



Computer Architecture

Spring 2020

Hamed Farbeh

farbeh@aut.ac.ir

Department of Computer Engineering

Amirkabir University of Technology

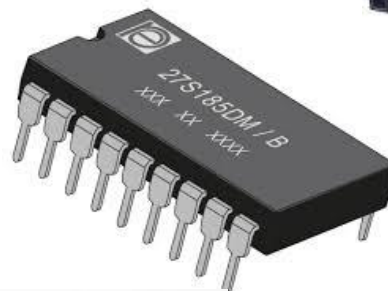


Copyright Notice

Lectures adopted from

- Computer Organization and Design: The Hardware/Software Interface, 5th edition, David A. Patterson, John L. Hennessy, MK pub., 2014
 - Chapter 5: Large and Fast: Exploiting Memory Hierarchy

Memories



Memories Criteria

- Speed (Read and Write)
- Capacity
- Volatility
- Power/Energy (Read and Write)
- Endurance
- Retention Time
- Reliability
- Compatibility
- ...

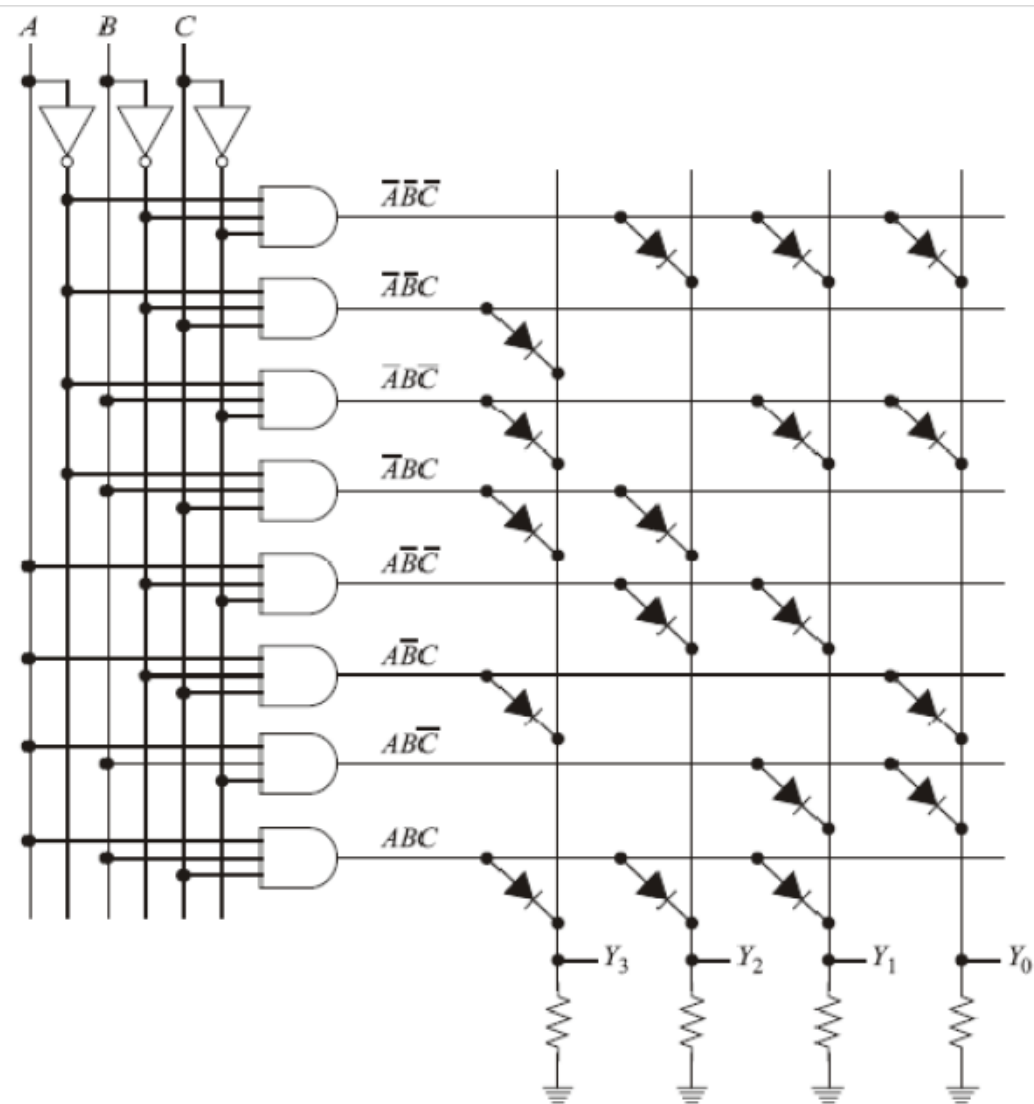
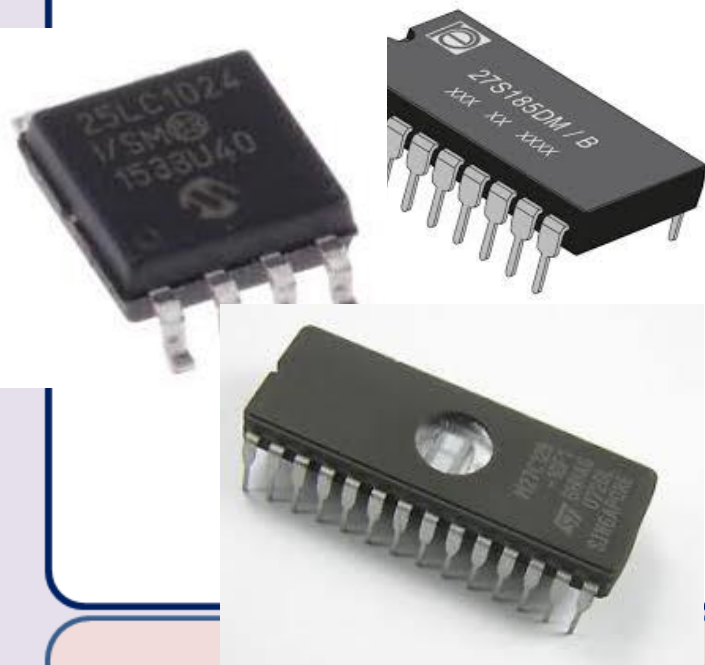


