

# NATIONAL LABS: When SCIENCE



Dr. Michael Turner, Chief Scientist at Argonne National Laboratory, answers questions about the role of national laboratories, large-scale computing, and other aspects of the SciDAC program.

**SciDAC Review:** *How have your experiences at Argonne National Laboratory (ANL), Fermi National Accelerator Laboratory (Fermilab), and the National Science Foundation (NSF) shaped your vision about the unique role of national laboratories in the U.S. research enterprise?*

**Dr. Turner:** The university research community is most important at the NSF. That community shapes the vision for the programs and receives more than 80% of NSF's dollars. It is a community of individual investigators—professors and their graduate students—and it is very good at discovery science; in fact, it has been called the Great Discovery Machine. I saw firsthand how well the NSF model works and how important funding the university research community is. But my experiences at ANL and Fermilab and at NSF made me realize that there are some kinds of science where it “takes a village” to make breakthrough discoveries.

National laboratories, like ANL and Fermilab, can assemble larger groups, cross discipline boundaries more easily, have more coherence, and focus longer than university research groups can. This makes the national labs the place to address complex issues, such as the intersection of computing and biology, climate change, and sustainable energy production. Unlike academia and private enterprise, national labs can work on a five- to ten-year research plan, manage large teams of interdisciplinary specialists, and pursue use-inspired basic research to respond to national needs. The national labs can also build and operate big facilities, like Fermilab's Tevatron—the most powerful accelerator in the world—or ANL's Advanced Photon Source (APS), which provides the most brilliant x-ray beams in the U.S. Scientific user facilities like these are large, expensive, and complicated, and so require dedicated and skilled staff to operate effectively. These facilities attract the brightest minds, young and old and from around the world, to pursue scientific

research at the cutting edge with unique instruments that cannot be replicated at every university.

Beyond providing world-class facilities, national laboratories offer a place where university researchers can work in larger and more diverse groups. Fermilab not only provides the accelerator but also the infrastructure for more than one thousand physicists from more than one hundred institutions to come together to do cutting-edge experiments in particle physics, from proton–antiproton collisions on the Collider Detector at Fermilab and the DZero Experiment on the Tevatron Collider, to neutrino oscillations at the Main Injector Neutrino Oscillation Search (MINOS) and mapping the Universe at the Sloan Digital Sky Survey (SDSS). Another example is the Midwest Center for Structural Genomics, located at ANL, which takes advantage of the APS and has recently deposited its 500th structure in the Protein Data Bank; 400 of these were within just the past three years. Funded by the National Institutes of Health (NIH), it is a consortium made up of scientists from ANL, the European Bioinformatics Institute, Northwestern University, the University of Toronto, Washington University, University College London, the University of Virginia, and the University of Texas. These collaborations and cross-disciplinary approaches have improved technology and accelerated results. These examples illustrate some of the science that “takes a village,” and the national labs are able to create and maintain scientific villages that flourish and benefit the nation in both discovery science and use-inspired science.

*How do you see large-scale computing contributing to the missions of DOE and the basic research mission of the Office of Science?*

Many now refer to large-scale computing and simulation as a third branch of science. Large-scale computing has become indispensable in almost all areas of science. By enhancing researchers' capabilities to do simulations, to compare theory to experiment, to understand complex theories, and to manipulate extremely large datasets, modern computing plays a crucial role in essentially all of DOE's missions, and is indeed essential to almost all frontier science today.

Take energy, for example. If you want to model the whole energy system, from energy inputs to energy outputs and impacts on society, you have to use large-scale computing. To find viable, sustainable energy solutions, you cannot just look at any one alternative all by itself. For nuclear energy, we need a new generation of reactors: we need to close the fuel cycle, we need more reactors that address the efficient use of uranium because it is not an unlimited resource, and we need reactors that are resistant to proliferation. Large-scale computing simulation can help us model and design such a sys-

---

# Takes a VILLAGE

tem. Currently, ANL is working on high-fidelity simulations of nuclear-energy systems. Together with nuclear engineers, scientists are developing a computational framework that integrates individual physics modules into a unified computational tool for simulation of sodium-cooled fast reactors, advanced thermal hydraulic modeling, and design of the reactor core framework.

