AIDA2020 and AIDA++

Graeme A Stewart

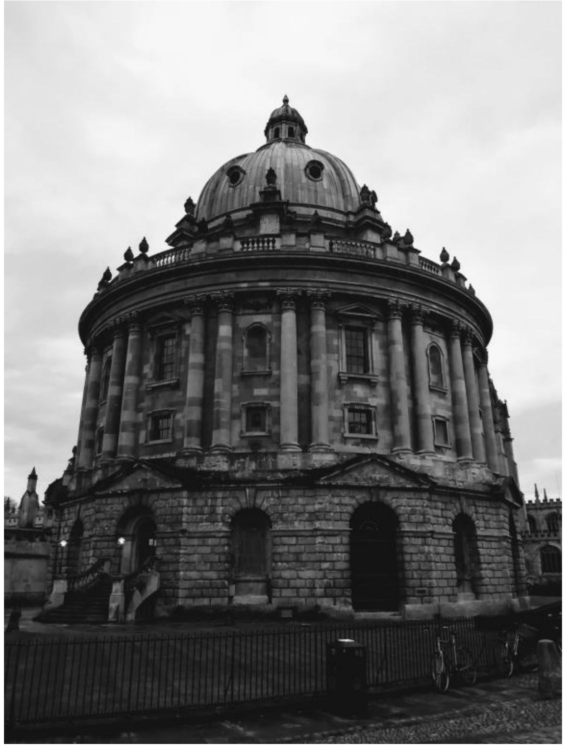
Based on the WP3 summary sides from Frank Gaede and Witek Pokorski

SFT Group Meeting, 2019-04-29

AIDA2020

- European Union Horizon 2020 Project
- **A**dvanced European **I**nfrastructure for **D**etectors at **A**ccelerators
- Four year project with €10M of EU funding
  - 2:1 ratio of beneficiary commitments (i.e. €20M from institutes)
- Now in a 1 year "no cost" extension period
- Many work packages on sensors (Si, Gas, Cryo Liquids), detector mechanics and cooling, data acquisition, test beams and irradiation facilities
- Also a work package on *Advanced Software* (WP3)
  - Coordinated by Witek and Frank

- 2019 Annual meeting held in Oxford at St Anne's College, 2-5 April
  - Gabriele, Graeme and Witek present for EP-SFT



WP3 Overview

- Advanced simulation and reconstruction for HEP
  - To address current challenges of multithreading, parallelization and vectorization

- Core software
  - Geometry
    - DD4hep and USolids extensions
  - Alignment and conditions data
  - Event Data Model and framework extensions
- Simulation
  - DDG4: Geant4 based simulation toolkit
- Reconstruction
  - Advanced tracking tools
  - Advanced particle flow algorithms

General Philosophy

- Develop software that is needed for projects: LHC upgrade, ILC and CLIC, FCC or neutrino programme
- Focus on generality and re-usability already in design of the software
  - Tools should address the current needs and at the same time be generic enough to be applicable for other/all HEP experiments
- Seek collaboration with HEP Software Foundation (HSF)
  - Some packages are already official HSF projects

- Partners:
  - CERN, DESY, LAL (Paris Sud, IN2P3), LLR (École Polytechnique), University of Manchester, University of Cambridge

WP3 Overview

Objectives

Task 3.1 Scientific coordination

- Coordinate and schedule the execution of the WP tasks
- Monitor the work progress (milestone and deliverable reports), follow-up on the WP budget and the use of resources
- Organise WP meetings

Task 3.2 Detector Description for HEP (DD4hep) and Unified Solids (USolids) extensions

- Extend USolids for vectorisation using Single Instruction, Multiple Data (SIMD) instructions and reviewed algorithms
- Define proper interfaces for use of USolids in Geant4, Root and Vector prototype
- Implement thread safety and alignment procedures in DD4hep

Task 3.3 Alignment and conditions data (test beam)

- Complete alignment toolkit with tight coupling to DD4hep for simulating the misalignment
- Provide alignment and conditions data for DD4hep

Task 3.4 Event Data Model (EDM) toolkit and framework extensions

- EDM toolkit for efficient creation of Event Data Models in C++ with high performance I/O
- Implementation of parallel algorithm scheduling mechanisms in HEP frameworks

Task 3.5 DDG4 (Detector Description Geant 4): Geant4 based simulation toolkit

- Modular and flexible simulation toolkit based on DD4hep and Geant4
- Application to LC and FCC

Task 3.6 Advanced Tracking Tools

- Development of advanced parallel algorithms for track finding and fitting in AIDA Tracking Tool toolkit (aidaTT)
- Application to LHC and LC

Task 3.7 Advanced particle flow algorithms

- Development of advanced particle flow and pattern recognition algorithms in PandoraPFA (particle flow algorithms toolkit)
- Application to LHC, LC and neutrino experiments

- List of Tasks and Deliverables

4nd Annual Meeting WP3 session

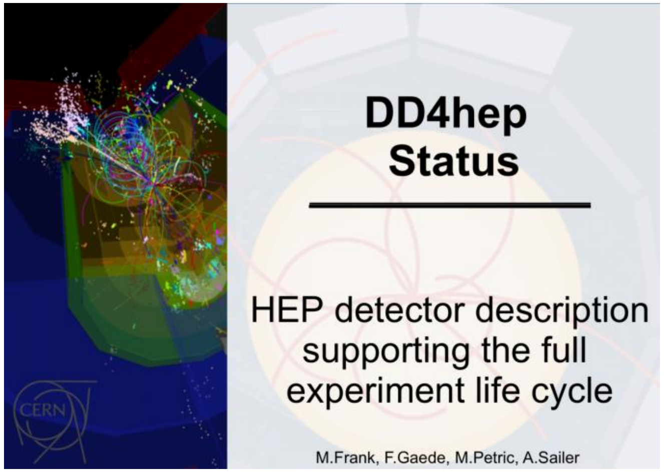
Introduction	*Frank-Dieter Gaede et al.*	
Seminar Room 10, St Anne's College	09:00 - 09:10	
DD4hep	*Markus Frank*	
Seminar Room 10, St Anne's College	09:10 - 09:25	
DDG4	*Andre Sailer*	
Seminar Room 10, St Anne's College	09:25 - 09:40	
VecGeom	*Dr Gabriele Cosmo*	
Seminar Room 10, St Anne's College	09:40 - 09:55	
Alignement and Conditions Data	*Biljana Mitreska*	
Seminar Room 10, St Anne's College	09:55 - 10:10	
Update on PODIO	*Frank-Dieter Gaede*	
Seminar Room 10, St Anne's College	10:10 - 10:25	