Memory Technology

- Static RAM (SRAM)
 - 0.5ns – 2.5ns, \$2000 – \$5000 per GB
- Dynamic RAM (DRAM)
 - 50ns – 70ns, \$20 – \$75 per GB
- Magnetic disk
 - 5ms – 20ms, \$0.20 – \$2 per GB
- Ideal memory
 - Access time of SRAM
 - Capacity and cost/GB of disk

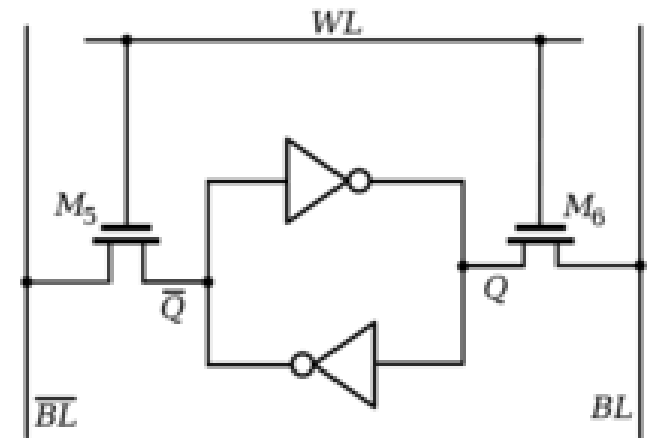
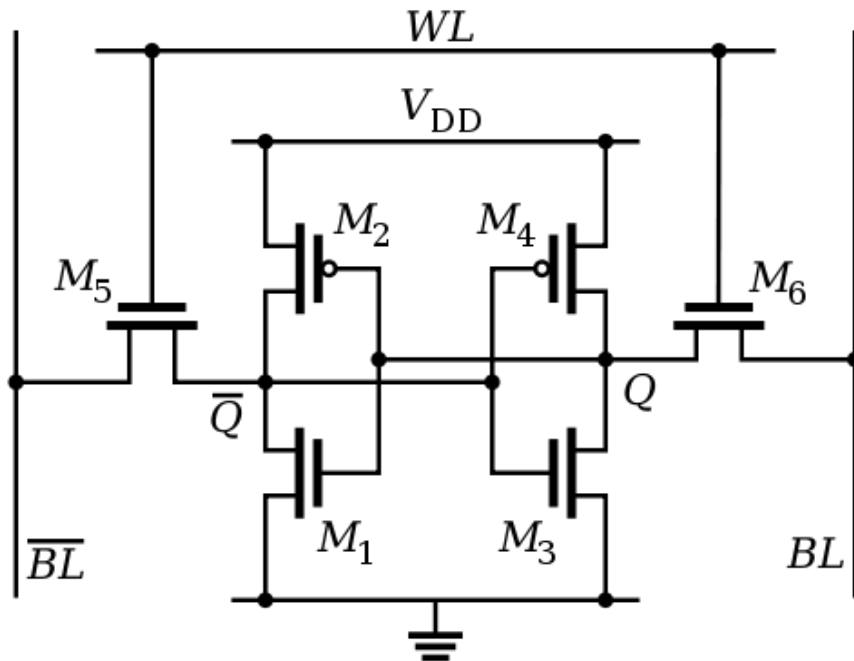
Memory Technology

