



# **CƠ SỞ DỮ LIỆU**

## **BÀI 1: TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU**

### **PHẦN 1**

- ⊙ Kết thúc bài học này bạn sẽ nắm được các kiến thức sau:
  - ⊙ Hiểu về khái niệm dữ liệu và cơ sở dữ liệu
  - ⊙ Giới thiệu về mô hình dữ liệu quan hệ
  - ⊙ Khái niệm hệ quản trị CSDL (DBMS) và hệ quản trị CSDL quan hệ (RDBMS)
  - ⊙ Giới thiệu về hệ thống client/server

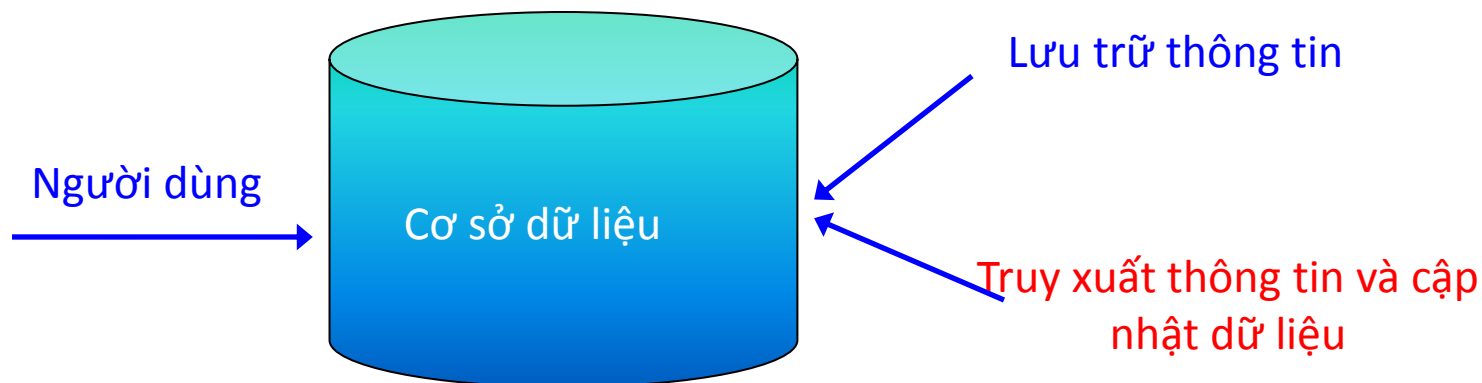


# **KHÁI NIỆM CƠ BẢN VỀ DỮ LIỆU VÀ CƠ SỞ DỮ LIỆU**

- **Dữ liệu (data)**
  - Là các thông tin của đối tượng (người, vật, một khái niệm, sự việc...) được lưu trữ trên máy tính.
  - Có thể truy nhập vào dữ liệu để trích xuất thông tin.
- **Dữ liệu được mô tả dưới nhiều dạng khác nhau** (các ký tự, ký số, hình ảnh, ký hiệu, âm thanh...). Mỗi cách mô tả gắn với một ngữ nghĩa nào đó.

- Dữ liệu về đối tượng có thể khác nhau, tùy thuộc vào ngữ cảnh.
- Ví dụ: dữ liệu về đối tượng sinh viên có thể khác nhau tùy vào mục đích quản lý:
  - Quản lý điểm: Tên, mã sinh viên, điểm môn 1, điểm môn 2, điểm môn 3
  - Quản lý nhân thân: Tên, địa chỉ, ngày sinh, quê quán, lớp

- Cơ sở dữ liệu (CSDL) = Tập hợp dữ liệu được tổ chức có cấu trúc liên quan với nhau và được lưu trữ trong máy tính.
- CSDL được thiết kế, xây dựng cho phép người dùng lưu trữ dữ liệu, truy xuất thông tin hoặc cập nhật dữ liệu



- CSDL được tổ chức có cấu trúc:
    - Các dữ liệu được lưu trữ có cấu trúc thành các bản ghi (record), các trường dữ liệu (field).
    - Các dữ liệu lưu trữ có mối quan hệ (**relation**) với nhau.
  - Khả năng truy xuất thông tin từ CSDL:
    - CSDL được cấu trúc để dễ dàng truy cập, quản lý và cập nhật
- ➔ Cần phải quản trị CSDL

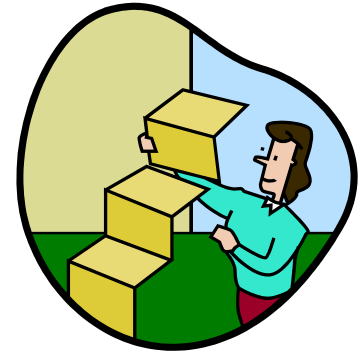
■ Ví dụ một CSDL:



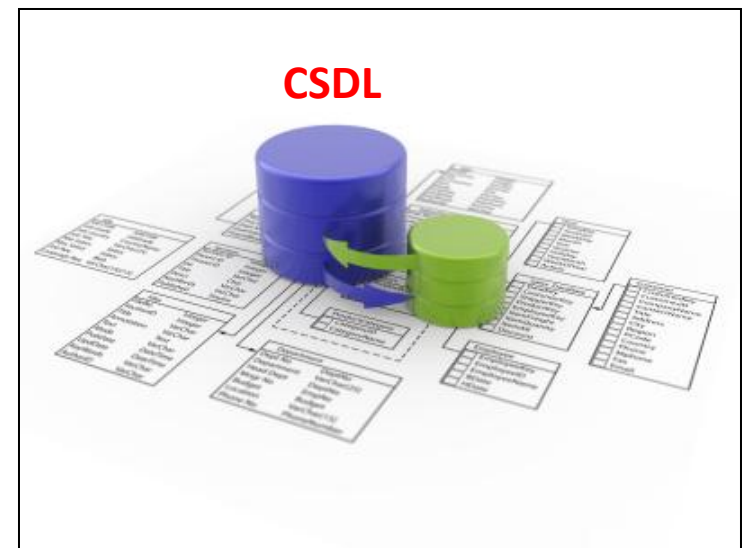


- **Quản lý dữ liệu:** là quản lý một số lượng lớn dữ liệu, bao gồm cả việc lưu trữ và cung cấp cơ chế cho phép **Thao tác** (thêm, sửa, xóa dữ liệu) và **Truy vấn** dữ liệu.
- 2 phương pháp quản lý dữ liệu:
  - Hệ thống quản lý bằng file
  - Hệ thống quản lý bằng CSDL

- Dữ liệu được lưu trữ trong các file riêng biệt
- Ví dụ: các chương trình lưu trữ thông tin bằng hệ thống các file dạng text
- Nhược điểm của việc quản lý bằng file:
  - Dư thừa và mâu thuẫn dữ liệu
  - Kém hiệu quả trong truy xuất ngẫu nhiên hoặc xử lý đồng thời
  - Dữ liệu lưu trữ rời rạc
  - Gặp vấn đề về an toàn và bảo mật



- Quản lý dữ liệu bằng CSDL giúp dữ liệu được lưu trữ một cách hiệu quả và có tổ chức, cho phép quản lý dữ liệu nhanh chóng và hiệu quả
- Lợi ích của hệ thống quản lý bằng CSDL:
  - Tránh dư thừa, trùng lặp dữ liệu
  - Đảm bảo sự nhất quán trong CSDL
  - Các dữ liệu lưu trữ có thể được chia sẻ
  - Có thể thiết lập các chuẩn trên dữ liệu
  - Duy trì tính toàn vẹn dữ liệu
  - Đảm bảo bảo mật dữ liệu

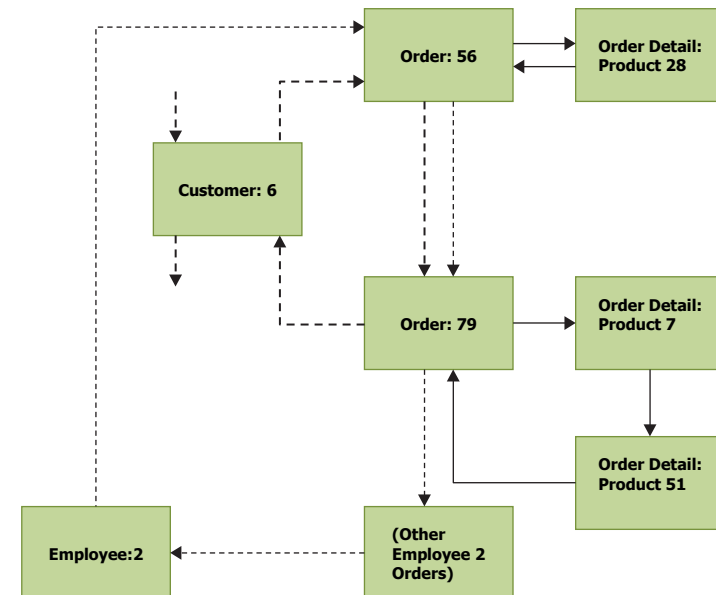


- Các CSDL có thể khác nhau về chức năng và mô hình dữ liệu (data model).
- Mô hình dữ liệu sẽ quyết định cách thức lưu trữ và truy cập dữ liệu.
- Tùy từng ngữ cảnh quan hệ giữa các thành phần dữ liệu trong CSDL, mô hình phức hợp được áp dụng để việc lưu trữ và truy xuất dữ liệu đạt hiệu quả cao nhất.

