FUTURE
CIRCULAR
COLLIDER

Physics and Detector Studies for FCC

Software and Computing

XV Latin American Symposium on High Energy Physics
Mexico City, Mexico

November 8, 2024

Gerardo Ganis and **Brieuc François** (CERN-EP)

FCC Software and Computing Mandate*

- Initial goal: support the software and computing needs of the FCC Feasibility Study
  - Recently extended to include the “pre-TDR” phase with the objective to **leave a *usable* system for proto-collaborations to start with**

Legend of the talk

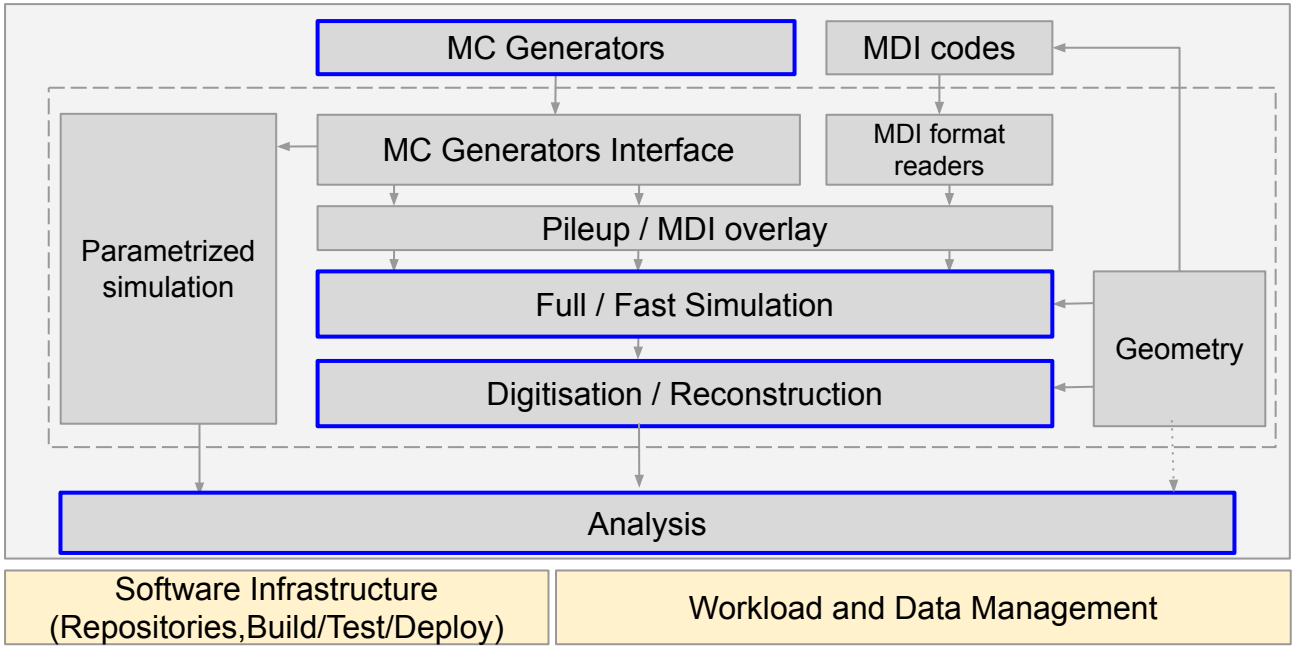
What has been done so far

What is ahead of us

Remarks

(*) Post-ESUPP 2019 mandate from FCC PED Coordination, review by dedicated Task Force 2021-2022

Typical workflows to support

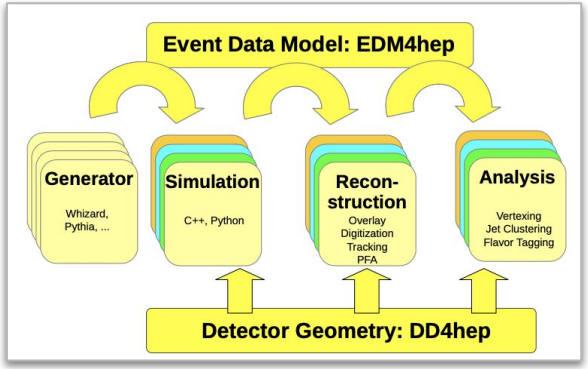


Key4hep, the common software vision

FUTURE CIRCULAR COLLIDER

FCC software is part and parcel of **Key4hep**

- Common software initiative created in 2019 to **optimise efforts** in the context of **future colliders studies**
  - Based on the belief that **software commonality between future experiments can be pushed further** than it was so far at LHC
  - **Unify communities**, contributions from CLIC, ILC, FCC, CEPC, EIC, …
  - Aim at providing a **complete set of tools** supporting generation, simulation, reconstruction, analysis, …



Current ingredients based on what was available in experiment-free way in 2019

- ❏ **Spack** package manager
- ❏ **Gaudi** framework, devel/used for (HL-)LHC
- ❏ **DD4hep** for geometry, adopted at LHC
- ❏ Podio based **EDM4hep** data model

Having Common building blocks is a must to leverage synergies and facilitates data openness and preservation!

