TIÊU ĐỀ CỦA BÀI BÁO, VIẾT CHỮ IN HOA ĐẬM

Nguyễn Văn Anha[[1]](#footnote-0), Trần Văn Hòab, Ngô Thị Bìnha,b

a Khoa Công nghệ thông tin, Trường Đại học Đà Lạt, Lâm Đồng, Việt Nam

bKhoa Công nghệ thông tin, Trường Đại học Tây Nguyên, Đắk Lắk, Việt Nam

Nhận ngày tháng năm

Chỉnh sửa ngày tháng năm | Chấp nhận đăng ngày tháng năm

Xuất bản trực tuyến ngày tháng năm

Tóm tắt

Phần này nêu nội dung nghiên cứu và các kết quả chính của bài báo; được soạn bằng mã Unicode, phông chữ Times New Roman, dài không quá 300 từ, cỡ chữ 10, in nghiêng, cách đều lề trái và lề phải 1 cm. Phần tóm tắt và từ khoá được viết bằng tiếng Việt và Anh. Phần tiếng Việt được đặt ở đầu bài báo và phần tiếng Anh được đặt ở cuối bài báo.

Từ khóa: Liệt kê từ 3 đến 6 từ khóa của bài báo, xếp theo thứ tự chữ cái đầu, phân cách các từ khóa bằng dấu chấm phẩy (;) soạn bằng mã Unicode, phông chữ Times New Roman, cỡ chữ 10, in thường.

### 3. Hệ cơ sở dữ liệu phân tán (Distributed database system\_ DDBS)

Một cơ sở dữ liệu phân tán là tập hợp nhiều cơ sở dữ liệu có liên đới logic và được phân bố trên một mạng máy tính. Từ đó ta định nghĩa: *Hệ quản trị cơ sở dữ liệu phân tán (DDMS) là một hệ thống phần mềm cho phép quản lý các cơ sở dữ liệu phân tán và làm cho việc phân tán trở nên “vô hình” đối với người sử dụng*.

Môi trường của hệ cở dữ liệu phân tán được mô tả qua hình 2

Trạm 5

Mạng truyền

dữ liệu

Trạm 5

Trạm 5

Trạm 5

Trạm 5

**Hình 2: Môi trường của hệ cơ sở dữ liệu phân tán**

### 4. Giao dịch và quản lý giao dịch:

**4.1. Các khái niệm**

Các đối tượng khác nhau của hệ không phải là các đối tượng độc lập nhau, chúng quan hệ bởi một tập các quan hệ gọi là một tập các ràng buộc toàn vẹn.

Trạng thái thoả mãn một tập các ràng buộc toàn vẹn gọi là trạng thái gắn bó.

Một CSDL ở trong một trạng thái nhất quán (gắn bó) nếu nó tuân theo tất cả các ràng buộc toàn vẹn (nhất quán) được định nghĩa trên nó.Sự thay đổi trạng thái xảy ra khi ta thực hiện các thao tác chèn, sửa hoặc xoá (gọi chung là cập nhật). Chúng ta cần đảm bảo rằng CSDL không bao giờ chuyển sang trạng thái không nhất quán. CSDL có thể tạm thời không nhất quán trong khi thực hiện giao dịch. điều quan trọng là CSDL phải trở về trạng thái nhất quán khi giao dich chấm dứt.

Một cơ chế cho phép duy trì sự gắn bó trong môi trường phân tán có sự cố phải là:

|  |  |
| --- | --- |
| STT | Phải thực hiện |
| 1 | Giao dich T bắt buộc phải được thực hiện một cách trọn vẹn |
| 2 | Nếu có sự cố diễn ra thì bắt buộc nó phải quay trở lại điểm xuất phát |

Muốn thực hiện những điều vừa nêu trong bảng trên, người ta dòi hỏi giao dịch phải có các đặc tính toàn vẹn như sau:

|  |  |
| --- | --- |
| STT | Đặc tính |
| 1 | Nếu một tiến trình bị sự cố trước khi kết thúc T nhưng lại sau các thao tác thay đổi cần thiết của T, trạng thái của hệ là gắn bó. |
| 2 | Nếu một tiến trình bị sự cố trước khi diễn ra các thay đổi của T, trạng thái của hệ là gắn bó. |
| 3 | Nếu một tiến trình bị sự cố giữa các thay đổi của T, trạng thái của hệ là không gắn bó. |

