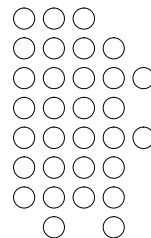


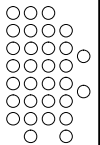
# Chương 1

## Giới thiệu Hệ quản trị cơ sở dữ liệu



## Nội dung chi tiết

- Định nghĩa DBMS
- Ba mức trừu tượng của dữ liệu
- Các đặc trưng của dữ liệu trong DBMS
- Kiến trúc DBMS
  - Ngôn ngữ giao tiếp
  - Quản lý giao tác (transaction manager)
  - Quản lý khôi phục (recovery manager)
  - Xử lý truy vấn (query processing)
  - Quản lý lưu trữ (data storage)
- Phân loại DBMS

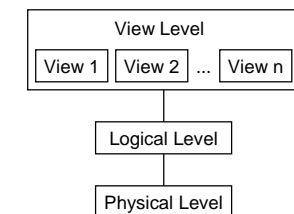


## Định nghĩa DBMS

- Là một hệ thống phần mềm cung cấp các công cụ để xây dựng và quản lý CSDL
  - Định nghĩa cấu trúc dữ liệu (DDL)
  - Cung cấp khả năng thao tác trên CSDL (MDL)
  - Hỗ trợ lưu trữ dữ liệu
  - Điều khiển truy xuất dữ liệu giữa nhiều người dùng

## Ba mức trừu tượng của dữ liệu

- Mức khung nhìn (View Level)
  - Mô tả 1 phần nào đó của toàn bộ hệ thống
- Mức luận lý (Logical level)
  - Những thông tin gì được lưu trữ trong CSDL
  - Có những mối quan hệ nào giữa những thông tin đó
- Mức vật lý (Physical Level)
  - Dữ liệu được lưu trữ như thế nào



## Ba mức trừu tượng của dữ liệu (tt)

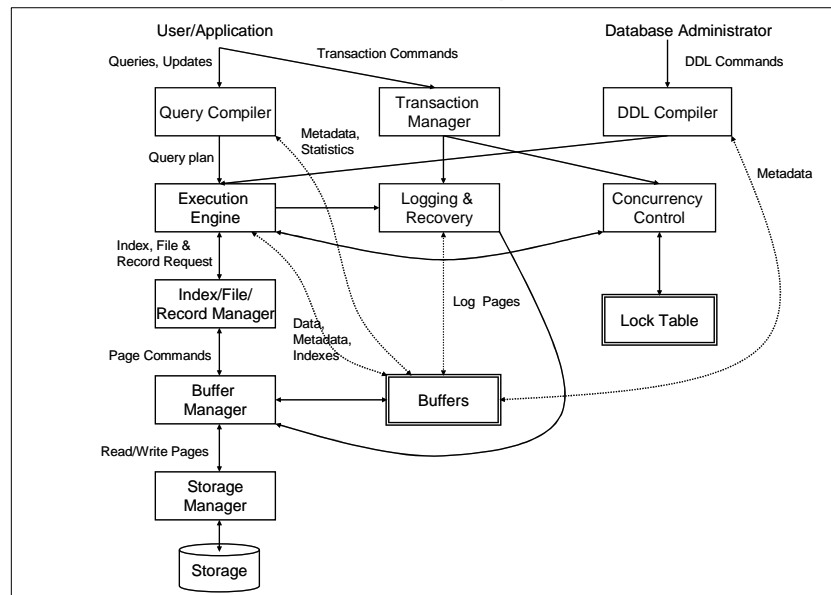
```
type customer = record
  customer-id: string;
  customer-name: string;
  customer-street: string;
  cusomter-city: string;
end;
```

- Mức vật lý
  - Mẫu tin customer là 1 block các byte liên tiếp nhau
- Mức luận lý
  - customer và các mối liên kết được định nghĩa
- Mức khung nhìn
  - Người sử dụng chỉ được truy xuất đến 1 phần thông tin của customer

