

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

CƠ SỞ DỮ LIỆU PHÂN TÁN

TỔNG QUAN CƠ SỞ DỮ LIỆU PHÂN TÁN

Ts. Phan Thị Hà

Học Viên Công nghệ BCVT, Phan Thị
Hà, hathiphan@yahoo.com

- ☞ Phạm Quang Chiến: chientpit.94@gmail.com - 01677657894
- ☞ Nguyễn Duy Tùng: tungnd.ptit@gmail.com – 0986585619
- ☞ **Bùi Văn Đạm : itbuivandam@gmail.com- 0982782663**
- ☞ **Trần Hoàng Linh : linth9x@gmail.com**
- ☞ **01655673374**
- ☞ **NHóm chiều thứ 6 kíp 4:**
- ☞ **Nguyentinh.ptit.94@gmail.com 01644626065.**

- ☞ **Lớp của Nam:**
- ☞ **Vũ Thanh Đạt - vuthanhdat.toank19@gmail.com - 01656693843**
- Nguyễn Thùy Dương – yeunuthongminh@gmail.com- 01644184933**
- Lớp CĐ:**
- Nhóm 7:**
- Nguyễn Đăng Hiếu- ngdanghieu2511@gmail.com- 01695102642**
- Nhóm 8:**
- Vũ Văn Hợp-hopvv1994@gmail.com-01678166116**

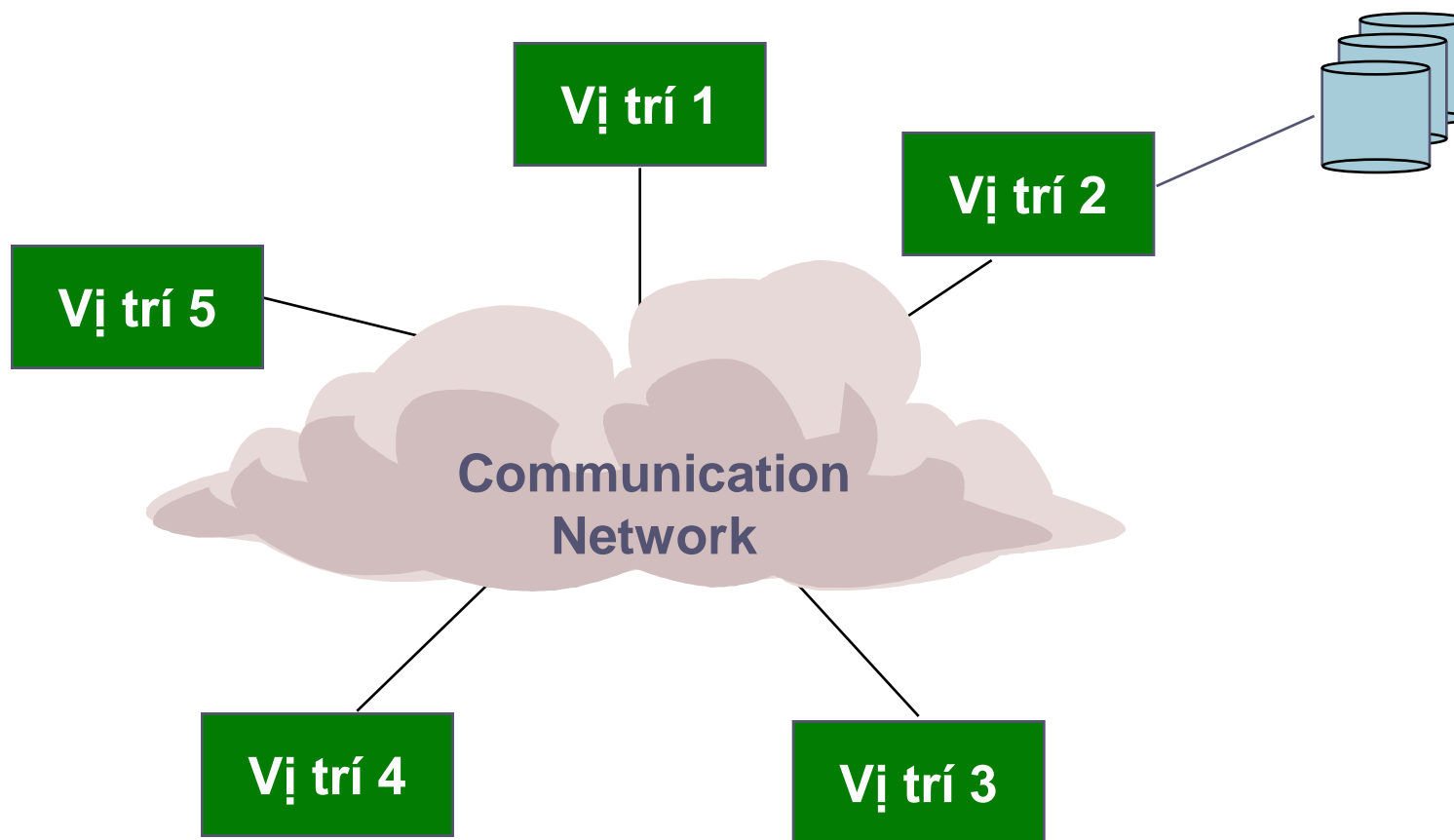
Nội dung

- ☐ 1. Giới thiệu
- ☐ 2. Môi trường phân tán
- ☐ 3. Hệ thống phân tán và Xử lý dữ liệu phân tán.
- ☐ 4. Hệ cơ sở dữ liệu phân tán là gì.
- ☐ 5. Các đặc điểm của cơ sở dữ liệu phân tán
- ☐ 6. Các mô hình CSDL
- ☐ 7. *Mô hình tham chiếu (Mô hình kiến trúc) cơ sở dữ liệu phân tán*
- ☐ 8. Hệ quản trị cơ sở dữ liệu phân tán.

