Simulation Campaign Update

Wouter Deconinck, for EPIC SimQA and CompSW
January 09, 2023

1 Slide Summary

- Latest reconstructed simulations exist on S3 at:
  - `S3/eictest/EPIC/RECO/22.11.2`
  - `S3/eictest/EPIC/RECO/22.11.3` (no geometry difference, just lower memory use)
- Physics jobs have been running for two weeks now, ~125k jobs
  - Primary platform for running simulations has been Open Science Grid
  - No significant operational issues due to software stability or performance
  - Reducing memory use below 2 GB allows for more than factor 2 increase in throughput
  - Interruption during Thanksgiving weekend due to S3 storage issues (impact of multiple days)
- Remaining jobs to be run:
  - fill the ‘holes’ in the production samples due to the S3 outage
  - djangoh DIS jobs: export to hepmc3 not possible, regenerating with djangoh 4.6.20
  - several sets exclusive events for which no hepmc3 files are available
- All brycecanyon will be rerun with working imaging calorimeter clustering

What is available? Number of files as of Jan 8, 2023

```
S3/eictest/EPIC/RECO/22.11.3/
├─ epic_arches
│  ├─ CI     4
│  ├─ DIS
│  │  └─ NC
│  │     └─ 5x41
│  │        └─ minQ2=100   1094
│  ├─ EXCLUSIVE
│  │  ├─ DIFFRACTIVE_JPSI_ABCONV
│  │  │  └─ Sartre
│  │  │     ├─ Coherent   9443
│  │  │     └─ Incoherent   3027
│  │  ├─ DIFFRACTIVE_PHI_ABCONV
│  │  │  └─ Sartre
│  │  │     ├─ Coherent   9192
│  │  │     └─ Incoherent   3215
│  │  ├─ DVCS_ABCONV
│  │  │  ├─ 10x100   1045
│  │  │  ├─ 18x275    985
│  │  │  └─ 5x41    453
│  │  ├─ TCS_ABCONV
│  │  │  ├─ 10x100
│  │  │  │  └─ hel_minus   2790
│  │  │  ├─ 18x275
│  │  │  │  ├─ hel_minus    390
│  │  │  │  └─ hel_plus    390
│  │  │  └─ 5x41
│  │  │     ├─ hel_minus    440
│  │  │     └─ hel_plus    440
│  │  └─ UPSILON_ABCONV     34
│  └─ SIDIS
│     ├─ Lambda_ABCONV   4489
│     └─ pythia6
│        ├─ ep_18x275
│        │  └─ hepmc_ip6
│        │     └─ radcor  58263
│        └─ ep_5x41
│           └─ hepmc_ip6
│              ├─ noradcor   9320
│              └─ radcor   7832
```

```
S3/eictest/EPIC/RECO/22.11.3/
└─ epic_brycecanyon
   ├─ DIS
   │  └─ NC
   │     └─ 5x41
   │        └─ minQ2=100   1094
   ├─ EXCLUSIVE
   │  ├─ DVCS_ABCONV
   │  │  ├─ 10x100   1045
   │  │  ├─ 18x275    985
   │  │  └─ 5x41    453
   │  └─ TCS_ABCONV
   │     ├─ 10x100
   │     │  └─ hel_minus   2790
   │     ├─ 18x275
   │     │  ├─ hel_minus    126
   │     │  └─ hel_plus    148
   │     └─ 5x41
   │        ├─ hel_minus    440
   │        └─ hel_plus    440
   └─ SIDIS
      ├─ Lambda_ABCONV   4492
      └─ pythia6
         ├─ ep_18x275
         │  └─ hepmc_ip6  51454
         │     └─ radcor  44740
         └─ ep_5x41
            └─ hepmc_ip6
               ├─ noradcor   9329
               └─ radcor   7856
```

