

CMS data quality monitoring: systems and experiences

L Tuura¹, A Meyer^{2,3}, I Segoni³, G Della Ricca^{4,5}

¹ Northeastern University, Boston, MA, USA

² DESY, Hamburg, Germany

³ CERN, Geneva, Switzerland

⁴ INFN Sezione di Trieste, Trieste, Italy

⁵ Università di Trieste, Trieste, Italy

E-mail: lat@cern.ch, andreas.meyer@cern.ch, ilaria.segoni@cern.ch, giuseppe.della-ricca@ts.infn.it

Abstract. In the last two years the CMS experiment has commissioned a full end to end data quality monitoring system in tandem with progress in the detector commissioning. We present the data quality monitoring and certification systems in place, from online data taking to delivering certified data sets for physics analyses, release validation and offline re-reconstruction activities at Tier-1s. We discuss the main results and lessons learnt so far in the commissioning and early detector operation. We outline our practical operations arrangements and the key technical implementation aspects.

1. Overview

Data quality monitoring (DQM) is critically important for the detector and operation efficiency, and for the reliable certification of the recorded data for physics analyses. The CMS experiment at CERN's Large Hadron Collider [1] has standardised on a single end to end DQM chain (Fig. 1). The system comprises:

- tools for the creation, filling, transport and archival of histogram and scalar monitor elements, with standardised algorithms for performing automated quality and validity tests on distributions;
- monitoring systems live online for the detector, the trigger, the DAQ hardware status and data throughput, for the offline reconstruction and for validating calibration results, software releases and simulated data;
- visualisation of the monitoring results;
- certification of datasets and subsets thereof for physics analyses;
- retrieval of DQM quantities from the conditions database;
- standardisation and integration of DQM components in CMS software releases;
- organisation and operation of the activities, including shifts and tutorials.

The high-level goal of the system is to discover and pin-point errors—problems occurring in detector hardware or reconstruction software—early, with sufficient accuracy and clarity to reach

good detector and operation efficiency. Toward this end, standardised high-level views distill the body of quality information into summaries with significant explaining power. Operationally CMS partitions the DQM activities in online and offline to *data processing*, *visualisation*, *certification* and *sign-off*, as illustrated in Fig. 1 and described further in subsequent sections. The CMS DQM supports mostly automated processes, but use of the tools is also foreseen for the interactive and semi-automated data processing at the CAF analysis facility [3].

2. Online DQM system

2.1. Data processing

As illustrated in Fig. 2(a), the online DQM applications are an integral part of the rest of the event data processing at the cluster at CMS Point-5. DQM distributions are created at two different levels, *high-level trigger filter units* and *data quality monitoring applications*.

The high-level trigger filter units process events at up to 100 kHz and produce a limited number of histograms. These histograms are delivered from the filter units to the storage managers at the end of each luminosity section. Identical histograms across different filter units are summed together and sent to a storage manager proxy server, which saves the histograms to files and serves them to DQM consumer applications along with the events.

The data quality monitoring applications receive event data and trigger histograms from a DQM monitoring event stream from the storage manager proxy at the rate of about 10-15 Hz, usually one application per subsystem. Events are filtered for the stream by applying trigger path selections specified by the DQM group. Each DQM application requests data specifying a