Framework extensions: LAL	*Hadrien Benjamin Grasland*
Seminar Room 10, St Anne's College	11:00 - 11:15
Advanced Tracking Tools	*Hadrien Benjamin Grasland*
Seminar Room 10, St Anne's College	11:15 - 11:30
Towards a parallel Marlin framework	*Remi Ete*
Seminar Room 10, St Anne's College	11:30 - 11:45
Advanced PFA Tools	*Steven Green*
Seminar Room 10, St Anne's College	11:45 - 12:00
Discussion: plans for AIDA++	
Seminar Room 10, St Anne's College	12:00 - 12:30

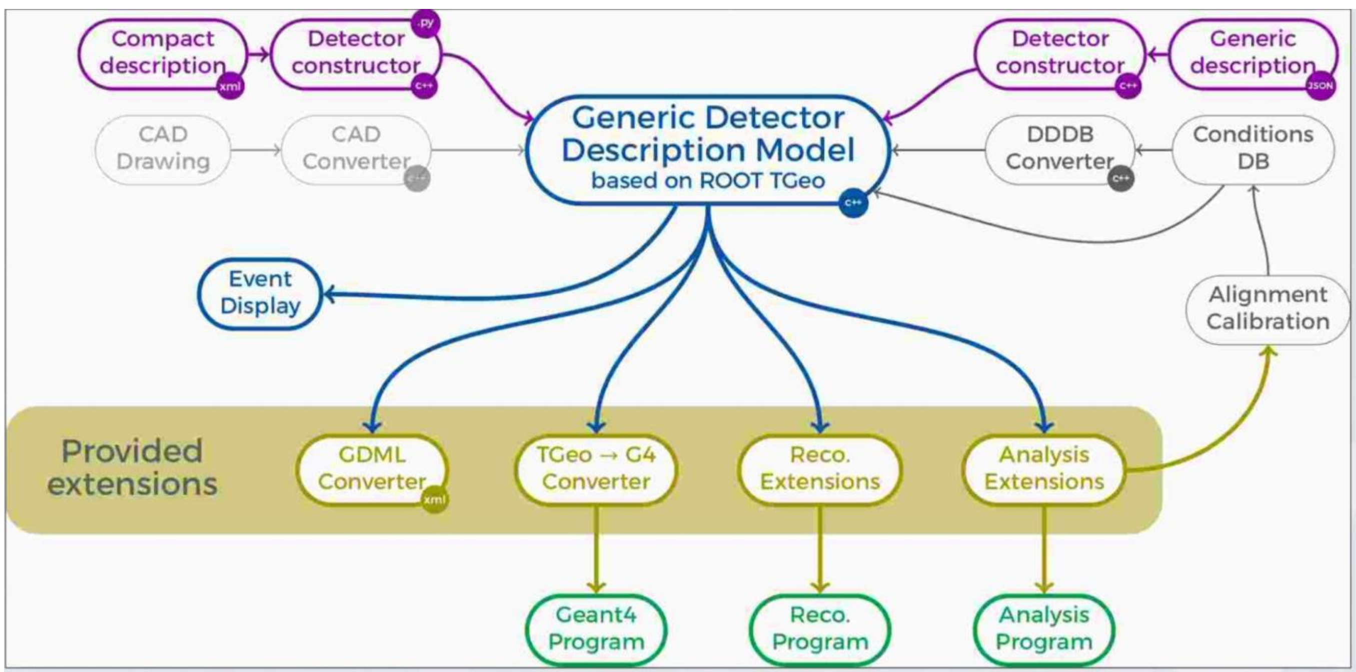
DD4hep - goals

Task 3.2

M.Frank



- **Develop a detector description**
  - **For the full experiment life cycle**
    - **detector concept development, optimization**
    - **detector construction and operation**
    - **"Anticipate the unforeseen"**
  - **Consistent description, with single source, which supports**
    - **simulation, reconstruction, analysis**
  - **Full description, including**
    - **Geometry, readout, alignment, calibration etc.**



NEW: Support for Optical Surfaces

- Requests from LHCb, FCC and SCTF
  - Placeholder concept developed by TGeo
  - Integration and validation in DD4hep ongoing
- Changes to
  - Surface objects => newly created in D4hep
  - Materials => added material properties
  - DDG4 => propagate surfaces and material properties to Geant4
- Iterative development collaboration between DD4hep and ROOT TGeo
- Ongoing development

