

Full simulation and reconstruction of multiple EIC concept-detectors

Status of the EIC Argonne Software Toolchain

Whitney Armstrong, David Blyth, Sergei Chekanov, Ian Cloët, Adam Freese,
Sereres Johnston, Mohammad Hattawy, José Repond
and the Argonne EIC Collaboration

Argonne National Laboratory
Funded by ANL LDRD

July 21, 2017



Outline

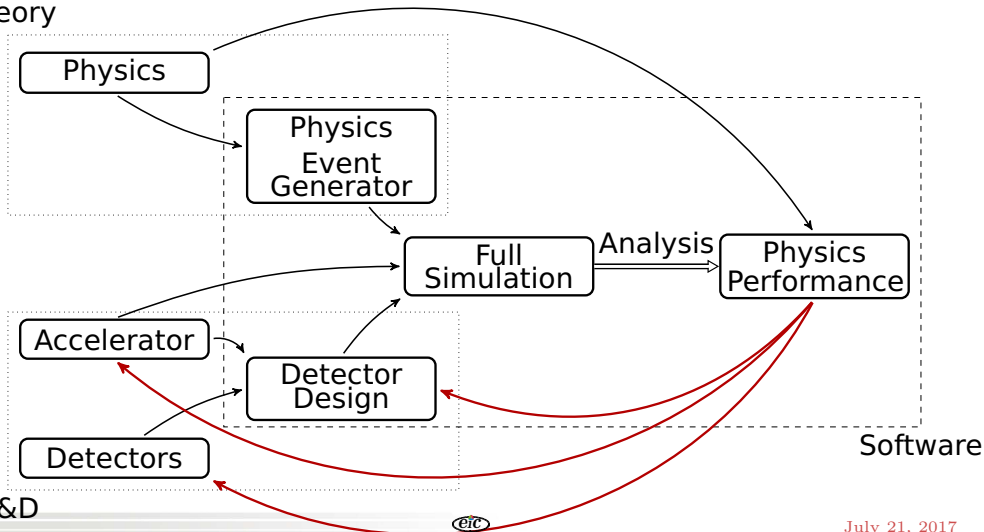
- 1 Overview of Argonne's EIC Effort
- 2 Introduction and Definitions
- 3 Critical Software Tools
 - HepSim
 - DD4hep and NPDet
 - Containers: Docker, Singularity, Shifter
- 4 Simulation and Reconstruction Data-Flows
- 5 Community Software for the Future



EIC effort at Argonne supported through ANL LDRD

Effort currently dominated by simulation and reconstruction software

Theory



Outline

- 1 Overview of Argonne's EIC Effort
- 2 Introduction and Definitions
- 3 Critical Software Tools
 - HepSim
 - DD4hep and NPDet
 - Containers: Docker, Singularity, Shifter
- 4 Simulation and Reconstruction Data-Flows
- 5 Community Software for the Future



Full Simulation and Reconstruction Tasks

Basic tasks:

- ① **Event Generation** - Produce the simulation input events
- ② **Detector Simulation** - Particle transport through detectors (Geant4)
- ③ **Digitization** - Turn *Sim Hits* into realistic hits
- ④ **Reconstruction** - Track, vertex, and trajectory reconstruction
- ⑤ **Performance Analysis** - Collection of benchmark analyses used to tune the overall design

EIC Argonne Software Toolchain (EAST)

Key software tools:

- HepSim
- DD4hep
- SLIC → ddsim which is part of lcgeo[†]
- lcsim

[†] work by D. Blyth highlights the usefulness of a **fixed data model**

The best software framework is ...

No Framework!

- We have collected and identified useful tools
- Pick and choose what works best for you
- Caveat: DD4hep is unavoidable (for good reasons)

more on this later...

