Software for Future Colliders

Alvaro Tolosa-Delgado (CERN)

XVI CPAN days, Madrid (Spain)

Nov. 19th, 2024

Personal experience

- 5 yr physics deg. (Valladolid/Zaragoza) + Master in nuclear physics (Madrid)

Personal experience

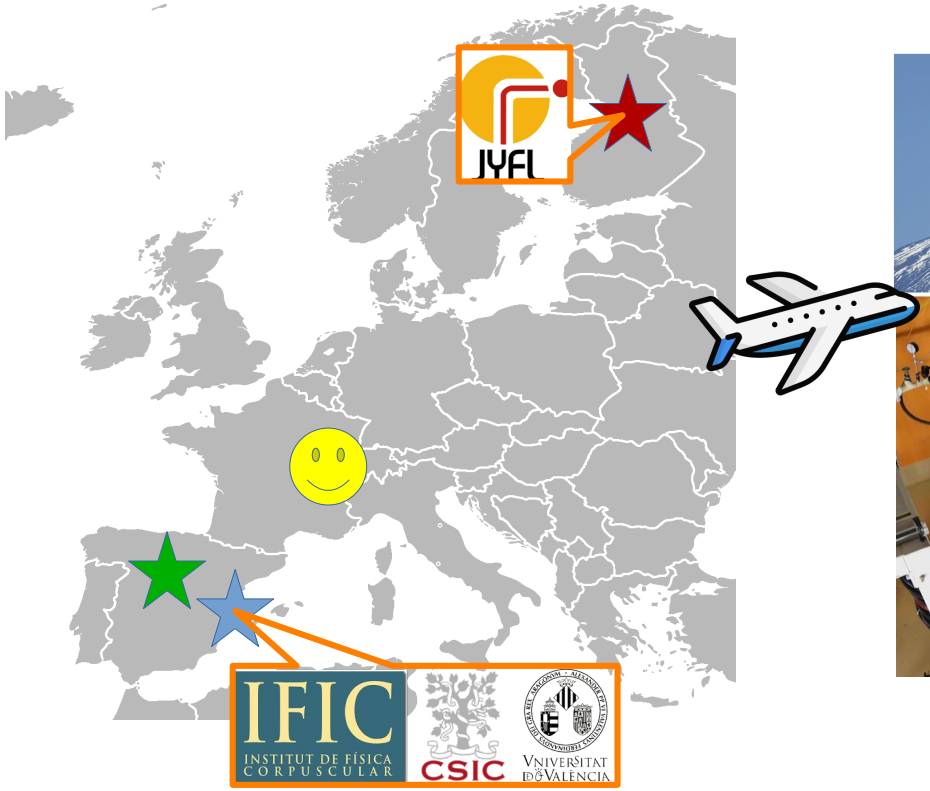
- 5 yr physics deg. (Valladolid/Zaragoza) + Master in nuclear physics (Madrid)
- PhD thesis about beta-delayed neutron emission within the BRIKEN project

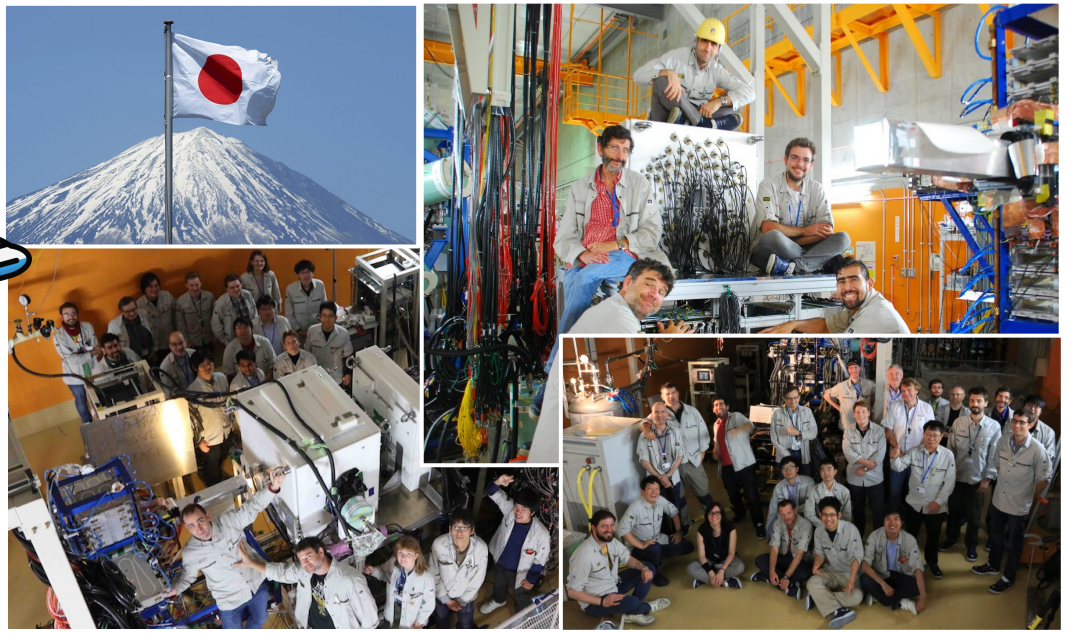
Personal experience

- 5 yr physics deg. (Valladolid/Zaragoza) + Master in nuclear physics (Madrid)
- PhD thesis about beta-delayed neutron emission within the BRIKEN project
- Postdoc at JYFL accelerator laboratory as part of MARA separator group

Personal experience

- 5 yr physics deg. (Valladolid/Zaragoza) + Master in nuclear physics (Madrid)
- PhD thesis about beta-delayed neutron emission within the BRIKEN project
- Postdoc at JYFL accelerator laboratory as part of MARA separator group
- Fellow at CERN, integrated into the FCC software group (now EP-FCC)



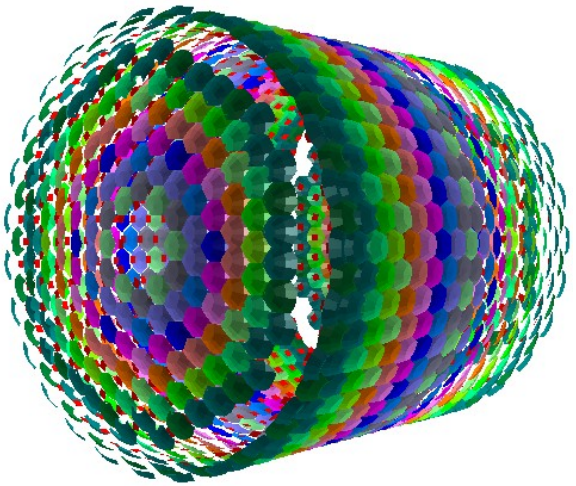


Personal experience

- I joined FCC team at CERN in 2022 as applied fellow (category II) to support detector simulations

Personal experience

- I joined FCC team at CERN in 2022 as applied fellow (category II) to support detector simulations
- I have work on two main subdetectors:
  - Array or RICH* Cells (ARC),
  - Implementation of geometry and optical properties
  - Co-supervision of master student who implemented the first step of the event reconstruction [link]
  - Creation of a new CLD option that include ARC

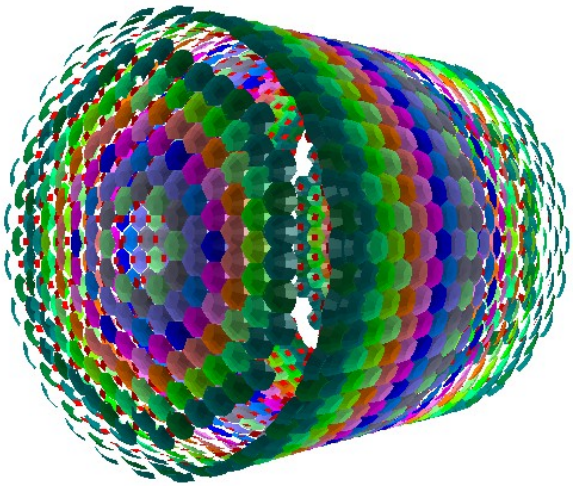


Full geometry of ARC

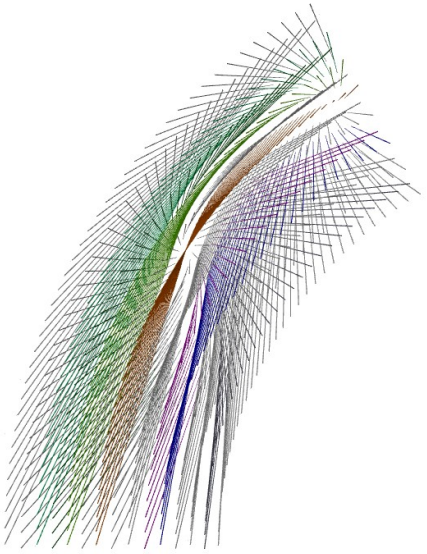
*RICH: Ring Image Cherenkov detector

Personal experience

- I joined FCC team at CERN in 2022 as applied fellow (category II) to support detector simulations
- I have work on two main subdetectors:
  - Array or RICH* Cells (ARC),
  - Implementation of geometry and optical properties
  - Co-supervision of master student who implemented the first step of the event reconstruction [link]
  - Creation of a new CLD option that include ARC
- Drift Chamber (for IDEA detector concept)
  - Implementation of 2 geometries
  - New data model for this subsystem in EDM4hep
  - Digitization algorithm which includes (ionization) cluster counting information for particle identification



Full geometry of ARC



Cross-section of the inner wires of the Drift Chamber

*RICH: Ring Image Cherenkov detector

Introduction

- The 2020 update of the European strategy for particle physics has identified an electron–positron Higgs factory as the highest priority collider after LHC

Introduction

- The 2020 update of the European strategy for particle physics has identified an electron–positron Higgs factory as the highest priority collider after LHC

- Several accelerators are proposed: ILC, CLIC, CEPC, FCC, IMCC...

