

GPU real-time processing in NA62 trigger system

R. Ammendola¹, A. Biagioni², S. Chiozzi³, P. Cretaro²,
S. Di Lorenzo^{4,5}, R. Fantechi⁵, M. Fiorini^{6,3}, O. Frezza²,
G. Lamanna⁷, F. Lo Cicero², A. Lonardo², M. Martinelli², I. Neri³,
P. S. Paolucci², E. Pastorelli², R. Piandani⁵, M. Piccini⁸,
L. Pontisso⁵, D. Rossetti⁹, F. Simula², M. Sozzi^{4,5} and P. Vicini²

1 INFN Sezione di Roma Tor Vergata, Roma, Italy

2 INFN Sezione di Roma, Roma, Italy

3 INFN Sezione di Ferrara, Ferrara, Italy

4 Università degli Studi di Pisa, Pisa, Italy

5 INFN Sezione di Pisa, Pisa, Italy

6 Università degli Studi di Ferrara, Ferrara, Italy

7 INFN LNF, Frascati, Italy

8 INFN Sezione di Perugia, Perugia, Italy

9 NVIDIA Corp., Santa Clara, CA, USA

E-mail: gianluca.lamanna@cern.ch

Abstract.

A commercial Graphics Processing Unit (GPU) is used to build a fast Level 0 (L0) trigger system tested parasitically with the TDAQ (Trigger and Data Acquisition systems) of the NA62 experiment at CERN. In particular, the parallel computing power of the GPU is exploited to perform real-time fitting in the Ring Imaging Cherenkov (RICH) detector. Direct GPU communication using a FPGA-based board has been used to reduce the data transmission latency. The performance of the system for multi-ring reconstruction obtained during the NA62 physics run will be presented.

1. Introduction

General Purpose computing on GPU (GPGPU) is nowadays widespread in scientific areas requiring large processing power such as computational astrophysics, lattice QCD calculations, image reconstruction for medical diagnostics. In a High Energy Physics experiments are using (i.e. ALICE [1]) or are considering to use (i.e. ATLAS [2]) GPUs for high level triggers.

Low level triggers could also benefit from GPUs computing power to build more refined physics-related trigger primitives and to design more selective condition in order to reduce the load to the higher level without sacrificing interesting physics events; Three main aspects must be carefully evaluated in order to establish the real-time capability of this kind of processors: computing power, total processing latency and its stability in time. In the present paper a description is provided of the GPU-based L0 trigger integrating in the experimental setup of the RICH detector of the NA62 experiment in order to reconstruct the ring-shaped hit patterns; we report and discuss results obtained with this system along with the algorithms that will be implemented.



2. The NA62 experiment, trigger and RICH detector

The NA62 experiment at CERN [3] has the goal of measuring the branching ratio of the ultra-rare decay of the charged kaon into a pion and a neutrino anti-neutrino pair.

The trigger is a key system to obtain such a result. In its standard implementation, the FPGAs on the readout boards of each sub-detector participating to the L0 trigger, compute simple trigger primitives. The maximum latency allowed for the synchronous L0 trigger was chosen to be rather large in NA62 (up to 1 ms). Such a time budget allows in principle the use of more complex but slower trigger. This would have the benefit of increasing the trigger selectivity for the $K^+ \rightarrow \pi^+ \nu \bar{\nu}$ process, and would allow the design of new trigger condition to collect additional physics processes.

The Ring Imaging Cherenkov (RICH) detector is a key element for particle identification in the NA62 experiment and one of the subdetectors responsible for the L0 trigger [4]. Its main purpose is to separate pions from muons in the momentum range from 15 GeV/c to 35 GeV/c and to measure the particle arrival time with better than 100 ps resolution [5]. Thanks to its fast response, it is used in the L0 trigger by providing a multiplicity count. However, much more information would be available from the RICH. Due to the loose relation between hits multiplicity and physics related quantities, evaluating this information at the online trigger stage would allow to build most effective trigger conditions.

