

NỘI DUNG

2.1 Kiến trúc cơ bản của một cơ sở dữ liệu phân tán

2.2 Các đặc điểm chính của hệ phân tán

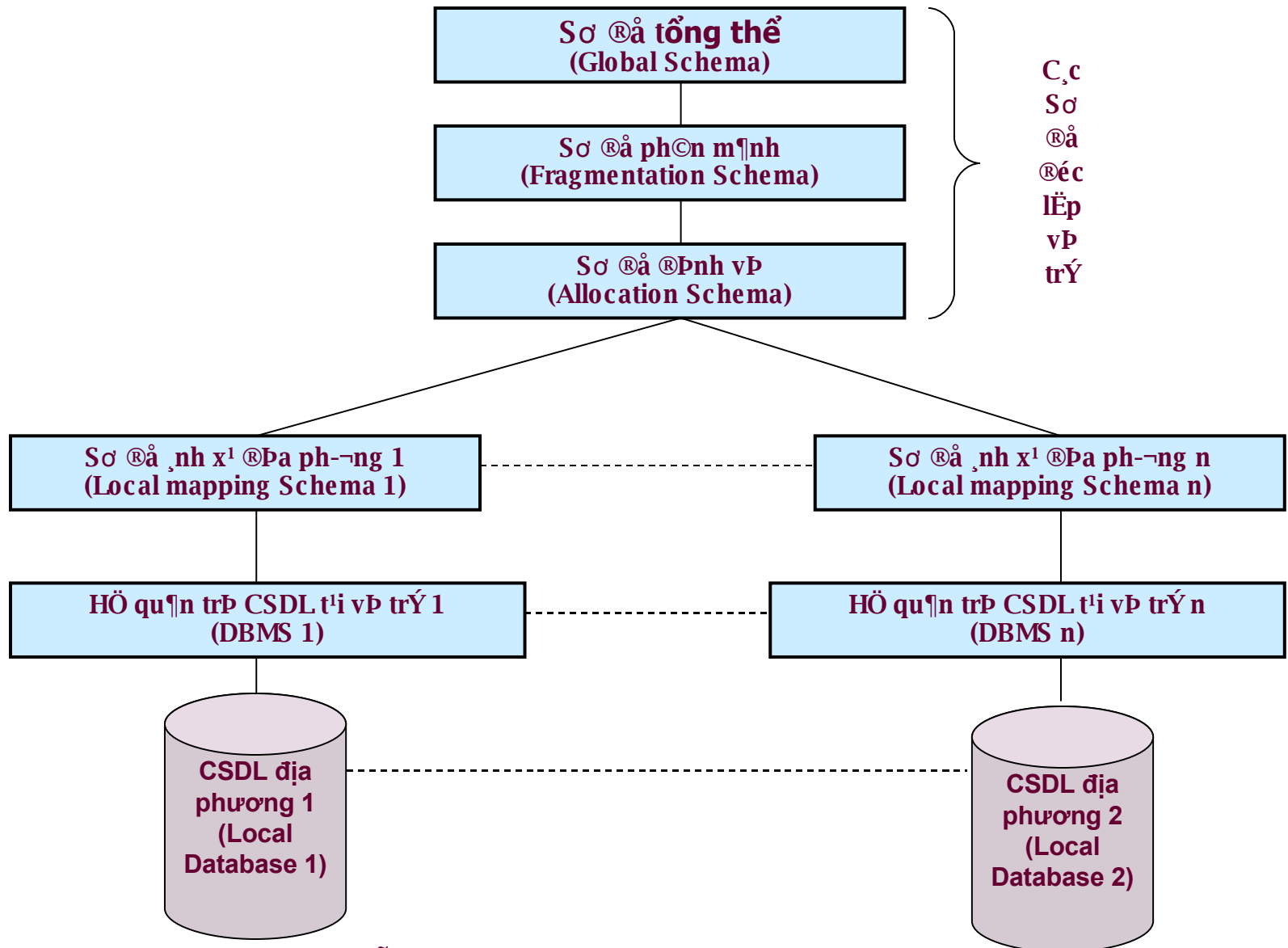
2.3 Trong suốt phân tán

2.4 Tổ chức hệ thống phân tán

MỤC ĐÍCH

Cung cấp cho người sử dụng thấy được các mức trong suốt khác nhau được cung cấp của một DDBMS.

