CRAY-3
SUPERCOMPUTER SYSTEMS

CRAY COMPUTER CORPORATION

Corporate Headquarters and Main Facility.

Printed Circuit Board Facility.

Cray Computer Corporation was established as an independent company on November 15, 1989, with corporate headquarters in Colorado Springs, Colorado. The spin-off agreement with Cray Research, Inc. provided for the transfer of currently owned assets, people and initial funding. A transfer of patents and cross-licensing of technology allows each company to pursue its own specific projects unencumbered by patent or technology conflicts. Cray Computer Corporation is a public company listed on the NASDAQ Exchange.

The company's mission is to design, manufacture, sell and support high-performance, general purpose scientific computers. The company's first product is the CRAY-3 supercomputer system.

The main facility in Colorado Springs houses corporate management, hardware and software designers, a gallium arsenide integrated circuit fabrication facility, manufacturing, and system testing. A separate facility, also in Colorado Springs, manufactures the CRAY-3 printed circuit boards.

Introducing the CRAY-3 Supercomputer Systems

The CRAY-3 is the first supercomputer to use gallium arsenide (GaAs) integrated circuits for all of its logic circuitry. The development of GaAs digital circuits was a fundamental step in enabling the CRAY-3 to attain the fastest clock cycle time available in a computer system (two nanoseconds).

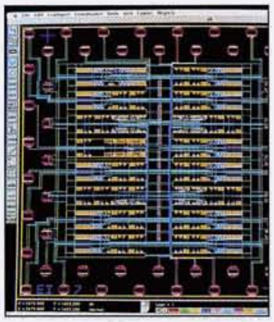
The CRAY-3 offers a balanced combination of high-speed vector processing, very fast scalar processing and the largest directly addressable memory available in a general purpose scientific computer (up to two gigawords). These features, combined with a highly parallel architecture, make the CRAY-3 the most powerful system available to the scientific and engineering communities.

The performance features of the CRAY-3 are a result of a unique synthesis of system architecture and hardware technology.

The CRAY-3 architecture is an evolutionary extension of the CRAY-2 architecture. The system cabinet illustrates the hardware technology required. All the logic and memory circuitry for the machine resides in the top eight inches of an octagon-shaped cabinet only 42 inches wide and 50 inches high. These top eight inches of the system cabinet contain one to 16 computational processors, one system management processor, up to two gigawords of common memory and up to 15 I/O modules.

The logic and memory circuitry are contained in three-dimensional modules only four inches square by one-quarter of an inch thick. Packaging the architecture in this small space allows for short signal paths throughout the system.

The combination of compact packaging and high-performance components was essential to the development of a balanced, powerful and high-speed system required by today's customers. It is unique to the CRAY-3.



CRAY-3 Design

The CRAY-3 achieves its high-performance processing capabilities with the use of GaAs logic circuitry, efficient packaging, liquid immersion cooling, multiple processors and very large common memory. These hardware technologies are then complemented and maximized by the elegant architecture and functional design of the CRAY-3.

