NERSC Science Gateways: Deep Sky Provides a Portal into the Data Universe

BERKELEY LAB
LAWRENCE BERKELEY NATIONAL LABORATORY
Computing Sciences

Programmers at the Department of Energy's National Energy Scientific Research Computing Center (NERSC) are working with science users to design custom web browser interfaces and analytics tools, a service called "science gateways," which will make it easier for them to share their data with a larger community of researchers. DeepSky is just one of several science gateways at NERSC.

After a decade of archiving observations from the Palomar-QUEST Consortium and the Near Earth Asteroid team, NERSC holds over 14 million images covering more than 20,000 square degrees of sky, making this one of the largest troves of ground-based celestial images available.

The archive also contains 10-200 observations of particular portions of the sky with cadences ranging from minutes to years, giving astronomers a valuable tool for studying changes in the night sky.

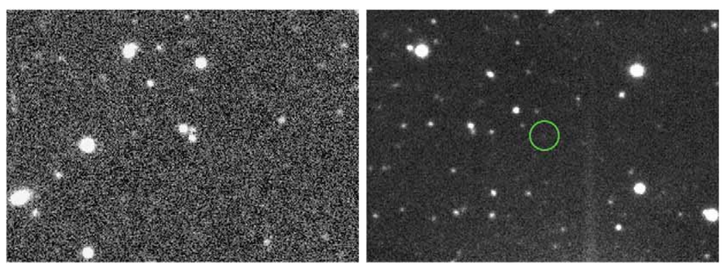
All of these images are now incorporated into the DeepSky Repository, which currently holds more than 14 million processed astronomical images on disks in the NERSC Global File System. These images, coupled with a database that kept track of the processing and image details (location in the sky, depth, quality, etc.) are now available through a variety of scripts on NERSC systems and soon will be available through a web interface to astronomers worldwide.

