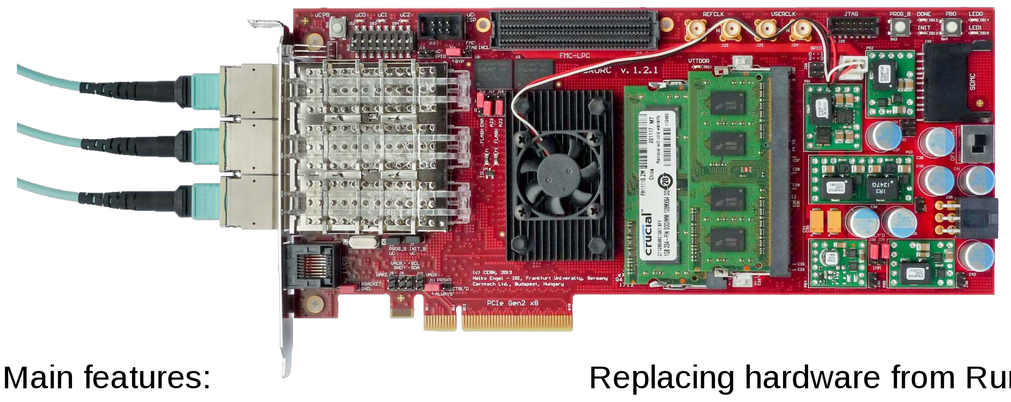
The ALICE HLT Readout Upgrade for Run2

Heiko Engel and Udo Kebschull for the ALICE Collaboration

Infrastructure and Computer Systems for Data Processing (IRI), Goethe-University Frankfurt

C-RORC Overview

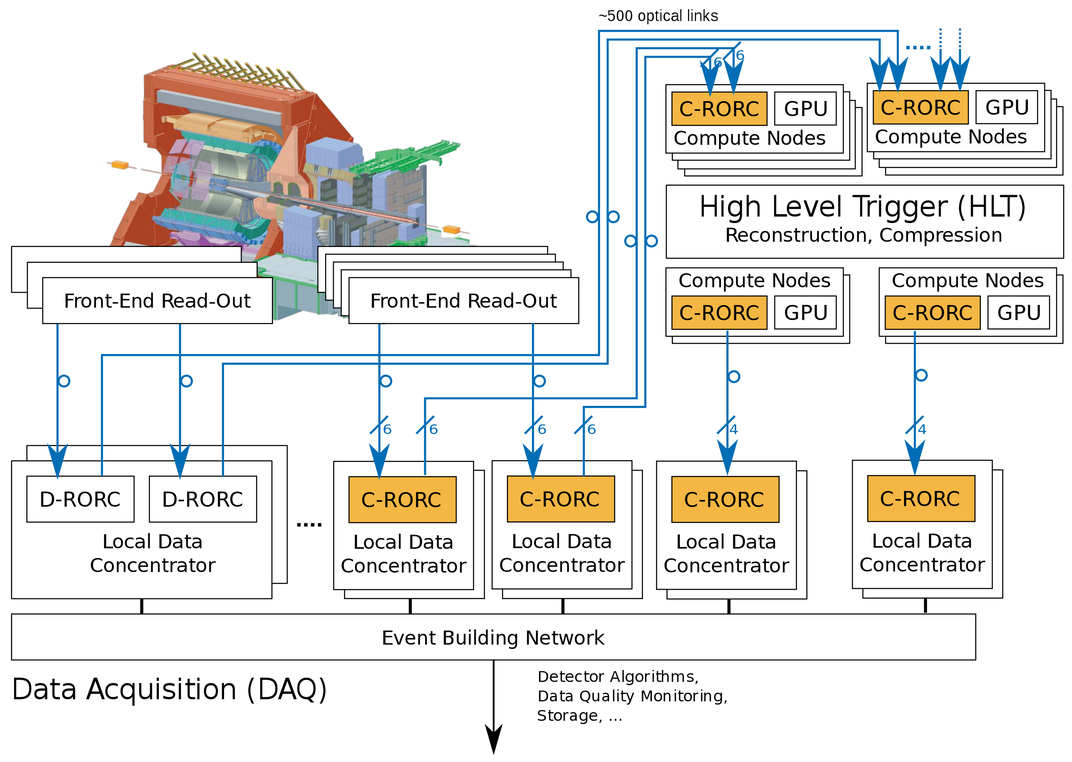
ALICE DAQ & HLT Common Read-Out Receiver Card (C-RORC)
- joint production of 370 boards together with ATLAS
- installation at CERN in Q3/2014
- used in production systems since start of Run2



