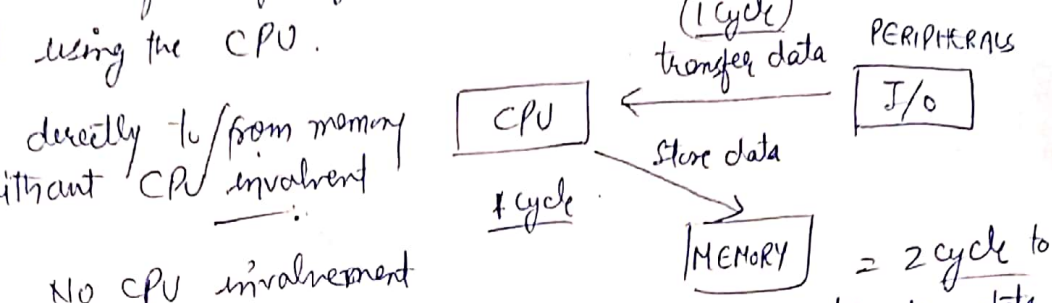
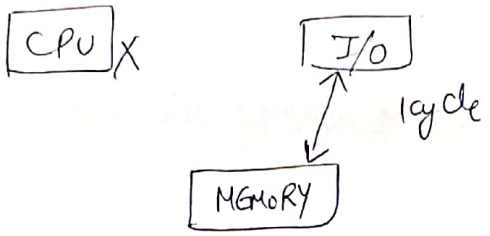
① DMA / DIRECT MEMORY ACCESS (Designed by INTEL)

- DMA provides direct access to Memory
- For Fast operation, DMA works as a bus master
- DMA is a method of transferring data b/w Peripherals & memory without using the CPU.
- Transfers data directly to/from memory without CPU involvement



- In DMA, No CPU involvement



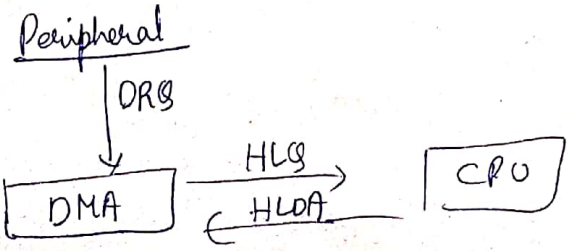
DMA DATA TRANSFER

- It is data transfer technique directly b/w memory & I/O without CPU intervention.
- Under the supervision of extra h/w called DMA Controller
- Fastest type of data transfer technique among Parallel Groups

WORKING

① Device (Peripheral) which wants to send data to memory first has to send DMA request (DRQ) to DMA Controller

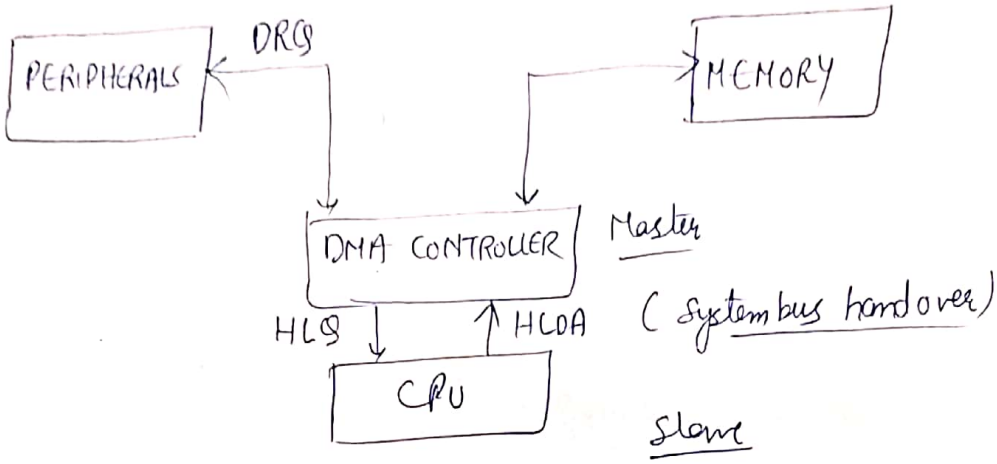
② DMA controller sends ——→ HLQ (Hold request) to CPU & waits for the CPU to send hold acknowledgment signal (HLDA)



