BCN1043

COMPUTER ARCHITECTURE & ORGANIZATION

By
Dr. Mritha Ramalingam

Faculty of Computer Systems & Software Engineering
mritha@ump.edu.my

AUTHORS

- **Dr. Mohd Nizam Mohmad Kahar** (mnizam@ump.edu.my)
- **Jamaludin Sallim** (jamal@ump.edu.my)
- **Dr. Syafiq Fauzi Kamarulzaman** (syafiq29@ump.edu.my)
- **Dr. Mritha Ramalingam** (mritha@ump.edu.my)

Faculty of Computer Systems & Software Engineering

Learning outcomes

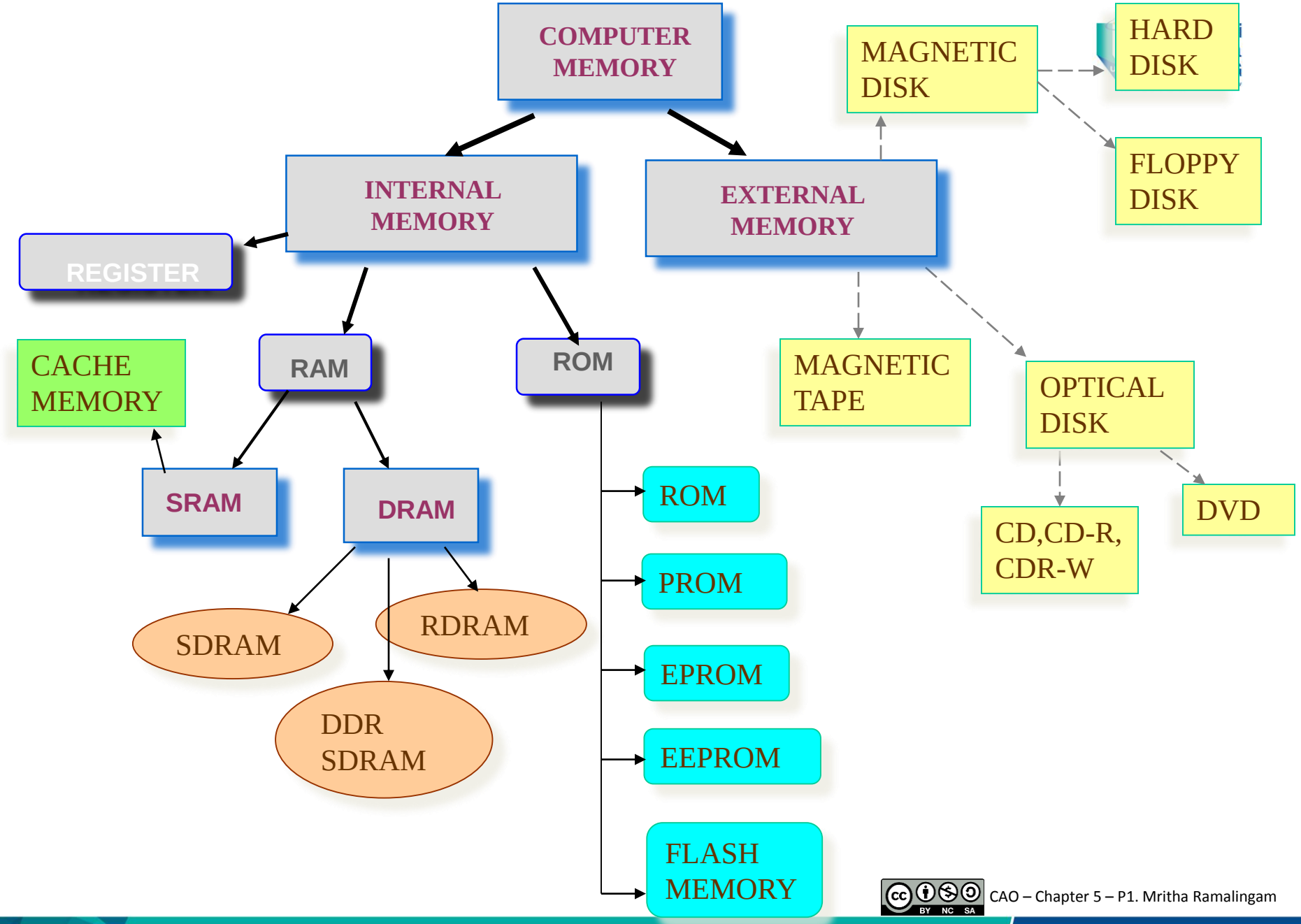
- Able to identify the main types of memory technology (e.g., SRAM, DRAM, Flash, magnetic disk) and their relative cost and performance.
- Able to describe the principles of memory management and how the use of memory hierarchy (cache, virtual memory) is used to reduce the effective memory latency.
- Able to describe the reason for and use of cache memory
- Able to explain the effect of memory latency on running time.

