Contribution ID: **537** Type: **Oral**

NaNet: a Configurable Network Interface Card for Trigger and DAQ Systems

Wednesday, 12 October 2016 12:45 (15 minutes)

In order to face the LHC luminosity increase planned for the next years, new high-throughput network mechanisms interfacing the detectors readout to the software trigger computing nodes are being developed in several CERN experiments.
Adopting many-core computing architectures such as Graphics Processing Units (GPUs) or the Many Integrated Core (MIC) would allow to reduce drastically the size of the server farm, or to develop completely new algorithms with improved trigger selectivity. NaNet project goal is the design and implementation of PCI Express (PCIe) Network Interface Cards (NICs) featuring low-latency, real-time data transport towards CPUs and nVIDIA GPU accelerators.
Being an FPGA-based NIC, NaNet natively supports a number of different link
technologies allowing for its straightforward integration in diverse experimental setups. One of the key features of the design is the capability of managing the network protocol stack in hardware, thus avoiding OS jitter effects and guaranteeing a deterministic behaviour of the communication latency.
Furthermore, NaNet integrates 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 the host node. On a per experiment basis different solutions can be implemented, e.g. data compression/decompression and reformatting or merging of event fragments. NaNet accomplishes zero-copy networking by means of a hardware implemented memory copy engine that follows the RDMA paradigm for both CPU and GPU, supporting the nVIDIA GPUDirect RDMA protocol. The RDMA engine is assisted by a proprietary Translation Look-aside Buffer based on Content Addressable Memory performing virtual-to-physical memory address translations. Finally, thanks to its PCIe interface NaNet can be configured either as Gen2 or Gen3 x8 PCIe endpoint.
On the software side, a Linux kernel device driver offers its services to an application level library, which provides the user with a series of functions to: open/close the device; register and de-register circular lists of receiving buffers (CLOPs) in CPU and/or GPU memory; manage software events generated when a receiving CLOP buffer is full (or when a configurable timeout is reached) and received data
are ready to be consumed. A configuration of the NaNet design, featuring four 10GbE channels for the I/O and a PCIe x8 Gen3 host interface, has successfully been integrated in the CERN NA62 experiment to interface the readout of the RICH detector to a GPU accelerated server performing multi-ring pattern reconstruction. Results will be then sent to the central L0 processor, where the trigger decision is made taking into account information from other detectors, within the overall time budget of 1 ms. We will describe two multi-rings pattern recognition algorithms we developed specifically to exploit the many-core parallelism of GPUs and discuss the results we obtained during the NA62 2016 data taking.

Tertiary Keyword (Optional)

Parallelization

Secondary Keyword (Optional)

Trigger

Primary Keyword (Mandatory)

Network systems and solutions

Primary authors: LONARDO, Alessandro (Universita e INFN, Roma I (IT)); BIAGIONI, Andrea (Universita e INFN, Roma I (IT)); Mrs LO CICERO, Francesca (INFN Roma); LAMANNA, Gianluca (Istituto Nazionale Fisica

Nucleare Frascati (IT)); PONTISSO, Luca (Universita di Pisa & INFN (IT)); MARTINELLI, Michele (INFN); Mr FREZZA, Ottorino (INFN Roma); Mr CRETARO, Paolo (INFN Roma); VICINI, Piero (Universita e INFN, Roma I (IT)); AMMENDOLA, Roberto (Universita e INFN Roma Tor Vergata (IT))

Co-authors: ROSSETTI, Davide (Universita e INFN, Roma I (IT)); Mrs PASTORELLI, Elena (INFN Roma); NERI, Ilaria (Universita di Ferrara & INFN (IT)); SOZZI, Marco (Universita di Pisa & INFN (IT)); FIORINI, Massimiliano (Universita di Ferrara & INFN (IT)); Mr PAOLUCCI, Pier Stanislao (INFN Roma); FANTECHI, Riccardo (Universita di Pisa & INFN (IT)); PIANDANI, Roberto (Universita di Pisa & INFN (IT)); Mr FRANCESCO, Simula (INFN Roma); DI LORENZO, Stefano (Universita di Pisa & INFN (IT))

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

Session Classification: Track 1: Online Computing

Track Classification: Track 1: Online Computing