Software for PED studies

Introduction. Status of Things

FCC Physics Workshop 2023
Krakow, Poland

January 24, 2023
G Ganis, CERN-EP

S&C sessions

24/1/2023

09:00 → 10:35 **Software and Computing**

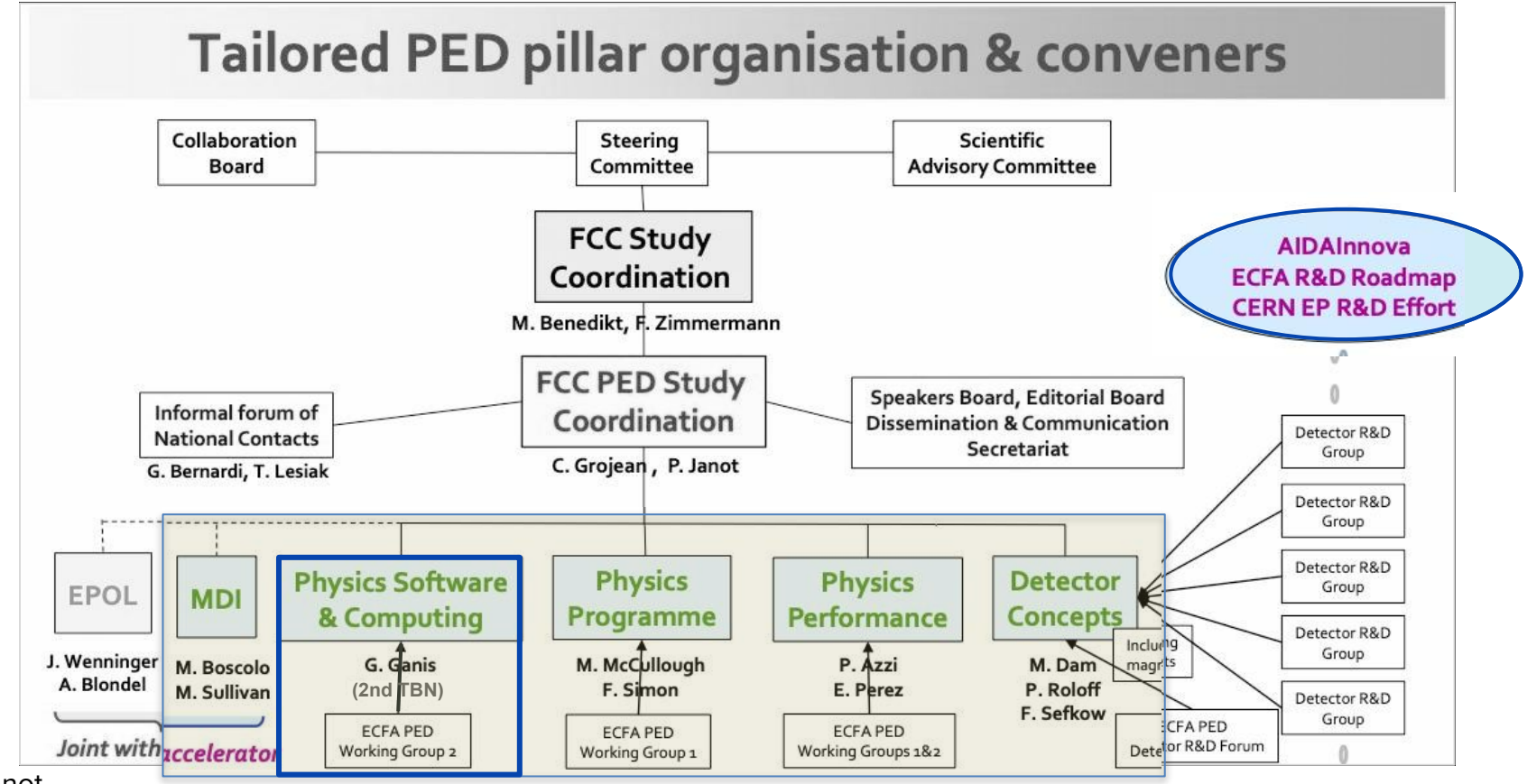
- 09:00 **Introduction and statuts of things**
  Speaker: Gerardo Ganis (CERN)
- 09:20 **Key4hep and FCC software**
  Speaker: Gerardo Ganis (CERN)
- 09:45 **Analysis and visiualisation** ¶
  Speaker: Juraj Smiesko (CERN)
- 10:10 **Tools for Monte Carlo productions**
  Speaker: Andre Sailer (CERN)

