Workload handling software has broad potential to maximize use of available supercomputing resources

July 8 2015, by Karen Mcnulty Walsh

The PanDA workload management system developed at Brookhaven Lab and the University of Texas, Arlington, has been integrated on the Titan supercomputer at the Oak Ridge Leadership Computing Facility at Oak Ridge National Laboratory.

With the successful restart of the Large Hadron Collider (LHC), now operating at nearly twice its former collision energy, comes an enormous increase in the volume of data physicists must sift through to search for new discoveries. Thanks to planning and a pilot project funded by the offices of Advanced Scientific Computing Research and High-Energy Physics within the Department of Energy's Office of Science, a

remarkable data-management tool developed by physicists at DOE's Brookhaven National Laboratory and the University of Texas at Arlington is evolving to meet the big-data challenge.

The workload management system, known as PanDA (for Production and Distributed Analysis), was designed by high-energy physicists to handle data analysis jobs for the LHC's ATLAS collaboration. During the LHC's first run, from 2010 to 2013, PanDA made ATLAS data available for analysis by 3000 scientists around the world using the LHC's global grid of networked computing resources. The latest rendition, known as Big PanDA, schedules jobs opportunistically on Titan—the world's most powerful supercomputer for open scientific research, located at the Oak Ridge Leadership Computing Facility (OLCF), a DOE Office of Science User Facility at Oak Ridge National Laboratory—in a manner that does not conflict with Titan's ability to schedule its traditional, very large, leadership-class computing jobs.

This integration of the workload management system on Titan—the first large-scale use of leadership class supercomputing facilities fully integrated with PanDA to assist in the analysis of experimental high-energy physics data—will have immediate benefits for ATLAS.

"Titan is ready to help with new discoveries at the LHC," said Brookhaven physicist Alexei Klimentov, a leader on the development of Big PanDA.

The workload management system will likely also help meet big data challenges in many areas of science by maximizing the use of limited supercomputing resources.

"As a DOE leadership computing facility, OLCF was designed to tackle large complex computing problems that cannot be readily performed using smaller facilities—things like modeling climate and nuclear

fusion," said Jack Wells, Director of Science for the National Center for Computational Science at ORNL. OLCF prioritizes the scheduling of these leadership jobs, which can take up 20, 60, or even greater than 90 percent of Titan's computational resources. One goal is to make the most of the available running time and get as close to 100 percent utilization of the system as possible.

"But even when Titan is fully loaded and large jobs are standing in the queue to run, we are typically using about 90 percent of the machine averaging over long periods of time," Wells said. "That means, on average, there's 10 percent of the machine that we are unable to use that could be made available to handle a mix of smaller jobs, essentially 'filling in the cracks' between the very large jobs."

As Klimentov explained, "Applications from high-energy physics don't require a huge allocation of resources on a supercomputer. If you imagine a glass filled with stones to represent the supercomputing capacity and how much 'space' is taken up by the big computing jobs, we use the small spaces between the stones."

A workload-management system like PanDA could help fill those spaces with other types of jobs as well.

New territory for experimental physicists

While supercomputers have been absolutely essential for the complex calculations of theoretical physics, distributed grid resources have been the workhorses for analyzing experimental high-energy physics data. PanDA, as designed by Kaushik De, a professor of physics at UT, Arlington, and Torre Wenaus of Brookhaven Lab, helped to integrate these worldwide computing centers by introducing common workflow protocols and access to the entire ATLAS data set.

But as the volume of data increases with the LHC collision energy, so does the need for running simulations that help scientists interpret their experimental results, Klimentov said. These simulations are perfectly suited for running on supercomputers, and Big PanDA makes it possible to do so without eating up valuable computing time.

The cutting-edge prototype Big PanDA software, which has been significantly modified from its original design, "backfills" simulations of the collisions taking place at the LHC into spaces between typically large supercomputing jobs.

"We can insert jobs at just the right time and in just the right size chunks so they can run without competing in any way with the mission leadership jobs, making use of computing power that would otherwise sit idle," Wells said.

In early June, as the LHC ramped up to 13 trillion electron volts of energy per proton, Titan ramped up to 10,000 core processing units (CPUs) simultaneously calculating LHC collisions, and has tested scalability successfully up to 90,000 concurrent cores.

"These simulations provide a clear path to understanding the complex physical phenomena recorded by the ATLAS detector," Klimentov said.

He noted that during one 10-day period just after the LHC restart, the group ran ATLAS simulations on Titan for 60,000 Titan core-hours in backfill mode. (30 Titan cores used over a period of one hour consume 30 Titan core-hours of computing resource.)

"This is a great achievement of the pilot program," said De of UT Arlington, co-leader of the Big PanDA project.

"We'll be able to reach far greater heights when the pilot matures into

daily operations at Titan in the next phase of this project," he added.

The Big PanDA team is now ready to bring its expertise to advancing the use of supercomputers for fields beyond high-energy physics. Already they have plans to use Big PanDA to help tackle the data challenges presented by the LHC's nuclear physics research using the ALICE detector—a program that complements the exploration of quark-gluon plasma and the building blocks of visible matter at Brookhaven's Relativistic Heavy Ion Collider (RHIC). But they see widespread applicability in other data-intensive fields, including molecular dynamics simulations and studies of genes and proteins in biology, the development of new energy technologies and materials design, and understanding global climate change.

"Our goal is to work with Jack and our other colleagues at OLCF to develop Big PanDA as a general workload tool available to all users of Titan and other supercomputers to advance fundamental discovery and understanding in a broad range of scientific and engineering disciplines," Klimentov said. Supercomputing groups in the Czech Republic, UK, and Switzerland have already been making inquiries.

Provided by Brookhaven National Laboratory