

Cray X-MP
Cray T3D

The CRAY X-MP
Series of Computer Systems



CRAY

For close to a decade, Cray Research has been the industry leader in large-scale computer systems. Today, about 70 percent of all supercomputers installed worldwide are Cray systems. They are used in advanced scientific and research laboratories around the world and have gained strong acceptance in diverse industrial environments. No other manufacturer has Cray Research's breadth of success and experience in supercomputer development.

The company's initial product, the CRAY-1 Computer System, was first installed in 1976 and quickly became the standard for large-scale scientific computers — and the first commercially successful vector processor. For some time previously, the potential advantages of vector processing had been understood, but effective practical implementation had eluded computer architects. The CRAY-1 broke that barrier, and today vectorization techniques are used commonly by scientists and engineers in a wide variety of disciplines.

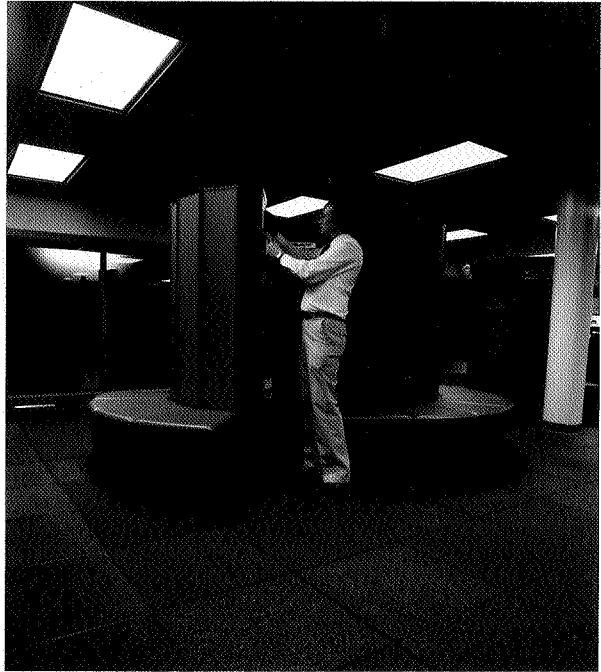
The field-proven CRAY X-MP Computer Systems now offer significantly more power to solve new and bigger problems while providing better value than any other systems available. Large memory size options allow a wider range of problems to be solved, while innovative multiprocessor design provides practical opportunities to exploit multitasking, the next dimension of parallel processing beyond vectorization. Once again, Cray Research, has moved supercomputing forward, offering new levels of hardware performance and software techniques to serve the needs of scientists and engineers today and in the future.

Introducing the CRAY X-MP Series of Computer Systems

Announcing expanded capabilities to serve the needs of a broadening marketplace: the CRAY X-MP Series of Computer Systems. The CRAY X-MP Series now comprises nine models, ranging from a uniprocessor version with one million words of central memory to a top-end system with four processors and a 16-million-word memory. Today's CRAY X-MP line is a field-proven technology; Cray Research introduced the dual-processor CRAY X-MP in 1982 and expanded the series to include one- and four-processor models in 1984.

The flexible CRAY X-MP multiprocessor configurations allow users to employ multiprogramming, multiprocessing and multitasking techniques. The multiple-processor architecture can be used to process many different jobs simultaneously for greater system throughput, or it can apply two or more processors to a single job for better program turnaround time. The combination of multiprocessing and vector processing provides a geometric increase in computational performance over conventional scalar processing techniques.

The CRAY X-MP system design is carefully balanced to deliver optimum overall performance. Fast long and short vector processing is balanced with high-speed scalar processing, and both are supported by powerful input/output capabilities. Cray Research software has been developed to ensure easy access to these performance features. The result is that users can realize maximum throughput for a variety of job mixes and programming environments.



A diversity of applications

The spectrum of applications for CRAY X-MP Computer Systems ranges from the subatomic to the celestial. Whether calculating charge densities of atoms or the aerodynamics of spacecraft, CRAY X-MP Computer Systems offer

new opportunities for research and discovery.

The CRAY X-MP's many configuration options give users the freedom to tailor systems to meet specific needs. In a business, university or government laboratory, in basic or applied research, Cray systems can be adapted to meet the most varied and demanding

computational requirements. As the marketplace for supercomputers has grown in size and diversity, Cray Research has provided new supercomputing performance capabilities.

Applications for Cray systems are well established in numerous high technology fields. The ability to run realistic simula-

tions of complex phenomena and to process enormous amounts of data quickly have made CRAY X-MP systems the standard for securing the most accurate, detailed, enlightening and profitable results. The following pages illustrate real-life applications for which the CRAY X-MP has proven invaluable.

CRAY

Each X-MP CPU offers gather/scatter and compressed index vector instructions. These instructions allow for the vectorized processing of randomly organized data, which previously was performed by scalar processing.

Complementing the power of the X-MP Series is a new generation of I/O technology. Cray's DD-39 and DD-49 disk drives offer 1200-Megabyte (Mbyte) capacity and very fast sustained transfer rates (9.8 Mbyte/sec for a DD-49, 5.9 Mbyte/sec for a DD-39). In addition, Cray's Solid-state Storage Device (SSD) provides up to 1024 Mbytes of very fast random-access secondary MOS memory. When connected to a four-processor CRAY X-MP through two 1000-Mbyte/sec channels, it provides a maximum aggregate transfer rate of 2000 Mbyte/sec.

A wide variety of applications programs for solving problems in industries such as petroleum, aerospace, automotive, nuclear research and chemistry are available for operation on CRAY X-MP computers. Thus, scientists and engineers can use X-MP systems and industry standard

codes to solve a wide range of problems. Additionally, software developed for the CRAY-1 can be run on all models of the CRAY X-MP Series, thus protecting user software investment.

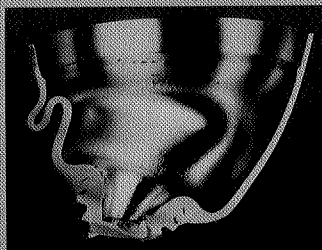
From both a hardware and software standpoint, the CRAY X-MP can be integrated easily into a user's existing computer environment. Hardware and software front-end interfaces for other manufacturers' equipment are available. And the CRAY X-MP requires a minimum of floor space, occupying just 112 square feet (11 square meters) in its maximum configuration, including the Solid-state Storage Device.

Cray computers offer the most powerful and cost-effective computing solutions available today for advanced scientific applications — both for experienced supercomputer users with the most demanding computing requirements and for newer users whose research needs now require supercomputer power. The CRAY X-MP features one or more powerful CPUs, a very large central memory, exceptionally fast computing speeds and I/O throughput to match. As the supercomputer marketplace broadens, the CRAY X-MP Series of Computer Systems will evolve to meet users' expanding computing requirements.

Structural analysis

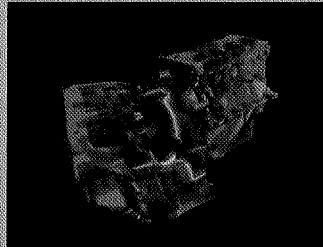
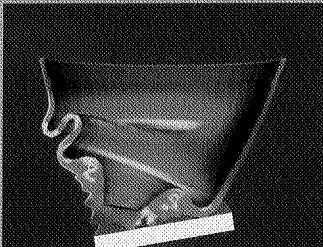
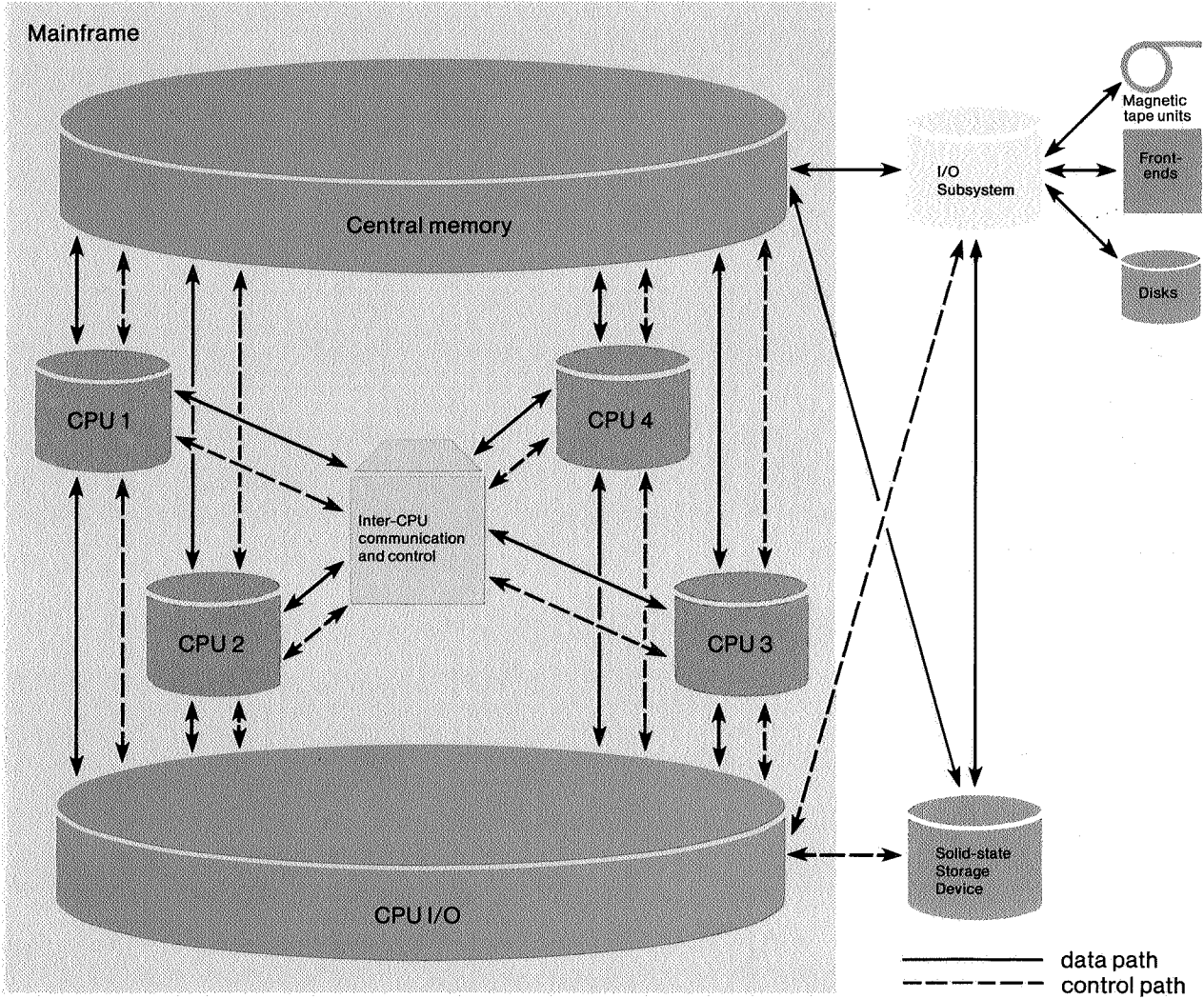
Finite element analysis is a mathematical method for calculating the effects of temperature- and pressure-related stress on physical structures. The aerospace, automotive and civil engineering industries rely on the method to conduct engineering design and analysis. Using

CRAY X-MP systems, scientists can rapidly evaluate structures too large or complex to be analyzed adequately any other way. The result is improved engineering efficiency and more structurally sound and lightweight components.



Experimental and computational results of the impact of a nose cone against an angled solid surface. (Left) Actual results and (right) the numerical simulation of this highly nonlinear process. The striking similarity between the two confirms the validity of the computational approach. (Credit: Lawrence Livermore and Sandia National Laboratories.)

CRAY X-MP multiprocessor system organization



Deformed powertrain assembly illustrating strain energy density distribution under first mode of vibration. (Credit: ©1984, Ford Motor Co.)

CRAY

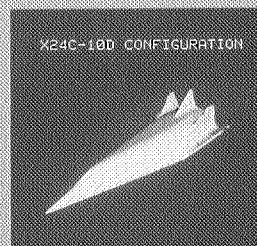
System overview



Aerodynamic simulation

Airplane designers have long relied on wind tunnel tests to evaluate the aerodynamics of airplanes and airplane sections. But wind tunnel testing requires the time-consuming and costly construction of physical test models. CRAY X-MP supercomputers enable

airplane designers to evaluate designs mathematically and to modify designs faster and more cost-effectively than they could by relying solely on wind tunnel tests. In recent years, the auto industry has also begun enjoying the benefits of aerodynamic testing via supercomputer.



(Left) Full model of a Martin-Marietta X24C-10D Lifting Body. (Right) A transverse flow plane (perpendicular to the flight direction) just behind the cockpit. The coloration shows the variation in Mach number. (Far right) Flow plane at the tail of shuttle-like aircraft.

The CRAY X-MP/4 computers

The top-of-the-line CRAY X-MP/4 computer systems offer an order of magnitude greater performance than the original CRAY-1. They are configured with eight or sixteen million 64-bit words of ECL bipolar memory and provide a maximum memory bandwidth 16 times that of the CRAY-1. Central memory has a bank cycle time of 38 nanoseconds (nsec) and is shared by four identical CPUs with a clock cycle time of 9.5 nsec. The X-MP/4 mainframe is the familiar 12-column 270° arc chassis with the same electrical requirements as the CRAY-1.

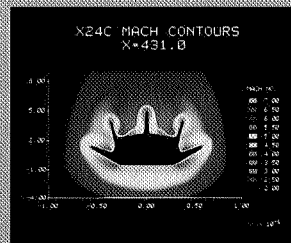
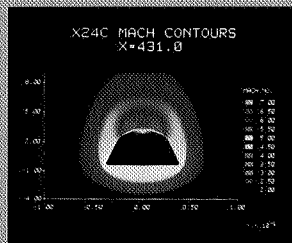
Each of the four CRAY X-MP/4 processors has scalar and vector processing capability and can access all of central memory. The CPUs may operate independently on separate jobs or may be organized in any combination to operate jointly on a single job.

The raw computational power of the CRAY X-MP/4 systems is augmented by the powerful input/output and data-handling capabilities of the CRAY I/O Subsystem (IOS). The IOS is integral to all CRAY X-MP computers and enables fast, efficient data access and processing by the CPUs.

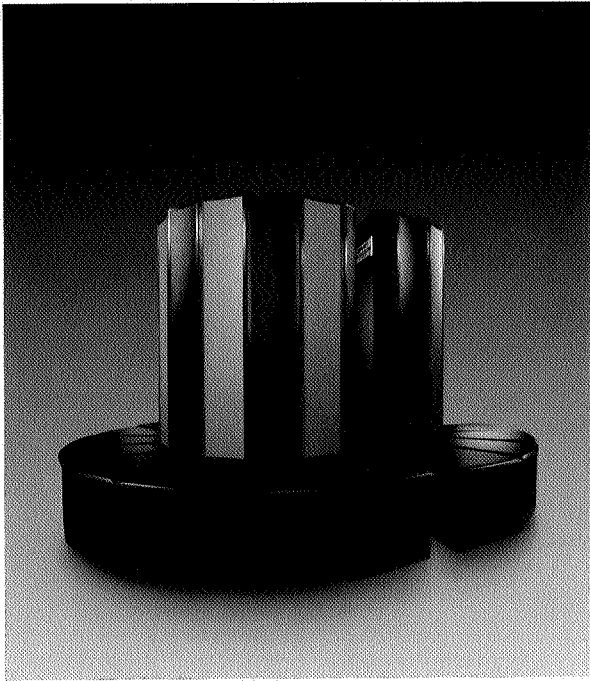
Cray Research's DD-49 disk drive matches the power of the X-MP/4 models, offering 1200-Mbyte capacity, sustained transfer rates of 9.8 Mbyte/sec and very fast access times (2 milliseconds).

In addition to high-capacity, fast access disk technology, the field-proven SSD offers up to 1024 Mbytes of very fast random-access secondary MOS memory. The SSD connects to the CRAY X-MP/4 systems through two very high-speed channels with a maximum aggregate transfer rate of 2000 Mbyte/sec. The SSD, in conjunction with the X-MP/4 multiprocessor architecture, enables users to fully exploit existing applications and to develop new algorithms to solve larger and more sophisticated problems in science and engineering — problems that could not be attempted before due to computational or I/O limitations.

These cross-sections represent three-dimensional airflows as depicted by the Navier-Stokes equations. Their computational requirements make the Navier-Stokes equations the most difficult to solve, but they are also the most accurate. (Credit: W.L. Hankey, S.J. Scherr, J.S. Shang, Air Force Wright Aeronautical Laboratories.)



CRAY



The CRAY X-MP/2 computers

The field-proven CRAY X-MP/2 models have become the established price and performance leaders in the supercomputer industry. The new X-MP dual-processor systems offer up to four times the memory and require only half the electrical power of the original CRAY X-MP/2 systems. Overall throughput is typically three to five times that of a CRAY-1.

The CRAY X-MP/2 systems are available with four, eight or sixteen million 64-bit words of shared MOS central memory, providing a maximum memory bandwidth four times that of the CRAY-1. Each CPU has a 9.5 nsec clock cycle time and memory bank cycle time of 76 nsec. The CRAY X-MP/2 models consist of eight vertical columns arranged in a 180° arc.

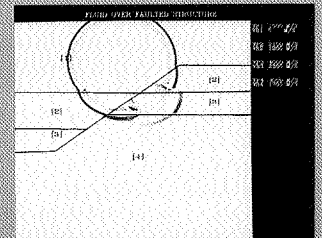
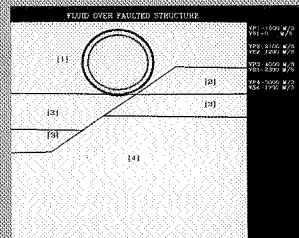
As with the X-MP/4 systems, the CRAY X-MP/2 CPUs can operate independently on different programs or can be harnessed together to operate on a single user program.

CRAY X-MP/2 computers incorporate the same I/O Subsystem and SSD hardware as the X-MP/4 models. One SSD channel, with a total transfer rate of 1000 Mbyte/sec, connects the optional SSD to the mainframe. Typically, the system is configured with DD-49 disk drives.

Geological exploration

Inducing a shock in the ground and recording sound waves reflected back to the surface is a method scientists use to "see" underground structures. The method is called reflection seismology and can indicate the presence or absence of petroleum and other resource deposits. However, the amount

of data needed to profile a large volume of earth accurately can be immense, and the required analyses are staggeringly complex. CRAY X-MP systems can perform detailed analyses on these large amounts of data in a timely and cost-effective way, saving petroleum companies time and money.

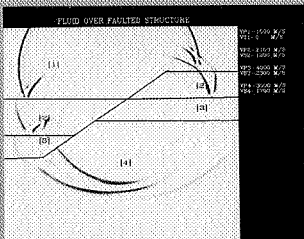
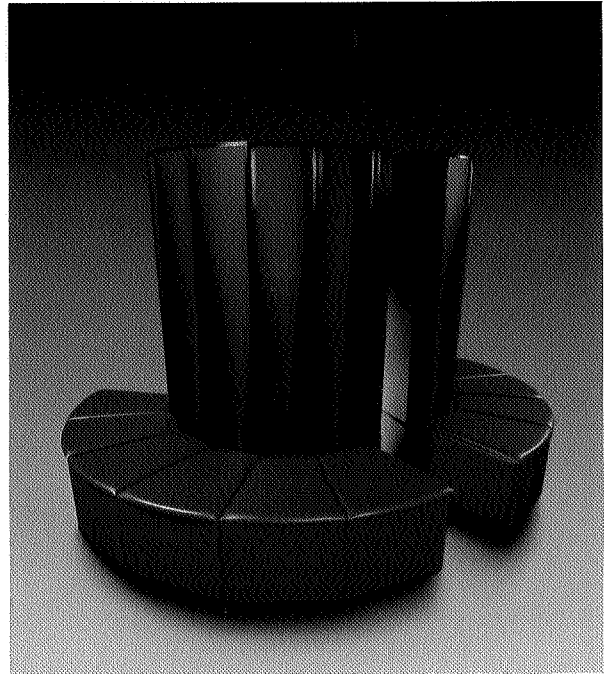


The CRAY X-MP/1 computers

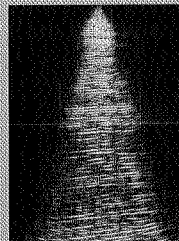
The CRAY X-MP/1 models combine a single CRAY X-MP CPU with one, two, four or eight million 64-bit words of static MOS memory. Memory bandwidth is four times that of the CRAY-1. Single processor CRAY X-MP systems typically provide the user with 1.5 to 2.5 times CRAY-1 power at a comparable cost. The CRAY X-MP/1 CPU has a 9.5 nsec clock cycle time, and a memory bank cycle time of 76 nsec. The X-MP/1 mainframe is a six-column, 135° arc chassis requiring the same electrical power as the X-MP/2.

CRAY X-MP/1 models use the same I/O Subsystem and support the same range of Solid-state Storage Device models as the CRAY X-MP/2 models. Typically, the X-MP/1 is configured with DD-39 disk drives.

With the availability of a wide range of applications software and its superior price/performance characteristics, the entry-level CRAY X-MP is particularly appropriate for the first-time supercomputer customer.



Color variable density displays of a seismic wave propagation from a fluid through a faulted structure. The full elastic equation was used, demonstrating the conversion of P waves (red) into S waves (blue). (Credit: Dan Kosloff and Moshe Reshet, University of Tel Aviv)



Common source seismic record after depth migration. It contains 96 traces, each six seconds long and sampled at an interval of four milliseconds, producing a total of 144,000 samples. Hundreds of these records are required to study subsurface geology. (Credit: Geo-Quest International, Inc.)

