

CƠ sở dữ liệu

IS222



CƠ SỞ DỮ LIỆU

Nội dung chính

- I. *Mở đầu: Một số khái niệm*
- II. Mô hình CSDL
- III. Mô hình thực thể - liên kết
- IV. Mô hình quan hệ
- V. Phụ thuộc hàm và Các dạng chuẩn hóa
- VI. Đại số quan hệ

Chương 1. Tổng quan về Hệ CSDL

■ Mở đầu

- 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

1.1 Mở đầu

□ Thông tin và dữ liệu

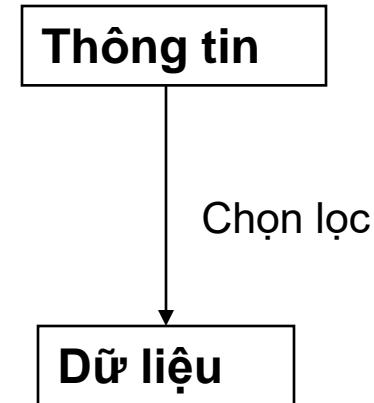
▪ Thông tin:

- Kinh doanh
- Thời tiết
- Giáo dục
- ...

▪ Dữ liệu (Data)

- Khách hàng: Tên, địa chỉ, số điện thoại
- Sinh viên: Họ tên, masv, địa chỉ

Dữ liệu là những sự kiện (thông tin) có thể ghi lại được và có nghĩa, có mục đích sử dụng.



1.1 Mở đầu

❑ Cơ sở dữ liệu (Database): *Một tập hợp những dữ liệu có liên quan với nhau được lưu trữ trong máy tính đáp ứng nhu cầu khai thác của nhóm người sử dụng với các mục đích xác định nào đó.*

❖ *Các tính chất:*

- Một CSDL biểu diễn một phần của thế giới thực (thế giới nhỏ - miniworld), được cập nhật phản ánh sự thay đổi của thế giới nó biểu diễn.
- Một cơ sở dữ liệu là một tập hợp dữ liệu liên kết với nhau một cách logic và mang một nghĩa nào đó.
- Một cơ sở dữ liệu được thiết kế và được phổ biến cho một mục đích riêng, có nhóm người sử dụng.

1.1 Mở đầu

❖ Tại sao cần CSDL ?

- Dư thừa dữ liệu và không nhất quán,
- Nhu cầu truy cập dữ liệu
- Vấn đề toàn vẹn dữ liệu
- Vấn đề tính nguyên tố của giao tác
- Vấn đề nhiều người dùng tương tranh
- Vấn đề an toàn