Total number of files: 416377
Total size: 42 TB

```
S3/eictest/EPIC/RECO/22.11.2/
├─ epic_arches
│  ├─ DIS
│  │  ├─ CC
│  │  │  ├─ 10x100
│  │  │  │  ├─ minQ2=100   1917
│  │  │  │  └─ minQ2=1000
│  │  │  ├─ 18x275
│  │  │  │  ├─ minQ2=100   6166
│  │  │  │  └─ minQ2=1000   5682
│  │  │  └─ 5x41
│  │  │     └─ minQ2=100    884
│  │  └─ NC
│  │     ├─ 10x100
│  │     │  ├─ minQ2=1   2675
│  │     │  ├─ minQ2=10   1162
│  │     │  ├─ minQ2=100    545
│  │     │  └─ minQ2=1000    833
│  │     ├─ 18x275
│  │     │  ├─ minQ2=1   6969
│  │     │  ├─ minQ2=10   6792
│  │     │  ├─ minQ2=100   6628
│  │     │  └─ minQ2=1000   6602
│  │     └─ 5x41
│  │        ├─ minQ2=1   1259
│  │        ├─ minQ2=10   1105
│  │        └─ minQ2=100   1160
│  └─ SINGLE
```

```
S3/eictest/EPIC/RECO/22.11.2/
└─ epic_brycecanyon
   ├─ DIS
   │  ├─ CC
   │  │  ├─ 10x100
   │  │  │  ├─ minQ2=100   1920
   │  │  │  └─ minQ2=1000   1775
   │  │  ├─ 18x275
   │  │  │  ├─ minQ2=100   6169
   │  │  │  └─ minQ2=1000   5664
   │  │  └─ 5x41
   │  │     └─ minQ2=100    885
   │  └─ NC
   │     ├─ 10x100
   │     │  ├─ minQ2=1   2545
   │     │  └─ minQ2=10   1214
   │     ├─ 18x275
   │     │  ├─ minQ2=1   7019
   │     │  ├─ minQ2=10   6795
   │     │  ├─ minQ2=100   6634
   │     │  └─ minQ2=1000   6614
   │     └─ 5x41
   │        ├─ minQ2=1   1260
   │        ├─ minQ2=10    978
   │        └─ minQ2=100   1154
   └─ SINGLE
```

Operational details for this and future productions

Condor job scheduling:

- Using input events on JLab xrootd server:
  - HepMC3 WriterRootTree conversion
  - Read-only public access at dtn-eic.jlab.org
  - Used for up to ~10k simultaneously active TCP connections without any issues
- Using built-in S3 transfer output files
  - Jobs themselves unaware of S3
  - Single TCP connection at end of job
  - Reconstructed output at 25 kB/event; about 80 MB/s for 10k jobs at at 3 s/event
  - Working with OSG and Condor on S3 transfer resiliency to dtn01 and now eics3
  - Currently treating S3 output resiliency as operational limitation
- Load leveling settings:
  - max_jobs = 500, max_idle = 100

Other operational details: total of O(1M) jobs

- <1% failure rate for jobs, by frequency:
  - $ Failed hand shakes when reading xrootd (limited to certain sites: Syracuse U)
  - $$ OOM killer when > 2 GB; suspected config issue (limited to certain sites: UNL)
  - $$ Geant4 navigation errors in MRICH Fresnel lens
  - $$$ Stalls or 503 of S3 output file transfer
  - $$ podio::~MCParticles segfault in eicrecon
  - $$$ Jobs go awol and stop reporting
- Failed jobs resubmitted, complete on 2nd run (except for geant4 navigation error)
- Currently avoiding following OSG sites:
  - GPGrid, Crane, IU-Jetstream2-Backfill, SU-ITS-CE2, SU-ITS-CE3

Current production requirements

```
SIDIS: 412512 core-hours
 pythia6: 401949 core-hours
   10x100: 68942.1 core-hours
    noradcor: 68942.1 core-hours
   18x275: 299066 core-hours
    noradcor: 156833 core-hours
    radcor: 142233 core-hours
   5x41: 33940.2 core-hours
    noradcor: 19286.1 core-hours
    radcor: 14654.1 core-hours

EXCLUSIVE: 122426.7 core-hours
 DIFFRACTIVE_JPSI: 38154.8 core-hours
 DIFFRACTIVE_PHI: 55919.1 core-hours
 DVCS: 9731.32 core-hours
 TCS: 18555.3 core-hours

SINGLE (1M): 4635.79 core-hours
 3to50deg: 2124.77 core-hours
 130to177deg: 1301.02 core-hours
 45to135deg: 1210 core-hours

DIS: 259859 core-hours
 CC: 73743.2 core-hours
    10x100: 9246.4 core-hours
     minQ2=1000: 4243.85 core-hours
     minQ2=100: 5002.55 core-hours
    18x275: 29100.3 core-hours
     minQ2=1000: 13950.6 core-hours
     minQ2=100: 15149.7 core-hours
    5x41: 2197.1 core-hours
     minQ2=100: 2197.1 core-hours
    10x275: 26272.9 core-hours
     minQ2=1000: 12784.3 core-hours
     minQ2=100: 13488.6 core-hours
    5x100: 6926.45 core-hours
     minQ2=1000: 2716.51 core-hours
     minQ2=100: 4209.94 core-hours
 NC: 186116 core-hours
    10x100: 24992.8 core-hours
     minQ2=1000: 6742.12 core-hours
     minQ2=100: 5933.93 core-hours
     minQ2=10: 6121.88 core-hours
     minQ2=1: 6194.85 core-hours
    18x275: 65939.5 core-hours
     minQ2=1000: 16678.8 core-hours
     minQ2=100: 15881 core-hours
     minQ2=10: 16347.8 core-hours
     minQ2=1: 17031.9 core-hours
    5x41: 8606.16 core-hours
     minQ2=100: 2760.59 core-hours
     minQ2=10: 2706.75 core-hours
     minQ2=1: 3138.82 core-hours
    10x275: 63340.1 core-hours
     minQ2=1000: 15808.9 core-hours
     minQ2=100: 15312.9 core-hours
     minQ2=10: 15586 core-hours
     minQ2=1: 16632.3 core-hours
    5x100: 23237.6 core-hours
     minQ2=1000: 5778.19 core-hours
     minQ2=100: 5378.98 core-hours
     minQ2=10: 5865.21 core-hours
     minQ2=1: 6215.25 core-hours
```

