LOGO

CHƯƠNG 1
TỔNG QUAN VỀ AN TOÀN CSDL

TS. Trần Thị Lượng
* Khoa An toàn thông tin *

MỤC TIÊU

- **Kiến thức:**
  - Hiểu và trình bày được các kiến thức cơ bản về CSDL, DBMS, hệ CSDL, thiết kế CSDL, SQL, ..., các tấn công hàng đầu vào CSDL
- **Kỹ năng:**
  - Thực hành được một số câu lệnh SQL cơ bản trong Oracle, SQL Server, MySQL

TÀI LIỆU THAM KHẢO

- [1] TS. Nguyễn Nam Hải, TS. Lương Thế Dũng, ThS. Trần Thị Lượng, *Giáo trình An toàn cơ sở dữ liệu,* Học viện Kỹ thuật Mật mã, 2013.
- [2] David M. Kroenke, David Auer, *Database Concepts,* 7th edition, Pearson, 2014.
- Alfred Basta, Melissa Zgola, *Database Security*, Cengage Learning US, 2011.
- [3] Bhavani Thuraisingham, *"Database and applications security – integrating information security and data management",* Auerbach 2005.
- [4] Amichai Shulman, *Top Ten Database Security Threats*, Imperva.

NỘI DUNG

1. Một số khái niệm trong CSDL
2. Thiết kế CSDL
3. Ngôn ngữ SQL
4. Kiến trúc DBMS
5. Các yêu cầu bảo vệ CSDL

NỘI DUNG

1. Một số khái niệm trong CSDL
2. Thiết kế CSDL
3. Ngôn ngữ SQL
4. Kiến trúc DBMS
5. Các yêu cầu bảo vệ CSDL

MỘT SỐ KHÁI NIỆM TRONG CSDL

Câu hỏi:

- *CSDL là gì?*
- *CSDL khác dữ liệu ở chỗ nào?*
- *DBMS là gì? Cho ví dụ*

Một số khái niệm trong CSDL

CSDL: là một tập hợp dữ liệu và một tập các quy tắc tổ chức dữ liệu chỉ ra các mối quan hệ giữa chúng.

DBMS: là hệ thống phần mềm cho phép quản lý, thao tác trên CSDL, tạo ra sự trong suốt phân tán với người dùng.

- ***Ví dụ:*** Access, Foxpro, MySQL, SQL, Oracle, DB2, SyBase…PostgreSQL

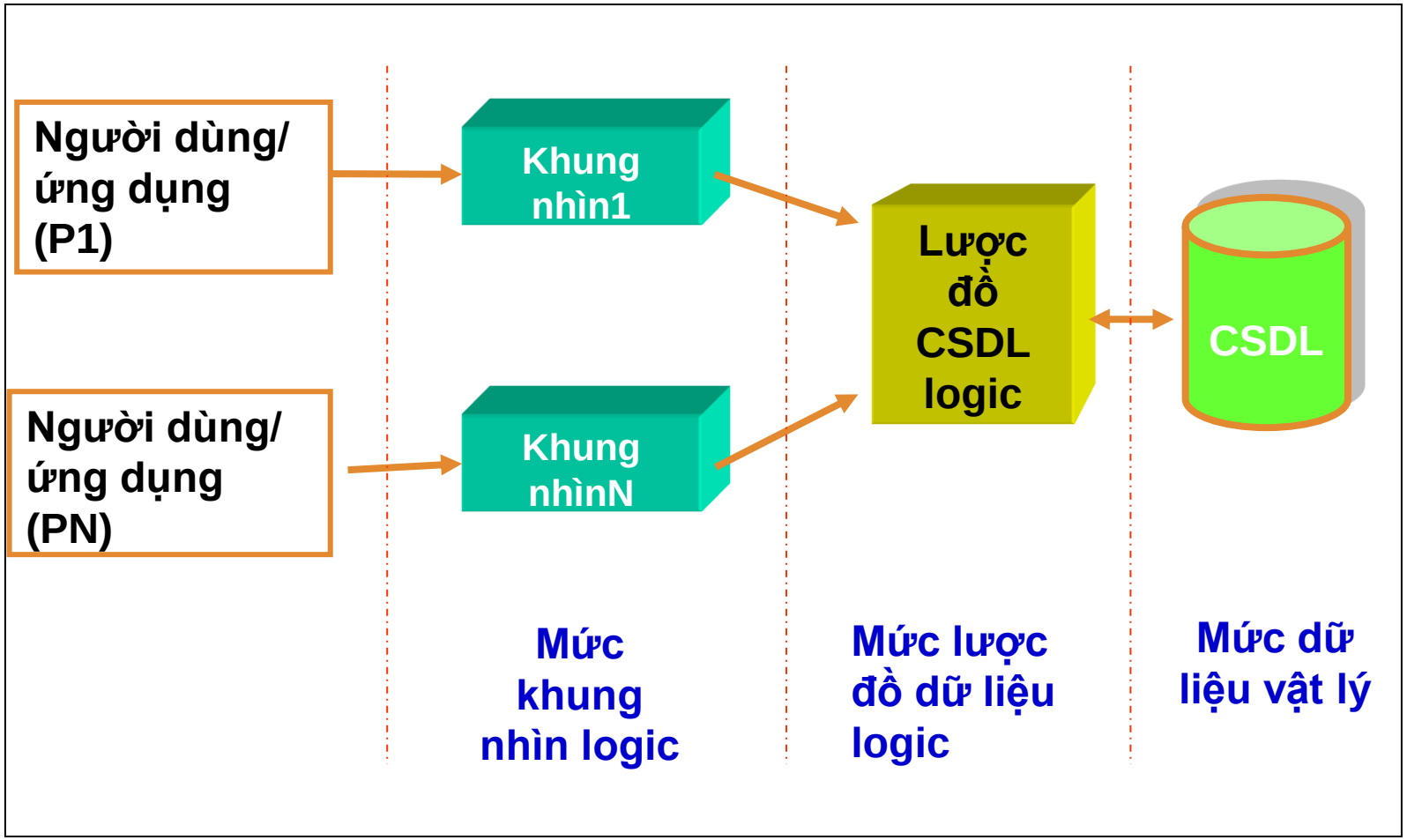
Một số khái niệm trong CSDL

- ***Mô hình logic:*** phụ thuộc vào DBMS (ví dụ mô hình quan hệ, mô hình phân cấp, mô hình mạng, mô hình hướng đối tượng)
- ***Mô hình khái niệm:*** độc lập với DBMS.
  - ***Ví dụ***: mô hình quan hệ thực thể (E-R) là một trong các mô hình khái niệm phổ biến nhất, được xây dựng dựa trên khái niệm thực thể.

