

DISPLAY

IEEE

# SPECTRUM

TOMORROW'S TECHNOLOGY TODAY

AUGUST 2004

www.spectrum.ieee.org

# THE UNRULY POWER GRID

STUDIES OF ELECTRICAL NETWORKS COME TO THE SAME CONCLUSION: BIG BLACKOUTS ARE INEVITABLE

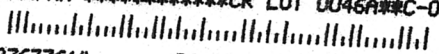
CHINA'S MOUNTING TECH TRASH PROBLEM P. 13

MORE COLORFUL COLOR TVS P. 17

DISPLAY UNTIL: 3  
www.ieee.org \$4.9



#BXNFWN \*\*\*\*\*CR LOT 0046A\*\*C-074



#28367761# 082004 500 61 IE08-2

SMITHSONIAN ASTROPHYS OBSER

LIBRARY MS 56

60 GARDEN ST

CAMBRIDGE MA 02138-1516

P000049

744



IEEE

# CONTENTS

## SPECIAL REPORT ASTRONOMY

ASTRONOMY AND COSMOLOGY are being transformed by computers. In this pair of articles, we consider two of the main reasons: a supercomputer simulation that aims to show how the universe evolved, and an Internet-based project to give researchers access to the entire spectrum of astronomical research.

### 28 Computing the Cosmos

The most detailed computer model of the universe ever made is helping researchers answer some of the most fundamental questions in cosmology.

By Alexander Hellemans & Madhusree Mukerjee

### 35 Downloading the Sky

The Virtual Observatory will let astronomers effortlessly pluck the exact information they need from the many trillions of bytes of research data. By Jonathan C. McDowell

**“Virgo Consortium’s goal is to simulate the evolution of the entire universe, from just after the Big Bang until now.”**

# 28

Cover Story

### 22 The Unruly Power Grid

Chaos theory and nonlinear dynamics suggest that widespread blackouts will be an inevitable fact of life. By Peter Fairley

Manufacturing

### 40 Supertubes

The amazing electrical and optical properties of carbon nanotubes promise fast, low-power integrated electronic and optoelectronic circuits.

By Phaedon Avouris

ON THE COVER:

Photography by Robert Llewellyn

Photo Illustration: John Ritter

This Page: NASA

Astronomers and computer scientists are building the world's best telescope—and it's all online  
BY JONATHAN C. MCDOWELL

# DOWNLOADING THE SKY

**THE YEAR IS 2007, AND I'M SITTING AT HOME**, drinking a cup of tea, and observing a galaxy millions of light-years away. No, I don't have a fortune in cutting-edge astronomical instruments, just a personal computer and a reasonably fast Internet connection.

Scattered across the screen is a handful of images, each showing that same galaxy but at a different wavelength. The visible-light image, a five-year-old photo from the twin Keck telescopes on Mauna Kea in Hawaii, shows the classic galactic pinwheel, spiral arms twisting out from a dense, starry center. In the infrared image, captured just a few seconds ago by a mountaintop telescope in Arizona, the galaxy looks more like a series of concentric rings, the telltale signs of dust-filled regions where stars are born. A radio image from a space-based telescope also shows a bright ring, but in this case it signifies the energy thrown off by countless exploding stars. Seen in the X-ray portion of the spectrum, the galaxy's rings are completely lost, replaced by a bright central core—probably a black hole.

As I superimpose the different images, I spot something peculiar: a faint, curved wisp of infrared gas next to a bright X-ray star. Zooming in, I realize that the shock wave from a supernova explosion has smashed into a gas cloud and triggered the formation of a batch of baby stars. My fingers tremble as I dash off a message to order up a new set of images....

NASA

It's all a dream now, unfortunately. When I pore over data on my computer nowadays, even at work, I see the same information I've been chewing over for weeks or months. No instant access to new data, no effortless comparing of multiple views of the universe. Though all those other images may exist in the public domain, they're stored away in vast databases at research institutions around the world, locked up in computers that speak different languages, use different data-storage formats, and even identify the same celestial bodies by different names. Getting those images takes days or weeks of fiddling and analysis—no astronomer can pull all those streams of data together in an easy way.

