

Muon g-2 reconstruction and analysis framework for the muon anomalous precession frequency

K S Khaw on behalf of the Muon g-2 collaboration

University of Washington, Department of Physics, Box 351560, Seattle, WA 98195, USA

E-mail: khaw84@uw.edu

Abstract. The Muon g-2 experiment at Fermilab, with the aim to measure the muon anomalous magnetic moment to an unprecedented level of 140 ppb, has started beam and detector commissioning in Summer 2017. To deal with incoming data projected to be around tens of petabytes, a robust data reconstruction and analysis chain based on Fermilab's *art* event-processing framework is developed. Herein, I report the current status of the framework, together with its novel features such as multi-threaded algorithms for online data quality monitor (DQM) and fast-turnaround operation (nearline). Performance of the framework during the commissioning run is also discussed.

1. Introduction

The Muon g-2 experiment [1] at Fermilab aims to measure the muon anomalous magnetic moment, a_μ to an unprecedented precision of 140 ppb after several years of data taking. In the experiment, a longitudinally-polarized 3.1 GeV/c μ^+ beam will be delivered by an accelerator complex (after upgrading and modifying the existing antiproton ring) at Fermilab. This beam will be injected into the muon storage ring, which is a superconducting dipole magnet, through a superconducting inflector magnet. After traversing one-quarter of the ring, the beam is “kicked” by three pulsed magnetic kickers onto the intended storage orbit. There are four electrostatic quadrupoles installed symmetrically around the ring to focus the beam vertically. Due to the $B = 1.45$ T dipole magnetic field in the ring, the muon will undergo cyclotron motion with frequency ω_c and the spin of the muon will precess in the magnetic field with frequency ω_s . The difference between the two is the anomalous precession frequency, $\omega_a = \omega_s - \omega_c$, and it can be extracted by fitting the time modulation of the number of positrons (coming from muon decay) above an energy threshold detected by the calorimeters [2, 3] installed around the inside of the ring, with a fit function $N(t) = N_0 e^{-t/\gamma\tau_\mu} [1 + A \cos(\omega_a t + \phi)]$. Here, N_0 is the normalization factor, $\gamma\tau_\mu$ the dilated lifetime of the muon, A the average asymmetry of the decay and ϕ the initial phase. About 400 NMR probes are installed around the ring to measure the magnetic field around the ring. The field is expressed in terms of free proton's Larmor frequency in the field after convoluted with the muon beam distribution, $\tilde{\omega}_p$. a_μ can then be extracted using the formula $a_\mu = \frac{g_e \omega_a}{2} \frac{m_\mu}{\omega_p} \frac{m_\mu}{m_e} \frac{\mu_p}{\mu_e}$, where $\frac{g_e}{2}$ is the electron g-factor, $\frac{m_\mu}{m_e}$ the muon-to-electron mass ratio and $\frac{\mu_p}{\mu_e}$ the proton-to-electron magnetic moment ratio. This article will focus on the data reconstruction and analysis framework for the ω_a frequency measurement.

