



ĐIỀU KHIỂN CÁC GIAO DỊCH ĐỒNG THỜI PHÂN TÁN

Nhóm 8

Nhóm 8

Trương Văn Thắng – B18DCCN633

Phạm Văn Vinh – B18DCCN691

Nguyễn Thế Dũng – B18DCCN101

Vũ Đức Thắng – B18DCCN634

Trần Trung Đức – B18DCCN183

Nguyễn Ngọc Huy – B18DCCN268

Trần Đỗ Minh – B18DCCN411

Trần Quang Huy – B18DCCN272

Cao Tuấn Nghĩa – B18DCCN446



Tổng quan 01

Phân loại các cơ chế
điều khiển đồng thời 02

Điều khiển 03

Mục Lục

04 Thuật toán ZPL tập
quyền

05 Thuật toán ZPL phân
tán

06 Thảo luận


01

Tổng quan





Tại sao phải điều khiển đồng thời phân tán

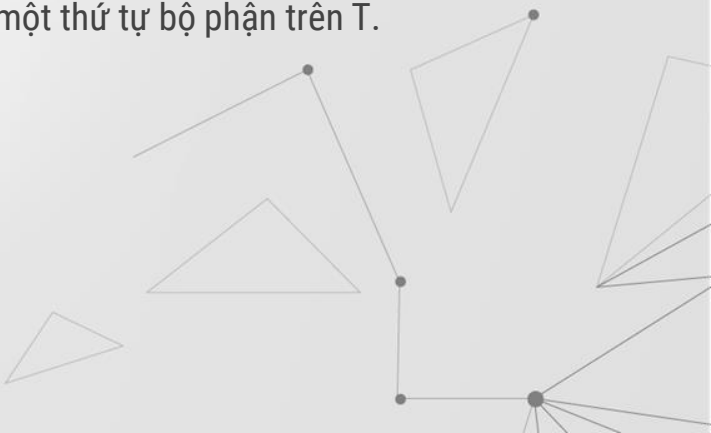
- Các giao dịch tiến hành song song làm tang hiệu năng hệ thống nên phát sinh vấn đề phức tạp
 - Điều khiển các giao dịch đồng thời phân tán cần quan tâm đến tính biệt lập(isolation) và tính nhất quán(consistency)
 - Cần có cơ chế đảm bảo tính chất nhất quán toàn cục và hiệu quả thực thi toàn hệ thống
 - Lý thuyết tuần tự (Serializability Theory) là nền tảng cơ sở để xây dựng cho các thuật toán điều khiển phân tán.
- 



Lý thuyết khả tuần tự lịch biểu

-Nếu việc thực thi đồng thời các giao tác làm cho cơ sở dữ liệu ở vào trạng thái như khi thực hiện tuần tự theo một thứ tự nào đó thì các vấn đề như cập nhật bị thất lạc sẽ được giải quyết. Đây là điểm quan trọng của lý luận về tính khả tuần tự.

-Một lịch S (schedule) được định nghĩa trên tập giao dịch $T = \{T_1, T_2, \dots, T_n\}$ và xác định thứ tự thực thi đan xen lẫn nhau các thao tác trong giao dịch (lịch biểu tuần tự), khi đó các sự cố tranh chấp chắc chắn không xảy ra và trong cơ sở dữ liệu sẽ được một kết quả nào đó. Lịch biểu có thể được mô tả như là một thứ tự bộ phận trên T .



Một lịch đầy đủ (complete schedule) :

- Là lịch định nghĩa thứ tự thực hiện cho tất cả các thao tác trong miền biến thiên của nó.