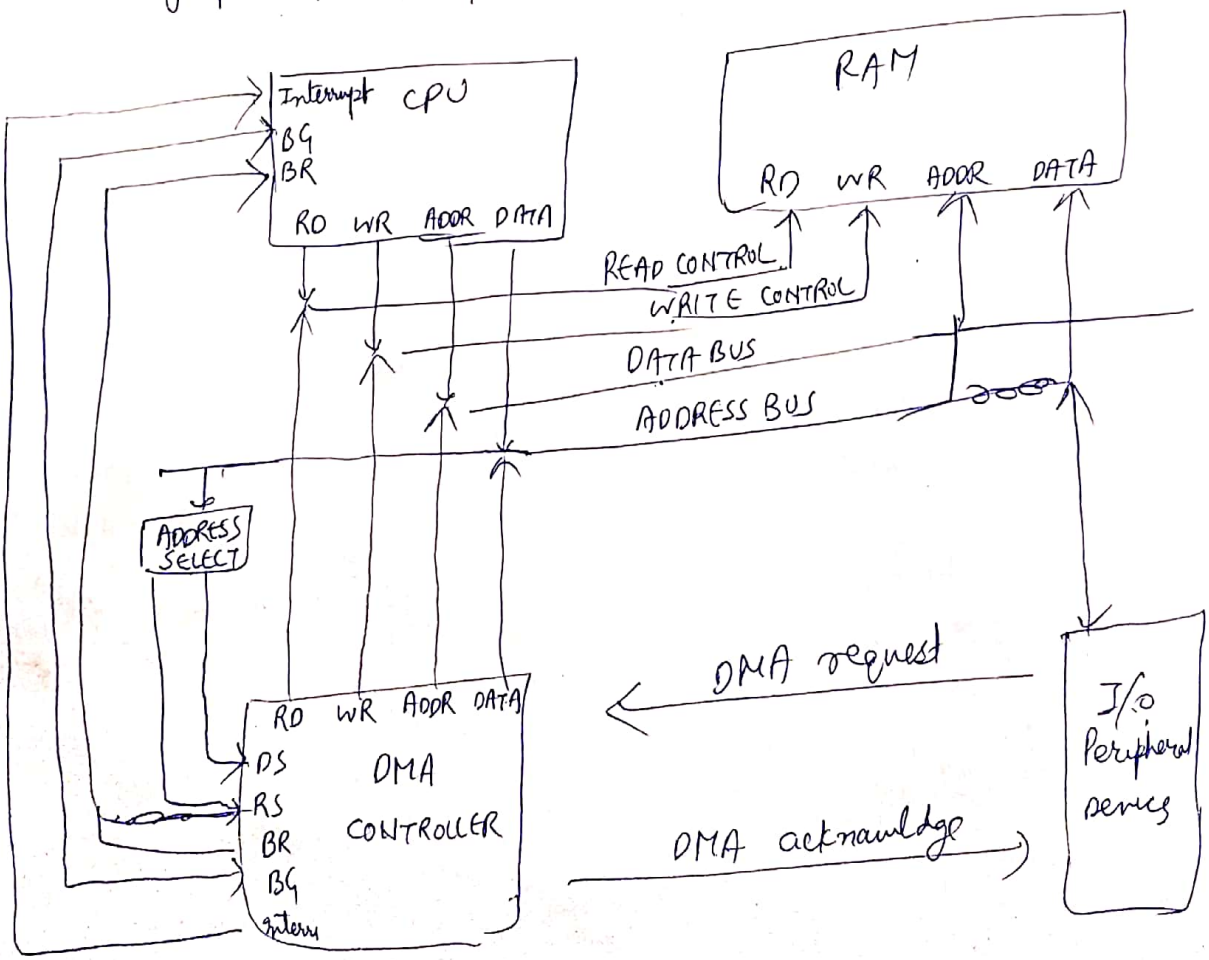
③ When DMA controller get HLDA signal Through System Bus, it becomes the Master & CPU becomes the slave

