General Updates for June 2023 Train

Sakib Rahman 15 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 or the jlab login node osg-eic through scilogin.jlab.org
 - ssh <username>@login.collab.ci-connect.net -Y
 - ssh <username>@scilogin.jlab.org -Y and then ssh osg-eic -Y
- Cloned https://github.com/eic/job_submission_condor and set up your S3 writing credentials
- Execute submit_csv.sh script 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

Dataset https://github.com/eic/c/simulation_campai gn_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>
- epic_brycecanyon is the default detector config starting in June. Additional configs need to be requested.

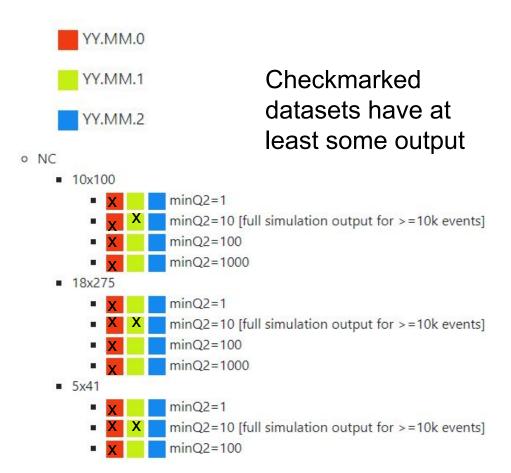
Conductor's Log Book

- To view the output from the campaign in the S3 storage:
 - Install minio client curl https://dl.min.io/client/mc/release/linux-amd64/mc --create-dirs -o ~/bin/mc
 - Do this ONCE: ~/bin/mc config host add S3 https://eics3.sdcc.bnl.gov:9000 eicS3read eicS3read
 - Then you can use ~/bin/mc Is 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
- <u>Documentation under construction</u>

Trains

Outputs should be available for epic_brycecanyon for DIS-NC (pythia8) on S3/eictest/EPIC/RECO/<version> and S3/eictest/EPIC/FULL/<version>

Results from 23.06.0 and 23.06.1



Charters/Taxis

Aiming for >=2M events for each SIDIS dataset for various beam configs

Both full and reco output available at S3/eictest/EPIC/RECO/23.06.1 and S3/eictest/EPIC/FULL/23.06.1 respectively for about 250 K events as of now

Results from 23.06.1

Unresolved Submit Errors

Some jobs staying perpetually in hold if they are restarted too many times. May be something do with condor config?

14:49:21 Job was held.

The job restarted too many times
Code 26 Subcode 0

Task List

- Create list of default datasets [Done]
 Create input preprocessing policy [Done pending final approval)
- 3) Xrootd and S3 mirroring work in progress
 - We are planning to start testing writing to xrootd directly
- 4) Sync condor configuration between ośg-eic.jlab.org (JĽAB) and login.collab.ci-connect.net (PATh) submit node