

Tổng quan về HQT CSDL

Nội dung chi tiết

- * *Yêu cầu về Dữ liệu trong CSDL*
- * *Khái niệm Hệ Quản trị CSDL*
- * *Kiến trúc của một HQT CSDL*
- * *Phân loại HQT CSDL*

Yêu cầu về Dữ liệu trong CSDL

*** Dữ liệu trong CSDL cần thể hiện qua 3 cấp độ sau**

❖ Mức quan niệm (Conceptual Level)

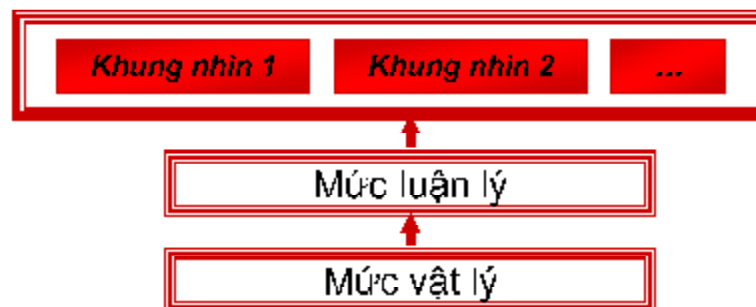
- ⊛ Mô tả 1 phần của CSDL mà 1 đối tượng người dùng được quyền tiếp cận

❖ 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 trên thiết bị lưu trữ



Yêu cầu về Dữ liệu trong CSDL (tt)

*** Dữ liệu trong CSDL cần có các đặc trưng sau**

❖ Ít hoặc không trùng lặp dữ liệu

❖ Chia sẻ cho nhiều người sử dụng → Truy xuất đồng thời

❖ An ninh, bảo mật

❖ Khôi phục khi có sự cố

❖ 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 cần thay đổi tổ chức vật lý đang lưu trữ dữ liệu của mức quan niệm ấy

*** Vì vậy cần có một hệ thống quản lý hiệu quả Dữ liệu trong CSDL**

- * ***Yêu cầu về Dữ liệu trong CSDL***
- * ***Khái niệm Hệ Quản trị CSDL***
- * ***Kiến trúc của một HQT CSDL***
- * ***Phân loại HQT CSDL***

Khái niệm DBMS

- * ***Là hệ thống phần mềm cung cấp các công cụ để xây dựng, khai thác và quản lý CSDL***
 - ❖ Định nghĩa cấu trúc dữ liệu (DDL – Data Definition Language)
 - ❖ Cung cấp khả năng thao tác trên CSDL (DML – Data Manipulation Language)
 - ⊕ Cập nhật Dữ liệu
 - ⊕ Truy vấn Dữ liệu
 - ❖ Hỗ trợ lưu trữ dữ liệu
 - ❖ Điều khiển truy xuất dữ liệu đồng thời
 - ❖ Bảo mật và an ninh Dữ liệu
- * ***DBMS là đối tượng nghiên cứu của môn học***

Nội dung chi tiết

- * ***Yêu cầu về Dữ liệu trong CSDL***
- * ***Khái niệm Hệ Quản trị CSDL***
- * ***Kiến trúc của một HQT CSDL***
- * ***Phân loại HQT CSDL***

Các thành phần chính của DBMS



Chào diện lập trình

- * **DBMS cung cấp giao diện lập trình để sử dụng với một ngôn ngữ lập trình CSDL**
 - ❖ Vd : SQL Server cung cấp ngôn ngữ Transaction-SQL (T-SQL)
- * **Ngôn ngữ bao gồm**
 - ❖ Ngôn ngữ định nghĩa dữ liệu : Giúp người dùng ra lệnh cho DBMS tạo ra các cấu trúc Dữ liệu của CSDL (cách tổ chức dữ liệu và mối liên hệ giữa các đối tượng dữ liệu)
 - ❖ Ngôn ngữ thao tác dữ liệu : giúp người dùng tích lũy, hiệu chỉnh và khai thác Dữ liệu

