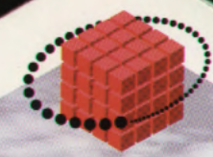
CRAY, CRAY Y-MP, and UNICOS are registered trademarks, and CRAFT, CRAY C90, CRAY T3D, CrayTools, MPP Apprentice, Xbrowse, and UNICOS MAX are trademarks of Cray Research, Inc. The Cray TotalView debugger is based on the TotalView debugger from Bolt, Beranek and Newman Inc. TotalView is a trademark of Bolt, Beranek and Newman Inc. Development of the CRAFT programming model was partially supported by the Advanced Research Projects Agency under Agreement #MDA972-92-H-0002, dated 21 January 1992. CDC is a trademark of Control Data Systems, Inc. DECchip is a trademark of Digital Equipment Corporation. Ethernet is a trademark of Xerox Corporation. IBM is a trademark of International Business Machines Corporation. UNIX is a trademark of UNIX System Laboratories, Inc. X Window System is a trademark of the Massachusetts Institute of Technology.

MCPB-1330893 ©1993 Cray Research, Inc.

CRAY RESEARCH, INC.
655A Lone Oak Drive
Eagan, MN 55121

CRAY T3D THE RIGHT TOOL AT THE RIGHT TIME



"The CRAY T3D has the highest bisection bandwidth and the lowest end-to-end internode latency of any commercial MPP in the market today. This superior communication performance, combined with a global shared address space, greatly simplifies programming."

William Dally, Professor of Electrical Engineering and Computer Science, Massachusetts Institute of Technology

CRAY T3D—the right system for performance

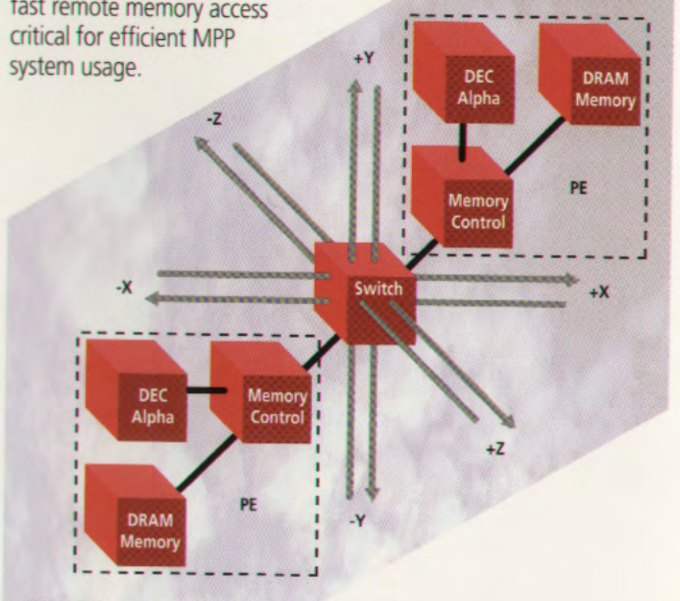
Cray Research has been designing the world's leading supercomputers for the past two decades. We drew upon this experience to bring customers the first MPP system based on true supercomputing technology. The CRAY T3D system integrates industry-leading microprocessors with a Cray Research-designed system interconnect and high-speed synchronization mechanisms to produce a balanced, scalable MPP. The system software and programming models exploit the hardware to sustain high performance on user applications. Here are the underlying design features that empower the CRAY T3D system to deliver its performance:

High-speed microprocessors

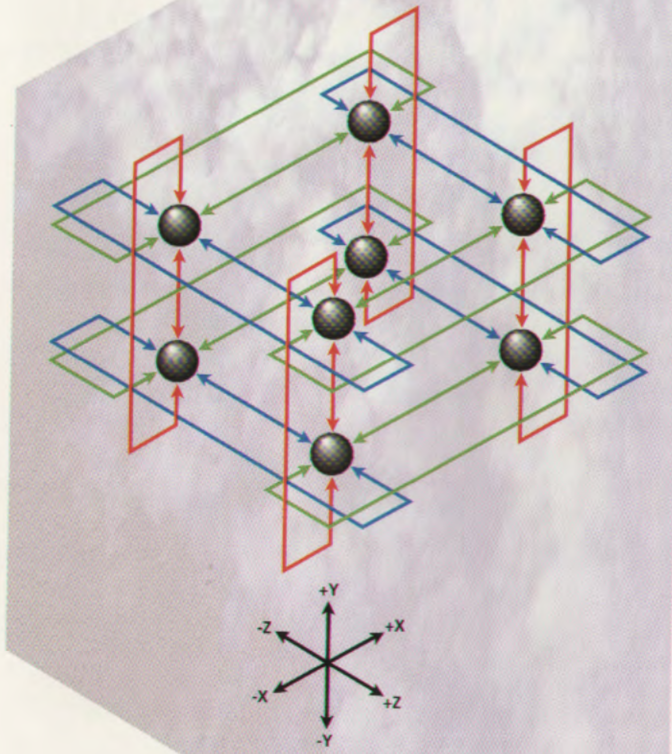
The CRAY T3D system incorporates a state-of-the-art microprocessor—the DECchip 21064 (more familiarly known as the DEC Alpha) from Digital Equipment Corporation—capable of 150 MFLOPS peak performance. This reduced instruction set computing (RISC) microprocessor is cache-based, has pipelined functional 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. Each processing element (PE) in the CRAY T3D system comprises the DEC Alpha microprocessor, local memory, and Cray Research-designed support logic. The PE is the basic computational unit in the CRAY T3D system's multiple instruction multiple data (MIMD) architecture. A CRAY T3D system node consists of two PEs sharing the Cray Research-designed switch and network support logic.

