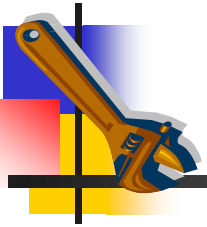




Chương 1



GIỚI THIỆU VỀ CƠ SỞ DỮ LIỆU





Nội dung

- 1.1. Các khái niệm và các định nghĩa cơ bản
- 1.2. Cơ sở dữ liệu
- 1.3. Hệ quản trị cơ sở dữ liệu
- 1.4. Hệ cơ sở dữ liệu
- 1.5. Các giai đoạn phát triển của CSDL
- 1.6. Cách tiếp cận cơ sở dữ liệu
- 1.7. Quá trình phát triển cơ sở dữ liệu
- 1.8. Kiến trúc cơ sở dữ liệu
- 1.9. Kiến trúc cơ sở dữ liệu client/server 3 tầng

Trần thị Sơn; 19 tuổi; quê quán: Bình Định

Đơn vị hành chánh

Sinh viên

Thái Lan; dân số 60 triệu người

Tỉnh Bình Thuận; cách Tp.HCM 150km

Quốc gia

Lào; dân số 30 triệu người

Trung Quốc; 1,2 tỷ người

Tỉnh Long An; Cách Tp.HCM 50km

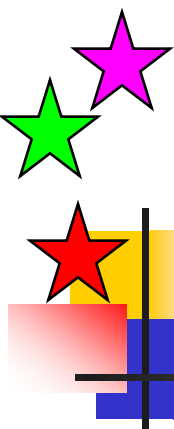
Nguyễn thị Mến; 20 tuổi; Quê quán: Tiền Giang

Tỉnh Đồng Nai; cách Tp.HCM 30km

Việt nam; dân số 80 triệu người

Lê Văn Lúa; 21 tuổi; quê quán: Long An

Trần Thị Kim Chi



Các khái niệm và các định nghĩa cơ bản

Data (dữ liệu) và information (thông tin)

- **Data:** sự biểu diễn của các đối tượng và sự kiện (văn bản, hình ảnh, âm thanh,...) được ghi nhận, có ý nghĩa không rõ ràng và được lưu trữ trên các phương tiện của máy tính.
 - Dữ liệu có cấu trúc: số, ngày, chuỗi ký tự, ...
 - Dữ liệu không có cấu trúc: hình ảnh, âm thanh, đoạn phim, ...
 - **Information:** dữ liệu đã được xử lý để làm tăng sự hiểu biết của người sử dụng.
- Phân biệt giữa data và information??



Các khái niệm và các định nghĩa cơ bản

Dữ liệu

1	10273	Nguyễn Văn Hoà	CDTH7	20
2	00298	Nguyễn Minh Tâm	CDTH7	19
...				
151	50542	Hồ Xuân Phương	TCTH33	18
152	50075	Lê Việt Dũng	CNTH34	20

Thông tin: dữ liệu trong ngữ cảnh

STT	Mã sinh viên	Họ và tên sinh viên	Lớp	Tuổi
1	10273	Nguyễn Văn Hoà	CDTH7	20
2	00298	Nguyễn Minh Tâm	CDTH7	19
...				
151	50542	Hồ Xuân Phương	TCTH33	18
152	50075	Lê Việt Dũng	CNTH34	20

Database System

Các khái niệm và các định nghĩa cơ bản

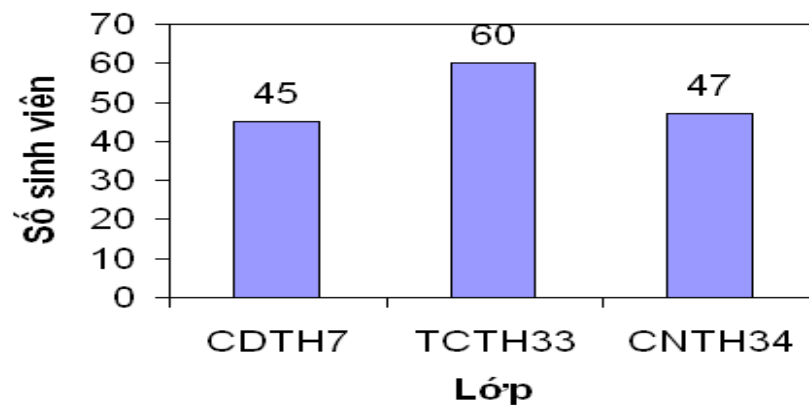
■ Ví dụ: Dữ liệu

Baker, Kenneth D.	324917628
Doyle, Joan E.	476193248
Finkle, Clive R.	548429344
Lewis, John C.	551742186
McFerran, Debra R.	409723145

Thông tin

Class Roster			
Course:	MGT 500	Semester:	Spring 2010
	Business Policy		
Section:	2		
Name	ID	Major	GPA
Baker, Kenneth D.	324917628	MGT	2.9
Doyle, Joan E.	476193248	MKT	3.4
Finkle, Clive R.	548429344	PRM	2.8
Lewis, John C.	551742186	MGT	3.7
McFerran, Debra R.	409723145	IS	2.9
Sisneros, Michael	392416582	ACCT	3.3

Thông Tin: Dữ liệu được tổng hợp / xử lý



**DỮ LIỆU
(DATA)**

XỬ LÝ

**THÔNG TIN
(INFORMATION)**



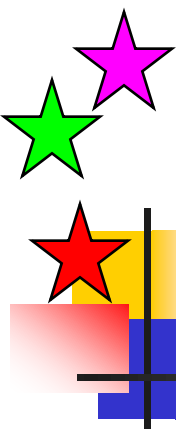
Các khái niệm và các định nghĩa cơ bản

- **Siêu dữ liệu (metadata):** mô tả các tính chất hoặc các đặc điểm của dữ liệu người dùng cuối. Các đặc tính là định nghĩa dữ liệu, cấu trúc dữ liệu, qui tắc/ràng buộc.

Siêu dữ liệu cho Sinh_viên

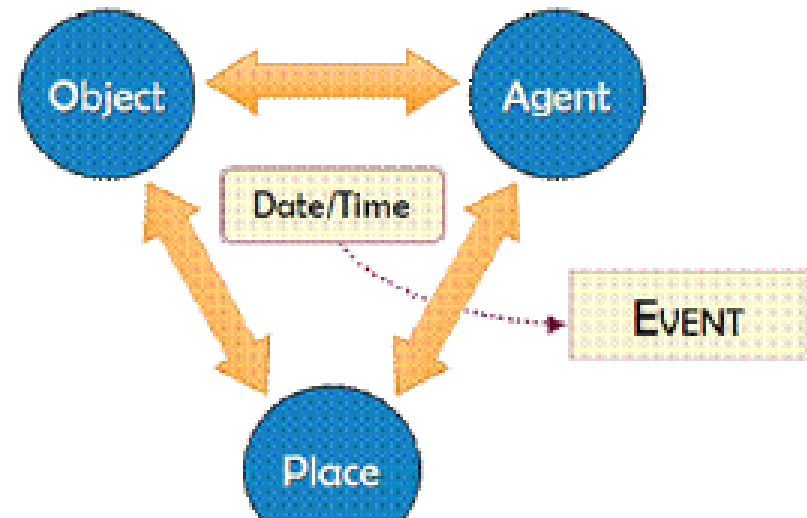
Siêu dữ
liệu

Data Item		Value	
Name	Type	Length	Description
MaSV	Char	5	Mã sinh viên
Hoten	Char	30	Họ và tên sinh viên
Lop	Char	7	Lớp học
Tuoi	smallint		Tuổi
02345	Tran Hoang		DH1H9B 19



Mô hình hóa dữ liệu

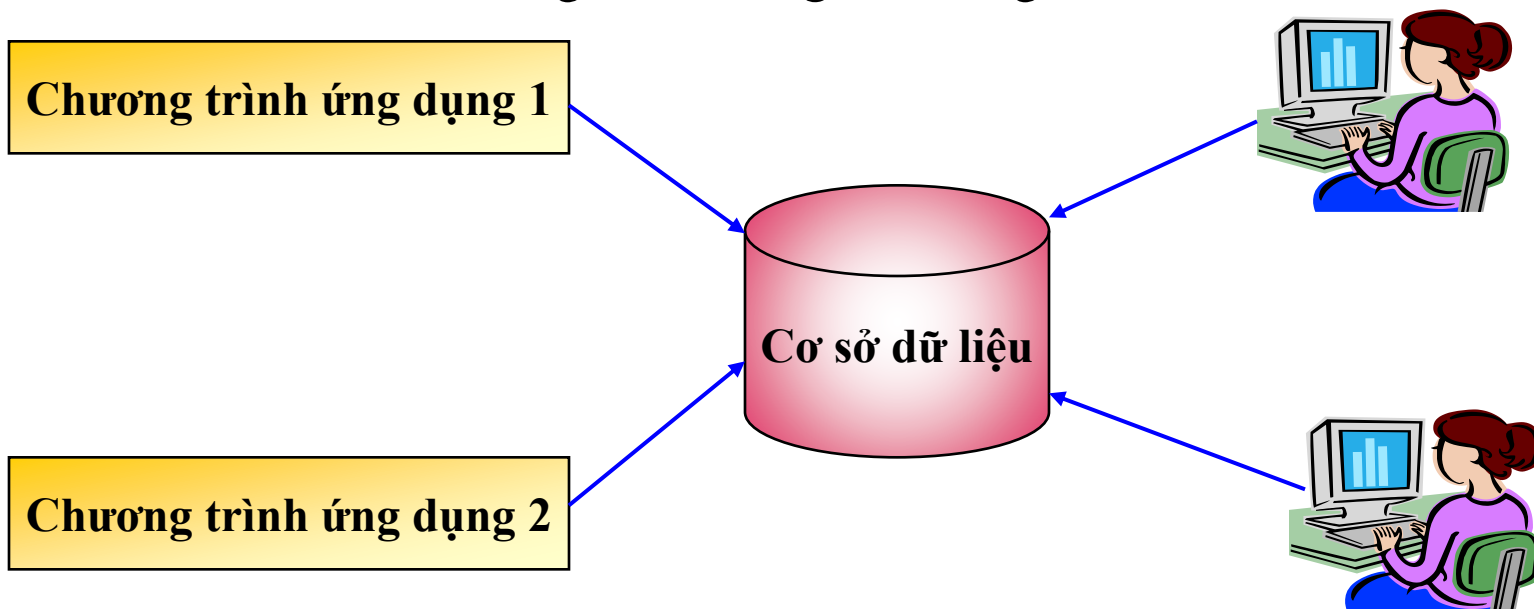
- Là các hệ thống đồ họa (graphical system) được dùng để nắm bắt bản chất và mối quan hệ giữa dữ liệu
- Các mô hình dữ liệu thông dụng:
 - Mô hình thực thể kết hợp
 - Mô hình dữ liệu quan hệ
 - Mô hình hướng đối tượng



Database System

Cơ sở dữ liệu là gì?

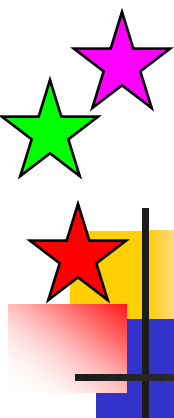
- ▶ Cơ sở dữ liệu là một tập hợp dữ liệu có liên quan luận lý với nhau chứa thông tin về một tổ chức nào đó, được lưu trữ trên máy tính theo một hệ thống và được dùng chung đáp ứng nhu cầu khai thác thông tin của người dùng.



**Các hệ thống chương trình
ứng dụng khai thác**

Trần Thi Kim Chi

Người sử dụng khai thác 9



Ví dụ về một CSDL

SINHVIEN

<u>MASV</u>	TEN	MALOP	TENLOP	SISO	<u>MAMH</u>	TENMH	DIEM
TCTH01	Sơn	TCTHA	TCTH32A	80	THVP	Nhập môn TH	8
TCTH02	Bảo	TCTHB	TCTH32B	65	CSDL	Nhập môn TH	6
TCTH03	Hà	TCTHA	TCTH32C	82	CTDL	Nhập môn TH	7
TCTH01	Sơn	TCTHA	TCTH32A	80	THVP	Cấu trúc dữ liệu	8
TCTH02	Bảo	TCTHB	TCTH32B	65	CSDL	Cấu trúc dữ liệu	6
TCTH03	Hà	TCTHA	TCTH32C	82	CTDL	Cấu trúc dữ liệu	7