- ROM
- PROM
- EPROM
- EEPROM



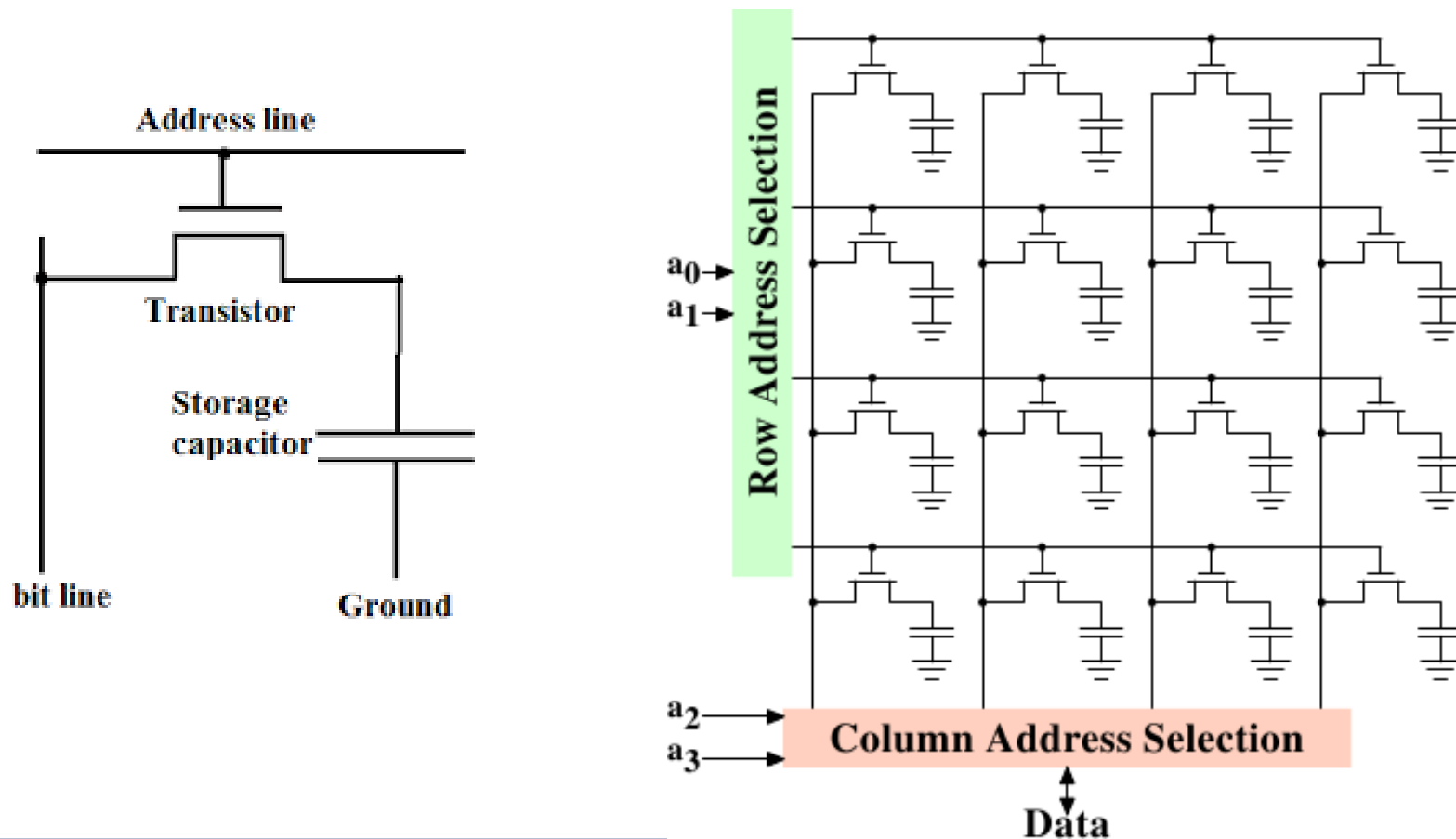
Memory Technology

■ Static RAM (SRAM)



Memory Technology

■ Dynamic RAM (DRAM)



Principle of Locality

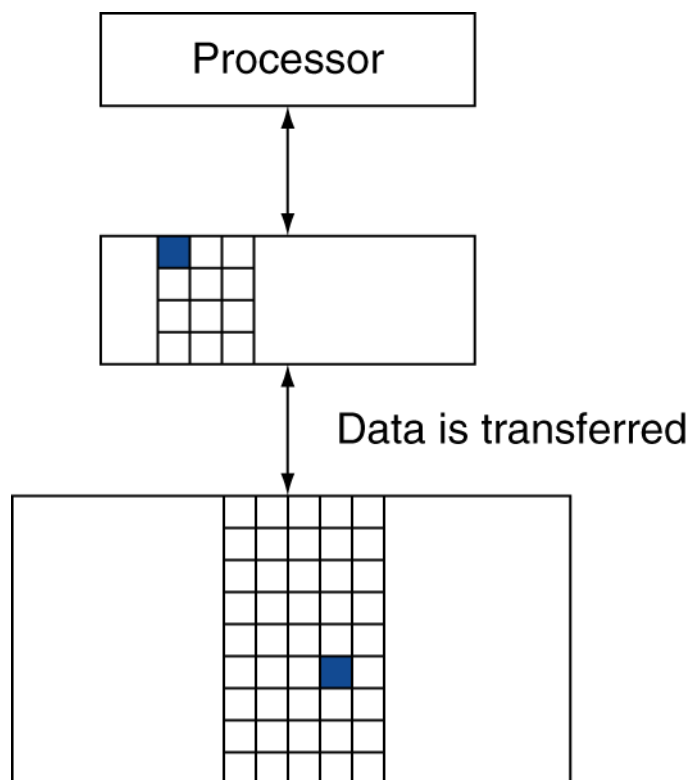
- Programs access a small proportion of their address space at any time
- Temporal locality
 - Items accessed recently are likely to be accessed again soon
 - e.g., instructions in a loop, induction variables
- Spatial locality
 - Items near those accessed recently are likely to be accessed soon
 - E.g., sequential instruction access, array data

Taking Advantage of Locality

- Memory hierarchy
- Store everything on disk
- Copy recently accessed (and nearby) items from disk to smaller DRAM memory
 - Main memory
- Copy more recently accessed (and nearby) items from DRAM to smaller SRAM memory
 - Cache memory attached to CPU

Memory Hierarchy Levels

- Block (aka line): unit of copying
 - May be multiple words
- If accessed data is present in upper level
 - Hit: access satisfied by upper level
 - Hit ratio: hits/accesses
- If accessed data is absent
 - Miss: block copied from lower level
 - Time taken: miss penalty
 - Miss ratio: misses/accesses
 $= 1 - \text{hit ratio}$
 - Then accessed data supplied from upper level



Memory Hierarchy Levels

