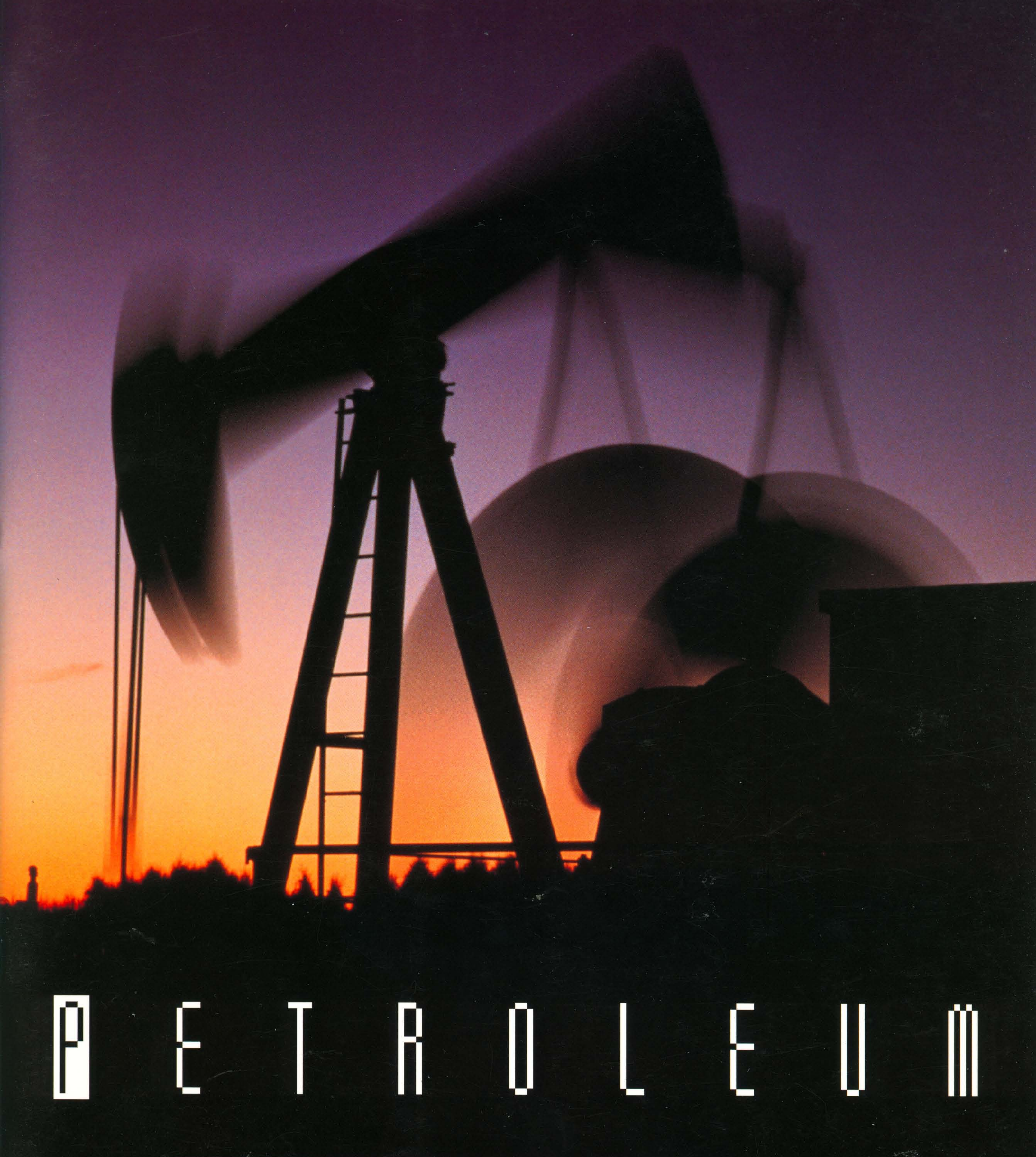
CRAY CHANNELS

VOLUME 15, NUMBER 3 • A CRAY RESEARCH, INC. PUBLICATION



PETROLEUM

Announcing CRAY T3D massively parallel and CS6400 SPARC server systems

CRAY CHANNELS

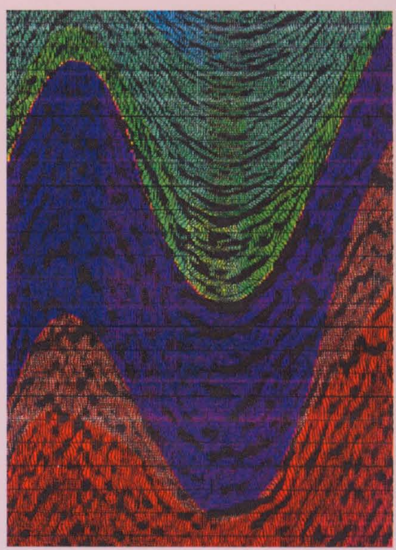
In this issue

Nearly two decades have passed since the energy crisis of the 1970s heightened the world's awareness of the value of petroleum. In the intervening period, the petroleum industry has invested heavily in efforts to maximize production of current oil fields. The threat of potential scarcity underscored the need to fully exploit technology such as seismic processing to pinpoint deposits and reduce exploration costs. Lowering these costs in turn lessens the pressure on the monies needed to sustain high standards of living.

This issue of CRAY CHANNELS contains several articles that discuss seismic exploration and how supercomputing accelerates the process of petroleum discovery. This issue also introduces two of Cray Research's new computer systems, the CRAY T3D massively parallel processing (MPP) system and the CRAY SUPERSERVER 6400, a SPARC-compatible product of Cray Research Superservers, Inc., a subsidiary of Cray Research. In addition, this issue includes an article on computer simulation of computational fluid mixing in the chemical process industries. Our regular departments include reports on Cray Research's new CF90 and Cray Standard C programming environments, as well as version 2.0 of Cray Research's UniChem software and the announcement of our new anonymous ftp information server.

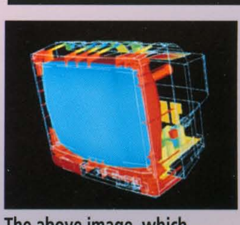
Since the CRAY-1 system was introduced in 1976, Cray Research has helped the petroleum industry break computational barriers. With the introduction of the CRAY T3D system, Cray Research continues its tradition of placing powerful computational tools in the hands of petroleum engineers. Engineers can now distill terabytes of 3-D prestack seismic imaging data into useful information for oil extraction. Tomorrow, Cray Research will continue to help petroleum engineers solve the industry's next generation of grand-challenge problems.

Features



18

Departments



The above image, which appeared uncredited in the last issue, was provided by Samsung Electronics Company.

CRAY CHANNELS is a publication of the Cray Research, Inc. Marketing Communications Department. Published three times per year, it is intended for users of Cray Research computer systems and others interested in the company and its products. Subscription inquiries and address changes should be sent to Department D, CRAY CHANNELS, Cray Research, Inc., 655D Lone Oak Drive, Eagan, Minnesota 55121.

Volume 15, Number 3

Editorial staff
Ken Jopp, managing editor
Chris Kruell, associate editor
Molly Sachse, associate editor
Rich Brueckner, Gayla Conners, Heidi Hagen, and Jeff Paulson, contributing editors

Design and production
Barbara Cahlander, electronic production artist
Eric Hanson, graphic designer
James Morgan, art director

PETROLEUM

Cover photo: Stock Imagery, Inc.

From infeasible to practical

CRAY T3D system makes production-quality 3-D prestack depth migration a reality

Cray Research unveiled it at the 1993 meeting of the Society of Exploration Geophysicists. A major oil company is one of the first customers committed to it, with other oil companies expected to follow suit. And 3-D prestack seismic imaging is one of the first application development initiatives for it. No mere string of coincidences, these factors demonstrate what Cray Research has believed for some time: the petroleum industry, long a major customer of Cray Research technology, stands to gain even more from Cray Research's newest product offering—the CRAY T3D massively parallel processing (MPP) system.

A "grand challenge" application for the petroleum industry, 3-D prestack depth migration in a production environment has eluded the capabilities of all previous supercomputing platforms. The lack of a production-environment MPP, overwhelming I/O demands (input data sets on the order of hundreds of gigabytes—even terabytes—of data), and compute-intensive algorithms were insurmountable barriers. Not anymore. The CRAY T3D system makes 3-D prestack depth migration possible and changes the qualifiers—from research to production, from two-dimensional to three-dimensional, from poststack to prestack, and from time migration to depth migration. Exploration geophysicists can meet the requirements of full 3-D imaging in production environments, modeling steep dips in complex subsurface geologies with strong lateral velocity variations. Already, observed turnaround times on a 128-processor-element (PE) CRAY T3D system at Cray Research enabled analysts to image a typical marine survey containing 40 million traces into a 409,600-grid-point vertical image plane in two hours. After working extensively with Cray Research on a 3-D prestack seismic depth migration problem, one major oil company, for the first time ever, is making a major commitment to this kind of production-quality processing.

From research to production

In general, the CRAY T3D system promises a three-week turnaround on problems with 30 to 40 million input seismic traces, as encountered in many 3-D surveys. This kind of timeframe was unthinkable in the past. The CRAY T3D system also establishes a new level of price/performance. It costs less than the parallel-vector supercomputer that would be required for these 3-D surveys while delivering the same or greater performance. And because the CRAY T3D system matches fast microprocessor speed with fast I/O, fast memory access, and capable software, it is the most balanced MPP architecture available.

Workstation technology has been used in earlier-generation MPP systems to provide loosely coupled, multicomputer environments. And while individual workstations boast high-speed computing in the mega- and gigaflops-range, clusters of such workstations are insufficient for production-class problems because the channels connecting the workstations are slow.

The CRAY T3D system runs Cray Research's UNICOS operating system, a proven petroleum production environment that Cray Research customers have used for years. The CRAY X-MP system made 2-D prestack and 3-D conventional modeling possible. The CRAY Y-MP and CRAY C90 systems improved the processing speed and production potential of 3-D dip-move-out and 3-D poststack imaging and targeted 3-D prestack imaging. With the CRAY T3D MPP system, petroleum companies can expect to conduct 3-D prestack imaging and direct lithology characterization in a more cost-effective and timely way than ever before.

In addition to having access to an MPP system based on proven technology, petroleum industry users will benefit from high-bandwidth I/O and unsurpassed high-speed interprocessor communications.

High-bandwidth I/O

The flexibility and capacity of I/O channels in the CRAY T3D system helps make this a truly balanced architecture. Now the real data acquired from seismic field surveys can be read into the machine as many times as necessary and processed at a peak throughput rate in the gigabytes-per-second range. Based on Cray Research's Model E I/O technology, the CRAY T3D system's high-bandwidth I/O subsystem provides access to the full range of Cray Research disk, tape, and network peripherals. Extensive networking connectivity to a variety of protocols through HIPPI, FDDI, and Ethernet media is supported. This infrastructure becomes increasingly important as 3-D processing calls for 3-D visualization of results.

High-speed interprocessor communications

Because of the extraordinarily fast connections between its processor elements (PEs), the CRAY T3D system surpasses the capabilities of workstation clusters. A very fast bidirectional 3-D torus system interconnect network ensures short connection paths and high bisection bandwidth. With peak interprocessor communication rates of 300 Mbytes per second in every direction through the torus, resulting in up to 76.8 Gbytes per second of bisection bandwidth for a 1024-processor system, this design allows the extremely fast remote memory access critical for efficient MPP system usage.

Based on technology not found in any other MPP system, this high-speed interprocessor communication is wrapped around commodity state-of-the-art microprocessors—the DECchip 21064 (more familiarly known as the DEC Alpha) from Digital Equipment Corporation, which is capable of 150 MFLOPS peak performance. This reduced instruction set computing (RISC) microprocessor is cache-based, has pipelined functional units, issues multiple instructions per cycle, and supports IEEE standard 64-bit floating-point arithmetic. Each processor has its own local DRAM memory with a capacity of either 16 or 64 Mbytes. The Alpha microprocessor can process in 32-bit words as well. Cray Research software takes advantage of this capability—a 32-bit version of the Fortran compiler and libraries which will further enhance performance for signal processing applications.

In the CRAY T3D system, memory is physically distributed among processors, but is globally addressable. Any microprocessor can address any memory location in the system. The high-level programming models developed for the CRAY T3D system take advantage of this shared distributed memory—together with the high performance, low latency interconnect technology—to support high performance on applications with fine, medium, and coarse-grained parallelism.

Reservoir simulation

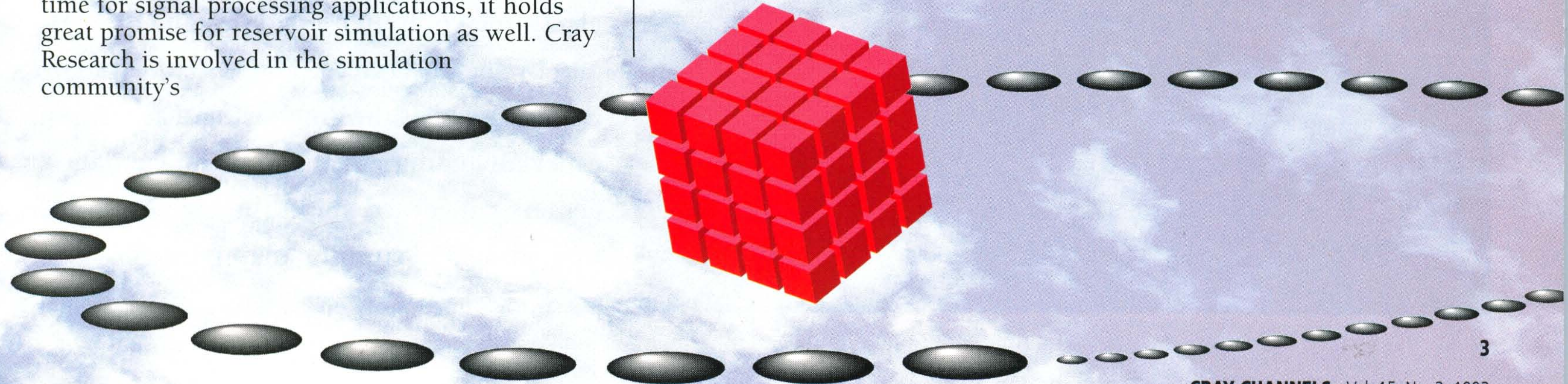
While the CRAY T3D system is making breakthroughs in performance and turnaround time for signal processing applications, it holds great promise for reservoir simulation as well. Cray Research is involved in the simulation community's investigation of development efforts for a parallel solver for MPP systems. With power, memory capacity, and a price/performance ratio unsurpassed by any other supercomputer or workstation, the CRAY T3D system will be critical as users increase the number of grid blocks within oil reservoir models, capture all the necessary physics and fluid flow through porous media, and transfer that data into a numerical simulation. Indeed, the CRAY T3D system opens the door to field-scale execution of many types of recovery mechanisms—depletion mode, double-porosity formulations, and thermal and displacement models, for example. Current models of these mechanisms are limited significantly by memory constraints and impractical execution times.

Through geostatistics, the oil industry is looking at statistically creating geological realizations that can be captured in a numerical model. To fully appreciate the benefits of geostatistics, a very large number of numerical grids are necessary, and the power and memory of the CRAY T3D system are essential.

Maintaining the customer connection

Cray Research hardware and software engineers worked closely with industry-recognized experts and key customers to create the robust, reliable, sharable, and easy-to-administer CRAY T3D system. Cray Research remains actively involved in several collaborative projects with customers, prospects, and applications software vendors. The company has signed a letter of agreement with a major oil company to demonstrate the functionality of the CRAY T3D system and a 3-D prestack depth migration application by first quarter 1994. In addition, Cray Research has active petroleum benchmarking studies under way with a number of customers. Finally, the company has established a program through which specific customer sites will receive the earliest CRAY T3D systems and partner with Cray Research on MPP applications development. This program targets reservoir modeling as one of the key MPP application areas for parallel code conversion, as well as seismic processing and new refinery simulation techniques.

The CRAY T3D system is the most usable, highest-performing MPP system geared for a production environment. High-bandwidth I/O and high-speed interprocessor communications make it especially valuable in the attack on grand challenge problems in the petroleum industry. This is why Cray Research calls the CRAY T3D MPP system "the right tool at the right time."



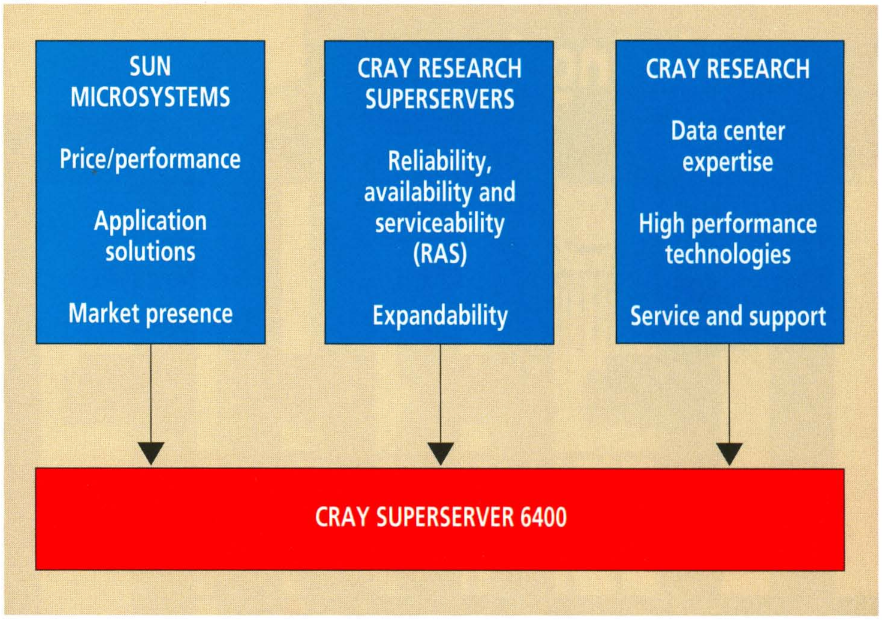
A new enterprise server for business and technical computing

Synonymous with powerful scientific and technical computing for over 20 years, Cray Research has just expanded its horizons.

Through its Beaverton, Oregon-based subsidiary, Cray Research Superservers (CRS), Cray Research introduced the world's fastest and most expandable SPARC/Solaris-compatible server, the CRAY SUPERSERVER 6400 system. The CS6400 system addresses commercial markets and applications not typically associated with previous Cray Research systems, such as relational database management systems (RDBMS), transaction processing monitors, fourth generation languages, computer-aided software engineering (CASE) tools, electronic computer-aided design (ECAD), and financial services. The CS6400 system is also aimed at the rightsizing market, especially commercial and technical data centers concerned about the high cost of upgrading and running mainframe systems.

The CS6400 system, developed by CRS under a Cray Research and Sun Microsystems technology exchange agreement, runs Sun's Solaris UNIX SVR4 operating environment and is a binary-compatible upward extension of Sun's product line. The CS6400 system leverages SPARC technology dominance in the RISC market—a 1992 International Data Corporation study showed that SPARC technology holds a 57 percent market share for the UNIX RISC market.

"The SPARC architecture is a powerful foundation for cost-effective scalability," said Scott McNealy, Sun chairman and chief executive officer. "The Sun and CRS product lines will form a price-performance continuum from SPARCclassics to SPARCcenter servers to the CS6400 product. Sun users who require more power on the network can move to the binary-compatible CS6400 system with no migration problems."



The CRAY SUPERSERVER 6400 system combines the strengths of Cray Research, Sun Microsystems, and Cray Research Superservers to provide a powerful enterprise server.

Strong price-performance with reliability, availability, and serviceability

The CS6400 system delivers features to the open systems environment that data center users have enjoyed for years. The strong price-performance of Sun products with built-in reliability, availability, and serviceability (RAS) features makes the CS6400 system the first of its kind in the high-end, open systems server arena.

With up to 64 processors, the CRAY SUPERSERVER 6400 is the world's fastest SPARC/Solaris system

One example of the RAS features is the ability of the system to automatically reboot, isolate a fault, and reconfigure itself, and its "hot swap" capabilities allow a failed module to be removed and replaced while the system is still running. Systems also can be upgraded while online.

The expandable system scales with customer processing requirements. The CS6400 system is configured with 4 to 64 SuperSPARC RISC microprocessors (initially at 60 MHz), 256 Mbytes to 16 Gbytes of central memory, 1.3 Gbyte/s peak data bandwidth, and more than 2 Tbytes of online disk storage. Key features of Cray Research's supercomputing environment—sophisticated tape management, high-performance I/O, networked batch processing, system management software, program debugging tools and high-performance compilers—are being incorporated.

Expandable system fits into existing networks

Electricité de France (EDF), the world's largest electrical utility, acquired a 16-processor CS6400 system with 2 Gbytes of central memory, which will be upgraded to 32 processors and 4 Gbytes of memory. EDF, a long-time Cray Research customer, plans to use the CS6400 for scientific, engineering, and data management applications. "We were looking for a high-performance UNIX server to support a variety of applications and a large number of users, and the CS6400 system was our choice," said Michel Pavard, head of EDF's data processing and applied mathematics unit.

SICAN, a leading German microelectronics firm, has ordered a 48-processor system. The system will be used for highly scalar ECAD applications, as well as for data management. "We are targeting the CS6400 system for a number electronic design codes including commercially available programs and those we've developed in-house," said Hans Weinerth, chairman and chief executive officer for SICAN. "The CS6400 is an outstanding option for us since most of our codes now run on our more than 50 Sun systems. We know that these same codes will run without recompilation on the CS6400. In effect, the CS6400 allows us to upgrade each of our workstations with one single purchase."

With U.S. pricing beginning at under $400,000 for the 4-processor version, and at $2.5 million for the top-of-the-line 64-processor system, the CS6400 is an affordable solution that provides more capacity and capability for thousands of applications.

Aiming for TERAFLOPS in seismic processing

A. J. Berkhout,
Delft University of Technology, The Netherlands

The use of seismic migration to visualize the 3-D subsurface is one of the most valuable exploration and production technologies in the petroleum industry. It is also incredibly complex—and expanding quickly. We can expect in the coming decades that subsurface usage will include a variety of geo-activities with far-reaching public policy implications. These activities (Table 1) likely will include

- Use of the underground infrastructure to transport information, goods, and people
- Exploitation of geo-energy
- Management of groundwater reservoirs
- Storage of hazardous waste and energy residues

Table 1 (below). The geoscience market.

Geo-energy	Groundwater	Geo-technics
Extraction of energy	Water supply by groundwater	Extraction of ores and industrial minerals
Subsurface storage of energy	Subsurface storage of water	Facilities concerning the infrastructure
Subsurface storage of energy residues	Subsurface purification of polluted water	Man-made fractions of reservoirs

To optimally regulate and use the earth's subsurface, policy makers and players in the geoscience market need accurate information about the in-situ condition of soil and rock; information not only about the depth and shape of the geologic layering (geometric properties), but also about the deformation behavior of the subsurface and its fluid flow properties.

An important part of international geoscience research and development is devoted to scientific methods that "invert" surface and borehole measurements into a detailed 3-D model of the subsurface that contains all parameters of engineering interest. Technical geoscience is a macrodiscipline that generates technology to optimally utilize the subsurface in terms of economics and ecology. It is based on the three disciplines of geophysical imaging, geological characterization, and subsurface engineering. Satisfactory solutions to today's technical geoscience problems can be expected only if research in all three disciplines is fully integrated. Ideally, geophysical imaging and geological characterization should be carried out iteratively, using surface as well as borehole measurements. An even larger scientific challenge is to incorporate fluid flow measurements in this iterative process, ultimately leading to accurate multiphase flow models that properly take into account small-scale reservoir heterogeneities. This means that leading geoscientific research can only be expected from large-scale research centers, where high-quality research is carried out in an integrated manner. It also means that those research centers require very advanced information technology tools related to visualization, supercomputing, and mass storage.