An ninh và Bảo mật

- * **Bộ phận bảo đảm An toàn và Bảo mật thực hiện hai nhiệm vụ**
 - ❖ Kiểm soát, quản lý tốt những người dùng hợp pháp của hệ thống → Bảo mật Dữ liệu
 - ❖ Ngăn chặn các tấn công của những đối tượng tin tặc → An ninh Dữ liệu

Khôi phục sau sự cố

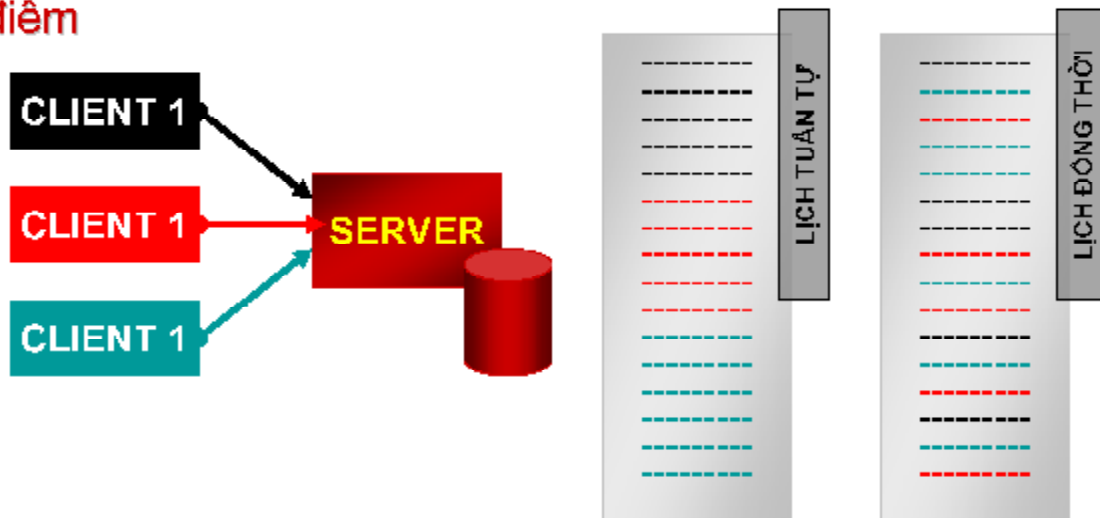
- * ***Đảm bảo tổn thất, sai sót về dữ liệu là ít nhất có thể***
- * ***Nhật ký***
 - ❖ Để CSDL được bền vững (durable), mọi thay đổi lên CSDL phải được ghi nhận lại (trong Nhật ký)
 - ❖ Log manager – ghi chép nhật ký
 - ⊛ Đảm bảo ghi nhận đầy đủ và chính xác mọi thay đổi trên CSDL
 - ❖ Recovery manager – 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)

Xử lý truy xuất đồng thời

- * ***Đối tượng của thành phần quản lý giao tác là các Giao tác***
 - ❖ Giao tác là một nhóm các hành động mà hoặc được thực hiện hết toàn bộ, hoặc không thực hiện bất kỳ hành động nào trong nhóm ấy.

* Điều khiển đồng thời

- ❖ Bộ lập lịch (scheduler) - có nhiệm vụ lập 1 lịch thực hiện từ n giao tác không tách biệt về thời gian sao cho kết quả không vi phạm tính nhất quán của CSDL.
- ❖ Cơ chế khóa (lock) trên 1 đơn vị dữ liệu nào đó → ngăn 2 giao tác cùng thao tác lên 1 đơn vị dữ liệu ấy tại cùng 1 điểm



Xử lý truy xuất đồng thời (tt)

