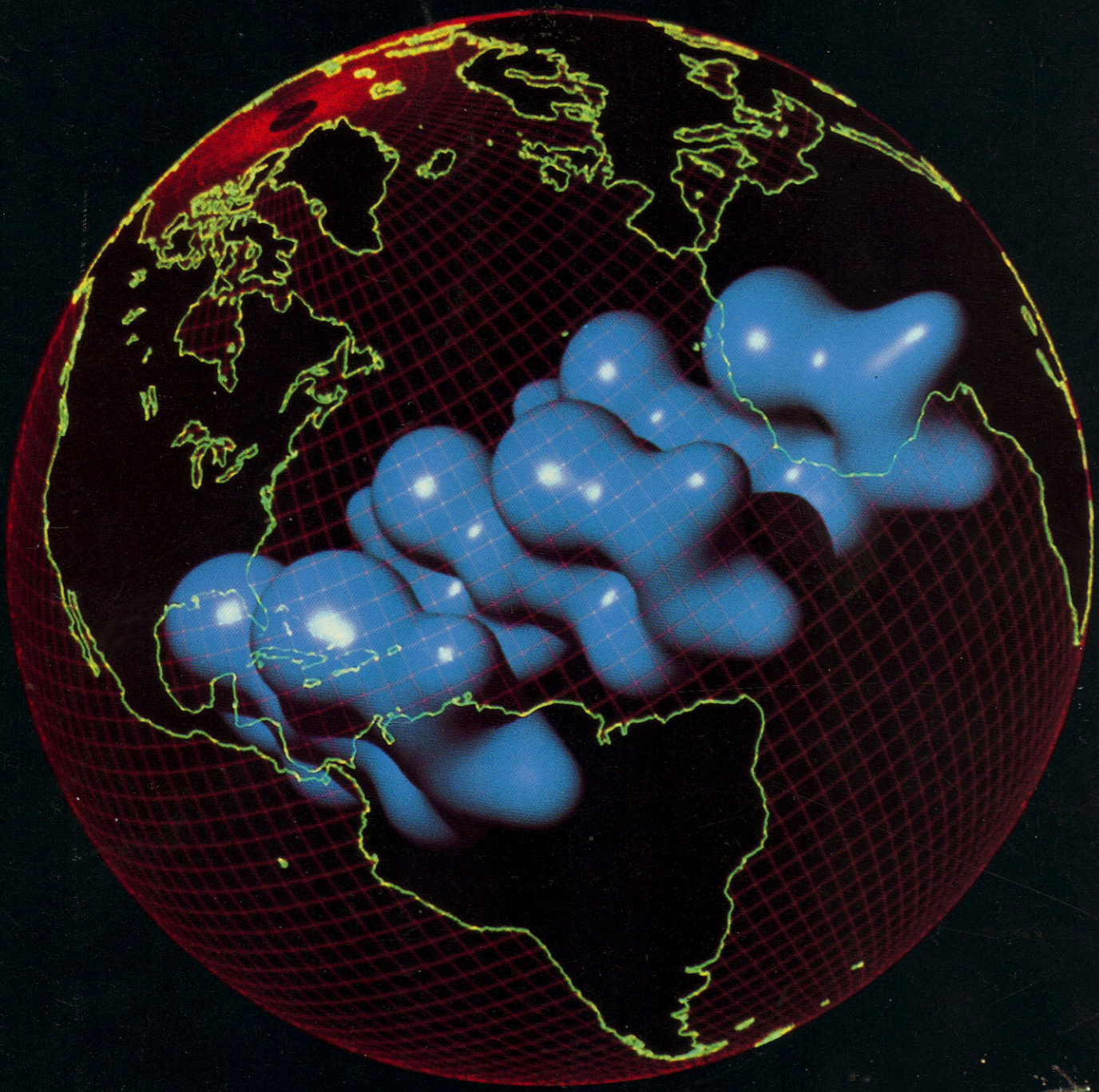
CRAY CHANNELS

SUMMER 1988 · A CRAY RESEARCH, INC., PUBLICATION

A universe of applications



Announcing
the Cray Extended Architecture series of computer systems

CRAY CHANNELS

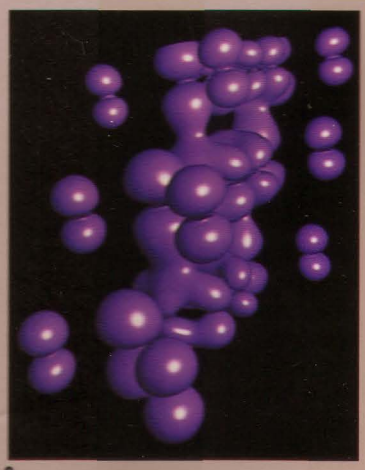
In this issue

A decade ago few people would have predicted the impact that supercomputers would have on our lives. Today, researchers around the world are applying Cray systems to diverse and complex problems. Cray systems are helping users design and test automobiles, search for natural resources, investigate atomic nuclei and stellar supernovae, predict weather, and build aircraft — to name only a few applications. Cray systems give new meaning to the term "general-purpose" by demonstrating their outstanding value in a wealth of disciplines.

In this issue of CRAY CHANNELS, scientists from our Industry, Science & Technology Department assess the impact Cray systems are having on their areas of expertise. They examine the broadening use of supercomputers within traditional application areas and reveal emerging trends. In this issue we also introduce the Cray Extended Architecture series of computer systems, a new line of supercomputers that provides an upward migration path through the use of CRAY Y-MP technology, along with larger memory options and improved price/performance over the CRAY X-MP series. The regular departments offer new system orders and some unique Cray system applications, including modeling dolphin sonar and mapping the skies.

With just one Cray system, users can solve a wide variety of problems. For example, engineers in the auto industry are using Cray systems to test crashworthiness, to optimize components, and to compute complex internal and external airflows. In the aerospace field, Cray systems are being used to design engines, to analyze structural responses to stress, to design electronics systems, and to simulate wind-tunnel testing. Whether they are used to design vehicles or predict tomorrow's weather, supercomputers are positively affecting our lives. Automobiles and aircraft are safer, quieter, and more reliable; electronic components are smaller, faster, and more reliable; weather predictions are more accurate; and our understanding of the world is more complete. Areas such as financial analysis and artificial intelligence, among others, show promise of evolving into everyday applications that will benefit from high-performance supercomputing. Though the impact of supercomputers has been phenomenal in the last decade, it is clear that the changes have only just begun.

Features



8

Departments

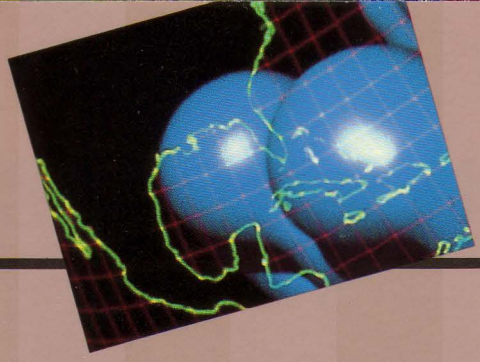
CRAY CHANNELS is a quarterly publication of Cray Research, Inc., intended for users of Cray computer systems and others interested in the company and its products. Please mail subscription requests, feature story ideas, and news items to CRAY CHANNELS at Cray Research, Inc., 608 Second Avenue South. Minneapolis, MN 55402.

Volume 10, Number 2

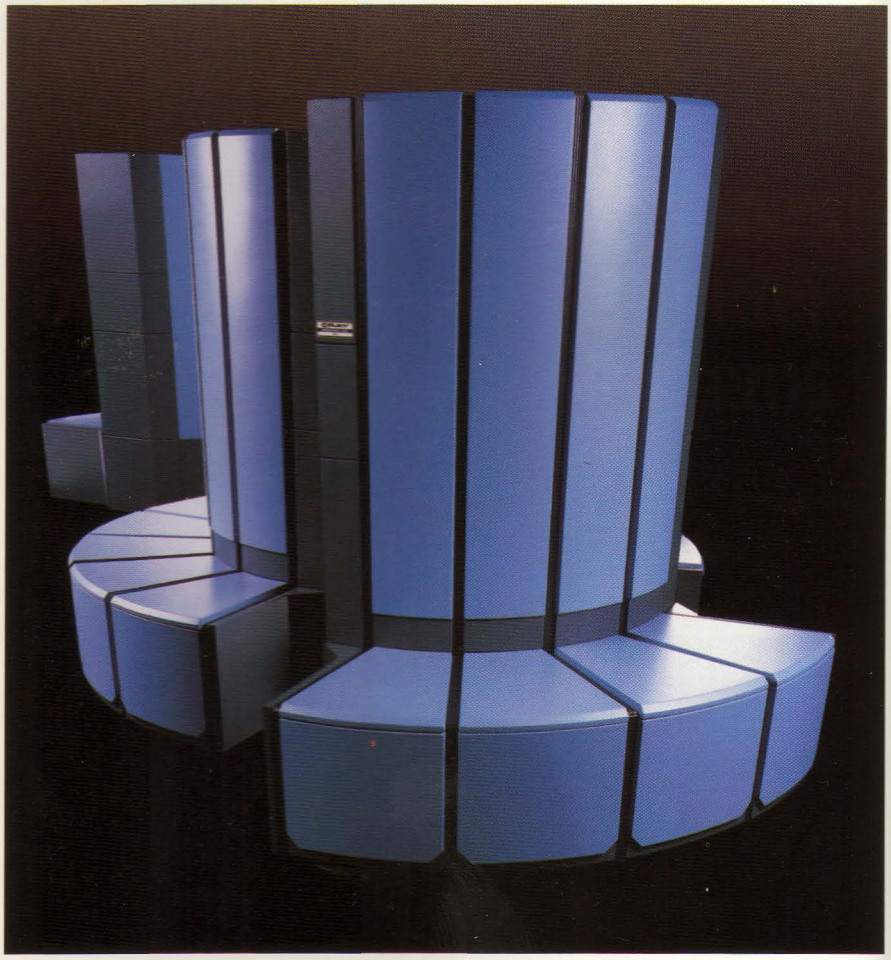
Editorial staff
Tina Bonetti
Ken Jopp
Elizabeth Knoll
John Swenson

Design and production
Barbara Cahlander
Eric Hanson
James Morgan
Cynthia Rykken

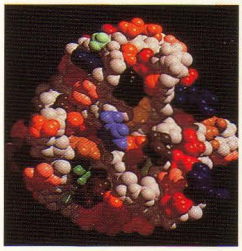
On the cover: Researchers use Cray systems for a universe of applications — from examining molecules to mapping global climate patterns. This illustration merges a grid of Earth used for climate and weather research with a molecule of tetrathiafulvalene, a component of an organic, electrically conductive crystal. Global grid courtesy of Tom Bettge, National Center for Atmospheric Research. Molecular model provided by Erich Wimmer, Cray Research, Inc.



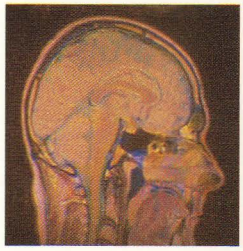
Introducing the Cray Extended Architecture series of computer systems



From bacterial enzyme modeling (top) to medical imaging (bottom), Cray EA systems can solve a universe of applications.



(Courtesy of Michael James, University of Alberta)



Cray Research has combined the advanced architecture of the CRAY Y-MP system and the field-proven technology of the CRAY X-MP series to create a new line of computer systems — the Cray Extended Architecture (EA) series. The result is a broad product family that provides users with a migration path to Cray Research's latest technology products, and provides larger memory options and improved price/performance over the CRAY X-MP series.

"The Extended Architecture systems have a threefold significance," said John Rollwagen, chairman and chief executive officer of Cray Research. "First, they allow our users to migrate from the CRAY X-MP systems to the CRAY Y-MP architecture, while preserving their investments in application software. Second, their larger memories allow users to solve even more complex problems. And third, they are priced at a level that represents outstanding value to our customers."

The new series implements the CRAY Y-MP and CRAY X-MP instruction sets, allowing users to run programs developed for CRAY X-MP and CRAY Y-MP systems. Users can switch processing modes (X-MP or Y-MP) through commands to the system. Like the CRAY X-MP series, the most widely used supercomputer line to date, the EA series will solve a diverse range of problems, from structural optimization to electronic component design. Users can perform weather analyses, compute the dynamics of flow around objects, study molecular dynamics, and predict the flow of fluids in petroleum reservoirs — to name only a few applications.

Balanced power

Cray Research's Extended Architecture systems offer up to four times the memory of existing CRAY X-MP systems at approximately the same prices. With the CRAY X-MP EA/216, Cray Research offers doubled processing power at a lower cost, as compared to the CRAY X-MP/116. CRAY X-MP EA models are available with one, two, or four processors and with memory sizes of 4, 16, 32, or 64 million words, while the CRAY Y-MP/832 system offers eight processors and 32 million words of memory. Most models in the Extended Architecture series can be upgraded to include additional CPUs and larger memory. The same high-performance software runs on all models of the Extended Architecture series.

The system design of the Extended Architecture series is balanced carefully to deliver optimum overall performance. It builds on the best features of the CRAY X-MP systems, including fast long and short vector processing, high performance I/O, gather/scatter and compressed index vector instructions, flexible hardware chaining, and the ability to connect to a wide variety of other vendors' systems. The CRAY X-MP EA systems also represent the successful blending of 2500-macrocell arrays with 16-gate array technology, resulting in reduced component count, higher overall component integration, and improved reliability. The end result of EA system design is exceptional speed and high overall system throughput in a balanced and practical computing system.

Multiprocessor configurations in the EA series allow users to employ multiprogramming, multiprocessing, and multitasking techniques, which can reduce dramatically the time needed to solve the largest and most complex computational problems. The multiple central processing units may operate independently and simultaneously on separate jobs, or may be applied in any combination to operate jointly on a single job for faster program turnaround. Multiprocessing and vector processing combine to provide a substantial increase in computational performance over conventional scalar processing techniques.

	CRAY Y-MP/8	CRAY X-MP EA/4	CRAY X-MP EA/2	CRAY X-MP EA/1	CRAY X-MP EA/se
Mainframe					
CPUs	8	4	2	1	1
Memory (64-bit words)	32M (bipolar)	16, 32, or 64M	16, 32, or 64M	16, 32, or 64M	4 or 16M
6-Mbyte/sec channels	8	4	4	2 or 4	1
100-Mbyte/sec channels	8	4	2	1 or 2	1 or 2
1000-Mbyte/sec channels	2	1 or 2	1	1	N/A
Solid-state Storage Device					
Memory size (Mwords)	128, 256, or 512	32, 128, 256, or 512	32, 128, 256, or 512	32, 128, 256, or 512	N/A
I/O Subsystem	*One IOS	Two IOSs			
I/O processors	4	7 or 8	2, 3, or 4	2, 3, or 4	2, 3, or 4	2 or 3
Disk storage units	2-24	4-48	2-24	2-24	2-24	2-8
Magnetic tape channels	1-8	2-16	1-8	1-8	1-8	1-4
Network interfaces	1-7	2-14	1-7	1-7	1-7	1-3
Buffer memory (Mwords)	4-32	8-64	4-32	4-32	4-32	2-32

N/A signifies option is not available on the model.
*One IOS is standard with the CRAY Y-MP/832 system; a second IOS can be configured as an option.

Extended Architecture system configuration options.

Peripheral devices

Cray Research also has introduced a new model of SSD solid-state storage device. The new SSD-5I storage device is a 128-million-word SSD located within the cabinet of the Cray I/O Subsystem (IOS). By sharing the IOS cooling and power supplies, cost to the user is reduced compared to stand-alone models of the same capacity; floor space requirements are reduced as well. SSD models with 128, 256, or 512 million words of memory are available in stand-alone cabinets. An SSD model with 32 million words of memory also can be configured within the IOS cabinet that accompanies the mainframe. The SSD is not available for use with CRAY X-MP EA/14se or CRAY X-MP EA/116se systems.

Also complementing the computational power of the Extended Architecture series is Cray Research's top-of-the-line DS-40 disk subsystem. Its four disk storage units provide a total storage capacity of 20.8 Gbytes, and each drive offers sustained transfer rates of up to 9.6 Mbytes/sec. Cray's DD-49 and DD-39 disk drives also are supported, each drive offering a storage capacity of 1.2 Gbytes and sustained transfer rates of 9.6 and 5.9 Mbytes/sec, respectively.

Networking products also are available to integrate Cray Extended Architecture systems into existing computing environments. Cray Research offers the HSX-1 high-speed external communications channel, providing full duplex point-to-point communication (up to 100 Mbytes/sec) with very fast devices over distances of up to 70 feet (22 meters). Cray Research also offers network interfaces that compensate for differences in channel width, word size, logic levels, and control protocols between other vendors' front-end systems and Cray Extended Architecture systems. The Network Systems Corporation (NSC) HYPERchannel, Computer Network Technology LANlord, and similar network adapters enable Extended Architecture systems to communicate with multiple front-end computer systems. Digital Equipment Corporation offers a VAX Supercomputer Gateway that provides a high-performance direct connection between the VAXcluster environment and Cray Extended Architecture systems.

Software

All Cray Extended Architecture computer systems come with state-of-the-art software including UNICOS, which offers a widely accepted program development environment that complements the advanced computational power of Extended Architecture systems. The Cray operating system COS also is supported, allowing COS programs developed in the earlier CRAY X-MP environment to use up to 16 million words of memory on an Extended Architecture system. The Guest Operating System feature of COS allows users to run UNICOS simultaneously within COS. Standard software also includes an automatic scalar optimizing and vectorizing Fortran compiler, automatic scalar optimizing and vectorizing C and Pascal compilers, extensive library routines, program- and file-management utilities, debugging aids, a Cray assembler (CAL), PSL LISP, and a wealth of third party and public-domain applications. In addition, Ada, Common Lisp, and SIMSCRIPT compilers are under development.

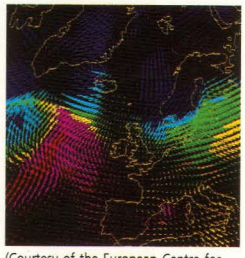
Parallel processing is supported by Cray Research software and hardware, and will be enhanced with a future release of Cray Research's latest generation Fortran compiler, CFT77. The new CFT77 release automatically will recognize parallel aspects of users' Fortran programs and automatically allocate those program segments for independent and concurrent execution on the system's processors. All multiprocessor Extended Architecture systems, including the CRAY Y-MP system, will be able to take advantage of this "autotasking" feature. CRAY-2 systems will support autotasking as well.

A commitment to value

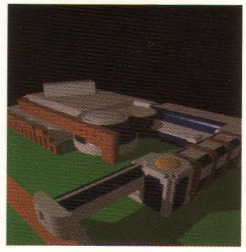
Cray Research has developed a comprehensive array of support services to meet customer needs. During site planning and throughout the lifetime of an installation, hardware engineering and system software support are provided locally and through technical centers throughout the company. All components undergo strict inspection and reliability testing prior to shipment.

The Cray Extended Architecture series represents the best of Cray Research technology and workmanship. It also marks another step in the company's commitment to providing product enhancements, efficient upward migration, and real value to customers.

Cray EA systems are ideal for rapidly processing vast amounts of data required for weather and climate analysis (top). Cray EA systems can generate architectural models (bottom) and "fly-through" movies of structures.



(Courtesy of the European Centre for Medium-Range Weather Forecasts)



Crashworthiness and beyond

Structural analysis on Cray systems

Gregory Clifford, Cray Research, Inc.

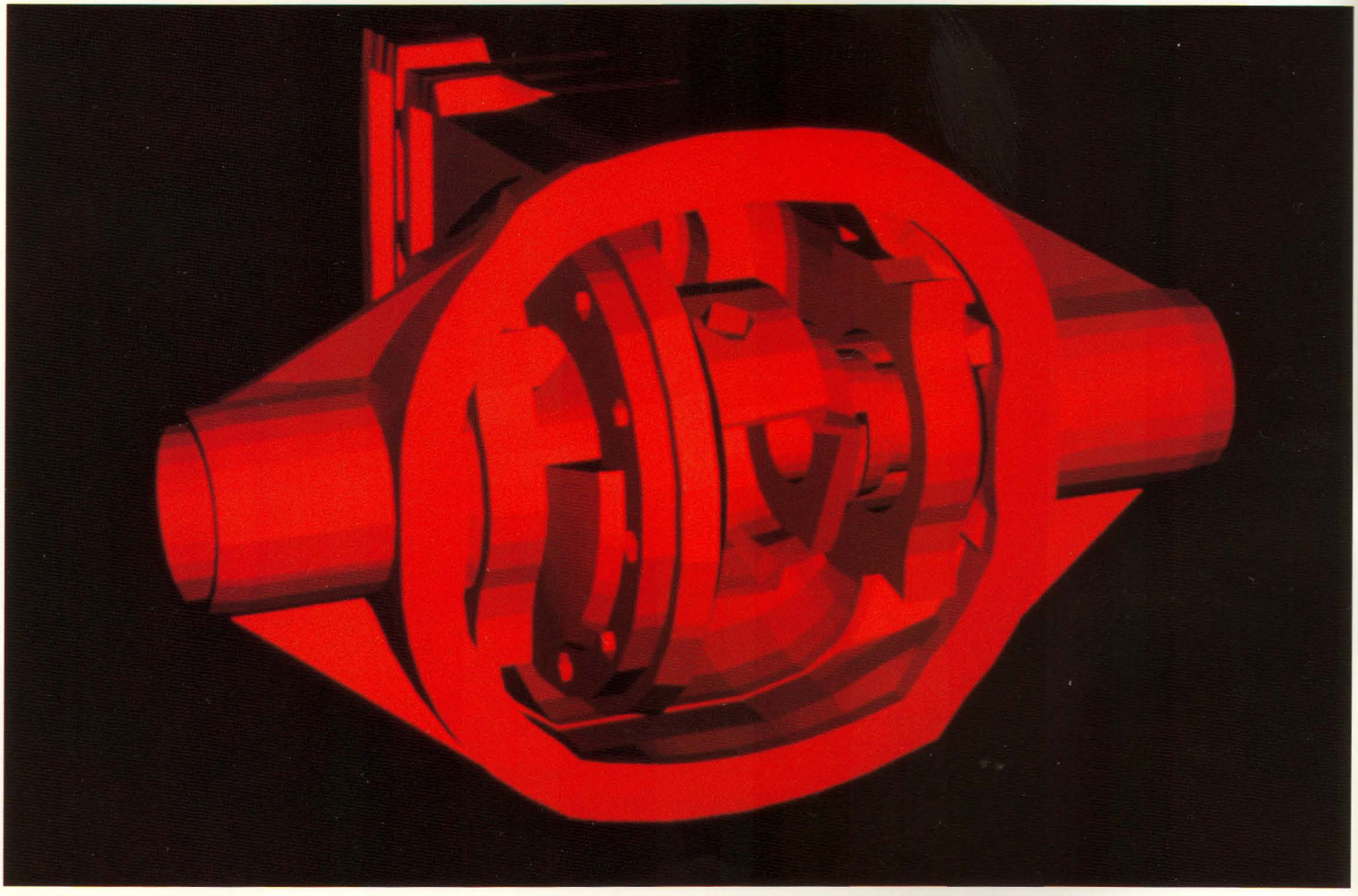
Computational structural analysis is a cost-effective test tool used by designers and engineers in many industries. It has been part of standard practice in the automotive and aerospace industries since the 1960s, and now is used by virtually every manufacturing industry to ensure the structural integrity of products and components. The long history, variety of analyses that can be performed, and established nature of the structural field make it a unique supercomputer application area.

Software support

A structural analysis computer environment includes a well-developed combination of hardware, software, and engineering experience, which represents an investment of millions of dollars. Any new component introduced into such an environment must be able to interface with the existing system without interrupting productivity. Cray computers are versatile enough to integrate easily into virtually any computer environment. Also, the large number of customers performing structural analysis on Cray systems has proven invaluable when evaluating new customer needs. Almost every possible combination of hardware and software has been accommodated at existing Cray sites. New customers can gain from the experiences of others, learning from previous obstacles that have been encountered and overcome.

Cray Research works closely with software vendors to ensure that the applications software Cray users need is available and running efficiently. All of the major analysis codes are available and being used on Cray computer systems. For example, five versions

Automobile differential structure model. Credit: General Motors.



of NASTRAN run on Cray systems. Most analyses are performed using large, general-purpose, finite element programs such as NASTRAN, ABAQUS, and ANSYS, which have been developed by established third-party vendors. The size of the programs and their general-purpose nature require balanced computer performance. Not only is a fast CPU required, but also fast and efficient I/O capability.

Increasing demands

During the past 25 years, the size of computer models has increased steadily, making more accurate analysis possible. However, the computer time required for analysis also increases with model size. Problems that would have been impractical a few years ago now are executed routinely in production environments. This demand for increasingly accurate analysis, and hence more computing power, has made Cray systems necessary tools in large analysis environments.

An example of increasing demand for computing power is illustrated in the computer models shown in Figure 1. The top figure is a model of the Space Shuttle solid rocket motor (SRM) joint from Morton Thiokol. This model contains 5236 nodes and represents a one-degree segment of the cylindrical joint. This static analysis, containing GAP elements, took 2400 CPU seconds on a CRAY X-MP system running the ANSYS analysis package. Based on this experience, more complex models were created, such as the middle image in Figure 1, which is a six-degree segment of a cylindrical joint containing 10,472 nodes. The ANSYS code required 18.4 CPU hours on the CRAY X-MP system to complete analysis of this fully nonlinear model, which included plasticity. The initial linear iteration took nine minutes of CPU time on a CRAY X-MP system, as compared to 115 minutes on an IBM 3084 computer. A total of 180 iterations were needed for convergence during 19 load steps.

This analysis, a pioneering effort in Morton Thiokol's use of supercomputers, demonstrated the value of detailed analysis. A more recent model represents a 12-degree segment of the joint (bottom). The analysis is estimated to exceed 100 CPU hours on a CRAY X-MP system. Although 100 hours is a large amount of computer time, computer simulation is the only means to obtain the required information, and no other system could have provided the information in less time.

Crashworthiness simulation

The automotive industry has a tremendous need to reduce the number of tests necessary to verify vehicle crashworthiness and meet government and internal requirements. The goal of crash analysis is to compute the response of an automobile hitting a rigid wall at 35 mph, a very sophisticated analysis that involves complicated material and geometric nonlinearities. The analysis begins at the moment of impact and continues until the velocity of the vehicle is zero.