Introduction

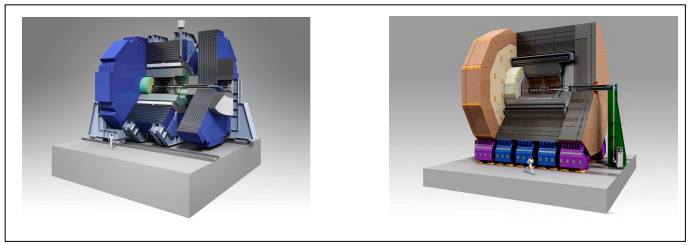
- The 2020 update of the European strategy for particle physics has identified an electron–positron Higgs factory as the highest priority collider after LHC

- Several accelerators are proposed: ILC, CLIC, CEPC, FCC, IMCC...

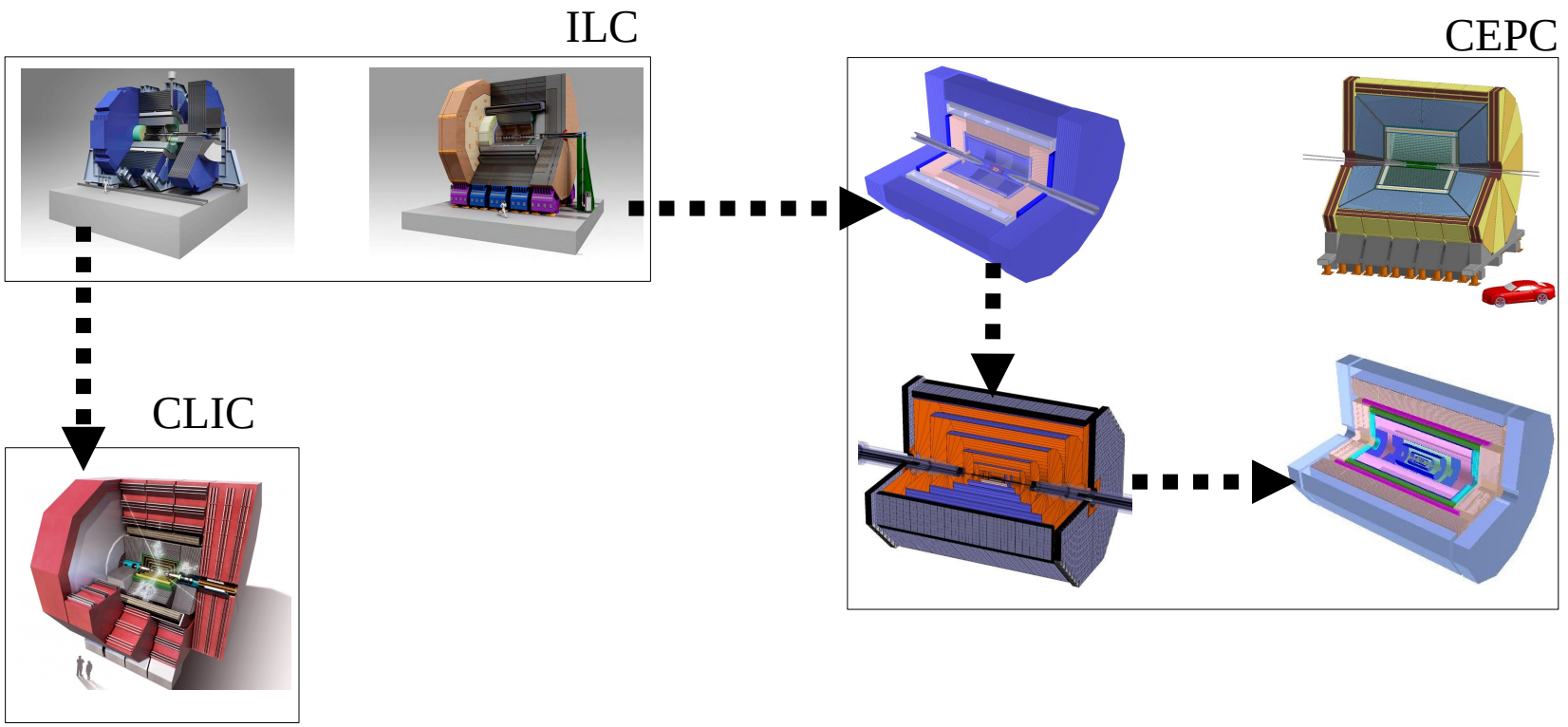
- Different detector concepts are being developed, where **simulation is the key to identify the ones that maximize detector performance for the physics cases**
  - See Rebeca's talk [link] about physics cases, and Marçà's talk [link] about hardware

