SUMMARY OF SIMULATION AND RECONSTRUCTION PROCESSING

Mousumi Datta

Hampton University

Minerva Data Management Review

April 8, 2014

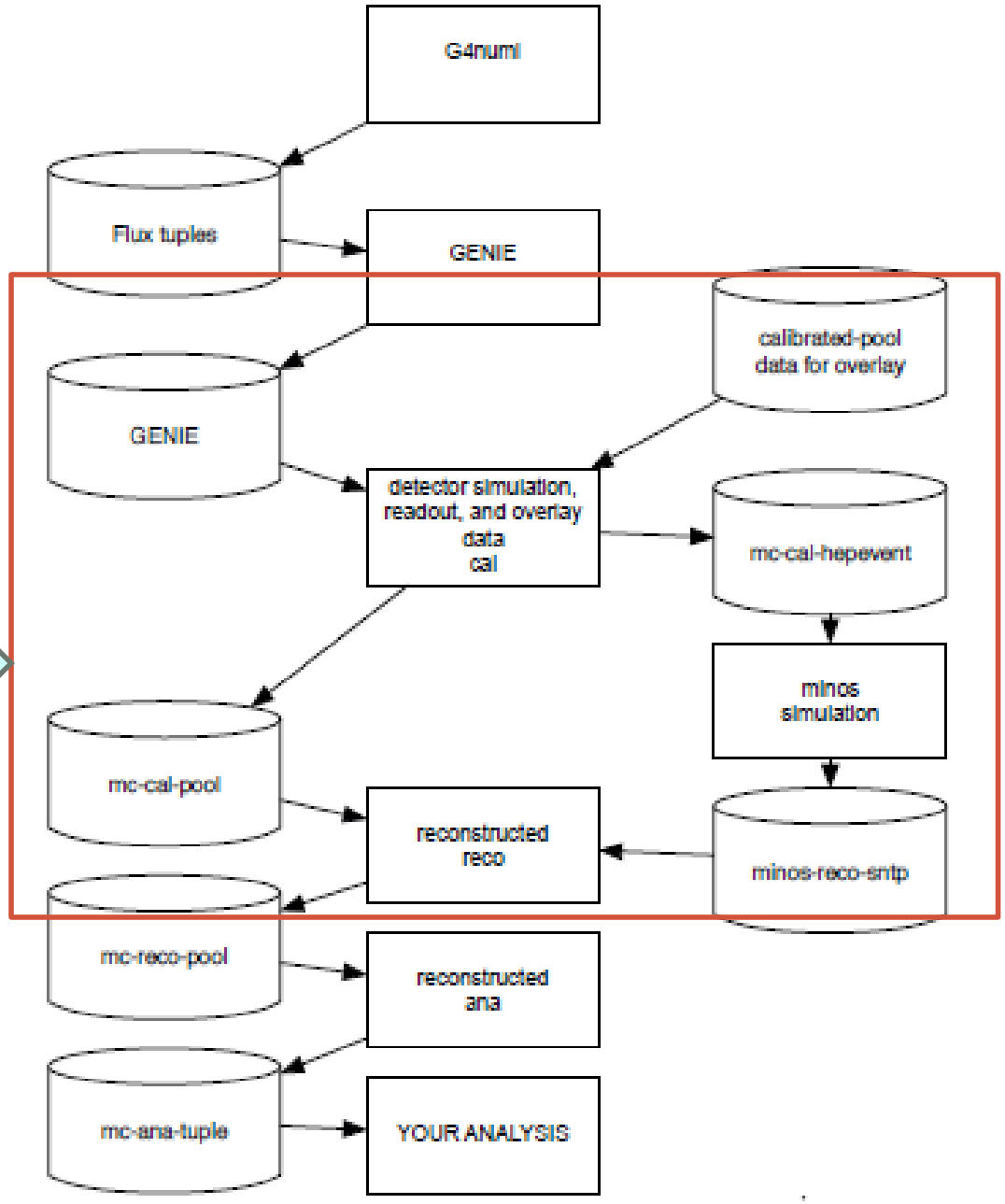
Overview

- Simulation and reconstruction
  - Even generation using “Genie” covered by J. Park
  - Three stage process
    - Minerva Simulation : “cal”
    - Minos Simulation and reconstruction : “minos”
    - Reconstruction : “reco” (for both simulation and data)
  - Different MC samples
    - “Central value” for different run conditions: playlists 1-12, 13A-E
    - Systematic variations samples
  - Processing time and disk space usage
- Scripts used