3. GPU computing at L0

In a standard approach for GPU computing, data from a detector reach a Network Interface Card (NIC) which copies them periodically on a dedicated area in the PC RAM. Here, a sufficient data load of buffered events is usually prepared and then copied to GPU memory through the PCI express bus. The PC on which the GPU card is plugged, i.e. the host, has the role of starting the GPU “kernel”.

In contrast with respect to the applications for which GPUs have been originally designed, in triggering the computing latency is a very important issue. The total latency is essentially composed by two parts: the time to bring the data on the GPU processor and the effective time of processing. With the aim of decreasing the transport latency (and its fluctuations) for L0 applications a direct data transfer protocol from a custom FPGA-based NIC to the GPU has been used. Dedicated pattern recognition algorithms have been developed to exploit the GPU parallel structure. Both these aspects will be discussed below.

NaNet is a low-latency NIC with GPUdirect capability developed at INFN Rome division. Its design comes from the APENet+ card logic [6] and, in the current implementation, it provides 10 Gb/s communication (NaNet-10) [7]. NaNet board is able to exploit the GPUDirect peer-to-peer (P2P) and RDMA capabilities of NVIDIA Tesla GPUs injecting into their memory an UDP input data stream from the detector front-end. With respect to the previous version (NaNet-1 with 1 Gb/s links) NaNet-10 board is able to cope with higher data rate and allow to receive data from 4 TEL62 RICH readout board [4], providing on-the-fly data processing to perform tasks as data de-compression and event fragments coalescing.

A fast ring finding algorithm is a crucial point to allow the use of RICH in on-line trigger selection. The processing time must be small enough to cope with high input data rate. The parallel structure of the GPUs helps in increasing the computing throughput both allowing the parallel processing of several event concurrently and exploiting pattern recognition parallel algorithms. We focused in particular on two multi-rings standalone pattern recognition algorithms based only on geometrical considerations (as no information from other detectors is available at this level). In case of the first algorithm, named Histogram, the XY plane is divided into a grid and a histogram is created with distances from the grid points and the PMT hits. Rings are identified looking at distance bins whose contents exceed a predefined threshold value. The second algorithm, i.e. Almagest [8], is based on Ptolemy Theorem. Both these algorithms

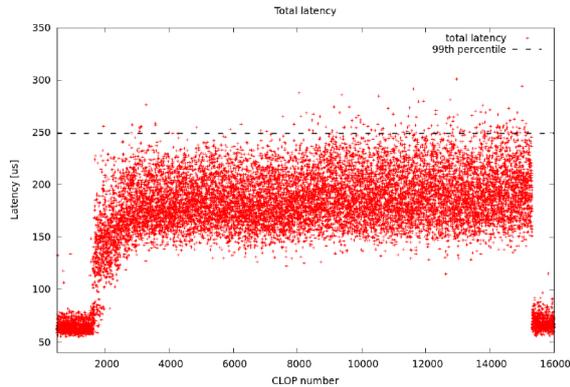


Figure 1. Total latency during an entire burst.

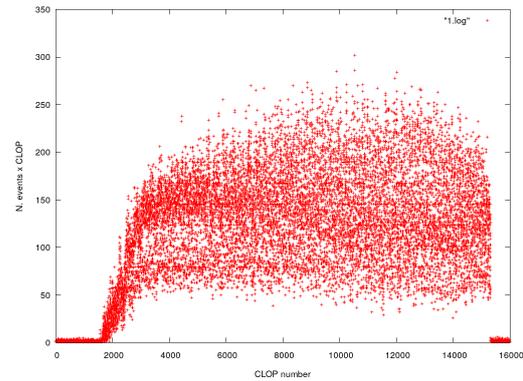


Figure 2. Number of events per CLOP during the burst.

are used for pattern matching, once the number of rings and points belonging to them have been found, it is possible to apply Crawford's method [9] to obtain centre coordinates and radii with good spatial resolution.

4. Results

The whole system, equipped with the NaNet-10 board, has been tested during the normal NA62 data taking, with data coming from the four RICH TEL62 readout boards.