## ■ Các mô hình:

- Mô hình dữ liệu file phẳng (Flat file)
- Mô hình dữ liệu mạng (Network model)
- Mô hình dữ liệu phân cấp (Hierarchical model)
- Mô hình dữ liệu quan hệ (Relational model)
- Mô hình dữ liệu hướng đối tượng (Object-Oriented model)

Customer ID	Company Name	Contact First Name	Contact Last Name	Job Title	City	State
6	Company F	Francisco	Pérez-Olaeta	Purchasing Manager	Milwaukee	WI
26	Company Z	Run	Liu	Accounting Assistant	Miami	FL



- Mô hình này chỉ dùng cho các CSDL đơn giản
- CSDL dạng file phẳng thường là file kiểu văn bản chứa dữ liệu dạng bảng

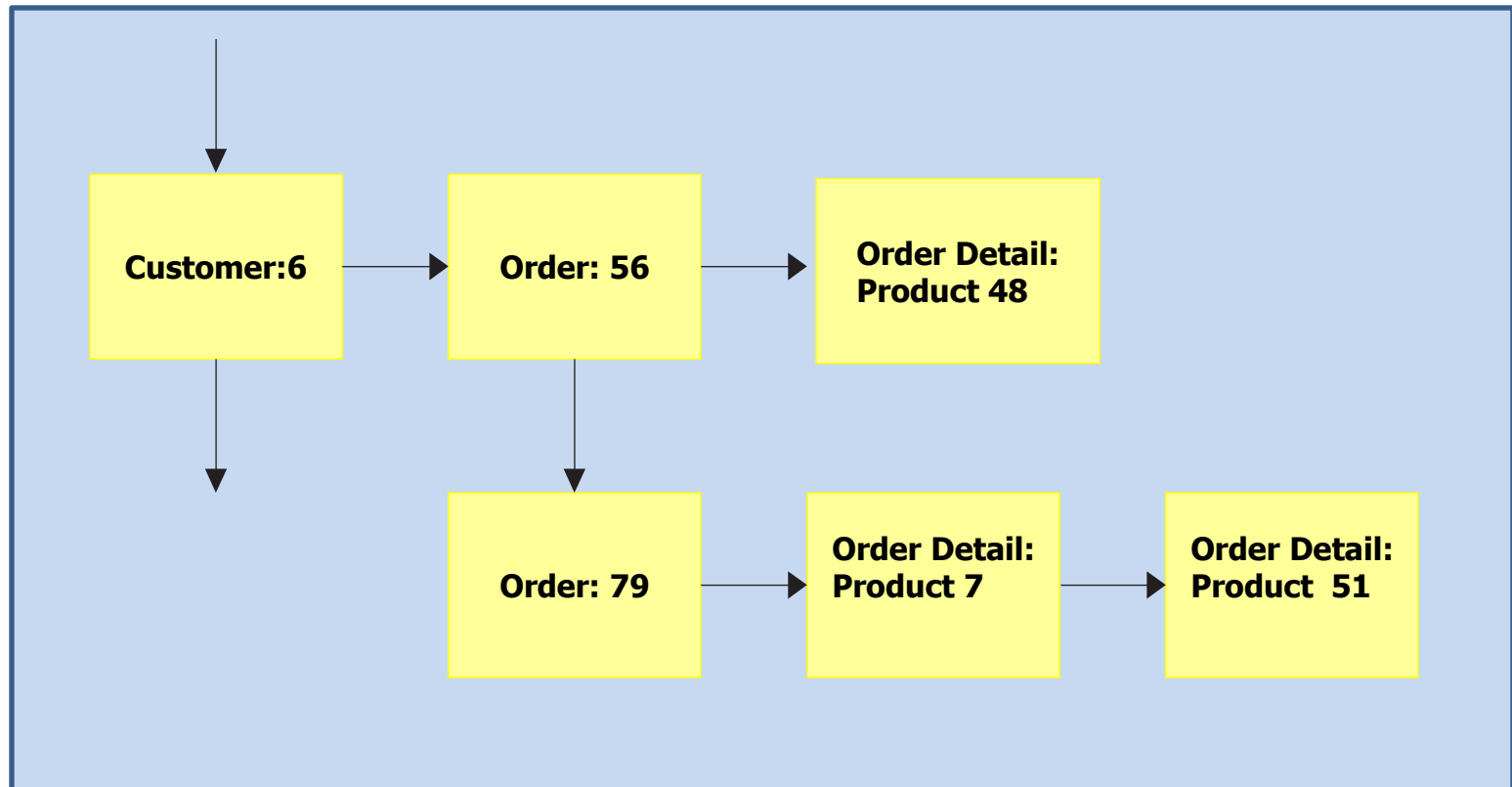
- ❑ Ví dụ: một file phẳng thể hiện thông tin về Customer (Khách hàng) dưới dạng bảng của công ty Northwind Traders

