

Unit-1: Database and Database users:

- Ankit Pangani

DATE

* Data vs Information

- Data are the raw facts that are generated after some experiment, observation or analysis.
It can be texts, numbers or symbols written in papers, or it can be bits or bytes inside the memory of electronic devices.

- Data itself do not provide any meaning but after processing it becomes information.

Note: → Data is input & information is output
→ Data is raw and information is processed
→ Data doesn't depend on information but ^{not} vice-versa.

* Database

- It is the collection of data organized in some specific manner. Example: University database for maintaining information about students name, courses and grades in university.

- Database is usually controlled by a database management system (DBMS).

* Database Management System (DBMS)

- The database, its processing methods and the set of rules and conditions to be followed is collectively known as DBMS.

- It is a way of controlling database i.e. store and manipulate the data.

- It is basically a computerized record keeping system whose main goal is to store and retrieve data in a convenient and efficient ~~process~~ manner.

Ex: Oracle, MySQL, SQL servers, MS Access etc

* Some applications of DBMS.

- o Banking → To store information about customers, their account numbers, balance etc.
- o Airlines → For reservations and schedule information.
- o Telecommunications → To keep records of customers ex: call records, bills, balance left etc.
- o Universities → To keep records of students
- o Manufacturing → To store, track, orders & items
- o Online Shopping → Store product and customer information

* File Management System (Flat File system) (Traditional filing system)

- File Management System (FMS) also called Flat File System or traditional filing system. It stores data in a plain text file and allows access to single files or tables at a time.
- A flat file is a file that contains ~~the~~ records and in which record is specified in a single line. There is no structural relationships and the data are 'flat' as in paper.

→ (Limitations of flat-file systems)

(also can be written as advantages of DBMS over File Management system by changing languages.)

◦ Data Redundancy and Inconsistency

Redundancy is the repetition of data. ie. each data may have more than a single copy. Ex: address and telephone number of a customer may appear in a file that consists of savings-accounts records and in a file that consists of checking-account records. This redundancy leads to higher storage and access cost. In addition, it may lead to data inconsistency, ie., the various copies of the same data may no longer agree. Ex: changed customer address may be reflected in savings-accounts records but not elsewhere in system.

◦ Difficulty in accessing data:

It is not easy to access data in a desired or efficient way. If a user wants information in a specific manner, then he requires creating a program for it. Ex: say if admin of a college wants information like his name, roll number, marks ~~etc~~ then program for it is written but if he wants the record of students whose score is more than 80 %, then he require to create a different program for it.

◦ Data Isolation

Because data are scattered in various files, and files may be in different formats, writing new application programs to retrieve the appropriate data is difficult.

- o Security Problems

Poor data security is the most threatening problem in File processing system. There is very less security as anyone can easily modify and change the data.

- o Wastage of Labor and space

File processing system needs lots of copied data in different files that cause wastage of labour and space.

- o Integrity and Atomicity Problem

If a computer is subjected to failure, it is very crucial to retrieve the data.

→ Advantages of Flat-File System

Despite of the number of drawbacks, they are beneficial in many situations and are widely used:

- o Flat files are relatively quick and easy to setup and use
- o Easy to understand and implement
- o Less hardware and software requirements
- o Less skill is required to handle
- o Best for small databases

(Advantages of database m.s. over traditional).
DATE Advantages

* Database management system (Purpose / Functions)

Modern database management systems have a number of advantages over traditional approaches to data base maintenance. Some of them are:

- Provides multi user interface.
- Reduce data redundancy and inconsistency.
- Easily access data as well as restricting unauthorized access.
- Maintain integrity and atomicity.
- Enforce/Improve security.
- Provide efficient backup and recovery.
- Share data easily.
- Up-to date information.

Explain above in simple language. (next few pages) ^{after}

* Disadvantages of DBMS

- Problems associated with centralization.
- Increased costs.
- Complexity of backup and recovery.
- Frequent upgrade/replacement cycle.
- Management and Maintenance.
- Database failure.

Explain above in simple language.

File based systemDatabase system

- | | |
|---|---|
| i) Here the data and program are dependent. | i) Data and program are independent of each other. |
| ii) It causes data redundancy. The data may be duplicated in different files. | ii) It controls data redundancy. The data appear only once in the system. |
| iii) Causes data inconsistency. Data in different files may be different. | iii) Data are always consistent. Because data appeared only once. |
| iv) Data cannot be shared because data is distributed in different files. | iv) Data can be shared easily because data is stored at one place. |
| v) Provides poor security. | v) Provides advanced security. |
| vi) It is less complex system. and requires more space & memory. | vi) It is very complex system. and requires less space and memory. |
| vii) Does not provide concurrency facility and atomicity functionality. | viii) Provides |
| ix) Difficult to maintain and hardware cost is less than database. | x) Easy to maintain and hardware cost is more. |

Characteristics of the Database Approach

In past, traditional file processing approach was used in which each user defines and implements the files needed. But in database approach, a single repository ^{by a programming} maintains data that is defined once and then accessed by various users repeatedly.

