



Cơ sở dữ liệu

CHƯƠNG 1: TỔNG QUAN VỀ HỆ CSDL (tiếp)

Nguyễn Thị Hậu
Đại học Công nghệ - Đại học Quốc Gia Hà Nội
nguyenhau@vnu.edu.vn

Chương 1. Tổng quan về Hệ CSDL

- Một số khái niệm
- 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
- Một số khái niệm: Mô hình, Lược đồ, Thể hiện
- Ngôn ngữ CSDL

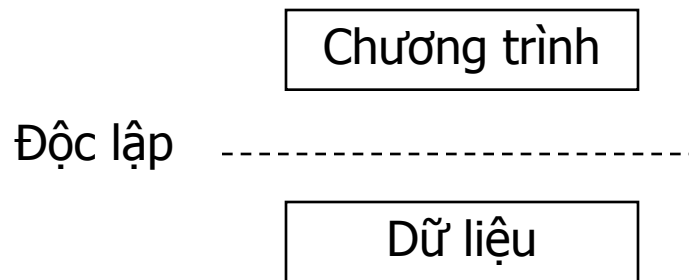
1.3 MỘT SỐ ĐẶC TÍNH CỦA CSDL

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

TÍNH ĐỘC LẬP

- Dữ liệu và chương trình có tính độc lập tương đối (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)



MỘT SỐ ĐẶC TÍNH KHÁC

- Tính trừu tượng:
 - Cho phép trình bày dữ liệu ở một mức trừu tượng nào đó, nhằm che bớt những chi tiết lưu trữ vật lý 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

Hỗ trợ nhiều 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

1.4 NGƯỜI SỬ DỤNG CSDL

Quản trị viên, Thiết kế viên

- 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,

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,...

1.5 CÁC KHÁI NIỆM: MÔ HÌNH DỮ LIỆU, LƯỢC ĐỒ, THỂ HIỆN

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

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 - liên kết (ER)

2. Mô hình mức thể hiện (*mức logic*)

- Dùng 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 dữ liệu 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 phân cấp

Mức 1:

KQua	
DiemTH	DiemLT

Mức 2:

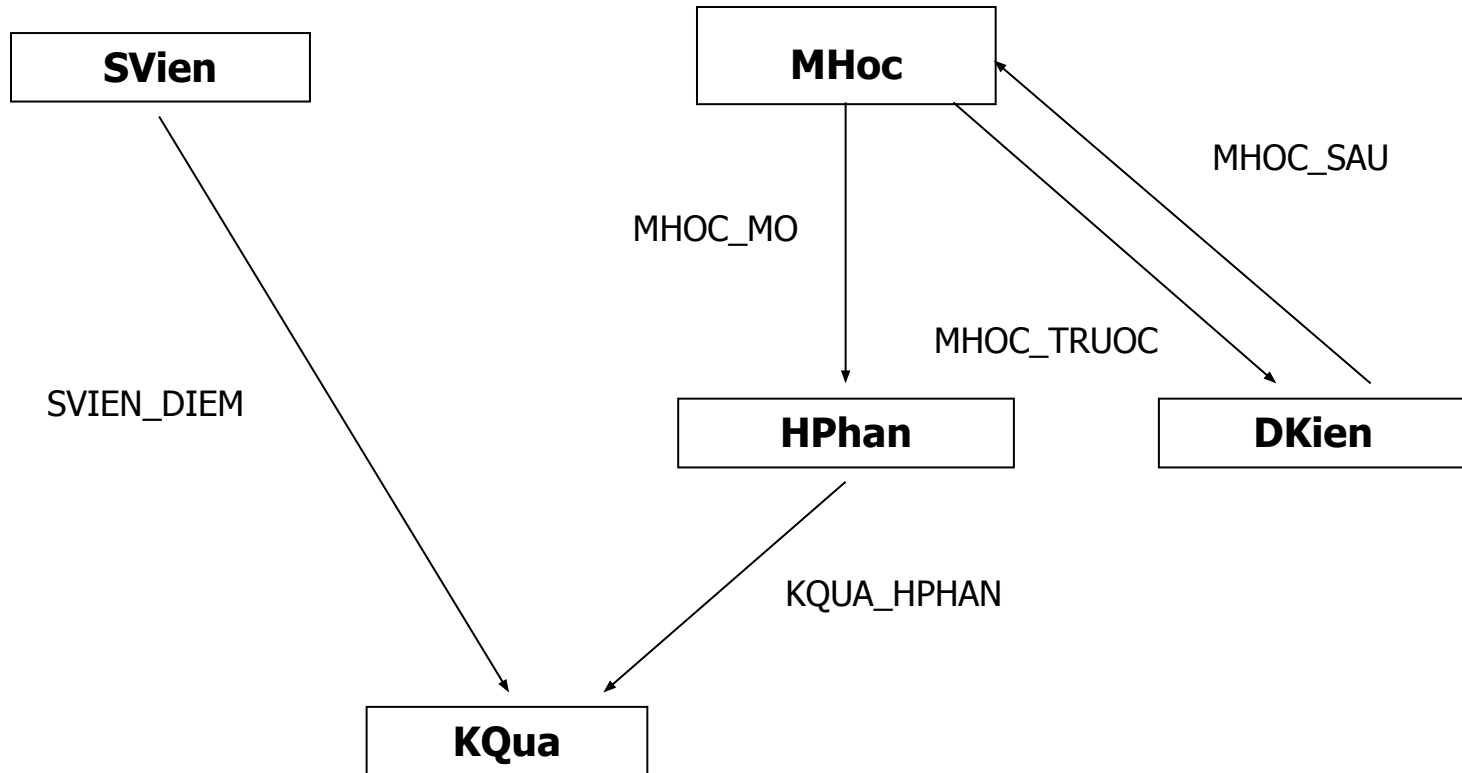
HPhan	
TenHP	SLuong

SVien		
TenSV	Lop	Nganh

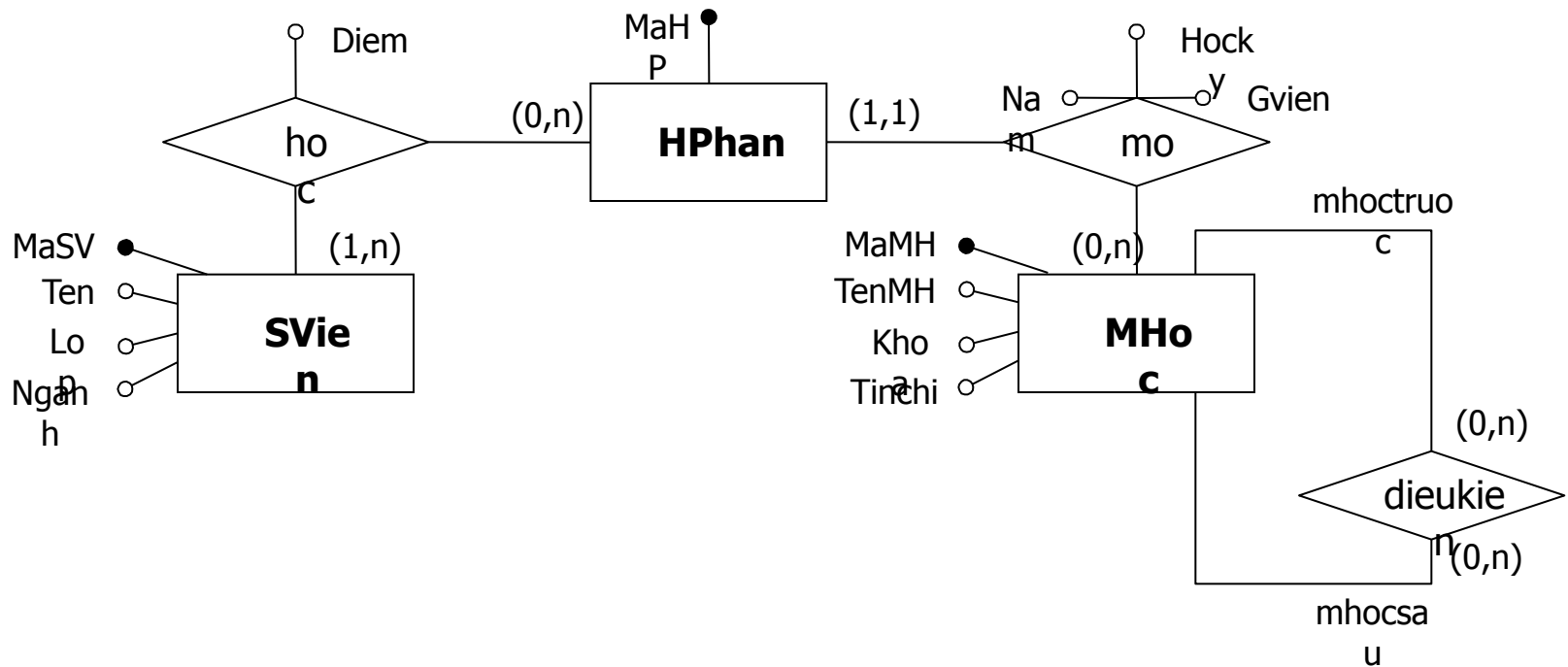
Mức 3:

MHoc		
TenMH	Khoa	TinChi

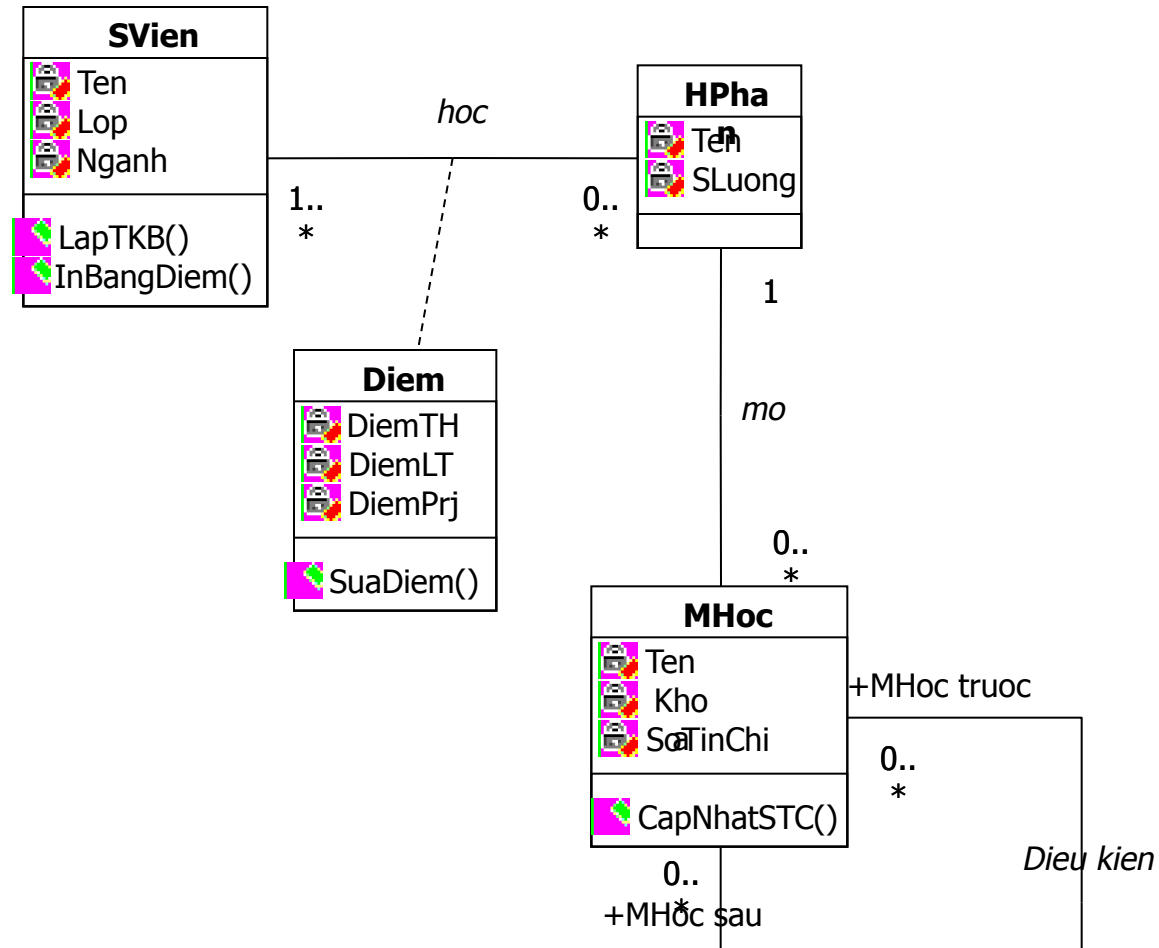
Ví dụ mô hình mạng



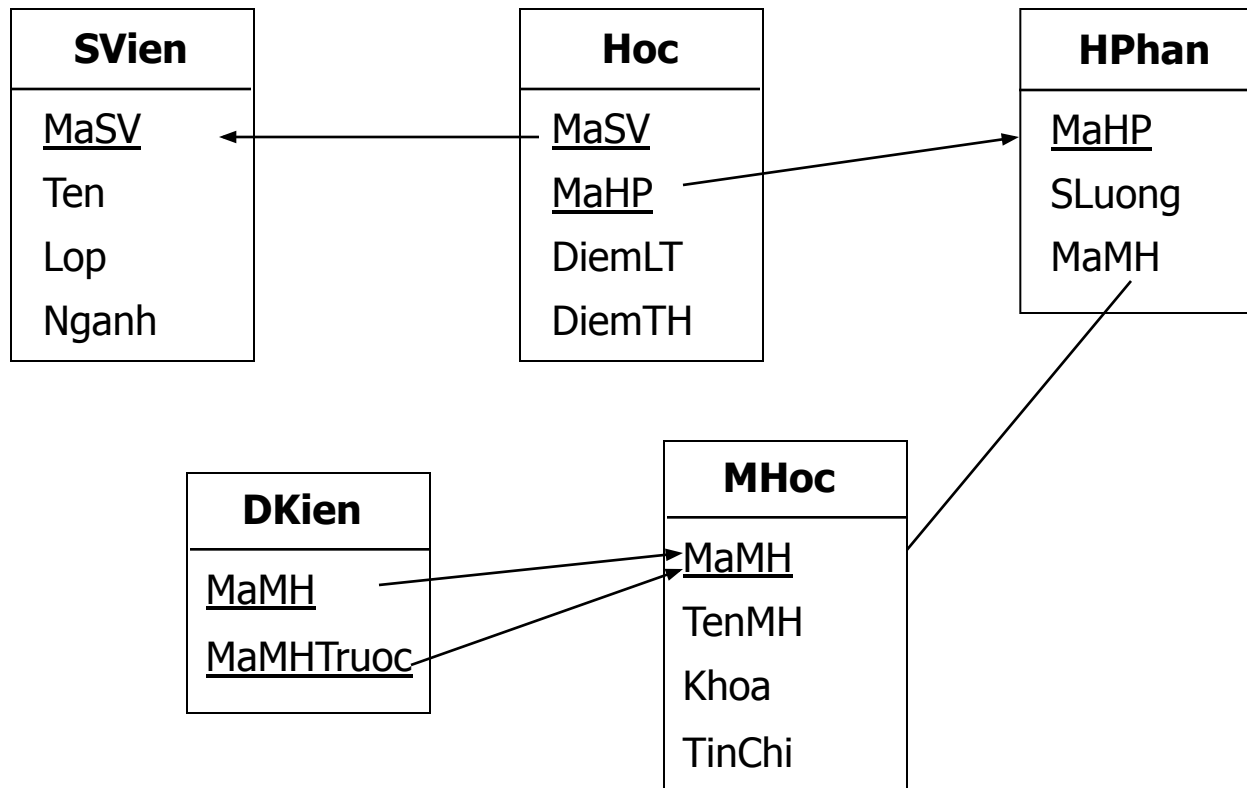
Ví dụ mô hình ER



Ví dụ mô hình đối tượng