Main features:
- Xilinx Virtex-6 FPGA
- PCIe x8 Gen2 interface
- 12 optical links up to 6.6 Gbps
- 2x DDR3 SO-DIMM up to 2x 8GB

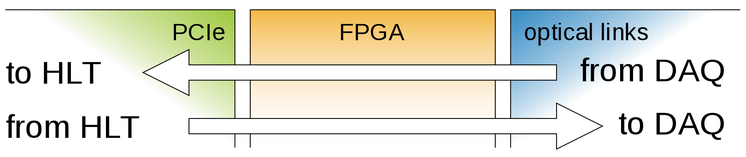
Replacing hardware from Run1:
- common hardware platform
- increased link density
- up-to-date interfaces
- extended processing capabilities

ALICE Readout Architecture

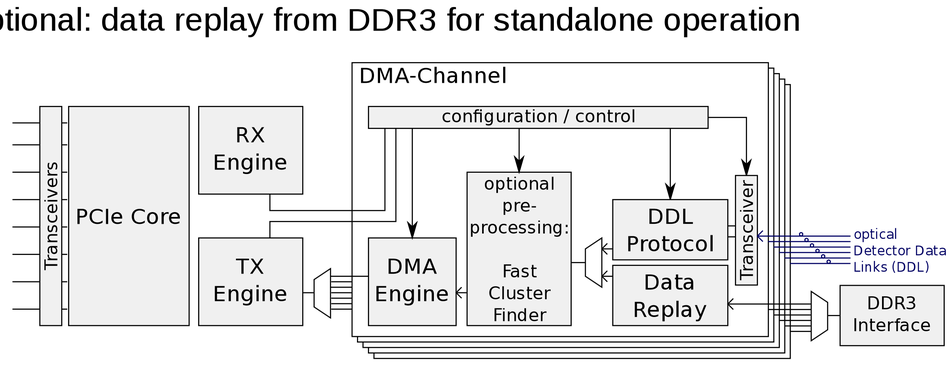


ALICE HLT C-RORC Firmware

- Two main firmware schemes:

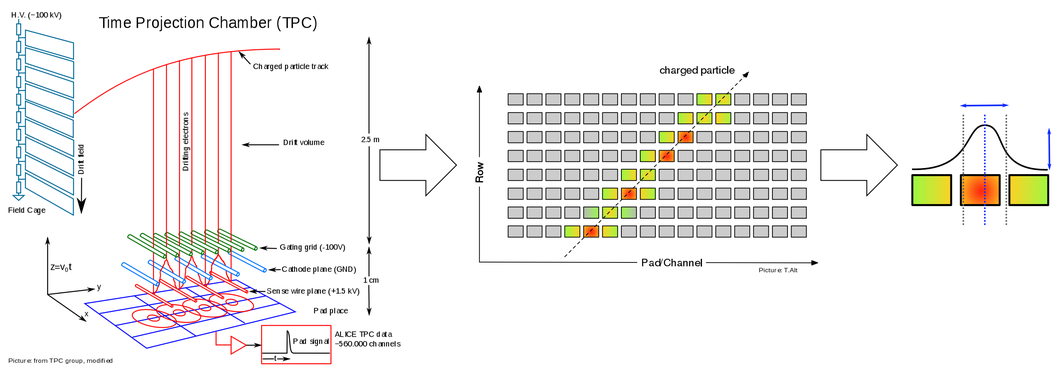


- optical interface: Detector Data Link (DDL)
  at several different link rates, selectable at runtime:
  2.125 Gbps / 4.0 Gbps / 4.25 Gbps / 5.3125 Gbps
- up to 12 independent DMA channels
- optional: data replay from DDR3 for standalone operation

