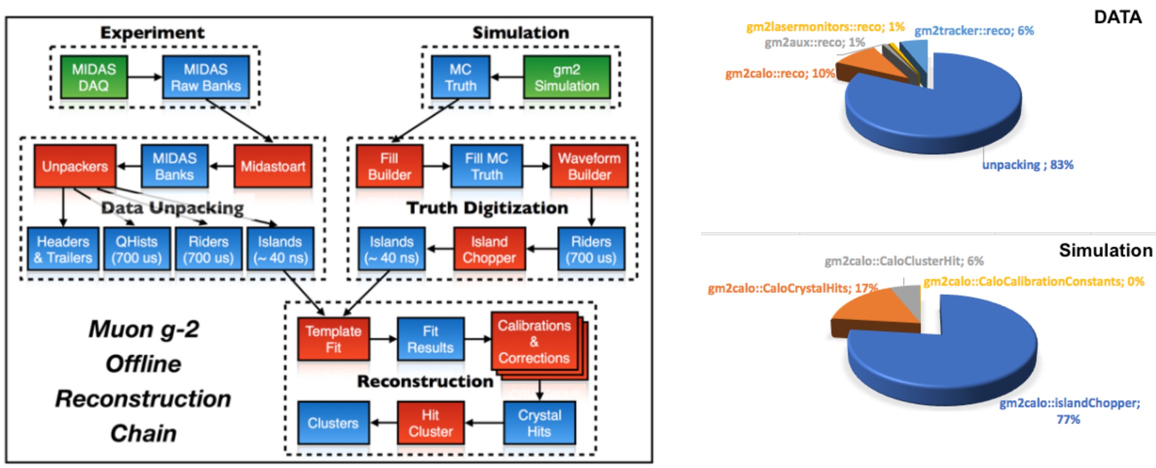
Muon $g-2$ Disk Space Requests for 2018

March 11, 2018

The goal of the Muon $g-2$ collaboration for 2018 is to run the experiment and collect physics data of at least the same statistics of the BNL experiment, with first results ready for the 2019 winter conferences.

The Muon $g-2$ experiment uses MIDAS for data acquisition, a software developed at PSI and TRIUMF [1]. The MIDAS DAQ backend machines collect data from a large number of detectors and systems: the 24 calorimeters, the laser calibration monitoring system, the 3 tracking stations, the 4 fiber harps, the Inflector Beam Monitoring System (IBMS), the T-zero (T0) detector, the quadrupoles and the three kickers. The incoming data rate is expected to be approximately 20 GB/s. To reduce the amount of space, the frontend machines of the calorimeters pre-select the pulses from the 1296 calorimeter crystals recorded over a period of 700 μs, using a threshold-based algorithm (so-called "island chopping") on twelve GPUs [2]. The raw, compressed data is stored in MIDAS files and is estimated to require about 750 TB of disk space in 2018. In order for the users to perform analysis, the raw files are converted offline to the art/ROOT format. This task is done by a production team using the official software releases, distributed by CVMFS, to unpack (read the raw MIDAS file and store the information as an *art* event structure) and reconstruct (process the raw data using algorithms, such as pulse fitting or clustering) the raw data on the FermiGrid [3]. For the end of December 2018, the $g-2$ collaboration expects to have stored about 600 TB of reconstructed files which would correspond to $3\times$ BNL statistics. The $g-2$ Italian group is planning to select from these files the information needed for the measurement of ω_a with the so called "T-Method" technique (*i.e.*, by performing a fit of the number of positrons detected *vs.* time). This information, which comprises the energy and the time of hits in the calorimeters' crystals, the energy and space-time coordinates of the reconstructed clusters and the data from the laser calibration system, corresponds to about 10% of the reconstructed data. We plan to store it in ROOT files, and estimate that the amount of space needed for this in 2018 is about 60 TB. It will be of great importance to store this data in Italy, because having local access to the data will improve the analysis performance. Moreover, the ROOT format could be analyzed without the *art* framework, making it more accessible. Besides data, we plan to save simulation files, which are necessary for testing the algorithms and for the absolute energy calibration of the calorimeters. Similar to data, simulation files are produced by the production team, and both the truth and reconstructed information are stored as *art* files. In 2018, we expect 20% of muons stored will be simulated and simulated events need 50% more disk space for truth information, which corresponds to an occupancy of 200 TB. Therefore, in 2018 the Muon $g-2$ offline reconstruction and simulation chain, depicted in Fig. 1, will provide about 600 TB of (reconstructed) data plus 200 TB of simulated events. We plan to reduce the data occupancy by at least a factor of 10 by



(a) (b)

Figure 1. a) Muon $g-2$ official offline reconstruction and simulation chain [2]. From these outputs, which are in the *art* format, we plan to extract the essential information for the measurement of ω_a (*i.e.*, fit results, crystal hits and clusters) and store it as ROOT TTrees in INFN's disks.
b) Example of the space required by the *art* event branches of the reconstructed (top) and simulated (bottom) data files.

storing the essential information for the measurement of ω_a in the INFN disks, so that the amount of space required is about 80 TB. The files can be transferred from Fermilab's machines to INFN's storage space or, as an alternative, the ROOT files can be generated by running the reconstructed dataset through an analyzer on the INFN-Pisa grid and then store the output locally.

In addition, since the gain calibration is the main task of the Italian group, we would like to install the $g-2$ and dependent software (which is available in CVMFS) for exercising the whole gain calibration chain from raw data to full reconstruction. For this task, it is required to store few days of raw data (~ 20 TB) and a copy of the databases. The offline framework interfaces with the latter for reading the run information needed for data reconstruction, and for writing/reading the calibration and correction constants [2] (~ 500 GB/year).

In conclusion, to perform analysis and to test/improve the whole gain calibration offline procedure for 2018 we need about 100 TB of INFN storage space.

A summary of the expected space resources for the Muon $g-2$ experiment in the years 2018 (6 months or running), 2019 (9 months of running) and 2020 (3 months of running) are shown in Table 1. The numbers for 2018 are obtained using realistic machine conditions so far achieved for which we expect 2 months to collect the BNL statistics. For 2019 and 2020 the projections have been computed using the "nominal operation" conditions: 12 Hz average muon spill rate that comprises four fill of 700μs each, separated by 11 ms and a beam rate of 120k muon/second [4]. Running in nominal operation conditions will provide 20 times the BNL statistics (*i.e.*, $\sim 20 \times 7 \times 10^9$ detected positrons) in about 12 months.

Year	N. detected positrons	BNL Statistics	Raw Data [TB]	Full Reconstructed Data [TB]	Simulated Data [TB]
2018 (6 m)	21×10^9	3	750	600	200
2019 (9 m)	105×10^9	15	1250	1100	1000
2020 (3 m)	35×10^9	5	400	350	400

Table 1: Expected space resources required for data storage in year 2018, 2019 and 2020.

References

[1] MIDAS project, `https://midas.triumf.ca/MidasWiki/index.php/Main_Page` [accessed 2013-03-05]

[2] K. S. Khaw [Muon $g-2$ collaboration], *"Muon $g-2$ reconstruction and analysis framework for the muon anomalous precession frequency"* 18th International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT 2017) Seattle, WA, USA, August 21-25, 2017", arXiv:1710.07839 [physics.ins-det].

[3] R. Fatemi, K.S. Khaw, L. Li, A. Lyon, *"Data Production for the Muon g-2 experiment"*, DocDB4638 [2016].

[4] J. Grange, *et al.*, [Muon $g-2$ collaboration], *"Muon (g-2) Technical Design Report"*, [2015].