Software and Computing: Synergies between FCC & Linear Machines, Community Needs

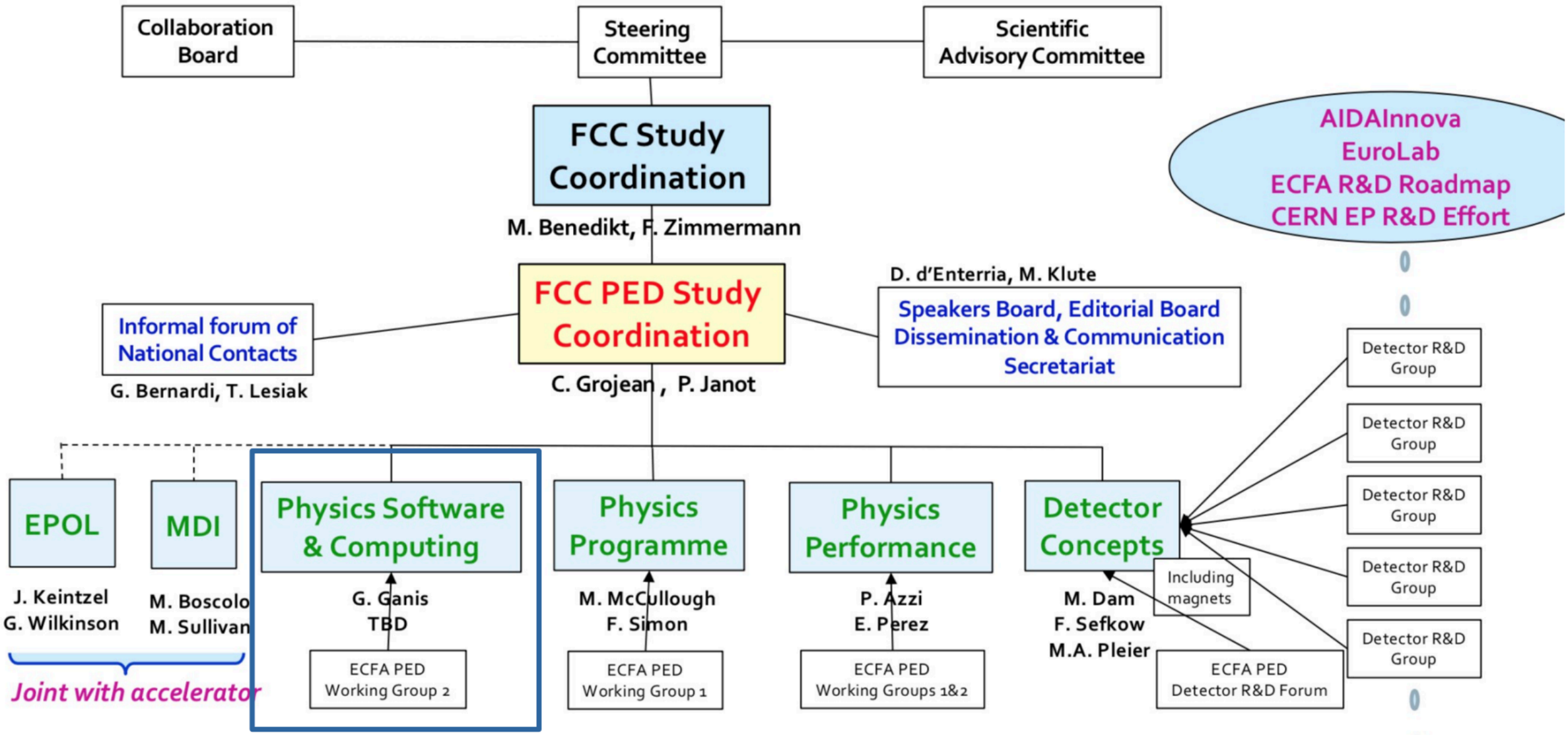
Lindsey Gray

2nd USFCC Workshop

25 March 2024

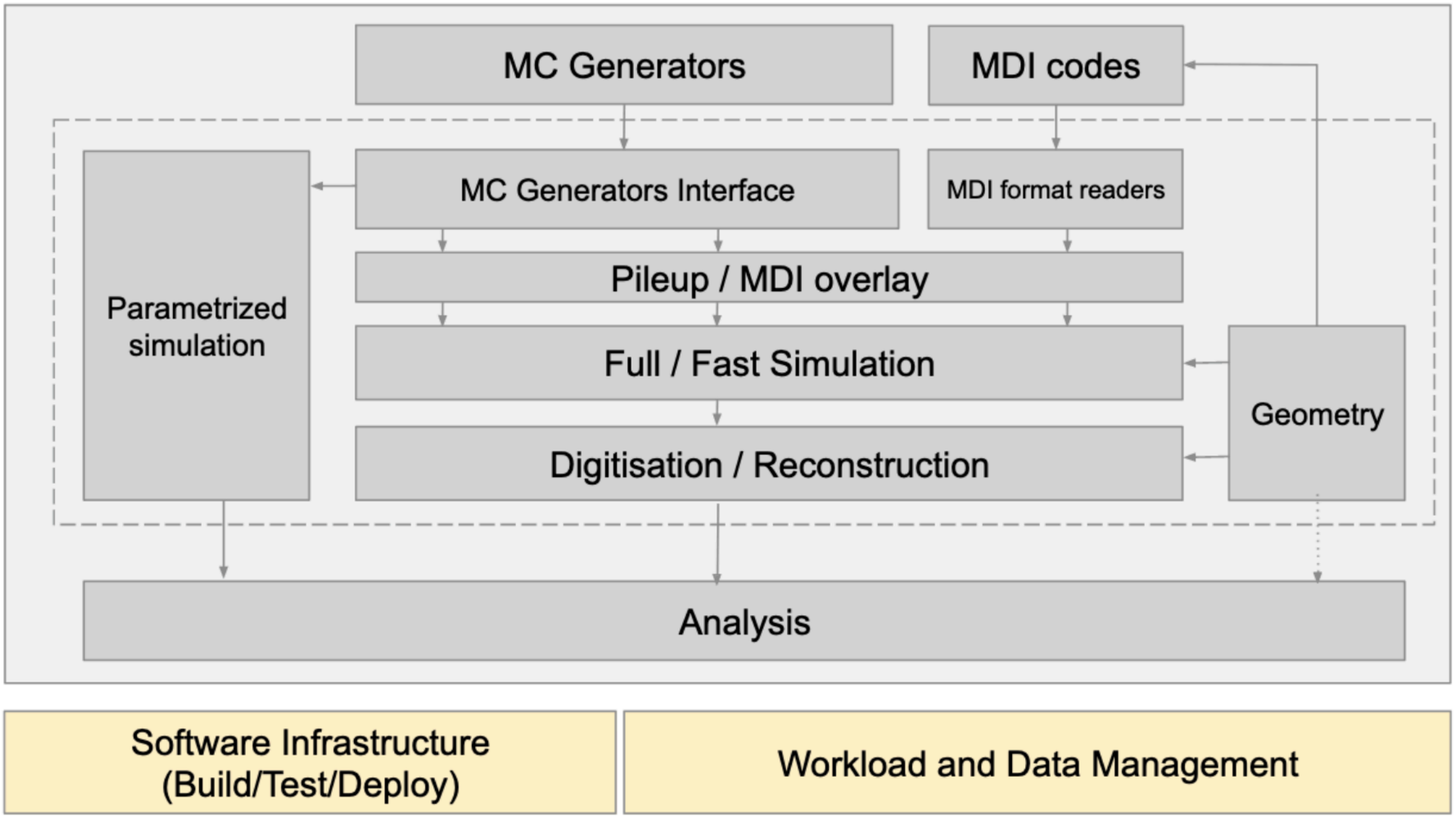
Software and Computing in Wider FCC Context

- Physics Software & Computing area develops, maintains, and provides support for software tools used by all PED study subgroups
  - Difficult to clarify common themes, serves wide variety of reconstruction and analysis needs
  - Interaction and contribution across PED is critical to developing good software