Toolkit Users

Increasing interest in the HEP community

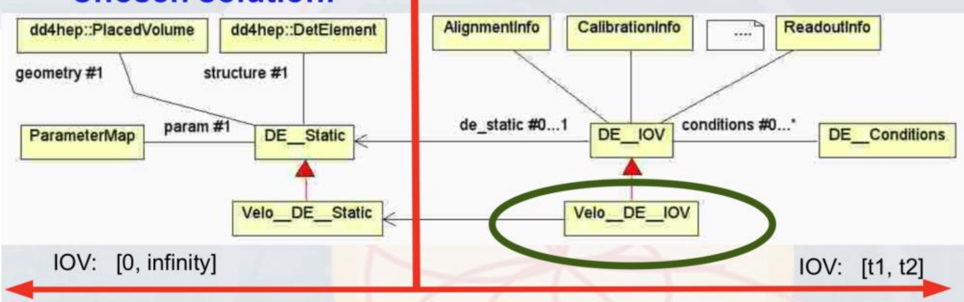
- ILC — F. Gaede et al.
- CLICdp — A. Sailer et al.
- SiD — D. Protopopescu et al.
- FCC-eh — P. Kostka et al.
- FCC-hh — A. Salzburger et al.
- FCC-ee — O. Viazlo (CLD design), N. Alipour, G. Voutsinas
- SCTF — Super-Tau-Charm Factory designs (Novosibirsk, Bejing)
- EIC — Evaluation considered/started (W. Armstrong et al.)

- LHCb — LHCb Upgrade for Run III (B.Couturier et al.)
- CMS — Evaluation for upgrade started (202x) (Y.Osborne et al.)
- CALICE — Calorimeter R&D, started

Conditions data

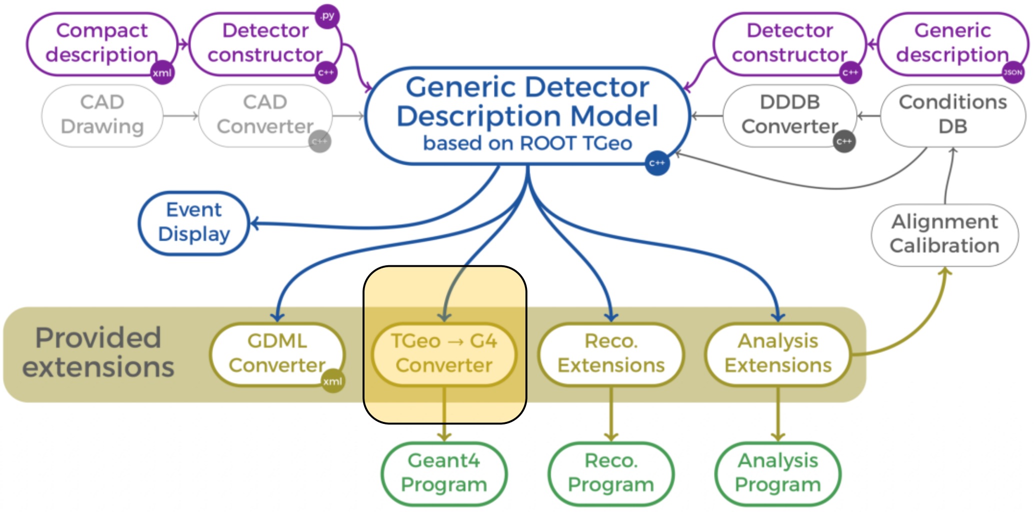
- New approach to provide aggregated conditions for each object (detector element)
  - handles both static data and time dependent data
  - supports multi-threaded processing without locking

- Chosen solution:



- IOV dependent projection is our new "detector element"
  - Facade provides functionality
- *Backwards* and *forward* compatible
  - Any number of facades possible on the same data

DD4hep/DDG4 - goals



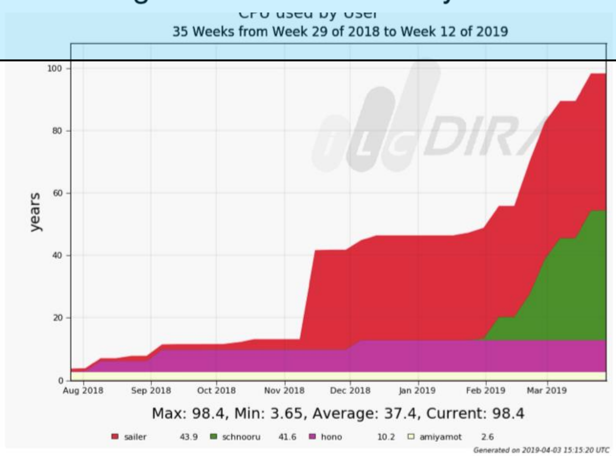
- In-memory translation of geometry from TGeo to GEANT4
  - Materials, Solids, Limit sets, Regions
  - Logical volumes, placed volumes and physical volumes
- External configuration:
  - Plug-in mechanism
  - Property mechanism to configure plug-in instances
  - Supports configuration via XML, Python or ROOT-AClick
- Use plug-in mechanism to configure: Generation, Event Action, Tracking Action, Stepping Action, SensitiveDetector, PhysicsList. . .
- Provides MC truth handling with record reduction
- DDG4 is highly modular

DD4hep/DDG4 - new development

- Optical Surfaces
- DDSIM python program ported from LCGEO to DD4HEP
- Possibility to simulate all events in one file
  - Useful when number of events varies from file to file
- Bug fixes: input readers
- Keeping up with GEANT4

- DDSIM python executable is now part of DD4hep
- Get steering file `ddsim --dumpSteeringFile > mySteer.py`
  - Steering file includes documentation for parameters and examples
  - The python file contains a `DD4hepSimulation` object at global scope
  - Configure simulation directly from command line

- DDG4 based simulation used by ILD, CLIC, FCCee in production
- CLICdp published note on Performance of CLICdet in December
  - Big activity in November to produce final samples for this note
  - Physics analysis productions starting
- FCC: Input to CDR with CLD detector
- ILD: Working on ILD Design Report