## Các đặc trưng của dữ liệu

- Có tính ít dư thừa dữ liệu
- Chia sẻ cho nhiều người sử dụng
- Có tính an toàn, bảo mật
- Cho phép truy xuất đồng thời
- Có thể khôi phục khi có sự cố
- Có tính độc lập
  - Vật lý (physical)
    - Khi thay đổi tổ chức vật lý của CSDL thì không làm ảnh hưởng đến chương trình đang khai thác CSDL đó
  - Luận lý (logical)
    - Khi thay đổi mức quan niệm không làm ảnh hưởng đến các chương trình đang khai thác CSDL

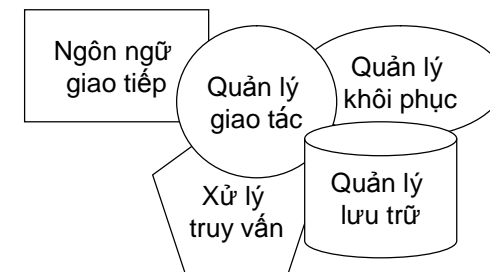
# Kiến trúc của DBMS



Hệ Quản Trị CSDL - Khoa CNTT - ĐH CNTT Tp.HCM

7

# Các thành phần chính của DBMS

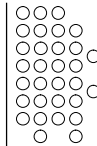


Hệ Quản Trị CSDL - Khoa CNTT - ĐH CNTT Tp.HCM

8

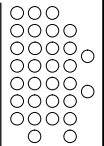
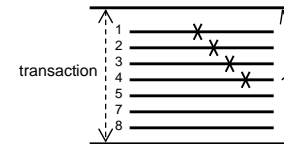
## Ngôn ngữ giao tiếp

- DBMS cung cấp giao diện lập trình để sử dụng với một ngôn ngữ lập trình CSDL
  - SQL Server: Transaction SQL (T SQL)
  - Oracle: PL/SQL
- Ngôn ngữ bao gồm
  - Định nghĩa dữ liệu
  - Thao tác dữ liệu



## Quản lý giao tác

- Thành phần quản lý các giao tác có ảnh hưởng đến CSDL
  - Giao tác là một nhóm các hành động mà nếu thực hiện được thì phải thực hiện hết tất cả các hành động trong giao tác đó, ngược lại xem như không thực hiện hành động nào

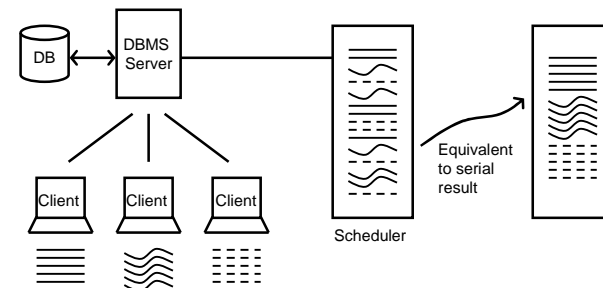


## Quản lý giao tác (tt)

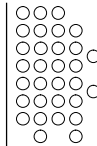
- Nhật ký
  - Để CSDL được bền vững (durable), mọi thay đổi lên CSDL phải được ghi nhận lại
- Log manager – ghi chép nhật ký
  - Đảm bảo CSDL vẫn nguyên vẹn khi có sự cố xảy ra
- Recovery manager – khôi phục
  - Dựa vào nhật ký để phục hồi lại CSDL về trạng thái nhất quán trước đó
    - Trạng thái thỏa tất cả các RBTV của CSDL đó

## Quản lý giao tác (tt)

- Điều khiển đồng thời
  - Bộ lập lịch (scheduler)- có nhiệm vụ lập 1 lịch thực hiện từ  $n$  giao tác được kích hoạt đồng thời
  - Cơ chế khóa (lock)- ngăn 2 giao tác cùng thao tác lên 1 đơn vị dữ liệu tại 1 điểm

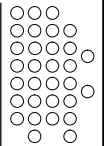


## Quản lý giao tác (tt)



- Giải quyết deadlock
  - Vì sử dụng cơ chế khóa nên các giao tác sẽ phải tranh giành tài nguyên
  - Tình huống “không một giao tác nào có thể thực hiện được công việc của mình”
    - Các giao tác chờ đợi lẫn nhau để được cấp phát tài nguyên
  - Thành phần quản lý giao tác sẽ phải can thiệp vào
    - Rollback
    - Abort