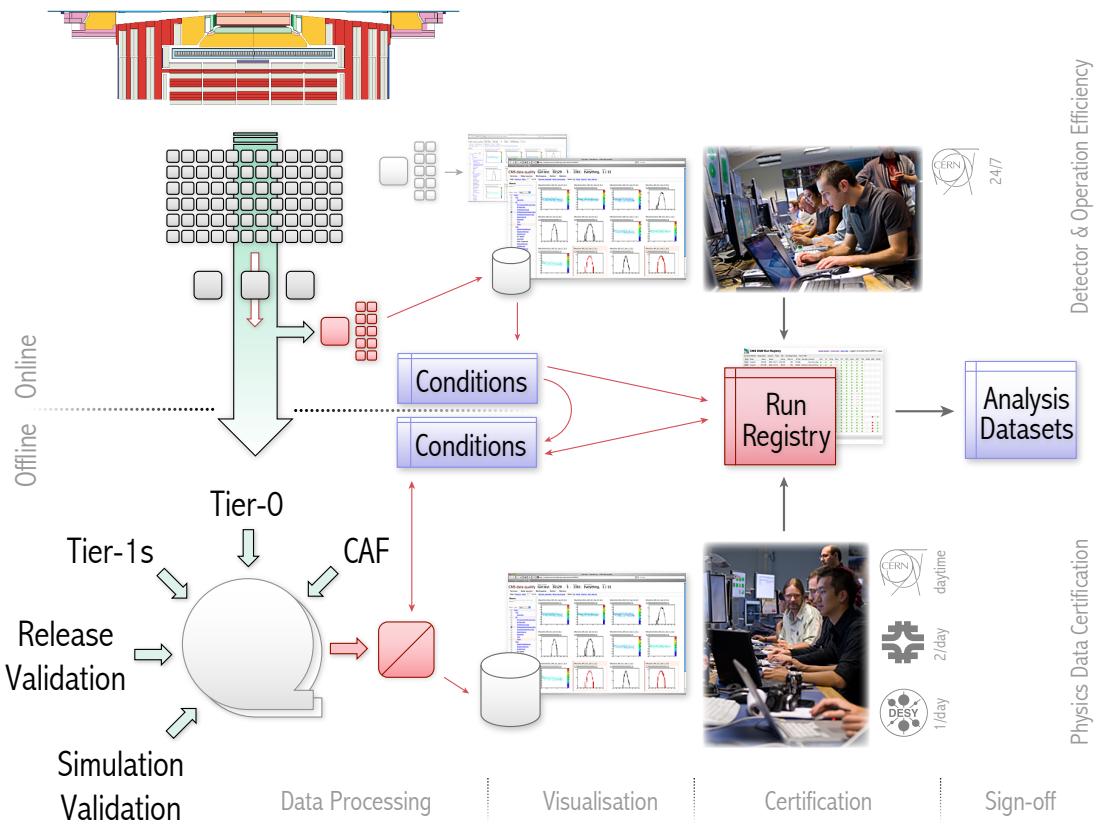


Figure 1. DQM system overview.

subset of those paths as a further filter. There is no special event sorting or handling, nor any guarantee to deliver different events to parallel DQM applications. The DQM stream provides raw data products only, and on explicit request additional high level trigger information.

Each application receives events from the storage manager proxy over HTTP and runs its choice of algorithms and analysis modules and generates its results in the form of *monitoring elements*, including meta data such as the run number and the time the last event was seen. The detector level algorithms include for example checks for hot, cold or otherwise bad channels, data integrity, noise and pedestal levels, occupancy, timing problems, reconstructed quantities, trigger issues, and detector-specific known problems. The applications re-run reconstruction according to the monitoring needs. The monitor element output includes reference histograms and quality test results. The latter are defined using a generic standard quality testing module, and are configured via an XML file.

2.2. Visualisation

All the result monitor element data is made available to a central DQM GUI for visualisation in real time [2]. The data includes alarm states based on quality test results. During the run the data are also stored to a ROOT file [4] from time to time. At the end of the run the final archived results are uploaded to a large disk pool on the central GUI server. There the files are merged to larger size and backed up to tape. The automatic certification summary from the online DQM step is extracted and uploaded to the run registry and on to the condition database (see section 4), where it can be analysed using another web-based monitoring tool, WBM [6]. Several months of recent DQM data is kept on disk available for archive web browsing.