Ví dụ về một CSDL

SINHVIEN

<u>MASV</u>	TEN	MALOP
TCTH01	Sơn	TCTHA
TCTH02	Bảo	TCTHB
TCTH03	Trang	TCTHA

LOP

<u>MALOP</u>	TENLOP	KHOA
TCTHA	TCTH32A	CNTT
TCTHB	TCTH32B	CNTT
TCTHC	TCTH32C	CNTT

KETQUA

<u>MASV</u>	MAMH	DIEM
TCTH01	THVP	8
TCTH01	CSDL	6
TCTH01	CTDL	7
TCTH02	THVP	9
TCTH02	CSDL	8
TCTH03	THVP	10

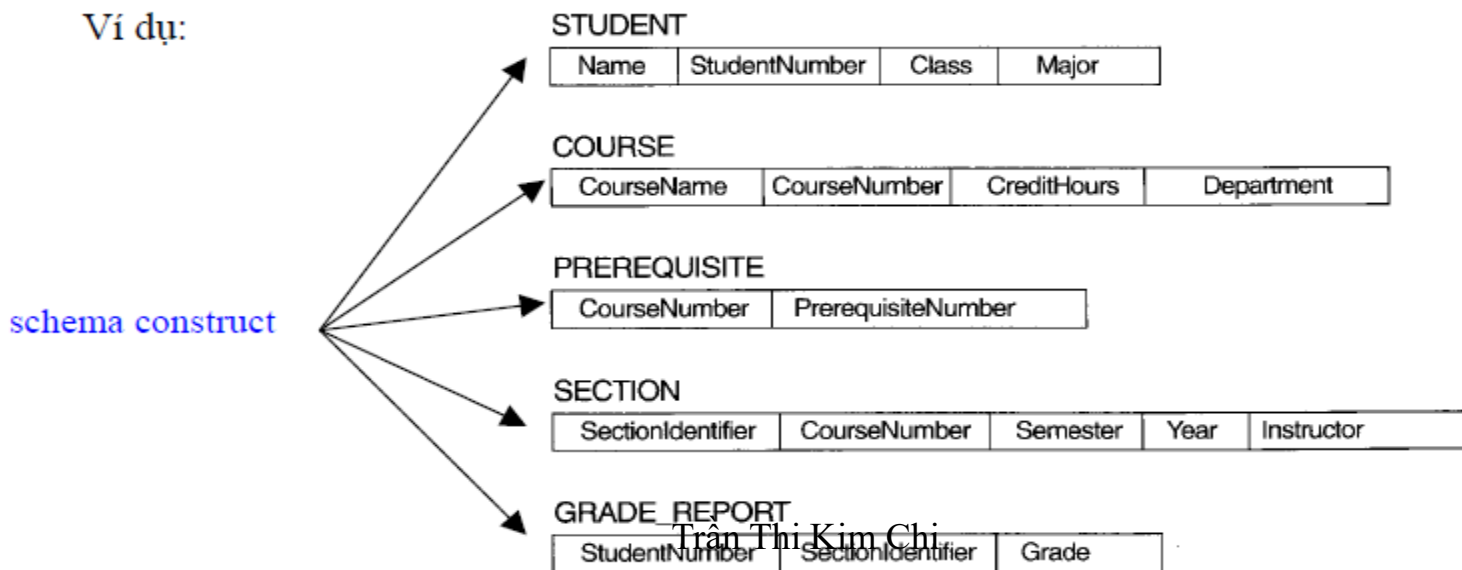
MONHOC

<u>MAMH</u>	TENMH	TINCHI
THVP	Nhập môn TH	4
CSDL	Cấu trúc dữ liệu	4
CTDL	Toán rời rạc	3

Lược đồ (Schema) cơ sở dữ liệu

- **Lược đồ cơ sở dữ liệu (Database Schema):** là biểu diễn của cơ sở dữ liệu, bao gồm cấu trúc cơ sở dữ liệu và những ràng buộc trên dữ liệu.
- **Sơ đồ của lược đồ cơ sở dữ liệu (Schema Diagram):** Là lược đồ cơ sở dữ liệu được biểu diễn thông qua sơ đồ.

Ví dụ:





Lược đồ (Schema) cơ sở dữ liệu

- **Thể hiện cơ sở dữ liệu (Database Instance):** Là dữ liệu thực sự được lưu trữ trong cơ sở dữ liệu ở thời điểm hiện tại. Database Instance cũng được gọi là trạng thái của cơ sở dữ liệu (**database state**)

Ví dụ:

STUDENT	Name	StudentNumber	Class	Major
	Smith	17	1	CS
	Brown	8	2	CS

COURSE	CourseName	CourseNumber	CreditHours	Department
	Intro to Computer Science	CS1310	4	CS
	Data Structures	CS3320	4	CS
	Discrete Mathematics	MATH2410	3	MATH
	Database	CS3380	3	CS

SECTION	SectionIdentifier	CourseNumber	Semester	Year	Instructor
	85	MATH2410	Fall	98	King
	92	CS1310	Fall	98	Anderson
	102	CS3320	Spring	99	Knuth
	112	MATH2410	Fall	99	Chang
	119	CS1310	Fall	99	Anderson
	135	CS3380	Fall	99	Stone

Trần Thị Kim Chi



Đặc điểm của Cơ Sở Dữ Liệu

- **Tính tự mô tả:**

- Hệ CSDL không chỉ chứa bản thân CSDL mà còn chứa định nghĩa đầy đủ (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
- Các chương trình ứng dụng có thể truy xuất đến nhiều CSDL nhờ thông tin cấu trúc được lưu trữ trong catalog



Đặc điểm của Cơ Sở 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



Đặc điểm của Cơ Sở Dữ Liệu

- **Persistent – Thường trú:**

- Dữ liệu được lưu trữ trong bộ nhớ ổn định như đĩa cứng, server. Khi dữ liệu không cần dùng nữa thì có thể xoá hay sao lưu lại.

- **Independent - Interrelated – Độc lập - Tương tác:**

- Dữ liệu được lưu trữ như những đơn vị riêng biệt và được kết nối với nhau để tạo 1 tổng thể chung
- Database vừa chứa thực thể và cả mối quan hệ giữa các thực thể

- **Shared – Chia sẻ và xử lý tiến trình:**

- Database có thể có nhiều người dùng và nhiều người dùng có thể sử dụng cùng 1 database tại cùng 1 thời điểm.
- Cho phép các tiến trình đồng thời bên trong DBMS (Bài toán đồng thời (concurrency problem))



Đặc điểm của Cơ Sở Dữ Liệu

■ Tính trừu tượng (Data Abstraction)

- 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
- Tính trừu tượng hóa giúp bảo đảm tính độc lập chương trình và dữ liệu.
- Trừu tượng hóa dữ liệu – Mô hình dữ liệu
 - Đối tượng
 - Thuộc tính của đối tượng
 - Mối liên hệ

■ Hỗ trợ nhiều cách nhìn cơ sở dữ liệu

- Đáp ứng yêu cầu đa người dùng, giúp người dùng có thể xem dữ liệu theo những yêu cầu khác nhau



Chức năng chính của CSDL

- **Định nghĩa cấu trúc:** khai báo tập tin hay quan hệ + kiểu dữ liệu
- **Cập nhật dữ liệu:** thêm, sửa, xóa dữ liệu
 - Nhập dữ liệu liên quan đến sinh viên , môn học, điều kiện, ...
- **Truy vấn:** xem dữ liệu
 - Cần phải học môn học nào trước môn học CSDL?
 - Liệt kê sinh viên có điểm lớn hơn 8 trong môn CSDL năm 1997



Chức năng chính của CSDL

- **Báo cáo:** in ra bảng điểm, theo định dạng văn bản, tên sinh viên, tên môn học, học kỳ, năm học, điểm
- **Thêm, xóa, sửa dữ liệu**
 - Tạo một học phần mới
 - Nhập điểm cho sinh viên thi môn Nhập môn tin học
- **Cập nhật cấu trúc, lược đồ**
 - Tạo quan hệ mới chứa thông tin về giáo viên
 - Thêm thuộc tính địa chỉ vào quan hệ SVIEN



Ưu điểm của Cơ sở Dữ liệu

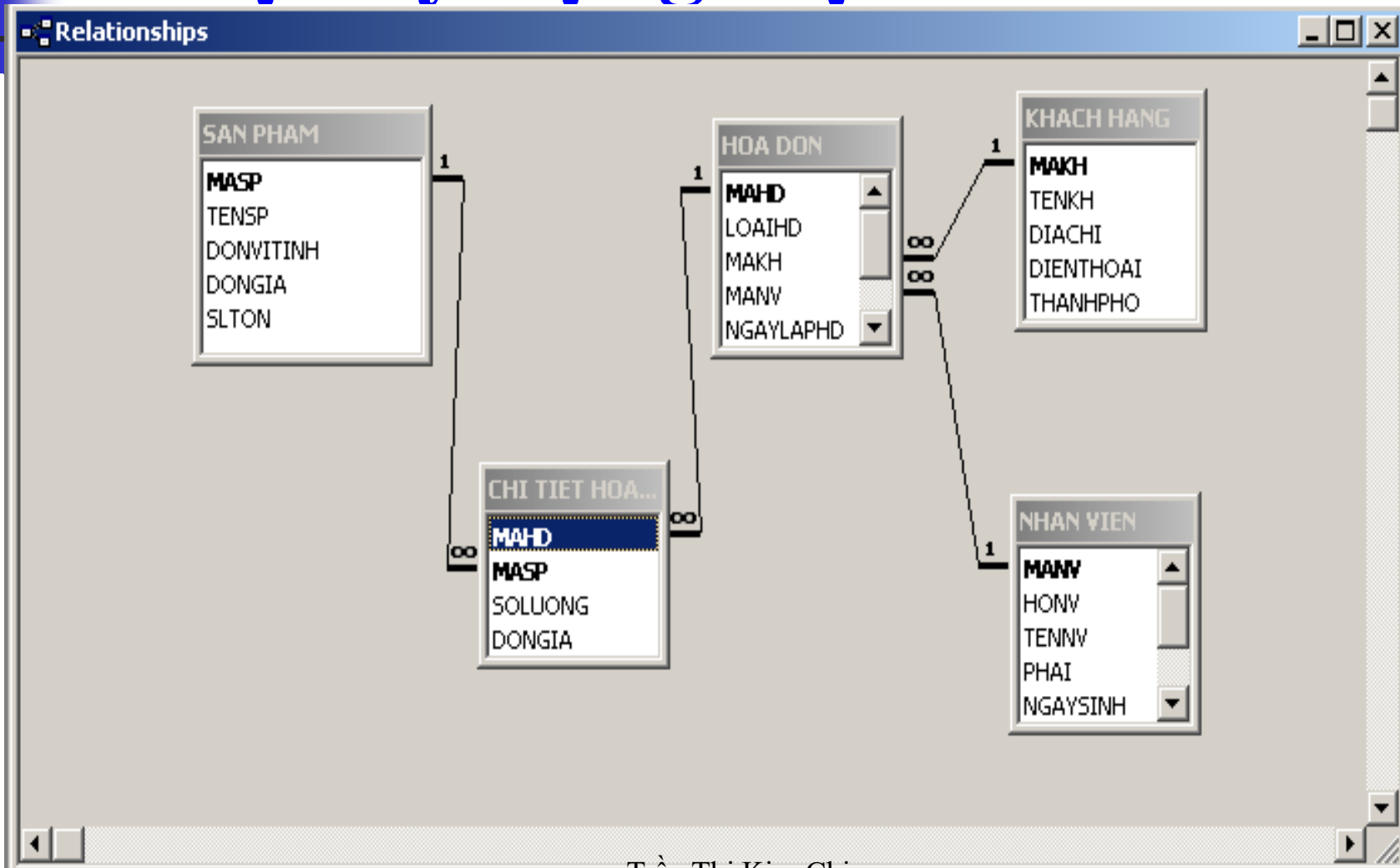
- Giảm sự trùng lặp thông tin xuống mức thấp nhất do đó bảo đảm được tính nhất quán và toàn vẹn dữ liệu.
- Độc lập dữ liệu-chương trình. Đảm bảo **dữ liệu có thể được truy xuất theo nhiều cách khác nhau.**
- **Khả năng chia sẻ thông tin** cho nhiều người sử dụng và nhiều ứng dụng khác nhau.
- Nâng cao tính nhất quán (data consistency)/toàn vẹn dữ liệu (data integrity).
- Giảm chi phí bảo trì chương trình.
- Bảo mật (security)
- Chép lưu (backup) và phục hồi (recovery)
- Điều khiển tương tranh (concurrency control).



Ví dụ xây dựng một CSDL

- Để quản lý hóa đơn bán hàng ta cần xây dựng một CSDL gồm các bảng sau :
- **SanPham:** Mô tả các thông tin về sản phẩm như Masp, Tensp, Donvitinh, DongiaMua, Slton.
- **KhachHang:** Mô tả các thông tin về khách hàng như Makh, Tenkh, Diachi, DienThoai.
- **Nhanvien:** Mô tả các thông tin về nhân viên như Manv, Honv, Tennv, Phai, Ngaysinh, Diachi, DienThoai, Hinh.
- **Hoá đơn:** Mô tả các thông tin về hóa đơn như Mahd, LoaiHD, Makh, Manv, NgaylapHD, NgayGiaoNhanHang, DienGiai.
- **Chitiethoadon:** Mô tả các thông tin chi tiết của từng hóa đơn như Mahd, Masp, Soluong, DongiaBan.

