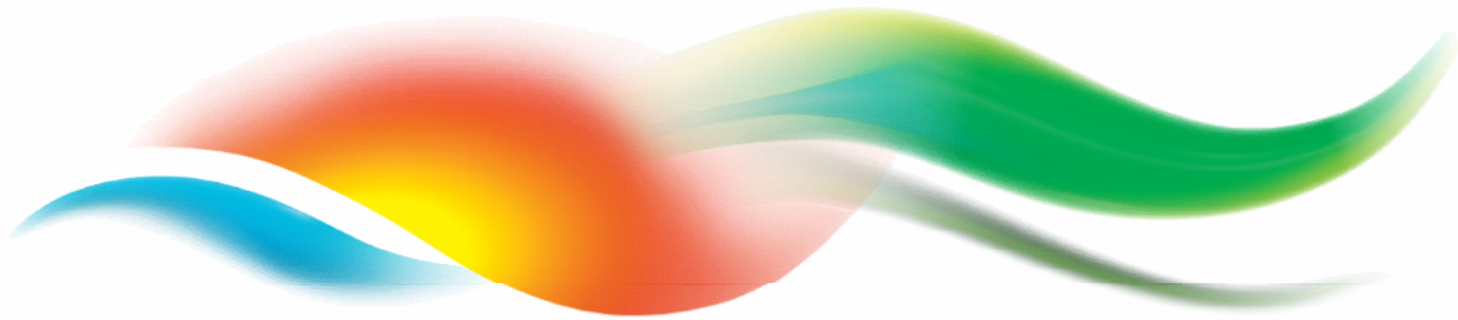




**FPT POLYTECHNIC**



Bài 1:

T NG QUAN V C S D LI U

---

[www.poly.edu.vn](http://www.poly.edu.vn)

[hoclaptrinhweb.com](http://hoclaptrinhweb.com)

# Mục tiêu bài học hôm nay

- Giới thiệu khái niệm dữ liệu và cơ sở dữ liệu (CSDL)
- Các phương pháp quản lý dữ liệu và các cấu trúc
- Giới thiệu các mô hình dữ liệu khác nhau
- Hệ quản trị CSDL (DBMS) và hệ quản trị CSDL quan hệ (RDBMS)

## ■ D ị u (data)

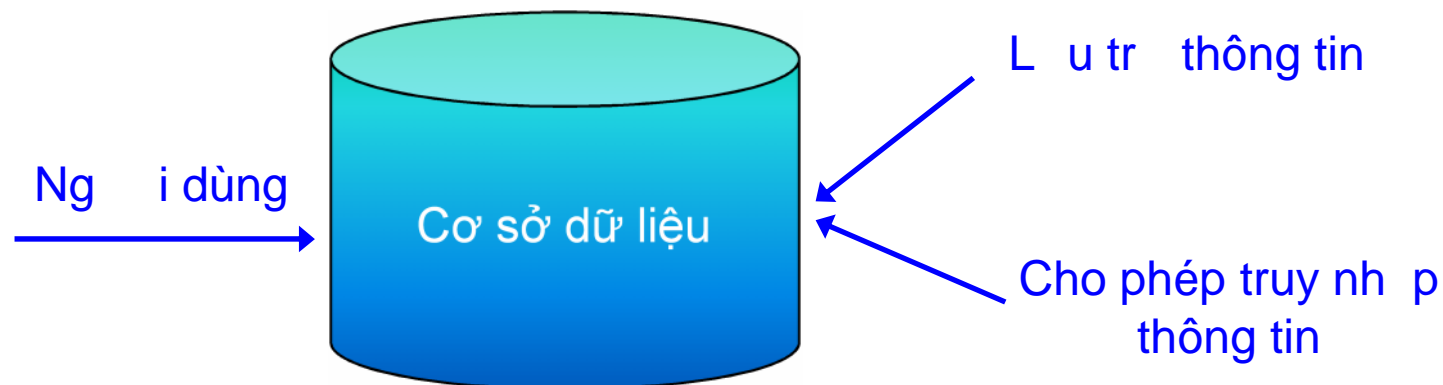
- Là các thông tin c a i t ã ng (ví d : ã g i, v t, m t khái ni m, s vi c...) c l u tr ã trên máy tính.
- Có th truy ã nh p vào d ị u tr í ch xu t ra các thông tin.

## ■ D ị u c mô t d i ã nh i u d ã ng khác nhau (các ký t , ký s , hình ã nh, ký hi u, âm thanh...). M i cách mô t ã nh v y g ã n v i m t ã ng ã ngh a ã ã o ó.

- D li u v i t ng có th khác nhau, tùy thu c vào ng c nh.
- Ví d : d li u v i t ng sinh viên có th khác nhau tùy vào m c ích qu n lý:
  - Qu n lý i m: Tên, mã sinh viên, i m môn 1, i m môn 2, i m môn 3
  - Qu n lý nhân thân: Tên, a ch , ngày sinh, quê quán, l p

# C s d li u (CSDL)

- CSDL (Database) = T p h p d li u c t ch c có c u trúc liên quan v i nhau và c l u tr trong máy tính.
- CSDL c thi t k , xây d ng cho phép ng i dùng l u tr d li u, truy xu t thông tin ho c c p nh t d li u



## ■ CSDL c t ch c có c u trúc:

- Các d li u l u tr có c u trúc thành các b n ghi (record), các tr ng d li u (field).
- Các d li u l u tr có m i quan h (relational) v i nhau.