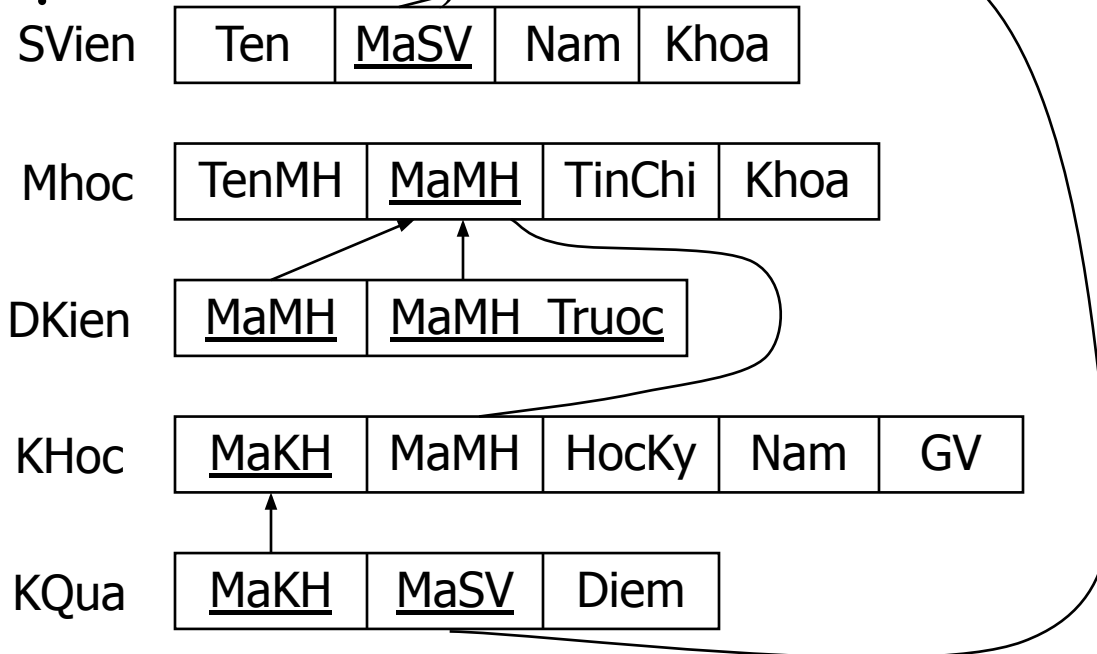


Ví dụ mô hình quan hệ



Lược đồ

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



Thẻ 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	