The latency measurement is done by exploiting an hardware clock on the FPGA synchronized with the NA62 general clock through a TTC [10] interface. The resolution of this clock is 25 ns. The clock counter is reset at each start of burst (the burst is a bunch of particles produced by the protons extracted from the SPS. The burst duration is about 6 s each 14 s, the duty cycle depends on the SPS extraction scheme).

In fig.1 the total latency including all the components is shown for an entire burst. In fig.2 the number of events per CLOP is shown for an entire burst. The main source of the width of the latency distribution is due to fluctuation in the number of events buffered before the processing: in order to better exploit the parallel structure of the GPU the data are processed in buffers collected in 350 μ s, called CLOP (Circular List of Persistent buffers). The latency for processing of a single datagram in the NaNet-10 board, since its receiving in the 10GbE MAC to the completion of the DMA operation towards the GPU memory, accounts for less than 2 μ s. In 2016, at 1/3 of the nominal intensity, each CLOP contains in average about 150 events (fig.2). Nevertheless, the processing time per event, in the present version of the computing kernel, is quite stable below 1 μ s (using a single GPU). The interaction with the present central trigger processor (L0TP) needs a further stage of synchronization. The send of the GPU results, codified in a format compatible with the present NA62 L0 trigger, is done directly from the GPU memory with a low latency RDMA, still through the NaNet-10 board.

5. Conclusions

The GPU system performed as expected in the NA62 2016 run. In the next physics run, the GPU system can be used to provide high quality primitives to increase the purity of the main triggers and to give handle to build selective triggers to enlarge the NA62 trigger menu. Current results for the processing time per event hints to the further development of GPU algorithms and/or to the use of multiple GPU devices to reach an order of magnitude speed-up to cope with the experiment nominal intensity. Additional details on the results shown here can be found in [11].

Acknowledgments

This work was partially funded by the ExaNeSt project (H2020 grant no. 671553). S. Di Lorenzo, M. Fiorini, G. Lamanna, I. Neri, R. Piandani, L. Pontisso, M. Sozzi thank the GAP project, partially supported by MIUR under grant RBFR12JF2Z Futuro in ricerca 2012.

References

- [1] D. Rohr, S. Gorbunov, A. Szostak, M. Kretz, T. Kollegger, T. Breitner, T. Alt, Journal of Physics: Conference Series 396 (1) (2012) 012044.
- [2] A. T. Delgado, P. C. Muno, J. A. Soares, R. Gonalo, J. Baines, T. Bold, D. Emeliyanov, S. Kama, M. Bauce, A. Messina, M. Negrini, A. Sidoti, L. Rinaldi, S. Tupputi, Z. D. Greenwood, A. Elliott, S. Laosooksathit, An evaluation of gpus for use in an upgraded atlas high level trigger, ATLAS-DAQ-PROC-2015-061, (November 2015).
- [3] G. Ruggero on behalf of NA62 Collaboration, in KAON2016 Proceedings
- [4] B. Angelucci *et al.*, JINST **9** (2014) no.01, C01055. doi:10.1088/1748-0221/9/01/C01055
- [5] D. Aisa *et al.*, Nucl. Instrum. Meth. A **766** (2014) 1. doi:10.1016/j.nima.2014.06.089
- [6] R. Ammendola, A. Biagioni, O. Frezza, F. Lo Cicero, A. Lonardo, P. S. Paolucci, D. Rossetti and F. Simula *et al.*, J. Phys. Conf. Ser. **396** (2012) 042059.
- [7] R. Ammendola *et al.*, JINST **11** (2016) no.03, C03030. doi:10.1088/1748-0221/11/03/C03030
- [8] G. Lamanna, Nucl. Instrum. Meth. A **766** (2014) 241. doi:10.1016/j.nima.2014.05.073
- [9] J. F. Crawford, Nucl. Instrum. Meth. **211** (1983) 223. doi:10.1016/0167-5087(83)90575-6
- [10] S. Baron Timing, trigger and control (TTC) systems for the LHC, <http://ttc.web.cern.ch/ttc/intro.html>
- [11] A. Lonardo, Proceedings of TWEPP2016 conference (to be published)