Commissioning and First Experiences of the ALICE
High Level Trigger

Timm M. Steinbeck for the ALICE HLT collaboration
Kirchhoff Institute of Physics, University Heidelberg, im Neuenheimer Feld 227, D-69120
Heidelberg, Germany
E-mail: timm.steinbeck@kip.uni-heidelberg.de

Abstract. For the ALICE heavy-ion experiment a large computing cluster will be used to
perform the last triggering stages in the High Level Trigger (HLT). For the first year of operation
the cluster consisted of about 100 multi-processing nodes with 4 or 8 CPU cores each, to be
increased to more than 1000 nodes for the coming years of operation. During the commissioning
phases of the detector, the preparations for first LHC beam, as well as during the periods of first
LHC beam, the HLT has been used extensively already to reconstruct, compress, and display
data from the different detectors. For example the HLT has been used to compress Silicon
Drift Detector (SDD) data by a factor of 15, lossless, on the fly at a rate of more than 800 Hz.
For ALICE’s Time Projection Chamber (TPC) detector the HLT has been used to reconstruct
tracks online and show the reconstructed tracks in an online event display. The event display
can also display online reconstructed data from the Dimuon and Photon Spectrometer (PHOS)
detectors. For the latter detector a first selection mechanism has also been put into place
to select only events for forwarding to the online display in which data has passed through
the PHOS detector. In this contribution we will present experiences and results from these
commissioning phases.

1. Introduction
The ALICE High Level Trigger (HLT) [1] is the highest trigger in the ALICE trigger hierarchy,
processing events before they are passed to permanent storage. It has been designed with three
primary goals in mind:

Trigger Mark events for storage on the basis of a full online reconstruction and analysis.
Select Select regions-of-interest by physics characteristics within events for an even more
efficient selective readout.
Compress Online compression of marked and selected event data without loss of physics
information to make better use of the available storage capacity.

In order to achieve these goals, the HLT consists of a large PC farm. Currently about 100
PCs are available, but up to about 1000 nodes are possible for later stages. On these nodes a
full online analysis will be performed on an event-by-event basis. The results of this analysis
will be made available in the form of Event-Summary-Data (ESD).
2. HLT Architecture

ALICE data from the experiment is sent from the Front-End-Electronics to the data acquisition (DAQ) via optical links, the Detector-Data-Links (DDLs). The DAQ duplicates the data by optical splitters and a copy is then sent by further DDLs to the HLT. In the HLT the links are terminated in special-purpose PCI-X HLT Read-Out-Receiver-Cards (H-RORCs). The H-RORCs write the received data directly in the main memory of their respective Front-End-Processor (FEP) host computer. From there a dedicated software package [2], [3] running on the cluster takes control of the data. This framework is responsible for transporting and aggregating the data inside the cluster. For this purpose the software consists of independent components, coupled by efficient communication interfaces, which can be plugged together in varying configurations depending on the requirements presented by the detector hardware as well as the reconstruction and analysis tasks desired.

For the online reconstruction the processing components are arranged in multiple hierarchy levels, typically corresponding to the hardware layout of the detector. In each step the data size is reduced, so that less data needs to be transported to the next component in the processing pipeline. Each step also aggregates and merges data from multiple components in the prior stage until at the final stage the fully reconstructed event is available in ESDs. At several stages of the processing hardware co-processors can be used to accelerate the processing compared to pure software processing on the CPUs. This can even be done before the data initially enters the FEPs in the H-RORCs. These cards are equipped with FPGAs that can execute the first level of reconstruction, e.g. cluster finding for the TPC detector, already before the data is written to the host FEP’s main memory.

3. Hardware Status

The current status of the HLT cluster hardware is as used during the runs in 2008. 87 Front-End-PCs are installed with a total of 384 CPU cores and 696 GB RAM, the nodes being dual Opteron nodes with either dual or quad-core CPUs. These FEPs are connected to detectors via 384 DDL links terminating in 174 H-RORCs in the nodes. Each node currently hosts two H-RORCs with most of the H-RORCs receiving data via two DDLs. On these FEPs the first level of reconstruction will be executed, if it is not done on the H-RORCs already. In addition to the FEPs 16 computing PCs are installed which currently share all subsequent processing steps between them. These nodes together contain 128 CPU cores and 256 GB of RAM, in this case all of them having quad-core Intel Xeon CPUs.