COMPUTER MEMORY

1. Storage System & Technology
2. Memory Hierarchy
3. Memory Organization and Operations
4. Cache Memories

Storage System & Technology

1. **Location**
2. **Capacity**
3. **Unit of Transfer**
4. **Access method**
5. **Performance**
6. **Physical type**
7. **Physical characteristics**
8. **Organization**


COMPUTER MEMORY
INTERNAL MEMORY
EXTERNAL MEMORY
REGISTER
CACHE MEMORY
RAM
ROM
SRAM
DRAM
SDRAM
RDRAM
DDR SDRAM
ROM
PROM
EPROM
EEPROM
FLASH MEMORY
MAGNETIC DISK
HARD DISK
FLOPPY DISK
MAGNETIC TAPE
OPTICAL DISK
CD,CD-R, CDR-W
DVD

Storage System & Technology

1.Location: Based on location,

memory can be classified as

- CPU memory
- Internal memory
- External memory

2.Capacity: Based on the amount of data that can be stored in memory

1.Word

2.bytes

3.Unit of transfer: number of bits per time period

4. Based on Access Method

1. Random access:
   - Each location - unique physical address
   - Any order access
   - E.g main memory
2. Sequential access:
   - no unique address
   - Read in a sequence
   - High variable Access times
   - E.g tape drive

3. Direct access

- unique addresses
- E.g disk drives

4. Associative access

- A variation of RAM
- Accessing based on their contents instead of their location
  - highly fast
- E.g few cache memory

5. Performance

- Access time
- Memory cycle time
- Transfer rate

6. Physical Type

- Semiconductor memory or RAM
- Magnetic Disk & Tape
- Optical disk

7. Physical Characteristics

- Volatile and non-volatile
- Erasable and nonerasable

8. Organisation

- Physical arrangement of bits to form words

Random Access Memory - RAM

- **SRAM** (Static RAM) and **DRAM** (Dynamic RAM).
- **Static RAM (SRAM)** holds its data without external refresh, as long as electricity is supplied to the circuit.
- **Dynamic RAM (DRAM)** must be refreshed many times per second in order to hold its data contents.
- Strength of SRAMs:
  - **Simplicity:** SRAMs don't require external refresh circuitry or other work in order for them to keep their data intact.
  - **Speed:** SRAM is faster than DRAM.
- Weakness of SRAMs:
  - **Cost:** SRAM is, byte for byte, several times more expensive than DRAM.
  - **Size:** SRAMs take up much more space than DRAMs (which is part of why the cost is higher).