Automotive crash simulation has progressed dramatically in the past few years. Three years ago, a full-body crash simulation was considered intractable. In July 1986, a full-body simulation was completed successfully for Peugeot using a CRAY X-MP system

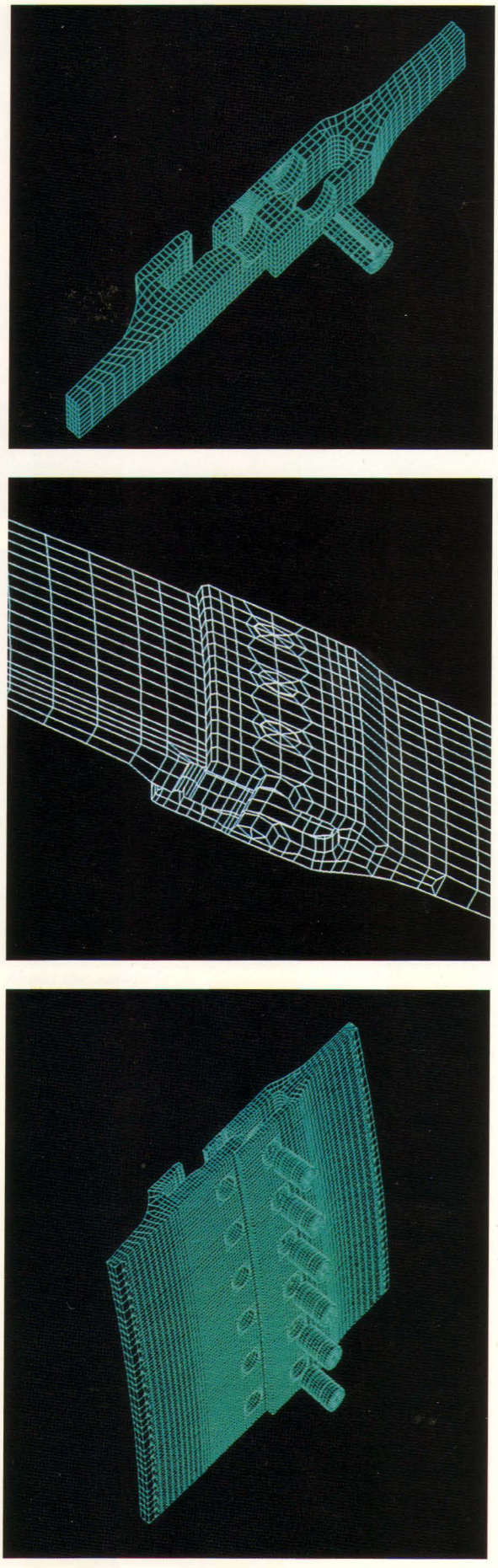
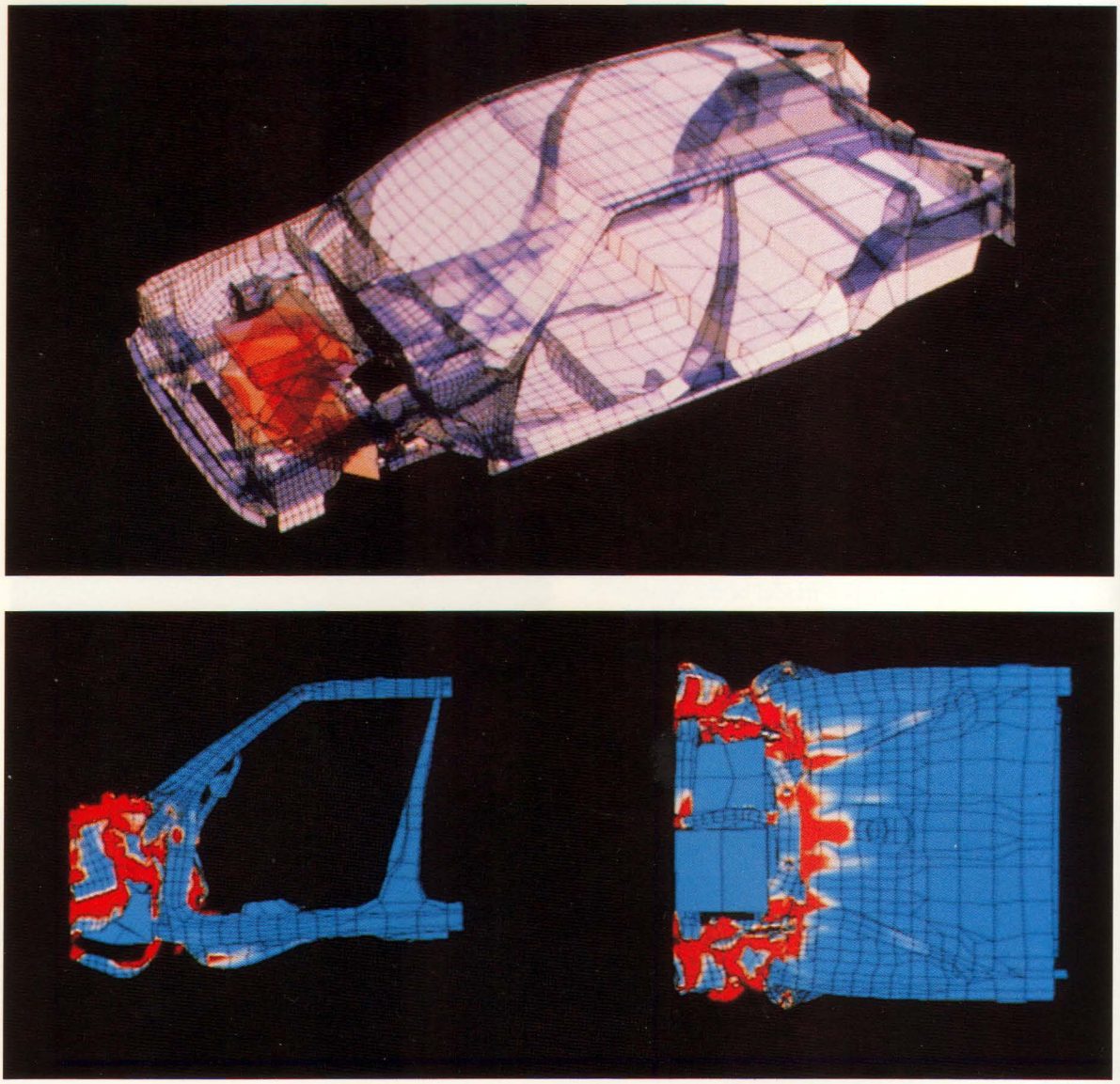


Figure 1. Increasingly complex structural models of the Space Shuttle solid rocket motor (SRM) joint. (Top) SRM model contains 5236 nodes and represents a one-degree segment of the cylindrical joint. (Middle) A six-degree segment of a cylindrical joint containing 10,472 nodes. (Bottom) This more recent model represents a 12-degree segment of the joint.

Figure 2. Peugeot BX car crash simulation. Top image shows the undeformed geometry, and the bottom image shows the deformed shape at 80 milliseconds.



Computer simulation provides information that cannot be obtained from crash tests.

and an explicit finite element program, PAM-CRASH.[1] The results compared extremely well with the experimental data (Figures 2 and 3), demonstrating the potential for full-body crash simulation. The use of an explicit finite element program, as opposed to the more traditional implicit methods, was a key component in the success of the analysis. The explicit method has proven its effectiveness in short-duration, high-energy type analyses such as crash simulation.

Once the automotive industry recognized the potential for computer crash simulation using explicit programs, new crash analysis applications rapidly spread throughout the industry. Since the Peugeot analysis in 1986, all automotive manufacturers have begun using explicit finite element programs to perform crash simulations, coinciding with a sudden increase in the number of Cray systems used in the automotive industry. Since early 1987, BMW, Daimler Benz, Honda, Peugeot, and Volkswagen have added Cray computers to their CAE environments. As of May 1, 1988, Cray Research has received new system orders from Toyota and Nissan.

Currently, finite element models for front-end crash simulations contain approximately 10,000 elements (primarily four-noded shells), including a very detailed model of the main structural components of the vehicle front. During the crash, the engine hits the wall and is pushed into the lower dashboard. The engine's interaction with the rest of the structure is a critical component of the crash simulation and requires the use of contact surface algorithms in the analysis programs. To perform an analysis of this magnitude on one processor of a CRAY X-MP system requires roughly 10 CPU hours. The explicit programs are well-suited for vectorization and run at over 70 MFLOPS on a CRAY X-MP system.

New crash scenarios

Numerous successful crash simulations have been performed on Cray computers.[2,3,4] However, the automotive industry is still in the early stages of incorporating computer simulation into the design process for crashworthiness. The initial analyses simulated the response of existing automobiles, and were intended to verify the computer methods against physical test data. Analyses are growing in sophistication as engineers gain more experience and confidence

in crash simulation. The full potential of crash simulation has yet to be realized.

Up to this point, crash simulations had represented only front-end impact against a rigid wall. Several other crash scenarios are worth consideration. The National Highway Traffic and Safety Administration (NHTSA) estimates side-impact collisions account for close to one-third of all highway crashes and serious injuries. The NHTSA has proposed stricter side-impact standards, which they estimate would save 1200 lives per year in the United States. Research into side-impact computer simulation already has begun in Europe, where government regulations have set new standards for side-impact safety. Also, the same analysis techniques apply to automotive rear-impact collisions, aircraft hard-landing simulations, and many other dynamic events of short duration.

The goal of crash simulation is to provide engineers with tools to investigate new design concepts without having to crash-test vehicles. Although computer simulation is not intended to replace physical testing, computer simulation reduces the number of tests needed and shortens the design cycle, resulting in tremendous cost savings. Additionally, computer simulation provides information that cannot be obtained from crash tests. For example, a computer simulation can model the interactions of the engine and lower dashboard, as well as the energy absorbed by the side rails. Computer simulation provides the flexibility to experiment cost effectively with new ideas, which ultimately should lead to the design of safer vehicles. As crashworthiness simulation continues to evolve, the number of elements in structural models is expected to increase five- to tenfold. This will provide more accurate simulations of the complicated interactions in the front structure of vehicles. Although an increase in elements will lead directly to an increase in computer time, multitasking should reduce the elapsed time for the larger models. With the CRAY Y-MP/832 system's eight processors, the potential to make the large models practical already exists.

New application areas

Computer power and improved software are opening several new structural analysis areas. Design optimization is an important new analysis field that involves using computers to calculate automatically an efficient design based on given constraints. For example, analysis might seek to minimize the weight of a component without exceeding an allowable stress or deformation. Although commercial optimization software has been available for several years, it only now is gaining acceptance in production environments.

The computer-aided engineering environment is changing constantly. As UNIX-based workstations become standard in the analysis environment, networking with Cray systems running under the UNICOS operating system allows flexibility and ease of use. This makes Cray systems effective in a wider variety of application areas.

Computer graphics, which has always been a strength of Cray computer systems, is emerging as another application area. This application typically has not been utilized fully in structural analysis because the COS operating system has been viewed as a batch processing environment. The increasing size of today's models brings greater need for high-quality graphics. Cray computers and the UNICOS environment provide cost-effective graphics postprocessing.

The finite element method has a long, successful history as a structural analysis tool. The computational needs of engineers, combined with the versatility of Cray systems, have resulted in a dramatic increase in the number of Cray systems applied to structural analysis. Because of the availability and power of supercomputers, problems that previously were not practical are solved today during routine production schedules.

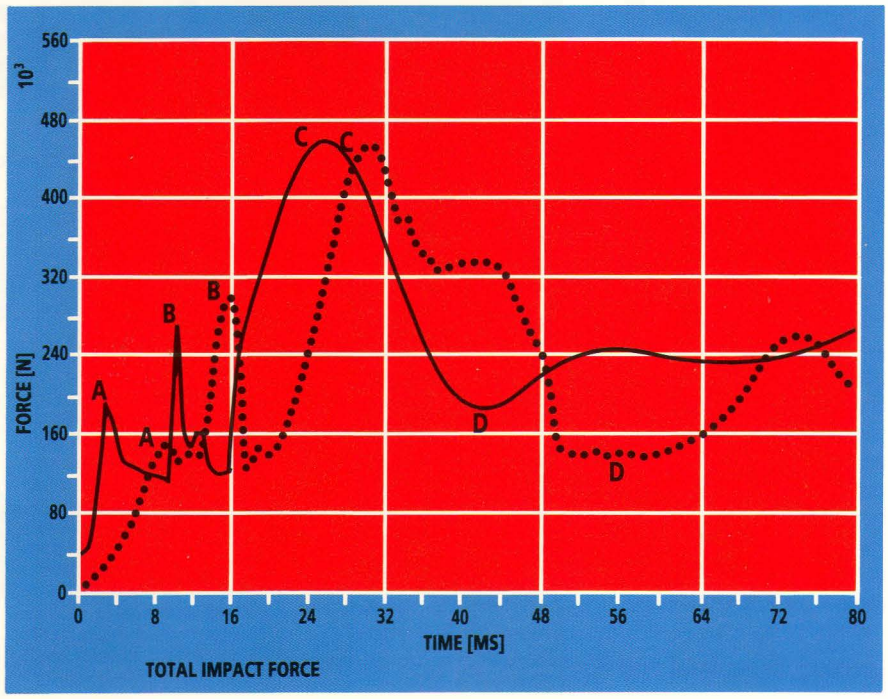


Figure 3. Curves indicate impact force of car versus time for the Peugeot BX crash simulation. Dotted line represents experimental data and solid line represents numerical data. The maximum force values attained before the engine hits the wall are represented at A and B. The force of the engine impacting the wall is shown at C. The movement of the engine/gearbox and its supporting structure is denoted at D.

About the author

Gregory Clifford is manager of the structural applications group in Cray Research's Industry, Science & Technology Department. Before he joined Cray Research in 1983, he had worked at Honeywell, Control Data Corporation, and System Development Corporation. Clifford earned B.S. and M.S. degrees in civil engineering from the University of Minnesota in 1977 and 1979, respectively.

References

1. J. F. Chedmail, et al., "Numerical Techniques, Experimental Validation and Industrial Applications of Structural Impact and Crashworthiness Analysis with Supercomputers for the Automotive Industries," *Proceedings of the International Conference on Supercomputer Applications in the Automotive Industry*, October 1986.
2. Scharnhorst, T., "Simulation of Frontal Car Crash on the CRAY X-MP," *Science and Engineering on Cray Supercomputers, Proceedings of the Third International Symposium*, September 1987.
3. Benson, D. J. and J. O. Hallquist, "DYNA3D, A Computer Code for Crashworthiness Engineering," XV International Finite Element Congress, 1986.
4. Le-The, H., et al., "Crash Simulation of Vehicles Using the Explicit Finite Element Method," *Finite Element News*, April 1988.

Chemistry by simulation

Erich Wimmer and John Mertz, Cray Research, Inc.

Most materials and products common today are in some way derived from a knowledge of chemistry. Consider the fibers and colors in clothing, the building materials in houses, the fertilizers and pest controls in agriculture, as well as pharmaceuticals, health care products, and fuels. Even the unique materials that characterize the information age come from chemical processes. Pure silicon and gallium arsenide crystals are doped carefully to achieve unique semiconducting properties; thin glass fibers are fabricated for high-bandwidth telecommunication and computer networks; compact disks, magnetic disks, and tapes store music and information; and fluorocarbon fluids cool today's supercomputers.

Looking back: trial and error

All these products of applied chemistry were developed from a desire to formulate new molecules, compounds, and materials, to manipulate and engineer their properties, and to understand their interactions with the environment. Until recently, scientists followed an ancient procedure of trial and error for developing new materials. Chemists would synthesize compounds, test them for the desired properties, use successful compounds and discard unsuccessful ones, then test more compounds.

Scientists today can approach problems differently. They know in principle that the answers exist in the laws of quantum, statistical, and classical mechanics. These theories essentially were complete in the mid-1920s; however, scientists also recognized immediately that the equations were much too complex to be solved exactly. Nevertheless, this theoretical achievement constituted a major step forward, because less than 100 years earlier people believed that the attempt to use mathematical methods for chemical problems was an "aberration" and "contrary to the spirit of chemistry," as philosopher August Comte stated in 1830.

During the past 50 years, many sophisticated approaches have been developed to solve the equations of quantum mechanics and statistical mechanics in an approximate, but sufficiently accurate, way to produce meaningful answers. Earlier, these techniques could be applied only to very small, idealized systems of mostly academic interest. The computational effort was too monumental to apply these approaches to complex real-world problems. Today, the power and speed of supercomputers enable chemists to employ these approaches to increasingly large systems of technological importance.

A technological revolution

In 1985, the world's largest chemical company, E.I. DuPont de Nemours and Company, acquired the first supercomputer in the chemical industry, a CRAY-1 system. In early 1986, the first supercomputer exclusively dedicated to research in the health sciences, a CRAY X-MP system, was installed at the National Cancer Institute near Washington, D.C. In 1987, the Research Institute of Scripps Clinic acquired a CRAY X-MP/14se supercomputer for its research efforts in molecular biology. In that same year DuPont upgraded its supercomputer capacities substantially by moving from a CRAY-1 to a CRAY X-MP/24 system with an SSD solid-state storage device. Also in 1987, Kodak joined the supercomputer consortium at the University of Illinois, followed by Eli Lilly, an Indianapolis-based pharmaceutical company, and the Amoco Research Center in Naperville, near Chicago. In early 1988, Exxon Research and Engineering acquired a CRAY X-MP supercomputer for research efforts including computational chemistry. Increasingly, existing Cray customers such as petroleum companies and electronics companies like NTT and Toshiba are developing computational chemistry applications. Major pharmaceutical companies are evaluating seriously the use of Cray supercomputers for their research and development. Supercomputer simulations are becoming an integral part of industrial chemical research, signaling the beginning of a technological revolution.

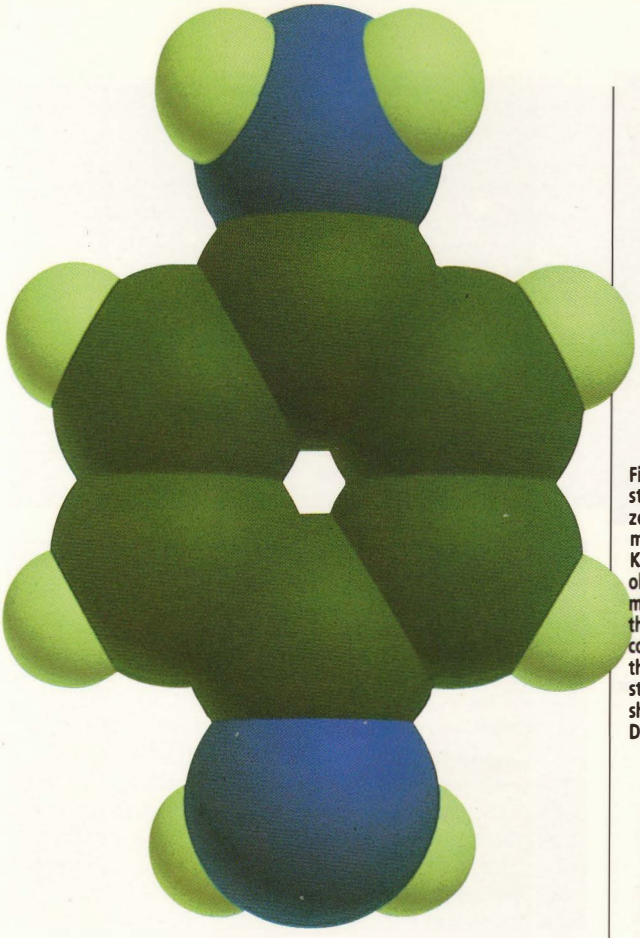
The supercomputer simulation led quickly to the decision not to pursue the costly and time-consuming experimental route to develop an isomerization catalyst.

Predicting the performance of catalysts

Quantum mechanical calculations and molecular dynamics simulations are two major application areas within the broad range of computational chemistry applications. Kevlar® originates from DuPont's research efforts in high-performance polymer fibers. Kevlar is an aramid fiber of great commercial demand, and stands out because of its unique mechanical properties. The important monomer required for the production of Kevlar is para-diaminobenzene, which is synthesized in an expensive process that involves toxic intermediates. An alternative method of synthesis would be highly desirable.

Theoretically, one could start with the meta-isomer, which is easier to synthesize, and transform it into the para-isomer through an isomerization process, which requires a catalyst. Before embarking on the development of such a catalyst, a process that can cost millions of dollars, it would be highly desirable to anticipate the catalyst performance under optimal conditions. To predict performance, one needs to know the thermodynamic stability of each isomer of diaminobenzene — the energy required to transform one isomer into another. Usually, these "heats of formation" are found in tables or determined experimentally. However, in the case of diaminobenzenes, the experiments cannot be performed easily within required accuracy. In fact, to obtain the necessary thermodynamic data, a catalyst first must be developed.

To overcome this lack of information, David A. Dixon from DuPont's Experimental Station carried out precise quantum mechanical calculations.



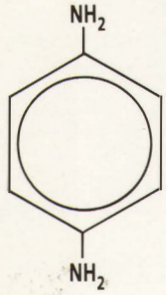


Figure 1. Three-dimensional structure of para-diaminobenzene. This isomer is the critical monomer in the synthesis of Kevlar. The structure was obtained by fully quantum mechanical calculations using the program GRADSCF on a Cray computer system. For reference, the two-dimensional chemical structure of the para-isomer is shown above. Image courtesy of David A. Dixon, DuPont.

Examining the chemical structure of these compounds, the quantum mechanical complexity is not immediately apparent. However, a crucial factor in predicting the heat of formation in molecules is the distribution of electrons. In this case, each molecule has 58 electrons, each interacting with all the other electrons and with all of the positively charged nuclei. The motion of the electrons is governed by the quantum mechanical Schrodinger equation. Its approximative solution leads to a set of complex integro-differential equations known as Hartree-Fock equations. On the order of 50 million six-dimensional integrals need to be calculated before the equations can be solved iteratively.

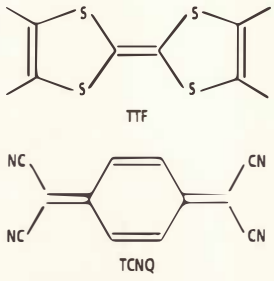
To proceed, a starting distribution for the electrons is estimated. The first set of equations can be set up and solved using this distribution. The resulting solutions, corresponding to a rearrangement of the electrons, define a more accurate set of equations. The process is then iterated until the electronic distribution does not change from one step to the next. At this point, "self-consistency" is reached and the total energy, which determines the heat of formation, can be calculated. However, an additional complication occurs: the exact position of the atomic nuclei also must be determined. To make this determination, one needs to assume a reasonable starting geometry, calculate all necessary integrals, and solve the self-consistency problem. Then, one must not only evaluate the total energy, but also all forces on all atoms (that is, the gradients of the total energy, which requires the solution of 500 million integrals) and move the atoms to reduce the total energy. Finally, at equilibrium, the forces on all atoms vanish and the ground-state structure of the molecule is found. In the present case, the program GRADSCF, designed and written by Andrew Komornicki of Polyatomics Research Institute, has been employed. To obtain accurate relative energies, the molecular frequencies must be calculated, which requires evaluation of the second derivative of the energy. This process is extremely demanding, with over 2.2 billion integrals to compute. From the complexity of the procedure, it is apparent that such a calculation requires the power and speed of a Cray supercomputer to obtain the necessary accuracy within reasonable time.

The optimized geometry for the para-isomer is shown in Figure 1. The calculated relative heat of formation leads to an interesting and somewhat unexpected result: the para-isomer is the least stable (3.0 kcal/mole less stable then the meta-isomer), while the ortho-isomer lies energetically in between the para- and meta-isomers. This result implies that the planned isomerization catalyst would yield less than 1 percent of the desired para-diaminobenzene at room temperature, and only 8 percent if operated at fairly elevated temperatures. This concentration is too low to be commercially viable. The supercomputer simulation led quickly to the decision not to pursue the costly and time-consuming experimental route to develop an isomerization catalyst.

Understanding metallic organic polymers

The notion that organic polymers could have metallic conductivity seems rather surprising. Following DuPont's pioneering work in the 1960s, it was not until 1972 that a crystal was discovered that exhibited relatively good metallic conductivity. D. O. Cowan and J. P. Ferraris at Johns Hopkins University and A. J. Heeger and A. F. Garito at the University of Pennsylvania discovered the organic, conductive crystal, which was made from two types of organic molecules (Figure 2), tetrathiafulvalene (TTF) and tetracyanoquinodimethane (TCNQ).

Not only are the conducting properties of this compound novel and exciting, but also its crystal structure is quite unusual. In the crystalline TTF-TCNQ solid, an electronic charge transfers from the TTF molecule to the TCNQ molecules, resulting in TTF^+ and $TCNQ^-$ ions. Normally, one would expect that in an ionic crystal the positively and negatively charged ions would be arranged such that negative ions surround the positive ions, and vice versa, as is the case in rock salt. Surprisingly, despite the Coulombic repulsion between ions of the same charge, TTF molecules are stacked on top of each other, as are the TCNQ molecules, forming the crystal structure shown in Figure 2.



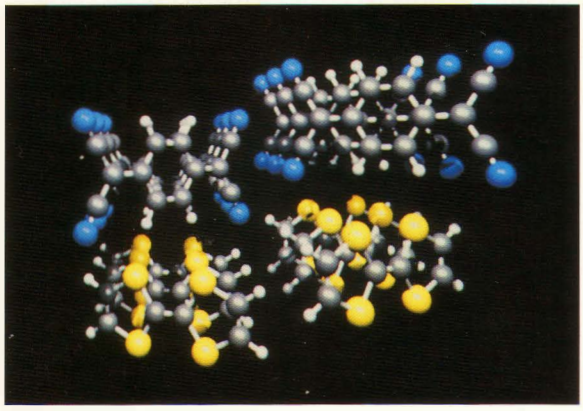


Figure 2. (Above) Molecular structures of tetrathiafulvalene (TTF) and tetracyanoquinodimethane (TCNQ); (right) stacking of molecules as observed in the TTF-TCNQ crystal. Despite their electrostatic repulsion, positively charged TTF molecules are stacked on top of each other, as are the negatively charged TCNQ molecules.

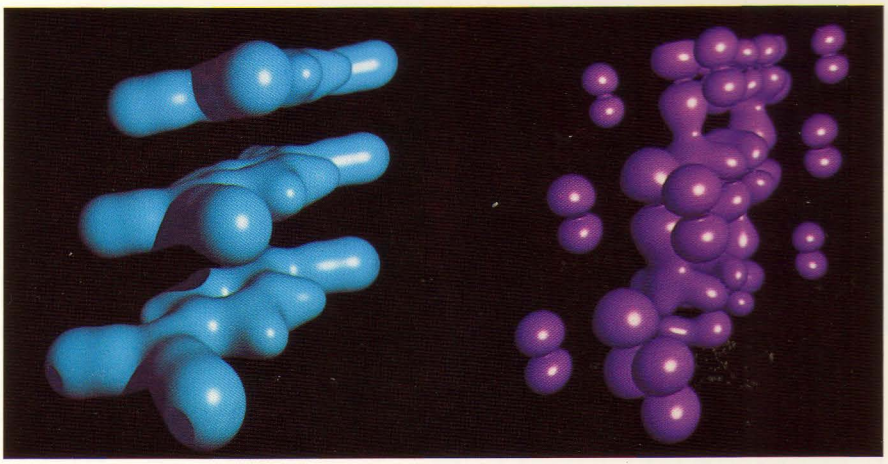


Figure 3. (Left) Iso-electron density contours in a TCNQ trimer; (right) electron density in the frontier orbital responsible for the intermolecular bonding. Note the connections between the different molecules. Graphics by Gray Lorig, Cray Research, using the OASIS ray-tracing program

Until today, computational approaches had failed to provide a quantitative picture of this intriguing intermolecular interaction. Progress in theoretical/computational approaches and the performance of today's supercomputers have changed this situation. In a collaborative effort between Shih-Hung Chou and A. J. Freeman (Northwestern University), J. Stevens and P. C. W. Leung (3M Science Research Laboratory), B. Delley (Paul Scherrer Institute, Switzerland), and E. Wimmer (Cray Research, Inc.), the puzzle of the crystal structure was solved by carrying out first-principles quantum mechanical calculations using a CRAY X-MP computer system and the DMOL program.