Soon, though, we'll be able to. An international collaboration of astronomers and computer scientists is now piecing together the means to connect all those dispersed stores of data—many trillions of bytes' worth, collected over the last several decades by hundreds of ground-based and orbiting observatories in thousands of archives. Their efforts will create, in effect, the world's biggest and best telescope. Known as the Virtual Observatory, or VO, it will allow astronomers, as well as students and the general public, to easily locate and download research data over the Internet. The VO will also serve as a grid computing network, giving researchers, regardless of location or resources, the equivalent of a supercomputer on their desktops, for comparing billion-record archives or running large-scale simulations [for more on cosmological simulations, see "Computing the Cosmos" in this issue].

The VO will transform how we view the universe. With our eyes, we can see only a tiny fraction of the light that makes up the night sky. But astronomical objects shine in every portion of the electromagnetic spectrum—optical, infrared, radio, X-ray, gamma ray, and more [see box, "The Spectrum of Astronomy"]. Each band of light reveals distinct physical processes. For example, infrared radiates from the cold gas and dust clouds around forming stars, while X-rays are generated by matter cooling in the fireball of a supernova. Only by fusing together these different clues can we get deep insights into the underlying processes driving our universe [see photos, "The All-Seeing Eye"].

Ultimately, the Virtual Observatory will alter the course of discovery. Astronomers will no longer be confined to working with one or two types of instruments, and they'll be freed from the tedious searching and gathering together of data that accompany current efforts. By allowing rapid comparisons of enormous quantities of disparate data, the VO will make it possible to get comprehensive views of large-scale processes at work in the universe, shedding light on some of the most fundamental questions: how did the universe evolve? When did the stars first form?

How many different kinds of galaxies are there? By giving researchers the means to comb quickly through enormous databases of images and catalogs and then compare the results, the observatory will also let them pinpoint rare events, such as the sudden, quick gamma-ray burst that occurs when certain stars die. Computer scientists are likewise betting that they can apply the cutting-edge technologies developed for the VO to other undertakings—from drug discovery to aerospace design—that require moving and manipulating huge amounts of data.

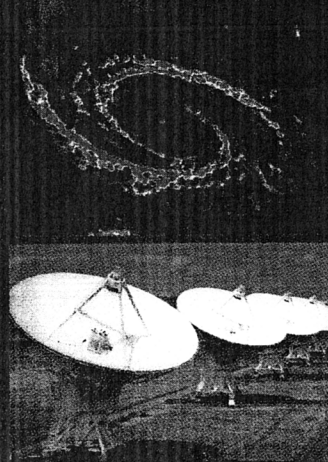
The VO encompasses a patchwork of projects organized under the International Virtual Observatory Alliance. The alliance includes more than 200 astronomers and computer scientists in at least 13 countries. In the United States, the VO effort is led by astronomers Alex Szalay at Johns Hopkins University in Baltimore and Roy Williams at the California Institute of Technology in Pasadena, assisted by computer scientist Jim Gray at the Microsoft Bay Area Research Center in San Francisco. The Europeans are led by Peter Quinn at the European Southern Observatory in Garching, Germany, and Françoise Genova at the Stellar Data Center in Strasbourg, France.

Through the VO's various working groups, the scientists are hammering out standards to make the archives interoperable, outlining the necessary IT infrastructure, and defining the VO's scientific goals. Compared with advanced astronomical instruments, which can cost several hundred million dollars to build and launch, the VO is operating on a shoestring: about US \$30 million over five years.

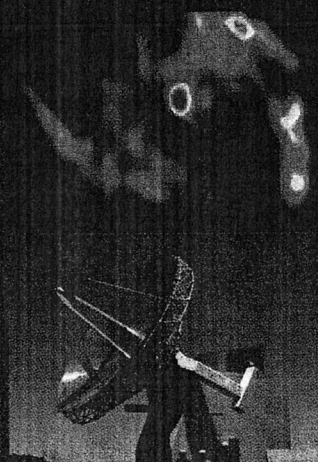
The main pieces of a working global system are expected to be in place within two years. An early demonstration offered a tantalizing glimpse of what's possible: in 2002, astronomer Bruce Berriman and colleagues at NASA's Infrared Processing and Analysis Center, based at Caltech, used a VO matching algorithm to com-

## THE SPECTRUM OF ASTRONOMY

Different kinds of light give us different clues about the universe and call for different kinds of telescopes. The Virtual Observatory will link all these types of data.



