

OPEN ACCESS

Experience with the custom-developed ATLAS offline trigger monitoring framework and reprocessing infrastructure

To cite this article: Valeria Bartsch 2012 *J. Phys.: Conf. Ser.* **396** 012006

View the [article online](#) for updates and enhancements.

You may also like

- [A trigger simulation framework for the ALICE experiment](#)
F Antinori, F Carminati, A Gheata et al.
- [B-physics and LVL1 di-muon trigger in the ATLAS experiment at the LHC](#)
Theodota Lagouri, Naoko Kanaya and Attila Krasznahorkay
- [Implementation of a triage monitoring program for internal exposure to short-lived radionuclides in Israel—challenges and recommendations](#)
R H Shukrun, L Epstein, J Koch et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Experience with the custom-developed ATLAS offline trigger monitoring framework and reprocessing infrastructure

Valeria Bartsch for the ATLAS TDAQ collaboration

University of Sussex, BN1 9RH, UK

E-mail: V.Bartsch@sussex.ac.uk

Abstract. The offline trigger monitoring consists of the data quality assessment which is done shortly after a data taking run has finished and the analysis of events where no trigger decision could be made. A reprocessing system tests changes to the trigger software and configuration to ensure their smooth deployment and online usage. This note explains the activities performed to provide a flawless operation of the trigger system and to determine the quality of the recorded data.

1. Introduction

Powerful and sensitive monitoring is vital for such a complex system as the ATLAS trigger. Any occurring processing failures, misbehaviour of selection algorithms and data defects must be discovered immediately and made known to the relevant experts. Events for which the software-based trigger system cannot make a decision are automatically sent into the debug stream, reprocessed and analysed offline. Any data not usable for physics analysis must be flagged. A complex system of data quality assessment has been developed which was used very successfully in 2011 and 2012. This is based on a first processing of a subset of data a few hours after recording, complemented by further monitoring performed during the bulk data processing approximately 48 hours later.

There is a steady stream of bug fixes and improvements to trigger algorithms. In addition the trigger configuration changes constantly due to the luminosity increases experienced in the first years of the Large Hadron Collider (LHC) operation. These changes need to be validated in the custom-developed reprocessing framework before they can be put into online operation.

2. ATLAS Trigger System

The three level trigger system [1] reduces event rate from up to 40 MHz delivered by the detector to about 400 Hz of the bunch crossings, while maintaining high efficiency for physics processes of interest. In 2011 the level 1 (L1) trigger, realised as a hardware trigger, has reduced the event rate down to ≤ 60 kHz based solely on muon and calorimeter information with a latency of $2.5 \mu\text{s}$. The following two software-based higher level steps collectively known as the High Level Trigger (HLT) consists of the level 2 (L2) and the event filter (EF). A positive L2 trigger decision based on only a part of the information contained in a whole event initiates the readout of the full event record and its transfer to a node of the EF farm. The L2 output rate in 2011

was about 4-5 kHz with an average latency of about 40 ms. The EF applies optimised offline analysis algorithms to data. Events accepted by the EF are routed to permanent storage. The average latency of the EF is about 4 s at an EF output rate of about 300-400 Hz.

3. Offline Data Quality Assessment

The offline data quality (DQ) assessment is implemented in the Data Quality Monitoring Framework (DQMF) which is described elsewhere [2]. The DQ assessment is based on an automated analyses of appropriate histograms created and filled during the standard raw data reconstruction. The offline DQMF framework displays those on a web page for visual inspection. It is also possible to implement and automatically apply algorithmic tests on the histograms in order to derive the DQ assessment results. The DQ distributions can be overlaid and compared with a reference. Possible automatic algorithms to evaluate the DQ are checks on the RMS, Mean, as well as simple Fermi fits, and Kolmogorov tests. The optimisation of the mean and the threshold values for the automatic checks is an iterative process so that the assessment becomes more adequate with time as experience with data is accumulated.