High-speed interprocessor communications

The PEs are connected by a very fast bidirectional 3-D torus system interconnect network. This topology ensures short connection paths and high bisection bandwidth (the maximum rate at which one half of the system can exchange data with the other half). 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, this design allows the extremely fast remote memory access critical for efficient MPP system usage.



CRAY T3D compute node



Three-dimensional torus network

Latency hiding

The CRAY T3D system design supports three major latency hiding mechanisms that augment the high-performance system interconnect. (Latency is the time taken to service a request, deliver a message, or perform some other function that is independent of the size or nature of the operation.) These latency hiding mechanisms are the prefetch queue (PFQ), read ahead, and Block Transfer Engine (BLT). The PFQ and read-ahead features are designed to hide latency at the level of a few words, while BLT is designed to hide latency at the level of hundreds or thousands of words.

Fast synchronization

Microprocessor design is focused on the requirements of the workstation marketplace. However, the synchronization requirements of large-scale MPP systems go significantly beyond those of workstations. Hence, Cray Research has developed a range of synchronization mechanisms, including Barrier/Eureka Synchronization, Fetch-and-Increment Register, Atomic Swap, and Messaging Facility. With these mechanisms, we accommodate both Single Instruction Multiple Data (SIMD) and Multiple Instruction Multiple Data (MIMD) programming styles.

High-bandwidth, parallel I/O

The high-bandwidth I/O subsystem, based on Cray Research's Model E I/O subsystem technology, provides access to the full range of Cray Research disk, tape, and network peripherals. Peak performance in the gigabytes per second range can be achieved.

Scalability

The CRAY T3D system is designed to be scalable in all aspects; from 32 to 2048 processors, from 4.2 to over 300 GFLOPS peak performance, from 512 Mbytes to 128 Gbytes of memory, and terabytes of storage capacity. This scalability is used not only for improving system throughput by increasing the number of job streams within the system, but also for significantly increasing the performance of individual user jobs—which many would consider the true test of scalability for a highly parallel system.

"Having all three emerging programming standards supported on the system will help us get more work done on this MPP system. We'll be able to mix and match these methods—even within an application—on the CRAY T3D."

Harold Trease, Staff Member, Theoretical and Applied Physics Division, Los Alamos National Laboratory, Los Alamos, New Mexico

CRAY T3D—the right system for productivity and ease of use

High-level programming models and the supporting development tools define the programmability and ease of use of high-performance computer systems. Users of the CRAY T3D system can choose from powerful Fortran or C programming environments. Each environment includes an industry-leading optimizing compiler, advanced performance analysis tools, programming tools with industry standard visual interfaces, and high-performance scientific and I/O libraries. The following hardware and software characteristics enable Cray Research to deliver unprecedented MPP productivity to the user.

Shared distributed memory

Cray Research recognized early in the design process that significant programming and ease-of-use benefits follow from being able to view the memory as a shared distributed resource. 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. Memory is physically distributed among processors, with either 16 or 64 Mbytes of DRAM memory in each PE. 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—to support high performance on applications with fine, medium, and coarse-grained parallelism.

CRAFT—the Cray Research Adaptive Fortran programming model

A number of styles for programming MPP systems have evolved in the industry, defined as much by the limitations as by the strengths of earlier-generation MPP hardware designs. Freed from these constraints by the global addressability and high-performance communication and synchronization of the CRAY T3D system, Cray Research's CRAFT programming model incorporates both the traditional data parallel and message passing MPP programming styles with the newer work sharing capability. As a result of CRAFT's broad support of programming techniques, parallel work can be distributed at all levels within a program, from the subroutine, to the individual loop, to an array. Parallelism can be expressed explicitly through message passing library calls, or implicitly through data parallel constructs and directives, taking advantage of the compiler to distribute data and work.

