



Chapter

5

Data Resource Management

Foundation Concept

Database Definition

- A usually large collection of data organized especially for rapid search and retrieval
- A database is an integrated collection of logically related data elements.
- A database consolidates records previously stored in separate files into a common pool of data

Foundation Concept

Database Management System:

- While a database itself is a collection of several related files.
- *The program used to build databases, populate them with data, and manipulate the data is called a **database management system (DBMS)**.*

Foundation Concept

- Problems with the traditional file environment (files maintained separately by different departments)
 - Data redundancy:
 - Presence of duplicate data in multiple files
 - Data inconsistency:
 - Same attribute has different values
 - Program-data dependence:
 - When changes in program requires changes to data accessed by program
 - Lack of flexibility
 - Poor security
 - Lack of data sharing and availability

Foundation Concept

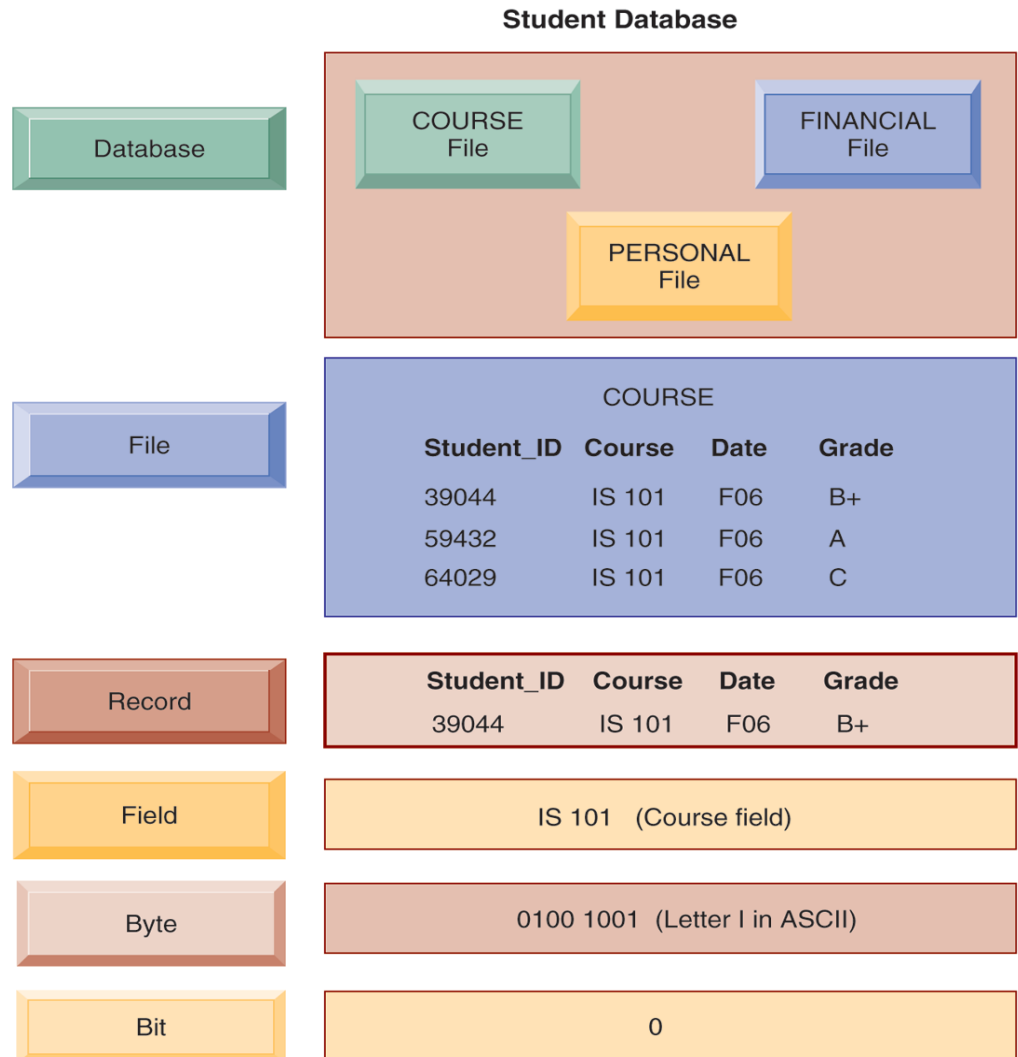
- **Information Organization**

- Database: Group of related files
- File: Group of records of same type
- Record: Group of related fields
- Field: Group of characters as word(s) or number
 - Describes an **entity** (person, place, thing on which we store information)
 - **Attribute:** Each characteristic, or quality, describing entity
 - E.g., Attributes Date or Grade belong to entity COURSE

Foundation Concept

- THE DATA HIERARCHY

A computer system organizes data in a hierarchy that starts with the bit, which represents either a 0 or a 1. Bits can be grouped to form a byte to represent one character, number, or symbol. Bytes can be grouped to form a field, and related fields can be grouped to form a record. Related records can be collected to form a file, and related files can be organized into a database.