Organizing the seismic data

Information consists of organized data and scientific functions that meaningfully manipulate the data. Seismic field data is five-dimensional (two source coordinates, two receiver coordinates, and one time ordinate): if this 5-D dataset is not organized properly, the description of any prestack seismic process becomes very complex. This is one reason seismic processes have been formulated in the much simpler midpoint domain. However, it turns out that the data matrix **P** is an elegant and practical way to present prestack seismic data. It is easy to understand because of its geometrically oriented structure, and, above all, it is directly available for a concise formulation of complex mathematical manipulations (Figure 1). For instance, if the matrix operator $\mathbf{F}_1$ is applied to data matrix **P**, then

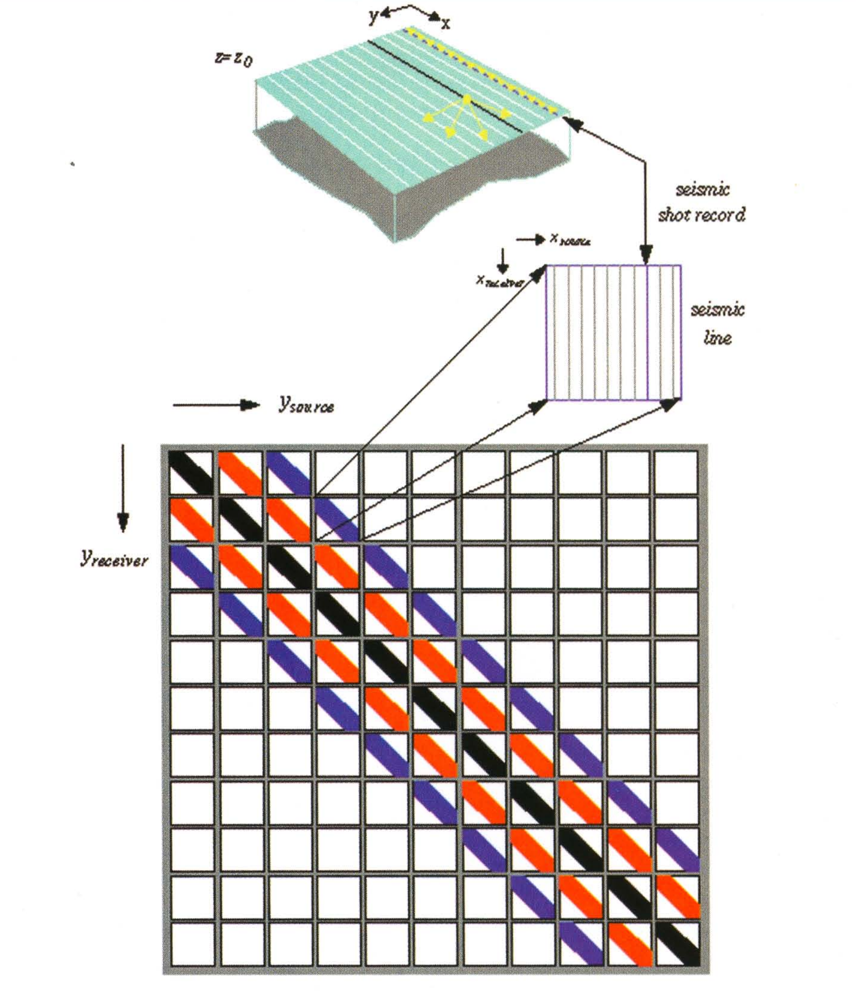
$$\mathbf{Q}_1 = \mathbf{F}_1\mathbf{P}, \qquad (1a)$$

means that all measurements are filtered along the receiver coordinates(x_r, y_r). Similarly,

$$\mathbf{Q}_2 = \mathbf{P}\mathbf{F}_2 \qquad (1b)$$

means that all measurements are filtered along the source coordinates(x_s, y_s).

Formulations 1a and 1b are generic and may represent basic seismic processes such as prestack Radon transformation and prestack multichannel deconvolution (parts of preprocessing), and prestack wave field extrapolation (part of migration).



Multiline measurement matrix; for $Y_{source} \neq Y_{receiver}$ the source has an offset with respect to the receiver line.

Figure 1. The data matrix for 3-D seismic measurements. One column represents one multistreamer shot record. For a set of 2-D parallel lines all submatrices off the diagonal are zero. For a full 3-D dataset all submatrices and all elements in each submatrix contribute. For an irregular dataset the columns are irregularly filled.

It is important to realize that **P** is determined for each temporal Fourier (or Laplace) component and thus may be seen as the result of one monochromatic seismic survey. Typically, as many as 500 to 1000 components may exist. As all components are independent, each **P** matrix may be processed fully independently of the others. This property has two important consequences: the very large amount of travel-time-related time shifts that must be applied in seismic processing are replaced by simple multiplications; and different frequency components may be processed by different processors without any interaction (large grain parallelism).

Forward model of seismic data

The principal physical processes that occur in the subsurface during each seismic experiment can be summarized by the following (Figure 2):

☐ Downward propagation

$$\mathbf{P}^+(z_m) = \mathbf{W}^+(z_m, z_o)\mathbf{P}^+(z_o), \qquad (2a)$$

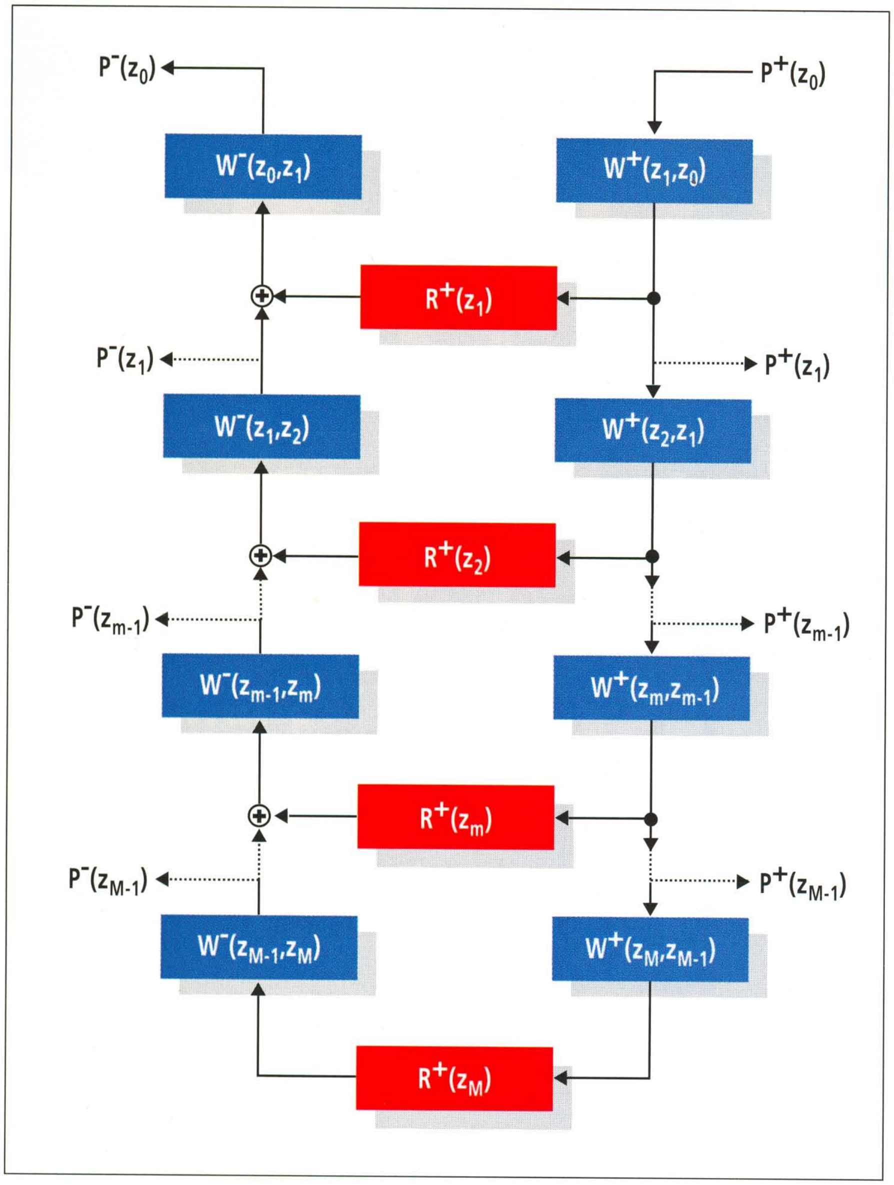


Figure 2. The seismic processes in the subsurface in terms of propagation operators $\mathbf{W}^+$ (down) and $\mathbf{W}^-$ (up), and reflection operators $\mathbf{R}^+$.

where $\mathbf{P}^+(z_o)$ defines the wave fields that leave the surface z_o, and $\mathbf{P}^+(z_m)$ defines the wave fields that arrive at depth level z_m (incident wave fields); $\mathbf{W}^+(z_m, z_o)$ is the downward propagation matrix that defines the propagation property between each gridpoint of z_o and z_m.

☐ Reflection

$$\mathbf{P}^-(z_m) = \mathbf{R}^+(z_m)\mathbf{P}^+(z_m), \qquad (2b)$$

where $\mathbf{P}^-(z_m)$ represents the reflected wave fields that leave depth level z_m due to inhomogeneities at z_m; $\mathbf{R}^+(z_m)$ is the reflectivity matrix that defines the angle dependent reflection property at each gridpoint of z_m.

☐ Upward propagation

$$\mathbf{P}^-(z_o) = \sum_m \mathbf{W}^-(z_o, z_m)\mathbf{P}^-(z_m), \qquad (2c)$$

where $\mathbf{P}^-(z_o)$ represents the reflected wave fields that arrive at the surface z_o; $\mathbf{W}^-(z_o, z_m)$ is the upward propagation matrix that defines the propagation property between each gridpoint of z_m and z_o.

Expressions 2a, 2b, and 2c can be integrated to the WRW model:

$$\mathbf{P}^-(z_o) = \sum_m [\mathbf{W}^-(z_o, z_m)\mathbf{R}^+(z_m)\mathbf{W}^+(z_m, z_o)]\mathbf{P}^+(z_o). \quad (3a)$$

In case of primary reflections we should take in (3a)

$$\mathbf{P}^+(z_o) = \mathbf{S}^+(z_o), \qquad (3b)$$

where each column of $\mathbf{S}^+(z_o)$ defines the downgoing source wavefield at a source position. Ideally $\mathbf{S}^+(z_o)$ equals the unity matrix.

Seismic numerical processes

Analyzing the numerical processes that are being applied in the seismic discipline, one may distinguish two principal functionalities:

☐ Simulation of synthetic measurements
☐ Inversion of field measurements

Simulation is based on a user-specified model of the subsurface as well as an algorithm that predicts the geophysical phenomena of interest—seismic waves—in that model. Hence, simulation may generate synthetic measurements for a given subsurface model and a given data acquisition geometry.

Inversion is the other way around: the real measurements are given, and the subsurface is unknown. Unfortunately, determination of the subsurface model from measurements defines an ill-posed problem, as seismic measurements only represent a projection of the subsurface. Therefore, strategy is a key issue in seismic inversion, and success depends largely on the preparation of the input (preprocessing). If the seismic field measurements are decomposed into different wave field components (along versus below the surface, primaries versus multiples, and P-waves versus S-waves), the unstable joint inversion of all superimposed components can be replaced by the robust sequential inversion of the individual components. In the seismic discipline, joint inversion of all wave field components is based on the two-way wave equation; sequential inversion of the individual wave field components is based on the one-way wave equation.

Another important aspect of the inversion strategy is the parameterization of the subsurface. In the context of the seismic method, the subsurface can be described in terms of reflectivity, velocities and density, and rock and pore parameters.

Reflectivity is a boundary-related property, generally dependent on the incident wave field (angle of incidence). It shows the structural properties of the subsurface and, with sufficient resolution, the depositional patterns. If the inversion process aims at reflectivity, the propagation effects between the surface and each reflecting boundary must be removed from the data. This is done in the seismic migration process.

Velocities (c_p, c_s) and density (ρ) are layer-related properties. If we consider macrolayers, the "macrovelocity model" implicitly defines the propagation operators to be used in migration, making the inversion process that aims at a macrovelocity model very important. As information about the macrovelocity model is largely presented in the seismic travel times, the involved forward relationships are very nonlinear. On the other hand, if we consider local velocities and density, then the detailed spatial changes (due to fine layering) define the reflectivity of the related seismic boundaries. For small relative changes across a boundary there is a linear relationship between the contrast parameters $[\Delta c_p/\bar{c}_p, \Delta c_s/\bar{c}_s, \Delta\rho/\bar{\rho}]$ and the reflectivity. However, for large relative changes the relationship is significantly more complex and clearly nonlinear, particularly for large angles of incidence.

Note that the global macromodel and the detailed contrast model together define the full elastic model in terms of c_p, c_s, and ρ. Inversion for the full elastic model implicitly involves two forward relationships: velocities and density to reflectivity, and reflectivity to measurements. The first relationship is defined by the boundary conditions and determines the net amplitude-versus-offset (AVO) properties in the data; the second relationship is defined by the propagation operators (macromodel-based **W** operators) and determines travel times in the data.

What is gained by extending the inversion objective from a reflectivity model to an elastic model? Not much if the elastic model is an end product. Given the physical and empirical relationships between the elastic model (in terms of c_p, c_s, and ρ) and the geological model (in terms of rock and pore parameters), a final inversion process should be designed that translates the seismic information into geological information. We refer to this inversion process as seismic reservoir characterization.

Summarizing, seismic inversion can be subdivided into three hierarchical processing steps:

- ☐ Decomposing the data into different wave field components (preprocessing)
- ☐ Imaging for reflectivity (migration)
- ☐ Estimation of rock and pore parameters (characterization)

Migration may be considered the nucleus of modern seismic processing technology. The challenge of the 1990s is to develop an economic solution to the 3-D prestack migration problem that includes estimation of the macrovelocity model.

TERAFLOPS

"Rather than concentrate on the conservative, ray-tracing-based, space-time algorithms only, the industry should make significant investments in the long-term solution given by the parallel, wave-theory-based, space-frequency algorithm."

Alternative numerical solutions

After perfect preprocessing, the model of the seismic response related to the inhomogeneities at depth level z_m, can be represented by

$$\mathbf{P}_m^-(z_o) = \mathbf{W}^-(z_o, z_m)\mathbf{R}^+(z_m)\mathbf{W}^+(z_m, z_o). \qquad (4a)$$

Note from expression (4a) that migration involves removal of the influence of propagation operators $\mathbf{W}^+(z_m, z_o)$ and $\mathbf{W}^-(z_o, z_m)$ from the seismic data at z_o:

$$\mathbf{R}^+(z_m) = \mathbf{F}^-(z_m, z_o)\mathbf{P}_m^-(z_o)\mathbf{F}^+(z_o, z_m), \qquad (4b)$$

where $\mathbf{F}^+$ and $\mathbf{F}^-$ are the inverses of $\mathbf{W}^+$ and $\mathbf{W}^-$, respectively. In the migration literature, the direct inversion process (4b) is referred to as "downward extrapolation." Note from (4b) that downward extrapolation involves a convolution process along the source coordinates (application of matrix operator $\mathbf{F}^+$) and along the detector coordinates (application of matrix operator $\mathbf{F}^-$). Looking at the size of 3-D data matrix $\mathbf{P}^-(z_o)$, which is several gigabytes, and the hundreds of depth levels, downward extrapolation involves many thousands of teraflops per frequency component.

In practice, $\mathbf{P}_m^-(z_o)$ is not available, and the downward extrapolation process (4b) is applied to the full seismic data set $\mathbf{P}^-(z_o)$ as given by (3). This means that after downward extrapolation to depth level z_m, a selection process must be applied to recover $\mathbf{R}^+(z_m)$ from the extrapolation result. In the migration literature this selection process is referred to as "application of the imaging principle."[1]

Summarizing, seismic migration consists of three numerical processes:

- ☐ Computation of forward operators $\mathbf{W}^+$ and $\mathbf{W}^-$ for a prespecified macrovelocity model
- ☐ Determination of the related inverse operators $\mathbf{F}^+$ and $\mathbf{F}^-$ followed by downward extrapolation
- ☐ Application of the imaging principle

A confusingly large variety of migration algorithms in the seismic industry represents all kinds of variations of the above three numerical processes. Fortunately, all significant developments today aim at depth migration with integral operators. Focusing on the integral-based depth-migration methods, they can be subdivided into two main categories:

- ☐ Nonrecursive, in which the operators $\mathbf{W}^+$ and $\mathbf{W}^-$ are represented by global, time-domain

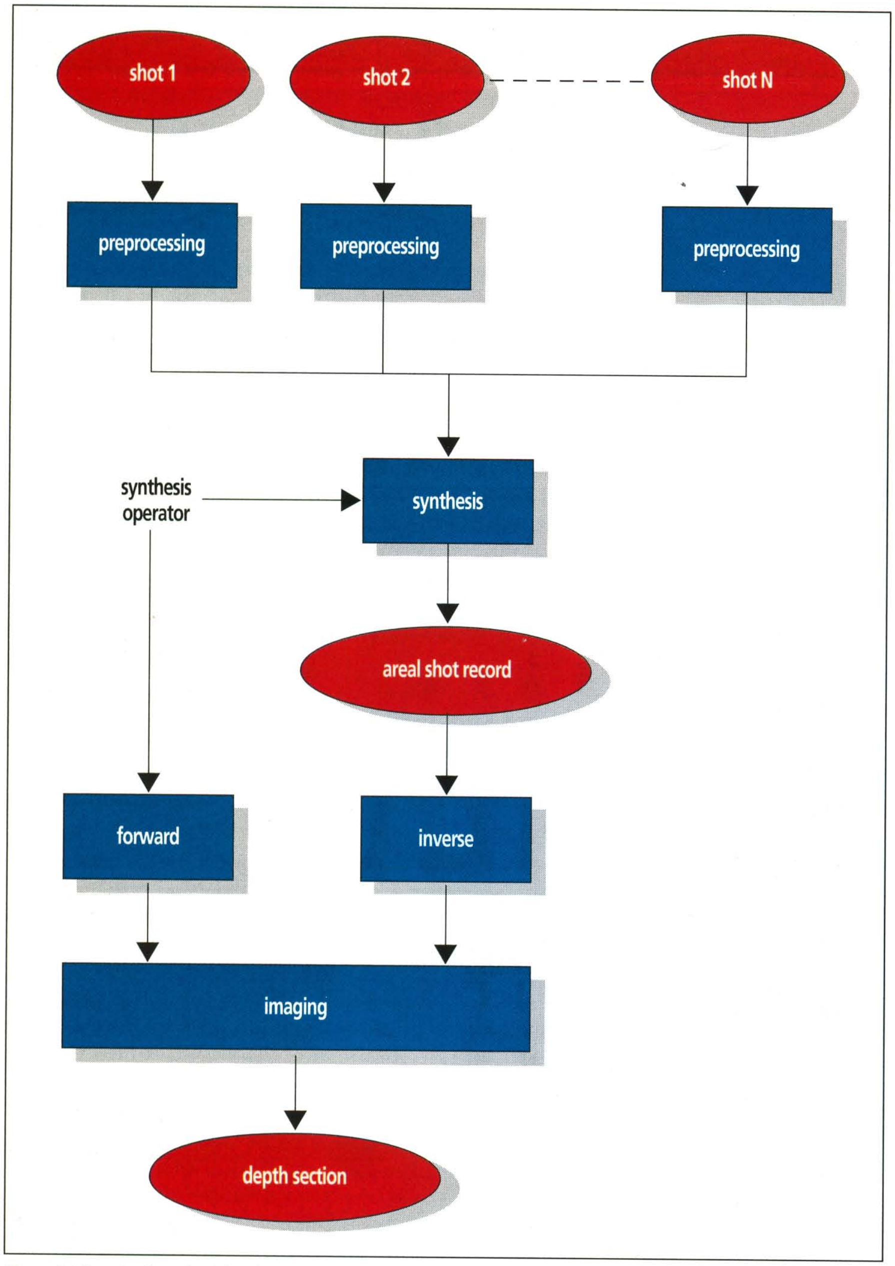


Figure 3 (above). The principle of areal shot record migration. The synthesis operator represents the areal source. Both the forward and inverse wave field extrapolation process occurs in the (x, y, ω) domain by spatial convolutions.

Table 2a (below). Data volumes in 3-D seismic acquisition. Table 2b (below right). CPU times in 3-D seismic processing.

data unit	number of shots	number of channels/shot	total data volume
trace	1	1	8 kbyte
record	1	10^3	8 Mbyte
line	10^3	10^3	8 Gbyte
survey	10^6	10^3	8 Tbyte

data unit	number of flops	number of flops/s	total CPU time
trace	10^9	giga	1 second
		tera	1 millisecond
record	10^{12}	giga	17 minutes
		tera	1 second
line	10^{15}	giga	283 hours
		tera	17 minutes
survey	10^{18}	giga	31 years
		tera	283 hours

Kirchhoff operators,[2] each operator being computed by a ray-tracing process in the macromodel

- Recursive, in which the operators $\mathbf{W}^+$ and $\mathbf{W}^-$ are represented by local, frequency-domain Kirchhoff operators,[3] each operator being defined directly by a local macrovelocity value that occurs in the macromodel

More specifically, nonrecursive methods make use of travel time tables (from each surface grid point to each subsurface gridpoint), and recursive methods make use of an operator table (the entry being $k = (\omega/c)$). Assuming precomputed tables, nonrecursive time-domain methods are faster than recursive frequency-domain methods. This explains why the industry has opted for 3-D prestack migration algorithms that use nonrecursive time-domain extrapolation operators. However, apart from the theoretical property that monochromatic wave theory operators are potentially more accurate than broadband ray theory operators, recursive methods also have a significant practical advantage over nonrecursive methods. This becomes clear if one considers that the macromodel must be updated many times before the migration result is optimal. For nonrecursive methods, this means recomputation of the travel time tables and the operators each time. For recursive methods, a new macromodel does not mean that the operator table has to be recomputed; only the order in which the operators will be chosen during the extrapolation process is changed. Actually, one could generate a few standard tables with carefully optimized operators and use those for all migration jobs.

