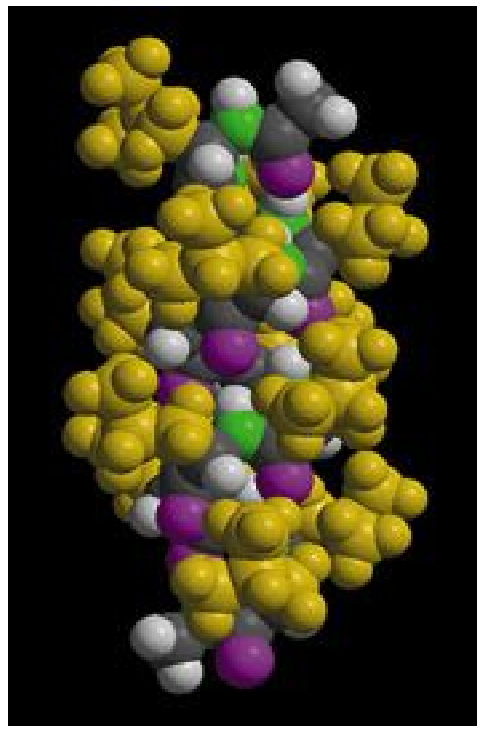
NASnews
January – February 1999 • Volume 4, Number 1

In this issue...

Biomolecular Simulation Reveals How Cells May Integrate Proteins



A molecular dynamics simulation using the NAS CRAY C90 showed how a randomly-coiled protein fragment can use the cell membrane to fold itself into a helix, shown here.

If living organisms exist elsewhere in the solar system, they are likely to be composed of cells bounded by membranes analogous to those of terrestrial cells, some biologists hypothesize. On Earth, cell membranes perform such essential functions as holding in the cell's contents and regulating the absorption of fuel and raw materials. But predicting how extraterrestrial cell membranes might work is difficult, because researchers can't precisely explain how the membranes of even the simplest Earth cells discriminate between useful and toxic molecules, or how certain proteins assume key-like shapes that allow them to "pick the cell's locks" and insert themselves into the membrane.

Now, however, a team at NASA Ames Research Center has used supercomputer-based simulations and cutting-edge visualization techniques to solve one small part of this puzzle. In a recently published paper, Christophe Chipot and Andrew Pohorille of the

Ames Exobiology Branch describe a large molecular dynamics simulation of a protein fragment undergoing folding at a membrane-like interface. The simulation demonstrated for the first time how such a molecule can use the interface to fold itself into a helix capable of inserting itself into the membrane.

To The Article...

NASnews

▶ Main Menu

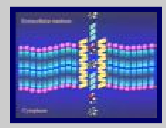
January – February 1999 • Volume 4, Number 1

In this issue...

Biomolecular Simulation Reveals How Cells May Integrate Proteins

by Wade Roush

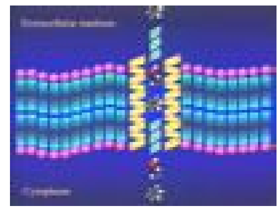
Within This Article...



If living organisms exist elsewhere in the solar system, they are likely to be composed of cells bounded by membranes analogous to those of terrestrial cells, some biologists hypothesize. On Earth, cell membranes perform such essential functions as holding in the cell's contents and regulating the absorption of fuel and raw materials. But predicting how extraterrestrial cell membranes might work is difficult, because researchers can't precisely explain how the membranes of even the simplest Earth cells discriminate between useful and toxic molecules, or how certain proteins assume key-like shapes that allow them to "pick the cell's locks" and insert themselves into the membrane.

Now, however, a team at NASA Ames Research Center has used supercomputer-based simulations and cutting-edge visualization techniques to solve one small part of this puzzle. In a recently published paper, Christophe Chipot and Andrew Pohorille of the Ames Exobiology Branch describe a large molecular dynamics simulation of a protein fragment undergoing folding at a membrane-like interface. The simulation demonstrated for the first time how such a molecule can use the interface to fold itself into a helix capable of inserting itself into the membrane. The simulation covered a mere 50 nanoseconds in the life of the fragment, a polypeptide "undecamer" consisting of 11 amino acid units, but it separated that fleeting moment into 25 million time steps–making it the largest analysis of its kind when it was carried out last year on a Hewlett-Packard workstation at the Exobiology Branch and a CRAY C90 supercomputer at the NAS Facility.

"Protein folding is the Holy Grail of structural biology," explains Pohorille. "All proteins must undergo some folding to enter the membrane, but we know very little about the sequence of events that lead to folding and translocation," he says. "The result we got was one of the nicest things we could have hoped to find."



RealVideo clip:
28.8, 56, ISDN

Video Credits

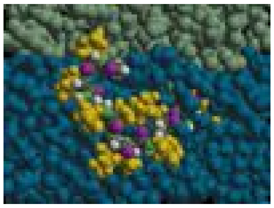
Their results are described in a paper published in the November 27 issue of the *Journal of the American Chemical Society.* At the same time, NAS data visualization consultant Vee Hirsch helped Chipot, Pohorille, and colleague Michael New turn sections of the numerical simulation into a 10-minute video showing crucial moments in the undecamer's voyage into the membrane. "The animation turned out to be quite revealing," Pohorille says. "We learned about things that would have been hard to pick out just by looking at the numbers," such as the way the undecamer undergoes a Titanic-like pivot before sinking head-first into the interface. In March 1998, the presentation won third prize among 106 films competing at the International Festival of Scientific Animation in Nancy, France.

