

KHÁI NIỆM CSDL & HỆ QUẢN TRỊ CSDL

GIỚI THIỆU

1. Giới thiệu về dữ liệu, cơ sở dữ liệu
2. Khái niệm cơ bản hệ quản trị CSDL
3. Một số hệ quản trị CSDL phổ biến
4. Mục tiêu của hệ quản trị cơ sở dữ liệu
5. Thuật ngữ RDBMS

DỮ LIỆU LÀ GÌ?

- Dữ liệu là các thông tin của đối tượng (người, vật, một khái niệm, sự việc...) được lưu trữ trên máy tính.
- 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. VD: 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 sinh viên: Tên, mã sinh viên, địa chỉ, ngày sinh, quê quán, lớp.

CƠ SỞ DỮ LIỆU (DATABASE)

Cơ sở dữ liệu:

- Cơ sở dữ liệu (CSDL) là 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.



QUẢN LÝ DỮ LIỆU

Quản lý dữ liệu:

- 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.
- Có 2 phương pháp quản lý dữ liệu:
 - ✓ Hệ thống quản lý bằng file: Dữ liệu được lưu trữ trong các file riêng biệt.
 - ✓ Hệ thống quản lý bằng CSDL.
- Các **mô hình** CSDL:
 - ✓ 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)



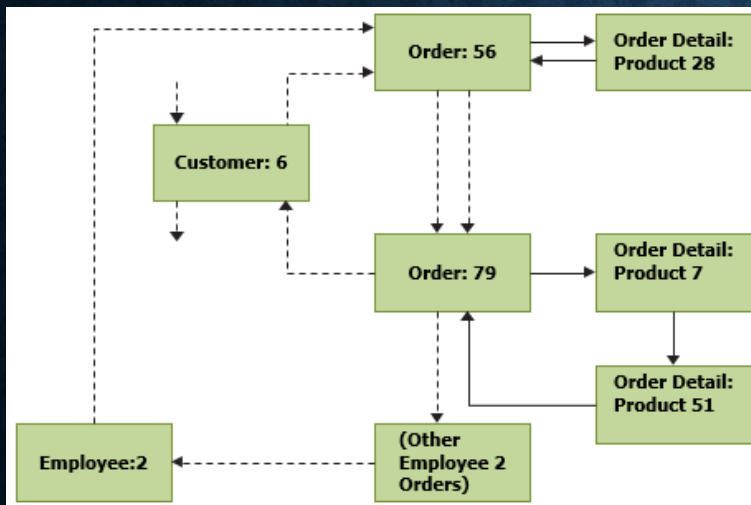
MÔ HÌNH DỮ LIỆU FILE PHẪNG

Mô hình dữ liệu phẳng:

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.

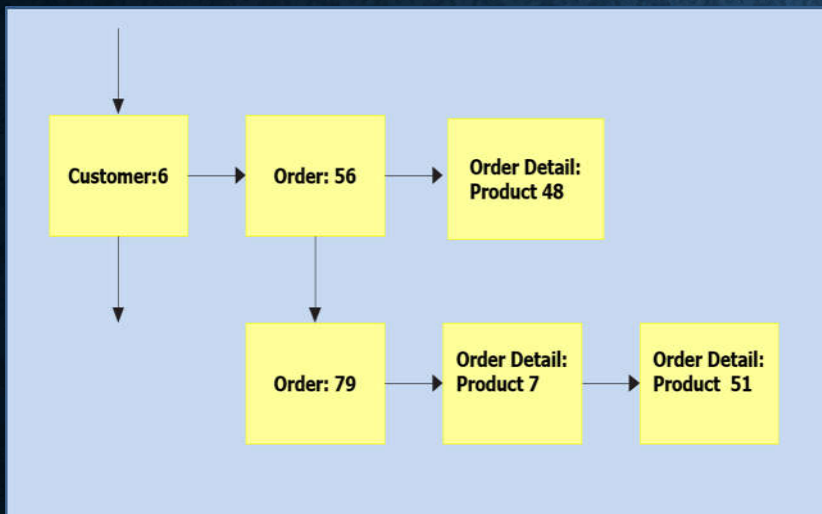
MÔ HÌNH DỮ LIỆU MẠNG



Mô hình dữ liệu mạng:

- 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.
- 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

MÔ HÌNH DỮ LIỆU PHÂN CẤP

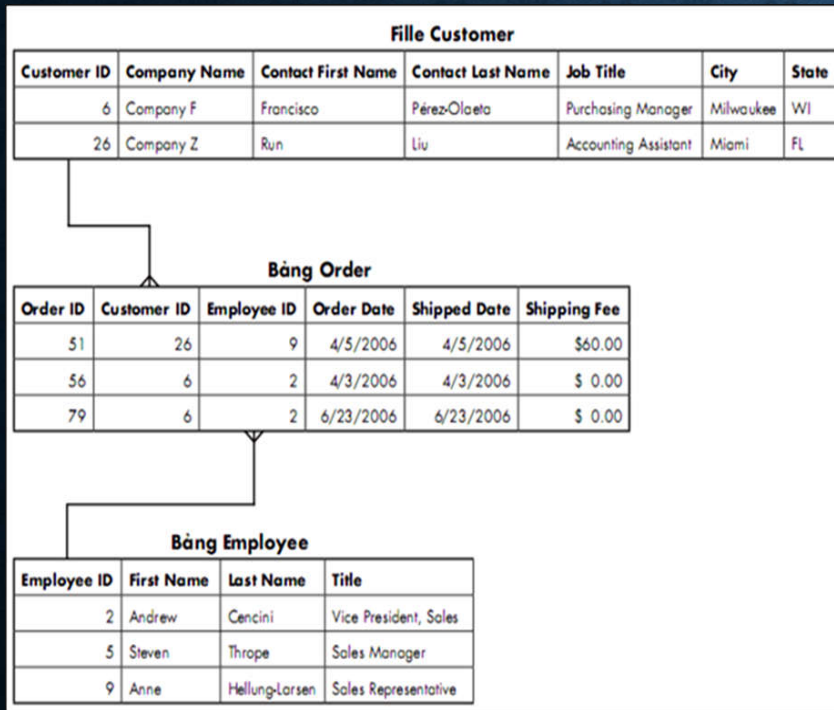


Mô hình dữ liệu phân cấp:

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

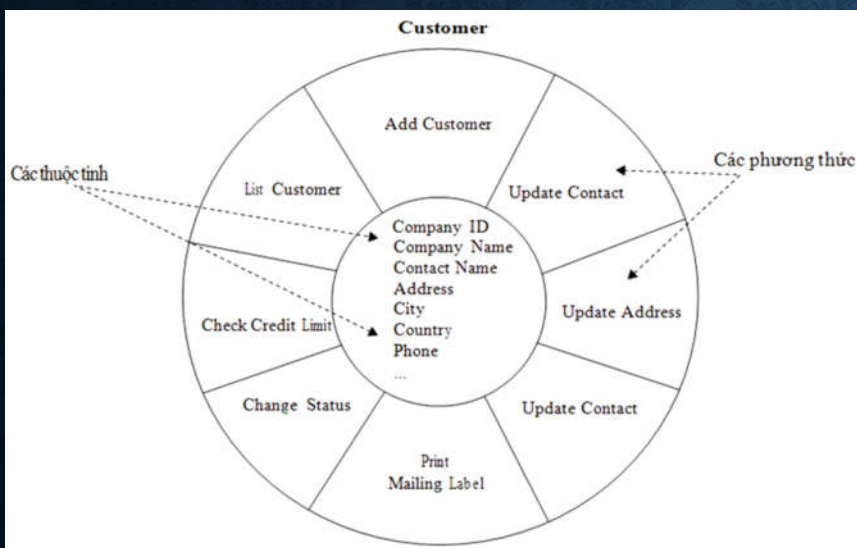
MÔ HÌNH DỮ LIỆU QUAN HỆ

Mô hình dữ liệu quan hệ:



- Trong mô hình dữ liệu quan hệ, không có các liên kết vật lý. 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.
- Có các toán tử để thao tác trên các hàng của bảng.

