

Software & Computing

Mike Hance (UC Santa Cruz)

HFCC Software & Computing

- FCC Software and Computing
 - Focusing on summaries and news from this week
 - Highlighting current US efforts and future opportunities
- Future Work
 - Priorities for HFCC S&C efforts
 - HFCC S&C activities in FY27
 - Getting involved

FCC S&C Today

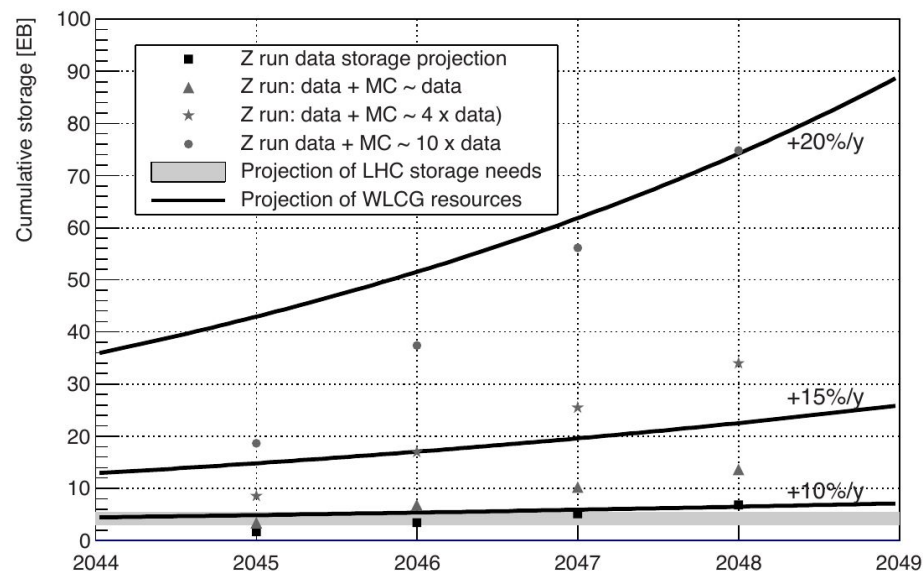
(as seen in talks this week)

FCC-ee Resource Projections

[Feasibility study](#) estimated CPU and disk needs for FCC-ee.

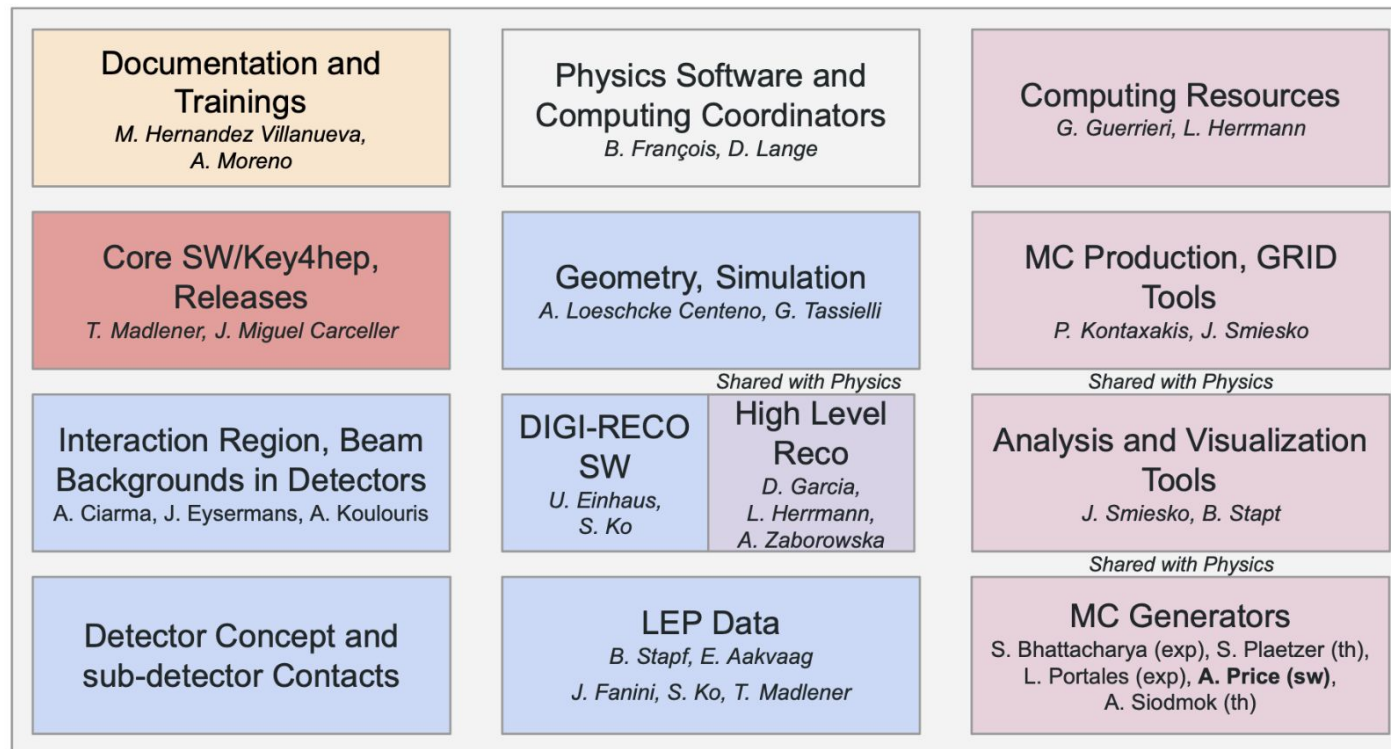
- Completely dominated by Tera-Z program
- Storage requirements are comparable to LHC!