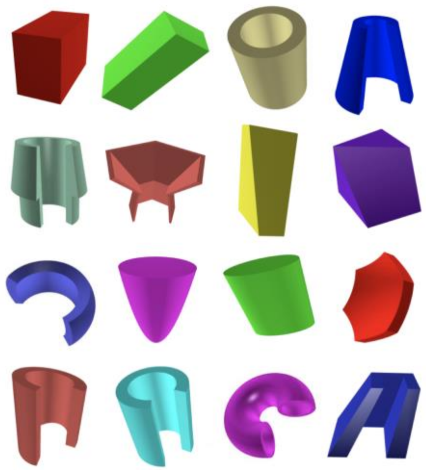


AIDA2020

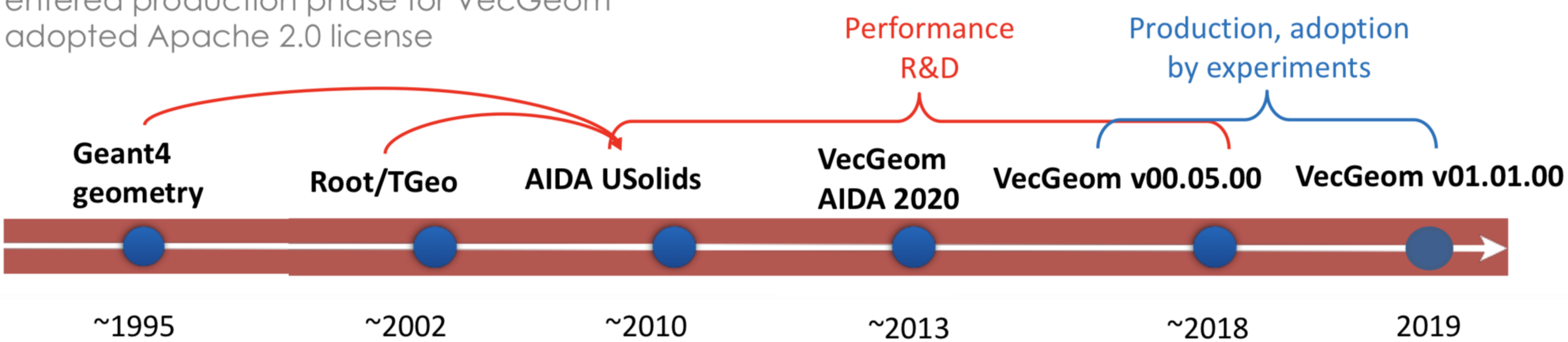
Usolids/VecGeom - goals

Task 3.2

G.Cosmo



- AIDA project aiming to unify Geant4 and Root geometry algorithms
  - merge code base
  - pick best implementation and increase code quality
  - improve performance and increase long term maintainability
- Extended scope in **VecGeom**
  - encompass parallelism/vectorization
  - multi-architecture/multi-platform support
  - provide advanced navigation features
- Old initial USolids implementation phased out in 2018
  - entered production phase for VecGeom
  - adopted Apache 2.0 license

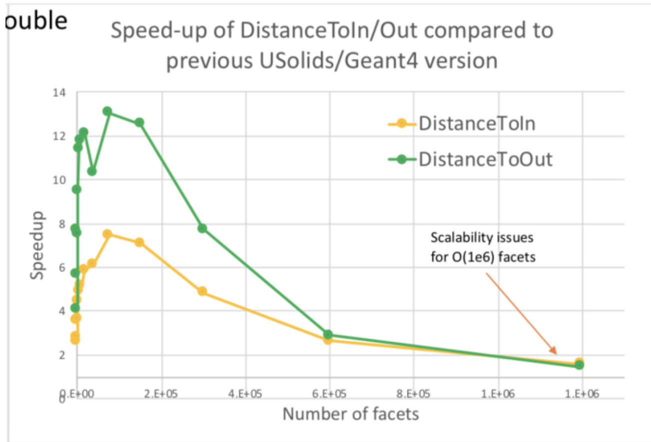


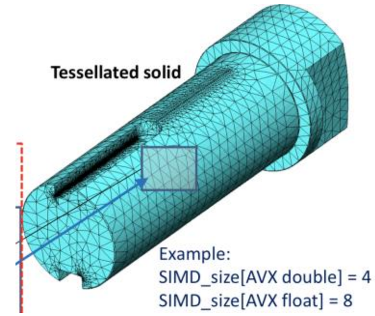
USolids/VecGeom – new development

Task 3.2

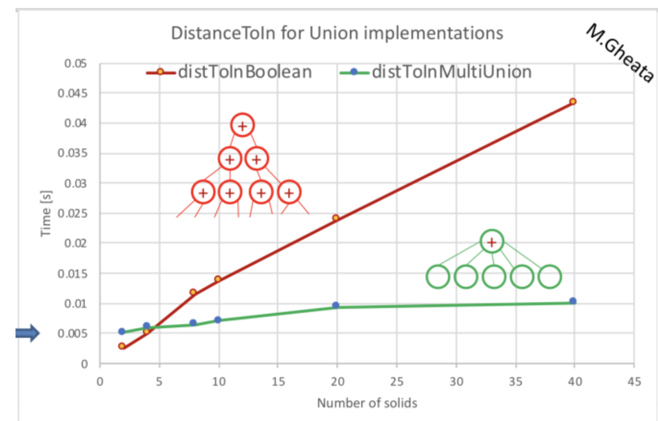
G.Cosmo

- Improved tessellated solids performance by exploiting SIMD instructions
  - O(10) speed-up

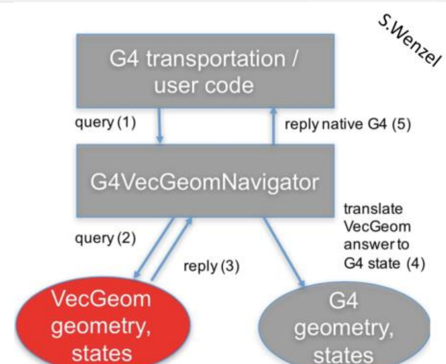




- Mutli-unions re-implemented in VecGeom using technique based on Bounding Volume Hierarchies (BVH)
  - 2x – 4x speed up
  - VecGeom implements automatic conversion of classic Boolean union of volumes to the new multi- union structure