MÔ HÌNH DỮ LIỆU HƯỚNG ĐỐI TƯỢNG



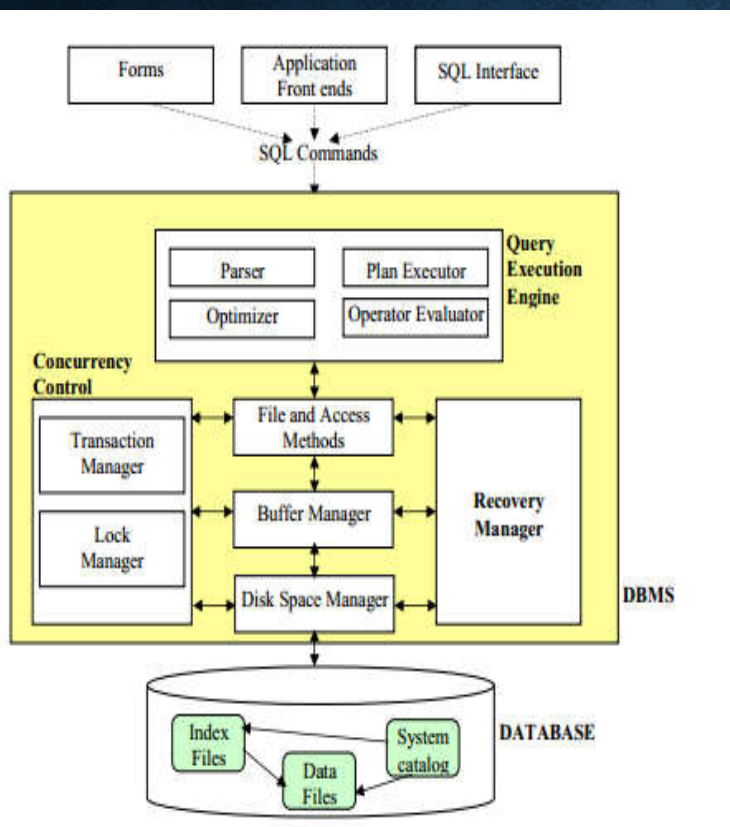
Mô hình dữ liệu hướng đối tượng:

- Ra đời vào khoảng đầu năm 90, dựa trên cách tiếp cận của **phương pháp lập trình hướng đối tượng**.
- CSDL bao gồm các đối tượng:
 - ✓ Mỗi đối tượng bao gồm các thuộc tính, phương thức (hành vi) của đối tượng.
 - ✓ Các đối tượng trao đổi với nhau thông qua các phương thức.
 - ✓ Một đối tượng có thể được sinh ra từ việc thừa kế từ đối tượng khác, nạp chồng (@override) phương thức của đối tượng khác...