Run	Process	Number of events	DELPHES		FULL	
			Storage (PB)	CPU (HS06)	Storage (PB)	CPU (HS06)
Z	$q\bar{q}$ $\ell^+\ell^-$	1500 G	12.5	2.2 k	1650	2 M
		225 G	0.275	12	40	40 k
W	W^+W^-	60 M	$\sim 10^{-3}$		0.075	72
HZ	HZ	500 k	$\sim 10^{-5}$		$\sim 10^{-3}$	~ 1
	VBFH	16 k	$\sim 10^{-6}$		$\lesssim 10^{-3}$	
top	$t\bar{t}$	500 k	$\sim 10^{-5}$		$\sim 10^{-2}$	~ 1
	HZ	90 k	$\sim 10^{-6}$		$\lesssim 10^{-3}$	
	VBFH	23 k	$\sim 10^{-6}$		$\lesssim 10^{-3}$	
Total		1725 G	13	2.2	1690	2 M



FCC Physics Software & Computing

David Lange (Wednesday)



Legend

Onboarding

Core Software

Detectors

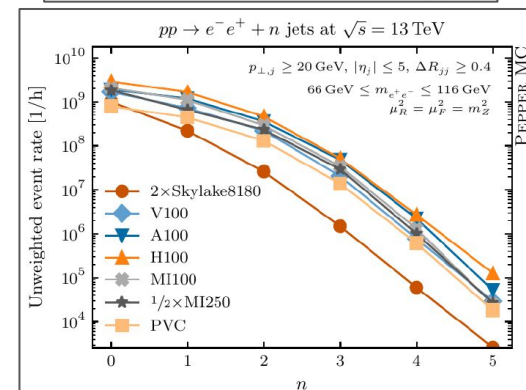
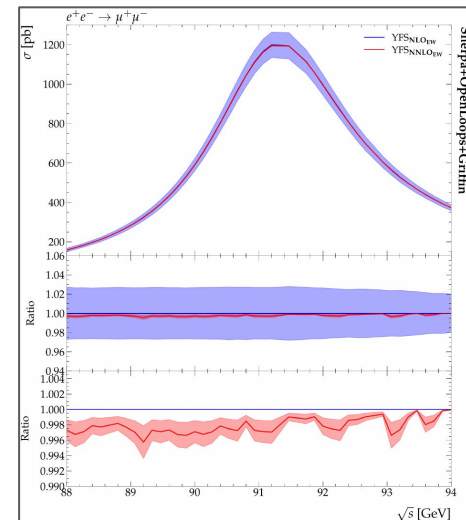
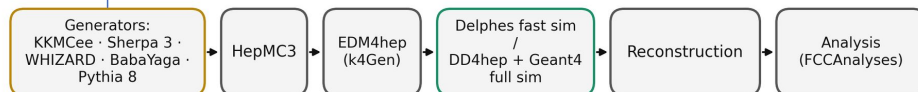
Physics

MC Generators

Sapta Bhattacharya (Thursday)

- Progress in Calculators:
 - NNLO_{EW} at the Z pole
 - $ee \rightarrow ZH$ at NLO_{EW} matched to QED shower
 - N^3LO at top threshold
- Progress in Generators:
 - KKMCee, PanScales, BabaYaga, Whizard all extending to N(N(N))LO / NNLL for flagship processes
 - Sherpa 3 provides full NLO EW workflows, YFS resummation @ NNLO_{EW}
 - Phase-space sampling with normalizing flows
 - Increased GPU support for LHC workhorses
- Next challenge: integration into FCC production

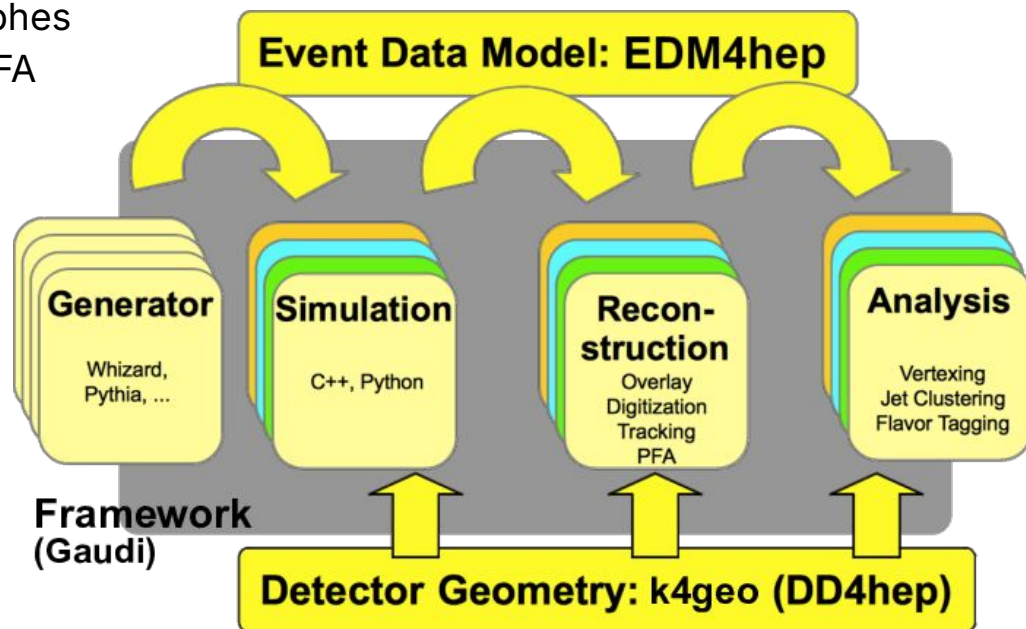
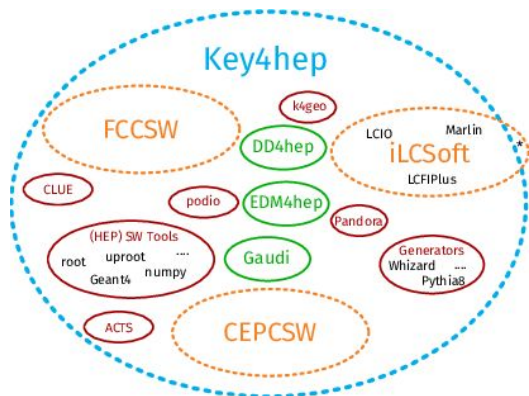
GPU parton-level generation on HPC (DOE/NSF allocations)



