GEMC Status, Docker and off-site simulations

R. Johnston, S. Lee, P. Moran, M. Ungaro

CLAS12 Simulation Software Distribution

home | examples

The CLAS12 software packages are distributed using docker images.

Quickstart

Use the following command to run the clas12 software image:

```
docker run -it --rm jeffersonlab/clas12simulations:iprod bash
```

Software content

The clas12simulations docker image contains:

- "deep" generators
- gemc with the clas12 geometry
- CLARA
- Coatjava
- the *CLAS12_BIN*, *CLAS12_LIB*, *CLAS12_INC* dirs and environment variables

Current production version (tag "iprod" or "latest" or no tag): ← tag 1.1

- clasdis, generate-dis, dvcsgen generator executables
- gemc 4.3.0
- Clara 4.3.3
- Coatjava 5.7.4
- ced 1.06

Tag 1.0:

- gemc 4a.2.5
- Clara 4.3.3
- Coatjava 5b.7.1
- ced 1.06

Run Configurations to be released in new container (and we'll demo these gcards)

- Run group A Spring 2018:
  - Central detector shifted 19.4mm upstream
  - target (LH2) at (0, 0, -19.4) mm
  - HTCC shfted 9.5mm upstream
  - FT On configuration
  - FMT present
  - LTCC sectors: 2 (N2), 3 (N2), 5 (old C4F10), 6
  - Torus polarity: -1, 1, -0.75, 0.75
  - Solenoid polarity: -1
  - Beam Current: from 5 to 75 nA

rga-spring.gcard

- Run group A Fall 2018:
  - Central detector shifted 30 mm upstream
  - target (LH2) at (1.2, 1.1, -30) mm
  - HTCC shfted 19.5mm upstream
  - FT On configuration
  - FMT not present
  - LTCC sectors: 3 (50% C4F10), 5 (N2)
  - Torus polarity: -1, 1,
  - Solenoid polarity: -1
  - Beam Current: from 5 to 75 nA

rga-fall.gcard

Run Configurations to be released in new container (and we'll demo these gcards)

- Run group K First part:
  - Central detector shifted 30 mm upstream
  - target (LH2) at (1.2, 1.1, -30) mm
  - HTCC shfted 19.5mm upstream
  - FT On configuration
  - FMT not present
  - LTCC sectors: 3 (50% C4F10), 5 (N2)
  - Torus polarity: 1,
  - Solenoid polarity: -1
  - Beam Current: from 5 to 75 nA

rgk-1.gcard

- Run group B Winter 2019:
  - Central detector shifted 30 mm upstream
  - target (LD2) at (1.2, 1.1, -30) mm
  - HTCC shfted 19.5mm upstream
  - FT On configuration
  - FMT not present
  - LTCC sectors: 3 (C4F10), 5 (C4F10)
  - Torus polarity: -1,
  - Solenoid polarity: -1
  - Beam Current: from 5 to 75 nA

rgb-winter.gcard

Friday Demo: advanced docker usage

- Use of different experiment gcards: RGA, RGB, etc
- How to make geometry changes: different targets, FTOn, FTOff
- How to use the docker image to test your own gemc version.

Production and Development Version

Production:

- 4.3.0: **COATJAVA: 5.7.4, JLAB_VERSION: 2.3**:
  - Updated DC geometry using latest survey (May 18 Entry in DB)
  - Fixed bug that prevented material name from being displayed in the GUI
  - 3d cartesian field map support
  - new geant4 version: 10.4.p02
  - 51 micron tungsten shield (for bst) surrounding the target
  - calorimeters: reading ecal effective velocity from CCDB
  - change htcc time offset table to use the same used in reconstruction
  - Tony Forest: Added polarized target geometry/material and cad volume.

