Memory Module (CM04)

HTM-320-A
CRAY T90 Series Systems
Last Modified: November 1997

Figures

Tables

Record of Revision

March 1997

Original printing.

November 1997

Revision A removes references to asynchronous SRAM chips and modifies Figure 2 to include TZ0, Figure 6 to show CPUs 0-3 in the B1 square, and Figure 9 to show bank numbers.

CM04 Module Description

Cray Research designed the CM04 memory module to replace the CM03 memory module in a CRAY T90 series system. The performance of the CM04 varies significantly from that of the CM03. The CM04 has half the bandwidth of the CM03.

The CM04 module has 2 sections that each hold 16 banks of memory. Each section of memory can be accessed by four CPUs, and the memory array in each section can provide about 2.25 words per clock period. Thus, on average, each CPU can receive about 0.56 memory words per clock period.

Like the CM03 memory module, a write reference requires a 2-packet transfer, but a read reference requires only 1 packet.

Error detection and correction is done on the CPU module.

Capacity

The CM04 module has twice the capacity of the CM03 module. The CM04 has 2 sections. Each section holds 16 banks. Each bank holds 2 million words, for a total of 64 million words.

The memory chip is a 4M SRAM that is used as 2 million locations by 2 bits. The chips are assembled in four 20 memory chip stacks, for a total of 80 chips per replaceable stack. Previous versions of the chip stack held 40 chips. Each 20-chip quarter-stack comprises the lower or upper data bits of one bank, half of the check bits for that bank, and a spare chip.

Memory Organization

Central memory is organized into sections, subsections, and banks. Remember, however, that a CRAY T94 system does not have subsections. All CRAY T90 series computers normally have 8 sections of memory. Memory may be degraded with the SCE program by logically eliminating failing sections, subsections, and banks until time is scheduled to replace the failing module.

Section

A memory section is the range of components that can be the destination of a request from a CPU through a single path. The components that compose the section are spread out over the eight memory modules that compose the module stack in CRAY T916 and CRAY T932 systems. The components that compose the section in a CRAY T94 system are all located on one module. In fact, each memory module in a CRAY T94 system has 2 full sections on it.

In a CRAY T94 system, the CPU uses the section bits of the address to steer the memory reference from a particular connector to the appropriate memory module. In CRAY T916 and CRAY T932 systems, the section bits are used to steer the reference to the appropriate network module.

The section is made up of 8 subsections in all models except the CRAY T94 system, which has no subsections.

Subsection

There is a separate path from the network module to each subsection. The network module decodes the subsection bits and steers the memory reference to the appropriate memory module after any subsection conflicts that may exist have been resolved. Each subsection contains 8 or 16 banks.

A bank contains 2 million 76-bit words (64 data and 12 check bits). A bank always resides on one module.

Memory Configurations

The three types of mainframes (CRAY T932, CRAY T916, and CRAY T94) normally have 8 sections of memory.

- A CRAY T94 system includes only one memory module stack; the module stack contains either two or four memory modules. Refer to page 14 (“CRAY T94 Memory”) for a more detailed explanation.
- A CRAY T916 system includes two memory module stacks; each module stack contains either four or eight memory modules. Refer to page 19 (“CRAY T916 Memory”) for a more detailed explanation.
- A CRAY T932 system can include two or four memory module stacks; each module stack contains either two, four, or eight memory modules. Refer to page 24 (“CRAY T932 Memory”) for a more detailed explanation.

Table 1 shows the various memory configurations of the three types of systems. The numbers in this table do not reflect partitioned or degraded memory.

Table 1. Memory Configurations by System Type

Model	CRAY T94	CRAY T916	CRAY T932
Sections	8 or 4	8 or 4	8 or 4
Subsections/section	1	8 or 4	8 or 4 or 2
Banks/subsection	16	8	16
Total banks/system	128 or 64	512 or 256	1,024 or 512 or 256

Memory Partitioning and Degradation

When you degrade memory, you are bypassing areas in memory where failures occur so that the system can continue to operate. By degrading memory, you are forcing selected section, subsection, and bank bits. Therefore, the memory address referenced by the CPU is different from the address received by memory. For specifics and examples, refer to “Memory Degradation” in the *SCE User Guide*, publication number HDM-069-C.

Software address mapping through the maintenance channel controls memory configuration. The System Configuration Environment (SCE) provides the interface for selecting sections, subsections, banks, and groups in a particular configuration to control memory degrades and/or logical memory partitions.

Memory Address Mapping

Memory addressing depends on the configuration of the CRAY T90 series system. The CP, network, and memory modules determine how memory addressing occurs. Table 2 describes the address map that the CM04 module uses in the CRAY T90 series CPU. (The CPU is capable of addressing up to bit 34 for accessing up to 16 subsections, but the hardware currently uses only bits 0 through 33, with 8 subsections.) This address map accommodates future CRAY T90 series systems and is capable of addressing up to 16 Gwords of memory. The “CRAY T94 Memory,” “CRAY T916 Memory,” and “CRAY T932 Memory” subsections include address maps for the specific CRAY T90 series systems.

Table 2. Address Map

Address Bits	Function
0 – 2	Section select
3 – 5	Subsection select
6 – 9	Bank select
10 – 33	Word select [a]

[a] SCE uses bits 10 and 11 as group profile and select bits 0 and 1 when partitioning memory.

Module Components

The CM04 memory module contains the central memory that is common to all processors in a CRAY T90 series mainframe. The module consists of a printed circuit board (PCB) with memory stacks and logic options on board 1 and logic options on board 2. Actually, the PCB is composed of two PCBs laminated together. The term *board 1* refers to one side of the PCB, and *board 2* refers to the other side of the PCB. Memory stacks contain the memory chips that store the data. Refer to Figure 1 and Figure 2 for drawings of the boards.

Memory modules are arranged in stacks within the mainframe. This is referred to as a **memory module stack**. Do not confuse this with a **memory stack**, which is a stack of memory chips that is mounted on board 1 of the memory module.

Logic Options and Connectors

There are five types of logic options on the CM04 module:

- 16 MK options (8 per section), which interface with the network or CPU modules
- 16 MJ options (8 per section), which steer the address, data and control to one of four 4-bank groups
- 48 MI options (24 per section), which steer address and data to one of four banks in the group
- 1 MZ option for maintenance functions
- 1 TZ option for clock fanout

There are three types of logic connectors:

- 8 ZA connectors, between the network or CPU modules
- 8 ZC connectors, between the memory stacks in one of the two sections on the module
- 8 ZD connectors, between the memory stacks in the other section on the module

NOTE: ZC and ZD connectors are identical. The reason for using different names is for ease in distinguishing sections.

Figure 1. CM04 Memory Module - Board 1

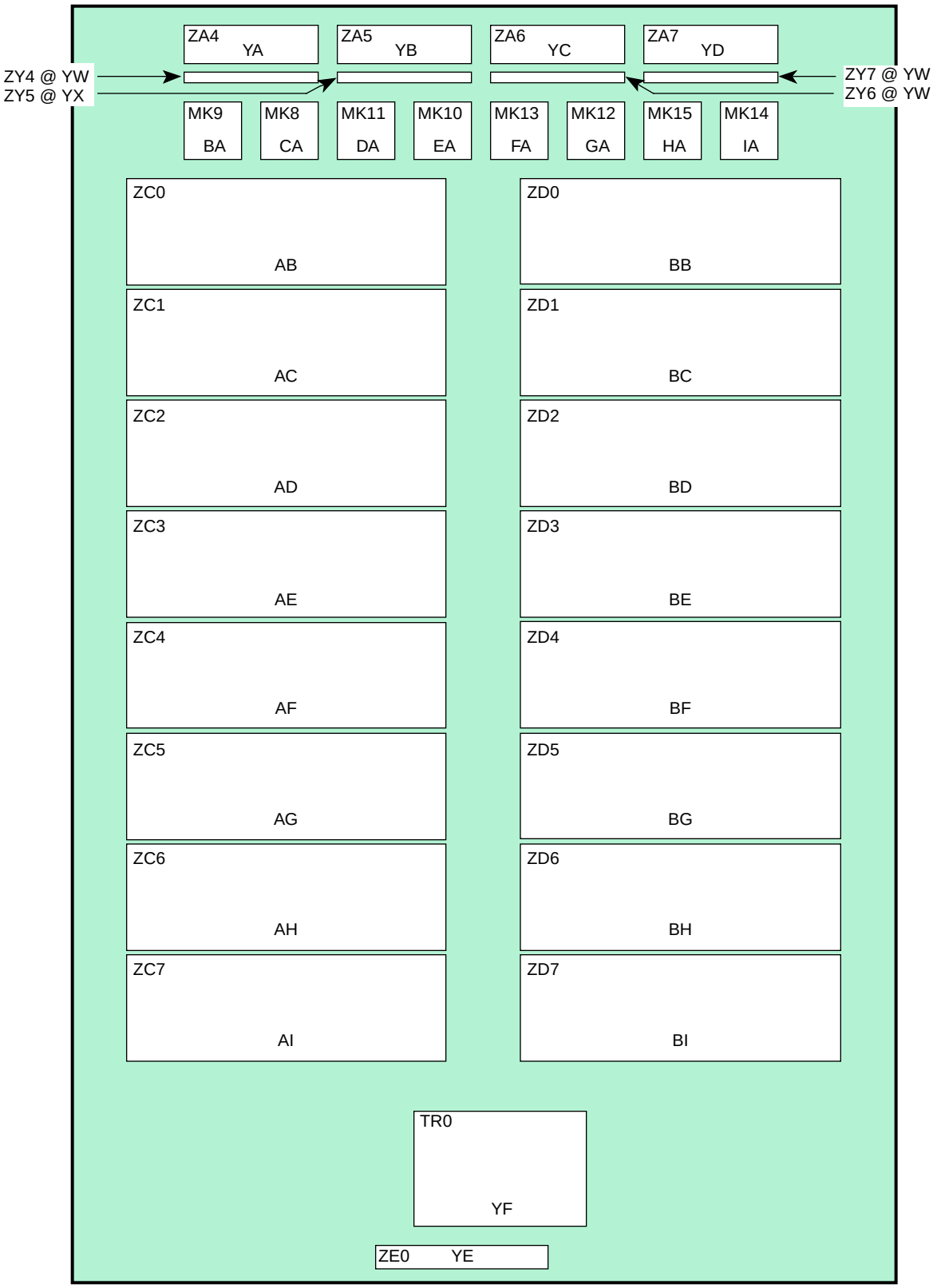
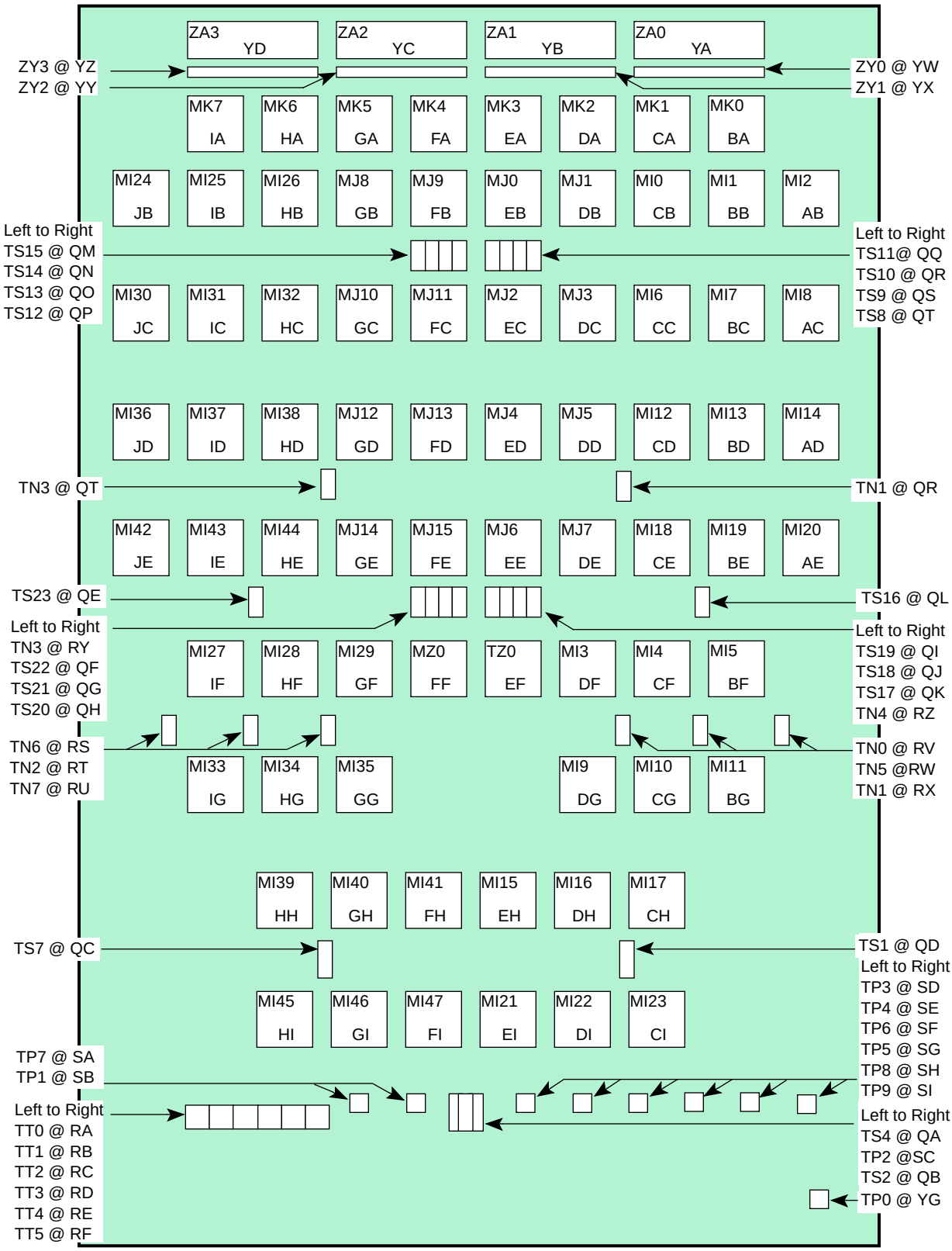


Figure 2. CM04 Memory Module - Board 2



Memory Stacks

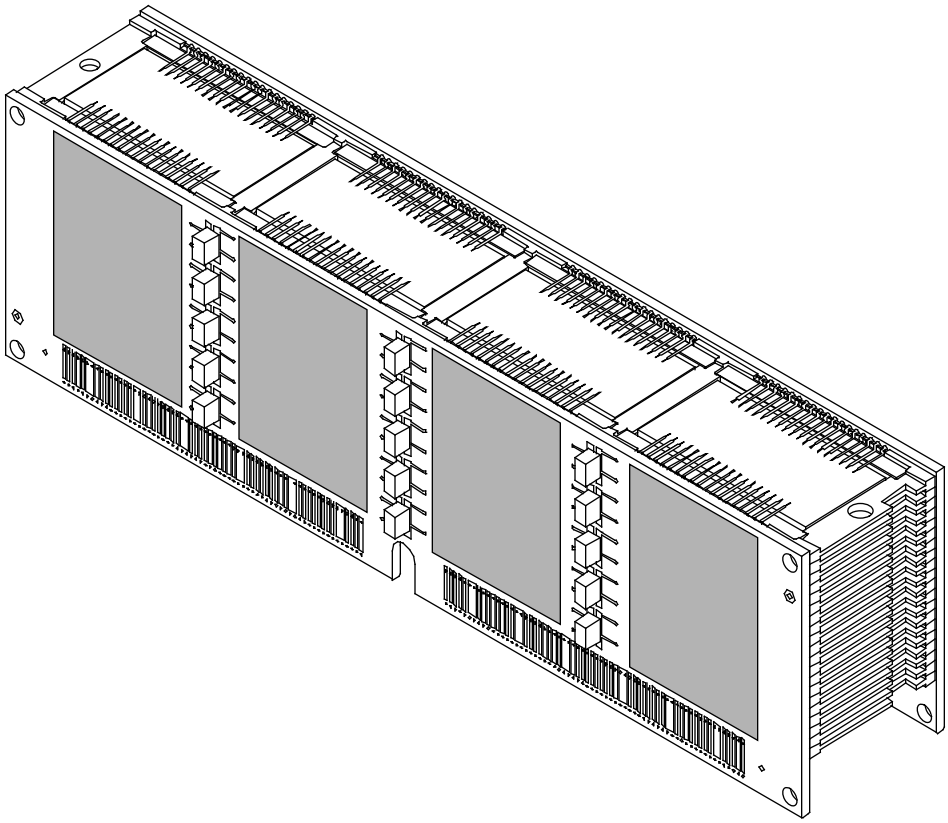
There are 16 memory stacks in each memory module. Each memory stack has 4 quarter-stacks. These quarter-stacks hold either the upper or lower 32 data bits, plus 6 check bits of 1 bank. Each quarter of the memory stack has 19 memory chips for data and check bits, and 1 spare chip for that quarter, for a total of 20 memory chips. Each memory chip has 4 million memory cells. The memory chip is used as a 2 meg by 2 bit part. Refer to Table 3 for the bank layout of a stack.

