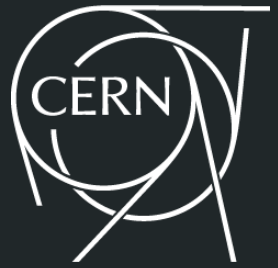


HSF and EP R&D Plans

Graeme Stewart

2021-01-18

HEP Software Foundation

Graeme, with HSF Coordination and WG Conveners

HEP Software Foundation Origins

- Organisation that started in 2015 with the aim of facilitates *cooperation* and *common efforts* in High Energy Physics software and computing internationally
- Inspired by the challenges faced by the field in the coming decades
  - From the ambitious physics programmes that we have
    - Run 3 upgrades for ALICE and LHCb
      - New detectors, far high data rates being pushed into the software pipelines
    - HL-LHC for ATLAS and CMS
      - Pileup of up to 200, trigger rates increase by ~x10
    - To say nothing of Belle II, DUNE, FCC, ...
  - From computing technology evolution
    - The ‘free lunch’ has been over for more than a decade
    - We are still learning how to effectively use parallelism
      - GPUs are becoming fairly ubiquitous, along with other hardware (TPUs, FPGAs)

Potted History of the HSF

- 2015 HSF gets started
  - Discussion on packaging tools and licenses are some of the first activities
- 2016 Charged by WLCG to produce a vision of computing for 2020s
- 2017 Community White Paper is written
  - 310 authors from 124 institutes, 14 chapters covering HEP software and computing matters
- 2018 First workshops digesting the outcomes of the CWP
  - WLCG-HSF Workshop in Naples
  - HSF-LPCC Workshop on Event Generators
  - PyHEP group starts up at Sofia 2018 pre-CHEP workshop
- 2019
  - Setup new working groups in Analysis, Simulation, Reconstruction
  - Significant involvement in European Strategy process
- 2020
  - Moved to the successful working group model in other areas
  - Formalised role at the LHCC, involvement in HL-LHC software and computing review

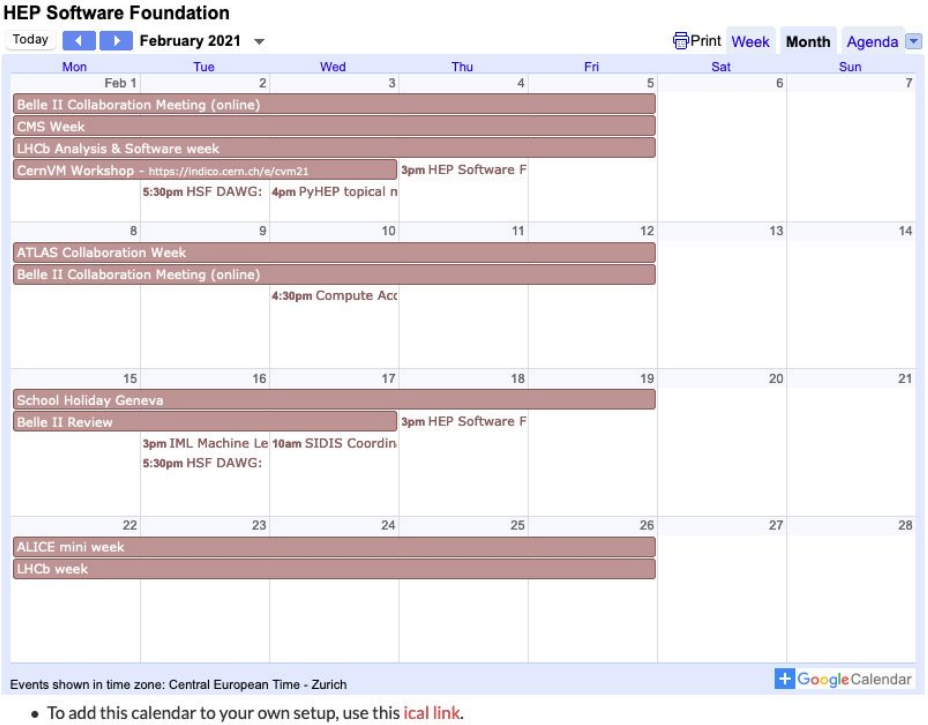
Coordination Meetings

- Our “beating heart” as the HSF!
  - Pass news, review upcoming activities, discuss matters where everyone could contribute
  - Open to all, with coordination and conveners largely attending
- We moved to bi-weekly in 2020
  - On odd-numbered weeks (use even-weeks if there is an exceptional need)
- We have continued to plan on bi-weekly meetings for 2021
  - https://indico.cern.ch/category/7970/
  - All in the HSF calendar...

