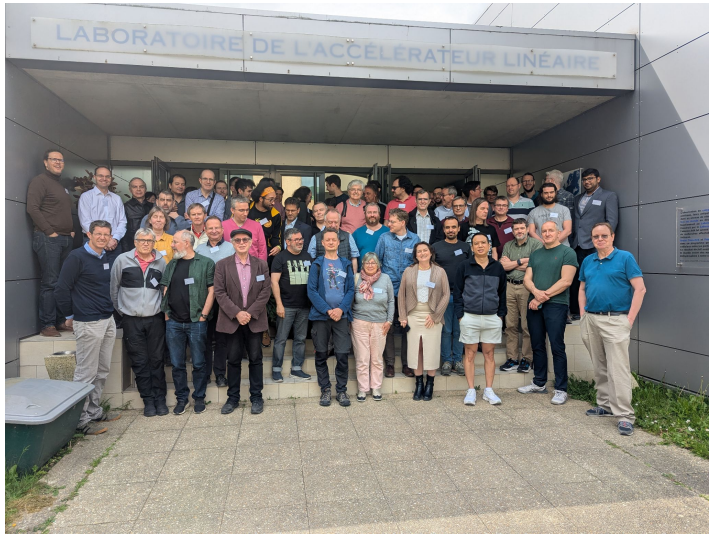
Community Software and Software Sustainability

Graeme Stewart (CERN)

PUNCH Lunch Seminar, 26 June 2025

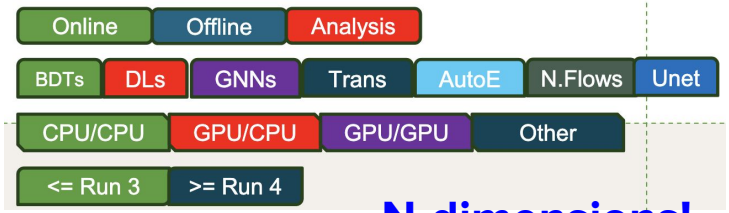
WLCG/HSF Workshop in Orsay

- Worldwide LHC Computing Grid and HEP Software Foundation workshop
  - Tradition now to bring together the distributed computing and the HEP software communities
- Latest workshop was in May 2025, at IJCLab in Orsay
  - Covering many aspects of computing and Software, including...
    - **AI and heterogeneous resources**
    - **Common software**
    - WLCG facilities
    - Analysis at HL-LHC
    - WLCG operations
    - **Sustainability**
- I will give an overview here of the **bold** topics



AI Overview

- AI has been disruptive, highly beneficial and very broadly applicable
  - Cheaper: ultra fast high-fidelity simulation
  - Better: x100 improvement in background rejection
  - New: anomaly detection with auto-encoders
  - It also has, and is, moving very fast
    - New techniques come all the time
    - Areas of applicability expand
  - AI will impact our computing resources
    - Need to prepare for this
    - Providing the correct resources and (grid) accounting for them



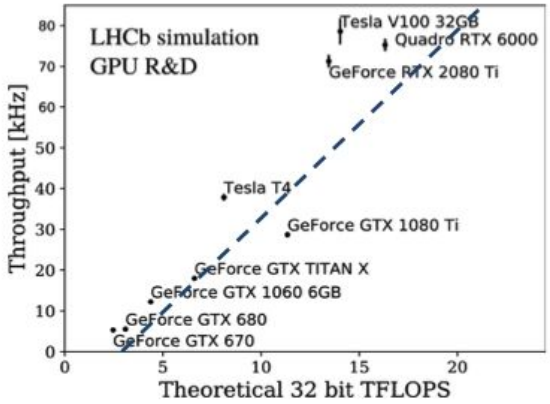
N-dimensions!

- **1st axis: application domain**
  - **Online, offline, analysis**
- **2nd axis: AI technology**
  - **GNNs, GANs, Fully connected, …**
- **3rd axis: resource domain**
  - **Training: laptop, server, GPU, many servers and GPUs**
  - **Inference: laptop, WLCG nodes, GPUs, FPGA, …**
- **4th axis: experiment**
- **5th axis: when?**

To be followed up in the “Heterogeneous architectures in WLCG” workshop Dec 3-5

Heterogeneous computing

- What computing that we do on CPUs could be envisioned on GPUs in 5 or 10 years?
  - We have some outstanding examples of success: ALICE TPC reconstruction, LHCb Allen L1 trigger
- What could be next?
  - Event generation
  - Detector simulation
- KPI: event throughput per unit cost
- Heterogeneous computing is challenging and needs training
  - Not too much interest from students in C++/legacy codes
    - ***Long time to train people, struggle to retain***
  - Can we live with single/reduced precision?
    - NGT - Openlab "Optimising Floating Point Precision" Workshop next week
  - Not many experts in our field (CUDA, Alpaka, ...)
  - Distill experience and expertise in languages, compilers, platforms, etc.