The CPU leaves the control bus over bus & ack the HLDA signal to DMA controller

④ Now CPU is in HOLD state, DMA controller has to manage operations over buses b/w CPU, Memory & I/O devices



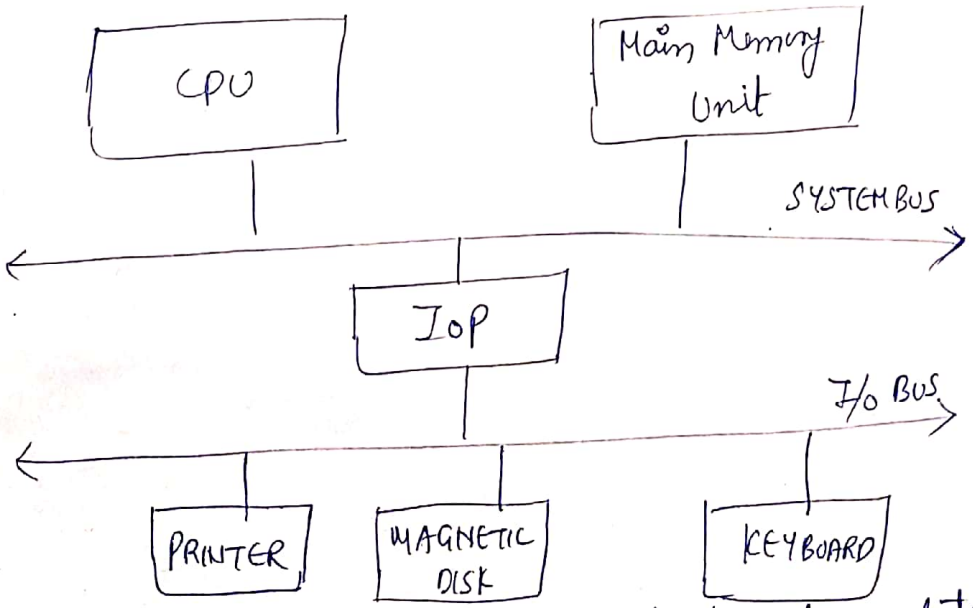
⑤ Many h/w system use DMA tech including disk drive controllers, graphics card, n/w card & sound cards.



DS - DMA SELECT
RS - REGISTER SELECT

INPUT - OUTPUT PROCESSOR

- The DMA mode of data transfer reduces CPU's overhead in handling I/O operation. It also allows Parallelism in CPU & I/O operations. Such Parallelism is necessary to avoid wastage of valuable CPU time while handling I/O devices whose speeds are much slower as compared to CPU.
- The concept of DMA operation can be extended to relieve the CPU further from getting involved in with the execution of I/O operations.
- This gives rise to the development of Special Purpose Processor called Input - output Processor (IOP) or IO channel
- The I/O Processor is just like a CPU that handles the details of I/O operations. It is more equipped with facilities than those available in typical DMA controller.
- The IOP can fetch & execute its own instructions that are specifically designed to characterize I/O transfers.