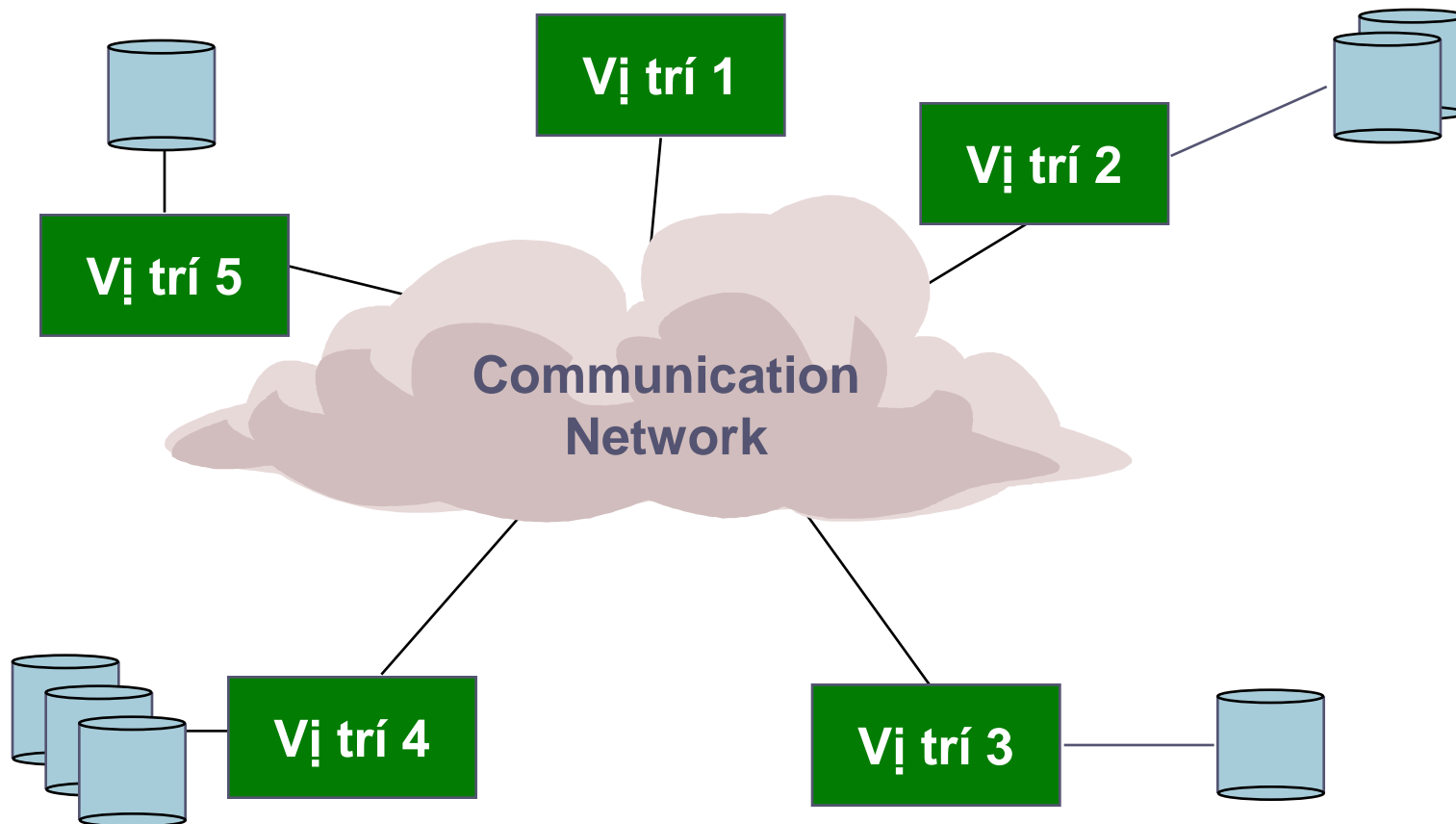
1. Giới thiệu

- ***Tại sao phải phân tán? - Phân tán cái gì?***
- Các tổ chức kinh tế có nhiều trụ sở phân tán ở nhiều vị trí địa lý khác nhau
- Làm thế nào để có thể quản lý các luồng dữ liệu và sử dụng chung dữ liệu, chương trình của các trụ sở trên những vị trí khác nhau đó
- => Nhu cầu dùng chung dữ liệu hợp nhất trên nhiều vị trí địa lý khác nhau đó chính là nhu cầu về hệ thống phân tán

