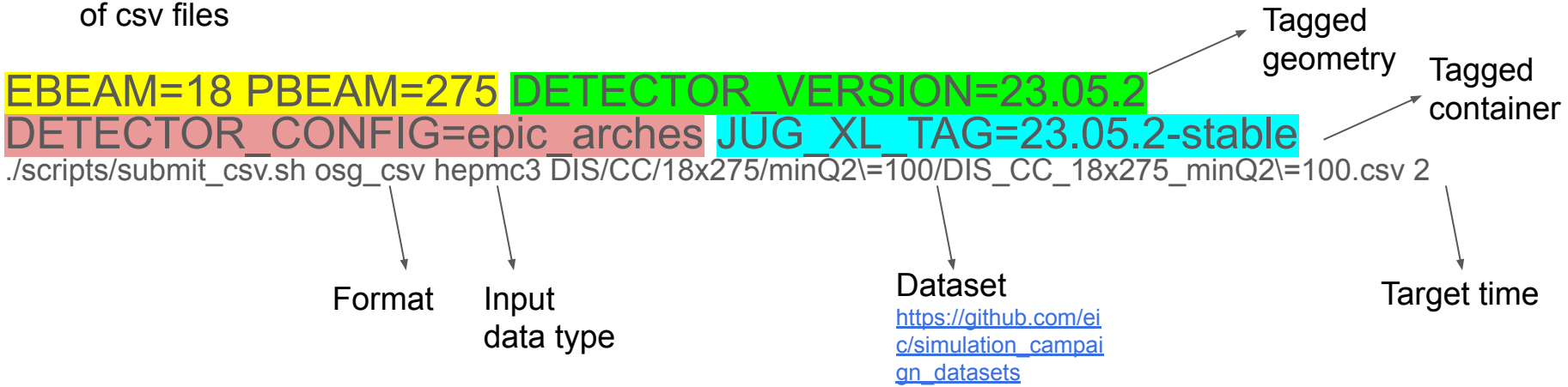
General Updates for May 2023 Train

Sakib Rahman
8 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



- 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 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

Results from 23.05.02

Reco output should be available for epic_brycecanyon for the following on S3/eictest/EPIC/RECO/23.05.2

DIS (pythia8):
- NC
- CC

EXCLUSIVE:
- DEMP
- DVCS_ABCONV
- TCS_ABCOV

SIDIS (pythia6):
- Noradcor

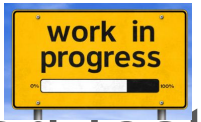
Backgrounds:
- Beamgas+Synrad (10k events)

```
jug_xl+> rahmans1@login:~$ ~/bin/mc tree S3/eictest/EPIC/FULL/23.05.2
S3/eictest/EPIC/FULL/23.05.2
├─ epic_arches
│  └─ DIS
│     └─ NC
│        ├─ 10x100
│        │  ├─ minQ2=10
│        │  └─ minQ2=100
│        ├─ 18x275
│        │  ├─ minQ2=10
│        │  └─ minQ2=100
│        └─ 5x41
│           ├─ minQ2=10
│           └─ minQ2=100
└─ epic_brycecanyon
   └─ DIS
      └─ NC
         ├─ 10x100
         │  ├─ minQ2=10
         │  └─ minQ2=100
         ├─ 18x275
         │  ├─ minQ2=10
         │  └─ minQ2=100
         └─ 5x41
            ├─ minQ2=10
            └─ minQ2=100
```

Only ran DIS NC for epic_arches

Outstanding Tasks

1) Create list of default datasets
2) Create input preprocessing policy (PR under discussion 

https://github.com/eic/epic-prod/pull/12)
3) Xrootd and S3 mirroring 

    - We are planning to start testing writing to xrootd directly