CRAY CHANNELS

SUMMER 1991 · A CRAY RESEARCH, INC., PUBLICATION

Announcing the CRAY Y-MP4E computer system

CRAY CHANNELS

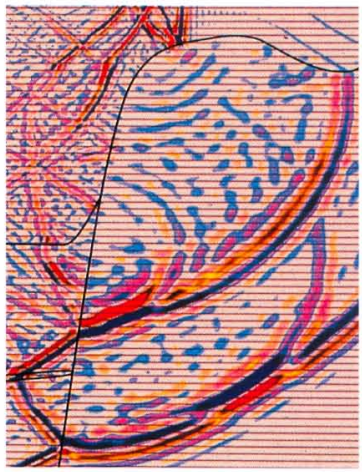
In this issue

Because petroleum is a finite resource that underlies much of the global economy, it must be managed with care. Such management requires extensive technological resources to optimize both exploration and production. Cray Research supercomputers have played a central role in managing this resource during the past decade. Cray Research systems enable exploration geophysicists to process vast amounts of seismic data to locate and map reservoirs. Reservoir engineers use Cray Research systems to model fluid behavior to help design efficient recovery strategies.

This issue of CRAY CHANNELS covers a range of seismic and reservoir modeling applications. This issue also introduces the CRAY Y-MP4E computer system, Cray Research's newest supercomputer, and includes a report from the Toshiba Corporation on device simulation for integrated circuit design. Our regular departments cover CVT, a new visualization tool; petroleum application software packages; and an article from *Chemical and Engineering News* on supercomputer-aided research on the HIV virus.

Since ARCO Oil & Gas Company became Cray Research's first petroleum industry customer in 1981, the industry has continued to demand Cray Research computer systems for large-scale scientific processing. Petroleum companies now account for more than 10 percent of Cray Research's installed base of supercomputers worldwide. As the capabilities of Cray Research systems have grown, geophysicists and petroleum engineers have expanded their use of the systems to solve larger, more complex seismic and reservoir modeling problems. These expanded capabilites have helped ensure that the resources of petroleum companies are used most efficiently and that petroleum remains an available and affordable resource.

Features



8

Departments

CRAY CHANNELS is a quarterly publication of the Cray Research, Inc., Marketing Communications Department, Tina M. Bonetti, Director. It is intended for users of Cray Research computer systems and others interested in the company and its products. Please mail feature story ideas, news items, and Gallery submissions to CRAY CHANNELS at Cray Research, Inc., 1440 Northland Drive, Mendota Heights, Minnesota 55120. Subscription inquiries and address changes should be sent Attention Dept. D.

Volume 13, Number 2

Editorial staff
Ken Jopp, editor
Mary Miller, associate editor
Vicky Frank, contributing editor
Heidi Hagen, contributing editor

Design and production
Barbara Cahlander, typesetter, production artist
Eric Hanson, graphic designer
James Morgan, art director
Cynthia Rykken, graphic designer

PETROLEUM

ANNOUNCING

The CRAY Y-MP4E computer system

You are challenged to turn your best ideas into better solutions. Now the power to make a difference is more accessible than ever before.

For the past 15 years, top scientists and engineers have relied on Cray Research supercomputers to turn their best ideas into better solutions. To bring this capability to more users than ever before, Cray Research announces the CRAY Y-MP4E computer system, a supercomputer that delivers the performance of previous four-processor CRAY Y-MP systems at a significantly lower cost. The CRAY Y-MP4E system is based on the same new technologies used in the CRAY Y-MP8E, CRAY Y-MP8I, and CRAY Y-MP2E systems. The CRAY Y-MP4E system is an affordable, easy-to-use supercomputer that enables engineers and scientists to bridge the gap between inspiration and innovation.

The CRAY Y-MP4E system offers unmatched performance with two or four central processing units (CPUs) and 32 or 64 million words of central memory. To provide the highest possible performance for a full spectrum of applications, the system features a balanced architecture and all new I/O and SSD technology. This arcitecture provides sustained processing rates of more than one billion floating-point operations per second on a variety of applications.

Input/output subsystem

The CRAY Y-MP4E system uses the new input/output subsystem (IOS) technology introduced in 1990 with the CRAY Y-MP2E system. With support for high-performance DD-60 disk drives, the new IOS allows users to access over one-half terabyte of data at twice the transfer rate of earlier CRAY Y-MP systems.

To increase the system's production workload capacity, the CRAY Y-MP4E system has an aggregate I/O bandwidth of over 2000 Mbytes/sec to peripheral devices, including Cray Research's SSD solid-state storage device. This large bandwidth allows users to access more peripheral devices and perform more simultaneous activities than previously possible. CRAY Y-MP4E systems support 1000-Mbyte/sec and 200-Mbyte/sec channels.

The 1000-Mbyte/sec channels transfer data between central memory and the SSD. The 200-Mbyte/sec, full-duplex channels transfer data between central memory and the IOS or between the IOS and an optional SSD.

Solid-state storage device

The CRAY Y-MP4E system offers new SSD technology that provides high-speed, reliable data storage at half the cost of previous SSD systems. Using VLSI chips and increased system integration, the SSD is available with up to 512 Mwords of storage capacity. This large capacity allows users to solve larger problems and increase I/O throughput.

The CRAY Y-MP4E system communicates with the SSD through one or two 1000-Mbyte/sec channels. The SSD is connected to the IOS through

up to three 200-Mbyte/sec channels. These connections enable data to be transferred directly between an IOS and the SSD without passing through central memory, thereby increasing overall system performance.

Disk drives

Cray Research offers fast, reliable mass storage devices that give users access to more data at faster rates than was possible with previous disk drives. The CRAY Y-MP4E system supports all current Cray Research disk storage devices including the DD-60 and DD-61 disk storage units.

Software

Cray Research emphasizes performance-oriented, feature-rich software products that enhance the hardware capabilities of its systems. Cray Research provides the most complete body of system software available on any supercomputer system.

The CRAY Y-MP4E system runs Cray Research's UNICOS operating system, which is based on the UNIX System V operating system from UNIX System Laboratories, Inc. The UNICOS system is an interactive operating system that excels in performance, functionality, portability, and connectivity. The UNICOS operating system includes the UNICOS Storage System, a powerful file management system. When running the UNICOS Storage System, the CRAY Y-MP4E system can satisfy requests from multiple supercomputers over gigabit networks while providing service to smaller systems, workstations, and personal computers. The CRAY Y-MP4E system can function as a file server and perform scientific processing simultaneously.

Cray Research also offers the most powerful compilers in the industry, including the CF77 Fortran compiling system, and Cray Standard C, Cray Ada, and Pascal compilers.

Cray Research supports more than 600 application programs used by industrial engineers and scientists to solve problems in the aerospace, automotive, chemical, petroleum, and other high-technology and manufacturing industries. Cray Research systems running these programs can be used to shorten product design cycles, optimize manufacturing processes, and solve basic engineering and research problems.

High reliability

The CRAY Y-MP4E supercomputer provides high system reliability along with high performance. System quality begins with a design process that integrates quality and reliability into every system component. All components undergo strict inspection and checkout prior to assembly. Before shipment, all CRAY Y-MP4E computer systems undergo rigorous operational and reliability tests.

To maximize system availability, Cray Research has developed advanced system support tools including the new System Maintenance and Remote Test Environment (SMARTE), which provides continuous error detection and isolation. SMARTE schedules all on-line diagnostic activity, automatically reports errors, and provides a common X Window System interface to all on-line diagnostics, concurrent maintenance tools, and expert systems analysis.

CRAY Y-MP4E system highlights

- ☐ Up to four processors
- ☐ 6-nanosecond clock cycle
- ☐ Very Large Scale Integration (VLSI) gate-array circuits
- ☐ Flexible hardware chaining for vector operations
- ☐ Gather/scatter and compressed index vector support
- ☐ Flexible processor clustering for multitasking applications
- ☐ Four parallel memory ports per processor
- ☐ Dedicated registers for efficient interprocessor communication and control
- ☐ SECDED memory protection

Model	CPUs	Central memory (Mwords)	IOCs	Optional SSD (Mwords)
CRAY Y-MP4E/232	2	32	1 - 3	128, 256, or 512
CRAY Y-MP4E/264	2	64	1 - 3	128, 256, or 512
CRAY Y-MP4E/432	4	32	1 - 3	128, 256, or 512
CRAY Y-MP4E/464	4	64	1 - 3	128, 256, or 512

CRAY Y-MP4E computer system configuration options.

Real supercomputing within reach

The CRAY Y-MP4E computer system makes supercomputing affordable to a broad range of users. It provides the performance of previous four-processor CRAY Y-MP systems but costs less to acquire, install, operate, and maintain. Several factors lower the ownership costs:

- ☐ The system can be air cooled in a standard computer room environment with commercial air conditioning. A water cooling option also is available.
- ☐ The system runs on commercially available electrical power; no motor generator set is required.
- ☐ The system has fewer component interconnections than previous systems, simplifying installation.
- ☐ The system's high reliability and enhanced maintenance features enable customers to select from a range of maintenance options, from round-the-clock, on-site coverage to on-call hardware service. Customers who choose remote access for service receive an additional discount.

The CRAY Y-MP4E system offers outstanding performance and functionality at a lower cost than previous comparable systems. As with all CRAY Y-MP systems, it excels in a wide range of applications and allows users to solve problems that could not reasonably be attempted on other computer systems. Backed by Cray Research's unmatched experience with total supercomputing solutions, the CRAY Y-MP4E system gives users the power to bridge the gap between inspiration and innovation.

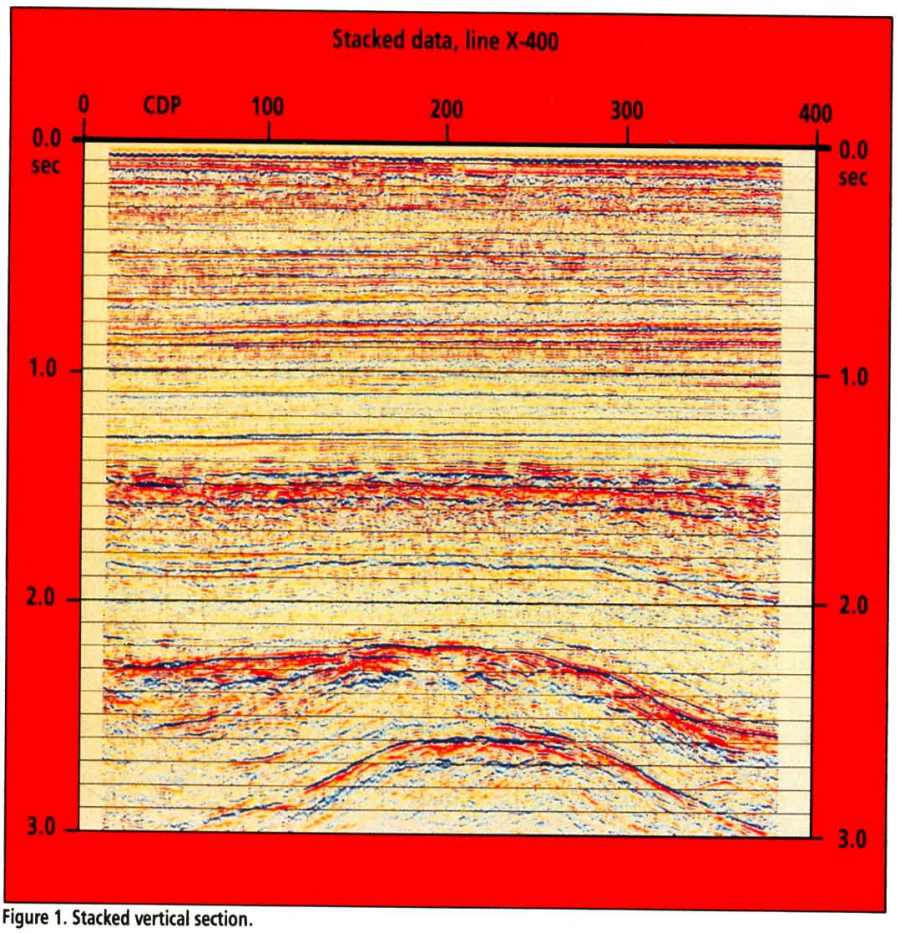


Figure 1. Stacked vertical section.

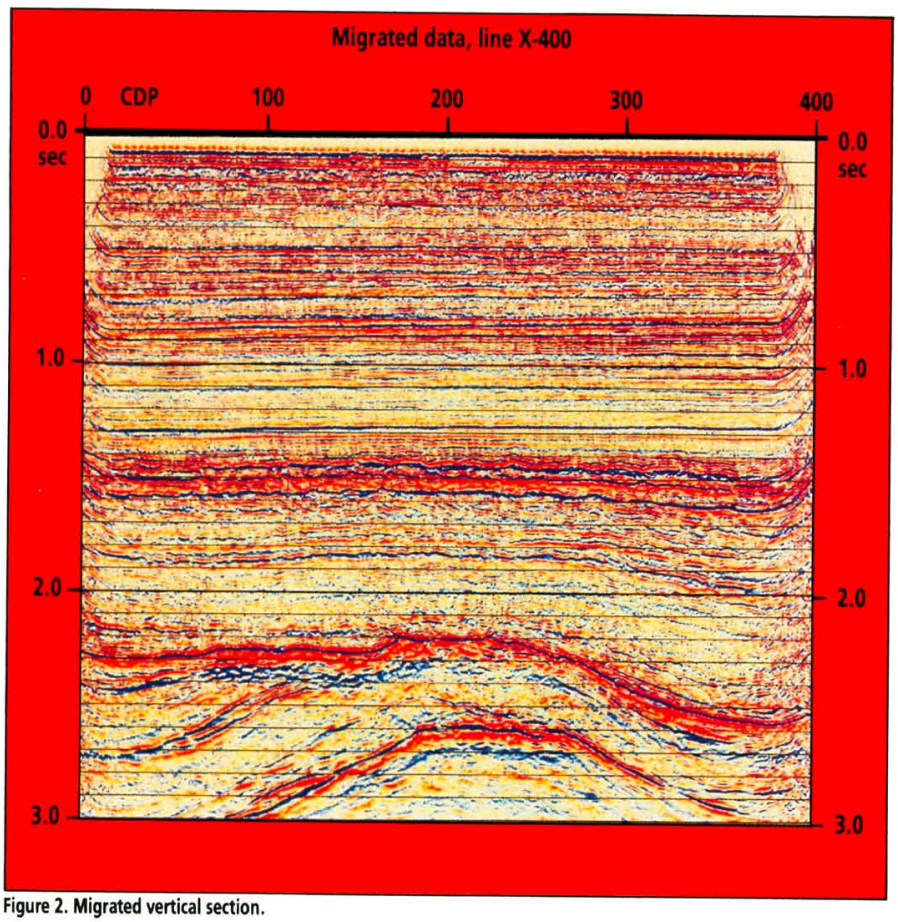


Figure 2. Migrated vertical section.

3-D time and depth migration at gigaflops speed

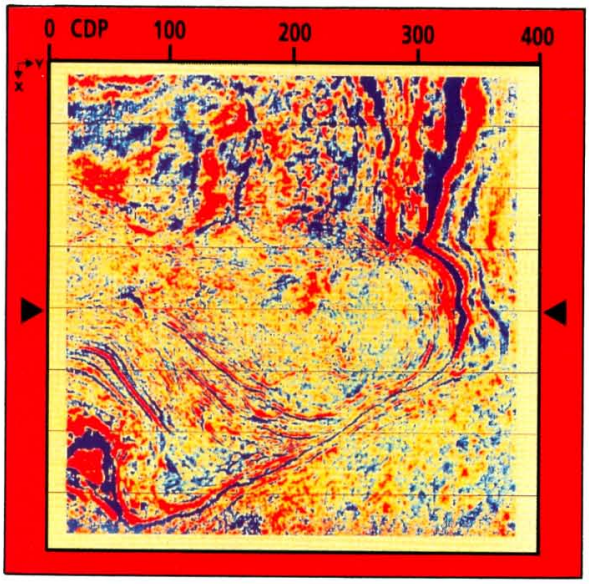
Meinhard Holling, Prakla-Seismos AG, Hannover, Germany
Klaus Ketelsen, Cray Research GmbH, Munich, Germany

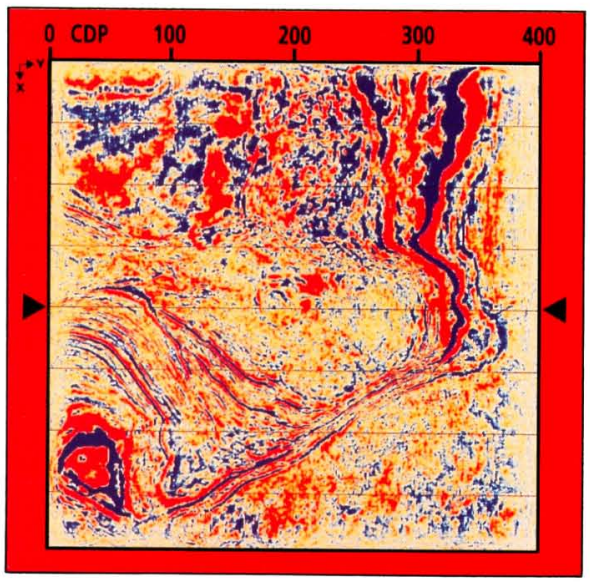
Three-dimensional (3-D) data migration is one of the more computationally intensive seismic data processing problems. Seismic data typically are collected by reflection seismology, a process that induces a pulse at the Earth's surface and records reflections back to the surface on arrays of receivers (geophones). Processing these data provides information about velocities in the subsurface and makes geological events and structures visible. This enables researchers to "view" geological structures for possible deposits of petroleum and natural gas.

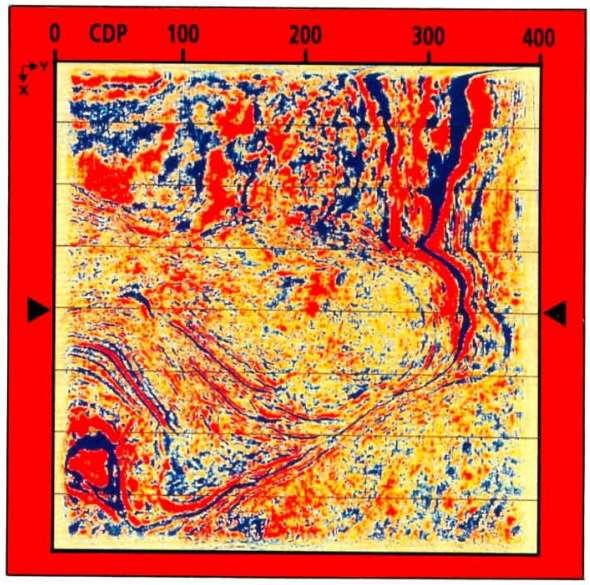
Prakla-Seismos AG is a leading international geophysical contractor involved in all aspects of modern processing and interpretation techniques. Seismologists are running new generation 3-D migration techniques implemented in Prakla-Seismos' processing package GEOSYS on the CRAY Y-MP2/232 system at the Geophysical Data Center of Prakla-Seismos in Hannover, Germany.

