CRAY Y-MP EL Computer Systems Hardware Troubleshooting and Maintenance Manual

DAS - Disk Array Subsystem
DAC - Disk Array controller
ESDI - Enhanced serial Drive interface
PE - peripheral enclosure
DIEB - Disk Enclosure Box
9 DD 3s (maybe 10)

Small computer system Interface

SCSI
(scuzzy)

Application specific IC
ASIC
100K gates

CMM-xxxx-PR2

Cray Research, Inc.

→ DAS
VME BUS controller
(BUS master)
↓
DAC
↓

PREFACE

To be provided.

CONTENTS

FIGURES

FIGURES (continued)

TABLES

1 SYSTEM OVERVIEW

The CRAY Y-MP EL computer system is a completely self-contained system. This means that the central processing unit (CPU), input/output subsystem (IOS), and all peripherals are contained in one easily moved cabinet. The cabinet can be moved easily to most general office environments equipped with a 200- to 250-Vac power source and be operational.

One to four CPU boards can be installed in the CRAY Y-MP EL system cabinet. This cabinet also holds four memory boards, an IOS contained within a VMEbus-based assembly, and the optional peripherals requested by the customer. Another area in the frame contains the small computer system interface (SCSI) subsystem. The SCSI subsystem consists of a cartridge-type streaming tape drive, a 780-Mbyte hard disk drive, and an optional 8-mm helical scan tape drive. This SCSI subsystem is analogous to the expander chassis that formed part of the CRAY X-MP computer system. The SCSI is used to install any new releases of software as they become available. It also has the capability of performing data transfers and being the system backup and boot device.

The primary frame is constructed so that another standard frame can be bolted to the primary frame to expand the system. It is possible to connect as many as three secondary frames to the primary frame; thus, the CRAY Y-MP EL system can contain as many as 16 IOSs and additional peripheral devices.

Figure 1-1 and Figure 1-2 can be used to locate the four cabinets that can be configured and the various subassemblies within the CRAY Y-MP EL system.

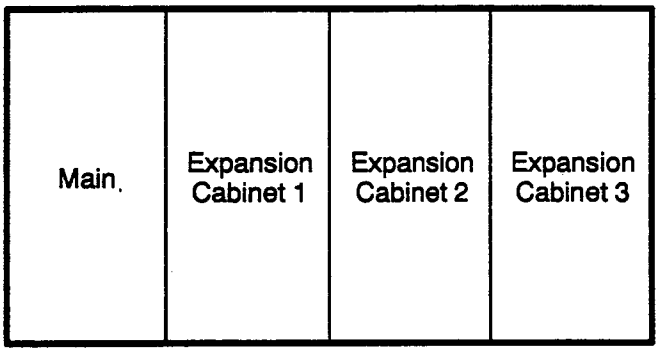


Figure 1-1. Left Side View (Maximum configuration)

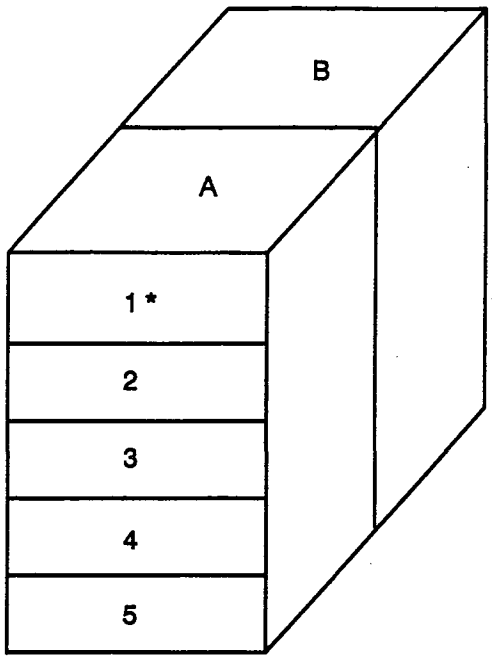


Figure 1-2. Front-Right View, Cabinet A

All of the frames forming the CRAY Y-MP EL system are air cooled by integrated fans at both the top and bottom of the frame. This form of cooling is referred to as vertical cooling.

The peripheral devices available on the CRAY Y-MP EL system and their characteristics are listed below.

- DS-3 Disk Storage System
  - 5.25-inch format Winchester disk drive
  - Holds 1.5 Gbytes of unformatted data per drive
  - Transfer rates are between 1 and 1.75 Mbytes/s
  - Average seek time is 15 milliseconds
  - Mean time between failures (MTBF) is estimated at 150,000 hours
  - DC-3 controller controls up to four enhanced serial drive interface (ESDI) disk drives
  - Each disk controller provides error correction code (ECC) and media defect management

- DD-3 low performance disk unit
  - 1.5 Gbytes; 5.25 inch ESDI drive
  - 2.75-Mbyte/s transfer rate
  - Packaged ten drives per drawer with a shared power supply
- DS-4 disk storage system
  - DD-4 medium performance disk unit
  - 3.0 Gbyte 8-inch two-head IPI dual port disk drives
  - 9.34 Mbytes/s transfer rate
  - Packaged two disk drives per drawer with one per drive
  - DC-4 controller controls intelligent peripheral interface (IPI-2) containing two IPI channels; each channel can handle two DD-4 drives
- DAS-2 disk array subsystem
  - Intelligent disk array controller
  - Bank of eight 1.5-Gbyte ESDIs for storage plus one for parity and one spare
  - Hardware striping used to distribute data evenly across all drives
  - Sustained transfer rate of 13 Mbytes/s
  - Has 12-Gbyte unformatted capacity
- Tape units
- 9-track tape drive subsystem
  - One TCU-2 tape controller unit
  - One TD-2800 bpi (NRZI), 1600 bpi (PE), 6250 bpi (GCR) 9-track low profile tape drive, 125 ips
- Cartridge tape drive (EXABYTE) models
  - EX-1 2.3 Gbyte, EXB-200 246K/s 8-mm tape drive
  - EX-2 5.0 Gbyte, EXB-850 500K/s 8-mm tape drive

- TD-3 .5-inch cartridge tape drive; 3480-type tape
- DR-1 removable disk system
  - Dual 5.25-inch disks in easily removable cartridges designed to protect data from damage during transport
- DR-2 removable IOS disk drive
- TD-2 autoloading low profile 9-track tape drive
- TD-3 .5 inch cartridge; 3480-type tape drive

Customers select peripheral devices according to their requirements, some of which are dictated by the system requirements.

System Configurations

Three system configuration examples are offered to fill the specific needs of the customer. In many cases, the configuration defines the peripheral options selected. The three configurations are: remote seismic option, departmental system, and file server system. The characteristics of each of these configurations are listed below.

- Remote seismic system (refer to Figure 1-3).
  - Low overall system cost
  - 9-track tape capability
  - 2 to 4 VMEbus channels
  - Minimum disk space required (10 to 20 Gbytes)
  - Single 19-inch frame
- Departmental supercomputer system (refer to Figure 1-4).
  - Compatible with office environment
  - Compact size
  - Low power consumption
  - Minimal heat dissipation
  - Quiet operation
  - Easy to network
  - 4 to 8 VMEbus channels
  - Customer-specified disk capacity
  - 32 to 64 Gbytes of ESDI disk drive storage
  - Higher performance DAS or IPI drives
  - Tape backup system
  - 1 or 2 19-inch frames

- File server system (refer to Figure 1-5).
  - Computer room environment
  - Significant disk storage space (200 to 400 Gbytes)
  - 8 to 16 VMEbus channels
  - Up to 4 racks per system

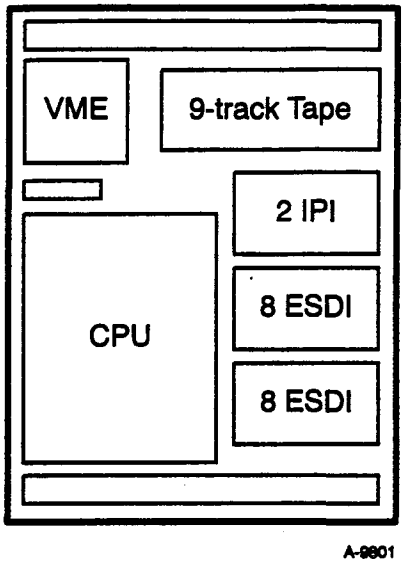


Figure 1-3. Remote Seismic System

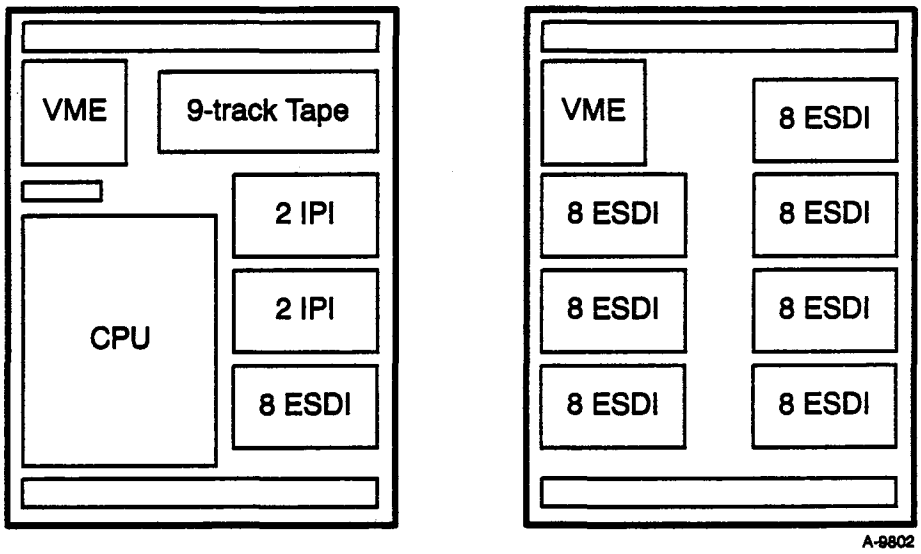


Figure 1-4. Departmental System

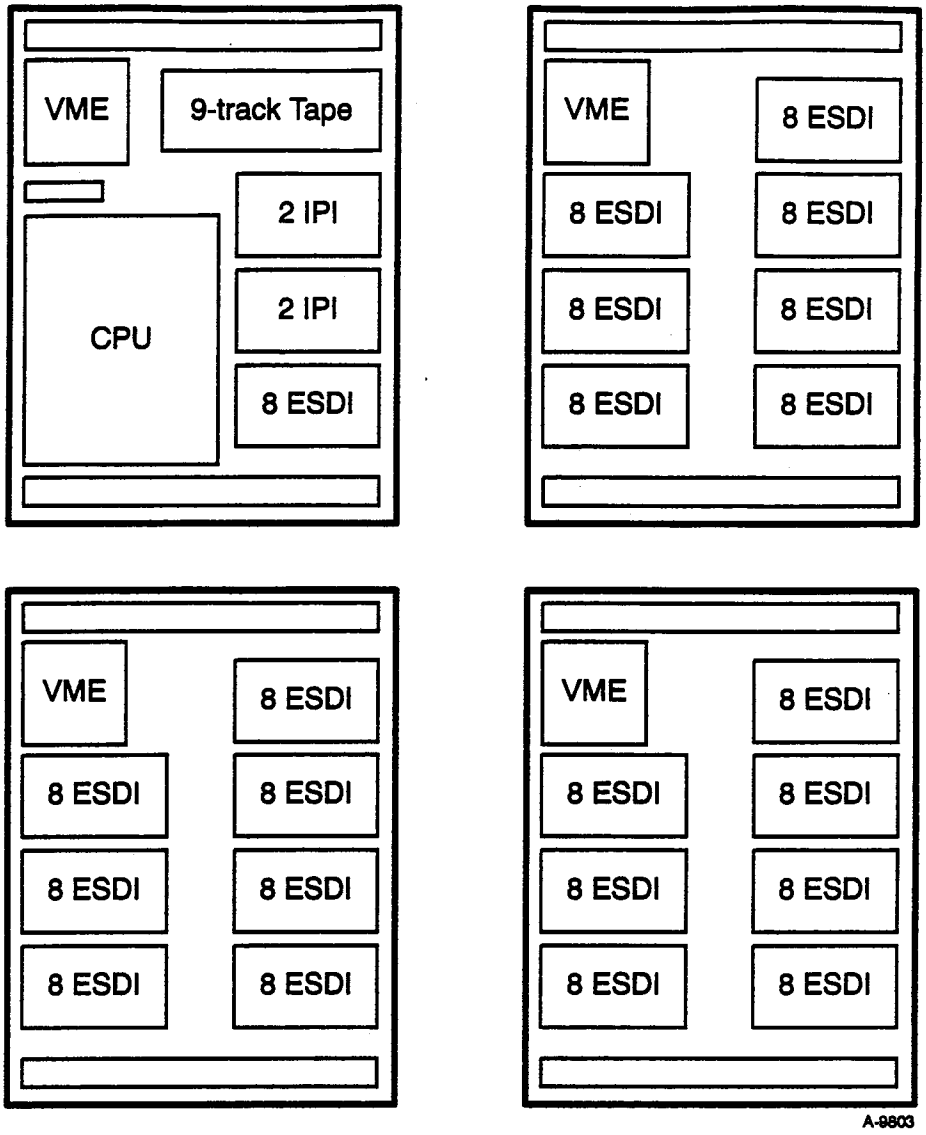


Figure 1-5. File Server System

CPU Overview

The CPU in the CRAY Y-MP EL computer system is constructed on a single 16 x 22 inch printed circuit (PC) board. This PC board contains all of the logic associated with the CRAY Y-MP EL system CPU. Very large scale integration (VLSI) solid-state technology enables a relatively small PC board to contain an entire CPU.

The VLSI chips used in the CPU are application-specific integrated circuits (ASICs). They are constructed using complementary metal oxide semiconductors (CMOS). The ASICs are available in two package sizes; one has 299 pins and one has 223 pins. The internal construction

of the ASIC forms 100,000 undefined gates. This massive number of gates is contained in a 2 x 2 inch or a 1.88 x 1.88 inch package, which consumes an average of 5 watts at + 5 volts.

The CPU contains nine ASIC chip types; there are 23 ASIC chips per CPU. Figure 1-6 shows the chip layout for the CPU of the CRAY Y-MP EL system.

The nine types of ASICs that form the CPU are:

- 1 arbiter (AR) ASIC, used for memory access control and for inter-CPU synchronization.
- 1 address and scalar (AS) ASIC, containing the address registers and address functional units and the scalar registers and scalar functional units.
- 2 channel control (CC) ASICs. Each CC supports two Y1 channels, which are the 40 Mbyte/s channels that connect to the VMEbus sybsystem. The CC option is also used for control support.
- 8 data switch (DS) ASICs, which perform the major data steering between memory, channel, and functional units; they contain the vector registers and B/T registers.
- 4 execution unit (EU) ASICs. Each EU contains all of the vector and floating-point functional units, except the floating-point reciprocal. All of these functional units are fully pipelined, but only one functional unit can be active per EU at any one time. Vector mask (VM) operations can only be performed on EU3.
- 1 memory control (MC) ASIC, which provides memory address generation for a maximum of 512 Mwords of memory. The MC also performs operand and program range error detection.
- 4 memory data (MD) ASICs, which perform single-error correction/double-error detection (SECDED †) and generate check bits. The MDs also support any memory maintenance instructions.
- 1 processor control (PC) ASIC. This device contains eight 32-word instruction buffers(IB), performs shared register (SR) access control, and supports 7 shared register clusters. Also included are the instruction issue control circuits, including the instruction and functional unit scoreboard, and I/O interrupt handling circuits.

† Hamming, R. W. "Error Detection and Correcting Codes." Bell System Technical Journal. 29.2 (1950): 147–160.

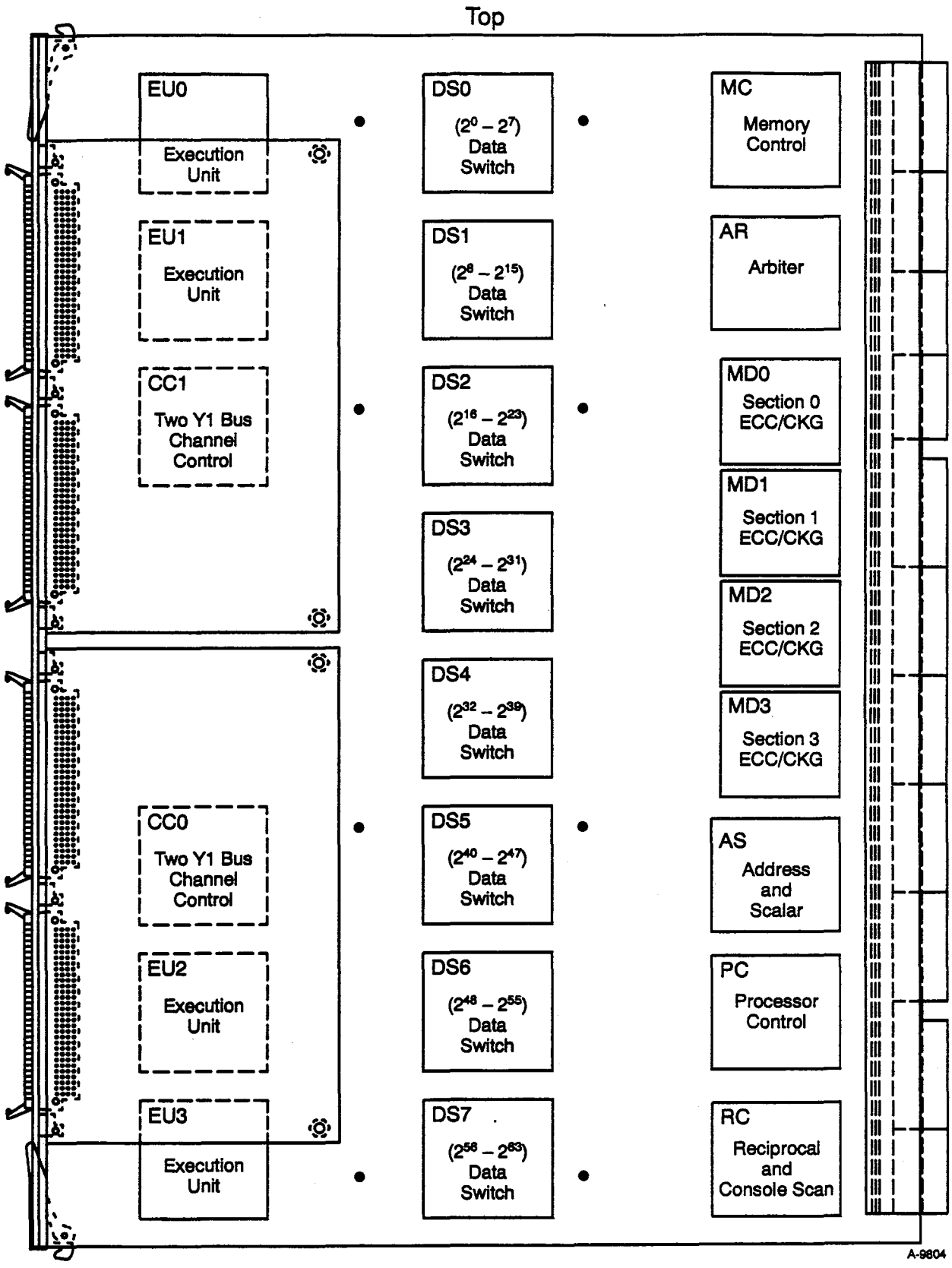


Figure 1-6. CPU/Channel Module

- 1 reciprocal and console (RC) ASIC. The RC performs floating-point reciprocal approximations and contains CONBUS interfaces, scan control and clock control hardware, and console registers.

The interconnection of these CPU components is shown by the block diagram in Figure 1-7.

The CRAY Y-MP EL system is designed to contain up to four independent CPUs. However, the four CPUs can work in conjunction by using shared registers. The CPU runs on a 30 nanosecond (ns) clock. Each CPU connects to the system backplane using 1230 signal pins.

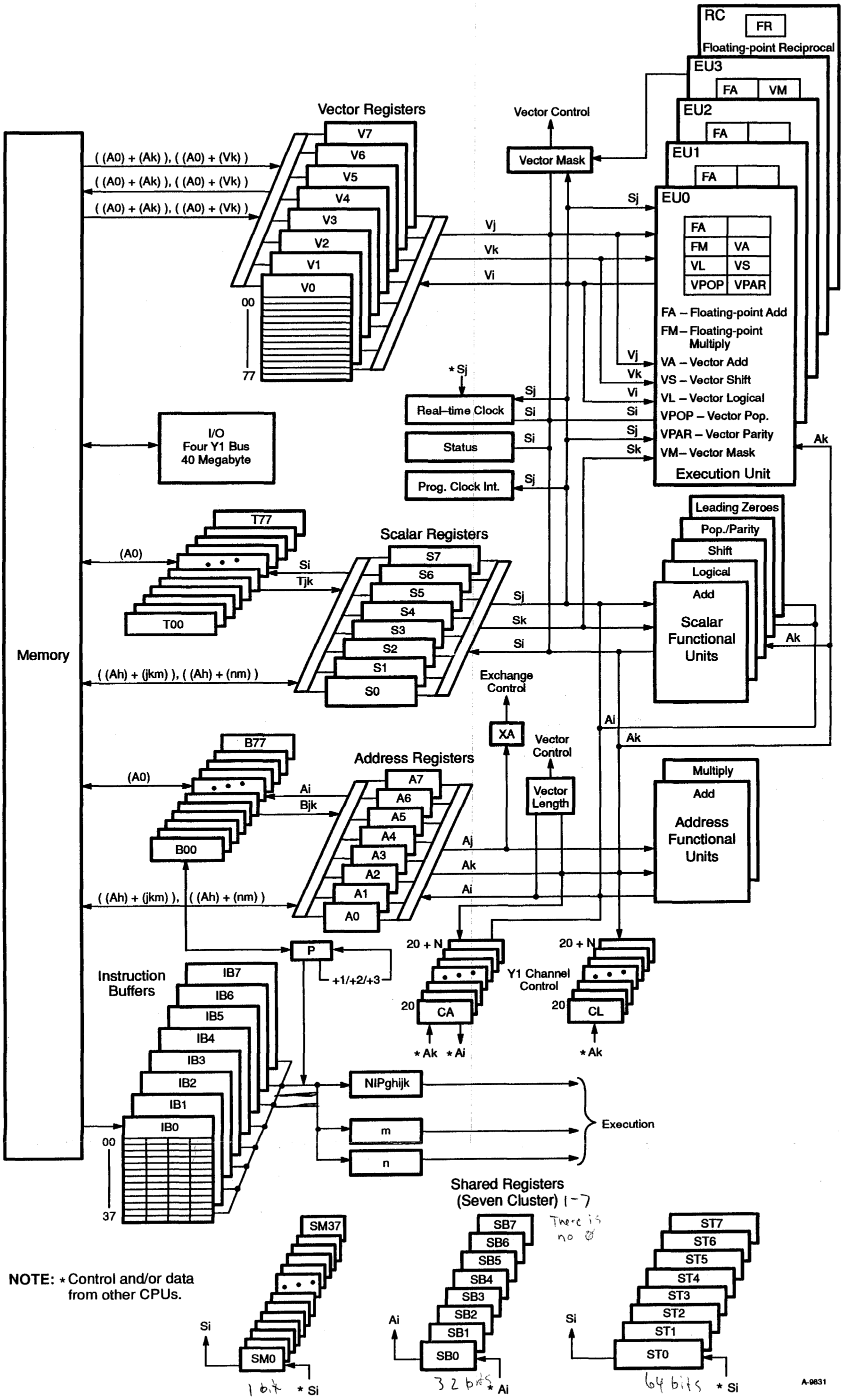


Figure 1-7. CRAY Y-MP EL Block Diagram

Memory Overview

Central memory within the CRAY Y-MP EL system is contained on four PC boards. Each of these boards is 16 x 22 inches (the same size as the CPU board) and is composed of two ASIC types, with a total of nine ASICs per board and a specific number of 1M x 4 dynamic random access memory (DRAM) integrated circuits. The number of DRAMs depends upon the customer's choice of available memory options. The CRAY Y-MP EL system can be ordered with either 64 Mwords or 128 Mwords of central memory. When the 64 Mword option is selected, the memory module contains 16 Mwords of memory supplied by 288 DRAMs. This is a half-populated module. The larger memory, 128 Mwords, consists of 576 DRAMs mounted on each module, creating 32 Mwords per module and 128 Mwords per system. This option uses fully populated modules.

32 mw = half populated 1mx4
25 address bits used

64 mw = half populated 1mx4
26 address bits used

128 mw = fully populated 1mx4
27 address bits used

The layout of the memory board is shown in Figure 1-8. The two ASIC types used on the memory board are:

- 2 memory array control (MAC) ASICs support the four memory functions. These ASICs contain an address crossbar which allows access to memory from each of the four CPUs as well as refresh address counters for the local refresh function. The MAC ASIC also contain DRAM address and control circuitry, which provides control to all MADs.

- 7 memory array data (MAD) ASICs. These contain a data crossbar that connects the four CPUs to the 16 banks contained on each memory board. The MAD ASICs handle a portion of the 72-bit memory data word.

The CRAY Y-MP EL system central memory contains a total of 64 banks spread across the four modules. Each module contains 16 banks and is considered a memory section. These 16 banks are separated into lower and upper banks on each board. Thus, a half-populated memory board uses only the lower bank, but still retains the full 16 banks of memory. This means that a fully populated memory module uses both upper and lower banks.

Figure 1-9 shows the chassis locations of the eight boards that form the CRAY Y-MP EL computer system. Note that this represents a top view, with the front of the chassis at the bottom of the diagram.