Alternative data configurations

If we consider a 3-D seismic survey consisting of N surface grid points with M traces per grid point, then the total data volume amounts to NM traces, and the data matrix contains NM complex-valued Fourier elements. In straightforward prestack migration, this entire data volume is migrated by "numerically lowering the sources and detectors into the subsurface" with the aid of downward extrapolation (4b). By dividing the entire dataset into subsets, such as common shot gathers or common offset sections, the individually migrated subsets can be inspected for consistency

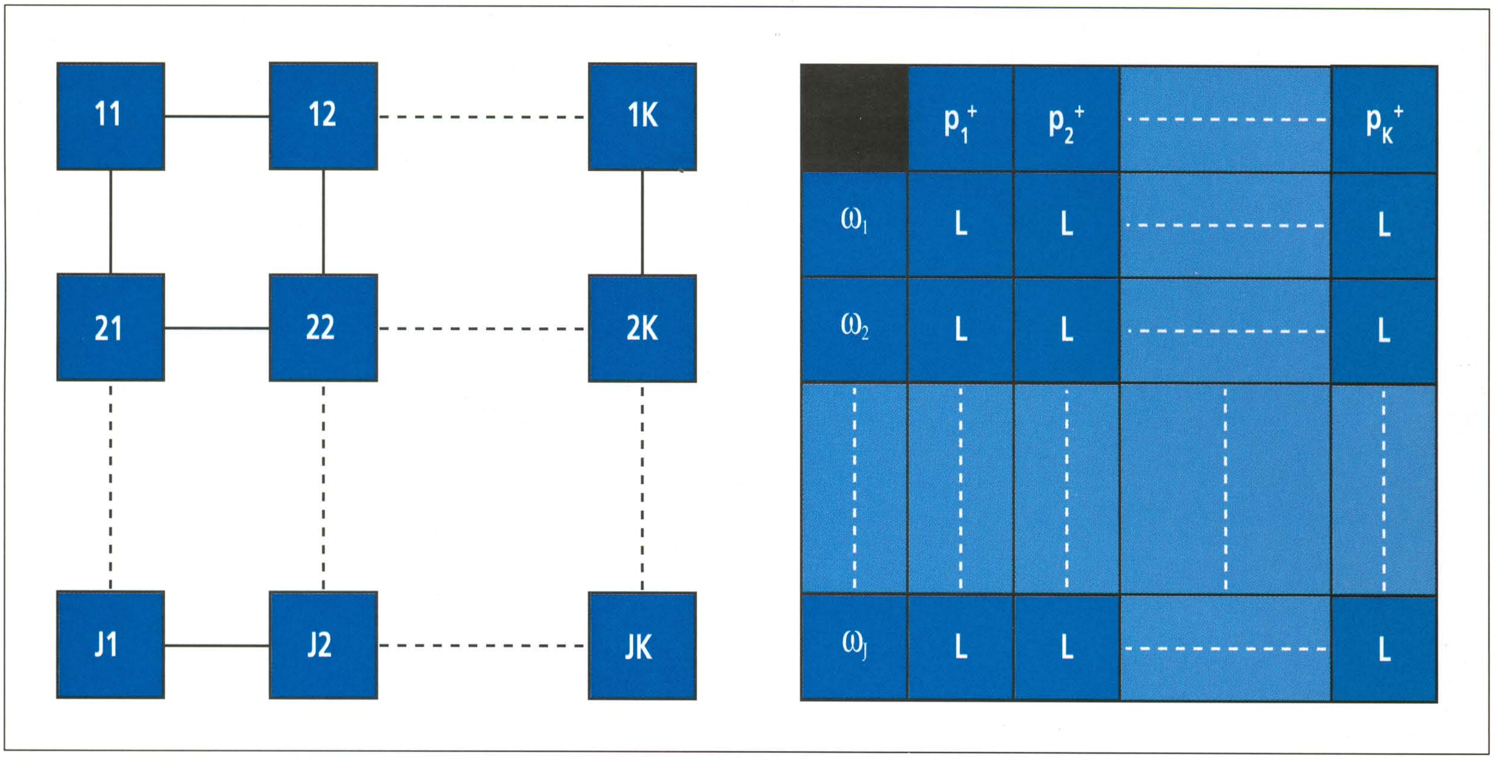


Figure 4 (above). The MPP subsystem architecture in terms of *J* rows and *K* columns, allowing a double parallelization of the migration process. Each of the *JK* processors contains *L* complex-valued data samples.

and optionally combined in some data-adaptive way. Before applying the gigantic computational process on those subsets, consider the following transformations:[4]

- ☐ All point sources are combined into *K* areal sources (synthesis)
- ☐ Each areal source response is resampled to *L* data samples (regularization)

Next, each regularized areal source response is independently migrated by a full 3-D areal shot record migration process as shown in Figure 3. If we choose monochromatic plane waves for the areal sources, then the original *NM* complex-valued Fourier elements (generally on an irregular spatial grid) are transformed into *KL* complex-valued data samples (defined on a regular spatial grid).

For most geologically realistic macrovelocity models a plane wave does not stay a plane wave during downward extrapolation. The downward extrapolation algorithm must honor this important property.

The MPP procedure

MPP systems have become necessary to seismic migration because the large volumes of data acquired (Table 2a) require higher computing capacities (Table 2b). If we consider the MPP subsystem as a matrix of processors, then we will assign one areal source to the processors in one column, and we will assign one frequency component to the processors in one row (Figure 4a). Following the plane wave example, the local memory of each processor should contain

- ☐ A data table (Figure 4b), consisting of *L* complex numbers (unique for each processor) and
- ☐ An operator table (Table 3a), consisting of several hundred operators (unique for the processors of one row, independent of the macrovelocity model)
- ☐ A reference table (Table 3b), linking each subsurface gridpoint to an appropriate position in the operator table (the same for all processors, depending on the macrovelocity model only).

When a seismic trace is read and Fourier transformed by the host system, each frequency component is transferred to the assigned MPP processor. When the frequency components of the next trace arrive, each processor can already start the regularization and the synthesis process. Hence, when the last trace of the 3-D data set has arrived, all regularized areal source responses are available for migration (Figure 5).

Next, using the operator table and the reference table, the (x, y, ω) recursive areal shot record migration process will start according to Figure 3. At each depth level, imaging means a weighted average of the extrapolation results of the *J* processors in one column. We end up with *K* migration results per depth level. Those results are directly available for display. Figure 6 is an example of such a display. For a given *y*-value an *x*-*z* cross section is shown. If all events line up along the p^+ coordinate, the macrovelocity model is correct, and the *K* migration results can be superimposed without any adjustment.

The proposed 3-D MPP procedure has been illustrated with plane wave areal sources. This is just one of several attractive alternatives. In particular, illumination of local spots by focused areal sources may generate new solutions in the areas of velocity analysis and AVO inversion. Note that the processes in each MPP processor are fully independent of the others (coarse grain parallelism). Note also that if the macrovelocity model is changed, only the reference table need be changed. Current key issues include the number and distribution of temporal frequencies and areal sources.

The previous discussion demonstrates that major improvements in seismic imaging technology are to be expected. These improvements will require a tremendous increase in computational power. New algorithms, as discussed in this article, need to be evaluated and fine-tuned today in order to perform tomorrow. The yield may be significant: true amplitude 3-D prestack migration algorithms that also include velocity estimation may advance the quality of the subsurface images beyond expectation.

Conclusions

Significant developments in seismic migration are based on 3-D integral operators and

ω_j/c	operator
ω_j/c_{min}	x x - - - - x
⋮	⋮
ω_j/c_{max}	x x - - - - x

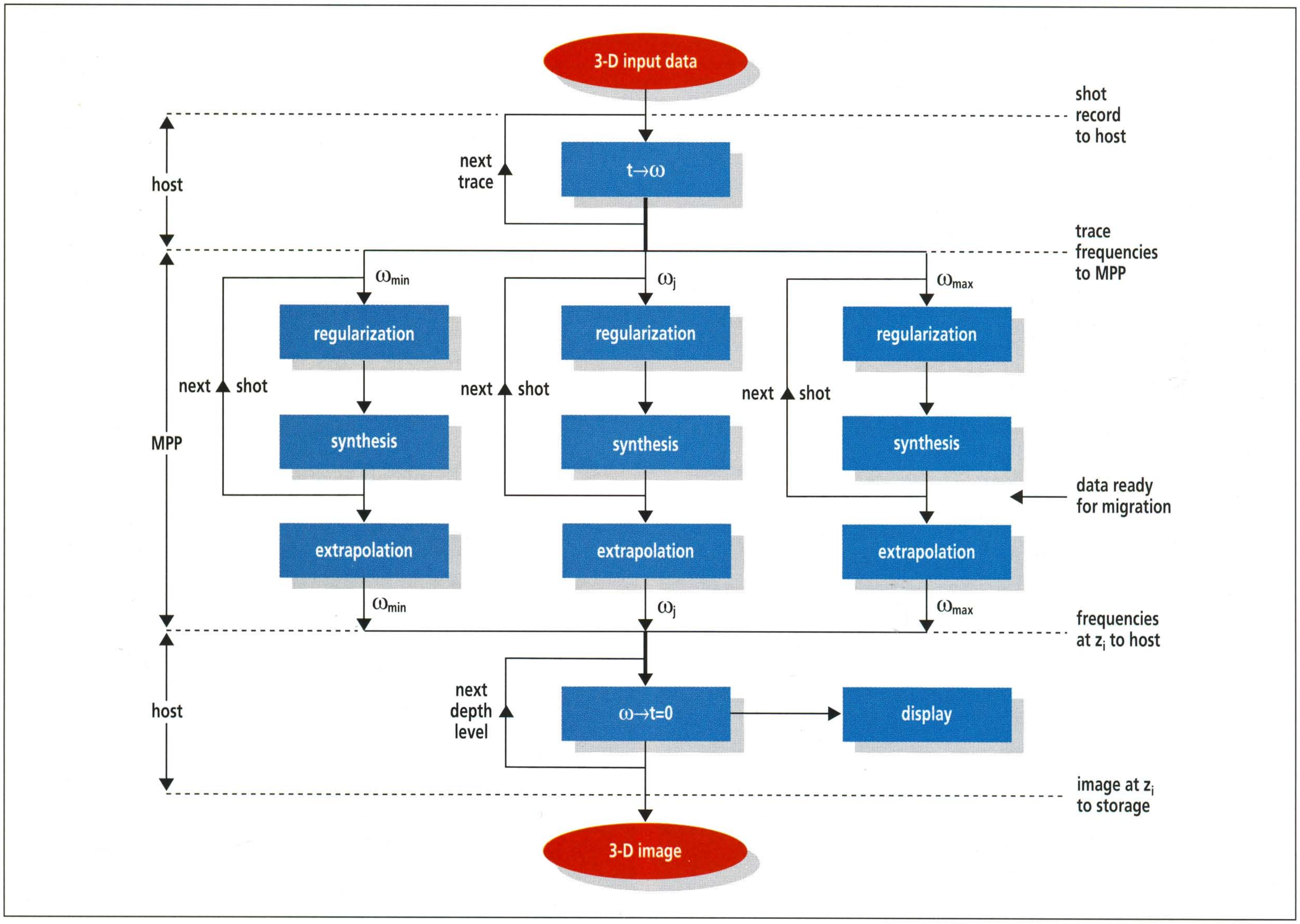
a

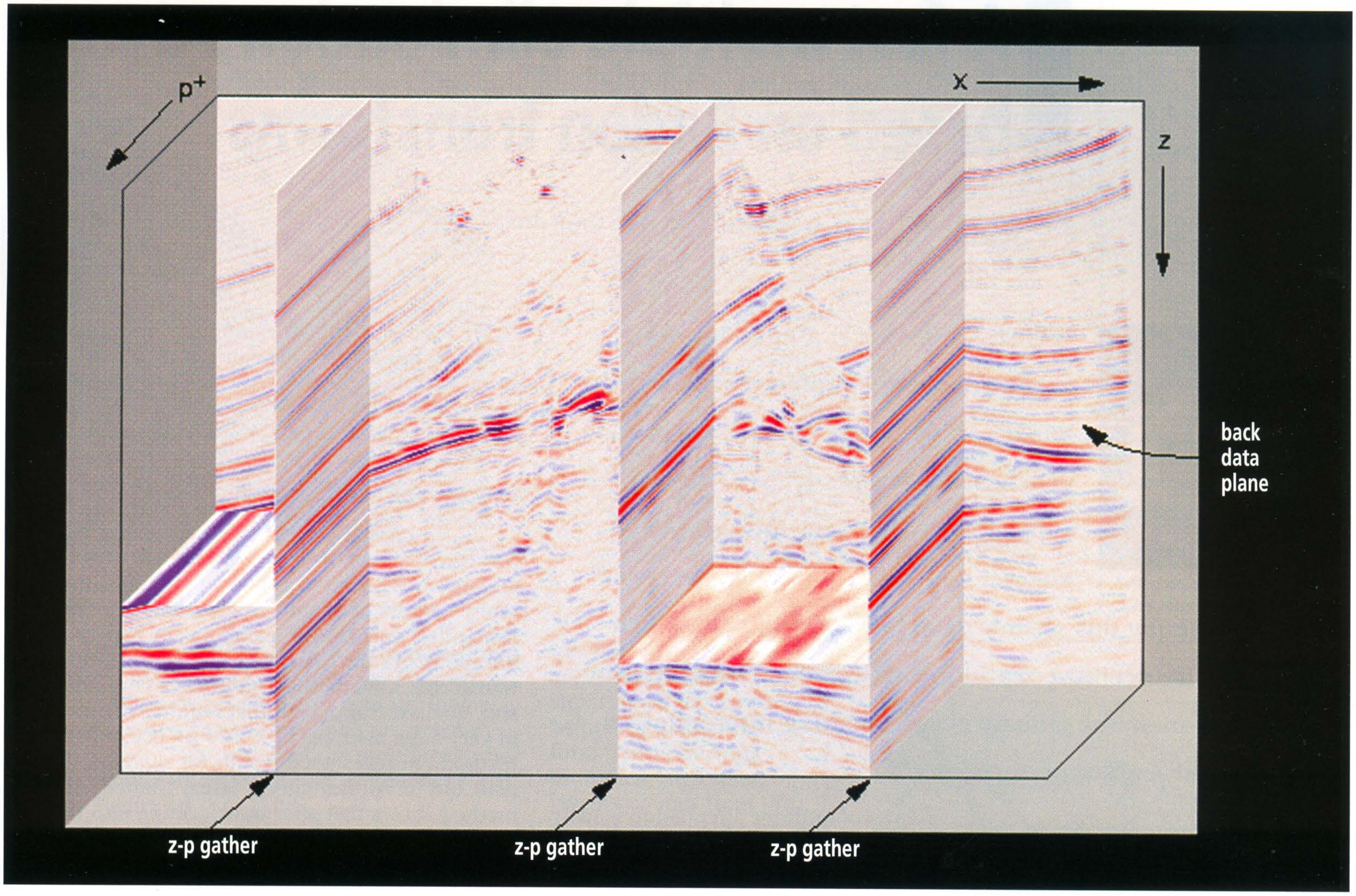
gridpoint	velocity
1	c_1
2	c_2
⋮	⋮
last	c_{last}

b

Table 3 (left). In addition to the data table, each processor contains a monochromatic operator table and a reference table. The reference table links each gridpoint to the proper operator. Note that if the macrovelocity model is changed, then only the reference table need be changed.

Figure 5 (below). MPP processing for 3-D prestack depth migration. Here, the processing diagram is shown for one column of processors only (each column migrates the frequency components of one areal source response). At the end of the synthesis process each processor contains *L* regularized data samples, and the actual "double" parallel migration process can start.





can be divided into two main categories: nonrecursive time-domain and recursive frequency domain. Nonrecursive time-domain methods use broadband ray theory operators and are therefore economically attractive on current vector computers and popular in the industry. Recursive frequency domain methods use monochromatic wave theory operators. Theoretically these operators are superior to ray-tracing-based operators but, unfortunately, require long processing times on current vector computers.

It is the author's opinion that, rather than concentrate on the conservative, ray-tracing-based, space-time algorithms only, the industry should make significant investments in the long-term solution given by the parallel, wave-theory-based, space-frequency algorithm, whereby

- Each row of processors is responsible for one frequency component, and each column of processors is responsible for one areal source.
- Macromodel-dependent travel time tables are replaced by a macromodel-independent operator table.

Future imaging solutions should integrate seismic migration and macromodel estimation. The use of parallel operator tables and reference tables in the MPP subsystem, together with fast display facilities in the host system, facilitates user-controlled updating of the macromodel during the migration process at each depth level.

Figure 6. Cross section (x-z) of migrated areal shot records for several p^+ values (plane wave areal sources). For this example the macromodel is correct: all events line up along the p^+ coordinate.

Partnerships between petroleum companies, university research centers, and the suppliers of leading edge computational tools are critical if we expect to keep pushing the limits of seismic imaging.

About the author

A. J. Berkhout heads the Seismics and Acoustics Laboratory at the Delft University in the Netherlands. His international experience includes 12 years with Royal Dutch Shell. In the last 15 years, he has written several books and hundreds of papers on the subject of imaging. He is director of the successful DELPHI consortium and serves as a consultant to the petroleum industry. In 1990, he became a member of the Royal Netherlands Academy of Arts and Sciences. In 1993, he was appointed an honorary member of the Society of Exploration Geophysicists.

References

1. Claerbout, J., *Fundamentals of Geophysical Data Processing,* McGraw-Hill, New York, 1976.
2. Schneider, W. A., "Integral formulation for migration in two and three dimensions," *Geophysics,* Vol. 43, No. 1, 1978.
3. Berkhout, A. J. and D. W. Van Wulfften Palthe, "Migration in terms of spatial deconvolution," *Geophysical Prospecting,* Vol. 27, No. 1, 1979.
4. Berkhout, A. J., "Areal shot record technology," *Journal of Seismic Exploration,* Vol. 1, No. 2, pp. 251-264, 1992.

A fast multiple elimination scheme based on Radon transforms

Panos G. Kelamis, Saudi Arabian Oil Company, Saudi Arabia

In the search for underground oil, which requires accurate representations of the earth's structure, oil company scientists rely heavily on seismic data to locate oil deposits, determine their size, and suggest how they can best be harvested. When gathering this seismic data, however, scientists often encounter the problem of multiple reflections. These reflections distort and mask the primary events, making interpretation not only difficult but erroneous.

Techniques of multiple attenuation usually are based on the moveout difference between primaries and multiples (velocity discrimination), dip difference, periodicity of the multiples, and difference in the frequency content between primary and multiple events. Another approach to surface multiple elimination is based on the auto-convolution of the data itself.[1] For datasets with a velocity difference between primaries and multiples, Radon transform-based multiple suppression schemes are gaining popularity over the conventional predictive deconvolution and frequency-wavenumber approaches. The increased resolution in the model space obtained via the forward Radon transform is the key element in the multiple elimination process. Thus, by inverse-mapping the multiples and subtracting them from the original common-midpoint (CMP) gather, a primaries-only gather is obtained. This article discusses the concept of Radon transform, outlines a simple, cost-effective methodology for multiple attenuation on a CRAY Y-MP 8E system, and demonstrates its effectiveness with field datasets.

Concept and methodology

Although Radon transform is a data-decomposition scheme, in terms of mathematics it is commonly referred to as a mapping transformation. Many authors have studied the problem of mapping time-offset data (linear, parabolic, or hyperbolic) into points in the transform space (forward Radon transform). One method proposes a compute-intensive stochastic approach that requires the inversion of very large matrices (100,000 by 100,000).[2] This method produces good results on synthetic and real data, but its applicability is rather limited. Another method shows that after transforming into the frequency domain, the Radon operators can be computed for each frequency by a simple linear inversion involving small-sized matrices.[3] This method requires the application of normal-moveout (NMO) prior to the forward Radon transform and directly maps time-offset parabolae into the time-moveout plane. The NMO correction can be avoided by applying a t-squared stretch and mapping the data into the time-velocity plane.[4] Figure 1 depicts the concept of multiple elimination via Radon transforms using a synthetic NMO-corrected CMP gather.

In Figure 1a the input gather is shown with two primary (flat) events and two multiples (curved). The same gather after Radon transform is shown in Figure 1b. Note the separation of events in the model space, in which the linear and curved

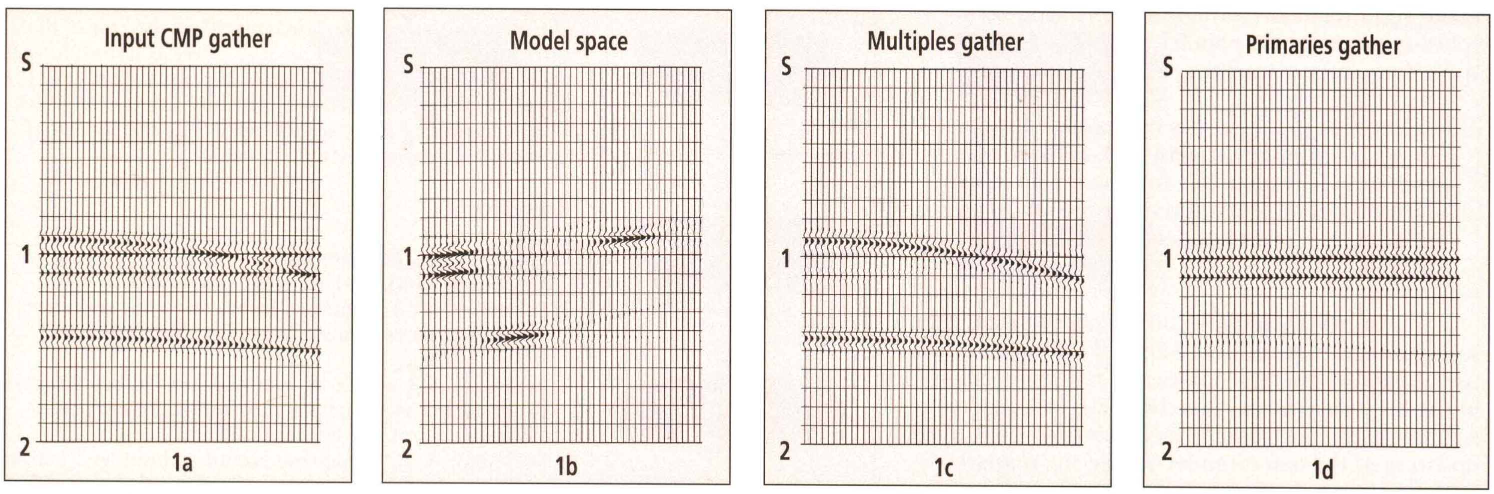
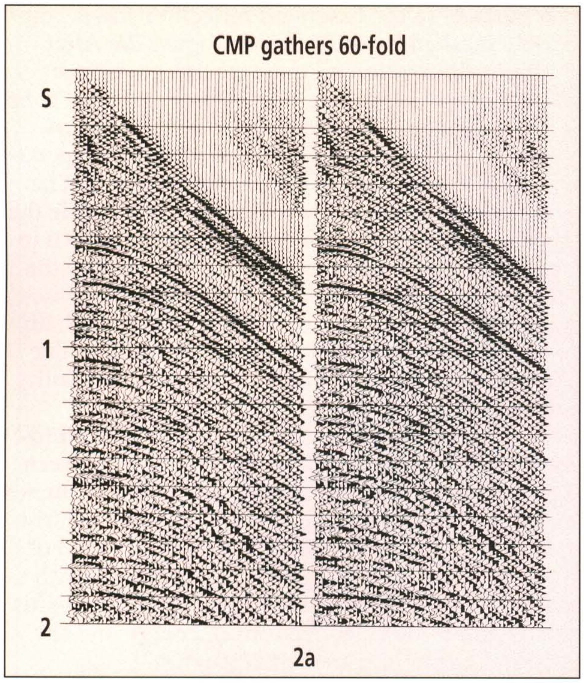
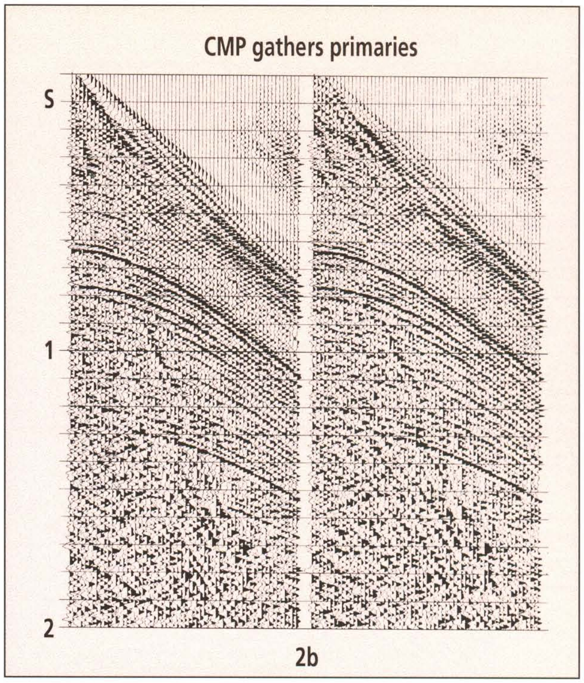


Figure 1. Concept of Radon-based multiple elimination. (a) Input CMP gather after NMO with two primary events (flat) and two multiples. (b) The result of the forward Radon transform; the events are separated. (c) The estimated multiples. (d) The primaries-only gather obtained by subtraction of Figure 1c from Figure 1a.