- Về hình thức, một lịch biểu đầy đủ S_T^c được định nghĩa trên một tập giao dịch $T = \{T_1, T_2, \dots, T_n\}$ là một thứ tự bộ phận $S_T^c = \{\Sigma_T, \alpha_T\}$

- Ví dụ:

T_1 : Read(x)	T_2 : Read(x)
$x \leftarrow x + 1$	$x \leftarrow x + 1$
Write(x)	Write(x)
Commit	Commit

Một lịch biểu đầy đủ S_T^c trên $T = \{T_1, T_2\}$ có thể được biểu diễn như sau:

$$S_T^c = \{\Sigma_T, \alpha_T\}$$

Trong đó:

$$\Sigma_1 = \{R_1(x), W_1(x), C_1\}$$

$$\Sigma_2 = \{R_2(x), W_2(x), C_2\}$$

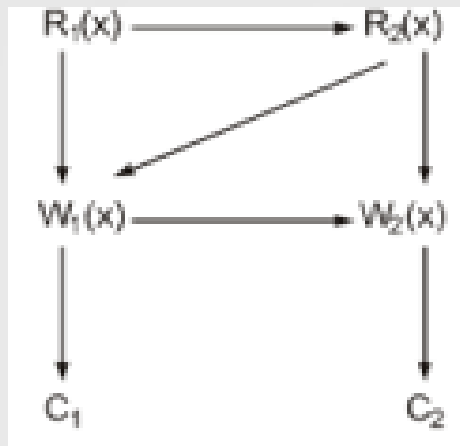
Vì vậy:

$$\Sigma_T = \Sigma_1 \cup \Sigma_2 = \{R_1(x), W_1(x), C_1, R_2(x), W_2(x), C_2\}$$

và

$$\alpha_T = \{(R_1, R_2), (R_1, W_1), (R_1, C_1), (R_1, W_2), (R_1, C_2), (R_2, W_1), (R_2, C_1), (R_2, W_2), (R_2, C_2), (W_1, C_1), (W_1, W_2), (W_1, C_2), (C_1, W_2), (C_1, C_2), (W_2, C_2)\}.$$

- Có thể đặc tả như một DAG hình dưới đây



• Lịch là tiền tố (Prefix) của lịch đầy đủ được định nghĩa như sau: Cho một thứ tự bộ phận $P = \{\Sigma, \alpha\}$. Nói rằng $P' = \{\Sigma', \alpha'\}$ là một tiền tố của P nếu :

1. $\Sigma' \subseteq \Sigma$
2. $\forall e_i \in \Sigma', e_1 \alpha' e_2$ khi và chỉ khi $e_1 \alpha e_2$
3. $\forall e_i \in \Sigma',$ nếu $\exists e_j \in \Sigma$ và $e_j \in e_i$ thì $e_j \in \Sigma'$.

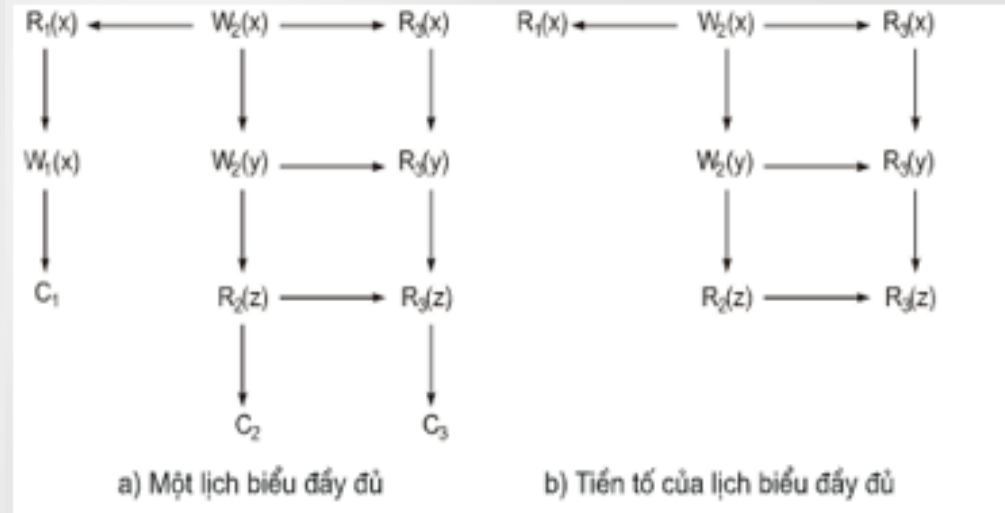
Ví dụ:


Xét ba giao dịch sau đây:

T1: Read(x) Write(x) Commit

T2: Write(x) Write(y) Read(z) Commit

T3: Read(x) Read(y) Read(z) Commit





-Lịch tuần tự: Nếu trong một lịch biểu S, các thao tác của các giao dịch khác nhau không được thực hiện xen kẽ, nghĩa là các thao tác của mỗi giao dịch xảy ra liên tiếp

-Ví dụ Xét một lịch như sau:

$S = \{W2(x), W2(y), R2(z), C2, R1(x), W1(x), C1, R3(x), R3(x), R3(y), R3(z), C3\}$.

