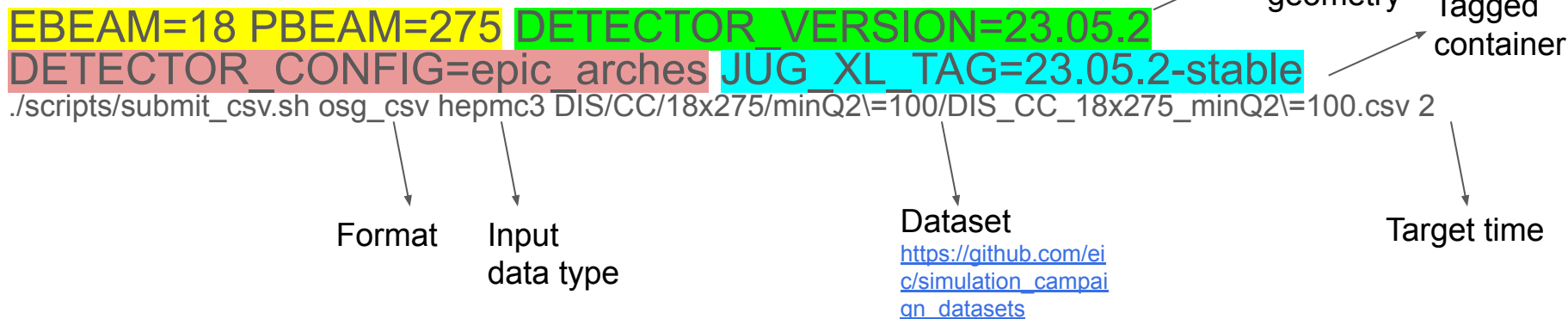# General Updates for May 2023 Train

Sakib Rahman
1 June 2023

Conductor's Log Book

- Created account on Open Science Grid and subscribed to EIC project
- Connected to osg login node login.collab.ci-connect.net
- Cloned https://github.com/eic/job_submission_condor
- Within the job submission folder the following script is executed for each dataset provided in the form of csv files
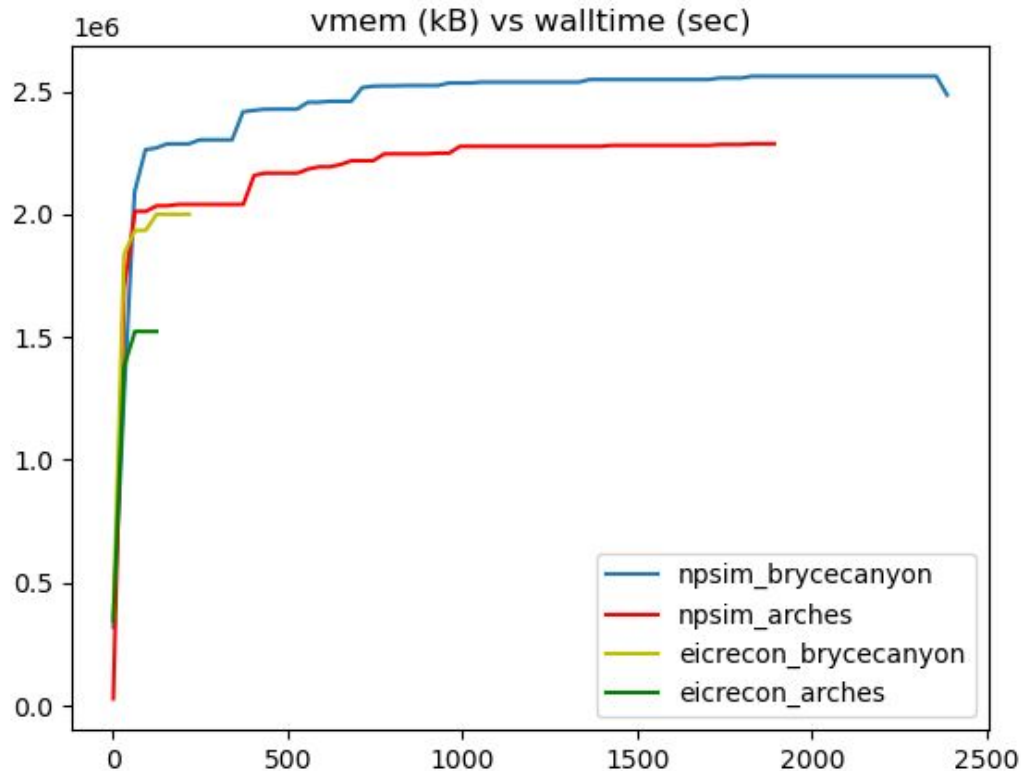
EBEAM=18 PBEAM=275 DETECTOR_VERSION=23.05.2 DETECTOR_CONFIG=epic_arches JUG_XL_TAG=23.05.2-stable

./scripts/submit_csv.sh osg_csv hepmc3 DIS/CC/18x275/minQ2\=100/DIS_CC_18x275_minQ2\=100.csv 2