Frozen
Installed on farm
Installed on docker/singularity
(offsite resources)

In development:

- 4.3.1 (By March 2019):
  - FTOF Time resolution updated based on data
  - Option SAVE_SELECTED, RERUN_SELECTED to save RNG state for certain particles, detector
  - Option SAVE_ALL_ANCESTORS to save complete particles hierarchy in output (evio2root also updated)
  - Gemc configuration written to evio file with json strings ➡ SOON
  - gcards for rg-a different run-periods ➡ SOON
  - gcards for rg-b different run-periods ➡ SOON
  - gcards for rg-k different run-periods ➡ SOON
  - ec, pcal digitization removed obsolete constants
  - moved ftof shield in the correct position

Installed on farm, dynamic
Not installed on docker

GITHUB > CLAS12 DOCKER > OSG, MIT Connection is completed and operational

1. Changes to the Docker configuration trigger the docker image creation
2. The docker image creation trigger the singularity image creation
3. The singularity image is distributed to CVMFS

GITHUB > CLAS12 DOCKER > OSG, MIT Connection is completed and operational

1. Changes to the Docker configuration trigger the docker image creation
2. The docker image creation trigger the singularity image creation
3. The singularity image is distributed to CVMFS

What is CVMFS?

CernVM File System (CernVM-FS): a scalable, reliable and low-maintenance software distribution service. The software is available *on demand*.
Read-only file. Files and directories are hosted on standard web servers and mounted in the universal namespace /cvmfs.

GITHUB > CLAS12 DOCKER > OSG, MIT Connection is completed and operational

1. Changes to the Docker configuration trigger the docker image creation
2. The docker image creation trigger the singularity image creation
3. The singularity image is distributed to CVMFS

What is CVMFS?

CernVM File System (CernVM-FS): a scalable, reliable and low-maintenance software distribution service. The software is available *on demand*. Read-only file. Files and directories are hosted on standard web servers and mounted in the universal namespace /cvmfs.

What does that mean?

If your institution has or can install CVMFS (trivial, secure), there is NO OVERHEAD in using the CLAS12 software in your farm. You would have everything installed and ready to go FOR FREE.

GITHUB > CLAS12 DOCKER > OSG, MIT Connection is completed and operational

Submission Script

Working toward ideal scenario: collaborator submitting CLAS12 simulation jobs from her/his sofa, anywhere in the world.

Output File

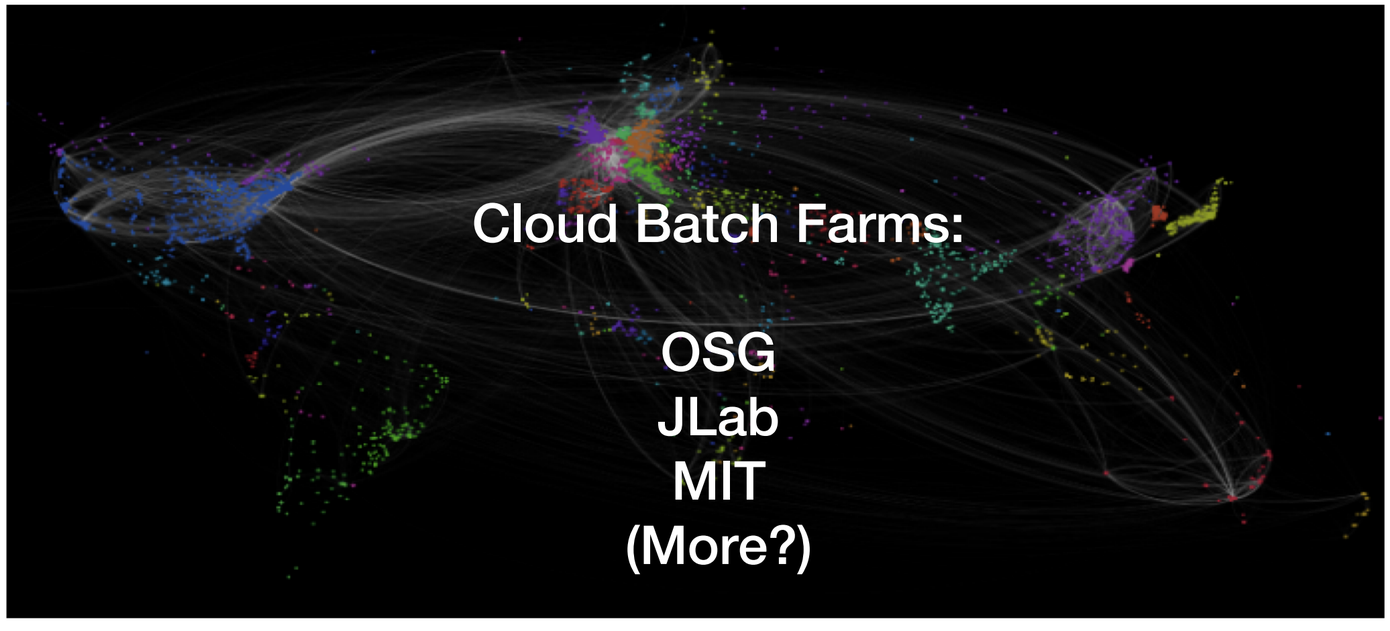
At JLab: Currently: 10 GB / s connection.
FY2020: could be 100 GB connection.

