Database Systems: Design, Implementation, and Management

Lesson 1

Objectives

In this lesson, you will learn:

- The difference between data and information
- What a database is, the various types of databases, and why they are valuable assets for decision making
- The importance of database design
- How modern databases evolved from file systems

Objectives (cont'd.)

- About flaws in file system data management
- The main components of the database system
- The main functions of a database management system (DBMS)

Introduction

- Good decisions require good information derived from raw facts
- Data is managed most efficiently when stored in a database
- Databases evolved from computer file systems
- Understanding file system characteristics is important

Why Databases?

- Databases solve many of the problems encountered in data management
  - Used in almost all modern settings involving data management:
    - Business
    - Research
    - Administration
- Important to understand how databases work and interact with other applications

Data vs. Information

- **Data** are raw facts
- **Information** is the result of processing raw data to reveal meaning
- Information requires context to reveal meaning
- Raw data must be formatted for storage, processing, and presentation
- Data are the foundation of information, which is the bedrock of **knowledge**

Data vs. Information (cont'd.)

- **Data**: building blocks of information
- Information produced by processing data
- Information used to reveal meaning in data
- Accurate, relevant, timely information is the key to good decision making
- Good decision making is the key to organizational survival

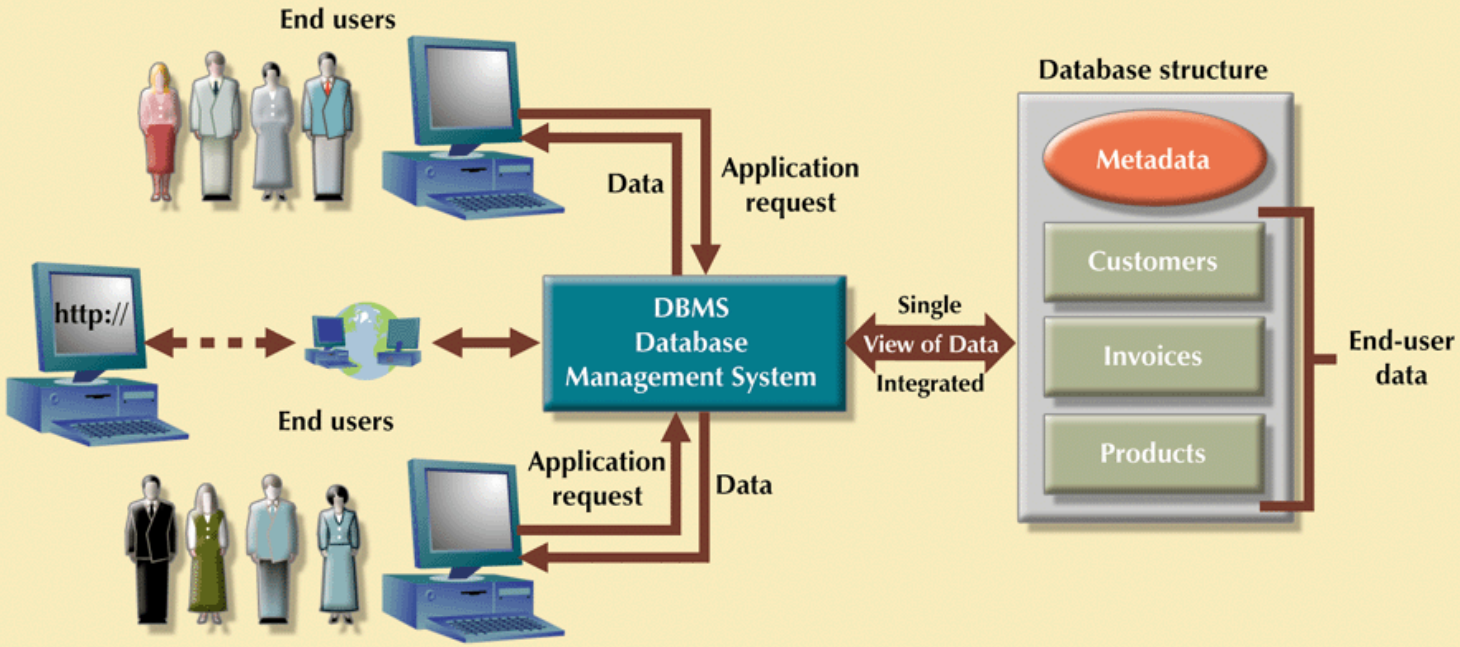
Introducing the Database

- Database: shared, integrated computer structure that stores a collection of:
  - End-user data: raw facts of interest to end user
  - **Metadata**: data about data
    - Provides description of data characteristics and relationships in data
    - Complements and expands value of data
- **Database management system (DBMS)**: collection of programs
  - Manages structure and controls access to data

