



EIC Software Working Group Meeting 2021-07-28

The Software and Computing WG Conveners: Andrea Bressan (University of Trieste and INFN), Dmitry Romanov (Jefferson lab), Sylvester Joosten (Argonne National Laboratory), Whitney Armstrong (Argonne National Laboratory), Wouter Deconinck (The University of Manitoba)



Oversimplified software stack

DD4hep: Geant4 geometry, detector plugin library, wrappers to run Geant4

Juggler: Digitization and reconstruction software (based on Gaudi with Podio-based data model and ACTS for tracking)

Gaudi: Generic open project for building event processing frameworks. Enables modern task-based concurrent execution in a heterogeneous computing environment. Used by ATLAS and LHCb.

ACTS: Experiment-independent tracking toolkit (ACTS' geometry constructed from DD4hep via plugin)

Podio: Robust data model definition to cross the boundaries between the tools



DD4Hep handling geometry problem



DD4Hep community



Automated workflows at eicweb

GitLab server (eicweb.phy.anl.gov)

- continuous integration
- dedicated build cluster

Runs automatically on each user commit, executing workflows running multiple tests, benchmarks and analysis

Automated containers

Both Docker and Singularity images are created nightly or on demand (commit) providing:

- reproducibility,
- production level images
- latest updates for those working locally



Benchmarks, documentation, conterization





Calorimetry WG

Realistic HCAL Hybrid electron endcap calorimeter with crystal

- Ready to study impact of magnet on HCAL:
 - V Realistic HCAL geometry
 - Solenoid & Helmholtz
 - W HCAL clustering and energy calibration
- ECAL system well-developed:
 - Barrel ECAL:
 - Barrel SiW imaging calorimeter
 - Barrel hybrid SiW + WSciFi calorimeter
 - Electron-endcap ECAL:
 - Crystal calorimeter
 - Glass calorimeter
 - Wybrid electron endcap for baseline
 - Primize geometry implementation
 - WSciFi calorimeter
- Geometry:
 - Finalize/validate geometry for "baseline" setup

BECAL with support

Forward shashlik

calorimetry

Far-forward & Far-backward WGs



- FF being worked on by Alex Jentsch
 - Magnets and detectors updated to latest design
 - W Flip IR orientation to right-handed coordinate system - "big flip" being prepared, right now (tentative merge before Monday)

• FB

IR implementation

Ω

PID WG

- dRICH (Christopher Dilks, Chao Peng)
 - ✓ ► Base geometry plugin ready
 - mirror alignment issue being solved by Christopher
- mRICH (Murad Sarsour, Whitney Armstrong)
 - V received baseline realistic geometry (implemented in fun4All) from Murad
 - ✓ implement realistic detector in DD4hep
 - Fix issue with optical photons getting trapped in Fresnel lens
- DIRC (Grzegorz Kalicy, Dmitry Romanov)
 - Converted geometry (some polishing needed)
 - **W** Refactor the code for DD4hep
 - Make initial validation benchmark
- TOF (Zhenyu Ye)
 - LGAD implementation with realistic services
- GridPix (Sanghwa Park)
 - Maintain implementation

mRICH geometry (w/ optical surfaces) complete

PID delegate: Zhenyu Ye S&C WG contact: Dmitry Romanov¹¹

Tracking WG

Outer LGAD layer not part of the "0-0-0" setup

Barrel staves as in ITS2 TDR

- 🚧 silicon tracker
 - Material validation (Shujie Li)
 - Were Vertex layers to be changed to cylindrical geometry
 - Validate barrel geometry
 - X Support cones
- Image: Second Stress (Second Stress Second Se
 - 🔽 also have barrel µRWEL

V GEM

0

Disks are wedges with sensitive layer and average material backing. Needs better constraints from WG

> Tracking delegate: Matt Posik S&C WG contact: Sylvester Joosten ¹²

Reconstruction Status

V Calorimetry

0

- Algorithms
 - Simple Clustering
 - Island Clustering (2D)
 - V 2+1D Clustering
 - V Topological Clustering (3D)
- Clustering benchmarks

🚧 PID

- Algorithms
 - Fuzzy-K ring clustering
 - MRICH clustering (Murad)
 - M DIRC clustering
- X PID benchmarks

774 Far Forward & Far Backward

- X Integrate B0 with tracker
- Matrix transform for Roman Pot & OMD reconstruction
- X low Q2 tagger