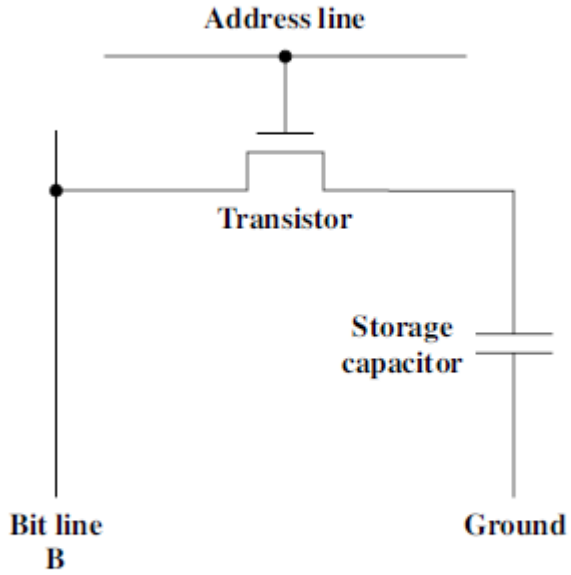
DRAM vs. SRAM

Dynamic RAM

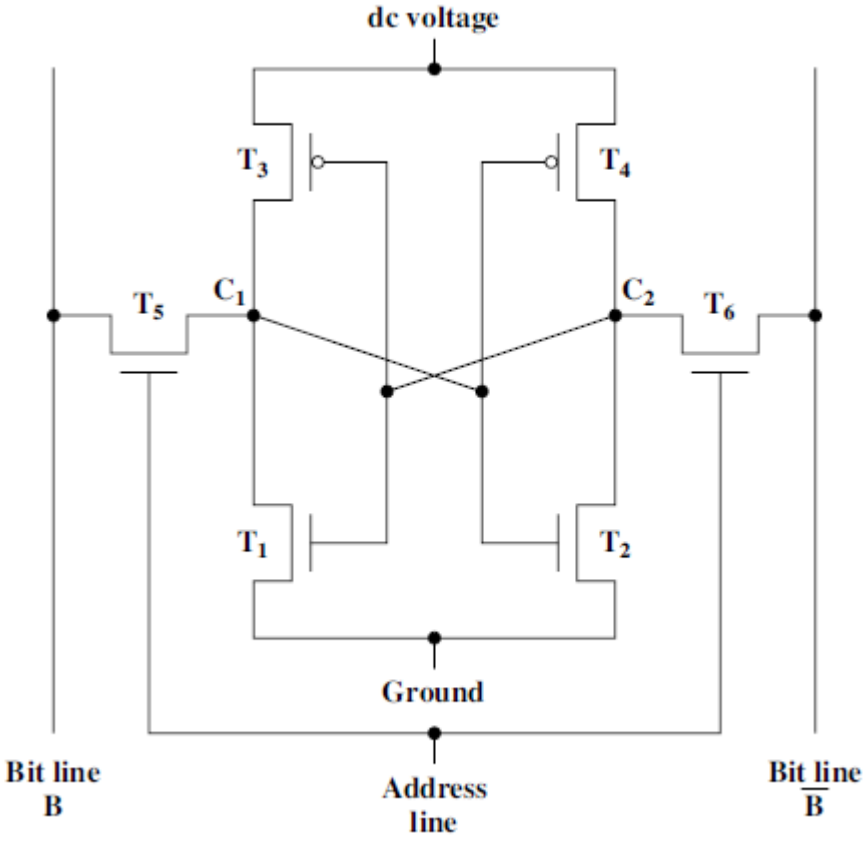
- Bits stored as charge in capacitors
- Charges leak
- Need refreshing even when powered
- Simpler construction
- More dense (smaller cell-more cell per unit area)
- Less expensive
- Need refresh circuits
- Slower access time
- Low power
- Application: Main memory

Static RAM

- Bits stored as on/off switches
- No charges to leak
- No refreshing needed when powered
- More complex construction
- Less dense (less cell per unit area)
- More expensive
- Does not need refresh circuits
- Faster access time
- Consumes more power
- Application: Cache memory