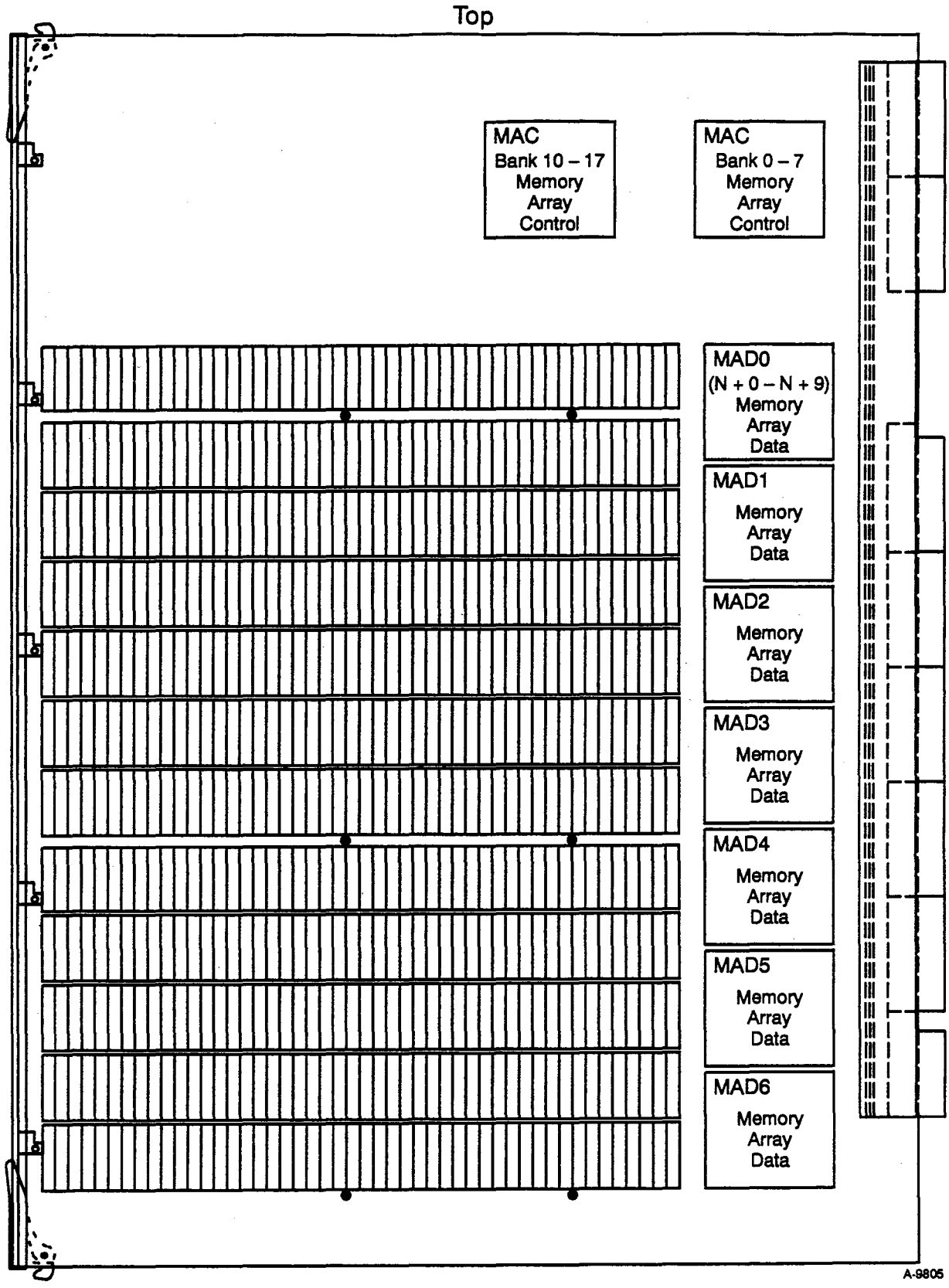


Figure 1-8. Memory Module

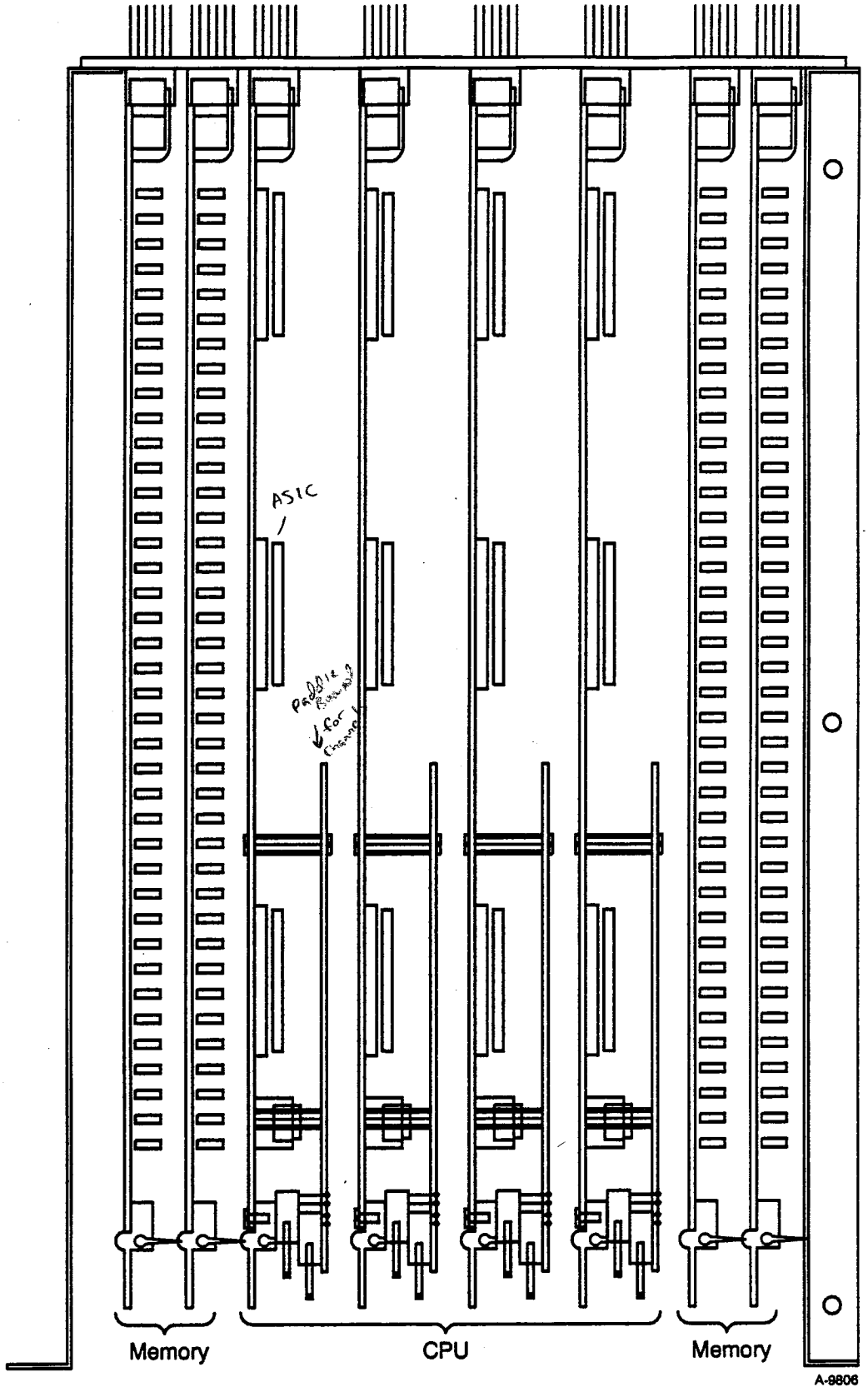
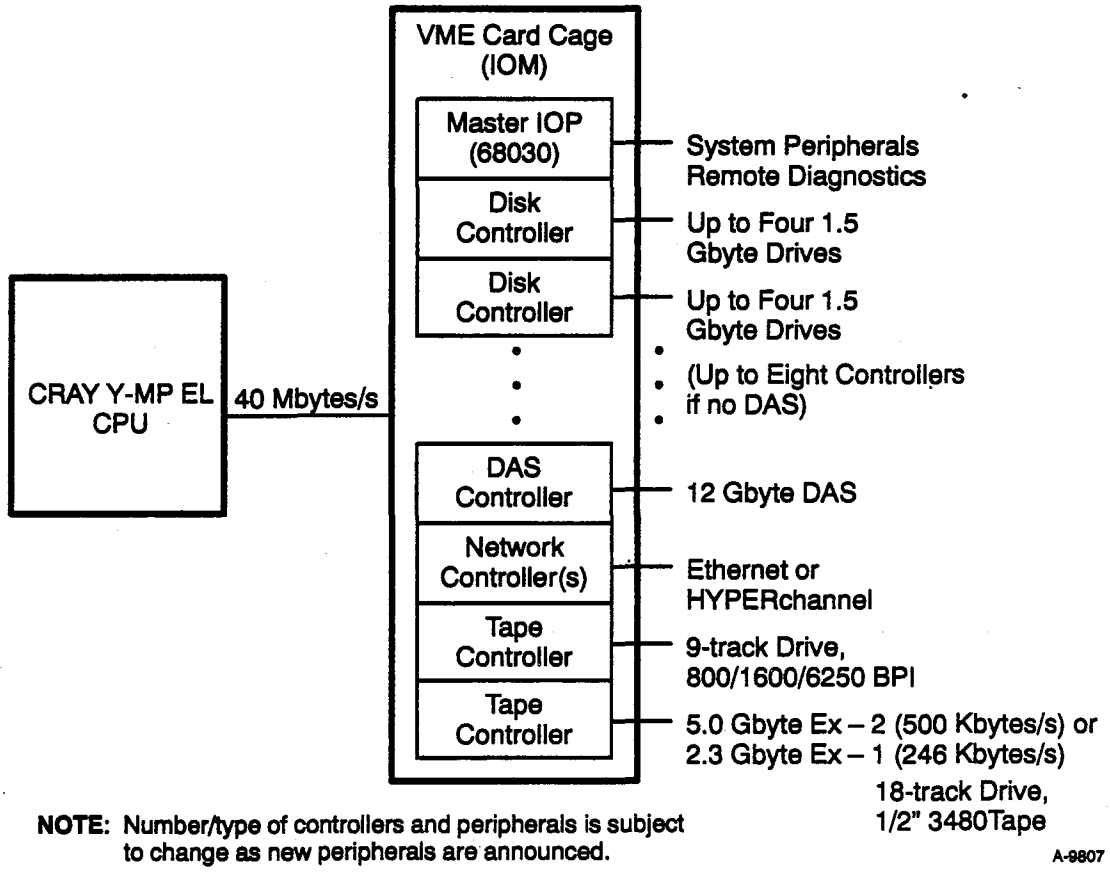


Figure 1-9. Chassis, Top View

Input/Output Subsystem Overview

The IOS for the CRAY Y-MP EL system was selected to provide customers with the maximum choice of peripheral equipment. The IOS is a VME-based system that communicates with the CPU via the Y1 bus (a 40-Mbyte/s channel) using a Cray Research, Inc. (CRI) proprietary module, the input/output buffer board (IOBB). The IOBB is the only CRI proprietary board within the IOS. All other functions within the IOS are performed using vendor-supplied VME boards.

The restrictions of the IOS configuration require use of a 68030-type processor that is capable of processing 32 bits, and as many as eight peripheral controllers to handle data transfers. A possible IOS configuration is represented in Figure 1-10. The types of controllers used are defined by the customer's choice of system peripherals.



NOTE: Number/type of controllers and peripherals is subject to change as new peripherals are announced.

Figure 1-10. IOS Configuration

There are several other considerations pertaining to the VMEbus system. The standard VME mechanical chassis is a 19-inch rack mount chassis that is air cooled, supports standard 64 x 160 mm VME boards, and requires a 750-watt power supply. The backplane is a modular design based on a 10-slot system. The backplane configurations include:

- 10-slot, 10-slot option
- 10-slot, 6-slot, 4-slot option
- 6-slot, 4-slot, 6-slot, 4-slot option

2 MEMORY

The memory portion of the CRAY Y-MP EL computer system consists of four modules. These modules currently are provided in two types: a fully populated module and a half populated module. Both module types are constructed on 16 x 22 x .093 inch printed circuit (PC) boards consisting of 16 circuit layers. These 16 layers are:

- 1 top pad
- 1 bottom pad
- 4 ground (Vss) layers
- 4 power layers: 5-volt application-specific integrated circuit (ASIC); 5-volt dynamic random access memory (DRAM)
- 6 signal layers

The logic portion of a memory PC board comprises two types of ASICs and a group of DRAM chips. There are nine ASICs on each memory board: two memory array control (MAC) ASICs and seven memory array data (MAD) ASICs. The number of DRAM chips on each memory PC board is determined by the memory size selected by the customer. Currently, the CRAY Y-MP EL system can be supplied with either a 64-Mword memory or a 128-Mword memory.

The 64-Mword memory option is a half-populated memory board that contains 288 DRAM chips. Each of these DRAM chips is a 1M x 4 memory chip with a 70 nanosecond (ns) access time. The same DRAM chip is used with the 128-Mword memory, but there are 576 memory chips mounted on the fully populated memory board. Refer to Figure 2-1 for a diagram of a fully populated memory module.

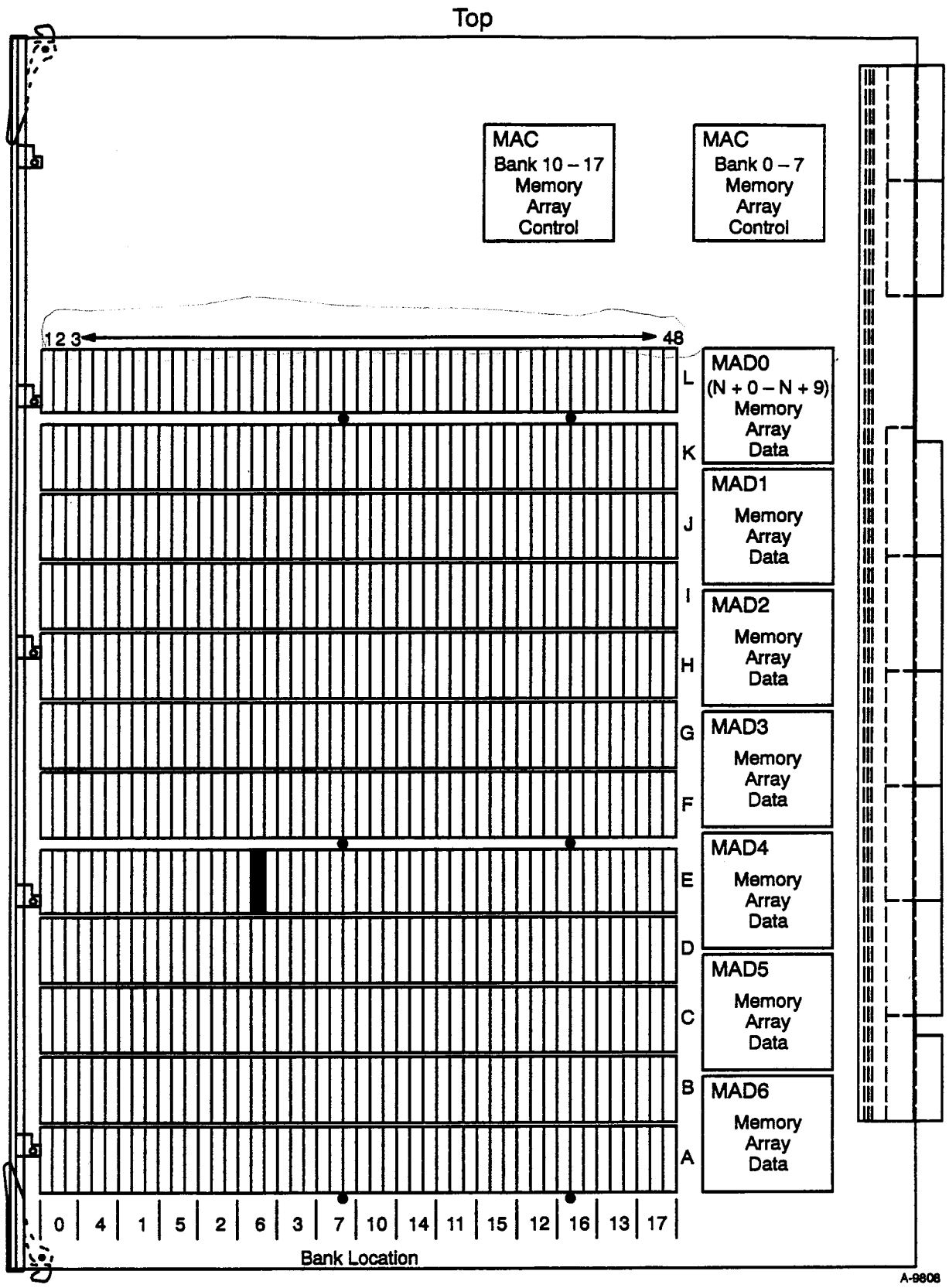


Figure 2-1. Memory Module

Memory is divided into 16-banks for addressing whether the module is half populated or fully populated. The fully populated module contains an upper 16 banks and a lower 16 banks of addressable memory; the half populated module uses only the lower 16 banks. There are 16 banks of memory on each memory module, providing the mainframe with a total of 64 banks of memory. The memory is also separated into sections; each memory module is one section. This arrangement provides the mainframe with four memory sections, each consisting of 16 banks.

The addressing method used by the CRAY Y-MP EL system is a 32-bit address scheme, of which 27 bits (0 through 26) are used. The bits are assigned as shown in Figure 2-2. Bits 2^0 and 2^1 are used to select the appropriate section and bits 2^2 through 2^5 are used for bank selection.

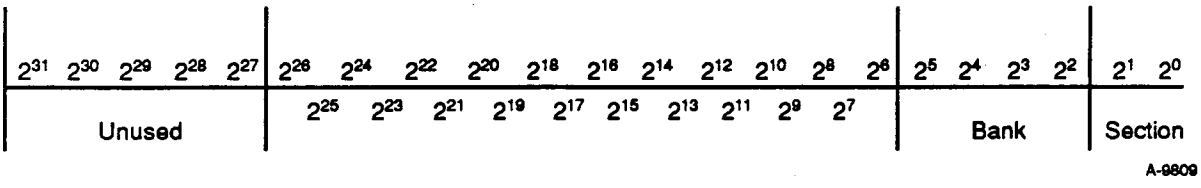


Figure 2-2. Address Bit Assignments

The remainder of the address bits, with one exception, are used to select the actual memory location. The exception is bit 2^{26}, which is used as the upper/lower bank select bit. This bit is only necessary on a fully populated 128-Mword memory module. The internal addressing scheme, represented by bits 2^6 through 2^{25}, uses a row and column scheme. The odd numbered bits are used to count the rows of memory, while the even numbered bits are used to locate the column. It is possible to determine the exact DRAM chip that contains a failing bit by first decoding bits 2^0 and 2^1 to determine which section is in error. This corresponds to a specific memory module. The locations of the memory sections is shown in Figure 2-3, which is a representation of the eight-slot mainframe card cage. All of the memory modules are interchangeable, so swapping the suspected failing module with a good module can help isolate the failure.

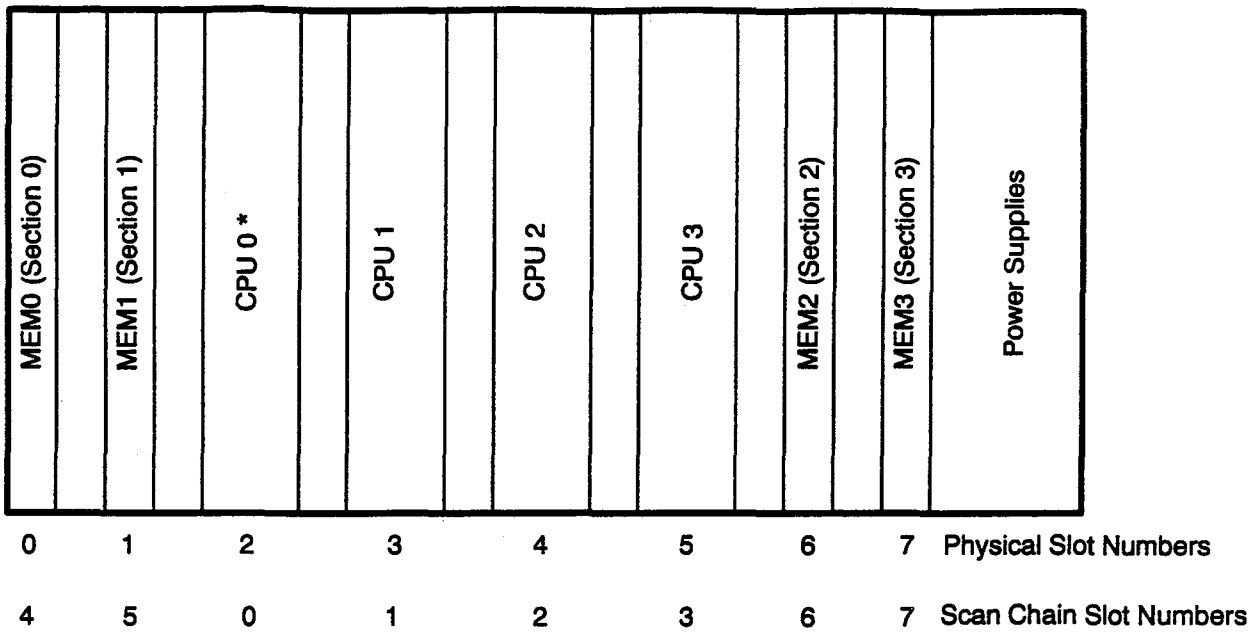


Figure 2-3. Mainframe Chassis, Front View

The next step in the failure isolation procedure is to decode bits 2^2 through 2^5 from the failing address. Once the failing bank has been identified, use Figure 2-1 to identify the location of the failing bank on the PC board. Next, locate the failing data bit in the center section of Table 2-1 as well as the column representing the failing bank. The bank selection will partially depend on the condition of address bit 2^{26}. If 2^{26} is zero, the failing bank is located to the right of center. When 2^{26} is one, the failing bank is located to the left of center. Once the two starting coordinates have been located, draw lines to the left or right from the failing data bit and down under the failing bank. The point at which these two lines intersect represents the board coordinates of the failing chip.

In Figure 2-1, representing the actual PC board, it is possible to locate the chip that requires replacement. In this figure, note that at the right side of the chip locations are labeled positions A through L, moving bottom to top. Along the top, the chips are numbered left to right from 1 to 48. Finally, along the bottom of the chips, the bank locations are indicated. Each bank contains three chips in each of the rows A through L, and the banks are **not** numbered consecutively. Using these locator coordinates, apply the board coordinate determined on Table 2-1 , and the physical location of the failing chip is defined.