Một số khái niệm trong CSDL

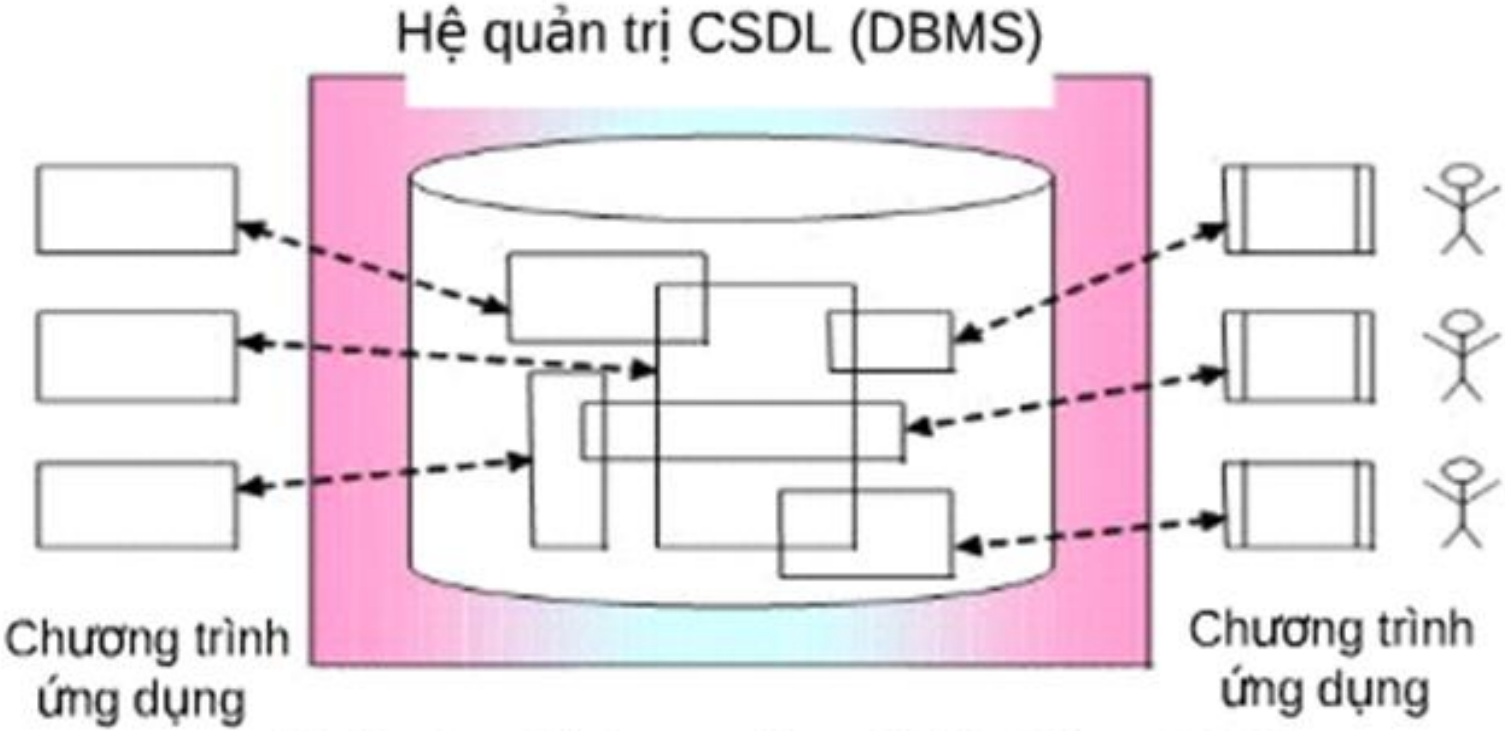
- ***Lược đồ dữ liệu vật lý***: mô tả cấu trúc lưu trữ dữ liệu trong các file trên bộ nhớ ngoài. Dữ liệu → bản ghi, con trỏ.
- ***Lược đồ dữ liệu logic:*** mọi dữ liệu trong CSDL được mô tả bằng mô hình lôgíc của DBMS. Các dữ liệu và quan hệ của chúng được mô tả thông qua ngôn ngữ DDL của DBMS.
- ***Khung nhìn logic:*** phụ thuộc các yêu cầu của mô hình logic và các mục đích của ứng dụng. Khung nhìn logic mô tả một phần lược đồ CSDL logic. Sử dụng DDL để định nghĩa các khung nhìn logic, DML để thao tác trên các khung nhìn này.

Các mức mô tả dữ liệu



HỆ CƠ SỞ DỮ LIỆU

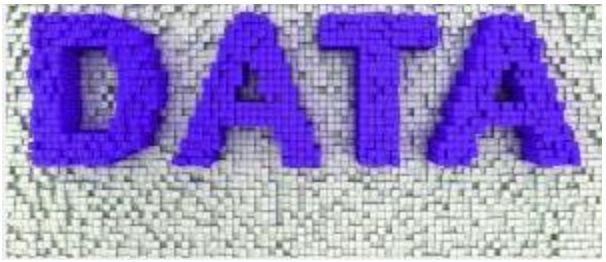
- Một hệ CSDL gồm 4 thành phần:
  - Dữ liệu
  - Phần cứng
  - Phần mềm
  - Người dùng



Hình : Sơ đồ lược giản về hệ thống CSDL

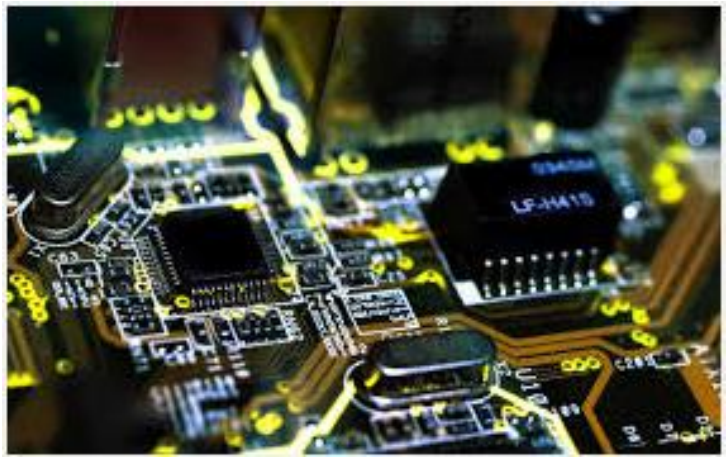
HỆ CƠ SỞ DỮ LIỆU

- **Dữ liệu:** có hai đặc trưng chính
  - *Tính tích hợp:* CSDL là nơi tập hợp nhiều hồ sơ và nó được loại bỏ đến mức tối đa các dư thừ dữ liệu
  - *Tính chia sẻ*: CSDL là nơi cho phép nhiều người sử dụng truy cập đồng thời