The bits for a bank are always spread out over 2 memory stacks. Therefore, it takes 2 memory stacks to make up 4 complete banks. Refer to Figure 3 for an illustration of a CM04 memory stack.

Table 3. Memory Stack Bit Layout

Bank N+14	Bank N+10	Bank N+4	Bank N
2 Bits	2 Bits	2 Bits	2 Bits
2 Bits	2 Bits	2 Bits	2 Bits
2 Bits	2 Bits	2 Bits	2 Bits
2 Bits	2 Bits	2 Bits	2 Bits
2 Bits	2 Bits	2 Bits	2 Bits
2 Bits	2 Bits	2 Bits	2 Bits
2 Bits	2 Bits	2 Bits	2 Bits
2 Bits	2 Bits	2 Bits	2 Bits
2 Bits	2 Bits	2 Bits	2 Bits
2 Bits	2 Bits	2 Bits	2 Bits
2 Bits	2 Bits	2 Bits	2 Bits
2 Bits	2 Bits	2 Bits	2 Bits
2 Bits	2 Bits	2 Bits	2 Bits
2 Bits	2 Bits	2 Bits	2 Bits
2 Bits	2 Bits	2 Bits	2 Bits
2 Bits	2 Bits	2 Bits	2 Bits
2 Check Bits	2 Check Bits	2 Check Bits	2 Check Bits
2 Check Bits	2 Check Bits	2 Check Bits	2 Check Bits
2 Check Bits	2 Check Bits	2 Check Bits	2 Check Bits
2 Spare Bits	2 Spare Bits	2 Spare Bits	2 Spare Bits

Figure 3. CRAY T90 Series Memory Stack (ST04)



Spare Chips

There are four spare memory chips in each memory stack, one for each quarter of the stack. A bad memory chip can be flawed out, and the data in the remainder of the stack can be shifted down to the next chip, with the last chip shifting its bits to the spare chip for that quarter-stack. The flawing is accomplished with the SCE maintenance program. For a detailed explanation of spare chips and the flawing of memory chips, refer to “Chip Flawing” on page 28 of this document.

CRAY T94 Memory

CRAY T94 memory is organized by sections and banks. A fully configured system has 8 sections of memory. Each section contains 16 banks for a total of 128 banks. Figure 4 shows the memory organization in a CRAY T94 system.

Figure 4. CRAY T94 Memory Organization

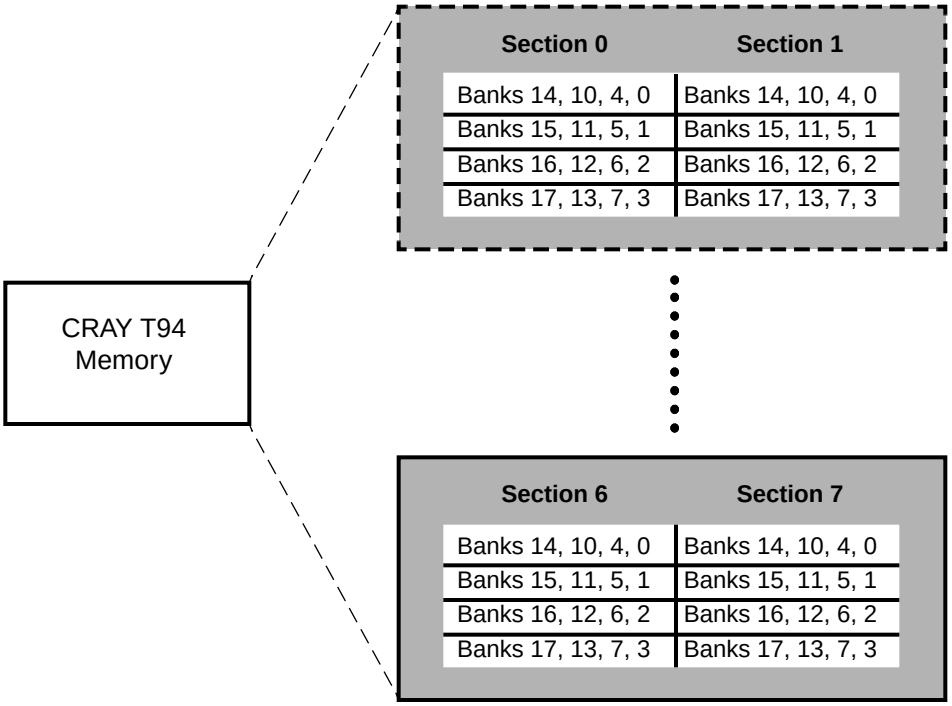


Figure 5 shows the addressing map for a fully configured CRAY T94 system. It shows the module types that determine the section select, bank select, and word select. Bits 28 through 31 are presently not used; however, these bits may be used in future systems to address up to 4 Gwords of memory.

Figure 5. CRAY T94 Memory Addressing (No Partitioning or Degradation)

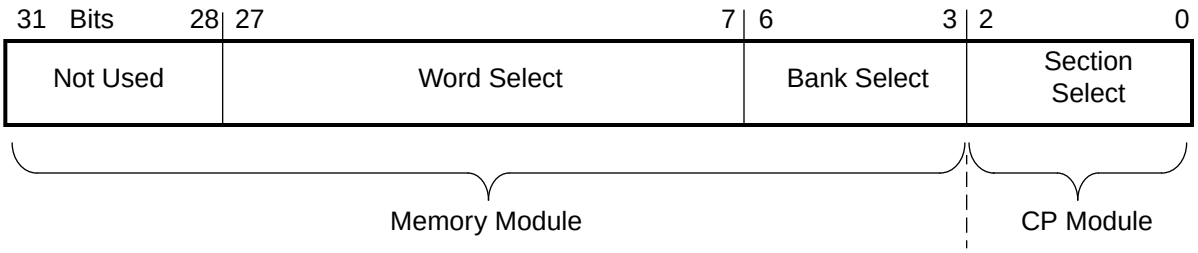


Table 4 lists the available CRAY T94 configurations. A CRAY T94 system has one memory module stack that contains either two or four memory modules. The memory modules connect directly to the CP modules.

Table 4. CRAY T94 System Configuration

Module Counts			Configuration			
Processor	Network	Memory	Sections	Subsects	Banks	MWords
1 to 4	0	4	8	1	128	256
1 to 4	0	4	8	1	64	128
1 to 4	0	2	4	1	64	128
1 to 4	0	2	4	1	32	64

Figure 6 shows the interconnections between the CP modules and memory modules.

Assuming the first reference is to section 0 and the memory references are sequential, the first two references go to sections 0 and 1 (bank 0) on the memory module at location C2, the next two references go to sections 2 and 3 on the memory module at location C4, and so on, until each module in the stack receives two references. References nine through sixteen follow the same pattern but address bank 1; references 17 through 24 address bank 2, and so on, until all banks (128 in a fully configured system) have been referenced.

Figure 6 also shows the section layout in a CRAY T94 system. Each memory module in a CRAY T94 system comprises 2 sections of memory.

Figure 6. CRAY T94 CPU-to-memory Interconnections

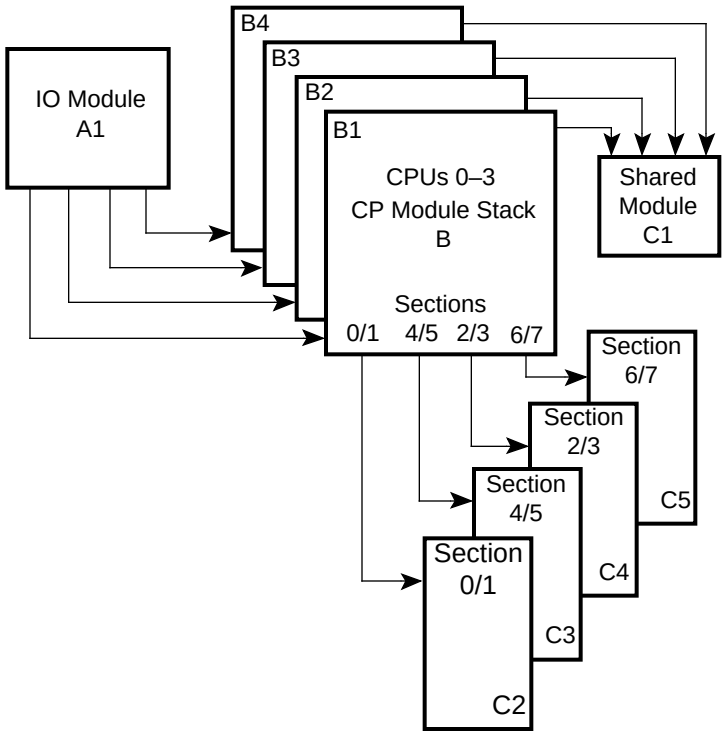


Figure 7 shows the stack layout on the memory module and the bit layout in the memory stack for the CRAY T94 system.

Figure 7. CRAY T94 Memory Module Bank and Bit Layout

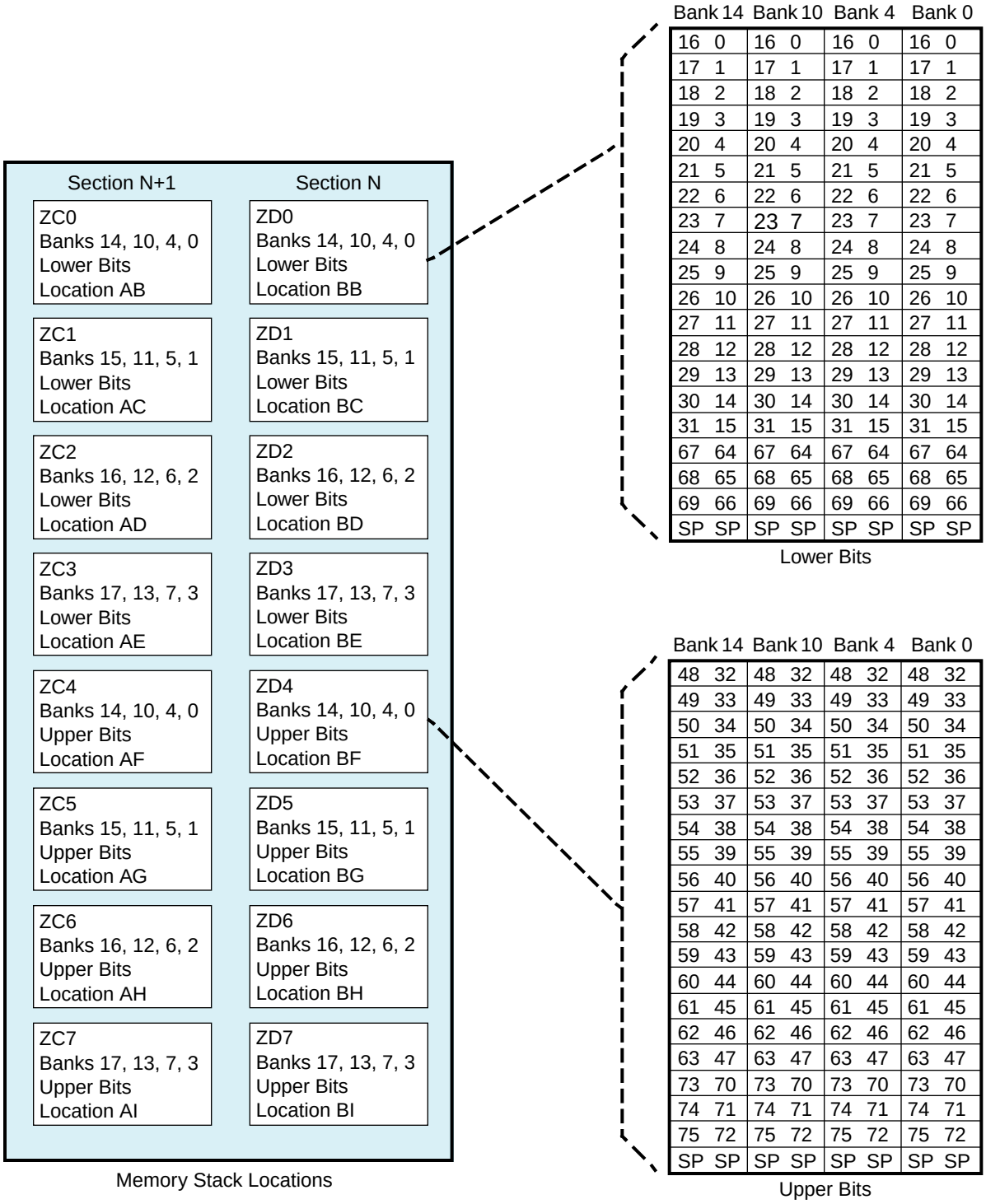
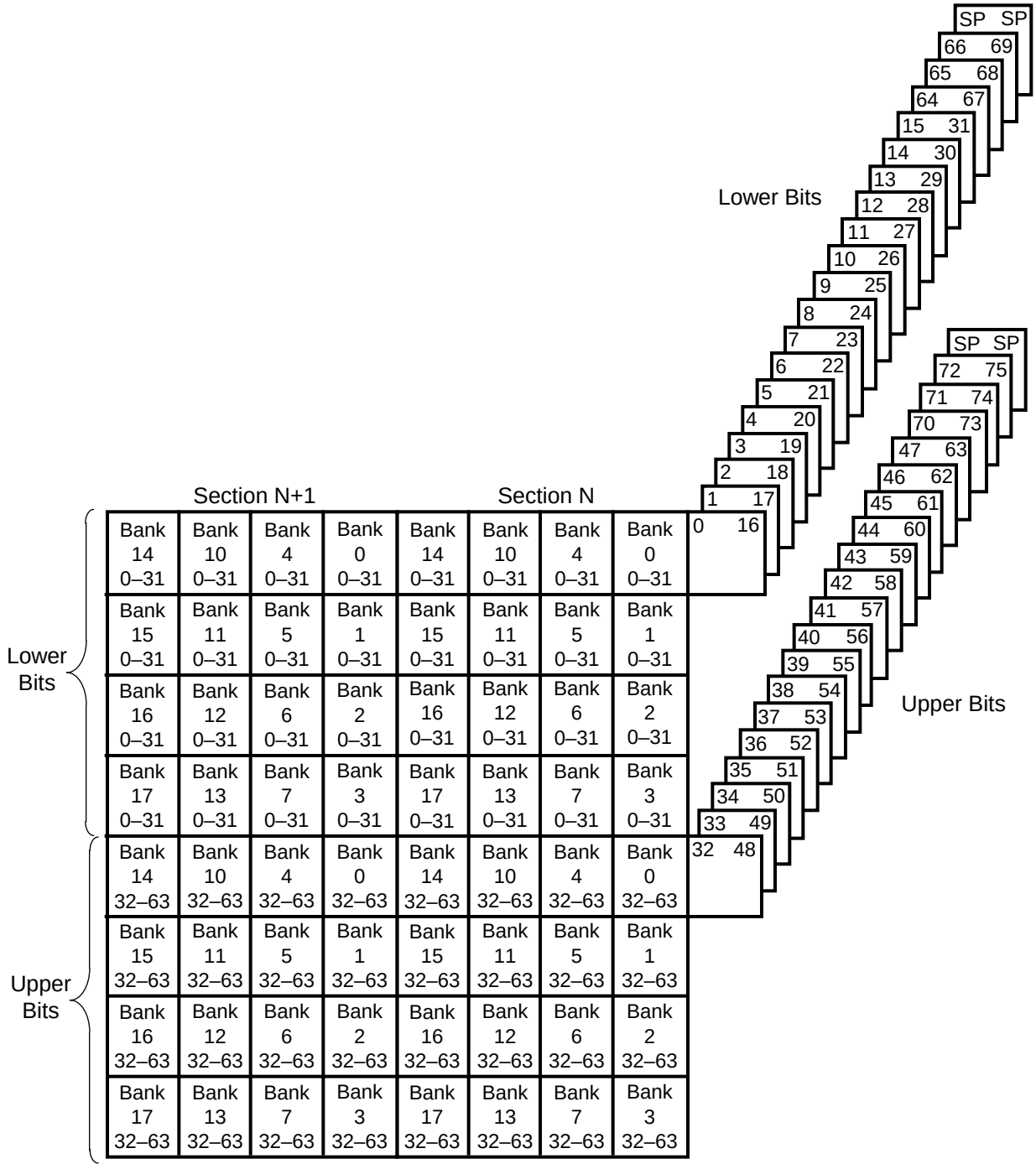