- CLAS12 Singularity image distributed to OSG and MIT nodes
- Submit jobs successful to OSG and MIT
- First draft of scripts to submit jobs to OSG
- Ongoing work for seamless experience for collaborators

JLab iFarm (Slurm)

JLab OSG Node (Condor)

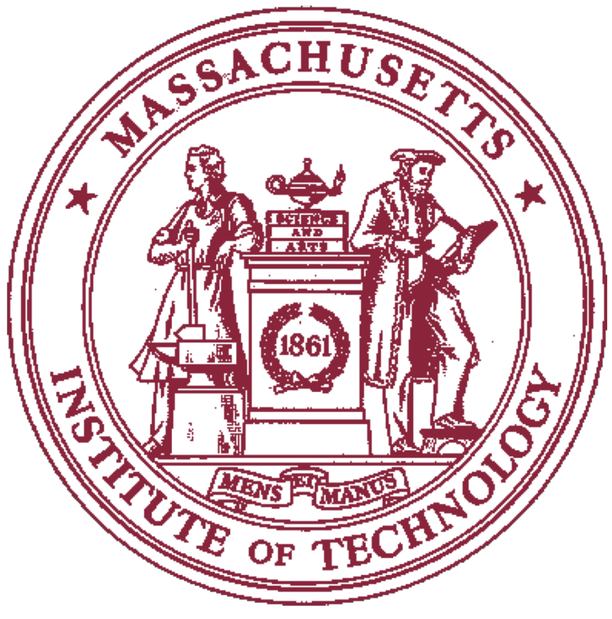
MIT Farm (Slurm / Condor)



