The ALICE High Level Trigger

M. Richter†*, T. Alt‡, H. Helstrup§, V. Lindenstruth‡, C. Loizides||, G. Øvrebekk†, D. Röhrich†, B. Skaali¶, T. Steinbeck‡, R. Stock||, H. Tilsner‡, K. Ullaland†, A. Vestbø†, T. Vik¶ and A. Wiebalck‡
for the ALICE collaboration

† Department of Physics and Technology, University of Bergen, Norway

‡ Kirchhoff Institute of Physics, University of Heidelberg, Germany

§ Faculty of Engineering, Bergen University College, Norway

|| Institute for Nuclear Physics, University of Frankfurt, Germany

¶ Department of Physics, University of Oslo, Norway

Abstract

The ALICE experiment at LHC will implement a High Level Trigger System for online event selection and/or data compression. The largest computing challenge is imposed by the TPC detector, requiring real-time pattern recognition. The main task is to reconstruct the tracks in the TPC, and in a final stage combine the tracking information from all detectors. Based on the physics observables selective readout is done by generation of a software trigger (High Level Trigger), capable of selecting interesting (sub)events from the input data stream. Depending on the physics program various processing options are currently being developed, including region of interest processing, rejecting events based on software trigger and data compression schemes.

The system entails a very large processing farm, designed for an anticipated input data stream of 25 GB/s. In this paper we present the architecture of the system and the current state of the tracking methods and data compression applications.

INTRODUCTION

The ALICE experiment described in [1] will investigate Pb-Pb collisions at a center of mass energy of about 5.5 TeV per nucleon pair and p-p collisions at 14 TeV. The detectors are optimized for charged particle multiplicities of up to $dN_{ch}/d\eta$ of 8000 in the central rapidity region.

The whole system has to handle event sizes of up to 100 MByte per event. Several subdetectors contribute to that data volume. Among them the Time Projection Chamber (TPC), which is the main central tracking detector, will produce data of up to 75 MB per event for central Pb-Pb collisions. The experiment is designed to run at 200 Hz event rate for central Pb-Pb collisions and up to 1000 Hz for p-p collisions. The overall event rate is limited by the foreseen bandwidth to permanent storage of 1.25 GB/s. With no further reduction the ALICE TPC can only accumulate central Pb-Pb events up to 20 Hz. Higher event rates are possible by either online event selection and/or data compression. Both applications require a real-time analysis of the detector information with a latency of the order of a few hundred ms. To accomplish the pattern recognition tasks at an incoming data rate of up to 25 GB/s, a massive parallel computing system, the High Level Trigger (HLT) system, is being designed [2].

* Matthias.Richter@ift.uib.no

Figure 1 shows the integration of the HLT into the data flow of the ALICE experiment. The raw data is transferred via optical fibers from the detector front-end to the DAQ system. The *ReadOut Receiver Cards* of the DAQ system (D-RORC) read the data into the local data concentrators and send an exact copy of the data to the HLT. The data stream is received by the HLT-RORCs, which are interfaced to the receiving nodes through the internal PCI-bus. In addition to different communication interfaces, the HLT-RORC provides an FPGA co-processor for the data intensive local tasks of the pattern recognition.

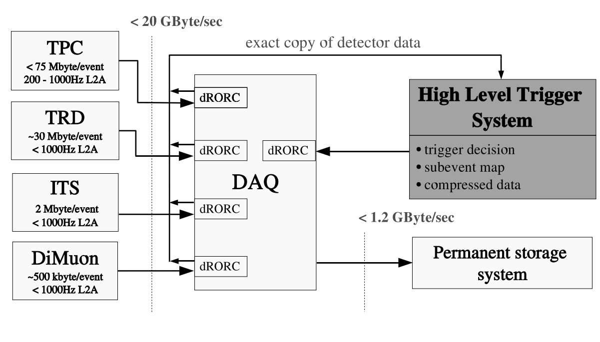

Figure 1: Integration of the HLT system into the Data-flow architecture of the ALICE experiment . The detector raw-data is duplicated by the DAQ and sent as exact copy to HLT system.

Internally the HLT adopts a tree like structure allowing parallel processing of the data stream as much as possible. The result from the processing on one layer (e.g. track segments on sector level) will be merged at a higher layer (sector merging and track fitting). Finally all local results are collected from the sub-detectors and combined on a global level where the complete event can be reconstructed and trigger decision can be issued.

The result of the HLT which can be a trigger decision, a subevent map or even compressed data is sent back to the DAQ which treats the HLT as a specific detector with some additional information. The information provided by the HLT will allow the DAQ to reduce the data rate below its limit.

