

## EP R&D Software Projects and the EIC

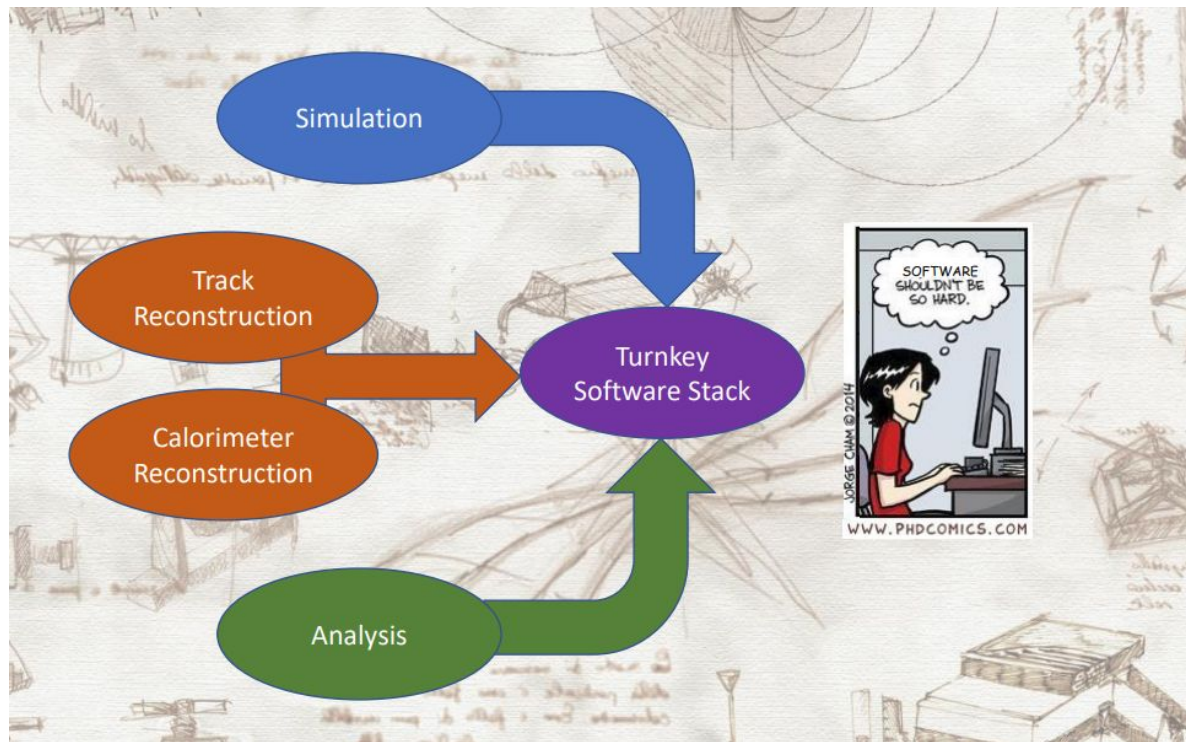
Graeme Stewart and André Sailer, for the WP7 team



# EP R&D Software Projects

- Faster Simulation
- Track Reconstruction
- Calorimeter Reconstruction
- Efficient Analysis Facilities
- Turnkey Software Stack

Full details in [Graeme's talk from yesterday](#)



# Should focus on **eAST** and **HPDF**

eAST: <https://github.com/eic/east>

“[...] ease leveraging new and rapidly evolving computing technologies, [...] a common and integrated approach for fast and full detector simulations in Geant4 with a plug and play modular approach.”

→ Sounds like good synergy with the Fast Simulation task

HPDF: High Performance Data Facility

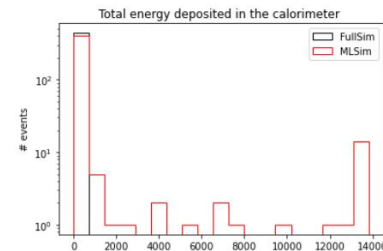
→ High Performance should mean efficient

