

UNIT-04

NORMALIZATIONS

Refer PPT shared in Group

Problems on Normalization

Refer lecture notes

INTRODUCTION TO FILE ORGANIZATION & DBMS

FILE SYSTEM

- File: File is a collection of records which are logically related to any object. Record value can in any form like data.
- For eg. : Each students records which having values of Roll no, Name, Class. For arranging data we use file.
- For eg.: files of bank"s customer, files of department, files of stock records etc. Files are recorded on secondary storage such as magnetic disks, magnetic tables and optical disks

TYPES OF FILES:

Physical file:

- Physical file concern with actual data that is stored.
- It stores description about how the data is to be represented.

Logical file

- Logical file: do not contain data.
- They contain a description of records that are found in one or more physical files.
- A logical file is a view or representation of one or more physical files.

Special character file:

- At the time of file creation we insert some special characters in file.
- For eg: Control + z for end of a file which having ASCII value 26

ACCORDING TO RECORDS TYPES OF FILES:

1. Fixed length record file:

- a. Every record in this file has same size(in bytes). Record having value set, in the fixed length record file, memory block are assign in same size.
- b. For eg., if the size for a record is assigned 30 bytes to each then records in this type are stored like as below,
- c. Advantage: records are stored in fixed distance of memory block, so fast searching for a particular record is done.
- d. Disadvantage: Memory blocks are unnecessarily used when record size is small as compared to assigned memory block. This useless memory block increases size of file.



2. Variable length record file:

- a. Every record in this file has variable size (in bytes). Memory block are assign for a file records are in variable size. Different records in the file have different sizes. As per size of records value, memory blocks are used.
- b. Advantage: Memory used efficiently for storing record. Whatever exact size of record that much size of memory block occupies in memory in this kind of records. Because of less memory they can move, save or transfer from one location to other in fast manner. Record 1 Record 2 Record 3 Record 4 Record 5 Record 6
- c. Disadvantage: Access for record is slower as compared to fixed length record file due to varying size of a record

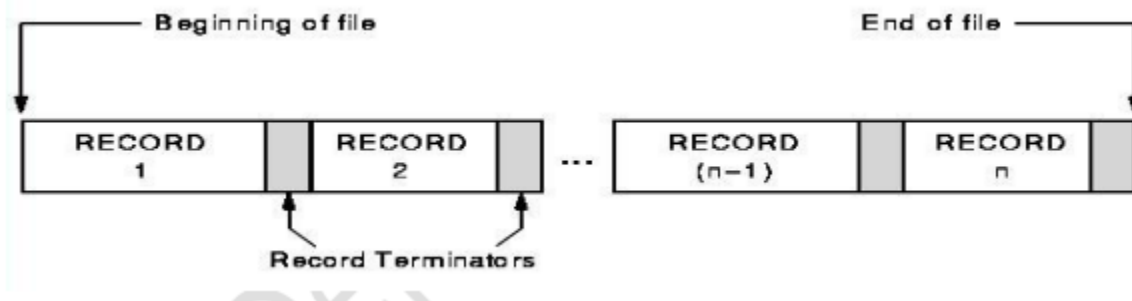


FILE ORGANIZATION:

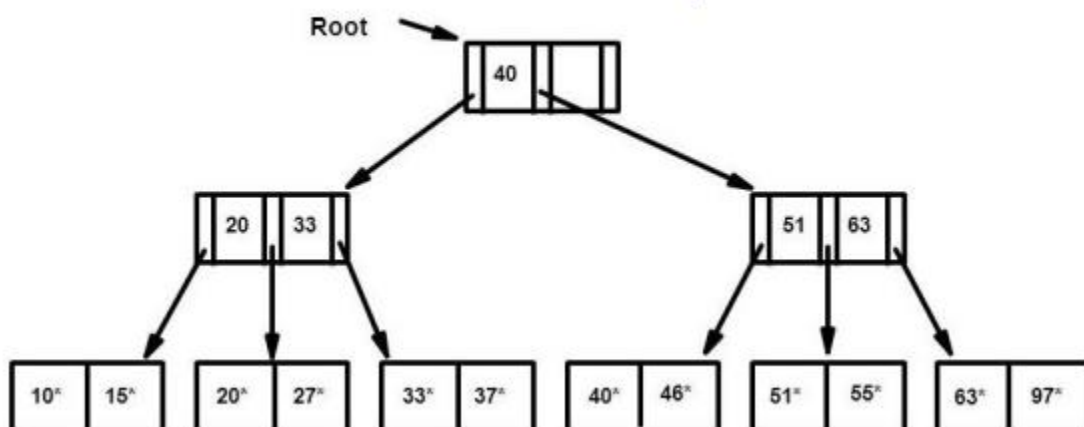
File organization refers to the logical relationships among various records that constitute the file, particularly with respect to the means of identification and access to any specific record. In short, storing the files in certain order is called file organization.