Role and Advantages of the DBMS

- DBMS is the intermediary between the user and the database
  - Database structure stored as file collection
  - Can only access files through the DBMS
- DBMS enables data to be shared
- DBMS integrates many users' views of the data

FIGURE 1.2 The DBMS manages the interaction between the end user and the database



Role and Advantages of the DBMS (cont'd.)

- Advantages of a DBMS:
  - Improved data sharing
  - Improved data security
  - Better data integration
  - Minimized **data inconsistency**
  - Improved data access
  - Improved decision making
  - Increased end-user productivity

Types of Databases

- **Centralized database**: data located at a single site
- **Distributed database**: data distributed across several different sites
- **Operational database**: supports a company's day-to-day operations
  - Transactional or production database
- **Data warehouse**: stores data used for tactical or strategic decisions

TABLE 1.1 Types of Databases

PRODUCT	NUMBER OF USERS			DATA LOCATION		DATA USAGE		XML
	SINGLE USER	MULTIUSER		CENTRALIZED	DISTRIBUTED	OPERATIONAL	DATA WAREHOUSE	
		WORKGROUP	ENTERPRISE					
MS Access	X	X		X		X		
MS SQL Server	X[3]	X	X	X	X	X	X	X
IBM DB2	X[3]	X	X	X	X	X	X	X
MySQL	X	X	X	X	X	X	X	X*
Oracle RDBMS	X[3]	X	X	X	X	X	X	X

* Supports XML functions only. XML data are stored in large text objects.

Why Database Design Is Important

- **Database design** focuses on design of database structure used for end-user data
  - Designer must identify database's expected use
- Well-designed database:
  - Facilitates data management
  - Generates accurate and valuable information
- Poorly designed database:
  - Causes difficult-to-trace errors

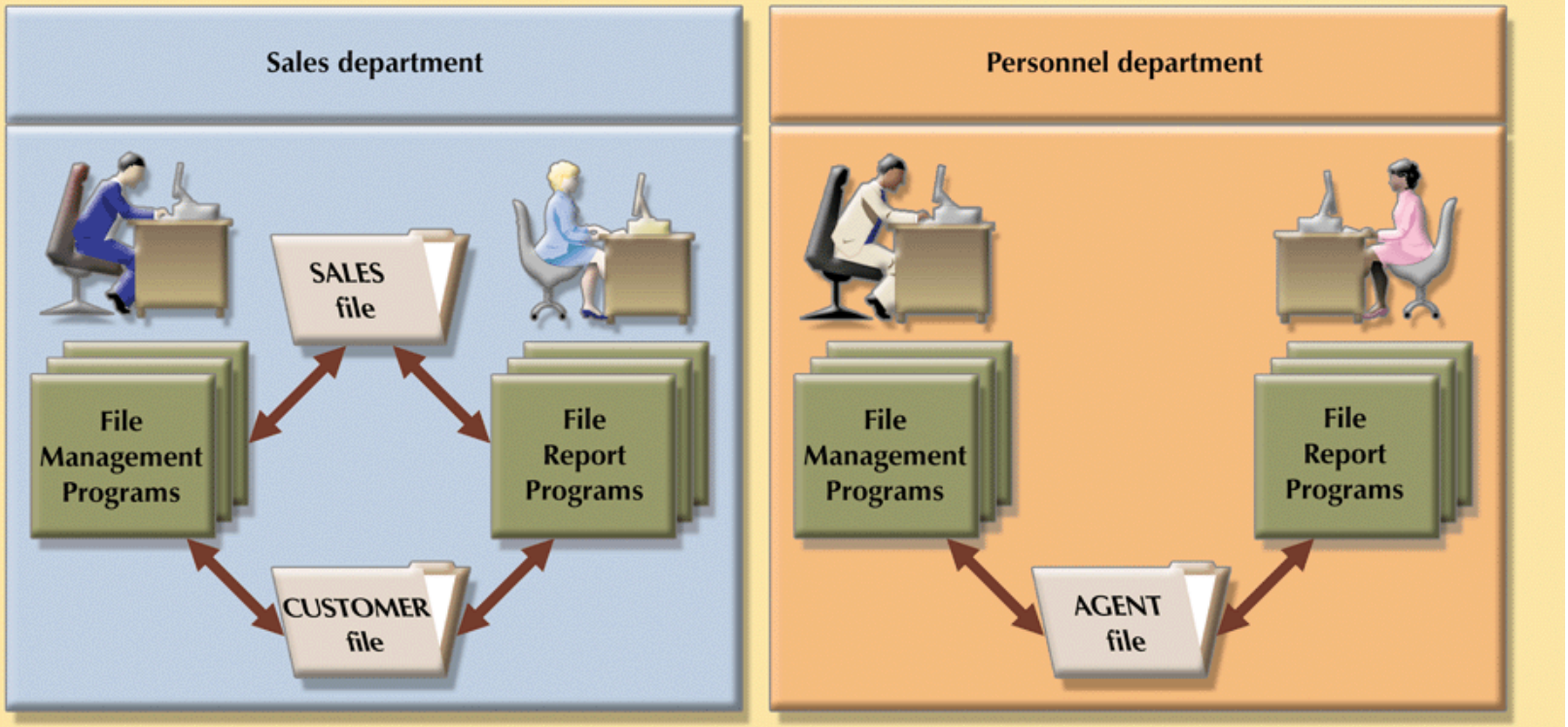
Evolution of File System Data Processing