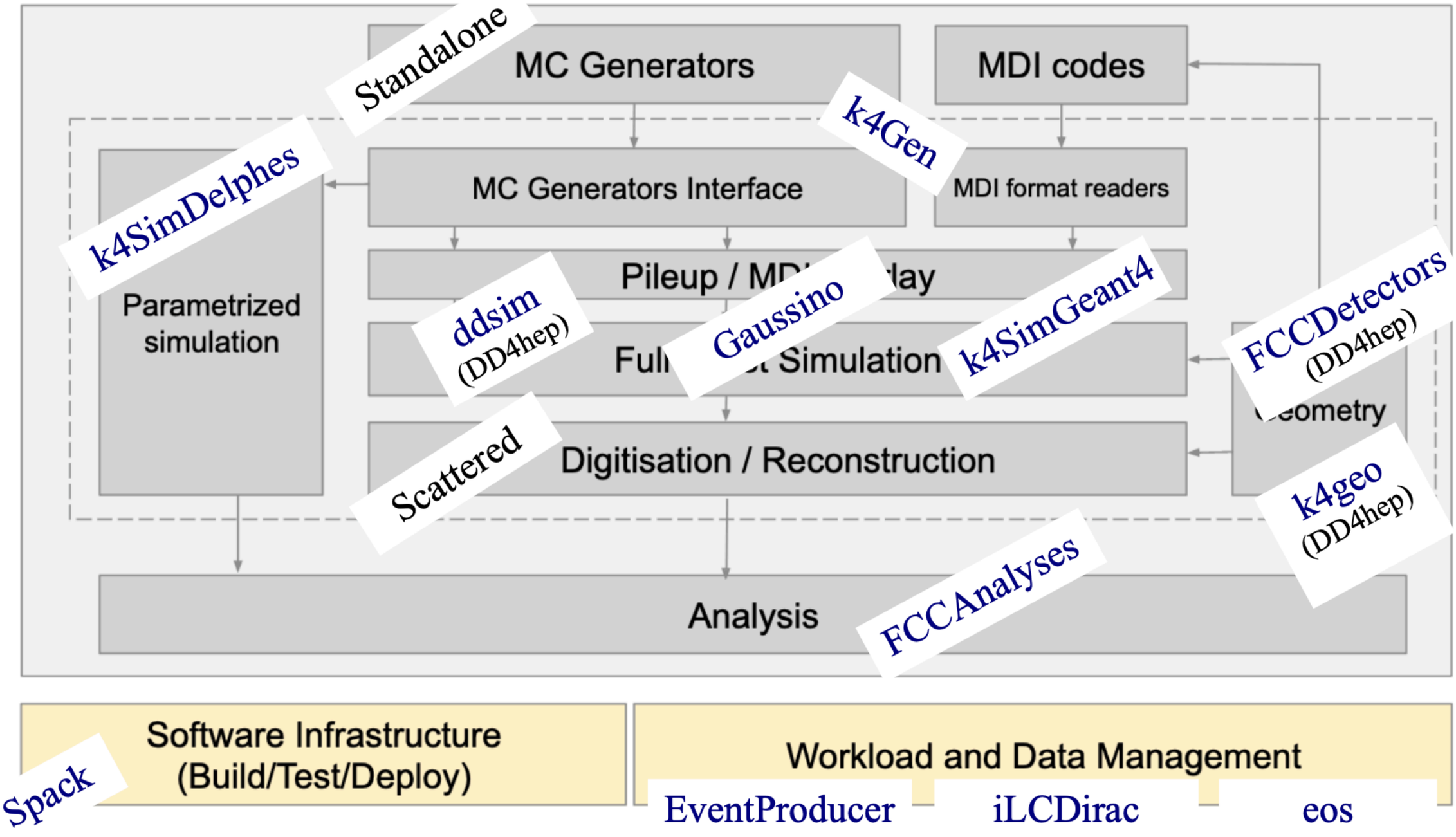
HEP Software Stack

- The typical workflow FCC Software wants to support



Fermilab

FCC-focused Software Sub-projects



Fermilab

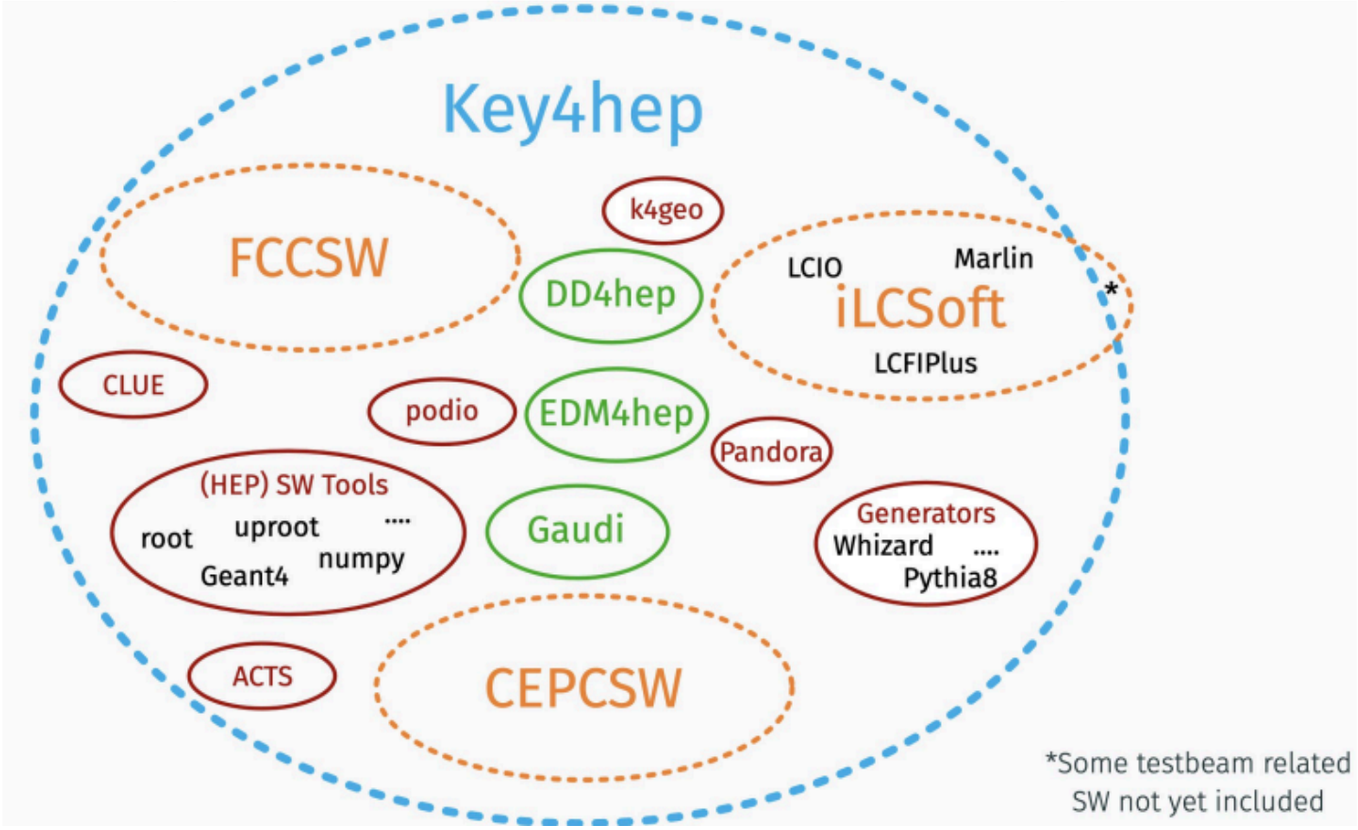
FCC Software Requirements

- Want to simultaneously
  - Plan for delivering common, modern, and high-performance software for the future
  - Deliver something presently for the feasibility studies and initial detector R&D
    - This is prioritized for now
- During pre-construction phase software is a critical resource
  - However, finding developers is difficult due to specific profile of skilled programmer who is also a capable physicist who understands the intended uses of the software
  - Need to share developments as widely as possible to manage wasted effort
  - A shared software infrastructure (key4hep) is necessary to function in these constraints
- Should note that needs and specialization evolve with operating experiments
  - Building to the specific needs of a built detector is at odds with flexibility and R&D
  - Serving all the detailed needs of a specific experiment demands reduced versatility
  - Code re-use more difficult after hyper-optimization to specific experiment

- key4hep ecosystem desires to meet needs and deform smoothly to future requirements

Quick overview of key4hep ecosystem

- Key ingredients in Key4hep aim to maximize synergies
  - Common algorithm orchestration framework: Gaudi (LHCb, ATLAS)
  - Common data format for algorithm input/output: edm4hep
  - Common detector geometry construction tool: DD4hep (plug-and-play, CMS, LHCb)
  - A set of packages of general interest is provided through the Spack package manager
    - Available on cvmfs through /cvmfs/sw.hsf.org/key4hep/setup.sh
  - Strategy: re-use/adapt existing solutions whenever possible
  - Multiple OS supported (CentOS 7, Alma 9, Ubuntu 22.04)



Fermilab