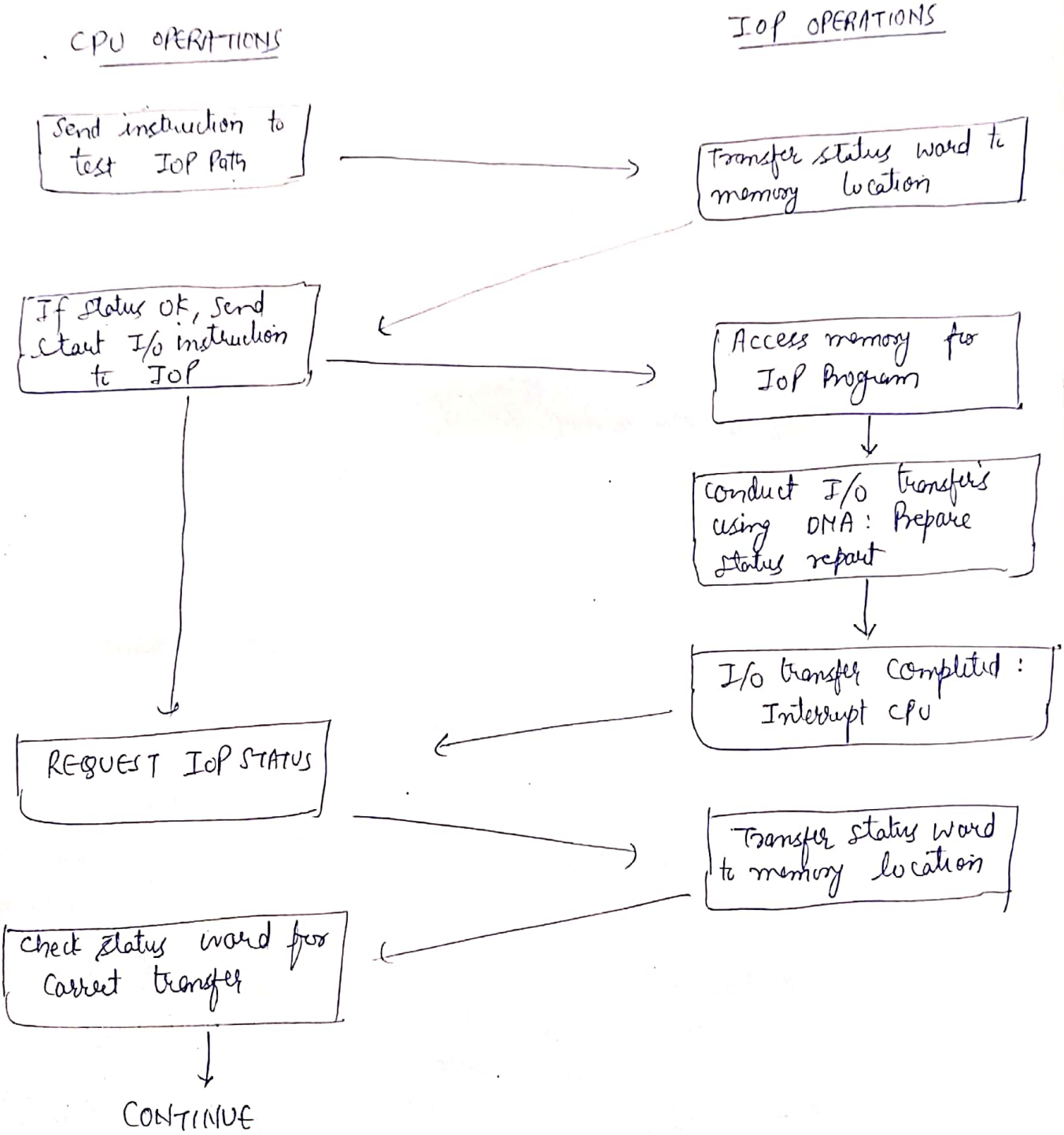
- The IOP is a specialized Processor which loads & stores data into memory along with execution of I/O instructions.

Advantages - ① The I/O devices can directly access the main memory without the intervention by Processor in I/O based systems
② It is used to address the Problems that arise in DMA method.