2.1 KIẾN TRÚC CƠ BẢN CỦA CSDL PHÂN TÁN



Kiến trúc tham khảo cho CSDL phân tán

2.1 Kiến trúc cơ bản của một cơ sở dữ liệu phân tán

a. Sơ đồ tổng thể (*Global Schema*):

- Xác định tất cả các dữ liệu sẽ được lưu trữ trong cơ sở dữ liệu phân tán cũng như các dữ liệu không được phân tán ở các trạm trong hệ thống.
- Sơ đồ tổng thể được định nghĩa theo cách như trong CSDL tập trung.
- Trong mô hình quan hệ, sơ đồ tổng thể bao gồm định nghĩa của tập các quan hệ tổng thể (*Globale relation*) .

2.1 Kiến trúc cơ bản của một cơ sở dữ liệu phân tán

b. Sơ đồ phân đoạn (fragment schema):

- Mỗi quan hệ tổng thể có thể chia thành một vài phần không giao nhau gọi là phân đoạn (fragment).
- Có nhiều cách khác nhau để thực hiện việc phân chia này
- Sơ đồ phân đoạn mô tả các ánh xạ giữa các quan hệ tổng thể và các đoạn được định nghĩa trong sơ đồ phân đoạn (fragmentation Schema),
- Các đoạn được mô tả bằng tên của quan hệ tổng thể cùng với chỉ mục đoạn. Chẳng hạn, R_i được hiểu là đoạn thứ i của quan hệ R .

2.1 Kiến trúc cơ bản của một cơ sở dữ liệu phân tán

c. Sơ đồ định vị (allocation schema):

- Các đoạn là các phần logic của một quan hệ tổng thể được định vị vật lý trên một hay nhiều trạm.
- *Sơ đồ định vị* xác định đoạn dữ liệu nào được định vị tại trạm nào trên mạng.
- Tất cả các đoạn được liên kết với cùng một quan hệ tổng thể R và được định vị tại cùng một trạm j cấu thành ảnh vật lý quan hệ tổng thể R tại trạm j.
- Do đó ta có thể ánh xạ một-một giữa một ảnh vật lý và một cặp (quan hệ tổng thể, trạm).
- Các ảnh vật lý có thể chỉ ra bằng tên của một quan hệ tổng thể và một chỉ mục trạm.

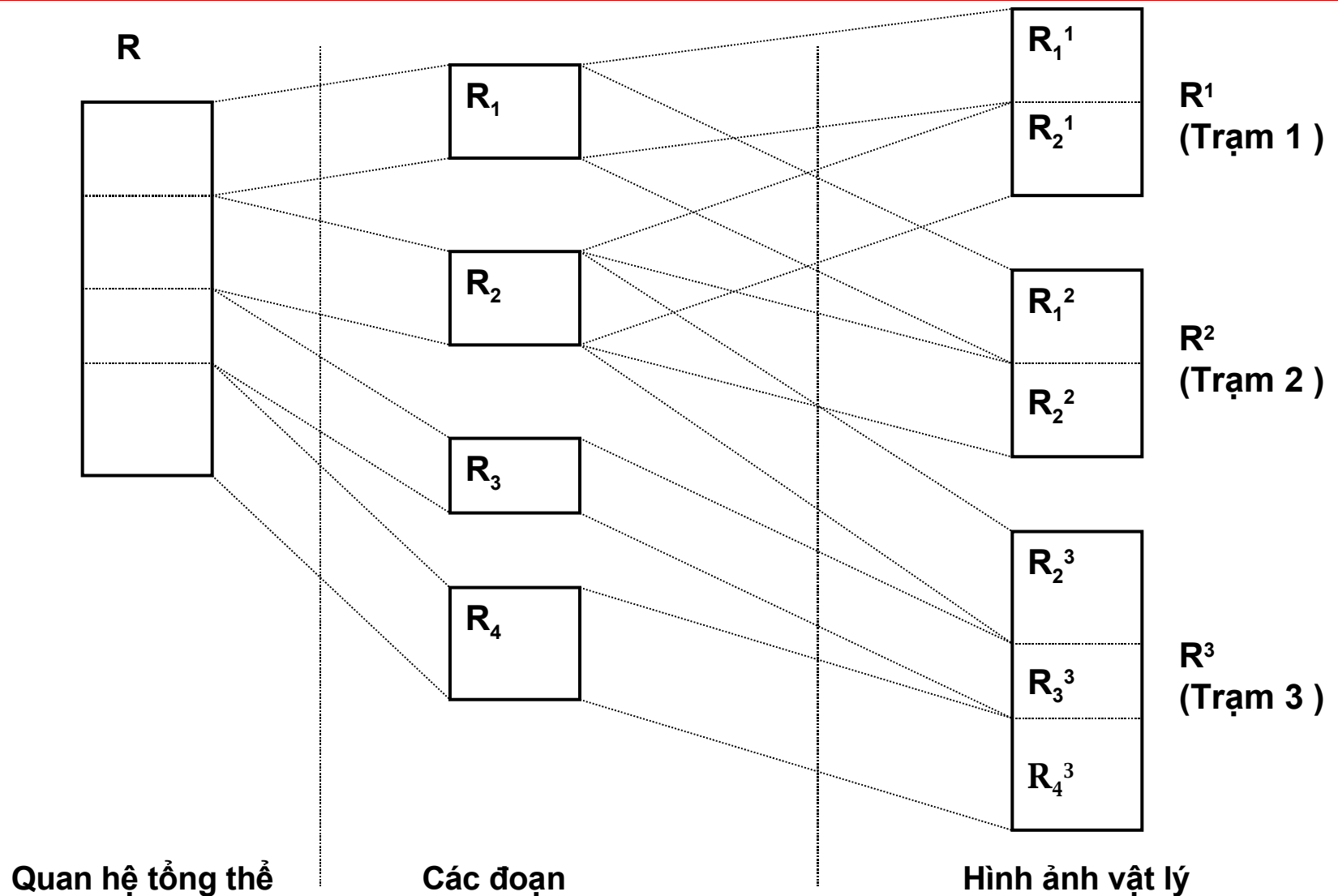
2.1 Kiến trúc cơ bản của một cơ sở dữ liệu phân tán