The calculations reveal that the crystal structure mainly is determined by the bonding interactions between TCNQ molecules, and to a lesser extent by the bonding between adjacent TTF molecules. The calculations simulated the first steps of the crystal formation by bringing together two neutral TCNQ molecules. Surprisingly, this neutral TCNQ dimer assumes an intermolecular distance and a parallel displacement very close to that found in the extended crystal. If one then adds two electrons to the TCNQ dimer, it does not fall apart from the Coulombic repulsion, as one might expect. On the contrary, its binding becomes even tighter. A similar behavior is found in the TCNQ trimer. Figure 3 shows the calculated electronic charge in a TCNQ dimer, as well as the "frontier orbital," which is responsible for the TCNQ-TCNQ binding. Similarly, the neutral TTF dimer is predicted to be stable, yet removal of electrons destroys this stability. Hence, the TCNQ-TCNQ interaction holds the crystal together, whereas the positively charged TTF stacks can exist only in a crystalline environment. In addition, the calculations predict a transfer of 0.65 electrons per TTF-TCNQ pair, which agrees well with experimental data. This achievement demonstrates that novel theoretical/computational techniques, combined with the power of Cray supercomputers, enable scientists to simulate technologically important conducting polymers to understand the intriguing intermolecular interactions and their consequences for electronic properties and macroscopic properties such as crystal structure.

Computational molecular biology

The triumphs of molecular biology have contributed immensely to the quality of modern life. Certainly, the development of critical drugs and vaccines has achieved the greatest impact, but molecular biology also has given scientists respect for the incredible complexity of biomolecular systems. Buried within this complexity are insights that can aid in the design and discovery of new drugs. Cray supercomputers now have become important tools to scientists as they attempt to unveil some of this complexity.

One method of probing biomolecules is called molecular dynamics. In this approach, which is aimed at large biomacromolecules, the interactions between atoms are described through parameterized force fields rather than full quantum-mechanical interactions. This method allows the investigation of a large number of atoms because it reduces computational demands. One difficulty, however, is determining the necessary parameters. In the past, this was done mostly in an intuitive and empirical way, leading to inconsistent quality.

Today, scientists commonly carry out rigorous quantum-mechanical calculations on small subsystems, determine the forces and force constants from these first-principles results, and then build a consistent force field in a systematic way. Once the force field parameters are determined, the problem of molecular vibrations and motion is reduced to classical mechanics and the equations of motions can be integrated numerically. During the integration procedure, data are collected and the properties of interest are calculated. One such property that recently has received much attention is the relative free-energy difference between two molecular systems, for example between an enzyme complexed with drug A, and the same enzyme complexed with drug B. The free energy is important because it is the thermodynamic measure of stability. As free energy lowers, the stability of the chemical system increases. Therefore, if the enzyme-drug A complex is lower in free energy than the enzyme-drug B complex, the former will be more stable. Knowledge of relative stabilities allows predictions to be made about the effectiveness of one drug over another.

In a recent investigation, scientists from the University of California, San Francisco, reported their methods to calculate free energy.[1] They demonstrated the utility of their method by investigating the free energies of solvation in amino acid side chains, nucleic acid bases, and a few other organic molecules. Their results were consistently within 10 percent of experimental values. In addition, they cited one case in which traditional techniques incorrectly predicted one system to be more stable than another, but when they used the free energy of solvation as a measure of stability, they reproduced the experimental observations.

U. C. Singh of the Research Institute of Scripps Clinic is using free-energy methods to investigate and verify experimental work performed on the anticancer drug methotrexate (MTX) (Figure 4). This drug helps to slow the growth of cancerous cells by inhibiting a biochemical cycle important to the production of DNA. The drug breaks the chain in this cycle by binding to the enzyme dihydrofolate reductase (DHFR) (Figure 5), thereby inhibiting it from performing normal functions. Normally, DHFR is used in the reduction of the vitamin folic acid, but folic

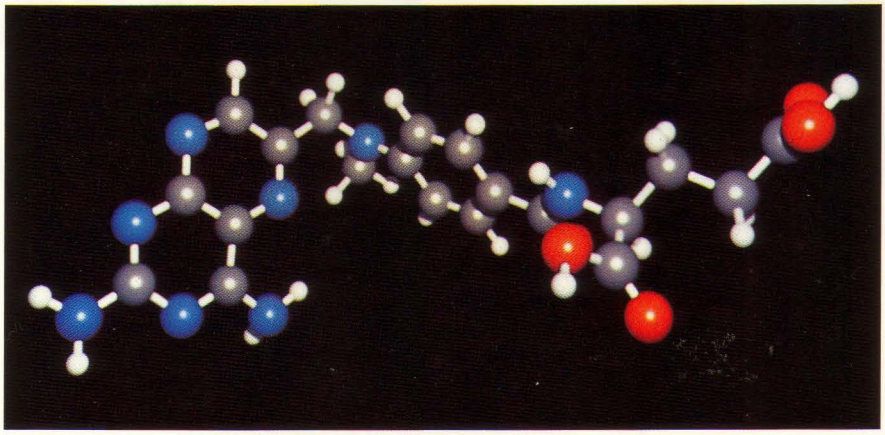


Figure 4. Three-dimensional structure of the anticancer drug methotrexate.

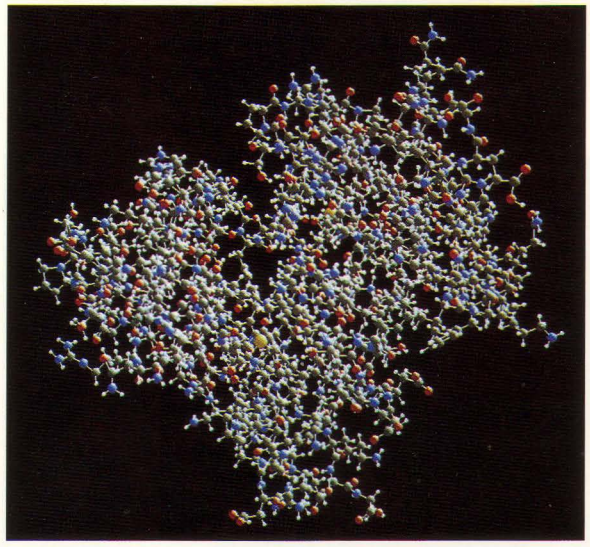


Figure 5. Three-dimensional structure of dihydrofolate reductase (DHFR) as used in the molecular dynamics study. This enzyme, needed in the production of DNA, is blocked by the anticancer drug methotrexate, thus inhibiting the growth of cancer cells.

acid analogues, like MTX, are not reduced by DHFR; they bind to it. Singh is attempting to find out why MTX binds to DHFR, whereas folic acid, which is chemically very similar, does not.

Some have proposed that a proton exchange exists between MTX and DHFR, creating an ion-pair bridge that causes the drug to bind to the enzyme. Detailed experimentation has shown that the free energy due to the ion-pair is not enough to account for the binding behavior of MTX. The technique used in this experiment is called site-directed mutagenesis. One amino acid in DHFR was replaced by a mutant amino acid. Singh simulated this experiment on a Cray supercomputer and obtained the same results. In addition, computational techniques allowed him to explore the origin of the results and predict the catalytic mechanisms of DHFR. Given the recent successes and the potential of free-energy methods, scientists are excited about performing simulations with supercomputers. As simulation techniques become more standardized and supercomputers improve in performance, molecular dynamics simulations of biomolecular systems will become an integral part of any in-depth research efforts, academic as well as industrial.

Grand challenges

Today, we barely are scratching the surface of all the potential benefits of large-scale simulations, particularly in chemistry and materials science. For example, DuPont is saving millions of dollars with just a few well-formulated and carefully executed quantum mechanical calculations on relatively simple molecules. In materials research, the current state of the art in electronic structure calculations enables researchers to study systems such as organic conductors, yet many questions remain open. Today's molecular dynamics methods and the free-energy perturbation technique are capable of correctly reproducing experimental data. Thus, they are proving to be viable theoretical/computational concepts, but the majority of biomacromolecular systems still await computational simulation, and grand challenges such as the protein-folding problem still are unsolved.

Step by step, scientists will come closer to the goal of simulating the incredibly rich hierarchy of chemical systems, from simple diatomic molecules of oxygen and nitrogen to the delicate balance of chemical structures and reactions in living cells. These systems involve complex ensembles of macromolecules, each comprising hundreds of thousands of atoms that obey the same laws of quantum and statistical mechanics as the diatomic molecules. Today, the first rewards are becoming apparent, scientifically as well as commercially. These successes are producing confidence to further refine and advance computational methods with expectations that the most exciting breakthroughs are yet to come. ▬

Acknowledgments

The authors would like to thank for many fruitful discussions David A. Dixon, DuPont Experimental Station; Peter Leung and John Stevens, 3M Science Research Laboratory; Shih-Hung Chou and Arthur J. Freeman, Northwestern University; and U. Chandra Singh, Research Institute of Scripps Clinic. Figures 2 through 4 were created with the OASIS ray-tracing package, written by Gray Lorig of Cray Research. Figure 5 was created with a graphics package written by Georg Temnitschka of Cray Research.

About the authors

Erich Wimmer is manager of the chemistry group in Cray Research's Industry, Science & Technology Department. He received a B.S. degree in chemical engineering from the Institute of Technology in Vienna, Austria, in 1974, and received a Ph.D. degree in chemistry from the same institution in 1977. Before joining Cray Research in 1985, he spent three years as a research associate at Northwestern University in A.J. Freeman's group. He also was an associate professor of theoretical chemistry at the University of Vienna.

John Mertz is a senior scientist in the chemistry group in Cray Research's Industry, Science & Technology Department. He received a B.S. degree in chemistry from the University of Minnesota in 1983, and received a Ph.D. degree in chemical physics from the University of Illinois in 1987. His activities and research interests include supercomputer applications in biology, polymer science, and condensed matter physics.

Reference

1. Bash, P. A., U. C. Singh, R. Landridge, and P. A. Kollman, "Free Energy Calculations by Computer Simulations," *Science*, Vol. 236, 1987, pp 546-568.

New applications for CFD

Kent Misegades, Cray Research, Inc.

Computational fluid dynamics (CFD) deals with the simulation of fluid flows through the computational solution of the equations of fluid motion. CFD uses essentially the same set of equations to describe the flow of many different fluids. Researchers use CFD to model the flow of air, water, blood, burning gases, reacting liquids, steam in powerplants, planetary atmospheres, slurries, ocean currents, liquified foods, and molten metal, plastic, and glass. Although CFD is so complex that most problems of interest defy exact solution, Cray systems are powerful enough to make CFD a practical tool for many industrial research and production applications.

Aerospace industry applications

Research in the aerospace industry represents the most developed industrial use of CFD. Industrial aerospace engineers use Cray systems to solve problems such as

- ☐ Wing aerodynamics
- ☐ Wing/fuselage airflow interaction
- ☐ Engine nacelle airflows
- ☐ Turbomachinery (internal) airflows
- ☐ Combustion chamber airflows
- ☐ Liquid and solid rocket fuel flows
- ☐ Re-entry vehicle gas dynamics
- ☐ Blast wave/vehicle interaction
- ☐ Turbulence research

The work of professors Antony Jameson and Tim Baker at Princeton University is among the more significant aerospace research supported by Cray Research. Since 1985 Jameson and Baker have been developing an Euler method for solving the airflow about arbitrary geometries using unstructured, tetrahedronal meshes generated with the Delauney triangulation method. The first results of this effort were published in 1986 and described the airflow about a Boeing 747.[1] Although the memory requirements for this code are higher than those of conventional codes, the run times are comparable: 1 hour on a CRAY X-MP system and 14 hours on a Convex C-1 system. The main advantage of this method is its enhanced ability to handle complex geometries.

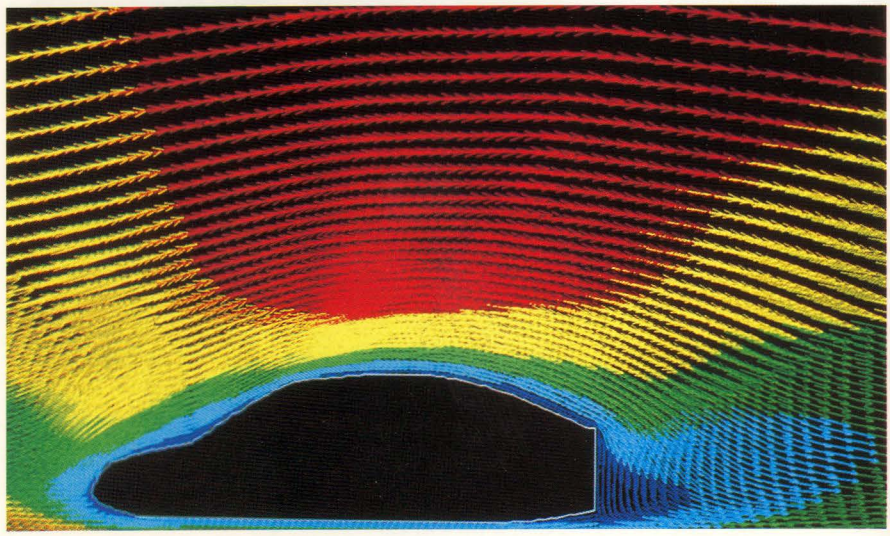


Figure 1. FIDAP results for external automobile aerodynamics.

More recent results using the same program show the airflow about a McDonnell-Douglas F-15 fighter plane, including flow through the intake.[2] These computations were performed on Cray systems in Mendota Heights, Minnesota.

Although the aerospace industry is the most experienced user of CFD technology, many other industries also use the technology. In some of these cases, the computer requirements exceed those of current methods in the aerospace industry.

Automotive industry applications

Although the use of CFD in the automotive industry is limited, it will significantly affect the industry in the near future as commercial software developed for non-aerospace users is introduced into automotive laboratories. Existing and potential CFD applications in the automotive industry include

- ☐ Brake cooling
- ☐ External aerodynamics

- ☐ Fuel injection spray flows
- ☐ Internal combustion
- ☐ Manifold airflows
- ☐ Metal casting
- ☐ Passenger compartment heating, ventilation, and cooling
- ☐ Water jacket flows
- ☐ Plastic injection molding
- ☐ Radiator/engine compartment airflows
- ☐ Tire molding

The program FIDAP was used recently to solve the three-dimensional, steady, laminar airflow about a passenger car, including the effects of a fixed ground plane beneath (Figure 1). This case required approximately 120 million words of memory and approximately one hour of computing time on a CRAY X-MP system. A vehicle's external aerodynamics usually is difficult to simulate. Two main factors account for the difficulty: the geometry is complex, which makes mesh generation difficult; and some of the physical phenomena, such as turbulence and large regions of flow separation, are difficult to describe precisely.

The field of combustion simulation makes aerodynamics seem simple. A recent computation of the "cold," or noncombusting, intake and compression strokes of a conventional internal combustion engine required only four million words of memory but approximately 20 hours of computing time on a CRAY X-MP system (Figure 2). The FMCS code from AVL List GmbH in Graz, Austria, was used. Although a better understanding of the internal flow pattern just before ignition is important, the ultimate goal is to compute the actual combustion process. Such a capability will be useful for analyzing engine performance and predicting the type and amount of pollution produced.

Plastics industry applications

Most of the publicity given to supercomputer use in chemical research focuses on molecular and atomic modeling. However, industrial researchers also are interested in an equally computationally intensive application: modeling the fluid flows that characterize much of chemical engineering. Computational modeling can be applied to many types of chemical engineering problems, including

- ☐ Atmospheric, water, and soil pollution
- ☐ Chemically reacting fluid flows
- ☐ Composite fiber orientation
- ☐ Extrusion
- ☐ Fluid mixing, drying, and separation
- ☐ Thermoplastic and reaction injection molding
- ☐ Multiphase flows
- ☐ Thin-film coating

Chemical engineers using computational methods would like to model chemically reacting fluid flows, but capable software is not yet available. This situation likely will change in the next few years, based on known activities of software developers. However, applications in polymer processing already are nearing maturity. Polymer processing applications include the simulation of extrusion, injection molding, and thin-film coating.

Figure 2. Intake flow in an internal combustion engine modeled with the FMCS software package.

Extrusion modeling

Extrusion is a process in which a polymer melt stream is shaped by being forced through a die. The two primary concerns of engineers working in this area are die swell, which is the expansion of the elastic material after it leaves the die to take on its final shape, and the detailed fluid flow patterns inside the die.

Dennis J. Coyle of General Electric Corporate Research and Development has conducted extrusion research using the program FIDAP on a CRAY-2S system. Figure 3 shows the fluid flow within a spiral mandrel die consisting of a rectangular channel that is wrapped two revolutions around a cylindrical slit.[3] The primary objective in the design of such extrusion dies is to find the proper geometry to produce a uniform melt profile. The figure shows that an essentially linear pressure drop exists from the top to the bottom of the die and that the circumferential pressure distribution at the die's orifice is constant. This situation ensures a constant mass flow rate and thickness of material out of the die.

Injection molding

In contrast to the continuous process of extrusion, injection molding involves the intermittent process of plastic injection, packing, cooling, and part ejection with subsequent shrinkage and warpage. Three major software packages available on Cray systems are being used to simulate the injection molding of thermoplastic polymers: MOLDFLOW, POLYFLOW, and C-FLOW. Two types of organizations use these codes. Polymer manufacturers use them to demonstrate the superior properties of their materials and to assist customers in producing finished parts. Mold designers and manufacturers use them to design effective molds (with associated gates, runners, and cooling channels) and to determine effective process setup parameters, including injection pressure, temperature, and feed rates.

MOLDFLOW, which simulates the mold-filling process, and MOLDTEMP, which handles the

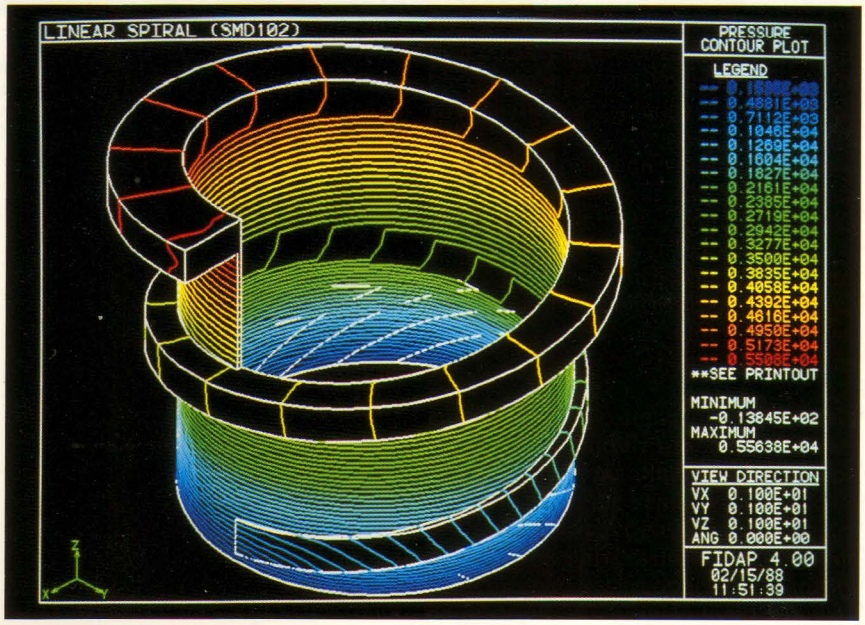
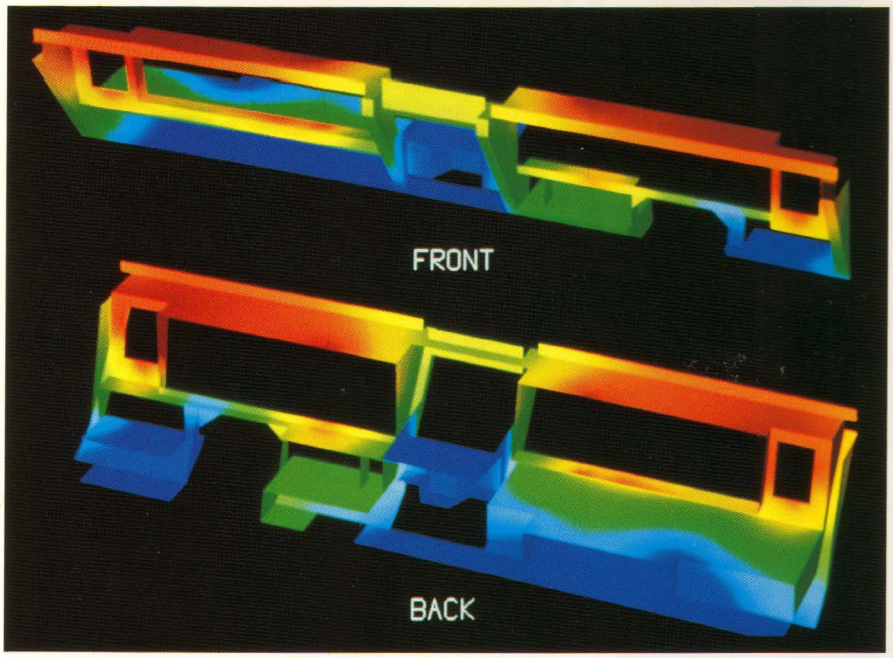


Figure 3. Spiral mandrel die flow modeled with the FIDAP package.



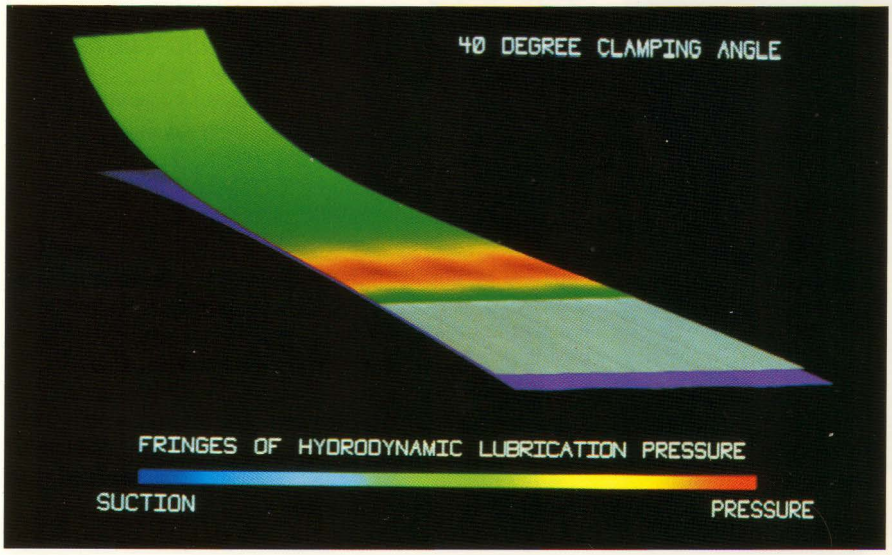


Figure 4 (Top). Automobile dashboard injection molding modeled with the C-FLOW package. High pressure areas are shown in red, low pressure areas in cyan.

Figure 5 (Bottom). Simulation of liquid film deposition on gravure cylinder.

cooling channel of molds, are products of Moldflow Pty., Ltd., of Melbourne, Australia. Moldflow pioneered the use of computer simulation for injection molding, and now has made the first supercomputer installations of its software on Cray systems. The finite element program POLYFLOW is a product of the Belgian firm POLYFLOW, S.A. It is designed primarily for simulating viscous and viscoelastic flow problems including polymer and rubber processing, food rheology, drilling, and glass furnaces. It was developed particularly for non-Newtonian fluid flow problems typical of polymer processing, and is now being used on Cray systems in the petrochemical industry.

The C-FLOW package is the product of Advanced CAE Technology of Ithaca, New York. This company, associated with the Cornell Injection Molding Project (CIMP) Consortium, has developed a suite of programs for mold filling (C-FLOW), mold packing (C-PACK), and mold cooling (C-COOL). Figure 4 shows an automobile dashboard injection process modeled with C-FLOW. A CRAY X-MP system computed the simulation of the 5-second process in 175 seconds, compared to 17,500 seconds on a DEC MicroVAX II. Graphical output was generated using the MOVIE.BYU graphics package, running on the same CRAY X-MP system under the UNICOS operating system.

Film coating

The application of a film of one material to a substrate of a different material is a problem that occurs in many industries. Examples include the manufacture of paper and magnetic storage media. Unique to modeling this process computationally is the need to model the free surface, which is the boundary of the fluid flow not defined by a solid wall. The free surface typically is controlled by surface tension and normally is not known beforehand. Its final shape often is computed iteratively, which adds an additional nonlinearity to the already nonlinear equations of fluid motion. Few commercially available programs can handle free-surface flows, the most notable exceptions being FIDAP and FLOW-3D.

Cray Research's Industry, Science & Technology Department recently assisted Stephan F. Kistler, of the 3M Memory Technologies Group Laboratory in St. Paul, Minnesota, in modeling the application of a liquid film to a rotating gravure cylinder. As shown in Figure 5, the liquid is spread across the grooved surface of the cylinder by a flexible blade, the end of which is pressed against this cylinder. The structural deformation of the blade for various clamping angles, the metered flow rate, the distribution of flow in the grooves, and the spacing between the blade and the gravure cylinder are among the predicted variables. They were computed in a coupled manner with a 3M-proprietary finite element program. The objective of Kistler's computer experiments was the optimization of blade configuration, gravure cylinder design, and operating conditions. These computations and the image in Figure 5 were executed on a CRAY X-MP system running the UNICOS operating system.

Metal industry applications

Manufacturers and users of raw materials related to metals processing have explored CFD for a variety of applications, including plasma processing for chemical vapor deposition (CVD), low pressure plasma deposition (LPPD), and vacuum arc remelting (VAR); strip casting applications such as roll casting and twin-belt casting; welding-related phenomena such as convection in three-dimensional arc and laser weld pools; and casting problems such as foundry mold filling, metal solidification in molds, panel cracking during steel ingot casting, and solidification during continuous casting.[4]

Figure 6 shows the results of a recent project to study the solidification of metals in foundry molds. Jon Dantzig of the University of Illinois conducted this research. Dantzig first modified the FIDAP program to account for latent heat release, then used it to simulate the solidification history of molten steel in a hammer mold. The computation predicted that a pool of molten steel would solidify in the mold itself, and not in the riser, above the hammer. This situation would render the final part porous

and weak. In effect, casting research via CFD enables an engineer to see through molten metal, a capability not available with experimental methods.

The computation took about two hours on a CRAY X-MP system running the UNICOS operating system. The final computations for the figures were set up and run on a CRAY X-MP system at Cray Research's computer center in Mendota Heights, Minnesota, via a modem and lap-top microcomputer (operated by Dantzig from his hotel room during a recent conference on casting simulation)!

Conclusions

Although industrial CFD applications have matured the most in the aerospace industry, much work still needs to be done before certain seemingly simple tasks, such as computing viscous compressible airflow about a wing using a Navier-Stokes method, can be performed cost-effectively and reliably.

