

Các khái niệm về Hệ Cơ Sở Dữ Liệu

- ❑ Cơ sở dữ liệu (CSDL)
- ❑ Các mô hình dữ liệu
- ❑ Hệ quản trị CSDL - DBMS

Các định nghĩa

- Dữ liệu (data): các sự kiện, văn bản, hình ảnh, âm thanh, phim ảnh có ý nghĩa được ghi lại và lưu trữ trên máy tính
- Cơ sở dữ liệu (database): một tập hợp có tổ chức các dữ liệu có liên quan luận lý với nhau
- Thông tin (information): dữ liệu đã được xử lý, hữu ích trong việc ra quyết định

Bảng dữ liệu thí dụ

Dữ liệu được đặt trong ngữ cảnh

Class Roster

Course: MGT 500
Business Policy

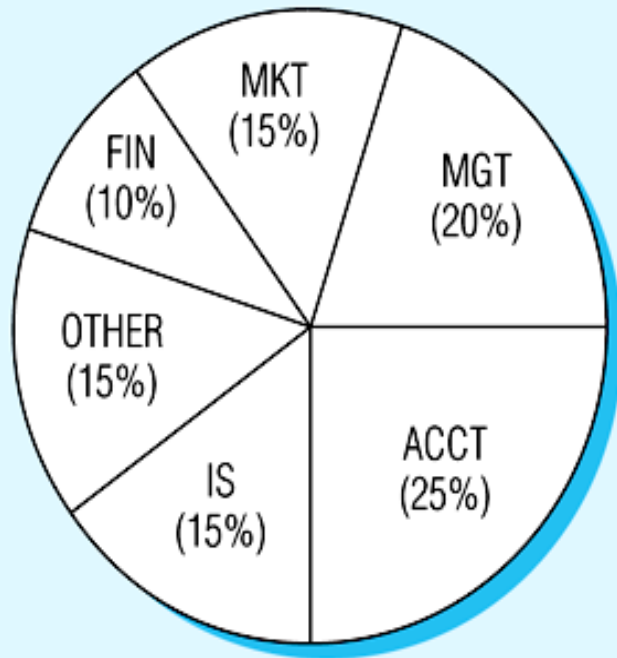
Semester: Spring 200X

Section: 2

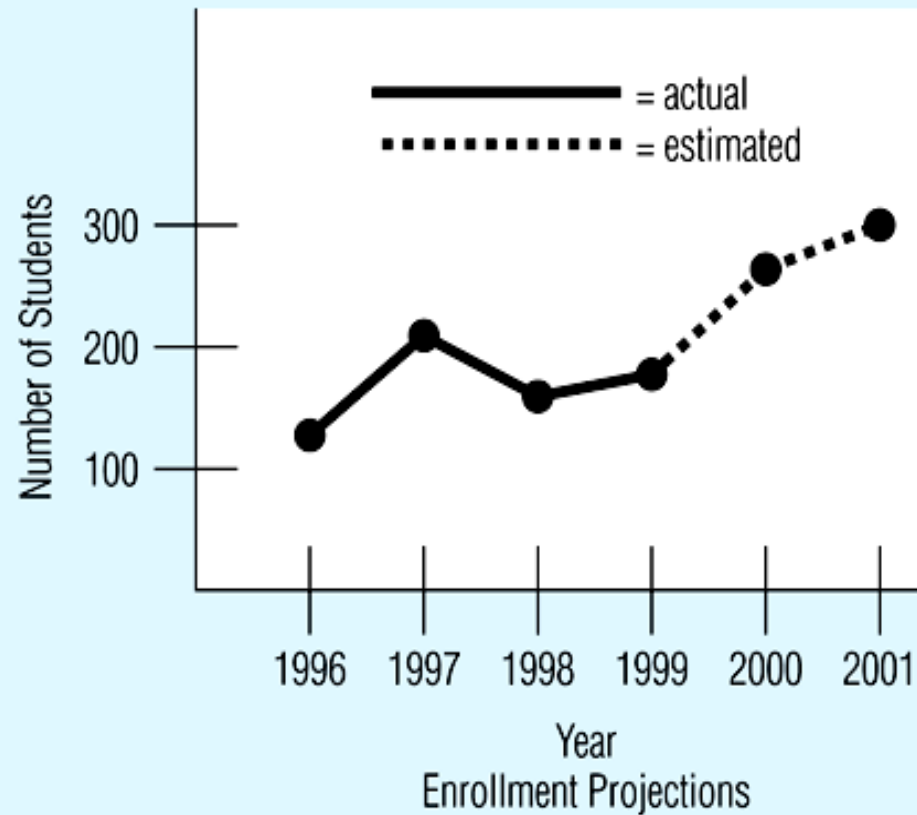
<u>Name</u>	<u>ID</u>	<u>Major</u>	<u>GPA</u>
Baker, Kenneth D.	324917628	MGT	2.9
Doyle, Joan E.	476193248	MKT	3.4
Finkle, Clive R.	548429344	PRM	2.8
Lewis, John C.	551742186	MGT	3.7
McFerran, Debra R.	409723145	IS	2.9
Sisneros, Michael	392416582	ACCT	3.3

Dữ liệu đã được thống kê

Các thông tin cần thiết để đọc, để hiểu



Percent Enrollment by Major (199X)