Overview of Status

Generators (Matrix Element integrators and others)

- Status of generator integration, matrix element generators, beam and machine backgrounds (e-group)
  - Usual LHC event generators, Whizard, BabaYaga, KKMCee, Guinea-pig, fluka, …
  - Beam spectra modeling (CIRCE2) handled within generator programs
  - Tight feedback loop between detectors < - > beam < - > physics analysis (fun work!)

- dd4sim can read hepmc2/3, hepevt, pairs, stdhep, et al.; so easy to integrate new generators into workflows

- Given the precision desired by both linear and circular e+e- machines there are a wealth of generator level studies to pursue
  - These studies can inform detector and MDI design considerations
  - Effects of ISR/radiative return, beam energy spread, crossing angle, machine backgrounds on final result accuracy (difficult to measure with delphes)
  - Needs to be studied and double checked well in advance of building detectors, machine
  - Multiple talks in parallels, lightning talks in this direction

Full Detector Simulations

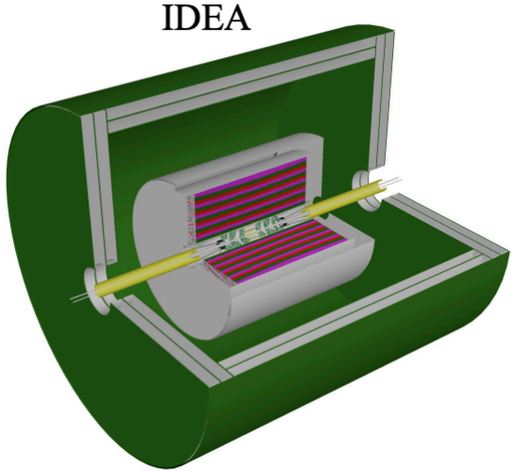
- Implementing detector geometries in k4geo is the current primary focus
  - Updated beam pipe with smaller radius recently added
  - IDEA implementation in DD4hep almost complete
    - Multiple implementations of calorimeter being integrated
  - CLD available with complete reconstruction, and detector variants
  - ALLEGRO has a complete fullsim geometry available
    - Working on fleshing out simplifications
  - There is a fourth IR but no detector envisioned for it yet. Do we want a fourth detector?
    - What should it focus on if we do want such a detector?

simulation e-group

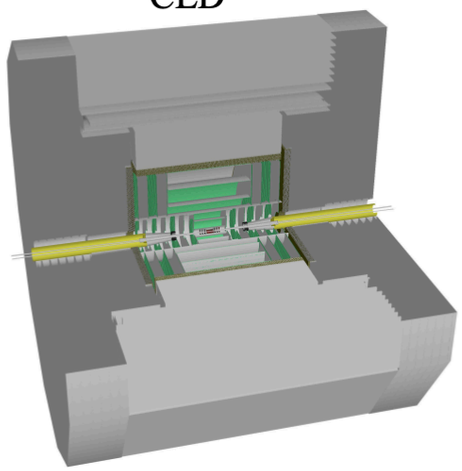
detectors e-group

- Work proceeding now towards towards digitization and reconstruction
  - Interesting technical work here sitting between software and detector implementation!