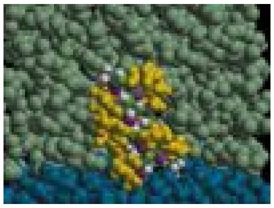
Deconstructing a 'Miracle'

Pohorille started scientific life not as an exobiologist but as a theoretical physicist, applying quantum mechanics to isolated biological molecules. In order to understand the large assemblies of molecules that make up living organisms, he eventually turned to large-scale computer simulations that also employ statistical mechanics. Once computers grew powerful enough to calculate the myriad atomic forces acting within and between complex molecules such as polypeptides, it became possible to simulate the behavior of biomolecules over time, Pohorille explains. Simulations involving many molecules could eventually reveal clues about how the rich but lifeless soup of organic chemicals present on the early Earth gave rise to the first protocells–and about whether the same thing could happen on other planets.

One mystery about Earth cells has to do with the remarkable shape-shifting abilities of so-called "transmembrane" proteins, polypeptides that lodge themselves in the cell membrane in order to carry out jobs such as transporting nutrients, waste products, and biochemical signals through the cell wall. In a polypeptide molecule, the amino-acid subunits are linked together along a backbone of carbon and nitrogen atoms, with "side chains" including hydrogen, oxygen, nitrogen, and other atoms branching out from the backbone. In water, the backbone units of most transmembrane proteins freely swivel with respect to one another, giving the molecules irregular, coiled structures. Once inserted into the membrane, however, most transmembrane proteins are twisted into a regular helical pattern called an alpha-helix. Scientists presume that this rigid shape, which is stabilized by bonds between hydrogen atoms on the side chains, makes the molecules more soluble in lipid membranes.

In the past, all biologists could say about the transformation of a randomly coiled chain in water into an ordered alpha-helix in the membrane was that, as Pohorille puts it, "somewhere, a miracle happens." Over the last decade, he says, simulations in the Exobiology Branch have suggested that this "somewhere" is near the cell's interface with the outside environment. But fundamental questions, such as whether folding occurs before or after a polypeptide enters the membrane, remained unanswered.





A simple polypeptide molecule immersed in water (blue) approaches and then penetrates a wall of hexane (green). More details...

Anatomy of a Simulation

Researchers studying these questions must resort to computer simulations, since X-ray crystallography, nuclear magnetic resonance, and other standard methods for observing the structures of large biomolecules don't work when the molecules are enmeshed in a cell membrane. In simulations, computers calculate how a system of molecules with properties similar to those of real proteins and cell membranes should change over time under the laws of classical mechanics. In Chipot and Pohorille's simulation, for example, an undecamer called Poly-L-Leucine stands in for longer polypeptides, and a slab of hexane represents the membrane, which in actuality is made up of a double layer of lipid molecules. For studies of folding, hexane is "a fairly good approximation" of the lipid bilayer, Pohorille explains. Both are nonpolar, meaning that their electrical charge is uniformly distributed, so that they change their structure only slightly in response to the electric fields of dissolved molecules. Water molecules, by contrast, are strongly polar due to the fact that their two hydrogen atoms, which carry considerable positive charges, are attached to the central oxygen atom, which carries a balancing negative charge.

Chipot and Pohorille's simulation centered on an aquarium-like box with 1,380 water molecules on one side and 409 hexane molecules on the other. (Like oil and water, hexane and water do not mix.) The simulation began with a randomly coiled undecamer on the water side of the interface, with one end of the molecule about 5 angstroms away from the hexane. Once these initial conditions had been specified, Von Neumann (the NAS Facility's CRAY C90) and the Hewlett-Packard K-9000 workstation belonging to the Exobiology Branch performed their parts of the simulation automatically. A program called COSMOS (COmputer Simulation of MOlecular Systems) calculated the forces acting between each molecule and its nearest neighbors. It then predicted how far each molecule would move in response to these forces over a period of two femtoseconds (a femtosecond is a millionth of a billionth of a second). After calculating each molecule's new coordinates, COSMOS started over, calculating the next set of movements. This iterative process was repeated 25 million times, covering a total of 50 nanoseconds.

A 'Coupled Process'

The outcome of the simulation "was by no means a foregone conclusion," Pohorille says. In fact, the scientists on his team "were taking bets on whether [the undecamer] was going to fold or not," he says. If it did, they wondered, would the event occur at the interface, as predicted by one common model of folding transmembrane proteins, or in the nonpolar hexane? And would the alpha helix arise in a linear way, with folding progressing from one end of the molecule to the other (as Pohorille bet it would), or would the change be less orderly?

The undecamer's complex behavior surprised everyone. As Chipot's animation revealed, the molecule moved to the interface in less than one nanosecond and floated parallel to it for much of the rest of the simulation. Folding and insertion into the membrane then occurred simultaneously, with the hexane providing a helping hand. In water, the hydrogen atoms on the undecamer's side chains are normally bonded with other hydrogen atoms on the polar water molecules, preventing the side chains from bonding with one another to promote folding. At the hexane interface, however, the nonpolar hexane molecules protected the side chains from this competition, allowing the side chains to bond to one another.

But the resultant folding wasn't linear, the simulation showed. Instead, the hydrogen bonds broke and reformed intermittently all across the undecamer. Moreover, the side chains didn't always bond across the distance needed to form an alpha-helix; sometimes the bonded side chains were closer together along the backbone, making the undecamer into a different kind of helix called a 3_{10}-helix.