- Interfacing VecGeom navigator to Geant4
  - will be done in two steps
    1. Simultaneous existence of Geant4 and VecGeom geometry with necessary synchronization/translation of states/objects
    2. full integration with no intermediate layer



The alignment

- Best determination of the position of detector elements → best performance of the system
- Track based alignment provides the highest precision of the detector position

The BACH alignment toolkit

- Tools for performing track-based alignment
- Developed during the original AIDA project and extended in AIDA 2020
- Extended to use the DD4hep toolkit

Meeting milestones

- PCMAG hardware project is delayed
- Milestone 89 met through application of BACH alignment package to LHCb beam Telescope


AIDA-2020-MS89
AIDA-2020
Advanced European Infrastructures for Detectors at Accelerators
Milestone Report

Application of alignment toolkit to an external tracker

C. Burr (UNIMAN) *et al*

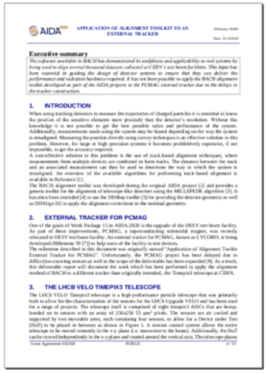
21 December 2018

The AIDA-2020 Advanced European Infrastructures for Detectors at Accelerators project has received funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement no. 654168.

This work is part of AIDA-2020 Work Package 3: **Advanced software**.

The electronic version of this AIDA-2020 Publication is available via the AIDA-2020 web site <http://aida2020.web.cern.ch> or on the CERN Document Server at the following URL: <http://cds.cern.ch/search?p=AIDA-2020-MS89>

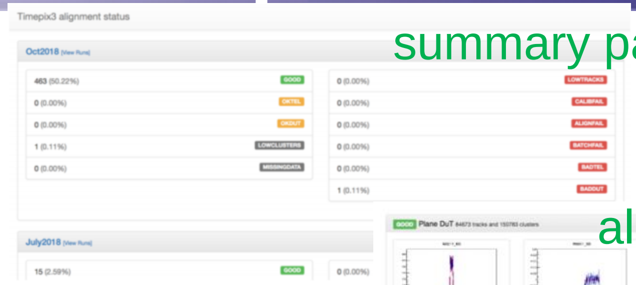
Copyright © CERN for the benefit of the AIDA-2020 Consortium



Alignment – new development

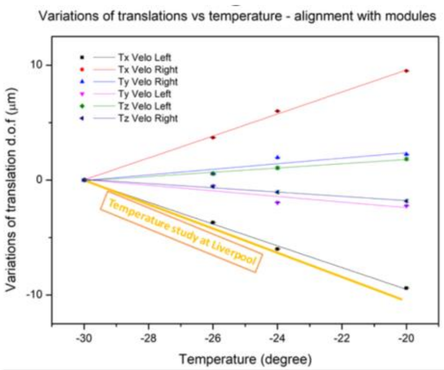
- web application created using Python's flask framework to help manage the large number of runs and check the alignment status

summary page

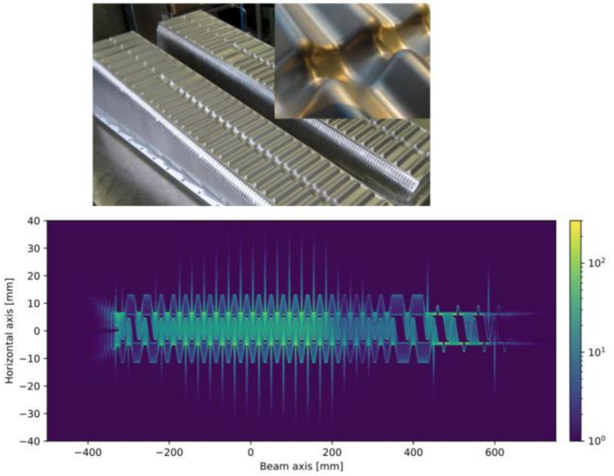


alignment quality

- alignment evaluation was used to determine the variation of the VELO half position as function of the temperature



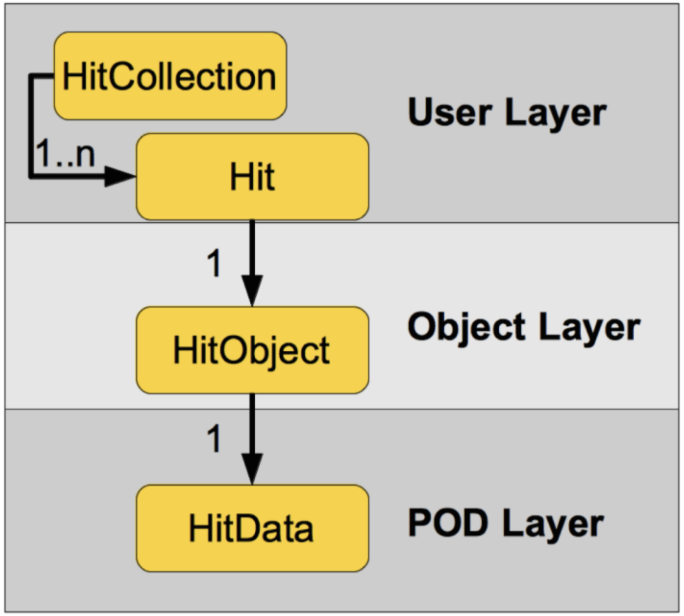
- VELO alignment using material interaction studied on a simulation sample
    - Create a 2D map of the VELO using secondary vertices of interaction with the VELO material



