



Chương 1 Tổng quan về CSDL

TS. Võ Đức Hoàng

Khoa Công nghệ thông tin

Trường Đại học Bách khoa - Đại học Đà Nẵng

Nội dung chi tiết



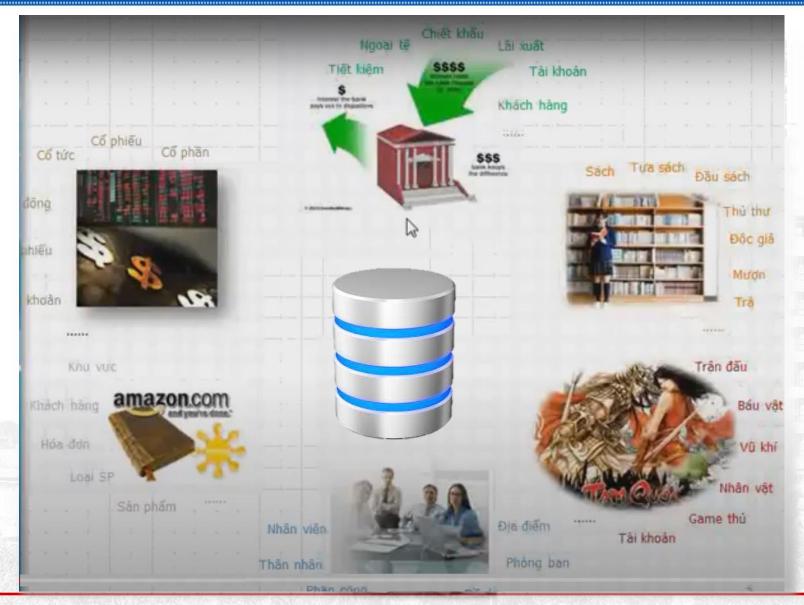


- Giới thiệu
- Quá trình phát triển
- Một số đặc tính của CSDL
- Người sử dụng CSDL
- Kiến trúc của HQT CSDL
- Các tính năng của HQT CSDL
- Các khái niệm
- Ngôn ngữ CSDL

Ví dụ





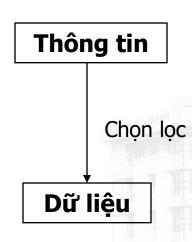


Giới thiệu





- Ví dụ
 - Kinh doanh
 - Ngân hàng và tài chính
 - Giáo dục
 - Hành chính
 - Giải trí
 - ...
- Dữ liệu (Data)
 - Một mô tả hình thức về thông tin và hoạt động
 - Tên, địa chỉ, số điện thoại của khách hàng
 - Báo cáo doanh thu
 - Đăng ký học phần









CƠ SỞ DỮ LIỆU





Programs = Algorithms + Data Structure

Phát triển Các thuật toán

Vậy trước đó, tổ chức dữ liệu bằng gì?





ų: ứng dụng

- Quản lý nhân sự
- Quản lý hàng không

Phát triển Các cấu trúc lưu trữ

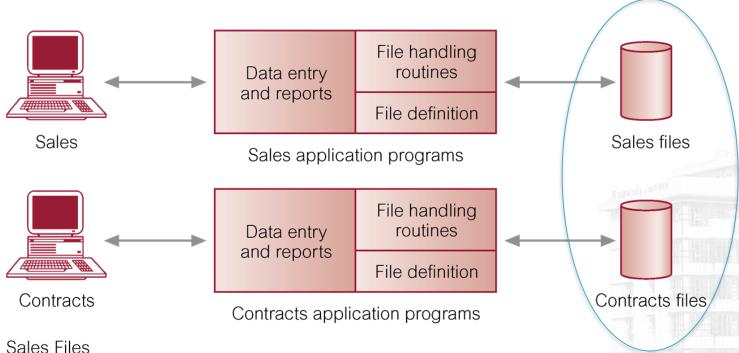


Phù hợp cho các ứng dụng

- Có khối lượng dữ liệu lớn
- Không có tính toán phức tạp

Xử lý dữ liệu trên hệ thống tập tiến





PropertyForRent (propertyNo, street, city, postcode, type, rooms, rent, ownerNo)

PrivateOwner (ownerNo, fName, IName, address, telNo)

Client (clientNo, fName, IName, address, telNo, prefType, maxRent)

Contracts Files

Lease (leaseNo, propertyNo, clientNo, rent, paymentMethod, deposit, paid, rentStart, rentFinish, duration)

PropertyForRent (propertyNo, street, city, postcode, rent)

Client (clientNo, fName, IName, address, telNo)

Hạn chế của Hệ thống tập tin





- Dữ liệu tách biệt và không chia sẻ
 - Mỗi ứng dụng sử dụng một hệ thống tập tin riêng
 - Dữ liệu không thể chia sẻ giữa các ứng dụng

- Sự trùng lắp
 - Dữ liệu bị trùng lắp giữa các ứng dụng
 - Tốn không gian lưu trữ
 - Có thể không nhất quán

Hạn chế của Hệ thống tập tin



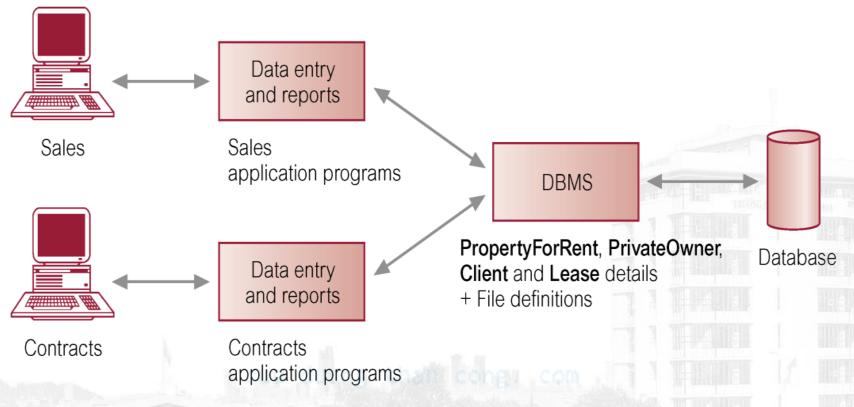


- Phụ thuộc dữ liệu
 - Các cấu trúc tập tin được định nghĩa cố định trong chương trình
- Định dạng tập tin không tương thích
 - Các tập tin được định dạng theo một ngôn ngữ lập trình nào đó -> khó chia sẻ
- Câu truy vấn bị cố định trước
 - Mỗi chương trình truy xuất được phát triển cố định cho một câu hỏi