Outline

- 1 Overview of Argonne's EIC Effort
- 2 Introduction and Definitions
- 3 Critical Software Tools
 - HepSim
 - DD4hep and NPDet
 - Containers: Docker, Singularity, Shifter
- 4 Simulation and Reconstruction Data-Flows
- 5 Community Software for the Future



HepSim

A database and interface for detector simulations

<http://atlaswww.hep.anl.gov/hepsim/>

- Developed at ANL by S. Chekanov
- **Generic tool** used for various projects: LHC, FCC, CLIC, ILC, and **now EIC**

HepSim is an excellent collaboration platform

A documentation and organization tool

- Stores and documents input MC data (produced from HPC or simple EG).
- Stores and documents **full simulation**
- Stores input configurations (geometry, digitization and reconstruction parameters)
- Stores tagged container of full software tool-chain (**New**)

HepSim - Generated Events

The screenshot shows the HepSim website interface. The top navigation bar includes links like 'Get involved', 'Full Search', 'Experiments', 'Manual', 'Mirrors', 'Tools', 'About', and 'Login'. The main header features the 'HepSim' logo and the tagline 'Repository with Monte Carlo simulations for particle physics'. A sidebar on the left lists various collision types and energies, such as $p \rightarrow p$ at 8 TeV, 13 TeV, 14 TeV, 27 TeV, 33 TeV, and 100 TeV, as well as $e^+ \rightarrow e^-$ at 250 GeV, 380 GeV, 500 GeV, 1 TeV, and 3 TeV. The main content area displays information for the 'gev35ep_lepto6ard_dislowq2' dataset, including its name, collision type (e-p), CM energy (0.035 TeV), entry ID (276), topic (SM), generator (LEPTO/ARIADNE), calculation level (LO+PS+hadronisation), process (DIS events at $Q^2 > 1 \text{ GeV}^2$ and $W^2 > 4 \text{ GeV}^2$), total events (25000000), number of files (500), cross section, luminosity, format (ProMC), download URL, status (Available), mirrors, EVGEN size (15.896 GB), tags, and simulation details for both fast and full simulations. The fast simulation is 40.46 GB, while the full simulation is 499 / 12.53 GB. The website also features a search bar and a list of recent updates in a box on the right.

atlaswww.hep.anl.gov/hepsim/info.php?item=276

Get involved Full Search Experiments Manual Mirrors Tools About Login

HEP.ANL.GOV

HepSim

Repository with Monte Carlo simulations for particle physics

Information about "gev35ep_lepto6ard_dislowq2" dataset

Name: gev35ep_lepto6ard_dislowq2
Collisions: e-p
CM Energy: 0.035 TeV
Entry ID: 276
Topic: SM
Generator: LEPTO/ARIADNE
Calculation level: LO+PS+hadronisation
Process: DIS events at $Q^2 > 1 \text{ GeV}^2$ and $W^2 > 4 \text{ GeV}^2$
Total events: 25000000
Number of files: 500
Cross section (σ): $4.376\text{E}+05 \pm 1957.1871 \text{ pb}$
Luminosity (L): 57.1245 pb^{-1} (or) 0.0571 fb^{-1} (or) $5.712\text{E}-05 \text{ ab}^{-1}$
Format: ProMC
Download URL: http://mc1.hep.anl.gov/web/hepsim/events/ep/35gev/lepto6ard_dislowq2/
Status: Available
http://eicsim01.jlab.org/hepsim/events/ep/35gev/lepto6ard_dislowq2/
http://mc.hep.anl.gov/asc/hepsim/events/ep/35gev/lepto6ard_dislowq2/
http://portal.nersc.gov/project/m1758/data/events/ep/35gev/lepto6ard_dislowq2/
EVGEN size: 15.896 GB
Tags:
Fast simulation:
Full simulation:
Fast/Full size: 40.46 GB
Record slimmed: No
Events weighted: No
Submission time: Wed May 17 16:30:14 CDT 2017
Updated on: Thu Jun 29 13:11:58 CDT 2017

• Jun.29, 2017: rfull058 tag with improved tracking strategy from D.Blyth
• Jun.20, 2017: rfull057 tag with alternative tracking strategy from D.Blyth

rfull058 | Info
499 / 12.53 GB
06/28/2017

rfull057 | Info
484 / 15.50 GB
06/19/2017

rfull056 | Info
496 / 12.43 GB
05/17/2017

- Stored MC events and description of how they were generated
- List of **all full simulations and reconstructions**
- Full EG details and configuration stored
- MC Events stored for repeated use (systematic comparisons, optimization, ...)