3-D migration schemes

3-D migration is an indispensable processing step when exploring areas with salt domes and folded belts. Even if collected along two-dimensional lines, seismic wave fields have a three-dimensional nature and contain reflected and diffracted energy from offline structures. The growing demands for better resolution and for the extraction of all information available in recorded data have led to the development of a variety of 3-D migration methods. The two-pass approach is







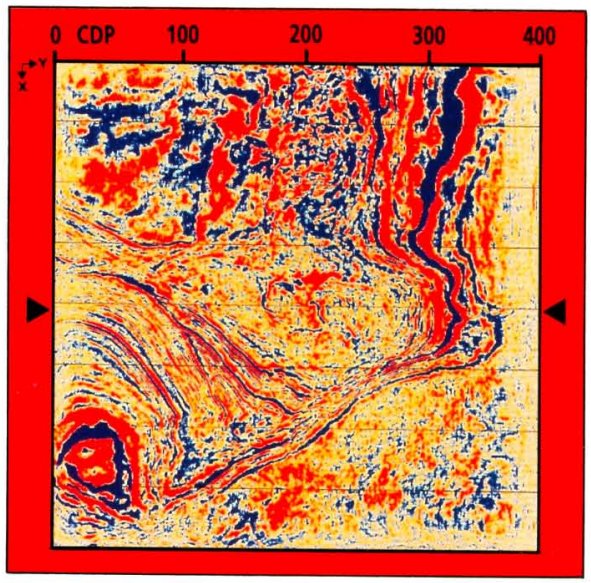


Figure 3. The process of 3-D migration: initial stacked data at 2500 ms (left), partially migrated data up to 500 ms (right) and 1500 ms (below left), and migrated data at 2500 ms (below right).

based on a full separation of the 3-D migration operator in two orthogonal directions, while the one-pass approach relies on a splitting technique of the full 3-D downward continuation operator.

Because the 3-D migration operator is separable completely and exactly only for homogeneous media, the two-pass method has limited value in areas with subsurface velocity variations. The fact that data management is easier in the two-pass mode, and that fast migration methods can be implemented easily with this method made it a viable compromise before true 3-D migration schemes could be run in reasonable time on powerful vector supercomputers.

The one-pass approach, on the other hand, is applicable for arbitrary inhomogeneous media. A proper imaging of very complex structures can be achieved only by a one-pass 3-D depth migration.

Actual 3-D data processing, however, requires considerable cost for data handling. By using parallelized algorithms and a special I/O organization, the one-pass approach has been implemented for efficient 3-D migration of large volumes of seismic data. Results of a real-world example are shown in Figures 1 to 3.

Organization scheme of the time migration

In the one-pass time migration, downward continuation is achieved in one pass through the 3-D volume. Splitting techniques are used to separate the computation in the *x* and *y* directions. To compute a migrated level, all time slices below the current slice have to be processed in the *x* and *y* directions, respectively. In older processing schemes this meant reading the partially migrated data from disk, computing the data, and writing it back to disk. This kind of implementation makes the 3-D migration very I/O intensive. Multitasking by this older method is not very efficient, because even a one-CPU version of the algorithm would be I/O bound.

Due to the large memory available on Cray Research systems, two major changes were made in the migration program. First, a complete layer now can be processed instead of only parts of a layer, as in the previous version. As a result, data can be read from disk in large chunks, which eliminates the seek time of the disk drives. Second, the migration scheme has been changed to reduce I/O at the expense of

memory use. Looking at the migration algorithm, we can see that six slices are involved in computing a new, partially migrated slice.

tj(new) {tj(old), tj+1(old), tj+2(old), tj+1(new), tj+2(new) }

t A partially migrated time slice
i Current level to be calculated
j Current time slice i < j < last level

After computation, tj + 2(new) will not be used any more for this step. In older schemes it was written to disk at this point. In the new scheme, tj(new) stays in memory instead of being written to disk. Delayed by three slices, the computation of the *y* part of layer *i* is performed.

tj+3(new) {tj+3(old), tj+4(old), tj+5(old), tj+4(new), tj+5(new) }

With 12 slices in main memory, two computations can be carried out in one step, and I/O is reduced by a factor of two.

Figure 4 shows an example of computing two levels in parallel. Both levels are computed in the *x* and *y* directions simultaneously; that is, a total of four blocks is computed concurrently. All six slices belonging to one block are displayed in the same color. The colored slices reside in central memory; the rest are located on disk. Because the computations in the four blocks are independent of each other, they can be executed on different CPUs in a multiprocessor environment. The scheme described in Figure 4 can keep up to four CPUs busy. To run the 3-D migration on an eight-processor CRAY Y-MP system, four or even eight levels with eight or even sixteen calculations have to be computed in parallel.

Even with the reduced I/O scheme, it is difficult to keep the program CPU bound. The program uses four disk channels in parallel, and each is driven by asynchronous I/O. An average disk speed of up to 24 Mbytes per second is achieved during the total run.

Number of layers	Number of CPUs	Memory (Mwords)	Transferred Data (Gbytes)	Average transfer rate (Mbytes)	Wall clock time (min)	MFLOPS
.5	1	4	647	24	450	173
1	2	7	324	24	220	355
2	2	13	163	14	187	418
2	4	13	163	23	116	675*
4	4	27	82	14	98	816*
4	8	27	82	22	61	1280
8	8	55	42	13	52	1591

Table 1. 3-D migration performance with various hardware configurations (*interpolated).

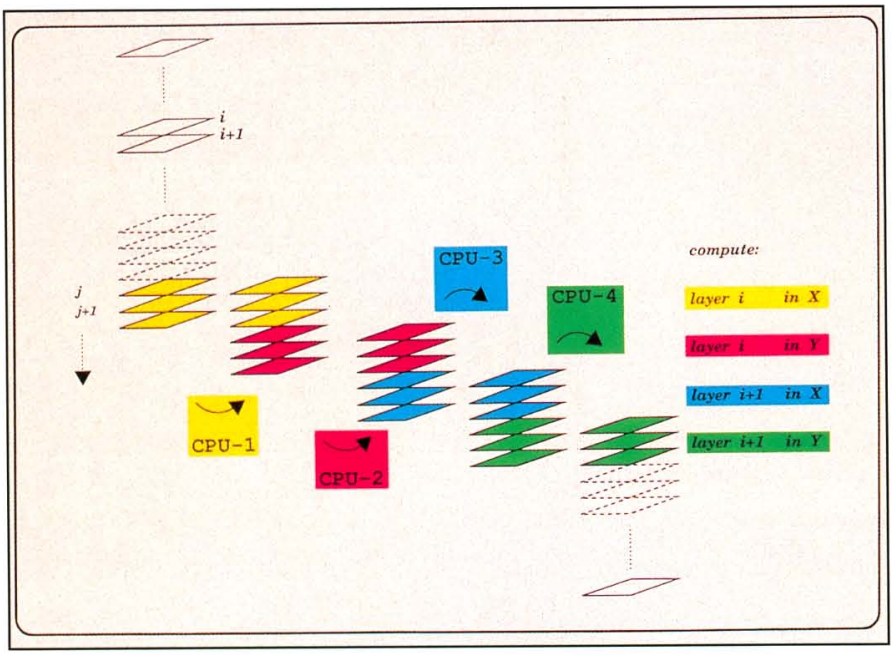


Figure 4. Memory and CPU organization.

The test case

The test case for the 3-D time migration was an off-shore seismic survey of the North Sea. The survey was carried out to investigate the stratigraphic distribution of jurassic sandstones. The size of the survey is 800 x 400 CDPs (common depth points) or bins and a processing length of five seconds, that is, 1250 layers for the actual sampling interval of 4 ms. This results in a dataset of 1.6 gigabytes, if the data is stored in 32-bit format. Within this job, 104 million tridiagonal equation systems have to be set up and solved with a matrix size of 400, and 52 million systems with a matrix size of 800. Figures 1 and 2 show a comparison of stacked and migrated data of a vertical section. The position of this section is marked by an arrow in the time slices of Figure 3.

In general, stronger events, such as bigger changes in the reflection coefficients, can be seen throughout the migrated section. Greater color contrasts indicate these stronger events. In the salt dome region, in the lower part of the section, some events with a steep dip appear in the migrated section that cannot be seen in the stacked one. Obviously the migration process is most critical in areas with large dips.

Figure 3 shows how the events move during migration. Starting at the top left, stacked data of the layer at 2500 ms can be seen. The two intermediate displays show this layer partially migrated while the total volume is fully migrated up to 500 and 1500 ms. The lower right display shows layer 2500 finally migrated.

As in the sections, stronger events can be seen in time slices. Some events come into focus during the migration progress. Particularly in the middle part, steeply dipping horizons move (or migrate) until they become focused in their real positions.

Performance of the 3-D migration

The 3-D time migration program is designed to run on different configurations of Cray Research systems. The number of layers that can be handled concurrently is calculated automatically depending on the available memory, the number of CPUs, and the number of CDPs in the seismic dataset.

Table 1 shows some of the performance numbers obtained when running our test dataset on various configurations. "Number of layers" refers to the number of layers handled concurrently. The .5 in the first row means that the *x* and *y* directions are calculated in different passes. This may become

necessary if extremely large 3-D volumes have to be processed.

As demonstrated, it is possible to run a 3-D time migration in less than one hour on a CRAY Y-MP8 supercomputer. The first production dataset procured with the new time migration version came from a large on-shore survey with 1.25 million CDPs and a processing length of 1500 samples. The migration kernel ran in 21 hours elapsed time on Prakla-Seismos' CRAY Y-MP system. One layer could be processed concurrently. The job used 26 Mwords of memory and worked with a dataset of eight gigabytes on disk.

This speed has enabled researchers at Prakla-Seismos to develop a new generation of 3-D migration processing. Migrations in the target areas now can be run several times with different velocities or processing parameters, a capability in high demand with Prakla-Seismos' customers. Another positive aspect is turnaround time. Fast processing is a basic prerequisite when drilling crews stand by waiting for crucial results.

The high performance of the migration algorithm has encouraged researchers at Prakla-Seismos to plan additional improvements such as:

- ☐ Handling of steep dips. Dips of up to 80 degrees will be handled correctly.
- ☐ Additional splitting in diagonal directions. In addition to splitting in the *x* and *y* directions, the algorithm will split in the two diagonal directions, resulting in four passes for each layer.

Each of these two changes in the migration algorithm will double the required CPU time.

The 3-D depth migration program is designed similarly to the time migration program. It also can run on different Cray Research system configurations. Due to the algorithm, production jobs performing depth migration are much more computationally intensive than comparable time migration jobs. Depending on job parameters, depth migration requires three to eight times more CPU time.

On-line display of the results

At the 1990 SEG convention in San Francisco, Cray Research showed an on-line 3-D migration. A workstation at the exhibit was connected to a CRAY Y-MP8 supercomputer in Eagan, Minnesota, via a T-1 link. As soon as the migration of a layer had been completed on the CRAY Y-MP system, the layer was displayed on a workstation. An interface program looked into the working files of the migration, extracted the data, and sent them to the Metaplot graphics package for display. Users also could select a slice at a fixed time and display the partially migrated result whenever a layer was completed. This procedure was used to create the displays in Figure 3.

Interactive control of the migration process

The combination of a fast 3-D migration and online displays opens new possibilities for 3-D migration. Users can not only display the just completed layer, but also look at the whole migration volume at any time. The fully and partially migrated data, as well as the migration progress, can be observed periodically. As soon as it appears that migration parameters, such as velocities, have to be changed, the migration job can be suspended. The migration process periodically writes checkpoints to disk or tape. After parameters have been changed, the migration can be set back one or two checkpoints and the job restarted.

The Prakla-Seismos/ Cray Research team achieved 1.59 GFLOPS when eight layers were calculated concurrently on the CRAY Y-MP system. The code has been assimilated into Prakla-Seismos' GEOSYS production system and is used in seismic applications routinely.

Winning the Gigaflop Performance Award

To recognize researchers who have achieved high levels of computing performance, Cray Research instituted a Gigaflop Performance Award program in 1989. The 3-D time migration team received a Gigaflop Performance Award in 1990. Table 1 shows that the Prakla-Seismos/Cray Research team achieved 1.59 GFLOPS when eight layers were calculated concurrently on the CRAY Y-MP system. The code has been assimilated into Prakla-Seismos' GEOSYS production system and is used in seismic applications routinely.

Conclusion

The new migration techniques described above have turned a demanding batch application into an interactively controlled process. As demonstrated, the availability of large memory allows typical I/O-bound programs to become CPU bound, even in a multiprocessor environment. The high performance provided by Cray Research supercomputers opens the possibility of running new algorithms, such as 3-D depth migration, even on large datasets, which was not possible in a reasonable time on previous generations of computer systems. However, the challenges of exploration seismology are by no means diminished by the ability to process 3-D migration on supercomputers. Geophysicists now point to the need for teraflops performance to address the problem of "pre-stack" 3-D depth migration. ▬

Acknowledgments

The authors thank the staff at Prakla-Seismos, especially Emil Hinrichs, for support during the preparation of this work; BEB Erdgas und Erdöl, Hannover, Germany, for supplying the dataset for the test case; and Interactive Network Technologies, Apple Valley, Minnesota, for permission to use Metaplot to create the seismic displays.

About the authors

Meinhard G. Holling is a senior geophysical researcher at the Prakla-Seismos data processing center in Hannover, Germany. After receiving his degree in mathematics and physics in 1978, Holling worked on interpretation software at Sattlegger Engineering for Applied Geophysics before he moved to Prakla-Seismos in 1985, where he researches and develops seismic processes. Holling specializes in 3-D migration with Prakla-Seismos' GEOSYS package.

Klaus Ketelsen is a sales support analyst with the marketing and sales support group at Cray Research GmbH, Munich, Germany. After receiving his degree in electrical engineering from the University of Hannover in 1978, he worked in the petroleum industry for nine years. At Cray Research, Ketelsen specializes in developing and optimizing codes for seismic customers, in particular Prakla-Seismos in Hannover.

Pseudo-spectral algorithms and seismic modeling

Maynard S. Redeker and Christopher P. Ross,
Oryx Energy Company, Richardson, Texas

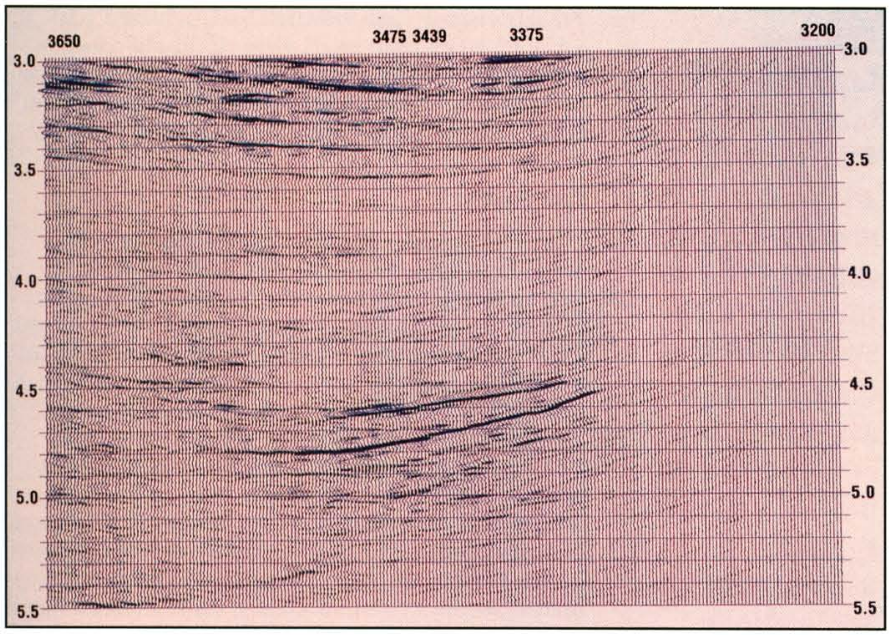


Figure 1 (above). Seismic reflection profile Line B, windowed from 3000 to 5500 ms. Relative amplitude processing illustrates two strong reflectors (bright spots) at 4600 ms and 4700 ms.

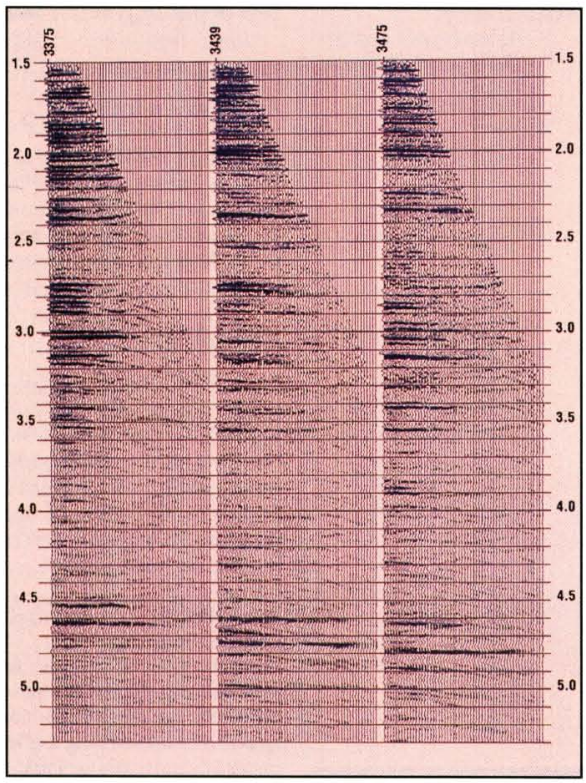


Figure 2 (right). Three NMO corrected CMP gathers over the bright spots in Figure 1. The reflectors of interest are at 4600 and 4700 ms. The location of each CMP is annotated in Figure 1.

The origins of seismic modeling can be traced back to the 1880s, when the first effective seismometers were built to measure movement in the Earth's surface caused by earthquakes. Although seismology still is used for earthquake detection, a derivative of seismology, exploration geophysics, uses seismic modeling extensively to evaluate hydrocarbon prospects. Seismic modeling for these advanced applications requires a high degree of accuracy, which has been obtained with the pseudo-spectral modeling algorithm on a Cray Research supercomputer.

The CMP (common midpoint) gather method is used to obtain subsurface data for seismic modeling in petroleum exploration. This method begins with a controlled explosion that generates wave motions reflecting and refracting off geological interfaces encountered by the wave front in the subsurface. To record these wave motions accurately enough to reconstruct a large area of the subsurface, over 200 colinear receivers are positioned from a few hundred to over 15,000 feet from the source. By repeating this process in a line at select source-receiver positions, a seismic reflection profile can be constructed. Simple geometry usually determines the paths the reflected energy has taken from source to receiver. Source-receiver pairs then are sorted into groups that have common reflection points in the subsurface. For flat subsurface interfaces, these common reflection points should be midway between the source and the receiver. Hence the term common midpoint (CMP) gather for this method of data acquisition.