- Ký hiệu R_i để chỉ đoạn thứ i của quan hệ tổng thể R
- Ký hiệu R_i^j để chỉ ảnh vật lý của quan hệ tổng thể R tại trạm j
- Tương tự như vậy, bản sao của đoạn i thuộc quan hệ R tại trạm j được ký hiệu là R_i^j

d. Sơ đồ ánh xạ địa phương (Local mapping schema):

- Thực hiện ánh xạ các ảnh vật lý lên các đối tượng được thực hiện bởi hệ quản trị cơ sở dữ liệu địa phương
- Tất cả các đoạn của một quan hệ tổng thể trên cùng một trạm tạo ra một ảnh vật lý

2.1 Kiến trúc cơ bản của một cơ sở dữ liệu phân tán



Các đoạn và hình ảnh vật lý của một quan hệ tổng thể

2.1 Kiến trúc cơ bản của một cơ sở dữ liệu phân tán

Ba yếu tố được suy ra từ kiểu kiến trúc này là:

- a. Tách rời khái niệm phân đoạn dữ liệu với khái niệm định vị dữ liệu.
- b. Biết được dữ liệu dư thừa
- c. Độc lập với các DBMS địa phương

Ba yếu tố này tương ứng với ba mức trong suốt tương ứng

2.1 Kiến trúc cơ bản của một cơ sở dữ liệu phân tán

a. Tách rời khái niệm phân đoạn dữ liệu với khái niệm định vị dữ liệu.

- **Phân đoạn dữ liệu**, bao gồm những công việc mà người lập trình ứng dụng làm việc với quan hệ tổng thể, phân chia quan hệ tổng thể thành các đoạn.

- Thông qua tính **trong suốt phân đoạn** (*fragmentation transparency*) người lập trình sẽ nhìn thấy được những đoạn dữ liệu bị phân chia như thế nào.

- **Định vị dữ liệu** lại liên quan đến các công việc của người sử dụng và người lập trình ứng dụng tại trên các đoạn dữ liệu được định vị tại các trạm.

- Thông qua tính **trong suốt vị trí** (*location transparency*) người lập trình sẽ biết được vị trí của các đoạn dữ liệu trên các trạm.

2.1 Kiến trúc cơ bản của một cơ sở dữ liệu phân tán

b. Biết được dữ liệu dư thừa:

- Người lập trình ứng dụng có thể biết được dư thừa dữ liệu ở các trạm.
- Trên hình vẽ trên, chúng ta thấy rằng hai ảnh vật lý R_2 và R_3 có trùng lặp dữ liệu. Do đó các đoạn dữ liệu trùng nhau có thể tránh được khi xây dựng các khối ảnh vật lý.

c. Độc lập với các DBMS địa phương

Tính chất này còn được gọi là ***trong suốt ánh xạ địa phương*** (*local mapping transparency*), cho phép chúng ta khảo sát các vấn đề về quản lý CSDL phân tán mà không cần phải hiểu rõ mô hình dữ liệu của DBMS địa phương đang sử dụng.