CRAY

CRAY X-MP design

The CRAY X-MP Series design combines high-speed scalar and vector processing with multiple processors, large and fast memories and high-performance I/O. The result is exceptional speed and high overall system throughput. Innovative architecture and technologies built into the CRAY X-MP make such performance a practical reality.

Processors

Each CRAY X-MP processor offers very fast scalar processing with high-speed processing of long and short vectors. Additionally, multiprocessor models enable the user to exploit the extra dimension of multitasking.

The scalar performance of each processor is attributable to its fast clock cycle, short memory access times and large instruction buffers. Vector performance is supported by the fast clock, parallel memory ports and flexible hardware chaining. These features allow simultaneous execution of memory fetches, arithmetic operations and memory stores in a series of linked vector operations. As a result, the processor design provides high-speed and balanced vector processing capabilities for short and long vectors characterized by heavy register-to-register or heavy memory-to-memory vector operations.

The overall effective performance of each processor executing typical user programs with interspersed scalar and vector codes (usually short vectors) is ensured through fast data flow between

scalar and vector functional units, short memory access time for vector and scalar references and short start-up time for both scalar and vector operations. As a result, CRAY X-MP computers offer high performance using the standard FORTRAN compiler, without the need for hand-coding or algorithm restructuring.

On all models, a second vector logical unit is used to provide twice the execution speed of bit-level logical operations in each CPU.

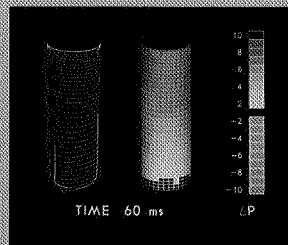
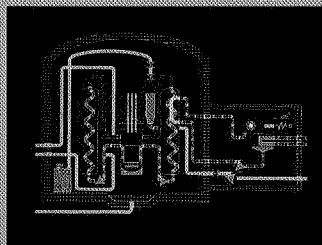
Each X-MP processor also includes instructions for the efficient manipulation of randomly distributed data elements and conditional vector operations. Gather/scatter instructions allow for the vectorization of randomly organized data, and the compressed index instruction allows for the vectorization of unpredictable conditional operations. With these features, CPU performance can be improved by a factor of five for program segments dependent on the manipulation of sparse matrices.

Central memory

Depending on the model, one to sixteen million 64-bit words of directly addressable memory is available with the CRAY X-MP Series. Options for field upgrade of memory are available on all models. The large memory sizes enable users to solve larger problems than before without the need for out-of-memory techniques. CRAY X-MP memory features single-bit error correction, double-bit error detection (SECCDED) logic.

Nuclear energy research

Computer simulation of nuclear power plants requires the most advanced computer systems available. Only supercomputers such as the CRAY X-MP provide the computing power needed to simulate the intricate fluid flow, heat transfer and neutronics phenomena that characterize today's nuclear power plants.



(Far left) A pressurized water nuclear reactor. The reactor core, primary and secondary heat exchange loops and containment system are shown. (Left) A cylindrical section of a nuclear reactor core with a calculated three-dimensional pressure field. The vertical color scale indicates the pressure drop (ΔP) in the vertical direction.

CRAY X-MP mainframe highlights:

- Four processors sharing 8 or 16 million words of ECL bipolar memory with the X-MP/4, or
- Two processors sharing 4, 8 or 16 million words of MOS memory with the X-MP/2, or
- One processor with 1, 2, 4 or 8 million words of MOS memory, on the X-MP/1
- 9.5 nsec clock cycle
- 38 nsec (on X-MP/4) or 76 nsec (on X-MP/1 and X-MP/2) memory bank cycle time
- SECCDED memory protection
- Four parallel memory ports per processor
- Flexible hardware chaining for vector operations
- Second vector logical unit
- Gather/scatter and compressed index vector support
- Flexible processor clustering for multitasking applications
- Dedicated registers for efficient interprocessor communications and control

The CRAY X-MP multiprocessor systems share a central memory organized in interleaved memory banks that can be accessed independently and in parallel during each machine clock period. Each X-MP processor has four parallel memory ports connected to central memory: two for vector

fetches, one for result store and one for independent I/O operations. Thus, each processor of a CRAY X-MP system has four times the memory bandwidth of a CRAY-1. Ensuring high efficiency, the multiport memory has built-in conflict resolution hardware to minimize delays and maintain the integrity of simultaneous memory references to the same memory bank.

The interleaved and efficient multiport memory design, coupled with the short memory cycle time, provides high-performance memory organization with sufficient bandwidth to support high-speed CPU and I/O operations in parallel.

Multiprocessors and multitasking

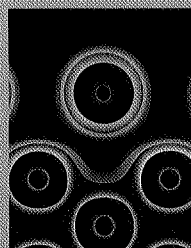
The CRAY X-MP multiple-CPU configurations have made Cray Research the recognized leader in multiprocessing. They continue to offer users the opportunity to process jobs faster than with single CPUs by using either multiprocessing or multitasking techniques.

Multiprocessing allows several programs to be executed concurrently on multiple CPUs of a single mainframe. Multitasking is a feature that allows two or more parts of a program (tasks) to be executed in parallel sharing a common memory space, resulting in substantial throughput improvements over serially executed programs. Performance improvements are in proportion to the number of tasks that can be constructed for the program and the number of CPUs that can be applied to the separate tasks.

Computational physics

In certain fields of physics, such as quantum chromodynamics and condensed matter physics, experimentation is difficult if not impossible. But by tapping the CRAY X-MP's extraordinary processing power, physicists can experiment on mathematical models of atomic and sub-

atomic structures and thus refine their theories faster than would be possible by any other means.



Charge density contours for an atomic overlayer of cesium on tungsten. Using a Cray system for computation and graphics generation, physicists have investigated the electronic structures of these materials and obtained results impossible to determine analytically. (Credit: Arthur J. Freeman, Henri J. F. Jansen, Erich Wimmer, Northwestern University.)

CRAY

When executing in multitasking mode, all processors are identical and symmetrical in their programming functions; no CPU is dedicated to any one function. Any number of processors (a cluster) can be dynamically assigned to perform multiple tasks of a single job. In order to provide flexible and efficient multitasking capabilities, special hardware and software features have been built into the systems. These features allow one or more processors to access shared memory or high-speed registers for rapid communication and data transmission between CPUs. All of these capabilities are made available through library routines which can be accessed from FORTRAN. In addition, hardware provides built-in detection of deadlocks within a cluster of processors.

Experience shows that multitasked applications running on CRAY X-MP/2 computers can realize speed increases of 1.8 to 1.9 times over single-processor CRAY X-MP execution times; speed increases of 3.5 to 3.8 times have been achieved with the CRAY X-MP/4 systems.

Input/output processing

For super-scale problems requiring extensive data handling, Cray has developed hardware that ensures computing power is not held captive by I/O limitations. The architecture of the IOS, with its parallel data paths and direct access to main memory, results in a very high I/O bandwidth with a minimum of interference to computation.

Input/output highlights:

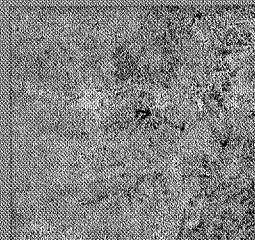
- 6-Mbyte, 100-Mbyte and 1000-Mbyte channels
- I/O Subsystem with:
 - Parallel disk streaming capabilities, one controller per disk cabinet
 - I/O buffering for disk- and tape-resident datasets
 - Software support for parallel disk striping
 - Buffer memory-resident datasets
 - High-performance disk drives
 - High-performance on-line tape handling
 - Front-end system communication with IBM, CDC, DEC, Honeywell, Data General and Sperry computer systems
 - Linkage to workstations such as Apollo™ and Sun™ via Network Systems Corporation (NSC) network adapters

The I/O Subsystem (IOS) is an integral part of the CRAY X-MP design and acts as a data distribution point for the X-MP mainframe. The IOS handles I/O for a variety of front-end computer systems and peripherals such as disk units and plug-compatible IBM Series 3420 and 3480 tape subsystems. The IOS includes two, three or four interconnected I/O processors, each with its own local memory, and a common buffer memory.

Image processing

Earth-imaging satellites, space probes and medical imaging technologies generate tremendous amounts of data. However, the data must often be processed extensively to be useful. The pictures created by digital imaging technology are composed of millions of tiny

dots called pixels. Processing the information contained in these pixels in a practical time-frame requires the processing speed of a CRAY X-MP. For everything from scanning the Earth for resources to tracking down deadly tumors, CRAY X-MP systems are ideal for the most sophisticated image processing applications.



Buffer memory is solid-state secondary storage, accessible by all of the I/O processors in the IOS. With its 8, 32 or 64 Mbytes of static MOS memory, it provides I/O buffering of data to and from the peripheral devices. It can also be used to store user datasets, thus contributing to faster and more efficient data access by the CPUs.

Complementing and balancing CRAY X-MP computing speeds are the DD-39 and DD-49 disk drives, high density (1200-Mbyte) magnetic storage devices. The DD-39 can sustain a data transfer rate of 5.9 Mbyte/sec with an average access time of 18 milliseconds (msec); the DD-49 can sustain a rate of 9.8 Mbyte/sec with an average access time of 16 msec. These disks are the fastest available, and when combined with the data handling and buffering capability of the IOS, they provide unsurpassed I/O performance. From 2 to 32 disk drives can be connected to an I/O Subsystem for up to 38 gigabytes of total disk storage. Typically, DD-49 disks are configured on the CRAY X-MP/4 and CRAY X-MP/2 and DD-39 disks are configured on the CRAY X-MP/1.

Effective disk transfer rates can be increased further by the use of optional disk striping techniques. When specified, striping causes system software to distribute a single user dataset across two to five disk drives, depending on the device type. Successive disk blocks are allocated

cyclically across the drives and consecutive blocks can thus be accessed in parallel. The resultant I/O performance improvements are in proportion to the number of disk drives used. DD-49 disks may be striped two or three wide; DD-39 disks may be striped two to five wide.

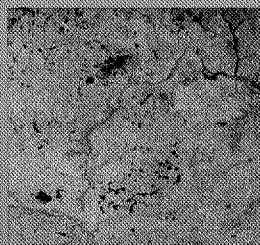
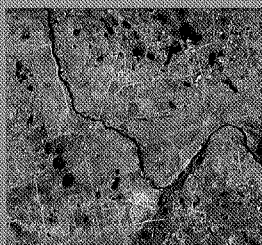
The CRAY X-MP supports three channel types, identified by their maximum transfer rates: 6 Mbyte/sec, 100 Mbyte/sec and 1000 Mbyte/sec. Depending on the X-MP model, two or four 6-Mbyte channels and one to four 100-Mbyte channels are connected to each system. The 100-Mbyte channels are available for transferring data between the I/O Subsystem and central memory and/or to the SSD.

Solid-state Storage Device

The optional Solid-state Storage Device (SSD) is a very fast random-access device suited for use with the CRAY X-MP. The SSD in conjunction with multiprocessor architecture allows the development of algorithms to solve larger and more sophisticated problems in science and engineering.

The SSD is used as a fast-access device for large prestaged or intermediate files generated and manipulated repetitively by user programs. Datasets may be assigned to the SSD by a single Cray Operating System (COS) control statement without modification of the user program.

Three Thematic Mapper images of regions in the Midwest. (Left) An area west of Kansas City, Kansas. (Right) A close-up of the Minneapolis/St. Paul area, and (far right) an area southwest of the Twin Cities.



Each image was produced by Cray Research's CSADIE software on a CRAY X-MP from a 336 million-byte Thematic Mapper database.

CRAY

System performance is significantly enhanced by the SSD's exceptionally high transfer rates and short data access times. Up to 1024 Mbytes of rapid-access MOS memory may be configured on an SSD. Transfer rates of 100 to 1000 Mbyte/sec per channel and access times of less than 25 microseconds are achievable between the SSD and an X-MP mainframe. The SSD offers significant potential for performance improvement on I/O-bound applications, and thus allows users to develop new algorithms that would not otherwise be practical with traditional disk I/O.

SSD highlights:

- Memory size of 256, 512 or 1024 Mbytes
- Support for:
 - Two 1000-Mbyte channels for linkage to CRAY X-MP/4
 - One 1000-Mbyte channel for linkage to CRAY X-MP/1 or X-MP/2
- SECCDED memory protection
- Software support to allow existing programs to use the SSD without program modification
- Direct data path (100-Mbyte channel) between SSD and IOS

An SSD can also be connected to the I/O Subsystem. This connection enables data to be transferred between the IOS and the SSD directly, without passing through central memory.

On the CRAY X-MP/4, support is provided to link the SSD to the mainframe via two 1000-Mbyte channels. For linkage to the X-MP/1 and X-MP/2 models, one 1000-Mbyte channel is used.

Physical characteristics

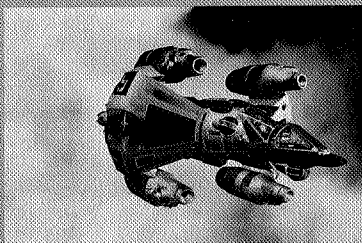
The CRAY X-MP is extremely compact; keeping wire lengths short minimizes signal propagation times. The elegant and compact CRAY X-MP/1 mainframe consists of six vertical columns arranged in a 135° arc that occupies 32 square feet (3 square meters) of floor space. A CRAY X-MP/2 model consists of eight vertical columns arranged in a 180° arc that occupies 43 square feet (4 square meters) of floor space. And a CRAY X-MP/4 system is composed of 12 vertical columns arranged in a 270° arc and requires just 64 square feet (6 square meters) of floor space.

The accompanying I/O Subsystem is composed of four vertical columns in a 90° arc and occupies 24 square feet (2.3 square meters) of floor space. The IOS can be positioned up to 19 feet (5.8 meters) from the mainframe.

Graphics

For many supercomputer applications, graphics are needed to display meaningfully the large amounts of data produced. But graphics is itself a unique application. Computer graphics has revolutionized commercial animation. Whether in motion pictures, advertisements or the latest rock video,

computer graphics transport viewers to worlds made not of real objects, but of digital information. CRAY X-MP supercomputers provide the speed and memory needed to generate the most complex and convincing visual displays. With Cray systems, animators can create motion picture sequences without using sets or props



Scene from "The Last Starfighter" motion picture - Gunstar moving through Rylosian clouds. (Credit: Digital Scene SimulationSM by Digital Productions, Los Angeles, California, U.S.A. © 1985. All rights reserved.)

System configuration options

	X-MP/1	X-MP/2	X-MP/4
Mainframe			
CPUs	1	2	4
Bipolar memory (64-bit words)	N/A	N/A	8 or 16M
MOS memory (64-bit words)	1, 2, 4 or 8M	4, 8 or 16M	N/A
6-Mbyte channels	2 or 4	4	4
100-Mbyte channels	1 or 2	2	4
1000-Mbyte channels	1	1	2
I/O Subsystem			
I/O processors	2, 3 or 4	2, 3 or 4	4
Disk storage units	2-32	2-32	2-32
Magnetic tape channels	1-8	1-8	1-8
Front-end interfaces	1-7	1-7	1-7
Buffer memory (Mbytes)	8, 32 or 64	8, 32 or 64	64
Solid-state Storage Device			
Memory size (Mbytes)	256, 512 or 1024	256, 512 or 1024	256, 512 or 1024

N/A signifies option is not available on the model

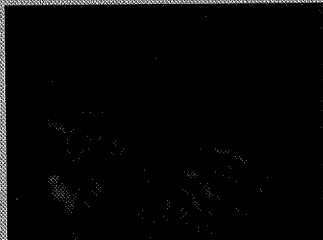
The optional SSD consists of four columns arranged in a 90° arc occupying 24 square feet (2.3 square meters) and is connected to the mainframe through one or two short aerial bridgeways, depending on model.

High-speed 16-gate array integrated logic circuits are used in the CRAY X-MP CPUs. These logic circuits, with typical 300 to 400 picosecond propagation delays, are faster and denser than the circuitry used in the CRAY-1. CRAY X-MP/4

memory is composed of ECL bipolar circuits; CRAY X-MP/1 and CRAY X-MP/2 memory is composed of static MOS components.

The dense concentration of components requires special cooling techniques to overcome the accompanying problems of heat dissipation. A proven, patented cooling system using liquid refrigerant cooling maintains the necessary internal system temperature which contributes to high system reliability and minimizes the requirement for expensive room cooling equipment.

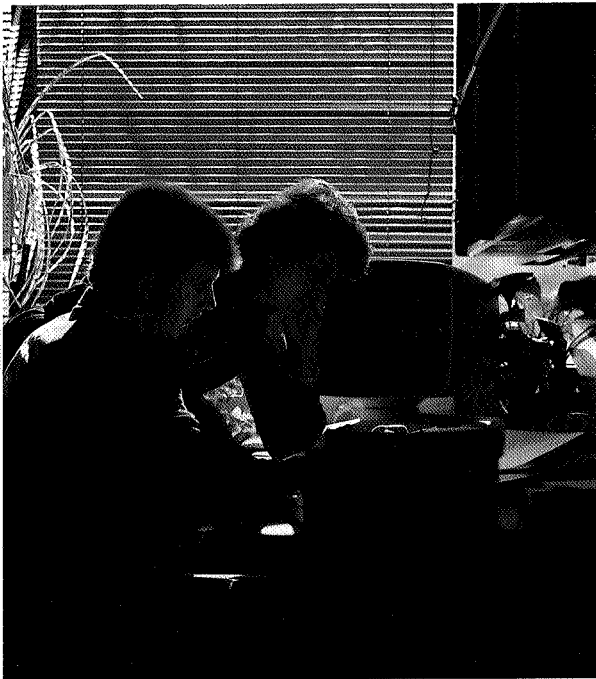
A terrain mapping of the San Francisco Bay area developed for real-time emergency assessment. A 7-gigabyte database and a ray-tracing algorithm were used to prepare the image. (Credit: * Patrick Weidhaas, Lawrence Livermore National Laboratory, 1983.)



Characters from "The Adventures of Andre and Wally B." generated on a CRAY X-MP. (Credit: * 1984: Lucasfilm.)

CRAY

CRAY X-MP software



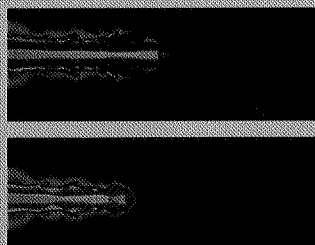
A full range of system and applications software compatible with that provided on the CRAY-1 computer systems is available for the CRAY X-MP systems. This software includes the efficient Cray Operating System (COS), an auto-vectorizing ANSI 78 Cray FORTRAN compiler, extensive FORTRAN and scientific library routines, program and dataset management utilities, debug aids, a selection of compilers, a powerful Cray assembler (CAL) and a wealth of third-party and public-domain application codes.

The operating system, the FORTRAN compiler and library programs are designed to allow users to take advantage of the vectorizing, multiprocessing and multitasking features of the CRAY X-MP systems. Multitasking is a technique whereby an application program can be partitioned into independent tasks that can execute in parallel on a multiprocessor CRAY X-MP system. Two methods can be used: FORTRAN callable subroutines to explicitly define and synchronize tasks at the subroutine level, or a FORTRAN preprocessor to identify DO loops whose independent iterations may be dispatched to separate processors. The first method (macrotasking) is best suited to programs with large tasks running with dedicated processors. The second method (microtasking) is beneficial for programs with any size tasks running in either a dedicated or a production environment.

Computational fluid dynamics

Fluid flow characterizes physical processes ranging from the circulation of gases in the atmosphere to the emission of supersonic jets from galaxies. Although the equations describing fluid physics were identified early in the 19th century, the development of super-

computers made possible the accurate modeling of complex three-dimensional flow fields. Today, CRAY X-MP systems are used for state-of-the-art fluid flow modeling in studies of coating flows, petroleum reservoir simulations and research in the atmospheric and astrophysical sciences.



Evolutionary images of supersonic gas jets boring their way through other gases that are 100 times (left) and 10 times (right) denser. In simulating supersonic gas jets, the combination of color imaging with a Cray system enables one to probe in detail both the dynamics and internal physics of the gas flows. (Credit: Michael Norman, Larry Smarr and Karl-Heinz Winkler.)

The Cray Operating System efficiently delivers the full power of the hardware to both batch and interactive users. The operating system, which is distributed between central memory and the IOS, effectively manages high-speed data transfers between the CRAY X-MP and peripherals such as disks, SSD and on-line magnetic tapes. Standard system software is also offered for interfacing the CRAY X-MP Computer System with other vendor's operating systems and with networks. This is described further under "System Integration". COS also includes a variety of utility programs that assist in program development and maintenance.

Cray's FORTRAN compiler fully meets the ANSI 78 standards while offering a high degree of automatic scalar and vector optimization within these standards. The Cray compiler permits maximum portability of programs between different Cray systems and accepts many nonstandard constructs written for other vendor's compilers. There is no need for using nonstandard vector syntax to produce vectorized object code. The compiler is fully supported by highly optimized FORTRAN and scientific library routines for maximum performance from the CRAY X-MP Series computers.

The success of the CRAY-1 stimulated the development of a wide variety of third-party and public domain application programs, which are now available on CRAY X-MP computers. Major applications codes are offered for the CRAY X-MP in fields such as computational fluid dynamics, mechanical engineering, nuclear safety, circuit design, seismic processing, image processing, molecular modeling and artificial intelligence.

