Introducing the Database

- **Data versus Information**
  - Data constitute building blocks of information
  - Information reveals meaning of data
  - Good, timely, relevant information key to decision making
  - Good decision making key to organizational survival
- **Database Management System (DBMS)**
  - Manages Database structure
  - Controls access to data
  - Contains query language

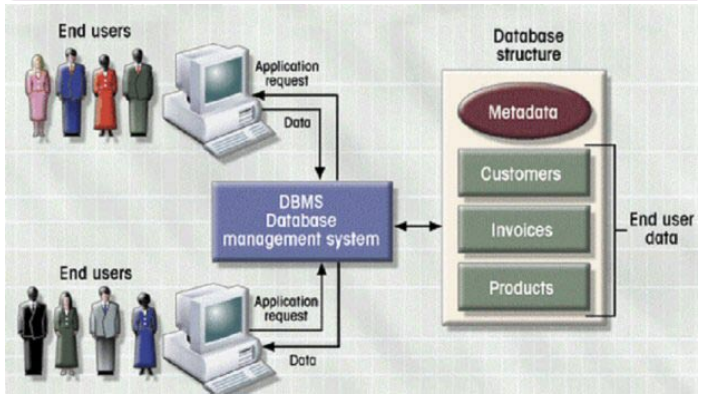
 1

Importance of DBMS

- It helps make data management more efficient and effective.
- Its query language allows quick answers to ***ad hoc*** queries.
- It provides ***end users*** better access to more and better-managed data.
- It promotes an integrated view of the organization's operations "big picture."
- It reduces the probability of inconsistent data.

 2

Figure 1.2: The DBMS Manages the Interaction between the End User and the Database



 3

Why Database Design Is Important?

- A well-designed database facilitates data management and becomes a valuable information generator.
- A poorly designed database is a breeding ground for uncontrolled data redundancies.
- A poorly designed database generates errors that lead to bad decisions.
- Poorly Designed databases are "self-correcting". Organization fails! ☹
- You can make a big salary! ☺

 4

File Systems - Historical Roots

- First applications focused on clerical tasks
  - Order Entry
  - Work Scheduling
  - Accounting and Payroll
- Requests for information quickly followed
- File systems developed for needs
  - Data organized according to expected use
  - Data Processing (DP) specialists computerized manual file systems
- File System's characteristics facilitate complex database understanding

Figure 1.3: Contents of the CUSTOMER File

C_NAME	C_PHONE	C_ADDRESS	C_ZIP	A_NAME	A_PHONE	TP	AMT	REN
Alfred A. Ramas	615-844-2573	218 Fork Rd., Babs, TN	36123	Leah F. Hahn	615-882-1244	T1	$100.00	3/12/99
Leona K. Dunne	713-894-1238	Box 12A, Fox, KY	25246	Alex B. Alby	713-228-1249	T1	$250.00	5/23/99
Kathy W. Smith	615-894-2285	125 Oak Ln, Babs, TN	36123	Leah F. Hahn	615-882-2144	S2	$150.00	1/5/98
Paul F. Olowski	615-894-2180	217 Lee Ln., Babs, TN	36123	Leah F. Hahn	615-882-1244	S1	$300.00	9/20/99
Myron Orlando	615-222-1672	Box 111, New, TN	36155	Alex B. Alby	713-228-1249	T1	$100.00	12/4/99
Amy B. O'Brian	713-442-3381	387 Troll Dr., Fox, KY	25246	John T. Okon	615-123-5589	T2	$850.00	8/29/99
James G. Brown	615-297-1228	21 Tye Rd., Nash, TN	37118	Leah F. Hahn	615-882-1244	S1	$120.00	3/1/99
George Williams	615-290-2556	155 Maple, Nash, TN	37119	John T. Okon	615-123-5589	S1	$250.00	6/23/99
Anne G. Farriss	713-382-7185	2119 Elm, Crew, KY	25432	Alex B. Alby	713-228-1249	T2	$100.00	11/9/99
Olette K. Smith	615-297-3809	2782 Main, Nash, TN	37118	John T. Okon	615-123-5589	S2	$500.00	2/18/99

