

CO558 Databases an Introduction

- Introductions,
- Module plan & reading list & assessment
- Lecture 1 – Introduction to Relational Databases

Introductions

- Your Name?
- What course are you on?
- A fascinating fact about you?
- What job do you want to do?



Module content

Fundamental theories of relational database design and management.

- Relational data analysis and entity relationship modelling,
- Techniques of data normalisation to 3NF
- The use of SQL DML (Structured Query Language : Data Manipulation Language) to query a relational database,
- Introduction to database transaction processing and concurrency control.

Learning outcomes

On successful completion of the module:

1. Construct a relational data model (entity relationship diagram)
2. Apply the techniques of data normalisation to 3NF
3. Use SQL DML to query a relational database
4. Explain the role of concurrency control and database recovery in a large-scale database environment

The Database Environment Lecture1: 26/09/2016

Modern Database Management

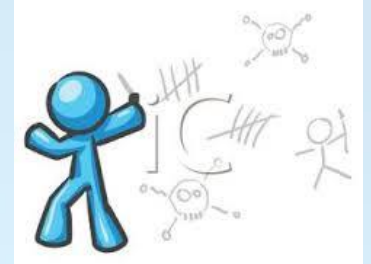
11th Edition

Jeffrey A. Hoffer, V Ramesh and Heikki Topi

Lesson Learning Outcomes

- Concisely define key terms regarding a Relational Database.
- Name several limitations of conventional file processing systems.
- Explain at least 10 advantages of the database approach.
- Identify several costs and risks of the database approach.

- Can you tell me about the history of databases?



- Get yourself into 2's and find out one significant fact about the history of databases. When you have found one fact write it on the board.

Evolution of DB Systems

- Flat files - 1960s - 1980s
- Hierarchical – 1970s - 1990s
- Network – 1970s - 1990s
- Relational – 1980s - present
- Object-oriented – 1990s - present
- Object-relational – 1990s - present
- Data warehousing – 1980s - present
- Web-enabled – 1990s - present

Types of Databases

Relational	Amazon Aurora, Oracle, Microsoft SQL Server, PostgreSQL, MySQL and MariaDB
Key Value	DynamoDB, Raik, Redis, Aerospike and Azure Table Storage
Column-Orientated	HBase, Cassandra, MonetDB, IBM Informix, Apache Flink and Google Clodata
Document-orientated	MongoDB, CouchDB, Azure Document DB and JSON ODM
Graph	Neo4J, ArangoDB, Trinity, AllegroGraph and Bigdata

Break time – 20 minutes



Important Terms!

Can you tell me what we mean by:

- **Data?**
- **Information?**
- **Database?**
- **Entity**
- **Metadata?**

Figure 1-1a Data in Context (Information)

Class Roster

Course: MGT 500 Semester: Spring 200X
Business Policy **Large volume of facts, difficult to interpret or make decisions based on**

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

Table 1-1 Metadata

Descriptions of the properties or characteristics of the data, including data types, field sizes, allowable values, and documentation

Table 1-1 Example Metadata for Class Roster

<i>Data Item</i>			<i>Value</i>		
Name	Type	Length	Min	Max	Description
Course	Alphanumeric	30			Course ID and name
Section	Integer	1	1	9	Section number
Semester	Alphanumeric	10			Semester and year
Name	Alphanumeric	30			Student name
ID	Integer	9			Student ID (SSN)
Major	Alphanumeric	4			Student major
GPA	Decimal	3	0.0	4.0	Student grade point average

Entities

An entity is a person, place, object, event or concept. It is things/objects in our database system which the organisation wishes to maintain. E.g.