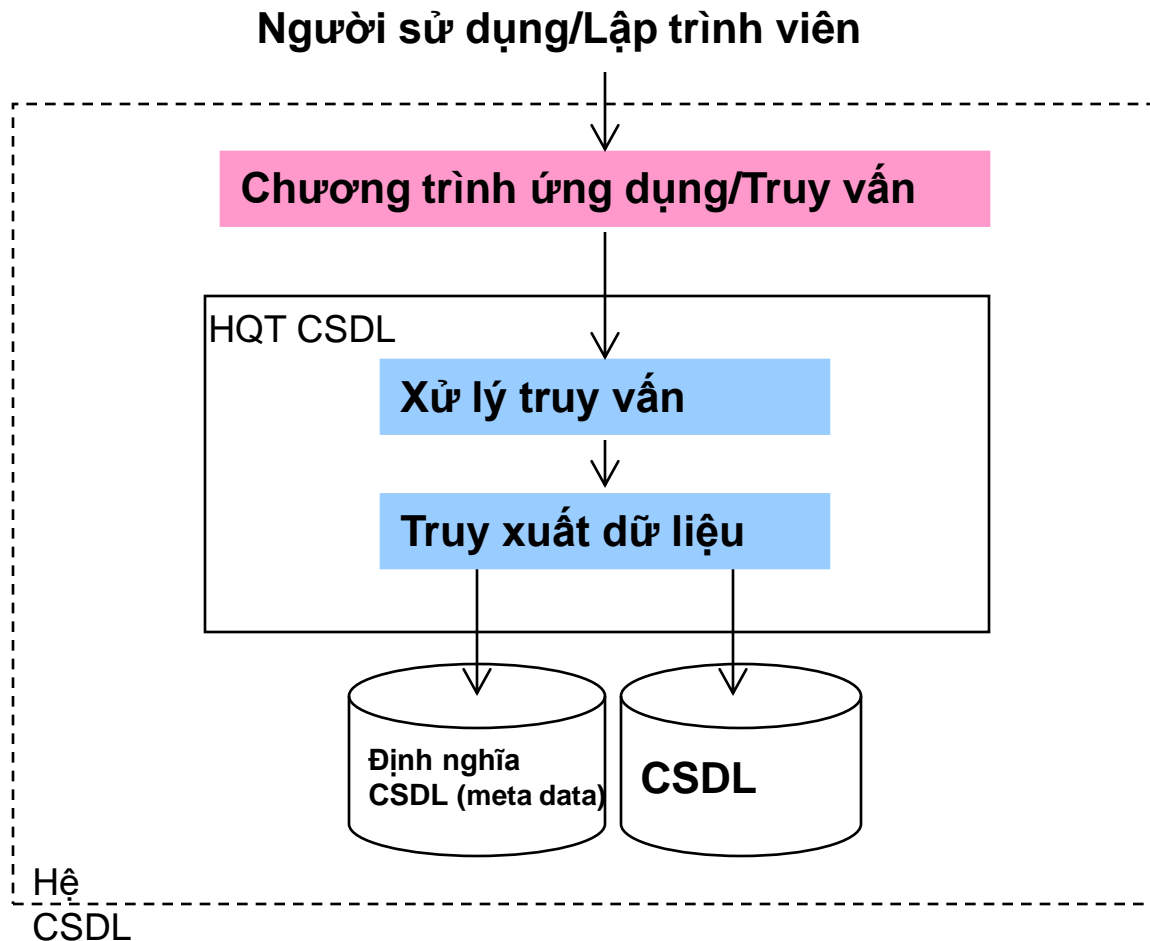
1.1 Mở đầu

□ *Hệ quản trị CSDL (Database Management System)*

- Tập hợp các chương trình cho phép người sử dụng tạo ra, bảo trì và khai thác CSDL; tức là phần mềm cho phép **định nghĩa, xây dựng** và **thao tác** xử lý dữ liệu
 - Định nghĩa CSDL: đặc tả các kiểu dữ liệu, cấu trúc mô tả chi tiết về dữ liệu, các ràng buộc của dữ liệu trong CSDL.
 - Xây dựng CSDL: lưu trữ dữ liệu trên bộ nhớ phụ
 - Thao tác: truy vấn, cập nhật và sinh báo cáo.

1.1 Mở đầu

- Hệ CSDL (Database System)



Một ví dụ về CSDL: Quản lý Đào tạo

- *Các đối tượng:* Sinh viên, Môn học, Học phần, Ngành..
- *Mối quan hệ:*
 - Mỗi sinh viên có thể học một số học phần trong mỗi kỳ
 - Mỗi môn học được dạy trong một hoặc nhiều học phần xác định
 - Mỗi môn học có thể có các môn điều kiện
 - Mỗi học phần do một giáo viên giảng
 - Mỗi môn học do một khoa quản lý
 - Mỗi sinh viên thuộc một ngành nào đó
 - Mỗi ngành do một khoa quản lý.

Một ví dụ về CSDL: Quản lý Đào tạo

SINHVIEN			
MaSv	Hoten	Lop	MaNganh

MONHOC			
MaMH	TenMH	SoTC	MaKhoa

MONDK	
MaMH	MaMonDK

HOCPHAN			
MaHP	MaMH	Hocky	Giaovien

NGANH	
MaNg	TenNg

KHOA		
MaKhoa	TenKhoa	MaNganh

DIEM		
MaSv	MaHP	Diem

Một ví dụ về CSDL: Quản lý Đào tạo

Định nghĩa CSDL: Xác định cấu trúc các bảng, Các kiểu dữ liệu của các trường,

Xây dựng CSDL: Dữ liệu được ghi trong các tệp như thế nào, các quan hệ giữa các bản ghi,

Thao tác CSDL: Cập nhật và khai thác dữ liệu

Một ví dụ về CSDL: Quản lý Dự án

NHANVIEN	HONV	TENDEM	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

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

Một ví dụ về CSDL

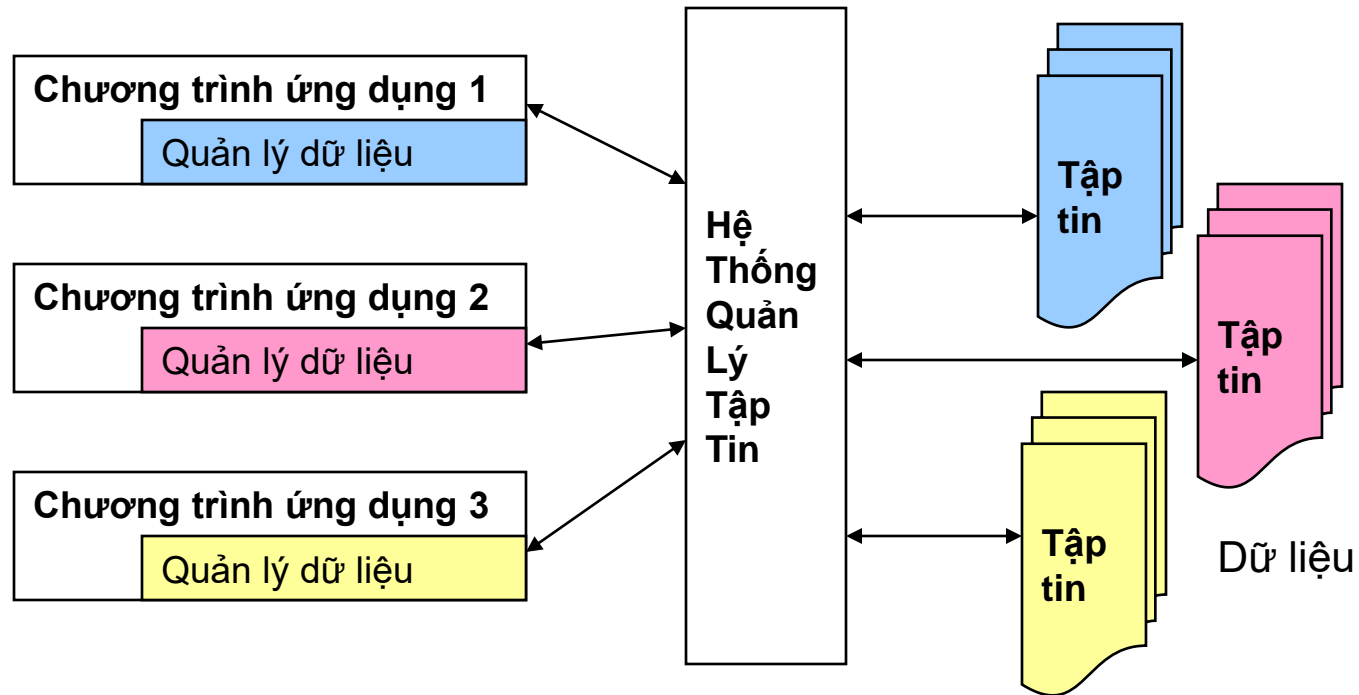
- Định nghĩa CSDL: Cấu trúc bảng, bao gồm các thành phần dữ liệu và kiểu dữ liệu tương ứng
- Xây dựng CSDL: Đưa dữ liệu vào các bảng
- Xử lý CSDL:
 - Thực hiện các truy vấn: “Cho biết những nhân viên thuộc phòng nào đó”
 - Thực hiện các phép cập nhật: thêm nhân viên, chuyển phòng của một nhân viên.
 - Xóa một nhân viên,...

