Deployment of the ATLAS High-Level Trigger

To cite this version:

HAL Id: in2p3-00145425
http://hal.in2p3.fr/in2p3-00145425
Submitted on 10 May 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Abstract—The ATLAS combined test beam in the second half of 2004 saw the first deployment of the ATLAS High-Level Trigger (HLT). The next steps are deployment on the pre-series farms in the experimental area during 2005, commissioning and cosmos tests with the full detector in 2006 and collisions in 2007. This paper reviews the experience gained in the test beam, describes the current status and discusses the further enhancements to be made. We address issues related to the dataflow, integration of selection algorithms, testing, software distribution, installation and improvements.

Index Terms—Cluster, high-energy physics, high-level triggers, software integration, testbeam, triggering.

I. INTRODUCTION

The ATLAS experiment [1] at CERN will use a 3-level trigger system to help identifying new physics phenomena generated by LHC proton-proton interactions. The program includes a search for the Higgs boson, super-symmetry and other new phenomena. The ATLAS detector is composed of specialized sub-detectors to register the properties of the decaying particles: an inner detector inside a magnetic field of 2 T measuring trajectories, a Calorimeter to measure energy and finally a muon spectrometer.

The First-Level Trigger (LVL1) is directly connected to the detector front-end electronics of the calorimeter and muon detectors. Fast algorithms implemented in custom hardware are used for LVL1 event selection. This trigger level also defines Regions of Interest (RoIs) in the detector where interesting physics signatures were found. Data of accepted events are sent out into the Data Acquisition system (DAQ) via Read-out Drivers (RODs) and are made available to the High-Level Triggers (HLT) through ∼1600 read-out buffers (ROBs), housed in ∼200 read-out systems (ROS) that implement the dataflow between these units and the former. The LVL1 trigger has to cope with the high input rate of the experiment (40 MHz), being designed to have a maximum output rate of 75 kHz, upgradeable to 100 kHz.

The RoIs found by LVL1 are used as seeds for the Second-Level Trigger (LVL2). The RoI information is transmitted to the LVL2 by means of a custom hardware component known as the RoI Builder. This component collects information from different LVL1 modules, concatenates and sends it to the LVL2 Supervisor (L2SV). The latter receives the LVL1 result and fans it out to one of the ∼1000 LVL2 Processing Units (L2PUs). L2PUs contain software encoded algorithms that will perform further event selection.

By only looking at data in LVL1 RoIs, it is possible to reduce the amount of data transferred to the LVL2 processors to less than 2% of the complete event of ∼1.3 MB, to achieve a further background rejection of ∼30 times. LVL2 selection algorithms request data from a variable numbers of RoIs, typically 1 or 2. An RoI spans on average 18 ROBs when located in the calorimeter, but only a maximum of 3 ROBs if the object of interest lies on the Muon detectors. If an event is accepted by LVL2, details of the processing at this trigger level, the LVL2 Result, is appended to the event stream and used by the Event Filter to proceed with the analysis.

The last trigger level is the Event Filter (EF). After a LVL2 accept, the full event data is assembled by special computing nodes (Subfarm Inputs, or SFIs) and redirected to specialized processing farms, where more elaborate filtering and monitoring algorithms are used. This level reduces the LVL2 output rate ∼3.2 kHz to only ∼200 Hz. If the event is accepted, it is recorded to permanent storage, via Sub-Farm Output nodes (SFOs) for later offline analysis. The final event stream will also contain a summary of the processing executed at this trigger level. The flow of triggers and data as explained in this text is depicted in Fig. 1. The messages are numbered indicating the order in which an event is processed inside the ATLAS trigger.

A. Time and Hardware Requirements

At LVL2, the total average processing time per event is expected to be ∼10 ms [2]. Thus, considering the LVL1 output rate, LVL2 will require a processing farm with a capacity equivalent to 1 000 CPUs. In this configuration, each processing node
should deliver a trigger decision at a rate of $\sim 100$ Hz, requiring an input bandwidth of 2.6 MB/s, based on the 2% estimate of data transfer per event processed by LVL2.

The total expected average processing time per event in the EF is about 1 s. Considering the LVL2 output rate, the EF will require a processing farm of 3 200 CPUs. The input bandwidth of every processing node in this configuration will be 1.3 MB/s.