25/1/2023

11:00 → 13:05 **Joint Detector/Software**

- 11:00 **Strategy and plans for detector software**
  Speaker: Brieuc Francois (CERN)
- 11:25 **ARC: progress update and plans towards full simulation**
  Speaker: Martin Tat (University of Oxford)
- 11:50 **Modelling signal digitisation for test calorimeters: the CALICE experience (remote)**
  Speaker: Vincent Boudry (LLR, CNRS, École polytechnique, Institut Polytechnique de Paris)
- 12:15 **Modelling signal digitisation for trackers ~~(remote)~~**
  Speaker: Riccardo Farinelli (Universita e INFN, Ferrara (IT))
- 12:40 **Performance of an ALICE ITS3-like vertex detector for FCC-ee (remote) and progress on the IDEA vertex detector implementation in full simulation**
  Speakers: Armin Ilg (University of Zurich) , Leila Freitag (University of Zürich)

S&C relation with other groups



P Janot

S&C goals for the workshop

- Assess status of things and required workflows in view of the FSR
  - In particular wrt detector concept implementation

- Trigger a discussion on the missing pieces; consolidate the strategy to address them

- Further engage community participation

FCC S&C approach

Driving consideration: *one software (ecosystem) for all*

- Initially was *all FCC cases (hh, ee, eh)*, now is *all future collider projects* (key4hep)
  - Crucial is to allow for evolution (modular approach)