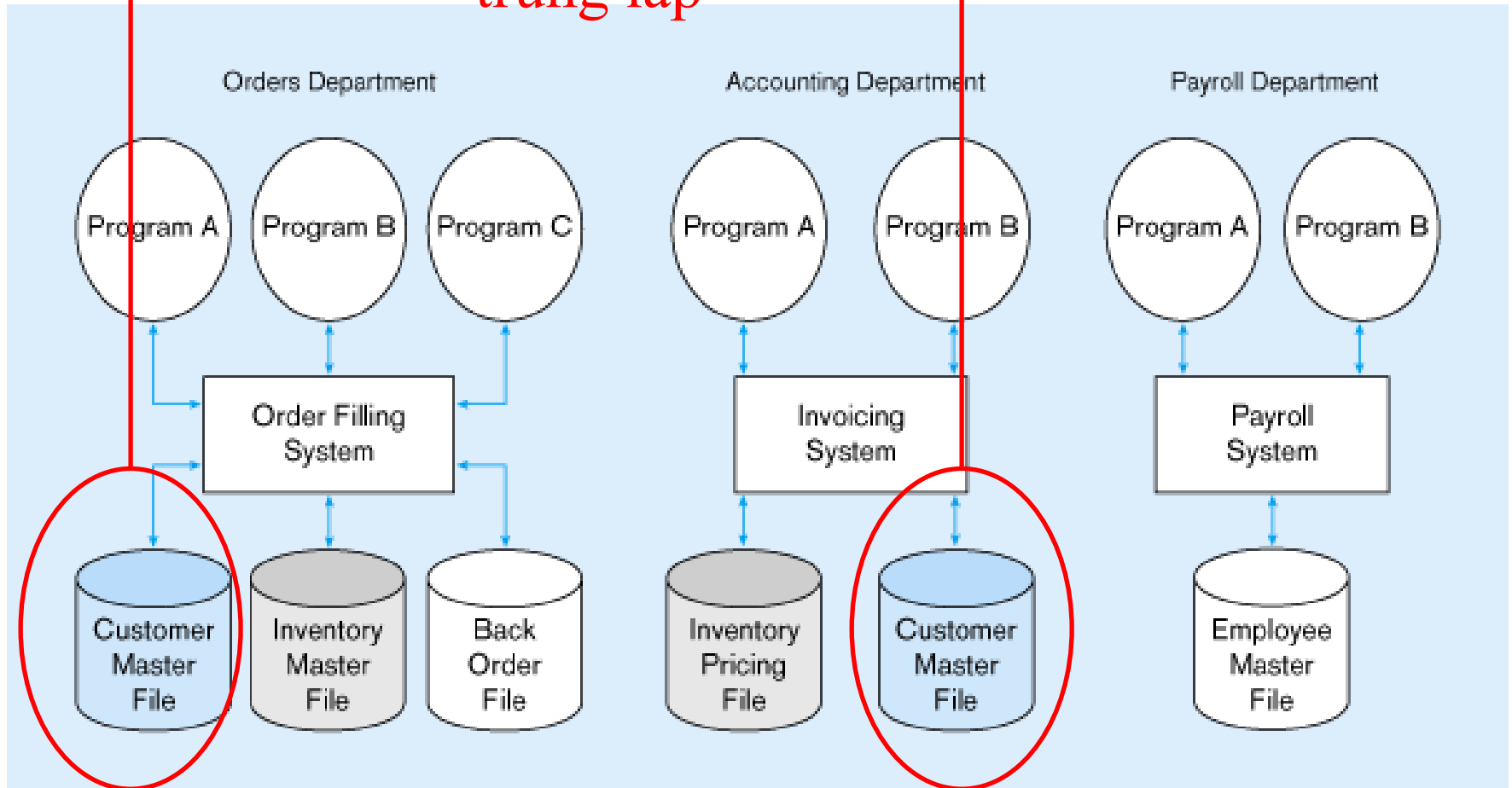


Vấn đề Quản lý Dữ liệu

- Xử lý tập tin
- Tiếp cận CSDL

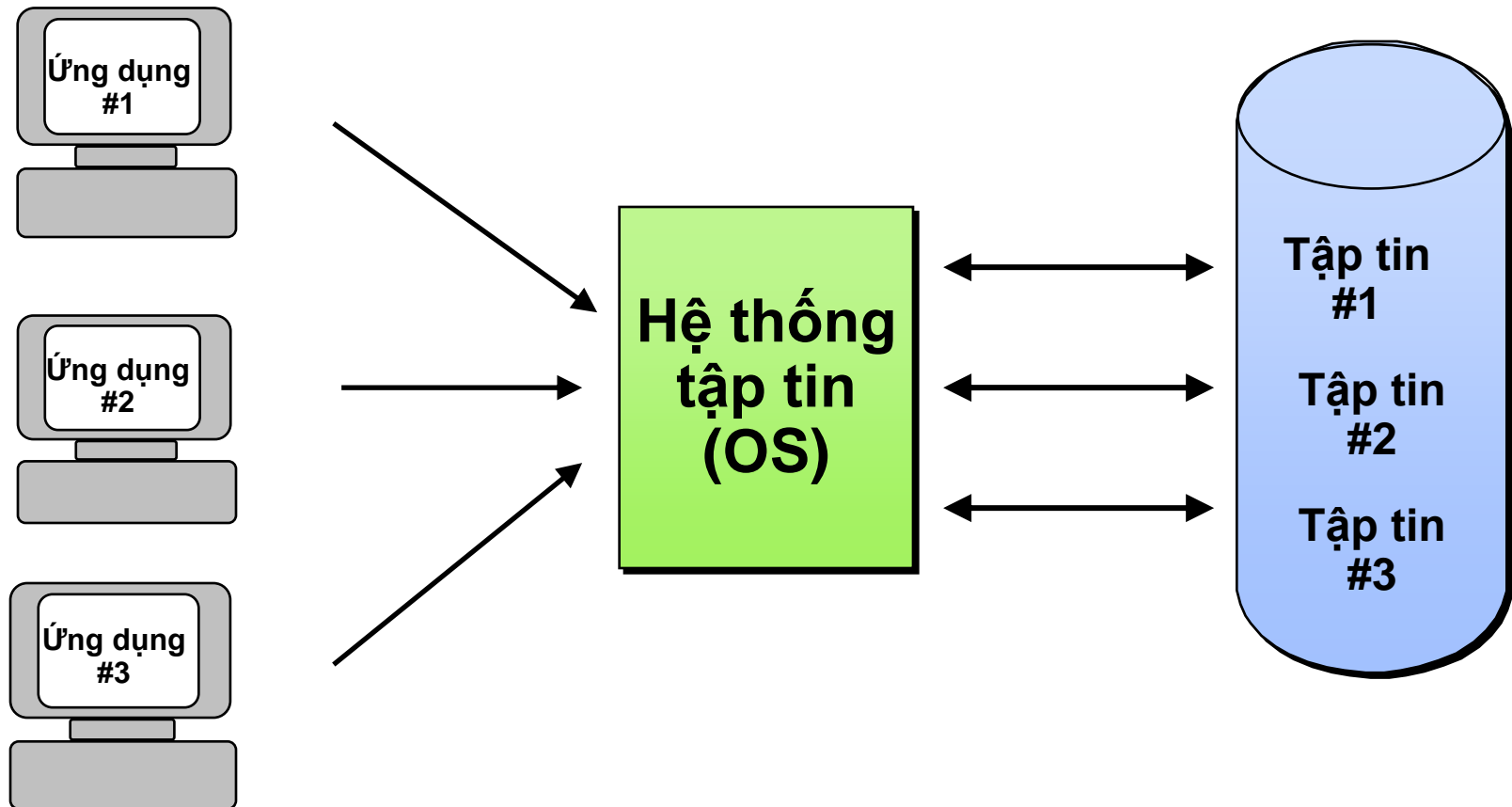
Hệ thống xử lý tập tin ở một công ty

Dữ liệu
trùng lặp



Hệ thống Xử lý tập tin

- Dữ liệu được lưu trữ dưới dạng nhiều tập tin riêng biệt



Hạn chế của phương pháp xử lý tập tin

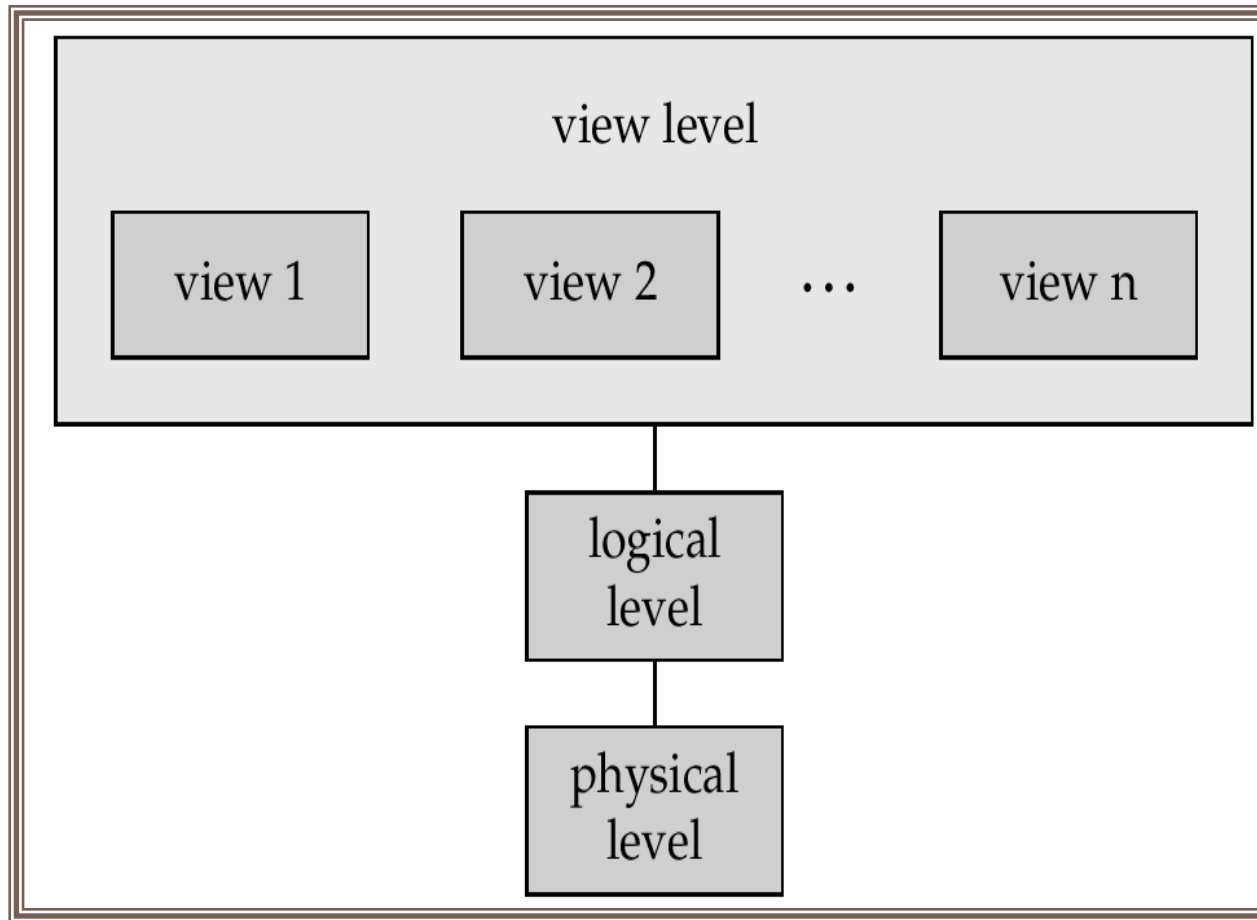
- Sự phụ thuộc của chương trình vào dữ liệu
- Sự dư thừa và trùng lặp dữ liệu
- Khả năng chia sẻ dữ liệu bị hạn chế
- Thời gian phát triển phần mềm kéo dài
- Chi phí bảo trì lớn