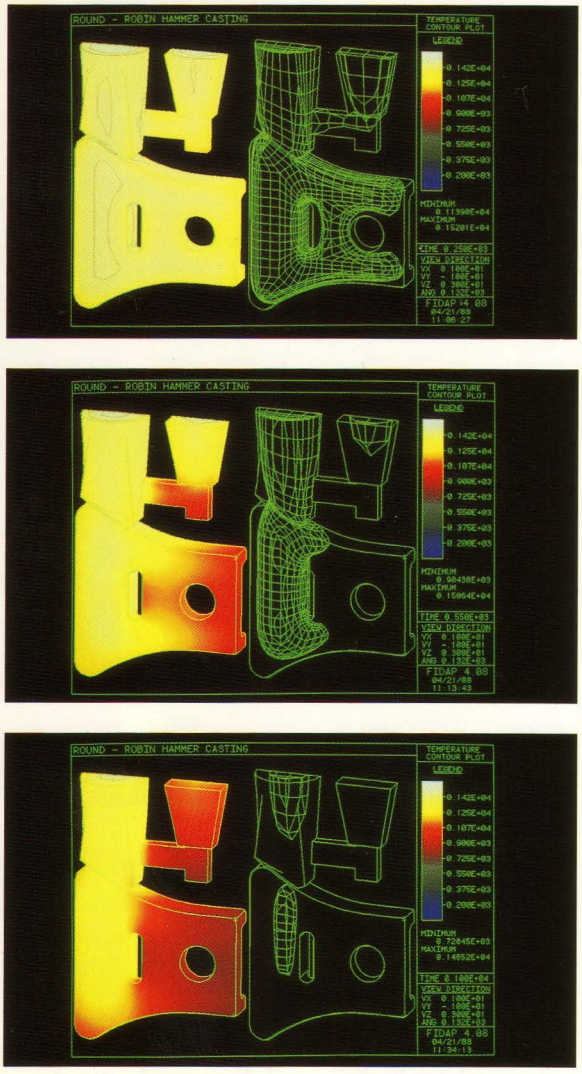
Non-aerospace industries provide numerous examples of CFD applications. Their physics and geometric complexity require supercomputing resources that equal if not exceed those in place today in the aerospace industry. Capable commercial CFD software has become widely available during the past four years, including many packages designed for non-aerospace industries. These packages typically include graphics pre- and postprocessing software, which is essential for engineers inexperienced in the use of CFD.

In the next several years more commercial software also will become available to the aerospace industry. The CFD applications group at Cray Research is working with a number of organizations developing such software, to ensure maximum performance and capability to the CFD user. Weak areas of CFD still exist, however. Along with the general problem of inadequate general turbulence models, CFD codes need to provide many other capabilities to make them more globally useful. Needed capabilities include

- ☐ Mesh generators for complex geometries
- ☐ Better interfaces to non-CFD software (CAD/CAM, structural analysis, heat transfer)
- ☐ Combustion models for subsonic and supersonic flows
- ☐ Three-dimensional free surfaces
- ☐ Multiphase fluid flows
- ☐ General reacting fluid flows
- ☐ Latent heat release for metal solidification
- ☐ General droplet spray models, droplet evaporation
- ☐ Better documentation, training, and support from software vendors

While CFD continues to mature in certain application areas, new and more sophisticated analyses will be attempted and eventually incorporated into commercial software aimed at a broad spectrum of industries. Supercomputers capable of meeting the turnaround demands of engineers are crucial to the continued proliferation of CFD applications in industrial research and production. The high performance, large memories, interactive operating systems, and unequaled price/performance of Cray systems assure that CFD will be an increasingly cost-effective tool for many industries.

Figure 6. Metal solidification simulation conducted with the FIDAP package. The grid shows the solidification front moving through the mold. Note the pool of molten steel separate from that in the riser in the bottom image. This situation, predicted by simulation, would result in an unacceptable part.

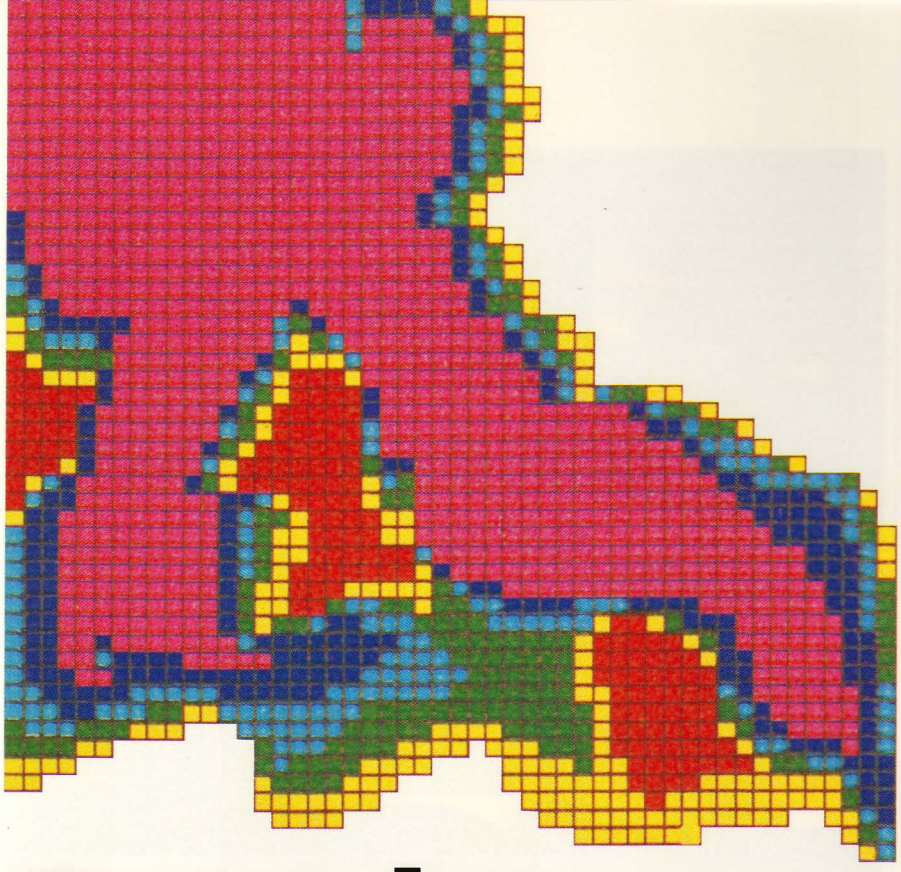


About the author

Kent Misegades is a senior applications analyst in Cray Research's Industry, Science & Technology Department. He joined the company in 1984 as senior applications analyst for computational fluid dynamics, and in 1987 became manager of the CFD applications group. Before joining Cray Research, he was an aerodynamicist in the Theoretical Aerodynamics Department at Dornier GmbH, West Germany. He is a graduate of the Diploma Course in Fluid Dynamics of the von Karman Institute for Fluid Dynamics, Brussels, Belgium. He received a B.S. degree in mechanical engineering in 1979 from Auburn University in Alabama.

References

1. Jameson, A. and T. J. Baker, "Calculation of Inviscid Transonic Flow over a Complete Aircraft," AIAA-86-0103, January 1986.
2. Baker, T. J., "Development and Trends in 3D Mesh Generation," *Proceedings of the 1988 Transonic Symposium*, NASA Langley Research Center, April 19-21, M. Salas, Ed., 1988.
3. Coyle, D. J., "Three-Dimensional Computation of Polymer Processing Flows," *Proceedings of the First FIDAP Users Conference*, 13-15 Sept., Evanston, IL, Michael Engelman, Ed., 1987.
4. Kou, S. and R. Mehrabian, Eds., Modeling and Control of Casting and Welding Processes, *Proceedings of the Third Conference on Modeling of Casting and Welding Processes*, The Metallurgical Society, Inc., 1986.



Petroleum reservoir modeling

good to the last drop

Gene Shiles, Cray Research, Inc.

Reservoir simulation is a tool used by the petroleum industry to examine management and recovery strategies. This tool relies on numerical techniques to predict the flow of fluids in a petroleum reservoir. Petroleum engineers use simulation when deciding where to drill additional wells, how to maximize the amount of oil or gas produced from a well or group of wells, and how to choose and apply enhanced recovery methods. The use of reservoir simulation helps ensure economical and efficient extraction of the world's available, and unfortunately dwindling, supply of oil and gas.

Making practical use of reservoir simulations requires fast, preferably vector, computers with large memories, such as Cray computer systems. Field-size simulations typically run in one to several hours on Cray systems, whereas similar simulations can require days or even weeks to run on mainframes and minicomputers. The petroleum industry's demand for computing power will continue to increase as the industry strives to improve the accuracy of reservoir simulation results and to simulate larger and more complex reservoirs.

Reservoir simulation

Hydrocarbon fluids (oil and gas) are found in the pore spaces of rock formations, which typically are composed of sandstone or limestone. Such formations, or reservoirs, also contain water and are usually tens of feet thick and thousands of feet wide, though they can be hundreds of feet thick and several miles wide. Reservoirs typically are found at depths of one to several miles.

The finite difference methods generally used require that the reservoir be divided into grid blocks of constant porosity and permeability. Porosity is the volume of pore space within the reservoir that is available for fluid; permeability is the interconnection of pore spaces that allows fluid to flow through the reservoir. A simulator for modeling large or heterogeneous reservoirs and two- or three-dimensional geometries can require many thousands of grid blocks to describe the reservoir accurately.

Fast computers with large memory capacity are needed to solve the matrix equations describing pressure and saturation of each type of fluid in each block. The computational requirements are complicated by the possible presence of several liquid and gaseous hydrocarbon species with differing properties, the need to simulate various enhanced recovery methods such as chemical or thermal, and the interaction of the reservoir with the wells. Initially, multiple simulations must be run to match the reservoir's field history to obtain a more accurate description of the reservoir. The improved reservoir description then can be used to predict future behavior more confidently. Very fast computers make it practical to carry out these multiple simulations and to compare alternative management strategies.

Grid-size effect

The use of small grid blocks to improve simulator resolution increases computing requirements enormously. Historically, the limitations on computing speed and memory space have restricted grid block volumes to thousands of cubic feet. This limitation on the resolution of simulators made it impossible to capture the effects of small-scale heterogeneities, which can be on the scale of centimeters. In particular, the averaging necessary to assign a single permeability value in the larger blocks can give very misleading results.

Figure 1 shows the results of a simulation performed on a computer-generated reservoir model at the Schlumberger Doll Research Center in Ridgefield, Connecticut.[1] The simulation modeled an enhanced recovery process called water flooding. During this process, water is injected at injector well(s) to increase production at producer well(s).

The injector well is at the upper left corner, and the producer well is at the lower right corner of each image. The larger grid-block simulation gives an overly optimistic estimate of oil production. The

simulation with smaller grid blocks is more accurate and shows more trapped or bypassed oil in the pore spaces (shown in the figures by regions of low water saturation). This indicates that less oil will be produced for a given production cost. Clearly, such a result justifies the use of more computing power to obtain finer resolution. However, the more accurate result will not always show a lower rate of production. Whether it shows a lower or higher rate depends on the process simulated and the reservoir model used.

Chemical flood simulator

UTCHEM is a three-dimensional compositional chemical flood simulator developed and continually improved by the Department of Petroleum Engineering at the University of Texas at Austin. This code has been highly optimized for vector operation on Cray systems. A compositional simulator is designed to handle various fluids with differing flow, thermodynamic, and, in this case, chemical properties. The inclusion of these factors represents a large increase in complexity over the simpler and more widely used black-oil simulators, which describe only three fluids: gas, liquid oil, and water.

Chemical flooding is an enhanced oil recovery (EOR) method in which a chemical detergent, or surfactant, is injected into a reservoir to help release trapped oil. A polymer also is injected to improve the flow properties of the resulting fluid mixture. The process generally includes water injection as a driving mechanism. Tracer materials, usually small amounts of radioactive substances, also are included among the injection fluids to help engineers track the fluid movements. Chemical flooding is a tertiary process undertaken when much oil remains in a reservoir's pore spaces but secondary processes, such as water injection, no longer produce oil economically.

The large number of possible fluids, the complexity of the chemical and thermodynamic processes, and the need for large numbers of grid blocks for an adequately detailed reservoir description require enormous computing power for field-size simulations. Simulation is extremely important in planning and managing tertiary recovery processes because the actual processes are very expensive to perform and will perform well only within a small range of the design parameters. The performance depends not only on the process, but also on the fluids that are in place and on the particular rock and geometric structure of the reservoir. In most cases, the design parameters of the tertiary recovery processes can be determined in advance only by a series of simulations.

Big Muddy pilot project

In 1973 Conoco, Inc., initiated a chemical flooding pilot test at the Big Muddy oil field in Wyoming to examine the effectiveness of the process. A sandstone formation with an average porosity of 19 percent and a net thickness of 65 feet was chosen for the study. The test area extended horizontally for about one acre. Previous water injection in this part of the reservoir had left residual oil occupying about 32 percent of the reservoir's total pore space. Initial work included tracer injection to determine fluid flow patterns, which was followed by surfactant and polymer injection. The total pilot project lasted until 1978, and all data produced, such as the amounts and types of materials injected, the tracer results, and the amount of oil produced, were recorded carefully.

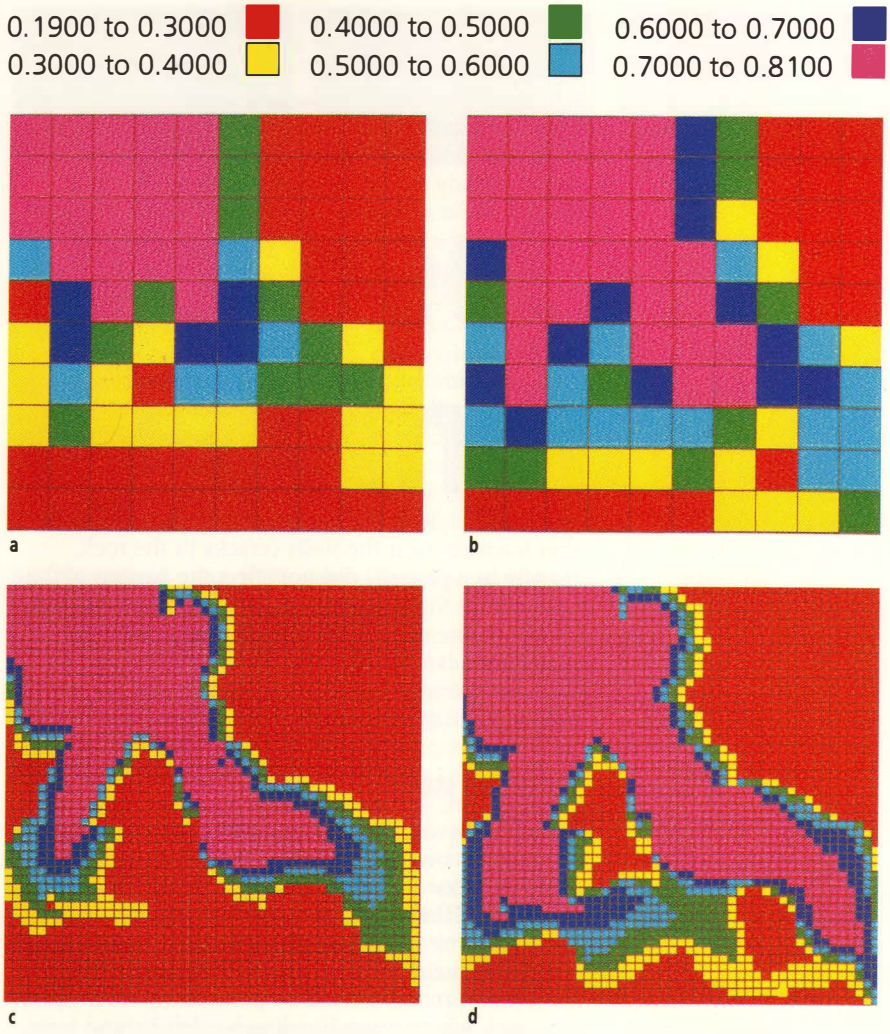


Figure 1. Simulated water saturation (a) at 1270 days and (b) at 2080 days using large grid blocks. Simulated water saturation (c) at 1270 days and (d) at 2080 days using small grid blocks.

Engineers in the Department of Petroleum Engineering at the University of Texas at Austin recently simulated this pilot test with the UTCHEM simulator, using the CRAY X-MP/24 supercomputer at the UT System's Center for High Performance Computing.[2] This series of simulations was run to learn more about chemical flooding effects on residual oil recovery efficiency, to validate and improve the UTCHEM simulator, and to predict the results one would obtain if similar processes were performed on a larger reservoir segment.

Simulator input was determined as much as possible from the actual field data and supporting physical property data. Because the reservoir description always will be incomplete due to the inaccessibility of much of the reservoir, the input was adjusted on succeeding simulations until results, such as tracers and oil produced, matched the pilot project results as closely as possible. The simulation indicated the presence of four homogeneous layers (constant permeability) in the pilot project area, rather than the three layers initially assumed. An overlying thin fourth layer with high permeability had to be included in the model to match the field results.

The computation times increased from 800 seconds for the initial three layers to 2400 seconds for the four-layer model. The difference in computation time was due to the larger number of grid blocks used and the smaller time step required for stability in the solution. These computation times indicate that the study would have been impractical on a smaller scalar computer, particularly when one considers that the code was highly optimized for vector operation. A typical minicomputer, for example, would require two or three days for a single simulation of the four-layer model, and a typical mainframe would require 15 to 20 hours.

Simulation, including tracer studies, revealed detailed information regarding chemical effects, especially on the main oil-recovery mechanism. Other results also indicated that the horizontal extent of the simulated area had to be extended slightly to include the complete flow patterns, and that fractures near the wells (cracks in the rock, usually in a pattern) did not affect the pattern of fluid flow within the reservoir. Crossflow, or flow perpendicular to the major flow direction in a reservoir, sometimes can significantly effect an EOR process. The simulation indicated that some crossflow mechanisms were more important than others in this study.

North Sea study

Researchers at the British Petroleum (BP) Research Centre at Sunbury, UK, recently simulated surfactant flooding using an in-house modified version of the UTCHEM simulator called BPOPE.[3] The study used reservoir parameters typical of the North Sea and included variations in polymer mobility control. The surfactant process should be particularly suitable for this area, where water flooding has left behind large amounts of residual light oils.

The study would have been impractical on a smaller scalar computer.

The North Sea area has unique features, such as wide well spacings, low oil viscosities, limited loading and storage facilities on platforms, and economic limits on the number of wells that can be used as injectors. Also, some of the North Sea reservoirs have been flooded with sea water, so that variations in temperature and salinity may alter the results of chemical processes. The guidelines used in other parts of the world cannot be applied without modification in this area, and the BP study is intended to provide some new guidelines for conducting simulations and for performing enhanced recovery processes.

The economic advantage in combining secondary recovery with tertiary recovery, in what may be called surfactant-aided water flooding, is a major incentive for developing a simulation capability for the North Sea area. The extent of such an advantage will have to be evaluated by simulation for individual cases.

The work discussed here was accomplished on BP's CRAY X-MP/12 system, with typical run times of one hour (BP has since upgraded to a CRAY X-MP/24 system). More detailed results are given in Reference 3.

Because of the wide well spacing and low oil viscosity, low-concentration surfactant injected over long periods was expected to yield the best results. Simulation supported this assumption when appropriate surfactant materials and procedures were used. Economic incentives exist to omit the polymer for a low-concentration surfactant process. Simulation showed that including a polymer could increase production significantly, so the decision whether or not to use a polymer will have to be based on the economics of each recovery project.

History matching of laboratory experiments gave information on phase behavior, dispersion, and gravity segregation. Additional information on the effects of physical and numerical dispersion was obtained, including grid-block size effects on simulator results (smaller grid blocks indicated more recovery). The simulations also demonstrated the possibility of designing systems less sensitive to the effects of dispersion.

Discussion

Many of the reservoir simulators used today are highly optimized for use on vector machines, and others currently are being optimized. Multiprocessing also is becoming more important. Wall-clock time can be improved with a relatively routine application of the Cray microtasking feature, although this strategy usually is not optimal with reservoir simulator codes. Achieving the speedup potential of multiprocessing will require the modularization of simulator codes, and parallelization of the algorithms. This is relatively simple on some parts of a simulator, but difficult on others, particularly the solver codes, although work is proceeding in this area. Solvers and simulators now are being written with built-in vectorization and multiprocessing. ▬

Acknowledgments

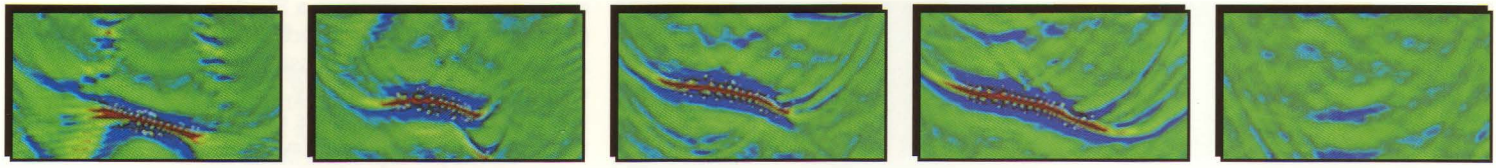
The author thanks Gary A. Pope and Kamy Sepehrnoori of the Department of Petroleum Engineering at the University of Texas, Austin, John W. Barker of the British Petroleum Company Research Centre, Sunbury, and Yogeshwar Sharma of Cray Research, Inc., Dallas, for providing and discussing the examples, and C. D. White, now with Shell Oil, for permission to include Figure 1.

About the author

Gene Shiles is a senior applications analyst in the Industry, Science & Technology Department of Cray Research, Inc. He received a B.S. degree in physics from Northern Illinois University and a Ph.D. degree in theoretical physics from the University of California at Santa Barbara. Prior to joining Cray Research in 1986, he worked for six years developing computer applications in the petroleum industry.

References

1. Sharma, Y., private communication. Related results, with emphasis on the procedures for averaging permeability values, are given in White, C. D. and R. N. Horne, SPE 16011, presented at the Ninth SPE Symposium on Reservoir Simulation, February 1987; and Lasseter, T. J., J. R. Waggoner, and L. W. Lake, *Reservoir Characterization*, L. W. Lake and H. B. Carroll Jr., Eds., Academic Press, New York, 1986, pp. 545-560.
2. Saad, N., G. A. Pope, and K. Sepehrnoori, submitted to the Society of Petroleum Engineers, November 1987.
3. Barker, J. W. and F. J. Fayers, 4th European Symposium on Enhanced Oil Recovery, Hamburg, West Germany, October 1987.



Prospecting by computer

Bill Kamp and Olivier Lhemann, Cray Research, Inc.

Two technological developments are revolutionizing seismic data processing. One is the availability of computers that can perform in minutes tasks that used to take days or even weeks to perform. Cray systems, in particular, have been instrumental in cutting the turnaround time needed to process seismic data. The second revolutionary development is the integration of engineering workstations into seismic processing environments. As a result, plots that used to take one-half to two days to complete now take one to two seconds to complete. These developments make possible interactive execution of many types of seismic processing that traditionally were executed in batch mode.

Seismic velocity analysis is one of the more important tasks that now can be processed interactively. Seismic velocity analysis can include many of the traditionally separate batch methods of the past, and if performed on a sufficiently fast computer, can include methods such as modeling and prestack depth migration.

Seismic velocity analysis

Seismic data are generated by an explosion, called a seismic shot, set off at the surface of the earth. Receivers, called geophones, record the echo from the shot as it bounces back to the surface from various underground levels. Velocity analysis uses the recorded echo to determine the velocity of sound inside the earth. Once the velocity, which varies from place to place inside the earth, is determined, computers remove the distortions in the data created by the velocity variations.

The most desirable type of velocity analysis is iterative prestack depth migration. However, this method also uses the most computer time and disk space. This method of seismic data processing enables seismic interpreters to test the validity of their velocity models effectively and completely. A tradeoff exists, however. The more accurate the method used to resolve seismic data, the greater the amount of computer time needed, and the more important the accuracy of the velocity model.

Less data actually are moved over the line if the system is truly interactive.

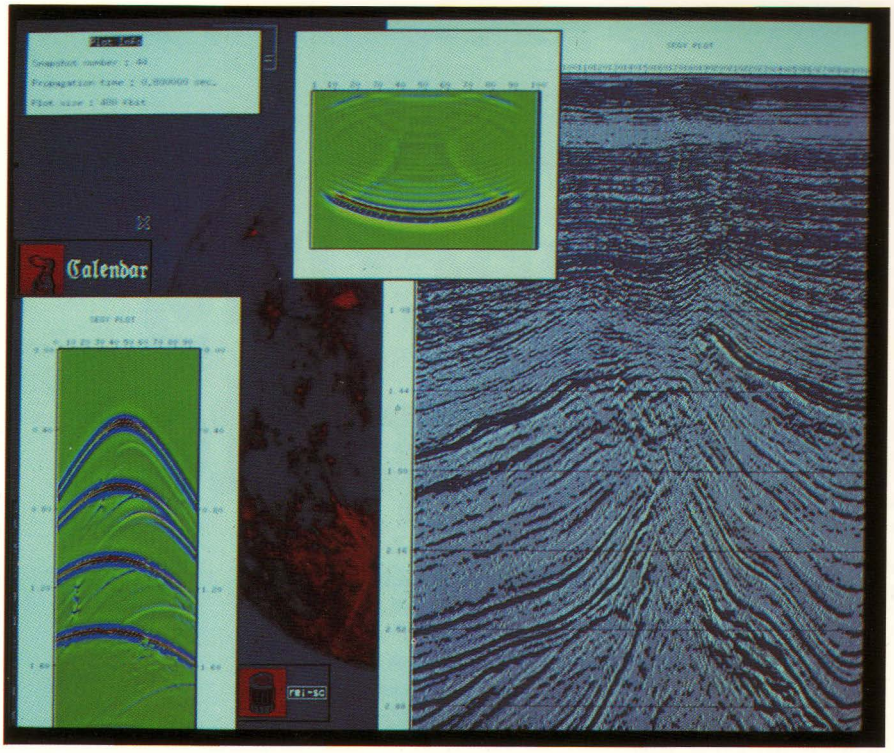


Figure 1. Windowing lets interpreters compare the output of a model side-by-side with field data. Clockwise from upper left: time slice plot of a three-dimensional finite difference acoustic seismic model, plot of a field data set, and plot of a full elastic wave equation model. Elastic model courtesy of Geophysical Systems.

Iterative prestack velocity methods

Windowing systems and Cray Research's interactive UNICOS operating system make a series of powerful techniques available to petroleum engineers, or interpreters. During normal velocity analysis, an interpreter selects reflections from velocity contrasts in the earth, called events, and discerns the depth and velocity of the events from the time delay of the events for each shot-receiver pair. Typically this method of velocity analysis is performed in batch mode, which means the interpreter submits a job to the mainframe computer and one-half to two days later the plot returns. At that point the interpreter chooses the events and a technician reads the marks on the plots and types the choices into a computer file that is used for continuing processing.

Under Cray Research's UNICOS operating system, the interpreter has several tools that eliminate the drudgery of this task. The UNICOS operating system allows the interpreter to hold several shots, called an aperture, in memory for faster access. The extra memory also allows sufficient I/O buffering for fast access to the data. With a Cray system running the UNICOS operating system and a workstation running the UNIX operating system, an interpreter can use the seamless UNIX environment to process the data graphically and interactively on the workstation, while keeping the data on the Cray system where it can be used most effectively.

The UNICOS operating system also makes possible portable programs and portable programmers. Because the same operating system works on the workstation and the supercomputer, only one class of programmers needs to be trained. The seamless environment makes transfer of data between computers so easy that graphics data can be sent directly to a Cray system. The interpreter can select events on the workstation screen. The results of those events are entered automatically into a file that is sent to the Cray system for further processing. This convenience not only eliminates the possibility of transcription errors, but also dramatically speeds up the iterative velocity analysis cycle. Plots come back in seconds instead of one-half to two days. This not only produces a quantitative speedup in the velocity analysis process, but also a qualitative change in the methods used.

Modeling

Modeling traditionally has been reserved for postmortems of seismic processing because models too often have not represented the field data accurately. Thus, the most popular method of modeling has been ray-tracing, which accounts for only the arrival times of an event, without considering the wave equation effects in more than an ad hoc manner. The lack of fit between modeling and field data has been the result of an inability to run a sufficient number of models and an inability to compare model data to field data effectively.