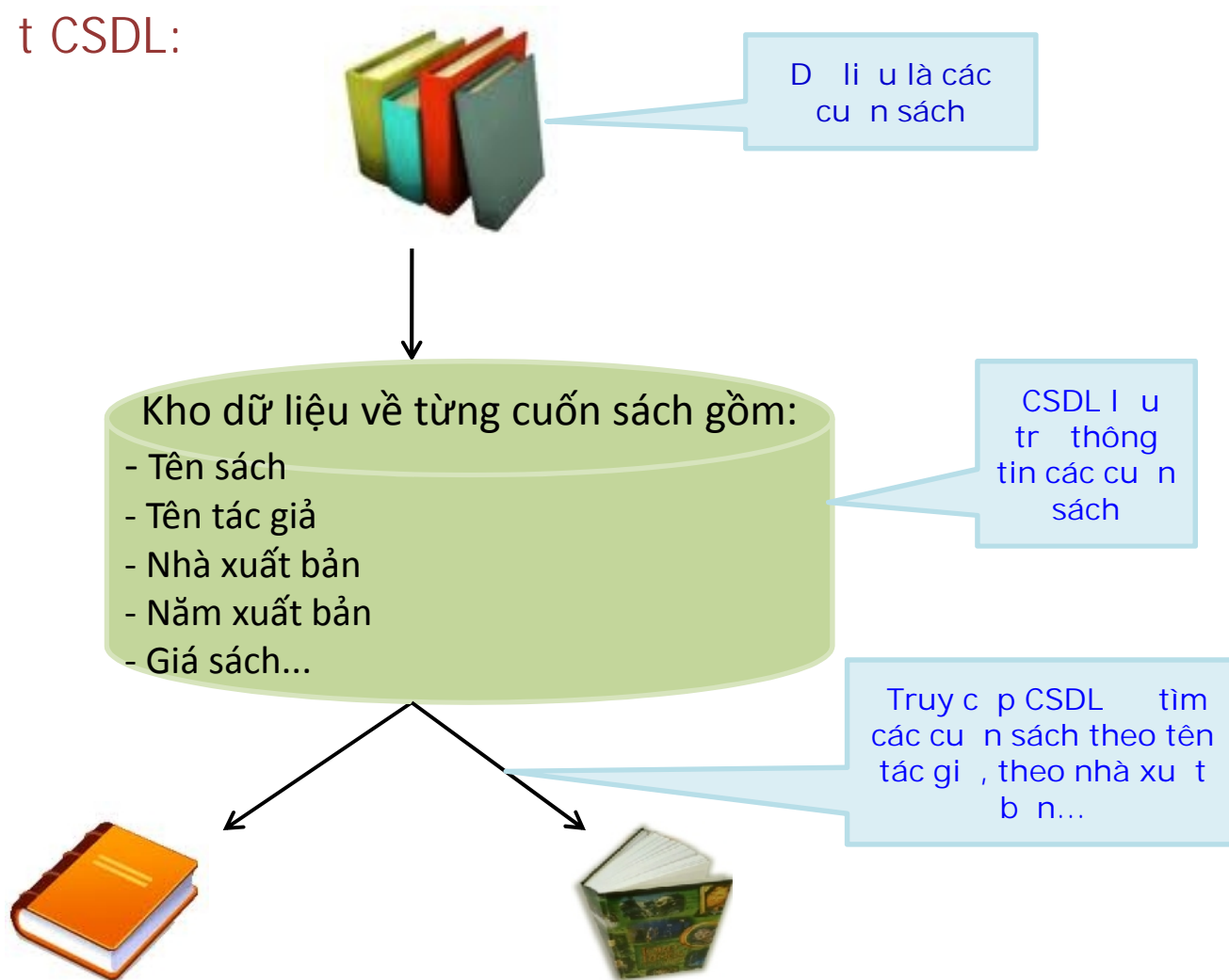
## ■ Kh n ng truy xu t thông tin t CSDL:

- CSDL c c u trúc d dàng truy c p, qu n lý và c p nh t d li u

➔ C n ph i qu n tr CSDL

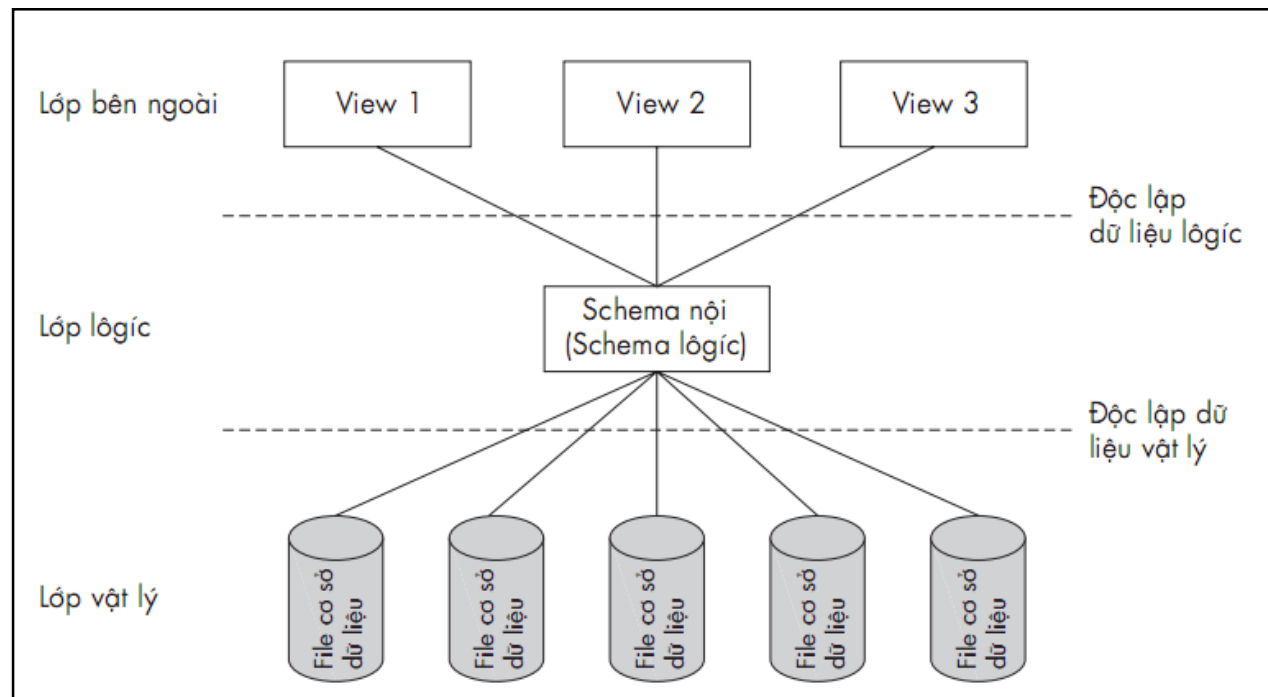
# Dữ liệu và Cơ sở dữ liệu

## ■ Ví dụ minh họa CSDL:



# Các tính trong CSDL

- CSDL cung cấp khả năng trừu tượng hóa dữ liệu thông qua các lớp. Bao gồm 3 lớp: Lớp vật lý, Lớp logic, Lớp bên ngoài.
- Sự phân biệt giữa các lớp tạo nên 2 tính chất: tính độc lập dữ liệu vật lý và tính độc lập dữ liệu logic.





# Các tính trong CSDL

## ■ Lập v t lý:

- Lập v t lý ch a toàn b các file d li u
- Ng i dùng CSDL không nh t thi t ph i n m c c u trúc t ch c c a các file d li u v t lý khi s d ng m t c s d li u.

# Các tính trong CSDL

- Logic (còn gọi là Schema):
  - Cấu trúc dữ liệu trừu tượng tạo thành tập hợp lý.
  - Logic có thể chia thành hai loại: các bảng hai chiều, một chiều phân cấp theo tính chất của công ty hay một vài cấu trúc khác

# Các tính trong CSDL

## ■ Lập bên ngoài:

- Bao gồm các khung nhìn (view) và các g i chung là subschema
- Khung nhìn: là khái niệm cho phép người dùng quan sát dữ liệu theo những cách khác nhau trong khi dữ liệu lưu bên dưới đang vật lý là duy nhất.

# Các tính trong CSDL

## ■ clip dữ liệu vật lý:

- Khả năng thay đổi cấu trúc file vật lý của một cơ sở dữ liệu mà không làm gián đoạn việc dùng đang truy cập vào các quá trình đang diễn ra cũng là clip dữ liệu vật lý
- Việc phân biệt clip vật lý với clip lô-gic tỏ ra sự clip dữ liệu vật lý

# Các tính trong CSDL

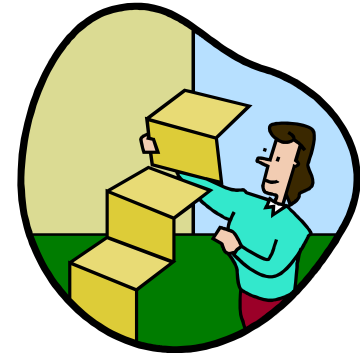
## ■ Các lập luận logic:

- Khả năng tạo ra các thay thế lập luận logic mà không làm gián đoạn dòng dữ liệu và các quá trình đang diễn ra cũng là các lập luận logic.
- Việc phân biệt giữa lập luận logic và lập luận bên ngoài tạo ra tính lập luận logic.