B. The ATLAS High-Level Trigger Event Processing Framework

Both HLT subsystems, i.e., the LVL2 and EF use off-line software components as base for the event selection mixed with specialized algorithms created for running at these trigger levels, optimized for speed. A thin interface, the Steering Controller (SC) [3], binds the offline ATHENA/GAUDI [4] software environment to the HLT framework. Slightly different implementations of the SC are available for LVL2 and EF. Event selection runs in LVL2 in multiple, concurrent threads of execution, while the EF is process based.

In both cases, multiple algorithms are scheduled on a per-event basis by common steering software. The steering controls the execution sequence of algorithms based on the seed received, i.e., in LVL2, it uses the LVL1 result, while in EF, the LVL2 result.

II. Functional Description of the Trigger at 2004’s Combined Testbeam (CTB)

An experimental setup (shown in Figs. 2 and 3) to verify the functionality of the different detector components was put in place at CERN, starting early in June 2004. The setup was composed of different detector prototypes in final or almost final version and computing infrastructure running the latest Trigger and DAQ software to analyze and monitor data produced by beams from CERN’s Super Proton Synchrotron (SPS).

The order in which outcoming particles will interact in the final ATLAS detector was mimicked during the tests, although the setup resembles that of a fixed target experiment. The beams produced by the SPS passed through a magnet housing prototypes of the Pixel and Silicon Tracker (SCT) detectors, followed by a Transition Radiation Tracker (TRT) module. These Inner Detector modules were followed by a Liquid Argon (LArg) and Scintillating Tiles (hadronic) calorimeter prototypes and finally components of the Muon Detector.
As depicted in Fig. 4, the LVL1 Central Trigger Processor (CTP) [5], was connected to the detector Read-out Drivers (RODs) controlling the dataflow of accepted events. The system could be triggered either by:

- coincidence scintillators (hodoscope), for muons;
- interesting objects in the calorimeters or muon chambers, using the normal trigger logic;
- signals coming from the SPS timing electronics, via the LVL1 Timing, Trigger, and Controls module (TTC);

The trigger precedence was configured via hardware switches. The several trigger sources were combined using a standard LVL1 Local Trigger Processor module (LTP), which also acknowledged signals from the busy logic of detectors. A NIM fan-out module was used to deliver the LVL1 accept signal to the detector readout. These adjustments allowed the system to run with very few modifications and to maximize the number of standard LVL1 components that could be tested.

A. HLT Operation During ATLAS’s Testbeam

Although primarily intended for hardware tests, the HLT testbeam crew was able to test its software components together with the combined setup at this last opportunity before ATLAS commissioning. Being this the first occasion in which the HLT software was deployed together with other trigger components, a limited set of goals was established:

- Test the software interfaces between the dataflow and physics analysis components;
- Evaluate the modeled production cycle: offline tests, online deployment;
- Test the integration between the HLT sub-systems (LVL2 and EF) and between LVL1 and the former. In other words, if LVL1 could trigger event analysis at LVL2 and that in turn could seed the reconstruction at the EF as planned.

The LVL2 system for the testbeam was composed of a prototype of a custom hardware component connecting the LVL1 trigger and LVL2, known as the RoI Builder (RoIB), and 5 single-processor Intel Xeons running at 2.4 GHz and equipped with 1 Gb of memory each. One of the processors functioned as a L2SV, other three as L2PUs. The fifth machine was used for a dataflow application used to transmit the LVL2 to the EF, the so called pseudo-ROS (pROS), which has the responsibility of buffering the LVL2 result and acts as a normal detector ROS as seen from the event building subsystem. The in situ EF farm was composed of 4 nodes identical to those in LVL2. Other external EF farms were deployed throughout the testbeam period (see [6]) but will not be covered in this text. Every local EF node was running one Event Filter Dataflow Manager (EFD) and up to two processing tasks performing the event selection.

The dataflow as foreseen for final ATLAS was implemented. The decision returned by the L2PUs was forwarded to the Data Flow Manager (DFM) which initiated event building in the Sub-Farm Input processor (SFI). The SFI then sent the complete event to one out of the four EFDs.

