Performance quality monitoring system for the Daya Bay reactor neutrino experiment

Y B Liu on behalf of the Daya Bay collaboration

Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China

E-mail: `liuyb@ihep.ac.cn`

Abstract. The Daya Bay Reactor Neutrino Experiment started running on September 23, 2011. The Performance Quality Monitoring system (PQM) has been developed to monitor the detector performance and data quality. The main feature is its ability to efficiently process multi-data-stream from three experimental halls. The PQM processes raw data files from the Daya Bay data acquisition system, generates and publishes histograms via a graphical web interface by executing the user-defined algorithm modules, and saves the histograms for permanent storage. That the whole process takes only around 40 minutes makes it valuable for the shift crew to monitor the running status of all the sub-detectors and the data quality.

1. Introduction

The Daya Bay experiment is designed to determine precisely the neutrino mixing angle θ_{13} with a sensitivity better than 0.01 in the parameter $\sin^2 2\theta_{13}$ at 90% confidence level. Three experimental halls (EHs) were built to contain eight functionally identical antineutrino detectors (ADs) [1]. Before October 18, 2012, two ADs were located in EH1 and one in EH2 (the near halls). Three ADs were positioned in the far hall, EH3. After the installation of the last two ADs, 2 and 4 ADs are located in each near hall and the far hall, respectively. In each experimental hall there are also muon detectors, including the inner (IWS), outer (OWS) water shields, and a resistive plate chamber (RPC) tracker. With the 6-AD configuration, the Daya Bay collaboration reported the first observation of non-zero $\sin^2 2\theta_{13}$ [2], and has recently updated the result of $\sin^2 2\theta_{13} = 0.089 \pm 0.010(\text{stat.}) \pm 0.005(\text{syst.})$ [3] in a three-neutrino framework.

During physics data taking, events from different experimental halls are separately assembled in raw data format and stored into persistent files on the online disk array by the data acquisition system (DAQ). Event rates of 8-AD configuration for EH1, EH2, EH3 are about 1.2 kHz, 1.0 kHz, and 600 Hz, respectively. The total number of raw data files for each day is about 320, approximately 1 GB volume per file. The online histogram presenter of the DAQ provides the hit-maps for sub-detectors. In addition, high-level histograms are necessary to be produced for monitoring the running status of all sub-detectors and the data quality (DQ).

The Performance Quality Monitoring system (PQM) is developed to utilize the offline software to process multi-data-stream from three experimental halls, generates more user-defined histograms, including some high-level ones containing the calibration and reconstruction information, and thus plays an important role in offline monitoring for the experiment. Even with the total event rate of around 3 kHz, the PQM can handle the whole process around 40 minutes.

In this article, we present the offline computing environment at Daya Bay in Section 2. In Section 3, the PQM design is outlined. The performance and the information display of the PQM are shown in Section 4. We give a short summary in the last section.

2. Offline computing environment at Daya Bay

There are 11 servers in the offline computing environment at Daya Bay, including a file server, a data transfer server, an offline database server, a web server, two PQM servers (one for backup), and the rest of 5 servers forming user farms. The Portable Batch System (PBS) is implemented to allocate computational jobs among the available computing resources.

The file server provides disk array for onsite data processing. To prevent interruptions to the DAQ, the data transfer server is responsible for copying raw data files to the offline file server. The offline database server copies the information that is necessary for offline analysis from the online database to the offline database. The PQM server is responsible for the implementation of the functionality of the PQM for monitoring the detector performance and the DQ. This will be described in detail in Section 3. The web server is used to display user-defined histograms in analysis algorithm modules.

The Daya Bay offline software (NuWa) [1] using Gaudi [4] as the underlying software framework is available in the offline computing environment. In this framework, reconstruction algorithms have been developed to reconstruct the energy and the vertex of the antineutrino event from the charge pattern of the PMTs. One can find the details of the calibration and reconstruction methods in [1]. By default, the latest calibration constants in the offline database are loaded into the reconstruction algorithms, reflecting the best understanding of the detectors.