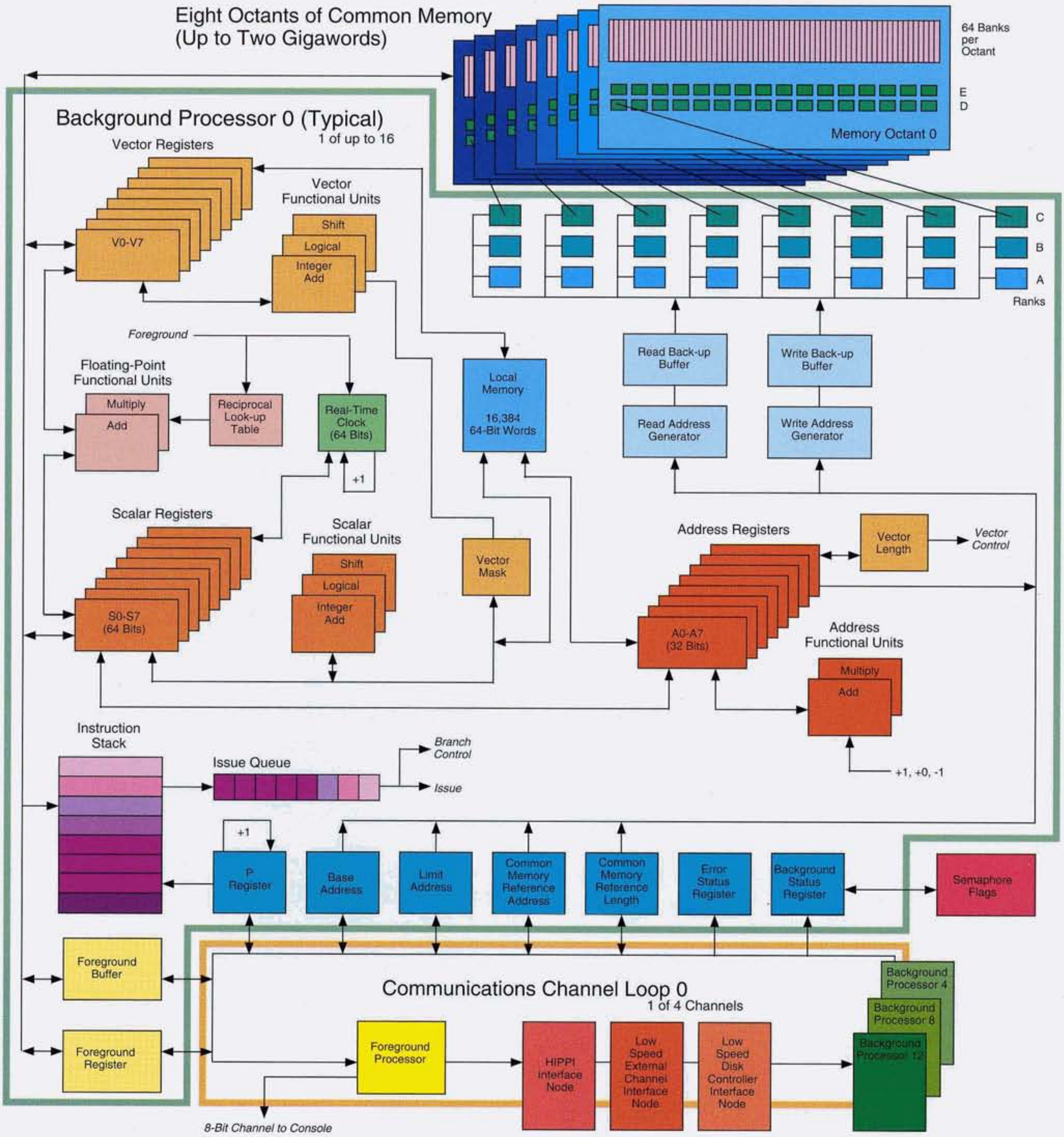
Background Processors

The parallelism in the CRAY-3 extends beyond the multiprocessor features of the system. Each of the background processors consists of three sections: the computation section, the control section and high-speed local memory. A broad mixture of scalar and vector arithmetic and logical operations can take place at the same time in the computation section. Instructions can issue every clock period. Computation instructions execute register-to-register to allow them to operate at the maximum rate possible. The control section supports the parallel operation of the multiple functional units in the computation section. The high-speed local memory is used to temporarily store scalar and vector data during computations. The peak performance of each CRAY-3 background processor is one gigaflop.

The computation section of the background processors contains registers and functional units associated with address, scalar and vector processing. Two integer arithmetic functional units are employed in address processing. Three functional units are dedicated solely to scalar processing, and two floating-point functional units are shared with vector operations. Two additional functional units are dedicated to vector operations allowing CRAY-3 systems to issue one result per clock period in vector mode.

Features of the Computation Section

- Twos complement integer and signed magnitude floating-point arithmetic
- Address and arithmetic registers
  - Eight 32-bit address (A) registers
  - Eight 64-bit scalar (S) registers
  - Eight 64-element vector (V) registers with 64 bits per element
- Address functional units
  - Add/subtract
  - Multiply
- Scalar functional units
  - Logical
  - Shift
  - Integer
    - Add/subtract
    - Population/parity
    - Leading zero count
- Vector functional units
  - Logical
  - Shift
  - Integer
    - Add/subtract
    - Population/parity
    - Leading zero count
    - Compressed iota
- Floating-point functional units
  - Add/subtract
  - Multiply/reciprocal/square root
- Scatter and gather vector operations to and from common memory



Eight Octants of Common Memory
(Up to Two Gigawords)
64 Banks per Octant
E
D
Memory Octant 0
C
B
A
Ranks
Background Processor 0 (Typical)
1 of up to 16
Vector Registers
V0-V7
Vector Functional Units
Shift
Logical
Integer Add
Foreground
Floating-Point Functional Units
Multiply
Add
Reciprocal Look-up Table
Real-Time Clock (64 Bits)
+1
Local Memory
16,384 64-Bit Words
Read Back-up Buffer
Write Back-up Buffer
Read Address Generator
Write Address Generator
Scalar Registers
S0-S7 (64 Bits)
Scalar Functional Units
Shift
Logical
Integer Add
Vector Mask
Address Registers
A0-A7 (32 Bits)
Vector Length
Vector Control
Address Functional Units
Multiply
Add
+1, +0, -1
Instruction Stack
Issue Queue
Branch Control
Issue
+1
P Register
Base Address
Limit Address
Common Memory Reference Address
Common Memory Reference Length
Error Status Register
Background Status Register
Semaphore Flags
Foreground Buffer
Foreground Register
Communications Channel Loop 0
1 of 4 Channels
Foreground Processor
HIPPI Interface Node
Low Speed External Channel Interface Node
Low Speed Disk Controller Interface Node
Background Processor 4
Background Processor 8
Background Processor 12
8-Bit Channel to Console

Each background processor employs an identical, independent control section of registers and instruction buffers for instruction issue and control. Each control section has memory, base and limit registers for program relocation and protection. The instruction buffers (eight segments with 16 words per segment) contain the instructions to be executed. A 64-bit real-time clock is synchronized with the foreground processor's 32-bit real-time clock at system start-up. Each clock is advanced by one count each clock period. Semaphore flags are used to support application of multiple processors to a single program. These flags are one-bit registers which provide interlocks for common access to shared memory fields. A background processor is assigned access to one semaphore flag by a field in the status register. The background processor has instructions to test, branch, set and clear a semaphore flag.

Features of the Control Section

- 136 basic instruction codes
- Eight instruction buffers
- 32-bit Program Address register
- 32-bit Base Address register
- 32-bit Limit Address register
- 32-bit Status register
- 64-bit Real-Time Clock
- Multiple semaphore flags which provide interlocks for multi-tasking

In addition to the registers and functional units, each background processor incorporates 16,384 64-bit words of high-speed local memory which is assigned by the compilers and used as fast scratch memory during computations. This design minimizes the number of common memory access calls required and improves overall performance of the system. Local memory accesses take four clock periods and can overlap accesses to common memory.

