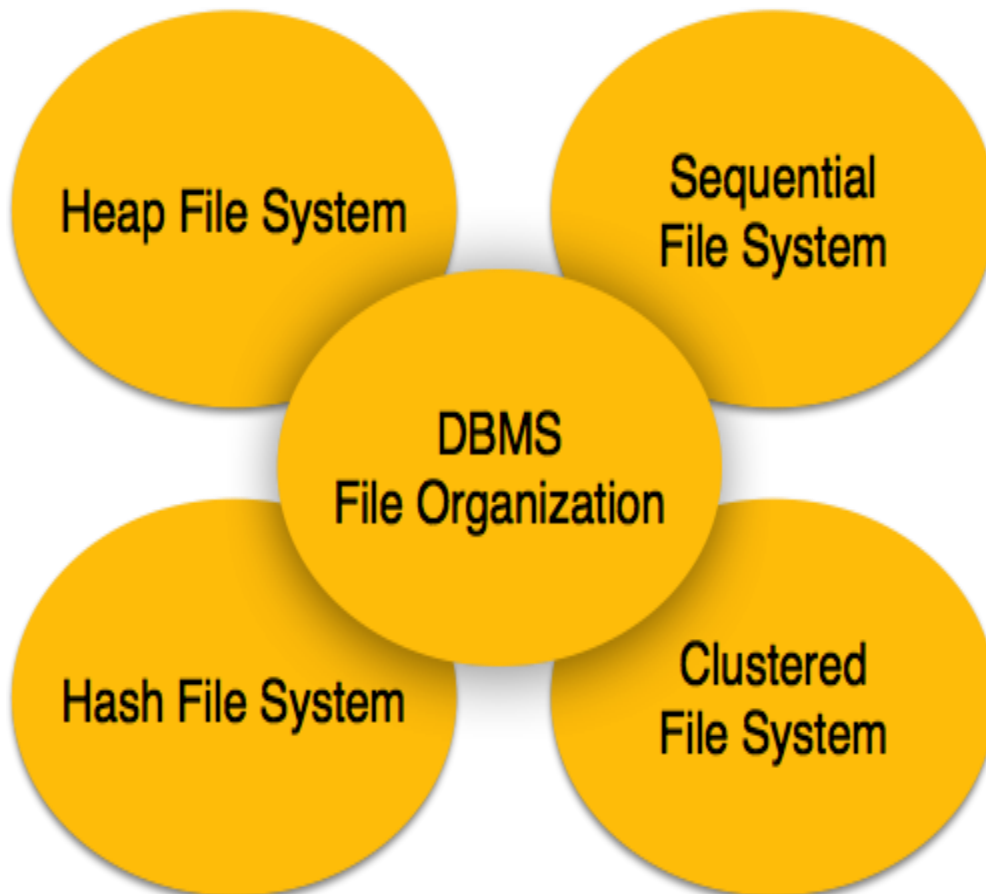


DBMS File Structure

Relative data and information is stored collectively in file formats. A file is sequence of records stored in binary format. A disk drive is formatted into several blocks, which are capable for storing records. File records are mapped onto those disk blocks.

File Organization

The method of mapping file records to disk blocks defines file organization, i.e. how the file records are organized. The following are the types of file organization



[Image: File Organization]

- **Heap File Organization:** When a file is created using Heap File Organization mechanism, the Operating Systems 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 software to manage the records. Heap File does not support any ordering, sequencing or indexing on its own.
- **Sequential File Organization:** Every file record contains a data field (attribute) to uniquely identify that record. In sequential file organization mechanism, records are placed in the file in the 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.
- **Hash File Organization:** This mechanism uses a Hash function computation on some field of the records. As we know, that file is a collection of records, which has to be mapped on some block of the disk space allocated to it. This mapping is defined that the hash computation. The output of hash determines the location of disk block where the records may exist.

- **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 a same disk block, that is, the ordering of records is not based on primary key or search key. This organization helps to retrieve data easily based on particular join condition. Other than particular join condition, on which data is stored, all queries become more expensive.

File Operations

Operations on database files can be classified into two categories broadly.

- **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 significant role. Other than creation and deletion of a file, there could be several operations, which can be done on files.

- **Open:** A file can be opened in one of two modes, read mode or write mode. In read mode, operating system does not allow anyone to alter data it is solely for reading purpose. Files opened in read mode can be shared among several entities. The other mode is write mode, in which, data modification is allowed. Files opened in write mode can be read also but cannot be shared.
- **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.
- **Read:** By default, when files are opened in read mode the file pointer points to the beginning of file. There are options where the user can tell the operating system to where the file pointer to be located at the time of file opening. The very next data to the file pointer is read.
- **Write:** User can select to open files in write mode, which enables them to edit the content of file. 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 allowed doing so.
- **Close:** This also is most important operation from operating system point of view. When a request to close a file is generated, the operating system removes all the locks (if in shared mode) and saves the content of data (if altered) to the secondary storage media and release all the buffers and file handlers associated with the file.

The organization of data content inside the file plays a major role here. Seeking or locating the file pointer to the desired record inside file behaves differently if the file has records arranged sequentially or clustered, and so on.

Source:

http://www.tutorialspoint.com/dbms/dbms_file_structure.htm