Cray Research provides support for the ongoing process of converting and maintaining applications software on the CRAY X-MP Series. A comprehensive directory of available programs is published by the Cray Applications Software Library Service.

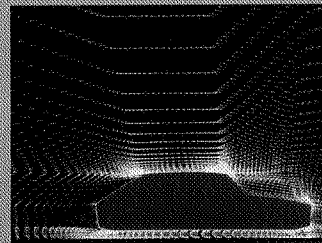
The above-mentioned software teamed with an ISO Level 1 Pascal compiler, a sort package, a C compiler and many other software tools and products, provides users with the software they need to use the CRAY X-MP to its fullest capabilities.

Software highlights:

- An efficient multiprogramming and multitasking operating system
- High-performance I/O management
- Versatile system utility programs
- An auto-vectorizing and optimizing ANSI 78 FORTRAN compiler
- Highly optimized scientific libraries
- C compiler
- ISO Level 1 PASCAL
- A sort package
- A wide variety of major application programs

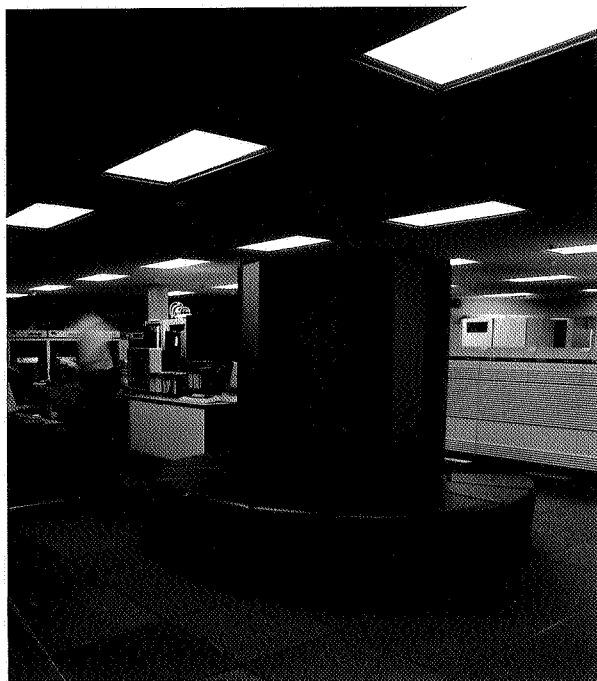


Two-dimensional flow around a fast-back automobile. The image was produced by solving the two-dimensional Navier-Stokes equations.
(Credit: Dornier GmbH, West Germany.)



CRAY

System integration



CRAY X-MP Series computers are designed to be connected easily to one or more front-end computer systems. Thus, a CRAY X-MP computer can be added into an existing configuration so that the end user continues to work in a familiar computer environment but now has access to a considerably greater computational resource. Jobs can be submitted from a front-end to the CRAY X-MP for processing and results returned to the user on the originating front-end or optionally to a different front-end. Data can be transferred readily between any front-end system and the X-MP, with data conversion and reformatting handled automatically by software.

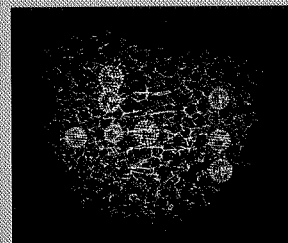
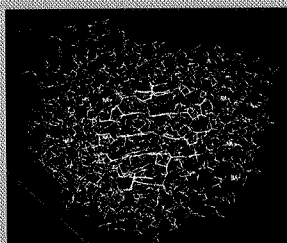
Cray Research offers hardware interfaces that connect the CRAY X-MP I/O Subsystem to a wide variety of front-end equipment, including IBM, CDC, DEC, Data General, Sperry and Honeywell. Additionally, the I/O Subsystem may be connected to one or more Network Systems Corporation HYPERchannel™ adapters for those installations wishing to configure their CRAY X-MP in a high-speed local area network.

Cray Research provides software interface support for a variety of front-end systems. Station software runs on the front-end system and provides the logical connection between other vendors' equipment and CRAY X-MP computers. Standard Cray software is available for the following: IBM MVS and VM, CDC NOS and NOS/BE, DEC VAX/VMS, Data General RDOS and AT&T UNIX™. Station software for Sperry and Honeywell operating systems is currently available from third-party sources.

Molecular science

Computer simulation is an invaluable tool for studying molecular motion, which can occur in a matter of picoseconds (trillionths of a second). Using CRAY X-MP systems, scientists can simulate atomic and molecular events and gain insight into chemical reaction rates,

catalytic mechanisms, properties of synthetic polymers and the shapes of biological molecules millions of atoms long. The detailed and highly iterative mathematics involved in modeling such systems demands the computational capability of the CRAY X-MP



Support and maintenance

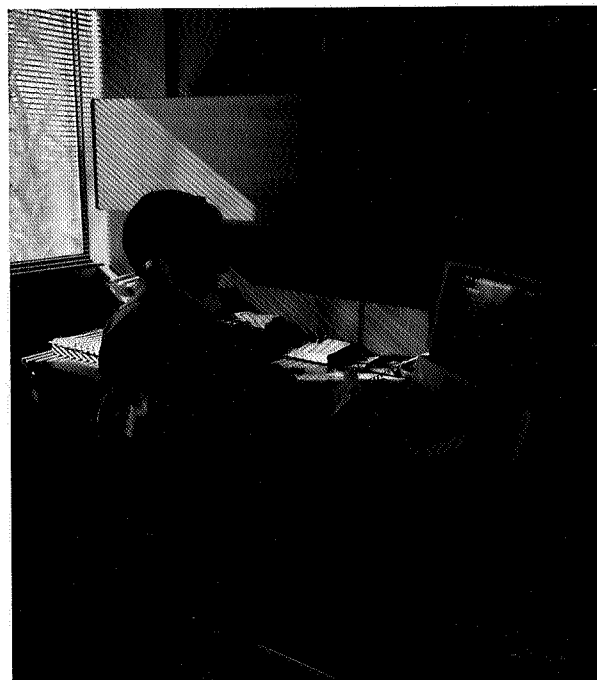
Customer support

Cray Research has developed a comprehensive array of support services to meet customer needs. From pre-installation site planning through the life of the installation, ongoing on-site engineering and system software support is provided. Additional assistance is available from technical centers throughout the company.

Cray Research provides comprehensive documentation and offers customer training on-site or at Cray training facilities. Cray Research's responsive customer support program results from extensive accumulated experience in the supercomputer business and from a strong customer orientation.

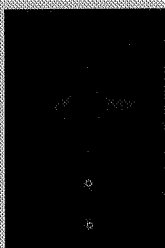
CRAY X-MP reliability and maintenance

Cray Research recognizes the need for high system reliability while maintaining a high level of performance. The use of higher-density integrated circuits, an overall higher level of component integration and an increased cooling capacity, all ensure that X-MP system reliability exceeds that of the CRAY-1. Components used in CRAY X-MP computers undergo strict inspection and checkout prior to assembly into a system. All CRAY X-MP Series computers undergo rigorous operational and reliability tests prior to shipment.



Preventive maintenance techniques identify potential problems before they affect system performance. Diagnostics can be invoked locally at the customer's site or remotely by Cray Research technical support personnel. The Cray maintenance philosophy is to repair and replace modules on-site with minimum system downtime and highest system availability.

DNA molecules with water and sodium ions. (Far left) DNA-water-counterion system. (Left) DNA with full surface display of the sodium ions. A 106-picosecond molecular dynamics simulation was performed, and for each step a complicated evaluation of forces involving all 2733 atoms was performed. (Credit: Computer Graphics Laboratory, UCSF.)



Distribution of individual electrons in an oxygen molecule, O₂. One electron has a high probability of being found between the atoms in a "bonding orbital" (far left), while the other electron avoids the region between the atoms representing an "anti-bonding orbital" (left). Near the atomic nuclei, there is a small region where the

probability of finding an electron is very high, as illustrated by the spiked regions in the graphs. The power of a CRAY X-MP is required in solving the kind of quantum-mechanical equations used to form these images.

CRAY

CRAY X-MP design detail

Mainframe

CRAY X-MP single- and multiprocessor systems are designed to offer users outstanding performance on large-scale, compute-intensive and I/O-bound jobs.

CRAY X-MP mainframes consist of six (X-MP/1), eight (X-MP/2) or twelve (X-MP/4) vertical columns arranged in an arc. Power supplies and cooling are clustered around the base and extend outward.

Model	Number of CPUs	Memory size (millions of 64-bit words)	Number of banks
CRAY X-MP/416	4	16	64
CRAY X-MP/48	4	8	32
CRAY X-MP/216	2	16	32
CRAY X-MP/28	2	8	32
CRAY X-MP/24	2	4	16
CRAY X-MP/18	1	8	32
CRAY X-MP/14	1	4	16
CRAY X-MP/12	1	2	16
CRAY X-MP/11	1	1	16

Hardware features:

- 9.5 nsec clock
- One, two or four CPUs, each with its own computation and control sections
- Large multiport central memory
- Memory bank cycle time of 38 nsec on X-MP/4 systems, 76 nsec on X-MP/1 and X-MP/2 models
- Memory bandwidth of 25-100 gigabits, depending on model
- I/O section
- Proven cooling and packaging technologies

A description of the major system components and their functions follows.

CPU computation section

Within the computation section of each CPU are operating registers, functional units and an instruction control network — hardware elements that cooperate in executing sequences of instructions. The instruction control network makes all decisions related to instruction issue as well as coordinating the three types of processing within each CPU: vector, scalar and address. Each of the processing modes has its associated registers and functional units.

The block diagram of a CRAY X-MP/4 (opposite page) illustrates the relationship of the registers to the functional units, instruction buffers, I/O channel control registers, interprocessor communications section and memory. For multiple-processor CRAY X-MP models, the interprocessor

communications section coordinates processing between CPUs, and central memory is shared.

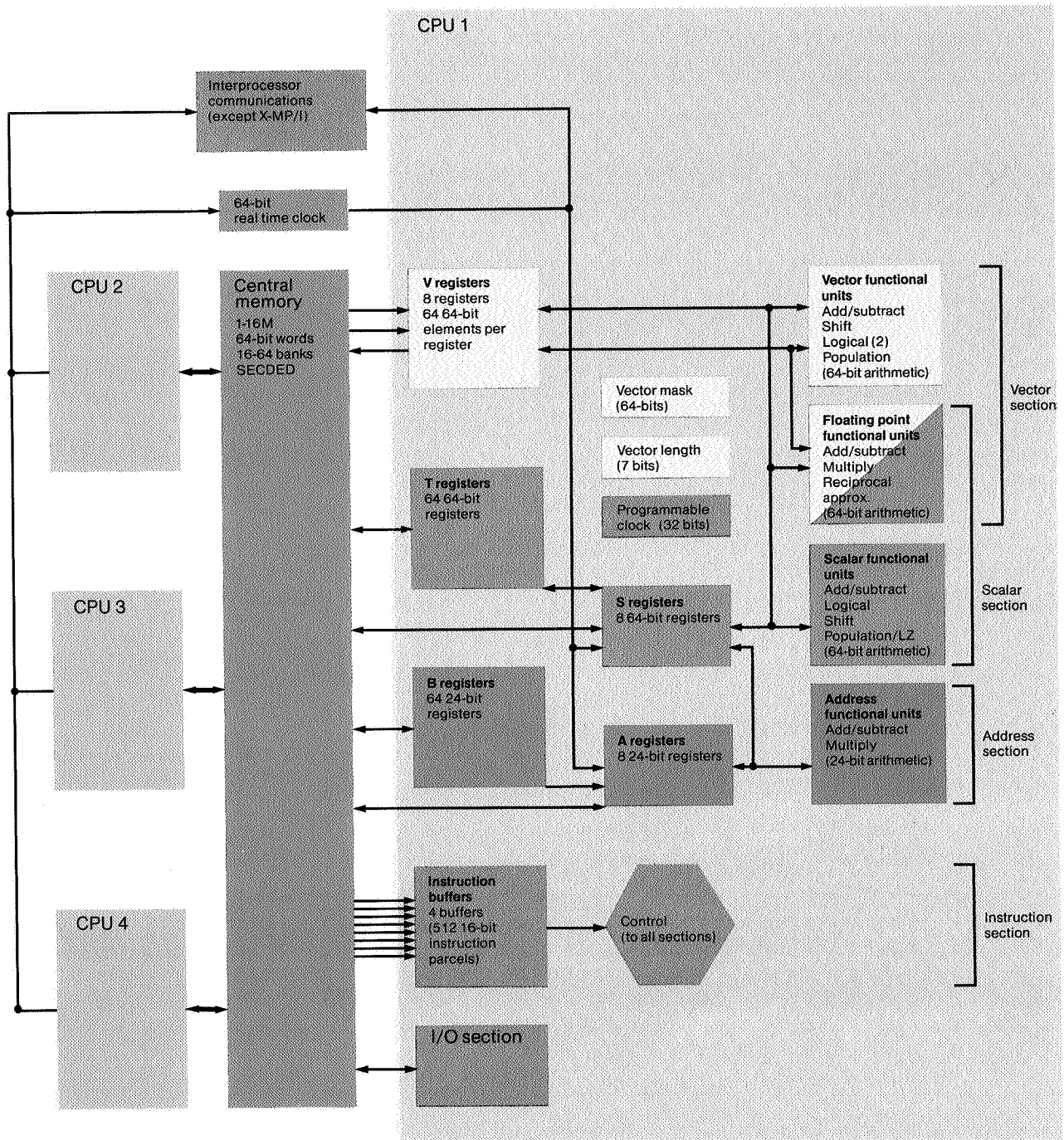
Registers

The basic set of programmable registers is composed of:

- Eight 24-bit address (A) registers
- Sixty-four 24-bit intermediate address (B) registers
- Eight 64-bit scalar (S) registers
- Sixty-four 64-bit scalar-save (T) registers
- Eight 64-element (4096-bit) vector (V) registers with 64 bits per element

The 24-bit A registers are generally used for addressing and counting operations. Associated with them are 64 B registers, also 24 bits wide. Since the transfer between an A and a B register takes only one clock period, the B registers assume the role of data cache, storing information for fast access without tying up the A registers for relatively long periods.

CRAY X-MP system organization



The 64-bit S registers are used for floating-point, logical and some integer and character operations. The 64-bit T registers act as cache memory for the S registers. Typically, the B and T registers are used for storing local variables within subroutines.

Each of the eight V registers is actually a set of sixty-four 64-bit registers. The V registers are used for vector operations. Successive elements from a V register enter a functional unit in successive clock periods. The effective length of a vector register for any operation is controlled by a program selectable vector length (VL) register. The vector employed in any calculation need not contain exactly 64 elements. A vector mask (VM) register allows for the logical selection of particular elements of a vector.

In addition to the operating registers, the CPU contains a variety of auxiliary and control registers. These generally are not accessible to a programmer.

Addressing

Instructions that reference data do so on a word basis. Branch instructions, on the other hand, reference parcels within words; the lower two bits of an

address identify the location of an instruction parcel in a word. Significantly, the destination of a jump can be any instruction parcel in a four-million-word instruction segment; word alignment is not required.

The expanded addressing capability in the 8- and 16-million-word systems is accomplished by using 24-bit direct word addressing of data elements while retaining 24-bit parcel addressing for instruction references. In addition there is a mode that allows the execution of a program that is compatible with conventional 22-bit data addressing.

Hardware supports separation of memory segments for each user's data and program, thus facilitating concurrent programming.

Instruction set

The comprehensive CRAY X-MP instruction set features over 100 operation codes and provides for both scalar and vector processing. Most instructions occupy 16 bits (one parcel); certain branch instructions and memory reference operations occupy 32 bits (two parcels).

Floating-point instructions provide for addition, subtraction, multiplication and reciprocal approximation. The reciprocal approximation instruction enables CRAY X-MP computers to have a completely segmented divide operation using a floating-point divide algorithm.

Integer addition, subtraction and multiplication are provided for by the hardware. An integer multiply operation produces a 24-bit result; an addition or subtraction produces either a 24-bit or a 64-bit result. An integer divide is accomplished through a software algorithm using floating-point hardware.

The instruction set includes Boolean operations for OR, AND, exclusive OR and for a mask-controlled merge operation. Shift operations allow for the manipulation of 64-bit or 128-bit operands to produce a 64-bit result. Similar 64-bit arithmetic capability is provided for both scalar and vector processing.

A programmer may index throughout memory in either scalar or vector processing mode. This full indexing capability allows matrix operations in vector mode to be performed on rows, columns, diagonals and, in general, on any set of data that is stored in memory with regular spacing between elements with no performance degradation relative to sequentially stored data elements. With gather/scatter, a vector of indices may be used to reference a random pattern of data in memory. Additionally, a compressed or dense index may be generated containing only those items that correspond to some testable condition.

Instructions for population, parity and leading zero counts (scalar only) return bit counts based on register contents.

Instructions for population, parity and leading zero counts (scalar only) return bit counts based on register contents.

Programmable clock

A 32-bit programmable real-time clock that has a frequency of 105 MHz, corresponding to an increment of 9.5 nsec, is a standard feature of CRAY X-MP Series computers. This clock allows the operating system to force interrupts to occur at a particular time or frequency.

Data structure

CRAY X-MP internal character representation is in ASCII with each 64-bit word able to accommodate eight characters.

All integer arithmetic is performed in 24-bit or 64-bit 2's complement mode. Floating-point numbers (64-bit quantities) consist of a signed magnitude binary coefficient and a biased exponent. The unbiased exponent range is:

$$2^{-20000}_8 \text{ to } 2^{+17777}_8,$$

or approximately

$$10^{-2466} \text{ to } 10^{+2466}$$

An exponent greater than or equal to 2^{+20000}_8 is recognized as an overflow condition and causes an interrupt if floating point interrupts are enabled.

Functional units

Instructions other than simple transmit or control operations are performed by hardware elements known as functional units. Each functional unit specializes in implementing algorithms for a specific portion of the instruction set and operates independently of the other units. A functional unit performs its operation in a fixed time called the functional unit time. No delays are possible once the operands have been delivered to a functional unit.

All functional units have one-clock-period segmentation. As a result, information arriving at or moving within the unit is captured and held in a new set of functional unit registers at the end of every clock period. New pairs of operands can thus enter the functional unit each clock period even though the unit may require more than one clock period to complete the calculation.

Functional units can operate concurrently so that, in addition to the benefits of pipelining (each unit can be driven at a result rate of one per clock period), there is also parallelism across the units.

The functional units can be thought of as forming four groups: address, scalar, vector and floating-point (see next page). The first three groups act in conjunction with one of the three primary register types to support address, scalar and vector modes of processing. The fourth group, floating-point, can support either

scalar or vector operations and accepts operands from or delivers results to scalar or vector registers accordingly.

The exchange sequence

Instruction issue is terminated by the hardware upon detection of an interrupt condition. All memory bank and functional unit activity is allowed to finish. To switch execution in order to handle the interrupt, the CRAY X-MP executes an exchange sequence. This causes program parameters for the next program to be exchanged with current information in the operating registers. Each program in the system has associated with it a 16-word block called an exchange package containing the parameters used in its execution sequence. Only the address and scalar registers are maintained in a program's exchange package.

Exchange sequences may be initiated automatically upon occurrence of an interrupt condition or may be initiated voluntarily by the software.

CPU intercommunication section

The CRAY X-MP CPU intercommunication section, present on CRAY X-MP multiprocessor systems, comprises five (CRAY X-MP/4) or three (CRAY X-MP/2) clusters of

CRAY X-MP CPU functional units

	Register usage	Time in clock periods
Address functional units		
Addition	A	2
Multiplication	A	4
Scalar functional units		
Addition	S	3
Shift-single	S	2
Shift-double	S	3
Logical	S	1
Population, parity and leading zero	S	3 or 4
Vector functional units		
Addition	V	3
Shift	V	3 or 4
Full vector logical	V	2
Second vector logical	V	4
Population, parity	V	5
Floating-point functional units		
Addition	S and V	6
Multiplication	S and V	7
Reciprocal approximation	S and V	14

shared registers for interprocessor communication and synchronization. Each cluster of shared registers consists of eight 24-bit shared address (SB) registers, eight 64-bit shared scalar (ST) registers and thirty-two one-bit synchronization (SM) registers.

Under operating system control, a cluster may be allocated to zero, one, two, three or four processors, depending on system configuration. The cluster may be accessed by any processor to which it is allocated in either user or system (monitor) mode. Any processor in monitor

mode can interrupt any other, and cause it to switch from user to monitor mode. Additionally, each processor in a cluster can asynchronously perform scalar or vector operations dictated by user programs. The hardware also provides built-in detection of system deadlock within the cluster.

Real-time clock

Programs can be precisely timed with a 64-bit real-time clock shared by the processors that increments once each 9.5 nsec.

CPU control section

Each CRAY X-MP CPU contains its own control section. Within each of these are four instruction buffers,

each with 128 16-bit instruction parcels, twice the capacity of the CRAY-1 instruction buffer. The instruction buffers of each CPU are loaded from memory at the burst rate of eight words per clock period.

The contents of the exchange package are augmented to include cluster and processor numbers. Increased data protection is also made possible through a separate memory field for user programs and data. Exchange sequences occur at the rate of two words per clock period on the CRAY X-MP.

