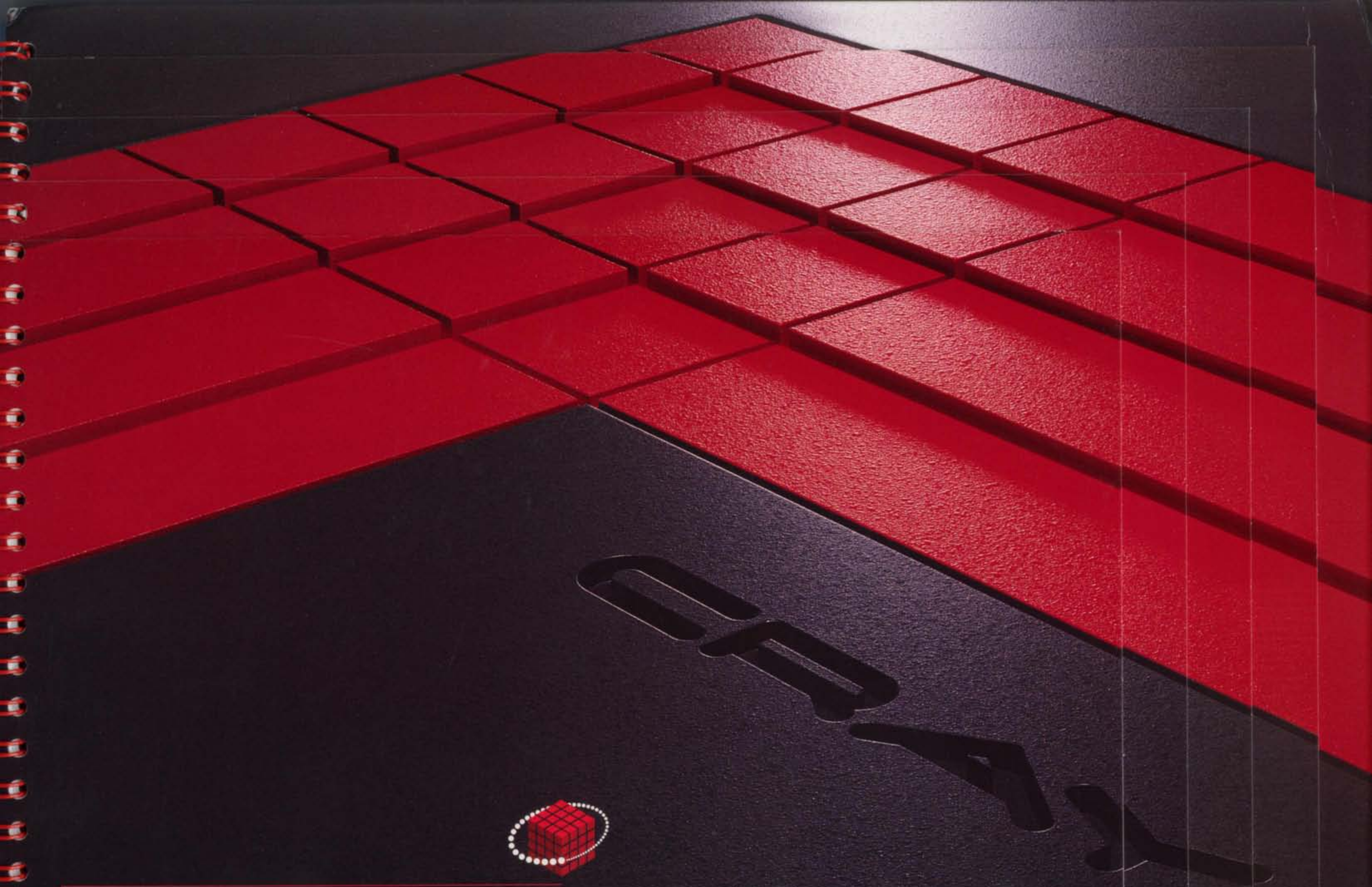


The CRAY T3E Series

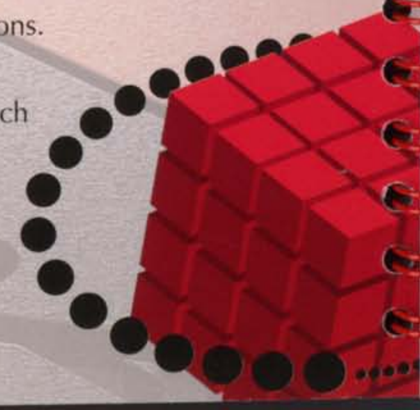
Scalable power that works

Scalable computing has reached a new level.

A new level where a single, powerful system grows with your applications. Where performance and price/performance scale steadily from 16 to 2048 processors. Where each feature and function is designed for larger problems, faster solutions, and a greater return on your investment.

A new level where true scalability is available from under $1 million. Where price no longer keeps budget-conscious organizations and departments from investigating scalable algorithms and developing production applications. Where university-based R&D centers, government laboratories, and automotive, petroleum, aerospace, weather, and environmental research organizations of all sizes have access to the most powerful computer on the planet.

A new level where flexibility, configurability, and reliability reign.



The CRAY T3E series

Introducing the CRAY T3E series, the world's most powerful and flexible scalable parallel systems. Across a range of configurations from 16 to 2048 processors with corresponding peak performance levels from gigaflops to teraflops, CRAY T3E systems deliver superior scalable performance and price/performance on technical applications in scientific and engineering R&D and industrial production computing.