*What are some areas at the frontiers of science where modeling and simulation might enable fundamental breakthroughs?*

The Large Hadron Collider (LHC) at CERN in Geneva, Switzerland, which will start up next year, will probe deeper into matter than ever before. In the effort to understand the collisions of subatomic particles and learn the secrets of the innermost workings of nature, large-scale computing will be key for sorting through the massive amounts of data, picking events for further analysis, and helping us test theoretical explanations. As partners in the ATLAS experiment at CERN, ANL will be working with Indiana University and the University of Chicago to handle data flowing from the new collider by developing innovative computing infrastructure and techniques for extremely large-scale data handling and complex domain-specific collaboration with research physicists. Without GRID computing it is doubtful that discoveries could be made.

Another example comes from the biological side. Computing plays a central role in understanding the complex nature of biological systems. One of the grand challenges is to sequence an organism's DNA, which is a set of incredibly long molecules, and to rapidly interpret the sequence—that is, to determine where the coding sequences for proteins begin and end, a process called “annotation.” James Watson's DNA was sequenced in about six months at a cost of one million dollars. Think of what we could do if we could sequence the genome of a human in a matter of a few days for a few hundred dollars. ANL researchers are developing a server that retrieves a new genome or variant from the National Center for Biotechnology Information (NCBI) database and produces an annotated genome automatically, substantially reducing annotation time from weeks to hours. This capability will lead to a better understanding of protein families and help provide a foundation for comparative genomics.

Four exponentials are driving the information landscape: computing speed, data storage, network bandwidth, and sensor complexity are all doubling every 18 months or so. The big challenge is to harness these exponentials to accelerate science breakthroughs. One simple example is better utilization of the computing power that exists. Right now we are rapidly approaching petascale computing, but I don't think we are cur-

rently getting full scientific use out of terascale computing. Effectively programming computers with hundreds of thousands of processors is not an easy task, and today, science speed does not match bench speed. The great challenge is to turn petaflops into mega-discoveries.

*What role or roles do you see for computation in pursuing the path of connecting quarks with the cosmos?*

The mother of all examples is simulating the Universe. We are really good at this when all we consider is gravity, but if there were only gravity, we would not be here! Hydrodynamics, the interaction of the atoms to form stars, is very challenging to handle. Doing so is critical for connecting the predictions of the early Universe for the initial conditions that seeded the galaxies and clusters of galaxies to what we see today. Observations largely involve light and stars, and we have to understand how the matter in the Universe forms stars and lights up. Without that link it is hard to close the circle.

Another area, which is not unique to quarks and cosmos, involves very large datasets. If you look at the cosmic microwave background, you can resolve an enormous number of independent pixels on the sky, tens of millions, if not a billion. Trying to manipulate that data, understand the correlations, and distribute that wealth of information is taxing to the largest computers. Much of the deluge of data in science today is being driven by sensors; for example, CCD cameras, which are following Moore's law and doubling their number of resolution elements every 18 months, now have a billion pixels. To understand how dark energy has shaped the large-scale structure of the Universe, we will need to analyze datasets tens of petabytes in size, such as those generated by gigapixel CCD cameras.

*Many people in the computing community use the Connecting Quarks with the Cosmos report as an example of what is missing in the field. Can you say a little about what was required to rally people around the “11 big questions” in the report?*

We used the 11 big questions to give some coherence to a diverse and exciting set of activities. We used these big questions to communicate the importance and excitement of what was already going on. I think articulating what scientists do is something that will become more and more important as time goes on.

There is great support in this country for funding basic research, in part, I think, because people see the links, even if they don't quite see how the links work, between basic research, competitiveness, economic well-being, national security, and health. In some fields, you don't have to work

## Biography in Brief: Dr. Michael Turner

**Dr. Michael S. Turner** is the Rauner Distinguished Service Professor at the University of Chicago, Staff Scientist at the Fermi National Accelerator Laboratory, and since January 2007, the Chief Scientist at ANL. He received his B.S. in physics from the California Institute of Technology (1971) and his Ph.D. in physics from Stanford University (1978). He came to the University of Chicago in 1978 as an Enrico Fermi Fellow and joined the faculty in 1980. From 1998 to 2001 he was the first Scientific Spokesperson for the Sloan Digital Sky Survey. Dr. Turner was instrumental in establishing the Kavli Institute for Cosmological Physics at the University of Chicago in 2001, and from 2003 to 2006, he served as the Assistant Director of the

National Science Foundation for Mathematical and Physical Sciences. Dr. Turner is a Fellow of the American Physical Society, the American Association for the Advancement of Science, and the American Academy of Arts and Sciences, and he is a member of the National Academy of Sciences.

