ePIC Production Working Group General Update

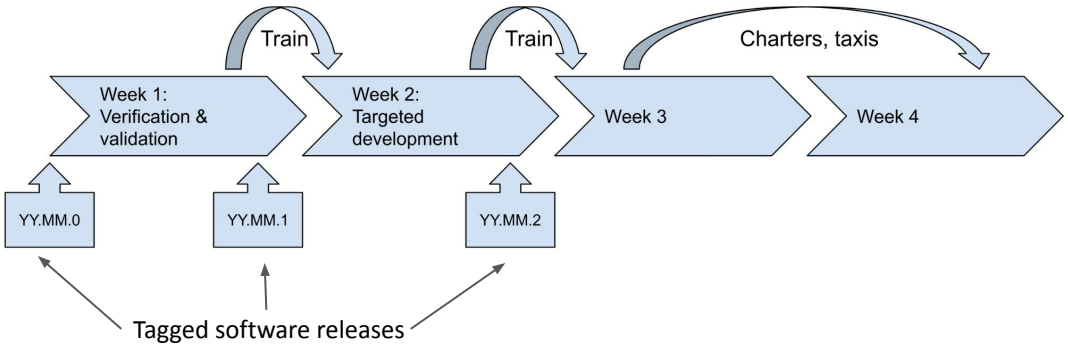
ePIC Production WG

Monthly simulation productions since May 2022

Simulation Campaign Strategy

Objectives

1. Achieve **continuous deployment** of the software used for detector and physics simulations
2. Ensure **regular updates** of simulation productions for detector and physics studies, and for geometry and algorithm development
3. Implement **timely validation and quality control** for simulation productions on datasets that require substantial time and resources



Train: Major central campaign at a fixed (monthly) schedule

Charter: Special interest runs for working groups

Taxi: Bespoke runs for individual users

See “Simulation Production Strategy” document

Production on OSG

- Central production occurs on the **Open Science Grid** (OSG)
  - Submitted primarily from JLab but can be submitted from any OSG submit host

- Jobs can run opportunistically or on dedicated nodes
  - Jobs can run internationally! (e.g **Digital Research Alliance of Canada**)

- Interface with the **PATh collaboration** (Partnership to Advance Throughput Computing) via biweekly meetings.

Latest Completed Production Campaigns (24.02.0, 24.02.1, 24.03.1)

For live campaign updates, follow the firehose mattermost channel.

To see what was run in past campaigns, review our campaign history pages for reconstructed output files and full geant4 simulation output files.

Update is provided at the end of the campaign via email/web. To learn how to access files on xrootd, review our FAQ page.

To learn about our simulation and analysis framework, review the tutorials on the ePIC Collaboration Landing Page.

Example:
Listing and viewing Reconstructed Outputs for a particular dataset

xrdfs root://dtn-eic.jlab.org ls
/work/eic2/EPIC/RECO/24.03.1/epic_craterlake/DIS/NC/18x275/minQ2=10

where the different segments indicate server address, base address, detector config, physics process and beam properties respectively.

The corresponding event generator files will be at
/work/eic2/EPIC/EVGEN/DIS/NC/18x275/minQ2=10
and geant4 simulation output (for small subset) will be at
/work/eic2/EPIC/FULL/24.03.1/epic_craterlake/DIS/NC/18x275/minQ2=10

Copy files locally using xrdcp or open them in root directly

Current Campaign (24.04.0)

We have organized a group of datasets (Physics Processes and Backgrounds) that will be run every campaign (moving gradually towards version control for everything). Besides regular production campaign datasets, we will accommodate charter requests from Physics Working Groups depending on availability of resources after approval by Analysis and Software coordinators (AC/SCs).

March Problems

- Low throughput from part of February through March
  - Ultimately lead to some uncompleted work

- A lot of debugging work behind the scenes
  - Big thanks to Kurt Strosahl (JLab) who dug in with many OSG folks over the course of a week

- Ultimately, it came down to a central collector issue confounded by some tokenization and authentication problems. Some suboptimal memory requests (3 GB vs the standard 2 GB per “slot” provisioning)
  - The memory request inefficiency likely understood. Expect fixes shortly…
  - Additional real life factors hampered diagnosis and repair

- The OSG is a complex set of systems enabling our large scale simulations across disparate compute resources
  - All these systems must work in unison
    - Taking lessons learned from March and implementing new systems/protocols to increase robustness

Simulation Campaigns for TDR

- Working with the PWGs to revise the r**eference list of physics processes** and r**elated MC samples** to be included in the simulation campaigns for the TDR:
  - Made good progress in meeting on Feb. 14: Preliminary list.
  - Continued discussion on March 13 with additional feedback on exclusive, diffractive, and tagging processes.