Cách giải quyết: Cơ sở dữ liệu

- Định nghĩa: CSDL là một tập hợp dữ liệu được tổ chức, lưu trữ dưới dạng có cấu trúc và được sử dụng (chia sẻ) bởi các hệ thống ứng dụng.
 - ▣ Tập trung các dữ liệu dùng chung
 - ▣ Toàn bộ dữ liệu được quản lý bằng một hệ thống thống nhất
 - ▣ Dữ liệu được lưu trữ theo một hình thức tiện lợi và tiêu chuẩn

**Cần một hệ quản trị cơ sở dữ liệu
(DBMS: Database Management System)**

Các cách nhìn khác nhau về dữ liệu



Các mức độ trừu tượng trong một DBMS

Các đặc điểm của cách tiếp cận CSDL

- Một kho dữ liệu được định nghĩa một lần, được duy trì và truy xuất bởi nhiều người
- Bản chất tự mô tả của một hệ CSDL: Một DBMS catalog lưu trữ mô tả về CSDL (meta-data)
- Phần mềm DBMS làm việc với nhiều ứng dụng CSDL
- Sự tách biệt giữa chương trình và dữ liệu, sự trừu tượng của dữ liệu
- Hỗ trợ nhiều cách nhìn về dữ liệu
- Chia sẻ dữ liệu và xử lý giao dịch nhiều người dùng

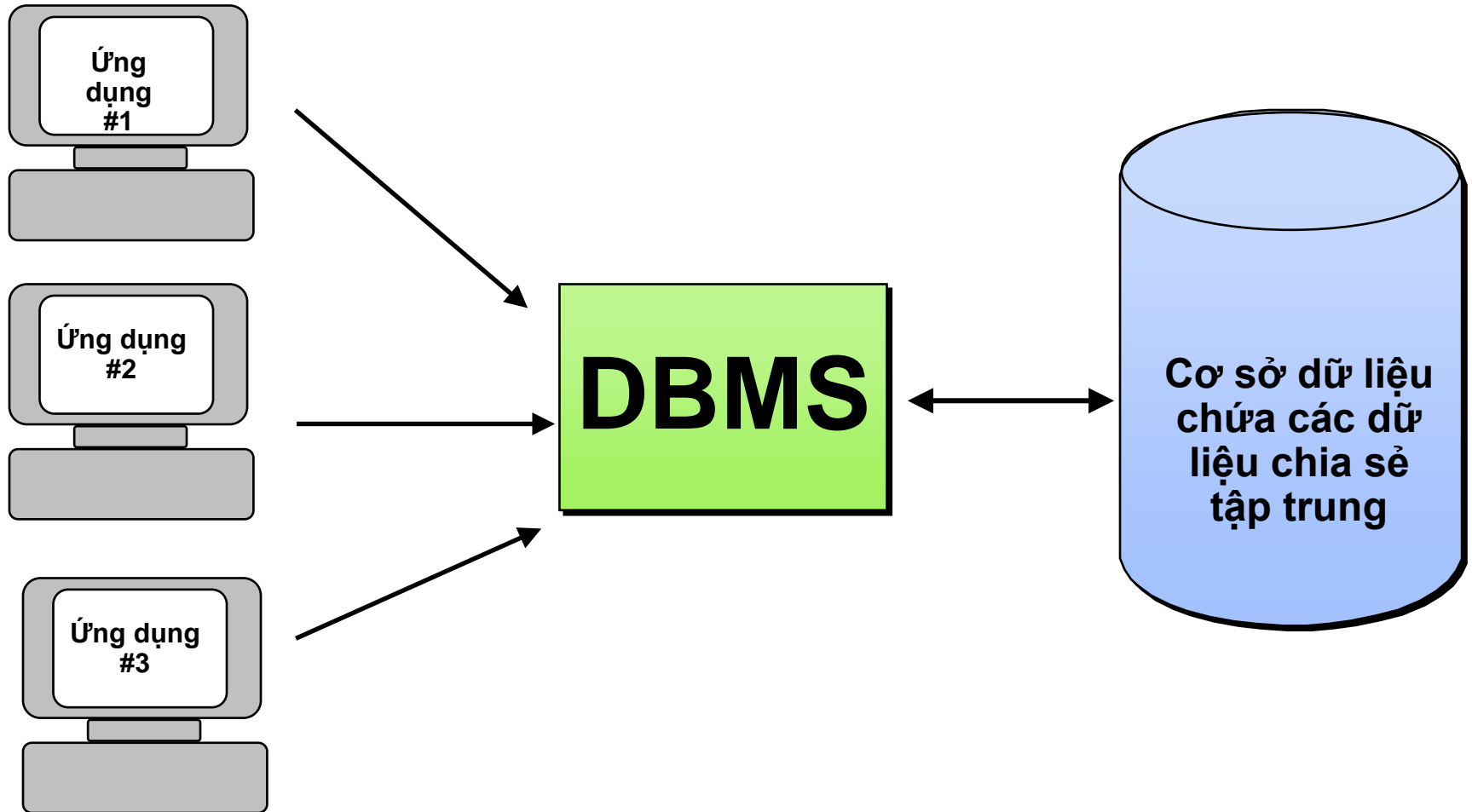
Sự cần thiết của DBMS

- CSDL đặt ra những vấn đề cần phải giải quyết:
 - ▣ Tính chủ quyền
 - ▣ Cơ chế bảo mật hay phân quyền khi khai thác CSDL
 - ▣ Giải quyết tranh chấp trong quá trình truy nhập dữ liệu
 - ▣ Phục hồi dữ liệu khi xảy ra sự cố
- Phần mềm có khả năng giải quyết các vấn đề trên được gọi là hệ quản trị CSDL (DBMS)