Figure 8 further illustrates the chip and bit layout of the CM04 memory module.

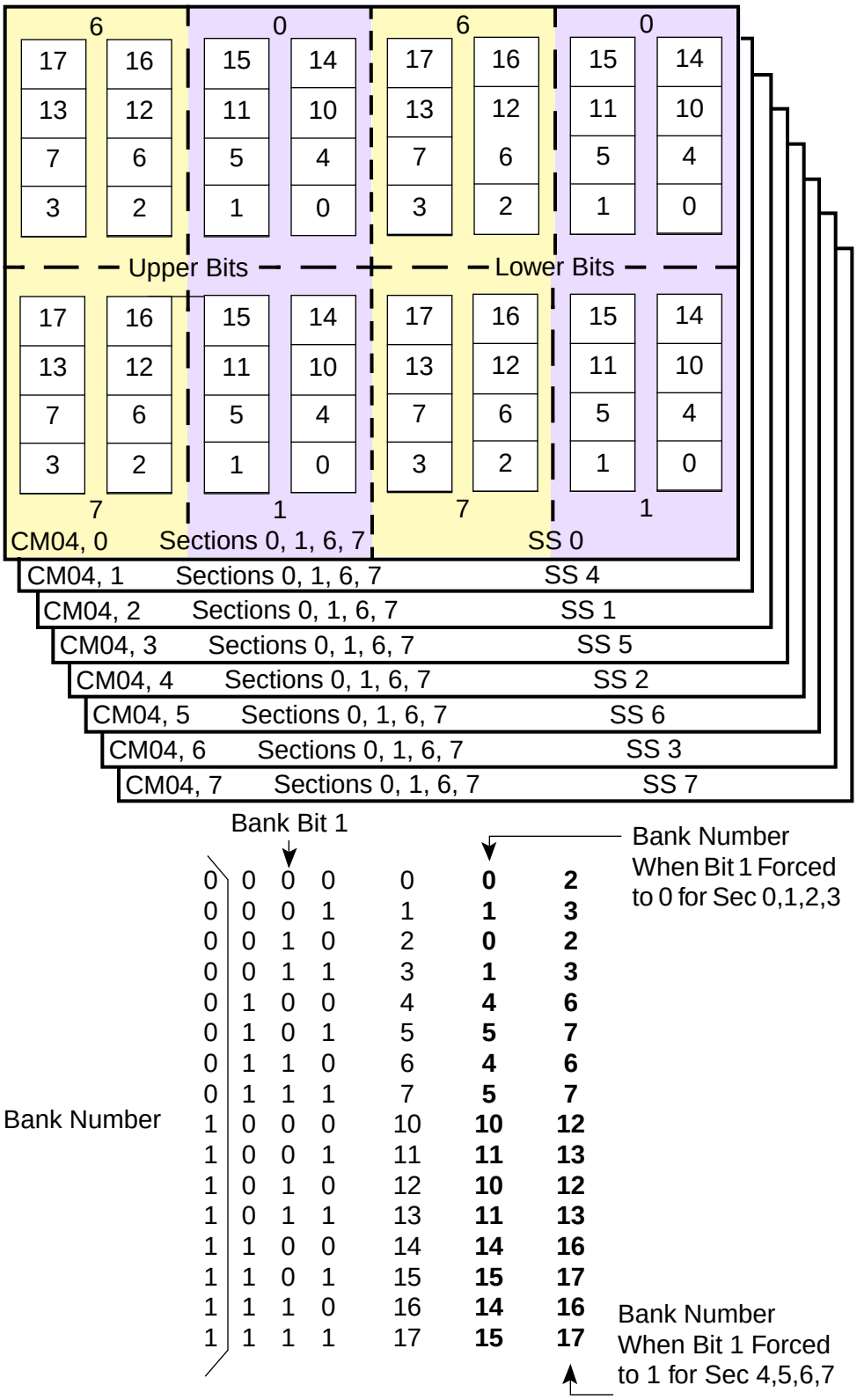
Figure 8. Basic Stack and Bit Layout for CM04



CRAY T916 Memory

The memory in a CRAY T916 system is organized by sections, subsections, and banks. A fully configured system has 8 sections of memory. Each section has 8 subsections and each subsection has 8 banks for a total of 512 banks. Figure 9 shows the memory organization in a CRAY T916 system. Upper and lower bits are in the same positions in all configurations of CRAY T90 series systems; however, the CM04 is partitioned into four sections in the CRAY T916 system. Each memory module stack contains four sections on each module of that stack. Each module is a separate subsection. Because the CRAY T916 system incorporates only 8 banks per subsection, bank bit 1 is manipulated to force access to restricted banks. Figure 9 illustrates one of the two memory module stacks in a CRAY T916 system, sections 0, 1, 6, and 7. Sections 2, 3, 4, and 5, which are located on the other stack but not shown, can be substituted in their places.

Figure 9. CRAY T916 Memory Organization



The CP and memory modules determine which section, bank, and word to address. Figure 10 shows the addressing map for a fully configured CRAY T916 system. It also shows the module type that determines the section select, subsection select, bank select, and word select. Bits 30 through 33 are presently not used; however, these bits may be used in future systems to address up to 16 Gwords of memory.

Figure 10. CRAY T916 Addressing (No Partitioning or Degradation)

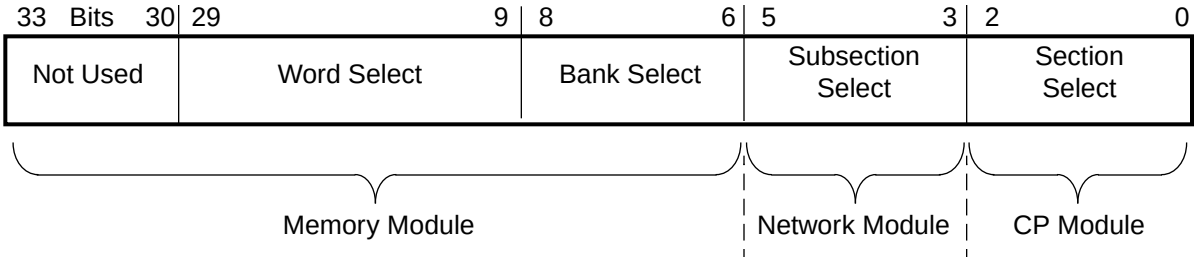


Table 5 lists the various configurations of CRAY T916 systems. A CRAY T916 mainframe has two memory module stacks that consist of four or eight memory modules each. The memory modules connect to the network modules, which connect to the CP modules. Connections between network and CP modules are actually made through the system interconnect board (SIB).

Table 5. CRAY T916 System Configuration

Module Counts			Configuration			
Processor	Network	Memory	Sections	Subsects	Banks	MWords
4 to 8	4	8	8	4	256	512
4 to 8 [a]	4	8	8	2	128	256
4 to 8 [a]	2	8	4	4	128	256
4 to 8 [a]	2	4	4	2	64	128
4 to 8	4	16	8	8	512	1,024
4 to 8 [a]	4	8	8	4	256	512
4 to 8 [a]	2	8	4	8	256	512
4 to 8 [a]	2	4	4	4	128	256
8 to 16	8	16	8	8	512	1,024
8 to 16 [a]	8	8	8	4	256	512
8 to 16 [a]	4	8	4	8	256	512
8 to 16 [a]	4	8	4	4	128	256

[a] This indicates a degraded or partitioned system; the system is not usually sold with this configuration.

Figure 11 shows the logical interconnections between the CP modules and memory modules in a CRAY T916 system. Assuming the first reference is to section 0 and the memory references are sequential, reference one goes to section 0, subsection 0 at module location L1. Reference two goes to the same module but to section 1. References three through six go to sections 2, 3, 4, and 5 and subsection 0 on the memory module at location H1. References seven and eight go to sections 6 and 7, subsection 0 on the memory module at location L1. The next eight references follow the same sequence for subsection 1 but reference the modules at locations L3 and H3. This referencing pattern continues until all banks have been addressed.

Figure 11. CRAY T916 CPU-to-memory Interconnection

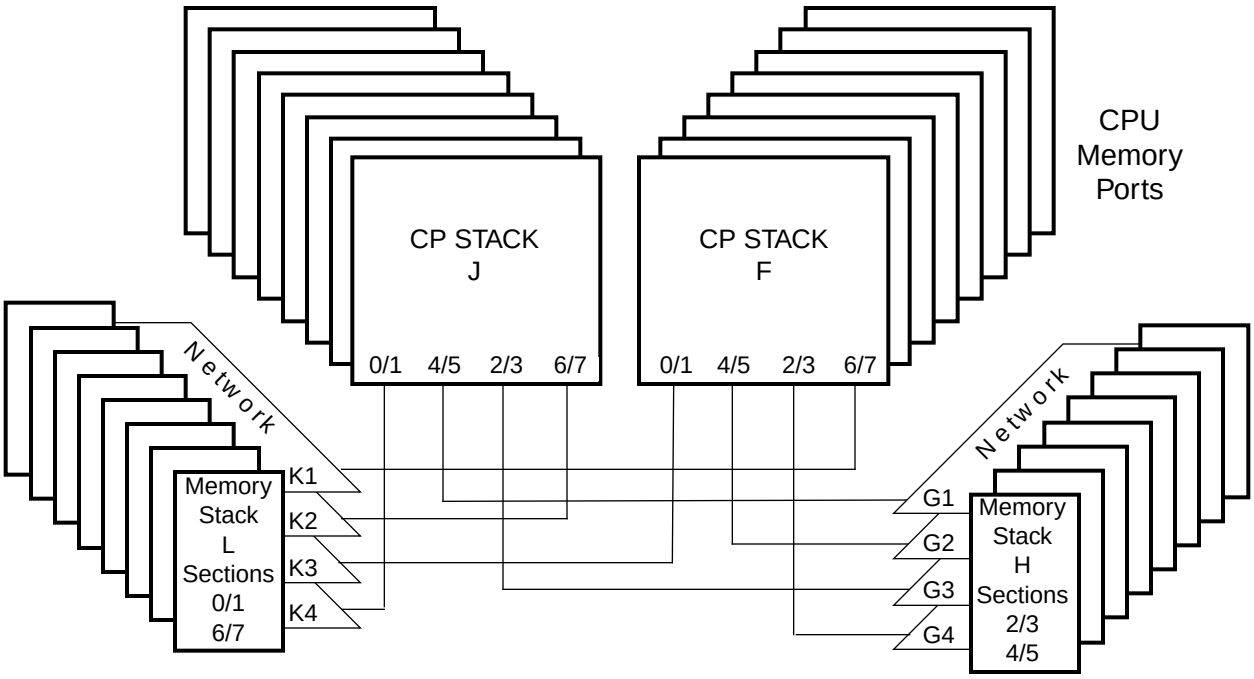
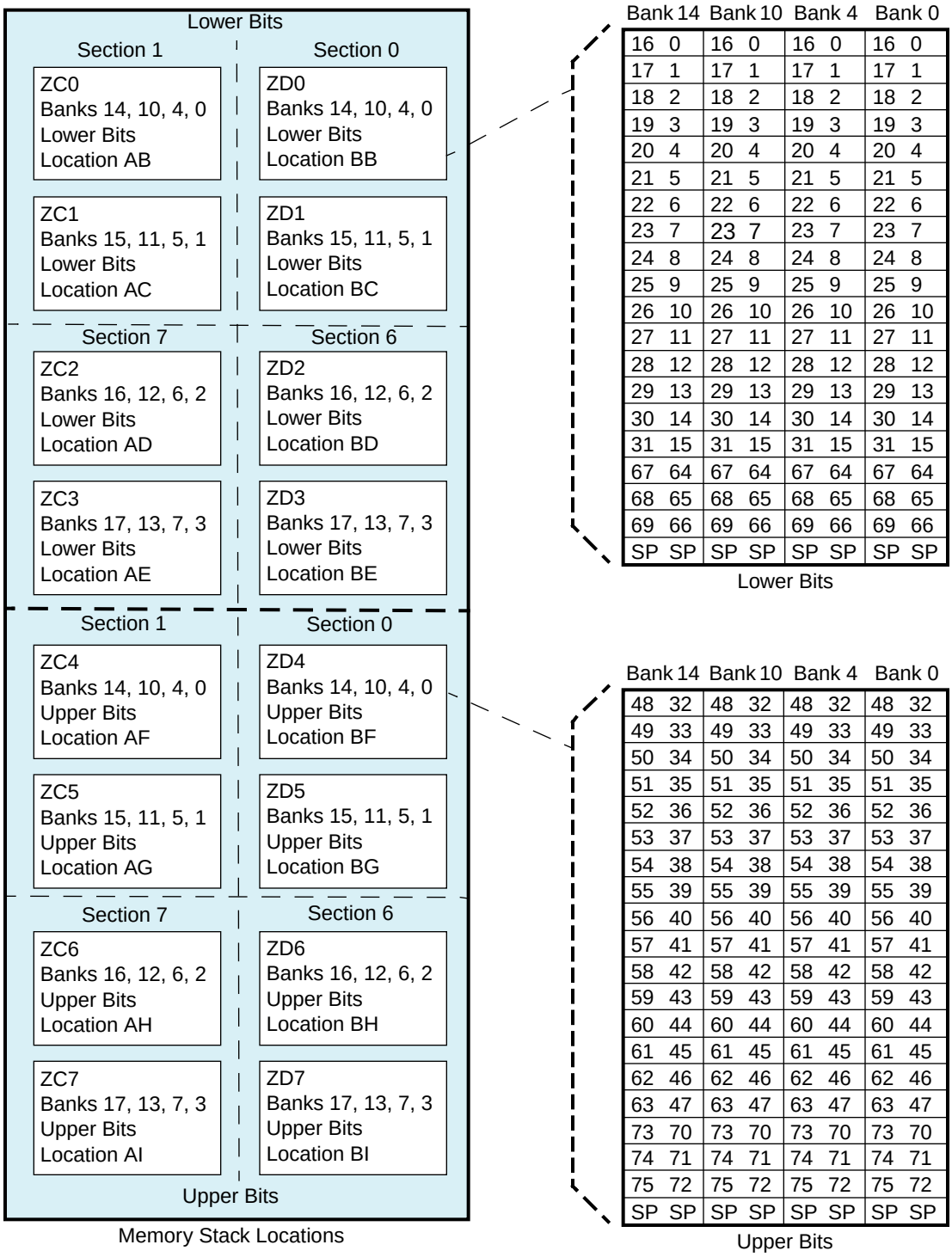