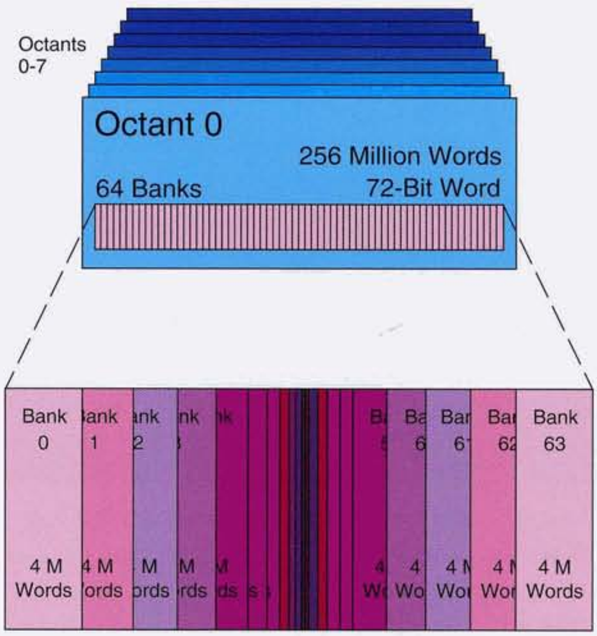
Features of the Local Memory Section

- 16,384 64-bit words
- Access time of four clock periods
- Used as fast scratch memory during computations
- Register accesses can overlap common memory accesses
- Provides temporary storage of vector segments
- Reduces the number of common memory access calls required, improving overall performance

Common Memory

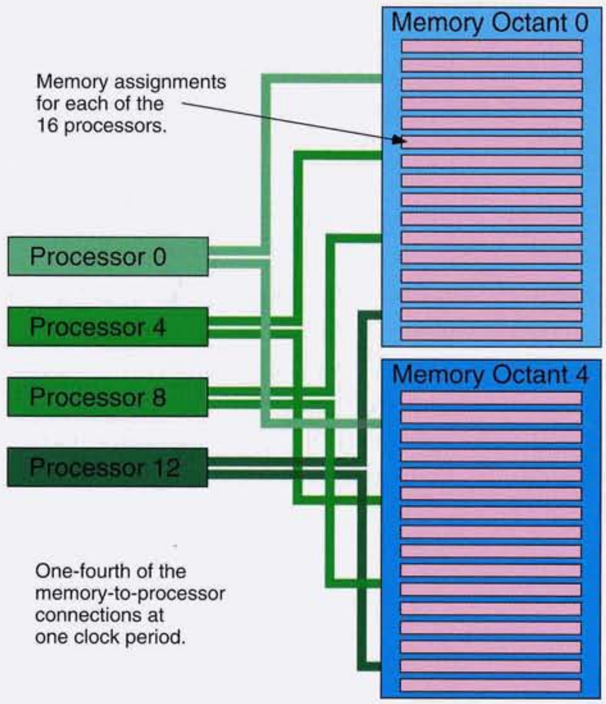
The CRAY-3 has the largest directly addressable high-speed memory available in a general purpose scientific computer (up to two gigawords). This vast memory resource is directly addressable by an application program. Such a large common memory allows the individual user to run programs that would be impractical to run on any other computer system.

Common memory is arranged in octants of 64 banks each, providing up to 512 interleaved banks for an eight-octant machine. Each word consists of 64 data bits and eight error correction bits. A memory bank can utilize 16 background processor memory access ports. High-speed, silicon SRAM CMOS chips are used in all versions of the CRAY-3. Total memory bandwidth is 128 gigabytes per second with a peak burst transfer rate of one gigaword per second per processor using two ports to memory—a total peak burst rate of 16 gigawords per second.



Features of Common Memory

- Up to two gigawords available
- 72-bit words (64 data bits, eight correction bits)
- Up to 512 memory banks
- Each memory bank can utilize up to 16 background processor memory access ports
- Bi-directional memory access ports
- Bandwidth of 128 gigabytes per second
- High-speed SRAM CMOS memory technology used on all machines