The main parameter determining event shape in a full wave equation model is the source signature, also called the source wavelet, which designates the type of source. Changing the wavelet changes the shape and time position of events. On small computers, the choice of wavelet has been severely restricted for reasons of model size and available computer time. A more accurate wavelet enlarges the model and requires increased CPU time. Supercomputers make the matching of wavelets to field data feasible, so that a high-bandwidth model can be run fast enough for an interpreter to iterate it until finding a correct wavelet match to the field data. A good match is necessary not only for processing field data via modeling, but also for validating seismic inversion methods.

A workstation interface to Cray systems, primarily under UNICOS, makes feasible the iterative process of wave matching. Comparing paper plots does not allow for a correct match. But windows such as those shown in Figure 1 allow an interpreter to compare the output of a model side-by-side with the field data. This capability, along with quick model turnaround, allows the iterative matching of wavelets necessary for full acceptance of modeling in the velocity analysis portion of the seismic processing stream.

Figure 1 shows two types of models on a workstation display: a three-dimensional finite difference acoustic model in the upper screen and an integral equation model in the lower left-hand corner. The field data are plotted on the right. The three-dimensional model was developed at Cray Research and runs interactively on CRAY-2 systems, which allows several iterations per session for fit to field data. The integral equation model, from Geophysical Systems, models a flat-layered earth with a fully elastic wave equation; this model also can be run interactively. Thus, one model accounts for the three-dimensional effects of a velocity model, while the other model accounts for the effects of the full wave equation, that is, for shear-to-compressional wave conversion and attenuation. Both can be run interactively, making practical the wavelet matching techniques necessary for the use of modeling in velocity analysis.

Prestack depth migration

Prestack depth migration is another seismic processing method used extensively in the petroleum industry. But only the most advanced users apply it to velocity analysis, an application that requires a graphics-based workstation linked directly to a supercomputer running the stacking program interactively.

Figure 2 shows a result created with the PREMIG migration program from GeoQuest International. The figure shows a migration of a 96-trace, 600-sample shot that runs in less than five seconds. The three side-by-side plots show the results of three velocity models. The geology of the area indicates that the interfaces between regions should be flat with a fault (break) of the interface on the right. Thus, the migration of the data on the right and left pictures can be ruled out; the migration as presented in the middle picture appears to be most valid. Note that because this earth has dip (the interfaces are slanted), the more traditional normal moveout (NMO) stacking would not image the earth as accurately. In fact, NMO stacking cannot stack the single shot as shown here because NMO stacking needs much more data redundancy.

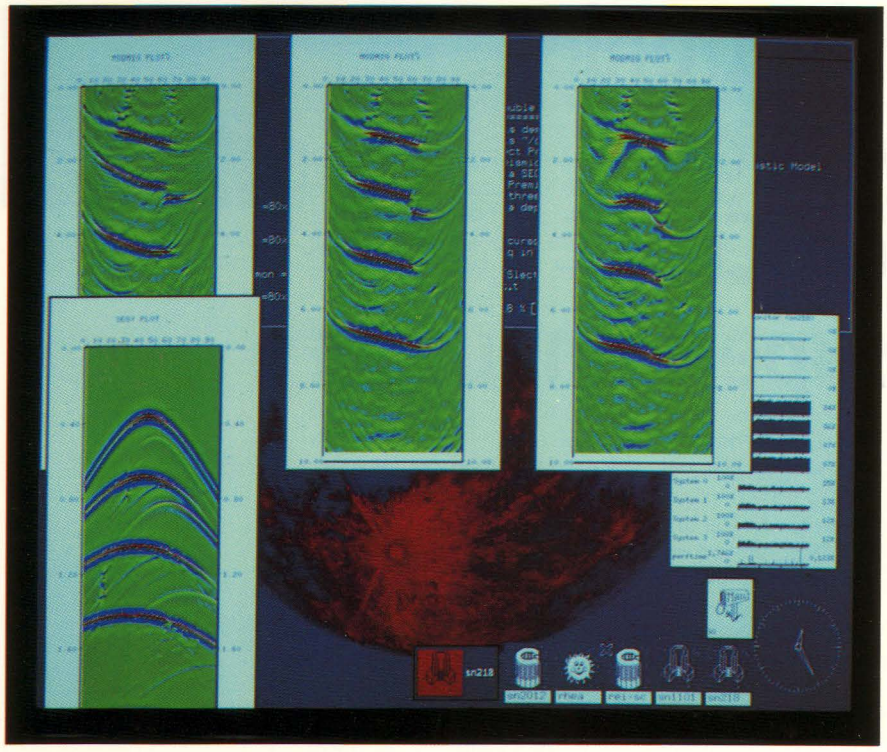
Many other methods of prestack migration are available, including Stolt, Kirchoff, and wave-equation datuming. The method used here is not unique, but it is new and exciting when used in this way. Interpreters now can turn around the data fast enough to use the depth migration algorithm in the all-important velocity analyses. New applications for this and other more CPU-intensive tools, such as inversion, guarantee that supercomputers will continue to play a central role in seismic processing.

Case study: Amoco

Amoco engineers have been using the CRAY-2 computer system at the Minnesota Supercomputer Institute to apply these and other methods. An AT&T T1 line from Minneapolis to Tulsa extends the Ethernet network from the CRAY-2 system to workstations at Amoco running the UNIX operating system. Graphical data are transferred through the line at speeds approaching the local Ethernet speeds of about 200 Kbytes/sec. Response time is the same as that for local users. Moving seismic data sets is not feasible with these techniques, so seismic data are preprocessed into graphical data on the CRAY-2 system before being transferred.

When looking at a three-dimensional volume of data, transferring two-dimensional slices of it through the line and examining them interactively is a cost-effective approach. The method may seem counterintuitive given the widely held notion that interactive processing requires more bandwidth than batch processing. But our experience indicates that less data actually are moved over the line if the system is truly interactive.

In a truly interactive system, graphics, rather than data, are moved over the line, and the user aborts the interactive program when the results are not acceptable. This option enables users to abort a group of plots if the first few are unacceptable, a resource-saving option not available on batch systems.



In this environment, Amoco engineers execute prestack depth migration, two- and three-dimensional seismic modeling to fit field data, and two-and three-dimensional seismic modeling to explore underground mechanisms of wave propagation.

Figure 2. One shot of input data from a Society of Exploration Geophysicists Y-tape (left), and prestack depth migrations of the input data with a velocity model that is (left to right) too high (fast), correct (matches synthetic plot), and too low (slow). Images courtesy of GeoQuest International.

Resource savings

Supercomputers and engineering workstations integrated into the computing environment have brought about a qualitative change in seismic data processing methodology. For example, prestack depth migration now is an option for velocity analysis, whereas the method previously was available only in selected batch cases. Supercomputer power also can be used in simulated annealing methods of velocity model perturbation to achieve more accurate results. Cray systems save users time and other resources by enabling them to use complicated methods more often and to use traditional batch methods interactively.

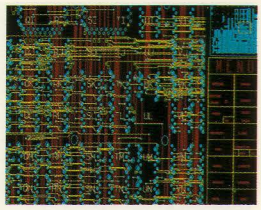
About the authors

Bill Kamp is manager of petroleum applications in the Industry, Science & Technology Department of Cray Research. Prior to joining Cray Research in 1986, he was on the research staff at Phillips Petroleum and at Standard Oil of Ohio (now British Petroleum North America). Kamp received a B.S. degree in mathematics from the University of Minnesota in 1971 and a Ph.D. degree, also in mathematics, from the University of Minnesota in 1975.

Olivier Lhemann joined Cray Research in 1986 and is a senior analyst in the Industry, Science & Technology Department. He has an engineering degree from the Ecole Nationale Supérieure des Télécommunications, Paris, France, and a Ph.D. degree in computer science from the University of Houston.

Electronic design on Cray supercomputers

Erwin Liu and Vince Vasquez, Cray Research, Inc.



Advances in semiconductor fabrication technology have steadily shrunk the size of electronic transistors. As a result, electronic circuit designers are packing increasing numbers of transistors onto single-chip integrated circuits (ICs). To design and manufacture these complex circuits quickly and cost-effectively, design and process engineers are turning to computer simulation as an alternative to physical prototyping.

Simulation is gaining in importance largely because electronic products often have narrow market windows and short life cycles. A report from the marketing research firm McKinsey & Company states that in a high-growth market with five-year product life cycles, shipment of a product six months late can reduce profits by 33 percent over the life of the product.

To manage IC and system design efficiently, the industry has developed a hierarchical design methodology in which a design is decomposed recursively into smaller pieces. Powerful engineering workstations can simulate these individual pieces cost-effectively. Workstations, however, generally cannot simulate in a reasonable time an entire design once the individual pieces have been verified. Now, Cray systems can meet the need for cost-effective turnaround of entire design simulations.

Digital circuit design

In digital design, a system design team uses a product specification to develop a rough architecture composed of a number of large functional blocks, such as the data path, memory, control logic, and clock. They then use a functional/logic simulator to simulate the system comprising these functional blocks. A functional/logic simulator models the operation of functional circuits with well-defined input/output pins that assume only discrete logic values. The internal composition of these functional models is disregarded. The high-level models then are decomposed into smaller blocks that may be modeled either by other functional models or at the gate level. However, several problems have emerged as larger numbers of transistors have become incorporated into circuit designs.

A set of 20 experiments required more than two months of CPU time on a minicomputer (but) less than three hours on a CRAY X-MP/48 system.

The fundamental problem is the sheer complexity of the design task. Fortunately, a number of solutions have arisen. For example, designers are using functional simulation more extensively to help decrease the level of detail that must be considered at each level of design. In addition, unified simulation environments are available to better track designs and manage data. Cray Research's UNICOS operating system, based on AT&T's UNIX system, and Cray Research's support of the TCP/IP protocol for communications enable Cray systems to be integrated easily into most electronic-design environments. A Cray system can be used as an integrated point accelerator for the same programs that run on workstations. The performance advantage provided by Cray systems greatly simplifies the design process because an engineer need not be concerned with using a different simulator when more performance is required.

Another problem created by the new levels of circuit complexity is the corresponding complexity of large systems that incorporate the new circuitry. Functional/logic simulation helps reduce the amount of detail needed at a particular design stage, but large system simulations still require extraordinary computing resources. Engineers at Apple Computer are addressing this challenge by simulating future products in their entirety on Apple's CRAY X-MP system. The system's fast processors enable engineers to exhaustively investigate many options.

Incorporating analog circuitry into digital circuit designs further complicates the designs. Current estimates predict that designs typically contain 60 percent pure digital, 10 percent pure analog, and 30 percent combination analog and digital circuitry. Most CAD tools can simulate either digital or analog circuits, but not both. When mixed-mode simulators mature, however, designers will have an environment in which they can model on a single tool designs comprising functional, gate, and analog circuits. This type of simulator will require the high-performance of Cray systems, which can run digital and analog algorithms quickly and precisely.

Analog circuit design

Analog circuit designers face a set of problems slightly different from those faced by digital circuit designers. Analog circuit design presents problems in which voltage waveforms must be characterized and understood in great detail. Modeling input/output with simple discrete states, as in digital simulations, does not provide enough accuracy for analog circuit design, which must account for numerous parameters, such as noise, gain, coupling, input mismatch, slew rate, and frequency response. The primary application is circuit simulation, which models the operation of electronic devices such as transistors, capacitors, and resistors.

In addition, traditional breadboarding techniques using discrete components can be inadequate for design verification because a design on a monolithic IC can have very different operating characteristics from breadboarded components. Verification during design is desirable because debugging a manufactured analog chip is extremely difficult, requiring accurate control and measurement of circuit voltages and currents.

A group of analog circuit designers at Hughes Aircraft recently faced a difficult analog circuit design challenge. Their designs typically require three years and four iterations before entering production. With a need to integrate 10,000 bipolar transistors on a single chip, these engineers decided to investigate full-chip circuit simulation to verify the design before manufacturing. By using a Cray supercomputer, they reduced the design cycle to 15 months and designed a product that worked the first time as modeled.

Memory design also requires extensive circuit simulation because of the large number of analog components involved. In fact, it is estimated that 700 to 1000 hours on a Cray system are required to design a large memory chip. Immense competition in the memory marketplace, small market windows, and low profit margins increase the importance of reducing the design cycle.

Testability and reliability

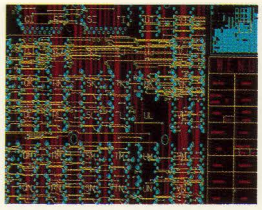
As the number of transistors on a chip increases, testing chips before they are shipped also becomes increasingly difficult. Small dust particles can fall on a chip and disable a small portion of it, challenging designers to determine how well input test vectors can detect a minor disabling of the chip. Fault simulators help make this measurement, but fault simulation is computationally very intensive. This application was a primary motivation for Fairchild Semiconductor Corporation's acquisition of a Cray system, which the company uses for gate-array application specific integrated circuit (ASIC) design.

Production reliability is a crucial element in maintaining a competitive product. A primary method used to increase product reliability is extensive simulations across a matrix of operating and manufacturing parameters such as ambient temperature, process parameters, and operating voltage levels. This added requirement clearly increases the strain on computational resources and design-cycle pressures.

Device and process design

As the dimensions of electronic devices enter the submicron range, the cost and difficulty of developing the next generation of semiconductors increases greatly. Process engineers will require three-dimensional process and device simulators to explore alternative fabrication processes before making physical prototypes. Most such simulators today can model only in one or two dimensions.

Completely simulating and characterizing fabricated transistors using an engineering workstation and conventional two-dimensional process and device simulators would take days. Cray supercomputers substantially reduce this long simulation time, making such simulations practical tools. As a result, engineers can develop much better fabrication processes by running simulations with fine-tuned process parameters. Researchers at AT&T Bell Laboratories are using the labs' CRAY X-MP system for process and device simulations. For example, they are using device simulation to study the latchup characteristics of CMOS fabrication processes used in-house and to change process parameters to enhance latchup immunity.



Preliminary studies indicate that the computational requirement of the more accurate technique will be several times that of the conventional methods. In a separate effort, Cray Research and University of Illinois investigators simulated a one-micron high-electron mobility transistor (HEMT) using the three moments of the Boltzmann equation. HEMT is a high-speed transistor structure fabricated with gallium arsenide and aluminum gallium arsenide. A set of 20 experiments required more than two months of CPU time on a minicomputer. The same experimental results were obtained in less than three hours on a CRAY X-MP/48 system. This study has led to better understanding and more accurate prediction of the operating characteristics of HEMT.

Moreover, as device dimensions continue to shrink, and device structures become more complex, quantum physics rather than classical physics will govern the operating characteristics of devices. A device simulator based on the laws of quantum mechanics would be impractical to use without the processing power of a supercomputer.

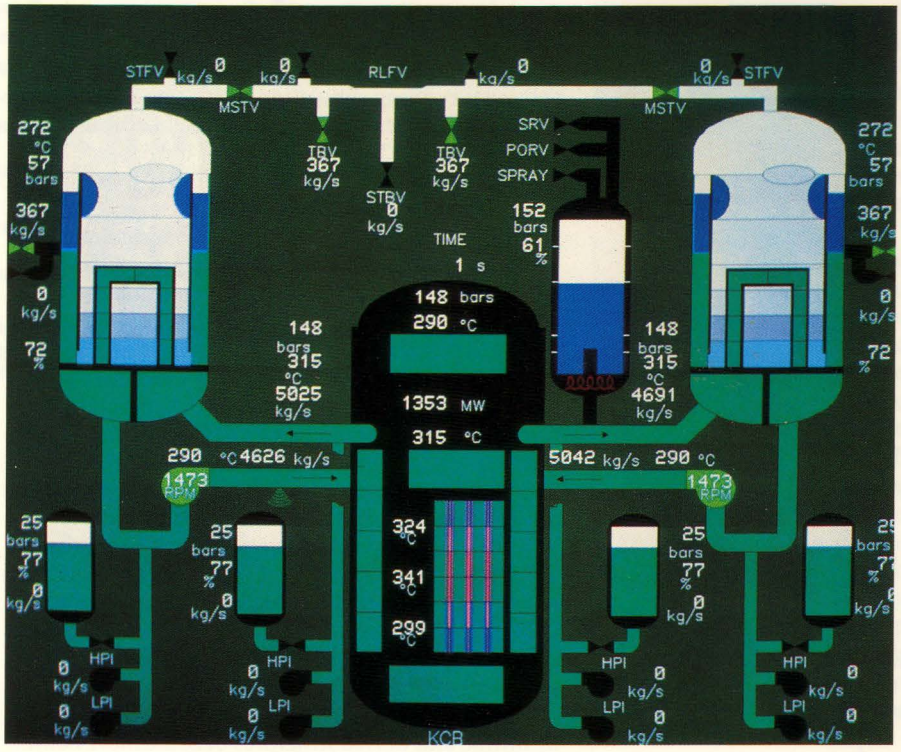
Eliminating the bottlenecks

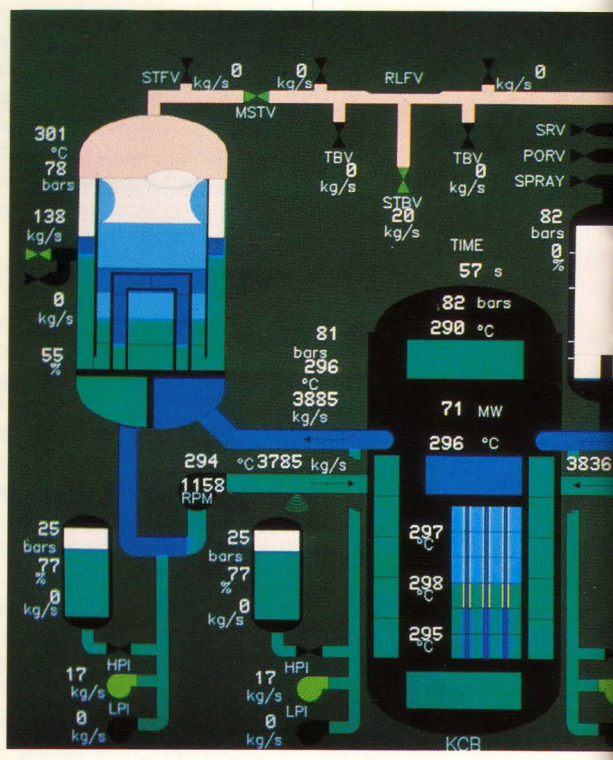
From digital, analog, memory, and process design to testability and reliability analysis, computational bottlenecks can undermine the bottom-line effectiveness of a new product offering. Supercomputing is an effective way to break through design bottlenecks and help ensure the profitability of new components. Cray supercomputers closely integrated into electronic design environments provide an important tool to accommodate the complexity and quick design turnaround required in commercial integrated circuit development. ▬

About the authors

Erwin Liu is a senior electronics application analyst in the Industry, Science & Technology Department of Cray Research. He received his bachelor of applied science and M.S. degrees in electrical engineering from the University of Ottawa and the University of Calgary, respectively. He received his Ph.D. degree in computer science from the University of Calgary in 1987. From 1984 to 1986, Liu interned at the Fairchild Laboratory for Artificial Intelligence Research and at Schlumberger Palo Alto Research.

Vince Vasquez is an electronics applications analyst at Cray Research. He received his B.S. degree in electrical engineering and computer science in 1983 from the University of California, Berkeley, and his M.S. degree in electrical engineering in 1984 from Carnegie-Mellon University. Before joining Cray Research in 1986, he was a product engineer at Hewlett-Packard and a CAD software engineer at Digital Equipment Corporation.





Designs for safety

Stephen R. Behling, Cray Research, Inc.

The nuclear energy industry applies Cray systems to diverse problems, including reactor design, safety analysis, and planning for the efficient use of uranium fuel. For example, Westinghouse Electric Corporation, a major worldwide vendor of reactors and fuel, has been using Cray systems for all aspects of reactor design, fuel optimization, and safety analysis. In addition, U.S. national laboratories rely on Cray systems to develop computer codes that model reactor operations. In France, where over 75 percent of electrical energy is provided by nuclear power, Cray systems are used by the reactor vendor (La Société Framatome), the utility (Électricité de France), and the government regulatory and research agencies (Commissariat à l'Energie Atomique).

Best-estimate safety analysis

In addition to developing computer programs for licensing analyses, which are based on very conservative assumptions, the U.S. Nuclear Regulatory Commission (NRC) has developed programs to provide accurate, or "best-estimate," analyses of the events during a reactor accident. Best-estimate models must accurately account for the nuclear physics of fission, the thermodynamics of boiling water, the thermal hydraulics of flow through pumps, valves, pipes, and other complex geometries, the temperature of the reactor structure and fuel, and the transport of energy between these structures and the water.

The NRC has developed two best-estimate codes to model these processes: TRAC-PF1/MOD1 and RELAP5/MOD2. The TRAC code was developed at the Los Alamos National Laboratory and the RELAP5 code was developed at the Idaho National Engineering Laboratory (INEL). These codes divide a reactor into fluid volumes, or nodes, and connect the nodes by flow paths. Mass and energy conservation equations for a two-phase fluid (liquid water and steam) are solved in the nodes and two-phase momentum equations are solved in the flow paths.

Large, fast computers are critical for improving the precision of reactor models because precision depends on the number of nodes used in the models. Prior to the availability of Cray systems, models of entire power plants had to be limited to less than about 200 fluid nodes. However, researchers at Texas A&M University recently performed some transient analyses on a CRAY X-MP system using a 2500-node model of a nuclear reactor that they constructed for the TRAC code. Comparison of this model against less-detailed models will help determine the circumstances under which the smaller models provide sufficient accuracy.

The Nuclear Plant Analyzer

The Nuclear Plant Analyzer, or NPA, is a program developed by the NRC at INEL to provide sufficient display capabilities so that the best-estimate TRAC and RELAP5 codes can be used as power plant simulators. These codes then can be used to train operators, just as flight simulators are used to train

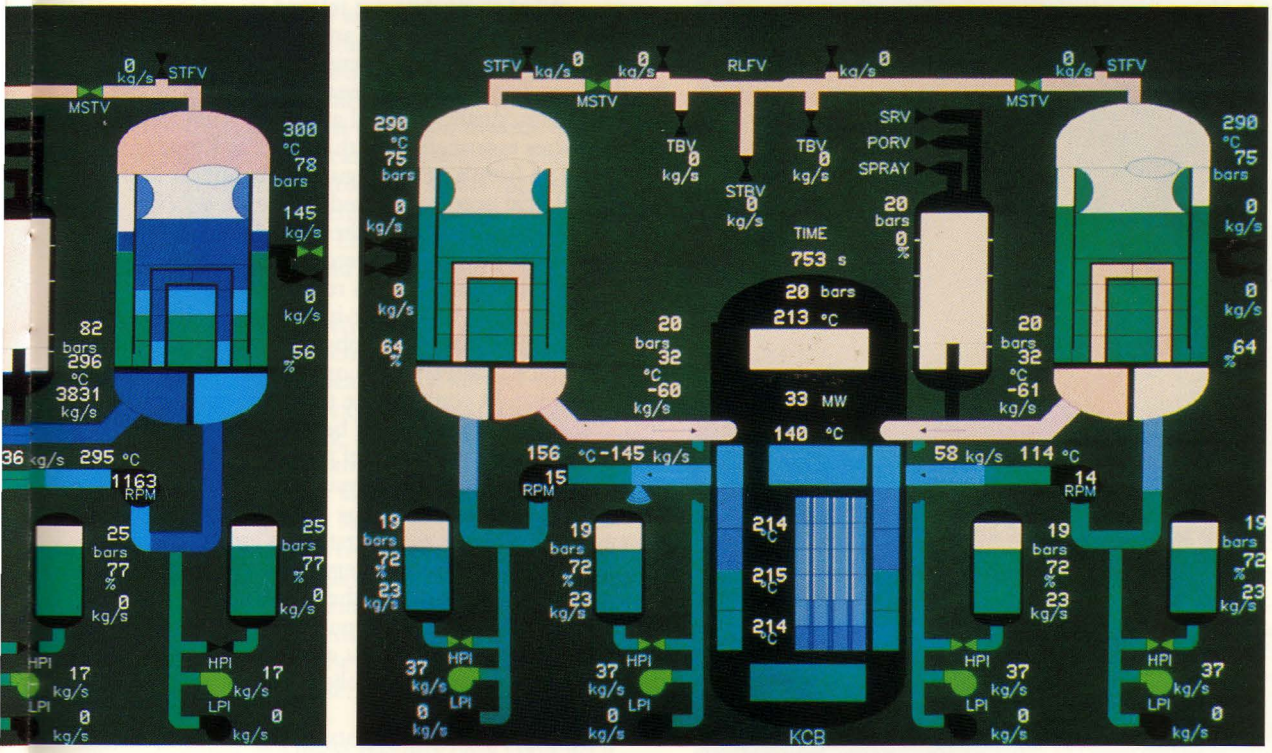


Figure 1. NPA displays of the Borssele reactor, a 452-megawatt electric (1388-megawatt thermal) two-loop pressurized water reactor in the Netherlands. The data and the image template were produced at INEL. Shown are the reactor vessel, the two piping loops with coolant pumps, the pressurizer, the two steam generators, some of the piping that directs steam to the turbine generators, some of the piping that returns cooled water from the condenser, and the emergency coolant systems that inject water into the system if coolant is lost due to a pipe break or other cause.

Bright green represents an open valve or a pump in operation. Black represents a closed valve or a pump not in operation. Other colors represent liquid water cooler than the boiling point (dark green), liquid water at the boiling point (dark blue), steam at the boiling point (white), mixtures of liquid and steam (intermediate shades of blue), and steam hotter than the boiling point (pink).

airplane pilots. This capability is being pursued because Cray systems now can run complex models in real-time or better.

Figure 1 was generated with the NPA software converted to run on a CRAY X-MP system under the UNICOS operating system. The left image shows the reactor in a normal operating condition. The liquid in the primary systems (reactor vessel, piping, and the inside of the U-tubes in the steam generators) is subcooled except in the pressurizer, where a small amount of water is boiled to maintain the system pressure. On the secondary sides of the steam generators, liquid water enters the steam generators and boils when it contacts the hot steam generator tubes. The steam then leaves toward the turbines where it is condensed and returned to the steam generators. At this time it is assumed that a small break occurs in the left piping loop between the coolant pump and the reactor vessel.

The middle image shows the state of the reactor 57 seconds after the break occurs. The pressurizer (right center) has emptied, water is boiling in the loop piping, the high-pressure injection (HPI) emergency coolant system has activated, and the reactor power has been shut down. The power has decreased from the initial 1353 megawatts, but 71 megawatts of power from the decay of radioactive fission products must be removed to prevent the fuel from heating up and becoming damaged.

The right image shows the state of the reactor 753 seconds after the break. Much of the water in the system has boiled to steam. The fission product decay power has decreased to 33 megawatts. The pressure has decreased enough to open the valves connected to the accumulation tank (just above the HPI pumps) and flood the system with water. The calculation proceeded until the system was refilled.

Future safety

The NPA has the potential to help train plant operators and engineers, simulate planned tests, and — if the codes could run much faster than real time — help mitigate the consequences of accidents. If the NPA had been available at Chernobyl, for example, it would have predicted the condition that resulted in the accident and the release of radioactive material. Similarly, if operators at Three Mile Island Unit 2 had been trained with real-time simulations of various reactor transients in two-phase conditions, they most probably would have had enough knowledge of their plant to prevent the reactor damage that occurred following the rather innocuous initiating events.

As real-time capabilities develop, we can expect that supercomputers and detailed, accurate computer programs will be used to train power plant operators, provide realistic simulations for drills, and plan tests of operational changes. Computer hardware and software now are available to improve significantly our ability to demonstrate the safety of nuclear power plant operations. ▬

About the author

Stephen R. Behling is a nuclear science and engineering applications analyst in Cray Research's Industry, Science & Technology Department. He received a B.S. degree from the University of Wisconsin at Madison in 1974 and worked at the Idaho National Engineering Laboratory from 1975 until 1986, when he joined Cray Research. At INEL he worked on various aspects of reactor safety, including development of the RELAP4 computer code, analysis of experimental systems with the RELAP5 code, analysis of the Three Mile Island accident, and program management functions.