Cách tiếp cận CSDL







PropertyForRent (propertyNo, street, city, postcode, type, rooms, rent, ownerNo)

PrivateOwner (ownerNo, fName, IName, address, telNo)

Client (clientNo, fName, IName, address, telNo, prefType, maxRent)

Lease (leaseNo, propertyNo, clientNo, paymentMethod, deposit, paid, rentStart, rentFinish)

Đặc điểm của Dữ liệu lưu trên CSDL





- Dữ liệu lưu trữ độc lập với chương trình
- Dữ liệu được quản lý tập trung
- Dữ liệu được chia sẻ cho nhiều ứng dụng
- Dữ liệu được bảo đảm an toàn
- Dữ liệu ít dư thừa

Định nghĩa Cơ sở dữ liệu



Cơ sở dữ liệu (Database) là một tập hợp các dữ liệu được tổ chức theo một cấu trúc chặt chẽ, nhằm phục vụ (chia sẻ) cho nhiều mục tiêu khác nhau một cách có chọn lọc.

Ví dụ :

- Cơ sở dữ liệu nhân viên,
- cơ sở dữ liệu hàng hóa ...
- Một CSDL biểu diễn một phần của thế giới thực (thế giới thu nhỏ)
- CSDL được thiết kế, xây dựng, và lưu trữ với một mục đích xác định, phục vụ cho một số ứng dụng và người dùng
- Tập ngẫu nhiên của các dữ liệu không thể xem là một CSDL

Định nghĩa Cơ sở dữ liệu





NHA	NVIEN	HONV	TENLOT	TENNV	MANV
		Tran	Hong	Quang	NHQuang
		Tran	Thanh	Tam	TTTam

Tập ngẫu nhiên của các dữ liệu không thể xem là một CSDL

DEAN	TENDA	MADA	DDIEM_DA
	IBM Leader	1	VUNG TAU
	Photoshop	2	NHA TRANG
	Management	3	TP HCM

PHANC	ONG	MA_NVIEN	SODA	THOIGIAN
		NHQuang	1	32.5
		NHQuang	2	7.5
		TTTam	3	40.0

Ví dụ về một CSDL





NHANVIEN	HONV	TENLOT	TENNV	MANV	NGSINH	MA_NQL	PHG
	Tran	Hong	Quang	987987987	03/09/1969	987654321	4
	Nguyen	Thanh	Tung	333445555	12/08/1955	888665555	5
	Nguyen	Manh	Hung	666884444	09/15/1962	333445555	5
	Tran	Thanh	Tam	453453453	07/31/1972	333445555	5

DEAN	TENDA	MADA	DDIEM_DA	PHONG
	San pham X	1	VUNG TAU	5
	San pham Y	2	NHA TRANG	5
	San pham Z	3	TP HCM	5
	Tin hoc hoa	10	HA NOI	4

CSDL

Quản lý đề án công ty

PHANCONG	MA_NVIEN	SODA	THOIGIAN
	123456789	1	32.5
	123456789	2	7.5
	666884444	3	40.0
	453453453	1	20.0

Ví dụ về một CSDL





	DEAN	TENDA	MADA	DDIEM_DA
100		IBM Leader	1	VUNG TAU
		Photoshop	2	NHA TRANG
		Management	3	TP HCM

THOIGIAN 32.5 7.5

40.0

15

- Quản lý đề án của một công ty
 - Định nghĩa CSDL
 - Cấu trúc bảng, bao gồm các thành phần dữ liệu và kiểu dữ liệu tương ứng
 - Xây dựng CSDL
 - Đưa dữ liệu vào các bảng
 - Xử lý CSDL
 - Thực hiện truy vấn: "Cho biết các nhân viên thuộc phòng 5"
 - Thực hiện phép cập nhật: "Chuyến nhân viên Nguyễn Thanh Tùng sang phòng số 1"

Định nghĩa Hệ quản trị cơ sở dữ liệu





- HQT CSDL (DataBase Management System)
 - Hệ thống phần mềm, cung cấp công cụ:
 - Xây dựng CSDL
 - Thao tác trên dữ liệu của CSDL
 - Kiểm soát việc truy xuất dữ liệu
 - Cho phép định nghĩa, xây dựng và xử lý dữ liệu
 - Định nghĩa khai báo bộ khung, các mô tả chi tiết về dữ liệu
 - Xây dựng lưu trữ dữ liệu lên bộ nhớ phụ
 - Xử lý truy vấn, cập nhật và phát sinh báo cáo

Định nghĩa Hệ quản trị cơ sở dữ liệu





Một số Hệ quản trị CSDL phổ biến

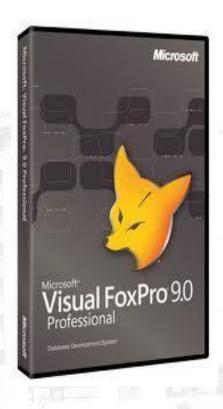








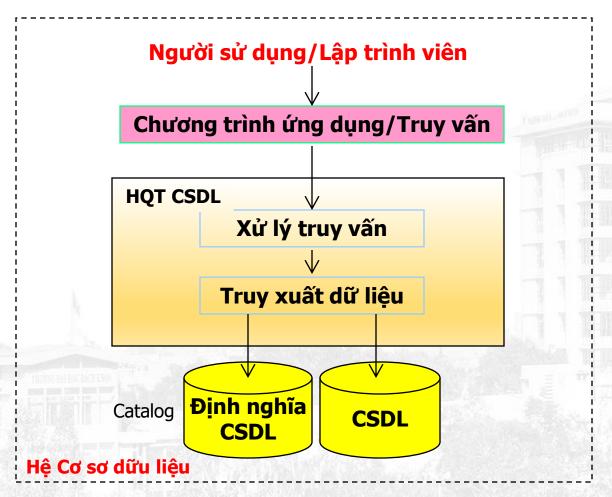




Định nghĩa – Hệ Cơ sở dữ liệu

IT FACULTY

Hệ cơ sở dữ liệu (Database System)