All the events were accepted by default at the L2PU and Processing Task levels because in this way detectors groups could keep all data for later analysis of hardware problems. To ensure continuous data taking, in case of problems with the selection software, one of the L2PUs was equipped with a dummy version of the HLT event selection framework, that would force an accept decision without loading any data from the detector readout or executing any real calculations. The LVL2 result information from the other two L2PUs contained the relevant summaries of standard LVL2 processing that was used by the EF algorithms and monitoring tools. The events recorded by the Sub-Farm Output processor (SFO) also contained the detailed EF result. The events were finally stored on the CERN mass storage facility. This setup is shown in Fig. 5.

B. Algorithms Running in the HLT Framework

The L2PU in the test beam was configured to run track finding algorithms for the Pixel, SCT, TRT and muon detectors. The algorithms were scheduled by the HLT steering framework [7] using many software components from the ATLAS offline detector description, Event Data Model and infrastructure software.

Because of the nature of the tests executed in this Combined Testbeam (CTB), the various different parts of the LVL1 hardware were being tested, switched on, off or simply re-configured. The LVL1 Result contents therefore could not be used as a seed for LVL2. In absence of RoI information for all detectors, a
software simulation of the RoI data was used to initiate the event selection process. In an initial phase the raw data decoding software of a LVL2 muon selection algorithm was commissioned with beam and cosmic data. The raw data decoding software was in a second phase complemented with the full muon track fitting algorithm “muFast” [8], using alignment and calibration data. The obtained event/track features were encoded in the LVL2 result record, which was sent together with the LVL2 decision to the EF farm.

At EF, the event selection was further refined with the “TrigMoore” [6] track-finding algorithm. TrigMoore’s event reconstruction was “seeded” with the LVL2 result information. For both, LVL2 and EF, histograms allowed to monitor the selection process. The histograms were sent from the processing units to a histogram collection facility from where display programs could retrieve them for graphical presentation. Fig. 6 shows e.g., the distribution of track sagitta values obtained at LVL2.

Before installation at the testbeam site, the software was extensively tested in LVL2 and EF emulators with both simulated and previously recorded data, as devised in the development strategy defined by the HLT (see also [3]). Fully functional LVL2 and EF systems were also put together before final deployment at the CTB, at off-site testbeds, what ensured a smoother integration with the rest of the testbeam components at run time.

### III. EXPERIENCES

The complete setup described was used as part of the Trigger and DAQ chain of the 2004’s ATLAS Combined Testbeam (CTB) in several occasions, including extended data taking periods lasting several hours with good results. Millions of events have passed through LVL2 and EF components in this first full integration exercise. The LVL1 Result trigger path into the LVL2 system, composed by the RoIB and the L2SV worked without problems during the whole period. The data was transmitted successfully from the LVL1 subsystem into the single L2SV and then distributed to one of the three L2PUs. The L2PUs and EF nodes executing HLT algorithms worked reliably.

The algorithms running at the L2PUs, seeded by the L2SV, were able to take data from the ROB’s, transform it into higher-level objects and apply algorithmic work. The LVL2 decision and processing log were reused at the EF level to
monitor and confirm LVL2 analysis. The transmission path of the LVL2 Trigger result, using the pROS, was extensively tested and proved to work. The data were recorded reliably by the SFO and made available for offline studies.

A. Data Quality

Occasional detector data corruption was a source of difficulties during the test period causing crashes in the data conversion process. At these moments, either L2SV (on L2PU problems) or EFD (on PT problems) would time-out and let the event be recorded on disk. The data sets which contained corrupted events where investigated on the off-site testbeds and the problems were analyzed. The necessary protections were added to the data conversion modules. After these changes the system ran smoothly. Data samples taken during this period were of extreme importance to debug the detector readout and the HLT system.

B. Configuration and Software Installation

Because of the changing conditions at the testbeam, the HLT algorithms had to have their configuration changed quite frequently. The text-based configuration system from the Athena offline environment was not flexible enough in this operation mode and required specialist intervention.

The software installation was normally carried out by system administrators and, because of the varying nature of the data being collected, mostly due to new test conditions or different detector readout parameterization, the HLT testbeam crew was obliged to frequently patch the installed software release. This proved to be flexible enough for development, but non-scalable for setups with a larger number of computing nodes (networked filesystem problems were observed sometimes) and should be avoided in future.

