UNIVERSITAS INDONESIA
Veritas, Probitas, Iustitia
FAKULTAS ILMU KOMPUTER

An Overview to File Organization

CSF2600700 - BASIS DATA

Acknowledgements

This slide is a modification to supplementary slide of
“Database System”, 6th edition, Elmasri/Navathe, 2011: **Chapter 16 Disk Storage, Basic File Structures, Hashing, and Modern Storage Architectures** and Unit 2 Storage and file structure slides used in Basis Data Lanjut course in Fakultas Ilmu Komputer, Universitas Indonesia

File Organization: An Overview

A **record** representing an entity of Employee with name = 'Aaron, Ed'

The database is stored as a **collection of files**. Each file is a **sequence of records**. The illustration shows an example for a sequential/ordered file.

A **file** representing a relation named EMPLOYEE

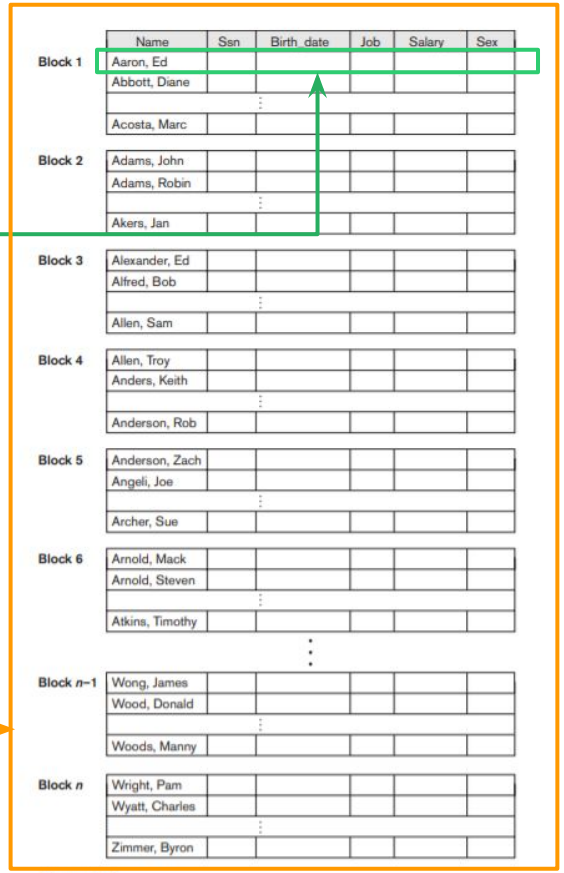

Figure 16.7
Some blocks of an ordered (sequential) file of EMPLOYEE records with Name as the ordering key field.

File Organization: An Overview

Each record consists of a collection of related data values or items, where each value is formed of one or more bytes and corresponds to a particular field of the record. Records usually describe **entities** and their **attributes**.

Fixed-length records

If every record in the file has exactly the **same size** (in bytes)

Variable-length records

If different records in the file have **different sizes**

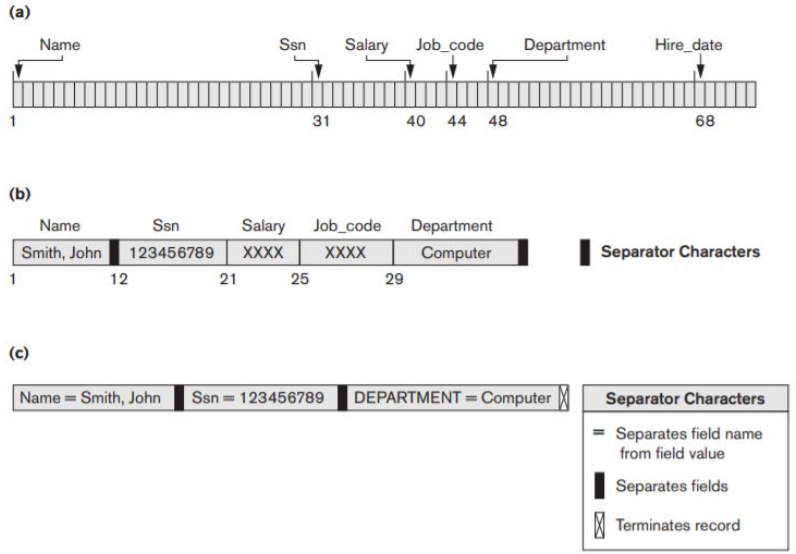

Figure 16.5
Three record storage formats. (a) A fixed-length record with six fields and size of 71 bytes. (b) A record with two variable-length fields and three fixed-length fields. (c) A variable-field record with three types of separator characters.