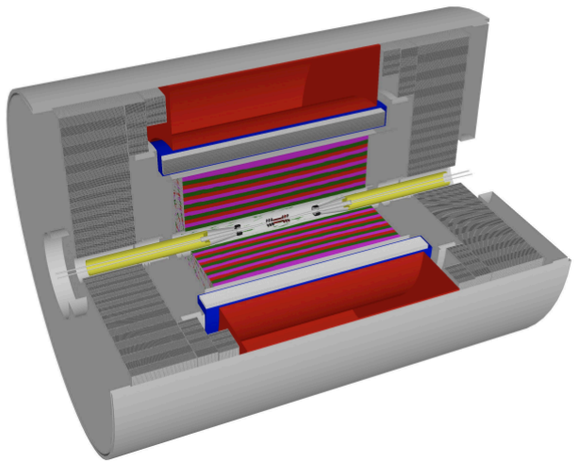
IDEA



CLD



ALLEGRO

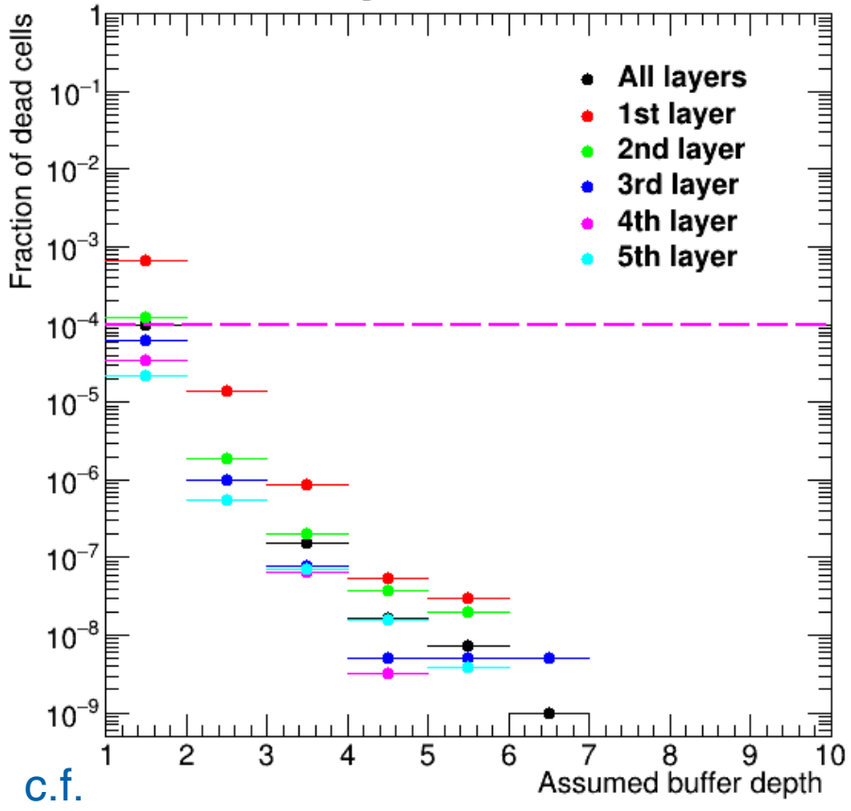


A fourth detector?

Fermilab

Background Studies and Synergies with ILC/C3

- Beam-induced backgrounds important at e+ e- machines (guinea-pig)
  - At FCC Synchrotron radiation dominates, at linear machines incoherent pairs mdi e-group
  - See talk by C. Lawson this workshop for cross-collider comparison
- Hadrons photoproduction background workflows need to be integrated with key4hep stack
  - Very low energy ($M_{aa} < 2$ GeV) production not described well by pythia, whizard
  - Custom generator toolchain (T. Barklow, et al.) needs update to Whizard3, pythia8



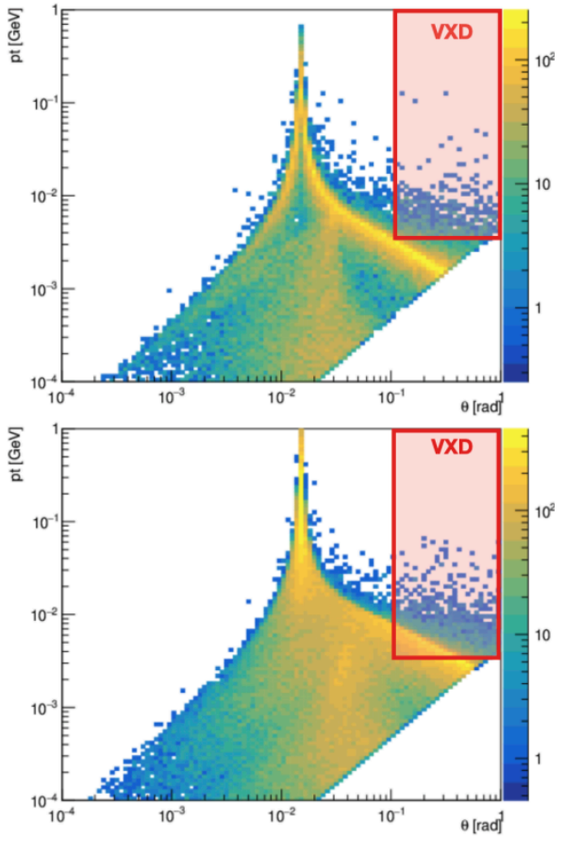
c.f.