The main characteristics of the database approach versus the file processing approach are:

- i) Self-describing nature of a database system
- ii) Insulation between programs and data, and data abstraction
- iii) Support for multiple views of data
- iv) Sharing of data and multiple transaction processing

Self-describing nature of a database system

The database contains not only the database itself but also a complete description. The definition is stored in DBMS catalog which contains information of each data item, like their structure, type and storage format. Information stored in catalog is called meta-data.

Latest database systems don't require meta data, as data is stored as self-describing data which includes every information about data in one structure. These database systems are known as NoSQL systems.

- Insulation between programs and data, and data abstraction. (hides the data)

In traditional file processing, the structure of data files is embedded in the application programs, so any changes to the structure of a file may require changing all programs that access the file. But in DBMS, access programs do not require such changes in most cases. The structure of data files is stored in DBMS catalog separately from the access programs. We call this property program-data separation.

- Support of multiple views of data.

A database typically has many types of users, each of whom may require a different perspective or view of database. A multi-user DBMS whose users have a variety of distinct applications must provide facilities for defining multiple views.

- Sharing of data and multi-user transaction processing

A DBMS must allow multiple users to access the database at the same time. This is essential if data for multiple applications is to be integrated & maintained in a single database.

* Actors on the scene.

In large organizations, many people are involved in the design, use and maintenance of a large database with hundreds or thousands of users. The people whose jobs involve the day-to-day use of a large database are called the actors on the scene.

i) Database Administrators (DBA)

DATE [] [] [] []

In database environment, the primary resource is database itself & secondary is DBMS & related software. Administering these resources is responsibility of DBA. He/she authorizes access to database, coordinating and monitoring its use. He/she is also accountable for problems such as security breaches & poor system response time.

- ii) Database designers: They are responsible for identifying the data to be stored in the database & for choosing appropriate structures to represent & store this data. [create a design that meets all requirements]
- iii) End users: People whose jobs require access to the database for querying, updating & generating reports.
casual end users, novice end users, sophisticated end users, ^{standardized} _{engineers}
- iv) System Analysts & Application Programmers (software engineers)

* Workers Behind the Scene.

Those who work to maintain the database system environment but who are not actively interested in the database contents as a part of their daily job are called workers behind the scene. They include the following categories.

i) DBMS System designers and implementors

They design and implement the DBMS modules and interfaces as a software package. A DBMS is a very complex software consisting many components including query language processing, handling data recovery and security etc. It must interface with other system software such as operating system & compilers for various programming languages.

classmate

PAGE [] []

ii) Tool developers:

They design and development tools which are software packages that facilitate database modeling and design and improve performance. Tools are optional packages that are often purchased separately. In many cases, independent software vendors develop and market these tools.

iii) Operators and maintenance personnel

They are also called system administrators and are responsible for the actual running and maintenance of the hardware and software environment for the database system.

Note: System analysis & Application

→ Determines the requirements for end users, especially naïve users and casual end users.

They should be familiar with full range of capabilities of DBMS.

* Advantages of Using the DBMS Approach:

(already written before, explanation here)

o Reduce data redundancy and inconsistency

The file based management system contains multiple files that are stored in many different locations in a system or even across multiple systems. Which may lead to multiple copies of data and causes data redundancy.

This is prevented in database, as there is a single repository which is defined once and any change in it is reflected immediately. Because of this, there is no chance of data duplication and inconsistency.

o Restricting Unauthorized access

In DBMS, only authorized users are allowed to access the database and they can easily access data once they are authorized. Their identification is based on their username and password.

Unauthorized users cannot access the database under any circumstances as it violates the integrity constraint.

o Providing Multiuser Interface

Many types of users with varying levels of technical knowledge use database. DBMS provides a

Variety of user interface: like mobile ~~apps~~ ^{YAGE}, Classmate websites, computer apps, query language, API and so on

So that users can access data easily from anywhere and any interface device.

DATE

o Provide efficient backup and Recovery

It is responsible for recovery of data in case the computers fails in the middle of a complex ~~up~~ data transfer process. The recovery system is responsible for making sure that the data is backed up and restored to the state it was paused before or ~~to~~ to the failure state.

o Enforce/Improve Security

Data security is a vital concept on a database. DBMS prevents the leakage of any data from the database by using advanced security layers and by preventing unauthorized access of data.

o Up to date information.

As soon as the user update is applied to the database repository, all other users can immediately see the update. This availability of up to date information is very essential for many transaction processing applications such as banking or reservation systems.

o Economics of scale:

The DBMS approach reduces the amount of wasteful overlap between activities of data-processing personnel. This enables ^{classmate} the whole organization to invest in more

powerful processors, storage devices, rather than having each department purchase its own low performance equipment. This reduces overall costs of operation and management.

Questions asked from this Chapter (board exam)

- Q. What are the advantages of using DBMS over traditional filing system? (2076 → 3 marks) (2073 → 5 marks)
- Q. Who is data administrator? What are the main functions of database administrator? (2074 → 5 marks)
- Q. Difference between Database Manager and database administrator? (2073 → 5 marks)
- Q. What is data encryption? (2073 → 5 marks)
- Q. Difference between Database Management system and file based System. (very important)
- Q. Why do we need DBMS? Discuss drawbacks of file system and advantages of DBMS.