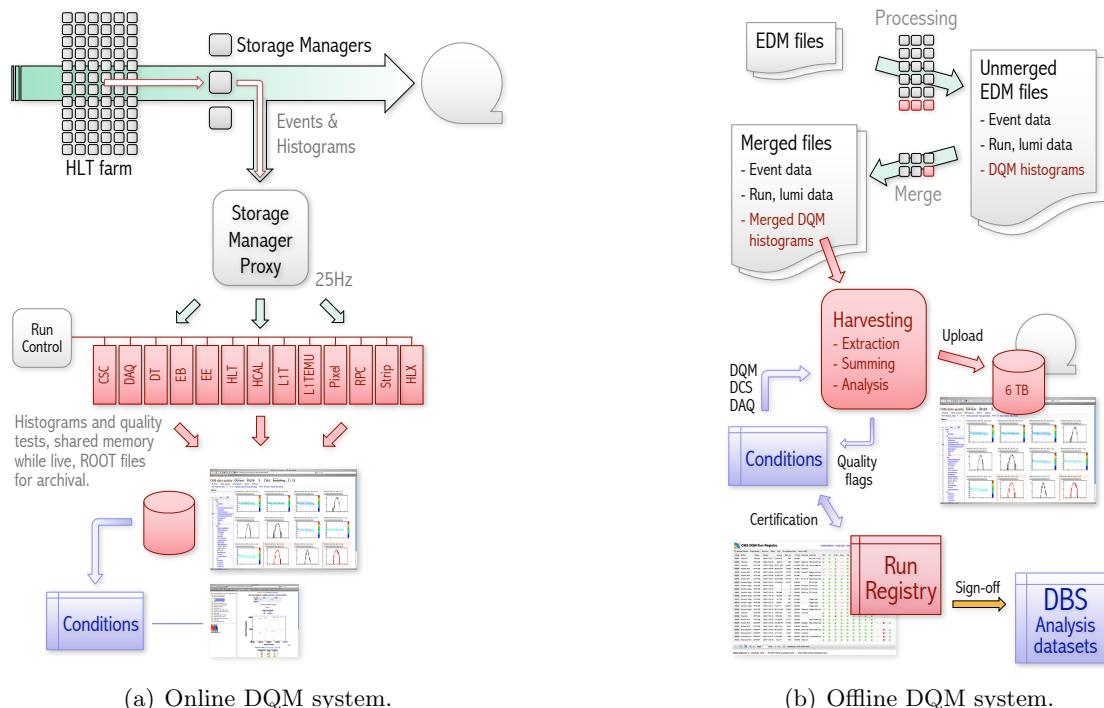


Figure 2. DQM workflows.

2.3. Operation

Detector performance groups provide the application configurations to execute, with the choice of conditions, reference histograms and the quality test parameters to use and any code updates required. Reviewed configurations are deployed into a central replica *playback integration test system*, where they are first tested against recent data for about 24 hours. If no problems appear, the production configuration is upgraded. This practice allows CMS to maintain high quality standard with reasonable response time, free of artificial dead-lines.

The central DQM team invests significantly in three major areas: 1) to integrate and standardise the DQM processes, in particular to define and enforce standard interfaces, naming conventions and the appearance and behaviour of the summary level information; 2) to organise shift activities, maintain sufficiently useful shift documentation, and train people taking shifts; and 3) to support and consult the subsystem DQM responsibles and the physicists using the DQM tools.

All the data processing components, including the storage manager proxy, the DQM applications and the event display, start and stop automatically under centralised CMS run control [5]. The DQM GUI and WBM web servers are long-lived server systems which are independent of the run control. The file management on the DQM GUI server is increasingly automated.

3. Offline DQM systems

3.1. Data processing

As illustrated in Fig. 1, numerous offline workflows in CMS involve data quality monitoring: Tier-0 prompt reconstruction, re-reconstruction at the Tier-1s and the validation of the alignment and calibration results, the software releases and the simulated data. These systems vary considerably in location, data content and timing, but as far as DQM is concerned, CMS has standardised on a single two-step process for all these activities, shown in Fig. 2(b).

In the first step the histogram monitor elements are created and filled with information from the CMS event data. The histograms are stored as *run products* along with the processed events to the normal output event data files. When the CMS data processing systems merge output files together, the histograms are automatically summed together to form the first partial result.

In a second *harvesting step*, run at least once at the end of the data processing and sometimes periodically during it, the histograms are extracted from the event data files and summed together across the entire run to yield full event statistics on the entire dataset. The application also obtains detector control system (DCS, in particular high-voltage system) and data acquisition (DAQ) status information from the offline condition database, analyses these using detector-specific algorithms, and may create new histograms such as high-level detector or physics object summaries.