Tagged geometry

Tagged container

Format

Input data type

Dataset
https://github.com/eic/simulation_campaign_datasets

Target time

- Jobs when submitted goes into idle queue before running. Failed jobs go into hold queue. Use hold_review_and_release.sh to monitor reasons for common job failures and resubmit. To release all jobs on hold at once use condor_release <username>
- Running both epic_arches and epic_brycecanyon in May train. Should be the case until we converge to a single config.
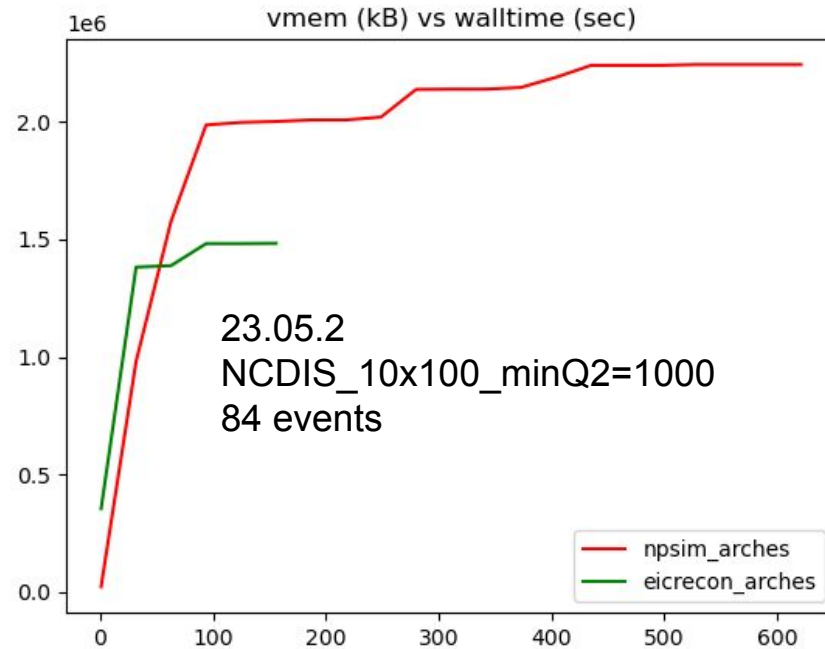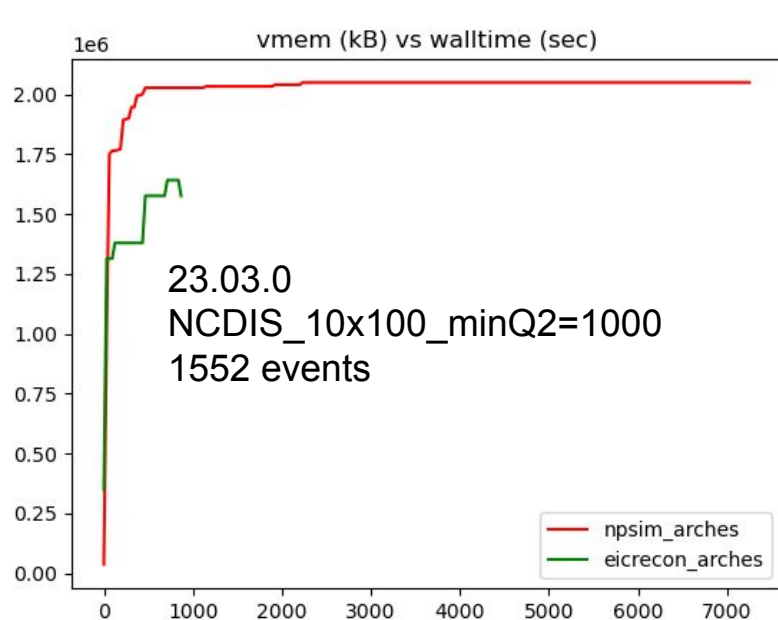
Conductor's Log Book

- To view the output from the campaign in the S3 storage:
  - Install minio client
  - Do this ONCE: ./mc config host add S3 https://eics3.sdcc.bnl.gov:9000 eicS3read eicS3read
  - Then you can use ./mc ls S3/eictest/EPIC or ./mc tree S3/eictest/EPIC to view content
  - Xrootd mirroring to jlab started. Location: /work/eic2/EPIC
- For the majority of the jobs we only store RECO output. For a small number of datasets, we are storing FULL sim results (roughly 10000 events). So, far we have been submitting duplicate jobs with a modified template that stores full sim output. Deprioritize using condor_prio command after first 200 jobs in the array are queued in RUN status. Then kill the jobs when the RUN jobs reach DONE status.
- Analyze process monitor logs (https://github.com/HSF/prmon) to understand variation in resource usage between campaigns
- Web summary page under construction
- Documentation under construction

Brycecanyon vs Arches in 23.05.2

vmem (kB) vs walltime (sec)

Legend:
- npsim_brycecanyon
- npsim_arches
- eicrecon_brycecanyon
- eicrecon_arches

NCDIS_18x275_minQ2=1
191 events

Arches in 23.03.0 vs Arches in 23.05.2

vmem (kB) vs walltime (sec)

23.03.0
NCDIS_10x100_minQ2=1000
1552 events

npsim_arches
eicrecon_arches

vmem (kB) vs walltime (sec)

23.05.2
NCDIS_10x100_minQ2=1000
84 events

npsim_arches
eicrecon_arches

Need to consider node variability

Default Data Sets

Reco output should be available for epic_brycecanyon for the following on S3.

DIS (pythia8):
    NC
    CC
EXCLUSIVE:
    DEMP
    DVCS_ABCONV
    TCS_ABCOV

Only ran DIS NC for epic_arches

Backgrounds

Just 1 merged background file available. Contains beamgas and synchrotron radiation.
Queued today. Output should be available at some point tomorrow on S3.

```
jug_xl+> rahmans1@login:~$ ~/bin/mc ls S3/eictest/EPIC/EVGEN/BACKGROUNDS
[2023-05-01 13:41:29 CDT]   17GiB STANDARD bgmerged_ep_noradcor.10x100_q2_10_100_run001_n_10000.hepmc
[2023-06-01 12:23:19 CDT]      0B BEAMGAS/
[2023-06-01 12:23:19 CDT]      0B SYNRAD/
```