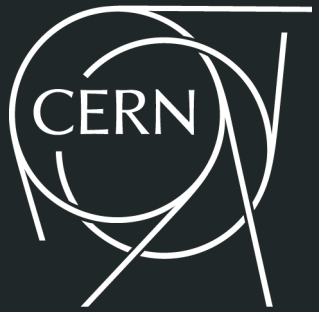


Software Highlights and Review Preparation

Graeme Stewart (CERN EP-SFT) and Liz Sexton-Kennedy (FNAL)

LHCC Referees, 2021-03-02

Software Highlights - 2021 Plans of Work

- SFT Projects' 2021 plans of work have been presented
  - ROOT, https://indico.cern.ch/event/996294/
    - Many meetings with stakeholders and experiments in advance
  - Geant4 and Simulation R&D, https://indico.cern.ch/event/996264/
    - Requirements gathered from, e.g., technical forum meeting(s)
  - SPI and Key4hep, https://indico.cern.ch/event/996262/
  - CernVM, https://indico.cern.ch/event/1008722/
- For HSF we also did some planning for 2021
  - Details about plans that we have in general and also from working groups
  - Nothing cast in stone, but good to present these to the community and experiments
- Additional comments and input, of course, very welcome for all plans

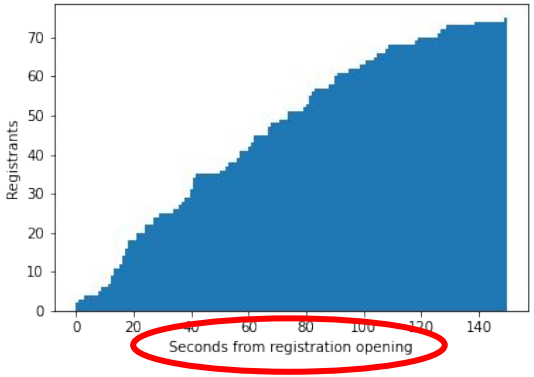
Software Highlights - New releases

- ROOT
  - v6.22/06 released 1 December 2020
    - Important bug fixes for the LHC experiments, OS X Big Sur and Apple's M1 supported
- Geant4
  - G4 10.7 released 4 December 2020
    - New VecGeom, templated field classes with better diagnostics, new tasking system, CMake improvements for modular builds, data structure optimisations (5-7% speedup!)
    - Improved GFlash for fast simulation, many physics model improvements (see backup)
  - 10.7p1 released 5 February 2021
- CernVM and CVMFS
  - CernVM 4.5, 27 January 2021 (EL 7.9)
  - CernVM-FS 2.8.0, 1 February 2021
    - OS X Big Sur, WSL2, parallelised garbage collection, template transactions
- SPI
  - LCG_99 release rolled out December to January
    - Python3 now the default, ROOT 6.22/06 based

Project Highlights

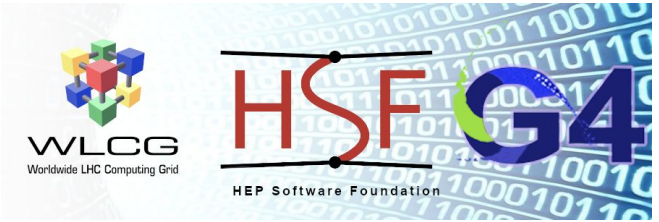
- ROOT
  - Merged llvm9 - full official C++17 support
    - Makes ROOT 6.24/00 imminent
  - Many papers coming at vCHEP: RNTuple, RANLUX++, distributed analysis
  - Effort: new RooFit developer; loosing ML people; RNTuple development held back by tension with CVMFS (same developer)
  - See backup for additional items
- SPI and Key4hep
  - Spack prototype now consistent with LCG ROOT build
  - Python3 is the default
  - devARM builds and more CUDA builds now available; support for AdePT
  - Gaudi tested in the nightly builds
- CernVM
  - 2021 Workshop held 1-2 February (virtual NIKEF)
  - Almost 100 participants

Training Events



- Second iteration of the C++ course ran 18-22 January
  - 75 places taken in 2½ minutes!
  - 134 people on the waiting list
  - Next course planned for August
  - Work ongoing to carpentry-fy this kind of training material
    - Try to have material suitable for self-study and sustainable training (any expert can teach)
    - Rebase the material to emphasise best modern practices
- This grew out of the Training Hackathon that happened at the end of 2020
- Github CI/CD Training course ran 16-20 February
  - 200 people registered - people seemed to be very happy with the course
  - Still trying to understand the best format for interactive work

Other News and Meetings



- HSF/WLCG Workshop took place in November (19-24), about 100 active participants most days
  - Sessions on WG updates (PyHEP, Training), detector simulation, generators and general R&D
- The new Compute Accelerator Forum is proving very popular (attracting around 70 participants; overview talk in February GDB)
  - Meetings planned now up to the summer
- HSF has joined forces with JLab and BNL colleagues to organise the Software and Computing Roundtable
  - Common topics for the community, with strong nuclear physics links
- GSoC projects have been proposed again, with the HSF acting as in umbrella organisation
  - 37 proposals (cf. 48 last year)
  - Why?
    - Remote interaction exhaustion?
    - Change in format, only half the coding hours?
  - We still believe this to be a useful resource, of course
- Setup an HSF sponsored Linux4Science discussion area
  - Due to recent RHEL announcements
  - https://github.com/HSF/Linux4Science/discussions/
  - Particularly to interact with other labs (not CERN/Fermilab) and smaller experiments