The final histograms are used to calculate efficiencies and checked for quality, in particular compared against reference distributions. The harvesting algorithms also compute the preliminary automatic *data certification decision*. The histograms, certification results and quality test results along with any alarms are output to a ROOT file, which is then uploaded to a central DQM GUI web server.

The key differences between the various offline DQM processes are in content and timing. The Tier-0 and the Tier-1s re-determine the detector status on real data using full event statistics and full reconstruction, and add higher-level physics object monitoring to the detector and trigger performance monitoring. The Tier-0 does so at time scale of a day or two whereas the Tier-1 re-processing takes from days to weeks. On CAF the time to validate alignment and calibration quantities varies from hours to days. The validation cycle of simulated data reflects the sample production times, and varies anywhere from hours for release validation to weeks on large samples. The validation of simulated data differs from detector data in that entire datasets

are validated at once, rather than runs, and that the validation applies to a large number of additional simulation-specific quantities.

3.2. *Visualisation*

As in the case of online, the DQM results from offline processing are uploaded to the central DQM GUI server with a large disk pool. There the result files are merged to larger size and backed up to the tape; recent data is kept cached on disk for several months. The automatic certification results from the harvesting, called *quality flags*, are extracted and uploaded to the run registry. From there the values are propagated to the condition database and the dataset bookkeeping system DBS as described in section 4.

CMS provides one central DQM GUI web server instance per offline activity, including one public test instance for development. All online and offline servers provide a common look and feel and are linked together as one entity. They give the entire worldwide collaboration access to inspect and analyse all the DQM data at one central location. The GUI will offer in its final form all the capabilities needed for shift and expert use for all the DQM activities. We emphasise it is custom-built for the purpose of efficient interactive visualisation and navigation of DQM results; it is not a general physics analysis tool.

3.3. *Operation*

In all the offline processing the initial histogram production step is incorporated in the standard data processing workflows as an additional execution sequence, using components from standard software releases. The harvesting step implementation currently varies by the activity. The Tier-0 processing system has fully integrated an automated harvest step and upload to the DQM GUI. For other data we currently submit analysis jobs with the CMS CRAB tool [7] to perform the harvest step; the histogram result file is returned in the job sandbox which the operator then uploads to the GUI. This is largely a manual operation at present.

