The New York Times https://www.nytimes.com/2011/04/26/science/26planetarium.html

Digging Deeper, Seeing Farther: Supercomputers Alter Science

By John Markoff
April 25, 2011

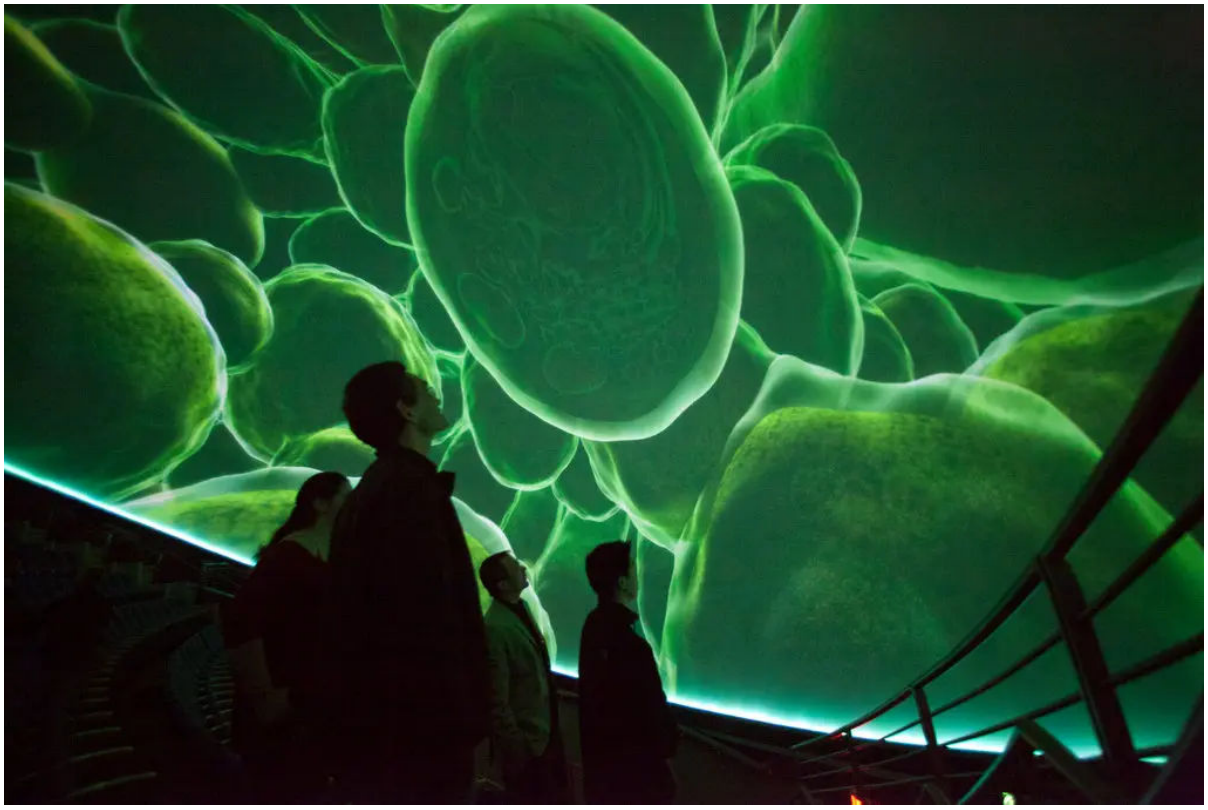
SAN FRANCISCO — Inside a darkened theater a viewer floats in a redwood forest displayed with Imax-like clarity on a cavernous overhead screen.

The hovering sensation gives way to vertigo as the camera dives deeper into the forest, approaches a branch of a giant redwood tree, and then plunges first into a single leaf and then into an individual cell. Inside the cell the scene is evocative of the 1966 science fiction movie “Fantastic Voyage,” in which Lilliputian humans in a minuscule capsule take a medical journey through a human body.

There is an important difference — “Life: A Cosmic Journey,” a multimedia presentation now showing at the new Morrison Planetarium here at the California Academy of Sciences, relies not just on computer animation techniques, but on a wealth of digitized scientific data as well.