Trước khi cập nhật (thay đổi)

Trong khi cập nhật (thay đổi)

Sau khi cập nhật (thay đổi)

Gắn bó

Không gắn bó

Gắn bó

T

**Hình 3. Ba giai đoạn của một giao dịch**

Độ tin cậy hay khả tín (reliability) muốn nói đến khả năng tự thích ứng (resiliency) của một hệ thống đối với các loại sự cố và khả năng khôi phục lại từ những sự cố này. Một hệ thống tự thích ứng sẽ dung nạp được các sự cố hệ thống và có thể tiếp tục cung cấp các dịch vụ ngay cả khi xảy ra sự cố. Một DBMS hồi phục được là DBMS có thể chuyển sang trạng thái nhất quán ( bằng cách quay trở lại trạng thái nhất quán trước đó hoặc chuyển sang trạng thái nhất quán mới) sau khi gặp sự cố.

Quản lý giao dịch (transaction management) lo giải quyết các bài toán duy trì dược CSDL ở trong trình trạng nhất quán ngay cả khi có nhiều truy xuất đồng thời và khi có sự cố.

*Một giao dịch là một đơi vị tính toán nhất quán và đáng tin cậy. Vì thế về mặt trực quan, một giao dịch nhận một CSDL, thực hiện một hành động trên CSDL và sinh ra một “bản” CSDL mới gây ra một dich chuyển trạng thái. Điều này tương tự như điều mà câu vấn tin thực hiện, ngoại trừ trường hợp nếu CSDL nhất quán trước khi thực hiện giao dịch.*

**4.2 Tình huấn kết thúc giao dịch:**

Một giao dịch luôn luôn kết thúc ngay cả khi có xảy ra sự cố. Nếu giao dịch có thể hoàn tất thành công tác vụ của nó, chúng ta nói giao dịch đó uỷ thác (commit). ngược lại nếu một giao dịch phải ngừng lại khi chưa hoàn tất công việc, chúng ta nói nó bị huỷ bỏ (abort). Ngoài ra DBMS có thể huỷ bỏ một giao dịch, chẳng hạn do bị khoá cài (deadlock). Khi một giao dịch bị huỷ bỏ, quá trình thực thi sẽ ngùng và tất cả mọi hành động đã được thực hiện đều phải được hồi lại (undo), đưa CSDL trở về trạng thái trước khi thực hiện giao dịch. Quá trình này gọi là cuộn ngược (rollback).

Vai trò quan trọng của uỷ thác biểu hiện ở hai mặt. Thứ nhất lệnh uỷ thác thông tin cho DBMS biết rằng tác vụ của giao dịch đó bây giờ cần đượcphản ánh vào trong CSDL, qua đó làm cho các giao dịch đang truy xuất các mục dữ liệu đó có thể thấy được chúng. Thứ hai, điểm mà giao dịch uỷ thác là một điểm “không đường về”. Kết quả của một giao dịch đã uỷ thác bây giờ được lưu cố định vào CSDL và không thể hồi lại được.

**4.3. Đặc trưng hoá các giao dịch:**

Các giao dịch đều đọc và ghi một số dữ liệu. Điều này được dùng làm cơ sở nhận biết một giao dịch. Các mục dữ liệu được giao dịch đọc cấu tạo nên tập đọc RS (Read set) của nó. Tương tự, Các mục dữ liệu được giao dịch ghi được gọi là tập ghi WS (Write set). Chú ý rằng tập đọc và tập ghi của một giao dịch không nhất thiết phải tách biệt. Cuối cùng hợp của ttập đọc và tập ghi của một giao dịch tạo ra tập cơ sở BS (Base set), nghĩa là .