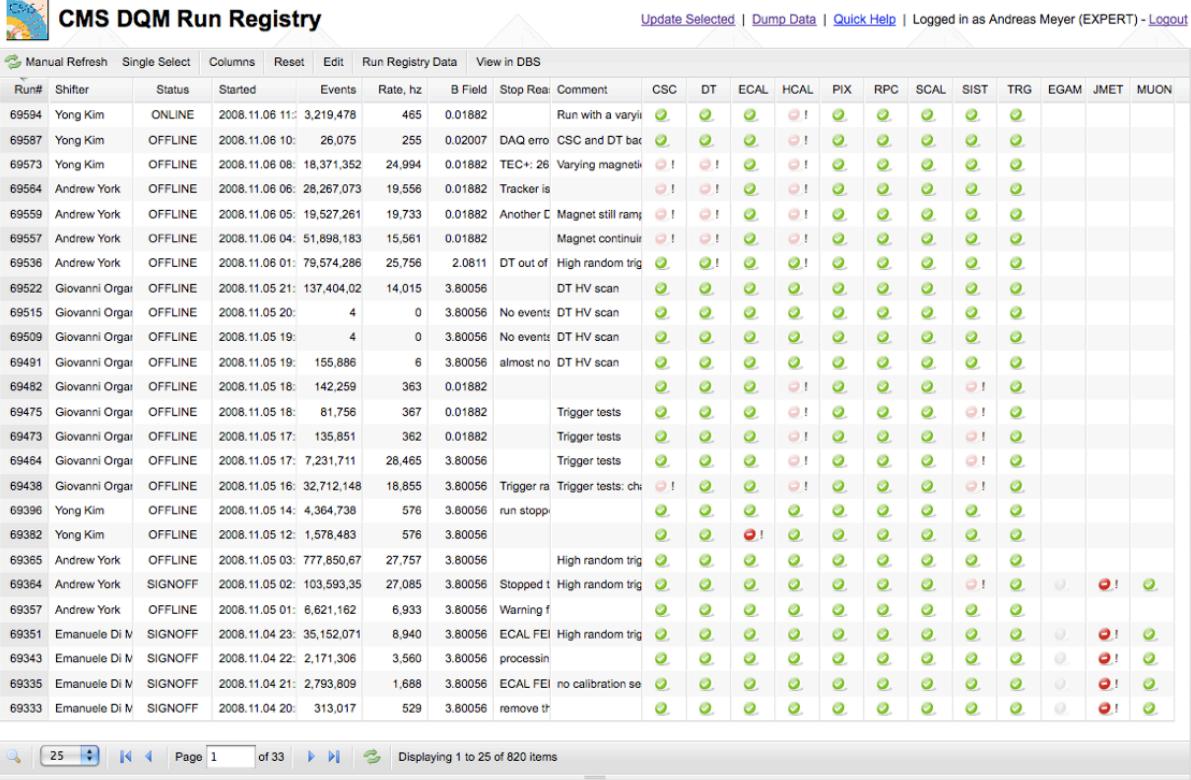
4. Certification and sign-off workflow

CMS uses a *run registry* database with a front-end web application as the central workflow tracking and bookkeeping tool to manage the creation of the final physics dataset certification result. The run registry is both a user interface managing the workflow (Fig. 3), and a persistent store of the information; technically speaking it is part of the online condition database and the web service is hosted as a part of the WBM system.

The certification process begins with the physicists on online and offline shift filling in the run registry with basic run information, and adding any pertinent observations on the run during the shift. This information is then augmented with the automatic data certification results from the online, Tier-0 and Tier-1 data processing as described in the previous sections. This results in basic major detector segment level certification which accounts for DCS, DAQ and DQM online and offline metrics, and in future may also include power and cooling status. For each detector segment and input one single boolean flag or a floating point value describes the final quality result. For the latter we apply appropriate thresholds which yield binary “good” or “bad” results. We label the result “unknown” if no quality flag was calculated.

Once the automatic certification results are known and uploaded to the run registry, the person on shift evaluates the detector and physics object quality following the shift instructions on histograms specifically tailored to catch relevant problems. This person adds any final observations to the run registry and overrides the automatic result with a manual certification decision where necessary.

The final combined quality result is then communicated to the detector and physics object groups for confirmation. Regular sign-off meetings collect the final verdict and deliver the agreed result to the experiment. At this stage the quality flags are copied to the offline condition



The screenshot shows a web-based interface for the CMS DQM Run Registry. The top navigation bar includes links for 'Update Selected', 'Dump Data', 'Quick Help', and 'Logout'. The main content is a table with the following columns: Run#, Shifter, Status, Started, Events, Rate, hz, B Field, Stop Reason, Comment, CSC, DT, ECAL, HCAL, PIX, RPC, SCAL, SIST, TRG, EGAM, JMET, and MUON. The table lists 820 items, with the first few rows shown below:

Run#	Shifter	Status	Started	Events	Rate, hz	B Field	Stop Reason	Comment	CSC	DT	ECAL	HCAL	PIX	RPC	SCAL	SIST	TRG	EGAM	JMET	MUON
69594	Yong Kim	ONLINE	2008.11.06 11: 3,219,478	465	0.01882		Run with a varyin	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD
69587	Yong Kim	OFFLINE	2008.11.06 10: 26,075	255	0.02007	DAQ erro	CSC and DT bac	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD
69573	Yong Kim	OFFLINE	2008.11.06 08: 18,371,352	24,994	0.01882	TEC< 26	Varying magnetic	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD
69564	Andrew York	OFFLINE	2008.11.06 06: 28,267,073	19,556	0.01882	Tracker is	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD
69559	Andrew York	OFFLINE	2008.11.06 05: 19,527,261	19,733	0.01882	Another C	Magnet still ramp	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD
69557	Andrew York	OFFLINE	2008.11.06 04: 51,898,183	15,561	0.01882		Magnet continuu	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69536	Andrew York	OFFLINE	2008.11.06 01: 79,574,286	25,756	2.0811	DT out of	High random trig	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69522	Giovanni Orga	OFFLINE	2008.11.05 21: 137,404,02	14,015	3.80056		DT HV scan	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69515	Giovanni Orga	OFFLINE	2008.11.05 20: 4	0	3.80056	No events	DT HV scan	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69509	Giovanni Orga	OFFLINE	2008.11.05 19: 4	0	3.80056	No events	DT HV scan	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69491	Giovanni Orga	OFFLINE	2008.11.05 19: 155,886	6	3.80056	almost no	DT HV scan	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69482	Giovanni Orga	OFFLINE	2008.11.05 18: 142,259	363	0.01882			GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69475	Giovanni Orga	OFFLINE	2008.11.05 18: 81,756	367	0.01882	Trigger tests		GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69473	Giovanni Orga	OFFLINE	2008.11.05 17: 135,851	362	0.01882	Trigger tests		GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69464	Giovanni Orga	OFFLINE	2008.11.05 17: 7,231,711	28,465	3.80056	Trigger tests		GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69438	Giovanni Orga	OFFLINE	2008.11.05 16: 32,712,148	18,855	3.80056	Trigger ra	Trigger tests: ch	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69396	Yong Kim	OFFLINE	2008.11.05 14: 4,364,738	576	3.80056	run stopp		GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69382	Yong Kim	OFFLINE	2008.11.05 12: 1,578,483	576	3.80056			GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69365	Andrew York	OFFLINE	2008.11.05 03: 777,850,67	27,757	3.80056		High random trig	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69364	Andrew York	SIGNOFF	2008.11.05 02: 103,593,35	27,085	3.80056	Stopped t	High random trig	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69357	Andrew York	OFFLINE	2008.11.05 01: 6,621,162	6,933	3.80056	Warning f		GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69351	Emanuele Di M	SIGNOFF	2008.11.04 23: 35,152,071	8,940	3.80056	ECAL FEI	High random trig	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69343	Emanuele Di M	SIGNOFF	2008.11.04 22: 2,171,306	3,560	3.80056	processin		GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69335	Emanuele Di M	SIGNOFF	2008.11.04 21: 2,793,809	1,688	3.80056	ECAL FEI	no calibration se	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	
69333	Emanuele Di M	SIGNOFF	2008.11.04 20: 313,017	529	3.80056	remove th		GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	GOOD	

Below the table, a status reference legend is provided: GOOD (green circle), BAD (red circle), NOTSET (grey circle), and EXCLUDED (yellow circle).

Figure 3. DQM run registry web interface.

database and to the dataset bookkeeping system (DBS) [8]. The flags are stored in conditions as data keyed to an interval of validity, and are meant for use in more detailed filtering in any subsequent data processing, longer-term data quality evaluation, and correlation with other variables such as temperature data. In the DBS the quality flags are used to define convenience *analysis datasets*. The flags are also accessible in the CMS data discovery interface [9], a web interface to browse and select data in the DBS (Fig. 4, “data quality” column).

Some trivial trend plots of the key monitor element metrics have recently been generated automatically. We plan to extend this to more comprehensive interactive trend plotting of any selected histogram metric, and are working on common design for convenient access and presentation of trends over time.

5. Organisation and operation

Online shifts take place 24/7 during detector operation at the CMS Point-5 in Cessy. Offline DQM shifts are carried out in day time at the CMS centre [10] on the main CERN site. The shift activities are supported by regular remote shifts, two shifts per day at Fermilab and one shift per day at DESY, at the local CMS centre [11]. Standard shift instructions, as illustrated in Fig. 5, have been fully exercised.