Nội dung chi tiết





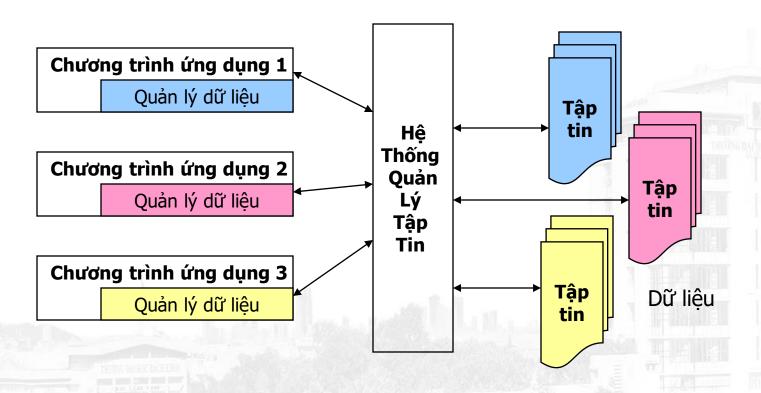
- Giới thiệu
- Quá trình phát triển
- Một số đặc tính của CSDL
- Người sử dụng CSDL
- Kiến trúc của HQT CSDL
- Các tính năng của HQT CSDL
- Các khái niệm
- Ngôn ngữ CSDL

Quá trình phát triển





Tập tin (File)



Quá trình phát triển (tt)





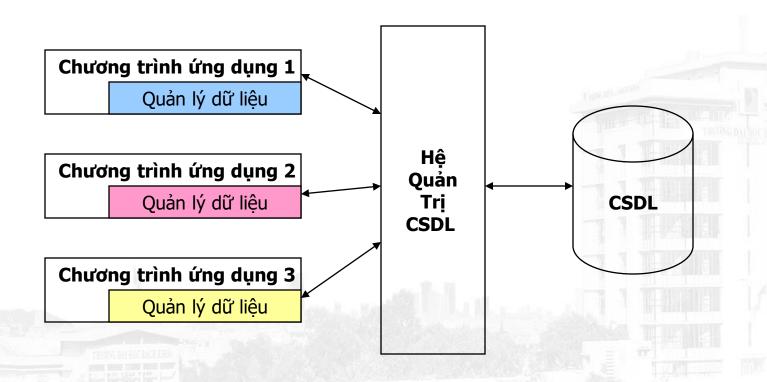
- Hạn chế
 - Dữ liệu bị trùng lắp và dư thừa
 - Thiếu tính nhất quán giữa các dữ liệu
 - Khó khăn trong việc truy xuất
 - Việc chia sẻ dữ liệu bị hạn chế
 - Khó khôi phục

Quá trình phát triển (tt)





Cơ sở dữ liệu (Database)



Nội dung chi tiết





- Giới thiệu
- Quá trình phát triển
- Một số đặc tính của CSDL
 - Tính tự mô tả
 - Tính độc lập giữa chương trình và dữ liệu
 - Tính trừu tượng dữ liệu
 - Tính nhất quán
 - Tính đa khung nhìn
- Người sử dụng CSDL
- Kiến trúc của HQT CSDL
- Các tính năng của HQT CSDL
- Các khái niệm
- Ngôn ngữ CSDL

Tính tự mô tả





- Hệ CSDL không chỉ chứa bản thân CSDL mà còn chứa định nghĩa đầy đủ (mô tả: cấu trúc + ràng buộc) của CSDL
- Thông tin mô tả được lưu trong <u>Catalog</u>
 - Chứa các thông tin về cấu trúc tập tin, kiểu và dạng thức lưu trữ của mỗi thành phần dữ liệu và những ràng buộc dữ liệu
- Dữ liệu trong catalog gọi là meta-data (data of data)
- Thông tin lưu trong catalog được sử dụng bởi :
 - Hệ quản trị CSDL
 - Người lập trình CSDL
- Các CTƯD có thể truy xuất đến nhiều CSDL nhờ thông tin cấu trúc được lưu trữ trong catalog

Tính độc lập





 Vì định nghĩa về cấu trúc CSDL được lưu trữ trong catalog nên khi có thay đổi nhỏ về cấu trúc ta ít phải sửa lại chương trình

Chương trình
Độc lập

Dữ liệu

Tính trừu tượng





- Hệ CSDL cho phép trình bày dữ liệu ở một mức trừu tượng cho phép, nhằm che bớt những chi tiết lưu trữ thật của dữ liệu
- Trừu tượng hóa dữ liệu
 - Mô hình dữ liệu
 - Đối tượng
 - Thuộc tính của đối tượng
 - Mối liên hệ

Tính nhất quán





- Lưu trữ dữ liệu thống nhất
 - Tránh được tình trạng trùng lắp thông tin
- Có cơ chế điều khiển truy xuất dữ liệu hợp lý
 - Tránh được việc tranh chấp dữ liệu
 - Bảo đảm dữ liệu luôn đúng tại mọi thời điểm
- Ví dụ đặt vé máy báy:
 - Cùng 1 chuyến bay nhưng có thể đặt vé và ghế ngồi ở nhiều đại lý khác nhau nhưng không bao giờ trùng vị trí.

