**Assigment:**

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

Khi lưu trữ trên file thì chúng ta gặp phải 1 số vấn đề sau:

-         Bộ nhớ không đủ để lưu trữ nếu dữ liệu quá lớn.

-         Dữ liệu phải chuyển qua lại giữa bộ nhớ chính và bộ nhớ phụ nên nếu dùng hệ điều hành 32bit thì không đủ.

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

-         Không đảm bảo tính nhất quán.

-         Khi gặp sự cố thì khôi phục lại dữ liệu 1 cách khó khăn.

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

Trong khi đó chúng ta dùng hệ thống CSDL có những lợi ích sau:

-         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.

-         Truy cập đồng thời và khôi phục khi gặp sự cố.

Khi không cần lưu trữ thông tin hay dữ liệu gì thì không nên dùng hệ CSDL.

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

* 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. 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.

* 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 (người quản trị dữ liệu không cần quan tâm nhiều đến mức vật lý cũng như sự thay đổi của mức này)  
  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

**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).

* 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. Lược đồ ý niệm/logic(conceptual/logical schema) là định nghĩa các mô hình quan hệ( các bảng). 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?

Trách nhiệm của DBA:

-         Thiết kế lược đồ logic và lược đồ vật lý.

-         Điều khiển bảo mật và phân quyền, dữ liệu luôn được toàn vẹn.

-         Đảm bảo cho dữ liệu luôn trong tình trạng sẵn sang và phục hồi lại khi có sự cố xảy ra.

-         Sự điều chỉnh dữ liệu cần phải tối ưu hơn.

DBA cần phải hiểu về tối ưu hóa câu truy vấn vì tầm quan trọng của DBA là rất lớn. Cần phải am hiểu trong nhiều mảng liên quan đến lập trình để đảm bảo cho hệ thống CSDL luôn được hoạt động tốt và am hiểu về tối ưu câu truy vấn để có thể tối ưu hệ thống CSDL 1 cách tốt nhất khi cần thiết.

**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

Ông A nên mua DBMS có các tính năng sau:

-         Tiện ích bảo mật(nên có vì không có chức năng này dữ liệu của ông A có thể sẽ bị tấn công).

-         Khôi phục dữ liệu sau sự cố( Cần thiết để tránh mất dữ liệu khi gặp sự cố).

-         Cơ chế khung nhìn(Phải có để ông A dễ dàng sử dụng).

Kiểm soát đồng thời với ngôn ngữ truy vấn có thể không cần vì chỉ có 1 mình ông A sử dụng và ngôn ngữ truy vấn ông A cũng không cần biết vì đó là công việc của DBA quản lý hệ CSDL đó.

**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.

Cấu trúc của một DBMS gồm 4 phần sau:

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

-         Mức ý niệm (conceptual level):

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

o   Che dấu chi tiết vật lý

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

-         Mức ngoại (external level):

o   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 ?

1.      Transaction( giao tác) là một tập các hành động thực hiện cùng một chức năng và chúng chỉ có thể cùng thành công hoặc thất bại.

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.