Data Production for the Muon g-2 experiment

Renee Fatemi, Kim Siang Khaw, Liang Li, Adam Lyon

November 29, 2016

1 Production Workflow

This document outlines the workflow of the data production of the Muon g-2 experiment. There are two different production chain: simulation and DAQ.

1.1 Simulation production workflow

Simulation chain involves generating Geant4-based simulated data files, digitization of the truth information and reconstruction of the digitized information. Interaction between the gm2 instance and jobsub and SAM in the simulation chain is summarized in Fig. 1.

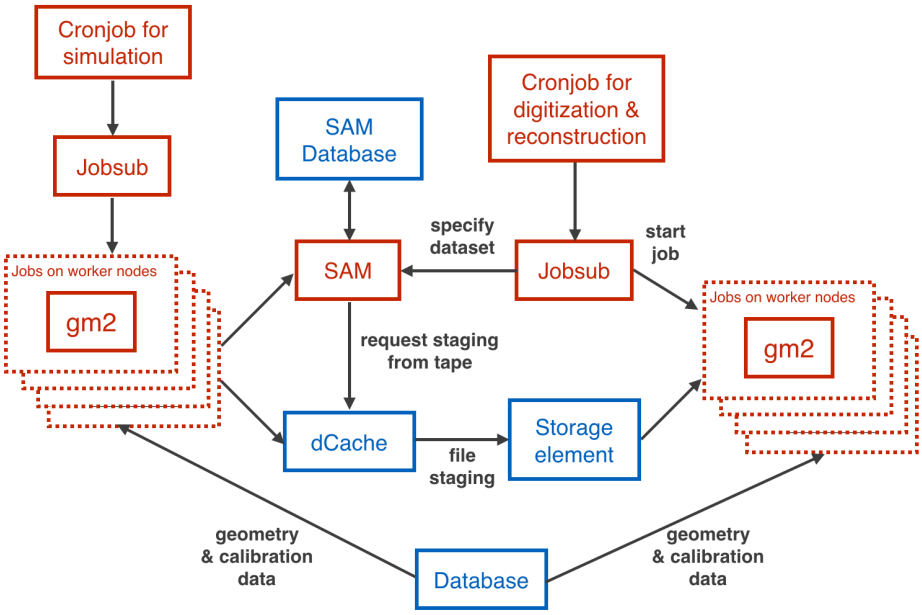

Figure 1: Workflow for the simulation production.

The following describes the basic steps in the simulation production:

1. A cronjob is setup to submit jobs to the FNAL grid at a specific time interval.

2. Worker nodes then execute the submitted scripts to generate simulated data. Database may or may not be used for the simulation.

3. Metadata of the generated data files are communicated to SAM data handling system and the files are transferred to FNAL permanent storage area.

4. Another cronjob independent of Cronjob1 is setup to submit jobs to the FNAL to digitize and reconstruct the simulated data.

5. Worker nodes then specify SAM dataset to be digitized and reconstructed.

6. The reconstructed data are then stored in the permanent storage area.

1.2 DAQ production workflow

Interaction between the gm2 instance and jobsub and SAM in the DAQ chain is summarized in Fig. 2.

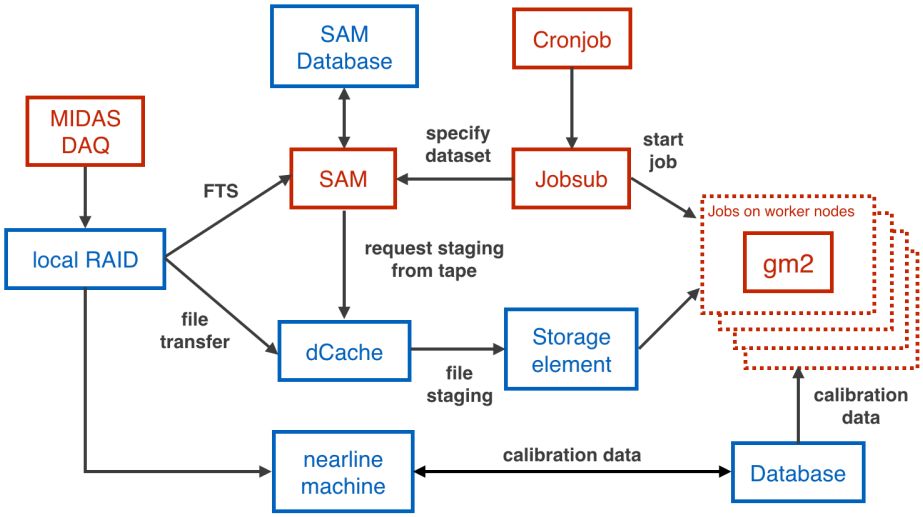

Figure 2: Workflow for the DAQ production.

The following describes the basic steps in the DAQ production:

1. MIDAS DAQ outputs raw data and stores them into local RAID storage.

2. A backend machine running FTS transfers raw files to the permanent storage area while communicates with SAM regarding the metadata of the files.

3. At the same time, a nearline machine analyzes specific calibration runs and extracts calibrations from these runs. All the constants are stored in the database.

4. A cronjob is setup to submit jobs to unpack and reconstruct the DAQ data.
5. Worker nodes then specify SAM dataset to be unpacked and reconstructed.
6. The reconstructed data are then stored in the permanent storage area.

2 SAM Metadata: Metadata definition and dataset definition

How to insert metadata using FTS, art, etc. What are the metadata we want to insert, etc.

3 Workflow of the data production: Implementation

mock data handling, type of runs, working schedule, deadlines, etc

4 Versioning of the scripts/codes (related to releases)

5 POMS: What do we know so far

Reference slides:
https://indico.fnal.gov/getFile.py/access?contribId=8&resId=0&materialId=slides&confId=12120