

Technology

Volume 18 | Issue 16 | 30 | Aug. 30, 2004

[Previous](#) | [Next](#)

The Democratization of Supercomputing

Costs are down, performance is up; supercomputing power has come to the people | [By Bennett Daviss](#)

Courtesy of Oak Ridge National Laboratory



The genome has been read. The proteome has been opened. As a result, research problems have gotten more difficult. Fortunately, access to the tools that help investigators rise to those new challenges is quickly becoming easier.

How much easier? Ask Charles Taylor, a biomechanical engineer at Stanford University. To model the flow of blood in human arteries, Taylor needs to solve as many as 10 million nonlinear partial differential equations at once, "and because the equations are nonlinear, we need to do a few subiterations of each," Taylor says. "That definitely requires supercomputing."

Taylor is creating a digital tool that will allow physicians to translate MRI and CT scans into blueprints of each patient's unique vascular architecture, enabling surgeons to plan and then simulate various treatments so they can determine which will be most effective for any given individual. Once, Taylor would have had to wait for time on giant machines such as the "Thunder" supercomputer at the Lawrence Livermore National Laboratory (LLNL) in California. But today he doesn't even have to leave his desk.

Login [:: Go](#)

Are you a new user?
[Please register](#)

Tools

[Comment](#)

[E-mail article](#)

Vertical Electrophoresis Systems



Mini-PROTEAN® 3 System

Small gel format
for fast results

Criterion™ System

Medium gel format
for speed and
resolution

PROTEAN® II and PROTEAN® Plus Systems

Large gel format
for maximum
resolution

BIO-RAD

[Daily News](#)

[Upfront](#)

[Feature](#)

[Research](#)

[Technology](#)

[BioBusiness](#)

[Supplements](#)

[Archives](#)

[Print Renewals](#)

Search

[:: Go](#)

[Advanced Search](#)

adsumo
LIFE SCIENCE CAREERS



Readers' Choice Awards

Instead, he can test his models in a few hours using no more than half of the CPUs on the 128-processor Silicon Graphics (SGI) computer cluster that serves his department. "If we used all the processors, we could run a bigger problem or the same problem faster," he says, "but that full power is usually more than we need."

Taylor is one of thousands of life scientists reaping the benefits of what can be called the democratization of high-performance computing. More machines with more speed and power are coming within the grasp of more and more scientists. IBM's two "Blue Gene" prototypes currently hold the fourth and eighth spots on the "Top 500" list of the world's fastest computers, a semi-annual global survey compiled by computer scientists at the Universities of Tennessee and Mannheim. IBM will deliver a 64-rack version capable of 360 trillion floating-point logic operations (teraflops) at once, to the LLNL around the end of this year.

"For massive problems such as modeling protein folding, the Blue Gene already is able to calculate in one day about a nanosecond of atomic and molecular activity," says Ajay Royyuru, senior manager of IBM's computational biology center. "Five years ago, a state-of-the-art machine would require a month or more to accomplish the same task."

The Oak Ridge National Laboratory (ORNL) in Tennessee, partnered with Cray in Seattle and SGI in Mountain View, Calif., is about to begin building a new supercomputer able to blaze through 50 teraflops, 25% faster than Japan's Earth Simulator, which holds the current speed record. Optimists project that the new US machine could pass speeds of 100 teraflops before 2007. And the machine will be available to any researcher who can satisfy an independent review panel of the need for its use.

Even distributed computing, the most democratic form of high-powered data processing, is offering the promise of more widespread usability. The idea of yoking together stray PCs that can spend their idle moments working problems has long been attractive. As long as each task is simple and self-contained, distributed computing has been a cheap ticket to working a large number of similar, relatively simple problems. But getting machines of different speeds with different processors and operating systems to handle the same commands in the same way has long been a headache.