**4.4. Các tính chất của giao dịch**

Các khía cạnh nhất quán và khả tín của giao dịch là do bốn tính chất sau:

**4.4.1. Tính nguyên tử (atomicity)**

Tính nguyên tử liên quan đến sự kiện là một giao dịch được xử lý như một đơn vị hoạt tác. Chính vì thế mà các hành động của giao dịch, hoặc tất cả đều hoàn tất hoặc không một hành động nào hoàn tất. Điều này cũng được gọi là tính chất “được ăn cả, ngã về không” (all-or-nothing). Tính nguyên tử đòi hỏi rằng nếu việc thực thi giao dịch bị cắt ngang bởi một loại sự cố nào đó thì DBMS sẽ chịu trách nhiệm xác định công việc cần thực hiện đối với giao dịch để khôi phục lại sau sự cố bằng cách hoặc nó sẽ được kết thúc bằng cách hoàn tất các hành động còn lại, hoặc có thể được kết thúc bằng cách hồi lại tất cả các hành động đã được thực hiện

**4.4.2. Tính nhất quán (consistency):**

Tinh chất này chỉ đơn giản là tính đúng đắn của nó. Nói cách khác, một giao dịch là một chương trình đúng đắn, ánh xạ CSDL từ trạng thái nhất quán này sang trạng thái nhất quán khác. Việc đảm bảo tính nhất quán là mục tiêu của các cơ chế điều khiển đồng thời.

**4.4.3: Tính biệt lập (isolation):**

Biệt lập là tính chất của các giao dịch, đòi hỏi mỗi giao dịch phải luôn nhìn thấy CSDL nhất quán. Nói cách khác, một giao dịch đang thực thi không thể làm lộ ra các kết quả của nó cho những giao dịch khác đang cùng hoạt động trước khi nó uỷ thác.

Có một số lý do cần phải nhấn mạnh đến tính biệt lập, một là phải duy trì tính nhất quán qua lại giữa các giao dịch. Nếu hai giao dịch đồng thời truy xuất đến một dữ liệu đang được một trong chúng cập nhật thì không thể bảo đảm rằng giao dịch thứ hai sẽ đọc giá trị đúng.

**4.4.4: Tính bền vững (dulability)**

Muốn nói đến tính chất của giao dịch, bảo đảm rằng một khi giao dịch uỷ thác, kết quả của nó được duy trì cố định và không bị xoá ra khỏi CSDL. Vì thế DBMS bảo đảm rằng kết quả của giao dịch sẽ vẫn tồn tại dù có xảy ra sự cố hệ thống. Tính bền vững đưa ra vấn đề để khôi phục CSDL (database recovery), nghĩa là cách khôi phục CSDL về trạng thái gắn bó mà ở đó mọi hành động đã uỷ thác đều được phản ánh.

Với việc giới thiệu về những khái niệm giao dịch ở trên, chúng ta cần xem lại mô hình kiến trúc mà cụ thể là mở rộng vai trò của bộ phận theo dõi hoạt động phân tán (distributed execution monitor). Bộ phận này bao gồm hai đơn vị: Bộ quản lý giao dịch (transection manager\_TM) và bộ xếp lịch (scheduler\_SC). Bộ TM chịu trách nhiệm điều phối việc thực hiện các thao tác CSDL của các ứng dụng. Ngược lại bộ SC chịu trách nhiệm cài đặt một thuật toán điều khiển đồng thời cụ thể nhằm đồng bộ hoá việc truy xuất đến CSDL.

Phần thứ ba tham gia vào việc quản lý các giao dịch phân tán là bộ quản lý khôi phục cục bộ (local recovery manager) có mặt tại mỗi vị trí. Chức năng của chúng là cài đặt các thủ tục khôi phục khôi phục tại chỗ nhằm đưa CSDL về trạng thái gắn bó sau khi xảy ra sự cố.

Mỗi giao dịch đều xuất phát từ một vị trí và chúng ta gọi là vị trí khởi nguồn (originating site). Việc thực hiện các thao tác CSDL của một giao dịch được điều phối bởi TM tại vị trí khởi nguồn của giao dịch..

Bộ quản lý giao dịch cài đặt một giao diện cho các ứng dụng, bao gồm năm lệnh : begin\_transection, read, write, commit, abort. Xử lý mỗi lệnh này trong một hệ thống DBMS phân tán có nhân bản sẽ được từu tượng ở mức độ nhất định. Để cho đơn giản, chúng ta tạm bỏ qua việc xếp lịch cho các giao dịch đồng thờicũng như những chi tiết truy xuất dữ liệu vật lý của bộ xử lý dữ liệu . Những giả thiết này cho phép chúng ta tập trung vào giao diện của TM.