Because reflections for shorter ray paths (reflected energy paths from source to receiver) arrive before reflections for longer ray paths, velocity data must be incorporated into the processing to correct the differences. After velocity corrections (normal move out, or NMO) have been applied, the data are "stacked" by summing all the source-receiver traces in a CMP gather and dividing the data by the sum of the traces. Stacking the data improves the signal quality and generates a single trace that approximates normal incidence at the common midpoint. The normal incidence approximation assumes that the source and receiver are at coincident locations and the elastic energy from the source travels along the same path (down and up) to and from the geological reflector. All stacked CMP traces then are plotted together to create a CMP stacked reflection seismic profile (Figure 1). Reflections from subsurface layers then can be interpreted and time-mapped. At this point, localized subsurface structures that appear to contain hydrocarbons can be identified and pursued.

While the CMP method just described is effective on simple, theoretical data, it does not take into account more complex subsurface areas that make acquisition, data processing, interpretation, and modeling much more difficult. To account for these complexities, more efficient computational algorithms must be coupled with supercomputers, such as Cray Research systems, to reduce the complexities and make effective imaging of the subsurface possible.

Seismic modeling

Even with the advances of supercomputing, subsurface distortion and ambiguity of the elastic energy's

ray path still occur, and they affect the ability to focus and interpret the image. Often after the data have been reduced as much as possible, the subsurface remains "out of focus." The geophysicist then must interpret the data as is, hopefully correctly correlating seismic data and subsurface geology. Seismic modeling is one method that assists the geophysicist in identifying the correct correlation.

Table 1 reflects the four broad categories of data and seismic modeling available to the geophysicist. For each of these quadrants, several algorithms can be used to obtain the best solution. Choosing the best type and method depends on the specific geography; often this decision is a matter of balancing computer resources with algorithm sophistication to get a "good enough" answer in a reasonable time.

For advanced seismic applications, a "good enough" answer may not be good enough, however. Since the acceptance and proliferation of the CMP method, geophysicists have been trying to glean more information from the seismic data than just a structural image. Stratigraphic knowledge, details about the interstitial pore fluids contained in various geologic formations, is extremely important in detecting the presence of hydrocarbons. An inversion process is used to extract such information from seismic reflection data. Amplitude versus offset (AVO) is one type of inversion that in some formations can detect oil and gas pore fluids.

The problem

AVO inversions on prestacked data require high quality signals on all CMP gather source-receiver pairs (offsets). When data from certain offsets are degraded by various phenomena, the inversion process becomes inaccurate and misleading. Figure 2 presents recently recorded CMP gathers from the Gulf of Mexico without NMO corrections. On the event at 4.6 and 4.7 seconds on the far offsets for each CMP record there is a loss of reflected energy; interference from a salt diapir is suspected. Since the AVO inversion process depends on data from the far offsets, the data must be qualified before inversion. In this case, the prestack statigraphic method referred to in Table 1 is the most appropriate for this type of geological framework.

Salt diapir model

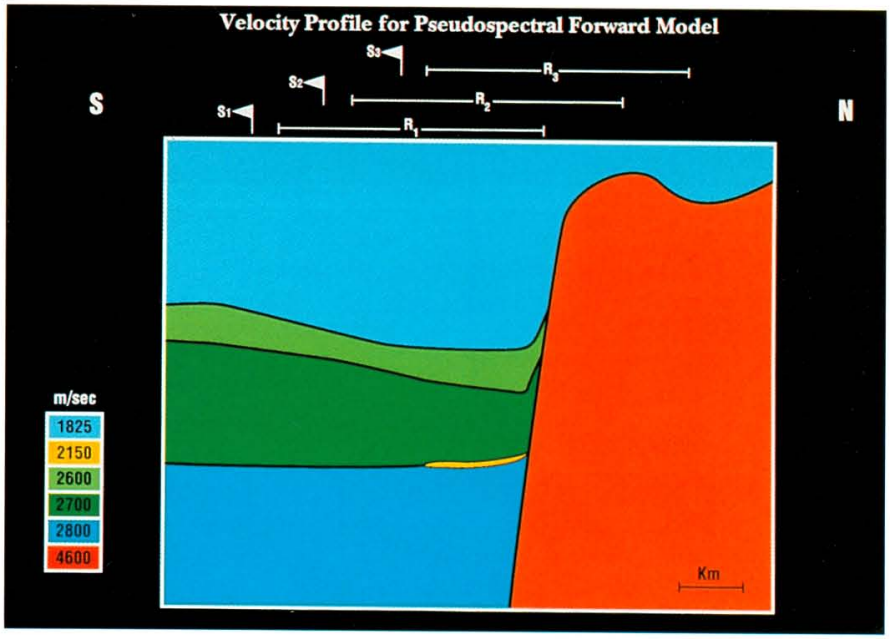
Figure 3 illustrates a modeled geologic cross-section with a large salt diapir structure shown in orange. The model is 9 km wide and the objective sand (shown in yellow) is 5 km deep. Because salt is less dense than the overlying sediments, viscous salt has flowed and thrusted upward through the sediment layers. Salt intrusions such as this often upturn sediments, which in turn trap oil and gas accumulations. In the figure, the yellow gas sand is upturned and trapped against the salt diapir. Because this gas sand is 40 m thick and has a lower velocity and density than the surrounding sediments, it should have a very distinguished reflection and therefore be relatively easy to identify on the modeled results. Synthetic CMP gathers will be shot from source locations (S1, S2, S3) into receiver spreads (R1, R2, R3) as annotated at the top of the model in Figure 3.

The algorithm

The algorithm used to model the synthetic CMP is a pseudo-spectral wave equation. In this case, the choice of an algorithm was based on a balance of computer resources and accuracy. The pseudo-spectral

Table 1. Data type vs. desired modeling.

	Type of modeling	
Data	Structural modeling prestack data	Stratigraphic modeling prestack data
	Structural modeling poststack data	Stratigraphic modeling poststack data



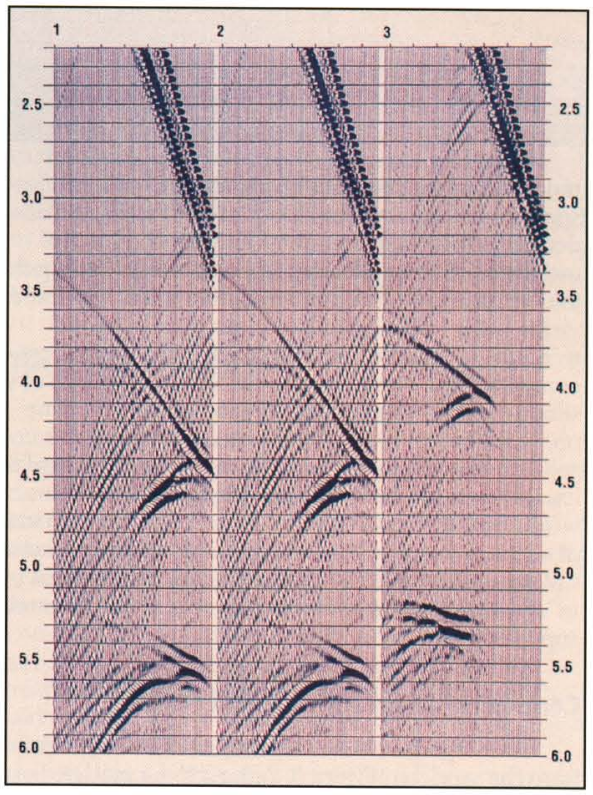


Figure 3 (above). Velocity cross-section used for PSMA modeling. One bright spot has been inserted at the appropriate depth (yellow horizon).

Figure 4 (left). Three PSMA CMP gathers generated using the velocity profile in Figure 3 and PSMA. These events have not been NMO corrected. Note that record 3 is missing data on the far offset traces. The reflective nature of the bright spot also is apparent just before the attenuated traces.

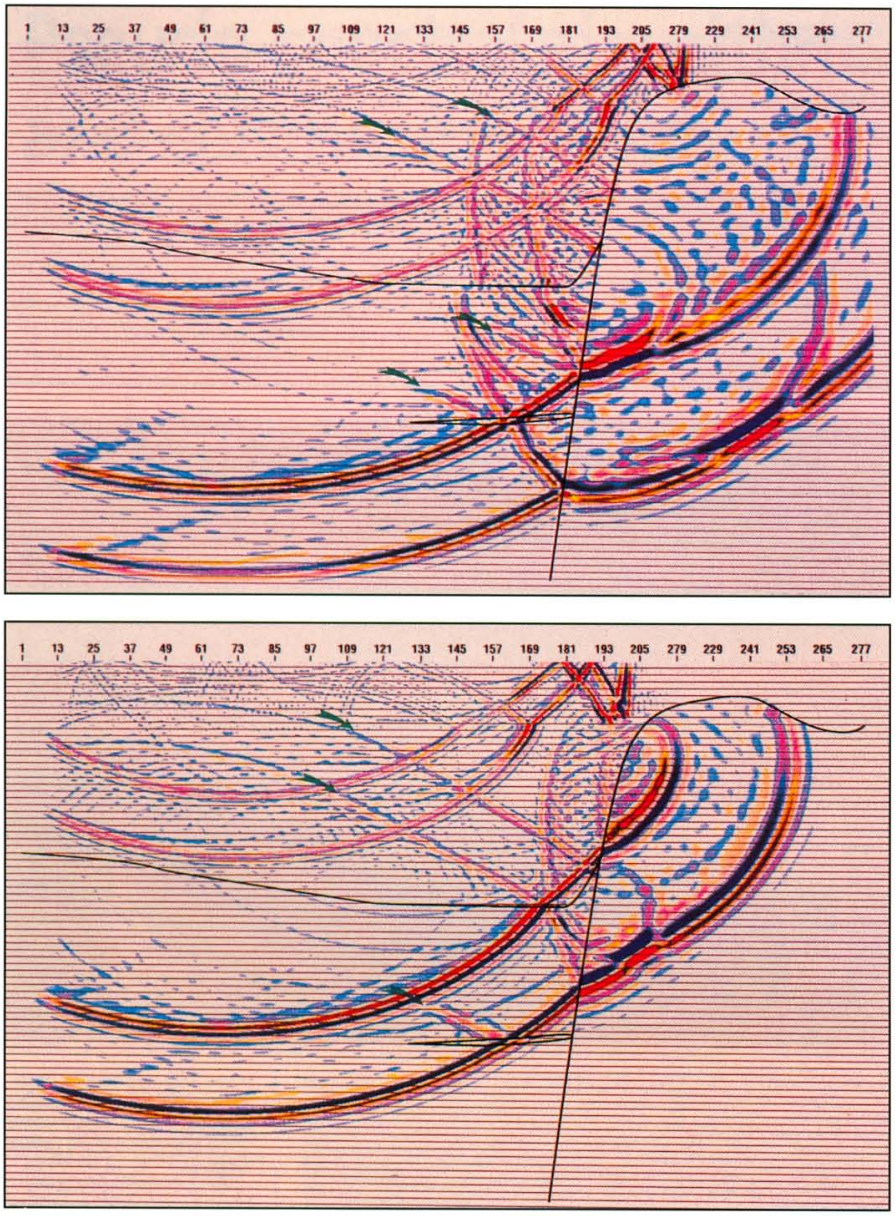


Figure 5. Wave field snapshots at 2.6 (top) and 3.0 (bottom) seconds. The primary incident wavefront has just passed through the bright spot, and the resultant upgoing reflection is apparent (lowermost arrow). Arrows denote reflected upgoing energy. Note the lack of coherent upgoing energy inside the salt diapir.

modeling algorithm (PSMA) was chosen for both speed and accuracy. The PSMA uses efficient spatial derivative calculations in the FFT domain that are more precise than other algorithms that do not compute the derivatives in this domain. Additionally, the PSMA yields higher resolution than other wave equation techniques, allowing more detailed analysis with less computational overhead. The full-elastic PSMA results shown in Figure 4 required 15 minutes of CPU time on a CRAY X-MP/48 system with 32 Mwords of solid-state storage. Attempting to replicate these results using a comparative algorithm (finite-difference) would take upwards of 6 hours of CPU time.[1] As part of a comprehensive modeling process, many synthetic CMP records should be generated. To complete this in a reasonable time, the power of a Cray Research supercomputer is required.

Comparing actual and modeled results

Figure 4 shows the synthetic CMPs generated from the model in Figure 3. Although the modeled CMPs have not had NMO corrections, the energy on the far offsets is diminished at 5.3 seconds for record 3. This is strikingly similar to the actual CMP shown in Figure 2. Obviously, the reflected energy on the far traces has been attenuated, discrediting any data that may be considered for AVO inversion.

Figure 5 shows wavefield snapshots at 2.6 and 3.0 seconds. Snapshots show how the energy radiates away from the source and is reflected or refracted by subsurface geology. The cross-sectional model shown in Figure 3 has been overlaid for orientation. Primary and first-order multiple downgoing wavefronts have the largest amplitudes and are therefore easily identified. Arrows denote the upgoing reflected data in both figures. The lack of coherent upgoing energy inside the salt mass, energy that would be recorded on the far offsets, further reinforces the attenuated energy on the far traces of the modeled and actual CMP data.

Conclusions

In this example, a gas sand adjacent to a salt diapir cannot be subjected to AVO inversions because of degradation to the far offsets of the example CMP. PSMA generated similar results to confirm suspicions of salt dome interference with the data. These results are made possible by the efficient use of computational algorithms such as the pseudo-spectral wave equation algorithm coupled with the computational resources of the Cray Research supercomputer. Together they provide the geophysicist with highly accurate subsurface modeling at reasonable computing costs and turnaround time. ▬

Acknowledgments

The authors thank TGS for the Gulf of Mexico data and Oryx Energy Company for time and resources.

About the authors

Maynard S. Redeker is the manager of Geophysical Technology at Oryx Energy Company (formerly Sun Exploration and Production Company). He previously was manager of Geophysical Services. He received his B.A. degree in physics from Emporia State University in 1966, his Ph.D. degree in physics from the University of Kansas in 1970, and graduated from the SMU Executive Excellence Program in 1989. He holds several professional patents and has written several publications related to seismic prospecting methods.

Christopher P. Ross is a senior professional geophysicist at Oryx Energy Company. Since joining Oryx in 1988, his primary duties have been hydrocarbon fluid and thickness prediction using HCI/AVO methods on the Gulf of Mexico and international lease acreage. He previously was a geophysicist for Amoco Production Company and a division geophysicist for Schlumberger Well Services. He received a B.S. degree in geophysics from the Virginia Polytechnic Institute and State University in 1982, an M.S. degree in applied physics from the University of New Orleans in 1984, and is a Ph.D. candidate in geophysics at Georgia Institute of Technology, working on problems associated with AVO.

Reference

1. Kosloff, D. D., and E. Baysal, "Forward Modeling by a Fourier Method," *Geophysics*, Vol. 47, No. 10, pp. 1402-1412, 1982.

A new parallel iterative linear solution method for large-scale reservoir simulation

J. R. Wallis, J. A. Foster, and R. P. Kendall,
Western Atlas Software, Houston, Texas

Numerical reservoir modeling has become a standard production tool throughout the petroleum industry. Numerical models of recovery operations give petroleum engineers valuable information about the likely behavior of fluids beneath the Earth's surface. The accuracy of such models depends on many factors, including the resolution of the computational grid used to divide a given reservoir into discrete spatial volumes. The dimensions of the largest black-oil models have reached a few hundred thousand grid blocks. However, the practical limit for field-scale enhanced oil recovery (EOR) models is only a few tens of thousands of grid blocks. Models of this magnitude strain the processing capabilities of the current generation of sequential, albeit vector, supercomputers. EOR applications, therefore, pose a serious computational challenge, especially when high spatial resolutions and implicit formulations are required. Whereas the computational demands of a reservoir model grow in direct proportion to the number of grid blocks, they grow at least quadratically with the number of implicit unknowns. Moreover, the quadratic growth with implicit unknowns occurs only where approximate factorization suffices for the linearized flow equation; otherwise, the growth is cubic.

Many computer platforms include a largely untapped resource that can be used to help meet these computational demands: the parallel processing capability of multiprocessor systems. This capability remains largely untapped due to a lack of reservoir simulation software that effectively exploits parallelism. Numerical linear algebra algorithms constitute a long-standing bottleneck in reservoir simulation. Robust linear solvers make parallel processing difficult because they are recursive and have fine granularity. The solution of the linearized conservation equations required during each Newton step of simulation can consume as much as 90 percent of the processor time required to process a timestep in EOR applications. With the exception of the well equations, the remaining computational steps in simulation are naturally parallel. Consequently, our research has focused on the linear solution step. Although the algorithms described here can be applied to the multiple-unknown-per-cell case, we address only the one-unknown-per-cell case: the pressure equation. Complete details of the algorithm are presented in Reference 7. We ran several test problems using this algorithm on an eight-processor CRAY Y-MP system and obtained excellent parallel processing results.

The pressure equation

The reservoir pressure equation is a significant challenge in reservoir simulation because of its elliptic nature. The seven-point finite difference discretization of this equation in 3-D gives rise to a sparse banded matrix structure. This research has focused on an algorithm for the pressure equation that includes generalized conjugate residual (GCR) acceleration[1], nested factorization preconditioning[2], and residual constraints[3,4]. These features are robust in a variety of simulation applications. In particular they can be adapted to models with faults and models to which well equations have been coupled directly. Adaptation of this procedure to parallel processing was based also on domain decomposition[5] and two-way ordering.

A parallel GCR implementation

The GCR algorithm has been reformulated so that the preconditioning is applied to the previous A-orthogonal direction vector rather than the residual. This typically allows only a one-time calculation of the residual and solution iterate, and thus reduces the amount of fine-grain computation. The parallel GCR algorithm is defined as follows:

Choose $\vec{x_o}$	Initial guess	
Set $\vec{r_o} = b - A\vec{x_o}$	Initial residual	
Set $\vec{s_o} = \vec{r_o}$	Initialize polynomial argument	
Define $\|\| \vec{r_o} \|\|_2^2 = (\vec{r_o}, \vec{r_o})$	Initial residual norm	
Set I = 0	Index of latest computed residual	
For i = 1 step 1 until convergence DO		
$\vec{u_i} = M^{-1} s_{i-1}$	Preconditioning	
$\vec{v_i} = A\,\vec{u_i}$	Matrix-vector multiply	
$\alpha_m^i = -\frac{(\vec{s_m}, \vec{v_i})}{(\vec{s_m}, \vec{s_m})}$ $m = 1, 2, \ldots, i-1$	Orthogonalization coefficients	
$\vec{q_i} = \vec{u_i} + \sum_{m=1}^{i-1} \alpha_m^i \vec{q_m}$	New direction vector	
$\vec{s_i} = \vec{v_i} + \sum_{m=1}^{i-1} \alpha_m^i \vec{s_m}$	New A-direction vector	
$\omega_i = \frac{(\vec{r_o}, \vec{s_i})}{(\vec{s_i}, \vec{s_i})}$	Step length	
$\|\| \vec{r_i} \|\|_2^2 = \|\| \vec{r_{i-1}} \|\|_2^2 - \omega i(\vec{r_o}, \vec{s_i})$	Residual norm	(1)
If $\frac{\|\| \vec{r_i} \|\|_2}{\|\| \vec{r_o} \|\|_2} <$ tolerance		
$\vec{x_i} = \vec{x_i} + \sum_{m=l+1}^{i} \omega_m \vec{q_m}$	Update solution vector	(2)
$\vec{r_i} = \vec{r_i} - \sum_{m=l+1}^{i} \omega_m \vec{s_m}$	Update residual vector	(3)
$\|\| \vec{r_i} \|\|_2^2 = (\vec{r_i}, \vec{r_i})$	Update residual norm	(4)
I = i	Index of computed residual	
If $\frac{\|\| \vec{r_i} \|\|_2}{\|\| \vec{r_o} \|\|_2} <$ tolerance stop		(5)
end if		

Reformulated parallel GCR alogrithm.