CLAS12 MC at MIT/Offsite

CLAS 12 Collaboration Meeting

March 6, 2019

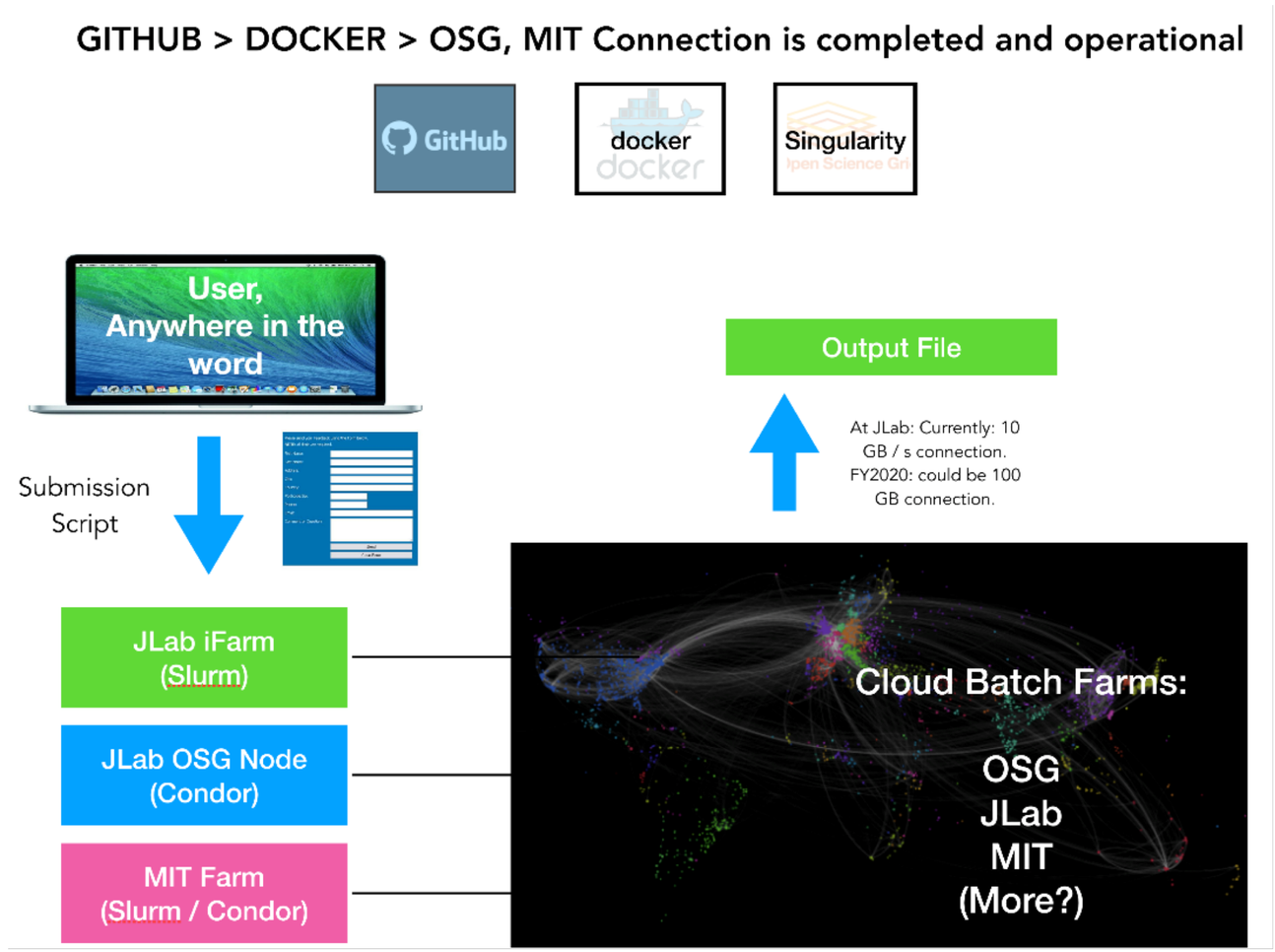


R. Johnston, S. Lee,

P. Moran, M. Ungaro

Overview

- MIT Group
- MIT Resources & Capabilities
- Current Project Status
- Example Simulation
- Current Project Work
- Path Forward



MIT CLAS12 Group

Graduate Students

- Andrew Denniston (RG-M)
- Bobby Johnston (RG-A)
- Sangbaek Lee (RG-A)
- Patrick Moran (RG-A)
- Jackson Pybus (RG-M)
- Efrain Segarra (RG-B)
- Rey Torres (RG-B)

Principal Research Scientist

- Dr. Douglas Hasell

Research Associates

- Dr. Axel Schmidt
- + one to be hired

Faculty

- Or Hen
- Richard Milner

Scientific Focus

- CLAS6 data mining
- CLAS12:
  - High momentum nucleons & EMC effect
  - Deeply virtual exclusive processes