Figure 12 further illustrates the memory module layout for a CRAY T916 system. Each memory module stack handles 4 sections of memory. Each memory module within the stack is 1 subsection for each of the 4 sections. Each subsection has 8 banks instead of 16 banks as in the CRAY T932 and CRAY T94 systems. The individual memory stacks work the same way in the CRAY T916 system as they do in the CRAY T94 and CRAY T932 systems.

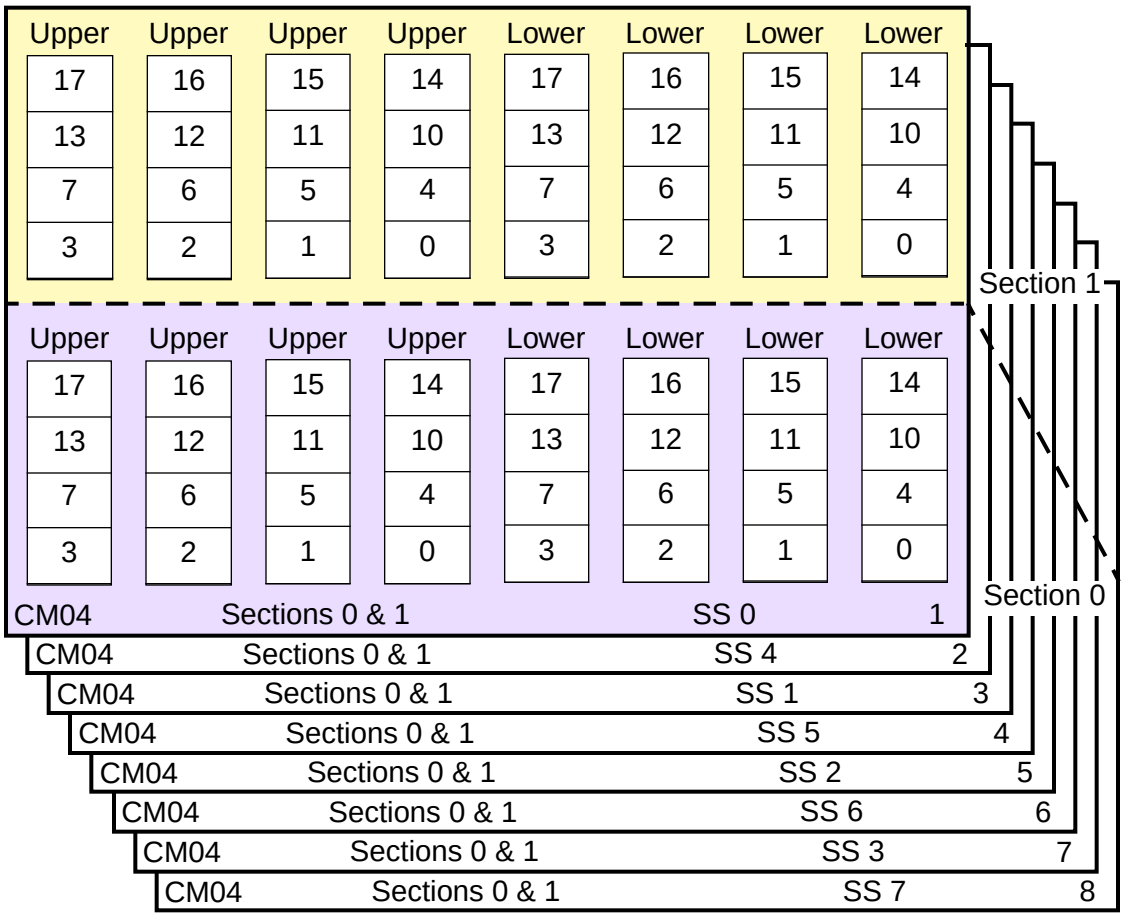
Figure 12. CRAY T916 Memory Module Bank and Bit Layout



CRAY T932 Memory

CRAY T932 memory is organized by sections, subsections, and banks. A fully configured system has 8 sections of memory. Each section has 8 subsections and each subsection has 16 banks, for a total of 1,024 banks. Figure 13 shows memory organization in a CRAY T932 system. Figure 13 illustrates only one of the four quadrants of a CRAY T932 memory. Each memory module within the memory module stack is 1 subsection for each of 2 sections.

Figure 13. CRAY T932 Memory Organization



The CP, network, and memory modules determine which section, subsection, bank, and word in memory to address. Figure 14 shows the addressing map for a fully configured CRAY T932 memory. It also shows the module type that determines the section select, subsection select, bank select, and word select. Bits 31 through 34 are presently not used; however, these bits may be used in future systems to address up to 32 Gwords of memory.

Figure 14. CRAY T932 Addressing Map

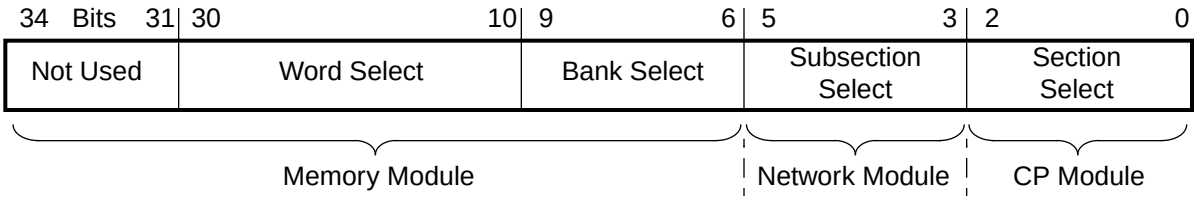


Table 6 lists the various configurations of CRAY T932 systems. A CRAY T932 system has two or four memory module stacks with each module stack containing either two, four, or eight memory modules. The memory modules connect to the network modules, which connect to the CP modules. The system interconnect board (SIB) connects these memory modules to the network modules and to the CP modules.

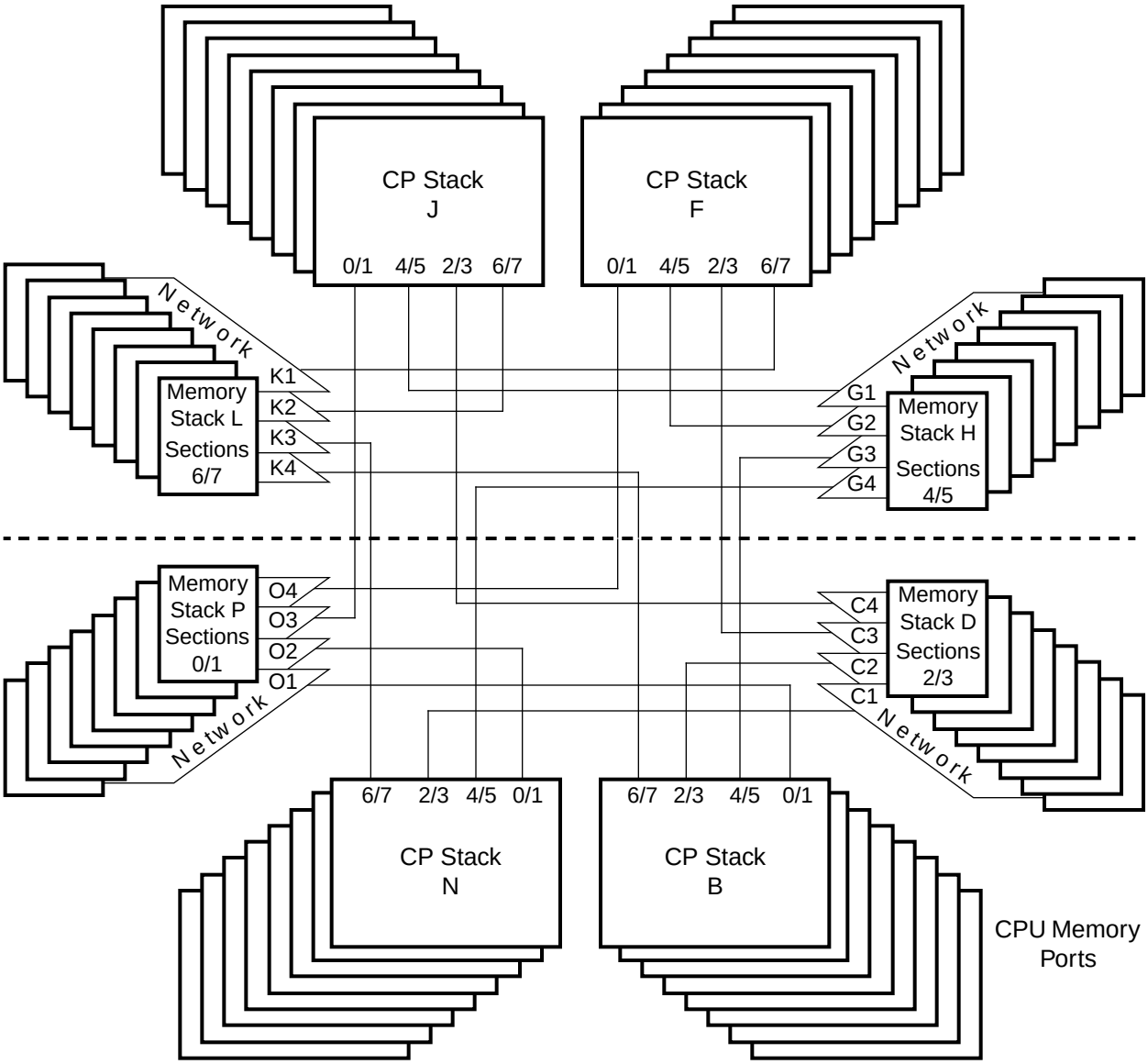
Table 6. CRAY T932 System Configuration

Module Counts			Configuration			
Processor	Network	Memory	Sections	Subsects	Banks	MWords
8	4	32	8	8	1024	2,048
8	4	16	8	4	512	1024
8 [a]	4	8	8	2	256	512
8	8	8	8	2	256	512
8 [a]	8	8	8	2	128	256
8 [a]	4	4	4	2	128	256
8 [a]	4	4	4	2	64	128
8 to 16	8	16	8	4	512	1,024
8 to 16 [a]	8	16	8	4	256	512
8 to 16 [a]	8	8	8	2	256	512
8 to 16 [a]	4	8	4	4	256	512
16	16	16	8	4	512	1,024
16 [a]	16	16	8	4	256	512
16 [a]	16	8	8	2	256	512
16 [a]	8	8	4	4	256	512
16 to 32	16	32	8	8	1024	2,048
16 to 32 [a]	16	32	8	8	512	1,024
16 to 32 [a]	16	16	8	4	512	1,024
16 to 32 [a]	8	16	4	8	512	1,024

[a] This indicates a degraded or partitioned system; the system is not usually sold with this configuration.

Figure 15 shows the logical interconnections between the CP modules and memory modules in a CRAY T932 system. Assuming the first reference is to section 0 and the memory references are sequential, references one and two go to sections 0 and 1, subsection 0 at module location P1. References three and four go to sections 2 and 3, subsection 0 at module location D1. References five and six go to sections 4 and 5, subsection 0 at module location H1. References seven and eight go to sections 6 and 7, subsection 0 at module location L1. This sequence continues as a descending spiral through the module stacks and memory subsections until all subsections and banks have been addressed.

Figure 15. CRAY T932 CPU-to-memory Interconnections



Chip Flawing

Each memory stack has 4 spare memory chips, one for each quarter of the stack. This feature enables you to flaw out a bad memory chip from that quarter of the memory stack and shift the data up to the spare chip. Use the maintenance program SCE to perform chip flawing. Refer to “Spare Chip Memory Management” in the *SCE User Guide*, publication number HDM-069-C, for information on how to enter or remove a flaw.

