

Chương 1

Tổng quan về CSDL

Khoa HTTT, Trường CNTT&TT

Nội dung

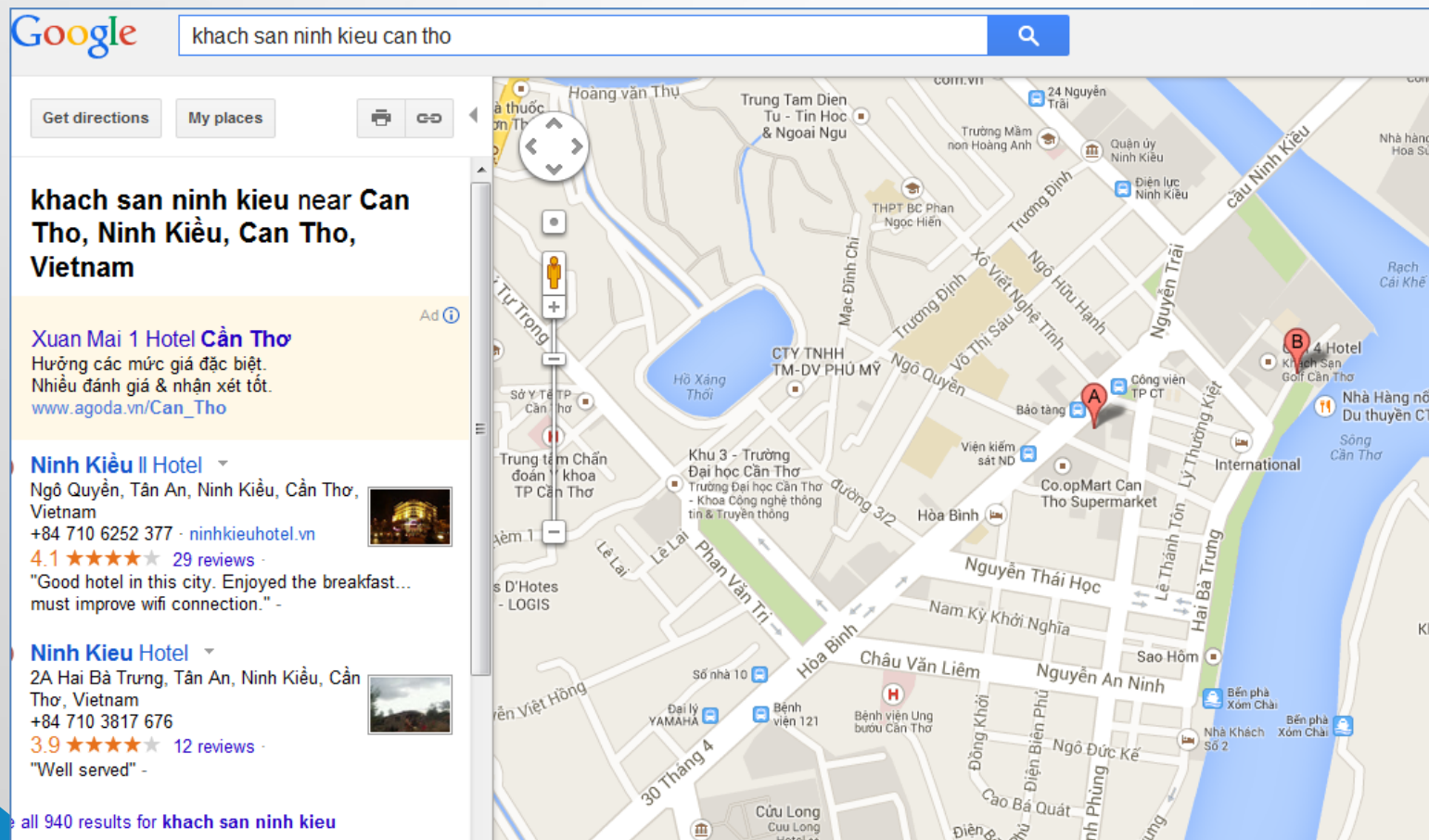
- Khái niệm chung
- Khung nhìn dữ liệu
- Các mô hình của CSDL
- Ngôn ngữ cơ sở dữ liệu
- Quản lý các giao dịch (transaction)
- Quản lý lưu trữ
- Người quản trị CSDL và Người sử dụng CSDL
- Hệ quản trị CSDL