Central memory

CRAY X-MP central memory can be one, two, four, eight or 16 million words (depending on model). A Cray word is composed of 64 data bits and eight check bits. Central memory is shared by the CPUs on multiprocessor systems and is arranged in interleaved banks. The interleaved memory banks enable extremely high transfer rates through the I/O Section and provide low read/write times for vector processing. All banks can be accessed independently and in parallel during each machine clock period. Based on a 9.5 nsec clock period, bank cycle time is 38 nsec on CRAY X-MP/4 computers and 76 nsec on CRAY X-MP/1 and X-MP/2 MOS memory models. The table on page 18 indicates memory size and banking arrangements for X-MP computers.

Each processor of the X-MP product line has four parallel memory ports; three for vector and scalar operations and one for I/O. The multiport memory has built-in conflict resolution hardware to minimize delays and maintain the integrity of all memory references to the same bank at the same time.

All CRAY X-MP models provide a flexible hardware chaining mechanism for vector processing. This feature enables a result vector to be used at any time as an operand in a succeeding operation. Also, vector chaining to and from memory is possible.

Consider the vector triad operation

$$A(I) = B(I) + S * C(I)$$

where S is a scalar, B and C are two input vectors, and A is the output

Model	Channel type		
	1000-Mbyte	100-Mbyte	6-Mbyte
CRAY X-MP/4	2	4	4
CRAY X-MP/2	1	2	4
CRAY X-MP/1	1	1 or 2	2 or 4

vector. The multiple memory access ports on X-MP systems enable two operands to be read and one to be written simultaneously. Thus, the reads of B and C, the multiply, the add and the write into A will all chain together and execute in parallel. In general, the CRAY X-MP enables memory block transfers to the B, T and V registers in parallel with vector arithmetic operations.

In addition, the CRAY X-MP provides hardware support for vector conditionals. Gather/scatter operations (chainable from other vector memory fetches and stores) and compressed-index generation facilitate and speedup execution of various conditional vector operations realized from ordinary user programs. All CRAY X-MP computers allow execution of two vector logical operations of the same type at the same time.

Input/output section

The I/O section of the CRAY X-MP may be equipped with a variety of high-performance channels for communicating with the mainframe, the I/O Subsystem and the Solid-state Storage Device. The latter two devices are high-speed data transfer devices designed to support CRAY X-MP processing speeds.

CRAY X-MP computers support three channel types identified by their maximum transfer rates as 6-Mbyte, 100-Mbyte and 1000-Mbyte channels. The table above indicates channel support capabilities on CRAY X-MP systems.

Input/output

I/O Subsystem

The power of the CRAY X-MP is enhanced by the I/O Subsystem (IOS). The IOS, with its multiple I/O processors (IOPs), acts as a data concentrator and data distribution point for the CRAY X-MP mainframe. A minimum of two IOPs is configured on X-MP/1 and X-MP/2 systems, and four IOPs are standard on the X-MP/4. A maximum of four IOPs is possible on all CRAY X-MP Computer Systems. The IOS handles I/O for a variety of front-end computer systems and for peripherals such as disk units and magnetic tape units. A direct-access path is also available between the IOS and the SSD.

One I/O processor is always designated as a master processor and is used for communication with all front-end computer systems and for controlling maintenance peripherals. Typically, one or two I/O processors can be used for controlling disk storage units. IOPs are linked to central memory via one or two 100-Mbyte channels.

When there are three or four I/O processors in an IOS, one can be designated for block multiplexer control. The block multiplexer IOP supports many concurrent data streams, and up to 48 tape units at a time may be configured and active. The tape units supported are IBM-compatible 9-track, 200 IPS, 1600/6250 BPI devices and IBM Series 3480 tape cartridge subsystems. They are connected to the IOP by one to eight block multiplexer channels.

IOS buffer memory is a separate independent storage unit composed of 8, 32 or 64 Mbytes of MOS integrated circuits. For an X-MP/4, buffer memory must be 32 Mbytes or larger. The IOPs connect to buffer memory through 100-Mbyte ports. Buffer memory is SECDED-protected and is field-upgradable.

The I/O Subsystem IOPs, buffer memory and controllers are mounted in four columns arranged in a 90° arc with power supplies hidden by benchlike extensions arranged around the outside of the base. This cabinet may be positioned up to 19 feet (5.8 meters) from the mainframe.

Solid-state Storage Device

The Solid-state Storage Device is available in sizes of 256, 512 or 1024 Mbytes of on-line storage; memory is made of MOS semiconductors and is fully field-upgradable. The SSD is used as an exceptionally fast-access disk device. Datasets are identical to those on disk storage, providing portability and flexibility. Storage on the SSD is allocated as with disk storage; just one job control language statement is required for each dataset assigned to the SSD. Software features allow for SSD resource management including automatic overflow to disk, if required. User data access time can be as little as 25 microseconds.

On the CRAY X-MP/4, the SSD is connected to the mainframe through two 1000-Mbyte channels, while on the X-MP/1 and X-MP/2 systems, this connection is via one 1000-Mbyte channel. SSD memory is fully equipped with SECDED logic.

The SSD cabinet closely resembles the IOS. It is made of four vertical columns arranged in a 90° arc mounted in a bench-like base.

DD-39 and DD-49 disk drives

Cray's very high performance DD-39 and DD-49 magnetic storage disk drives support the data capacity and transfer speed requirements of the largest CRAY X-MP computers. The DD-39 has a capacity of 1200 Mbytes and can sustain a data transfer rate of 5.9 Mbyte/sec. The DD-49 also has a capacity of 1200 Mbytes, however, it can sustain a rate of 9.8 Mbyte/sec.

Up to 32 disk drives may be configured on a CRAY X-MP system. A combination of DD-39 and DD-49 drives may be configured on the same system.

Front-end interfaces

CRAY X-MP computers are interfaced to front-end computer systems through the I/O Subsystem. Up to seven front-end interfaces, identical to those used in the CRAY-1, can be accommodated. Users may also elect to supply Network Systems Corporation (NSC) channel adapters in place of one of the front-end interfaces, thus enabling interfacing to many systems. The hardware connection between CRAY X-MP systems and Apollo workstations is via NSC HYPERchannel.

Cray Research currently provides front-end interface support for IBM, CDC, DEC, Sperry and Honeywell systems. Front-end interfaces compensate for differences in channel widths, word size, logic levels and control protocols between other manufacturers' equipment and the CRAY X-MP.



The CRAY X-MP Series of Computers — a family of supercomputers that offers flexibility for the broad and growing range of science and engineering computational needs at all levels.

State-of-the-art technology, outstanding price/performance, flexible and balanced system design and a commitment to customer support with the resources to provide it — these are the reasons that Cray Research computer systems remain the large-scale computational tool of choice the world over.

The equipment specifications contained in this brochure and the availability of said equipment are subject to change without notice. For the latest information, contact your local Cray Research sales office.

CRAY-1 and SSD are registered trademarks and CRAY X-MP is a trademark of Cray Research, Inc.

HYPERchannel is a registered trademark of Network Systems Corporation.

Apollo and DOMAIN are registered trademarks of Apollo Computer, Inc.

Sun Workstation is a registered trademark and Sun Microsystems is a trademark of Sun Microsystems, Inc.

UNIX is a trademark of AT&T Bell Laboratories.



Corporate Headquarters
608 Second Avenue South
Minneapolis, MN 55402
612/333-5889

Domestic sales offices

Albuquerque, New Mexico
Atlanta, Georgia
Beltsville, Maryland
Boston, Massachusetts
Boulder, Colorado
Chicago, Illinois
Dallas, Texas
Detroit, Michigan
Houston, Texas
Laurel, Maryland
Los Angeles, California
Minneapolis, Minnesota
Pittsburgh, Pennsylvania
Pleasanton, California
Rochester, New York
Seattle, Washington
Tampa, Florida
Tulsa, Oklahoma

International subsidiaries

Cray Research (UK) Limited
Bracknell, Berkshire, U.K.

Cray Research GmbH
Munich, West Germany

Cray Research France S.A.
Paris, France

Cray Research Japan, Limited
Tokyo, Japan

Cray Research S.R.L.
Milan, Italy

Cray Canada Inc.
Toronto, Canada

The CRAY X-MP Series of Computers

Cray Research, Inc.



Introducing the CRAY X-MP Series of Computer Systems

Now, Cray Research announces an answer to your expanding computational needs — the CRAY X-MP Series of Computer Systems. The CRAY X-MP Series, with its major innovations in architecture and technology, offers overall system throughput up to five times that of a CRAY-1 S/1000 CPU, and a maximum burst rate up to eight times that of the CRAY-1 for specific cases. At the same time, software compatibility has been maintained between the CRAY X-MP and the CRAY-1 to protect user software investment.

The CRAY X-MP is a powerful multiprocessor system. The mainframe features two identical Central Processing Units (CPUs) and a multipoint memory. The dual CPUs allow for both multiprocessor jobs and concurrent independent uniprocessor jobs while sharing a two- or four-million word bipolar Central Memory. Four parallel memory access ports per processor provide over eight times the total usable memory bandwidth of the CRAY-1.

The CRAY X-MP Computer System, with its 9.5 nsec clock cycle time, is the fastest general-purpose computer system commercially available today. The X-MP is capable of an overall instruction issue rate of over 200 million instructions per second (MIPS). Computation rates of over 400 million 64-bit floating point operations (MFLOPS) are possible, and combined arithmetic/logical operations can exceed 1000 million operations per second (MOPS).

A new high-performance peripheral device has also been developed for use on the CRAY X-MP and S Series systems. The new Solid-state Storage Device (SSD), with its exceptionally high transfer rates, can be used as a fast-access disk device for large datasets generated and manipulated repetitively by user programs. It can also be used by the system for temporary storage of system programs. The SSD is available with 64, 128, or 256 million bytes of storage. Complementing the SSD and enabling its high performance is a broadband channel capable of a maximum burst transfer rate of 10 gigabits per second. Performance improvement factors of 50 to 100 are anticipated over disk units for short random or long sequential transfers. Transferring a million word dataset requires only about 50 milliseconds, including system overhead.

The I/O Subsystem, which is an integral part of the CRAY X-MP System, also contributes to the new system's outstanding performance. The I/O Subsystem offers parallel disk drive capabilities, I/O buffering for disk-resident and Buffer Memory-resident datasets, on-line tape handling, and efficient front-end system communication. Up to 8 million words of Buffer Memory can be configured on the I/O Subsystem, enabling faster and more efficient data access and processing by the CPUs.

The CRAY X-MP... a major new computational resource available now. In the future, Cray Research will emphasize development of multiprocessing architecture as an important technique for increasing processing power.

CRAY X-MP hardware features

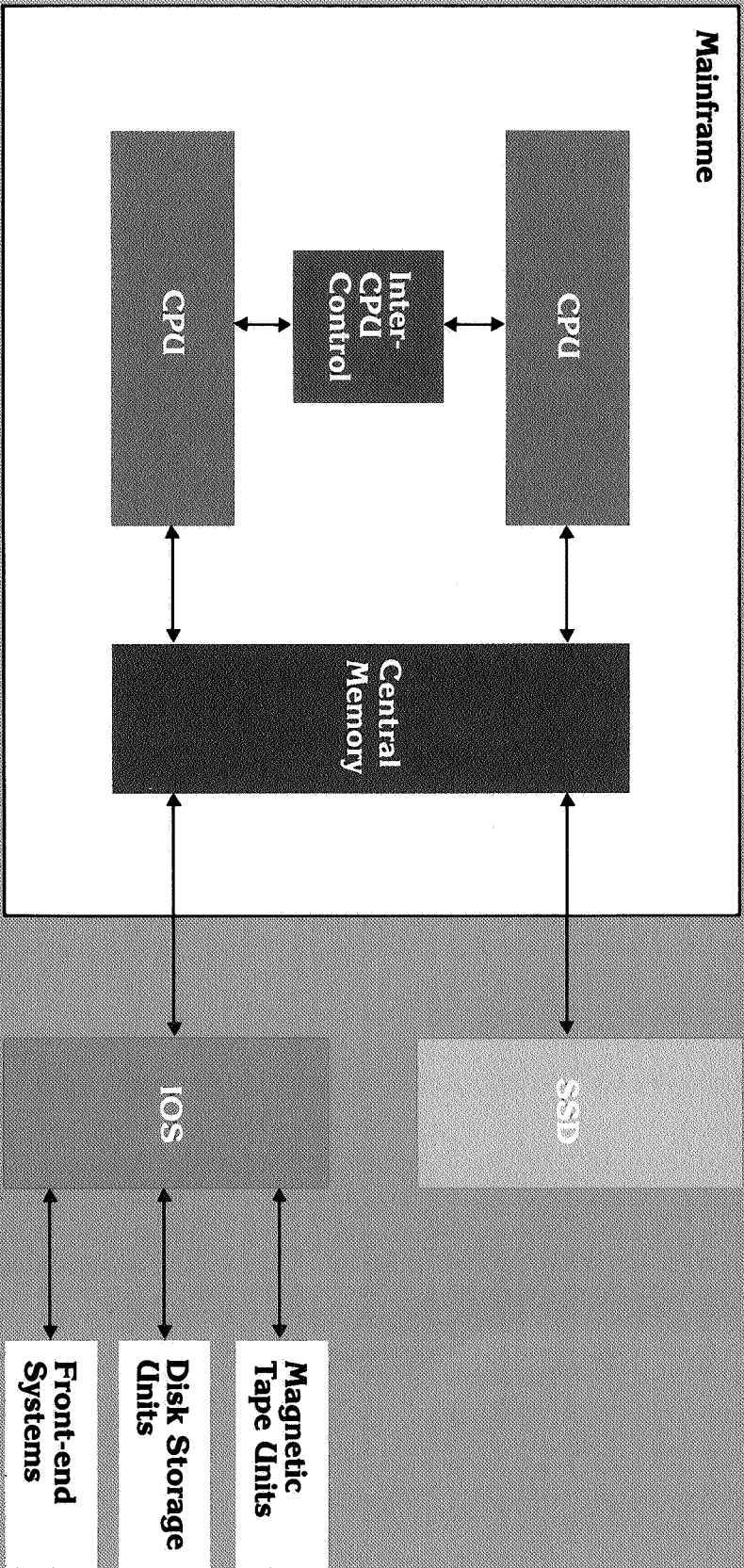
Throughout the CRAY X-MP CPUs, 16-gate array integrated circuits are used. These circuits, which are faster and denser than the circuitry used in the CRAY-1, contribute to a clock cycle time of 9.5 nanoseconds and a memory bank cycle time of 38 nanoseconds. Proven cooling and packaging techniques have also been used on the CRAY X-MP to ensure high system reliability.

The CRAY X-MP's four parallel memory access ports per processor, combined with the improved clock cycle time, means that the CRAY X-MP has more than eight times the total usable memory bandwidth of the CRAY-1.

The high performance of the CRAY X-MP is evident in both scalar and vector modes. Scalar performance is improved through the faster clock, short memory access time, and larger instruction buffers, while vector performance is improved through a combination of faster clock, parallel memory ports and hardware automatic 'chaining' features. These new features allow simultaneous memory fetch, arithmetic, and memory store operations in a series of related vector instructions. Either long or short vector operations, characterized by heavy register usage or heavy memory references, use these features to advantage.

CRAY X-MP software features

The innovative hardware features of the CRAY X-MP are supported by the standard Cray Research software. The CRAY Operating System (COS) supports concurrent independent uniprocessor jobs and multiprocessing of a single job. New techniques extending the multiprocessing capabilities of the CRAY FORTRAN Compiler (CFT) are also being explored.



CRAY X-MP Overall System Organization

The CRAY X-MP Models 22 and 24

The CRAY X-MP Models 22 and 24 are composed of the following basic hardware components.

Two Central Processing Units (CPUs), each with:

- 13 functional units
- A, B, S, T, and V operational registers as in the CRAY-1
- 4 instruction buffers
- 4 concurrent memory ports

Central Memory composed of:

- Either 2M or 4M 64-bit words arranged in 16 or 32 banks, respectively, in 12 columns

Input/output channel configuration featuring the following:

- 4 6-Mbyte/sec I/O control channels
- 2 100-Mbyte/sec channels for transferring data between the IOS and Central Memory
- 1 1250-Mbyte/sec channel for transferring data between the SSD and Central Memory

An I/O Subsystem (IOS) identical with that for a CRAY-1 S

Series System composed of:

- 2, 3, or 4 high-speed I/O Processors
- 8M, 32M, or 64M bytes of I/O Buffer Memory
- 1 to 8 DCL-4 Disk Control Units
- 2 to 48 DD-29 Disk Storage Units (32 if Block Multiplexer Channel Controller is also configured)
- 1 to 4 BMC-4 Block Multiplexer Channel Controllers
- 1 to 16 Block Multiplexer Channels, which can support user-supplied on-line magnetic tape units
- Operator consoles

An optional Solid-state Storage Device (SSD) with:

- 64M, 128M, or 256M bytes of memory arranged in 16, 32, or 64 banks

A Peripheral Expander providing maintenance functions

Power and cooling equipment

One standard, two optional front-end interfaces

Model	X-MP/22	X-MP/24				
Mainframe						
CPU's	2	2				
Bipolar memory (64-bit words)	2M	4M				
6 Mbyte/sec channels	4	4				
100 Mbyte/sec channels	2	2				
1250 Mbyte/sec channels	1	1				
I/O Subsystem						
I/O Processors	2	3	4			
DCU-4 Disk Control Units	1-4	1-8	1-12	1-4	1-8	1-12
DB-29 Disk Storage Units	2-16	2-32*	2-48*	2-16	2-32*	2-48*
Block Multiplier Channel Controllers	1-4	1-4	1-4	1-4	1-4	1-4
Block Multiplier Channels	1-16	1-16	1-16	1-16	1-16	1-16
Front-end Interfaces	1-3	1-3	1-3	1-3	1-3	1-3
Buffer Memory Size (bytes)	8M, 32M, or 64M					
Storage Storage Devices**	64M, 128M, or 256M					
Memory Size (bytes)	64M, 128M, or 256M					

*16 fewer Disk Storage Units can be configured if Block Multiplier Channel Controllers are configured.
 **Optionally one per CRAV X-MP.

The CRAY X-MP is a powerful computer system ideal for execution of multiprocessor jobs and concurrent independent uniprocessor jobs. With its advanced design and improved performance, the CRAY X-MP offers:

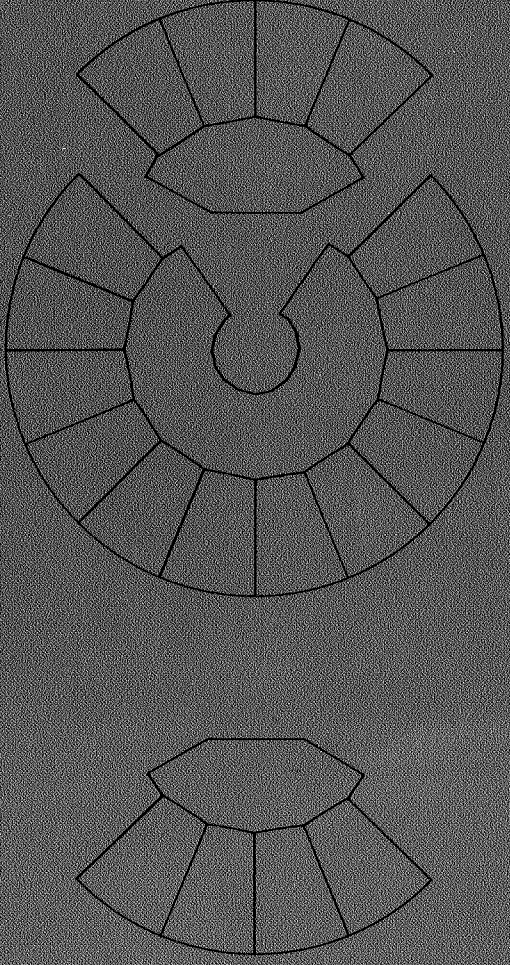
- Overall system throughput up to five times that of a CRAY-1 S/1000 CPU on many jobs, with a maximum burst rate up to eight times that of the CRAY-1 for specific cases
- Two identical Central Processing units sharing a Central Memory of up to four million 64-bit words
- Four parallel memory access ports per processor providing over eight times the total usable memory bandwidth of the CRAY-1
- Four instruction buffers with a combined capacity of 512 16-bit instruction parcels, twice the capacity of those on the CRAY-1
- Operational registers and functional units that are among the features providing compatibility with the CRAY-1
- Hardware support for partitioning of memory fields into data and program areas
- The new high-performance Solid-state Storage Device (SSD) which, with its transfer rate of up to 10 gigabits/second, can be used as an exceptionally fast-access disk device
- An integral I/O Subsystem that efficiently performs input/output functions between the mainframe, peripheral devices, and the front-end systems and has a sustained transfer rate of 40 Mbytes/second between the mainframe and the I/O Subsystem.
- Software that takes advantage of the unique CRAY X-MP hardware features while remaining compatible with that of the CRAY-1
- Compact size—just 100 square feet of floor space required for the mainframe
- Proven component and cooling technologies designed for high reliability

- Two Central Processing Units
- I/O Subsystem
- Central Memory with either 2M or 4M 64-bit words
- Field upgradability
- One 1250 Mbyte channel for SSD linkage
- Two 100 Mbyte channels for I/O Subsystem data transfers
- Four 6 Mbyte channels for I/O control
- Up to 16 Block Multiplier Channels
- Up to 48 DD-29 drives each with a 600 Mbyte capacity
- Optional Solid-state Storage Device (SSD)

SSD

Mainframe

IOS



Software for the CRAY X-MP

The processing potential of the CRAY X-MP Series has stimulated the development of new system and user software techniques. Cray Research is committed to providing users with full and easy access to the power of the new CRAY X-MP.