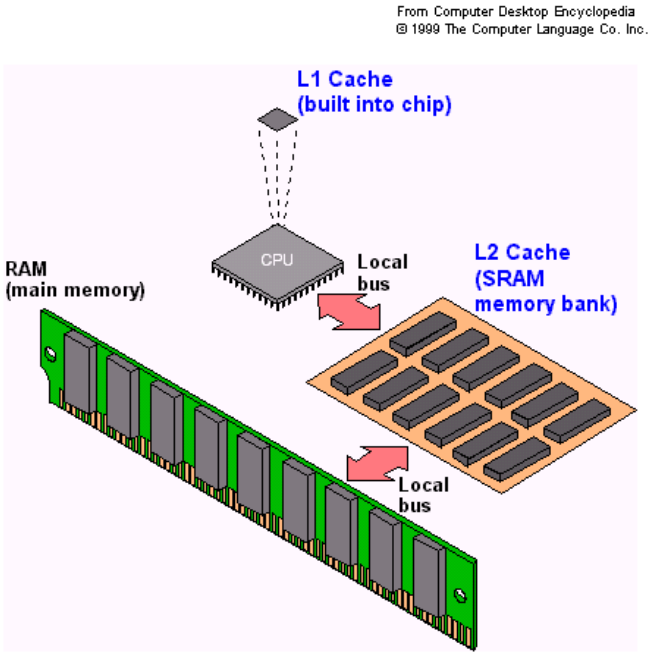
(a) Dynamic RAM (DRAM) cell



(b) Static RAM (SRAM) cell

RAM (cont…)

- DRAM is used for system memory.
- SRAMs are used for level 1 cache and level 2 cache memory, for which it is perfectly suited (as the cache memory needs to be very fast, and not very large).



Source: http://images.slideplayer.com

Read-only memory - ROM

- There are five basic ROM types:
  - Read-only memory (ROM)
  - Programmable ROM (PROM)
  - Erasable PROM (EPROM)
  - Electrically EPROM (EEPROM)
  - Flash memory
- Each type has unique characteristics, but they have two things in common:
  - **nonvolatile** → it is not lost when electricity is removed.
  - Data stored in these chips is either **unchangeable** or requires a special operation to change (unlike RAM, which can be changed as easily as it is read).

ROM (cont…)

- **ROM** - cannot reprogram or rewrite. Creating ROM chips is time-consuming and very expensive in small quantities.
- **PROM** - can be written once. chips can be bought inexpensively and coded using a special tool called a **programmer**.
- **EPROM** - can be rewritten many times. Erasable using ultraviolet (UV) light.
- **EEPROM**
  - Changing the contents does not require additional dedicated equipment.

ROM (cont…)

- Electric field can return the electrons in the cells of an EEPROM to normal. This erases the targeted cells of the EEPROM, which can then be rewritten. EEPROMs are changed 1 byte at a time, which makes them versatile but slow.
- Hence, **Flash memory**, a type of EEPROM that uses **in-circuit wiring** to erase by applying an electrical field to the entire chip or to predetermined sections of the chip called **blocks**. Flash memory works much faster than traditional EEPROMs because it writes data in chunks, usually 512 bytes in size, instead of 1 byte at a time.

- **Characteristics of ROM**
  - Permanent type of memory storage.
  - It is nonvolatile. (Data will not lose if the power is turn off)
  - Data stored in ROM is either unchangeable or required a special operation to change it.
  - Example of application in computer: BIOS

- **Types of ROM**
  - ROM
  - PROM
  - EPROM
  - EEPROM
  - Flash Memory

- ROMs
  - ➔Data is “wired in” during fabrication at a chip manufacturer’s plant
  - ➔Permanent pattern of data that cannot be change.
  - ➔Not possible to write new data into it.

- PROMs
  - ➔Programmable ROM
  - ➔Data can be written once by the user employing a PROM programmer
  - ➔Useful for small production runs
- EPROM
  - ➔Erasable PROM
  - ➔Programming is similar to a PROM
  - ➔Can be rewritten many times.
  - ➔Erase using ultraviolet (UV) light.

