ELEC 263 COMPUTER ARCHITECTURE AND ORGANIZATION

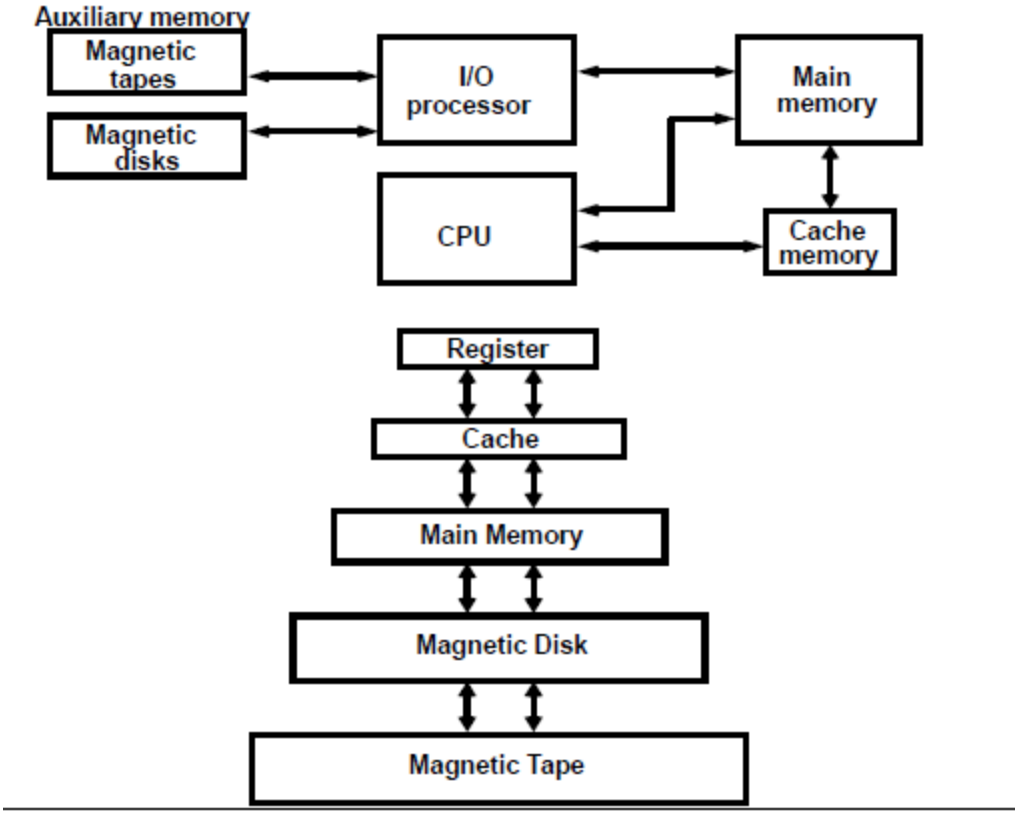
Dr. Mohamed Al-Meer
485-2899
almeer@qu.edu.qa, almeerqatar@gmail.com

CHAPTER 12:
MEMORY ORGANIZATION

12-1 Memory Hierarchy

- The memory unit that communicates with directly with CPU is called main memory.
    - Not enough storage space.
    - SRAM or DRAM
    - Contains programs and data currently needed are stored here
- Devices that provide backup storage are called auxiliary memory.
    - Most common are magnetic disks and tape drives.
    - Used to store system programs, large data files, backup data
    - Not urgently needed data are stored here.
- Total memory capacity of a computer can be visualized as a hierarchy of components.
    - Ranges from 1- slow but high capacity to 2- relatively faster and less capacitive main memory to 3- smaller and fast cache memory.
    - At bottom of hierarchy you can find slow magnetic tapes(removable files)
    - at middle you can find magnetic disks (backup storage)
    - then main memory which can communicate with CPU and auxiliary memories.
    - At top of pyramid resides the cache memory. Used to increase speed of processing. Compensate of speed difference between CPU and main memory. Usually in cache segments of programs and data frequently

accessed are stored to benefit from high speed access by CPU not found in main memory.

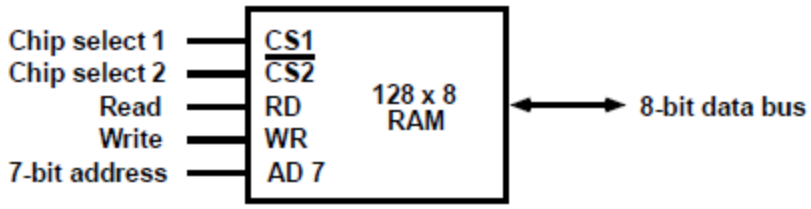


12-Main Memory

- Main memory is central storage unit in computers
- Fast and relatively large type of memory and used to store programs and data used in program execution
- RAM: integrated circuit chips (Static or Dynamic)
    - SRAM consists of flip flops as storage media. Stored data remains valid as long as power is applied to the unit
    - DRAM stores data as form of electric charges in small capacitors. Capacitors are provided by CMOS transistors. Needs refreshing periodically as charges on small capacitor discharge soon (need electronic control unit for that).
    - DRAM compared to SRAM offer reduced power consumption and larger capacity. But SRAM are faster.
- ROM is different type of main memories. Used to store programs and data that does not change at all (programs, tables, etc.)
    - Used ROMs for storing bootstrap loaders (start loading operating systems).
    - ROMs are used to startup any computer.

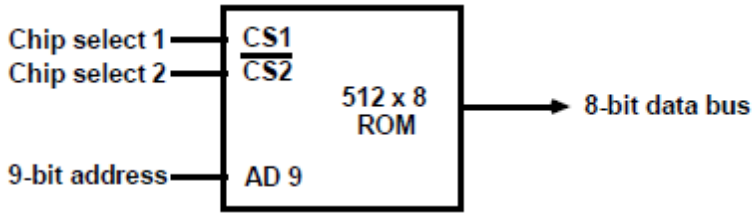
- ROM and RAM are available in different sizes. And usually we have to combine many chips to increase size.