1. Begin\_transection: Đây là một chỉ điểm (indicator) cho TM rằng một giao dịch mới đang bắt đầu. TM sẽ ghi nhận điều này, chẳng hạn ghi lại tên giao dịch, ứng dụng gốc, vân vân.
2. Read: Nếu mục dữ liệu x được lưu cục bộ, giá trị của nó được đọc và chuyển cho giao dịch. Bằng không, TM sẽ chọn một bản sao của x và yêu cầu chuyển bản sao đó cho giao dịch.
3. Write: TM điều phối việc cập nhật giá trị x tại mỗi vị trí có chứa nó.
4. Commit: TM điều phối việc cập nhật vật lý của toàn thể CSDL có chứa bản sao của mỗi mục dữ liệu mà một lệnh write trước đó đã đưa ra.
5. Abort: TM kiểm tra lại rằng không có tác động nào của giao dịch được phản ánh trong CSDL.

Khi cung cấp các dịch vụ này, TM có thể giao tiếp với SC và bộ xử lý dữ liệu tại cùng vị trí hoặc khác vị trí.

Bộ xử lý giao dịch TM

Bộ xếp lịch SC

Bộ phận theo dõi

hoạt động phân tán

Xếp lịch

các thỉnh cầu

Với các bộ xử lý

dữ liệu

Với các SC khác

đến các bộ xử lýdữ liệu

Begin\_transectoin,

Read, Write,

Commit, Abort Kết quả

Hình 4. Mô hình chi tiết của bộ phận theo dõi hoạt động phân tán

Với các TM khác

### 

### Điều khiển đồng thời phân tán:

Điều khiển đồng thời giải quyết các tính chất biệt lập và nhất quán của giao dịch. Cơ chế điều khiển đồng thời phân tán của một hệ quản trị CSDL phân tán bảo đảm rằng tính chất nhất quán của CSDL. Nếu các giao dịch đã nhất quán nội tại (nghĩa là không vi phạm bất kỳ ràng buộc nhất quán nào), thế thì để có thể đạt được mục tiêu này, cách đơn giản nhất là thực hiện lần lượt từng giao dịch một. Rõ ràng điều này làm giảm lưu lượng hệ thống. Mức độ đồng thời (nghĩa là số lượng các giao dịch hoạt động cùng một lúc) có lẽ là một tham số quan trọng nhất trong các hệ phân tán. Do vậy cơ chế điều khiển đồng thời cố gắng tìm ra một phương án thích hợp, vừa duy trì được tính nhất quán của CSDL, vừa duy trì được mức độ đồng thời cao.

Chúng ta có thể chỉ ra rằng chức năng cơ bản của bộ điều khiển đồng thời là tạo ra một lịch biểu khả tuần tự để thực hiện các giao dịch đang chờ đợi.

Có một số cách phân loại các phương pháp điều khiển đồng thời. Một chế độ hiển nhiên là phân tán CSDL. Một số thuật toán đã được đề xuất đòi hỏi có một CSDL nhân bản hoàn toàn, còn một số khác có thể hoạt tác trên các CSDL phân hoạch hoặc nhân bản một phần. Có hai lớp thuật toán điều khiển đồng thời là lớp dựa trên các truy xuất độc quyền đến dữ liệu dùng chung (Khoá chốt) và lớp thuật toán cố gắn sắp xếp thứ tự thực hiện giao dịch theo một tập quy tắc (nghi thức). Các lớp thuật toán này dựa trên hai quan điểm khác nhau:

Quan điểm bi quan: Cho rằng có nhiều giao dịch sẽ trương tranh với nhau

Quan điểm lạc quan: cho rằng không có quá nhiều giao dịch trương tranh với nhau

Vì vậy chúng ta sẽ xếp các cơ chế điều khiển đồng thời thành hai nhóm lớn: Các phương pháp điều khiển đồng thời lạc quan và Các phương pháp điều khiển đồng thời bi quan . Các thuật toán bi quan đồng bộ hoá việc thực hiện đồng thời của các giao dịch trước khi thực hiện chúng , trong khi đó các thuật toán lạc quan để việc đồng bộ hoá các giao dịch cho đến khi chúng kết thúc.