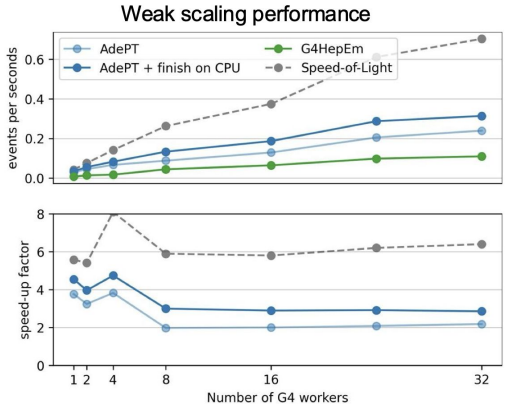


To be followed up in the “Heterogeneous architectures in WLCG” workshop Dec 3-5

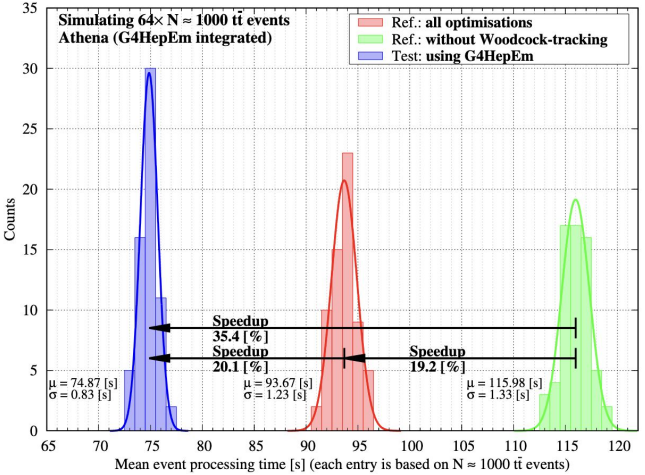
Common software - Simulation with G4HepEm, AdePT

Accelerate HEP simulations of electrons, positrons and photons:

- G4HepEm: specialized tracking for $e-$, $e+$, and γ on **CPU**
  - Provides a significant (20%) speedup for the Run 4 ATLAS/CMS detector simulations
- AdePT: offloading electromagnetic showers to **GPU**
  - Still very much R&D, some promising indications of speed-ups, but not proven yet in full use cases
  - Geometry bottleneck (VecGeom) is being worked on