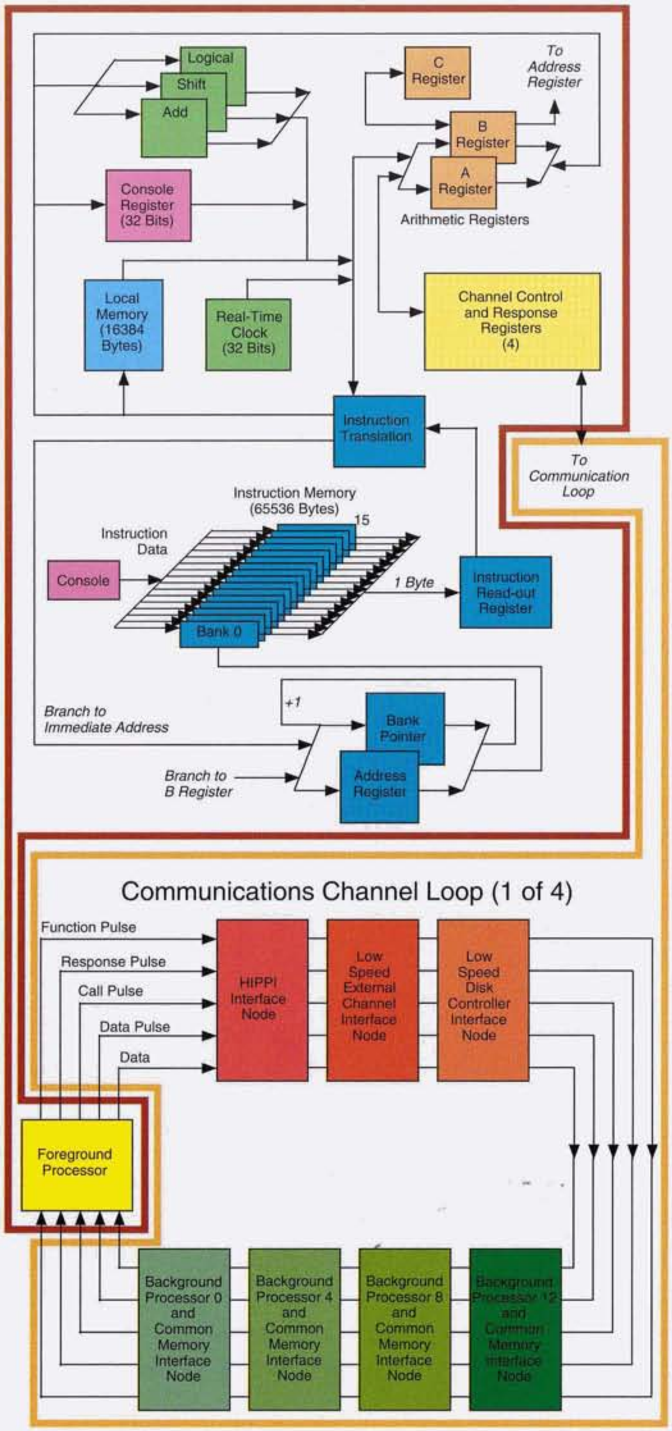
Foreground System

The foreground processing system provides input, output and overall system management. The foreground system includes a 32-bit, two-nanosecond CPU, with its own registers and memory. It operates in parallel with the background processors.

The foreground system monitors the system components via the foreground communication channels. It handles all I/O interrupts and transfers, freeing the background processors to asynchronously perform the computations associated with the user applications.

System communication occurs through four high-speed synchronous data channels (one gigabyte per second each). These channels interconnect the background processors, foreground processors, disk control units and host interfaces. Each foreground communication channel connects to four background processors and one group of I/O controllers.

The majority of foreground processor activity involves data transfer between common memory and external devices. The system provides a mixture of 12 megabytes per second low-speed interfaces and 100 megabytes per second High Performance Parallel Interfaces (HIPPI) to accommodate the data transfer needs of the user.

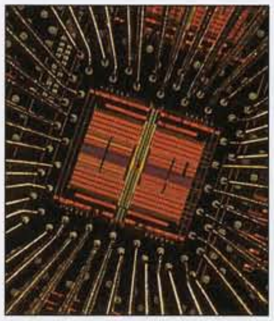


External I/O Interfaces

A fully-configured CRAY-3 system can have up to 15 interface modules. Control circuitry in the interface modules provides for the use of low-speed devices (six megabytes per second), high-speed devices (12 megabytes per second) and HIPPI channels (100 megabytes per second). One interface module can provide all three types of data transfer modes, or single modules can be dedicated to several of one kind. A single HIPPI interface module can provide dynamic switching (software controlled) between four 32-bit channel pairs, or two 64-bit channel pairs, or two 32-bit and one 64-bit channel pairs.

Disk Subsystems

The CRAY-3 system supports Redundant Arrays of Inexpensive Disks (RAID) for high volume and very fast sequential transfer rates. RAID units are connected via high-speed, IEEE standard, HIPPI channels with 32-bit and 64-bit HIPPI controllers for both source and destination. The HIPPI channels use a 40-nanosecond clock for a burst bandwidth of 100 megabytes per second.



CRAY-3 Technology

The technological innovations in the CRAY-3 include gallium arsenide (GaAs) digital logic; 69-layer, three-dimensional modules; and direct contact liquid immersion cooling using a clear, odorless, inert fluorocarbon.

GaAs Logic

The CRAY-3 is the first supercomputer to use gallium arsenide integrated circuits for all of its logic circuitry. The use of GaAs die was a key factor in enabling the CRAY-3 to achieve the fastest clock cycle time of any computer system currently in existence.

Component Packaging

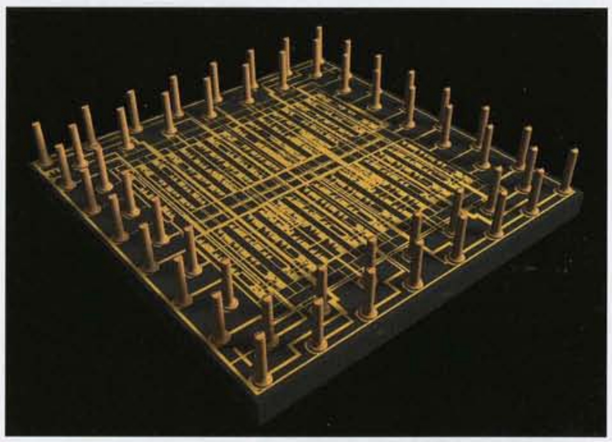
The laws of physics demand that very high-speed electronic circuits must have short path lengths. With a clock speed of 500 Mhz, the CRAY-3 required greater creativity and efficiency of packaging than had ever before been attempted.

The CRAY-3 logic and memory circuitry is packaged in up to 336 removable modules, each containing up to 1,024 GaAs integrated circuit die. Total integrated circuit population in a 16-processor CRAY-3 is over 142,000 die, of which 36,864 are for common memory. This packaging results in a GaAs gate density of approximately 96,000 gates per cubic inch.