3. Implementation of the PQM

The PQM is designed to efficiently process all events in all raw data files as soon as possible. It communicates with the offline database to get the latest raw data file list and submit jobs to the PBS to process them in parallel. At the end of each job, some selected histograms are printed for web display. The shift crew can compare the new histograms with standard ones to report the sub-detector status and the DQ.
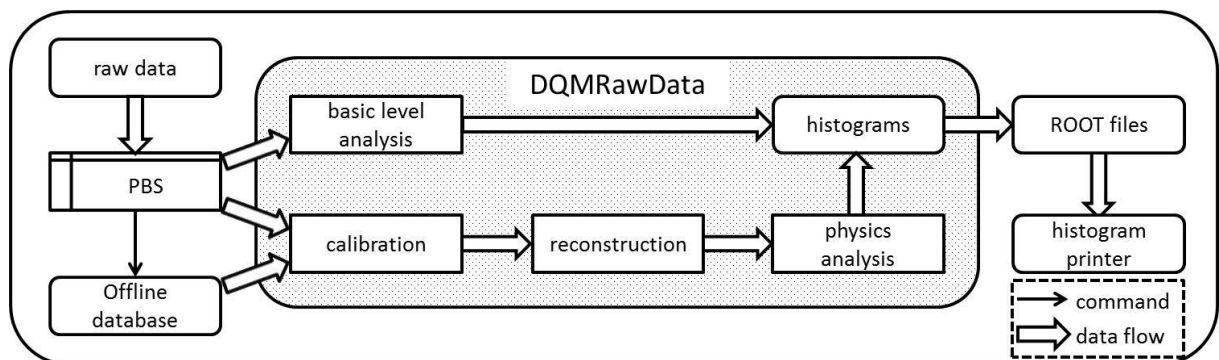

Figure 1. The data flow of the PQM job.

3.1. Data processing chain

To manage the data flow of the PQM as shown in figure 1, a control script in Python language has been developed and runs in a background mode on the PQM server. When the DAQ closes a raw data file, the offline database will be updated with a new record copied from the

online database. A copy of the file will be transferred to the file server in less than 5 minutes, assuming a size of ($\sim$1 GB). When the data transfer is completed, the status of this file is set as TRANSFERRED in both online and offline database.

The control script queries the offline database at a fixed time interval (10 seconds). If a new record is found and the data status returned by the query shows that the corresponding raw data file is already on the file server, a job for processing the new file is submitted to the PBS. The PQM job uses a recently manually tagged NuWa version to reconstruct data with the latest calibration constants in the offline database, runs analysis algorithms, and fills the results into user-defined histograms. All of the histograms are dumped into a ROOT [5] file at the finalization step of the algorithms, which is then merged with the accumulated ROOT file for the same run. After that, a histogram printer in C++ language is executed taking the merged ROOT file as input to print selected histograms into figures. The control script also checks if the job is finished at a 10-second interval by detecting an empty TXT file created at the end of the job. For finished jobs, the control script transfers the corresponding figures to the PQM disk for web display (see Section 4). When the run ends and all the corresponding raw data files are processed, the control script saves the accumulated ROOT file on the PQM disk for permanent storage.

3.2. Analysis modules in PQM

Currently four modules are developed for the PQM to analyze data and fill user-defined histograms in the NuWa framework: one for histograms of electronics channel information, one for the histograms of calibrated PMT information of ADs and the water shields, one for the histograms of the reconstruction for the ADs, and the other one for the histograms of the RPCs. The histograms displayed by the web interface are shown in table 1. The detailed description of the variables listed in table 1 can be found in [1, 6, 7]. It provides the flexibility to add more algorithm modules as well as to fill more user-defined histograms.

Table 1. The histograms produced by analysis modules in PQM for monitoring the detector performance and the DQ.

Detector Unit	Level	Histograms
PMT	Basic	Mean of ADC, TDC and preADC, RMS of ADC, TDC and preADC, ΔADC, dark noise, dark rate and TDC vs channel ID, hit rates, ADC sum vs. trigger types, ADC sum, number of blocked triggers vs. run time, number of channels, trigger types
	Reconstruction	y vs. x, z vs. radius, energy, energy vs. radius, energy vs. z, event rate vs. radius, event rate vs. z
RPC	Basic	Map of patch rate, map of module rate, map of channel multiplicity, distributions of time between module and system triggers, maps of layer efficiency, distribution of layer efficiency, electronics error types vs. time, maps of channel hit count, channel hit count vs. channel ID, distribution of number of layers per readout, distribution of number of modules per system trigger, maps of layer singles rate, distribution of layer singles rate, module and system trigger rates vs. time.

Since a different configuration of detectors is adopted for EH3, all the analysis modules take the strategy of creating histograms dynamically for each detector, and even for each PMT channel to reduce the output file volume and to improve the processing efficiency. If one event read from the raw data file is the first one for the current detector, a new set of histograms for this detector are created. Otherwise, the event is analyzed and filled into existed histograms.