6. Experience

CMS has commissioned a full end to end data quality monitoring system in tandem with the detector over the last two years. The online DQM system has now been in production for about a year and the full offline chain has just been commissioned: we have recently completed the first

Run	Type	Events	Store	Start of run	End of run	dataset	# Files	Size	LFNs	Transfer	Data Quality	Run Summary
70687	data	0	0	1226390992	0	/Calo /Commissioning08-PromptReco-v2 /RECO	1	3.0MB	cff plain details cff wParents py wParents	cmssrm.fnal.gov srm-cms.cern.ch gridka-dCache.fzk.de	<input checked="" type="checkbox"/> -Pixel_Global <input checked="" type="checkbox"/> -L1T_Global <input checked="" type="checkbox"/> -JetMET <input checked="" type="checkbox"/> -Muon	
						/Calo /Commissioning08-v1 /RAW						
						/Cosmics /Commissioning08-PromptReco-v2 /RECO						
						/Cosmics /Commissioning08-v1 /RAW						

Figure 4. DBS discovery page displaying quality flags.

full cycle of certification and sign-offs. DQM for the less structured alignment and calibration activity at the CAF exists but a fair amount of work remains.

In our experience it takes approximately one year to commission a major component such as online or offline DQM to production quality. Shift organisation, instructions, tutorials and supervision are major undertakings. Significant amounts of effort are needed in various groups to develop the DQM algorithms, and centrally to standardise and integrate the workflows, procedures, code, systems and servers. While we find only modest amounts of code are needed for the core DQM systems, on the balance there is a perpetual effort to optimise histograms to maximise sensitivity to problems, to standardise the look and feel and to improve efficiency through better documentation, and a battle to promote sharing and use of common code against natural divergence in a collaboration as large as CMS.

CMS has so far focused on commissioning a common first order DQM system throughout the entire experiment, with the aim of having an effective basic system ready for the first beam. We believe we have successfully achieved this goal and will address second order features in due course.

CMS is very pleased with the DQM visualisation served using web technology and operating shifts from the CMS centres. Remote access to all the DQM information, especially offsite real-time live access to the online as been greatly beneficial and appreciated. Together the CMS centres and remote access have been essential and practical enabling factors to the experiment.

Acknowledgments

The authors thank the numerous members of CMS collaboration who have contributed the development and operation of the DQM system. Effective data quality monitoring is a truly

