CRAY T90 SERIES SYSTEM OVERVIEW

Figures

Figures (continued)

Tables

Purpose of this Document

This document is designed to introduce the reader to new technologies and components used in CRAY T90 series computer systems. It provides a basic summary of many subjects, most of which are described in more detail in other documents that are part of the CRAY T90 series hardware and maintenance documentation set. Users should read through this overview document to obtain a general understanding of the CRAY T90 series systems before reading the detailed documents in the set. Note that the CRAY T90 series hardware and maintenance documentation set is also available online; refer to "Online Documentation" on page 68 for more information.

CRAY T90 Series Systems

The CRAY T90 series computers are powerful, large-scale, general purpose parallel-vector supercomputer systems. They deliver high processing rates with their multiple central processing units (CPUs), large common memory, and fast 2-ns/500-MHz clock speed.

There are three models of CRAY T90 series supercomputers: the CRAY T94, CRAY T916, and CRAY T932 systems. The numbers identify the maximum number of CPUs in each system.

CRAY T90 series computers use two types of chassis. The CRAY T94 system uses a small chassis; the CRAY T916 and CRAY T932 systems use a larger chassis. Both types of chassis house various configurations of logic and memory modules, as well as power supplies and other components. Both chassis contain tanks in which these components are submersed in a dielectric coolant. Figure 2 shows the small chassis and Figure 3 shows the large chassis. The large chassis contains a physical boundary that divides the tank into halves.

CRAY T90 series systems can easily be scaled up or down, either by physically adding or removing CPUs and memory stacks or by logically upgrading or downgrading the number of CPUs and memory modules. However, the CRAY T916 system cannot be upgraded to a CRAY T932 system because the two systems use different system interconnect board (SIB) modules, and it is not possible to "swap out" the SIB on site. The SIB board is described on page 29.

CRAY T90 series computer systems have two modes of operation: a C90 compatibility mode and the Triton mode (TRI bit in the exchange packet). The A and B registers in CRAY T90 series systems have been expanded to 64 bits, and several new instructions have been added to take advantage of other architectural differences. You must recompile software written for earlier CRI machines before it will run in Triton mode. Registers are forced to 32 bits for compatibility when operating in C90 mode. Refer to "CRAY T90 and CRAY C90 Differences" on page 11.

CRAY T94 System

The CRAY T94 computer system consists of a variety of standard and optional equipment. Each computer system may be configured to meet customer needs and requirements.

A standard CRAY T94 computer system consists of the following components (illustrated in Figure 1):

- Mainframe chassis (one to four CPUs)
- 64- or 128-Mwords of memory
- Combined IOS and SSD (600-, 700-, or 800-series chassis)

  Most systems will ship with either the 600- or the 800-series chassis.
- HEU-T90 heat exchanger unit
- HVDC high-voltage DC device (40 kW)
- Support system – support system chassis, MWS, OWS, and laser printer
- Disk drives and other peripheral equipment

The components in the CRAY T94 systems, except disk drives and peripherals, are described in more detail later in this document.

The number of disk drives and other peripheral equipment used with the CRAY T94 computer system depends on individual customer needs. The disk drive configuration for a CRAY T94 computer system will usually include DE-60 and DE-100 disk enclosures or ND-12 and ND-14 ND Series Network Disk Arrays.

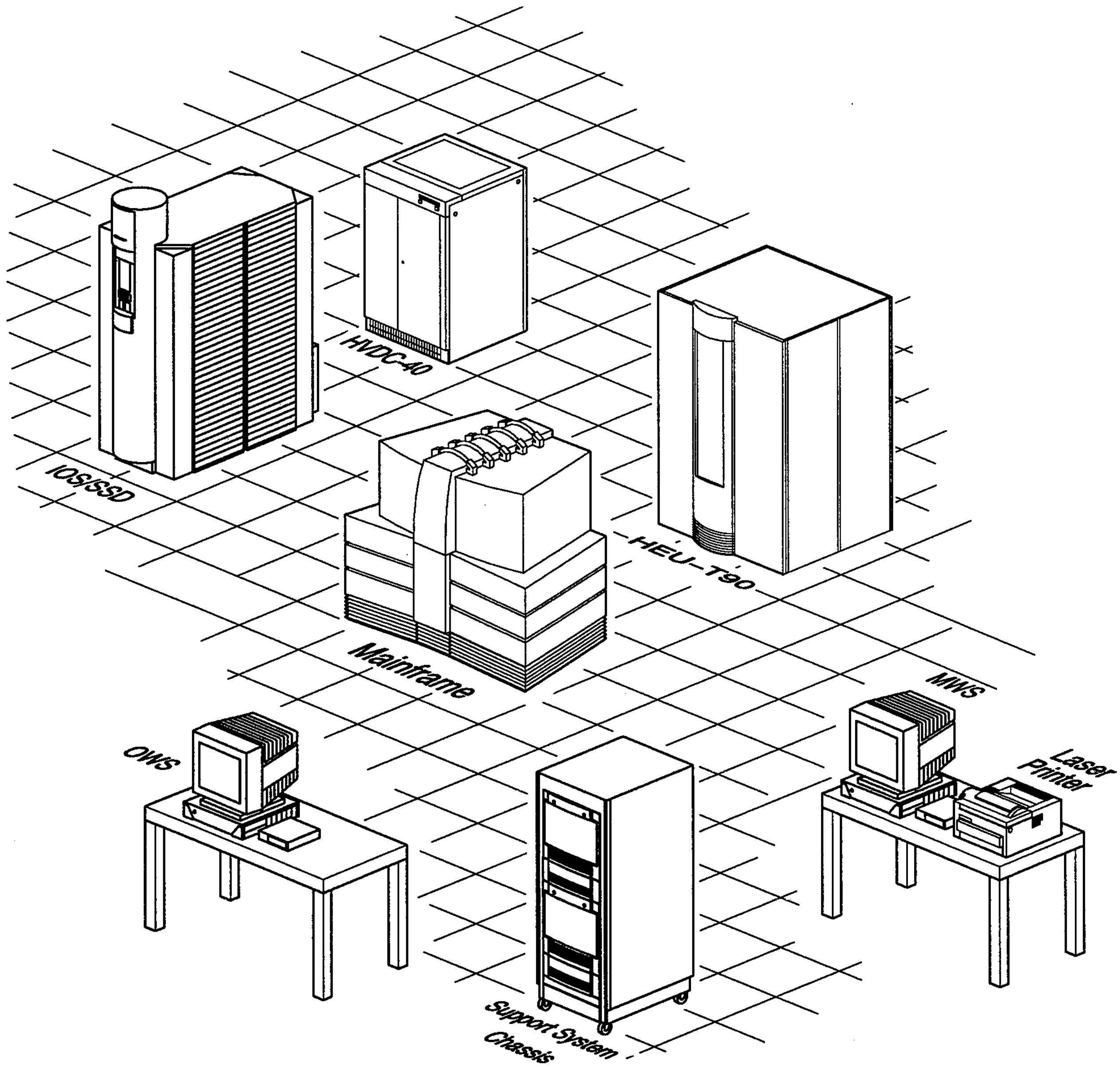


Figure 1. CRAY T94 Computer System

CRAY T916 System

A standard CRAY T916 computer system consists of the following components:

- Mainframe chassis (4 to 16 CPUs)
- 256- or 512-Mwords of memory
- One or two IOS-Es (600-, 700-, or 800-series chassis)

  Most systems will ship with one or two 700-series chassis.
- HEU-T90 heat exchanger unit
- HVDC high-voltage DC device (160 kW)
- Support System – support system chassis, MWS, OWS, and laser printer
- Disk drives and other peripheral equipment

In the CRAY T916 system, the system modules occupy one half of the chassis (refer to Figure 3). The other half remains empty, with no power or cooling connections.

Each 700-series IOS/SSD-E chassis requires a dedicated HEU.

CRAY T932 System

A standard CRAY T932 computer system consists of the following components:

- Mainframe chassis (8 to 32 CPUs)
- 512- Mwords or 1-Gwords of memory
- One to four IOS-Es (600-, 700-, or 800-series chassis)

  Most systems will ship with one to four 700-series chassis.
- Two HEU-T90 heat exchanger units
- Two HVDC high-voltage DC device (160 kW)

- Support System – support system chassis, MWS, OWS, and laser printer
- Disk drives and other peripheral equipment

CRAY T932 systems that contain fewer than 32 CP modules or memory modules can be upgraded by adding more CP modules or memory modules.

In the CRAY T932 system, the two halves of the chassis have separate power and cooling connections. This enables personnel to power off one half of the chassis and drain the dielectric coolant, while the other half continues to operate.

Each 700-series IOS/SSD-E chassis requires a dedicated HEU.

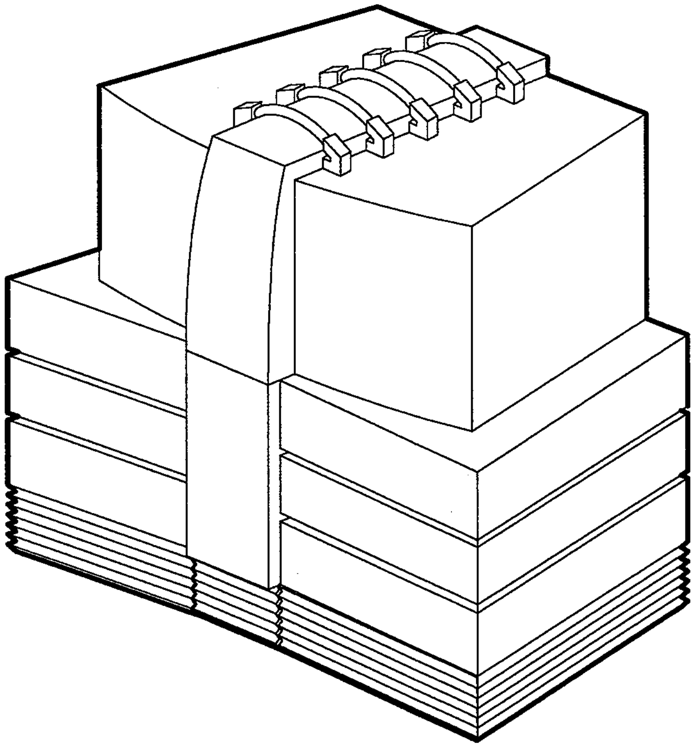


Figure 2. CRAY T94 Mainframe Chassis

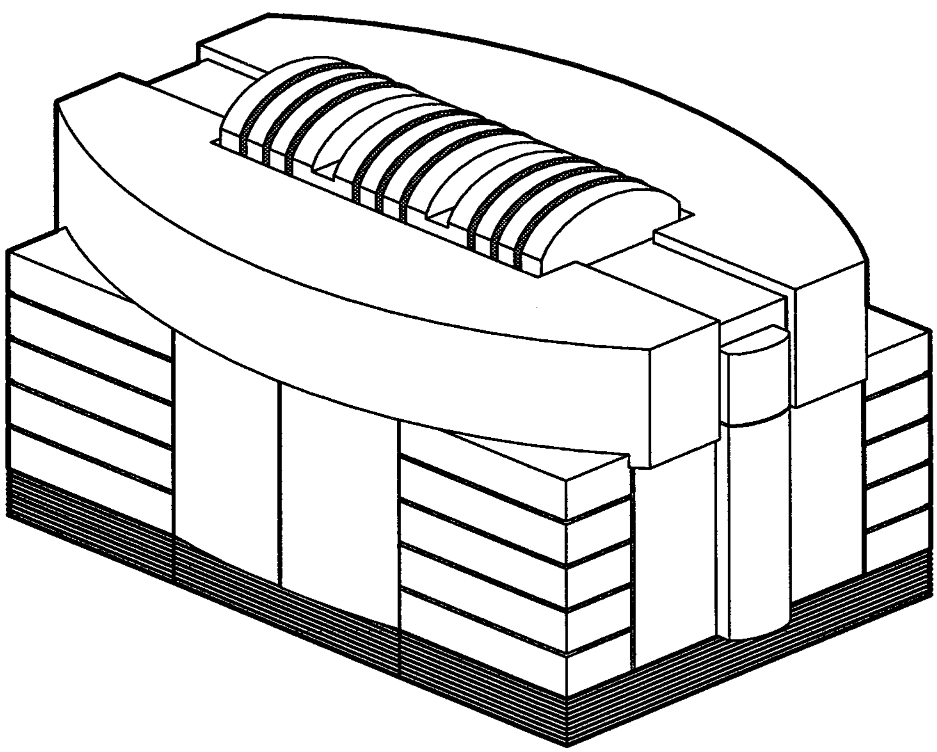


Figure 3. CRAY T916 and CRAY T932 Mainframe Chassis

CRAY T90 and CRAY C90 Differences

The CRAY T90 series systems have a base compatibility with previous CRI mainframes. However, they also have unique features that provide higher performance, increased reliability, and support for larger memories.

Major CPU Differences

- Expanded the A and B registers to 64 bits.
- Expanded the exchange package to 32 words, and changed the format. Refer to Figure 4.
- Five exit addresses in addition to the exchange address. This allows user jobs to exchange to other user jobs and monitor jobs to exchange to other monitor jobs thus reducing system overhead.
- Altered the memory management scheme. Memory management now uses logical address tables (LATs) rather than base addresses and limit addresses. Each LAT is based on a 40-bit logical space and a 35-bit physical space, allowing direct access as many as 32 Gwords of memory.
- Added scalar cache to each CPU for A and S register load/store operations. Added new instructions to manage cache and cache data use. Cache is organized into 8 pages, with 8 lines per page and 16 words per line, for a total cache size of 1,024 words.
- Expanded the shared B and T registers to 16 registers per cluster. Expanded the communication semaphores to 64 per cluster.
- Expanded the number of clusters to accommodate all CPU and configuration options. The maximum number of clusters available in a CRAY T94 or T916 system is 18 clusters. A total of 36 clusters are available in a CRAY T932 system.

CPU Scalar Differences

Several new instructions use the 64-bit A registers for logical and shift operations in order to reduce loading on the S registers. An instruction that is preceded by a 005400 instruction is specific to CRAY T90 series systems.

Other CPU scalar differences:

- Expanded the address adder functional unit to 64 bits wide and expanded the address multiplier to 48 bits wide.
- Enabled access to the bit matrix multiply functional unit from scalar registers.

CPU Vector Differences

- Added several functional units: Iota, Compress, and Expand (all are 64 bits wide). The new integer multiply functional unit is 40 bits wide. The bit matrix multiply functional unit is standard.
- Added instructions for the new functional units.
- Separated the second vector logical functional unit from the floating multiply functional unit, which eliminates a hold issue condition.
- Enabled simultaneous parallel operation of up to two gather operations and one scatter operation.
- Expanded vector mask capabilities used in the new iota and expand instructions. Modified internal logic to reduce issue conflicts.

Other CPU Differences

- Created chunk chaining of common memory read operations for vector registers. The entire vector is no longer needed for memory read operations. Chaining proceeds when eight contiguous memory operands are available instead of requiring the whole vector from memory before starting processing.
- The exchange address is no longer lost on a deadstart exchange.
- A read followed by a write to the same A*k* address will not necessarily occur in order in memory.
- The performance monitor now supports acquiring information about cache use.

I/O Differences

- Interrupt status can be read for sets of I/O channels.
- One or two LOSP channels are connected to the OWS/MWS in addition to the maintenance channels.
- Machine boot, deadstart, and deaddump functions are driven directly from the OWS/MWS.
- The OWS and MWS support platform is changed. OWS and MWS capabilities are now symmetrical; any differences are due to software.

System Configuration Differences

The System Configuration Environment (SCE) program is used to configure common memory. The SCE program establishes the number of sections, subsections, and banks in a particular system or system configuration.

SCE is also used to degrade or partition memory. By degrading memory, the system can continue operating, or maintenance can be performed in the presence of memory system failures. Systems can be degraded to a level in which one CPU is addressing one bank of memory in each of four memory sections (a total of two memory modules with only one-eighth of the memory on each of those modules in use).

By partitioning the system, parts of the I/O system can be configured to direct I/O traffic to any quarter or half of the system. The other memory configuration settings also apply to any I/O traffic through a CPU.

All configuration/degrade selections apply symmetrically within a processor; for example, selecting four-bank mode within a CPU means that all requests on all memory ports will use four-bank addressing. It is not possible during normal operation to get four-bank addressing through one port and eight-bank addressing from the remaining ports. Other configuration features are listed below.

- System configuration is performed via software. Software can logically configure modules in to or out of the system. Modules can be physically absent from the machine while the system operates normally.
- The system can be logically partitioned, allowing different operating systems to be run concurrently.

- Separate power and cooling connections in CRAY T932 systems allow one half of the mainframe to be powered down, while the other half is still operating.

- Mainframe major components (common memory, I/O, and clusters) can be configured into separate partitions. Each CPU can be assigned to a partition.

Other Hardware Features

- Wire mat has been replaced with printed circuit boards called a SIB (System Interconnect Board).

- Power supplies are arranged in an N+1 configuration to keep the system operational in the event of a power supply failure.

- All modules in the mainframe, with the exception of the I/O module, support Boundary Scan testing. This enables field personnel to test the interconnects on the module and the interconnects from module to module.