Introduction. Future collider landscape

ILC

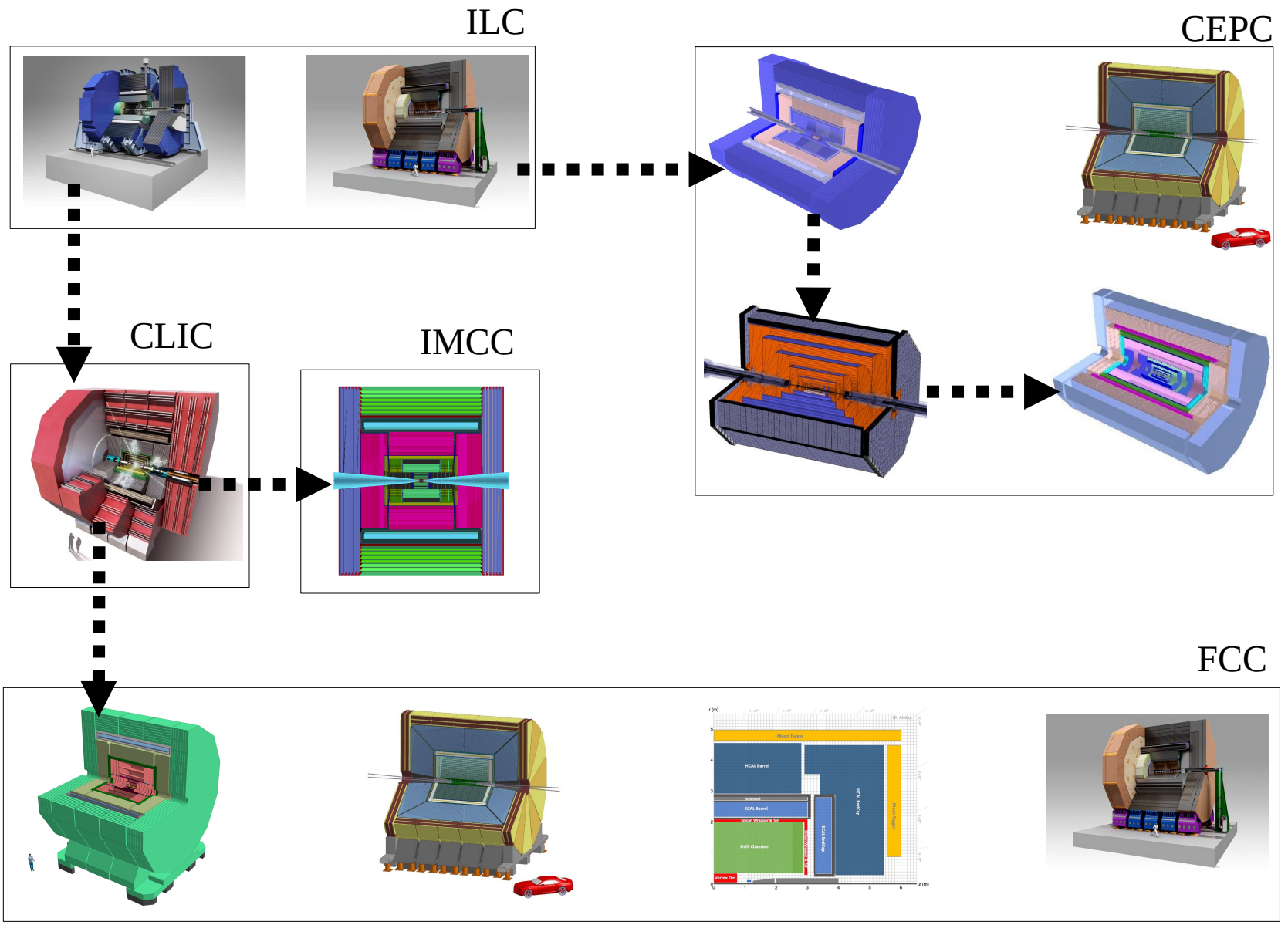


Introduction. Future collider landscape



The arrows show the historical evolution of some detector concepts

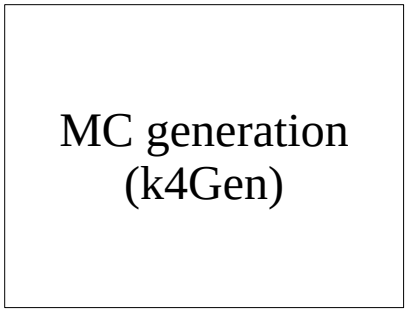
Introduction. Future collider landscape



The arrows show the historical evolution of some detector concepts

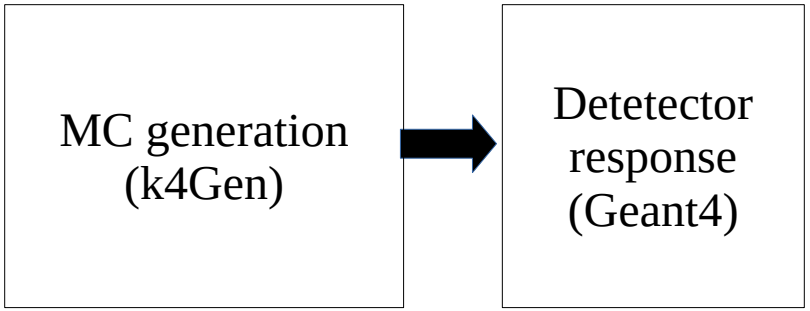
Introduction. Full simulation chain

- Full chain consist in the following steps:
  - Monte Carlo (MC) particle generation



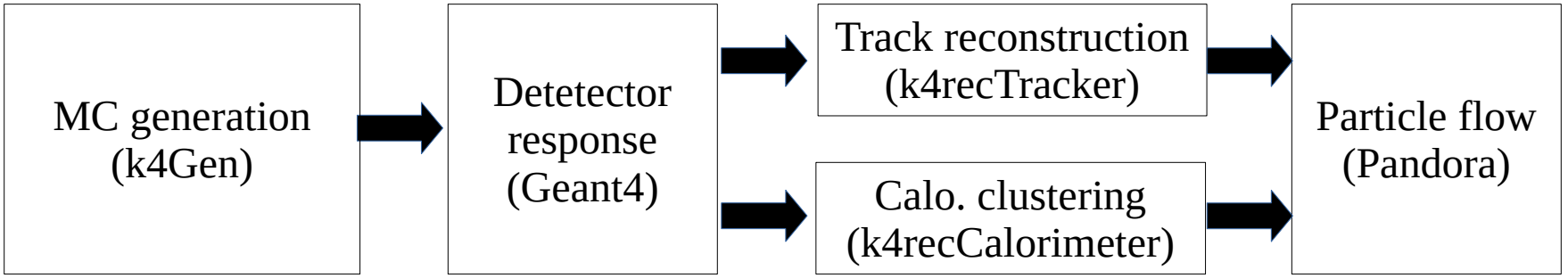
Introduction. Full simulation chain

- Full chain consist in the following steps:
  - Monte Carlo (MC) particle generation
  - Detector response simulation (Geant4)



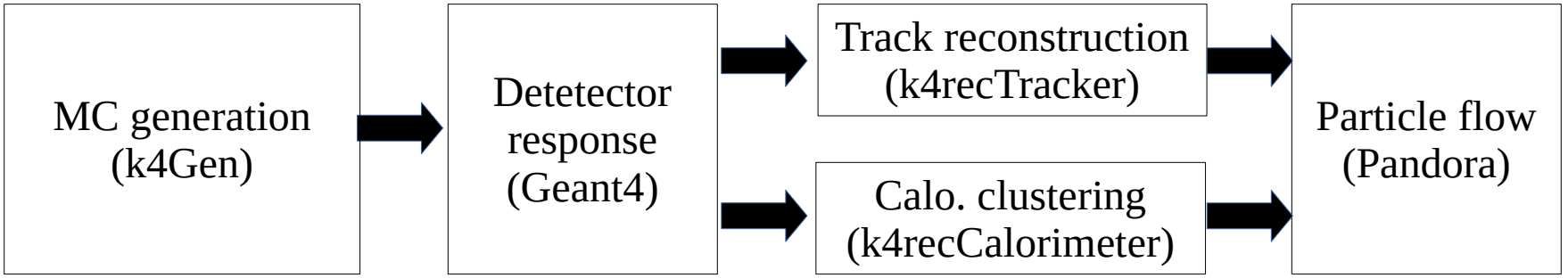
Introduction. Full simulation chain

- Full chain consist in the following steps:
  - Monte Carlo (MC) particle generation
  - Detector response simulation (Geant4)
  - Event reconstruction
    - Track reconstruction, calorimeter cell clustering
    - Particle flow



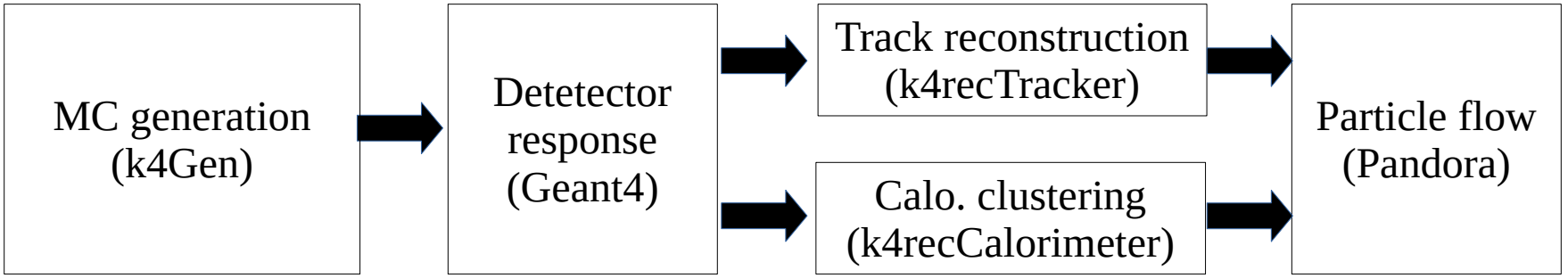
Introduction. Full simulation chain

- Full chain consist in the following steps:
  - Monte Carlo (MC) particle generation
  - Detector response simulation (Geant4)
  - Event reconstruction
    - Track reconstruction, calorimeter cell clustering
    - Particle flow
  - Analyses may focus on each step or include all of them



Introduction. Full simulation chain

- Full chain consist in the following steps:
  - Monte Carlo (MC) particle generation
  - Detector response simulation (Geant4)
  - Event reconstruction
    - Track reconstruction, calorimeter cell clustering
    - Particle flow
  - Analyses may focus on each step or include all of them
- If detailed detector geometry is not available, analyses use parametric simulations (mainly based on Delphes), instead of the simulation and reconstruction chain



Key4hep

- Key4hep [link] is a collection of consistently built and configured packages serving the future collider community to perform detector simulations
  - Prebuilt binaries are distributed via CVMFS, accesible from GRID infrastructure