As an example, if data bit 2^{43} is reported as failing with an address of 312764132_8, locate the failing DRAM chip.

continued pg 2-6

Table 2-1. DRAM Locator Chart

Banks H A26 = 1																DRAM				Banks L A26 = 0																	
17	16	15	14	13	12	11	10	7	6	5	4	3	2	1	0	1M x 4				0	1	2	3	4	5	6	7	10	11	12	13	14	15	16	17
L 46	L 40	L 34	L 28	L 43	L 37	L 31	L 25	L 22	L 16	L 10	L 4	L 19	L 13	L 7	L 1	0	9	23	28	K 1	K 7	K 13	K 19	K 4	K 10	K 16	K 22	K 25	K 31	K 37	K 43	K 28	K 34	K 40	K 46
L 47	L 41	L 35	L 29	L 44	L 38	L 32	L 26	L 23	L 17	L 11	L 5	L 20	L 14	L 8	L 2	1	10	21	30	K 2	K 8	K 14	K 20	K 5	K 11	K 17	K 23	K 26	K 32	K 38	K 44	K 29	K 35	K 41	K 47
L 48	L 42	L 36	L 30	L 45	L 39	L 33	L 27	L 24	L 18	L 12	L 6	L 21	L 15	L 9	L 3	2	11	22	31	K 3	K 9	K 15	K 21	K 6	K 12	K 18	K 24	K 27	K 33	K 39	K 45	K 30	K 36	K 42	K 48
J 46	J 40	J 34	J 28	J 43	J 37	J 31	J 25	J 22	J 16	J 7	J 4	J 19	J 13	J 7	J 1	3	8	20	29	I 1	I 7	I 13	I 19	I 4	I 10	I 16	I 22	I 25	I 31	I 37	I 43	I 28	I 34	I 40	I 46
J 47	J 41	J 35	J 29	J 44	J 38	J 32	J 26	J 23	J 17	J 8	J 5	J 20	J 14	J 8	J 2	4	13	19	24	I 2	I 8	I 14	I 20	I 5	I 11	I 17	I 23	I 26	I 32	I 38	I 44	I 29	I 35	I 41	I 47
J 48	J 42	J 36	J 30	J 45	J 39	J 33	J 27	J 24	J 18	J 9	J 6	J 21	J 15	J 9	J 3	5	14	17	26	I 3	I 9	I 15	I 21	I 6	I 12	I 18	I 24	I 27	I 33	I 39	I 45	I 30	I 36	I 42	I 48
H 46	H 40	H 34	H 28	H 43	H 37	H 31	H 25	H 22	H 16	H 10	H 4	H 19	H 13	H 7	H 1	6	15	18	27	G 1	G 7	G 13	G 19	G 4	G 10	G 16	G 22	G 25	G 31	G 37	G 43	G 28	G 34	G 40	G 46
H 47	H 41	H 35	H 29	H 44	H 38	H 32	H 26	H 23	H 17	H 11	H 5	H 20	H 14	H 8	H 2	7	12	16	25	G 2	G 8	G 14	G 20	G 5	G 11	G 17	G 23	G 26	G 32	G 38	G 44	G 29	G 35	G 41	G 47
H 48	H 42	H 36	H 30	H 45	H 39	H 33	H 27	H 24	H 18	H 12	H 6	H 21	H 15	H 9	H 3	32	41	55	60	G 3	G 9	G 15	G 21	G 6	G 12	G 18	G 24	G 27	G 33	G 39	G 45	G 30	G 36	G 42	G 48
F 46	F 40	F 34	F 28	F 43	F 37	F 31	F 25	F 22	F 16	F 10	F 4	F 19	F 13	F 7	F 1	33	42	53	62	E 1	E 7	E 13	E 19	E 4	E 10	E 16	E 22	E 25	E 31	E 37	E 43	E 28	E 34	E 40	E 46
F 47	F 41	F 35	F 29	F 44	F 38	F 32	F 26	F 23	F 17	F 11	F 5	F 20	F 14	F 8	F 2	34	43	54	63	E 2	E 8	E 14	E 20	E 5	E 11	E 17	E 23	E 26	E 32	E 38	E 44	E 29	E 35	E 41	E 47
F 48	F 42	F 36	F 30	F 45	F 39	F 33	F 27	F 24	F 18	F 12	F 6	F 21	F 15	F 9	F 3	35	40	52	61	E 3	E 9	E 15	E 21	E 6	E 12	E 18	E 24	E 27	E 33	E 39	E 45	E 30	E 36	E 42	E 48
D 46	D 40	D 34	D 28	D 43	D 37	D 31	D 25	D 22	D 16	D 10	D 4	D 19	D 13	D 7	D 1	36	45	51	56	C 1	C 7	C 13	C 19	C 4	C 10	C 16	C 22	C 25	C 31	C 37	C 43	C 28	C 34	C 40	C 46
D 47	D 41	D 35	D 29	D 44	D 38	D 32	D 26	D 23	D 17	D 11	D 5	D 20	D 14	D 8	D 2	37	46	49	58	C 2	C 8	C 14	C 20	C 5	C 11	C 17	C 23	C 26	C 32	C 38	C 44	C 29	C 35	C 41	C 47
D 48	D 42	D 36	D 30	D 45	D 39	D 33	D 27	D 24	D 18	D 12	D 6	D 21	D 15	D 9	D 3	38	47	50	59	C 3	C 9	C 15	C 21	C 6	C 12	C 18	C 24	C 27	C 33	C 39	C 45	C 30	C 36	C 42	C 48
B 46	B 40	B 34	B 28	B 43	B 37	B 31	B 25	B 22	B 16	B 10	B 4	B 19	B 13	B 7	B 1	39	44	48	57	A 1	A 7	A 13	A 19	A 4	A 10	A 16	A 22	A 25	A 31	A 37	A 43	A 28	A 34	A 40	A 46
B 47	B 41	B 35	B 29	B 44	B 38	B 32	B 26	B 23	B 17	B 11	B 5	B 20	B 14	B 8	B 2	64	66	68	70	A 2	A 8	A 14	A 20	A 5	A 11	A 17	A 23	A 26	A 32	A 38	A 44	A 29	A 35	A 41	A 47
B 48	B 42	B 36	B 30	B 45	B 39	B 33	B 27	B 24	B 18	B 12	B 6	B 21	B 15	B 9	B 3	65	67	69	71	A 3	A 9	A 15	A 21	A 6	A 12	A 18	A 24	A 27	A 33	A 39	A 45	A 30	A 36	A 42	A 48

1. Decode the address. Refer to Figure 2-4.

   The reported error is in section 2, lower bank 6.

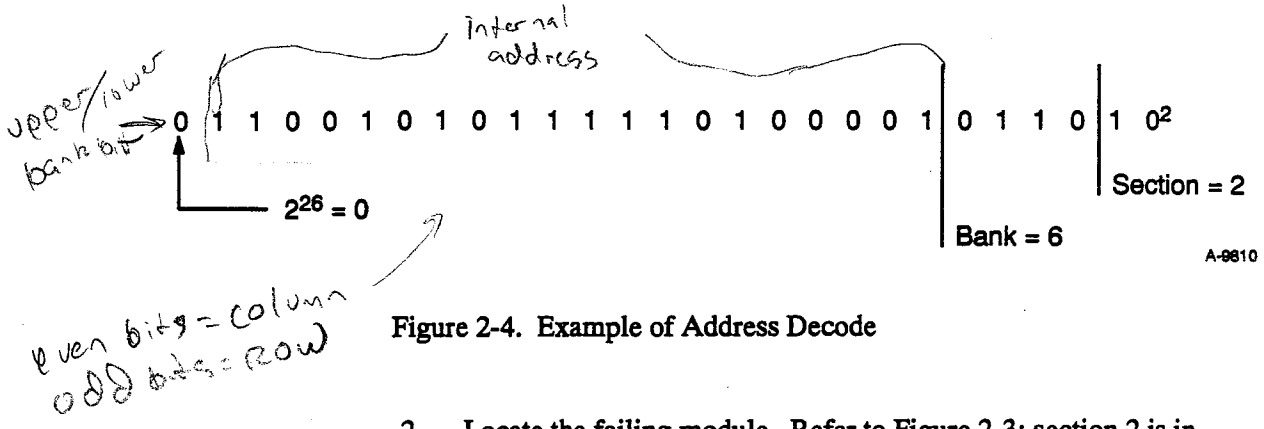


Figure 2-4. Example of Address Decode

2. Locate the failing module. Refer to Figure 2-3; section 2 is in slot 6.

3. Using Table 2-1, locate data bit 2^{43} in the center section and lower bank 6 in the top row. Draw lines to the right from bit 2^{43} and down from bank 6. These two lines should intersect at a point labeled E17.

4. Using Figure 2-1, locate point E17. To the right of the DRAM chips, locate row E. From the left end of row E, count to the seventeenth DRAM chip (shaded on the diagram). To check, follow column 17 down and verify that the failing chip is in bank 6.

Because the DRAM chip used in the CRAY Y-MP EL system memory is a 1M x 4 chip, it is expected that a single-bit error will rapidly escalate into a multiple-bit error. For this reason, it is important to repair single-bit errors as soon as possible.

Other specifications pertaining to the DRAM chip are:

- ZIP package .4 x 1.03 inch
- 5 clock period bank access time
- Standby power = 5 milliwatts at 5 volts
- Active power = 550 milliwatts at 5 volts

The rest of the memory PC board is made up of two types of MAC and MAD ASICs. The MAC ASIC chip supports four memory operations:

- Read, a normal DRAM read operation, lasts 5 clock periods.
- Write, a normal DRAM write operation, lasts 5 clock periods.

- Refresh, which uses row address strobe (RAS) to refresh data, lasts 5 clock periods.

- Read/Modify/Write (RMW) is used during an exchange, and uses normal DRAM read operations followed by a normal DRAM write operation to the same address. RMW lasts 10 clock periods.

The refresh control for the DRAM chips is located on the CPU and memory PC boards.

Other functions of the MAC ASIC include:

- Connects any of the four processors to any of the 16 banks through an address crossbar

- Controls both address and control signals

- Holds a refresh address counter

- Controls all of the MAD ASICs

The MAD ASICs perform the following functions:

- Each handles a specific portion of the 72-bit memory data word

- Each contains a data crossbar which connects any of the four processors to any of the 16 banks

The interconnection between the MAD, MAC, and DRAM chips is represented in Figure 2-5. Note that MAC 0 controls operations for banks 0 through 7, and MAC 1 controls banks 10 through 17. The content of the entire data bus is presented to the seven MAD ASICs and each receives its assigned bits. The bits assigned to each MAD ASIC are in no specific order. If a group of 10 or 11 bits of data are reported as failing, it could in fact be a single MAD that is failing. Included in this section are Figure 2-6, the internal block diagram of a MAD ASIC, and Figure 2-7, the internal block diagram of the MAC ASIC, indicate how an ASIC is designed. When an ASIC fails, the failing chip is removed from the PC board, discarded, and replaced by a new ASIC. No internal repairs are possible.

Other important characteristics of CRAY Y-MP EL system memory modules are:

- Each memory module contains 16 Mwords of memory if half populated and 32 Mwords if fully populated

- Operates with a 30 ns clock period

- Air cooled
- Connected to the backplane by two types of connectors:

  120 pins x 7 connectors = 840 pins
  90 pins x 1 connector = 90 pins
  930 pins

- Connected to the individual CPUs using one port per CPU; all ports are read/write ports
- There is no hardware error logger. Single and double bit errors are reported to the operating system by bits being set within the exchange package.
- 270 watts per module = 1080 watts per system
- Voltage = + 5 volts

2^2	2^1	2^0	Command	Action
0	0	x	No Command	None
0	1	0	Read-modify-write	DRAM Read Cycle Followed by Write Cycle to Same Location (Exchange)
0	1	1	Refresh	DRAM Read Cycle with Address from Refresh Address Counter
1	0	0	Write	DRAM Write Cycle
1	0	1	Range Error Write	None
1	1	0	Read	DRAM Read Cycle
1	1	1	Range Error Read	None

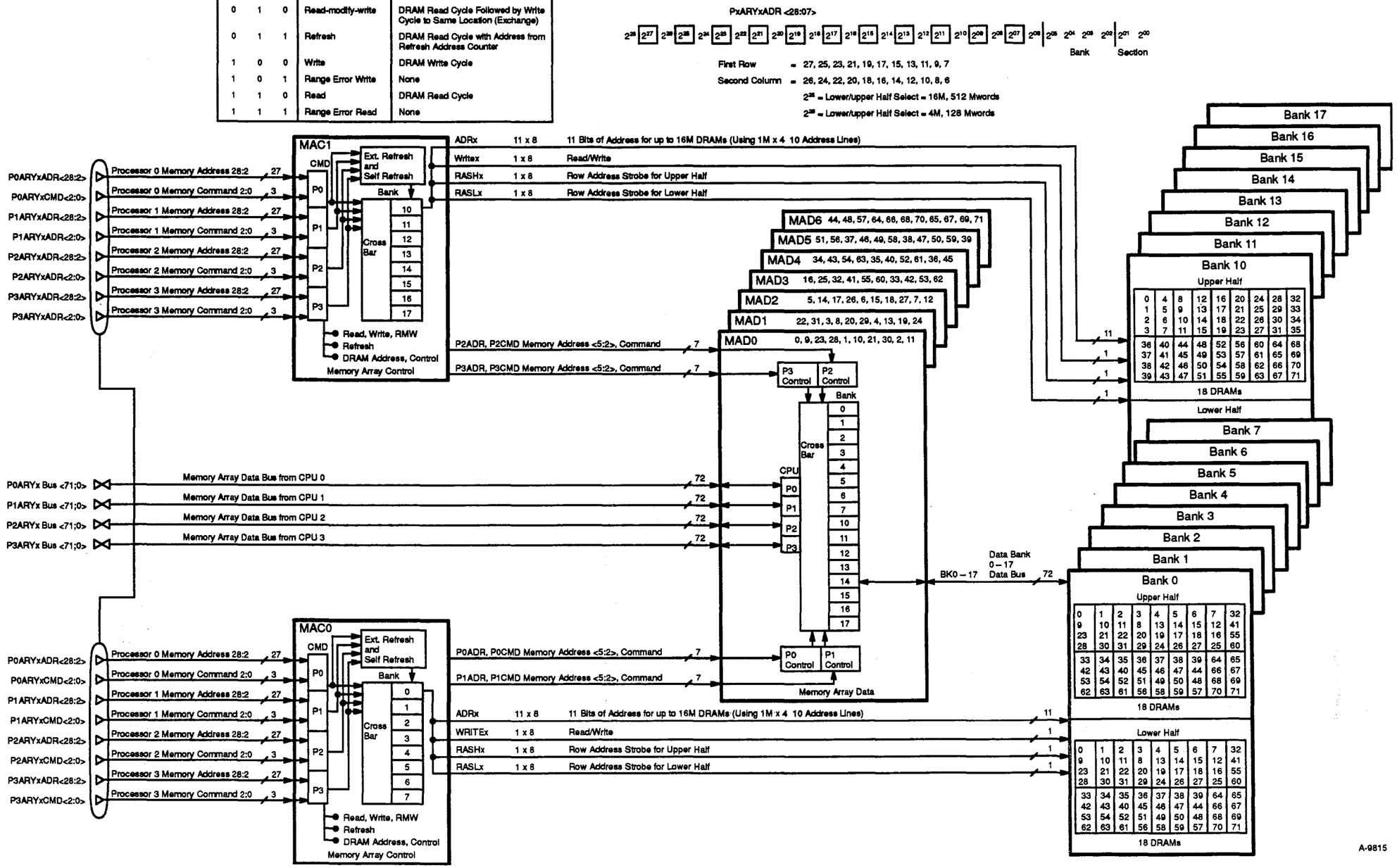


Figure 2-5. Memory Block Diagram

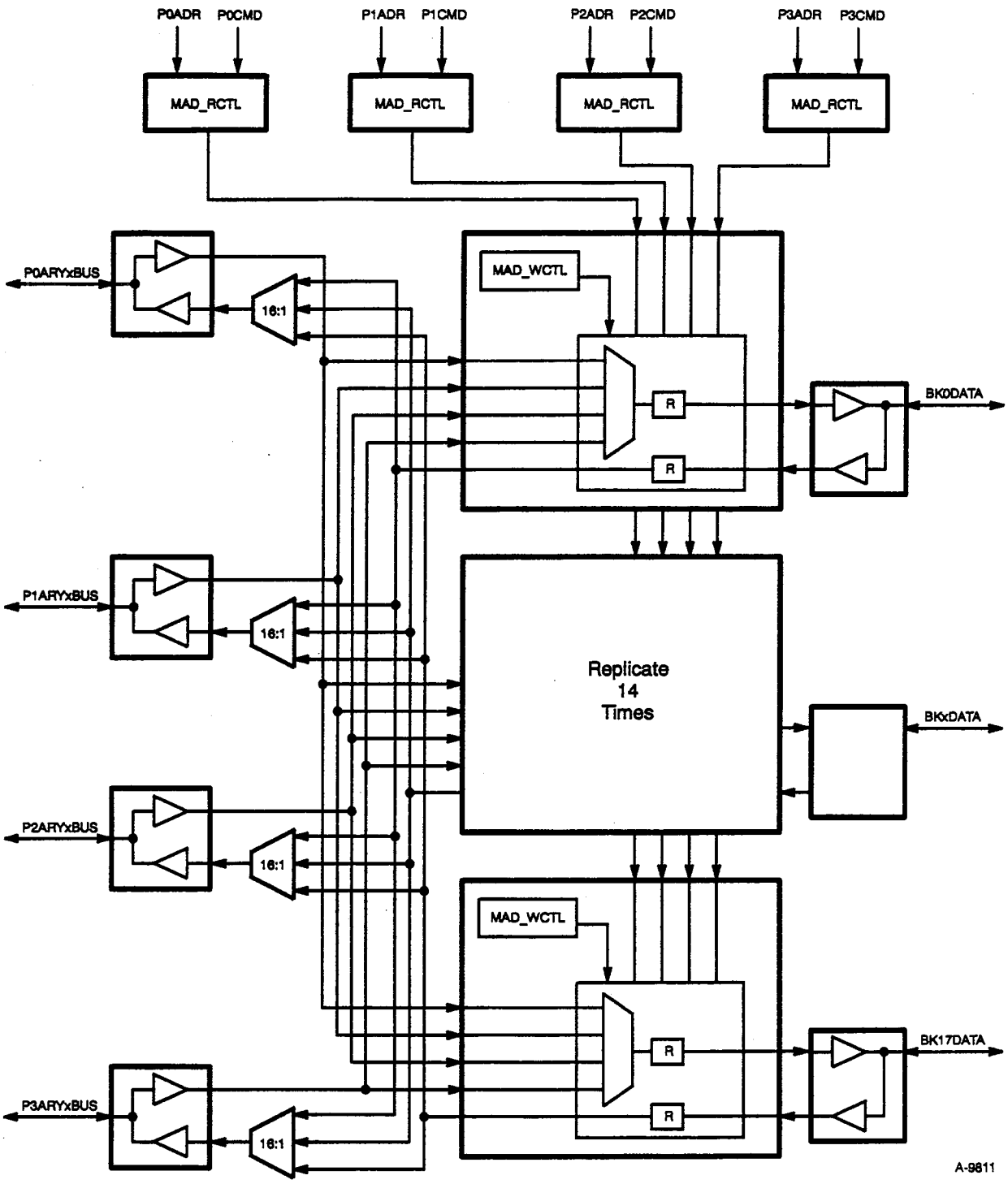


Figure 2-6. Memory Array Data Block Diagram

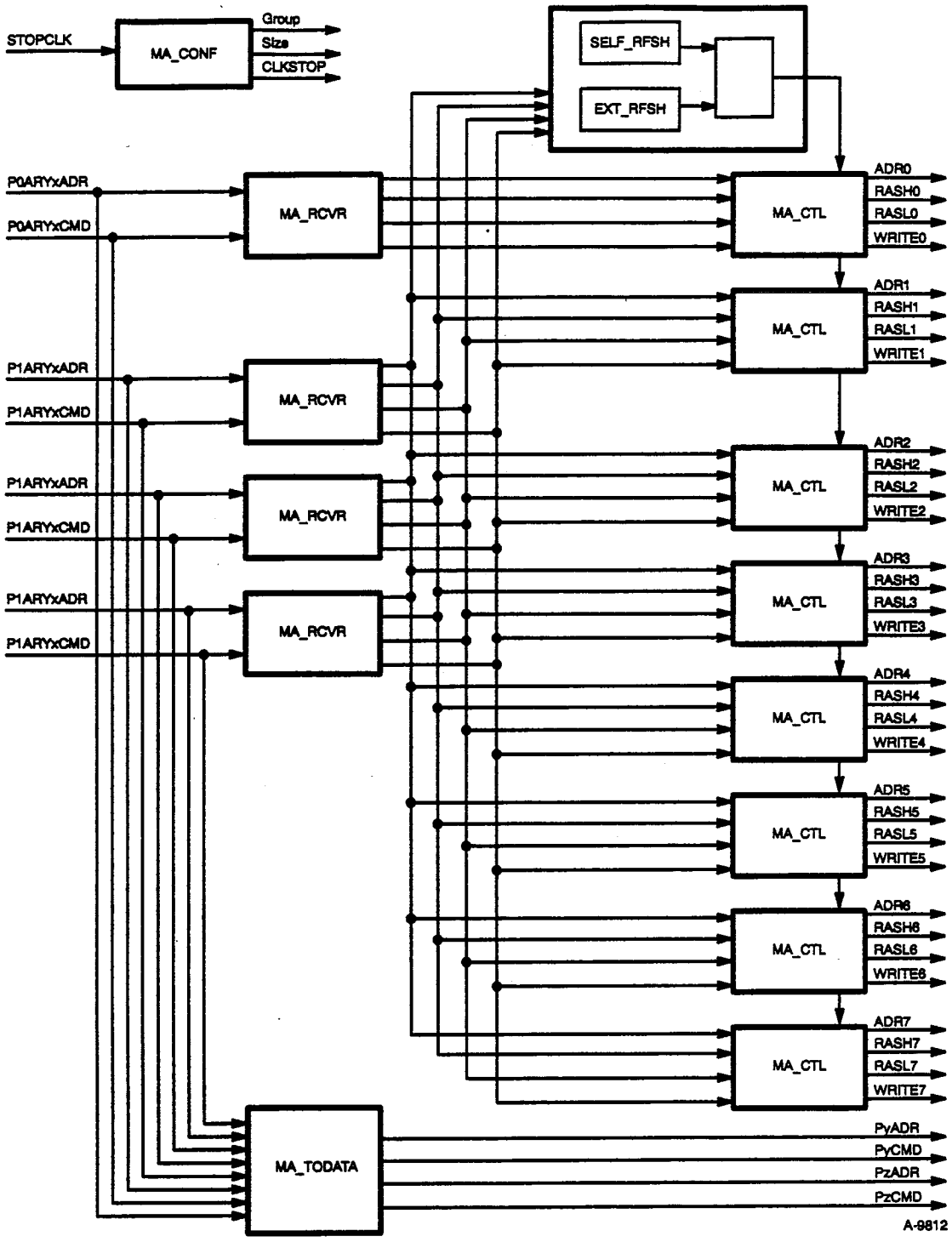


Figure 2-7. Memory Array Control Block Diagram

3 CPU

The CRAY Y-MP EL computer system central processing unit (CPU) is a single printed circuit (PC) board module that contains all the registers and functional units normally used in a computer mainframe. This miniturization is made possible by using very large-scale integration (VLSI) complimentary metal oxide semiconductive (CMOS) application-specific integrated circuits (ASICs). Twenty-three of these ASICs are mounted on a single 16 x 22 x .093 inch PC board.

The CMOS ASICs are provided in two sizes, both 1-micron, two-layer devices with 100,000 undefined gates:

- 2.08 x 2.08 inch chip with 299 pins
- 1.88 x 1.88 inch chip with 233 pins

The power consumed by these chips averages less than 5 watts per ASIC, operating at +5 volts.

The PC board used to form the CPU module is manufactured of 16 separate layers:

- 1 top pad
- 1 bottom pad
- 4 ground (Vss) layers
- 4 power (Vdd) layers
- 6 signal layers

This module uses an average of 160 watts of power per CPU module. Therefore, the maximum of four CPU modules in a system consumes 640 watts. The CPU, like the entire CRAY Y-MP EL system, is an air cooled device.

The following nine types of ASIC chips reside on the CPU module:

- 1 arbiter (AR) ASIC controls memory access and arbitrates all memory conflicts. It also provides inter-CPU synchronization.

- 1 address and scalar (AS) ASIC contains all of the address registers and the address functional units (FUs), as well as the scalar (S) registers and scalar FUs.

1 CC ASIC = 2 Y1 channels

- 2 channel control (CC) ASICs control or support the functions of two Y1 channels. Each of the Y1 channels is capable of 40-Mbyte/s transfers and each connects to a VME subsystem. ~~It is also possible to use two Y1 channels in conjunction to provide one high performance parallel interface (HIPPI) channel for 100-Mbyte/s transfers~~). One other function of the CC chip is console support.

- 8 data switch (DS) ASICs steer data between memory, the selected channel, and the required FUs.

- 4 execution unit (EU) ASICs contain all of the vector and floating-point FUs, with the exception of the floating-point reciprocal FU. This format provides full pipelining to the EU chips and allows each to work on a different problem independently of the others. The restrictions which pertain to the EU chips are:

  - Only one FU in each EU chip can be operating at any one time

  - Only the EU3 ASIC is capable of executing vector mask (VM) instructions (146 – 147, 175 instructions)

- 1 memory control (MC) ASIC provides address generation, which can support up to 512 Mwords of memory. The other function of the MC chip is to provide both operand and program range error detection.

- 4 memory data (MD) ASICs perform single-error correction, double-error detection (SECDED) functions, including check bit generation on the read and write memory data. The MD chips are also used to support all of the memory maintenance instructions.

- 1 processor control (PC) ASIC contains the CPU instruction buffers. The CRAY Y-MP EL system CPU uses eight instruction buffers, each of which is 32 words wide. Instruction issue control and I/O interrupt handling control also reside on the PC chip. Part of the issue control function is a resource scoreboard, also resident on the PC. Shared register access control is also performed on the PC ASIC; the CRAY Y-MP EL system supports seven shared register clusters.

- 1 reciprocal and control (RC) ASIC contains the floating-point reciprocal FU, the CONBUS interfaces, the scan and clock control, and the control registers.

These units interconnect in the way shown on the the CRAY Y-MP EL system block diagram, Figure 3-1. A more specific diagram showing some of the signal paths can be found in Figure 3-2, a block diagram of the CPU bus. This diagram also shows the internal contents of each of the ASICs located on the CPU module. Another diagram of interest is shown in Figure 3-3, which shows the actual location of the various ASICs.