EDM Toolkit - goals

- new EDM toolkit developed in AIDA2020 in the context of FCC allowing re-use by different HEP groups
- solving the problems of overly complex, deep object hierachies with many virtual function calls and non-optimal I/O performance
- exploiting C++ objects <-> Plain Old Data structures duality
- implementation
  - the three PODIO layers

- user layer (API):
  - handles to EDM objects (e.g. **Hit**)
  - collection of EDM object handles (e.g. **HitCollection**).
- object layer
  - transient objects (e.g. **HitObject**) handling vector members and *references* to other objects
- POD layer
  - the actual POD data structures holding the persistent information (e.g. **HitData**)



EDM Toolkit – new development

LCIO EDM in PODIO: pLCIO

- implemented almost **complete LCIO EDM** in PODIO
  - dedicated package **pLCIO**: https://stash.desy.de/projects/IL/repos/plcio

benchmarking of LCIO vs pLCIO ongoing

Ongoing work:

- implementation of direct binary I/O making use of array-of-POD
- implementation of HDF5 I/O layer

And also:

- moved Github repository to: https://github.com/aidasoft/podio
- added some standard templates from iLCSoft for release notes, issues, etc.
- addressed a few issues needed for MS90
  - add `CollectionBase::size()` method
  - implement **vector member** streaming
- some minor bugs and issues fixed

Framework Extensions - goals

- Task 3.4: Extend frameworks for **parallel event processing**
  - Initially aimed for common Gaudi/Marlin/Pandora solution
  - …but Gaudi went its own way before AIDA-2020 started
- LAL decided to refocus on **parallel condition handling**
  - No condition handling in Gaudi itself → effort duplication
  - Singleton storage must be re-thought for parallelism
  - An ATLAS-specific take on this exists, not generalizable
  - Proposal: Provide generic components at Gaudi level

Framework Extensions - status

- Enumerated known use cases for condition data
- Devised a generally applicable scheme
- Proved viability and performance via prototyping
- Investigated shortcomings of Gaudi infrastructure

However…

- It is very hard to get any meaningful change into Gaudi
  - Wildly diverging visions across experiments
  - …so experiment-specific code almost always preferred

So…

- Solution: Take this work in experiment-specific direction
  - Re-use API design work from this project
  - Drop dependency on Gaudi infrastructure improvements
  - Implement everything as experiment-specific components
  - Support legacy LHCb conditions, DDCond being worked on

Deliverable 3.5, Milestone 91 Postponed

Advanced tracking tools - goals

- Task 3.6: Common developments on track finding & fitting
  - Initially envisioned in the context of aidaTT toolkit
  - …but since then, ACTS was announced and released

A Common Tracking Software (ACTS) released as open source software

- ACTS is a deep rework/generalization of ATLAS Run2 code
  - Turns battle-tested code into an experiment-agnostic library
  - Enables new use cases like efficient multithreading
  - Used in FCC studies, TrackML, interest from LC community
- Consensus was that contributing to ACTS is best

Advanced tracking tools – news (ACTS)

Task 3.6

F.Gaede, H. Grasland

- Finished:
  - Implemented and integrated multi-threaded validation
  - Improved API const-correctness for easier multithreading
- Ongoing:
  - Optimizing key components (Runge-Kutta, Kalman Filter...)
  - Numerical analysis study of ACTS floating-point code ← presented at CHEP2018
  - Analyzing Belle 2 perf. in view of piecewise integration
- Optimized interpolated magnetic field map
  - Being reviewed, will provide 10x faster random field lookups
- Spack packages for ACTS are now available
  - Installing dependencies & building has never been easier

DESY ACTS activity:

- started to look into ACTS and adopting it to the LC tracking code by implementing the MarlinTrk interface, that has been developed in the first AIDA project and that is also implemented by aidaTT

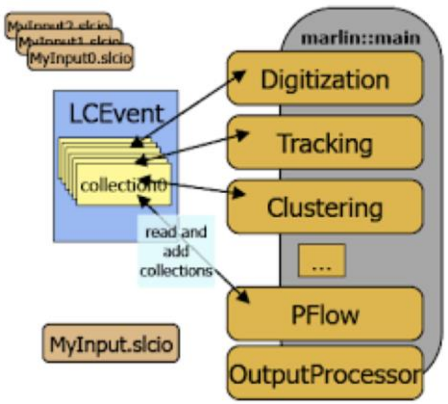
- it is at this moment unclear what will be possible with ACTS in the last year of AIDA2020
- in any case we will follow the development closely and **start to contribute to ACTS**

Deliverable 3.7 Postponed

The Marlin framework

A **M**odular **A**nalysis and **R**econstruction for **LIN**ear collider

- An event processing framework for HEP
- Event data model: LCIO
- Main usage is reconstruction
  - Post simulation framework
  - Pre analysis framework



Motivation for a multi-threaded version

- Limitation to multi-process parallelism
  - **Duplication of process memory** (detector geometry, user data, etc ...)
    → LHC showed strong limitation there: CMS tot mem (2018) $\sim$ 8Gb
  - **Duplication of IO calls** (fread/fwrite)
    → Concurrent access to cache by different processes
    → Reduce data caching: speedup application
- Clumsy map-reduce strategy at process level
  - Splitting of hundreds of files
  - Requires understanding of file splitting
  - Re-combination can be tricky (event numbering)
  - **Introduces additional data managment**
- **Duplicated application startup time for each process**
  - Geometry loading
  - Background map loading
  - Potential database access

Framework Extensions MarlinMT - news

Task 3.4

R. Ete

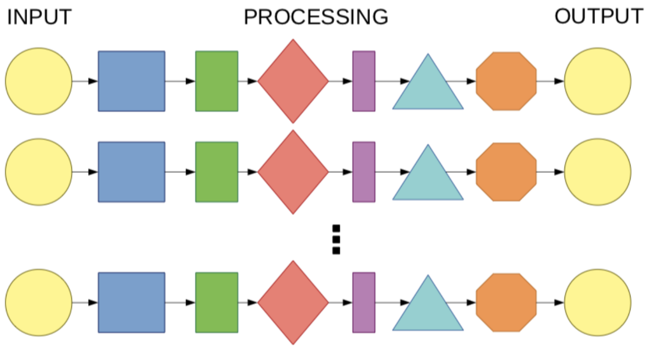
Current status

First compiling and running version available on Github !

