

Rick Stevens: Connecting Computing to Science
by Michael Feldman
HPCwire Editor

Perhaps the two most important technologies of the 21st century will be information technology and biotechnology. Certainly they have become the most rapidly expanding domains of our era. The advancements in devices such as microarray biochips, medical imaging, and mass spectrometers have created a wealth of biological data to be analyzed. The result is that, increasingly, biological problems now require large scale computing. In a sense, life science has become a sub-domain of information science.

Expressions such as bioinformatics, computational biology and systems biology are being used to describe this new integration. And research organizations are actively exploring problems within the intersection of biology and computer science.

At the Department of Energy's (DOE) Argonne National Laboratory, the Computing and Life Sciences (CLS) directorate is attempting to synergize these two technologies to fulfill the department's mission. At Argonne, the integration of computational science with systems biology is designed to help build basic scientific knowledge, solve environmental problems related to energy production, and develop and manage new energy sources.

Heading the CLS directorate is Rick Stevens, a man who seems perfectly suited for the type of interdisciplinary work that the organization is doing. There, he is able to indulge his deep interests in algorithms, math and science, especially biological science.

As a scientist, Stevens is hard to categorize. In fact, he himself is not a great believer in distinct scientific disciplines. According to him, calling yourself a biologist, a chemist or a computer science is a just way people self-identify with a community. But these disciplines have become a rather artificial way to view the world. There are just people and problems, he says.

"I've always been interested in trying to connect computing to science," says Stevens. "But I'm not that interested in computing for computing's sake."

As a kid, Stevens was very much attracted to computing as it was portrayed on Star Trek. In the 23rd century, computers were things you used to do exciting things, like computing wormhole trajectories. In the 21st century, we'll have to be satisfied with sub-warp applications. But that still leaves plenty to do.

According to Stevens, putting biology and computing under the same lab directorate is a kind of experiment. By forging these cross-cultural relationships, they want to see if the sum is greater than the parts. Since Stevens is personally aligned with this intersection of computing and biology, to him the challenges are not only some of the most interesting problems in the world, but also are just great fun.

As one might imagine, the life of the head of an DOE lab directorate can be rather intense. It's not unusual for Stevens to be up at 5:00 AM. At that ungodly hour, he tries to pound out a little code, which he mostly writes in C, Perl, Python or Mathematica. He says he's also learning a little UPC.

"If I spend a couple of hours in the morning writing code, I'm a much more cheerful person the rest of the day," notes Stevens.

He spends the remainder of the day managing the lab: cheerleading the staff, working with the funding agencies, and planning the direction of the lab work. He tries to reserve some time for himself to reflect on the big picture and think about the future.

But when things get onerous at the lab, Stevens retreats to his other job -- as a professor of computer science at the University of Chicago. There you'll find him working with his five PhD students. With one exception they are all working on projects in computational biology.

Stevens seems to get the most out of both his roles. He says Argonne is a great place to get things done. It's a very high energy, very focused environment, and the people are extremely supportive. "We think of it as a cross between a university and a start-up company," he says. On the other hand, he also enjoys the teaching culture and more free-wheeling atmosphere of the university. There, he's able to wander off and follow his interests, wherever they take him.

But one of the big advantages of working for the DOE is the access to big iron. As one of the department's leadership computing centers, Argonne is on a select list to receive the latest cutting-edge supercomputers. It is expected to get a 100 teraflop IBM Blue Gene machine sometime this year. In 2008, the lab is looking to deploy a 500 teraflop system. Stevens says the lab is on a trajectory to get a sustained petaflop and even beyond.

High-end capability supercomputing systems for life sciences have traditionally focused on biochemical modeling at the level of atoms and molecules. But, according to Stevens, that misses the complexity of the organism and interactions of the ecosystem. Lately, he has become interested in applying petascale power to systems biology problems. For example, modeling microbial soil habitats is a vast computational undertaking, but promises to help us understand one of the most complex and important ecosystems on the planet.

Another promising use of petascale systems involves building models of cells that incorporate genetic information. This will allow scientists to predict a cell's response to different environment and substrates, and perform computational what-if questions to understand design tradeoffs in natural or man-made biological systems. For example, this type of application could be used to model highly efficient ethanol-producing microorganisms for different nutritional substrates.

"To understand the dynamics of how something works, you have to execute a simulation on a computer," explains Stevens. "There's no other way to do it. So in many ways, doing theory in biology is going to be equivalent to doing these complex simulations. That's an insight that is just starting to hit lots of people."

The computing power required to pursue some of these problems already exists today. As teraflop systems become available to more people, the opportunity for scientists to do interesting systems biology is exploding. While the hardware continues to become more accessible, building the models is the hard part.

"We don't have enough people with a background in computing and mathematics and, at the same time, have a background in biology, to actually wire these two things together," he says. "Most bioinformatics programs are too superficial. Because of that we have a lack of models."

The lack of expertise in computational biology may be holding back the field, but futurist and inventor Ray Kurzweil probably considers that problem just background noise. If there's anyone more bullish than Rick Stevens on the potential of computer science and biology, it's Kurzweil. His notion of "Singularity" is the predicted outcome of merging immense computational power with human beings, precipitating "a technological change so rapid and profound it represents a rupture in the fabric of human history." Not surprisingly, Kurzweil's prediction of a transhumanist world draws its share of controversy.

"Ray has a hugely optimistic vision of where humanity could go," says Stevens "What he's try to say is that the future could be so unbelievably cool that we should all want to get there. He's thinking in terms of exponentials and trying to understand the effects of extrapolation. The question is: How good are we at predicting the outcome of complex questions where there are underlying exponential drivers, like Moore's Law?"

"Is there merit to this view of the world?" continues Stevens. "Well, probably. It's been well understood that people have a hard time thinking in exponentials. This is a classical futurist viewpoint, whether it is understanding population increases, global warming, pollution or whatever. People are very bad at making predictions. They tend to over estimate in the near term and under estimate in the long term. What this means is that Ray could very well be correct that in 10 to 20 years, the convergence of these underlying technologies will enable many, many things to be different. Now, if he just simply said that, I don't think anyone would disagree."

The fact is, Kurzweil is making a more precise prediction: achieving Singularity in 2045. According to Stevens, that's where he gets sort of quasi-theological. What Kurzweil is essentially arguing is that the technological juggernaut will take us to this brave new world regardless of the specific technologies in play. In other words, it's not a function of Moore's Law, network bandwidth, storage capacity, bioimaging technology, microarray chips or any number of rapidly growing technologies; it's the exponential rate of technology itself.

Stevens sums it up as follows: "To solve problems you need three things -- time, money and ideas. If you have two, you can compensate for the other one. Kurzweil collapses time and money because of exponential processes. What's left are the ideas."