CRAFT also maximizes programmer productivity. With CRAFT, programmers are not locked into a programming method that fits only a SIMD or MIMD architecture. Rather, they can work in the programming style that best fits their applications or, indeed, any portions of them—a choice not previously available to users. The work sharing style lets users who are familiar with large common memories view their applications as they did before. With support for work sharing, data parallel, and message passing styles, CRAFT provides an optimal application development platform that also enables straightforward porting of many existing MPP applications. As an extension to Fortran 77, CRAFT incorporates useful Fortran 90 features such as array syntax and array intrinsics.

Applications development for CRAY T3D systems can start on CRAY Y-MP, CRAY C90, or CRAY EL90 systems.

The CRAY T3D Emulator helps programmers using the CF77 Fortran programming environment to map, develop, and test codes for the CRAY T3D system on Cray Research parallel vector systems.

The C programming model

The C programming model provides portability by using a highly-optimized Parallel Virtual Machine (PVM) implementation of message passing. Users will recognize the familiar standard interface.

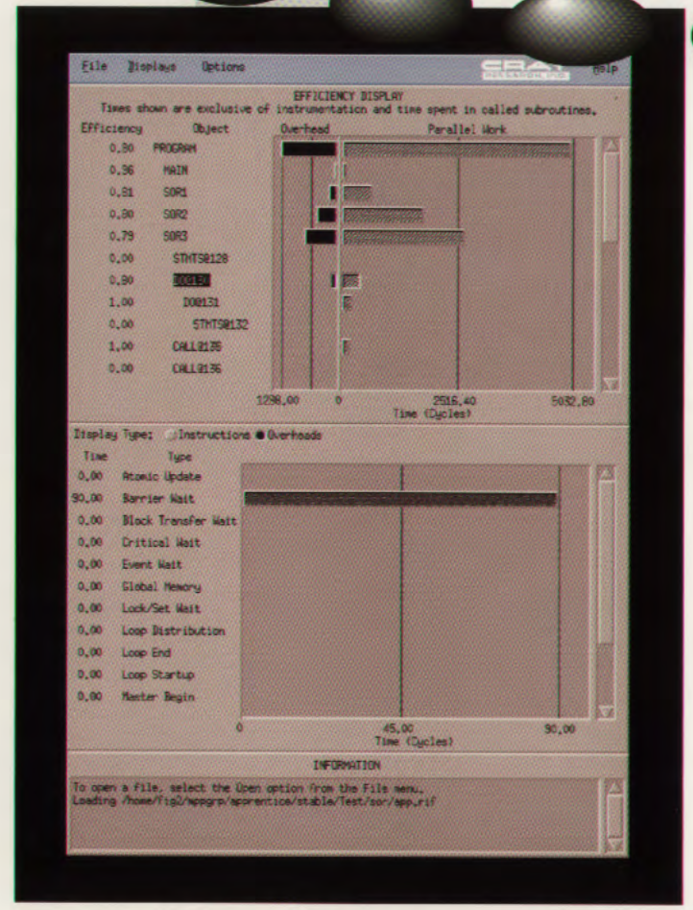
Development tools

Cray Research has extensive experience developing application support tools for parallel systems. For the CRAY T3D system, the CrayTools development suite comprises a performance analyzer called MPP Apprentice, the Cray TotalView debugger, and a range of programming utilities.

MPP Apprentice is a visually-based tool for analyzing the performance of applications running on the CRAY T3D system. This tool includes a display of routine and code region execution efficiency, and an instruction display providing hardware performance statistics (including floating point unit counts, cache hits, and system interconnect delays). MPP Apprentice does more than display performance data. It interprets the information and provides recommendations to improve performance. Interoperation with the Xbrowse source analysis tool relates information back to the original source code.

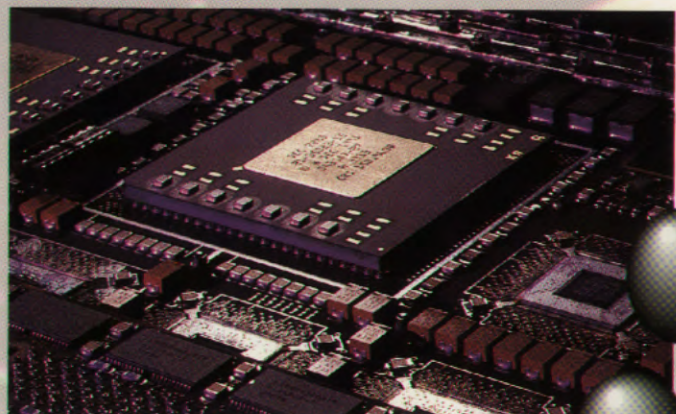
Cray TotalView is an advanced, symbolic, source-level debugger. It provides programmers with a window-oriented interface to use in symbolic debugging of multiprocessed applications. Cray TotalView allows users to control and display information about the individual processes of their MPP applications, and provides a higher-level view to control debugging at the application level.