Community Calendar

- Service we run for the community
- It is the go-to place to check that events you want to host don't clash with others
- Add this to your own calendar via iCal
- Add your events...
  - Ask coordination team (or Graeme)

Future HSF and Community Events



- To add this calendar to your own setup, use this ical link.

Working Groups

- Working groups provide the bulk of the HSF's substantive activity
- Focused on one area, though these are often cross-cutting
- Different working groups adopt different modes of working
  - This seems pretty naturally connected to their different areas of interest
  - In particular, Training and PyHEP have a very large engagement outwith the HSF and HEP software development 'core', so focus on larger events
  - Autonomy for each working group is actually a good thing
    - The convenors are experts in the areas they work in
  - Refresh of convenors in 4 groups for 2021
    - 1-2-3 or year terms

Working Group Convenors in 2021 (new)

Data Analysis
- Chris Burr (LHCb)
- Allison Hall (CMS)
- Teng Jian Khoo (ATLAS)

Detector Simulation
- Ben Morgan (ATLAS+Geant4)
- Krzysztof Genser (Mu2e+Geant4)
- Kevin Pedro (CMS)

Frameworks
- Chris Jones (CMS)
- Kyle Knoepfel (Neutrino expts.)
- Attila Krasznahorkay (ATLAS)

Physics Generators
- Andrea Valassi (LHCb)
- Efe Yazgan (CMS)
- Josh McFayden (ATLAS)

PyHEP
- Eduardo Rodrigues (LHCb)
- Jim Pivarski (CMS)
- Ben Krikler (LZ)

Reconstruction and Software Triggers
- Dorothea vom Bruch (LHCb)
- Andreas Salzburger (ATLAS)
- David Lange (CMS)

Software Developer Tools and Packaging
- Serhan Mete (ATLAS)
- Marc Paterno (DUNE)
- Mircho Rozodov (CMS)

Training
- Sudhir Malik (CMS)
- Meirin Oan Evans (ATLAS)
- Michel Hernandez Villanueva (Belle II)

Working Group Plans in 2021

- Detailed plans are being made by each working group now... and we will be presenting them at this Thursday's Coordination Meeting

Working Group	**Early Planned Activity in 2021 (*italics* means things likely to be touched on, but TBD)**
Data Analysis	Metadata handling for analysis
Detector Simulation	*Fast simulation ML, Accelerator R&D*
Frameworks	Offload to accelerators
Physics Generators	*Negative event weights, Accelerator R&D*
PyHEP	Topical meetings throughout 2021 (Python module of the month!); PyHEP 2021
Reco and Triggers	*Accelerator based reconstruction and optimisation, Real time reconstruction/analysis*
Tools and Packaging	Spack, performance analysis tools
Training	Run training following on from very successful 2020, develop material from hackathon

