University of Central Florida
CGS 2545
Database Concepts

DEPARTMENT OF ELECTRICAL ENGINEERING & COMPUTER SCIENCE
COMPUTER SCIENCE DIVISION

File Structure

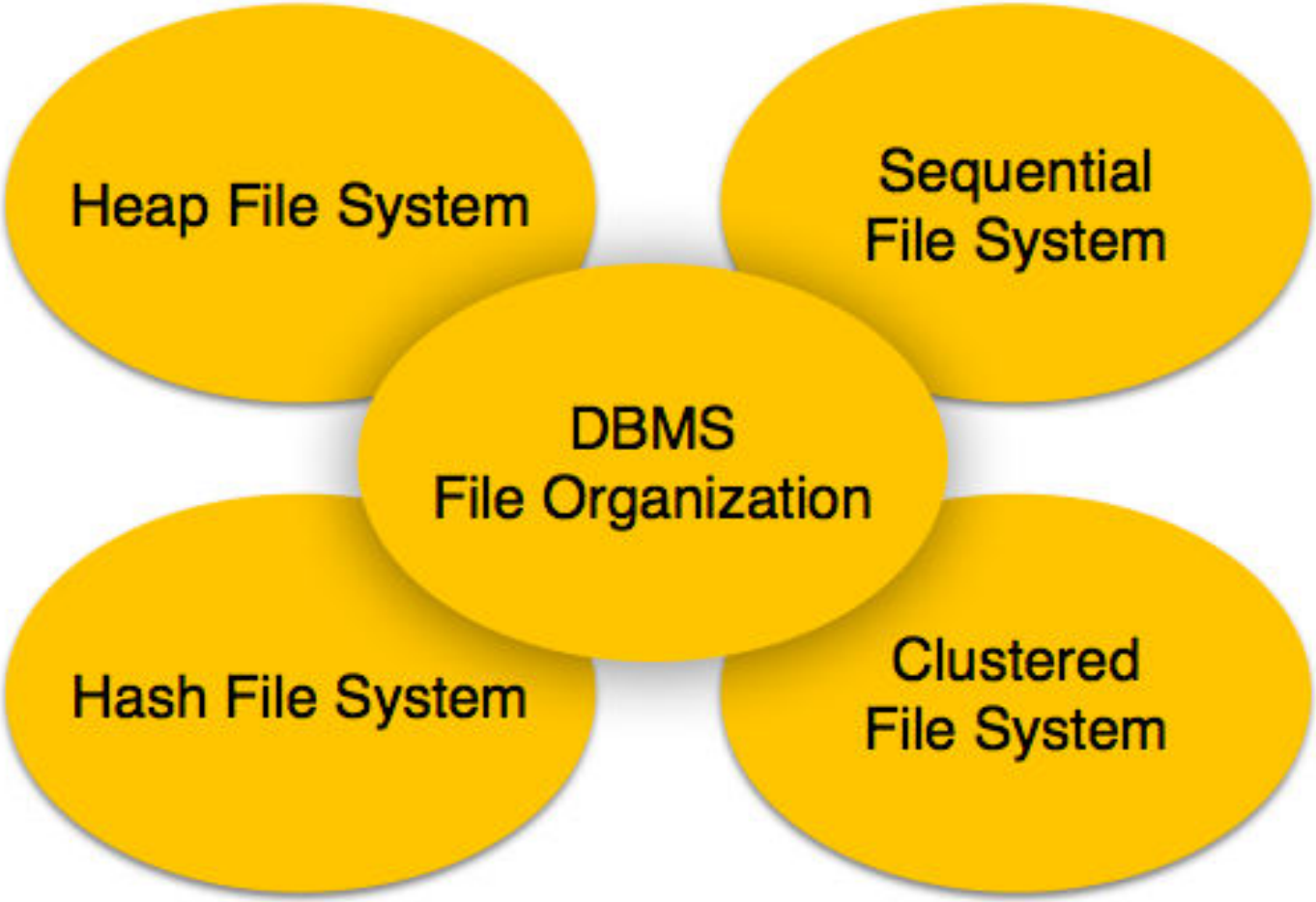
- Relative data and information is stored collectively in file formats.
- A file is a sequence of records stored in binary format.
- A disk drive is formatted into several blocks that can store records.
- File records are mapped onto those disk blocks

File Structure

- File Organization
  - defines how file records are mapped onto disk blocks.
  - There are four types of File Organization to organize file records
    - Heap File Organization
    - Sequential File Organization
    - Hash File Organization
    - Clustered File Organization

File Structure

- File Organization



File Structure

- File Organization
  - Heap File Organization
    - When a file is created using Heap File Organization, the Operating System allocates memory area to that file without any further accounting details.
    - File records can be placed anywhere in that memory area.
    - It is the responsibility of the software to manage the records.
    - Heap File does not support any ordering, sequencing, or indexing on its own.

File Structure

- File Organization
  - Sequential File Organization
    - Every file record contains a data field (attribute) to uniquely identify that record.
    - In sequential file organization, records are placed in the file in some sequential order based on the unique key field or search key.
    - Practically, it is not possible to store all the records sequentially in physical form.

File Structure

- File Organization
  - Hash File Organization
    - Hash File Organization uses Hash function computation on some fields of the records.
    - The output of the hash function determines the location of disk block where the records are to be placed.

File Structure

- File Organization
  - Clustered File Organization
    - Clustered file organization is not considered good for large databases.
    - In this mechanism, related records from one or more relations are kept in the same disk block, that is, the ordering of records is not based on primary key or search key.

File Structure

- File Operations
  - Operations on database files can be broadly classified into two categories
    - Update Operations
    - Retrieval Operations
  - Update operations change the data values by insertion, deletion, or update.
  - Retrieval operations, on the other hand, do not alter the data but retrieve them after optional conditional filtering.
  - In both types of operations, selection plays a significant role.

File Structure

- File Operations
  - Other than creation and deletion of a file, there could be several operations, which can be done on files.
    - Open
    - Locate
    - Read
    - Write
    - Close
  - The organization of data inside a file plays a major role here.
  - The process to locate the file pointer to a desired record inside a file various based on whether the records are arranged sequentially or clustered.

File Structure

- File Operations
  - **Open**
    - A file can be opened in one of the two modes
      - read mode
      - write mode
    - In read mode, the operating system does not allow anyone to alter data.
    - data is read only
    - Files opened in read mode can be shared among several entities.
    - Write mode allows data modification.
    - Files opened in write mode can be read but cannot be shared.

File Structure

- File Operations
  - **Locate**
    - Every file has a file pointer, which tells the current position where the data is to be read or written.
    - This pointer can be adjusted accordingly.
    - Using find (seek) operation, it can be moved forward or backward.

File Structure

- File Operations
  - **Read**
    - By default, when files are opened in read mode, the file pointer points to the beginning of the file.
    - There are options where the user can tell the operating system where to locate the file pointer at the time of opening a file.
    - The very next data to the file pointer is read.

File Structure

- File Operations
  - **Write**
    - User can select to open a file in write mode, which enables them to edit its contents.
    - It can be deletion, insertion, or modification.
    - The file pointer can be located at the time of opening or can be dynamically changed if the operating system allows to do so.

File Structure

- File Operations
  - **Close**
    - This is the most important operation from the operating system's point of view.
    - When a request to close a file is generated, the operating system
      - removes all the locks (if in shared mode),
      - saves the data (if altered) to the secondary storage media, and
      - releases all the buffers and file handlers associated with the file.