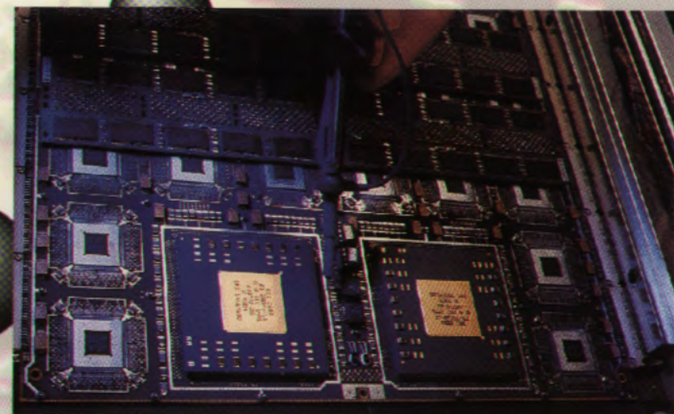
The programming environments also include optimized scientific and I/O libraries to aid in the development of high-performance applications.



MPP Apprentice is an advanced tool for parallel performance analysis. Detailed information is provided at the program, subroutine, loop, and statement level. This screen shows the program efficiency window indicating parallel processing overhead.

A D V A N C E D   T E C H N O L O G I E S



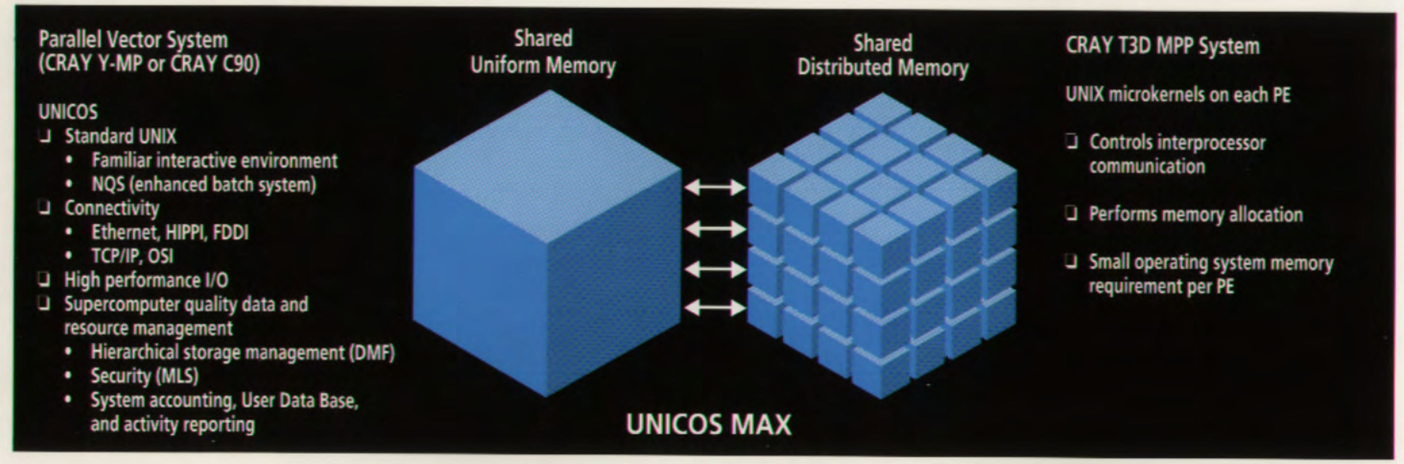


Top left, DEC Alpha microprocessor. This industry-leading RISC chip, capable of 150 MFLOPS peak performance, is at the heart of every PE.

Top right, supercomputer-class communication circuitry. Surrounding every DEC Alpha chip is an array of Cray Research-designed support logic, which maximizes system interconnect bandwidth.

Bottom left, CRAY T3D wire mat. The system's highly ordered interconnect technology minimizes latency within the 3-D torus architecture. Inset, modular connector technology maximizes system scalability.

Bottom right, robotic component placement system. To enhance system reliability, automated systems scan components and mount them on processor modules using innovative tab tape technology.



CRAY T3D—the right system for production environments

Previous generations of MPP systems have fallen far short of reaching their potential performance on real applications. Such MPP systems and their peripherals reveal their workstation heritage by providing loosely coupled, multicomputer environments. In contrast, the CRAY T3D system, coupled with the full power of Cray Research's UNICOS operating system, provides users with a supercomputer-class, production-oriented, multiprocessor environment.

