Chương 1.

1. Định nghĩa Cơ sở dữ liệu (database)?

Một tập hợp các dữ liệu có liên quan đến nhau trong một lĩnh vực cụ thể

1. Các chức năng của hệ quản trị CSDL

Dùng điều khiển thông tin,tạo,cập nhật,và duy trì cho cơ sở dữ liệu

1. Ngôn ngữ định nghĩa dữ liệu là gì? Cho ví dụ.

DDL (Data Definition Language) là tập hợp các câu lệnh thay đổi cấu trúc của bảng. Tất cả các thay đổi đều sẽ được lưu vĩnh viễn trong cơ sở dữ liệu

Ví dụ : CREATE , ALTER , DROP , TRUNCATE , RENAME

1. Ngôn ngữ thao tác dữ liệu là gì? Cho ví dụ

DML (Data Manipulation Language) là tập hợp các lệnh dùng để sửa đổi cơ sở dữ liệu. Nó chịu trách nhiệm cho các thay đổi về hình thức trong cơ sở dữ liệu. Tất cả các thay đổi không thể lưu vĩnh viễn trong cơ sở dữ liệu. Có nghĩa là chúng ta có thể khôi phục

Ví dụ : SELECT , INSERT , UPDATE , DELETE

1. Ngôn ngữ kiểm soát dữ liệu và ngôn ngữ kiểm soát transaction? cho ví dụ.

Ngôn ngữ kiểm soát dữ liệu dùng để quản lí quyền truy cập và kiểm soát quyền đối với dữ liệu

Ví dụ : GRANT và REVOKE.GRANT

Ngôn ngữ kiểm soát transaction

TCL (Transaction Control Language) chỉ có thể sử dụng với các lệnh DML.

Một số câu lệnh trong TCL:

* Commit: Câu lệnh được dùng để lưu tất cả các giao dịch vào cơ sở dữ liệu.
* Rollback: lệnh được dùng để hoàn tác các giao dịch chưa được lưu vào cơ sở dữ liệu.
* Savepoint: lệnh được dùng để chuyển giao dịch trở lại một điểm nhất định mà không cần quay lại toàn bộ giao dịch.

1. Nhiệm vụ của công cụ kiểm soát, điều khiển truy cập vào CSDL?

-Duy trì tính nhất quán của CSDL

-Khôi phục CSDL khi có sự cố   
-Phát hiện và ngăn chặn sự truy cập không cho phép

1. Các loại đối tượng sử dụng CSDL?

+Người quản trị CSDL(DBA-DataBase Administrator)

+Người thiết kế CSDL(DataBase designer)

+Người dùng cuối

+Phân tích viên hệ thống và lập trình viên ứng dụng

1. Hệ quản trị CSDL (DBMS) là gì?

Một gói / hệ thống phần mềm giúp cho việc tạo và duy trì cơ sở dữ liệu trên máy tính một cách thuận lợi.

1. Kể tên một vài DBMS thông dụng

SQL-Server, DB2, Paradox, Informix, Oracle, …

1. Một hệ CSDL?

Bao gồm một DBMS cùng với một CSDL. Đôi khi có thể bao gồm cả chương trình ứng dụng

1. Siêu dữ liệu (meta-data)?

Hệ CSDL không chỉ lưu CSDL mà còn lưu trữ định nghĩa về cấu trúc dữ liệu và các ràng buộc trong catalog. Thông tin được lưu trữ trong catalog được gọi là siêu dữ liệu.

1. Ưu điểm của DBMS

Dữ liệu độc lập và truy cập hiệu quả

Giảm thời gian thực thi ứng dụng

Tính toàn vẹn dữ liệu (data integrity) và bảo mật

Quản lí dữ liệu thống nhất

Truy cập đồng thời , và phục hồi dữ liệu khi có sự cố

1. Mô hình dữ liệu (data model) là gì?

Một tập hợp các khái niệm dùng để mô tả cấu trúc của cơ sở dữ liệu.

1. Database designer có nhiệm vụ?

- Hiểu yêu cầu của người dùng và tạo một thiết kế đáp ứng yêu cầu.

- Xác định dữ liệu cần lưu trong CSDL và cấu trúc CSDL

1. System Analyst có nhiệm vụ?

Phân tích viên hệ thống xác định yêu cầu của người dùng và xây dựng  
 bản đặc tả cho những giao tác đáp ứng yêu cầu người dùng.