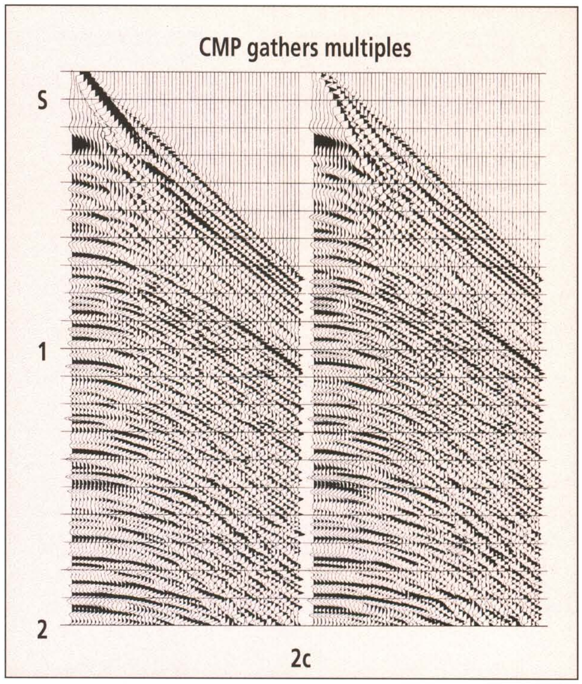


Figure 2. (a) CMP gathers from land data, which show the coherent hyperbolic multiple energy that dominates the record at later times. (b) The estimated multiples after the application of Radon transform technology. (c) Primaries-only gathers obtained by subtraction of Figure 2b from Figure 2a; the enhancement of the primary reflection is evident at about 1.260 seconds.

events of the input gather are now transformed into well-defined foci. The gather shown in Figure 1c is obtained after the primary events in the Radon domain are muted and the multiples are inverse-mapped. Finally, subtraction of the multiples gather from the input data yields the primaries-only CMP gather in Figure 1d.

In practice, the computation of the forward Radon operator involves the inversion of a complex matrix for each frequency of interest for every CMP gather. The size of the matrix is directly related to the fold of the data and the number of velocities (moveouts). Additionally, the values of the complex matrix depend only on the acquisition geometry (offsets). Thus, for a regular survey geometry, such as for marine data, the matrix inversion per frequency is performed only once for the first gather, saved, and then applied to the data. For land data, however, missing traces and irregular geometry dictate that each Radon operator per frequency be computed independently for each gather. For high-fold datasets, this matrix inversion can be quite slow. Consequently the whole Radon multiple-elimination process becomes very expensive and thus impractical to implement in a routine seismic data processing flow.

This problem is overcome by introducing a preprocessing step prior to the application of the forward Radon transform.[5] In this step, the multiplicity (fold) of the data is reduced, and the offset distribution (geometry) is regularized. By reducing the fold, the task of matrix inversion is performed very efficiently. The geometry regularization enables each Radon operator per frequency to be computed only once (for the first gather). The computed Radon operators for all frequencies are then kept in memory. Thus, application of the Radon transform on the subsequent gathers involves just a complex matrix-vector multiplication operation. Partial stacking of CMP gathers is used to obtain fold reduction and geometry regularization.[5] This technique points out the necessity of applying a noise suppression procedure before the Radon transform when removing multiples. High noise levels also may limit the resolution of events in the transform domain and thus degrade the performance of the Radon multiple elimination process. Partial stacking not only offers fold reduction and geometry regularization, it improves the signal-to-noise ratio, resulting in better-focused events in the model space.

The Radon-based multiple elimination methodology described previously is implemented on the Saudi Arabian Oil Company's CRAY Y-MP8E system. The input data consists of CMP gathers with or without NMO correction since both methods can be applied. After partial stacking, a singular value decomposition (SVD) algorithm is used for the complex matrix inversion during the forward Radon transform for every frequency but only for the first gather. The Radon operators for all frequencies are then stored in memory and applied on the subsequent CMP gathers as very fast matrix-vector multiplications. Depending on the desired fold, after partial stacking (usually 30 or 60), the number of time samples per trace, and the number of input gathers, the whole Radon multiple elimination process is performed in a few minutes on the Cray Research system.

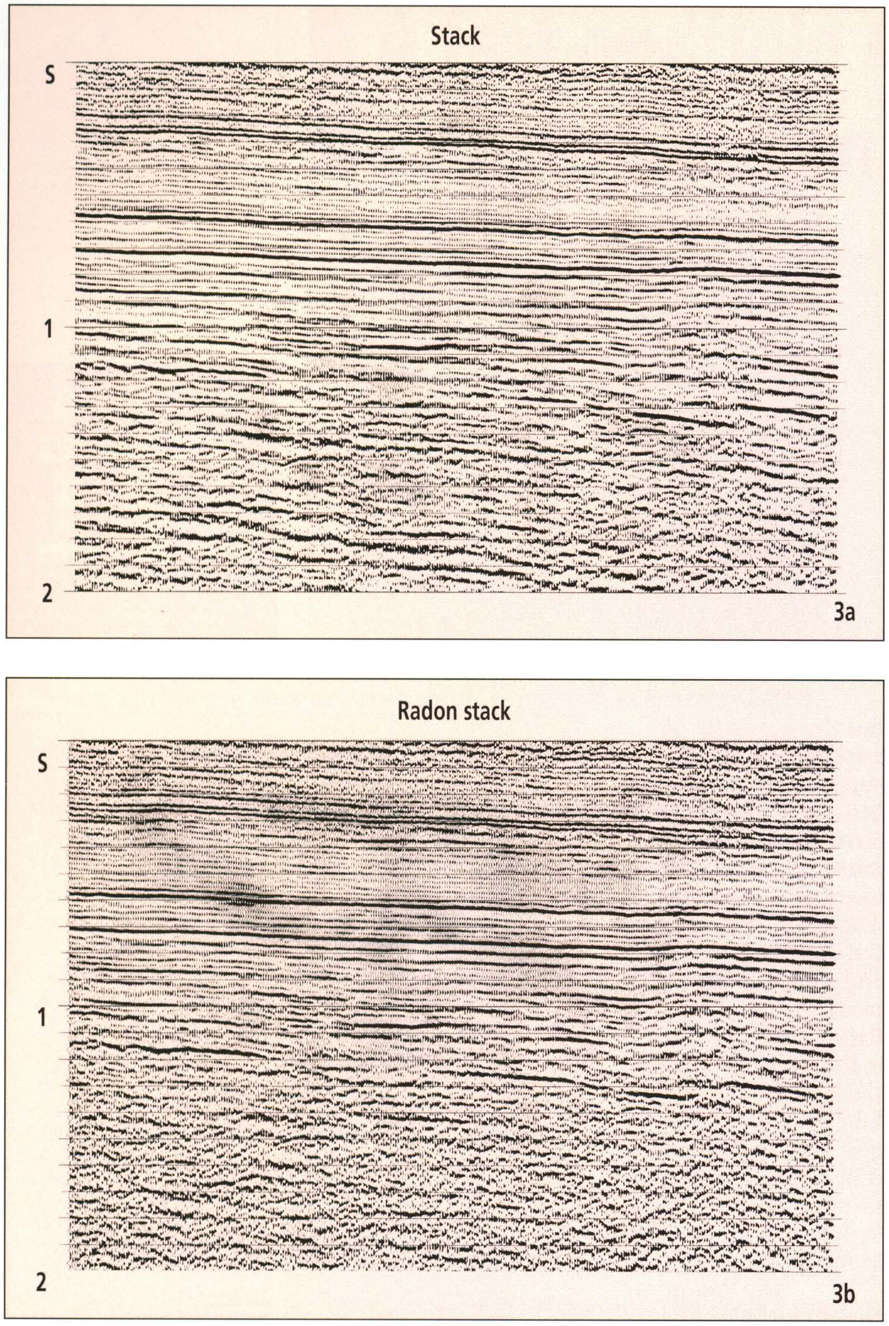


Figure 3. (a) Conventional stack with no multiple elimination. (b) The same data as in Figure 3a after Radon-based multiple elimination; the basement reflection is now clearly identified.

Field data examples

Next, the performance of the Radon-based multiple elimination scheme is shown with two land datasets from the Arabian peninsula. In the first example, Radon multiple elimination improves the definition of the basement, which is associated with faulting and masked with multiple energy. The input CMP gathers are depicted in Figure 2a, showing how the coherent hyperbolic multiple energy dominates the records at later times. The estimated multiples are shown in Figure 2b. The low frequency at early times is due to a t-squared stretch during the forward transform. The primaries-only gathers are obtained by subtraction and are shown in Figure 2c, highlighting the reduction of the multiple energy and the enhancement of a primary reflection at about 1.260 seconds. This primary corresponds to the basement reflection and is completely masked by multiples in Figure 2a. After multiple elimination all that remains are primary reflection energy and the noise cone. The noise cone will be attenuated by stacking while the primary reflection will be further enhanced. The corresponding stacked sections are depicted in Figure 3. The conventional stack is shown in Figure 3a, while the stack with Radon multiple elimination is shown in Figure 3b. The increased resolution offered by the Radon stack leads to a better mapping of the basement reflection which now is clearly identified, dipping regionally from left to right. In addition, due to the reduction of multiples, the basement faults now can be traced easily.

The second case of Radon multiple elimination is shown in Figure 4. The target is between 2.0 and 2.3 seconds. The presence of strong multiples in the conventional stack (Figure 4a), particularly at the right part of the section, makes any kind of interpretation highly suspicious. The stack with Radon multiple elimination (Figure 4b), highlights the overall reduction of multiple energy and the clear definition of the event of interest.

Conclusions

This article describes a methodology for multiple attenuation based on Radon transforms and its implementation on a CRAY Y-MP 8E supercomputer. By combining partial stacking and forward/inverse Radon transforms, the whole process of multiple elimination is equivalent to complex matrix-vector multiplication. Thus, a very fast and practical algorithm is readily available for use in a routine seismic data processing flow. The performance of this algorithm, implemented on the Saudi Arabian Oil Company's CRAY Y-MP 8E system, was demonstrated with two field datasets from the Arabian peninsula.

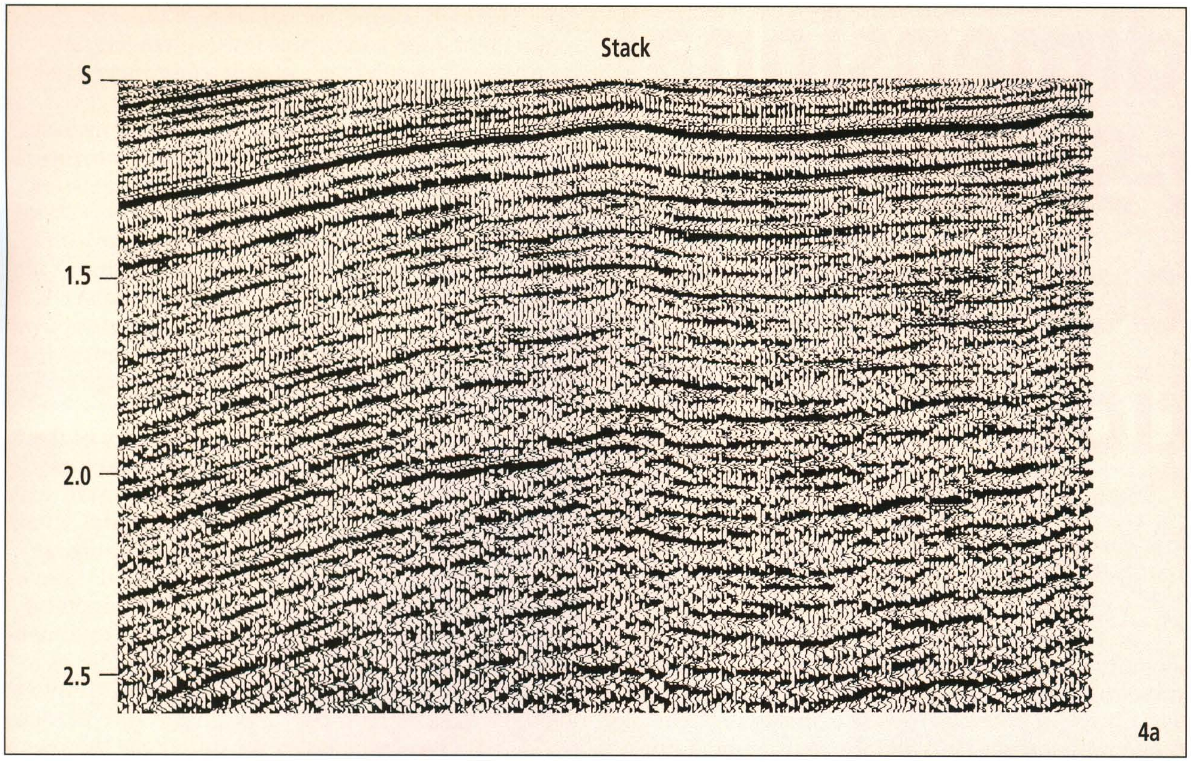
This Radon-based multiple attenuation scheme can easily be combined with the surface-related multiple elimination approach.[4] Successive application of surface-related multiple elimination and the Radon-based scheme should result in multiple-free seismic sections with clear subsurface images.

Acknowledgments

The author thanks the Saudi Arabian Oil Company for encouraging this work and granting permission to publish this paper.

About the author

Panos G. Kelamis received a B.S. degree (Hon) in physics from the University of Athens, Greece in 1977. He received MSC and DIC degrees in geophysics from Imperial College of the University of London. In 1982 he was awarded a Ph.D. degree in geophysics from the University of Alberta in Edmonton for his research in generalized ray theory and seismic attenuation. After graduation, Kelamis joined the research and development group of Western Geophysical in Houston as a senior research geophysicist. He continued his career with Dome Petroleum in Calgary working in



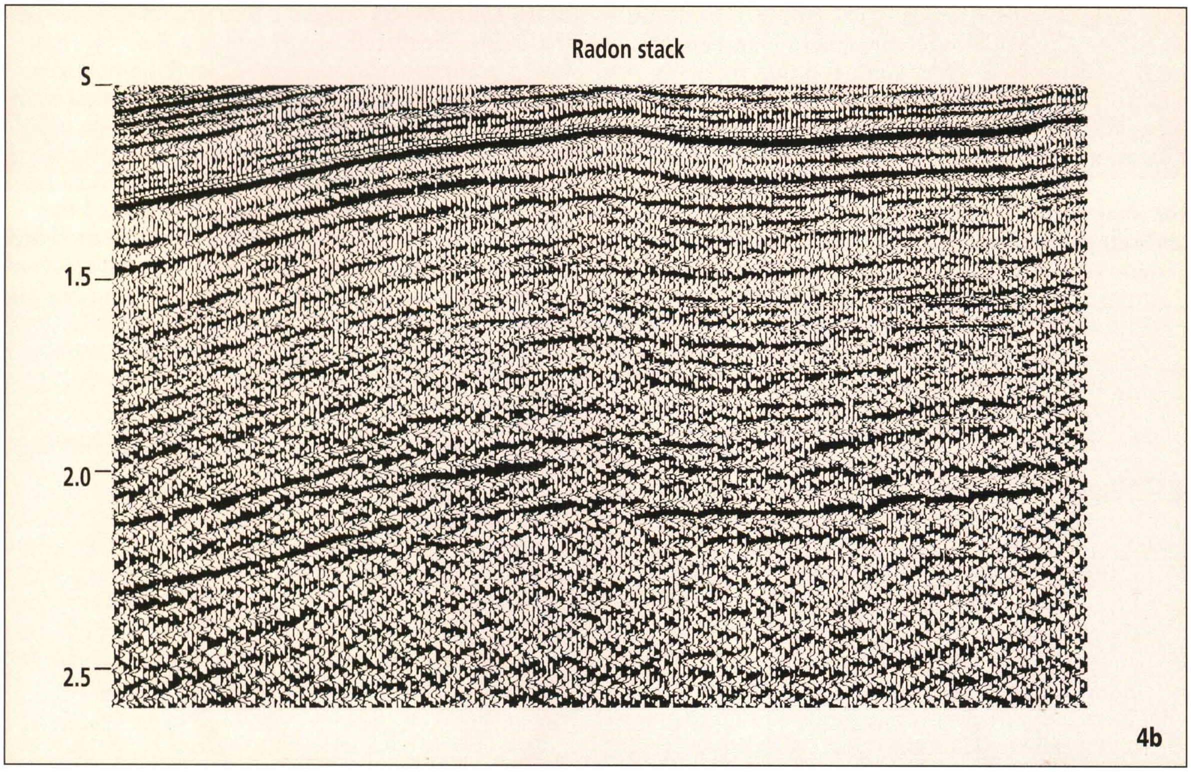


Figure 4. (a) Conventional stack with no multiple elmination; the presence of strong multiples dominates the section. (b) The same data as in Figure 4a after Radon-based multiple elimination. The resolution has improved in the zone of interest (2.0 to 2.3 seconds) and the overall reduction of the multiple energy is highlighted.

migration and imaging techniques. In 1985 he joined the research and development division of the geophysical department of the Saudi Arabian Oil Company in Dhahran where he is currently head of the field development unit.

References

1. Verschuur, D. J., A. J. Berkhout, and C. P. A. Wapenaar, "Adaptive surface-related multiple elimination," *Geophysics,* Vol. 57, pp. 1166-1177, 1992.
2. Thorson, J. R. and I. F. Claerbout, "Velocity-stack and slant-stack stochastic inversion," *Geophysics,* Vol. 50, pp. 2727-2741, 1985.
3. Hampson, D., "Inverse velocity stacking for multiple elimination," *Journal of the Canadian Society of Exploration Geophysicists,* Vol. 22, pp. 44-55, 1986.
4. Yilmaz, O., "Velocity-stack processing," *Geophysical Prospecting,* Vol. 37, pp. 357-382, 1989.
5. Kelamis, P. G., E. F. Chiburis, and S. S. Shahryar, "Radon multiple elimination, a practical methodology for land data," *Expanded Abstracts,* 60th Meeting, Society of Exploration Geophysicists, pp. 1611-1613, 1990.

Three-dimensional Remez-Soubaras poststack depth migration

Jean-Yves Blanc, Irène Huard, Compagnie Générale de Géophysique, Massy, France
André Gomez, Roland Piquepaille, Cray Research France, Les Ulis, France

Three-dimensional poststack depth migration provides an accurate image of the earth's subsurface but requires compute-intensive codes with heavy I/O. Researchers at Compagnie Générale de Géophysique (CGG) and the petroleum support team of Cray Research France wanted to process very large datasets (typically exceeding 1.5 million traces) in a production environment—and ensure high performance. They achieved this by using CGG's GEOVECTEUR seismic package to implement the 3-D Remez-Soubaras poststack depth migration algorithm on a CRAY C90 supercomputer.

Recently developed algorithms improve the imaging of complex geologies (steep dips, faults, and salt domes, for example) in the earth's subsurface. The high computational costs of these methods make them good candidates for implementation on high-end parallel vector supercomputers, such as CRAY C90 systems.

This new class of 3-D algorithms includes F-XY explicit poststack depth migration schemes, which downwardly extrapolate wavefields using spatial convolution with finite-length filters. These filters usually are computed with nonlinear least-squares-based methods or Taylor expansions (Figure 1). In the 3-D Remez-Soubaras poststack depth migration method, these filters are computed by variants of the Remez exchange algorithm, which is faster than most previous methods and optimal for the L-∞ norm.

In Figure 1, the circular symmetry of the operator in the wavenumber domain is used to synthesize the operator in the space domain by a Chebychev recursion using the 17-point McClellan 2-D convolutional operator. The Remez-Soubaras method replaces the 17-point McClellan 2-D convolutional operator by a 17-point Laplacian operator in the shape of a cross—the sum of two one-dimensional second-derivative operators. This allows a more precise synthesis and a more efficient implementation, as no true 2-D convolutions are needed.

Figure 2 illustrates the difference between the Remez-Soubaras and a conventional split operator on a horizontal slice of a 3-D operator response. The Remez-Soubaras algorithm is a near-optimal way to compute operators to achieve the necessary accuracy up to a given specific dip. Compared with less accurate conventional methods, like finite difference schemes, the higher computational costs are kept within a reasonable limit, especially considering the much higher quality images obtained for complex tectonics. In Figure 2, the salt wall modeled

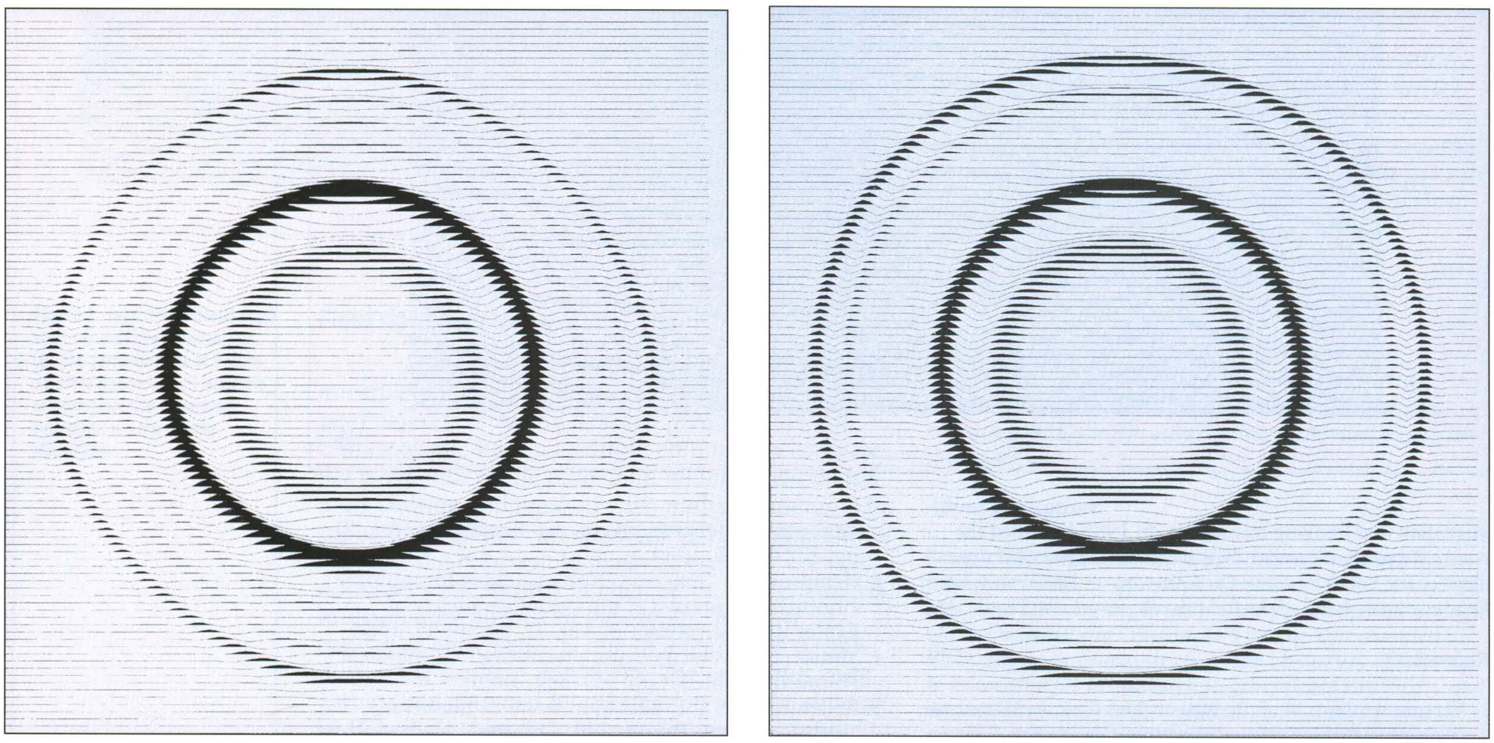


Figure 1. Conventional split operator (left) and Remez-Soubaras operator (right). Comparison of horizontal slices of 3-D operator response. A 30° (inner ring) and a 60° (outer ring) depth slice are superimposed. Anisotropy is evident in the conventional approach where the theoretical circular response of the migration operator is distorted at 45° and 135°.