Các thuật toán điều khiển đồng thời phân tán: Điều khiển đồng thời bằng khoá chốt, Điều khiển đồng thời bằng thời dấu, Điều khiển đồng thời lạc quan

### 6. Độ khả tín của hệ quản trị CSDL phân tán

Một hệ CSDL phân tán khả tín là hệ thống có thể tiếp tục xử lý các yêu cầu của người sử dụng gây cả khi hệ thống nền không khả tín. Nói khác khác ngây cả khi các thành phần của môi trường tính toán phân tán ngừng hoạt động, hệ quản trị CSDLphân tán khả tín phải có khả năng tiếp tục thực hiện các yêu cầu của người sử dụng mà vẫn không vi phạm tính nhất quán của CSDL.

**6.1. Hệ thống, trạng thái và sự cố:**

Trong ngữ cảnh của độ khả tín, thuật ngữ hệ thống (system) muốn nói đến một cơ chế được cấu tạo bởi một tập các thành phần và tương tác với môi trường bằng cách đáp ứng với các kích thích từ môi trường nhờ một mẫu hành vi nhận rađược

Thành phần 2

Thành phần1

Thành phần 3

Đáp ứng

Kích thích

hệ thống

môi trường

Hình 5. Sơ đồ hệ thống

Có nhiều cách để mô hình hoá sự trương tác giữa phần cứng và phần mềm

trong một hệ thống máy tính. Một phương pháp là xử lý mã lệnh như là bản thiết kế của một hệ thống từu trượng, các thành phần của nó là các đối tượng phần cứng và phần mềm mà chúng được thao tác trong quá trình thực hiện chương trình. Một cách khác là đặc tả trường minh mỗi thành phần phần cứng và phần mềm, và để bản thiết kế trình bày cách tương tác của các thành phần này với nhau.

Trạng thái ngoại giới của hệ thống có thể được định nghĩa là sự đáp ứng mà hệ thống đưa ra đối với kích thích bên ngoài, vì vậy có thể nói rằng hệ thống thay đổi các trạng thái ngoại giới theo các kích thích được lập từ môi trường. Chúng ta có thể định nghĩa tương tự về trạng thái nội tại của một hệ thống . cũng có thể định nghĩa trạng thái nội tại là hợp của các trạng thái ngoại giới của các thành phần tạo ra hệ thống. Và hệ thống cũng thay đổi trạng thái nội tại của nó để đáp lại kích thích từ môi trường.

Hành vi của hệ thống khi đáp ứng với tất cả các kích thích khả hữu từ môi trường cần được trình bày trong một đặc tả chính thức về hành vi của nó. Đặc tả chỉ rõ hành vi hợp lệ của mỗi trạng thái hệ thống .

Mọi sự lệch lạc của hệ thống khác với hành vi được nói đến trong đặc tả đều được xem là sự cố. Trong một trạng thái bất khả tín, hệ thống có thể chuyển sang một trạng thái nội tại không tuân theo đặc tả. Các dịch chuyển ra khỏi trạng thái này cuối cùng sẽ gây ra sự cố hệ thống, những trạng thái nội tại như thế gọi là trạng thái *sinh lỗi*, phần trạng thái không đúng được gọi là một *lỗi* trong hệ thống. Một lỗi trong trạng thái nội tại của các thành phần hoặc trong thiết kế của hệ thống được gọi là sai sót trong hệ thống. Vì vậy sai sót sẽ sinh ra lỗi và dẫn đến sự cố hệ thống

Hình 6. Nguồn gốc của các sự cố hệ thống

Sai sót bền

Thiết kế sai

Thành phần không ổn định

Môi trường

Không ổn định

Sai lầm

điều hành viên

Lỗi bền

Lỗi ngắt quảng

Lỗi thoáng qua

Sự cố

hệ thống

**6.2. Độ khả tín và độ khả dụng:**

- Độ khả tín (reliability) muốn nói đến xác suất mà hệ thống đang xét không trải qua bất kỳ sự cố trong một khoảng thời gian đã cho. Điển hình nó được sử dụng để mô tả các hệ thống không sửa chữa được (như các máy tính dùng trong không gian), hoặc trong các hoạt động quan trọng đến nỗi mọi thời gian dừng lại để sử chữa đều không được chấp nhận

