

---

# Database Systems

# Data

---

- Data refers to the facts & figures concerning:
  - ✓ PEOPLE
  - ✓ EVENT
  - ✓ ORGANISATION
  - ✓ OBJECT
- The modern data sources/types include multimedia data, sensor's data, data on web and so on.
- To get most out of data, we require tools than can simplify the task of **managing the data** and **extract useful information** out of it in **timely** fashion.

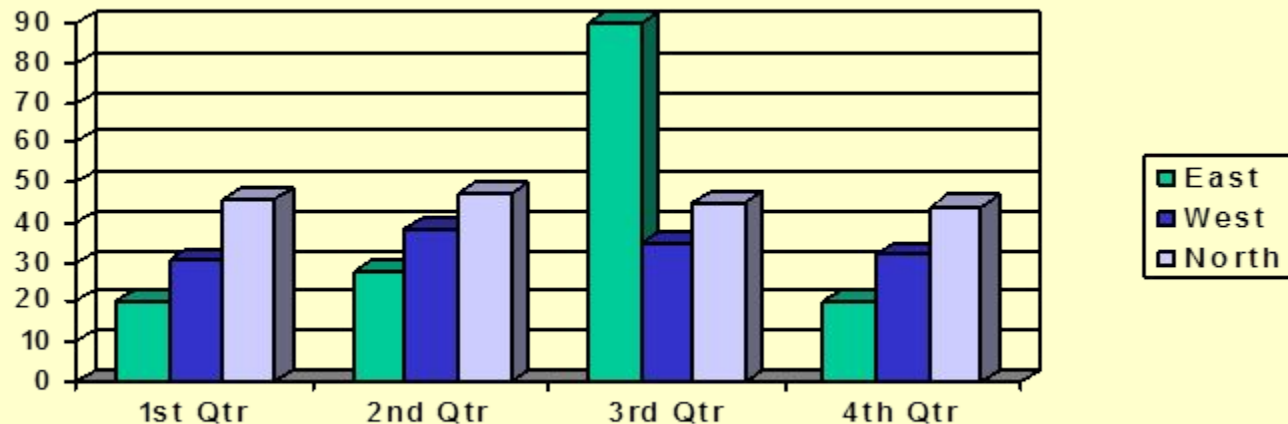
# Data

- Large volume of facts, difficult to interpret or make decisions based on

| Class Roster       |                            |              |             |
|--------------------|----------------------------|--------------|-------------|
| Course:            | MGT 500<br>Business Policy | Semester:    | Spring 200X |
| Section:           | 2                          |              |             |
| <u>Name</u>        | <u>ID</u>                  | <u>Major</u> | <u>GPA</u>  |
| Baker, Kenneth D.  | 324917628                  | MGT          | 2.9         |
| Doyle, Joan E.     | 476193248                  | MKT          | 3.4         |
| Finkle, Clive R.   | 548429344                  | PRM          | 2.8         |
| Lewis, John C.     | 551742186                  | MGT          | 3.7         |
| McFerran, Debra R. | 409723145                  | IS           | 2.9         |
| Sisneros, Michael  | 392416582                  | ACCT         | 3.3         |

# Data vs Information

- Information is derived from data OR Information is the processed data
- Information may be presented either Textually or Graphically or both



- The representation and amount of information depends upon the requirements of a user.
- When data is processed and organized into a form needed for its application, it is called an **Information Product (IP)**.