Nội dung chi tiết

- Mở đầu
- **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

1.2 Quá trình phát triển

- Tập tin (trước 1960)

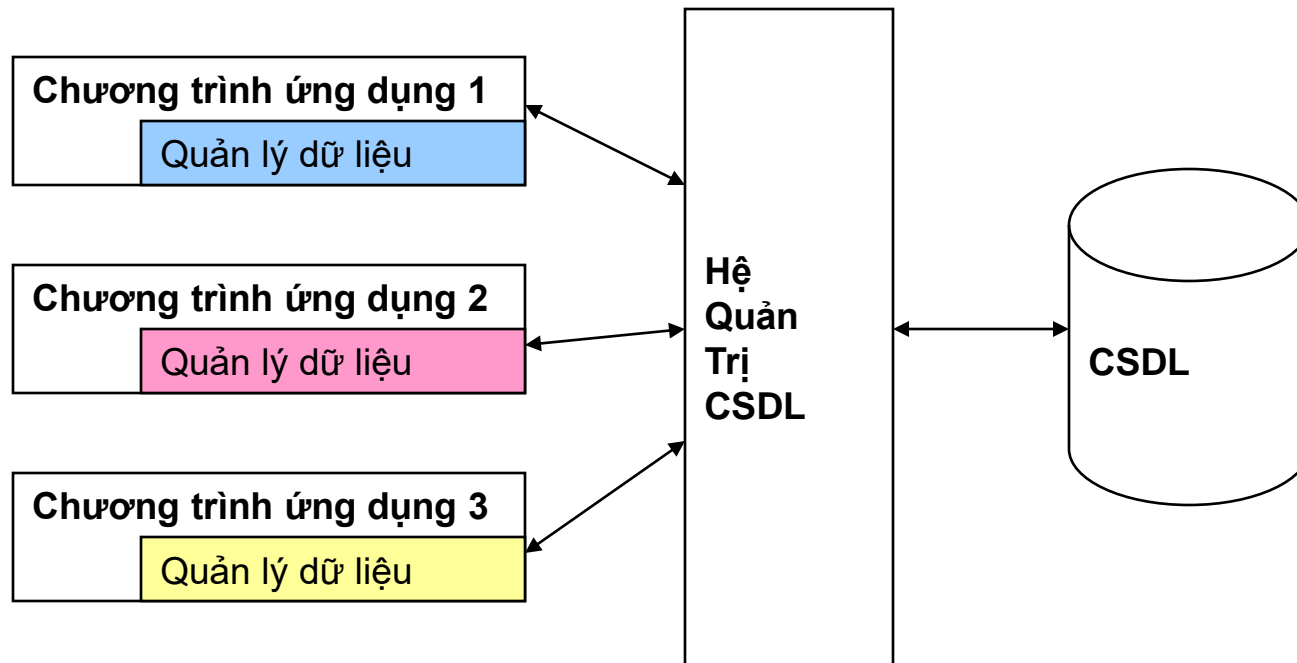


1.2 Quá trình phát triển

- *Hạn chế*
 - Dữ liệu bị trùng lặp 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

1.2 Quá trình phát triển

- Cơ sở dữ liệu (Database)



1.2 Quá trình phát triển

- 1960: Hệ QTCSDL đầu tiên ra đời, dựa trên mô hình phân cấp : IMS (Information Management System) của IBM; IDS (Integrated Data Store) dựa trên mô hình mạng
- 1976: HQTCSDL theo mô hình quan hệ đầu tiên xuất hiện (System – R) của IBM; Từ 1980 nhiều HQTCSDL mới ra đời: Oracle, DB2, Access, Foxpro, Paradox
- Từ 1990 Xuất hiện HQTCSDL hướng đối tượng
- Hệ quản trị CSDL theo mô hình NoSQL

Nội dung chi tiết

- Mở đầu
- 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
 - Hỗ trợ nhiều khung nhìn
 - Chia sẻ dữ liệu và nhiều người dùng
- Người sử dụng CSDL
- Kiến trúc của HQT CSDL
- Các tính năng của HQT CSDL
- Ngôn ngữ CSDL

1.3- Một số đặc tính của CSDL

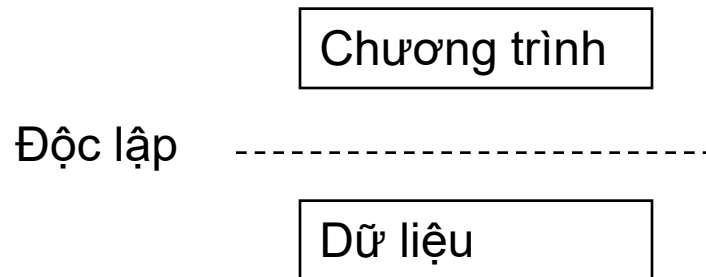
■ Tính tự mô tả

- Hệ CSDL không chỉ chứa bản thân CSDL mà còn chứa định nghĩa(mô tả) của CSDL
- Các định nghĩa được lưu trữ trong catalog : 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 (siêu dữ liệu)
- Các Chương trình ứng dụng có thể truy xuất đến nhiều CSDL nhờ thông tin cấu trúc được lưu trữ trong catalog

1.3 - Một số đặc tính của CSDL

- 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



1.3 - Một số đặc tính của CSDL

■ Tính trừu tượng:

- cho phép trình bày dữ liệu ở một mức trừu tượng nào đó, nhằm che bớt những chi tiết lưu trữ thật của dữ liệu.