- EEPROM
  - Electrically erasable PROMs - It contents are erased by applying a specific electrical voltage
  - Can be written to many times while remaining in a system
  - Does not have to be erased first
  - Program and erase by individual bytes
  - Used in systems for development, personalization, and other tasks requiring unique information to be stored
- Flash Memory
  - Similar to EEPROM in using electrical erase
  - Fast erasures, erase by block
  - Higher density than EEPROM

Magnetic Disk

- The disk is a circular platter constructed of nonmagnetic material called the substrate, coated with magnetizable material.
- Data is recorded onto and later read from the disk using a conducting coil, the head.
- Substrate used to be aluminium.
- Now glass
  - Improved surface uniformity
    - Increases reliability
  - Reduction in surface defects
    - Reduced read/write errors
  - Better stiffness (maintain data)
  - Better shock/damage resistance (avoid surface damage)

Physical Characteristics

- Fixed(rare) or movable head
- Removable or fixed
- Single or double (usually) sided
- Single or multiple platter

Fixed/Movable Head Disk

Fixed head

- One read write head per track
- Heads mounted on fixed ridged arm

Movable head

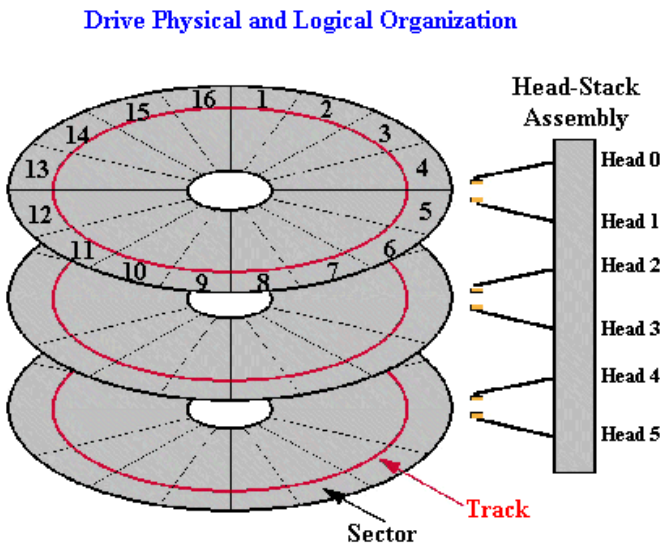
- One read write head per side
- Mounted on a movable arm

Multiple platter

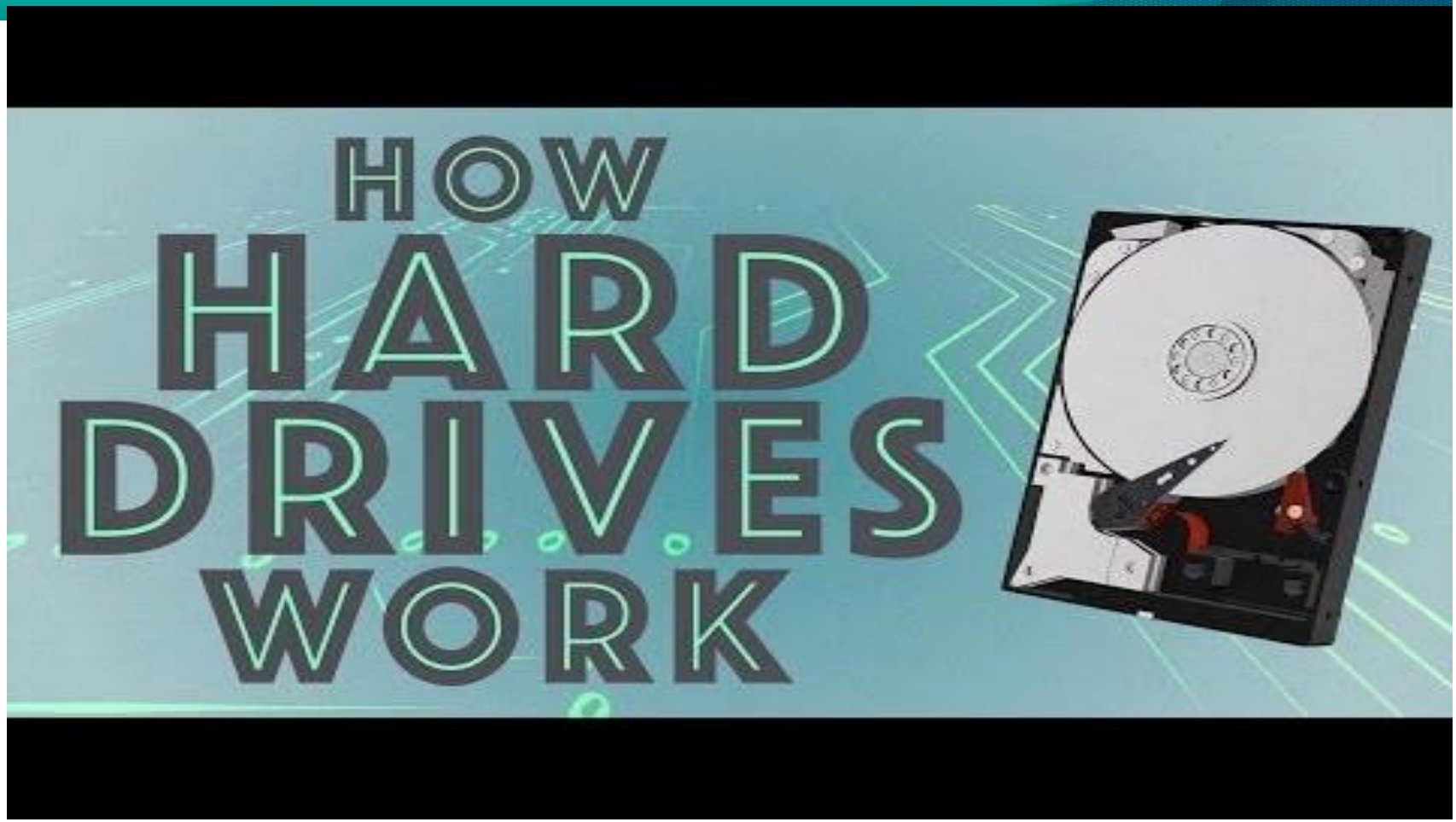
- One head per side
- Heads are joined and aligned
- Aligned tracks on each platter form cylinders
- Data is striped by cylinder
  - ➔reduces head movement
  - ➔Increases speed (transfer rate)