HepSim - Detector Information

The screenshot shows the HepSim website interface. The browser address bar displays `atlaswww.hep.anl.gov/hepsim/detectorinfo.php?id=sieic5`. The website has a navigation bar with links: [Get involved](#), [Full Search](#), [Experiments](#), [Manual](#), [Mirrors](#), [Tools](#), [About](#), and [Login](#). The main content area is titled **HepSim** and **Repository with Monte Carlo simulations for particle physics**. It features a sidebar on the left with buttons for various particle physics processes and energies, such as $p \rightarrow p$ (8 TeV, 13 TeV, 14 TeV, 27 TeV, 33 TeV, 100 TeV), $e^+ \rightarrow e^-$ (250 GeV, 380 GeV, 500 GeV, 1 TeV, 3 TeV), and $\mu^+ \rightarrow \mu^-$ (1 TeV, 5 TeV, 10 TeV, 20 TeV, 40 TeV). The main content area displays information about the **"sieic5" detector**. It includes a **Summary** section with details like Name (sieic5), Level (Geant4 simulation and full event reconstruction), Description (summary), Calibrations (view), Tracking (view), 3D View (eye icon), GeoManager (GeoManager icon), and Last modified (May 16, 2017). There is a **Reconstruction tags** section with a tag list: `rfull056`, `rfull057`, `rfull058`, and `rfull059`. Below that is a **Detector geometry files** section listing files for HEPREP, GDML, JSON, LCDD, and Pandora. A **Full detector** section provides a download link for `sieic5.zip`. A 3D visualization of the detector geometry is shown on the right, with axes labeled x , y , and z . The x -axis ranges from -533 to 533, and the y -axis ranges from -433 to 433. The detector is depicted as a central green rectangular region surrounded by blue and red layers, all within a larger grey rectangular frame. The text 'SiEIC5' is visible above the 3D visualization. At the bottom of the page, there is a 'Back' button and the EIC logo.

- Full detector geometry available in a few formats
- Calibration and reconstruction algorithm configuration parameters
- List of full reconstructions using detector

HepSim - Keeping track of reconstructions

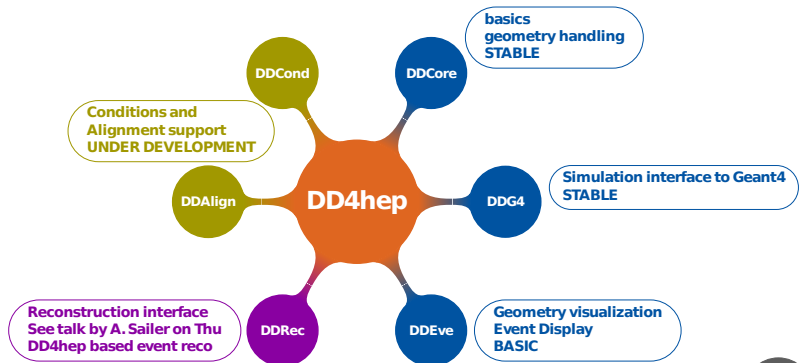
The screenshot shows the HepSim website interface. The browser address bar displays 'atlaswww.hep.anl.gov/hepsim/taginfo.php?id=rfull058'. The page title is 'HepSim' with the subtitle 'Repository with Monte Carlo simulations for particle physics'. A sidebar on the left contains navigation links for various particle physics processes and energies, including $p \rightarrow p$ (8 TeV to 109 TeV), $e^+ \rightarrow e^-$ (259 GeV to 3 TeV), $\mu^+ \rightarrow \mu^-$ (1 TeV to 40 TeV), $e^- \rightarrow p$ (318 GeV to 35 GeV), and 'Misc.'. The main content area is titled 'Information about the "rfull058" production tag'. It includes a 'Summary' section with details: Name: rfull058, Detector: sieic5, Description: SiEIC detector based on "all-silicon" concept, Experiment: EIC, Simulation: fpadsim-1.34 on OSG: SLIC version 5.0.1 with Geant 10.3p1, Reconstruction: fpadsim-1.34 on OSG: lcsim 3.4, slicPandora, pandoraPFA. Compared to rfull057, this version has improved track strategy developed by D.Blyth. In addition, speed of track reconstruction was increased by a factor 4-5 compared to the previous versions. Submission time: Tue Jun 20 10:00:55 CDT 2017, Updated on: -, Submitter: S.Chekanov, D.Blyth. An 'Available data' section shows a 'Find data:' button with a list icon and the tag 'rfull058'. A 'Comment' section contains a paragraph: 'This is a reduced version of SiD with 2.5T solenoid field, smaller ECAL/HCAL and muon system. This version is identical to SiEIC1, but min number of track hits was moved from 7 to 5, to get high efficiency in the forward. It has 25 layers of ECAL. This version was done using Geant 10.0.3p1. The Z size of this detector was extended by 1m by D.Blyth. Compared to rfull057, this version has improved track reconstruction (faster by a factor 3-4). This version is based on lcsim-3.4 Designed for ANL LDRD.' A 'Back' button is located below the comment section. At the bottom, a footer states 'HepSim: A data catalog with HEP predictions © 2014-2017. HEP Division, ANL, US Department of Energy' and the EIC logo is visible.