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

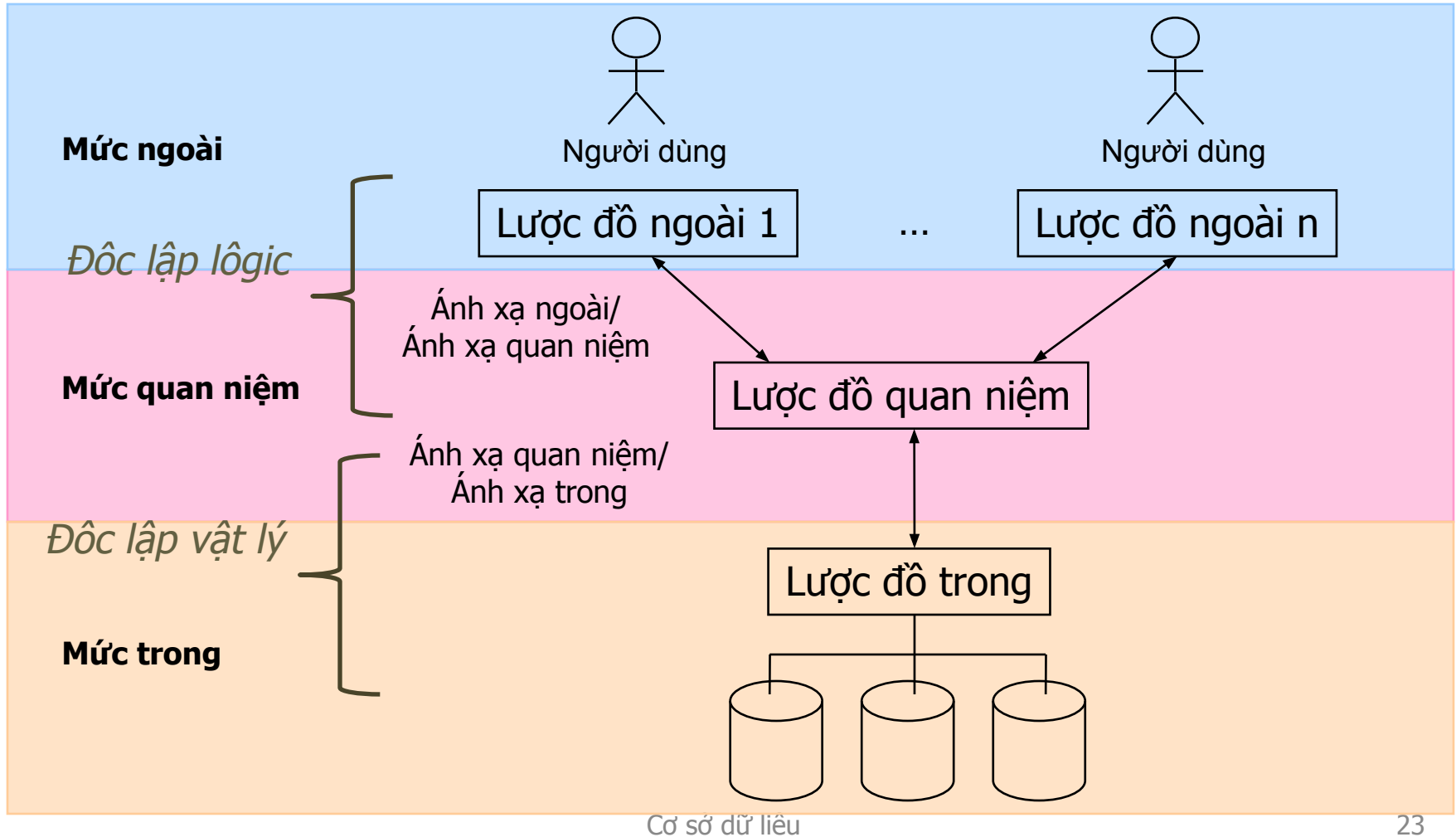
SVien	Ten	MaSV	Nam	Khoa
	Son	17	1	CNTT
	Bao	8	2	CNTT

