

DQM online cluster processes the live data and produces histograms that are used to identify problems during the run

- ✓ runs DQM framework, which is part of CMS software framework
- ✓ all detector subsystems have their own applications, which rely on the DQM framework
- ✓ the subsystems produce real-time histograms that are transferred to the DQM GUI over the local network
- ✓ histograms are periodically archived as ROOT files for a continuous backup
- ✓ new software is deployed on demand twice a week so all the subsystems keep their software up to date
- ✓ to ensure stability, all software changes are verified in an online replica cluster for a 24hr period

Improvements:

- ✓ full re-implementation of the Data Acquisition System
- ✓ online framework publishing capabilities
- ✓ improvements of the monitoring system of the online cluster

Online archive area collects the ROOT files periodically produced by the online applications

- ✓ each ROOT file has several time-based versions to recover from potential crashes

Improvements:

- ✓ archive the ROOT files before collecting them to save up to 50% additional space
- ✓ automatic clean up of the area to avoid interruptions

The online DQM GUI collects, stores, and visualizes the results from the online applications

- ✓ independent customizations of the detector subsystems (see the offline GUI)
- ✓ flexible schema for the deployment of new software
- ✓ DQM team maintains a replica setup where all the code is tested before it reaches production

Real-time alarm at the CMS detector is used to indicate critical problems during data taking

- ✓ maximizes good quality data during data-taking
- ✓ a dedicated location in the GUI maps all the critical problems found
- ✓ if a critical problem occurs, the DQM GUI triggers an audio alarm for the shift crew

Improvements:

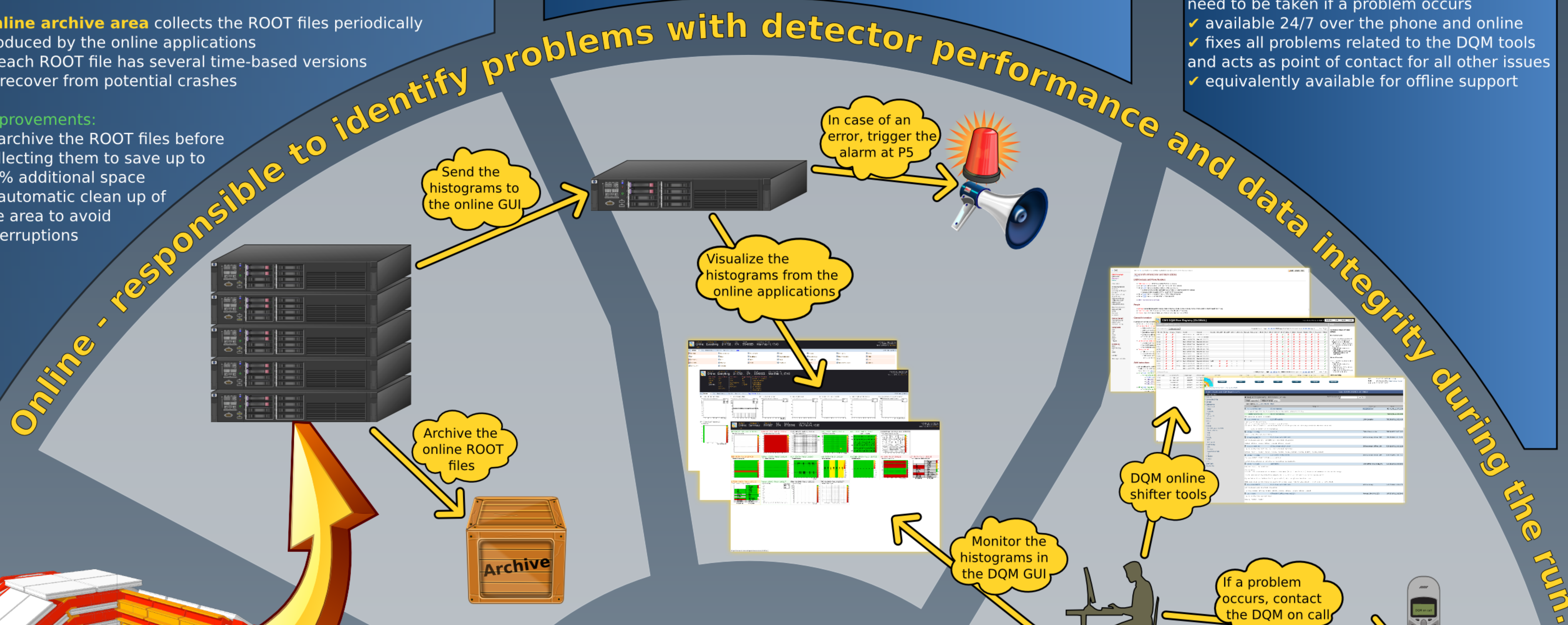
- ✓ easy suppression of the individual errors without interaction from a DQM expert