Khái niệm về dữ liệu

- Dữ liệu ở khắp mọi nơi xung quanh ta ...
 - Dữ liệu về phim
 - Dữ liệu của thư viện
 - Dữ liệu liên quan cho thuê xe, đặt chỗ máy bay
 - Dữ liệu về thẻ bảo hiểm, bệnh viện
 - Dữ liệu của các nhà mạng (Vinaphone, ...)
 - Hệ thống thông tin địa lý
 - ...

Khái niệm về dữ liệu

- Dữ liệu ở khắp mọi nơi xung quanh ta ...



Khái niệm về dữ liệu

- Dữ liệu ở khắp mọi nơi xung quanh ta ... trong công ty:
 - Khách hàng
 - Sản phẩm
 - Đơn đặt hàng
 - Hóa đơn
 - ...

HOÁ ĐƠN BÁN HÀNG

Đơn vị : CÔNG TY TNHH DƯỢC PHẨM HOÀNG CHƯƠNG

Mẫu số 01-TT

HOÁ ĐƠN BÁN HÀNG

Ngày 04/05/2008

Số HĐ: 72053

Tk nợ: 131

Tk có:

Đơn vị bán hàng: Công ty TNHH Dược phẩm Hoàng Chương - Kho CTY -

Điện thoại: (08) 9143786, (08) 914.3787. MST: 0 3 0 1 1 4 6 9 0 1

Họ tên người mua hàng: Cty TNHH DP Hồng Phương

Địa chỉ: 135A Nguyễn Tất Thành P8 TP. Cà Mau. Số tài khoản:

Hình thức thanh toán: Thời hạn thanh toán

Lý do xuất: Cty TNHH DP Hồng Phương

STT	TÊN HÀNG	QUICÁCH	SỐ LÔ	DATE	NƠI SX	SL	ĐƠN GIÁ	THÀNH TIỀN
1	Vidoca		010108	31/01/2011		10	4.285	42.850
2	Vidoca		010108	31/01/2011		10	4.285	42.850
3	Nazovel		410508	31/05/2010		5	25.715	128.575
4	Pemecar		01/120508	31/05/2010		10	38.190	381.900
5	Fenroxil 500 mg		040309	31/03/2012		10	70.476	704.760
6	Forvita		070707	31/07/2010		10	20.000	200.000
7	Fonclar		081208	31/12/2010		70	68.571	4.799.970
8	Eurpharton H45 viên Lafon		170908	30/09/2010		20	34.286	685.720
				//				
				//				

Tổng tiền hàng: 6.966.625 VND

Tiền chiết khấu: 0 VND

Thuế suất GTGT: 5% Tiền thuế GTGT: 348.331 VND

Tổng cộng tiền thanh toán: 7.314.956 VND

Số tiền viết bằng chữ: bảy triệu ba trăm mười bốn nghìn chín trăm năm mươi sáu đồng

Thủ trưởng đơn vị
(Ký, họ tên)Kế toán trưởng
(Ký, họ tên)Phụ trách cung tiêu
(Ký, họ tên)Người nhận hàng
(Ký, họ tên)Gám đốc
(Ký, họ tên)

Làm thế nào lưu trữ dữ liệu ?