Key4hep

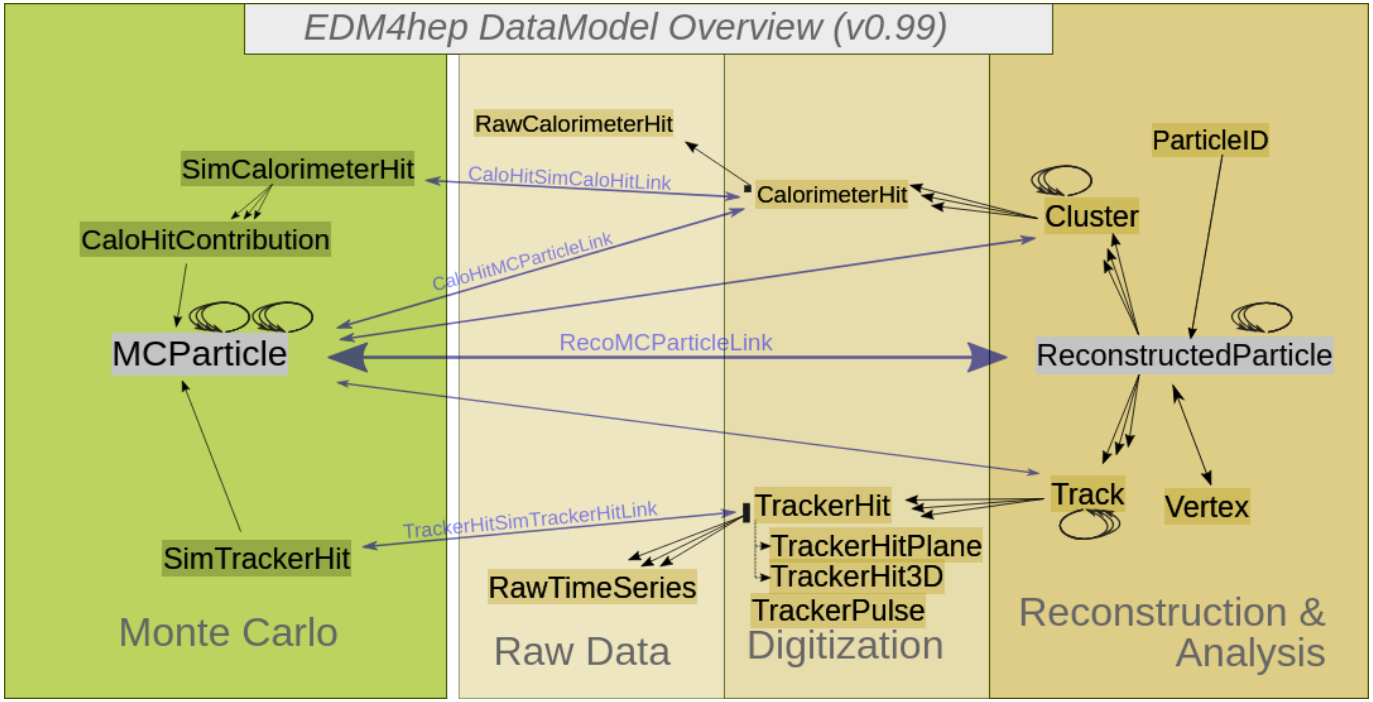
- Key4hep [link] is a collection of consistently built and configured packages serving the future collider community to perform detector simulations
  - Prebuilt binaries are distributed via CVMFS, accesible from GRID infrastructure
- The main packages are:
  - Detector geometry description: DD4hep (LHCb, CMS)
  - Data format for input/output: EDM4hep
  - Algorithm framework: Gaudi (LHCb, ATLAS)
  - HEP packages (ROOT, Geant4) and other repos dedicated to specific tasks

Key4hep

- Key4hep [link] is a collection of consistently built and configured packages serving the future collider community to perform detector simulations
  - Prebuilt binaries are distributed via CVMFS, accesible from GRID infrastructure
- The main packages are:
  - Detector geometry description: DD4hep (LHCb, CMS)
  - Data format for input/output: EDM4hep
  - Algorithm framework: Gaudi (LHCb, ATLAS)
  - HEP packages (ROOT, Geant4) and other repos dedicated to specific tasks
- In addition, key4hep include an extensive collection of algorithms and geometries
  - This allows a fast exploration of many combinations of subsystems with a minimal workload while fostering reproducibility
  - Subdetector geometries and algorithms can be exchanged with minimal effort thanks to DD4hep, EDM4hep, and Gaudi

Key4hep. EDM4hep

- EDM4hep is the event data model of Key4hep
  - Used to Input/Output data by all components, allowing interoperability
- Data model automatically generated from a YAML file (by podio)
  - Many classes, collections (vectors) and all possible connections among them
    - It is not frozen, new classes can be added if needed (e.g., drift chamber)
  - Schema evolution: can read 'old data' with newer versions (from v1.0 onwards)



Key4hep. Gaudi

- Framework used to analyze data through the execution of a chain of algorithms acting each event
  - The event loop, memory and data IO are controlled by the framework
  - Configured via a Python steering file, while algorithms are implemented in C++

Key4hep. Gaudi

- Framework used to analyze data through the execution of a chain of algorithms acting each event
  - The event loop, memory and data IO are controlled by the framework
  - Configured via a Python steering file, while algorithms are implemented in C++
- Algorithms follow a functional programming model, where each takes EDM4hep collections as input and produces new EDM4hep collections as output
  - This ensures reproducibility, thread safety, and seamless integration into the global workflow

Key4hep. Gaudi

- Framework used to analyze data through the execution of a chain of algorithms acting each event
  - The event loop, memory and data IO are controlled by the framework
  - Configured via a Python steering file, while algorithms are implemented in C++
- Algorithms follow a functional programming model, where each takes EDM4hep collections as input and produces new EDM4hep collections as output
  - This ensures reproducibility, thread safety, and seamless integration into the global workflow
- External packages interfaced through dedicated converter/wrapper algorithms
  - Wrappers for MC Generators, Geant4, Delphes inherited from FCCSW
  - k4MarlinWrapper allows reuse of iLCSoft algorithms
  - Recent additions: k4CLUE, clustering algorithm developed for CMS HGCAL
  - Under development: k4GaudiPandora, k4ActsTracking, …

Key4hep. DD4hep

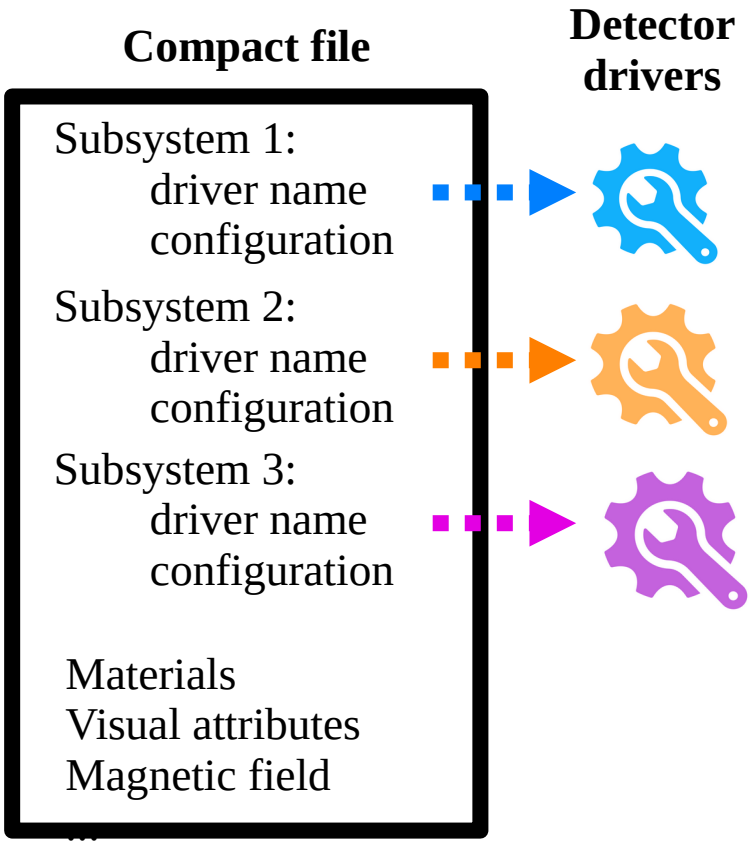
- DD4hep is a comprehensive suite that manages the detector description, including geometry and other tools necessary for simulation and reconstruction

Key4hep. DD4hep

- DD4hep is a comprehensive suite that manages the detector description, including geometry and other tools necessary for simulation and reconstruction
- Detector geometry description is the primary input for performing simulations, as Geant4 provides built-in physics, and DD4hep the mechanisms to record and write simulation data

Key4hep. DD4hep

- DD4hep is a comprehensive suite that manages the detector description, including geometry and other tools necessary for simulation and reconstruction
- Detector geometry description is the primary input for performing simulations, as Geant4 provides built-in physics, and DD4hep the mechanisms to record and write simulation data
- Each detector is described by a text file, called compact file, that contains
  - Specific configuration of each subdetector system, and a global configuration
  - Each system configuration links to a detector driver, which actually builds the geometry in memory according to the given specific configuration



Key4hep. DD4hep

- DDSim (part of DD4hep) is chosen as the way to interact with Geant4 (for now)
  - Supports many MC generator output formats (stdhep, hepmc, hepevt)
    - Provides features needed for analysis (e.g., vertex smearing and boost)
    - Particle guns (Geant4 and DD4hep) are also available
  - Highly configurable via plugin mechanism: Physics List, Sensitive Detector Actions…
    - Many plugins are provided by DD4hep, which will cover most of the use cases
    - Custom plugins can be easily developed and used instead