**RADIO WAVES**, the longest wavelength, are detected by large ground-based dishes like the Very Large Array in New Mexico. The radio sky is dominated by gas clouds in our galaxy and by the amazing relativistic jets of matter escaping from quasars in distant galaxies.



**SUBMILLIMETER RADIATION**, a new domain for astronomers, is studied from satellites and from the highest, driest spots on Earth, like the Submillimeter Array on the summit of Hawaii's Mauna Kea. In this band, astronomers study complex molecules in the coldest, darkest clouds in our galaxy, as well as the faint glow of the Big Bang.



**FAR INFRARED LIGHT** can be seen only from space observatories like the recently launched Spitzer Space Telescope. Most far infrared sources are glowing dust grains clustered around the birthplaces of stars and planets.

pare tens of millions of entries from two of the larger data archives, the visible-light Sloan Digital Sky Survey and the infrared Two Micron All Sky Survey. By looking for objects that are bright in the infrared but invisible in the optical, the hallmark of the elusive brown dwarf star, they quickly narrowed down the list to several hundred thousand objects and then to just a handful. When the search was finished, they'd identified a new brown dwarf.

Szalay, whose specialty is galaxy formation and who helped design the database architecture for the Sloan Digital Sky Survey, points to the VO as evidence of a broader trend now reshaping science. "Traditionally, science was entirely phenomenological and descriptive," he says. "Now, the quantity of scientific data is so enormous that dealing with data is a whole new discipline in itself—this is happening in every branch of science. You need to combine information management, computer science, new statistical approaches, and your own domain-specific expertise, whether that's astronomy, or genomics, or oceanography, or business."

**ASTRONOMERS ARE DROWNING IN DATA.** To take one example, the ambitious Sloan Digital Sky Survey is using ground-based telescopes equipped with digital cameras to record a quarter of the sky with unprecedented accuracy and depth; its latest release of images and related data totaled six terabytes. The newest catalog of star positions, published by the U.S. Naval Observatory in Washington, D.C., contains more than a billion stars, while a single observation from the Hubble Space Telescope can easily swallow several gigabytes. In the next few years, as the archives from these and other instruments continue to swell and a number of new large-scale projects come online, the total amount of data is expected to double every year or two. Not surprisingly, astronomers are already worried about how

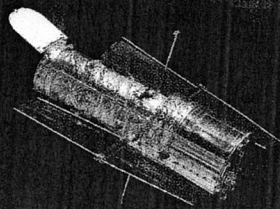
to find what they need. Similarly crushing tides of data await lots of other people in other fields, whether they work in multinational corporations or in large, geographically dispersed research projects.

This abundance of astronomy data is actually fairly new. Decades ago, observations were made for a project and then thrown away—if you wanted to ask another question about the same star or galaxy, you went back to the telescope. The launch of the first space telescopes in the 1970s, such as the Einstein Observatory and the International Ultraviolet Explorer, changed that. Their high cost convinced researchers, and funders, that the data were too precious to lose.

In the process, archival astronomy was born. Researchers quickly learned that data gathered for one star could be reused to study other stars; hundreds of papers were generated in this way. Nowadays, astronomers, more than other scientists, tend to share their data. Those who observe on a NASA space facility, for example, get exclusive access to their data for only one year—after that, anyone can download it and look for discoveries that the original observer may have missed. [For a glimpse of how modern astronomy is done, and how the VO might help, see box, "An Astronomer's Life."]

To take the next step and make any data set available to any astronomer anywhere in the world will mean solving a number of major challenges. These include resolving differences in data format, defining a query language for accessing the VO, creating the computational infrastructure, figuring out how to keep the VO up to date as new data sets are created, and, of course, getting all the players to agree on the many software standards and protocols.