The DQ distributions of the trigger signatures can be roughly divided into

- transverse energy, E_T , or transverse momentum, p_T , distributions. E_T or p_T values are often used as threshold values in the trigger chains, thus the distributions of these parameters are vital for the DQ assessment. The thresholds of the trigger chains change with increasing luminosities. The reference plots for the E_T or p_T distributions need to be updated whenever the thresholds change.
- pseudorapidity, η and azimuthal angle, ϕ , distributions of the number of trigger objects at different trigger levels can quickly identify any geometry dependence of the trigger efficiency. The distributions can be used to determine the exact location of the deviating parameter. A further diagnosis might relate the location of the deviation to a detector issue, e.g. a dysfunctional part of the calorimeter, or an algorithm problem. In case of a significant change in the observed parameter two dimensional η - ϕ maps can be used to diagnose the issue.
- invariant mass distributions. Dimuon invariant mass distributions show a clear J/ψ peak which can be picked up on trigger level. This distribution is clearly the most sophisticated distribution to be used on trigger level and is quite independent on the trigger configuration or the running conditions.
- multiplicity distributions. The number of tracks can be checked against a reference. The number of tracks depends on the luminosity and therefore the reference needs to be updated whenever there is a major luminosity change.

In case of any issue with the trigger operation or trigger algorithms appropriate DQ defects need to be set. The defects are either tolerable or intolerable. Defects can be vetoed when selecting the data used in physics analysis. Intolerable defects are usually recorded when a defect is clearly an issue for most physics analysis. Tolerable defects are used when the perceived impact on physics analysis is much smaller.

In 2011 most of the defects set for the HLT were tolerable defects. Only 3% of data collected had an intolerable defect, see Fig. 1. The percentage of the tolerable defects is much higher. The high percentage reflects the fact that tolerable defects were used to document information about the trigger performance in the DQ defects. They are available for use in lists of good data taking periods for physics analyses if required and can safely be ignored in most cases.

4. Trigger Reprocessings

The trigger selection algorithms can be run offline to validate changes in the HLT, namely: software changes in the release (e.g. bug fixes), major changes of the trigger menu (e.g. new

L1			HLT						
Muon	Calo	CTP	electron	photon	muon	tau	jet	b-jet	missing E_T
99.0	100	99.8	99.3	99.3	100	99.9	98.6	99.9	99.3
Luminosity weighted relative fraction of good trigger data quality delivery during 2011 stable beams in pp collisions at $\sqrt{s}=7$ TeV between 13 March and 30 October (in %).									

Figure 1. Luminosity weighted relative fraction of good quality data delivery by the various components of the ATLAS trigger system during LHC fills with stable beams in proton-proton collisions, and after switching the tracking detectors on. Data taking period between March 13th and October 30th, 2011. The slight inefficiency in the HLT jet trigger is due to a misconfiguration associated with specific calorimeter conditions that occurred during two short data taking periods. However, all high p_T jets were still successfully triggered and recorded.

chains or thresholds) and condition changes (e.g. improved alignment or calibration). Small changes are assessed by a nightly test on a single computing node, with memory usage test and a count of the number of events accepted. However, if the proposed change influences data taking (e.g. if it is planned to put a new release into operation) the changes are tested by a trigger reprocessing to rerun the HLT decision. In addition the reprocessed data are reconstructed to obtain the offline objects and to allow a more thorough validation of the reprocessing. Sometimes the reconstruction step is skipped in case of trivial changes. The data from the trigger reprocessings and reconstructions are validated with similar tools as those used for the usual DQ assessment described in Sect. 3. The reprocessings are run over about one million events of a dataset collected with a very loose trigger selection, thus delivering a large enough data sample to validate the changes. The event selection by the HLT has to be deterministic otherwise changes between the trigger reprocessing to test and the reference are expected.

Although currently only computing resources within the PANDA framework [3] local to CERN are used, the framework allows the utilisation of distributed resources in future. This move has become necessary because of the larger datasets expected due to the increasing luminosities.

5. Debug Stream Treatment

Events for which the trigger could not make a decision within the allocated time end up in the debug stream. The presence of an event in a debug stream usually indicates a problem in some aspect of the online system. However, presence of events in the debug stream does not necessarily imply a misbehaviour at collision conditions; timeouts may be caused at the presence of a very busy event and in that case the timeout ensures the system robustness. At the current phase of the experiment the debug stream is a way to spot problems and weaknesses of the online system. Most of the debug stream events are caused by misconfigurations. Although misconfigurations do not happen very often, they can send a large number of events to the debug stream in a very short time. Other sources of debug stream events are timeouts or errors at L2 or EF. For most runs (where a run is a short data taking period often corresponding to an proton fill in the LHC) the percentage of events in the debug stream is lower than one per mille. Having dedicated tools to deal with the debug stream aims at identifying problems as soon as possible and reducing the turn-around time for fixing these problems. To achieve that, there is a documented procedure followed by an offline trigger expert in all events of the debug stream of each run. The goal is to achieve quasi real-time handling of the debug stream.