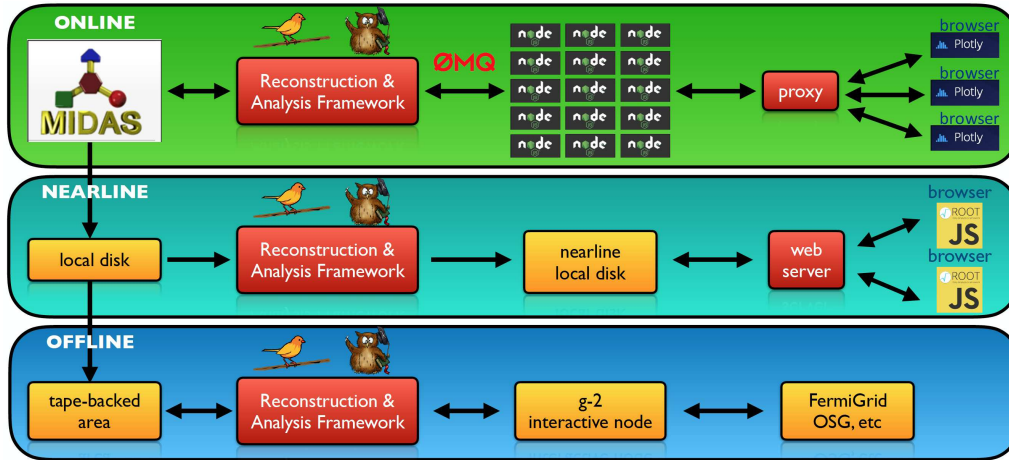


Figure 1. Data flow from online to nearline to offline for the Muon g-2 experiment.

2. Data processing flow

The Muon g-2 event building process is done using a modern data acquisition system, MIDAS [4] developed at PSI and TRIUMF. Data from twenty-four calorimeters, three tracking stations, four fiber harps, Inflector Beam Monitoring System (IBMS), T-zero (T0) detector, three kickers and so on are streamed into a backend machine running MIDAS DAQ. It is worth noting that due to the high incoming data rate (20 GB/s), calorimeter frontend machines are equipped with NVIDIA K40 GPUs to pre-select over-threshold short pulses from 54 PbF₂ crystals per calorimeter each recorded at 800 MSPS in 700 μ s long raw waveforms. The DAQ system is triggered by an FC7-based clock and control system which takes input from the Fermilab accelerator trigger system. The average rate at which muon bunches are injected (each bunch is defined as a *fill* in subsequent paragraphs) and hence the DAQ rate is 12 Hz. The raw data is stored in the MIDAS binary format in a 40 TB RAID6 local disk before being transferred to a tape-backed area in Fermilab for permanent storage.

The data quality monitoring (DQM) system for the experiment is a custom-built framework built on top of software stacks of MIDAS, *art* [5], ZeroMQ [6], Node.js [7] and Plotly [8]. Communication between MIDAS and *art* is enabled using a custom-built module in *art* called *midastoart*. It is developed based on the MIDAS *mserver* utility. The reconstructed and analyzed data are then aggregated using the Node servers, as shown in Fig. 1(top). Finally, the data are plotted using Plotly, an online data analytics and visualization tool. A reverse proxy is also implemented to balance the query traffic.

In addition to the DQM, the experiment also has a nearline operation (see Fig. 1 (middle)) that provides live feedback regarding the quality of the data files and the quality of the reconstructed data. It also produces the latest calibration data for the DQM and the offline data processing. It operates on a multi-core machine utilizing the availability of the multi-threaded modules. Reconstructed and analyzed data are stored in the *art*/ROOT [9] format in the nearline local disk. Histograms of high-level physics variables are also stored in separate ROOT files. Users can browse the histograms through a web server running JSROOT [10].

Once the raw data files are transferred to the tape-backed area, they will be processed using a combination of FermiGrid and Open Science Grid (OSG) as shown in Fig. 1 (bottom). A web interface named POMS [11], which was initially developed by a Fermilab offline production team, is utilized to automate and to book-keep the production process.

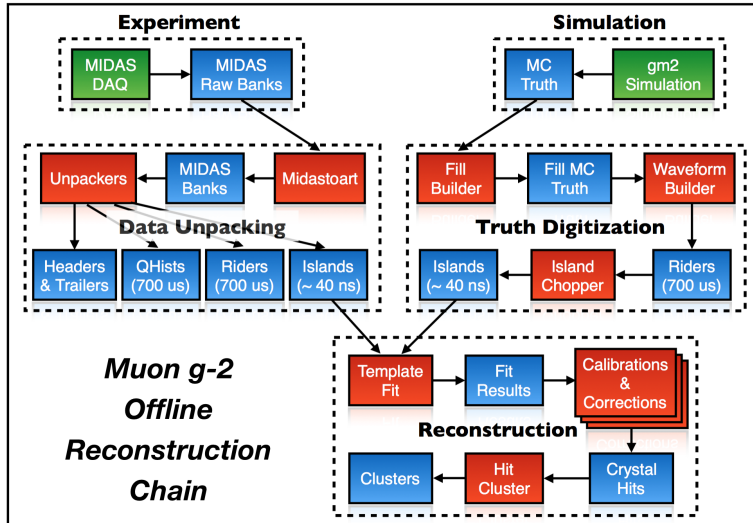


Figure 2. Data reconstruction and analysis chain implemented in the *art* framework for the ω_a measurement.

