

A decorative graphic on the left side of the slide, consisting of a network of white lines and circles on a teal background, resembling a circuit board or a neural network.

# WEEK 3

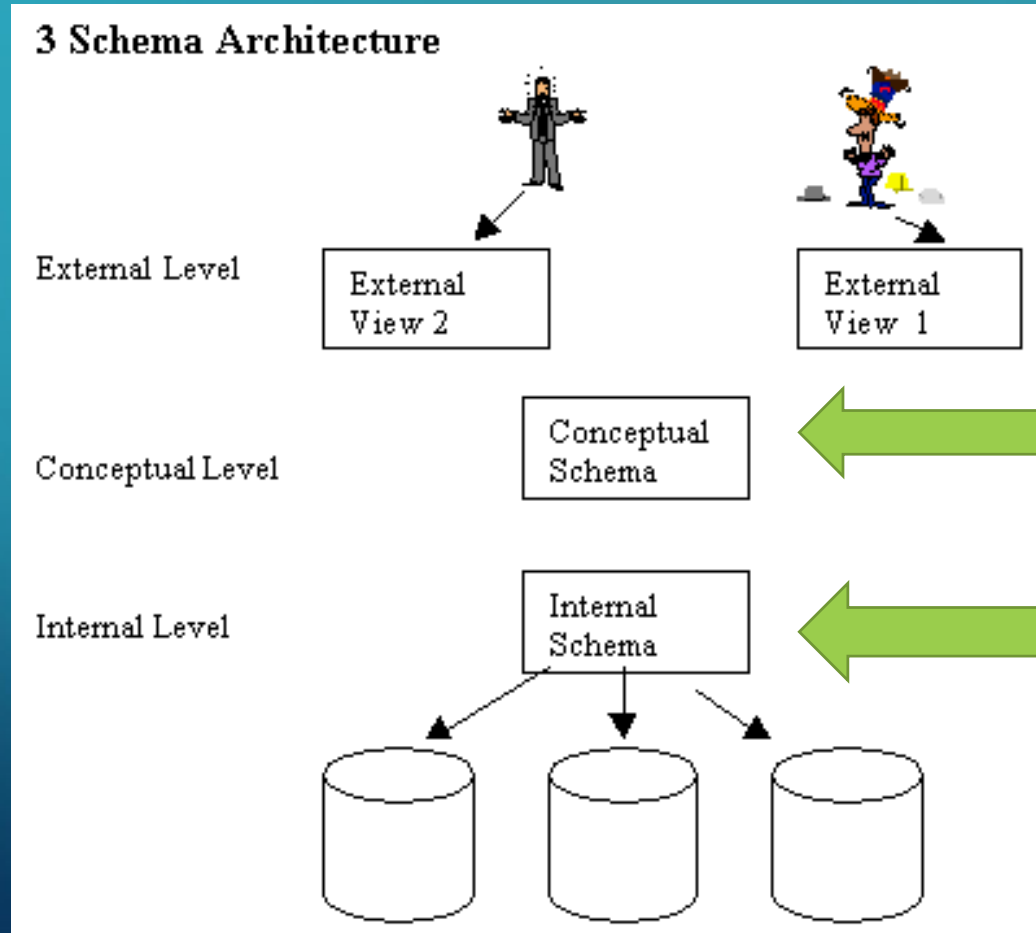
## THE PHYSICAL STORAGE OF A DATABASE

CS3319

# STUDENT OBJECTIVES

- Upon completion of this video, you should be able to:
  - Give 3 reasons why we store the database on secondary storage
  - List at least 4 different storage devices.
  - Give advantages and disadvantages of each device
  - Distinguish between the Primary Organization of the records and the Secondary Organization of the records.

# REMEMBER THE 3-SCHEMA ARCHITECTURE



In week 2, we talked about this level, this is where the relational database resides

NOW we will talk about this level, when we physically write the data onto a storage device

# STORAGE MEDIUM – PHYSICALLY STORING THE DATABASE ON THE COMPUTER

- Primary Storage → CPU (Main Memory and faster Cache memory)

- Secondary Storage

- Magnetic Disks *e.g. hard drives*

- 闪存* • Flash Memory (Solid State Drive SSD)

- Tertiary Storage

- Optical Disk (CD-ROM, DVDs) (declining because of decrease in cost and increase in capacity of magnetic disks)
  - Tapes

# THE PROBLEM:

- Why is a database normally stored on a hard disk rather than in faster main memory? THERE ARE THREE MAIN REASONS, what are they?
  - **Cost** : ram is expensive
  - **Volatility** : if power down, data may lost
  - **Usually an entire database is too large to fit into main memory**  
main memory is far too small.

# STORAGE OF DATABASES

- Typically on a **hard drive**
  - Advantage – Online (available at anytime)
  - Disadvantage – Slowish, Expensivish
- Also, still need some **offline** storage such as **tapes** but quite slow → (start at minute 2) how offline storage works with tapes and autoloaders (NASA uses tapes to archive databases)
  - Advantage – Cheap, *light, and unemitted.*
  - Disadvantage – Slow
- Also **flash memory (SSD)**
  - Advantage – Fast
  - Disadvantage – Have to overwrite the whole block, limited number of writes
- Which is better? <https://www.youtube.com/watch?v=YQEjGKYXjw8&t=115s>

# PHYSICAL DATABASE DESIGN

- Picking an organization for the data on the hard disk to maximize efficiency!
- Must decide on
  - 1. THE PRIMARY FILE ORGANIZATION:
    - How the file records are physically placed on the disk (ordering of the primary attribute)?
    - How the records can be accessed?
  - 2. SECONDARY ORGANIZATION
    - How can we efficiently access OTHER attributes than the primary attribute?