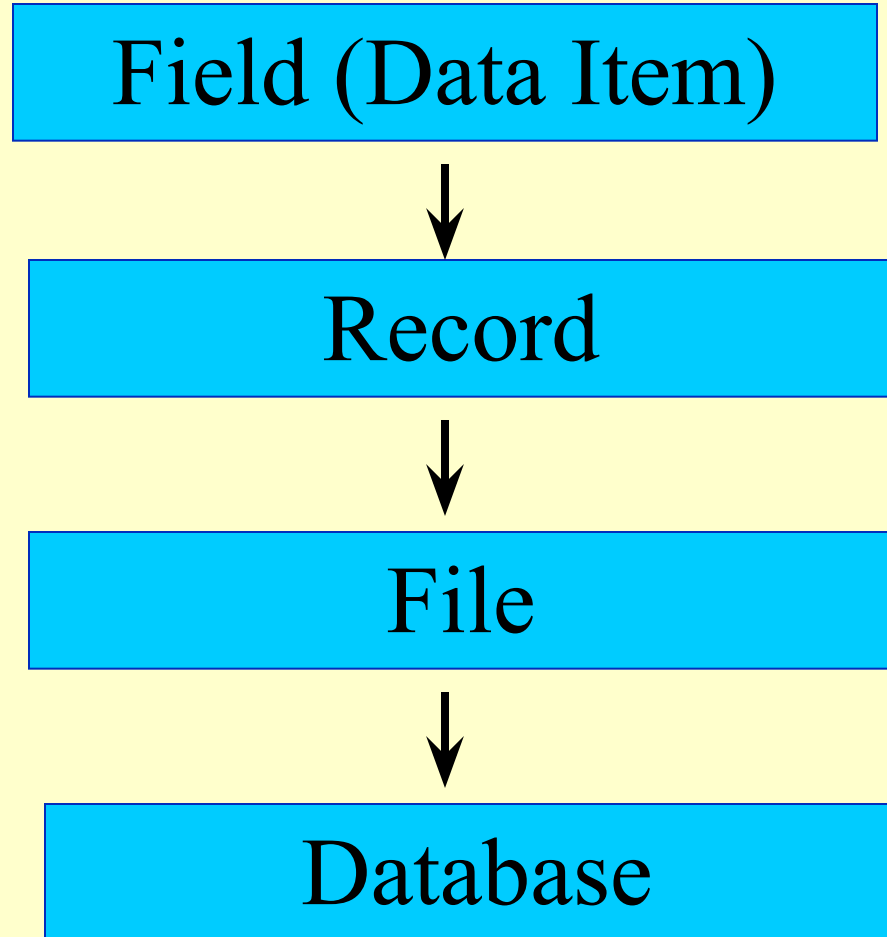
# Data vs Information

---

- Accurate, reliable, relevant and timely information is key to good decision making.
- Information is now increasingly viewed as an organizational asset (Redman, 2008) which can be used to improve organizational performance and help an organization gain a competitive advantage in the marketplace.

# **Data Hierarchy**

---



# Database

- Database is a shared collection of logically related data, and a description of this data, designed to meet the information requirements of multiple users in an organization.
- Database System can simply be regarded as a computerized record keeping system.
- Database System is a system whose overall purpose is to maintain data and make the information available to its users as per their requirements.
- The user of the system can perform the operations like:
  - Defining structure and types of data
  - Data manipulation (Insert, Delete, Update, Query etc.)
  - Enforcing security restrictions

# Database

---

- Database is an organized collection of **data** (and **metadata**) about **entities** and the **relationships** among these entities.
- **Metadata:** Data that describes properties or characteristics of other data OR it is “data about data”.
- The description of data is also known as **system catalog** or **data dictionary**.
- Metadata allows database designers & users to understand what data exists and what the data mean?
- Metadata describes the domain (set of possible values) for a data item.



# Metadata Example

| Data Item |              | Value  |     |     |                  |
|-----------|--------------|--------|-----|-----|------------------|
| Name      | Type         | Length | Min | Max | Description      |
| Course    | Alphanumeric | 30     |     |     | Course ID & Name |
| Section   | Integer      | 1      | 1   | 9   | Section Number   |
| Semester  | Alphanumeric | 10     | 1   | 8   | Semester & Year  |
| Name      | Alphanumeric | 30     |     |     | Student name     |
| ID        | Integer      | 9      |     |     | Student ID       |
| Major     | Alphanumeric | 4      |     |     | Student Major    |
| GPA       | Decimal      | 3      | 0.0 | 4.0 | Student GPA      |

# Examples of Database Applications

---

- Databases are everywhere ...
  - Student / Employee's Information System
  - Library System
  - Hotel / Airline Reservation System
  - Billing System
  - Stock / Inventory System
  - Payroll System

# **Record Keeping Techniques**

---

1. Manual Record Keeping.
2. Computerized Record Keeping
  - a) File Processing System.
  - b) Database System or Integrated Database Environment. (The paradigm shift)

# Case Study

---

- Consider a College System consisting of the following offices:
  - Admission Office.
  - Academics Office.
  - Exam Office.
- Each Office is maintaining its own set of files for its day to day operations.
- Some of the files used in the system are Student's File, Faculty File, Course File, Correspondence File etc.

