

RDMA optimizations on top of 100 Gbps Ethernet for the upgraded data acquisition system of LHCb

Tuesday, May 23, 2017 2:54 PM (18 minutes)

The LHCb experiment will be upgraded in 2018-2019 to change its operation to a triggerless full-software readout scheme from Run3. This results in increasing the load of the event building and filtering farm by a factor of 40. The farm will need to be able to handle all the 40 MHz rate of the particle collisions. The network of the data acquisition system is facing with a target speed of 40 Tb/s, aggregated by 500 nodes. It requires the links to be capable of delivering the data with at least 100 Gbps speeds per direction.

Three solutions are being evaluated: Intel® Omni-Path Architecture, 100G Ethernet and EDR InfiniBand. Intel® OPA and EDR IB runs by Remote Direct Memory Access. Ethernet uses TCP/IP or UDP/IP by default, which involves significant CPU load. However, there are solutions to implement RDMA-enabled data transfer via Ethernet as well. These technologies are called RoCE (RDMA over Converged Ethernet) and iWARP. We present first measurements with such technologies on 100 Gbps equipment in respect of the data acquisition use-case.

Primary author: Mr VONEKI, Balazs (CERN)

Co-authors: Dr NEUFELD, Niko (CERN); Mr VALAT, Sebastien (CERN)

Presenter: Mr VONEKI, Balazs (CERN)

Session Classification: R3-Trigger and data acquisition systems(3)

Track Classification: Trigger and data acquisition systems