List of Shift Histograms

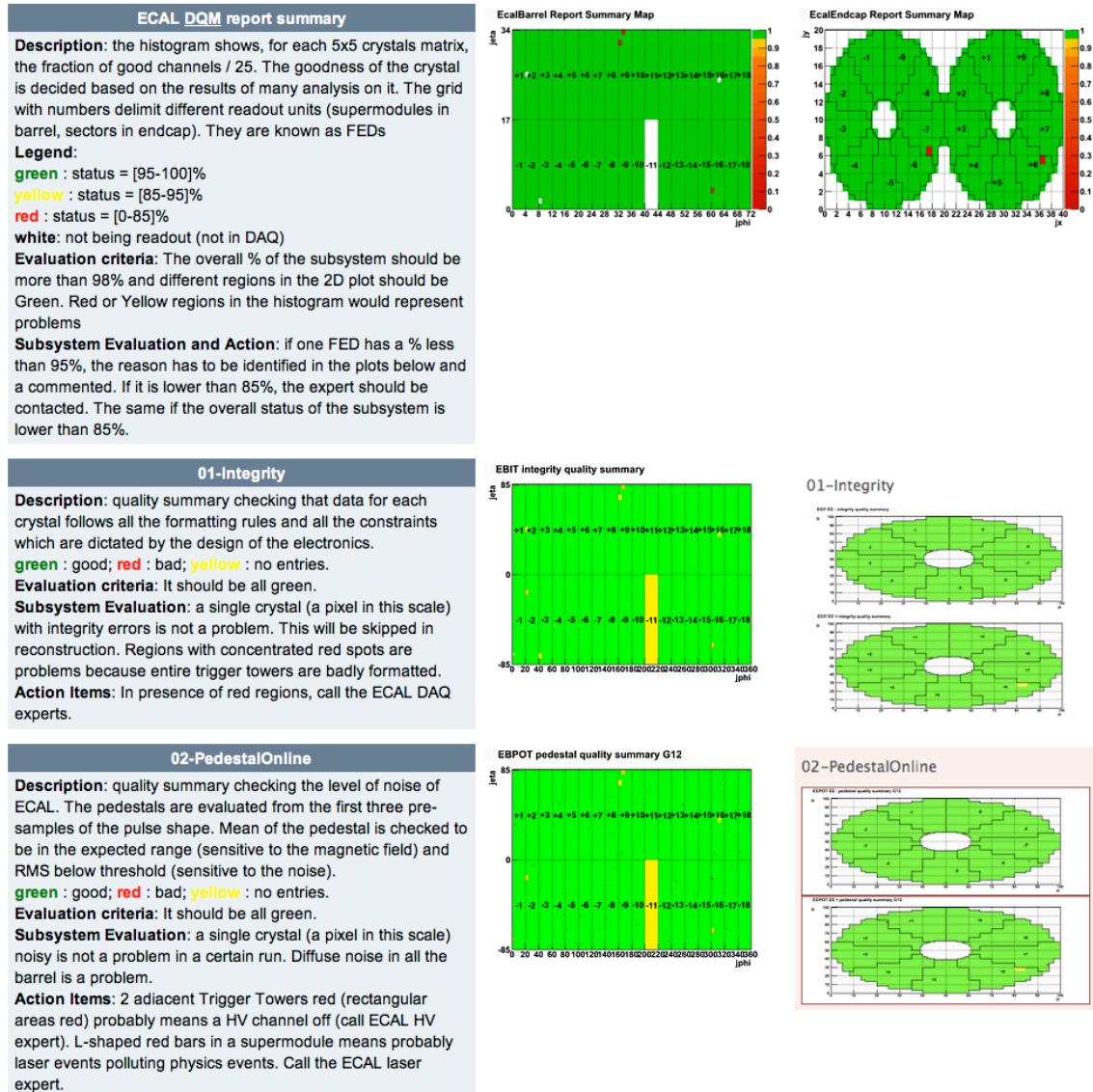


Figure 5. Example shift instructions.

collaborative effort involving a lot of people from several other projects: the trigger, detector subsystems, offline and physics software, production tools, operators, and so on.

References

- [1] CMS Collaboration, 1994, *CERN/LHCC 94-38*, “Technical proposal” (Geneva, Switzerland)
- [2] Tuura L, Eulisse G, Meyer A, 2009, *Proc. CHEP09, Computing in High Energy Physics*, “CMS data quality monitoring web service” (Prague, Czech Republic)
- [3] Kreuzer P, Gowdy S, Sanches J, et al, 2009, *Proc. CHEP09, Computing in High Energy Physics*, “Building and Commissioning of the CMS CERN Analysis Facility (CAF)” (Prague, Czech Republic)
- [4] Brun R, Rademakers F, 1996, *Proc. AIHENP'96 Workshop*, “ROOT – An Object Oriented Data Analysis Framework” (Lausanne, Switzerland); see also <http://root.cern.ch>
- [5] Bauer G, et al, 2007, *Proc. CHEP07, Computing in High Energy Physics*, “The Run Control System of the CMS Experiment” (Victoria B.C., Canada)

- [6] Badgett W, 2007, *Proc. CHEP07, Computing in High Energy Physics*, “CMS Online Web-Based Monitoring and Remote Operations” (Victoria B.C., Canada)
- [7] Spiga D, 2007, *Proc. CHEP07, Computing in High Energy Physics*, “CRAB (CMS Remote Anaysis Builder)”, (Victoria B.C., Canada)
- [8] Afaq A, Dolgert A, Guo Y, Jones C, Kosyakov S, Kuznetsov V, Lueking L, Riley D, Sekhri V, 2007, *Proc. CHEP07, Computing in High Energy Physics*, “The CMS Dataset Bookkeeping Service” (Victoria B.C., Canada)
- [9] Dolgert A, Gibbons L, Kuznetsov V, Jones C, Riley D, 2007, *Proc. CHEP07, Computing in High Energy Physics*, “A multi-dimensional view on information retrieval of CMS data” (Victoria B.C., Canada)
- [10] Taylor L, Gottschalk E, 2009, *Proc. CHEP09, Computing in High Energy Physics*, “CMS Centres Worldwide: a New Collaborative Infrastructure” (Prague, Czech Republic)
- [11] Gottschalk E, 2009, *Proc. CHEP09, Computing in High Energy Physics*, “Collaborating at a Distance: Operations Centres, Tools, and Trends” (Prague, Czech Republic)