



The Growing NOAO Science Archive

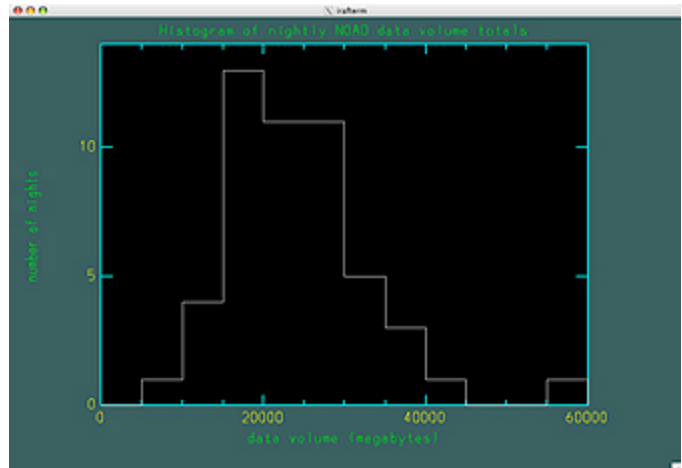
Rob Seaman & Irene Barg

Since the NOAO Science Archive first began operating in April 2002, many milestones have been achieved. NOAO Survey Program data now comprise 1.5 terabytes. These holdings represent one of the founding archives of the National Virtual Observatory, and the volume of NOAO Survey data is projected to grow by about one terabyte per year.

The NOAO Data Cache Initiative (DCI) has been an important transitional stage from the venerable Save the Bits (STB) safestore to the next-generation NOAO End-to-End system, which is currently being commissioned to manage our telescopes' raw and pipeline-processed data among our six geographically distributed sites. The DCI holdings contain data reaching back to August 2004, and number about 2.3 million compressed files and more than 19 terabytes. For comparison, the STB system utilizes 10,000 exabyte tapes that hold about five million images and 40 terabytes dating back to 1993.

Today's larger cameras generate 40 percent more images per year than those of a decade ago, with two to three times more pixels to store and serve. So while the NOAO archive has collected 60 terabytes of data over the past 15 years, we expect it to double in size over the next five years (just before the flood of pixels from the next generation of wide-field imagers).

With the commissioning of the NOAO Extremely Wide-Field Infrared Imager (NEWFIRM), the archive has seen new nightly records set with each new lunation. The histograms in the two figures below cover about a two-month sample from NOAO telescopes in both hemispheres. A "typical" night at NOAO adds about 2,500 images and 25,000 megabytes to the archive holdings, but with a large variance



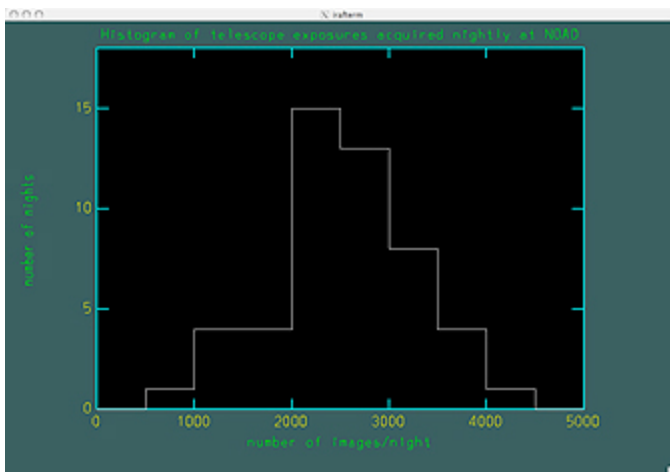
Histogram of nightly NOAO data volume totals

depending on the evolving instrumentation mix, age of the Moon, the weather, and, of course, on the particular observing programs scheduled.

The 2008A semester saw NOAO's busiest night yet at 56,027 megabytes. On a clear, dark night with both Mosaic cameras and NEWFIRM operating, this record may soon be eclipsed. When regular pipeline operations commence, the record will be obliterated. This is but a small taste of what is to come with the Dark Energy Camera at the Blanco telescope and the One-Degree Imager at WIYN.

These data volumes are reported as compressed using the familiar gzip algorithm. For NOAO data, this averages to a compression ratio of about 57 percent. NOAO has been evaluating FITS tile compression using the Rice algorithm as an alternative. Rice provides a compression ratio in the mid-40 percent range, and is a much faster algorithm. Most significant, the FITS tiling convention leaves the image headers readable. Software such as Bill Pence's CFITSIO library (see heasarc.gsfc.nasa.gov/docs/software/fitsio/fpack) can read and write such compressed images on the fly, raising the possibility that the first copy of the original data can be written already compressed and that the file need never be explicitly uncompressed afterward. IRAF support for Rice compression is being tested.