# Manual System - Drawbacks

---

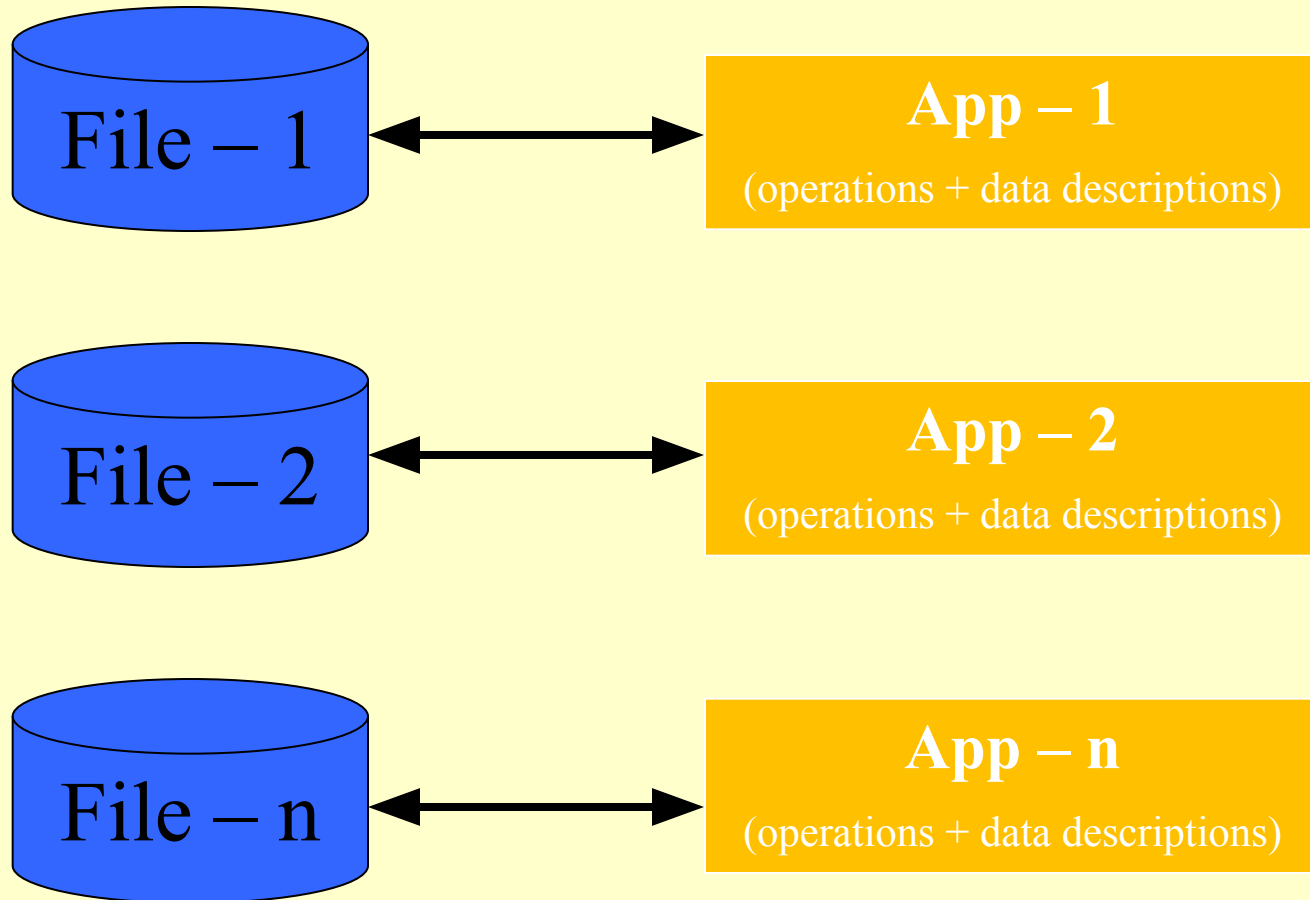
- High data volume
- Not reliable
- Inefficient
- Duplication of data
- Inconsistency
- A lot of data movement is required
- The System can't answer complex queries involving multiple departments.