Opportunities ahead

- ❏ Thread safety
- ❏ Framework homogenization
- ❏ Keep assessing choices made so far
- ❏ Monitor and incorporate emerging tools
- ❏ …

Nothing is written in stone! Things can be changed (minimizing disruption)

Full support by ECFA, AIDA, CERN EP R&D

Deliverables already used in large scale productions

Kick-off meetings: Bologna, Hong Kong

Weekly working meetings

Monte Carlo Generators

FUTURE CIRCULAR COLLIDER

Relevant generators available through Key4hep, including LEP ones (still state-of-art in some cases)

- KKMCee (v5), Whizard, Sherpa, Herwig3, BabaYaga, MadGraph5, Pythia8, …
- Some ported from **Fortran to C++** (e.g. KKMC, S. Jadach, FCCW 2022)

Interface to Key4hep processing chains (k4Gen)

- Readers for the relevant formats (e.g. hepmc3)
- Integration of **MC generation directly in Gaudi**
  - GEN(-SIM-)DIGI-RECO in one config
  - Only Pythia8 so far

Python based module for **automatic MC configuration file generation**: k4GeneratorsConfig

- Sherpa, Whizard, MadGraph, KKMC, Pythia

A lot of work needed in the long run…

- Higher order corrections integration, improved Beam energy spread (BES), ISR, beamstrahlung treatments (FCC-ee. EPJ. Plus 136, 911 (2021))
- MC Generators should be provided as **'modern', open software packages**
  - Versioning, building, automatic testing, with documentation
  - Some generator code modernization needed
- Validation of 'LHC' generators for e^+e^- (e.g. MadGraph)
- Key4hep specific
  - Generic, a posteriori, treatment of BES for generators not providing it
  - Extend the list of generators supported by k4Gen

Full Simulation Needs

FUTURE CIRCULAR COLLIDER

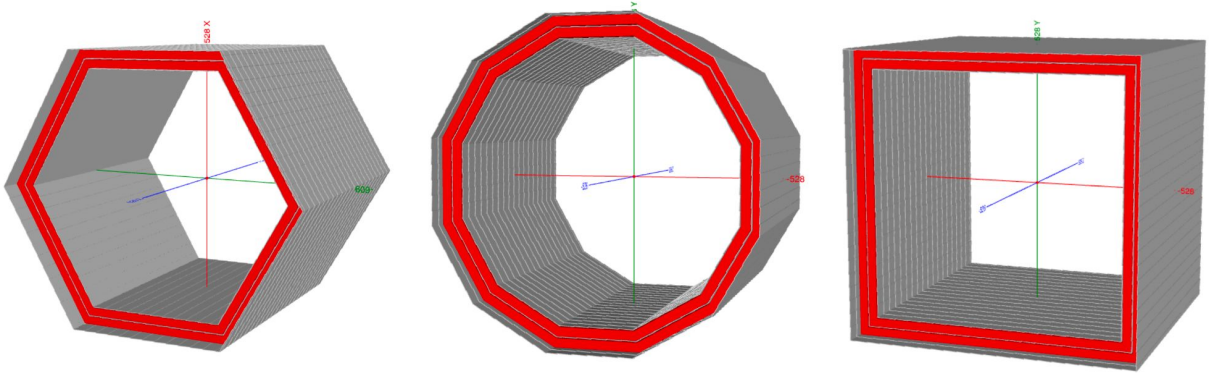
Full Simulation for all FCC detectors: **urgent and critical need!**

- FCC **physics reach** estimated mostly with parametrised simulation
  - Results need further **consolidation**
- Provide **feedback to detector R&D teams** with realistic detector performance estimations

Flexibility is a must

- Unlike for operating experiments, FCC **detector layouts are** (and will be for a long time) **continuously evolving**
- Thorough **detector optimization** campaigns require to **try many different configurations**
  - Limited number of real prototypes
  - Difficult to have test beams with all sub-detectors

Different IDEA µ-RWELL muon system options obtained by changing a single parameter



Courtesy of Mahmoud Ali

Full Simulation Status

FUTURE CIRCULAR COLLIDER

Where do we stand with FCC-ee Full Sim?

