**Câu 1. Trình bày khái niệm cơ sở dữ liệu phân tán ? anh/chị hãy cho biết các loại truy xuất cơ sở dữ liệu phân tán ?**

CSDL phân tán là một tập hợp dữ liệu về mặt logic của cùng một hệ thống nhưng được phân tán trên các site của mạng máy tính;

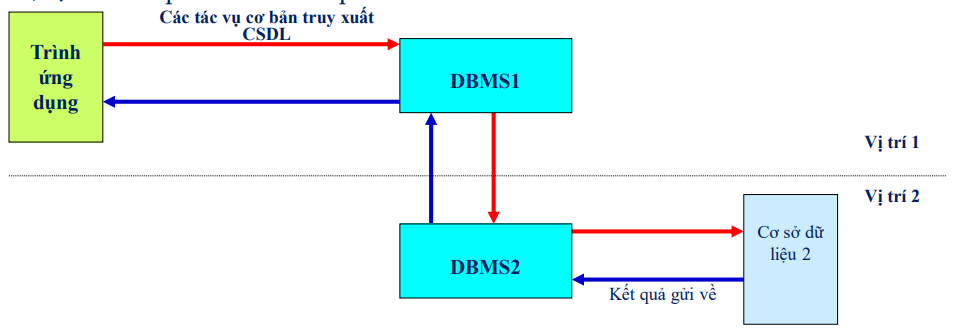
Định nghĩa này nhấn mạnh 2 khía cạnh quan trọng:

* Phân tán: Dữ liệu không cư trú tại cùng một site. (Đó là khác biệt cơ bản so với CSDL tập trung).
* Tương quan logic: dữ liệu có một số thuộc tính gắn kết chúng lại với nhau. (Điều này có thể phân biệt CSDL phân tán với một tập hợp cơ sở dữ liệu tập trung hoặc các file được lưu trữ tại các site khác nhau của mạng máy tính).

**\***Các loại truy xuất CSDL phân tán:

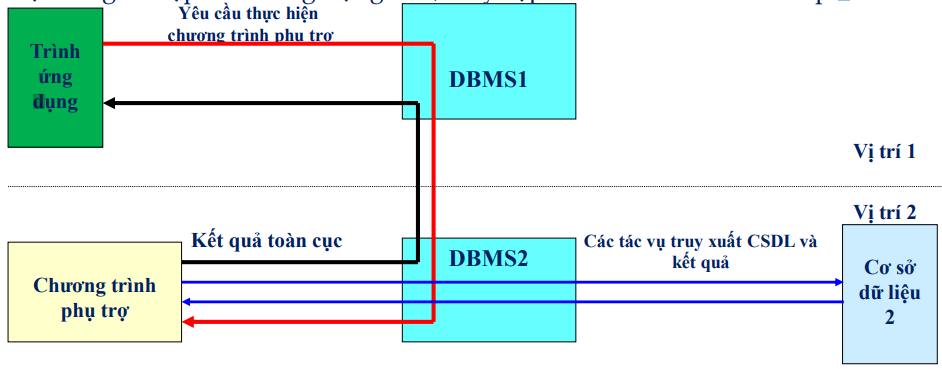
a) Truy cập từ xa thông qua tác vụ cơ bản:

* Ứng dụng đưa ra yêu cầu truy cập CSDL đề cập đến dữ liệu từ xa
* Yêu cầu này được hệ quản trị CSDL phân tán tự động định tuyến đến site chứa dữ liệu.
* Sau đó, tại site đó phát hành kết quả trả về



b) Truy cập từ xa thông qua một chương trình phụ trợ

* Ứng dụng yêu cầu thực hiện một chương trình phụ trợ tại site từ xa
* Chương trình phụ trợ do người lập trình ứng dụng viết, truy cập CSDL từ xa và trả kết quả cho ứng dụng yêu cầu



Câu 2. Anh/chị hãy trình bày sơ đồ thiết kế tổng thể cơ sở dữ liệu phân tán ? cho ví dụ ?