Atmospheric modeling a global view

Eric J. Pitcher, Cray Research, Inc.

Forecasting the weather for the next several days, assessing the influence of the constant interplay between the atmosphere and the oceans on the coming winter's weather, and predicting the climatic consequences of increasing carbon dioxide in the atmosphere all depend upon the lightning speed and processing power of Cray supercomputers.

Tomorrow's weather

Since the introduction of the first electronic computers in the 1940s, numerical weather prediction has taxed the most powerful of each successive generation of systems. That tradition continues today at major weather centers around the world where Cray systems are used, such as the European Centre for Medium-Range Weather Forecasts (ECMWF), the Canadian Meteorological Centre, and the Air Force Global Weather Central. Numerical weather prediction is the process of solving the equations that govern the behavior of the atmosphere subject to specified initial and boundary conditions. The equations describe the fundamental laws of thermodynamics, conserve the total mass of the atmosphere, and keep track of water in its three phases. The Navier-Stokes equations, with some modifications to remove unwanted motions such as sound waves, predict the changing flow patterns.

To predict weather computationally, forecasters first must have detailed knowledge of the atmosphere at some instant; that is, they must have a global "snapshot" of all the required meteorological variables. The observations that constitute the initial state come from a variety of sources. For example, radiosonde balloons are released every 12 hours at several hundred stations around the world. The balloons' instrument packages telemeter temperature, pressure, and moisture content to the release points. Tracking a radiosonde in flight permits the construction of a vertical profile of wind speed and wind direction. Ships at sea, instrumented buoys, and high-flying aircraft also provide information. Geostationary and orbiting satellites complement the Earth-based observation network and help fill gaps in the data, especially over the oceans. Automated procedures have been developed to infer wind vectors from the tracking of cloud elements on successive satellite photographs. Satellites measure the outgoing long-wave radiation from the atmosphere and Earth. The atmospheric temperature as a function of height may be deduced from this radiation.

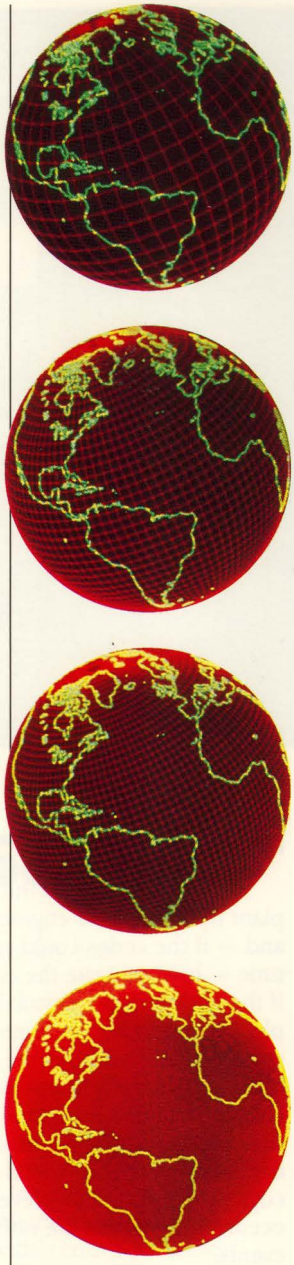


Figure 1. Climate and forecasting model resolutions as depicted by NCAR's CRAY X-MP/48 system. The upper images typify model resolutions used for climate studies, and the lower images show resolutions more typical of weather forecasting models. Image courtesy of Tom Bettge of NCAR.

All of these data are relayed to a global telecommunications network and into the major numerical weather forecasting centers worldwide. Before the data can be used by a forecast model, they undergo preliminary quality control and analysis. Analysis accomplishes essentially two things. First, it interpolates the data from a very irregular distribution to the network of grid points expected by the model (Figure 1). Second, the analysis makes the fully three-dimensional picture of the atmosphere self-consistent. In other words, motion fields, as represented by the wind vectors, must be in near-balance with the pressure fields. Otherwise, small observational errors could produce relatively large oscillations within the model that would contaminate the forecast.

Global analyses can be produced every six hours, with the forecast model providing the necessary "first-guess" for the analysis. The analysis in turn provides the initial state for the subsequent six-hour forecast, which becomes the first guess for the following analysis.[1] In addition to this very short-range forecast required for the next analysis, the model is time-stepped forward several days. The output of this calculation is distributed to regional forecast offices and used by meteorologists to produce public, aviation, and marine forecasts (Figure 2).

Objective assessments of the computer forecast models show that their accuracy has increased steadily over the years. For example, a five- to six-day forecast produced by the ECMWF model in 1987 was more accurate than a three-day forecast in 1977.[2] Accuracy has been boosted by a range of factors including improvements in analysis procedures, more realistic physics, and more refined model resolutions. The availability of supercomputers such as Cray systems has spurred substantial research in numerical weather prediction, and has enabled the benefits of such research to be incorporated into operational forecasts. A forecast is useful only if it is timely. Often, only one to two hours of wall-clock time may be devoted to the actual forecast. Thus, the leading forecast centers must have access to the most powerful machines in existence to further advance the science of numerical weather forecasting.

Is the climate changing?

The year is 2050. The temperature is 95°F in New York City with a relative humidity of 95 percent. The heat wave began in June and will last until the end of September. Alaska's climate is relatively benign, and the state experiences a thriving commercial fishing industry. Farmers are emigrating from the south-central United States, which has become a permanent dust bowl, to western Canada, where the growing season extends two months longer than it does today. The farms of western Canada, along with those of the Soviet Union, provide the majority of the world's food supply.[3]

The above scenario may not actually develop in detail, but it is consistent with results that are emerging from recent global climate models. Figure 3 shows a simple schematic diagram. Climate models include physics similar to that used in weather forecast models, but climate models typically are of lower resolution than forecast models and focus on longer-

term processes (average weather).[4] The hypothetical scenario for the next century is based upon results from lengthy simulations in which carbon dioxide concentrations have been doubled from their present values. The atmospheric content of carbon dioxide has been increasing steadily in recent decades, and this increase can be attributed at least in part to the increasing use of fossil fuels. As a "greenhouse gas," carbon dioxide traps outgoing heat from Earth's surface and atmosphere. Thus, any increase would be expected to generate a warmer climate.

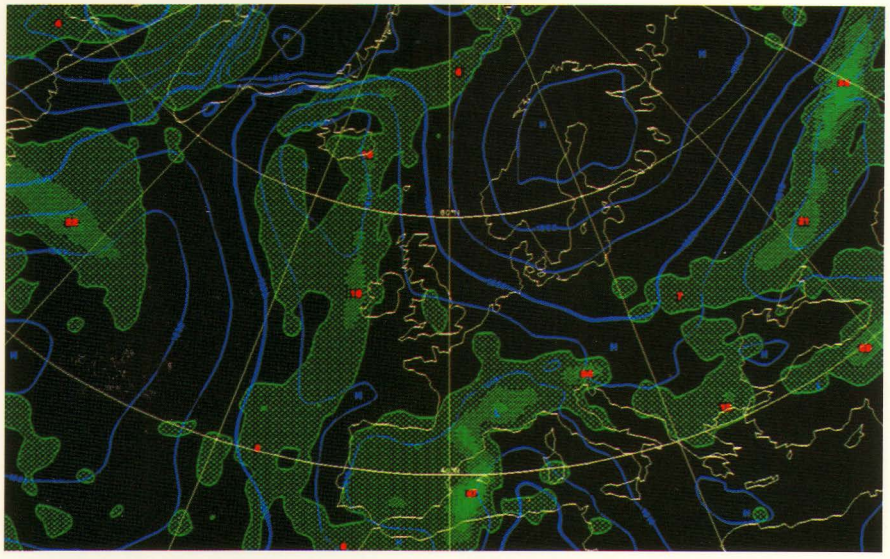
Such experiments cannot ignore the influence of the oceans, which constitute a vast heat reservoir that could extract heat from a warmer atmosphere and lead to a less-than-expected, or at least delayed, carbon dioxide warming. Moreover, the ocean circulations may be so altered that they introduce other secondary effects to the atmosphere. At the National Center for Atmospheric Research (NCAR), scientists Warren Washington and Gerald Meehl are using coupled atmosphere-ocean models to simulate Earth's climate system. These models consume extraordinary amounts of computer time on NCAR's CRAY-1 and CRAY X-MP/48 systems, as well as on the Cray systems at the magnetic fusion energy center at the Lawrence Livermore National Laboratory.

These scientists recently have completed two 30-year simulations with their coupled model. In one simulation, the carbon dioxide content of the atmosphere was doubled at the outset, while the carbon dioxide in the second simulation was permitted to increase gradually at the rate of 1 percent per year. One purpose of these experiments is to determine whether the oceans could act as a climatic buffer and delay warming by extracting heat from the atmosphere. Another purpose is to determine whether the pattern of climate changes evolves differently in the simulation with gradually increasing carbon dioxide compared to the simulation in which carbon dioxide is doubled instantaneously. Washington and Meehl are in the process of analyzing massive amounts of model-generated data to gain further insight into a problem that could affect all of humanity for generations to come.

Future challenges

Whether Cray supercomputers are used to forecast global weather patterns or to generate a detailed simulation of a single thunderstorm, they have been an essential part of the environmental sciences for many years. The forecast models of the coming decade will improve in horizontal and vertical resolution. The execution of such models within stringent time limits will require multiple processors with access to a large central memory.

The coupled atmosphere-ocean models are a first attempt to treat the entire Earth as a system. Contemporary issues such as global warming, acid rain, tropical deforestation, and desertification are multidisciplinary and will be addressed under the sponsorship of several national and international programs. Future climate models will include the many complex and subtle interactions with the biosphere. Such models will require supercomputers of unprecedented power to address these problems effectively.



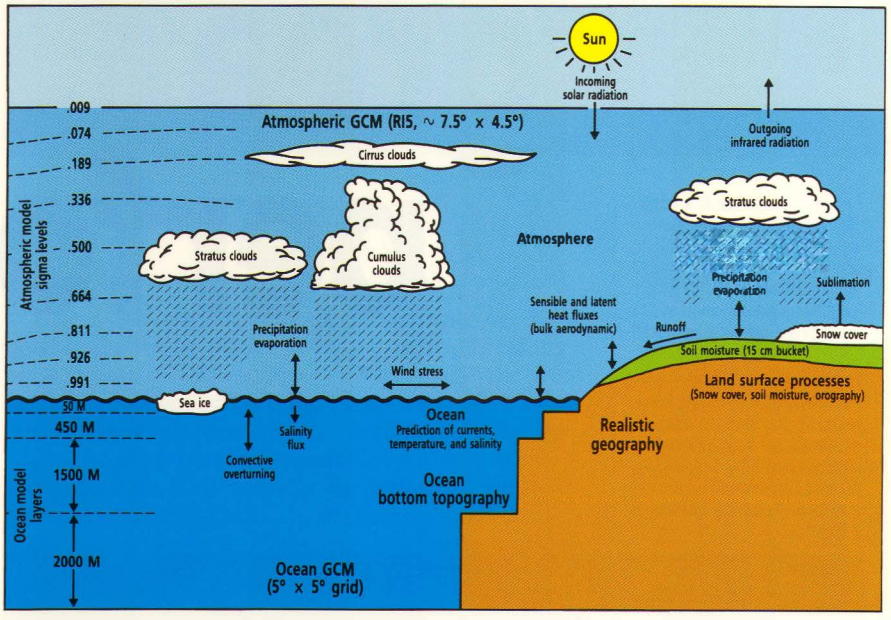


Figure 2 (Top). Forecast of sea level pressure (solid contours) and precipitation (stippled regions) from the ECMWF operational model. Image courtesy of ECMWF.

Figure 3 (Bottom). Characteristics of NCAR's coupled ocean-atmosphere climate model showing the interplay of many physical factors. Image courtesy of Gerald Meehl of NCAR.

About the author

Eric Pitcher is a senior meteorological analyst in Cray Research's Industry, Science & Technology Department. He received a Ph.D. degree in atmospheric science from the University of Michigan in 1974. Prior to joining Cray Research in 1987, he was on the faculty of the Rosenstiel School of Marine and Atmospheric Science of the University of Miami. Pitcher also has worked at the National Center for Atmospheric Research and the National Meteorological Center.

References

1. Simmons, A. J., "Numerical Prediction," *Anomalous Atmospheric Flows and Blocking, Advances in Geophysics*, Academic Press, 1986.
2. Dent, D., and R. Gibson, "Multitasking the ECMWF Weather Forecasting Model," *CRAY CHANNELS*, Winter, 1988, pp. 2-7.
3. Bretherton, F., "The Global Change Program," Sixth Annual SCD Users Conference, NCAR, January 11-14, 1988.
4. Washington, W. M., and C. L. Parkinson, *An Introduction to Three-Dimensional Climate Modeling*, University Science Books, 1986.

Small steps toward great performance

Tony Warnock, Cray Research, Inc.

Figure 1. Matrix-oriented example shows gain resulting from restructuring loops to minimize overhead, (a) a set of arrays dimensioned 8-by-4-by-8-by-4, (b) loops rewritten in more normal form. Resulting speedup is 7.3 to 1.

```
      DO 600 I = 1, 8                          ! loop setups       1
      DO 500 J = 1, 4                          ! loop setups       8
      DO 400 K = 1, 8                          ! loop setups      32
      DO 300 L = 1, 4                          ! loop setups     256
      C(I, J, K, L) = 0
      DO 200 M = 1, 8                          ! loop setups    1024
      DO 100 N = 1, 4                          ! loop setups    8192
      C(I, J, K, L) = C(I, J, K, L) + A(I, J, M, N) * B(N, M, K, L)
100   CONTINUE
200   CONTINUE
300   CONTINUE
400   CONTINUE
500   CONTINUE
600   CONTINUE
```

a

```
      DO 100 I = 1, 32                         ! loop setups       1
      DO 200 J = 1, 32                         ! loop setups      32
      C(I, J) = 0
100   CONTINUE
200   CONTINUE
      DO 500 K = 1, 32                         ! loop setups       1
      DO 400 J = 1, 32                         ! loop setups      32
      DO 300 I = 1, 32                         ! loop setups    1024
      C(I, J) = C(I, J) + A(I, K) * B(K, J)
300   CONTINUE
400   CONTINUE
500   CONTINUE
```

b

To take full advantage of the unmatched performance of Cray Research hardware, programmers should consider some basic performance factors when designing programs. Four of the more important factors affecting code performance are algorithm selection, overhead, vectorization or parallelization, and memory contention.

The following examples of the effects these factors have on performance are derived from real-life codes developed by Cray system users. The classification of examples is somewhat arbitrary because in reality performance factors interact. Nevertheless, addressing these factors appropriately will improve throughput and code performance.

Algorithm selection

The most important factor determining the speed of a code is the algorithm used. A bad algorithm, however beautifully coded, will not perform well. For example, a set of 1024 numbers and a set of 16,384 numbers were sorted using two different methods. The first algorithm was a vectorized version of Batcher's odd-even parallel exchange. Execution time for this sort is proportional to N^2, where N is the number of elements to be sorted. The algorithm often

is discussed in articles and books on parallel processing. The articles suggest that this method's parallelism makes it an efficient method. However, this is true only if the number of processors equals the number of elements to be sorted. The second algorithm used was a scalar version of *quicksort*, a method proportional to *N*log(N)* for *N* elements. The superiority of quicksort is evident, as illustrated by the following results. The speedup made possible by using the better algorithm differentiates between a nearly impossible computation and a trivial one.

Elements	Odd-even	Quicksort	Ratio
1024	350 ms	13 ms	27 to 1
16,384	86,817 ms	260 ms	334 to 1

Overhead reduction

An algorithm often can be coded correctly in Fortran in many different ways, and the run times may vary greatly. For example, when multiplying two matrices, a user usually does not care how many indexing operations are done, but does care that the appropriate multiplications and additions of the matrix elements are completed. Some overhead is necessary to start any computation, but if overhead dominates total run time, the code is faulty. The following examples illustrate some types of overhead reduction.

The mechanism to set up a Fortran DO-loop or similar structure is time-consuming. For example, the number of trips must be computed, various counters must be initialized, and loop invariants must be calculated. One hallmark of efficiently coded nested DO-loops is the order of the loops, which has been chosen to minimize the number of loop startups. The following two matrix-oriented examples illustrate the gain resulting from restructuring loops to minimize overhead. In the first example, a set of arrays dimensioned 8-by-4-by-8-by-4 was used, a form suggested by the physics of the problem. The arrays were used as 32-by-32 complex matrices in a complex matrix multiply. The loops are nested six levels deep in the four-dimensional formulation, as opposed to three levels in a two-dimensional form (Figure 1a).

The code has 9513 loop startups, compared to 32,768 trips through the inner loop consisting of one multiply and one add. On Cray computers, loop startups are important because, as shown in the matrix multiply example, the inner loop executes in approximately one clock tick per trip. Figure 1b shows the loop rewritten in a more normal form.

The revised code has a total startup count of 1080, including the extra loop to zero the result array. Differences in timing for 1000 sets are shown below.

Before	22.441 ms
After	3.082 ms
Speedup	7.3 to 1

For comparison with scalar machines, a run was made with the vectorization turned off, as shown in the following timings. The problem was run on a CRAY X-MP/416 computer system running CFT 1.14. Other computers and compilers will show slightly different speedups, so mileage will vary.

```
      DO 40 L = 1, 512                                  !    1 startup
      DO 30 I = 1, 3                                    !  512 startups
      DO 20 J = 1, 3                                    ! 1536 startups
      C(L, I, J) = 0
      DO 10 K = 1, 3                                    ! 4608 startups
      C(L, I, J) = C(L, I, J) + A(L, I, K) * B(L, K, J)
10    CONTINUE
20    CONTINUE
30    CONTINUE
40    CONTINUE
```

a

```
      DO 3 J = 1, 3                                     ! 1 startup
      DO 2 I = 1, 3                                     ! 3 startups
      DO 1 L = 1, 512                                   ! 9 startups
      C(L, I, J) = 0
1     CONTINUE
2     CONTINUE
3     CONTINUE
      DO 40 K = 1, 3                                    ! 1 startup
      DO 30 J = 1, 3                                    ! 3 startups
      DO 20 I = 1, 3                                    ! 9 startups
      DO 10 L = 1, 512                                  !27 startups
      C(L, I, J) = C(L, I, J) + A(L, I, K) * B(L, K, J)
10    CONTINUE
20    CONTINUE
30    CONTINUE
40    CONTINUE
```

b

Figure 2. A set of 512 3-by-3 complex matrix multiplies, (a) original code with loop length of 512 and a 3-by-3 matrix multiply inside, (b) rewritten loop with 512 in the innermost position. Resulting speedup is 6 to 1.

Before	21.884 ms
After	16.673 ms
Speedup	1.3 to 1

The speedup in scalar mode is not nearly as great as the vector-mode speedup for several reasons. Most importantly, the multiply and add cannot be overlapped in straight scalar code. On a CRAY X-MP system, the multiply time is seven clock periods and the add is six clock periods in scalar mode. In vector mode, both multiply and add are performed jointly in one clock period. (This is 13 times faster than the scalar speed.) In scalar mode, the time for the work is comparable to the loop setup overhead, whereas in vector mode the work is done much faster, so overhead dominates.

The next example (Figure 2a) includes a set of 512 3-by-3 complex matrix multiplies. The small size of the matrix indicates that overhead could be important for inefficiently coded Fortran. The code's original form had a loop of length 512 with a 3-by-3 matrix multiply inside. The code segment has 6657 loop startups for 13,284 trips through the inside (useful) statement.

The loop was written to place 512 in the innermost position (Figure 2b). Even with the extra loops to zero the matrices, the loop startups are 100 times fewer than with the first code, as illustrated below.

Form	Time (ms)	Rate (MFLOPS)
512-by-3-by-3-by-3	5.220	19.4
3-by-3-by-3-by-512	.858	118.2

Rewriting the loops has provided a six-fold speedup. The speedup is not proportional to the startup reduction time, since for the complex matrix multiply, the arithmetic work in the loop is not small compared to the loop overhead. That is, a complex

```
      DO 10 I = 1, 6
      C(I) = 0.0
10    CONTINUE
      DO 30 I = 1, 6
      DO 20 J = 1, 6
      C(I) = C(I) + A(I, J) * B(J)
20    CONTINUE
30    CONTINUE
```

a

```
      DO 10 I = 1, 6
      C(I) = 0.0
10    CONTINUE
      DO 30 J = 1, 6
      DO 20 I = 1, 6
      C(I) = C(I) + A(I, J) * B(J)
20    CONTINUE
30    CONTINUE
```

b

Figure 3. Code examples optimized to increase performance, (a) original code, (b) reversed loops, (c) unrolled loops, (d) batched protons.

```
      DO 10 I = 1, 6
      C(I) = A(I, 1) * B(1) + A(I, 2) * B(2) + A(I, 3) * B(3)
     &     + A(I, 4) * B(4) + A(I, 5) * B(5) + A(I, 6) * B(6)
10    CONTINUE
```

c

```
      DO 20 I = 1, 6
      DO 10 K = 1, 1024
      C(K,I) = A(K, I, 1) * B(K, 1) + A(K, I, 2) * B(K, 2) + A(K, I, 3) * B(K, 3)
     &       + A(K, I, 4) * B(K, 4) + A(K, I, 5) * B(K, 5) + A(K, I, 6) * B(K, 6)
10    CONTINUE
20    CONTINUE
```

d

multiply takes four real multiplies and two real adds, while a complex add requires two real adds. Each trip through the inner loop performs one complex multiply and one complex add, which together count as eight operations.

Vectorization

A code is vectorized when the compiler determines that the code can be executed using the computer's vector hardware. A code running in vector mode generally has less overhead than one running in scalar mode. For example, loop counters on a Cray system running in vector mode are tested on every 64th element rather than on each element. Roughly only 1/64 times as many CPU instructions are needed to execute a code. Vectorization on Cray systems exploits parallelism at the CPU level by allowing for continuous or even concurrent execution on various functional units. Each functional unit, such as the adder, multiplier, or shifter, can operate independently of the other fundamental units (if they are given independent operands). In addition, each functional unit is pipelined, meaning that it is divided into independent segments, each of which can be executed as soon as it has the operands from the previous segment. That is, if a functional unit is fed a continuous stream of data from a vector register, then the functional unit will produce nearly one word of output at every tick of the basic clock.

Furthermore, most Cray models support chaining. In this process, as output of one functional unit begins to fill a vector register, that same output can be used immediately as an input operand for another functional unit. Thus, multiple vector operations can be performed concurrently using the same vector data. Peak vector performance is achieved when a user can write loops that allow the CPU to keep as much of its vector hardware as busy as possible. Often, higher levels of parallelism can be found at the user's program level.

Vector performance

On Cray computer systems, significant performance gain is achieved if a scalar loop can be manipulated to run in vector mode. For example, the table below illustrates the vector-versus-scalar speed for evaluating several library functions.

Function	Scalar time (ms) (1000 calls)	Vector time (ms) (1000 calls)	Ratio
sine	190.0	25.4	7.5 to 1
exp	177.2	15.5	11.4 to 1
log	197.4	20.3	9.7 to 1
sqrt	156.4	8.6	18.2 to 1
ranf	102.6	6.0	17.1 to 1

The following examples (Figures 3a-d) are from a code designed to follow protons in a circular storage ring. At each step the proton's orbit is updated by multiplying a 6-by-6 matrix into a six-element vector. The inner loop of the problem appears in Figure 3a.

The operations in loop 20 are known as reduction, or dot product, operations. This code does not appear to be vector code, since loop 20 cannot be executed in parallel for all values of *J*. However, on Cray computer systems, the compilers use special code that evaluates reductions using primarily vector instructions. In the above example, even better performance (using only vector instructions) can be obtained simply by interchanging the loop indices (Figure 3b).

Loop 20 contains two floating point operations (one multiplication and one addition) and three memory references. *(A(I, J)* and *C(I)* are loaded and *C(I)* is stored; *B(J)* is invariant in loop 20.) Best performance is achieved by maximizing the ratio of floating point operations to memory references. Therefore, outer loop 30 is unrolled (Figure 3c).

Loop 10 has 11 floating point operations and seven memory references. This helps ensure that memory accesses will not limit performance. Figure 3c represents the best manipulation of the given data structure, although the vector length of loop 10 is only six.

The program's data structure can be rearranged to obtain better performance. A batch of protons (1024 in the example) may be run at one time. The outer loop structure may be retained and an inner loop may be added over the batch index *K* (Figure 3d). The results are summarized below.

	Speed (MFLOPS)	Ratio to original	Ratio to previous
Original code	4.2	1.0 to 1	1.0 to 1
Reversed loops	20.0	4.8 to 1	4.8 to 1
Unrolled loops	60.9	14.5 to 1	3.0 to 1
Batched protons	155.9	37.1 to 1	2.6 to 1

This example illustrates two important points: first, even a small problem can benefit substantially from vectorization; and second, more work is required of the user at each step in the optimization. An approximate 5-to-1 speedup is obtained by reversing indices, but the last 2.6-to-1 speedup requires substantial rewriting of the code.

Memory contention

To illustrate the effect of memory traffic on performance, consider the problem of multiplying two *N*-by-*N* real matrices (canned library routines for matrix multiplication, such as those in SCILIB, should be used when possible). The matrix multiply loops are shown in Figure 4a, in which matrix *C* is assumed to be already zeroed.

Loop order is an important performance consideration. Three essentially different forms exist, depending on the position of the *K* (summation) loop. If the *K* loop is the innermost loop, we recognize this as a reduction loop. Since reductions do not execute at full vector speed, they should be avoided when a choice exists. Another reason to avoid reductions in this case is that the memory accesses of *A(I, K)* have a stride or skip distance equal to the first dimension of *A*. *(B(K, I)* is a simple successive-memory location access.) If the leading dimension of *A* is a multiple of the number of the computer's independent memory banks, delay may occur due to memory conflicts. The example in Table 1 is run with arrays dimensioned 256-by-256 to amplify the effect of bank conflicts, and with arrays dimensioned 257-by-256 to eliminate the conflicts.

If the *I* loop or *J* loop is located on the inside, *C(I, J)* is a vector being defined in the inner loop. This form of loop is easy to handle, and the code will run in vector mode. Positioning the *J* loop on the inside causes three memory references with a nonunit stride to occur each time through the loop. If no bank conflicts existed, the loops would be essentially identical. The size used in the example causes the most problems. A loop running over the second index will have an odd stride that causes no conflicts. For dimension size 256-by-256, the case with the *I* loop in the inside clearly is the best. The difference between the *K-J-I* and the *J-K-I* loops is due to second order effects. The result is that the *K-J-I* code runs approximately six times faster than the *K-I-J* code.

Even better performance may be obtained by unrolling the *K* loop *M* times (for simplicity, assume that *N* is a multiple of *M*). The *K-J-I* loop or the *J-K-I* loop could be unrolled on *K*. However, unrolling the *K-J-I* loop turns out to be the more efficient option. The original form of this loop has two floating-point operations and three memory references. Denote this ratio by *R(K-J-I)*:

$$R(K\text{-}J\text{-}I) = 2/3$$

In most cases, it is best to perform operations such as multiplication and addition; memory traffic is just overhead. As memory is slow compared to arithmetic, increasing the floating-point-to-memory ratio improves performance. Rewrite the loop by unrolling *M* times (Figure 4b).