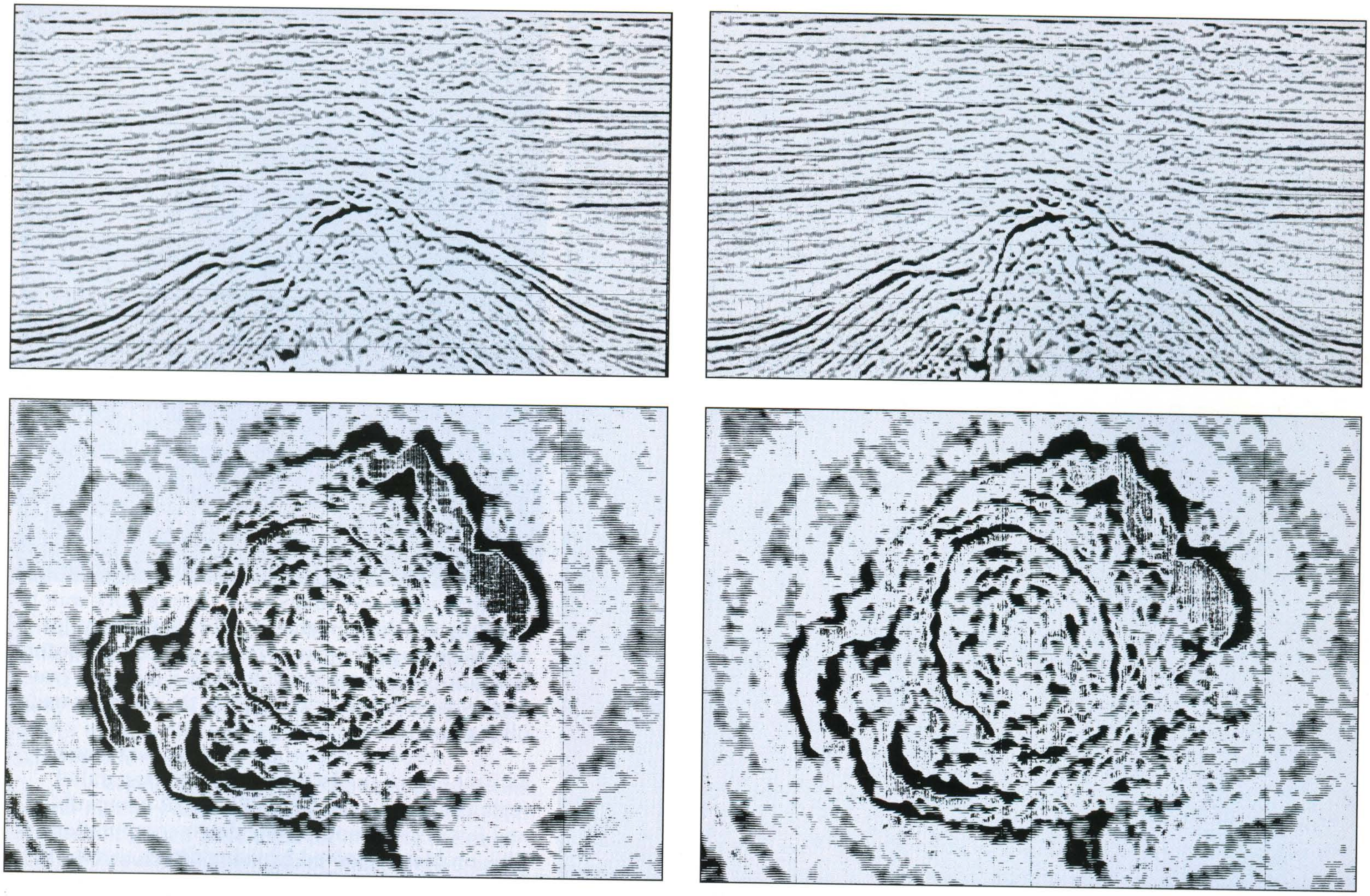


Figure 2. 3-D depth migration example on a steeply dipping salt dome. Conventional split operator (left) and Remez-Soubaras operator (right).

by a conventional split method leaves the steep flanks of the salt dome undermigrated. This new method produces much better images. Figure 3 shows a 2-D section from the final production migration produced on the CRAY C90 system, with the velocity model superimposed in color.

Implementation on the CRAY C90 system

For real-size 3-D datasets, main memory cannot support an in-core scheme, even on a fully configured CRAY C90 supercomputer. A realistic migration code must be implemented using an out-of-core scheme. Effective parallel implementation has been achieved with 16 CPUs on a CRAY C90 system and Cray Research's SSD solid-state storage device. Large-grain parallelism is used to accommodate the parallel model and to keep the dual-pipe vector units as busy as possible. This also allows for the design of a very efficient I/O strategy.

Autotasking, Cray Research's automatic parallel processing capability, was chosen for several reasons: portability and maintenance (directives are inserted into the code to increase both readability and ease of use on other machines), the ability to run batch jobs in a production environment, and no modification of GEOVECTEUR management of memory.

The image is obtained by extrapolating in depth the wavefield for every frequency. In other words, each frequency is contributing independently to the imaged volume. The obvious strategy for parallelizing this algorithm is to distribute frequencies among the processors, which perform computations independently, as described in Figure 4 (*nz* being the number of extrapolation steps, *nfreq* the number of frequencies, and *njump* the number of interpolation steps between two extrapolations).

While it is a sophisticated algorithm, the computation of the set of migration operators does not represent a significant runtime; it typically accounts for less than one percent of the elapsed time for a migration on a real dataset (even less for a very large survey). Therefore, it is not critical to run these computations concurrently, even if this part of the algorithm intrinsically presents large-grain parallelism opportunities.

The loop over frequencies is distributed among the processors where computations are done independently. To obtain high performance and avoid memory contentions due to highly referenced shared variables, most of the variables are private (leading to a very demanding memory use). Due to problem size and the corresponding out-of-core strategy, frequencies are read and written back to disk. Fortunately, those I/Os can be done in parallel, allowing full exploitation of the CRAY C90 system's powerful I/O capabilities.

For a given depth, the imaging stage consists of each frequency adding its contribution to the image slice. Because frequencies are distributed among the processors, each image update must be

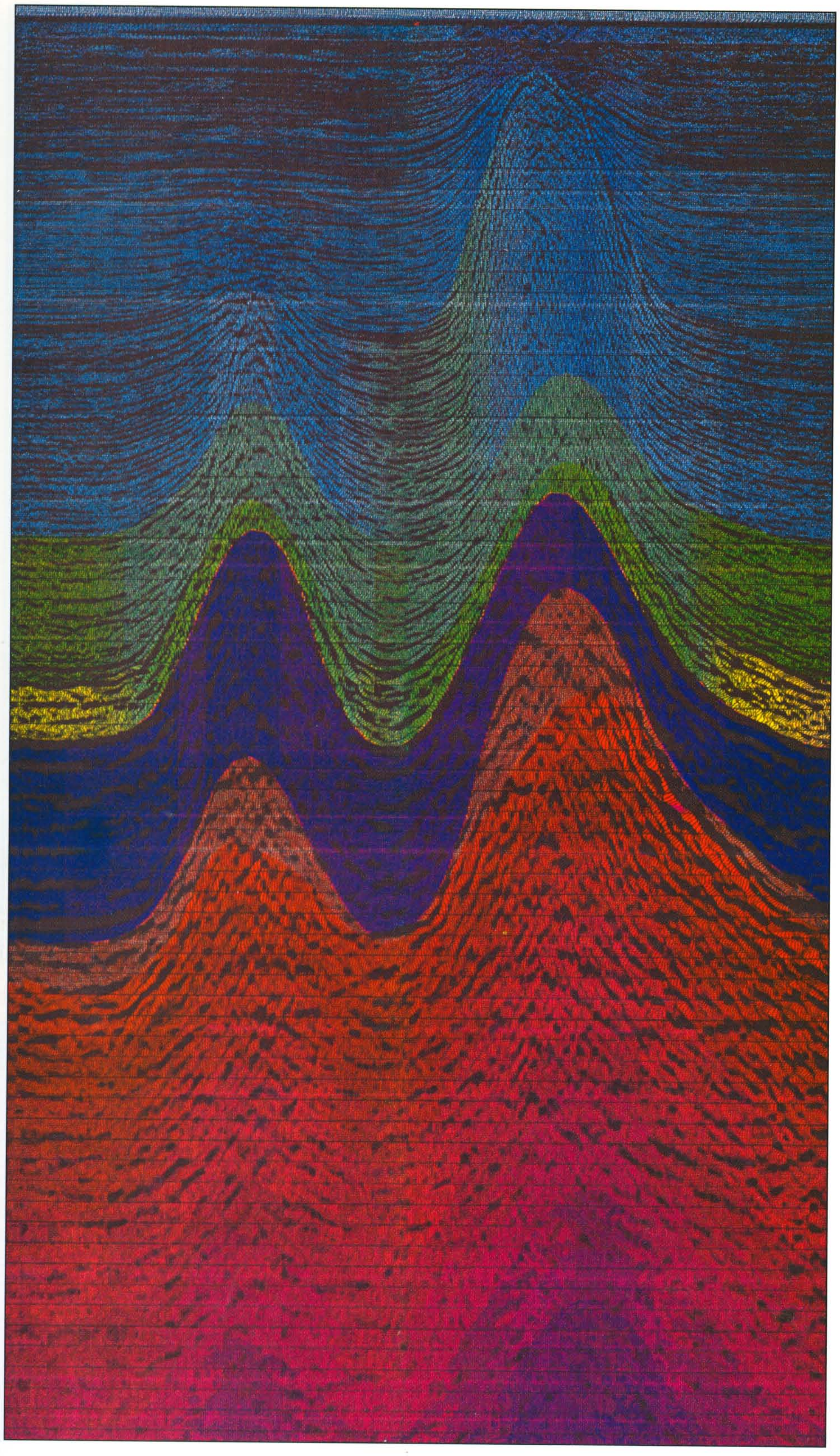


Figure 3. Migrated section.

operated in a guarded zone to assure data coherency. This critical part of the code demands a very fast device to handle those sequential I/Os. Cray Research provides this I/O support in the SSD, which was used as a secondary memory. Once the image (for the current depth) is entirely computed by adding the contribution of every frequency (i.e., every processor has updated the image slice stored in the SSD under the previously described guarded zone), the image slice is written back to disk. This is done outside the critical region, while another frequency spool is prepared for calculations, allowing the use of a slower I/O support.

Saving the extrapolated field back to disk (see Figure 3) is not only mandatory due to our out-of-core scheme (when looping over depth, you need the previously extrapolated field to migrate the image to the next depth step), it will be mandatory for a checkpoint-restart implementation as well. This could be an interesting feature for very large datasets.

Although optimization can be criticized as highly machine dependent, it is necessary to achieve very high performances and low turnaround times for this kind of algorithm. Though fully vectorized, the Remez-Soubaras migration is memory bound. The most important memory reference bottleneck occurs when the processors simultaneously try to access migration operators (stored as a long vector) by indirect addressing. This is due to data mapping on memory banks. The way processors access the operators is data dependent (indirect addressing is done using the frequency/velocity ratio, and velocities are different in each point of the grid). One way to decrease the memory wait is to duplicate these operators. This is efficient for two reasons: performance increases are significant (megaflops rates are multiplied by a factor of two on the original vector code), and the memory tradeoff is reasonable.

Another important point is data management. Most of the I/O is performed in parallel regions and must be GEOVECTEUR-compatible and fast enough not to slow down computations. This problem is solved by using direct access I/O (library locked) with well-formed requests and by using more than half of the SSD as ldcache (user and system caches are bypassed). Fortunately, large datasets allow us to deal with huge record lengths, thus minimizing disk drive seek time.

Figure 5 shows the speedups obtained. These are global speedups, including the pre- and postprocessing of vector sequential code. The speedup for the migration itself (the largest parallel region) is much higher: 14.7 out of 16 in the standalone version. This means that the granularity of computational kernels is well suited to the CRAY C90 system, our parallel I/O strategy is quite effective, and the migration code is not slowing down the Autotasking process.

Performance on a real case

The dataset processed was an offshore North Sea seismic survey, acquired by a consortium of British Petroleum, Ranger, AGIP, Fina, Enterprise, Santos, and British Gas and carried out to investigate a salt dome. The survey's size was 1470 x 1060 bins, and processing length was seven seconds with a sampling interval of 4 ms. Extrapolations were performed every 25 m, with an interpolation step of 5 m. The frequency range was 2 - 70 Hz, leading to about 540 frequency layers. The operators were designed so that migration would be accurate for dips up to 70°.

This dataset represents around 50 Gbytes of disk space (data is stored in 64 bits to allow faster computations by avoiding format translation).

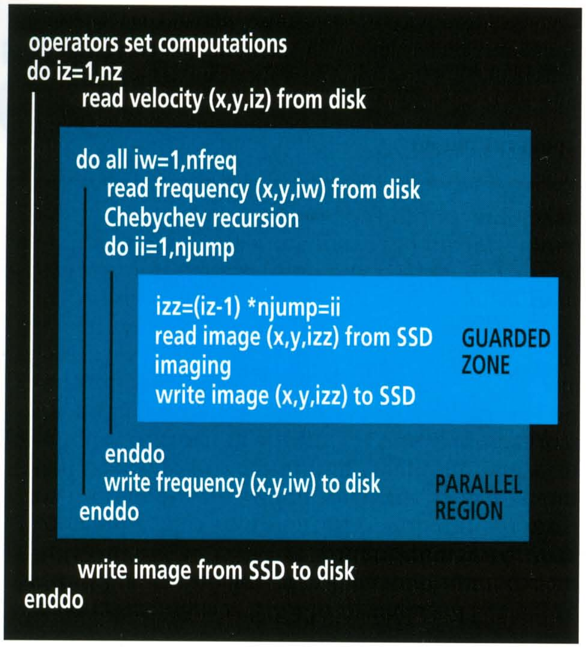


Figure 4 (left). Remez-Soubaras migration's parallel strategy on the CRAY C90 system.

To support our I/O strategy, we used a file system with 21 Cray Research DD-60 disk drives to store input data temporarily and output results. In addition, we used 12 Cray Research DD-60 striped disk drives to store the migration's critical files on a separate file system.

Besides the disks, this run used 230 Mwords of real memory and nearly one Gword of SSD (700 Mwords of which were used for caching the most critical files and 100 Mwords of which were used for secondary data segment (SDS)).

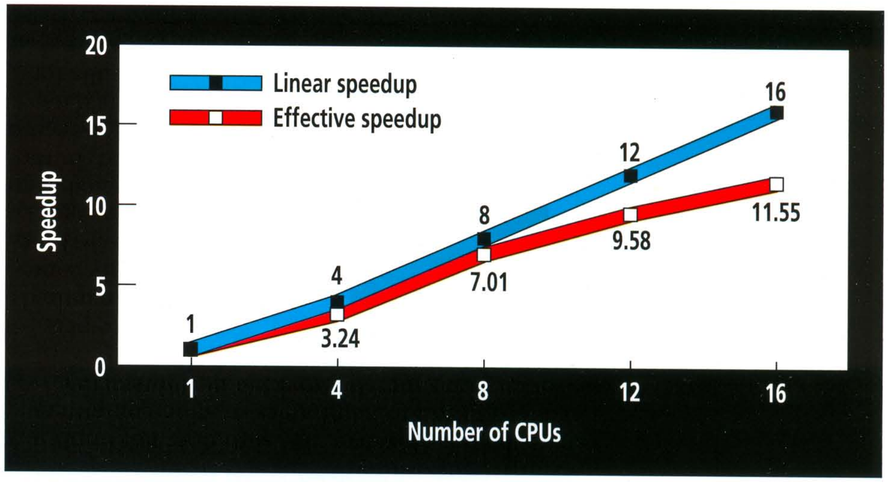
Including input and output, the total run lasted 26.4 hours, while the parallel execution took slightly less than 22 hours. The critical parallel region represented 1,080,000 CPU seconds, which means that the ratio of parallelism was about 13.7.

The parallel region was executed at a sustained rate of about 6 GFLOPS. This ratio of parallelism is remarkable if we take into account the huge amount of I/O requested by this job. More than 6 Tbytes were exchanged between the disks and the memory, leading to an average transfer rate of about 80 Mbytes/s, thanks to the efficiency of ldcache on the SSD. The SDS transfers were even larger, with more than 40 Tbytes exchanged with a sustained transfer rate in excess of 480 Mbytes/s.

Conclusion

The high performance obtained on the CRAY C90 supercomputer allows us to run new 3-D migration schemes on very large datasets in a reasonable timeframe. Using a high-end supercomputer with smart parallelization transforms an I/O-bound problem into a memory-bound one. This kind of algorithm is thus a promising candidate for implementation on Cray Research's newest supercomputer, the CRAY T3D massively parallel supercomputer.

Figure 5 (below). Global speedup for the Remez-Soubaras migration.



Acknowledgments

The authors thank British Petroleum, Ranger, AGIP, Fina, Enterprise, Santos, and British Gas for providing a publishable dataset. Special thanks to Ian F. Jones and Robert Soubaras of CGG for their help in algorithm design and dataset preprocessing, and everybody at Cray Research, Eagan, especially Don Lee, Paul Helvig, Lori Gilbertson, and Brian Kuznia, who helped make this big run a success.

About the authors

Irène Huard is a research geophysicist at CGG, designing migration codes for high-end computers. She received a Ph.D. degree in geophysics from the Ecole des Mines de Nancy.

Jean-Yves Blanc is a research computer scientist at CGG, in charge of parallel processing. He received a Ph.D. degree in applied mathematics from the Institut National Polytechnique de Grenoble in 1991. At CGG, he specializes in designing and optimizing 3-D seismic codes for parallel machines.

Since joining Cray Research France in 1992, André Gomez has been involved in software research, development, and optimization of aerodynamic and structural analysis codes. He specializes in optimizing codes for seismic customers, in particular CGG. He received an M.S. degree in theoretical physics from the Paris VI University in 1982 and advanced studies diplomas in astrophysics and geophysics in 1983 and 1984, respectively.

Roland Piquepaille manages petroleum industry accounts at Cray Research France. Originally a civil engineer, he began working in the computer industry 20 years ago. He joined Cray Research in 1983, and he helped CGG design the I/O components of GEOVECTEUR.

References

1. Hale, D., "3-D Depth Migration via McClellan Transformations," *Geophysics*, Volume 56, Number 11, pp. 1778-1785, November 1991.
2. Jones, I. F., "Choice of Algorithm for Poststack One Pass 3-D Depth Migration," Extended Abstracts of Papers of the 55th EAEG Meeting, Stavanger, Norway, 1993.
3. Soubaras, R., "Explicit 3-D Migration Using Equiripple Polynomial Expansion and Laplacian Synthesis," 62nd Annual International SEG Meeting, New Orleans, 1992.
4. Li, Z., "Compensating Finite Difference Errors in 3-D Migration and Modeling," *Geophysics*, Volume 56, pp. 1650-1670.
5. Holberg, O., "Toward Optimum One-Way Wave Propagation," *Geophysical Prospecting*, Volume 37, pp. 925-958, 1988.

Supercomputers as superintegrators

a new perspective for E&P

Richard D. Chimblo, Saudi Arabian Oil Company, Saudi Arabia

At the core of every exploration and production (E&P) organization lies a common goal—to sell more oil. The principles of supply and demand, of course, have a tremendous influence over this simple goal. So, too, does the technology that supports exploration and field development programs. As oil companies try to minimize costs and maximize production, they must reexamine technology while strengthening their relationships with geophysical contractors and university consortia. Supercomputers must become superintegrators of the vast amounts of information used by decision makers. By viewing supercomputing technology in this new light, oil companies can meet the world's energy needs cost-effectively in the twenty-first century.

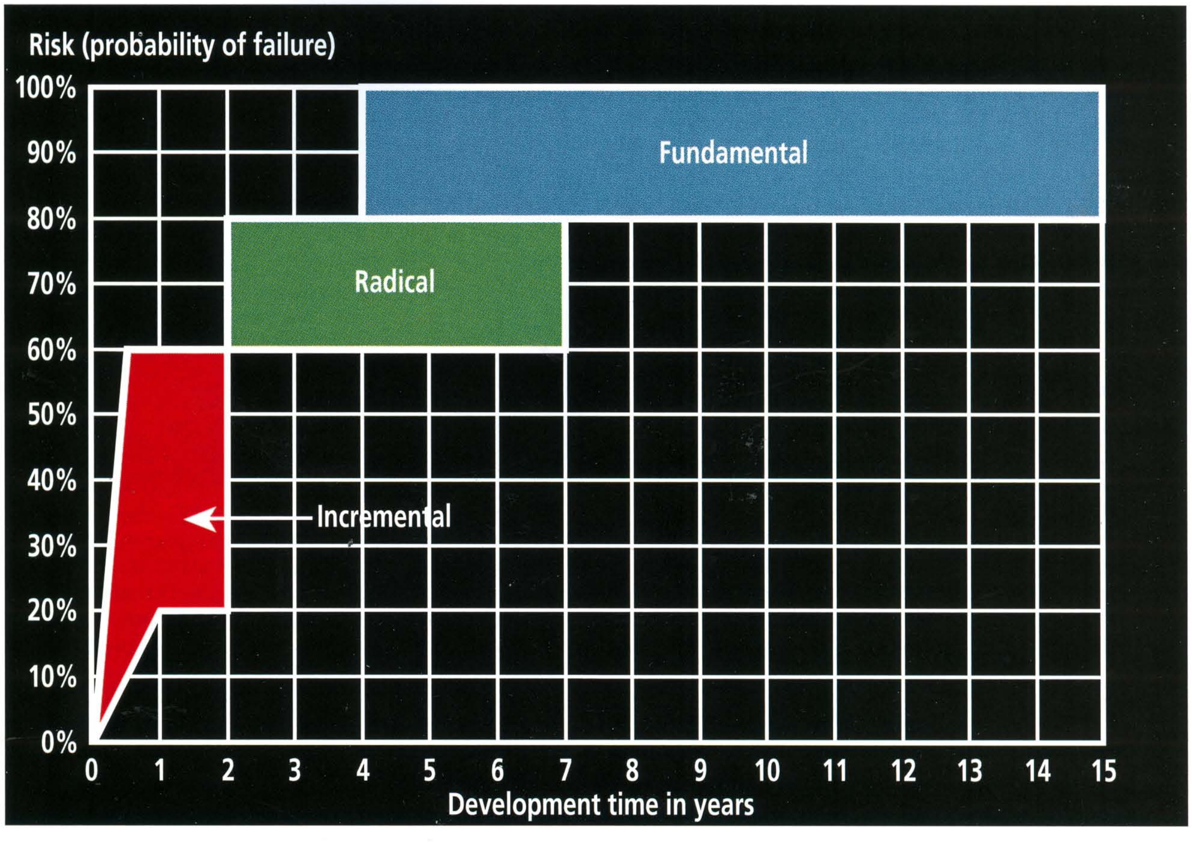


Figure 1. Types of R&D.

