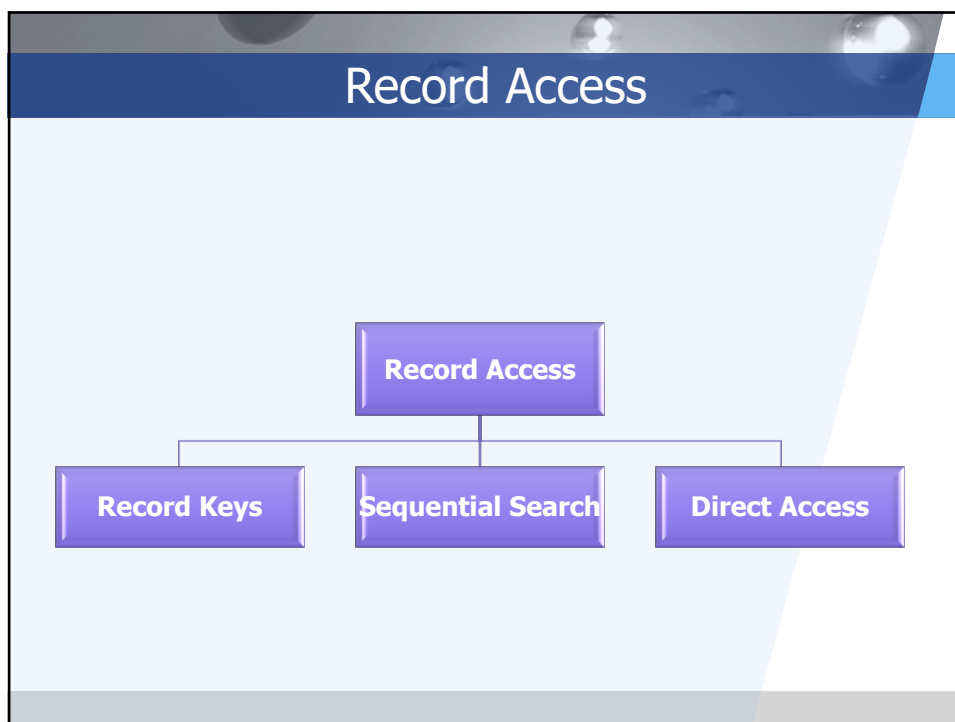


## Managing Files of Records

Lecture No. 4

7

7



8

## 1- Record Keys

- ❖ **Key:** a subset of the fields in a record used to uniquely identify the record.
- ❖ **Primary Key:** A key that uniquely identifies a record.
- ❖ **Secondary Key:** Other keys that may be used for search
- ❖ In general not every field is a key
- ❖ Keys correspond to fields, or combination of fields, that may be used in a search

9

## 1- Record Keys

Primary Key	Secondary Key
must identify records uniquely	Does not identify records uniquely
It is not dataless	It is not dataless
Has a canonical form	Has a canonical form
Ex. Student ID	Ex. Student Name

10

## 2- Sequential Search

- ❖ Search for a **record matching a given key**
- ❖ Look at records **sequentially** until **matching record** is found.
- ❖ The work required to search sequentially for a record in a file with **n records** is proportional to **n**: It takes at most **n comparisons**; **n/2** on average.
- ❖ Time is in  **$O(n)$**  for n records.

11

## 2- Sequential Search

- ❖ To **improve the performance of sequential search**, use **record blocking**
- ❖ **Record Blocking**: by reading in a block of several records all at once and then processing that block of records in memory.

12

## 2- Sequential Search

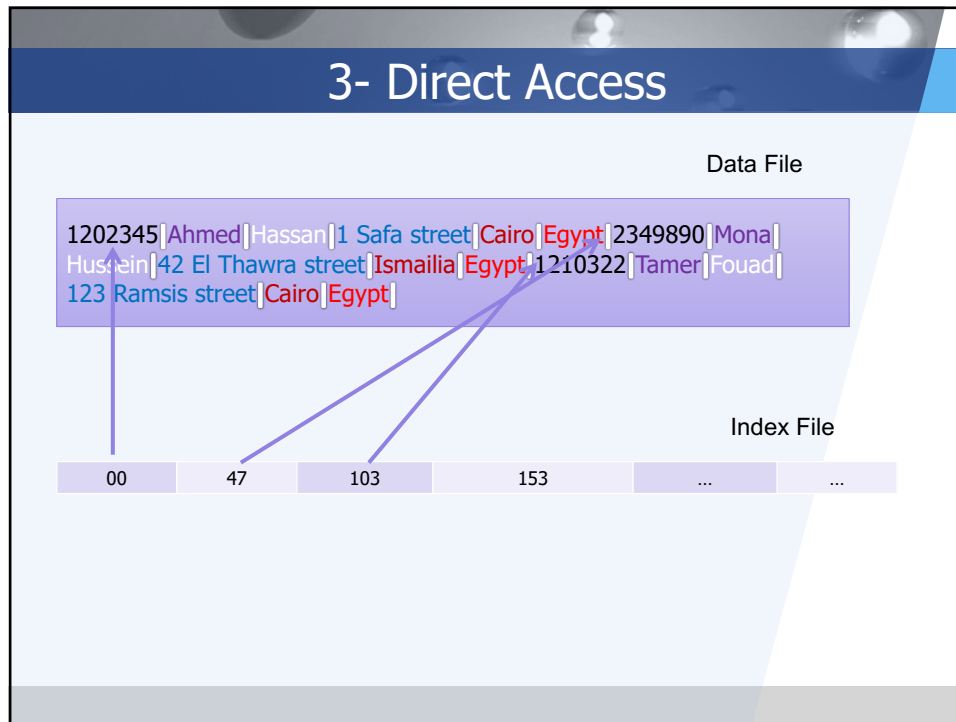
- ❖ It is appropriate for :
  - ASCII files in which you are searching for some pattern.
  - Files with few records (ex. 10 records)
  - Files that hardly ever need to be searched (ex. Tape files)
  - Files in which you want all records with a certain secondary key value, where a large number of matches is expected.

13

## 3- Direct Access

- ❖ Being able to seek directly to the beginning of the record.
- ❖ Time is in  $O(1)$  for n records.
- ❖ Possible when we know the Relative Record Number (RRN)
- ❖ First record has RRN 0, the next has RRN 1, etc.

14



15

### 3- Direct Access

- ❖ Requires records of **fixed length**.
  - RRN=30 (31st record)
  - Record length = 101 bytes
  - Byte offset =  $30 \times 101 = 3030$
- ❖ Now, how to go directly to the byte 3030 in the file
  - By seeking

16

## Header Record

- ❖ A record placed at the **beginning** of a file that is used to **store information about the file contents** and the **file organization**.
- ❖ **Ex:** the length of data records, the date and time of the file's most recent update, the name of the file, and so on.
- ❖ The header record usually has a **different structure & different size** than the data records in the file.