UNICOS MAX

Sharing the CRAY T3D system among a large user community is the responsibility of UNICOS MAX, Cray Research's distributed MPP operating system. UNICOS MAX is a complete, multiuser operating system based on emerging and de facto industry standards. This powerful supercomputing environment enhances application development, system interoperability, and user productivity.

UNICOS MAX is a distributed operating system—functions are divided between the CRAY T3D system's PEs and the UNICOS operating system on the CRAY Y-MP or CRAY C90 host. By minimizing the amount of operating system software running on a PE, software overhead is reduced, releasing most of the PE local memory for application data storage and processing.

The CRAY T3D system can be shared among large numbers of users with applications program partitions ranging in size from two PEs to the whole system. An extensive set of system administration capabilities for managing system resources is provided. Interoperability with multiple computers on a network is assured through conformance to the following established and de facto industry standards:

- ❑ Operating systems (UNIX System V, BSD UNIX, POSIX 1003.1)
- ❑ Languages (Fortran 77, C)
- ❑ Networks (HIPPI, FDDI, Ethernet)
- ❑ Protocols (RPC, OSI, TCP/IP)
- ❑ Distributed tools (PVM, RQS/NQS, NQE)

Scalable heterogeneous supercomputing
The CRAY T3D system combines the best of both worlds—the proven production capabilities of parallel vector and the power of MPP. This scalable heterogeneous system enables customers to choose from a variety of system solutions and allows them to enhance their MPP or parallel vector capabilities incrementally as required.

This architecture offers production quality I/O, proven software tools, high-speed networking, and the ability to distribute a single application or mix of jobs between a parallel vector environment and a closely coupled, highly parallel environment. With this total computational capability, the CRAY T3D and CRAY Y-MP or CRAY C90 systems provide the right tool at the right time.

"With the heterogeneous architecture of the CRAY T3D front-ended by the Cray parallel vector system, the Cray MPP will give our users the robust, easy-to-use environment we've come to expect from Cray Research."

Larry Eversole, Deputy Manager, Supercomputing, Jet Propulsion Laboratory, Pasadena, California

System Configurations

Multi-cabinet CRAY T3D Model	Processing Elements (PEs)	Total Memory (Gbytes)	Peak Performance (GFLOPS)	Cooling
MCA32	32	0.5 or 2	4.8	Air or water
MCA64	64	1 or 4	9.6	Air or water
MCA128	128	2 or 8	19.2	Air or water
MC128	128	2 or 8	19.2	Water
MC256	256	4 or 16	38.4	Water
MC512	512	8 or 32	76.8	Water
MC1024	1024	16 or 64	153.6	Water
MC2048	2048	32 or 128	307.2	Water

Single-cabinet CRAY T3D Model	Processing Elements (PEs)	CRAY T3D Memory (Gbytes)	CRAY T3D Peak Performance (GFLOPS)	Host CPUs	Host Memory (Mwords)	Cooling
SC128	128	2 or 8	19.2	1 - 4	64	Water
SC256	256	4 or 16	38.4	1 - 4	64	Water

Consistent macroarchitecture across future MPP generations
Future Cray Research MPP systems will share a consistent hardware and software high-level architecture (macroarchitecture) with the CRAY T3D system. As a result, customers will see additional return on their application development investments as programs written for the CRAY T3D system transfer easily to future systems.

Reliability
Reliability has been designed into the CRAY T3D system at all levels. At the software level, UNICOS MAX is an extension of Cray Research's mature, robust, and well-proven UNIX-based operating system, UNICOS. At the hardware level, supercomputer technology already proven in the CRAY Y-MP and CRAY C90 series systems and redundant PEs and power supplies within the CRAY T3D system provide high overall system reliability.

A wide variety of configurations
CRAY T3D system configurations are available to meet a wide range of customer needs. Existing CRAY Y-MP and CRAY C90 system customers can enhance their computer facilities by integrating a CRAY T3D multi-cabinet (MC) or multi-cabinet, air-cooled (MCA) configuration with their existing Cray Research systems. For a stand-alone solution, the CRAY T3D SC systems integrate up to 256 PEs with up to four CRAY Y-MP host CPUs in a single cabinet.

"Using the emulator, simulator, and compiler, we've gotten five of our major chemistry and biomedical codes readied for the CRAY T3D. Our goal is to move our broad base of applications, including CFD, chemistry, and structures, to the CRAY T3D/C90 platform."