2. Môi trường cơ sở dữ liệu tập trung



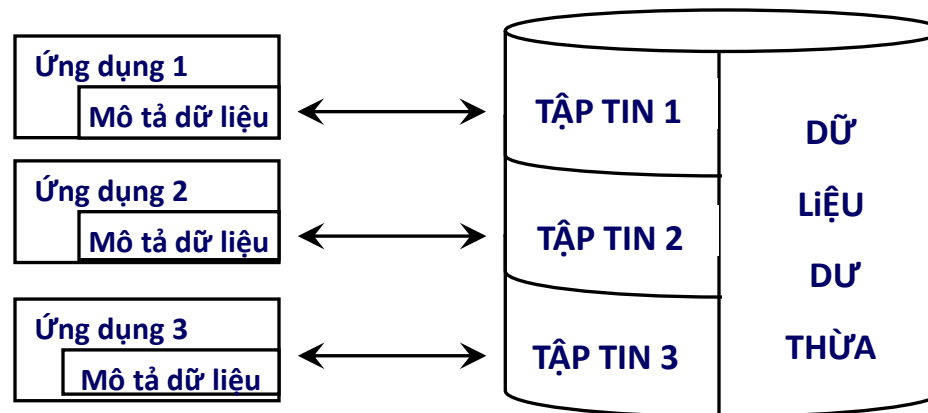
Môi trường cơ sở dữ liệu phân tán



3. Hệ thống phân tán và Xử lý dữ liệu phân tán.

❑ *Xử lý dữ liệu truyền thống*

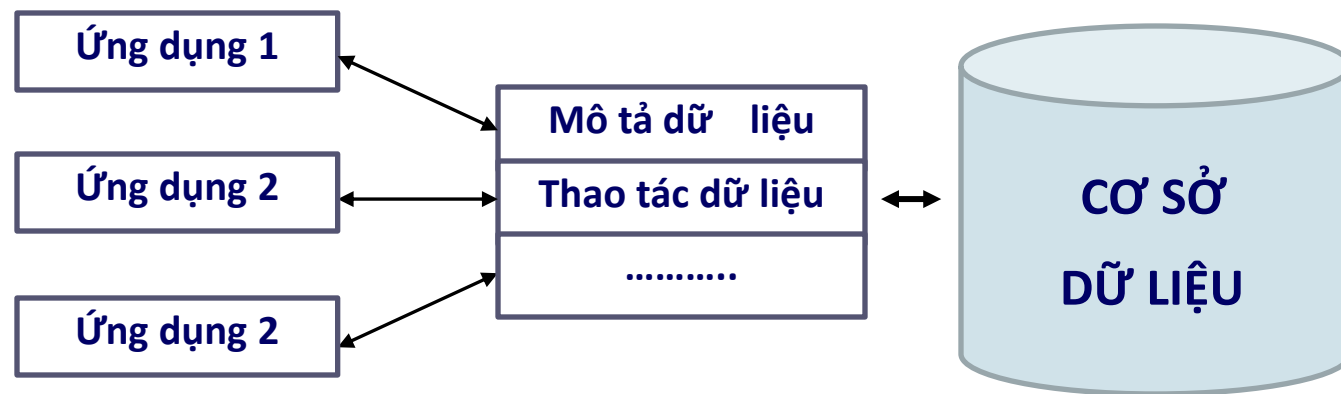
- Mỗi ứng dụng định nghĩa một (nhiều) tệp dữ liệu.
- Ứng dụng khác nhau, tổ chức dữ liệu khác nhau
- Ứng dụng và tổ chức dữ liệu phụ thuộc lẫn nhau.
- Dư thừa dữ liệu



Hình 2.2: Xử lý dữ liệu truyền thống

❑ *Xử lý dữ liệu tập trung:*

- Độc lập dữ liệu
- Tính bất biến ứng dụng đối với sự thay đổi cấu trúc lưu trữ và các chiến lược truy nhập dữ liệu.
- Không dư thừa dữ liệu



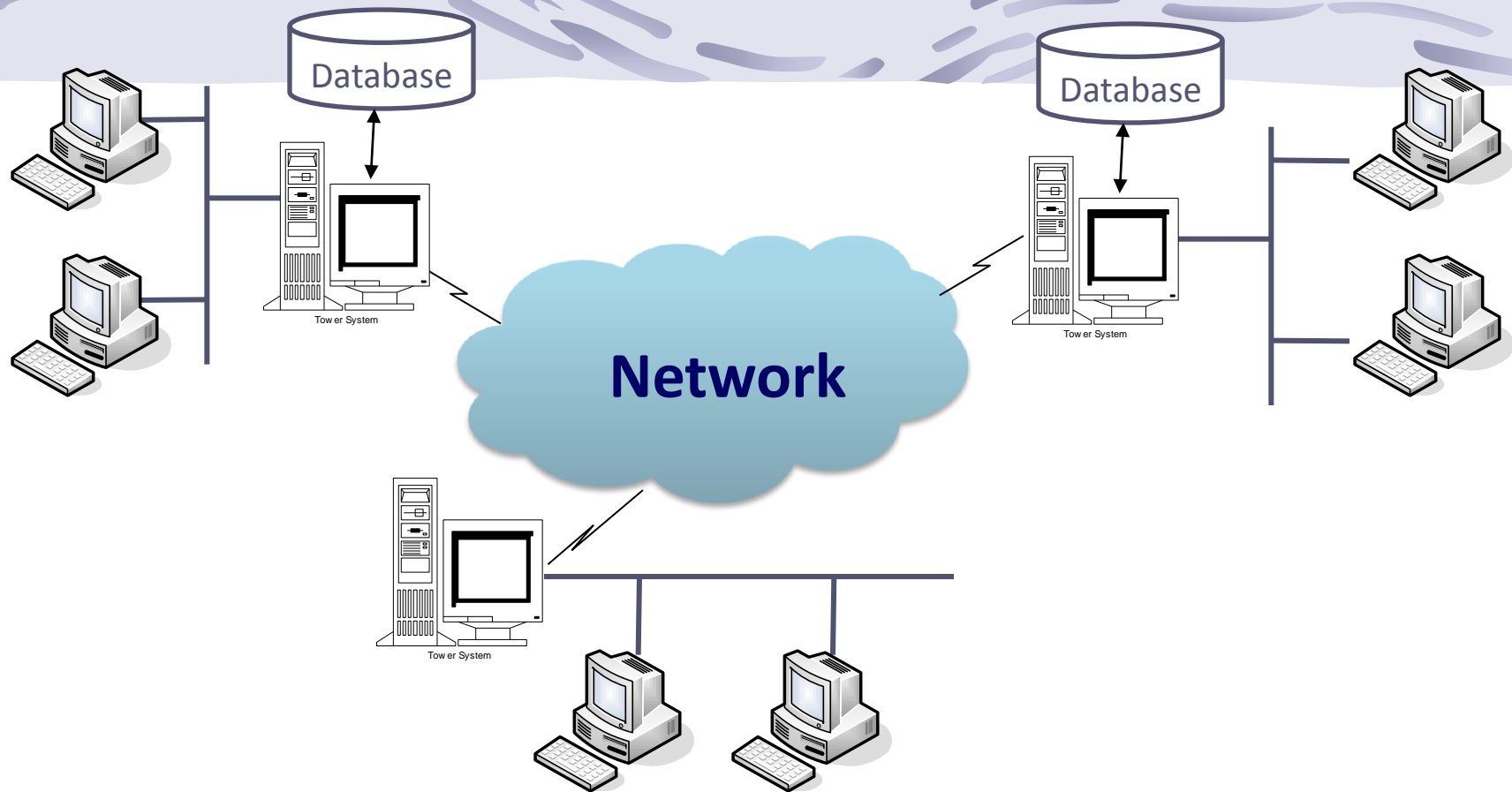
Hình 2.3. Xử lý cơ sở dữ liệu

❑ *Khái niệm xử lý phân tán*

- Hệ xử lý phân tán là một tập các phần tử xử lý tự trị được kết nối với nhau bởi một mạng máy tính và thực hiện những công việc theo chức năng.

❑ Có hai khái niệm xử lý phân tán.

- Mô hình tính toán Client/Server. Client truy nhập trực tiếp và xử lý dữ liệu trên Server.
- Một giao dịch được chia ra xử lý trên nhiều hệ thống. Máy trung tâm sẽ giám sát và quản lý các tiến trình.