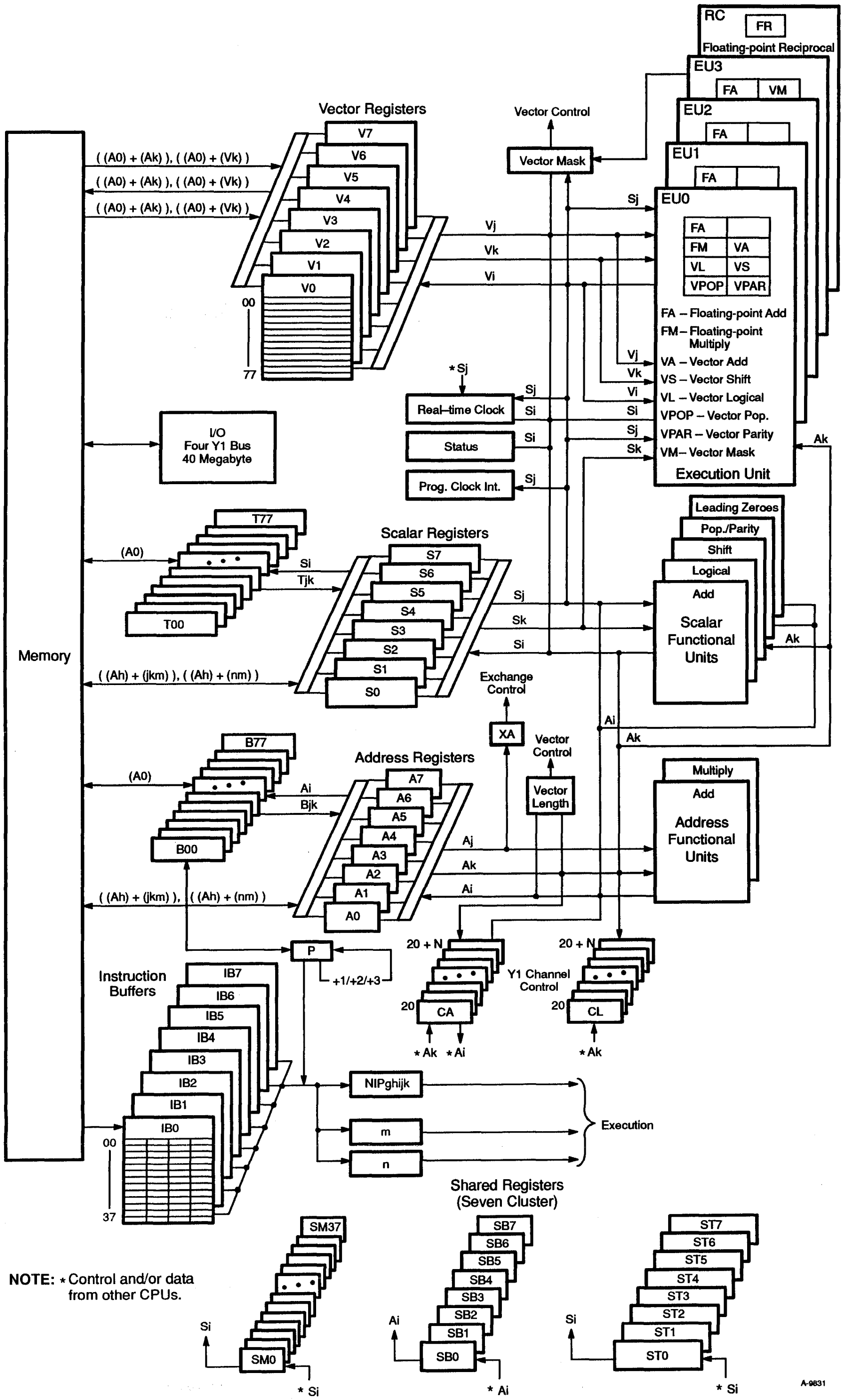


Figure 3-1. CRAY Y-MP EL Block Diagram

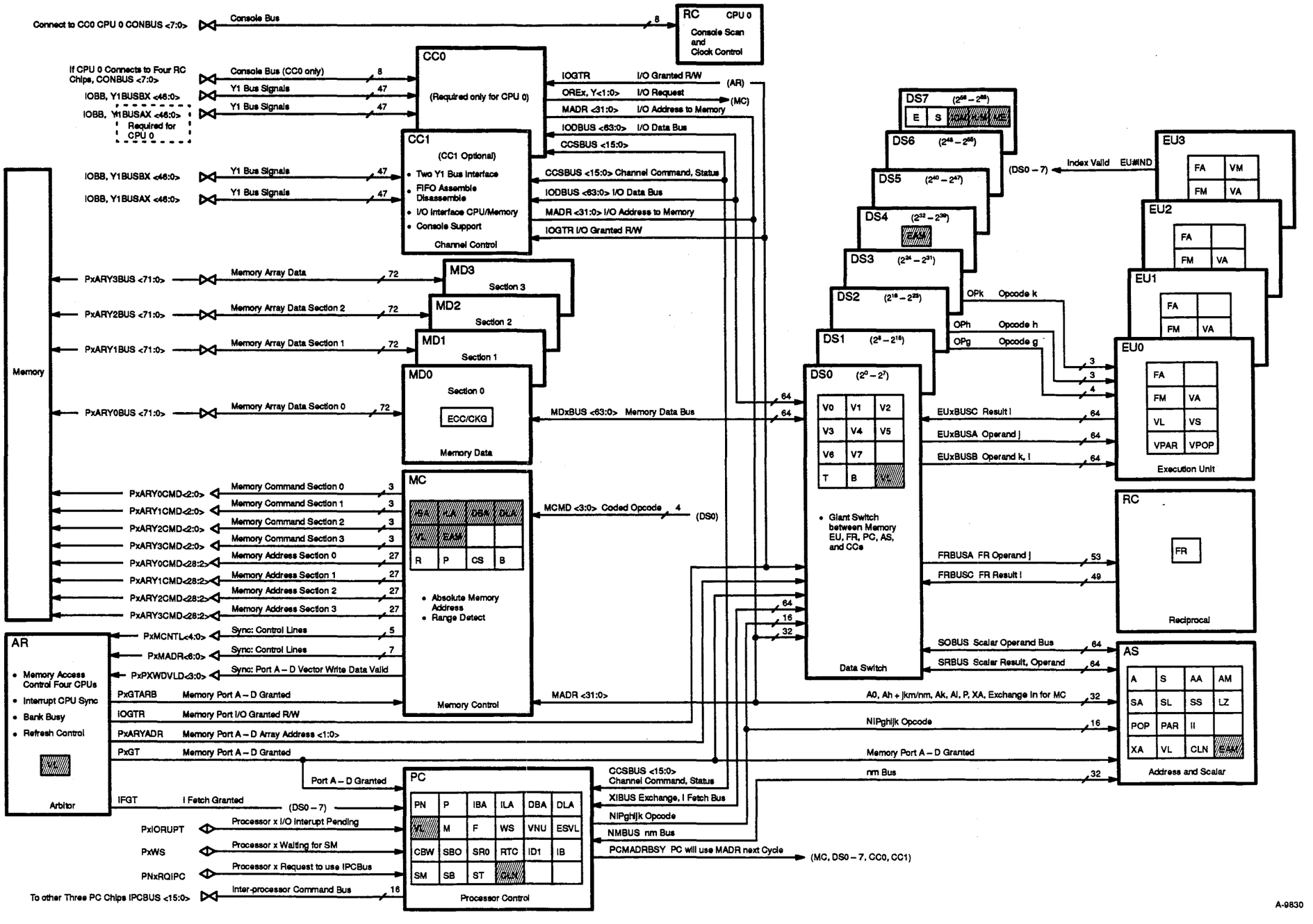


Figure 3-2. CPU Bus Block Diagram

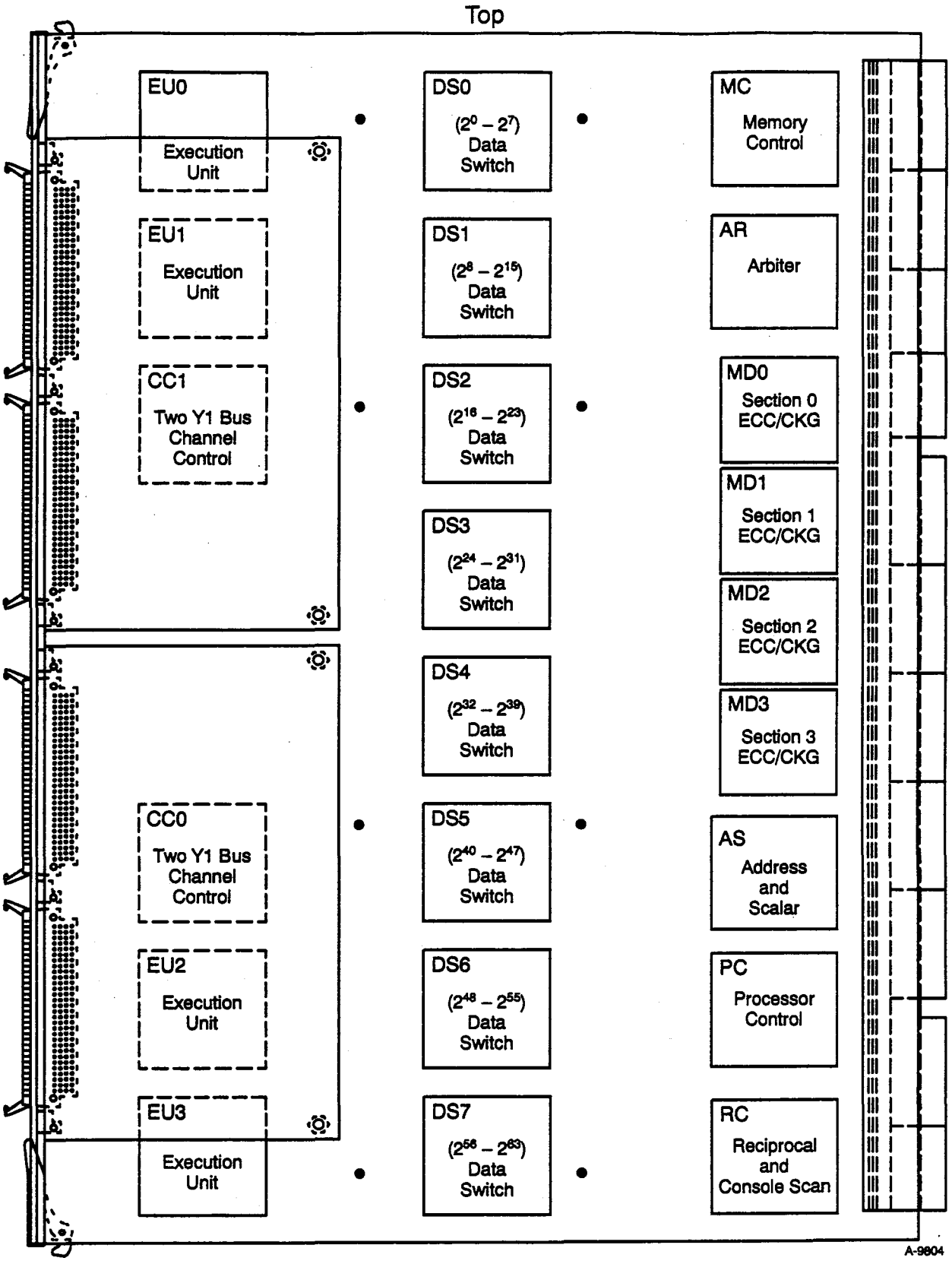


Figure 3-3. CPU/Channel Module

The representative diagram of the exchange package used in the CRAY Y-MP EL system, shown in Figure 3-4, is useful for troubleshooting. The exchange package may contain pertinent information about the state of the CPU at the point of failure. This information can be extracted by comparing the condition of the flag bits, mode bits, and the error word bits with the respective charts shown under the exchange package in Figure 3-4. In some cases, useful information can also be found in the A registers or the S registers, also shown in Figure 3-4. The usefulness of the A or S register information depends on the individual diagnostic used.

The CRAY Y-MP EL system supports one to four CPUs. Each of these CPUs can support up to four VME subsystems. The CPU module is completely self contained and plugs into the mainframe backplane using 11 connectors that provide 1230 signal pins.

When a CPU is indicated as the faulty unit in an incident, the only on-site repair performed is the replacement of the CPU module.

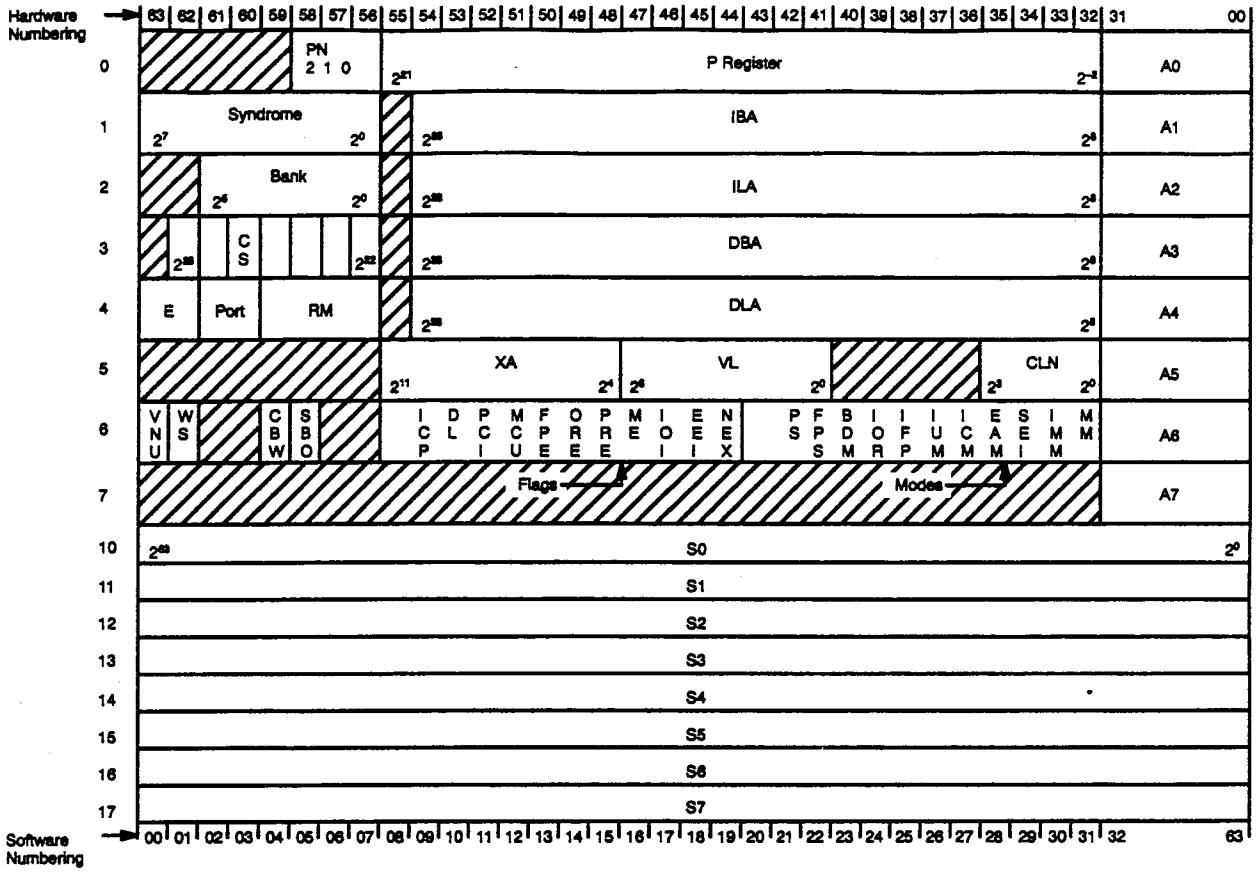
The CPU module connects to the mainframe memory via four 72-bit bidirectional ports. Each of these ports connects to a separate memory section. Each of the CPU modules in a CRAY Y-MP EL system contain a copy of all memory and shared register reservations, to reduce the possibility of conflicts. Each CPU has to check its local request registers, not the request registers of all of the CPUs.

Support for multiple CPUs in the mainframe includes such items as:

- Shared memory
- Shared registers
- Shared I/O channels
- Deadlock detection
- Shared deadstart paths

However, the CRAY Y-MP EL system does not currently support:

- Performance monitors
- High-speed (HISP) channels
- Very high-speed (VHISP) channels



Word 6		Flag Bits	Set Flag	Cause Exchange
55		Reserved		
54	ICP	Interrupt from Internal CPU	MM'	MM'
53	DL	Deadlock Interrupt	MM' and IMM	MM' and IMM
52	PCI	Programmable Clock Interrupt (staged)	MM'	MM'
51	MCU	MCU Interrupt	MM'	MM'
50	FPE	Floating-point Error Interrupt	MM and IMM	MM and IMM
49	ORE	Operand Range Error Interrupt	MM' and IMM	MM' and IMM
48	PRE	Program Range Error Interrupt	MM' and IMM	MM' and IMM
47	ME	Memory Error Interrupt	Always	Always
46	IOI	I/O Interrupt (staged)	MM'	MM'
45	EEI	Error Exit Interrupt	MM' and IMM	Always
44	NEX	Normal Exit Interrupt	MM'	Always

Word 6		Mode Bits
63	VNU	Vectors Not Used
62	WS	Waiting on Semaphores
59	CBW	Concurrent block Write (CRAY Y-MP EL only)
58	SBO	Scalar and block Overlap (CRAY Y-MP EL only)
43		Reserved
42	PS	Program State
41	FPS	Floating-point Status
40	BDM	Bidirectional Memory
39	IOR	Interrupt Operand Range Error
38	IFP	Interrupt Floating-point Error
37	IUM	Interrupt Uncorrectable Memory
36	ICM	Interrupt Correctable Memory
35	EAM	Extended Address Mode (32-bit)
34	SEI	Selected External Interrupt
33	IMM	Interrupt Monitor Mode
32	MM	Monitor Mode

E – Error Word 4						
63	U	Uncorrectable Memory Error				
62	C	Correct Memory Error				
P- Port				RM Read Mode		Port Usage
66 10	66 10	66 10	66 10	55 98	55 76	
A 00	B 01	C 10	D 11	00	01	Exchange
00	01	10	--	00	10	A or S
--	--	--	D xx	01	01	I/O Single
--	--	--	D xx	01	11	I/O Block
00	01	10	--	10	00	B or T
--	--	--	D xx	10	11	Fetch
00	01	10	--	11	00	Vector Stride
00	01	10	--	11	10	V Gather/Scatter

A-9829

Figure 3-4. Exchange Package

4 INPUT/OUTPUT SUBSYSTEM

The input/output subsection (IOS) of the CRAY Y-MP EL computer system is designed to act as a preprocessor and interface between the mainframe section [central processing unit (CPU) and memory] and the various customer-selected peripheral devices. The purpose of this section is to describe devices incorporated within the IOS as well as their functions.

Basic Architecture

The IOS used in the CRAY Y-MP EL system is a VMEbus-based device. This VME is a modular design, and thus is easily adapted to customer requirements. The IOS is available in three different backplane configurations:

- 10-slot + 10-slot
- 10-slot + 6-slot + 4-slot
- 6-slot + 4-slot + 6-slot + 4-slot

Cray Research, Inc. (CRI) uses a VME backplane that supports the double-height 233 mm (6U) x 160 mm printed circuit (PC) board, which fits into a standard 19-inch rack. The IOS is powered by a 1000-watt power supply and is air cooled.

The CRAY Y-MP EL system can support up to four independent IOSs for each installed CPU. Thus, a fully enhanced system containing four CPUs can support 16 IOSs. Each of the IOSs communicates with its associated CPU via a 40-Mbyte/s channel, designated the Y1 bus.

IOS 0 is required and must reside in the primary cabinet. This primary IOS, designated the multiplex input/output processor (MIOP), consists of:

- Type 68030-processor based processor board
- Local memory
- Control store

The primary IOS also includes the following devices that are connected to the small computer system interface (SCSI):

- Winchester disk drive
- Cartridge tape drive (.25-inch)
- Communications port for the operators console
- Remote diagnostic facility
- Optional helical-scan tape system

Many peripheral products can be connected to an IOS. Some of these are:

- Networks using TCP/IP
  - HYPERchannel
  - Fiber-distributed data interface (FDDI)
  - Ethernet
- Several varieties of disk drives
- Several varieties of tape drives

For a more detailed list of peripherals supported by the IOS, refer to Section 5 of this manual.

Each of these devices must be driven by a VMEbus-compatible device controller. Cray Research recommends and supports several of these controllers, but the customer may provide other controllers. The IOS backplane can contain up to 8 of these peripheral controllers (in the 10-slot configuration). However, each IOS must contain an I/O processor (IOP) board and a CPU channel communications interface board.

The IOP selected for the CRAY Y-MP EL system IOS is a type 68030-based board supplied by Heurikon Corporation. This IOP supports the SCSI, allows I/O computation, and controls IOS-to-CPU communications. Communication between the IOS and CPU takes place via a CRI board called the input/output buffer board (IOBB).

IOBB

The IOBB is used to provide buffer memory for the IOP and a communications interface between the IOS and the CPU. IOS-to-CPU communication takes place in the following manner: the IOP generates interrupts to initiate a CPU request, and the CPU generates interrupts to initiate a peripheral read or write operation or to terminate the CPU request. These interrupts are called I/O task control blocks (IOTCB), of which there are two types: I/O IOTCB and console IOTCB. Both of these types have the same format with minor variations (refer to Appendix B).

Once a channel interrupt has been received, the IOP initiates an IOTCB, sends an IOTCB pending interrupt to the CC application-specific integrated circuit (ASIC) on the CPU, and waits for the CPU to acknowledge the interrupt. When the interrupt is received by the CPU, it is handled by the CC ASIC, which acts as the IOTCB processor within the CC ASIC. Refer to Appendix B for illustrations and descriptions of the channel transfer sequences.

The IOBB also detects and reports errors. The Y1 bus enables parity transfers between the CPU and IOBB. The IOBB can thus detect and report parity errors, of which there are two types: errors during IOTCB fetch and errors during IOTCB execution. If a parity error occurs during the IOTCB fetch operation, the CC ASIC is not allowed to execute the IOTCB.

Errors that occur during IOTCB execution can also be of two types: command channel errors or data channel errors. In both cases, the execution of the IOTCB continues to completion, then the YC ASIC initiates a retry. If this retry is not successful, an error message (IOTCB execution error) is returned to the MIOP. At this point, the operating system determines the procedure to follow for further error resolution.

Figure 4-1 shows how the IOBB fits into the communication path employed by the CRAY Y-MP EL system. It must be remembered that the IOBB is a slave device to the IOP connected to it via the VMEbus. Refer to Figure 4-2 for a general block diagram of the IOBB.

Each of the Y1 busses has a channel pair number assigned to it. Because the Y1 bus is bidirectional, it must be considered a channel pair. These channel assignments are as follows:

- CPU 0
  - CC0 - channels 20/21
  - CC0 - channels 22/23
  - CC1 - channels 24/25
  - CC1 - channels 26/27

- CPU 1
  - CC0 - channels 30/31
  - CC0 - channels 32/33
  - CC1 - channels 34/35
  - CC1 - channels 36/37

- CPU 2
  - CC0 - channels 40/41
  - CC0 - channels 42/43
  - CC1 - channels 44/45
  - CC1 - channels 46/47

- CPU 3
  - CC0 - channels 50/51
  - CC0 - channels 52/53
  - CC1 - channels 54/55
  - CC1 - channels 56/57

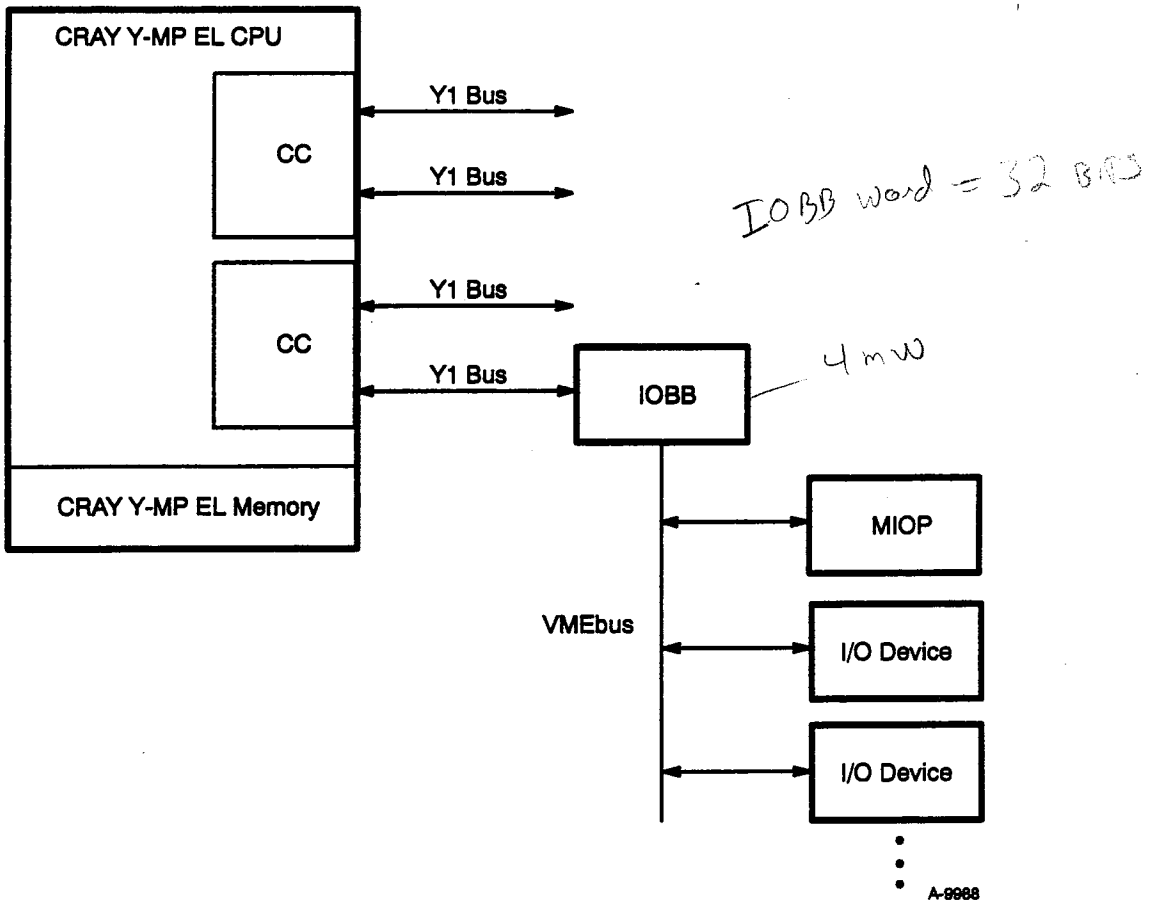


Figure 4-1. IOBB Communication

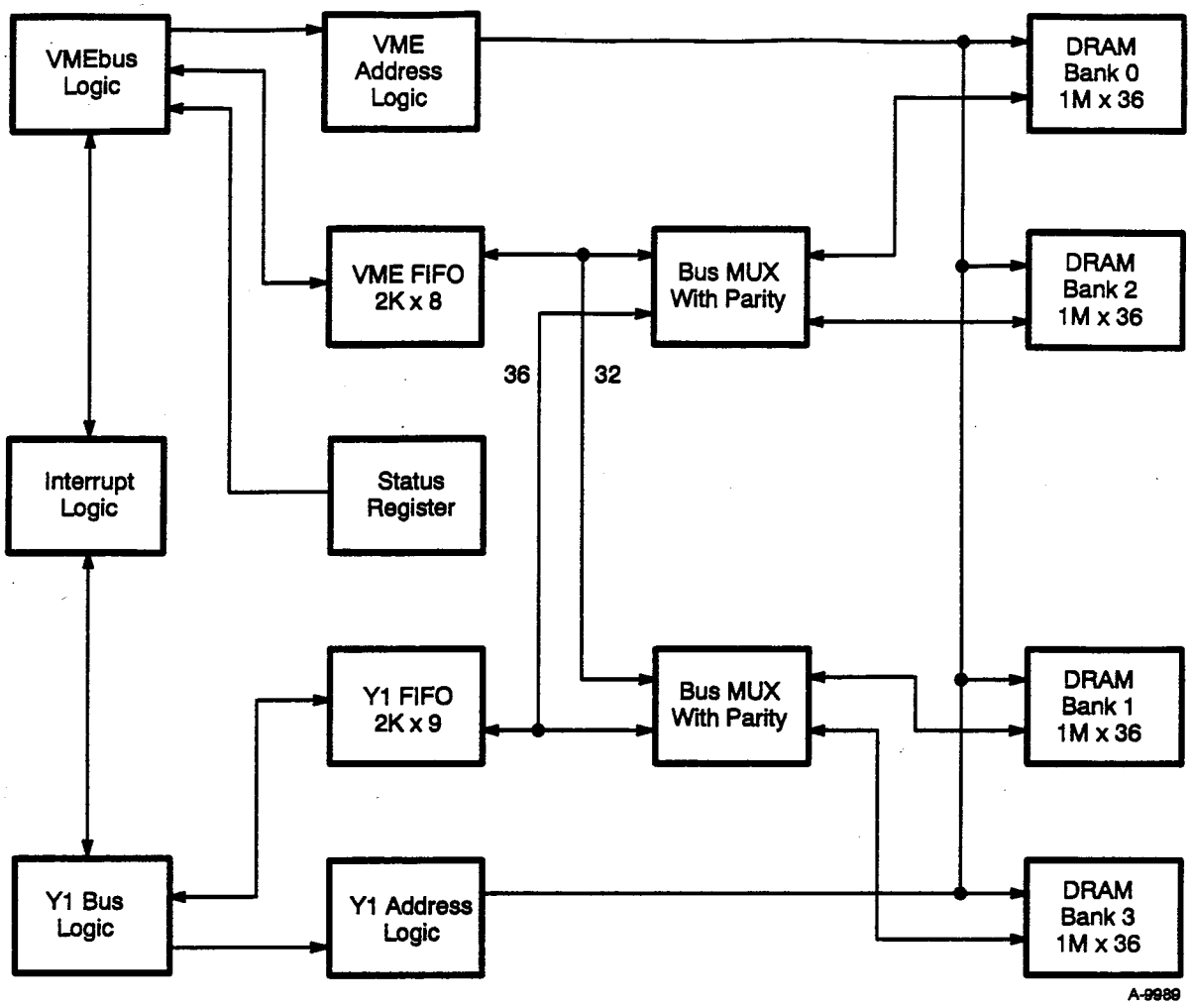


Figure 4-2. IOBB Block Diagram

Communications

No discussion of an IOS would be complete without an introduction to the communications channel. The communications channel in the CRAY Y-MP EL system is called the Y1 bus. The Y1 bus is capable of 40-Mbyte/s transfer rates; these data transfers exist as data bursts of 4, 32, 64, or 128 thirty-two-bit words. The actual data signal is a differentiated value that provides a high level of noise immunity. Refer to Appendix B for a general diagram of the Y1 bus, as well as a table that lists the exact pin assignments.