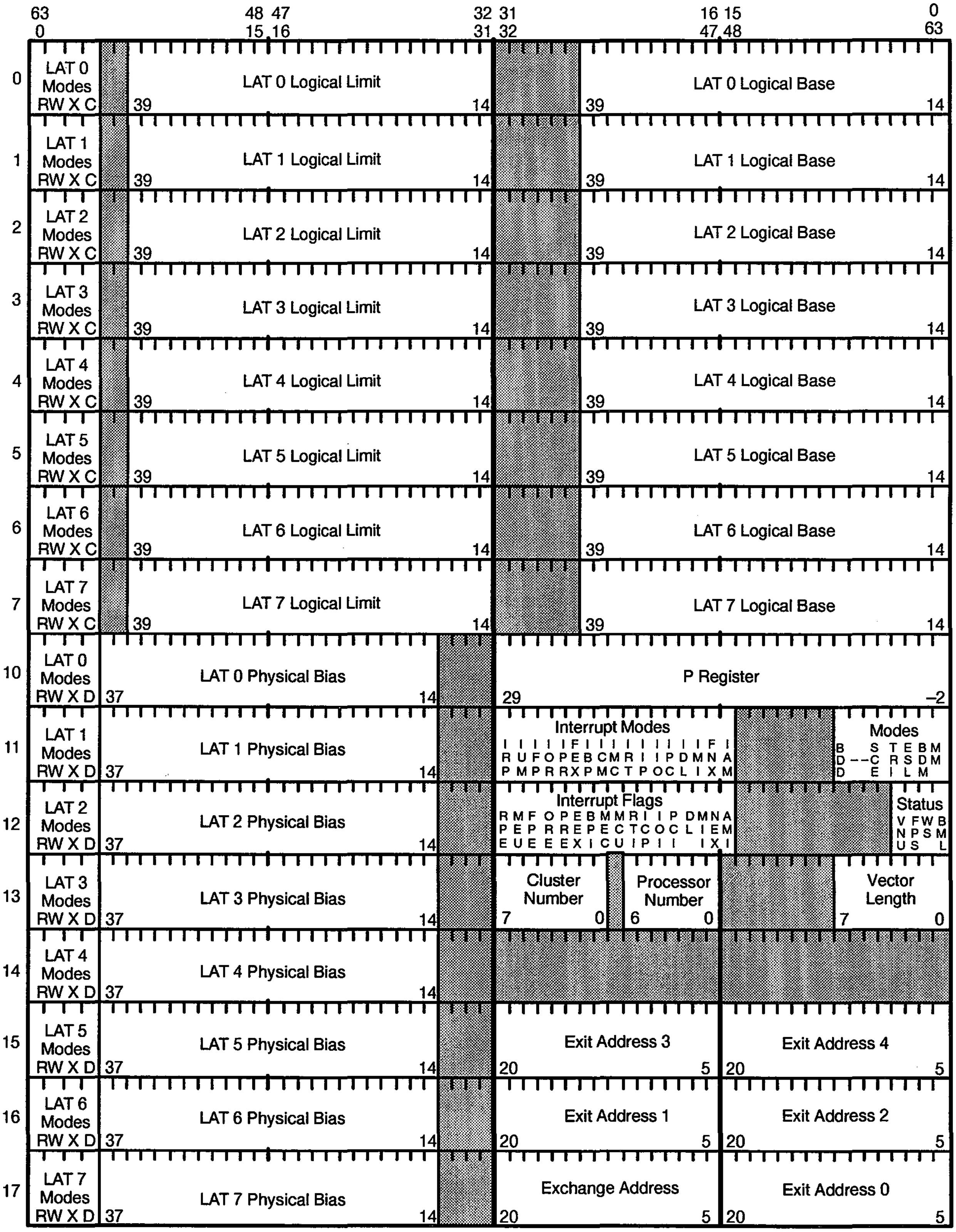
- The CRAY T90 series systems use the single-byte correction/double-byte detection (SBCDBD) error-correction scheme. With SBCDBD, single-byte errors can be detected and corrected; double-byte errors can be detected, but not corrected. Errors involving 3 or more bytes cannot be detected reliably. [A byte is defined as 2 data bits that are stored on the same memory chip (bits 0 and 16 on one chip, bits 1 and 17 on another chip, and so on).]

Memory Control Differences

The memory scheme of the CRAY T90 series system differs from the memory scheme of earlier systems. In earlier systems, Cray Research uses a memory system referred to as a *parallel access system.* In this type of memory, all CPUs are involved in every memory reference.

The CRAY T90 series system uses a *network access system.* In this system, the requesting CPU arbitrates the access in its own ports only. Once the arbitration is complete, the reference leaves the CPU, and no other CPU ever gets involved in the arbitration process. The network module and the memory module must resolve access conflicts to the memory bank.

The network and memory modules work in a *handshaking* fashion. There are first-in-first-out (FIFO) buffers and arbitrators on these modules. If a FIFO buffer for an arbitrator is full because of an unresolved conflict, that buffer will not send a Resume signal to its supplier. The supplier then starts filling up its FIFOs. This process continues all the way back to the CPU. When the CPU FIFOs are full, the system stops issue. The buffer sizes minimize the number of hold issues caused by memory backup.



Words 20 – 27: A Registers 0 – 7
Words 30 – 37: S Registers 0 – 7

Figure 4. CRAY T90 Exchange Package

Logical Address Table (LAT)

The CRAY T90 series computer system uses a system of memory management called logical address translation (LAT) tables. The LAT tables replace the traditional base address registers and limit address registers, which were too restrictive. LAT tables allow for multiple address spaces.

Each LAT has a 16-K block size and is based on a 40-bit logical address space and a 38-bit physical address space. This means that the maximum logical address is 1 teraword and that space is allocated and checked in 16-K blocks. Because of the 38-bit physical address space, the LAT implementation has a 256-Gword limit. However, the CRAY T90 series system supports 34 address bits for a total of 16 Gwords of memory.

The exchange package provides space for eight LAT entries. There are hardware limitations in the memory control logic that allow for a maximum of three LAT entries. There can be three entries with read and/or write permission and three entries with execute permission. If the user violates these limits the LAT information could be lost or ignored. If a LAT table has no mode bits set, the rest of the entry is ignored.

The CRAY T90 series system does not return memory requests in a predetermined order. This means that the order of the LAT entries into the hardware cannot be determined from their order in the exchange package. Because of this, there is no way to determine which LAT table was lost if mode limits are exceeded.

Refer to Table 1 for a descriptions of the various LAT fields.

Table 1. LAT Fields

Field	Description
Logical base	The first logical address of this LAT
Logical limit	The last logical address +1 of this LAT
Physical bias	A logical-to-physical mapping constant (Physical_Base_Address = Physical_Bias + Logical_Base Address)
Modes	Bits that define the use of this LAT (R)ead (W)rite (X)execute (C)achable

LAT Calculations

LAT calculations are based on three fields: logical base address, logical limit address, and physical bias. Unlike previous systems, there is no physical base address.

The physical bias filed maps logical addresses to a physical space in memory. The physical bias is calculated by the operating system and the user has no control over it. Physical bias is calculated using the following equation:

Physical Bias = Physical Base Address – Logical Base Address

Physical Base address = Physical Bias + Logical Base Address

NOTE: Because there is no relation between physical space in memory and logical space, the physical bias could be a negative number.

The programmer sets the logical base address and logical limit address in the exchange package. Refer to Figure 5 for an example of the CRAY T90 series memory mapping. In this example, three *pages* or areas of memory are mapped.

Page 1 has a physical base address of 0 and a logical base of 30K. Using the formula to calculate physical bias yields a physical bias of –30K. If a requested address was 35K and you wanted to map that to a physical space in memory, you would add that address to the physical bias, and the results would be at address 5K in memory.

The other two pages can be calculated the same way as the first page. Note that when the physical bias is 0, the request address always equals the physical address.

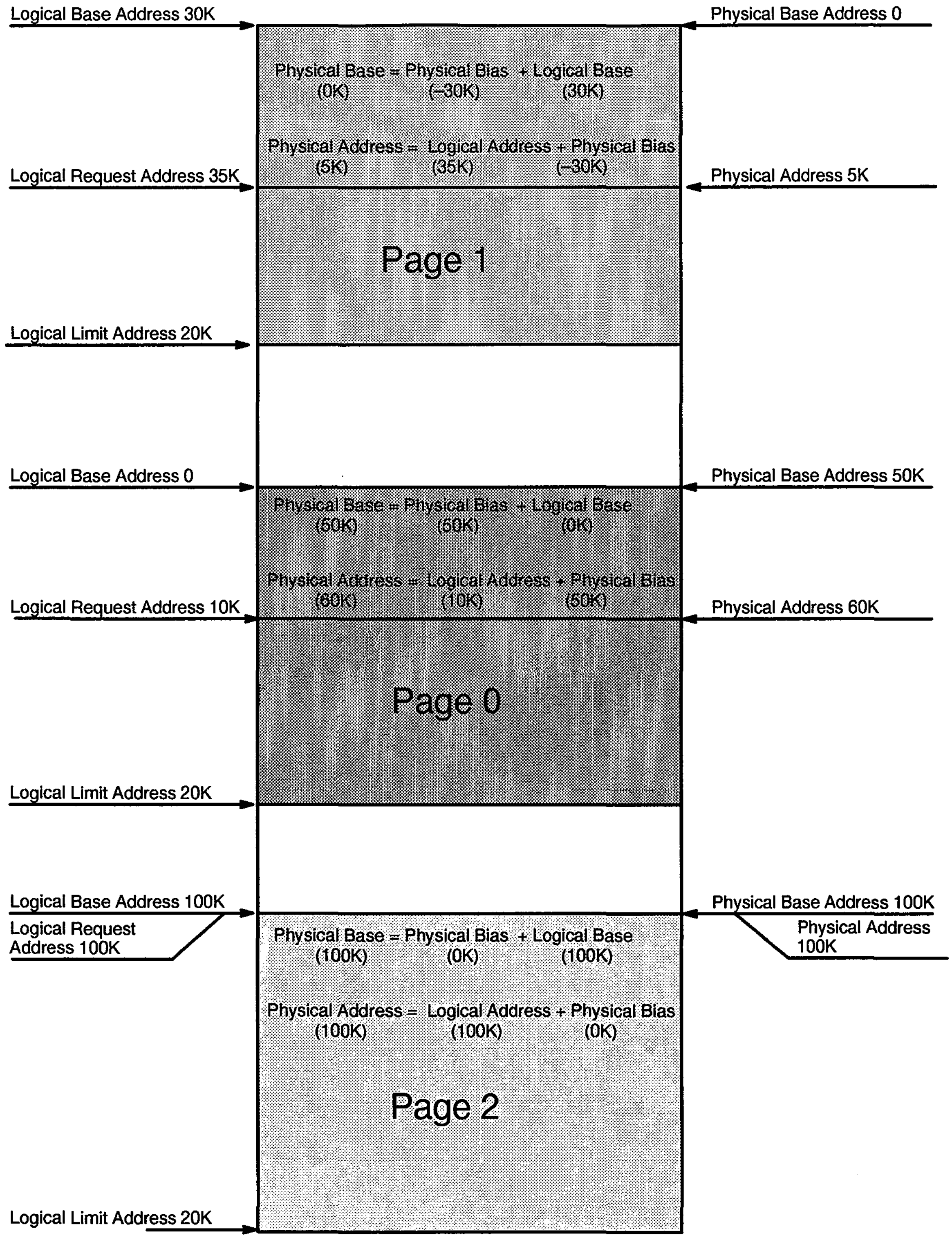
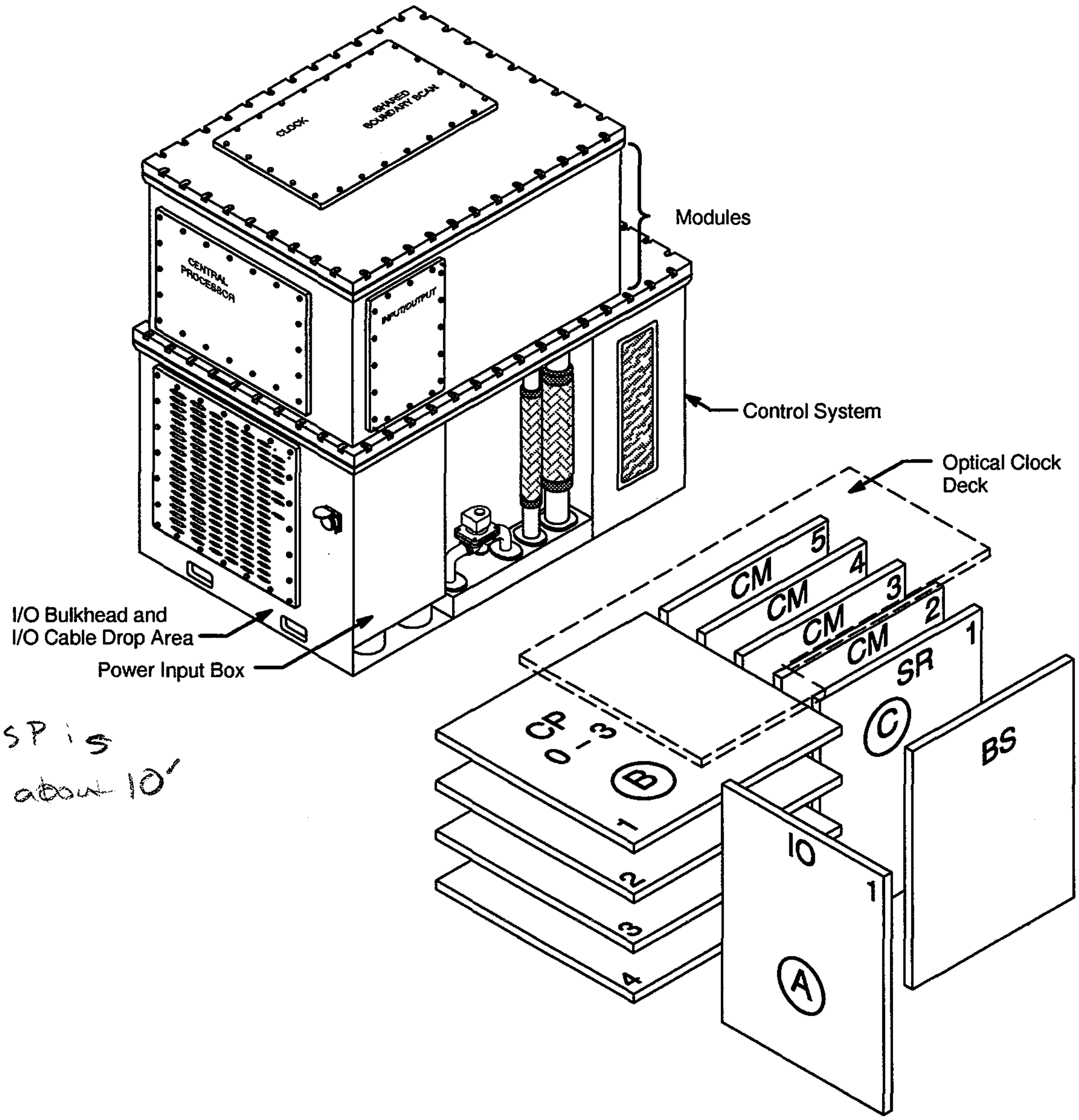


Figure 5. LAT Memory Mapping

Chassis Maps

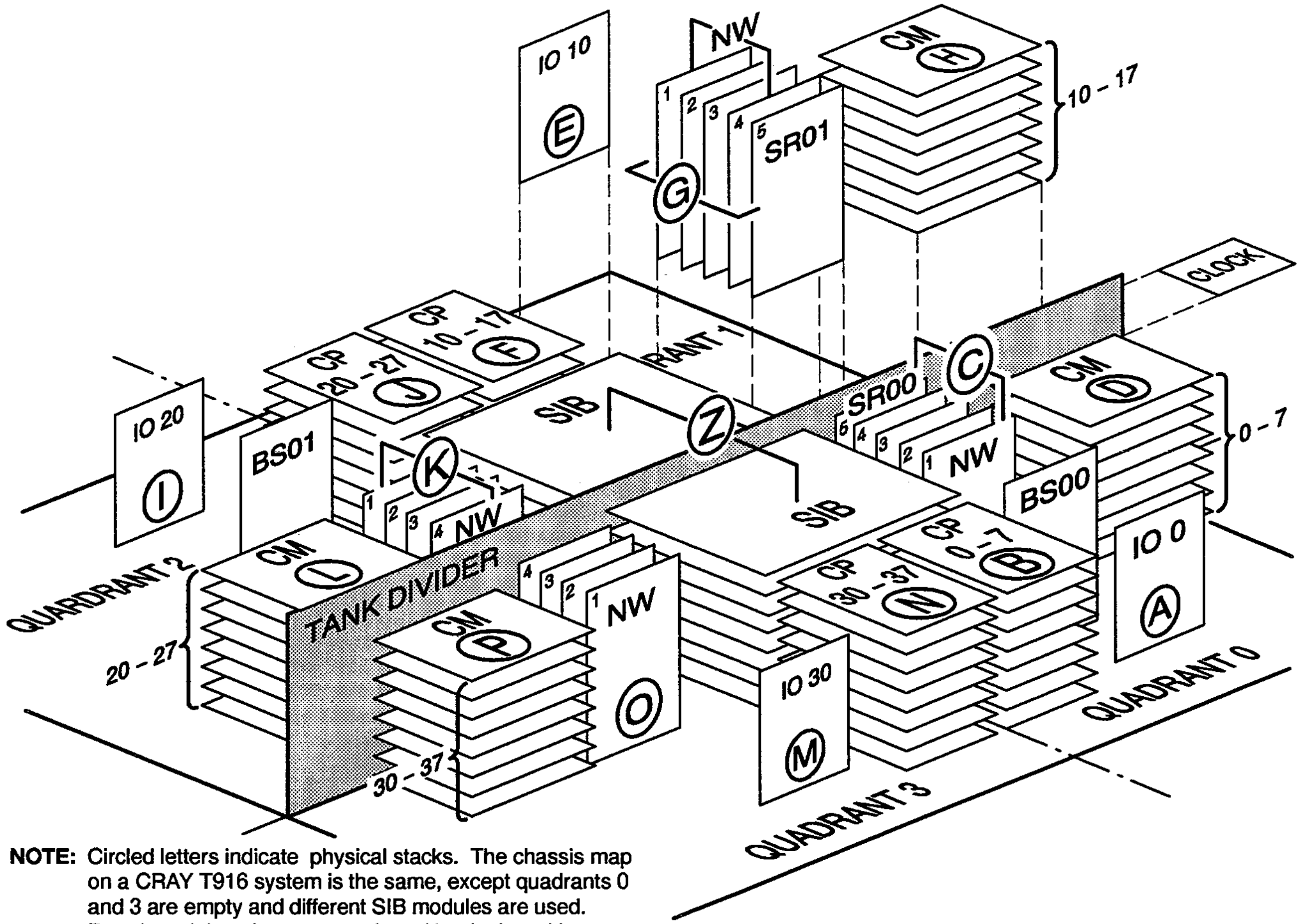
Chassis and module maps for CRAY T90 series mainframes are shown in Figure 6 and Figure 7.



VHISP is now about 10'

Physical Stack	Module Type	Module Number
A	I/O	IO 0 (to CPs 0, 1, 2, 3)
B	CPU	CPs 0 – 3 (CP 0 is normally on the top of the stack.)
C	Common memory and shared	SR 0 and CMs 1 – 4

Figure 6. CRAY T94 Chassis and Module Map



NOTE: Circled letters indicate physical stacks. The chassis map on a CRAY T916 system is the same, except quadrants 0 and 3 are empty and different SIB modules are used. Board stack locations are numbered beginning with number 1. For example, CP0 is normally in location B1.