# **File Processing System**

- File-based systems were an early attempt to computerize the manual filing system.
- In the traditional file processing system, the existing manual system is automated by focusing on the data processing needs of the individual departments instead of treating the organization as whole.
- Each application will have its own set of **Private Files** designed to meet the needs of a particular department.

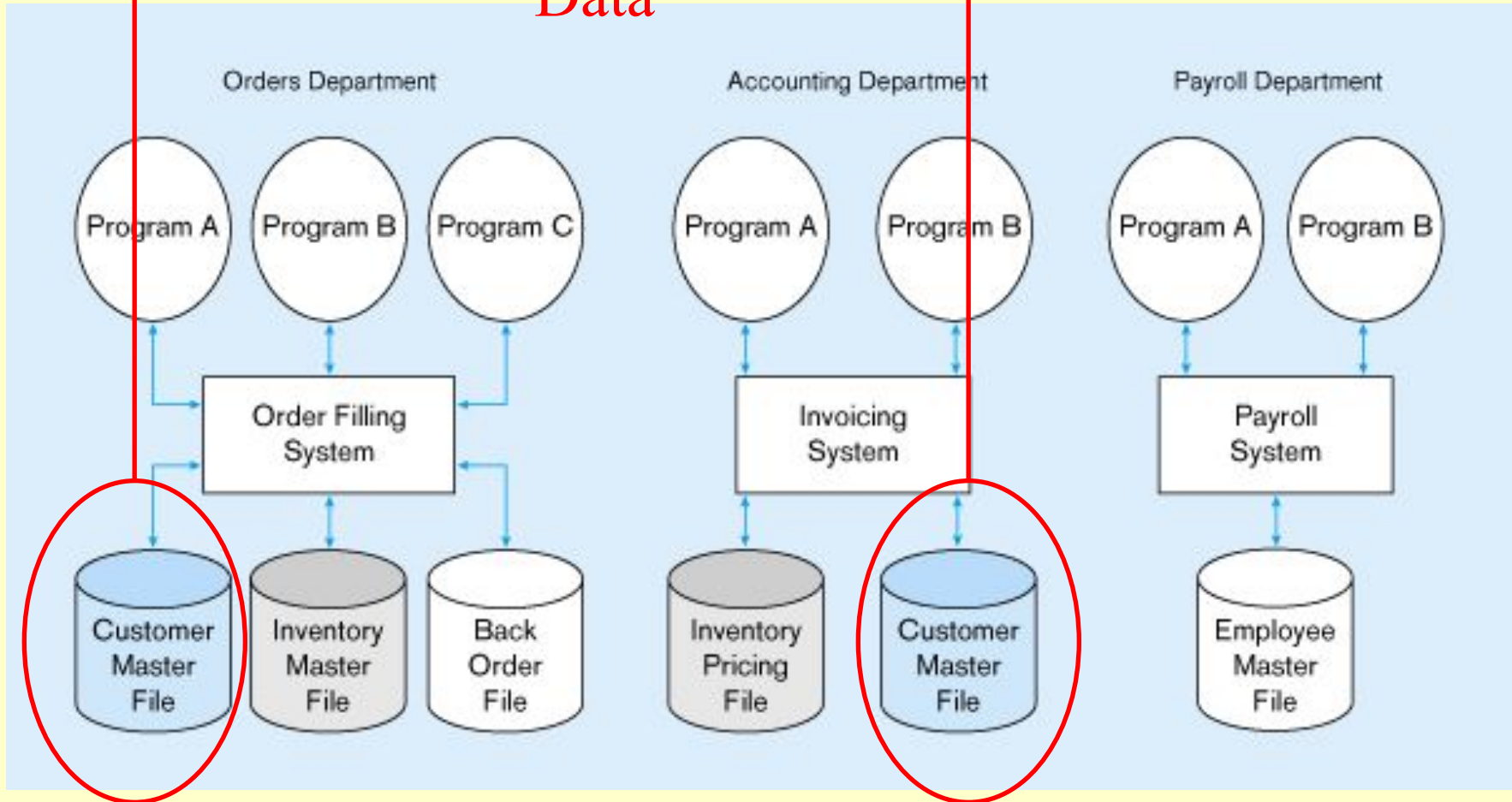
# File Processing System

---



# File Processing System

Duplicate  
Data





# **Drawbacks of File System**

1. Redundancy of Data
2. Inconsistent Data
3. Poor Enforcement of Standards
4. Excessive Program Maintenance
5. Limited Data Sharing
6. Inflexible
7. Lack of provision for security
8. Limited recovery from failure