Still the other projects might be of interest

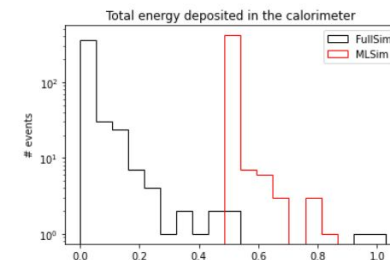
# Developments for Fast Simulation

- Tools facilitating use of fast-sim models implemented into Geant4
- ML-based fast simulation
  - Demonstrated generalisation on sample geometries, for different energies and incident angles
  - Fast adaptation step allows to tune to any geometry
  - LWTNN and ONNX for inference within Geant4
  - Network optimization to reduce memory footprint of inference
- ML4Sim spin-off series of meetings
  - Machine learning for simulation discussion forum
  - Hybrid format of topical meetings and lectures
  - Discussions cover successful and unsuccessful ML tests
  - Keep in touch [mailing list](#)!

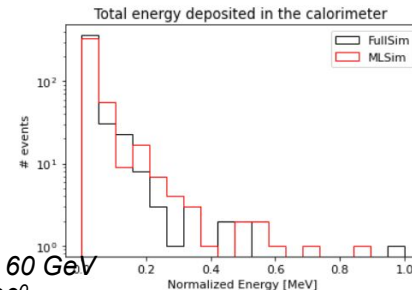
0 iteration



After 100 iterations



After 1000 iterations



Energy 60 GeV  
Angle 20°

3mn on CPU, local machine (4 cores, 16Gb memory)

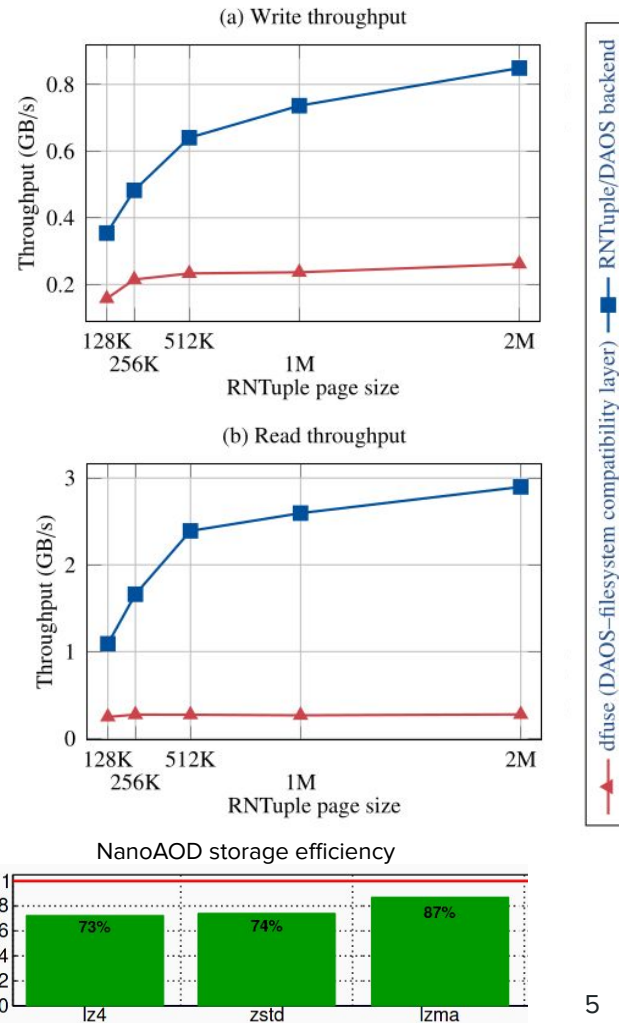
# Developments for Efficient Analyses

## Goals

- Full exploitation of modern storage hardware
  - SSDs, Fast Networks, Object Stores
- developing programming models that boost scientists' productivity
- helping to design specialist Analysis Facilities specifically targeting this workflow

