The CBM Experiment @ FAIR - New challenges for Front-End Electronics, Data Acquisition and Trigger Systems

Walter F.J. Müller for the CBM collaboration

1 GSI, Darmstadt
E-mail: w.f.j.mueller@gsi.de

Abstract. The 'Compressed Baryonic Matter' (CBM) experiment at the new 'Facility for Antiproton and Ion Research' (FAIR) in Darmstadt is designed to study the properties of highly compressed baryonic matter produced in nucleus-nucleus collisions in the 10 to 45 A GeV energy range. One of the key observables is hidden ($J/\psi$) and open ($D^0$, $D^\pm$) charm production. To achieve an adequate sensitivity extremely high interaction rates of up to $10^7$ events/second are required, resulting in major technological challenges for the detectors, front-end electronics and data processing. The front-end electronics will be self-triggered, autonomously detect particle hits, and output hit parameter together with a precise absolute time-stamp. Several layers of feature extraction and event selection will reduce the primary data flow of about 1 TByte/sec to a level of 1 GByte/sec. This new architecture avoids many limitations of conventional DAQ/Trigger systems and is for example essential for open charm detection, which requires the reconstruction of displaced vertices, in a high-rate heavy ion environment.

Introduction

The exploration of the phase diagram of strongly interacting matter is the leading theme in heavy-ion physics. Of particular interest is the transition from hadronic to partonic degrees of freedom which is expected to occur at high temperatures and/or high baryon densities. The experiments at SPS, RICH, and LHC explore the regime of low baryon density and high temperature, and thus probe the conditions important in the development of the early universe. The planned 'Compressed Baryonic Matter' (CBM) experiment at the new 'Facility for Antiproton and Ion Research' (FAIR) in Darmstadt focuses at the regime of high baryon density and modest temperatures, realized in nature for example in neutron stars. It is designed to study the properties of highly compressed baryonic matter produced in nucleus-nucleus collisions in the 10 to 45 A GeV energy range. The key observables are

- Low-mass vector mesons. The in-medium spectral function of short-lived vector mesons will be measured directly via their decay into dilepton pairs.
- Charm production. The effective masses of D mesons are expected to be modified in dense matter, which leads to a change of the relative abundance of charmonium and D mesons. The anomalous suppression of charmonium due to the screening effects of a partonic medium allows to probe the onset of QGP formation. It is thus important to measure the production of open charm ($D^0$, $D^\pm$) as well as hidden charm ($J/\psi$).
Strangeness production.
Event-by-event fluctuations.

A more detailed account of the scientific program and the experimental setup of CBM is given in a separate contribution to these proceedings [1], a full description is given in the recent 'CBM Technical Status Report' [2].

Data Acquisition and Event Selection

Many of the signatures pursued with the CBM experiment are based on rare processes. To achieve an adequate sensitivity, the detector systems are designed to operate at interaction rates of up to 10 MHz for A-A collisions and up to several 100 MHz for p-p and p-A collisions. It is the task of the data acquisition and event selection system to identify the candidate events for the physics signals under study and send them to archival storage. The most challenging aspect is here the measurement of open and hidden charm production in heavy ion collisions down to very low cross sections. The D mesons will be identified via the displaced vertices of their hadronic decays, the decision for selecting candidate events thus requires tracking, primary vertex reconstruction, and secondary vertex finding in the STS. In addition, the system has to be configurable to handle a wide range of physics signals, ranging from D and $J/\psi$ in A-A collisions over low-mass dileptons to $\Upsilon$ in p-p and p-A collisions.

The conventional system design with triggered front-end electronics allows to keep the event information for a limited time, usually a few $\mu$s, in the front-end electronics while a fast first level trigger decision is determined from a subset of the data. Upon a positive trigger decision, the data acquisition system transports the selected event to higher level trigger processing or archival storage. A system with such a fixed trigger latency constraint is not well matched to the complex algorithms needed for a D trigger, especially in the case of heavy ion interactions, where the multiplicities and thus the numerical effort needed for a decision varies strongly from event to event.

The concept adopted for CBM will use self-triggered front-end electronics, where each particle hit is autonomously detected and the measured hit parameters are stored with precise timestamps in large buffer pools. The event building, done by evaluating the time correlation of hits, and the selection of interesting events is then performed by processing resources accessing these buffers via a high speed network fabric. The large size of the buffer pool ensures that the essential performance factor is the total computational throughput rather than decision latency. Since we avoid dedicated trigger data-paths, all detectors can contribute to event selection decisions at all levels, yielding the required flexibility to cope with different operation modes.

In this approach we have no physical trigger signal which prompts a data acquisition system to read a selected event and transport it to further processing or storage. The role of the data acquisition system is to transport data from the front-end to processing resources and finally to archival storage. The event selection is done in several layers of processing resources, reminiscent of the trigger level hierarchy in conventional systems.