# Drawbacks of File System

---

## 1. Redundancy of Data

- Redundancy means duplication of data.
- Since applications in FPS are independent so unplanned duplicate data is there.
- Same information is needed to be kept at different places. It costs time and money.
- Same data may have to be input several times to update all the occurrences.
- Redundancy results in data anomalies (Insert, Update, Delete)

# Drawbacks of File System

---

## 2. Inconsistent Data.

- Redundancy leads to inconsistency.
- The same data stored at different places may disagree with each other if an operation has not yet been performed against all the occurrences.
- The use of different programming environment and platforms lead to Heterogeneous Environment.

## 3. Poor Enforcement of Standards

- Organization wide enforcement of standards is poor. This leads to the following types of inconsistencies:
  - a) **Synonym**: Using different names for same data item  
Example: Stu-Id, Reg-No.
  - b) **Homonym**: Using same name for different data items.  
Example: “course” for a single subject or entire program.

# Drawbacks of File System

---

## 4. Excessive Program Maintenance

- Physical structure and storage of files are stored in the application code. Against any change in the data, all the application using that data are needed to be changed. Such applications are called **Data Dependent Applications**.
- File descriptions are stored within each application program that accesses a given file (a collection of related records). Any change to a file structure requires changes to the file description for all programs that access the file.
- Often difficult to locate all programs affected by change
- Errors are often introduced when making changes.
- **Incompatible file structures.**

# Drawbacks of File System

---

## 5. Limited Data Sharing:

- As each application has its own private files so little opportunity to share data with others.
- The filing system works well when we have to interact with each file separately. However, it breaks down when we have to **cross-reference** or process the information in the files. This makes it hard to answer queries involving multiple files.

