MUON DETECTOR CALIBRATION IN THE ATLAS EXPERIMENT: DATA EXTRACTION AND DISTRIBUTION

E. Pasqualucci, S. Falciano, A. De Salvo, A. Di Mattia, F. Marzano, L.Luminari, M.Curti,
INFN, Roma, Italy
D. Orestano, INFN, Roma3, Italy,
S. Kolos, B. Martin, C. Meirosu, CERN, Switzerland,
S. Gadomski, University of Bern, Switzerland

Abstract

In the Atlas experiment, fast calibration of the detector is vital to feed prompt data reconstruction with fresh calibration constants. We present the use case of the muon detector, where a high rate of muon tracks (small data size) is needed to accomplish MDT calibration requirements.

The ideal place to get data suitable for muon detector calibration is the second level trigger, where the pre-selection of data by the first level trigger allows to select all and only the hits from isolated muon tracks and to add useful information to seed the calibration procedures. The online data collection model for calibration data is designed to minimize the use of additional resources, without affecting the behavior of the trigger/DAQ system. Collected data are then streamed to remote Tier 2 farms dedicated to detector calibration. Measurements on the preseries of the ATLAS TDAQ infrastructure and preliminary tests on the wide area data distribution path are shown, proving the feasibility of the system.

MDT CALIBRATION

The precision chambers of the ATLAS Muon Spectrometer are built with the *Monitored Drift Tube (MDT)* technology. The requirement of high accuracy and low systematic error can only be accomplished if the calibrations are known with an accuracy of some mm. The relation between the drift path and the measured time (the so called *r-t relation*) depends on many parameters (T, hit rate, gas composition, thresholds,...) varying with time. It has to be measured from the data without the use of an external detector, using the *auto-calibration* technique. It relies on an iterative procedure applied to the same data sample, starting from a preliminary set of constants.

Calibration requirements

The required precision requires a large amount of non-parallel tracks crossing a region, called *calibration region*, i.e. the region of the MDT chamber sharing the same r-t relation. The evaluated number of required tracks is 10^4 to 10^5 per region.

A stringent requirement is related to the maximum latency allowed before starting data reconstruction after data taking (about 24 hours). This condition implies online selection of muon tracks with a minimum data acquisition rate. This rate can be calculated from simple parameters:

- The detector calibration granularity, i.e. the number of calibration regions.
- The data collection time.
- The number of needed tracks per region.

The number of calibration regions has to be a compromise between the need of keeping things simple and not requiring too high statistics (that would also imply a very large CPU processing time) and the need of reducing systematic errors by separately calibrating regions where parameters may take very different values. Since there are 1200 chambers, each of them made of two multi-layers, a good guess for calibration granularity can be obtained dividing each multi-layer into 4 wire segments, so giving 9600 calibration regions. Since each muon crosses 3 chambers, i.e. 6 regions (see Fig. 1), we have 1600 calibration towers. Taking into account the number of required tracks per region, we get a data rate between 300 Hz and 3 kHz.

Critical Issues

Extracting data online from the trigger/DAQ system implies additional requirements to the extraction system. Indeed, it must not affect regular data taking, neither adding any latency to the trigger, nor introducing error conditions in case of failure.

These requirements can be summarized as follows:

- The data collection scheme must fit in the actual trigger/DAQ architecture.
- The data fragments must contain only interesting data.
- The overhead of tasks providing calibration data collection must be negligible.
- The required bandwidth must be already available in the trigger/DAQ system.
- Some flexibility is desirable, in order to provide data streaming, data pre-selection and possibly seeding of calibration procedures.

DATA READOUT

In the ATLAS muon detector, MDT chambers are read out by on-chamber ADC and TDC; data are then concentrated into a CSM (Chamber Service Modules).

Each set of 6 (or 8) CSM, i.e. 6 or 8 chambers, is connected via optical links to a Read-Out Driver (ROD).

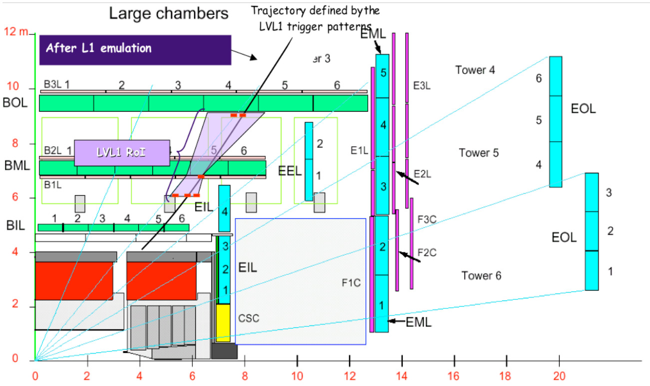

Figure 1: Data organization in the ATLAS muon spectrometer.