Physical Stack	Module Type	Module Number
A	I/O	IO 0 (to CPs 2, 3, 4, 5)
B	CPU	CPs 0 – 7
BS0	Boundary Scan	BS 0
BS1	Boundary Scan	BS 1
C	Networks and Shared	NWs 0 – 3 and SR 0
D	Memory	CMs 0 – 7
E	I/O	IO 10 (to CPs 12, 13, 14, 15)
F	CPU	CPs 30 – 37
G	Networks and Shared	NWs 10 – 13 and SR 1
H	Memory	CMs 10 – 17
I	I/O	IO 20 (to CPs 22, 23, 24, 25)
J	CPU	CPs 20 – 27
K	Network	NWs 20 – 23
L	Memory	CMs 20 – 27
M	I/O	IO 3– (to CPs 32, 33, 34, 35)
N	CPU	CPs 10 – 17
O	Network	NWs 30 – 33
P	Memory	CMs 20 – 27
Z	System Interconnect Boards	SIs or SJs 0 – 7

Figure 7. CRAY T916 and CRAY T932 Chassis and Module Map

Module Types

Eight types of modules make up the CRAY T90 series system:

- CPU module (CP)
- Memory module (CM)
- I/O module (IO)
- Shared module (SR)
- Network module (NW)
- Boundary scan module (BS)
- System interconnect board (SI or SJ)

NOTE: The SIB is actually a board, not a module. All of the eight module types listed, except the SIB, are field replaceable units (FRUs).

A single printed circuit (PC) board is used with all CRAY T90 series modules. Unlike CRAY C90 style modules, the CRAY T90 series PC boards are not mounted on a coldplate. Because the system is submerged in dielectric coolant, module coldplates are not needed. The dielectric coolant flows around the PC boards.

Central Processor (CP) Module

The CP module contains one CPU. The CRAY T90 series system can contain from 1 to 32 CPUs, depending on the chassis type. CP modules are arranged in a stack of up to four modules in a CRAY T94 chassis. CP modules are arranged in stacks of up to eight modules in CRAY T916 and CRAY T932 chassis. The CRAY T916 chassis contains two physical stacks and the CRAY T932 chassis has four stacks.

CRAY T90 series CPUs have double vector and floating-point functional units and a double pipeline for operands to and from memory. The double pipeline allows two elements or operand pairs to access two functional units in 1 clock period. The CPU is instruction compatible with the C90 instruction set, and will run in a C90 mode of operation.

Figure 8 shows a block diagram of the CPU.

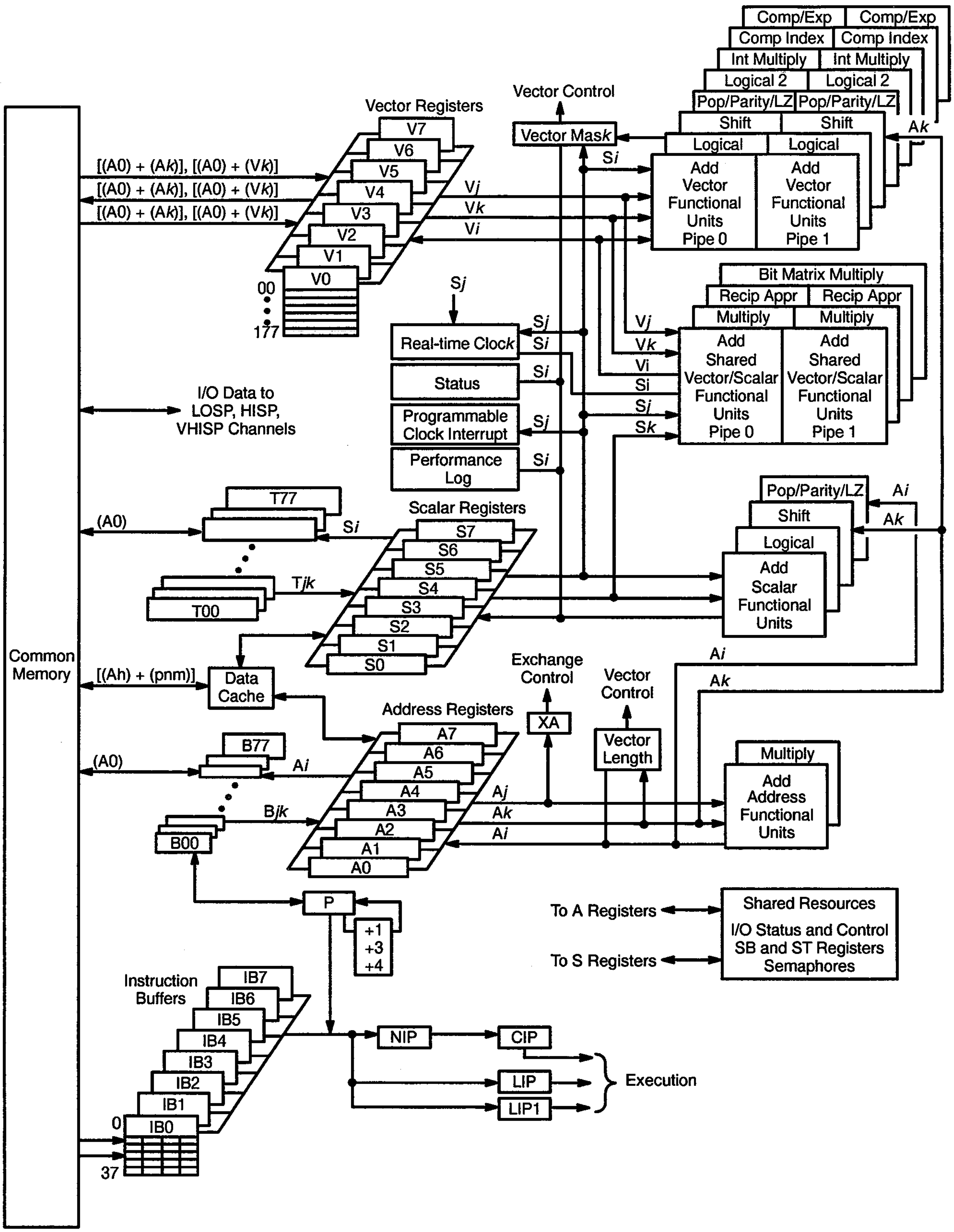


Figure 8. CPU Block Diagram

Common Memory (CM) Module

The CM module contains central memory that is common to all CP modules. Like the CP modules, the memory modules are arranged in stacks. The CRAY T94 chassis contains one stack of up to four CM modules. The CRAY T916 chassis contains up to two stacks of eight modules, and the CRAY T932 chassis contains up to four stacks of eight modules. Central memory is organized by sections, subsections, and banks. However, CRAY T94 systems do not have subsections.

Each CM module consists of a single printed circuit board with memory stacks on one surface and logic options on the other surface. Memory stacks (Figure 9) contain 40 SRAM chips stacked in two columns of 20. When fully populated with 4-Mbit chips, a CM module contains 32 Mwords of memory. Memory stack connectors (MSCs) connect memory stacks to the memory module. The MSC is an EZIF connector that requires the application of electric current to insert or remove the memory stack from the module.

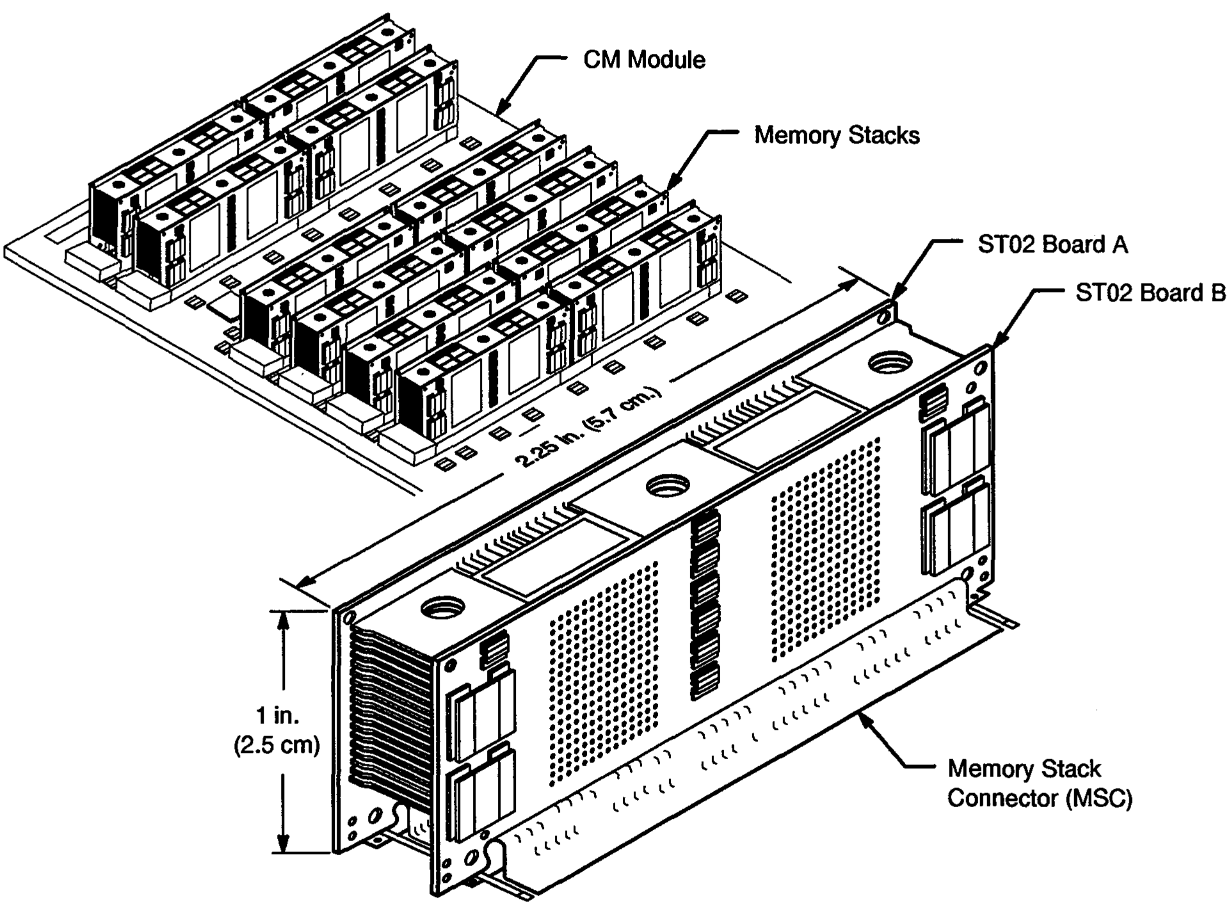


Figure 9. Memory Stack

Each memory module has 16 memory stacks. Each memory stack consists of two printed circuit boards (ST02 A and ST02 B) with forty 4-Mbyte SRAM chips. The stack has two columns of 20 chips each. Nineteen chips in each column store data. The top chip functions as the spare chip.

In CRAY T916 and CRAY T932 systems, each module stack connects to four network modules. The network modules enable any CPU to access any area of common memory. They also provide connections for memory access control. In CRAY T94 systems, a module stack connects directly to the CP modules; therefore, the final level of memory interconnecting and routing takes place on the memory modules.

I/O Module

The I/O module provides an interface between system modules and external channels. The I/O module connects directly to four of the eight modules in a CP stack. Table 2 lists the maximum number of I/O modules and channels that each CRAY T90 series system supports.

Table 2. Maximum I/O Modules and Channels Configuration

CRAY T90 System	I/O Modules	LOSPX Channels	LOSP Channels	HISP Channels	VHISP Channels
CRAY T932	4	16	32	32	16
CRAY T916	2	8	16	16	8
CRAY T94	1	4	8	8	4

Each I/O module is divided into four logical groups or quadrants. Each quadrant is logically connected to one of the four CPUs that make physical connection to the I/O module. Each quadrant on the module can handle transfers for the following channels:

- Two LOSP channels
- Two HISP channels
- One VHISP channel
- One LOSPX channel

Each quadrant defines LOSPX channel characteristics. Quadrant position on the I/O module determines the type of a given LOSPX channel. LOSPX channel assignments are as follows: quadrant 0 is the maintenance channel; quadrant 1 is the support channel; quadrant 2 is the error logger channel; and quadrant 3 is the PINT channel.

LOSP, HISP, VHISP, and LOSPX channels exit the I/O module from 1 of 16 "fuzz-button" connectors. Refer to "Fuzz-button Connectors" on page 41 for more information and an illustration.

All I/O channels can operate simultaneously, sharing memory bandwidth with individual CPUs. All CPUs can perform data transfers on all channels. Any CPU within the system can request the services of any LOSP or VHISP channel. Because a direct connection between the requesting CPU and receiving I/O module may not exist, any I/O command is always initially sent to the SR module. The SR module can either access the I/O module or it can send the command to the remote SR module in a CRAY T932 system. The SR module sends the command to an I/O channel through the CP module that lies between the SR and I/O module where the channel resides.

The CRAY T932 system contains up to four I/O modules, with one of the four I/O modules using its maintenance channel port. Two maintenance channels may be configured, one on each side of the tank, if the mainframe is configured for more than one logical system.

Shared (SR) Module

The shared module contains the following shared resources: 16 SB registers, 16 ST registers, and 64 semaphore registers per cluster. Each shared module supports 18 clusters. The following list shows the number of clusters available for each system:

- CRAY T94 systems .... 18 clusters
- CRAY T916 systems ... 18 clusters
- CRAY T932 systems ... 36 clusters

CRAY T94 and CRAY T916 systems contain one shared module, CRAY T932 systems contain two. The shared module contains all of the shared register logic and some of the I/O control logic. The shared module synchronizes CPUs when setting the real-time clock (RTC). The shared module also maintains configuration data that specifies which CPUs are assigned to each cluster group.

Network (NW) Module

The network module provides an interface between CM and CP modules. There are four network modules per stack of CM modules. Each network module is capable of steering data from eight sources to eight destinations. The network module is also responsible for subsection arbitration. CRAY T94 systems do not contain NW modules.

Boundary Scan (BS) Module

The BS module performs the following functions:

- Serves as an interface for all boundary scan operations
- Serves as the system interface for continuity line information and control between the continuity line sensors and the monitoring system
- Acts as a substitute maintenance interface (in place of an I/O module) in the CPU tester only

CRAY T94 and CRAY T916 systems use one BS module, and CRAY T932 systems use two. In a CRAY T932 system, a BS module is located in each half of the mainframe, with each BS module connecting to 44 modules (88 in the system). Each BS module has connectors to support 48 system modules.

System Interconnect Board (SIB)

The CRAY T90 series system has no wire mat. The system interconnect board (SIB) replaces the wire mat that was used in previous Cray Research computer systems. All module interconnections are accomplished with an SIB and two types of EZIF module connectors.

The SIB passes data and control signals between modules that connect to it. CRAY T916 and CRAY T932 computer systems contain eight SIBs; CRAY T94 systems do not contain any SIBs.

CRAY T916 and CRAY T932 systems use two different SIB module types: SJ and SI modules. CRAY T916 systems use the SJ module type, which is approximately one half the size of an SI module. CRAY T932 systems use the SI module type. The SI is a single module that is present on both sides of the tank divider in order to connect to module stacks on both sides of the system.

Each SIB module consists of two printed circuit boards bonded together with a nonconductive adhesive. SIBs have no logic components and no interconnections between the two printed circuit boards. The SIB is not a field replaceable unit.

Optical Clock Module

The optical clock module generates and sends an optical clock signal to all modules in the mainframe chassis. The optical clock module consists of two printed circuit boards: an oscillator board and a laser board, as shown in Figure 10. The oscillator board contains the crystal oscillators and integrated circuits. The laser board contains the optical transmitter modules (OTMs), which house the lasers. The optical clock module has its own power supply.

In CRAY T94 systems, the optical clock module is mounted on a clock deck and is located on top of the card cage (Figure 10). The clock deck is a field replaceable unit (FRU) if the coupler fails; however, if the optical clock module fails, then the module becomes the FRU instead of the clock deck. In CRAY T916 and CRAY T932 systems, the optical clock module is located between memory module stack D and memory module stack H and is labeled Clock Box.

Clock Signal Path

Figure 10 shows the path of the optical clock signal. The oscillators create a differential electronic clock (ECL level) signal to drive the OTMs. Four oscillators on the optical clock module provide three different clock frequency levels: 475, 500, and 525 MHz (slow, normal, and fast). The normal clock frequency runs at 500-MHz or 2 nanoseconds. Two normal frequency oscillators are available for redundancy purposes.

The OTMs convert the differential electronic clock signal received from the oscillators into an optical signal. The OTMs contain the lasers that send the optical signals off the optical clock module through fiber-optic couplings. Fiber-optic waveguides transfer the optical signals to the optical receiver on mainframe modules. Specifications for OTM lasers used in CRAY T94 systems differ from those used in CRAY T916 and CRAY T32 systems.

Couplers split clock signals coming from an OTM and send them to modules in the mainframe chassis. There are two levels of couplers on each optical clock module. A 2-by-2 coupler is used in CRAY T94 systems; it has two inputs (one from each OTM) and two outputs (only one output is used). This coupler sends the optical clock signal to a 1-by-8 coupler. The 1-by-8 coupler splits the signal and sends it to 8 different system modules.

The optical receiver on each module converts the signal back to an electrical signal and sends it to a clock distribution option on the module. The clock distribution option then distributes the electrical clock signal to all options on the module.

