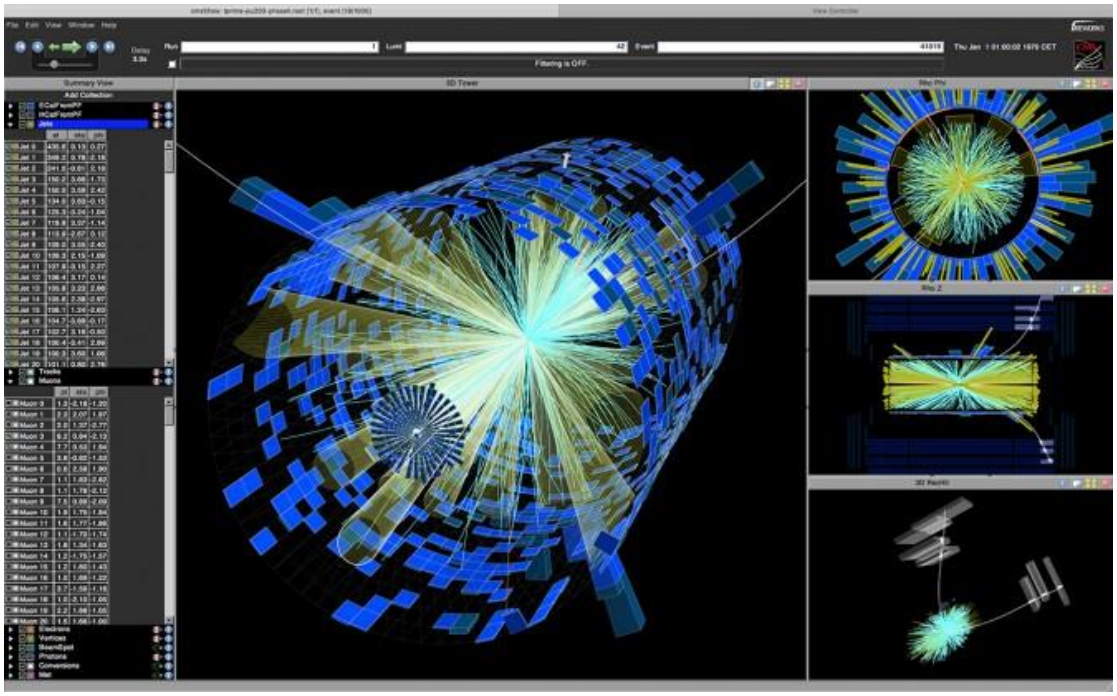
Optimizing Data Analysis Software for High Energy Physics

CAC

Project description

In Elementary Particle Physics, the key theory that has long guided the quest to understand the fundamental building blocks of nature and their interactions is the Standard Model of Particle Physics. The CMS and ATLAS experiments at CERN's Large Hadron Collider (LHC) took a huge step forward in this quest when they announced the detection of the Higgs boson, which won a Nobel prize for the theorists who had predicted it. The next major phase in CERN's evolution will be the High-Luminosity LHC (HL-LHC). Data collection is projected to start ~2026 and continue into the 2030's. The goal is to search for physics beyond the Standard Model by recording 10 times as much data from 100 times as many collisions as were used to discover the Higgs boson. This will require significant R&D advances in the software for acquiring, managing, processing and analyzing HL-LHC data to realize the scientific potential of the upgraded accelerator and detectors. The National Science Foundation awarded a five-year, $25 million grant to the Institute for Research and Innovation in Software for High Energy Physics (IRIS-HEP) to help in meeting these challenges. An IRIS-HEP research team at Cornell, led by physics professor and director of the Laboratory of Elementary-Particle Physics Peter Wittich, is responsible for developing parts of the high-performance analysis software that will be required to reduce "time to insight" and maximize the HL-LHC physics potential.



A simulation of data that would be obtained from the High-Luminosity Large Hadron Collider in Geneva, Switzerland. Cornell researchers, in collaboration with the Cornell Center for Advanced Computing (CAC), are part of IRIS-HEP, an NSF-funded project that is developing computer software to help analyze the data. Cornell is receiving $1 million for its role in IRIS-HEP, which is funded through August 2023.

CAC services

To develop computational software capable of handling the tsunami of data expected from the upgraded CERN, Peter Wittich decided to collaborate with Cornell's Center for Advanced Computing. Wittich, CAC senior research associate Steve Lantz, and project team members Dan Riley and Tres Reid are focused on improving the performance of a particular part of the data analysis software toolbox. "Specifically, the piece of software is the part that takes individual energy deposits in the detector and plays 'connect the dots' to reconstruct trajectories, or tracks, of the charged particles," says Wittich. "With this upgrade at LHC, there will be a large increase

CLIENT

Peter Wittich

Physics

SERVICES

- Code Parallelization
- Performance Analysis
- Data Analysis
- Algorithm Exploration on GPUs

in the number of proton-proton collisions per event – from 40 to around 200. That means many more particles flying around and more confusion in trying to get the trajectories understood." CAC's expertise in code parallelization and performance analysis have played a significant role in enhancing the software required for the high-luminosity era of the LHC. Lantz has run analyses of the group's tracking and clustering codes using Intel Advisor and VTune Profiler, and helped the team understand nuances of multithreaded and vectorized performance in Intel processors. He routinely collaborates on a variety of project-related tasks with researchers from the University of California, San Diego, Princeton, and Fermilab. Lantz also lectures regularly on vectorization techniques at an annual summer school for high energy physicists. Future plans include the exploration of novel particle-tracking algorithms on GPUs.

Results

The Cornell project team and their collaborators are in the process of having their main software product, dubbed "mkFit", integrated into the suite of data analysis tools that is available to CMS and CERN members worldwide. The code's methodology and physics validation results are detailed in the paper "Speeding up Particle Track Reconstruction using a Parallel Kalman Filter Algorithm," published in the *Journal of Instrumentation* (Lantz et al., 2020).

The High-Luminosity LHC project is led by CERN with the support of an international collaboration of 29 institutions, including the United States, Japan, and Canada. Peter Elmer, computational physicist at Princeton and CERN, and NSF's lead investigator for IRIS-HEP, says the Institute is an "intellectual hub" for software R&D that allows scientists to explore HL-LHC data and make new discoveries. Cornell is pleased to play its part in this endeavor. The LHC upgrade is slated for completion in 2026.

The LHC is the most powerful accelerator ever built. The accelerator sits in a tunnel 100 meters underground. The LHC pushes protons or ions to near the speed of light. It consists of a 17-mile ring of superconducting magnets with a number of accelerating structures that boost the energy of the particles along the way.

ABOUT CAC

The Cornell University Center for Advanced Computing is located in Ithaca, NY. With a professional staff of computational and data scientists, cyberinfrastructure specialists, and a breadth of PhD-level expertise in astronomy, biology, computer science, information science, mathematics, and physics, CAC is creating tomorrow's software, data and computing solutions today.

CAC INQUIRIES

Susan Mehringer
Associate Director for Consulting
Cornell University Center for Advanced Computing
shm7@cornell.edu
www.cac.cornell.edu

PHYSICS

Peter Wittich
Professor and Director, Laboratory of Elementary Particle Physics
Cornell University
wittich@cornell.edu

Portions of this case study were made available by Tom Fleischman, Cornell Chronicle

Copyright © 2021 Cornell University