Sơ đồ tổng thể (Global Schema):

* Xác định tất cả dữ liệu được chứa trong CSDL phân tán như các CSDL không được phân tán.
* Sơ đồ tổng thể có thể định nghĩa theo cách tương tự như trong CSDL không được phân tán.
* Trong mô hình CSDL quan hệ (relational), sơ đồ tổng thể bao gồm định nghĩa các quan hệ tổng thể (global relations).

Câu 3. Trình bày các tính chất của giao tác phân tán ? Với trình tự thực hiện giao tác phân tán sau, anh/chị hãy cho biết kết quả của x và y ?

**Tính nguyên tử**

Tính nguyên tử của một giao tác là sự thực hiện trọn vẹn mà không một giao tác nào được chen vào.

Khi thực thi một giao tác thì hoặc là các hành động của giao tác đó được thực hiện hoặc là không một hành động nào được thực hiện cả.

Tính nguyên tử đòi hỏi rằng nếu việc thực thi giao tác bị cắt ngang bởi một loại sự cố nào đó thì DBMS sẽ chịu trách nhiệm xác định những công việc của giao tác để khôi phục lại sau sự cố.

Có 2 chiều hướng thực hiện:

* hoặc nó sẽ được kết thúc bằng cách hoàn tất các hành động còn lại,
* hoặc có thể kết thúc bằng cách hồi lại tất cả các hành động đã được thực hiện.

02 lý do cơ bản làm cho giao tác không hoàn thành: ¬ giao tác tự huỷ bỏ (transaction aborts) ¬ hệ thống bị sự cố (system crashes)

Tại sao giao tác tự hủy? ¬

Do yêu cầu của bản thân giao tác hoặc của người sử dụng nó. ¬ Do sự ép buộc của hệ thống: ♣ quá tải hệ thống ♣ bị kẹt trong một khoá gài (deadlock).

**Tính bền vững - Mục đích:**

* Để bảo đảm rằng mỗi khi giao tác uỷ thác, kết quả của nó sẽ được duy trì và không bị xoá ra khỏi CSDL.
* DDBMS có trách nhiệm bảo đảm kết quả của giao tác và ghi vào CSDL.
* Tính bền vững được sử dụng như là một điều kiện để khôi phục dữ liệu (database recovery), nghĩa là cách khôi phục CSDL về trạng thái nhất quán mà ở đó mọi hành động đã uỷ thác đều được phản ánh.

**Tính tuần tự (Serializability):**

**Mục đích:**

* Quản lý vấn đề thực hiện đồng thời của hai hoặc nhiều giao tác.

**Yêu cầu:**

* Nếu có nhiều giao tác thực hiện đồng thời, thì kết quả phải như nhau nếu nó được thực hiện tuần tự trong cùng một thứ tự đó.
* Nếu một hệ thống có tính điều khiển đồng thời, người lập trình có thể ghi lại giao tác như khi nó thực hiện một mình.

**Tính biệt lập**

**Mục đích:** Tính chất này để ngăn ngừa sự hủy bỏ dây chuyền (cascading abort - Còn gọi là hiệu ứng domino).

**Chú ý:**

* Một giao tác đang thực thi không thể đưa ra các kết quả của nó cho những giao tác khác đang cùng hoạt động trước khi nó uỷ thác.
* Nếu một giao tác cho phép những giao tác khác sử dụng những kết quả chưa hoàn tất của mình trước khi uỷ thác, rồi sau đó nó quyết định huỷ bỏ, thì mọi giao tác đã đọc những giá trị chưa hoàn tất đó cũng sẽ phải được huỷ bỏ nếu không xâu mắt xích này dễ dàng tăng nhanh và gây ra những phí tổn đáng kể cho DDBMS.