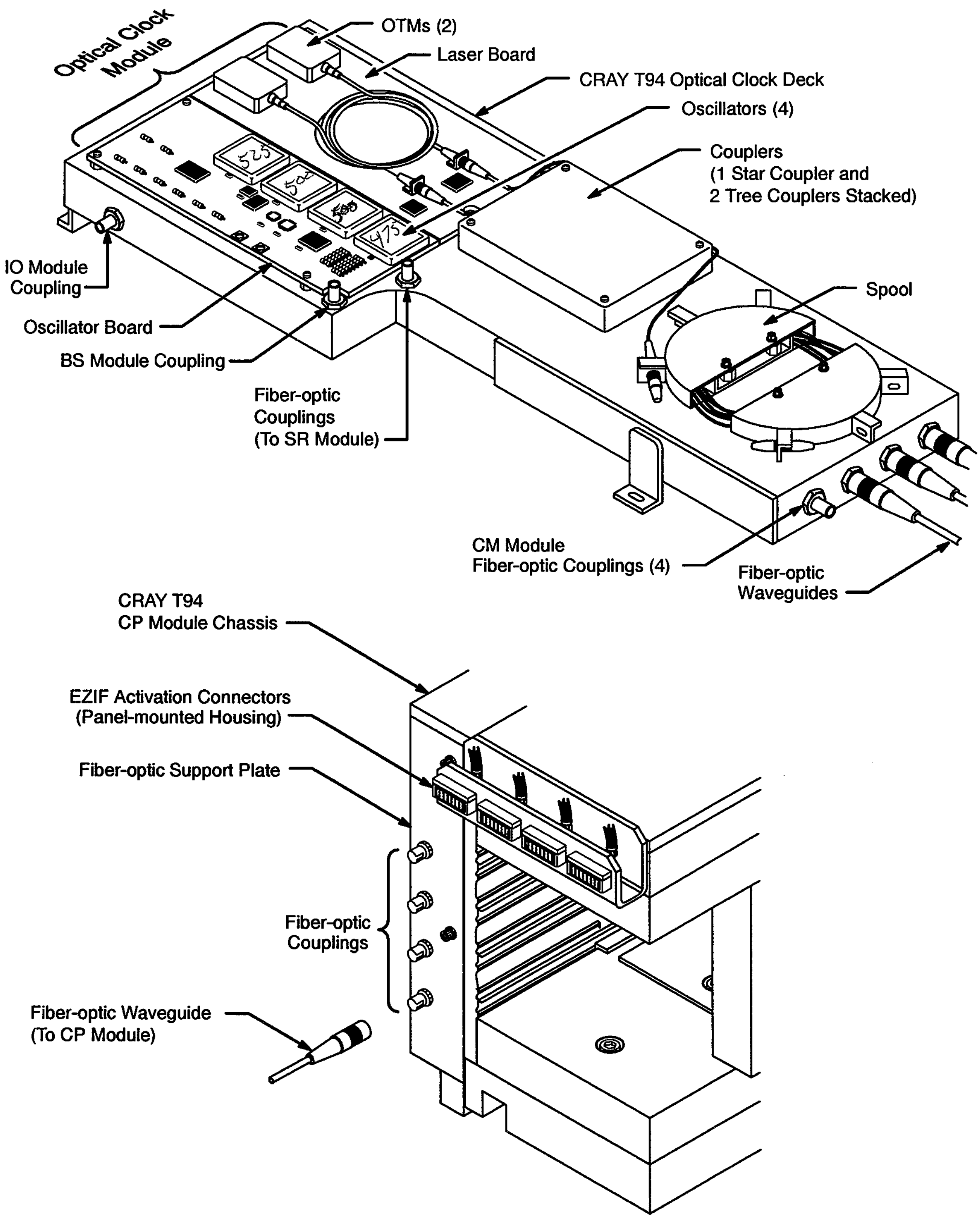
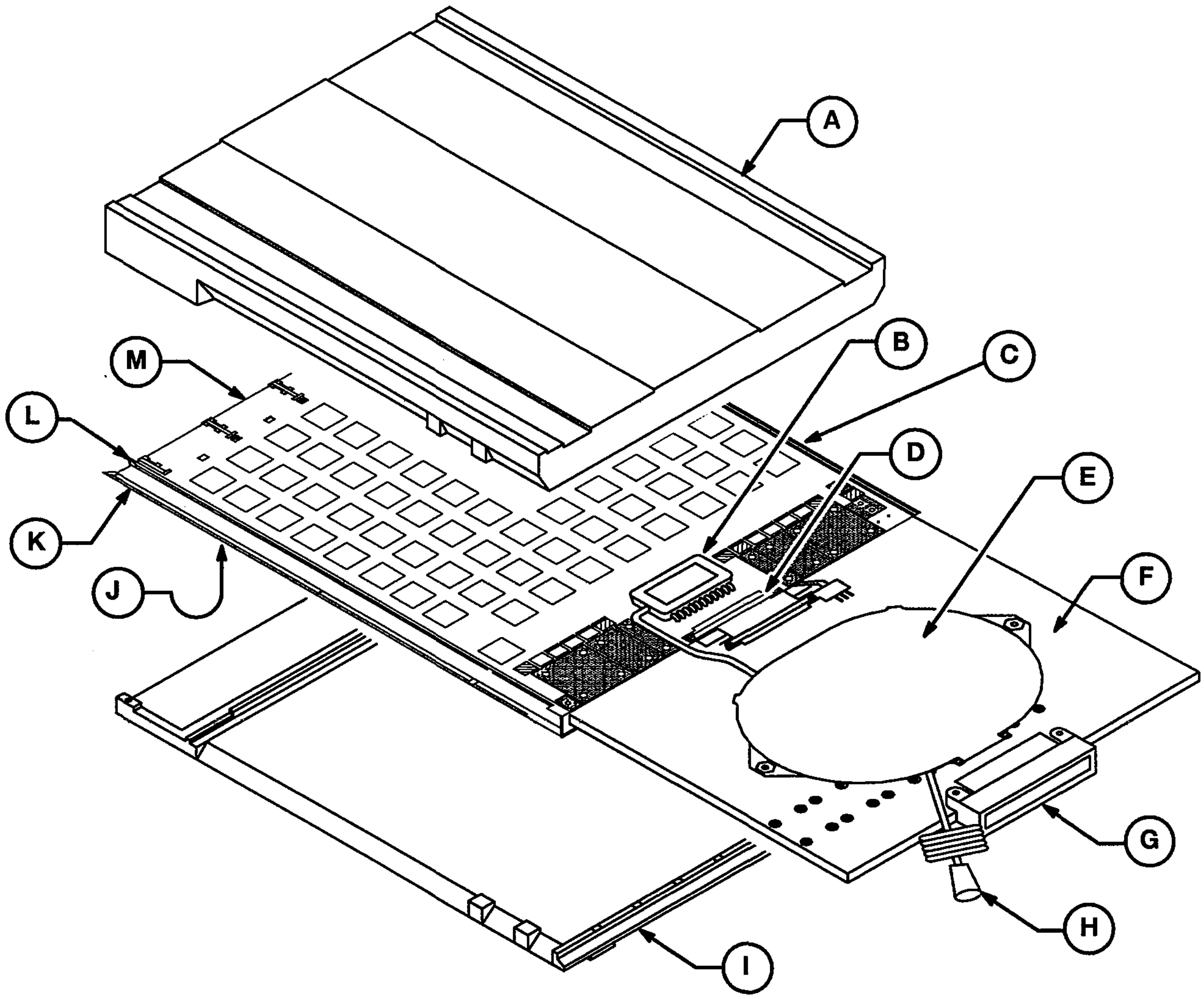


Figure 10. Optical Clock Signal Path

Module Assembly Components

Figure 11 illustrates CP module components. This illustration is provided to show the basic components that are part of all mainframe modules. Sizes of various components differ between modules.



A	Flow Block, Board 1	**H**	Fiber-optic Coupler
B	Optical Receiver	**I**	Flow Block, Board 2
C	PC Board Edge Shim	**J**	PC Logic Board 2
D	Maintenance Connector Flex Assembly	**K**	Outer Rail
E	Fiber-optic Spool Assembly	**L**	Inner Rail
F	Voltage Regulator Board Assembly	**M**	PC Logic Board 1
G	Maintenance Connector		

Figure 11. CP Module Assembly Components

Module Shipping Containers

Figure 12 illustrates the module shipping container. The ESD bag and its contents are placed under a vacuum to remove all the air; then, nitrogen is forced through the ESD bag to remove any remaining air and moisture. The vacuum-sealed ESD bag and contents are packaged in an ESD box, which fits into an aluminum shipping case. The final shipping case has not been defined at this time; it is not shown in Figure 12.

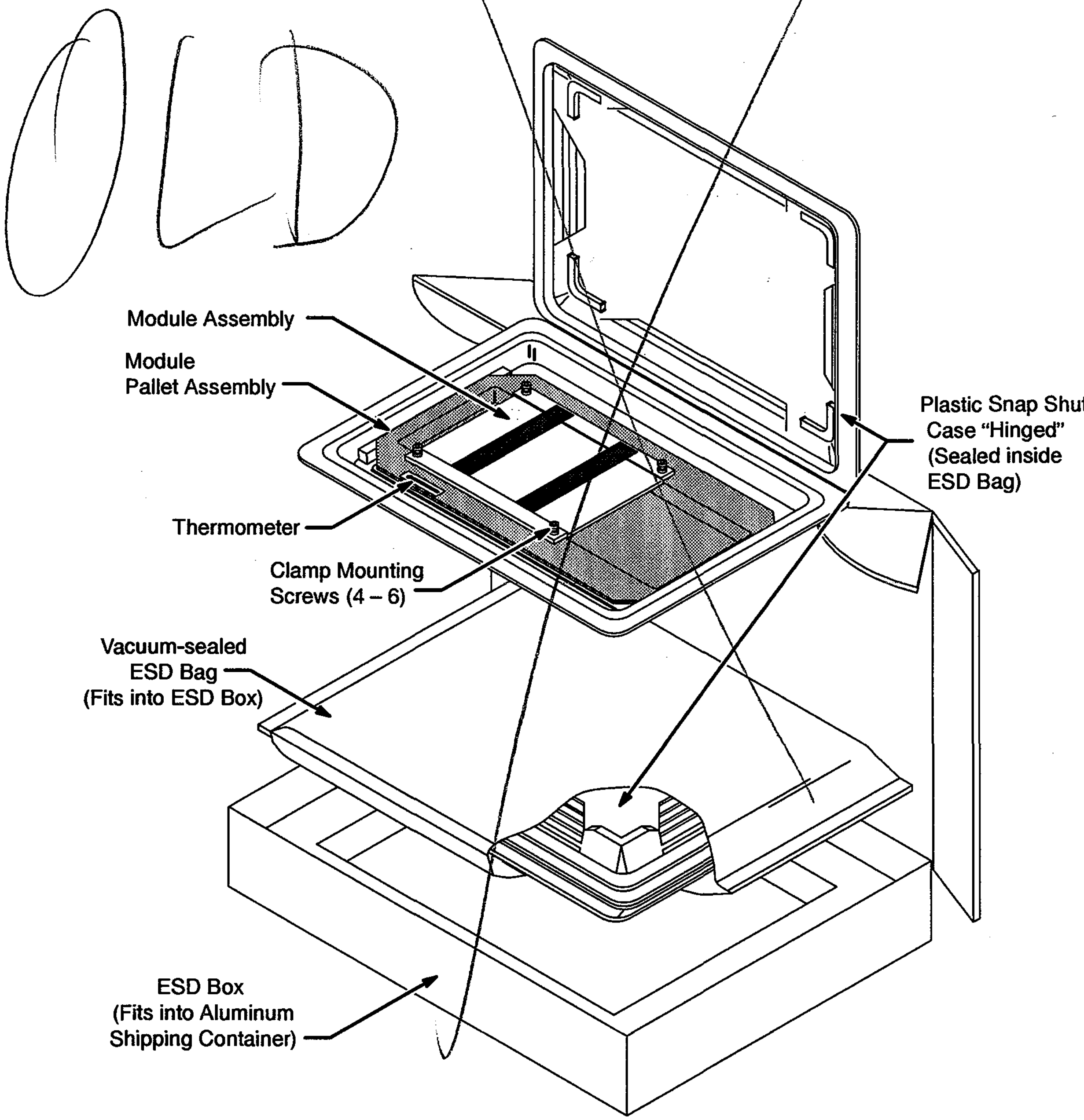


Figure 12. Module Shipping Container

The thermometer attached to the module pallet assembly should be checked before breaking the seal on the ESD bag. The module should reach room temperature, within + or – 5 °F/, before you open the sealed ESD bag. Temperature change should be no more than 10 °F per hour.

PC Board Technology

The CRAY T90 series modules are built with 22-, 52-, and 54-layer printed circuit boards. NW and CP boards have 54 layers. The BS module has one 22-layer board. Boards used in other modules in the system have 52 layers.

The 52-layer printed circuit boards consist of the following three individually fabricated and tested boards:

- One 8-layer printed circuit board (BDA), which contains the following features:
    - Two power/ground pairs
    - Two surface layers/ground pairs
- Two 22-layer printed circuit boards (BD1 and BD2), which contain the following features:
    - Five vertical-horizontal signal pairs
    - Five power-ground pairs
    - Two surface layers/ground pairs

The primary purpose of the 8-layer PCB is power distribution and 'Z' directional interconnects. This board serves as the core of the 3-board assembly. The two 22-layer PCBs serve as the logic signal transmission boards, and they provide some power/ground distribution. The cell arrays, connectors, and capacitors are mounted to these boards.

Module-to-Module Connectors

CRAY T90 series systems use the following types of module-to-module connectors:

- OIMs D/D and D/L
- SIMs

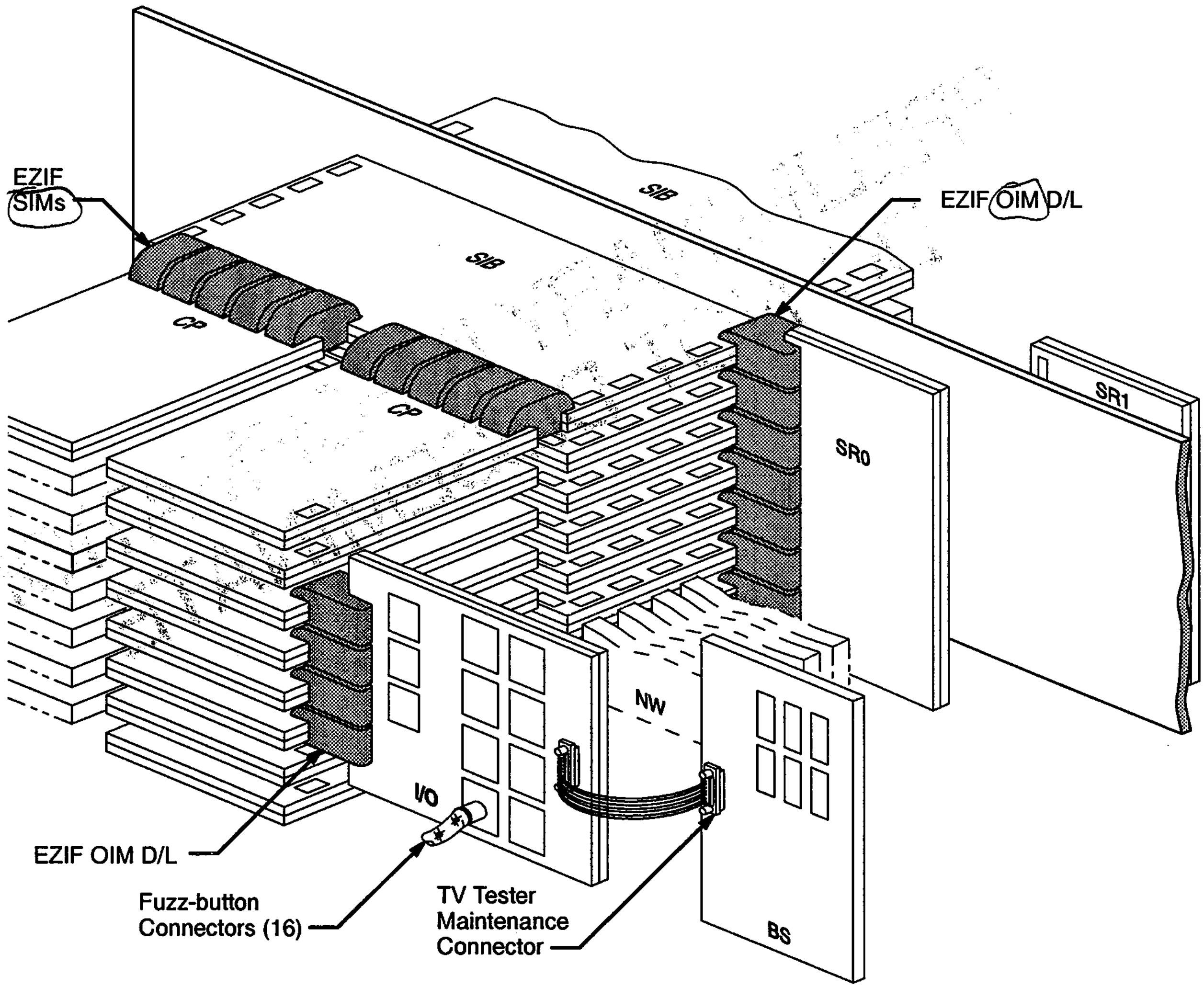


Figure 13. Module-to-Module Connector Types

OIM Connectors

Orthogonal interconnect module (OIM) connectors are used to connect modules that are linearly rotated 90 degrees to each other.

There are two types of OIM connectors:

- OIM D/D – direct-direct OIM
- OIM D/L – direct-lateral OIM (also known as OIM direct-slide-by)

Modules that use the OIM D/D connector plug into each end of the connector. With OIM D/L connectors, one module plugs into one end of the connector, while the other module must slide through the entire row of connectors before seating in place.

OIM D/L connectors are used in CRAY T916 and CRAY T932 systems. OIM D/L connectors are also used between the four CP modules and the I/O module in CRAY T94 systems.

SIM Connectors

Straight interconnect module (SIM) connectors are used to connect modules that are oriented in the same plane. All SIM connectors are the direct-direct type, denoted SIM D/D.

T-rail Connector Assemblies

OIM and SIM connectors attach to T-rail assemblies as shown in Figure 14. A printed circuit board is mounted to each T-rail assembly. Each CM/CP T-rail assembly contains five OIM D/D EZIF connectors, which provide interconnect paths between the CM, CP, and SR modules. The IO/CP T-rail assembly contains four OIM D/L EZIF connectors, which provide interconnect paths between the I/O and CP modules.

CRAY T94 systems contain four CM/CP T-rail assemblies and one IO/CP T-rail assembly. All T-rail assemblies in CRAY T94 systems are accessed and replaced through the CP module hatch. Each quadrant in CRAY T916 and CRAY T932 systems contains eight CM/NW, NW/SIB, and CP/SIB EZIF T-rail assemblies and one IO/CP T-rail assembly.

If a module is removed from a stack or if a stack is not fully populated, a flow block must be installed. The flow block maintains the even distribution of Flourinert liquid throughout the system.