HỆ CƠ SỞ DỮ LIỆU

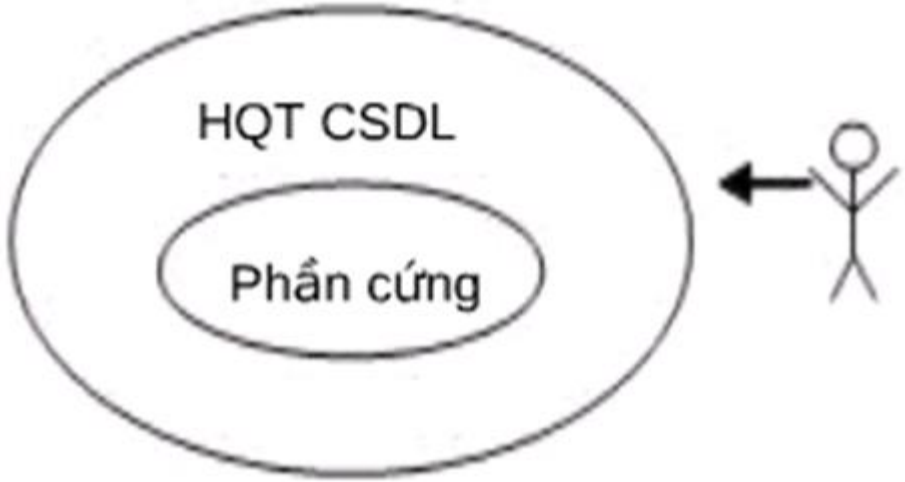
- **Phần cứng** của hệ CSDL gồm:
  - *Bộ nhớ ngoài*: đĩa từ, đĩa cứng, khối vào/ra, ổ đĩa, …
  - *Bộ xử lý và bộ nhớ trong*, card mạng/modem, …



HỆ CƠ SỞ DỮ LIỆU

❖Phần mềm (DBMS):

- Đứng trung gian giữa phần cứng và người dùng CSDL.
- Chức năng cơ bản: tạo lớp vỏ bọc phần cứng đối với người dùng.



HỆ CƠ SỞ DỮ LIỆU

❖ Người dùng:

- *Lớp thứ nhất – Lập trình viên CSDL:* là người viết chương trình ứng dụng sử dụng CSDL thông qua một ngôn ngữ như: C++, PHP, ASP, v.v.
- *Lớp thứ hai – Người dùng cuối:* sử dụng chương trình đã lập sẵn (chương trình ƯD hoặc một phần của DBMS) để giao tiếp với CSDL.
- *Lớp thứ ba – quản trị viên CSDL (DBA):* là người làm công tác quản trị CSDL.

HỆ CƠ SỞ DỮ LIỆU

- **Người dùng:** *các lớp người dùng an toàn*
  - *Nhân viên an toàn*: kiểm soát an toàn cho CSDL thông qua các quyền truy nhập, các chính sách an toàn, ...
  - *Kiểm toán viên*: chịu trách nhiệm kiểm tra các yêu cầu kết nối và các câu hỏi truy nhập, kiểm tra vết kiểm toán nhằm phát hiện ra các xâm phạm vào CSDL.
  - *Nhân viên sao lưu/phục hồi:* chịu trách nhiệm sao lưu cơ sở dữ liệu, phục hồi hệ thống cơ sở dữ liệu khi gặp sự cố

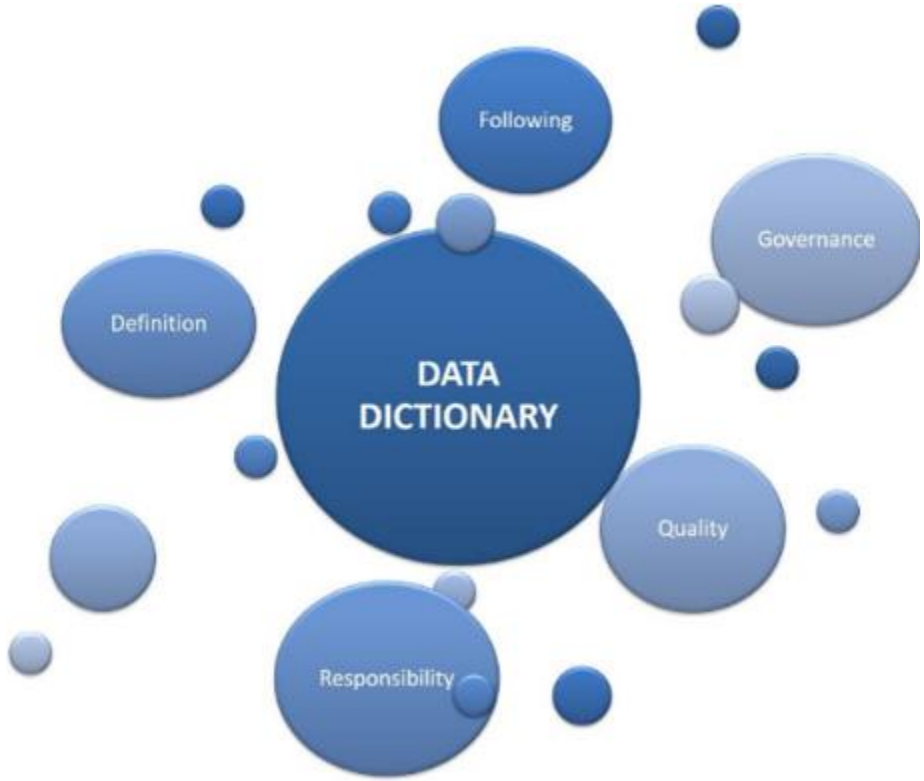
Một số khái niệm trong CSDL...

❖ Câu hỏi:

- *Từ điển dữ liệu (Data dictionary) là gì?*

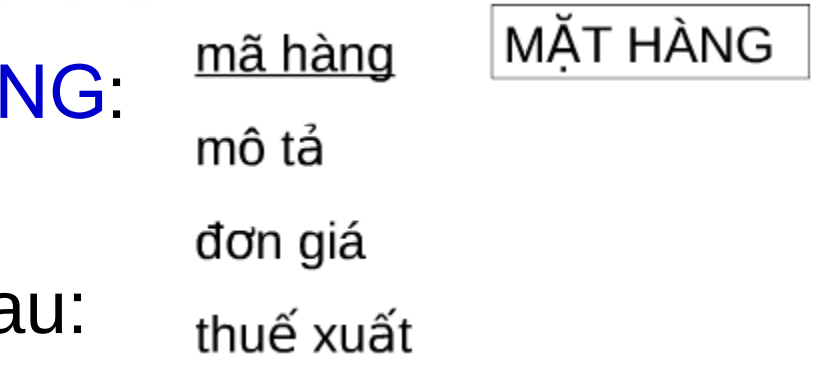
❖ Trả lời:

- *Nhằm làm rõ hơn các khái niệm, các thực thể và thuộc tính trong CSDL.*



Từ điển dữ liệu

- Ví dụ: tập thực thể MẶT HÀNG:



- Sẽ có từ điển dữ liệu như sau:

Thực thể	MẶT HÀNG
Tên khác	Hàng, sản phẩm, hàng hóa.
Mô tả	Hàng là những thứ được mua với số lượng khác nhau từ các nhà cung ứng, được lưu trữ trong kho và bán cho khách hàng.