Key4HEP

David Lange (Wednesday)

- Partners: FCC, EIC, Muon Collider, ILC, CLIC, etc
- “Turnkey” software solution for future colliders, based on existing tools:
 - Framework: Gaudi
 - Simulation: DD4HEP, Geant4, Delphes
 - Reconstruction: ACTS; Pandora PFA
 - Data Model: EDM4HEP, PODIO
 - Generators, analysis tools
 - Wrappers for legacy code

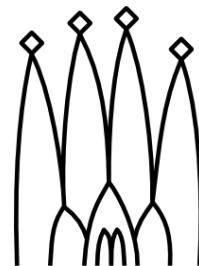


GAUDI Framework

David Lange (Wednesday)

Charles Leggett (Wednesday)

- “Framework for frameworks”
 - Mature (“but not geriatric”) toolkit (~30 years old)
 - Modular algorithms, tools, and services
 - Concurrency support
 - Plugin architecture
 - Active developer community (ATLAS and LHCb)
- Used in Key4HEP, but more can be done to fully integrate it
 - Improve modularity
 - Provide better support for EDM4Hep, DD4Hep, etc
 - Rework configuration management



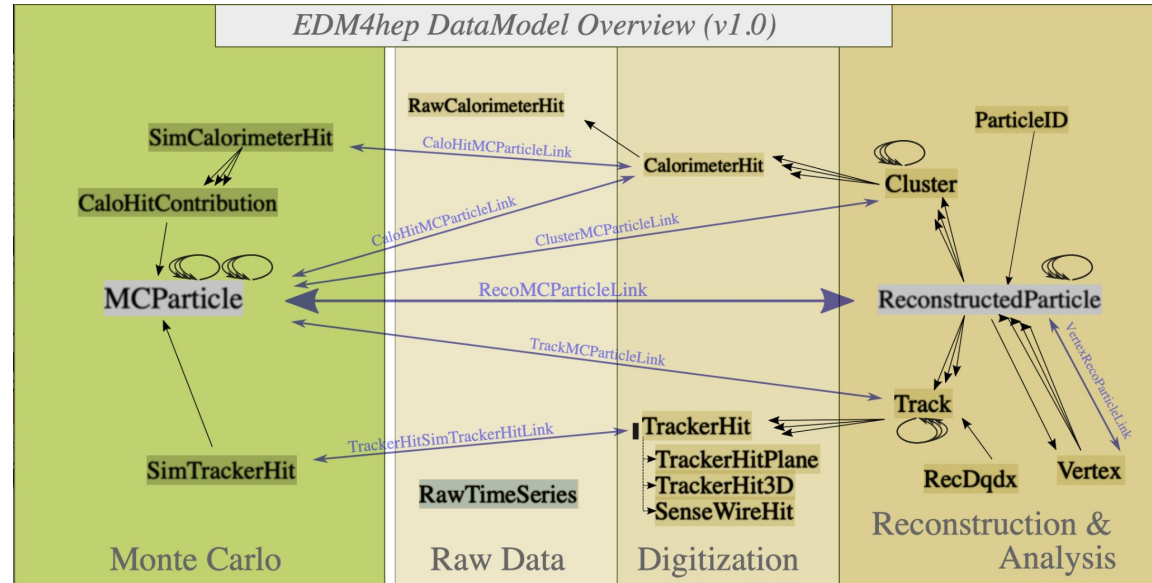
Data Model: EDM4HEP

David Lange (Wednesday)

Giovanni Marchiori (Wednesday)

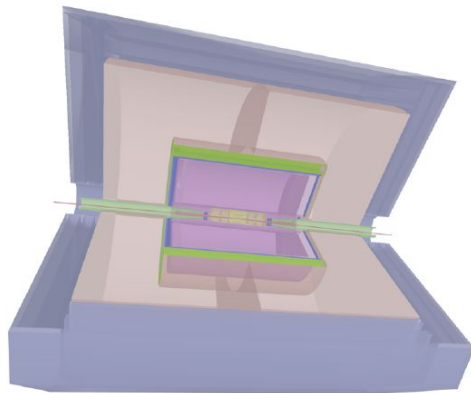
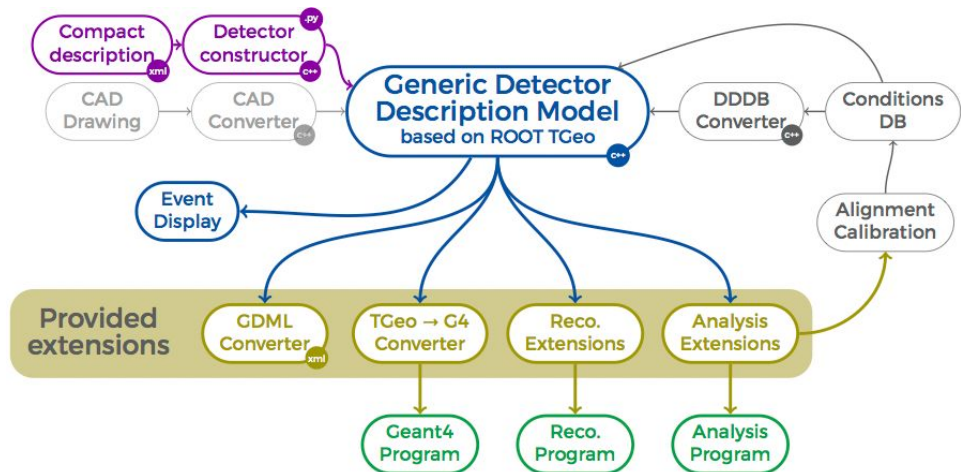
Data model (EDM4HEP) and creation/manipulation software (PODIO) were developed to support Key4HEP. (**Not** an EDM used by LHC experiments.)