The power and flexibility of this new generation come from global system scalability. Every CRAY T3E system delivers scalable performance through a combination of the world's fastest microprocessor and the industry's best interprocessor communications, the world's first truly scalable operating system, and breakthrough I/O technology that sets new standards for scalable I/O. In any configuration, every function and feature contributes to balanced scalable performance. Every price point offers tremendous room for growth with a corresponding payback in performance, throughput, and productivity.

In many system designs, scalable typically means "expandable hardware"—more processors and memory can be added to the system. But as users seek to attack larger problems, get faster solutions, and realize a return on investment, the limitations of expandable systems become apparent: interconnect bandwidth, operating system, and I/O bandwidth cannot keep up with the demands of expanding workloads; they become bottlenecks that compromise productivity.

With the CRAY T3E series, Cray Research's plan to be the first to deliver sustained teraflops performance by the end of the decade across a wide range of configurations and price points remains on track. This new generation redefines scalable performance and price/performance to deliver scalable power that works.

Liquid-cooled CRAY T3E system.

Scalable power that works

CRAY T3E scalable parallel systems build on the successful architecture and proven scalability of our first-generation scalable system. First delivered in 1993, the CRAY T3D system made Cray Research the global revenue leader in the technical market for highly scalable systems. We were first to report results for a scalable system with more than 1000 processing elements; a 1024-PE CRAY T3D system set performance and scalability records on the full set of NAS Parallel Benchmarks.

Researchers using the CRAY T3D system value its performance, reliability, and ease of use. Users around the world have put it to work on a wide range of established and emerging technical applications from long-term global climate prediction to petroleum exploration to traffic simulation.

Case studies throughout this brochure highlight some of this real-world work already accomplished on scalable systems from Cray Research.

Now, with the CRAY T3E series, Cray Research introduces global system scalability in a wider range of configurations with improved price/performance.

Phillips Petroleum Company

The easy-to-find sources of oil in the United States have already been explored and well developed. To keep a competitive commercial edge, U.S. petroleum companies now are turning to sites that are more structurally complex—located under and around salt bodies in the ocean, for example.

Phillips Petroleum Company of Bartlesville, Oklahoma, used its CRAY T3D massively parallel supercomputer to help assess the commercial value of the first subsalt oil field, dubbed "Mahogany," slated for development in the Gulf of Mexico. One specific technique Phillips used is 3-D prestack seismic depth migration, a complex term for turning huge volumes of raw, unprocessed data into a coherent image that accurately depicts the situation below the earth's surface. Incredibly compute-intensive because of the amount of seismic data that must be processed (as much as 25 million records, or traces) to create an image of a complex geology, 3-D prestack depth migration on previous computing platforms often exhausted the patience of geophysicists needing these images in a fast-enough-to-be-useful timeframe.

Able to distribute pieces of the problem across 128 powerful processors working in parallel in their CRAY T3D system, members of Phillips' Project IMAGE eliminated a lot of the

Scalable performance

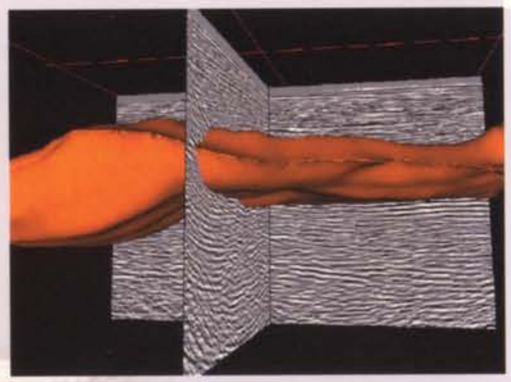
For top performance, we integrated the world's fastest microprocessor into a refined architecture that reflects our unmatched experience in designing the world's most powerful computer systems. CRAY T3E scalable parallel systems use the DECchip 21164 (DEC Alpha EV5) from Digital Equipment Corporation, capable of 600 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 32-bit and 64-bit floating-point arithmetic. CRAY T3E processing elements (PEs) include the DEC Alpha microprocessor, local memory, and performance-accelerating control logic designed by Cray Research engineers and fabricated in low-cost CMOS. As faster versions of the EV5 become available, CRAY T3E system customers will be able to upgrade their systems for additional performance gains.

Each PE has its own local DRAM memory with a capacity of from 64 Mbytes to 2 Gbytes. A shared, high-performance, globally addressable memory subsystem makes these memories accessible to every PE in a CRAY T3E system.

approximations that were necessary with older, more conventional seismic imaging methods. The team obtained higher quality images of large exploration areas in less time.

"3-D prestack depth migration on the T3D improves our ability to unravel the picture," said Kay Wyatt, Geophysical Senior Research Associate at Phillips. "The nature of the seismic data is very amenable to a massively parallel computer architecture. We can distribute the many millions of seismic traces in one 3-D survey for imaging across the many processors of the T3D system. The technical expertise we receive from Cray Research helps ensure we get the most from this system.

"We wanted to provide production 3-D prestack depth migration capability for large areas, not just smaller targeted areas as we had done before the arrival of the CRAY T3D system. We needed to push a lot of data through the system, and we didn't feel we could do this with existing computers. Without the T3D, we'd be limited to imaging very small subsets of the 3-D surveys, the value of which is greatly reduced. We're now able to see much more clearly and feel more comfortable with development decisions."



