ATLAS Online Data Quality Monitoring

C. Cuenca Almenar[a] *, A. Corso-Radu[b], H. Hadavand[c], Y. Ilchenko[c], S. Kolos[b], K. Slagle[b], A. Taffard[b]

[a]Department of Physics, Yale University
P.O. Box 208120, New Haven, CT 06520-8120, USA

[b]Department of Physics & Astronomy, 4129 Frederick Reines Hall
University of California, Irvine, CA 92697-4575, USA

[c]Department of Physics, Southern Methodist University
Dallas, Texas 75275-0175, USA

Every minute of an ATLAS data taking session, the monitoring framework serves several thousands of physics events to the monitoring data analysis applications and handles millions of histogram updates coming from many different applications. This framework also executes over forty thousand advanced data quality checks for a subset of those histograms and displays those histograms together with the results of all those checks on several dozens of monitors installed in main and satellite ATLAS control rooms.

The Data Quality Monitoring Display (DQMD) is the visualization tool for the automatic data quality assessment. It is the interface through which the shift crew and the experts can validate the quality of the data being recorded or processed, be warned of problems related to data quality, and identify the origin of such problems. This tool allows great flexibility for visualization of histograms, with an overlay of reference histograms when applicable, configurations used for automatic checking of those histograms, and the results. The display configuration is stored in a database, that can be easily created and edited with the Data Quality Monitoring Configurator (DQMC) tool.

The first months of collisions data taking turned into a very successful experience for the monitoring framework and translated into several improvements to easy usability and efficient information transfer. A description of the design and implementation of the DQMD and DQMC will be presented, as well as the performance of the monitoring framework during the first ATLAS run and the recent upgrades concerning alarm handling and parameters finding.

1. INTRODUCTION

The ATLAS experiment is one of the general-purpose particle physics detectors designed and built to record the proton-proton collisions provided by the Large Hadron Collider (LHC) at CERN, Geneva. The LHC will produce proton head-on collisions with center-of-mass energy of 14 TeV at a rate of 40 MHz, rate at which ATLAS has to be able to analyze and filter the information provided by its approximately one hundred and forty million channels.

Due to the complexity of the ATLAS experiment, a framework for automatic data quality

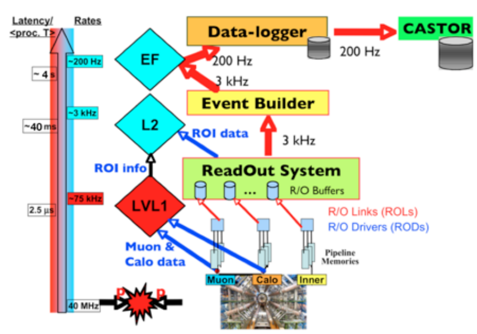

Figure 1. The data acquisition architecture: data flow makes data fragments available to trigger processors and to mass storage.

*corresponding author e-mail: Cristobal.Cuenca@cern.ch

assessments of incoming data and a visualization tool for easy identification of problems are essential. A highly scalable distributed Data Quality monitoring framework (DQMF) has been developed and is being used to monitor the quality of the data as well as operational conditions of hardware and software elements of the detector, trigger, and data acquisition systems. [2].

2. THE DATA QUALITY MONITORING FRAMEWORK

The Data Quality Monitoring Framework is a data driven distributed and scalable framework to monitor data quality both online and offline.

Single data quality tests are defined by DQParameters. Each DQParameter specifies what input histogram(s) to use, what algorithm and parameters (DQAlgorithm) to apply and the thresholds to classify the result (DQResult) as good or bad. All the DQParameters are grouped in different DQRegions, and these DQRegions can be themselves combined in more general DQRegions, thus forming a hierarchy, DQTree. The mechanism to combine the results of the subparameters and subregions is specified in the configuration in the same way that the algorithms are specified for each DQParameters. In this structure, each subdetector of ATLAS is described as a DQRegion that will host several tiers of DQRegions and DQParameters to check the performance of its hardware elements and data being taken. Finally, the DQAgents are the applications that execute specific analysis algorithms on various types of data (e.g. histograms, messages, etc) and produce color coded DQResults relaying the quality of data according to a particular configuration specified by the user.

Given the diversity of information sources and destinations, the framework implements the input and output as plug-ins. With this solution, data can be read from the configuration data base, from some information services or from a ROOT file. The same way, the output can be sent to a conditions data base, and information services or a ROOT file.