Đa khung nhìn





- Hệ CSDL cho phép nhiều người dùng thao tác lên cùng một CSDL
- Mỗi người đòi hỏi một cách nhìn (view) khác nhau về CSDL
- Một view là
 - Một phần của CSDL

hoặc

- Dữ liệu tổng hợp từ CSDL

Nội dung chi tiết



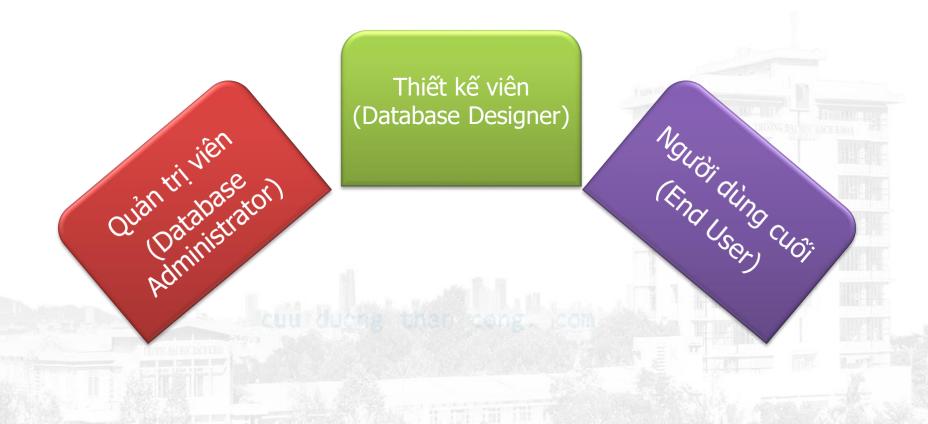


- Giới thiệu
- Quá trình phát triển
- Một số đặc tính của CSDL
- Người sử dụng CSDL
 - Quản trị viên (Database Administrator DBA)
 - Thiết kế viên (Database Designer)
 - Người dùng cuối (End User)
- Kiến trúc của HQT CSDL
- Các tính năng của HQT CSDL
- Các khái niệm
- Ngôn ngữ CSDL

Người sử dụng CSDL







Quản trị viên





DBA – Database Administrator

- Có trách nhiệm quản lý hệ CSDL
 - Cấp quyền truy cập
 - Điều phối việc sử dụng
 - Giám sát việc sử dụng
 - Tìm kiếm & đảm bảo phần cứng, phần mềm liên quan
- Chịu trách nhiệm cho các vấn đề liên quan đến
 - Tính an toàn của hệ CSDL
 - Khả năng đáp ứng của hệ CSDL

Thiết kế viên





- DBD Database Designer
- Chịu trách nhiệm về vấn đề:
 - Quyết định những dữ liệu nào cần được lưu trữ
 - Lựa chọn cấu trúc phù hợp để lưu trữ dữ liệu
- Trao đổi với người dùng để nắm bắt được những yêu cầu và đưa ra một thiết kế CSDL thỏa yêu cầu này
- Có thể là thành viên của
 - Nhóm DBA quản lý các CSDL
 - Các nhóm khác sau khi việc thiết kế hoàn tất (nhóm lập trình,...)

Người dùng cuối





End Users

- Người ít sử dụng (Casual users)
 - Ít khi truy cập CSDL, nhưng cần những thông tin khác nhau trong mỗi lần truy cập và dùng những câu truy vấn phức tạp
 - Người quản lý
- Người sử dụng thường xuyên (Naïve users)
 - Thường xuyên truy vấn và cập nhật CSDL nhờ vào một số các chức năng đã được xây dựng sẵn và kiếm thử
 - Nhân viên
- Người sử dụng đặc biệt (Sophisticated users)
 - Thông thạo về HQT CSDL, tự xây dựng những truy vấn phức tạp cho công việc
 - Kỹ sư, nhà khoa học, người phân tích,...
- Người sử dụng cá nhân (stand-alone users)
 - Sử dụng các ứng dụng truy cập HQT CSDL để phục vụ công việc cá nhân
 - Người sử dụng chương trình quản lý thuế cá nhân

Nội dung chi tiết





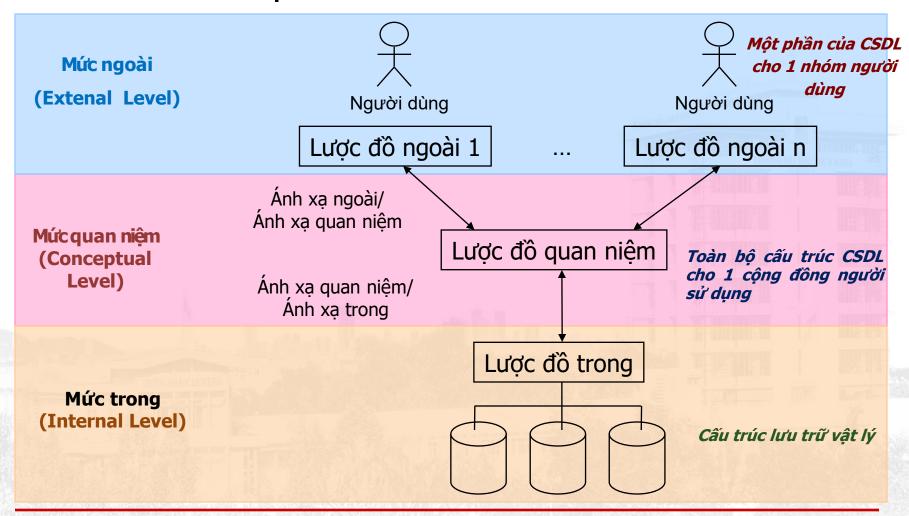
- Giới thiệu
- Quá trình phát triển
- Một số đặc tính của CSDL
- Người sử dụng CSDL
- Kiến trúc của HQT CSDL
- Các tính năng của HQT CSDL
- Các khái niệm
- Ngôn ngữ CSDL

Kiến trúc của HQT CSDL





Kiến trúc 3 lược đồ



Kiến trúc của HQT CSDL (tt)