1. Application Programmer có nhiệm vụ?

Lập trình viên cài đặt đặc tả thành chương trình, kiểm tra, dò lỗi, ghi  
sưu liệu và bảo trì những giao tác này

1. Lược đồ (schema) là gì?

Là sự mô tả về một tập hợp của dữ liệu cụ thể , sử dụng mô hình dữ liệu được cho

1. Phân loại mô hình dữ liệu

- Mô hình dữ liệu mức cao (mức khái niệm): cung cấp các khái niệm gần với cách người dùng cảm nhận về dữ liệu.

Vd. Mô hình thực thể kết hợp là mô hình dữ liệu mức khái niệm

- Mô hình dữ liệu mức thấp (mức vật lý): mô tả cách dữ liệu được lưu trữ trên máy tính.

Vd. File trên máy tính được mô tả bởi định dạng file, chỉ mục, …

- Mô hình dữ liệu mức logic: mức trung gian giữa 2 mức trên để người dùng có thể hiểu nhưng cũng gần với cách dữ liệu được tổ chức trên đĩa cứng. Mô hình này dấu đi chi tiết lưu trữ dữ liệu trên đĩa nhưng có thể cài đặt trên máy tính.

Vd. Mô hình quan hệ, mô hình phân cấp, mô hình mạng là các mô hình dữ liệu mức logic.

1. Mô hình dữ liệu mức khái niệm?

Mô hình dữ liệu mức cao (mức khái niệm): cung cấp các khái niệm gần với cách người dùng cảm nhận về dữ liệu.

1. Mô hình dữ liệu mức vật lý ?

Mô hình dữ liệu mức thấp (mức vật lý): mô tả cách dữ liệu được lưu trữ trên máy tính.

1. Mô hình dữ liệu mức logic ?

Mô hình dữ liệu mức logic: mức trung gian giữa 2 mức trên để người dùng có thể hiểu nhưng cũng gần với cách dữ liệu được tổ chức trên đĩa cứng. Mô hình này dấu đi chi tiết lưu trữ dữ liệu trên đĩa nhưng có thể cài đặt trên máy tính.

1. Kiến trúc 3 mức trừu tượng của một hệ CSDL? Công dụng của nó?

Theo kiến trúc ANSI-PARC, một CSDL có 3 mức biểu diễn

Kiến trúc 3 mức giúp tách biệt các ứng dụng người dùng với CSDL vật lý.

1. Mức vật lý trong kiến trúc 3 mức của 1 hệ CSDL?

Mức vật lý: Sử dụng mô hình dữ liệu vật lý để mô tả cấu trúc lưu trữ vật lý của CSDL

1. Mức logic trong kiến trúc 3 mức trừu tượng của 1 hệ CSDL?

Mức logic/quan niệm: Dấu đi chi tiết về cấu trúc lưu trữ vật lý. Dùng mô hình dữ liệu logic để mô tả cái gì được lưu trữ trong CSDL và mối quan hệ giữa các dữ liệu đó.

1. Mức view trong kiến trúc 3 mức trừu tượng của 1 hệ CSDL?

Mức ngoài/ view: Mô tả một phần của CSDL cho một nhóm người dùng quan tâm và dấu đi phần còn lại của CSDL khỏi nhóm người dùng đó.

1. Phân loại Tính độc lập dữ liệu ?

Tính độc lập vật lí và tính độc lập logic

1. Tính độc lập vật lý?

là khả năng thay đổi lược đồ vật lý mà không phải thay đổi lươc đồ quan niệm. Vì vậy lược đồ ngoài cũng không cần phải thay đổi.

1. Tính độc lập logic là gì?

là khả năng thay đổi lược đồ quan niệm mà không phải thay đổi lược đồ ngoài hay chương trình ứng dụng.

**Assigment:**

1. Why would you choose a database system instead of simply storing data in operating system files? When would it make sense not to use a database system?

Uư điểm của database system

Dữ liệu độc lập và truy cập hiệu quả

Giảm thời gian thực thi ứng dụng

Tính toàn vẹn dữ liệu (data integrity) và bảo mật

Quản lí dữ liệu thống nhất

Truy cập đồng thời , và phục hồi dữ liệu khi có sự cố

Nhược điểm file based system :

+giảm khả năng sử dụng lại

+ Có vấn đề về truy cập dữ liệu , tính riêng biệt của dữ liệu

+Khi nhiều người truy xuất cùng một lúc thì sẽ gây ra nhiều vấn về

+Vấn đề về bảo mật

+Vấn đề về tính toàn vẹn dữ liệu

Đối với những ứng dụng nhỏ , ứng dụng chỉ có một người dùng dữ liệu ít(đơn giản) thì chỉ cần lưu dữ liệu trong file của hệ điều hành của máy tính

1. What is logical data independence and why is it important?

khả năng thay đổi lược đồ quan niệm mà không phải thay đổi lược đồ ngoài hay chương trình ứng dụng.

