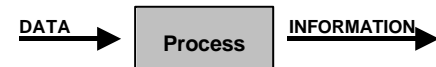


Database Environment

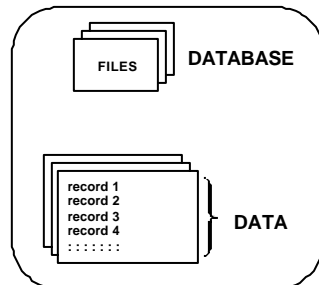
Basic Terminologies

- Data are facts concerning things such as people, objects or events.
- Information is data that have been processed and presented in a form suitable for human interpretation, often with the purpose of revealing trends or patterns.
- Process involves acquisition, storage, manipulation, retrieval and distribution.



Database Concept

- A database is a shared collection of logically related data, designed to meet the information needs of multiple users in an organization.



Centralized Database

- All data are located at a single site in a Centralized Database
- Advantage:
 - ◆ Provide greater control over accessing and updating data.
- Disadvantage:
 - ◆ Vulnerable to failure
- Examples:
 - ◆ Personal Computer Database, Central Computer Database, Client/Server Databases

Distributed Database

- A single logical database that is spread physically across computers in multiple locations.

Heterogeneous Databases

- Different computers and operating systems, different data models, different DBMS may be possible.
- Comprise of the following conditions:
 - ◆ The computer operating systems used at each location are the same or highly compatible.
 - ◆ The data models used at each location are the same.
 - ◆ The DBMS used at each location are the same or highly compatible.
 - ◆ The data at the various locations have common definitions and formats.

File Processing Approach

- A traditional approach to information system design.
- Focuses on the data processing needs of individual departments in the organization.

Disadvantages

- Uncontrolled Redundancy
- Inconsistent Data
- Inflexibility
- Limited Data Sharing
- Poor Enforcement of Standards
- Excessive Program Maintenance

Uncontrolled Redundancy

- Duplication of data.
- Valuable storage space is wasted
- The same data may have to be input several times to update all occurrences of a data item
- Inconsistencies

Inconsistent Data

- When the same data are stored in multiple locations, inconsistencies are inevitable.
- Inconsistencies in stored data are one of the most common sources of errors in computer applications.
- They lead to inconsistent documents and reports and undermine the confidence of users in the integrity of the information systems.

Inflexibility

- Inflexible and cannot easily respond to requests for a new or redesigned product.
- Leads to considerable frustration on the part of the users, who cannot understand why the computer system cannot give them information in a new format when they know it exists in the applications files.

Limited Data Sharing

- Each application has its own private files, and users have little opportunity to share data outside of their own applications.
- The same data may have to be entered several times to update files with duplicate data.
- In developing new applications, the designer often cannot exploit data containing in existing files; instead new files are designed that duplicate much of the existing data.

Poor Enforcement of Standards

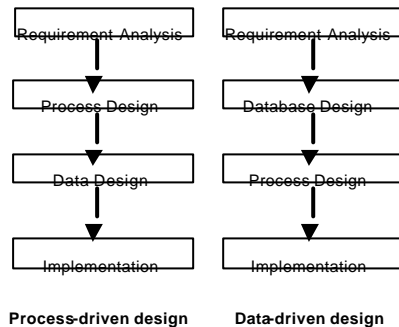
- The responsibility for system design and operation has been decentralized.
- Two types of inconsistencies may result from poor enforcement of standards: Synonyms and Homonyms.
 - ◆ A synonym results when two different names are used for the same data item (e.g. student number and matriculation number)
 - ◆ A homonym is a single name that is used for two different data items (e.g. in a bank the term balance might be used to designate a checking account balance in one department and a savings account balance in a different department)

Excessive Program Maintenance

- Descriptions of files, records, and data items are embedded within individual application programs.
- Any modification to a data (such as change of data name, data format, or method of access) requires that the program (or programs) also be modified.

Database Approach

- Emphasizes the integration and sharing of data across the organization.



