Throttling Studies for the CBM Self-triggered Readout

Xin Gao
GSI Helmholtzzentrum für Schwerionenforschung GmbH, Germany
Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai, China
E-mail: x.gao@gsi.de

David Emschermann, Jörg Lehnert, Walter F.J. Müller
GSI Helmholtzzentrum für Schwerionenforschung GmbH, Germany
E-mail: d.emschermann@gsi.de, j.lehnert@gsi.de, w.f.j.mueller@gsi.de

The Compressed Baryonic Matter experiment (CBM) will study rare probes in a heavy-ion environment at high interaction rates of up to 10 MHz. The observation of detached vertices requires a topological trigger, which is too complex to be defined in the hardware, and fully realized in software. Excluding conventional, latency-limited trigger architectures, CBM opted for a free-running readout. The primary beam is delivered by a slow extraction synchrotron. To be able to operate the experiment at highest interaction rates, despite beam intensity fluctuations, a time-based throttling mechanism is under study. SystemVerilog based simulations of the readout system show the feasibility of the examined throttling algorithm to reduce the number of incomplete events, particularly in presence of intensity fluctuations.
1. Introduction

As a fixed-target experiment, CBM is designed to cope with very high interaction rates up to 10 MHz. To achieve the high rate capability, CBM will be equipped with fast and radiation hard detectors employing self-triggered front-end electronics and a free-streaming readout system. The Data Acquisition (DAQ) system will transport raw data with up to 1 TB/s to a large scale computer farm (First-Level Event Selector, FLES), where online event building and event selection are performed [1][2].

The primary beam is delivered by a slow extraction synchrotron. However, significant beam intensity variations on short time scales are expected. Such fluctuations can lead to a sudden increase of occupancy in the detector. Such overload situation may temporarily exceed the readout bandwidth and the capacity of front-end buffers, thus leading to data losses and consequently to incomplete events. Therefore, the throttling mechanism is one promising solution to achieve a high quality data stream in CBM. Currently, the CBM DAQ system is in the prototyping phase. The research of the throttling algorithm in this article is the base of future firmware realization.

2. Throttling system

This study is based on the CBM Silicon Tracking System (STS), which is the main tracking system closest to the target with a readout bandwidth of up to 50 MHits/s per front-end ASIC. The readout electronics (see figure 1) comprises 14,400 STSXYTER ASICs, populating 1,800 Front-End Boards (FEB-8), interfacing to about 600 GBTx Readout Boards (ROB-3), connecting to about 80 Common Readout Interface cards (CRI) connected to a central Timing and Fast Control system (TFC). All of the components are under the supervision of the Experiment Control System (ECS) [1][3].

![Figure 1: Hierarchy of the readout tree of the STS subsystem](image)

Figure 2 is the throttling hardware functional diagram. Various parts of the throttling logic are implemented in the CRI and TFC firmware and in the front-end ASIC. Each STS-XYTER comprises 128 readout channels as input. Each channel has a FIFO of 8 words. The throttling functionality is implemented as follows. A FIFO-almost-full flag is asserted once 7 elements are
Throttling for self-triggered DAQ

Xin Gao

The ASIC counts the number of channels with almost-full FIFOs. The channel threshold triggering an alert is programmable [4]. If exceeded, the ASIC reports an almost full condition to the CRI by sending an alert frame. In CRIs, the Alert Unit (AU) receives the alert frame, and transfer this alert signal to FIFO Full Indicator (FFI). Meanwhile, the AU resets ASIC status bits immediately and thus guarantees proper arbitration of incoming alerts and the resulting action. The alert signal is locked in FFI and released by timeout or throttling reset. Through FFI and AU, busy ASIC information is saved in CRIs temporarily. Collecting the busy information from all CRIs, the TFC decides if CBM as a whole should be throttled. The CRIs receive the throttling decision from the TFC and propagate system dependent throttling instructions to the ASICs.