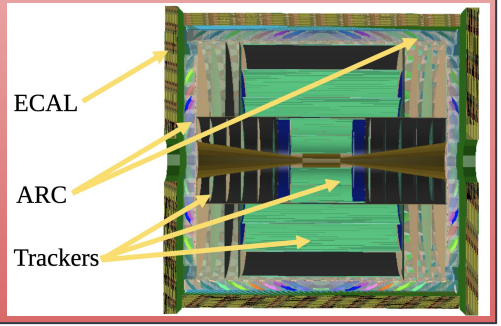
- **IDEA**: first **full geometry** model completed
- **ALLEGRO**: **calorimeters with reconstruction** available
- **CLD**: mature geometry, reconstruction including **particle flow** available (coming from CLICdet)
  - Needs further validation/optimization

Benefits from **DD4hep plug&play**: detector variants

- **CLD with a PID detector**: Array of RICH Cells (ARC)
- **CLD with ALLEGRO ECAL** to exercise PFlow
- An **ILD** version was made **for FCC-ee**
  - MDI and vertex detector taken from CLD
- **ALLEGRO tracking system taken from IDEA** for now
  - Need a dedicated muon system (no return yoke)

Important developments/studies needed

- IDEA **drift chamber detailed digitizer and tracking**
  - Evaluate impact of beam background
- **ALLEGRO particle flow** to optimize calo granularity
- CLD with ARC: gain from PID, loss on PFlow?
- **Tune against test beam data**
- Beam background to Full Sim interface
- G4 simulation in Gaudi
- And much more…



Courtesy of Alvaro Tolosa Delgado

Exciting developments at the boundary between software and detector physics!

[Full Sim meetings]
[Tutorial]
[Webpage]
[E-group]
[Geometries]

Analysis

FUTURE CIRCULAR COLLIDER

- FCC RECO or Delphes outputs are in **ROOT based EDM4hep format**
  - Primarily meant to be manipulated through the PODIO layer (C++ or Python bindings)
  - Also possible to analyze with ‘plain’ (Py)ROOT (but less user friendly, less stable)

While everyone can use their favourite **analysis framework**, FCCSW provides a **solution**: FCCAnalyses

- ❑ Used by most FCC analyzers so far
- ❑ **RDataFrame** (RDF) + **RDataSource** to keep EDM4hep/PODIO functionalities
- ❑ Automatized **central samples management**
- ❑ Skimming, filtering, histogramming, plotting, CERN HTCondor job submission
- ❑ Helper tools, **ML support** (ONNX, TMVA, XGB), ...

What is ahead of us?

- ❑ **Extend distributed computing** support (e.g. Slurm)
  - ❑ FCCAnalyses on the Grid (ILCDirac)
- ❑ **Improve** functionalities (esp. **plotting**) and code organization
- ❑ **Extend ML support**
- ❑ Improve RDF interface
- ❑ Add new interfaces (e.g. Combine)

Bringing LHC expertise will be instrumental!

Considering the project's timeline, **complete paradigm shifts can also be explored!**

- ❑ **EDM4hep** files can now also be **analyzed** with the **Julia language** (link)
- ❑ Part of JuliaHEP, a general effort proposing a **solution to the “two-language problem”** (C++/Python)

FCC-ee and the LEP data

FUTURE CIRCULAR COLLIDER

- FCC-ee and LEP share some center of mass energies
- Data from ALEPH, DELPHI and OPAL are still available but difficult to manipulate
- **Migration of LEP data to EDM4hep would**
  - Bring a **great added value** to
    - FCC-ee: get **access to still useful data**
    - LEP: get **better data preservation perspectives**
  - Enable the application of the **same algorithms on simulated and real data**
    - More realistic evaluations of the techniques under study
  - Provide opportunities for future (EW/Higgs factories) **PhD students to publish with real data**
- Ongoing TECH project to investigate **migration of ALEPH data to EDM4hep**
  - Extract data in ASCII form, running ALEPH code on SLC4/SLC6 (Singularity + CernVM-FS)
  - Injection in EDM4hep, **analysis with *standard* FCC analysis techniques**
  - Main **difficulty**: recover detailed information about the data, **not fully documented**
    - **Opportunity** for **LEP experts** to help!
  - **More details in the talk at the 4th DPHEP workshop, Oct. 2024, CERN**