Service Contributions

- Backward Angle Neutron Detector
- CLAS12 Monte Carlo Simulations at MIT

Available Computing Clusters

	Pool Name	Main Users	Total Resources	Status for CLAS12 MC
MIT Pools	**Tier 2**	**LNS Research Computing**	12,000 cores, 2 GB RAM / core	**Operational**
	Tier 3	**LNS Research Computing**	600 cores, 2 GB RAM / core	**Testing**
	EAPS	**MIT Geosciences**	~ 3,000 cores	**Configuring**
Other Pools	**OSG**	**Various**	~ 4 million core hours per day (2017)	**Testing**
	Holyoke	**UMass Amherst, Various**	~ 100,000 cores	**Configuring**

MIT Access to Resources

SubMIT from LNS allows access

- Batch submission interface to computing resources
- Uses HTCondor to manage jobs
- One interface, Four computing clusters:
  - MIT CMS Tier-2
  - MIT CMS Tier-3
  - Engage at MIT Earth, Atmospheric, and Planetary Sciences (EAPS)
  - Connection to Open Science Grid (OSG)

Other Clusters not on SubMIT

  - Massachusetts Green High Performance Computing Center (MGHPCC aka Holyoke)

Current Project Status

CLAS12 MC Runs on SubMIT

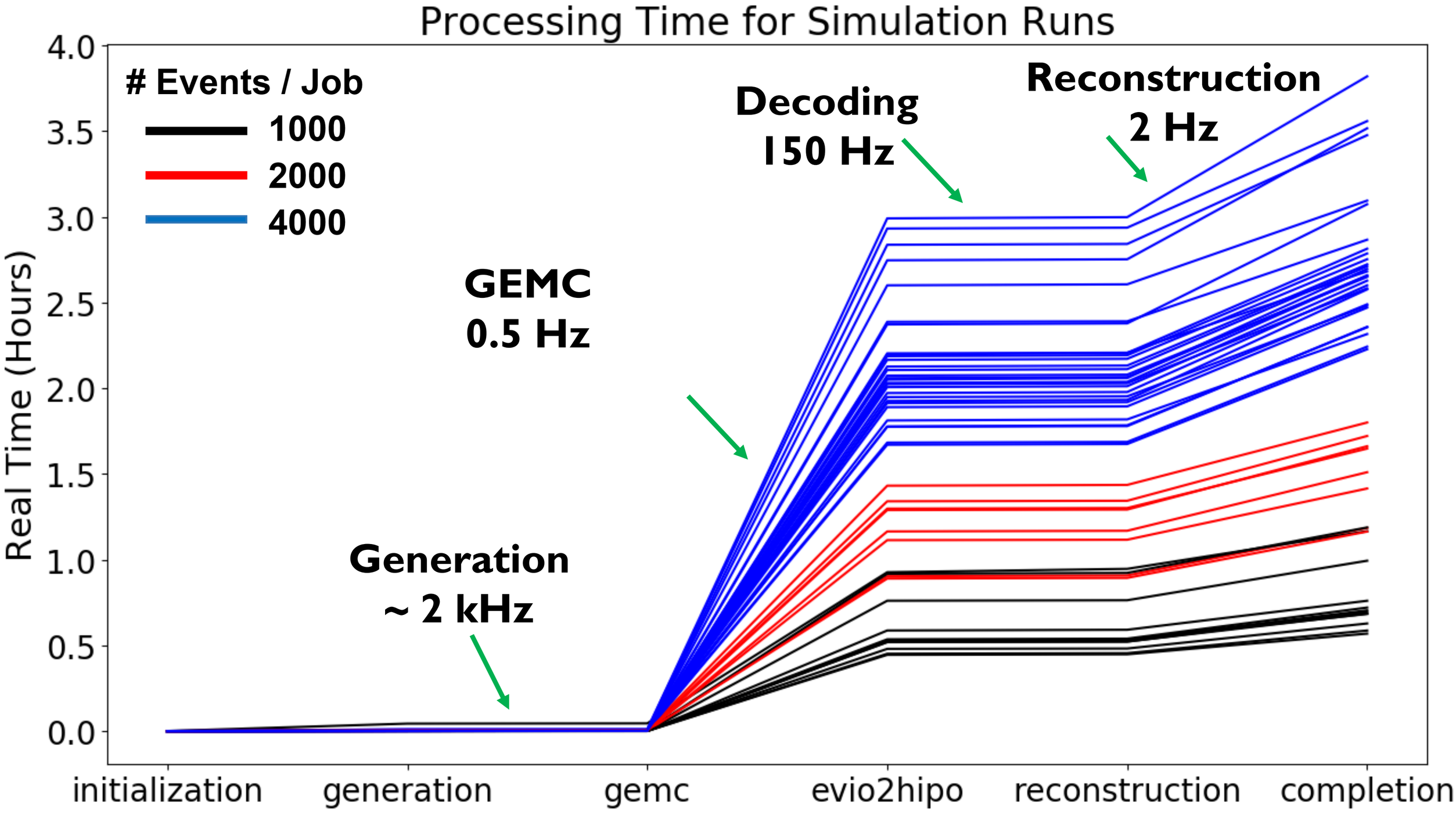
Completed steps:

- ✓ Supports manual job submission for registered users
- ✓ Fully functional on Tier 2
- ✓ Beginning configuring & testing on other pools
- ✓ Currently jobs run “opportunistically”, CLAS12 does not have priority yet. Our priority in future will increase when nuclear physics funding becomes available.

Example Simulation Run

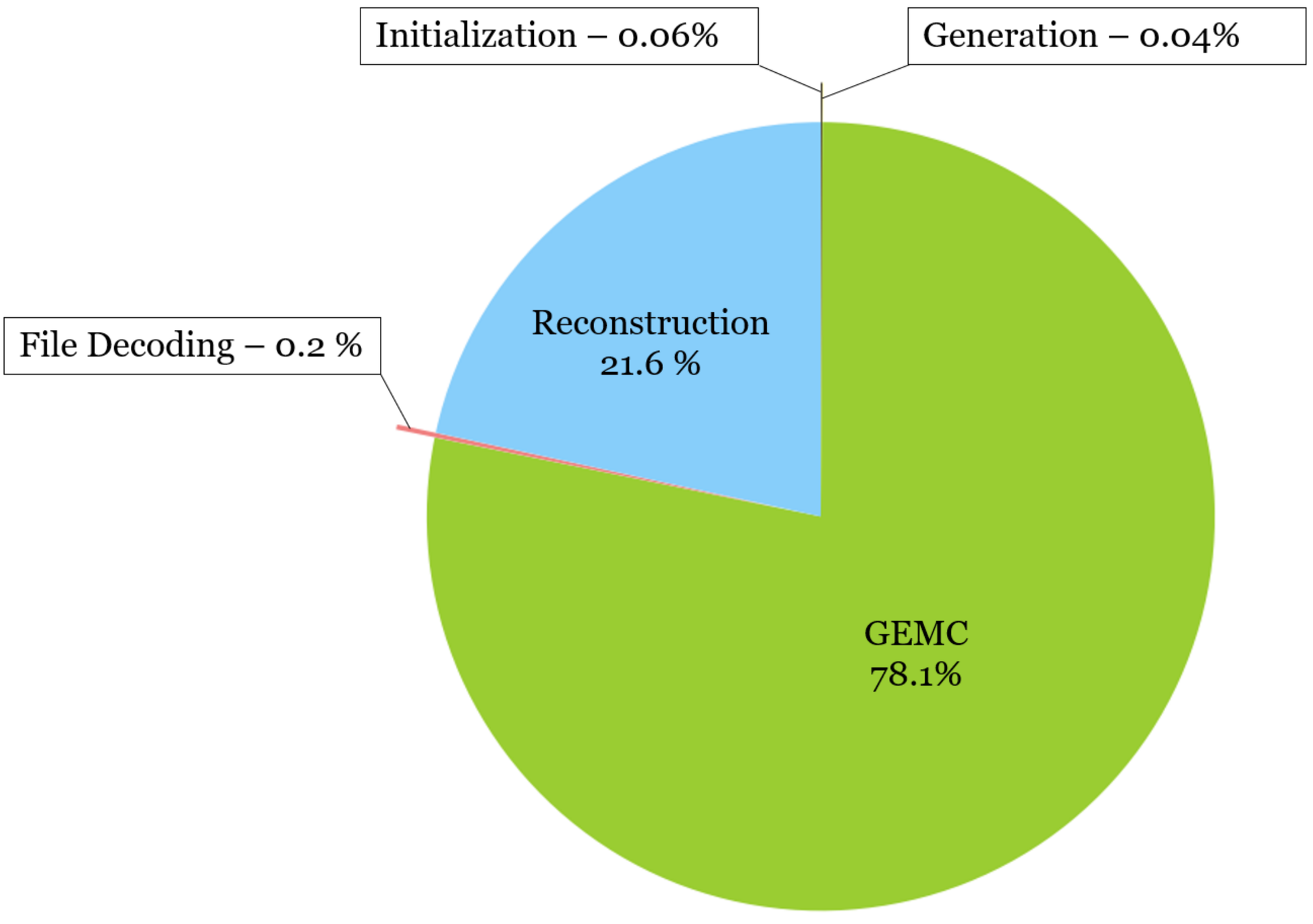
- Simulated 260,000 events using clasdis generator
  - theta between 20 and 25 degrees
- Submission breakdown:
  - 50 jobs with 4,000 events/job
  - 20 jobs with 2,000 events/job
  - 20 jobs with 1,000 events/job
- Docker image "production":
  - gemc 4.3.0
  - Coatjava 5.7.4
  - ced 1.06
- Ran on Tier 2; requested 1 core, 2 GB Memory for each job
- Each job currently includes the following:
  - Event generation
  - GEMC
  - File decoding (EVIO2HIPO)
  - Reconstruction