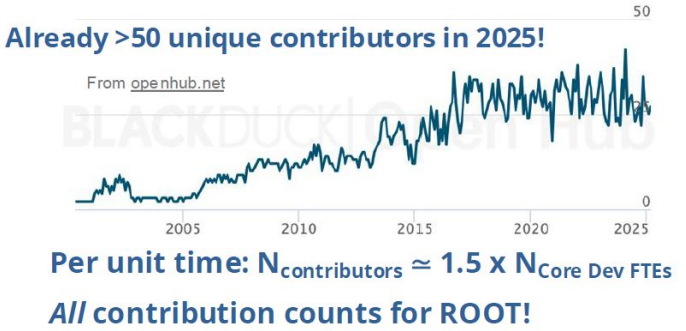


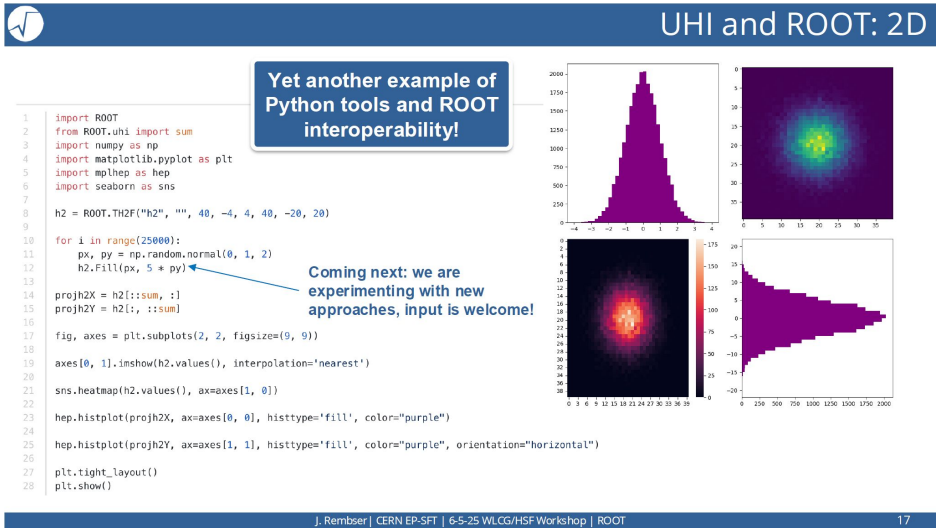
~1.4x speedup in CMS geometry



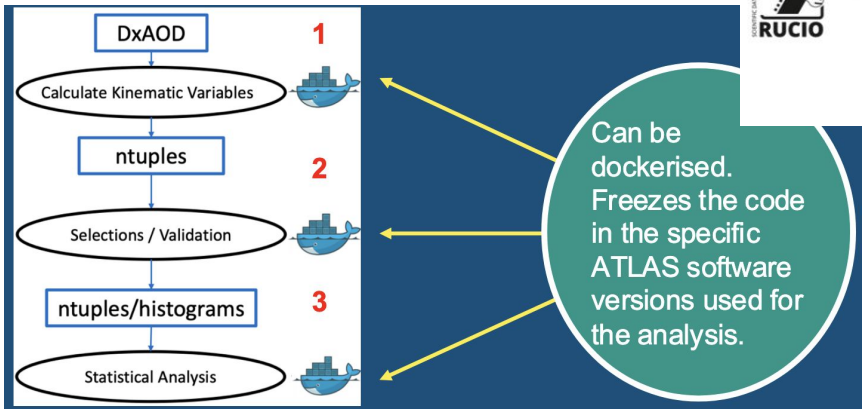
Common software - ROOT and RooFIT

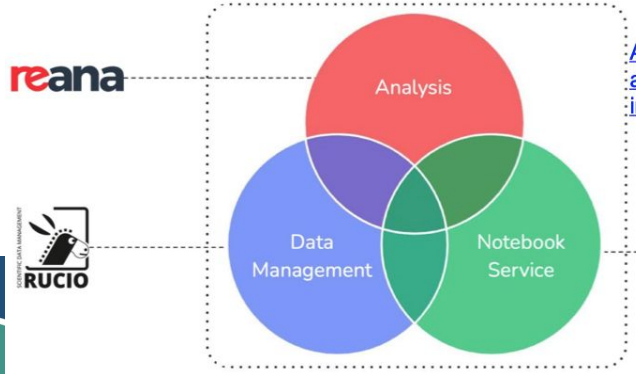
- **2025 Programme of Work proceeding well**
- Steady increase in number of issues closed over the past year
- **1st stable RNTuple format released in 6.34.00**
- New histograms under development
  towards a thread-safe & fast class (for ROOT7)
- RFile will replace TFile (WIP)
- A lot of work on interoperability (SOFIE, for ML)
- **Python ecosystem seeing a lot of attention**
- Histograms to honour Scikit-HEP's proposed
  UHI protocol
- **Many significant performance enhancements in RooFit - for CPUs and GPUs**
  - E.g., Automatic Differentiation available via Clad





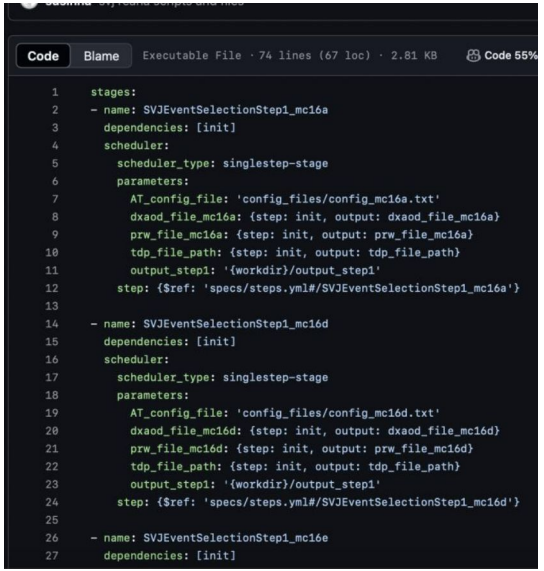
Common software - Virtual Research Env. user experience





VRE@CERN

- ✅ Available at the click of a button
- ✅ Data persistence
- ✅ No account restrictions
- ✅ Change easily the software stack
- ❌ Access >200 cores
- ❌ Access to batch
- ❌ Access to GPUs
- ✅ Integration with Data Lake tech.
- ✅ Integration with REANA