The HLT system will run on a computing farm of about 500 dual processor nodes. The farm is designed to be completely fault tolerant avoiding all single points of failure, except for the unique detector links. A generic communication framework has been developed based on the publisher-subscriber principle, which allows an arbitrary connectivity metric of processing elements across the underlying network interface [3].

ONLINE EVENT RECONSTRUCTION

The main processing task is the reconstruction of the tracks in the TPC, and in a final stage combine the tracking information from all detectors. Given the uncertainties of the anticipated particle multiplicities, different approaches are being considered for the TPC track reconstruction.

The conventional approach of TPC track reconstruction consists of a Cluster Finder and a subsequent Track Follower. In a first step the Cluster Finder reconstructs the cluster centroids from the generated two-dimensional charge distributions in the TPC pad-row planes. Together with the position of the pad-row-planes the centroids are interpreted as three-dimensional space points along the particle trajectories, and serve as an input for the Track Follower which connects the space points into track segments. A final helix-fit of the track segments provides the track parameters and thus the kinematic properties of the particles. Such an approach has been implemented and evaluated on simulated ALICE TPC data. The algorithms were originally developed for the STAR L3 trigger [4] and consist of a straight-forward center-of-gravity calculation of cluster centroids, and a Track Follower which applies conformal mapping on the space points. The latter enables the circular tracks to be fitted by a linear parametrization, and thereby significantly reducing the computational requirements. The overall measured performance of the reconstruction chain represented by the tracking efficiency for an event sample of multiplicity density $dN_{ch}/d\eta$=4000 and the integrated tracking efficiency is shown in figures 2 and 3.

For higher multiplicities the observed tracking performance deteriorates significantly (figure 3). This is due to the increasing detector occupancy which give rise to a significant amount of overlapping clusters. In such a scenario the Cluster Finder fails to recontruct the cluster centroids due to its incapabilities of deconvoluting overlapping charge distributions. In this case information about the tracks is needed *prior* to reconstructing the cluster centroids in order to fit the individual distributions to a known shape. This can be done since the cluster shape to a good approximation can be described by the track parameters, and together with the knowlegde of the number of tracks contributing to a given cluster deconvolution can be done based on a two-dimensional Gauss-fit. Such an approach has been evaluated by applying an implementation of the Hough Transform on the raw ADC-data, and subsequently fitting the clusters to a two-dimensional Gauss-function based on the found track candidates. However, in the current implementation the fitting procedure is unstable because of a relative high number of *false* candidates produced by the Hough Transform. Various approaches are currently being investigated in order to overcome this problem.

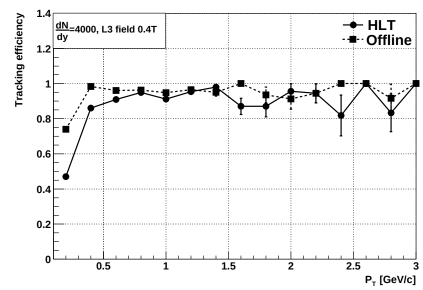

Figure 2: Performance of the HLT tracking algorithms compared to the Offline reconstruction chain. Tracking efficiency for a multiplicity of $dN_{ch}/d\eta$=4000.

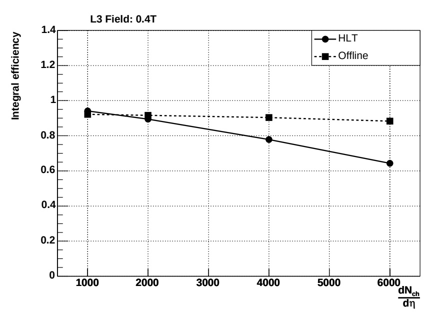

Figure 3: Integrated tracking efficiency as a function of multiplicity.

TPC DATA COMPRESSION

The option to compress the data online provides a method that can improve the physics capabilities of the experiment in terms of statistics, even without performing selective readout. Standard comression techniques such as entropy coding and Vector Quantization were studied on real NA49 and simulated ALICE TPC data and the results are presented in detail in [5]. The results show that compression ratios $\geq$ 50% are achievable. [1]

Even better data compression can however be achieved by using compression techniques which are highly adapted to the system itself. Such methods exploit the fact that the relevant information is contained in the reconstructed cluster centroids and the track charge depositions. These parameters can be stored as deviations from a model, and if the model is well adapted to the data the resulting bitrate needed to store the data will be small. Since the clusters in the TPC critically depend on the track parameters, the reconstructed tracks and clusters can be used to build such efficient data models.