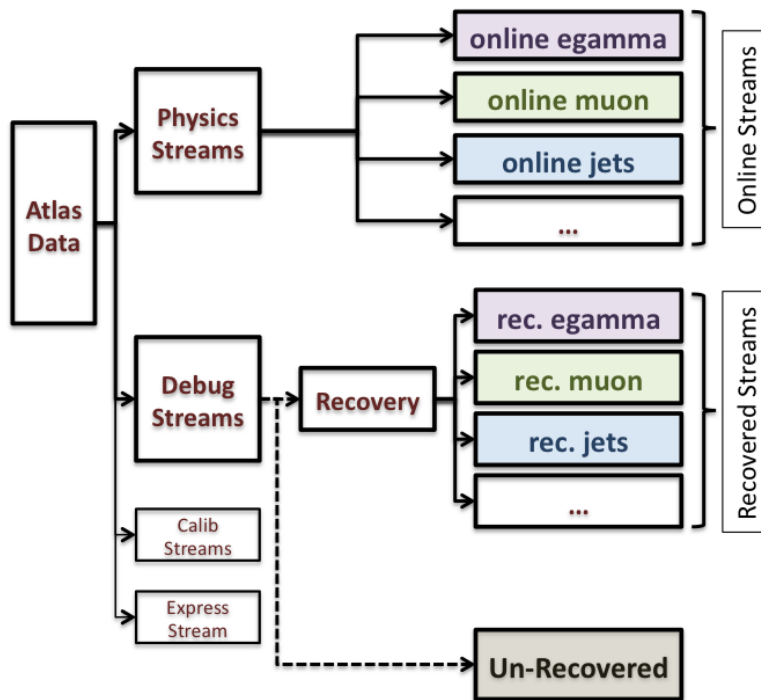


Figure 2. ATLAS data streams as recorded by the HLT. The data streams can be roughly divided into physics streams, calibration streams, debug streams, the express stream and several streams dedicated to special purposes. Whereas the physics, calibration and the express stream can be immediately used for physics analyses, calibration of the detector and a quick assessment of the recorded data, the debug stream is caused by a missing HLT decision. A recovery framework attempts to recover the debug stream dataset automatically. If the recovery is successful the data is streamed in the same manner as the physics streams. For most runs the percentage of events in the debug stream is lower than one per mille. For less than 0.1% of the debug stream data the recovery is not successful.

The raw data of the debug stream events is recorded and the trigger decision is rerun offline. If the recovery is successful the data streams obtained online (egamma, muon, jets) are recorded as pictured in Fig. 2. It is up to the physics groups to decide how to deal with the recovered data streams.

A thorough analysis of the debug stream events before and after recovery aims to find trigger chains with errors, correlation between detector and data acquisition issues, and non-recoverable events. Errors in the trigger chains are followed up and lead to improved trigger algorithms. Issues with the non-recoverable events are fixed quickly and improved for the next data run. Only 0.1% of the debug stream events remain unrecovered.

6. Summary

The motivation for developing a custom offline trigger monitoring and reprocessing system is to provide feedback on any issues with the ATLAS trigger system which do not become apparent during data taking. Any misbehaviour causing issues in the data quality is recorded.

In 2010 and 2011 the thorough offline data quality checks were challenging because of the changes of the instantaneous luminosity provided by the LHC which causes constant changes of

the ATLAS trigger configuration. Although the algorithms and threshold of the trigger needed to be constantly adapted only about three percent of all data had an intolerable defect resulting from the ATLAS trigger system. This is achieved by the strict validation of any trigger changes before they are used online with the help of trigger reprocessings. In addition after each run has finished events with a failed HLT decision are analysed in detail and recovered for physics analyses.

In 2012, after about two years of data taking the offline monitoring framework and reprocessing infrastructure have evolved into an essential and reliable component of the experiment operation.

References

- [1] S. George [ATLAS Collaboration], PoS ICHEP **2010** (2010) 487.
- [2] S. Kolos, A. Corso-Radu, H. Hadavand, M. Hauschild and R. Kehoe, “A software framework for Data Quality Monitoring in ATLAS,” J. Phys. Conf. Ser. **119** (2008) 022033.
- [3] T. Maeno [ATLAS Collaboration], “PanDA: Distributed production and distributed analysis system for ATLAS,” J. Phys. Conf. Ser. **119** (2008) 062036.