Case Study: Exposing Cosmic Transients with Deep Sky

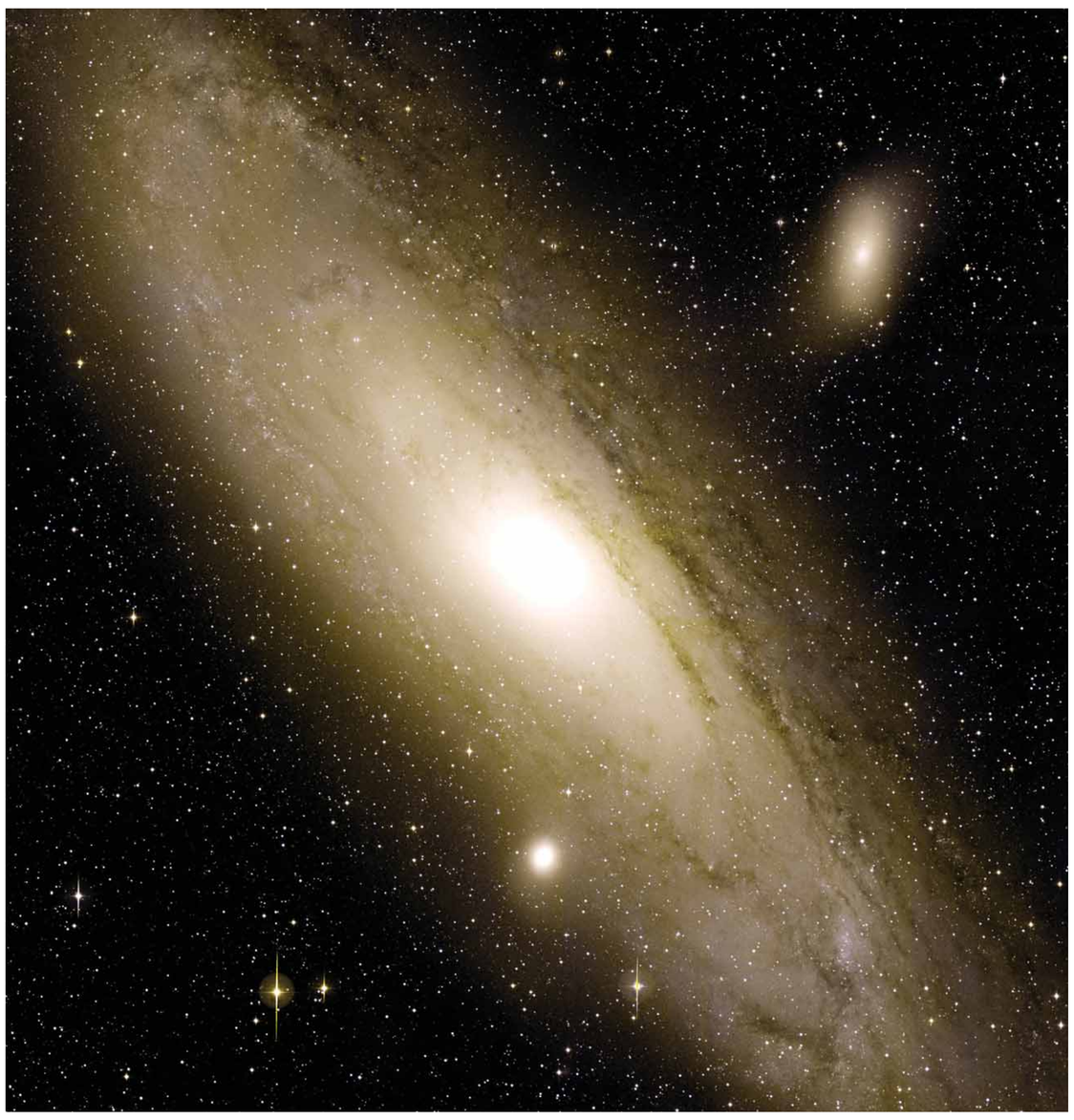
One project that is currently reaping enormous benefit from the DeepSky repository is the Palomar Transient Factory (PTF). This innovative survey combines the power of a wide-field telescope, a high-resolution camera, high-performance network and computing to identify candidate supernovae, active galactic nuclei (AGN) and variable stars.

Every night the PTF camera — a 100-megapixel machine mounted on the 48-inch Samuel Oschin Telescope at Palomar Observatory— automatically snaps pictures of the sky, then send those images to NERSC for archiving. Computers running machine-learning algorithms in the Real-time Transient Detection pipeline scour the PTF observations for sources that change in brightness or position, by comparing the new observations those in the NERSC archive. Once an interesting event is discovered, machines will send the coordinates to telescopes at Palomar and around the world for follow-up observations.