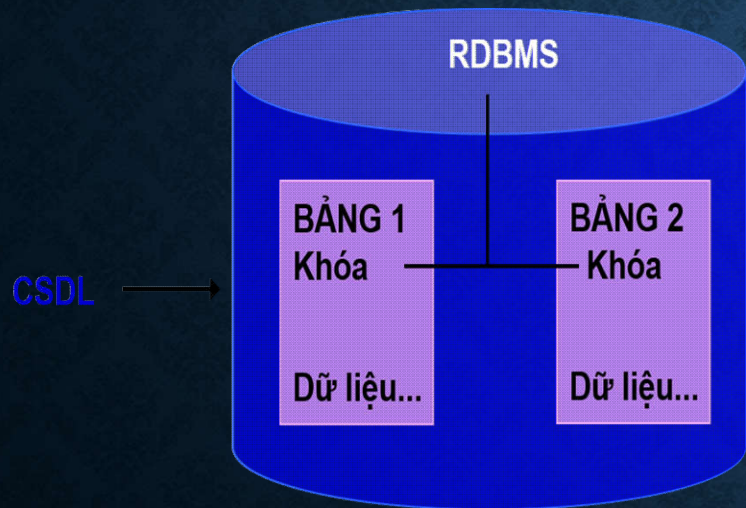
HỆ QUẢN TRỊ CƠ SỞ DỮ LIỆU

Hệ quản trị dữ liệu - DBMS:

- Database Management System – DBMS là phần mềm hay hệ thống được thiết kế để quản trị một cơ sở dữ liệu. Cụ thể, chúng hỗ trợ khả năng lưu trữ, sửa chữa, xóa và tìm kiếm thông tin trong một CSDL.
- Chúng sử dụng ngôn ngữ truy vấn cấu trúc (Structured Query Language - SQL)
- Kiến trúc của HQTCSDL gồm 2 phần chức năng:
 - ✓ Bộ quản lý lưu trữ: Storage manager.
 - ✓ Bộ xử lý truy vấn: Query Processor



TÓM LƯỢC HỆ QUẢN TRỊ CSDL



- Các mô hình CSDL đề cập đến các hình thức tổ chức lưu trữ và truy cập dữ liệu.
- 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

MỘT SỐ HQT CSDL PHỔ BIẾN

- MySQL
- Oracle (11g, 12c)
- **PostgreSQL***
- SQL Server (2005, 2008, 2010,...)
- DB2
- Infomix
- MongoDB (No-SQL)
- ...

TÍNH CHẤT

- **Tính sẵn sàng dữ liệu** (data availability): dữ liệu được tổ chức sao cho mọi người dùng có thể truy cập dễ dàng theo chức năng và nhiệm vụ của họ.
- **Tính toàn vẹn dữ liệu** (data integrity): dữ liệu lưu trữ trong cơ sở dữ liệu là đúng đắn, đáng tin cậy.
- **An toàn dữ liệu** (data security): Chỉ những người dùng được phép mới có thể truy cập dữ liệu. Nếu nhiều người dùng truy cập chung một mục dữ liệu cùng lúc thì hệ quản trị cơ sở dữ liệu không cho phép họ thực hiện những thay đổi gây mâu thuẫn dữ liệu.
- **Độc lập dữ liệu** (data independency): hệ quản trị cơ sở dữ liệu phải cho phép tất cả mọi người dùng được phép lưu trữ, cập nhật và rút trích dữ liệu hiệu quả mà không cần nắm chi tiết về cấu trúc của cơ sở dữ liệu được biểu diễn và cài đặt.

THUẬT NGỮ RDBMS

- **Database:** Một cơ sở dữ liệu là một tập hợp các bảng dữ liệu, với dữ liệu có liên quan.
- **Table:** Một bảng là một ma trận dữ liệu. Một bảng trong một cơ sở dữ liệu trông giống như một bảng tính đơn giản.
- **Column:** Một cột chứa cùng một kiểu dữ liệu, ví dụ như tên khách hàng.
- **Row:** Một hàng (row, entry, record) là một nhóm dữ liệu có liên quan.
- **Primary Key:** Một giá trị key không thể xuất hiện hai lần trong một bảng. Với một key, bạn có thể tìm thấy phần lớn thông tin trên một hàng.
- **Foreign Key:** Khóa ngoại, sử dụng để liên kết dữ liệu các bảng với nhau
- **Compound Key:** Một Compound Key (hay composite key) là một key mà gồm nhiều cột, bởi vì một cột là không duy nhất.
- **Index:** Một chỉ mục trong một cơ sở dữ liệu tương tự như chỉ mục trong một cuốn sách.

QUESTIONS & ANSWERS