Communication between the MIOP, the other IOSs, and the maintenance workstation (MWS) takes place via an RS-422 serial interface that can sustain transfers of 1.2 Mbits/s. This is a serial interface; therefore, any point that is defective eliminates the string. Essentially, if the MIOP is

down, the entire system is down. Refer to Figure 4-3 for an illustration of this daisy chain. This figure also shows the communication path from the MIOP to the required system peripherals via the SCSI interface. These system peripherals are the .25-inch cartridge tape drive and the Winchester hard disk drive. ~~The helical scan tape drive is optional, as is the system console.~~ The system console used with the CRAY Y-MP EL system is a Wyse model 60 using an RS-232 interface to the MIOP.

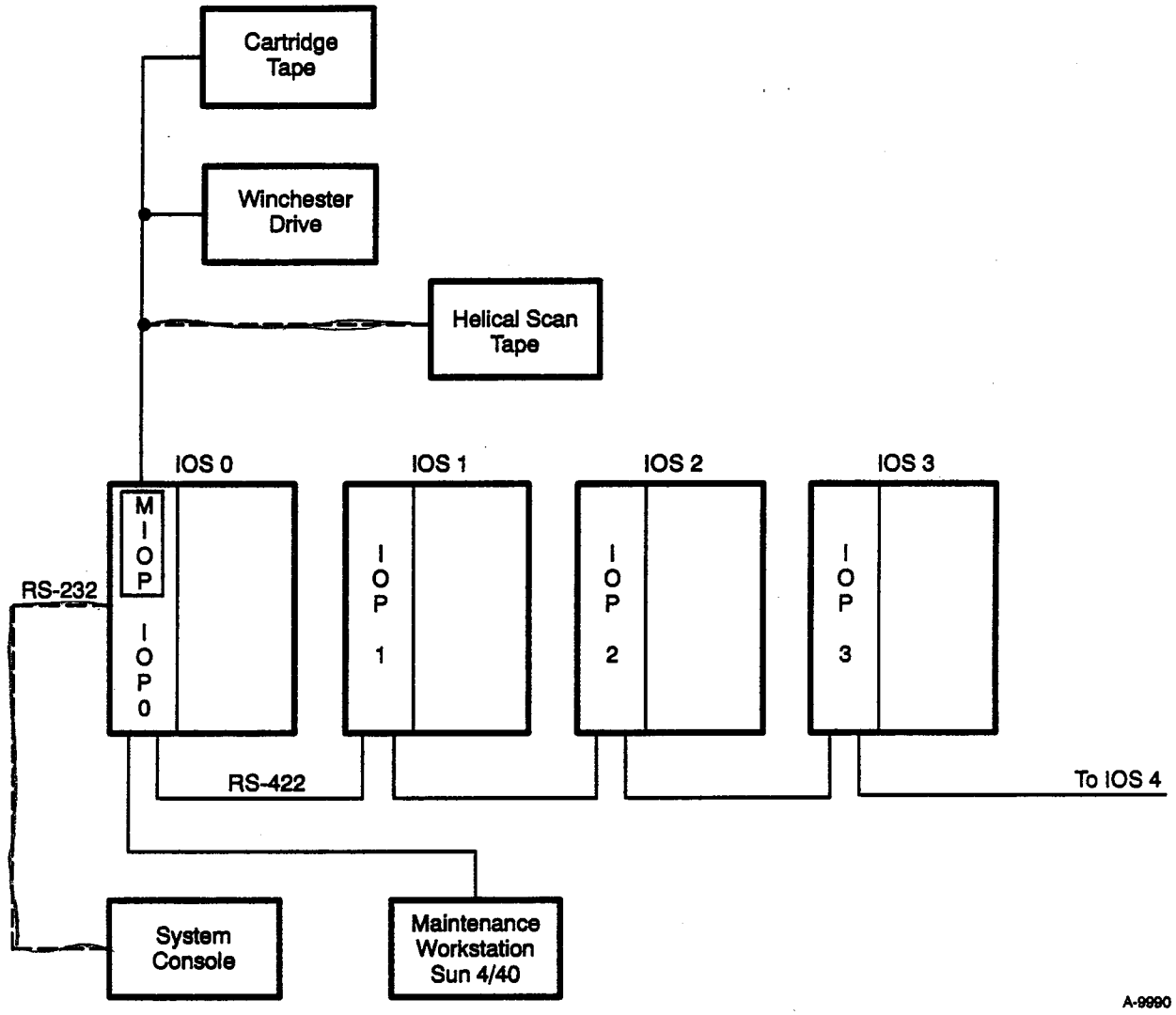


Figure 4-3. IOP Daisy Chain Configuration

IOP 0 must be in the daisy chain and it must be the first in the chain

IOS-supported Peripheral Controllers

The design of the IOS allows many different controllers to be installed. Consequently, many different peripheral devices can be connected to the CRAY Y-MP EL system. The types of controllers supported by Cray Research include:

- Enhanced serial drive interface (ESDI) disk controller with one to four 1.3-Gbyte drives
- IPI-2 disk controller with one or two 2.7-Gbyte drives
- Disk array subsystem (DAS) controller that can control 10 to 40 Gbytes of data
- Network controllers
  - Ethernet
  - FDDI
  - HYPERchannel
- Pertec controller with a 125 ips 9-track tape drive
- SCSI controller for square-cartridge tape drive (18-track)

Refer to Figure 4-4 for a configuration diagram of these devices. To gain a more complete understanding of the individual controllers, refer to the Original Equipment Manufacturer (OEM) manuals supplied with the system. The specific controllers are:

- DC-3 ESDI disk controller
- DC-4 IPI-2 disk controller
- DAS-2 disk array subsystem
  - 10 ESDI disk drives
- IFI1 FDDI
- HC1 HYPERchannel interface
- HI1 High performance parallel interface (HIPPI)
- SI1 VMEbus SCSI host bus
- TC-2 9-track tape drive interface (Pertec)

Refer to Appendix C for more specifications.

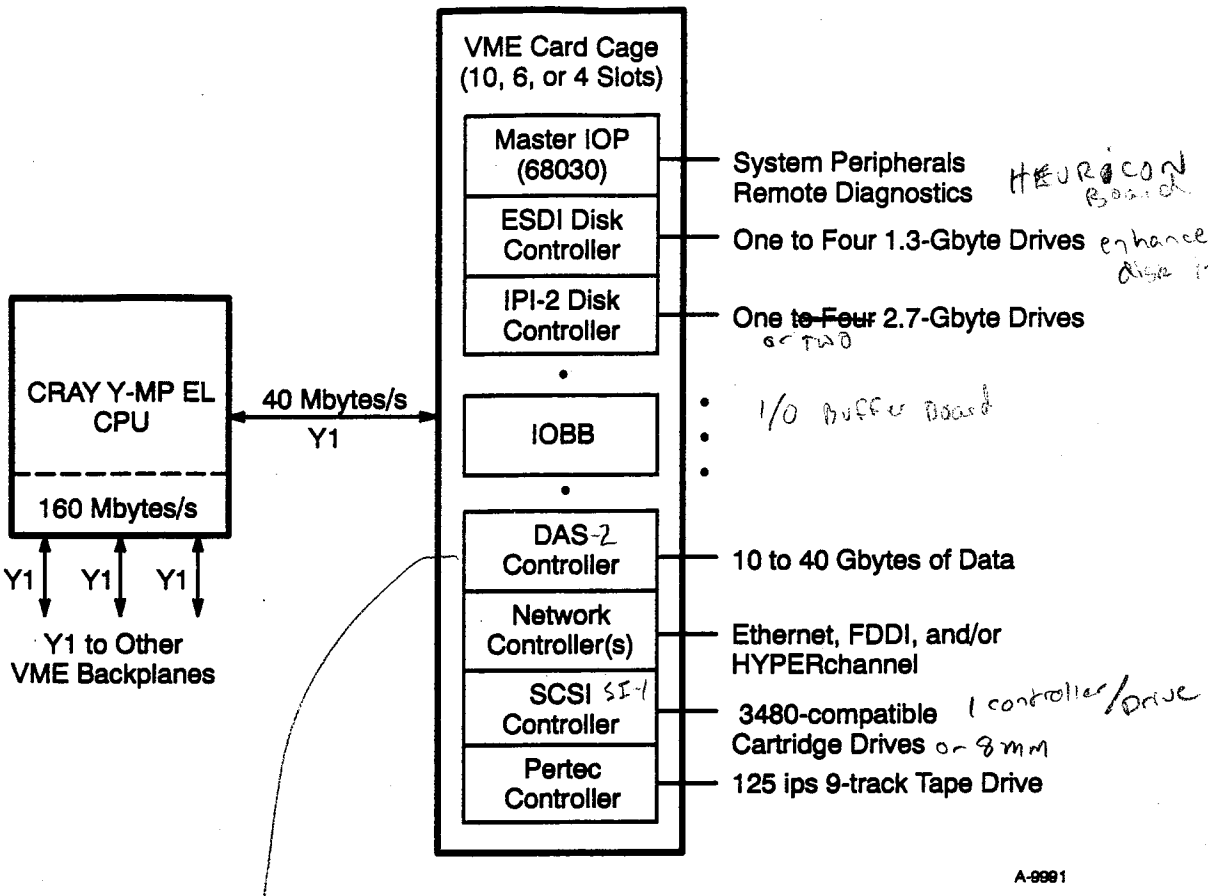
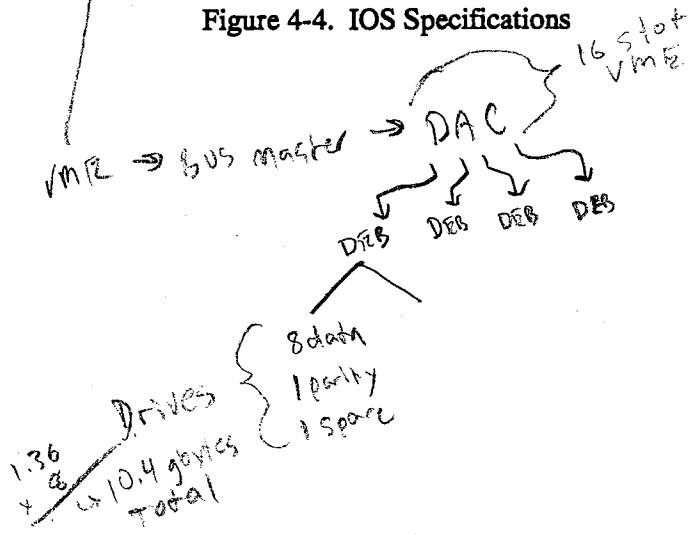


Figure 4-4. IOS Specifications



5 PERIPHERAL DEVICES

The CRAY Y-MP EL computer system is designed with customer peripheral requirements in mind. The CRAY Y-MP EL system can accommodate a wide range of peripheral devices with a diverse amount of storage and communication capabilities. Because of this wide range of capabilities, in most cases it is best to refer to the original equipment manufacturer (OEM) manual for specific information on the equipment. A list of available peripheral devices and their characteristics follows.

- DD-3 Winchester disk drive
  - High-speed access
  - High-speed data transfer rate (2.753 Mbytes/s)
  - Enhanced serial drive interface (ESDI) - industry standard
  - High mean time between failure (MTBF) (150,000 hours, MIL STD 2.17)
  - 1.321 Gbytes of formatted storage
  - No maintenance
- DD-4 IPI-2 (intelligent peripheral interface) disk drive
  - High-speed data transfer rate (9.34 Mbytes/s)
  - High MTBF (150,000 hours, MIL STD 2.17)
  - 2.7 Gbytes of formatted storage
- TD-2 9-track tape drive (covered in detail in this section)
- TD-3 18-track tape drive (covered in detail in this section)
- EX-2 8mm cartridge tape drive
  - High density (1638 tracks/in.; 74 Mbits/sq. in.)
  - Large capacity cartridge (5 Gbytes)
  - Peak transfer rate of 4 Mbytes/s
  - Sustained transfer rate of 500 Kbytes/s
  - MTBF is 40,000 hours
  - Intended as system backup device

- DAS-2 disk array sybsystem
  - 1 disk array controller (DAC)
  - 10 ESDI DD-3 drives forming one bank
  - 8 data, 1 parity, and 1 spare drive
  - 10.4 Gbytes of storage per bank
  - Sustained transfer rate of 15 Mbytes/s
  - Internal media defect management

- DAS-M disk array controller with multiplexer option installed

- DEB-2 disk array bank
  - 8 DD-3 data drives
  - 1 DD-3 parity drive
  - 1 DD-3 spare drive
  - 1 to 4 DEB-2s connect to 1 DAS-M

To obtain a more detailed description of these devices, refer to Appendix C for the design specifications.

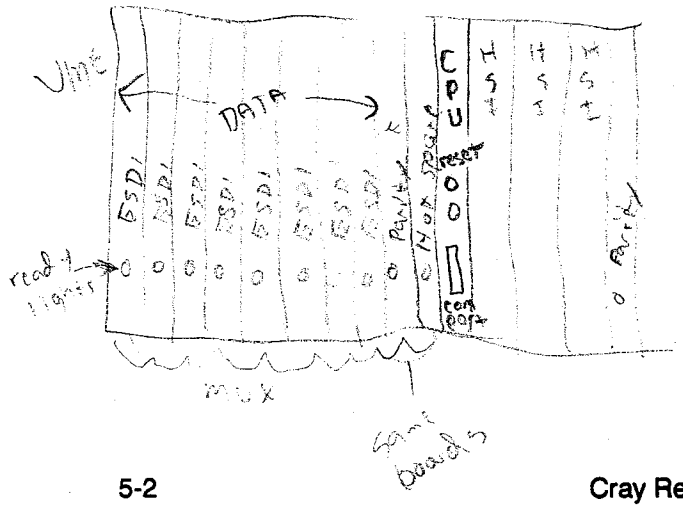
All of the peripheral devices are mounted in the cabinets available. The input/output subsystem (IOS) that controls the peripheral is normally included in the same cabinet. Figure 5-1 shows a typical layout of the various components within the system. The exact placement of any component is determined by the system configuration.

The majority of the CRAY Y-MP EL system peripherals are considered field replaceable units (FRUs). This means that when one of the devices breaks down, it is completely replaced. There are no field repairable parts inside the peripherals, except in the 9-track tape drive (TD-2), the 18-track tape drive (TD-3), and the disk array subsystem (DAS-2) controller. These components are described in the following subsections.

DAS-M - can handle 4 DEBs multiplexed. each ESDI board handles 4 drives

4 DEBs

DAC



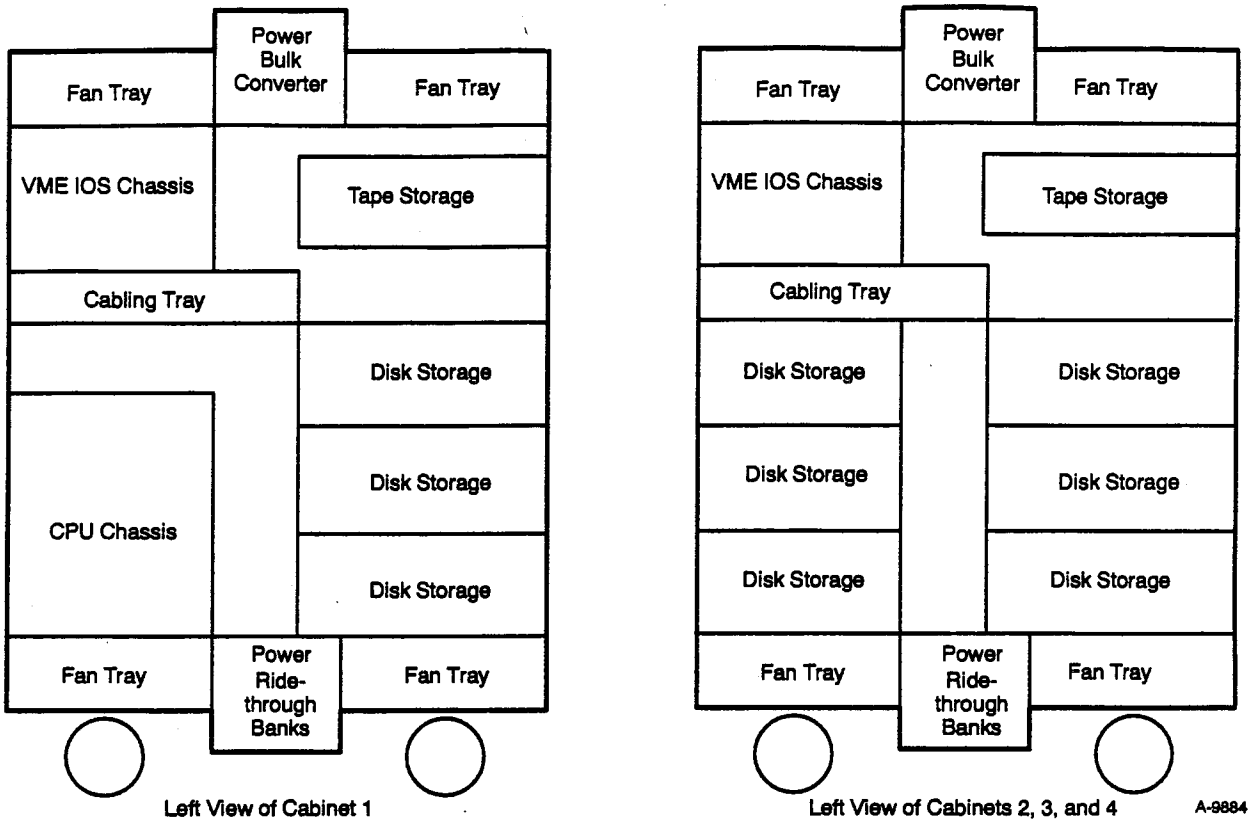


Figure 5-1. Example of Cabinet Layout for the CRAY Y-MP EL System

TD-2 Tape Drive

The TD-2 tape drive (hereafter referred to as the TD-2) is a 9-track tape drive that handles open reel 0.5-in. tape. Any reel up to 10.5 in. fits on the unit. The TD-2 is manufactured by Storage Tek Manufacturing, Ltd. of London as Model 9914. It is designed to handle four types of data encoding:

- NRZI - non-return to zero indicated (800 bpi)
- PE - phase encoded (1600 bpi)
- GCR - group coded recording (6250 bpi)
- DPE - double phase encoded (3200 bpi)

The TD-2 is controlled by the small computer system interface-1 (SI-1) controller card in the VMEbus that connects to a Pertec cache interface within the TD-2. The Pertec interface is one of the FRUs included in the TD-2. Other FRUs in the TD-2 are as follows:

- Data control board (DCB)
- Digital data processor board (DDP)

- Analog data processor board (ADP)
- Servo control board (SCB)
- Power supply board (PSB) (contains two fuses)
- Hub sensor board (HSB)
- In-chute sensor board

Failures on any of these FRUs can be located by using the onboard diagnostics. These diagnostics are run by using the front panel switches on the TD-2. Other than a tape device that indicates failure, there are no provisions for maintaining the TD-2 via remote diagnostics. Faults must be isolated on site.

For complete information on how to run and interpret these diagnostics, refer to the Storage Tek user/diagnostic manual (# 95121796) supplied with the system. Because several combinations of front panels, switches, and diagnostic programs are available, Storage Tek service and diagnostic manuals must be used as references.

When the failure is diagnosed to the FRU level, the FRU must be replaced using the procedures provided in the Storage Tek servicing manual (# 95121797) provided with the system. It should be noted that there are no individual tape-path FRUs. Instead, the entire tape path subassembly is removed in case of failure. When you remove or insert the tape path subassembly, ensure that the Tacho roller is not misaligned on its pivot shaft. The Tacho roller and the heads are the only components that can be damaged, and can cause the TD-2 to malfunction.

As mentioned, the other components that can be damaged during normal removal and replacement are the read/write heads. These heads are thin film heads and are useless if they are scratched. As a precaution when you remove or install the tape path subassembly, cover the heads with an adhesive bandage. Place the cloth pad over the face of the heads and press the adhesive to the sides. Ensure that the adhesive does not contact the face of the heads. After you install the tape path subassembly, clean the heads to ensure proper operation.

The TD-2 Model 9914 is very sensitive to power line voltages. When you install the TD-2, be sure to check the AC power source before inserting the power cord. If the power source measures less than 115 Vac, be sure to set the input power switch to 100 Vac. In all other instances, set the input power switch to match, as closely as possible, the actual supply voltage.

TD-3 Tape Drive

The TD-3 tape drive (TD-3) is also supplied by Storage Tek as the Model 4220 cartridge tape subsystem. The TD-3 is an 18-track tape drive that is IBM 3480 compatible.

The TD-3 comprises several subsystems. These subsystems include:

- Tape transport
- Servo electronics
- Disk drive electronics
- Read/write electronics
- Pneumatics
- Operator control panel
- DC power supply

The disk drive subsystem provides the TD-3 with both functional and diagnostic microcode. The floppy disk drive uses a 3.5-in., 720-Kbyte format, and is controlled by a single integrated circuit (IC) mounted on the system board (SB). Generated signals are sent to or from the drive through the bottom card (BT). The BT contains only foil runners (no active components); therefore, if there is a failure involving the floppy drive, the failing FRU is probably the drive itself or the SB.

As is true in any electro-mechanical device, three types of failure can occur in the TD-3: power-level electrical, mechanical, and signal-level electronic.

Power circuitry for the TD-3 is designed to accommodate either 100 to 120 Vac or 200 to 240 Vac, switch selectable from 47 to 63 Hz, single phase. From this input, the TD-3 supplies + 5 Vdc, +/- 12 Vdc, and + 24 Vdc to the various subassemblies. The only location in the TD-3 where AC power voltages exist is at the DC power supply board (DCPS). All voltages internal to the TD-3 are DC voltages. Refer to Figure 5-2 to locate the power supply.