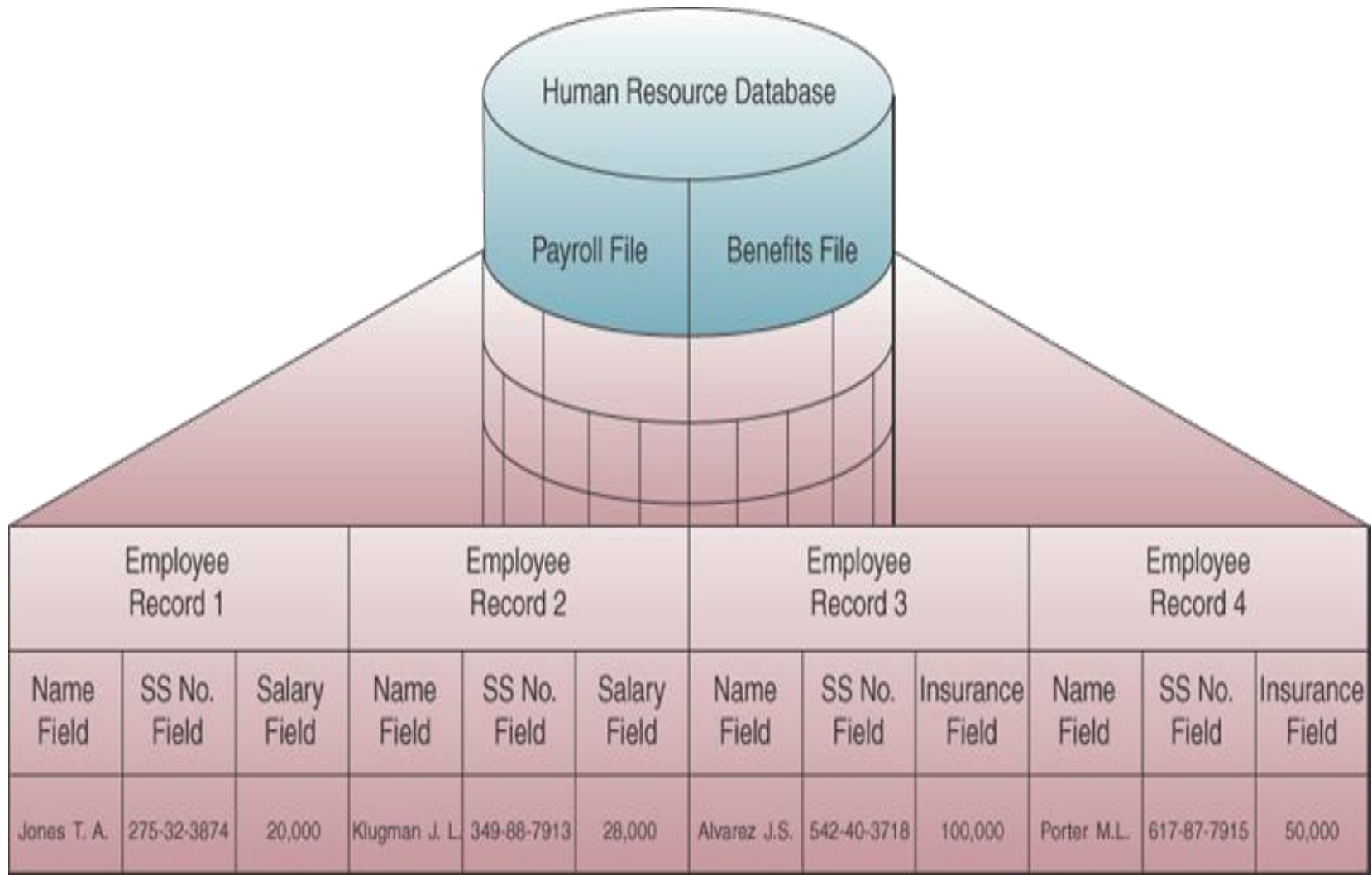
Logical Data Elements

- Character
 - A single alphabetic, numeric, or other symbol
- Field or data item
 - Represents an attribute (characteristic or quality) of some entity (object, person, place, event)
 - Examples: salary, job title
- Record
 - Grouping of all the fields used to describe the attributes of an entity
 - Example: payroll record with name, SSN, pay rate

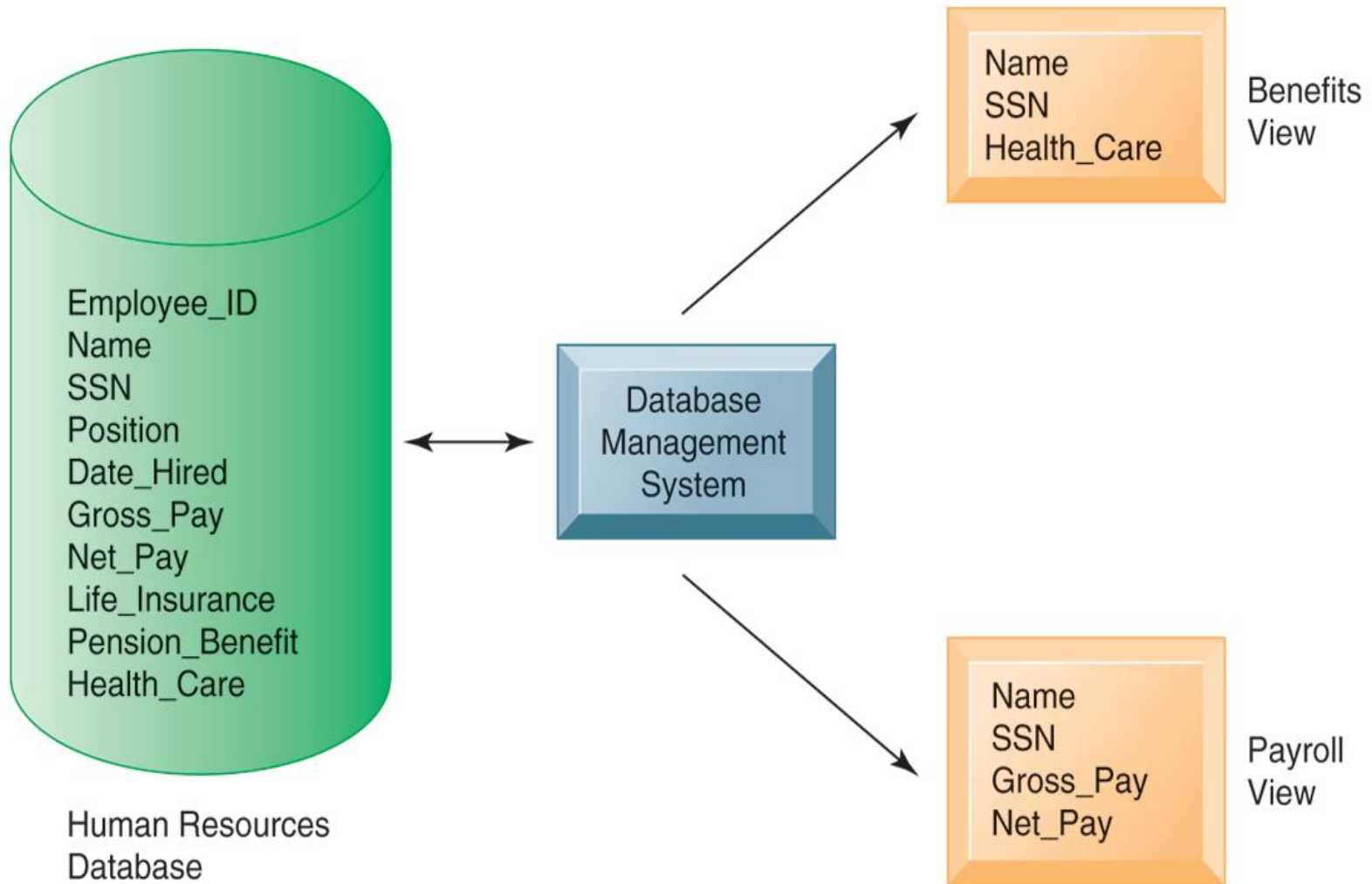
Logical Data Elements

- File or table
 - A group of related records
- Database
 - An integrated collection of logically related data elements

