EIC Software Meeting with Key4Hep

EIC Software: Key4Hep

Wednesday Sep 21, 2022, 11:00 AM → 1:00 PM US/Eastern

Benedikt Hegner (BNL), Markus Diefenthaler (Jefferson Lab)

Description We will discuss the collaboration with the Key4Hep project.

We will use **Zoom** for the remote meeting:

- https://jlab-org.zoomgov.com/j/1614875218?pwd=RFRPcGlNM3BaS0pQaDhxS3JURkdJZz09
- Meeting ID: 1614875218
- Password: 925723

Live Notes

11:00 AM → 12:00 PM **Status and Plan**

- 11:00 AM **EIC Software: Status and Plans** (15m)
  Speaker: Dr Markus Diefenthaler (Jefferson Lab)
- 11:15 AM **Discussion** (15m)
- 11:30 AM **Key4Hep: Status and Plans** (15m)
  Speaker: Benedikt Hegner (CERN / Stony Brook University)
- 11:45 AM **Discussion** (15m)

12:00 PM → 1:00 PM **Possible EIC-Key4Hep Projects**

Discussion Items

- 12:00 PM **Discussion** (1h)

OVERVIEW

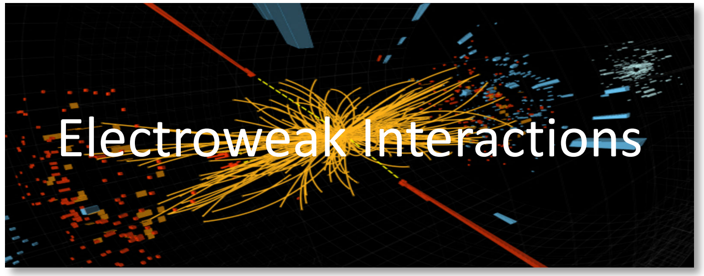
with a focus on community

Nuclear Physics

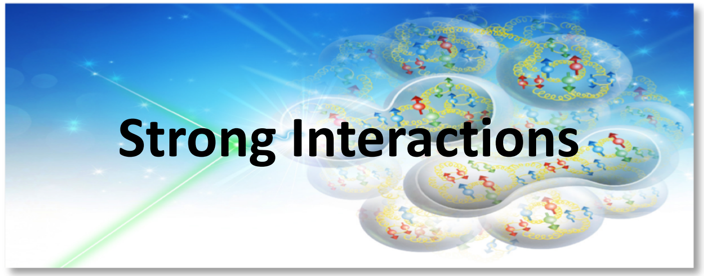
Further exploration of the Standard Model