-Lịch S là tuần tự vì tất cả các thao tác của T2 được thực hiện trước các thao tác của T1 và các thao tác của T1 được thực hiện trước các thao tác của T3.

-Có thể biểu diễn mối liên hệ thứ bậc giữa các thực thi giao dịch là:

$T2 \propto_S T1 \propto_S T3$ hoặc $T2 \rightarrow T1 \rightarrow T3$





-Tương đương tương tranh giữa hai lịch:

+Hai lịch S1 và S2 được định nghĩa trên cùng một tập giao tác T được gọi là tương đương

+Hai lịch S1 và S2 tương đương theo thứ tự thực thi tương đối của các thao tác tương tranh trong lịch được gọi là tương đương tương tranh

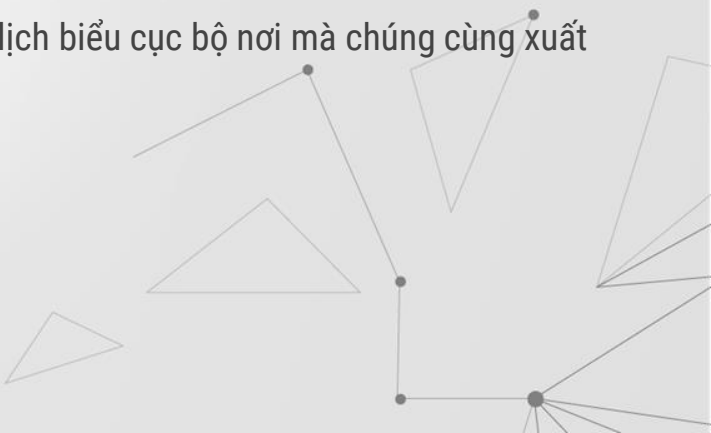
-Lịch khả tuần tự: Một lịch S được gọi là khả tuần tự nếu và chỉ nếu có tương đương tương tranh với một lịch tuần tự

-Lịch khả tuần tự một bản(One – Copy Serializable): các lịch có thể duy trì được tính nhất quán tương hỗ được gọi là khả tuần tự một bản

+Một lịch biểu toàn cục khả tuần tự một bản phải thỏa mãn những điều kiện sau:

1. Mỗi lịch biểu cục bộ phải khả tuần tự.

2. Hai thao tác tương tranh phải có cùng thứ tự tương đối trong tất cả các lịch biểu cục bộ nơi mà chúng cùng xuất hiện





02

Phân loại các cơ chế điều khiển đồng thời

Các tiêu chuẩn phân loại

Các tiêu chuẩn phân loại gồm:

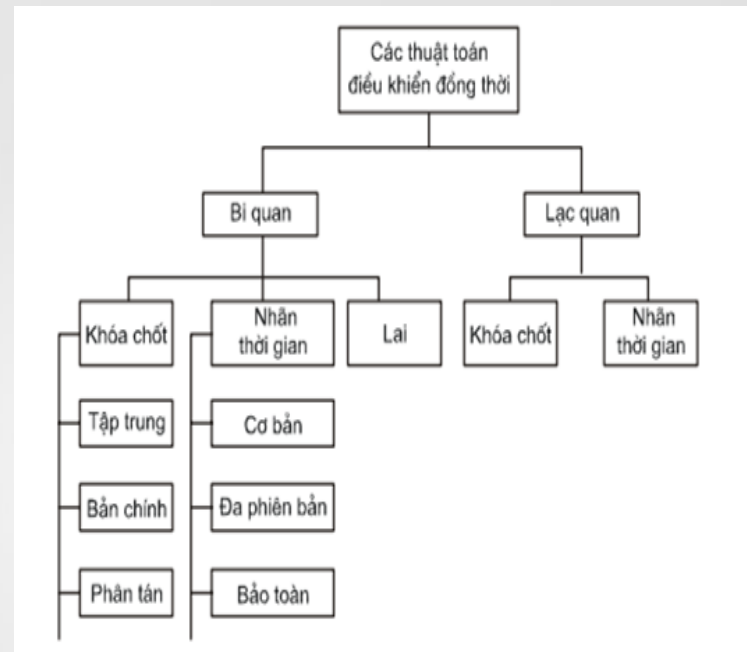
- Theo chế độ phân tán CSDL.
- Theo topo mạng.
- Theo tiêu chuẩn thông dụng nhất: theo sự nguyên thủy đồng bộ hóa (synchronization primitive).