- RAM chips consist of a number of address pins, bidirectional data pins, and some control pins.
    - Bidirectional means data can be initiated from memory to the outside or from outside to memory chip. (Constructed from 3 states buffers?).
    - Next figure shows 128 words of 8 bits per word RAM chip. This requires 7 address bits and 8 bidirectional data bits.
    - Next figure shows a RAM chip that have 2 chip enables CS1=1 and CS2#=0. Those called chip select controls and they enable the chip to function if selected.
    - Another 2 controls RD and WR. When R=1 then chip in a read status such that it provided data into data pins from location specified by address. If WR=1 then it is in write status and it accepts data from data bits into the chip's storage location specified by address.



CS1	CS2	RD	WR	Memory function	State of data bus
0	0	x	x	Inhibit	High-impedence
0	1	x	x	Inhibit	High-impedence
1	0	0	0	Inhibit	High-impedence
1	0	0	1	Write	Input data to RAM
1	0	1	x	Read	Output data from RAM
1	1	x	x	Inhibit	High-impedence

- ROM chips are constructed the same way. It has address bits, one direction data bits, CS control pin and/or RD control pin.
    - Next figure shows 512 words ROM chip each with 8 data bits. See how address pins and storage locations are related.

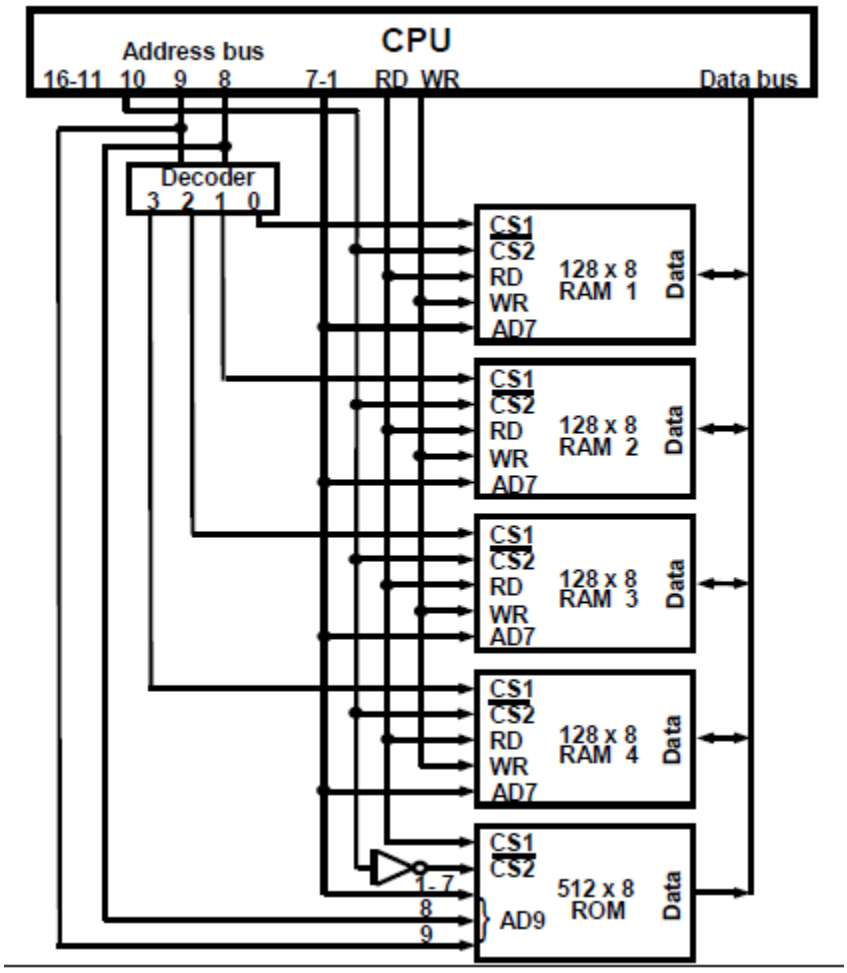


    - Internal binary cells of ROM occupy much less space than RAM.

Memory Address Map and Connection to CPU

- Memory Address Map is “address space assignment to each memory chip”.
- Assume computer system with 512 bytes of RAM and 512 bytes of ROM as shown in next figure. The memory address map is shown in next table.

Component	Hexa address	Address bus									
		10	9	8	7	6	5	4	3	2	1
RAM 1	0000 - 007F	0	0	0	x	x	x	x	x	x	x
RAM 2	0080 - 00FF	0	0	1	x	x	x	x	x	x	x
RAM 3	0100 - 017F	0	1	0	x	x	x	x	x	x	x
RAM 4	0180 - 01FF	0	1	1	x	x	x	x	x	x	x
ROM	0200 - 03FF	1	x	x	x	x	x	x	x	x	x



- RAM chips are 128 words each so they need 7 address lines.
- ROM chips are 512 words each so they need 9 address lines.
- For RAM selection A10 is always 0 whereas for ROM selection A10 is always 1.
- Memory chips are connected to CPU through address and data busses. RAM chips stores 128 by 4 = 512 bytes. ROM chip stores 512 bytes alone.
- A10 selects ROM chip if it is 1. And it selects all RAM chips if it is 0.
- RD and WR control pins are connected to RD and WR pins in corresponding ROM and RAM chips to initiate read or write operation from CPU.
- Address range from 0 to 511 is assigned for RAM chips while address range from 512 to 1023 is assigned to ROM chip.