- Độ khả dụng (availability) Nói đến xác suất hệ thống còn hoạt động theo đặc tả của nó tại một thời điểm t đã cho. Một số sự cố có thể xảy ra trước thời điểm t, nhưng nếu chúng đã được sửa chữa thì hệ thống vẫn khả dụng tại thời điểm t. rõ ràng là độ khả dụng đề cập đến các hệ thống có thể sửa chữa được

**6.3 Sự cố trong hệ quản trị CSDL phân tán**

**6.3.1 Nguyên do của sự cố:**

Các sự cố mềm chiếm hơn 90% các sự cố của phần cứng. Có ba nghiên cứu số liệu thống kê đã chỉ ra phần trăm các loại sự cố như sau:



Hệ thống 17%

Phần cứng 18%

Phần mềm 25%

Môi trường 15%

Bảo trì 25%

(b) Dữ liệu Tandem



Phần cứng 57%

Hệ thống 14%

Phần mềm 12%

Môi trường 17%

(a) Dữ liệu SLAC



Phần cứng 32,3%

Phần mềm 44,3%

Hệ thống 17,5%

Không biết 5,9%

(c) Dữ liệu bộ chuyển mạch 5ESS

**6.3.2.Các loại sự cố ( Có 4 loại sự cố)**

**- Sự cố giao dịch :** Giao dịch có thể bị sự cố vì một số lý do. Sự cố có thể do một lỗi trong giao dịch, gây ra bởi dữ liệu nhập không đúng hoặc do phát hiện ra một khoá cài tiềm tàng hoặc hiện hữu. Hơn nữa một số thuật toán điều khiển đồng thời không cho phép một giao dịch tiến hành hoặc thậm chí là chờ đợi nếu dữ liệu mà giao dịch muốn truy xuất đang được một giao dịch khác truy xuất. Điều này cũng có thể được xem là một sự cố. Cách tiếp cận thông thường cho sự cố giao dịch là huỹ bỏ giao dịch, sau đó dặt lại CSDL về trạng thái của nó trước khi khởi động giao dịch này.

- **Sự cố vị trí (Hệ thống):** Sự cố này có thể từ sự cố phần cứng (bộ xử lý, bộ nhớ chính, nguồn điện , vân vân)hoặc từ sự cố phần mềm (có bọ trong hệ điều hành hoặc hệ quản trị CSDL). Điểm quan trọng là một sự cố hệ thống luôn làm mất nội dung của bộ nhớ chính. Vì thế dẫn đến việc đánh mất mọi phần của CSDL đang có trong vùng nhớ đệm. Tuy nhiên CSDL lưu trong bộ nhớ thứ cấp được xem là an toàn và chính xác. Nói theo thuật ngữ CSDL phân tán, sự cố hệ thống thường được gọi sự cố vị trí vì chúng làm cho vị trí đó ngưng hoạt động và những vị trí khác trong hệ thống phân tán không thể tiếp cận được nó.

**- Sự cố vật liệu (đĩa):** Sự cố này liên quan đến sự cố của các thiết bị lưư trữ thứ cấp dùng để lưu trữ CSDL. Những sự cố này có thể do lỗi của hệ điều hành cũng như những sai sót phần cứng, chẳng hạng như vỡ đầu đọc hoặc hư hại bộ điều khiển. Điểm quan trọng theo quan điểm độ khả tín của DBMS đó là tất cả hoặc một phần CSDL trên thiết bị đó được xem như bị huỷ hoại hoặc không thể truy xuất được. Ghép đôi đĩa và duy trì các bản lưu của CSDL là những kỹ thuật thông thường để giải quyết loại vấn đề này.

**- Sự cố đường liên lạc ( truyền thông):**

Ba sự cố trên đều có trong hệ tập trung lẫn phân tán, còn sự cố truyền thông chỉ có trong các hệ phân tán. Sự cố này là do lỗi trong các thông báo, các thông báo vô trật tự, thông báo bị thất lạc (hoặc không phân phối thông báo) và sự cố đường liên lạc.