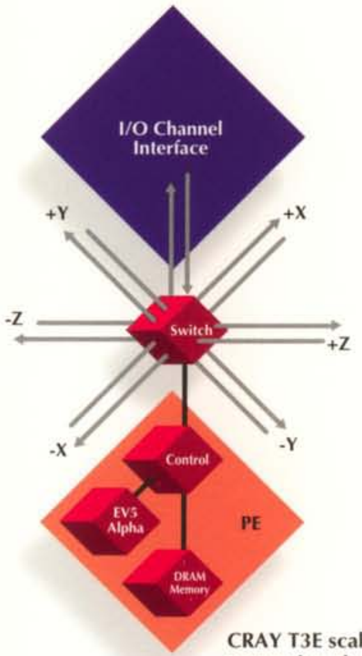
3-D representation of the tubular salt sheet above the Mahogany subsalt discovery. The salt is shown in its correct depth position with intersecting seismic profiles. Image created using Phillips' Model Building Software from the prestack depth migrated seismic volume processed using the CRAY T3D system.

Scalable interprocessor communications

PEs in the CRAY T3E system are connected by a high-bandwidth, low-latency bidirectional 3-D torus system interconnect network six times faster per PE than that used in our CRAY T3D system. We added adaptive routing to allow messages on the interconnect network to be rerouted around temporary "hot spots." Interprocessor data payload communication rates are 480 Mbytes per second in every direction through the torus; in a 512-PE CRAY T3E system, bisection bandwidth exceeds 122 Gbytes per second.

Latency hiding and synchronization are integrated in 512 off-chip memory-mapped E-registers. The latency of remote memory access can be hidden by using E-registers as a vector memory port, allowing a string of memory requests to flow to memory in a pipelined fashion, with results coming back while the PE works on other data. A range of synchronization mechanisms accommodates both Single Instruction Multiple Data (SIMD) and Multiple Instruction Multiple Data (MIMD) programming styles. Building on experience gained from our first-generation scalable system, we enhanced the barrier mechanisms within the network and increased the length and type of messaging facilities in the CRAY T3E system.

I/O channels are integrated into the 3-D torus network and increase in number with system size. As PEs are added to a CRAY T3E system, interprocessor and I/O bandwidth scale accordingly. This means that scalable applications run as efficiently on larger configurations as they do on smaller ones.



CRAY T3E scalable processing element

Edinburgh Parallel Computing Centre (EPCC)

A better way to orchestrate traffic flows is what software developers at Quadstone Ltd. had in mind when they designed the Paramics traffic simulation package. Located in Edinburgh, Scotland, Quadstone is demonstrating the benefits of modeling traffic flow computationally as an aid to traffic management and urban planning. On the CRAY T3D system, Quadstone analysts are using microscopic modeling techniques (in which vehicles are represented individually) rather than traditional macroscopic methods (in which vehicles are modeled collectively as if traffic flowed as a continuous fluid). It takes highly parallel supercomputers to make this microscopic methodology practical for realistic situations of concern to policymakers.

The United Kingdom Department of Transport recently commissioned a study of traffic on the London ring highway, the M25, now being widened. The traffic management consultancy SIAS Ltd., also in Edinburgh, estimates that a reduction by 20 percent in congestion delays through the Scottish trunk road network would produce about a £250 million ($400 million) annual savings for the Scottish economy. Quadstone analysts are using the Paramics package to simulate traffic management on the highway using programmable speed-limit signs. Another project involves predicting the effects on traffic of a proposed bridge to be built in Wideracren, Scotland. And Quadstone is using Paramics to anticipate the effects of constructing a proposed shopping mall

Scalable operating system

To support the scalability of the CRAY T3E system, we created UNICOS/mk, a scalable version of our UNICOS operating system. UNICOS/mk is distributed among the CRAY T3E system's PEs, not replicated on each. This distribution of operating system functions provides a global view of the computing environment—a single-system image—that allows administrators to manage a systemwide suite of resources as a single entity. As a result, systemwide resource management and utilization are greatly simplified compared to most other parallel computing environments.

UNICOS/mk is divided into "servers," which are distributed among the processors of a CRAY T3E system. Local servers process operating system requests specific to each user PE. Global servers provide systemwide operating system capabilities, such as process management and file allocation. The result is the world's first scalable operating system.

In addition to the user PEs that run applications and commands, CRAY T3E systems include dedicated system PEs which run the global UNICOS/mk servers. Because global services are provided by system PEs and not replicated throughout the CRAY T3E system, UNICOS/mk efficiently scales, with full functionality, to service from tens to thousands of PEs with minimal overhead.

Through standard UNIX operating system utilities and commands and the same interface as our UNICOS operating system, UNICOS/mk provides a familiar operating environment for users and administrators alike. Scalability is enhanced through 64-bit addressing, which allows the operating system to support large data sets and file systems and thousands of users.

in Aberdeen. Planners hope to identify likely bottlenecks in traffic flow and determine alternative routing schemes.

"Using the highly parallel CRAY T3D supercomputer, we can model complex traffic flows considerably faster than realtime," explained Mark Smith, Quadstone marketing director. "Twelve times realtime is a realistic possibility, which would give us the likely congestion an hour from now in just five minutes, given current ground positions from remote sensors. Alternative routes then could be suggested to motorists via programmable message signs along the highway.

"Microscopic methods were held back by the lack of computing power. The CRAY T3D not only lets us tackle more complex problems, but Cray's standard software environment lets us do development work on workstations and move the code directly to the supercomputer for computationally intensive production runs. At this point we can model 120,000 vehicles on 30,000 kilometers of road at three times realtime rates on 32 nodes of the CRAY T3D. The Paramics-supercomputer combination opens many new possibilities for addressing practical traffic planning and management problems."

Visual interface to the Paramics traffic modeling program. In combination with highly parallel computing systems, the program can give traffic managers and city planners valuable insight into traffic flow patterns and routing options.