Simulation Run Statistics



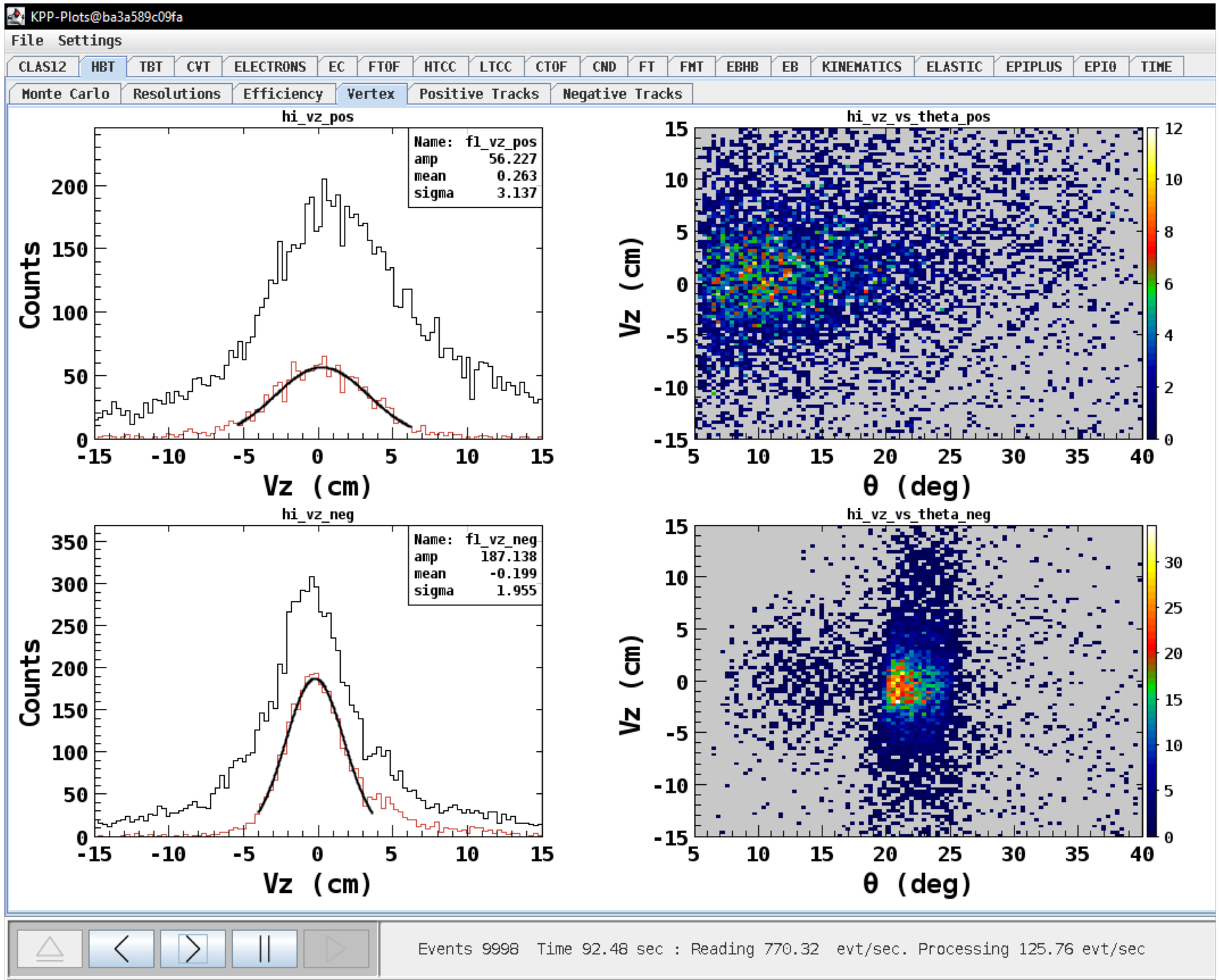
Simulation Run Statistics

Breakdown of Runtime



Simulation Run Output

Sample output from 10,000 events in kpp-plots



Current Work – Submission Framework

All user has to do is submit a configuration file

User Side

Framework

CLAS12 user submits a configuration file from terminal

Config. file written as record to SQL database

In Progress

Example Config. File

```
group:  rgaDIS                          # project description
user:  mungaro                          # username
nevents:  10000                         # number of events each job
generator: clasdis                      # one of clasdis, dvcs, disrad
genOptions:  --t 20 25                  # generator option
gcards: /jlab/work/clas12.gcard         # gcard gemc option
jobs: 100                               # number of jobs
project:  osg.clas12MC                  # OSG project
luminosity:  0                          # percent of 10^35 luminosity
tcurrent: -100                          # percent of torus current
pcurrent: -100                          # percent of solenoid current
```

Record parsed into submission files, passed to HTCondor on SubMIT

In Progress

Simulation runs and returns output files

