



Contribution ID: 12

Type: Oral

Integration of FPGA RDMA into the ATLAS Readout with FELIX in High Luminosity LHC

Wednesday, 21 September 2022 09:40 (20 minutes)

The FELIX system is used to interface the front-end electronics and the commodity hardware in the server farm. FELIX is using RDMA through RoCE to transmit data from its host servers to the Software Readout Driver using off-the-shelf networking equipment. In the current version of FELIX, RDMA communication is implemented using software on both ends of the links. Improvements of the data throughput as part of the High Luminosity LHC upgrade, by implementing RDMA support in the front-end FELIX FPGA, have been tested. Now, a version of FELIX that uses the FPGA implementation of RDMA is being proposed and demonstrated.

Summary (500 words)

The FELIX (Front-End Link eXchange) system is used for interfacing the front-end detector electronics to the readout system and the high-level trigger farm. The system is based on a custom FPGA board which receives data from the front-end detector electronics via optical links and outputs data via a PCIe interface to a host computer which manages processing and relaying the data further to the readout system. The host computer uses the RDMA (Remote Direct Memory Access) support offered by network interface cards with RoCE (RDMA over Converged Ethernet) support to transmit data further towards the readout systems over an Ethernet network. A possible improvement is the implementation of RDMA support in the FELIX FPGA itself. This would mean not using the PCIe interface and the host computer anymore in the data path from front-end detector electronics to the readout system.

The proposed FPGA implementation of RDMA has already been developed, tested and presented at TWEPP 2021. The results of the performance testing showed that the FPGA implementation can reach the expected performance of an RDMA link.

The next step is integrating the FPGA implementation of the RDMA protocol in the FELIX system, thus offering a working alternative for the data path to the one currently implemented in FELIX. This means implementing changes to three components of the FELIX system. First, the FELIX FPGA firmware, where the FPGA RDMA implementation needs to be paired with a elink subscription management controller. Second, the modifications to netio-next that have already been implemented needs to be integrated in felix-star. Moreover, felix-star itself needs to be adjusted accordingly to work with the modified netio-next. And third, a custom interface for Software Readout Driver (Software ROD) needs to be implemented, so that the Software ROD can receive data via the new data path. The main challenge in this is the difficulty of integrating the newly developed components into an existing large and complex system like FELIX.

In the process of developing this demonstrator, the Xilinx Alveo platform was also used. Xilinx Alveo boards are FPGA accelerator platforms with PCIe interfaces and multiple 100Gbps network interfaces. More precisely, Alveo U50 and Alveo U250 boards have been used, alongside the Xilinx Runtime Library (XRT), in order to simplify the implementation of the communication between the host PC the FPGA board, and the hardware design running in the FPGA.

Primary authors: VASILE, Matei (IFIN-HH (RO)); ILIC, Nikolina (University of Toronto (CA))

Presenter: VASILE, Matei (IFIN-HH (RO))

Session Classification: Programmable Logic, Design Tools and Methods

Track Classification: Programmable Logic, Design Tools and Methods