Another significant difference between this implementation and the standard GCR implementation is the use of Equation (1) to compute the residual norm recursively and thus avoid fine-grain parallel computation inherent in the direct evaluation of the norm of the residual using Equations (2), (3), and (4) at each iteration. The calculation of the residual norm using Equation (1) may induce rounding errors, but Equations (4) and (5) assure the accurate determination of convergence.

Preconditioning

Preconditioning presents the acceleration step of the algorithm with a problem easier to solve than the original one. The robustness of the solver is related directly to the choice of preconditioner. For natural ordering (left to right and top to bottom), the pressure equation matrix is block tridiagonal, for example

$$A = \begin{bmatrix} P_1 & D_1 & & \\ E_2 & P_2 & D_2 & \\ & E_3 & P_3 & \end{bmatrix} \qquad (6)$$

The matrix blocks themselves correspond to planes of grid blocks. Matrix *A* can be factored approximately into lower and upper factors:

$$A \approx \begin{bmatrix} L_1 & U_1 & & \\ E_2 & L_2 & U_2 & \\ & E_3 & L_3 & U_3 \end{bmatrix}$$

$$\begin{bmatrix} I & U_1^{-1} L_1^{-1} D_1 & & \\ & I & U_2^{-1} L_2^{-1} D_2 & \\ & & I & \end{bmatrix} \qquad (7)$$

Preconditioning is introduced through the particular choice of the approximate factors L_iU_i into which each plane P_i is decomposed. The choices investigated during the course of this research for the approximation LiUi are nested factorization and red-black modified incomplete factorization with level one infill, MILU(1).[1] For our implementation of the nested factorization algorithm, a row-sum condition is imposed on the approximate factorization

$$L_i U_i \vec{e} = P_i \vec{e} - E_i U_{i-1}^{-1} L_{i-1}^{-1} D_{i-1} \vec{e}. \qquad (8)$$

In the conventional nested factorization method, Equation (8) is replaced with a column-sum condition and L_i and U_i are completely specified. Hence we designate the row-sum variant with unspecified choice of L_i and U_i as modified nested factorization (MNF).

Constraints

Constraints have been used for decades to improve the convergence of linear solvers. The basic idea is to produce approximate solutions that are deficient in the lowest frequency error components. An alternative strategy is to force the preconditioning to be an exact inverse of matrix *A* on a subspace spanned by the lowest frequency error components. Let

$$C = \begin{bmatrix} \vec{e} & & & \\ & \vec{e} & & \\ & & \ddots & \\ & & & \vec{e} \end{bmatrix}, \qquad (9)$$

where $\vec{e}$ is the vector whose entries are unity with length equal to the dimension of the P_i, that is, the plane size of the model. The number of columns of

C is the number of submatrix blocks A_i. Modified nested factorization satisfies

$$LUC = AC. \quad (10)$$

Thus $U^{-1} L^{-1}$ is an exact inverse of matrix A on the subspace spanned by the columns of C.

Domain decomposition

Domain decomposition has been applied to many problems in science and engineering. The idea is illustrated in Figure 1.

The reservoir grid is divided into regions labeled A_i, separated by strips or boundaries labeled B_i. For our applications, the domains are groups of contiguous vertical planes; each boundary consists of a single vertical plane. For the decomposition of the reservoir depicted in Figure 1, the linearized flow equations can be factored approximately as

$$\begin{bmatrix} A_1 & & & & Q_1 \\ & A_2 & & & Q_2 \\ & & A_3 & & Q_3 \\ & & & A_4 & Q_4 \\ S_1 & S_2 & S_3 & S_4 & B \end{bmatrix} \approx \begin{bmatrix} L_1 & & & & \\ & L_2 & & & \\ & & L_3 & & \\ & & & L_4 & \\ \widetilde{S}_1 & \widetilde{S}_2 & \widetilde{S}_3 & \widetilde{S}_4 & L_c \end{bmatrix} \begin{bmatrix} U_1 & & & & \widetilde{Q}_1 \\ & U_2 & & & \widetilde{Q}_2 \\ & & U_3 & & \widetilde{Q}_3 \\ & & & U_4 & \widetilde{Q}_4 \\ & & & & U_c \end{bmatrix} \quad (11)$$

where for $i = 1, \ldots, 4$

$$L_i U_i \approx A_i$$
$$\widetilde{S}_i = S_i U_i^{-1}$$
$$\widetilde{Q}_i = L_i^{-1} Q_i$$
$$L_c U_c \approx B - \sum_{i=1}^{4} \widetilde{S}_i \widetilde{Q}_i$$

How do we obtain the approximate factorizations L_iU_i for A_i and L_cU_c for the capacitance matrix? The approximate factors L_iU_i are obtained from the MNF method; the constraint equation is used to compute L_c and U_c.

Ordering

To avoid more than one domain solution per preconditioning step, the planes are given a two-way ordering depicted in Figure 2. Each domain is ordered from the center outward to the boundaries; however, in the two-domain case both domains are ordered from the outside toward the center boundary plane. The domain solutions involve only one forward solution and one backward solution.

Summary

The computational strategy for the pressure equation can be summarized as follows.

Preconditioning step

Domain decomposition

- ☐ Use strip method of domain decomposition.
- ☐ Subdivide grid into n domains and $n-1$ boundaries.
- ☐ Order domains first, then order boundaries.
- ☐ Use two-way ordering on domains.
- ☐ Factor domains approximately using modified nested factorization.
- ☐ Compute approximation to capacitance matrix.
  - a) Block tridiagonal structure and satisfying planar constraints or
  - b) Block diagonal structure
- ☐ Factor this boundary matrix approximately using modified nested factorization.

Matrix vector-multiply step

- ☐ Subdivide entire grid into N domains of approximately equal size (no boundaries).
- ☐ Perform partial matrix-vector multiply associated with each domain.

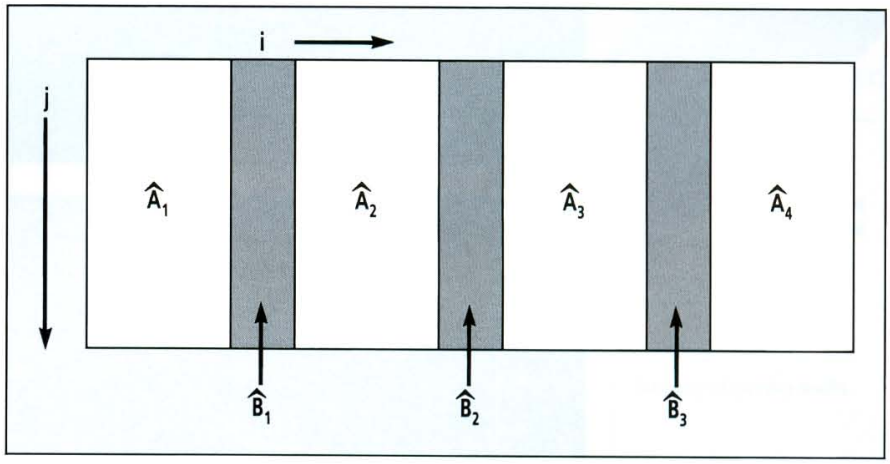


Figure 1. Division of reservoir for domain decomposition.

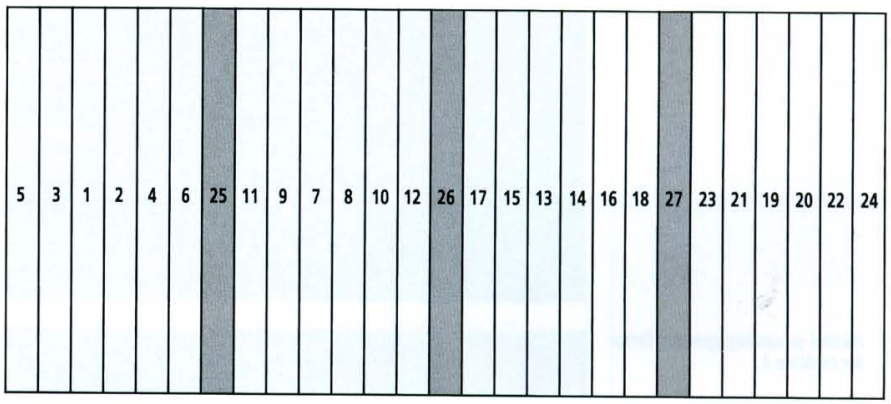


Figure 2. Two-way ordering scheme used in domain decomposition.

Acceleration step

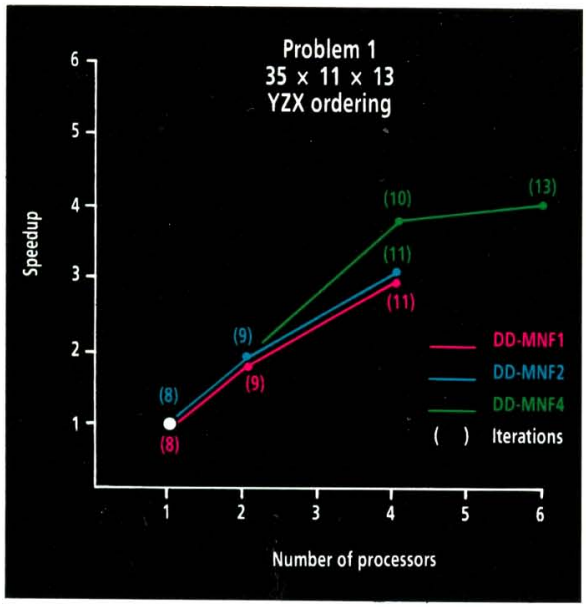
- ☐ Use new implementation of GCR.
- ☐ Subdivide entire grid into N domains of approximately equal size (no boundaries).
- ☐ Compute as many whole dot products per processor as possible.
- ☐ Compute partial dot products and linked triads (SAXPYs) associated with each domain.
- ☐ Accumulate partial dot products.

Results

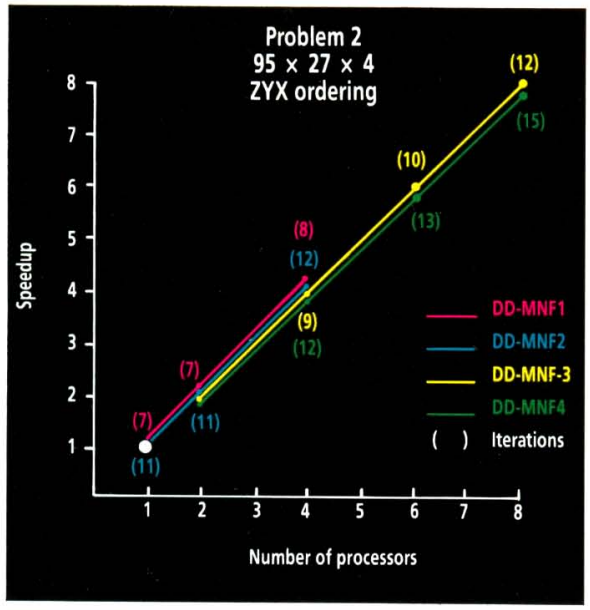
Four preconditioners were used on several test problems:

- ☐ DD/MNF1. This preconditioner subdivides the grid using domain decomposition. Domains are

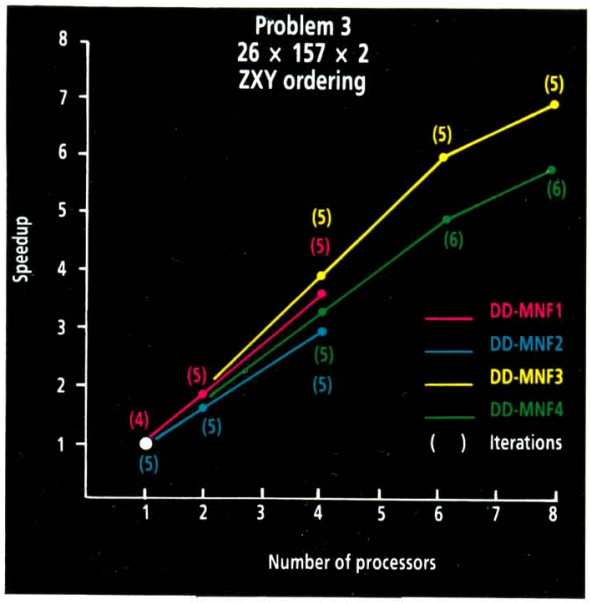
Parallel processing speedup factor for Problem 1.



Parallel processing speedup factor for Problem 2.



Parallel processing speedup factor for Problem 3.



ordered first — plane-by-plane using two-way ordering; then the boundary planes are ordered. Incomplete factorization is performed using modified nested factorization on both domains and coupled boundary submatrix (capacitance matrix). Modified nested factorization MNF1 uses nested factorizaton at the 3-D level and at the 2-D level and satisfies a planar row-sum condition rather than a planar column-sum condition.

- ☐ DD/MNF2. Same as DD/MNF1, except that incomplete factorization is performed using nested factorization on both domains and coupled boundary submatrix (capacitance matrix). Modified nested factorization MNF2 uses nested factorization at the 3-D level and red-black MILU(1) at the 2-D level.
- ☐ DD/MNF3. Same as DD/MNF1 except that boundary planes are not coupled.
- ☐ DD/MNF4. Same as DD/MNF2 except that boundary planes are not coupled.

Problem 1

This example first was presented as Problem 3 during the SIAM/SPE Linear Algebra Comparison Project held in January 1986. It is based on a field-scale model of an actual oil reservoir and its supporting aquifer. Two neighboring reservoirs are included in the model but are represented as three-phase aquifers. The reservoir grid has the dimensions 35 by 11 by 13, but only 2898 cells are active.

Problem 2

This example is based on the Third SPE Comparison Project.[6] The reservoir fluid is a retrograde condensate; the displacement mechanism is gas cycling. The grid of the original problem has been refined to the dimensions 95 by 27 by 4, and a fault has been introduced at $y = 9$.

Problem 3

This example is from a Middle Eastern water flood study of an undersaturated oil reservoir. The grid consists of two layers of 26 by 157 cells each for a total of 8164 active grid blocks.

Test results

Test cases of 5005, 8164, and 10,260 grid blocks were run on a dedicated CRAY Y-MP8/832 system using a convergence criterion of

$$\frac{\| \vec{r}_i \|_2}{\| \vec{r}_0 \|_2} < 10^{-3}$$

The results of the test problems show speedups ranging from 1.95 to 1.98 for two processors and from 3.1 to 3.9 for four processors. Six processors and eight processors for problems 2 and 3 yielded speedups ranging from 5.2 to 5.9, and 6.3 to 7.6, respectively. It is interesting to compare a four-processor, four-domain timing with one using one processor and one domain to obtain the effective speedup. In the case of problem 2 using DD/MNF2 (one domain) and DD/MNF4 (four domains) an effective speedup of 3.45 is obtained. For problem 3 the same comparison yields an effective speedup of 3.21. In addition, the algorithm can sustain peak performance of 850 to 900 MFLOPS on large test cases.

Observations and recommendations

Our results suggest several rules of thumb regarding orderings and numbers of domains.

Orderings

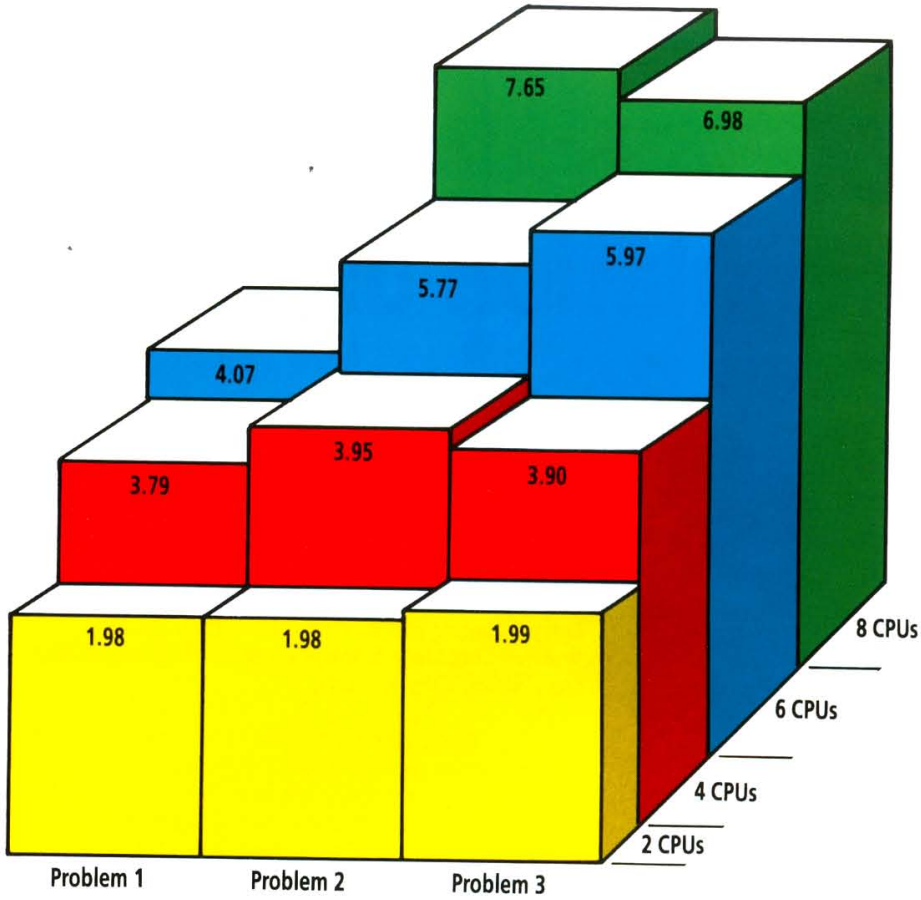
- ☐ To obtain the best vectorization, order the grid first along the largest grid dimension.
- ☐ To obtain the best parallelization, order the grid first along the two smallest grid dimensions and last along the largest grid dimension.
- ☐ To obtain the best convergence, order the grid first along the direction of largest transmissibility, then along the direction of next largest transmissibility, and finally along the direction of smallest transmissibility. Often this corresponds to ordering in ascending order of grid dimensions.

