



Contribution ID: 162

Type: Poster

NaNet: a Configurable NIC Bridging the Gap Between HPC and Real-time HEP GPU Computing

Wednesday, 24 September 2014 17:55 (1 minute)

NaNet is a FPGA-BASED PCIe Network Interface Card with GPUDirect capability featuring a configurable set of channels: standard 1/10GbE and custom 34Gbps APElink and 2.5Gbps optical with deterministic latency KM3link.

GPUDirect feature combined with a transport layer offload module and a data stream processing stage makes NaNet a low-latency NIC suitable for real-time GPU processing.

We will describe NaNet architecture and its performances, and present two use cases for it: the GPU-based low-level trigger for the RICH detector in NA62 experiment and the on-/off-shore data link for KM3 underwater neutrino telescope.

Summary

Although GPGPU is widely accepted as an effective approach to high performance computing, its adoption in low-latency, hard real-time processing systems, like low level triggers in HEP experiments, still poses several challenges.

GPUs show a rather deterministic behaviour in terms of processing latency once input data are available in their internal memories, but assessment of the real-time features of a whole GPGPU system takes a careful characterization of all subsystems along data stream path. In our analysis we identified the networking subsystem as the most critical one because of the relevant fluctuations in its response latency.

To overcome this issue, we designed NaNet, a FPGA-based PCIe Network Interface Card (NIC) featuring a configurable set of network channels and capable of receiving and sending data directly to and from Nvidia Fermi/Kepler GPU internal memories without intermediate buffering on host memory (GPUDirect).

The design includes a transport layer offload module with cycle-accurate deterministic latency, with support for UDP and custom KM3link and APElink protocols, added to eliminate host OS intervention on data stream and thus avoiding a possible source of jitter.

NaNet design currently supports both standard - 1GbE (1000Base-T) and 10GbE (10Base-R) - and custom - 34Gbps APElink and 2.5Gbps deterministic latency KM3link - channels, but its modularity allows for a straightforward inclusion of other link technologies.

An application specific module operates on input/output data streams, performing processing on them with cycle-accurate deterministic latency (e.g. to perform decompression and to rearrange data structures in a GPU-friendly fashion before storing them in GPU memory).

We will describe NaNet architecture and its latency/bandwidth characterization for all supported links and present NaNet usage in the NA62 and KM3 experiments.

The NA62 experiment at CERN aims at measuring the branching ratio of the ultra-rare charged kaon decay into a pion and a neutrino/antineutrino pair.

The ~ 10 MHz rate of particles reaching the detectors must be reduced by the multilevel trigger down to a \sim kHz rate, manageable by the data storage system. First level (L0) is implemented in dedicated hardware performing rough selections on their output reducing ~ 10 times the data stream rate to match the ≤ 1 MHz event target rate within 1ms time budget.

A GPU-based L0 trigger for the RICH detector using NaNet is being integrated in a parasitic mode in the experimental setup; this will allow assessing the real-time features of the system, leveraging on GPU relevant computing power to implement more selective trigger algorithms.

The KM3 experiment aims at detecting high energy neutrinos through an underwater Cherenkov telescope with a volume of the order of 1 km³.

In this context, NaNet is charged with two main tasks: first, global clock and synchronization signals delivery to the off-shore electronic system; second, reception of underwater devices data through optical cables. Fundamental requirement for the experiment is having a known and deterministic latency between on- and off-shore devices.

Results of NaNet performances in both experiments will be reported and discussed.

Primary authors: LONARDO, Alessandro (Universita e INFN, Roma I (IT)); BIAGIONI, Andrea (INFN); Dr ROSSETTI, Davide (NVIDIA Corp); Dr PASTORELLI, Elena (INFN sezione di Roma); LO CICERO, Francesca (INFN sezione di Roma); Dr SIMULA, Francesco (INFN sezione di Roma); TOSORATTO, Laura (INFN); Dr PONTISSO, Luca (INFN); Dr MARTINELLI, Michele (INFN sezione di Roma); Dr FREZZA, Ottorino (INFN sezione di Roma); Dr PAOLUCCI, Pier Stanislao (INFN Sezione di Roma); VICINI, Piero (INFN Rome Section); AMMENDOLA, Roberto (INFN)

Co-authors: COTTA RAMUSINO, Angelo (Universita di Ferrara (IT)); Dr AMELI, Fabrizio (INFN Sezione di Roma); SIMEONE, Francesco (INFN); LAMANNA, Gianluca (Sezione di Pisa (IT)); Dr NERI, Ilaria (Universita di Ferrara); SOZZI, Marco (Sezione di Pisa (IT)); FIORINI, Massimiliano (Universita di Ferrara (IT))

Presenter: LONARDO, Alessandro (Universita e INFN, Roma I (IT))

Session Classification: Second Poster Session

Track Classification: Trigger