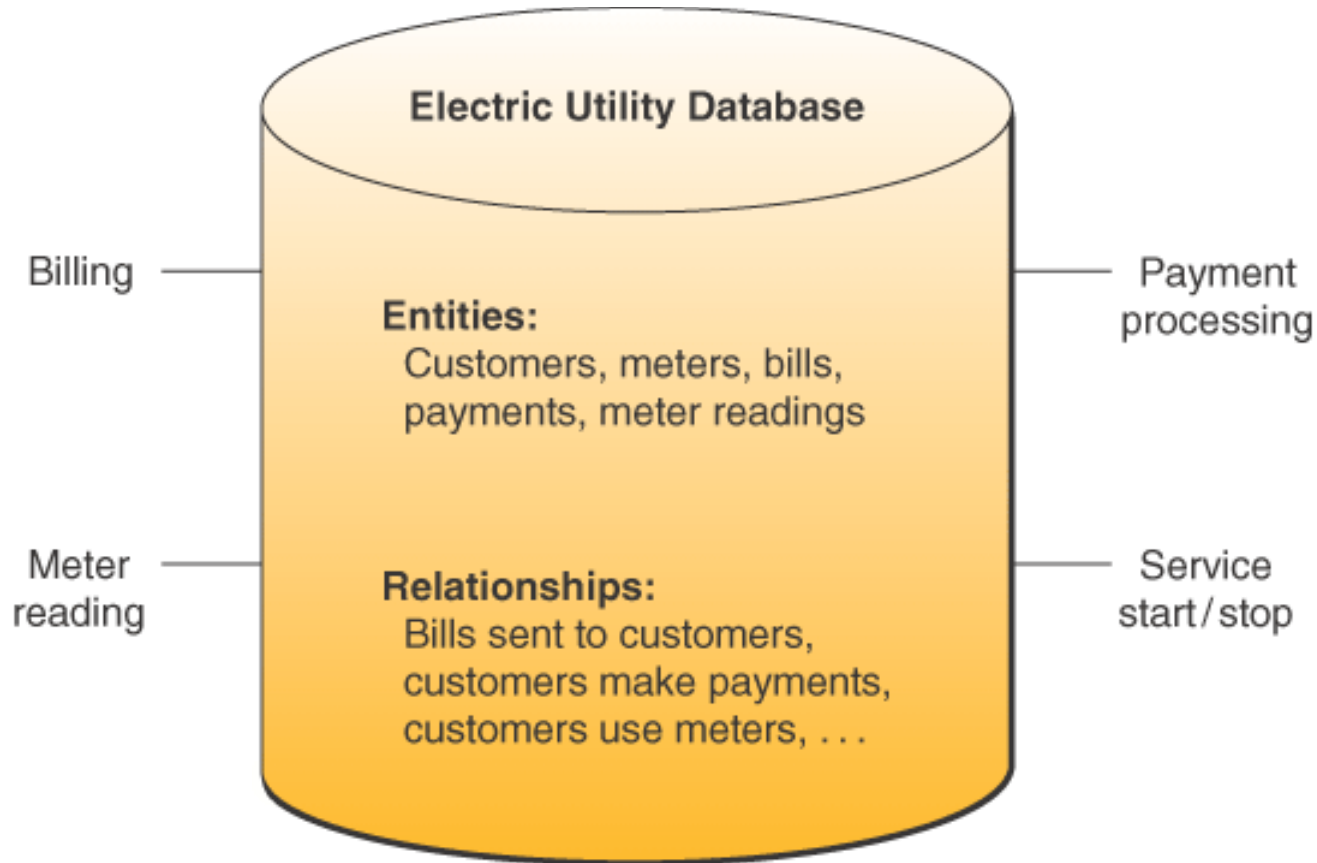
Human Resource Database



Human Resource Database



Electric Utility Database



Relational DBMS

- Relational DBMS
 - Represent data as two-dimensional tables called relations or files
 - Each table contains data on entity and attributes
- Table: grid of columns and rows
 - Rows : Records for different entities
 - Fields (columns): Represents attribute for entity
 - Key field: Field used to uniquely identify each record
 - Primary key: Field in table used for key fields
 - Foreign key: Primary key used in second table as look-up field to identify records from original table

RELATIONAL DATABASE

SUPPLIER

Columns (Attributes, Fields)

Supplier_Number	Supplier_Name	Supplier_Street	Supplier_City	Supplier_State	Supplier_Zip
8259	CBM Inc.	74 5 th Avenue	Dayton	OH	45220
8261	B. R. Molds	1277 Gandolly Street	Cleveland	OH	49345
8263	Jackson Composites	8233 Micklin Street	Lexington	KY	56723
8444	Bryant Corporation	4315 Mill Drive	Rochester	NY	11344

Rows
(Records,
Tuples)

Key Field
(Primary Key)

Relational DBMS

SUPPLIER

Columns (Attributes, Fields)

Supplier_Number	Supplier_Name	Supplier_Street	Supplier_City	Supplier_State	Supplier_Zip
8259	CBM Inc.	74 5 th Avenue	Dayton	OH	45220
8261	B. R. Molds	1277 Gandolly Street	Cleveland	OH	49345
8263	Jackson Composites	8233 Micklin Street	Lexington	KY	56723
8444	Bryant Corporation	4315 Mill Drive	Rochester	NY	11344

Rows
(Records,
Tuples)

Key Field
(Primary Key)

PART

Part_Number	Part_Name	Unit_Price	Supplier_Number
137	Door latch	22.00	8259
145	Side mirror	12.00	8444
150	Door molding	6.00	8263
152	Door lock	31.00	8259
155	Compressor	54.00	8261
178	Door handle	10.00	8259

Primary Key

Foreign Key

BASIC OPERATION OF RDBMS

PART

Part_Number	Part_Name	Unit_Price	Supplier_Number
137	Door latch	22.00	8259
145	Side mirror	12.00	8444
150	Door molding	6.00	8263
152	Door lock	31.00	8259
155	Compressor	54.00	8261
178	Door handle	10.00	8259