- Based on LCIO (ILC), informed by contributions from FCC, EIC
- Significant recent update to v1.0
- User extensions provide flexibility prior to addition of new data types into standard data model
- Can support non-collider data models (e.g. TPC readout) with PODIO "Frames"



DD4HEP is the detector description toolkit used by LHCb, CMS, and ILC.

- XML description language with C++/python “detector constructors”
- Interface to Geant4
- Want to skip full-sim? Use FCC/ILC detector cards in delphes instead, with EDM4HEP output



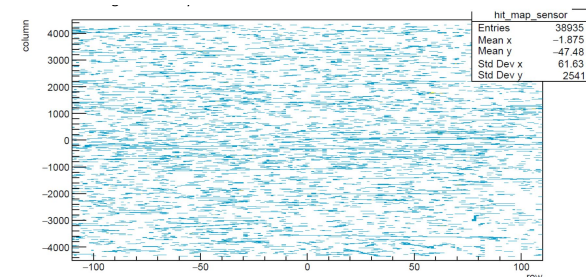
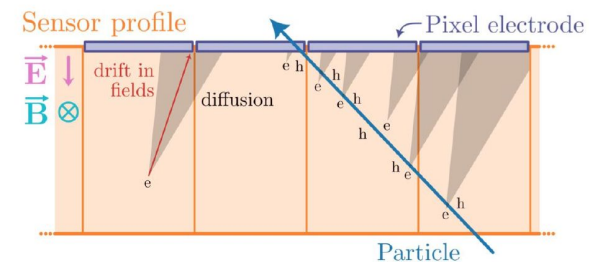
IDEA

Simulation and Digitization

Giovanni Marchiori (Wednesday)

Sasha Paramonov (Tuesday)

- Simulation progress:
 - Delphes and Geant4 simulation increasingly supported
 - Full-sim support for all detector concepts is a priority
 - dd4sim now supports multithreading
 - Improved tools for validation and regression tests
- Significant progress in digitization in past year:
 - Tracking algs now available in Key4HEP:
 - Planar Sensors: Gaussian smearing; Projection; LUT
 - Drift chamber: DCHDigi_v0{2,3} for wire-level digi hits, being adapted for straws
 - Calorimeters:
 - Si-W and scintillator-SiPM digitizers being ported from ILC SW
 - Existing digitizers for ALLEGRO, IDEA, muon tagging
 - More under development for Grainita/ALLEGRO LAr
 - Extending digitization outputs to include detector hits for TDAQ data rate studies



Reconstruction - tracking 1/2

- Two-step problem: track finding (pattern recognition), track fitting (refining track parameter, estimating uncertainties)

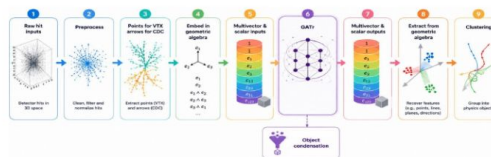
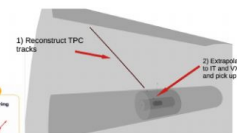
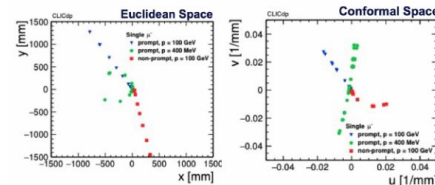
Pattern recognition:

Without ML:

- Conformal mapping:** Circles in xy plane become straight lines in conformal space
 - Available in key4hep ([ConformalTracking](#)), does not work for drift chamber hits (drift isochrones require alternative pattern recognition) ([more details](#))
 - Used by CLD (full-Si detector), usable for ALFA and for vertex+Si-wrapper in e.g. ALLEGRO, IDEA..
- ACTS:** experiment-independent toolkit for track reconstruction in HEP
 - Available in key4hep ([k4ActsTracking](#)), same limitations as conformal tracking, work ongoing to develop dedicated pattern matching for drift chambers (Hough-transform) ([more details](#))
- Hybrid** pattern recognition for ILD@FCC: TPC-seeded ([Clupatra](#)), build tracks from TPC hits, search for Si hits near tracks. New alg under R&D to combine Clupatra & conformal tracking ([more details](#))

- With ML:** **Geometric Graph Track Finder**, embeds hits from subdetectors in latent space and clusters them. Available in key4hep ([GGTFTTrackFinder](#)), model trained on IDEA

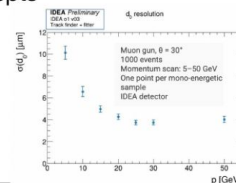
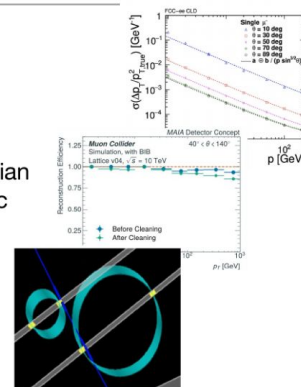
- More details in Andrea's [talk](#) at FCC week



Reconstruction - tracking 1/2

Reconstruction - tracking 2/2

- Track fitting:
 - **Conformal tracking:** apply Kalman filter to tracks found from pattern recognition. Can be used for Si-based trackers. In key4hep ([ConformalTracking](#)), in use by **CLD** ([more details](#))
 - **ACTS:** provides KF-based fitting, but also Gaussian Sum Filter for electron tracks (strongly non-Gaussian energy loss from bremsstrahlung), and Global-chi2 fitter for DCH tracks (work ongoing on Deterministic Annealing Fitter). In key4hep ([k4ActsTracking](#)), first validation with **MAIA** ([more details](#))
 - **GenFit2:** framework developed for Panda/Belle-II, implements DAF, resolves left-right ambiguities. Available in key4hep ([GenfitTrackFitter](#)), included in ALLEGRO and IDEA reco chain ([more details](#))
- Vertexing:
 - **Least-squares vertex fitter** code in [Delphes](#) ([VertexFit](#)) being migrated to Gaudi for **ALLEGRO/IDEA** ([details](#))
 - Full-suite of algs in [LCFIPlus](#) developed for ILC and available via Marlin wrapper (used by **CLD** and **ILD**) ([details](#))
 - **Tear-down PV finder, build-up SV finder, V0 fitting and selection.** Being ported to key4hep, reusing VertexFit for every trial fit ([details](#)), also for other detector concepts
 - Potential other candidates: [ACTS](#)
- **Benchmarking/validation:** algorithm developed and recently made available in [k4DetectorPerformance](#) ([TrackingValidationPlots](#)) (included as option in **ALLEGRO** reco chain)