🚧 Tracking

Algorithms

0

0

0

0

- Proof-of-concept working again!
- Yes Finalize tracking for baseline
 - Tracking benchmarks
 - X Incorporate B0
- Setup realistic vertex reconstruction
- **W** Tracking Benchmarks
 - Masic benchmarks working
 - Tracking with realistic background

🚧 Global

- Event builder (produces ReconstructedParticle)
 - V Dummy event builder to test reco chain
 - implementation of full realistic event builder
- Finalize data model
- Interval on full reconstruction

Ways to get involved NOW (from 🚧 to 👷 to 🔽)

- Develop test cases for reconstruction (e.g. <u>here</u> or <u>here</u>)
- Get familiar with running code locally using containers
- Implement material geometry (often without coding)
- Clearly document detailed geometry in issue tracker

Resource Utilization

Software & Computing Working Group						
Inclusiv	/e	Activel a TM. Neutral current location 3 bins in G2: 5-10, 10-100, 10-00, 10-20, 10-20, 10-30, 10-20, 10-20, 10-20, 10-5-1 May need fine limiting at deficient energies. Also dedicated samples for Q2+1 and with QED reliation withched off. Estimated local 100M	fast simulation pythia8	ful simulation far élt, prohably noed different dehictler setup		
ep Charge current sample, at 18x275		15M at 18x275	sarre as above	full simulation for all, probably need different detector setup 5.4 electrons for at		
SIDIS	christopher j dika@gmeil.com			4x100M (tame as generator level) ~4M generated ~3M (might be the same as standard SIDIS sample, but might want to try cut.		
		4M at top energy (18x100)		tul for all		
Address 2014 Venthing Group 1 hains and prostrational (1927), 153 10, 163 11, 1 hains and prostrational (1927), 153 10, 163 11, 1 hains and the second second second second second 1 hybrid hains and and prostrational to higher and and prostrational (1920), 100 100 100 100 100 1 hybrid hains and and prostrational for the prostration of the prostrational (1920) (1920), 100 100 100 1 hybrid hains and the prostrational (1920), 100 100 2 hybrid hains and the prostrational	tpoge@bit.gov	$\begin{array}{l} \text{shar Datasets} \\ -Q^{2}2<1 \rightarrow N=38400M (can be reduced to -5M assuming an efficient (flar) \\ +1 < CO_{2}<10 \rightarrow N=38400M \\ +1 < CO_{2}<10 \rightarrow N=38400M \\ +100 < Q^{2}<100 \rightarrow N=38400M \\ +100 < Q^{2}<100 \rightarrow N=3340M (NC) & 33420M (NC) & 33420M (NC) \\ +2Q^{2}>100 \rightarrow N=3340M (NC) & N=3340M (NC) & 3450M (NC) \\ +Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC) \\ +2Q^{2}>100 < N=1 > 3450M (NC) & 3450M (NC)$	Process at event-level anyples through that simu pikey Dophos)	- 012 + 1 -> N = 3x1M - 14 012 < 10 -> N = 3x1M - 11 = 012 < 10 -> N =		
Unlead with the point games at this point. We may request some particle gue samples in conjunction with detector WGs execution regime interesting receipt DVCB - 00 - BHI EpiC guerratio, standard energy withing (16:27), 10:305, 64(1)	Darie Schreiggangskacht sehntere Autografiselt			- 10 < 0.0 < 100 0 × 100 0 × 10 345M - 100 < 0.42 < 1000 > N = 342M (VC) & 342M (VC) - 0.12 > 1000 > N = 3x1M (NC) & 3x1M (OC) - 11M events per energy setting, more specific info		
	Vice praspon ac sk portas Junice por sposten gant gov. tornas Junice gov tornas Junice gov teores Junice gov repose track genet com			wit be added gradually for all channels – events per G2 region, etc		
The following estimates include 14 eta bins with 20 momentum bins per eta bin. Optimization of tracking devictor with charged eigns.					20 x 1.5M	
Trackin	a				12 x 3M	
servatione zerverse of an error Detector emailed a with electronic. The electronic electronic emailed and the electronic of the electronic simulations a 200 bins x 10k events Simulations events from generator lavel for physics states on hearthmain simulation. This attimute includes 2	9			12 x 3M	12 x 3M	
systems x 2 B-fields x 3 different simulations x 300 bins x 10k events						
PID Working Group Calorimetry Working Group	seguences and a					
nECAL, nHCAL			Pythia6 532 events, for 1x and Qr2 res		electrons, photons and places , 50% per everyy point	
Calorin	netry		Pythiali SNI events, for a and Q*2 rev		alectrone, photoes and places, 50 k per energy pe	
NECAL			Pythia8 5M events			
Far-Forward Working Group We simply worth need a ton. Our group is not simulate physics, but dirapity making sure the detector implementations work and produce sepected social contains and smaring. I have put distanger bounds on how neary particle guns events are needed to an how neary particle guns events are needed	ejentsch@bril.gov		0	1M	10 x 1M	
For the time being we do not need simulations of physics events with contrail detectors (bremsstrahlung is not visible there anyway), and will make only stand-alone runs.	and the second s					