Elias Mettner: LCWS 2023

Dimitris Ntounis: LCWS 2023

FCC pairs / Occupancy

		Z	WW	ZH	$t\bar{t}$
1	Pairs/BX	1300	1800	2700	3300
10^{-6}	O_{max}(VXDB)	70	280	410	1150
10^{-6}	O_{max}(VXDE)	23	95	140	220
10^{-6}	O_{max}(TRKB)	9	20	38	40
10^{-6}	O_{max}(TRKE)	110	150	230	290



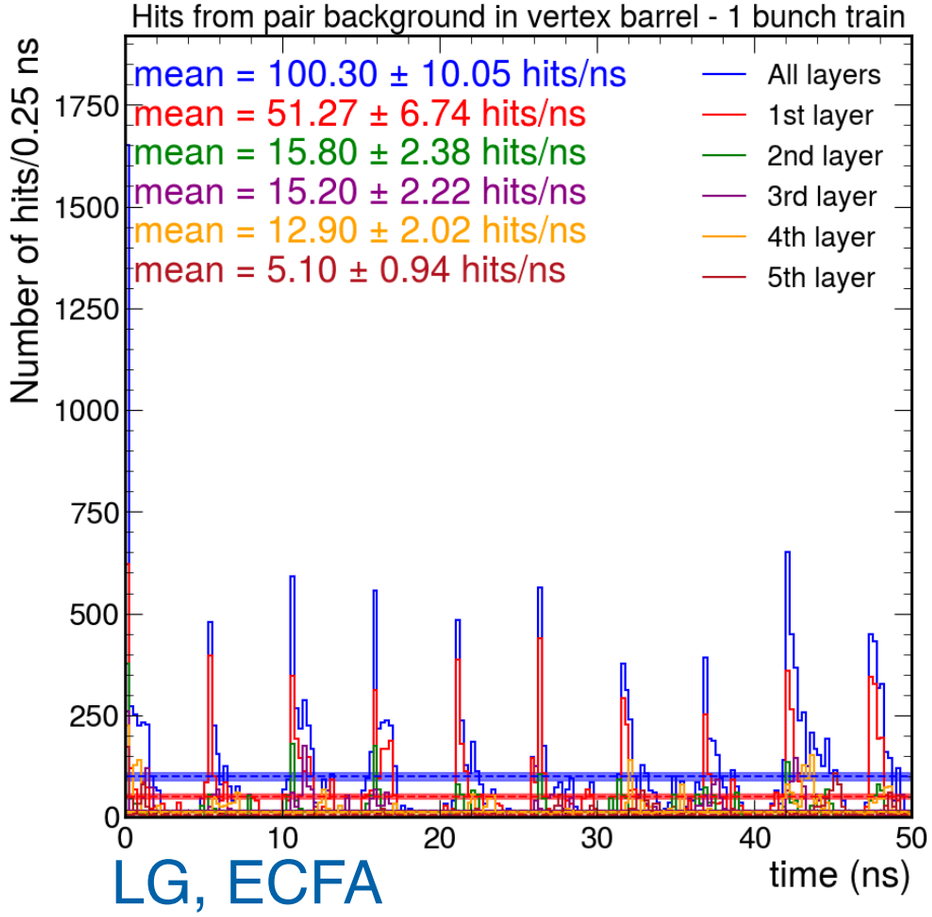
Ciarma et al.