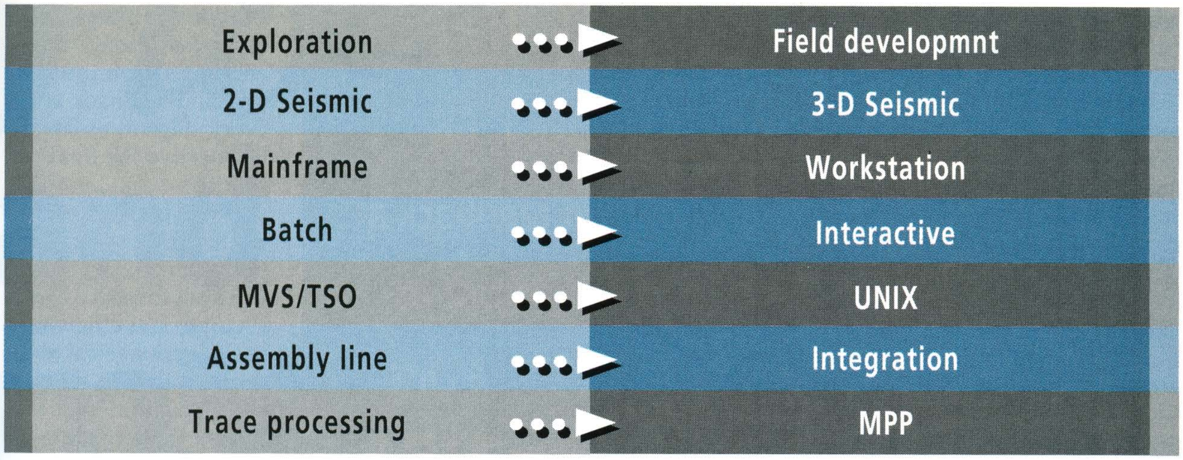


Figure 2. Current directions in exploration and production computing technology.

Geobusiness

The concept of geobusiness was introduced at the Cray Research conference, "Economic and Technical Impact of Supercomputing in the Petroleum Industry," held in Venice, Italy, in March 1993. Geobusiness is the economic perspective of the geophysical and geological technology employed in today's search for and recovery of hydrocarbons. Without this perspective and its cost-effective application, the goal of maximizing profits while minimizing costs is lost.

Geobusiness includes the effective utilization of people, resources, and computer technology. To meet E&P objectives, the Saudi Arabian Oil Company and others are adopting a new technology strategy:

- ☐ Optimize current computer resources
- ☐ Evaluate cost-effectiveness of emerging technologies
- ☐ Actively participate in research consortia

With this strategy, research and development have changed from budget-item liabilities to activities and operations tied closely to the corporate business strategy of finding and producing hydrocarbons. The new role of E&P technology is to help companies make more money. To do this, technology must take on a new perspective as the practical application of science.

A practical framework for R&D

Figure 1 shows the three types of R&D in terms of risk and development time.

- ☐ *Incremental R&D* is the clever exploitation of existing technology. It has a high probability of success and immediate application. At this level of research, oil companies can operate and still focus on current E&P problems using their own resources.
- ☐ *Radical R&D* is the commercialization of emerging technology for a specific business objective. With a moderate probability of success and application, radical R&D is best suited for industry contractors, vendors, and service companies that must maintain their competitive edge.
- ☐ *Fundamental R&D* is the introduction of new concepts or technologies that may have broader application. Because it is not well understood, this technology has a low probability of success and uncertain application within the industry for which it was originally planned. To maintain access to this long-term R&D, oil companies must take an active role in industry-sponsored university and research consortia.

Focusing resources

The Saudi Arabian Oil Company continues to focus on the business of finding and producing hydrocarbons. By engaging exclusively in incremental R&D and directing company resources into areas of low risk and high reward, the company can focus more attention and resources on the information integration and decision making required for profitable exploration and field development. Having adopted a systematic approach to the identification, evaluation, and implementation of new technology, we can achieve the optimum blend of geophysical contractor services and consortia participation.

This focusing of resources is only the process. Let us consider what technology will be needed to improve oil company operations during the next decade. Figure 2 offers some obvious shifts experienced in the area of E&P activities and computing technology. The shift from exploration to field development is a consequence of lower oil prices and higher finding costs. New discoveries generally are smaller, and the cost of locating these marginal reserves is higher. These facts have caused a reduction in the number of seismic crews, exploration drilling rigs, and other activities directed toward finding new commercial oil and gas fields. The geophysical technology that supports this shift from exploration to field development is seen in the recent advances and increased use of 3-D seismic surveys.

Maturing 3-D seismic technology

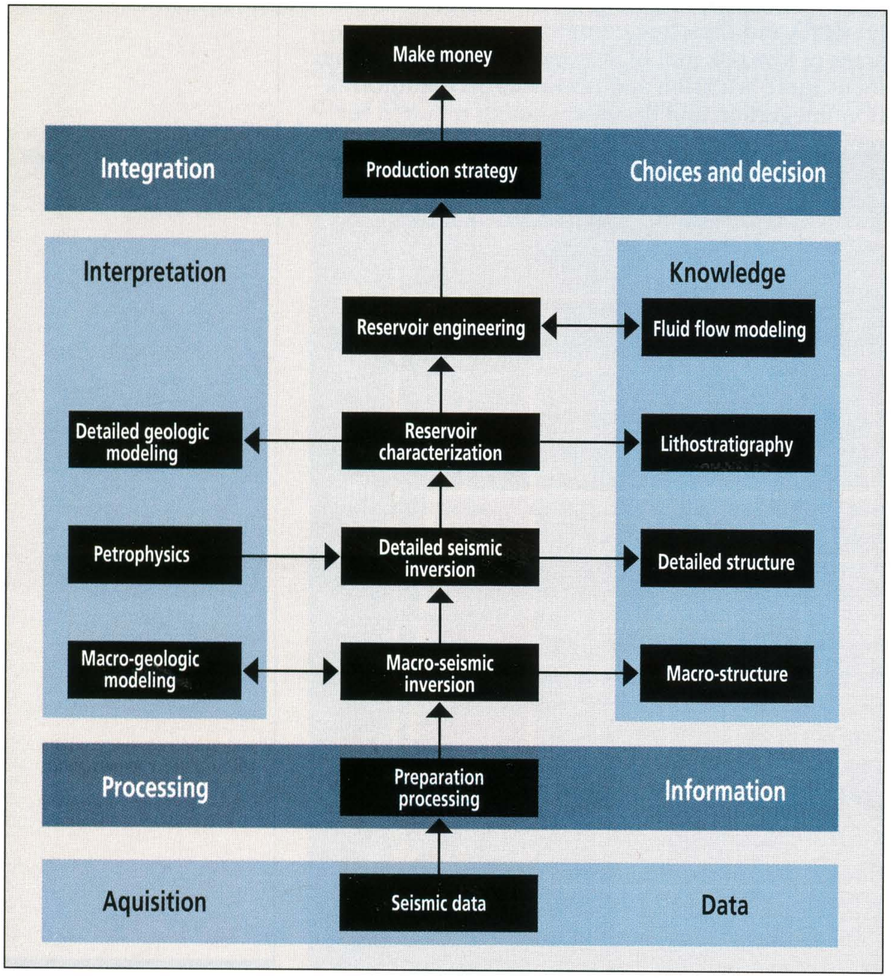
Initially targeted at improved reservoir delineation in field development activities, 3-D seismic technology now is being applied to the exploration of geologically complex structural and stratigraphic areas. Unlike 2-D seismic technology, which only provides vertical profiles or cross-sections of the earth's substrata, 3-D seismic technology offers detailed horizontal "map" views of structure, faults, and reservoir characteristics. Because of the greater need to interact with this 3-D data to form decisions regarding its processing, imaging, and interpretation, there has been a major shift from mainframe to workstation applications. However, this shift does not imply a reduced need for mainframes or supercomputers. Workstations merely provide a new interactive window in which to view our data.

Today, the emphasis is on integration and feedback at every stage in order to improve the final results.

Supercomputer power is needed to perform the once-undreamed-of calculations required to model large seismic volumes using highly computationally intensive applications such as 3-D prestack depth migration. But more accurate images, models, and maps of the reservoir constructed from seismic data are only part of what is needed. In the past, the assembly line approach was used to record seismic data in the field, transport it to the data center for processing, and interpret the results for selecting drilling locations. Today, the emphasis is on integration and feedback at every stage in order to improve the final results.

Integration starts with a system of people and technology working toward a more accurate and efficient use of geophysical and geological data. Figure 3 offers a model for this system that recognizes maximized profits through optimal reservoir management as the ultimate objective of any E&P organization. This system will have to put large volumes of information to work to allow us to understand the structural and stratigraphic architecture of large oil fields—their reservoir characteristics and their fluid flow models. This system requires the power of supercomputers—not just for migration and inversion of seismic data, but also for information integration, processing, and rapid retrieval from knowledge bases. With these capabilities, the system can perform "pattern" recognition, providing the complete picture of an oil field in structure, stratigraphy, lithology, production history, and expected performance.

Figure 3. Exploration and production seismic integration.



Information, processing, and superintegration

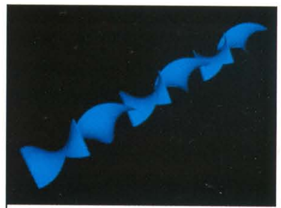
This concept of E&P seismic integration incorporates information management and synergism. More importantly, however, it is result-driven from the top down. In their new role, supercomputers will be superintegrators of vast amounts of information. Bigger and faster CPUs, interactive workstations, and the most sophisticated applications will not help us if we can't locate and integrate all available data. Leveraging supercomputers for integrated seismic processing is only one side of the coin. These systems also must be used to organize huge volumes of data so that the information can be fully perceived and analyzed to support the Saudi Arabain Oil Company's exploration and field development projects.

About the author

Richard D. Chimblo received an M.S. degree in geophysics from the University of Tulsa, Oklahoma, in 1975 and a Ph.D. degree in administration and management sciences from Columbia Pacific University, California, in 1991. From 1969 through 1980, he was an exploration systems supervisor with Amoco Production Company, engaged in defining and implementing seismic trace processing software, interactive interpretation systems, and training. In 1980, he joined Saudi Aramco Exploration as chief explorationist, directing seismic acquisition, processing, and interpretation activities for both prospect generation and regional studies. Since 1989, he has been chief geophysicist of field operations, processing, and R&D. He currently is directing the Saudi Arabian Oil Company's geophysical technology activities for exploration, field development, and information management.

Flow and mixing with Kenics static mixers

Designing optimal mixer configurations for the process industries, which rely to a great extent on common devices such as static mixers for pipeline mixing and dynamic mixers for agitated tanks, has been difficult because of limitations imposed by experimentation. Computer simulations on Cray Research systems, however, bypass some of these difficulties and give engineers new insight into mixer design.



Static mixers

The KM inline static mixer, manufactured by Chemineer Kenics, Inc., consists of a number of elements of alternating 180° helices, as shown in Figure 1. The elements are positioned such that the leading edge of one element is perpendicular to the trailing edge of the next element. The length of each element is one-and-one-half tube diameters. This type of static mixer is used under laminar flow conditions such as mixing polymers or food products such as peanut butter and chocolate.

The High Efficiency Vortab (HEV) static mixer, also manufactured by Chemineer Kenics, Inc., consists of an array of vortex-generating tabs mounted in a pipe, as shown in Figure 2. The HEV is used both for liquid-liquid and gas-gas mixing, as in the wastewater industries or in smokestacks.

Most experimental work on static mixers has concentrated on establishing design guidelines and pressure drop correlations, and the number of investigations of the flow and the mixing mechanisms has been limited. Recent advances in computational fluid mixing (CFM) have

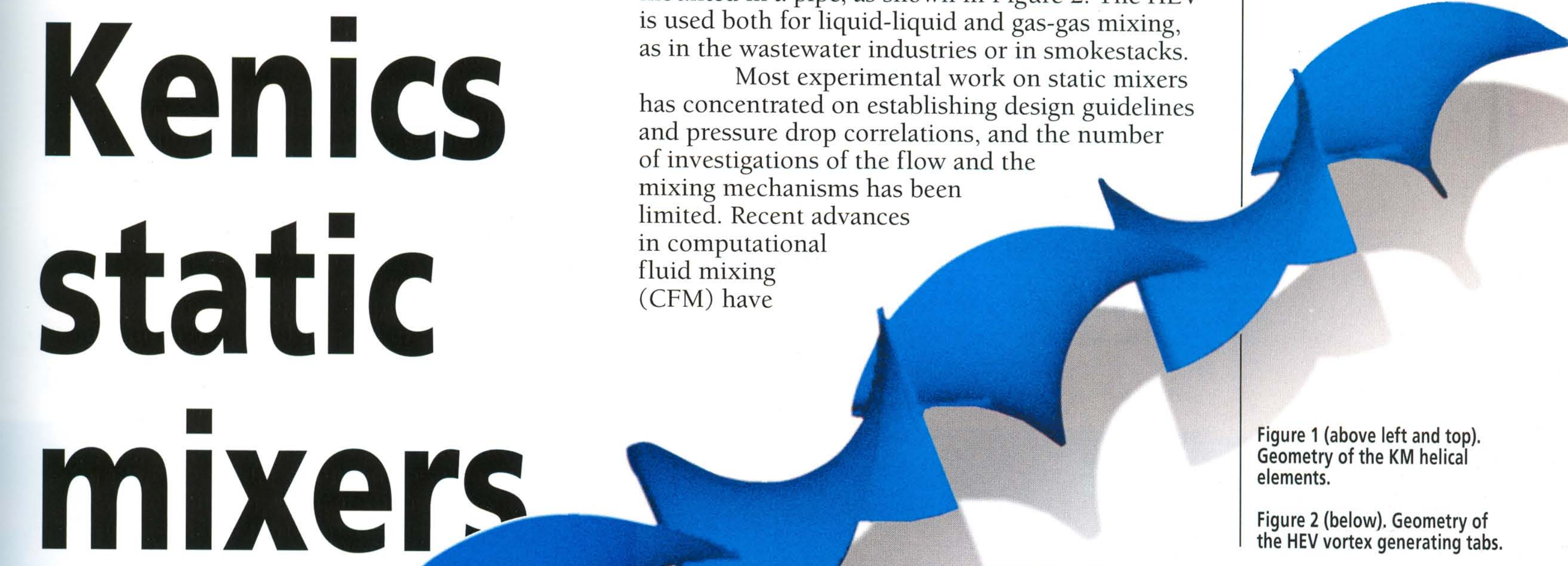
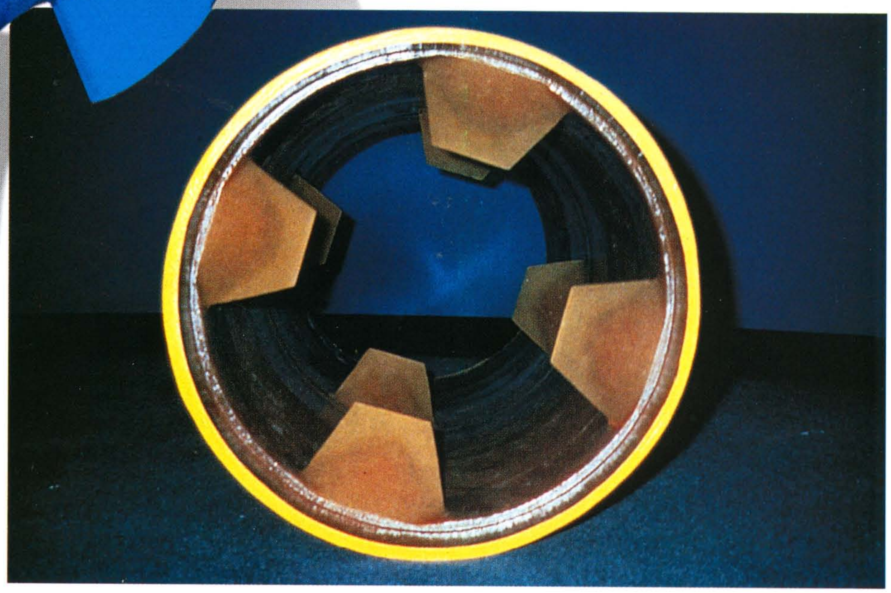


Figure 1 (above left and top). Geometry of the KM helical elements.

Figure 2 (below). Geometry of the HEV vortex generating tabs.



André Bakker, Chemineer, Inc., Dayton, Ohio
Richard LaRoche, Cray Research, Inc.

made computer simulations a useful tool in static mixer design and analysis. These simulations explore the possibilities CFM offers in the analysis of mixing and provide insight into the mixing mechanism. The FLUENT V4.21 software package from Fluent, Inc. helps analyze the static mixers' flow pattern, pressure drop, and mixing characteristics.

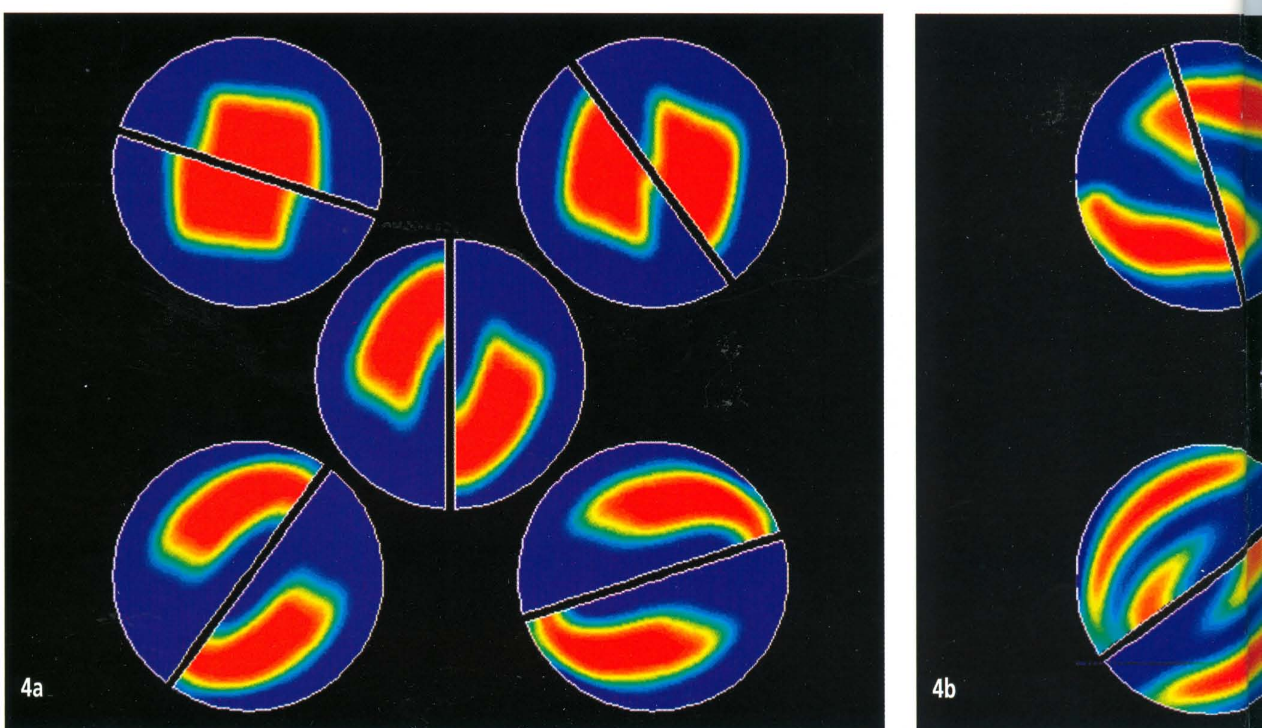
The numerical grids used to model the mixers were generated with FLUENT PreBFC V4.01 and exported to FLUENT V4.21 for the flow and mixing computations. For the laminar flow KM mixer, the model used the Reynolds number Re = 10; for the turbulent flow HEV mixer, the model used the Reynolds number Re = 100,000. For the turbulent flow conditions, the Reynolds stress model (RSM) also was used. The QUICK differencing scheme was used in all calculations.

Initial calculations were performed with 100,000 grid nodes on a Hewlett-Packard HP-750 workstation. The converged solutions were exported to the CRAY C90 system, where the grid was doubled in two directions, resulting in 400,000 nodes. Final calculations then were performed on a CRAY C90 system. These 400,000-node problems require 104 Mwords of core memory and are among the largest FLUENT problems ever run. Problems of this level of grid density are not possible on today's workstations. Typical calculation times are five and nine CPU hours on the CRAY C90 system for the KM and HEV mixers, respectively. These calculations required approximately 400 iterations with the algebraic multigrid solver option to reduce residuals to 0.001 for convergence.

There was very little difference between the flow field results at 100,000 nodes and at 400,000 nodes. Due to numerical diffusion, the species mixed too fast with the 100,000-node grid.

Figure 3 (above). Inlet concentration of the chemical species for the KM mixer.

Figure 4a, 4b, 4c, (right). Concentration profiles at various intersections of the first, second, and sixth helical element (18°, 55°, 90°, 126°, 162°).



The 400,000-node grid gave a much more accurate prediction of the mixing rate.

To evaluate the KM mixer's mixing mechanism, which consists of a series of helical mixing elements, the transport of two chemical species was calculated. Figure 3 shows the center of the inlet as 100 percent red and the outside of the inlet as 100 percent blue. The results are shown in Figure 4 as a series of raster plots. The plots show the concentration fields of the chemical species after the mixing mechanism has passed through 18°, 54°, 90°, 126°, and 162° rotation in the first, second, and sixth mixing element, respectively.

Figure 4 also shows how the red core coming from the inlet is split into two red islands, which are stretched and move outward. The blue, which was on the outside in the inlet of the element, is split in two semi-circular filaments, which are moved toward the inside of the element. Similar stretching and folding processes occur in the next elements. In the inlet of the third element, the blue species is on the inside, meaning that the concentration field has flipped inside out. This process of splitting, stretching, folding, and flipping inside out repeats itself every two elements, until the fluids are mixed. By the time the end of the sixth element is reached (Figure 4), the species concentrations are nearly uniform.

The pressure drop across the elements was calculated with the correlation proposed in the Kenics design guides. The predicted pressure drop was within 10 percent of the experimentally found pressure drop.

Figure 5 shows particle streaklines behind one of the HEV turbulent mixer's tabs. The streaklines show a strong circulation flow in the wake of the tab. The vortex, attached to the wall of the tube and not to the tabs, lies parallel with the tab and then bends to a longitudinal vortex with a center close to the tip of the tabs.