UNICOS/mk complies with the POSIX 1003.1 (system services) and POSIX 1003.2 (commands and utilities) standards. But UNICOS/mk does more than ensure a familiar, open operating environment through compliance with industry standards. It also incorporates our more than 10 years of experience in high-performance parallel UNIX operating systems and includes significant enhancements for scalability; system availability; resource management; and scalable, high-bandwidth I/O.

UNICOS/mk is easy to install, use, and administer. Automatic resource management utilities such as the Unified Resource Manager, User Database, and Data Migration Facility minimize the need for manual system monitoring. When operator intervention is required, administrators can monitor system utilization and activity with a variety of monitoring tools. A visual installation tool helps administrators customize their systems and allows them to enable only the features they need.

UNICOS/mk supports our scalable I/O architecture, which provides I/O capability that scales with system size and ensures superior I/O capacity for high-performance, scalable applications. UNICOS/mk delivers this performance through software innovations including

- ❑ Distribution of file management services: File server functions are distributed using local file server assistants to provide maximum efficiency and performance.

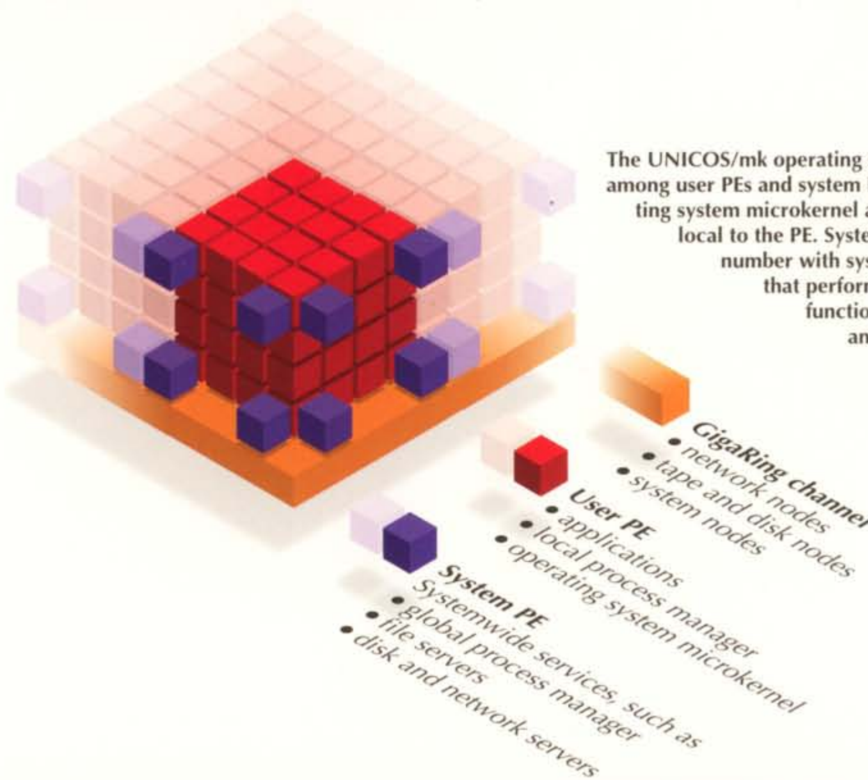
Pittsburgh Supercomputing Center

A solid understanding of how a human being sees could lead to computers that see, too. Computer scientists used a CRAY T3D massively parallel processing system to run the first computer model to mimic in great detail the process the human brain uses to see. Able to establish a strong connection between their simulation and observed data from neuroscientists and brain researchers, the computer scientists hope this breakthrough work can be used to replicate the human vision process in a host of artificial vision systems.

Using 256 processors working in parallel in the CRAY T3D system at the Pittsburgh Supercomputing Center, University of Texas at Austin computer scientist Risto Miikkulainen and Ph.D. candidate Joseph Sirosh were able to successfully model the neural network—billions of interconnected neurons—governing human vision. "We couldn't have done this 10 years ago because the technological capability to simulate these large neural networks didn't exist," said Sirosh. "You need to model at the scale of tens of thousands of neurons, hundreds of millions of connections. Machines like the T3D are absolutely crucial to this work. The large memory allows us to store all of the possibilities for neuron-to-neuron connections. Because the algorithms we're using are highly parallel to begin with, it's fairly straightforward to spread the work across the many processors of the T3D and allow them to work simultaneously on this huge computational problem."

❑ Distio (I/O centrifuge): A single distio request from one PE can generate parallel I/O transfers that distribute data among some or all of the PEs used by a given program.

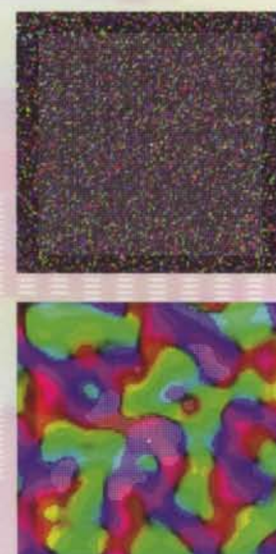
❑ Multiple global file servers: File system management is distributed among multiple system PEs, which allows full use of the parallel disk paths supported on CRAY T3E systems.



The UNICOS/mk operating system distributes its functions among user PEs and system PEs. Each user PE hosts an operating system microkernel and servers that manage requests local to the PE. System PEs, which also increase in number with system size, host global servers that perform systemwide operating system functions, such as process management and file allocation. This distributed design makes UNICOS/mk the world's first scalable operating system.