File Organization: An Overview (Cntd.)

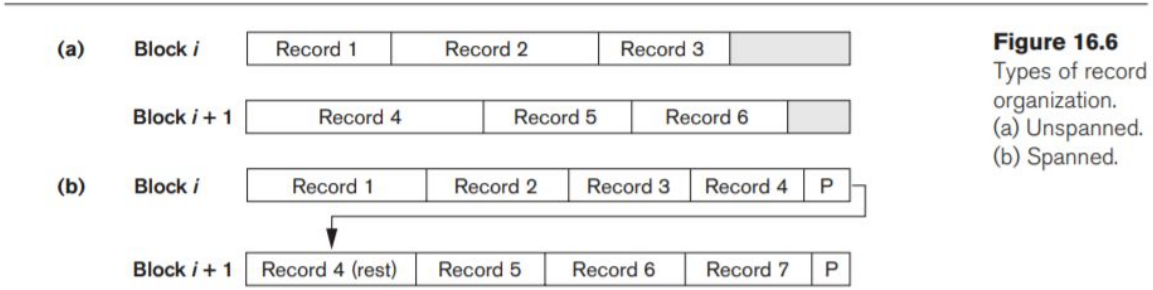

Figure 16.6
Types of record organization.
(a) Unspanned.
(b) Spanned.

The records of a file **must be allocated to disk blocks** because a block is the unit of data transfer between disk and memory.

When the block size is larger than the record size, each block will contain numerous records, although some files may have unusually large records that cannot fit in one block (must use spanned record organization)

Unspanned

If records are not allowed to cross block boundaries

Spanned

If records can span more than one block

Some Primary Files Organizations

Records of each relation **may be stored in a separate file**

Heap

A record can be placed anywhere in the file where there is space

Sequential/Ordered

Store records in sequential order, based on the value of the search key of each record

Hashing

A hash function computed on some attribute of each record; the result specifies in which block of the file the record should be placed

Multitable clustering file organization

Store records of several different relations in the same file

Examples: Heap/Pile/Unordered File

In this simplest and most basic type of organization, records are placed in the file **in the order in which they are inserted**, so new records are inserted at the end of the file

Shows an unordered file for an EMPLOYEE relation.

In the below figure, we can see a sample of heap file organization for EMPLOYEE relation which consists of 8 records stored in 3 contiguous blocks, each blocks can contains at most 3 records.

	Name	EID	Address	Birthdate	Salary
block 1	Adams John	....	...	...	...
	Melinda, Perkin				
	Raymond, Wong				
block 2	Alan, Delon				
	Bill, Clinton				
	Nancy, Davies				
block 3	Jay, Shana				
	Son, Nguyen				

Examples: https://www.slideshare.net/macjones25/file-organisation

Examples: Ordered/Sequential/Sorted File

We can physically order the records of a file on disk based on the values of one of their fields—called the **ordering field**

Shows an ordered file with Name as the ordering key field (assuming that employees have distinct names).

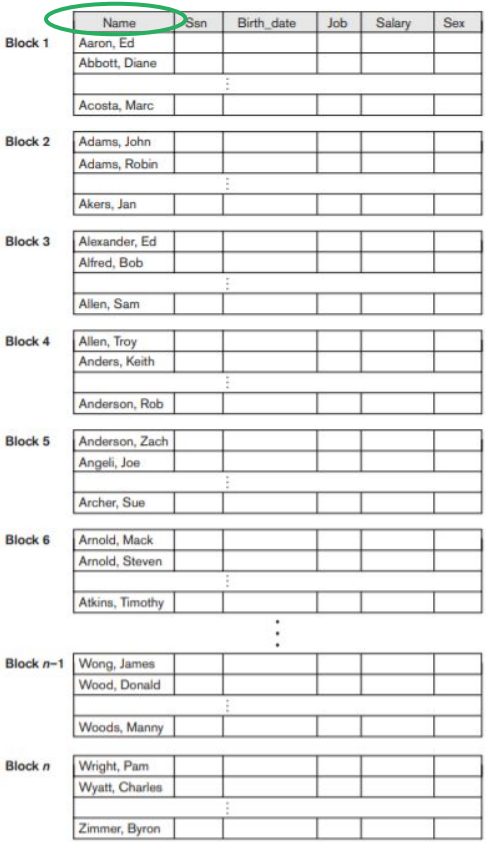

Figure 16.7
Some blocks of an ordered (sequential) file of EMPLOYEE records with Name as the ordering key field.