- Person: employee, student, patient
- Place: City, State, Country
- Object: Machine, Building, Automobile
- Event: Sales, Registration, Renewal
- Concept: Account, Course, Department

Example - Beginning Design Process

Pine Valley Furniture Company

- Manufactures high-quality furniture
- Distributes furniture to retailers
- Pine Valley has several product lines of furniture
- Customers submit orders via Telephone, Mail, Email and Fax.
- The company employs 100 staff

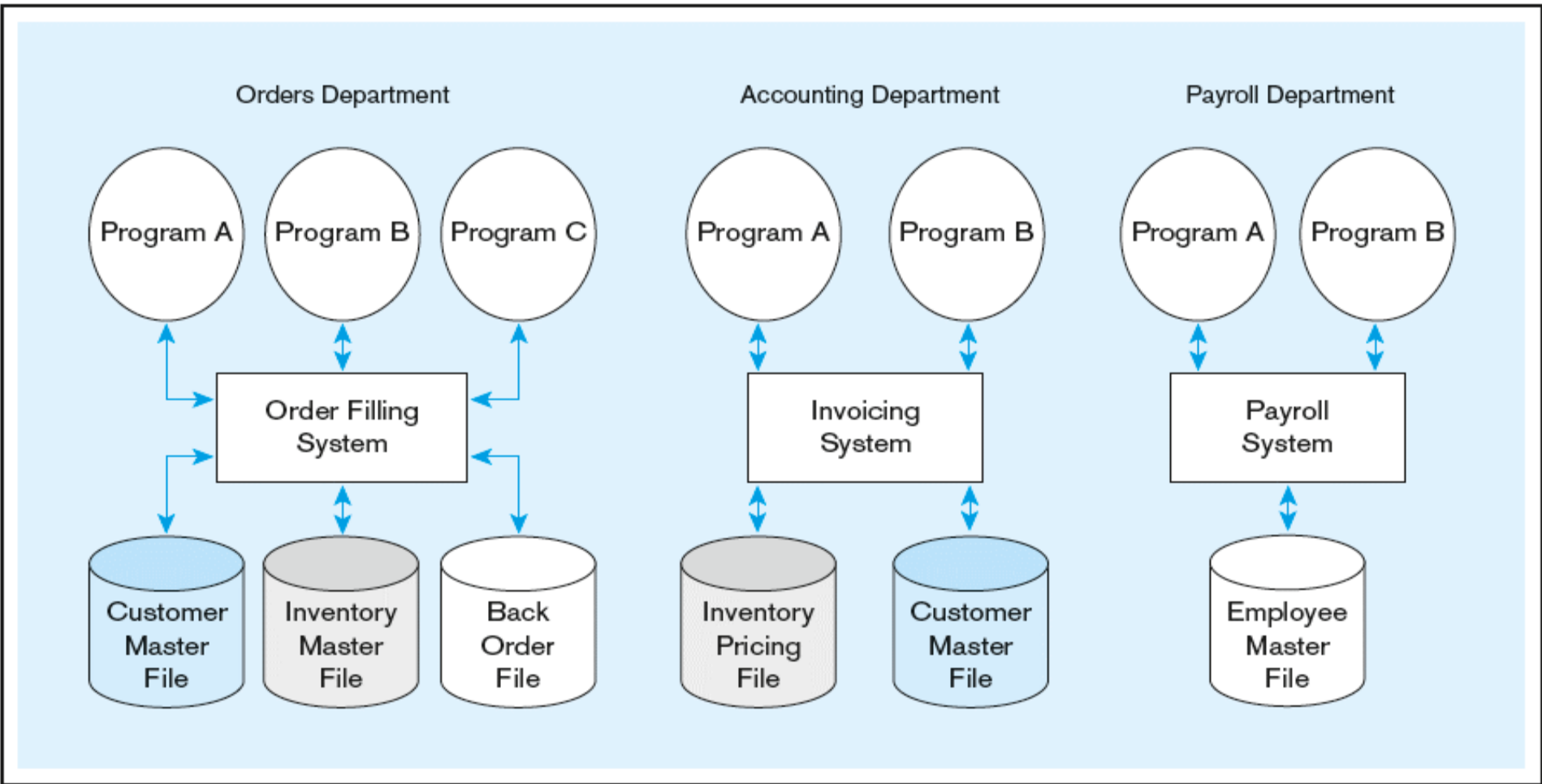
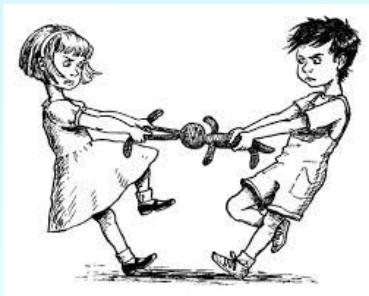