DKien	MaMH	MaMH_Truoc
	COSC3380	COSC3320
	COSC3380	MATH2410
	COSC3320	COSC3380

1.6 KIẾN TRÚC 3 MỨC (LƯỢC ĐỒ) CỦA HỆ CSDL

Kiến trúc 3 lược đồ

Kiến trúc ba mức - theo ANSI/SPARC



Các mức mô tả CSDL

- **Mức trong** (lược đồ trong)
 - Mô tả cấu trúc lưu trữ vật lý của CSDL trong máy tính
- **Mức quan niệm** (lược đồ quan niệm)
 - Cách nhìn của một nhóm người dùng về CSDL
 - Mô tả dữ liệu gì được lưu trong CSDL và quan hệ của chúng
- **Mức ngoài** (lược đồ ngoài)
 - Cung cấp các khung nhìn (view) của người dùng
 - 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

External view 1

sNo	fName	lName	age	salary
-----	-------	-------	-----	--------

External view 2

staffNo	lName	branchNo
---------	-------	----------

Conceptual level

staffNo	fName	lName	DOB	salary	branchNo
---------	-------	-------	-----	--------	----------

Internal level

```
struct STAFF {  
    int staffNo;  
    int branchNo;  
    char fName [15];  
    char lName [15];  
    struct date dateOfBirth;  
    float salary;  
    struct STAFF *next;  
};  
index staffNo; index branchNo;
```

/* pointer to next Staff record */
/* define indexes for staff */

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

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 hợp lệ
 - 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

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

Các tính năng khác

- Chuẩn hóa
 - Cho phép DBA định nghĩa và bắt buộc áp dụng một chuẩn thống nhất cho mọi người dùng
- 유연 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

Phân loại HQT CSDL

- ***Theo mô hình :***
 - mô hình mạng,
 - mô hình phân cấp,
 - mô hình quan hệ,
 - mô hình đối tượng,...
- ***Theo số người dùng:***
 - một người dùng,
 - đa người dùng
- ***Theo sự phân tán của dữ liệu:***
 - tập trung,
 - phân tán (thuần nhất hoặc không thuần nhất)