The planetarium show is a visually spectacular demonstration of the way computer power is transforming the sciences, giving scientists tools as important to current research as the microscope and telescope were to earlier scientists. Their use accompanies a fundamental change in the material that scientists study. Individual specimens, whether fossils, living organisms or cells, were once the substrate of discovery. Now, to an ever greater extent, researchers work with immense collections of digital data, and the mastery of such mountains of information depends on computing power.

The physical technology of scientific research is still here — the new electron microscopes, the telescopes, the particle colliders — but they are now inseparable from computing power, and it is the computers that let scientists find order and patterns in the raw information that the physical tools gather.



DUET "Life: A Cosmic Journey" relies not just on computer animation techniques, but on a wealth of digitized scientific data as well. Jim Wilson/The New York Times

Computer power not only aids research, it defines the nature of that research: what can be studied, what new questions can be asked, and answered.

“The profound thing is that today all scientific instruments have computing intelligence inside, and that’s a huge change,” said Larry Smarr, an astrophysicist who is director of the California Institute for Telecommunications and Information Technology, or Calit2, a research consortium at the University California, San Diego.

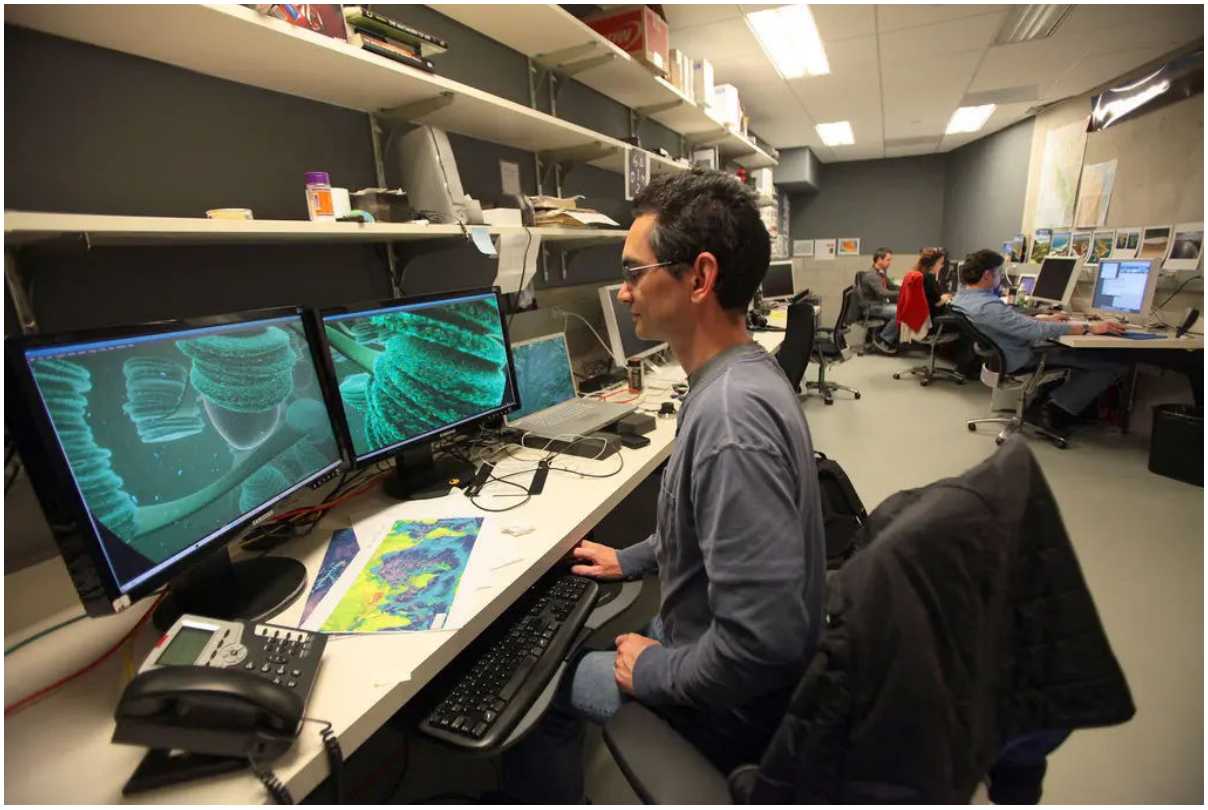
In the planetarium’s first production, “Fragile Planet,” the viewer was transported through the roof of the Morrison, first appearing to fly in a graceful arc around the Renzo Piano-designed museum and then quickly out into the solar system to explore the cosmos. Where visual imagery was once projected on the dome of the original Morrison Planetarium using an elaborate home-brew star projector, the new system is powered by three separate parallel computing systems which store so much data that the system is both telescope and microscope. From incomprehensibly small to unimaginably large, the computerized planetarium moves seamlessly over 12 orders of magnitude in the objects it presents. It can shift “from subatomic to the large-scale structure of the universe,” said Ryan Wyatt, an astronomer who is director of the planetarium.