Loop order	Vector rate of 256-by-256 (MFLOPS)	Scalar rate (MFLOPS)	Vector-to-scalar ratio	Vector rate of 257-by-256 (MFLOPS)
I-J-K	35.3	4.6	7.7 to 1	66.6
J-I-K	35.3	4.4	8.0 to 1	66.9
I-K-J	23.1	5.2	4.4 to 1	133.5
K-I-J	22.7	5.2	4.4 to 1	132.2
K-J-I	126.7	5.7	22.2 to 1	134.4
J-K-I	128.3	5.7	22.5 to 1	132.3

Table 1. Results of varied loop order and amplified and eliminated bank conflicts.

Unrolling the *K-J-I* loop on *K* is more efficient than unrolling the *J-K-I* loop for two reasons. First, the former loop has $1 + N/M + (N/M) * N$ loop startups while the latter has $1 + N + (N/M) * N$ loop startups. Second, in the former loop, the indices $K + 1, \ldots, K + M - 1$ are computed only N/M times, while in the latter loop these indices are computed $(N/M) * N$ times. The *K-J-I* version of the matrix multiply has *R*(unrolled *M* times) $= 2M/(M+2)$. For $M = 1$, we obtain $R(M = 1) = 2/3$, as before, but for $M = 6$ we have $R(M = 6) = 12/8$. Thus, the floating point operation to memory reference ratio is 2.25 times larger for the unrolled case.

```
      DO 30 J = 1, N
      DO 20 K = 1, N
      DO 10 I = 1, N
      C(I, J) = C(I, J) + A(I, K) * B(K, J)
10    CONTINUE
20    CONTINUE
30    CONTINUE
```

a

```
      DO 30 K = 1, N, M
      DO 20 J = 1, N
      DO 10 I = 1, N
      C(I, J) = C(I, J) +  A(I, K) * B(K, J) + A(I, K + 1) * B(K + 1, J)
     +...+A(I, K + M-1) * B(K + M-1, J)
10    CONTINUE
20    CONTINUE
30    CONTINUE
```

b

Figure 4. Matrix multiply loops, (a) original code in which matrix *C* is assumed to be already zeroed, (b) loop rewritten by unrolling *M* times. Resulting floating point operation to memory reference ratio is 2.25 times larger for the unrolled case.

Conclusion

Seemingly small details such as loop ordering and array dimensioning may have a large effect on code performance. Instead of wasting 80 percent or more of code time waiting on memory access, a simple rewriting of the loop structure can effect a five- or sixfold improvement in speed. Because the ideas presented are not difficult to apply, the extra effort is worth the gain. ■

About the author

Tony Warnock is a senior consultant at Cray Research. He has supported computational efforts at the Los Alamos National Laboratory for eight years. Before joining Cray Research he was a professor of mathematics for nine years at Sul Ross State University in Alpine, Texas. Warnock received a B.S. degree in chemistry at Sul Ross, and received M.S. degrees in chemistry and computer science and a Ph.D. degree in computer science from the University of Wisconsin, Madison.

CORPORATE REGISTER

New customers join Cray's worldwide market; first order received for CRAY Y-MP system

The **Toyota Motor Corporation** of Japan has ordered a CRAY X-MP/18 computer system, Cray Research announced in March. Toyota, which has been using a Japanese-made supercomputer for several years, is the third Japanese automaker to acquire a Cray system, and the first Japanese automaker to purchase a second supercomputer. According to Toyota officials, the Cray system selection is a result of overall performance, availability of application software, and maintenance servicing. It will be used for crash analysis, structural analysis, and computational fluid dynamics. The computer will be the 11th Cray system in the auto industry, and the 12th Cray system in Japan. The system is scheduled for installation in the second half of 1988.

In March Cray Research announced that **Sun Company**, the 11th largest U.S. petroleum company, had ordered a CRAY X-MP/48 computer system with SSD solid-state storage device. Sun is an energy resources company that explores for and produces crude oil and natural gas, refines crude oil, and markets refined petroleum products. The supercomputer will be used primarily for advanced seismic processing for onshore and offshore exploration. The system also will process data for reservoir modeling, log analysis, and other scientific computing applications. The system was installed in the first quarter of 1988 at the Sun Exploration and Production Technology Center in Dallas, Texas.

Also in March, Cray Research announced that **Leibniz Rechenzentrum** (LRZ) of the Bavarian Academy of Sciences in Munich, West Germany, ordered a CRAY X-MP/24 system, which will be upgraded to a four-processor Cray system in 1989. Leibniz Rechenzentrum is a computer center located in Munich, which serves the major universities of Munich and provides computer time for the Bavarian Academy of Sciences. The system will be used for basic research and engineering applications, and will serve the supercomputing needs of nine Bavarian universities. The system will be installed in the fourth quarter of 1988.

Los Alamos National Laboratory in Los Alamos, New Mexico, ordered two CRAY Y-MP computer systems to be installed in the second half of 1988, Cray Research announced in April. The systems will displace several Cray systems at the computer facility, which currently has eight Cray computer systems in operation. "We are pleased that Los Alamos, a pioneer in supercomputing with the first CRAY-1 system, will now be one of the first to take advantage of our newest generation of supercomputers," said John Rollwagen, chairman of Cray Research.

The Swiss government procurement agency has ordered Cray computer systems for two Swiss Federal Institutes of Technology located in Zurich and Lausanne. A two-processor CRAY X-MP system was installed at the **Eidgenössiche Technische Hochschule Zurich** in the second quarter of 1988. It will be used for a broad range of science and engineering projects. A

CRAY-2 computer system will be installed at the **Ecole Polytechnique Federale de Lausanne** in the third quarter of 1988. The supercomputer will be used for science and engineering research, including plasma physics and hydrodynamics simulation. Cray Research announced the order in April.

In June Cray Research announced that the **Korea Advanced Institute of Science and Technology** (KAIST) ordered a CRAY-2S/4-128 computer system. The system, which will be purchased, will be installed in Seoul at KAIST's System Engineering Research Institute during the third quarter of 1988, pending export license approval. This order is Cray Research's first supercomputer sale to an Asian country other than Japan. The system will be used to support scientific research for universities and research institutes throughout the Republic of Korea.

Westinghouse Electric Corporation, acting as prime contractor for the United States Department of Energy, ordered a CRAY X-MP/48 computer system with SSD solid-state storage device. The purchased system will be installed in the third quarter of 1988 at the **Bettis Atomic Power Laboratory** (BAPL) in West Mifflin, Pennsylvania. DOE/Bettis is a new customer for Cray Research. Cray Research announced the order in June.

Cray Research forms Industry, Science & Technology Department

Cray Research has consolidated its applications and industry management functions to form the Industry, Science & Technology Department.

As part of its mission, the group will continue to foster development and optimization of application software by third-party vendors for use on Cray systems. "Our mission also includes building closer cooperation between universities and industry," said Carlos Marino, director of the department.

The Industry, Science & Technology Department also is responsible for providing Cray Research with guidance on industry's future requirements in supercomputing, and will develop and implement industry programs to address those needs.

Cray Research moves Colorado Springs sales office

The Central Region of Cray Research has moved its Colorado Springs sales office, formerly located at 5345 Centennial Boulevard in Colorado Springs. The new office is located at 102 S. Tejon St., Suite 1100, Colorado Springs, CO, 80903; telephone: (719) 578-3339. David Guthrie, who formerly headed Cray Research's St. Louis, Missouri office, will direct the Colorado Springs office.

Cray Research's Central Region has maintained an office in Colorado Springs since May of 1986 to service customers in the area. The Central Region, headquartered in Boulder, is responsible for sales and service of the company's products in a 15-state area, including Colorado. This sales office relocation is unrelated to the company's development of a manufacturing facility in Colorado Springs.

Cray Research selects Colorado site for CRAY-3 facility

Cray Research has chosen Colorado Springs, Colorado, as the location for a new manufacturing facility. The company has purchased an existing manufacturing facility, which will be adapted for volume production of the CRAY-3 computer system. The new facility would complement Cray Research's manufacturing and development facilities in Chippewa Falls, Wisconsin.

"Colorado Springs was chosen for its size, location, climate, and available work force," said John Rollwagen, chairman of Cray Research. "A large number of high technology companies have facilities in and around the Colorado Springs area and there are several major universities and research facilities that can be good resources for potential employees."

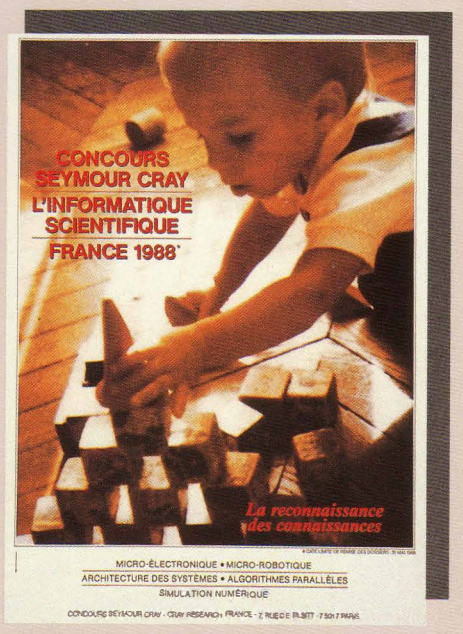
At the company's annual meeting in May, Seymour Cray said that the small size of the computer's components will require automated manufacturing techniques. Colorado Springs production of the CRAY-3 system is scheduled to begin in 1990.

Contest honors research excellence

Cray Research France S.A. has announced the 1988 Seymour Cray Research Proposal Competition, a contest to acknowledge excellence in research by French scientists. The 1988 competition includes five disciplines: microelectronics, computer architecture, numerical simulation, parallel algorithms, and microrobotics.

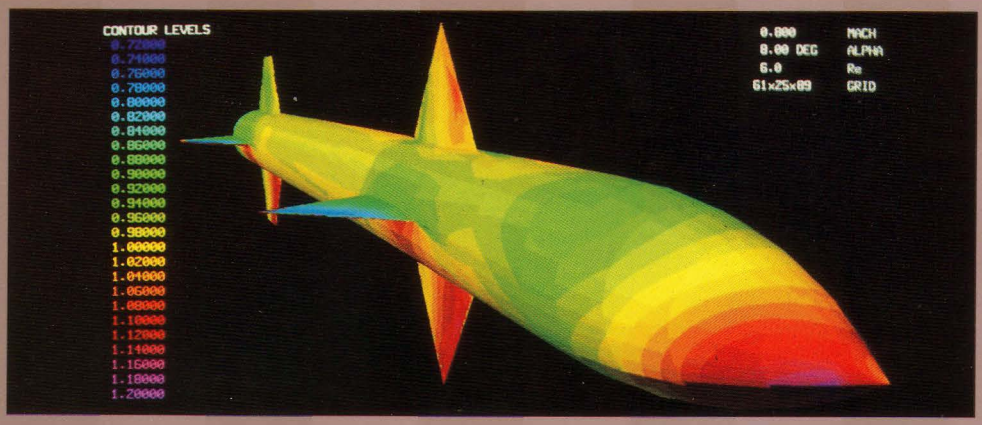
Cray Research France S.A. also announced winners of the 1987 Seymour Cray Award, which covered three disciplines: microelectronics, computer architecture, and numerical simulation. Although no first prize was awarded, second prize was awarded to Bernard Hennion, an engineer at the French National Telecommunications Research Center, for his work, "A Contribution to an Electrical Simulation of VLSI Circuits." Two third prizes were awarded. One was awarded to Jean-François Colonna, engineer and software expert at the Ecole Polytechnique, for his work, "Picture: a Working Tool." The other went to Marianna Braza, Toulouse engineer and research expert at the National Center for Scientific Research, and Hieu Ha Minh, professor at the National Polytechnique Institute in Toulouse and member of the American Institute of Aeronautics and Astronautics, for their work, "Numerical Methods in Laminar and Turbulent Flow."

A special prize for original applications in industry was awarded to Paul de Casteljau, an engineer at Automobiles Citroen for his work, "Contribution to Research in Algorithms for Industrial Computers." An encouragement prize went to research engineer Dominique d'Humierès and research expert André Clouqueur, both from the Ecole Nationale Supérieure, for their work, "RAP 1, a Cellular Automaton Machine for Fluid Dynamics."



An independent jury develops the competition rules, reviews submissions, and determines the award recipients. Cray Research France limits its participation in the contest to financing and logistical support. Winners of the 1988 contest will be chosen in November.

APPLICATIONS UPDATE



Pressure distribution over a generic missile as illustrated by PLOT3D. Data generated by BAUD 57, a code developed by Cray Research.

PLOT3D presents CFD data sets

PLOT3D is an interactive graphics program that runs on CRAY-2 computer systems under the UNICOS operating system in conjunction with a Silicon Graphics IRIS 3000 series workstation. The program, which is written in Fortran, displays computational fluid dynamics grids and flow solutions. It allows analysis of CFD results for solution quality and flow characteristics.

Two files are read into PLOT3D: the grid file, which consists of a three-dimensional coordinate at each grid point, and the Q file, which includes density, momentum, and energy at each point. Flow quantities computed from these variables include pressure, Mach number, enthalpy, temperature, velocity, and vorticity.

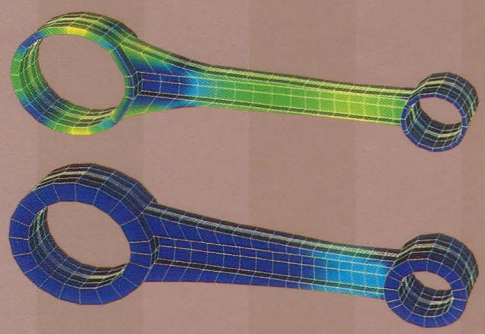
Contour, carpet, vector, and grid plots can be made in any subset of the computational grid. The code handles multiple grid solutions resulting from zonal, overlapped grid, or grid refinement methods. Plots of shock locations and particle traces also can be generated.

For more information about using PLOT3D with Cray computer systems, contact Pieter G. Buning, RFT: 258-2, NASA Ames Research Center, Moffett Field, CA 94035; telephone: (415) 694-5194.

SUPER-DOT solves optimization problems

SUPER-DOT is an enhanced version of the modified method of feasible directions algorithm from the ADS optimization program, and runs on CRAY X-MP, CRAY X-MP EA, CRAY Y-MP, and CRAY-2 computer systems under the UNICOS and COS operating systems. Available from Engineering Design Optimization, Inc., SUPER-DOT is compatible with the COPES control program, which makes it particularly easy to couple user-supplied analysis programs with optimization. The code minimizes a nonlinear function subject to a set of nonlinear inequality and equality constraints. The program is optimized to take advantage of the Cray operating systems.

SUPER-DOT is a subroutine that is called from a user-supplied calling program. A user provides an initial array of design variables that are changed in search of the minimum or maximum objective function that satisfies the constraints. When function values or gradient information is needed, control is returned to the calling program. The user then evaluates the needed information, calling SUPER-DOT again in sequence until the optimum is found. If the user cannot provide gradient information, this is calculated internally by



Shape optimal design for General Motors connecting rod performed by SUPER-DOT.

finite difference. This is the most common case, in which a user wishes to use optimization without major modifications to the analysis program.

The initial release of SUPER-DOT is operational, and in the near future a version will be released that will make full use of parallel processing during finite difference gradient calculations, which is the most computationally intensive part of the optimization process. Also, a new one-dimensional search routine will be added that makes optimum use of parallel processing. The efficiency gains obtained over serial processing are problem-dependent, but efficiency is expected to improve by a factor of 4 to 7 for problems of significant size and complexity.

SUPER-DOT is designed to be a robust numerical optimizer for problems of approximately 30 design variables, though problems with several hundred design variables have been solved. Structural optimization problems that use SUPER-DOT as a subroutine in conjunction with a structural analysis code and user-supplied gradients often involve well over 100 design variables and several thousand constraints.

For more information about using SUPER-DOT with Cray computer systems, contact Garret Vanderplaats, Engineering Design Optimization, Inc., 1275 Camino Rio Verde, Santa Barbara, CA 93111; telephone: (805) 967-0058.

FIDAP models fluid flow

The Fluid Dynamics Analysis Package (FIDAP) numerically simulates fluid flow and heat transfer, based on the finite element method. Revision 4.0 now is available to run on CRAY-2, CRAY X-MP, CRAY X-MP EA, and CRAY Y-MP systems under the UNICOS, COS, and CTSS operating systems.

FIDAP can be applied to a wide variety of incompressible fluid flow problems. Some application areas include electronics, environmental studies, and biomedical science, as well as the automotive, nuclear, chemical, metal-forming, and food and beverage industries. The package can model two-dimensional, axisymmetric simulations, as well as three-dimensional, steady-state, or transient simulations of incompressible fluid flows in complex, arbitrary geometries.

An advantage of the finite element method is its inherent flexibility in treating arbitrarily complex flow domains and boundary conditions. Designers can model unstructured grids to study areas of interest in greater detail without the need for excessive grid points throughout the entire flow domain. The finite element method also allows natural and correct imposition of boundary conditions on curved boundaries. The method's elegant mathematical formulation can derive comprehensive error estimates and determine accurate simulations within user-prescribed tolerances.

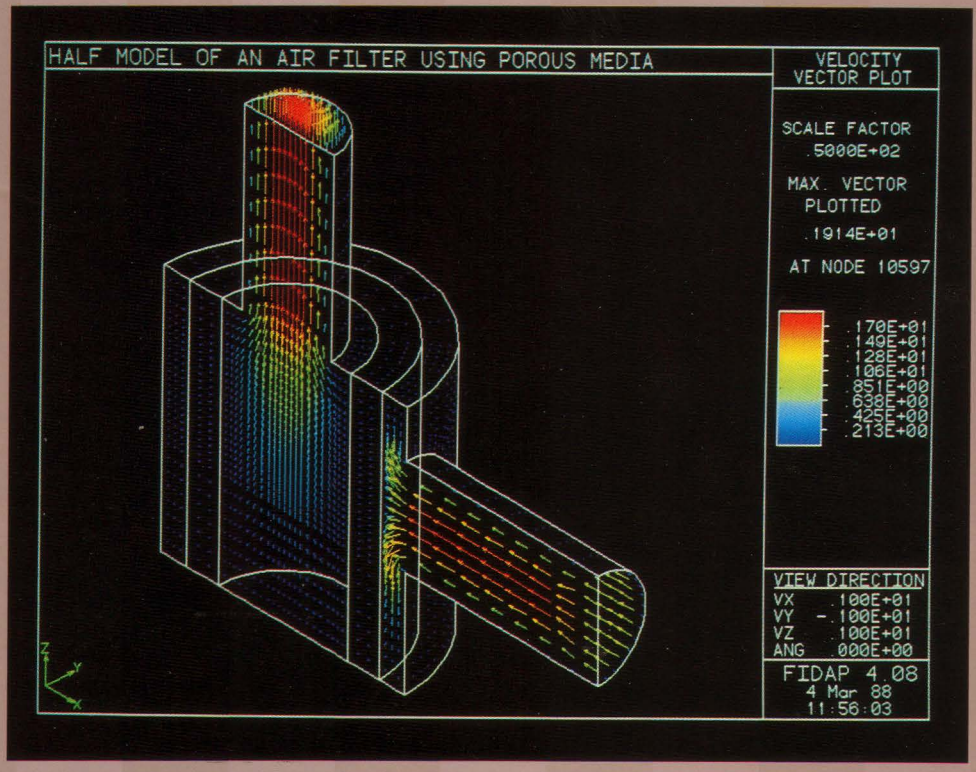
The accuracy of the finite element method in flow visualization and thermal performance prediction makes FIDAP a useful complement to experimental testing. FIDAP software can be used to

- ☐ Investigate implications of design changes within relatively short time spans
- ☐ Provide detailed information on the velocity, temperature, and pressure fields throughout the flow domain
- ☐ Eliminate disturbances to the flow and thermal fields due to measurement devices
- ☐ Simulate severe operating conditions that are impossible to reproduce experimentally
- ☐ Reduce time and materials costs
- ☐ Reduce the range of conditions over which physical testing is required

FIDAP operates as a collection of interdependent modules. The package includes the FIPREP preprocessor input generator and the FIPOST graphics postprocessor. Some key program features of the FIDAP system include

- ☐ Boundary conditions that include specified velocities, temperatures, turbulent kinetic energy, and turbulent dissipation; applied surface stresses, heat fluxes, turbulent kinetic energy, and dissipation fluxes; conductive and radiative heat transfer; a slip velocity boundary condition; a free or moving fluid interface with or without surface tension effects; and a turbulent wall-function boundary condition
- ☐ Volumetric forces that can be applied to the fluid include body forces, heat sources or sinks, Coriolis force, and centrifugal force
- ☐ Constant or temperature-dependent models for fluid properties including viscosity, specific heat, conductivity, volume expansion coefficient, emissivity, surface tension, slip coefficient, permeability, and convective heat-transfer coefficient
- ☐ Non-Newtonian models, such as power law and Bingham-yield stress-type models
- ☐ A wide range of time integration and solution techniques
- ☐ An element library that includes linear and quadratic isoparametric quadratic bricks, tetrahedrons, and triangular prisms for three-dimensional analysis

For more information on using FIDAP software on Cray computer systems, contact Michael Engelman, Fluid Dynamics International, 1600 Orrington Avenue, Suite 400, Evanston, Illinois, 60201; telephone: (312) 491-0200.



Simulation of the flow through a truck air filter using FIDAP.

USER NEWS

Supercomputer investigates supercontinent breakup

At about the time dinosaurs began to roam the Earth, most of Earth's land area was contained in one supercontinent, according to today's geologists. However, when it comes to explaining how this supercontinent split into six separate continents, opinions differ.

John Baumgardner, a geophysicist at the Los Alamos National Laboratory (LANL) in New Mexico, has created a model on a CRAY X-MP system to explain the formation and breakup of the early supercontinent. "Our project basically involves building a solid-Earth simulator capable of expressing the three-dimensional dynamic behavior of Earth's mantle and lithosphere," he says.

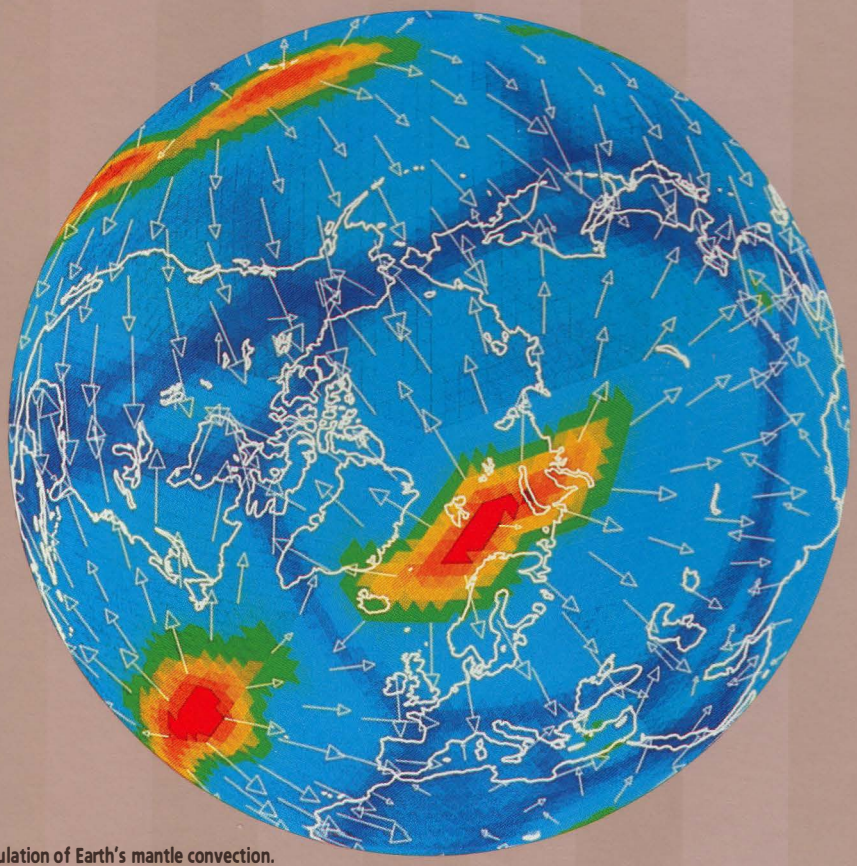
The mantle is the silicate portion of Earth between its core and crust, a volume that is 1800 miles thick and composed of hot, slow-moving rock. The lithosphere is the cold outer skin of Earth, which is relatively rigid compared to Earth's deeper interior. The lithosphere is divided into approximately 20 plates. Geologists believe that the motions of the continents are related to their connection to these mobile lithospheric plates.

"It was somewhat of a sensation that anyone could do such three-dimensional mantle convection calculations."

About 25 years ago, the widespread acceptance of the theory of plate tectonics caused a major revolution in the earth sciences. Says Baumgardner, "Even though that happened over two decades ago, we still do not have a clear understanding of the driving mechanism for the plates. The forces generally are ascribed to convection in Earth's mantle, but beyond that we have been able to say surprisingly little."

Baumgardner's objective is to develop a finite element computer model to simulate the dynamics of Earth and to explore the connections between the mantle's behavior and the geological expressions at Earth's surface. "We are seeking to model Earth as a whole system to gain insight into the fundamental physics responsible for seafloor spreading, mountain building, volcanoes, and earthquakes," he says.

When Baumgardner first developed this three-dimensional code in 1982 as part of his doctoral thesis at the University of California, professional literature said that computers capable of this research were at least a decade away. "The combination of the Cray system and a special multigrid solution technique enabled us to make large advances in performing these whole-mantle simulations," he says. "At professional geophysics meetings, it was



Simulation of Earth's mantle convection.

somewhat of a sensation that anyone could do such three-dimensional mantle convection calculations."

And modeling mantle convection for the entire Earth is no small task. He adds, "We use the finite element model to divide Earth's mantle into 338,000 cells, or chunks of material." To do this he uses a special mesh based on a regular icosahedron, a solid polyhedron with 20 equal triangular faces. When the icosahedron is projected onto a sphere, it divides the spherical surface into 20 equal spherical triangles. Baumgardner connects the midpoints of the sides of these spherical triangles with great-circle arcs to subdivide each of the 20 triangles into four subtriangles. This process is repeated until the desired resolution is reached. "A major benefit of this scheme is that it provides an almost uniform triangular mesh on the sphere," he says. The Cray system simulates eons of slow mantle movement by taking the data from the 338,000 cells and calculating the mutual interactions through 10,000 or more time steps. Each time step requires 400-500 million arithmetic operations.

"There haven't been many machines around that could do such calculations," says Baumgardner. When the project began, a Cray system was the only computer that could tackle such a task in a reasonable amount of time. The program is well vectorized to perform efficiently on a CRAY X-MP system. "The code runs at about 100 MFLOPS per processor on the CRAY X-MP system, and typical calculations require on the order of two to five hours. We solve a system of one million simultaneous equations in about two CPU seconds on the CRAY X-MP system," he says.

So far, the model has provided some significant insights into the behavior of Earth's mantle. "We observe features that correlate with actual observations," Baumgardner says. "We typically see upwelling from the base of the mantle in the form of narrow plumes." In most of the calculations, the downwelling occurs in the form of narrow sheets that are actually portions of the cold upper boundary layer that peel off and sink. "This corresponds to the lithospheric slabs that we know sink into Earth. As far as style of convective flow in the mantle, we commonly see six to ten convection cells in the simulations," he says.

Next, Baumgardner hopes to use the added power of two new CRAY Y-MP systems soon to be installed at LANL to add more physical features to the model such as variable viscosity to represent the lithospheric plates more realistically, as well as regions of buoyant continental crust. "By incorporating these new features, we are hoping to identify some of the physics responsible for supercontinent formation and breakup. We are scrambling to make efficient use of the eight-processor machines."