In view of these observations, orderings that are best for parallelism and rapid convergence typically do not yield the best vectorization. Pressure solution preconditionings DD/MNF2 and DD/MNF4 are recommended because their convergence is less sensitive to having the largest direction of transmissibility ordered first. These preconditionings generally require only that this direction be among the first two of ordering.

To obtain the best load balancing among processors, the grid dimension of the last direction to be ordered should be of the form $N * (M + 1) - 1$ where N is the number of domains (equal to the number of processors) and M is any positive integer.

Number of domains

The number of domains generally affects the rates of convergence, particularly for the preconditionings DD/MNF3 and DD/MNF4 (uncoupled boundaries). For these preconditionings, each domain should consist of at least eight planes. Boundary planes are more tightly coupled the closer they are to each other. Thus, the minimum size of the last ordered grid dimension for 4, 6, and 8 domains is 35, 53, and 71 respectively. From the standpoint of vectorization, the DD/MNF2 and DD/MNF4 methods exhibited the best results on the CRAY Y-MP system. The primary performance limitation on the system, memory contention, diminishes when going from the CRAY X-MP/416 system to the CRAY Y-MP/832 system. ▬



Summary of speedup results.

Acknowledgments

The authors thank ARCO Oil & Gas, British Petroleum, Convex, Cray Research, Digital Equipment, IBM, Petrobras, Statoil, and Texaco. Special thanks to Frank Kampe of Cray Research for his technical support and to Peggy Henry for her skillful manuscript preparation.

About the authors

J. R. Wallis has been involved in the research and development of numerical linear algebra software for more than 10 years. He developed linear solver software packages used in many commercial and proprietary reservoir simulators used by major oil companies. He received a Ph.D. degree in mathematics from the University of Houston.

J. A. Foster has worked in software design and optimization for more than 11 years, including vectorization and parallelization on Cray Research, Convex, Alliant, Control Data, and IBM systems. He received a B.S. degree in computer science from the University of Houston.

R. P. Kendall has 17 years of experience in the development and application of numerical reservoir simulators. This experience ranges from conducting major simulation studies for ARAMCO to the supervision of the development of commercial and proprietary reservior simulators. He received a Ph.D. degree in mathematics from Rice University.

References

1. Elman, H. C., "Iterative Methods for Large, Sparse, Nonsymmetric Systems of Linear Equations," Yale Research Report No. 229, 1982.
2. Appleyard, J. R. and I. M. Cheshire, "Nested Factorization," Paper SPE 12265 presented at the 7th SPE Symposium on Reservoir Simulation, pp. 315-324, 1983.
3. Watts, J. W., "A Method of Improving Line Successive Overrelaxation in Anisotropic Problems — A Theoretical Analysis," *Society of Petroleum Engineering Journal*, pp. 105-118, March 1971.
4. Wallis, J. R., R. P. Kendall, and T. E. Little, "Constrained Residual Acceleration of Conjugate Residual Methods," Paper SPE 13536, presented at the Eighth SPE Symposium on Reservoir Simulation, Dallas, Texas, 1985.
5. Killough, J. E. and M. F. Wheeler, "Parallel Iterative Linear Equation Solvers: An Investigation of Domain Decomposition Algorithms for Reservoir Simulation," Paper SPE 16021, presented at the Ninth SPE Symposium on Reservoir Simulation, San Antonio, Texas, February 1-4, 1987.
6. Kenyon, D. E. and A. Behie, "Third SPE Comparative Solution Project: Gas Cycling of Retrograde Condensate Reservoirs," Paper SPE 12278, Proceedings of the Seventh SPE Symposium on Reservoir Simulation, 1983.
7. Wallis, J. R., J. A. Foster, and R. P. Kendall, "A New Parallel Iterative Linear Solution Method for Large Scale Reservoir Simulation," Paper SPE 21209, presented at the eleventh SPE Symposium on Reservoir Simulation, 1991.

Large-scale reservoir simulation on Cray Research supercomputers

Larry C. Young and K. Hemanth Kumar,
Reservoir Simulation Research Corporation, Tulsa, Oklahoma
Gene Shiles, Cray Research, Inc.

Computer simulation is the only viable way to forecast the performance of petroleum reservoirs. Simulation of a reservoir requires that the reservoir be divided by a grid into a number of discrete elements. Prior to the work described here, a typical "large" reservoir simulation included about 50,000 grid points. But grids of this size are not sufficient to resolve many flow phenomena that are of interest in reservoir management. We describe here a problem with more than one million grid blocks. Practical run times were achieved by combining the efficiency of the MORE reservoir simulator with the computational power of a Cray Research supercomputer.

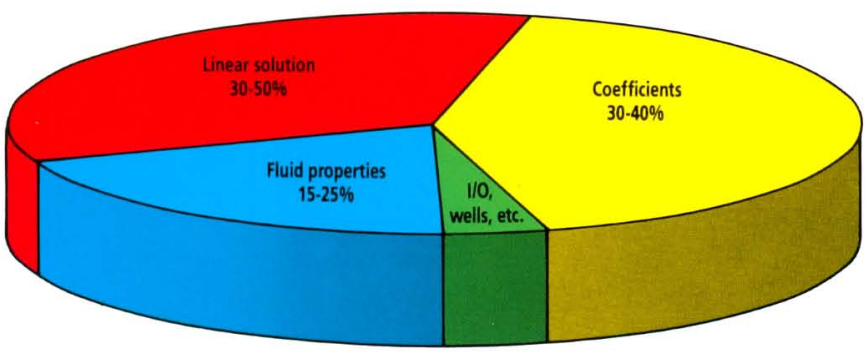
Petroleum reservoir simulation

Petroleum reservoir simulation uses numerical methods to predict the flow of fluids in a reservoir. This technique is used to examine various management and recovery strategies. It is particularly useful in choosing locations for drilling wells, maximizing production of gas or oil, and determining the parameters and evaluating the economics of various recovery methods.

Oil and gas are found in porous rock formations such as sandstone or limestone in which water also is present. These rock formations, called reservoirs, are located at depths of one to several miles and are typically tens of feet in thickness and thousands of feet in horizontal extent (often "pancake" shaped). Large reservoirs can be hundreds of feet thick and cover many square miles. To qualify as a reservoir, the rock layer must also be permeable; that is, fluid must be able to flow between the microscopic pore spaces.

Simulating flow in a reservoir requires the formulation of partial differential equations that describe mass and energy conservation and fluid flow in response to pressure differences and gravity. These nonlinear partial differential equations are discretized, using finite difference methods, and linearized. Initial and boundary conditions are imposed, and the unknown values of pressure and the amount of each component are calculated at increasing values of the discrete time levels, thus simulating the flow of fluids in the rock. The finite difference methods require that the reservoir be divided into grid blocks of constant porosity and permeability. Reliable simulations in large or heterogeneous reservoirs and two- or three-dimensional geometries can require many thousands of grid blocks. The fast computing speeds and large memories of supercomputers such as Cray Research computers are required for the practical solution of the resulting matrix equations for pressure and composition in each grid block.

Figure 1. Typical CPU usage in a black oil simulation.



The need for large-scale reservoir simulation

The black oil simulator is adequate for many reservoir studies and is the most widely used type of reservoir model. It accounts for three components: gas, liquid oil, and water. Gas may move in and out of solution in the oil in response to pressure changes. The compositional simulator provides a more general treatment in which several hydrocarbon fluid species with differing flow, thermodynamic, and chemical properties are described, and equations of state determine the fractions of each species that are liquid and gas at the local values of pressure and temperature. The compositional problem has much greater computational requirements than the black oil problem and is used when the black oil model is inadequate. The black oil simulator and problem described here are discussed in detail in Reference 1.

Supercomputers have made it practical to run larger black oil models than were practical to run 10 years ago. Full field models can represent global reservoir effects that cannot be represented using cross sections, average patterns, or small reservoir segments. High-resolution models have been useful for simulating in detail effects such as viscous fingering and the impact of small-scale heterogeneity. Even with supercomputers, the grid size for these problems has been limited by the available computational resources.

Full field models with 30,000 to 50,000 grid blocks have been useful in reservoir management. However, in large field studies, 50,000 blocks often are not adequate to allow more than one or two grid blocks between each adjacent pair of wells. Consequently, full field models frequently have used pseudorelative permeability functions to account for the lack of grid resolution. If 5 to 10 blocks could be used between each well pair, numerical dispersion could be reduced to acceptable levels, and the inherent uncertainties associated with pseudofunctions could be avoided.

The black oil model formulation reported here used the MORE reservoir simulator developed by Reservoir Simulation Research Corporation. This simulator seeks a balance between minimizing the number of calculations and maximizing the calculation rate. This approach allows the solution of a black oil problem with more than a million grid points in a reasonable computation time and with reasonable memory usage. The MORE simulator is based on a generalized compositional solution algorithm, which facilitates a modular program design and allows a large degree of flexibility within one basic simulator. The present work builds on the development of a high performance equation of state (EOS) compositional simulator.[2] Since the original simulator code used a generalized approach, development of a high-performance black oil simulator was accomplished by developing a black oil PVT module for the simulator.

Black oil computations

The major computational tasks in a black oil reservoir simulator are

- ☐ Fluid property calculations
- ☐ Coefficient calculations
- ☐ Linear equation solution

Figure 1 shows the breakdown of the computation time for a typical black oil simulation. For a black oil simulator, the solution of the linear equations usually takes somewhat more time than the fluid property or coefficient calculations. In contrast, for an EOS compositional simulator, 70 to 80 percent of the computational time is spent in the fluid property calculations, due to the need to evaluate equations of state for each component, in each grid block, for each time step. For the black oil simulator, all three steps are important and must be performed efficiently. Amdahl's Law illustrates an important requirement for the overall efficiency of a program. If F_{vs} is the ratio of the vector calculation rate to the scalar calculation rate, then the overall speedup relative to scalar calculation for an entire program is

$$F = \frac{1}{1 - f_v + f_v/F_{vs}} \qquad (1)$$

where f_v is the fraction of the calculation that will run in vector mode. For example, if the vector/scalar ratio, F_{vs}, is 10, then the overall speedup of codes that are 90, 95, and 98 percent vectorized will be 5.3, 6.9, and 8.5, respectively. A large portion of the code must be vectorized to approach the potential speedup of 10.

The calculation rate and the speedup due to vectorization are very important. However, another major factor is frequently overlooked. The total time to solve a problem is

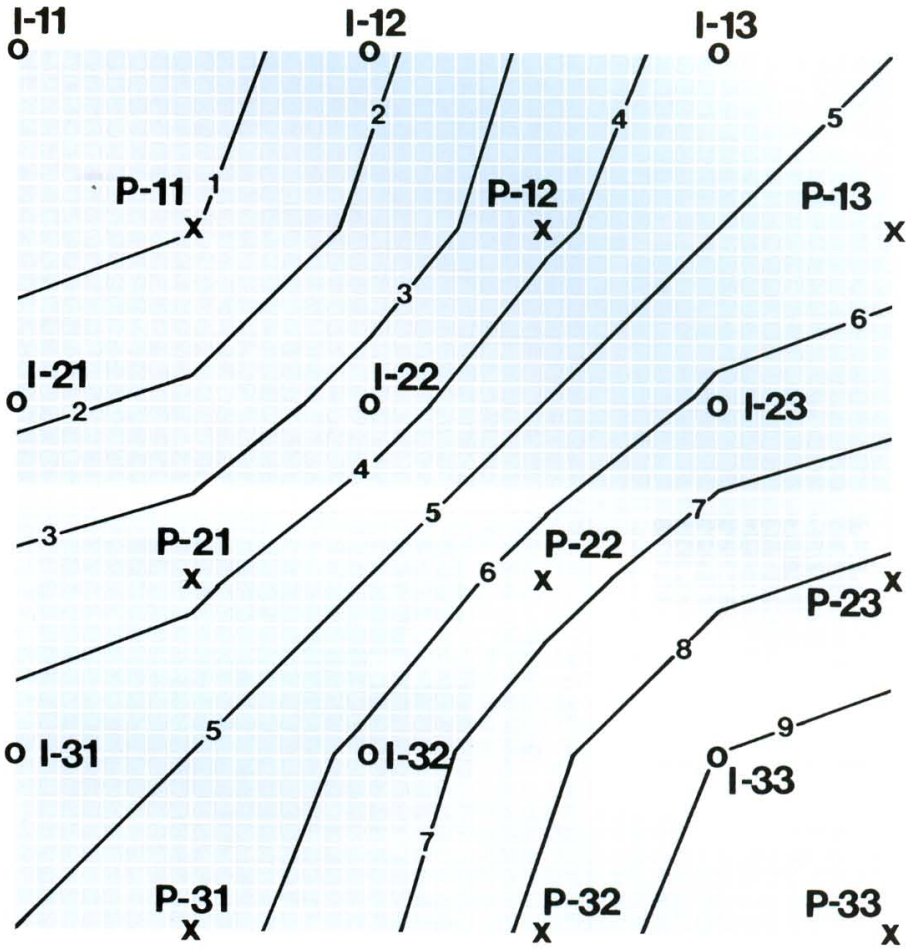
$$t = N_t/R \qquad (2)$$

where N_t is the total number of calculations and R is the average calculation rate. The time can be reduced not only by increasing the calculation rate, but also by minimizing the number of calculations. Although the preceding discussion emphasizes vector calculations, similar relationships apply for parallel and vector calculations.

Efficient computations also require that vectors be as long as possible (in a reservoir simulator, loops over grid blocks provide the only opportunity for long vectors). Minimization of conditional statements (IF conditions) and indirect addressing also are important. Memory requirements are an important consideration for very large problems. References 1 and 2 describe in detail how these considerations were handled in the work described here.

Test problem description

Several problems were designed not only to test the performance of the black oil simulator on large problems, but also to determine how the performance varies as the problem size increases. A series of waterflood problems, with a systematic variation in problem size, was created. In a waterflood, water is injected into one or more injection wells to help recover the oil at the production wells. A five-spot pattern often is used, where four injection wells form the corners of a square (in areal view), with a production well at the center. The smallest problem in this work represents a reservoir segment containing 6.25 80-acre five-spots (25 quadrants), with an 8-by-8 grid in each quadrant. Ten vertical layers, of varying thickness and permeability, were included. To make the problem realistic, areal permeability variations also were used. The porosity was kept at a uniform 20 percent. Figure 2 shows an areal view of the grid with the well placement and the distribution of total permeability-thickness.



Level	kh
9	4560.
8	4320.
7	4080.
6	3840.
5	3600.
4	3360.
3	3120.
2	2880.
1	2640.

x - production well
o - injection well

Figure 2. Well placement and areal variation of total permeability — thickness.

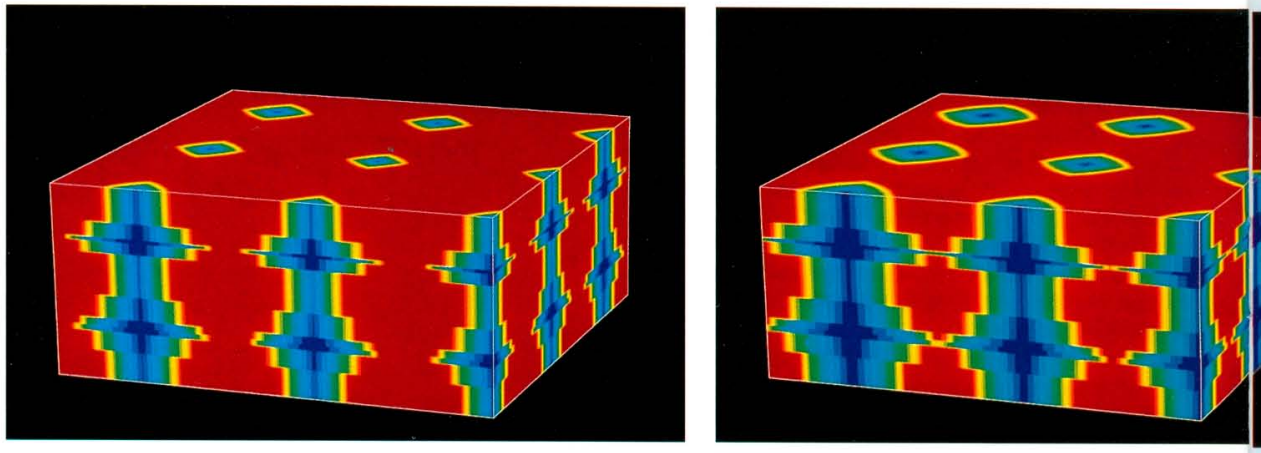
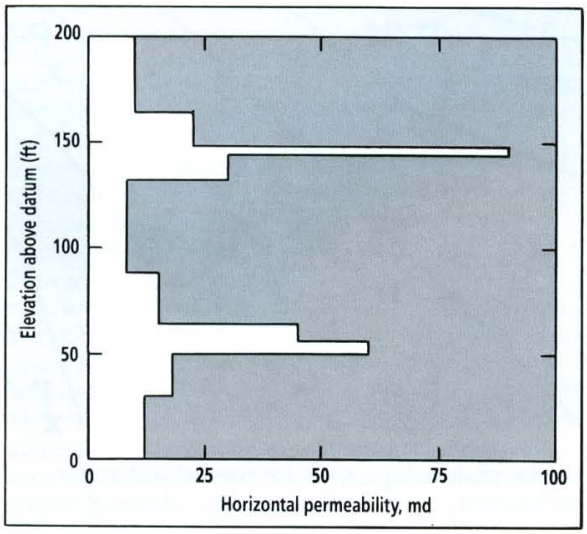


Figure 3. Vertical distribution of permeability (total kh = 3600 md-ft).



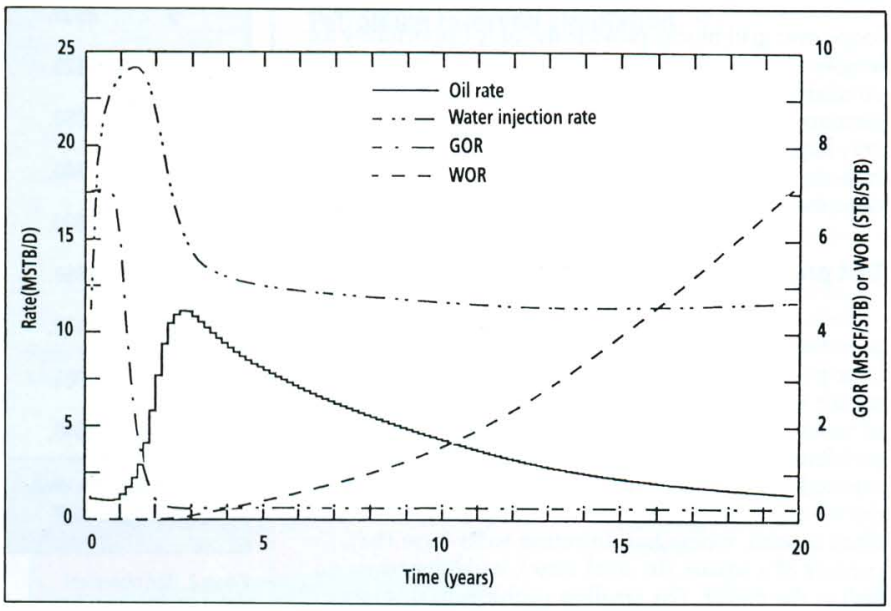


Figure 4. Calculated performance for test problem.