The modules are three-dimensional structures measuring 121 mm by 107 mm by 7 mm (about four inches square by one-quarter of an inch thick). Nine printed circuit boards make up the module sandwich and contain a total of 69 electrical layers. Circuit connections are made in all three dimensions within the module. X-y traces are as small as 0.048 mm (a human hair averages 0.070 mm). Z-axis connections are made with approximately 14,000 gold-plated, beryllium-copper twist-pin jumpers per module. The logic signal jumpers, which make up the bulk of the z-axis connections, are only 0.122 mm in diameter.

Cooling

The CRAY-3 is cooled by direct contact immersion in an inert liquid fluorocarbon, technology similar to that employed in the CRAY-2. However, the high-power density of the CRAY-3 (up to 640 watts per cubic inch)

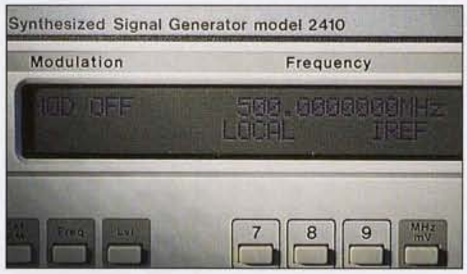


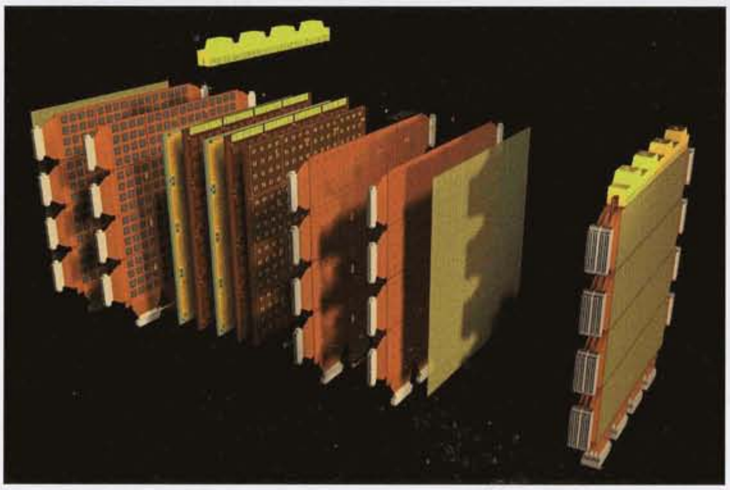
GaAs die used in the CRAY-3 are only 3.835 mm square. They are mounted unpackaged to the printed circuit boards using 0.076 mm gold leads ultrasonically welded to 52 die bonding pads.

required further development of this cooling mechanism to provide adequate heat removal. A key factor in the design of the cooling mechanism for the CRAY-3 was the narrow channels between the module layers through which the cooling fluid must flow. The channels are a maximum of 300 microns from the back surface of a die to the adjacent printed circuit board.

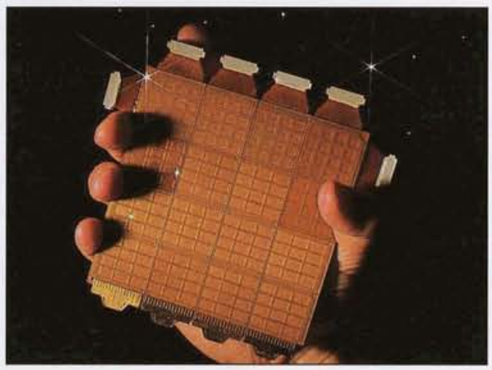
The system tank is carefully engineered to ensure that all of the coolant flows between the module layers, rather than between modules, where it comes in direct contact with the die and gold jumpers for maximum heat transfer. Operating temperatures throughout the computer circuits average 30° Celsius with a total temperature rise of the coolant across the cooling loop of only 5° Celsius. This allows for low thermal shock and hence long-term reliability of the modules.

The CRAY-3 modules are continually bathed in a precisely controlled flow of a clear, inert fluorocarbon. The presence of this liquid can only be visually detected by an occasional bubble, or when the tank is being filled.





The CRAY-3 modules are a multi-layer sandwich of printed circuit boards containing 69 electrical layers and four layers of GaAs die in a vertical space of only one-quarter of an inch.



CRAY-3 System Configurations

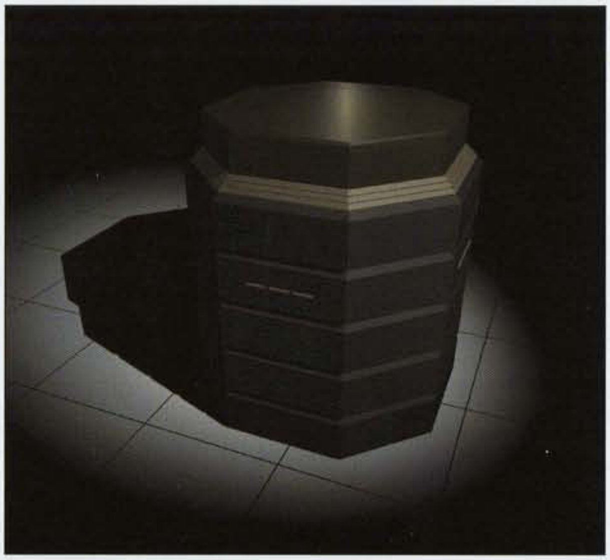
The modular design of the CRAY-3 cabinet allows considerable flexibility in configuring systems to customers' needs. This design also allows CRAY-3 systems to be upgraded in the field. The accompanying table lists some of the available configurations.

Actual configurations of CPUs and memory can be specified by the customer. This allows customers to tailor the machine to their budget and problem-solving requirements.

