Processing, Calibration and Reprocessing of ATLAS Data from LHC Collisions at 7 TeV

Michael Böhler[*]**, on behalf of the ATLAS Collaboration**

Deutsches Elektronen-Synchrotron (DESY)

E-mail: michael.boehler@cern.ch

The ATLAS experiment has successfully recorded, reconstructed and analyzed millions of collision events delivered by the LHC at an unprecedented centre-of-mass energy of 7 TeV. The involved large-scale data processing operations worked remarkably well from an early commisioning period that gradually evolved to a stable operation mode aiming at physics. The early commissioning samples were promptly processed at the CERN Tier-0 computer farm in all ATLAS data formats: from raw to reconstructed to a variety of reduced datasets. These outputs were used to improve the reconstruction software and derive state-of-the-art calibration and alignment constants.

These improvements were then used for a series of centralized data reprocessings using worldwide distributed computing on the Grid. The reprocessing exercises were done coherently for real and simulated data samples and provided the high-quality data sets required for the publication of the first ATLAS physics results. By April 2010, the prompt reconstruction was stabilized and a mechanism to pre-calibrate the detector run-by-run was put in place. New data samples produced by the Tier-0 are now directly usable for physics analysis and can be combined with the earlier reprocessed data, real and simulated.

35th International Conference of High Energy Physics - ICHEP2010,
July 22-28, 2010
Paris France

---

[*]Speaker.
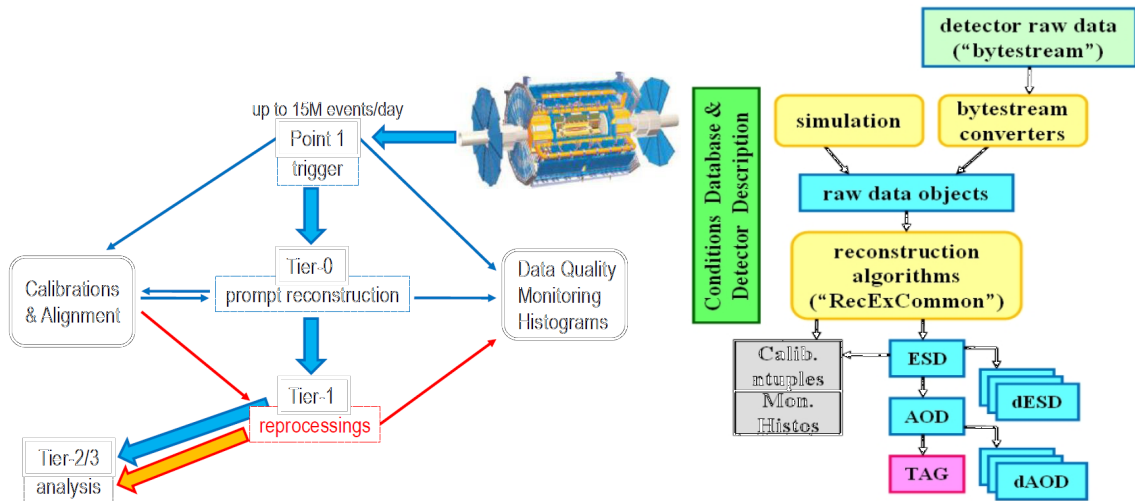
Figure 1: a) ATLAS data flow   b) ATLAS data formats produced at Tier-0

1. Data Processing

The processing, calibration and reprocessing of ATLAS[1] data is summarised in the AT-LAS data flow shown in Fig. 1a). All events accepted by the ATLAS trigger system are processed in a so-called prompt reconstruction at a CERN-based computing farm (Tier-0). During the prompt reconstruction dedicated data streams are provided: an express stream and several calibration streams.

These streams are used to derive offline calibration constants. All calibration constants are validated and uploaded to the conditions database. During a 36 hours time period e.g. the calorimeters and tracking detectors check for hot channels and mask them, if needed. In addition, an offline beam spot is determined using the offline reconstruction based on the express stream data. The beam spot position and width are provided for five luminosity blocks grouped (one luminosity block ≈ 120 sec. data taking). These are then input to the reconstruction of the physics streams.

Up to the conference, the ATLAS experiment has successfully recorded data corresponding to a total integrated luminosity of 338 nb^{-1}[2].

2. ATLAS Event Data Model

The ATLAS Event Data Model foresees different data formats with different levels of detail and different event selection criteria: e.g. detailed reconstructed objects are written in Event Summary Data files (ESD) for detector and reconstruction performance studies. Smaller Analysis Object Data files (AOD) are designed for physics analysis. Other dedicated formats with further data reduction and filtering are provided as well (see Fig 1b)). TAG, dESD and dAOD files are provided for further data reduction. TAG allows to select individual events quickly. dESD/dAOD have analysis-specific event filtering. These data formats are produced during the reconstruction of the physics streams at the Tier-0. They are then distributed to Tier-1 and Tier-2 centers all over the world.

3. Frozen Tier-0

During the early commissioning phase the reconstruction software had to be updated frequently. After stabilizing the software, the release installed at Tier-0 was frozen for 2010 proton-proton data-taking period. Only updates which did not change the physics content of the data were allowed to be added to the Tier-0 software, for instance monitoring software updates. This allows new data to be immediately combined with already (re)processed data samples and simulated data from the latest (re)processing campaign. If updates are neccessary or new releases are available raw data from previous data taking periods have to be re-processed (red arrow Fig. 1 a)). These reprocessing jobs are run on the world-wide distributed computing Grid at different Tier-1 centres using exactly the same setup for Monte-Carlo samples and raw data.

References

[1] Aad, G and Abat, E and Abbott, et al., *Expected performance of the ATLAS experiment: detector trigger and physics*, CERN, Geneva, 2009, arXiv:0901.0512[hep-ex]

[2] The ATLAS Collaboration, *Luminosity Determination Using the ATLAS Detector*, ATLAS-CONF-2010-060, CERN, Geneva, 2010