Đơn vị: CÔNG TY TNHH DƯỢC PHẨM HOÀNG CHƯƠNG Mẫu số 01-TT

Đơn vị: CÔNG TY TNHH DƯỢC PHẨM HOÀNG CHƯƠNG Mẫu số 01-TT

Đơn vị: CÔNG TY TNHH DƯỢC PHẨM HOÀNG CHƯƠNG Mẫu số 01-TT

HOÁ ĐƠN BÁN HÀNG

Ngày 04/05/2008

Số HD: 72053
Tk ng: 131
Tk có:

Đơn vị bán hàng: Công ty TNHH Dược phẩm Hoàng Chương - Khe CTY -
Điện thoại: (08) 9143786, (08) 914.3787 MST: 0 3 0 1 1 4 6 9 0 1
Họ tên người mua hàng: Cty TNHH DP Hồng Phương
Địa chỉ: 135A Nguyễn Tất Thành, P.8 TP. Cà Mau Số tài khoản:
Hình thức thanh toán: Thời hạn thanh toán:
Lý do xuất: Cty TNHH DP Hồng Phương

STT	TÊN HÀNG	QUICÁCH	SỐ LÔ	DATE	NƠI SX	SL	ĐƠN GIÁ	THÀNH TIỀN
1	Vidoca		010108	31/01/2011		10	4 285	42 850
2	Vidoca		010108	31/01/2011		10	4 285	42 850
3	Nazovel		410508	31/05/2010		5	25 715	128 575
4	Pemecar		01/120508	31/05/2010		10	38 180	381 900
5	Fonixdi 500 mg		040809	31/03/2012		10	70 476	704 760
6	Fonvita		070707	31/07/2010		10	20 000	200 000
7	Fonclar		061208	31/12/2010		70	68 571	4 799 970
8	Eurphaton H45 viên Lato		170908	30/09/2010		20	34 288	685 720
			//					
			//					

Cộng tiền hàng: 6 866 825 VND
Tiền chiết khấu: 0 VND
Thuế suất GTGT: 5% Tiền thuế GTGT: 348 331 VND
Tổng cộng liên thanh toán: 7 314 656 VND

Số tiền viết bằng chữ: bảy triệu ba trăm mười bốn nghìn chín trăm năm mươi sáu đồng

Thủ trưởng đơn vị (Ký, họ tên) Kế toán trưởng (Ký, họ tên) Phụ trách cung tiêu (Ký, họ tên) Người nhận hàng (Ký, họ tên) Giám đốc (Ký, họ tên)



Làm thế nào lưu trữ dữ liệu ?

- Làm thế nào sắp xếp dữ liệu trong một máy tính?
 - . Scan các hóa đơn
 - . Copy các dữ liệu vào trong các bảng trong word
 - . Copy các dữ liệu vào trong các sheet trong Excel
 - . Copy các dữ liệu vào trong một cơ sở dữ liệu

Làm thế nào lưu trữ dữ liệu ?

- Quan sát ...

Đơn vị : CÔNG TY TNHH DƯỢC PHẨM HOÀNG CHƯƠNG Mẫu số 01-TT

HOÁ ĐƠN BÁN HÀNG

Ngày 04/05/2009 Số HD: 72053
TK ng: 131
TK có:

Đơn vị bán hàng: Công ty TNHH Dược phẩm Hoàng Chương - Kho CTY -
Điện thoại: (08) 9143786, (08) 914.3787 MST: 0 3 0 1 1 4 6 9 0 1

Họ tên người mua hàng: Cty TNHH DP. Hồng Phương
Địa chỉ: 135A Nguyễn Tất Thành P8 TP. Cà Mau Số tài khoản:
Hình thức thanh toán: Thời hạn thanh toán
Lý do xuất: Cty TNHH DP. Hồng Phương

STT	TÊN HÀNG	QUICÁCH	SỐ LÔ	DATE	NƠI SX	SL	ĐƠN GIÁ	THÀNH TIỀN
1	Vidoca		010108	31/01/2011		10	4 285	42 850
2	Vidoca		010108	31/01/2011		10	4 285	42 850
3	Nazovel		410508	31/06/2010		5	25 715	128 575
4	Pemecar		01/120508	31/06/2010		10	36 190	361 900
5	Fonroxil 500 mg		040309	31/03/2012		10	70 476	704 760
6	Fonvita		070707	31/07/2010		10	20 000	200 000
7	Fonclar		061208	31/12/2010		70	68 571	4 799 970
8	Eurpharton H45 viên Lafon		170908	30/09/2010		20	34 296	685 720
				///				
				///				