Ví dụ xây dựng một CSDL





Các đối tượng dùng CSDL

- **Các chuyên viên tin học (người lập trình-Database Designer)** biết khai thác CSDL, đây là những người có thể xây dựng các ứng dụng khác nhau phục vụ cho những yêu cầu khác nhau trên CSDL.
- **Những người sử dụng không chuyên (End User)** về lĩnh vực tin học và CSDL, do đó CSDL cần có các công cụ để giúp cho người sử dụng không chuyên có thể khai thác hiệu quả CSDL
- **Những người quản trị CSDL(Database Administrator):** là những người hiểu biết về tin học, các hệ quản trị CSDL và hệ thống máy tính. Họ là người tổ chức CSDL (khai báo cấu trúc CSDL, ghi nhận các yêu cầu bảo mật cho các dữ liệu cần bảo vệ...), do đó họ cần phải nắm rõ các vấn đề về kỹ thuật để có thể phục hồi dữ liệu khi có sự cố. Họ là những người cấp quyền hạn khai thác CSDL nên họ có thể giải quyết các vấn đề tranh chấp dữ liệu

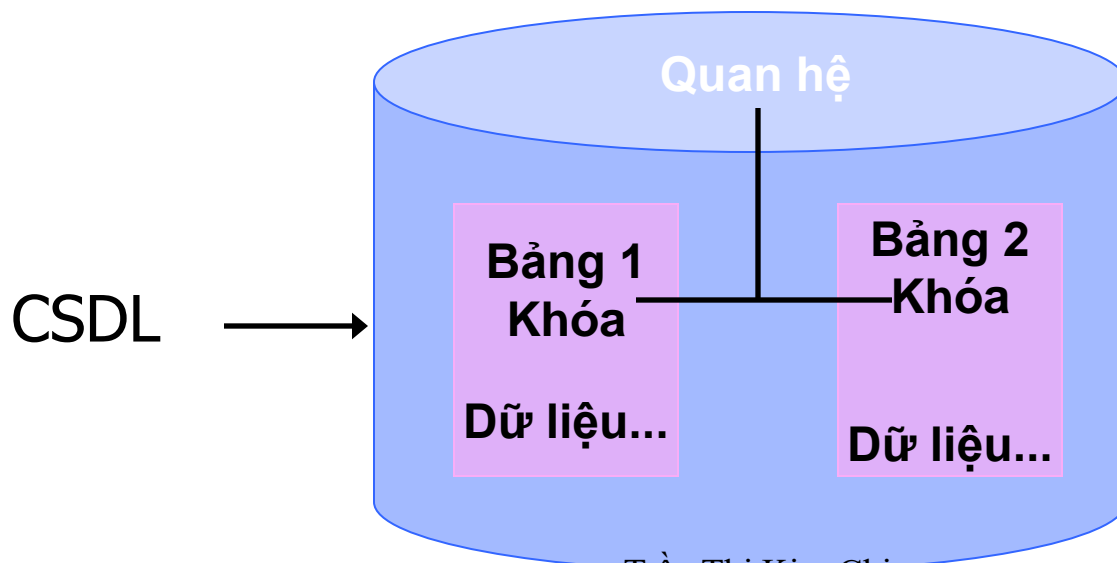


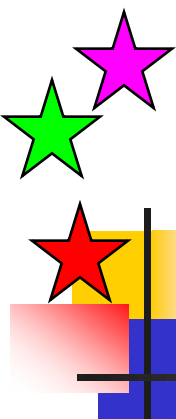
Hệ quản trị CSDL

- **Hệ quản trị cơ sở dữ liệu (Database Management System, viết tắt là DBMS):** là một tập chương trình hay một phần mềm giúp cho người sử dụng tạo ra, duy trì và khai thác CSDL.
- Hệ quản trị CSDL hỗ trợ khả năng lưu trữ, sửa chữa, xóa, xem và tìm kiếm thông tin trong CSDL.
- **Tính chất của các hệ quản trị CSDL (DBMS)**
 - Tính sẵn sàng, chia sẻ dữ liệu
 - Toàn vẹn, chuẩn hóa
 - Bảo mật
 - Độc lập, sao lưu dự phòng
 - Cung cấp nhiều giao diện, uyển chuyển, giảm thời gian
- Các hệ quản trị CSDL phổ biến như: Access, FoxPro, MySQL, SQL Server, Oracle, DB2

Hệ quản trị CSDL quan hệ (Relation Database Management System)

- Một hệ quản trị CSDL quan hệ (RDBMS) là một hệ quản trị CSDL được xây dựng trên mô hình quan hệ.
- Một CSDL quan hệ là một CSDL được chia nhỏ thành các đơn vị logic gọi là bảng, các bảng có quan hệ với nhau trong CSDL.





Sự khác nhau giữa DBMS và RDBMS

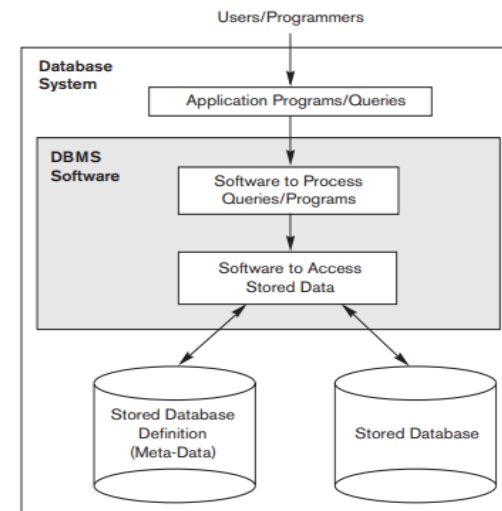
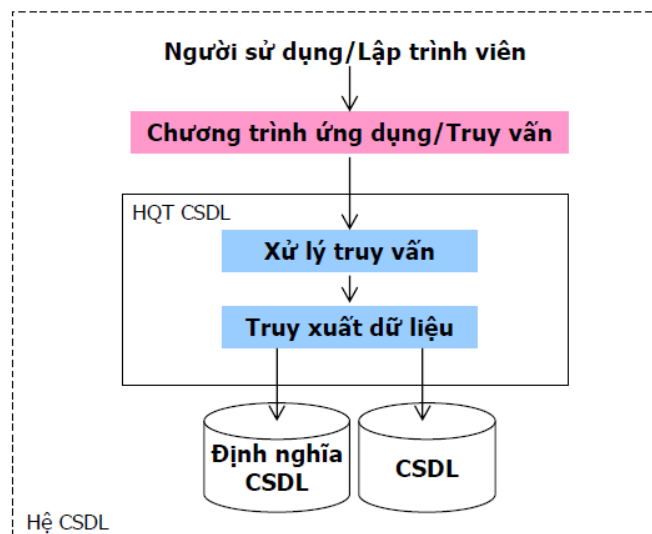
DBMS	RDBMS
Nó không cần chứa dữ liệu trong cấu trúc bảng và các mục dữ liệu không có quan hệ với nhau.	Trong RDBMS, cấu trúc bảng là bắt buộc và giữa chúng có mối quan hệ với nhau. Các mối quan hệ này làm cho người dùng có thể áp dụng và quản lý các quy tắc kinh doanh vào các đoạn mã.
DBMS chỉ lưu trữ và thao tác với số lượng dữ liệu nhỏ.	RDBMS có thể lưu trữ và thao tác trên một số lượng lớn dữ liệu.
DBMS kém đảm bảo hơn RDBMS.	RDBMS đảm bảo hơn DBMS.
Nó là 1 hệ thống đơn người dùng.	Nó là 1 hệ thống đa người dùng
Hầu hết các DBMS không hỗ trợ kiến trúc client/server.	Nó hỗ trợ kiến trúc client/server.

Hệ CSDL

- **Hệ cơ sở dữ liệu (hệ CSDL):** Người ta thường dùng thuật ngữ hệ cơ sở dữ liệu để chỉ một CSDL và hệ quản trị CSDL để truy cập CSDL đó thông qua các phần mềm ứng dụng.
- Mục đích chính của một hệ CSDL là cung cấp cho người dùng một cách nhìn trù tượng về dữ liệu (có nghĩa là hệ thống che dấu những chi tiết phức tạp về cách thức thao tác dữ liệu và bảo trì dữ liệu).



Hệ Cơ Sở Dữ Liệu

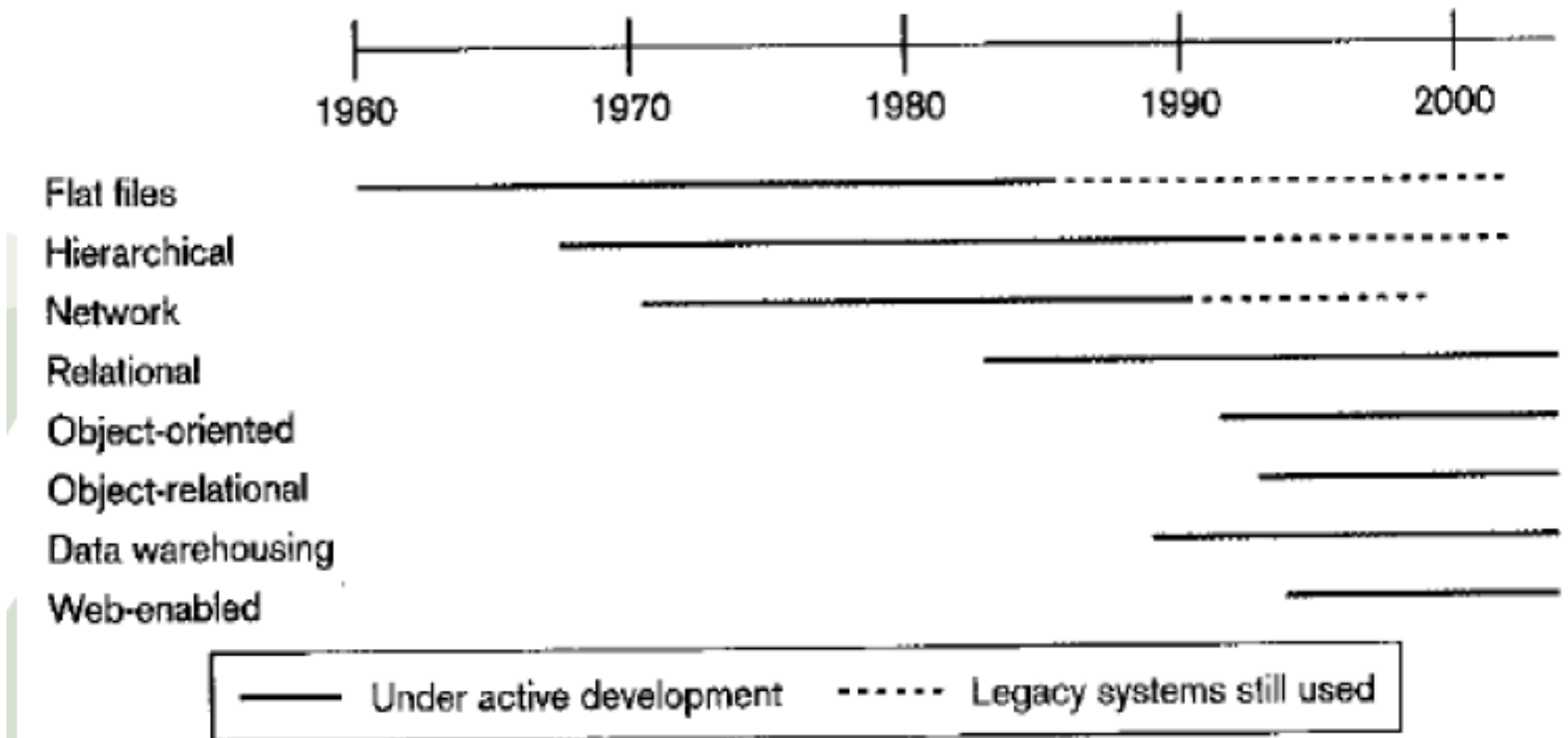




Các giai đoạn phát triển của hệ CSDL

- ❖ Hệ thống xử lý tập tin theo lối cũ
 - ❖ Hệ thống tập tin (*flat file*): 1960 - 1980
- ❖ Các hệ thống CSDL
 - ❖ Hệ CSDL phân cấp (*hierarchical*): 1970 - 1990
 - ❖ Hệ CSDL mạng (*network*): 1970 - 1990
 - ❖ Hệ CSDL quan hệ (*relational*): 1980 - nay
 - ❖ Hệ CSDL hướng đối tượng (*object-oriented*): 1990 - nay
 - ❖ Hệ CSDL đối tượng - quan hệ (*object-relational*): 1990 - nay