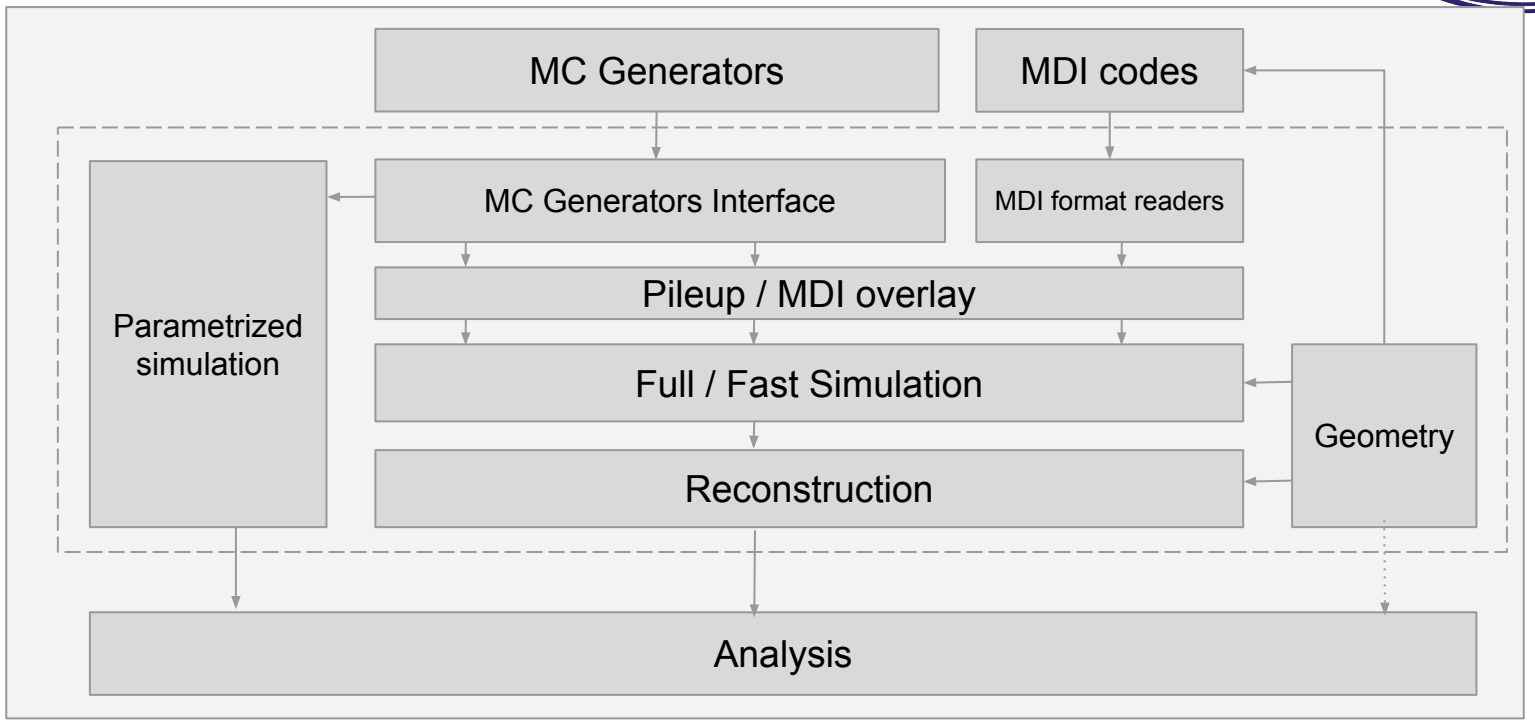
Adopted strategy

- Adapt existing solutions
  - Mostly from LHC and running experiments, but also look at ongoing R&D projects (AIDA, …)

Priorities somehow differ from those of running experiments

- Privilege low-barrier of entry, agile support for detector concept evolution
  - Correctness, easy-of-use, commonality, interoperability
- Nevertheless, ability to adopt best practices, integrate lessons and performance accelerators from running experiments might be instrumental for the FSR

Typical workflows to support



Software Infrastructure (Build/Test/Deploy)

Workload and Data Management

FCCSW: status of things

- Monte Carlo generators ..... ✓≈
  - General purpose + old LEP available
  - Treatment of BES, x-angle, … still in progress
- MDI interfaces ..... ≈
  - GuineaPig prototype available, start investigation of SR
- Fast simulation (Delphes) .... ✓
- Full simulation and reconstruction ..... ≈
  - Single components available (machine elements, simplified vertex+DC, LAr calo, …)
  - CLD (reconstruction through iLCSoft wrappers)
  - IDEA standalone, integration scheduled
- Distributed computing ..... ✓≈
  - iLCDirac instance set up (CERN, CNAF storage)
- Analysis ..... ✓≈
  - FCCAnalysis (EMD4hep, ROOT::RDF)
- Documentation ..... ✓≈

FCCSW: status of things

5th FCC Physics Workshop

- Monte Carlo generators ..... ✓≈
  - General purpose + old LEP available
  - Treatment of BES, x-angle, ... still in progress
- MDI interfaces ..... ≈
  - GuineaPig prototype available, start investigation of SR
- Fast simulation (Delphes) .... ✓
- Full simulation and reconstruction ..... ≈
  - Single components available (machine elements, simplified vertex+
  - CLD (reconstruction through iLCSoft wrappers)
  - IDEA standalone, integration scheduled

  S&DC, Wed, 11h00
- Distributed computing .... ✓≈
  - iLCDirac instance set up (CERN, CNAF storage)

  A Sailer, Tue, 10h05
- Analysis .... ✓≈
  - FCCAnalysis (EMD4hep, ROOT::RDF)

  J Smiesko, Tue, 9h40
- Documentation ..... ✓≈

Monte Carlo Generators @ key4hep

- All relevant generators available
  - Latest version of KKMC (v5) available w/ several improvements (S. Jadach, FCCW 2022)
- Repository of ready-to-use codes
  - (Software-wise) sanitized: build, tested (with provide test suite)
- Interfaces to key4hep processing chain (k4Gen)
  - Readers for relevant data formats
  - Set of auxiliary tools related to MC particles (filters, smearing, particle guns, …)
  - Possible additions:
    - Tools for crossing-angle, event mixing, event overlap
    - Generic treatment of BES, packet length, … for generators not providing them