- Quản lý dữ liệu: là quản lý m t s l ng l n d li u, bao g m c vi c l u tr và cung c p c ch cho phép Thao tác (thêm, s a, xóa d li u) và Truy v n d li u.
- 2 ph ng pháp quản lý dữ liệu:
  - H th ng quản lý b ng file
  - H th ng quản lý b ng CSDL

# Quản lý dữ liệu lưu trữ file

- Dữ liệu được lưu trữ trong các file riêng biệt
- Ví dụ: các chương trình lưu trữ thông tin bằng hình ảnh, âm thanh, văn bản, các file dữ liệu text
- Những thách thức của việc quản lý lưu trữ file:
  - Đa dạng và mâu thuẫn dữ liệu
  - Khó khăn trong truy xuất thông tin, hoặc xử lý thông tin
  - Dữ liệu lưu trữ rườm rà
  - Giảm thiểu rủi ro an toàn và bảo mật

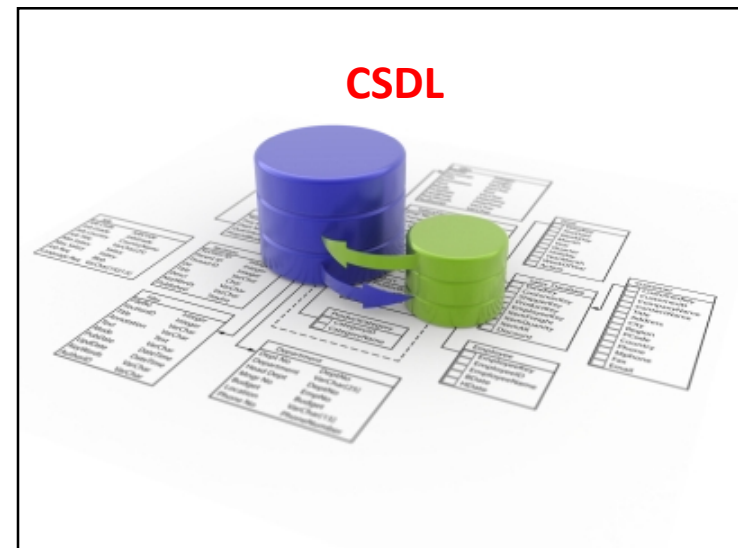


# Quản lý dữ liệu bằng CSDL

- Quản lý dữ liệu bằng CSDL giúp dữ liệu có 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 dữ liệu 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





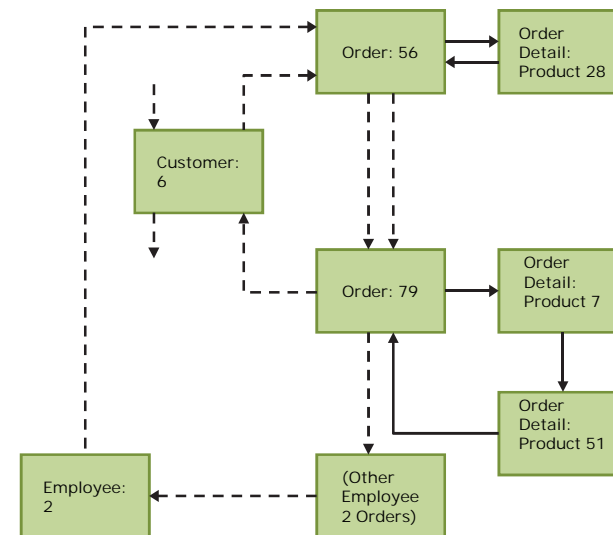
- Các CSDL có thể khác nhau về chức năng và mô hình dữ liệu (data model).
- Mô hình dữ liệu sẽ quy định cách thức lưu trữ và truy cập dữ liệu.
- Tùy từng ngôn ngữ quan hệ giữa các thành phần dữ liệu trong CSDL, mô hình phân cấp áp dụng vì cách lưu trữ và truy xuất dữ liệu tối ưu cao nhất.

# Các mô hình CSDL

## ■ Các mô hình:

- Mô hình dữ liệu phân cấp (Hierarchical model)
- Mô hình dữ liệu quan hệ (Relational model)
- Mô hình dữ liệu mạng (Network 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 dữ liệu file phông

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

# Mô hình dữ liệu file phông

- Ví dụ : mô tả file phông thể hiện thông tin mô tả Customer (Khách hàng) đã đi đăng ký công ty Northwind Traders chuyên cung cấp các mô tả hàng thực phẩm