Customer ID	Company Name	Contact First Name	Contact Last Name	Job Title	City	State
6	Company F	Francisco	Pérez-Olaeta	Purchasing Manager	Milwaukee	WI
26	Company Z	Run	Liu	Accounting Assistant	Miami	FL

- Tổ chức theo hình cây, mỗi nút biểu diễn một thực thể dữ liệu.
- Liên hệ dữ liệu thể hiện trên liên hệ giữa nút cha và nút con. Mỗi nút cha có thể có một hoặc nhiều nút con, nhưng mỗi nút con chỉ có thể có một nút cha.
- Do đó mô hình phân cấp thể hiện các kiểu quan hệ:
  - 1-1
  - 1-N



- Ví dụ: mô hình dữ liệu phân cấp của CSDL Northwind

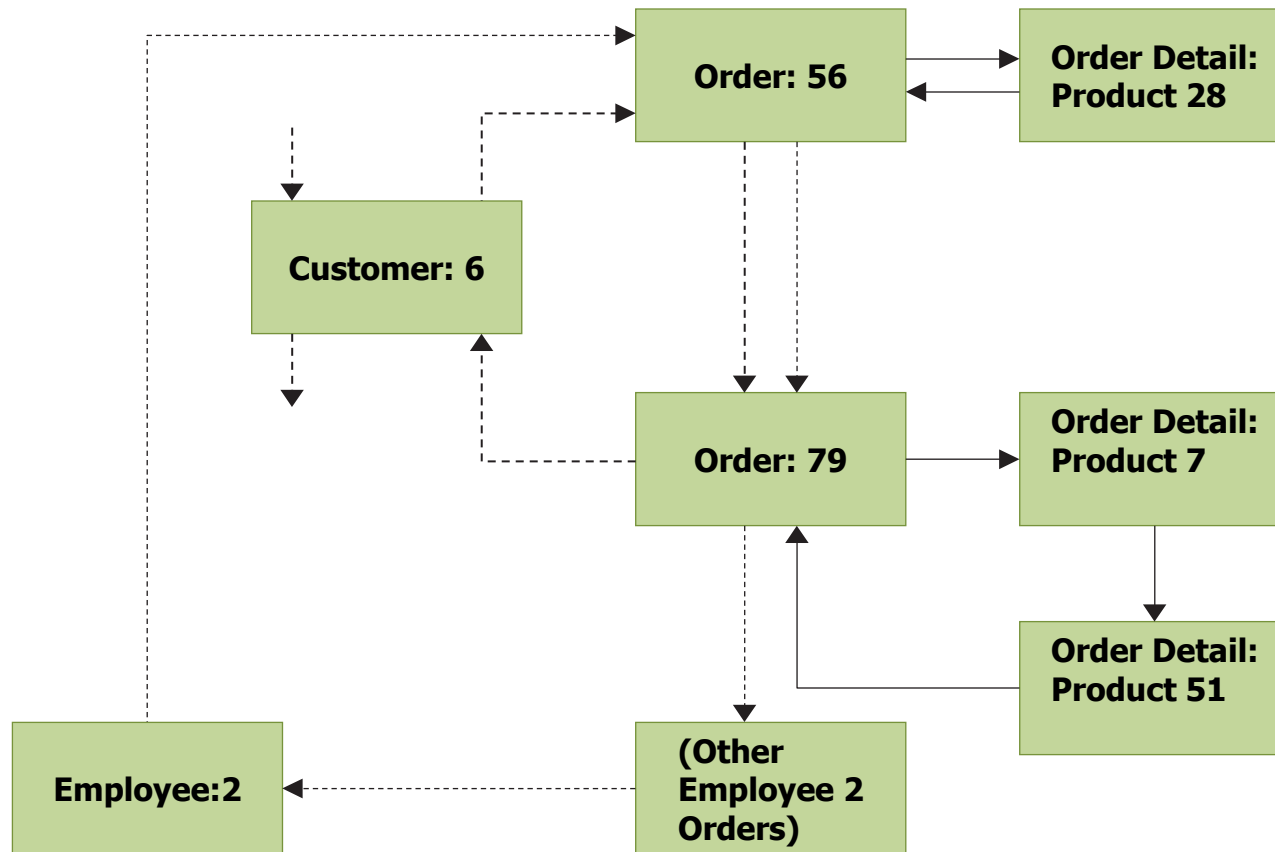


## ■ Hạn chế:

- Một nút con không thể có quá một nút cha -> Không biểu diễn được các quan hệ dữ liệu phức tạp

- Cách tổ chức:
  - Các file riêng biệt trong hệ thống file phẳng được gọi là các *bản ghi* . Tập hợp bản ghi cùng kiểu tạo thành một kiểu thực thể dữ liệu.