3. THE DATA QUALITY MONITORING DISPLAY

The Data Quality Monitoring Display is a graphical interface implemented using a QT-ROOT library developed by BNL. The strengths of this implementation are that performance is scalable, and the complete functionality of ROOT histograms is readily available for an interactive display.

The display allows easy navigation between Regions and Parameters providing great flexibility for visualization of Results produced by DQMF. For any particular data acquisition configuration, all the available DQRegions and DQParameters are organized in a tree where each element is colored according to the result of the algorithm applied. For each DQRegion, the status color is inferred from the status of underlying DQParameters taken with some weight coefficients defined in configuration that determine importance of those Parameters for that Region. Each Region being defined to represent a particular set of Parameters and/or Regions stands in the tree as a branch node with a summary Result status for that set.

For any Parameter, DQMD displays histograms used for the check (overlaid by corresponding reference histograms if applicable), configuration parameters of the Algorithms used to run the DQMF Algorithm, history graph of the data quality results. The Display also supports visualization of the Results in graphical form, i.e. a hardware view of the detector to easily detect faulty channels or modules by eye. The graphical representation of the detector and its subsystems allows for navigation functionality. The display provides the shift crew with a checklist before the final assessment of the data is saved to the database, a list of experts to contact in case of specific problems, and actions to perform in case of failure.

Moreover, specific information regarding each DQParameter is provided in the configuration and shown in the display, as well as instructions to the shifters about the actions to be taken in case of problems.

4. THE DATA QUALITY MONITORING CONFIGURATOR

Data quality layouts allow for easier understanding of the status of the subsystems and faster navigation throughout the DQTree, and in particular to problematic regions. Layouts are defined and configured together with the tree structure and the automatic test parameters. Clicking on each part of the layout brings the appropriate subtree or result.

However, designing these layouts and translating them into configuration language is not trivial. An application has been provided for this reason: the Data Quality Monitoring Configurator, DQMC.

DQMC, Figure 2, allows experts to create new configurations or modify the existing configuration database by specifying various layouts and shape parameters of the detector sub-system until it reaches a desired look. Configuration database is written in platform-independent extensible markup language (XML). XML allows storing hierarchically structured data and provides the ability to validate it with respect to the schema. Using the set of basic shapes and layouts, experts can easily create advanced graphical representation of any detector sub-system or any part of the sub-system and then save it to the configuration database for future use in the Data Quality Monitoring Display.

The application graphical interface is divided in three parts. The left hand side window shows the DQTree as it would be displayed in DQMD. However, the symbols by the names of each tree element indicate whether a proper layout has been configured. Clicking on any element of the tree will display, in the central window, the layout associated to this element. This layout can be modified with the controls on the right hand side panel.

5. CURRENT STATUS AND CONCLUSIONS

So far, the framework handles about 20 DQAgents and more than 75000 DQParameters organized in more than 15000 DQRegions. In normal running conditions, more than 150000 new results are generated every minute.

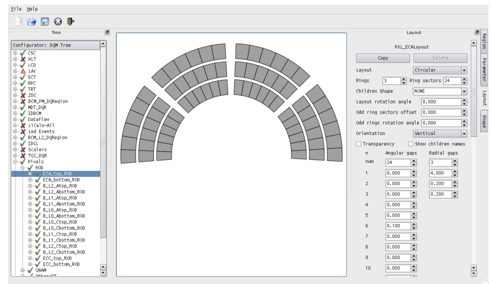

Figure 2. Data Quality Monitoring Configurator

Since deployment of the first prototype at the end of 2008, the feedback provided by system experts and shifters has resulted in many upgrades, specially in visualization tools. DQMC is currently used by most systems to generate layouts and DQMD is the tool used in most desks in the ATLAS control room to ensure good data quality taking and chase down any issues that might arise.

The Data Quality Monitoring Framework has been successfully commissioned and improved, proving to be able to meet the stringent ATLAS requirements and very useful to ensure good data quality.

This same framework is reused offline to assess DQ and set offline DQ flags. The good runs lists for the first physics results were generated using this offline DQ assessment. Currently, DQMF is actively being used to ensure good data taking with collisions runs at 7 TeV. It will not discover new physics, but it does ensure we are taking and analyzing good data for our first discoveries.

REFERENCES

1. The ATLAS Collaboration, ATLAS DAQ, EF, LVL2, and DCS, 1998 *CERN/LHCC/98-16*
2. Y. Ilchenko *et al*, Data Quality Monitoring Display for ATLAS experiment at the LHC, 2010 *J. Phys.: Conf. Ser.* 219 *022035*