Select Part_Number = 137 or 150

SUPPLIER

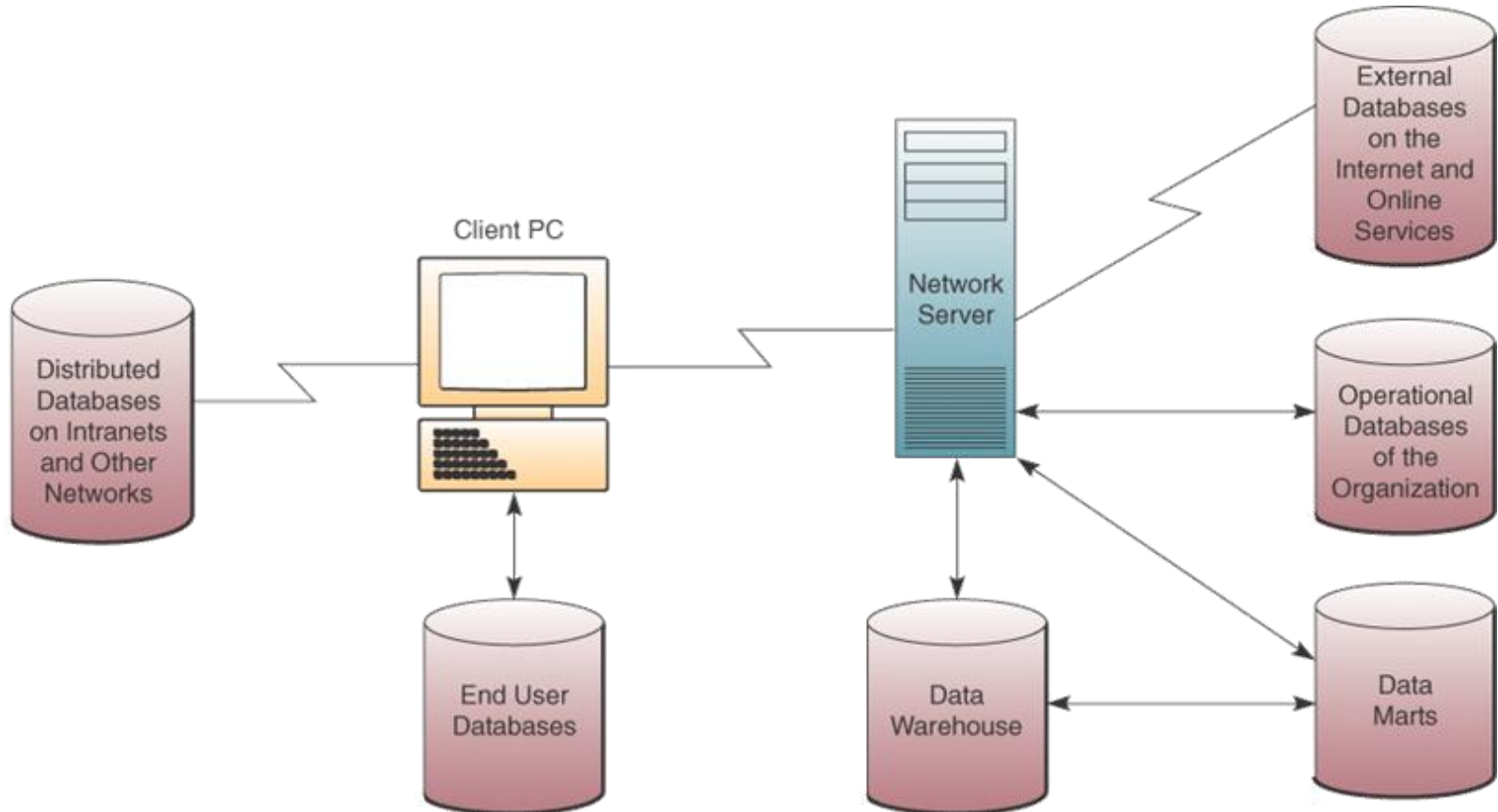
Supplier_Number	Supplier_Name	Supplier_Street	Supplier_City	Supplier_State	Supplier_Zip
8259	CBM Inc.	74 5 th Avenue	Dayton	OH	45220
8261	B. R. Molds	1277 Gandolly Street	Cleveland	OH	49345
8263	Jackson Components	8233 Micklin Street	Lexington	KY	56723
8444	Bryant Corporation	4315 Mill Drive	Rochester	NY	11344

Join by Supplier_Number

Part_Number	Part_Name	Supplier_Number	Supplier_Name
137	Door latch	8259	CBM Inc.
150	Door molding	8263	Jackson Components

Project selected columns

Types of Databases



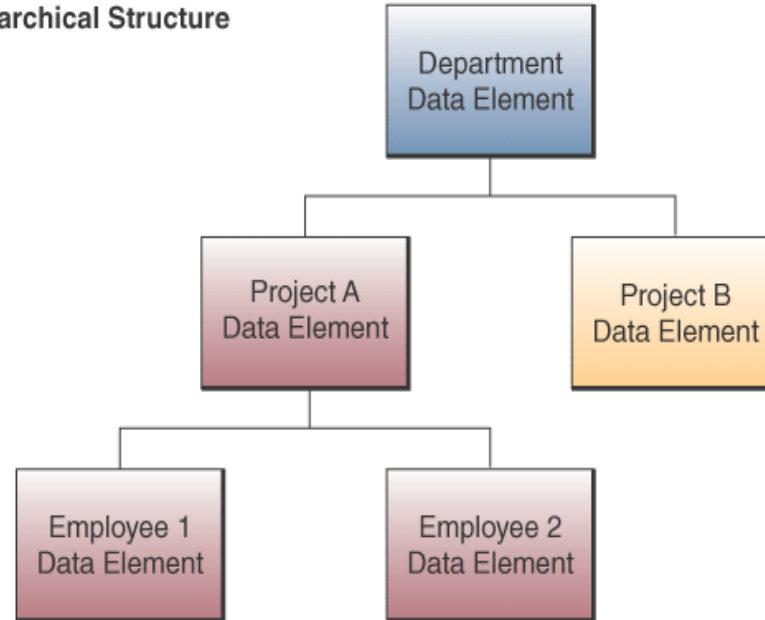
Database Structures

- Common database structures...
 - Hierarchical
 - Network
 - Relational
 - Object-oriented
 - Multi-dimensional

Hierarchical Structure

- Early DBMS structure
- Records arranged in tree-like structure
- Relationships are one-to-many