But Baumgardner doesn't mind the scramble. "I find it exciting to do the large-scale simulations that the Cray system makes possible. These machines have opened a door to numerical exploration only dreamed of just a few years ago."

Cray system illustrates brain tissue

Two neurobiologists at the University of California, San Diego, (UCSD) are using a CRAY X-MP system and high-voltage electron microscope data to visualize brain tissue and explore the effects of Alzheimer's disease, the fourth largest cause of death in America. About 2.5 million American adults suffer from this progressive, degenerative disease that impairs memory, thinking, and behavior.

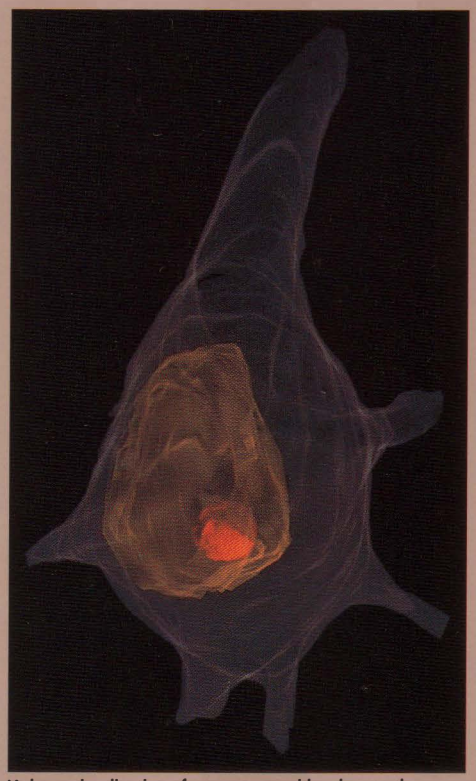
"We are trying to explain how nerve cells live, interact, and sustain their electrical and chemical interactions," says Mark Ellisman, professor of neurosciences and director of the laboratory for neurocytology at UCSD. "Our mission is to understand where the earliest changes are occurring in neurons that appear affected or degenerated from Alzheimer's disease," he says.

He and colleague Steve Young, a specialist in the department of psychiatry at UCSD, are studying the brain by combining techniques of dissection, analysis, reconstruction, and computer simulation. Ellisman and Young have analyzed brain tissue from patients in various stages of Alzheimer's disease. These samples were collected by Robert Terry of UCSD and Suzanne Mirra of Emory University School of Medicine and the Veterans Administration Medical Center in Atlanta, Georgia.

"First, we create serial section reconstructions by cutting about 100 sections similar to how you slice a loaf of bread," explains Ellisman. "Then we examine those sections under a high-voltage electron microscope. This allows us to look at thicker sections than you normally could look at, and allows us to go through larger volumes of material with fewer slices."

Next, the team enters data into a microcomputer by tracing onto a digitizing tablet photographic images created by the electron microscope. "We trace profiles — boundaries of specific structures — and we name them," says Ellisman. "We are examining changes in the relationship between different cell structures that result from the disease. Boundaries of the cell, the nucleus, and other structures are delimited in order to create a skeletonized reconstruction."

Then data from the microcomputer are offloaded to the CRAY X-MP system at



Volume visualization of a neuron provides clues to the causes of Alzheimer's disease. This neuron's surface membrane was rendered using polygons to connect data from the brain slices. The distribution of surfaces of interior structures is also being studied. Photo courtesy of Mark Ellisman and Steve Young of UCSD, and Brandyn Webb of SDSC.

the San Diego Supercomputer Center (SDSC). Ellisman and Young are taking advantage of supercomputer expertise at SDSC to generate more sophisticated renderings and representations of cell structure.

The Cray system enables the researchers to simulate surfaces across gaps between contour slices. "But we're not just painting a skin on the neuron. Many more internal surfaces must be rendered, and many layers of transparency must be represented with different colors," Ellisman says. "By using this kind of computer graphic technology, especially the powerful technology fostered by being linked to a Cray computer, we can look at larger volumes of data and zoom through and review structural relationships over large areas," he explains. "This couldn't be done any other way."

By classifying brain structures and the progression of changes that occur as cells die, Ellisman and Young hope to better understand the causes of brain degeneration. However, Ellisman points out that this project is only in its initial stages. "We are just exploring what can and can't be done and trying to streamline our efforts so we can start working more rapidly with more neurons."

By the end of the year the team hopes to reconstruct a whole panel of neurons. Says Ellisman, "We hope that from the renderings we are capable of doing on the Cray system, we will be able to add detail, make further observations, and then simulate degeneration."

Cray system looks back

Supercomputers and historians. Sound like an unusual match? Not so, according to Vernon Burton of the University of Illinois. Burton, a professor of history, is using a CRAY X-MP system to run an automated record-linkage program, the first of its kind to run on a supercomputer.

"Anybody interested in history has to be interested in numbers," he says. Burton is analyzing United States census data collected since 1790 to develop a database of historical information. The research project is called nominal record linkage, which is a process of tracing and linking an entire population's progress through time. "For a long time, scholars have used this census data for different purposes, but they haven't been able to deal with the data in a systematic way until computers became available in the sixties and seventies," Burton says.

In the early 1970s, he began examining historical data with a mainframe computer. He used historical databases and handwritten census data from the Civil War era to study black and white families. Burton's findings are documented in his book, *In My Father's House are Many Mansions*, which was nominated for the Pulitzer Prize in 1985. Although he found the mainframe to be a useful tool for historical research, he was stifled by the lack of power. Because the data-entry procedure monopolized the system, he was asked to merge his databases after midnight. "Another problem was data linkage," he adds. "I wanted to link people from the 1790 census through the 1910 census, and link people from the farm, slave, and manufacturing censuses to the population census."

Then Burton heard about the possibility of working simultaneously with different databases using special software on a Cray computer system. Using the CRAY X-MP system at the National Center for Supercomputing Applications in Urbana-Champaign, Illinois, Burton can perform nominal record linkage, store census data from large regions, and graphically display demographic patterns. "The speed is incredible," he says. "The size of the problem was so large that it really could not be done on the other computers that were available." He has joined forces with a linguist, Atsushi Fudaka, who has developed portable algorithms for nominal record linkage. The next hurdle to jump is data entry. As of now, data are entered manually into the computer. "In the future, optical scanners will help us enter data straight from microfilm," says Burton.

The added power of a Cray system allows Burton to think big. His goal is to develop a national historical database. By analyzing census data from most everyone who ever has lived in the United States, Burton hopes to study the nature of American society, democracy, social mobility, and land ownership.

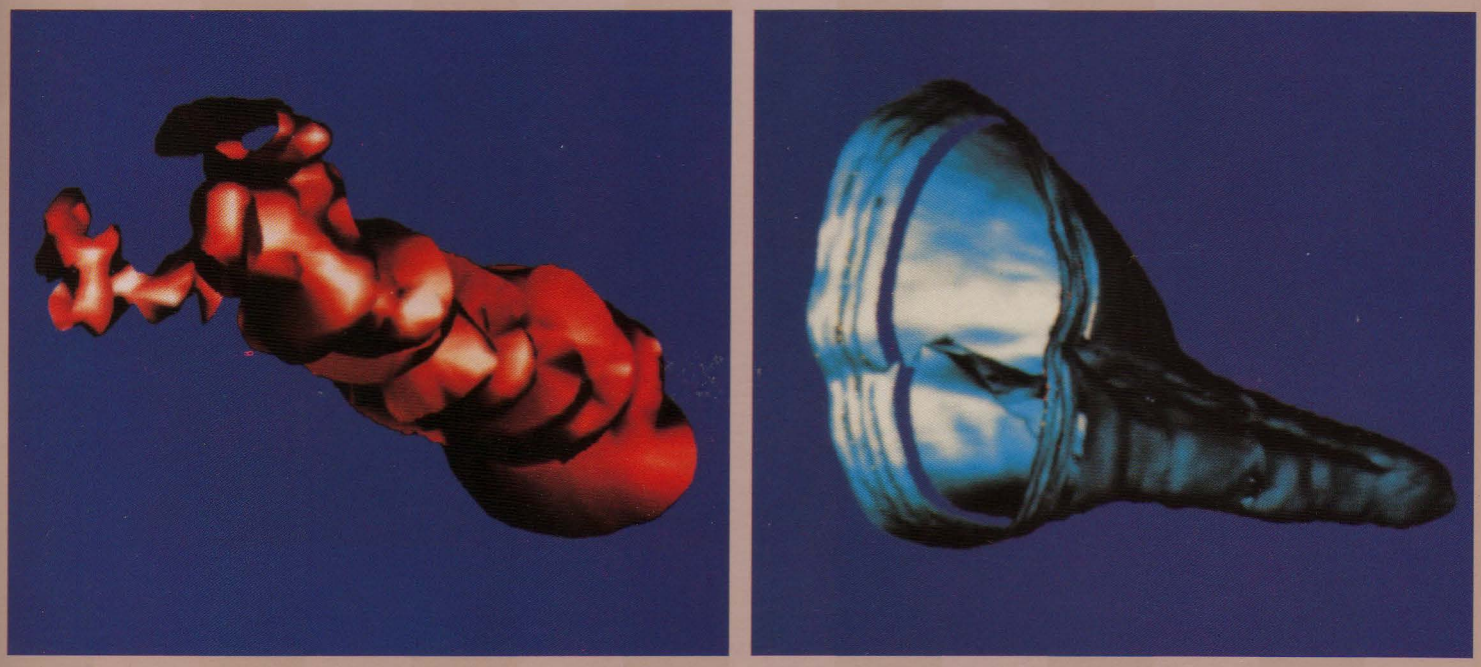
Graduate history students have applied to work on nominal record linkage research with Burton. One student, Terence Finnegan, is studying lynching patterns throughout the nineteenth century South. "Previously, historians have studied lynching in great detail, but one of the problems with these studies is that historians haven't been able to apply their theories beyond state and regional boundaries. I'm attempting to do that," says Finnegan. "That's why the Cray system is important to me — because I'm dealing with the entire South, 11 states, and approximately 1000 counties."

Both Finnegan and Burton agree that supercomputers could change the way historical research is done. "I think it's exciting to use the Cray system and supercomputers in general for humanistic studies," says Finnegan. "As historians begin to see the utility of looking at large amounts of information, they will be able to extend their conclusions beyond political phenomena to broad movements of people in regions or the nation as a whole." Burton adds, "The speed of a supercomputer fosters creativity because it allows historians to explore questions that were previously too complex to answer."

Fishing for answers about dolphin sonar

The sounds produced by dolphins have long baffled marine biologists. Although scientists know that dolphins use some form of sonar to locate food and navigate the seas, they do not know how dolphins produce and project their narrow sound beams.

Ted Cranford, a marine biology doctoral student at the University of California at Santa Cruz, is using a Cray computer system and medical imaging techniques to find out how dolphins produce and transmit sounds. Cranford is fascinated by the acute sensing ability of dolphins. "From the length of a football field away, a dolphin can find a tangerine-sized sphere," he says. "A blindfolded dolphin can discriminate between different types of metal, even if they are the same shape or size."



Sections of dolphin anatomy involved in sound production, as rendered by MOVIE.BYU software, (left) right posterior-lateral view of melon core, (right) right posterior-lateral view of skin on the head of a spinner dolphin.

Although dolphins create various sounds, Cranford is most interested in the short, clicking sounds that emanate from their foreheads. Originally, some researchers have resorted to dissection in their attempts to understand the relationships between acoustic fat, air spaces, and valves in a dolphin's forehead. "They basically took a dead animal, cut it open, and described what they saw," says Cranford. "Well, if you think about a sound-production system that creates a very narrow beam nine or ten degrees wide, the geometric structure of the anatomy must be very important. When a dolphin is cut open, these relationships are disrupted."

When Cranford began to study with marine biologist Ken Norris, Norris asked him to find out about the complex geometric structure of the forehead. "I was able to go to the University of California, San Francisco, Medical School where doctors Goldberg and Cann let me use an x-ray computed tomography (CT) scanner to get images of slices of the head." Cranford scanned slices as thin as 1.5 millimeters in some areas. "Over the course of a normal dolphin's head, I had 130-140 slices," he says. "I found some structures that had not been described before. I've since found similar structures in every dolphin I have examined.

"Next, I took the CT data to the University of California, San Diego Medical School's Quantitative Morphology Lab, and entered the CT data into a computer by using a digitizing pad. This process reduced the data to a series of topographic lines," he explains. Then, at the San Diego Supercomputer Center (SDSC), Cranford used a CRAY X-MP system running the MOVIE.BYU graphics program to create pictures from the wire-frame images of topographic lines.

"I learned about structures, geometric relationships, and variations in tissue density that had not been revealed before."

"The process involved taking each pair of 'wires' from these images and having the Cray system draw polygons between them using the mosaic part of MOVIE.BYU. The movie part of MOVIE.BYU rendered perspectives of the objects and added colors. This was accomplished with SDSC consultants and graphics guru Steve Lamont," Cranford explains, adding that the size of this problem required the use of a Cray system. He was able to depict side, front, and back views of three parts of a dolphin's head: the skin, the melon (a fatty organ in the forehead), and a "rabbit ear" structure that may help generate sound in the core of a dolphin's melon.

"Through my research I learned about structures, geometric relationships, and variations in tissue density that had not been revealed before," Cranford says. He found two projections in the dolphin's forehead above the nasal passages. These projections, called dorsal bursae, resemble tiny rabbit ears. Cranford suspects that dolphins create sounds with the dorsal bursae by using muscles to push air through the small air passage between the two projections. As air is forced between the dorsal bursae, the tough ridged tissue below may act as a trumpet player's embouchure, regulating sound repetition rate. The fatty tissue in the dolphin's melon helps focus the sound beam. "Focusing probably occurs in stages as described by Norris," says Cranford. "Air spaces and other tissue planes also probably contribute to the focusing phenomena."

Cranford has won two best-paper awards for describing his work, has toured Europe lecturing about this discovery, and has stirred interest among marine biologists around the world. Now he is attempting to explain structural similarities and sound production between dolphins, sperm whales, and other toothed whales. He suspects all toothed whales create sounds with similar dorsal bursae structures and fatty tissues in their foreheads.

He hopes to obtain a grant of supercomputer time to create a map of an entire dolphin head upon which to model sound propagation. Although he has attempted this feat with a ray-tracing program, he expects the wave equation to provide more accurate results. Cranford plans to use programs that have been developed by Moshe Reshef, a consultant to Cray Research from Tel Aviv University in Israel. With these equations and information collected

from 20 dolphins of varied types, shapes, and sizes, Cranford also hopes to understand more about dolphin hearing by looking at how sound is propagated into a dolphin's head.

Cranford adds, "Because medical images contain so much information and the wave equation is computationally intensive, it requires a computer with large memory and enough speed to calculate where the sounds will go. It looks as though using a supercomputer like a Cray system is the only way to attempt this model."

CRAY-2 system performs star search

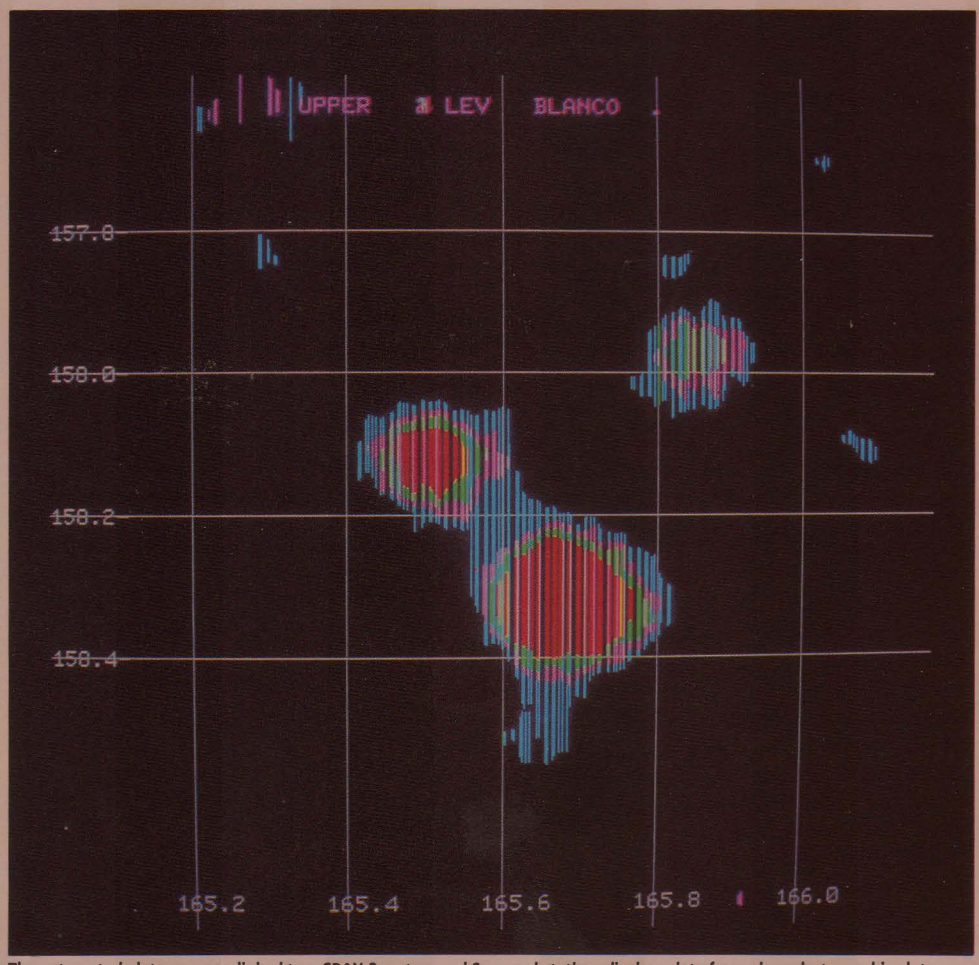
The sky is no limit for a team of astronomers at the University of Minnesota. By linking the university's CRAY-2 computer system and a high-speed measuring machine used to scan photographic plates, they are analyzing the sky statistically to establish a library of information about billions of stars and millions of galaxies.

"An astronomical data set of this magnitude has never existed," says Roberta Humphreys, director of the project, explaining that up to now, astronomers have been working with scant data sets — catalogs from the 1920s and 1930s that were compiled by astronomers who manually counted the stars from photographs. According to Humphreys, the new catalog will have the potential to revolutionize many areas of astrophysics. "A catalog of this completeness will provide vital information for the understanding of the structure of the entire universe," she says.

The university's automated plate scanner was developed for U of M professor Willem Luyten in the late 1960s. Luyten's creation originally was designed to measure star motions, but has been redesigned to scrutinize pairs of glass photographic plates of the universe, recording position, brightness, size, and shapes of stars and galaxies. The measuring machine can scan a pair of plates simultaneously. Plates in two colors, blue and red, are scanned.

The scanner's lasers skim these plates to identify and compare data. The university has 935 pairs of plates that were taken in the early 1950s at Mount Palomar Observatory. The set is one of few in existence, and covers about 77 percent of the sky. The astronomers hope to catalog about one-half of the sky, which will provide a solid statistical base for their studies.

"The CRAY-2 system is perfect for this," says research associate Robert Pennington, explaining that this type of research was never before possible. "On an average day we are generating about 600 Mbytes of raw data," he says. This correlates to 600,000 images that must be classified as stellar or nonstellar, then parameterized as stars or galaxies — an astronomical data problem. Says Humphreys, "With this computer, the possibility of scanning and digitizing two-thirds of the sky became a reality for us. We realized it could be done — and in a reasonable amount of time."



The automated plate scanner, linked to a CRAY-2 system and Sun workstation, displays data from glass photographic plates of the universe. The graphic output will help researchers determine the number of stars in the universe and their spatial and density distributions.

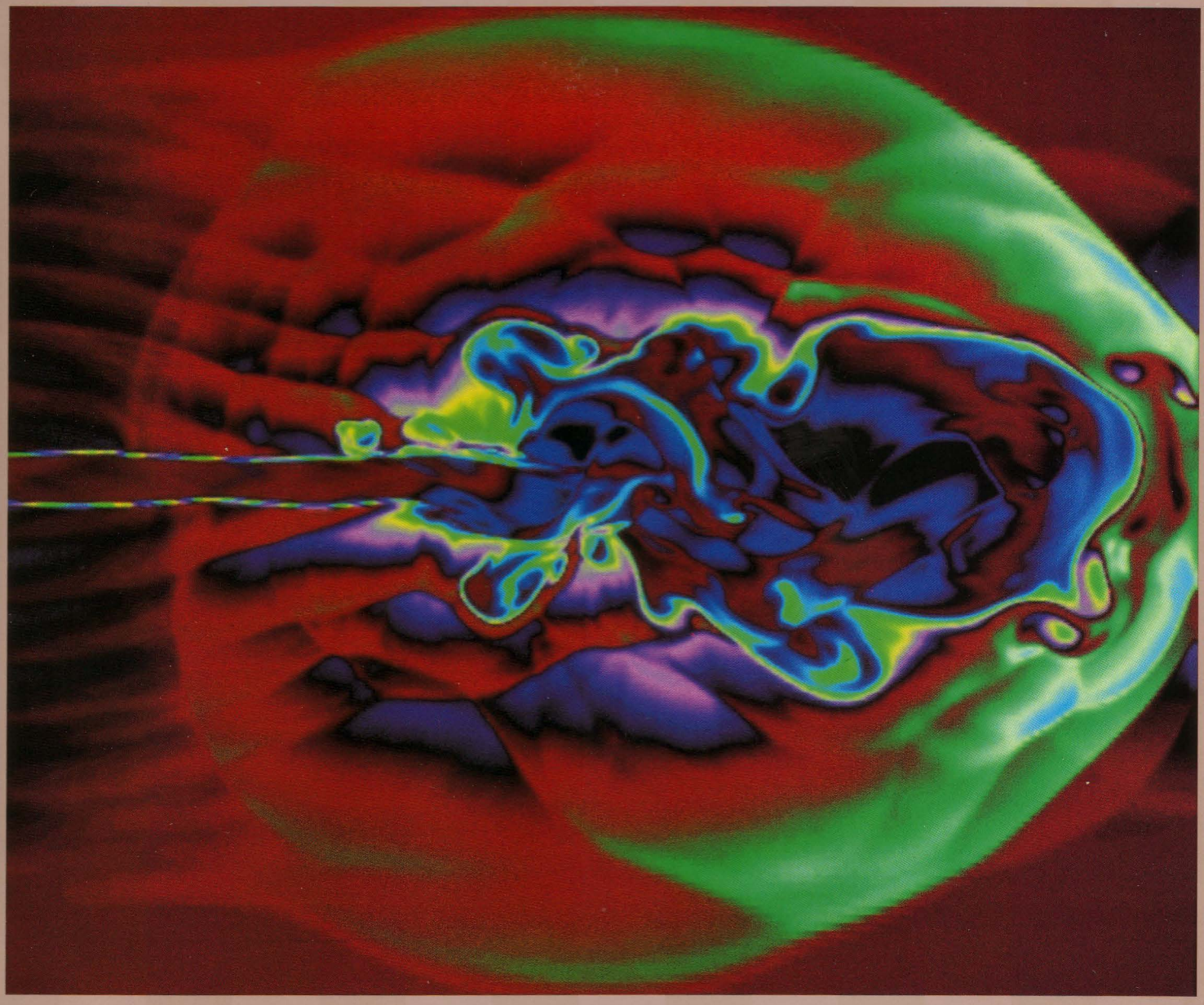
"With this computer, the possiblility of scanning and digitizing two-thirds of the sky became a reality for us."

To leap the hurdle of data, Pennington and computer programmer Bill Zumach have developed two levels of software. The first sorts, then produces images of the raw data. The second tier of software discriminates between stars and galaxies. The code is being written in Fortran and should vectorize well, according to Pennington. A Sun Workstation displays data from each pair of plates digitized by the scanner.

With the scanner linked to the Cray system and Sun Workstation, the astronomers can determine the number of stars and their spatial and density distributions. "These counts of stars, primarily as a function of their brightness and location in the galaxy, can then be used in combination with a mathematical model of the galaxy. This information can be used to determine the structure of the Milky Way and the distribution of galaxies in the universe," says Humphreys.

When the catalog is completed in about two years, it will contain data for more than a billion stars in our galaxy, and about 4 million galaxies. Information will be stored on computer tapes and made accessible through NASA's Goddard Space Flight Center, a clearinghouse of astronomical data, so that other researchers can reach for the stars.

GALLERY



"Instabilities in Supersonic Flow" shows the development of the fundamental kink mode of instability in a supersonic gas jet. The image is from an animated sequence of 400 images. The animation was computed in about 15 hours on the CRAY X-MP/48 computer system at the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign. The principal investigators were David Clarke, Donna Cox, Phil Hardee, and Michael Norman. Gallery submissions may be sent to CRAY CHANNELS at the address inside the front cover.

Worldwide business centers

Corporate Headquarters
608 Second Avenue South
Minneapolis, MN 55402
(612) 333-5889, telex 6879144

Government Marketing/
Government Relations Office
National Place
1331 Pennsylvania Avenue, Suite 1331 N
Washington, DC 20004
(202) 638-6000

Domestic

Central Region Office
3375 Mitchell Lane
Boulder, CO 80301
(303) 447-8855
Albuquerque
Chicago
Colorado Springs
Detroit
Minneapolis/St. Paul
St. Louis

Eastern Region Office
4041 Powder Mill Road, Suite 600
Calverton, MD 20705
(301) 595-5100
Atlanta
Bernardsville
Cincinnati
Darien
Huntsville
Nashua
New York
Pittsburgh
Tampa
Washington, D.C.

Southern/Petroleum Region Office
5847 San Felipe, Suite 3000
Houston, TX 77057
(713) 975-8998
Dallas
Tulsa

Western Region Office
3130 Crow Canyon Place
Suite 400
San Ramon, CA 94583
(415) 866-0600
Huntington Beach
Los Angeles
Salt Lake City
Scottsdale
Seattle
Sunnyvale

International

Cray Asia/Pacific Inc.
Suite 1003
Sutherland House
3 Chater Road Central
Hong Kong
(5)-214669, telex 89765

Cray Canada Inc.
4211 Yonge Street, Suite 210
Willowdale, Ontario
Canada M2P 2A9
(416) 229-2729

Cray Research (Australia) Pty. Ltd.
Victorian Technology Centre
283 Normandy Road
Port Melbourne, Australia 3207
(03)-647-9833, telex 32111

Cray Research France S.A.
7 Rue de Tilsitt
75017 Paris, France
(01)-47-66-01-55, telex 660568

Cray Research GmbH
Perhamerstrasse 31
8000 Munich 21, West Germany
(089)-56014-0, telex 5213211

Cray Research Japan, Limited
Ichibancho Eight One Building
6-4 Ichibancho
Chiyoda-Ku, Tokyo 102, Japan
(03)-239-0711, telex 28893

Cray Research S.A.E.
Pasco de la Castellana
28046 Madrid, Spain
(01) 459-0150, telex 48998

Cray Research S.R.L.
Via Vivaio, 15
20122 Milan, Italy
(02)-70-98-27, telex 315403

Cray Research (Suisse) S.A.
Route de Pre-Bois 20
P.O. Box 534
1215 Geneva-15, Switzerland
(022) 980-091, fax 4122980516

Cray Research (UK) Ltd.
Cray House
London Road, Bracknell
Berkshire RG12 2SY
United Kingdom
(0344)-485971, telex 848841

CRAY RESEARCH, INC.