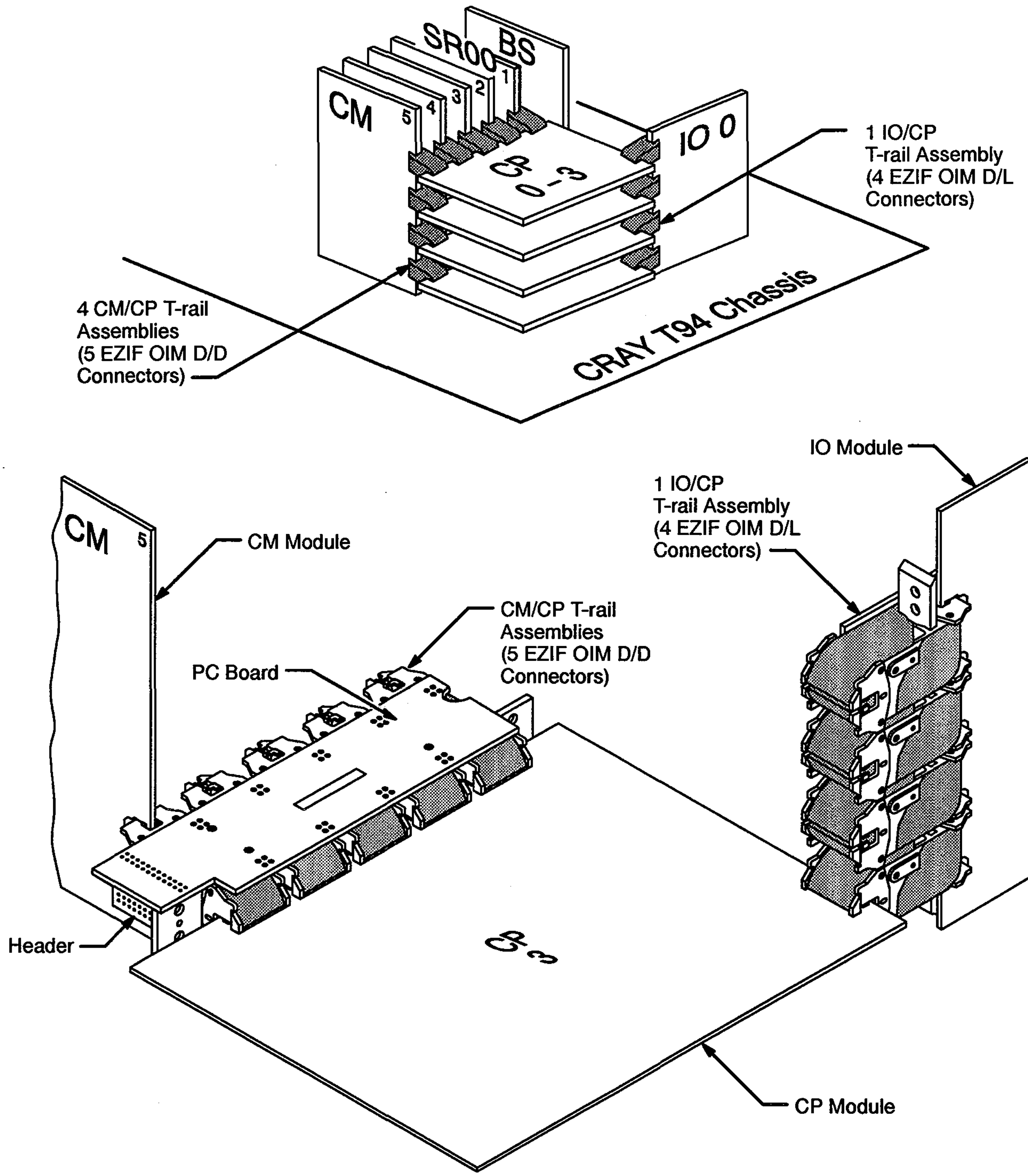


Figure 14. OIM T-rail Assemblies

Flex Circuit

Flex circuits are made of copper leads that are encapsulated in layers of a flexible insulating film. The flex circuit technology was employed in CRAY T90 series systems to accommodate higher signal speeds and the orientation of modules on different planes. The pin and socket connector technology used with prior CRI mainframe systems does not provide the electrical characteristics needed to run at the higher speeds used in CRAY T90 series systems.

OIM/SIM Connector Operation

Each OIM and SIM connector has an embedded heater-actuator spring (HAS) assembly to open and close the flex circuit jaws of the connector. Figure 15 illustrates the HAS assembly.

An EZIF heater controller device is designed to open and close different OIM and SIM connectors. The controller applies 12-Vdc current to maintain a preset temperature to each connector on a T-rail assembly. The heater controller device also monitors the status of each T-rail connector to prevent damage to the heater elements in the HAS. The heater controller has a secondary mode that can be used to apply current to the secondary heater element in the HAS, in case the primary heater element fails.

NOTE: The secondary heater can be used only once (to remove a module). Once the module has been removed, the faulty connector must be replaced.

When 12-Vdc current is applied to the HAS, the internal resistance causes the temperature of the primary heater element to rise gradually. Heat is transferred by conduction from the heater element to the actuator. The actuator begins to expand at a temperature of 70 °C (158 °F), which starts to force open the inner and outer springs in the HAS. The opening and closing of the HAS causes the OIM/SIM connector to open and close. The HAS will open fully after current has been applied for approximately 2 minutes, when the temperature of the actuator reaches 100 °C (212 °F).

The actuator applies a force to the springs only when opening the connector. When current to the heater is turned off, the heater and actuator cool, and the outer spring forces the HAS to close.

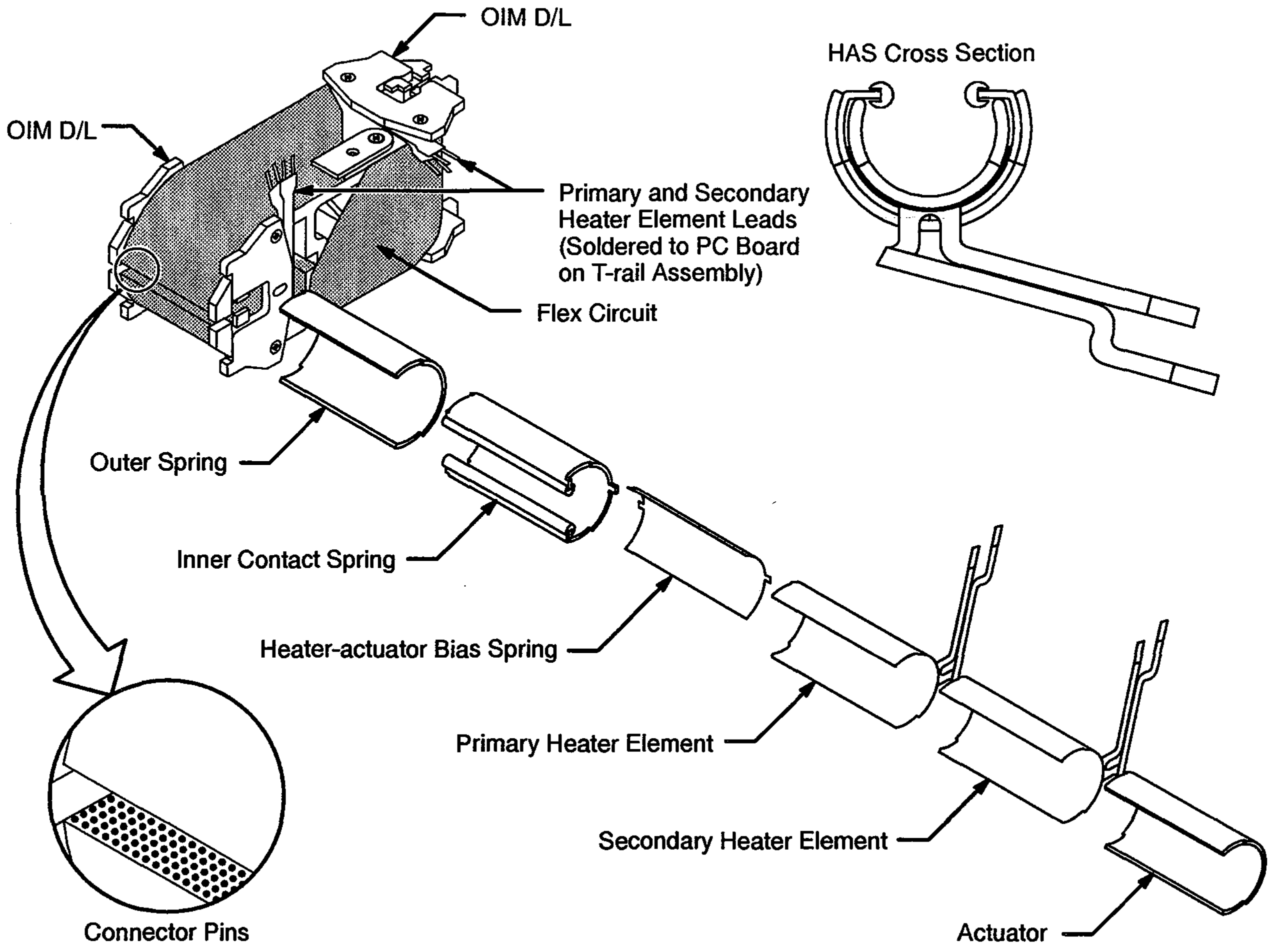


Figure 15. Heater-actuator Spring (HAS) Assembly

The external 12-Vdc temperature controller plugs into a heater-board assembly inside the chassis tank. Each heater board is labeled for the appropriate module type and location. From the heater board assembly, 12-Vdc wire harnesses connect the header on the T-rail assembly.

On CRAY T94 systems, the wire harness from the heater board assembly housing connects to a housing on the EZIF interconnect board (Figure 16). A harness from the EZIF interconnect board plugs into the header on the T-rail PC board. On CRAY T916 and CRAY T932 systems, the wire harness from the panel-mounted housing connects directly to the header on the T-rail PC board.

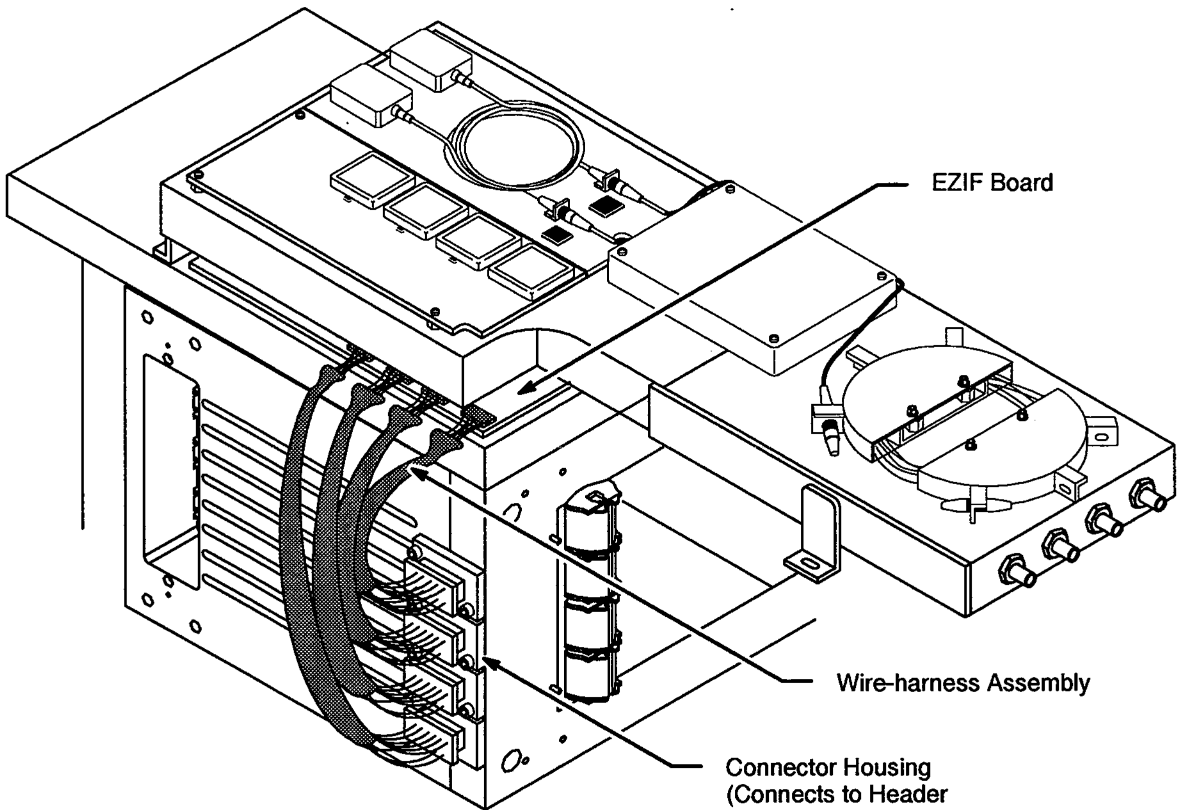


Figure 16. HAS Wire-harness Assembly

Fuzz-button Connectors

Data, address and control signals pass to or from the I/O module via a series of fuzz-button connectors. Figure 17 illustrates a fuzz-button cable assembly. There are 16 fuzz-button connectors on the I/O module; each connector has 514 pins. The cabling from the fuzz-button connector divides into several micro-D connectors, which connect to the I/O bulkhead. The entire fuzz-button cable assembly resides within the tank.

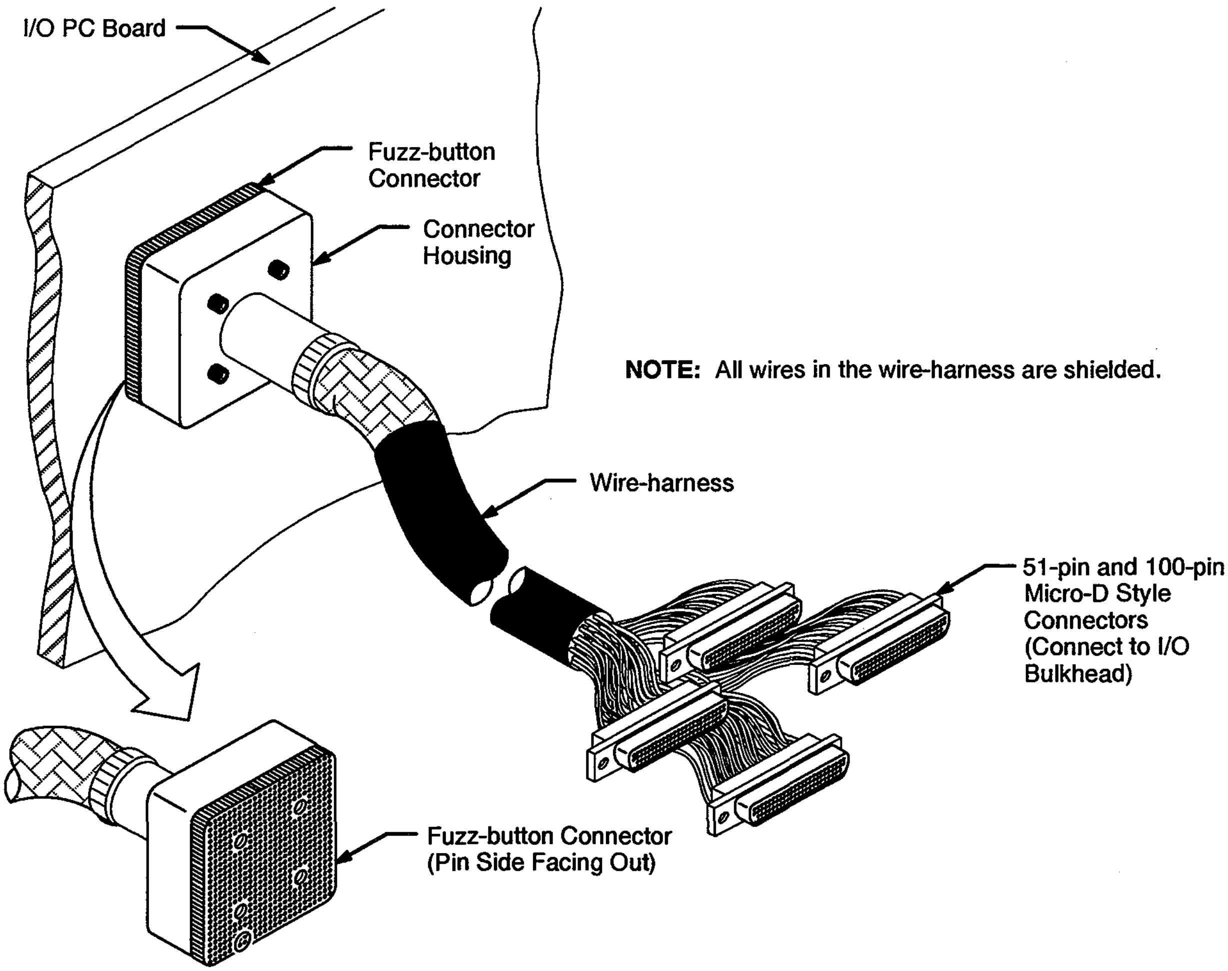


Figure 17. Fuzz-button Connector Assembly

Maintenance Connector

Each SR, I/O, CP, NW, and CM module has a maintenance connector. The maintenance connector is located on the power regulator of each module. The maintenance connector is used to send and receive the following signals to and from the module:

- Boundary scan signals – TCLK, TM, TDI, TDO, and RCLCK
- Boundary scan cable ground
- OIM and SIM connector interlock signals
- Continuity line signals
- Test-point inputs for connectors and test-point probing
- Thermal diode to output module temperature status

NOTE: The OIM maintenance connector on the BS module is used only in a test vehicle; it is called a TV maintenance connector.

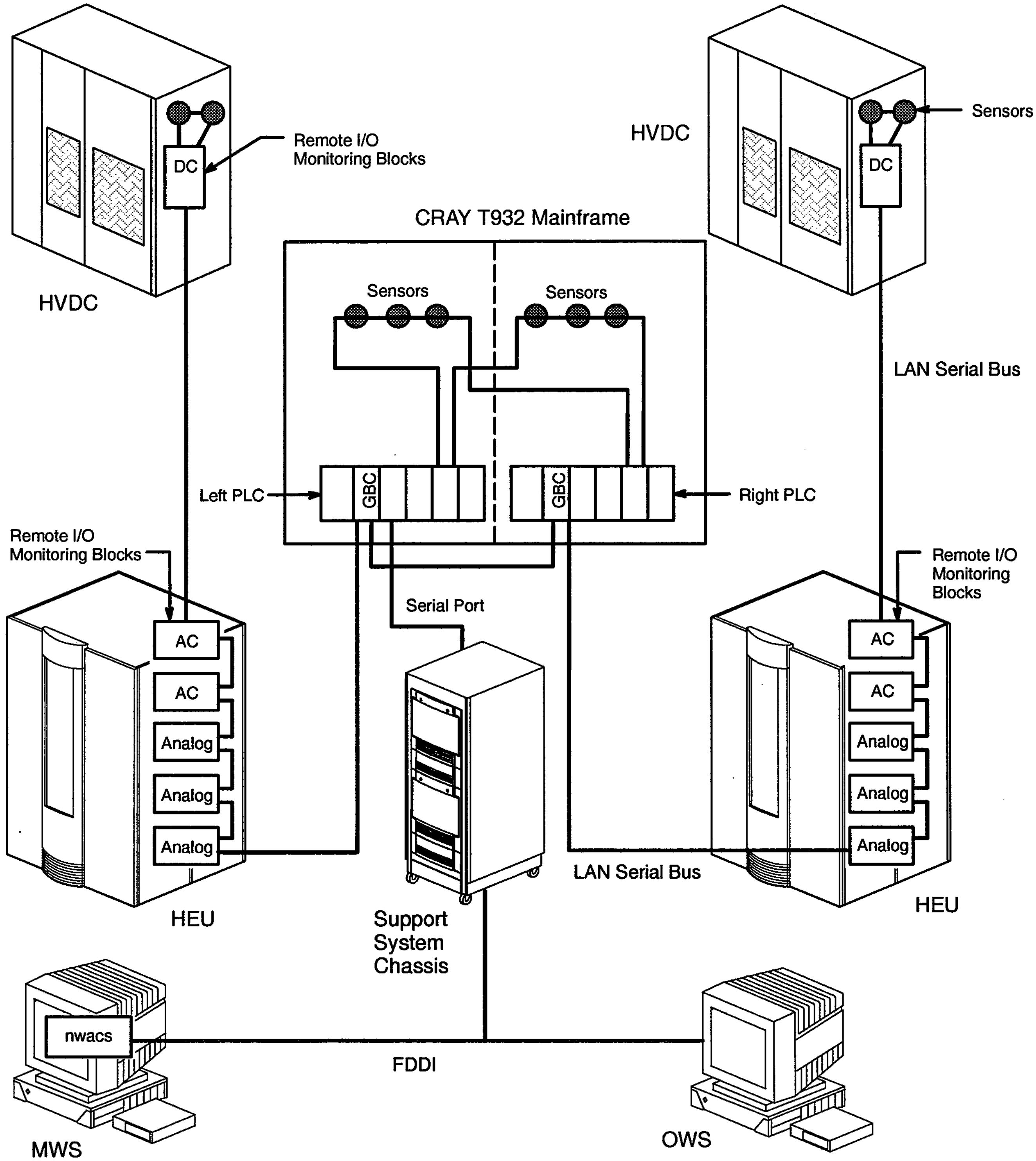
Thermal diodes embedded in options sense module temperature. The diodes send an analog signal through the maintenance connector to the PLC (programmable logic controller) in the control system. High-temperature signal warnings are sent to the PLC's memory where the `nwacs` program residing on the MWS can access this status information. The status information is available to the MWS by way of an RS-232 port from the PLC to an RS-232 card in the support system VME chassis.