In this paper, the so-called “Stop” strategy is investigated. Upon a throttling decision, the strategy is to stop accepting new hits in the channel FIFOs for a time long enough to fully read out these FIFOs and restart accepting hits afterwards.

3. Simulation environment

The simulation structure is a closed-loop model, comprising a hit generator, a data flow model and a result evaluation. As the core, the data flow model implemented in the Questa framework using the System Verilog language. It calls Linux shells to invoke the front and back stages realized in C++/ROOT. In the hit generator, one event is one collision. In order to keep the initial simulation simple, the event size is random with a uniform distribution, which means the average number of hits per event is constant. Also for initial simplification, for each event, hits have a uniform distribution on all detector channels. This can later be extended to a more complex model based on real physics simulations. The hit rate equals the event rate multiplied by event size. For a more intuitive understanding of throttling results, the normalized hit rate is defined as the ratio of hit rate to readout bandwidth. The model implements 32 ASICs. The total simulated time is 10 ms, enough to cover order of hundreds of throttling cycles. The hit generator only includes valid hits after pileup correction. Pileup signals which cannot be distinguished by the front-end electronics are removed.

In the data flow model shown in figure 2, the maximum number of 5 readout elinks/ASIC are active [4], resulting in 50 MHits/ASIC/s as bandwidth limit. The corresponding drain time of the...
ASIC is 20.48 $\mu$s. Throttling parameters are the channel thresholds for alerts per ASIC and the fraction of ASICs reporting alerts. In this article, the alert threshold is $1/3$ out of 128 almost full channels. Throttling is triggered when half of the ASICs report busy alerts.

If 95% of hits in one event are saved, the event is considered as restorable or good. We compared the absolute number of good events in the same simulation time with and without throttling when the normalized hit rate is increased from a small fraction of the readout bandwidth up to twice the bandwidth limit.

4. Simulation results

4.1 Simulation with stable beam intensity

Assuming a stable beam intensity, the event rate is only subject to Poissonian fluctuations. As shown in figure 3, green and blue lines represent "expected all events" and "all events", respectively. The "all events" are the event counts of the hit generators, while the "expected all events" are calculated by $R_{hit} \times T / S_{event}$. The hit rate $R_{hit}$ is acquired by multiplying the normalized hit rate and readout bandwidth. $T$ denotes total simulated time. The average event size $S_{event}$ equals 160 hits (5 hits/ASIC * 32 ASICs) in the simulations. The "all events" generated by the hit generator meets the "expected all events" in this model. Without throttling, the good events go quickly down to 0 after bandwidth limitation. However, we can see the improvement of data quality with "Stop" strategy. When the hit rate exceeds the bandwidth limit, a large fraction of the available bandwidth is still used for almost complete, good events.

![Figure 3: Simulation with stable beam intensity](image)

4.2 Simulation with realistic beam intensity fluctuation

Figure 4a is the realistic beam intensity structure measured in beam time in March 2019 [5]. The measurement time resolution is 20 $\mu$s. The simulated event rate is proportional to the beam intensity and obeys a Poisson process within each 20 $\mu$s time slot. Similar improvement of data quality with throttling can be observed in figure 4b. Without throttling, the high intensity fluctuations lead to event losses already at rates below the bandwidth limit. The Stop strategy significantly improves the number of good events over all hit rates, and provides some advantage already at rates below the bandwidth limit.
5. Conclusion

This time-based throttling without event information, is designed to adapt data acquisition to the high interaction rate of the CBM experiment with beam intensity fluctuation. The self-triggered front-end electronics will asynchronously deliver time-stamped data messages on activation of the respective detector channel. Meanwhile, when busy conditions in the front-end buffers are detected, the throttling system will disable the hit data taking of the whole experiment until the buffers of front-end electronics are emptied. Ideally, events saved into the large computer farm are almost complete. According to our simulation both in stable and fluctuating beam structure, the probability of good event reconstruction is significantly improved with the applied throttling mechanism.

References