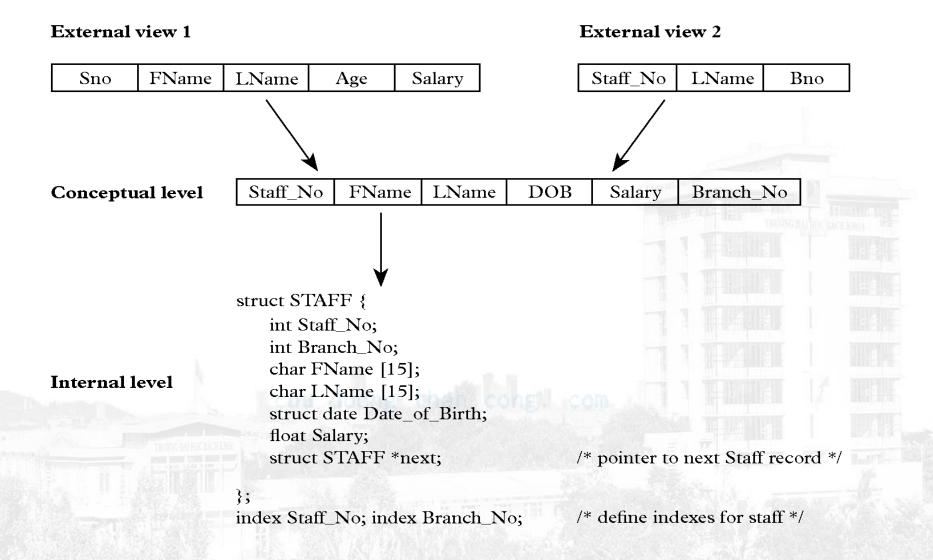


- Mức trong (lược đồ trong)
 - Mô tả cấu trúc lưu trữ vật lý CSDL
- Mức quan niệm (lược đồ quan niệm)
 - Mô tả cấu trúc của toàn thể CSDL cho 1 cộng đồng người sử dụng, gồm thực thể, kiểu dữ liệu, mối liên hệ và ràng buộc
 - Che bớt các chi tiết của cấu trúc lưu trữ vật lý
- Mức ngoài (lược đồ ngoài)
 - Còn gọi là mức khung nhìn (view)
 - Mô tả một phần của CSDL mà 1 nhóm người dùng quan tâm đến
 và che dấu phần còn lại của CSDL đối với nhóm người dùng đó

Ví dụ về 3 mức nhìn về CSDL







Kiến trúc của HQT CSDL (tt)



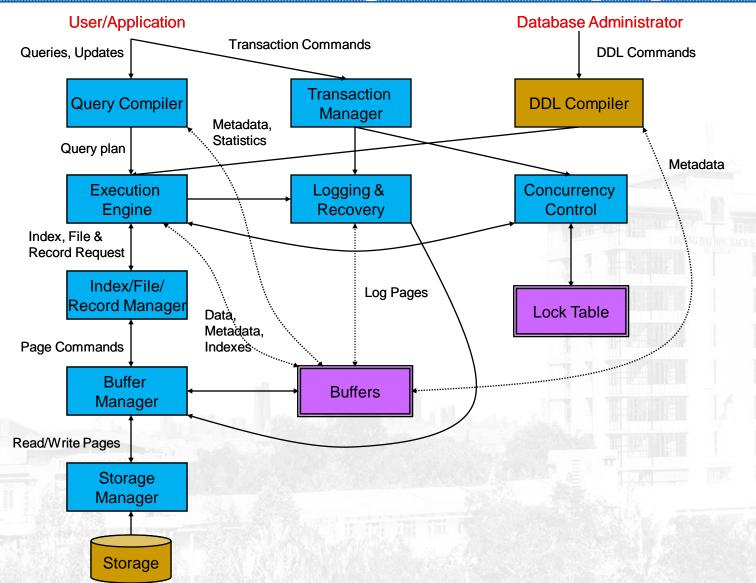


- Độc lập dữ liệu
 - Độc lập logic
 - Khả năng thay đổi lược đồ quan niệm (VD: Thêm/xóa thực thể) mà không thay đổi lược đồ ngoài hoặc các chương trình ứng dụng
 - Độc lập vật lý
 - Khả năng thay đổi lược đồ trong (VD: thay đổi cách thức lưu trữ dữ liệu) mà không làm thay đổi lược đồ quan niệm cũng như lược đồ ngoài

Kiến trúc của HQT CSDL (tt)







Nội dung chi tiết





- Giới thiệu
- Quá trình phát triển
- Một số đặc tính của CSDL
- Người sử dụng CSDL
- Kiến trúc của HQT CSDL
- Các tính năng của HQT CSDL
- Các khái niệm
- Ngôn ngữ CSDL

Các tính năng của HQT CSDL





- · Hạn chế sự trùng lắp dữ liệu
 - Tích hợp các nhu cầu dữ liệu của người dùng để xây dựng một CSDL thống nhất
- Chia sẻ dữ liệu
 - Trong môi trường nhiều người dùng, các HQT phải cho phép truy xuất dữ liệu đồng thời
- Hạn chế những truy cập không được phép
 - Từng người dùng và nhóm người dùng có một tài khoản và mật mã để truy xuất dữ liệu

Các tính năng của HQT CSDL (📆)

- Cung cấp nhiều giao diện giao tiếp với dữ liệu
 - Các ngôn ngữ truy vấn
 - Giao diện ngôn ngữ lập trình
 - Form và giao diện dòng lệnh
 - Giao diện hướng menu và Giao diện ngôn ngữ tự nhiên
- Đảm bảo các ràng buộc toàn vẹn (RBTV Integrity Constraints)
 - Là những qui định cần được thỏa mãn để đảm bảo dữ liệu luôn phản ánh đúng ngữ nghĩa của thế giới thực
 - Một số RBTV Có thể được khai báo với HQT và HQT sẽ tự động kiếm tra.
 - Một số RBTV khác được kiểm tra nhờ ứng dụng
- Khả năng sao lưu dự phòng khi gặp sự cố
 - Có khả năng khôi phục dữ liệu khi có sự hư hỏng về phần cứng hoặc phần mềm

Các tính năng của HQT CSDL (H)

Các tính năng khác

- Chuẩn hóa
 - Cho phép DBA định nghĩa và bắt buộc áp dụng một chuẩn thống nhất cho mọi người dùng
- Uyển chuyển
 - Khi nhu cầu công việc thay đổi, cấu trúc CSDL rất có thể thay đổi, HQT cho phép thêm hoặc mở rộng cấu trúc mà không làm ảnh hưởng đến CTƯD
- Giảm thời gian phát triển ứng dụng
- Khả năng cung cấp thông tin mới nhất
- Cân đối được sự tiết kiệm đầu tư về chia sẻ dữ liệu