Hình 2.1 Hệ thống xử lý phân tán

❑ *Hệ thống xử lý phân tán*

Mạng máy tính được cài đặt các cơ sở dữ liệu và các phần mềm hệ thống, tạo khả năng cho nhiều người sử dụng truy nhập chia sẻ nguồn thông tin chung.

=> Hệ thống xử lý phân tán đáp ứng nhu cầu thu thập, lưu trữ, xử lý và trao đổi thông tin của các tổ chức kinh tế, xã hội có các hoạt động trên phạm vi rộng lớn. Yêu cầu có độ tin cậy cao, đáp ứng nhanh. Hệ cơ sở dữ liệu phân tán là nhân tố quan trọng, làm cho quá trình xử lý phân tán dễ dàng và có hiệu quả.

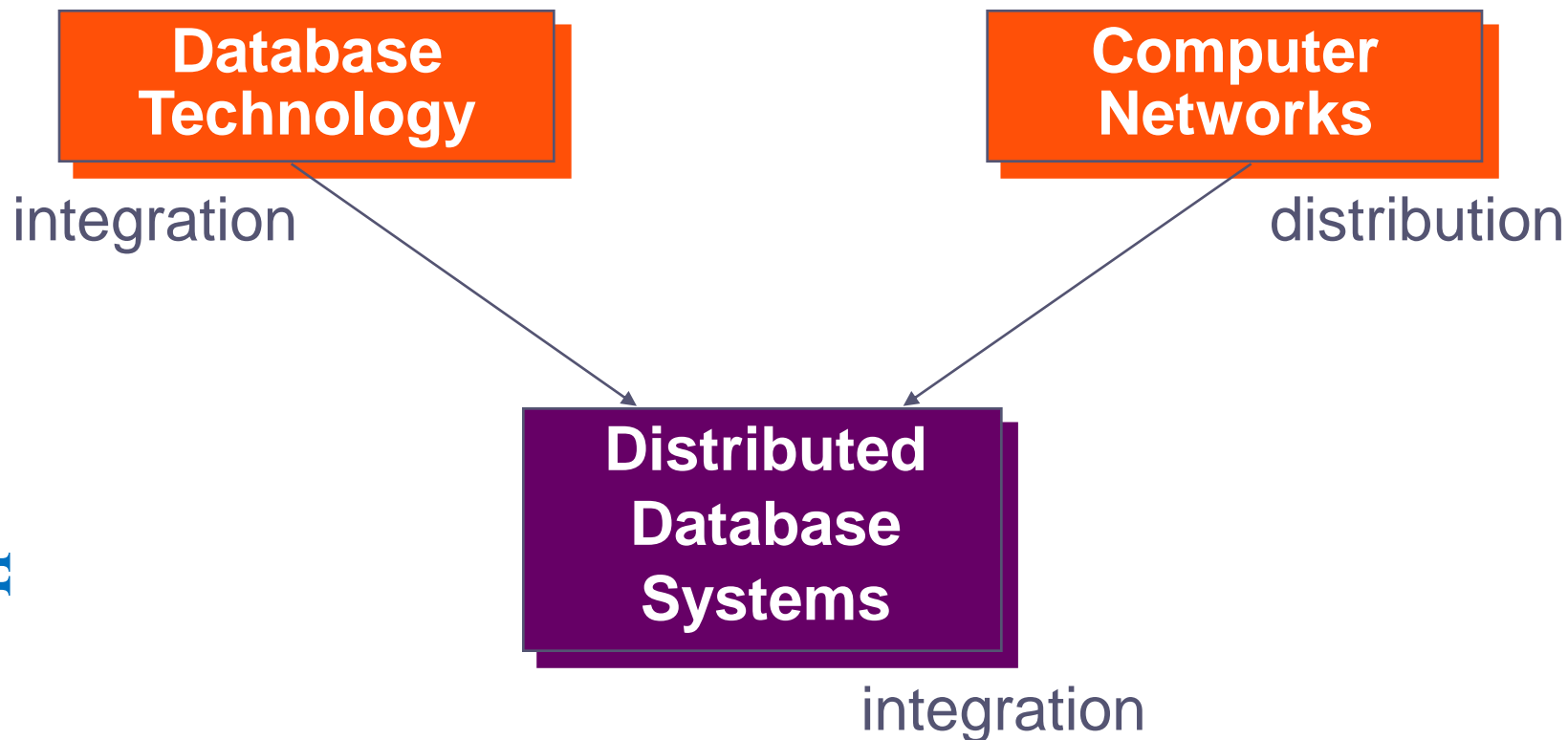
❑ *Các đối tượng phân tán bao gồm*

- Phân tán chức năng (Logic xử lý phân tán): Các chức năng xử lý đc phân tán.
- Phân tán dữ liệu: Dữ liệu được phân tán trên nhiều vị trí khác nhau, nhưng ràng buộc và liên quan đến nhau
- Điều khiển phân tán: Điều khiển các giao dịch khác nhau được phân tán trên nhiều vị

trí

4. Cơ sở dữ liệu phân tán

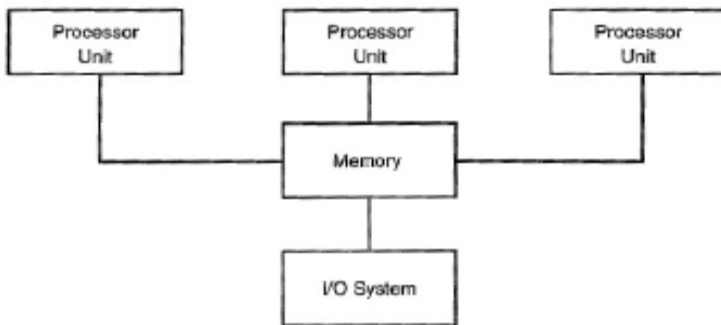
- ❑ Là một tập các cơ sở dữ liệu có quan hệ với nhau về mặt logic và được phân tán trên một mạng máy tính.
- ❑ Hệ QTCSDL phân tán là hệ thống phần mềm quản trị CSDLPT và làm cho sự phân tán trong suốt với NSD.
- ❑ Khái niệm CSDLPT nhấn mạnh hai khía cạnh:
 - *Tính phân tán*: dữ liệu phân tán trên nhiều vị trí. cơ sở dữ liệu địa phương (Local)
 - *Mối quan hệ logic*: Dữ liệu trên các vị trí phụ thuộc, ràng buộc lẫn nhau.