■ Tính nhất quán:

- Lưu trữ dữ liệu thống nhất: tránh được tình trạng trùng lặp 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

■ Chia sẻ dữ liệu và nhiều người dùng

- Nhiều người sử dụng, quản lý tương tranh
- Đảm bảo tính đúng đắn đối với các thao tác

1.3 - Một số đặc tính của CSDL

■ Các 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
- Khung nhìn (view) là :
 - ✓ Một phần của CSDL hoặc
 - ✓ Dữ liệu tổng hợp từ CSDL

Nội dung chi tiết

- Mở đầu
- 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)
- Các khái niệm
- Kiến trúc của HQT CSDL
- Các tính năng của HQT CSDL
- Ngôn ngữ CSDL

1.4 Người sử dụng CSDL

○ Quản trị viên (DBA)- quản lý hệ CSDL

- ✓ Cấp quyền truy cập CSDL
- ✓ Điều phối và giám sát việc sử dụng CSDL

○ Thiết kế viên: Chịu trách nhiệm :

- Lựa chọn cấu trúc phù hợp để lưu trữ dữ liệu
- Quyết định những dữ liệu nào cần được lưu trữ,
- Liên hệ với người dùng để nắm bắt được những yêu cầu và đưa ra một thiết kế CSDL thỏa yêu cầu này,
- Có thể là 1 nhóm các DBA quản lý các CSDL sau khi việc thiết kế hoàn tất

1.4 Người sử dụng CSDL

Người dùng cuối

- *Người ít sử dụng* (người quản lý)
 - 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 sử dụng thường xuyên* (nhân viên)
 - Thường xuyên truy vấn và cập nhật CSDL nhờ vào một số các chức năng đã được xây dựng sẵn
- *Người sử dụng đặc biệt*
 - Thông thạo về HQT CSDL, tự xây dựng những truy vấn phức tạp cho công việc
 - Kỹ sư, nhà khoa học, người phân tích,...

Nội dung chi tiết

- Mở đầu
- Quá trình phát triển
- Một số đặc tính của CSDL
- Người sử dụng CSDL
- **Các khái niệm:** Mô hình, Lược đồ, Thể hiện
- Kiến trúc của Hệ CSDL
- Các tính năng của HQT CSDL
- Ngôn ngữ CSDL

1.5 - Các khái niệm

a. Mô hình dữ liệu

- Mô hình dữ liệu (Data Model) bao gồm
 - Tập hợp các khái niệm biểu diễn cấu trúc CSDL:
 - Kiểu dữ liệu
 - Các liên kết,
 - Các ràng buộc
 - Các phép toán xử lý dữ liệu

1.5 - Các khái niệm

a. Mô hình dữ liệu

1. Mô hình mức cao (*mức quan niệm*)

- 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 thực thể...

2. Mô hình cài đặt (*mức logic, thể hiện*)

- 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.

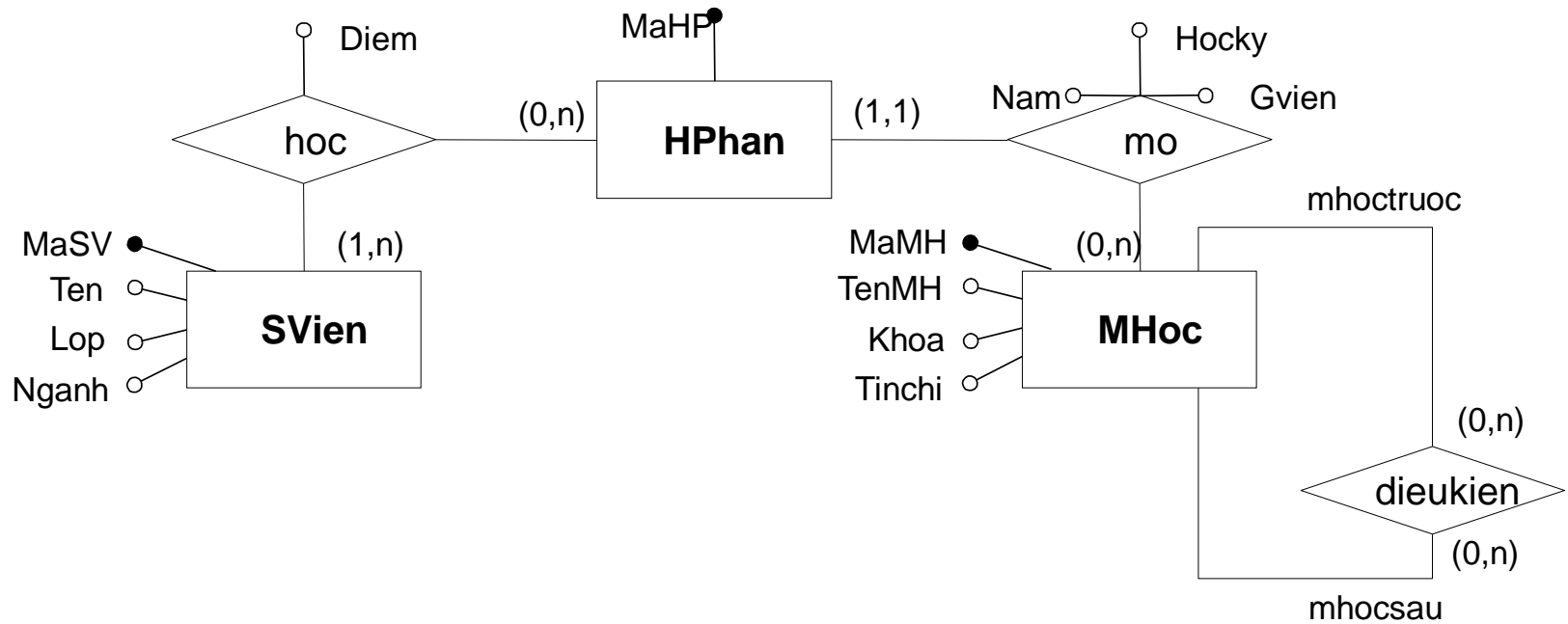
Dùng các khái niệm: thực thể (entity), thuộc tính, mối liên kết (quan hệ) giữa các thực thể