The Gigabit Ethernet network connecting all nodes is already installed with the full network infrastructure necessary for the planned future upgrades of the system during later project phases. This also applies to the services infrastructure. Both are installed and ready for heavy-ion running, which will have very high CPU processing requirements.

4. Software Status

Not all of ALICE’s detectors are currently being used in the HLT. Up until the LHC shutdown in September the following detectors were in use with the HLT: Dimuon, FMD, PHOS, SDD, TPC, and TRD. For most of these both first (e.g. cluster finding for TPC) and second level reconstruction (e.g. track finding for TPC) was tested and used successfully. Fig. 1 shows an event in the TPC which has been fully reconstructed, including cluster finding and tracking, by the HLT reconstruction. For SDD no reconstruction was done as only a lossless reformatting and compression of the detector’s raw data was foreseen. This online data compression was also tested and found to work at the required processing rates.

A sophisticated infrastructure for integrating trigger algorithms working on the fully reconstructed data has also been put in place. This allows multiple trigger algorithms to be easily inserted into an HLT configuration and the results from the different trigger algorithms
to be combined into a coherent global HLT decision. This decision does not have to be a simple accept/reject for the whole event, but can also include region of interest information to select only parts of an event for final storage by the DAQ. Trigger algorithms are placed in standard software components like for the reconstruction, only the data type for the output is different and the base class used for the trigger implementation is more specialized.

4.1. Processing Rates
Reconstruction of event data has been benchmarked for some of the detectors contributing to the HLT’s input data. An initial benchmark was done of the data transport framework to determine the limits this imposes. Here the limit in an example configuration was measured at 8 kHz. As the framework was designed and optimized to transport large data blocks efficiently with low CPU overhead at comparatively low rates due to ALICE’s heavy ion focus, this value is acceptable. As can be seen, with this rate the reconstruction rates are still primarily limited by the data processing itself and not significantly by the transport of the data. For PHOS, Dimuon, and TPC first and second level reconstruction has been performed on simulated pp-data. In these tests the following event and data rates, respectively, have been achieved:

\textbf{Dimuon}: 1.2 kHz / 160 MB/s  
\textbf{PHOS}: 2.1 kHz / 80 MB/s  
\textbf{TPC}: 600 Hz / 70 MB/s

4.2. Calibration
The ALICE HLT is also able to produce calibration data for use in event reconstruction. Here two types of calibration data can be distinguished. The first of these is produced by running detector calibration algorithms inside the HLT. The output of this is offline calibration data which can be forwarded to the ALICE Offline Conditions Database (OCDB). From there it can be used in the following offline reconstruction of events for which this calibration is applicable. The second type of calibration is an online calibration in which the HLT calibrates itself. In this
mode online calibration algorithms are run on different kinds of input data. These types of input data can include normal physics events, special intermediate calibration events, e.g. laser events in the ALICE TPC, or even live values retrieved from the ALICE Detector Control System (DCS). The results obtained by analysing this calibration input data can then be directly used by the online reconstruction components running inside the HLT processing chain.

4.3. HLT Interfaces

![Figure 2. Interfaces of the HLT to other ALICE online systems.](image)

The HLT has interfaces to all other online systems in ALICE as illustrated in Fig. 2. From the DAQ the HLT receives the raw data from the detectors via optical splitters. HLT output comprised of internally produced data (e.g. online data compression) and HLT decisions in turn is forwarded from the HLT via optical links back to the DAQ for storage. DCS delivers monitored parameter values to the HLT for use in online calibration and the HLT can write back internally calculated values to DCS. From the Experiment Control System (ECS) commands are received and status is delivered for inclusion in the global ALICE run control system. Calibration data is exchanged with the ALICE Offline Systems OCDB. These interfaces are described in more detail in [4]. Finally raw as well as reconstructed data from the HLT can be used to feed online event displays dynamically. This allows for a quick and direct overview of the events in the HLT and the participating detectors. All these online systems have been tested and used successfully before and during the commissioning and LHC first beam phases in 2008.