Monte Carlo (MC) Production Flow Diagrams

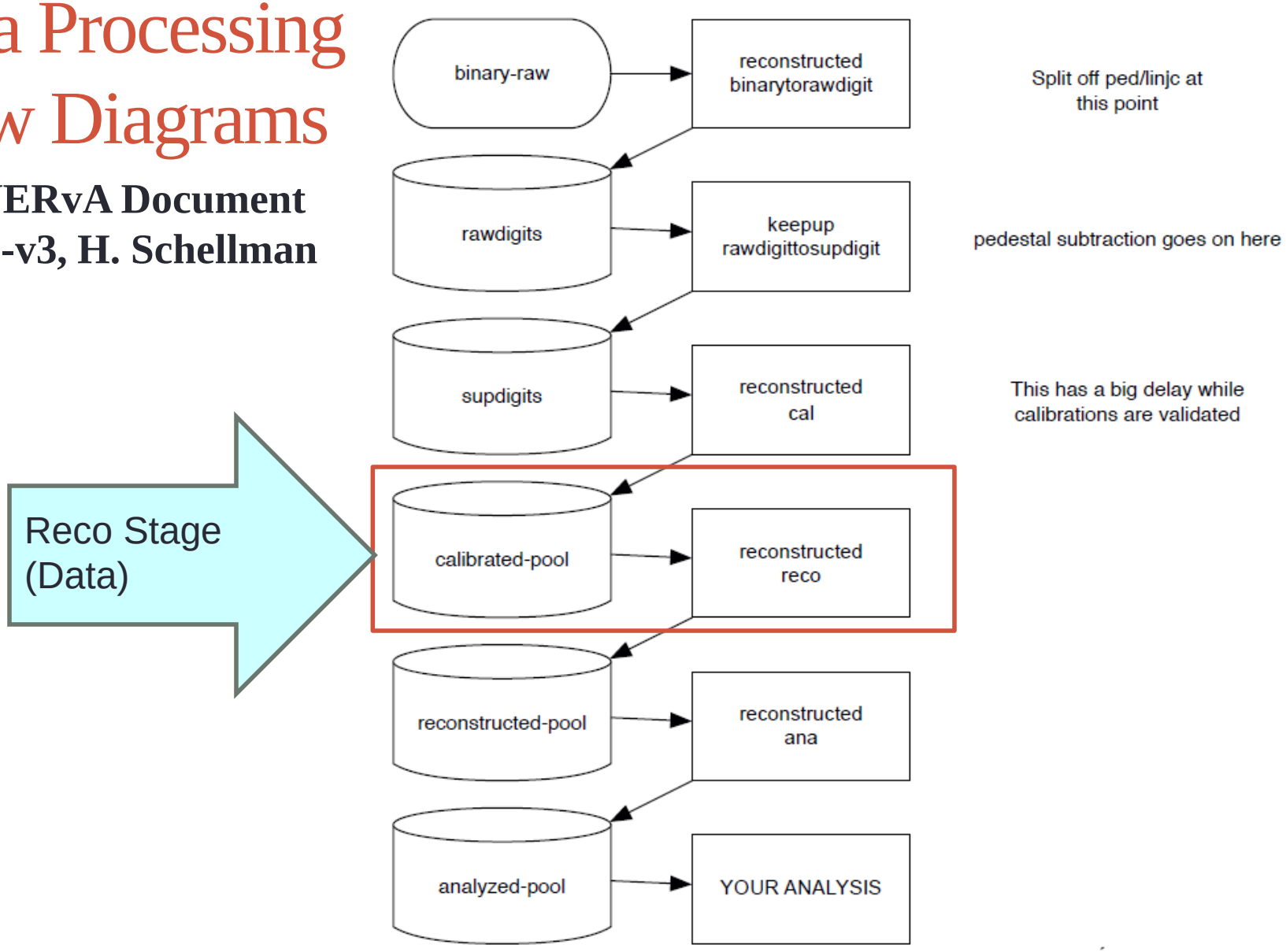
- **MINERvA Document 9773-v3, H. Schellman**

Cal, Minos & Reco Stages (MC)



Data Processing Flow Diagrams

- **MINERvA Document 9773-v3, H. Schellman**



Different MC Samples for Low Energy (LE) Run

MC	Target POT (data POT ~4e+20)	Typical POT/job	# cal file sets	# minos file sets	# reco file sets
Central value	~40e+20	1.0e+17	40,000	40,000	40,000
Systematic Variation	~8e+20 Per variation x 15 = 120e+20	1.0e+17	120,000	120,000	120,000

- Run dependent "central value" (CV) MC
  - For ~15 data taking periods (playlists) : using data overlay from playlist
    - Genie samples generated with input flux (LE/pME/pHE) and (FHC/RHC)
  - Goal: ~x10 of data POT
- Variation samples for systematic studies
  - ~15 variations
  - Goal: ~x2 of data POT

Processing Time and Disk Space Usage

	Cal	Minos	Reco
LE MC : Disk Space	~350GB Per 1e+20 POT	~34GB Per 1e+20 POT	~500GB Per 1e+20 POT
LE MC : Processing Time (1 job → 1e+17 POT)	Run time 3-4 hours per job	Run time 1 hour per job	Run time 3-4 hours per job

- Total ~35 TB for CV MC, ~100 TB for Variation MC
  - The output from cal stage is backed up and removed from disk right after the validation of the outputs from cal+minos+reco processing
- For "minervapro" typical processing time in grid for 1e+20 POT MC cal+minos+reco ~24 hours
  - Includes wait time on grid
- Disk space usage for data in LE run : 7TB

Scripts Used

Running jobs:

- Tools/ProductionScripts/mc_scripts/ProcessMC.py
  - Tools/ProductionScripts/py_classes/MCCalStage.py
  - Tools/ProductionScripts/py_classes/MCMinosStage.py
  - Tools/ProductionScripts/py_classes/MCRecoStage.py
  - Tools/ProductionScripts/py_classes/MCStandardRun.py
- Tools/ProductionScripts/data_scripts/ProcessData.py

Validation and copying to official areas:

- Tools/ProductionScripts/mc_scripts/copyMCFiles.py
- Tools/ProductionScripts/data_scripts/summarizeProduction.py
- Tools/ProductionScripts/data_scripts/copyProcessedData.py

- Outputs from cal+minos+reco stages are first written to a temporary location (currently in /minerva/data/users/minervapro/..)
  - Moved to production area after validation (if metadata exists, exit code in the log file etc.)