Thuộc tính	Mã hàng:	Là một số dùng để phân biệt mặt hàng này với mặt hàng kia. Giá trị có dạng 0001>9999
	Mô tả :	Mô tả mặt hàng gồm qui tắc và hình dáng. Loại ký tự chuỗi gồm 100 ký tự. Có thể có giá trị rỗng.
	Đơn giá:	Đơn giá hiện tại của mặt hàng. Có loại dữ liệu số với 2 số thập phân, có giá trị từ 10 đến 50, mặt nhiên là 0.
	Thuế xuất:	Tỷ suất thuế bán của mặt hàng được ghi dưới dạng phần trăm. Có loại dữ liệu số, có giá trị từ 0 đến 99. Giá trị mặt nhiên 0.

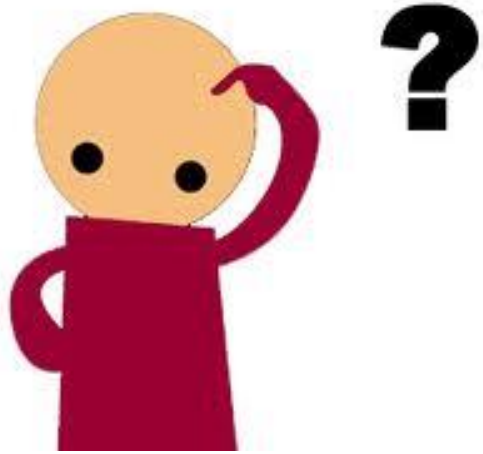
Một số khái niệm trong CSDL...

❖ Câu hỏi:

- *Có mấy mô hình xử lý CSDL?*

❖ Trả lời:

- Có 3 thành phần trong mô hình xử lý CSDL gồm: ƯD, DBMS, Dữ liệu.
- Vị trí của 3 thành phần này sẽ quyết định mô hình đó thuộc loại nào.
- Có 3 mô hình chính:
  - *Mô hình CSDL tập trung*
  - *Mô hình CSDL phân tán*
  - *Mô hình CSDL Client/Server*

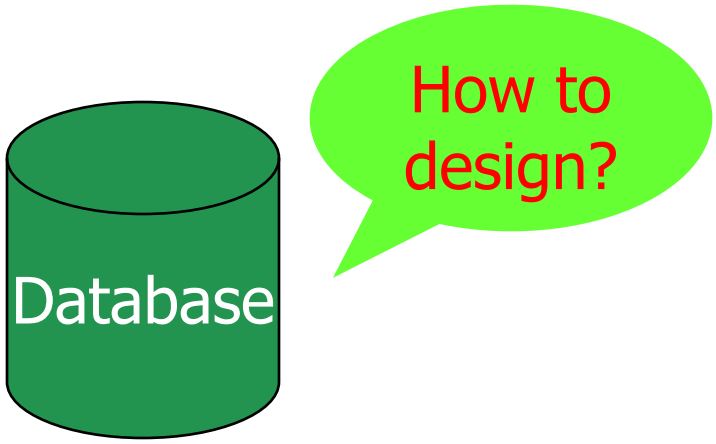


NỘI DUNG

1. Một số khái niệm trong CSDL
2. Thiết kế CSDL
3. Ngôn ngữ SQL
4. Kiến trúc DBMS
5. Các yêu cầu bảo vệ CSDL

THIẾT KẾ CSDL

- Câu hỏi:
  - *Các bước thiết kế một CSDL là gì?*



THIẾT KẾ CSDL

1. Đặc tả vấn đề

Phân tích đặc tả để xác định dữ liệu yêu cầu và mối liên quan giữa chúng để xây dựng mô hình thực thể kết hợp

2. Xây dựng mô hình thực thể quan hệ (ER)

Áp dụng quy tắc biến đổi mô hình thực thể kết hợp thành lược đồ CSDL.