  - Các kiểu thực thể kết nối với nhau thông qua mối quan hệ cha-con.
  - Mô hình dữ liệu mạng biểu diễn bởi một đồ thị có hướng, và các mũi tên chỉ từ kiểu thực thể cha sang kiểu thực thể con.

- Ví dụ: Mô hình dữ liệu mạng của CSDL Northwind



## ■ Tính chất:

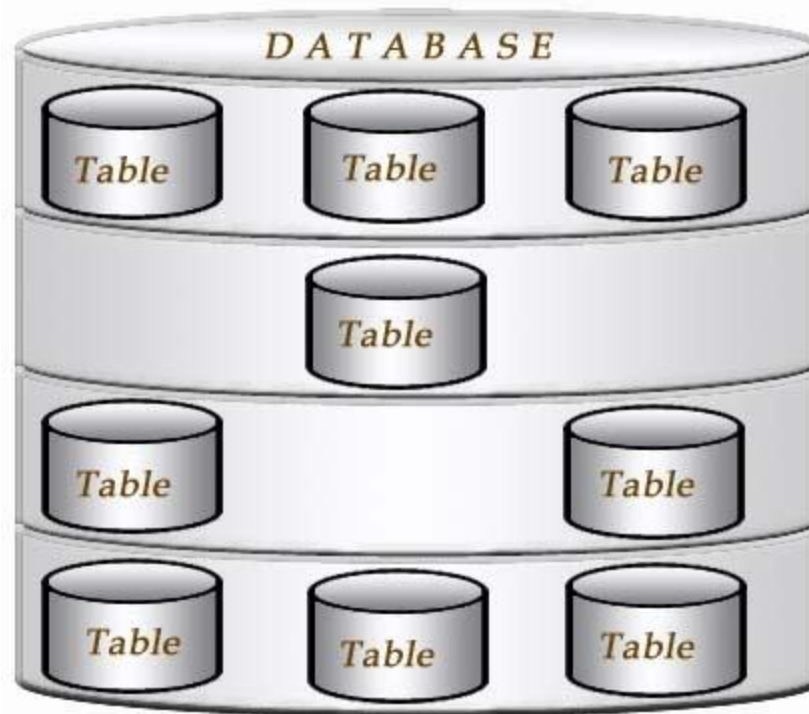
### ● Ưu điểm:

- Dễ biểu diễn mô hình
- Diễn đạt được các liên hệ dữ liệu phức tạp

### ● Nhược điểm:

- Truy xuất chậm
- Không thích hợp với các CSDL có quy mô lớn

- ❑ Dữ liệu trong database được lưu trữ vào các bảng



- Dữ liệu được biểu diễn dưới dạng bảng với các hàng và các cột:
  - CSDL là tập hợp các bảng (còn gọi là quan hệ)
  - Mỗi hàng là một bản ghi (record), còn được gọi là bộ (tuple)
  - Mỗi cột là một thuộc tính, còn được gọi là trường (field)
- Dữ liệu trong hai bảng liên hệ với nhau thông qua các cột chung.

