ATLAS Draft Cover

Commissioning of the ATLAS Reconstruction Software with First Data

Version: 1.1

To be submitted to: IEEE NSS Conference Record

Corresponding editor:

W. Lampl (walter.lampl@cern.ch)

Comments are due by: Nov 12th, 2009

ATL-SOFT-PROC-2009-008
23 November 2009

Commissioning of the ATLAS Reconstruction Software with First Data

Walter Lampl, University of Arizona
On behalf of the ATLAS Collaboration

***Abstract*—Looking towards first LHC collisions, the ATLAS detector is being commissioned using all types of physics data available: cosmic rays and events produced during a few days of LHC single beam operations. In addition to putting in place the trigger and data acquisition chains, commissioning of the full software chain is a main goal. This is interesting not only to ensure that the reconstruction, and monitoring chains are ready to deal with LHC physics data, but also to understand the detector performance in view of achieving the physics requirements. The recorded data have allowed us to study the ATLAS detector in terms of efficiencies, resolutions, channel integrity, alignment and calibrations. They have also allowed us to test and optimize the sub-systems reconstruction as well as some combined algorithms, such as combined tracking tools and different muon identification algorithms. The status of the integration of the complete software chain will be presented as well as the data analysis results.**

***Index Terms*—ATLAS, Reconstruction, Commissioning**

I. Introduction

ATLAS is a general-purpose experiment at the Large Hadron Collider (LHC) at CERN. It is designed to cover the largest possible range of physics expected from the LHC. The ATLAS detector will start to take physics data when the LHC starts operating at the end of 2009. The last years have been dedicated to the commissioning of the experimental setup. This includes the detector hardware, the data-acquisition system, the trigger and the software and computing system for reconstruction and analysis of the data taken. ATLAS had two sources of particles for detector commissioning: Cosmic rays and during a short period in September 2008 single-beam data from LHC.

II. The ATLAS Detector and Reconstruction Software

A. Detector systems in ATLAS

FIGURE 1 shows a schematic view of the ATLAS detector with the subsystems labelled. The overall size of the detector is 44m in length and 22m in diameter. It has a total weight of about 7000 tons. The size is dominated by the Muon Spectrometer and its superconducting air-core toroid providing a magnetic field of 0.5 Tesla. The muon trajectory is measured with high precision in the bending direction by Monitored Drift Tubes (MDT). Resistive Plate Chambers (RPC) in the barrel region and Thin Gap Chambers (TGC) in the end-caps are used for triggering.

Inside the muon system are the calorimeters. ATLAS uses sampling calorimeters for hadronic as well as electromagnetic calorimeters. The hadronic barrel consists of scintillating tiles inside an iron absorber structure, covering $\eta < 1.7$. The hadronic end-caps (HEC) and the EM calorimeters use liquid Argon as the active material. The HEC covers the range $1.5 < \eta < 3.2$ and uses copper as absorber. The Electromagnetic Calorimeters provide coverage up to a pseudo-rapidity of 3.2. The absorber is made of accordion-shaped Lead sheets. It has three longitudinal compartments supplemented by a pre-shower detector. The forward region is covered by a dedicated forward calorimeters up to η=4.9.

The magnetic field for tracking is provided by a 2 Tesla solenoid magnet. The Inner Detector consists of three subdetectors: A Pixel Detector around the interaction point followed by a Semiconductor Tracker and a Transitions Radiation Tracker at outer radii.

A complete discussion of the ATLAS Detector can be found in [1].

B. The ATLAS Reconstruction Framework

THE reconstruction software framework developed by the ATLAS collaboration is called *ATHENA*. The algorithmic part is mostly written in C++ with a few components in Python and FORTRAN. The job configuration is done in Python.

Athena is a flexible framework that is used not only for reconstruction but also for simulation, monitoring, analysis and to compute calibration constants.

Athena distinguishes between *data objects* that hold only data and have no algorithmic capabilities and *Algorithms* and *AlgTools* that do computations. Dedicated stores are used to pass data objects from one Algorithm to the next. Data Objects can be written out to disk using the POOL-ROOT technology.

The ATLAS Computing Model [2] foresees the following file types:

- RAW Raw data coming from the detector or produced by simulation. Simulated raw data is stored in POOL files while real data produced by the ATLAS DAQ is in a *byte-stream* format.
- ESD Event Summary Data, is the most comprehensive data type in ATLAS. It contains object reconstructed from raw input like calorimeter cells energies or tracks as well as physics objects.
- AOD Analysis Object Data, contains reconstructed physics objects like electrons or jets.
- DPD Derived Physics Data, contains more condensed information for specific analyses.
- TAG A database or a ROOT file allowing quick identification of events inside an ESD or AOD.