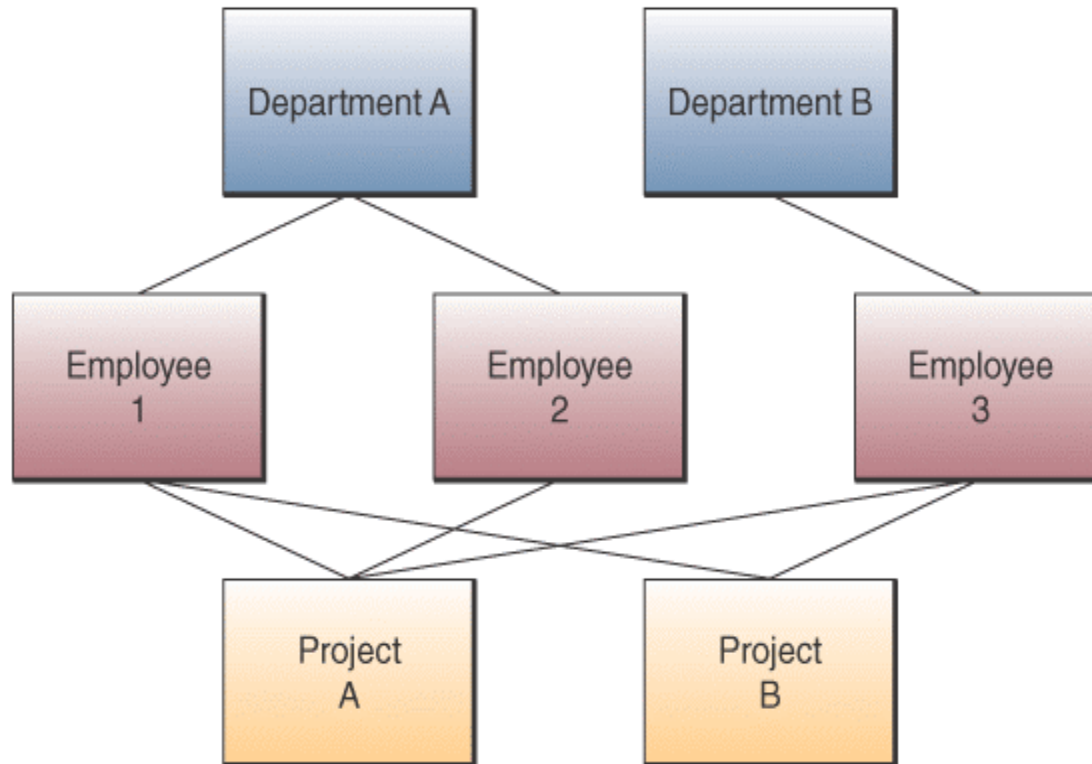
Hierarchical Structure



Network Structure

- Used in some mainframe DBMS packages
- Many-to-many relationships

Network Structure



Relational Structure

- Most widely used structure
 - Data elements are stored in tables
 - Row represents a record; column is a field
 - Can relate data in one file with data in another, if both files share a common data element

Department Table

Deptno	Dname	Dloc	Dmgr
Dept A			
Dept B			
Dept C			

Employee Table

Empno	Ename	Etitle	Esalary	Deptno
Emp 1				Dept A
Emp 2				Dept A
Emp 3				Dept B
Emp 4				Dept B
Emp 5				Dept C
Emp 6				Dept B

Relational Operations

- Select
 - Create a subset of records that meet a stated criterion
 - Example: employees earning more than \$30,000
- Join
 - Combine two or more tables temporarily
 - Looks like one big table
- Project
 - Create a subset of columns in a table

Multidimensional Structure

- Variation of relational model
 - Uses multidimensional structures to organize data
 - Data elements are viewed as being in cubes
 - Popular for analytical databases that support Online Analytical Processing (OLAP)

Multidimensional Model

		Denver					
		Los Angeles					
		San Francisco					
		West					
East		February		March			
		Actual	Budget	Actual	Budget		
Sales	Camera						
	TV						
	VCR						
	Audio						
Margin	Camera						
	TV						
	VCR						
	Audio						

		Profit					
		Total Expenses					
		Margin					
		COGS					
Sales		East		West			
		Actual	Budget	Actual	Budget		
TV	January						
	February						
	March						
	Qtr 1						
VCR	January						
	February						
	March						
	Qtr 1						

		April					
		Qtr 1					
		March					
		February					
January		Actual		Budget			
		Sales	Margin	Sales	Margin		
TV	East						
	West						
	South						
	Total						
VCR	East						
	West						
	South						
	Total						

		April					
		Qtr 1					
		March					
		February					
January		Sales		Margin			
		TV	VCR	TV	VCR		
East	Actual						
	Budget						
	Forecast						
	Variance						
West	Actual						
	Budget						
	Forecast						
	Variance						

Object-Oriented Structure

An **object** consists of

Data values describing the attributes of an entity

Operations that can be performed on the data

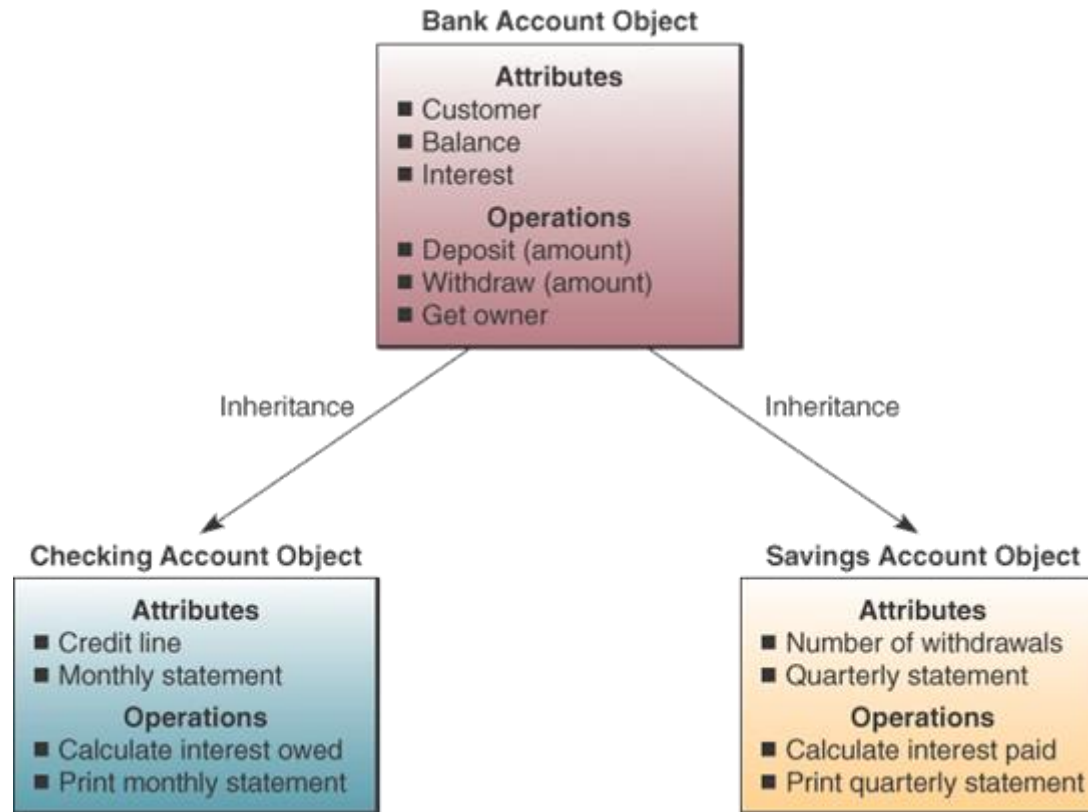
Encapsulation

Combine data and operations

Inheritance

New objects can be created by replicating some or all of the characteristics of parent objects

Object-Oriented Structure



Source: Adapted from Ivar Jacobsen, Maria Ericsson, and Ageneta Jacobsen, *The Object Advantage: Business Process Reengineering with Object Technology* (New York: ACM Press, 1995), p. 65.
Copyright © 1995, Association for Computing Machinery. By permission.