Dark matter searches

Electroweak symmetry breaking



Deeper understanding of QCD



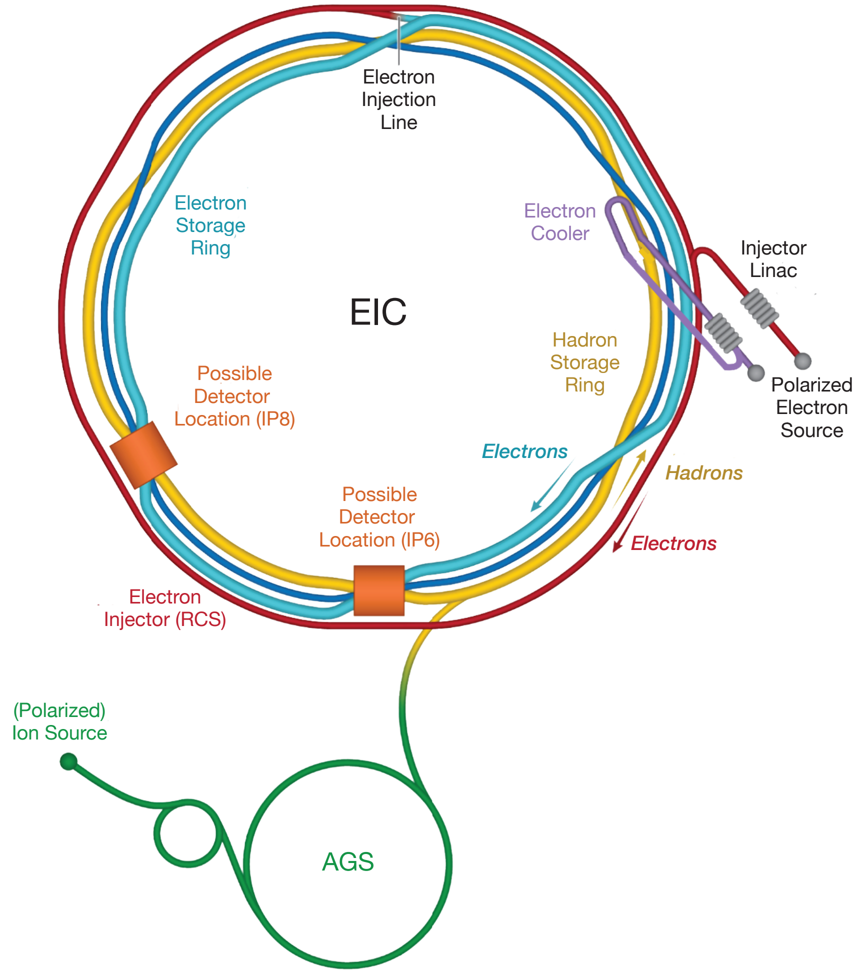
Mission of Nuclear Physics

- discover, explore, and understand all forms of nuclear matter

Frontiers in Nuclear Physics

- One of the enduring mysteries of the universe is the nature of matter—what are its basic constituents and how do they interact to form the properties we observe? The largest contribution by far to the mass of the matter we are familiar with comes from protons and heavier nuclei.
- Although the fundamental particles that compose nuclear matter—quarks and gluons—are themselves relatively well understood, exactly how they interact and combine to form the different types of matter observed in the universe today and during its evolution remains largely unknown.

The Electron-Ion Collider (EIC)



Frontier accelerator facility in the U.S.

- **World's first collider of**:
  - Polarized electrons and polarized protons,
  - Polarized electrons and light ions (d, ^{3}He),
  - Electrons and heavy ions (up to Uranium).
- The EIC will enable us to embark on a **precision study of the nucleon and the nucleus at the scale of sea quarks and gluons**, over all of the kinematic range that are relevant.
- Jefferson Lab and BNL will be host laboratories for the EIC Experimental Program. Leadership roles in the EIC project are shared.
- More than 1360 accelerator, experimental, and theoretical physicists from 267 institutions in 36 countries are part of the **EIC User Group**.
- The experimental community is moving towards the formation of **EPIC**, the first collaboration to support the realization of the EIC project detector.

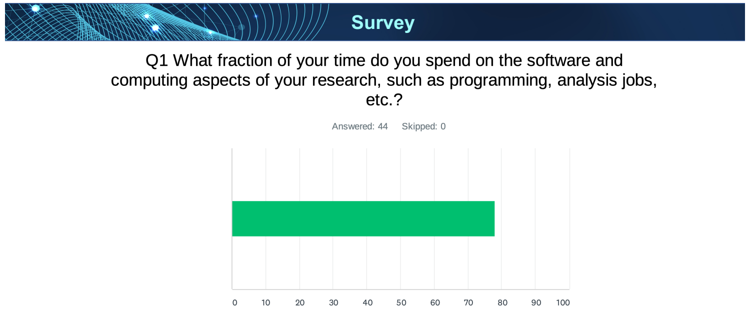
EIC Software Efforts

2016 – 2020	**EIC Software Consortium (ESC)**
2018 – now	**EICUG Software Working Group (SWG)**
2019 – 2021	**Yellow Report Initiative**
2021 – 2022	**Detector Collaboration Proposals**
2022 – now	**EPIC Computing and Software (CompSW) Working Group** **EPIC Simulation, Production, and QA (SimQA) Working Group**

Our Vision for Software & Computing at the EIC

“The purpose of computing is insight, not numbers.” Richard Hamming (1962)

Software & computing are an integral part of our research:



Survey among NP Ph.D. students and postdocs in preparation of ”Future Trends in NP Computing”

- **Goal** We would like to ensure that scientists of all levels worldwide can participate in EIC analysis actively.
- **User-Centered Design**: To achieve this goal, we must develop simulation and analysis software using modern and advanced technologies while hiding that complexity (turnkey) and engage the wider community in the development.