2.2 CÁC ĐẶC ĐIỂM CHÍNH CỦA HỆ PHÂN TÁN

2.2.1 *Chia sẻ tài nguyên*

- Được thực hiện qua mạng truyền thông.
- Mỗi tài nguyên cần phải được quản lý bởi một chương trình có giao diện truyền thông
 - Các tài nguyên có thể được truy nhập, cập nhật một cách tin cậy và nhất quán.

Quản lý tài nguyên: bao gồm

- Lập kế hoạch dự phòng
- Đặt tên cho các lớp tài nguyên
- Cho phép tài nguyên được truy nhập từ nơi này đến nơi khác
 - Ánh xạ tên tài nguyên vào địa chỉ truyền thông.

2.2 CÁC ĐẶC ĐIỂM CHÍNH CỦA HỆ PHÂN TÁN

2.2.2 Tính mở

Tính mở của hệ thống phân tán là tính dễ dàng mở rộng phần cứng của nó. Một hệ thống được gọi là có tính mở thì phải có các điều kiện sau:

- Hệ thống có thể tạo nên bởi nhiều loại phần cứng và phần mềm của nhiều nhà cung cấp khác nhau.
- Có thể bổ sung vào các dịch vụ dùng chung tài nguyên mà không phá hỏng hay nhân đôi các dịch vụ đang tồn tại.
- Tính mở được hoàn thiện bằng cách xác định hay phân định rõ các giao diện chính của một hệ và làm cho nó tương thích với các nhà phát triển phần mềm.
- Tính mở của hệ phân tán dựa trên việc cung cấp cơ chế truyền thông giữa các tiến trình và công khai các giao diện dùng để truy nhập các tài nguyên chung.

2.2 CÁC ĐẶC ĐIỂM CHÍNH CỦA HỆ PHÂN TÁN

2.2.3 Khả năng song song

- Hệ phân tán hoạt động trên một mạng truyền thông có nhiều máy tính, mỗi máy có thể có một hay nhiều CPU.
- Có thể thực hiện nhiều tiến trình trong cùng một thời điểm. Việc thực hiện tiến trình theo cơ chế phân chia thời gian (một CPU) hay (nhiều CPU).

Khả năng làm việc song song trong hệ phân tán được thể hiện qua hai tình huống sau:

- Nhiều người sử dụng đồng thời đưa ra các lệnh hay các tương tác với các chương trình ứng dụng.
- Nhiều tiến trình Server chạy đồng thời, mỗi tiến trình phải đáp ứng yêu cầu từ các Clients._

2.2 CÁC ĐẶC ĐIỂM CHÍNH CỦA HỆ PHÂN TÁN

2.2.4 Khả năng mở rộng

Khả năng mở rộng của một hệ phân tán được đặc trưng bởi tính không thay đổi phần mềm hệ thống và phần mềm ứng dụng khi hệ được mở rộng.

Yêu cầu cho việc mở rộng không chỉ là mở rộng phần cứng, về mạng mà nó trải trên các khía cạnh khi thiết kế hệ phân tán.

Ví dụ: tần suất sử dụng trên mạng đột ngột. Để tránh tình trạng tắc nghẽn xảy ra khi chỉ có một Server và phải đáp ứng các yêu cầu truy nhập các file đó. Người ta nhân bản các file trên một Server khác và hệ thống được thiết kế sao cho việc thêm Server được dễ dàng. Một số giải pháp khác là sử dụng Cache và các bảng sao dữ liệu.

2.2 CÁC ĐẶC ĐIỂM CHÍNH CỦA HỆ PHÂN TÁN

2.2.5 Khả năng thứ lỗi

Việc thiết kế khả năng thứ lỗi các hệ thống máy tính dựa trên hai giải pháp sau:

- Dùng khả năng thay thế để đảm bảo sự hoạt động liên tục và hiệu quả.
- Dùng các chương trình hồi phục dữ liệu khi xảy ra sự cố.