Object-Oriented Structure

Used in object-oriented database management systems (OODBMS)

Supports complex data types more efficiently than relational databases

Examples: graphic images, video clips,
web pages

Evaluation of Database Structures

Hierarchical

- Works for structured, routine transactions

- Can't handle many-to-many relationship

Network

- More flexible than hierarchical

- Unable to handle ad hoc requests

Relational

- Easily responds to ad hoc requests

- Easier to work with and maintain

- Not as efficient/quick as hierarchical or network

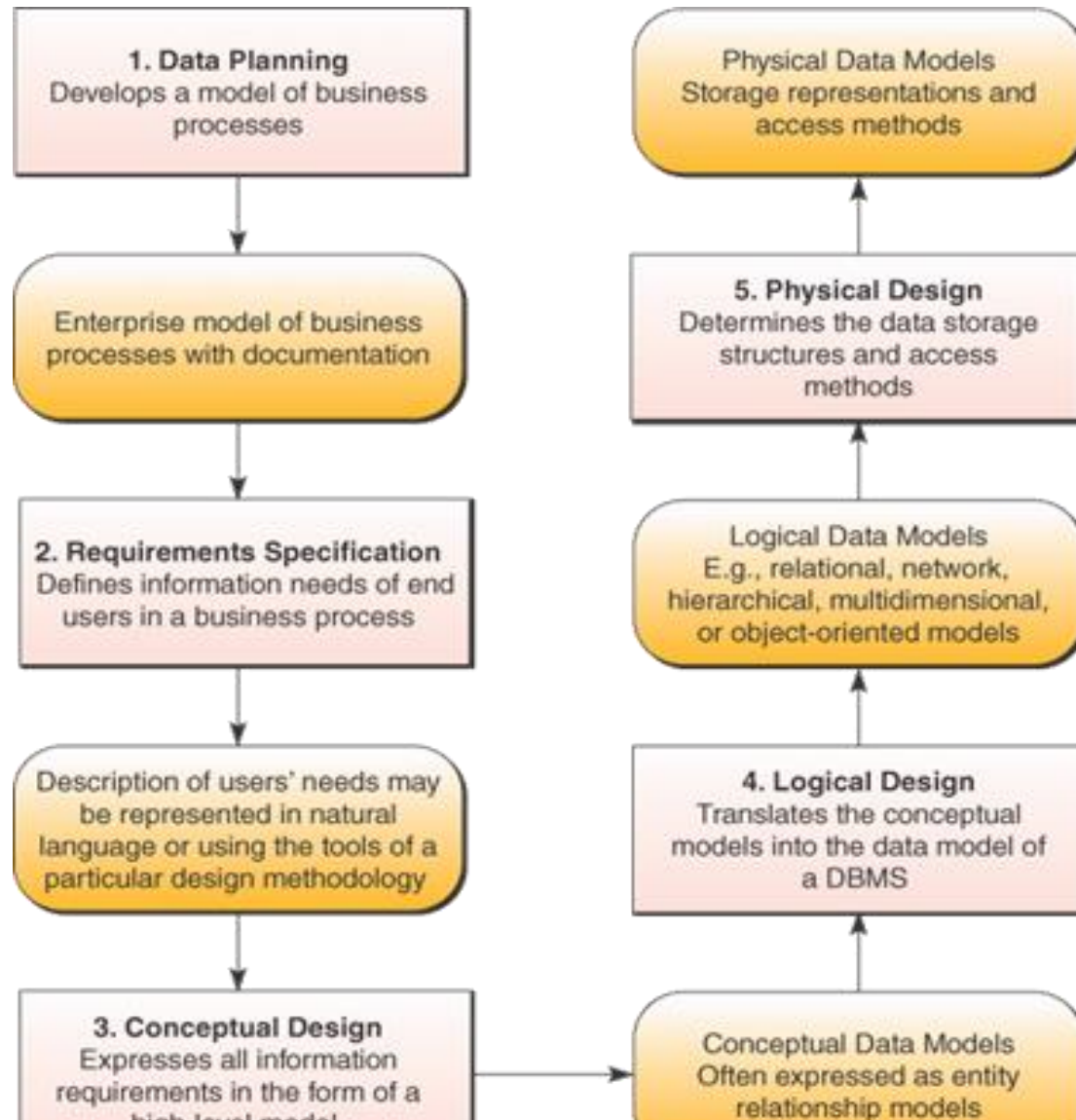
Database Development

- Database Administrator (DBA)
 - In charge of enterprise database development
 - Improves the integrity and security of organizational databases
 - Uses Data Definition Language (DDL) to develop and specify data contents, relationships, and structure
 - Stores these specifications in a data dictionary or a metadata repository

Data Dictionary

- A data dictionary
 - Contains data about data (metadata)
 - Relies on specialized software component to manage a database of data definitions
- It contains information on..
 - The names and descriptions of all types of data records and their interrelationships
 - Requirements for end users' access and use of application programs
 - Database maintenance
 - Security

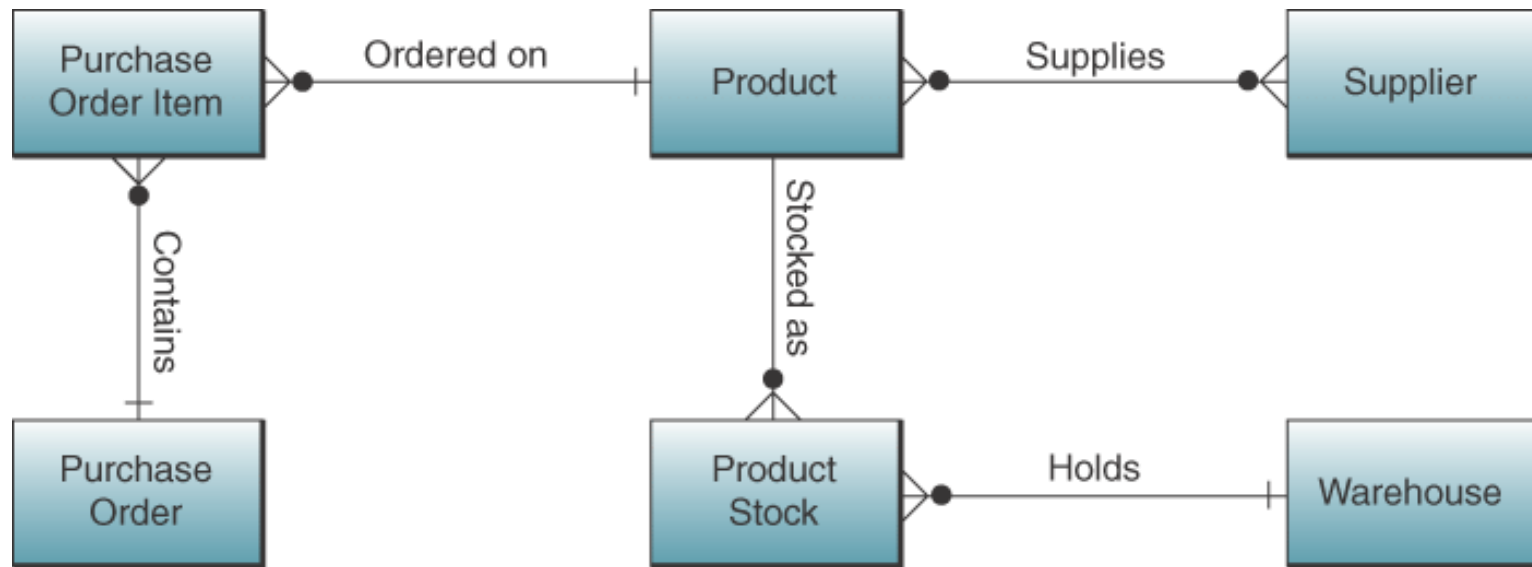
Database Development



Data Planning Process

- Database development is a top-down process
 - Develop an enterprise model that defines the basic business process of the enterprise
 - Define the information needs of end users in a business process
 - Identify the key data elements that are needed to perform specific business activities (entity relationship diagrams)