- Stores all information associated with the full simulation chain.
- Coming soon: A link to singularity/docker image with **exactly the same software** used for each simulation/reconstruction!

DD4hep

The result of a study from the *Advanced European Infrastructures for Detectors at Accelerators* ([EU AIDA 2020](#)) initiative.

Structure and packages

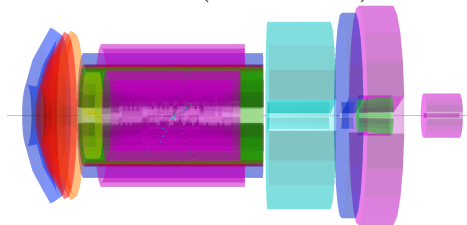


- Thoughtfully designed
- Interface to Geant4
- Single source of geometry
- Simple geometry hook for algorithm development
- Full concept detector described in human readable text file
- Can be easily used in a root script

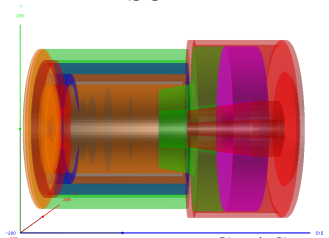
Nuclear Physics Detector Library (NPDet)

NPDet is a collection of parameterized detectors (using DD4hep) which can be used to construct full concept detectors in a single text file.

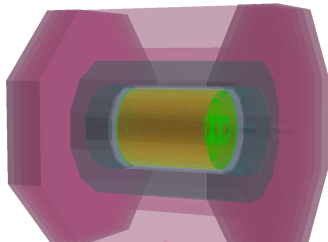
JLEIC (S. Johnston)



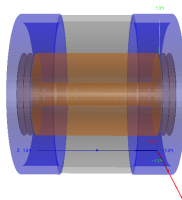
SOLID



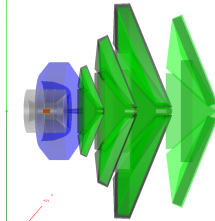
SiEIC



eRHIC



CLAS12



Software Containers

What are Containers?

- Containers are *Operating-system-level virtualization* (not a VM)
- Important concepts/lingo
 - “Image” : a snapshot of the software (thing which is downloaded or built)
 - “Container” : An instance of an “image” (disposable when done running).
- **docker** (under the hood and docker-hub for storage)
- **singularity** (distribution tool for ease of use)
- **shifter** (containers for HPC) - Currently investigating to exploit the coming **exascale computing at ALCF** (theta and aurora)

Simple example using singularity (pulling docker image):

```
$ singularity pull docker://argonneeic/evochain:rfull058
$ ./evochain_rfull058.img
```