2.2 CÁC ĐẶC ĐIỂM CHÍNH CỦA HỆ PHÂN TÁN

2.2.6 Đảm bảo tin cậy và nhất quán

Hệ thống yêu cầu độ tin cậy như:

- . Bí mật của dữ liệu
- . Các chức năng khôi phục hư hỏng phải đảm bảo
- . Ngoài ra các yêu cầu của hệ thống về tính nhất quán

cũng thể hiện ở chỗ: không có mâu thuẫn trong nội dung cơ sở dữ liệu

2.3 TRONG SUỐT PHÂN TÁN

2.2.7 Tính trong suốt

Tính trong suốt của một hệ phân tán được hiểu như là việc che khuất đi các thành phần riêng biệt của hệ đối với người sử dụng và những người lập trình ứng dụng.

Các loại trong suốt trong hệ phân tán:

d. Trong suốt phân đoạn (fragmentation transparency)_

e. Trong suốt về vị trí (location transparency)_

f. Trong suốt ánh xạ địa phương (local mapping transparency)_

g. Không trong suốt (no transparency)_

2.3 TRONG SUỐT PHÂN TÁN

a. Trong suốt phân đoạn (*fragmentation transparency*):

Khi dữ liệu đã được phân đoạn thì việc truy cập vào CSDL được thực hiện bình thường như là chưa bị phân tán và không ảnh hưởng tới người sử dụng.

Ví dụ: Xét quan hệ tổng thể **NCC (Id, Tên, Tuổi)**

và các phân đoạn được tách ra từ nó:

NCC1 (Id, Tên, Tuổi)

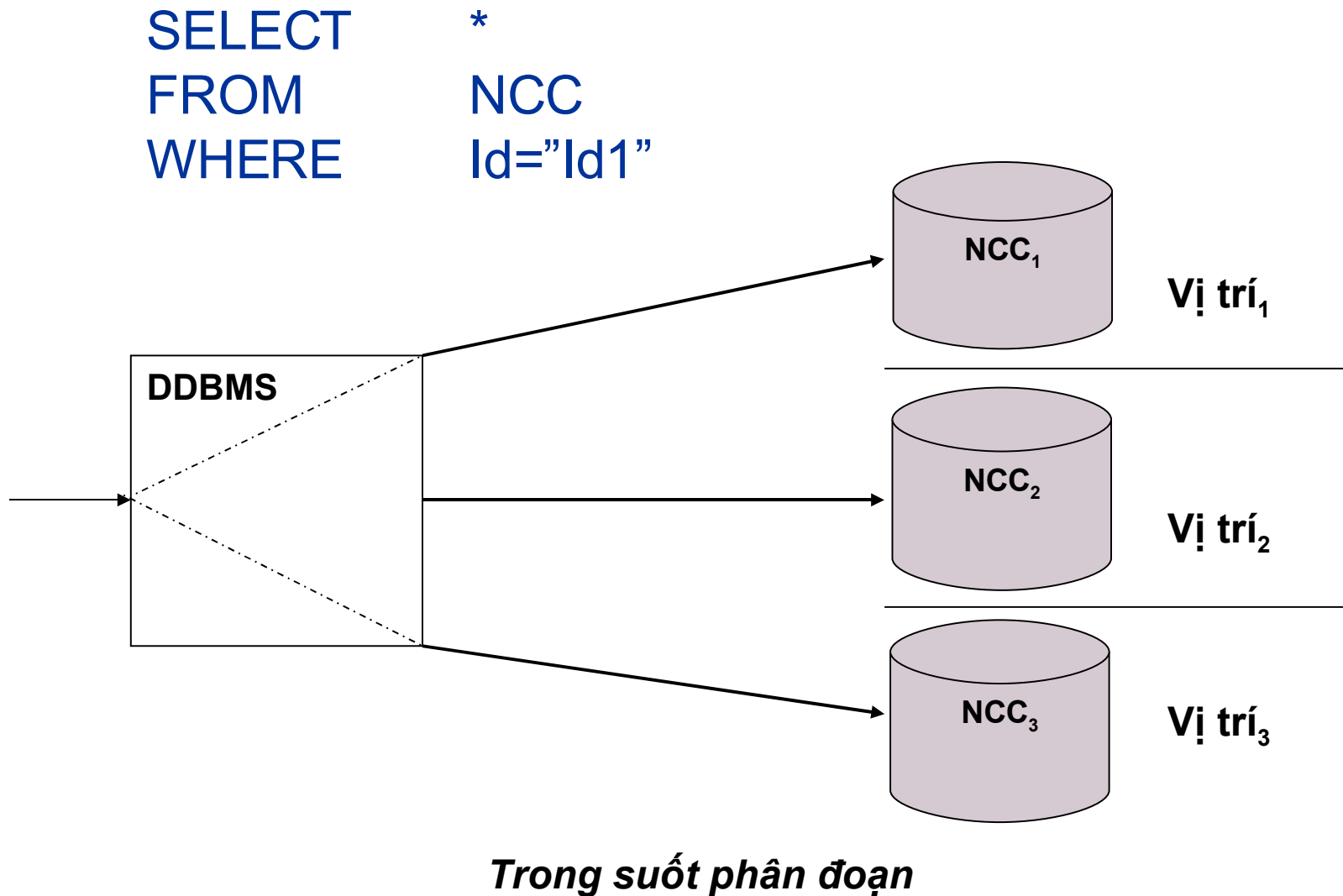
NCC2 (Id, Tên, Tuổi)