Cộng tiền hàng: 6 966 825 VND
Tiền chiết khấu: 0 VND
Thuế suất GTGT: 5% Tiền thuế GTGT: 348 331 VND
Tổng cộng tiền thanh toán: 7 314 956 VND

Số tiền viết bằng chữ: bảy triệu ba trăm mười bốn nghìn chín trăm năm mươi sáu đồng

Thủ trưởng đơn vị (Ký, họ tên) Kế toán trưởng (Ký, họ tên) Phụ trách cung tiêu (Ký, họ tên) Người nhận hàng (Ký, họ tên) Giám đốc (Ký, họ tên)

Dữ liệu
khách hàng

Dữ liệu
hóa đơn

Chi tiết
hóa đơn

Làm thế nào lưu trữ dữ liệu ?

- Chuyển đổi dữ liệu vào các bảng

Dữ liệu hóa đơn

soHD	ngay	tonghang	VAT	Tong VAT	tongHD
72053	4/5/2009	6 966 625	0.05	348331	7 314 956

Chi tiết hóa đơn

tenhang	quycach	solo	ngay	nois x	SL	Dongi a	Thanh tien
Fonvita		070707	31/07/2010		10	20000	200 000
...							

Dữ liệu khách hàng

hoten	diachi	taikhoan
Cty TNHH DP Hồng Phương	135A Nguyễn Tất Thành P8, TP Cà Mau	

- Quan sát ...
 - Các dữ liệu **tonghang**, **tongVAT**, **tongHD**, **thanhtien** được tính : không cần thiết lưu giữ chúng, chúng có thể được tính lại khi cần
 - Không thể hình thành lại đúng hóa đơn gốc ban đầu : ai là khách hàng của hóa đơn, chi tiết hóa đơn là của hóa đơn nào ?

Làm thế nào lưu trữ dữ liệu ?

- Dữ liệu dư thừa với tham chiếu ...

Dữ liệu hóa đơn

soHD	hoten	ngay	VAT
72053	Cty THHH...	4/5/2009	0.05

Chi tiết hóa đơn

soHD	tenhang	quycach	solo	ngay	nois x	SL	Dongia
72053	Fonvita		070707	31/07/2010		10	20000
	...						

Dữ liệu khách hàng

hoten	diachi	taikhoan
Cty TNHH DP Hồng Phương	135A Nguyễn Tất Thành P8, TP Cà Mau	

- Quan sát
 - Nhiều chi tiết hóa đơn có cùng sản phẩm : các thông tin sản phẩm lặp lại nhiều lần.

=> Tách thông tin của sản phẩm trong một bảng riêng.

Làm thế nào lưu trữ dữ liệu ?

- Dữ liệu không dư thừa...

Dữ liệu hóa đơn

soHD	hoten	ngay	VAT
72053	Cty THHH...	4/5/2009	0.05

Chi tiết hóa đơn

soHD	tenhang	SL	Dongia
72053	Fonvita	10	20000
	...		

Dữ liệu khách hàng

hoten	diachi	taikhoan
Cty TNHH DP Hồng Phương	135A Nguyễn Tất Thành P8, TP Cà Mau	

Dữ liệu sản phẩm

tenhang	quycach	solo	ngay	noisx
Fonvita		070707	31/07/2010	
...				

=> CSDL đầu tiên ...

Cơ sở dữ liệu (Database)

- Định nghĩa

CSDL là một tập dữ liệu có cấu trúc, được tích hợp, được chia sẻ và được thiết kế nhằm đáp ứng nhu cầu của nhiều người dùng.

- Ví dụ:

- CSDL cho phép quản lý nhân sự, sinh viên, môn học, đăng ký, ... của một trường ĐH
- CSDL của các hệ thống đặt chỗ máy bay, khách sạn, ...
- CSDL cho các tài khoản của một ngân hàng, ...