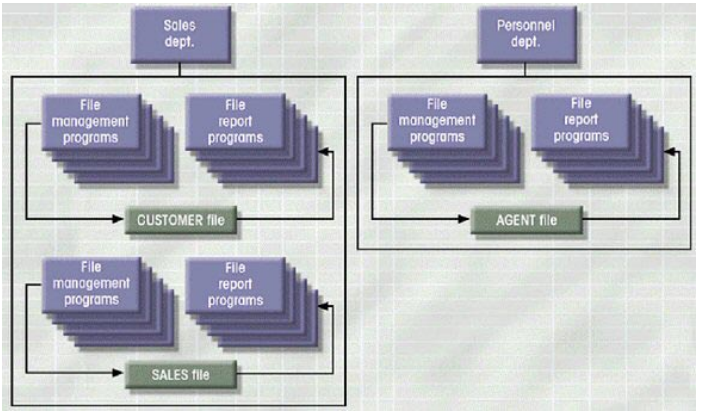
C_NAME = Customer name
C_PHONE = Customer phone
C_ADDRESS = Customer address
C_ZIP = Customer ZIP code

A_NAME = Agent name
A_PHONE = Agent phone
TP = Insurance type
AMT = Insurance policy amount, in thousands of $
REN = Insurance renewal date

Data Organization

- **Data Value** (Cells)
  - Contents of a field contained in a record
  - "Raw Facts" that can be recognized
- **Fields** or **Attributes** (Columns)
  - Group of characters representing something
- **Records** or **Entities** or Tuples (Rows)
  - Collection of related fields
- **Tables** or **Entity Set** (File)
  - Collection of related records

Figure 1.5: Simple File System



File System Environment

- **Information System approach**
  - **Bottom-Up Process**
  - **Individual departments design and control data**
  - **Organization needs must utilize this dissimilar data**
- **Disadvantages**
  - **(a) uncontrolled redundancy -- duplicate data**
  - **(b) inconsistencies – lack of data integrity**
  - **(c) inflexibility -- can't respond to new requests**
  - **(d) limited data sharing -- increases data duplication**
  - **(e) poor enforcement of standards -- decentralized file design**
  - **(f) excessive program maintenance -- modify data means change to program**

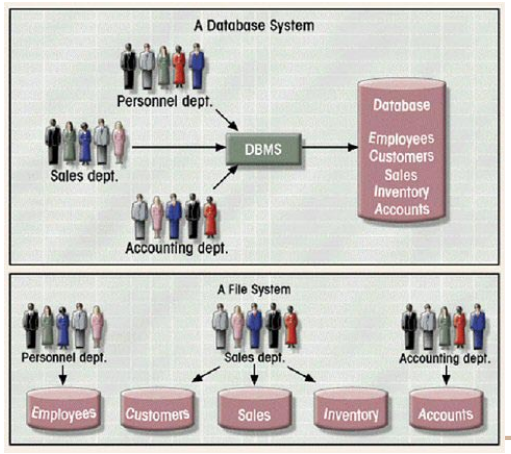
 9

File System Critique

- **File System Data Management**
  - **File systems require extensive programming in a third-generation language (3GL, i.e. COBOL).**
  - **As the number of files expands, system administration becomes difficult.**
    - **Making changes**
    - **Security difficult to program**
  - **Ad hoc queries impossible**
- **Structural and Data Dependence**
  - **Structural Dependence**
    A change in any file's structure requires the modification of all programs using that file.
  - **Data Dependence**
    A change in any file's data characteristics requires changes in all data access programs.

 10

Figure 1.6: Database vs. File Systems



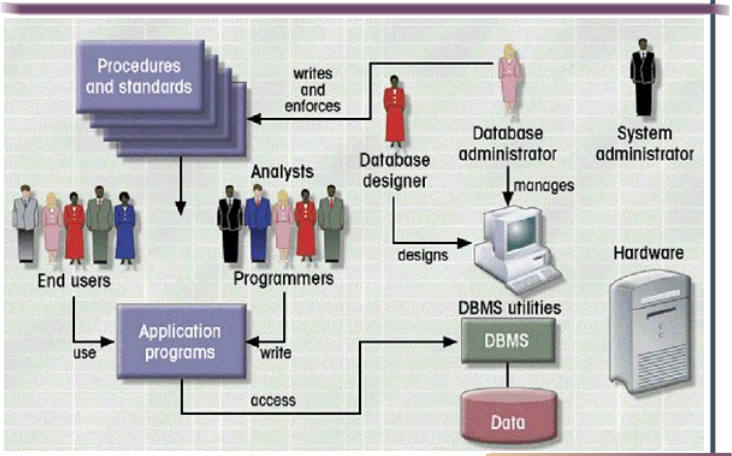
 11

Database Systems

- **Hardware = PC, Web Server, Mainframe**
- **Software = Access, MS SQL, Oracle, IBM DB2**
- **People**
  - **Systems administrators**
  - **Database administrators (DBAs)**
  - **Database designers**
  - **Systems analysts and programmers**
  - **End users**
- **Procedures**
  - **Instructions and rules that govern the design and use of the database system**
- **Data**
  - **Collection of facts stored in the database**