Hai lỗi đầu tiên thuộc trách nhiệm của mạng máy tính, còn các thông báo bị thất lạc hoặc không phân phối được thường là hậu quả các sự cố đường truyền hoặc sự cố vị trí đến. Nếu đường truyền bị hư, ngoài vấn đề làm mất các thông báo tuyền qua, nó cũng có thể phân cắt mạng thành hai hoặc nhiều nhóm tách rời. Tình huống này được gọi là phân hoạch mạng, lúc đó các vị trí trong mỗi phân hoạch có thể vẫn tiếp tục hoạt động. Khi đó việc thực hiện các giao dịch cần truy xuất đến nhiều phân hoạch trở thành một vấn đề quan trọng. Duy trì tính nhất quán tương hỗ của CSDL là một vấn đề có ý nghĩa quan trọng, đặc biệt nếu CSDL được nhân bản trong nhiều phân hoạch.

Nếu các thông báo không phân phối được, ta giả thiết rằng mạng sẽ không làm gì, nó sẽ phân phối cho các nơi khi dịch vụ được tái lập lại và sẽ không thông tin cho tiến trình gửi biết rằng thông báo không thể phân phối đi được. Nói vắn tắt, thông báo chỉ đơn giản là bị thất lạc. Chúng ta đưa ra giả thiết này bởi vì nó biểu diễn sự mong chờ ít nhất từ mạng và trao trách nhiệm giải quyết những sự cố này cho DBMS phân tán.

Một kết quả từ giả thiết này là trách nhiệm phát hiện ra rằng một thông báo không phân phối được sẽ để lại cho chương trình ứng dụng (trong trường hợp này là các DBMS phân tán).Việc phát hiện này tạo ra hướng giải quyết là dùng bộ đếm thời gian và cơ chế quá hạn để theo dõi xem đã qua bao nhiêu lâu kể từ khi vị trí gửi không nhận được hồi đáp của vị trí đến. Thời gian quá hạn nhất thiết phải được đặt lớn hơn khoảng thời gian trể truyền tải tối đa của một thông báo trên mạng.

**6.4. Nghi thức khả tín cục bộ:**

**6.4.1. Các vấn đề kiến trúc:**

CSDL ổn định được lưư trên thiết bị ổn định, thiết bị ổn định thường được cài đặt bằng phương tiện các đĩa từ ghép đôi với khả năng lưu các bản dữ liệu giống nhau luôn cần phải gửi nhất quán tương hỗ.

Bộ quản lý đệm BM gửi một số dữ liệu mới được truy xuất trong vùng đệm, điều này làm tăng hiệu quả truy xuất, phần CSDL đang hiện diện trong vùng đệm gọi là CSDL khả hủy. LRM (bộ quản lý khôi phục cục bộ)thực hiện các thao tác thay cho giao dịch chỉ trên CSDL khả hủy mà về sau sẽ được ghi lại vào CSDL ổn định.

Bộ quản lý

khôi phục cục bộ

Bộ quản lý

vùng đệm CSDL

Vùng đệm

CSDL

CSDL

khả hủy

Read

Write

CSDL

ổn định

Fetch

Flush

Bộ nhớ chính

Bộ nhớ thứ cấp

Read

Write

Hình 7. Giao diện giữa LRM và BM

Khi LRM muốn đọc một trang dữ liệu dùm cho một giao dịch, chính xác là cho một thao tác của giao dịch. Nó đưa ra lệnh **fetch,** nộp vào trang mà nó muốn đọc. BM kiểm tra xem trang đó cá trong vùng đệm hay không và nếu có trong đó thì trao cho giao dịch đang yêu cầu; nếu không, nó sẽ đến CSDL ổn định và đọc trang đó vào chỗ còn trống trong vùng đệm, nếu không còn chỗ trống, nó chọn một trong những trang đệm và ghi trở lại vào bộ nhớ ổn định và đọc trang đang yêu cầu vào chỗ đó. LRM ghi lại trang đệm nhờ lênh **Flush.**

**6.4.2. Các thông tin khôi phục**

Ta giả thiết rằng chỉ xảy ra các sự cố hệ thống, khi xảy ra sự cố hệ thống, CSDL khả hủy sẽ bị mất. Vì thế DBMS phải duy trì một số thông tin về trạng thái của nó vào lúc xảy ra sự cố, nhờ đó có khả năng đưa CSDL trở về trạng thái hiện hữu của nó khi có sự cố. Chúng ta gọi đây là thông tin khôi phục.