File Customer

Customer ID	Company Name	Contact First Name	Contact Last Name	Job Title	City	State
6	Company F	Francisco	Pérez-Olaeta	Purchasing Manager	Milwaukee	WI
26	Company Z	Run	Liu	Accounting Assistant	Miami	FL

Bảng Order

Order ID	Customer ID	Employee ID	Order Date	Shipped Date	Shipping Fee
51	26	9	4/5/2006	4/5/2006	\$60.00
56	6	2	4/3/2006	4/3/2006	\$ 0.00
79	6	2	6/23/2006	6/23/2006	\$ 0.00

Bảng Employee

Employee ID	First Name	Last Name	Title
2	Andrew	Cencini	Vice President, Sales
5	Steven	Thrope	Sales Manager
9	Anne	Hellung-Larsen	Sales Representative



- ❖ Relational database
- ❖ Table – Bảng
- ❖ Column – Cột
- ❖ Row - Hàng (bản ghi)
- ❖ Primary Key – Khoá chính
- ❖ Unique Key – Khoá duy nhất
- ❖ Index – Chỉ mục
- ❖ Quan hệ một- một (1-1)
- ❖ Quan hệ một – nhiều (1-N)
- ❖ Quan hệ nhiều nhiều (N-N)



# **CƠ SỞ DỮ LIỆU**

## **BÀI 1: TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU**

### **PHẦN 2**

# HỆ QUẢN TRỊ CSDL VÀ HỆ QUẢN TRỊ CSDL QUAN HỆ

- Hệ quản trị CSDL (DataBase Management System – DBMS) là các phần mềm giúp tạo các CSDL và cung cấp cơ chế lưu trữ, truy cập theo các mô hình CSDL.
- Ví dụ:
  - SQL Server, Microsoft Access, Oracle là các hệ quản trị CSDL điển hình cho mô hình quan hệ.
  - IMS của IBM là hệ quản trị CSDL cho mô hình phân cấp
  - IDMS là hệ quản trị CSDL cho mô hình mạng

- Những lợi ích DBMS mang lại:
  - Quản trị các CSDL
  - Cung cấp giao diện truy cập để che dấu các đặc tính phức tạp về mặt cấu trúc tổ chức dữ liệu vật lý
  - Hỗ trợ các ngôn ngữ giao tiếp. Ví dụ:
    - Ngôn ngữ mô tả, định nghĩa dữ liệu – DDL
    - Ngôn ngữ thao tác dữ liệu – DML
    - Ngôn ngữ truy vấn dữ liệu có cấu trúc – SQL
  - Có cơ chế an toàn, bảo mật cao

- Hệ quản trị CSDL quan hệ (Relational DataBase Management System = RDBMS)
- RDMBS là một dạng DBMS được sử dụng phổ biến nhất, trong đó tất cả dữ liệu được tổ chức chặt chẽ dưới dạng các **bảng** dữ liệu.
- Tất cả các thao tác trên CSDL đều diễn ra trên các bảng.

- ❑ Rất nhiều người dùng tham gia vào hệ thống RDBMS:
  - ❖ Người quản trị CSDL ([DataBase Administrator](#))
  - ❖ Người thiết kế CSDL ([DataBase Designer](#))
  - ❖ Người phân tích hệ thống ([System Analyst](#))
  - ❖ Người lập trình ứng dụng ([Application Programmer](#))
  - ❖ Người thiết kế và triển khai CSDL ([DBMS Designer and Implementer](#))
  - ❖ Người dùng cuối ([End User](#))