Those strictures may now be easing as well. The exemplar of a problem unfit for distributed computing has been protein folding: The processors devising the model need to exchange data constantly to make collective sense of the changing structure, a communication feasible only among processors that are close to each other and tightly linked. But now physicist Vijay Pande at Stanford University is making even that problem work across a grid of 165,000 PCs scattered across more than 100 countries. By using proprietary software techniques to

break down the model into minute time steps and manage the network, Pande's "folding@home" project has modeled 10 proteins since beginning work in October 2000. Pande hopes to make his techniques available to other researchers before long.

NEW DEVELOPMENTS Giant supercomputers and transnational grids aren't the only forms of computing to welcome more life-science researchers. Over the last few years, four developments have put more raw computational ability in the hands of more scientists who are spending no more, and often less, money for it: the rise of cluster computing, the development of lower-cost 64-bit processors, vendors' embrace of the Linux operating system, and the widespread availability of off-the-shelf components able to handle the massive demands of today's research projects.

Courtesy of Sun Microsystems



SUN JAVA WORKSTATION W110Z

According to Jack Dongarra, professor of computer science at the University of Tennessee and a founder of the Top 500 list, "A striking change is that we're seeing a huge increase in the number of clusters. Of the top 500 machines in June 2004, 287 have Intel processors. Six months ago, that number was 189. The increase is fueled by the increased capacities and low cost of Intel's processors, mainly Pentium, the same chips used in home PCs, and also some its newer Itanium processors."

Clusters are collections of plug-in units such as "blades," the circuit boards or tube-like servers with processors and memory aboard, or what Lorilyn Mears, Sun Microsystem's market segment manager for life sciences, calls "pizza boxes." Each has a processor, interconnecting hardware, and usually some memory, and can be slipped into a rack with other modules. With management software such as Sun's new Grid Engine 6, the modules can work together even using chips of varying architectures. When more computing power is needed, researchers can buy more plug-ins; units unused by one researcher

are available to others.

Clusters are filling the gap between cheap but cumbersome distributed computing and powerful but demanding supercomputers. "There's a certain amount of red tape in using a supercomputer," says Erec Stebbins, an assistant professor at Rockefeller University, New York. "You have to make sure all the software you need will function appropriately on that particular machine." He adds that sometimes users wait weeks for their turn in line.

Stebbins is happy using his university's cluster instead, testing molecular inhibitors of potential biological agents such as *Salmonella* and *Yersinia*. With his 18-node cluster, each node sporting an AMD Athlon 1.66 GHz processor, he can screen about three million compounds in two weeks. "We can quickly look for certain properties in molecules that might work as a drug. When we're modeling how each of these molecules might dock with pathogenic proteins, we farm those jobs out to our cluster. We have to cut back on accuracy to finish the screening within a given time period, but it's absolutely essential that we have a cluster at least this size." The price: about half of what the same computing power cost just 18 months ago.

FALLING PRICES In tandem with hardware costs, the price of 64-bit processing has been falling. "Two years ago, 64-bit processing was only in the realm of the most expensive machines," notes Marc Rieffel, senior manager of research and development at Paracel, a computer systems manufacturer in Pasadena. "Now, with the advent of AMD's Opteron chip and Intel's upcoming EM64T Nocona processor, the amount of memory you can address with 64-bit applications is virtually limitless by today's standards, while you still can run standard 32-bit applications on 64-bit systems." Previously, 32-bit hardware has been limited to addressing only about four gigabytes of memory at a time.

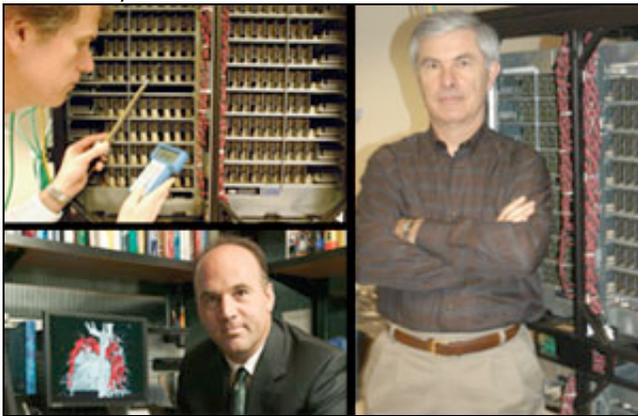
And then there's Linux, the open-source operating system. Until recently, IT managers had only two options, both unattractive: Pay a premium for site licenses to use a proprietary operating system such as Microsoft Windows and hope that the software the lab needed would run on it; or skip the fees and adopt the freely available Linux system, which required a good deal of tinkering and customization.