- Track progress on including related MC samples and their versioning in Production WG meeting on April 11.

Reminder: Criteria for MCEGs to be Included in Production

1) Must not duplicate effort. Need to have reference generator for each process.
2) Must be in hepmc3.tree.root format.
3) Must be version-tracked in a publicly accessible repository: Source code, steering files, run cards, etc. Follow the input preprocessing guidelines.

File Nomenclature and Organization

Organization of files	Example
<physics processes>/<generator repository release tag>/<electron momentum>x<proton momentum>/q2_<minimum q2>to<maximum q2>/<generator repository release tag>_<physics processes>_<electron momentum>x<proton momentum>_q2_<minimum q2>to<maximum q2>_run<index>.hepmc3.tree.root	DIS/NC/pythia6.428-1.0/10x100/q2_10to100/pythia6.428-1.0_DIS-NC_10x100_q2_10to100_run001.hepmc3.tree.root

Upcoming Changes

- Follow Software & Computing Weekly Meetings for news.
  - Meetings include now updates on **major software stack changes, notable merges,** and other news from the WGs: [March 20], [March 27]
- Software release for April will be announced in Software & Computing Weekly Meeting on April 10. Stay tuned for exciting updates!

- Production WG is actively working on Rucio for scientific data management, i.e., make it easier to find specific simulations for detector or physics studies.

- We have a large-scale simulation effort for ePIC. We welcome help with running the many productions and automating the simulation campaigns.

Backup

Physics in Production (Unversioned)

Dataset	Expected Corehours	Generator
DIS NC 10x100 [minQ2=1, minQ2=10, minQ2=100, minQ2=1000] 18x275 [minQ2=1, minQ2=10, minQ2=100, minQ2=1000] 5x41 [minQ2=1, minQ2=10, minQ2=100]	 49k 98k 24k	Pythia8
DIS CC 10x100 [minQ2=100, minQ2=1000] 18x275 [minQ2=100, minQ2=1000] 5x41 [minQ2=100]	 21k 47k 7k	Pythia8
EXCLUSIVE TCS ABCONV [10x100 (hel minus), 18x275(hel minus/plus), 5x41(hel minus/plus)	 45k	?
EXCLUSIVE UCHANNEL PI0 UCHANNEL RHO	 1k 0.3k	?
EXCLUSIVE DVCS ABCONV	 14k	?

Physics in Production (Unversioned)

Dataset	Expected Corehours	Generator
SINGLES 3to50degrees(e-, e+, pi-, pi+, pi0, kaon-, kaon+, gamma, proton) 45to135degrees(e-, e+, pi-, pi+, pi0, kaon-, kaon+, gamma, proton) 130to177degrees(e-, e+, pi-, pi+, pi0, kaon-, kaon+, gamma, proton) etaScan(e-, mu-, gamma)	 29k 24k 23k 9k	

Physics in Production (Versioned)

Dataset	Expected Corehours	Generator
SIDIS 10x100 [q2_0to1] 18x275 [q2_0to1] 5x41 [q2_0to1] 10x275 [q2_0to1]	 75k 144k 46k ?	pythia6-eic 1.0.0
LAMBDA	?	LambdaGen pythia8.306-1.0
D0	?	D0Gen pythia8.306-1.0
DVMP	0.3k	DVMPdataset EpIC1.0.0-1.0
DVMP LAGER	?	LAGER
DEMP	22k	DEMPGen 1.0.0
EXCLUSIVE DIFFRACTIVE PHI ABCONV PHOTOPRODUCTION JPSI PHOTOPRODUCTION JPSI ABCONV	 51k ? ?	SARTREdataset sartre-1.39-1.0

Backgrounds In Production (Unversioned)

Dataset	Expected Corehours	Generator
MERGED (10x100)	0.06k	?

Backgrounds In Production (Versioned)

Dataset	Expected Corehours	Generator
PROTON BEAMGAS [275 GeV, 100 GeV]	33k	pythia8.306-1.0
ELECTRON BEAMGAS	?	?