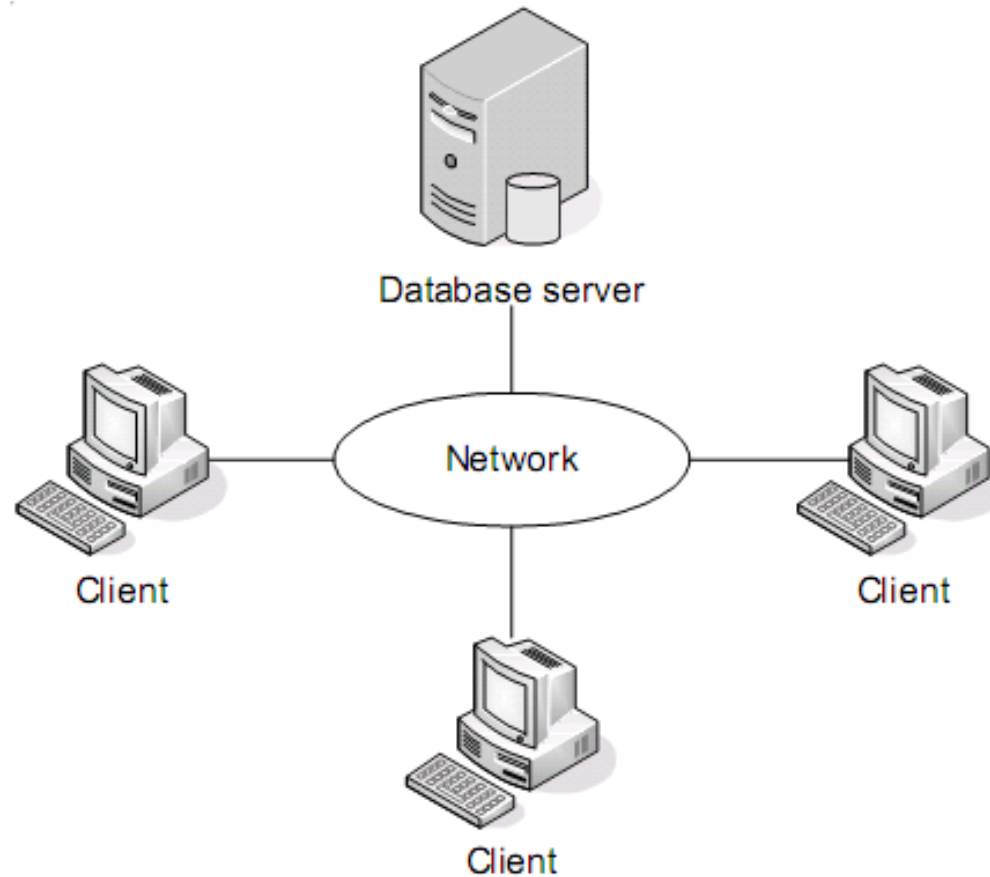


- ☐ Dễ dàng định nghĩa, duy trì và thao tác dữ liệu lưu trữ
- ☐ Dễ dàng trích xuất dữ liệu
- ☐ Dữ liệu được chuẩn hóa và được bảo vệ tốt
- ☐ Nhiều nhà cung cấp cung cấp phần mềm
- ☐ Dễ dàng chuyển đổi giữa nhà cung cấp và nhà triển khai
- ☐ RDBMS là các sản phẩm trưởng thành và ổn định