3. Lược đồ CSDL

Lược đồ CSDL xây dựng theo hướng phân tích thiết kế

THIẾT KẾ CSDL

- Nhắc lại về mô hình ER

NỘI DUNG

1. Một số khái niệm trong CSDL
2. Thiết kế CSDL
3. Ngôn ngữ SQL
4. Kiến trúc DBMS
5. Các yêu cầu bảo vệ CSDL

Ngôn ngữ SQL

❖ ***Câu hỏi:***

- *Sự khác nhau giữa SQL, MySQL và SQL Server?*

Ngôn ngữ SQL

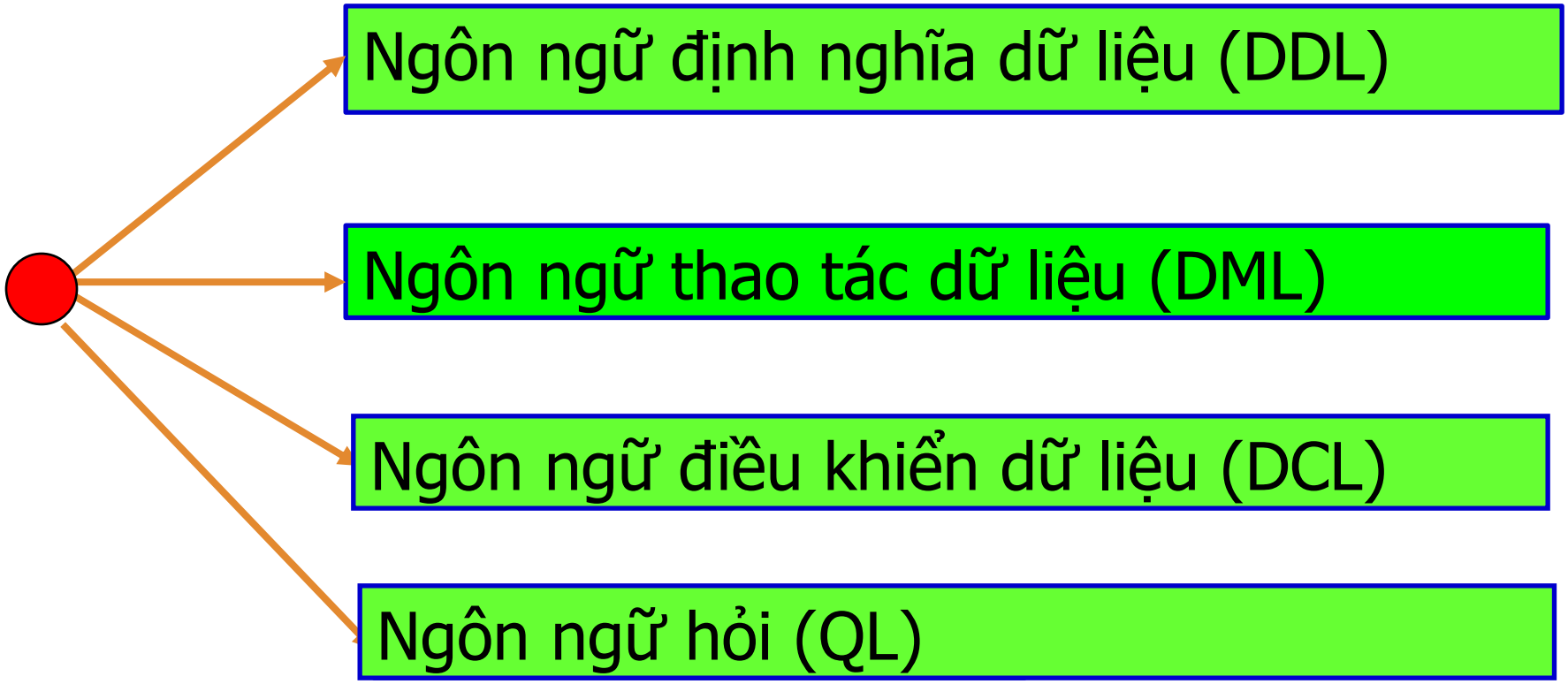
SQL (Structured Query Language) - Ngôn ngữ truy vấn cấu trúc - là một chuẩn của *ANSI* (American National Standards Institute) về truy xuất các hệ thống CSDL.

- Có thể thực thi các câu truy vấn SQL trên CSDL như: Select, insert, update, delete,…
- SQL hoạt động với hầu hết các chương trình CSDL như MS Access, DB2, Informix, MS SQL Server, Oracle, Sybase v.v...

Ngôn ngữ SQL

- *Transact-SQL* (T-SQL): là ngôn ngữ SQL mở rộng dựa trên SQL chuẩn của ISO và ANSI được sử dụng trong SQL Server.
- *PL-SQL* (Procedural-SQL)
  - Là mở rộng ngôn ngữ hướng thủ tục của Oracle
  - PL/SQL kết hợp SQL với các hàm, thủ tục của ngôn ngữ chương trình có cấu trúc như: IF…THEN, WHILE, và LOOP.

CÁC NGÔN NGỮ CON CỦA SQL



CÁC NGÔN NGỮ CON CỦA SQL

- *DDL (Data definition language):* Là ngôn ngữ máy tính để định nghĩa lược đồ CSDL logic.
- Các lệnh DDL quan trọng nhất của SQL là:
  - CREATE TABLE - tạo ra một bảng mới.
  - ALTER TABLE - thay đổi cấu trúc của bảng.
  - DROP TABLE - xóa một bảng.
  - CREATE INDEX - tạo chỉ mục (khóa để kiếm - search key).
  - DROP INDEX - xóa chỉ mục đã được tạo.

CÁC NGÔN NGỮ CON CỦA SQL

- *DDL (Data definition language):*
  - *Ví dụ:* Lệnh ***Create*** sau sẽ tạo ra một table tên ***Employees***

```
CREATE TABLE Employees(
    EmpID   int NOT NULL,
    Name    varchar(30) NOT NULL,
    Salary  numeric(10),
    Contact  varchar(40) NOT NULL
    )
```

CÁC NGÔN NGỮ CON CỦA SQL

❖ *DDL (Data definition language):*

- Lệnh ***Alter:***

```
ALTER TABLE  Employees
ADD  email   varchar(40) NULL
```

- Lệnh ***Drop*** sau đây sẽ hoàn toàn xóa table khỏi database nghĩa là cả định nghĩa của table và data bên trong table đều biến mất (khác với lệnh Delete chỉ xóa data nhưng table vẫn tồn tại).

```
DROP TABLE  Employees
```

CÁC NGÔN NGỮ CON CỦA SQL

- *DML (Data manipulation language):*
  - Là họ các ngôn ngữ máy tính được người dùng sử dụng để tìm kiếm, chèn, xóa và cập nhật dữ liệu trong một CSDL.
  - Ví dụ về DML như các câu lệnh của SQL: SELECT, INSERT, UPDATE, DELETE

CÁC NGÔN NGỮ CON CỦA SQL

- *DML (Data manipulation language):*
  - **Select**

```
SELECT EmpID, Name
FROM Employees WHERE (EmpID = 10)
```

  - **Insert**

```
INSERT INTO Employees
VALUES (101, ‘Lan', ‘HN’,’lan@yahoo.com’)
```

  - **Update**

```
UPDATE Employees SET  Name = ‘Minh'
WHERE  EmpID = 101
```

  - **Delete**

```
DELETE FROM  Employees
WHERE EmpID = 101
```

CÁC NGÔN NGỮ CON CỦA SQL

- *DCL (Data control language):*
  - Là ngôn ngữ điều khiển dữ liệu
  - Sử dụng hai từ khóa là: GRANT và REVOKE
- *QL (Querying languge):*
  - Là ngôn ngữ hỏi
  - Chính là lệnh Select trong DML