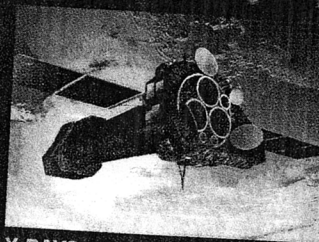
**ASTRONOMICAL DATA TAKE MANY FORMS,** depending on the instrument that collected them and the format and medium they were stored in. Some records aren't digitized; a lot of radio



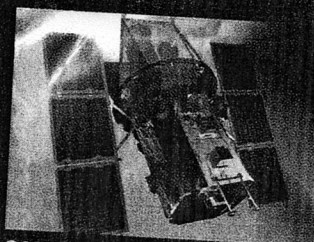
**VISIBLE, or OPTICAL, and NEAR INFRARED LIGHT** are the traditional meat of astronomy. Our eyes evolved to see the light from the sun that can traverse the Earth's atmosphere; on mountaintops above the clouds, modern detectors can see a little way into the infrared as well. In these wave bands, mainstream stars and gas clouds reveal their glory in images from Hawaii's Keck telescopes and the Hubble Space Telescope, as well as via binoculars from your backyard.



**ULTRAVIOLET LIGHT** forms the signature of hot, young stars and the violent winds flowing from them. The beautiful new images from NASA's small GALEX satellite reveal the ultraviolet skeletons of galaxies, which will be merged with Spitzer's infrared view to complete our picture of star formation and evolution.



**X-RAYS** are light that is beyond ultraviolet. Each X-ray photon packs hundreds of times the punch of the ultraviolet radiation that gives you a sunburn, but Earth's atmosphere shields us from them. Space telescopes like NASA's Chandra and the European Space Agency's XMM-Newton provide access to the X-ray sky, which shines with the light of exotic objects such as exploding stars and star-swallowing black holes.



**GAMMA RAYS** are the high-energy frontier, with satellites, balloons, and ground-based telescopes beginning to reveal their secrets. Such gamma-ray eyes are still rather fuzzy, but the mysterious gamma-ray bursts, the most violent explosions in the universe, are motivating rapid advances in the field. Stay tuned for the launch of the Swift gamma-ray burst satellite in September.

# AN ASTRONOMER'S LIFE

Here's how a typical astronomer's research life goes (using mine as an example): The first thing I do is think of a question that's worth answering—say, when two galaxies collide and merge, it triggers an enormous burst of newly formed stars in the center, but what's happening on the outskirts of the collision? Next I find what other astronomers have already figured out about this or similar problems.

I then devise an experiment that might answer the question; the assumption is that if I look at this particular collision of galaxies, I can learn about all colliding galaxies. I plan carefully what data I want to collect, and then I request some time on a ground-based telescope (in which case I head up to an Arizona mountaintop with a week's worth of

groceries) or ask NASA to point one of its space-based telescopes at my chosen patch of sky.

I spend the next six months analyzing those few nights of data. If they match what I'd expected, that's good, but I still have to spend a few months figuring out how I might have messed up. The chances of finding something really unusual are small: nine times out of ten, you're just mistaken. So a good astronomer is paranoid: the universe is lying to me, my telescope is lying to me, my software is lying to me. Was there a bug in the software? Did the CCD camera on the telescope have hot pixels that created bright spots in the image? Is the thing I think I saw in the sky just an artifact created by a nearby star passing in front of the galaxy? Once I've ruled all these things out, I pub-

lish the results in enough detail that my colleagues either are convinced that I'm right or can go out and try to prove that I'm wrong.

The Virtual Observatory will help enormously at both the front and back ends of this process. It will make it relatively trivial to find out what other research has been done, so that I have a better idea what kinds of new observations I need to make. And it will also allow me to compare the data I've gathered against other types of data, the better to explain the phenomenon I've captured. Sadly, it won't make me any faster at writing papers, but it may well prevent getting that dreaded note from a referee, informing me that I could've just as easily answered my question using existing data from such-and-such obscure satellite.

—J.C.M.

astronomy data, for example, are still on analog nine-track tape. Some smaller observatories don't even archive their data; instead, researchers take home whatever raw data they collect. VO collaborators hope that as the virtual observatory comes online and begins to prove its worth, the data laggards will devote the resources necessary to create or upgrade their databases.