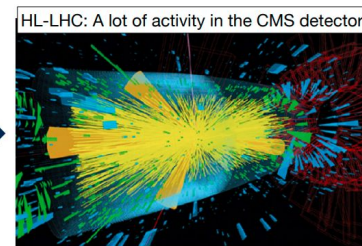
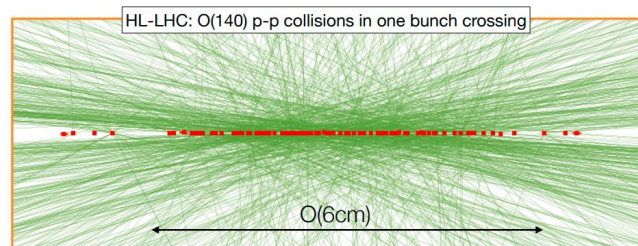
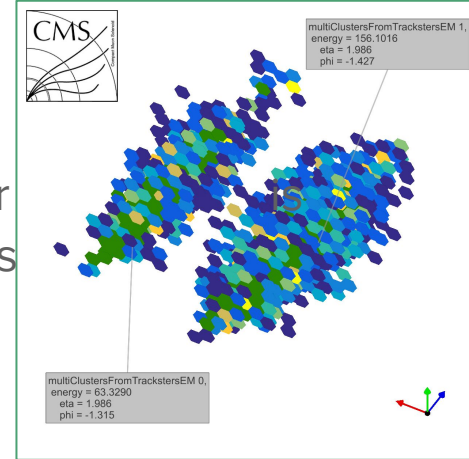
## Developments:

- ROOT::RNTuple: major upgrade of the event data file format and access API
  - 10-20% smaller files, order of magnitude better throughput
  - Closely integrated with analysis workflows
  - Support for code & data evolution over decades
- Joint EP R&D, CERN openlab, Intel and HPE activity on optimal use of object stores
  - Order of magnitude better throughput with dedicated support in the ROOT I/O layer
- Transparent Caching for Analysis Facilities
  - Targets iterative data exploration on the same dataset in an analysis facility
  - Plugged into the RNTuple I/O pipeline, cache on the fly. No extra user code required.



# Calorimeter Reconstruction

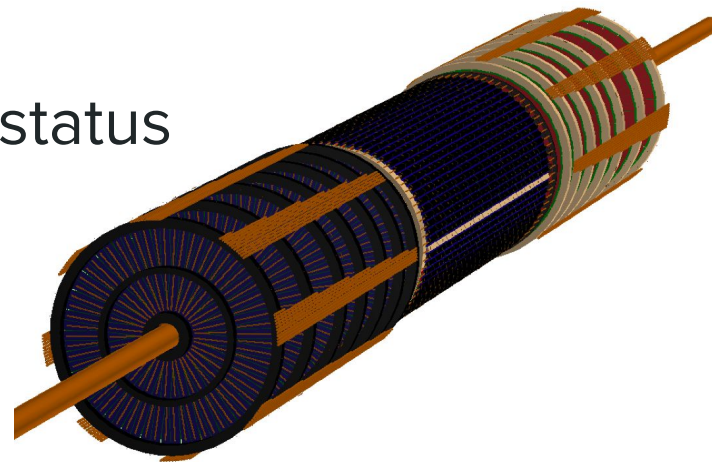
- Particle shower reconstruction in high-granularity calorimeter very interesting and crucial task in high-density environments
  - Fertile ground for new techniques and algorithms: clustering, machine learning, graph theory, and modern computer architectures
- TICL (The Iterative Clustering) is a modular framework integrated and under development in CMSSW
  - processing calo rec-hits (x, y, z, t, E) and returning particle properties and probabilities
  - In a nutshell: grouping 2D layer clusters (CLUE) into 3D clusters (Tracksters) iteratively to reconstruct different particle species
- New flexible and GPU friendly framework can be reused in other (future) experiments using high-granularity calorimeters (not sure what is planned for EIC)





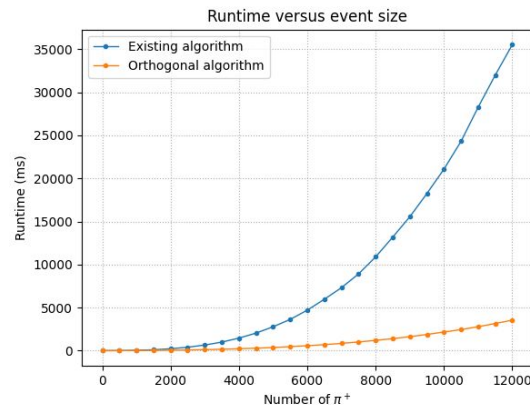
# ACTS EP R&D Tracking project status

- ACTS is an experiment independent toolkit for track reconstruction
- Modern architecture and code, unit tested, continuous integration
- Addressing similar challenges found for calorimeter reco
  - pileup, overlapping clusters, cpu/memory scaling