Finally, as folding proceeded, one end of the undecamer occasionally penetrated into the hexane; whenever that occurred, folding accelerated. When the molecule completely folded into an alpha-helix, it turned from a parallel to a perpendicular position relative to the interface and sank completely into the hexane. Once submerged, it would then suddenly rotate back to the surface.

Pohorille and Chipot (who is now with the Centre National de la Recherche Scientifique in France) concluded from their observations that peptide folding doesn't occur solely on the water-membrane interface. Instead, some penetration into the membrane is required. "Folding helps translocation, and translocation helps further folding," says Pohorille. "It's a coupled process." Furthermore, polypeptides undergoing folding may not neatly wind themselves into typical alpha-helices, but may instead form 3_{10}-helices as intermediates. "People have thought and postulated about this nonuniform folding concept, but this is the first time it's been seen in computer simulations," he adds.

Study Generates 'Much to Think About'

While the study gives researchers investigating terrestrial transmembrane proteins much to think about, the results may also prove interesting for exobiologists, Pohorille believes. "Even the simplest functions of a cell are performed by molecules with precise structures and precise locations in space," he explains. "The challenge in studies of the origin of life is to explain how simple molecules present in the protobiological milieu could organize themselves structurally and spatially to perform these functions." Pohorille continues: "Our simulations demonstrate a clear example of spontaneous structural and spatial organization of such simple molecules–the peptide becomes an alpha-helix and moves from water into the membrane."

Once inserted into the membrane, a molecule like the undecamer would be able to take on basic functions, Pohorille speculates. An influenza virus protein called M2, for example, is the subject of the most recent simulations in the Pohorille group. The four identical alpha-helices in the protein span the membrane of an infected cell and transport protons across it. This is an essential step in a normal cell's energy-gathering process, though the virus carries it out for its own purpose.

"To be cautious, one has to say that what we saw in the undecamer simulation was an example of how proteins may fold at the interface," Pohorille concludes. "However, this also revealed some features about folding that were not known before. It at least shows what is possible."

For more information,

- Ames Exobiology Branch
- send email to Andrew Pohorille.

Video Credits:
"Folding And Translocation Of The Undercamer Of poly-L-leucine Across The Water-Hexane Interface: A Molecular Dynamics Study"
A film by: Christophe Chipot, Michael H. New, and Andrew Pohorille
A joint effort of: Laboratoire de Chimie Théorique and NASA Ames Research Center
Visualization by: Gail Felchle and Vee Hirsch
Narration by: Maria Farmer
Coordinated by: Christophe Chipot
Acknowledgements: Centre National de la Rechreche Scientifique and NASA Exobilogy Program

NASnews
▶ Main Menu
January – February 1999 • Volume 4, Number 1

SGI Origin2000 Operations Plan Unveiled

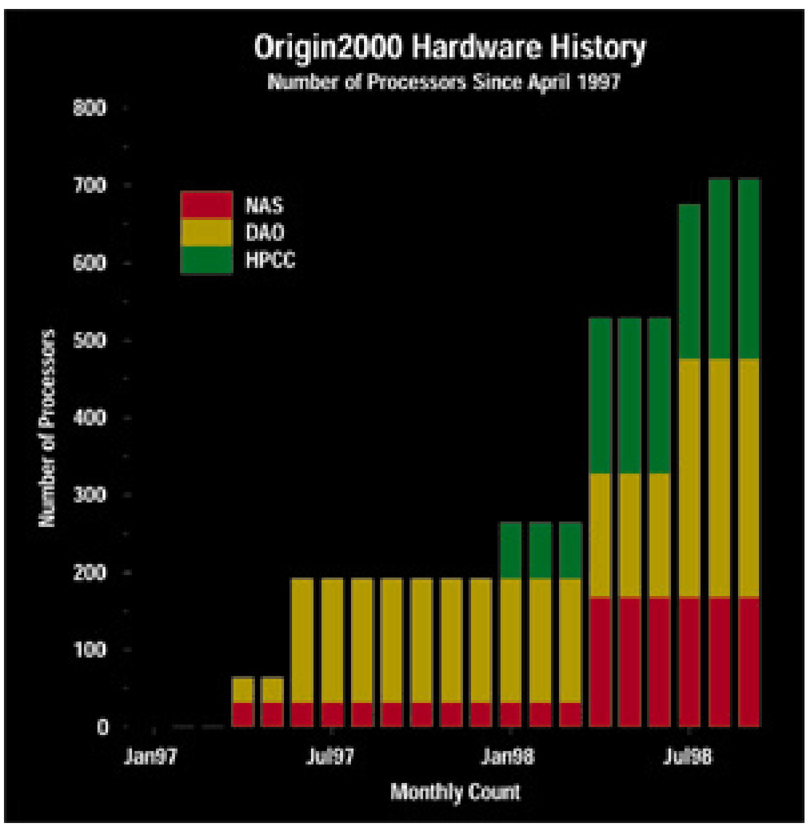
by Wade Roush

Becoming the caretaker for a flock of new parallel-processing computers with hundreds of processors between them might seem like an overwhelming logistical challenge. Yet that was the task faced last fall by the NAS parallel systems group, which maintains five of the 12 Silicon Graphics Inc. (SGI) Origin2000 computers now housed at the NAS Facility (NAS's high speed processor group now maintains the other seven systems). To meet the challenge, the group recently devised a new operations plan covering the Origin2000s managed by the NAS Systems Division.