The scope of **the DQM online shift** is to identify problems with detector performance and data integrity during the run

- ✓ shifts take place at the CMS detector
- ✓ dedicated DQM shifts are scheduled 24/7 in data taking periods
- ✓ 3 shifts per day, each takes 8 hours
- ✓ the online shifter monitors a subset of histograms for every detector and reconstruction objects in the DQM GUI
- ✓ the tools of the online DQM shifter are: the DQM GUI, the twiki pages, online Run Registry and the ELog

Twiki Pages, Run Registry, ELog (find the definitions in the offline chain)

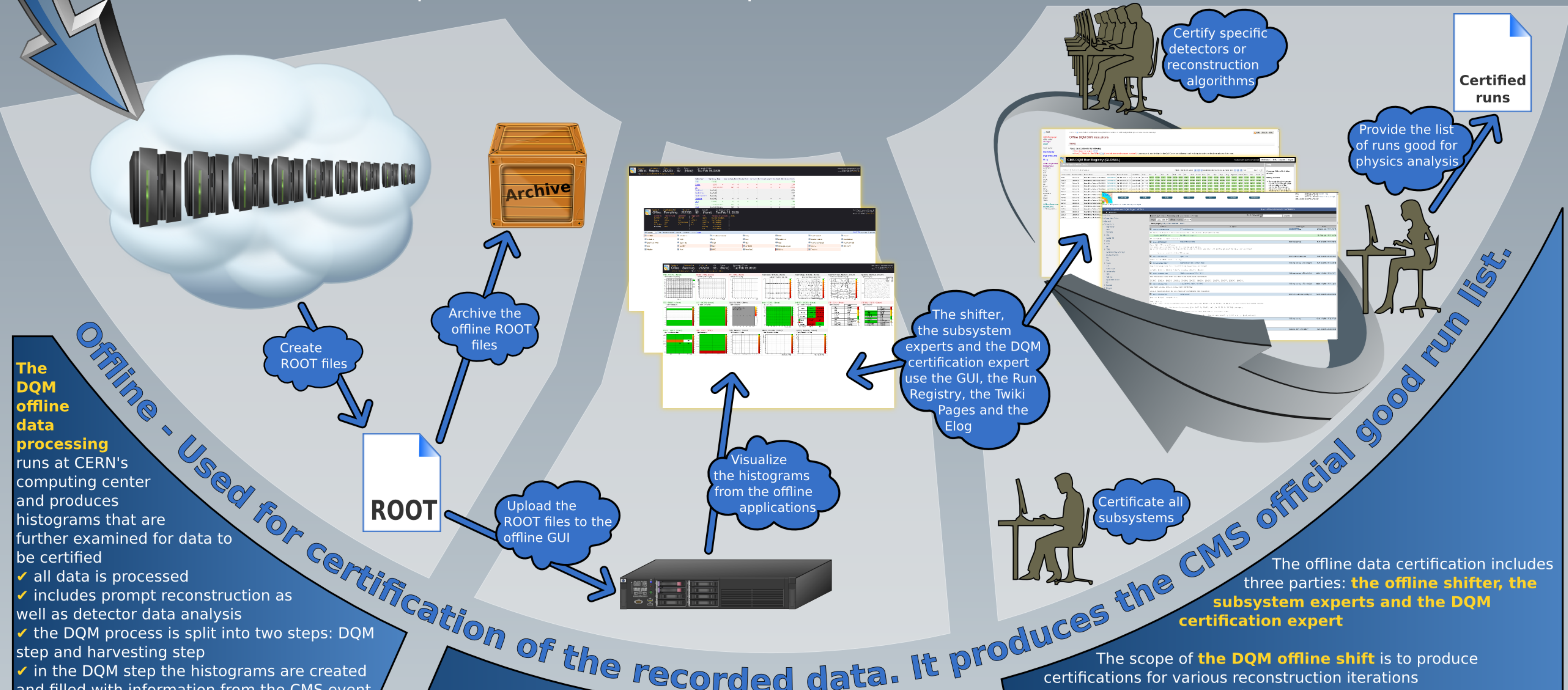
- DQM on call** coordinates the actions that need to be taken if a problem occurs
- ✓ available 24/7 over the phone and online
 - ✓ fixes all problems related to the DQM tools and acts as point of contact for all other issues
 - ✓ equivalently available for offline support