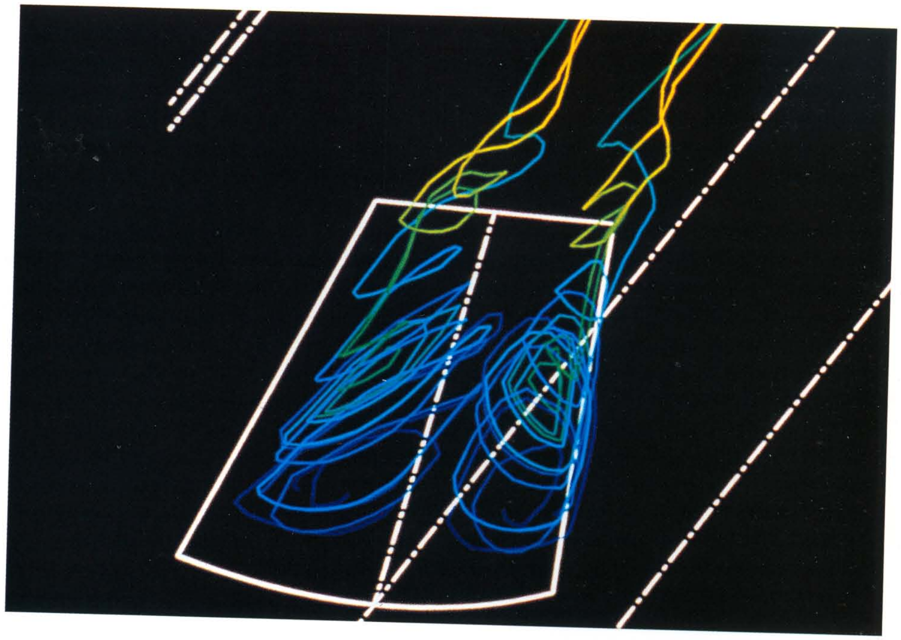
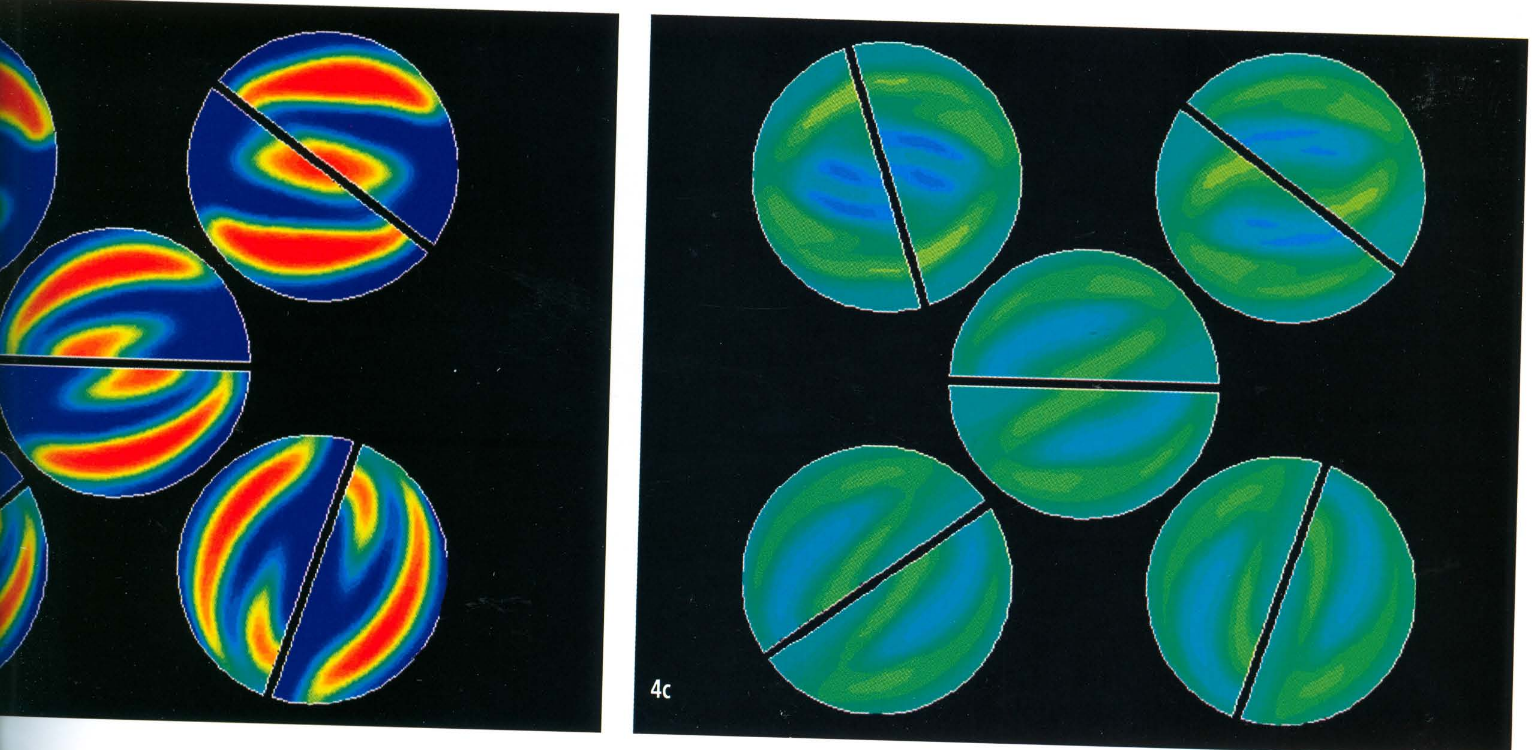


Figure 5 (above). Particle streaklines in the wake of the tab.

W. J. Gretta, in his master's thesis, "An Experimental Study of the Fluid Mixing Effects and Flow Structure due to a Surface Mounted Passive Vortex Generating Device," investigated the flow pattern as generated by the tabs using a combination of hot wire anemometry, hydrogen bubble visualization, and dye visualization.[1] Figure 6 shows the flow pattern according to Gretta. He discovered that the tabs not only generate a pair of counterrotating, longitudinal vortices but shed hairpin vortices as well. The smaller hairpin vortices move downstream with the larger longitudinal vortices.

The generation of these hairpin vortices is a transient process. Since the current model does not take time-dependent effects into account, these vortices could not be modeled explicitly. However,



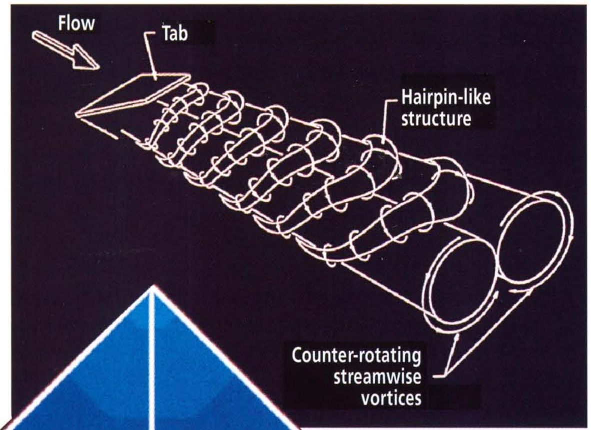


Figure 6 (right). The flow field behind the HEV tabs according to Gretta.

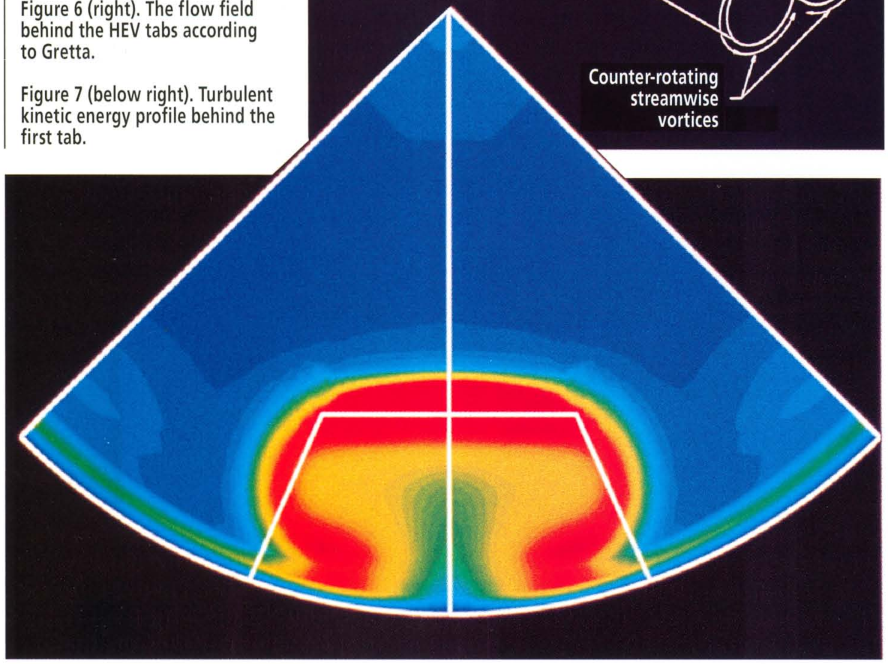


Figure 7 (below right). Turbulent kinetic energy profile behind the first tab.

the hairpin vortices do show up in the CFM results as regions with a large turbulence intensity. Figure 7 shows the turbulent kinetic energy in a plane directly behind the first tab. Red denotes regions with a large turbulence intensity. This plot shows that there is a region with a large turbulence intensity surrounding the vortex, where the hairpin vortex would otherwise be found. The hairpin vortex is generated in the high shear region at the edge of the tab. In the steady-state model used here, high shear increases the production of turbulent kinetic energy.

The mixing of a tracer fluid was studied to determine the HEV mixer's efficiency. The tracer fluid was injected at two positions—the center of the tube and a point in front of a tab. The total concentration of tracer fluid in the tube was 1.25 percent of the total fluid volume. Figure 8 shows the concentration field in a plane through the center of the tabs. Red denotes regions with large concentrations of tracer fluid, and blue denotes low concentration regions. The injection in front of the tab bends off when it hits the tab and is blended almost immediately in the turbulent wake of the tab. The injection in the center persists almost undisturbed until halfway between the two tabs. There, the turbulence intensity generated by the vortex is large enough to blend the material in the center. These results indicate that it is not just the longitudinal vortex which controls the blending; the hairpin vortex and random turbulence contribute significantly to the mixing.

The state of the art in CFM allows for the modeling of flows and mixing of chemical species in complex geometries such as those of static mixers. Computer simulations, though time-intensive, are much faster than extensive experimentation when engineers are trying to optimize the geometry of the mixing elements for a variety of operating conditions, fluid viscosities, and equipment size.

Grid independence for the flow pattern is achieved with fewer grid nodes than for the results of the species mixing. The large-memory capability and fast processing speed of the CRAY C90 system made the 400,000-node problems possible.

With the KM helical elements, mixing occurs through a combination of flow splitting and shearing at the junctions of successive elements and a stretching and folding mechanism within the elements. The concentration field looks like it is flipped inside out after two elements: material originally at the wall is in the core, and vice versa. This makes the KM element an excellent radial mixing device, applicable in a variety of laminar mixing applications.

The HEV mixers generate a complicated vortex system, consisting of transient hairpin vortices and steady longitudinal vortices. The model correctly predicted the longitudinal vortices. The transient hairpin vortices showed up as regions of high turbulent kinetic energy. Future work will concentrate on more grid dependency studies and on comparing various alternative geometries.

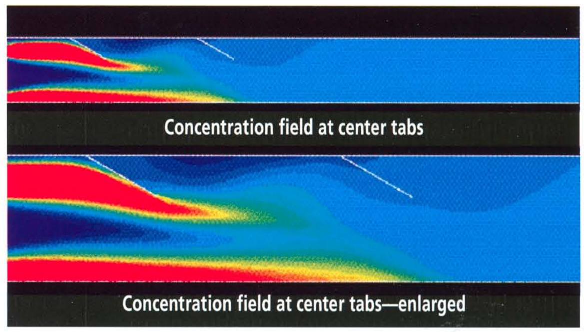


Figure 8 (far right). Side view of the concentration field in the HEV mixer.

Reference

1. Gretta, W. J., *"An Experimental Study of the Fluid Mixing Effects and Flow Structure due to a Surface Mounted Passive Vortex Generating Device,"* master's thesis, Lehigh University, Bethlehem, Pennsylvania, 1990.

About the authors

André Bakker is principal research engineer at Chemineer, Inc. in Dayton, Ohio. He is specialized in modeling mixing problems using CFD. He holds both an engineering degree and a Ph.D. degree in applied physics from the Delft University of Technology in the Netherlands.

Richard LaRoche is a senior chemical engineer in the Engineering Applications Group at Cray Research, Inc., where he focuses on the application of CFD to process industry problems. He received his B.S. degree in chemical engineering from Montana State University and his M.S. and Ph.D. degrees in chemical engineering from the University of Illinois at Urbana-Champaign. He has previously held positions at Dow Chemical and Penn State University.

CORPORATE REGISTER

CRAY T3D system debuts with strong initial orders

Cray Research's first massively parallel processing (MPP) system, the CRAY T3D supercomputer (see article, p. 2), attracted nine initial orders from U.S. and international customers in the government, university, and commercial sectors. These first customers include a major petroleum company, a U.S. university, and a large, well-known Japanese industrial firm.

The Pittsburgh Supercomputing Center installed and accepted a 32-processor CRAY T3D prototype, scheduled to become a 512-processor production system in early 1994. "There is a continued need for increased computational power to attack scientific problems of great social importance," said Ralph Roskies, co-scientific director of the Pittsburgh center. To achieve that goal, the CRAY T3D system, already operational, will work in tandem with a CRAY C90 system, a parallel-vector supercomputer with 16 processors and a peak theoretical speed of 16 billion calculations per second.

Cray Research technicians began installing the CRAY T3D system August 23 at the Westinghouse Energy Center (WEC) in Monroeville, where the center's hardware is maintained by WEC staff. "The CRAY T3D system is the 11th supercomputer we've installed at the site over the years, including two other MPPs," says James Kasdorf, director of supercomputing for Westinghouse Electric Corporation. "This one has come up more quickly and smoothly than any before, which is really remarkable for the first one in the field."

NASA's Jet Propulsion Laboratory/Caltech (JPL/Caltech), Pasadena, California, is scheduled to receive a 256-processor system in fourth quarter 1993. And **Ecole Polytechnique Fédérale de Lausanne (EPFL),** the Swiss Federal Institute of Technology, signed a preliminary agreement to acquire a 256-processor system in early 1994. As part of Cray Research's Parallel Applications Technology Program (PATP), agreements with PSC, JPL/Caltech, and EPFL call for the systems to be available not only for the organizations' users, but for collaborations with Cray Research to develop targeted software applications for the CRAY T3D system.

The Arctic Region Supercomputing Center, a national facility located at the University of Alaska Fairbanks, ordered a 128-processor CRAY T3D system for installation in first quarter 1994.

The Commissariat à l'Energie Atomique (CEA), Division des Applications Militaires—the French Atomic Energy Commission—ordered a 128-processor CRAY T3D system that will be closely coupled with a CRAY M92 system already onsite at CEA/Limeil. The CRAY T3D system will be used for research in fundamental and laser physics and for electrodynamic and aerodynamic studies.

The European Centre for Medium-Range Weather Forecasts (ECMWF) ordered a CRAY T3D system that will be closely coupled with a CRAY Y-MP2E parallel-vector supercomputer system also ordered by ECMWF. Both systems are scheduled to be installed at the Centre's Reading, England, operations in mid-1994. ECMWF will use the CRAY T3D system to develop new forecasting models that take advantage of MPP technology and to run the Centre's ensemble prediction system. This system is used to generate forecasts provided to 18 national weather services throughout Europe.

The research division of **Electricité de France (EDF),** the world's largest electrical utility, based in Clamart, France, will be an early customer for Cray Research Superservers' new CRAY SUPERSERVER 6400 (CS6400) system, introduced October 25 (see article, p. 4). A 16-processor system with two Gbytes of central memory was shipped in November, and the system will be upgraded to 32 processors and four Gbytes of memory in 1994. A long-time Cray Research customer, EDF plans to use the CS6400 for scientific, engineering, and data management applications serving users over EDF's extensive computer network. Cray Research Superservers is a subsidiary of Cray Research, Inc. EDF also ordered a CRAY C98 supercomputer, its eighth Cray Research system since 1981. EDF will use its new eight-processor system from Cray Research's expanded CRAY C90 series for structural analysis, electromagnetics, climatology, nuclear physics, and fluid dynamics simulations related to power transmission, new nuclear power plant design, and power system analysis.

SICAN, a German electronics firm based in Hannover, ordered a CRAY SUPERSERVER 6400 (CS6400) system and is scheduled to receive it in the first quarter of 1994. SICAN initially will receive a 20-processor system that will be upgraded later in 1994 to a 48-processor system. SICAN provides design services for German and U.S.-based industries that use applications-specific integrated circuits and full custom designs. The CS6400 system will be used for highly scalar electronic computer-aided design applications, as well as for data management. The CS6400 runs the Solaris operating environment and is binary compatible with the full line of workstations and servers from Sun Microsystems, Inc.

The Faculty of Computer Science and Information Systems at the **Universiti Teknologi Malaysia (UTM)** ordered a CRAY Y-MP EL supercomputer, to be installed in the newly established Advanced Computing Laboratory at its Jalan Semarak campus in Kuala Lumpur. This supercomputer, the first Cray Research system to be installed in Southeast Asia, will be used for large-scale simulation in a wide variety of scientific disciplines. It also forms the delivery platform for a series of training courses that will be offered through the Cray Training Centre being set up jointly by UTM and Cray Research within the Advanced Computing Laboratory.

The U.S. Commerce Department's **National Oceanic and Atmospheric Administration** awarded a five-year, $46 million contract to Cray Research to lease a CRAY C916 supercomputer. The National Weather Service's National

Meteorological Center (NMC) will use the supercomputer to process data from around the globe into graphical and alphanumerical outputs from atmospheric models. Operating in excess of 15 GFLOPS, the system will take four hours to run an extremely complex model the NMC uses to assemble its daily mid-range forecasts. Domestic and international aviation information, precipitation forecasts for agriculture, and maritime forecasts are some of the needs the CRAY C916 system will address.

Freddie Mac installed its second CRAY Y-MP EL supercomputer system at its Reston, Virginia, facilities. The company's financial analysts will use this system to develop and test new financial models and algorithms and for disaster recovery backup for the company's first Cray Research system. The original system was installed in 1992 at the McLean facilities and is used by Freddie Mac's Dealer Services Division for day-to-day financial modeling and analysis related to its $465 billion total servicing portfolio of mortgage-backed securities.

The **Hyundai Motor Company** installed a CRAY Y-MP 4E supercomputer at its Passenger Car Engineering Center in Ulsan, Republic of Korea. This is the first supercomputer and the first Cray Research system purchased by Hyundai. The automaker will use the supercomputer for structural and mechanical engineering applications related to new car design, crashworthiness testing, external airflow simulation, and engine combustion. The Hyundai order also includes a license for CRI/*Turbo*Kiva software, a Cray Research-designed engine combustion simulation software package that models fluid flows—air intake, combustion, and exhaust characteristics—and provides three-dimensional graphic simulations of internal combustion engines used in passenger cars, trucks, and off-road vehicles.

Ford Motor Company ordered a second CRAY C916 system, to be installed at its Engineering Computer Center in Dearborn, Michigan. The system will be used to accelerate the design analysis process using three primary applications: computational fluid dynamics for improved aerodynamic design; structural analysis for reducing noise, vibration, and harshness; and crash analysis for enhanced vehicle safety. The system also is slated for manufacturing applications such as the simulation of sheet metal forming of car components. Ford was the first commercial purchaser of the CRAY C916 system.

Cray Research received its first Asia Pacific order for the company's top-of-the-line supercomputer when the **Korean Institute of Science and Technology (KIST),** in Taejon, Korea, ordered a CRAY C916 system. Scheduled for installation at KIST's Systems Engineering Research Institute, the new supercomputer will be used primarily by Korean universities and industrial research organizations for structural and mechanical engineering, weather forecasting and environmental research, chemistry and biological applications, computational fluid dynamics, graphics and image processing research, and petroleum exploration.

The **CNRS (National Center for Scientific Research),** Paris, France, ordered a CRAY C98 supercomputer, the eight-processor model in the expanded CRAY C90 series. The French equivalent of the U.S. National Science Foundation, CNRS oversees all academic research in the country. CNRS will use the CRAY C98 system for a variety of scientific projects including climate modeling, physics, materials science, and chemistry research.

From PCs to supercomputers, agreement makes programming easy, compatible

Cray Research, Inc. and Absoft Corporation announced an agreement to develop software tools that would allow scientists and engineers to write programs on personal computers and then run the programs on more powerful systems, including Cray Research supercomputers. Under the joint agreement, Absoft will create, market, and distribute versions of Cray Research's CF90 compiler for high-performance desktop computers. Initial targets for the Absoft versions of the CF90 compiler include: PCs based on Intel486 and Pentium processors, and the forthcoming series of PowerPC-based Macintoshes. The initial versions of these products are scheduled for release in mid-1994 and will be marketed by both Cray Research and Absoft, as well as Absoft's worldwide network of more than 500 commercial software resellers.

CF90 programming environment 1.0 sets new standard for scientific programming

Fortran 90 is the latest standard version of the Fortran programming language, the most widely used language in scientific computing. The new standard includes enhanced performance and ease-of-use features while retaining full compatibility with Fortran 77 to ensure portability of Fortran 77 codes and preserve existing application investments. Cray Research's CF90 programming environment includes a robust, fully compliant Fortran 90 compiler, the CrayTools programming toolkit, and the CrayLibs suite of optimized scientific libraries.

The CF90 compiler is the first full native Fortran 90 compiler for supercomputers. This new compiler is based on proven code generation and instruction scheduling technology to enhance reliability.

Because Fortran 90 is a superset of Fortran 77, the CF90 compiler also supports the full Fortran 77 language, as well as extensions to Fortran 77 supported by release 6.0 of Cray Research's CF77 compiler. Programmers can incorporate into their Fortran 77 codes as many or as few of the Fortran 90 features as they choose. They need not learn the new standard in its entirety, but can concentrate on those features most appropriate for their applications.

The new standard's portability creates a natural bridge between heterogeneous architectures, such as those of Cray Research's parallel vector and massively parallel systems. Release 1.0 of the CF90 programming environment supports the CRAY C90, CRAY Y-MP, and CRAY EL parallel vector computer systems (running release 7.0, 7.C.3 or later of Cray Research's UNICOS operating system). Subsequent releases will support Cray Research Superservers SPARC-based systems and CRAY T3D systems.

On multiprocessor Cray Research systems, the CF90 compiling system automatically partitions programs for execution across multiple processors, requiring no intervention by the programmer. However, programmers can optionally insert directives to control partitioning explicitly. Unlike previous Cray Research Fortran compiling systems, in which Autotasking was incorporated into the preprocessor, the CF90 compiler itself includes the Autotasking capability. This arrangement enhances code reliability and eliminates the need for source code translation, making it easier for programmers to relate optimized code back to their original source code. In addition to Autotasking, the CF90 compiling system supports implicit data parallel expressions, including array syntax and array intrinsics.

The CF90 compiling system uses state-of-the-art PDGCS optimization technology for aggressive code restructuring, which maximizes vectorization and parallelization of user programs.

Code performance is further enhanced by CrayLibs, a collection of high-performance scientific libraries, such as BLAS levels 1, 2, and 3, LINPACK/EISPACK, LAPACK, and new fast Fourier transforms (FFTs). These libraries serve as building blocks for developing high-performance code, delivering to users near peak speed from their Cray Research systems. The CF90 compiling system automatically detects and replaces user code with calls to these optimized library routines.