The total permeability-thickness varied by a factor of two. Figure 3 shows the vertical distribution of the permeability.

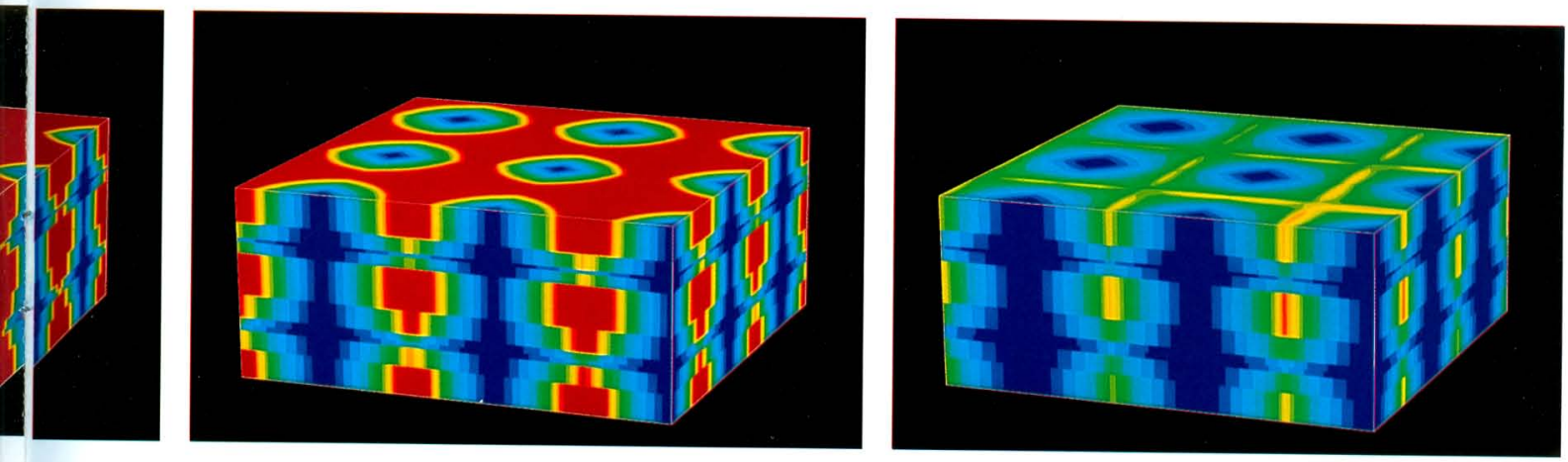
The initial reservoir pressure was 750 psi, and the initial gas saturation was 10 percent. These initial conditions represent a reservoir that has undergone some primary depletion (production utilizing the natural pressure in an undisturbed reservoir) prior to waterflooding. Water was injected throughout the 20-year simulation. The bottom hole pressures were maintained at 3000 psi for injection wells and 200 psi for producers. An effective wellbore radius of one foot was used for all wells. During the first two years of waterflooding, a gas phase was present. After two years, all the gas went into solution and a single liquid hydrocarbon phase was present. The overall performance of the segment is presented in Figure 4. Water saturation contours at 1, 2, 5, and 20 years are presented in Figure 5.

Larger problems were created systematically by treating the smallest problem as a symmetry element of the larger problems. The larger problems contained 4, 16, and 64 symmetry elements, respectively. Table 1 gives the overall size of the various problems. This method of expanding the problem simulates a reservoir study that considers a reservoir segment initially, then is expanded to model a full field.

Results

A 20-year simulation was calculated for all four problems. The problems were solved on a single processor of a 128-Mword CRAY Y-MP8 system. Table 1 shows that the simulator required only about 66 words of memory per grid point, so even the largest problem could be run in central memory. For every problem, the number of time steps was 244 and the number of Newton-Raphson iterations was 337. The total number of linear solver iterations varied from 4869 to 4874 and averaged about 20 per time step. Due to the manner in which the problem was expanded, a constant number of linear solver iterations was to be expected for the line successive over-relaxation method.[3] For the larger problems, the calculated results of the small problem were replicated in each symmetry element, as expected.

Using vector mode, the total computation time for each of the four problems was 0.44, 1.7, 6.6, and 26.1 minutes, respectively. When normalized by the number of grid points and time steps, these values correspond to times of 6.5, 6.3, 6.2, and 6.2 microseconds per grid point per time step, respectively. The decrease of the normalized times with problem size is consistent with a consideration of vector startup time. The variation is small because the vectors are long in all cases. For the CRAY Y-MP system, these results clearly show that the performance does not decrease with increasing problem size. We also ran the smallest



problem in scalar mode, and determined an overall vectorization of 98 to 99 percent.

For the smallest problem, detailed statistics were gathered with the aid of the hardware performance monitor on the CRAY Y-MP system. Table 2 shows the normalized computation time, normalized number of operations, and the computation rate for each of the major sections of the program. These statistics demonstrate that the simulator minimizes the number of calculations while maximizing the calculation rate. Roughly half of the floating-point operations were performed by the linear solver. The number of operations corresponds to about 18 per grid point per iteration of the solver. The linear solver also executed at the highest computation rate. The coefficient calculations and fluid properties required fewer operations, but ran more slowly. The lower computation rates reflect the amount of indirect addressing for table lookups and reordering for fluid property calculations.

The overall performance of the simulator is quite impressive. The short computation time is combined with the stability of a fully coupled, fully implicit well treatment. As a result, the relatively heterogeneous and finely gridded test problem was solved with time steps of about one month. The real proof of the approach was the solution of a full 20-year simulation with more than one million grid points. The Cray Research supercomputer used to run the tests has eight processors. Some preliminary parallel simulations with the Autotasking automatic parallel processing feature of Cray Research's CF77 Fortran compiling system have reduced the total computation time to just under 10 minutes. With further parallelization effort, we expect to reduce the computation time for this large problem to about 5 minutes. This work shows that the CRAY Y-MP system and the MORE simulator can solve very large problems in a reasonable time.

Figure 5. Water saturation contours at (left to right) one, two, five, and twenty years. Blue represents water.

Table 1. Size of test problems.

Grid	Number of blocks	Production wells	Injection wells	Memory (Mwords)
41 x 41 x 10	16,810	9	9	1.3
81 x 81 x 10	65,610	25	36	4.4
161 x 161 x 10	259,210	100	121	16.7
321 x 321 x 10	1,030,410	400	441	66.0

	Normalized time microsec/(gp)/(ts)	Normalized operations operations/(gp)/(ts)	Calculation rate (MFLOPS)
Fluid properties	1.44	101	70
Coefficients	2.01	241	120
Linear solvers	2.15	358	167
Other	0.90	25	28
Total/average	6.50	724	111

Table 2. Computational statistics for the 41-by-41-by-10 problem.

About the authors

Larry C. Young is president of Reservoir Simulation Research Corporation. Young received a B.S. degree in chemical engineering from Texas Tech University, and M.S. and Ph.D. degrees in chemical engineering from the University of Washington. He was a member of technical program committees for the 1984-88 annual meetings of the Society of Petroleum Engineers and program committees for the 1983 and 1987 SPE Reservoir Simulation Symposium. He also chaired the program committee for the 1985 Reservoir Simulation Symposium.

K. Hemanth Kumar is vice president of Reservoir Simulation Research Corporation. He has been involved in the area of reservoir simulation for the last six years. Kumar received a B. Tech. degree in chemical engineering from the University of Madras in India, and M.S. and Ph.D. degrees in chemical engineering from the University of Oklahoma. He is a member of the Society of Petroleum Engineers.

Gene Shiles is a senior applications specialist in the Industry, Science & Technology Department at Cray Research, Inc. He received his B.S. degree in physics from Northern Illinois University and Ph.D. degree in theoretical physics from the University of California at Santa Barbara. Prior to joining Cray Research in 1986, he spent six years on various computer applications in the petroleum industry. He is a member of the Society of Petroleum Engineers.

References

1. Young, L. C. and K. Hemanth Kumar, "High-Performance Black Oil Computations," Paper SPE 21215, presented at the SPE Symposium on Reservoir Simulation, Anaheim, California, February 17-20, 1991.
2. Young, L. C., "Full-Field Compositional Modeling on Vector Processors," *SPERE*, February 1991, pp. 107-114.
3. Varga, R. S., *Matrix Iterative Analysis*, Prentice-Hall, Inglewood Cliffs, New Jersey, 1962.

POSC welcomes Cray Research

Cray Research actively supports standards in the computer industry. This support includes participation in various standards organizations, helping to formulate and draft standards as user needs evolve. Cray Research understands the value of standards in creating efficient and transparent computing networks, and has made support of such environments a strategic objective. Within the petroleum industry, Cray Research has extended its support for networking by becoming an active member of the Petrotechnical Open Software Corporation (POSC). The goal of POSC is to minimize or eliminate many information management problems by providing common data-application, software-hardware, and user-application communications. Following, POSC president Dan Turner describes the specific objectives of the organization.

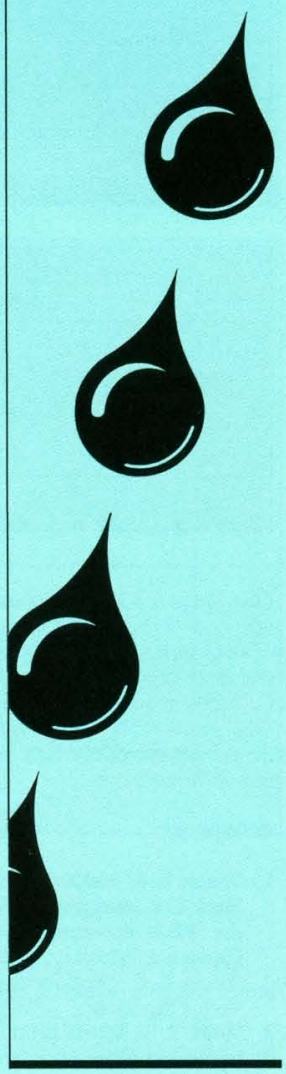
Dan Turner,
Petrotechnical Open Software Corporation,
Houston, Texas

The Petrotechnical Open Software Corporation is proud to welcome Cray Research as the first computer manufacturer to join the organization as a member. Cray Research's membership in POSC reflects the company's commitment to the petroleum industry and its continuing commitment to standards and open systems. Cray Research long has been a supporter of open systems, as demonstrated by the company's continuing enhancement of the UNICOS operating system, networking products, and other technologies that support industry standards.

POSC is a not-for-profit organization incorporated in 1990 to develop an industry standard, open systems software integration platform for upstream (exploration and production, or E & P) technical applications in the petroleum industry. The platform POSC is developing will serve as an interface for petrotechnical software applications, database management systems, workstations, and users. Specifically, POSC is developing an integrated E&P data model, an application program interface for data access, a common "look and feel" user interface, and a set of test suites that will enable software developers to evaluate their offerings against selected industry standards.

The founding sponsors of POSC include BP Exploration, Chevron Corporation, Société Nationale Elf Aquitaine, Mobil Corporation, and Texaco Inc. Membership has grown to include, along with Cray Research, other petroleum companies, major E&P service companies, large and small software/hardware vendors, and research institutions.

With the advent of sophisticated data modeling tools, good database management systems, and a desire to solve technical and business problems using an interdisciplinary approach, the upstream businesses have a new and more focused incentive to solve their chronic information management problems. As a result, the industry is looking to the open systems environment to help take advantage of the opportunities appearing throughout the computing world. The rapid development of high-performance computing hardware, the acceptance of UNIX operating systems, distributed computing across wide and local area networks, and the evolution of a standard graphics user interface provide a basis to enable engineers and scientists to concentrate on the problems of the oil business rather than on the problems of data processing. An open systems environment will facilitate the efficient application of these new technologies.



The currently defined offerings of POSC will be licensed and distributed to all interested organizations. The offerings generally will be packaged as a set of releases and delivered over a three-year timeframe. These releases are grouped as follows:

- ☐ *Base Standards* will provide the E&P industry with a definition of a base computing Application Environment Profile, including relevant open computing system standards and selected E&P data exchange and subroutine call standards.
- ☐ The *E&P Data Model* will provide the E&P industry with a conceptual and detailed logical data model.
- ☐ *Data Access* will contain the Application Program Interface for data access.
- ☐ The *User Interface* will contain the E&P User Interface Style Guide for multidiscipline and multivendor applications.
- ☐ The *Development Toolkit* will include an extended library of basic productivity tools that will aid the E&P applications developer.

Since its incorporation, the pace at POSC has been rapid. Project teams for the Base Standards, E&P Data Model, and the Data Access releases are currently in place. Staffing is underway for the remaining projects. Two special interest groups (SIGs) have been formed: the Data Model SIG and the Base Standards SIG. The first POSC Request for Technology (RFT), for the E&P data model, has been published and sent to the industry at large. A Request for Comment (RFC) on base standards also has been published. Additional SIGs, RFTs, and RFCs are planned for 1991.

It is exciting to note that Cray Research, the world leader in supercomputers, is the first computer manufacturer to join POSC. Much of the impetus for supercomputer power historically has come from the seismic and reservoir simulation requirements of the E&P segment of the petroleum business. Through its membership, Cray Research will become more directly involved in helping the E&P industry apply and benefit from the latest computing and networking technologies. ▬

About the author

Dan Turner was named president and CEO of POSC when it was founded in November 1990. He has been involved in technical computing for the upstream functions of the petroleum industry for over 35 years, having been Division Manager, Information Systems, at Exxon Production Research Company and Director, Information Systems for BP Exploration. He holds degrees in electrical engineering and mathematics from the University of Houston.

Interactive device simulation at Toshiba

Yuichiro Yoshida, Akio Oka, and Kazutoshi Iizuka,
Toshiba Corporation, Kawasaki, Japan

Rapid advances in semiconductor technology have increased the importance of device simulation for chip manufacturers such as Toshiba. As the size of VLSI (very large scale integration) devices has become increasingly microscopic, it has become more difficult to define the electrical characteristics of semiconductor devices through conventional experimentation methods. For this reason, Toshiba relies extensively on an interactive device simulation system that runs on its CRAY Y-MP computer system.

A scientific visualization system for supercomputers

One result of the recent efforts made by the information system division at Toshiba to assist the semiconductor division is an interactive device simulation system, developed by the authors, that provides Toshiba engineers and researchers with a high-performance distributed system. In Toshiba's interactive device simulation system, the CRAY Y-MP supercomputer, installed at Toshiba's Semiconductor Design Center, running under the UNICOS operating system, communicates with a graphics workstation through a high-speed network. The system allows raw simulation data to be transferred instantaneously to the graphics workstation, providing the researcher with immediate visualization of the transferred data.

The interactive device simulation system developed at Toshiba allows concurrent interactive job control of the device simulator on the CRAY Y-MP system and real-time graphic representation of simulated results. The interactive device simulation system thus provides a powerful tool for visual simulation that enables engineers and researchers to develop new semiconductor devices quickly and efficiently and may offer a substitute for physical experiments.

Numerical simulation using the supercomputer is one of the key technologies used today for the speedup of research and new product development. Along with the supercomputer, a scientific visualization facility is indispensable for a quick understanding of the massive amounts of numerical results obtained from a simulation. Rosenblum has defined the stages of scientific visualization systems as postprocessing, tracking, and steering.[1] In the systems recently developed by supercomputer manufacturers for scientific visualization on their machines, the supercomputer carries out both the numerical simulation and graphical image processing from the simulation data. Only graphical images are transferred to a graphic display such as a frame buffer through a high-speed network.

These systems can perform both postprocessing and tracking. Tracking is used for nonlinear or time-dependent problems, whose simulation data are calculated step by step. In tracking, processing and transfer of graphical images follow each numerical simulation step. By using this method, the simulation can be carried out while monitoring the calculated results simultaneously. For the visualization of large-scale problems in fields such as computational fluid dynamics involving 200,000 to 300,000 nodes, these systems certainly provide the prerequisite foundation of power and speed.

A job is submitted in either batch or time-sharing mode, allowing quick and interactive transformations of graphical images. MPGS, the Multipurpose Graphic System developed by Cray Research, uses the graphics workstation to create images from the simulation data, supporting interactive operations of graphical images.

In the interactive simulation system, raw simulation data is transferred to the graphics workstation instantaneously through the highest-speed network to allow graphical images to be processed by the workstation locally. The interactive transformation of graphical images by rotation, zoom, and translation,

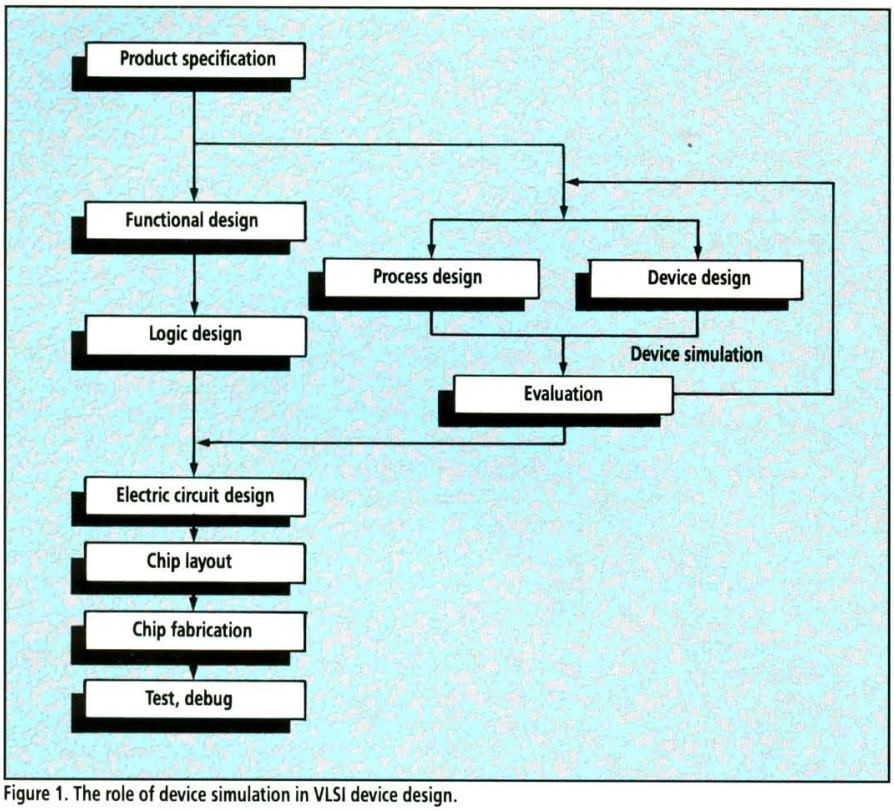


Figure 1. The role of device simulation in VLSI device design.