4. Performance and information display

The PQM is required to provide more information about the running status of all sub-detectors in 3 EHs than the online histogram presenter provided by the DAQ. The data transfer from the DAQ takes less than 5 minutes. The job can be submitted to the PBS about 10 seconds after the data transfer is finished. Running NuWa job modules takes around 30 minutes. In the meanwhile, some user-defined histograms produced by job modules are saved in ROOT files. They are printed and published via the web interface in around 2 minutes. The job failure rate is less than 1% during nomarl data taking. The typical latency of the display of the PQM plots is 40 minutes after a new raw data file is closed by the DAQ. The most time-consuming part is running the calibration and reconstruction algorithms. Since all algorithms are configurable in the PQM job, it is easy to remove them and only show the basic-level information of the detectors to reduce the latency to less than 20 minutes. This special configuration of PQM jobs was usually implemented during the commissioning period, since the onsite experts could see the detector status as soon as possible.

A web interface, as shown in figure 2, running on the web server is utilized to display the selected user-defined histograms. On the main page (see figure 2(a)), two options are provided for users: one is for showing histograms for the current runs of each site, named as 'realtime', the other one is for all the previously processed runs, named as 'history'. For the latter, one can choose the run number in a drop-down menu listing the site ID and run type, i.e. 'Physics', 'FEEDiag' and so on. On the 'realtime' page (see figure 2(b)), site buttons are available for different sites: 'DayaBay', 'LingAo' and 'Far'. The 'realtime' page refreshes at a fixed time interval of 1 minute. For each site, the user can also use detector buttons to view histograms for different sub-detectors. On the 'history' page, there are only detector buttons, because this page is for the run selected by the user, and the specific run is for only one site.

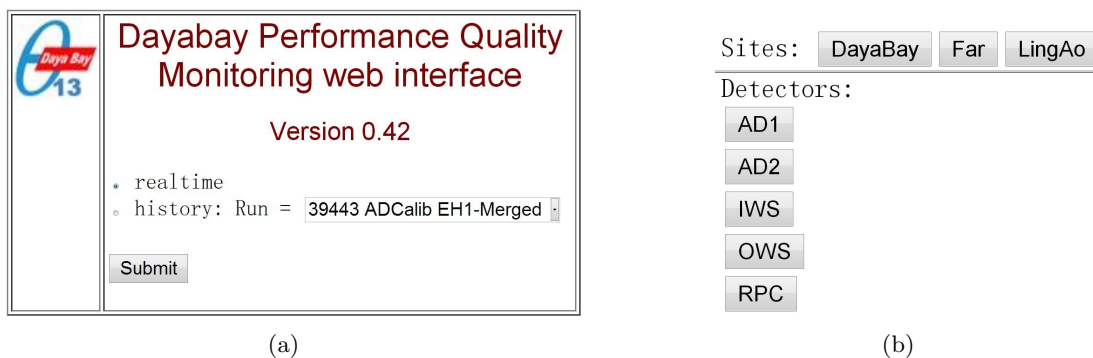(a)   (b)

Figure 2. Web interface for the PQM.

Several example histograms are given here. Figure 3 shows the number of blocked triggers in one event as a function of time. Generally, if a cosmic-ray muon passes through the detector and induces a shower with extremely large energy deposit, overflow of the electronics occurs, and the local trigger and the corresponding data package can be blocked. The blocked trigger number is recorded and read out. If the fraction of the blocked trigger number over the total trigger
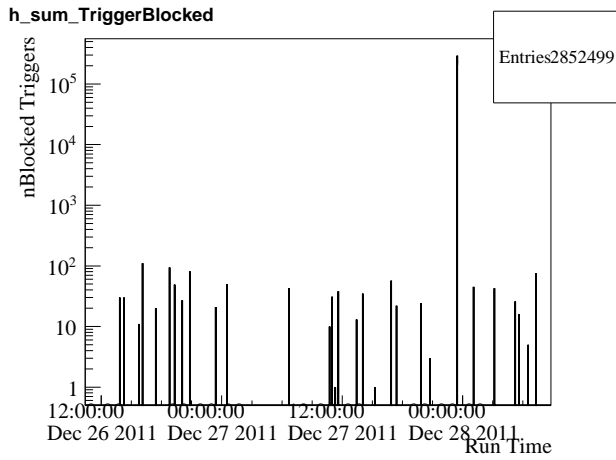
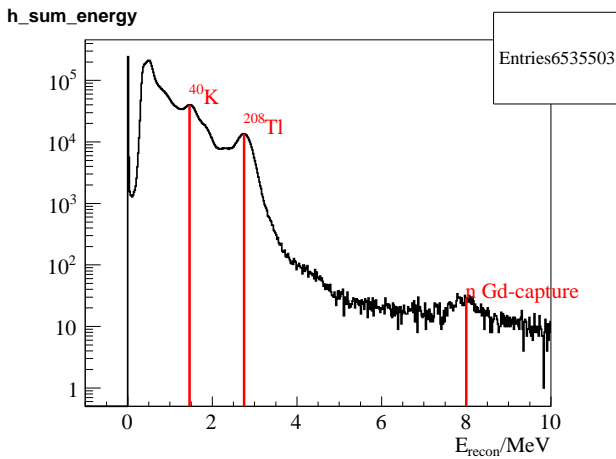**Figure 3.** The number of blocked triggers in one event vs. run time.