Fig. 1 shows the data organization for the muon spectrometer. The level-1 trigger has been designed to identify high momentum muon candidates. Starting from the hits in the trigger chambers, it defines a Region of Interest (RoI), defined as a *trigger tower* as shown in the figure plus the "nearest" one – i.e. the one closest to the level-1 hit pattern. The RoI is then used to input the 2nd level trigger with data from trigger and precision chambers to perform local data reconstruction. It is then natural to organize the detector readout into trigger towers.

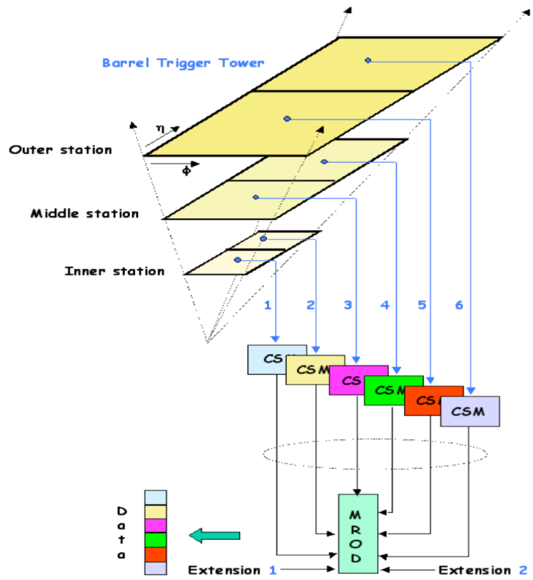

Figure 2: Data readout organization for the muon barrel

In figure 2, the readout organization is shown. Each MROD receives data from the chambers in the same trigger tower. Only data coming from a RoI (i.e. two MROD) are transferred to the 2nd level trigger, thus minimizing the data throughput from the DAQ system to the 2nd level trigger. The fragments of accepted events are then collected by ReadOut Systems (ROS) and sent to an event building farm node (Sub Farm Input, SFI). The full events are then analyzed by the event filter and sent to an SFO (Sub Farm Output) process to be staged on disk.

In principle, data can be extracted at any level, both using data sampling and monitoring facilities and creating ad-hoc systems. On the other hand, the required rate is not achievable after the 2nd level trigger, whose overall accept rate is of the order of 1 kHz. On the other hand, getting data at ROD level, where rate is high (75 to 100 kHz) but there is no knowledge of the RoI makes no sense, since it would imply complete event building at the required rate.

At level 2, the accepted muon rate is about 12 kHz and only data from the RoI are moved. In addition, the muon trigger algorithm deals only with data related to the candidate muon track. Since the algorithm reconstructs the track, the 2nd level trigger is the ideal place to extract muon data selecting all and only the hits in the track and adding the fit parameters in order to seed the calibration procedures to allow fast convergence. A first evaluation of the extracted data size is about 800 bytes per event.

Data extracted from the trigger can then be concentrated in a calibration server and made available to be distributed to remote calibration farms. Data concentration can happen in one or two steps, either directly sending data from the level-2 nodes to the calibration server or sending data to the local file/boot servers in the level-2 racks and, in a second step, to the calibration server. The former option allows not using local server resources, the latter improves the flexibility of the system and allows optimized use of the network.

SYSTEM ARCHITECURE

The general system architecture is shown in figure 3. The 2nd level trigger algorithms run in a farm of about 500 processors, divided in 20 racks. 25 nodes in a rack are booted from a local disk/file server. Each node runs 3 level 2 Processing Units.

In the two level option, data prepared in level 2 PU are sent to a collector in the local file server, grouped in multi-event packets and sent to a global collector, which writes them to disk. The throughput to each local server is about 480 kB/s and the global throughput is about 9.6 MB/s.

Collected data can then be sent to remote calibration farms (Tier 2 in the LHC computing model) for processing. In the proposed architecture, final data destination can be easily decided on the basis of the calibration region. We foresee to send barrel events to Italy, having Munich as a backup farm, using the Atlas Data Manager to transfer data through Tier 0 and Tier 1 to a Tier 2 farm, and endcap events to U.S. (Michigan) through the UltraLight network.