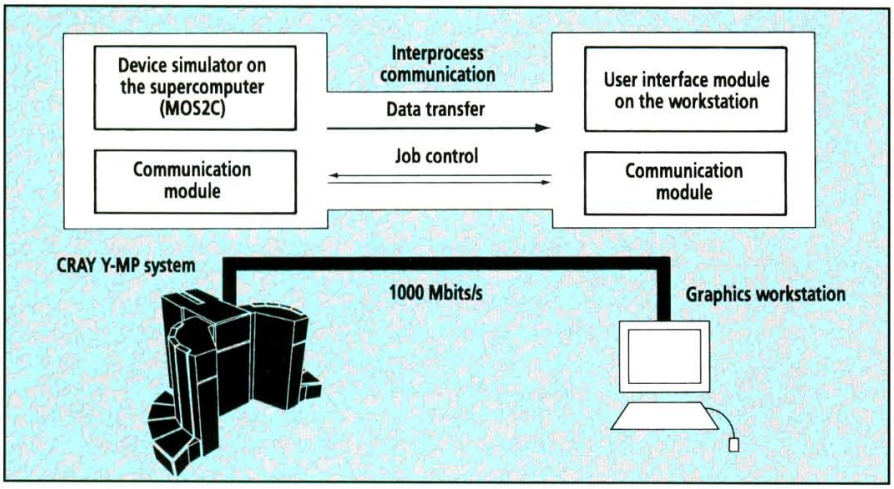


Figure 2. Software and hardware of the interactive device simulation system.

for example, can be performed quickly and easily on the interactive simulation system. One of the most attractive features of an interactive simulation system is the instantaneous verification of the computed results by graphical images, and the ability to control or steer the simulator on the supercomputer at the same time. The operating system on the CRAY Y-MP computer system and the graphics workstation allow interactive control of the simulator on the Cray Research system through the use of interprocess control facilities.

Problem size limitations have not yet been reached. The interactive simulation system is very efficient when used to perform parametric studies and to check the influence of parameters in the analysis of complex problems. The interactive simulation system provides researchers with a very sophisticated, yet easy-to-use analytical tool in a high-speed, network-based supercomputing environment. Phillips,[2] Crabel,[3] and Haber[4] also have reported solutions which incorporate interactive visualization systems in their network-based supercomputing environments using Cray Research supercomputers.

Device simulation

Device simulation is performed to evaluate the electrical characteristics of a semiconductor device.[5,6] The role of device simulation in the design of a VLSI device is shown in Figure 1. Unknown variables in device simulation include potential, electron density, and hole density. The fundamental equations involved in the simulation consist of drift-diffusion equations. For device simulation, the authors use MOS2C, a two-dimensional, two-carrier (electron and hole) device simulator developed at Toshiba.[7] This code uses the finite difference method for discretization of fundamental equations, employing the Shafetter-Gummel scheme in the finite difference discretization. To solve discretized nonlinear equations, decoupled and coupled methods are used. MOS2C is highly vectorized to achieve very good performance on the supercomputer.

Hardware and software configuration

Figure 2 shows the hardware and software contained in the interactive device simulation system. The CRAY Y-MP supercomputer system is connected to a graphics workstation through a high-speed network operating at 1000 Mbits/sec. The software for the interactive device simulation system consists of the following four modules:

- ☐ The device simulator MOS2C on the CRAY Y-MP supercomputer
- ☐ The communication module on the CRAY Y-MP system, which accepts control commands to steer MOS2C from the graphics workstation and sends computational results to the graphics workstation
- ☐ The communication module on the graphics workstation, which sends control commands to the CRAY Y-MP system to steer MOS2C and accepts computational results from the CRAY Y-MP system
- ☐ A user interface and graphic module on the workstation

The interactive device simulation system supports real-time graphic representations of simulation data. Visual results include three-dimensional stereographs, two-dimensional contours, vector plots, doping

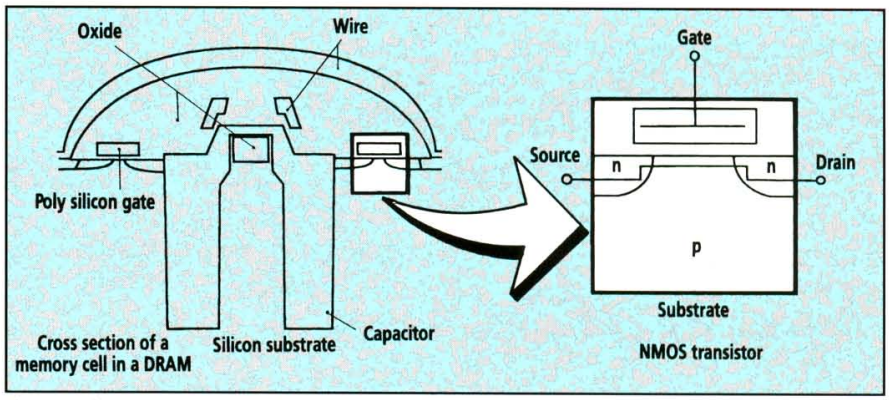
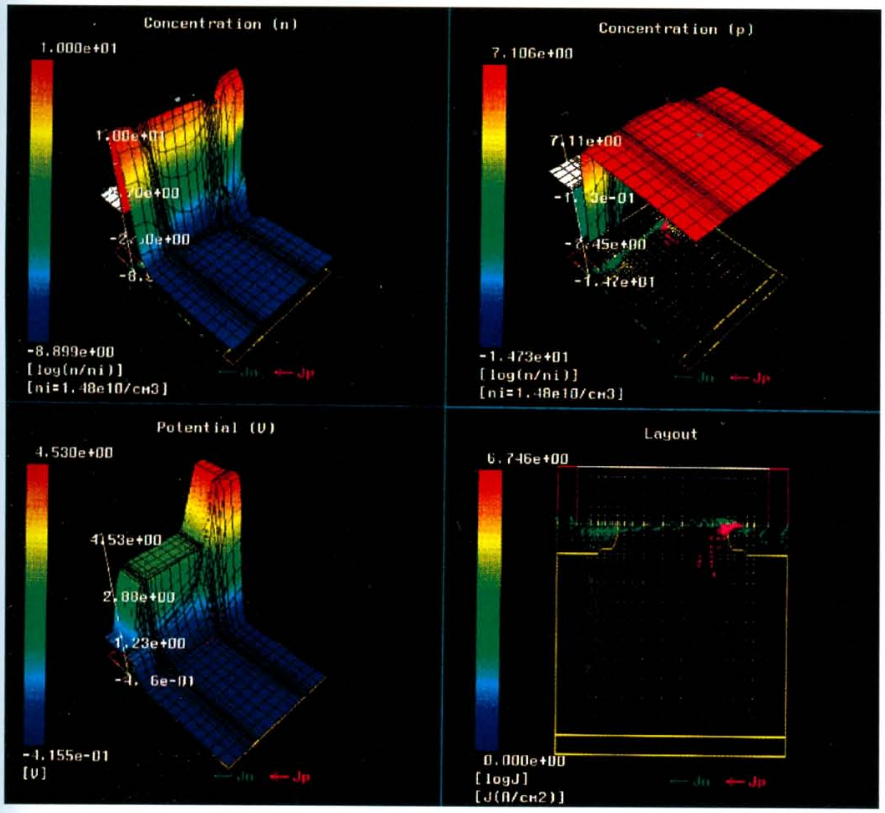
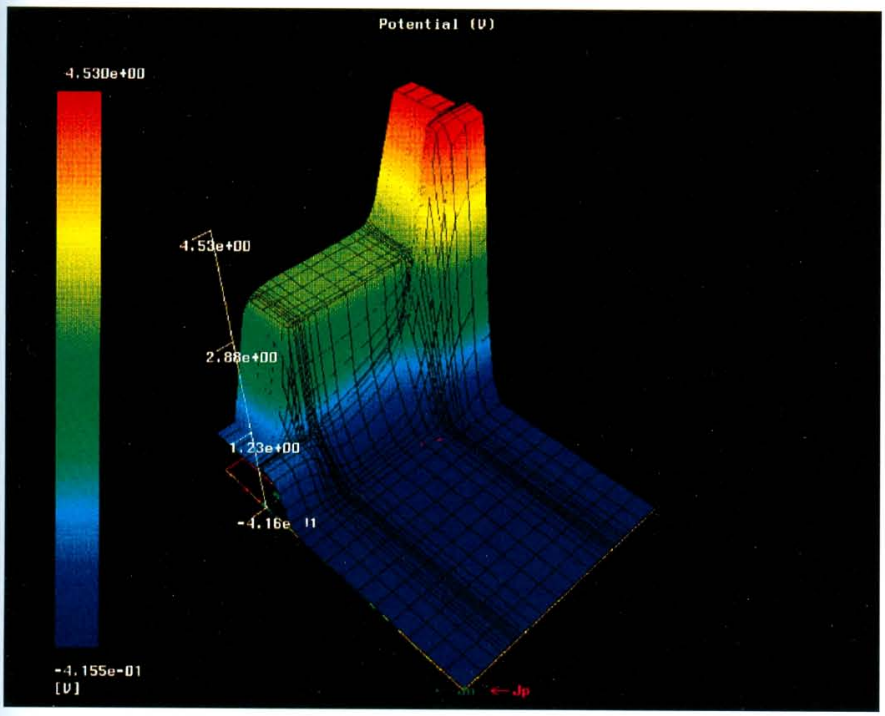


Figure 3. Schematic of a memory cell in a DRAM.





profile graphs, terminal current voltage diagrams, and convergence monitor diagrams. On the workstation, the graphics window can be divided into four portions to display graphics for 13 simulation parameters. Interactive control of the device simulator allows commands such as "pause," "restart," and "stop." After the simulation has been completed, an animation of the results, hard copy of the graphic images in color, or a video recording of the results using a real-time scan converter and a video controller can be obtained.

Figure 4 (top). The main window of the interactive simulator divided into four portions, clockwise from upper left: electron distribution, hole distribution, vector plot, and potential distribution.

Figure 5 (bottom). Potential distribution in the NMOS transistor.

An NMOS transistor in a DRAM

Figure 3 shows a memory cell in a DRAM. The NMOS transistor works like a switch in the memory cell. In this analysis, voltages on gate and drain are controlled. To evaluate static characteristics between terminal voltages and terminal currents of the NMOS transistor, voltages are increased incrementally. The graphical results are displayed in Figure 4. Figure 5 is a stereograph of the potential distribution.

The elapsed time for this analysis, which includes time for transfer of all simulation results and time for graphics, was only one minute compared to hours with conventional mainframe-based postprocessing systems. While parametric studies are carried out using MOS2C, only minimal effort is necessary for data modification. Therefore, the interactive device simulation system offers an efficient method to evaluate the influence of parameters very quickly. The authors hope that the interactive device simulator will enable designers and researchers at Toshiba to gain a better understanding of computed results through visualization and thereby help to shorten design cycles. ■

Acknowledgments

The authors thank Mr. Wada, manager of the ULSI research center; Mr. Kinoshita, manager, and Mr. Kobori of the semiconductor division; and Mr. Nakakubo and Mr. Fukui, manager, of the Information System Division at Toshiba Corporation for their technical assistance and helpful discussions. The authors express their appreciation to Mr. Morota, system engineer at Toshiba CAE Systems, Inc., for his dedicated work in the development of the graphics module.

About the authors

Yuichiro Yoshida is the deputy manager of the information systems division at Toshiba Corporation working on device simulation. He received his M.E. degree from Yokohama National University.

Okio Oka is the system engineer of the information systems division at Toshiba, working on computer graphics. He received his B.S. degree from Waseda University.

Kazutoshi Iizuka is the manager of the information systems division at Toshiba. He is responsible for computer-aided engineering systems in electrical engineering, working on logic simulation and layout systems. He received his B.S. degree from the University of Tokyo.

References

1. Rosenblum, L. J., "Scientific Visualization at Research Laboratories," IEEE Computer Society, August, pp. 69-70, 1989.
2. Phillips, R. L., "Distributed Visualization at Los Alamos National Laboratory," *ibid.*, pp. 70-77.
3. Crabel, B. and L. H. Carol, "Visualization Tools at Lawrence Livermore National Laboratory," *ibid.*, pp. 77-84.
4. Haber, R. B., "Scientific Visualization and the Rivers Project at the National Center for Supercomputing Applications," *ibid.*, pp. 84-89.
5. Kurata, M., *Numerical Analysis for Semiconductor Devices*, Lexington Books, 1982.
6. Dang, R., editor, *Process Device Simulation Technique*, Sangyo Tosho, 1988.
7. Wada, T. and R. Dang, Proceedings of NASECODE 4 Conference, p. 108.

CORPORATE REGISTER

100th CRAY Y-MP system installed at Pfizer

Pfizer Inc., the fifth largest pharmaceutical company in the United States, has installed its first supercomputer, a CRAY Y-MP2E system. The new system, installed at Pfizer Central Research in Groton, Connecticut, is the 100th CRAY Y-MP system installed by Cray Research at a customer site. Pfizer scientists plan to use the system to model potential drug molecules and simulate the interactions of these agents with the biological targets believed to be associated with diseases. Pfizer's new supercomputer will be used by scientists located in Groton and, through a high-speed link, by scientists at the company's laboratories in Sandwich, England.

Apple Computer has ordered a CRAY Y-MP2E supercomputer to be installed at Apple's main campus in Cupertino, California. The new system will replace Apple's first Cray Research supercomputer, a CRAY X-MP system installed in 1986. The new supercomputer will be used for heat flow analysis to simulate and determine the flow of heat, air, and dust particles within computer cabinets, speech recognition analysis in the development of a speech recognition system for future computer products, and analog circuit simulation for circuitry design. Apple also will use the system to streamline its manufacturing processes, the majority of which are automated and run by Apple personal computers.

The Centro Supercalcolo Piemonte (CSP), a new supercomputer center located in Turin, Italy, has ordered a CRAY Y-MP2E system. The center is funded by and will serve a consortium that includes the University of Turin, the Polytechnical Institute of Turin, and private industry in the Piemonte region.

Stichting Academisch Rekencentrum Amsterdam (SARA), the Netherlands national supercomputer center, has ordered a CRAY Y-MP supercomputer for academic research in science and engineering. Cray Research also has established research partnerships with SARA and the National Computing Facilities (NCF), a foundation in the Netherlands that promotes the use of computation in the sciences. The research partnership will provide $1.75 million to academic researchers in the Netherlands over the next four years. $750,000 will be given to SARA and devoted to research on operating system software, and $1 million will be administered by NCF to support the advancement of new supercomputer applications in various scientific fields, including economics, chemistry, and physics.

Universidad Nacional Autonoma de Mexico (UNAM) has ordered a CRAY Y-MP4/432 system, the first Cray Research system to be installed in Latin America. UNAM is one of the world's largest universities, with approximately 270,000 students on 19 campuses located throughout Mexico. The system, located at the university's new supercomputer center on UNAM's main campus in Mexico City, will be one of the world's 10 largest supercomputers installed at a university. UNAM has 23 scientific institutes and research centers, each devoted to the study of a different discipline. Eighteen

of the institutes are located in Mexico City, three in Cuernavaca, and two in Ensenada. The CRAY Y-MP system will be used primarily by researchers at the institutes and centers; researchers at Cuernavaca and Ensenada will connect to the system through a satellite link. The CRAY Y-MP supercomputer will support research in chemistry, structural engineering, astronomy, physics, and environmental science. The system will be used to simulate earthquake shock on earthen dams and skyscrapers, and model the pollution in Mexico City.

Eidgenössische Technische Hochschule (ETH), a polytechnical institute located in Zürich, Switzerland, has ordered a CRAY Y-MP supercomputer to replace its CRAY X-MP system. ETH will use the CRAY Y-MP system for general academic research, including mechanical engineering, electronics, and computational chemistry. ETH also ordered Cray Research's UniChem interactive chemistry software environment.

The leading computer services bureau in the Netherlands, **Informatica Centrum Voor Infrastructuur en Milieu (ICIM),** has ordered a CRAY Y-MP2E supercomputer for ICIM's headquarters in Rijswijk. ICIM provides high-level services for environmental studies, major public works projects, and traffic analysis in the Netherlands. ICIM will use its CRAY Y-MP2E system for a variety of tasks, including monitoring and forecasting the effects of transportation on the environment, supporting various civil engineering projects, and analyzing and predicting changes in the waterways and coastline in the Netherlands. Because much of the Netherlands is below sea level, the country is particularly concerned with coastal protection. To this end, the CRAY Y-MP2E supercomputer will be used to model and predict the effects of the North Sea tidal propagation and to study the impact of human activities on water movement at sea and in estuaries and rivers.

The German Climate Research Center (DKRZ), the primary climate research facility for united Germany, has ordered a CRAY Y-MP4E system for the organization's facilities in Hamburg, Germany. The new supercomputer will be networked to an existing CRAY-2 system DKRZ installed in 1987. The CRAY Y-MP4E system will be used primarily to conduct global climate research and analyze weather-related environmental issues such as global warming.

Fiat Auto S.P.A. has ordered a CRAY Y-MP4 system for the company's facility in Turin, Italy. The new system replaces a CRAY X-MP/14se system that was installed in 1989. The new CRAY Y-MP supercomputer will be used for automotive research and design, including structural analysis, computational fluid dynamics, and crash simulations.

Isuzu Motors, Ltd. in Tokyo, has ordered a CRAY Y-MP2E supercomputer and an SSD solid-state storage device.

Isuzu, a new customer for Cray Research, plans to use its CRAY Y-MP2E system to streamline its computer-aided engineering environment to better use computer simulation and analysis in its design and engineering processes. The company's specialty is the design and production of buses, trucks, and other large vehicles; Isuzu will use its Cray Research system for applications such as structural analysis, crash simulation, and engine combustion analysis.

Marion Merrell Dow Inc. has ordered a CRAY Y-MP2E system for its Research Institute headquarters in Cincinnati, Ohio. The company will use its new CRAY Y-MP2E system for macromolecular modeling of enzymes and protein receptors in the design of pharmaceutical products. The CRAY Y-MP2E system in Cincinnati will be shared with company researchers at labs in Strasbourg, France; Gerenzano, Italy; and Kansas City, Missouri, the company's global headquarters. The supercomputer also will be networked to high-performance graphics workstations located at the Cincinnati facility via a Fiber Distributed Data Interface (FDDI) local area network. FDDI provides 100 megabits (one million bits) of data transfer per second and allows for high-speed visualization of computational chemistry problems.

Cray Research offers Cray Visualization Toolkit 1.0

Continuing its commitment to industry standards, Cray Research now offers Cray Visualization Toolkit (CVT) 1.0. This toolkit enables users to run applications on Cray Research systems through their workstation interfaces. As a result, most applications that run on Cray Research systems have the same look and feel as the most common workstation environments. With a simple binary code installation, CVT delivers

- ☐ Release X11R4 of the X Window System
- ☐ Sun Microsystem's XView Toolkit (OPEN LOOK)
- ☐ Open Software Foundation's (OSF) Motif 1.1 Toolkit
- ☐ Silicon Graphics, Inc., Distributed Graphics Library (DGL)

These tools allow Cray Research applications to generate graphics and user interfaces easily, making Cray Research systems even easier to use and making users more productive.