Entity Relationship Diagram



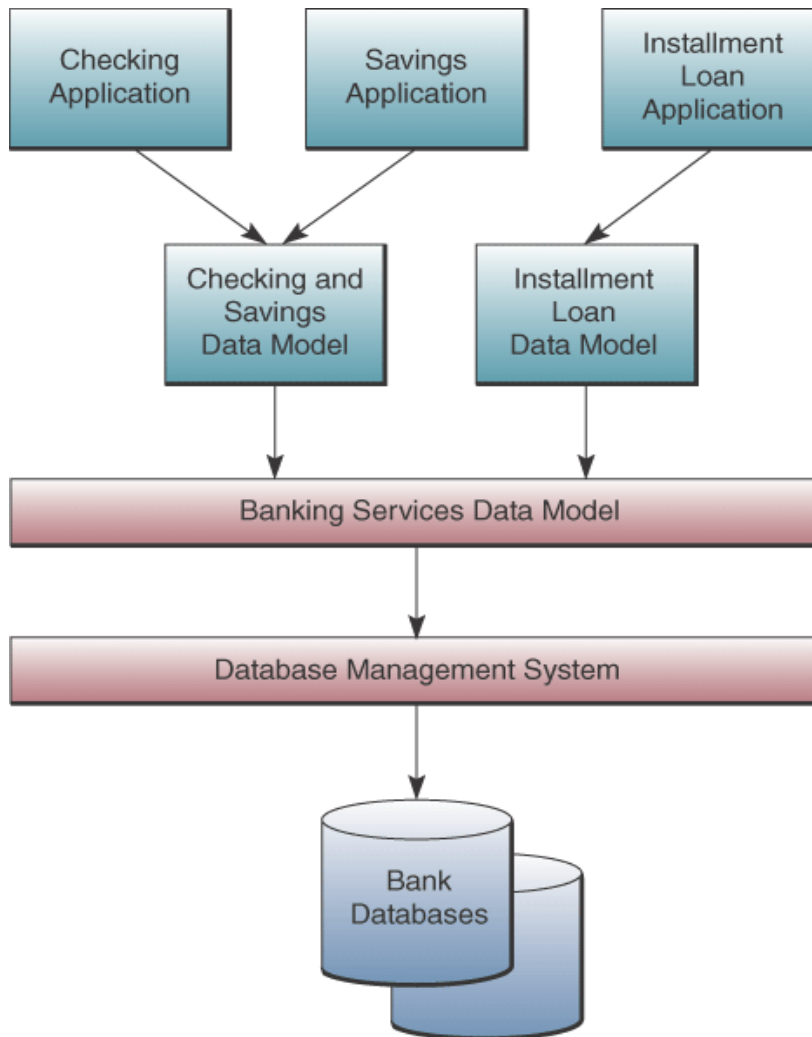
Database Design Process

- Data relationships are represented in a data model that supports a business process
- This model is the **schema** or **subschema** on which to base...
 - The physical design of the database
 - The development of application programs to support business processes

Database Design Process

- Logical Design
 - Schema - overall logical view of relationships
 - Subschema - logical view for specific end users
 - Data models for DBMS
- Physical Design
 - How data are to be physically stored and accessed on storage devices

Logical and Physical Database Views



Logical User Views

Data elements and relationships (the subschemas) needed for checking, savings, or installment loan processing

Data elements and relationships (the schema) needed for the support of all bank services

Software Interface

The DBMS provides access to the bank's databases

Physical Data Views

Organization and location of data on the storage media

Data Resource Management

- Data resource management is a managerial activity
 - Uses data management, data warehousing, and other IS technologies
 - Manages data resources to meet the information needs of business stakeholders

Operational Databases

- Stores detailed data needed to support business processes and operations
 - Also called **subject area databases (SADB)**, **transaction databases**, and **production databases**
 - Database examples: customer, human resource, inventory

Distributed Databases

- Distributed databases are copies or parts of databases stored on servers at multiple locations
 - Improves database performance at worksites
- Advantages
 - Protection of valuable data
 - Data can be distributed into smaller databases
 - Each location has control of its local data
 - All locations can access any data, any where
- Disadvantages
 - Maintaining data accuracy

Distributed Databases

- Replication
 - Look at each distributed database and find changes
 - Apply changes to each distributed database
 - Very complex
- Duplication
 - One database is master
 - Duplicate the master after hours, in all locations
 - Easier to accomplish

External Databases

- Databases available for a fee from commercial online services, or free from the Web
 - Examples: hypermedia databases, statistical databases, bibliographic and full text databases
 - Search engines like Google or Yahoo are external databases