Michael Levine and Ralph Roskies, Scientific Directors of the Pittsburgh Supercomputing Center (PSC)

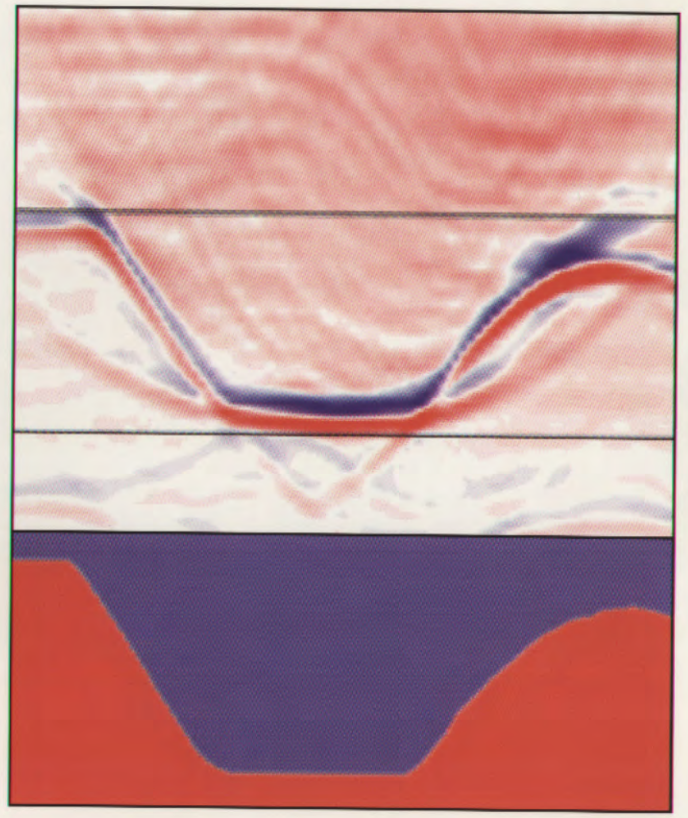
Application initiatives for the CRAY T3D system

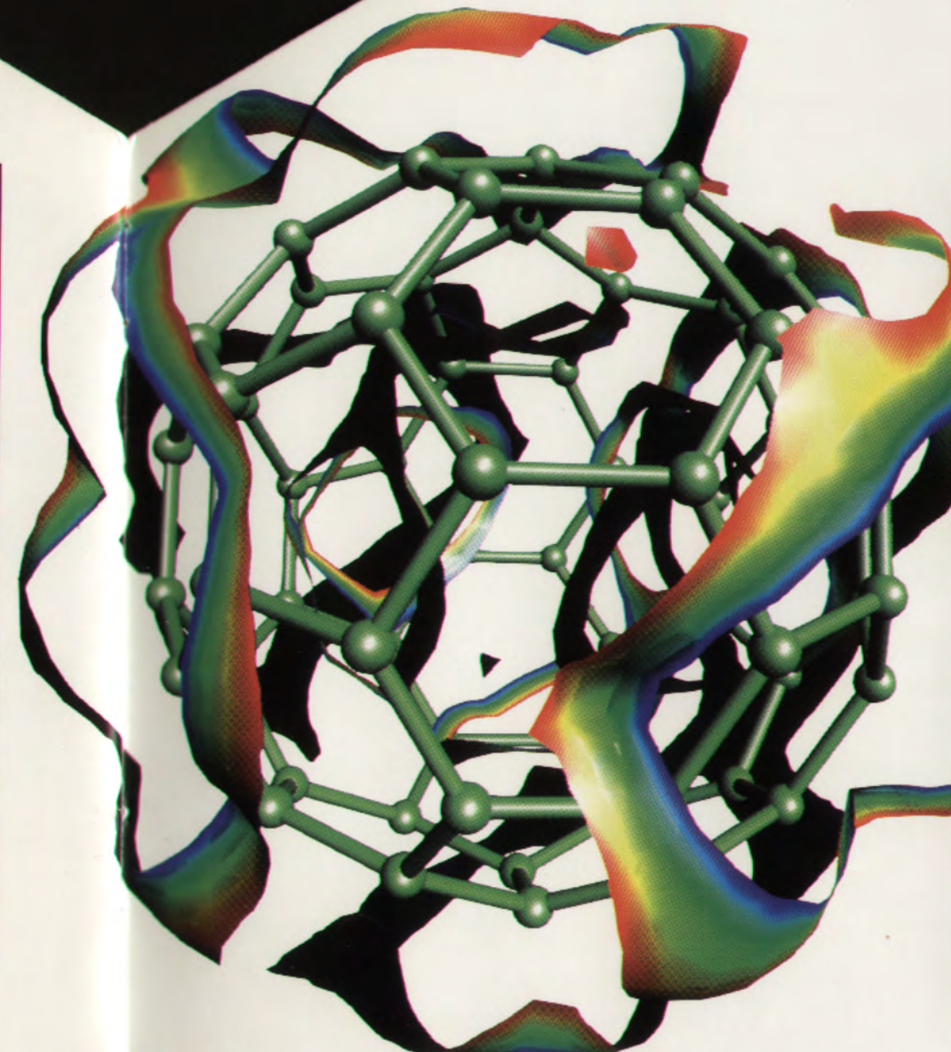
Today, few commercially available applications run on existing MPP systems. Cray Research and the CRAY T3D system will change that. In partnership with customers and application vendors, Cray Research is developing a wide selection of application software for CRAY T3D systems. Our first initiatives focus on the following key areas:

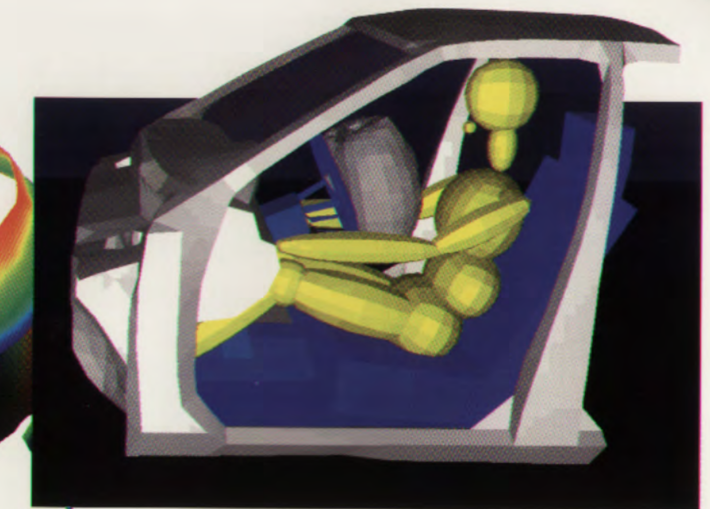
- 3-D prestack seismic data processing for petroleum exploration
- Atmospheric modeling for weather prediction and climate research
- Computational fluid dynamics and structural analysis for the aerospace, automotive, chemical, and semiconductor industries
- Computational chemistry for drug design and materials science applications
- Computational electromagnetics
- Combustion modeling for engine design

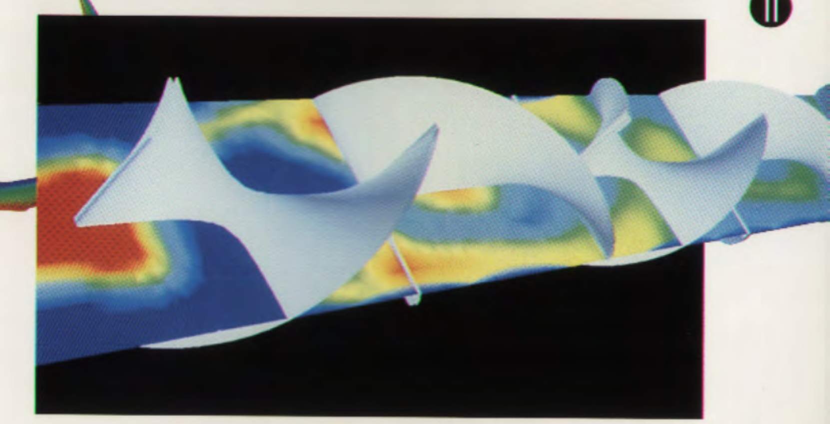
To develop applications technology for massively parallel processing, Cray Research has established the Parallel Applications Technology Program in partnership with a worldwide network of leading research institutions and industrial organizations. Members currently include the Pittsburgh Supercomputing Center and the California Institute of Technology. Others will be joining this partnership shortly.

Cray Research, the participating centers, and many third-party vendors are working together to leverage MPP technology to extract the greatest potential from the CRAY T3D system on existing widely used codes, as well as new interdisciplinary applications.



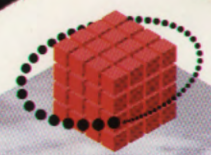






Cray Research's applications-development partnerships concentrate on several areas that show promise for significant performance gains with massive parallelism. These areas include prestack seismic data processing for petroleum exploration (far left); computational chemistry for the analysis of new pharmaceuticals and materials, such as the Buckminsterfullerene (C_{60}) molecule (above left); structural analysis for industrial applications, such as vehicle crashworthiness studies (top right); and computational fluid dynamics (CFD) for industrial applications, such as chemical mixing (above right).