## 6. Inflexible

## 7. Lack of provision for security

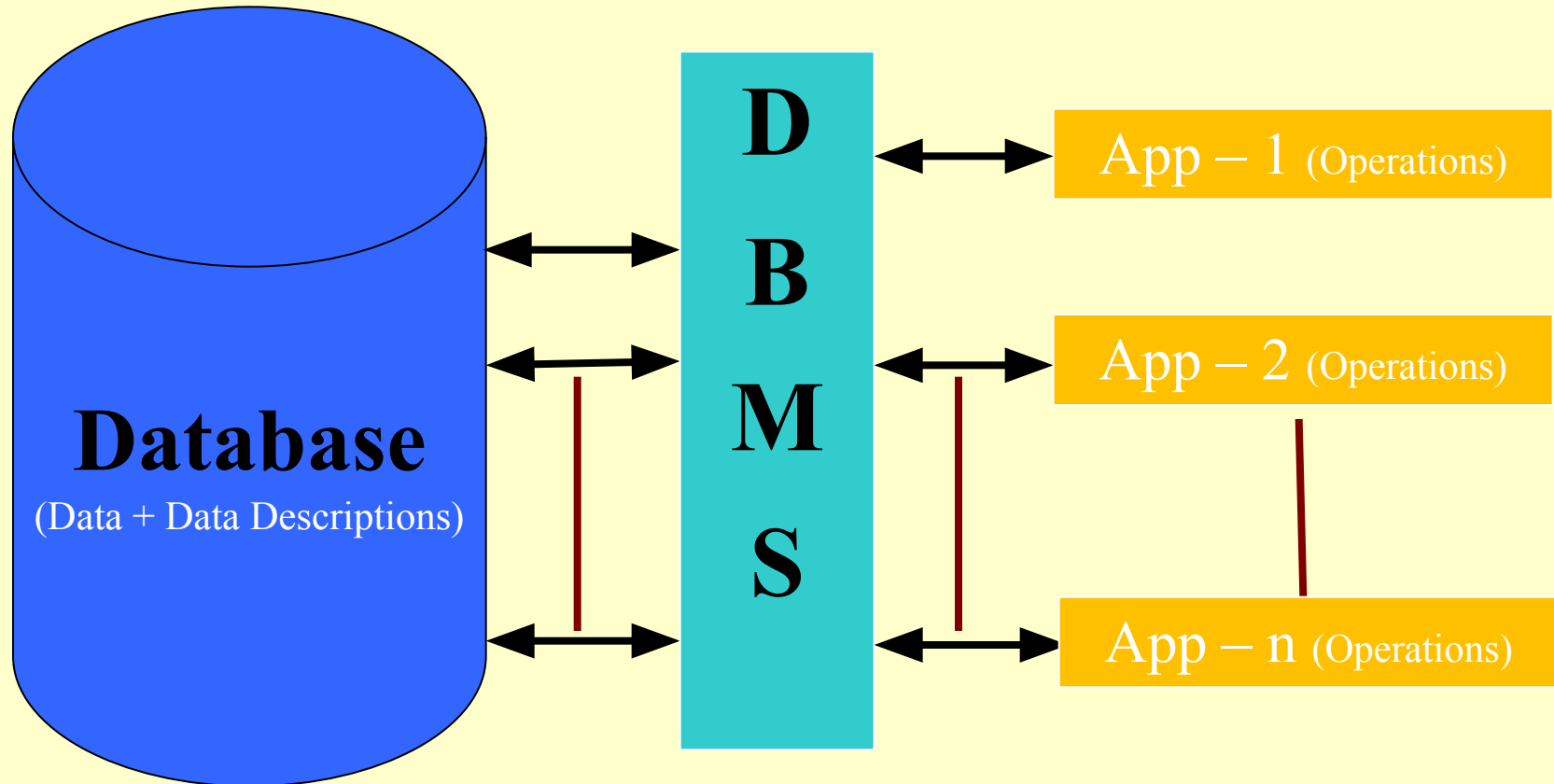
## 8. Limited recovery from failure

# Integrated Database Environment(IDE)

---

- The limitations of file-based approach can be attributed to the following factors:
  - Redundancy of data.
  - Data dependence: the definition of data is embedded in the application programs, rather than being stored separately and independently.
- In contrast to file processing system, the IDE has a single large repository of data.
- Emphasizes the integration & sharing of data throughout the organization.
- The organization wide requirements are analyzed as a whole and there is no longer concept of MY FILE or Private Files.
- A Database Management System (DBMS) is a software designed to assist in maintaining and utilizing large collections of data.
- By storing data in a DBMS rather than as a collection of files, we achieve enormous advantages.