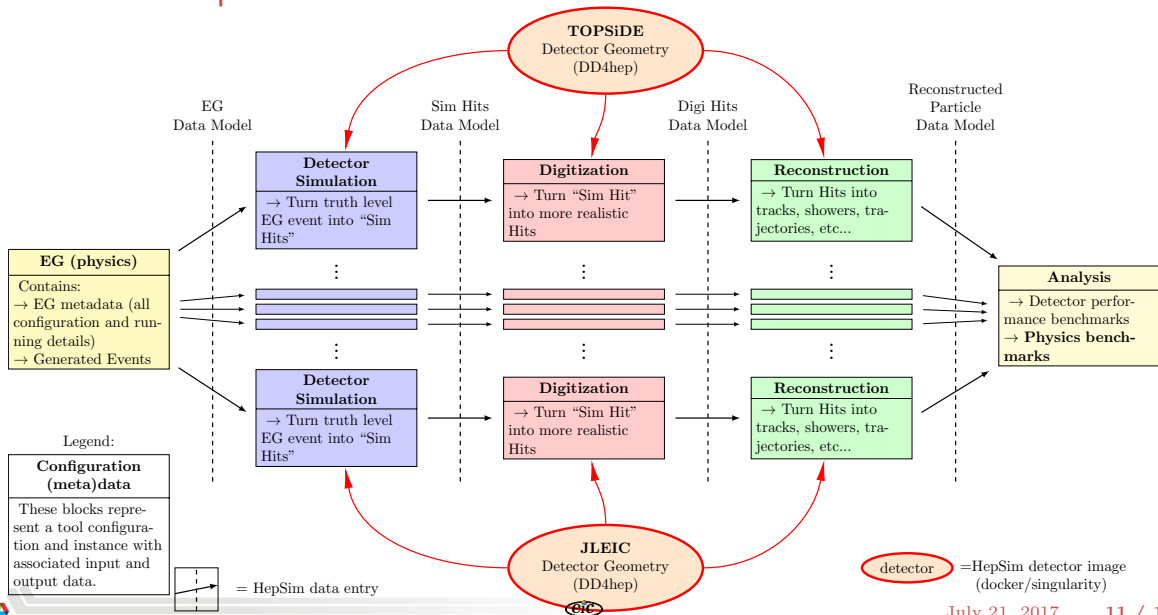
Now you are running an interactive shell in a container of the image with the **full software chain that was used now at your finger tips!**

Outline

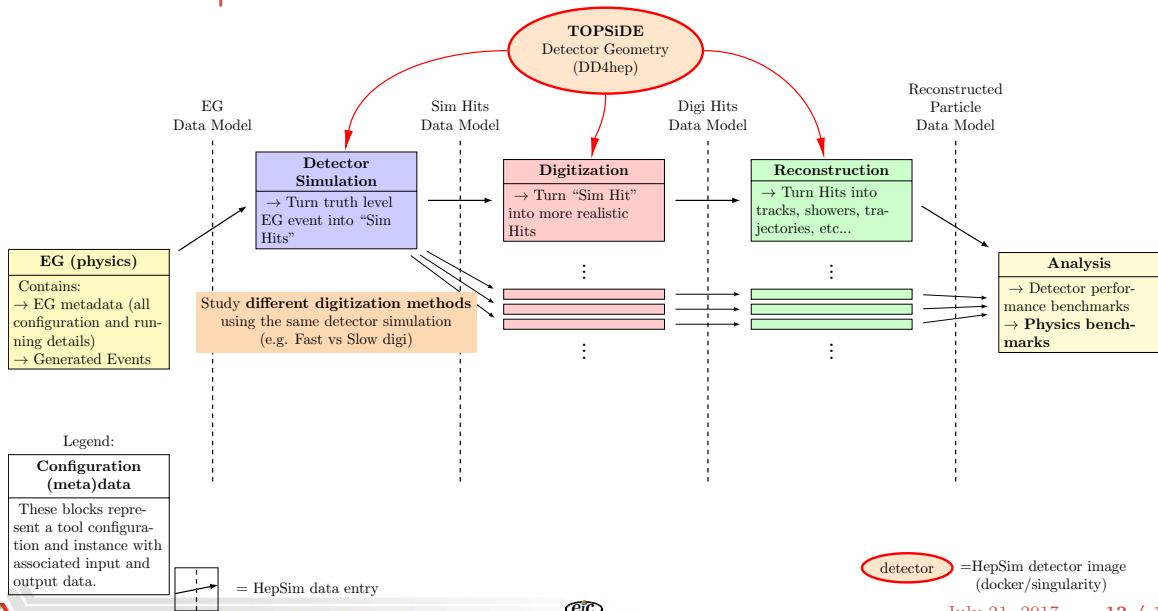
- 1 Overview of Argonne's EIC Effort
- 2 Introduction and Definitions
- 3 Critical Software Tools
 - HepSim
 - DD4hep and NPDet
 - Containers: Docker, Singularity, Shifter
- 4 Simulation and Reconstruction Data-Flows
- 5 Community Software for the Future