I/O configurations will also be determined by the needs of each specific customer. Each I/O module can accommodate multiple channels and speeds. HIPPI interfaces are available for network and RAID disk connections. In addition, for customers with installed Cray Research, Inc. systems interfaces are available to support the following network connections and disk systems: Computer Network Technology equipment, Network Systems Corporation equipment, and Cray Research, Inc. DD-49 and DD-40 disk drives.

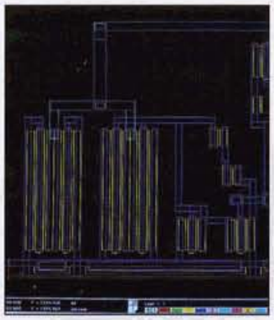
	Number of CPUs	Megawords of Memory	Available I/O Modules
CRAY-3/1-256	1	256	1
CRAY-3/2-256	2	256	1
CRAY-3/4-512	4	512	3
CRAY-3/4-1024	4	1024	3
CRAY-3/4-2048	4	2048	3
CRAY-3/8-1024	8	1024	7
CRAY-3/8-2048	8	2048	7
CRAY-3/16-2048	16	2048	15

CRAY-3 Four-Processor system cabinet.



CRAY-3 Eight-Processor system cabinet.

Our mission is to design, manufacture, sell and support high-performance, general purpose scientific computers.



CRAY-3 Applications Environment

Top-end supercomputers are used to solve scientific problems that are not computationally tractable on less powerful machines. But to do this the machines and their users need the applications software and support which can effectively utilize the inherent power and performance of the machine.

Application Environment

Scientists developing codes in such diverse areas as energy research, image processing, seismic modeling, computational chemistry and structural analysis require tools that enable them to apply all the computing power of the CRAY-3 to their particular problems. Cray Computer Corporation is committed to providing the application environment they need.

Users connect to the CRAY-3 by using standard protocols. Familiar UNIX editors are used to create or modify their programs. Powerful compiler and multiprocessing tools are used to prepare executable programs that extract all the power of their CRAY-3. Finally, program output can be displayed using many connectivity options. When debugging is required a rich visual debugger speeds the task.

Applications

User Fortran and C applications that meet existing ANSI standards will easily convert to the CRAY-3. In most cases users will only need to move the source codes to the CRAY-3, recompile and execute code.

The large memory of the CRAY-3 (up to two gigawords) gives users the opportunity to attack demanding problems much more efficiently than on a machine with a small memory.

Scaling up a prototype solution to provide a production program need not involve elaborate programming to do work-arounds because of limited memory. With the large memory of the CRAY-3 the user will most likely need to simply increase the parameters or dimensions of his code, proceed directly to debugging and run the production program. The production program remains flexible and is easily modified since there is no complex I/O scheme obscuring the simplicity of the original prototype program. Similarly, production programs and third-party codes can be easily ported without massive rewriting, speeding the entire process and increasing productivity.

CRAY-3 software gives users the tools needed to prepare and run an application. Cray Computer Corporation provides the performance tools to enable aggressive users to capture the top performance potential in their applications. Flow-tracing tools locate the most time-consuming areas of a code. These tools also help the user analyze the multiprocessing performance and data flow.

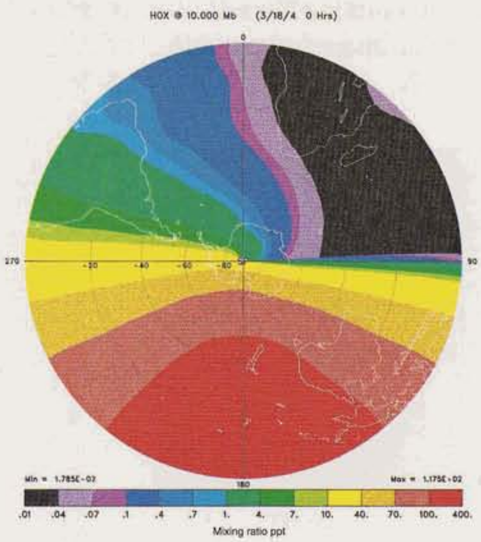
Accounting tools are also available to profile the system resources required by an application. With these tools the user can locate the high-leverage opportunities for optimization, identify the kind of change required and then go on to increase vectorization, parallelism or I/O effectiveness. With a more effectively optimized code the user gets improved turnaround—the scientific job gets done more quickly.

Third-Party Applications

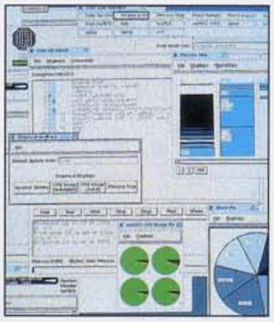
CRAY-3 software, with the standard UNIX environment and language support, provides an atmosphere that facilitates moving third-party applications from other supercomputer systems, particularly from the CRAY-2. The CRAY-3 is a natural extension of the CRAY-2 architecture, building upon the library of instructions previously available.

A large repertoire of highly portable third-party applications is available for supercomputers and includes codes for reactor safety, computational chemistry, structural analysis, computational fluid dynamics and much more. Third-party investments that made their codes available on the CRAY-2 will provide a simplified path for bringing codes to the CRAY-3. As the CRAY-3 product matures, Cray Computer Corporation will further respond to customer application needs by working with the vendors of third-party codes.

A CRAY-3 Four-Processor system solving large-scale climatology models at the National Center for Atmospheric Research.



The mixing ratio in parts per trillion at 10 millibars (mb) of the radicals OH and HO_2 at 0000 UTC over the southern hemisphere, as solved by a global tropospheric chemical tracer model on the CRAY-3 at the National Center for Atmospheric Research.