The mechanical processes of the TD-3 are contained primarily on the tape deck. These mechanical processes deal with loading, unloading, and moving the tape after the cartridge is inserted into the deck. Figure 5-3 and Figure 5-4 show the location of the mechanical FRUs. When a mechanical problem occurs, the problem could be caused by another nonmechanical subassembly. For instance, the servo circuitry could be providing values to the mechanical assemblies that emulate a mechanical problem.

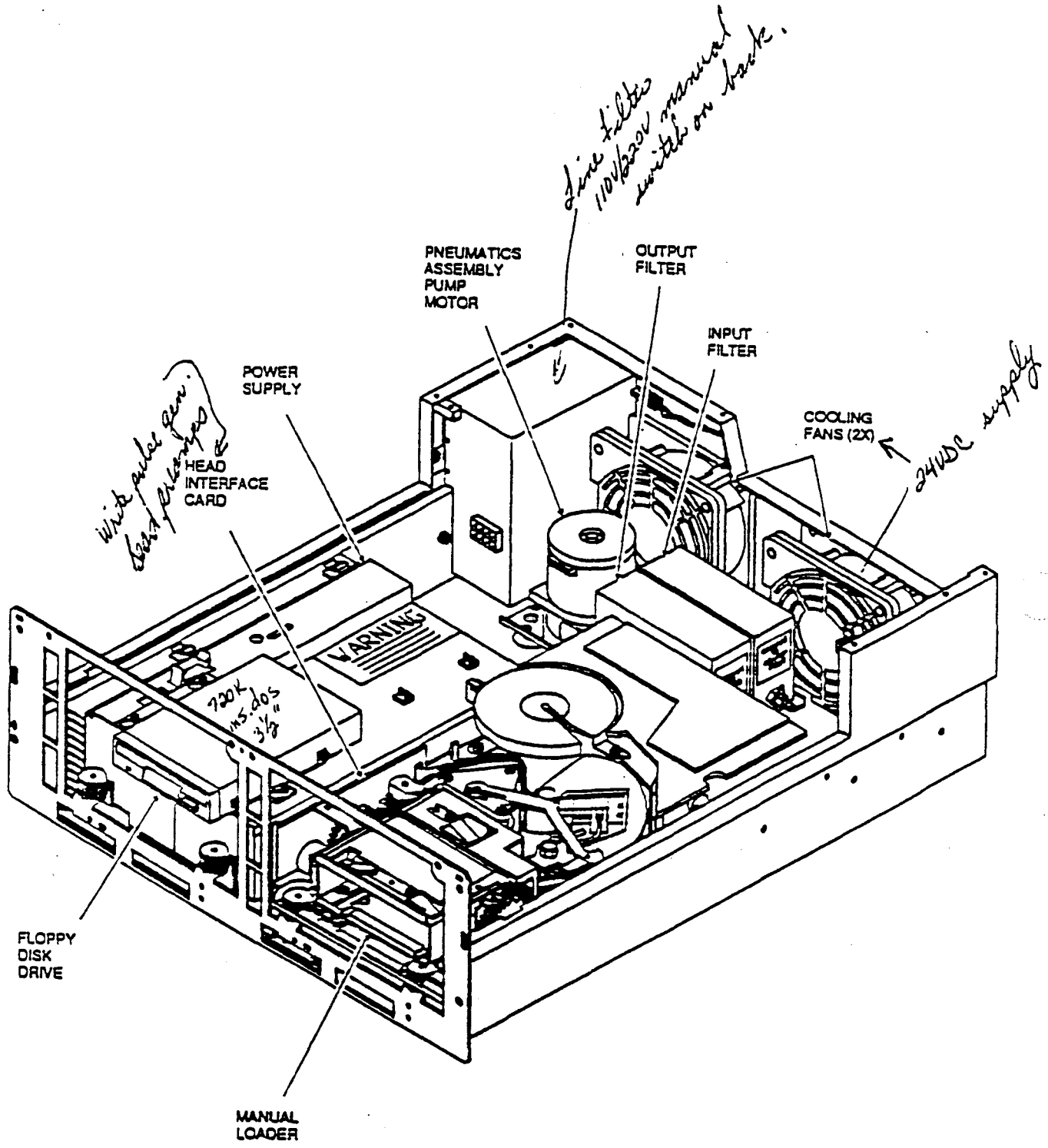


Figure 5-2. TD-3 Tape Drive (Power Supply)

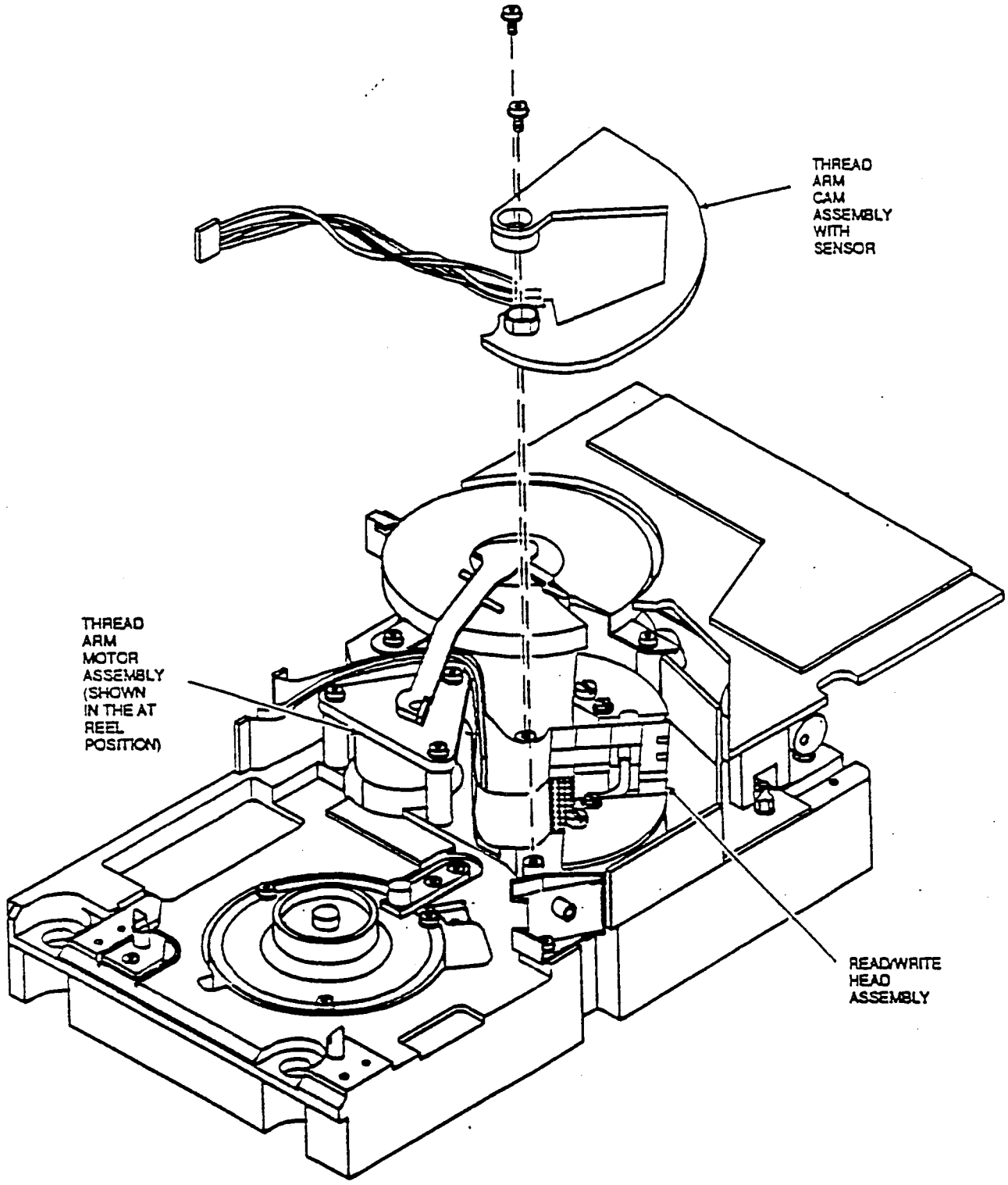


Figure 5-3. TD-3 Tape Drive (FRUs)

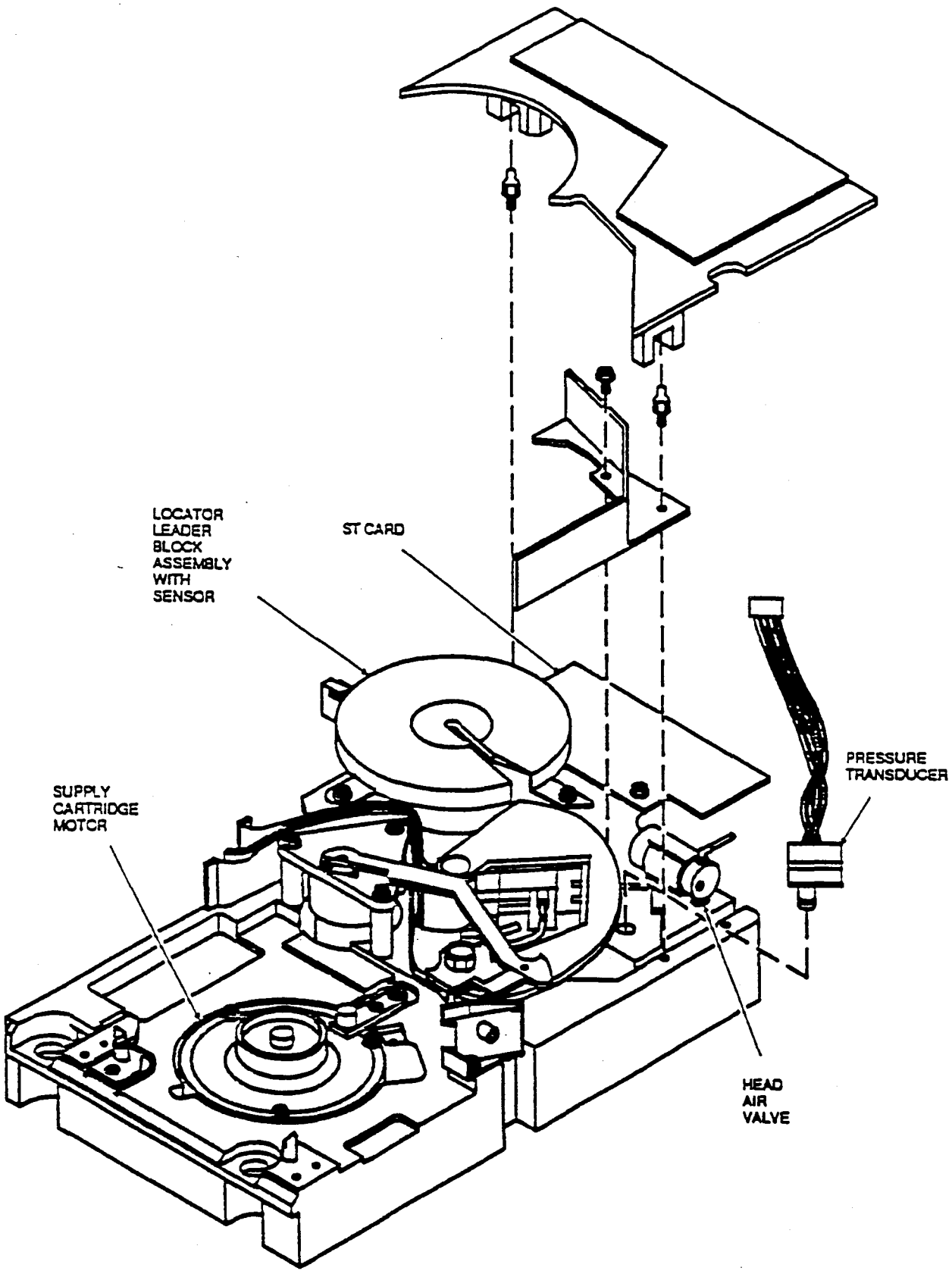


Figure 5-4. TD-3 Tape Drive (FRUs)

The TD-3 cannot be diagnosed from a remote location. A failing FRU within the TD-3 must be identified by using the front panel to load, run, and read the diagnostics and the error codes associated with them. Figure 5-5 shows the front panel of the TD-3. Note that there are four buttons on the panel, two of which are important to loading and running diagnostics: the scroll and the select buttons.

The scroll button advances the display through a menu, one step at a time. The menu is shown in Figure 5-6. When you reach the desired location in the menu, use the select button to select (lock in) the item. Then, push the select button a second time to run the selected item.

For a complete list and a more detailed explanation of the individual diagnostics available for the TD-3, refer to the Storage Tek manual for the Model 4220 that is supplied with the system. That manual contains an extensive explanation of the diagnostics, the error codes as returned to the operator display, and a complete FRU guide.

The following printed circuit (PC) boards comprise the signal-level electronics portion of the TD-3. Figure 5-7, Figure 5-8, and Figure 5-9 show the locations of these PC boards.

- Capacitor (CP) card
- Servo monitor (SM) card
- Servo power (SP) card
- Read interface(RI) card
- System (SB) card

Note that the PC boards reside on or under a tip-up lid on the top of the TD-3 (except the SCSI board) and cannot be accessed unless the TD-3 is extended from the rack. Follow the procedure provided in Section 10 of this manual to extend the TD-3.

The RI card controls generation of clock pulses for each track of data via three phase-locked loop (PLL) oscillators. The RI board also contains the write data encoders and peak detectors.

The SB card is the heart of the TD-3. It contains a 68010 microprocessor that processes host commands, provides I/O interface control, and controls the physical processes relating to tape operations. The SB card also includes a write formatter (WF) IC, a 2-Mbyte buffer, buffer FIFOs, an 8031 bus processor, deskew buffers, a floppy disk controller, and read data formatter devices. Be aware that all operations in the TD-3 are linked to the SB. When you troubleshoot the TD-3, remember that the SB may be faulty.

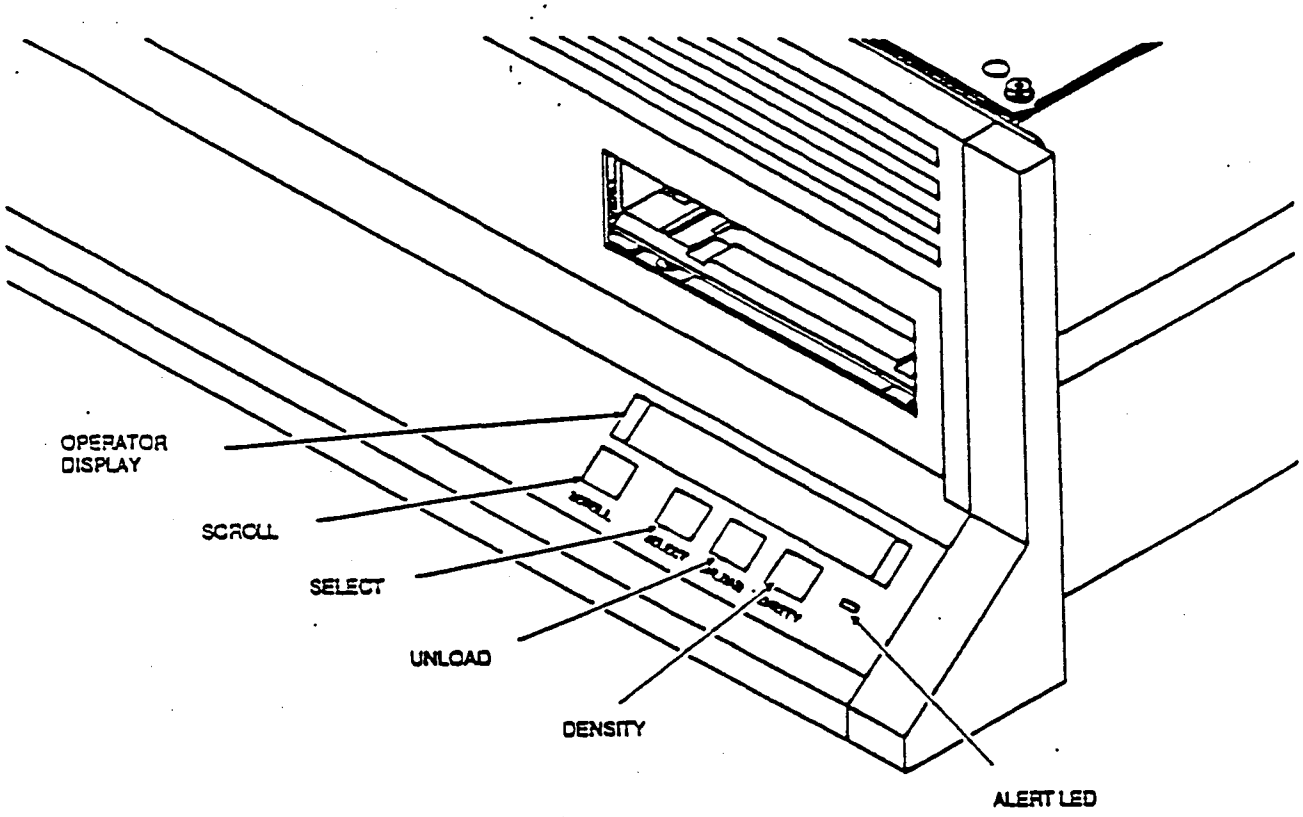


Figure 5-5. TD-3 Tape Drive Front Panel

Offline Menus

```
[OF:DIAG:      *]
[+SUBSYS STATUS]
      [*Online Request]
      [*Exit      ]
[+SET OR DISPLAY]
      [**CONFIGURATION]
            [> > Set Diags:None]
            [> > Diag EC LVL:0]
            [> > Set Host: SCSI]
            [> > Lang: English]
            [> > Set Part: Auto]
            [> > Set Sync Tm: 2]
            [> > Passkey: Enab]
            [> > Sync Nego:Enab]
            [> > Set BRN:   00]
            [> Exit      ]
      [*Retension Tape]
      [*Write Tape Mark]
      [**Set Scsi ID: 0]
      [**Set Lun:   0]
      [**Set Clean:  16K]
      [*Exit     ]
[+DIAGNOSTICS   ]
      [*Diag: Test 1   ]
      [*Diag: Test 2   ]
      [*Diag: Test 3   ]
      [*Diag: Test 4   ]
      [*Diag:Load/Unld]
      [**Diag: Loop 1]
      [Diag: Test Disp]
      [*Loader Sensors]
      [*Disply Pressure]
      [*Disply Tension]
      [*Exit      ]
[+SAVE TRACE   ]
      [*Save Trc: All]
      [*Exit     ]
```

Online Menus

```
[          *]
[ON:IDLE:     *]
[+SUBSYS STATUS]
      [*Offline Request]
      [*Exit     ]
```

Figure 5-6. Basic 4280 Menus with a SCSI Interface

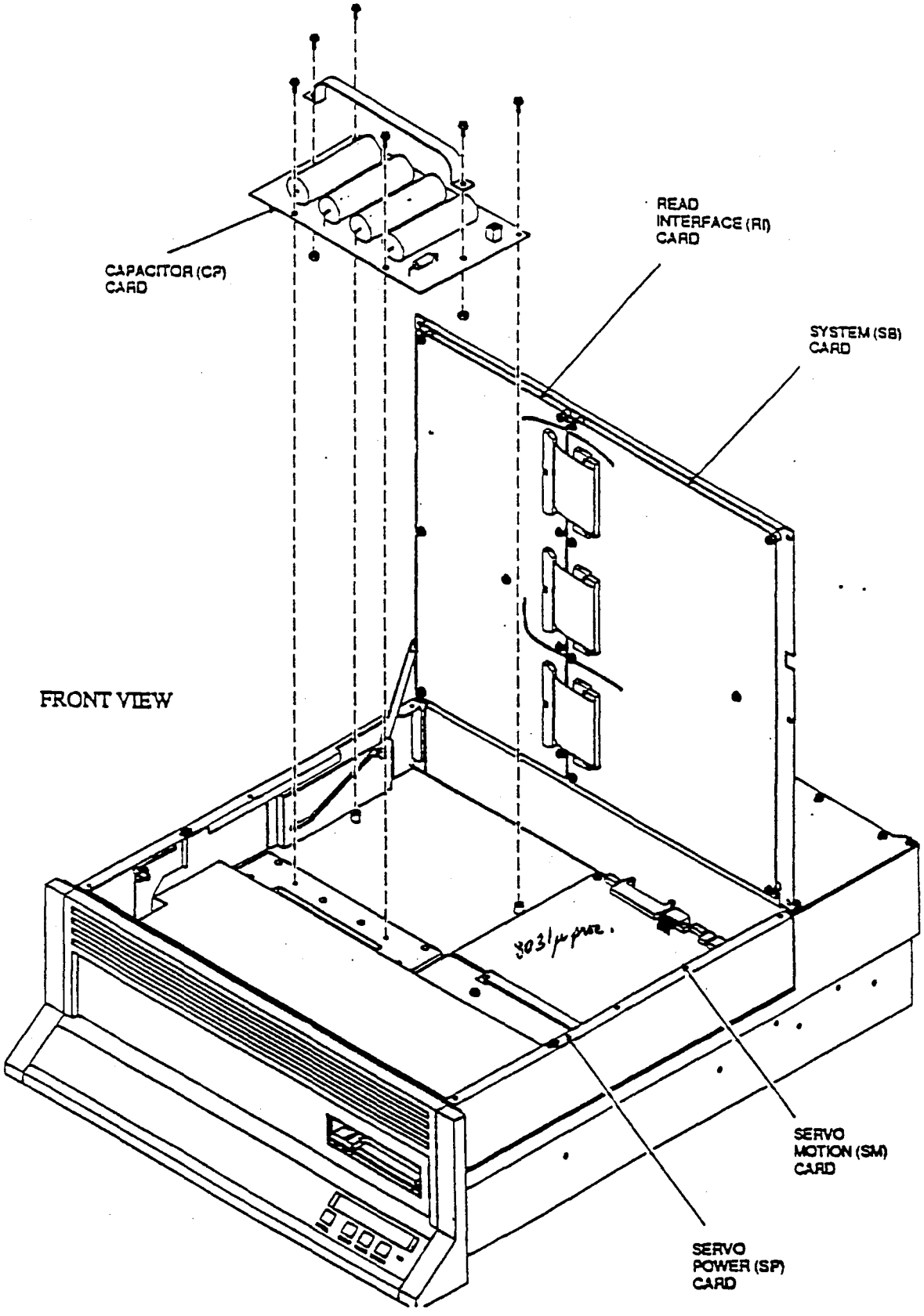


Figure 5-7. TD-3 Tape Drive Controller Module FRUs (Front View)

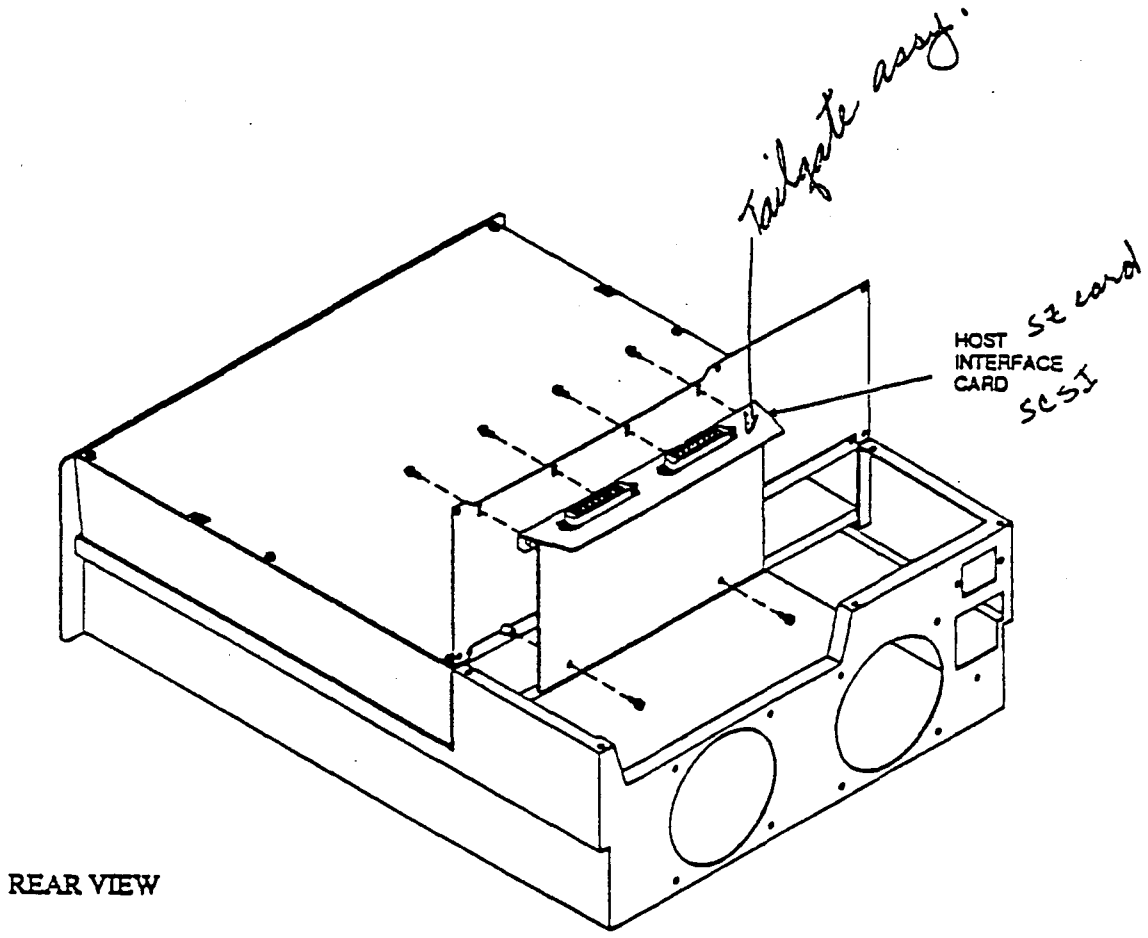


Figure 5-8. TD-3 Tape Drive Controller Module FRUs (Rear View)