This data compression technique requires the full reconstruction of the tracks and the event prior to the compression. Clusters can then be assigned to the reconstructed tracks and are described in terms of track parameters and a small deviation. In order to maintain a good tracking efficiency for reconstruction of the event from the compressed data, a selection of the remaining clusters must be kept. Figure 4 illustrates the process. Some of the remaining tracks do obviously not originate from the primary vertex and can be dismissed while other track segments might belong to the track of a particle coming from the vertex. An additional cluster analyzer has to be trained to distinguish between the two types of clusters.

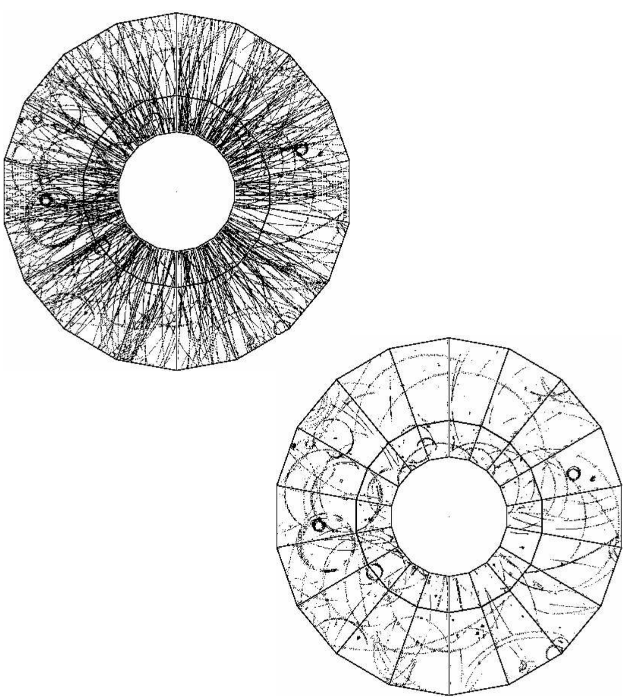

Figure 4: A TPC event display illustrating both the adopted compression technique and the power of the track reconstruction. Left above: before assigning clusters to reconstructed tracks and removing them from the display. Right below: after the removal.

[1]Compression ratio is calculated as *Compressed size*/*Original size* $\times$100 [%]

Such a data compression scheme has been evaluated on simulated ALICE TPC data, and indicates that compression ratios of 10-15% can be achieved. Figure 5 illustrates the impact on the tracking efficiency and relative transverse momentum resolution for an event sample of multiplicity density $dN_{ch}/d\eta$=1000. The overall loss in tracking efficiency is $\sim$1.5% and no significant impact on the momentum resolution is observed.

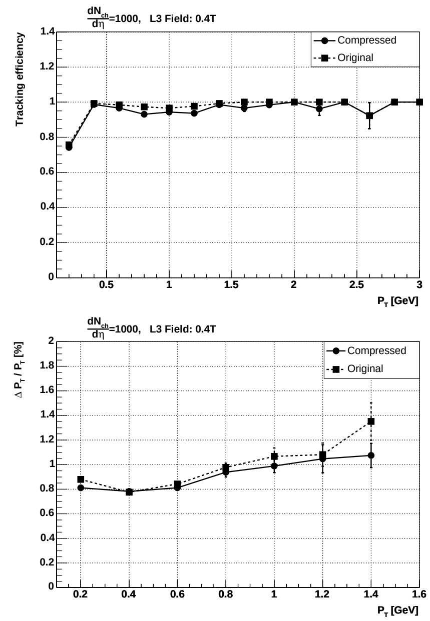

Figure 5: Impact of the TPC data compression scheme on the tracking efficiency and relative transverse momentum. The achieved compression ratio is 12%.

SUMMARY

Focusing on the TPC the ALICE HLT system is designed to increase the readout and storage of relevant physics events by a factor of 10. The current tracking performance shows that a sufficient event reconstruction within the central Pb-Pb event rate of 200 Hz will be achievable for multiplicity densities of $dN_{ch}/d\eta \leq$4000. For higher densities cluster deconvolution based on track parameters becomes necessary.

Efficient data compression schemes indicate that compression ratios of 10-15% are achievable with an insignificant impact on the Offline tracking efficiency.

REFERENCES

[1] ALICE Collaboration CERN/LHCC/1995-71.

[2] ALICE Collaboration CERN/LHCC/2003-062.

[3] Steinbeck T *et al* 2002 *IEEE Trans. Nucl. Sci.* **49** 455.

[4] Adler C *et al* 2003 *Nucl. Instrum. Methods* **A499** 778

[5] Berger J *et al* 2002 *Nucl. Instrum. Methods* **A489** 406