FIGURE 1-2 Old file processing systems at Pine Valley Furniture Company

Limitations of file processing for Pine Valley

Data Dependency

Any change is duplicated through structures and descriptions



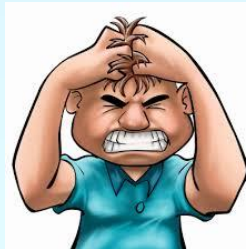
Data Sharing

No centralized control of data



Duplicate Data

Different systems/programs have separate copies of the same data



Changes difficult to implement

Simple changes can have unknown consequences

Why relational?

- **Data consistency and conforming to ACID properties**
- **Enforcement of standard**
- **Based on mathematical principles on Set theory – data is structured**
- **Data accessibility straightforward SQL queries**
- **Data Constraints – rules that cannot be violated**

Enterprise Data Model

Pine Valley Furniture Company

- Manufactures high-quality furniture
- Distributes furniture to retailers
- Pine Valley has several product lines of furniture
- Customers submit orders via Telephone, Mail, Email and Fax.
- The company employs 100 staff

Identifying the Entities

- A company has several departments.
- Each department has a supervisor and a number of employees.
- An Employees is assigned to one departments.
- One employee is assigned to a project. A project can have a number of employees working on it.

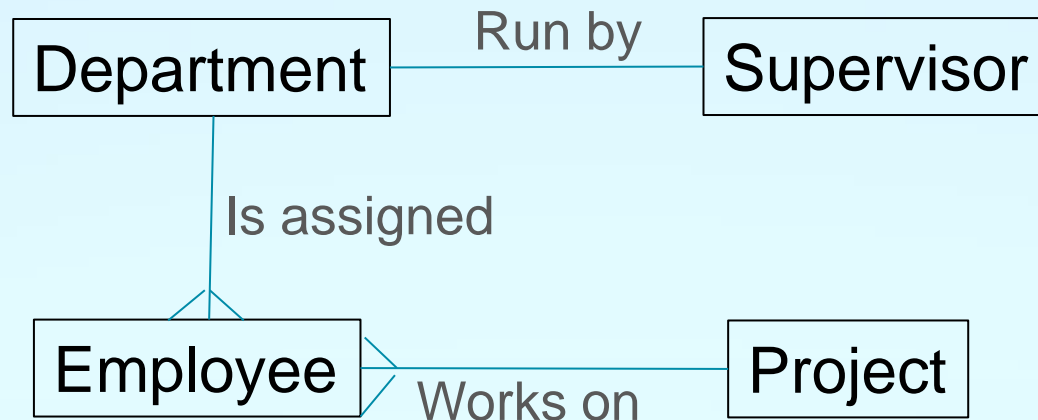


Figure 1-3 Segment from enterprise data model

Figure 3

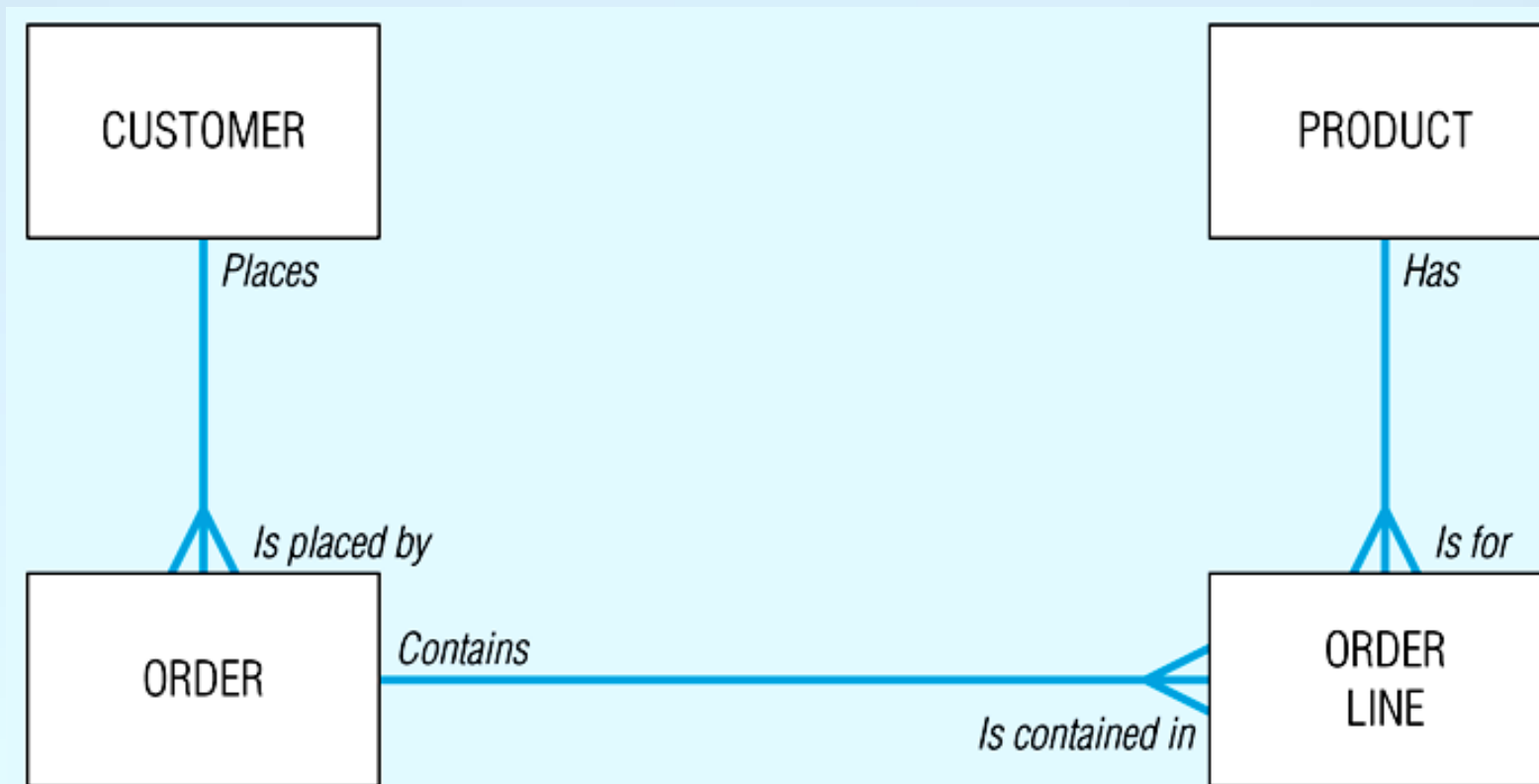


Figure 1-3 Segment from enterprise data model
Figure 3

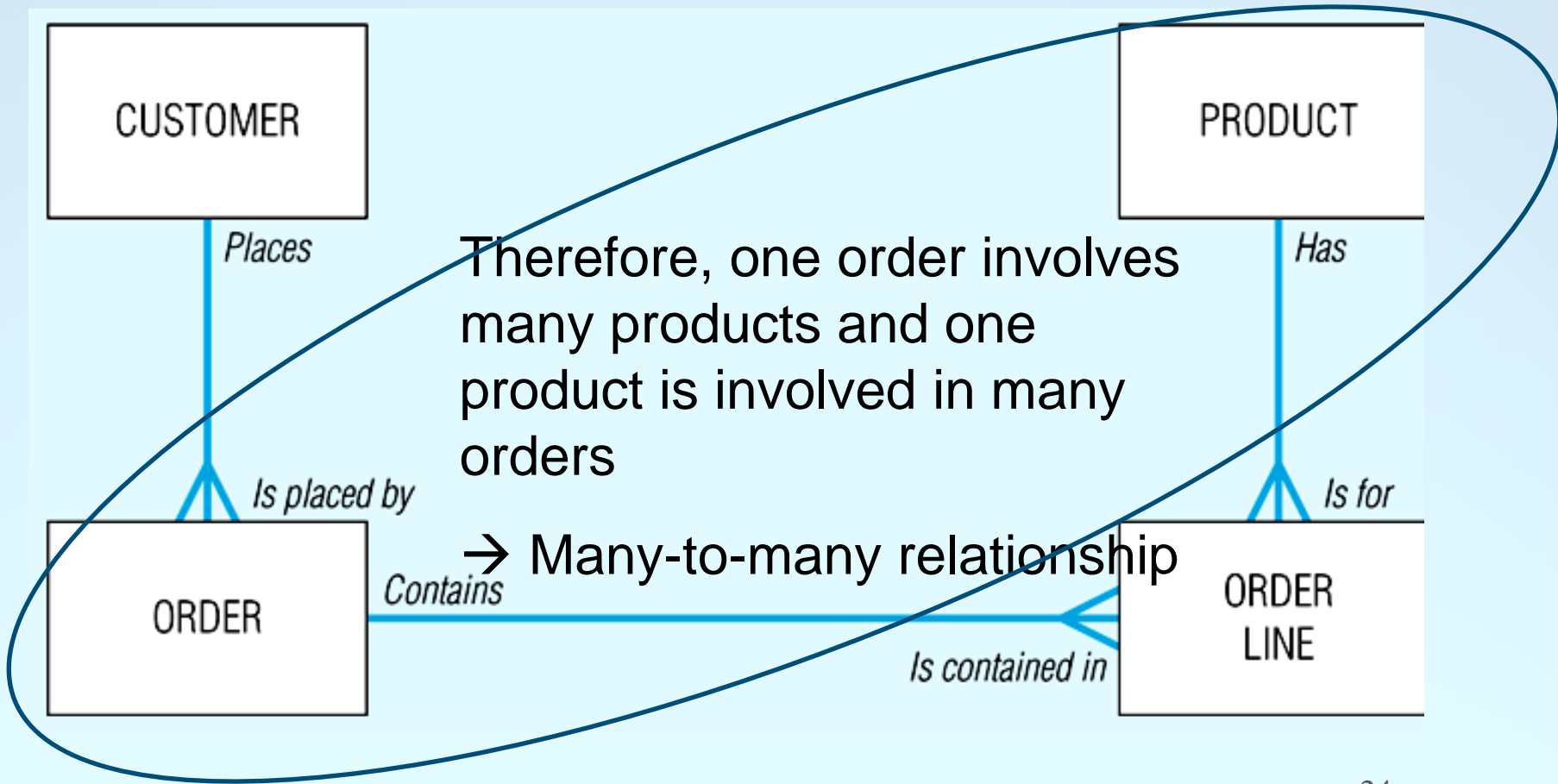


Figure 1-4 Order, Order_Line, Customer, and Product tables

Relationships established in special columns that provide links between tables

The screenshot displays four tables in Microsoft Access:

- ORDER_t : Table**

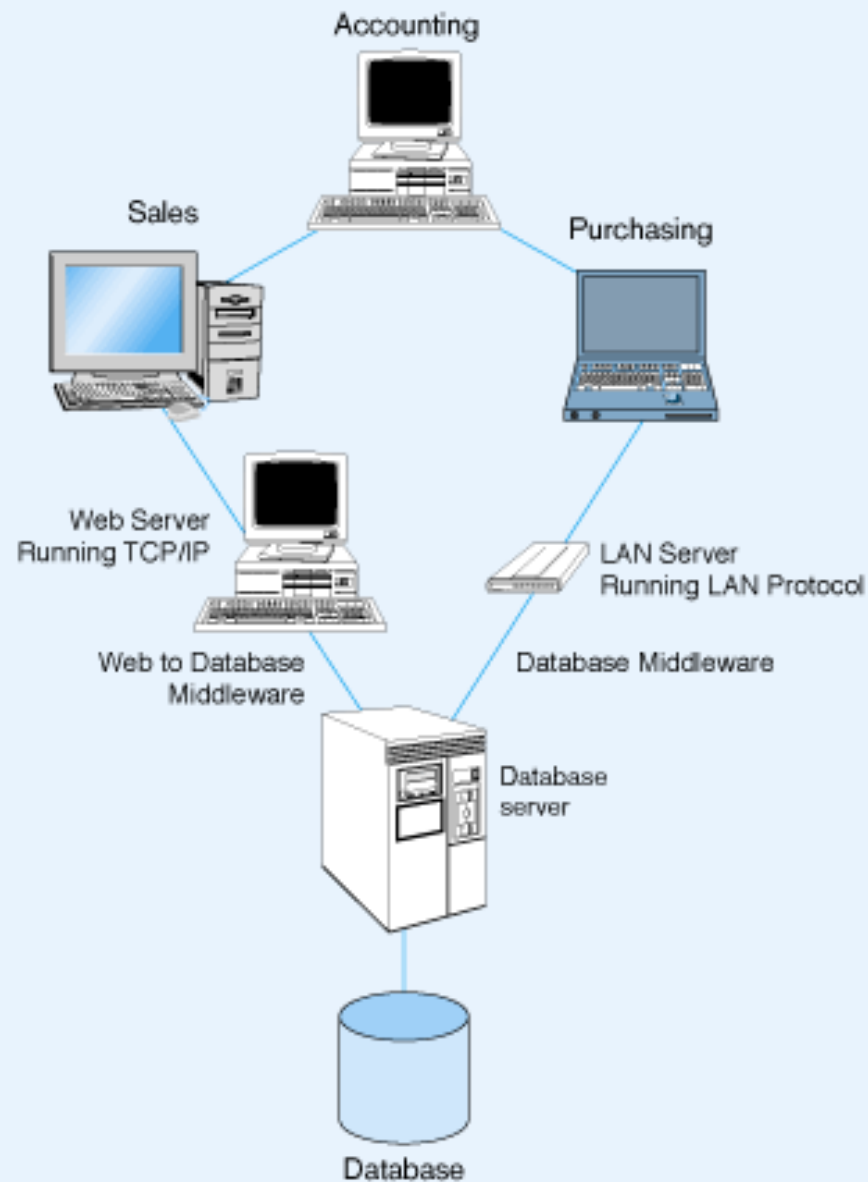
Order_id	Order_Date	Customer_ID
1	9/8/2000	4
2	10/4/2000	3
3	7/19/2000	1
4	11/11/2000	6
5	7/28/2000	4
6	8/17/2000	6
7	8/17/2000	3
(AutoNumber)		0
- Order line 1 : Table**

Order_ID	Product_ID	Quantity
1	2	2
1	10	2
2	3	1
2	8	2
2	14	2
3	6	2
4	3	1
4	7	3
5	6	2
5	10	1
0	0	0
- PRODUCT_t : Table**

Product ID	Product Description	Product Finish	Unit Price	On Hand
1	End Table	Cherry	\$175.00	8
2	Coffe Table	Birch	\$200.00	4
3	Computer Desk	Oak	\$375.00	5
4	Entertainment Center	Maple	\$650.00	3
5	Writer's Desk	Oak	\$325.00	0
6	8-Drawer Dresser	Birch	\$750.00	5
7	48" Bookcase	Walnut	\$150.00	5
8	48" Bookcase	Oak	\$175.00	2
9	96" Bookcase	Walnut	\$225.00	4
10	96" Bookcase	Oak	\$200.00	4
11	4-Drawer Dresser	Oak	\$500.00	3
12	8-Drawer Dresser	Oak	\$800.00	2
13	Nightstand	Cherry	\$150.00	5
14	Writer's Desk	Birch	\$300.00	2
- CUSTOMER_t : Table**

Customer ID	Customer Name	Address	City	State	Postal Code
1	Contemporary Casuals	1355 S Hines Blvd	Gainesville	FL	32601-
2	Value Furniture	15145 S.W. 17th St.	Plano	TX	75094-
3	Home Furnishings	1900 Allard Ave	Albany	NY	12209-
4	Eastern Furniture	1925 Beltline Rd.	Carteret	NJ	07008-
5	Impressions	5585 Westcott Ct.	Sacramento	CA	94206-
6	Furniture Gallery	325 Flatiron Dr.	Boulder	CO	80514-