Error Correction

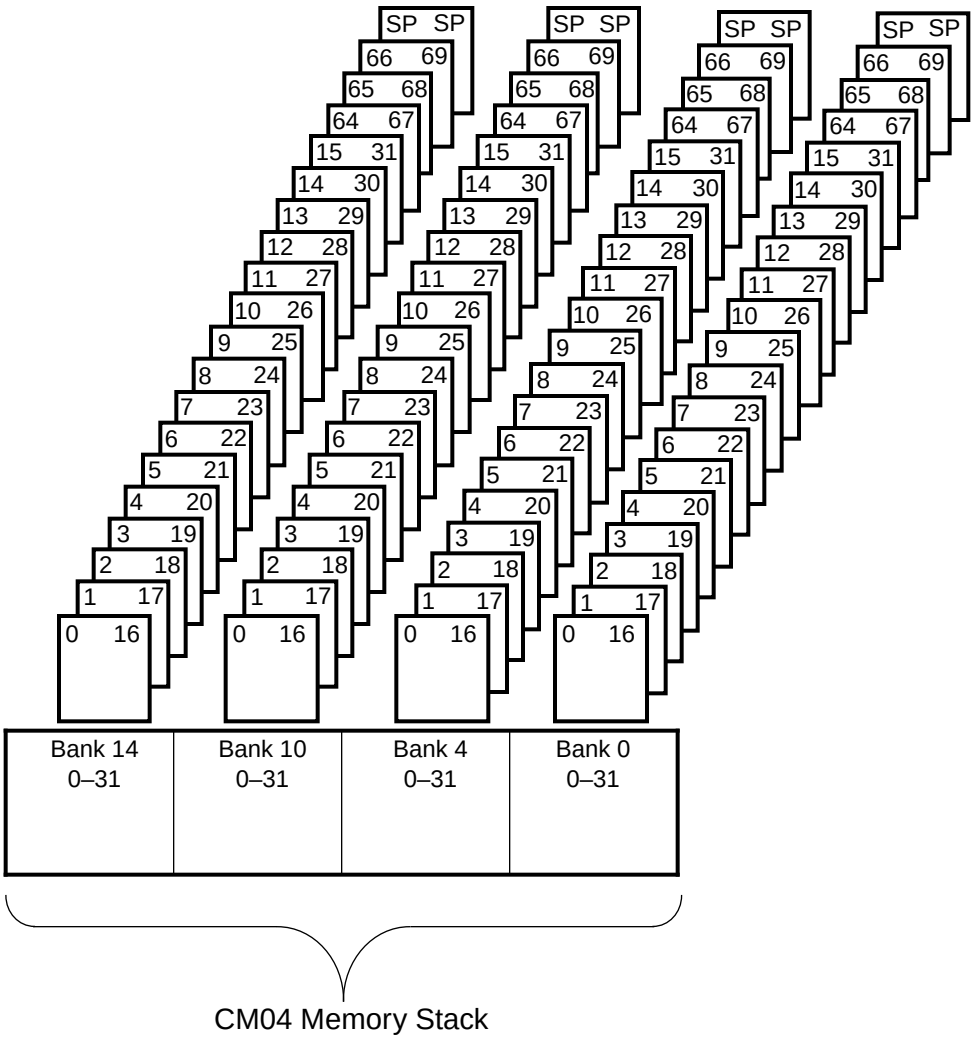
Error detection/correction is done on the CPU module. The CPU module must be able to accommodate a number of configurations: the CM02 module, which handles 2 bits per memory chip; the CM03 module, which handles either 2 or 4 bits per memory chip and 1 or 2 banks per half-stack; and the CM04 module, which handles 2 bits per memory chip and 1 bank per quarter-stack.

The CPU module is capable of correcting single-bit errors. It can correct 2 single-bit errors in the same half-bank if they are on the same memory chip. The 2 correctable **data bits** on 1 memory chip are 16 bits apart. For example, bits 0 and 16 are on 1 memory chip. Bits 1 and 17 are on another memory chip. Bits 0 and 16 are correctable **or** bits 1 and 17 are correctable, but not bits 0 and 1, or 0 and 17, and so on. The 2 correctable **check bits** on 1 chip are only 3 bits apart, bits 70 and 73 for example. Refer to Figure 16.

Memory Stack

Each memory stack handles half the bits of a word for 4 banks, and each quarter of the stack handles half the bits for banks n, n+4, n+10, or n+14. Refer to Figure 16 for an illustration of how the bits are **physically** laid out.

Figure 16. Memory Stack Bit Layout

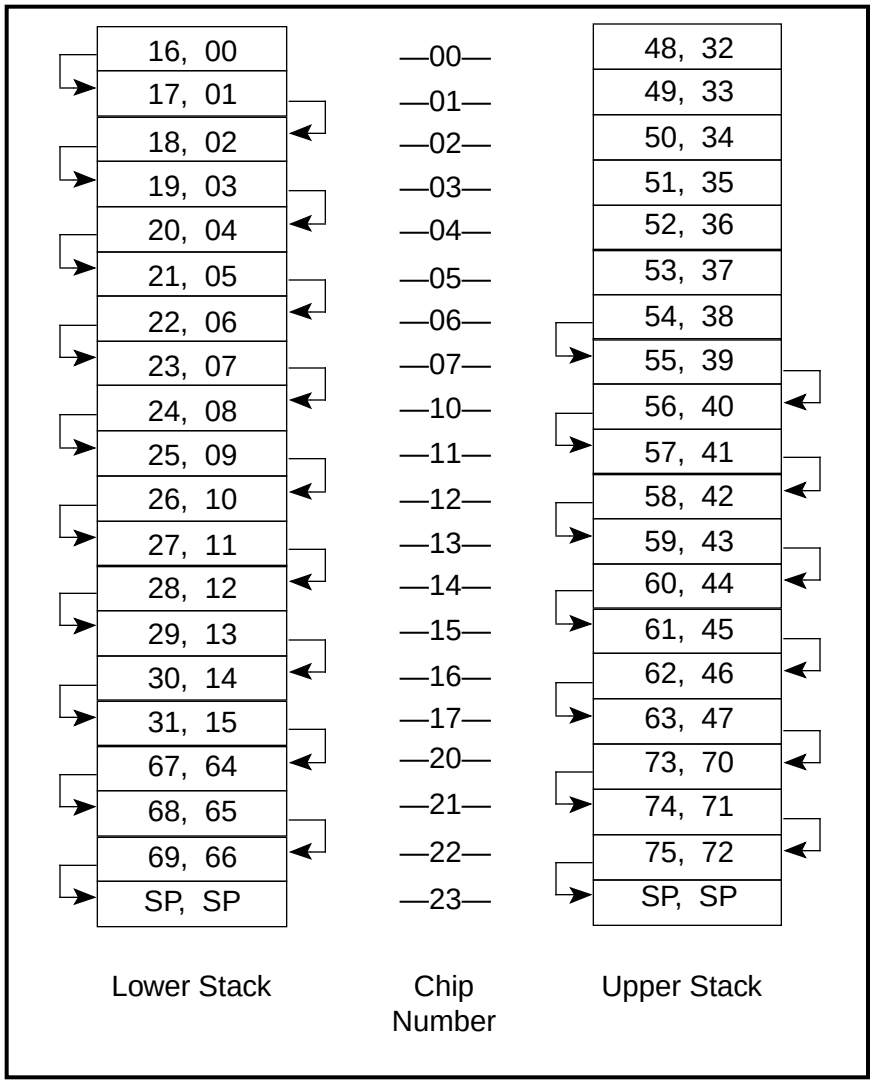


Each memory quarter-stack contains 20 memory chips: 19 chips for data and 1 spare chip. The spare chip on each quarter-stack accommodates one flaw for the 19 data chips in that quarter-stack. If, for example, the memory chip that handles bits 0 and16 for bank 0 is flawed out, then there are no remaining flaws for this quarter of the stack.

Bit-shift Pattern

Once a chip is flawed out, its bits are shifted on the MI options to the next chip **for that bank.** Refer to Figure 17. Notice on the **lower bits quarter-stack** that chip 0's bits are shifted to chip 1, chip 1 is shifted to chip 2, and so on, until chip 23, the spare chip, is reached. Furthermore, if chip 6's bits on the **upper bits quarter-stack** are shifted in turn from chip to chip until the spare is used, chips 0 through 5 are not affected. The **upper bits quarter-stack** functions identically to that of the lower bits quarter-stack.

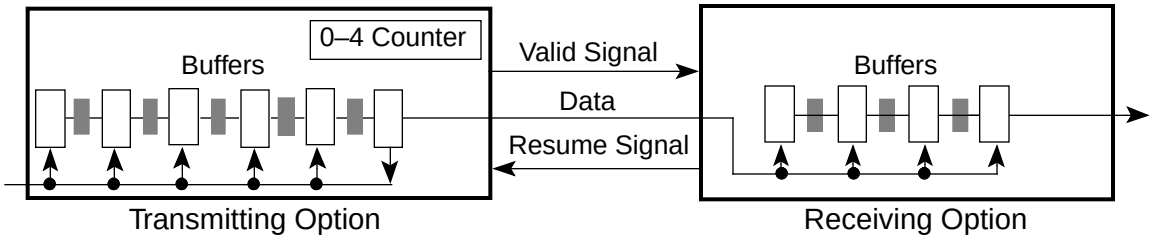
Figure 17. Memory Stack Bit-shift Pattern



Memory Overview

The communication protocol, or *handshaking*, that occurs between a CP module and memory module and between the options on the memory module uses Valid and Resume signals. Each transmitting option contains a counter. Each receiving option contains buffers. The transmitting option can send as many Valid signals to the receiving option as there are buffers in that option. With each Valid signal sent, the counter in the transmitting option increments by one. When the reference advances to the next stage, the receiving option sends a Resume signal to the transmitting option. The counter in the transmitting option then decrements by one. Figure 18 illustrates the handshaking that occurs between options.

Figure 18. Handshaking between Transmitting and Receiving Options

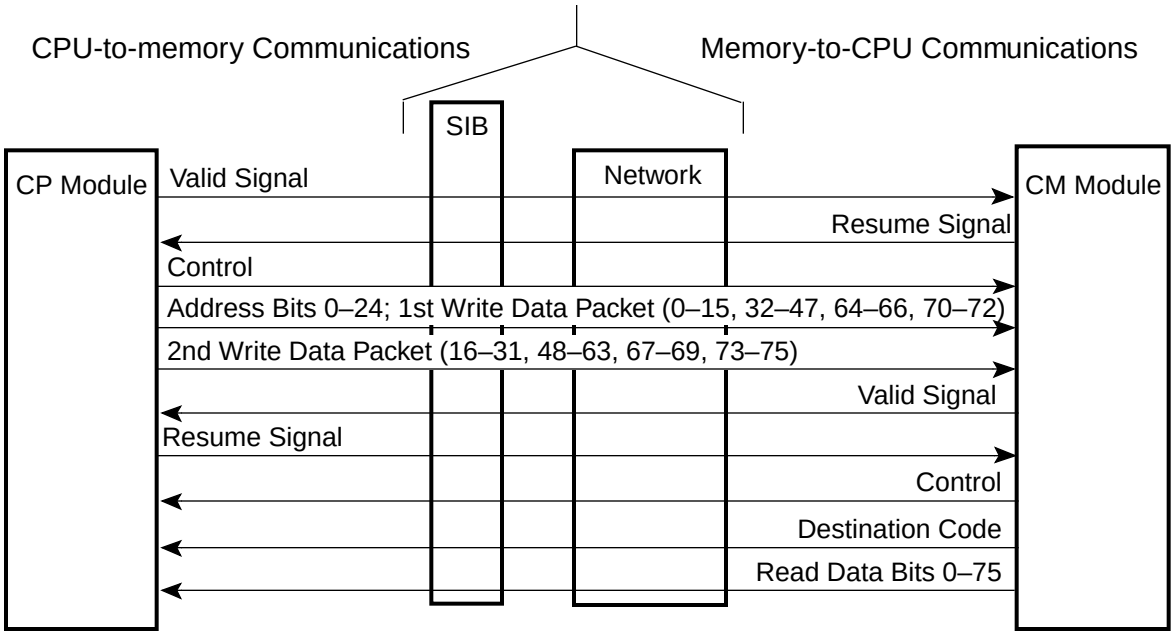


CP Module to Memory Module Communications

Figure 19 shows communications between a CP module and memory module during a memory reference. When a CPU references memory, it sends Valid signals to memory followed by control signals and then address and data. Control information includes **steering bits**, a **Write Reference** signal, and a **Reference Code**. The steering bits direct data to the appropriate bank, the Write Reference signal notifies memory whether the reference is a read or a write reference, and the Reference Code notifies memory of the type of reference coming from the CPU. During a read reference, the CPU also sends a **Destination Code** that gets buried into the first packet of the write data field. This Destination Code is sent back to the CPU with the data to ensure that the reference arrives at the appropriate location.

During a write operation, the CPU transfers data to memory in 2 data packets. Each data packet is 38 bits wide; 2 packets equal 1 data word. The data packets are sent in 2 clock periods (CPs). The first data packet is sent in CP 1, and the second data packet is sent in CP 2. During a read operation, the full 76-bit word is transferred to the CPU in 1 packet.

Figure 19. CPU-to-memory Communications

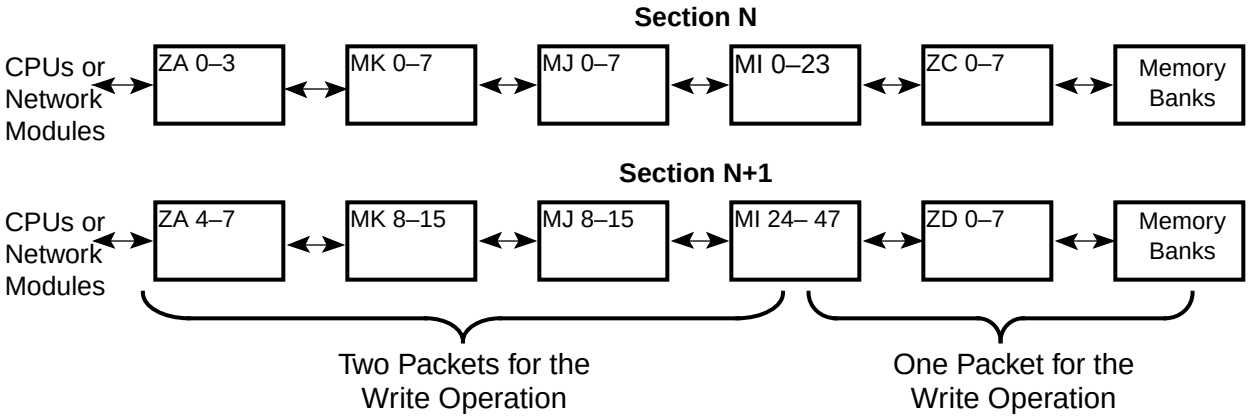


Memory Module Operations

Control, address, and data pass through one of eight ZA connectors to arrive on the memory module. Each ZA connector is associated with one CPU access path. During a memory write operation, the ZA connectors distribute control, address, and data to the MK options. The MK options buffer the control, address, and data and send it to the MJ options, which multiplex it and select its bank group destination. The MI options steer the control, address, and data to a specific bank within the bank group.

On a read from memory, the ZA connectors distribute the control, address, and destination code to the MK options. The MK options direct the address to the MJ options that select the bank group to address. The MI options steer the address to the specific bank within the bank group. Figure 20 is a block diagram of the memory module that shows the path through the options on the memory module. The following subsections explain the function of each option.