Benefits of the Database Approach

- Minimal Data Redundancy
- Consistency of Data
- Integration of Data
- Sharing of Data
- Ease of Application Development

Minimal Data Redundancy

- Data files are integrated into a single, logical structure.
- Improve performances
- Aware of redundancy

Consistency of Data

- By controlling data redundancy, we greatly reduce the opportunities for inconsistency.
 - ◆ For example, if each address is stored only once, we cannot have disagreement on the stored values.
- When controlled redundancy is permitted in the database, the database system itself should enforce consistency by updating each occurrence of a data item when a change occurs.

Integration of Data

- Data are organized into a single, logical structure, with logical relationships defined between associated data entities.
- Easy for users to relate one item of data to another.

Sharing of Data

- Permit multiple users to share a database concurrently

Ease of Application Development

- Greatly reduces the cost and time for developing new business applications as programmer is relieved from the burden of designing, building, and maintaining master files.
- Data are independent of the application programs that use them.
- Either data or the application programs that use the data can be changed without necessitating a change in the other factor.

Cost of the Database Approach

- New, Specialized Personnel
- Need for Explicit Backup
- Interference with Shared Data
- Organizational Conflict

New, Specialized Personnel

- Need to hire train individuals
 - ◆ To maintain the new database software
 - ◆ To develop and enforce new programming standards
 - ◆ To achieve the highest possible performance
 - ◆ To manage the new staff

Need for Explicit Backup

- Restoring damaged files and in providing validity checks on crucial data.
- To ensure that data are accurate and available whenever needed
- Either database management software or additional procedures have to provide these essential capabilities.

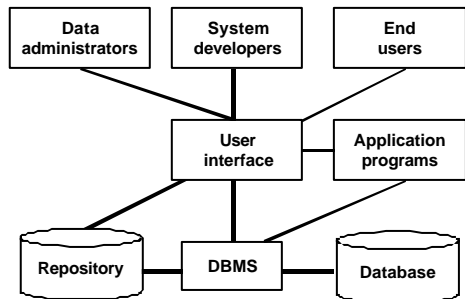
Interference with Shared Data

- The concurrent access to shared data via several application programs can lead to some problems.
- Example:
 - ◆ When two concurrent users both want to change the same or related data, inaccurate results can occur if access to the data is not properly synchronized.
 - ◆ When data are used exclusively for updating, different users can obtain control of different segments of the database and lock up any use of the data (deadlock).

Organizational Conflict

- A shared database requires a consensus on data definitions and ownership as well as responsibilities for accurate data maintenance.
- Experience has shown that conflicts on how to define data, data length and coding, rights to update shared data, and associated issues are frequent and difficult managerial issues to resolve.

Components of the Database Environment



Components of the Database Environment

- Repository
 - ◆ Centralized knowledge base containing all data definitions, screen and report formats and definitions of other organizations and system components.
- Database Management System (DBMS)
 - ◆ Commercial software system used to create, maintain and provide controlled access to the database and repository.

Components of the Database Environment

- Database
 - ◆ A shared collection of logically related data, designed to meet the information needs of multiple users in an organization.
- Application Programs
 - ◆ Computer programs that are used to create and maintain the database and provide information to users.

Components of the Database Environment

- User Interface
 - ◆ Languages, menu and other facilities by which users interact with various system components.
- Data Administrators
 - ◆ Persons who are responsible for the overall information resources of an organization.

Components of the Database Environment

- System Developers
 - ◆ Persons such as system analysts and programmers who design new application programs.
- End Users
 - ◆ Persons throughout the organization who add, delete and modify data in the database and who request or receive information from it.

Review Question

1. Discuss the characteristics of traditional file processing system. Why is the system criticized?
2. How did the database approach eliminate problems of file processing system?
3. Describe the components of a database system, with the aid of a diagram.
4. Explain why data redundancy is so common in traditional application systems.
5. Where are data definitions maintained in each of the following environment?
 - ◆ Traditional file processing system
 - ◆ Database system