Cray Research software will evolve through planned stages to support multiprocessing capabilities and improved user access to the unique features of the CRAY X-MP, thus achieving greater processing speed of user code. Software products, including the CRAY Operating System (COS), FORTRAN Compiler (CFT), Cray Assembler Language (CAL), and the associated libraries have been enhanced to take advantage of the CRAY X-MP hardware features. All CRAY-1 source code and almost all binary code is upward compatible with the CRAY X-MP. The only exception is binary code that uses the vector functional unit recursion capability available on the CRAY-1.

The CRAY Operating System (COS), by providing the same user interface to both the CRAY X-MP and CRAY-1, enables a smooth migration path to the higher capacity CRAY X-MP systems. COS treats the multiple processors of the CRAY X-MP symmetrically, that is, COS and user code may execute on either processor.

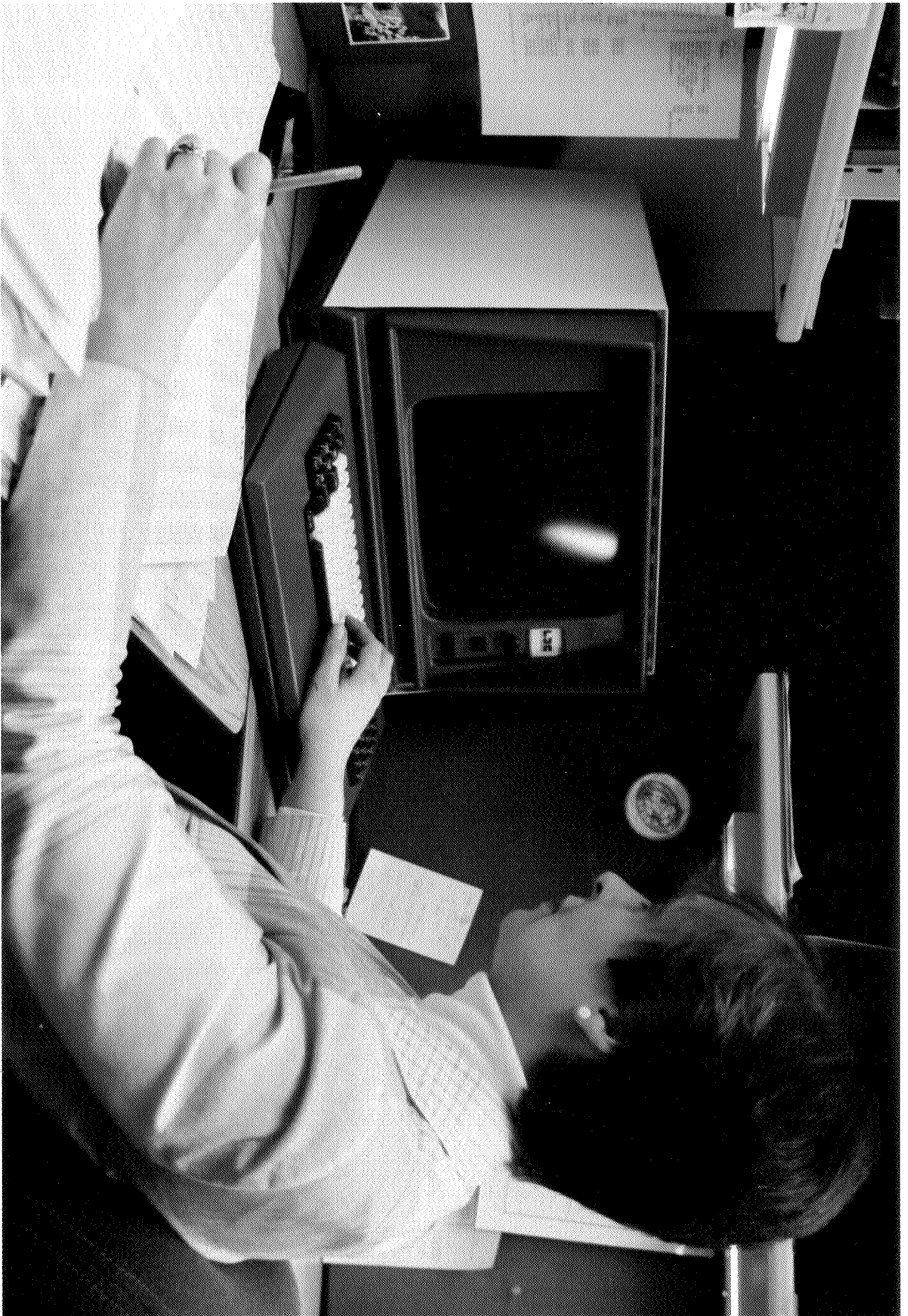
The CFT library allows user partitioning of an application into concurrently executed tasks. Techniques are also being explored for automatic compiler partitioning of a program for multiprocessing.

Special multiprocessor communications and control instructions are also available through the CRAY Assembler Language (CAL) and enable maximum exploitation of the hardware features of the CRAY X-MP.

New software also supports the Solid-state Storage Device (SSD) and I/O Subsystem Buffer Memory so that to users, SSD and Buffer Memory appear like disks. That is, temporary datasets, employed by user jobs, may reside wholly or partially within the SSD or IOS Buffer Memory, resulting in significant reductions of I/O wait time. Use of SSD or Buffer Memory resident datasets does not require changes to user code or to the subroutine libraries; all logical I/O requests are device-independent.

Software Summary

- CFT, a vectorizing and optimizing ANSI '77 FORTRAN compiler
- The FORTRAN subroutine library
- A scientific subroutine library
- A Cray Applications Software Library of public domain software offered as a service
- The Cray Assembly Language (CAL), providing access to all hardware capabilities
- COS, a multiprogramming and multiprocessing operating system
- A variety of system utility programs
- Interface software for IBM MVS operating system and CDC NOS and NOS/BE operating systems offered as a service
- SSD resident and IOS Buffer Memory resident datasets



The Design of the CRAY X-MP

The CRAY X-MP Mainframe

At the heart of the CRAY X-MP Series mainframe are two identical CPUs, each of which, through various hardware enhancements, is even more powerful than the CRAY-1 uniprocessor. Synchronization of the processors is achieved through clusters of shared registers in the CPU Intercommunication Section and through shared Central Memory.

The CRAY X-MP mainframe is a composite of several key hardware features: a 9.5 nsec clock, two CPUs each with its own computation and control sections, a CPU intercommunication section, a single multipoint, bipolar Central Memory, and an I/O section.

The elegant and compact CRAY X-MP mainframe consists of 12 vertical columns arranged in a 270° arc; each column houses two chassis holding up to 72 modules. Power supplies and cooling are clustered around the base and extend outward to provide seating for maintenance personnel.

Physical Characteristics of the CRAY X-MP Mainframe

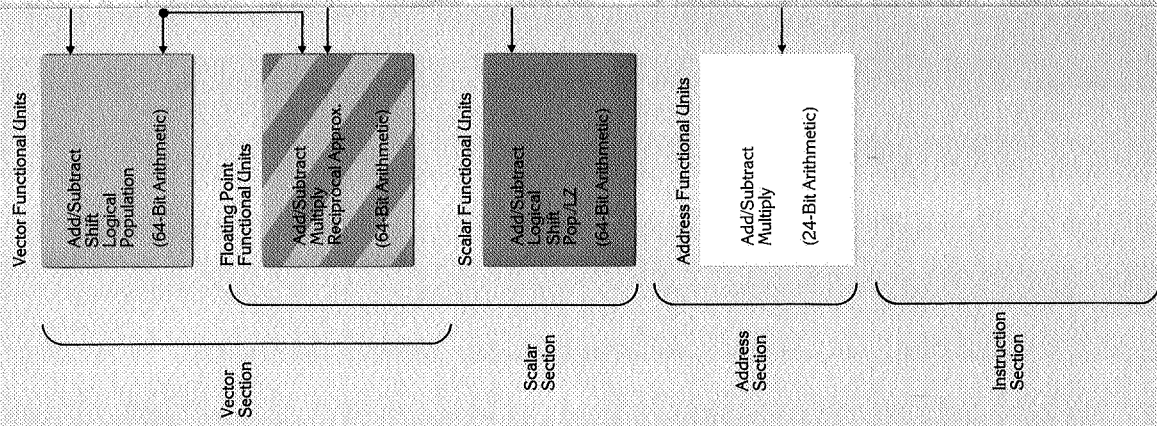
- 100 square feet of floor space for mainframe
- 5.25 tons mainframe weight
- 9.5 nanosecond clock period
- Liquid refrigerant cooling
- 400 Hz power from motor generators

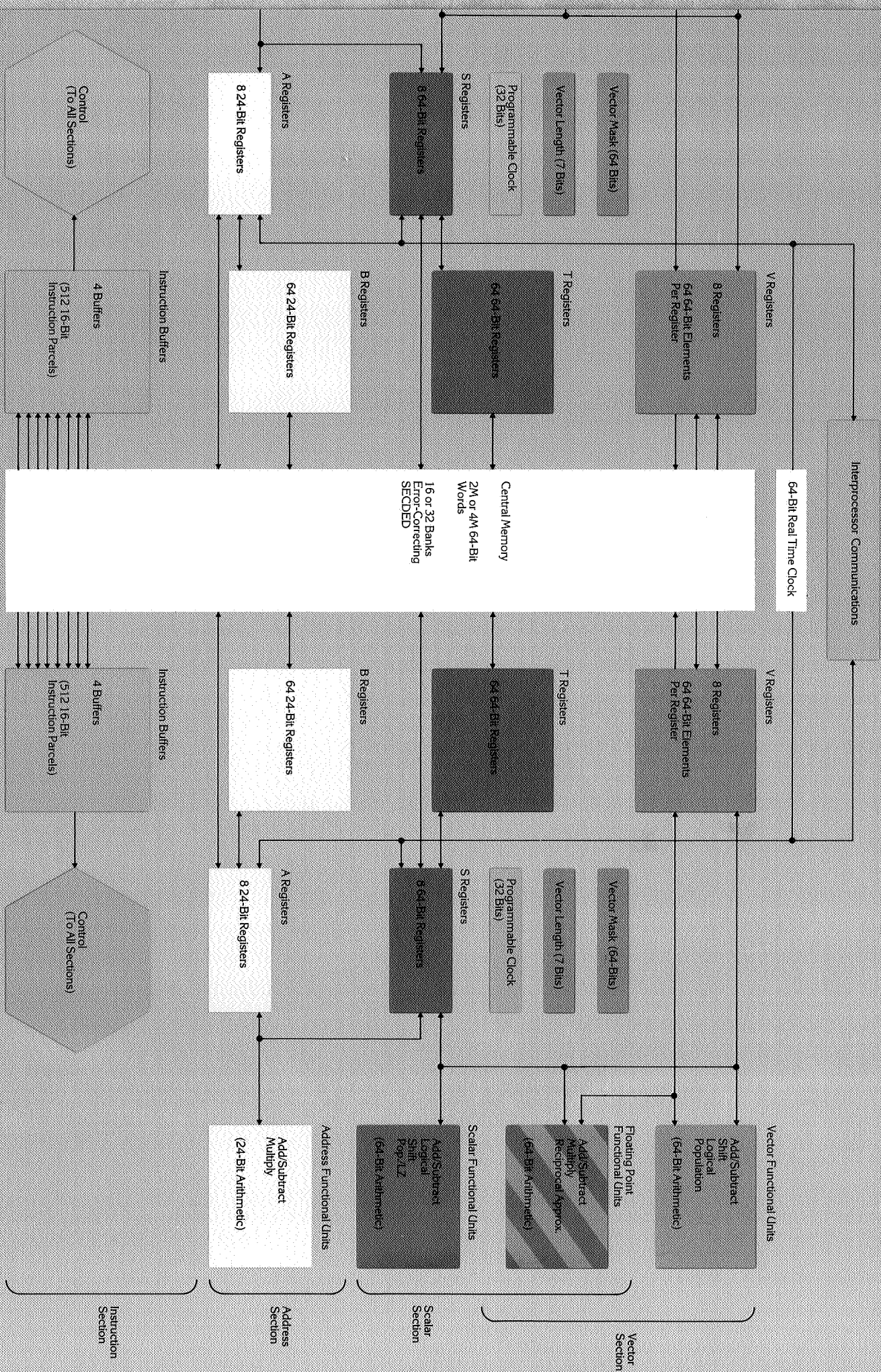
CPU Computation Section

All CRAY X-MP arithmetic operations are bit-compatible with those on the CRAY-1. In fact, the CRAY X-MP's CPUs feature the same set of registers and functional units as the CRAY-1 uniprocessor. However, as explained later, the multiple memory ports available to each CPU in the CRAY X-MP greatly increase the speeds of memory-to-register and register-to-memory transfers over those of the CRAY-1.

CPU Computation Section Summary

- Integer and floating-point arithmetic
- 2's complement integer arithmetic
- Signed magnitude floating-point arithmetic
- Address, scalar, and vector processing modes
- 13 functional units:
 - Vector add/subtract
 - Vector shift
 - Vector logical
 - Vector population count
 - Floating point add/subtract
 - Floating point multiply
 - Floating point reciprocal approximation
 - Scalar add/subtract
 - Scalar shift
 - Scalar logical
 - Scalar population and leading zero
 - Address add/subtract
 - Address multiply
- Eight 24-bit address (A) registers
- Sixty-four 24-bit intermediate address (B) registers
- Eight 64-bit scalar (S) registers
- Sixty-four 64-bit intermediate scalar (T) registers
- Eight 64-element vector (V) registers with 64 bits per element
- 32-bit Programmable clock





CPU Control Section

Each CRAY X-MP CPU contains its own control section. Within each of these are four instruction buffers, each with 128 16-bit instruction parcels, twice the capacity of the CRAY-1 instruction buffers. The instruction buffers of each CPU are loaded from memory at the burst rate of 8 words per clock period. The contents of the exchange package have been augmented to include cluster number and processor number. Increased protection of data is also made possible through a separate memory field for user programs and data.

Exchange sequences occur at the rate of two words per clock period on the CRAY X-MP.

CPU Control Section Summary

- Four instruction buffers, each holding 128 16-bit instruction parcels
- 128 basic instruction codes
- Exchange sequence mechanism
- Instruction buffers loaded at 8 words per clock period
- Normal and interprocessor interrupt handling
- Separate program and data field protection in memory

CPU Intercommunication Section

The CRAY X-MP CPU intercommunication section comprises three clusters of shared registers for interprocessor communication and synchronization. Each cluster of shared registers consists of eight 24-bit shared address (SB) registers, eight 64-bit shared scalar (ST) registers, and thirty-two 1-bit synchronization (SM) registers. Under operating system control, a cluster may be allocated to both, either, or none of the processors. The cluster may be accessed by any processor to which it is allocated in either user or system mode.

A 64-bit real-time clock is shared by the processors.

CPU Intercommunication Section Summary

- 3 clusters of intercommunications registers, each with:
 - 8 24-bit shared address (SB) registers
 - 8 64-bit shared scalar (ST) registers
 - 32 1-bit synchronization (SM) registers
- A 64-bit real-time clock

Central memory

The CRAY X-MP processors share a single bipolar Central Memory of 2M or 4M 64-bit words that supports the requirements of large-scale applications. Memory is arranged in 32 banks for 4 million word systems and in 16 banks for 2 million word systems. These interleaved memory banks enable extremely high transfer rates through the I/O section and provide low read/write times for vector processing. Finally, the short bank cycle time (38 nanoseconds) is well-suited to high-performance scalar and vector applications. A major feature of the CRAY X-MP is its four parallel memory access ports per processor, which include two ports for vector reads, one for vector writes, and one for I/O. This notable hardware enhancement provides the CRAY X-MP with over eight times the memory bandwidth of the CRAY-1.

The CRAY X-MP hardware also provides a flexible hardware chaining mechanism for vector processing. This feature enables a result vector to be used at any time as an operand in a succeeding operation. Also, vector chaining to memory as well as from memory is now possible.

Consider the vector triad operation

$$A(I) = B(I) + S * C(I)$$

where S is a scalar, B and C are two input vectors, and A is the output vector. The CRAY X-MP's multiple memory access ports enable two operands to be read and one to be written simultaneously. In general, the CRAY X-MP enables memory block transfers to the B, T, and V registers in parallel with vector arithmetic operations.

I/O transfers occur at a 2-word-per-clock-period rate, concurrent with CPU memory activities.

Central Memory Summary

- 2M or 4M words of bipolar IC memory arranged in 16 or 32 banks, respectively
- Shared access from the two CRAY X-MP processors
- 4 clock periods (38 nanoseconds) bank cycle time
- 4 memory access ports per CPU
- 64 data bits and 8 error correction bits per word
- Single-bit error correction, double-bit error detection (SECEDED)

Memory Transfer Rates

Source/Destination	Words per clock period	Total maximum system transfer rate (Mbits/sec)
B, T, V	6	40,420
A, S	1	6,730
Instruction buffers	8	53,890
I/O	2	13,470

I/O Section

The I/O Section of the CRAY X-MP mainframe, shared by the two CPUs, may be equipped with a variety of high-performance channels for communicating with the mainframe, the I/O Subsystem, and a Solid-state Storage Device (SSD). The CRAY X-MP supports three channel types identified by their maximum transfer rates as 6 Mbytes/sec, 100 Mbytes/sec, and 1250 Mbytes/sec.

Four 6-Mbyte/sec channels are available for communication with the mainframe. In addition, two 100-Mbyte/sec channels are provided. At least one of the 100-Mbyte/sec channels and one of the 6-Mbyte/sec channels must be connected to the I/O Subsystem. The I/O Section is also equipped with a single 1250-Mbyte/sec SSD channel.

To increase CPU efficiency and encourage parallel I/O processing, no peripherals such as disk units are attached directly to the mainframe.

I/O Section Summary

- Four 6-Mbyte/sec channels for communication with the mainframe
 - 16 data bits, 3 control bits, and 4 parity bits
- Two 100-Mbyte/sec channels for data transmissions to/from the I/O Subsystem
 - 64 data bits, 3 control bits, and 8 check bits in each direction
- One 1250-Mbyte/sec channel for use with the SSD
 - 128 data bits and 16 check bits in each direction

I/O Subsystem

The power of the CRAY X-MP is enhanced by the I/O Subsystem (IOS). The IOS with its multiple I/O processors, acts as a data concentrator and data distribution point for the CRAY X-MP mainframe. It can handle I/O for a variety of front-end computer systems and for peripherals such as disk units and user-supplied magnetic tape units.

One of the four I/O processors is always designated as a master processor and is used for communication with all front-end computer systems and for controlling maintenance peripherals. One to three of the I/O processors can each be used for controlling 16 DD-29 Disk Storage Units. Each DD-29 has a capacity of 600 Mbytes. Either one or two of these can be connected to a 100 Mbyte/sec channel between disks and Central Memory. When there are three or more I/O processors in an IOS, one can be designated for block multiplexer control. This IOP supports up to 8 concurrent data streams and up to 64 configurable tape units, 32 of which may be active or assignable at a given time. The tape units supported are IBM-compatible 9-track, 200 IPS, 1600/6250 BPI devices.

The IOS Buffer Memory consists of 8M, 32M, or 64M bytes arranged in 8 or 16 banks, depending on size. It is equipped with single-bit error correction, double-bit error detection (SECDED). Buffer Memories can be upgraded in the field.

The I/O Subsystem is housed in a cabinet that complements the CPU cabinet. Modules comprising Buffer Memory, I/O Processors, and controllers are mounted in four columns arranged in a 90° arc.

I/O Subsystem Summary

- Two to four I/O Processors
- 12.5 nsec clock period
- 8, 32, or 64 Mbytes of Buffer Memory
- Up to 48 600 Mbyte disk storage units
- Optional Block Multiplexer Channels for user supplied tape units
- One to three Cray Research Front-End Interfaces or user-supplied Network Systems HYPERchannel Adapters
- Operator consoles
- A Peripheral Expander and associated maintenance peripherals
- 1.5 tons weight
- 10 square feet of floor space
- Liquid refrigerant cooling
- 400 Hz power from motor generators

Solid-state Storage Device (SSD)

Complementing the CRAY X-MP and designed with its demanding throughput requirements in mind, is the new Cray Research Solid-state Storage Device (SSD). The SSD is available in sizes of 64, 128, or 256 million bytes of on-line storage.

Memories are fully field-upgradable from the smallest to the largest sizes offered. Using the latest memory chip technology, the SSD greatly reduces the access and transfer times over that for conventional rotational storage devices.

The SSD connects to the CRAY X-MP mainframe through the specially designed 1250 Mbyte/sec channel so that theoretically the hardware can transfer 8 million bytes of data in 8 milliseconds between the SSD and the mainframe.

The SSD cabinet closely resembles the cabinet of the I/O Subsystem. Similar design to that of the mainframe is used in the power supplies and the liquid refrigerant cooling system. Depending on existing capacities, a site may require additional power and cooling equipment.

Modules are arranged in 16 banks for 64 Mbyte systems, in 32 banks for 128 Mbyte systems and in 64 banks for 256 Mbyte systems. Transfer block sizes are a minimum of 64 words. The memory is fully equipped with single-bit error correction, double-bit error detection (SECDED) logic.

