Data Quality Monitoring of the CMS Tracker

Suchandra Dutta, on behalf of the CMS Tracker Collaboration
National Central University, Taipei, Taiwan
E-mail: Suchandra.Dutta@cern.ch

Abstract. The Data Quality Monitoring system for the Tracker has been developed within the CMS Software framework. It has been designed to be used during online data taking as well as during offline reconstruction. The main goal of the online system is to monitor detector performance and identify problems very efficiently during data collection so that proper actions can be taken to fix it. On the other hand any issue with data reconstruction or calibration can be detected during offline processing using the same tool. The monitoring is performed using histograms which are filled with information from raw and reconstructed data computed at the level of individual detectors. Furthermore, statistical tests are performed on these histograms to check the quality and flags are generated automatically. Results are visualized with web based graphical user interfaces. Final data certification is done combining these automatic flags and manual inspection. The Tracker DQM system has been successfully used during cosmic data taking and it has been optimised to fulfill the condition of collision data taking. In this paper we describe the functionality of the CMS Tracker DQM system and the experience acquired during proton-proton collision.

1. Introduction
The Data Quality Monitoring (DQM) in an experiment ensures that the detector components and trigger system work optimally during data taking and certifies recorded data for physics analysis. It is one of the most important tools and must be robust as it needs to function throughout the life time of the experiment. It is required to be very efficient to spot problems in the detector and trigger systems at the earliest possible stage of data acquisition and to issue alerts for the relevant experts so that immediate actions can be taken.

In the Compact Muon Solenoid (CMS) at the Large Hadron Collider (LHC) at CERN, the DQM is done in a common infrastructure and individual sub-detector or trigger DQM systems are developed based on this infrastructure. We shall discuss here the DQM system developed for the Silicon Strip Tracker (SST) of CMS. The SST consists of 15 K microstrip silicon detector modules arranged in 10 cylindrical barrel layers and 12 end-cap disks on forward and backward sides. These detectors are read out by about 10 million electronic channels. The complexity of the SST DQM system reflects the complexity of the Tracker system. It was essential to develop a few SST specific DQM tools to ensure optimal functionality of the SST.

2. Data Quality Monitoring Scheme
The DQM framework in CMS is embedded in the CMS software environment (CMSSW) [1] and provides the basis for creation, transportation and visualization of monitoring information. Monitoring is done in the form of histograms, numbers or strings. The Object Oriented analysis framework ROOT [2] is used for this purpose.
The CMS Data Quality Monitoring system has three components as shown in figure 1. The
histograms are created and filled by accessing event information **Producer** modules.
Histograms are accessed by the **Consumer** module which performs further analysis and produces
summary histograms and/or numbers. The module uses the **Quality Test** tool available in the
DQM infrastructure to perform statistical tests on histograms comparing them with a reference.
The reference could be a histogram or a set of values like range, mean, r.m.s etc. The deviation
from the reference is classified as **Warning** or **Error** depending on the severity and alarms are
generated.
The histograms are visualised in web based Graphical User Interfaces (GUI) which is discussed
in detail in section 4.
The data quality in CMS is monitored in real-time (online) during data taking as well as in
various stages of the offline processing of event data. During data acquisition, apart from a
few specific cases, a fraction of events are processed to have quick feedback on the detector
performance. The same DQM application is used during offline reconstruction, within a few
hours of data taking, where complete event statistics is available and fully reconstructed and
calibrated objects are accessible. The main purpose here is to check the data in detail to spot
reconstruction as well as calibration problems. The output from both online and offline DQM
applications is visualised on the web GUI and also archived for future reference.

3. Monitoring Information and Organisation
In the SST DQM system [3], a number of **Producer** modules have been developed. Each producer
module deals with a specific type of event information and the histograms are defined and
filled accordingly. The monitoring information, accessed from different levels of reconstruction,
consists of
- **Raw Data Monitoring**: the basic event information with different error codes which arise
during data unpacking
- **Digitisation Monitoring**: digitisation characteristics, namely amplitude, hottest and coolest
  strips and number of digitised signals etc.
• Cluster Monitoring: cluster properties like charge, position, width, noise, etc.
• Track Monitoring: reconstructed track parameters such as momentum, track angles, fit qualities, number of hits used in track fit etc.
• Cluster Monitoring with Track: the properties of on-track and off-track clusters are studied

Histograms are defined for individual detector modules as well as globally depending on the nature of the information. The granularity of the SST is directly reflected in the number of histograms (∼ 300 K). The geometrical structure of the tracker is followed to create a tree like folder structure to store histograms. Histograms related to individual tracker components are placed inside the corresponding folder in order to facilitate quick and easy access.

A dedicated consumer module has been developed in the SST DQM system. One of it’s tasks is to create summary histograms combining information from individual detector level histograms. For example, the mean values of a given histogram type in the lowest level of the tree structure is accessed to fill new histograms at geometrically higher levels. Summary histograms are very efficient to locate a possible problem and save the time needed to go through individual detector level histograms. The consumer accesses the Quality Test flags from different levels of the SST and finally combines them to produce a global flag. This is used in data certification.