doi : 10.18429/JACoW-eeFACT2022-TUZAT0203

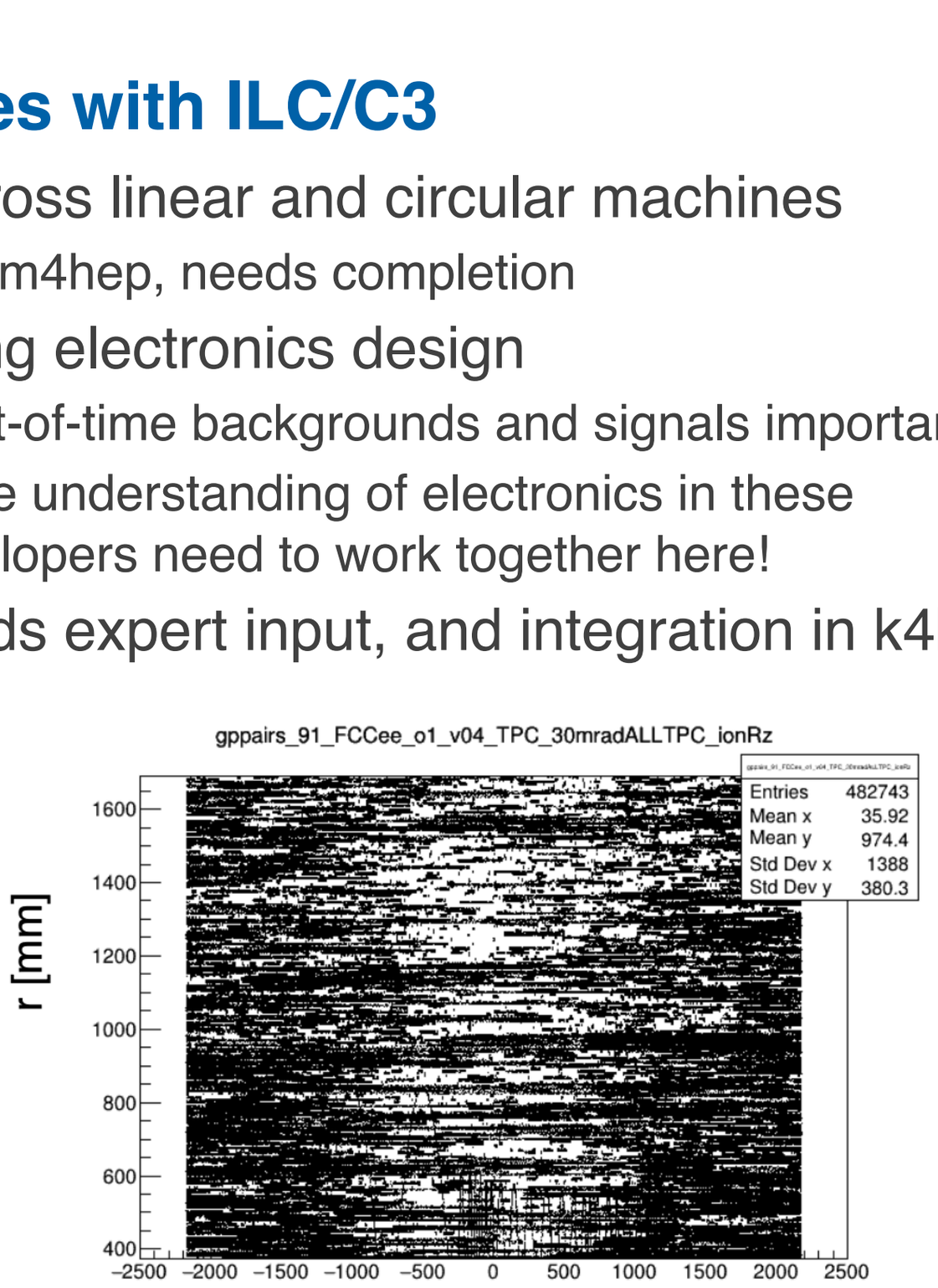
Fermilab

Background Studies and Synergies with ILC/C3

- Overlay (pileup) mixing working well across linear and circular machines
  - But it is not yet completely integrated with edm4hep, needs completion
- This is a critical capability for determining electronics design
  - On digitization work: accurate handling of out-of-time backgrounds and signals important
  - It behooves us to quickly achieve an accurate understanding of electronics in these cases, detector designers and software developers need to work together here!
- All baseline components are there, needs expert input, and integration in k4h



LG, ECFA



D. Jeans LCWS 2023

Fermilab

Analysis Software

- FCCAnalyses: maintained by FCCSW team based on RDataFrame
  - Handles bookkeeping across multiple axes, maintains suite of high level variable definitions
  - Plotting / weights / systematics handling provided
  - All inputs are edm4hep by default

analysis e-group

- Centrally produced Delphes samples are provided for all to use
  - /eos/experiment/fcc/ee/generation/DelphesEvents/winter2023/IDEA/
  - CLD, ALLEGRO, IDEA fullsim in or planned for production as designs completed

- Ongoing work:
  - Documentation
  - Better integration of edm4hep/podio functionality via RDataSource
  - Thorough logging for easier debugging and understanding
  - Multiple entry points depending on preferred software environment stack
  - Scientific python ecosystem (awkward array, and family), but not often requested