Sirosh and Miikkulainen's simulation is the first model of the primary visual cortex—responsible for most of the complex neurological processes required to bring us a colorful, three-dimensional world—to simulate what is now believed to be the "self-organizing" process of the brain's billions of interconnected neurons. In the visual cortex, neurons that respond to vertical lines—trees or telephone poles in the visual field, for example—are grouped together in patches called orientation columns. Other orientation columns respond to horizontal lines, or 45° lines. Still other neuron groupings form according to whether they respond predominantly to the right or left eye.

"Neurons that reinforce each other wire together, and unnecessary connections are eliminated," said Miikkulainen. "Such filtering eliminates redundancies and allows the brain to efficiently process the massive amounts of visual information impinging on the retinas. Modeling this filtering process on the T3D allows us to predict how lateral connections organize in the visual cortex and what visual knowledge is stored in them. It suggests that neural modeling will play a vital role as vision research presses on to explore these higher-level processes."



Using oriented light spots as input, Sirosh and Miikkulainen's simulation modeled how orientation preference and neuron-to-neuron connection patterns develop in the cortex.

Initially (top left), the orientation preferences were random, and lateral connections covered almost the entire map. After a series of several thousand inputs, the neurons organize (bottom left) into orientation columns, and lateral connections link areas of similar orientation preference. These patterns agree in key features with maps obtained by experimental imaging.

Programming environments

To simplify the porting and development of high-performance applications, Cray Research programming environments provide an integrated set of the most powerful application development tools available: optimizing compilers, high-performance parallel libraries, advanced debugging and performance analysis tools, and the most flexible and complete parallel programming models in the industry. These tools are designed to complement and enhance each other, resulting in a complete, easy-to-use programming environment that facilitates the development of scalable applications.

For more than a decade, Cray Research has been the recognized leader in optimizing compilers. This tradition continues with the Fortran 90, C, and C++ compilers for the CRAY T3E system. These compilers go beyond code optimization to provide access to the most comprehensive parallel programming models in the industry. Each compiler supports explicit parallel programming via the Cray Message Passing Toolkit. The toolkit contains high-performance implementations of the standard PVM and MPI libraries, providing fast point-to-point and global message passing communication.

Explicit parallel programming capabilities are further extended through the innovative Cray Shared Memory Library, which provides data passing or one-sided communications. This library delivers the fastest interprocessor communication in the industry.

Fortran programmers also can benefit from the simplicity of implicit parallel programming, with system software based on the CRAFT and HPF parallel models. CRAFT is an implicit parallel model developed by Cray Research for the CRAY T3D system to provide users with flexible

EPFL (Ecole Polytechnique Fédérale de Lausanne)

At EPFL, Directview allows near-realtime video compression on the CRAY T3D system and the visualization of the obtained decompressed images on a workstation.

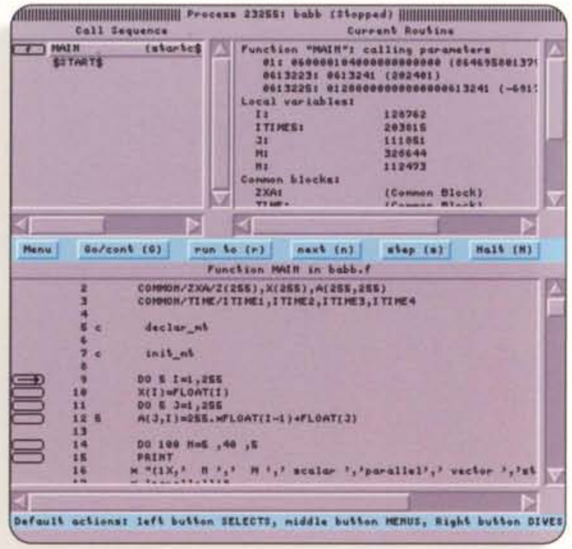
The digital age is bringing into focus the full potential of television and other visual media. Just as compact discs supplanted vinyl records, television and movies are headed in the digital direction as well. Digital representation—encoding images in discrete bits of data—is much more reliable and flexible than the analog waves that carry TV images today. Copying a set of numbers yields the exact same set of numbers on the first, fiftieth, and five-hundredth generations.

The challenge comes in trying to store, manage, and deliver all of this digitized data, these massive strings of numbers, as quickly as necessary (30 images per second for TV) with available technologies and without exorbitant costs. Advances in asynchronous transfer mode (ATM) technology, which allows voice, data, and video transmission across a single physical connection, and improvements in fiber optic cables suggest that telephone lines will be an important mode for digital television and videotelephony in the future. Still, the sheer volume of data in these video streams will remain prohibitive in terms of both time and expense.

Part of the solution is data compression—reducing these huge video streams to data levels that existing and emerging connections can handle. The ultimate question is how to achieve meaningful data work distribution among processors and high-level control over data distribution. The HPF implicit parallel model addresses many of these same issues and is an open model that provides increased application portability.