Rapid turnaround of data for the physics analysis and to start the work on publications:

- **Goal**: Analysis-ready data from the DAQ system.
- **Compute-detector integration** with AI at the DAQ and analysis level.

Community Document on Software Vision

EIC SOFTWARE:
Statement of Principles

1 We aim to develop a diverse workforce, while also cultivating an environment of equity and inclusivity as well as a culture of belonging.

2 We will have an unprecedented compute-detector integration:

- We will have a common software stack for online and offline software, including the processing of streamed data and its time-ordered structure.
- We aim for autonomous alignment and calibration.
- We aim for a rapid, near-real-time turnaround of the raw data to online and offline productions.

3 We will leverage heterogeneous computing:

- We will enable distributed workflows on the computing resources of the worldwide EIC community, leveraging not only HTC but also HPC systems.
- EIC software should be able to run on as many systems as possible, while supporting specific system characteristics, e.g., accelerators such as GPUs, where beneficial.
- We will have a modular software design with structures robust against changes in the computing environment so that changes in underlying code can be handled without an entire overhaul of the structure.

4 We will aim for user-centered design:

- We will enable scientists of all levels worldwide to actively participate in the science program of the EIC, keeping the barriers low for smaller teams.
- EIC software will run on the systems used by the community, easily.
- We aim for a modular development paradigm for algorithms and tools without the need for users to interface with the entire software environment.

5 Our data formats are open, simple and self-descriptive:

- We will favor simple flat data structures and formats to encourage collaboration with computer, data, and other scientists outside of NP and HEP.
- We aim for access to the EIC data to be simple and straightforward.

6 We will have reproducible software:

- Data and analysis preservation will be an integral part of EIC software and the workflows of the community.
- We aim for fully reproducible analyses that are based on reusable software and are amenable to adjustments and new interpretations.

7 We will embrace our community:

- EIC software will be open source with attribution to its contributors.
- We will use publicly available productivity tools.
- EIC software will be accessible by the whole community.
- We will ensure that mission critical software components are not dependent on the expertise of a single developer, but managed and maintained by a core group.
- We will not reinvent the wheel but rather aim to build on and extend existing efforts in the wider scientific community.
- We will support the community with active training and support sessions where experienced software developers and users interact with new users.
- We will support the careers of scientists who dedicate their time and effort towards software development.

8 We will provide a production-ready software stack throughout the development:

- We will not separate software development from software use and support.
- We are committed to providing a software stack for EIC science that continuously evolves and can be used to achieve all EIC milestones.
- We will deploy metrics to evaluate and improve the quality of our software.
- We aim to continuously evaluate, adapt/develop, validate, and integrate new software, workflow, and computing practices.

The "Statement of Principles" represent guiding principles for EIC Software. They have been endorsed by the international EIC community. For a list of endorses, see https://eic.github.io/activities/principles.html

EPIC Software

EIC community asks for commonality and one software stack.

The design of the **modular simulation and reconstruction toolkit** for the development of the EPIC detector and the EPIC science program is based on the **EIC Software: Statement of Principles** and a **decision-making process involving the wider EIC community.**

Simulation

MC Event Generators — **Detector Simulations in Geant4** — **Readout Simulation (Digitization)** — **Reconstruction in JANA2** — **Physics Analyses**