Nội dung chi tiết





- Giới thiệu
- Quá trình phát triển
- Một số đặc tính của CSDL
- Người sử dụng CSDL
- Các tính năng của HQT CSDL
- Kiến trúc của HQT CSDL
- Các khái niệm
 - Mô hình dữ liệu (Data Model)
 - Lược đồ CSDL (Database Schema)
 - Thể hiện CSDL (Database Instance)
- Ngôn ngữ CSDL

Mô hình dữ liệu





- Bao gồm
 - Các khái niệm biểu diễn dữ liệu
 - Các phép toán xử lý dữ liệu
- Phân loại
 - Mô hình mức cao
 - Mô hình cài đặt
 - Mô hình mức thấp

Mô hình dữ liệu (tt)





Mô hình mức cao

- Cung cấp các khái niệm gần gũi với người dùng
- Mô hình phải tự nhiên và giàu ngữ nghĩa
- VD: mô hình thực thể kết hợp (ER), mô hình đối tượng...

Mô hình cài đặt

- Đưa ra các khái niệm người dùng có thể hiểu được nhưng không quá xa với cách dữ liệu được tổ chức thật sự trên máy tính
- VD: mô hình quan hệ, mô hình mạng, mô hình phân cấp

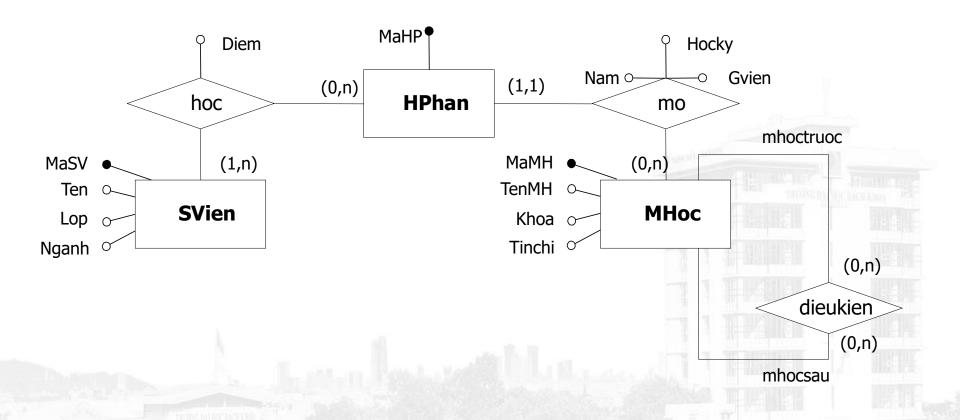
Mô hình mức thấp (mô hình vật lý)

Đưa ra các khái niệm mô tả chi tiết về cách thức dữ liệu được
 lưu trữ trong máy tính