Control System

The CRAY T90 series control system monitors various conditions within system components and, if necessary, shuts the system down if a monitored condition deviates from the predefined range. The control system is based on GE Fanuc 90-70 programmable logic controllers (PLCs) and the GE Fanuc Genius I/O System. Figure 18 is a block diagram of the CRAY T932 control system.

The control system consists of the following components:

- GE Fanuc 90-70 programmable logic controllers (PLCs)
- GE Fanuc Genius I/O System and communications system – analog, digital, and 115-Vac I/O remote monitoring blocks
- GE Fanuc Genius I/O System hand-held monitor
- MWS software – nwacs program
- Sensors



NOTE: A PLC in each side of the CRAY T932 mainframe controls half of the system. If one PLC malfunctions, the other PLC can be used to control the entire computer system. CRAY T94 and CRAY T916 systems each have a primary and secondary PLC, but both systems have only one HEU and one HVDC.

Figure 18. CRAY T932 Control System Block Diagram

Programmable Logic Controllers

Two programmable logic controllers (PLCs) monitor and control the computer system. Two PLCs (primary and standby) are used for system reliability and redundancy purposes. Each PLC is identical except that the redundant PLC does not have an analog output card, which enables power supply margining. If the primary PLC fails, the backup PLC controls the computer system until the primary PLC is repaired and operating.

The PLCs and related components (power supply, CPU, state logic processor, and other modules) are mounted in a nine-slot VMEbus rear-mount rack. This rack can hold ten PLC components: one power supply and nine other modules. Figure 19 shows the PLC.

The control system operates on a local area network (LAN) in which the PLCs are the central components (Figure 18). Each component within the system is daisy chained and communicates through the LAN serial bus.

On CRAY T94 systems, the two PLCs (primary and standby) are located next to each other on the right side of the mainframe. On CRAY T916 and CRAY T932 systems, the PLCs are located in the bottom of the chassis next to the I/O bulkhead. One is in quadrant 0 and the other is in quadrant 2.

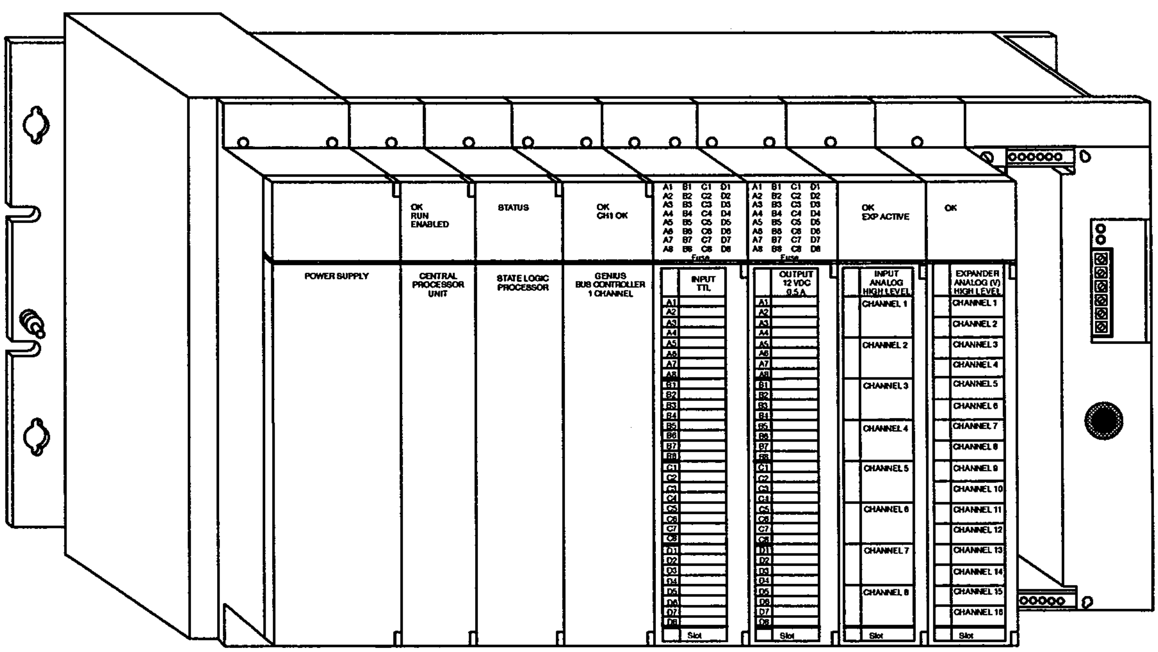


Figure 19. Programmable Logic Controller

I/O Remote Monitoring Blocks

Both PLCs receive data from the remote I/O remote monitoring blocks located in the HVDC and HEU units, and from the sensors located in the mainframe (Figure 18). However, the remote I/O monitoring blocks receive commands only from the primary PLC, unless the primary PLC does not report to the remote I/O monitoring blocks for three bus scan cycles. After three scan cycles, the remote I/O monitoring blocks switch over to the standby PLC.

The Genius bus controller (GBC) is a module in the PLC that provides the interface and control to the I/O remote monitoring blocks. The GBC sends I/O information along the VME backplane of the PLC to the CPU module. Any commands from the CPU module in the PLC are routed through the GBC to the remote monitoring blocks.

Sensors

The control system operation relies on the information sent back from the sensors, which are located in the mainframe chassis, HVDC, and HEU. Sensors monitor temperature, pressure, power bus voltages, power supply currents, and dielectric-coolant fluid levels and flow rates. The sensor location determines the routing of the sensor information.

The sensors located within the HEU and HVDC route information to the I/O remote monitoring blocks located on the device. The information is routed through the control system to the GBC modules in each PLC.

Status Displaying Software

The MWS runs a software program called `nwacs` to display the status information of the control system. The `nwacs` program is used to control the system; for example, to adjust voltage levels, dielectric-coolant flow rates, and control power-up and power-down processes.

Hand-held Monitor

The hand-held monitor (HHM) is used to troubleshoot and configure the control I/O system components. The hand-held monitor plugs into the GBC and I/O remote monitoring block monitor ports. Figure 20 shows the HMM.

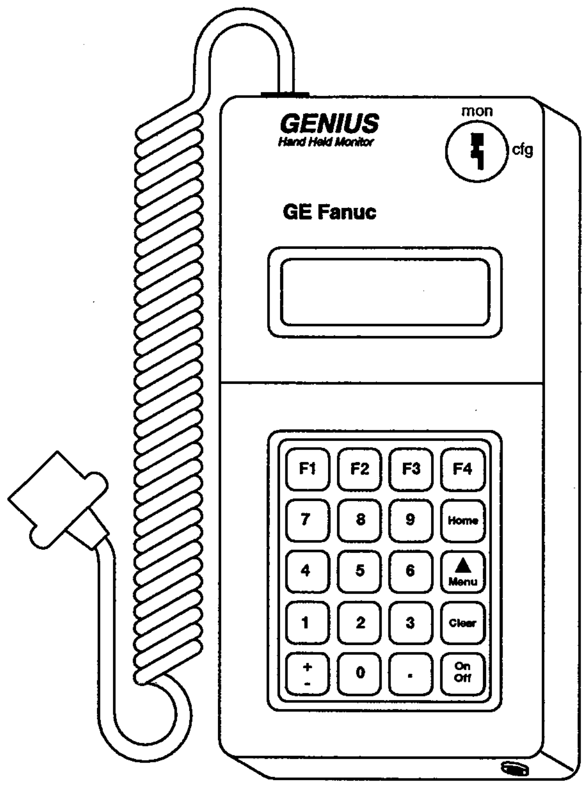


Figure 20. Hand-held Monitor

Support System

The support system is the point of first-level access, control, and status of a CRAY T90 series computer system.

The support system consists of the following components:

- Support system chassis
- MWS and OWS workstations
- Laser printer

Figure 21 shows the interconnection of the support system chassis components.

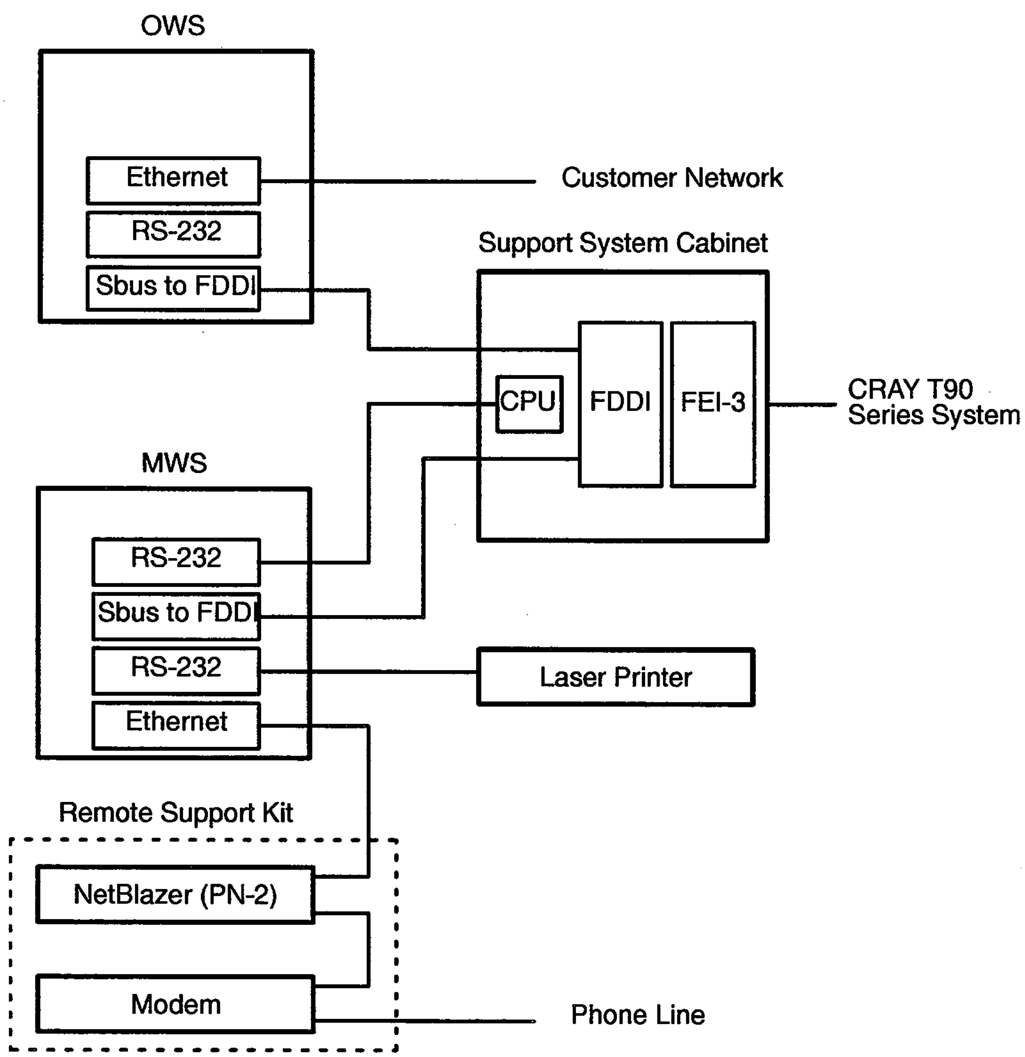


Figure 21. Support System Interconnections Block Diagram

Support System Chassis

The support system cabinet is a cabinet enclosure that contains up to two 20-slot VME chassis, one stacked above the other. CRAY T94 and CRAY T916 systems use only the upper chassis (the lower chassis is empty). CRAY T932 systems use both the upper and lower VME chassis. Figure 23 illustrates the support system chassis.

Each VME chassis is configured with a Motorola CPU, up to eight FEI-3 two-board sets, and an FDDI controller. A Motorola serial interface module provides eight RS-232-D serial channels on the CPU to provide maintenance connections to system disks and to the control system. A fan tray that contains three rotary fans provides airflow to cool the boards and power supplies in the VME chassis. Figure 22 shows the board layout in the support system chassis.

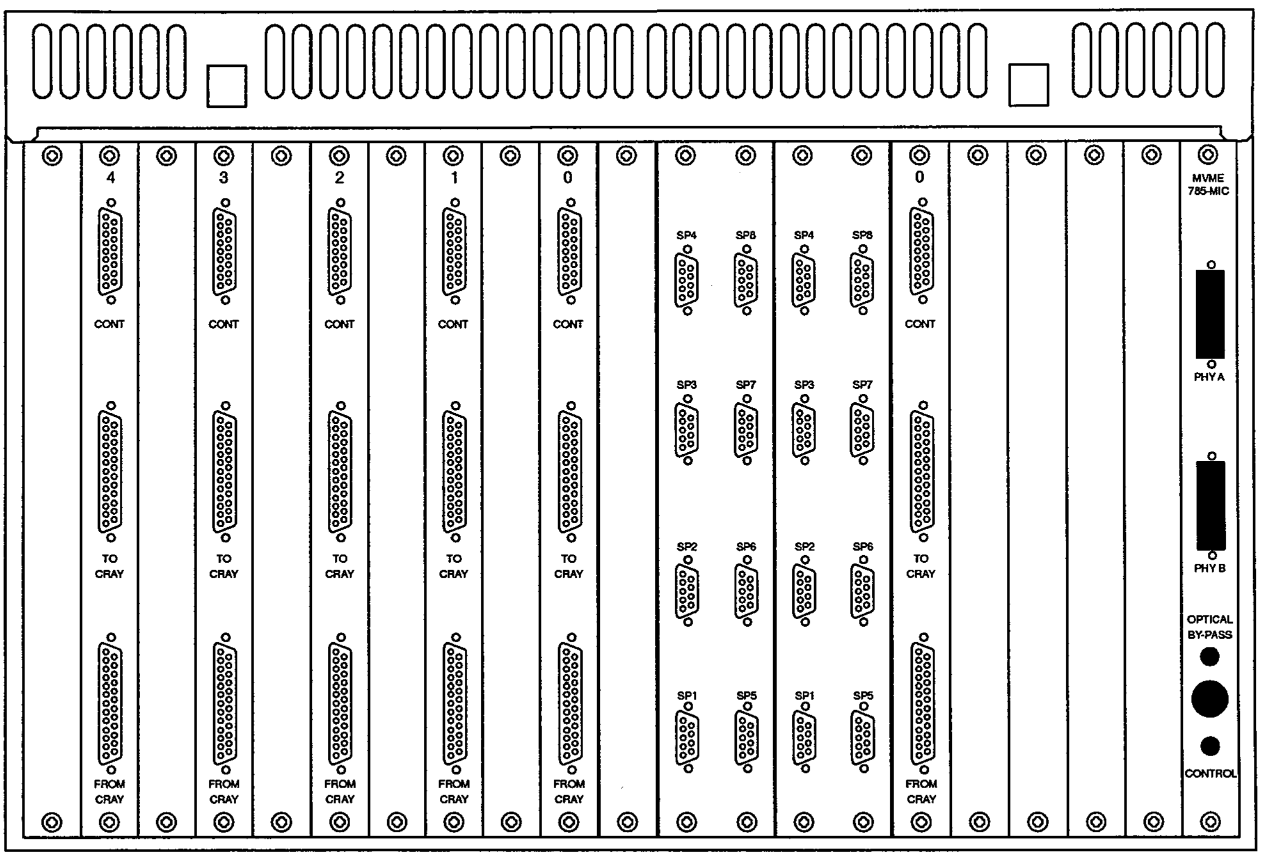


Figure 22. Board Layout for the Support System Chassis

The FDDI network is a 100-Mbit/s fiber-optic local area network that connects the MWS and OWS to the VMEbus interface in the support system cabinet. FDDI is a full-duplex network used as a high-performance interconnection among computers and peripheral equipment and as a high-speed backbone network for medium-performance local area networks (LANs).