Các giai đoạn phát triển Database





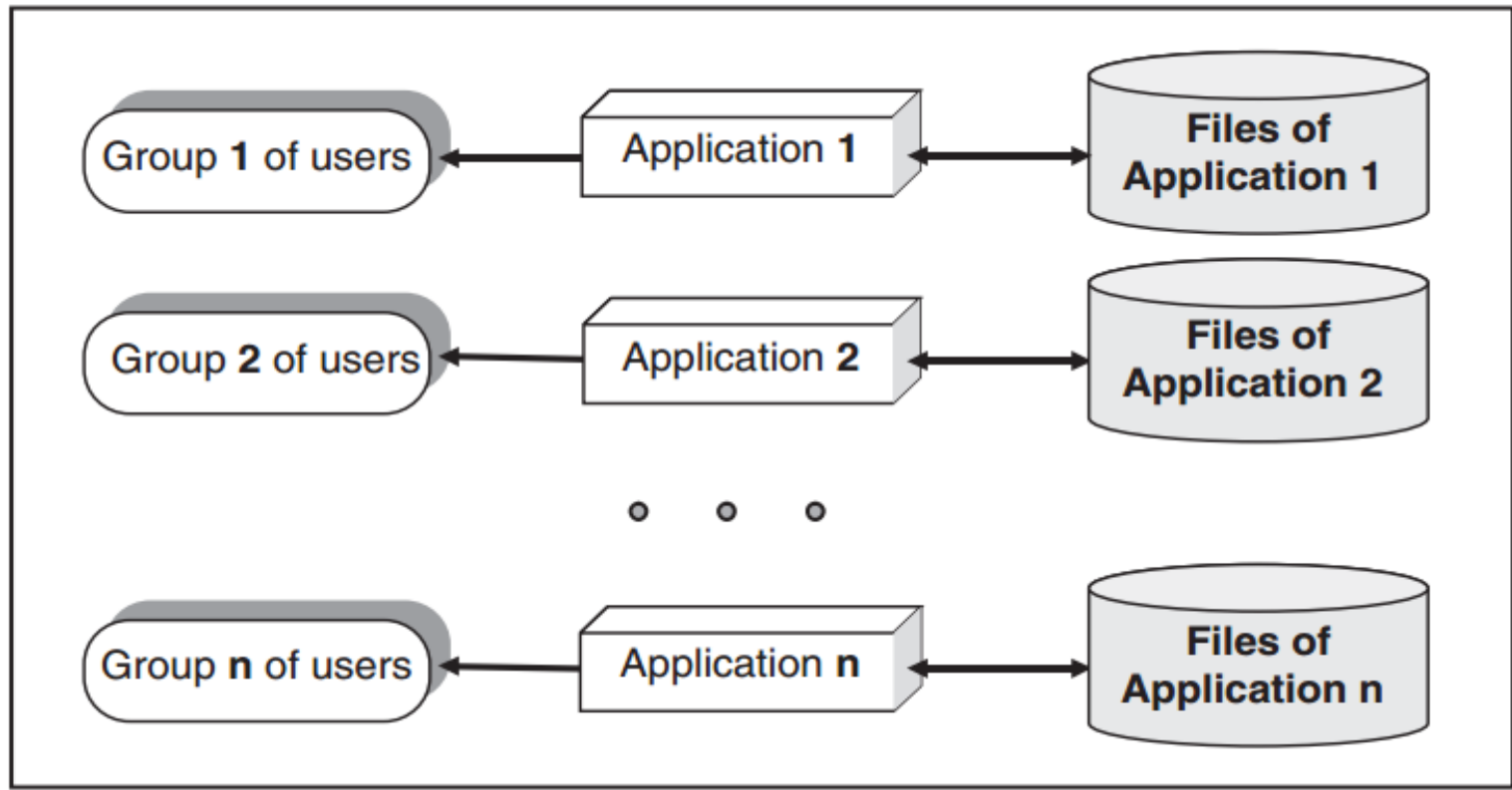
Hệ thống xử lý tập tin theo lối cũ

Hệ thống xử lý tập tin (*file processing system*)

- Các hệ thống mà trong đó dữ liệu được lưu trữ trên các file riêng biệt.
- Trong hệ thống này, một nhóm các file được lưu trữ riêng trên một máy tính và có thể được truy cập bởi một điều hành viên. Các tập tin dữ liệu được lưu trữ trong các thư mục (*folder*).



Hệ thống xử lý tập tin theo lối cũ



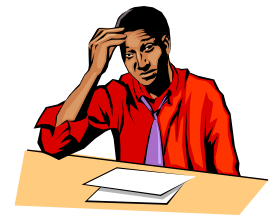
Hệ thống xử lý tập tin theo lối cũ

■ Ưu điểm

- Triển khai nhanh
- Khả năng đáp ứng nhanh, kịp thời

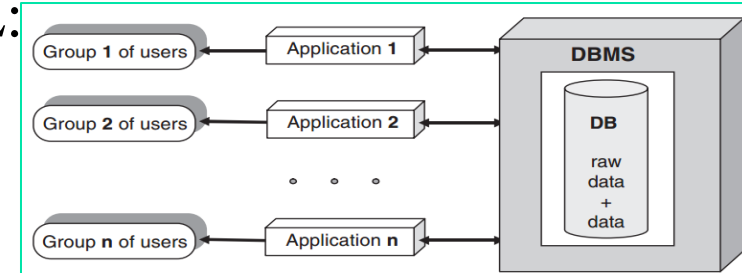
■ Nhược điểm

- Trùng lặp dữ liệu → lãng phí, dữ liệu không nhất quán
- Chi phí cao
- Chia sẻ dữ liệu kém
- Khó định dạng
- Có vấn đề về tính toàn vẹn dữ liệu



Các tiếp cận CSDL

- Các CSDL được dùng để lưu trữ dữ liệu một cách hiệu quả và có tổ chức sao cho quản lý được nhanh chóng và dễ dàng.
- Các ưu điểm của các hệ thống CSDL:
 - Giảm bớt sự dư thừa dữ liệu
 - Đảm bảo sự nhất quán dữ liệu
 - Dữ liệu lưu trữ có thể được chia sẻ, nhiều người dùng
 - Đảm bảo sự độc lập giữa dữ liệu và chương trình ứng dụng
 - Nhiều khung nhìn (multi-view) cho các đối người dùng khác nhau
 - Toàn vẹn dữ liệu
 - Bảo mật dữ liệu





Các mô hình CSDL

- Các CSDL có thể khác nhau về chức năng và mô hình của dữ liệu.
- Một mô hình dữ liệu mô tả một bộ chứa dữ liệu, xử lý và truy xuất dữ liệu từ bộ chứa.

Mô hình phân cấp



Mô hình mạng



Mô hình quan hệ

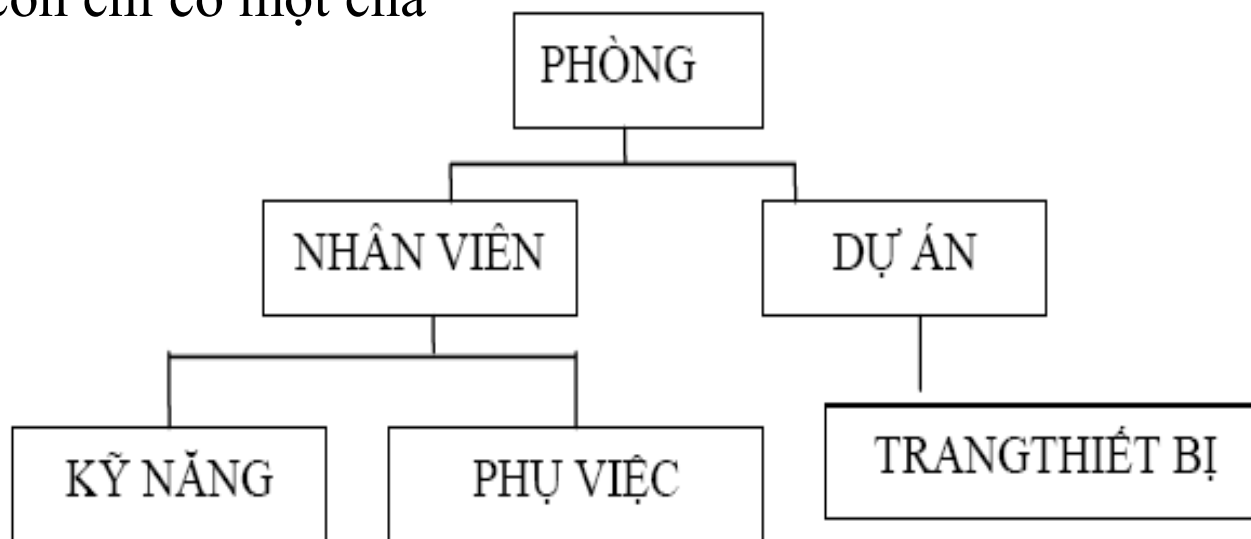




Mô hình phân cấp

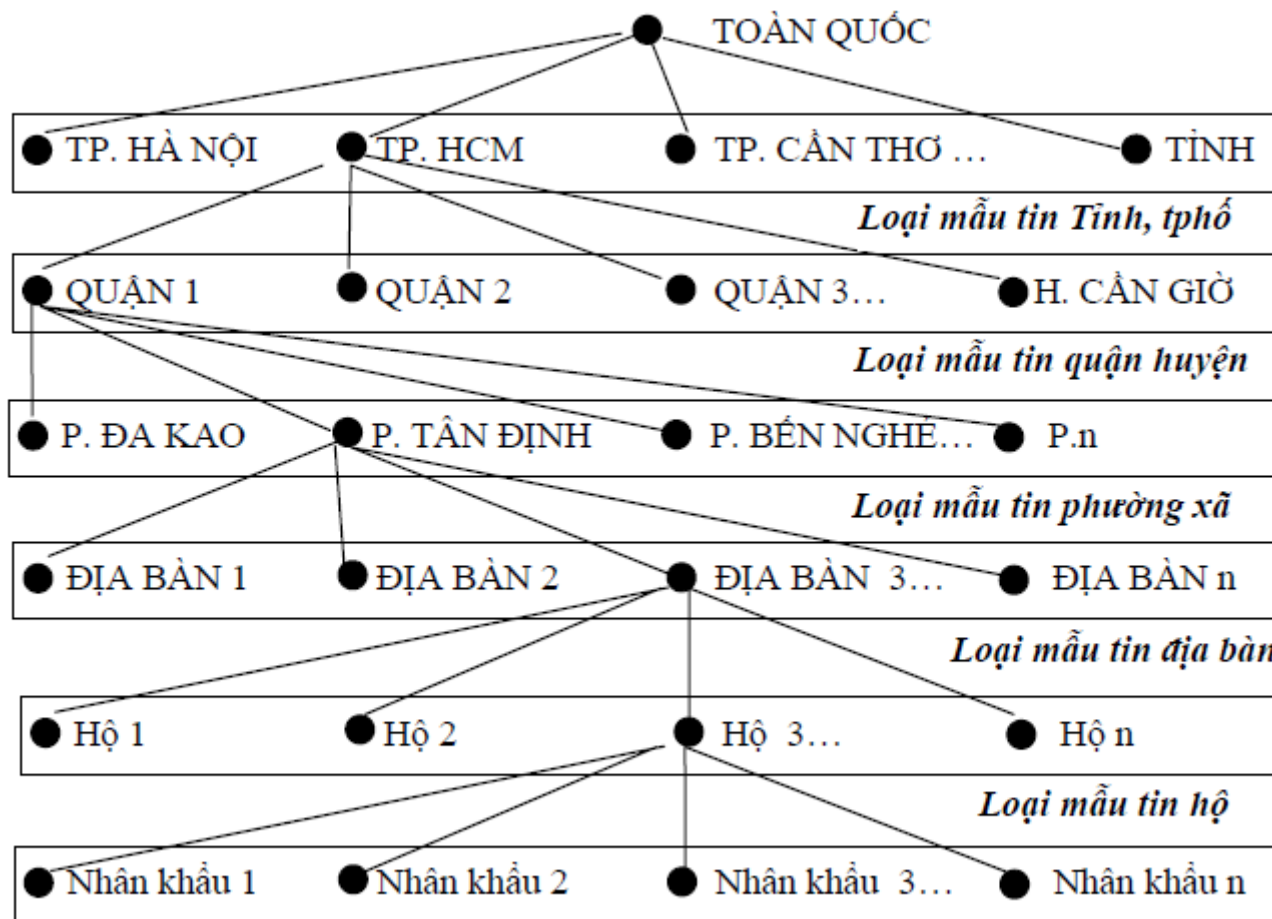
Hierarchical model

- Mô hình CSDL phân cấp được biểu diễn dưới dạng cây và các đỉnh của cây là các bản ghi. Các bản ghi liên kết với nhau theo mối quan hệ cha-con.
 - Một cha có nhiều con
 - Một con chỉ có một cha



Mô hình phân cấp

Hierarchical model





Mô hình phân cấp

Hierarchical model

- **Ưu điểm:**

- Thể hiện dễ dàng quan hệ 1-N.
- Việc phân chia dữ liệu dễ thể hiện, đảm bảo an toàn dữ liệu
- Tính độc lập của chương trình và các dữ liệu được đảm bảo

- **Nhược điểm:**

- Không thể hiện được mối quan hệ M-N
- Trong một hệ thống phân cấp, dữ liệu được tổ chức như trên dẫn đến khó sửa đổi dữ liệu.