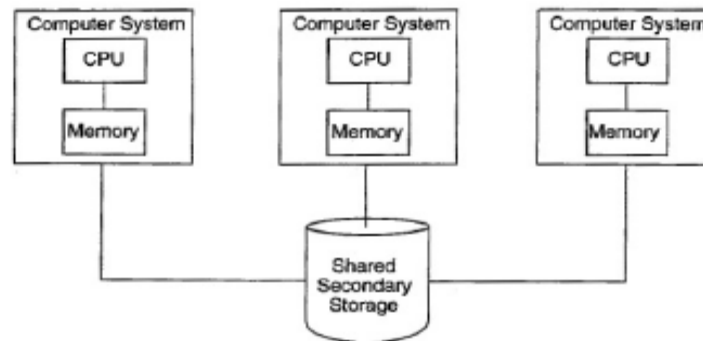
Sự cần thiết của CSDLPT

- ✎ Sự phát triển của cơ cấu tổ chức
- ✎ Giảm chi phí truyền thông
- ✎ Hiệu quả công việc (xử lý song song)
- ✎ Độ tin cậy sẵn sàng

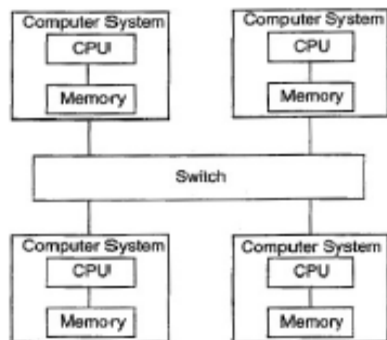
Các hệ thống không phải là DDBS



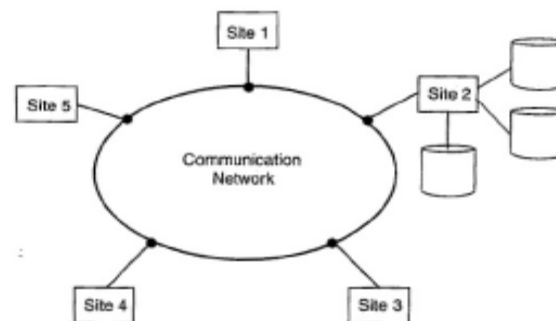
Shared Memory



Shared Disk



Shared Nothing



Central Databases

5. Đặc điểm cơ sở dữ liệu phân tán :Điều khiển tập trung

□ Điều khiển tập trung

- CSDL tập trung: đảm bảo tính độc lập của dữ liệu. tránh tổn thất thông tin. Sự điều khiển được thực hiện bởi Người QTCSDL
- CSDL phân tán: không nhấn mạnh điều khiển tập trung. Sự điều khiển được thực hiện bởi Người CSDL toàn cục (Global Database Administrator) và Người CSDL địa phương-cục bộ (Local Database Administrator). Người CSDL địa phương có quyền độc lập vị trí.

Đặc điểm cơ sở dữ liệu phân tán: Độc lập dữ liệu

□ ***Độc lập dữ liệu***

- Tập trung: Độc lập dữ liệu : Các Ctud không bị thay đổi khi thay đổi tổ chức lưu trữ vật lý của dl. Thể hiện thông qua kiến trúc nhiều mức, các mức này có những mô tả các nhau về dl và những ax biến đổi giữa các mức
- Trong phân tán: (Cũng quan trọng), Là sự trong suốt phân tán (Distribution Transparency), nghĩa là tính đúng đắn của các trình ứng dụng không bị ảnh hưởng bởi sự di chuyển dữ liệu từ một vị trí này đến một vị trí khác.

=> Tuy nhiên, tốc độ thực hiện của các trình ứng dụng bị ảnh hưởng khi phải trao đổi dữ liệu giữa các vị trí.

Đặc điểm CSDL phân tán: Giảm dư thừa dữ liệu

□ Giảm dư thừa dữ liệu

-Trong CSDL tập trung, sự dư thừa dữ liệu được giảm thiểu, tránh sự không nhất quán dữ liệu.

-Trong CSDL phân tán, sự dư thừa theo ngữ nghĩa sau:

- Tăng tính cục bộ: dữ liệu có thể được nhân bản tại tất cả các vị trí=>ứng dụng cục bộ thực hiện nhanh hơn.
- Tăng tính sẵn sàng: Dữ liệu được nhân bản tại vt nào đó, Khi một vị trí có sự cố thì sẽ không làm ngưng sự thực hiện của các ứng dụng ở những vị trí khác .

❑ *Khi nào sử dụng “dự thừa dữ liệu” nhân bản có nhiều ưu điểm hơn???*

- Sự nhân bản dữ liệu dựa vào hai loại ứng dụng cơ bản: Ứng dụng chỉ đọc và ứng dụng cập nhật.

- Sự nhân bản giúp cho các ứng dụng chỉ đọc được thực hiện nhanh hơn, nhưng các ứng dụng cập nhật thực hiện chậm hơn vì phải cập nhật tại các vị trí

=> Như vậy, sự nhân bản dữ liệu sẽ là một ưu điểm nếu hệ thống có rất nhiều ứng dụng chỉ đọc và có rất ít ứng dụng cập nhật. Trong trường hợp ngược lại thì sự nhân bản

Đặc điểm CSDL phân tán: Cải tiến hiệu năng

❑ *Cải tiến hiệu năng*

- *Hiệu năng của CSDL PT được cải tiến dựa vào hai điểm:*

1. Khả năng phân mảnh CSDL khái niệm và cục bộ hoá dữ liệu. Có ưu điểm:

- *Vì mỗi vị trí chỉ xử lý một phần cơ sở dữ liệu, sự tranh chấp CPU và các dịch vụ vào/ra không nghiêm trọng như trong môi trường tập trung.*
- *Tính cục bộ làm giảm thời gian trễ truy nhập từ xa.*

Lợi ích của việc phân mảnh và phân tán dữ liệu hợp lý sẽ làm giảm tranh chấp và giảm chi phí truyền thông.

Đặc điểm CSDL phân tán: Cải tiến hiệu năng

2. *Tính song song của hệ thống phân tán* có thể được khai thác để thực hiện

- Song song liên truy vấn: là khả năng thực hiện nhiều truy vấn tại cùng thời điểm

Nội truy vấn song song là phương pháp tách một truy vấn đơn thành các truy vấn con và mỗi truy vấn con được thực hiện tại các trạm khác nhau, truy nhập các phần khác nhau của cơ sở dữ liệu phân tán.