* Giải quyết deadlock

- ❖ Vì sử dụng cơ chế khóa nên các giao tác sẽ phải chờ khi cần truy xuất 1 đơn vị dữ liệu đang bị khóa.
- ❖ Tình huống chờ vĩnh viễn mà vẫn không được truy xuất đơn vị dữ liệu bị khóa gọi là Deadlock (khóa chết)
 - ⊛ Các giao tác chờ đợi lẫn nhau để được cấp phát tài nguyên và không giao tác nào có thể hoàn tất
- ❖ Thành phần quản lý giao tác sẽ phải can thiệp vào
 - ⊛ Hoặc hủy bỏ một trong các giao tác gây deadlock
 - ⊛ Hoặc ngăn chặn từ trước để không bao giờ xảy ra deadlock

Xu 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 truy vấn**
 - ❖ **Query parser**
 - ⊗ Xây dựng cấu trúc phân tích câu truy vấn dưới dạng cây
 - ❖ **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 (Siêu dữ liệu) về cấu trúc của CSDL, đặc biệt là lược đồ của CSDL
 - ❖ **Chỉ mục**
 - ⊗ Giúp cho việc tìm kiếm Dữ liệu được nhanh chóng

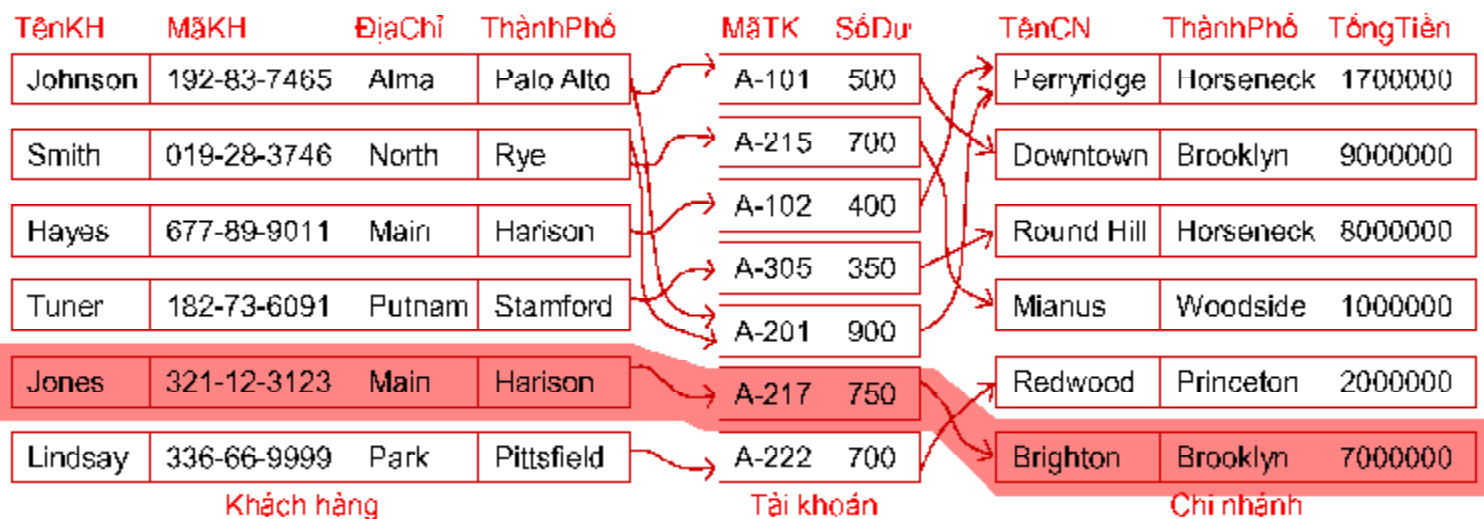
- * ***Yêu cầu về Dữ liệu trong CSDL***
- * ***Khái niệm Hệ Quản trị CSDL***
- * ***Kiến trúc của một HQT CSDL***
- * ***Phân loại HQT CSDL***