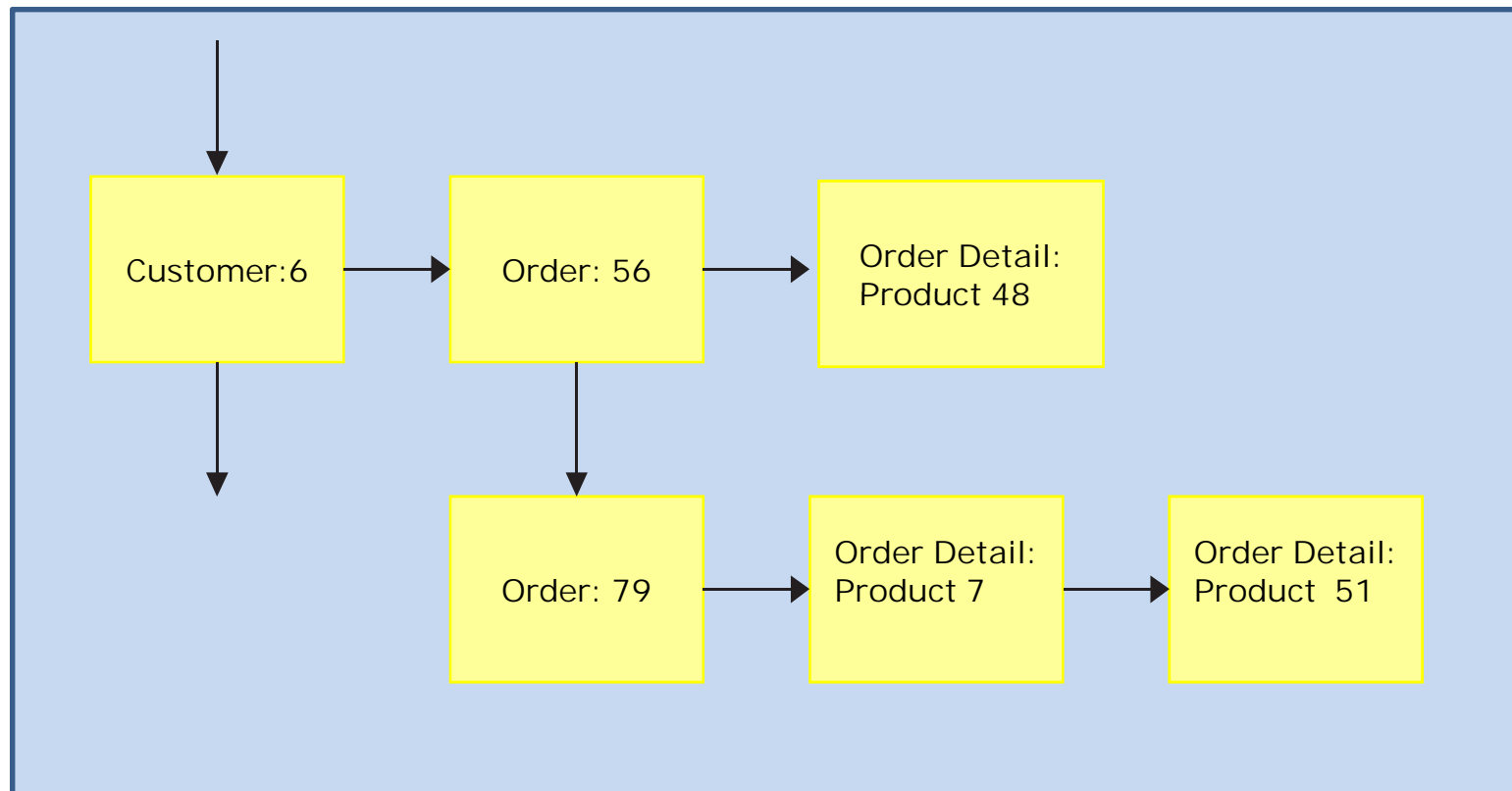
Customer ID	Company Name	Contact First Name	Contact Last Name	Job Title	City	State
6	Company F	Francisco	Pérez-Olaeta	Purchasing Manager	Milwaukee	WI
26	Company Z	Run	Liu	Accounting Assistant	Miami	FL

# Mô hình 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

# Mô hình dữ liệu phân cấp

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



# Mô hình dữ liệu phân cấp

## ■ Hình thức :

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

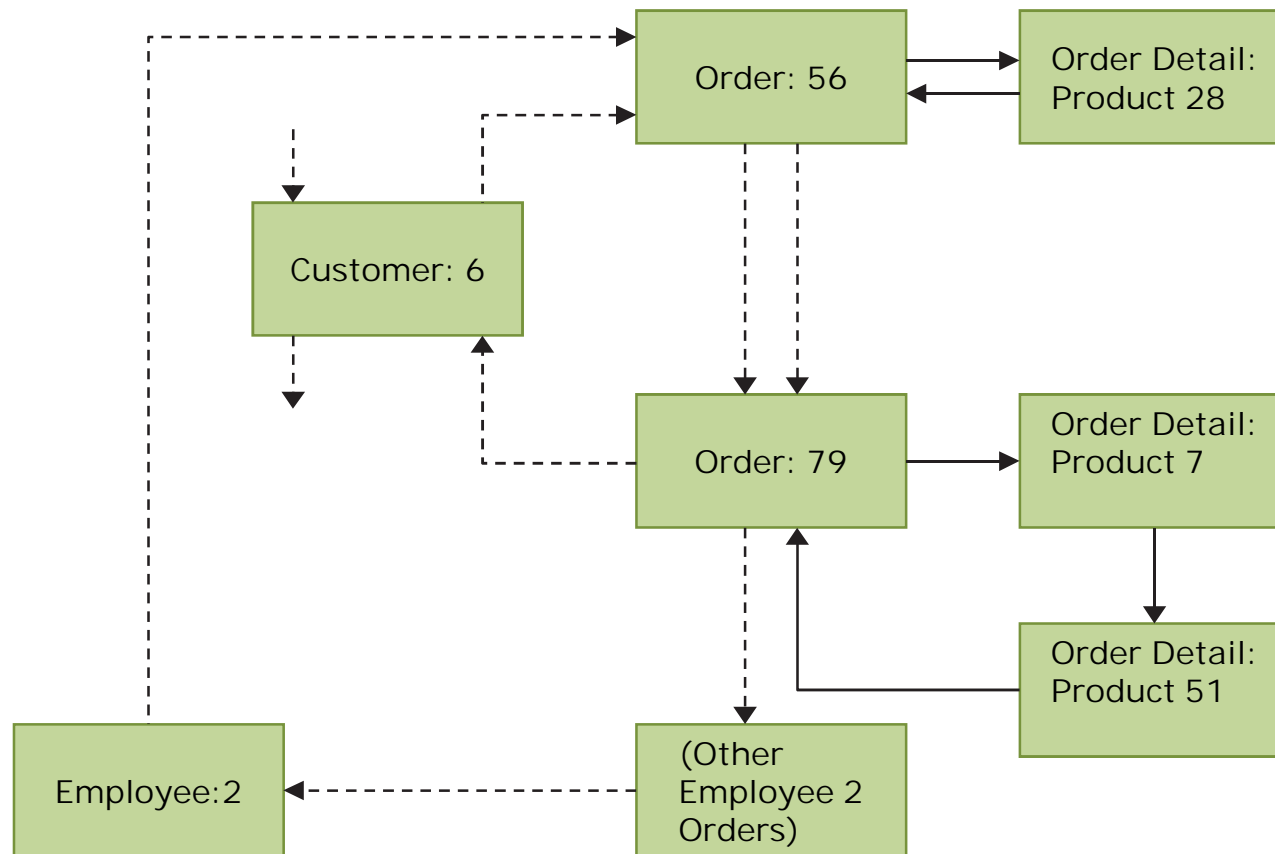
## ■ Cách tổ chức:

- Các file riêng biệt trong hệ thống file phẳng có nghĩa 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 bất kỳ một thực thể có hướng, và các mối tên chính thực thể cha sang kiểu thực thể con.



# Mô hình dữ liệu m

- Ví dụ : Cấu trúc mô hình m của Northwind



## ■ Tính chất:

### ● Ưu điểm:

- Dễ biểu diễn mô hình
- Dễ nhận thấy các liên hệ dữ liệu phức tạp

### ● Nhược điểm:

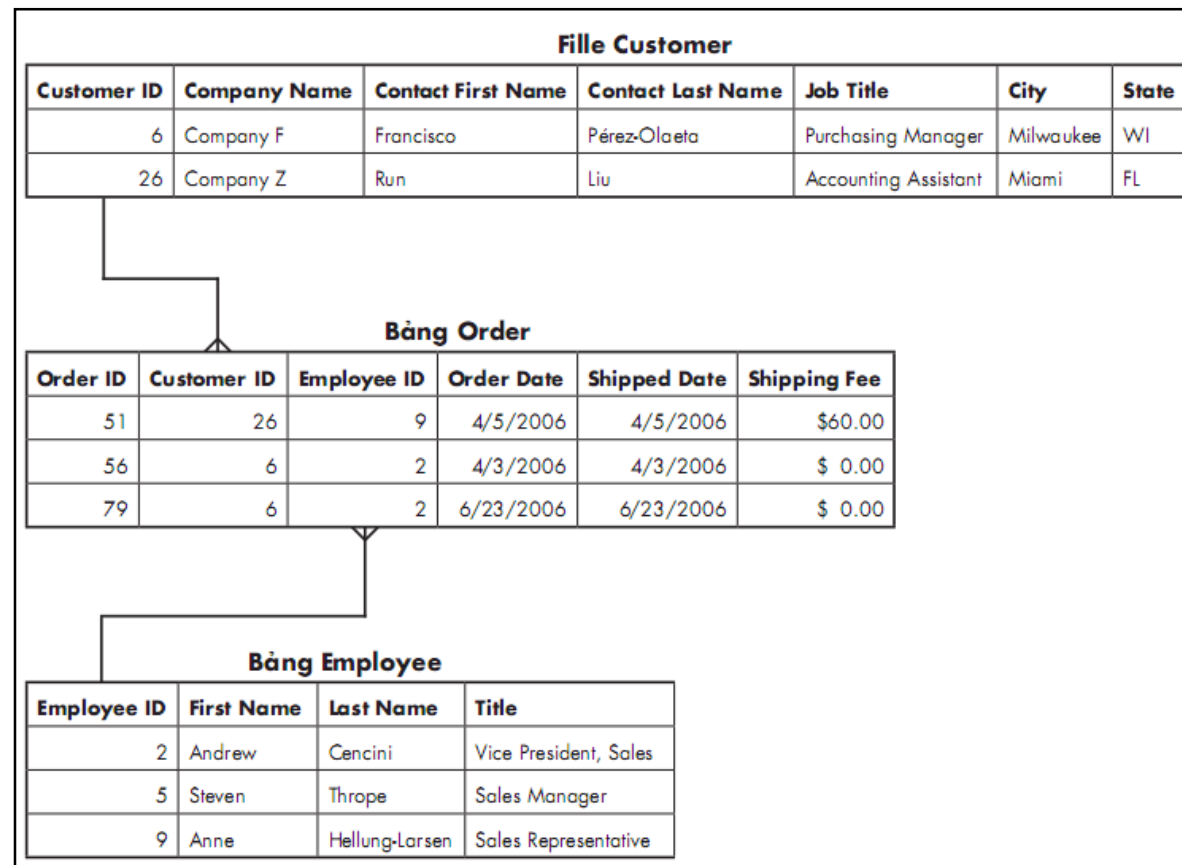
- Truy xuất chậm
- Không thích hợp với các CSDL có quy mô lớn.