Đặc điểm CSDL phân tán: Cải tiến hiệu năng

- Phân mảnh và nhân bản dữ liệu sao cho khai thác tốt nhất tính song song.
- Trường hợp rất nhiều NSD cùng truy nhập vào một quan hệ chỉ được đọc, một số site yêu cầu xử lý song song, thì càng nhiều bản sao càng tăng tính sẵn sàng.
- Tăng bản sao làm giảm khối lượng thông tin trao đổi giữa các site. Tuy nhiên tính phức tạp tăng khi thực hiện cập nhật dữ liệu, phải đảm bảo tất cả các bản sao trên mạng phải thỏa các điều kiện toàn vẹn dữ liệu.

Đặc điểm CSDL phân tán: Cải tiến hiệu năng

Tóm lại, việc nhân bản dữ liệu sẽ làm tăng hiệu quả các phép đọc, tăng tính sẵn sàng trong các giao tác đọc dữ liệu. Vấn đề điều khiển các giao tác tương tranh có sử dụng kỹ thuật bản sao trong các cơ sở dữ liệu phân tán rất phức tạp. Việc quản lý có thể đơn giản hóa bằng việc sinh ra các bản sao sơ cấp quan hệ.

Đặc điểm : Độ tin cậy giao dịch phân tán

- ❑ Độ tin cậy giao dịch phân tán được cải thiện, vì nhân bản hạn chế được các lỗi trên vị trí riêng lẻ.
- ❑ Lỗi của cục bộ, hoặc lỗi truyền thông, không ảnh hưởng đến hoạt động bộ hệ thống.
- ❑ Nếu một số dữ liệu không thể truy nhập được, các giao dịch phân tán vẫn có thể truy nhập được tới phần khác trong cơ sở dữ liệu phân tán.

Đặc điểm : Độ tin cậy giao dịch phân tán

- ❑ Giao dịch là một đơn vị tính toán cơ bản, nhất quán và tin cậy, bao gồm một chuỗi các thao tác cơ sở dữ liệu được thực hiện từ trạng thái CSDL nhất quán này sang trạng thái nhất quán khác, ngay cả khi có một số giao dịch được thực hiện đồng thời và khi xảy ra lỗi.
- ❑ Khi hệ thống có lỗi, hệ QTCSDL đảm bảo việc thực thi đồng thời các giao dịch, không vi phạm tính nhất quán, với điều kiện là giao dịch được thực hiện chính xác, nghĩa là tuân theo các qui tắc toàn vẹn của CSDL.

Đặc điểm CSDL phân tán: Dễ dàng mở rộng

Dễ dàng mở rộng

- ❑ Dễ dàng tăng kích thước dữ liệu, bằng cách tăng khả năng lưu trữ và xử lý của mạng. Phụ thuộc vào chi phí phân tán. Tuy nhiên, khả năng mở rộng hệ thống dễ dàng mang tính kinh tế, chi phí giảm.

Đặc điểm CSDL phân tán:

Tính toàn vẹn, phục hồi và điều khiển tương tranh

□ Tính toàn vẹn, phục hồi và điều khiển tương tranh

- Tính toàn vẹn, phục hồi và điều khiển tương tranh có quan hệ chặt chẽ với nhau thông qua các giao tác.

- Mỗi một giao tác là một nguyên tố, nghĩa là một dãy các thao tác hoặc được thực hiện hoàn toàn, hoặc không được thực hiện hay chỉ thực hiện một phần.

- Như vậy giao tác nguyên tố là phương tiện để đạt được tính toàn vẹn dữ liệu, vì các giao tác đảm bảo tất cả các thao tác biến đổi cơ sở dữ liệu từ trạng thái toàn vẹn này sang trạng thái toàn vẹn khác.

Đặc điểm CSDL phân tán:

Tính toàn vẹn, phục hồi và điều khiển tương tranh

- Tác động sự cố kỹ thuật và tương tranh có ảnh hưởng đến tính nguyên tố của giao tác.
- Sự cố kỹ thuật làm cho hệ thống dừng giao tác đang thực hiện, vì vậy vi phạm phải mang tính nguyên tố.
- Sự tương tranh của các giao tác cho phép một giao tác nào đó quan sát một trạng thái chuyển tiếp không toàn vẹn, tạo ra bởi một giao tác khác trong quá trình thực hiện nó.

Đặc điểm CSDL phân tán:

Tính toàn vẹn, phục hồi và điều khiển tương tranh

- ❑ Sự phục hồi liên quan đến đảm bảo tính nguyên tố của giao tác khi có sự cố về kỹ thuật, khi có một số vị trí (site) tham gia thực hiện giao tác có thể bị hỏng.
- ❑ Điều khiển tương tranh nhằm đảm bảo tính nguyên tố của giao tác khi xuất hiện sự tương tranh giao tác. Vấn đề điều khiển tương tranh trong môi trường phân tán phức tạp hơn nhiều trong môi trường tập trung.

Đặc điểm CSDL phân tán:

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:
 - a. Trong suốt phân đoạn (fragmentation transparency)
 - b. Trong suốt về vị trí (location transparency)
 - c. Trong suốt ánh xạ địa phương (local mapping transparency)