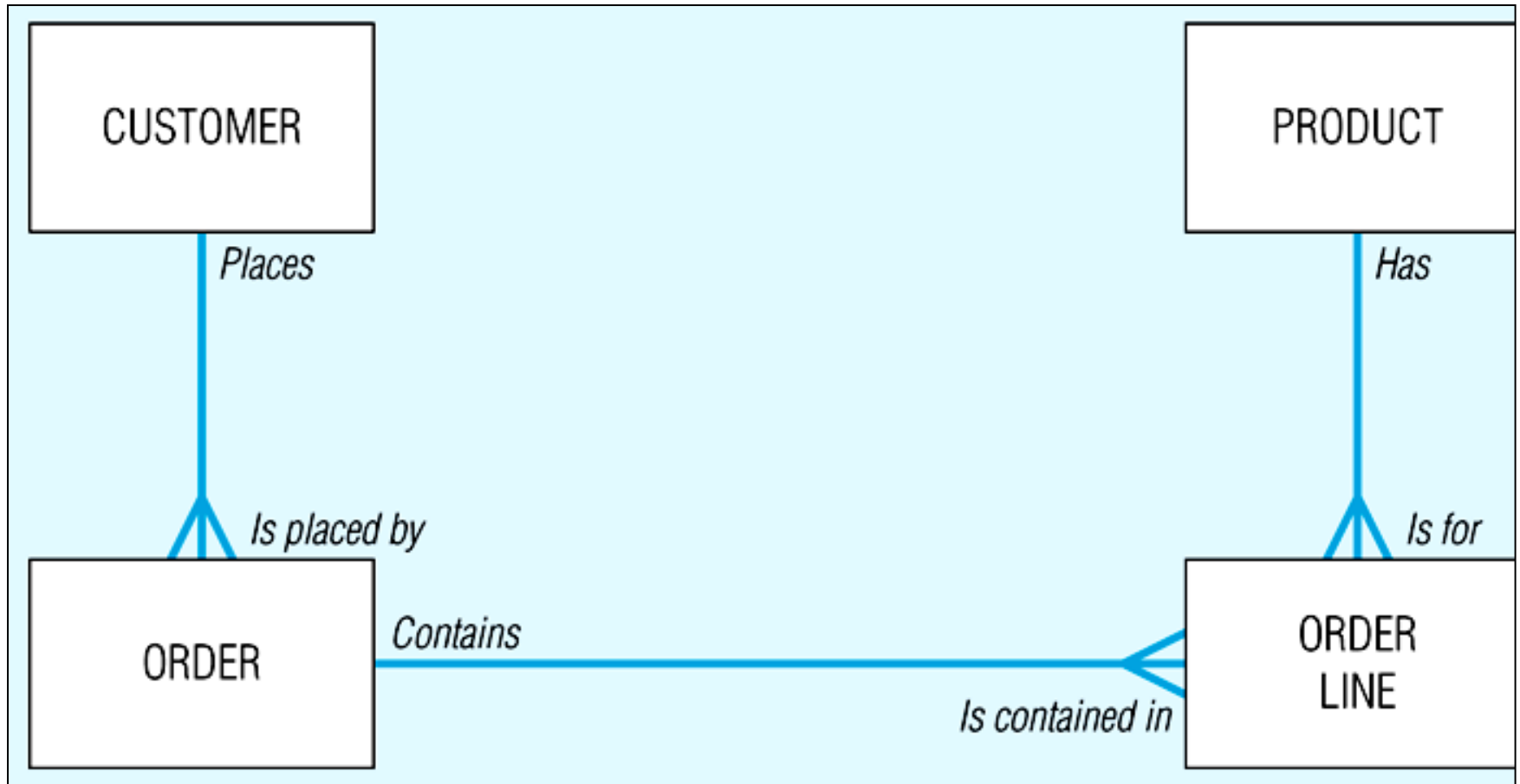
Hệ quản trị CSDL

- Cung cấp các công cụ hỗ trợ tích cực cho các nhà phân tích, thiết kế và người khai thác CSDL
- Trợ giúp cho quá trình định nghĩa, xây dựng và xử lý CSDL cho các ứng dụng khác nhau
- Một số DBMS: MS SQL Server, Oracle, MySQL, Sybase, DB-2, ...

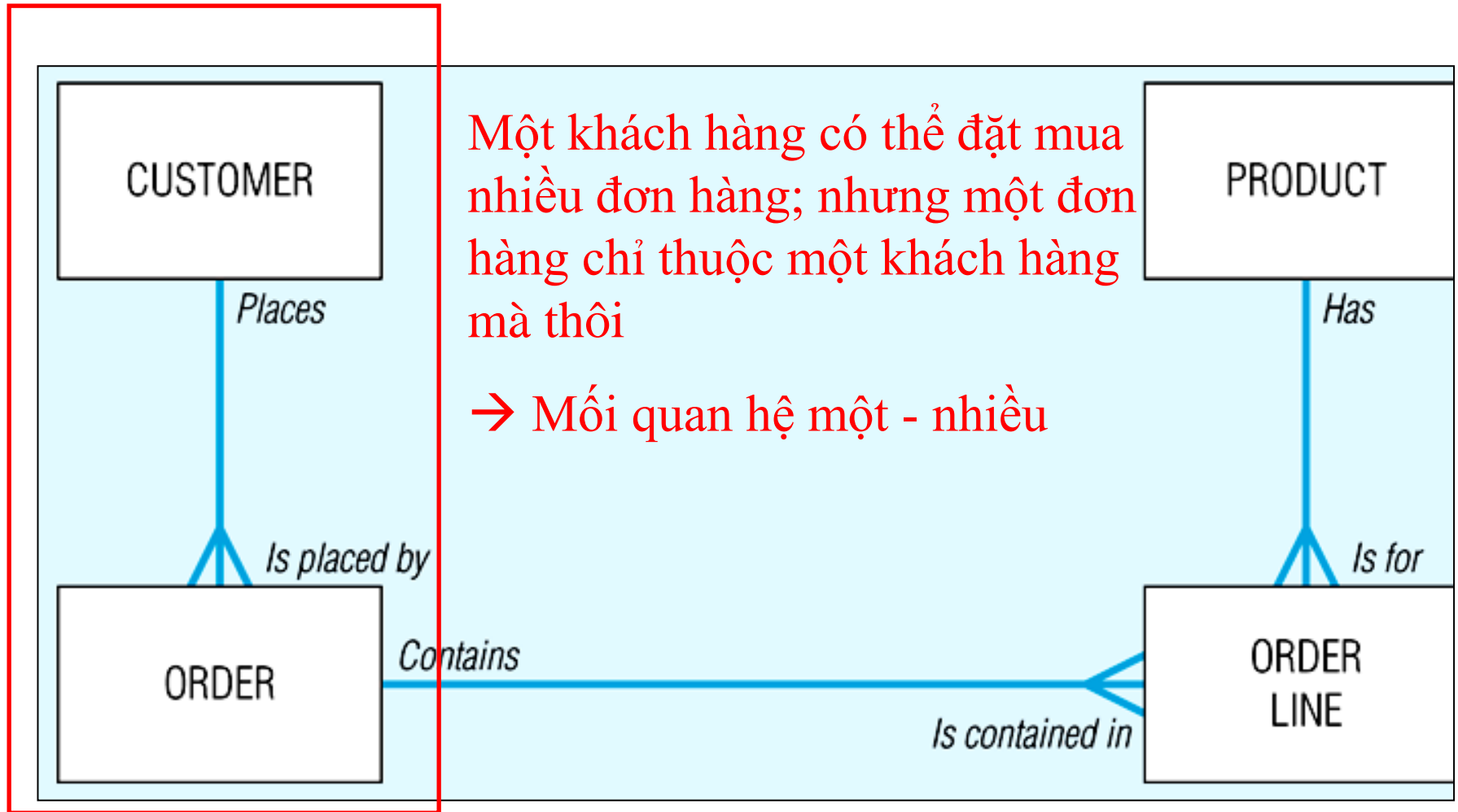
Hệ quản trị cơ sở dữ liệu (tt)