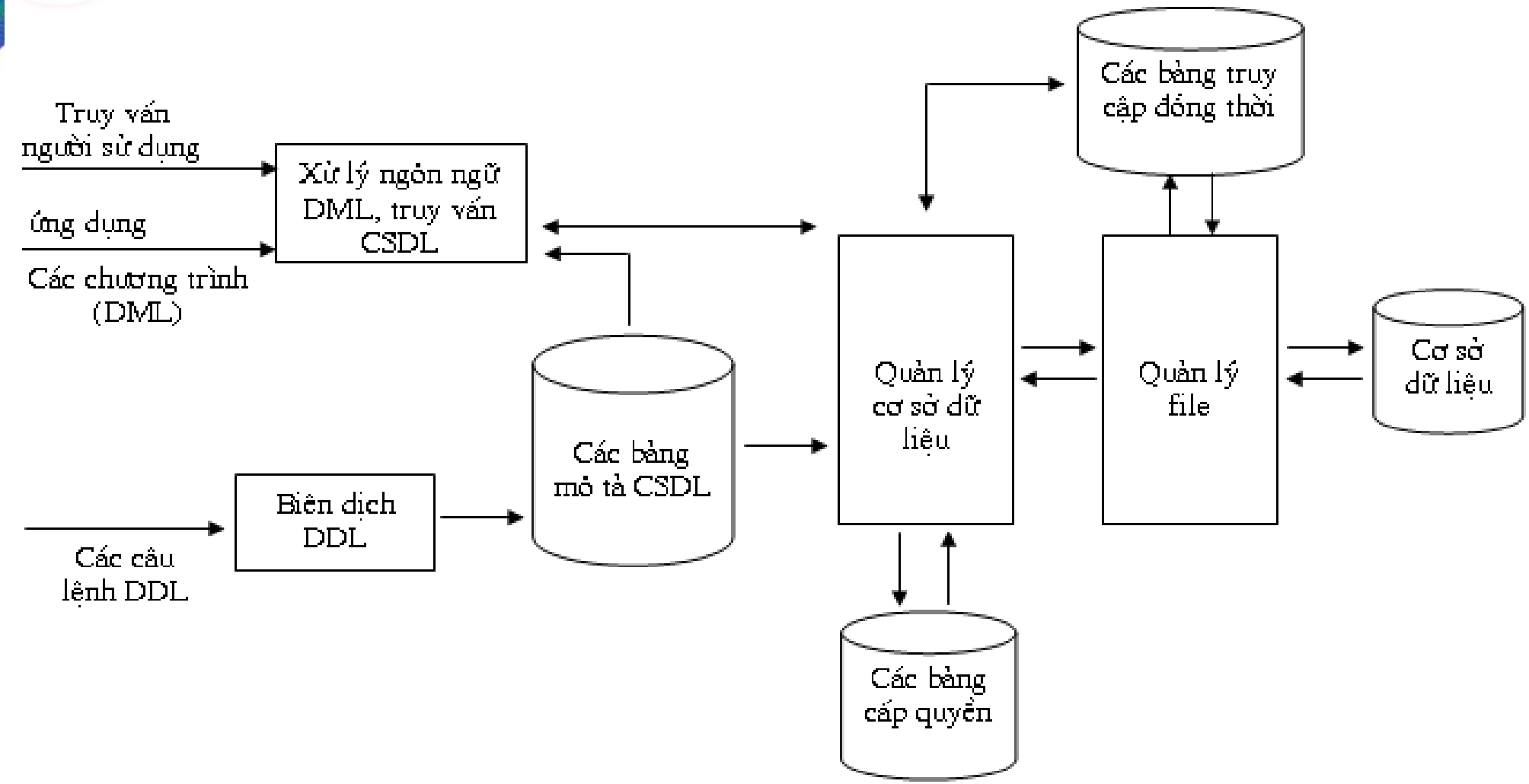
NỘI DUNG

1. Một số khái niệm trong CSDL
2. Thiết kế CSDL
3. Ngôn ngữ SQL
4. Kiến trúc DBMS
5. Các yêu cầu bảo vệ CSDL

Kiến trúc DBMS

- Một DBMS thông thường bao gồm nhiều modul tương ứng với các chức năng sau:
  - Trình biên dịch DDL (DDL Compilation)
  - Trình biên dịch ngôn ngữ DML(DML Compiler)
  - Bộ xử lý truy vấn (Querying Language)
  - Bộ quản lý CSDL - DBMS
  - Bộ quản trị file
- Tập hợp dữ liệu hỗ trợ các modul này là:
  - Các bảng mô tả CSDL
  - Các bảng cấp quyền
  - Các bảng truy nhập đồng thời

Kiến trúc của một DBMS



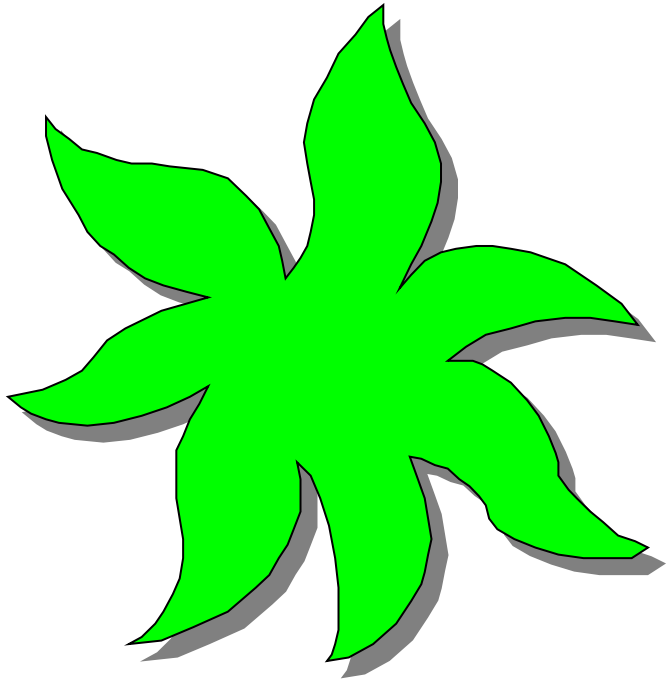
Kiến trúc của DBMS

NỘI DUNG

1. Một số khái niệm trong CSDL
2. Thiết kế CSDL
3. Ngôn ngữ SQL
4. Kiến trúc DBMS
5. Các yêu cầu bảo vệ CSDL

CÁC YÊU CẦU BẢO VỆ CSDL

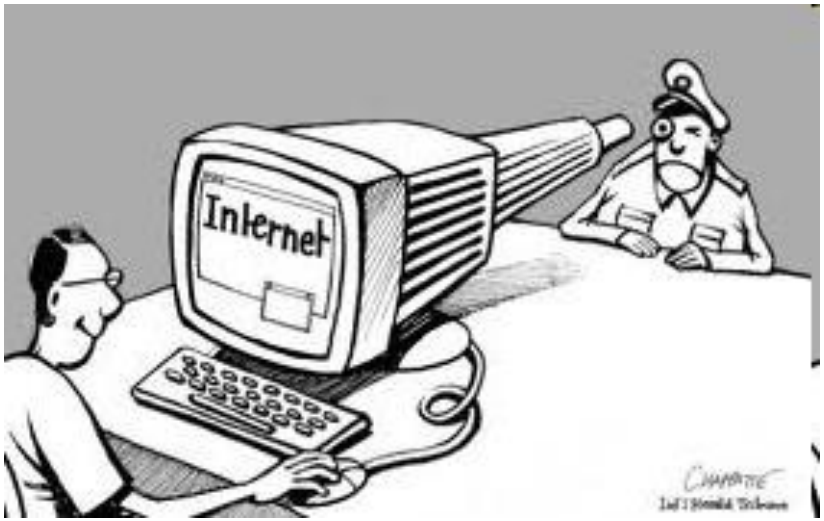
- Bảo vệ chống truy nhập trái phép
- Bảo vệ chống suy diễn
- Bảo vệ toàn vẹn CSDL
- Toàn vẹn dữ liệu thao tác
- Toàn vẹn ngữ nghĩa của dữ liệu
- Khả năng lưu vết và kiểm tra
- Xác thực người dùng
- Bảo vệ dữ liệu nhạy cảm
- Bảo vệ nhiều mức