Connections w/
- Physics Programme
- Physics Performance
- Phenomenologists
- Key4hep
- ECFA

Beam and MDI-related backgrounds

- FCCee complex interaction region combined with high statistical precision requires deep level of understanding of the detector backgrounds (*)
  - Only achievable with integration in experiment software
- Several codes for modeling the processes (EPJ+ contribution)
  - Codes not always in public repositories, outputs in different, non-standard formats
- Supercode w/ interfaces to relevant codes more difficult to get than foreseen
  - Difficult access to codes, difficult to run them (exception is GuineaPig)
  - Still need code owners to produce list of particles for a given configuration
  - Injection in the framework standardised w/ the help of dedicated readers
- Consistent description of the relevant geometry elements
  - Requires interplay between detector and machine geometry formats (e.g. CAD)

A Ciarma
Tue 18h10

Connections w/
- MDI study group, Physics Performance, Detectors

Parametrized Simulation

- Current studies are based on Delphes
  - Includes a tracking system, w/ magnetic field, calorimeters and a muon system
    - TrackCovariance, dEdx, ParticleDensity: enable realistic algorithms for vertexing, b-tagging, …
    - Effect of magnetic field, granularity of calorimeters, sub-detector resolutions
  - Interfaced to standard file formats (e.g. EDM4hep, HepMC)
  - Palette of the relevant detector concepts for e^+e^- available
- Acts as tool to understand the impact of a particular performance parameter and understand where to focus

Connections w/
- Physics Performance
- Detector Concepts
- ECFA WG2

For reviews and usage examples see
1st ECFA simulation workshop 1-2 Feb 2022

Detector software

- Everything having to do with sub-detectors
  - Geometry description, digitisation, reconstruction
- Situation reviewed in depth at the DC kick-off workshop in June 2022
- Open discussion on a wide range of topics, including
  - CAD to DD4hep conversion
  - Geometry consistency tools
  - Role of fast simulation
  - Geant4 limitations (missing processes / decays, …)
  - Reconstruction tools (sub-detector specific, Particle Flow, …)
  - Figure of merit for PF (Jet resolution? Pi0 resolution? …)
- Dedicated joint session Software-Detector tomorrow

S&DC, Wed, 11h00

Analysis: FCCAnalysis

- Based on RDataFrame, new ROOT paradigm developed for (HL-)LHC
  - Python framework with C++ backend
  - Bridges the gap with LHC involved people
- Runs on EDM4hep, non FCC specific
  - Prototype of generic analysis framework
  - Possible inclusion in key4hep under discussion

J Smiesko, Tue, 9h40

Connections w/
- Physics Performance
- Key4hep

Analysis configuration
4 **python** scripts to configure:

1. Samples to run over
2. Functions/algorithm to call
3. Event selection
4. Plotting configuration

Common utility functions, algorithm, etc…
C++ library

Common interface code
Sample database, RdataFrame, plotting
Python

Reminder: resource needs for the FSR

See EPJ+ contribution

- The run at Z peak sets the scale
  - ≈10^{12} evts, 3-6 EByte storage, 10 MHS06 CPU (≈ current ATLAS yearly needs)
- These numbers are similar to the ones expected for (HL-)LHC
  - Do not expect issues for operations in 2040 and beyond
- For the FSR the situation is different
  - Analysis at Delphes level are possible (see $B_c \rightarrow \tau^+ \nu_\tau$)
  - Full simulation of all components require 10^3-10^4 times more
- Techniques of overcome this limitations are required
  - E.g. interplay of full and parametrized simulation
- Need to understand role of improvements in (fast) simulation of EM calorimetry
  - Geant4/GFlash or Machine Learning / GAN
  - New simulation accelerating techniques