One consequence of using self-triggered front-end electronics is a much higher data flow coming from the front-ends on the detector. For CBM a data rate of about 1 TByte/sec is expected. However, communication cost is currently improving faster over time than processing cost, an observation sometimes termed Gilder’s law, making such a concept not only feasible but also cost effective.

Overall Architecture
The communication and processing needed between the front-end electronics, generating digitized detector information, and the archival storage, where the complete context of selected candidate events is recorded, can be structured and organized in several ways. The current
baseline solution for CBM is guided by two principles: processing is done after event building and it is done in a structured processor farm. It is well adapted to the type of processing needed in the CBM experiment and leads to a straightforward and modular architecture.

A logical data flow diagram is shown in Fig. 1, indicating the data sources and processing elements as boxes and every form of interconnection networks as ovals. The main components are:

- **Front-end electronics (FEE):** The front-end detects autonomously every particle hit and sends the hit parameters together with a precise timestamp and channel address information over the concentrator network (CNet) to a buffer pool. A rough estimate for the data volume generated by a detector channel can be deduced from typical CBM operation and detector parameters: 10 MHz interaction rate, 10% occupancy for central collisions, a ratio of 1/4 for minimum bias to central multiplicity, and a typical cluster size of 3 fired electronics channels per physical particle hit gives a channel count rate of about 750 kHz. Assuming 8 byte per hit yields a data flow of about 6 MB/sec and channel. For a typical FEE unit with 16 channels this results in a data rate of 100 MB/sec which can be transported over a single GBit serial link.

- **Clock and time distribution (TNet):** The timestamps of each hit are used to associate hits with events but also in drift and flight time measurements. Thus not only a time scale, in practice a frequency, but also information about the absolute time has to be communicated to all front-end units. The most stringent requirements come from the detectors involved in time-of-flight measurements, where the contribution from the clock jitter should be below 25 ps sigma. The most straightforward approach is to distribute a common clock frequency and to provide a mechanism for broadcasting information with clock cycle precise latency to...
all units. The minimal required functionality is a global clock reset at the begin of the measurement, or alternatively, distribution of tick marks every second as provided by the planned campus-wide frequency and time normal. The TNet is thus a dedicated broadcast network, connecting a central controller logically with all front-end units. The last hop to the front-end units may be implemented with the part of the CNet infrastructure, as indicated by the connection of TNet to CNet in Fig. 1.

- **Concentrator Network (CNet):** The role of the concentrator network is to collect the data from the individual front-end units and aggregate the traffic on a set of high speed links which connect the detector with the area where the data buffers and the data processing is located. A rough estimate for the total data rate is 1 TB/sec which could be finally transported off the detector with about 1000 links with 10 Gbps each. The simplest implementation of the CNet is a collection of independent concentrator trees, one for each high speed link. However, a better load balancing or an appropriate degree of failure tolerance is likely to call for a more connected topology.

In addition to the hit data transport from front-end to buffering, other communication tasks like control traffic or time distribution can be handled by the CNet infrastructure. Such an integrated approach is especially useful in conjunction with low-cost optical links.

- **Active Buffers:** The next stage in the data flow is a large buffer pool. The units are indicated as magenta boxes in Fig. 1. They are dubbed 'active buffers' because the data is not only stored but potentially also reformatted and reorganized. They are also hand-over points between different types of networks, thus logically separating them and allowing to use different technologies in CNet, BNet, and PNet.

- **Build Network (BNet):** The data arrives from the detector in about 10³ parallel streams, each reflecting a small section of a detector sub-system. For the event selection processing, the information of an event has to be assembled in a farm node. This data reorganization is performed by the build network and the active buffers.

In a conventional system, a trigger already defines the context of an event, so all further data processing and transport can be organized in terms of events starting at the FEE. In our case, the FEE sends a stream of time-stamped hits, and it is one of the tasks of the data processing, to first identify at what times interactions occur, and in a second step, to associate the hits with those events. This event tagging processing can be done before or after data traverses the BNet. In the first case, event tagging is handled in the active buffers, and the entities being assembled in the BNet transfers are indeed events. In the second case, only the timestamp information is available, and it is thus natural to assemble all the data of a time interval. A strict event-by-event approach would lead at the nominal Au+Au interaction rate of 10 MHz to a message rate of 10¹⁰ messages per second with an average message size of 100 Bytes. However, because the transfer latency is uncritical, it is possible to choose a bigger dispatching unit, either an interval of events, or in the simplest case, a time interval containing a significant number of events. This aggregation reduces the message rate, increases message size, and because event size fluctuations average out, also yields a more uniform message size distribution. A reasonable choice is an event interval of about 100 events or equivalently a time interval in the order of 10 µs.

The simplest solution is a time interval based build logic with a shaping and traffic scheduling setup similar to the one developed for the LHCb first level trigger [3]. More involved solutions, which support building event intervals and the suppression of event incoherent backgrounds, are currently being investigated.

