# Bài tập Chương 6

## Bài tập 6.1

* **Eager Centralized (Sao chép Sốt sắng Tập trung):** Hệ thống đặt vé máy bay trung tâm. Có một master database duy nhất để xử lý tất cả các giao dịch đặt vé (ghi), đảm bảo tính nhất quán mạnh mẽ ngay lập tức. Các văn phòng đại lý trên toàn cầu có các bản sao chỉ đọc (read-only) để truy vấn nhanh thông tin chuyến bay, và chấp nhận được độ trễ khi đọc dữ liệu mới nhất.
* **Eager Distributed (Sao chép Sốt sắng Phân tán):** Hệ thống quản lý kho hàng thời gian thực cho một chuỗi bán lẻ. Mỗi cửa hàng có thể bán hàng (cập nhật số lượng tồn kho) và cần thấy ngay lập tức số lượng tồn kho chính xác từ các cửa hàng khác để thực hiện chuyển kho. Tính nhất quán mạnh và khả năng ghi tại nhiều nơi là rất quan trọng.
* **Lazy Centralized (Sao chép Lười biếng Tập trung):** Danh mục sản phẩm (product catalog) của một trang web thương mại điện tử. Văn phòng chính cập nhật thông tin sản phẩm trên một master database. Các thay đổi này sau đó được truyền từ từ đến các máy chủ web trên toàn cầu. Người dùng có thể chấp nhận xem thông tin sản phẩm cũ hơn vài phút.
* **Lazy Distributed (Sao chép Lười biếng Phân tán):** Lịch làm việc (calendar) hoặc ứng dụng ghi chú cộng tác (collaborative note-taking app) như Google Docs. Mỗi người dùng có thể cập nhật bản sao của riêng mình (kể cả khi ngoại tuyến). Các thay đổi sau đó được đồng bộ hóa với những người dùng khác, và hệ thống sẽ thực hiện giải quyết xung đột (reconciliation) nếu có nhiều người cùng sửa một mục.

## Bài tập 6.2

1. **Nhóm Dịch vụ Khách hàng (Customer Service Group):**

* Loại dữ liệu sao chép: Toàn bộ các bảng CUSTOMER, CLIENT-ORDER, ORDER, ORDER-LINE và STOCK.
* Giao thức kiểm soát bản sao: Sao chép sốt sắng phân tán (Eager Distributed Protocol). Do khối lượng công việc 80% là truy vấn, việc có bản sao cục bộ giúp đọc nhanh. 20% cập nhật (đặt hàng, thanh toán) yêu cầu tính nhất quán mạnh (strong mutual consistency) để đảm bảo dữ liệu (như số lượng tồn kho) là chính xác ngay lập toàn bộ cluster. Giao thức này cho phép bất kỳ nhân viên nào cũng có thể thực hiện cập nhật và đảm bảo tính nhất quán.

1. **Nhóm Phân tích của Ban Quản lý (Management’s Analysis Group):**

* Loại dữ liệu sao chép: Các bảng ORDER, ORDER-LINE, STOCK, và các dữ liệu tổng hợp từ chúng.
* Giao thức kiểm soát bản sao: Sao chép lười biếng tập trung (Lazy Centralized Protocol). Cơ sở dữ liệu phân tích đóng vai trò là “slave”, nhận dữ liệu từ các cơ sở dữ liệu hoạt động tại các nhà kho (đóng vai trò là “master”). Việc phân tích không yêu cầu dữ liệu thời gian thực, do đó việc sao chép trễ (lazy) là phù hợp và không ảnh hưởng đến hiệu năng của hệ thống giao dịch chính tại các nhà kho.

## Bài tập 6.3

Một phương pháp quản lý phân tán cho đồ thị sao chép (replication graph) có thể được thực hiện như sau:

1. **Lưu trữ đồ thị cục bộ:** Mỗi site Si duy trì một phần của đồ thị sao chép, chỉ chứa các giao dịch và các cạnh liên quan đến các hoạt động tại Si hoặc được Si biết đến.
2. **Trao đổi thông tin đường đi (Path Information Exchange):** Khi một giao dịch Ti tại site Si cần tương tác với site Sj, Si sẽ gửi thông tin về các đường đi có liên quan trong đồ thị cục bộ của nó cho Sj.
3. **Phát hiện chu trình phân tán:** Khi Sj nhận được thông tin đường đi từ Si, nó sẽ cập nhật đồ thị cục bộ của mình. Nếu việc thêm các cạnh mới tạo ra một chu trình, một xung đột tiềm tàng đã được phát hiện.
4. **Cơ chế phối hợp:** Khi một giao dịch Ti muốn commit tại primary site của nó, site này sẽ khởi tạo một “giao thức kiểm tra chu trình” (cycle-check protocol). Nó gửi các phụ thuộc của Ti đến tất cả các site tham gia khác. Mỗi site kiểm tra cục bộ và báo cáo lại. Nếu bất kỳ site nào phát hiện ra chu trình, Ti sẽ bị hủy bỏ (abort). Giao thức này hoạt động tương tự như 2PC nhưng để kiểm tra tính khả tuần tự thay vì đồng ý commit.

## Bài tập 6.4

1. **Phân bổ phiếu bầu (Votes) và Quorum:**

* Đối với dữ liệu x (tại site 1, 2, 3): Tổng số phiếu V(x) = 3. Gán V1=1, V2=1, V3=1. Một cặp quorum hợp lệ là Vr(x) = 2 và Vw(x) = 2. (Kiểm tra: 2+2 > 3 và 2 > 3/2).
* Đối với dữ liệu y (tại site 2, 3, 4): Tổng số phiếu V(y) = 3. Gán V2=1, V3=1, V4=1. Một cặp quorum hợp lệ là Vr(y) = 2 và Vw(y) = 2.

1. **Phân hoạch mạng cho x (Các site {1, 2, 3}):**

* Một giao dịch cập nhật x cần Vw(x) = 2 phiếu.
* Trong bất kỳ trường hợp phân hoạch nào (ví dụ: {1} | {2, 3} hoặc {1, 2} | {3}), chỉ có phân vùng chứa ít nhất 2 trong 3 site (tức là phân vùng đa số - majority partition) mới có thể thu thập đủ quorum để kết thúc giao dịch. Phân vùng thiểu số sẽ bị khóa.

1. **Phân hoạch mạng cho y (Các site {2, 3, 4}):**

* Tương tự như x.
* Chỉ có phân vùng chứa ít nhất 2 trong 3 site {2, 3, 4} mới có thể thu thập đủ quorum Vw(y) = 2 để kết thúc một giao dịch cập nhật y.