

CƠ SỞ DỮ LIỆU PHÂN TÁN

NGUYỄN THỊ THANH THỦY

1

Nội dung môn học

- Giới thiệu
- Thiết kế CSDL phân tán
- Điều khiển dữ liệu phân tán
- Xử lý truy vấn phân tán
- **Xử lý giao dịch phân tán**
- Nhân bản dữ liệu
- Tích hợp CSDL – Các hệ thống đa CSDL
- Các hệ thống CSDL song song
- Quản lý dữ liệu ngang hàng (Peer-to-Peer)
- Xử lý dữ liệu lớn
- NoSQL, NewSQL và Polystores
- Quản lý dữ liệu Web

2

Nội dung

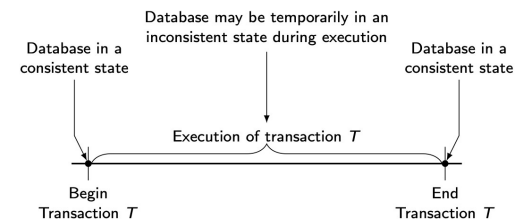
- **Xử lý giao dịch phân tán**
 - Điều khiển đồng thời phân tán
 - Độ tin cậy phân tán

3

Giao dịch

Giao dịch là một tập hợp các hành động thực hiện các chuyển đổi nhất quán về trạng thái hệ thống trong khi vẫn duy trì tính nhất quán của hệ thống.

- Trong suốt đồng thời
- Trong suốt lỗi



4

Đặc tính của giao dịch

Begin_transaction

...
Read
Read
...
Write
Read
...

Commit

- Tập đọc (RS)
 - Tập các mục dữ liệu được đọc bởi một giao dịch
- Tập ghi (WS)
 - Tập các mục dữ liệu có giá trị được thay đổi bởi giao dịch
- Tập cơ sở (BS)
 - $RS \cup WS$

5

Nguyên tắc giao dịch

- **Tính nguyên tử**
 - Tất cả hoặc không có gì
- **Tính nhất quán**
 - Không vi phạm các ràng buộc toàn vẹn
- **Tính cô lập**
 - Những thay đổi đồng thời không nhìn được → tuần tự hóa
- **Tính bền vững**
 - Các cập nhật được cam kết về tính bền vững

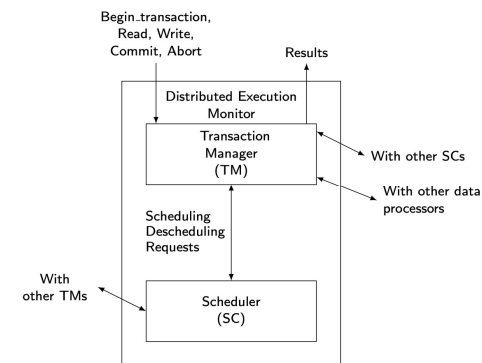
6

Giao dịch cung cấp...

- Thực thi *tính nguyên tử* và *tính tin cậy* khi có lỗi.
- Thực thi *chính xác* khi có nhiều người cùng truy nhập.
- Quản lý chính xác các *bản sao*.

7

Kiến trúc bộ quản lý giao dịch phân tán



8

Nội dung

- Xử lý giao dịch phân tán
 - Điều khiển đồng thời phân tán
 - Độ tin cậy phân tán

9

Điều khiển đồng thời

- Thực hiện đồng bộ hóa các giao dịch đồng thời sao cho tính nhất quán của cơ sở dữ liệu được duy trì, cũng như đạt được mức độ đồng thời tối đa.
- Thực thi tính **cô lập**
- Những bất thường:
 - Mất bản cập nhật
 - Ảnh hưởng của một số giao dịch không được phản ánh trên cơ sở dữ liệu.
 - Truy xuất không nhất quán
 - Một giao dịch, nếu nó đọc cùng một mục dữ liệu nhiều lần, thì phải luôn đọc được cùng một giá trị.

10

Khả năng tuần tự hóa trong CSDLPT

- Hai lịch sử cần được xem xét:
 - Lịch sử cục bộ
 - Lịch sử toàn cục
- Để các giao dịch toàn cục có thể **tuần tự hóa** được, cần hai điều kiện:
 - Mỗi lịch sử cục bộ phải được tuần tự hóa → **khả năng tuần tự hóa cục bộ**
 - Hai hoạt động xung đột phải có cùng thứ tự tương đối trong tất cả các lịch sử cục bộ nơi chúng xuất hiện cùng nhau → **khả năng tuần tự hóa toàn cục**

11

Tính không tuần tự hóa toàn cục

| | |
|---|---|
| T_1 : Read(x) $x \leftarrow x - 100$ Write(x) Read(y) $y \leftarrow y + 100$ Write(y) Commit | T_2 : Read(x) Read(y) Commit |
|---|---|

- x được lưu tại Trạm 1, y được lưu tại Trạm 2
- LH_1, LH_2 có thể tuần tự hóa riêng (trên thực tế là nối tiếp), nhưng hai giao dịch này không thể tuần tự hóa toàn cục.

$$LH_1 = \{R_1(x), W_1(x), R_2(x)\}$$

$$LH_2 = \{R_2(y), R_1(y), W_1(y)\}$$

12