Ví dụ mô hình ER



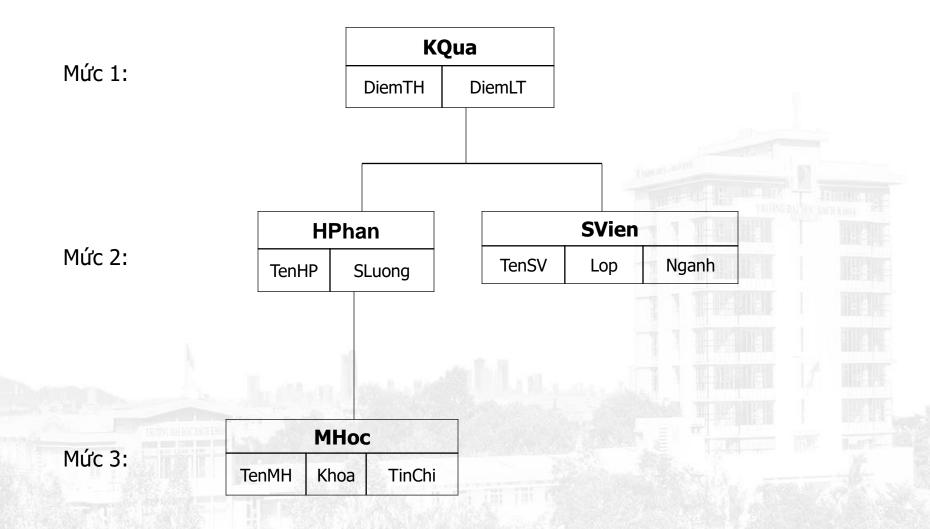




Ví dụ mô hình phân cấp



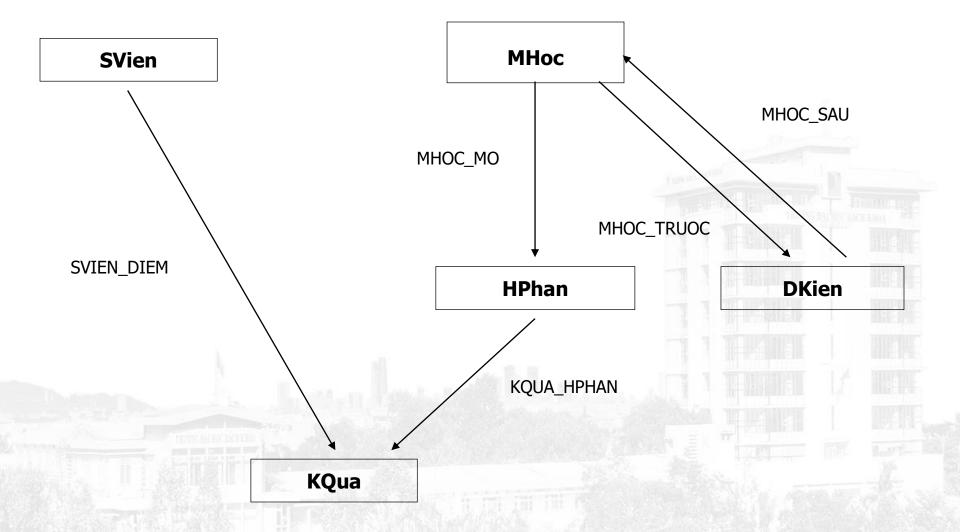




Ví dụ mô hình mạng



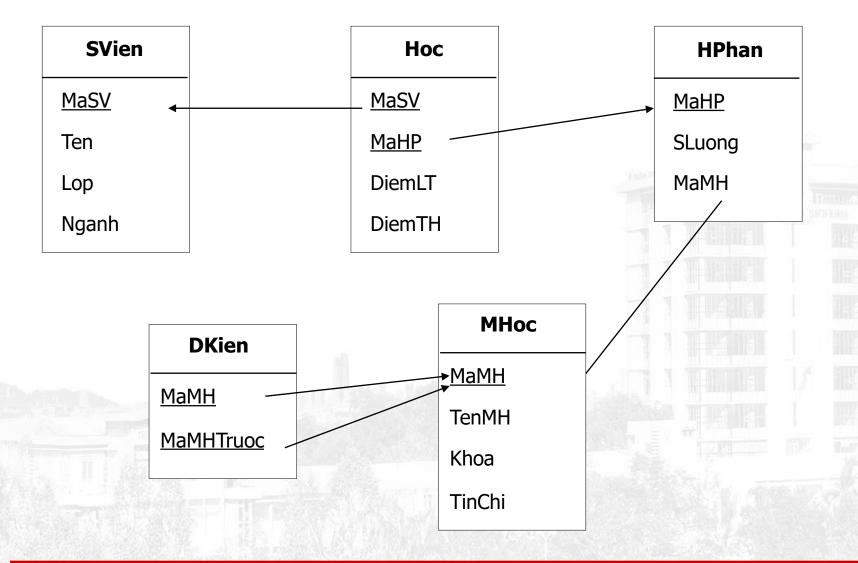




Ví dụ mô hình quan hệ



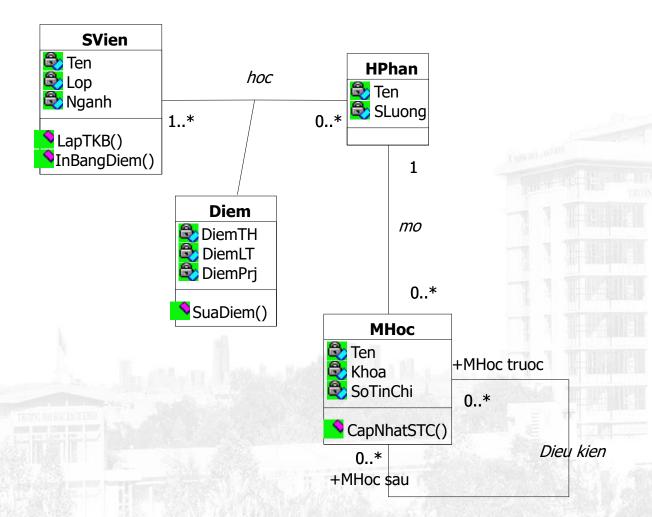




Ví dụ mô hình đối tượng





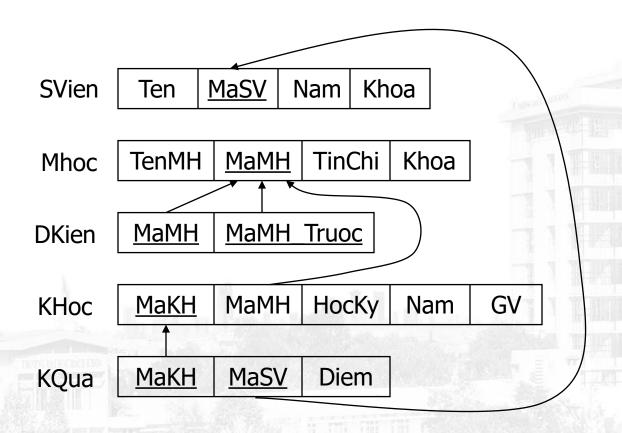


Lược đô CSDL





Là các mô tả về cấu trúc và ràng buộc trên CSDL



Thể hiện





- Là dữ liệu hiện thời được lưu trữ trong CSDL ở một thời điểm nào đó
- Còn gọi là tình trạng của CSDL

Mhoc	TenMH	МаМН	TinChi	Khoa
	Nhap mon tin hoc	COSC1310	4	CNTT
	Cau truc du lieu	COSC3320	4	CNTT
	Toan roi rac	MATH2410	3	TOAN
	Co so du lieu	COSC3380	3	CNTT

KQua	MaSV	MaKH	Diem	
	17	112	8	
	17	119	6	
	8	85	10	
	8	92	9	
	8	102	8	
	8	135	10	
	0	133	10	

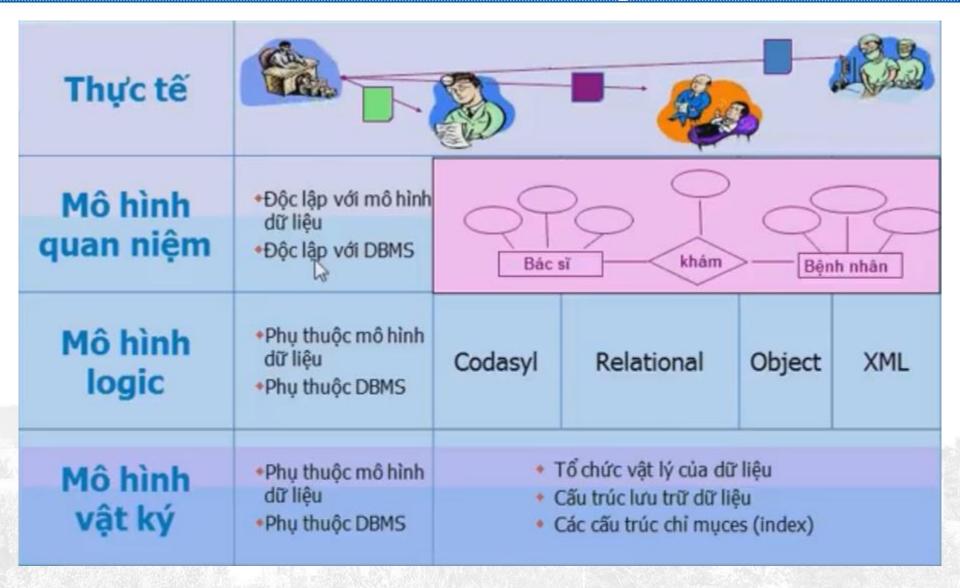
SVien	Ten	MaSV	Nam	Khoa
	Son	17	1	CNTT
	Bao	8	2	CNTT