Figure 1-5
Client/server
system for a
Furniture
Company



Break time – 10 minutes



Database Management System

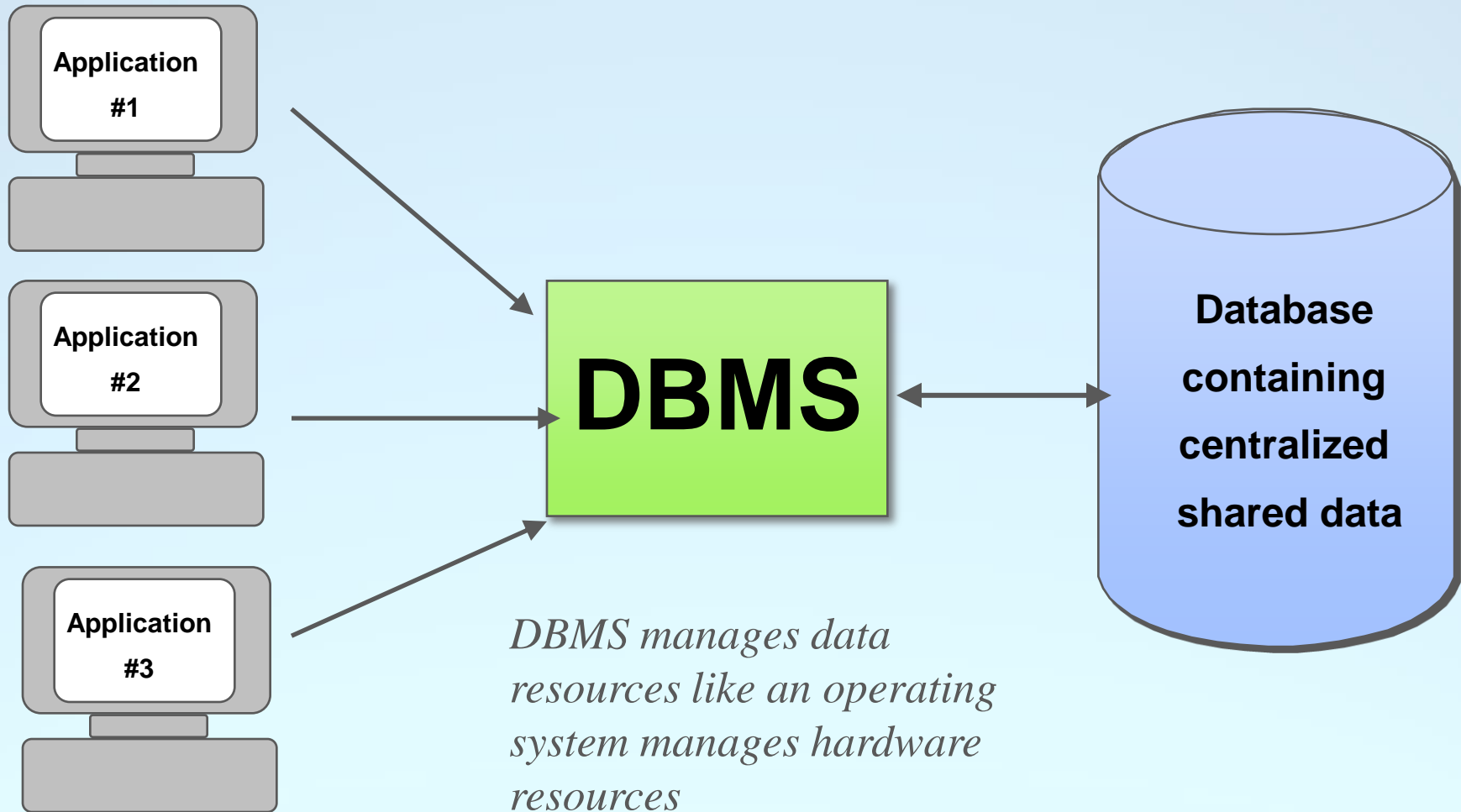
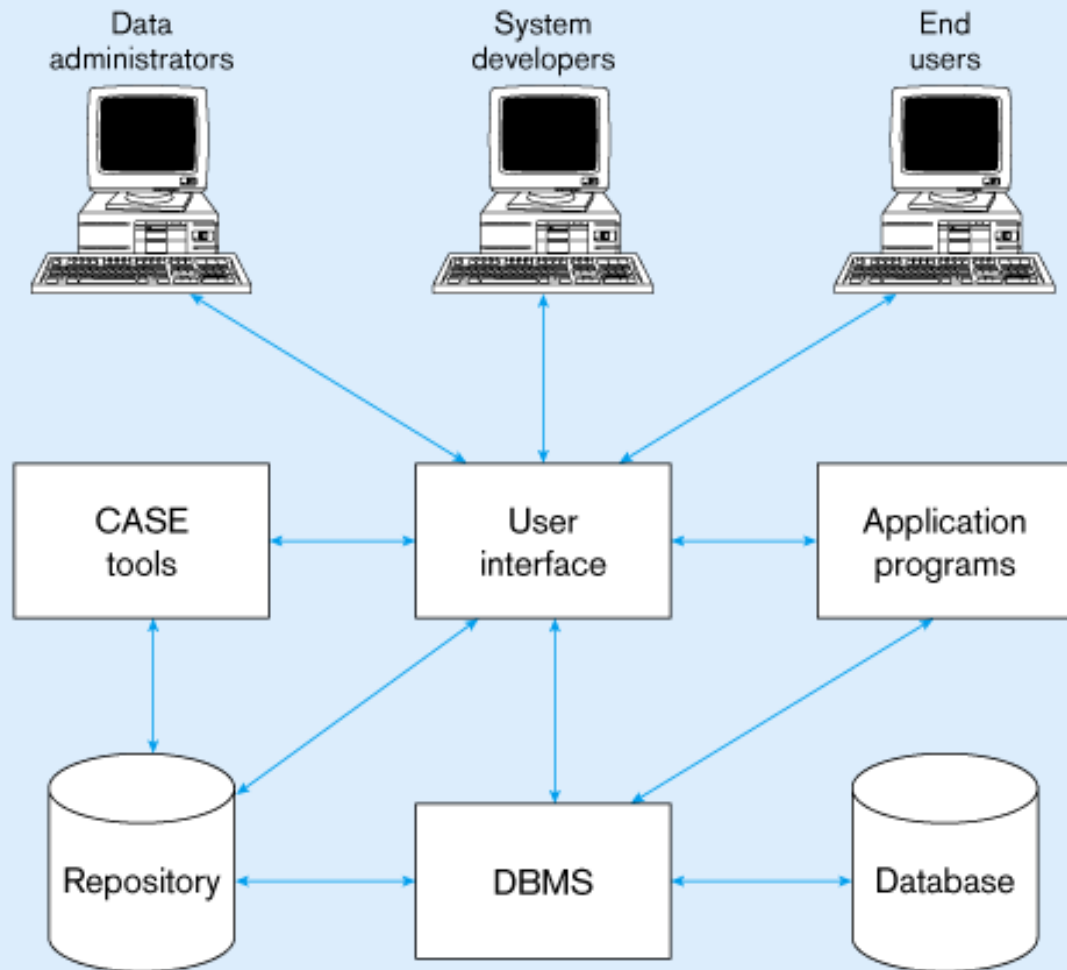


Figure 1-10
Components
of the
database
environment



Costs and risks of the database approach

- Up-front costs:
 - Installation Management Cost and Complexity
 - Conversion Costs
- Ongoing Costs
 - Requires New, Specialized Personnel
 - Need for Explicit Backup and Recovery
- Organizational Conflict
 - Old habits die hard

Modern Database Management
11th Edition

Jeffrey A. Hoffer, V Ramesh and Heikki Topi

wayne.palmer@bucks.ac.uk

kompel.campion@bucks.ac.uk