Fig. 1 indicates that source as well as destination of a BNet transfer is an active buffer. They implement the protocol used on the BNet and are responsible for organization of the data flow, in particular for traffic shaping and appropriate scheduling of transfers. Because
the actual traffic seen by the BNet can be controlled to a large degree and adapted to a given networking technology, it is assumed that the BNet can be based on a commercial off-the-shelf (COTS) technology. Plausible candidates are Ethernet, Infiniband, or the emerging Advanced Switching Interconnect (ASI).

Fig. 1 only indicates the logical data flow, not a concrete network topology. It is possible to merge the functionality of the two active buffer layers, one interfacing CNet to BNet and one interfacing BNet to PNet, into a single entity, leading to a system with half as many BNet physical ports and bidirectional traffic on all BNet links. Also, it is possible to factorize the network in several ways, which allows to build the BNet with a set of medium-sized switches and thus to exploit the usually significantly better price/port ratio of smaller switches.

- **Processing Resources:** The first level of event selection processing has to handle the full event rate, and depending on how many detector sub-systems are involved in the decision, a substantial fraction of the total data volume. A very rough estimate shows, that processing a data flow on the scale of a TByte/sec is likely to require a computational bandwidth on the scale of $10^{15}$ operations/sec. With today's technology, the most promising approach is a hybrid system using a combination of hardware processors, implemented with programmable logic components like FPGA's, and software processors, implemented with commodity PC's. The kernels of algorithms which allow highly parallel execution are done in hardware processors, the rest in software processors. The aim is to execute most of the operations on hardware processors, which offer the best price/performance ratio for computational bandwidth, but to keep most of the code volume on software processors, which offer much easier program development and maintenance.

Since programmable logic devices are essentially arrays of simple structures, it is expected that they scale well and thus density and speed will improve as the underlying silicon technology evolves. The development for software processors over the relevant time scale till the design freeze of the CBM data processing is likely to be more complex. The evolution of single processor speed has reached apparent limits of complexity and power dissipation, calling for changes in concepts and architectures. This trend is already apparent in recent developments like the compute ASIC for IBM's Blue Gene system or the Sony-Toshiba-IBM Cell processor. It is also expected, that conventional fixed instruction set processor and programmable logic concepts will be coupled, resulting new forms of configurable computing, like processors with dynamically adaptable instruction sets. The current R&D efforts are based on physically separated hardware and software processors because these are the mainstream products available today and the near future. The overall architecture is, however, easily adaptable to more integrated forms of configurable computing.

- **Processing Network (PNet):** The processing resources are grouped in farm nodes. Each farm node is organized around a local processing network which provides the communication between the associated processing resources and active buffers, which act as central data repository and as gateway to the BNet. The PNet is thus structured into many local networks.

The number of hardware and software processors aggregated to one farm node is determined by the amount of resources needed to efficiently handle all the algorithms needed for an event selection decision. Each hardware processor will have a dedicated configuration to execute a particular algorithm, a total about a dozen different configurations may be needed. A rough estimate gives, that an about equal number of software processors is needed for a balanced load of the whole system.

It is expected that the resources of a farm are concentrated in a crate or are at least in close proximity. The PNet can therefore use technologies designed for short distance chip-to-chip and board-to-board interconnects, plausible candidates are from today's perspective PCIe express or ASI. As stated already for the BNet, Fig. 1 only indicates the logical data
flow, not a concrete network topology. The PNet can be a homogeneous, single technology, switch based star network as suggested by the figure, but many other topologies are possible.

- **High-level Network (HNet):** The task of the event selection processing described up to now is to perform a first reduction, similar to the ‘level 1 trigger’ in conventional systems. A further reduction will be needed to reduce the data volume to a level suitable for archival storage. The HNet provides the connection to this high-level computing.

**Summary**

The measurement of open charm is the key factor shaping the architecture of front-end electronics, data acquisition, and event processing in CBM, leading to an architecture with self-triggered front ends without a conventional trigger.

The key R&D areas for CBM are the development of self-triggered front-end electronics with adequate output bandwidth, of a data processing system which allows to combine the advantages of hardware processors built from programmable logic and software processors, and last but not least, of highly efficient feature extraction and event selection algorithms adapted to such a processing environment.

**Acknowledgements**

This work is supported by the EU Integrated Infrastructure Initiative Hadron-Physics Project under Contract number RII3-CT-2004-506078.

**References**

[1] P. Senger, contribution to these proceedings
[2] Compressed Baryonic Matter Experiment - Technical Status Report. January 2005. http://www.gsi.de/documents/DOC-2005-Feb-447-1.pdf
[3] D. Atanasov, I. Kisel V. Lindenstruth, H. Müller, D. Altmann, A. G. Elias, F. V. dos Santos: A scalable 1 MHz Trigger Farm Prototype with Event-Coherent DMA input, Proceedings of the 13th Real Time Conference 2003, Valencia Spain