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
 - ❖ Tập trung / Phân tán
 - ❖ Hệ chịu lỗi tốt (fault-tolerant)

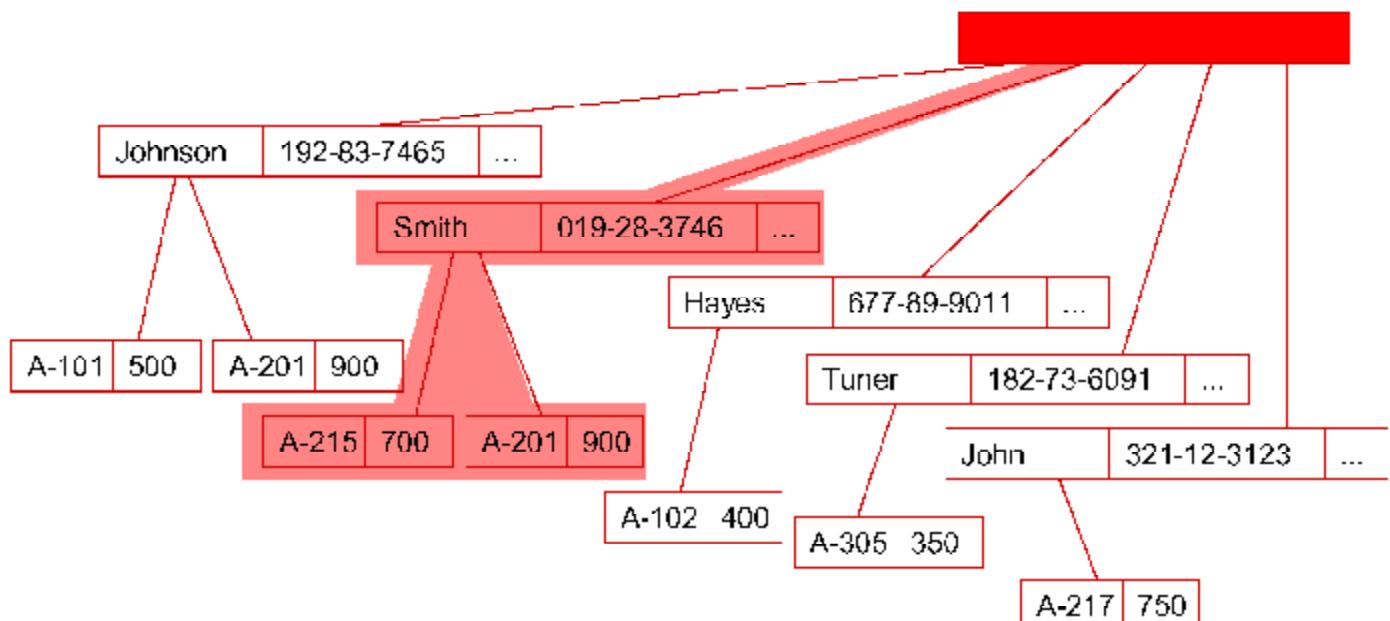
Mô hình mạng

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



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**



Mô hình quan hệ

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

TênKH	MãKH	ĐịaChí	ThànhPhố	MãTK	MãTK	SốDư
Johnson	192-83-7465	Alma	Palo Alto	A-101	A-101	500
Smith	019-28-3746	North	Rye	A-215	A-215	700
Hayes	677-89-9011	Main	Harison	A-102	A-102	400
Tuner	182-73-6091	Putnam	Stamford	A-305	A-305	350
Jones	321-12-3123	Main	Harison	A-217	A-201	900
Lindsay	336-66-9999	Park	Pittsfield	A-222	A-217	750
Johnson	192-83-7465	Alma	Palo Alto	A-201	A-222	700

Khách hàngTài 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**

- * *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>
 - ❖ Thành phần (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*

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>

```