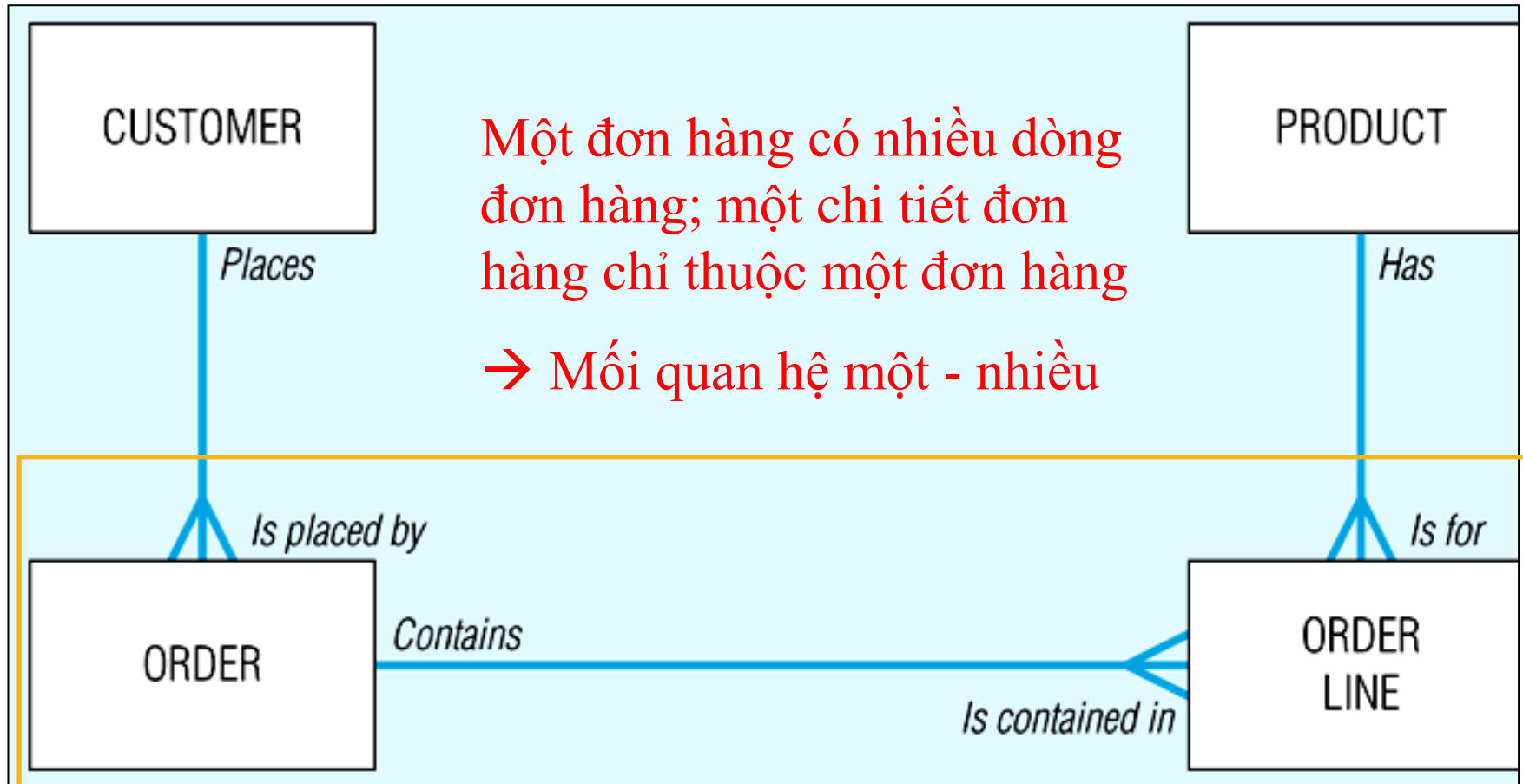
Mô hình dữ liệu thí dụ



Mô hình dữ liệu thí dụ (tt)



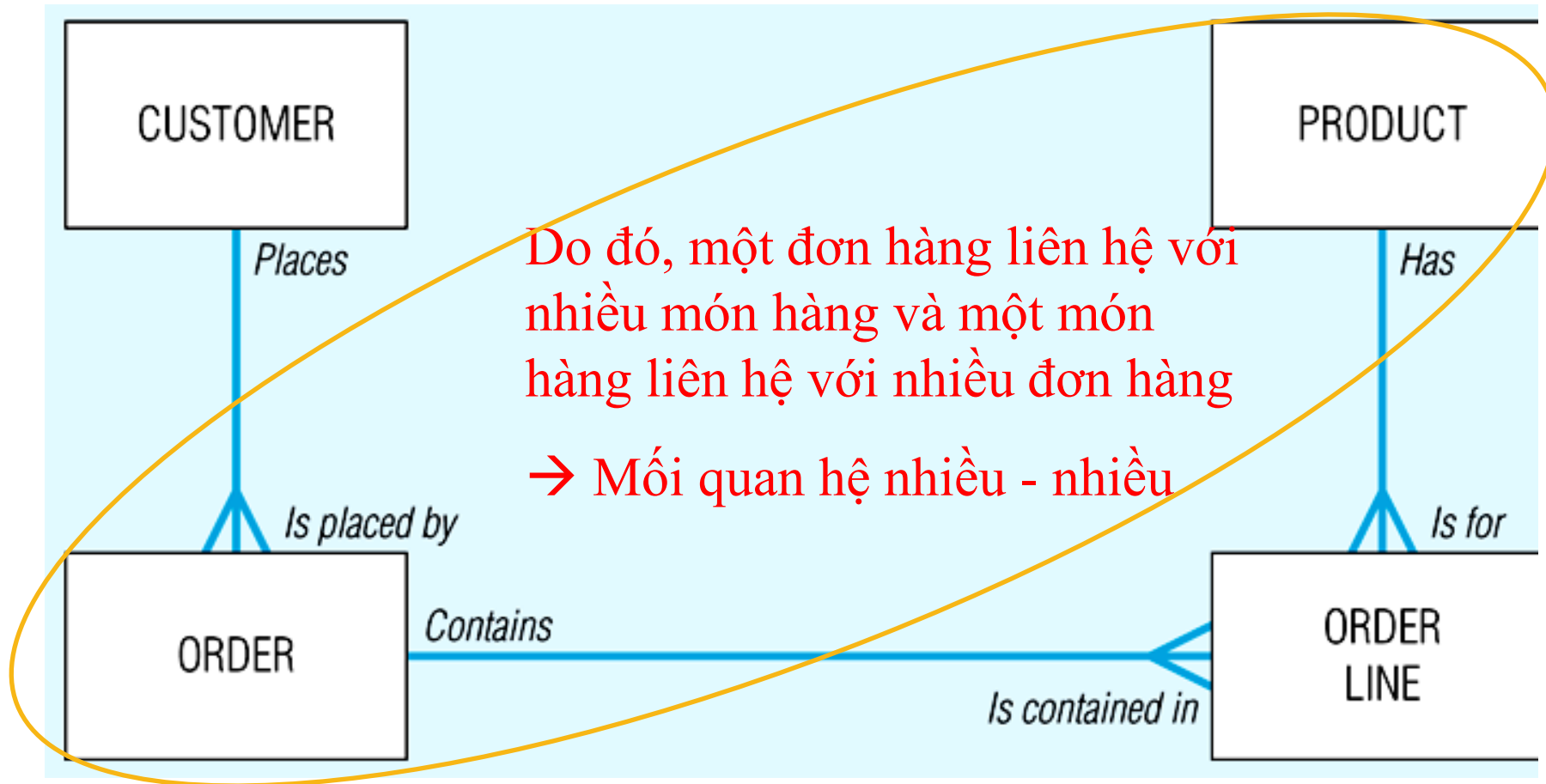
Mô hình dữ liệu thí dụ (tt)



Mô hình dữ liệu thí dụ (tt)



Mô hình dữ liệu thí dụ (tt)



Mô hình dữ liệu thí dụ (tt)

Các bảng dữ liệu Khách hàng, Đơn hàng, Chi tiết đơn hàng và Món hàng

The screenshot displays four tables in Microsoft Access:

- ORDERS Table:** Contains columns **Order_id** (primary key), **Order_Date**, and **Customer_ID** (foreign key to CUSTOMERS).
- ORDERITEMS Table:** Contains columns **Order_ID** (foreign key to ORDERS), **Product_ID** (foreign key to PRODUCTS), and **Quantity**.
- PRODUCTS Table:** Contains columns **Product ID** (primary key), **Product Description**, **Product Finish**, **Unit Price**, and **On Hand**.
- CUSTOMERS Table:** Contains columns **Customer ID** (primary key), **Customer Name**, **Address**, **City**, **State**, and **Postal Code**.

