



Contribution ID: 100

Type: not specified

Front-End Rdma Over Converged Ethernet, real-time firmware simulation

Monday, 4 September 2023 16:30 (20 minutes)

Several physics experiments are moving (or are evaluating the possibility to move) towards new acquisition models. The tendency is to leave the hardware trigger system in favour of a complete or partial acquisition of the front-end data paired with a powerful online software event discrimination. Hardware trigger systems usually have to deal with a tight latency budget due to the narrow readout buffering. To reduce the selection inefficiencies resulting from the adoption of not optimal trigger algorithms due to the limited time budget and online computing resources, the main trigger schema is going to be revisited. The traditional first trigger level is going to be replaced by a hardware pre-processing of the data stream followed by a software online selection [1,2,3].

In a DAQ system a large fraction of CPU resources is engaged in networking rather than in data processing. The common network stacks that take care of network traffic usually manipulate data through several copies performing expensive operations. Thus, when the CPU is asked to handle networking, the main drawbacks are throughput reduction and latency increase due to the overhead added to the data transmission process. Networking with zero-copy can be achieved by adding a RDMA layer to the network stack and making dedicated hardware take care of the burden of the stack handling.

The main goal of the RDMA implementation in the detector front-end electronics is to move up the adoption of clever networking protocols to the data producer. Therefore, it is the front-end electronics that could take care of initiating the RDMA transfer towards the computing farm. In such a way it is possible to eliminate the point-to-point connection between the front-end and the back-end leaving the freedom of switching dynamically the routing to the computing nodes according to their processing availability. By appropriately choosing the network protocol for RDMA it is also possible to obtain a two-fold benefit. The possibility of adopting commodity hardware makes the DAQ system reduce reliance on custom hardware and it exploits all the advantages of a mature technology. In this way, the DAQ system gains in scalability and easiness of maintainability.

RoCE is the natural choice as it is the only industry-standard Ethernet-based RDMA solution with a multi-vendor ecosystem. In this work the main firmware block needed for the realisation of the RoCE endpoint has been implemented and verified. A real-time firmware simulation of the RoCE network stack has been developed where real network packets are exchanged between free-running Systemverilog code and the host machine via a TUN/TAP device which emulates a connection with a physical device (FPGA). The second part is devoted to show the verification process of the modified RoCE stack using the tools developed so far such as the novel simulation framework. The lightweight RoCE will be a stripped down version of the already verified firmware allowing the deployment on FPGAs with a low resource pool possible target devices could be rad-hard FPGAs used in front-end detector boards.

[1] LHCb Collaboration, LHCb Trigger and Online Upgrade Technical Design Report, CERN-LHCC-2014-016, LHCb-TDR-016

[2] CMS Collaboration, The Phase-2 Upgrade of the CMS Data Acquisition and High Level Trigger Technical Design Report, CERN-LHCC-2021-007, CMS-TDR-022

[3] Ryu, FELIX: The new detector readout system for the ATLAS experiment, JINST Volume 898, 2017

Primary author: BORTOLATO, Gabriele (Universita e INFN, Padova (IT))

Co-authors: BERGNOLI, Antonio (INFN sezione Padova); MENGONI, Daniele (Universita e INFN, Padova); MONTECASSIANO, Fabio (INFN sezione Padova); PAZZINI, Jacopo (Padova University, INFN); BELLATO, Marco (INFN)

sezione Padova); ZANETTI, Marco (Universita e INFN, Padova (IT)); MIGLIORINI, Matteo (Universita e INFN, Padova (IT)); VENTURA, Sandro (INFN sezione Padova); TRIOSI, andrea (University of Padova)

Presenter: BORTOLATO, Gabriele (Universita e INFN, Padova (IT))

Session Classification: A3