Key4hep. DD4hep

- DDSim (part of DD4hep) is chosen as the way to interact with Geant4 (for now)
  - Supports many MC generator output formats (stdhep, hepmc, hepevt)
    - Provides features needed for analysis (e.g., vertex smearing and boost)
    - Particle guns (Geant4 and DD4hep) are also available
  - Highly configurable via plugin mechanism: Physics List, Sensitive Detector Actions…
    - Many plugins are provided by DD4hep, which will cover most of the use cases
    - Custom plugins can be easily developed and used instead
  - Monte Carlo particle history is created automatically
  - When using EDM4hep, the sim hits include a relation to the MC particle that created it
  - Configuration from steering file, command line, or Geant4 UI command macro file

Opportunities for ECR in software development

Areas of work associated with different roles:

- Pure software development,
  - For core developers, contribute to ongoing software R&D
    - Gaudi framework
    - Event Data Model
    - Computing infrastructure
  - For advanced users
    - Implement or extend new components of key4hep packages (e.g. DD4hep plugins, including detector geometries, Gaudi algorithms, MC generators)

Opportunities for ECR in software development

Areas of work associated with different roles:

- Pure software development,
  - For core developers, contribute to ongoing software R&D
    - Gaudi framework
    - Event Data Model
    - Computing infrastructure
  - For advanced users
    - Implement or extend new components of key4hep packages (e.g. DD4hep plugins, including detector geometries, Gaudi algorithms, MC generators)
- Analysis of physics cases
  - Bridge for researchers doing analysis in running HEP experiments

Opportunities for ECR in software development

Areas of work associated with different roles:

- Pure software development,
  - For core developers, contribute to ongoing software R&D
    - Gaudi framework
    - Event Data Model
    - Computing infrastructure
  - For advanced users
    - Implement or extend new components of key4hep packages (e.g. DD4hep plugins, including detector geometries, Gaudi algorithms, MC generators)
- Analysis of physics cases
  - Bridge for researchers doing analysis in running HEP experiments
- Detector R&D
  - Broad spectrum of activities related to running beam tests and simulations of the detector prototypes

Opportunities for ECR in software development

Possibilities to join efforts

- CERN has a large variety of programs for different entry levels:
  - Research graduates (PhD), fellow program (categories I and II)
    - Two general calls per year, and asynchronous openings on specific topics
  - PhD programs, Marie Curie PhD program [link]
  - Experienced graduates, Technical student intership and Quest program [link]

Opportunities for ECR in software development

Possibilities to join efforts

- CERN has a large variety of programs for different entry levels:
  - Research graduates (PhD), fellow program (categories I and II)
    - Two general calls per year, and asynchronous openings on specific topics
  - PhD programs, Marie Curie PhD program [link]
  - Experienced graduates, Technical student intership and Quest program [link]
- Many other institutions contribute to future collider activities (physics analysis, DRD), so it is better to contact either them directly or though a contact person from CERN
  - At the moment, some countries apply an upper limit to the fraction of the FTE be spent on future colliders (e.g., France limit is 50% for postdoc positions)

Community engagement

Join upcoming events

- FCC workshop meeting on January 2025 at CERN
- Key4hep tutorials happening in Europe next year
- DRDs are organizing schools, which is a great opportunity to learn about instrumentation and the current activities of the DRD**s**

Community engagement

Join upcoming events

- FCC workshop meeting on January 2025 at CERN
- Key4hep tutorials happening in Europe next year
- DRDs are organizing schools, which is a great opportunity to learn about instrumentation and the current activities of the DRDs

The CERN group EP-FCC has regular meetings about software and physics analysis, join to learn the details about the ongoing work and to let yourself being known

- Detector full simulation, biweekly [link]
- Physics cases analyses, biweekly [link]
- Physics performance, monthly [link]
- General software meeting, monthly [link]

Community engagement

Join upcoming events

- FCC workshop meeting on January 2025 at CERN
- Key4hep tutorials happening in Europe next year
- DRDs are organizing schools, which is a great opportunity to learn about instrumentation and the current activities of the DRD

The CERN group EP-FCC has regular meetings about software and physics analysis, join to learn the details about the ongoing work and to let yourself being known

- Detector full simulation, biweekly [link]
- Physics cases analyses, biweekly [link]
- Physics performance, monthly [link]
- General software meeting, monthly [link]

Contact the DRD group of your interest [link]

You can contact me directly or through mailing list [link] for further details

Summary

- Software for future colliders is an active field
  - Integrating existing tools and being the test-bed of new ones (R&D)
- There are many opportunities in software for future colliders
  - It is a broad field, aiming to include core software developers, advanced users and physics analysis experts
- There will be positions at CERN and other institutions supporting future colliders
  - Including different entry levels, from undergraduate to PhD graduates
- CERN will make up a decision about the next collider on 2026
  - Until then, positions may be only partially devoted to future colliders

Backup slides

Software for Detector Simulation in Future Colliders

- **Key4hep** is a collaborative, community-maintained software. See J. Smiesko talk [link]
- There are two types of simulation based on the level of detail in the physics modeling:
  - Parametric simulation (mainly Delphes-based)
    - Detector responses are summarized in "Delphes cards," available for FCC, CLIC, ILC, and MuonCollider detectors [link]
    - Allows for quick estimates of physics reach and helps set detector requirements
      - Seamless generation of samples for many different detector scenarios
    - Recently extended to include more accurate tracking, cluster counting, and time-of-flight, useful for example for machine learning-based flavor tagging
  - Full simulation (Geant4-based)

This talk will focus on Full Simulation (where most of the work ahead lies)

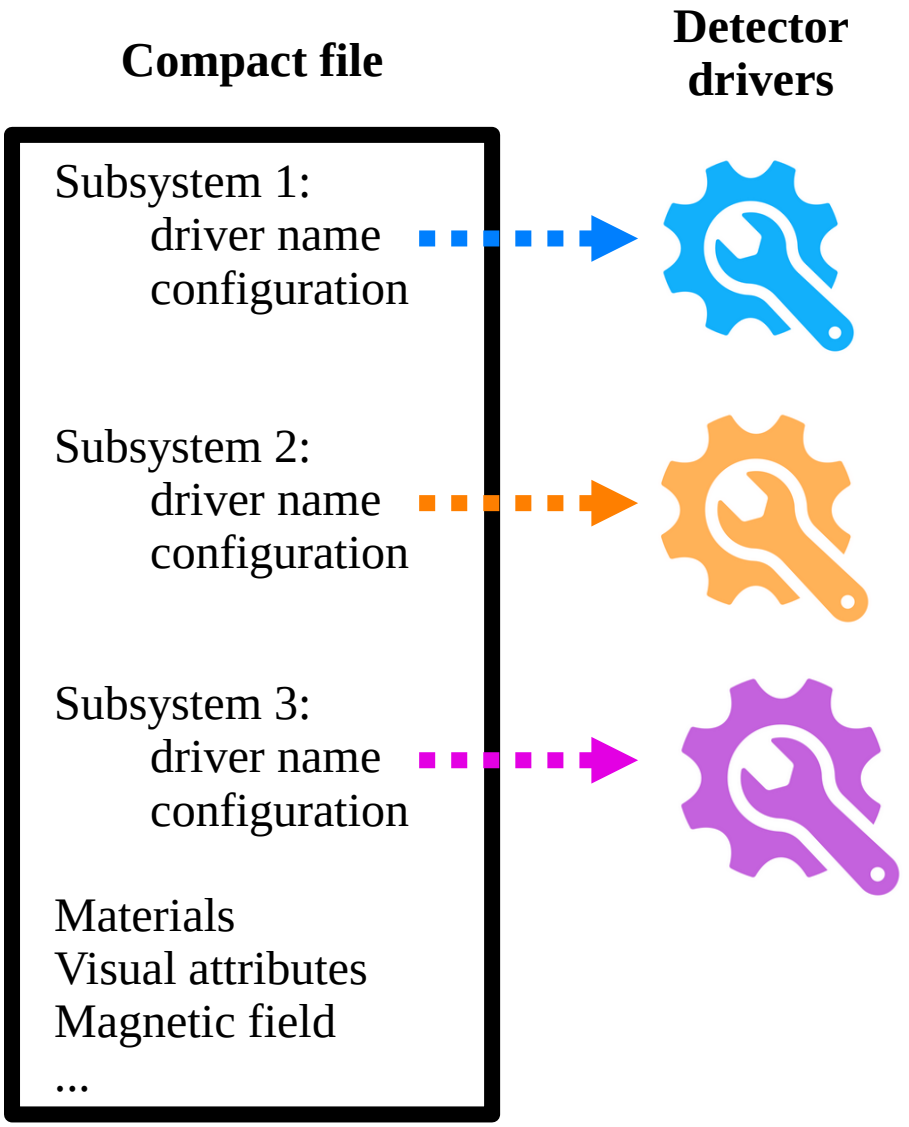
Software for Detector Simulation in Future Colliders