In interviews with NAS News and at a technical forum last fall, parallel systems group lead Mary Hultquist outlined the plan, which affects those five machines (Turing, Hopper, Steger, Evelyn, and Piglet). The plan's main components cover job management systems, scheduling policies, techniques for minimizing interference between jobs, configuration standardization, and effective performance metrics.

A coordinated strategy is needed, Hultquist explains, because the 1999 operational period (which began October 1) marks the first time that users are being charged for access to designated Origin2000 machines through a formal accounting system like the one governing the Cray supercomputers at NAS. "We were expecting two to three times the number of users [for the new operational period] starting October 1st, given that more than a hundred projects had already been approved for the Origin2000s," Hultquist says. "In preparation, we tried to get the systems to the point where we were confident about the performance of both the hardware and the software."

Current Systems History
The Origin2000 buildup began when the Information Technology (IT) Program and NASA's Data Assimilation Office (DAO) purchased Turing, a 64-processor machine, in April 1997. The DAO bought two 64-processor Origin2000s of its own in June 1997, and five more machines with a total of 136 processors in July 1998. Meanwhile, in January 1998, the High Performance Computing and Communications (HPCC) Program supplied funds for the purchase of Hopper (with 64 processors) and Evelyn (with 8 processors). Then in April, using joint funding from IT and HPCC, the division acquired Heisenberg, Steger, and Piglet, with 264 processors among them. (In November, the parallel systems group joined Heisenberg and Steger into a single machine, which retains the name Steger.) "Lo and behold, we are at nearly 700 processors on 12 systems in a little over 18 months," Hultquist marvels (see graph below).



Two of the systems are currently off-limits to most outside users. Evelyn, formerly used as a front-end system to manage access to the other machines, is now being used as a testbed for the next-generation

NASA Metacenter and the Information Power Grid. "Evelyn is open for some developers right now, and will be open for users as soon as we're confident enough that it's stable," Hultquist says. Piglet, with eight processors, serves as the group's permanent testbed and development system.

That leaves Turing, Hopper, and Steger, which make up what the group calls the NAS "Origin cluster." Turing acts as the cluster's front end, handling logins, compiling, and small debugging jobs, Hultquist explains. Hopper and Steger are compute servers, handling batch and interactive-batch jobs only, and the latter is now available to any user submitting a job requiring up to 256 processors (see sidebar).

New Job Management Policies

PBS, the Portable Batch System (see NAS News, September-October '98), is used for job management throughout NAS computer systems, including the Origins, Hultquist says. The Origin cluster is currently running PBS version 1.1.11, but version 2.0 will be placed on the machines soon.

A single scheduler controls all of the cluster machines. Scheduling and charges are based on nodes (the building blocks of Origin systems) containing two processors and 512 megabytes of memory. User charges are based on the fewest number of nodes required to meet both memory and processor requirements. For instance, a three-processor job needing 8 gigabytes of memory would be charged for 16 nodes, but so would a 31-processor job needing 1 gigabyte of memory, Hultquist explains. Jobs that exceed their allocated resources will automatically be killed, she warns. The group is also working on ways to deter users from using unreasonable numbers of processors or excessive amounts of runtime on the front end, Turing.

Under the new operations plan, scheduling priority on the Origin cluster varies to fit user needs. During "prime time," 5:00 a.m. to 6:00 p.m. Pacific Time on weekdays, shorter jobs take priority. Longer jobs have priority during non-prime time, 6:00 p.m. to 5:00 a.m. Monday through Friday and all day on weekends and holidays. During prime time, jobs can run for no longer than two hours of wall-clock time, while non-prime time jobs can run for up to eight hours.

Home filesystems, which are limited to 200 gigabytes, currently reside on Turing. To keep Turing from becoming overloaded, all home filesystems will be moved to external file servers at NAS, which are better suited for this type of traffic, according to Hultquist. "We are trying to reduce the number of single points of failure. We don't want one thing to blow up everything else," she says.

Beating the Interference Problem

In order to minimize one problem endemic to shared-memory computers -- the tendency of jobs running simultaneously to conflict with each other by attempting to use the same processors or memory units -- the parallel systems group is experimenting with the "nodemask" library provided by SGI and with a "dynamic nodemask" system written at NAS.

Nodemask advises the Origin2000 operating system to keep a given job running only on specified processors. Dynamic nodemask defines a unique nodemask for each job -- in effect, packing jobs into the available processors most efficiently. Nodemask is not a foolproof answer to the interference problem, admits Hultquist, because the program only has advisory powers and does not manage memory. "But," she says, "it's the best we can do right now." With the emergence of the 256-processor architecture, new methods combining nodemask and Miser queues (provided by SGI) are also being explored.

NAS is also considering a more drastic step to beat the interference problem: physical partitioning, the temporary creation of "hard boundaries" between groups of modules in the Origin2000 (a module consists of eight processors). Under physical partitioning, each partition runs its own operating system and its own jobs, just as if it were a stand-alone computer. The Origin2000s are "a new kind of beast," says NAS Systems Division consultant Jim Taft, and partitioning may be the only sure-fire way to prevent interference between jobs and assure that the runtime for a given job is reproducible from run to run.

One drawback to partitioning, according to Hultquist, is that it requires a reboot of the system, often a lengthy process. Moreover, only one job at a time can run in a partition, meaning that fewer jobs are completed daily. "I think Jim and I agree that physical partitioning is not the answer for the entire cluster," she says. "We're looking at taking a piece of it." The new nodemask with Miser queues may be a good compromise, Hultquist notes.