6. Các mô hình CSDL

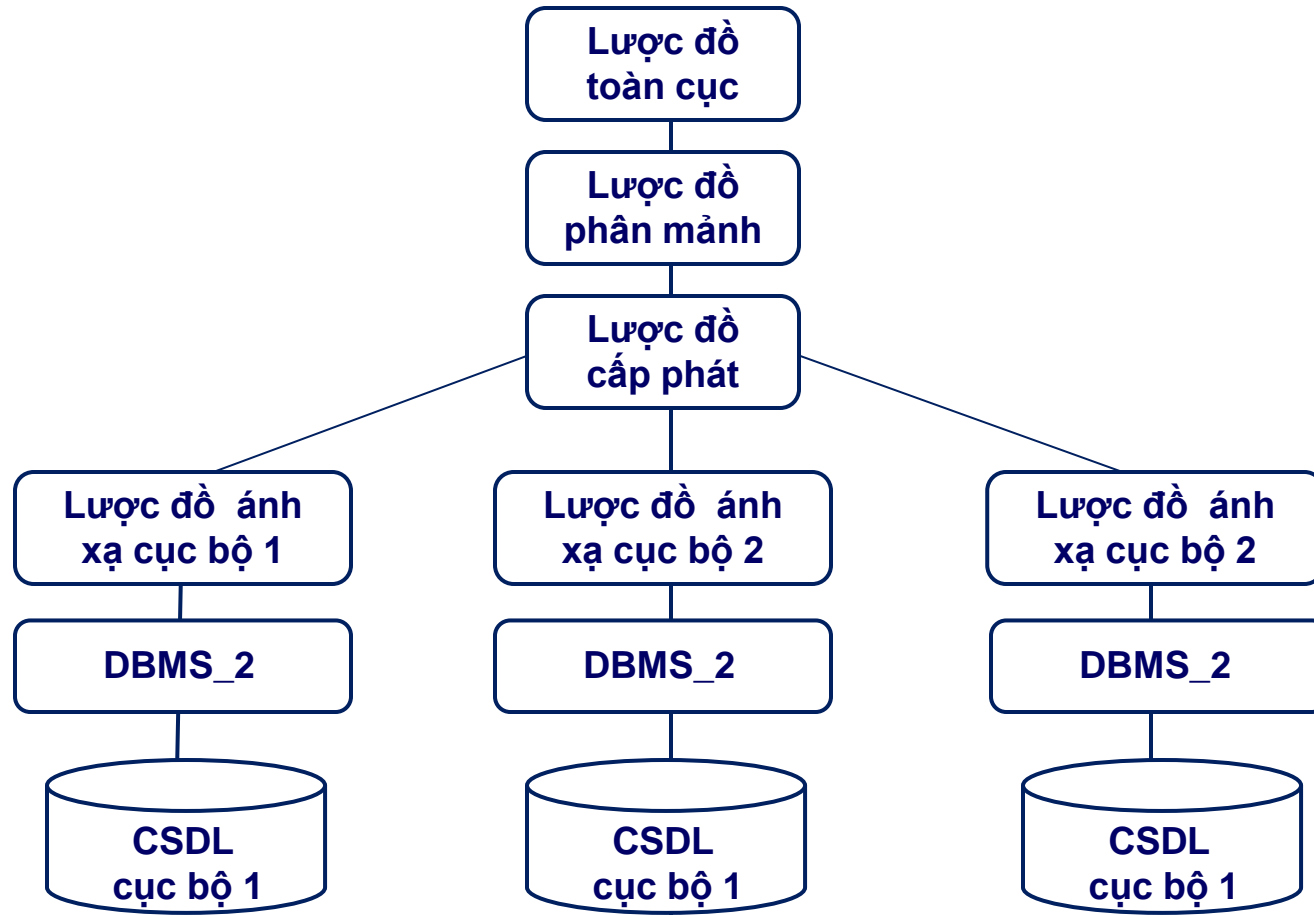
- ☛ MH CSDL tập trung: UwD+HQTCSDL tập trung trên 1 máy (bộ vi xử lý)
- ☛ MH File-Server
 - Máy cá nhân: PMUD+phần mềm csdl
 - Máy Server: File DL vật lý
- ☛ MH Client-Server
 - Máy cá nhân: PMUD+phần mềm csdl
 - Máy Server: CSDL+phần mềm csdl
- ☛ MH Phân tán: CSDL nằm trên nhiều máy khác nhau

7. Mô hình kiến trúc của cơ sở dữ liệu phân tán

(Mô hình tham chiếu)

- ❑ Mô hình kiến trúc cơ sở dữ liệu phân tán dựa trên sự mở rộng của mô hình kiến trúc ba mức của cơ sở dữ liệu tập trung ANSI/SPARC.
- ❑ Mô hình kiến trúc cơ sở dữ liệu phân tán bao gồm lược đồ tổng thể, lược đồ phân mảnh và lược đồ cấp phát.

Mô hình tham chiếu cơ sở dữ liệu phân tán



Hình 2.4 Mô hình tham chiếu của cơ sở dữ liệu phân tán

Lược đồ toàn cục (Mô hình tham chiếu toàn cục)

- ❑ Lược đồ toàn cục định nghĩa tất cả dữ liệu được chứa trong CSDL phân tán như trong CSDL tập trung.
- ❑ Lược đồ toàn cục được định nghĩa chính xác như định nghĩa lược đồ cơ sở dữ liệu tập trung.
- ❑ Tuy nhiên, mô hình dữ liệu lược đồ toàn cục cần phải tương thích với việc định nghĩa các ánh xạ tới các mức của cơ sở dữ liệu phân tán.
- ❑ Sơ đồ toàn cục bao gồm thông tin về các thực thể, tập các phụ thuộc hàm và mối quan hệ giữa các thực thể .

Lược đồ phân mảnh (Mô hình tham chiếu phân mảnh)

- ❑ Các mảnh: Quan hệ toàn cục có thể chia thành nhiều quan hệ con được gọi là các mảnh, tách biệt với nhau
- ❑ Lược đồ phân mảnh: Là ánh xạ giữa quan hệ toàn cục và các phân mảnh. Ánh xạ này là mỗi quan hệ một-nhiều.
 - Các mảnh được chỉ ra bằng tên của quan hệ toàn cục với một chỉ số (chỉ số phân mảnh)
 - Các mảnh có thể được cài đặt tại một hay nhiều vị trí khác nhau trên mạng.