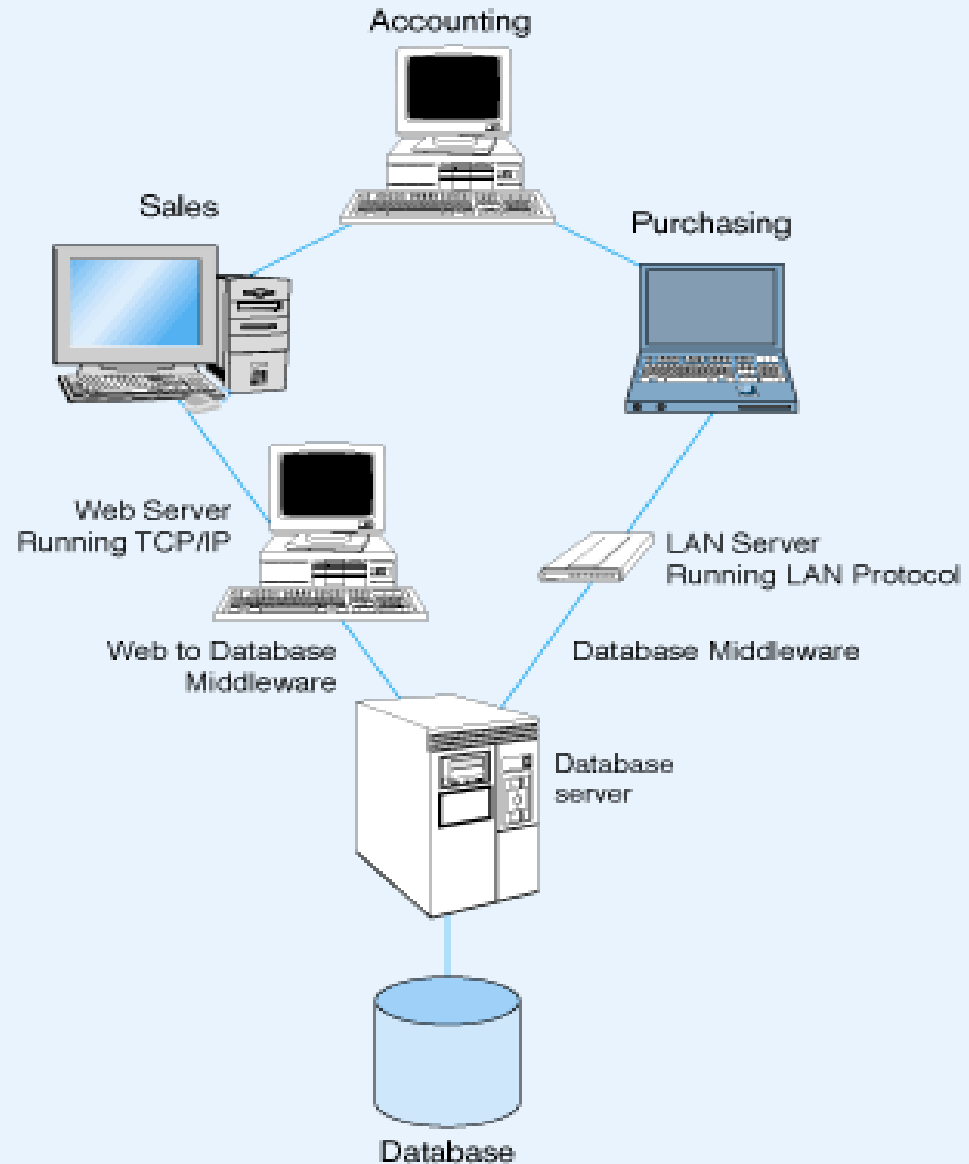
Relationships are indicated by yellow arrows:

- From **ORDERS** to **ORDERITEMS** on **Order_ID**.
- From **ORDERITEMS** to **PRODUCTS** on **Product_ID**.
- From **ORDERS** to **CUSTOMERS** on **Customer_ID**.

Các mối quan hệ được thiết lập trên các cột đặc biệt liên kết giữa các bảng

Mô hình dữ liệu thí dụ (tt)

Kiến trúc
client/server
hiện thực hệ
thống CSDL



Mô hình dữ liệu thí dụ (tt)

Chức năng chương trình ứng dụng: **nhập** dữ liệu mới, **cập nhật**, **xoá**, **đọc** dữ liệu để hiển thị

PVFC Customer Invoice

Customer ID: 2
Customer Name: Value Furniture
Address: 15145 S.W. 17th St.
Plano, TX 75094

Order Number: 1006
Order Date: 10/24/2000

Product_ID	Product Description	Finish	Quantity	Unit Price	Extended Price:
7	Dining Table	Natural As	2	\$800.00	\$1,600.00
5	Writer's Desk	Cherry	2	\$325.00	\$650.00
4	Entertainment Center	Natural M	1	\$650.00	\$650.00
Total					\$2,900.00

Phạm vi ứng dụng CSDL (tt)