# Mô hình dữ liệu 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 gọi là bộ (tuple)
  - Mỗi cột là một thuộc tính, còn gọi là trường (field)
- Dữ liệu trong hai bảng liên hệ với nhau thông qua các cột chung.
- Có các toán tử thao tác trên các hàng của bảng.

# Mô hình dữ liệu quan hệ

- Ví dụ : mô hình dữ liệu quan hệ trong CSDL Northwind gồm 3 bảng:  
Customer, Order, Employee

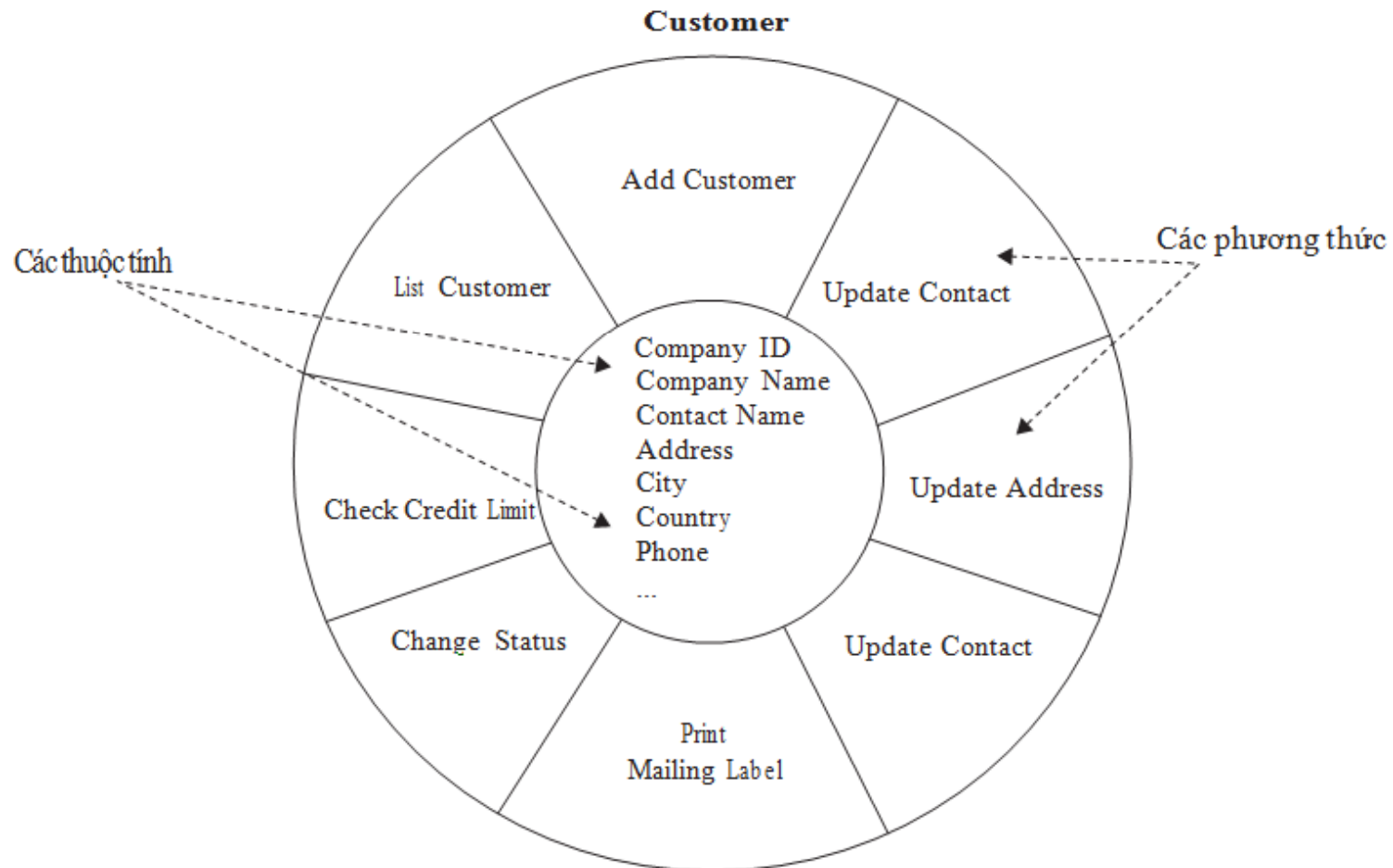


## Mô hình dữ liệu hướng đối tượng

- Ra vào kho hàng luôn luôn 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 sinh ra một vài đối tượng khác, nên phép (hay nhúng) phương thức của đối tượng khác...

