

GPU for triggering in HEP experiments

R. Ammendola[†], A. Biagioni^{*}, P. Cretaro^{*}, S. Di Lorenzo[‡], R. Fantechi^{‡§}, M. Fiorini^{||}, O. Frezza^{*}, G. Lamanna[¶],
F. Lo Cicero^{*}, A. Lonardo^{*}, M. Martinelli^{*}, I. Neri^{||}, P. S. Paolucci^{*}, E. Pastorelli^{*}, R. Piandani[‡], L. Pontisso[‡],
D. Rossetti^{**}, F. Simula^{*}, M. Sozzi^{‡§}, and P. Vicini^{*}

Abstract—Over the last few years the GPGPU (General-Purpose computing on Graphics Processing Units) paradigm represented a remarkable development in the world of computing. Computing for High-Energy Physics is no exception: several works have demonstrated the effectiveness of the integration of GPU-based systems in high level trigger of different experiments. On the other hand the use of GPUs in the low level trigger systems, characterized by stringent real-time constraints, such as tight time budget and high throughput, poses several challenges. In this paper we focus on the low level trigger in the CERN NA62 experiment, investigating the use of real-time computing on GPUs in a synchronous system. Our approach aimed at harvesting the GPU computing power to build in real-time refined physics-related trigger primitives for the RICH detector, as the the knowledge of Čerenkov rings parameters, allows to build stringent conditions for data selection at trigger level. Latencies of all components of the trigger chain have been analyzed, pointing out that networking is the most critical one. To keep the latency of data transfer task under control, we devised NaNet, an FPGA-based PCIe Network Interface Card (NIC) with GPUDirect capabilities. For the processing task, we developed specific multiple ring trigger algorithms to exploiting the parallel architecture of GPUs to increase the processing throughput in order to sustain the high event rate. Results obtained during the first months of 2016 NA62 run are presented and discussed.

I. A GPU-BASED TRIGGER FOR NA62 EXPERIMENT'S RICH

IN High Energy Physics experiments the realtime selection of the most interesting events is of paramount importance because of the collision rates which do not give the possibility to save all the data for offline analysis. For this purpose, different trigger levels are usually used to select the most meaningful events. The low level trigger usually requires low and (almost) deterministic latency and its standard implementation is on dedicated hardware (ASICs or FPGAs). The GAP project aims at studying the usage of Graphic Processing Units (GPUs) in order to build refined physics-related trigger primitives with a offline quality and therefore leading to a net improvement of trigger conditions and data handling. While GPUs execution times are rather stable, also I/O tasks have to guarantee real-time features along the data stream path, from detectors to GPU memories. In order to reduce and control the latency due to data transfer, a dedicated Network Interface

Card (NIC) has been designed and developed within the INFN funded project NaNet.

A. NaNet architecture

The design of a low-latency, high-throughput data transport mechanism for real-time systems is mandatory in order to bridge the front-end electronics and the software trigger computing nodes of High Energy Physics Experiments [1]. NaNet, being an FPGA-based NIC, natively supports a variety of link technologies allowing for a straightforward integration in different experimental setups. Its key characteristics are i) the management of custom and standard network protocols in hardware, in order to avoid OS jitter effects and guarantee a deterministic behaviour of communication latency while achieving maximum capability of the adopted channel; ii) a processing stage which is able to reorganize data coming from detectors on the fly, in order to improve the efficiency of applications running on computing nodes; iii) data transfers to or from application memory are directly managed avoiding bounce buffers.

NaNet-1 was developed in order to verify the feasibility of the project; it is a PCIe Gen2 x8 network interface card featuring GPUDirect RDMA over GbE.

NaNet-10 is a PCIe Gen2 x8 network adapter implemented on the Terasic DE5-net board equipped with an Altera Stratix V FPGA featuring four SFP+ cages [2].

B. Trigger setup and implementation

In the RICH detector, Čerenkov light is reflected by a composite mirror with a focal length of 17 m focused onto two separated spots equipped with ~ 1000 photomultipliers (PM) each. The final system consists of 4 GbE links to move primitives data from the readout boards to the GPU_L0TP (see Fig. 1). Data communication between the readout boards (TEL62) and the L0 trigger processor happens over multiple GbE links using UDP streams. The overall time budget for the low level trigger comprising both communication and computation tasks is of 1 ms, so a deterministic response latency from GPU_L0TP is a strict requirement. Refined primitives coming from the GPU-based calculation will be then sent to the central L0 processor, where the trigger decision is made taking in account informations from other detectors.