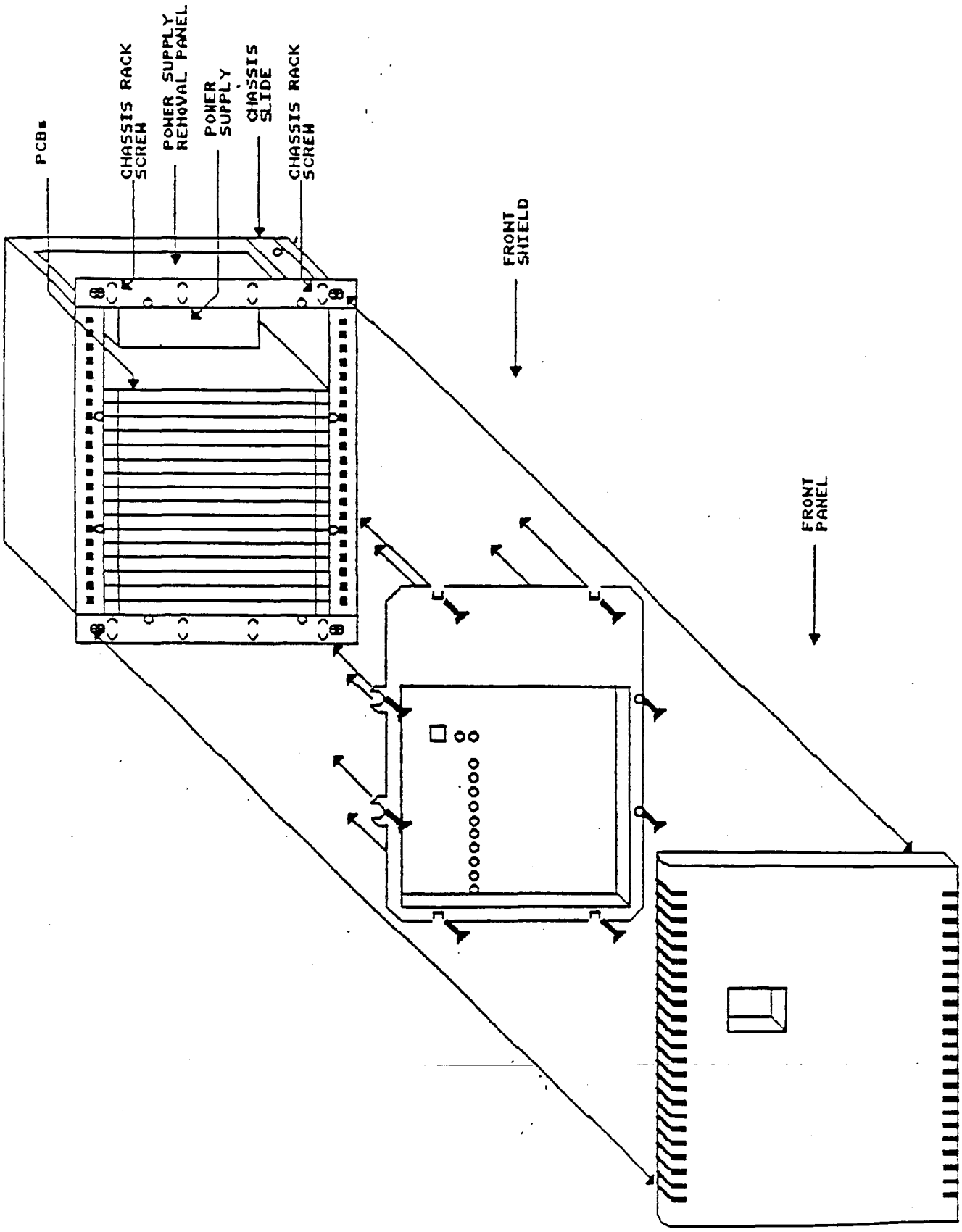


Figure 5-9. TD-9 Chassis

Disk Array Controller

The third non-FRU device in the CRAY Y-MP EL system is the disk array controller (DAC) unit supplied by Maximum Strategy, Inc. The DAC is an intelligent disk controller capable of storing and retrieving data from one to four banks of eight standard ESDI disk drives in parallel. The DAC also controls an additional drive for each bank of eight used to store a parity bit for fault protection as well as an operational standby (hot spare) drive used to replace a faulty drive from the bank.

The drives supported by the DAC are 5.25-in. ESDI serial data drives, capable of transferring data at a rate of 16 Mbytes/s.

Each of the ten disk drives in the bank are connected to a disk interface and data buffer PC board in the DAC. The DAC also contains a parity PC board, separate from the parity drive disk interface PC board, and a CPU PC board. The CPU contains a high-speed interface (HSI) unit. The DAC is designed to interface with up to four banks of ESDI drives. If more than one bank of drives is to be controlled, an additional HSI PC board must be installed for each bank. A multiplexer board must also be installed to link the four banks to the one controller.

Figure 5-9 indicates which components must be removed to access the PC boards, and Figure 5-10 shows the location of each of the PC boards within the DAC.

The DAC has a small cut-out on the front panel (refer to Figure 5-9) in which the READY LED, a RESET button, and RS-232 serial communication (COM) port are visible. The READY LED is located on the CPU card in slot 12, and is a visual indicator of the state of the DAC. When the READY LED is glowing solid red, it indicates that the DAC is offline. During a normal power-on sequence or a reset sequence, the LED glows red as the disk drives spin up. The normal time for this sequence should not exceed 120 seconds. If the LED has not turned green after 120 seconds, it indicates a problem with the DAC or with one of the disk drives. Note in Figure 5-9 that each of the disk interface PC boards has an LED visible through the front shield. As the disk drive associated with the interface board spins up, this LED is red. When the disk drive is up to speed and when the internal diagnostics for the disk interface PC board have run, the LED turns green.

If the CPU READY LED is red, no communication is possible with the DAC. If any one of the data buffer/disk interface LEDs is red, either the disk drive failed to reach speed or the internal diagnostics for that interface board failed. In either case, the next step is to initiate disk drive diagnostics.

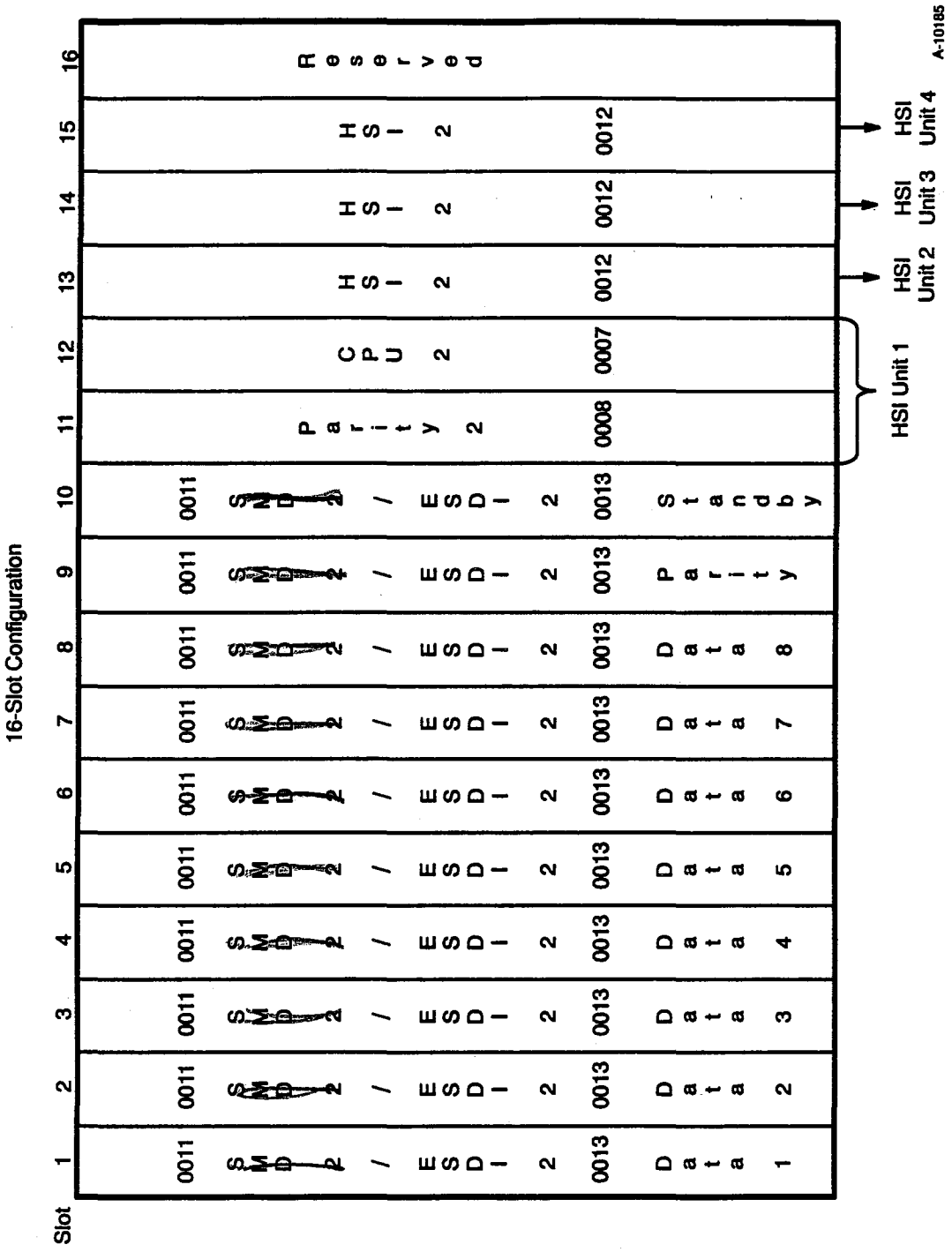


Figure 5-10. DAC Printed Circuit Board Locations

During normal operation, the RESET button should not be used. Instead, the reset function should be initiated by software through the IOS. However, during a power-up sequence you may use the RESET button once if the READY LED fails to turn green within 120 seconds.

The COM port located on the front of the DAC is a serial RS-232 connection that can be used for offline diagnostics or for online system configuration. These functions require test equipment not normally supplied, either on site or at the service center. If you need more information regarding the use of the COM port, refer to the Maximum Strategy documentation supplied with the system or contact Hardware Product Support.

Diagnostics

The peripheral devices associated with the CRAY Y-MP EL system are FRUs. There is only one online diagnostic test, named olcfdt, that checks the functionality of the peripherals. This test is explained in the *Cray Research Entry Level (EL) Computer System UNICOS Online Diagnostic Maintenance Manual*, publication number SPM-1025. The only response of the olcfdt diagnostic to UNICOS is pass or fail. Likewise, there are no offline diagnostics, in place or planned, to deal specifically with the CRAY Y-MP EL system peripherals.

Disk drives for the CRAY Y-MP EL system are driven by intelligent controllers that perform media defect (flaw) management without operator/technician intervention.

6 DIAGNOSTICS

If the CRAY Y-MP EL computer system does not operate properly, a customer employee is the first person notified. This notification usually occurs via a fault code provided by the UNICOS operating system. The customer employee then follows a procedure, defined by the maintenance contract, to gather error information and notify the Cray Research Service Center.

When the service center has been notified, the problem becomes the responsibility of the Cray Research field engineer, who troubleshoots the system using the tools that have been provided for that purpose. Among these tools is a complete set of Cray Research diagnostic tests, both online and offline, that are accessed by remote service when available or by on-site intervention. The purpose of this section is to provide a general description of these diagnostics.

Online Diagnostics

The online diagnostics available for the CRAY Y-MP EL system are listed in Table 6-1. Instructions for running these diagnostics as well as more detailed descriptions of these diagnostics are contained in the *Cray Research Entry Level (EL) Computer System UNICOS Online Diagnostic Maintenance Manual*, publication number SPM-1025. These diagnostic tests are designed to function as a submitted job under UNICOS, and can be used while user jobs are concurrently submitted into the queue.

Table 6-1. Online Diagnostics

Name	Test Description	Test Sections		Monitor	Notes
		Section Select	Function		
olcmon	Online confidence monitor			olcmon	Accepts and interprets commands
olcfpt	Online comprehensive floating-point test				
olcm	Online central memory test	1	Memory storage and scalar path test	olcmon	
		2	Memory storage and T register path test		
		3	Memory storage and B register path test		
		4	Memory storage and V register path test		Uses only the first vector logical unit
		5	Memory storage and V register path test		Uses both vector logical units
		6	Random data test		
		7	Memory conflict test		
olcrit	Online comprehensive random-instruction test			olcmon	Generates both random instructions and random data

Table 6-1. Online Diagnostics (continued)

Name	Test Description	Test Sections		Monitor	Notes
		Section Select	Function		
olcsvc	Online comprehensive scalar and vector comparison test			olcmon	Tests V registers, functional units (FUs) and paths; S registers, FUs, and paths; A registers, FUs, and paths must be functional
olibuf	Online instruction buffer test			olcmon	Tests data, in-stack, and out-of-stack jumps
olsbt	Online semaphore, shared B, and shared T test			olcmon	
olcfdt	Online confidence test for mass storage devices				

A group of diagnostic tests designed to be used in devices that have been removed from normal system operation is also available. These programs are still run as online programs, but are known as down device programs. They provide testing for individual central processing units (CPUs) and peripheral equipment. A list of the down device programs is provided in Table 6-2.

Two utility programs are also available as online maintenance devices. These utilities are:

- olhpa - A hardware performance analyzer that analyzes and reports the hardware errors and statuses reported in the system error log.
- runsequence - A utility used with the crontab command to perform automatic test sequencing.

More detailed information on both the down device tests and the utility programs is contained in the *Cray Research Entry Level (EL) Computer System UNICOS Online Diagnostic Maintenance Manual*, publication number SPM-1025.

Table 6-2. Down Device Programs

Name	Test Description	Monitor	Notes
oldmon	Down CPU monitor	self	
unitap	Online magnetic tape test		

Offline Diagnostics

In most cases, the online diagnostic tests should lead you to a logical swap and repair operation. However, for those instances when the online diagnostics do not identify the problem, there is a full suite of offline diagnostic tests available. The major disadvantage of using offline tests to diagnose a problem is that the system must be offline; that is, no user jobs can be running and the operating system must be turned off.

When the system has been taken offline, the product specialist can call up any of the offline diagnostics listed in Table 6-3, Table 6-4, and Table 6-5. There are two separate levels of diagnostics listed: level 0 and level 1. The level 0 tests are used as dead machine tests, while the level 1 tests are used as failure-specific tests.

All of the listed offline diagnostic tests are run under the mainframe maintenance environment (MME) on the maintenance workstation (MWS). The MME is a menu-driven program designed to work in all Cray Research computer systems. For detailed information on how to use MME and what it does, refer to the *CRAY Y-MP EL Offline Diagnostic User Guide*, publication number CDM-xxxx-PR1. This manual also contains a complete description of all the available offline diagnostic tests, including standard locations, restrictions, and the type of monitor used.

Table 6-3. Offline Diagnostics - Level 0 Tests

Name	Test Description	Notes
INTA	Checks integer instructions 030*ijk* through 033*ijk*	
YEXCH	Tests exchange package register	Tests exchange paths
YTD0	Tests instruction buffers, parcel 0	Tests ability to do fetch
YTD1	Tests instruction buffers, parcel 1	
YTD2	Tests instruction buffers, parcel 2	
YTD3	Tests instruction buffers, parcel 3	

(Scan diagnostic) scan pool command buffer

Table 6-4. Offline Diagnostics - Monitors

Name	Description	Notes
YM8	Runs all processors with same diagnostic	
YMI	Interrupt-driven monitor	
YMM	Basic monitor	
YMS	Interrupt-driven, master/slave monitor	One diagnostic in all processors, reported by master
YSMI	Multi-function monitor	
YSMY	Multi-function monitor	Replaces all other monitors except M8

Table 6-5. Offline Diagnostics - Level 1 Tests

Name	Test Description	Test Sections		Monitor	Notes
		Section Select	Function		
BAAT	Basic address-adder test	One section only		OWN	Tests 030*ijk* and 031*ijk* instructions
CRIDY	Multiple-CPU random-instruction test	One section only		YSMI	
EEUTE	Execution unit (EU) test	One section only		YSMY	
YAAB	Address register basic test	Condition 1	022*ijk* instructions	MM or YSMY	
		Condition 2	020*ijk* instructions		
		Condition 3	021*ijk* instructions		
		Condition 4	030*ijk* instructions		Ai = Aj + Ak
		Condition 5	030*ijk* instructions		Path test
		Conditions 6 – 10	Special case instructions		030*i*0*k*, 030*ij*0, 031*i*00, 031*i*0*k*
		Condition 11	020 and 021 instructions		Walking a 1 through the unused *jk* fields
YAHT	Ah addressing test	0 – 9		MS or YSMY	
		7	Will not run unless CPU 0 enabled in location CPUS		
		8	Automatically disabled if MS not used		
YAMB	Address multiply basic test	0	K = 1; J = sliding ones	MM	
		1	j = 1; k = sliding ones		
		2	Predetermined operands for enables and satisfies		

Table 6-5. Offline Diagnostics - Level 1 Tests (continued)

Name	Test Description	Test Sections		Monitor	Notes
		Section Select	Function		
YAMB (cont.)		3	Predetermined operands for carries		
		4	Random increasing operands		
		5	Path test		
YARS	Address (A) and scalar (S) register add and multiply test	0	A register subtract	YSMY, MI, M8	All sections use toggle numbers of a 1, a 3, a 5, and a 4-bit random number in sequence
		1	A register add		
		2	A register multiply		
		3	S register subtract		
		4	S register add		
		5	S register multiply		
YAVE	Vector register test			MI, YSMY	
YBRT	Block transfer register test	0	B register basic and block transfer	MS, YSMY	Sections 0, 1, and 2 do not use B00 for return jumps; can detect erratic B00 operation
		1	T register basic and block transfer		
		2	V register basic and block transfer		
		3	B, T, and V registers comprehensive block transfer		

Table 6-5. Offline Diagnostics - Level 1 Tests (continued)

Name	Test Description	Test Sections		Monitor	Notes
		Section Select	Function		
YBTAS	B to A register transfer or T to S register transfer test	0	Tests 025*ijk*	MI, YSMY	
		1	Tests 025*ijk*		
		2	Tests 075*ijk*		
		3	Tests 024*ijk*		
		4	Tests random combinations of the four instructions		
YCMPY	Compatibility mode test			OWN	Determines CRAY Y-MP compatibility to CRAY Y-MP EL
YEJT	Exchange jump test	0	Tests A0 through S7 registers using scalar memory instructions	OWN	
		1	Tests (mn + Ah + DBA), (A0 + Vj + DBA), and (A0 + Ak + DBA) address adders		Uses sliding ones address pattern
		2	Tests DBA range error		
		3	Tests DLA range error		
		4	Tests instructions ability to set/clear mode bits		
		5	Tests 076*ijk*, 077*ijk*, and 0014*j*3 instructions		

Table 6-5. Offline Diagnostics - Level 1 Tests (continued)

Name	Test Description	Test Sections		Monitor	Notes
		Section Select	Function		
YEJT (cont.)		6	Tests for operand and program range errors		
		7	Tests address adders using random operands		
YFPT	Floating-point units test			Mi, M8, YSMY	Tests floating-point add, multiply, and reciprocal functional units
YJPT	Jump test	0	Tests 006*ijkm* instruction	OWN	Only runs in CPU selected as master
		1	Tests A0 conditional branch instructions		S4 = condition counter; A0 = failing data pattern
		2	Tests S0 conditional branch instructions		A1 = condition counter; S0 = failing data pattern
		3	Tests P to B00 path		A1 = condition counter; S0 = bits in error; S1 = actual data; S2 = expected data
		4	Tests B00 to P		A1 = condition counter; S0 = bits in error; S1 = actual data; S2 = expected data
		5	Bxx to P		A1 = Bxy; S0 = bits in error; S1 = actual data; S2 = expected data

Table 6-5. Offline Diagnostics - Level 1 Tests (continued)

Name	Test Description	Test Sections		Monitor	Notes
		Section Select	Function		
YJPT (cont.)		6	Various IBA jumps		
		7	Times in-buffer jumps; confirms that no fetch was done		A1 = condition counter; A7 = loop counter; B00 = address of failing code
		8	Parcel 1 of instruction in one instruction buffer; parcel 2 in different instruction buffer		A1 = condition counter (equals instruction buffer under test); A7 = loop counter
YPAVE	Memory addressing, conflicts, and port select tests	0	1 and 0 data; address limited to 11,000 words	MI, M8, or YSMY	
		1	Random data; address limited to 11,000 words		
		2	Random data; address unlimited		
		3	Random data; gather/scatter		
		4	Abort on operand range error		
YPRTC	Real-time/programmable clock test	0	RTC data integrity test	MI or MS	
		1	RTC propagate carry test		
		2	RTC path test		
		3	Basic PCI mechanism test		Forces time-out condition
		4	Basic CCI and DCI mechanism test		
		5	Basic ICD register test		

Table 6-5. Offline Diagnostics - Level 1 Tests (continued)

Name	Test Description	Test Sections		Monitor	Notes
		Section Select	Function		
YPRTC (cont.)		6	PCI II register timing test		
		7	RTC fanout test		
YSAB	Scalar register basic test	0	Checks zeroes, ones, and alternate patterns	MM	Tests S registers and S adders
		1	Checks compliment patterns mn to S		
		2	Tests 060 instructions		Predetermined operands
		3	Path test		
		4	Tests 023*ij*0 instruction		
		5	Tests 071*i*0*k* instruction		
		6	Tests 040 and 041 instructions		
YSAS	Scalar adder test	0	Address add	MI	
		1	Scalar add		
		2	Address multiply		
		3	Address subtract		
		4	Scalar subtract		
		5	Population count and leading zero		
		6	A to S floating-point		
		7	A to vector length (VL) register		

Table 6-5. Offline Diagnostics - Level 1 Tests (continued)

Name	Test Description	Test Sections		Monitor	Notes
		Section Select	Function		
YSCL	Scalar logical basic test	Condition 0	042 instruction	MM	
		Condition 1	043 instruction		
		Condition 2	044 instruction		Conditions 2 through 7 are tested with no S0 operands
		Condition 3	045 instruction		
		Condition 4	046 instruction		
		Condition 5	047 instruction		
		Condition 6	051 instruction		
		Condition 7	050 instruction		
		Condition 10	044 instruction		Conditions 10 through 15 are tested with S0 operands
		Condition 11	045 instruction		
		Condition 12	046 instruction		
		Condition 13	047 instruction		
		Condition 14	051 instruction		
		Condition 15	050 instruction		
YSCS	Scalar shift test	0	Shift all S registers left and right by a count of 1, 2, 4, 10, 20, and 40	MM or YSMY	
		1	Shifts an S register left and right by all bits of an A register 1, 2, and 4 until A is negative		
		2	Double shift left and right; random shift count and data		

Table 6-5. Offline Diagnostics - Level 1 Tests (continued)

Name	Test Description	Test Sections		Monitor	Notes
		Section Select	Function		
YSCS (cont.)		3	Selectable parcel; random shift instruction with random data and shift counts		
YSEM	Shared and semaphore register test	0	Shared B/T address and control tests	MS or YSMY	
		1	Shared B/T data test		
		2	Semaphore set/clear test		
		3	Semaphore master/slave or single CPU test		
		4	Multi-CPU deadlock timing test		
		5	Semaphore master/multiple slave CPUs		
YSMT	Check bit generation and memory error reporting test	0	Tests correct SECDED generation	YSMI	
		1	Tests exchange package fields associated with memory errors		
		2	Memory errors using A and S registers		
		3	Memory errors using B and T registers		
		4	Memory errors using V registers		
		5	Memory errors on exchange		

Table 6-5. Offline Diagnostics - Level 1 Tests (continued)

Name	Test Description	Test Sections		Monitor	Notes
		Section Select	Function		
YSMT (cont.)		6	Memory errors on fetch		
		7	Memory errors for I/O single and block references		
		8	Proper error reporting with back-to-back memory errors		
		9	Tests 001501 instruction		
		10	Tests error bits in status register		
YSR3	Register and scalar/vector instruction (3-parcel) conflicts test			YSMI	
YVBT	Vector register basic test	0	Scalar to vectors	MI or M8	Contains some special error locations:
		1	Vector logical		2001 - failing section
		2	Vector add		2002 - failing condition
		3	Vector shift		2003 - failing subcondition
		4	Vector mask/ compressed index		2004 - failing vector
		5	Vector population/ parity		2005 - failing element
		6	Second vector timing		2006 - vector length
YVPT	Vector path test	0	Vector register test	MI or M8	
		1	Vector path test		
		2	Different paths		

Table 6-5. Offline Diagnostics - Level 1 Tests (continued)

Name	Test Description	Test Sections		Monitor	Notes
		Section Select	Function		
YVPT (cont.)		3	Bit counter		
		4	140 - 177 instruction test		
YVSG	Gather/scatter test	0	Gather using port B	MI	
		1	Scatter using port C		
		2	Gather using port B		Port A conflicts
		3	Scatter using port C		Port A conflicts
		4	Gather using port B		Port C conflicts
		5	Scatter using port C		Port B conflicts
		6	Gather using port A		Port B conflicts
		7	Scatter using port C		Port A and port B conflicts
		10	Gather using port A		Port B and port C conflicts
		11	Scatter using port C		Port A and port B conflicts
		12	Gather using port A		Port B chaining and port C conflicts
		13	Scatter using port C and gather using port B		

7 POWER AND CONTROL SYSTEMS

To be provided.