+Quan điểm bi quan.

+Quan điểm lạc quan.

→ Sử dụng các tiêu chuẩn trên, ta phân loại được hai phương pháp điều khiển đồng thời đó là:

- Điều khiển đồng thời bi quan
- Điều khiển đồng thời lạc quan

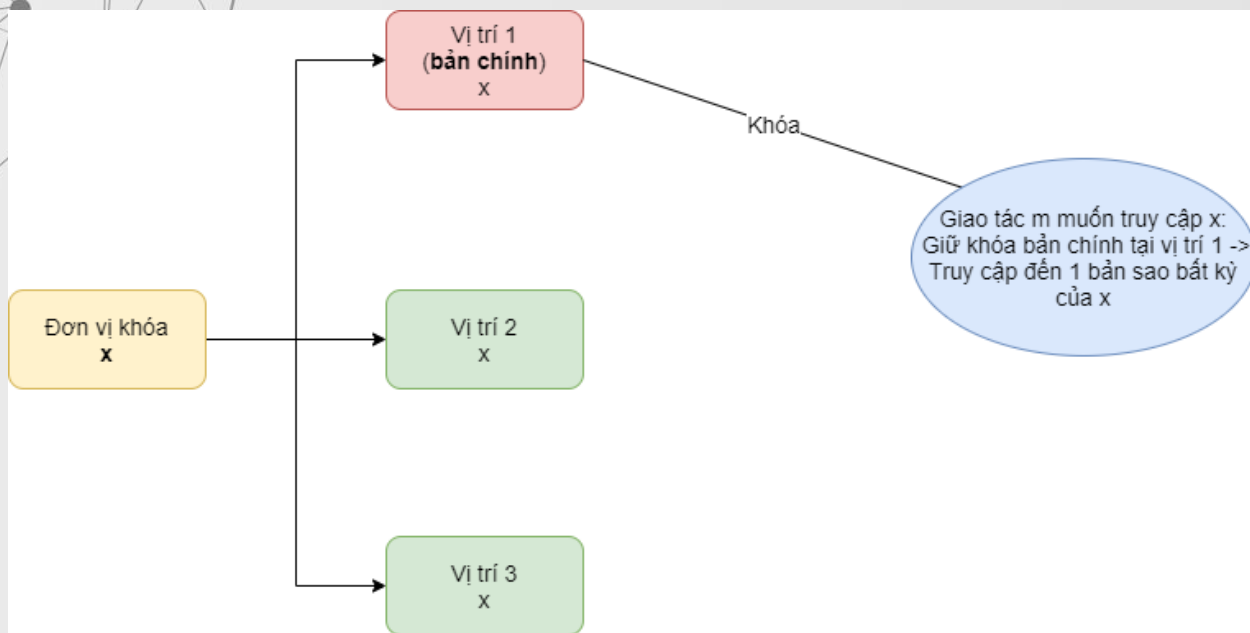




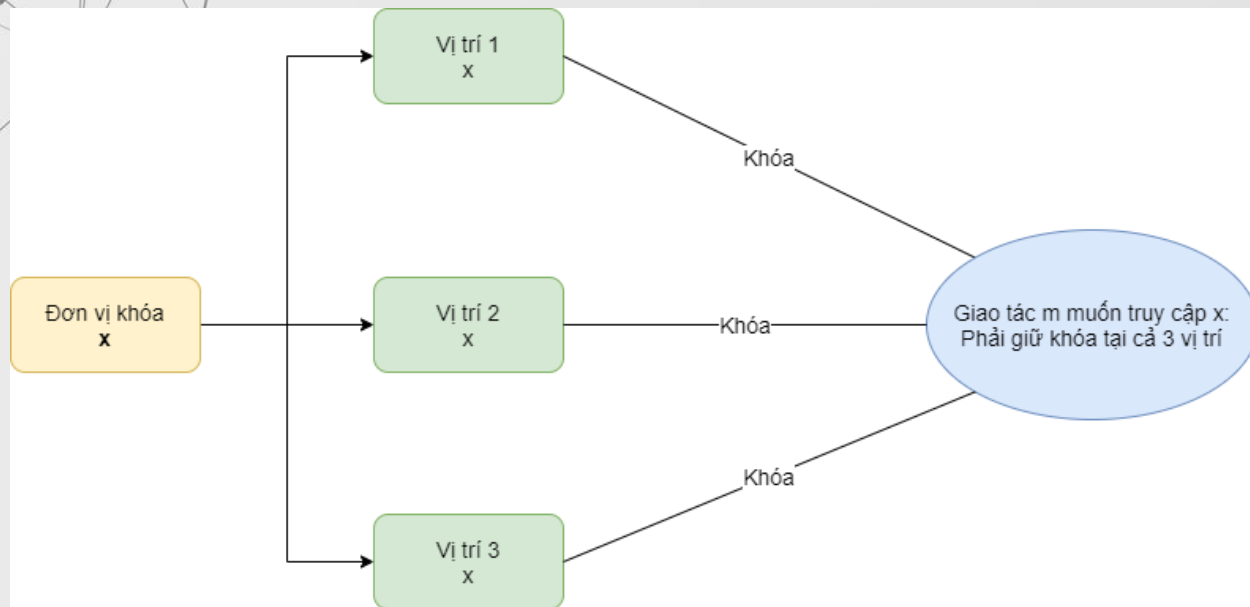
Các thuật toán dựa trên khóa chốt

- Đồng bộ giao tác bằng khóa chốt vật lý hoặc khóa logic.
- Kích thước của phần CSDL được áp dụng khóa gọi là độ mịn khóa (locking granularity) nhưng hiện ta tạm bỏ qua điều này, coi kích thước của phần CSDL được áp dụng khóa là 1 đơn vị khóa (lock unit).
- Các TT dựa theo khóa chốt được phân loại theo vị trí thực hiện hoạt động quản lý khóa:
 - +Lỗi khóa tập quyền
 - +Lỗi khóa bản chính
 - +Lỗi khóa phân tán

-Lỗi khóa bản chính



-Lỗi khóa phân tán





Các thuật toán dựa theo thứ tự nhãn thời gian

- Tạo ra 1 thứ tự thực hiện giao tác đảm bảo được tính chất nhất quán và tính tương hỗ giữa các vị trí trong mạng
- Gán nhãn thời gian cho cả thao tác và mục dữ liệu trong CSDL
- Phân thành 3 loại sau:
 - +Cơ bản (Basic)
 - +Đa phiên bản (Multiversion)
 - +Bảo toàn (Conservative)



Các thuật toán lai

- Phối hợp thuật toán dựa theo khóa chốt với nhãn thời gian nhằm cải thiện:
 - + Hiệu quả
 - + Mức độ hoạt động đồng thời
- Chưa được cài đặt trong các hệ quản trị CSDLPT thương mại hay thử nghiệm.



03

Điều khiển



Các thuật toán điều khiển đồng thời bằng khóa chốt

Một số định nghĩa:

Declare-type

Operation: Các toán tử Begin_transaction, Read, Write, Abort, hoặc Commit.

Dataltem: Một mục dữ liệu trong cơ sở dữ liệu phân tán.

TransactionId: Mã là duy nhất của mỗi giao dịch.

DataVal: Một giá trị có kiểu dữ liệu cơ bản (số nguyên, số thực...).