DKien	MaMH MaMH_Tru	
	COSC3380	COSC3320
	COSC3380	MATH2410
	COSC3320	COSC3380

Mô hình hóa nhiều cấp







Nội dung chi tiết



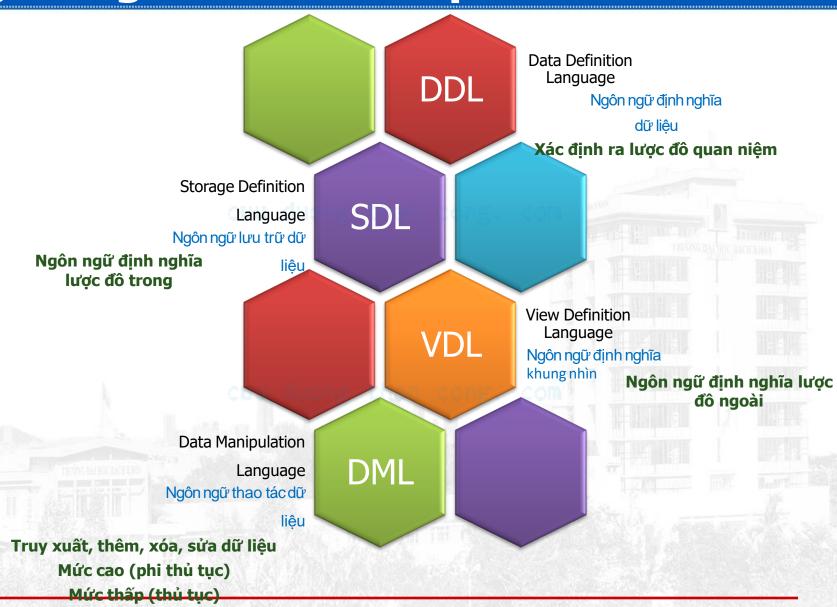


- Giới thiệu
- Quá trình phát triển
- Một số đặc tính của CSDL
- Người sử dụng CSDL
- Các tính năng của HQT CSDL
- Kiến trúc của HQT CSDL
- Các khái niệm
- Ngôn ngữ CSDL

Ngôn ngữ Cơ sở dữ liệu







Ngôn ngữ CSDL





- Ngôn ngữ định nghĩa dữ liệu (DDL Data Definition Language)
 - Xác định ra lược đồ quan niệm
- Ngôn ngữ lưu trữ dữ liệu (SDL Storage Definition Language)
 - Ngôn ngữ định nghĩa lược đồ trong
- Ngôn ngữ định nghĩa khung nhìn (VDL View Definition Language)
 - Ngôn ngữ định nghĩa lược đồ ngoài
- Ngôn ngữ thao tác dữ liệu (DML Data Manipulation Language)
 - Cho phép truy xuất, thêm, xóa, sửa dữ liệu
 - Mức cao (phi thủ tục)
 - Mức thấp (thủ tục)

Ngôn ngữ CSDL (tt)





- Ngôn ngữ thao tác dữ liệu
 (DML Data Manipulation Language)
 - Cho phép truy xuất, thêm, xóa, sửa dữ liệu
 - Mức cao (phi thủ tục)
 - Mức thấp (thủ tục)

Câu hỏi





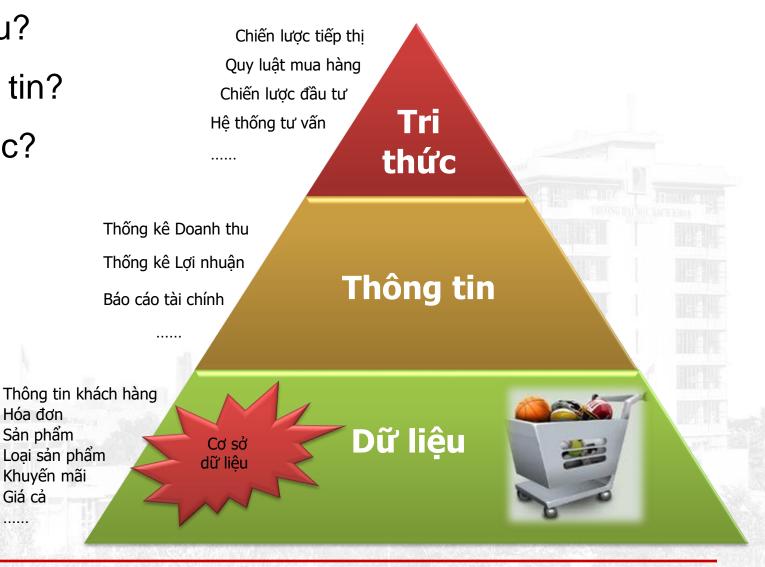
- Khi nào nên dùng và không nên dùng cách tiếp cận CSDL?
- Các vấn đề ảnh hưởng
 - Đầu tư cao cho phần cứng, phần mềm, đào tạo
 - Chức năng chính chỉ gồm định nghĩa và xử lý dữ liệu
 - Tốn chi phí cho vấn đề bảo mật, điều khiển trùng lấp thông tin, chức năng phục hồi và toàn vẹn dữ liệu
- Không phù hợp với ứng dụng :
 - Nhỏ, có tổ chức dữ liệu ít và tốt, không có nhu cầu thay đổi
 - Có yêu cầu về xử lý thời gian thực nghiêm ngặt
 - Không có nhiều người truy cập dữ liệu đồng thời

Câu hỏi





- Dữ liệu?
- Thông tin?
- Tri thức?



Giá cả