(3)

CPU – IOP COMMUNICATION

- The communication b/w CPU & IOP may take different forms depending on the particular computer considered. The sequence of operations during CPU & IOP communication is shown in figure



1. The CPU sends a test I/O instruction to IOP to test the IOP Path.
2. The IOP responds by inserting a status word in Memory location

NOTE - Word in CA is an ordered set of bytes or bits that is the normal unit in which Inform may be stored.

3. The CPU refers to the status word in Memory. If everything is in order, the CPU sends the start I/O instruction to start the I/O transfer

4. The IOP access memory for IOP program.

The CPU can now continue with another program while the IOP is busy with the program. Both programs refer to memory by means of DMA transfer

5. When the IOP terminates the execution of its program, it sends an interrupt request to the CPU

6. The CPU then issues a read I/O instructions to read the status from the IOP

7 The IOP transfers the status word to memory location

The status word indicates whether the transfer has been completed satisfactorily or if any error has occured during the transfer

MEMORY ORGANIZATION

- Memory unit is an essential component in any digital computer as it is needed for storing programs & data

MEMORY HIERARCHY - The memory hierarchy system consists of all storage devices employed in a computer system from slow but high capacity auxillary memory to relatively faster main memory to even smaller & faster cache memory.

a) **Cache memory** - A very high speed memory called a cache is sometimes used to increase the speed of processing by making current programs & data available to the

CPU at a rapid rate.

- The Cache memory is employed in computer systems to compensate for the speed differential b/w main memory access time & Processor.
- The Cache is used for storing segments of Programs currently being executed in the CPU & temporary data frequently needed in the Present calculations

I/O Processor manages data transfer b/w auxiliary memory & main memory whereas Cache is concerned with transfer of infon b/w main memory.

- The overall Goal of using a memory hierarchy is to obtain the highest Possible average access speed while minimizing total cost of entire memory system

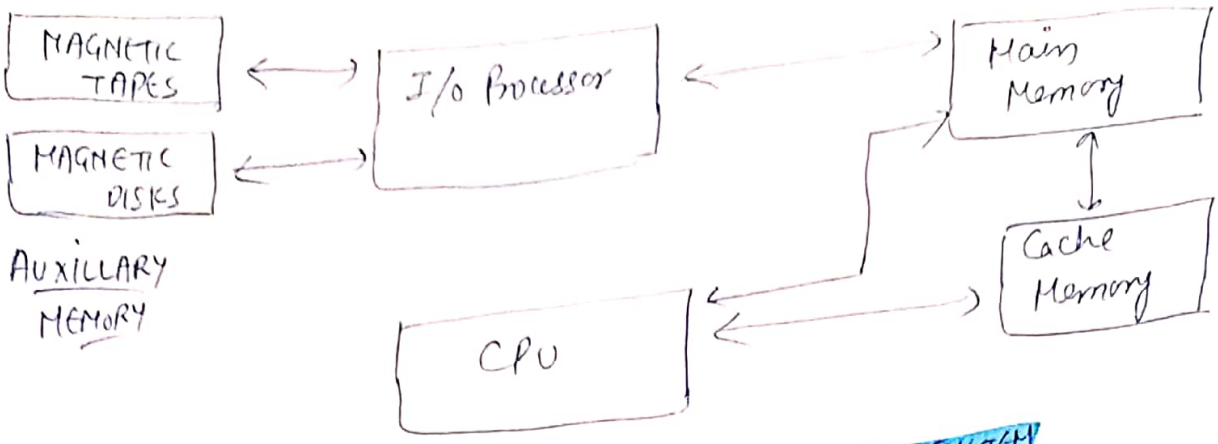
AUXILIARY MEMORY - has large storage capacity, is relatively expensive but has low access speed.

Cache Memory	Auxiliary Memory
• holds those Parts of Program & data that are most heavily used.	• holds those Parts that are not Presently used by CPU

- CPU has direct access to both Cache & main memory but not to auxiliary memory.

Note - Transfer from auxiliary to main memory is usually done by means of DMA.