Software for FCC: simulation & reconstruction - Giovanni Marchiori - 10

Reconstruction

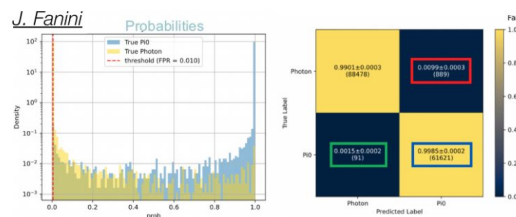
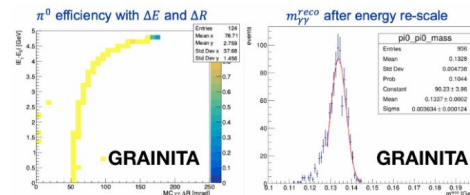
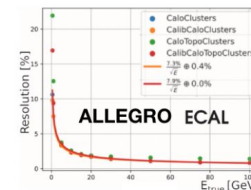
Giovanni Marchiori (Wednesday)

Reconstruction - tracking 1/2

Reconstruction - tracking 2/2

Reconstruction - clustering

- Algorithms implemented for clustering of hits with sufficient deposited energy
- fixed-size clustering** ([CreateCaloClustersSlidingWindowFCCee](#))
finding local maxima of total energy in **sliding window**
- topological clustering** ([CaloTopoClusterFCCee](#)) of adjacent cells with S/B above given thresholds
- CLUE** ([k4Clue](#)): fast and parallel, density-based, weighted clustering algorithm - used e.g. for studies of the GRAINITA ECAL for **ALFA**
- Higher-level algorithms developed for e.g MVA-based **cluster energy regression**, **particle ID** (single photon vs $\pi^0 \rightarrow \gamma\gamma$) (**BDT**, **transformer**), ..



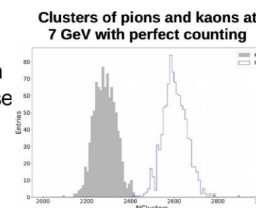
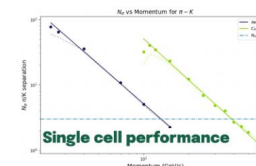
Reconstruction - tracking 1/2

Reconstruction - tracking 2/2

Reconstruction - clustering

Reconstruction - higher-level

- **Global event reconstruction** (tracks+clusters => particle candidates with 4-momentum and PID)
 - **Particle-flow** (Pandora & ML-based) => see [talk](#) by Dolores tomorrow
- **PID algorithms**
 - **ARC**: algorithm ([ARCA1g](#)) for reconstruction of **Cherenkov** angle from track direction and hit positions under development ([details](#)). Simulations done at single-cell level, moving to full detector, to extract tabulated resolutions for π , K and p as a function of particle momentum and pseudorapidity for particle-flow studies
 - **dN/dx**: total track dN/dx can be computed from n(Clusters) of associated track hits and track length in **gaseous detectors**. More realistic dN/dx to be available after DCH digitiser v03 ready with digital pulse and peak-finding algorithm implemented for estimation of primary ionisation clusters
- **Jet reconstruction and tagging**
 - **clustering**: [ClusterJet](#) - basically a wrapper for FastJet
 - ML-based **flavour tagging**: transformer-based algorithm for **CLD** implemented in key4hep ([k4MLJetTagger](#)), expect more to come in the future with dedicated working group now established. See also tomorrow's [talk](#) by Gregor

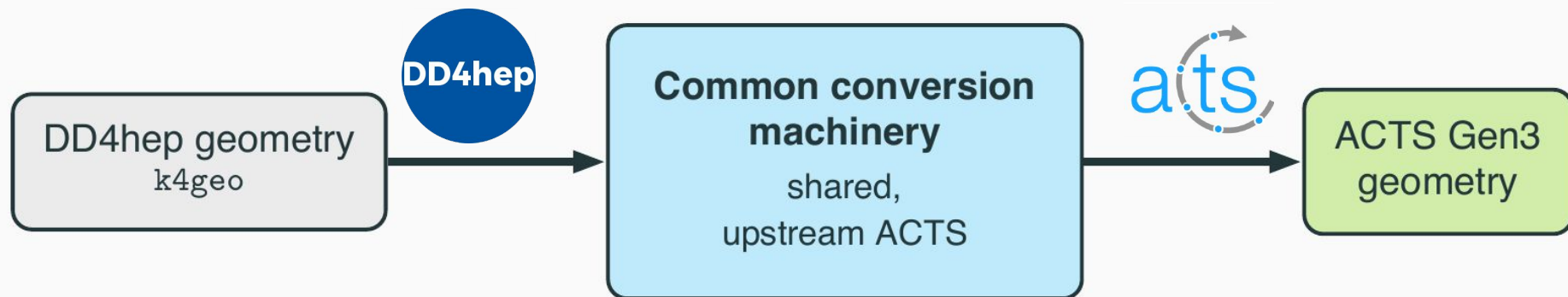


Reco, Tracking, Geometry

Vakho Tsulaia (Wednesday)

Key4Hep provides a common platform, but does not automatically enable all tools to work together:

- DD4HEP→ACTS Geometry Conversion
 - Translating detector models from DD4HEP into ACTS "Gen 3" via standard APIs
- Also exploring alignment/conditions/geometry integration





