



[signal to background](#)

June 27, 2014

Getting the jump on big data for LSST

Efforts are already underway to ensure that the data the Large Synoptic Survey Telescope collects will be ready to be mined for scientific gold.

By [Lori Ann White](#)

On the first night the Large Synoptic Survey Telescope points its 8.4-meter mirror toward the exquisitely dark skies over Chile—probably in the year 2022—its 3.2-billion-pixel camera will record 30 trillion bytes of data about the contents of the universe. On the second night, LSST will do it again. It will do it again and again, collecting unprecedented amounts of data every clear night for 10 years.

By the end of its proposed mission, the LSST camera, designed and constructed at SLAC National Accelerator Laboratory, will have captured a full picture of the southern sky hundreds of times over.

Scientists around the world will search trillions of bytes of LSST data for information about the universe on all scales. They will look for asteroids in the Earth's backyard; map the Milky Way Galaxy; and study dark energy, the name given to whatever is causing the acceleration of the expansion of the entire universe.

But getting those terabytes of raw data from the camera processed, polished and to the researchers' computers in a usable form will be no small task. Cutting-edge computer applications will need to hold the data and mine it for scientific discoveries. These processing and database applications must work together flawlessly.

Jacek Becla, technology officer for scientific databases at SLAC, leads the group at SLAC constructing the LSST database. Their design recently passed a "stress test" intended to determine whether the software could put more resources to effective use as more was asked of it.

“We have a very solid prototype,” Becla says. “I’m actually quite confident we’ll be ready for LSST. We just have to stay focused.”

The LSST processing software, which is being developed by a collaboration led by the Associated Universities for Research in Astronomy, has also proven itself through an ongoing series of “data challenges.” In these challenges, the software is used to analyze data from previous astronomical studies, including nine years of data from the Sloan Digital Sky Survey and a total of 450 nights of data collected over five years by the Legacy Survey at the Canada-France-Hawaii Telescope. The results of the challenges are compared with results from the original surveys, which can highlight bugs and verify that the software does what it’s been written to do.

“These challenges have been very successful,” says LSST Director Steven Kahn. “They’ve already proved crucial algorithms are as good as—and in some cases better than—the software originally developed for the data.”

To help spread the wealth, scientists have made all LSST software open-source.

“The idea was to create software that’s available to the entire astrophysics community,” Kahn says. The Hyper Suprime-Cam, an 870-megapixel camera recently installed and commissioned on Japan’s Subaru Telescope, is already using an early version of LSST’s processing software.

Meanwhile, Becla wants the database technology to be available to anyone who can put it to good use. “There have already been a lot of inquiries about the software: from Germany, from Brazil, from the United Kingdom,” he says.

US financial support for the LSST construction comes from the National Science Foundation, the Department of Energy and private funding raised by the LSST Corporation, a non-profit 501(c)3 corporation formed in 2003, with its headquarters in Tucson, Arizona.

Kahn says he sees their work as an indication that the worlds of “big data” and “high performance”—or supercomputing—are converging.

“You need high-performance computing to run dark energy simulations; you have the big data you must compare the simulations to; and you have the big database to store the data,” he says. “LSST is a fantastic example.”

Like what you see? Sign up for a [free subscription](#) to symmetry!

Copyright © 2014 symmetry