Further complicating things is that different archives can refer to the same object by different names. The International Astronomical Union, based in Paris, oversees the naming of celestial objects (and no, you can't pay to have a star named after you), but that doesn't prevent other unsanctioned designations from popping up. So when comparing astronomical catalogs covering two types of wavelengths, researchers must also typically check an object's position. Such double-checking fails, however, if the object in question is visible (and therefore recorded) only at one of the wavelengths.

Tracking down data sets, which can take weeks or even months, became somewhat more straightforward in 1996, with the creation of the online SIMBAD service. Run by the Stellar Data Center, SIMBAD lets researchers call up a list of papers that cite a celestial object, plus its other names, its position, and a few other numbers. What SIMBAD doesn't tell you is where the data are archived; nor does it return actual data with which you can do actual science.

For that, you need the VO. With it, a user in Chicago will be able to sit at her computer, type in a data request—say, all the brightness information at all different colors of the spectrum for quasar PG1407+265—and then wait for the data to come in. Behind the scenes, her query may be processed by a Web portal at Caltech, which in turn searches several archives, including one based in Strasbourg that lists star locations and another in Cambridge, Mass., that knows the stars' X-ray intensities. After the searches, the Web portal gathers up all the results and replies to the user.

As this scenario suggests, the Virtual Observatory is a distributed system, much like the Internet itself. To link its disparate parts—to "federate" them, as astronomers say—the VO is being built around registries. A registry is basically an online catalog of what is in each archive, indexing the virtual sky by position and wave band; it is continually updated to incorporate new data and new archives. Functionally, a VO registry is like the domain name

servers that point to things on the Internet. Prototype VO registries are already running at Johns Hopkins, the University of Illinois, and Caltech. The Data Inventory Service, created by Thomas McGlynn and colleagues at NASA Goddard Space Flight Center, calls on these registries (and eventually others) to locate data based on an object's position or name (see <http://heasarc.gsfc.nasa.gov/vo/>).

The VO registries will also point to Web-based programs, known as Web services, which will allow data from those archives to be processed. Astronomers already use various software tools for analyzing and filtering their data, but such programs are designed to run on local workstations, using locally stored data. A Web service, by contrast, is accessed through the Internet, and the user may not even know it is running. The VOStat service, for example, lets users run many types of statistical routines on their data; the user doesn't need to worry about having the latest statistical software.

Sifting through these disparate databases is eased by past attempts at data standardization, such as the Flexible Image Transport System (FITS) format. Invented by radio astronomers back in the 1970s to exchange data on magnetic tape, FITS has since been widely adopted by other astronomers, and FITS files can now be read by almost all astronomical software.

But FITS typically can't be read by mainstream software. The VO team therefore plans to supplement FITS files with eXtensible Markup Language descriptions of the data. Although XML is fast becoming the common text format for exchanging a wide variety of data on the Web and elsewhere, astronomers have been relatively late to embrace it. The VO's first use of XML is the VOTable format, developed by groups at the Stellar Data Center and Caltech, for exchanging tables and star catalogs.

Another problem with FITS is that it allows each group to make up the keywords that describe what the file contains; uninitiated astronomers have no way of deciphering these custom keywords. What's needed is a precise and universal vocabulary. For example, if I ask the VO for data about photon frequencies, I don't want data about stellar pulsation frequencies. The UCD (or Unified Content Descriptor), invented by the star-catalog experts at Strasbourg, is a first cut at defining such an unambiguous vocabulary. Initially, the VO will use UCDs to

render the software unable to process the data, leaving the user with no means to generate images at all.

**BEYOND FACILITATING DATA SEARCHES**, the Virtual Observatory will allow users to do large-scale computations from their desktops. Grid computing has been gaining momentum in scientific circles for years [see, for example, "Super Nets for Supercomputers," *IEEE Spectrum*, January 2001], but the VO Grid, envisioned as a sort of World Wide Web for computation, will push its technological limits. The grid will allow astronomers to run big programs—comparing catalogs of billions of stars or running multiple cosmic simulations in parallel—by yoking together the computing power of scattered, distant machines.

