu

ABSTRACT: Timing and Fast Control (TFC) system for the Compressed Baryonic Matter (CBM) experiment is being developed with focus on low and deterministic data transmission latency. This helps to minimize data corruption in the free-streaming Data Acquisition (DAQ) system during occasional data bursts caused by the expected beam intensity fluctuations. Proven in latency-optimized experimental data transport applications, the GBT-FPGA core is expected to positively contribute to the TFC system performance. In this work, the core has been integrated as the primary communication interface and its effect on transmission latency and quality of time distribution has been evaluated.

KEYWORDS: Control and monitor systems online; Detector control systems (detector and experiment monitoring and slow-control systems, architecture, hardware, algorithms, databases)

*Corresponding author.

Contents

1	Introduction	1
2	Background	2
3	Updated architecture	2
4	Evaluation setup	3
5	Measurement and results	4
6	Conclusion	5

1 Introduction

Being one of the experiments operating at the future Facility for Antiproton and Ion Research (FAIR), the Compressed Baryonic Matter (CBM) experiment aims to study strongly interacting matter at high baryon densities [1]. The detector will measure rare diagnostic probes with high precision at unprecedented interaction rates of up to 10 MHz. In order to accomplish that, the experiment will feature a free-streaming data acquisition system. Self-triggered front-end electronics (FEE) are expected to generate a total of up to 1 TB/s of timestamped experimental data. On the detector, GBTx-based readout cards concentrate the data onto optical links that transport the data to the data acquisition (DAQ) server room [2]. Here, the links terminate in Common Readout Interface (CRI) FPGA cards that process and aggregate the FEE data streams and pack the data into appropriate container structures that are transferred via PCIe into the DAQ entry node servers. They perform time-slicing and forward the data to a computer farm, where the First-Level Event Selector (FLES) system reconstructs the events and selects them for eventual storage [3].

The timing-related part of the role of the Timing and Fast Control (TFC) system is to distribute a common 40 MHz clock and a common time to the data readout tree. The common clock must be propagated to CRI boards and, subsequently, to the FEE layer. This ensures that front-end modules assign timestamps to the collected data at a common rate. Whereas FEE clocks do not need to be precisely aligned, since constant time offsets are calibrated out by FLES, their relative phases must be constant at all times, even after a full system restart. The common time in the form of a 64-bit counter value originates in the central TFC Master and is broadcast to all CRI endpoints. The CRI boards, if needed, adjust their local time counters. Both clock and time distribution mechanisms rely on latency-deterministic downstream (Master to Endpoint) communication in the system. Based on the configuration of the experiment, at least 200 CRI boards need to be synchronized and their relative time, in terms of clock phase and local time counter value, must remain stable within a range of 200 ps.

In addition to defining a time reference to the experiment, the TFC system provides a low-latency path for collecting status information from CRI boards and issuing control commands, e.g., throttling decisions to avoid congestion of the DAQ system [4].

2 Background

Currently, a functional prototype of the TFC system is used for beam tests with the mCBM experiment, a CBM full-system test setup [5]. The prototype has been implemented as a scalable hierarchical network, with a number of TFC Endpoints, connected to the Master node with bi-directional optical links. The purpose of the functional prototype was to provide the fundamental timing and command distribution functionality with limited performance, as this enables intermediate integration tests of the full experimental setup.

Even though the TFC prototype has been successfully integrated and tested, the quality of its time distribution and fast control capabilities is sub-optimal and depends on the latency performance of the transport link. The absolute amount of latency defines how fast the endpoints receive the throttling decisions, and deterministic delivery and execution of these commands reduces the amount of partially collected events from concurrent front-end links. On the other hand, the accuracy of time distribution directly depends on latency determinism of the downstream links, over which the common time information propagates. The purpose of the currently presented work is to ensure sub-clock determinism of the downstream data link, since this is a major prerequisite for accurate distribution of CRI time.