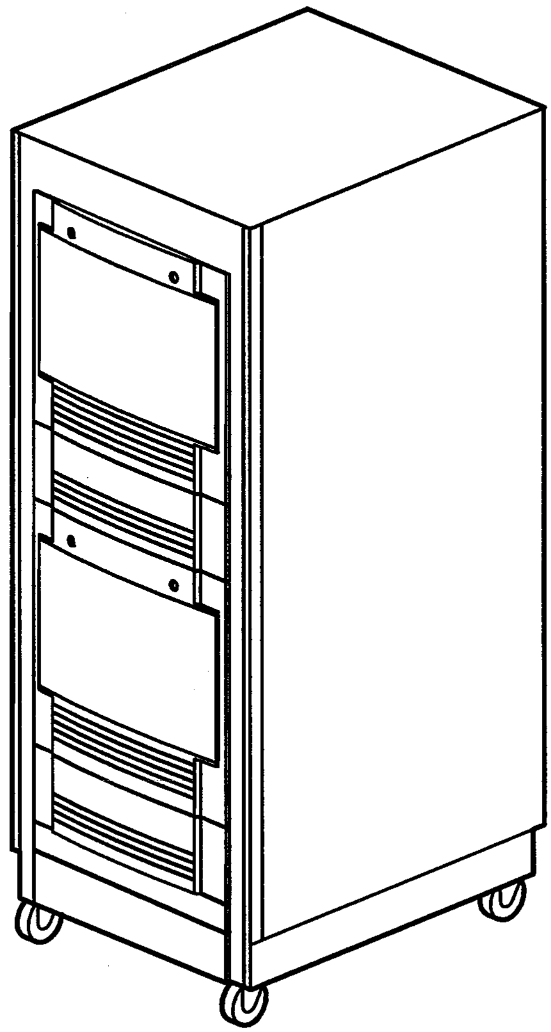


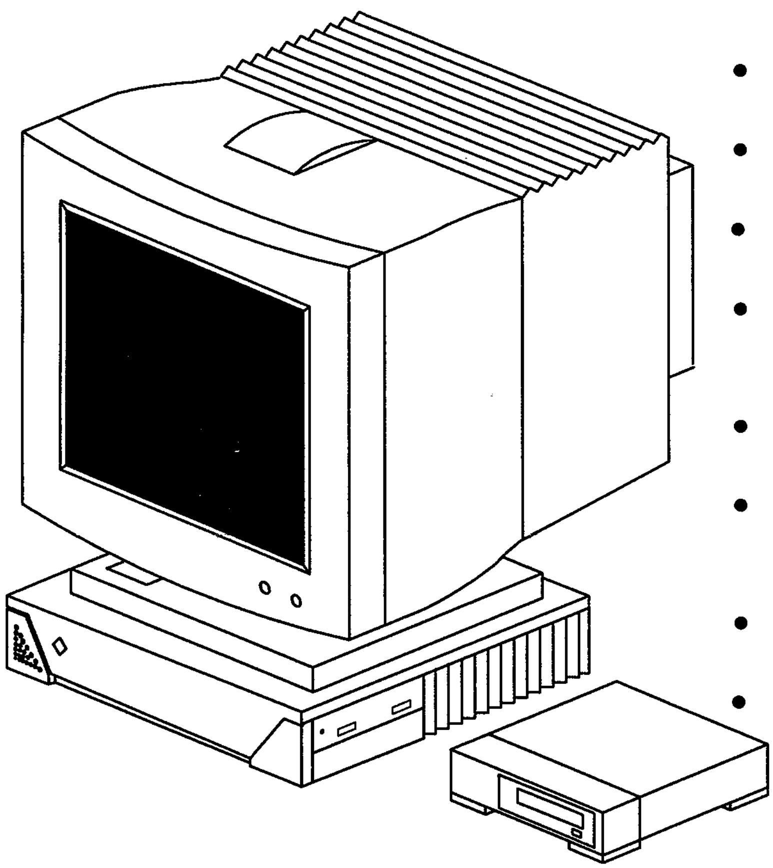
Figure 23. Support System Chassis

MWS and OWS Workstations

Cray Research uses Sun Microsystems, Inc. SPARCstation 5 workstations as the operator workstation (OWS) and the maintenance workstation (MWS) on CRAY T90 series systems. The MWS and the OWS consist of the same subassemblies and peripherals and are physically identical. For this reason, they have the same part number. Refer to Figure 24.

The FDDI node processor enables the MWS and OWS workstations and the 4-mm tape drive connected to the MWS to communicate with VMEbus-based devices and channels in the support system chassis.

To meet customer security requirements, Cray Research also offers a removable disk option. This option consists of a base unit and a disk module case that plugs into the base unit. With this option installed, the customer may remove the hard drive from the MWS and place it in a secure storage area when the hard drive is not in use.



- 70-MHz SPARC Version 8 CPU
- 32-Mbyte memory
- 1.05-Gbyte SCSI-2 hard drive
- 535-Mbyte SCSI-2 hard drive
- Internal CD-ROM player (SCSI-2 interface, double-speed)
- 17-in. color display
- Type 5 keyboard and optical 3-button mouse (not shown)
- FDDI node processor board
- Deskside 5-Gbyte, 4-mm tape drive

Figure 24. MWS/OWS Workstation

OWS Functions

The OWS is dedicated for the customer operator's use; the MWS is dedicated for use by Cray Research service personnel. However, in the event that one of these workstations fails, the other workstation could be used as a substitute until a replacement workstation is installed.

The OWS provides a dedicated workstation that customer operators and Cray Research analysts use to operate, administer, and monitor the CRAY T90 series system. The OWS is also used for system boot, dump, clear, and troubleshooting operations and for software support and upgrades.

MWS Functions

The MWS provides a dedicated platform for performing hardware maintenance, monitoring, and support of CRAY T90 series computer systems.

NOTE: The MWS, though connected to the CRAY T90 series system, is the property of Cray Research, Inc. The MWS is available exclusively for use by authorized service personnel.

The MWS performs the following functions:

- Providing system maintenance platform for performing maintenance and running diagnostics on all equipment in the computer system
- Configuring the mainframe
- Monitoring control system status
- Providing remote maintenance and control interface
- Detecting and logging system error information
- Providing an OWS-compatible platform
- Providing a platform for retrieving reference documents and for training

Laser Printer

A 300-dpi print quality laser printer is connected to an MWS port, but it is also accessible for printing from the OWS.

Support Channel

The support channel was designed to provide access to the operating system from the OWS. It is used for file system transfers and general administrative functions.

Each I/O module can have a single support channel; however, there can be only one support channel on either side of the tank divider. CRAY T932 systems can have two support channels and CRAY T94 and CRAY T916 systems can have one.

Support channels are not assigned to a particular CPU for data and control routing as are other I/O channels. SR module configuration function codes 10 through 26 are used to indicate which CPU is used to control routing. The route path selected on the SR module must match the path selected in the I/O module, which establishes data and control routing. No errors are returned if selections do not correspond.

Maintenance Features

Many new maintenance features have been built into CRAY T90 series systems, giving them an MTTI that is projected to be extremely high. Concurrent maintenance and redundancy are designed to greatly reduce customer downtime.

Many concurrent maintenance features are built into the system, such as spare memory chips, the ability to down processors, the use of sanity codes to reconfigure a system, and a physical boundary on CRAY T932 systems that allows half of the mainframe to remain operational while the other half is maintained.

The following list describes important maintenance features:

Maintenance Channel

In addition, there is also a maintenance channel that enables controlling the system configuration; through this channel, the service engineer can access up to 128 test points. The CRAY T90 series system also incorporates a boundary scan feature that enables the system to check the chip-to-chip and module-to-module interconnect paths and report which paths are failing. This feature also enables the system to read the chip type, revision level, and serial number of all the options.

No Hardware Switches

A CRAY T90 series system does not have hardware switches. All aspects of power-up and power-down, and diagnostic maintenance and monitoring processes are performed through the use of the monitoring and control system, which is controlled through the MWS.

System Interconnect Board

There is no longer a wire mat; it has been replaced with a printed circuit board called a system interconnect board (SIB). To reduce costs, only one type of SIB is used. Any changes to system configuration will be handled on the modules or with software.

Spare Memory Chips

Common memory modules contain spare memory chips. This allows us to flaw memory and provide normal memory operation, without physical maintenance, in the event of a memory chip failure.

Chip and Module Data

Chip type/revision and module type/serial number data is available from all modules by use of the boundary scan feature.

P Register

Each CPU's P register address can be read back to the OWS/MWS through the logic monitor.

Each CPU can be set to trap on a P register address, an issued instruction, or any test point.

System Configuration

System configuration information can be read from the mainframe.

Diagnostics

Instruction buffer load and dump functions are available for diagnostics.

Exchange load and dump functions are available for diagnostics.

Boundary Scan

Cray Research implements a boundary scan feature on CRAY T90 series systems. Boundary scan provides the following advantages over previous diagnostic and edge-connector test-vehicle techniques:

- Overcomes the difficultly or impossibility of accessing and probing pins and test points
- Shortens chip and module test and checkout times
- Reduces development and production times and costs
- Improves overall system quality and reliability

Boundary scan testing can test any interconnection in the machine, except memory chips pins (between boundary scan output and input cells) for proper connection. Boundary scan has the ability to detect shorts between option interconnections, shorts between interconnections and power/ground grids, or open foils in the interconnections. Boundary scan cannot test memory chips or test in the I/O channels because there are no boundary scan capabilities built into interconnections.

All modules in the mainframe support boundary scan. This enables most connections within the module and connections between the modules to be tested without diagnostics. Boundary scan also has a module and an option identification function that identifies the module and option type.

Boundary scan is run whenever a failure occurs that shuts down the system, and it is run to verify the integrity of the system after a repair procedure has been completed. Boundary scan requires control of the system and, therefore, cannot run simultaneously with the operating system. However, CRAY T932 systems can be degraded with boundary scan being run on one half of the system and the operating system running on the other half of the system.

Boundary Scan Register

Each option in CRAY T90 series mainframes contains a chain of shift-register based cells around the periphery of the option. These cells are the boundary scan register (BSR). There is one BSR location (or bit) for each input/output pad on the option. In addition, there are 16 additional bits used to store an option identification number.

In normal system operation, data passes between option pins and logic as if the BSR were not there. When boundary scan is run, test data is shifted along the BSR shift-register path, which includes each option on a module and all the interconnections between the options. Each option has a connection or pin for test data in (TDI) and test data out (TDO). Test data is shifted through the BSR path from TDI to TDO on each option by two control signals: Test Clock (TCLK) and Test Mode (TM).

Continuity Line Sensing

Every module in a CRAY T90 series mainframe has a long, continuous metal line that runs near all of the options on the module. This metal line is called the continuity line (or sometimes referred to as the C-line or burnline). The continuity line is connected to the maintenance connector on each module. If an area of a module overheats, a segment of the continuity line opens and error information is sent to the BS module and on to the control system.

A *sense signal* is continuously sent through the continuity lines on each module and is sampled by the BS module options three times near the end of every half-signal period (512 kHz). The continuity line sense signal is a 1-MHz square wave. A square wave is used instead of a ground reference so that if the burn wire shorts to ground, the short can be detected.

A *vote* is taken from the three samples, and the majority vote is used as the sensed value. This voting system removes glitches and prevents noise from causing an unnecessary shutdown of the system.

Continuity Line Errors

If the BS module fails to receive a toggle of the continuity line sense signal every half-signal period, the BS module notifies the control system that a continuity line error occurred. When the control system receives the error signal, it starts a countdown timer. If the error indication goes away before the timer times out, the timer is reset, and nothing happens. If the timer times out, the system performs a shutdown sequence.

Burn Mask

Control registers on the BS module enable and disable continuity line sensing. These registers function as a burn mask. There are 49 burn mask bits in the BS module: 1 bit for each BS option port that enables/disables continuity line sensing on the system modules and 1 bit to enable/disable continuity line sensing on the BS module. The system configuration environment (SCE) program enables and disables burn mask bits.

Sanity Code

Sanity code is a 6-bit data pattern that is used to control data flow and interconnecting processes between modules. Sanity code performs three primary functions:

- It establishes paths for sending maintenance commands and data
- It enables communications to/from module connectors to provide module-to-module communication
- It establishes a return path for reporting errors

Sanity code is used to ensure that a logical path exists to all modules configured within the system. Sanity codes control all signals that cross module boundaries. A module must receive sanity code from another module before it can respond to signals or commands from the other module. A module must also return a sanity code to the other module.

A module must receive sanity code from another module for 96 consecutive clock periods to establish a connection. Once the connection is established, the sanity code generator rebroadcasts the 6-bit sanity code pattern throughout the tree every 6 clock periods. If a module loses sanity code from a particular port, it still may be receiving sanity code from another port; however, all access from that port is now disabled.

A module must continue to receive valid sanity code in order to remain in an active state in a sanity tree structure. The absence of valid sanity code to all ports on a module causes that module to go into a master clear state; all outputs from those modules are then ignored.

Sanity Trees

A sanity tree is a logical structure that provides a flexible and reliable control network. A sanity tree is a sanity code distribution network that is defined when the mainframe is configured.

The actual path or structure of the sanity tree is not rigid; its structure is configured by the operator. Once a tree structure is established, it remains set at the same structure until it is changed or reconfigured. Remember that configuration of the mainframe defines the order of the tree structure.

Power and Cooling Equipment

CRAY T90 series computer systems use different power and cooling systems than those used with previous CRI computer systems. CRAY T90 series systems use a high-voltage direct-current (HVDC) cabinet that provides power to all components in the system. Motor-generator sets (MGs) and refrigeration condensing units (RCUs) are not used (an RCU is used if a 700-series IOS/SSD is part of the system).

Unlike previous CRI systems, CRAY T90 series systems do not use a refrigerant. This eliminates the requirement for an RCU. Eliminating the RCU has many benefits ranging from reduced power consumption (up to 15% on large computer systems) to increased cooling system reliability.

The following power and cooling components are described in this subsection. An overview of the control system begins on page 43.

- HEU-T90
- HVDC
- UPS

HEU-T90

Mainframe modules and power supplies are submersed in dielectric coolant. The heat exchanger unit (HEU-T90) routes dielectric coolant through the CRAY T90 series mainframe and IOS/SSD chassis to absorb heat generated by the modules and power supplies. After the dielectric coolant absorbs the heat, it flows back to the HEU-T90 where the heat is transferred to customer-supplied water. Figure 25 shows the HEU-T90.

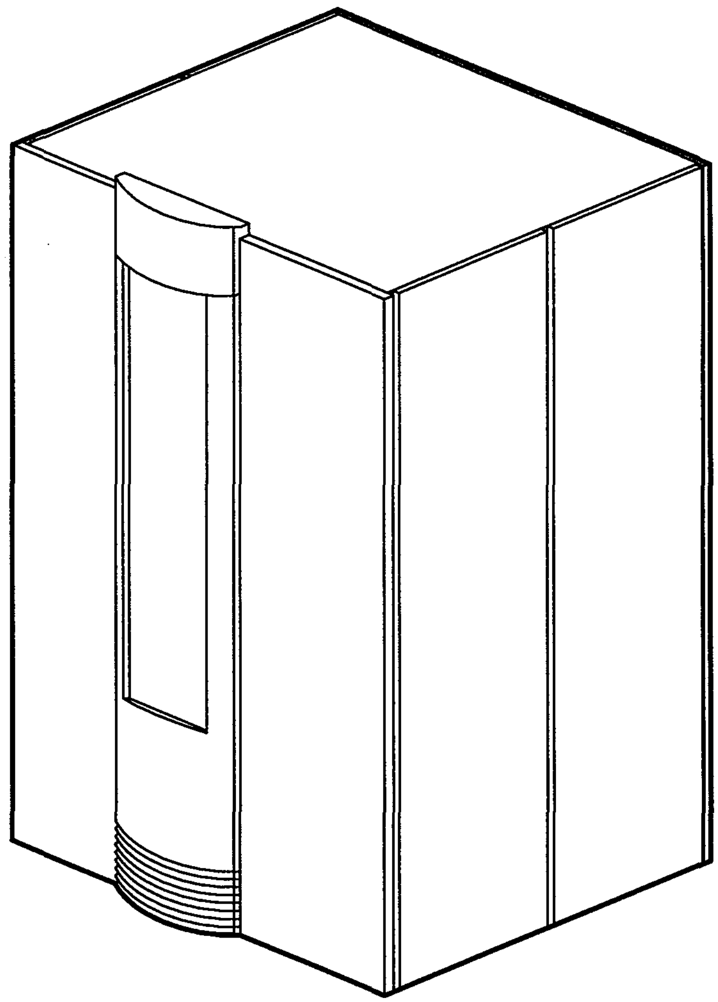


Figure 25. HEU-T90

CRAY T94 and CRAY T916 systems use a single HEU-T90. CRAY T932 systems require two HEU-T90s because of the physical boundary that allows one half of the machine to be powered off and drained while the other half is still functioning. Each HEU-T90 supports one-half of the chassis.

The HEU-T90 used with CRAY T94 systems has a smaller heat exchanger vessel than those in the HEU-T90s used with CRAY T916 and CRAY T932 systems. Other than a smaller heat exchanger vessel, the HEU-T90s shipped with each CRAY T90 series system are the same.

An RCU can be connected to an HEU-T90 when a 700-series IOS/SSD is used. In this situation, heat is exchanged to evaporative refrigerant (R-22) rather than to chilled water.

Control System

Control system monitors within the HEU-T90 ensure that temperatures and pressures fall within specific ranges. If the conditions are out of range, the control system either adjusts valves within the HEU-T90 to compensate for the out-of-range condition or shuts the computer system down to protect the equipment and computer room environment from damage. Refer to the "Control System" information that begins on page 43.

Level Sensors

Six level sensors monitor the level of dielectric coolant within the CRAY T90 series mainframe. Figure 26 shows level sensor locations in a CRAY T94 mainframe. These sensors are located on the same side as the supply and return lines and are used to fill and drain the mainframe. Two level sensors (primary and backup) measure the level of the dielectric coolant in the top area of the mainframe. Two additional level sensors (primary and backup) measure the dielectric-coolant level in the power supply area of the mainframe. Two more sensors in the drain area and cable drop area monitor the level of dielectric coolant present in the mainframe during the draining process.