CRAY-3 System Software

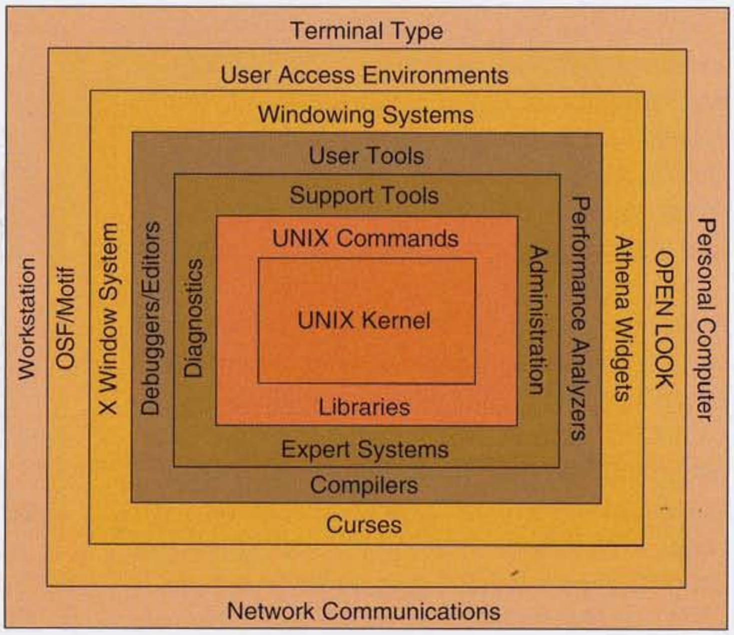
The operating system for the CRAY-3 is based on UNIX. The familiar UNIX environment gives users a head start as they attack their specific problems. The CRAY-3 provides an enhanced environment to support the diverse demands of top-end users—from high-performance I/O, when the user must process a great deal of data efficiently, to a standard batch environment to meet the needs of production codes; from job recovery to ensure the completion of long running codes, to multiprocessing tools to give the user access to all the power of the CRAY-3.

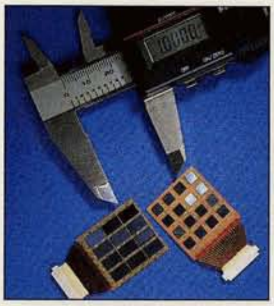
The system software package for the CRAY-3 also includes effective system administration and tunable scheduling tools.

The Fortran compiler complies with the ANSI 77 standards and has extensions that include compatibility with Cray Research, Inc. extensions, as well as parts of the Fortran 90 vector syntax. The compiling system automatically detects opportunities for parallel execution and generates code to take advantage of them.

The C compiler complies with the ANSI standard for the C language. It has extensions for vectorization and to allow use of multiple processors on a single application.

Network support is provided via TCP/IP, NFS and other standard protocols.





CRAY-3 Physical Specifications

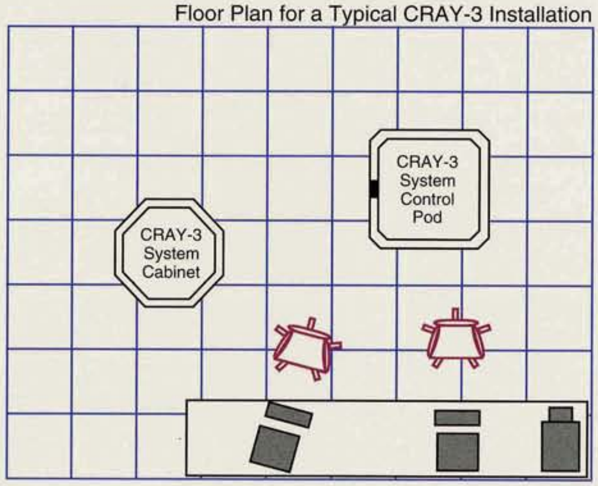
The CRAY-3 system cabinet is relatively small considering the amount of computing power packed inside. The cabinet for one-, two- and four-processor machines is 106.68 cm (42 inches) wide and 71.12 cm (28 inches) deep. The cabinet for eight- and 16-processor machines is a 106.68 cm (42 inch) octagon. All system cabinets extend 127 cm (50 inches) above the computer room raised floor. The cabinets are elegant in appearance with charcoal gray, matte-textured, gold-trimmed skins, and bronzed acrylic, see-through tops covering the modules.

The system control pod is 133.35 cm (52.5 inches) square and extends 140.46 cm (55.3 inches) above the computer room raised floor. Frequently used controls are hidden behind a bronzed acrylic panel. The entire top lid of the C-Pod can be raised electrically, giving access to further controls as well as the C-Pod electronics and system cooling components. All electrical and cooling system connections between the C-Pod and the system cabinet are hidden beneath the computer room raised floor. The C-Pod must be within eight to 15 feet of the system cabinet.

Basic Installation Specifications

- 1P, 2P and 4P system cabinets use 8.17 square feet of floor space and weigh about 2200 pounds in operation.
- 8P and 16P system cabinets use 12.25 square feet of floor space. The 8P weighs about 3900 pounds in operation. The 16P weighs about 7700 pounds in operation.
- The system control pod (two required for a 16P system, one for all other systems) uses 19.14 square feet of floor space and weighs about 3800 pounds in operation.
- Minimum footprint for a system cabinet and C-Pod installed in a computer room is 252 square feet. This includes space for interconnection cabling and piping.
- 45 KW of power is required per octant, 360 KW for a fully-configured system.
- One 100-KVA motor generator per 4P is required to supply the 3-phase wye, 400-Hz power to the system.
- MGs use 27 square feet of floor space and weigh about 7000 pounds.
- The liquid immersion cooling requires chilled water at 55° F for cooling the Fluorinert. The 1P and 2P systems require 50 gallons per minute. The 4P system requires 75 gpm, the 8P requires 135 gpm and the 16P requires 270 gpm.