In order to fulfill the requirements, the latency added to the muon 2nd level trigger must be negligible with respect to the processing time (O(10 ms)), the load on local servers must be negligible and data distribution channels to the remote farms on the WAN must sustain the data rate.

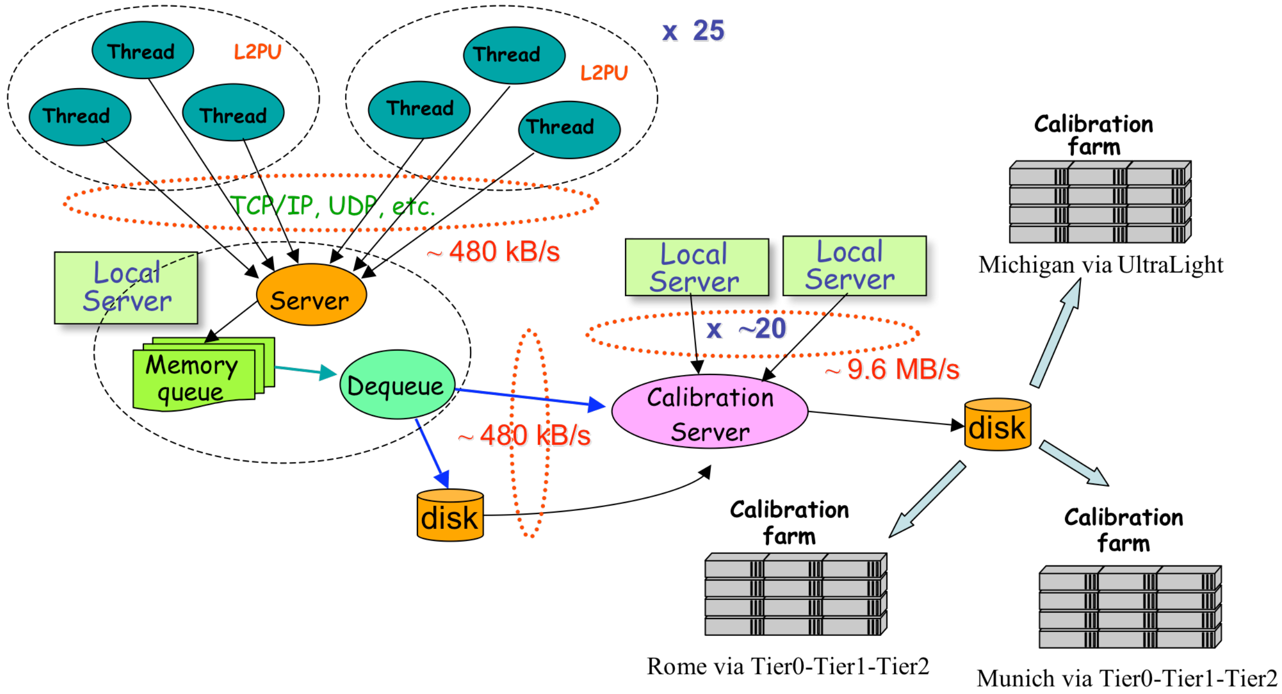

Figure 3: Data extraction and distribution architecture.

DATA EXTRACTION MEASUREMENTS

The data extraction procedure has been tested on the ATLAS TDAQ pre-series. A data producer emulates Level-2 Processing Units (L2PU) in 25 nodes in a rack; each node runs three L2PU. The 2nd level trigger is decoupled from the rest of the system by ring buffers: each L2PU writes muon pre-processed data to a ring buffer, a reader collects data from the buffers in the node and sends them to a data collector. If the buffer is full the event is discarded and the L2PU processes the next one. In this way, the only latency added to the trigger is due to data preparation (less than 200 μs) and no effect is due to backpressure (if any) of the data extraction system. The only other relevant parameter to be measured in the 2nd level nodes is the fraction of CPU time used by the data sender.

On the server side, measuring the CPU usage is mandatory since local server's resources, both in terms of I/O ad in terms of CPU usage must be mostly dedicated to server's specific tasks (NFS service, boot service). On the other hand, NFS service is not CPU expensive and the needed throughput for data extraction is negligible with respect to the available bandwidth.