IV. CONCLUSIONS AND OUTLOOK

The deployment of the ATLAS High-Level Triggers software in the 2004 Combined Testbeam has been presented. The objectives of this exercise were to verify the main design choices for this ATLAS subsystem and, in particular, the integration model between development and deployment of physics analysis algorithms. Good emphasis was given in testing the interfaces between the different trigger levels.

The HLT design is based on the extensive re-use of components from the ATLAS offline detector description, Event Data Model and infrastructure software. In practice, these components are compiled and distributed normally via the Offline software community and loaded into the trigger nodes using a set of special modules that mimic the Offline processing environment. This approach guarantees that the very same base of software components is used for triggering, reconstruction, analysis and simulation. In turn, it improves the overall understanding of the results obtained in any of these tasks and, specially, the offline reproducibility of trigger decisions.

The use of offline components for triggering naturally links development and deployment of software at the ATLAS High-Level Triggers. The production cycle starts inside the Offline software context, where the several tools that will be eventually deployed online are created. After extensive integration tests and having survived a reasonable number of real detector events, the tool is considered for inclusion into the HLT subsystem. These steps were followed during the testbeam exercise and proved to make a quite robust framework for trigger development.

While it seemed important to conduct offline tests prior to software deployment, the reproduction of the setups available at the testbeam location in outside testbeds was far from trivial. The testbeam machines had constant updates in the software installation which were difficult to reproduce. A centrally managed configuration system was also missed. The configuration options were, in some circumstances, difficult to reproduce outside the testbeam area, mainly due to the constraining text-file based technology used to configure the offline components running inside the HLT.

A new HLT configuration system should allow users and operators to set up, query and monitor software configurations for the different HLT components. The HLT software configuration will be integrated in the overall Trigger configuration system [9]. The system will use a database back-end to distribute the configuration information to (several thousands) HLT nodes. Organization of configuration information, integration in the overall system and scalability issues are presently being addressed.

The integration between the dataflow and offline components was tested during this exercise. Occasional corruption of data was observed. In most cases this effect was observed due to errors at the detector readout-drivers or associated electronics, because the different subdetector groups were conducting hardware tests in parallel. In some circumstances, these data anomalies also caused crashes in the HLT software. In general, it was felt that a greater amount of effort should be deployed in making trigger tools more resilient to corrupted detector data. Nevertheless, even on circumstances where an HLT processor would stop to work because of an error or a software problem, the dataflow system was still capable of saving the event for a posteriori analysis of the problem.

The LVL1 system was not sending out complete LVL1 results that could have been used for seeding the LVL2 system. Instead, a coincidence-based trigger for LVL1 was setup and that trigger impulse was used to initiate the execution at LVL2. The RoI based mechanism of LVL2 was, for these tests, replaced by fixed lookup on a number of ROBs and a rough estimate of the particle trajectory evaluated based on these data. The result summary, i.e., the LVL2 result was encoded and passed on to EF where a more accurate track-finding algorithm would refine the results of LVL2. The whole event data, including the detailed summaries of processing in each of the HLT subsystems were recorded to disk.

A large number of events flowed through the system without major problems, what points out that the dataflow subsystem seems mature enough for final deployment. The HLT development model based on the re-use of offline software components and extensive offsite testing worked as expected. The transmission of the LVL2 result as seed to EF algorithms has also been demonstrated. A few issues that need further development have been identified. They mainly concern the reproduction of setups,
configuration logging, monitoring and more robust error detection and handling.

ACKNOWLEDGMENT

The authors would like to thank colleagues working on the ATLAS data acquisition, detectors and offline for their enthusiasm and hard work during the Combined Test Beam.

