Meeting summary for DQM-Nearline (04/17/2025)

Quick Recap:

The meeting focused on discussing and developing a unified Data Quality Monitoring (DQM) system for the experiment, with various team members presenting their current approaches and tools. The team explored different options for integrating existing subsystem modules into a common framework, considering factors such as visualization methods, data storage, and user-friendliness. They agreed to implement an integrated DQM system while continuing to develop and compare alternative solutions, with the goal of having a functional system ready for commissioning.

Next Steps:

- **Michael to provide a working example of integrating subsystem code into the integrated DQM system and post it on the online DQM wiki.**
- **Subsystem teams to implement their DQM code into the integrated DQM system presented by Gianni and Michael.**
- **Mark to continue developing and testing the ROOT browser version of DQM.**
- **All teams to test both the integrated DQM system and their standalone versions.**
- **DQM group to reconvene after initial testing to discuss results and make further decisions on the DQM system.**
- **Gianni and Michael to assist subsystem teams with debugging and integration of their DQM code into the integrated system.**
- **All teams to work towards having a functional DQM system ready for the Mu2e commissioning.**

Topics Discussed:

Online Data Quality Monitoring Discussion
Kamal initiated a discussion on online data quality monitoring (DQM), highlighting the need for a collaborative framework among independent systems. The meeting will focus on real-time assessments of data quality, with contributions from experts from the different subsystems. Kamal also mentioned the importance of addressing the interface between online and offline DQM.

DQM Integration and Wiki Development

Kamal discussed the progress on integrated DQM, acknowledging contributions from Gianni and Michael. He emphasized the importance of documenting information in a wiki page and invited team members to add content as they move forward. The wiki will serve as a resource for future publications and discussions.

DQM System Overview and Requirements
Giani outlined the DQM systems, emphasizing the Grafana Web Display for time series metrics and the ART analyzer for generating root histograms. He mentioned the integration of metrics into Grafana and the decision to consolidate analyzer modules into a single DQM package. Additionally, he addressed the scheduling of DQM jobs and the necessity for separate data loggers for luminosity and trigger data.

DQM Configuration and Dashboard Development
Giani outlined the approach for DQM configuration, advocating for its integration within existing analyzer modules to avoid unnecessary repository duplication. He proposed using a JSON file for scheduling sequences and storing DQM configurations in MongoDB. Additionally, he highlighted the responsibility of each subsystem to develop their Grafana dashboards, focusing on collecting requests for data rates from various streams.

Discussion on Metrics Management and System Usage
Ryan Rivera explained that Grafana is a temporary solution for data visualization, while Epics is the approved system for storing persistent records. Andrew Edmonds highlighted the flexibility of adding metrics in the analyzer module, which can be directed to either Grafana or Epics depending on whether the data needs to be stored long-term. Rivera also mentioned that metrics can be dynamically turned on or off based on operational requirements.

CaloDQM Development Update
Paolo Girotti discussed the current status of the CaloDQM project, mentioning the online dispatcher and the offline DQM developed by Mark. He emphasized the importance of automating the data visualization process and proposed that the online and offline modules could potentially be combined for easier maintenance.

* **Future steps for automating processes and improving data access for monitoring.**

Discussion on DQM Integration and Development
D.J. Tedeschi discussed the integration of offline systems with online DQM processes, suggesting that offline automation could enhance the production of histogram files. Kamal called for focused discussions on the online DQM and the need to clarify user requirements for the tools. Giani raised concerns about understanding the intended users and usage of the offline tool being developed.

CRV Online DQM and ARTAC Monitoring Overview
Sam provided an overview of the CRV online DQM, detailing the architecture of the ARTAC online monitoring system. He highlighted the role of the dispatcher in managing data flow from

loggers to various art processes, which can operate independently. Additionally, he mentioned the use of a web server for visualizing data and the flexibility in customizing the web interface.

Discussion on DQM Architecture and Integration
Sam Grant raised the topic of DQM architecture, highlighting the importance of collaboration in deciding between a general tool and an expert-focused system. D.J. Tedeschi contributed by proposing that the system could function as a near-line solution for detailed file analysis. Andrew Edmonds then shared insights on their DQM art module, which aims to integrate with existing systems while maintaining flexibility.

* **The importance of user-friendly visualization tools for both experts and non-experts.**

DQM Functionality and User Experience
Stefano Miscetti described the DQM by the CRV as a powerful, standalone tool that enables users to access and compare data easily. Ryan Rivera pointed out that while experts will manage the framework setup, users should have the ability to customize their views and save them for personal or public use. The goal is to provide a user-friendly interface that allows for quick configuration changes.

Discussion on DQM-Nearline System Integration
Mark Vakulenko proposed enhancing the browser's interactivity, allowing users to change and execute code. Sam Grant explained that server access requires permissions, but users can view server activity via a VNC server. Ryan Rivera suggested prioritizing web security later, while Kamal noted the ongoing development requiring Kerberos for access.