**Các dữ liệu
thường thấy
từ một
CSDL cá
nhân**

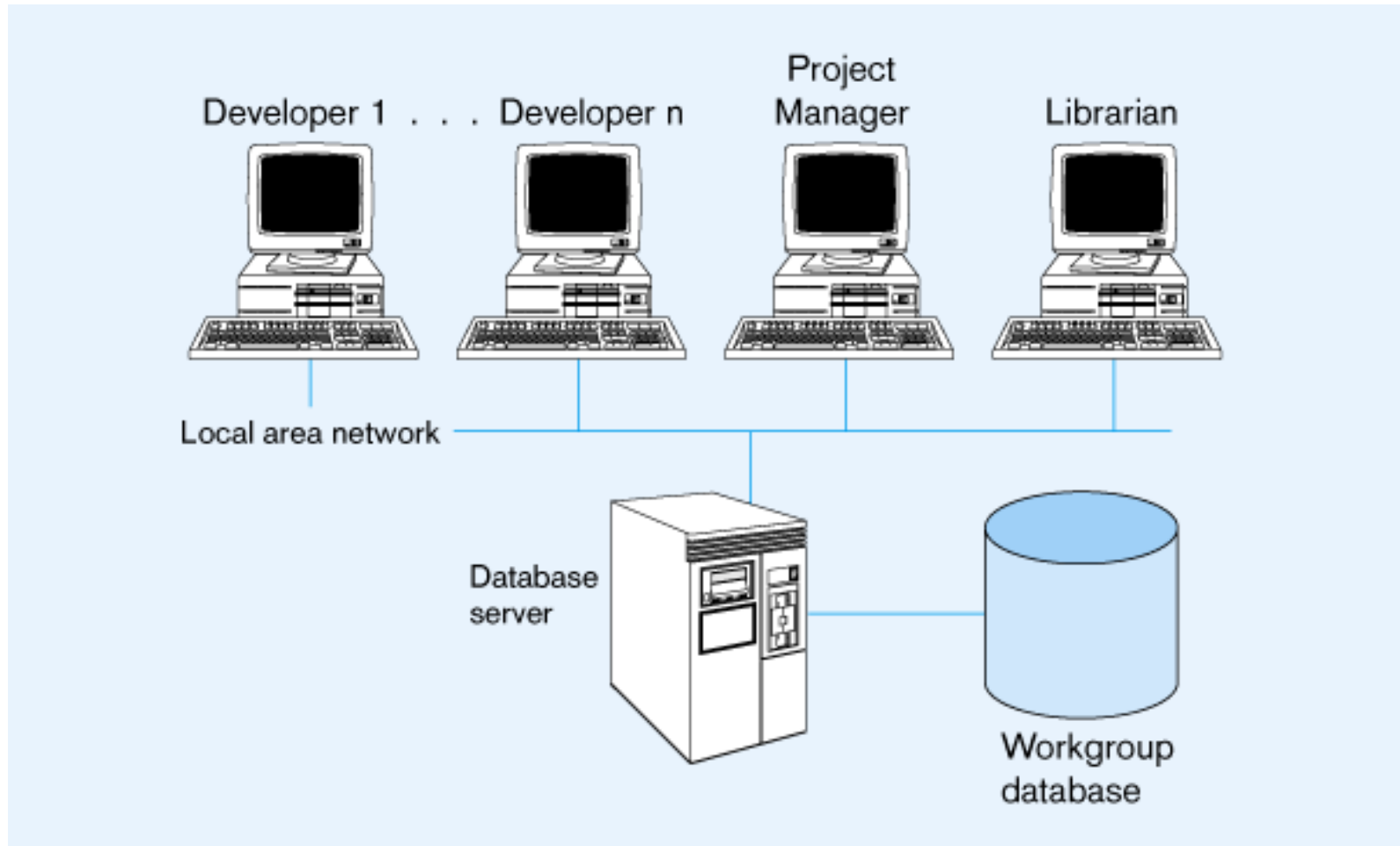
Customer

Customer Name: Multi Media, Inc.	
Address: 1000 River Road	
City: San Antonio	
State: TX	
Zip: 76235	
Phone: (219) 864-2000	
Next Contact Date: 10/17/2000	Time: 10:30 AM

Contact History for Customer

Date	Time	Contact	Comments
08/04/2000	10:00 AM	Roberts	Review proposal
08/19/2000	08:00 AM	Roberts	Revise schedule
09/10/2000	09:00 AM	Pearson	Sign contract
09/21/2000	02:00 PM	Roberts	Follow up

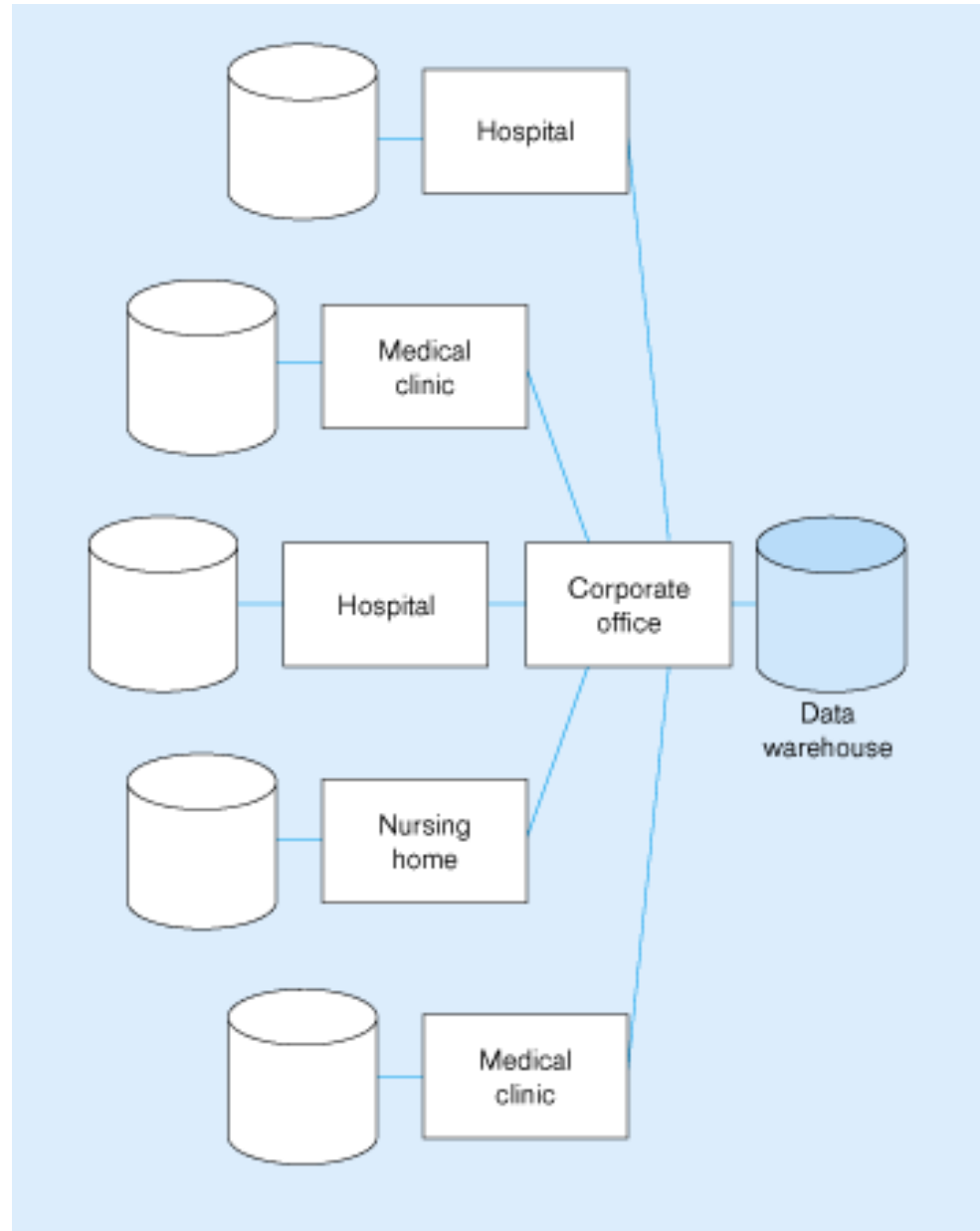
Phạm vi ứng dụng CSDL (tt)



CSDL nhóm với mạng cục bộ

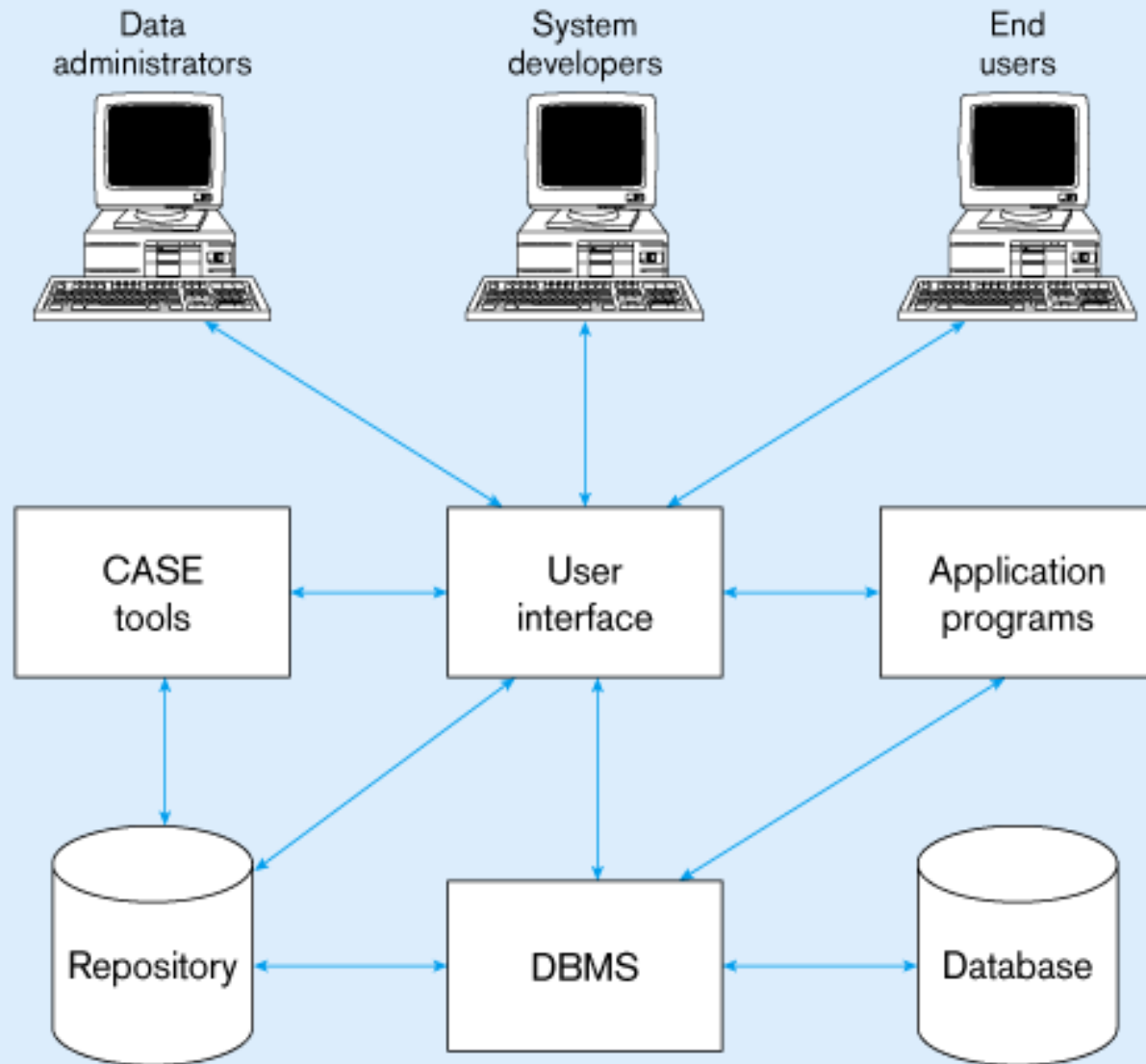
Phạm vi ứng dụng CSDL (tt)

