# New Chicago-Indiana computer network prepared to handle massive data flow

December 22 2006

Robert Gardner, Senior Research Associate in the University of Chicago's Computation Institute, stands in the blue glow of stacks of computer servers that are part of the MidWest Tier-2 Center, a joint effort with Indiana University. The Chicago-Indiana Tier-2 Center was created to help handle data flowing from the largest scientific experiment ever built, at CERN, the European particle physics laboratory in Geneva, Switzerland. Photo by Dan Dry

Massive quantities of data will soon begin flowing from the largest scientific instrument ever built into an international network of computer centers, including one operated jointly by the University of Chicago and Indiana University. The first phase of the Chicago-Indiana center, formally known as the MidWest Tier 2 Center, is now up and running, crunching test data in preparation for the real thing.

The Chicago-Indiana system is one of five Tier-2 (regional) centers in the United States that will receive data from one of four massive detectors at the Large Hadron Collider at CERN, the European particle physics laboratory in Geneva, Switzerland. When the new instrument begins operating late next year, beams of protons will collide 40 million times a second. When each of those proton beams reaches full intensity, each collision will produce approximately 23 interactions between protons that will create various types of subatomic particles.

"Understanding what's interesting and useful to record from those interactions is quite a challenge, because there is far more information than one is able to record for leisurely analysis," said James Pilcher, a Professor in Physics at the University of Chicago.

Frederick Luehring, a Senior Research Scientist at Indiana University, adds, "Even once the data is recorded, it will take years of careful sifting and sorting, which will require massive amounts of computing power to extract the final scientific results."

Pilcher and Luehring are among the physicists at 158 institutions in 35 nations who will harness the unprecedented power of the new collider in the ATLAS (A Toroidal Large Hadron Collider Apparatus) experiment at CERN. One of their goals will be to look for the long-sought Higgs boson, the theoretical particle that endows all objects in the universe with mass. The energy needed to create the Higgs boson is thought to be well within the capabilities of the Large Hadron Collider, Pilcher said. "If we don't see it, there's going to be a great deal of consternation," he said.

Another goal among physicists around the world is the search for evidence of supersymmetric particles, which could lead to the discovery of extra dimensions.

Physicists at Chicago and Indiana built components for the ATLAS particle detector with the search for the Higgs boson and supersymmetry in mind. The University of Chicago's Computation Institute, together with Indiana University's information technology services organization and Department of Physics, also collaborate on scientific grid computing projects that provide high-speed network computer power on demand, much the way a power grid provides electricity.

"In high-energy physics as in many disciplines, the computers and software used to analyze experimental data are now as vital to scientific success as the experimental apparatus that generate the data," said Ian Foster, director of the Computation Institute and a pioneer of grid computing. "This new Tier-2 center emphasizes the strengths that we have developed within the Computation Institute in creating and applying innovative computational infrastructure."

Luehring added, "Grid computing is the use of geographically distributed computing resources. Within ATLAS we have deliberately designed a tiered structure of computing resources spread throughout much of the world. All of these sites interconnect with each other using grid-computing techniques. In addition, grid-computing allows us to use other computing resources that are not fully dedicated to ATLAS or high-energy physics."

Data from the ATLAS experiment will first flow to Tier-0, the main computational center at CERN. Tier-0 will then transmit the data to 11 Tier-1 centers worldwide, including Brookhaven National Laboratory on Long Island, N.Y. Brookhaven will, in turn, distribute portions of the CERN data to the various Tier-2 centers.

The Chicago-Indiana Tier-2 center will serve physicists from around the world, said Robert Gardner, Senior Research Associate in the Computation Institute and the project's principal investigator. "It's really

driven not so much by where the physicists come from, but what their interests are," Gardner said. "Physicists will be able to submit jobs to this distributed network of centers and not worry about which center that their job is actually going to run on, because the data for their task will already be there," he said.

The Chicago-Indiana Tier 2 center is connected to a national computing infrastructure called the Open Science Grid, a national network dedicated to large-scale, computing-intensive research projects. "We run jobs from anyone who's participating in this Open Science Grid," Gardner said, whether the research involves particle physics, biology or some other topic.

Sites connected to the Open Science Grid include Fermi National Accelerator Laboratory in Illinois. Fermilab also is a Tier-1 center of the Large Hadron Collider's Compact Muon Solenoid experiment. As of Jan.1, 2007, Fermilab will be operated for the Department of Energy by Fermi Research Alliance, which consists of the University of Chicago and Universities Research Association Inc.

The initial set of computer servers, data storage and networking equipment of the multi-year project has been deployed in the basement of the Research Institutes building at the University of Chicago and at the Indianapolis campus of Indiana University, both of which will serve ATLAS data over the Open Science Grid. The sites will expand in January to bring the computing power equivalent to 300 personal computers to the national infrastructure via wide-area connections that can transfer data at 10 gigabits per second, which is like exchanging the music stored on an iPod in a second or two.

Nevertheless, the Tier-2 managers at both institutions regard the center as a single entity. "If our users apply for an account, we go through a security protocol that meets the common requirements for both

universities, nationally for the Open Science Grid and internationally for ATLAS, but we have them do that once. They don't have to do it twice," Luehring said. And if a hardware problem arises at Chicago, an Indiana technician may address the problem, or vice versa. It's a philosophy inspired by the culture of high-energy physics that physicists enjoy at CERN.

The Chicago-Indiana Tier-2 center is funded by annual $600,000 grants from the National Science Foundation. The project also was made possible by previous investments from the states of Illinois and Indiana in I-WIRE (Illinois Wired/Wireless Infrastructure for Research and Education), and I-Light (Indiana's high-speed fiber optic network for higher education and research).

"They're common infrastructure projects for the research community in the United States," Gardner said. "They're not necessarily for just one scientific purpose, but we're going to be early beneficiaries of these investments."

Source: University of Chicago