

# GORY DETAILS EXAMPLES

4/7/14

What actually goes on in an analysis job

# Example tuple creation

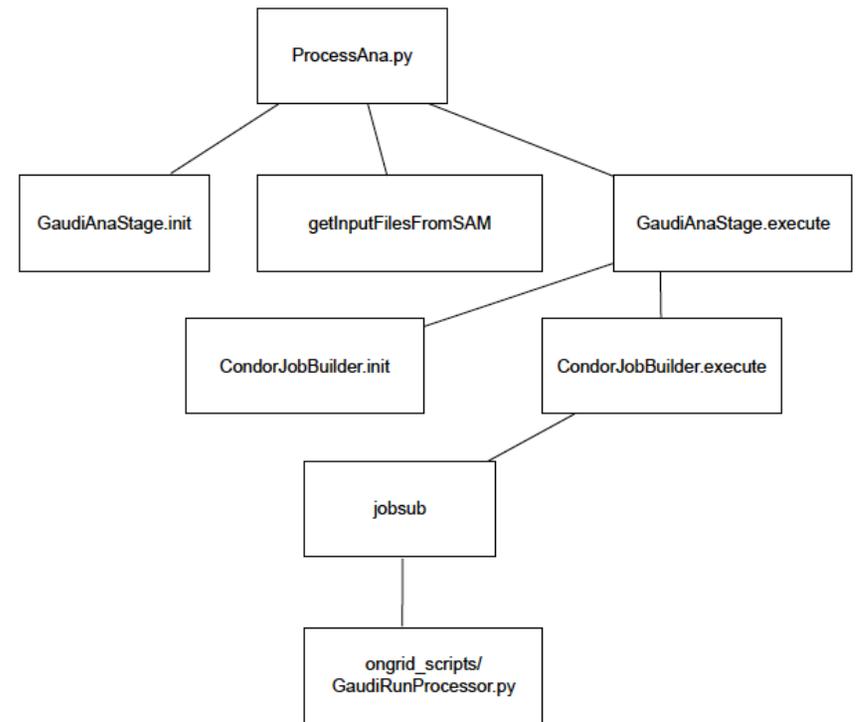
2

- Configure job runs on interactive node
- Creates a set of batch jobs
- Each batch jobs run a script on a fermigrid node
- Files come back to dcache/bluearc
- Check for completeness
- Normalization
  
- Two use cases
  - ▣ Official sample with direct access
  - ▣ Official sample with sam access

# Scripts used

3

- Productionscripts
  - ▣ ana\_scripts/ProcessAna.py
    - Sets up the job
  - ▣ py\_classes/GaudiAnaStage.py
    - More setup based on optior selected
  - ▣ ongrid\_scripts/GaudiRunProcessor.py
    - Actually does the execution on the grid node
- CondorUtils translates options for jobsub

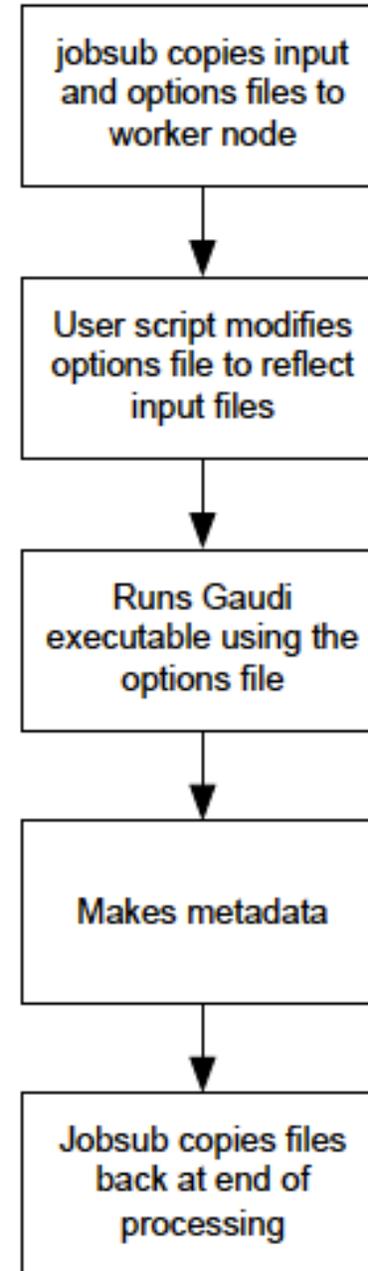


# Case 1: 'the old way'

## Using `-f` option

4

- ❑ Original submission job finds the data (using SAM as a catalog) and chops up into chunks for you.
- ❑ All chunks are in a single run.
- ❑ Use the `jobsub -f` option to have `jobsub` bring the files to grid node
- ❑ Subrun content stored in output file name