Note - Typical access time ratio b/w Cache & main memory is about 1 to 7. Example, a Typical Cache memory may have an access time of 100ns while main memory access time may be 700ns. For auxiliary it is 1000 times that of main memory.



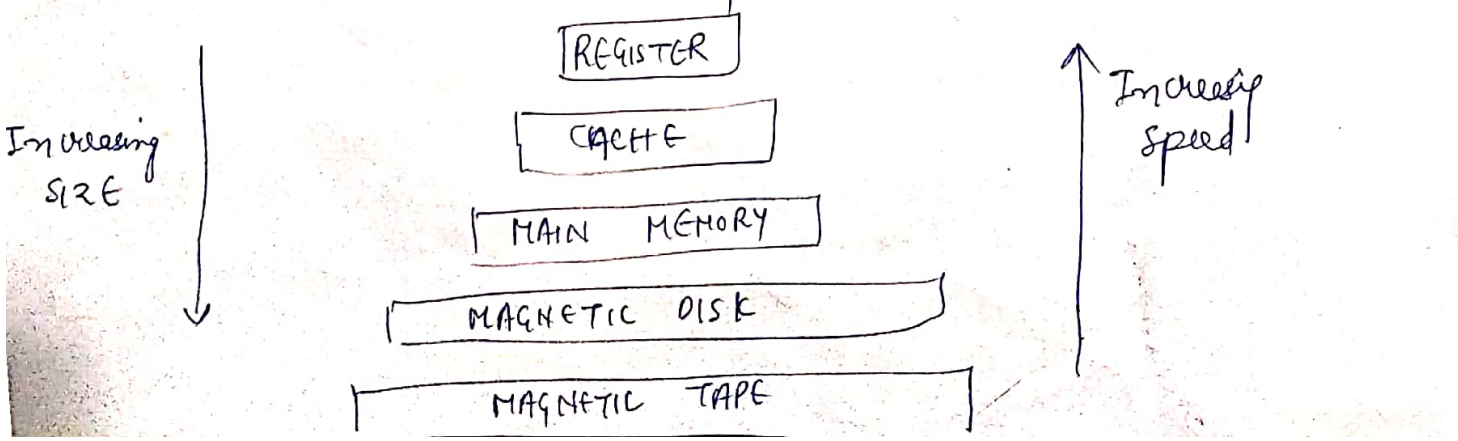
MEMORY HIERARCHY IN A COMPUTER SYSTEM

MAIN MEMORY – is the central storage unit in a computer system.
Large & fast memory used to store Programs & data.

- Technology used for the main memory is based on Semiconductor Integrated circuits.

RAM (volatile memory)

STATIC	DYNAMIC
• SRAM uses transistors to store single bit of data	• DRAM uses capacitor to store each bit of data
• SRAM structure is complex & expensive as compared to DRAM	• DRAM is simpler & less expensive than SRAM
• SRAM is faster than DRAM	• DRAM is slower than SRAM.
• SRAM used in cache memory	• DRAM used in main memory.