Solid-state Storage Device Summary

- 64 M, 128 M, or 256 M bytes arranged in 16, 32, or 64 banks
- Single-bit error correction, double-bit error detection (SECDED)
- 1250 Mbyte/sec maximum burst transfer rate, assuming 64 banks
- 1.5 tons weight
- 10 square feet of floor space
- Liquid refrigerant cooling
- 400 Hz power from motor generators

Interfaces to Front-end Computers

The CRAY X-MP is interfaced to front-end computer systems through the I/O Subsystem. Up to three front-end interfaces per I/O Subsystem, identical to those used in the CRAY-1, can be accommodated.

Front-end interfaces compensate for differences in channel widths, word size, logic levels, and control protocols, and are available for a variety of front-end systems.

Cray Research Front-End Interfaces

- CDC
- IBM
- Honeywell
- DEC
- Data General
- Systems Engineering Laboratories
- Univac

Users may also elect to supply a Network Systems NSC A130 Channel Adapter in place of one of the front-end interfaces.

Configurations

Flexibility in the choice of an initial configuration and the provision for upgradability to higher capacity systems are hallmarks of Cray Research's complete product family of CRAY X-MP and CRAY-1/S Computer Systems. The CRAY X-MP Series of Multiprocessor Computer Systems broadens the range of possible configurations.

The I/O Subsystem, which is a standard component of CRAY X-MP systems, can be configured in a variety of ways. In particular, the number of I/O Processors may vary from two to four and the amount of Buffer Memory from 8 Mbytes to 64 Mbytes.

Finally, the Solid-state Storage Device is offered for users with the requirement for mass memory of outstanding performance.

Upgradability is a key feature of the CRAY X-MP Series. In addition to upgrading to a maximum of four IOPs in an I/O Subsystem, Central Memory, I/O Subsystem Buffer Memory, and SSD memory are all field upgradable from the smallest to the largest sizes.

An SSD may easily be added to an installed CRAY X-MP system.

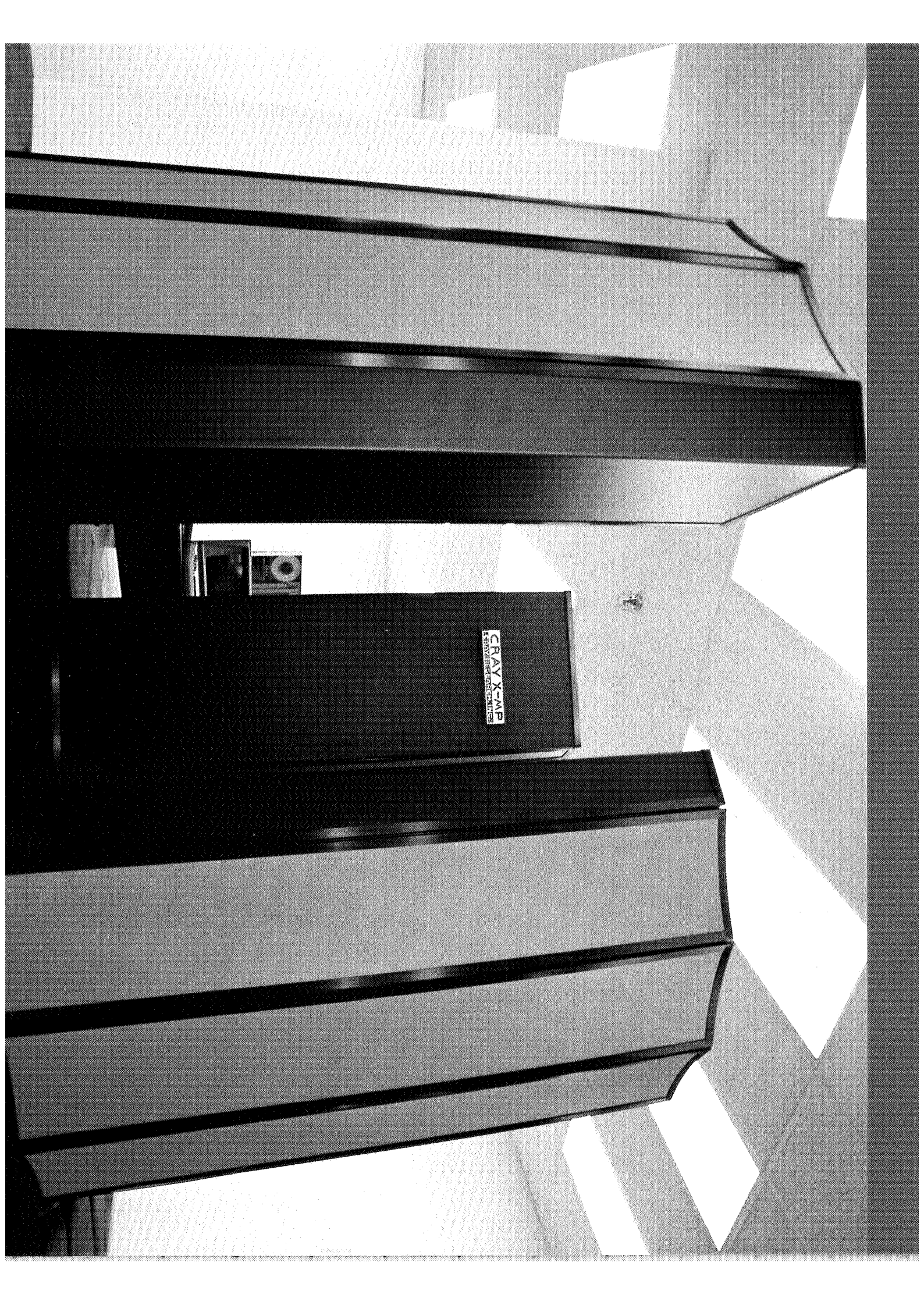
CRAY X-MP Maintenance

An extensive set of diagnostic programs is available to field engineers to aid in quickly identifying problem areas in the hardware in event of a failure. These diagnostics are accessed via operator consoles either locally or remotely attached to the I/O Subsystem for technical support.

Further onsite diagnosis to the component level occurs off-line from the mainframe via a sophisticated Cray Research module tester. This is consistent with the CRAY-1 maintenance philosophy of replacing and repairing modules onsite.

CRAY X-MP Reliability

The reliability of the CRAY X-MP, because of the reduced number of components and enhanced cooling system, will meet or exceed that of the CRAY-1, which is recognized as setting a standard in the industry.



CRAYX-MP
CRAY RESEARCH

Corporate Headquarters

CRAY RESEARCH, INC.
608 Second Avenue South
P.O. Box 154
Minneapolis, MN 55440

Sales Offices

Albuquerque, New Mexico
Atlanta, Georgia
Austin, Texas
Boston, Massachusetts
Boulder, Colorado
Chicago, Illinois
Dallas, Texas
Houston, Texas
Laurel, Maryland
Livermore, California
Los Angeles, California
Mountain View, California
Pittsburgh, Pennsylvania
Seattle, Washington
Silver Spring, Maryland

International Subsidiaries

Cray Research (UK) Limited
Wokingham, Berkshire, England
Cray Research GmbH
Munich, West Germany
Cray Research France, S.A.
Neully sur Seine, France
Cray Research Japan, Limited
Tokyo, Japan

Publications

Cray Research, Inc.
1440 Northland Dr.
Mendota Heights, MN 55120

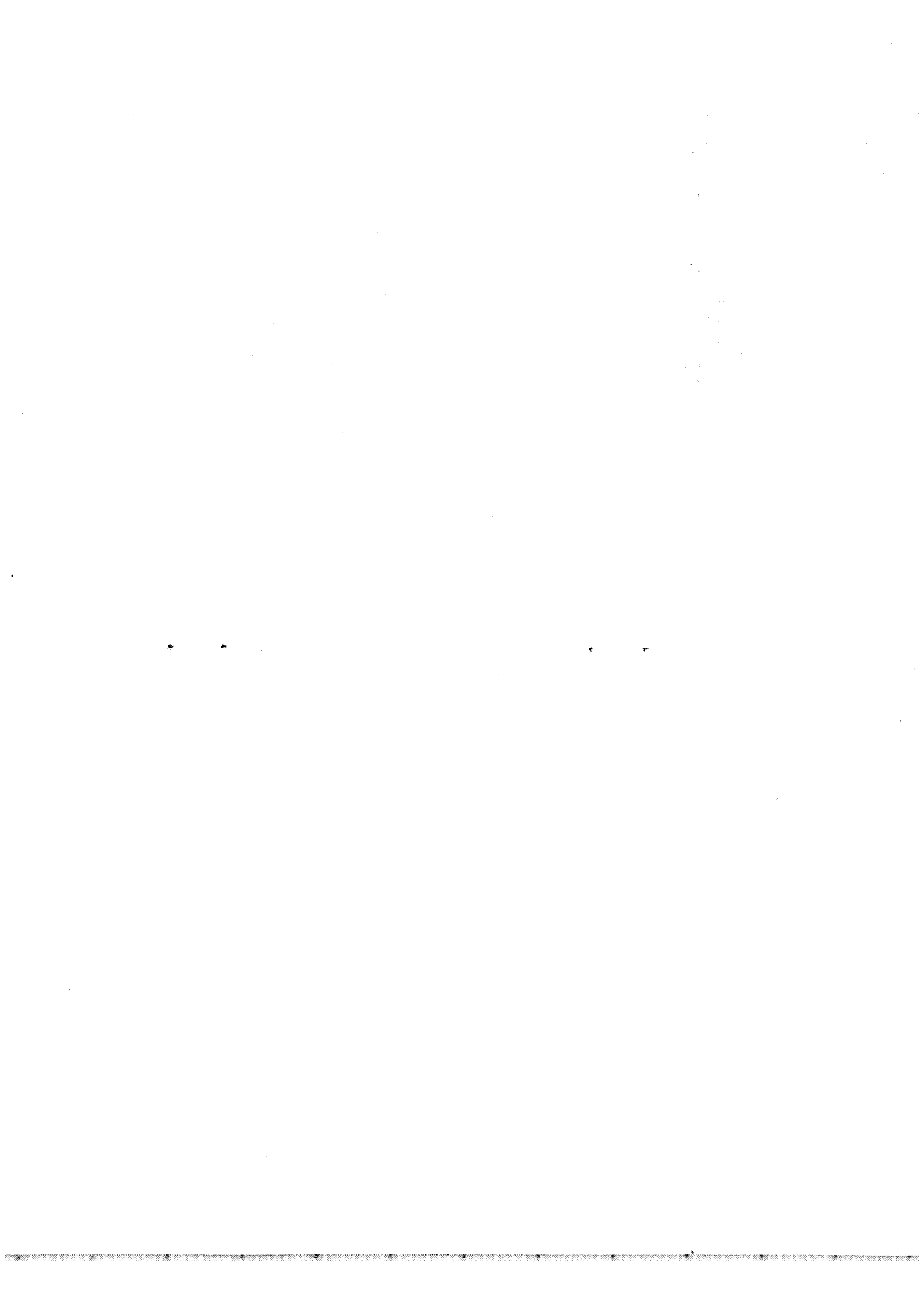
R.N.F.

**CRAY XMP:
A MULTIPROCESSOR SUPERCOMPUTER**

**Steve S. Chen
Christopher C. Hsiung
John L. Larson
Eugene R. Somdahl**

**Cray Research, Inc.
Chippewa Falls, WI**

Submitted for publication in Vector and Parallel Processors: Architecture, Applications, and Performance Evaluation, Myron Ginsberg, Editor, to be published by North Holland



CRAY XMP: A MULTIPROCESSOR SUPERCOMPUTER

Steve S. Chen
Christopher C. Hsiung
John L. Larson
Eugene R. Somdahl
Cray Research, Inc.
Chippewa Falls, Wisconsin

1. HISTORIC NOTE

Ever since its first delivery in 1976 to Los Alamos National Laboratory, the CRAY-1 computer has been the industry standard in very high speed computing. Intensive research and development work in various scientific fields has been made possible because of the use of supercomputers such as the CRAY-1. Among the fields that benefit from the use of CRAY computers are: aerodynamics, meteorology, climate modeling, seismology, reservoir simulation, cryptography, nuclear fusion research, nuclear power plant safety research, circuit design, structural analysis, particle physics, astronomy, animation, and human organ simulation.

The success of the CRAY-1 within the scientific computing community can be attributed to its innovative vector architecture, dense packaging, and advanced cooling technology. The CRAY-1 design employed many state of the art architectural features such as:

1. Pipelining in memory access and function units,
2. Utilization of vector registers and operations chaining,
3. Concurrent execution of multiple functional units,
4. Interleaved memory,
5. Instruction cache and lookahead,
6. Massive use of parallel logic to shorten the execution time of functional units.

The vector architecture introduced, at that time, a new era in high speed computing. The well balanced[1] and compact[2] design enhanced the performance of vector as well as scalar application codes.

There are many references available that discuss in great detail the architecture, physical characteristics and usage of the CRAY-1 computer[3,4,5].

In 1979, while Seymour Cray was leading the development of the CRAY-2, a separate effort, led by Steve Chen, and under the direction of Les Davis, was initiated within Cray Research to design a machine more powerful than the CRAY-1.

Several important decisions regarding design strategy were made.

To shorten the time for circuit design, the same 16-gate ECL gate arrays to be used in the CRAY-2 were chosen. However, other than components, there was virtually no similarity between the two design efforts.

Since the design requirements of the two projects were totally different, new electronic design rules needed to be defined.

To increase the packaging density and to shorten machine cycle time, new packaging techniques were employed and stringent design rules were used throughout the design process. For the first time at Cray Research, design rules and module temperatures were checked by CAD/CAM support software.

We had the option of using an exotic cooling technology to have denser packaging and hence a shorter clock period. Since we were experimenting with a new architecture and new component technology, a conservative decision was made to use an enhanced version of the CRAY-1 cooling technique.

We also had a choice between super vector speed and faster scalar performance. It was conceivable that adding more vector units would double or even quadruple vector speed for long vectors. Although it was necessary to improve vector speed, it was more important to improve the scalar speed and not to compromise the system throughput capability of the machine. A deliberate decision was made to pursue a multi-processor design instead of multiple vector units. Vector performance was increased, nevertheless, through other innovative means.

It was also decided that the development would be done within a small design team. Including logic designers, electronic, mechanical, CAD, software and application engineers, the entire design team had less than twenty people. The design and checkout of the prototype had to be completed in a very short timeframe before the market window was filled by other vendors.

It was less than three years from the inception of the project in mid 1979 to the completion of the checkout of the prototype CRAY XMP-2 in April, 1982. It proved once again that innovation and productivity are possible from a small team.

Shortly after the prototype checkout, the design of a four processor model began. In less than two years, the development, artwork, manufacture and checkout were completed. The four processor version, CRAY XMP-4, was demonstrated internally the end of April, 1984.

12. Frederickson, Paul, Hiromoto, Robert, and Larson, John, "A Parallel Monte Carlo Transport Algorithm Using a Pseudo-Random Tree to Guarantee Reproducibility," Los Alamos National Laboratory Report LA-UR-85-3184, submitted for publication in the Journal of Parallel Computing.
13. Kneis, Wilford, Industrial Real Time FORTRAN Standard, SIGPLAN Notices, July 1981, pp. 45-60.
14. Chen, Steve S., "Large-scale and High-speed Multiprocessor System for Scientific Applications - CRAY XMP Series," Proc. NATO Advanced Research Workshop on High Speed Computation, J. Kowalik, ed., Springer Verlag, Jutlich, West Germany, June 1983.
15. Hwang, Kai, and Briggs, Faye, Computer Architecture and Parallel Processing, McGraw-Hill, New York, 1984, pp. 714-731.
16. Multitasking User Guide, Cray Research, Inc., Publication SN-0222, January 1985.
17. de Forcrand, Philippe, and Larson, John, "Quantum Chromodynamics on the CRAY XMP/48 with SSD," Cray Channels, Winter 1985.
18. Edwards, Mickey, Hsiung, Christopher C., Kostof, Daniel D., and Resherf, Moshe, "Three Dimensional Seismic Forward Modeling, Part 1: Acoustic Case," submitted to the Journal of Geophysics.
19. Barton, John T., and Hsiung, Christopher C., "Multitasking the Code ARC3D," to appear in Proceedings of the GAMM Workshop, The Efficient Use of Vector Computers with Emphasis on Computational Fluid Dynamics, University of Karlsruhe, West Germany, March 13-15, 1985.
20. Larson, John, Sameh, Ahmed, Kizelyalli, Isik, Hess, Karl, and Widtger, Dave, "Two-Dimensional Model of the HEMT: (A Comparison of Computation with a Supermini and a Cray)," to appear in the Proceedings of the First Workshop on Large Scale Computational Device Modeling, NSF and IEEE Device Society, Naperville, Illinois, April 18-19, 1985.

A parallel effort was initiated by the continuation engineering team to employ the central processing unit (CPU) of the XMP, combined with the relatively inexpensive MOS memory technology, in developing a single processor model, the CRAY XMP-1. This project was completed by the end of 1983. However, in anticipation of market changes, this model was not announced until mid 1984, along with the four processor model.

Since overall performance of a machine can only be as fast as its slowest component[6], machine speed is much more than just MFLOPS (Millions of Floating point Operations per Second). High execution rates for certain special classes of codes (such as loops with long vectors) do not always lead to greater system throughput. In reality, large scale scientific problems benefit most from a balanced system design[7,8].

In most machines, I/O speed is the slowest component. To bridge the memory bandwidth difference between the mainframe and the slower I/O devices an optional Solid-state Storage Device (SSD) was designed by the continuation engineering team. The demonstrated balanced approach to system performance is one strength of Cray Research that people interested in MFLOPS rates often overlook.

2. ARCHITECTURE OVERVIEW

Housed inside the same physical chassis of the CRAY-1, the XMP mainframe has many distinctive features that advance the state of the art in high speed computing[15]. The CRAY XMP Series of computers is a range of nine compatible supercomputer models based on the XMP CPU. Configurations are available with one, two or four identical processors, and use two different memory technologies, either bipolar ECL or static MOS. The CRAY XMP Series is software upward compatible with earlier CRAY-1 systems and is supplemented by a significantly enhanced I/O capability. Figure 1 gives the overall system organization.

BIBLIOGRAPHY

1. Srinji, Vason P., and Asenjo, Jorge F., "Analysis of CRAY-1S Architecture," Proc. of the 10th Annual International Symp. on Computer Architecture, IEEE & ACM, 1983, pp. 194-206.
2. Hockney, R. W., and Jesshope, C. R., Parallel Computers, Adam Hilger Ltd., Bristol, 1981, pp. 69-95.
3. Johnson, Paul M., "An Introduction to Vector Processing," Computer Design, February 1978, pp. 89-97.
4. Kozdrowicki, Edward W., and Theis, Douglas J., "Second Generation of Vector Supercomputers," IEEE Computer, Vol. 14, No. 11, Nov. 1980, pp. 71-83.
5. The CRAY-1 Computer Systems, Cray Research, Inc., Pub. No. 22400088.
6. Worlton, Jack, "The Philosophy Behind the Machines," Conference on High-Speed Computing, Glenden Beach, Or., 1981, sponsored by Los Alamos and Lawrence Livermore National Labs.
7. Bucher, Ingrid Y., "The Computational Speed of Supercomputers," Proceedings of SIGMETRICS, 1983.
8. Worlton, Jack, "Understanding Supercomputer Benchmarks," Los Alamos Internal Report, March, 1984.
9. Larson, John L., "Multitasking on the CRAY XMP-2 Multiprocessor," IEEE Computer, Vol. 17, No. 7, July 1984.
10. Hsiung, Christopher C., and Butscher, Werner, "A Numerical Seismic 3-D Migration Model for Vector Multiprocessors," Parallel Computing, Vol. 1, No. 2, December 1984, pp. 113-120.
11. Hiromoto, Robert, "Results of Parallel Processing a Large Scientific Problem on a Commercially Available Multiple-Processor Computer System," Proceedings of the 1982 International Conference on Parallel Processing, IEEE Computer Society Press, August 1982, pp. 243-244.

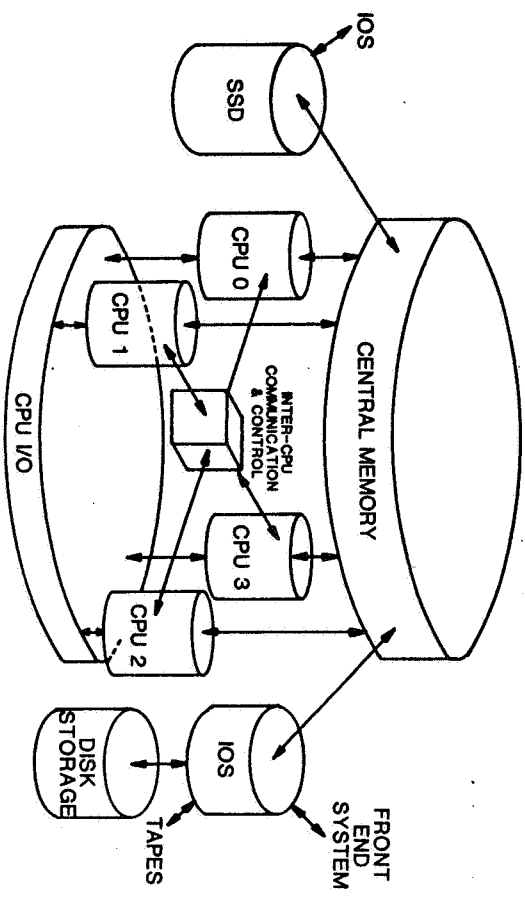


Figure 1: CRAY XMP-4 Overall System Organization

Additionally, it has set a new direction for supercomputing. With its multiple vector processors, it can simultaneously exploit two dimensions of parallelism, and, with its unsurpassed I/O capabilities, can be used in many application areas to solve problems which previously could not have been attempted.