Compute Accelerator Forum

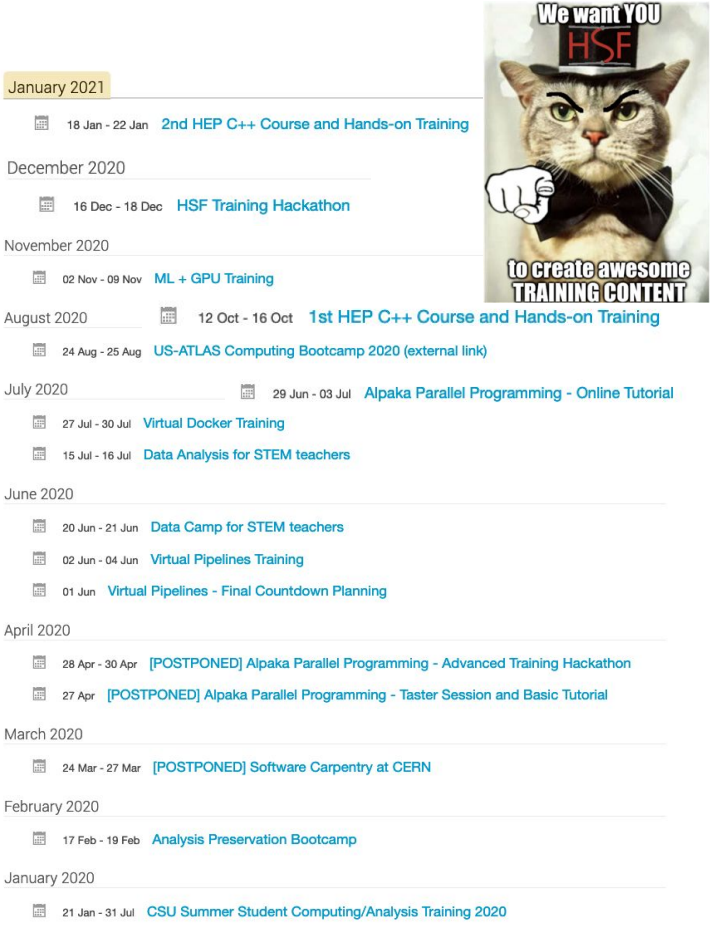
- Started in October last year
  - Cross-cutting talks on this important topic
    - Co-organised with openlab and SIDIS
  - Last year looked at
    - How do I get a GPU (access at CERN and in WLCG)?
    - Device abstraction in LHCb's Allen and ALICE's reconstruction code
    - Nsight introduction/tutorial from NVIDIA
  - Busy schedule planned for this year, https://indico.cern.ch/category/12741/...
    - Just looked at SYCL and HLS4ML
    - Detector Geometry on GPUs
    - Belle II GPU Reconstruction
    - Libcu++
    - More abstraction APIs (Alpaka, etc.)
- Topic suggestions welcome:
  - compute-accelerator-forum-organizers@cern.ch

+

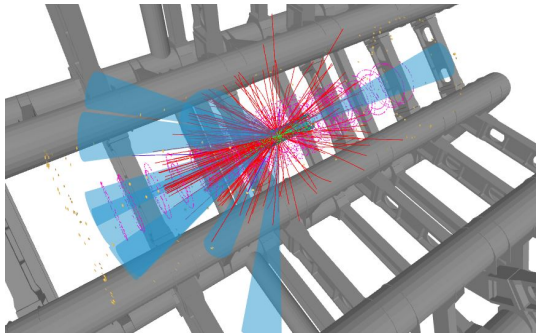
+

Training Development

- We have a big problem in our community to provide *scalable* and *sustainable* training
- So as well as the extremely active hosting of training events that happened last year
- We want to work on reusable modern training material
  - Following on from the hackathon
- C++ is a particular area where generic training material is weak
  - So we have now started to develop our own
  - https://github.com/hsf-training/hsf-training-cpp-webpage
  - Hope to complete this Basic Modern C++ course by the summer (or before!)
- *This is a particularly fruitful area for people to contribute too*



HSF Activities and GSoC



- We have a few other activity areas, which are quite variable in their activity
  - Quantum computing (mostly moved up to openlab and CERN quantum initiative)
  - Event displays (Phoenix is very active as a project)
  - iDDS (intelligent Data Delivery Service)
  - Differentiable computing
- Our most important contribution here is, without a doubt, *Google Summer of Code* and *Season of Docs*
  - Building on success year on year
  - 2021 should be no exception
  - N.B. This year's GSoC changes significantly the length of student projects
    - This will require more focused proposals
    - See recent messages from organisers , proposals by 15 February

Workshops

- Our workshop plans in 2020 were substantially changed by the Covid pandemic
  - We had to drop the visit to Lund and turn that into the May virtual workshop with WLCG
    - Focus on New Architectures, Portability, and Sustainability
  - Organised a second virtual workshop in November
    - Highlighted a few WG areas (Simulation, Generators), plus an open call for abstracts
  - N.B. Experience of running virtual workshops was reported to HEPiX and in EP Department Newsletter
- These were a success, but… pandemic continues and there is some sense of Zoom-fatigue
- Constraints from large software and computing events (see HSF calendar) in addition to obvious experiment weeks: vCHEP 17-25 May; ACAT 29 Nov - 3 Dec
- We decided *not* to have a long workshop in the first half of the year
  - We may go for a workshop in the Autumn (27 Sept - 1 Oct), TBD
  - Try to focus on ‘one-shot’ events, crossing WG boundaries as needed (e.g., I/O from the software, framework and site point of view)