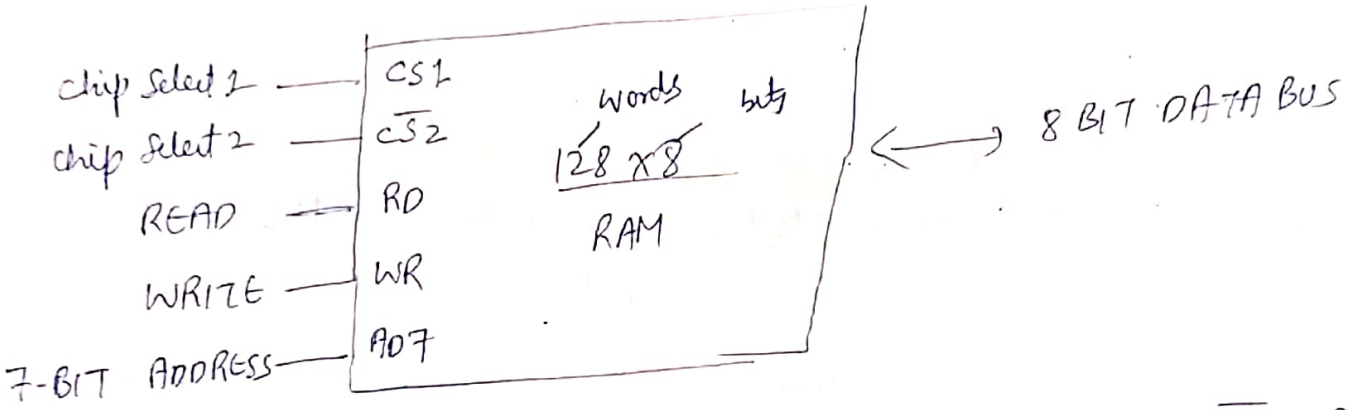


N – Non volatile memory

Needed for storing an initial program called Bootstrap loader

. Bootstrap loader is a program whose function is to start the computer S/W operating system when Power is turned on.

BLOCK DIAGRAM OF RAM CHIP



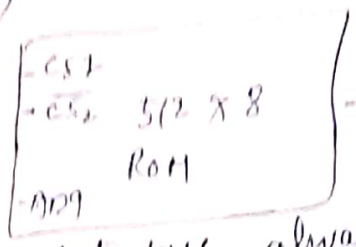
- The unit is in operation only when CS1 = 1 & $\overline{CS2}$ = 0
- AD specifies the address or location to read from or write to

CS1	$\overline{CS2}$	RD	WR	MEMORY FUNCTION	State of data bus
0	0	X	X	Inhibit (Prevent)	High Impedence
0	1	X	X	Inhibit	"
1	0	0	0	Inhibit	"
1	0	0	1	write	Input data to RAM
1	0	1	X	Read	O/P data from RAM
1	1	X	X	Inhibit	High Impedence

when CS1 = 1 & $\overline{CS2}$ = 0, the memory can be placed in a write or read mode

ROM CHIP BLOCK DIAGRAM

$2^9 = 512$ (9 bit addressing)



9 bit address → 8 Bit data bus

- ROM can only read, data bus always in O/P mode

RAM CHIP DESCRIPTION

- RAM capacity = 128 words of 8 bits each. So for 128 (2^7) words, 7 bits for addressing & 8 bit data bus
- Multiple Select lines to select the chip when multiple chips are available in micro computer
- The chip is in operation when CS1 = 1 & $\overline{CS2} = 0$. Bar on top of 2nd select indicates that the input is enabled when 0
- Bus in High Impedence State : when Sel lines are not enabled/ read - write inputs are not enabled.
- When WR input is enabled, data from data bus is stored in location specified by address bus
- When RD I/P is enabled, the selected byte is placed on the data bus

MEMORY CONNECTION TO CPU

- RAM & ROM chips are connected to a CPU through the data & address buses.
- Memory configuration : 512 bytes RAM + 512 bytes ROM = 1024 BYTE