Types of file organization:

1. Sequential file organization: Sequential file organization is easiest method. In this method files are stored one after the other in a sequential manner. This method is also called as Pile or sorted file. This method is fast & efficient for huge amount of data. Sorted file is inefficient as it takes time & space for sorting records.



2. Serial file organization: Serial file organization is also called as heap file. In this method, records are inserted at the end of file into the data blocks. There is no requirement of sorting data. When huge amount of data is to be inserted at a time in a organization, that time this method is suitable. Accessing of data is slower as compared to sorted file method.
3. Index Sequential Access Method (ISAM): ISAM method is advanced sequential file organization. In this method, index value is generated and mapped with every record. Using that index, accessing of record is done.



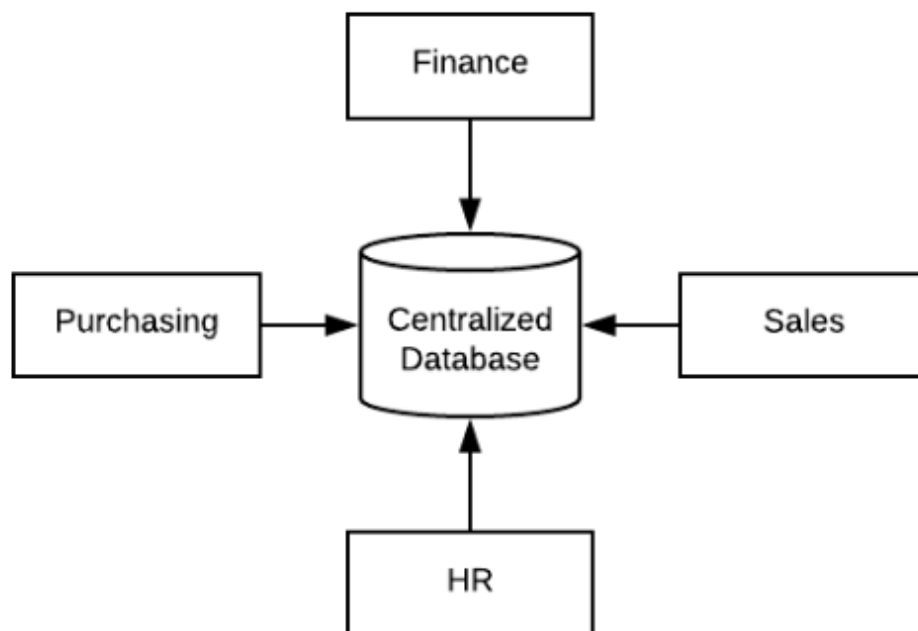
4. Random access / Direct access file organization: In this file organization records are stored randomly but accessed directly. To access a file

stored randomly, a record key is used to determine where a record is stored on the storage media. Magnetic and optical disks allow data to be stored and accessed randomly.

TYPES OF DATABASE MANAGEMENT SYSTEMS

1. Centralized Database Management System

- A centralized database is stored at a single location such as a mainframe computer. It is maintained and modified from that location only and usually accessed using an internet connection such as a LAN or WAN. The centralized database is used by organisations such as colleges, companies, banks etc.



- As can be seen from the above diagram, all the information for the organisation is stored in a single database. This database is known as the centralized database.

- Advantages
 - The data integrity is maximised as the whole database is stored at a single physical location. This means that it is easier to coordinate the data and it is as accurate and consistent as possible.
 - The data redundancy is minimal in the centralised database. All the data is stored together and not scattered across different locations. So, it is easier to make sure there is no redundant data available.
 - Since all the data is in one place, there can be stronger security measures around it. So, the centralised database is much more secure.
 - Data is easily portable because it is stored at the same place.
 - The centralized database is cheaper than other types of databases as it requires less power and maintenance.
 - All the information in the centralized database can be easily accessed from the same location and at the same time.
- Disadvantages
 - Since all the data is at one location, it takes more time to search and access it. If the network is slow, this process takes even more time.
 - There is a lot of data access traffic for the centralized database. This may create a bottleneck situation.
 - Since all the data is at the same location, if multiple users try to access it simultaneously it creates a problem. This may reduce the efficiency of the system.
 - If there are no database recovery measures in place and a system failure occurs, then all the data in the database will be destroyed.

Distributed Database

A **distributed database** is a collection of multiple interconnected databases, which are spread physically across various locations that communicate via a computer network.

Features

- Databases in the collection are logically interrelated with each other. Often they represent a single logical database.
- Data is physically stored across multiple sites. Data in each site can be managed by a DBMS independent of the other sites.
- The processors in the sites are connected via a network. They do not have any multiprocessor configuration.
- A distributed database is not a loosely connected file system.
- A distributed database incorporates transaction processing, but it is not synonymous with a transaction processing system.

Client/Server:

Client/server is developed to deal with various computing environments that have a large number of computers and servers connected together via a network. In this architecture, a Client is a user machine which provides the user interface and local processing capabilities. When any client requires additional functionality like database access, it can connect to Server that is capable of providing the functionality needed by the client. Basically Server is a machine that provides services to the Client i.e user machine.