The VO Grid will be built around existing facilities like the TeraGrid, a National Science Foundation-funded grid network created for academic researchers. The TeraGrid currently links supercomputers and related apparatus at nine U.S. sites, and it boasts a 6-petabyte magnetic tape archive, 1 PB of hard-disk storage, and 20 teraflops of computing power (that's 20 trillion calculations per second), soon to be doubled to 40 teraflops. The TeraGrid also has its own optical-fiber network, which can move 22 terabytes a day between sites and up to 430 TB a day within a single site.

Properly harnessed, the storage and processing power of the TeraGrid will let researchers analyze an entire 10-TB image survey (the data equivalent of the entire print collection of the U.S. Library of Congress) in just a few hours. An astronomer could rapidly determine how many galaxies are likely to host massive black holes at their centers by simply specifying the characteristic features of such a galaxy. Obtaining the answer to such a question would actually involve running a sequence of browsing and analysis programs—the Web services described earlier—residing at different sites across the Internet, and drawing on data sets stored at various data centers. But the user wouldn't need to know which machines were doing the work or whether the data were stored on several continents. Only the results—a list of possible galaxies—would be delivered to the astronomer, who could then refine the search by sending out further queries across the Web.

**ASTRONOMY HAS BEEN AT THE VANGUARD** of many advances in scientific technique, going back to Newton's calculus and continuing, more recently, with the discovery of new elements in stellar spectra. Over the next several years, as the Virtual Observatory begins to roll out more standards and Web services, it promises to open a new era in data exploration. If it is successful and becomes ubiquitous, the VO will continue to evolve, much as the Web continues to evolve through the World Wide Web Consortium.

Johns Hopkins's Szalay has predicted that the VO will have a democratizing effect on astronomy, allowing researchers from poorly funded organizations and countries to access the same data and computing resources that their richer colleagues now enjoy. Of course, the Virtual Observatory doesn't guarantee good science—it just makes it much, much easier to do. And the results will benefit not just astronomy but all of science and beyond.

**TO PROBE FURTHER**

The International Virtual Observatory Alliance's Web site, <http://www.ivoa.net/>, has information about the project and a list of participants. The U.S. National Virtual Observatory Web site, <http://us-vo.org/>, also describes and gives links to projects under development within the VO. A list of Web services currently offered by the VO is at <http://voservices.org/>.

For information about grid computing, see the TeraGrid's Web site, <http://www.teragrid.org>.



**THE ALL-SEEING EYE:** A composite image of the galaxy Centaurus A [shown at top] suggests a titanic explosion occurred there 10 million years ago. This image, which combines X-ray [blue], optical [yellow], and radio [pink and green] data, took Margarita Karovska of the Harvard-Smithsonian Center for Astrophysics and collaborators several days to create. The Virtual Observatory will be able to locate the relevant data and generate such images on the fly.

augment FITS keywords; eventually, they could become the sole means of describing a file's contents.

**MANY OF THE VO ASTRONOMERS** pride themselves on their computer savvy. Even so, they can be overwhelmed by the latest software techniques and jargon. VO computer scientists are equally lost in the zoo of celestial fauna, which includes such exotica as exoplanets, magnetars, and superclusters, to say nothing of the arcana of astronomical instrumentation. An "object" to astronomers may be an enormous physical thing in outer space, but to computer scientists familiar with "object-oriented design" it is an abstraction describing a concept in software.

A more serious ongoing debate revolves around how much and what kinds of new computing techniques to incorporate into the VO; the computer scientists lean toward the most cutting-edge technologies, whereas the astronomers worry whether the new ways will be useful and stable in the long run.

One such argument involves the use of so-called virtual data. At present, most astronomy archives store their data as calibrated images; the calibration takes into account deviations introduced by the instrumentation—hot pixels on the charge-coupled device (CCD) camera chip, say. With a virtual data system, archives would store only raw, uncalibrated data; each time a user would ask for a particular image, the data would be processed and calibrated, and the image created on the fly. Virtual data have the advantage of taking up less storage space and being easier to archive. On the other hand, such a system is fragile—a hardware change or failure may

ALL IMAGES: NASA