Organisational Engagement

- Notwithstanding that the HSF is an organisation of people in HEP and mostly in experiments and often in other projects…
- Who does the HSF engage with?
  - Experiments
  - Openlab & SIDIS
  - Nuclear physics
  - Funded R&D projects
    - IRIS-HEP
    - HEP-CCE
    - SwiftHEP & Excalibur-HEP
    - EP R&D
  - ESCAPE
  - ECFA and EPPSU
  - Snowmass

Nuclear Physics and the Software and Computing Roundtable

- Nuclear physics is a very close cousin of HEP
  - Growing size of collaborations
  - Electron-Ion Collider (EIC, will be built at BNL, USA)
  - FAIR Project (Darmstadt, Germany)
- We contributed to and heard from the Nuclear Physics Trends Workshop
- Concrete outcome is that we have joined the Software and Computing Roundtable (Agendas: 2020, 2021)
  - Covers a wide range of topics (v. similar to our HSF Software Forum, but also computing)
  - Started as a JLab meeting series, then joined by BNL, then by HSF at end of last year
  - Strong nuclear physics presence, so excellent for engagement with this community
  - Also busy, one meeting per month
    - Next month, Jakob will cover ROOT I/O and RNTuple!

LHCC and HL-LHC Review

- Regular engagement with LHCC
  - Meeting 4x year between referees and WLCG, LHC experiments and HSF (for common software) each time
    - Topics are decided by the referees, but we can make suggestions!
    - Recently has covered ROOT, Detector Simulation, Generators, plus a few general summary talks
  - Liz Sexton-Kennedy and Graeme are officially appointed as Software Liaisons to WLCG from the end of last year
- HL-LHC Review
  - We prepared an HSF document covering the important software areas last year
    - That was a lot of work (thank you!) and was well received by the reviewers
  - This year (Nov) we have the second phase of the review
  - Most work will go out to specific software projects (inc. ROOT and Geant4), but HSF (through Graeme and Liz) has a role to help in coordination

HSF Summary and Outlook...

- We did a lot in 2020
  - It was a challenging year for everyone, in many different ways
  - HSF has achieved reputation and recognition in HEP and is appreciated
  - This is a great outcome for a community based coalition-of-the-willing
- We are on a good path for 2021
  - Active engagement and creative thinking are the key to success
  - So of course corrections throughout the year are to be expected!
  - We have a great team of people in the HSF
  - *There are many areas where you can be involved and contribute!*

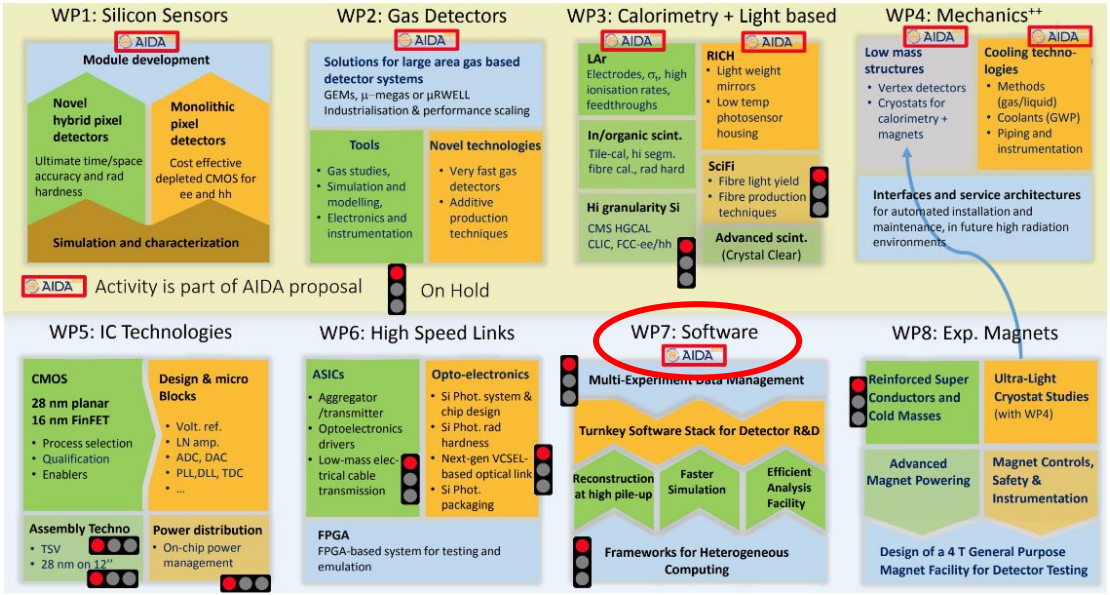
HSF Forum Mailing List: hsf-forum@googlegroups.com