## Xử lý truy vấn



- Biểu diễn câu truy vấn ở dạng ngôn ngữ cấp cao (SQL) và thực hiện câu truy vấn có hiệu quả
- Query compiler – biên dịch
  - Query parser
    - Xây dựng cấu trúc hình cây từ câu truy vấn
  - Query preprocessor
    - Kiểm tra ngữ nghĩa của câu truy vấn
    - Chuyển đổi cấu trúc cây sang ngôn ngữ đại số quan hệ
  - Query optimizer
    - Sắp xếp các phép toán nhằm mục đích tối ưu hóa câu truy vấn

## Quản lý lưu trữ

- Thành phần có nhiệm vụ điều khiển việc đọc/ghi dữ liệu qua lại giữa bộ nhớ và thiết bị lưu trữ
- Làm việc với các khái niệm
  - Tập tin dữ liệu
  - Từ điển dữ liệu
    - Lưu trữ các metadata về cấu trúc của CSDL, đặc biệt là lược đồ của CSDL
  - Chỉ mục

## Phân loại DBMS

- Theo mô hình dữ liệu
  - Mạng (Network)
  - Phân cấp (Hierarchical)
  - Quan hệ (Relation)
  - Hướng đối tượng (Oriented Object)
  - XML
- Theo kiến trúc
  - Đơn người dùng
  - Client/Server
- Khác
  - Multimedia
  - Hypertext



## Mô hình mạng

- Dữ liệu ~ tập các record
- Mỗi quan hệ giữa các dữ liệu ~ các link
- Các record được tổ chức thành tập các đồ thị tùy ý

| TênKH   | MãKH        | ĐịaChỉ | ThànhPhố   | MãTK  | SốDư | TênCN      | ThànhPhố  | TổngTiền |
|---------|-------------|--------|------------|-------|------|------------|-----------|----------|
| Johnson | 192-83-7465 | Alma   | Palo Alto  | A-101 | 500  | Perryridge | Horseneck | 1700000  |
| Smith   | 019-28-3746 | North  | Rye        | A-215 | 700  | Downtown   | Brooklyn  | 9000000  |
| Hayes   | 677-89-9011 | Main   | Harison    | A-102 | 400  | Round Hill | Horseneck | 8000000  |
| Tuner   | 182-73-6091 | Putnam | Stamford   | A-305 | 350  | Mianus     | Woodside  | 1000000  |
| Jones   | 321-12-3123 | Main   | Harison    | A-201 | 900  |            |           |          |
|         |             |        |            | A-217 | 750  | Redwood    | Princeton | 2000000  |
| Lindsay | 336-66-9999 | Park   | Pittsfield | A-222 | 700  | Brighton   | Brooklyn  | 7000000  |