SiteId: Định danh duy nhất cho vị trí.

-Dbop: Một bộ ba gồm

Opn: Operation

Data: Dataltem

Tid: TransactionId

-Dpmsg: Một bộ ba gồm

Opn: Operation

Tid: TransactionId

Result: DataVal

-Scmsg: Một bộ ba gồm

Opn: Operation

Tid: TransactionId

Result: DataVal

-Transaction ← Một bộ hai gồm

Tid: TransactionId

Body: thân giao dịch

-Message ← một thông báo cần được truyền đi.

OpSet: một tập các Dbop

SiteSet: một tập các SiteId

WAIT(msg:Message)

Begin

đợi một thông báo đến}

End



Thuật toán quản lý khóa cơ bản

- Khi giao dịch có yêu cầu đọc hoặc ghi vào một mục cơ sở dữ liệu chứa khóa, bộ quản lý giao dịch sẽ chuyển cho bộ quản lý khóa các thao tác đọc, ghi và các thông tin như mục dữ liệu cần truy xuất, định danh của giao dịch đưa ra yêu cầu...
 - Bộ quản lý sẽ thực hiện kiểm tra đơn vị khóa có chứa mục dữ liệu đã bị khóa hay chưa, tức là mục dữ liệu này đã được sử dụng bởi một giao dịch khác hay chưa
 - Bộ quản lý giao dịch sẽ được trả về các kết quả của các hoạt động. Khi giao dịch kết thúc nó sẽ giải phóng các khóa mà nó đã sử dụng và khởi tạo một giao dịch khác cũng đang chờ để truy cập vào khoản mục dữ liệu khác.
 - Thuật toán khóa chốt cơ bản là một thuật toán đơn giản. Tuy nhiên thuật toán không đồng bộ hóa sự thực thi giao dịch một cách hoàn toàn đúng đắn vì khi đưa ra các lịch biểu khả tuần tự, các thao tác khóa chốt và giải phóng khóa chốt cũng đòi hỏi phải được sắp xếp phối hợp.
- => Vì vậy có thể dẫn đến việc mất đi tính cô lập và tính nguyên tử tổng thể của giao dịch.

04

Thuật toán 2PL bản chính





Thuật toán 2PL bản chính

- Vấn đề của thuật toán 2PL tập trung là cả hệ thống phụ thuộc vào 1 vị trí LM => độ tin cậy không cao, có thể tắc nghẽn và hiệu năng thấp
- 2PL bản chính cải tiến giải quyết vấn đề trên: đặt LM tại 1 số vị trí trên mạng, mỗi bộ LM quản lý 1 tập đơn vị khóa . Thuật toán xử lý 1 bản sao của mỗi mục dữ liệu như bản chính. TM gửi yêu cầu xin cấp khóa đến các bộ LM.
- Bước trung gian từ 2PL tập trung đến 2PL phân tán



05

Thuật toán 2PL phân quyền



Mỗi vị trí đều có LM

1. TM điều phối yêu cầu các vị trí DM có tham gia vào giao dịch cấp khóa, và tham gia xử lý dữ liệu

2. Sau khi các DM hoàn thành các thao tác của giao dịch, TM điều phối giải phóng khóa

Khác nhau giữa 2PL tập trung và 2PL phân quyền

-TM điều phối trong thuật toán 2PL tập trung chỉ gửi yêu cầu xin khóa đến 1 vị trí LM, vị trí có bộ quản lý khóa trung tâm, còn trong 2PL phân tán TM điều phối phải gửi yêu cầu xin khóa đến tất cả các vị trí tham gia vào giao dịch

-Trong thuật toán 2PL phân tán, TM điều phối không gửi trực tiếp các thao tác xử lý đến bộ xử lý dữ liệu của vị trí khác, mà thông qua LM của các vị trí này, nghĩa là TM điều phối không chờ thông báo “yêu cầu khóa được chấp nhận”; các bộ xử lý dữ liệu cũng không gửi thông báo “kết thúc giao tác” đến TM điều phối, mà gửi đến các bộ quản lý khóa yêu cầu giải phóng khóa và thông báo cho TM điều phối.



06

Thảo Luận