CODE4HEP: FNAL LDRD to understand the applicability of CMSSW's architecture within key4hep ecosystem

- **stitched**: event processing framework extracted from CMSSW
 - Now supports read/write in EDM4HEP data formats
 - Interchangeable with Gaudi workflows for data processing
- **Multithreading consolidation in Geant4**
 - Various configurations now made thread local so that GEANT threading model can be adapted to modern threading models
 - k4Geo based simulation is coming soon, expect before end of year
- **EDM4HEP support in coffea for data analysis**
 - Pure-python complement to FCCAnalysis (RDataFrame-based)
 - Popular with LHC analyzers, strong support and performance at US analysis facilities

Impressive results, and good model for porting LHC tools into Key4HEP

Coffea support for EDM4HEP

Lindsey Gray (Thursday)

```
from coffea.nanoevents import NanoEventsFactory, EDM4HEPSchema
import awkward as ak
import numpy as np
import hist

events = NanoEventsFactory.from_root(
    "./ee_mumuH_rec_17090_90_again.root:events",
    schemaclass=EDM4HEPSchema,
).events()

E_cm = 240.0
Z_pole = 91.1876

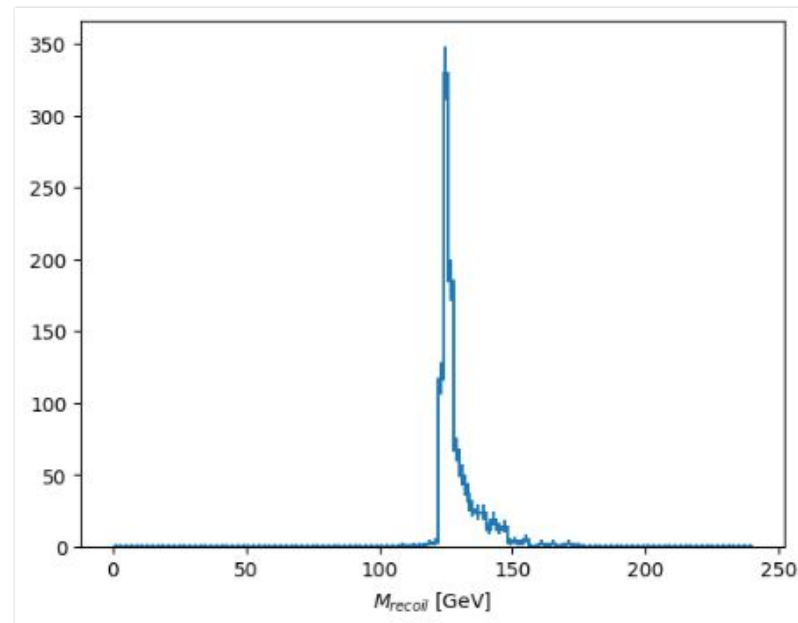
muons = events.PandoraPF0s[abs(events.PandoraPF0s.PDG) == 13]
dimuons = ak.combinations(muons, 2, fields=["mu1", "mu2"])
os_dimuons = dimuons[dimuons.mu1.charge != dimuons.mu2.charge]
os_p4 = os_dimuons.mu1 + os_dimuons.mu2
nearest_zpole = os_p4[
    ak.singletons(ak.argmax(abs(os_p4.mass2 - Z_pole**2), axis=1))
]
nearest_zpole = nearest_zpole[nearest_zpole.mass > 50] # get rid of low mass dimuons

recoil = np.sqrt(E_cm**2 + nearest_zpole.mass2 - 2 * nearest_zpole.energy * E_cm)

recoil_hist = (
    hist.Hist.new.Reg(120, 0, 240, name="recoil", label=r"$M_{recoil}$ [GeV]")
    .Double()
    .fill(ak.flatten(recoil, axis=1))
)

recoil_hist.plot1d()
```

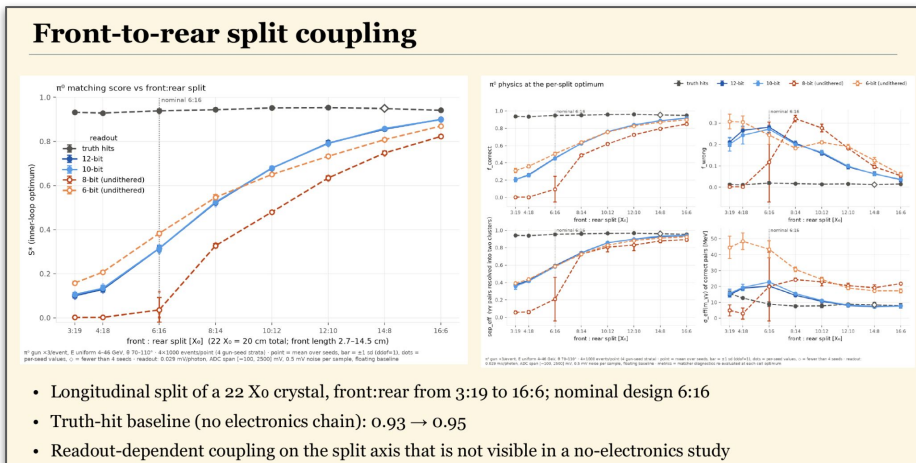
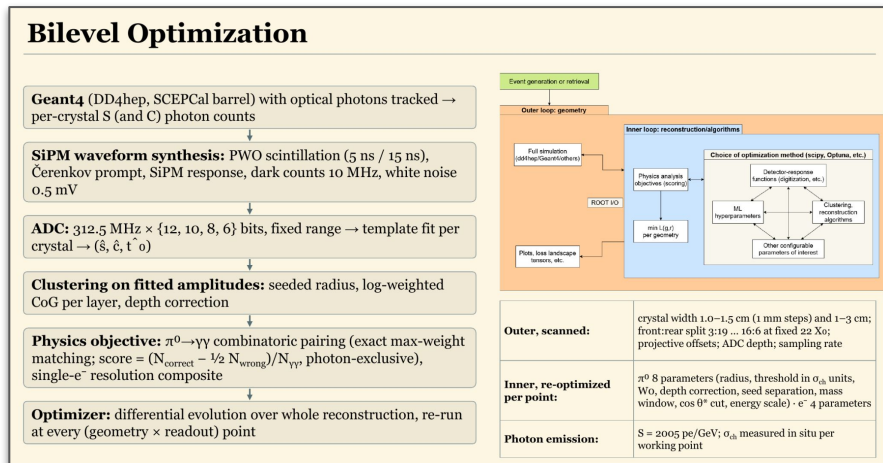
**This file was made
by stitched!**