C. Algorithms for multi-ring reconstruction on GPU

Taking the parameters of Čerenkov rings into account could be very useful in order to build stringent conditions for data

^{*}INFN Sezione di Roma, Italy.

[†]INFN Sezione di Tor Vergata, Italy.

[‡]INFN Sezione di Pisa, Italy.

[§]CERN, Switzerland.

[¶]INFN Laboratori Nazionali di Frascati, Italy.

^{||}INFN Sezione di Ferrara, Italy.

^{**}NVIDIA Corporation, U.S.A.

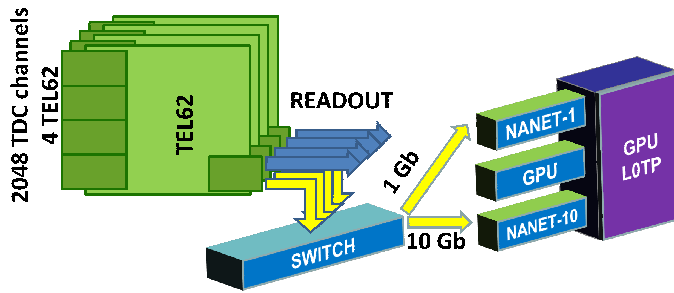


Fig. 1. Pictorial view of GPU-based Trigger.

selection at trigger level. This implies that circles have to be reconstructed using the coordinates of activated PMs.

We take in consideration two multi-rings pattern recognition algorithms based only on geometrical considerations (no other information is available at this level) and particularly suitable for exploiting the intrinsic parallel architecture of GPUs: i) the first based on histograms built with distances between the hits of physics event and points of a grid; ii) the Almagest algorithm that hinges on Ptolemy’s theorem, stating that when four vertices of a quadrilateral (ABCD) lie on a common circle, it is possible to relate four sides and two diagonals: $|AC| \times |BD| = |AB| \times |CD| + |BC| \times |AD|$. [3]

II. RESULTS

1) *2015 Run*: The GPU-based low level trigger included 2 TEL62 boards connected to a HP2920 switch and a NaNet-1 [4] board with a TTC¹ HSMC daughtercard plugged into a server made of a X9DRG-QF dual socket motherboard populated with Intel Xeon E5-2620 @2.00 GHz CPUs (*i.e.* Ivy Bridge architecture), 32 GB of DDR3 RAM and a Kepler-class NVIDIA K20c GPU. We tested the whole chain of the trigger system: the data events arriving by means of the GPUDirect RDMA interface within a configurable time frame, are gathered and then organized in a Circular List Of Persistent buffers (CLOP) in the GPU memory.

Events coming from different TEL62 need to be merged in the GPU memory before the launch of the ring reconstruction kernel. Results are reported in Fig. 2. The CLOP size measured as number of received events is on the X-axis and the latencies of different stages are on the Y-axis. The computing kernel implemented the histogram fitter with a single step (*i.e.* using an 8x8 grid only). Events coming from 2 readout boards, for a gathering time of 400 μ s, and parameters like events rate (collected with a beam intensity of 4×10^{11} protons per spill), a CLOP’s size of 8KB, time frame was chosen so that we could test the online behaviour of the trigger chain. The merge operation does not expose much parallelism, requiring instead synchronization and serialization. As a result it is an ill-suited problem to the GPU architecture. In operative conditions, the merging time only would exceed the time frame. The high latency of this task suggests to offload it to a dedicated implementation in the FPGA Data Processing stage [5].