Various parts of the system can introduce link latency variation. Unoptimized clock handling can cause fine sub-clock phase shifts, whereas non-deterministic components in the datapath, e.g., FIFOs, introduce larger latency uncertainties on the clock cycle level. The latency-optimized version of the GBT-FPGA core has been developed to guarantee latency determinism in both clock and data, which is essential in trigger and timing systems [6]. Latency variation between power-ups is not expected to be greater than 100 ps peak-to-peak with this core. Some sources of latency uncertainty are handled inside the GBT-FPGA core, while others have to be mitigated by the system developer. This makes the core a good candidate for enhancing link latency determinism.

3 Updated architecture

The updated FPGA gateway concept features a communication channel over a GBT link and is shown in figure 1. As the GBT-FPGA core implicitly manages bit- and word- alignment, the link-related complexity is hidden from the architecture layer, including additional clock domains. A 40 MHz system clock is used for both transmitter (Tx) and receiver (Rx) data frames. In the case of the Endpoint node, this clock is also extracted from the incoming GBT serial link.

The time information is represented in the system with a 64-bit timestamp and is periodically transmitted from the Master node to the Endpoint. This transmission is handled by the *Timing master* module that periodically captures the momentary timestamp value and sends it to Endpoint over the optical link. The received timestamp is detected by the *Timing endpoint* module and is used to adjust the local time counter upon link initialization, allowing all Endpoints to be synchronized.

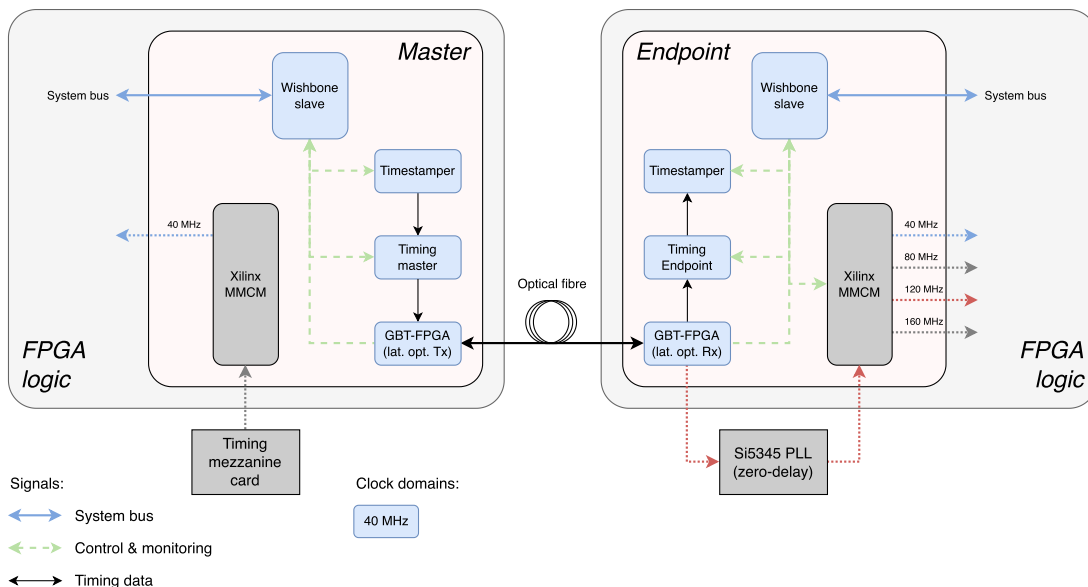


Figure 1. Proposed gateway architecture of the TFC upgrade with GBT-FPGA.

In order to fulfill this requirement, the downstream connection from the Master to Endpoints must be latency-deterministic.

In addition, the overall clock distribution scheme has been revised for all clocking components to have a deterministic input-to-output phase shift that is preserved between system restart cycles.

4 Evaluation setup

In order to evaluate the potential benefits of a data link based on GBT-FPGA to the TFC system, a minimal system has been built using the same hardware platform as is used in the existing prototype (see figure 2). It consists of a sender node, which represents the Master, connected with an optical fiber to a receiver, to emulate the Endpoint node. The nodes are implemented on BNL-712 boards, with the receiver board being equipped with a mezzanine card that provides an optical interface. Both boards are mounted into commercial server nodes that are accessible from a local Ethernet network.