Linux was "very much the hand-rolled cigarette, with the tobacco falling out the end and the paper never quite sticking properly," says James Cuff, group leader of applied production systems at the Broad Institute at Massachusetts Institute of Technology. "Three years ago, it was hard to find Linux support for new hardware such as disk controllers or network cards. Now manufacturers are catching up. The Linux kernel itself has improved in tasks such as memory scheduling, and major vendors are offering thoroughly supported, rock-solid Linux systems. Now we're sitting on a nice, stable platform."

As researchers and suppliers adopt an open-source operating system, they also are making more use of open-source hardware. Off-the-shelf processors, memory storage, and other parts, the so-called commodity components, have grown in speed and capacity such that they can substitute for the pricey, specially designed SCSI drives and ASICs (application-specific integrated circuits) that previously devoured equipment budgets.

"The PC-ization of high-performance computing has driven down price performance to an amazing degree," says Curt Van Tassell, a research geneticist at a US Department of Agriculture lab in Beltsville, Md. "Two years ago, a small lab like ours with a modest budget couldn't have dreamed of having the capacity that we have now."

Top images: courtesy of IBM; Bottom image: courtesy of Charles Taylor



BRINGING POWER TO THE PEOPLE: IBM

Research team member Shawn Hall (top left) and Bill Pulleybank, director of exploratory server systems for IBM Research (right), with the Blue Gene/L prototype. Bottom left: Stanford University's Charles Taylor.

Manufacturers get the point. SGI sales grew partly because of its proprietary chips and hardware; it is now overhauling some of its turnkey systems, replacing its proprietary UNIX variant and MIPS and IRIX processors with 64-bit Linux and Intel's Itanium 2 chips. Paracel, on the other hand, is sticking with proprietary components. The company's GeneMatcher package "is still orders of magnitude faster for certain kinds of genome comparisons, such as Smith-Waterman searches, than any commodity components," says Rieffel, "but it can't do everything."

IDLE "MACHOFLOPS" But the coiled power of those shiny new, low-cost components masks a seductive danger. "Some folks buy machoflops, the biggest, best hunk of iron they can afford," says John Reynders, vice president of informatics at Celera Genomics in South San Francisco. "It's wonderful to have that machine screaming away in the basement," he continues, "but if you don't have a software infrastructure that maps your application to that hunk o' iron, then

your asset is doing nothing but depreciating while your team rubs sticks and stones together to get something working."

Dongarra calls the state of high-performance software development "a crisis. We're far outpacing in hardware what we can do in software. We're using on our parallel machines a programming model that was developed in the 1970s for serial computers, and now some of our parallel machines have tens of thousands of processors."

The old model calls for a programmer to create the details of every step in the communication among those processors using a serial and sequential format, Dongarra points out. Working with parallel machines, he says, "We need to have a way to work at a higher level of abstraction. Today it's hard to maintain programs for these large new systems, it's difficult to understand what the programs are doing, it's tedious to debug the programs, and it's tricky to produce optimized implementations."

Promising to make the problem worse: One version of a software program typically is transferred from old machines to new ones, compounding the software's shortcomings as the same programs with the same limitations migrate to more powerful, complex, and massively parallel computers.

Dongarra and others blame, at least partly, the culture of professional science, where writing one's own programs is more than just a matter of pride and tradition. Many, if not most, noncommercial labs write their own code "and we probably always will because we're always tweaking and modifying software in our research programs, and that's not always possible with commercial products," says Fernando Pineda, associate professor of molecular microbiology and immunology, and biostatistics at the Bloomberg School of Public Health at Johns Hopkins University.