Khách hàng

Tài khoản

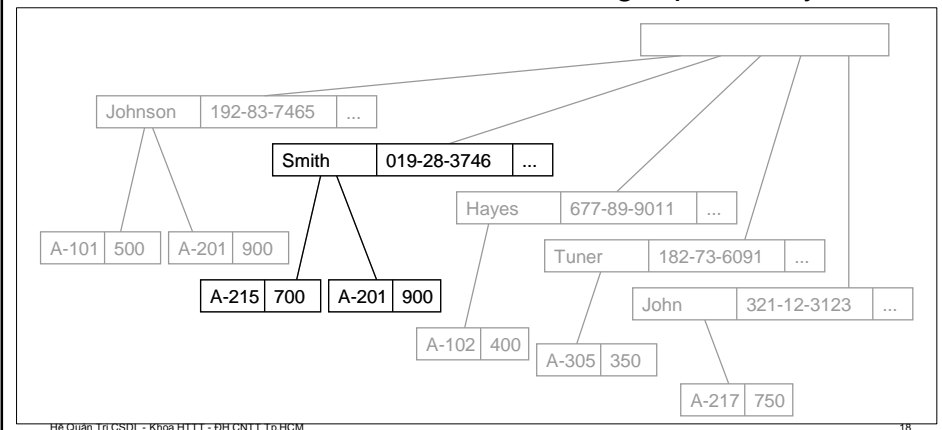
Chi nhánh

Hệ Quản Trị CSDL - Khoa CNTT - ĐH CNTT Tp.HCM

17

## Mô hình phân cấp

- Dữ liệu và mối quan hệ được biểu diễn bằng các record và link
- Các record được tổ chức dưới dạng tập các cây



Hệ Quản Trị CSDL - Khoa CNTT - ĐH CNTT Tp.HCM

18

## Mô hình quan hệ

- Sử dụng bảng 2 chiều (quan hệ) để biểu diễn cho dữ liệu và mối quan hệ

| TênKH   | MãKH        | ĐịaChỉ | ThànhPhố   | MãTK  |
|---------|-------------|--------|------------|-------|
| Johnson | 192-83-7465 | Alma   | Palo Alto  | A-101 |
| Smith   | 019-28-3746 | North  | Rye        | A-215 |
| Hayes   | 677-89-9011 | Main   | Harison    | A-102 |
| Tuner   | 182-73-6091 | Putnam | Stamford   | A-305 |
| Jones   | 321-12-3123 | Main   | Harison    | A-217 |
| Lindsay | 336-66-9999 | Park   | Pittsfield | A-222 |
| Johnson | 192-83-7465 | Alma   | Palo Alto  | A-201 |

Khách hàng

| MãTK  | SốDư |
|-------|------|
| A-101 | 500  |
| A-215 | 700  |
| A-102 | 400  |
| A-305 | 350  |
| A-201 | 900  |
| A-217 | 750  |
| A-222 | 700  |

Tài khoản

## Mô hình hướng đối tượng

- Dữ liệu và mối quan hệ được biểu diễn bằng khái niệm đối tượng
- Cấu trúc 1 đối tượng
  - Biến (variables)
  - Thông điệp (messages)
  - Phương thức (methods)
- Các khái niệm của CSDL hướng đối tượng
  - Lớp (class)
  - Định danh đối tượng (Object Identity)
  - Đóng gói (Encapsulation)
  - Kế thừa (Inheritance)
- Ngôn ngữ định nghĩa đối tượng - ODL
- Ngôn ngữ thao tác đối tượng - OML

## Mô hình hướng đối tượng (tt)

```
class Customer: public Person {  
    public:  
        d_Date member_from;  
        d_Long customer_id;  
        d_Ref<Branch> home_branch;  
};
```

```
class Person: public d_Object {  
    public:  
        d_String name;  
        d_String address;  
};
```

```
class Account: public d_Object {  
    private:  
        d_Long balance;  
    public:  
        d_Long number;  
        d_Set <d_Ref<Customer>> owners;  
        int find_balance();  
        int update_balance(int delta);  
};
```

## XML

- Dữ liệu và mối quan hệ được biểu diễn thông qua ngôn ngữ đánh dấu (markup language)
- Cấu trúc của dữ liệu XML
  - Thẻ <tag>
  - Element
    - 1 cặp thẻ mở/đóng <tag> </tag>
    - Dữ liệu dưới dạng text
- Ngôn ngữ truy vấn – Xquery, Xpath
- Ngôn ngữ định nghĩa – DTD, XML Schema (XSD)

## XML (tt)

```
<bank-2>
  <account account-number="A-401" owners="C100 C102">
    <branch-name>Downtown</branch-name>
    <balance>500</balance>
  </account>
  <customer customer-id="C100" accounts="A-401">
    <customer-name>Joe</customer-name>
    <customer-street>Monroe</customer-street>
    <customer-city>Madison</customer-city>
  </customer>
  <customer customer-id="C102" accounts="A-401 A-402">
    <customer-name>Mary</customer-name>
    <customer-street>Erin</customer-street>
    <customer-city>Newark</customer-city>
  </customer>
</bank-2>
```