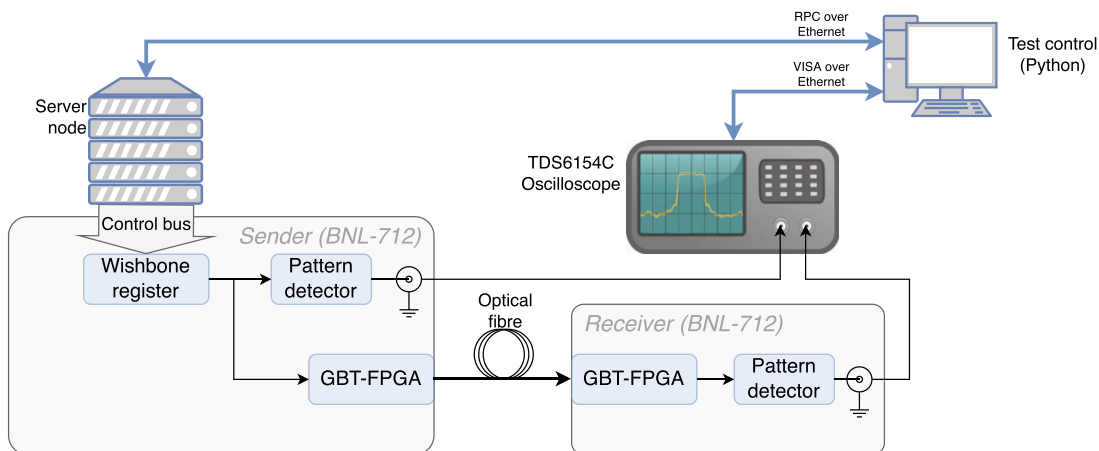


Figure 2. Link latency measurement setup.

Link latency is measured using an oscilloscope with probes connected to each of the boards. The measurement pulses are generated with pattern detectors, which are triggered by a pre-defined data pattern on the Tx and Rx ports of the sender and receiver boards, respectively.

While no significant control of the receiver board is required in this scheme for pulse generation, the sender relies on the external input for data pattern transmission. The data to be sent is programmed via a Wishbone register from a common Python script. This script handles a broad range of test control tasks and is interfaced with server- and board management infrastructure. With this centralized approach to test control, it is possible to thoroughly evaluate link latency variation on different time scales by running highly automated long-term tests.

Upon generation, latency measurement pulses from the sender and the receiver are captured by the oscilloscope and time skew between rising edges is measured. The test control loop is then closed by reading out the measurements into the control software.

5 Measurement and results

Two aspects of link latency variation have been investigated in the current work: latency distribution over extended periods of time without a reset and reset-to-reset stability. This has been accomplished by performing a number of long series (*runs*) of latency measurements (*samples*), with power cycling the hardware and reprogramming the FPGAs between the series (see figure 3). The worst-case reset-to-reset performance of the link can be evaluated this way.

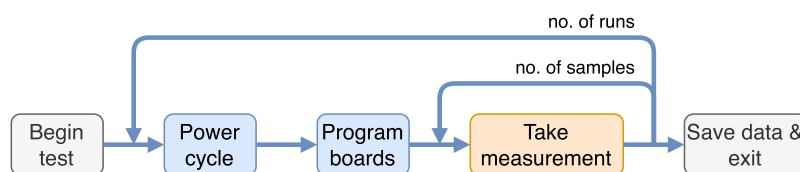


Figure 3. Procedure for automated measurement of link latency variation.

To set a reference for link latency performance, the data interface from the existing TFC prototype has been adapted for the evaluation setup. This way, a quantifiable improvement of link latency determinism can be derived from a direct comparison between the results obtained with the reference design and the setup based on a latency-optimized GBT link. The same test procedure has been run on both setups and it included 10 *runs*, 1000 *samples* each. With each latency measurement taking 4.35 s on average, executing the full test procedure has taken approx. 11 hours.