EP R&D - Software Work Package

Graeme and Jakob, plus task leaders and our fellows and students

CERN EP R&D Programme

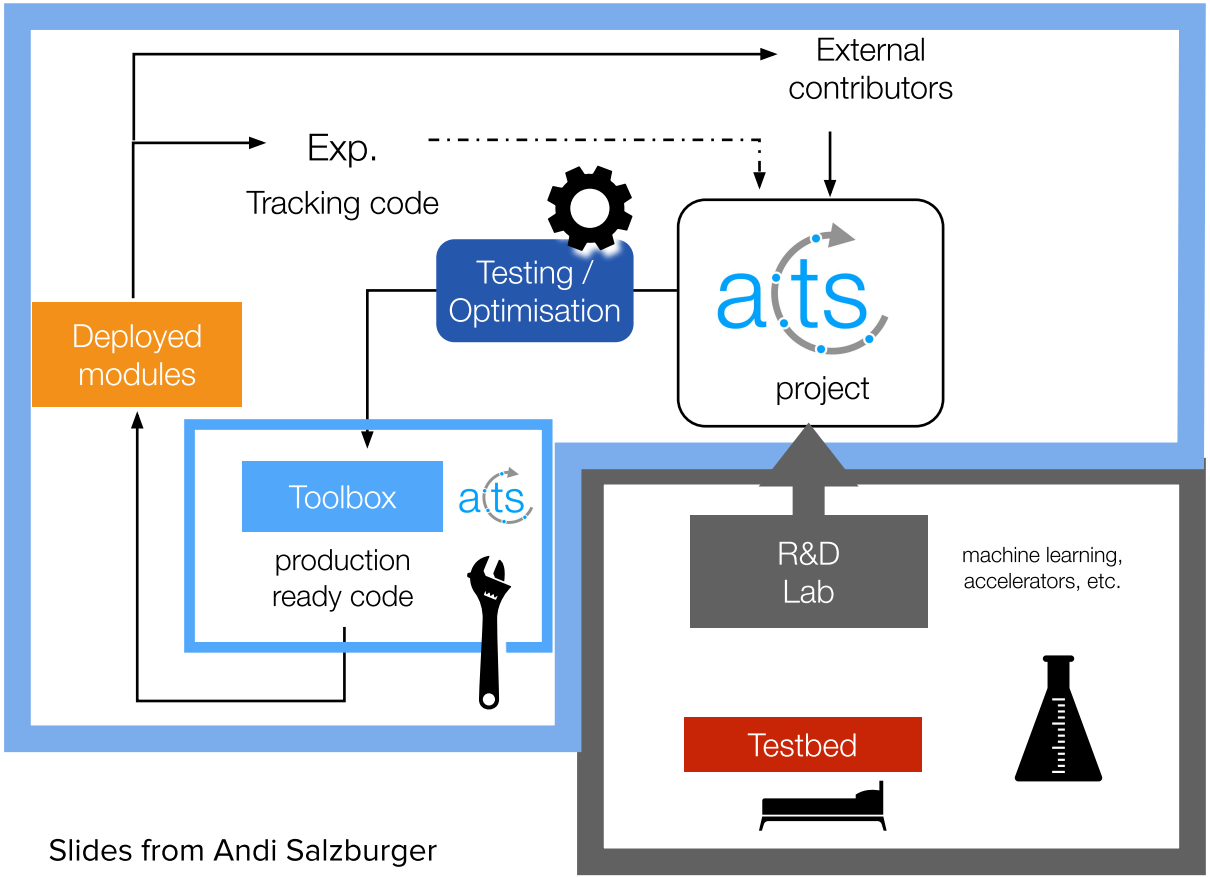
- Programme on R&D for new Detector Technologies
- Started planning in 2018, funded from 2020