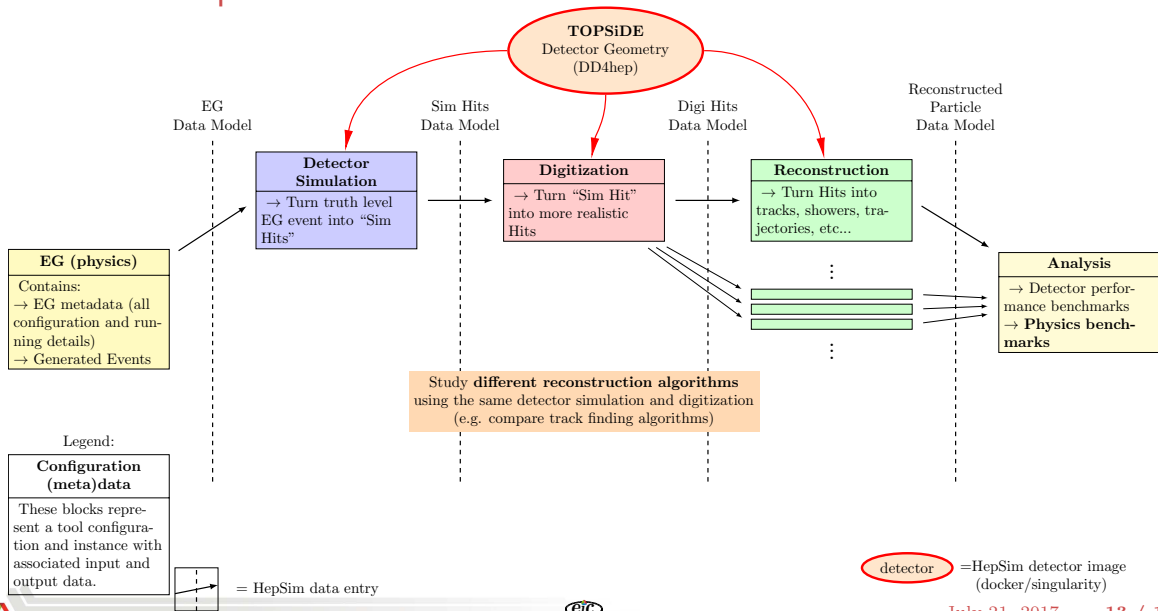
Data-flow Map



Data-flow Map



Data-flow Map



Outline

- 1 Overview of Argonne's EIC Effort
- 2 Introduction and Definitions
- 3 Critical Software Tools
 - HepSim
 - DD4hep and NPDet
 - Containers: Docker, Singularity, Shifter
- 4 Simulation and Reconstruction Data-Flows
- 5 Community Software for the Future



ESC - Geometry Description and Detector Interface

Requirements

- ✓ 1. The geometry information **should be the same in both simulation and reconstruction.**
- ✓ 2. Fast simulation systems should, as much as possible, **be able to use the common exchange format.**
- ✓ 3. The geometry system **should allow to include misalignment** and more general condition data.
- ✓ 4. (
- ✓ 5. (
- ✓ 10. Geometry exchange format should **allow clients to use a subset of the features clearly stating which are the optional ones.** We should support existing interested frameworks

EIC Software Consortium

✓ **EIC Argonne Software Toolchain meets all of the community's requirements**

- Prepared by EIC Software Consortium to identify EIC detector simulation software needs
- [ESC document link](#)

- ✓ 11. Some support for import from CAD should be foreseen.
 - ✓ 12. Geometry information **should have support for versioning.**
 - ✓ 8. I
 - ✓ 9. It should be possible to **change sensitivity attributes without changing other static aspects of the geometry.**
- case this is not possible , the additional libraries to manipulate hits should not depend on the simulation stack used to produce the hits.