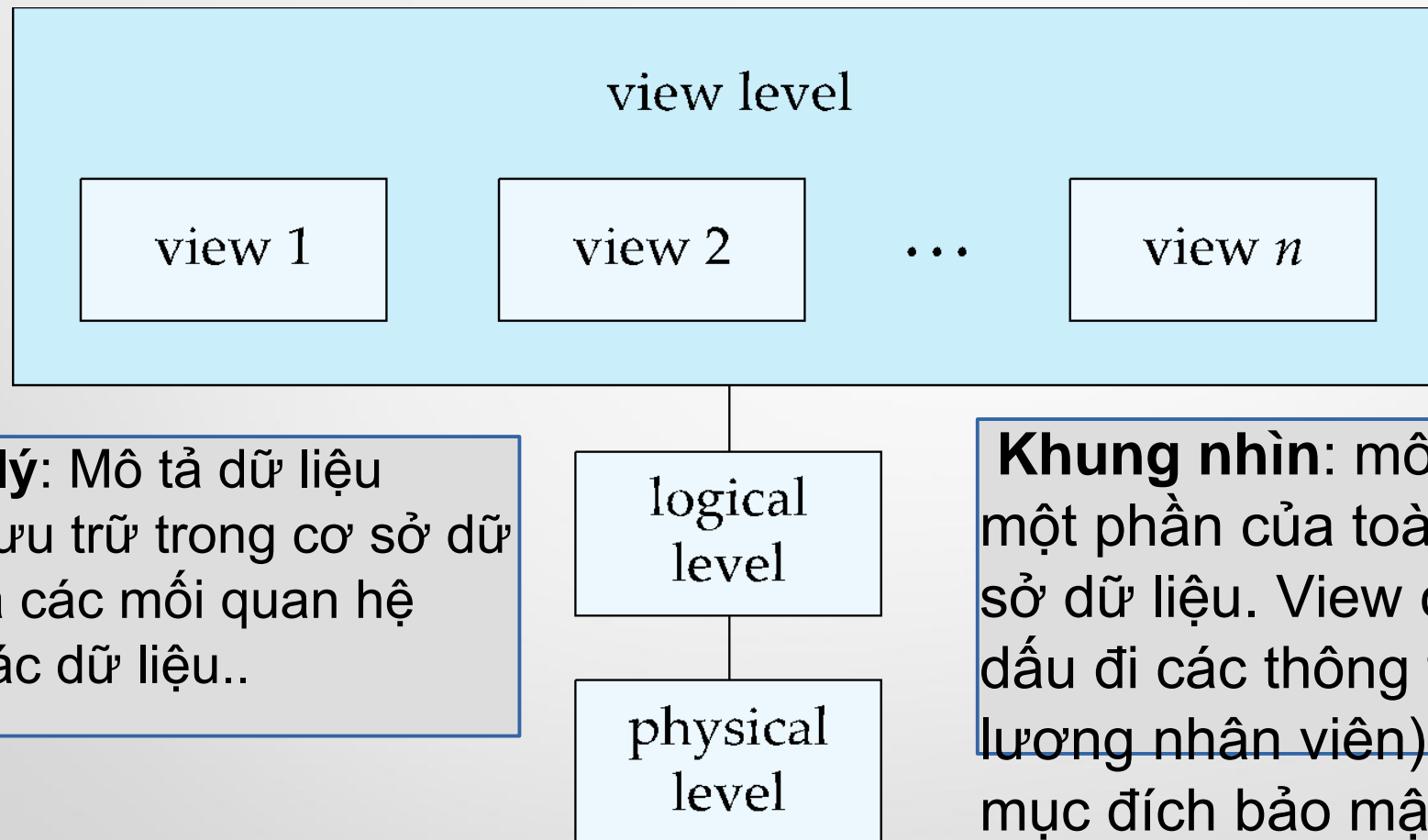
=> Việc quản lý các CSDL đặt ra các vấn đề phức tạp

=> Hệ quản trị CSDL (DBMS)