Figure 20. Option Path on the Memory Module

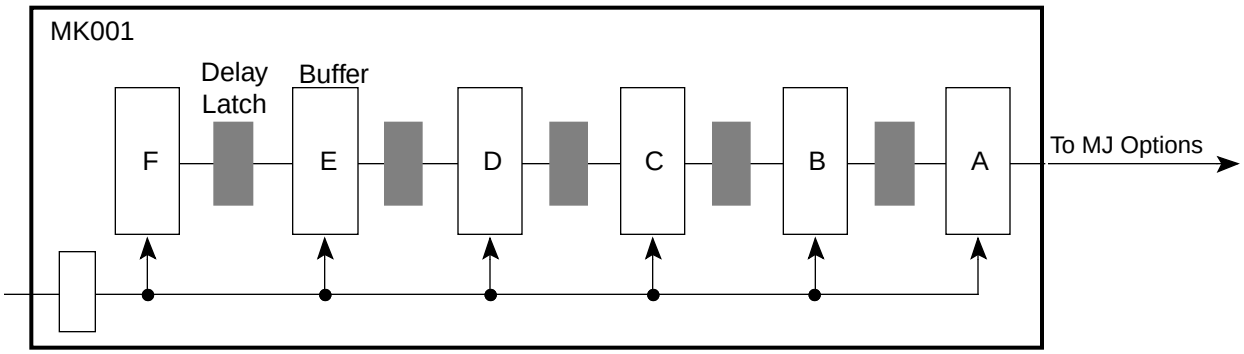


MK Option

The MK options are located adjacent to the ZA connectors (refer to Figure 1 and Figure 2); 2 MK options are associated with one ZA connector and are grouped by processor path, 2 MK options per processor path. Each option handles 32 bits of the data word, 6 of the check bits, and one-half of the address bits, steering bits, and Reference Code. The even-numbered option handles the lower bits (0 through 31 and half of the check bits), and the odd-numbered option handles the upper bits (32 through 63 and half of the check bits). Each MK option relays the data to 4 MJ options during a memory write operation and outputs the data to a ZA connector during a memory read operation.

Figure 21 is a block diagram of the MK option. The MK option provides a six-buffer relay between the CPU and memory and between memory and the CPU. Memory references arrive on the memory module from the processor and are latched in the MK options in a first-in-first-out order. The reference is advanced through the six buffers by a Valid signal from the CPU (CI option) and a Resume signal from the MK option. Memory references can be made as long as a buffer is available.

Figure 21. MK Option Block Diagram



MJ Option

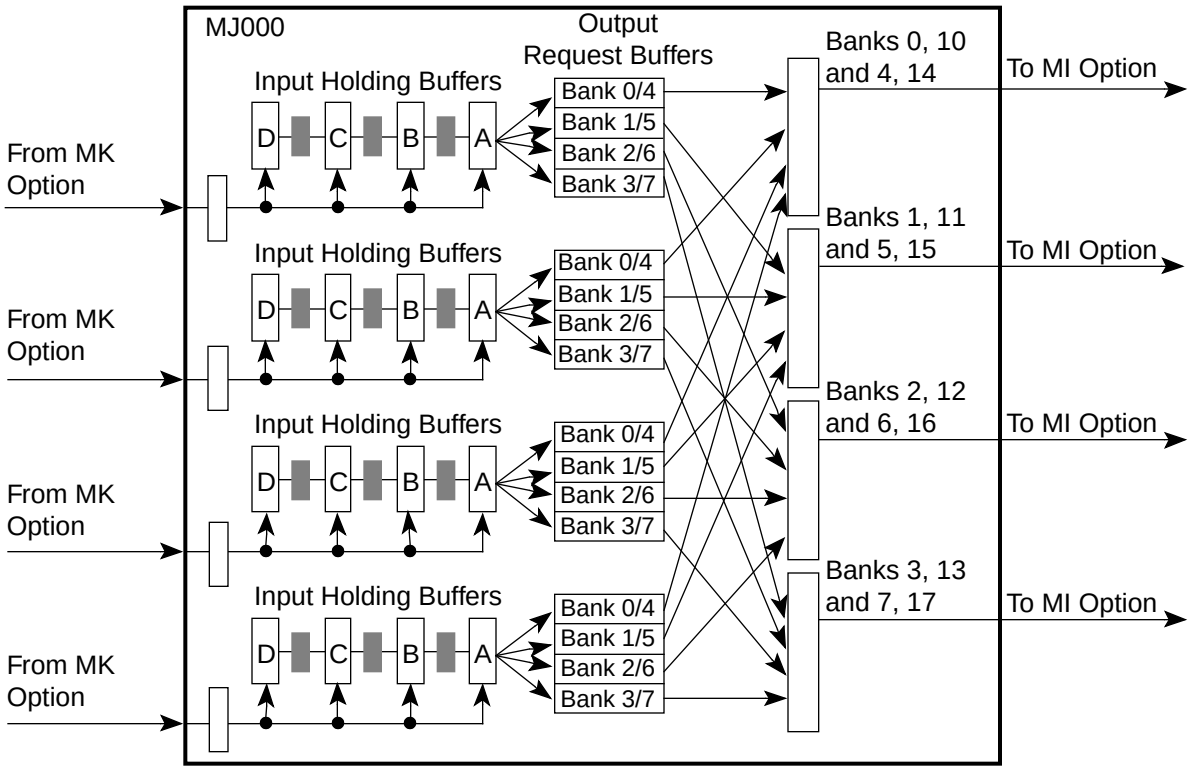
Each memory module contains 16 MJ options. The CM04 is made up of 2 sections, each of which utilizes 8 of the 16 MJ options. Each section is further divided into upper and lower bits. Four MJ options are used exclusively for the lower bits of a section, and 4 MJ options are used exclusively for its upper bits. The MJ options handle bank group and processor access. The priority is handled on a first-come, first-serve basis. When a reference gains access to a bank group or processor, that bank group or processor gets last priority the next time it is accessed. The MJ options have four input holding buffers that enable them to stack four memory references if the output request buffers are busy.

The MJ options receive 4 bank bits. Two of the bits (bank bits 0 and 1) determine which output buffer to select. The MJ option sends the other 2 bits to the MI options, which use it to determine the bank to address. Refer to Figure 22.

During a write reference, a group of 4 MJ options receives control, address, and data from 1 of 4 MK options, performs a 4 x 4 multiplex of the control, address, and data, and then sends it to 3 MI options, for half of the bits of a word. The other half of the bits of the word are handled by another MK option, 4 other MJ options and 3 other MI options. The MJ options receive data in 2 packets; the first packet contains its lower data bits, and the second packet contains its upper data bits. Do not confuse this with upper and lower bits that are handled on separate options.

During a read reference, the MJ option does a 4 x 4 multiplex of the **data** and **Destination Code** that comes from the MI options. **Bank bits**, from the MI option, are sent back to the CPU along with the data. The **return path** bits are used by the MJ option to determine the CPU port to which the reference data must be steered.

Figure 22. MJ Option Block Diagram



MI Option

The MI option provides all the control signals necessary to access 4 banks of memory. The MI option also provides a spare chip selection feature that enables the maintenance port to configure around a bad memory chip.

The MI option receives 7 bits of address, 7, 8, or 9 bits of packeted write data (14, 16, or18 bits total), 4 bits of reference code, and 4 steering and control bits from the MJ multiplexer. The packeted write data is actually broken down into 6 or 7 bits of write data and 1 or 2 bits of copies of selected write data that are used for the shift in the flaw operation.

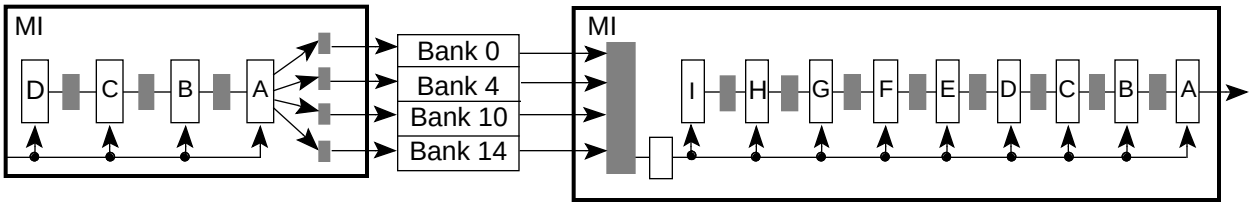
Each MI option controls approximately one-sixth of the data for a 4-bank group, which requires 6 MI options to work together on one bank group.

The write reference information coming into the MI option first goes into a 4-rank stack. If the bank requested by the reference in rank A of the stack is available, the reference information is gated directly out of that bank. If the requested bank is not available, the reference information gets held in rank A of the FIFO until the appropriate bank becomes available. The next reference in the stack can access its bank without having to wait the entire bank cycle time of the first

reference, unless the second reference is requesting the same bank as the first reference. If ranks A and B of the stack are both waiting for their banks, the reference in rank C can access its bank, if it is available. If ranks A through C are waiting and rank D's bank is available, the reference can proceed. Refer to Figure 23.

The MI option also receives 16 bits of readout data from each of the 4 banks, including 2 readout spare bits. There is a 9-rank FIFO in which the readout data and the corresponding reference information is stored before it is sent out to the MJ output multiplexer. The corresponding reference information includes steering, control, and destination code, which are all delayed on the MI option to ensure alignment with the appropriate readout data.

Figure 23. MI Option Block Diagram



Address Distribution

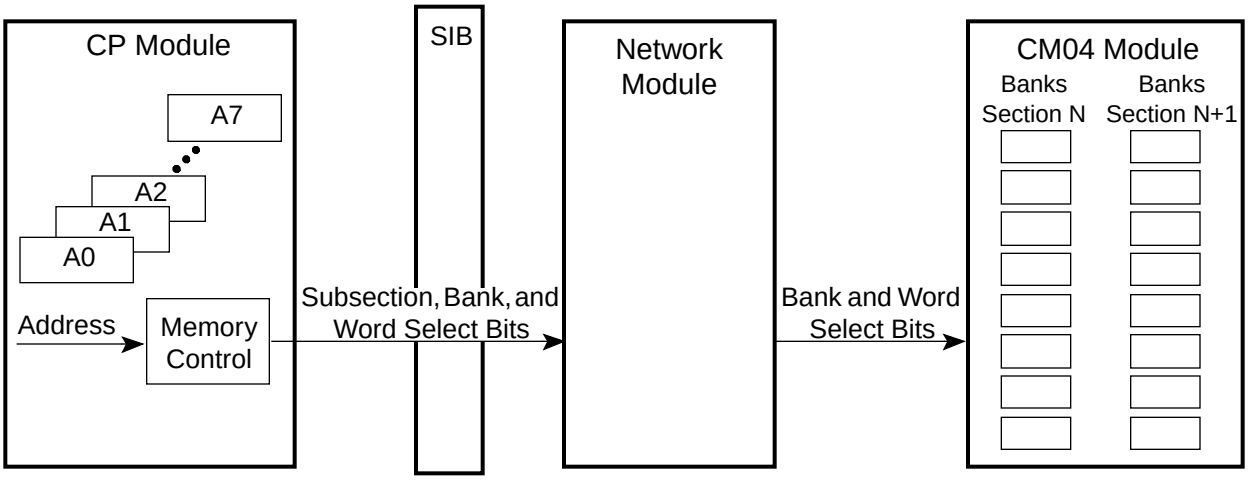
CRAY T90 series systems can use up to 35 address bits to determine which section, subsection, bank, and word in memory to address; however, not all address bits are used at this time (refer to "CRAY T94 Memory," "CRAY T916 Memory," and "CRAY T932 Memory" on page 14, page 19, and page 24 respectively, for information on addressing).

The CP module examines the lower 3 address bits (0 through 2) to determine which memory section receives the data and address. The remaining address bits (3 through 33) enter the network module through a connector. The network module examines bits 3 through 5 to determine which subsection receives the address and data (for CRAY T916 and CRAY T932 systems only). The remaining address bits enter the memory module, which then determines which bank and word receives the address and data.

As address bits are distributed through the CP, network, and memory modules, the modules strip off some of the address bits after they determine the section, subsection, bank, or word to address. Then, some of the bits are replaced with

Return Path bits that are used to steer the reference back to the CP module that sent the reference. Figure 24 shows the distribution of address bits through the CP, network, and memory modules.

Figure 24. Address Distribution



When address bits arrive on the memory module, the even-numbered MK options receive word address bits 0 through 11, and the odd-numbered MK options receive word address bits 12 through 24. Figure 27 is a block diagram that shows the flow of address bits through the memory module. Table 10, Table 11, and Table 13 show the address bit assignment for the MK, MJ, and MI options, respectively. As the address bits flow through each option, the options strip off the bits that are no longer needed to determine which word in memory to address. The MI option requires 21 address bits to determine the word select.

Memory Write Operation: A Block Diagram Description

When a CPU writes data into memory, it sends a **Valid** signal followed by bank bits and control signals and then word address bits and data. A data word is 76 bits: 64 data bits and 12 error-correction bits. The write reference arrives on the memory module through one of eight ZA connectors that are associated with a CPU access path. The ZA connectors pass the reference to the MK options, which pass the data to the MJ options and then to the MI options before they arrive on the memory stack.

Write Control Signals

A write reference uses the following control signals: **Bank Bits (0, 1, 2,** and **3**), **Subsection Bits (0, 1, 2,** and **3)**, a **Write Reference** signal, and **Reference Type (0, 1, 2,** and **3)** bits. The MJ and MI options use the bank bits to steer the address and data to the appropriate bank in memory. The **Subsection Bits** only pass through the memory module options; the network module used these bits to determine which memory module to address.

The **Write Reference** signal informs each memory module option that the reference is a write reference. This signal is needed because data is written to memory in packet form, with each option receiving 2 data packets. The first data packet contains the lower data bits, and the second data packet contains the upper data bits. If the **Write Reference** signal is equal to 1, the reference is a write reference and the memory module options can expect to receive 2 data packets. If the **Write Reference** signal is equal to 0, then the reference is a read reference and the options can expect to receive 1 data packet.

The CPU generates the **Reference Type** bits and sends them to the memory module with every reference. Table 7 lists the MI option input terms that are associated with the **Reference Type** bits and decodes these bits by CPU function. Refer to Figure 28.

Figure 27 shows the control signals associated with a memory write operation; however, the figure illustrates only the options associated with a write operation when banks 0, 4, 10, and 14 are being sent from processor 0. Sending a complete data word to memory requires 2 MK options, 8 MJ options, and 6 MI options.

Table 7. MI Option Reference Type Bit Decode

Function	ICD	ICC	ICB	ICA
Abort	0	0	0	0
CPU or I/O read	0	0	0	1
CPU write	0	0	1	0
Not valid function	0	0	1	1
Set spare chip config.	0	1	0	0
Not valid function	0	1	0	1
Read spare chip config.	0	1	1	0
Not valid function	0	1	1	1
I/O write	1	X	X	X

X = either a 1 or a 0.