On going and Future Developments

The Argonne Team is just getting started (we have actively working for roughly 6 months)

Currently working on

- Generic Track finding tool (S. Johnston)
- HepSim Container integration (D. Blyth)
- Event generator for deuteron observables (A. Freese, I. Cloët)
- Detector Benchmarks (M. Hattawy)
- Physics Benchmarks (**J. Repond - Next Talk**)

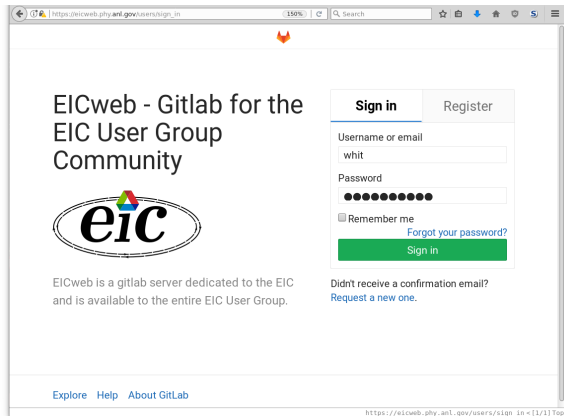


Please join us - All are welcome

- Weekly software meetings (Thursday afternoon)
- Bi-Weekly EIC meetings every (every other Friday)
- Create account and use the fast, eic-dedicated gitlab server (eicweb.phy.anl.gov)
- Subscribe to the mailing lists (email warmstrong@anl.gov to be added)

Links and References

- **HepSim**
- **EICweb** (eicweb.phy.anl.gov) - EIC dedicated gitlab server (publicly available to EIC UG)
- **Singularity** - install and start using EAST right away!
JLab users note: Singularity is already installed on the farm and all desktops.
- **DD4hep**
- **lcgeo**



DD4hep Presentations

Detector Simulations with DD4hep - Marko Petric

DD4hep Based Event Reconstruction - Andre Sailer

The FCC software: how to keep SW experiment independent - A. Zaborowska

Summary

EIC Argonne Software Toolchain (EAST)

- A **flexible, full, and easy-to-use** simulation and reconstruction toolchain is available now.
- **Collaboration tools** for the **EIC User Group** are available now.
- **We want to invite the entire EIC User Group to collaborate.**
 - **Contribute new EG data** (physics) – Let's see what detectors work best
 - Add detectors to **NPDet** detector library – Make your detector technology available
 - **Add reconstruction data** for a new concept detector
 - Write benchmarks (detector and physics) – Optimise your concept detector to physics
 - Suggest ideas for improvement! – **We want EIC UG feedback**

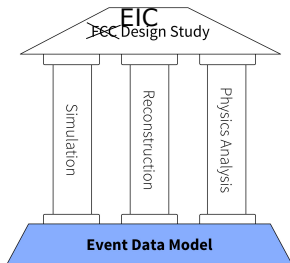
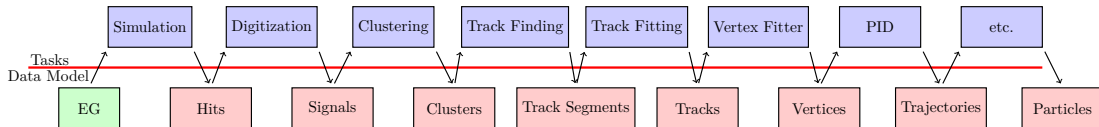
Please join us!

- Weekly software meetings (Thursday afternoon)
- Bi-Weekly EIC meetings every (every other Friday)
- Collaborate on eicweb (eicweb.phy.anl.gov)
- Communicate on mailing list (email warmstrong@anl.gov to be added)

Backup Slides



Why a Data Model?



The FCC software: how to keep SW
experiment independent - A. Zaborowska

- The **Data Model** is the boundaries of every task.
- A **Common** data model is the first step towards generic algorithms and tasks
- Challenge: Getting everyone to agree
- EAST initial data model: LCIO
- Note: *Data Model* does not mean *serialization tool*! It is just the data structures
- [podio](#) is a new tool which by default uses ROOT for serialization (new serialization libraries can be easily added)