- Good experience using the CERN VRE and rerunning an ATLAS analysis
  - Containerised workflow steps key

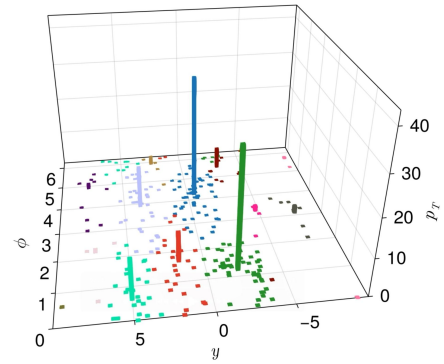
Sustainable Software

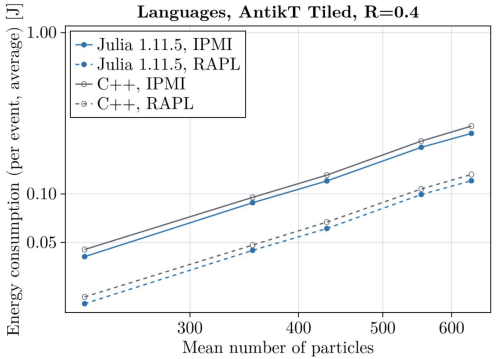
What do we mean by sustainable?

- Efficient Code (for machines)
  - Energy consumption to do a calculation should be as low as possible
  - Worldwide data centre power consumption is about 1.5%, but rising fast
    - Computing for LHC has a significant environmental footprint, which we want to minimise
- Sustainable Code (for humans)
  - Code needs to be maintained for decades
    - Experiment lifetimes are long
    - Preservation lifetimes for results are even longer
  - Code needs to be easy as possible to understand and maintain
  - Concentrate on our problems, reuse as much as possible

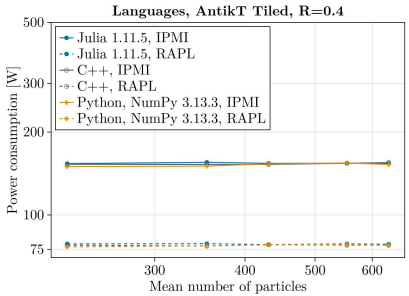
Sustainable software - Julia

- Jet reconstruction is a good “Goldilocks” problem to study
- Sequential jet reconstruction implemented natively in Julia
  - Tested alongside C++ and Python
- Excellent ergonomics in Julia
  - Expressive, compact code
  - Array operations and broadcast
  - Code composes very well
  - Macros allow spot optimisation





Julia can be faster (=use less energy) than C++ due to superior use of SIMD - many ergonomic/ecosystem advantages



CPUs use ~the same energy at 100%, irrespective of code source

```
@inbounds dij_min = dij[1]
@turbo for here in 2:n
    newmin = dij[here] < dij_min
    best = newmin ? here : best
    dij_min = newmin ? dij[here] : dij_min
end
```

Sustainable software: EVERSE Project

Engagement with 5 European Science Clusters ESCAPE covers HEP, plus Life Science, Environmental, Social, Photon and Neutron

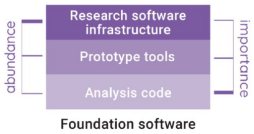
eosc | EVERSE

Establishing a Community

How to contribute to, and engage with EVERSE

Elements of EVERSE

- The Network (https://everse.software/network/)
- RSQkit (https://everse.software/RSQKit/)
- Software Reference model
- Training
- Recognition framework



Join Us

Any individual or organization that agrees with our vision statement is welcome to join the network