Plan Helps Prepare for 'Coming Load'

The parallel systems group has already implemented one big part of its new operations plan, by standardizing the operating systems and installed software on all of the Origin2000s. "There's nothing worse than finding out that [a particular] piece of software isn't there or that it's configured differently on this system than on that system," says Hultquist. Measures include cloning system disks to maintain software levels across systems, keeping one or more prior levels on hand, posting configuration diagrams and metrics online, and maintaining essential files using the Concurrent Versions System repository, by Cyclic Software . Posted metrics include CPU and memory utilization, number of batch jobs run, average time in queue, gross and net availability, number and type of failures, and time between failures.

With the transition to a production workload, the new operations plan should help the Origin2000s at NAS bear up under the additional load, says Hultquist, adding: "We're trying to make the best out of what we've got, given our resources, our architectures, and the number of users."

NASnews

▶ Main Menu   January – February 1999 • Volume 4, Number 1

Accessing Files on the NAS Facility Metacenters

by **Robert L. Hirsch**

The Cray supercomputers located at the NAS Facility have been grouped into a cluster, or metacenter, so that any system that is not busy can provide user-requested resources and accept and run waiting jobs–even those that were originally submitted on another system. (See "Making the Most of the New Cray Metacenter," May-June '98.)

Users are discovering that, to take advantage of this opportunity for better turnaround, they have to account for the fact that when they submit a job from one execution host, it might run on a different one.

In particular, each system (Eagle, VonNeumann, and Newton1-4, as a group) has its own set of user directories, independent of the others. If, for example, user Lee creates a project subdirectory /u/lee/projectX on Eagle, it exists only on Eagle, not on the other systems; so the /u/lee home directory on VonNeumann (VN) will not contain that subdirectory. A job script that refers to /u/lee/projectX will run as expected on Eagle, but if VonNeumann picks up the job, then the reference will generate a complaint: "No such file or directory."

A technique to take care of this situation is to use a special file path that makes all the different home filesystems visible from any execution host. On Eagle and VonNeumann, for example, the paths look like those shown in Figure 1.

File Path on Eagle	Refers To
/u/lee	Local user home directory
/R/eagle/u/lee	Symbolic link to the local directory
/R/vn/u/lee	Remote mount of the home directory on VN
/R/newton1/u/lee	Remote mount of the home directory on Newton1

File Path on VN	Refers To
/u/lee	Local user home directory
/R/eagle/u/lee	Remote mount of the home directory on Eagle
/R/vn/u/lee	Symbolic link to the local directory
/R/newton1/u/lee	Remote mount of the home directory on Newton1

Figure 1. Special file paths on Eagle and VonNeumann.

Using the Proper Executable

Something else that is necessary to take advantage of running on the first available host is having the right kind of executable for the machine running the job. An executable compiled on Eagle or VonNeumann (both CRAY C90 systems) run on the Newtons (CRAY J90s), so simply pointing to /R/eagle/u/lee/projectX/programX will work on Eagle and VonNeumann but not on any of the Newtons.

A job script that refers to /R/eagle/u/lee/projectX will find the project directory no matter which system it runs on.

One way to take care of this requirement is to prepare a local /u/lee/projectX directory on each machine (Eagle, VonNeumann, and Newton1), and make sure that each of them contains a programX executable that was compiled for that type of machine. Then, if the job script runs /u/lee/projectX/programX, it will be running the local version of the program regardless of the execution host.

The Origin2000 Cluster

The Turing-Hopper-Steger Origin2000 cluster (see "Converting Jobs From the CRAY C90 and J90 Systems to the SGI Origin2000s," November-December '98) has a related situation. The home directories are shared between these two systems, but each of them has a local /scratch1 and /scratch2 filesystem. Files created during a batch job on Hopper in its local /scratch1/lee directory are visible on Hopper, but they do not exist in Turing's;s local /scratch1/lee directory.

The solution is similar to that for the Cray cluster. On Turing, for example, the file path looks like that in Figure 2.

File Path on Turing	Refers To
/scratch1/lee	Local scratch directory
/cluster/turing/scratch1/lee	Symbolic link to local scratch directory
/cluster/hopper/scratch1/lee	Remote mount of Hopper's local scratch

Figure 2. Special file paths on Turing.

Extending local file paths to augmented file paths as illustrated makes it possible to extend your opportunities from using a single system to using any of the systems in a metacenter to run a job.

For more information, contact the NAS scientific consulting group at (650) 604-4444 or (1 800) 331-8737, or send email to **nashelp@nas.nasa.gov.**

NASnews

▶ Main Menu   January – February 1999 • Volume 4, Number 1

Grid, Challenges Sized Up at SC98 Conference

by Wade Roush

Returning to the city where it originated ten years ago, the annual IEEE/ACM supercomputing conference gave thousands of researchers an inside look at the latest commercial products and academic and government research in high-performance computing, and at recent experiments using networked grids of such systems. As its title promised, "SC98: High Performance Networking and Computing," held November 7-13 at Orlando's Orange County Convention Center, highlighted the growing inseparability between big iron and bandwidth -- that is, between supercomputers and the network connections and "metacomputing" software that allow them to work together.

Dozens of demonstrations and technical presentations at the conference proved that systems which enable distributed heterogeneous supercomputing have advanced well beyond the blueprint stage. In a demonstration of the Globus metacomputing toolkit at Argonne National Laboratory's exhibit booth, for example, researchers conferred with distant colleagues about a visualization of a micromachinery part that had been imaged just minutes before back at Argonne, while simultaneously participating in a massive battlefield simulation running on four separate computers in three states.