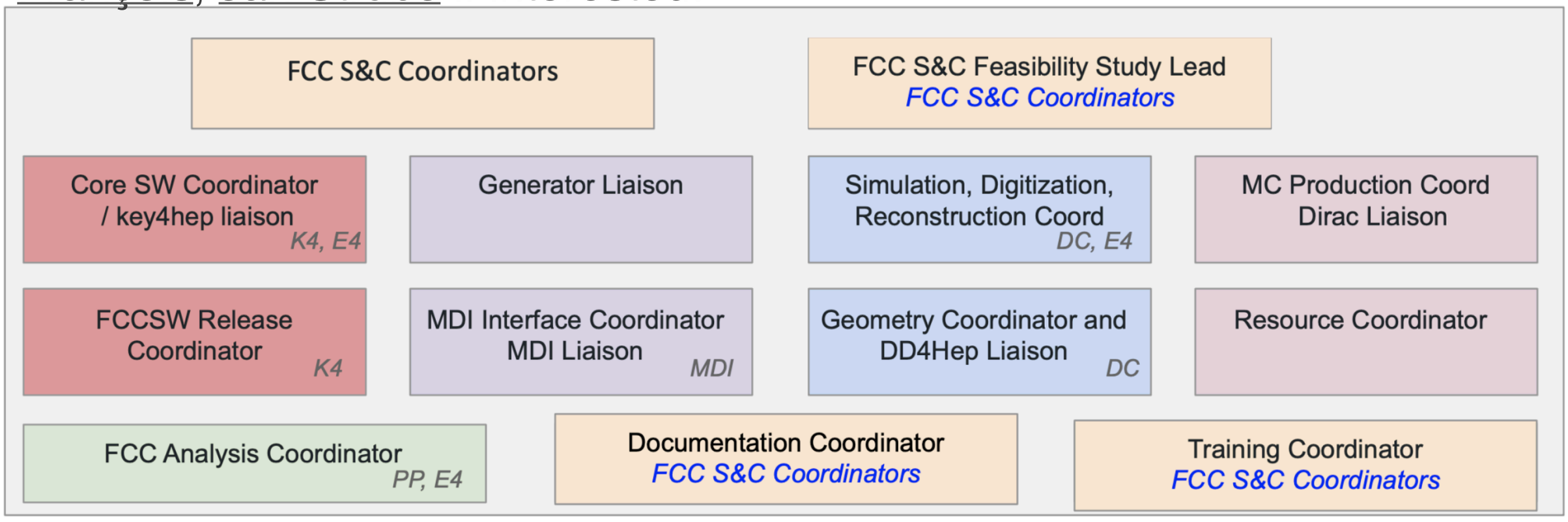
Needs for the Near-term Future

Computing Resources

- Proceeding towards physics sample production using grid resources
  - ILCDirac - current sites: CERN, Bari, CNAF (CPU and storage), Glasgow (storage only)
- Dedicated batch resource queues in CERN batch (group_u_FCC)
- Current CERN storage
  - 500 TB (~100 TB free) for central production
  - Supported by 200 TB for analysis
  - Groupwise allocations of 10-20 TB per group depending on need (expandable)
- Full Sim sample are in-flight: CLD most mature and ready for production
  - ALLEGRO and IDEA will be added as they are solidified in Full Sim
    - Will need to reassess disk space allocations to store detector variants
  - First productions will be minimal to address basic validation and initial analysis needs
    - It does not yet make sense to proceed to tera-Z scale productions
  - Need community input to determine a prioritized list of desired Full Sim samples
  - With this list we can determine storage and compute resource needs, plan for expansion
- This requires a tight feedback loop with the rest of the PED
  - Can possibly benefit from experience of prior large productions for ILC to predict needs better (via experience with dirac), but scale more similar to (HL-)LHC though

Present Personpower and Areas of Need

- FCC Software CERN Team: 1.5 FTE staff scientist + 2 FTE fellows total effort
- Additional contributions on a best-effort basis so far
  - This critical stage of detector work will require lots of expert contribution!
- Working to organize thrusts of effort: Contact Gerardo Ganis, Brieuc François, Jan Strube if interested!



- 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

Organizing Around the Feasibility Study Report

- All of this work focused towards the Feasibility Study Report (FSR)
  - There is ~ 1 year left to justify the FCC project as much as possible
  - Given tight personpower and wealth of creative ideas, we will have to prioritize aggressively
  - PED is working to achieve a prioritized list of software deliverables
    - We must deliver complete results at the expense of being exhaustive
      - Completely solid results are much easier to interpret than an array of nearly-complete results
    - It is very likely that not all detector concepts will be showcased for high level physics results
    - However, all detectors will have detailed performance studies included in the report

- Personpower is *critical* in this next year, the US could play a substantial role
  - Need help in coordination, sample production, core software and detector specifics
  - Technical developments especially necessary to bring Full Sim detectors to “life”
  - Physics validation and integration of test beam results necessary for solid justifications
  - Detector performance and then physics studies (and feedback cycles therein)

- There are few users giving feedback, this is critical to getting solid results
  - Detectors, simulation, and software is an area the US has a lot of expertise in and we can really help!

Conclusions

- Many open topics to pursue within FCC/ILC software and computing
  - FCC detectors need to be developed into complete FullSim implementations for the FSR
  - There are a variety of immediately interesting tasks:
    - Particle Flow studies for CLD under various detector configurations
    - Background overlay integration
    - Realistic detector digitization and electronics simulations
    - The FullSim physics analyses themselves!
  - Contact Brieuc François, Gerardo Ganis, Jan Strube if you would like to get involved
- The deep connection of key4hep to ILC software and the TDRs means that work can often have many resulting use cases
  - Physics studies can be used to understand multiple colliders, k4h makes the change from one to another more straightforward than ever before
  - Technical contributions have wide reaching impacts
  - Detector specialists can see the impact of a variety of experimental environments on their systems
- Please do not hesitate to get involved!
  - There's a lot of exciting work in the next 1 year and the years after that