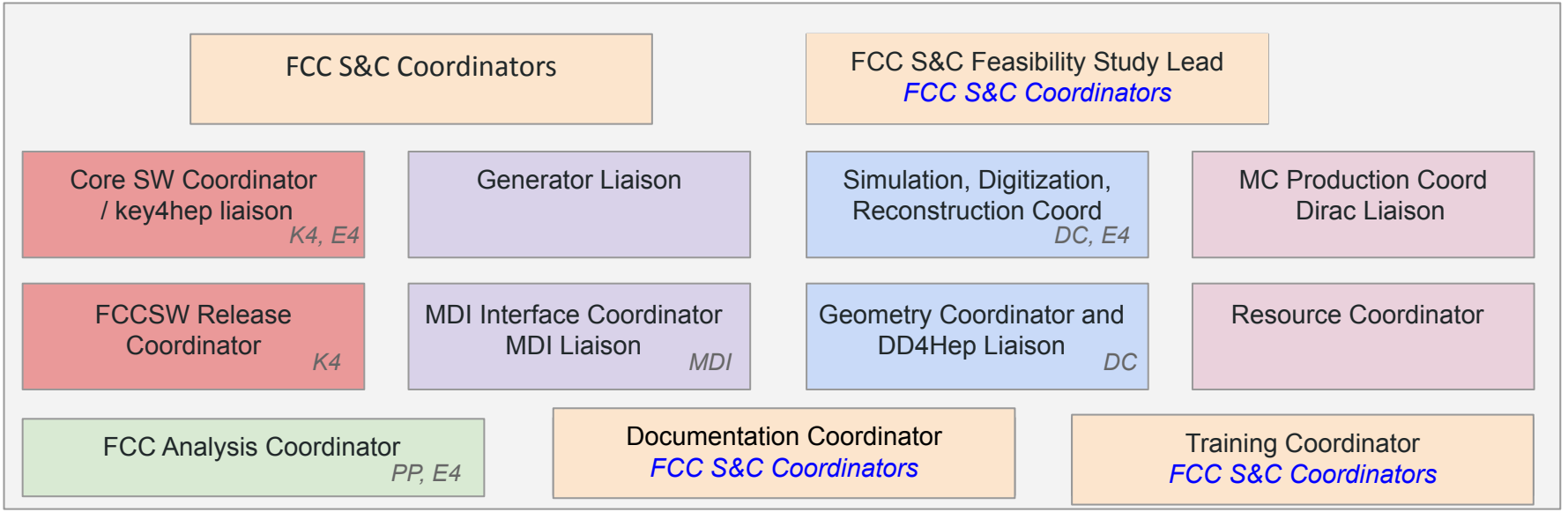
Outlook

FUTURE CIRCULAR COLLIDER

- Joining a project in **early stage is exciting**!
  - Many things to be done ➔ **you can really make a difference!**
  - **A lot of opportunities for high impact developments and studies**
    - Inject LHC techniques on analysis, reconstruction, geometry description, high-performance simulation, use of heterogeneous resources, …
    - Validate generic event generators developed for LHC, taking advantage of the acquired expertise to help understand these tools for e^+e^- physics. In the medium/long term, prepare the ground for the inclusion of the next theoretical calculations
- Towards the pre-TDR phase, we need to **enlarge the community**
  - **More** (expert) **users** (debug, contribute) and **dedicated developers**
  - Keep assessing the choice made and evaluate/incorporate new tools
  - More **structured approach** with well identified **responsibilities**
    - A lot of the work done so far is on a best-effort/interest basis
  - **More computing resources** available through the FCC virtual organization (GRID)
- The FCC Software is ready to be used and warmly welcomes new contributors!

Additional Material

FCC S&C suggested structure (TF '21-'22)



- 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

Sample Production and Resources

FUTURE CIRCULAR COLLIDER

Available samples

- Many Delphes samples available
  - CERN HTCondor
- **Grid submission** tool ready: ILCDirac
  - Used to produce **first Full Sim samples** with CLD (Higgs recoil mass and couplings)
- Currently trying to port **ALEPH data to EDM4hep**
  - Will make samples available centrally to enable analyses with real data (PhD's!)
- Everything done so far with 700 TB
  - **1 ‰ of LHC resource**

Upcoming Central Productions

- Full Sim getting ready ➜ resources allocated to FCC have to scale up (storage and CPU)
  - Precise estimation of the needs has to be done
- Assess/adapt/develop MC generators
  - FCC-ee statistical precision and clean environment will push the limits
- Set-up MC production campaign procedures
  - Pre-production, validation, …
  - Sample content per production step
    - GEN-SIM-DIGI-RECO-?
  - LHC experience highly valuable!