Track 5 Highlights
Sustainable and Collaborative Software Engineering

Giulio Eulisse, Eduardo Rodrigues, Liz Sexton-Kennedy
Topics

Sustainable Languages and Architectures
Sustainable CI and Build Infrastructure
Glance and Web based applications
Sustainable Analysis
MISC: MonteCarlos, Infrastructure and Simulation
Sustainable Frameworks

A total of 33 talks in 6 sessions.

A Theme for this CHEP: Increasing adoption of open source tools outside of HEP & NP
Theme of this Track: Sustainability through sharing of common tools

Fri 12th May Plenary Session - Conference Highlights: Track 5
Interoperability between different languages and different hardware platforms was a key topic of this session.

Two nice talks about Julia being a sweet spot between C++ performance and Python ergonomics, including one with a concrete example implementing FastJet algorithms.

“Julia is impressive, it’s easy to work with and fast.” -Graeme Stewart
Sustainable Languages and Architectures

This session also had 2 talks about engaging and training the community:

**User-Centered Design for the Electron-Ion Collider**

- **Reality:** On average, 78% of students’ and postdocs’ research time is devoted to software and computing.
- **Goal:** Enable active participation in physics analysis, regardless of career stage, beyond just students and postdocs.
- **Solution via User-Centered Design:** Engage community in development. Listen to users, then develop software.
- Started projects on User-Centered Design in ePIC.
- User-Centered Design is part of our software principles.

**Paradigm**
- No standard curricula for HEP students exists
- Not all HEP students can attend university-offered software courses
- HEP students in many cases don’t receive any programming training
- Students trained as physicists but asked to be data analysts

We need a unified, scalable, and sustainable software training framework powered by the entire community

Experiments need Cyberinfrastructure professionals and lifelong learners

IRIS-HEP is generating a wealth of training materials in addition to organizing training events

**Scalability and Sustainability**
- Most training modules are website built from easy-to-read source files
- Complete video walkthroughs!
- Lessons build on each other
- Enough verbiage for self-study

IRIS-HEP is generating a wealth of training materials in addition to organizing training events

**Conference Highlights: Track 5**
Sustainable CI and Build Infrastructure

A common trend across the experiments is orchestration & containers using kubernetes or nomad.

Jenkins still heavily used.

Gitlab / Github actions new-ish way of defining builds.

ALICE is using Hashicorp Nomad, Vault and Cosul

Juno has modernized it's software practices following HSF recommendations

CERN web services have been automated and more configuration is delegated to users

Fri 12th May Plenary Session - Conference Highlights: Track 5
Glance and Web based applications

The Glance project is providing IT services specific to the needs of the experiments.

Sharing infrastructure allows improvements to propagate. We heard about - synchronization, search, and infrastructure improvements.
Sustainable Analysis

Coffea = user interface for columnar analysis:
- **dask_awkward** fundamentally changes how we can describe analysis
- **dask_awkward** based analyses, via dask task graphs, are rendered into a general, complete, declarative analysis description language (ADL)
- Represents the culmination of ~4 years of R&D

Law = Luigi Analysis Workflow, in Python
- Large Scale End-to-End Analysis Automation over Distributed Resources
- Designed to fully decouple these 3 aspects:

Fri 12th May Plenary Session - Conference Highlights: Track 5
Sustainable Analysis

“Analysis of physics analysis”

- Idea to learn from Git repositories. Study limited to CMS
  ▶ user adoption of new versions
  ▶ most common function-call patterns
  ▶ decide if and when a feature can be deprecated
  ▶ discover which libraries are being used together, maybe motivate integrations

LbMCSubmit:

- New flexible & scalable request submission system for LHCb simulation
  - Based on specifications in YAML, with review, CI test, then full submission

Simulation data quality in LHCb now integrated with DQ monitoring web-based tool

Fri 12th May Plenary Session - Conference Highlights: Track 5
MISC: MonteCarlo, Infrastructure and Simulation

- **EPOS 4 General-purpose event generator for hadronic physics out since Oct. 2022**
  - RIVET now integrated into the EPOS analysis system for effective testing and best parameters tuning (required in the validation of release cycles).
- **GEMC - Can we define a geant4 simulation in its entirety from a database?**
- **LinearCollider IO Event Data Model turns 20!**
  - Long-term goal was to encourage interoperability of worldwide LC simulation/reconstruction/analysis programs
  - LC community moving to key4hep and EDM4hep in an adiabatic way (LCIO<->EDM4hep conversion available)
- **Darshan is a lightweight I/O characterization tool that captures concise views and entire traces (DXT) of applications’ I/O behavior**
  - Enhanced to also work in non-MPI contexts (HEP use cases)
  - Being used by ATLAS/CMS to study I/O workflows

Fri 12th May Plenary Session - Conference Highlights: Track 5
MISC: MonteCarlos, Infrastructure and Simulation

- Gauss-on-Gaussino = new version of Gauss (Gauss = the LHCb simulation framework) based on Gaussino with LHCb-specific additions
  - Gaussino = new core simulation framework, created by extracting experiment-independent components from Gauss

Gaussino can be used (1) as the core simulation toolkit for detector frameworks in HEP or (2) in a standalone mode as an ideal test bed for early detector studies.
Sustainable Frameworks

Key4hep is a turnkey system providing components & building blocks for many efforts in the future collider world.

Meld: Exploring the feasibility of a framework-less framework - for DUNE

Functional programming approach to Frameworks.
Sustainable Frameworks

The Open Data Detector Project is trying to establish the next generation detector modeling tools for algorithmic and computing R&D.

Tracking/Calo challenges w/in the Key4Hep ecosystem.

Open-sourcing and Publication of the Belle 2 software stack.

Impact?

- No negative consequences observed so far
- Software citations
- Google Summer of Code
- Job application support

Fri 12th May Plenary Session - Conference Highlights: Track 5
A Big Thank you to the CHEP organizers and all the participants for making this CHEP a truly enjoyable experience; it’s great to be back together again!