CÁC YÊU CẦU BẢO VỆ CSDL

❖ *Bảo vệ chống truy nhập trái phép*

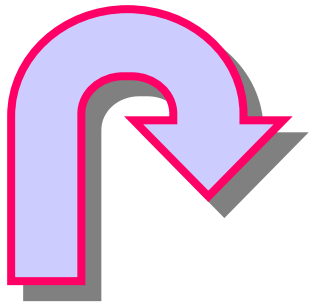
- Chỉ trao quyền cho những người dùng hợp pháp.
- Việc kiểm soát truy nhập cần tiến hành trên các đối tượng dữ liệu mức thấp hơn file: *bản ghi, thuộc tính.*

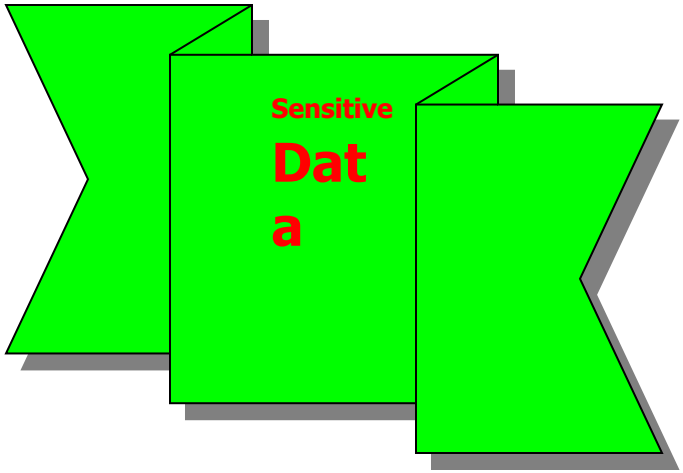


CÁC YÊU CẦU BẢO VỆ CSDL

❖ *Bảo vệ chống suy diễn:*

- Suy diễn là khả năng có được các thông tin bí mật từ những thông tin không bí mật (công khai).

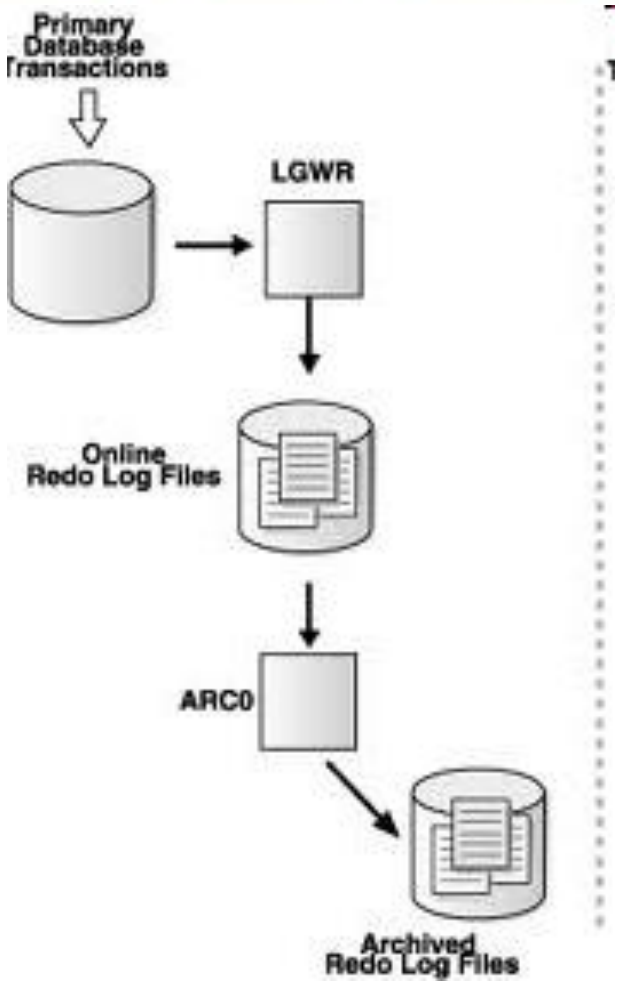




CÁC YÊU CẦU BẢO VỆ CSDL

❖ *Bảo vệ toàn vẹn CSDL*

- Bảo vệ CSDL khỏi những người dùng không hợp pháp, tránh sửa đổi nội dung dữ liệu trái phép.
- DBMS kiểm soát bằng các ràng buộc DL, thủ tục sao lưu, phục hồi và các thủ tục an toàn đặc biệt, ***file nhật ký***.



Oracle

CÁC YÊU CẦU BẢO VỆ CSDL

- ***Bảo vệ toàn vẹn CSDL***…
- Một số phương pháp đảm bảo toàn vẹn dữ liệu trong DBMS:
  - Kiểu dữ liệu (Data Type)
  - Không cho phép định nghĩa Null (Not Null Definitions)
  - Định nghĩa mặc định (Default Definitions)
  - Các thuộc tính định danh (Identity Properties)
  - Các ràng buộc (Constraints)
  - Các quy tắc (Rules)
  - Triggers
  - Các chỉ mục (Indexes)

