

ATLAS Data Quality Monitoring: Experience with First Collision Data

Peter Waller*, on behalf of the ATLAS Collaboration

University of Liverpool

E-mail: pwaller@hep.ph.liv.ac.uk

It is essential to get quick feedback at all stages of the collection and reconstruction of particle physics data, and to correctly record quality decisions to ensure that only good data are used to obtain physics results. The ATLAS data quality system provides prompt investigation of collected data, initial calibrations, and later reconstruction, and propagates the corresponding quality decisions to analysis users.

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

*Speaker.

Inner Tracking Detectors			Calorimeters				Muon Detectors			
Pixel	SCT	TRT	LAr EM	LAr HAD	LAr FWD	Tile	MDT	RPC	TGC	CSC
97.1	98.2	100	93.8	98.8	99.1	100	97.9	96.1	98.1	97.4

Figure 1: Luminosity weighted relative detector uptime and good quality data delivery during 2010 stable beams at $\sqrt{s} = 7$ TeV between March 30th and July 16th, in %.

1. Introduction

Between March and June 2010, ATLAS has recorded 243 nb^{-1} of pp collisions, with each subdetector maintaining above 93 % efficiency for usable physics data during stable beams (Fig. 1). Determining which data is suitable for physics analysis requires inputs from people and automated systems to aid the decision making process [1] [2].

2. Data Quality Flags

ATLAS datataking is segmented by run and luminosity block (LB). A run corresponds nominally to an LHC fill, and an LB is a length of time over which detector conditions are assumed to be constant. A decision comes in the form of a ‘traffic light’ coloured flag, whose evolution across LBs is stored in the conditions database. Red and green mean bad and good for analysis respectively, whilst yellow is a temporary value used to defer a decision to a later time. Other states exist which are only meaningful as diagnostic information. Before the flags become inputs to analyses, experts converge on the ‘good’ and ‘bad’ states.

Example decisions include those coming from the DQMF (Data Quality Monitoring Framework) [3], the DCS (Detector Control System) [4], and the control room shifter. Some of these flags can be seen in Figure 2. They are summarized into a single decision per physical subdetector region (such as ‘pixel detector barrel’ or ‘monitored drift tube endcap’) by the ‘Data Quality Calculator’. An expert reviews the automated decision and may decide to override it. Regions are summarized across subdetectors into ‘combined performance’ flags used to select ranges of good runs and luminosity blocks. Along with information about the LHC state, these flags can be used to build a ‘Good Run List’, an XML document describing ranges of data suitable for particular types of physics analyses.

3. Web-based Histogram Service

Events collected from the experiment are filtered into different streams for specific uses. There exists three types of stream: physics, calibration and express. The physics streams will ultimately end up in user analyses, the calibration stream is used to compute calibration constants and the express stream undergoes rapid processing (usually 1-2 hours after the data is recorded).

The histogram analysis framework produces, per active stream, 20,000 histograms (100 MB) per run, along with 700 histograms (6 MB) per ten luminosity blocks. Originally, these were

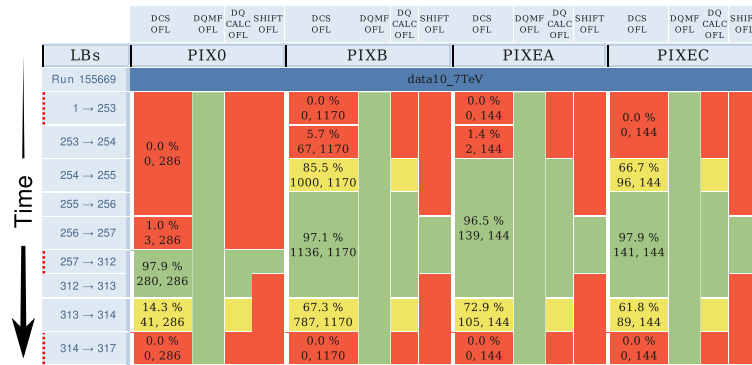


Figure 2: Data quality flags for regions of the pixel subdetector (Layer 0, Barrel, Endcap A and Endcap C) are shown for the different entries, DCSOFL, DQMFOFL, DQCALCOFL (from the Data Quality Calculator) and SHIFTOFL (the final expert decision). The percentage represents the fraction of the region considered functional, and the two numbers represent the number of modules working and the total number of modules respectively. The colour of the cells represents the decision made by each system. In this run, the pixel detectors are turned on after the LHC declares stable beams. This can be seen where the state transitions from bad to good, then the reverse happens at the end of the run when the beam is dumped.

all dumped as PNG images so that they could be served over the web to shifters and experts as static content, but the disk space and CPU usage per run were not sustainable. Given that the vast majority of histograms are not looked at for each run, it is better to generate the image files as they are requested and to cache them for a short time. This has been implemented and has led to a saving in disk space, whilst not impacting the load of the server.

4. Conclusion and Future

The ATLAS Data Quality Monitoring infrastructure has performed successfully, allowing rapid turn around between the recording and analysis of data. Small technical changes have been necessary to stay within resource limitations, especially with respect to conserving disk space. The current flagging system requires that the data quality group try to summarize any problem that may arise with effectively a single piece of information: whether the data is good or not. Other information about the state of data taking which may be useful to some analyses can become lost or hard to obtain. A proposal is under discussion to switch from the ‘traffic light’ colour code to a mechanism which records information about different ‘defects’ present at a given point in the data taking.

References

- [1] ATLAS Collaboration. Assessment of data quality in ATLAS. *Journal of Physics: Conference Series*, 119(4):042034, 2008.
- [2] ATLAS Collaboration. ATLAS offline data quality monitoring. *Journal of Physics: Conference Series*, 219(4):042018, 2010.
- [3] S Kolos, A Corso-Radu, H Hadavand, M Hauschild, and R Kehoe. A software framework for Data Quality Monitoring in ATLAS. *Journal of Physics: Conference Series*, 119(2):022033, 2008.
- [4] ATLAS Collaboration. Data quality from the Detector Control System at the ATLAS experiment. *Journal of Physics: Conference Series*, 219(2):022037, 2010.