It is, said Katy Börner, an Indiana University computer scientist who is a specialist in

scientific visualization, a “macroscope.” She uses the word to describe a new class of computer-based scientific instruments to which the new planetarium’s virtual and physical machine belongs. These are composite tools, with different kinds of physical presences that have such powerful and flexible software programs that they become a complete scientific workbench that can be reconfigured by mixing and matching aspects of the software to tackle specific research problems.

The planetarium’s macroscope is designed for education, but it could be used for research. Like any macroscope, its essence is its capacity for approaching huge databases in a variety of ways. “Macroscopes provide a ‘vision of the whole,’ ” Dr. Börner wrote in the March issue of The Communications of the Association for Computing Machinery, “helping us ‘synthesize’ the related elements and detect patterns, trends and outliers while granting access to myriad details.’ ” She said software-based scientific instruments are making it possible to uncover phenomena and processes that in the past have been, “too great, slow or complex for the human eye and mind to notice and comprehend.”

Computing is reshaping scientific research in a number of ways, Dr. Börner notes. For example, independent scientists have increasingly given way to research teams as cited by scientific papers in the field of high-energy physics that routinely have hundreds or even thousands of authors. It is unsurprising, in a way, since the Web was invented as a collaboration tool for the high-energy physics community at CERN, the European nuclear research laboratory, in the early 1990s. As a result research teams in all scientific disciplines are increasingly both interdisciplinary and widely distributed geographically.



The visualization studio in the basement beneath the planetarium. Jim Wilson/The New York Times

So-called Web 2.0 software, with its seamless linking of applications, has made it easier to share research findings, and that in turn has led to an explosion of collaborative efforts. It has also accelerated the range of cross-disciplinary projects as it has become easier to repurpose and combine software-based techniques ranging from analytical tools to utilities for exporting and importing data.

A macroscope need not be in a single physical location. To take one example, a midday visitor to the lab of Tom DeFanti, a computer graphics specialist, in the Calit2 building in San Diego is greeted by a wall-size array of screens that appears to offer a high-resolution window into a vacant laboratory somewhere else in the world. The distant room is a parallel laboratory at King Abdullah University of Science and Technology, in Thuwal, Saudi Arabia. Four years ago representatives of that university visited Calit2 and initiated a collaboration in which the American scientists helped create a parallel scientific visualization center in Thuwal connected to the Internet by up to 10 gigabits of bandwidth — enough to share high-resolution imagery and research.

Saudi researchers now have access to a software system known as Scalable Adaptive Graphics Environment, or SAGE, originally developed to permit scientists working far apart to share and visualize research data. SAGE is essentially an operating system for visual information, capable of displaying and manipulating images up to about one-third

of a billion pixels — as much as 150 times more than what can be displayed on a conventional computer display.

“The killer application is collaboration; that is what people want,” Dr. DeFanti said. “You can save so much energy by not flying to London that it will run a rack of computers for a year.”