3. Data reconstruction and analysis framework

The Muon g-2 offline reconstruction and analysis chain is depicted in Fig. 2. Each stage of data processing (data unpacking, digitization, and reconstruction) consists of a series of algorithms and data products. For example, for the simulation chain, the digitization stage starts with a *fill builder* which aggregate 10^4 muon decay events and converts them into a single g-2 fill event. A waveform building module with the implementation of SiPM responses and digitizer behaviors is used to simulated the digitizer waveforms. Finally, an *island chopping* module is used to filter out over-threshold short pulses (also called *islands*) in the 700 μ s long waveforms, similar to the island chopping done by the GPUs in the frontend machines. A typical reconstruction stage for the threshold-method (T-method) ω_a analysis starts with the pulse fitting (using a template pulse shape, χ^2 minimization is done using Eigen linear algebra library [12]) and ends with the hit clustering. Each algorithm will output a data product that serves as an input to the next algorithm. By having this modularity, it is very convenient to study alternative reconstruction algorithms developed by different collaborators.

The difference between simulated and experimental data in the reconstruction stage is that, in the simulation, all the PbF_2 crystals and the silicon photomultipliers (SiPMs) are assumed to be identical (e.g. in terms of crystal light yields and SiPM gains). Quality control of the crystals and the SiPMs have indicated that they could differ by about 20% each regarding light transmission and breakdown voltage. Hence in practice, additional steps have to be taken to equalize them as much as possible in the experiment. A gain calibration procedure using fast pulsed laser and neutral-density filters is implemented to extract gain calibration constants (from fitted pulse integral to the number of photo-electron, *npe*), and a laser monitoring system is used to monitor the SiPM gain fluctuations. Additional corrections involve monitoring the laser intensity fluctuation using pin diodes and photo-multiplier tubes (PMT). Finally the energy calibration constants (from *npe* to MeV) are extracted using either the minimum ionizing particle (MIP) signals from the muons or the end-point matching of the decay energy spectrum. All these calibration and correction constants are applied to the fitted energy in sequence, and the final number is a proxy for the particle energy. On the other hand, the hit time information is extracted from the template-fitting technique, and it has a resolution of about 20 ps. A pulse-pileup separation of 100% can be achieved for a time difference of larger than 5 ns between two

pulses.

The ω_a offline framework also has extensive interfaces to the database of the experiment for reconstruction and analysis purposes. They are the *MIDAS online database (ODB) parser* module and the *database service* module. The ODB is written into the raw MIDAS file and includes most of the run information needed for data reconstruction. A dedicated ODB parser based on the boost library is implemented to extract information like the digitizer channel to crystal number mapping and the number of enabled digitizer channels. Calibration and correction constants extracted by *art* modules are stored in a FNAL central database for g-2. A dedicated database service that can talk to both databases is implemented using libraries like *pqxx* (Postgres C++ API) [13] and *boost::asio* (URL request library) [14].

In addition to the feature mentioned above, we have also utilized the Intel Threading Building Blocks (TBB) to parallelize data processing within an *art* module. This parallelization is possible because our data products are in a nested structure and data from each calorimeter is independent from each other. We have implemented such parallelization in the unpacking and the reconstruction stages. A fill event processing rate of 8 – 10 Hz is achievable using a 16-core CPU machine, and such a speed up enabled us to run DQM and nearline operation without missing many events.