 12

Figure 1.7: Database System Environment



 13

Types of Database Systems

- **Number of Users & Scope**
  - **Single-user**
    - **Desktop database**
  - **Multiuser**
    - **Workgroup database (< 50 user)**
    - **Enterprise database (Used by many departments)**
- **Location**
  - **Centralized (DBMS resides at one site)**
  - **Distributed (Processing and storage)**
- **Use**
  - **Transactional (Production)**
  - **Decision support**
  - **Data warehouse**

 14

DBMS Functions

1. **Data Dictionary Management**
2. **Data Storage Management**
3. **Data Transformation and Presentation**
4. **Security Management**
5. **Multi-User Access Control**
6. **Backup and Recovery Management**
7. **Data Integrity Management**
8. **Database Access Languages and Application Programming Interfaces**
9. **Database Communication Interfaces**

 15

Database Model

Collection of logical constructs used to represent the data structure and the data relationships found within the database.

- **Conceptual models focus on *what* is represented rather than *how* it is represented.**
  - **Entity Relationship Diagram**
  - **Object Oriented Model**
- **Implementation models emphasis on *how* the data is represented in the database or on *how* the data structures are implemented.**
  - **Hierarchical Database Model**
  - **Relational Database Model**
  - **Object Oriented Database Model**

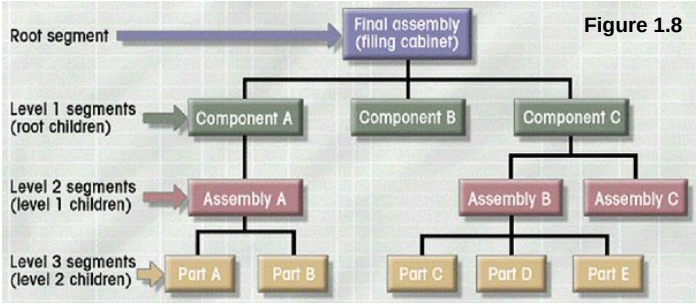
 16

Database Conceptual Model

- **Three Types of Relationships**
  - **One-to-many relationships (1:M)**
    - **A painter paints many different paintings, but each one of them is painted by only that painter.**
      - **PAINTER (1) paints PAINTING (M)**
  - **Many-to-many relationships (M:N)**
    - **An employee might learn many job skills, and each job skill might be learned by many employees.**
      - **EMPLOYEE (M) learns SKILL (N)**
  - **One-to-one relationships (1:1)**
    - **Each store is managed by a single employee and each store manager (employee) only manages a single store.**
      - **EMPLOYEE (1) manages STORE (1)**

Implementation Model: Hierarchical Database

- **Logically represented by an upside down tree**
  - **Each parent can have many children**
  - **Each child has only one parent**



Hierarchical Database

- **Advantages**
  - **Conceptual simplicity – Relationships defined**
  - **Database security – Uniform throughout system**
  - **Data independence – Data type cascaded**
  - **Database integrity – Child referenced to parent**
  - **Efficiency – Parent to Child (One to Many)**
- **Disadvantages**
  - **Complex implementation**
  - **Difficult to manage**
  - **Lacks structural independence**
  - **Applications programming and use is complex**
  - **Implementation limitations (Many to Many)**
  - **Lack of standards**

Implementation Model: Relational Database

- **Basic Structure**
  - **Relational DataBase Management Systems (RDBMS) allows operations in a human logical environment**
  - **The relational database is perceived as a collection of tables.**
  - **Each table consists of a series of row/column intersections.**
  - **Tables (or relations) are related to each other by sharing a common entity characteristic**
  - **The relationship type shown in a *relational schema***
  - **A table yields data and structural independence**
  - **Microsoft Access is a RDBMS**

Relational Database Model

Figure 1.11

Table name: AGENT

AGENT_CODE	AGENT_LNAME	AGENT_FNAME	AGENT_INITIAL	AGENT_AREACODE	AGENT_PHONE
501	Alby	Alex	B	713	228-1249
502	Hahn	Leah	F	615	882-1244
503	Okon	John	T	615	123-5589