Các mức trừu tượng (Abstraction levels)

- Để việc tìm kiếm được hiệu quả => người thiết kế thường dùng các cấu trúc dữ liệu phức tạp để biểu diễn dữ liệu trong CSDL
- Người phát triển dấu người dùng các chi tiết phức tạp qua nhiều mức trừu tượng để đơn giản việc tương tác với hệ thống:
 - Vật lý (Physic)
 - Mức luận lý (Logical)
 - Mức khung nhìn (View)

Các mức trừu tượng (Abstraction levels)



Luận lý: Mô tả dữ liệu được lưu trữ trong cơ sở dữ liệu, và các mối quan hệ giữa các dữ liệu..

Khung nhìn: mô tả chỉ một phần của toàn bộ cơ sở dữ liệu. View có thể dấu đi các thông tin (ví dụ lương nhân viên) cho mục đích bảo mật.

Vật lý: Mô tả cách một mẫu tin (ví dụ, khách hàng) được lưu trữ.

Mô hình dữ liệu

Data model

- Mô hình dữ liệu là một khái niệm được dùng để mô tả dữ liệu hoặc thông tin.
- Một mô hình dữ liệu bao gồm một tập các công cụ (thường là khái niệm, ngôn ngữ) dùng để mô tả :
 - Cấu trúc của dữ liệu được lưu trữ bên trong CSDL,
 - Những thao tác trên dữ liệu,
 - Ràng buộc dữ liệu.

Mô hình dữ liệu

Data model

- Cấu trúc dữ liệu:
 - Ngôn ngữ lập trình: mảng, cấu trúc, đối tượng.
 - CSDL: mô hình dữ liệu ở một mức cao hơn cấu trúc dữ liệu, ví dụ « mô hình quan niệm »
 - Ngôn ngữ định nghĩa dữ liệu DDL cho phép mô tả cấu trúc của dữ liệu trong CSDL

Mô hình dữ liệu

Data model

- Cấu trúc dữ liệu:
 - Ngôn ngữ lập trình: mảng, cấu trúc, đối tượng.
 - CSDL: mô hình dữ liệu ở một mức cao hơn cấu trúc dữ liệu, ví dụ « mô hình quan niệm »
 - Ngôn ngữ định nghĩa dữ liệu DDL cho phép mô tả cấu trúc của dữ liệu trong CSDL

Mô hình dữ liệu

Data model

- Những thao tác trên dữ liệu:
 - Ngôn ngữ lập trình: bất cứ điều gì có thể được lập trình.
 - CSDL: thao tác truy vấn, cập nhật dữ liệu
 - Ngôn ngữ thao tác dữ liệu DML cho phép mô tả các thao tác trên CSDL.
- Ràng buộc dữ liệu:
 - Các giới hạn hoặc điều kiện trên dữ liệu.
 - Ví dụ: một ngày của tháng có giá trị trong khoảng giữa 1 và 31.
 - Ngôn ngữ cho phép đặc tả các ràng buộc thường là một phần của DDL.

Mô hình CSDL

- Các mô hình dữ liệu phổ biến hiện nay
 - Mô hình quan hệ
 - Mô hình dữ liệu thực thể - Quan hệ (Entity-Relationship) (Chủ yếu dùng cho thiết kế CSDL)
 - Mô hình dữ liệu dựa trên đối tượng (Hướng đối tượng và đối tượng-quan hệ)
 - Mô hình dữ liệu bán cấu trúc (XML)
- Các mô hình dữ liệu cũ:
 - Mô hình mạng
 - Mô hình phân cấp

Ngôn ngữ cơ sở dữ liệu

- Cho phép người dùng/ứng dụng tương tác với CSDL
- Có 2 loại chính:
 - Ngôn ngữ định nghĩa dữ liệu (Data Definition Language – DDL): Định nghĩa lược đồ CSDL.
 - Ngôn ngữ thao tác dữ liệu (Data Manipulation Language – DML) : cập nhật và truy vấn CSDL.