4. Performance during the Summer 2017 commissioning run

During the 5-week commissioning run in Summer 2017, the event rate was about 0.15 Hz. The DAQ uptime was about 60% and the online DQM and nearline uptimes were close to 100%. Around 20 TB of data was recorded. The nearline data processing rate was limited by the ROOT I/O. To catch up with the DAQ rate, the ROOT compression level had to be reduced.

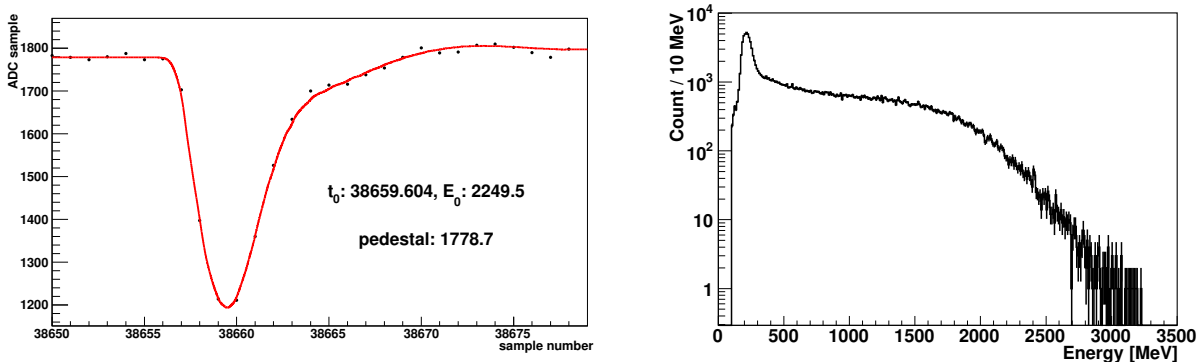


Figure 3. (left) Template fit of a digitized SiPM pulse. Red curve is the template pulse shape and black dots are the digitized samples for a typical positron impact. (right) Distribution of the reconstructed particle energy of a typical run.

As part of the upstream beamline that eliminates the proton and the pion beam was not ready during the commissioning run, the incoming beam was not purely muon but a mixture of the three species (p , π^+ , μ^+). An example of the fitted pulse is shown in Fig. 3(left). A typical energy distribution of the detected particles is displayed in Fig. 3(right). The low energy peak (≈ 200 MeV) is dominated by the $p/\pi^+/\mu^+$ MIP signals whereas the higher energy part is coming from mainly muon-decay positrons. A typical hit distribution is shown in Fig. 4(left). It is in good agreement with the simulation [15] where most of the hits are expected to be concentrating on the left-hand side of the calorimeter which is closer to the storage region of the beam. A preliminary result showing the modulation of the high energy positron rate is depicted in Fig. 4(right).

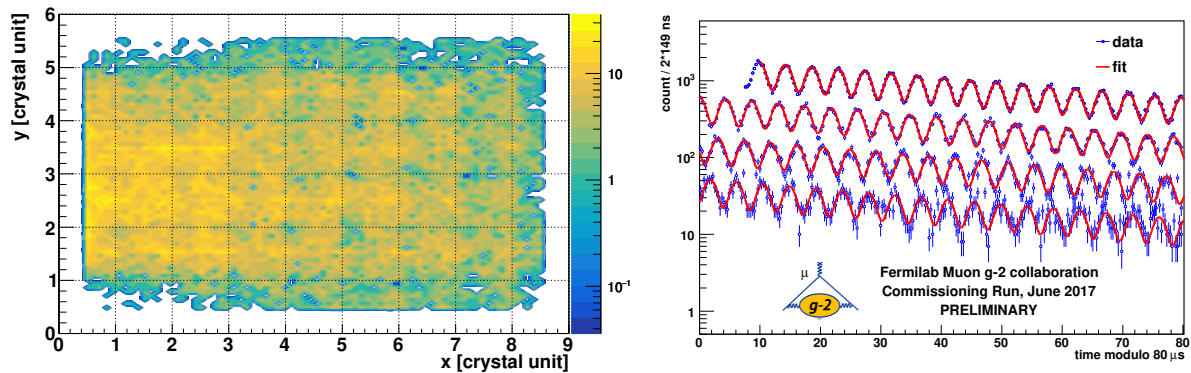


Figure 4. (left) Distribution of the reconstructed hit positions on calorimeters. (right) The number of positron with $E > 1.86$ GeV as a function of time detected by the calorimeters. The modulation is at the muon anomalous precession frequency.