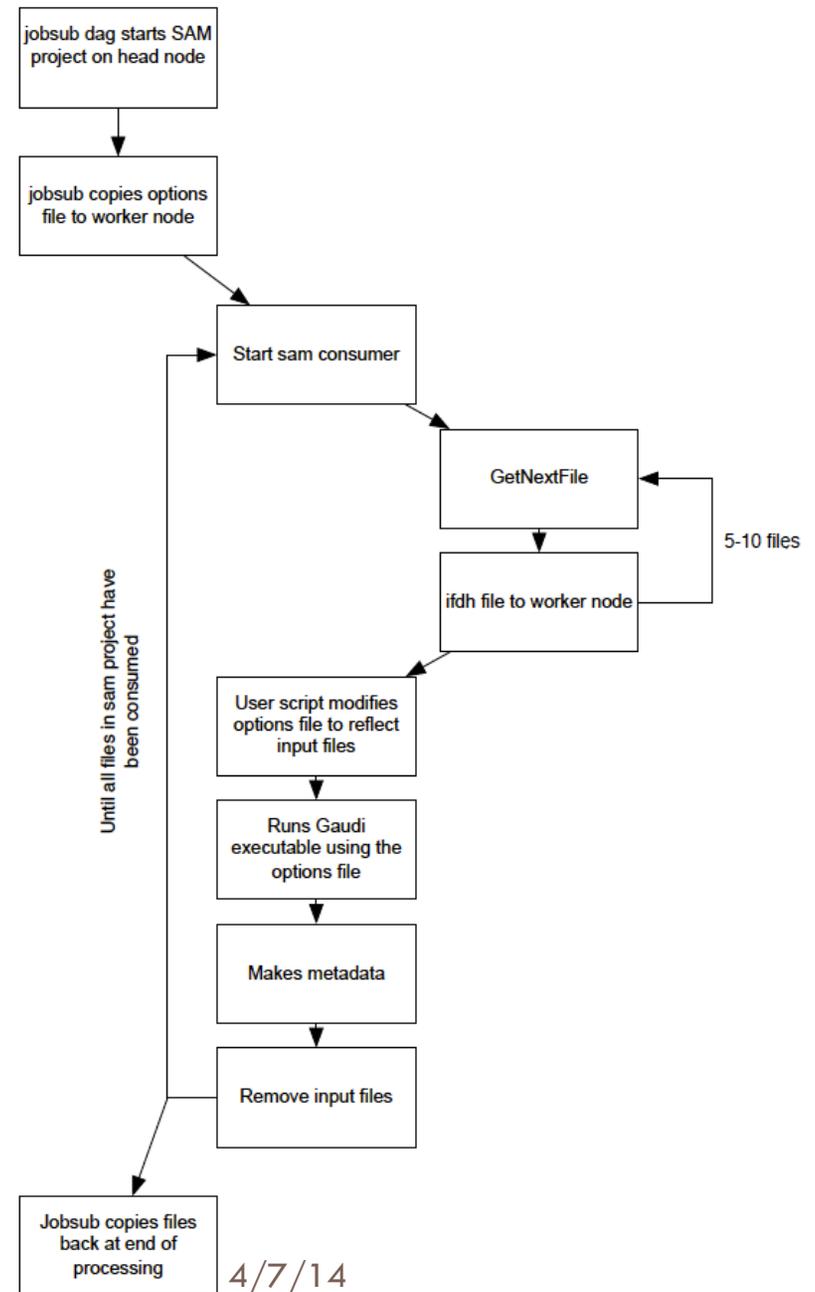


# Case 2: 'the new way'

## Using sam

5

- ❑ Loop in groups of 5-10 files
- ❑ One consumer per group to retain error handling
- ❑ File order not determined
- ❑ Creating unique output filename an issue
- ❑ Still under test



# A Plan

6

- High priority items for Minerva are:
- Getting all layers of production to make good metadata for samweb
  - ▣ Write instructions for users to convert to doing this themselves
- Getting dcache FTS dropbox to work so can write production directly to dcache
  - ▣ In testing – no monitoring at the moment
- Getting all layers of production/analysis to use dcache/sam catalog for input
  - ▣ Write instructions for users to convert to doing this.