The main software packages used to **perform physics simulations** are:

- Geant4: a Monte Carlo particle transport code
- EDM4hep: a collection of data structures used for persistency in memory
- DD4hep: A comprehensive framework that manages the detector description, including geometry and other functionalities necessary for simulation and reconstruction.
    - Detector **geometry description is the primary input** for performing simulations, as Geant4 provides built-in physics, and DD4hep the mechanisms to record and write simulation data

DD4hep for detector geometry description

- Each detector concept is described by a text file called **compact file**
- This text file contains:
  - specific configuration of each subsystem, such as its size, number of layers, materials, etc
  - global configuration, including list of materials, fields description, etc
- Each subsystem configuration links to a **detector driver**, which builds the detector geometry in memory according to the given configuration

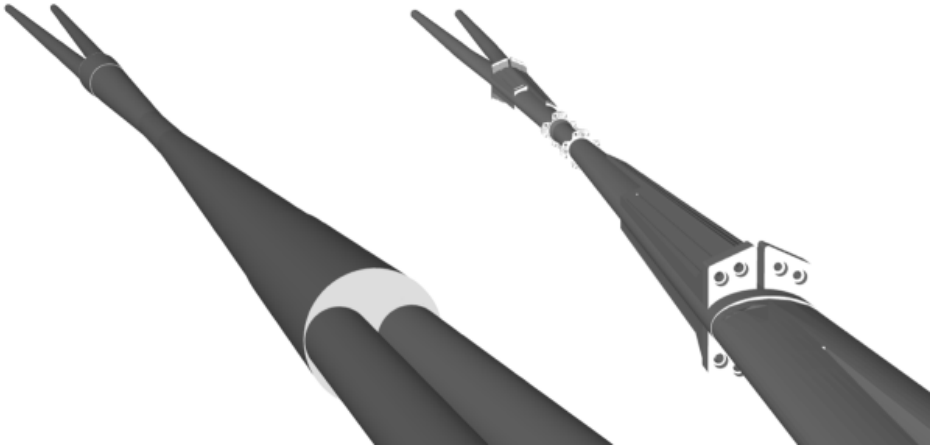


DD4hep for detector geometry description

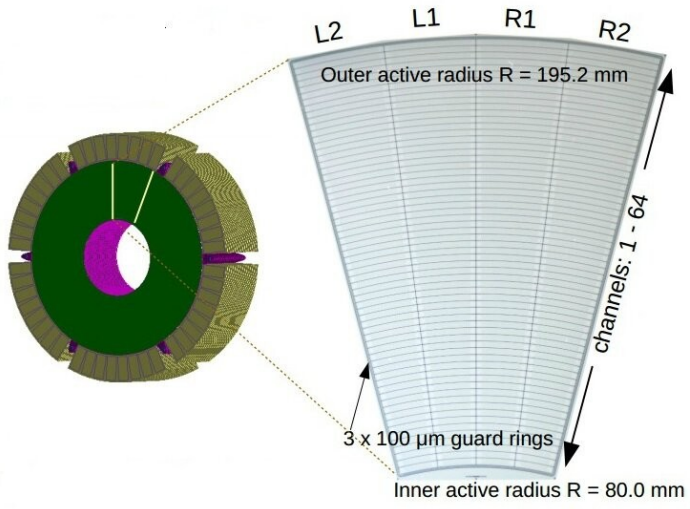
- DD4hep allows for easy swapping of subdetectors
- The **detector drivers may be shared by different detector concepts, changing only the configuration**
- This motivated that **all detector geometries for projects ILC, CLIC, and FCC are centralized in k4geo [link], a repository part of the Key4hep GitHub project**.
  - In the following, I will review the subdetectors hosted in k4geo
  - This is not an exhaustive list, as some subdetectors are hosted elsewhere (e.g. CEPCSW)
- The subsystems are grouped as:
  - Machine-Detector Interface (MDI)
  - Tracking Systems: Silicon-based and Silicon/Gas-based designs.
  - Calorimeter Systems: sampling calorimeters and dual-readout calorimeters.
  - Muon Systems
  - Particle Identification (PID) system

Machine-Detector Interface (MDI)

- MDI design is driven by the nature of the colliding beam
  - ILD and FCC have a centralized description of their MDI
- Beampipe structure is non sensitive, but detailed description is needed for acurate simulation results
  - baseline model is made up by simple geometrical shapes
  - detailed CAD design has been newly developed for FCC
- Lumimeter: a silicon-tungsten (SiW) calorimeter based on the one used by LEP-OPAL. At the moment the same detector driver is used by detectors at ILC, CLIC, and FCC.



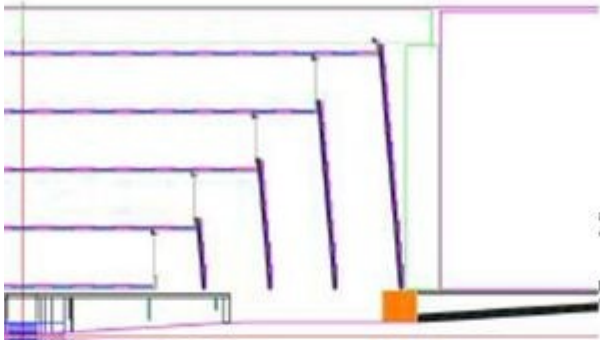
Shape-based model (left) and CAD model (right) of the FCC beampipe



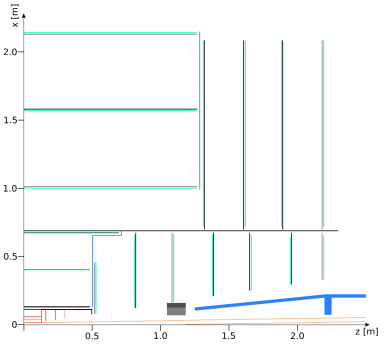
Lumimeter geometry

Full silicon-based tracker

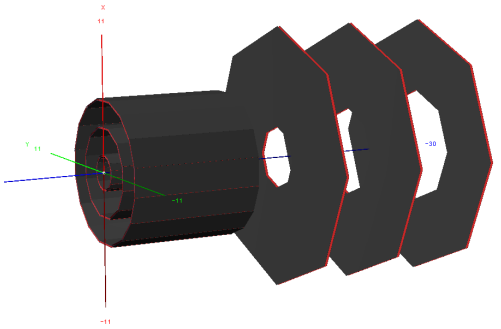
- The tracking system consists of a silicon vertex and tracker, covering both the barrel and endcap
- Two families of detector drivers are shared among different detector concepts:
  - SiD (ILC), CLICdet (CLIC), CLD (FCC)
  - ILD (ILC), CRD baseline (CEPC). TDR baseline (CEPC) is an evolution from the previous one.
- The detector driver builds the geometry as an assembly of radial or planar layers
- **Resizing tracking layers** is done by just **modifying a few lines in the compact file**
- The maturity of the geometry has paved the way for further developments, such as the ongoing integration of ACTS with the CLD tracking system



SiD (ILD)



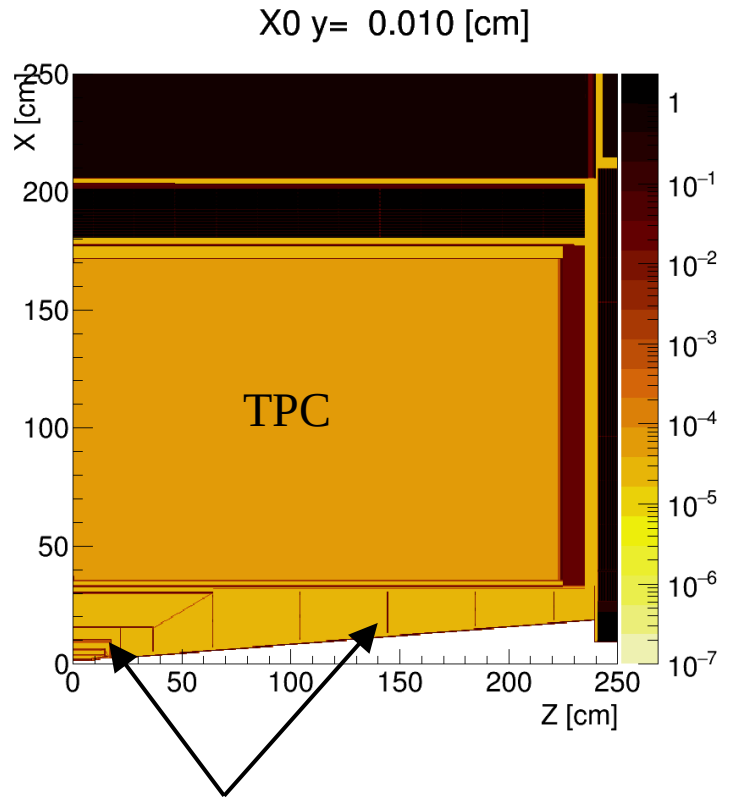
CLD (FCC)



CLD vertex (FCC)

Silicon + gaseous (Time Proj. Chamber) tracking

- The tracking system is made up by a silicon vertex and inner tracker, and a TPC as main tracker subsystem
- This combination was developed by ILD (TDR 2013) and adopted by CEPC baseline detector (CDR 2018),
- The geometry of the silicon vertex is more detailed than the one used by full silicon tracking subsystem
- The design of the TPC is an active project supported by the LC-TPC collaboration, and integrated within DRD1 of CERN.
- **A new version of the geometry has been developed for FCC**. See V. Schwan talk [link]. This version is now under study to evaluate the beam-induced backgrounds. See D. Jeans talk [link].