The CMS Data Quality Monitoring software: experience and future improvements

The Data Quality Monitoring (DQM) Software proved to be a central tool in the CMS experiment. Its flexibility allows its integration in several environments: Online, for real-time detector monitoring; Offline, for the final, fine-grained Data Certification. The central tool to deliver Data Quality information is a web site for browsing data quality histograms (DQM GUI). In this contribution the usage of the DQM Software in different environments and its integration in the CMS Reconstruction Software Framework (CMSSW) and in all production workflows are presented.

The Large Hadron Collider (LHC) at CERN smashes protons together at close to the speed of light. Some of the collision energy is turned into mass, creating new particles which are observed in the Compact Muon Solenoid (CMS) particle detector. CMS data is analyzed by physicists around the world to reconstruct a picture of what happened at the heart of these collisions. Data quality monitoring (DQM) is a crucial part of the experiment. Its purpose is to identify errors and problems in the detector hardware or reconstruction software. The ultimate goal is a stable detector leading to high quality reconstructed collision events. There are two main branches: **ONLINE** and **OFFLINE**



The DQM offline data processing runs at CERN's computing center and produces histograms that are further examined for data to be certified

- ✓ all data is processed
- ✓ includes prompt reconstruction as well as detector data analysis
- ✓ the DQM process is split into two steps: DQM step and harvesting step
- ✓ in the DQM step the histograms are created and filled with information from the CMS event data
- ✓ in the harvesting step the histograms are extracted from the event data files and summed together across the entire run to yield full event statistics
- ✓ the final results are output to ROOT files and uploaded to the offline GUI and archived

Improvements:

- ✓ CMS Framework switches to multi-core, multi-thread data processing

The offline DQM GUI collects, stores, and visualizes the results from the offline applications

- ✓ web based, OS independent, accessible from everywhere around the world
- ✓ uses key-value database called 'index' to store the serialized histogram data
- ✓ ROOT is used to render the data from the index to histograms
- ✓ shares the same code as the online GUI but looks different
- ✓ online and offline GUIs have their own set of workspaces, layouts and render plug-ins that define their unique look
- ✓ users from the detector subsystems can be defined/developed different workspaces, layouts and render plug-ins
- ✓ has rich web API that supports JSON format

Improvements:

- ✓ database key length grows from 64 to 128 bit
- ✓ create database for every new month of data taking
- ✓ improve server-client communication using JavaScript
- ✓ possible migration to new DB like Kyoto cabinet or leveldb
- ✓ move the rendering phase from the server to the client (users' browser)

- The scope of **the DQM offline shift** is to produce certifications for various reconstruction iterations
- ✓ operate at Meyrin, FNAL and DESY
 - ✓ 4 shifts per day, 6 hours each
 - ✓ offline DQM shifts can be scheduled outside of global data taking periods
 - ✓ the offline shifter monitors a subset of histograms for every detector and reconstruction objects in the DQM GUI
 - ✓ the tools of the DQM offline shifter are: the DQM GUI, the twiki pages, offline Run Registry and the Elog

Twiki Pages - describe how to examine the individual histograms in the GUI

Run Registry - web-based bookkeeping tool that collects all the significant runs and the data certifications

ELog - central logging system where all parties can leave comments, report issues, etc ...

The experts from the subsystems produce certifications for specific detectors or reconstruction algorithms

The DQM certification experts combine the results from the offline shifters and the subsystem experts to provide the list of runs good for physics analysis

References

- [1] The CMS Data Quality Monitoring software: experience and future improvements, F. De Guio et al, CHEP'13
- [2] CMS data quality monitoring: systems and experiences, L. Tuura et al, CHEP'09

Atanas Batinkov on behalf of the CMS collaboration