  1 × 512 BYTES ROM + 4 × 128 BYTES RAM

  $2^{10} = 1024$ → 10 address line used

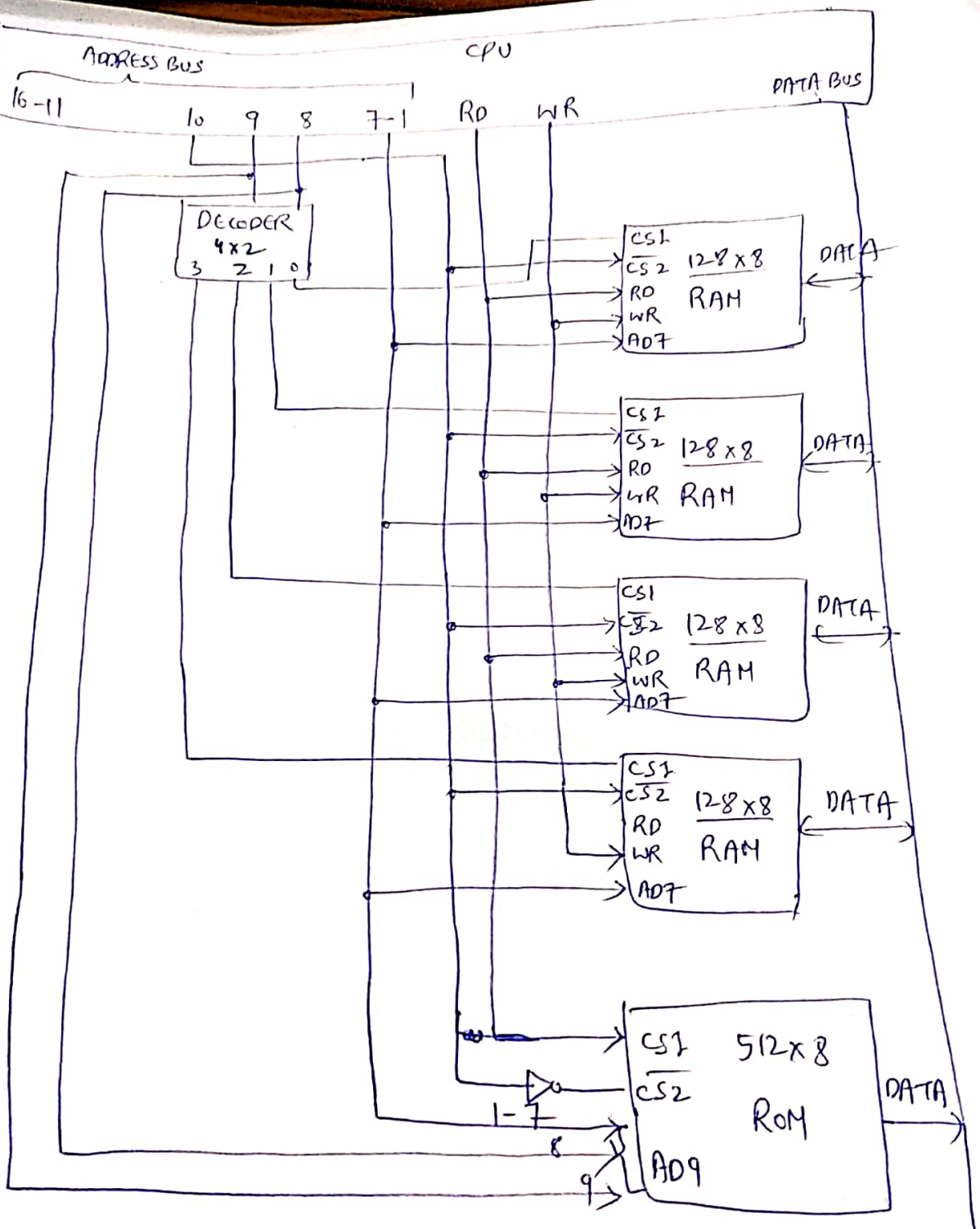
- Memory address Map
  - Address lines 9 8

0	0 — RAM 1
0	1 — RAM 2
1	0 — RAM 3
1	1 — RAM 4

  - Address line 10 ROM 1

- MEMORY CONNECTION TO CPU
  - 2 x 4 decoder : RAM select (CS 1)
  - Address line 10
    - RAM SELECT : $\overline{CS2}$ - 0
    - ROM SELECT : $\overline{CS2}$ - 1

10	9	8	RAM
0	0	0	— RAM 1
0	0	1	— 2
0	1	0	— 3
0	1	1	— 4
1	1	1	— ROM



MEMORY CONNECTION TO THE CPU