Event Producer Workload / Data Management

- Currently still using home-made solution (HTCondor, EOS, ...)
  - Served well CERN-based productions (CDR , Spring 2021)
- Inclusion of non-CERN resources desired
  - Main requirements: central file catalogue, replication, remote access
  - Major development for the in-house system
- iLCDirac: LC community DIRAC instance
  - Workload management, file catalogue used by LHCb, Belle II, BES III, JUNO, ILC/CLIC, ...
  - Already serving another VO (CALICE)
- FCC @ iLCDirac
  - Re-activated FCC VO
    - Associated CERN FCC resources to FCC VO (HTCondor, EOS area)
  - Added steering applications of interest for FCC workflows
  - Storage organisation based on LC and LHCb experience

Connections w/
- Key4hep
- Physics Performance

A Sailer, Tue, 10h05

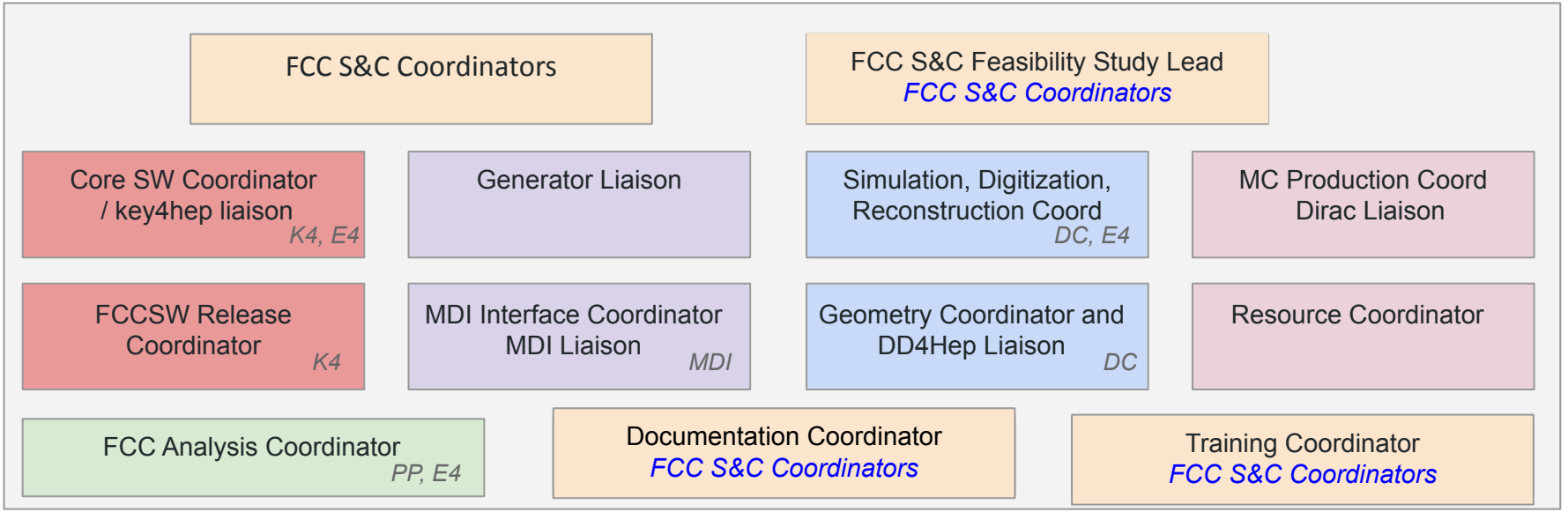
Documentation: new tutorials

- Hands-on event on 19-21 Oct 2022
- Focus on how to use the existing software to perform the most common workflows
- 6 tutorials covering all the spectrum
  - Event generation
  - Vertexing in FCCAnalyses
  - Visualisation
  - Modifying a DD4hep geometry
  - LAr Calorimeter full simulation
  - Gamma/Pi0 separation in full simulation with ParticleNet in FCCAnalyses
- Updated information on documentation website