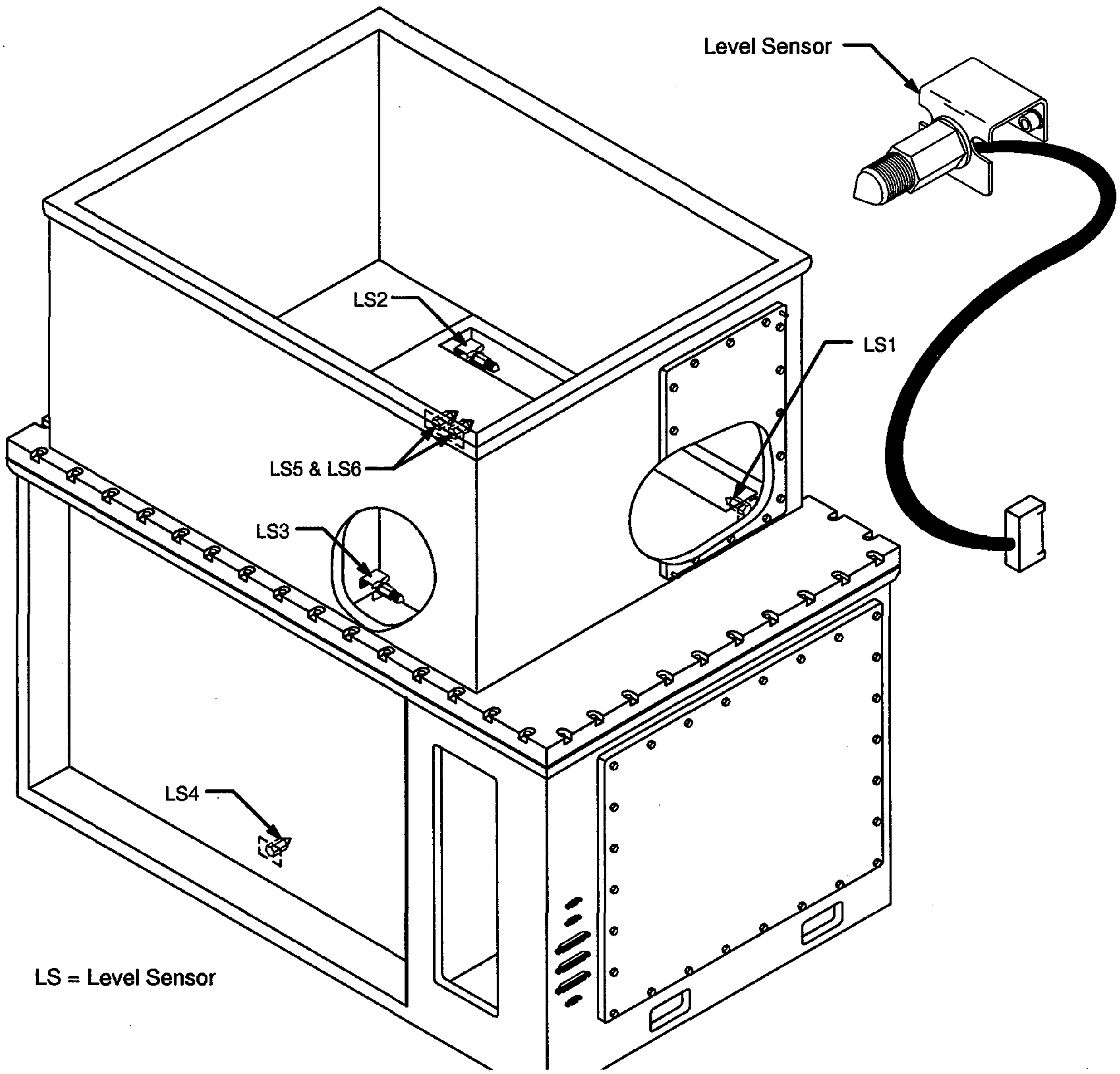


Figure 26. Level Sensor Location in a CRAY T94 Mainframe

HVDC

Power supplies in CRAY T90 series mainframes operate on 330-Vdc power. Commercial utility power at customer sites may vary from 200 Vac to 480 Vac at frequencies of 50 or 60 Hz. The high-voltage direct-current (HVDC) cabinet transforms any of these ranges of input voltages and frequencies. Figure 27 shows the HVDC systems. The HVDC transforms, rectifies, and regulates input power from the customer's utility to deliver conditioned 330-Vdc power to system components.

The following HVDC systems are used with CRAY T90 series systems:

- One HVDC-40 (40 kW) for CRAY T94 systems
- One HVDC-160 (160 kW) for CRAY T916 systems
- Two HVDC-160s for CRAY T932 systems

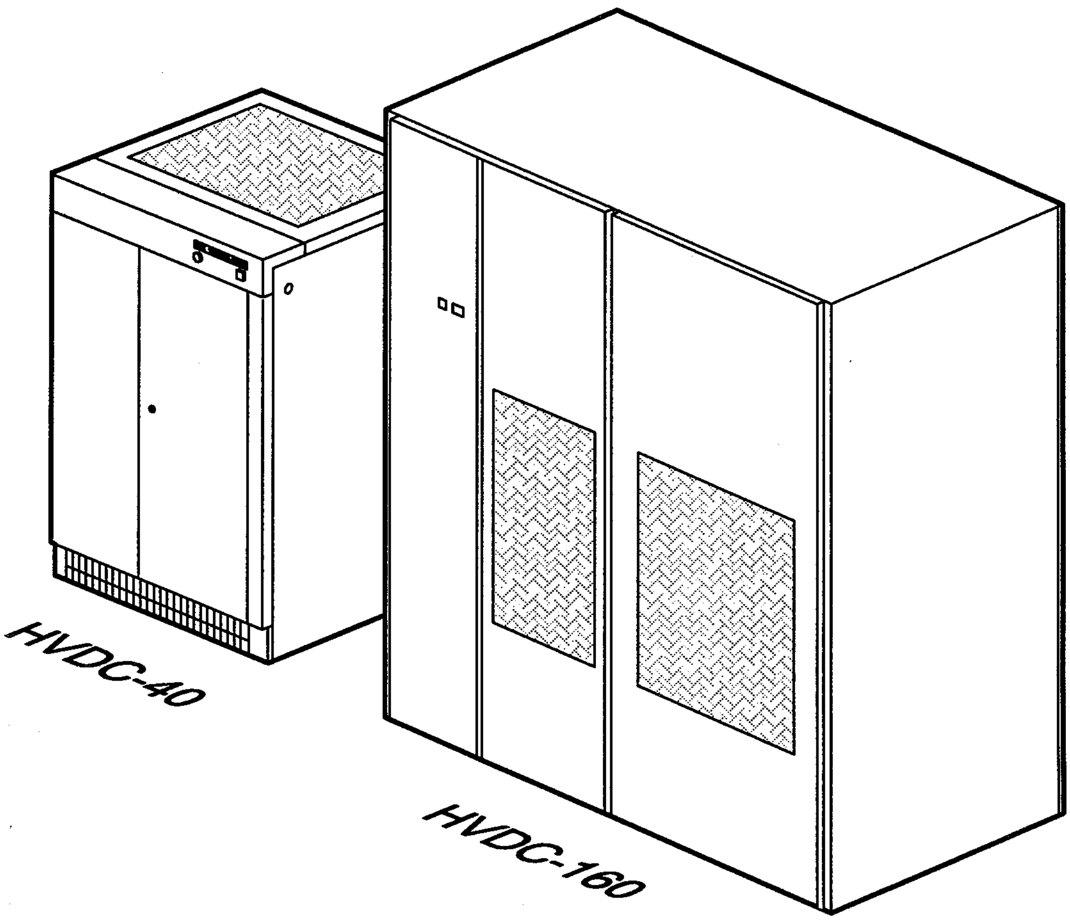


Figure 27. HVDC Cabinets

UPS

If CRAY T94 customers require power ride-through in excess of 80 milliseconds, CRI will provide an uninterruptible power supply (UPS). The capacity of the UPS is large enough to power the CRAY T94 mainframe, HEU-T90, IOS/SSD chassis, and peripherals. The UPS system for CRAY T94 systems is an optional piece of equipment that the customer may purchase.

CRI can supply a UPS with CRAY T916 and CRAY T932 systems, which is attached to the 160-kW HVDC. The UPS system is large enough to provide uninterruptible, conditioned power to all CRI system components, including the mainframe via the HVDC, workstations, IOS, peripherals, and the HEU(s).

The UPS used with a HVDC-160 is not a system-critical device. A bypass circuit on the UPS is available to enable service personnel to perform maintenance on the UPS while the CRAY T90 series system is in operation.

The UPS system used with the CRAY T90 series products is produced by EPE Technologies in the United States of America and Merlin Gerin in Europe. (Both EPE Technologies and Merlin Gerin are subsidiaries of Square D Corporation.)

The baseline UPS system supplied by CRI provides approximately 2 minutes of hold-up time. Customers have the option to buy additional battery cabinets to lengthen the hold-up time to the time they desire, with 9 minutes being a reasonable maximum. The UPS uses sealed, maintenance-free batteries that do not require any special environmental considerations. However, the batteries must reside in an ambient room temperature below 77 °F (25 °C).

Customer Use of Existing UPS Systems

Customers may connect their existing UPS systems to the HVDC-40 used with CRAY T94 systems. Customers cannot directly connect a UPS to a HVDC-160.

Customers may also connect other equipment within their facilities (such as network equipment and chilled water pumps) to the UPS that CRI provides, as long as the added load falls within the kW limitations of the UPS.

Service and Support for HVDC, UPS, and GE Fanuc Controller Equipment

HVDC equipment for CRAY T90 series systems will be self-supported by CRI. This means that CRI responds to all HVDC maintenance and service calls. CRI provides all repairs and preventive maintenance. Any HVDC situations that require escalated support should follow existing escalation procedures.

Trained CRI personnel will complete HVDC start-up as covered by the HVDC vendor's warranty. If CRI personnel are not available to provide HVDC start-up procedures, the HVDC vendor will complete start-up and charge CRI based on their time and material rates.

The HVDC-40 and HVDC-160 units require minimal preventive maintenance that consists of replacing air filters and verifying operating limits. The frequency for performing these procedures has not been defined at this time.

Preshipping Power and Cooling Equipment

HEUs are not preshipped. The only cooling system components that are preshipped are a set of chilled water connection flanges. These flanges connect flexible hoses from the HEU to the customer's chilled water loop.

CRI Site Planning does not plan on preshipping the HVDC-40 to customers in advance of the rest of the CRAY T94 system. CRI will preship the HVDC-160s used with CRAY T916 and CRAY T932 systems.

The HVDC-40 is quality checked with the CRAY T94 system in STCO and the entire system is shipped together. The HVDC-40 wiring consists of one input circuit and five output circuits. With proper preinstallation planning and preparation, the amount of time required to wire the HVDC-40 should be less than 4 hours. The installers should be able to power up CRAY T94 systems on the first day of system installation.

NOTE: A required MG set is preshipped to CRAY T90 series sites that have a 700-series IOS/SSD chassis.

Advantages to Using the HVDC and UPS

There are several advantages to using an HVDC rather than an MG set. The HVDC increases electrical efficiency and reduces required floor space and weight. The MGs were custom built specifically for CRI, which made the MG much more expensive than commercially available products like the HVDC and UPS.

The MG provides conditioned, isolated power only to the mainframe, IOS, and SSD. The other components in the system run on utility power, unless the customer provides conditioned power. The HVDC units provide conditioned, isolated power to all components in the system. The UPS significantly increases hold-up times. CRAY T90 series system availability should be increased because of the increased hold-up times and because conditioned power is supplied to all components in the system.

The UPS systems used with CRAY T90 series systems are standard products that are shipped with 5-week lead times. The custom-made MG set required a 26-week lead time.

IOS/SSD Equipment

The following types of IOS chassis are available with CRAY T90 series systems:

- 600 series with up to two I/O clusters and an SSD-E/128i
- 700 series with up to eight I/O clusters and four SSD-E sections
- 800 series with up to three I/O clusters and one SSD-E section

Figure 28 illustrates the 600- and 800-series IOS/SSD-E chassis. The first few CRAY T94 systems shipped in 1994 and 1995 will include a 600-series IOS/SSD; later systems will ship with 800-series IOS/SSDs. In both cases, the IOS chassis is cooled by the HEU-T90, eliminating the use of the RCU-8 or RCU-10.

A CRAY T916 system can support up to two IOS chassis, and a CRAY T932 system can support up to four IOS chassis.

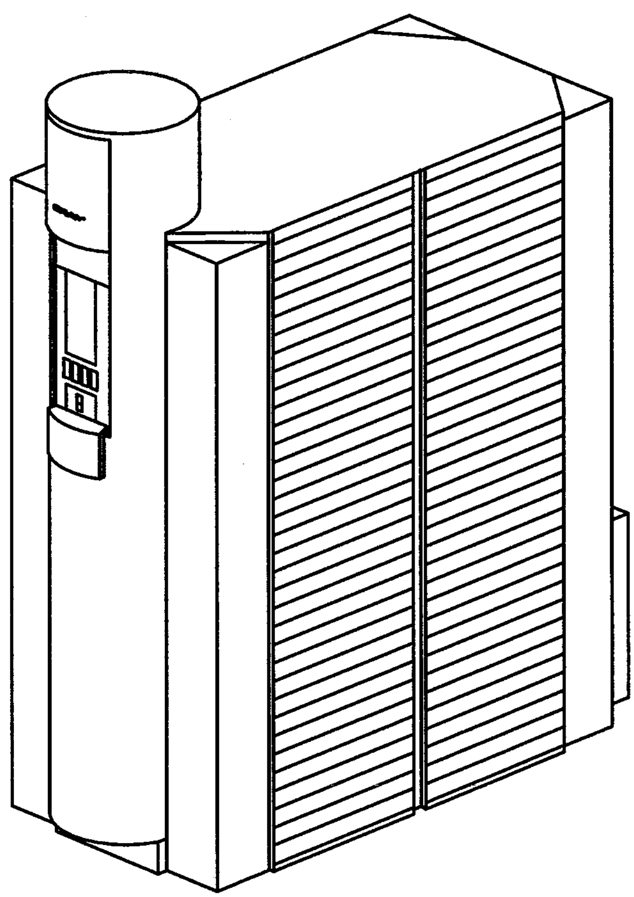


Figure 28. 600- or 800-series IOS/SSD-E Chassis

The 600- and 800-series cabinets have the same outer appearance. (The 600-series IOS/SSD is shipped as a single unit.) The 600-series IOS/SSD-E shipped with the CRAY T94 system uses 10 slots to house two I/O clusters and an SSD-E/128i. The 800-series IOS/SSD-E contains 32 slots to house three I/O clusters and a 1-section SSD-E.

CRAY T916 and CRAY T932 systems are likely to be configured with the 700-series IOS chassis. With the 700-series IOS, a water-cooled HEU is provided.

There may be cases in which CRAY T916 and CRAY T932 systems are configured with a 600- or 800-series IOS chassis. For CRAY T916 and CRAY T932 systems using an 800-series IOS, an RCU-8 or RCU-10 must be used because the HEU-T90 does not provide sufficient heat-rejection capacity. The rear of the 800-series IOS chassis contains different components to exhaust warm air and to connect to the power inlet and coolant hoses to the HEU.

Documentation and Information Sources

The following subsections describe the CRAY T90 series documentation sources.

Online Documentation

All CRAY T90 series documentation produced by the Hardware Publications and Training department (HPT) is available online through the online documentation interface (ODIE). ODIE provides a standard, easy-to-use interface for locating online information. HPT will be using the ODIE interface for all online documentation projects. CRAY T90 series online documentation will be available to students attending classes in Chippewa Falls. HPT plans to provide online documentation on CD-ROM to all CRAY T90 series sites.

Service and Support Plan

The Service Planning and Engineering department writes Service and Support Plans for each major Cray Research product. Logistics stocks these plans, and field personnel can order them as needed. Each plan is an index to the various service and support features for the product. The Service and Support Plans do not replace CRI technical instructions. They are intended to provide insight into CRAY T90 series site service and support functions.

CRAY T90 Series Mainframe Specifications

System Clock

Speed .... 2 ns

CPU Specifications

Number of CPUs .... 1 to 32

Number of registers per CPU:
- Address (A) registers 8
  - 64 bits each (32 bits in C90 mode)
- Scalar (S) registers 8
  - 64 bits each
- Intermediate address (B) registers .. 64
  - 64 bits each
- Intermediate scalar (T) registers ... 64
  - 64 bits each
- Vector (V) registers 8
  - 64 bits x 128 elements (two pipes of even and odd elements)
- Vector length (VL) register ....
  - 8 bits
- Vector mask (VM0, VM1) registers ...
  - 64-bits each
- Program address (P) register ....
  - 32 bits

Functional units per CPU:
- Address addition ....
- Address multiplication ....
- Scalar addition ....
- Scalar shift ....
- Scalar logical ....
- Scalar population/parity/leading zero .
- Vector addition ....
- Vector shift ....
- Vector population/parity/leading zero .
- Vector integer multiplication ....
- Full vector logical ....
- 2nd vector logical ....
- Vector population/parity/leading zero .
- Floating-point addition ....
- Floating-point multiplication ....
- Floating-point reciprocal approx. ....
- Bit Matrix Multiply ....
- Iota ....

Shared Resources

I/O section:

Channels per I/O module:
- 6-Mbyte/s LOSP channels 8
- 200-Mbytes/s HISP channels 8
- 1800-Mbyte/s VHISP channels 4
- LOSPX channels 4

Central memory:
- Word width .... 64 bits
- SBCDBD error correction .... 12 bits
- Memory size .... Up to 1 Gwords
- Number of banks .... Up to 512
- Number of modules .... Up to32
- Number of ports per CPU 8

Number of clusters .... Up to 36

Number of shared registers contained in each cluster:
- Shared address (SB) registers 16
  - 64 bits each
- Shared scalar (ST) registers 16
  - 64 bits each
- Semaphore (SM) registers 64
  - 1 bit each

Real-time clock (64-bits) ....

Physical Description

CRAY T94 system
- Height . . . . 51.75 in. (131.4 cm)
- Width . . . . 67.60 in. (171.7 cm)
- Depth . . . . 37.00 in. (94.0 cm)
- Weight . . . . . . 3,000 lbs (1,360 kg)

CRAY T916 system
- Height . . . . 85.5 in. (217 cm)
- Width . . . . 116.25 in. (295 cm)
- Depth . . . . 101.00 in. (257 cm)
- Weight . . . . . . 16,616 lbs (7,536 kg)

CRAY T932 system
- Height . . . . 85.5 in. (217 cm)
- Width . . . . 116.25 in. (295 cm)
- Depth . . . . 101.00 in. (257 cm)
- Weight . . . . . . 16,616 lbs (7,536 kg)

Access Requirement
- CRAY T94 system . . 36.0 in. (94.0 cm) on all four sides

Power Requirements

Input Voltage
- Varies (50/60 Hz. . . 200 Vac to 480 Vac)

CRAY T94 system:
- Two 330-Vdc power circuits
- One 120-Vac control circuit
- All supplied by one 40-kW EPE systems HVDC

Cooling

Submersion cooling in Fluorinert liquid. Heat exchange between the Fluorinert liquid and chilled water.

Reader Comment Form

Title: CRAY T90™ Series System Overview **Number: HGM-xxx-x**
Preliminary Information **December 1994**

Your feedback on this publication will help us provide better documentation in the future. Please take a moment to answer the few questions below.

For what purpose did you primarily use this document?

______Troubleshooting
______Tutorial or introduction
______Reference information
______Classroom use
______Other - please explain ____________________

Using a scale from 1 (poor) to 10 (excellent), please rate this document on the following criteria and explain your ratings:

______Accuracy ____________________
______Organization ____________________
______Readability ____________________
______Physical qualities (binding, printing, page layout) ____________________
______Amount of diagrams and photos ____________________
______Quality of diagrams and photos ____________________

Completeness (Check one)

______Too much information ____________________

______Too little information ____________________

______Just the right amount of information

Your comments help Hardware Publications and Training improve the quality and usefulness of your publications. Please use the space provided below to share your comments with us. When possible, please give specific page and paragraph references. We will respond to your comments in writing within 48 hours.

NAME ____________________
JOB TITLE ____________________
FIRM ____________________
ADDRESS ____________________
CITY ____________ STATE ________ ZIP ________
DATE ____________________

[or attach your business card]

CRAY RESEARCH, INC.

CUT ALONG THIS LINE

Fold

NO POSTAGE NECESSARY IF MAILED IN THE UNITED STATES

BUSINESS REPLY CARD

FIRST CLASS PERMIT NO 6184 ST. PAUL, MN

POSTAGE WILL BE PAID BY ADDRESSEE

CRAY RESEARCH, INC.

Attn: Hardware Publications and Training
890 Industrial Boulevard
Chippewa Falls, WI 54729

Fold

STAPLE