EVERSE: European Virtual Institute for Research Software Excellence

12

eosc | EVERSE

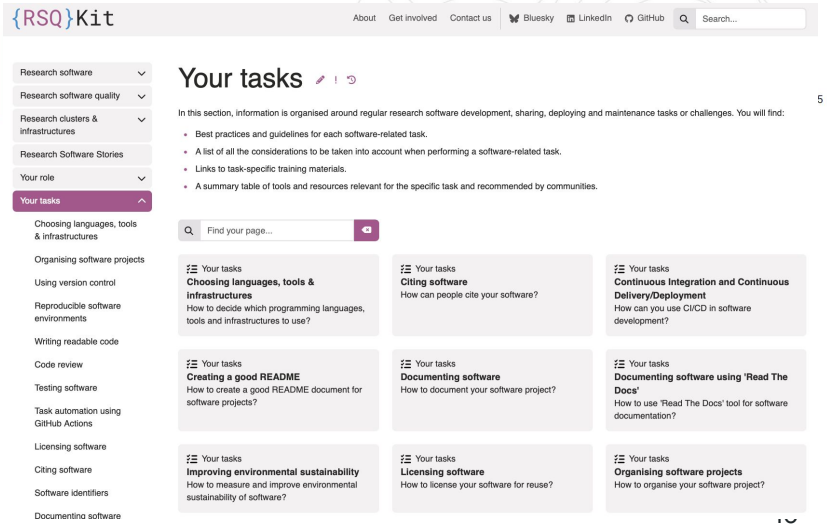
EVERSE

Paving the way towards a
European **V**irtual Institut**e** for **R**esearch **S**oftware **E**xcellence

EVERSE aims to create a framework for research software and code excellence, collaboratively designed and championed by the research communities, in pursuit of building a European network of Research Software Quality and setting the foundations of a future Virtual Institute for Research Software Excellence

- ensure research software curation, quality, preservation and adoption of best practices, by the Communities, for the Communities, build on collaboration with the five EOSC Science Clusters
- adopt a three-tier model for research software, i.e., analysis code, prototype tools and research software infrastructure, which captures the varying complexity of research software and its development, and can be used as a basis for research software excellence

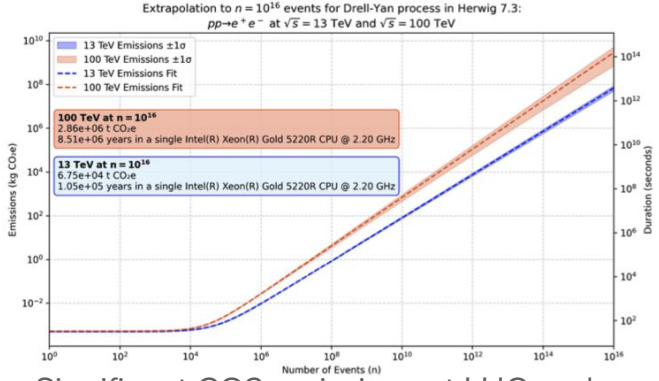
oth developers and software are essential components of our strategy to promote sustainable software practices



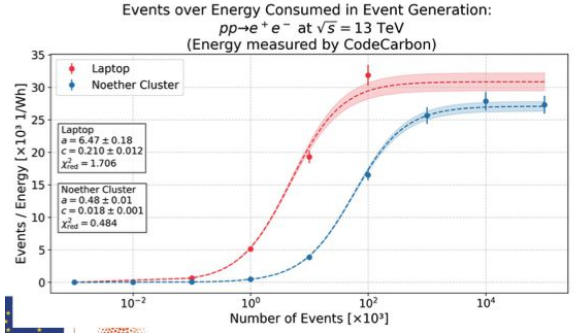
Green Software - Generator Benchmarks

- Goal is to understand the energy cost of computing at University of Manchester
- Drell-Yan events simulated in Herwig 7.3
  - Energy consumption grows linearly with number of events
- CO^2 Emissions vary widely per country
- Lifecycle environmental impact is very complicated
  - Vendor numbers are not standardised
  - Difficult to get estimates for modern equipment
  - Idea is to grow expertise in the community on understanding these impacts
    - Improve policies to lower impact

- Extrapolation to LHC scales (10^16 events)
  - LHC (13 TeV): 6.75e+04 t CO2e
  - FCC (100 TeV): 2.86e+06 t CO2e



Significant CO2 emissions at LHC scale workloads



Laptops are efficient for small workloads!

Outlook and next steps

HEP Software Foundation

- Software's importance can hardly be overstated!
- Software projects are making great strides towards HL-LHC
  - AI is becoming ubiquitous
  - Heterogeneous computing is hard, but there are successes and promising R&D
- HSF's mission is to help us to get to better, more sustainable and efficient software in the field
  - Will work closely with projects like EVERSE to define indicators and support tools that do that
  - VREs helping also with reproducibility
- Understanding sustainable software can be improved
  - Develop expertise in assessing the impacts of our software
  - Leverage important trends towards usability and efficiency (Eduardo's talk on Scikit-HEP)
  - Improve the community's understanding and level of skills (Michel's talk on training)

The HSF is bottom up - we want to hear your ideas and suggestions too!

Mailing List: hsf-forum@googlegroups.com; Website: https://hepsoftwarefoundation.org; Contact: hsf-steering@googlegroups.com