Several protocols have been used to build client and server applications, all of them have been tested both with one and two levels of data concentration. In all the cases, the same application is used as local and final calibration server; the local server, in the two levels architecture, packs data in 64 kB packets before sending them to the final server. The structure of the software used for the tests is shown in figure 4.

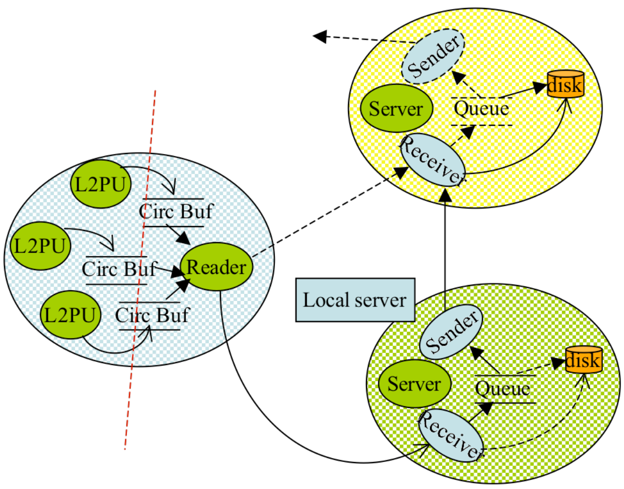

Figure 4. Structure of the software used for data extraction measurements.

In table 1, the CPU usage for TCP based applications in C++ and in Python and for applications built on top of the Atlas online monitoring library (based on CORBA) are shown for two levels. All the measurements have been performed on the pre-series of the Atlas TDAQ system, where a set of 30 2nd level trigger machines with a local server works in the Atlas TDAQ infrastructure environment. Data transfer from L2PU to local servers has been emulated using a level-2 rack plus a final server receiving data from the local one. The measurements on the final server have been performed generating packets of data as sent by local servers on 20 nodes and sending them to the calibration server.

Table 1. CPU usage

	TCP (C++)	TCP (Python)	Monitor
Sender	< 1%	< 1%	< 1%
Local server	1%	6%	2%
Final server	4%	14%	6%

The single data concentration level option is more complex. The simple TPC C++ option works after some optimization using about 10% of CPU on the final server. On the other hand the 2-levels option is simple and safe, having margins to accommodate requirements from other muon detectors (namely CSC).

REMOTE DATA DISTRIBUTION

Preliminary tests have been performed to check the feasibility of remote calibration. The data transfer between Tier 0 and Tier 1 at CNAF during Service Challenge 3 (SC3) has been measured.

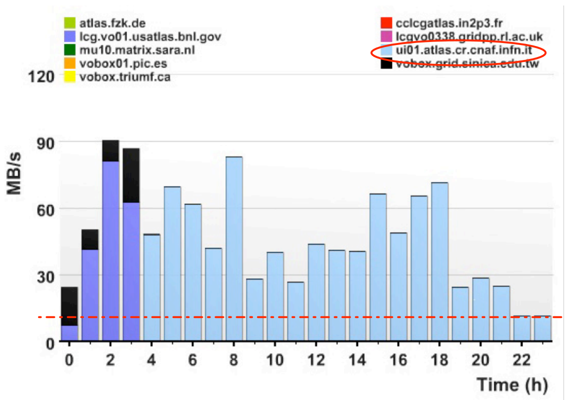

Figure 5. Data transfer tests between Tier 0 and Tier 1.

Though there is still no Quality of Service (QoS) implemented, the data transfer rate is always higher than the maximum required one (see figure 5). In addition, we have to consider a factor 2 of reduction on the requirement due to division of data sets between barrel and endcap data and a further reduction due to some event selection to balance the data flow between different parts of the detector. A further reduction factor could be obtained by on-the-fly data compression.

In figure 5 the result of a preliminary test of data transfer between Tier 0 and the Italian Tier 1 (CNAF) is compared with our maximum requirements (9.6 MB/s, dashed line).

CONCLUSIONS

A system to extract MDT data online from the Atlas 2nd level trigger system and to distribute them to remote calibration farms has been designed. The data extraction system has been tested extensively on the Atlas TDAQ pre-series in the Atlas infrastructure environment. Using the local file/boot servers in trigger racks to collect data from the L2PU and send them to the calibration server has been demonstrated to be simple and safe, independently from the data transfer protocol. Some very preliminary tests have been performed sending data from Tier 0 farm at CERN to the Italian Tier 1 in Bologna (CNAF), with encouraging results.