Development workforce

- Core team at CERN
  - Bootstrap, lead, and support essential activities
  - As of 1 February 2023
    - 1.5+ FTE staff, 2 FTE fellows (GG, B François, J Smiesko, A Tolosa Delgado, A Sailer)
  - Recently ‘lost’ historical members: C Helsens, V Volkl
- Task force for IDEA tracker (DC + Vertex) to DD4hep
  - L Capriotti, *INFN similfellow* (-> Q1/23)
  - A Ilg, *Univ Zurich*
- Contributions from other WG (Physics Performance, …)
  - E Perez, M Selvaggi, L Portales, …
- S&C Organisational chart
  - Foreseen by the reviewed S&C mandate (see next)
  - However, given the small group size, until now it was not applied …

FCC S&C suggested structure



- Core software group at CERN
- External contributions warmly encouraged
- Connection with other PED groups

PP	Physics Performance
DC	Detector Concepts
MDI	Machine Detector Interface
K4	Key4hep
E4	EDM4hep

Take away messages

- S&C plays a crucial role in PED studies
- Needs to adequate support by the FCC community
  - For concrete opportunities come to the IFNC session, Thursday 13h50
- Synergy with ‘stakeholder’ WG (Physics Performance, Detector Concepts, …) is essential to make the evaluation of the FCC physics potential as realistic as possible

Backup

Useful pointers

Project repositories

- GitHub: https://github.com/HEP-FCC, https://github.com/key4hep
- CernVM-FS: /cvmfs/sw.hsf.org, /cvmfs/fcc.cern.ch

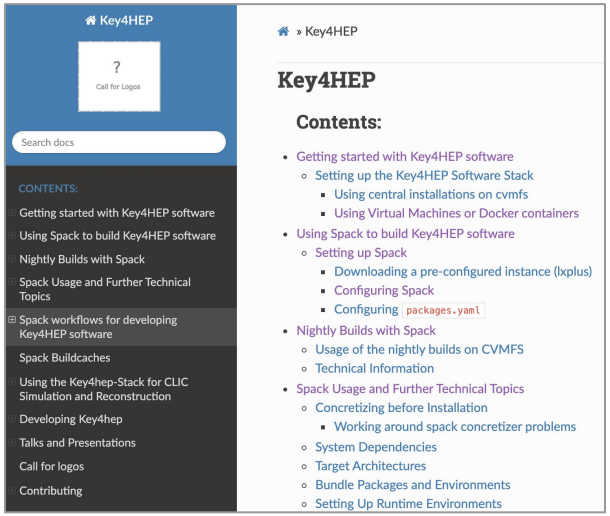
Forum: https://fccsw-forum.web.cern.ch/

Existing documentation: https://hep-fcc.github.io/fcc-tutorials/index.html

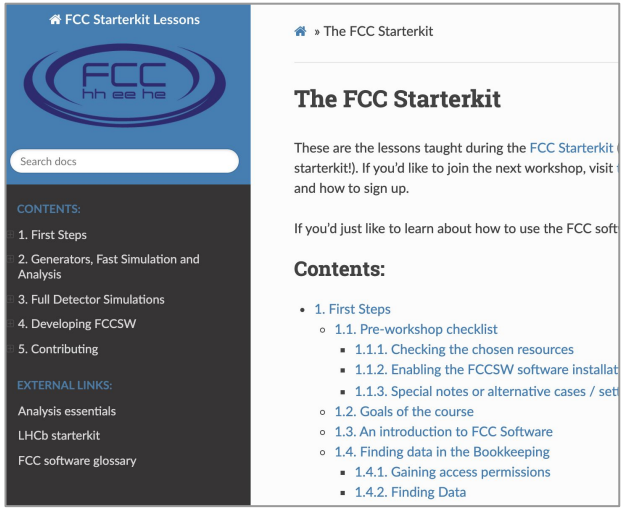
EPJ+ Software & Computing contributions (Part IV)

- Accelerator-related codes and their interplay with the experiment's software
- Online computing challenges: detector and readout requirements
- Offline Computing resources for FCC-ee and related challenges
- Key4hep, a framework for future HEP experiments and its use in FCC

Documentation, tutorials, …



Key4hep [GitHub Project](#)
[Main documentation page](#)
[Doxygen software documentation](#)



FCCSW [GitHub Project](#)
[Main documentation page](#)