# Mô hình dữ liệu hướng đối tượng

- Ví dụ mô hình dữ liệu hướng đối tượng "Customer"



- 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 tổ chức các CSDL và cung cấp cách lưu trữ, truy cập theo các mô hình CSDL.
- Ví dụ :
  - SQL Server, Microsoft Access, Oracle là các hệ quản trị CSDL hiện hành 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:

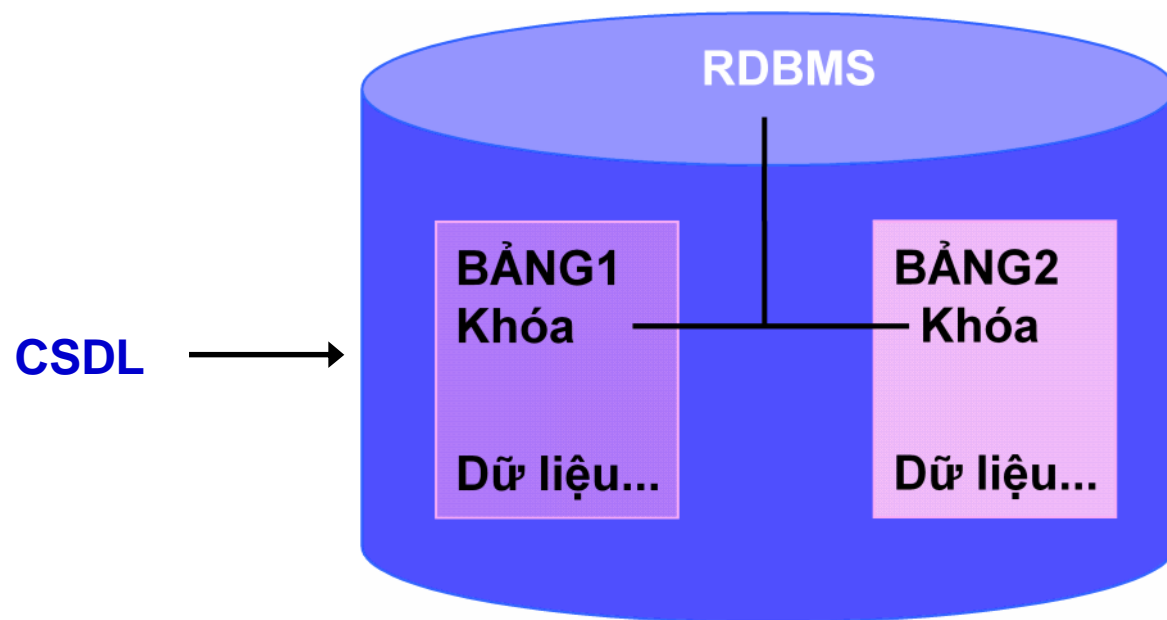
- Quản trị các CSDL
- Cung cấp giao diện truy cập để điều khiển các tính phức tạp về mặt cấu trúc tổ chức dữ liệu vật lý
- Hỗ trợ các ngôn ngữ giao tiếp. Ví dụ:
  - Ngôn ngữ mô tả, định nghĩa dữ liệu – DDL
  - Ngôn ngữ thao tác dữ liệu – DML
  - Ngôn ngữ truy vấn dữ liệu có cấu trúc – SQL
- Có các chức năng bảo mật cao



# Hệ quản trị CSDL quan hệ

- Hệ quản trị CSDL quan hệ (Relational DataBase Management System = RDBMS)
- RDBMS là một dạng DBMS có sử dụng phương tiện nhúng, trong đó tổ chức dữ liệu có thể thực hiện được các **biến đổi** dữ liệu.
- Thực hiện các thao tác trên CSDL được diễn ra trên các bảng.

# Hệ quản trị CSDL quan hệ



# Người dùng liên quan đến RDBMS

- R t n h i u n g i d ù n g t h a m g i a v à o h t h n g R D B M S n h :
  - N g i q u n t r C S D L (DataBase Administrator)
  - N g i t h i t k C S D L (DataBase Designer)
  - N g i p h â n t í c h h t h n g (System Analysts)
  - N g i l p t r ì n h n g d n g (Application Programmers)
  - N g i t h i t k v à t r i n k h a i C S D L (DBMS Designers and Implementers)
  - N g i d ù n g c u i (End User)



## T i sao l i t p trung vào CSDL quan h ?

- D ể dàng nh nh a, duy trì và thao tác d ữ li u l u tr .
- Trích xu t d ữ li u d ể dàng
- D ữ li u c chu n hóa và c b o v t t
- Nhi u nhà cung c p cung c p ph n m m
- D ể dàng chuy n i gi a nhà cung c p và nhà tri n khai
- RDBMS là các s n ph m tr ờ ng thành và n nh

- CSDL là tập hợp dữ liệu liên quan với nhau có cấu trúc và dạng cụ thể để trích xuất thông tin từ CSDL.
- Thành phần CSDL tạo ra các loại truy vấn CSDL: Truy vấn vật lý, Truy vấn logic và Truy vấn bên ngoài.
- Ban đầu dữ liệu lưu trữ rời rạc dưới dạng các file, gọi là mô hình dữ liệu file phân tán.
- Sau đó, các mô hình dữ liệu khác nhau được thiết kế cho phép mô tả cách thức lưu trữ dữ liệu và cách thức truy cập dữ liệu dạng

# Tổng kết bài học

- Hệ quản trị CSDL (DBMS) là tập 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 kỹ thuật ứng dụng RDBMS như: quản trị CSDL, thiết kế CSDL, phân tích và thiết kế ứng dụng, cài đặt CSDL, người dùng cuối.