If the reference is an I/O write reference, the MI option sends 3 of the **Reference Type** bits (ICA, ICB, and ICC) with 4 forced bits back to the CPU as the **Destination Code**. If the reference is a CPU write reference, then the MI option forces 7 **Destination Code** bits and passes them back to the CPU. The CPU uses the **Destination Code** bits to determine that the reference actually completed. Table 8 lists the Destination Code bits forced by the MI option. Figure 29 shows how the MI distributes the **Reference Type** bits and how the **Destination Code** bits are sent back to the CP module via the MJ and MK options.

Table 8. MI Option Destination Code Forced Bits

Function	MI01		MI00				
	OID	OIC	OIG	OIF	OIE	OID	OIC
Read/abort	IAB	IAA	IAE	IAD	IAC	IAB	IAA
I/O write	0	0	0	ICC	ICB	ICA	1
Processor write	0	0	0	0	0	1	0
Set bad bit (Reconfigure)	0	0	0	0	1	1	0

The CPU counts the number of references that it sent out and then checks this count against the number of references that actually completed. This information is sent to the J options on the CP module. The J options then determine if memory is quiet or if the CPU should generate a hold issue condition.

Write reference control signals pass through the MK options and enter the MJ options. Table 10 lists the control and data bits that leave each MK option. Two MK options are needed to direct all the control signals from one processor to memory.

The MJ options use the **Bank Bits** to steer the address and data to the appropriate bank in memory. The MJ options use **Bank Bits 0** and **1** to steer the reference to one of the four 4-bank groups. **Bank Bit 2** and **3** signals are passed along to the MI options where they are used to determine which bank within the bank group to address. Once the MJ options determine which bank group to address, they drop **Bank Bits 0** and **1** and create **Return Path Bits 0 and 1**. The **Return Path Bit** is a 2-bit code that is forced to either a 0 or a 1 to designate which processor connector the reference came from. The MI options pass the **Return Path Bits** back to the MJ options where they are used to steer the reference response back to the initiating CPU.

The other write reference control signals pass through the MJ options and enter the MI options. Table 11 lists the control, address, and data bits that leave each MJ option. Eight MJ options handle the control, address, and data for four processors. Options MJ000 through MJ007 handle control, address, and data for processor connectors 0 through 3; options MJ008 through MJ015 handle control, address, and data for processor connectors 4 through 7.

Twenty-four MI options handle control, address, and data for four processors. Options MI000 through MI023 handle control, address, and data for processor accesses 0 through 3; options MI024 through MI047 handle control, address, and data for processor accesses 4 through 7. Table 13 lists the control, address, and data received by options MI000 through MI023 for processor accesses 0 through 3. Table 13 lists the control, address, and data received by options MI024 through MI047 for processor accesses 4 through 7. The MI options generate write enable signals that control the memory stack. The **Chip Select** signal is forced. The **Write Enable** signal must be present for a write reference to complete.

Write Data Path

Figure 25 shows the write data path for processors 0 through 7. Refer to this illustration for detailed descriptions of the data flow between options on a memory module.

Data enters the MK options in packet form; the lower bits (0 through 31 and 64 through 69) arrive on the even-numbered MK options in 2 clock periods, and the upper bits (32 through 63 and 70 through 75) arrive on the odd-numbered MK options. Each MK option sends data to 4 MJ options. The MJ options distribute the data to the MI options, which direct it to the appropriate memory bank.

Write Completion

The terms listed below are sent back to the CPU that originated the write reference in order to inform the CPU that the write reference has completed.

Return Path bits 0 and 1 steer the reference response back to the proper CPU access port (ZA0 through ZA7). The bits are forced on the MJ option to the value of the port number after the reference leaves the MK options for the MJ options. The return path bits travel with the reference to the MI options and then return to the MJ options to steer the response to the proper MK options.

The **Valid** signal is sent back to the CPU to inform it that the reference has completed. It is originated by the MI option and sent back with the other response bits.

Destination Code bits are formed on the MI option and returned to the CPU with the other response bits. To learn how these bits are formed, refer to page 39.

The **Subsection** and **Bank** bits are also returned to the CPU with the other response bits.

Memory Read Operation: A Block Diagram Description

When a CPU reads data from memory, it sends a **Valid** signal, control signals, an address, and a destination code to the memory module. The **Valid** and control signals, address, and destination code arrive on the memory module through one of eight ZA connectors that are associated with a CPU access path. The read destination code occupies the first packet of what would be the write data. These signals flow through the MK, MJ, and MI options and steer the read reference to the designated location in the memory stack to retrieve the data.

Data leaves the memory stack and enters the MI options. The MI options send the data and control signals back to the CPU through the MJ and MK options.

Read Control Signals

Figure 29 shows the control signals associated with a memory read operation; however, the figure illustrates only the options associated with a read reference for banks 0, 4, 10, and 14 for processor 0.

A read reference uses the same control bits as a write reference to access the memory stack: **Subsection** bits 0 through 3, **Bank** bits 0 through 3, **Reference Type** bits 0 through 3, and a **Write Reference** signal (which should be zero on a read reference).

On the way back to the CPU, a read reference carries with it the following control signals: **Subsection** bits 0 through 3, **Bank** bits 0 through 6, and **Destination Code** bits 0 through 13.

On the way to the memory module, the subsection bits are stripped off by the network module and replaced with return information. They flow through the options on the memory module and are used on the way back by the network module to steer the reference back to the originating CPU. On the way to memory, bank bits 0 through 3 are used by the MJ and MI options to steer the reference to the correct bank. They are then returned with 3 extra bits forced, **Bank bits 4 through 6**, so that the CPU can determine from which bank the data came.

The CPU generates the **Reference Type** bits. The MI options pull the **Destination Code** out of the write data field on a read reference. Table 7 lists the terms associated with the **Reference Type** bits and decodes the bits by CPU function. When the MI option detects a read reference, it checks the write data field and sends the **Destination Code** back to the CPU with the return data. Table 8 lists the **Destination Code** bits that leave the MI options. The CPU uses the **Destination Code** to determine what to do with the data. Refer to the *CPU Module (CP02)* document, publication number HTM-003-A, for more information on how the CPU decodes the Destination Code bits.

The MI options also receive **Return Path** bits. The **Return Path** is a 2-bit code that steers the reference back to the CPU that sent the reference. Read reference control signals pass through the MI options and enter the MJ options. Table 12 and Table 13 list the control and data bits that leave each MI option.

The MJ options use the **Return Path** bits to determine which processor the reference came from; the MJ options then drop these bits and add 2 forced **Bank Bits.** The CPU uses the **Bank Bits** to determine which memory bank the data came from.

During a read reference, bank bits 0 through 6 are reported on options MJ0 through MJ4, and on MJ8 through MJ12. These bank bits are used by the CPU for error reporting. At this time, bank bits 4 through 6 (MJ2, MJ4, MJ10, and MJ12) are unused and forced to zeroes. Bank bits 0 and 1 are dropped on the way to memory, but they are re-created by forcing inputs on the MJ option, depending on the 4-bank group from which the reference is returning. Table 9 shows how these bank bits are forced on the inputs to MJ3 and MJ11.

Table 9. Bank Bits for Error Correction

MJ003, MJ0011	MJ003, MJ0011	Bank Number
Bank Bit 1	Bank Bit 0	
Forced 0 (IEM)	Forced 0 (IEN)	Banks 0, 4, 10, 14
Forced 0 (IFM)	Forced 1 (IFN)	Banks 1, 5, 11, 15
Forced 1 (IGM)	Forced 0 (IGN)	Banks 2, 6,12, 16
Forced 1 (IHM)	Forced 1 (IHN)	Banks 3, 7, 13, 17

Control signals leave the MJ options and enter the MK options. Table 11 lists the data and control signals that leave the MJ options. Table 10 lists the data and control signals that enter the MK options. The MK options send the data and control signals back to the CPU.

Read Data Path

Figure 26 shows the read data path for processors 0 through 7. Data leaves the memory stack as a 76-bit data word and enters the MI options. The spare bit is configured in the memory stack when a bad memory chip exists. If this bit is used, the MI options shift the bits to bypass the bad chip during a write operation and shift them back during a read operation. Refer to "Chip Flawing" on page 28 for more information. Data leaves the MI options and flows through the MJ and MK options before it leaves the memory module through the ZA connectors..

Table 10. MK Option Bit Assignments

Processor 0	
MK 000	MK 001
Processor 1	
MK 002	MK 003
Processor 2	
MK 004	MK 005
Processor 3	
MK 006	MK 007
Processor 4	
MK 008	MK 009
Processor 5	
MK 010	MK 011
Processor 6	
MK 012	MK 013
Processor 7	
MK 014	MK 015
Reference to Memory Signals	
Write Data 0–31, 64–69/Read Dest. 0–60	Write Data 32–63, 70–75/Read Dest 7–13
Address bits 0–11	Address bits 12–24
Subsection bits 0–01	Subsection bits 02–03
Bank bits 00, 02	Bank bits 01, 03
Reference type 0–01	Reference type 02–03
Write reference bit for lower bits	Write reference bit for upper bits
Valid in for lower bits	Valid in for upper bits
Resume from lower bits	Resume from upper bits
Reference from Memory Signals	
Read data 0–31, 64–69	Read data 32–63, 70–75
Subsection bits 0–01	Subsection bits 02–03
Bank bits 0–2	Bank bits 3–6
Destination bits 0–6	Destination bits 7–13
Valid out from lower bits	Valid out from upper bits
Resume in from lower bits	Resume in from upper bits

Table 11. MJ Option Bit Assignments

	Lower Bits				Upper Bits			
Proc. 0–3	MJ00	MJ01	MJ02	MJ03	MJ04	MJ05	MJ06	MJ07
Proc. 4–7	MJ8	MJ9	MJ10	MJ11	MJ12	MJ13	MJ14	MJ15
Write Data								
1st Data Packet	00–04	5–9	10–14	15 64–66	32–36	37–41	42–46	47 70–72
2nd Data Packet	16–20	21–25	26–30	31 67–69	48–52	53–57	58–62	63 73–75
Addr Bits	18–23	06–11	12–17	00–05	18–23	06–11	12–17	00–05
Ref Type	2–3	2–3	0–1	0–1	2–3	2–3	0–1	0–1
SS Bit			1	0			3	2
Bank Bit	0–3	0–3	0–1	0–1	0–3	0–3	0–1	0–1
Return Bit			1	0			1	0
Write Ref.	X	X	X	X	X	X	X	X
Read Data	0–9	10–15 64–66	16–25	26–31 67– 69	32–41	42–47 70–72	48–57	58–63 73–75
Destination	0–1	2–3	4–5	6	7– 8	9–10	11–12	13
Return Bit	0–1	0–1	0–1	0–1	0–1	0–1	0–1	0–1
SS Bit	0	1			2	3		
Bank Bit	2	3	4–5		6			
Force Bank Bit				0–1				

Table 12. MI Option Bit Assignments (Processors 0–3)

Processors 0–3						
Banks 0, 4, 10, 14	MI00 2CB	MI01 2BB	MI02 2AB	MI03 2DF	MI04 2CF	MI05 2BF
Banks 1, 5, 11, 15	MI06 2CC	MI07 2BC	MI08 2AC	MI09 2DG	MI10 2CG	MI11 2BG
Banks 2, 6, 12, 16	MI12 2CD	MI13 2BD	MI14 2AD	MI15 2EH	MI16 2Dh	MI17 2CH
Banks 3, 7, 13, 17	MI18 2CE	MI19 2BE	MI20 2AE	MI21 2EI	MI22 2DI	MI23 2CI
Write Data	0–3, 16–19, 4, 20	5–8, 21–24, 9, 25	10–14 26–30 15, 31	32–35 48–51 36, 52	37–40 53–56 41, 57	42–46 58–62 47, 63
Check Bits	64, 67	65, 68	66, 69	70, 73	71, 74	72, 75
Extra Copies for Logical Chip Sparing	15, 31	4, 20 64, 67	9, 25 65, 68	47, 63	36, 52 70, 73	41, 57 71, 74
Address Bits	14–20	7–13	0–6	14–20	7–13	0–6
Ref Type Bits	0–3	0–3	0–3	0–3	0–3	0–3
SSec Bits	0	1		0	1	
Bank Bits	2–3	2–3	2–3	2–3	2–3	2–3
Return Path	0	1		0	1	
Read Data	0–5 16–21	6–11 22– 27	12–15 64–66 Spare 28–31 67–69 Spare	32–37 48–53	38–43 54–59	44–47 70–72 Spare 60–63 73–75 Spare
Destination	0–4	5–6		7–11	12–13	
Return Path	0	1		0	1	
SSec Bits	0	1		2	3	
Bank Bits			2–3			2–3

Table 13. MI Option Bit Assignments (Processors 4–7)

Processors 4–7						
Banks 0, 4, 10, 14	MI24 2JB	MI25 2IB	MI26 2HB	MI27 2IF	MI28 2HF	MI29 2GF
Banks 1, 5, 11, 15	MI30 2JC	MI31 2IC	MI32 2HC	MI33 2IG	MI34 2HG	MI35 2GG
Banks 2, 6, 12, 16	MI36 2JD	MI37 2ID	MI38 2HD	MI39 2HH	MI40 2GH	MI41 2FH
Banks 3, 7, 13, 17	MI42 2JE	MI43 2IE	MI44 2HE	MI45 2HI	MI46 2GI	MI47 2FI
Write Data	0–3, 16–19, 4, 20	5–8, 21–24, 9, 25	10–14 26–30 15, 31	32–35 48–51 36, 52	37–40 53–56 41, 57	42–46 58–62 47, 63
Check Bits	64, 67	65, 68	66, 69	70, 73	71, 74	72, 75
Extra Copies for Logical Chip Sparing	15, 31	4, 20 64, 67	9, 25 65, 68	47, 63	36, 52 70, 73	41, 57 71, 74
Address Bits	14–20	7–13	0–6	14–20	7–13	0–6
Ref Type Bits	0–3	0–3	0–3	0–3	0–3	0–3
SSec Bits	0	1		0	1	
Bank Bits	2–3	2–3	2–3	2–3	2–3	2–3
Return Path	0	1		0	1	
Read Data	0–5 16–21	6–11 22– 27	12–15 64–66 Spare 28–31 67–69 Spare	32–37 48–53	38–43 54–59	44–47 70–72 Spare 60–63 73–75 Spare
Destination	0–4	5–6		7–11	12–13	
Return Path	0	1		0	1	
SSec Bits	0	1		2	3	
Bank Bits			2–3			2–3