5. Conclusions

In summary, the reconstruction and analysis framework for the ω_a analysis of Muon g-2 experiment is implemented into a highly-integrated system from online to offline using modern standardized tools. The raw detector data was first processed using GPU farms to reduce the incoming rate by a factor of 100, and the reconstruction and analysis of the data are performed using the algorithms implemented in the *art* framework. The reconstruction stage is a chain of modules performing pulse fitting, energy calibration and hit clustering. The system performed very well during the Summer 2017 commissioning run and is currently being optimized based on the data collected during the commissioning run.

6. Acknowledgments

The author thanks A. Fienberg for useful discussions in preparing this manuscript. This work was supported, in part, by the U.S. Department of Energy Office of Science, Office of Nuclear Physics under award number DE-FG02-97ER41020. This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

References

- [1] J Grange *et al.* [Muon g-2 Collaboration], “Muon (g-2) Technical Design Report,” arXiv:1501.06858 [physics.ins-det].
- [2] A T Fienberg *et al.*, “Studies of an array of PbF₂ Cherenkov crystals with large-area SiPM readout,” Nucl. Instrum. Meth. A **783**, 12 (2015) doi:10.1016/j.nima.2015.02.028 [arXiv:1412.5525 [physics.ins-det]].
- [3] J Kaspar *et al.*, “Design and performance of SiPM-based readout of PbF₂ crystals for high-rate, precision timing applications,” JINST **12**, no. 01, P01009 (2017) doi:10.1088/1748-0221/12/01/P01009 [arXiv:1611.03180 [physics.ins-det]].
- [4] MIDAS project, “MIDAS” [software], version 2.1, available from <https://bitbucket.org/tmidas/midas> [accessed 2017-10-18]
- [5] C Green, J Kowalkowski, M Paterno, M Fischler, L Garren and Q Lu, “The Art Framework,” J. Phys. Conf. Ser. **396**, 022020 (2012). doi:10.1088/1742-6596/396/2/022020
- [6] ZeroMQ project, “ZeroMQ” [software], version 4.2.2, Available from <http://zeromq.org/intro:get-the-software> [accessed 2017-10-18]
- [7] Node.js project, “Node.js” [software], version 6.11.4, Available from <https://nodejs.org/en/download/> [accessed 2017-10-18]
- [8] Plotly Technologies Inc., “Collaborative data science,” Plotly Technologies Inc., Montréal, QC, 2015, <https://plot.ly>

- [9] R Brun and F Rademakers, "ROOT: An object oriented data analysis framework," Nucl. Instrum. Meth. A **389**, 81 (1997). doi:10.1016/S0168-9002(97)00048-X
- [10] JavaScript ROOT project, "JSROOT" [software], version 5.2.1, Available from <https://root.cern.ch/root/html/doc/guides/JSROOT/JSROOT.html#javascript-root> [accessed 2018-04-09]
- [11] D Box *et al.*, "The FIFE Project at Fermilab : Computing for Experiments," PoS ICHEP **2016**, 176 (2016).
- [12] Eigen project, Eigen [software], version 3.3.4, Available from <https://bitbucket.org/eigen/eigen/> [accessed 2017-10-18]
- [13] pqxx project, "pqxx" [software], version 5.0.1, Available from <https://github.com/jtv/libpqxx> [accessed 2017-10-20]
- [14] J Torjo, "Boost.Asio C++ Network Programming," <http://www.packtpub.com/boost-asio-cplusplus-network-programming/book>
- [15] gm2ringsim project, "gm2ringsim" [software], version 7.0.4, Available from <https://cdcv.s.fnal.gov/redmine/projects/gm2ringsim> [accessed 2017-10-20]