AI for Detector Optimization

Wonyong Chung (Wednesday)

- Optimization over detector geometry $\rightarrow$ physics objects with agentic workflows to optimize ECAL geometry, sampling:



- Next steps: extend to full IDEA detector with agentic reasoning instead of Bayesian/grid scan.

The Funding Challenge

- We did not do well in the FY26 RCA for “Phase 1” projects
 - No one seem to know why, there are probably many reasons
 - My favorite: we are a large, mature AI community with years of successful R&D behind us.
 - To be “transformative” (as required by Genesis Mission) we may well need a moonshot project requiring AGI-scale resources.
- We are doing better in the “AI infrastructure” projects
 - Probably because of our experience working at scale in a distributed production environment.
 - We should leverage that.
 - Do FCC applications need AI services (IaaS, model repositories, MCP, etc)?
- One area of growth are industrial partnerships.
 - We have a lot to learn from their R&D teams.
 - In my experience, industrial partners are excited to work with “big science” projects.
 - Can we make a good business case?

In any case, **we can't sit this one out:**

we need Genesis resources, models, and tools that meet our needs.

- **The Fermilab Elastic Analysis Facility (EAF)**

Web-based, container-driven data platform for particle physics and machine learning workflows

- **JupyterHub Interface** (access to interactive environments: terminals, code editors, notebooks)
- **Integrated to Dash and HTCondor**, enable users to run large-scale parallel Python tasks
- **Access to high-performance resources**, including GPU instances (such as NVIDIA A100s) for heavy machine learning and data processing tasks

```
daniel@jupyter-daniel:~ x +
[daniel@jupyter-daniel ~]$ source /cvmfs/sw.hsf.org/key4hep/setup.sh -r 2026-04-08
AlmaLinux/RockyLinux/RHEL 9 detected
Setting up the Key4hep software stack release 2026-04-08 from CVMFS
Use the following command to reproduce the current environment:

    source /cvmfs/sw.hsf.org/key4hep/setup.sh -r 2026-04-08

If you have any issues, comments or requests, open an issue at https://github.com/key4hep/key4hep-spac/issues
Tip: A new -d flag can be used to access debug builds, otherwise the default is the optimized build
[daniel@jupyter-daniel ~]$
```



**Member of the Fermilab
FCC-ee user community**

Marc Andre Pleier (Thursday)

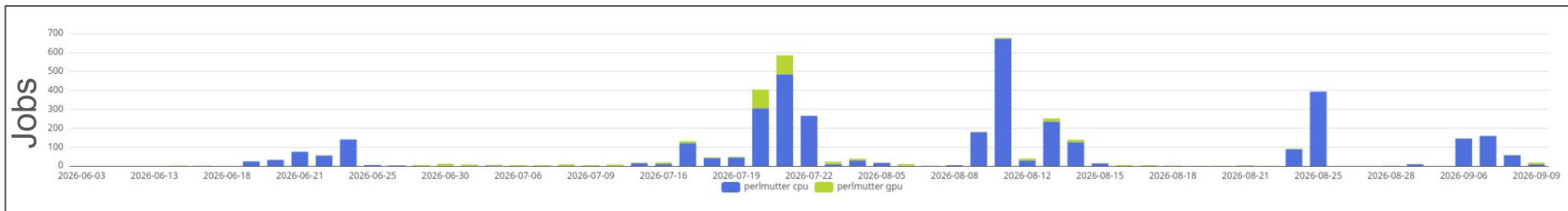
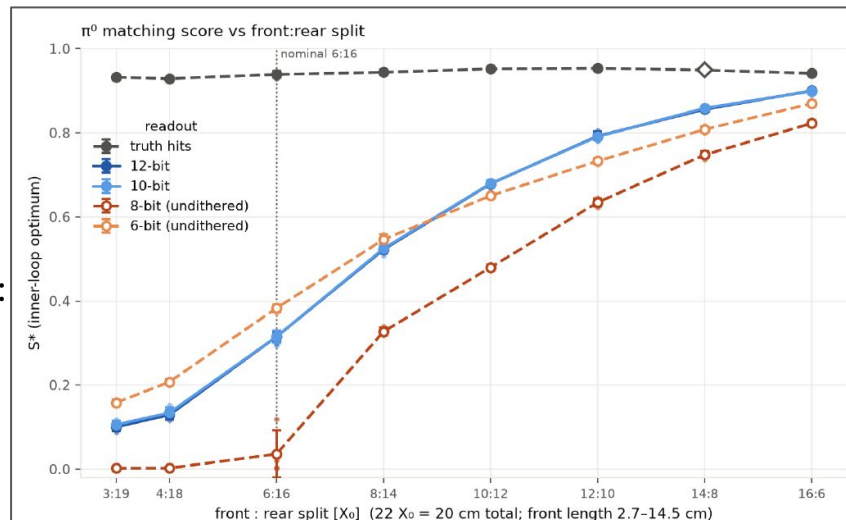
- **We have finally a BNL AFsetup ready!**
 - Currently, it includes:
 - 100 CPUs allocated for batch processing (with opportunistic access to thousands)
 - **>10 NVidia L40S GPU** for ML and vectorized computing
 - ~100 TBs of common disk space for sharing datasets
- Planning to grow it larger this year with more CPU and disk storage!
- Instructions to access these resources are available @ <https://docs.sdcc.bnl.gov/experiments/fcc/onboarding/>
- Two paths for access, full SCDF account or light weight federated account
 - People who want to get full SCDF account are invited to follow the process and