VD: mô hình quan hệ, mô hình mạng, mô hình phân cấp

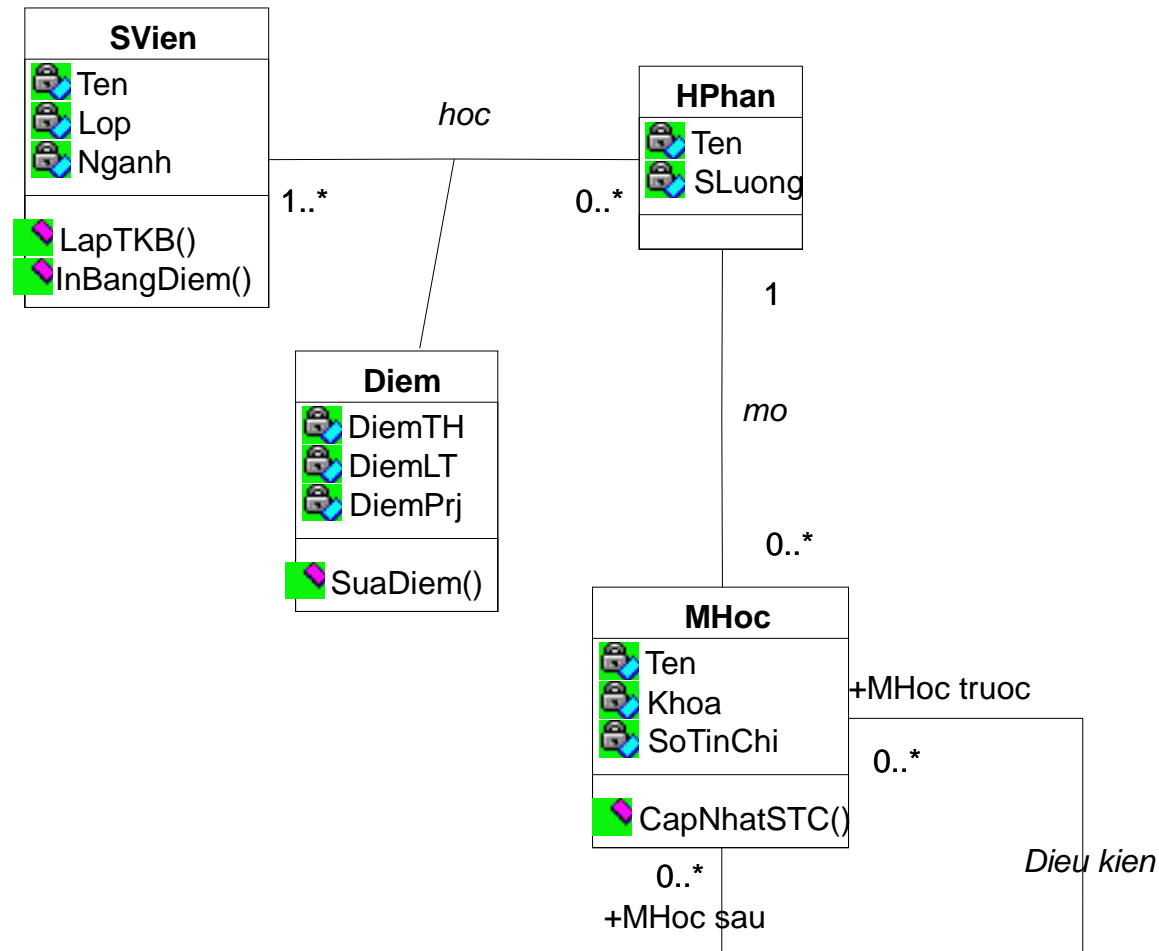
3. Mô hình mức thấp (*mức hình vật lý*)

- Đưa ra các khái niệm mô tả chi tiết về cách thức dữ liệu được lưu trữ trong máy tính: khuôn dạng bản ghi, cách sắp xếp, đường truy cập