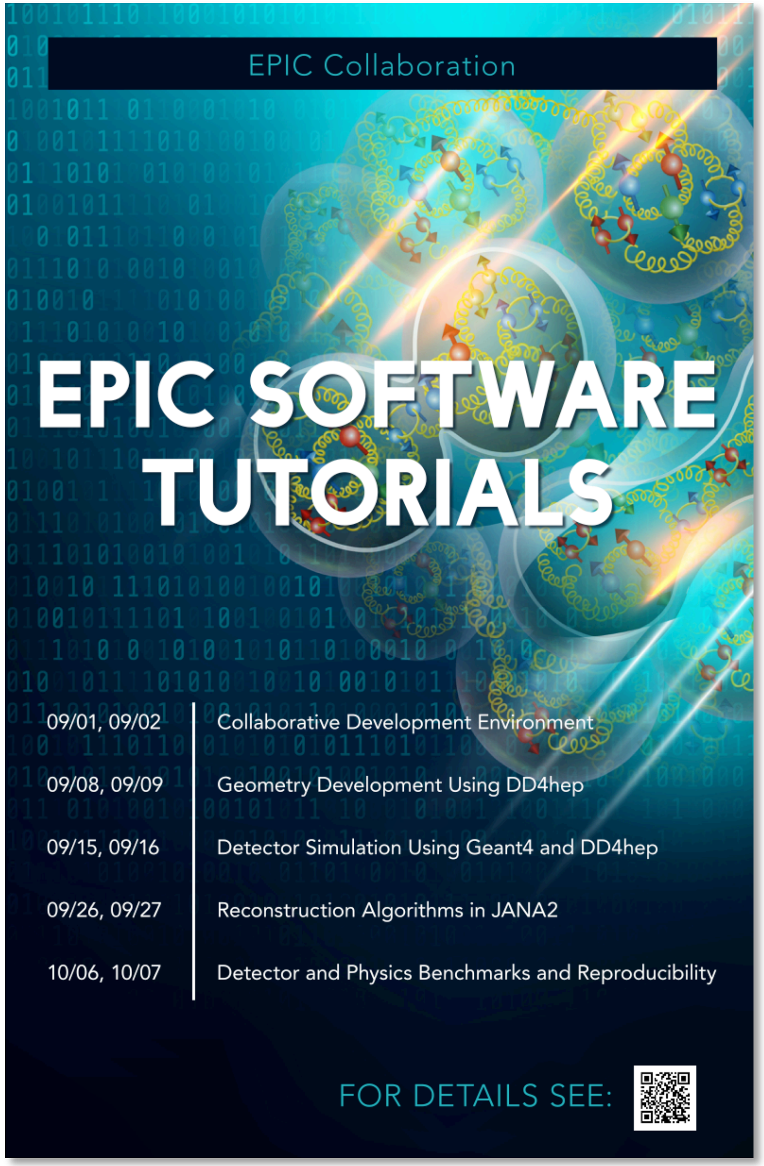
Edm4eic data model based on edm4hep and podio
Geometry Description and Detector Interface using DD4hep

Distributed Computing Approach for Large-Scale Simulations

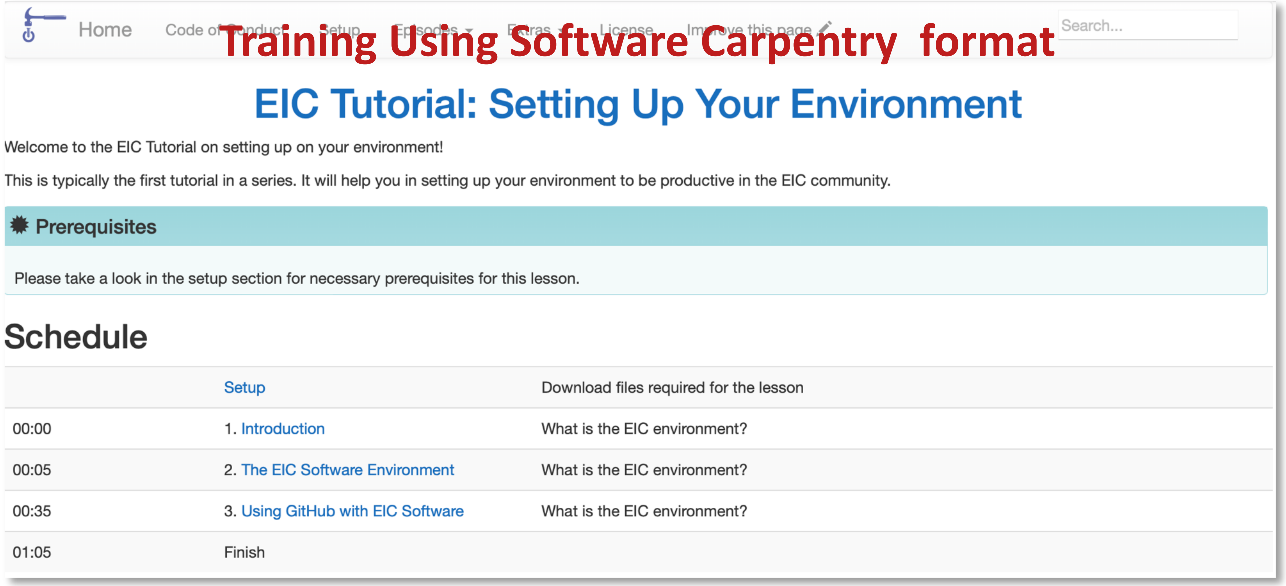
CI/eicweb for Detector and Physics Benchmarks and Reproducibility