- Implemented a simple number crunching test processor
- Running 9 Marlin processors on N threads with N files
  → scales perfectly !
- Very simple test setup but encouraging start for the project

Currently investigating:

- Histograming component:
  - ROOT `TH1::Fill` with locks in an AIDA wrapper ?
- Data conditions in MT environment:
  - Possible MT solutions: LCCD, DD4hep, others ?
- LCIO output writing
  - Currently: LCIOOutputProcessor (not MT suitable)
  - Output event pool + defered writing ? Split & merge strategy ?
- ILD reco chain issues in MT environment



Conclusion

- Parallel version of Marlin investigated
- First parallelization: MT over input LCIO files
- Investigated thread safety issues in dependencies:
  - ✔ LCIO: SIO layer, interface
  - ✔ streamlog: thread safe logging available
  - ✔ Geometry: geometry loading (DD4hep/Gear)
- Implemented a first working version of MarlinMT
  - ✔ Tested with simple processor

Outlook

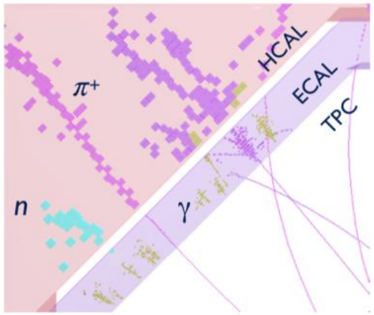
- Missing ingredients to be integrated
  - ✗ LCIO output data management
  - ✗ Histograming component
  - ✗ Conditions data handling
- Will investigate ILD reco chain bottlenecks

Advanced PFA - goals

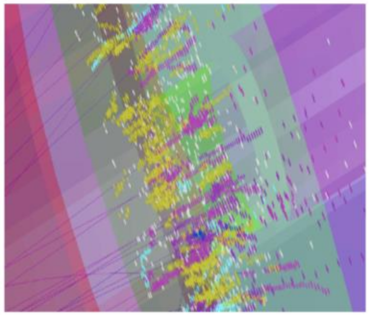
Task 3.7

S.Green

- Development of advanced particle flow and pattern recognition algorithms in PandoraPFA
- Application to LHC, LC and neutrino experiments

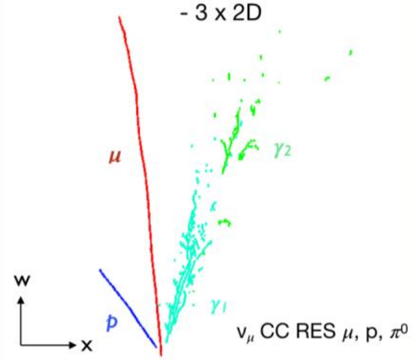


Typical ILC event topologies - 3D
NIMA.2009.09.009 NIMA.2012.10.038



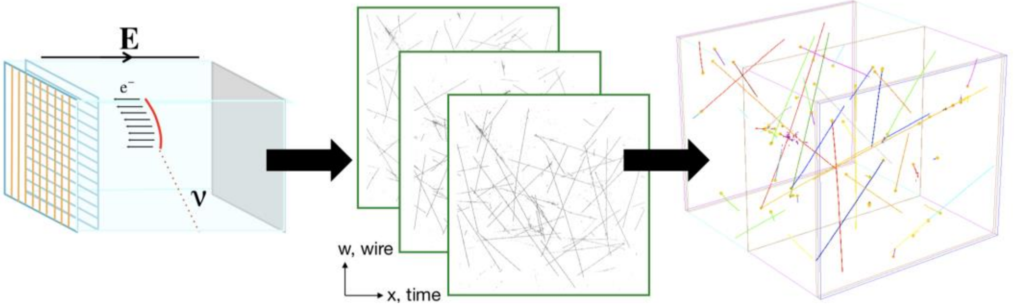
Typical showers in CMS HGCAL - 3D
LHCC-P-008

Typical LArTPC event topology - 3 x 2D

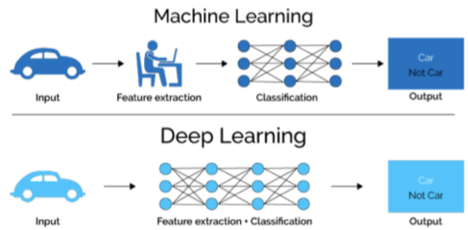


Example: Liquid Argon Time Projection Chambers (LAeTPCs):

- Aim of Particle Flow Algorithms:
  - Reconstruct the particle hierarchy produced from an interaction.
  - Produced 3D reconstructed particles from 3 x 2D images.



AIDA 2020 — Advanced PFA - news



- Embrace deep-learning approaches, which should be appropriate for LArTPC image analysis:
  - Combine power of multi-algorithm and deep-learning approaches.
  - Use deep learning to guide key steps in the Pandora algorithms.

- Full exploitation of Pandora for MicroBooNE, SBND, ProtoDUNE and DUNE physics programmes:
  - Assess performance of existing algs at ProtoDUNE and DUNE, including detailed studies of interaction hierarchies.
  - Use the data from ProtoDUNE to develop Pandora to provide the best possible reconstruction for DUNE.