The reconstruction uses calibration and alignment constants as well as detector control data stored in the ATLAS conditions

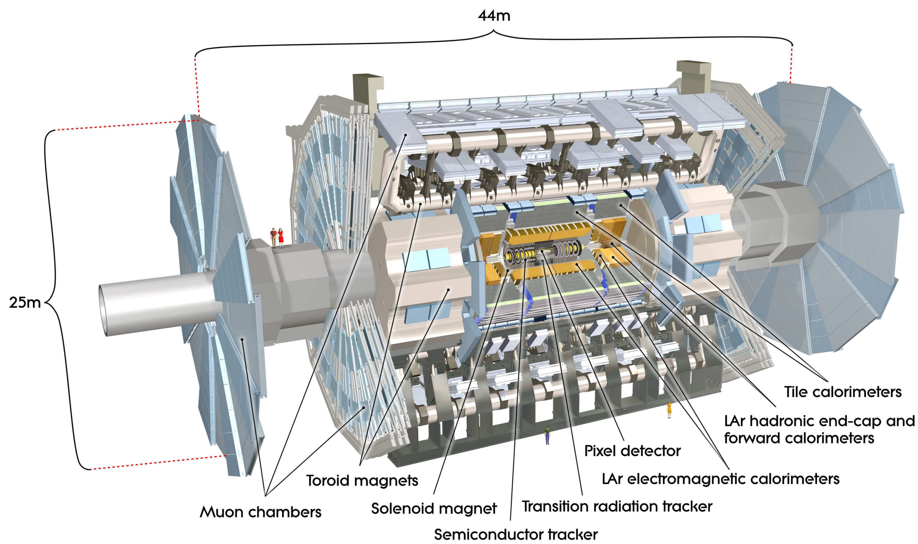

Fig. 1. Schematic overview of the ATLAS detector

database. Calibration constants can be derived from a variety of sources like dedicated pedestal or pulser runs or from physics data.

III. Commissioning of the ATLAS detector with data

THE installation of the ATLAS Detector in the cavern started in Summer 2004 with the bottom section of the Tile Calorimeter and finished in late Summer 2008. ATLAS was ready to take data at this point. Each detector element was commissioned with cosmic rays as soon as installed in the detector. Combined cosmic data-taking with modules of all major subsystems started in June 2007 using the combined ATLAS data acquisition and trigger system. Dedicated trigger streams were set up for cosmic data taking (see also figure 3).

A. *Cosmic data taking*

Several campaigns of combined cosmic data taking took place, involving several (sometimes all) ATLAS subdetectors. The longest was from September through November 2008. More than 200 million events (500 Terabytes or raw data) were recorded during this time with all subsystems participating. Figure 2 shows an event display of a cosmic muon passing close to the interaction point that was recorded in fall 2008. A similar cosmic data taking campaign happened in June 2009 with more than 90 million events recorded. Figure 3 shows the accumulation of data during this time with the magnetic field configuration indicated.

B. *Beam Splash Events*

During the start-up of the LHC in September 2008 several proton bunches at injection energy were shot on closed collimators located about 140m from ATLAS. Thus large particle showers were produced that traversed the detector along the beam axis. A very large fraction of active detector elements were hit by these showers. One beam-splash event

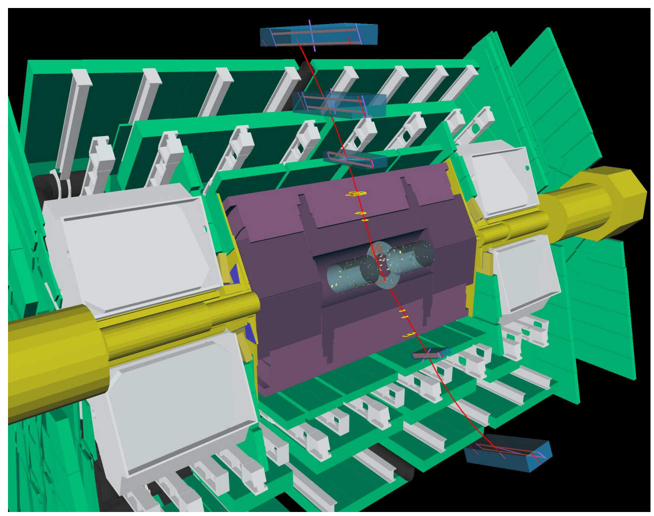

Fig. 2. Event display (VP1) of a cosmic muon traversing ATLAS close to the interaction point.