Training the EPIC Collaboration

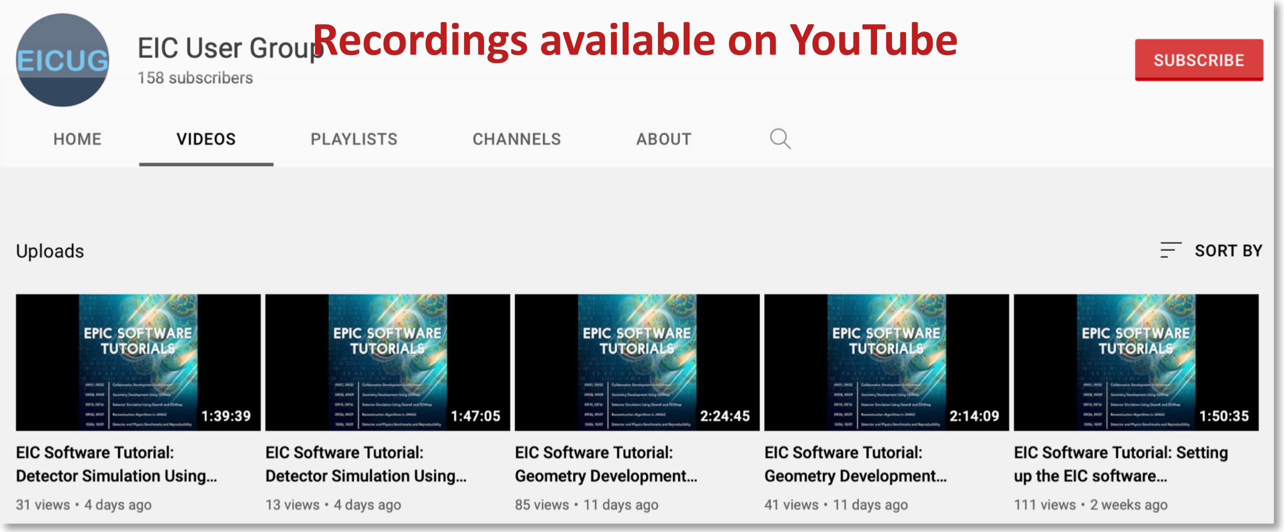
Driven by Wouter Deconinck



Training Using Software Carpentry format



Recordings available on YouTube



EPIC CompSW and EICUG SWG

EPIC CompSW

Sylvester's Picture of Salt & Pepper

EICUG SWG

- EPIC Collaboration and EIC project have strict timeline with well-defined deliverables from short to long term.
- This defines the scope and deliverables of EPIC CompSW.

- EICUG fosters the community.
- EICUG SWG is a forum for discussion and forward-looking projects.
- Ideal platform for cross-collaboration:
    - Experient and theory
    - EPIC and 2nd Detector
    - Interdisciplinary collaborations (data scientists, HEP, …)

EIC Software Community

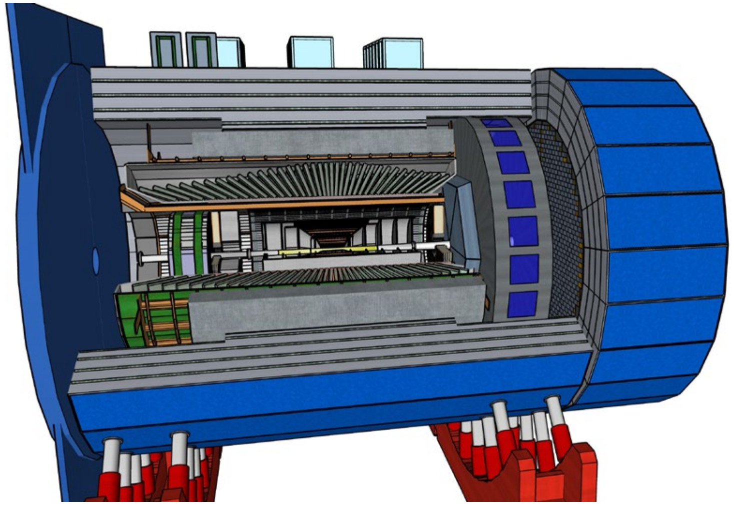
- Software in very early life stage:

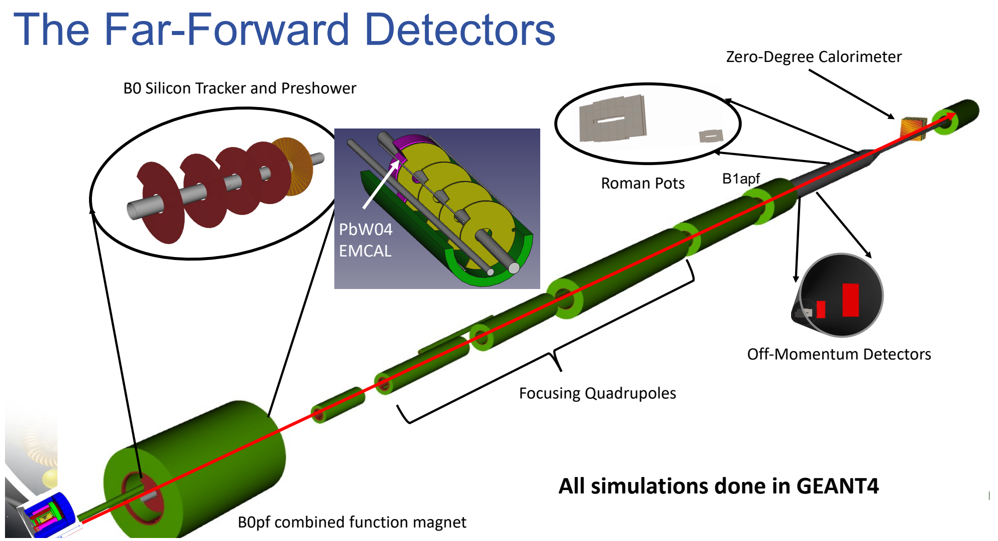
  8 We will provide a production-ready software stack throughout the development:
  - We will not separate software development from software use and support.
  - We are committed to providing a software stack for EIC science that continuously evolves and can be used to achieve all EIC milestones.
  - We will deploy metrics to evaluate and improve the quality of our software.
  - We aim to continuously evaluate, adapt/develop, validate, and integrate new software, workflow, and computing practices.

- **Focus on common software tools**
  - Avoid duplication of efforts.
  - Team up on challenges, e.g., heterogeneous computing.
- Work with community standards, e.g., HepMC3 for event generation.
- **Engage with the wider NHEP community**
  - AI/ML
  - **Detector collaborations**
    - EPIC and 2^{nd} detector efforts
  - Geant4 collaboration
  - HEP Software Foundation
  - **Key4Hep: ← Our Meeting Today**
    - Building up on ATHENA efforts
  - **Theory community**
    - MC event generators (MCEGs)

Priorities for Detector Design

- **Detector Simulations**
  - Validation of Geant4: Make test-beam setup and results available.
  - Detector design optimization using AI/ML.
  - Accelerate detector simulations:
    - Fast and accurate simulations using AI/ML, e.g., for simulation of calorimeters, Cherenkov detectors.
    - Fast simulations fully integrated into Geant4.
- **Reconstruction**
  - Accelerate reconstruction using AI/ML.
  - Reconstruction with far-forward detectors fully integrated.