Software Work Package (link)

- Turnkey Software Stack, Key4hep
  - Flexible and modern software stack, ready for physics studies for future experiments
  - Valentin Volkl working with André since January 2020
  - Planning report next week (25 Jan)
- Faster Simulation
  - Machine learning based fast, generic, simulation techniques
  - Dalila Salami joined SFT this month
  - Planning report in two weeks (1 Feb)
- Efficient Analysis
  - Design data structures and interfaces to support very high throughput analysis
  - Vincenzo Padulano and Javier Lopez-Gomez working with Jakob and Enric
  - Planning report in three weeks (8 Feb)
- Reconstruction at High Pileup...

project - Roadmap

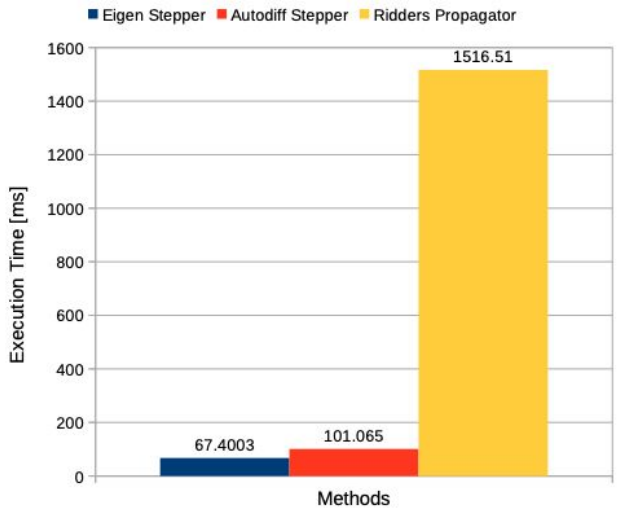


- Toolbox with production-ready code
  - e.g. ATLAS Run-3 vertexing
- R&D/Lab that feeds into toolbox, 2 main R&D lines
  - Tracking & ML
  - Tracking on GPUs
- TrackML & OpenDataDetector
  - Quasi-realistic testbeds

Slides from Andi Salzburger

current R&D projects ML

- Seed & track classification
  - DNN seed classifier, DNN ambiguity solver
- Auto-differentiation [WLCG/HSF Workshop, Nov 2021]
  - Automated Jacobian calculation using auto-diff
- NEW: ML Detector navigator
  - Learned navigation through detector using embedded space
- Bucketing, TrackNet [NeurIPS2021]
  - Approximate Nearest neighbourhood bucketing, clustering in learned metric space

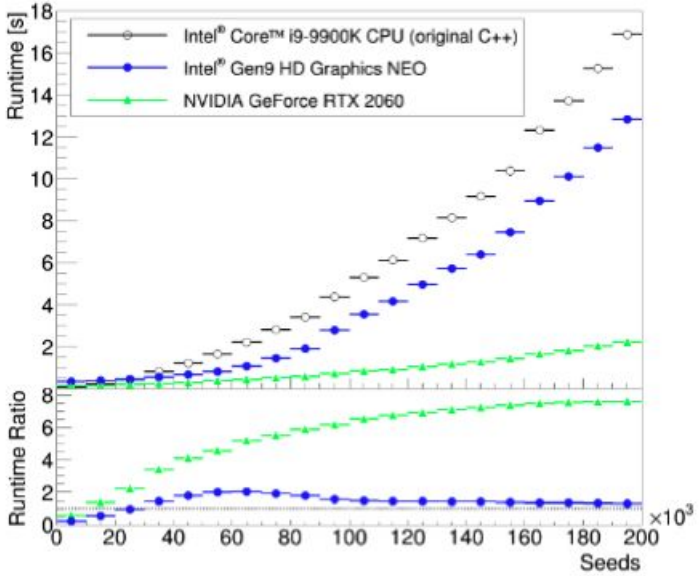


Slides from Andi Salzburger

current R&D projects GPU

[Overview talk: X. Ai, WLCG-HSF Workshop, Nov 2020]