HL-LHC Review Planning

HL-LHC Review Planning

- We have had a number of meetings to decide on how to prepare for the review (WLCG Project Lead, Software Liaisons, Computing Coordinators)
- We have converged on the following broad structure:
  - Introduction
  - Event Generators
  - Detector Simulation
  - Foundation and Core Tools
  - Analysis
  - DOMA
- We have not included reconstruction projects as, after discussion, none were felt to require the attention of this review

Introduction

- Probably a shorter document
  - Outline the selection of common software areas
  - Describe any important differences in the needs of each experiment from CSAs
  - Mention briefly software which is important, but we don’t think needs to be reviewed (e.g., software where HL-LHC scaling is not an issue, longevity is assured, alternatives are available)

Event Generators

- We will prepare this document with HSF Generator WG convenors and others as editors (experiments and theorists)
- Generators is a varied suite of different packages, which have relatively complex interactions
- However, already identifying some of the key ones for HL-LHC we have
  - Evtgen (managers are in LHCb, but work is independent)
  - MadGraph5_aMC@NLO
  - POWHEG (core + important processes?)
  - Pythia
  - Sherpa
- AA generators (HIJING, EPOS, Angantyr) still being discussed

Detector Simulation

- Geant4 remains the key piece of software used by the experiments, with a direct and highly significant resource consumption
  - Geant4 project will take primary charge here
  - Include geometry and detector description aspects
    - VecGeom, DD4hep
  - Development of any generic fast simulation
- R&D projects that look at doing parts of the simulation on accelerators we would also like to include
  - AdePT and Celeritas
  - However, even by November this is likely to still be very much R&D

Foundation Tools

- Will cover mainly the role of ROOT as a foundation layer
  - I/O system and its evolution
  - Geometry representation and event display
  - JSROOT
  - PyROOT
  - Maths
- We did consider Gaudi as a common LHCb/ATLAS framework, but we believe it is better reviewed later
  - Tied intimately to these two experiments, highly successful and not presenting any particular problems

Analysis

- Focus here on tools used at the end of the data processing chain
  - Analysis groups and users
- ROOT is the standard and carries the vast majority of current analysis
  - Roofit
  - Histograms
  - RDataFrame
  - Clearly the ROOT team lead this part of the document
- New Pythonic tools
  - Many tools with different realms of applicability, gaining traction and showing promise
  - Examples: Scikit-HEP project (pyhf, uproot, etc.), Coffea project, zfit fitting package
  - Supported strongly by IRIS-HEP and many community developers, from whom editors for this section will be drawn, along with the HSF PyHEP and Analysis convenors

DOMA

- Have discussed with the WLCG/DOMA community, identifying the key projects to cover
  - Rucio
  - File Transfer Service (FTS)
  - Storage interfaces and caching layers
    - CTA, dCache, Xcache, StoRM, xrootd, http TPC, etc.
  - Network technologies including monitoring and software defined networks
  - CVMFS
  - Token based authentication
- Currently starting to form an editorial team and to contact specific projects

General Remarks

- ROOT plays a really unique role and crosses boundaries
  - In two documents, to match the needs, but these will cross-reference each other a lot
  - Need a suitable place for the project specific inputs
- We are driving this process between WLCG (project lead and software liaison), the experiments (computing coordination) and the HSF
  - This is helped, of course, but strong cross over between software projects and experiments
- We will try to quickly identify if the requirements are well understood by all parties
  - A few mini-workshops will be organised
- We need to have drafts of reports ready in time for feedback
- This makes the timeline from now to 1 October quite tight (pilot beams in Sept.)
  - However, there should not be any significant problem in meeting the deadline

Backup

Geant4 10.7 Physics Highlights

- New general facilities based on GFlash for fast simulation models
- New EM model for polarized gamma elastic scattering
- New Coulomb scattering model for e+/e- based on very accurate Differential Cross-Sections using Dirac Partial Wave Analysis
- New thermal model of positronium decay to gammas
- New model for gamma-ray elastic interactions, able to account for molecular interference effects
- Extended Glauber-Gribov approach to cover heavy hadrons, for charm and bottom hadron-nuclear (elastic and inelastic) cross-sections
- Improved treatment of anti-baryon interactions in the Quark-Gluon-String (QGS) model
- Revised maximum energy of applicability of elastic and inelastic cross-sections for ions and anti-ions nuclear interactions
- New coalescence model, useful in particular for Cosmic Ray applications
- New utility to access hadronic processes, allowing to customise hadronic cross-sections per particle type

Additional ROOT Highlights

- Multi-prong HPC R&D with openlab
  - optimizing ROOT's next-gen I/O format RNTuple for Intel DAOS object store
  - optimizing RNTuple for HPC cluster I/O (Lustre)
- Ongoing integration of RNTuple into CMSSW for NanoAOD output, as early reality check
- Several “RDataFrame in production” studies / benchmarks from physics groups (e.g. https://indico.cern.ch/event/919839/contributions/4147763/attachments/2161997/3649407/cmgInteractive-Dec15-2020.pdf)
  - show the success of this approach for production analyses
- Dynamic compute libraries for RooFit, giving vectorization (and soon: GPU) acceleration factor > 10, part of upcoming v6.24
- distRDF now in ROOT, part pf v6.24: the (near-term) future of distributed analysis with ROOT, replacing PROOF in the RDataFrame era
- v6.24 will have a new architecture independent RanLux++ implementation, matching the performance of the original Intel-only assembler implementation, integrated into ROOT. Highly relevant for HEP and outside.