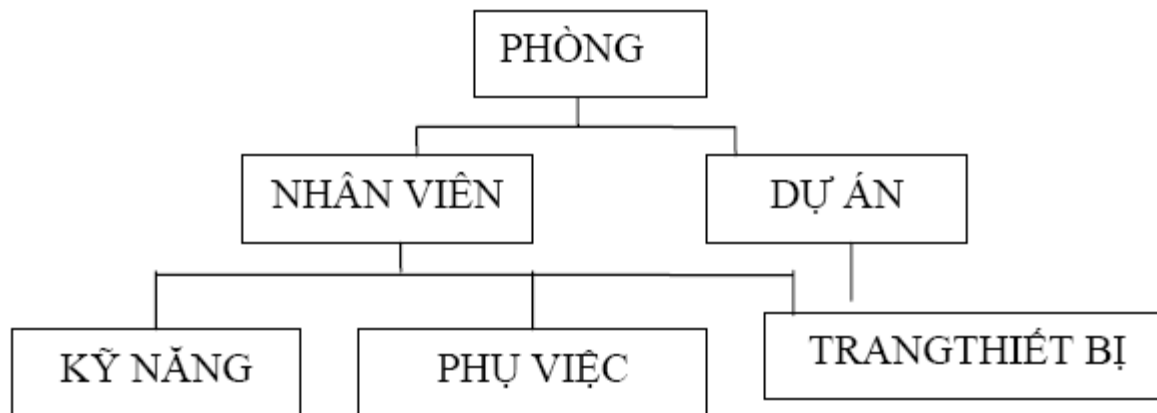


Mô hình mạng

Network model

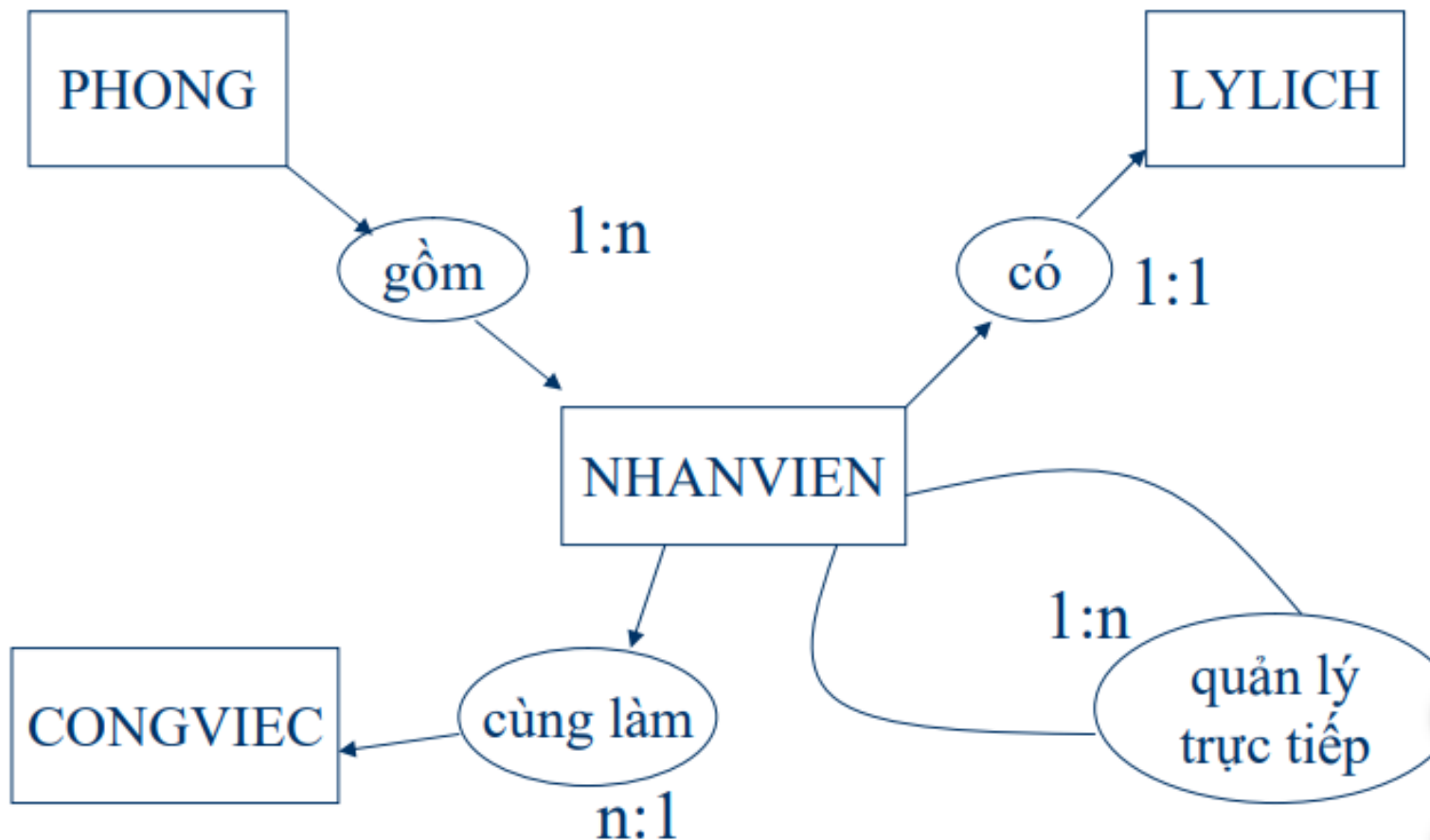
- Mô hình phân cấp là tập con của mô hình mạng.
- Mô hình mạng sử dụng kiến trúc cây phân cấp và cho phép các bảng con có thể có nhiều bảng cha.
- Dữ liệu được lưu trữ trong các bộ thay vì lưu trong định dạng cây phân cấp. Điều này giải quyết vấn đề dư thừa dữ liệu.

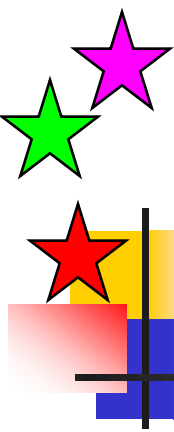
Ví dụ:



Mô hình mạng

Network model





Mô hình mạng

Network model

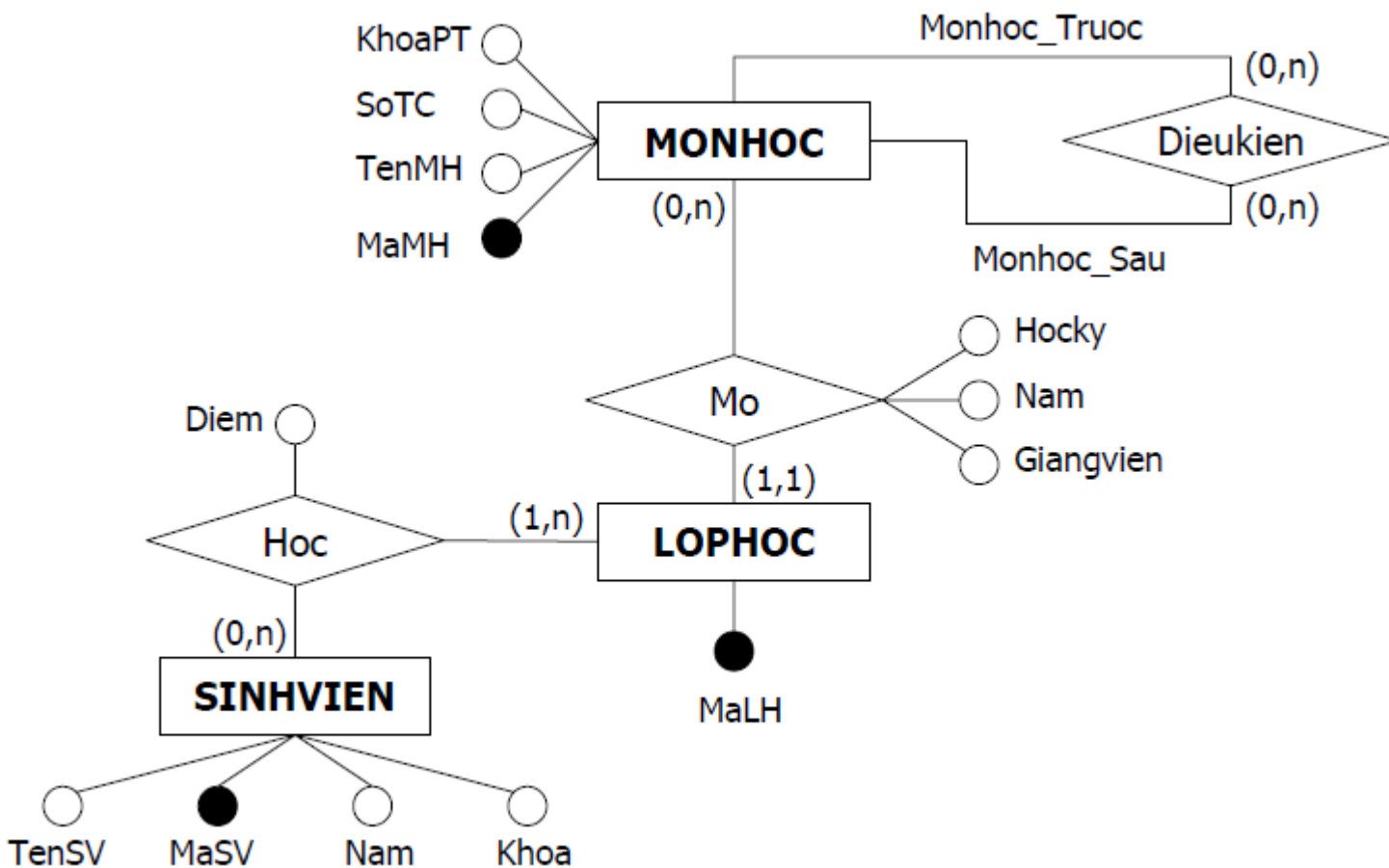
- **Ưu điểm:**

- Dễ thể hiện mối liên kết M-N
- Kiểu truy cập dữ liệu mềm dẻo hơn kiểu phân cấp
- Đơn giản, dễ sử dụng

- **Nhược điểm:**

- Không thích hợp trong việc biểu diễn CSDL lớn do hạn chế về khả năng diễn đạt ngữ nghĩa của dữ liệu, đặc biệt là các dữ liệu và mối liên hệ phức tạp của dữ liệu trong thực tế là rất hạn chế.
- Việc sửa đổi số liệu khó khăn.
- Với những lập trình viên, việc thiết kế CSDL khó.

Mô hình thực thể kết hợp



Mô hình quan hệ

Relational model

- Mô hình quan hệ không có các liên kết vật lý. Tất cả dữ liệu được lưu theo dạng bảng gồm các hàng và các cột.
- Các thao tác thực hiện trên các hàng của bảng.
- Sự kết nối giữa các bảng được mô tả logic bằng các giá trị được lưu trữ trong một trường chung.