Our programming environments also include visual programming and analysis tools that give developers detailed insight into code performance and structure, helping them write code more efficiently. These tools support the development of applications written in any combination of the CRAY T3E system's programming languages (Fortran 90, C, and C++) and programming models (MPI, PVM, Shared Memory Library, and CRAFT/HPF). These tools include

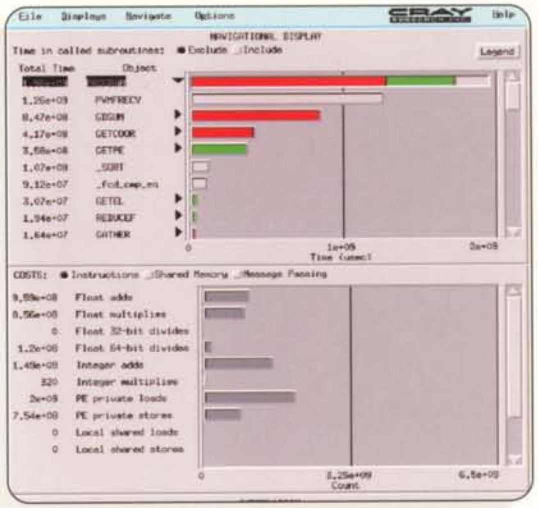
- ❑ The Cray TotalView Debugger, an advanced debugger designed for symbolic, source-level debugging of parallel applications. Cray TotalView allows users to control and display information about the individual processes of their parallel applications and provides a higher-level view to control debugging at the application level.



Above, the Cray TotalView Debugger, an advanced debugger designed for symbolic, source-level debugging of parallel applications.

reduction without compromising the quality that digitization provides—how to get the best, or highest, ratio of original image to compressed image.

On the 256-processor CRAY T3D system at the Swiss Federal Institute of Technology (EPFL), parallel implementation of digital image compression algorithms has allowed researchers to easily perform all of the simulations needed for optimization. For some algorithms, the CRAY T3D system compresses several images per second—25 images per second in realtime has been reached in some cases. What's more, EPFL researchers have developed a system on the CRAY T3D supercomputer that allows them to modify the algorithms interactively. They can see, very quickly and efficiently, the effects of various modifications on the quality of the compressed images. "The ability to perform this visual evaluation is crucial since there is no efficient mathematical criterion to evaluate the quality of an image," said EPFL professor Murat Kunt. "It's important to note that such visual evaluation is not possible with serial implementation of compression algorithms. The T3D represents a very important advantage over classical serial computers, and the T3E system will make possible visual evaluation for more complex algorithms."



MPP Apprentice, a state-of-the-art parallel performance analysis tool.

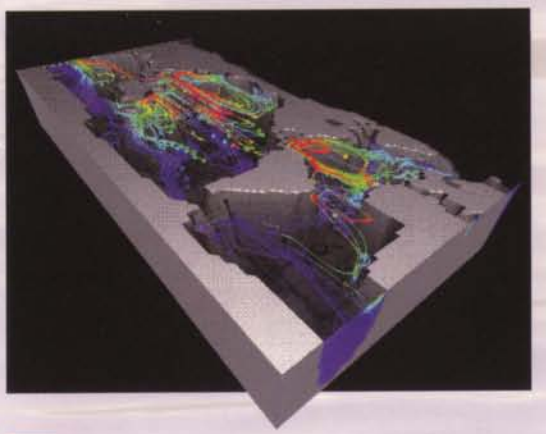
- The Cray Program Browser, an interactive environment for browsing and editing files and applications.
- MPP Apprentice, a state-of-the-art parallel performance analysis tool. MPP Apprentice not only displays performance data, it interprets the information and provides recommendations to improve performance. It provides information at the application level, relating all data directly back to the original source code using the program browser. The result is a program view scalable to thousands of PEs and able to accommodate full-length programs and very large data sets.

In addition, we offer a suite of optimized and parallelized mathematical and scientific libraries. These libraries serve as building blocks for developing high-performance parallel code. Our Fortran and C compilers automatically detect and replace user code with calls to the appropriate library routines.

German Climate Computing Center

If scientists can accurately predict the global climate 50, 10, even 5 years from now, world leaders can make informed decisions and, presumably, the best use of taxpayer dollars, to keep their countries at optimal agricultural and industrial levels and maintain the environmental health of the planet. Credible climate prediction depends on powerful computers. The CRAY T3D system has helped advance the technologies that improve long-term global climate prediction.

Computer simulations that realistically model all of the important climate components—ocean, air, land, and sea ice—as well as their interactions with each other, must run for incredibly long periods of time to produce the futuristic view scientists are looking for. The answers lie in combining ocean, air, land, and sea ice models in one experiment and simulating this model for hundreds and thousands of years at a time to capture the slow changes in the ocean. This combination, or "coupling," involves a lot of information, multiple experiments, and the ability of scientists to change model parameters from one run to the next to reflect the dynamic interchange among ocean, land, air, and ice. On the most powerful vector SMP systems available today, for example, a simulation of several years into the future can take many weeks to run. To provide decision makers with timely, practical information, scientists need results faster—especially when they want to tackle climate prediction for decades and centuries.



OPYC reveals a three-dimensional flow in the global ocean basin visualized by particle trajectories, where particles have been released at specified locations. Trajectories in blue indicate low temperature, red indicates high temperature.

Scalable I/O architecture

Breakthrough I/O technology available for the first time on the CRAY T3E system sets new standards for I/O scalability, performance, capacity, and configurability. Designed to work with all other components of the system for peak efficiency, this new architecture is designed to provide a balance between CPU, memory, and I/O demands.

The CRAY T3E system performs I/O through multiple ports onto one or more scalable GigaRing channels. This dual-ring I/O channel, with data in the two rings traveling in opposite directions, delivers high I/O data bandwidth and enhances reliability.

This high-performance I/O architecture scales to meet your needs by growing in step with your computing system resources. Regardless of size, all CRAY T3E systems may be configured with an I/O channel for every eight PEs (air-cooled) or 16 PEs (liquid-cooled). All I/O channels are accessible and controllable from all PEs. On the CRAY T3E system, each GigaRing channel has a maximum data payload bandwidth of 1 Gbyte/s and provides high-speed access to peripherals, networks, and other Cray Research systems.