2.1 THE XMP CENTRAL PROCESSING UNIT

Although built upon the basic architecture of the CRAY-1, see figure 2, the CRAY XMP processor is totally redesigned. All processors share a central memory of up to 16 million (64-bit) words, organized in up to 64 interleaved memory banks. All banks can be accessed independently and in parallel during each machine clock period. Each processor has four parallel memory ports (four times that of the CRAY-1) connected to the central memory: two for memory loads, one for memory stores and one for independent I/O operations.

The multiport memory has built-in conflict resolution hardware to minimize access delay and to maintain the integrity of all memory references from different ports to the same bank at the same time. The multiport memory design, coupled with a shorter memory cycle time, provides a high-performance memory organization with up to 16 times the memory bandwidth of a CRAY-1. The improved memory bandwidth balances the multiple-pipelined computing power of the CPU and the data streaming ability of the memory. For each processor, this capability coupled with the reduced clock period gives a performance speedup over the CRAY-1 of up to 4.

All processors are controlled synchronously by a central clock with a cycle time of 9.5 ns (vs. 12.5 ns of CRAY-1). I/O ports are also shared by all the processors. I/O can be initiated by any processor and any processor can field any I/O interrupt. Whenever a processor is in the operating system, for any reason, that processor will handle all I/O interrupts. If no processor is in the operating system, I/O interrupts are given to the initiating processor.

The scalar performance of each processor is improved through faster machine clock, shorter memory access time, larger instruction buffers (twice that of the CRAY-1 per processor), multiple data paths and, above all, multiple processors.

The vector performance of each processor is improved through faster machine clock, parallel memory ports and an hardware automatic 'flexible chaining' feature. The XMP design allows simultaneous memory fetches, a sequence of computations and memory store in a series of related vector operations.

Figure 3 illustrates the benefits of these features for a common vector computation in linear algebra. On the CRAY-1, this computation takes three chimes[3]. A chime is a chained operation time. The chimes consist of (load), (load, *, +), and (store). The single port to memory on the CRAY-1 prevents any further overlap of operations. The compiler must insure that the multiply and add instructions are issued at the proper time to catch the chain slot times for the computational chime. Each of the operations proceeds at a rate dictated by the CRAY-1 clock period.

Program 1-CPU 2-CPU 4-CPU

Program	Execution time			Actual Speedup (Theoretical Speedup)
	1-CPU	2-CPU	4-CPU	
PICF	72.3 sec	37.9 sec	20.7 sec	3.48(3.67)
SPECTRAL	333.7 sec	174.4 sec	94.0 sec	3.55(3.77)
GAMTEB	202.0 sec	103.0 sec	53.8 sec	3.75(3.88)
3DMICR	3.49 hr	1.85 hr	1.01 hr	3.45(3.85)
WILSON	7.65 days	n/a	2.03 days	3.77(4.00)
AC3D	4.6 hr	n/a	1.3 hr	3.50(3.80)
ARC3D	103.0 sec	55.0 sec	29.5 sec	1.87(1.98)
HESS	3.50 hr	1.81 hr	0.97 hr	1.93(1.98)

Maximum theoretical speedup based on degree of parallelism is given in parentheses.

Figure 15: Multitasking job speedup

The recorded speedups for these codes result from a combination of high percentage of parallelism, large granularity of tasks, and low synchronization overhead in the hardware and software. Other application codes will produce multitasking speedups which vary depending on the independence exploited and the multitasking programming style and mechanisms used.

6. CONCLUSION

The introduction of the CRAY XMP has set a new standard for high speed and large scale scientific computation. Its balanced and flexible architectural design, both in hardware and software, addresses the computational and I/O requirements of user programs, thus meeting the user need to solve larger and more sophisticated application problems.

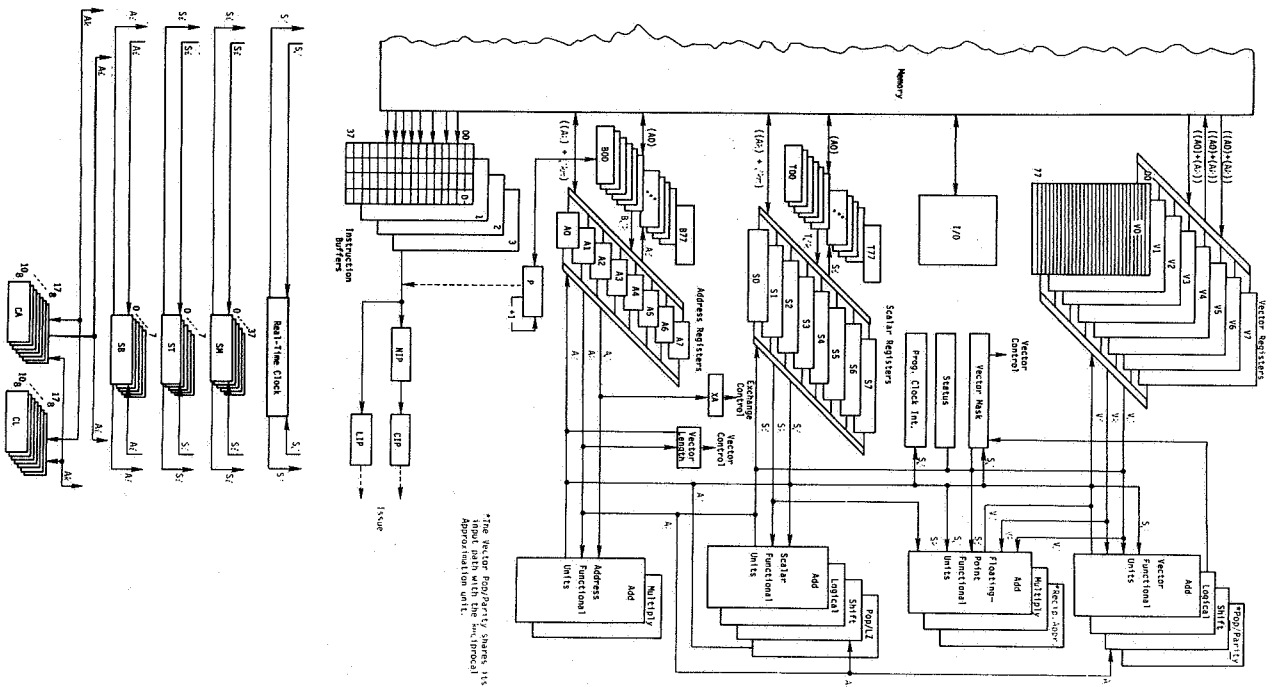


Figure 2: CRAY XMP CPU Block Diagram

time. Multitasking is applicable to 98 percent of the total execution time of this highly vectorized code.

WILSON is a lattice gauge program for measuring the force between quarks[17]. The program uses a 3 subgroup pseudo-heathath algorithm for lattice link updates, and a Wilson loop with Parisi improvement measurement algorithm applied every sweep. Parallelism occurs in the independent updating and measurement of lattice link values. Multitasking may be applied to 100 percent of the execution time of a production experiment involving a lattice of size $24 \times 24 \times 24 \times 48$ and 3500 sweeps.

AC3D is a seismic forward modelling code used to construct synthetic data by the solution of a 3-D wave equation[18]. The data are then compared to field results to determine a better model for the subsurface. A Fourier method is used which, because of its need for fewer grid points, is more efficient than a traditional finite-difference approach. Parallelism occurs in integrating independent spatial planes with FFTs. Multitasking accounts for 98 percent of the execution time of a $256 \times 256 \times 256$ size model with 1000 time steps.

ARC3D is a Reynolds-Averaged Navier-Stokes aerodynamics code using an implicit approximation algorithm[19]. Parallelism is exploited in processing independent grid planes. Multitasking is used in 99 percent of the execution time of a model experiment having a $30 \times 30 \times 30$ grid with 100 time steps.

HESS is a two-dimensional GaAs HEMT device simulation program for measuring steady-state current-voltage characteristics[20]. At each time step, the program solves a Poisson equation for the new potential, and uses an explicit scheme to update current and energy flux, electron concentration, and average energy. Parallelism occurs within each time step in the red/black SOR Poisson solution algorithm, and in the natural independence found in explicit schemes. Multitasking accounts for 99 percent of the execution time of a production experiment involving a 192×43 grid and 60000 time steps.

For each program, the maximum theoretical speedup attainable with no overhead may be computed. Let t_1 be the single CPU execution time. Then the highest speedup possible for p processors is given by

$$Sp = \frac{t_1}{t_1 * ((1-f) + (f/p))}$$

where f (the degree of parallelism) is the fraction of t_1 which is multitable.

CHAINING WITH MULTIPLE MEMORY PORTS

Workload (sec)	Wall clock time (sec)	Speedup 2-CPU 4-CPU
652.	328.	1.99
	171.	3.81

Figure 14: System throughput speedup

5.3.2 Multitasking

In certain applications, an execution speed which exceeds the capabilities of a single CRAY XMP processor is necessary. The multitasking abilities of the XMP enable the full power of the machine to be directed toward a timely solution of such large jobs. Multitasking can exploit the parallelism inherent in these codes from microscopic to macroscopic levels, and enhance either scalar or vector performance.

The benefits of multitasking are illustrated in figure 15 for several different application codes. PICF, a particle-in-cell program, is a scalar code which simulates the interaction of beams of plasma particles. See [11]. Independence occurs in the tracking of particles and the calculation of total charge distribution. The model experiment tracks 37,000 particles over 100 time steps. Ninety-seven percent of the total execution time of the program is spent in code which can be multitasked.

SPECTRAL is a short term weather forecasting code. The program is highly vectorizable. Here, independence occurs inside each time step at the outermost loop over latitudes. The model experiment has a global grid structure of 160 latitude and 192 longitude points, and simulates 200 time steps. Ninety-eight percent of the total execution time can be multitasked.

GAMTEB is a Monte Carlo code which transports gamma rays in a carbon cylinder. Parallelism occurs in the independent tracking of the original gamma rays and their offspring. This scalar program uses a technique [12] which allows reproducibility of results for codes whose execution flow is determined by a random number generator, irrespective of the number of tasks or processors. Multitasking accounts for 99 percent of the execution time of a model experiment involving one million original rays.

3DMIGR, was described earlier in the section on I/O performance speedup using SSD. Further speedup is obtained by exploiting the independence which occurs in the frequency domain at each depth level, after Fourier transformation over

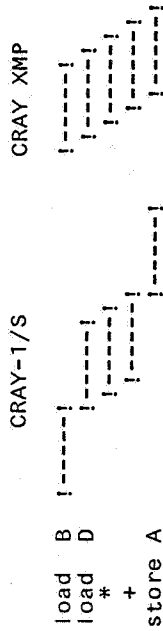


Figure 3: Vector computations

On the CRAY XMP, this computation takes only one chime. All operations are pipelined and proceed at a faster clock period rate. Additionally, the compiler has greater freedom in the scheduling of these and other supporting instructions since there is no fixed chain slot time. The elimination of fixed chain slot time and concurrent bidirectional memory access make the machine more amenable to the Fortran environment than the CRAY-1. As a result, the processor design provides higher speed vector processing capability for both long and short loops, characterized by heavy memory-to-memory vector operations.

There are also new features that support vector indirect addressing and vector conditional executions. The hardware gather/scatter unit and hardware support for compress/expand operations allow the vectorization of sparse matrix computations and certain Fortran loops with IF statements.

A unique distinction of the CRAY systems is that the vector and scalar units and controls are so intimately integrated that, physically, there is no clear scalar nor vector processor sections. This design philosophy requires tighter packaging, and allows for shorter vector pipe startup time and for faster data flow between scalar and vector units as required in the execution of typical user codes with interspersed scalar and vector code segments.

2.2 PROCESSOR COMMUNICATION AND MULTIPROCESSING

The identical processors are able to operate independently of one another and may execute different jobs simultaneously. For the processors to communicate with one another efficiently in a shared central memory environment, new hardware mechanisms are provided for interprocessor communications.

In the multiprocessor XMPs communication among processors is accomplished by clusters. Each cluster consists of eight 24-bit shared B registers, eight 64-bit shared T registers and thirty-two 1-bit semaphore registers. On the two processor XMP there are three of these clusters and on the four processor XMP there are five clusters. One cluster is typically reserved for operating system use and the others are available for user jobs.

The assignment of clusters is a function of the operating system. Any or none of the clusters may be assigned to a processor. However, only one cluster may be assigned to a processor at a time. When two or more processors are assigned to the same cluster, the cluster may be used for communication among those processors. The single processor XMP does not have clusters but the functionality can be simulated in memory for multitasking communication. On the multiprocessors, hardware arbitrates the access to a cluster during each machine cycle. When there is contention for the cluster, access is rotated among the contending processors on successive cycles.

Cluster Operations

1. The clustering of a processor may be interrogated.
2. The 24 bit values may be transferred between a processor's A registers and the shared B registers.
3. The 64 bit values may be transferred between a processor's S registers and the shared T registers.
4. A transfer may be made between the high order 32 bits of a processor's S register and all 32 semaphore bits simultaneously.
5. Single semaphores may be set or cleared.
6. A processor may wait for a single semaphore bit to be clear and then set the bit.

This last operation, the wait and set function on semaphore registers is a mechanism which includes hardware interlock. This interlock prevents a simultaneous wait and set operation on the same semaphore register from more than one processor. This basic control operation can be used to implement any of the other common software synchronization mechanisms.

3-D MIGRATION CODE (1-CPU)

EXECUTION TIME		SPEEDUP
with DISK	with SSD	
23.8 hr	3.58 hr	6.6

Figure 13: I/O speedup with SSD

5.3 MULTIPLE CPUS

The multiple processors of the XMP system are available to support the computational needs of users in two ways for increased throughput. The independence of user jobs is exploited in a multiprocessing environment for enhanced system throughput. The independence of user tasks belonging to a single job is exploited in a multitasking environment for enhanced personal throughput.

5.3.1 Multiprogramming

In batch mode, the operating system schedules independent user jobs for the processor resources. Jobs which wait for I/O or other reasons are rescheduled so that jobs which are ready to run may execute. This philosophy optimizes the utilization of the processor resources, and results in a system throughput speedup approaching the number of processors in the system.

A measure of the increase in system throughput is provided by the following example. Twenty copies of a vectorized program, each requiring 32.6 seconds of CPU time, are submitted simultaneously. The group of jobs represents a system workload of 652 seconds. If executed on a single XMP processor, the wall clock time is also 652 seconds. When two CPUs are used to simultaneously process the workload, the wall clock time, measured from when the first job starts execution until the last job completes execution, is now 328 seconds. The system throughput speedup is $652/328 = 1.99$. See figure 14. To process the workload on four CPUs, the wall clock time is 171 seconds, for a system throughput speedup of $652/171 = 3.81$.

5.2.1 Disk striping

Data transfer to or from disk can often be a bottleneck in out of core problems, resulting in diminished processor utilization and excessive I/O wait times. The limitation imposed by the transfer rate of a single disk is significantly enhanced through the disk striping capability in GRAY's IOS software, as described in a previous section.

Figure 12 illustrates the I/O times for an oil reservoir benchmark utilizing striped disks. Several runs were made, each with a different number of physical disks comprising the logical disk group.

Number of disks in stripe group	I/O time (sec)	Speedup
1	34.9	1.00
2	17.6	1.98
3	13.3	2.62

Figure 12: I/O speedup with disk striping

5.2.2 SSD Performance

The Solid-state Storage Device is a new secondary storage device that addresses the I/O needs of large scale scientific problems. An example of the I/O speedup attributed to the use of SSD is provided by a 3-dimensional seismic migration code, 3DMIGR. This code is used to determine the underground structure of the earth in oil exploration[10].

A model problem involving 200 x 200 traces, with 1024 time samples per trace, and 1000 depth levels is quite I/O bound when intermediate files reside on disk. The total computational requirement of the job is 1.5E12 floating point operations, while a total of 4.0E10 words are transferred to and from disk during execution. The total execution time for the code when run on one CPU with DD-29 disks is 23.8 hours. See figure 13. By assigning the intermediate files to the SSD, the I/O wait time is virtually eliminated. The job becomes computationally bound, and the total execution time for one CPU with SSD is 3.58 hours, a speedup of 6.6. This example is used again in the next chapter where multitasking will further improve this speedup.

This 'busy wait' operation has two significant advantages over most other hardware interprocessor exclusive access controls, such as, test and set, or compare and swap. First, it does not reference memory. Spin wait sampling of the same location in memory is avoided and hence disruption of memory accesses by other processors is eliminated. Second, the hardware can detect whether a processor is waiting. As a consequence the waiting processor can be selected at the time of interruption to do other useful work such as processing I/O interrupts. Furthermore, a primary deadlock condition is detectable by the hardware. When all processors which are assigned to a particular cluster are waiting on semaphores, that group of processors is deadlocked and the hardware issues a deadlock interrupt to each of them. The deadlock can then be resolved through the governing system software. This is useful not only in resolving a true deadlock situation but also in scheduling useful work. For example, when there are more tasks to be executed for a user code then there are physical processors, other tasks for that user code can be put into execution when the deadlock interrupt is detected.

The interprocessor communication mechanism allows the processors to send and acknowledge messages, and to synchronize activities in a timely way. The efficient mechanism enables the multiprocessors to execute the tasks of a single user code simultaneously in a coordinated manner (multitasking). This added dimension of parallel processing is at the Fortran outer loop level, on top of the vector processing at the inner loop level. The ability to simultaneously exploit these two levels of parallelism is unique to the GRAY systems. Figure 4 clarifies this multi-level parallelism.

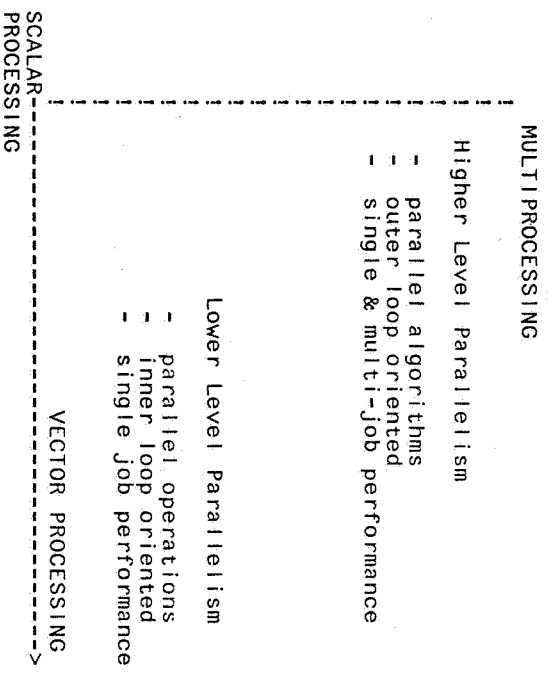


Figure 4: Two dimensions of parallelism

2.3 I/O SUBSYSTEM

The computational power of the XMP Series is complemented by powerful I/O capabilities. For problems requiring extensive data space and data movement, the I/O structure of the XMP ensures that computing power is not limited by I/O capability.

The I/O Subsystem (IOS) consists of two to four processors (IOPs). The IOS acts as a concentrator and data distribution point for the GRAY XMP mainframe. The IOS communicates with a variety of front-end computer systems (one to seven) and with peripherals such as disk units and magnetic tape units. A direct access path is also provided between the IOS and the SSD.

Associated with the IOS is a large buffer storage space for data transfer and for temporary scratch files. The IOS Buffer Memory is a solid-state storage device, using the same technology as the SSD, is accessible to all the I/O processors in the IOS and is available in 8, 32, or 64 Mbytes ca-

CODE	(1-CPU XMP)/GRAY-1 SPEEDUP
SSUM	1.36
SDOT	2.26
SAXPY	4.04
FOLR(*)	5.93 (3.40)
FOLRN(*)	7.32 (1.94)
GAITHER	2.55
SCATTER	2.48
MXM	1.33
MXMA	1.42
MINV	2.18
CFFT2	2.16
CRFFT2	2.12
RCFFT2	2.11

(*) New vector algorithms used on XMP only; the number in () indicates the speedup when vector algorithms are applied to both XMP and GRAY-1.

Figure 10: Scientific library (CAL)

CODE	(1-CPU XMP)/GRAY-1 SPEEDUP FACTOR
SGEFA	2.77
SGECO	2.75
SGESL	2.84
SGEDI	2.63
TRED2	3.08
TRED1	2.80
TRBAK1	3.27

Figure 11: General linear algebra (FORTRAN)

5.2 I/O PERFORMANCE

I/O performance is a good indication of the versatility of a machine in a real application environment. Performance gains on the XMP are achieved by addressing the I/O requirements in application codes. By using disk striping, Buffer Memory or SSD, several examples show significant speedups in this often neglected area.

```

DO 1 I=1,N
  A(I,J) = A(I(J)) + s*B(J)
1 CONTINUE

```

For this example, a compiler directive is needed to indicate that the subscripts are distinct and hence that the loop is indeed vectorizable.

```

(1-CPU XMP-4)/CRAY-1 SPEEDUP
      SHORT MEDIUM LONG
      VECTOR VECTOR VECTOR
      (N=8) (N=128) (N=1024)
      6.7 15.6 15.7

```

(Unit based on compiler generated code running on CRAY-1)

Figure 9: XMP hardware gather/scatter

The second set of benchmarks consists of assembly language scientific library routines. See figure 10. The vector length for the relative performance shown is 128 for the matrix operations, 8192 for the FFTs, and 4096 for the others. The code in each example is tuned for each machine. The speedups vary from 1.33, for those routines which use algorithms utilizing only one memory port, to 4.04 for the SAXPY operation. Other cases perform matrix computations, and Fast Fourier Transforms. Noteworthy are the first order linear recurrence routines for which new vector algorithms have been designed. Compared to the old method on the CRAY-1 the new versions show very impressive speedups: 5.93 for the full solution, and 7.32 if only the last value is required. The new algorithm may also be used on the CRAY-1 and a fair comparison shows speedups (in parentheses) in the expected range.