**Một kho dữ
liệu xí
nghiệp**



Các thành phần của môi trường CSDL (tt)

**Các
thành
phần
của môi
trường
CSDL**



Nghiên cứu CSDL: Thành tựu và Cơ hội

- Các ứng dụng CSDL mới:
 - ▣ Thương mại điện tử - Electronic Commerce
 - ▣ Hệ thống thông tin y tế
 - ▣ Thư viện điện tử và xuất bản kỹ thuật số
 - ▣ Thiết kế cộng tác – Collaborative Design
- Các xu hướng tác động tới việc nghiên cứu CSDL:
 - ▣ Dữ liệu đa phương tiện
 - ▣ Xa lộ thông tin (dữ liệu lớn)
 - ▣ Các kiến trúc lưu trữ và mạng

Nghiên cứu CSDL: Thành tựu và Cơ hội (tt)

- Hướng nghiên cứu mới:
 - ▣ Hỗ trợ Multimedia Objects
 - ▣ Trực quan hóa dữ liệu với người dùng cuối (visualization of data)
 - ▣ Phân phối và an toàn thông tin, ...
- Cách dùng mới đối với hệ CSDL:
 - ▣ Data Mining, Data Warehouses, OLAP, Digital Libraries, Repositories, Big Data, ...

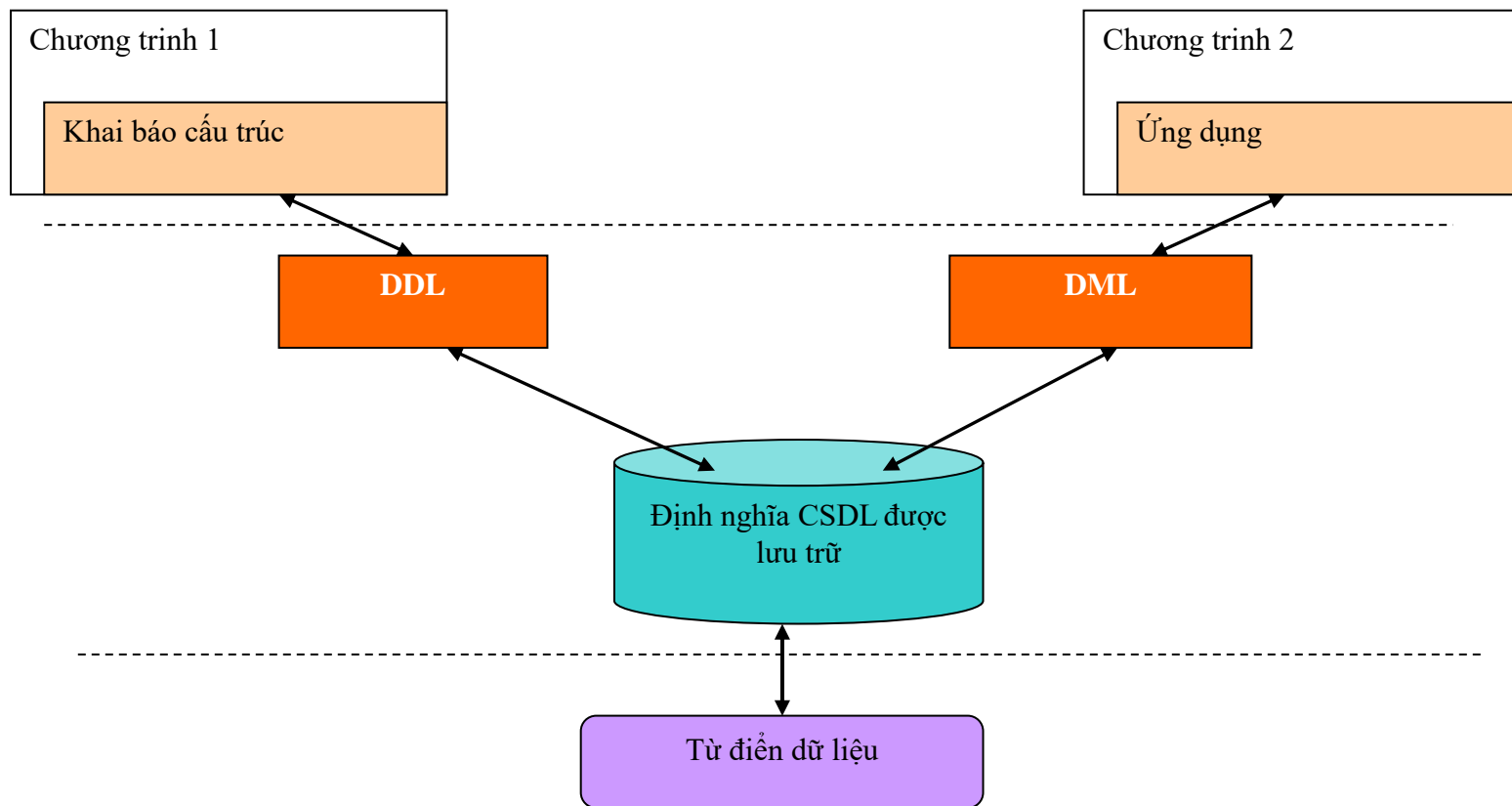
Các yêu cầu của một DBMS

- Khả năng đáp ứng nhanh
- Độ tin cậy cao
- Lưu lượng lớn
- Chu kỳ sống lâu
- An toàn

Phân loại DBMS

- Dựa trên mô hình dữ liệu được dùng:
 - ▣ Truyền thống: quan hệ, mạng, phân cấp
 - ▣ **Đang thịnh hành**: hướng đối tượng, quan hệ đối tượng
- Các sự phân loại khác:
 - ▣ Single-user và Multi-user
 - ▣ Tập trung và Phân tán

Mô hình 3 mức của DBMS



Các ngôn ngữ của DBMS

- Ngôn ngữ định nghĩa dữ liệu (DDL):
 - ▣ được sử dụng bởi những người thiết kế CSDL để chỉ rõ lược đồ mức quan niệm của một CSDL
- Ngôn ngữ thực thi dữ liệu (DML):
 - ▣ được dùng để chỉ rõ việc rút trích và cập nhật CSDL
 - ▣ Các lệnh DML (data sublanguage) có thể được nhúng trong một ngôn ngữ lập trình tổng quát (host language: C#, Java...)
 - ▣ có thể được dùng trực tiếp (query language) - stand-alone DML.

Quá trình hoạt động của một chương trình ứng dụng thông qua các tầng của CSDL