Link through AGENT code

Table name: CUSTOMER

CUS_CODE	CUS_LNAME	CUS_FNAME	CUS_INITIAL	CUS_AREACODE	CUS_PHONE	CUS_RENEW_DATE	AGENT_CODE
10010	Ramas	Alfred	A	615	844-2573	05-Apr-2002	502
10011	Dunne	Leona	K	713	894-1238	16-Jun-2002	501
10012	Smith	Kathy	W	615	894-2285	29-Jan-2001	502
10013	Olowski	Paul	F	615	894-2180	14-Oct-2002	502
10014	Orlando	Myron		615	222-1672	28-Dec-2002	501
10015	O'Brian	Amy	B	713	442-3381	22-Sep-2002	503
10016	Brown	James	G	615	297-1228	25-Mar-2002	502
10017	Williams	George		615	290-2556	17-Jul-2002	503
10018	Farriss	Anne	G	713	382-7185	03-Dec-2002	501
10019	Smith	Olette	K	615	297-3809	14-Mar-2002	503

 21

Relational Database Model

- **Advantages**
  - Structural independence
  - Improved conceptual simplicity
  - Easier database design, implementation, management, and use
  - Ad hoc query capability (SQL)
  - Powerful database management system
  - Most common DBMS used today
- **Disadvantages**
  - Substantial hardware and system software overhead
  - Possibility of poor design and implementation
  - Potential "islands of information" = local DB

 22

Conceptual Model: Entity Relationship

- E-R models are normally represented in an Entity Relationship Diagram (ERD).
- An entity is represented by a rectangle.
  - Usually a Noun or Object of the sentence.
- A relationship is represented by a diamond connected to the related entities.
  - Usually a Verb.
- An attribute is a characteristic of the entity.
  - Represented by ellipses connected to entity.
  - Usually Nouns

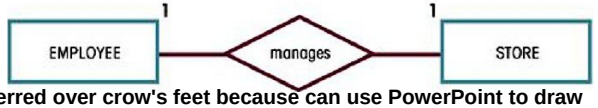
 23

Figure 1.13

A One-to-Many (1:M) Relationship: A PAINTER can paint many PAINTINGs; each PAINTING is painted by one PAINTER

A Many-to-Many (M:N) Relationship: an EMPLOYEE can learn many SKILLs; each SKILL can be learned by many EMPLOYEEs

A One-to-One (1:1) Relationship: an EMPLOYEE manages one STORE; each STORE is managed by one EMPLOYEE



Note: Preferred over crow's feet because can use PowerPoint to draw

Entity Relationship Model

- **Advantages**
  - Exceptional conceptual simplicity
  - Visual representation
  - Effective communication tool
  - Integrated with the relational database model
- **Disadvantages**
  - Limited constraint representation
  - Limited relationship representation
  - No data manipulation language
  - Loss of information content

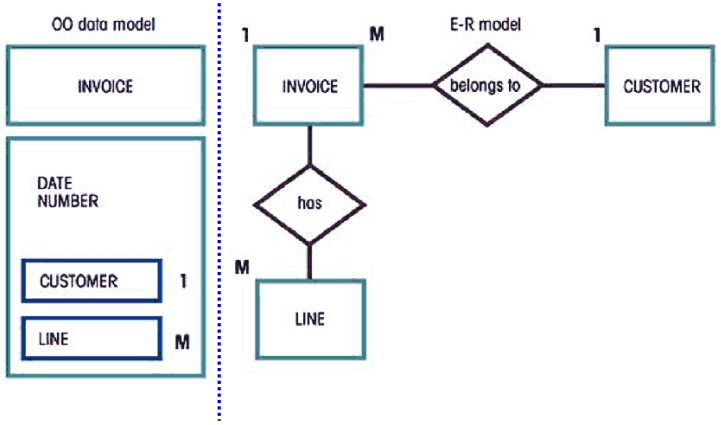
 25

Implementation Model: Object-Oriented DB

- **Basic Structure**
  - **Objects** are abstractions of actual entities.
  - **Attributes** are properties of an object.
  - A **Class** is a collection of similar objects with shared structure (attributes) and behavior (methods).
  - Classes are organized in a **class hierarchy**.
  - An object can **inherit** the attributes and methods of the classes above it.

 26

Figure 1.15: A Comparison: The OO Data Model and the ER Model



Object-Oriented Database Model

- **Advantages**
  - Visual presentation
  - Database integrity
  - Both structural and data independence
  - Object Oriented Method with Class Inheritance
- **Disadvantages**
  - Lack of Object Oriented Data Model standards
  - Complex navigational data access
  - Steep learning curve
  - High system overhead slows transactions

 28