



Contribution ID: 200

Type: **Poster**

## **NaNet: a flexible and configurable low-latency NIC for real-time trigger systems based on GPUs.**

*Tuesday, 24 September 2013 18:28 (1 minute)*

The adoption of GPUs in the low level trigger systems is currently being investigated in several HEP experiments.

While GPUs show a deterministic behaviour in performing computational tasks, data communication is the main source of fluctuations in the response time of such systems.

We designed NaNet, a FPGA-based NIC supporting 1/10GbE links and the custom 34 Gbps APElink channel. The design has GPUDirect RDMA capabilities, i.e. is able to inject the input data stream directly into the Fermi/Kepler class GPU(s) memory, and features a network stack protocol offloading engine.

We will provide a detailed description of the NaNet hardware modular architecture and a comparative performance analysis on the NA62 RICH detector GPU-based L0 trigger case study using the NaNet board and a commodity GbE NIC.

Figures of merit for the system when using the APElink and 10GbE links will also be provided.

### **Summary**

The adoption of GPUs in the low level trigger systems is currently being investigated in several HEP experiments.

In this context the main issue to be taken into account is the strict real-time requisite typical of such systems.

While GPUs show a deterministic behaviour in performing computational tasks once data are available to be processed in their own internal memories, when the low level trigger system is considered as a whole, it becomes soon evident that input data transfer from the detector readout system is the main source of fluctuations in the trigger response time.

Our approach to this problem is twofold.

First, we designed a NIC capable of injecting readout data directly from the links into the memories of Nvidia Fermi and Kepler class GPUs without any intermediate buffering or CPU operation, reducing data transfer latency and its fluctuations (GPUDirect RDMA being the commercial name for this feature).

Second, we offloaded the CPU from the network stack protocol management, implementing a dedicated engine in the NIC, to further reduce latency and avoid possible OS jitter effects.

We implemented these two features in the NaNet FPGA-based NIC: the first was inherited from the development of the APENet+ 3D HPC dedicated NIC, the second was realized adapting and integrating an open core developed by the FPGA vendor.

NaNet is flexible, supporting three different link technologies, namely GbE (1000BASE-T and 1000BASE-X), 10 GbE (10BASE-X) and the custom APElink channel implemented with 4 bonded LVDS lanes over QSFP+ cables and capable of 34 Gbps raw data bandwidth.

Beside this, being an FPGA based design, NaNet logic can be effectively tailored to different usage scenarios by adding dedicated custom logic blocks, e.g. performing compression or reshuffling on the data stream.

NaNet is currently being used in a pilot project within the CERN NA62 experiment aiming at the investigation of GPUs usage in the central Level 0 trigger processor.

We will provide a detailed description of the NaNet hardware modular architecture and a comparative performance analysis on the NA62 RICH detector GPU-based Level 0 trigger case study using the NaNet board and

a commodity GbE NIC.

Figures of merit for the system when using the APElink and 10 GbE links will also be provided, along with an outline of future project developments.

**Primary authors:** LONARDO, Alessandro (Universita e INFN, Roma I (IT)); BIAGIONI, Andrea (Universita e INFN, Roma I (IT)); ROSSETTI, Davide (U); Dr LO CICERO, Francesca (INFN); Dr SIMULA, Francesco (INFN); Dr TOSORATTO, Laura (INFN); Dr FREZZA, Ottorino (INFN); Dr PAOLUCCI, Pier Stanislao (INFN); VICINI, Piero (INFN Rome Section)

**Co-authors:** PANTALEO, Felice (Sezione di Pisa (IT)); LAMANNA, Gianluca (Sezione di Pisa (IT)); SOZZI, Marco (Sezione di Pisa (IT))

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

**Session Classification:** Poster