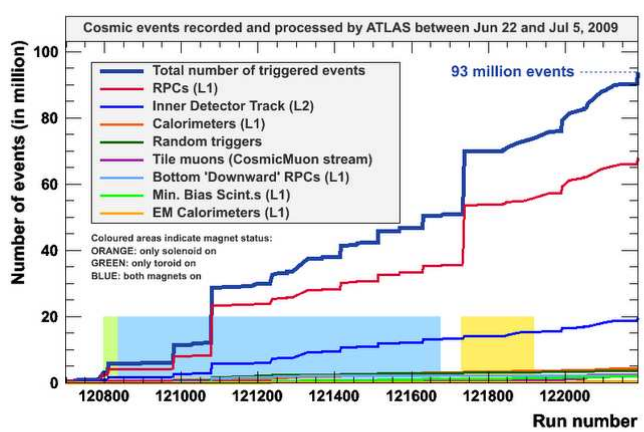

Fig. 3. Accumulated of cosmic events during the June 2009 cosmic run for each trigger stream. The color bars indicate the magnetic field configuration.

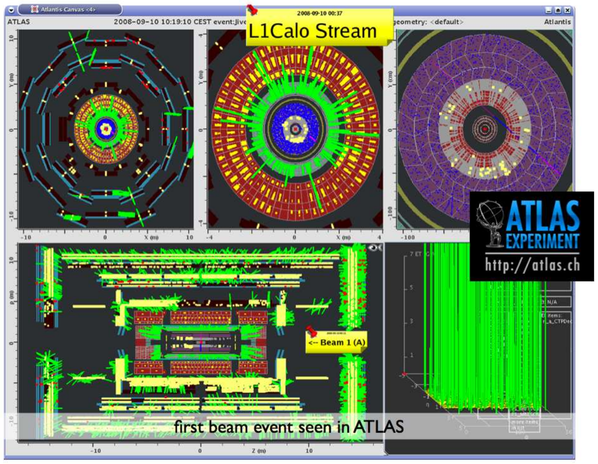

Fig. 4. Event Display (Atlantis) of a beam splash event in ATLAS

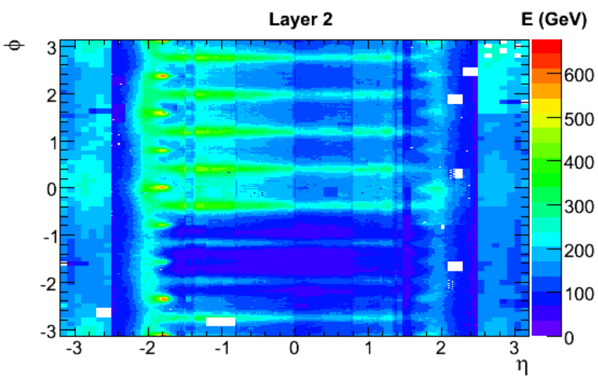

Fig. 5. Energy deposit in the second layer of the electromagnetic calorimeter accumulated over 100 beam-splash events. The non-uniform energy deposit can be explained by the non-uniform distribution of material between the collimator.

lead to 1000 TeV total energy deposit in the calorimeters. Such events are very useful for detector commissioning: They allow to identify dead detector elements and to verify the time synchronization between sub-systems.

Figure 4 shows an Atlantis event display of one beam-splash event. Figure 5 gives an eta-phi map of energy deposited in the middle layer of the Electromagnetic calorimeter. The plot was made by accumulating over 100 events. Only cells with an energy deposit bigger then five times the noise are taken into account. It's clearly visible that the energy deposit in the upper half of the detector (positive phi) is larger than in the lower part. This can be explained by the fact that the LHC machine is closer on the bottom of the tunnel, so particle going upwards have a longer path-length in the air of the tunnel and less in the surrounding rock than particles going downwards. The eight-fold modulation in phi is due to the eight-fold structure of the end-cap toroid magnet.

The cosmic and beam-splash data taken so far allow intial alignment and timing constants to be derived, measure the noise of detector elements and build maps of dead or affected regions.

IV. Software Commissioning

THE cosmic-ray and single-beam data that is taken during the commissioning phase of the detector is reconstructed using the regular ATLAS offline software. This serves two purposes: The first purpose is the commissioning the reconstruction software by exposing it to real data. The second purpose is supporting hardware commissioning. The output of the reconstruction is the basis of many studies of the detector performance. Providing reconstruction software for detector commissioning and commissioning the reconstruction software (with collisions in mind) are two separate goals. They have much in common but bear also some conflicts. The needs of cosmic reconstruction and detector commissioning require certain adaptions to the reconstruction model foreseen for LHC collisions:

- Unlike particles from LHC collisions, cosmic rays arrive asynchronosly with respect to the ATLAS readout clock. The event phase has to be found event-by-event, increasing the processing time. On top of this, the various detector elements are not yet perfectly timed-in and the relative timing of signals from different detector elements can change over time as the hardware installation progresses.
- Cosmic rays are not pointing. In order to reconstruct tracks, the tracking software needs to run with wider cuts, leading to increased CPU time per event.
- In order to understand and debug detector issues, detector experts need access to raw quantities that are often big in volume. This leads to an increased disk space consumption per event. For the purpose of commissioning special "Commissioning DPDs" have been introduced. They contain the full ESD information of events passing certain filter criteria ("skimmed ESD"). Similarly a specialized version of the tag file was introduced to simplify finding events interesting for commissioning.
- Data taking conditions during the commissioning phase are rather unstable. Subsequent runs can have very different configuration in terms of detector coverage. In many cases runs are taken with only one or a few subsystems participating. In this case only a fraction of the reconstruction software can run. The magnetic field configuration is also changing frequently.

Running the reconstruction software on real data exposed a number of problems and shortcomings that were not detectable with simulated data. All these issues were addressed meanwhile, leading to a better debugged and more usable code than was available two years ago.

V. Organized reconstruction campaigns

A. Prompt Reconstruction

ALL raw data coming out of the ATLAS Detector is reconstruction with little latency at the *Tier0*, a computing farm at the CERN site. The latency is typically 15 minutes, to ensure that all relevant detector-control data (voltages, temperatures, etc.) are available in the conditions database. The output of these reconstruction jobs are ESDs, AODs and monitoring histograms to allow the shift crew to assess the quality of the data taken. To help detector commissioning also specialized n-tuples are sometimes part of the Tier0 output.

For steady-state collision reconstruction it is foreseen to split the prompt reconstruction into an Express Stream that is reconstructed immediately and bulk processing with 24 hour delay that can take advantage of calibration constants derived from the express stream processing. This model was exercised on simulation data but not on cosmic data.

B. *Re-processing*

Particularly interesting subsets of the cosmic data ATLAS has taken so far have been re-processed several times with an updated version of the reconstruction software and improved calibration constants. These re-reconstruction campaigns were done at off-site computing centers, taking advantage of the concept of grid-computing. ATLAS has ten large computing facilities called *Tier1s* that are suitable for bulk reconstruction of raw data. These centers are spread over Europe, North America and Asia. The input data that is to be reconstructed is distributed among the Tier1s according to an agreed share.

The cosmic data taking in fall 2008 has been re-processed three times on the grid. The first time in January 2009, the second time around Easter 2009. The third reprocessing happened in summer 2009 and did not start from raw input data but used the ESDs from the previous reprocessing as input.

The data taking during a cosmic campaign in June 2009 was re-processed very shortly after the end of the data-taking campaign using the same software release that was used at the Tier0, only calibration constants were updated. This strategy is called "Fast Reprocessing".

VI. Patching and Testing Strategy

THE continuous nature of the reconstruction at the Tier0 requires a method to update the software used without disrupting the ongoing processing. In particular if problems arise quick patching is essential.

For the prompt reconstruction, patches are collected and automatically compiled twice per day. The compilation is followed by a set of tests on real events in various configurations. These test run over up to 10k events. The purpose of these test is to catch basic technical problems like segmentation violations, to monitor the resource requirements and to assess the impact of bug-fixes on the physics performance. The output is verified by comparing histograms to references. The compilation and testing sequence is completely automated. For new major releases a dedicated bug-hunting-campaign is performed, typically on about one million events.

Based on the result of the automated tests, patches are accepted or rejected. Once the number of patches becomes too large, a new patch-release is built. This involved a complete recompilation of all packages.

This patching strategy has been successfully used throughout the software commissioning period. The testing strategy was initially simpler and grew significantly in sophistication over the last year.

Re-processing campaigns use a fixed software version that is tested and signed-off before the campaign starts. Since reprocessing jobs are scattered over several computing centers, a dedicated site-validation is needed to ensure that all involved Tier1s give the same result on the same input.

VII. Conclusion

Commissioning of the reconstruction software is an integral part of the overall commissioning of the ATLAS experiment. Over the last years ATLAS gathered plenty of experience in reconstructing and analyzing cosmic ray data. The ATLAS reconstruction software is well tested on simulated collision data as well as on real data from cosmic runs. A patching and testing strategy is prepared to deal unexpected problem arising when reconstructing the first collision events.

We are looking forward to see first LHC collisions at the end of this year and are confident that the ATLAS reconstruction software is in good shape to reconstruct LHC data.

References

[1] The ATLAS Collaboration, G. Aad et al., The ATLAS Experiment at the CERN Large Hadron Collider, JINST 3 (2008) S08003.
[2] ATLAS Computing Technical Design Report (TDR) 2005 CERN-LHCC-2005-02