- Reasons for studying file systems:
  - Complexity of database design is easier to understand
  - Understanding file system problems helps to avoid problems with DBMS systems
  - Knowledge of file system is useful for converting file system to database system
- File systems typically composed of collection of file folders, each tagged and kept in cabinet
  - Organized by expected use

TABLE 1.2 Basic File Terminology

TERM	DEFINITION
Data	"Raw" facts, such as a telephone number, a birth date, a customer name, and a year-to-date (YTD) sales value. Data have little meaning unless they have been organized in some logical manner.
Field	A character or group of characters (alphabetic or numeric) that has a specific meaning. A field is used to define and store data.
Record	A logically connected set of one or more fields that describes a person, place, or thing. For example, the fields that constitute a record for a customer might consist of the customer's name, address, phone number, date of birth, credit limit, and unpaid balance.
File	A collection of related records. For example, a file might contain data about the students currently enrolled at Gigantic University.

FIGURE 1.5 A simple file system



Problems with File System Data Processing

- Summary of file system limitations:
  - Requires extensive programming
  - Cannot perform ad hoc queries
  - System administration is complex and difficult
  - Difficult to make changes to existing structures
  - Security features are likely to be inadequate

Data Redundancy

- File system structure makes it difficult to combine data from multiple sources
  - Vulnerable to security breaches
- Organizational structure promotes storage of same data in different locations
  - Islands of information
- Data stored in different locations is unlikely to be updated consistently
- **Data redundancy**: same data stored unnecessarily in different places

Data Redundancy (cont'd.)

- **Data inconsistency**: different and conflicting versions of same data occur at different places
- **Data anomalies**: abnormalities when all changes in redundant data are not made correctly
  - Update anomalies
  - Insertion anomalies
  - Deletion anomalies