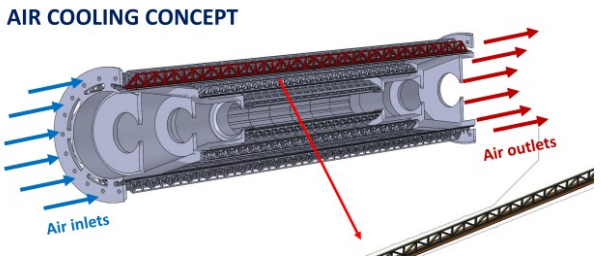


Vertex and inner silicon trackers

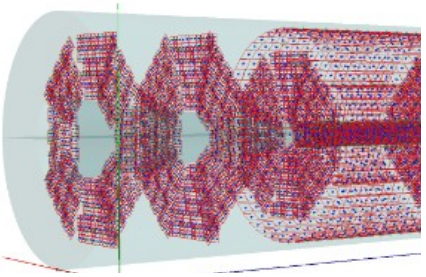
ILD cross section view of the vertex and TPC

Silicon + gaseous (Drift Chamber) tracking

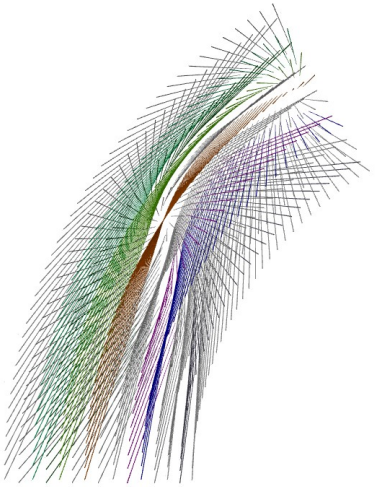
- This combination of silicon vertex and drift chamber is used by IDEA and ALLEGRO (FCC) detectors
- A **newly developed vertex detector driver**
  - Supports both single-sided flat and bent silicon layers
  - **Easy reuse** across different detector configurations.
  - Also used by the Silicon Wrapper subsystem
- The stereo drift chamber (DC)
  - **Geometry is fully implemented**
  - Geant4 **physics has to be tuned** and validated against beam test data for acurate simulations. See talk by W. Elmetenawee [link].
  - Preliminary background studies and development of new tracking algorithms are already using this implementation. See D. Garcia talk [link]



Vertex support structure



Vertex sensitive surfaces

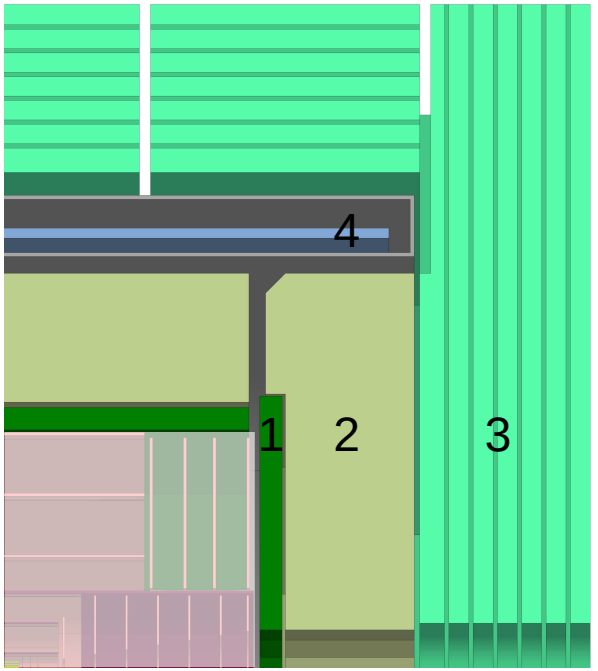


Cross-section of the wires (software)

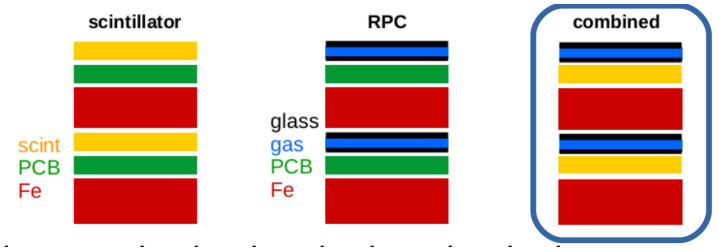
MEG II, a similar design of a DC

General sampling calorimeter

- Consist in a layered structure covering the barrel and endcap regions.
- The layer structure is given in the compact file
  - easily reused across different subsystems with minimal modifications
- The same detector driver is used to build the ECAL and HCAL of different detector concepts, including SiD (ILC), CLICdet (CLIC), and CLD (FCC)
- ILD has implemented calorimeters so that a single Geant4 run allows to simulate up to five configurations of different detector technologies. Some of these combinations are similar to those being explored in SiD.
- Baseline detector calorimeters for CEPC are built by specialized detector drivers that follow the aforementioned philosophy



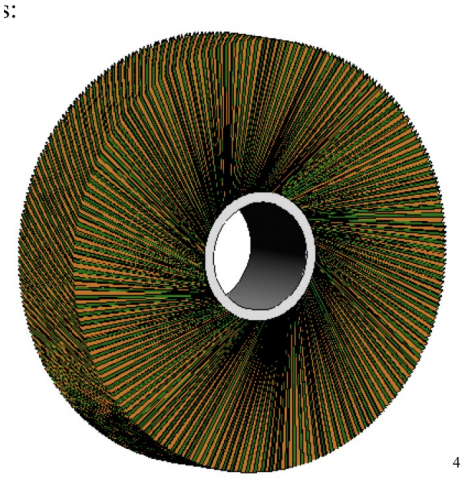
CLD outer systems:
1. ECAL
2. HCAL
3. Iron Yoke with RPC for muon ID
4. solenoid 2T



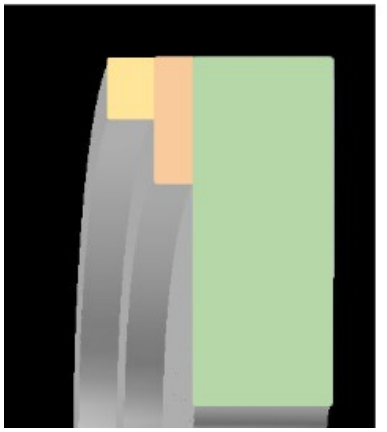
ILD calorimeter

Noble liquid ECAL + TileCal

- These calorimeters used by ALLEGRO (FCC) detector concept
- **Noble liquid ECAL** full detector geometry **is ready**.
  - Crosstalk effects implemented for the barrel region, ongoing work to extend them to the endcap. See F. Sopkova [link] and T. Li [link] talks.
  - First studies using the barrel region in CLD (FCC). See S. Sasikumar talk [link]
- **TileCal complete detector geometry is also ready.**
  - The segmentation has been updated to be projective in theta, with no significant changes in physics performance.
  - Current developments on optimizing the reconstruction steps.
  - See M. Mlynarikova talk [link]