More than a decade ago Dr. Smarr began building a distributed supercomputing capability he called the OptIPuter, because it used the fiber-optic links among the nation’s supercomputer centers to make it possible to divide computing problems as well as digital data so that larger scientific computing loads could be shared.

The advent of high-performance computing systems, however, created a new bottleneck for scientists, he said. “Over the past decade computers have become over a thousand times faster because of Moore’s Law and the ability to store information has gone up roughly 10,000 times, while the number of pixels we can display is maybe only a factor of two different,” he said.



SEAMLESS From a basement bank of servers, Cheryl Vanderbilt guides the planetarium display through over 12 orders of magnitude in the object it presents. Jim Wilson/The New York Times

To make it possible for visualization to catch up with accelerating computing capacity, researchers at Calit2 and others have begun designing display systems called OptIPortals

that offer better ways of representing scientific data.

Recently, the Calit2 researchers have begun building scaled-down versions called OptIPortables, which are smaller display systems that can be fashioned like Lego blocks from just a handful of displays, rather than dozens or hundreds. The OptIPortable displays can be quickly set up and moved, and Dr. DeFanti said his lab was now at capacity assembling systems for research groups around the world.

Within many scientific fields software-based instruments are quickly adding new functions as open-source systems make it possible for small groups or even individuals to add features that permit customization.

Cytoscape is a bioinformatics software tool set that evolved, beginning in 2001, from research in the laboratory of Leroy Hood at the University of Washington. Dr. Hood, one of the founders of the Institute for Systems Biology in Seattle, was a pioneer in the field of automated gene sequencing, and one of his graduate students at the time, Trey Ideker, was exploring whether it was possible to automate the mapping of gene interactions.

As complex a task as gene sequencing is, charting the multiplicity of interactions that are possible among the roughly 30,000 genes that make up the human chromosome is even more complex. It has led to the emergence of the field of network biology as biologists begin to build computer-aided models of cellular and disease processes.

“Very quickly we realized we weren’t the only ones facing this problem and that others were independently developing software tools,” Dr. Ideker said. The researchers decided to take what at the time was a large risk, and began to develop their code as an open-source software development project, meaning that it could be freely shared by the entire biological community. The project picked up speed when Dr. Ideker, who is now chief of genetics at the U.C.S.D. School of Medicine, merged his efforts with Gary Bader, a biologist who now runs a computational biology laboratory at the University of Toronto.



MACROSCOPE A long line last week at the planetarium waited to see a show powered by a whole new class of computer-based scientific instruments. Jim Wilson/The New York Times

The project picked up collaborators in the past decade as other researchers decided to contribute to it rather than develop independent tools. The project picked up even more speed because the software was designed so that new modules could be contributed by independent researchers who wanted to tailor it for specific tasks.

“We allowed what we called plug-ins back in 2001 — nowadays with Apple’s success you would call them an app,” he said. “There are a couple of hundred apps available for Cytoscape.” The project is now maintained with a $6.5 million grant from the National Institute of General Medical Sciences at the National Institutes of Health.

Tools like Cytoscape have a symbiotic relationship with immense databases that have grown to support the activities of scientists who are studying newer fields like genomics and proteomics. Gene sequencing led to the creation of Genbank, which is now maintained by the National Center for Biotechnology Information. And with a growing array of digital data streams, other databases are being curated — in Europe, for example, at the European Bioinformatics Institute, which has begun to build an array of new databases for functions like protein interactions. Cytoscape helps transform the disparate databases into a federated whole with the aid of plug-ins that allow a scientist to pick and chose from different sources.

For Dr. Börner, the Indiana University computer scientist, the Cytoscape model is a powerful one that builds on the sharing mechanism that is the foundation of the Internet

The idea, she said, is inspired by witnessing the power and impact of the sharing inherent in Web services like Flickr and YouTube. Moreover, it has the potential of being rapidly replicated across many scientific disciplines.

“You can now also share plug-in algorithms,” she said. “You can now create your own library by plugging in your favorite algorithms into your tool.”

A version of this article appears in print on , Section D, Page 1 of the New York edition with the headline: Digging Deeper, Seeing Farther: Supercomputers Alter Science