A. dos Anjos is with the Department of Physics, University of Wisconsin, Madison, WI 53706 USA.
S. Armstrong, D. Damazio, and T. Maeno are with the Brookhaven National Laboratory (BNL), Upton, NY 11973 USA.
J. T. M. Baines, D. Emel’yanov, J. Kirk, F. J.wickens, and M. Wielers are with the Rutherford Appleton Laboratory, Chilton, Didcot, U.K.
H. P. Beck, S. Gadomski, C. Haeberli, S. Kabana, V. Perez Reale, and E. Thomas are with the Laboratory for High Energy Physics, University of Bern, Switzerland.
P. C. Bee, C. Meessen, Z. Qian, and F. Touchard are with the Centre de Physique des Particules de Marseille, IN2P3-CNRS-Université d’Aix-Marseille 2, France.
M. Biglietti is with the Dipartimento di Fisica dell’Università degli studi di Napoli Federico II e I.N.F.N., Napoli, Italy.
M. Bosman, P. Casado, H. Garitaonandia, C. Osuna, C. Sánchez, E. Segura, and S. Sushkov are with the Institut de Física d’Altes Energies (IFAE), Universitat Autònoma de Barcelona, Barcelona, Spain.
M. Caprini is with the National Institute of Physics and Nuclear Engineering, Bucharest, Romania.
B. Caron, R. Moore, J. L. Pinfold, and R. Soluk are with the University of Alberta, Edmonton, Canada.
G. Cataldi and A. Venture are with the Dipartimento di Fisica dell’Università di Lecce and I.N.F.N., Lecce, Italy.
D. Cavalli and S. Resconi are with the Dipartimento di Fisica dell’Università di Milano and I.N.F.N., Milano, Milan, Italy.
M. Ciobotaru and C. Meirosu are with the Politechnica University of Bucharest, Bucharest, Romania.
G. Comune is with the Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824 USA.
G. Cron, N. Konstantinidis, E. Stefanidis, and M. Sutton are with the Department of Physics and Astronomy, University College London, London, U.K.
A. de Santo, S. George, R. Goncalo, G. Kilvington, and A. Lowe are with the Department of Physics, Royal Holloway, University of London, Egham, U.K.
A. Di Mattia, S. Falciano, L. Luminari, A. Nisati, E. Pasqualucci, and S. Rosati are with the Dipartimento di Fisica dell’Università di Roma “La Sapienza” e I.N.F.N., Rome, Italy.
M. Díaz-Gomez is with the Section de Physique, Université de Genève, Switzerland.
B. Epp and V. Ghete are with the Institut für Experimentalphysik der Leopold-Franzens Universität, Innsbruck, Austria.
R. Ferrari, A. Negri, D.A. Scannicchio, W. Vandelli, and V. Vercesi are with the Dipartimento di Fisica Nucleare e Teorica dell’Università di Pavia and INFN, Pavia, Italy.
A. Kazarov and I. Soloviev are with the Petersburg Nuclear Physics Institute, Gatchina, Russia.
A. Khomich is with the Lehrstuhl für Informatik V, Universität Mannheim, Mannheim, Germany.
S. Kolos, A. Lankford, S. Stancu, and G. Unel are with the University of California, Irvine, Irvine, CA 92664 USA.
A. Kootz is with the Fachbereich Physik, Bergische Universität Wuppertal, Germany.
J. Masik is with the Institute of Physics, Academy of Sciences of the Czech Republic, Prague, Czech Republic.
A. G. Mello and J. M. Seixas are with the Universidade Federal do Rio de Janeiro, COPPE/EE, Rio de Janeiro, Brazil and Universidade Federal do Rio de Janeiro, COPPE/EE, Rio de Janeiro, Brazil.
P. Moretti, F. Parodi, and C. Schiavi are with the Dipartimento di Fisica dell’Università di Genova and I.N.F.N., Genoa, Italy.
N. Nikitin and S. Sivoklovok are with the Institute of Nuclear Physics, Moscow State University, Moscow, Russia.
N. Panikashvili and S. Tarem are with the Department of Physics, Technion, Haifa, Israel.
S. Tapprogge is with the Institut für Physik, Universität Mainz, Mainz, Germany.
G. Usai is with the Dipartimento di Fisica dell’Università di Pisa e I.N.F.N, Pisa, Italy.
B. Venda Pinto is with CFNUL—Universidade de Lisboa, Faculdade de Ciências, Lisbon, Portugal.
S. J. Wheeler is with the University of Alberta, Edmonton, Canada and also with the University of California, Irvine, Irvine, CA 92664 USA.
W. Wiedenmann and G. Zobernig are with the Department of Physics, University of Wisconsin, Madison, WI 53706 USA.
Y. Yasu is with KEK, High Energy Accelerator Research Organization, Tsukuba, Japan.

REFERENCES