Slide from Alex Jentsch

R&D Towards Next-Generation Detector Simulations

Detector Simulation

- **EIC focused project**
- **Turn-key application**
- **Built on top of Geant4 for full and fast simulations**
- **With library of potential detector option**

Requirements

- Ease of **leveraging new and rapidly evolving** technologies:
  - AI/ML to accelerate simulations
  - Heterogeneous architectures:
    - AI/ML is the best near term prospect for using LCF/Exascale effectively.
- Ease of switching detector options
- Ease of switching between detailed and coarse detector descriptions

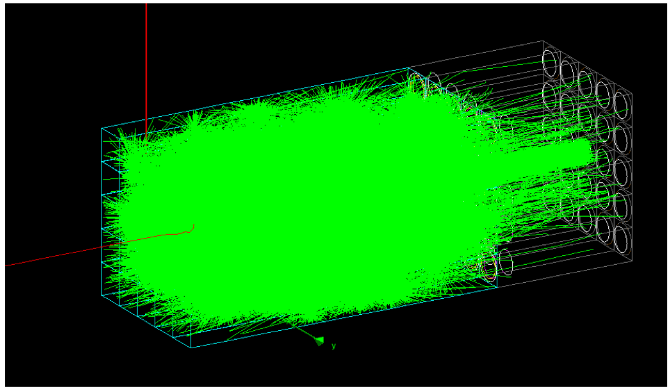
Project

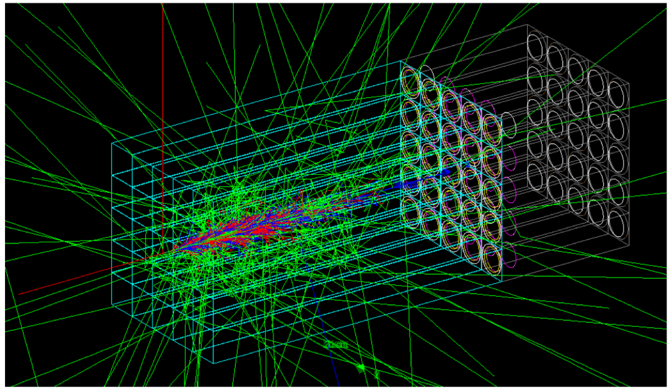
- Support for high concurrency heterogeneous architectures and fast simulations integrated with full detector simulations allows to leverage AI/ML in Geant4.
- Next phase in concurrent Geant4: Sub-event parallelism.
- **Integration in Key4Hep**: Python layer for eAST (fully controllable by UI commands)

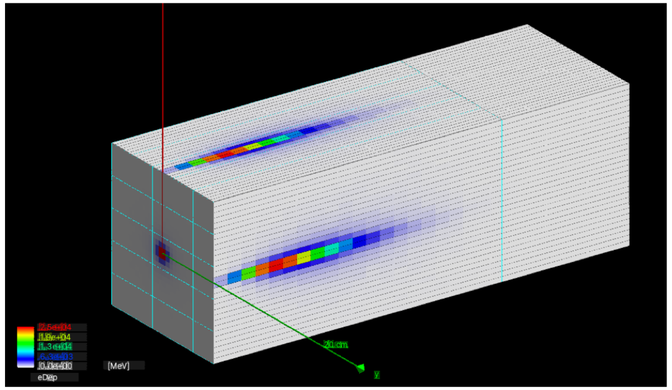
eAST example for using fast and full simultions

From Makoto Asai

5 GeV e^-







- Full simulation with optical photon transport to photo-multipliers
  - 18.41 s/event

x 154

- Full simulation without optical photon transport
  - 0.119 s/event

x 137

- Shower parameterization with GFlash
  - 0.00087 s/event

EIC Software Meeting with Key4Hep

EIC Software: Key4Hep

Wednesday Sep 21, 2022, 11:00 AM → 1:00 PM US/Eastern

Benedikt Hegner (BNL) , Markus Diefenthaler (Jefferson Lab)

Description We will discuss the collaboration with the Key4Hep project.

We will use **Zoom** for the remote meeting:

- https://jlab-org.zoomgov.com/j/1614875218?pwd=RFRPcGlNM3BaS0pQaDhxS3JURkdJZz09
- Meeting ID: 1614875218
- Password: 925723

Live Notes

11:00 AM → 12:00 PM **Status and Plan**

11:00 AM	**EIC Software: Status and Plans** **Speaker**: Dr Markus Diefenthaler (Jefferson Lab)	15m
11:15 AM	**Discussion**	15m
11:30 AM	**Key4Hep: Status and Plans** **Speaker**: Benedikt Hegner (CERN / Stony Brook University)	15m
11:45 AM	**Discussion**	15m

12:00 PM → 1:00 PM **Possible EIC-Key4Hep Projects**

Discussion Items

12:00 PM	**Discussion**	1h