1

ÔN TẬP CSDL – LÝ THUYẾT

\* Tài liệu: SV KHÔNG được sử dụng tài liệu

A - NỘI DUNG ÔN TẬP – LÝ THUYẾT VÀ BÀI TẬP

CHƯƠNG 4: MÔ HÌNH THỰC THỂ KẾT HỢP ER

- Ôn lại tất cả lý thuyết và bài tập của Chương 4

1. Thiết kế mô hình ER

- Thực thể (entity)

- Thuộc tính (attribute)

+ Thuộc tính đơn trị ; Thuộc tính đa trị ; Thuộc tính kết hợp; Thuộc tính suy diễn

(dẫn xuất)

+ Thuộc tính khoá

- Mối quan hệ (relationship)

+ Dạng bản số (min,max): (0,1), (1,1), (0,n), (1,n), đệ quy

- Thực thể yếu (weak entity)

- Vẽ lược đồ ER

CHƯƠNG 5: MÔ HÌNH DỮ LIỆU QUAN HỆ

- Ôn lại tất cả phần lý thuyết và bài tập Chương 5

- Ôn lại các khái niệm của mô hình quan hệ: thuộc tính, miền trị, bậc,…

1. Ràng buộc toàn vẹn

+ Siêu khoá ; Khoá dự tuyển; Khoá chính; Khoá ngoại

2. Cách chuyển đổi từ mô hình ER sang mô hình thiết kế quan hệ

+ Chuyển đổi thực thể mạnh

+ Chuyển đổi thực thể yếu

+ Chuyển đổi thuộc tính đa trị

+ Chuyển đổi tập mối quan hệ 1:1

+ Chuyển đổi tập mối quan hệ 1:N

+ Chuyển đổi tập mối quan hệ M:N

+ Chuyển đổi tập mối quan hệ bậc n

2

CHƯƠNG 2 + 3: NGÔN NGỮ SQL (kết hợp giữa lý thuyết về Ngôn ngữ

SQL trên lớp và phần bài tập thực hành trên phòng máy)

- Ôn lại tất cả phần lý thuyết và bài tập Chương 2 + Chương 3 \_Phần 1

1. Xem lại các kiểu dữ liệu

2. Lệnh thêm, xoá, sửa Database

3. Lệnh TẠO BẢNG, xoá, sửa bảng (table)

+ Cách tạo khoá chính (PRIMARY KEY), khoá ngoại (FOREIGN KEY) trong

bảng

+ Tạo các ràng buộc trong bảng: UNIQUE, DEFAULT,…

4. Lệnh INSERT, UPDATE, DELETE

5. Lệnh truy vấn dữ liệu (SELECT)

a) Truy vấn cơ bản

- Câu SELECT đơn giản

- Các hàm tập hợp trong câu lệnh SELECT: SUM, MAX, MIN, AVERAGE, COUNT

- Mệnh đề GROUP BY và HAVING trong câu lệnh SELECT

+ Chú ý: WHERE thực hiện trước GROUP BY

GROUP BY thực hiện trước HAVING

- Mệnh đề ORDER BY để sắp xếp (ASC tăng dần, DESC giảm dần)

b) Truy vấn lồng nhau

3

- Truy vấn lồng phân cấp:

+ Câu truy vấn con trả về 1 giá trị: sử dụng phép toán so sánh : =, <>, >, >=, <, <=

+ Câu truy vấn con trả về nhiều giá trị (1 danh sách): sử dụng IN , ALL, ANY (ALL và ANY

đi chung với các phép toán so sánh : =, <>, >, >=, <, <=)

- Truy vấn lồng tương quan: EXISTS, NOT EXISTS

c) Phép chia

d) Các dạng truy vấn khác

+ JOIN (INNER JOIN, OUTER JOIN,…)

+ CASE

e) Khung nhìn (View)

f) Các dạng câu truy vấn khác

CHƯƠNG 3 (tiếp theo): ĐẠI SỐ QUAN HỆ

- Ôn lại tất cả phần lý thuyết và bài tập Chương 3 (Phần 2)