- GPU Seeding algorithm
  - 1 CPU version, vs 2 CUDA versions, 1 SYCL version
- Prototypes: Runge-Kutta-Propagation, Kalman Filter
  - Still lacking proper geometry support
- NEW: GPU-friendly geometry project for Tracking
  - Acts/detray library to be released in Q1/2021
  - Non-polymorphic geometry library
- STARTING: Chain demonstrator
  - Clustering - Space Point Formation - Seeding



Comparison 1-core CPU vs GPU,
multi-core CPUs remain currently competitive

Slides from Andi Salzburger

Introduction

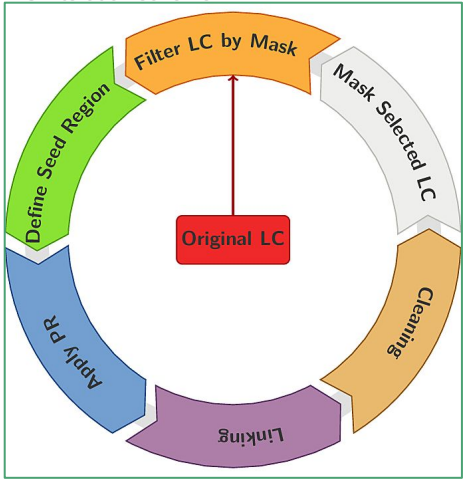
- Particle shower reconstruction in **high-granularity calorimeter** is very interesting task in high-density environments
  - Many showers tend to **overlap**
    - Typical situation at HL-LHC
  - Standard reconstruction algorithms using combinatorics are expected to fail due to **memory/timing explosion**
  - **Fertile ground** for new techniques and algorithms: clustering, machine learning, graph theory, and modern computer architectures
    - Planned and designed, taking into account the information from the tracking system and timing detectors
  - Development can profit from **experience** with Particle Flow techniques of CMS experiment
  - **New flexible framework** can be re-used in other (future) experiments using high-granularity calorimeters facing high pile-up conditions

Slides from Erica Brondolin

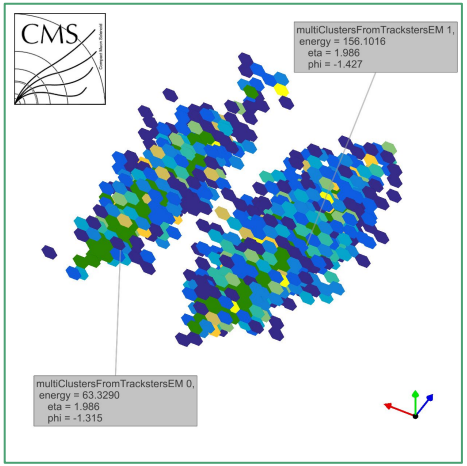
What is TICL?

- TICL (**The Iterative Clustering**) is a **modular framework** integrated and under development in CMSSW
- Main **purpose**: processing calo rechits (x, y, z, t, E) and returning particle properties and probabilities
- In a nutshell: grouping 2D layer clusters (CLUE) into 3D clusters (Tracksters) **iteratively** to reconstruct different particle species
- Important features:
  - No prior knowledge of CMSSW needed to contribute
  - Modules are designed such that new algorithms or techniques (e.g. Machine Learning) can be plugged on top easily
  - Algorithms are designed as swappable plugins, with heterogeneous architectures / portability in mind
  - Mostly geometry independent
- Documentation

TICL iteration scheme



Example of two close-by reconstructed tracksters



Slides from Erica Brondolin

Conclusion

- EP R&D got off to a decent start in the software work package in 2020
  - We were less affected than other WPs where lab work is more important
  - We also managed to procure some nice machines for the R&D work
    - Which are being hosted in the Meyrin data centre
- Excellent links from the R&D to other activities on software
- This year should be a very productive one!