Another highlight of the SC98 exhibit hall was NASA's own booth, where researchers from six of the agency's field centers demonstrated more than 70 projects dealing with scientific simulation and visualization and other applications of high-performance computing and networking (for details, see the official NASA SC98 web site at http://www.nas.nasa.gov/Services/Conferences/SC98).

"NASA participation in SC98 was outstanding and clearly provided a highly professional view of NASA research, on a par with the booths provided by the computing industry participants," said Lee Holcomb, NASA's chief information officer. "The display provided an opportunity for all NASA centers to participate in a single booth--providing a view of many of our emerging information technologies key to [NASA Administrator] Dan Goldin's vision of the Intelligent Synthesis Environment."

NASA joined more than 120 exhibitors whose leading-edge science and technology were presented. This look at highlights from some of the conference's technical sessions reveals some common issues among industry, government, and academic participants. Through many of the papers, invited talks, and expert panels, a few urgent, interlinked questions kept surfacing: Which scientific problems are beyond the ability of single machines or single institutions to solve, yet most in need of solution? What are the prospects for the technical advances that will make solving these problems possible? Can a seamless, global grid linking existing high-performance computers, massive data archives, and unique scientific instruments help, and how should it be constructed and operated?

Need for Increased Performance

In an invited "Challenges of the Field" talk, Department of Energy (DoE) undersecretary Ernie Moniz offered the plainest statement of the need for ongoing

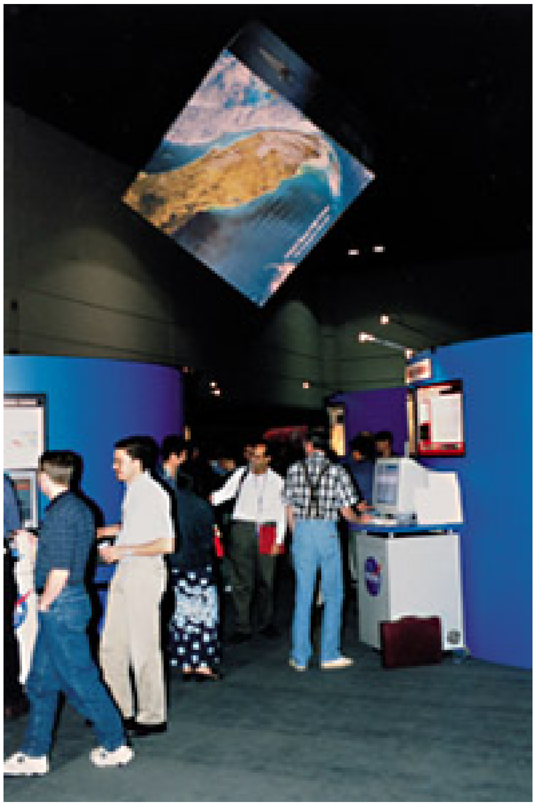
Jet Propulsion Laboratory researcher Charles Norton (left) describes a workstation-based demonstration to a booth visitor.

research in high-performance computing and networking.

DoE-supported studies of complex problems in the physical sciences such as nuclear fusion and fission, solar energy, global climate change, quantum physics, and advanced materials "all require continual advances in computer power," Moniz said. The department's biggest computational challenge, created by the international ban on underground nuclear tests, is to model the behavior of aging nuclear weapons. For that problem, the DoE needs computers capable of hundreds of teraFLOP/s (trillions of floating point operations per second) by 2004, Moniz said.

Exotic Applications

Other needs are even more immediate. High-performance computing could make or break the Spallation Neutron Source (SNS), a $1.3 billion facility under construction at Oak Ridge National Laboratory (ORNL), according to physicist John Cobb, a member of ORNL's Computing, Information and Networking Division. In the session "Exotic Applications as Drivers of Increased Performance," Cobb explained that in the SNS, neutrons knocked loose from a target by a burst of accelerated protons will penetrate dense materials and give information not revealed by other imaging techniques.

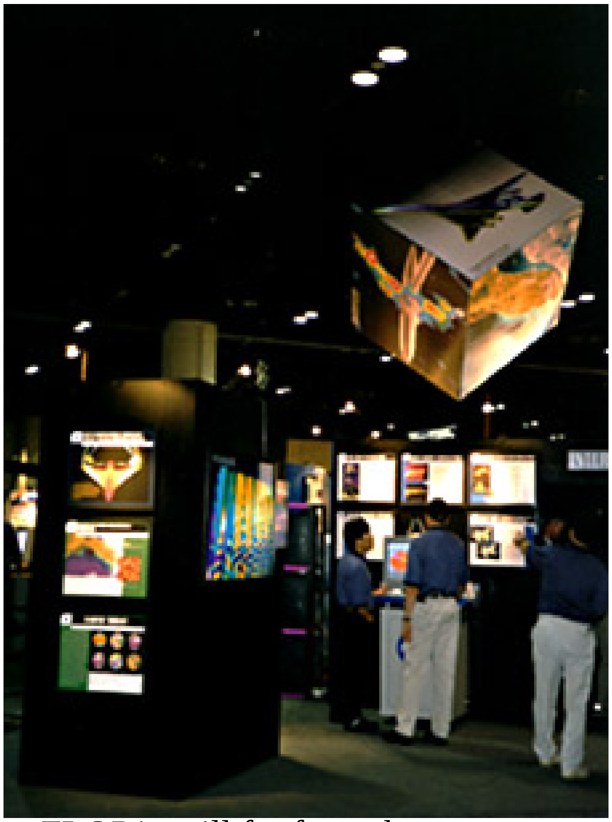


The symbolic centerpiece of the NASA exhibit (above) at SC98 in Orlando was the NASA Cube, a rotating fabric-and-metal structure bearing scientific visualizations and Web addresses from the six participating NASA centers. The image shown here is a true-color photo of Florida captured by NASA's SeaWiFS satellite, a global ocean color monitoring mission managed by Goddard Space Flight Center.