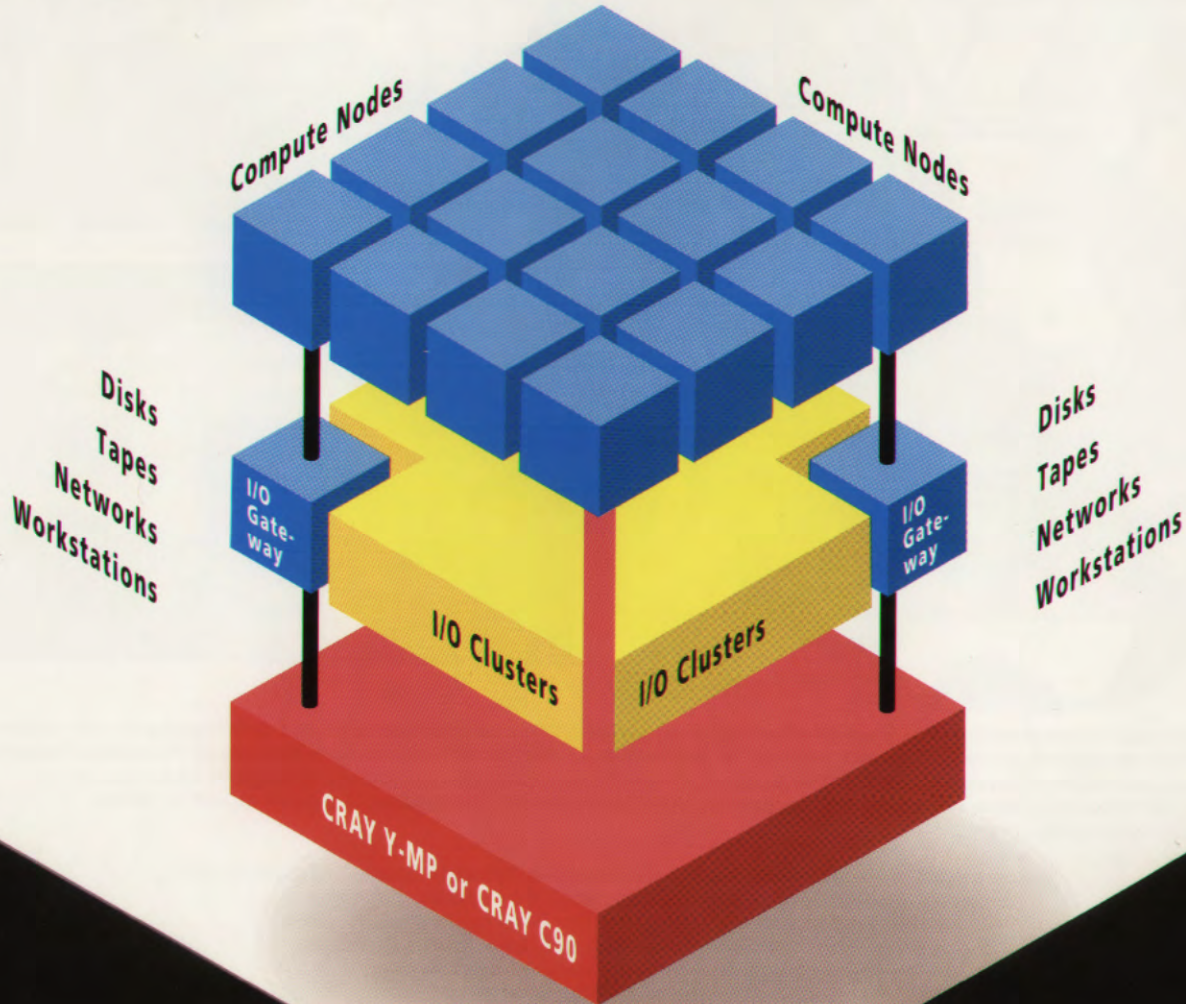
Seismic data and model provided by Jason Kao, Earth and Environmental Sciences Group, Cray Research. Image of Buckminsterfullerene molecule, with colored ribbon indicating highest occupied molecular orbital, provided by Cary Sandvig, UniChem Development Group, Cray Research. Image of flow through KM laminar flow static mixer courtesy Chemineer, Inc. Automobile crashworthiness model courtesy Livermore Software Technologies Corporation and Kia Motors Corporation.



> "Cray Research has mastered all enabling technologies required for the teraflops challenge: custom VLSI design, clock circuitry, advanced cooling, packaging, manufacturing, and software engineering."
>
> Michel Reymond, Director, Scientific Computing Center, Swiss Federal Institute of Technology (Ecole Polytechnique Fédérale, Lausanne/EPFL)

CRAY T3D—the right tool at the right time

The CRAY T3D system is the most usable, highest-performing MPP system available today that is geared for a production environment. With its efficient, shared distributed memory, high-speed system interconnect, and flexible programming model, the CRAY T3D system is the right tool for breaking the computational barriers of today and tomorrow.



Cray Research scalable heterogeneous architecture