NCC3 (Id, Tên, Tuổi)

Giả sử DDBMS cung cấp tính trong suốt về phân đoạn, khi đó ta có thể thấy tính trong suốt này được thể hiện như sau:

Khi muốn tìm một người có **Id="Id1"** thì chỉ cần tìm trên quan hệ tổng thể NCC mà không cần biết quan hệ NCC có phân tán hay không.

2.3 TRONG SUỐT PHÂN TÁN



2.3 TRONG SUỐT PHÂN TÁN

b. Tính trong suốt về vị trí (location transparency):

- Người sử dụng không cần biết về vị trí vật lý của dữ liệu mà có quyền truy cập đến cơ sở dữ liệu tại bất cứ vị trí nào.
- Các thao tác để lấy hoặc cập nhật một dữ liệu từ xa được tự động thực hiện bởi hệ thống tại điểm đưa ra yêu cầu.
- Tính trong suốt về vị trí rất hữu ích, nó cho phép người sử dụng bỏ qua các bản sao dữ liệu đã tồn tại ở mỗi vị trí. Do đó có thể di chuyển một bản sao dữ liệu từ một vị trí này đến một vị trí khác và cho phép tạo các bản sao mới mà không ảnh hưởng đến các ứng dụng.

2.3 TRONG SUỐT PHÂN TÁN

Ví dụ: Với quan hệ tổng thể R và các phân đoạn như đã nói ở trên nhưng giả sử rằng DDBMS cung cấp trong suốt về vị trí nhưng không cung cấp trong suốt về phân đoạn.

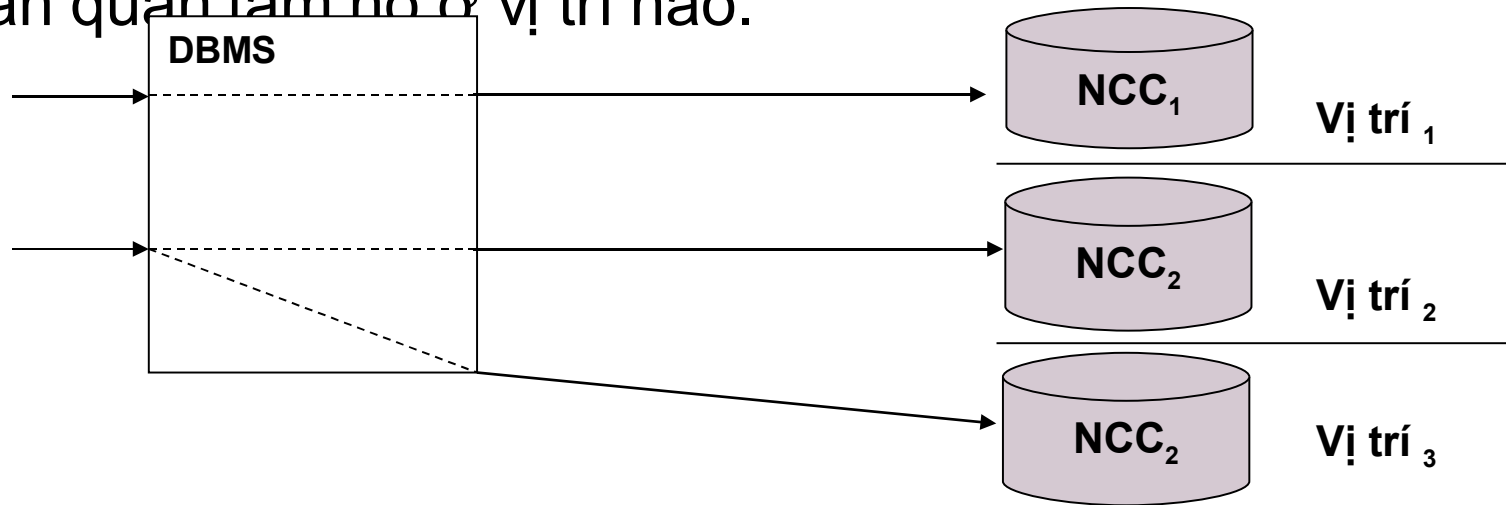
Xét câu truy vấn *tìm người có Id="ld1"*.

```
SELECT      *
FROM        NCC1
WHERE       Id="ld1"

            IF NOT #FOUND THEN
SELECT      *
FROM        NCC2
WHERE       Id="ld1"
```