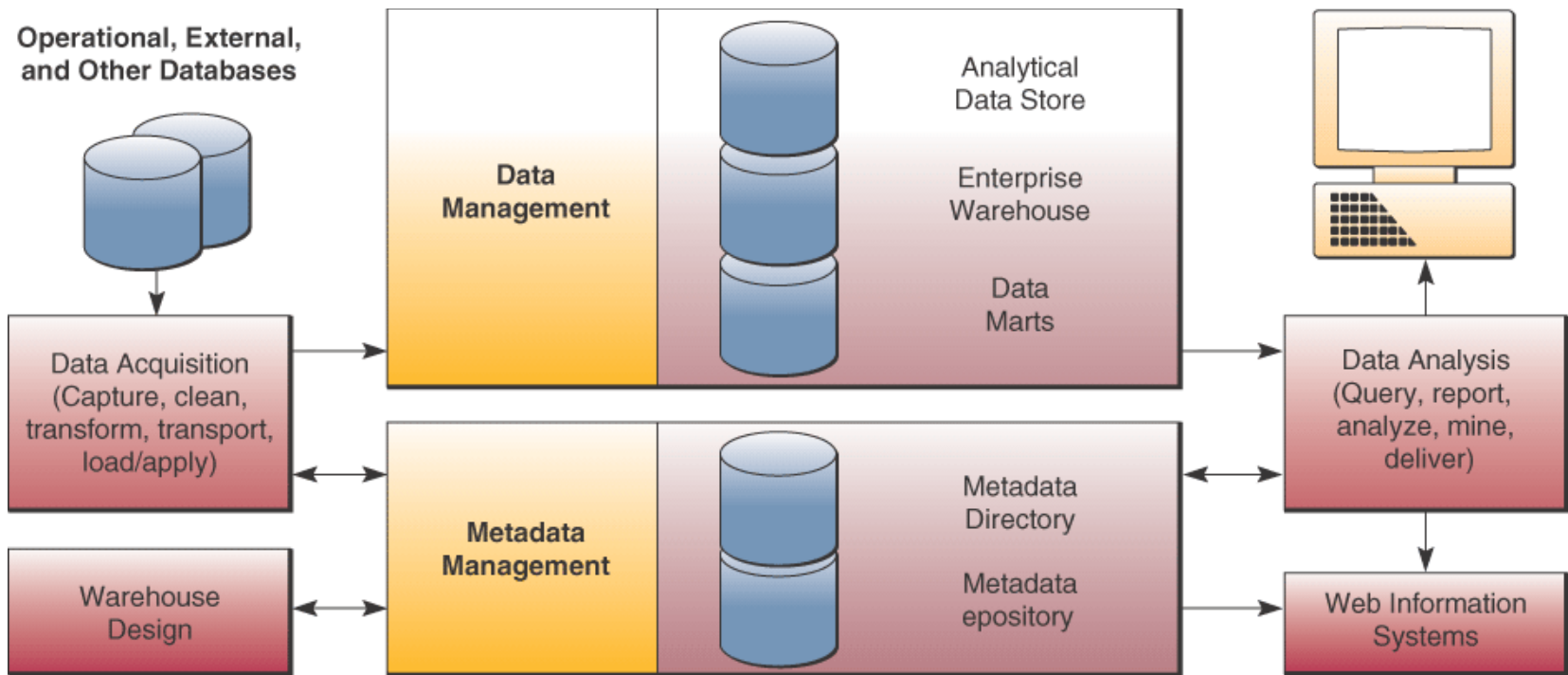
Hypermedia Databases

- A hypermedia database contains
 - Hyperlinked pages of multimedia
 - Interrelated hypermedia page elements, rather than interrelated data records

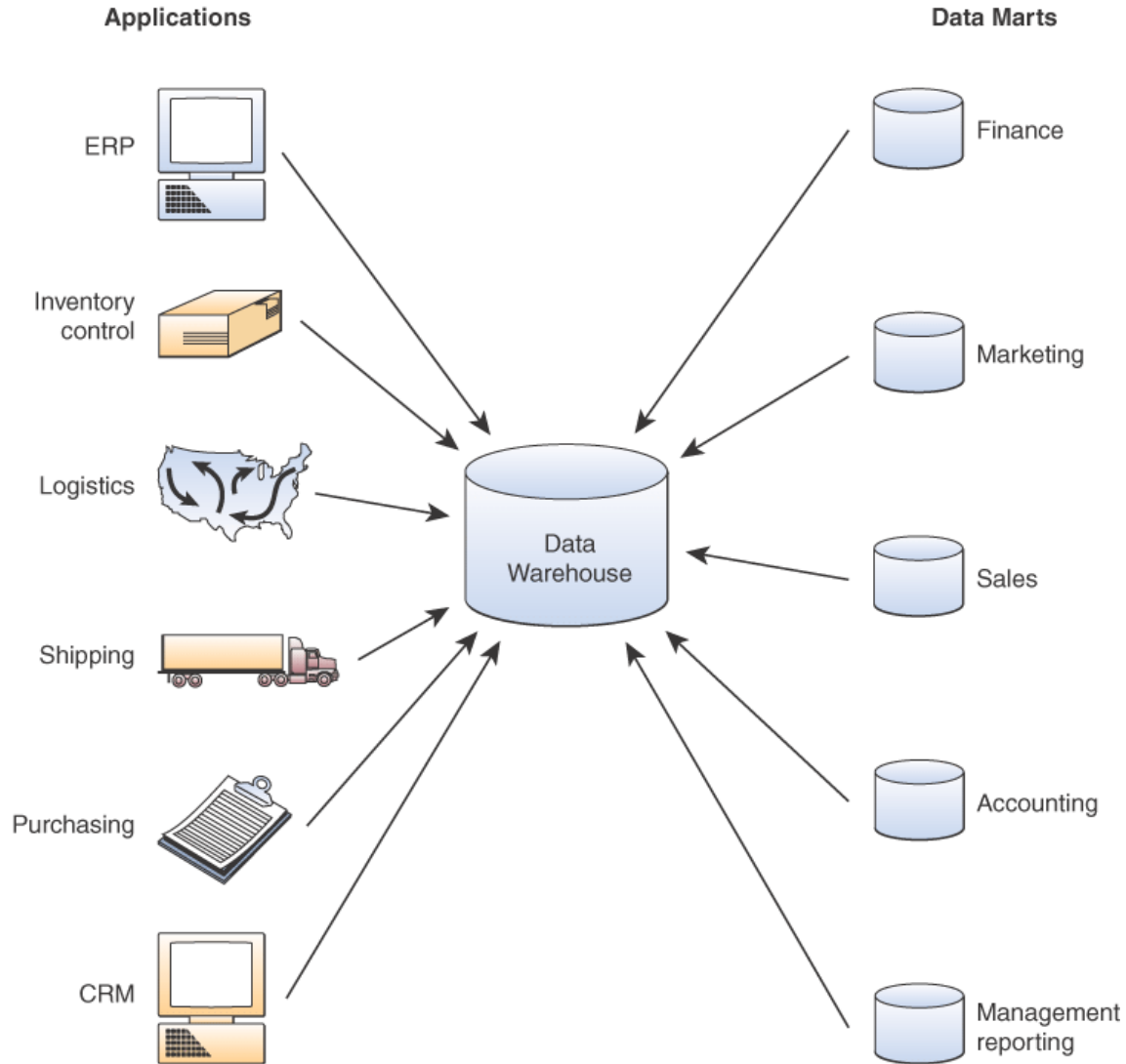
Data Warehouses

- Stores static data that has been extracted from other databases in an organization
 - Central source of data that has been cleaned, transformed, and cataloged
 - Data is used for data mining, analytical processing, analysis, research, decision support
- Data warehouses may be divided into data marts
 - Subsets of data that focus on specific aspects of a company (department or business process)

Data Warehouse Components



Applications and Data Marts



Data Mining

- Data in data warehouses are analyzed to reveal hidden patterns and trends
 - Market-basket analysis to identify new product bundles
 - Find root cause of quality or manufacturing problems
 - Prevent customer attrition
 - Acquire new customers
 - Cross-sell to existing customers
 - Profile customers with more accuracy