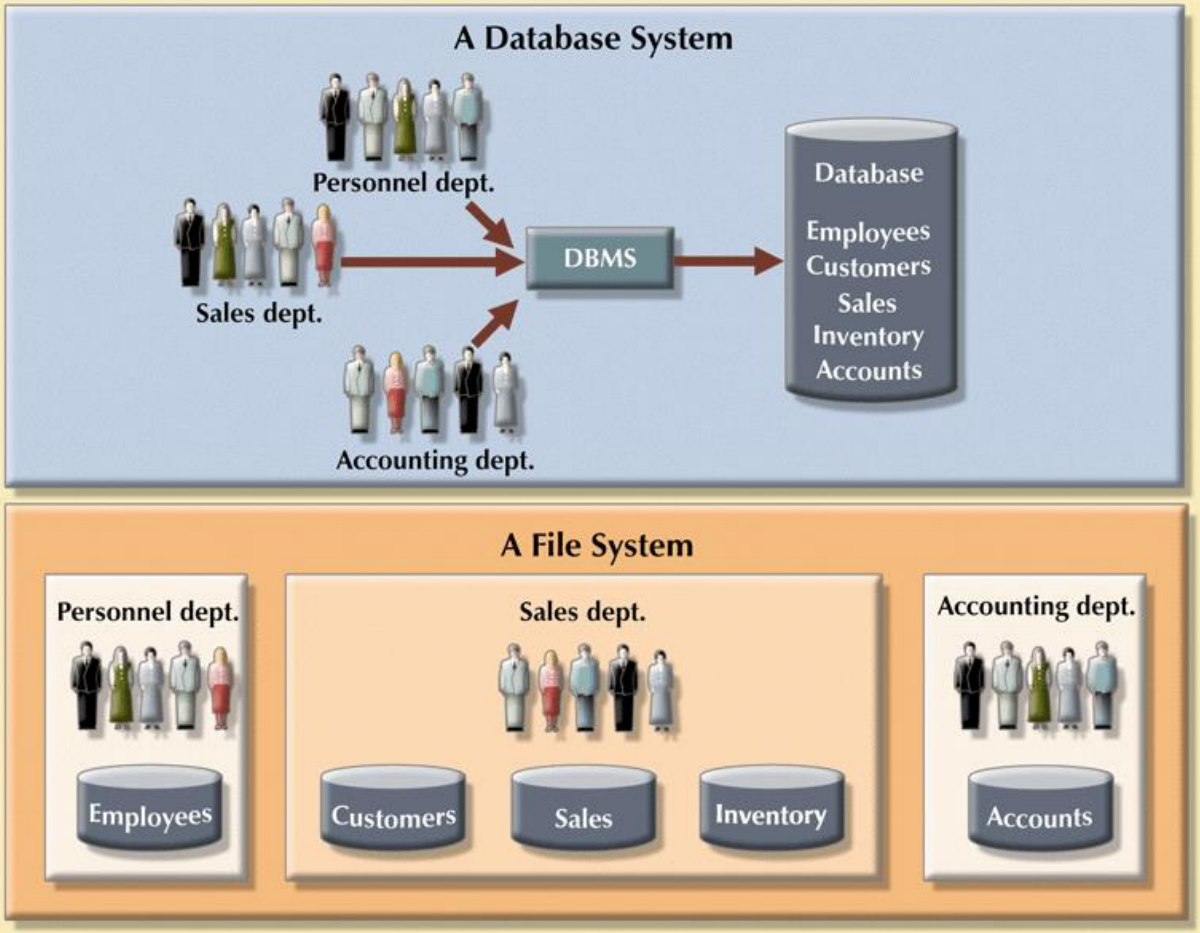
Lack of Design and Data-Modeling Skills

- Most users lack the skill to properly design databases, despite multiple personal productivity tools being available
- Data-modeling skills are vital in the data design process
- Good data modeling facilitates communication between the designer, user, and the developer

Database Systems

- Database system consists of logically related data stored in a single logical data repository
  - May be physically distributed among multiple storage facilities
  - DBMS eliminates most of file system's problems
  - Current generation stores data structures, relationships between structures, and access paths
    - Also defines, stores, and manages all access paths and components

