# Câu hỏi 1:

Bài tập chương Consistency and Replication

- Thực hiện sao lưu giữ liệu bởi vì:

+) Tăng tính sẵn sàng của hệ thống

+) Tăng hiệu năng, người dùng có thể đọc dữ liệu ở bản sao gần với mình nhất

+) Khả năng co giãn

# Câu hỏi 2:

a.

Có thỏa mãn tính thống nhất nhân quả bởi vì các thao tác ghi nhân quả (thao tác ghi a và thao tác ghi c) thống nhất thứ tự trên toàn cục

b.

Không thoả mãn tính thống nhất tuần tự bởi vì không tìm được thứ tự toàn tục để bảo toàn thứ tự cục bộ của P4 và P5

# Câu hỏi 3:

Conit là gì? Nếu đặt kích thước Conit quá lớn thì sẽ gây ra vấn đề gì? Tương tự với kích thước Conit quá nhỏ?

Conit: định nghĩa độ không thống nhất, từ đó đánh giá độ thống nhất, tập các đơn vị giữ liệu của kho dữ liệu phân tán, mức độ thống nhất được xác định theo nhiều góc độ khác nhau, về góc độ số lượng các thao tác đã được thực hiện nhưng chưa hoàn thành và các thao tác chưa được thực hiện trên phần tử dữ liệu, về sai khác của dữ liệu đã hoàn thành vs dữ liệu sẽ hoàn thành, bàn về kích thước conit nếu lớn -> số lần cập nhập nhiều, conit bé -> số lần cập nhập ít

Số lượng conit ít, kích thước lớn -> dẫn đến khả năng xuất hiện sai lệch cao -> chi phí cập nhập cao

Số lượng conit nhiều, kích thước nhỏ -> giảm khả năng xuất hiện sai lệch -> chi phí quản lý bản sao lớn

# Câu hỏi 4:

Vì trong một số trường hợp có thể thỏa mãn tính chất nhân quả nhưng lại không thỏa mãn tính chất tuần tự

VD: Có 4 tiến trình P4, P2, P3, P4 ở thời gian t1 P1 thực hiện ghi a, thời gian t2 P2 thực hiện ghi b, t3 P3 thực hiện đọc b, P4 thực hiện đọc a, t4 P3 thực hiện đọc a và P4 thục hiện đọc b.

Vd trên thỏa mãn tính thống nhất nhân quả nhưng không thỏa mãn tính chất thống nhất tuần tự.

# Câu hỏi 5:

Vấn đề xảy ra khi tiến trình khách đang truy cập vào một bản sao cố định giờ truy cập vào một bản sao khác

VD: một máy thực hiện thay đổi trên một bản sao, sau đó di chuyển đến một địa điểm khác -> bản sao tại đó chưa được cập nhập

Định nghĩa: Đảo bảo các tiến trình có sử dụng giữ liệu có một khunh hình thống nhất về dữ liệu

# Câu 6:

a.

Lựa chọn k - 1 bản sao, sẽ xác định bản sao thứ k sao cho khoảng cách trung bình của các bản sao đến client là nhỏ nhất.

b.

Nếu dữ liệu X nằm trên một server P, lượng truy cập của X nhỏ hơn del(X) thì sẽ xóa dữ liệu X tại

server P, nhưng đòi hỏi phải có thuật toán đảm bảo luôn có bản sao dữ liệu của X. Nếu lượng truy cập lớn hơn rep(X) thì sẽ sao chép cho các server khác, nếu nằm khoảng del(X) đến rep(X) thì sẽ giữ nguyên.

c.

|  |  |  |
| --- | --- | --- |
|  | sao lưu tích cực | quorum |
| ưu điểm | thống nhất mạnh do cập nhập tất cả bản sao | giảm chi phí vô ích vì chỉ cập nhập một số bản sao cố định |
| nhược điểm | chi phí cập nhập cao do phải cập nhất tát cả bản sao | thống nhất không mạnh |

**Câu 7:**

a.

Hướng dữ liệu: Hướng tới cung cấp một khung nhìn chung cho kho dữ liệu

trên toàn hệ thống vì vậy khi có thao tác cập nhập thao tác này cần được phổ biến cho các bản sao khác nhau.

Hướng người dùng: Đảm bảo các tiến trình sử dụng dữ liệu có một khung nhìn chung về dữ liệu, do đó chỉ cập nhập với các tiến trình nào có sử dụng dữ liệu các tiến trình khác thì thôi.

b. nhập

WWW: Sử dụng mô hình thống nhất hướng người dùng, do www chỉ đọc và không cập

CDN: Sử dụng mô hình thống nhất hướng dữ liệu, do đây là dịch vụ thời gian thực

tường xuyên cập nhập

c.

Nw + Nr = 2700 < N = 3000 -> không tránh được xung đột đọc ghi Nw > N/2 -> tránh được xung đột ghi ghi