Access Type	Institution	Features	Account Type
Full SDCC Access	BNL Users	SSH, Jupyter, Condor (CPU/GPU), Storage	SDCC Account
Federated Access	CERN / Fermilab / SLAC	Jupyter with terminal and Condor access	Federated Identity

Other US Computing Resources



- Several institutional clusters (e.g. MIT) + CERN Ixplus
- NERSC
 - Used for [VIBRATO](#) studies
 - ERCAP allocation on Perlmutter for FCC work:
 - 10k CPU hours + 10k GPU hours
 - Many hours remain this year, let's use them!
 - [Extensive documentation](#)
 - All users who start up work there have commented on how easy it is for them (and their agents) to get started!
 - For access, contact [Mike Hance](#)



Computing Opportunities

David Lange (Wednesday)

- CERN resources are limited: <1 PB disk, <10k HS23 compute.
 - National/institutional clusters enable studies, but no large-scale integration yet
- Starting rucio data federation: >1 PB online today, more in the works
 - <https://hep-fcc.github.io/fcc-sw-web/rucio>
- Restarting simulation campaign this fall:
 - DIRAC v8 middleware framework → rucio datasets
 - Excellent time for LHC physicists to start physics studies on FCC datasets!
- Integrating US compute/disk resource should be a high priority for US
 - Support MC campaigns
 - Provide local/fast access to FCC datasets for physics studies
 - Enable detector R&D groups to store/study data from test beams, lab tests, etc

Future Work

US HFCC S&C Priorities

1. **Integrate existing US computing and human resources** into international effort
2. **Support physics studies and detector R&D** by US physicists
3. Ensure that **lessons from (HL)LHC** are not lost for FCC
4. Exploit **AI to develop, operate, and do physics** at the FCC
5. Prepare for **transformational developments in HEP S&C** to meet the needs of future facilities

US HFCC S&C Priorities

1. **Integrate existing US computing and human resources** into international effort
2. **Support physics studies and detector R&D** by US physicists
3. Ensure that **lessons from (HL)LHC** are not lost for FCC
4. Exploit **AI to develop, operate, and do physics** at the FCC
5. Prepare for **transformational developments in HEP S&C** to meet the needs of future facilities
6. ... and don't ask Srini and Valentina for too much money

US HFCC S&C Projects

In addition to LDRD initiatives, HFCC is supporting targeted R&D and “operations” S&C projects for Higgs factories.

- FY26 focused on priorities 1 and 2
- FY27 proposal to continue and extend to priorities 3, 4, and 5:

Area	Work Package	Institution
Physics SW	Agentic AI for FCC Workflows	Princeton
Physics SW	AI-Driven Documentation & Knowledge Surfaces	Princeton
Core SW	Geometry and thread safety	LBNL
Core SW	Key4HEP configuration	LBNL
Computing	DDM and OSG	UMass
Computing	Computing Hardware	UMass

How to get involved

- Contact HFCC S&C coordinators (below): we're here to help!
- Join our [HFCC S&C monthly meeting](#) or the [general HFCC monthly meeting](#)
- Join us on the [US-FCC Mattermost Channel](#)
- Get started with FCC Software: <https://hep-fcc.github.io/FCCSW/>
- Use our computing facilities to test your ideas for Higgs Factories



[Sapta Bhattacharya](#)
(Physics SW)



[Charles Leggett](#)
(Core SW)



[Rafael Coelho Lopes de Sa](#)
(Computing)

Getting involved – Meetings and mailing lists

FCC PED indico: <https://indico.cern.ch/category/5251/>

Meeting	When	Indico Category	Mailing list
Full Simulation	Every other Wednesday alternating between 11:00 and 16:00 GVA	https://indico.cern.ch/category/16938/	https://groups-portal.web.cern.ch/group/fcc-ped-softwareandcomputing-full-simulation/details
Reconstruction	Every other Thursday 16:00	https://indico.cern.ch/category/19516/	https://groups-portal.web.cern.ch/group/fcc-ped-physicsgroup-highlevelreco/details
Key4hep	Weekly, alternating between 9:00 and 16:00	https://indico.cern.ch/category/11461/	https://groups-portal.web.cern.ch/group/key4hep-sw/details
LEP Data	Every other Monday 14:00	https://indico.cern.ch/category/20141/	https://groups-portal.web.cern.ch/group/fcc-ped-softwareandcomputing-lepdata/details
Documentation	Every other Friday 16:00	https://indico.cern.ch/category/21617/	https://groups-portal.web.cern.ch/group/fcc-ped-softwareandcomputing-documentation/details
Analysis and MC production	Every other Wednesday 14:00	https://indico.cern.ch/category/17971/	https://groups-portal.web.cern.ch/group/fcc-ped-softwareandcomputing-analysis/details

More mailing lists:

- Announcements: <https://groups-portal.web.cern.ch/group/fcc-ped-softwareandcomputing/details>
- Generators: <https://groups-portal.web.cern.ch/group/fcc-ped-softwareandcomputing-generators/details>
- MID: <https://groups-portal.web.cern.ch/group/fcc-ped-softwareandcomputing-mdi/details>

Conclusion

S&C for Higgs Factories has its own challenges and opportunities:

- Large data volume from Tera-Z runs
- Revolutionary changes in all aspects of software and computing hardware since LEP
 - Anticipate other transformative changes between now and 2040's!
- How to incorporate significant (and ongoing) R&D from (HL)LHC experiments while actively supporting R&D on FCC detectors?
 - Good example from FNAL team: targeted ports of SW into Key4HEP

How to get involved:

- Contact HFCC L2's with your ideas and interests
- Join our monthly meetings (next meeting in October)