From Bottom's Up Resource Estimates

- Identification of synergies
- Streamlining of run plans
- Prototyping and small runs

To Multi-TB Production Runs since June

• Primarily consumers: singles, (SI)DIS

Large scale ATHENA data productions

- Input: HepMC files preferred (mcconv developed for other formats)
- Full simulation with current detector model, all bells and whistles:
 - Typical: 0.25 to 3.0 s/event, <500MB RAM RSS, 30 kB to 750 kB output size/event
 - ScFi Barrel Ecal: example of how the impact of implementation on simulation timing was mitigated with judicious choice of intermediate non-physical volumes
 - Full ROOT files on S3 under <u>ATHENA/FULL/</u> (but likely only need reco files)
- Full reconstruction (<u>reconstruction_benchmarks/benchmarks/full</u>):
 - Calorimetry clustering (Ecal, ScFi, Hcal), tracking (up to inner tracker), RICH hits/digi
 - Reco ROOT files on S3 under <u>ATHENA/RECO/</u> and sci-xrootd.jlab.org
 - Working on jsroot and file browser support on sci-xrootd
- Full simulation: ~weekly repetition; reconstruction: every few days
- Written to work on any slurm batch system; performed at Compute Canada

compute | calcul

• Trial runs on OSG at the ~2k job scale for single particle events

Computing Resource Utilization Analysis

- **Single particle simulations** with General Particle Source CPU time per core, event size: both scale empirically as E^{0.85}
 - pi+: 20 ms/ev $\cdot E^{0.85}$, 3.4 kB/ev $\cdot E^{0.85}$
 - \circ neutron: 20 ms/ev \cdot E^{0.85}, 2.2 kB/ev \cdot E^{0.85}
 - \circ gamma: 17 ms/ev $\cdot E^{0.85}$, 5.7 kB/ev $\cdot E^{0.85}$
 - pi0: 17 ms/ev $\cdot E^{0.85}$, 6.1 kB/ev $\cdot E^{0.85}$
- Multi-particle simulations: Pythia8, DIS NC/CC Q² > 10 GeV²
 - 25 mRad, 5x41 GeV: 28.4 tracks/ev, 666 ms/ev, 186.1 kB/ev
 - 25 mRad, 18x275 GeV: 38.8 tracks/ev, 3010 ms/ev, 566.1 kB/ev
 - \circ Empirical scaling with s \cdot min(Q)
- Benchmarks on HS06 ~12 nodes
- Since these studies: slight increase in time per event in past weeks (scintillating fibers), to ~35 ms/ev · E^{0.85}

EIC AI/ML in ATHENA

- Current use of AI/ML
 - $\mathbf{V} = e/\pi$ PID with 3D shower profiles from imaging calorimeter in center barrel region.
- Near-term anticipated use:
 - MACTS: Track finding
 - PID: Pattern recognition in RICH, DIRCO
 - Calorimetry clustering (2D, 2+1D and 3D clustering)
 - M DNN-based fast simulation
 - M DNN-based detector optimization (Bayesian optimization)
 - M DNN-based reconstruction
- Implications on computing infrastructure:
 - Many exascale GPU accelerators, but lack of support in current software tools limited by IO/memory bandwidth

🗹 = working, 🚧 = in progress, 🗙 = planned

Full simulation/reconstruction team

Whitney Armstrong, Miguel Arratia, Wouter Deconinck, Sylvester Joosten, Jihee Kim, Chao Peng, Tomas Polakovic, Dmitry Romanov, Marshall Scott, Zhenyu Ye, Ziyue Zhang, Maria Żurek *...and a rapidly growing amount ATHENA collaborators!*

Software & Computing Conveners:

Whitney Armstrong, Andrea Bressan^(*), Wouter Deconinck, Sylvester Joosten, Dmitry Romanov