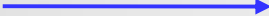
=> Không phải là 2 ngôn ngữ độc lập, là các thành phần của một ngôn ngữ CSDL, ví dụ ngôn ngữ SQL (Structured Query Language) : ngôn ngữ CSDL được sử dụng rộng rãi.

Quản lý giao dịch

- **Một giao dịch** là một tập hợp các hoạt động/ thao tác mà thực hiện một chức năng luận lý đơn giản trong một ứng dụng cơ sở dữ liệu.
- Điều gì xảy ra nếu hệ thống có sự cố ?
- Điều gì xảy ra nếu nhiều người dùng cập nhật đồng thời trên cùng một dữ liệu ?

Quản lý giao dịch

- Ví dụ 1: Xét giao dịch chuyển tiền từ tài khoản A sang tài khoản B



T
start
read A
$A = A - S$
write A
read B
$B = B + S$
write B
commit

Quản lý giao dịch

- Ví dụ 2: Xét 2 giao dịch cập nhật cùng dữ liệu:

T_1	T_2	BD
		$A = 10$
read A		
	read A	
$A = A + 10$		
write A		$A = 20$
	$A = A + 50$	
	write A	$A = 60$

$\Rightarrow A = ?$

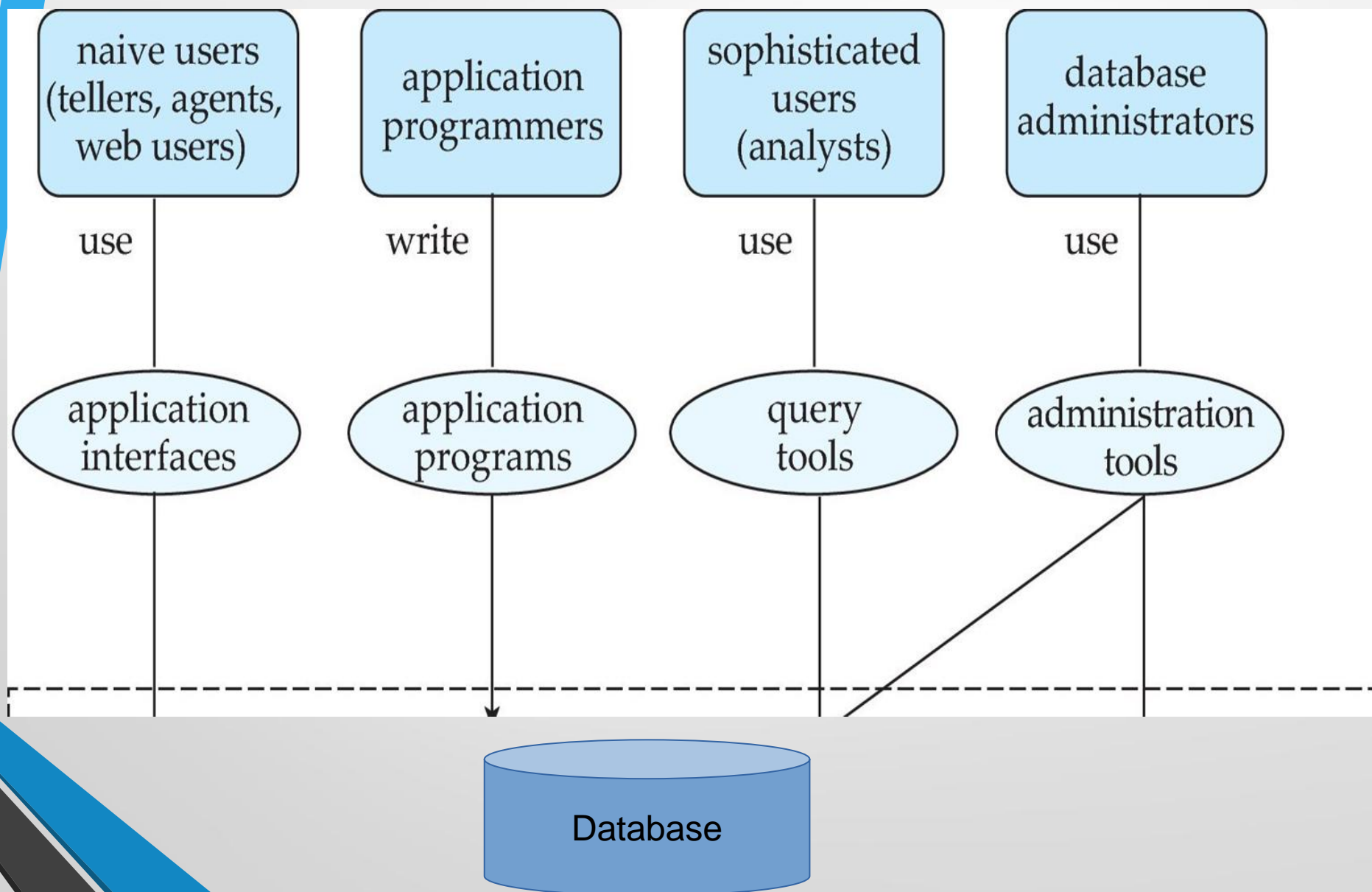
Quản lý giao dịch

- **Bộ phận quản lý giao dịch** đảm bảo CSDL vẫn ở trạng thái nhất quán (đúng) bất chấp lỗi hệ thống (ví dụ như mất điện và hệ điều hành bị treo) và sự thất bại của giao dịch.
- **Bộ phận điều khiển-cạnh tranh** điều khiển sự tương tác giữa các giao dịch đồng thời để bảo đảm tính bền vững của các cơ sở dữ liệu.

Quản lý lưu trữ

- Quản lý lưu trữ là một chương trình cung cấp giao diện giữa các dữ liệu cấp thấp được lưu trữ trong CSDL và các chương trình ứng dụng và truy vấn gửi tới hệ thống.
- Bộ phận quản lý lưu trữ đảm bảo các công việc/nhiệm vụ sau:
 - Tương tác với bộ phận quản lý tập tin
 - Lưu trữ, truy xuất và cập nhật dữ liệu hiệu quả

Người quản trị và người dùng CSDL



Hệ quản trị Cơ sở dữ liệu (Database Management System)

- Định nghĩa

Một DBMS là một hệ thống bao gồm một tập các chương trình (phần mềm) cho phép quản lý và truy xuất các CSDL: định nghĩa cấu trúc, truy vấn, cập nhật, sao lưu, phục hồi, ... các CSDL.

- Ví dụ:

- Nguồn mở: MySQL, PostgreSQL,...
- Thương mại: SQL Server, Oracle, DB2,...

Hệ quản trị Cơ sở dữ liệu (Database Management System)

- **Tổ chức dữ liệu**

DBMS quan hệ tổ chức dữ liệu trong các bảng và được lưu trữ trên đĩa. Nó tạo ra các cơ chế để đảm bảo truy cập nhanh dữ liệu;

- **Quản lý dữ liệu**

DBMS đảm bảo sự phát triển phù hợp của dữ liệu, nó kiểm tra các ràng buộc dữ liệu.

- **Truy cập dữ liệu**

DBMS cho phép truy cập dữ liệu đồng thời từ cả người dùng và từ các chương trình xử lý dữ liệu.

Hệ quản trị Cơ sở dữ liệu (Database Management System)

- **Bảo vệ chống lại rủi ro, tai nạn**

DBMS đảm bảo tính toàn vẹn và tính sẵn sàng của dữ liệu trong trường hợp xảy ra sự cố hoặc bị tấn công.

- **Quản lý truy cập đồng thời**

DBMS cho phép truy cập dữ liệu đồng thời từ hàng trăm hoặc thậm chí hàng ngàn người sử dụng. Nó kiểm soát chặt chẽ các thao tác đồng thời trên cùng dữ liệu.

- **Kiểm soát truy cập**

DBMS đảm bảo rằng chỉ người dùng có thẩm quyền có thể truy cập dữ liệu và thay đổi chúng.