**Nhóm 6:**

* Nguyễn Tuấn Vũ – 18133064
* Nguyễn Xuân Hiệu – 18133013
* Trần Thị Lệ Xuân – 18133066
* Trương Hùng Anh – 18133001

**Assigment:**

**Câu 1. Tại sao bạn lựa chọn hệ thống CSDL thay vì lưu dữ liệu trong file quản lý bởi hệ điều hành?**

Trả lời:

Vì khi lưu trữ trên file thường xảy ra các vấn đề:

+ Bộ nhớ file không đủ lưu được dữ liệu có dung lượng lớn, phức tạp

+ Gặp vấn đề trong bảo mật và kiểm soát truy cập

+ Khi gặp sự cố thỳ khôi phục dữ liệu một cách khó khăn

+ Với lưu trữ trên file thì mỗi câu truy vấn phải viết 1 chương trình đặc biệt.

+ Không đảm bảo tính nhất quán

Trong khi đó lợi ích khi dùng CSDL:

+ Độc lập dữ liệu và truy cập hiệu quả.

+ Giảm thời gian phát triển ứng dụng.

+ Hợp nhất dữ liệu và bảo mật.

+ Quản trị dữ liệu thống nhất.

+ Được nhiều người truy cập đồng thời và khôi phục khi gặp sự cố.

**Khi nào bạn không nên dùng hệ CSDL?**

Tuỳ vào nhu cầu sử dụng của người dùng. Nếu nhu cầu thấp, không cần truy cập đồng thời, tốc độ cao, thì việc sử dụng file sẽ phù hợp, ít tốn kém, phức tạp hơn

**Câu 2. Độc lập dữ liệu mức logic (Logical data independence) là gì? Tại sao nó quan trọng?**

Trả lời:

Độc lập dữ liệu logic (Logical data independence): Cho phép thay đổi lược đồ khái niệm mà không cần phải thay đổi lược đồ mức ngoài hoặc những chương trình ứng dụng. Chúng ta có thể thay đổi lược đồ khái niệm để mở rộng (thêm các trường dữ liệu, các bản ghi) hoặc thu nhỏ cơ sở dữ liệu (xóa các trường dữ liệu, các bản ghi) mà không ảnh hưởng đến mức vật lý.

Nó quan trọng vì những thay đổi này do người quản trị thay đổi cho phù hợp với sự thay đổi của dữ liệu thực tế liên quan đến hệ thống CSDL, view ra cho người dùng xem nên nó quan trọng hơn độc lập vật lý.

**Câu 3. Giải thích sự khác nhau giữa độc lập dữ liệu mức logic và độc lập dữ liệu mức vật lý (physical data independence)? Hãy cho ví dụ minh họa.**

Trả lời:

* Độc lập dữ liệu vật lý là các thay đổi trong mức vật lý để nâng cao hiệu năng lưu trữ mà không ảnh hưởng đến mức khái niệm
* Độc lập dữ liệu logic là người quản trị có thể thay đổi các cấu trúc trong mức khái niệm mà không ảnh hưởng đến mức vật lý.

Ví dụ: Một trường học có sơ đồ tổ chức: Hiệu trưởng -> Hiệu phó -> Giáo viên

Csdl vật lý: 30 file dữ liệu gồm : 1 hiệu trưởng, 2 hiệu phó, 28 giáo viên

Mức khái niệm: Gồm có họ tên, ngày sinh, giới tính, địa chỉ, SDT,…

Khi xóa 1 giáo viên, không ảnh hưởng đến mức khái niệm của trường học => Độc lập với dữ liệu vật lý

Khi trường học thay đổi xóa bớt cột ngày sinh , không ảnh hưởng tới 30 file của mức vật lý => độc lập dữ liệu logic

**Câu 4. Giải thích sự khác biệt giữa lược đồ ý niệm/logic (conceptual/logical schema), lược đồ vật lý/bên trong (Physical/internal schema) và lược đồ ngoài (external schema).**

Trả lời:

* Lược đồ ý niệm/logic (conceptual/logical schema) là định nghĩa các mô hình quan hệ.
* Lược đồ ngoài (external schema) là những bản ghi cụ thể, người dùng cũng như người quản trị CSDL sẽ thao tác với mô hình ngoài dùng ngôn ngữ truy vấn dữ liệu.
* Còn lược đồ trong (Physical/internal schema) là mô hình, cách thức lưu dữ liệu vật lý của hệ thống CSDL.

**Câu 5. Trách nhiệm của DBA. Giả sử là DBA không cần quan tâm đến việc thực hiện các câu truy vấn của chính DBA, thì DBA có cần hiểu về tối ưu hóa câu truy vấn không? Tại sao?**

Trả lời:

Trách nhiệm của DBA (Database administrator):

* Thiết kế lược đồ logic, lược đồ vật lí
* Xử lí bảo mật và uỷ quyền
* Dữ liệu luôn sẵn có, phục hồi dữ liệu
* Có thể điều chỉnh cơ sở dữ liệu khi cần phát triển
* Am hiểu dữ liệu và cách sử dụng dữ liệu của người dùng khác

DBA cần hiểu về tối ưu hóa câu truy vấn vì DBA là người tổ chức dữ liệu, cần đảm bảo hệ thống CSDL luôn hoạt động tốt và tối ưu.