Figure 4. The reconstructed energy distribution of all events for AD1.
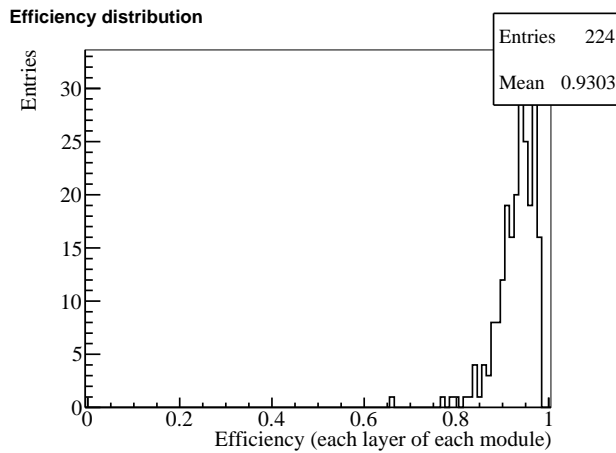

Figure 5. The efficiency of each layer for the RPCs.

number is larger than a pre-defined value, the shift crew should report to the DQ working group (DQWG).

Figure 4 presents the reconstructed energy spectrum for all triggers for an AD. The peaks of ^{40}K (1.46 MeV), ^{208}Tl (2.61 MeV), and the gamma peak from neutron capture on Gd ($\sim$8 MeV) are labeled in fixed positions in the histogram. The red line for ^{208}Tl is set to $\sim$2.75 MeV due to the energy non-linearity which is not corrected in the analysis algorithm. A relative

shift between the red lines and the observed peaks indicates problems in the data, e.g. incorrect energy scales read from the offline central database or unexpected behaviour of the PMTs.

The statistics of the efficiency of each RPC layer [6, 7] is displayed in figure 5. The efficiency is calculated in the algorithm of analysis modules for the PQM. The efficiency is a critical quantity to evaluate the performance of the RPCs used as an anti-coincidence detector for muons for the experiments. As shown in the figure, most of RPC layers are working with the efficiency of about 90%. The performance of the RPCs can change due to the gas pressure or other unexpected reasons. The RPC experts and the DQWG are informed if the RPC layer efficiency distribution distorts greatly by comparing with the standard ones.

There are also many other histograms to monitor the detector performance and the DQ. The shift crew can compare them with the corresponding standard figures to check the running status of the sub-detectors and the DQ. In the case that irregularities are observed, the shifter should inform the related experts and report to the DQWG.

A supernova monitoring program is under development to join the Supernova Early Warning System (SNEWS) [8], in which individual neutrino-sensitive experiments send supernova (SN) burst datagram to SNEWS coincidence computer at Brookhaven National Laboratory (BNL). Supernova burst neutrinos are important in studying the dynamics of core collapse and supernova explosion, neutrino properties, and many other interesting studies. The three-hall configuration design at Daya Bay can significantly reduce false SN alarms caused by the spallation background. We will have online and offline triggers for monitoring SN bursts. Given the short latency time of the PQM, the offline trigger for SN bursts will be added to the PQM job.

5. Summary
The onsite data processing is running smoothly. The PQM, which has been developed for the Daya Bay experiment, has the ability to provide user-defined histograms to the shift crew with a latency time of about 40 minutes. It satisfies the requirement for the data taking. By comparing the histograms with the standard ones, potential DQ problem can be found. The PQM will contribute to the SN monitoring program in the Daya Bay experiment in the near future.

References
[1] Daya Bay collaboration 2012 *Nucl. Instrum. Methods* A **685** 78
[2] Daya Bay collaboration 2012 *Phys. Rev. Lett.* **108** 171803
[3] Daya Bay collaboration 2013 *Chinese Physics* C **37** 011001
[4] Gaudi http://cern.ch/gaudi/
[5] ROOT http://root.cern.ch/drupal/
[6] Zhang Q M, He M, Xu J L, Zou J H, Ning Z, Zhao J, Lu H Q, Pěč V, Lebanowski L 2012 *J. Phys.: Conf. Series* **396** 022061
[7] Ning Z et al. 2013 *J. Instrum.* **8** T03007
[8] Wei H Y 2013 *Supernova Trigger in the Daya Bay Reactor Neutrino Experiment* 15th Int. Ws. on Advanced Computing and Analysis Techniques in Physics Research (Beijing, China)