SVIEN

<u>MASV</u>	TEN	MALOP
TCTH01	Sơn	TCTHA
TCTH02	Bảo	TCTHB
TCTH03	Trang	TCTHA

Tuple

Field

LOP

<u>MALOP</u>	TENLOP	SISO
TCTHA	TCTH32A	80
TCTHB	TCTH32B	65
TCTHC	TCTH32C	82

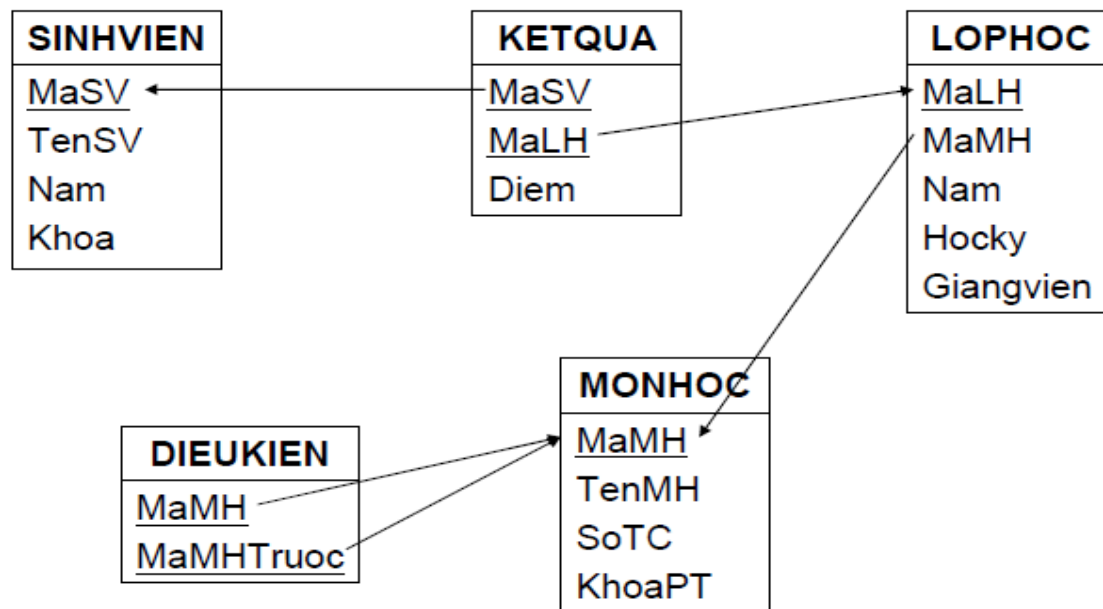
N

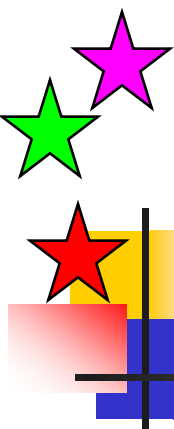
1

Mô hình quan hệ

Relational model

Mô hình quan hệ

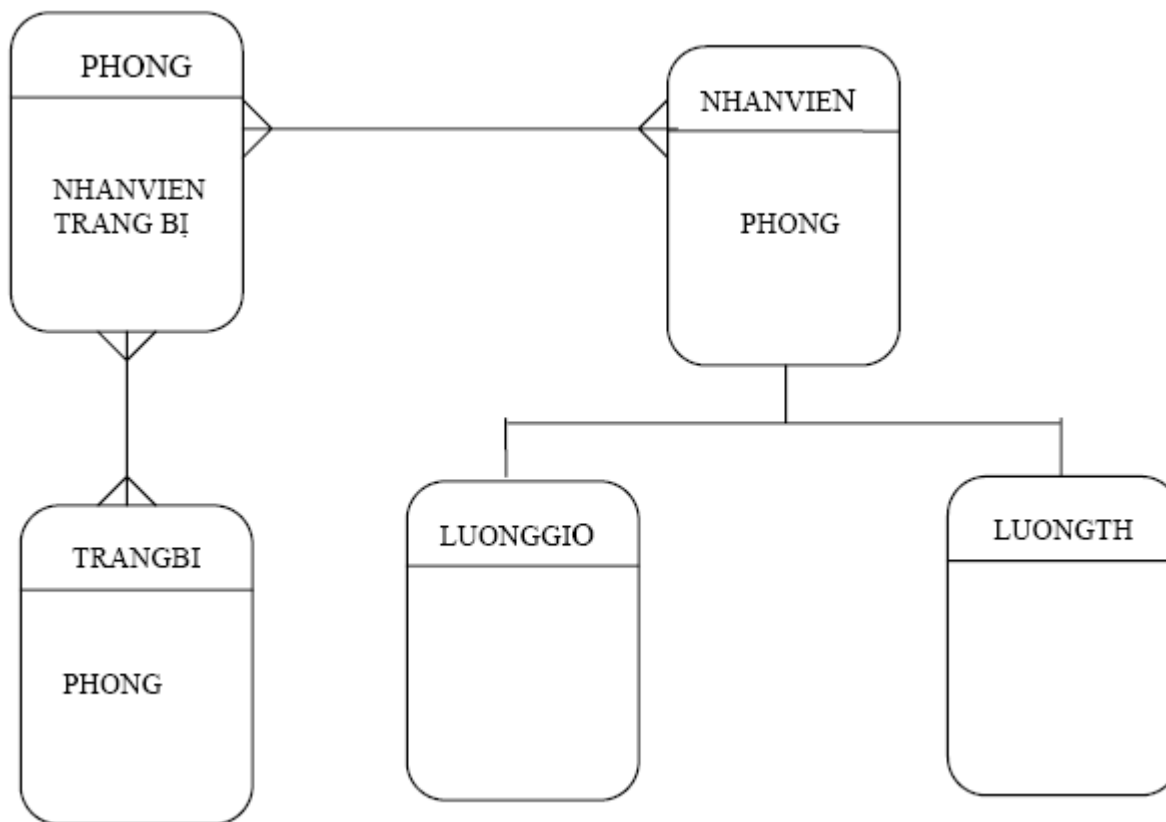




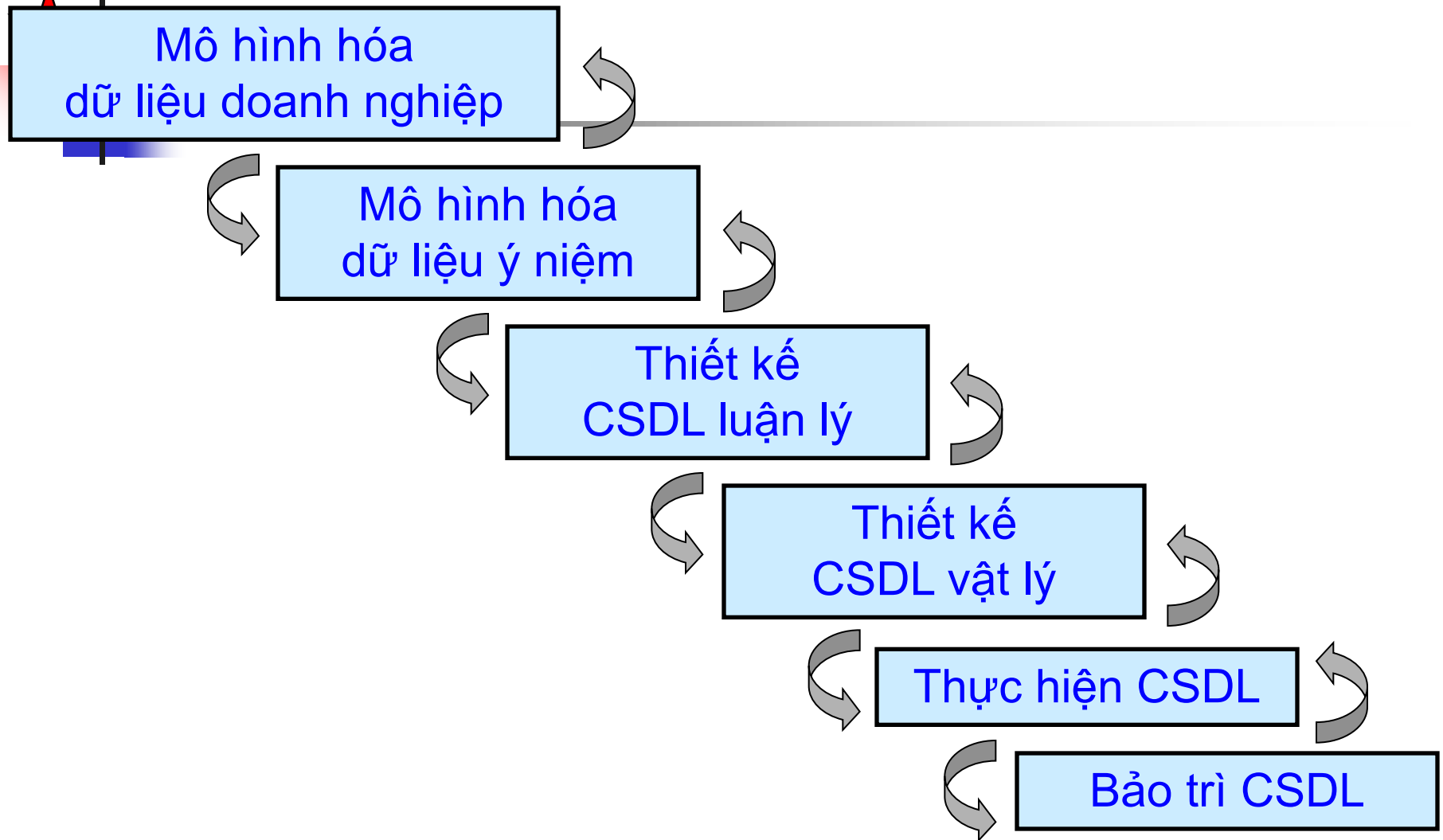
Mô hình hướng đối tượng

- Đối tượng có thể chứa các dữ liệu phức hợp như văn bản, hình ảnh, tiếng nói và hình ảnh động.
- Một đối tượng có thể yêu cầu hoặc xử lý dữ liệu từ một đối tượng khác bằng việc gửi đi một thông báo đến đối tượng đó.
- Mô hình hướng đối tượng biểu diễn một sơ đồ mới để lưu trữ và thao tác dữ liệu.
- Từ một đối tượng có thể sinh ra một đối tượng khác.

Mô hình hướng đối tượng



Các bước phát triển cơ sở dữ liệu



Các bước phát triển cơ sở dữ liệu

Mô hình hóa
dữ liệu doanh nghiệp

- Phân tích việc xử lý dữ liệu hiện tại.
- Phân tích toàn bộ các chức năng nghiệp vụ và các yêu cầu CSDL.

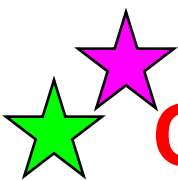
Mô hình hóa
dữ liệu ý niệm

Thiết kế
CSDL luận lý

Thiết kế
CSDL vật lý

Thực hiện CSDL

Bảo trì CSDL



Các bước phát triển cơ sở dữ liệu

Mô hình hóa
dữ liệu doanh nghiệp



Mô hình hóa
dữ liệu ý niệm



Thiết kế
CSDL luận lý



- Xác định phạm vi của các yêu cầu CSDL.
- Phân tích toàn bộ các yêu cầu dữ liệu của các chức năng nghiệp vụ dùng CSDL.
- Xây dựng sơ đồ liên kết thực thể bao gồm các thực thể và các mối liên kết giữa chúng, các thuộc tính và các ràng buộc.





Các bước phát triển cơ sở dữ liệu



Các bước phát triển cơ sở dữ liệu

Mô hình hóa
dữ liệu doanh nghiệp

Mô hình h
dữ liệu ý ni

- Định nghĩa CSDL trong DBMS.
- Chọn cấu trúc vật lý của dữ liệu: cấu trúc lưu trữ, chỉ mục, phương thức truy xuất, tổ chức tập tin.
- Thiết kế các chương trình xử lý CSDL.

Thiết kế
CSDL luận lý

Thiết kế
CSDL vật lý

Thực hiện CSDL

Bảo trì CSDL



Các bước phát triển cơ sở dữ liệu

Mô hình hóa
dữ liệu doanh nghiệp



- Mô hình hóa dữ liệu
- Viết và kiểm tra các chương trình xử lý CSDL.
 - Hoàn tất tài liệu về CSDL và tài liệu đào tạo.
 - Đào tạo người sử dụng.
 - Cài đặt CSDL và chuyển đổi dữ liệu từ các hệ thống trước.



Thiết kế
CSDL vật lý



Thực hiện CSDL



Bảo trì CSDL

Các bước phát triển cơ sở dữ liệu

Mô hình hóa
dữ liệu doanh nghiệp

Mô hình hóa
dữ liệu ý niệm

- Phân tích CSDL và các ứng dụng CSDL để bảo đảm thỏa mãn các yêu cầu thông tin.
- Tinh chỉnh CSDL để nâng cao hiệu suất.
- Đào tạo người sử dụng.
- Xác định các lỗi sai trong CSDL và các ứng dụng CSDL, phục hồi CSDL khi hư hỏng.

Thực hiện CSDL

Bảo trì CSDL

Kiến trúc 3 lược đồ

Mức ngoài

Sinh viên

(mãsv, họ tên, tên lớp)

Khung nhìn ngoài

Khung nhìn ngoài

Ánh xạ
mức ngoài-
mức ý niệm

Mức ý niệm

Lược đồ ý niệm

Sviên(mãsv, họ tên, mã lớp)

Lớp(mã lớp, tên lớp)

Ánh xạ mức ý niệm-
mức trong

Mức trong

Lược đồ trong

Qlsv.mdb

Cơ sở dữ liệu

Kiến trúc 3

```
create view Sinhviên(mãsv, họ tên, tên lớp)
as
select mãsv, họ tên, tên lớp
from sviên, lớp
where sviên.mãlớp = lớp.mãlớp
```

Mức ngoài

Sinhviên

(mãsv, họ tên, tên lớp)

Mức ý niệm

Sviên(mãsv, họ tên, mãlớp)

Lớp(mãlớp, tên lớp)

Mức trong

Qlsv.mdb

Lược đồ ý niệm

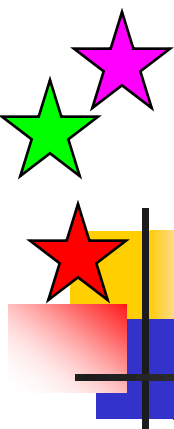
Lược đồ trong

Cơ sở dữ liệu

Ánh xạ
mức ngoài-
mức ý niệm

Ánh xạ mức ý niệm-
mức trong

Ánh xạ của DBMS
(tuyệt mật)



Kiến trúc database ba mức

1. Mức trong (mức vật lý - Physical level):

- Đây là mức lưu trữ CSDL. Tại mức này, vấn đề cần giải quyết là dữ liệu gì và được lưu trữ như thế nào? ở đâu (đĩa từ, băng từ, track, sector ... nào)? Cần các chỉ mục gì? Việc truy xuất là tuần tự (Sequential Access) hay ngẫu nhiên (Random Access) đối với từng loại dữ liệu.
- Những người hiểu và làm việc với CSDL tại mức này là người quản trị CSDL (Administrator), những người sử dụng (NSD) chuyên môn.

Kiến trúc database ba mức

1. Mức trong (mức vật lý - Physical level):

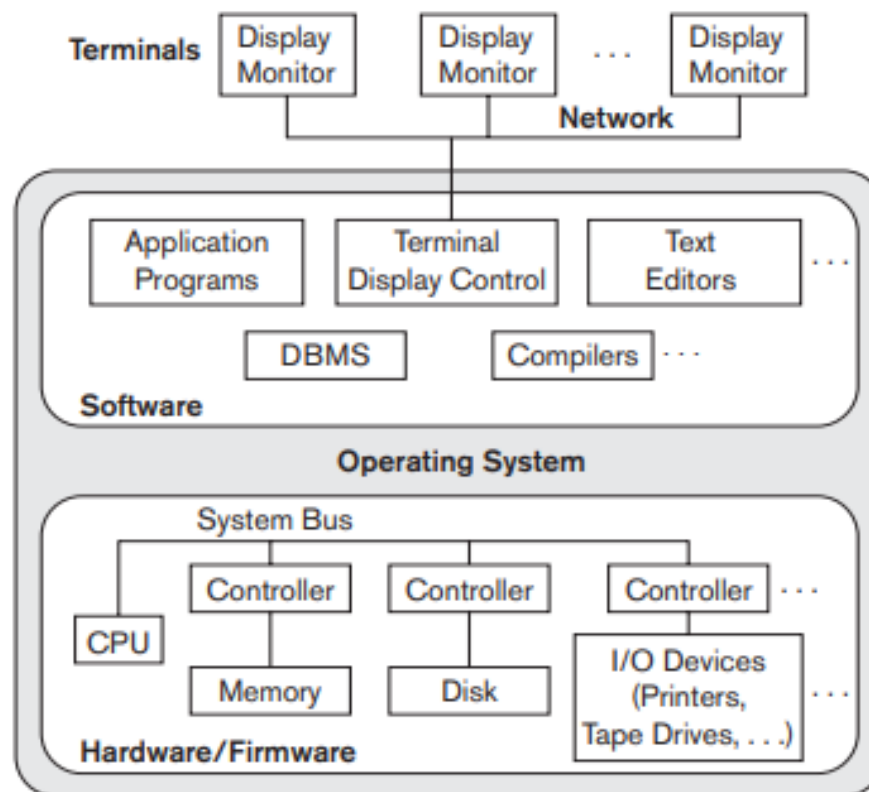
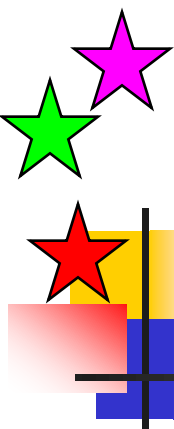


Figure 2.4
A physical centralized
architecture.