- large amount of work has gone into a recent note describing the CLIC detector performance using PandoraPFA
- wide range of studies: Particle ID Efficiencies, Performance of jet reconstruction, Performance of missing transverse energy reconstruction,Di-Jet mass separation

CLICdp-Note-2018-005
17 December 2018

A detector for CLIC: main parameters and performance

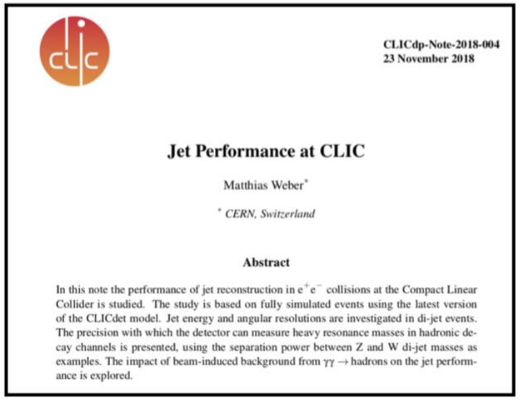
Dominik Arominski[a,b], Jean-Jacques Blaising[c], Erica Brondolin[a], Dominik Dannheim[a], Konrad Elsener[a], Frank Gaede[d], Ignacio García García[a,e], Steven Green[f], Daniel Hynds[a,1], Emilia Leogrande[a,*], Lucie Linssen[a], John Marshall[g,2], Nikiforos Nikiforou[a,3], Andreas Nürnberg[a,4], Estel Perez-Codina[a], Marko Petrič[a], Florian Pitters[a,h], Aidan Robson[i,5], Philipp Roloff[a], André Sailer[a,*], Ulrike Schnoor[a], Frank Simon[j], Rosa Simoniello[a,6], Simon Spannagel[a], Rickard Stroem[a], Oleksandr Viazlo[a], Matthias Weber[a,*], Boruo Xu[f]

On behalf of the CLICdp Collaboration

[a] CERN, Geneva, Switzerland, [b] Warsaw University of Technology, Warsaw, Poland, [c] Laboratoire d'Annecy-le-Vieux de Physique des Particules, Annecy-le-Vieux, France, [d] DESY, Hamburg, Germany, [e] IFIC, Universitat de Valencia/CSIC, Valencia, Spain, [f] University of Cambridge, Cambridge, United Kingdom, [g] University of Warwick, Coventry, United Kingdom, [h] Technische Universität Wien, Vienna, Austria, [i] University of Glasgow, Glasgow, United Kingdom, [j] Max-Planck-Institut für Physik, Munich, Germany

Abstract

Together with the recent CLIC detector model CLICdet a new software suite was introduced for the simulation and reconstruction of events in this detector. This note gives a brief introduction to CLICdet and describes the CLIC experimental conditions at 380 GeV and 3 TeV, including beam-induced backgrounds. The simulation and reconstruction tools are introduced, and the physics performance obtained is described in terms of single particles, particles in jets, jet energy resolution and flavour tagging. The performance of the very forward electromagnetic calorimeters is also discussed.



CLICdp-Note-2018-004
23 November 2018

Jet Performance at CLIC

Matthias Weber[*]

[*] CERN, Switzerland

Abstract

In this note the performance of jet reconstruction in e^+e^- collisions at the Compact Linear Collider is studied. The study is based on fully simulated events using the latest version of the CLICdet model. Jet energy and angular resolutions are investigated in di-jet events. The precision with which the detector can measure heavy resonance masses in hadronic decay channels is presented, using the separation power between Z and W di-jet masses as examples. The impact of beam-induced background from $\gamma\gamma \to$ hadrons on the jet performance is explored.

Summary of WP3

- AIDA-2020 Advance Software WP3 addressing core, simulation and reconstruction software for HEP
  - SFT group was involved in conceptualization, management and development
- All of the WP3 tasks are in a good state
  - Good progress on the software development - good communication between partners (regular phone meetings)
- All milestones and deliverables achieved on time
  - Two deliverables moved to year 5 and no problem foreseen to achieve them
    - In both cases external factors were a significant issue
- Year 5 can be useful
  - Polish code, improve documentation
  - Maintain the development community that has grown up

- Graeme's SAB observations that most successful projects were ones that had *excellent development* work in areas that were *focused* and offered a *clear migration path*
  - *Sometimes this took 2 rounds of AIDA support!*

AIDA++ (Tentative Name)

- AIDA2020 management team have been in discussions with Brussels about follow-up projects
- Target FP8 Call 5 “to foster the innovation potential of research infrastructures”, INFRAINNOV-04-2020
  - Innovation in light source technologies
  - **Innovation in detector technologies**
  - Innovation in accelerator technologies
- ***Deadline March 17, 2020***
  - Expression of Interest call May 2019
  - Open community meeting September 2019
    - Proposal team to form, scrutinise EoIs
  - New project could start end of 2020
- See these slides for a bit more background discussion
- For Advanced Software we plan a F2F meeting at CERN 17-18 June

Advanced Software++

- Address questions of *innovation* and *integrated activities*
- Software timelines not constrained in the same way as detectors
  - HL-LHC targeted developments certainly still possible
  - However, the path from development to robust production can be long
- Key drivers:
  - Increased event rates and complexity for new experiments
  - Need for high precision physics – standard model deviations
  - CPU pre-core improvements ~stalled
  - Accelerators are part of the landscape at computing centres (espc. HPCs)
  - New software techniques arise – machine learning
- Project ideas:
  - DD4hep and geometry/alignment tools – what's needed next?
  - Foundational libraries for low level (VecGeom/Core/Math)
  - Efficient reconstruction algorithms: ACTS, Advanced Pflow
  - EDMs for the accelerated landscape
  - Frameworks to bind them all together