4. Web Based Visualisation

![Figure 2](image_url)

**Figure 2.** SST specific page of the CMS DQM GUI where 7 groups of histograms are are shown using SST layout. The histograms represent overview of the SST performance. The reference histograms are superimposed in some cases and are shown as gray filled areas.

The visualisation of DQM histograms is achieved in web based Graphical User Interfaces (GUIs). The main advantage of the web based GUI is that one just needs a standard web browser without
the overhead of any specific software installation. The DQM histograms are thus available not only locally in CMS control rooms but also remotely, from anywhere in the world. The Grid authentication protocol is imposed to the CMS members to access the GUI.

A generic GUI with a common look & feel for all the subsystems has been developed which is used as the main visualisation tool in CMS [4]. A back-end server, based on CherryPy, a Python language web server framework [5], accesses histograms from DQM applications and make them available as images following the requests received from the web interfaces. The server provides navigation facilities for histograms and visualisation of colour coded alarms generated by the Quality Test. DQM sub systems are responsible to provide the layouts, in order to group histograms together for specific purposes. Additionally, a plug-in mechanism is implemented that allows one to specify custom drawing options for histograms as available in ROOT. There is also an option to overlay reference histograms. A snapshot of the SST specific page of the CMS Web GUI is shown in figure 2.

A specialised GUI has been developed in the SST DQM system [6] targeted at expert usage. Once a problem is detected by the shifters, SST experts can use it to study and understand the problem in detail. It is directly connected with the SST DQM Consumer unlike the CMS DQM GUI. Hence specific expert level actions can be taken. A snapshot of the alarm navigation page of the SST Expert GUI is shown in figure 3. One can follow the coloured tree structure to find out faulty detector modules and the reason of the failure.

![SiStrip Expert GUI : Alarm View](image.png)

**Figure 3.** The snapshot of the SST Expert GUI where the alarm navigation is shown. One can follow the coloured tree structure to find out faulty detector modules and the reason of the failure.

The usual colour notation is used where error and warning are represented in Red and Yellow colours and Green is used to indicate a good flag.

There is an option in the GUI to create a two dimensional synoptic view of the whole tracker as an interactive Scalable Vector Graphics image [7, 6] as shown in figure 4. The cylindrical layers are represented in rectangular boxes and the end-cap discs are represented as circular areas. Each pixel in these regions represent a detector module and they are painted by the resulting flags of the Quality Test. One can visualise the whole SST and easily identify faulty substructures, if any.
Figure 4. Two dimensional synoptic view of the SST painted by Quality Test flags. One can easily identify faulty region as shown in one of the end-cap discs.

5. Data Certification
In CMS, data certification is composed of manual and automatic components and is provided run by run. The certification procedure is described in a block diagram in figure 5. The automatic flag in a given run is derived from the Quality Test results. In the SST, a simple algorithm is used to combine individual results and an overall flag is created which represents the fraction of perfectly working detectors. The online and offline DQM shifters check selected set of histograms defined by the SST experts, compare them with reference and provide manual feedback for each run. Further manual input comes from the SST experts where SST specific tools are utilised to study the histograms in detail. One such tool is the Historic DQM [8] where the trend of mean

Figure 5. The certification procedure in CMS DQM system is represented in a block diagram.
or r.m.s of a few SST observables are plotted as a function of run numbers and the outliers are identified. Bookkeeping of automatic and manual flags together with data taking conditions is done in the CMS Run Registry [9] which has a database back-end.

6. Summary
The SST DQM system is performing excellently since data taking started in CMS. The operation has been very stable and the system provides key information about the SST performance and ensures goodness of data for physics analysis. Only 0.2% of data are marked bad for physics analysis by the SST DQM system which is mainly due to hardware failure and scheduled calibration activities. The system was prompt enough to spot them as problem or deviation from standard behaviour. The system proved to be robust against several orders of magnitude increase in LHC peak luminosity.

7. References
[1] CMS Physics Technical Design Report Volume I, Detector Performance and Software, CERN/LHCC/2006-001, CMS TDR 8.1, 2 February 2006
[2] ROOT, An Object-Oriented Data Analysis Framework, http://root.cern.ch
[3] Data Quality Monitoring of the CMS Experiment: the Tracker Case
S. Dutta
Nuclear Physics B Proceedings Supplement 197 (2009) p.271
[4] CMS data quality monitoring web service
L. Tuura, G Eulisse and A Meyer
2010 J.Phys:Conf Ser. 219 072055
[5] CherryPyA pythonic, object-oriented HTTP framework, 2009, http://cherrypy.org
[6] The CMS tracker Data Quality Monitoring expert GUI
F. M. Palmonari, S. Dutta
Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE
Digital Object Identifier: 10.1109/NSSMIC.2009.5401981
Publication Year: 2009 , Page(s): 602 - 604
[7] Use of Interactive SVG Maps and Web Services to Monitor CMS Tracker
M.S. Menea, G. Zito
Proceedings of IEEE-NPSS Real Time Conference, 2005
[8] History Plotting Tool of Data Quality Monitoring
D. Giordano , A.-C. Le Bihan, A. Pierro and M. De Mattia
Nuclear Ins. Methods in Physics Research A, NIMA50295
[9] CMS Run Registry : Data Certification Bookkeeping and Publication System
V. Rapsevicius
Proceeding of CHEP10