Floor Plan for a Typical CRAY-3 Installation



CRAY-3 Support and Maintenance

Few things we buy today are worth more than the support and service received after the sale. This becomes even more important with greater initial investment and product sophistication. Cray Computer Corporation is committed to offering the support our customers would expect when purchasing one of the most powerful supercomputer systems available today.

Hardware Support

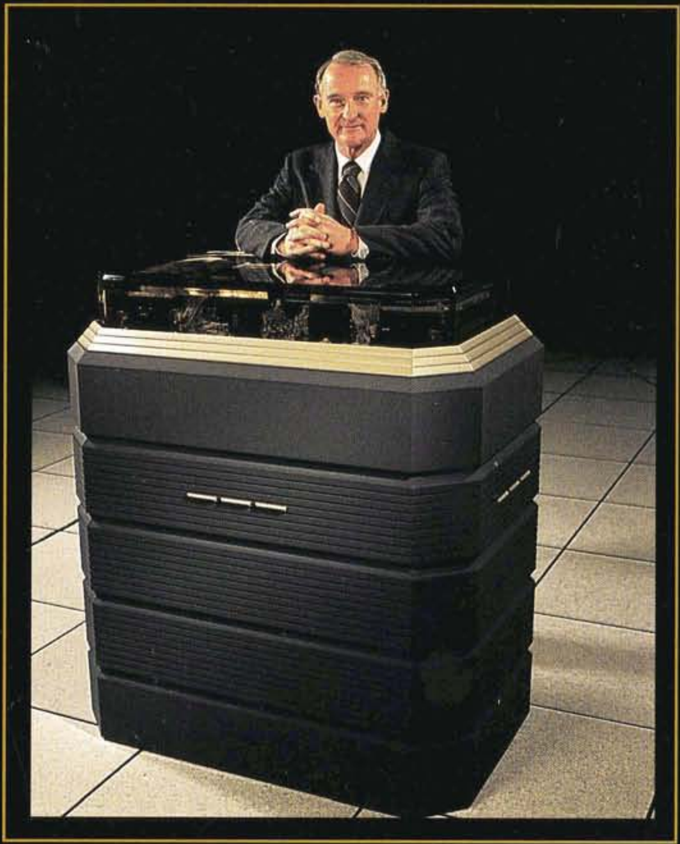
Our support begins with comprehensive pre-installation site planning. After installation a field support team takes the responsibility for maintaining proper operation of the machine. This support team will ensure correct operation through the use of on-line and off-line diagnostic tests. Preventive maintenance procedures are followed judiciously to maintain a high level of system performance. Any repairs that may become necessary are expedited by using a set of spare modules located right at the site, keeping downtime to a minimum.

Software Support

Cray Computer Corporation also provides on-site software support to assist customers in obtaining the maximum utilization of their CRAY-3. This software support team assists the customer in installing, debugging and tuning Cray Computer Corporation software products. They will also consult with the customer's application programmers in debugging, porting and optimizing customer applications. An on-line software problem data base is also maintained to help in the timely resolution of software problems.

Continuing Engineering

Another important aspect of support for a sophisticated machine like the CRAY-3 is that of continuing engineering. The continuing engineering group at Cray Computer Corporation provides the necessary technical engineering support for customizing each CRAY-3 to the specific requirements of individual customers. This includes ensuring that the machine will properly interface with the customer's workstation consoles, networks and data storage systems. The continuing engineering group also provides customers with product improvements and design enhancements over the life of the machine.



Seymour R. Cray has been the principal designer and developer of the CRAY-1, CRAY-2 and CRAY-3 supercomputer systems. Currently, Mr. Cray serves as the Chairman and Chief Executive Officer of Cray Computer Corporation and as a principal design consultant to the company.

Mr. Cray has been the leading architect of large scientific computers for more than 36 years. He is deservedly called "the father of supercomputing." Mr. Cray began putting computers together in 1951 for Remington Rand and Engineering Research Associates. His work continued at the Univac Division of Sperry Rand Corporation with the Univac 1100 series. In 1957 Mr. Cray and Bill Norris started Control Data Corporation. At CDC Mr. Cray was the principal architect in the design and development of the 1604, 6600 and 7600 computer systems. In 1972 Mr. Cray founded Cray Research, Inc. where he designed and developed the world's first supercomputer—the CRAY-1. Cray Computer Corporation was founded in 1989 as a spin-off from Cray Research, Inc. Mr. Cray is now busy with his next project—the CRAY-4.

CRAY COMPUTER CORPORATION

1110 Bayfield Drive
Colorado Springs, CO 80906

Mailing Address:

Post Office Box 17500
Colorado Springs, CO 80935

Phone: 719-579-6464
Fax: 719-540-4028

Editor and Design: Dean Tollefson, Artwork Department.
Photography: Steve Gombosi, Software Department.
Digital Color Renderings: Doug Cowles and Tom Sibert, Mechanical Engineering Department.
Additional Photography: Company headquarters building, inside front cover, Bob Hoeglund.
PCB facility, inside front cover, Don Jones.
Color Computer Graph on page 13: Courtesy of Mr. Stacey Walters, Atmospheric Chemistry Division, NCAR. Used by permission of the National Center for Atmospheric Research, University Corporation for Atmospheric Research under sponsorship of the National Science Foundation.
Lithography: ColorTek Printing, Colorado Springs, Colorado.

Printed August 1993.