Dr. Turner helped to pioneer the interdisciplinary field that has brought together cosmologists and elementary particle physicists to unravel the origin and evolution of the Universe and to understand the unification of the fundamental forces and particles of nature. His research focuses on the earliest moments of creation, and he has made seminal contributions to inflationary

cosmology, the understanding of particle dark matter (weakly interacting massive particles) and the large-scale structure of the Universe, the theory of Big Bang nucleosynthesis, and the study of dark energy, a term he coined to describe the driver for cosmic acceleration—the speeding-up of the expansion of the Universe. He believes that cosmic acceleration is the most profound mystery in all of science today, and it is the focus of his current research. The National Academy study *Connecting Quarks with the Cosmos*, which he chaired, identified opportunities at the intersection of astronomy and physics and has shaped the science investment in the U.S. and elsewhere around the world.

very hard to articulate the benefits; for example, everyone understands that life science research leads to cures for diseases. In the discovery sciences, where you can't promise the big invention, or the economic benefit (at least not in the foreseeable future), it is extremely important to articulate the excitement of the puzzles people are trying to solve.

The science in *Connecting Quarks with the Cosmos* is almost 100% discovery science, and the only way that you can try to make the public understand why it is important and why they should provide funding for it is to share the excitement of the adventure. In writing the report, we didn't change the direction of the field; what we did do was present the science at the interface of astronomy and particle physics in simple words so that everyone could appreciate how stunning the potential breakthroughs are. How did the Universe begin? What is the nature of the dark matter that holds everything together? What is causing the expansion of the Universe to speed up?

When considering discovery sciences, policymakers in Washington have a difficult job in figuring out how to allocate funds. There is very definitely a sense that some of the research this country does should be in discovery science or curiosity-driven science; historically, discovery science has produced the big breakthroughs, like quantum mechanics, which produce a sea change and dramatic economic return (the information age, for the example of quantum mechanics). Discovery science also attracts the next generation of scientists and engineers. Clearly, discovery science is part of the investment portfolio, but finding the right metric for making funding decisions is difficult. When you're talking about use-inspired science, it's much easier to see the benefit, and the metric for funding decisions is tied to the benefit. In discovery science the metric is related to the potential for breakthroughs, and so articulating the opportunities for discovery is essential.

*Can you explain how you see computation, experiment, and theory working together in the future?*

The short answer is "no."

This is one of the grand challenges for all of science. It took a long time in my field, astronomy, for people to understand that progress involves two distinct activities—experiment (we call it observation) and theory. Other fields came to realize more quickly that these are two separate activities.

In astronomy, it used to be that people built their instruments, made their observations, and then interpreted them. And that was a very, very simple model; everybody was a jack of all trades. It took a while for astronomers to see that specialization can have some real benefits, that in fact it takes a village. And now there is a new player involved—computation. I don't have to go out on a limb to say that we're not using resources optimally right now. And partly it's because of the four exponentials that are driving digital science; it is hard to keep pace with the rate of change in computation.

When you look at the grand challenge computing problems (for example, colliding black holes), it's going to take a team to solve the problem. And the team is going to involve not just discipline scientists, but also computational experts. We don't even have a name for them yet, but we recognize the need for people who can help with the optimal coding algorithms and visualization. We are just starting to learn how to assemble these teams to attack computational problems.

Here's another example of how we're just starting to learn how to tackle big computational problems. I was involved in the SDSS, a groundbreaking project with astronomers thinking about doing an astronomy experiment, rather than building an observatory that would last forever. The experiment was to map the large-scale structure of the Universe. But in our original budget, we forgot a small item—

software. And we were only the first of many projects to forget this item. We somehow assumed that the software for the data pipelines that would turn the 1s and 0s from the sky into spectra and images would be written by professors (and their grad students) late at night in their spare time. We created the software teams at significant additional cost and well into the project. This is an area of computation which is only going to increase. Computational professionals are extremely important to almost all big projects today. We don't have a name for them. We don't quite know where we'll train them, but it is clear that they are essential to computation transforming science.