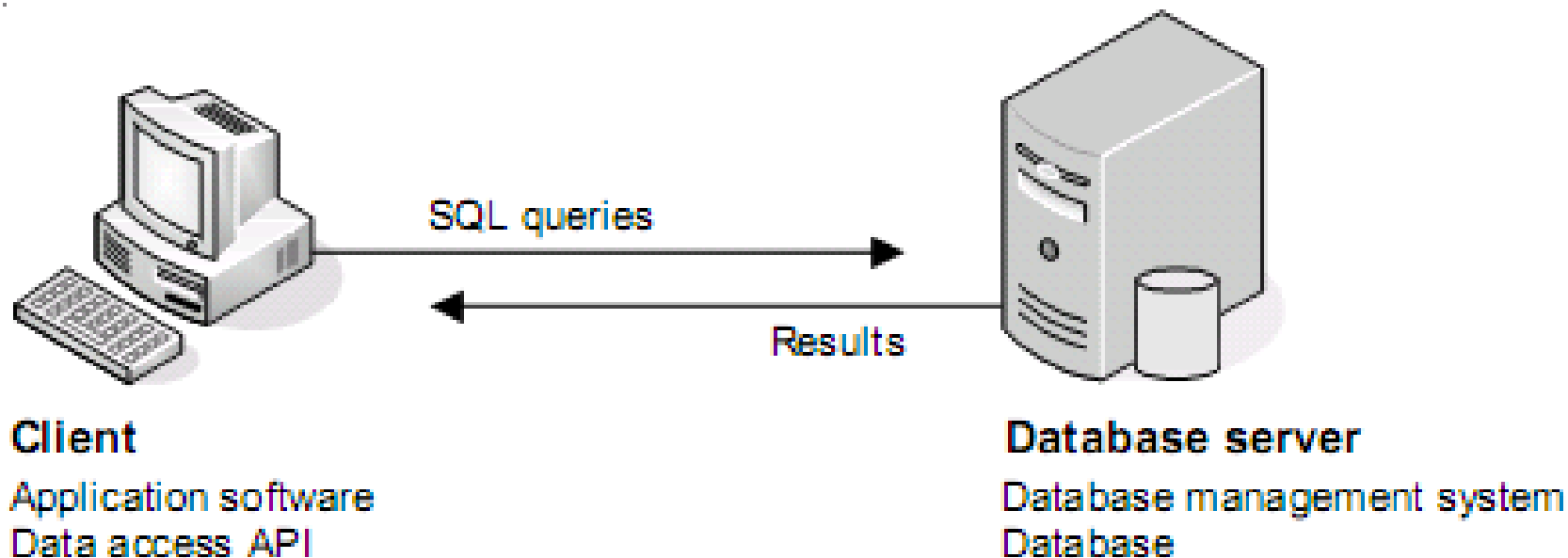


# CƠ SỞ DỮ LIỆU TRONG HỆ THỐNG CLIENT/SERVER

- ❑ Các thành phần phần cứng của hệ thống client/server

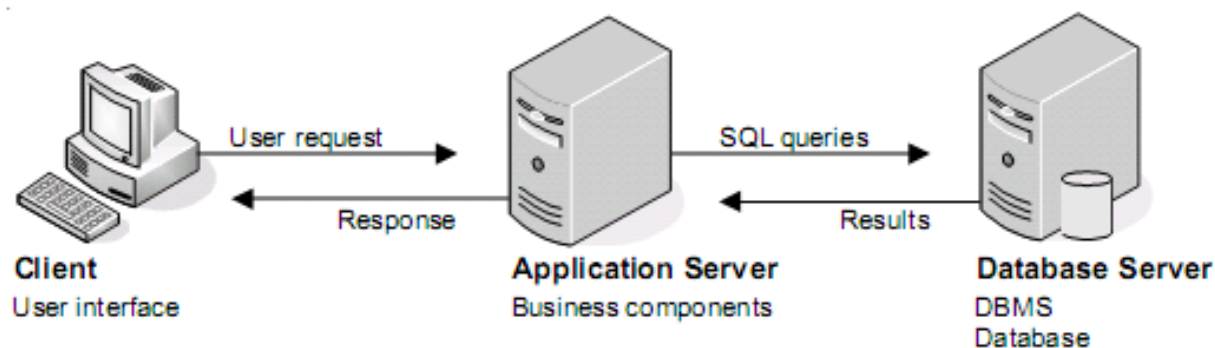


## ❑ Các thành phần phần mềm của hệ thống client/server

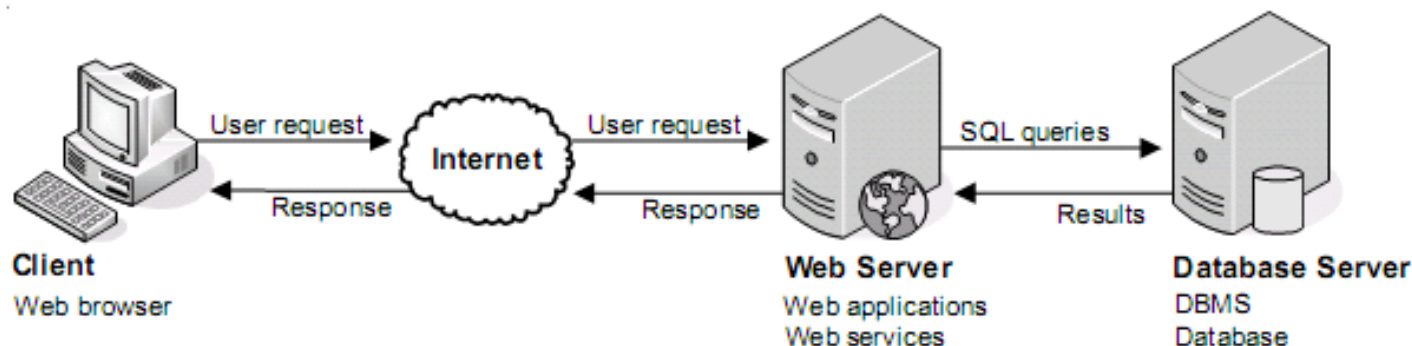


## □ Các kiến trúc khác

### A Windows-based system that uses an application server



### A simple web-based system



- **CSDL** là tập hợp dữ liệu liên quan với nhau được lưu trữ có cấu trúc. Người dùng có thể dễ dàng cập nhật dữ liệu hoặc trích xuất thông tin từ CSDL.
- Ban đầu dữ liệu lưu trữ rời rạc dưới dạng các file, gọi là mô hình dữ liệu file phẳng.
- Sau đó, các **mô hình dữ liệu** khác được thiết kế cho phép mô tả cách thức lưu trữ dữ liệu và cách thức để truy nhập dữ liệu dễ dàng



- **Hệ quản trị CSDL (DBMS)** là tập hợp các chương trình cho phép người dùng lưu trữ, cập nhật và trích xuất thông tin từ CSDL.
- **Hệ quản trị CSDL quan hệ (RDBMS)** là tập hợp các chương trình cho phép tạo và thao tác với CSDL quan hệ.
- Có nhiều đối tượng **người dùng RDBMS** như: quản trị CSDL, thiết kế CSDL, phân tích và thiết kế ứng dụng, cài đặt CSDL, người dùng cuối



**Cảm ơn**