2.3 TRONG SUỐT PHÂN TÁN

- Đầu tiên hệ thống sẽ thực hiện tìm kiếm ở phân đoạn NCC_1 và nếu DBMS trả về biến điều khiển #FOUND thì một câu lệnh truy vấn tương tự được thực hiện trên phân đoạn NCC_2
- ,...
- Ở đây quan hệ NCC_2 được sao làm hai bản trên hai vị trí₂ và vị trí₃, ta chỉ cần tìm thông tin trên quan hệ NCC_2 mà không cần quan tâm nó ở vị trí nào.

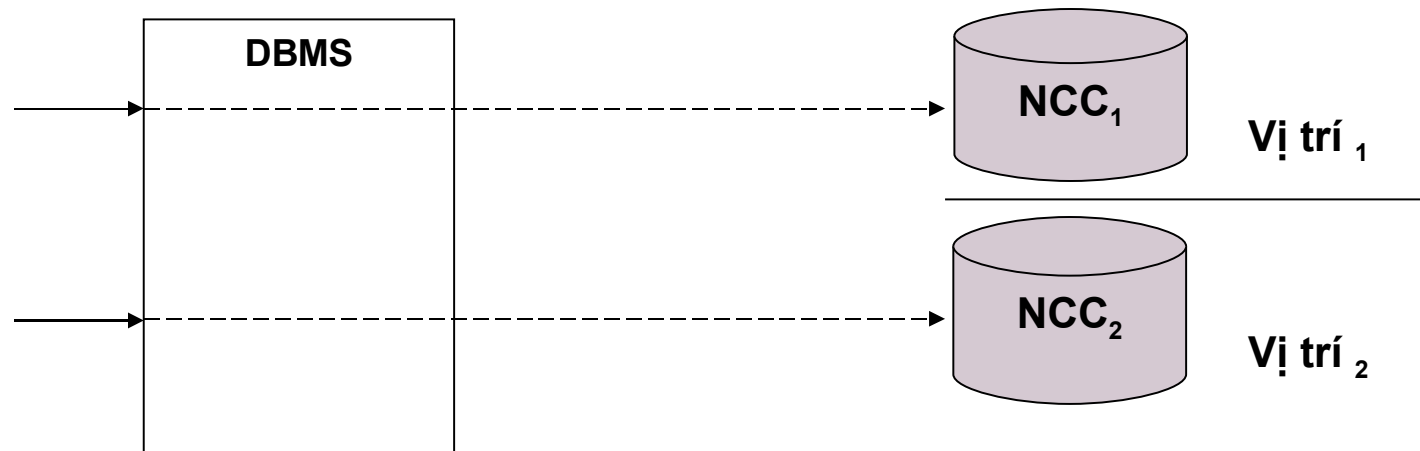


Sự trong suốt về vị trí

2.3 TRONG SUỐT PHÂN TÁN

c. Trong suốt ánh xạ địa phương (*local mapping transparency*):

- Là một đặc tính quan trọng trong một hệ thống DBMS không đồng nhất
- Ứng dụng tham chiếu đến các đối tượng có các tên độc lập từ các hệ thống cục bộ địa phương.
- Ứng dụng được cài đặt trên một hệ thống không đồng nhất nhưng được sử dụng như một hệ thống đồng nhất.



Sự trong suốt ánh xạ địa phương

CHƯƠNG 2. CÁC MỨC TRONG SUỐT



HẾT CHƯƠNG 2