With the X11R4 release of the X Window System, Cray Research supports diverse client applications, allowing users to run many X11R4 applications on their Cray Research systems from any workstation running the same version of the X Window System. It also provides the platform for the XView and OSF/Motif toolkits. These powerful software combinations create a seamless environment from workstation to Cray Research system.

By delivering the XView Toolkit, CVT provides software developers with the tools to create applications consistent with the OPEN LOOK Graphical Interface. OPEN LOOK is the graphical interface supported by UNIX System V release 4. Because XView is based on Sun Microsystems' SunView Window system, many popular applications are already available under XView.

By providing the OSF/Motif 1.1 Toolkit, CVT delivers a stable and widely available applications programming interface. Based on X11R4 Intrinsics, a toolkit framework provided with the X Window System, the OSF/Motif Toolkit provides application developers with compliance for X Window System applications and systems. The User Interface Language, an OSF/Motif component, also is included. This is an application development tool that allows rapid user interface design and prototyping. It simplifies the description and maintenance of user interfaces by allowing programmers to create the characteristics of the application interface independent of the application code itself.

By providing the Silicon Graphics Distributed Graphics Library (DGL), CVT delivers interactive 3-D graphics to the workstation while an application is running. This allows users to rotate, transform, and zoom the 3-D image. They also can change parameters during a simulation and conclude a simulation as soon as the relevant effect is observed. This level of visualization provides scientists and engineers with an extremely powerful and productive research tool.

CVT delivers a new level of productivity with the greatest possible portability, interoperability, and scalability of applications available today. Users can transfer their skills and experience from their workstations to Cray Research systems to increase productivity and the ability to solve the world's most challenging problems.

To use the Cray Visualization Toolkit 1.0, users must be running the UNICOS operating system version 5.1 or later and have a workstation. For more information, refer to the *Cray Visualization Toolkit (CVT) 1.0 Release Notice and System Installation Bulletin,* CV-1.00-UAN-RN, or contact your Cray Research representative.

APPLICATIONS UPDATE

MORE reservoir simulator available on Cray Research systems

The MORE reservoir simulation package from Reservoir Simulation Research Corporation is a comprehensive reservoir simulation system that runs on Cray Research computer systems under the UNICOS and COS operating systems. The MORE family of reservoir simulation tools includes

- ☐ MORE-EOS (Equation of State)
- ☐ MORE-BLACK OIL
- ☐ MORE-POST PROCESSOR
- ☐ MORE-GRAPHICS
- ☐ MORE-PVTx

The MORE-BLACK OIL and MORE-EOS simulators solve many types of problems easily and efficiently, including

- ☐ Dry gas
- ☐ Gas/oil
- ☐ Hydrocarbon/water
- ☐ Three-phase black oil
- ☐ Fully compositional problems

Using MORE, a reservoir can be described by flexible gridding in up to three dimensions, and fluids can consist of up to three phases and any number of components. Partitioning of the components between the oil and gas phases is governed by pressure-dependent K-values or a generalized two-parameter cubic EOS. The simulator is suited for many types of reservoir conditions and recovery methods, including

- ☐ Depletion, waterflooding, or immiscible gas injection in conventional black oil reservoirs
- ☐ Primary depletion of volatile oil or gas condensate reservoirs
- ☐ Injection of water or gas (hydrocarbon, nitrogen, or carbon dioxide) into volatile oil or gas condensate reservoirs
- ☐ Enhanced recovery from oil reservoirs by carbon dioxide or hydrocarbon gas injection

MORE is based on a generalized compositional algorithm. This method provides the framework for the seamless integration of the BLACK OIL and EOS capabilities. This formulation also prevents material balance errors from occurring during simulation. MORE provides flexible gridding options in one, two, or three dimensions. Two-dimensional nine-point and three-dimensional eleven-point difference options are available for adverse mobility ratio displacements.

Region arrays support the variation of the following quantities:

- ☐ Rock types
- ☐ PVT tables (black oil)
- ☐ Capillary pressure multipliers, such as Leverett J-Function
- ☐ Fluid-in-place and material balance
- ☐ Initialization

Equilibrium or nonequilibrium initialization may be selected. Initial composition gradients can be specified easily within one or more equilibration regions.

For the EOS option, phase behavior matching and parameter tuning can be obtained from the stand-alone program, MORE-PVTx, which features state-of-the art simultaneous regression for multiple fluids. Many commercial preprocessing or mapping programs can be used to provide input data for the simulator.

The latest release of MORE includes several enhancements:

- ☐ Adaptive implicit methods
- ☐ Horizontal wells
- ☐ Tubing head pressure
- ☐ Ability to split liquid from separator gas
- ☐ Fully integrated TECPLOT interface
- ☐ CO_2 solubility in water (MORE-EOS)
- ☐ Volume translation (MORE-EOS)
- ☐ Compaction options

For more information about using the MORE simulator on Cray Research systems, contact Sue Trippet, Reservoir Simulation Research Corporation, 2525 East 21st Street, Suite 205, Tulsa, OK 74114-1747; telephone: (918) 742-4330; or Gene Shiles, Cray Research, Inc., 655-E Lone Oak Drive, Eagan, MN 55121; telephone: (612) 683-3680.

Interactive DISCO available on Cray Research systems

IDISCO is an interactive seismic data processing system from CogniSeis Development, Inc., that operates in conjunction with the batch version of the DISCO processing system. IDISCO improves the quality and efficiency of seismic processing by combining the computational power of Cray Research supercomputers with the display flexibility and extensive user control available on workstations that run the X Window System interface. IDISCO is written for distributed processing; it harnesses the power of Cray Research supercomputers for accurate seismic processing in interactive environments.

IDISCO enables geophysicists to perform the iterative analysis needed for velocity analysis and many other seismic processing tasks. The core seismic algorithms, including FFT, convolution, and NMO, are optimized and highly vectorized to take advantage of the architecture of Cray Research systems. IDISCO uses the Motif graphic user interface with the X Window System. The program lets users run multiple programs or processes simultaneously and work with them in separate screen windows. The general operation of the window system is controlled by the Motif Window Manager.

The IDISCO software includes a new Interactive Monitor program and specific interactive applications, including

- ☐ Trace Flow Controller — controls the data flow through interactive processes
- ☐ Session Manager — an interactive session controller
- ☐ Pointing Dispatcher — performs inter-task communication
- ☐ Job Builder — implements graphical job set-up and display
- ☐ Parameter Editor — a smart parameter editor with on-line help and parameter validation
- ☐ Trace Widget — a Motif and X Window System trace display utility
- ☐ Graphics Routine — displays general-purpose X Window System graphics

For more information on using IDISCO with Cray Research computer systems, contact Glenn Cornelius, Executive Vice President, CogniSeis Development, Inc., 2401 Portsmouth Suite 270, Houston, TX 77095; telephone: (713) 526-3273; or Arvind Amin, Cray Research, Inc., Suite 800, LB 50, 5005 LBJ Freeway, Dallas, TX 75244; telephone: (214) 450-9500.

USER NEWS

Dynamical structure of proteins studied

This article was written by Stu Borman and is reprinted with permission from Chemical & Engineering News, *December 10, 1990, 68(50), pp 20-22. Copyright 1990, American Chemical Society.*

A team of theoretical chemists has developed a supercomputer-based method for probing the internal motion of proteins. The method, which is used to figure out how distant and seemingly unrelated parts of large proteins exert an influence on each other, has already led to a better understanding of HIV-1 protease, an essential enzyme in the replication of human immunodeficiency virus.

The method, developed by S. Swaminathan and David L. Beveridge of Wesleyan University, Middletown, Connecticut, and William E. Harte Jr., M. M. Mansuri, J.C. Martin, and I.E. Rosenberg of Bristol-Myers Squibb Co., Wallingford, Connecticut, is based on the idea that motions in one domain of a protein can affect motions in other domains [Proc. Natl. Acad. Sci. USA, 87, 8864 (1990)]. Such correlated motions can play an important role in protein functions.

Analyzing these motions is a two-step process, say the researchers. First, the motions of amino acid residues in the protein are calculated using molecular dynamics. Then a two-dimensional mapping technique is used to identify correlations between motions in different domains.

Researchers such as J. Andrew McCammon of the University of Houston and Martin Karplus of Harvard University have used molecular dynamics and mapping techniques to identify correlated motions between residues in an active site. However, this is the first time the techniques have been used to identify domain structure in proteins and to analyze long-range interactions between domains.

Beveridge and coworkers use the method to study HIV-1 protease dimer, an enzyme essential for the proliferation of the AIDS virus and thus a major focus of attention among researchers trying to find AIDS treatments. Their simulation, which included a solvation shell of 6990 water molecules along with 198 amino acid residues, required 100 hours of computer time on a CRAY Y-MP system at the Pittsburgh Supercomputer Center. "The capability of doing calculations of this scale was not available until the advent of supercomputers," says Beveridge.

HIV-1 protease dimer is the only HIV-associated protein whose detailed structure is known. This knowledge is based in part on x-ray crystallographic studies by Alexander Wlodawer of the National Cancer Institute, Frederick, Maryland. The simulation by Beveridge and coworkers takes this x-ray structure as its starting point. "The crystal structure provides you with a single, time-averaged structure of the molecule," says Beveridge, "but in its active form the molecule is undergoing internal dynamics."

The simulation predicts that motions in two distinct regions of HIV-1 protease are correlated with each other, although

CUG reports

Users of Cray Research computer systems established the Cray User Group (CUG) in 1977 to provide a forum for the exchange of ideas related to Cray Research systems and their applications. The group holds two general meetings each year. Its first meeting of 1991 was held April 22-26 and was hosted by the University of London Computing Center. Below, CUG president Mary Zosel, of the Lawrence Livermore National Laboratory, offers her comments on the meeting and other CUG-related business.

Cray User Group meetings have survived airline strikes and typhoon and tsunami warnings, but the months preceeding our London meeting produced some of the most anxious moments yet. Attendance was not a casualty, however. We enjoyed our largest European meeting ever with close to 300 people in attendance.

With the addition of nine new member sites in the past six months, our membership now stands at 181 sites from 21 countries. Organizationally we have added two new mutual interest groups (MIGs): the User Services MIG chaired by Jean Shuler from the Lawrence Livermore National Laboratory and the Environmental MIG with Gary Jensen from the National Center for Atmospheric Research as acting chair. Mass storage and accounting are two technical areas also generating much interest at CUG meetings.

Last year CUG asked Cray Research to address two important issues: plans for IEEE arithmetic support and availability of some software products on platforms other than Cray Research hardware systems. Cray Research since has announced its intention to convert to IEEE arithmetic, and at our most recent meeting representatives from the company announced that they are preparing plans and policy for support of some software products on workstations. CUG members will be working with Cray Research to develop these plans.

Two issues which have been outstanding for some time are approaching closure. There has been a request for some form of software sharing between CUG sites. The San Diego Supercomputer Center has taken an active lead in this area, and a practical new proposal was made at the London meeting. The plan is to maintain a directory at SDSC of available software packages and the sites to contact to obtain them.

The second issue is related to the difficulties many sites have had in obtaining the proper source licenses needed for the UNICOS operating system. A letter citing various problems has been sent to the president of UNIX System Laboratories with an invitation for them to send a representative to the next CUG meeting to discuss licensing procedures.

The last thing I would like to report to you is that CUG joined Cray Research in sponsoring an executive symposium in February, with about 50 representatives from customer sites attending. The purpose of the symposium was to provide a forum for the exchange of long-range planning information for computer centers. Presentations were made by Riaz Abdullah of Eli Lilly, Al Erisman of Boeing, Greg McRae of Carnegie Mellon University, and George Lindamood of the Gartner Group, along with representatives from Cray Research. The symposium included panel discussions about changes expected in computing by 1995 and the U.S. HPCC initiative. Possible future symposiums are being discussed.

For reference, I have included a list of members of the CUG board of directors and advisory council.

Bob Baynes, TAI	Operating systems SIC chair
Harvey Brock, NRL	Performance SIC chair
Gunter Georgi, Grumman	CUG secretary
Juergen Gottschewski, ZIB	Berlin local arrangements chair
Michael Heib, Daimler Benz	German CRI users representative
Carol Hunter, LLNL	Graphics SIC Chair
Michel Jaunin, EPFL	Latin CRI users representative
Gary Jensen, NCAR	Acting environmental MIG chair
Helene Kulsrud, IDA	Mass storage systems SIC chair
Christopher Lazou, ULCC	Past BOD member, past local arrangements chair
Claude Lecoeuvre, CEA-CEL	CUG director at large
David McWilliams, NCSA	CUG director at large
Sam Milosevich, Eli Lilly	Advisory council secretary
Reagan Moore, SDSC	Software tools SIC chair
Ken Neves, Boeing	CUG director at large
Steven Niver, Boeing	Past BOD member, LALC chair
Fran Pellegrino, Westinghouse	Operations SIC chair
Anna Pezacki, U. Toronto	Communications SIC chair
Robert Price, Westinghouse	Management SIC chair
Karen Sheaffer, SNLL	CUG vice president
Jean Shuler, LLNL	Newsletter editor, user services MIG chair
Charles Slocomb, LANL	Santa Fe local arrangements chair
Kyosuke Tsuruta, Toshiba	Japanese CRI users representative
Howard Weinberger, TAI	CUG treasurer, aerospace MIG chair
Mary Zosel, LLNL	CUG president

The CUG meeting schedule for the next few meetings includes fall 1991 in Santa Fe, New Mexico (September 23-27), spring 1992 in Berlin (April 6-10), fall 1992 in Washington D.C. (September 13-18), spring 1993 in Montreux, Switzerland (March 28-April 2), and fall 1993 in Kyoto, Japan (early October).

the regions are separated by a considerable distance. One domain, which the researchers named the "cantilever," acts through a "fulcrum" to cause motion in a "flap" region near the enzyme's active site.

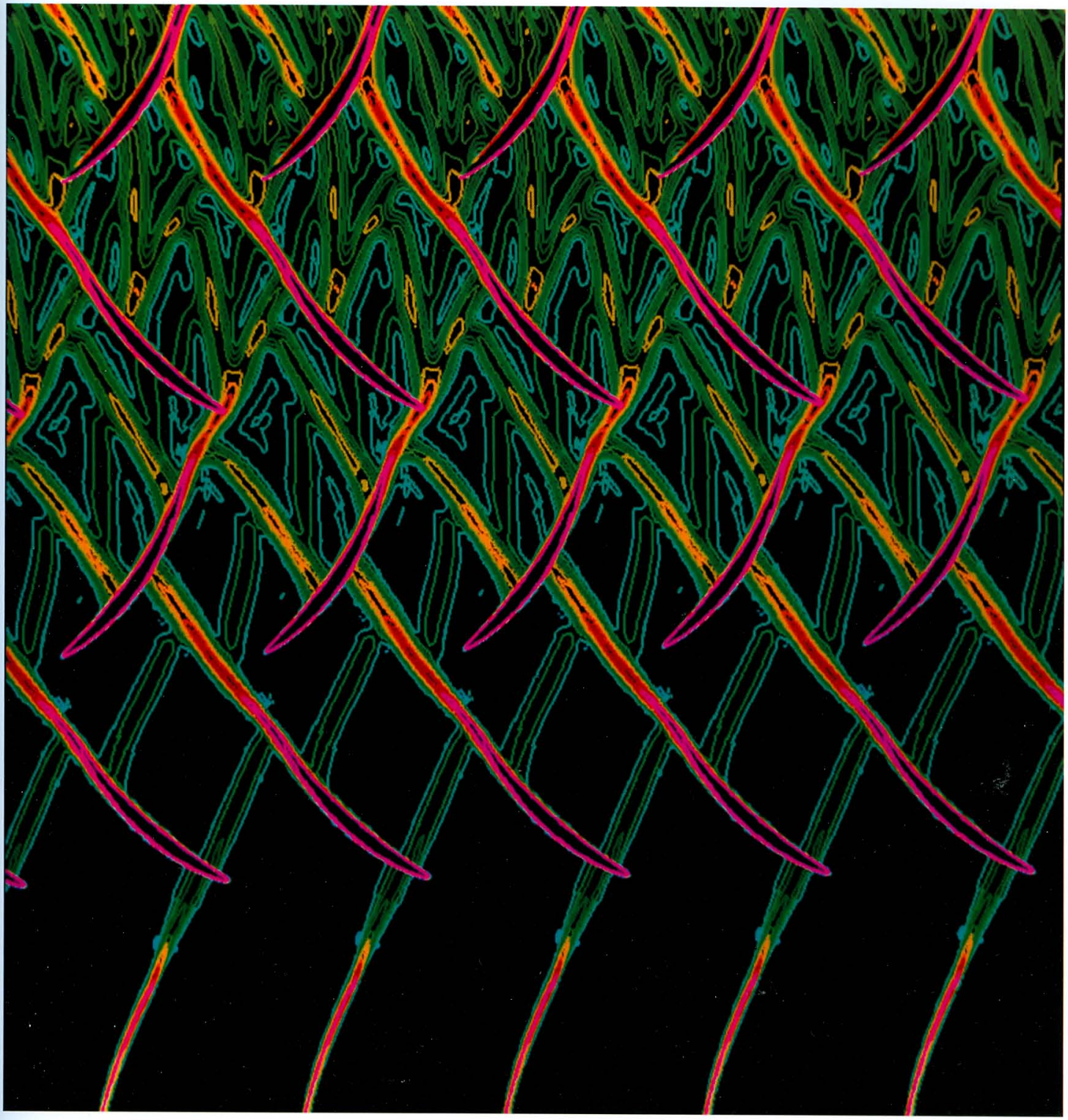
The flap's function is believed to involve guarding the entrance to the active site cleft or excluding water from the binding site. Domains such as the cantilever that affect the flap could also play a role in enzyme function and therefore could be targets for enzyme inhibitors, and potentially of use for development of AIDS therapies.

However, says Beveridge, "the idea of through-space interactions in proteins is not so much a breakthrough in AIDS research as much as a new way of looking at dynamical structure of proteins. It's a form of structure that originates uniquely in the dynamical motions."

He adds that "one would never know about this and would have no means to investigate it without large-scale molecular dynamics simulations, because no experiment is going to give you all this information on detailed atomic motions. Here, computational theory has a unique vantage point in elucidating the details of molecular motion."

The work could also bear on the mechanism of other protein processes, he suggests. "There are phenomena, such as allosteric inhibition of enzymes, where something binds far away from an active site and yet influences enzyme function. And protein folding is expected to involve nucleation of domains. What intrigues us now is the possibility that things like this go via correlated motions, but that remains to be established. This is now the subject of further studies."

GALLERY



Entropy contours show the viscous flowfield within a multistage compressor. The image is from research to improve the operation of aircraft turbomachinery. The model was run by Karen L. Gundy-Burlet and Man Mohan Rai on a CRAY Y-MP system at the NASA Ames Research Center.