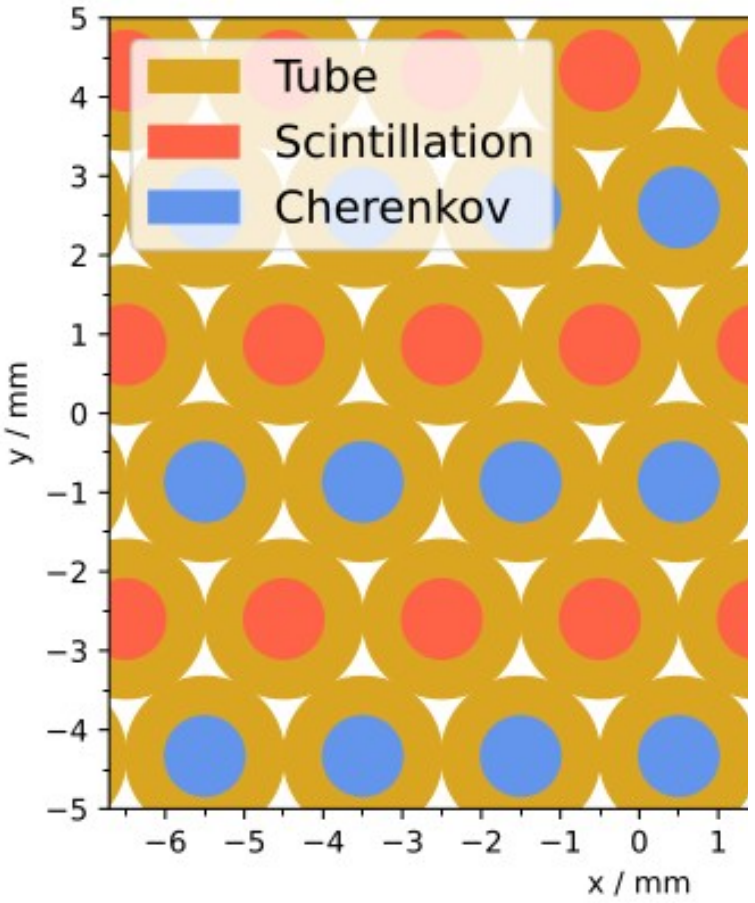
Noble liquid ECAL endcap



TileCal endcap

Dual-readout calorimeter

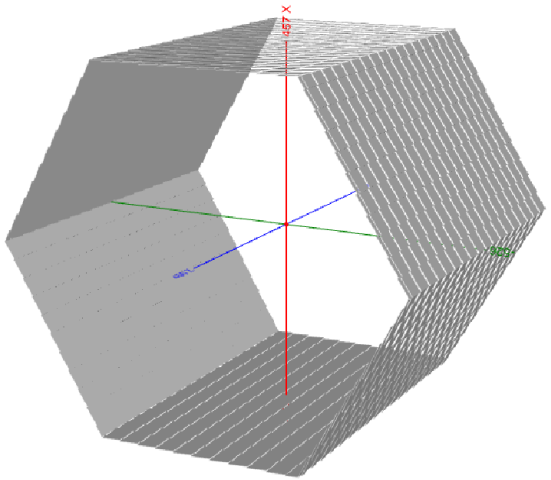
- This subdetector is so far specific to IDEA detector concept
- Two main approaches based on the absorber matrix:
  - In the monolithic option, the matrix fills the spaces between the fibers.
  - In the capillary option, each fiber is placed within a tube, leaving air gaps between the tubes.
- This is a complex detector concept, **requiring dedicated developments** such as:
  - Custom functionality to record optical photon information
  - A fast simulation for photon transport within the fibers.
- **Detector geometry and Geant4 physics models have been thoroughly validated**, see talk by R. Turra [link].

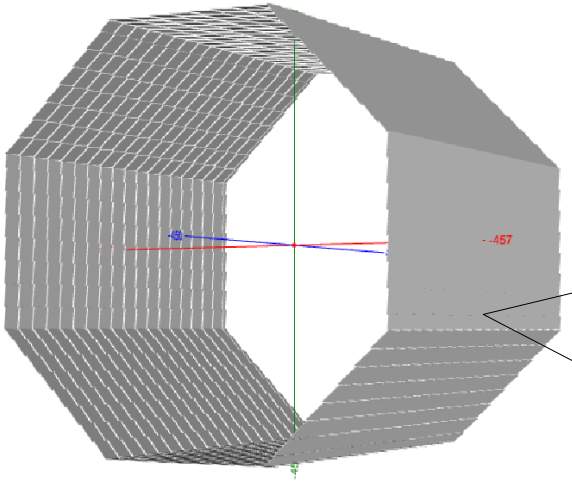


Cross-section of the DRC, capilary option [link]

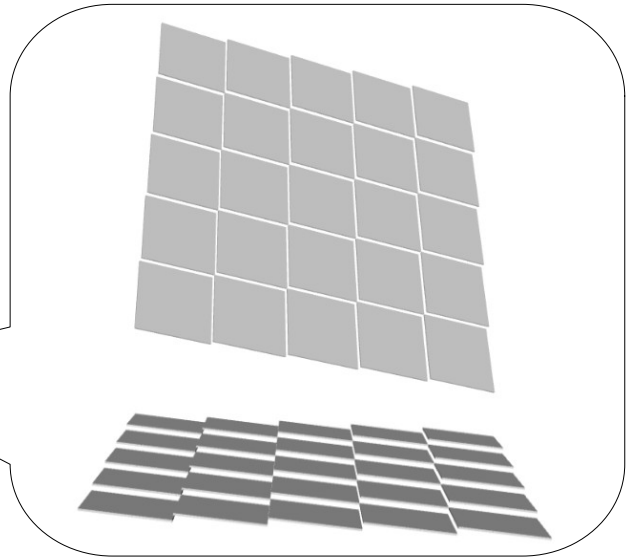
Muon systems

- Detector geometry of the instrumented iron yoke with **Resistive Plate Chambers** (RPC) is built by the same **generic calorimeter detector driver** described before.
- More detailed uRWELL-based systems added recently to k4geo
  - Polygonal geometries easy to tune (number of sides, sizes of the chamber, overlap area for hermeticity)
  - Driver used to build the IDEA muon system and pre-shower detector so far
  - More details in talk by R. Farinelli [link]





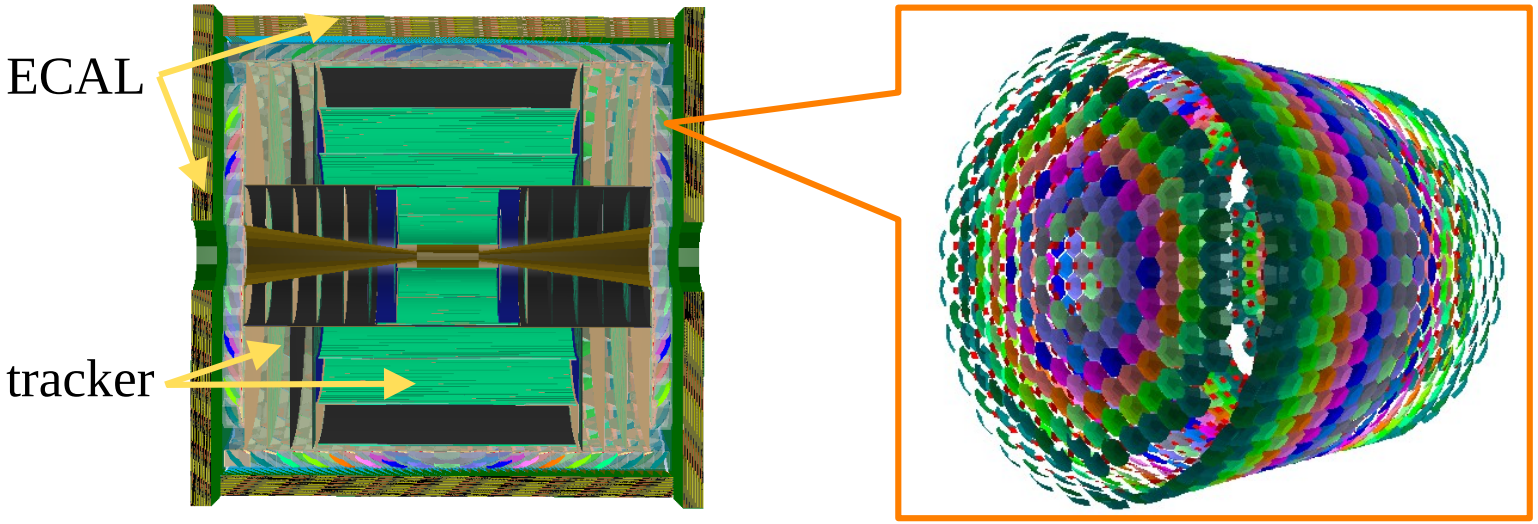
Detailed muon systems, based on uRWELL chambers



A custom overlap area is achieved by tilting the chambers

PID systems. ARC detector

- The **ARC** consist on an large array of **RICH cells** placed as in the picture below (only mirrors and sensors are visible for simplicity). The detector geometry, material description (including optical properties) and sensor readout is fully implemented in DD4hep
- **New CLD option** with a smaller tracker compared to baseline incorporates the ARC.
  - Tracking performance of this option was presented at LCWS 2024 by G. Sadowski [link].
- First ARC performance studies were presented by S. Pezzulo [link]



CLD with ARC (option 3) ARC