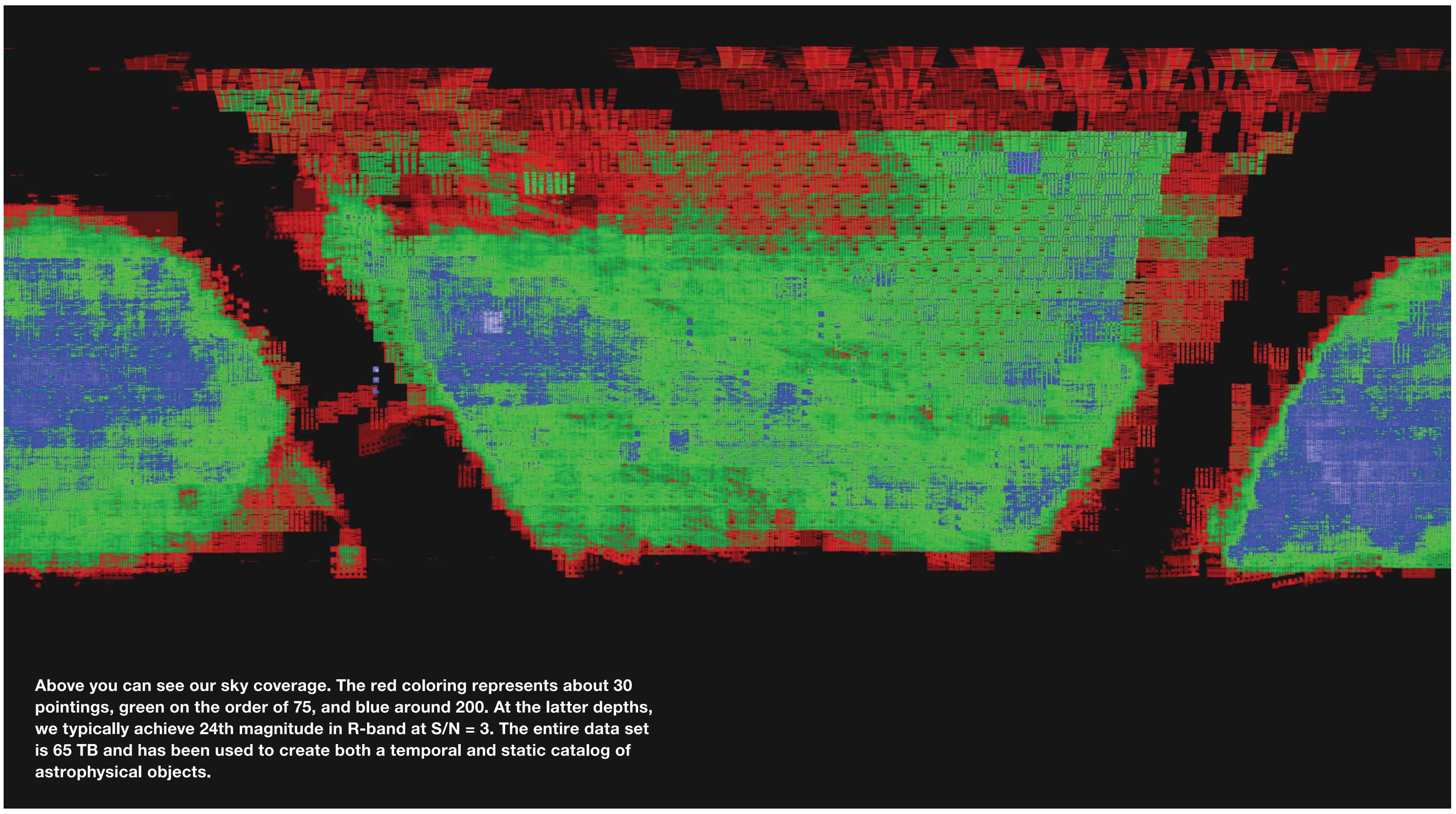
During the PTF commissioning phase alone, more than 51 transients were discovered with the help of DeepSky reference images. Since then, the team discovered 161 supernovae in just two months and over 5,000 variable stars. With a nine-year baseline of data, this repository has already proven to be an invaluable tool for identifying and classifying truly new transients.



The host galaxies of GRB071112c and the SNF20070825-001. Both are detected at 3-4 sigma at an R-band magnitude of 23.5.



This false color image of our nearest neighbor, the Andromeda galaxy, was created by layering over 400 images from observations made by PTF in February 2009. Extremely high quality images, such as this one, will be used as reference images from which new images taken later this year will be subtracted to find new transients. Fifty-one transients were detected in PTF commissioning data due to the depth of the reference images.



Above you can see our sky coverage. The red coloring represents about 30 pointings, green on the order of 75, and blue around 200. At the latter depths, we typically achieve 24th magnitude in R-band at S/N = 3. The entire data set is 65 TB and has been used to create both a temporal and static catalog of astrophysical objects.

U.S. DEPARTMENT OF ENERGY
Office of Science

CRD COMPUTATIONAL RESEARCH DIVISION

DEEP SKY

PALOMAR TRANSIENT FACTORY

NERSC NATIONAL ENERGY RESEARCH SCIENTIFIC COMPUTING CENTER