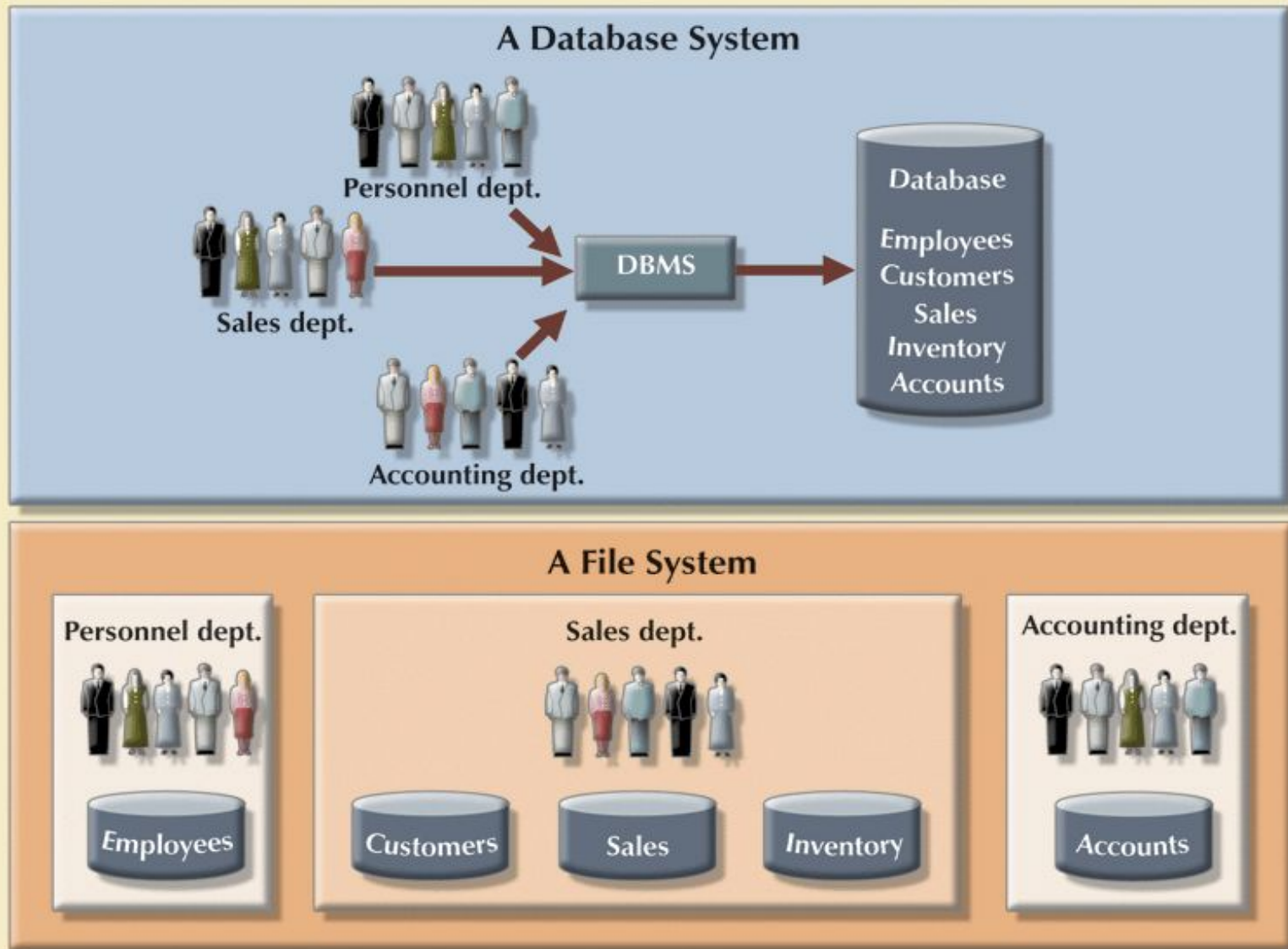
# Database System



*DBMS manages data resources like  
an operating system manages  
hardware resources*

**FIGURE  
1.6**

## Contrasting database and file systems





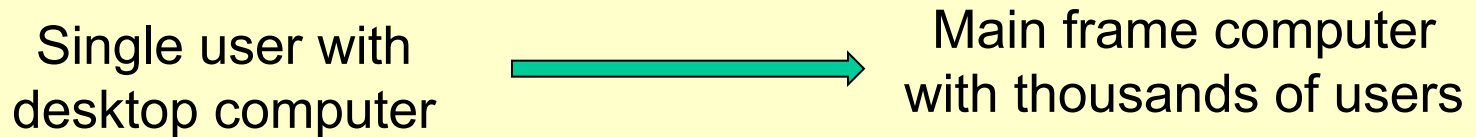
# Advantages of Database Approach

---

1. Reduced or Controlled Data Redundancy
2. Improved Data Consistency
3. Enforcement of Standards
4. Reduced Program Maintenance
5. Data Sharing
6. Data Integrity (Improved Data Quality)
7. Improved Security Restrictions
8. Improved Accessibility & Responsiveness
9. Improved Decision Making
10. Data Independence

# The Range of Database Applications

---



- Categories are:
  - Personal computer databases(Single User)
  - Workgroup databases(Multi-user )
  - Department databases (Multi-user )
  - Enterprise database (Multi-user )
  - Internet, Intranet & Extranet DBs (Multi-user )