Other libraries include high-performance Fortran I/O routines that take full advantage of the fast I/O of Cray Research computer systems. The CF90 environment provides transparent access to these I/O libraries through standard I/O interfaces, keeping source code standard and portable across platforms. The variety of routines gives users the flexibility to customize I/O according to the needs of their applications:

- ☐ Raw I/O provides large, fast data transfers to disk storage, bypassing UNICOS operating system buffering.
- ☐ Asynchronous I/O lets users overlap I/O and computations.
- ☐ Flexible file I/O (FFIO) provides data conversion from the proprietary formats of other vendors, eliminating the need for inefficient ASCII files for data transport. FFIO routines also allow programs to use real shared memory, eliminating I/O transfers altogether. Alternatively, users can use Cray Research's SSD solid-state storage device as an extended main memory, which also improves program performance by eliminating the need for I/O to peripheral devices.

The CF90 programming environment contains all of the Fortran 90 features that enhance user productivity by making code easier to write and maintain.

The CF90 environment includes the CrayTools programming toolset, a full set of advanced X Window System-based visual programming tools for debugging, source-code analysis, and performance analysis.

CrayTools incorporates expert systems that interpret performance data and offer suggestions to improve program performance. For instance, when programmers want to increase parallelism beyond the level provided automatically, they can use ATExpert, an expert system for parallel processing performance analysis. ATExpert predicts parallel performance without access to a dedicated system, identifies performance bottlenecks, and suggests actions to improve performance. Other analysis tools provide detailed information about I/O, memory use, vectorization, and algorithm performance.

Release 1.0 of the CF90 programming environment also includes xbrowse, a language-sensitive source code browser, and a pre-release version of the Cray TotalView debugger. This tool is a symbolic, source-level debugger with a window-oriented interface that allows users to control and display information about individual processes. It provides a high-level view to users to simplify parallel debugging at the application level. Cray TotalView will be the common debugger for all Cray Research and Cray Research Superservers systems: parallel vector, massively parallel, and SPARC.

The full CF90 programming environment is a complete and easy-to-use programming environment for scientists, engineers, and commercial application developers. To find out more, please contact your local Cray Research representative.

Cray Standard C programming environment 4.0

To help C programmers derive maximum benefit from their Cray Research systems, Cray Research has integrated its high-performance standard C compiler, powerful debugging and analysis tools, and optimized libraries into a complete, cohesive programming environment. Based on established, emerging, and de facto industry standards, the Cray Standard C programming environment provides the means to create optimized, portable code.

In addition to being a productive, integrated environment, the Cray Standard C programming environment contains a rich set of features that achieves unmatched performance, provides accessible analysis, and gives users portability of code and data across platforms and into the future.

The Cray Standard C programming environment includes the Cray Standard C 4.0 compiler and the latest versions of CrayTools and CrayLibs. Enhancements to the Cray Standard C 4.0 compiler include the following:

- ☐ cdbx now provides basic C++ support including name-demangling capabilities
- ☐ xbrowse is now C language sensitive and provides additional features such as
  - -Graphical call trees
  - -Loop marking
  - -Include file cross reference
  - -Tracing C global variables
- ☐ Cross-file inlining
- ☐ Character loop vectorization
- ☐ Conversion of conditional operators to fast scalar merge functions
- ☐ Automatic parallel processing of reduction loops
- ☐ Improved alias analysis offers better performance for loops with pointers
- ☐ Advanced loop-restructuring techniques including
  - -Loop unrolling for increased optimization
  - -Loop inversion
  - -Loop fusion
  - -Loop collapse
  - -Loop unwinding and peeling
- ☐ Enhanced parallel/vector dependency analysis

CrayTools includes cdbx, a visual interactive debugging environment that allows programmers to debug optimized code at the familiar source level, and xbrowse, a source-sensitive analysis tool that provides cross references, common block, and variable tracing capabilities. CrayTools also includes flexible listing tools and Tooltalk, a distributed message passing service based on RPC and UNIX sockets enabling independent applications, executing on the same or different machines, to communicate without having direct knowledge of each other.

CrayLibs consists of the libsci and libm libraries. The libsci scientific library, which provides highly optimized kernel subroutines for the Cray Fortran programming environment, also is accessible from C language programs. This library contains a wide selection of mathematical software, such as:

- ☐ Level 1, 2, and 3 Basic Linear Algebra Subroutines (BLAS)
- ☐ LINPACK, EISPACK, and LAPACK libraries for linear algebra and eigenvalue analysis
- ☐ FFT and signal processing routines
- ☐ Sparse linear algebra computations

Version 7.0 or later of the UNICOS operating system is required for using the power of the Cray Standard C programming environment 4.0. For further information, please contact your Cray Research representative.

APPLICATIONS UPDATE

Cray Research introduces CraySoft products for RISC systems

CraySoft products from Cray Research provide a Cray Research-compatible programming and production environment to users of desktop and server systems. These products are shrink-wrapped, workstation versions of the same products made available on Cray Research supercomputers. The first CraySoft product, the CraySoft Network Queuing Environment (NQE), is now available to users of RISC-based systems.

The CraySoft Network Queuing Environment (NQE) is a suite of software packages that provides enterprise-wide batch service. NQE comprises two components: the NQE Server and NQE Client. The NQE Server consists of the Cray Research-enhanced Network Queuing System, Network Load Balancer, and File Transfer Agent (FTA) for UNIX workstations. NQE Server provides a stable network batch queuing environment compatible with public domain NQS and supports destination selection, load balancing, and status of tasks across a batch complex. The FTA provides guaranteed synchronous or asynchronous outbound and inbound file transfers over ftp. The Network Queuing Client (NQC) allows users to submit jobs to the NQE server from their workstation.

CraySoft currently offers NQE for SPARC systems running Solaris 2.2, including Sun workstations and the CRAY SUPERSERVER 6400 product line. Support for IBM RS6000 (AIX), SGI (IRIX), HP (HP-UX), and DEC Alpha (OSF/1) systems will be available in the first half of 1994. Current RQS customers will receive credit toward NQE for their RQS platforms. Pricing for Solaris 2.2 systems is under $3000 for a 10-user server license with unlimited clients.

To order the CraySoft Network Queuing Environment or receive more information, contact Roy Mulvaney, CraySoft sales; telephone toll-free in the United States: 1-800/BUY-CRAY, or 612/683-3030; fax: 612/683-5508; or email: craysoft@cray.com.

Cray Research and TeamQuest begin system analysis service

Cray Research field sales representatives worldwide now offer a System Performance Service program to help customers identify and eliminate bottlenecks while fine tuning their supercomputing system. This program is a joint effort between Cray Research and TeamQuest Corporation, developer of the client/server-based performance analysis and capacity planning package Capacity Management Facility Baseline. System analysis, performance management, capacity planning, and system tuning are included in the program. For more information, please contact your Cray Research representative or Phil Hernick, Cray Research, Inc., telephone: 612/683-5747; email: phil.hernick@cray.com.

DELSI seismic data processing software available on Cray Research systems

At the 1993 European Association of Exploration Geophysicists (EAEG) conference and exhibition, a new seismic data processing software package that could substantially boost the petroleum industry's success in locating new oil and gas reserves was demonstrated on a CRAY Y-MP EL system.

Called DELSI, an acronym for DELft Seismic Inversion, the new software was developed by the Dutch research organization TNO (de nederlandse organisatie voor toegepast natuurwetenschappelijk onderzoek), headquartered in Delft, Holland.

DELSI software initially will incorporate two important technologies: a wave equation-based multiple elimination method, to eliminate unwanted "noise" and show a clearer "picture" of the hydrocarbon reservoirs, and a controlled illumination method based on areal shot record technology, which allows geophysicists to efficiently focus on the target zone. "The DELSI software is expected to be considerably more accurate in the discovery of oil and gas than today's technologies and will help the petroleum industry more readily locate these natural resources," said Max Mulder, head geophysicist at TNO.

Cray Research is assisting in DELSI software development through a 1991 joint software development agreement with TNO. The company also has been a DELPHI consortium member for the past seven years. The DELSI software will be commercially available on the full range of Cray Research systems in mid-1994.

UniChem 2.0 supports more workstations, has improved DGauss performance

Cray Research has announced the release of UniChem 2.0, an updated version of its popular quantum chemistry modeling package. The product provides a single, graphical interface to a variety of powerful quantum mechanics programs such as DGauss, CADPAC, and MNDO93. New features in UniChem 2.0 include:

- Expanded User Preferences such as atom property editing, launch options, and print options
- An Attacher's Toolkit 2.0 which allows users to add their own code to the software system and access in-house codes through the graphical user interface (GUI)
- New set-up options for the DGauss, CADPAC, MNDO93, and Gaussian 92 software packages

- DGauss 2.0 also offers significant performance improvements to the DGauss server code. For example, DGauss 2.0 time-to-solution performance has been increased by as much as a factor of two; numerical precision was increased by a factor of 1000, and memory requirements have been cut by as much as 30 percent.
- UniChem 2.0 also is available with X Window System graphics, allowing the GUI to be displayed on any workstation or terminal client supporting the X Window System protocol. While SGI hardware still is required to drive graphics, a single SGI system can act as a server to multiple X Window System terminals or workstations.

In time for the new release, Cray Research and Molecular Simulations, Inc. (MSI) announced a joint agreement allowing MSI to distribute UniChem. Under the terms of the agreement, MSI will focus on the new workstation version of UniChem 2.0, which was released in March, while Cray Research will concentrate on the supercomputer market. In addition, each company will provide first-line support to its UniChem customers.

The UniChem 2.0 release is available for use on any Cray Research system running the UNICOS operating system, release 6.0 or later. The workstation version of UniChem 2.0 will support Silicon Graphics IRIS workstations running IRIX 4.0 or later.

For more information about UniChem 2.0 on Cray Research systems, contact Mark Cole, telephone: 612/683-3688; email: mcole@cray.com.

HRB Systems' digital signal processing environment now available for Cray Research systems

HRB Systems, an E-Systems company, recently introduced ESY DSP, an enhanced digital signal processing environment for use on Cray Research systems. Using commercially available hardware and software, the ESY DSP package demonstrates that a general-purpose DSP flexible simulation system can also be a high-throughput production processing system.

ESY DSP software consists of five major components:

- Signal Processing Worksystem (SPW), a standards-based, general-purpose DSP system design toolkit with a block-flow-diagram user interface, from Comdisco Systems, Inc.
- Code Generation System (CGS), which generates the portable C code that implements the signal processing flow in the block flow diagram.
- Application Visualization System (AVS), a standards-based, general-purpose visualization design toolkit with a block-flow-diagram user interface, from AVS, Inc.
- SuperCGS, an enhancement to CGS which allows vector-based processing flows to transparently exploit the vector/parallel architectures of supercomputers
- SPW/AVS Link, a collection of tools that provides a seamless interface between the two major tools of ESY DSP and between the analyst and ESY DSP

ESY DSP hardware consists of three COTS components:

- Standards-based supercomputers for DSP
- Standards-based workstations for visualization
- Standards-based networks for transparent exchange of data and status and control between the supercomputer and graphics workstations

ESY DSP currently runs on CRAY Y-MP and CRAY Y-MP EL systems, as well as other supercomputers and workstations. ESY DSP also currently supports Ethernet and FDDI networks, with future support anticipated for HIPPI and/or ATM.

ESY DSP provides all the tools required for designing and running complex DSP/visualization applications in a distributed computing environment. Applications that ESY DSP has developed include a general-purpose communications system simulator and end-to-end signal processing for both real and simulated signals.

For more information about ESY DSP contact Russ Meyers, HRB Systems Inc., Mail Stop 911, PO Box 60, State College, PA 16804-0060; telephone: (800) 736-2797.

Cray Distributed Computing Environment Toolkit

The ability to distribute large, complex problems across heterogeneous computing environments is critical in solving today's most challenging scientific and engineering problems. By distributing the workload across several systems, users can use specific machines for what they do best, ensuring that they get the most from every resource on the network. For example, Cray Research supercomputers can be used to process large amounts of data, while workstations display the results. The ability to distribute the workload across systems is made possible through standards that enable these machines to communicate as efficiently as possible and make decisions about where best to run each part of a program. Cray Research now supports the emerging industry standard for multivendor distributed computing—the Open Software Foundation (OSF) Distributed Computing Environment (DCE).

In compliance with the OSF DCE standard, the Cray Research DCE Toolkit provides application developers with all the services needed to quickly and efficiently design and build distributed applications that operate across a variety of platforms. With the Cray Research DCE Toolkit, users can access all network resources and applications, regardless of which vendors' hardware or operating systems they are using or where the resources reside physically.

The Cray Research DCE Toolkit is written in C and uses standard interfaces for operating system services such as POSIX and X/Open guidelines. Conformance to these standards means that DCE can communicate with both UNIX and non-UNIX systems. Digital Equipment Corporation, IBM, Hewlett Packard, Sequent, and Sun Microsystems also support the OSF DCE standard.

The initial release of the DCE Toolkit includes client support for the core DCE components:

- Threads
- Remote procedure calls
- Naming
- Timing
- Security

This level of functionality enables Cray Research users to interact with other DCE servers to begin application development or porting of applications before full DCE client/server functionality is available. In mid-1994, Cray Research will introduce a DCE/DFS client/server product for Cray Research systems.

The DCE Toolkit runs on all Cray Research computer systems except CRAY-2 systems and requires the UNICOS 7.0 or UNICOS 7.C operating system.

For more information, please contact your Cray Research representative.

USER NEWS

Two oceanographers, one center win supercomputing awards

Cray Research congratulates oceanographers Robert Chervin, National Center for Atmospheric Research, and Albert Semtner, Naval Postgraduate School, who won the first Cray Research Information Technology Leadership Award for Breakthrough Computational Science. Kudos also were earned by the Pittsburgh Supercomputing Center, where researchers received a Computerworld Smithsonian Award for Science for simulating enzyme-DNA interaction using a CRAY Y-MP system.

Semtner and Chervin were recognized for their work in calculating, for the first time, the powerful effects of small-scale currents called "ocean eddies" on the global climate system. By tapping into the parallel processing ability of CRAY X-MP and CRAY Y-MP systems, Chervin and Semtner achieved what previously was unachievable—a high-resolution view of the earth's global ocean flow. The model revealed a long, looping feeder current that connects the Pacific, Atlantic, and Indian Oceans and helps to move heat and salt around the world.

Chervin and Semtner's work has major implications for future climate modeling. As computing power increases, scientists may be able to model the entire climate system—oceans, atmosphere, and landscape—and predict climate changes. Chervin and Semtner's ocean snapshot will become part of the permanent collection at the Smithsonian Institute's National Museum of American History.

Researchers John Rosenburg (University of Pittsburgh), Peter Kollman (University of California at San Francisco), and Robert Swendsen (Carnegie-Mellon University) used a CRAY Y-MP system at the Pittsburgh Supercomputing Center to simulate the interaction between DNA and an enzyme known as Eco RI endonuclease. Their research provides new knowledge about how proteins attach to segments of DNA.

Their award also recognizes the Pittsburgh Supercomputing Center's role in fostering an intellectual environment that allows multidisciplinary computational research teams to form. It was through the Center's series of seminars and workshops that Rosenburg teamed up with his co-researchers.

The Computerworld Smithsonian awards were established in 1988 to honor researchers from diverse disciplines who are using information technology to benefit society. The Cray Research Information Technology Leadership Award for Breakthrough Computational Science is awarded to researchers who, through the use of information technology, have made breakthrough advances in surmounting previously insurmountable obstacles.

Cray Research first company selected for Gordon Bell Prize competition

Cray Research is honored to be the first company ever chosen as a finalist for the Gordon Bell Prize competition. In 1993, Cray Research was chosen because entries submitted by customers, third-party vendors, and employees reported record-breaking time to solution for industrial problems on the CRAY C916 supercomputer.

Nine applications, including third-party codes, were submitted to the Gordon Bell Prize committee. Performance was obtained by balancing raw hardware speed; effective use of large, real, shared memory; compiler vectorization and Autotasking; hand optimization; asynchronous I/O techniques; and new algorithms. The highest performance for the submissions was 11.1 GFLOPS out of a potential peak performance of 16 GFLOPS for the CRAY C916 system. One program achieved a 15.45 speedup from the compiler. New I/O techniques hide tens of gigabytes of I/O behind parallel computations. Finally, new iterative solvers demonstrated times to solution on one CPU as high as 70 times faster than the best direct solvers.

Three computational fluid dynamics (CFD) applications were submitted. DRAG4D, which models unsteady, three-dimensional viscous incompressible flow for Nissan Motor Co., Ltd., achieved 6.87 GFLOPS on a CRAY C916 system using 60 Mwords of memory. The solver is more than 99 percent vectorized and 98 percent parallelized by the Cray Research CF77 compiling system. Nissan saves $400,000 per year on wind tunnel tests by using this CFD analysis. LANS3DUP, a three-dimensional steady/unsteady compressible Navier-Stokes/Euler code used by the Institute of Space and Astronautical Science in Japan, achieved 6.7 GFLOPS on 15 CPUs of a CRAY C916 supercomputer, representing a 12.27 speedup over the single-CPU run. And for MHD3D, a magnetohydrodynamics spectral code from NASA Goddard Space Flight Center which is used to study the evolution of the heliospheric plasma, a grid of 128^3 lattice points was solved. This corresponds to a flow Reynolds number of 1000, which is the largest computation that can be done conveniently today. This 128^3 grid required 97 Mwords of memory and 34 hours of CPU time on the CRAY C916 system and is not feasible on a workstation.

Two third-party computational chemistry applications demonstrated outstanding performance. MNDO91, a semiempirical code used to determine thermodynamic information such as heat of formation, was run on a large molecule, C_{960}, a member of the fullerene family. The problem used 150 Mwords of memory, ran in 39.5 minutes of elapsed time on the CRAY C916 system, and required 9.75 hours of CPU time. Based on elapsed time, the performance was 11.16 GFLOPS.

SUPERMOLECULE, the chemistry program formerly called DISCO, which originally was designed for multiple processors, performed at 7.19 GFLOPS on 16 processors of the CRAY C916 system.

Two seismic codes demonstrated very high levels of performance: a one-pass 3-D poststack depth migration, part of GECO-Prakla's GEOSYS production code; and the 3-D prestack Kirchhoff time or depth migration code written at Cray Research. The GEOSYS migration code's one-pass algorithm is implemented in Fortran; the code is fully vectorized, and Autotasking is used for multiple CPUs. The algorithm is a time-consuming finite-difference approach working in the frequency space domain. Fewer than 20 compiler directives are added to achieve multitasking. The migration is done in a cascade, with every CPU computing the result of one depth layer looping over all the frequencies. Since the CPUs do independent computations with the frequencies delayed by one for each successive CPU, more memory is required for every concurrently computed depth layer. Intermediate results cannot be held in memory, so asynchronous disk I/O on four channels runs parallel to computation.

For a test survey of 630 x 315 common depth points, a processing length of five seconds (or 1250 layers), and migration depth of 3.4 km, 10.1 GFLOPS was obtained on 16 processors of a CRAY C916 system in an elapsed time of 661 seconds while doing 12 Gbytes of I/O (or 18.5 Mbytes/s). This represents a speedup of 15.15 over the single-CPU performance.

The full 3-D Kirchhoff migration scheme was designed at Cray Research to study the feasibility of structural imaging of the earth on a general-purpose supercomputer. The algorithm has been optimized, vectorized, and microtasked. In fact, the Autotasking compiler achieves almost 100 percent parallelization. Performance reached almost 9 GFLOPS for one test problem; all problems ran in the 8-GFLOPS range. This means that a prestack migration can be completed in 30 hours for 10,000 shot gathers of prestack data for a 2 km by 2 km region on the surface and 500 grids in depth.

In the area of large-scale structural analysis, a 250,000 degree-of-freedom model of a proposed new design for the Space Shuttle liquid oxygen pump housing was modeled using the ANSYS code, a general-purpose, commercial finite-element analysis tool used in virtually every engineering discipline. Initially the analysis was done using 14 smaller subassembly models requiring a total of 29 separate runs on a CRAY X-MP system over about one week of elapsed time. On the CRAY C916 system, a single run of just over three hours was required to solve one large global model at a sustained rate of nearly 6 GFLOPS. During the analyses over 27 Gwords of data were transferred to disk in parallel with the computations while parallel processing in the equation solver effectively used multiple CPUs. Even greater throughput performance was obtained by running two separate analyses simultaneously, each using eight CPUs. A total of 45 Gwords of data were transferred to and from disk storage and both jobs completed in 4.5 hours, sustaining a combined rate of over 7 GFLOPS.

This Space Shuttle liquid oxygen pump housing model also won honorable mention in the Mannheim SuPar Cup '93 competition, which awards outstanding contributions in the field of parallel computing, for a novel application, a new parallel approach to an important algorithm, or a remarkable performance and/or speedup for a real application.

Finally, an acoustic radiation model of a car engine, a CFD application from a flow simulation, and a 10,000-equation computational electromagnetics problem were used to demonstrate the algorithmic advantage of new iterative solvers developed by scientists from Elegant Mathematics, Inc. and Cray Research.

For the car engine model, the estimated LU decomposition time for the 10,586 degree-of-freedom linear system is 3515 seconds, assuming a 900 MFLOPS, single-processor performance on a CRAY C90 system. The iterative solver required just 44.5 seconds to solve this problem to 10 orders of magnitude reduction in the error residual using only 12 iterations. On four processors, this time was further reduced to 23.6 seconds using an early implementation in which all parts were not yet parallelized. Compared to the idealized LU performance of 900 MFLOPS per CPU with perfect linear speedup, the new method is almost 80 times faster on one CPU and still over 30 times faster on four CPUs.

The CFD application demonstrated the algorithmic advantage of the new iterative solvers by comparing the nearly eightfold increase in the time for LU factorization, from 2500 seconds for a 10,472-equation-dense linear system to almost 20,000 seconds for a 20,682-equation-dense linear system, to the fourfold increase in iterative solution time, from 55 seconds to 210 seconds.

A 10,000-equation computational electromagnetics problem with over 700 right-hand sides demonstrated up to 2.7 GFLOPS performance on a four-processor CRAY C90 system for the iterative method. The method has also been used to solve 21K and 54K problems.

Sara Graffunder, senior director of applications at Cray Research, was invited to present this work at the Gordon Bell Prize Minisymposium during IEEE Supercomputing '93, in Portland, Oregon. Prize winners were chosen in three categories: performance, in which the submitted program must be shown to run faster than any other comparable engineering or scientific application; price/performance, in which the entrant must show that the performance of the application divided by the list price of the smallest system needed to achieve the reported performance is better than that of any other entry; and compiler parallelization, in which judges look for the combination of compiler and application that generates the greatest speedup.

Read all about it—online

Cray Research recently introduced an anonymous ftp public information service, ftp.cray.com, available to anyone with access to the Internet. This free service provides information about supercomputing and Cray Research to the worldwide Internet user community. Specific components include:

- ☐ Cray Research product announcements and press releases
- ☐ Information about Cray Research products
- ☐ Directory of application software available for Cray Research systems
- ☐ Articles of general interest to the supercomputing community
- ☐ Source code for some public domain programs that run on Cray Research systems

For more information about this new ftp service, telnet to ftp.cray.com, log in as anonymous, and enter your email address as the password.

GALLERY

A representation of the enzyme Eco RI endonuclease wrapping around DNA. The enzyme comprises two symmetrical subunits that correspond to the twofold symmetry of DNA's double helix. Each subunit kinks one of the DNA backbones and, like scissors, snips the DNA at the kink. Because of its ability to cut DNA at this precise location, Eco RI has become one of the biotechnology industry's most important tools for cloning DNA.

The calculations for the image were performed by the X-PLOR and AMBER software packages running on a Cray Research supercomputer at the Pittsburgh Supercomputing Center.

Image courtesy of John Rosenberg, University of Pittsburgh and the Pittsburgh Supercomputing Center.