1.8 NGÔN NGỮ CSDL

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 );
```

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”.

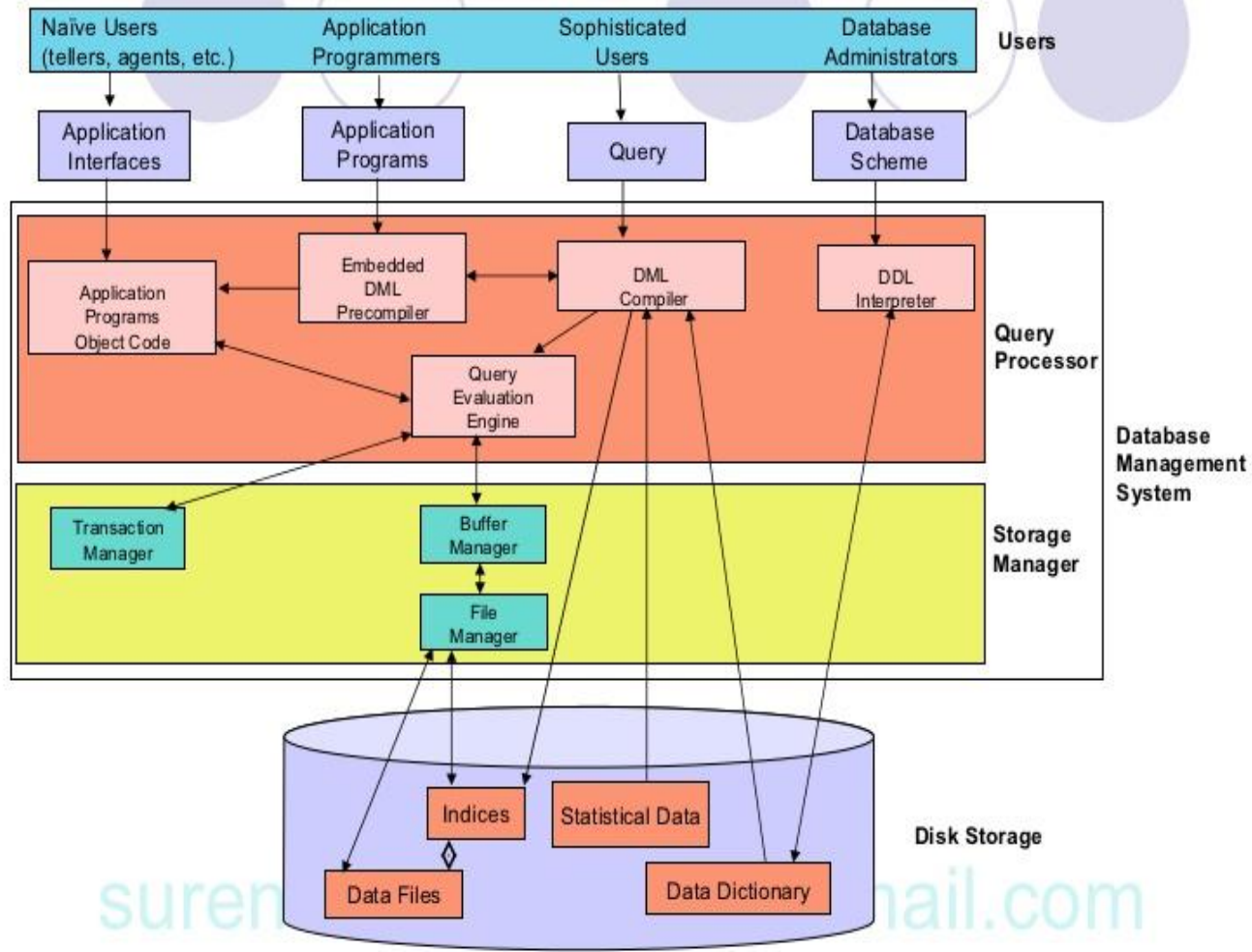
Ví dụ

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

Ngôn ngữ điều khiển giao dịch, Ngôn ngữ điều khiển dữ liệu

- ***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.
- ***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.

Database Management System Structure



Câu hỏi

1. Phân biệt các khái niệm: dữ liệu, cơ sở dữ liệu, hệ quản trị cơ sở dữ liệu, hệ cơ sở dữ liệu.
2. Thế nào là độc lập dữ liệu, độc lập logic và độc lập vật lý.
3. Khái niệm mô hình, lược đồ, thể hiện của cơ sở dữ liệu.
4. Các tính năng chính của HQT CSDL