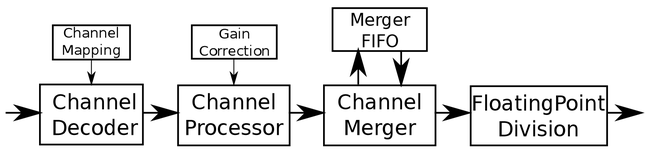


- custom DMA engine with scatter-gather-list support
- optional data preprocessing core: FastClusterFinder
- 4 different firmware images deployed on HLT cluster

Hardware Cluster Finding



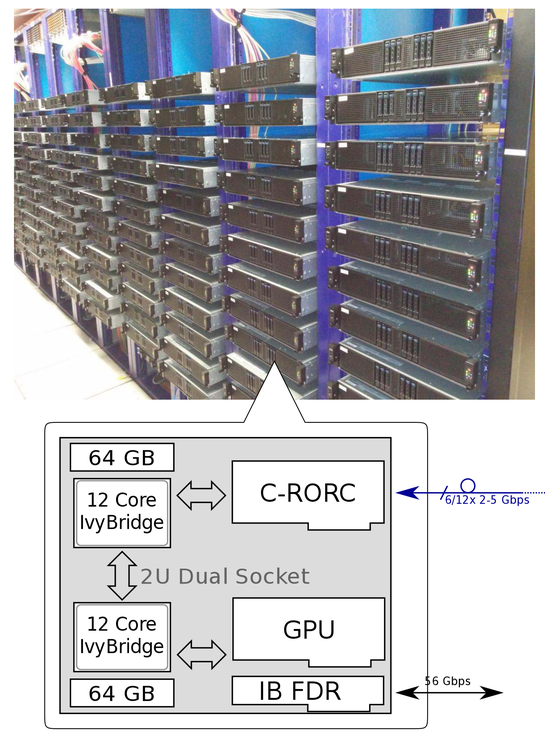
FPGA hardware cluster finder for Time Projection Chamber (TPC) data:
- find clusters in pad and time direction
- merge neighboring / separate overlapping clusters
- significantly faster than software cluster finding
- designed for full link bandwidth, only minor additional latency
- already essential part of Run1 HLT firmware
- improvements for Run2:
  - double throughput to handle DDL2 link speeds
  - support for upgraded TPC readout control units (RCU2)



ALICE High-Level Trigger

High-Level Trigger (HLT) primary functions:
- data compression
- online event reconstruction
- online calibration

- 180 identical worker nodes:
  - 2x 12 core CPU, 128 GB RAM
  - 2x SSD in RAID1
  - each with AMD FirePro S9000 GPU
  - Infiniband FDR 56 Gbps interconnect
  - Gigabit management network
- 74 nodes equipped with a C-RORC



C-RORCs integrated into HLT Publisher-Subscriber data transport framework:
- user space device driver for DMA and C-RORC access
- data source components receive events from C-RORCs
- event merging and distribution to processing groups across the cluster
- online AliRoot components for data processing
- reconstruction: FPGA cluster finding & GPU tracking
- compression: data format optimizations and Huffman encoding
- data sink components push processed event data into C-RORC for transmission to DAQ
- continuous link status monitoring via firmware error counters

Firmware Management & Monitoring

Node installation and automization

FOREMAN

- devide nodes into hostgroups
- define per hostgroup:
  - firmware type
  - firmware revision
  - firmware date

Software installation, state enforcement

puppet labs

- periodically check current firmware
- if necessary:
  - flash new firmware
  - reconfigure FPGA
  - re-initialize PCIe bus

Cluster monitoring

ZABBIX

- monitor FPGA temperature
- monitor PCIe interface status

SPONSORED BY THE Federal Ministry of Education and Research
05P15RFCA1

Contact: hengel@cern.ch