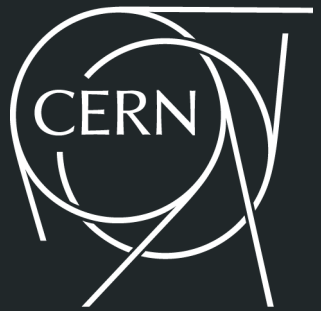
Software Update

Graeme Stewart, for EP-SFT and HSF



2019-09-10

Google Summer of Code & Season of Docs

- 33 slots from Google were used
  - 29 students successfully completed
  - Substantial influx of effort and broadens HEP standing in open source
- Google Season of Docs
  - New initiative to support a technical writer
  - ROOT project made a successful proposal and will work with a professional documentation writer

Event Generation WG

- Only two meetings (June 27, July 25) since the last LHCC report in July
- Further progress in ATLAS/CMS accounting analysis, but not yet completed
  - PRELIMINARY summary of MC event generation CPU for 2017 data:
    - ATLAS: MC gen 430Bs (17% of 2490Bs for MC gen+sim+rec)
    - CMS: MC gen 100Bs (9% of 1040Bs for MC gen+sim+rec)
  - MC gen as a fraction of all WLCG CPU usage: ATLAS~12%, CMS~5%
  - Generator (Sherpa vs MadGraph) explains part of the difference, not all
- MadGraph on GPU work at Argonne, based on 2008-2012 work in Japan
  - Very early stage – first benchmarking results expected after the summer
- Other interesting developments
  - Impact of negative weights discussed at the Les Houches workshop
  - ATLAS/CMS ttbar sample sharing is being considered
  - Proposal for PDGID extension

Reconstruction and Software Triggers WG

- Building stronger links with experiment reconstruction groups
  - Brings topics of common interest, such as GPUs
- Summer meeting programme:
  - Efficient data structures for accelerators and many-core architectures
  - Real-time Analysis and Partial event building
  - LHCb and ALICE trigger systems (followed similar meeting with ATLAS and CMS)
- Real-time analysis is a topic of increasing interest, request for webpage cataloging resources
- Intersting discussion with other communities (dark matter, gravitational waves) about synergies
  - Opportunity at the Joint ECFA-NuPECC-ApPEC Seminar (JENAS) at LAL in October
    - https://jenas-2019.lal.in2p3.fr/
  - Planning a general HSF talk, plus input from non-HEP communities about challenges and needs

Data Analysis WG

- Working group helping plan the pre-CHEP workshop in Adelaide
  - *Analysis Systems: From Future Facilities to Final Plots*
- Looking at everything we need to do analysis effectively in the future
  - From DOMA to the analysis software environment
- Planning Saturday afternoon of presentations
- Then Sunday morning of participative discussions focused on key challenge questions (to be finalised!):
  - Facilities - how to organise for HL-LHC?
  - Interactive Analysis - what are the needs and challenges?
  - Machine Learning - how to manage training and inference; what is the scale of training and who does it?
  - Analysis Model - what are viable models for the future? Key points to improve or change

ROOT and SWAN

- New version of ROOT, 6.18/00 released 25 June
  - One patch additional release 6.18/02 released 23 August
  - Many usability enhancements
    - Improved interplay with numpy in Python ecosystem
    - Reduced precision floats for smaller file sizes
    - Modernised analysis support with RDataFrame and RVec
  - Many additional performance improvements and bug fixes
  - Train the Trainer event held in June
    - Foster a training community of ROOT experts, armed with up-to-date best practice
- SWAN Service
  - Ongoing discussions on how to improve the service, IT and EP
  - Users workshop planned for 11 October - share use cases and discuss evolution

PyHEP

- First topical meeting planned this week on fitting tools
- PyHEP 2019 workshop in Abingdon, UK, 16-18 October, https://indico.cern.ch/e/PyHEP2019
  - Keynote presentation on the PyViz - open source visualization tools for Python - project, given by Philipp Rudiger, a member of the developers team.
  - Topical sessions on e.g. histogramming and statistics, including a talk and hands-on tutorials.
  - Lightning talks from participants.
  - Presentations following up from topics discussed at PyHEP 2018.

CernVM and CVMFS

- CernVM Users Workshop Debriefing
  - 2.5 days, >40 participants, 10% - 20% from outside HEP
  - Valuable user feedback on desired development priorities
    - Nightly build repositories publishing speed for Run 3 software developers
    - Kubernetes and user container support in anticipation of Run 3, 4 analysis workflows
  - Various external speakers (Cloudflare, Jülich HPC, Singularity, Red Hat)
    - Kicked off collaboration with Singularity on a CVMFS plugin for better HPC access, which is now released in Singularity 3.4
- CernVM 4 Support Update [July 2019]
- Several CVMFS patch releases [Aug-Sep 2019]
  - Various fixes and improvements for the S3 storage backend
- CernVM team is working towards the 2.7 winter release

LCG releases

- Version LCG_96 released early July
  - Based on ROOT v6.18
  - New entries: Gaudi v32r0, Garfield++
- Ongoing commissioning of version LCG_96a
  - ROOT v6.18/02, crucial fixes for Geant4/VecGeom and for handling of Python packages
  - Latest versions of several packages and generators
- Interest on the LCG stacks from smaller experiments increases
  - Recent new clients: NA61/SHINE, NA62
  - Manifested interest by COMPASS

Training

- Collaboration between HSF and FIRST-HEP fruitful
- Three core skills training events for HEP over the summer:
  - Argonne, https://indico.cern.ch/event/827231/
  - Princeton, http://codas-hep.org/
  - LBNL, https://indico.cern.ch/event/827232/
- Planning a Software Carpentries (https://software-carpentry.org/) event at CERN this year
  - 27-29 November, https://indico.cern.ch/event/834411/
- Appropriate C++ training is being discussed, no obvious solutions yet...

Frameworks Working Group

- Outcome of the last HSF/WLCG Workshop was renewed interest in progressing on software frameworks
  - Called for nominations from the community to restart this work
- We have just appointed...
  - Chris Jones, FNAL, CMS
  - Kyle Knoepfel, FNAL, Neutrino Community
  - Attila Krasznahorkay, CERN, ATLAS
- So we look forward to new activity here
  - Mandate: https://hepsoftwarefoundation.org/workinggroups/frameworks.html

CERN EP R&D

- New RDD budget line at CERN that will support R&D into detectors, including software
- Software work package will be able to support the following areas, 6 fellows in total:
  - Turnkey Software Stacks (optimised and integrated stack aimed at detector developments)
  - Reconstruction at High Pileup (trackers and high granularity calorimeters)
  - Faster Simulation (machine learning algorithms for speed-up)
  - Analysis Facilities (new ROOT RDataFrame and object store integration)
- Work ramps up next year and will run until at least 2024

Summary

- Active programme of software R&D in EP-SFT and all experiments
- HSF increasing communication between experiments
  - Working groups are active and meet regularly
    - New frameworks group
  - Forum for exchange of ideas
- EP R&D Programme will support new software efforts for the future
  - AIDA++ also in preparation, can provide mutual support
  - HSF working groups inform and provide the wider context
- Future events:
  - Pre-CHEP workshop, 2-3 November 2019 *Analysis Systems: From Future Facilities to Final Plots*
  - CHEP in Adelaide, key event for the community
  - *Latin American Workshop on Software and Computing Challenges in HEP*, November 20-23, Mexico City
  - Next WLCG/HSF workshop in May 2020, 3 hosting proposals