Nó quan trọng bởi về khi thay đổi dữ liệu sẽ không ảnh hưởng đến chương trình

3 Explain the difference between logical and physical data independence.

-Tính độc lập vật lí :

là khả năng thay đổi lược đồ vật lý mà không phải thay đổi lươc đồ quan niệm. Vì vậy lược đồ ngoài cũng không cần phải thay đổi.

-Tính độc lập logic

là khả năng thay đổi lược đồ quan niệm mà không phải thay đổi lược đồ ngoài hay chương trình ứng dụng.

4.What are the responsibilities of a DBA? If we assume that the DBA is never interested in running his or her own queries, does the DBA still need to understand query optimization? Why?

Công việc của DBA

+Thiết kế lược đồ vật lí và lược đồ logic , cũng như các phần được sử dụng của lược đồ ngoài

+Bảo mật và cho phép truy cập

+Dữ liệu sẵn có , và phục hồi dữ liệu khi có lỗi

+Điều chỉnh cơ sở dữ liệu(Database tuning) : người quản trị cơ sở dữ liệu có trách nhiệm phát triển(evolve), cụ thể là lược đồ logic và lược đồ vật lí , đảm bảo hiệu suất đầy đủ khi nhu cầu người dùng thay đổi

-Người quản trị cơ sở dữ liệu cần hiểu sự tối ưu truy vấn kể cả khi họ không hứng thú việc chạy truy vấn của họ bởi vì trách nhiệm của họ ( thiết kế và điều chỉnh cơ sở dữ liệu) có liên quan đến tối ưu truy vấn . Nếu người quản trị không hiểu sự cần thiết về hiệu suất của những câu truy vấn , làm thế nào để hệ quản trị cơ sở dữ liệu tối ưu và thực thi những truy vấn đó , quyết định thiết kế và điều chỉnh tốt sẽ không có .

**5.** Scrooge McNugget wants to store information (names, addresses, descriptions of embarrassing moments, etc.) about the many ducks on his payroll. Not surprisingly, the volume of data compels him to buy a database system. To save money, he wants to buy one with the fewest possible features, and he plans to run it as a stand-alone application on his PC clone. Of course, Scrooge does not plan to share his list with anyone. Indicate which of the following DBMS features Scrooge should pay for; in each case, also indicate why Scrooge should (or should not) pay for that feature in the system he buys.

1. A security facility.

Cần mua bởi vì McNugget không muốn share list của mình cho bất cứ ai

1. Concurrency control.

Không cần mua bởi vì chỉ có mỗi ông McNugget truy cập nên không cần đến tính năng truy cập đồng thời

1. Crash recovery.

Cần mua để khi mất dữ liệu vẫn có thể khôi phục.McNugget sẽ không bị mất dữ liệu nếu máy bị ngắt lúc anh ấy sử dụng thiết bị

1. A view mechanism.

Cơ chế xem là cần thiết . McNugget có thể sử dụng nó để phát triển “custom screens” để anh ấy có thể hiển thị nó lên mà không cần phải lặp lại đoạn truy vấn dài

1. A query language.

Cần thiết . McNugget có thể sử dụng để phân tích bệnh nhân của anh ấy . Cụ thể ngôn ngữ truy vấn được dùng để truy vấn thông tin bệnh nhân của anh ấy.

6. Describe the structure of a DBMS. If your operating system is upgraded to support some new functions on OS files (e.g., the ability to force some sequence of bytes to disk), which layer(s) of the DBMS would you have to rewrite to take advantage of these new functions?

Cấu trúc của DBMS :

Một lớp quản lí dung lượng lưu trữ trong ổ cứng (disk)

Một lớp quản lí dung lượng bộ nhớ ( main memory ) có thể sử dụng và đem dung lượng của ổ cứng cho memory khi cần

Một lớp hỗ trợ sự trừu tượng của files và cấu trúc chỉ mục(index structures)

Một lớp hỗ trợ các toán tử quan hệ

Một lớp phân tích cú pháp tối ưu hóa truy vấn và tạo ra kế hoạch thực hiện theo toán tử quan hệ

Lớp quản lí dung lượng lưu trữ trong ổ cứng(disk) , lớp quản lí dung lượng bộ nhớ (main memory) bị ảnh hưởng khi hệ điều hành update