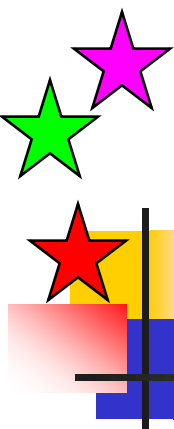


Kiến trúc database ba mức

Cấu trúc Dữ liệu trong Lược đồ Vật lý

Tên mục dữ liệu	Vị trí bắt đầu	Độ dài (bytes)
TEN	1	30
MASV	31	4
PHAI	35	4
MAKH	39	4

- Tổ chức vật lý các mẫu tin SVIEN
 - chiều dài của mẫu tin SVIEN = 42 bytes
 - được sắp xếp và lập chỉ mục trên MASV



Kiến trúc database ba mức

- ❖ **Mức quan niệm hay mức logic (conception level, logical level)**
 - CSDL mức quan niệm là một sự biểu diễn trừu tượng CSDL mức vật lý; hoặc ngược lại, CSDL vật lý là sự cài đặt cụ thể của CSDL mức quan niệm.
 - Trả lời câu hỏi cần phải lưu trữ bao nhiêu loại dữ liệu? Đó là những dữ liệu gì? Mối quan hệ giữa chúng như thế nào?



Kiến trúc database ba mức

Cấu trúc dữ liệu trong Lược đồ Quan niệm

SVIEN

<u>MASV</u>	HOTENSV	PHAI	MALOP
-------------	---------	------	-------

LOP

<u>MALOP</u>	TENLOP	SISO
--------------	--------	------

MONHOC

<u>MAMH</u>	TENMH	TINCHI	SOTIET
-------------	-------	--------	--------

KETQUA

MASV	MAMH	DIEM
------	------	------



Kiến trúc database ba mức

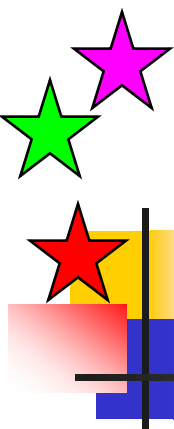
3. Mức ngoài hay mức nhìn (view level) :

- Đó là mức của người sử dụng và các chương trình ứng dụng. Làm việc tại mức này có các nhà chuyên môn, các kỹ sư tin học và những người sử dụng không chuyên.
- Người sử dụng hay chương trình ứng dụng có thể hoàn toàn không được biết về cấu trúc tổ chức lưu trữ thông tin trong CSDL, thậm chí ngay cả tên gọi của các loại dữ liệu hay tên gọi của các thuộc tính. Họ chỉ có thể làm việc trên một phần CSDL theo cách "nhìn" do người quản trị hay chương trình ứng dụng quy định, gọi là khung nhìn (View).



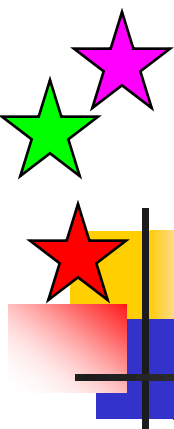
Các thành phần của môi trường DBMS

- Phần cứng (Hardware)
- Phần mềm (Software)
- Dữ liệu (data)
- Các thủ tục (procedure)
- Con người
 - Người quản trị CSDL (DBA–Database administration)
 - Người quản trị dữ liệu (DA – Data administration)
 - Người thiết kế CSDL (Database designer)
 - Người phát triển ứng dụng (application developer)
 - Người dùng cuối (end-user)



Kiến trúc DBMS

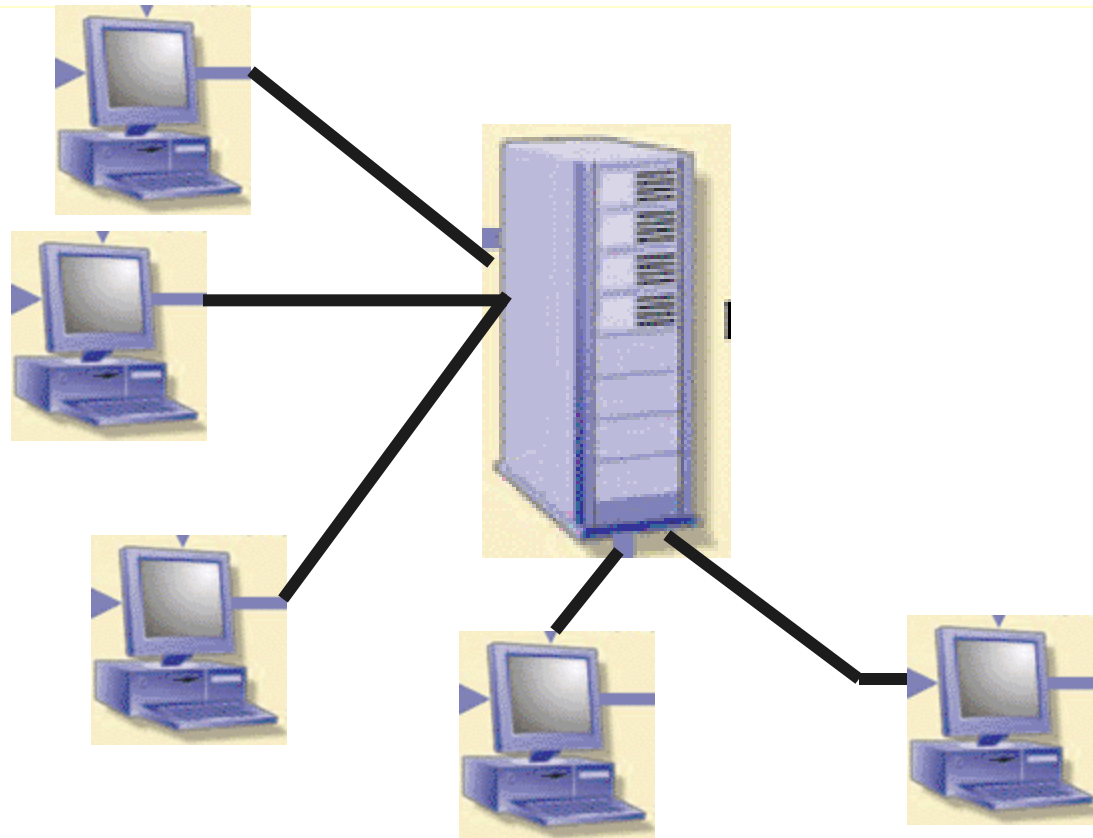
- Kiến trúc truyền thống
- Kiến trúc file – server
- Kiến trúc client-server



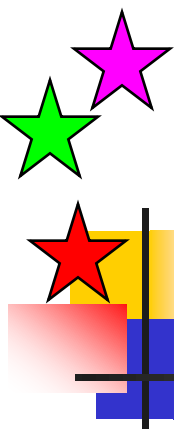
Kiến trúc truyền thống

- Còn gọi là hệ xử lý từ xa (teleprocessing)
- Gồm 1 máy tính CPU và 1 số trạm đầu cuối (terminal)
 - Tất cả xử lý tập trung trên cùng 1 máy tính.
 - Các trạm đầu cuối gửi yêu cầu về máy trung tâm

Kiến trúc truyền thông



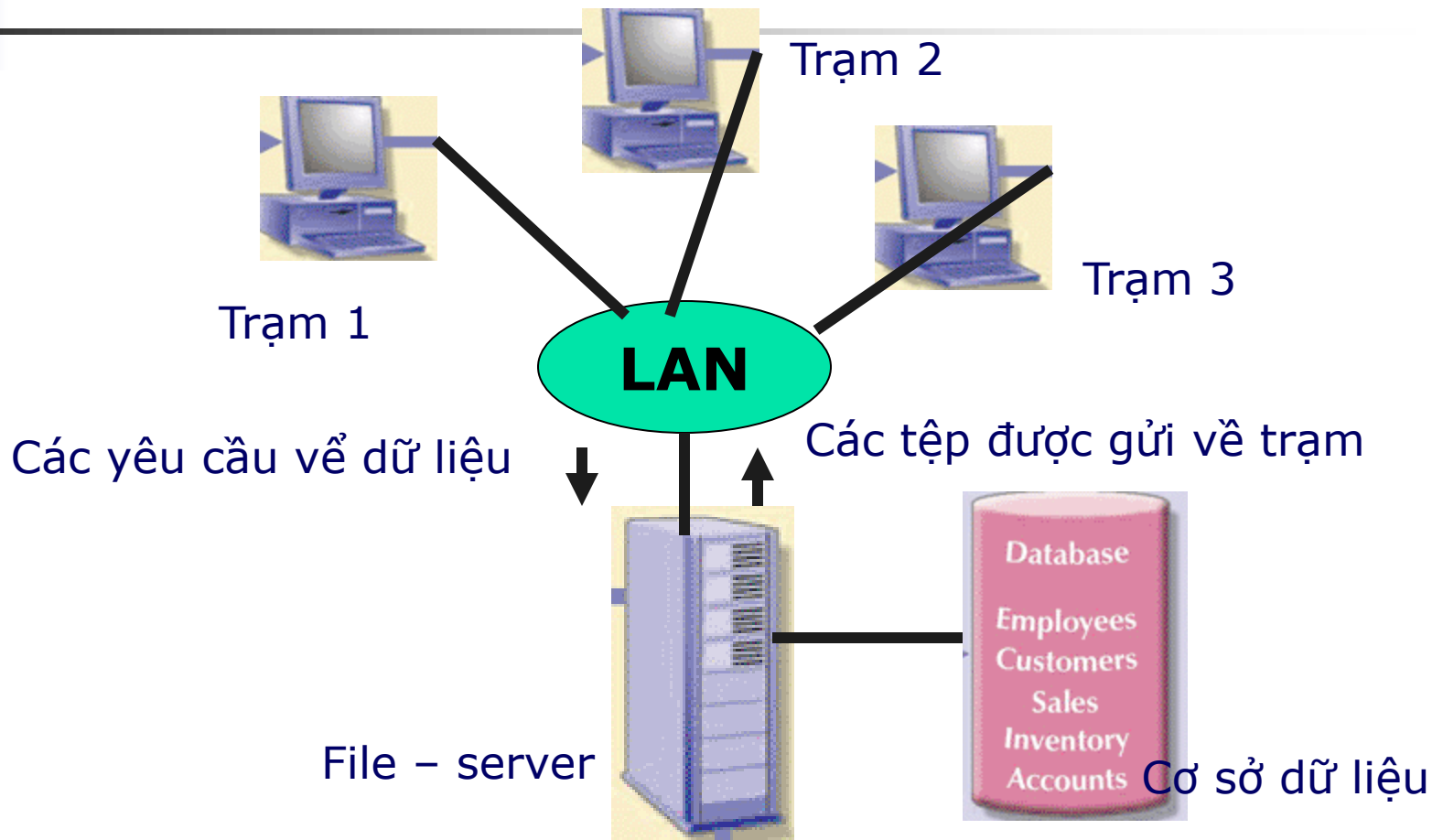
Database System



Kiến trúc file – server

- Việc xử lý không tập trung vào 1 máy trung tâm mà phân tán trên mạng.
- File-server lưu giữ các tệp dữ liệu mà các ứng dụng và hệ QTCSDL cần đến
 - Dữ liệu truyền trên mạng nhiều
 - Mỗi trạm phải cài đặt một bản sao đầy đủ của hệ QTCSDL
 - Tính nhất quán dữ liệu phức tạp hơn vì có nhiều hệ QTCSDL truy cập vào cùng tệp dữ liệu

Kiến trúc file – server



Kiến trúc file – server

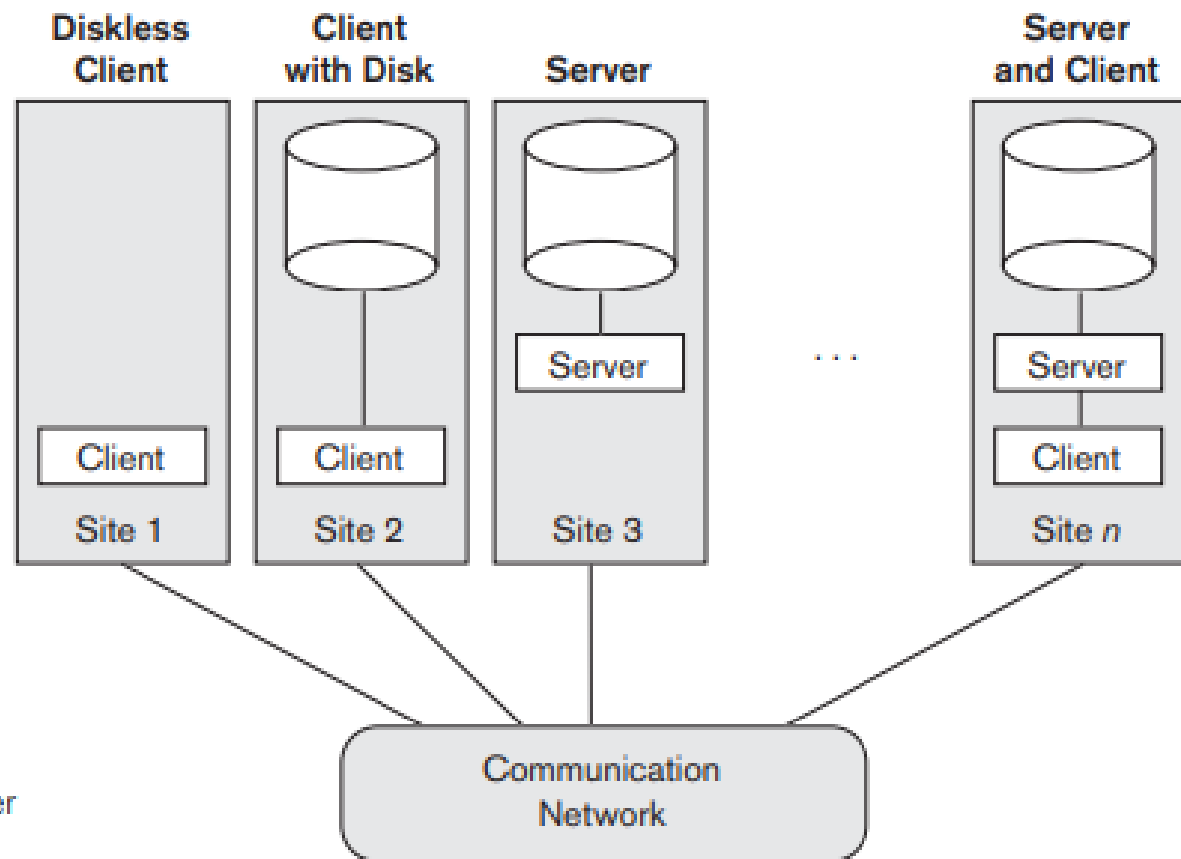


Figure 2.6

Physical two-tier client/server architecture.

