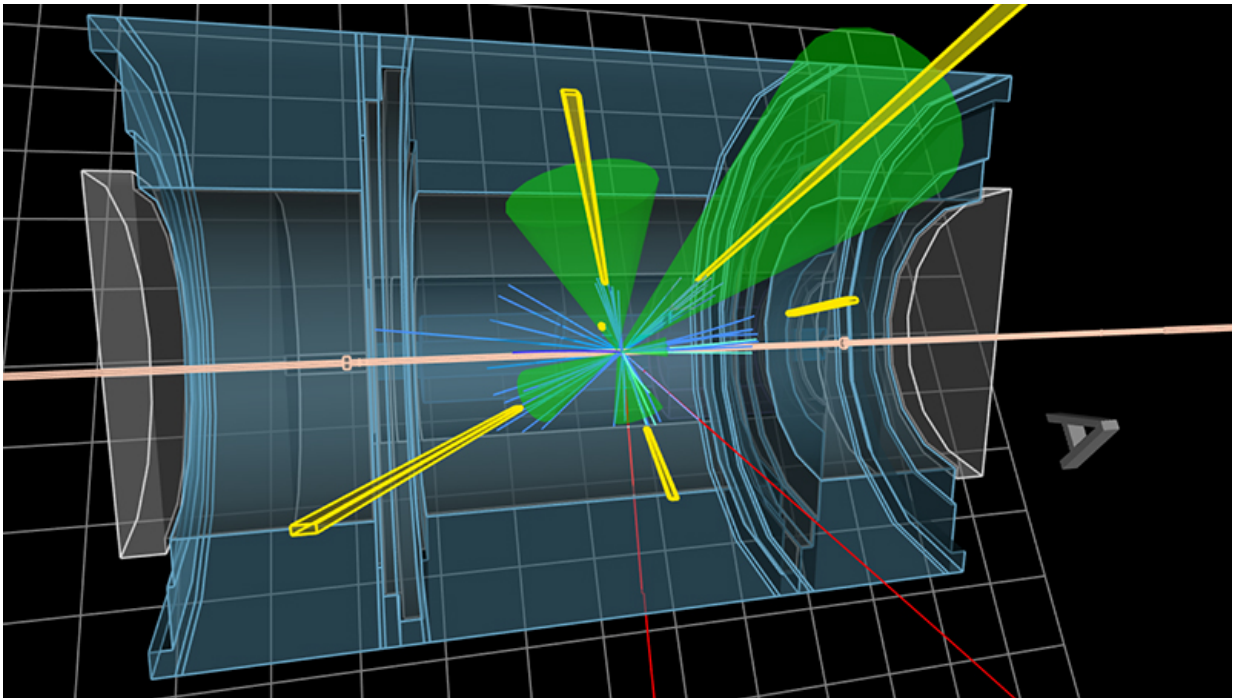


ALCF helps tackle the Large Hadron Collider's big data challenge

November 4 2015, by Jim Collins



A visualization of a simulated collision event in the ATLAS detector. This simulation, containing a Z boson and five hadronic jets, is an example of an event that is too complex to be simulated in bulk using ordinary PC-based computing grids.

Argonne physicists are using Mira to perform simulations of Large Hadron Collider (LHC) experiments with a leadership-class supercomputer for the first time, shedding light on a path forward for

interpreting future LHC data. Researchers at the Argonne Leadership Computing Facility (ALCF) helped the team optimize their code for the supercomputer, which has enabled them to simulate billions of particle collisions faster than ever before.

At CERN's Large Hadron Collider (LHC), the world's most powerful particle accelerator, scientists initiate millions of particle collisions every second in their quest to understand the fundamental structure of matter.

With each collision producing about a megabyte of data, the facility, located on the border of France and Switzerland, generates a colossal amount of data. Even after filtering out about 99 percent of it, scientists are left with around 30 petabytes (or 30 million gigabytes) each year to analyze for a wide range of physics experiments, including studies on the Higgs boson and dark matter.