TOTAL: 799434 core-hours = 91 core-years per configuration

Average OSG node running time (wall clock) about 25% longer than our benchmarks -> 114 core-years

Note: Not all data sets included (too many small exclusive datasets to list here); this is the dominant portion of it. A Djangoh EW DIS set of estimated 200k core-hours is the only unbenchmarked large part missing.

Lessons Learned: What took us so long?

Remember that we started from scratch

- Decision for software stack components only 6 months ago, August 2022 review
- Some components had a head start (some geometry descriptions from ATHENA)
- Other components developed from scratch

Main reasons for delay:

- Inconsistent treatment of units between python and C++ layer in EICrecon
  - E.g. 100 MeV specified in python, passed as ‘100’ to C++, interpreted as 100 GeV (where DD4hep/EDM4hep/Acts units)
- Imaging calorimeter reconstruction had been considered out of scope
- Unclear boundary between roles of DWGs (develop the reconstruction algorithms) and SWGs (develop reconstruction tools).

Lessons learned for future campaigns

- Enforce modularity; separate the reconstruction algorithms from the framework and its services
  - Share responsibilities between the detector and software experts
  - Validation of reconstruction algorithms
- Developing policies for software development to make it easier to develop software as a collaboration.
  - Bug fixes and new features in single development flow is not helping

Lessons Learned: What were other ‘quality of life’ issues?

Under our immediate control:

- Datasets in formats other than HepMC3 required more care than anticipated
  - Moving to HepMC3 ROOT trees for arbitrary forwards seeking via XRootD
  - **Several remaining event generators should be adapted to write HepMC3 files**, the standard in subatomic physics
  - Once in HepMC3, transformation to HepMC3 ROOT is a trivial conversion
- eic-shell container versioning adapted to handle Acts-20 roll-out
  - Acts-20 allows geometry specification independent of Acts version, big advance!
  - Container versioning aligned with geometry, e.g. 22.11-stable, 22.12-stable

Requiring interfacing with others:

- Reading events from XRootD server at JLab greatly improved management
  - Avoids need to copy input HepMC3
- Writing to S3 using condor functionality improved job stability
  - Worked through issues with OSG staff in their office hours
  - Bug fixes from us are flowing to htcondor
- Communication around infrastructure outages (e.g. S3) needs improvement
  - Discussed in ECCG with this in mind
  - Outage caused by personnel transition

Planning for the future

Simulation production campaign schedule for 2023:

- Large scale (full) production campaigns targeted for every quarter
    - Smaller campaigns in between
- Likely no duplication of these requirements for two configurations starting in April 2023
- Adoption of production workflow and scientific data management system in first or second quarter of CY 2023
- Likely increase in computational cost of simulations as fidelity increases

Why running jobs on OSG and not (only) BNL and JLab?

- Both host labs are providing computing resources are the order of 2k jobs slots dedicated to EIC. Why are we running on OSG?
- Until we have a dedicated discussion and decision on a production workflow and scientific data management framework, we decided to use the existing infrastructure with as few changes as possible (for well-defined workflow and to limit risk). For the single software stack that means running on OSG as we did during the proposal phase.
- Running on OSG gives us access to a level of computing that is an order of magnitude above what the host labs provide. In fact, we have to throttle simulation production because otherwise we risk overwhelming xrootd on the input side and S3 on the output side.