¹Timing, Trigger and Control (TTC) Systems for the LHC (<http://ttc.web.cern.ch/TTC/>)

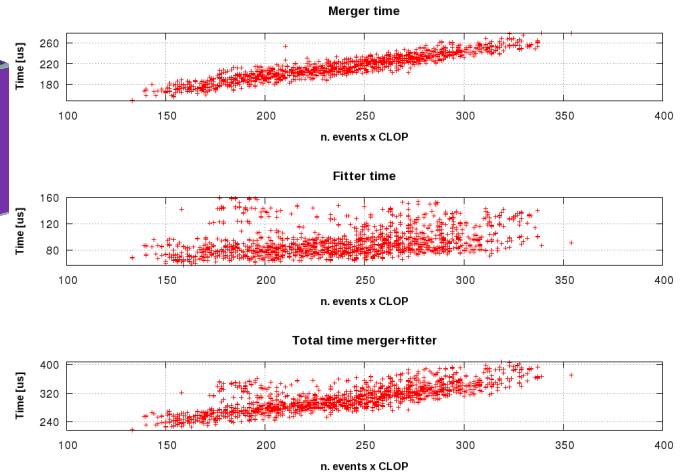


Fig. 2. Multi-ring reconstruction of events performed on K20c NVIDIA GPU.

2) *2016 Run*: To cope with the effective conditions, the NaNet-10 board, currently installed in the experiment, implements in firmware this merging stage. In this way the GPU will perform only the ring reconstruction, allowing to cope with the nominal events rate of 10 MHz. Besides the 10 Gb link provides enough bandwidth to manage the 4 read-out boards connected to the GPU_L0TP. During this Run we expect to keep working parasitically with respect to the standard trigger, collecting data and measuring throughput and total latency of the upgraded system. The primitives produced with the GPU-based trigger will be jointly used with the standard ones to build selective conditions for searching of very rare decay channels concurrently with the main data collection in NA62.

ACKNOWLEDGMENT

S. Di Lorenzo, R. Fantechi, M. Fiorini, 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] A. Lonardo, F. Ameli, R. Ammendola, A. Biagioni, A. Cotta Ramusino, M. Fiorini, O. Frezza, G. Lamanna, F. Lo Cicero, M. Martinelli, I. Neri, P.S. Paolucci, E. Pastorelli, L. Pontisso, D. Rossetti, F. Simeone, F. Simula, M. Sozzi, L. Tosoratto and P. Vicini, “NaNet: a Configurable NIC Bridging the Gap Between HPC and Real-time HEP GPU Computing,” *Journal of Instrumentation*, vol. 10, no. 04, p. C04011, 2015. [Online]. Available: <http://stacks.iop.org/1748-0221/10/i=04/a=C04011>
- [2] R. Ammendola, A. Biagioni, M. Fiorini, O. Frezza, A. Lonardo, G. Lamanna, F. Lo Cicero, M. Martinelli, I. Neri, P.S. Paolucci, E. Pastorelli, L. Pontisso, D. Rossetti, F. Simula, M. Sozzi, L. Tosoratto, and P. Vicini, “NaNet-10: a 10GbE network interface card for the GPU-based low-level trigger of the NA62 RICH detector,” *Journal of Instrumentation*, vol. 11, no. 03, p. C03030, 2016. [Online]. Available: [url=http://stacks.iop.org/1748-0221/11/i=03/a=C03030](http://stacks.iop.org/1748-0221/11/i=03/a=C03030)
- [3] G. Lamanna, *Almagest, a new trackless ring finding algorithm*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **766** (2014) 241 – 244. {RICH2013} Proceedings of the Eighth International Workshop on Ring Imaging Cherenkov Detectors Shonan, Kanagawa, Japan, December 2-6, 2013.

- [4] R. Ammendola, A. Biagioni, O. Frezza, G. Lamanna, A. Lonardo, F. L. Cicero, P. S. Paolucci, F. Pantaleo, D. Rossetti, F. Simula, M. Sozzi, L. Tosoratto, and P. Vicini, "Nanet: a flexible and configurable low-latency nic for real-time trigger systems based on gpus," *Journal of Instrumentation*, vol. 9, no. 02, p. C02023, 2014. [Online]. Available: <http://stacks.iop.org/1748-0221/9/i=02/a=C02023>
- [5] R. Ammendola, A. Biagioni, P. Cretaro, S. Di Lorenzo, R. Fantechi, M. Fiorini, O. Frezza, G. Lamanna, F. Lo Cicero, A. Lonardo, M. Martinelli, I. Neri, P. S. Paolucci, E. Pastorelli, R. Piandani, L. Pontisso, D. Rossetti, F. Simula, M. Sozzi and P. Vicini, "GPU-based Real-time Triggering in the NA62 Experiment," *arXiv* 2016. [Online]. Available: [url=https://arxiv.org/abs/x.y](https://arxiv.org/abs/x.y)