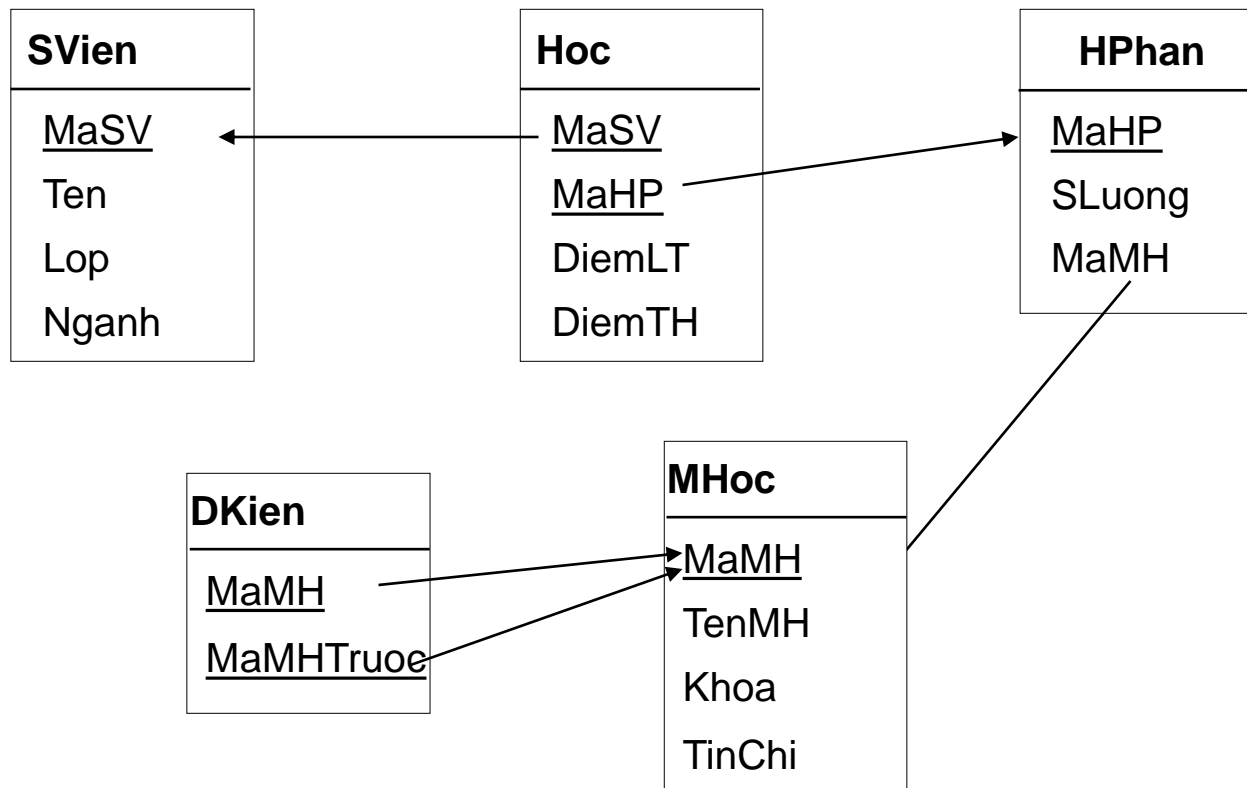
Ví dụ mô hình ER



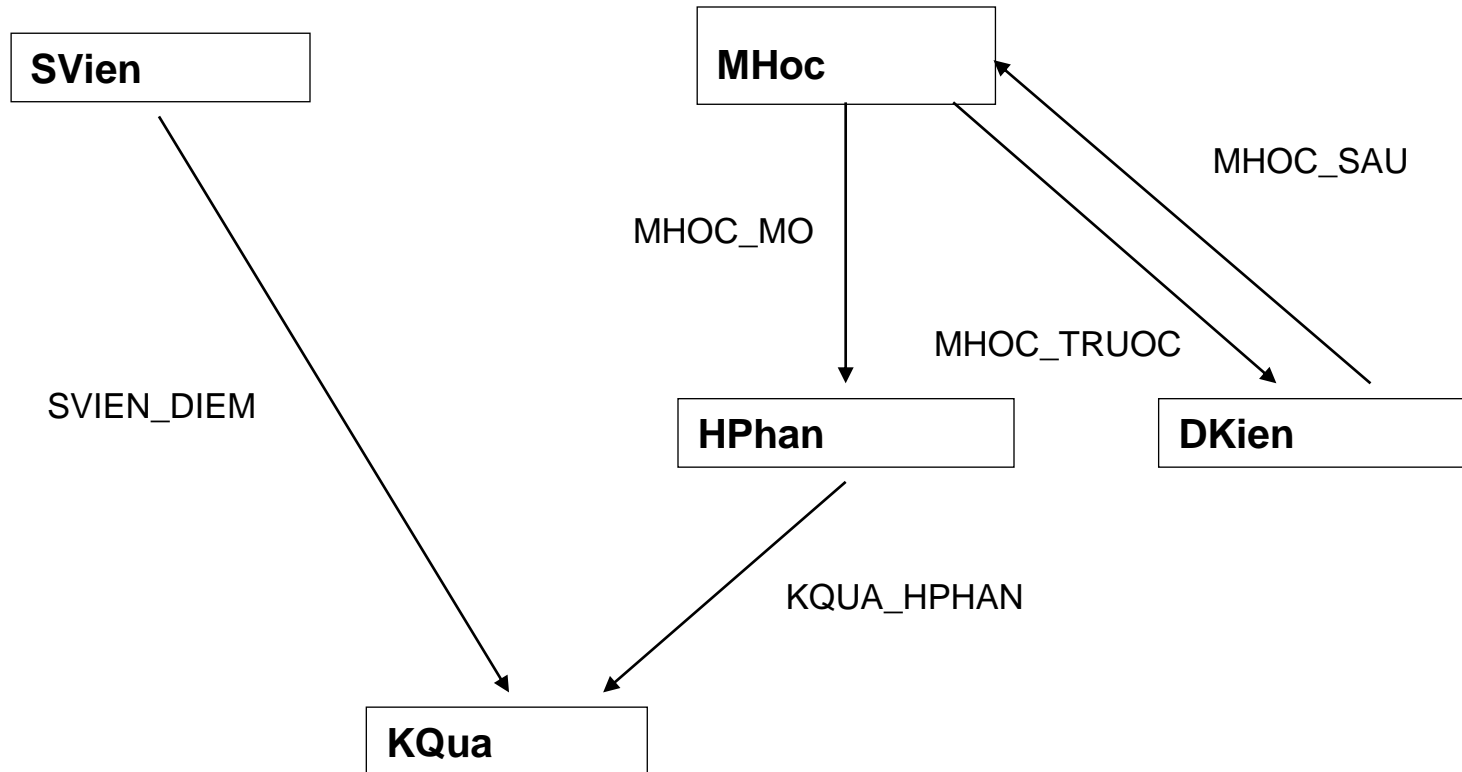
Ví dụ mô hình đối tượng



Ví dụ mô hình quan hệ



Ví dụ mô hình mạng



Ví dụ mô hình phân cấp

Mức 1:

KQua	
DiemTH	DiemLT

Mức 2:

HPhan	
TenHP	SLuong

SVien		
TenSV	Lop	Nganh

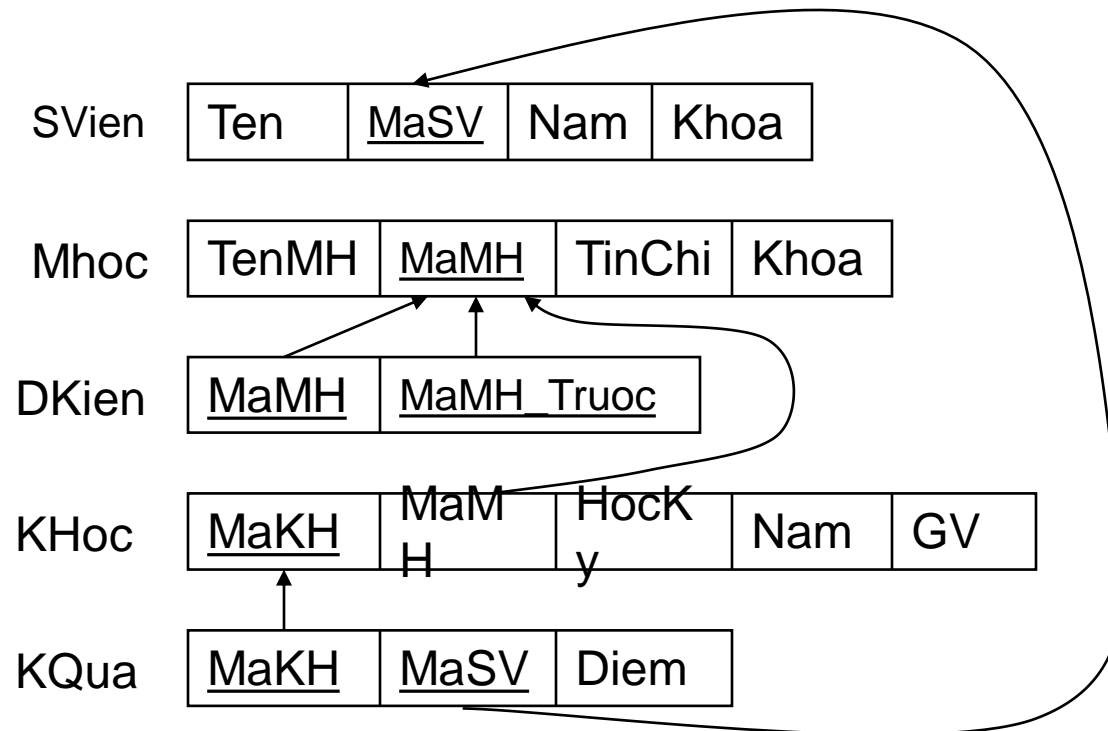
Mức 3:

MHoc		
TenMH	Khoa	TinChi

1.5 - Các khái niệm

b. Lược đồ

- Lược đồ CSDL (Database Schema)
 - Là các mô tả biểu diễn về CSDL(cấu trúc và ràng buộc trên CSDL)



1.5 - Các khái niệm

c. Thể hiện CSDL (*Database Instance*)

- Là dữ liệu hiện thời được lưu trữ trong CSDL ở một thời điểm nào đó
- Tình trạng của CSDL

Mhoc	TenMH	MaMH	TinChi	Khoa
	Nhap mon tin hoc	COSC131	4	CNTT
	Cau truc du lieu	COSC332	4	CNTT
	Toan roi rac	MATH241	3	TOAN
	Co so du lieu	COSC338	3	CNTT

0

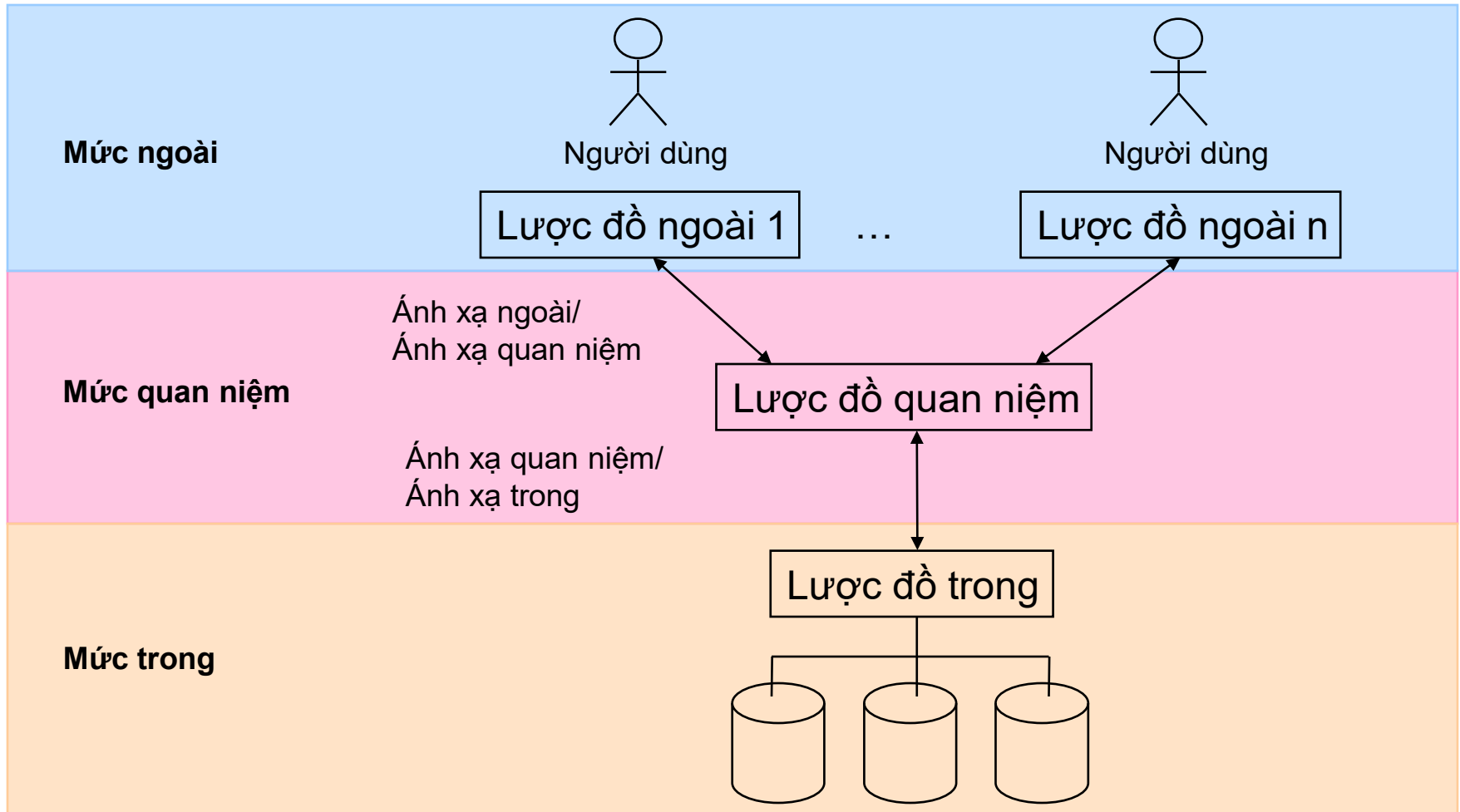
KQua	MaSV	MaKH	Diem
	17	112	8
	17	119	6
	8	85	10
	8	92	9
	8	102	8
	8	135	10