Results of the tests can be seen in figure 4. Mean and modal values are presented on the plots to represent a unified latency value within each *run*. The difference between these values indicates a non-Gaussian distribution of measured latency values. Nevertheless, they both are spread across approx. 12 ns in the case of the reference link. This spread represents reset-to-reset phase variation and greatly affects the overall latency variation of the link. In the case of the latency-optimized GBT-FPGA link, mean latency value stays within a 50 ps range. The results have also shown

that the in-run latency variation is more prominent than that between power-ups in the case of the latency-optimized GBT-FPGA link. The in-run standard deviation varies between 60 ps and 120 ps in different runs, and the overall latency variation of less than 500 ps peak-to-peak has been observed.

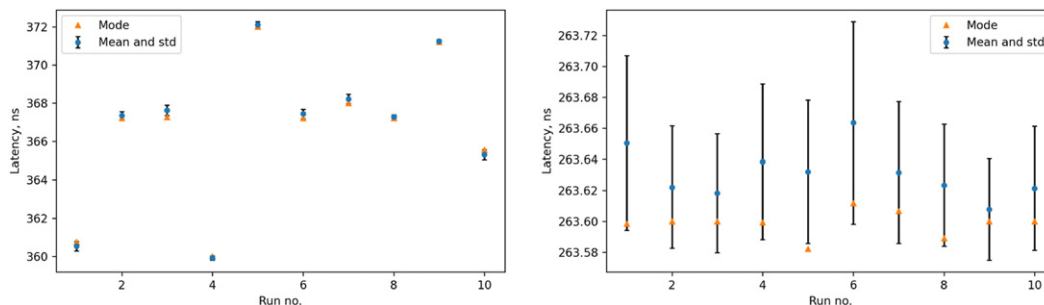


Figure 4. Latency variation with the reference link (left) and the latency-optimized GBT link (right).

6 Conclusion

A latency-optimized GBT-FPGA core has been evaluated for potential use in the TFC system of the CBM experiment, where link latency determinism is a major factor affecting system performance. By performing a series of latency measurements with power cycling the hardware in-between, the latency variation of the proposed link has been thoroughly studied in comparison with the links currently used in the experiment. Overall latency variation of the link has been observed to be less than 500 ps peak-to-peak. Although this performance leaves room for further work towards meeting the target synchronization accuracy of 200 ps between Endpoint nodes in the CBM experiment, the measurement results show that the new link guarantees sub-clock accuracy of time distribution to CRI boards in the current mCBM setup.

Prior study indicates greater potential with the GBT-FPGA interface for latency determinism [6]. Exploring this potential will be the focus of the further work towards fulfilling requirements of the CBM experiment on timing and fast control.

Acknowledgments

The project on which this report is based was funded by the German Federal Ministry of Education and Research under the funding code 05P21VKFC1. The responsibility for the content of this publication lies with the author. This work was also supported by the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement No. 824093 (STRONG-2020).

References

- [1] T. Ablyazimov, A. Abuhoza, R.P. Adak et al., *Challenges in QCD matter physics — the scientific programme of the compressed baryonic matter experiment at FAIR*, *Eur. Phys. J. A* **53** (2017) 60.
- [2] P. Moreira, R. Ballabriga, S. Baron et al., *The GBT project, proceedings of the Topical Workshop on Electronics for Particle Physics (TWEPP 2009)* (2009).
- [3] CBM collaboration, J. de Cuveland and V. Lindenstruth, *A first-level event selector for the CBM experiment at FAIR*, *J. Phys.: Conf. Ser.* **331** (2011) 022006.
- [4] X. Gao, D. Emschermann, J. Lehnert et al., *Throttling strategies and optimization of the trigger-less streaming DAQ system in the CBM experiment*, *Nucl. Instrum. Meth. A* **978** (2020) 164442.
- [5] V. Sidorenko, I. Fröhlich, W.F.J. Müller et al., *Prototype design of a timing and fast control system in the CBM experiment*, *2022 JINST* **17** C05008.
- [6] M.B. Marin, S. Baron, S.S. Feger et al., *The GBT-FPGA core: features and challenges*, *2015 JINST* **10** C03021.