The final set of benchmarks consists of FORTRAN library routines taken from LINPACK and EISPACK. See figure 11. The routines from LINPACK solve a general system of equations, while the EISPACK routines solve eigenvalue problems. The order of the matrices involved is 400. The performance speedup of the single XMP processor over the CRAY-1 varies from 2.80 to 3.27 for these examples.

capacities. All IOPs are connected to the Buffer Memory through 100 Mbytes/sec ports. The Buffer Memory has single-bit error correction and double-bit error detection (SECDED) logic, and is field upgradable. It provides a large I/O buffer area between the mainframe and the peripherals (up to one Mbytes each) and allows user files to be Buffer Memory resident, thus contributing to faster and more efficient data access and processing by the CPU's.

In the IOS, each high speed channel (HISP) for streaming data to central memory has a burst transfer rate of 100 Mbytes/sec. With software overhead, the sustained rate is 68 Mbytes/sec. One bi-directional channel is standard on all systems, and a second is optional.

The IOS can support parallel disk streaming, software disk striping, a direct path between disk and the SSD without interference to the mainframe, I/O buffering for disk resident and Buffer Memory resident files, and on-line tapes as well as front-end system communication.

2.4 THE SOLID-STATE STORAGE DEVICE

The SSD is an optional, large, very fast, CPU-driven random access secondary storage device designed as an integral part of the mainframe. It can be used like a disk to store non-permanent files.

The SSD has a range of storage capacities from 256 Mbytes to 1 Gbytes. Like the central memory, it has SECDED logic for error checking. The very high speed channel (VHISP) between the SSD and the central memory is capable of transfer rate of 1 Gbytes/sec, 100 times that of Cray's fastest disk, the DD-49. On the XMP-4, two VHISP channels can double this bandwidth. A transfer rate of 1.8 Gbytes/sec has been measured on that model. For example, a data block of 16 Mbytes can be transferred in 8.89 ms. This fast transfer rate, coupled with a short access time (less than 0.4 ms for a single request, more than 40 times faster than that of a DD-49, and 25 us for a list of requests) offer an attractive alternative to a large expensive bipolar central memory. The SSD configuration is field upgradable.

The SSD has four additional 100 Mbytes/sec channels that can be linked to the I/O subsystem. In all, the SSD can be used as a fast-access user device for large pre-staged or intermediate files generated and manipulated repetitively by user programs. Its speed versus disk I/O can significantly affect the performance of large, I/O limited, scientific application codes. See figure 5.

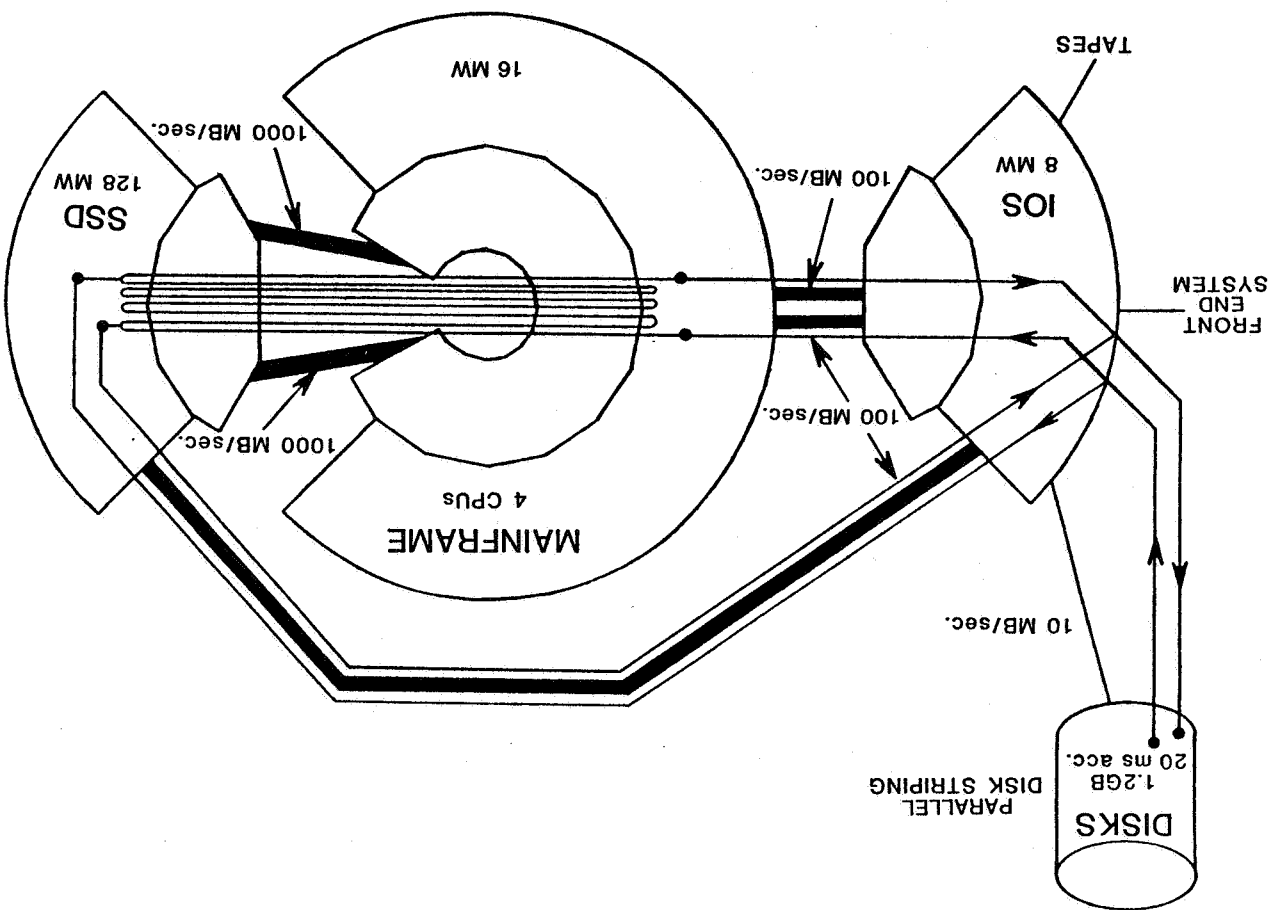


Figure 5: GRAY XMP Data Flow

The single XMP processor performance is demonstrated by comparing it with the CRAY-1 for several benchmarks sets. The first set consists of FORTRAN vector loops as shown in figure 7. The relative performance varies as a function of vector length from a typical speedup of 1.5 for short vectors (vector length = 8), to 2.5 speedup for medium vector length of 128, to 3.0 speedup for long vectors of length 1024. The best speedup occurs for the SAXPY operation (A=B+s*D) which produces a speedup of 4.0.

(1-CPU X)/1S SPEEDUP

	SHORT VECTOR (VL=8)	MEDIUM VECTOR (VL=128)	LONG VECTOR (VL=1024)
A = B	1.1	1.8	2.1
A = B+C	1.2	2.2	2.7
A = B*C	1.5	2.6	3.3
A = B/C	1.5	1.9	2.0
A = B+C*D	1.5	2.7	3.6
A = B+C*D	1.4	2.9	3.6
A = B+s*D	1.3	3.0	4.0
A = B+C+D+E	1.3	2.3	2.7
A = B+C+D*E	1.6	2.5	2.9
A = B*C+D*E	1.3	2.5	3.1
A = B+C*D+E*F	1.5	2.1	2.2
typical	1.5	2.5	3.0

(Unit based on compiler generated code running on CRAY-1)

Figure 8: Vector loop families

The XMP four processor system introduces new machine instructions and hardware support for executing gather and scatter operations in vector mode. These operations chain together with arithmetic and other memory instructions. The compiler is able to detect these constructs and generate vector code automatically. On the CRAY-1 and early GRAY XMP two processor systems, gather and scatter operations execute only in pseudo vector mode; for example, gathers are performed in a scalar fashion into a vector register. Figure 9 illustrates the performance speedup of the following sparse SAXPY, or SPAXPY, loop on a single XMP-4 processor versus the CRAY-1.

CFT treats this kind of vector merge as an unsafe optimization.

On the XMP-4 an index compression operation has been implemented in the hardware which allows the list of DO loop index values to be compressed into a dense array containing only the values selected under the condition. The scatter and gather functions can then be used to make the storage references related to these index values. The same number of elements is involved as would be in the scalar version of the loop but the loop now executes at vector speed. Also, only the elements selected are used for the arithmetic so no false errors can occur. CFT optimizes DO loops containing this type of conditional code automatically.

To facilitate multitasking, several major changes were made to CFT. Local variables of a subroutine are allocated on a stack. Related to this mechanism, the calling sequence is also amended. These changes allow CFT to produce reentrant object code, a first step toward multitasking.

In a multitasking environment, there is a need for a new kind of data scope, namely, at the task level. This scope allows data elements to be shared among the subroutines of a task, but to be private to each task for its duration. This is particularly useful if the same routine is used by different tasks. In order to support this new scope, a new COMMON statement called 'TASK COMMON' is provided in the CRAY FORTRAN language.

5. PERFORMANCE

There are several dimensions along which the CRAY XMP may be measured. The following sections investigate single and multiple processor, as well as I/O, performance.

5.1 SINGLE CPU

Many application codes require only the performance of a single, fast CPU. The architectural features of each XMP processor, as mentioned in previous sections, enable improved performance over the CRAY-1 either in FORTRAN or in Cray Assembly Language (CAL).

Examples of simple loops, and scientific FORTRAN and CAL library routines illustrate the speedups obtained in computationally intensive applications.

The SSD can also be used by the system for job swapping space and temporary storage of system programs, thus improving system performance.

Furthermore, the reduction of I/O time makes multiprocessing of large 3-D simulations attractive. High performance I/O and multitasking enable the user to explore new application algorithms for solving bigger and more sophisticated problems in science and engineering which could not be attempted before.

2.5 NEW GENERATION OF DISK TECHNOLOGY

Traditionally, disk speed has been the limiting factor for large I/O bound scientific problems. To complement and balance the CRAY XMP computing power, the CRAY DD-49 disk is available as a high density, dual port magnetic storage device, capable of reading and writing data at a burst rate of 12 Mbytes/sec. Each disk cabinet holds up to 1.2 Gbytes of formatted data on twin spindles. A sustained transfer rate of 10 Mbytes/sec is achievable, 2.5 times that of its predecessor, the DD-29.

Up to 32 DD-49 disks can be connected to an IOS for 38.4 Gbytes of disk storage. When combined with the disk striping and buffering capability of the IOS, these disks provide the XMP system with unsurpassed disk performance.

3. PHYSICAL CHARACTERISTICS

The CRAY XMP mainframe consists of 6 to 12 vertical columns arranged in an arc identical to that of the CRAY-1. The power required for a fully populated mainframe is about 30% higher than that of the CRAY-1. Accordingly, the machine requires 30% more cooling capacity. This is achieved through several enhancements. For example, a heat sink pad is added underneath each chip which attributes to a much better cooling effect for each individual chip. Also the use of drilled aluminum coldbar instead of a cast coldbar with internal tubing increases the efficiency due to less conductor interfaces.

One distinctive characteristic of the XMP lies in its packaging technology. The use of 16-gate gate arrays, eight times the integration of the 2-gate chips used in the CRAY-1, has great implications for packaging, and a faster machine cycle time.

The 16-gate gate array may consume four to five times the power of a 2-gate chip and thus creates heat dissipation problems if high wattage chips are concentrated in one locality. But, in order to shorten the signal travel time on the circuit, board foil, chips can not be placed too far apart either. To achieve a faster clock rate, stringent design rules also limit the number of fan-outs and load clustering. These difficult factors make chip placement and PC routing extremely challenging.

On the CRAY-1, two standard 6in x 8in circuit boards are mounted on two sides of a copper plate. The whole unit is called a module. Inter-module communication is done through twisted wires which are 3 to 4 feet in length. In order to shorten the signal traveling time, a tighter packaging technique is necessary. On the XMP, where a double module is used, two CRAY-1 like modules are sandwiched together. All four circuit boards may communicate through fixed locations via jumpers.

With 200 to 300 chips per module and thousands of latch-to-latch paths to check, the enforcement of design rules is no longer humanly possible without automation aids. The CAD support software checked and enforced design rules at every stage of the design process. Functional unit level logic simulation reduced the number of design errors before the machine was physically built. To keep the junction temperature of the chip to below 85C, power rules governing chip placement were also automated. The CAD effort smoothed the machine design process. This is of particular historical importance to Cray Research. The XMP project marks the first time this supercomputer manufacturer directly used a computer to design another. The complexity of the task required the use of a current generation of supercomputer to design the next generation.

4. SOFTWARE ENVIRONMENT

The CRAY-1 hardware was introduced in 1976 with only a minimum complement of available software. The early customers were sophisticated computer users willing to develop much of their own software to gain access to the performance of the CRAY-1. As the customer base broadened, the software available also grew. When the CRAY-XMP was introduced in 1982, a broad selection of software existed. The CRAY-XMP is upward compatible with its predecessor, so that software developed for the CRAY-1 migrated easily to the new hardware line. Since that time software development has continued with emphasis on optimization and the new capabilities of the XMPs.

4.3 THE CRAY FORTRAN COMPILER

The architecture is ideally suited to the FORTRAN DO loop structure. The Cray Fortran Compiler (CFT) automatically exploits the parallelism in this construct. No special syntax or subroutine calls are needed. This natural fit of the Cray vector architecture and the vectorizing compiler to the FORTRAN DO loop structure means that most codes can make use of the vector capabilities without reprogramming.

New hardware features allows CFT to vectorize two classes of code which previously were performed in scalar mode. These new features are a gather/scatter function and an index compression function. These functions are used by the compiler to vectorize codes which do not have a constant memory reference stride.

The first class involves indirect addressing, for example,

```
DO 1 I=1,N
  A(J(I)) = B(K(I)) + ...
1 CONTINUE
```

The storing of array 'A' indirectly addressed by J(I) is called a scatter operation and the corresponding loading of array 'B' indirectly addressed by K(I) is a gather operation. These functions appear in table look-up algorithms, translation codes, and in sparse matrix operations. Previously, the CFT compiler would identify the scatter/gather operations and generate calls to special optimized subroutines. With the hardware features, the scatter/gather functions can be included inline and be involved in the local code optimization.

The second class of code involves DO loops which contain conditional executions. For example,

```
DO 1 I=1,N
  IF (A(I) .NE. 0.) B(I) = B(I) / A(I)
1 CONTINUE
```

Once again there is no constant stride in the reference pattern. The CFT compiler optimized for previous systems would generate code which referenced and computed with all array elements but only changed the ones selected by the 'TRUE' condition.

This strategy has two drawbacks. First, since this vector mode implementation operates on all elements, scalar code may be faster when the selection is very sparse. Second, since the arithmetic is done on all elements, errors can occur for elements which should not have been selected; i.e., the condition may protect the code from dividing by zero.

4.1 THE CRAY OPERATING SYSTEM

The Cray Operating System (COS) is a mature system which executes on the full line of the earlier CRAY-1's and the current CRAY XMPs. This efficient operating system has both multiprocessing batch job management and interactive time-sharing capabilities. Both batch and interactive jobs have access to the same data files, interface with the outside world via stations on front-end processors, and use a common job control language.

COS supports up to 255 active user programs. These can be interactive jobs, batch jobs, or tasks of user multitasking jobs. Multitasking is invoked explicitly by user programs that create tasks to be run in parallel with other tasks within the job. Jobs compete for memory based on a priority and a dynamic aging scheduling system. Once in memory, jobs and tasks within jobs are scheduled together in a typical multiprocessing environment. A processor in a multiprocessing system is assigned on a first available basis to any of the jobs or tasks as they become ready on a round robin execution queue.

Cray Research and several major computer vendors have developed station software so that the familiar machine at the customer's site can be used as a front-end to the CRAY XMP. This enables the editing and data processing capabilities of the front-end system to be used to prepare and submit the jobs to the CRAY. This achieves a balanced separation of functions with the front-end handling the interfacing with the end user and the CRAY XMP doing the computationally intensive work. Interactive stations are available for several different manufacture's machines, and link the familiar terminal environments through to the interactive capabilities of COS on the CRAY XMP.

Much of the data used by jobs is staged to the CRAY XMP via the front-end stations. However, the low bandwidth of these front-end systems often limits the performance levels. The dataset management capability of COS is designed to support the full high-speed computational power of the CRAY XMP. The dataset manager provides efficient and flexible creation, use, and maintenance of temporary and permanent files. Although user programs can easily interface with the dataset manager, simple JCL directives allow programs to use different files without program modifications. The default allocation of files to physical disks can be overridden by JCL statements. This allows easy use of (physical or logical) devices of higher bandwidth at run time.

Multiple disks may be grouped together by COS using a software technique called striping (interleaving) where successive data blocks are distributed among the disks, effective-

TASK CONTROL

```
CALL TSKSTART ( TASKID, SUBNAME, ARGS )
```

creates a task with identification TASKID and entry point SUBNAME (ARGS), builds a stack, and enables the task for processor scheduling.

```
CALL TSKWAIT ( TASKID )
```

suspends the calling task until the task with identification TASKID has completed.

```
EVENT CONTROL
```

```
CALL EVPOST ( EVENT )
```

changes the status of the event variable, EVENT, to 'posted'.

```
CALL EVWAIT ( EVENT )
```

suspends the calling task until the status of the event variable, EVENT, is 'posted'.

```
CALL EVCLEAR ( EVENT )
```

changes the status of the event variable, EVENT, to 'cleared'.

```
LOCK CONTROL
```

```
CALL LOCKON ( LOCK )
```

suspends the calling task until the status of the lock variable, LOCK, is 'unlocked', then changes the status to 'locked'.

```
CALL LOCKOFF ( LOCK )
```

changes the status of the lock variable, LOCK, to 'unlocked'.

Figure 7: Key multitasking routines

ly multiplying the bandwidth by the number of disks in the group. The performance improvement is obtained by changes at the JCL level and need not affect the user program. See figure 6.

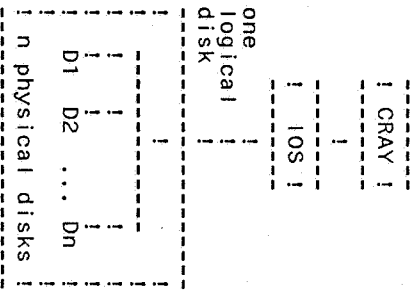


Figure 6: Disk striping

Furthermore, when a job uses files each of which is allocated on a different physical disk, access may proceed without contention. This is called disk streaming.

With two disk I/O processors (DIOPs) in the I/O subsystem, COS can sustain streaming of up to twelve DD-49s at a rate of 9 Mbytes/sec each. For file transfers involving multiple disks, COS can deliver an aggregate transfer rate of up to 108 Mbytes/sec.

Similar transfer rates can be achieved for individual files which are stored in the IOS's Buffer Memory. Through the high speed channel between the IOS and the SSD, files may be copied directly between disks and the SSD without involving the mainframe. This 'backdoor' approach totally eliminates central memory contention and the use of central memory space due to I/O. The staging of the data can be done either at the user program level or can be anticipated with a JCL level utility command.

4.2 THE MULTITASKING LIBRARY

A task is a unit of computation that can be scheduled. The main program is a task. Multitasking occurs when additional tasks belonging to the same job are created.

Multitasking is implemented at the FORTRAN level, where the user can write CALL statements to ask for multitasking functions. Many of the functions provided are similar to those found in the Industrial Real Time FORTRAN Standard[11]. The key multitasking routines are shown in figure 7, and provide the basic capabilities needed for task initiation, synchronization, and mutual exclusion.

Tasks are initiated with the TSKSTART statement by supplying a subroutine name and any necessary arguments. The TSKWAIT statement is used to wait on the completion of a generated task. As tasks execute concurrently, they may need to use quantities produced by other tasks. To ensure that these quantities are computed before they are used, the producing task may use the EYPOST routine to signal other tasks to proceed. Consuming tasks use the EYWAIT routine to listen for this signal. The EYCLEAR routine clears the signal.

Critical region protection is provided by the LOCKON and LOCKOFF routines. A task enters a critical region by turning a lock on, and exits by turning the lock off. Tasks which attempt to enter an occupied critical region wait until the lock is turned off.

The basic property of codes which can be multitasked is independence. This independence allows a partitioning of the program into tasks which may be executed in any order, or concurrently. Independence may be found at a low level in the iterations of a loop, or at a high level along geometric or other problem attributes which involve several subroutines. In general, the higher the level of independence exploited, the higher the performance speedup. Independence analysis considerations are described in [9,16].