1. Phép toán 1 ngôi

- Phép chọn ; Phép chiếu; Phép đổi tên ; Phép gán

2. Phép toán 2 ngôi

- Phép toán tập hợp: phép giao (INTERSECT), phép hội (UNION), Phép trừ (MINUS)

- Phép tích Đề -các

- Phép kết

+ Kết theta

+ Kết bằng

+ Kết tự nhiên: là kết quả của phép kết bằng nhưng sẽ bỏ đi thuộc tính thứ 2 trong

từng cặp thuộc tính có giá trị trùng nhau (các cặp thuộc tính trong các mệnh đề phải cùng

tên và cùng giá trị)

+ Kết ngoài (phép kết ngoài mở rộng bên trái, phép kết ngoài mở rộng bên phải,

phép kết ngoài mở rộng về hai phía)

- Phép chia

3. Hàm tính tổng hợp và gom nhóm:

+ COUNT, SUM, GROUP BY,…

4. Các thao tác cập nhật trên quan hệ: INSERT, UPDATE, DELETE

CHƯƠNG 7: RÀNG BUỘC TOÀN VẸN

- Ôn lại tất cả phần lý thuyết và bài tập Chương 7

1. Các đặc trưng của ràng buộc toàn vẹn

+ Bối cảnh

+ Nội dung: sử dụng ngôn ngữ hình thức

+ Bảng tầm ảnh hưởng

2. Phân loại RBTV

+ Bối cảnh là 1 quan hệ (1 bảng)

• RBTV miền giá trị

• RBTV liên bộ

• RBTV liên thuộc tính

• RBTV liên bộ , liên thuộc tính

+ Bối cảnh là nhiều quan hệ (nhiều bảng)

• RBTV tham chiếu

• RBTV liên bộ, liên quan hệ

• RBTV liên thuộc tính, liên quan hệ

• Thuộc tính tổng hợp

• Chu trình

B – BÀI TẬP

Bài 1. Cho lược đồ quan hệ sau:

KHACHSAN(MaKS, TenKS, DiaChi)

PHONG(SoP, MaKS, LoaiP, Gia)

DATPHONG(MaKS, MaKhach, NgayNhan, NgayTra, SoP)

KHACH(MaKhach, HoTen, DiaChi)

\* Ghi chú: các thuộc tính được gạch dưới liền nét là khoá chính, gạch dưới không liền nét là

khoá ngoại

Câu 1: Biểu diễn các ràng buộc toàn vẹn (RBTV) trong các bài tập sau, hãy khảo sát đặc trưng của

chúng (gồm bối cảnh, biểu diễn RBTV bằng ngôn ngữ hình thức, bảng tầm ảnh hưởng):

a) Mỗi khách sạn có một mã phân biệt với các khách sạn khác

b) Mỗi phòng thuộc về một khách sạn nào đó

c) Ngày nhận khác rỗng và ngày trả phòng phải >= ngày nhận phòng

d) Giá phòng khách sạn phải > 500000 cho phòng loại phòng “Cao cấp”

e) Mỗi khách có một mã số để phân biệt với các khách khác.

f) Mỗi khách sạn không được quá 100 phòng

g) Mỗi khách sạn phải có ít nhất một khách hàng đặt phòng

Câu 2: Viết lệnh tạo bảng (CREATE TABLE) cho tất cả các bảng trong Bài 1 với kiểu dữ liệu tuỳ

chọn, trong đó có cài đặt khoá chính, khoá ngoại và Ràng buộc toàn vẹn (RBTV)

Câu 3: Viết câu lệnh SQL cho các câu sau

5

a) Đưa ra danh sách Giá và và LoaiP của tất cả các phòng của khách sạn ABC

b) Liệt kê tất cả các khách của khách sạn ABC

c) Liệt kê thông tin khách sạn và tổng số khách trong từng khách sạn

d) Liệt kê danh sách tất cả các phòng ở khách sạn ABC có người ở

e) Liệt kê danh sách tất cả các phòng ở khách sạn ABC và tên khách đang ở trong phòng đó

nếu phòng đó có người ở

f) Liệt kê những phòng của khách sạn ABC không có người ở

g) Hãy cho biết tổng số phòng của mỗi khách sạn ở Hà Nội

h) Liệt kê các khách sạn có địa chỉ ở Hà Nội mà tất cả các khách hàng đều ở.