Kiến trúc file – server

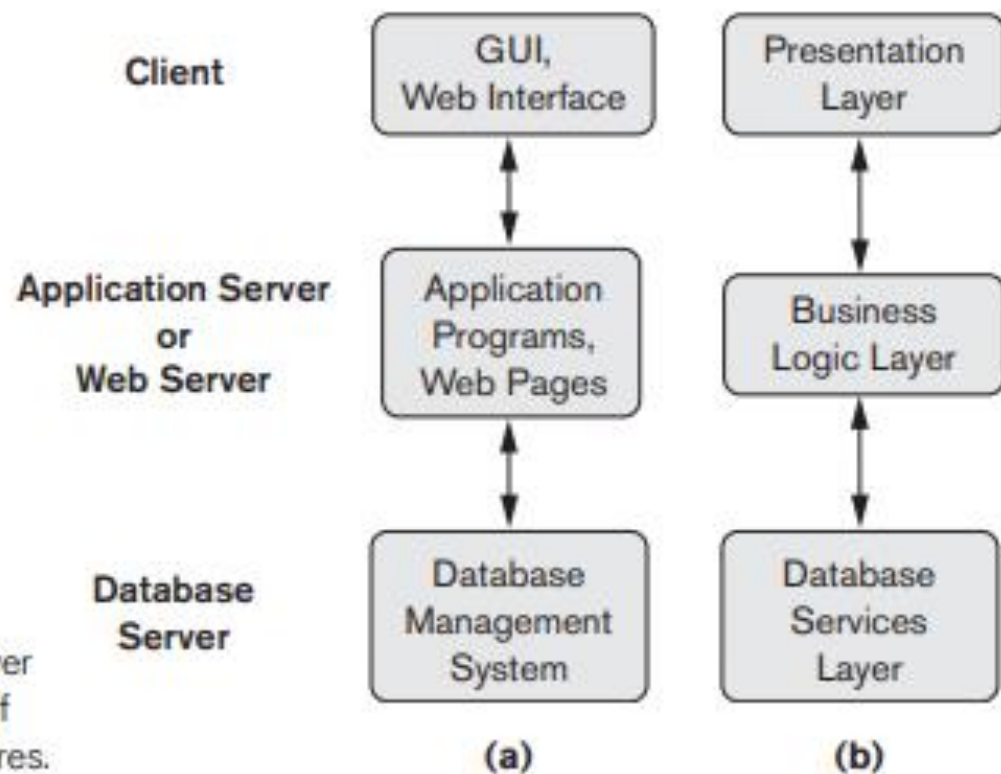


Figure 2.7

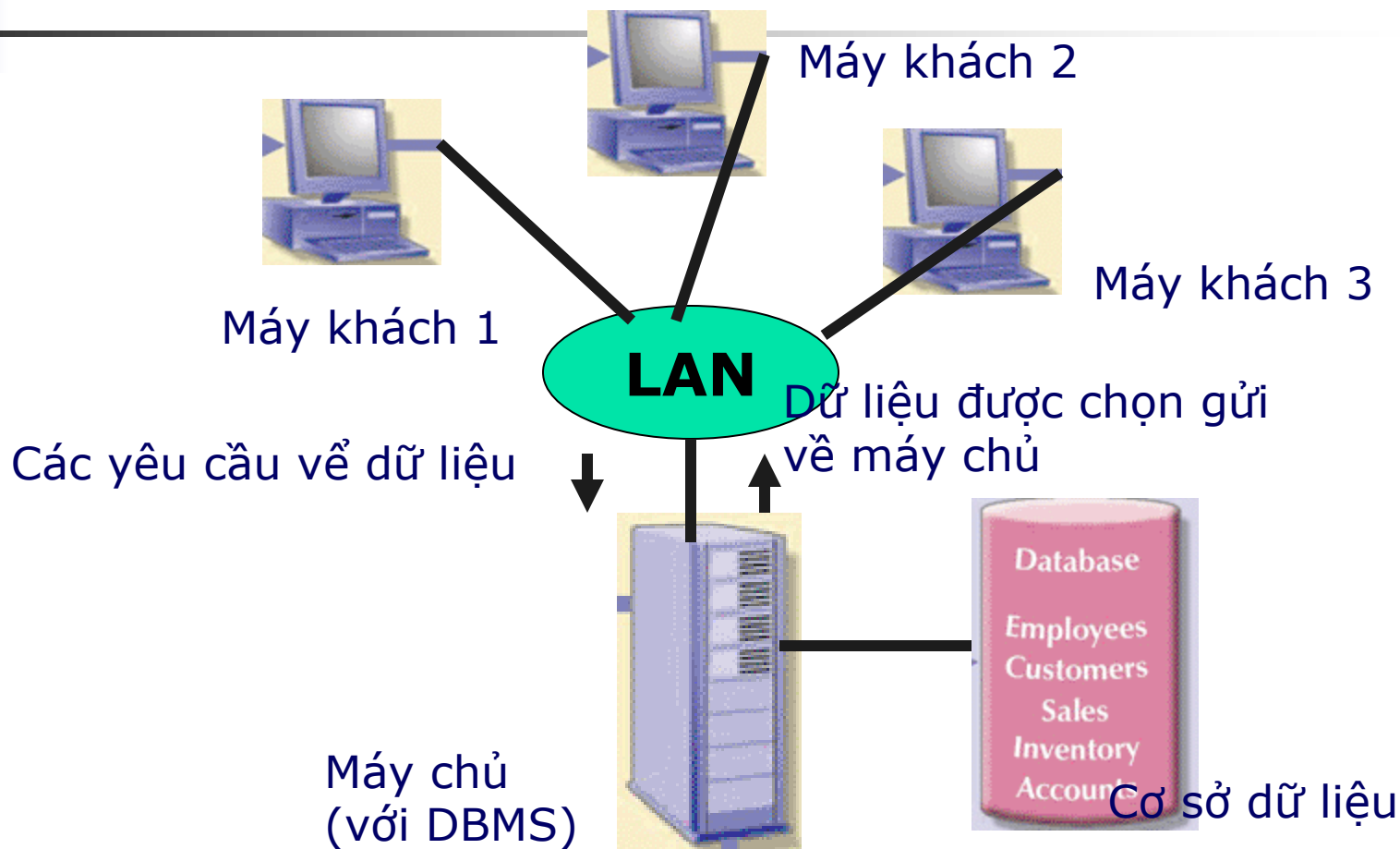
Logical three-tier client/server architecture, with a couple of commonly used nomenclatures.



Kiến trúc Client/Server

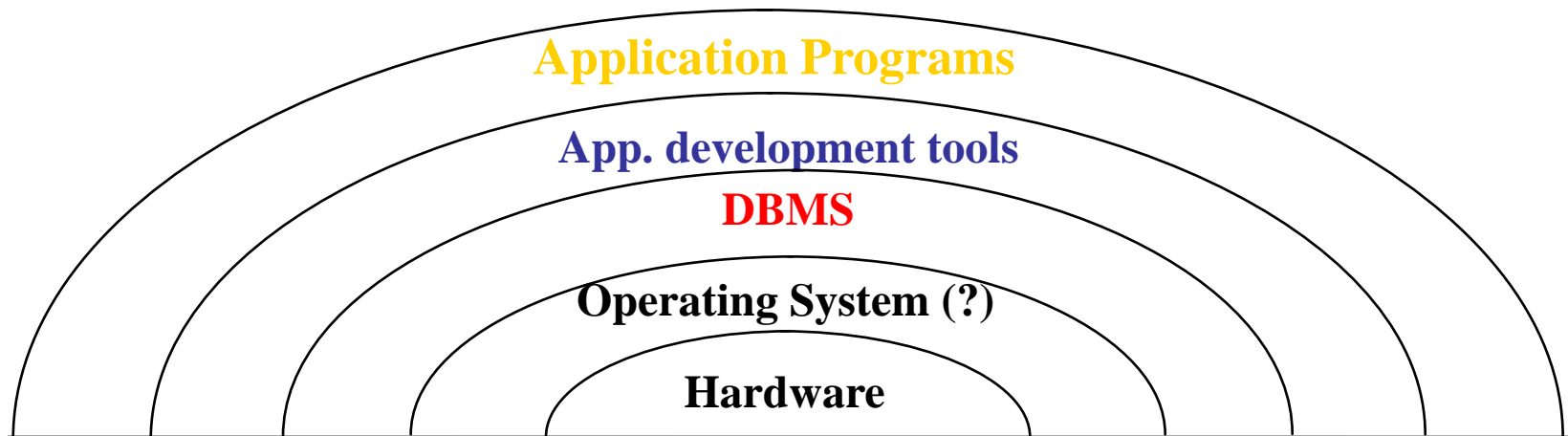
- Khắc phục được nhược điểm của 2 kiểu trên
- Các bộ phận phần mềm tương tác với nhau tạo nên 1 hệ thống:
 - Tiến trình client quản trị giao diện người dùng và ứng dụng logic. Tiến trình client nhận yêu cầu người dùng, kiểm tra, gửi thông điệp về server.
 - Tiến trình server tiếp nhận, xử lý yêu cầu, gửi trả kết quả về lại client.

Kiến trúc client – server





CSDL trong một hệ thống máy tính





Tóm tắt bài học

- Một CSDL là một tập dữ liệu có quan hệ được lưu trữ trong các bảng.
- Một hệ quản trị CSDL có thể được định nghĩa như một tập các bản ghi quan hệ và các chương trình có thể truy cập và thao tác trên các bản ghi.
- Một mô hình dữ liệu mô tả một bộ chứa dữ liệu và xử lý dữ liệu từ bộ chứa đó.
- DBMS là một tập các chương trình mà nó cho phép người dùng lưu trữ, cập nhật và trích rút thông tin từ một CSDL.
- Một hệ quản trị CSDL quan hệ (RDBMS) là một phần mềm cho phép tạo, thao tác, cập nhật trên một CSDL quan hệ.
- Một CSDL quan hệ được chia nhỏ thành các đơn vị logic gọi là bảng. Các đơn vị logic này có quan hệ với nhau trong CSDL.



Câu hỏi

- **Câu 1:** Phân biệt các thuật ngữ sau:
 - Cơ sở dữ liệu ?
 - Hệ quản trị cơ sở dữ liệu ?
 - Hệ cơ sở dữ liệu ?
- **Câu 2:** Cơ sở dữ liệu được ứng dụng rộng rãi trong cuộc sống. Hãy nêu một số ví dụ?
- **Câu 3:** Hãy nêu các chức năng chính của một CSDL và hệ quản trị cơ sở dữ liệu?
- **Câu 4:** Cho biết mục đích của việc mô hình hóa các cơ sở dữ liệu. Chỉ ra sự khác nhau giữa mô hình logic và mô hình vật lý.
- **Câu 5:** Vì sao các hệ cơ sở dữ liệu theo mô hình quan hệ được sử dụng phổ biến ?



Câu hỏi

HỎI : Mô hình dữ liệu được dùng để làm gì?

- Mô hình dữ liệu dùng để thiết kế CSDL.
- Khi thiết kế CSDL cần xác định :
 - Mỗi dữ liệu thể hiện một đối tượng có cấu trúc cụ thể
 - Mỗi quan hệ giữa các dữ liệu trong CSDL

HỎI : Khi thiết kế một CSDL cần xác định những yếu tố nào?

- VD : Khi xét hồ sơ của một học sinh ta cần xác định lưu trữ trực tiếp họ tên HS hoặc chỉ lưu mã HS hoặc lưu cả hai.



Câu hỏi

- HỎI : Em hãy cho biết có mấy loại mô hình cơ sở dữ liệu?
- a/ Mô hình logic :
 - Còn được gọi là mô hình bậc cao : Cho biết bản chất logic của biểu diễn dữ liệu, những cái gì được biểu diễn trong CSDL
- b/ Mô hình Vật lý :
 - Còn được gọi là mô hình bậc thấp : cho biết dữ liệu được lưu trữ như thế nào.
- Quá trình thiết kế thường được chia làm nhiều bước, cụ thể từ bước này sang bước khác. Quá trình này được thực hiện bằng cách chuyển từ mô hình logic sang mô hình vật lý.

Thank you