However, the fraction of protons that go astray must be kept low to avoid radioactivity, and the target itself must be able to withstand the immense energy of the 1-microsecond, 1-megawatt proton burst that the SNS will produce. Building a test accelerator to try various shielding materials would be too expensive, so project physicists must model the different possibilities computationally. "We 'cut copper' in 18 months," Cobb said. "We will gladly sacrifice a performance factor of two if you can give us an answer within six months."

It was only half in jest that Saul Perlmutter, an astrophysicist at Lawrence Berkeley National Laboratory (LBL), remarked in another talk that "the fate of the universe depends on super-computers." Perlmutter described an international project he is leading to determine, from measurements of supernova explosions, whether the universe will expand forever or will eventually come to a stop and then collapse. By comparing many images of the same chunk of the sky recorded over several weeks, Perlmutter's group can single out supernovas in distant galaxies and track changes in their luminosity over time. These patterns allow researchers to calculate the supernovas' intrinsic brightness. Comparing a supernova's intrinsic brightness to its apparent brightness gives its distance, which because the speed of light is finite is also a measure of how long ago the supernova occurred. The redshift or stretching-out of the supernova's light, moreover, indicates how quickly its galaxy is receding from our own.

Perlmutter's group used a CRAY T3E supercomputer at the National Energy Research Scientific Computing Center at LBL to compare dozens of supernova observations against each other and against theoretical simulations. They reached an unexpected conclusion: Supernovas that occurred 7 to 14 billion years ago have a lower redshift than would be expected if the universe were expanding at a constant rate, while those that occurred within the last 7 billion years have a higher redshift. The implication, which Perlmutter says "is making cosmologists rather uncomfortable," is that the expansion of the universe was slowing at one time but is now accelerating and will probably continue forever. (In its December 18, 1998, issue, *Science* magazine named Perlmutter's team's findings the "Breakthrough of the Year." See the LBL news brief.)



Call For Computer Architecture Research

While conference participants described many impressive new uses for supercomputers, the meeting was also pervaded by anxiety that the fundamental technical breakthroughs needed to make further advances possible are overdue. Machines developed under the DoE's Accelerated Strategic Computing Initiative have recently reached sustained performance of 1 to 4 teraFLOP/s, still far from the hundreds of teraFLOP/s called for by Moniz. Several speakers predicted that a radical advance comparable to the switch from vacuum tubes to transistors in the 1950s, or the development of the vector supercomputer by the late Seymour Cray in the 1970s, will be needed to close the gap.

In a panel discussion titled "Is Architecture Research Dead?", Tilak Agarwala, director of server architecture and system strategy for IBM, predicted that the limits of miniaturization of the silicon chip will be reached within several years, with no clear alternative material in sight. Burton Smith, chairman and chief scientist of Tera Corp., remarked that while techniques such as instruction-level parallelism are allowing CPUs to process more bits simultaneously, latency, the time required to retrieve those bits from memory, is a "narrowing bottleneck." But other panelists predicted that incremental improvements in existing technologies will offer routes around these problems over the next decade. Placing more CPUs on a single chip and increasing the bandwidth of communications between chips using optical interconnects and wavelength division multiplexing will be among these improvements, said Steve Wallach, co-founder of Convex Computer Corp.

All panelists agreed, however, that the federal government and today's dominant computer makers, such as Dell Computer, are spending too little on basic research for next-generation computing architectures.

"The U.S. government has completely dropped the ball" on funding for basic computer science research, Smith said. Discussion of ways to reverse that situation dominated several sessions at the conference.

The President's Information Technology Advisory Committee (PITAC; see http://www.ccic.gov/ac/) held a "town hall" meeting to gather comments on its draft report, issued last July, calling on the federal government to double the amount it spends on computer science to $2 billion per year. Advisory committee chair Ken Kennedy of Rice University said that most of the new money should be spent on long-term fundamental research on new computer architectures and more efficient software. "Our goal is to increase fundamental research at every agency," said Kennedy.

Programming the Grid
How to make the most of existing computational capabilities was a recurring question in many other technical sessions and research exhibits at the conference. Tying supercomputers together into "information grids," so that the overall supply of CPU time can be matched more efficiently to demand, was the answer promoted by many SC98 speakers and panelists. "We urgently need research on the nature of the scalability [of high-performance computing]," said Larry Smarr, president of the National Computational Science Alliance and director of the National Center for Supercomputing Applications at the University of Illinois, Urbana-Champaign.

Argonne National Laboratory's Ian Foster described one approach to scaling up supercomputing. MPICH-G, a "grid-enabled" version of the parallel programming standard known as the Message Passing Interface, modifies existing parallel software to run on multiple independent platforms. The system locates hosts, stages executables, submits multiple jobs, authenticates users, chooses a communications protocol, and controls output. "Our specific contribution is to replicate familiar single-machine programming models in a heterogeneous environment of high-performance computers," said Foster. The result is that existing software can be migrated to grid environments more easily.