8 UNICOS ADMINISTRATION

Because the CRAY Y-MP EL computer system has been designed to operate at unattended sites, it is possible that a hardware product specialist could be requested to perform some functions traditionally performed by a software product specialist. While this section can not convey the entire spectrum of software knowledge needed to administer a UNICOS system, it contains enough basic information to allow the hardware specialist to perform some of the basic functions of the system administrator. If you need more information than is provided here, Cray Research recommends that you attend the CRAY Y-MP EL system administrator course offered by Software Training. Another source of information that may be useful is the *UNICOS Installation Guide*, publication number SG-2112.

You need to refer to the latest System Installation Bulletin for the specific release of UNICOS you are working with. In order to run UNICOS 6.1 on the CRAY Y-MP EL system, the Heuricon HK68/V30 input/output processor (IOP) board must have a minimum hardware ECO level of 48. There is a Cray Research PROM chip supplied for installation on the IOP board. This PROM chip must be at revision (REV) level 5.3 to allow proper installation of UNICOS 6.1. You may also need the backup copy of the tape made by the system administrator when he/she installed the UNICOS version originally. If you are using a backup version of a tape, use caution so that no damage is done to the tape.

The release shipment for the current input/output subsystem (IOS) small computer system interface (SCSI) configuration consists of two tapes: an 0.25-in. cartridge tape and an 8-mm helical scan tape, along with pertinent release documents. At the site, there should also be a backup set of tapes. Use the procedure on page 3 of the *UNICOS 6.1 System Installation Bulletin for CRAY ELS Systems*, publication number EL-2.1-SIB, to install IOS software contained on the 0.25-in. cartridge tape. This procedure takes approximately ___ minutes.

The entire IOS and UNICOS installation will require approximately three hours to complete. During this time, the CRAY Y-MP EL system is without memory, so is removed from service. Reloading the UNICOS operating system should not be considered a valid troubleshooting process.

File manipulation, file recovery, dump recovery, and data recovery should not be attempted without proper training.

Tapes

ROOT	2
IROOT	1
USR	2
IOS	2 (both same

YEL

UNICOS

SIB

As of 11/25/91

Modified by Bill Webster

3 steps

Boot IOS

Boot IROOT, Build configuration

Boot configured root

reset - loads prom (very basic commands)

RT> mkfs c: (if drive not formatted then format c:)

load IOS tape into streaming drive (rst0)

RT> Tar -xvf rst0 (tape archive) (writes ios kernel to scsi drive)
↑
extract
(c = create)

RT> cp /install/config ^ /config
cp /install/config.usr ^ /sys/config.usr
cp /install/unicos.is ^ /sys/unicos.is
cp /install/boot ^ /bin/boot

RT> load (loads IOS kernel)

Software reset = call 040

IOS> /install/idiroot

load tape

IOS> /bin/boot

unicos
↓
install> cp sample_config ^ new_config

> vi new_config (change new_config to match system)

> bconfig new_config (tells you problems)

> exdf -o sys/config.usr < new_config

From wmw@tngstar.cray.com Wed Nov 13 13:09:03 1991
Return-Path: <wmw@tngstar.cray.com>
Received: from timbuk.cray.com by techops.cray.com
        id AA27328; 5.61/CRI-5.6b; Wed, 13 Nov 91 13:09:01 -0600
Received: from ansel.cray.com by timbuk.cray.com (4.1/CRI-MX 1.61)
        id AA03945; Wed, 13 Nov 91 13:08:58 CST
Received: by ansel.cray.com
        id AA04396; 4.1/CRI-4.4; Wed, 13 Nov 91 13:07:13 CST
Date: Wed, 13 Nov 91 13:07:13 CST
From: wmw@tngstar.cray.com (Bill Webster)
Message-Id: <9111131907.AA04396@ansel.cray.com>
To: parein
Subject: The SIB!
Status: OR

Pat,

How the hell are ya?!

Just sent the latest draft version of the SIB to you. Probably won't get to CF u

You don't happen to have a brief VMEbus overview I could steal?

And yes, is seems that Lauren will continue with the YEL after I'm gone. Methink

Keep those cards and letters coming...

Bill

From wmw@tngstar.cray.com Mon Nov 25 10:21:27 1991
Return-Path: <wmw@tngstar.cray.com>
Received: from timbuk.cray.com by techops.cray.com
        id AA18875; 5.61/CRI-5.6b; Mon, 25 Nov 91 10:21:24 -0600
Received: from ansel.cray.com by timbuk.cray.com (4.1/CRI-MX 1.6o)
        id AA17892; Mon, 25 Nov 91 10:21:19 CST
Received: by ansel.cray.com
        id AA06326; 4.1/CRI-4.4; Mon, 25 Nov 91 10:20:51 CST
From: wmw@tngstar.cray.com (Bill Webster)
Message-Id: <9111251620.AA06326@ansel.cray.com>
Subject: Latest ELS admin. notes
To: jawojh@paradise, stahar@els.cray.com, parein, mlm@cherry.cray.com
Date: Mon, 25 Nov 91 10:20:50 CST
X-Mailer: ELM [version 2.3 PL8-CRI]
Status: OR

John, Mark, Melissa and Pat,

Here is a copy of the latest ELS system administration otes I've
assembled. This set includes how to configure multiple IOS' and the
buffer memory resident(BMR) device. If you have any comments or ideas
where I need more details, please give me a call or send mail.

Bill

RT>
RT>load

Loading /IOS/IOS...
Loaded 253632 bytes to location 0x3010000.
RC> call 03010020
IFX Version 1.4
RTSCOPE 68020 v1.24

No FPU Present

==========

```
=================================================================================

Entry Level Systems IOS version 8.2 91/10/07 13:04:19


        SCSI winchester disk: 1 drive(s) found
e_init: pdebug 302e97c, xdebug 3287b0c
        Ethernet Controller: 1 controller(s) found
e_init: exit
        ESDI ctlr 0 unit 0 [E00], type = DD-2
        ESDI ctlr 1 unit 0 [E10], type = DD-2
        ESDI ctlr 2 unit 0 [E20], type = DD-2
        ESDI ctlr 3 unit 0 [E30], type = DD-2

        DAS  Controller 0 [D0]: 1 bank found
        HYPERchannel Controller: 1 controller(s) found

        Size of buffer pool: 3964928 bytes, 3872 kbytes
        Buffers allocated (0 bytes wasted):
         0:  buffer size:       116 bytes,  count: 1
         1:  buffer size:       128 bytes,  count: 16    (reserved)
         2:  buffer size:       256 bytes,  count: 3
         3:  buffer size:      1024 bytes,  count: 3
         4:  buffer size:      4096 bytes,  count: 30
         5:  buffer size:     32768 bytes,  count: 6
         6:  buffer size:     49152 bytes,  count: 2     (reserved)
         7:  buffer size:    131072 bytes,  count: 27

Architecture: XMS
This is the MASTER IOS.

Initializing mainframe...done
Turning on clock...done

IOS>boot

Booting Unicos 6.1

Zeroing memory...
Loading Unicos...

-- Starting Unicos --

==================================================
Switching console to Unicos.
To switch console back to the IOS, enter <CTRL>A
==================================================


Unicos/6.1: sn616-sin.1 (CRAY X-MP)

System gen time     = 09/26/91 13:50:01
UNICOS binary size = 461824 words
Memory Configured  = 16777152 words
Memory Allowed     = 16777152 words
 smem              = 474937 words
 xmem              = 15706624 words
Buffer pool size   = 559104 words (1092 buffers)
User memory avail  = 15543808 words

CPUs configured    = 1, started = 1 (0)
```

```
System-Call Timing On

uts/c1/os/sched.c-11: INFO      packtime calc'd  as 14 seconds.
uts/c1/os/sched.c-07: INFO      gfactor1_in set to 337 seconds.
uts/c1/os/sched.c-08: INFO      gfactor1_out calc'd  as 337 seconds.
uts/c1/os/sched.c-09: INFO      gfactor1_out limited to 0 seconds.
Tue Oct  8 17:02:42 CDT 1991
creating new DD devices
/etc/csaboots: Boot time Tue Oct  8 17:02:46 1991 written to /etc/csainfo

INIT: SINGLE USER MODE

This is a private computer facility.   Access for any reason must be
specifically authorized by the owner.  Unless you are so authorized,
your continued  access and any other use may  expose you to criminal
and/or civil proceedings.

noname# init 2

INIT: New run level: 2
Is the date Tue Oct  8 17:03:03 CDT 1991 correct? (y/n) y
gencat: /dev/dsk/usr Filesystem check bypassed
gencat: /dev/dsk/home Filesystem check bypassed
gencat: complete (1 secs.)
mount: Exec format error
mount: cannot mount /dev/dsk/core
/etc/coredd: /etc/mount /dev/dsk/core /core failed.
/etc/coredd: Manual intervention may be required.
dump partition not initialized
Startup output is being routed to /etc/rc.log.

Do you want to mkfs /tmp (y/n) ? y
Making and mounting file system tmp on /tmp.
Checking file system usr.
/usr: file system opened
/usr: super block fname usr, fpack usr
/usr: clean exit for clean file system
Mounting file system usr on /usr.
Mounting user file systems.
Mounting file system /proc.
Administrative clean up.
Starting system accounting.
Starting system activity collection.
Starting SYS1 daemons: errdemon cron.
Making binary hosts file:  /etc/hosts.bin: 4159 entries written
Configuring TCP/IP network interfaces: lo0e_close: cont=0
        e_init: Controller 0 - Ethernet address=0x02cf1fb00176
 en0.
Installing static routes:
add net default: gateway mh703-1
Starting TCP daemons: inetd.
Setting domain mapping tablesNFS ID mapping is disabled.
Starting NFS daemons:Starting SYS2 daemons:.
.
Executing /etc/rc.pst.
uts/c1/os/sys5.c-03: INFO       gfactor1_in set to 0 seconds.
uts/c1/os/sys5.c-06: INFO       gfactor1_out set to 0 seconds.

Console Login:
noname# shutdown 0

SHUTDOWN PROGRAM

Tue Oct  8 17:20:32 CDT 1991
```

```
Broadcast Message from root (console) on noname Tue Oct  8 17:20:33...
SYSTEM BEING BROUGHT DOWN NOW ! ! !

Qstop shutting down NQS subsystem.
) daemon stopped.
Sending SIGSHUTDN to all running processes.
Error logging stopped.
Sending SIGHUP to all running processes.
spget: security is not enabled!
Sending SIGKILL to all running processes.
Process accounting stopped.
Wait for 'INIT: SINGLE USER MODE' before halting.

INIT: New run level: S

INIT: SINGLE USER MODE

This is a private computer facility.   Access for any reason must be
specifically authorized by the owner.  Unless you are so authorized,
your continued  access and any other use may  expose you to criminal
and/or civil proceedings.

noname# sync
noname# sync
noname# sync
noname# IOS>IOS>mc
IOS>reset

syncing disk...done
Resetting VME..

HK68/V30  ver 01.04
)RTSCOPE 68020 v1.24

 No FPU Present


        SCSI winchester disk: 1 drive(s) found
        SCSI cartridge tape : 1 drive(s) found
        SCSI exabyte tape   : 1 drive(s) found

RT>


        The following is a listing of the /bin/boot script on the IOS disk:
#!
echo
echo Booting Unicos 6.1
echo
echo Zeroing memory...
fm 0 0x1000000 0 0 0 0
if -1 goto :done

echo Loading Unicos...
lu /sys/unicos.tim /sys/cfg.bak
if -1 goto :done

echo
echo -- Starting Unicos --
echo
)start


readswitch autoboot
if 0 goto :done
echo =-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
```

```
echo Switching console to Unicos.
echo To switch console back to the IOS, enter <CTRL>A
echo ===================================================
echo
 hswitch

:done
```

Disk Flawing for ELS systems:

The flaw maps for ELS systems is substantially different than those found in the X-MP and higher-end Y-MP platforms. Most importantly, UNICOS for ELS systems assumes flaw free media for all disk types: IPI-2, ESDI and DAS. Performance of the ELS systems because UNICOS does not have the overhead encvountered while maintaining flaws at the operating system level. The disk comtrollers for IPI-2, ESDI and DAS handle error correction at a lower level than the traditional CRI disk drives.

ESDI flawmaps and flaw commands:
ESDI drives have the factory flaw table written into the last cylinder by the manufacturer. This table is not used during normal disk operations. For ESDI drives each sector header contains flaw information. Each track contains one spare sector that will be used to slip a bad sector. Adjacent to the factory flaw table is a set of spare tracks which are used to map bad tracks.

The IOS command "dverify" reads each sector and determines is the sector is bad or good. If the sector is bad, dverify first attempts to slip the sector using the spare sector found with each track. If dverify cannot slip the sector it is then mapped out to a spare track from the trscks adjacent to the factory flaw table. All bad blocks found by dverify are unused by UNICOS. The IOS command "dformat" formats the entire drive (ESDI, IPI-2 and DAS types) using known paramters. Be careful when using dformat as all data will be lost.

Note that with IOS 8.2 the dformat command employs sector spiraling. This technique improves I/O throughput between 7 and 11 percent over the technique used in earlier versions of the IOS.

UNICOS file systems may reside on sectors 0 to 170100 on ESDI disk drives. Four disks are allowed per controller, and four controllers are permitted per IOS (total of 16 ESDI disks).

DAS flawmaps and flaw commands:
The Disk Array Subsystem is a set of 8 ESDI data drives, 1 parity and 1 stand-by drive controlled by a single DAS controller. The DAS controller uses the first 2 logical cylinders as system tables. Two tables are maintained: the raw flaw table (RFT) anf the growth error table (GET). The RFT resides on the last cylinder of the disk array; the GET is found in the first two cylinders of the disk array. These cylinders are not available or accessible by UNICOS.

The commands used for DAS flawing are:

ddisable	to disable defective drives in array
dflawr	reads the DAS flaw table
dflaww	writes the DAS flaw table
dformat	to format the disk array
dreplace	reconstructs data onto new DAS member

dwconfig	writes DAS configuration to controller
dverify	verifies disk array media
dinfo	gives disk array drive status

These commands are fully explained in your "CRAY XMS Systems IOS Reference Manual".

The dverify command, when run on a disk array (DAS) does a read-only sweep of the DAS media. Any detected errors are added to the growth error table (GET). Note that the errored blocks are still accessible until the dformat command id run on the disk array. The ddisable comamnd disbales a bad DAS member and will query for taking action to substitute the failed device with stand-by disk drive. The data from the failed drive is rebuilt onto the new DAS member from the remaining 7 data drives and the single parity drive. The dreplace command within the IOS software performs this rebuild.

The dformat command merges the RFT and GET tables and will attempt to move the data from bad sectors into good sectors. The dflawr reads the RFT and writes an ASCII copy to the IOS SCSI Winchester disk drive. The dflaww command reads the ASCII RFT file from the IOS SCSI disk and writes it bck to the RFT. The dinfo command gives DAS configuration and the current state of each DAS drive.

UNICOS file partitions can reside from sectors 0000000 through 1269114 on disk array subsystems. Sectors 1269114 through 1270051 are reserved for diagnostics. The drives of a disk array are formatedd into 512-byte blocks, one logical sector spans all eight (data) drives for a total of 4096 bytes.

Disk types on the YMP-EL
==========

```
DDESDI   /* old esdi drive */
DD3      /* new esdi drive */
DDLDAS   /* old DAS */
DDAS2    /* new DAS */
DD4      /* ipi-sabre 7 */
RD-1     /* removable esdi */
RD-2     /* removable SCSI */
DDBMR    /* buffer memory resident */
DDRAM    /* central memory resident */
```

General Performance Implications
---

1. Swap should not be configured on the same drive(s) with user home directories. Large process swaps could degrade interactive response.

2. If your typical site workload performs heavy of I/O, use a banded /tmp across esdi devices, and run batch jobs on the /tmp filesystem.

3. Configure /root and /usr on different controllers if possible to balance requests across controllers.

Use DAS for swapping and important single I/O stream applications that use large I/O sizes. DAS performance is optimal on large block transfers. Any remaining space on DAS can be filled with seldom used filesystems such as a backup root(bkroot), backup usr(bkusr), or extra tmp space.

5. Rotate slices of file systems (bands) across controllers & drives.

6. Split the /usr file system into a series of seperate, smaller file systems like /usr/adm, /usr/spool and /usr/dm. The /usr/spool file system is also a good candidate for banding.

7. If you have a BMR device, use it for swap device.

8. Stripe your swap device.

9. Locate your backup root file system on the same device as swap and /usr/spool.

Misc:

A maximum of 16 IPI drives are supported on the YMP-EL.

Each sector on CRAY ELS systems is 4096 bytes.

UNICOS imposes a limit of 15 disks per striped group.

The drives on the DAS are formatted in 512 bytes, one logical sector is spread across all eight (data) drives, 4096 bytes.

Configuring the IOS for XMS and Y-MP EL systems

```
#
#               IOS configuration file
#
#               Define architecture (XMS or YMP-EL)
#
ARCH=XMS
#
#       Define which slots in the system cabinet have cpu's (0,1,2,3)
#
#       We have a single XMS CPU in slot #0. If we had a multiple CPU
#       system with cpu's in slots 0 and 2 our entry would look like:
#       NCPUS=0,2
NCPUS=0
#
#
#       The master IOS is always IOS0. The 1274 value is the serial
#       number from the Heurikon 68030 CPU on the board. The serial
#       number is stored in NVR on the Heurikon board.
#
IOS0:   1274
#
#       Strategies are modules that deal with packets, communicate with
#       UNICOS and handle memory management with the IOBB board.
#
#       Strategy name
#       -----------
        /dev/disk
        /dev/console
        /dev/tape
        /dev/ethnet
#
#       Only the first (master) IOS has the /dev/console strategy.
#
#       Device drivers are loaded AFTER strategies and are device
#       specific.
#
#       Device name
```

```
#       -----------
        /dev/ct
        /dev/exb
        /dev/esdi
        /dev/das
        /dev/ether
#
#       A second IOS block of strategies and drivers follows for a second
#       IOS. We begin with the serial number of the IOS CPU.
#
IOS1:   1275
#
#       We begin with the serial number of the IOS CPU.
#
#       Strategy name
#       -------------
        /dev/disk
        /dev/tape
#
#
#       Device name
#       -----------
        /dev/ct
        /dev/exb
        /dev/ipi
#
#
# The possible device drivers are:
#
# /dev/bmr (BMR)
# /dev/ct (cartridge tape)
# /dev/exb (Exabyte tape)
# /dev/esdi (ESDI disk driver)
# /dev/das (array of ESDI disks)
# /dev/ether (Ethernet driver)
# /dev/ipi (IPI-2 disk driver)
# /dev/bm (block mux)
# /dev/telex (STK or Telex 9-track tape)
#
# end of IOS configuration file (called /config on IOS disk)
#

        CSL for EL systems
        ------------------
NOTE: EL disks are all sliced on disk block boundaries, NOT cylinder.
MIOP values:
        0-7 ESDI
        8   DAS
        9   BMR
        10-15 IPI

Unit values:
        0-3 ESDI
        0-1 DAS
        1   BMR (one BMR device per IOS)
        0-3 IPI

        Sample CSL definitions
        ----------------------

NFIG {

        AUTOGEN OFF;
        NLDCH = 778; /* for BMR device */

        /* physical devices */
```

```
/* Sample entry for esdi controller 0, drive 0 */

ESDI_00 := DDESDI MIOP 0 UNIT=0 (
        1228:iroot_e00          /*  12000 blocks */
        13228:tmp_e00           /* 100000 blocks */
        113228:home_e00         /*  56872 blocks */
);

/* Sample entry for esdi controller 0, drive 1 */

ESDI_01 := DDESDI MIOP 0 UNIT=1 (
        0:root_e01              /*  75000 blocks */
        75000:bkusr_e01         /*  95100 blocks */
);

/* Sample entry for esdi controller 1, drive 1 on second IOS */

ESDI_01 := DDESDI MIOP 1 UNIT=1  IOS = 1 (
        0[75000]:root_e01       /*  75000 blocks */
        75000:bkusr_e01         /*  95100 blocks */
);

/* Sample entry for esdi controller 1, drive 0 */

ESDI_10 := DDESDI MIOP 1 UNIT=0 (
        0:usr_e10               /*  95100 blocks */
        95100:bkroot_e10        /*  75000 blocks */
);

/* Sample entry for esdi controller 2, drive 0 */

ESDI_20 := DDESDI MIOP 2 UNIT=0 (
        0:tmp_e20               /* 100000 blocks */
        100000:home_e20         /*  70100 blocks */
);

/* Sample entry for esdi controller 3, drive 0 */

ESDI_30 := DDESDI MIOP 3 UNIT=0 (
        0:tmp_e30               /* 100000 blocks */
        100000:home_e30         /*  70100 blocks */
);

/* Sample entry for esdi controller 3, drive 0 */

ESDI_31 := DDESDI MIOP 3 UNIT=1 (
        0:tmp_e31               /* 100000 blocks */
        100000:home_e31         /*  70100 blocks */
);

/* Sample entry for DAS */

LDAS_00 := DDLDAS MIOP 8 UNIT=0 (
        0:dump_d00              /*  65536 blocks */
        65536:swap_d00          /* 163840 blocks */
        229376:dastmp_d00       /* 400000 blocks */
        629376:core_d00         /* 100000 blocks */
        729376:app_d00          /* 539738 blocks */
);

/* logical devices */

iroot := (              /* 12000 blocks */
        iroot_e00
);
root := (               /* 75000 blocks */
```

```
            root_e01
        );
        bkusr := (                  /* 95100 blocks */
            bkusr_e01
        );
        usr := (                    /* 90100 blocks */
            usr_e10
        );
        bkroot := (                 /* 75000 blocks */
            bkroot_e10
        );

        home := (                   /* 268400 blocks */
            home_e00
            home_e20
            home_e30
            home_e31
        );
        tmp := (                    /* 400000 blocks */
            tmp_e00
            tmp_e20
            tmp_e30
            tmp_e31
        );
        dump := (                   /*  40000 blocks */
            dump_d00
        );
        swap := (                   /* 163840 blocks */
            swap_d00
        );
        dastmp := (                 /* 400000 blocks */
            dastmp_d00
        );
        core := (                   /*  80100 blocks */
            core_d00
        );
        app := (                    /*  539738 blocks */
            app_d00
        );

        /* system devices */

        ROOTDEV = root;
        PIPEDEV = root;
        SWAPDEV = swap;
}
```

Configuring a BMR device for XMS and Y-MP EL systems:

1) In the IOS file /config, make sure the driver for BMR is activated. (i.e. uncomment /dev/bm)

2) In the UNICOS configuration(CSL) file, set NLDCH to 778.

3) In the UNICOS configuration(CSL) file, divide BMR into 3 parts.
   - Part 1 for interactive_nonhog SWAP.
   - Part 2 for interactive_hog + batch_nonhog SWAP.
   - Part 3 for LDCACHE of /root and /usr.

4) In the UNICOS configuration(CSL) file, add the BMR swap partitions to logical swap and add BMRDEV as a system device. Examples for all of these CONFIGURATION changes follow:

EXAMPLE CONFIGURATION CHANGES for BMR

```
CONFIG {
```

```
        AUTOGEN OFF;
        NLDCH = 778;

        /* physical devices */

        BMR := DDBMR MIOP 9 SIZE=32768  (
                0:bmr_part1             /*   2000 blocks */
                2000:bmr_part2          /*  14430 blocks */
                16430:bmr_part3         /*  16338 blocks */
        );

        /* logical devices */

        swap := (             /* BMR 16430 + X disk blocks */
                bmr_part1
                bmr_part2
                swap_xxx      /* disk portion, at least 100000 blocks */
        );
        bmr := (                 /*  16338 blocks */
                bmr_part3
        );

        /* system devices */

        SWAPDEV = swap;
        BMRDEV = bmr;
}
```

5) Cache root and usr filesystems. The commands to do this are:

```
    ldcache -l /dev/dsk/root -n 475 -s 21
    ldcache -l /dev/dsk/usr -n 303 -s 21
```

You can do this automatically in /etc/config/ldchlist.
for example:

```
unicos# cat /etc/config/ldchlist
#
#   SN616 - ldchlist - Edition 2  [Wed Apr  3 11:18:27 CST 1991]
#   Created by Configuration Generator Rev. 6.62
#
#
# Logical device cache table - used by /etc/rc
#
#       There are four fields per line, separated by white space.
#       1.  logical device
#       2.  cache type SSD|BMR|MEM
#       3.  number of cache units
#       4.  size in 4k blocks of each unit (usually a track size)
#
#
/dev/dsk/root61                 BMR     475     21
/dev/dsk/usr61          BMR     303     21
```

Sample ldcache (1m) output:

```
Mon Nov 11 11:16:08 EST 1991

T unit size   reads     writes       hits     misses    rate   name
- ---- ----  --------- ----------  ---------- -------  -----  --------

B  475   21      719784     118946      221211      985  99.56  /dev/dsk/root61
B  303   21      391674      16011       53028      747  98.61  /dev/dsk/usr61

   Cache to user     Cache to disk      Cache/disk ratio
```

Reads	Writes	Reads	Writes	Read	Write	Total	Name
714675	117270	15603	39339	45.8	3.0	15.1	/dev/dsk/root61
391643	15931	15015	9707	26.1	1.6	16.5	/dev/dsk/usr61

William Webster wmw@cray.com phone:(612)683.3831
Cray Research, Inc., 655A Lone Oak Drive, Eagan, MN 55121
(Alan Kay, Apple) The best way to predict the future is to invent it.

9 MWS ADMINISTRATION

To be provided.

10 FIELD REPAIR PROCEDURES

To be provided.

Cray Research, Inc.
Hardware Publications and Training
770 Industrial Boulevard
Chippewa Falls, WI 54729