# The Range of Database Applications

---

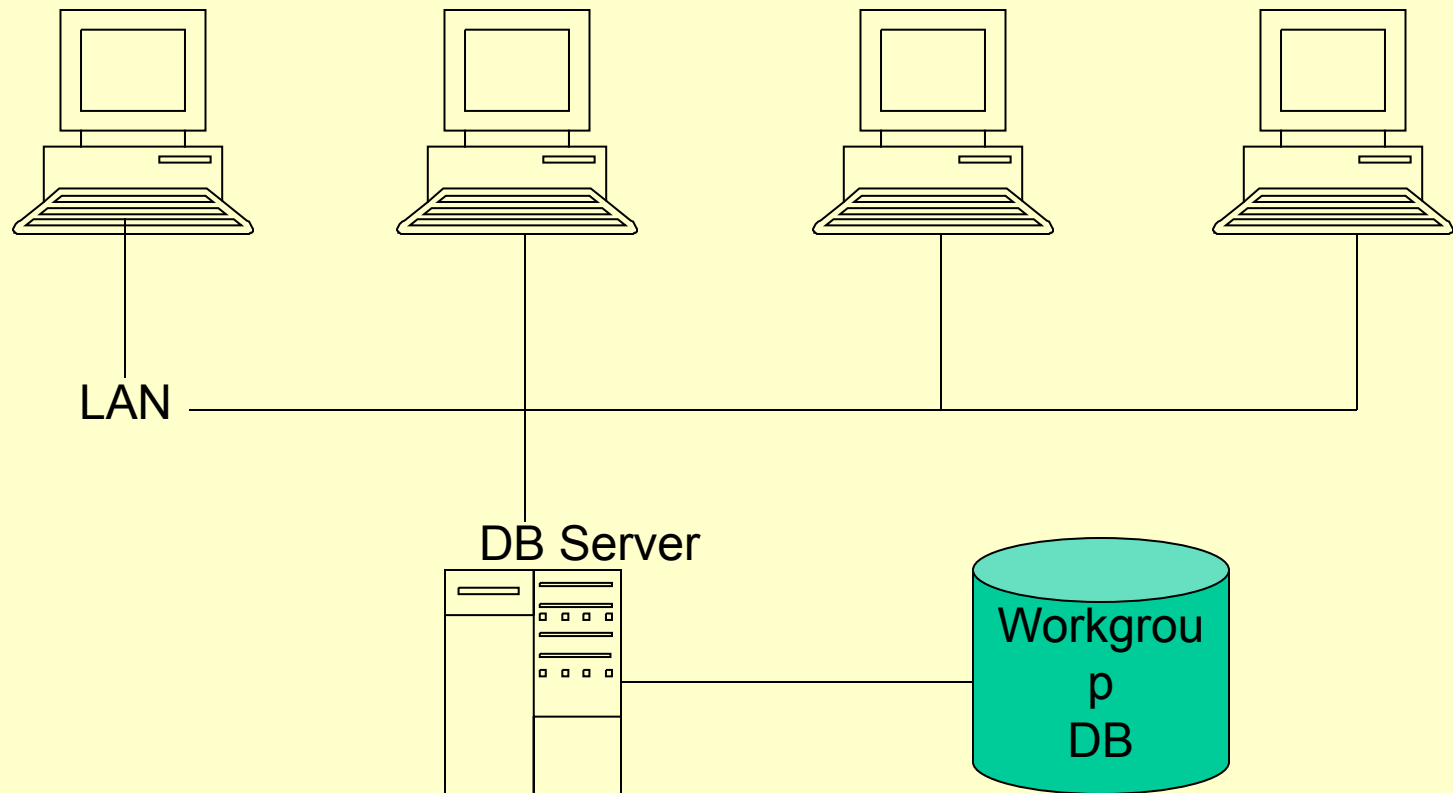
## Personal Computer Databases

- Support one user with a standalone PC
- E.g. a student's own database or a sales person's simple database

## Workgroup Databases

- Workgroup: relatively small group of people who collaborate on same project/application.
- A workgroup DB supports the collaborative efforts of a workgroup.
- Allows data sharing.
- Its model is shown on the next fig:

## Continued...



**Workgroup DB on LAN**  
(Method of Data Sharing)

# The Range of Database Applications

---

## Department Databases

- Department: Functional unit within an organization
- Department DBs support function & activities of a department.
- E.g. personnel DB to track employees, jobs, skills etc.

## Enterprise Databases

- DB scope is the entire organization
- To support organization-wide operations & decision making.

## Continued...