## Developments

- Parallelization for memory/datamodel, geometry, and algorithms
- GPU based algorithms
- New modules: Gaussian Sum Filter, new seed finding on KD-Trees
- Full chain tracking demonstrator established (TrackML, OpenDataDetector)
  - Large scale validation and tuning efforts ongoing
  - [OpenDataDetector](#) will be a proving ground for R&D that is OpenScience

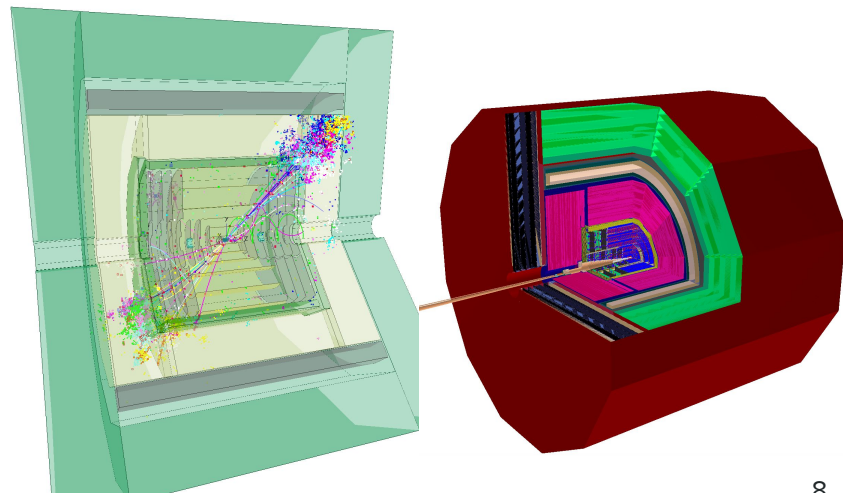
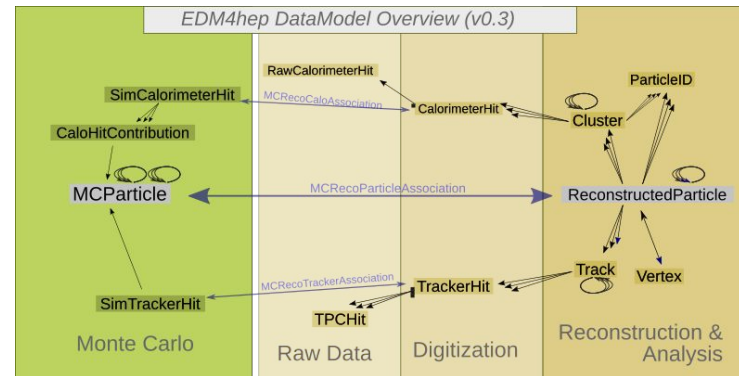
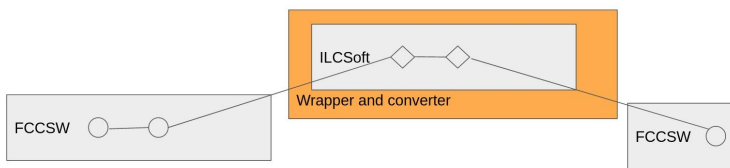


New seed finding strategy on based of KD-trees, achieving very similar tracking performance



# Key4hep: Turnkey Stack

- EP RnD Goals:
  - Provide a single software stack for FCC and CLIC detector studies: Simulation, Reconstruction, Analysis
  - Make software stack usable by other detector groups
  - Testbed for other software developed in WP7
- Common Software
  - Gaudi Processing Framework
  - EDM4hep Event Data Model using Podio
  - DD4hep Geometry Toolkit
  - Standard pieces, e.g., Geant4
    - Allows Fast Simulation developments to be integrated





# Conclusions

- WP7 is development basic solutions necessary for
  - Faster Simulations
  - Track Reconstruction
  - Calorimeter Reconstruction
  - Efficient Analysis Facilities
  - Framework for testing above (and beyond)
- Due to genericity of the solutions or their implementation in standard packages (Geant4, ROOT, Gaudi) synergy for the EIC project should be possible