The scalable I/O architecture is so flexible that most configuration designers will find they are nearly free from configurability constraints. For example, a single disk node on an I/O channel can provide access to up to 1.6 Tbytes of disk storage. On a CRAY T3E system, multiple I/O channels and I/O nodes can easily be configured to support up to quadrillions of bytes (petabytes) of disk capacity.

Today the CRAY T3D massively parallel processing (MPP) system gives these scientists the opportunity to put more processors to work simultaneously on a single global climate model, and thus complete the model faster, or to run more than one simulation in the same or less time than one model currently takes on a vector SMP system. In addition, scientists now can run their models at higher resolution (depicting climate components in ever-increasing detail) without compromising turnaround times; in the past, they often sacrificed resolution in favor of running the model as quickly as possible.

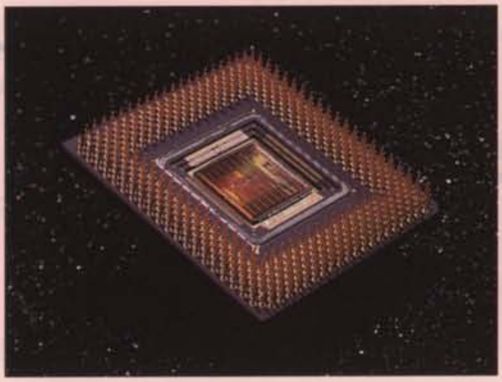
Researchers at DKRZ, the German Climate Computing Center, and computer specialists from Cray Research have moved a global circulation model already running on a CRAY C90 vector SMP system to the CRAY T3D system. A first production run of this parallelized model ran efficiently on a 32-processor CRAY T3D system. "This scientific experiment confirmed the reliability of Cray Research's MPP system," said Joseph Oberhuber, a modeler at DKRZ. "Porting the global circulation model is an important step toward finding a practical, affordable solution for computationally intensive ocean modeling."

If your needs call for maximum I/O performance rather than sheer capacity, the number of GigaRing channels and nodes can be configured to meet the most demanding workloads in the industry. The largest CRAY T3E system configurations are capable of up to 128 Gbytes/s of I/O bandwidth. Multiple channels can be configured to provide maximum disk throughput while maintaining the highest overall I/O bandwidth. Disk, tape, network, and system nodes can be configured in any number of combinations on multiple I/O channels, enabling you to precisely tailor a balanced I/O solution that meets your needs.

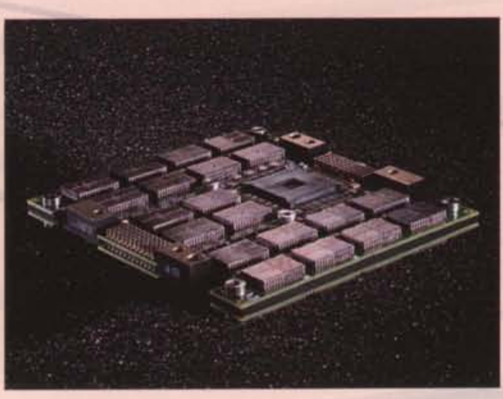
Each GigaRing I/O channel can be configured with multiple nodes of various types:

- ❑ A Multipurpose Node that accepts standard Sbus type controller cards to support FDDI, Ethernet, ATM, SCSI disk, and SCSI tape access.
- ❑ Disk drive nodes that provide connectivity to disk arrays and single disks over IPI and Fibre Channel Arbitrated Loop industry standard channels.
- ❑ Tape drive nodes that provide access to Block Mux and ESCON tape devices.
- ❑ Network nodes providing ANSI 32-bit or 64-bit HIPPI channels (100 or 200 Mbytes/s).

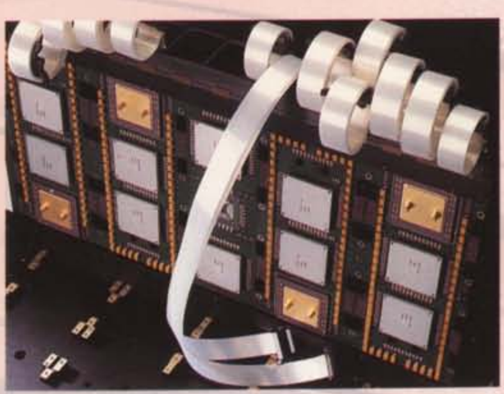
CRAY T3E System Technology Profile



DEC Alpha EV5 microprocessor, capable of 600 MFLOPS peak performance. CRAY T3E system PEs include this industry-leading RISC chip, local memory, and performance-accelerating control logic designed by Cray Research engineers.



Two memory daughtercards per PE form the basis of the physically distributed, logically shared memory of the CRAY T3E system. Two-high stacking (shown here) allows a storage capacity of 128 or 512 Mbytes on one memory daughtercard.

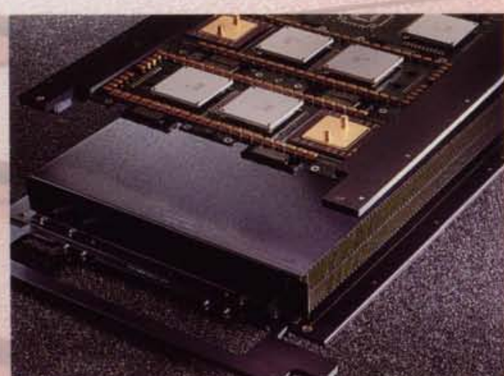


Motherboard with four CRAY T3E processing elements. Two PEs are shown on each side of the centrally located I/O circuits. The 3-D torus interconnect cables and the GigaRing I/O cables are shown at top.