**- Thông tin khôi phục về cập nhật tại chỗ:** Cập nhật tại chỗ sẽ làm mất giá trị trước đó của mục dữ liệu bị ảnh hưởng. Ta cần gửi đủ các thông tin cần thiết về những thay đổi trạng thái CSDL, tạo dể dàng cho việc khôi phục CSDL trở lại trạng thái gắn bó sau một sự cố, thông tin này được duy trì trong nhật ký CSDL . Vì vậy mỗi giao dịch cập nhật không chỉ thay đổi CSDL mà còn được ghi nhận trong nhật ký CSDL.

Trạng thái

ổn định cũ

Trạng thái

ổn định mới

Nhật ký CSDL

Thao tác

cập nhật

Hình 8. Thực thi các thao tác cập nhật

**6.5. Các nghi thức khả tín phân tán:**

Giống như các nghi thức khả tín cục bộ, các phiên bản phân tán nhằm duy trì tính chất nguyên tử và tính chất bền vững của các giao dịch phân tán thực hiện trên một số CSDL. Các nghi thức này tập trung giải quyết các vấn đề thực hiện phân tán các lệnh **begin\_transection, read, write, commit, recover.** Chúng ta giả thiết rằng tại vị trí nguồn của giao dịch có một tiến trình thực hiện các thao tác của nó. Tiến trình này được gọi là điều phối viên, điều phối viên trao đổi với các thành viên tại những vị trí có tham gia vào việc thực hiện các thao tác của giao dịch

**- Các thành phần của nghi thức khả tín phân tán:**

Kỹ thuật khả tín bao gồm ba nghi thức: uỷ thác, kết thúc và khôi phục. Các nghi thức ủy thác và khôi phục đặc tả cách thức thực hiện các lệnh ủy thác và khôi phục. Còn nghi thức kết thúc là như nhau đối với đối với các hệ phân tán. Giả sử rằng trong khi thực hiện một giao dịch phân tán, một trong các vị trí có tham gia vào việc thực hiện này có sự cố; chúng ta muốn các vị trí còn lại kết thúc giao dịch bằng cách nào đó. Kỹ thuật nhằm giải quyết tình huống này được gọi là nghi thức kết thúc.

### 7. Giải thuật đảm bảo gắn bó dữ liệu

Có nhiều loại giải thuật đảm bảo gắn bó dữ liệu. Ở đây ta đưa ra giải thuật hai pha tuyến tính (Linear two phase commit - 2PC). Trong đó các thành viên có thể trao đổi với nhau. Chúng ta giả thiết rằng thứ tự giữa các vị trí có tham gia vào việc thực hiện một giao dịch là 1, 2,…,N với điều phối viên là vị trí đầu tiên giải thuật này hoạt động như sau:

Điều phối viên gửi thông báo prepare đến thành viên 2. Nếu thành viên 2 chưa sẵn sàng ủy thác giao dịch, nó gửi thông báo biểu quyết hủy bỏ Vote-abort (VA) đến thành viên 3 và giao dịch bị hủy tại thời điểm này (hủy bỏ đơn phương của 2). Ngược lại nếu thành viên 2 đồng ý ủy thác, nó gửi thông báo vote-commit (VC) cho thành viên 3 rồi chuyển sang trạng thái READY. Quá trình này tiếp tục cho đến khi một biểu quyết uỷ thác đến được thành viên N. Đến đây kết thúc pha đầu tiên. Nếu N quyết định ủy thác nó gửi trở lại cho thành viên N-1 thông báo global-commit (GC); bằng không, nó gửi một thông báo hủy bỏ toàn cục global-abort (GA). Theo đó các thành viên chuyển sang trạng thái thích hợp (COMMIT hoặc ABORT) và làm lan truyền thông báo trở về điều phối viên

1

2

3

4

5

N

prepare

VC/VA

VC/VA

VC/VA

VC/VA

GC/GA

GC/GA

GC/GA

GC/GA

GC/GA

Pha 1

Pha 2

Hình 9. Cấu trúc giao tiếp 2PC tuyến tính

1. **Tác giả liên hệ**: Email: [nguyenvananh@dlu.edu.vn](mailto:nguyenvananh@dlu.edu.vn) (Vui lòng sử dụng email công vụ, nếu có) | Điện thoại: [↑](#footnote-ref-0)