i) Cập nhật đơn giá của tất cả các phòng đơn tăng thêm 5%

j) Liệt kê các khách sạn có số lượng phòng > 40

k) Liệt kê 20 khách sạn có nhiều khách đặt phòng nhất

Câu 4: Viết bằng đại số quan hệ cho câu a ◊ k của câu 2

Bài 2. Một hệ thống quản lý:

GIAOVIEN(MaGV, HoTen, NamSinh, DiaChi)

DETAI (MaDeTai, TenDeTai, TheLoaiDeTai)

SINHVIEN(MaSV, HoTenSV, NgaySinh, QueQuan, Lop, MaDeTai)

HUONGDAN(MaGV, MaSV, NamHuongDan)

Biết rằng sinh viên thực hiện chỉ thực hiện một và chỉ một đề tài, và một đề tài có nhiều sinh

viên tham gia. Mỗi năm mỗi giáo viên có thể hướng dẫn nhiều sinh viên và mỗi sinh viên có thể được

hướng dẫn bởi một hay nhiều giáo viên

Câu 1: Biểu diễn các ràng buộc toàn vẹn (RBTV) trong các bài tập sau, hãy khảo sát đặc trưng của

chúng (gồm bối cảnh, biểu diễn RBTV bằng ngôn ngữ hình thức, bảng tầm ảnh hưởng):

a) Mỗi giáo viên có một mã phân biệt với các giáo viên khác

b) Giáo viên phải sinh trước năm 1992 và địa chỉ không được để trống

c) Mỗi sinh viên có một mã phân biệt với các sinh viên khác

d) Mỗi sinh viên thực hiện 1 đề tài

e) Mỗi đề tài có một mã phân biệt với các đề tài khác, tên đề tài phải là duy nhất

f) Giảng viên hướng dẫn ít nhất một sinh viên

g) Giảng viên hướng dẫn không quá 5 sinh viên

Câu 2: Viết lệnh tạo bảng (CREATE TABLE) cho tất cả các bảng trong Bài 2 với kiểu dữ liệu tuỳ

chọn, trong đó có cài đặt khoá chính, khoá ngoại và Ràng buộc toàn vẹn (RBTV)

Câu 3: Hãy viết câu lệnh SQL để trả lời các câu hỏi sau:

a) Đưa ra thông tin về giáo viên có mã là ‘GV001’

b) Cho biết có bao nhiêu đề tài thuộc thể loại “Ứng dụng”

c) Cho biết giáo viên có mã là “GV012” đã hướng dẫn những sinh viên nào

d) Cho biết giáo viên có mã “GV012” đã hướng dẫn bao nhiêu sinh viên có quê quán là “Hà

Nội”

6

e) Cho biết tên những đề tài chưa có sinh viên nào thực hiện

f) Do sơ suất, thông tin ngày sinh của sinh viên “Nguyễn Văn Hùng”, quê quán “Hà Nam” đã

bị nhập không chính xác. Ngày sinh chính xác là “12/11/1991”. Hãy cập nhật thông tin này

g) Vì lý do khách quan, sinh viên “Lê Văn A”, quê quán “TPHCM” đã xin thôi học. Hãy xoá

toàn bộ thông tin liên quan đến sinh viên này

h) Cho biết những giáo viên tham gia hướng dẫn tất cả các sinh viên

i) Liệt kê thông tin các giáo viên hướng dẫn từ 5 sinh viên trở lên

j) Liệt kê thông tin giáo viên (mã giáo viên) và tổng số sinh viên mà giáo viên đã hướng dẫn

k) Cho biết 5 giáo viên hướng dẫn sinh viên nhiều nhất

Câu 4: Hãy viết đại số quan hệ cho các câu a ◊ k của câu 2