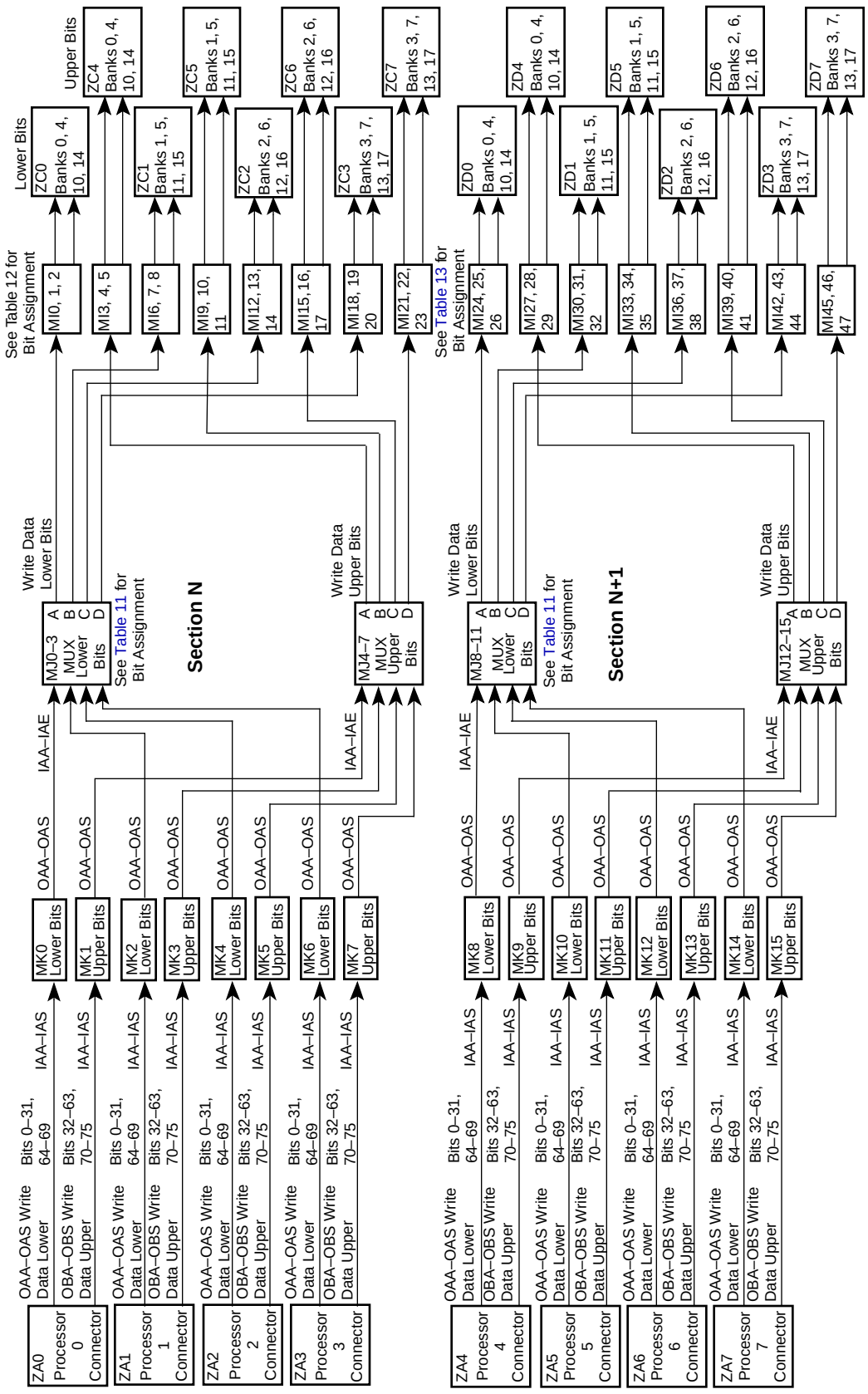
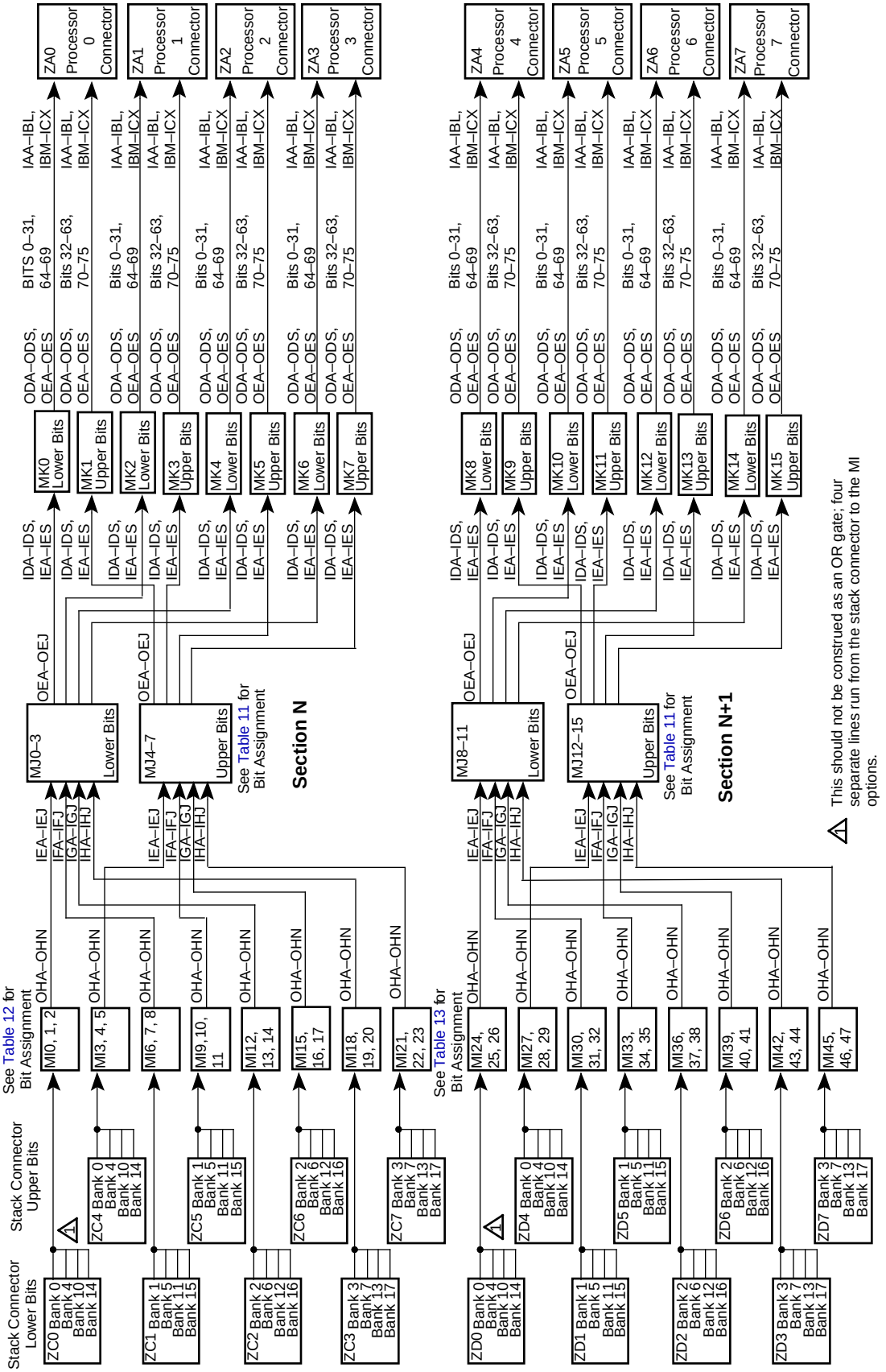


Figure 25. CM04 Memory Module Write Data Path

Figure 26. CM04 Memory Module Read Data Paths



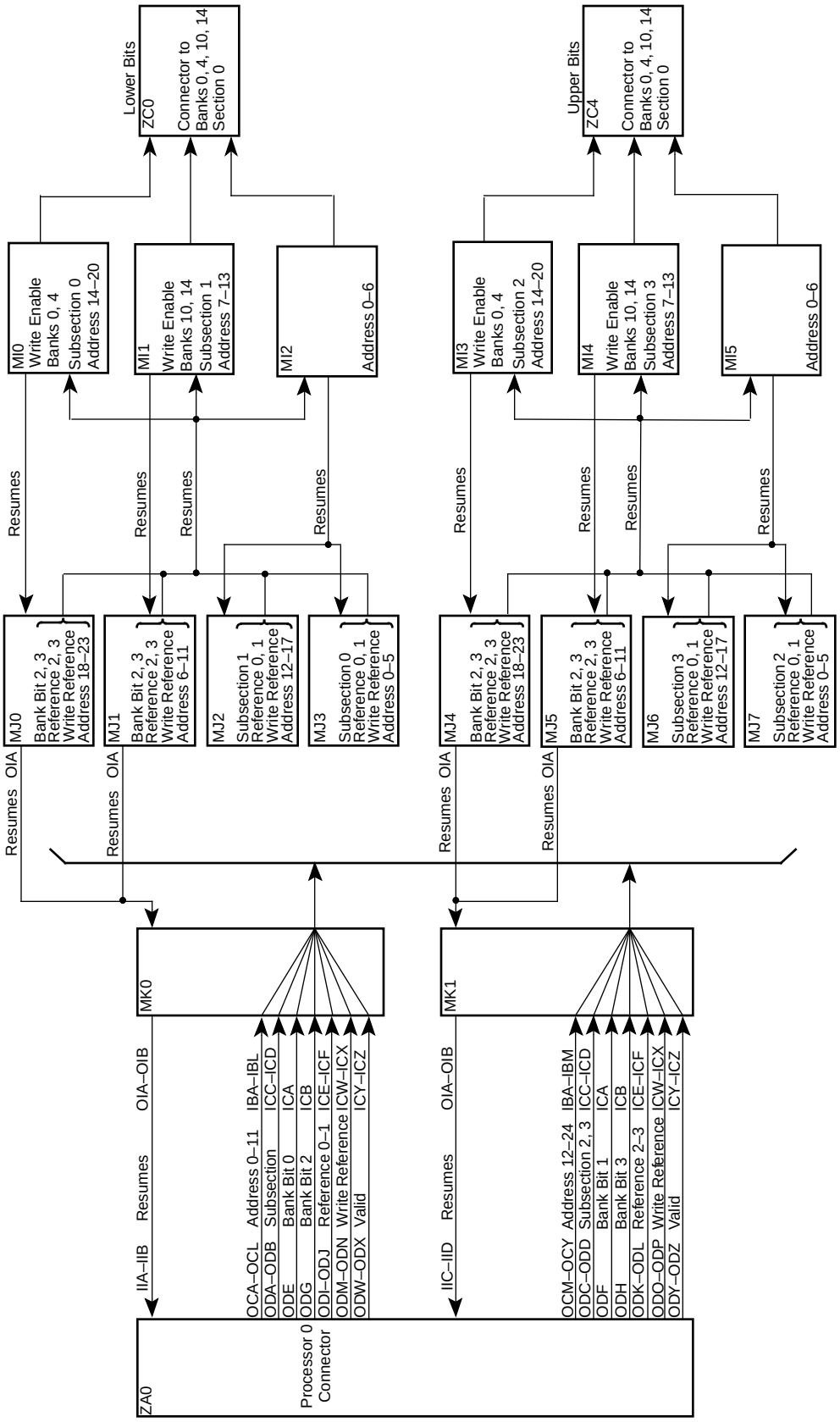


Figure 27. CM04 Memory Module Processor-to-memory Address and Control Signals for Processor 0 to Section N (Banks 0, 4, 10, and 14 Only)

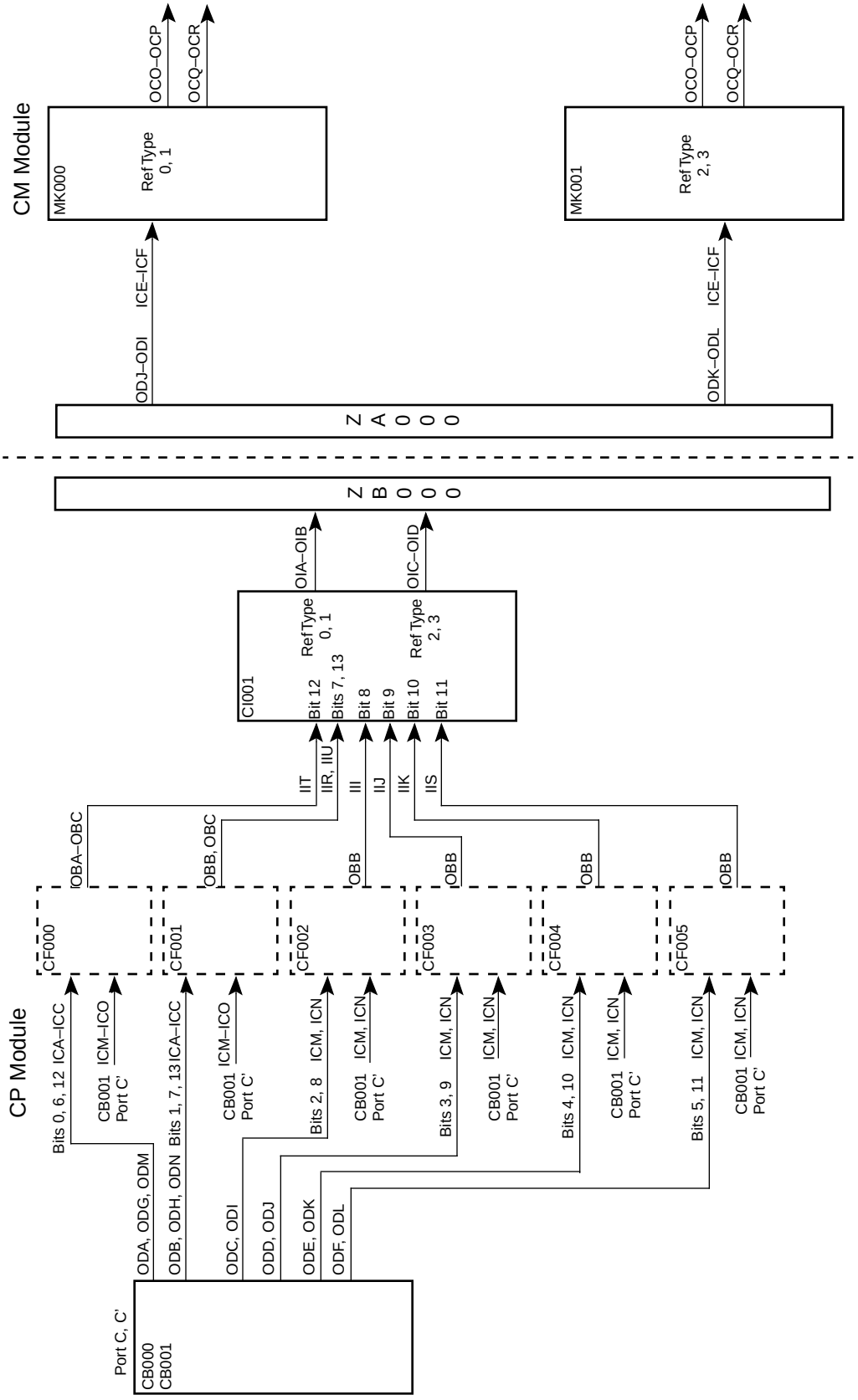


Figure 28. Reference Type Bit Generation on CI Option (CP02 Module)

Figure 29. Memory-to-processor Control Signals (Banks 0, 4, 10, and 14 to Processor 0 Only)