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

DKien	MaMH	MaMH_Truoc
	COSC3380	COSC3320
	COSC3380	MATH2410
	COSC3320	COSC3380

1.6- Kiến trúc 3 lược đồ của Hệ CSDL

- Kiến trúc 3 lược đồ (3 mức - theo ANSI/SPARC)



1.6- Kiến trúc 3 mức của Hệ CSDL

- **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

1.6- Kiến trúc 3 lược đồ của Hệ CSDL

- *Độc lập dữ liệu*

- Độc lập logic

- Khả năng thay đổi lược đồ quan niệm mà không thay đổi lược đồ ngoài hoặc các chương trình ứng dụng

- Độc lập vật lý

- Khả năng thay đổi lược đồ trong mà không làm thay đổi lược đồ quan niệm cũng như lược đồ ngoài

Nội dung chi tiết

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

1.7 - Các tính năng của HQT CSDL

1. Kiểm soát được tính dư thừa của dữ liệu

- Tích hợp các nhu cầu dữ liệu của người dùng để xây dựng một CSDL thống nhất

2. Chia sẻ dữ liệu

- Trong môi trường đa người dùng, các HQT phải cho phép truy xuất dữ liệu đồng thời

3. Hạn chế những truy cập không cho phép

- Từng người dùng và nhóm người dùng có một quyền (tài khoản và mật mã) nhất định để truy xuất dữ liệu

4. Cung cấp nhiều giao diện

- HQT cung cấp ngôn ngữ giữa CSDL và người dùng

1.7 - Các tính năng của HQT CSDL

5. 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 ràng buộc có thể được khai báo với HQT và HQT sẽ tự động kiểm tra.
- Một số ràng buộc khác được kiểm tra nhờ chương trình ứng dụng

6. 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

1.7 - Các tính năng của HQT CSDL

- 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 chương trình ứng dụng
 - Giảm thời gian phát triển ứng dụng
 - Tính khả dụng
 - Khi có một sự thay đổi lên CSDL, tất cả người dùng đều thấy được

Nội dung chi tiết

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

1.8- Ngôn ngữ CSDL

1. *Ngôn ngữ định nghĩa dữ liệu* (DDL – Data Definition Language)

- Xác định ra lược đồ quan niệm
- Được dùng bởi DBA và người thiết kế để xây dựng lược đồ.
- DBMS dịch các câu lệnh trong DDL thành mô tả xây dựng lược đồ

Ví dụ

```
CREATE TABLE employees (  
    id INTEGER PRIMARY KEY,  
    first_name CHAR(50) null,  
    last_name CHAR(75) not null,  
    date_of_birth DATE null );
```

1.8- Ngôn ngữ CSDL

2. ***Ngôn ngữ thao tác dữ liệu*** (DML – Data Manipulation Language)

- Cho phép truy xuất, thêm, xóa, sửa dữ liệu
- Mức cao (phi thủ tục): SQL
- Mức thấp (thủ tục): được nhúng trong một ngôn ngữ lập trình “vạn năng”.

Ví dụ

- Các câu lệnh trong SQL: SELECT, INSERT, UPDATE, và DELETE.
- SELECT id, last_name FROM employees

1.8- Ngôn ngữ CSDL

3. ***Ngôn ngữ điều khiển giao dịch*** (Transaction Control Language - TCL)

- Đảm bảo tính toàn vẹn dữ liệu khi thực hiện các tác vụ có sự thay đổi dữ liệu
- Các câu lệnh SQL tương ứng:
 - COMMIT, ROLLBACK, và SAVEPOINT.

4. ***Ngôn ngữ điều khiển dữ liệu*** (Data Control Language - DCL)

- Cung cấp các tính năng bảo vệ cho các đối tượng của CSDL
- Các câu lệnh SQL tương ứng:
 - GRANT và REVOKE.

Tài liệu tham khảo

- ***Giáo trình CSDL***
 - Chương 1,
- ***Database management system***
 - Chapter 1
- ***An introduction to Database System***
 - Chapter 1, 2