Storage and Disks
General Overview

- Relational model - SQL
  - Formal & commercial query languages
- Functional Dependencies
- Normalization

- Physical Design
  - Indexing
  - Query evaluation
  - Query optimization
  - ....

Application Oriented

Systems Oriented
Storage Media: Types

- **Cache** – fastest and most costly form of storage; volatile; managed by the computer system hardware.

- **Main memory**:
  - fast access (10s to 100s of nanoseconds; 1 nanosecond = $10^{-9}$ seconds)
  - generally too small (or too expensive) to store the entire database (but for some applications, this is changing)
  - **Volatile** — contents of main memory are usually lost if a power failure or system crash occurs.
  - But… CPU operates only on data in main memory
Storage Media: Types (cont.)

- **Disk**
  - Primary medium for the long-term storage of data; typically stores entire database.
  - random-access – possible to read data on disk in any order, unlike magnetic tape
  - Non-volatile: data survive a power failure or a system crash, disk failure less likely than them

- **Flash Memory**
  - no seeks
    - Cheap reads, expensive writes
  - experimental use for DB’s

- **NVM**
  - Coming soon
Traveling the hierarchy:
1. speed (higher=faster)
2. cost (lower=cheaper)
3. volatility (between MM and Disk)
4. Data transfer (Main memory the “hub”)
5. Storage classes (P=primary, S=secondary, T=tertiary)
Hard Disk Mechanism

- track \( t \)
- spindle
- sector \( s \)
- cylinder \( c \)
- platter
- rotation
- arm
- read-write head
- arm assembly
- **Read-write head**
  - Positioned very close to the platter surface (almost touching it)
- Surface of platter divided into circular **tracks**
- Each track is divided into **sectors**.
  - A sector is the smallest unit of data that can be read or written.
- To read/write a sector
  - disk arm swings to position head on right track
  - platter spins continually; data is read/written as sector passes under head
- **Block:** a sequence of sectors
- **Cylinder** $i$ consists of $i^{th}$ track of all the platters
Performance Measures of Disks

Measuring Disk Speed

- **Access time** – consists of:
  - **Seek time** – time it takes to reposition the arm over the correct track.
  - **(Rotational) latency time** – time it takes for the sector to be accessed to appear under the head.

- **Data-transfer rate** – the rate at which data can be retrieved from or stored to the disk.

Analogy to taking a bus:

1. Seek time: time to get to bus stop
2. Latency time; time spent waiting at bus stop
3. Data transfer time: time spent riding the bus
Random vs sequential I / O

- Ex: 1 KB Block
  - Random I/O: ~ 20 ms.
  - Sequential I/O: ~ 1 ms.

Rule of Thumb
- Random I/O: Expensive
- Sequential I/O: Much less ~10-20 times
**Mean time to failure (MTTF)** – the average time the disk is expected to run continuously without any failure.

- Typically 5 to 10 years
- Probability of failure of new disks is quite low, corresponding to a “theoretical MTTF” of 30,000 to 1,200,000 hours for a new disk

  - E.g., an MTTF of 1,200,000 hours for a new disk means that given 1000 relatively new disks, on an average one will fail every 1200 hours

- MTTF decreases as disk ages