The latest milestone challenging the archive team is commissioning the complex and robust systems needed to automatically ingest these unrelenting streams of science metadata into searchable databases layered on Virtual Observatory (VO) technologies. This VO infrastructure permits the NOAO Portal to provide secure access to a principal investigator's proprietary data. Watch as we keep on growing to meet the need.



Histogram of telescope exposures acquired nightly at NOAO

Accessing Astronomical Data via NOAO

Christopher Miller

The NOAO Data Products Program End-to-End system (E2E) is maturing into a fully-deployed operational phase. This E2E system moves data from all NOAO telescopes (on three mountaintops on two continents) into a storage archive where the data can be queried electronically for pipeline processing and user access. The users' entry point into the E2E system is the NOAO National Virtual Observatory (NVO) Portal, through which they access the NOAO imaging data as well as data from astronomical archives distributed all over the world.

Specific NOAO principal investigators (PIs) were asked to help beta-test our E2E system by accessing their raw and pipeline-reduced Mosaic and NOAO Extremely Wide-Field Infrared Imager (NEWFIRM) data from the past two semesters. Their use of the system provided much-needed input on how to improve it. We at DPP would very much like to thank them for their efforts.

Throughout this testing process, DPP has continued to improve the archive and its ability to move and manage image metadata. The NOAO Mosaic and NEWFIRM pipelines continue to process nightly data, which is then provided to the PIs for scientific validation. Finally, the NOAO NVO Portal was recently upgraded and includes new features for the astronomer:

- Updated Virtual Observatory (VO) archive data (SDSS, GALEX, CHANDRA, XMM, HST, ESO INT, NOAO)
- The ability to upload coordinate lists and find VO image data at those locations

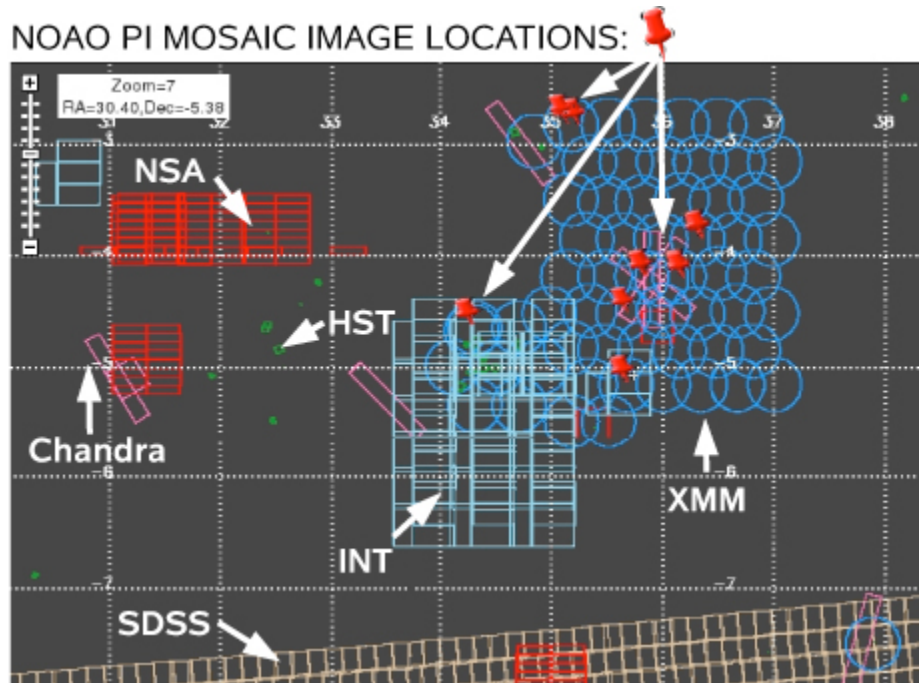


Figure: The NOAO NVO Portal showing a region of the sky containing VO data (HST, SDSS, XMM, Chandra, and NSA). The PI's data locations are shown as the pushpins.

- A new "Time View" to graphically see when data were taken
- A new "Query View" where users can query the ever-growing NOAO Archive (see article in this section)
- Safe and secure access to PI proprietary data
- No software to download or install—just open up your Firefox browser: www.nvo.noao.edu

In the near future, the E2E system will finish its "beta" phase and move into normal opera-

tions. When it is fully deployed, the community as well as PI astronomers will be able to discover and access raw and reduced data for the Mosaic and NEWFIRM instruments. As always, we are very much interested in your comments and suggestions (vohelp@noao.edu). If you have not already tried it, visit www.nvo.noao.edu and see what NOAO and the VO have to offer.