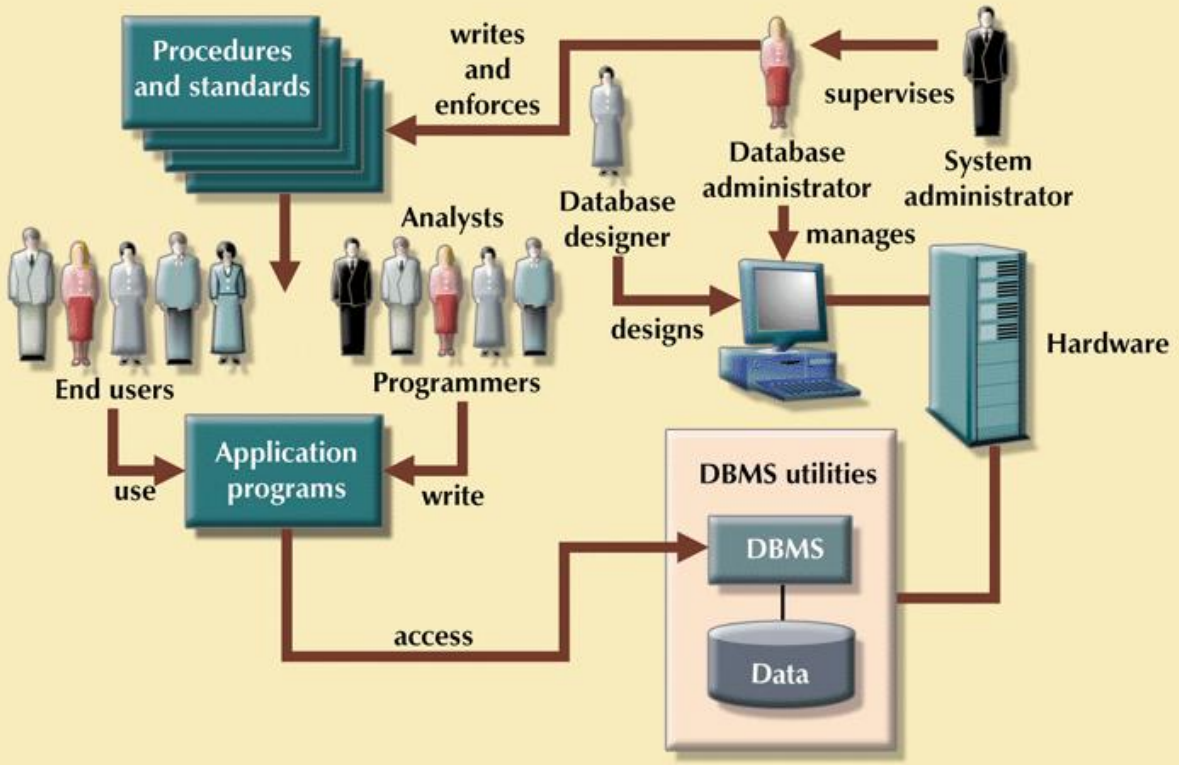
FIGURE 1.6 Contrasting database and file systems



The Database System Environment

- **Database system**: defines and regulates the collection, storage, management, use of data
- Five major parts of a database system:
  - Hardware
  - Software
  - People
  - Procedures
  - Data

FIGURE 1.7 The database system environment



DBMS Functions

- Most functions are transparent to end users
  - Can only be achieved through the DBMS
- Data dictionary management
  - DBMS stores definitions of data elements and relationships (metadata) in a **data dictionary**
  - DBMS looks up required data component structures and relationships
  - Changes automatically recorded in the dictionary
  - DBMS provides data abstraction and removes structural and data dependency

DBMS Functions (cont'd.)

- Data storage management
  - DBMS creates and manages complex structures required for data storage
  - Also stores related data entry forms, screen definitions, report definitions, etc.
  - **Performance tuning**: activities that make the database perform more efficiently
  - DBMS stores the database in multiple physical data files

DBMS Functions (cont'd.)

- Data transformation and presentation
  - DBMS transforms data entered to conform to required data structures
  - DBMS transforms physically retrieved data to conform to user's logical expectations
- Security management
  - DBMS creates a security system that enforces user security and data privacy
  - Security rules determine which users can access the database, which items can be accessed, etc.

DBMS Functions (cont'd.)

- Multiuser access control
  - DBMS uses sophisticated algorithms to ensure concurrent access does not affect integrity
- Backup and recovery management
  - DBMS provides backup and data recovery to ensure data safety and integrity
  - Recovery management deals with recovery of database after a failure
    - Critical to preserving database's integrity

DBMS Functions (cont'd.)

- Data integrity management
  - DBMS promotes and enforces integrity rules
    - Minimizes redundancy
    - Maximizes consistency
  - Data relationships stored in data dictionary used to enforce data integrity
  - Integrity is especially important in transaction-oriented database systems

DBMS Functions (cont'd.)

- Database access languages and application programming interfaces
  - DBMS provides access through a query language
  - **Query language** is a nonprocedural language
  - **Structured Query Language (SQL)** is the de facto query language
    - Standard supported by majority of DBMS vendors