Disk Performance Parameters

- Seek time - the time it takes to position the head at the right track
- Rotational delay - the time it takes for the beginning of the sector to reach the head
- Access time - the sum of the seek time and rotational delay
- Transfer time - time to read the block (sector) from the disk and transfer it to main memory



Source:http://img.watchstor.com



Optical Disk

- Data is stored on an optical medium (such as CD-ROM or DVD), and read with a laser beam.
- It's not practical for use in computer processing, but it is an ideal solution for storing large quantities of data very inexpensively, and more importantly, transporting that data between computer devices.
- Optical disk products:
  - ➔CD-ROM (Compact disk read-only memory)
  - ➔CD-R (CD Recordable)
  - ➔CD-RW (CD Rewritable)
  - ➔DVD (Digital Versatile Disk)

CD-ROM

- Originally for audio
- 650Mbytes giving over 70 minutes audio
- Polycarbonate coated with highly reflective coat, usually aluminum
- Data stored as pits and bumps
- Read by reflecting laser
- Constant packing density
- Constant linear velocity

CD-Writable

- WORM (Write Once, Read Many disks)
- Now affordable
- Compatible with CD-ROM drives

CD-RW

- Erasable
- Getting cheaper
- Mostly CD-ROM drive compatible

DVD

- Digital Versatile Disk
  - Will read computer disks and play video disks
- Multi-layer
- Very high capacity (4.7G per layer)
- Full length movie on single disk
  - Using MPEG compression
- Finally standardized
- Movies carry regional coding
- Players only play correct region films
- Can be “fixed”

Magnetic Tape

- Use the same reading and recording techniques as magnetic disk system.
- Data on the tape are structured as a number of parallel tracks running lengthwise.

- The first kind of secondary memory
- Still widely used
  - very cheap
  - very slow
- Sequential access
  - Data is organized as records with physical air gaps between records
  - One words is stored across the width of the tape and read using multiple read/write heads

Chapter 5

Chapter 5 will continue!