CRAY T3E I/O with GigaRing channels and various I/O nodes. Peripherals and network interfaces connect to an I/O channel via the I/O node. The rings of a channel operate independently, providing high bandwidth and dual and resilient data paths.

Liquid-cooled (LC) CRAY T3E system module. Two PE motherboards are mounted horizontally, one on top of and one underneath the LC coldplate. Cooling circuit is shown in red and blue.

Air-cooled (AC) CRAY T3E system modules include a PE motherboard and memory daughtercards. These modules are mounted vertically in the AC cabinet; air is drawn across the fins on the coldplate shown here.



Reliability, availability, serviceability

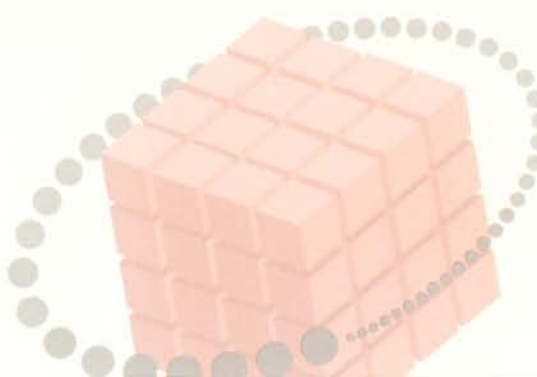
Air-cooled CRAY T3E system.

The hardware, software, and I/O subsystems of the CRAY T3E series are designed to ensure the highest possible levels of system availability for production computing.

Within the CRAY T3E system itself, failing components (PEs, links, or I/O channels) can be disabled individually. Support modules provide an average of one redundant PE for every 128 user PEs. In the event of a user PE failure, a redundant PE is configured into the system. A module with a failing user PE can be physically replaced by using the hot-swap mechanism to insert a module of working PEs. All of this can be done without rebooting the system.

For over 10 years, Cray Research's UNICOS operating system and programming environments have set the standard within the supercomputing industry for reliability and availability. To further enhance availability, UNICOS/mk is designed to exploit the redundancy features of the CRAY T3E system; it includes checkpoint/restart and other resiliency features that preserve work in progress if processing is interrupted by system failure or by an external circumstance such as power failure.

Each GigaRing channel is connected to the CRAY T3E system torus and its I/O subsystem via multiple ports. In the event that an I/O path becomes unavailable, GigaRing channels and the CRAY T3E system can be reconfigured to route data through alternate paths, ensuring the highest levels of system availability. This feature also makes it possible to perform hot swaps of I/O node modules without system downtime.



CRAY T3E—Scalable power that works

Power and performance from the world's fastest microprocessor.

Superior price/performance from competitively priced entry-level systems to affordable high-end solutions for the world's greatest challenges.

Breakthrough hardware, software, and I/O designs balanced to deliver the same efficient performance in every configuration and to quickly exploit scalable parallel applications developed for any scalable platform.

Hardware, software, and I/O configured to meet each customer's needs.

Low cost of ownership through improved reliability, availability, and serviceability.

The CRAY T3E system from Cray Research—scalable power that works.

System Configurations

Air-cooled models	User PEs[1]	Total Memory[2]	Peak Performance	Available I/O Bandwidth[3]
AC16 to AC128[4]	16 to 128	1 to 256 Gbytes	9.6 to 76.8 GFLOPS	2 to 16 Gbytes/s
Liquid-cooled models				
LC64 to LC2048[5]	64 to 2048	4 Gbytes to 4 Tbytes	38.4 GFLOPS to 1.2 TFLOPS	4 to 128 Gbytes/s

1. In addition to user PEs, CRAY T3E systems are configured with an average of one support module of PEs for every 16 modules of user PEs. Support PEs are used for global operating system services and to provide redundancy in the event of a user PE failure.

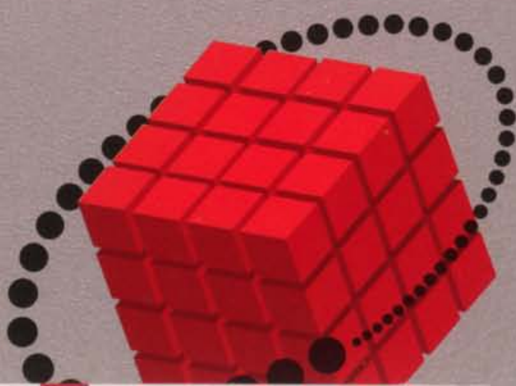
2. PEs may be configured with one of six memory sizes: 64, 128, 256, or 512 Mbytes, 1 Gbyte, or 2 Gbytes. Within a module, all PEs have the same memory size.

3. I/O channels are connected to two different points within the CRAY T3E system interconnect network. This provides maximum bandwidth and resiliency.

4. Air-cooled systems have from 1 to 6 cabinets and are configured in increments of 4 PEs. Each air-cooled module includes 4 PEs and one GigaRing interface.

5. Liquid-cooled systems have from 1 to 8 cabinets and are configured in increments of 8 PEs. Each liquid-cooled module includes 8 PEs and one GigaRing interface.

CRAY RESEARCH, INC.
655A Lone Oak Drive
Eagan, Minnesota 55121
U.S.A.



Research for the development of the CRAY T3E system and the GigaRing channel was partially supported by the Advanced Research Projects Agency under Agreement #MDA972-95-3-0032, dated 12 January 1995.

MCPB-1581095