*In your role as Chief Scientist at ANL, you have an opportunity to influence the future direction of the laboratory. What are some big opportunities you see coming down the road?*

One of the biggest problems that our country and the world face is sustainable energy production and use. Central to the question of sustainable energy production is climate change. It can't be separated from the energy problem, which is why I used the word "sustainable" energy. This is a very hard and critical set of problems and it's going to take not just one village, not just one national laboratory, but most of the DOE Office of Science laboratories, each playing a role in this very important endeavor.

ANL has some unique capabilities. We'll have a petascale computing facility designed to provide the computational science community with world-leading computing capability dedicated to breakthrough science and engineering. This will enable ANL and our collaborators to attack key scientific problems and expand the frontiers of discovery in areas like nuclear power, genetic and biological processes, and climate change. At ANL we have a long history in nuclear energy and design of reactors. We helped design the first nuclear reactors, and almost everyone sees nuclear energy as being part of the solution to this problem. This problem is going to involve policy analysis and the integration of economists and social scientists, and ANL again has some real capabilities here, not only at the laboratory but through its partnerships at the various universities—Northwestern University, the University of Chicago, and the University of Illinois. Energy is definitely one of the future directions of the laboratory, and we have to find our role in helping to be part of the cooperation between the villages that are working to solve this problem.

Another important area for ANL, because it's another thing that national laboratories can offer, is facilities. The APS, in my slightly biased opinion, is the most successful user facility that the Office of Science operates. We already mentioned the use of APS x-rays to solve protein structures. They also allow scientists to pursue new knowledge about the structure and function of materials, whether at the center of the Earth or from outer space, with far-reaching impacts on our technology, our economy, and our fundamental understanding of the materials that make up the world. Of course, a critical thing about all facilities is that if they are going to keep pushing the frontiers of science forward, they have to be upgraded. The upgrade path for the APS is extremely important because this light source enables so much science, both on the discovery side and on the use-inspired side.

There are two other facilities that ANL sees in its future. Nuclear physics has identified as its highest priority for a new facility a Facility for Rare Isotope Science to complement similar facilities being built in Japan and Germany, and I think ANL is uniquely positioned to lead this project forward for the nuclear science community. This project will make possible the kind of science you can express in simple words: completing the chart of all nuclides, understanding the properties of these nuclei, and measuring how these nuclei interact. This facility will be critical to astrophysics by measuring properties and cross sections of rare nuclei that played a role in producing the more common elements we see today. This facility may well have medical applications, as unstable nuclei have been useful in therapy and in diagnosis.

The other future facility is the International Linear Collider (ILC). Here we're going beyond the science that takes a village to the kind that takes a collection of villages. While the ILC is a Fermilab project, Fermilab needs help from other laboratories, both in the U.S. and around the world, to do the accelerator research and development. I'm sure that if Fermilab wins this project and it is built there, other labs—particularly ANL, the laboratory that's closest to Fermilab—will have an important role to play in this multibillion dollar facility.

The last example that I want to talk about is the one that I know least about. It's the interface between the biological sciences and the physical sciences. That's an incredibly broad frontier that cuts across lots of problems, from the energy problem (Can we learn from biological systems how to make energy? Can we harness biological systems to be more efficient at using the sun's energy?) to applications for medicine and the life sciences, and even to how life works. We want to build machines at the nanoscale. Can we learn how to build these machines from nature? Can we harness nature to build the machines for us? It's an incredibly fertile interface. And again, this is one of these areas that I think national laboratories are uniquely positioned to contribute to because it requires facilities, it requires an interdisciplinary approach, and it requires teams of scientists. I think that ANL's Center for Nanoscale Materials will play a leading role in these areas, and there will be exciting discoveries that will cut across the laboratory.

*The SciDAC program at DOE has formed partnerships between applied mathematics, computer science, and many science disciplines including biology, nuclear physics, astrophysics, fusion, climate, and nuclear energy. What advice might you give to the next generation of computational scientists on how to focus their energy on what is important?*

Science first! It was my strategic plan when I was the Assistant Director for Mathematical and Physical Sciences at NSF, and this very short strategic plan, embodied in two words, was often very helpful when one really got into the complicated details of trying to get things done. Sometimes when you are dealing with complicated organizations and partnerships and you are trying to attack important and exciting science problems, you get so immersed in the details that it is easy to lose sight of what you are actually trying to do—science! For the SciDAC program, science is first in the name, so make it first everywhere. ●