Also, a tradition of homegrown software nestles easily into a culture where such projects become the subject of academic papers or part of a case for tenure. But it leaves little of negotiable value in the expanding territory where academic or government research and commercial ventures overlap. "Software infrastructure isn't tracking because it's not cheap to make that investment," Reynders notes. But, as projects and hardware become more complex, the software culture of nonprofit science could begin shifting ever so slightly.

"The trend in our lab over the last 15 years has been toward an increasing use of validated commercial code," says Ananth Annapragada, associate professor of bioinformatics at the University of Texas School of Health Information

Courtesy of Paracel



Sciences in Houston. His research group is modeling the way air flows in human lungs, hoping to track the path and assimilation of inhaled medication.

Annapragada's 96-node cluster of Dell 32-bit processors crunches problems about as large as Taylor's vascular model, "but more than 70% of the simulations we run today involve commercial code," he says. It's not only easier than constantly rewriting the code, "but it's also easier to train new students, who can focus on the physics, rather than the numerics, of what we do." The gradually growing library of Linux-based programs from Paracel, SGI, and other vendors likely will accelerate that trend, observers say.



 **PARACEL CYCLONE**

INTERNODE COMMUNICATION

The lack of software must be remedied in tandem with another bottleneck, researchers say: shuttling data in and out of storage, as well as among multiple processors that need it. "If a processor is theoretically capable of doing, say, four trillion operations per second, perhaps it can achieve only five percent of that because it's waiting for data to come from somewhere before it can do its calculation," says Thomas Zacharia, director of ORNL's computer and mathematical sciences division. "We've realized that two areas needing attention are the speed of the interconnect switch as well as memory bandwidth."

Yet an effective solution has to go beyond those localized issues. "At Celera, we have about 100 terabytes of storage for every one teraflop of compute power," says Reynders. "One of the biggest challenges in biotechnology is dealing with heterogeneous data: 'How do I lay this bioinformatics data over a question I'm asking in chemoinformatics?' At the moment, data management has the most catching up to do." Reynders is intrigued by recent offerings such as the Netezza database architecture, and he holds some hope for Web-based services and grid technologies, "but those are still very nascent."

Vendors are on the case. Sun's "chip multithreading" architecture lets processors crunch data and move it around a system at the same time. SGI is developing a "storage area network" called the InfiniteStorage Shared Filesystem CXFS that enables networks of any size that encompass more than one operating system to store and retrieve data from one common source instead of having to format the

same data in different ways for different platforms.

SGI also is working to streamline image sharing by sending only pixels across a network, not the entire bulk of data needed to recreate the pixels anew at each site. Sun is developing a method to distribute images so that far-flung groups can view and alter them collaboratively in real time. In addition to issues of data security and regulatory compliance, Sun also is crafting what it calls the Java Enterprise System, which unites the operation of servers, databases, and operating systems in one package. Says Mears: "This will be the glue that connects high-throughput computing to high-performance computing."

Paracel is working with Platform Computing in Ontario, which makes a product called LSF (load-sharing facility) MultiCluster that automatically shuttles jobs among clusters that use the same hardware and software. Paracel expects to demonstrate one such "grid of supercomputers" before the year is over.

Such gains suggest that Taylor's dream of modeling a patient's arteries on a computer so that physicians can test various therapies in just a few hours could become reality within a decade. "Making the process fast and easy will make it available to more patients," he says. And that's the ultimate goal: The spreading democratization of high-performance computing in research today will make more effective treatments available to everyone tomorrow.

Bennett Daviss (bdaviss@the-scientist.com)

Vote

Please indicate on a 1 - 5 scale how strongly you would recommend this article to your colleagues?

Not recommended 1 2 3 4 5 Highly recommended

Please register your vote

Vote

© 2004, The Scientist LLC, All rights reserved.

[Previous](#) | [Next](#)