GUSTO Places First in Competition

First place in the High Performance Computing Challenge, an annual competition at the SC conferences for the best poster presentation or demonstration on leading-edge computing and networking techniques, went to a multi-institutional team for a related demonstration, titled "Innovative Wide-Area Applications on the GUSTO Grid Testbed." GUSTO, the Globus Ubiquitous Supercomputing Testbed Organization, was created to allow large-scale application experiments using Globus, a software toolkit that underlies MPICH-G and is designed to enable resource allocation, remote process creation, and uniform high-speed communication among distributed heterogeneous supercomputers. (Ames Research Center is a founding member of GUSTO.) In their demo, Foster and Carl Kesselman, co-principal investigators of the Argonne-based Globus Project, worked with collaborators at Argonne, the University of Southern California, Caltech, and other institutions to show that Globus can run multiple large-scale supercomputing applications simultaneously.

While judges watched, Globus juggled three computationally intensive problems, beginning with a collaborative visualization and discussion session on X-ray photographs of a silicon micromachinery part taken at Argonne's Advanced Photon Source. Globus also monitored a distributed, interactive war game involving 4000 tanks, aircraft, and other vehicles running on computers at Argonne, the University of California at Santa Cruz, and the Maui Supercomputing Center in Hawaii, while simultaneously overseeing a "high-throughput" quantum Monte Carlo simulation using a mix of six supercomputers and 50 workstations in a Condor pool at the University of Wisconsin.

At the NASA booth, meanwhile, NAS researcher Rupak Biswas and others also used Globus to demonstrate a version of the OVERFLOW computational fluid dynamics code that has been modified to run across multiple supercomputers. NASA's Information Power Grid project (see NAS News, September-October '98) has as its goals to advance systems like Globus so that they can be used to coordinate the numerical solution of aerospace design problems across computers and large databases at many NASA sites.

In the long run, this kind of distributed heterogeneous computing will be a resource for society at large, not just for the research establishment, some panelists predicted. "Eventually, billions or even trillions of devices will be connected to the Internet," said Mary Meada of the Defense Advanced Research Projects Agency's Information Office. "The network must scale up to this demand anytime, anywhere, securely and reliably." In the meantime, software for scaling up supercomputing, cutting-edge research on supercomputer architecture, pressing new applications for high-performance computing, and other hot topics from SC98 will be explored in upcoming issues of *NAS News*.

NASnews
January – February 1999 • Volume 4, Number 1 ▶ Main Menu

256-processor SGI Origin2000 System Announced at SC98 Conference

by Wade Roush

On November 11, Silicon Graphics Inc. (SGI) gave center stage at its SC98 exhibit booth in Orlando to NAS Systems Division consultant Jim Taft, who announced to a packed audience of onlookers the recent birth of the world's largest shared-memory parallel supercomputer.

Just days before, Taft told the audience, engineers from SGI had joined Steger and Heisenberg, two existing 128-processor Origin2000 computers at the NAS Facility, into a single, fully functional system. The result is a system with 256 processors that can directly and globally address all 64 gigabytes of memory in the system, setting a new record in its class of cache-coherent, Non-Uniform Memory Access (ccNUMA) computers. SGI configures standard Origin2000s with up to 128 processors.

The new machine, which kept the name Steger, performed remarkably well "right out of the box," Taft said. It completed a 35-million-point aerodynamic simulation using the OVERFLOW-MLP computational fluid dynamics code twice as fast as its 128-processor parents, with no modifications to the system or application software.

Two 128-processor Origin2000 systems at the NAS Facility have been joined into a single machine called Steger using a custom 256-processor meta-router interconnect. Cables cross between the two formerly independent machines over two bridges, visible here. *Photo by Tom Trower.*

Steger is configured with 64 gigabytes of system memory and 1.33 terabytes of disk space, making it by far the most powerful supercomputer at NAS. Taft explained that the Origin2000s' so-called "hypercube" architecture guarantees that even with 256 processors, memory-resident data can be passed between any two processors with no more than six quick intermediate hops, allowing Steger to digest the 35-million-point aerodynamic problem at a rate of 20.1 gigaflops, or 7.5 seconds per time step.

Taft's statistics show that Steger solved this problem 4.4 times faster than the 16-CPU CRAY C90 vector supercomputers at NAS, at about one-fifth the purchase price, giving it a cost-performance ratio of 21.5x over the C90. Under charging algorithms currently in place at NAS, users running OVERFLOW-MLP on Steger would see their computing costs drop by a factor of three, Taft says.

Taft's final comments looked to the future: "The real benefit and excitement of the 256-CPU Origin system is that NASA scientists can now solve problems that are an order of magnitude larger than previously possible on the C90 systems, at a fraction of the cost."

For more information on the 256-processor Origin2000 at the NAS Facility, send email to jtaft@nas.nasa.gov.

To Main Article...

In this issue...

Credits

Executive Editor: Bill Feiereisen

Editor: Jill Dunbar

Senior Writer: Wade Roush

Contributing Writers: Robert L. Hirsch

Image Coordinator: Eunah Choi

Special thanks to: Richard Anderson, Eunah Choi, Bob Ciotti, Judy Conlon, Archemedes DeGuzman, James Donald, Gail Felchle, Jim Fischer, Ian Foster, John Hardman, Vee Hirsch, Larry Hofman, Mary Hultquist, Bill Johnston, James Jones, Randy Kaemmerer, Art Lazanoff, Terry Nelson, Andrew Pohorille, Marcia Redmond, Cathy Schulbach, Jim Stobie, Jim Taft, Mark Tangney, Eugene Tu, Michael Wilson, Patti Yamakido