5. HLT Running in 2008

The HLT 2008 can be divided into three distinct periods, commissioning, cosmic data taking, and LHC first beam.

5.1. Commissioning

During this period the HLT was commissioned within the ALICE online data stream in three major steps. The first of these steps was receiving data via optical links from the optical splitters in the DAQ. This had been used before but was tested and verified extensively again. During the second step creation and application of trigger decision was tested. This involved creating trigger decisions in the HLT, consisting of full event accepts, full rejects, as well as partial readout specification. Decisions created by the HLT were then sent by the defined protocol to the DAQ where they were evaluated and executed. This full sequence of tasks could also be verified to
work. In the last step reconstructed event data was sent to the DAQ permanent storage via the DAQ output and retrieved successfully again in the offline framework. Fig. 3 illustrates the full path for both event decisions and HLT produced data. This full path could be verified to work during the commissioning.

![Figure 3. Flow of data through the ALICE HLT.](image)

5.2. Cosmic Data Taking
During cosmic data taking the HLT was running with several detectors in parallel. This was achieved by using multiple partitions, each partition running independently from the others for one of the detectors or a specific combination of detectors. Detectors running with HLT in this period were Dimuon, PHOS, TPC, TRD, and FMD. For these detectors online reconstruction could be performed with the detector data from the cosmic runs. In addition to this the interface to ALICE’s online display AliEVE was used extensively to provide an online monitor of the HLT reconstructed detector data in the AliEVE system. This proved very useful in tracing and analysing problems with the data analysis, mainly caused by subtle differences in simulated and real detector data.

5.3. LHC first beam
During the first LHC beam periods HLT was used only in a limited manner for online reconstruction as most detectors running with HLT were not yet switched on during the initial phases of injection. HLT, however, was used successfully with SDD to perform a full online data reformatting and compression to reduce the size of the output data by a factor of 15, at a rate of more than 800 Hz. Included in this was also a full rejection of the original detector raw data via the HLT trigger decision to the DAQ so that the stored SDD data was processed fully in the HLT. In addition PHOS was running for some time during the LHC injections and a first level reconstruction of the received data was done and the reconstructed events were shown on an AliEVE based event display in the control room. This setup also involved a first trigger mechanism in the HLT to forward only events to the online display which contained a sufficient amount of data in order to suppress empty events on the event display.

6. Conclusion and Outlook
The HLT was used successfully during the commissioning and data taking periods, both with cosmics and LHC beam injections, in 2008. Already the system has shown to be a valuable and essential tool even during these first phases. For 2009 the focus now lies on commissioning a larger HLT installation consisting of more PCs. This is needed to be ready for the extensive periods of pp and hopefully heavy-ion running starting 2009 through 2010 with their expected higher computing power requirements. On the software side the first physics triggers have to be finalized and integrated into the HLT trigger framework to be ready for use with first cosmics data, then LHC beam injection test data, and finally data from collisions.
References

[1] The ALICE Collaboration, ALICE Technical Design Report of the Trigger, Data Acquisition, High-Level Trigger, and Control System CERN/LHCC 2003-062, ALICE TDR 10, ISBN 92-9083-217-7, 2004.

[2] T. M. Steinbeck, V. Lindenstruth, H. Tilsner, “New experiences with the ALICE High Level Trigger Data Transport Framework”, CHEP04, Interlaken, Switzerland, 2004.

[3] A Modular and Fault-Tolerant Data Transport Framework (Dissertation), Timm M. Steinbeck, arXiv CoRR cs.DC/0404014 and http://www.ub.uni-heidelberg.de/archiv/4575, 2004.

[4] Heterogeneous Distributed Calibration Framework for the High Level Trigger in ALICE (Dissertation), Sebastian Robert Bablok, Dissertation at the Institute for Physics and Technology, University of Bergen, Norway, 2009.