To help tackle the considerable challenge of interpreting all this data, researchers from the U.S. Department of Energy's (DOE's) Argonne National Laboratory are demonstrating the potential of simulating collision events with Mira, a 10-petaflops IBM Blue Gene/Q supercomputer at the Argonne Leadership Computing Facility (ALCF), a DOE Office of Science User Facility.

"Simulating the collisions is critical to helping us understand the response of the particle detectors," said principal investigator Tom LeCompte, an Argonne physicist and the former physics coordinator for the LHC's ATLAS experiment, one of four particle detectors at the facility. "Differences between the simulated data and the experimental data can lead us to discover signs of new physics."

This marks the first time a leadership-class supercomputer has been used to perform massively parallel simulations of LHC collision events. The effort has been a great success thus far, showing that such

supercomputers can help drive future discoveries at the LHC by accelerating the pace at which simulated data can be produced. The project also demonstrates how leadership computing resources can be used to inform and facilitate other data-intensive [high energy physics](#) experiments.

Since 2002, LHC scientists have relied on the Worldwide LHC Computing Grid for all their data processing and simulation needs. Linking thousands of computers and storage systems across 41 countries, this international distributed computing infrastructure allows data to be accessed and analyzed in near real-time by an international community of more than 8,000 physicists collaborating among the four major LHC experiments.

"Grid computing has been very successful for LHC, but there are some limitations on the horizon," LeCompte said. "One is that some LHC event simulations are so complex that it would take weeks to complete them. Another is that the LHC's computing needs are set to grow by at least a factor of 10 in the next several years."

To investigate the use of supercomputers as a possible tool for the LHC, LeCompte applied for and received computing time at the ALCF through DOE's Advanced Scientific Computing Research Leadership Computing Challenge. His project is focused on simulating ATLAS events that are difficult to simulate with the computing grid.

While the LHC's big data challenge seems like a natural fit for one of the fastest supercomputers in the world, it took extensive work to adapt an existing LHC simulation method for Mira's massively parallel architecture.

With help from ALCF researchers Tom Uram, Hal Finkel, and Venkat Vishwanath, the Argonne team transformed ALPGEN, a Monte Carlo-

based application that generates events in hadronic collisions, from a single-threaded simulation code into massively multi-threaded code that could run efficiently on Mira. By improving the code's I/O performance and reducing its memory usage, they were able to scale ALPGEN to run on the full Mira system and help the code perform 23 times faster than it initially did. The code optimization work has enabled the team to routinely simulate millions of LHC collision events in parallel.

"By running these jobs on Mira, they completed two years' worth of ALPGEN simulations in a matter of weeks, and the LHC computing grid became correspondingly free to run other jobs," Uram said.

Throughout the course of the project, the team's simulations have equated to about 9 percent of the annual computing done by the ATLAS experiment. Ultimately, this effort is helping to accelerate the science that depends on these simulations.

"The datasets we've generated are important, and we would have made them anyway, but now we have them in our hands about a year and a half sooner," LeCompte said. "That, in turn, will help us get more results to conferences and publications at an earlier time."

As supercomputers like Mira get better integrated into the LHC's workflow, LeCompte believes a much larger fraction of simulations could eventually be shifted to high-performance computers. To help move the LHC in that direction, his team plans to increase the range of codes capable of running on Mira, with the next candidates being Sherpa, another event generation code, and Geant4, a code for simulating the passage of particles through matter.

"We also plan to help other high energy physics groups use leadership supercomputers like Mira," LeCompte said. "Our experience is that it takes a year or so to get to the minimum partition size, and another year

to run at scale."

Provided by Argonne National Laboratory

Citation: ALCF helps tackle the Large Hadron Collider's big data challenge (2015, November 4)
retrieved 25 April 2024 from

<https://phys.org/news/2015-11-alcf-tackle-large-hadron-collider.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.