CÁC YÊU CẦU BẢO VỆ CSDL

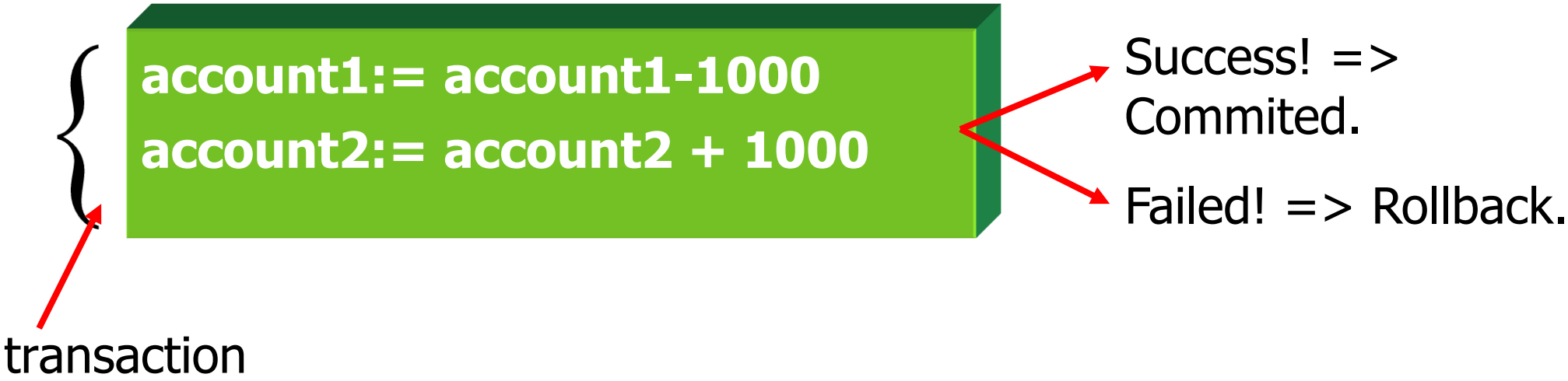
❖ *Toàn vẹn dữ liệu thao tác:*

- Yêu cầu này đảm bảo tính tương thích logic của dữ liệu khi có nhiều giao tác thực hiện đồng thời.
- *Một giao dịch (transaction):* là một loạt các hoạt động xảy ra được xem như một đơn vị công việc (unit of work) nghĩa là hoặc thành công toàn bộ hoặc không làm gì cả (all or nothing).

CÁC YÊU CẦU BẢO VỆ CSDL

❖ *Toàn vẹn dữ liệu thao tác:*

- *Ví dụ:* Chúng ta muốn chuyển một số tiền $1000 từ account1 sang account2:



CÁC YÊU CẦU BẢO VỆ CSDL

❖ *Toàn vẹn dữ liệu thao tác:*

- Ví dụ trong Oracle

Xem xét một CSDL ngân hàng. Khi khách hàng chuyển tiền từ tài khoản tiết kiệm vào tài khoản kiểm tra, giao dịch phải bao gồm 3 thao tác riêng: giảm số tiền tiết kiệm, tăng số tài khoản kiểm tra, ghi nhận giao dịch trong nhật ký.

Oracle phải bảo đảm rằng cả 3 lệnh SQL đều được thực hiện để bảo đảm tài khoản được cân đối đúng. Khi có điều gì đó ngăn cản một trong 3 lệnh trong giao dịch (như là phần cứng lỗi), thì những lệnh khác của giao dịch phải bị huỷ bỏ; quá trình này được gọi là rolling back. Nếu như lỗi xảy ra trong bất cứ thao tác cập nhật nào thì không thực hiện các thao tác cập nhật khác.

CÁC YÊU CẦU BẢO VỆ CSDL

❖ ***Toàn vẹn dữ liệu thao tác:*** Ví dụ trong Oracle

```
UPDATE savings_accounts
     SET balance = balance - 500
     WHERE account = 3209;
```

Decrement Savings Account

```
UPDATE checking_accounts
     SET balance = balance + 500
     WHERE account = 3208;
```

Increment Checking Account

```
INSERT INTO journal VALUES
     (journal_seq.NEXTVAL, '1B'
     3209, 3208, 500);
```

Record in Transaction Journal

```
COMMIT WORK;
```

End Transaction

Transaction Ends

Hình 1-5 minh hoạ ví dụ giao dịch ở ngân hàng

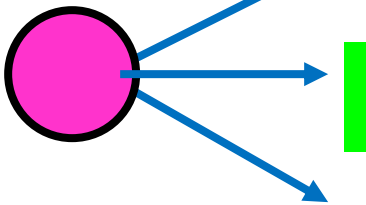
CÁC YÊU CẦU BẢO VỆ CSDL

❖ *Toàn vẹn ngữ nghĩa của dữ liệu:*

- Yêu cầu này đảm bảo tính tương thích logic của các dữ liệu bị thay đổi, bằng cách kiểm tra các giá trị dữ liệu có nằm trong khoảng cho phép hay không (đó là các ràng buộc toàn vẹn).
- *Ràng buộc (Constraints)* là những thuộc tính mà ta áp đặt lên một bảng hay một cột để tránh việc lưu dữ liệu không chính xác vào CSDL.

CÁC YÊU CẦU BẢO VỆ CSDL

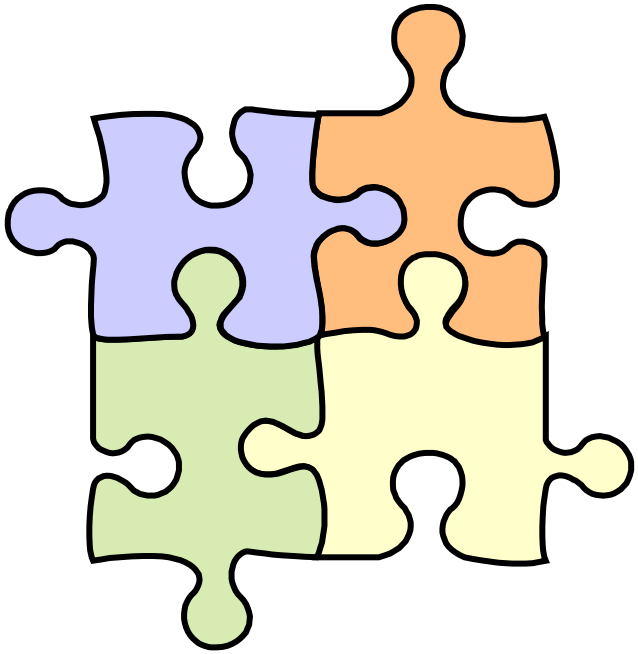
❖ ***Toàn vẹn ngữ nghĩa của dữ liệu:***



Ràng buộc khóa chính

Ràng buộc khóa ngoại

Ràng buộc kiểm tra



CÁC YÊU CẦU BẢO VỆ CSDL

- ***Toàn vẹn ngữ nghĩa của dữ liệu:***
  - Ví dụ về ràng buộc kiểm tra:

```
CREATE TABLE AT4
        (Col1   INT   PRIMARY  KEY,
         Col2   INT
         CONSTRAINT  limit_amount  CHECK
        (Col2   BETWEEN   0   AND  1000),
         Col3 VARCHAR(30)
         )
```

CÁC YÊU CẦU BẢO VỆ CSDL

- ***Bảo vệ nhiều mức (Multilevel Security):***
  - Dữ liệu được phân loại thành nhiều mức nhạy cảm.
  - *Mục đích*: là phân loại các mục thông tin khác nhau, đồng thời phân quyền cho các mức truy nhập khác nhau vào các mục riêng biệt.

CÁC YÊU CẦU BẢO VỆ CSDL

- ***Bảo vệ nhiều mức: Ví dụ***

User	C_{user}	Dept	C_{dept}	Salary	C_{salary}	TC
Bob	S	Math	S	10K	S	S
Ann	S	CIS	S	20K	TS	TS
Sam	TS	CIS	TS	30K	TS	TS

Security Levels on objects are called ***Classifications***.
Security Levels on subjects are called ***Clearances***.


QUESTION