**Câu 6. Ông A cần mua một hệ CSDL. Để tiết kiệm chi phí, ông A chỉ mua một hệ CSDL với số tính năng ít nhất có thể. Ông ta lập kế hoạch chỉ chạy nó một mình trên máy PC của ông ấy và không share thông tin với ai cả. Hãy cho biết tính năng nào trong các tính năng dưới đây của DBMS ông A mua nên có và tại sao:**

**+ Tiện ích bảo mật:**

**+ Kiểm soát đồng thời**

**+ Khôi phục dữ liệu sau sự cố**

**+ Cơ chế khung nhìn**

**+ Ngôn ngữ truy vấn**

Trả lời: DBMS ông A mua nên có các tính năng:

* Tiện ích bảo mật: Đảm bảo cho ông A không bị người khác tấn công hệ CSDL nhằm mục đích lấy cắp thông tin
* Khôi phục dữ liệu sau sự cố: Khi mà dữ liệu gặp phải sự cố bị mất ông A có thể khôi phục dữ liệu lại ban đầu
* Cơ chế khung nhìn: Tiện cho việc sử dụng một cách dễ dàng hơn

**Câu 7. Mô tả cấu trúc của một DBMS. Giả sử hệ điều hành của bạn được nâng cấp để hỗ trợ thêm một số chức năng về file (ví dụ khả năng cho phép lưu một chuỗi các bytes lên đĩa). Hãy cho biết lớp nào của DBMS bạn cần phải viết lại để có thể tận dụng ưu điểm của các tính năng mới đó.**

Trả lời:

Cấu trúc của một DBMS gồm 4 phần:

- Mức nội (internal level): mô tả cấu trúc lưu trữ vật lý của CSDL

- Mức ý niệm (conceptual level):

- Mô tả toàn bộ cấu trúc của cơ sở dữ liệu

+ Che dấu chi tiết vật lý

+ Tập trung vào mô tả dữ liệu, mối quan hệ, các ràng buộc

- Mức ngoại (external level):

Mô tả một phần của CSDL ứng với góc nhìn của một nhóm người dùng

Để có thể tận dụng ưu điểm của tính năng cho phép lưu một chuỗi các byte lên đĩa chúng ta cần phải viết lại mức ngoại.

**Câu 8. Trả lời các câu hỏi sau:**

1. Giao tác (transaction) là gì?
2. Tại sao một DBMS thực hiện xen kẽ các hành động của các giao dịch khác nhau thay vì thực hiện lần lượt từng giao dịch một ?
3. Một user phải chắc chắn điều gì để đảm bảo tính nhất quán giữa một giao dịch và CSDL ? Một DBMS nên chắc chắn điều gì để đảm bảo tính nhất quán giữa thực hiện đồng thời nhiều giao dịch và CSDL.
4. Giải thích về nghi thức khóa 2 giai đoạn nghiêm ngặt (the strict two-phase locking protocol).
5. Tính chất WAL là gì và tại sao nó quan trọng?

Trả lời:

1.Một giao tác (transaction) là một chuỗi một hoặc nhiều câu lệnh SQL được kết hợp lại với nhau thành một khối công việc. Các câu lệnh SQL xuất hiện trong giao tác thường có mối quan hệ tương đối mật thiết với nhau và thực hiện các thao tác độc lập.

2. DBMS thực hiện xen kẽ các transaction nhằm mục đích tận dụng tối đa thời gian trống của các chương trình trong hệ thống. Nếu chỉ thực hiện từng giao dịch một thì một số chương trình sẽ rảnh nhưng không có công việc làm, từ đó hiệu suất kém.

3. DBMS phải đảm bảo rằng giao dịch được thực thi toàn bộ và độc lập với các giao dịch khác. Một tính chất quan trọng của giao dịch trong DBMS là tính nguyên tử và có tính cô lập. Các giao dịch hoặc là thành công hoặc là thất bại hoàn toàn. Điều này đảm bảo tính nhất quán của dữ liệu.

4. Khóa hai-pha nghiêm ngặt sử dụng các khóa chia sẻ và khóa độc quyền để bảo vệ dữ liệu. Một giao dịch phải nắm giữ tất cả các khóa cần thiết trước khi thực thi và không giải phóng bất cứ khóa nào trước khi giao dịch kết thúc hoàn toàn.

5. Đặc tính WAL ảnh hưởng đến chiến lược logging trong DBMS. WAL là viết tắt của Write- Ahead Log, có nghĩa là mỗi một thao tác viết lên cơ sở dữ liệu đều phải được ghi vào log (trên đĩa) trước khi nó thực sự xảy ra trong cơ sở dữ liệu. Điều này bảo vệ được cơ sở dữ liệu nếu có sự cố xảy ra trong quá trình giao dịch đang thực hiện. Bằng việc ghi lại những thay đổi vào trong log trước khi những thay đổi này thực sự diễn ra, cơ sở dữ liệu biết được phải làm lại những thao tác nào để khôi phục sự cố. Ngược lại, nếu việc ghi vào log diễn ra sau thì cơ sở dữ liệu sẽ không thể phát hiện được có những thay đổi nào đã xảy ra trong quá trình khôi phục sự cố.