Working

Output files returned to a specifed directory

Next Steps

- Configuring other pools (EAPS, Holyoke)
- Optimizing events-per-job distribution
  - E.g. users asks for 10M events, what is the optimal number of jobs and events / core we should run?
- Organizing return of large simulation output data
- Establishing a web interface for job submission
- Using events from existing, user generated files

Outlook

- **MIT Group**
  - **Work covered here has occurred in last 3 weeks**
  - **Will continue work as described**
- **CLAS12 Collaboration & JLab**
  - **Configuring DB work in progress**
- **Resources Requested from DOE Nuclear Physics**
  - **1 full-time experienced MIT post-doc**
  - **0.5 FTE in MIT-LNS Research Computing Group**
  - **$50,000 for hardware:**
    - **Servers**
    - **Hard drives**
    - **Ethernet ports**

Other Notes

Software Demonstration on Friday

- 2:30 PM Friday, 3/8 – F113
- Updates on CLAS12 Docker usage
- Demonstration of current SubMIT software for job submission on OSG

Acknowledgements

- Christoph Paus, Max Goncharov, Benedikt Maier for ongoing help on using subMIT
- Marty Wise for ongoing help with JLab DB

Backup Slides

- **Backup Slides**

Backup Slides – SubMIT Monitoring

[robertej]