Lược đồ phân mảnh

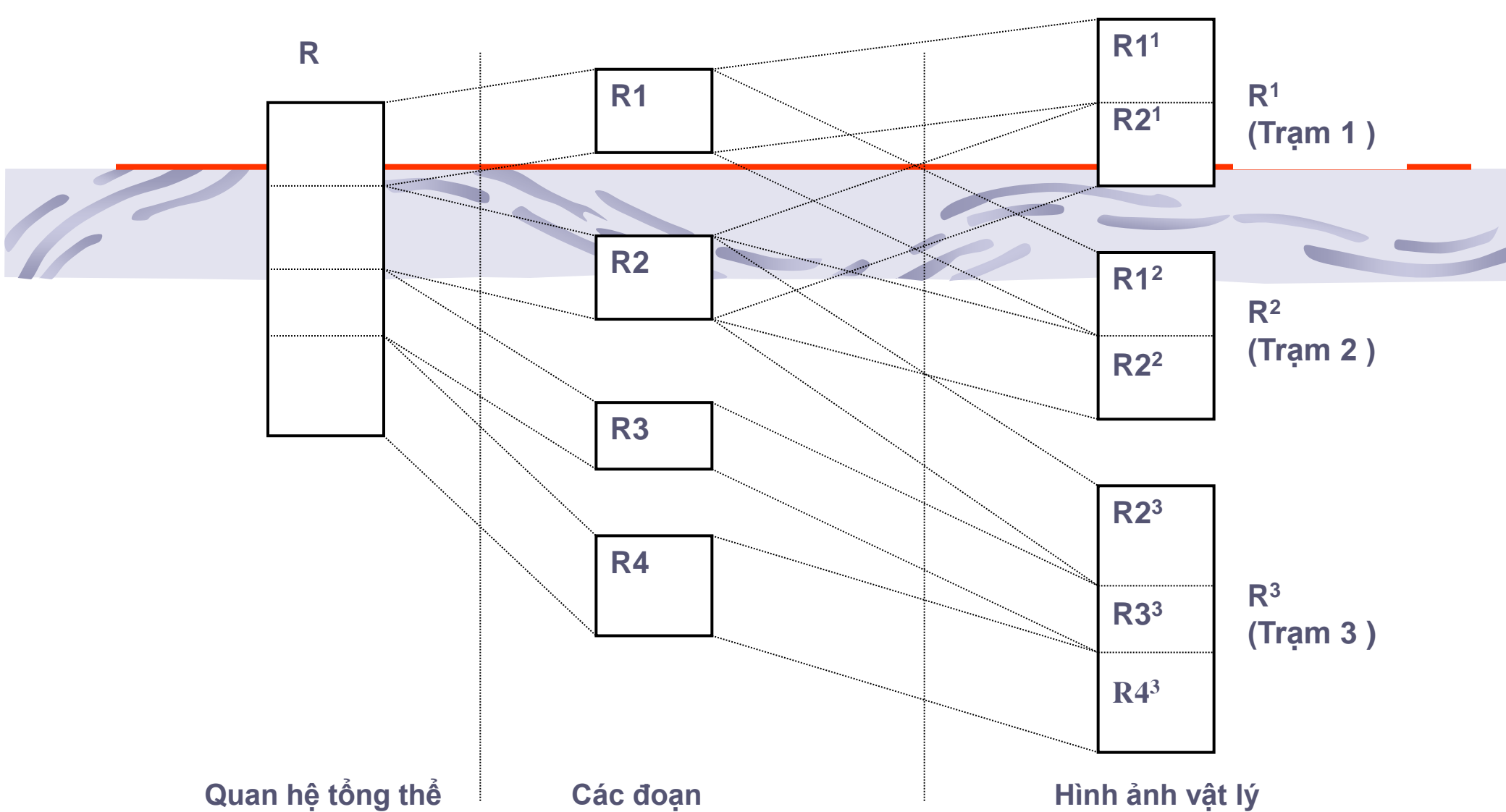
- ❑ Các kiểu phân mảnh dữ liệu bao gồm:
 - Phân mảnh ngang
 - Phân mảnh dọc
 - Sự kết hợp của phân mảnh ngang và phân mảnh dọc, phức tạp hơn
- ❑ Một mảnh được định nghĩa bằng một biểu thức đại số quan hệ, các toán hạng là các quan hệ toàn cục.

Lược đồ cấp phát

- ❑ Lược đồ cấp phát: chứa thông tin các mảnh được chứa trên những site nào (Các mảnh được chứa ở một hay nhiều site trên mạng).
- ❑ Các mảnh là các phần logic của một quan hệ tổng thể được định vị vật lý trên một hay nhiều trạm.
- ❑ Tất cả các mảnh 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, site). Các ảnh vật lý có thể chỉ ra bằng tên của một quan hệ tổng thể và một chỉ mục trạm.

Lược đồ cấp phát

- ❑ Ký hiệu R_{ij} là mảnh thứ i của quan hệ R trên site j .
- ❑ Quan hệ toàn cục R được phân thành 4 mảnh quan hệ con R_1, R_2, R_3 và R_4 , và được cấp phát trên 3 site của mạng máy tính, tạo nên ba mô hình vật lý:
 - Trên site 1: Một bản sao của R_1 và một bản sao của R_2 . Ký hiệu R_{11} và R_{21}
 - Trên site 2: Một bản sao của R_1 và một bản sao của R_2 . Ký hiệu R_{12} và R_{22}
 - Trên site 3: Một bản sao của R_2, R_3 và R_4 . Ký hiệu R_{23}, R_{33} và R_{43}



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

Lược đồ ánh xạ cục bộ

- Thực hiện ánh xạ các ảnh vật lý thành các đối tượng được thao tác bởi hệ quản trị cơ sở dữ liệu địa phương (Tất cả các mảnh của một quan hệ tổng thể trên cùng một trạm tạo ra một ảnh vật lý)
- Ánh xạ này được gọi là ánh xạ cục bộ, nó phụ thuộc vào kiểu của DBMS.
- (Trong ht đồng nhất lđ độc lập của mỗi site đc đn sử dụng cùng 1 mô hình như DBMS cục bộ, Trong hệ thống không đồng nhất tại mỗi site thì ánh xạ cục bộ khác nhau, khi đó lược đồ axcb còn dùng để phối hợp các kiểu khác nhau của các DBMS không đồng nhất)

Tính trong suốt trong Phân mảnh ,cấp phát dữ liệu, và ánh xạ địa phương

- ❑ Trong suốt trong phân mảnh: là mức trong suốt cao nhất, NSD chỉ cần làm việc trên mối quan hệ tổng thể (Quan hệ tổng).

Khi dữ liệu đc phân mảnh thì việc truy cập vào csdl bình thường như chưa bị phân tán và không bị ảnh hưởng

Ví dụ 1.1. Xét quan hệ tổng thể DAI_LY (ID, Họ tên, Tài khoản, Điện thoại, Địa chỉ) và các phân đoạn được tách ra từ nó:

DAI_LY 1 (ID, Họ tên, Tài khoản, Điện thoại, Địa chỉ)

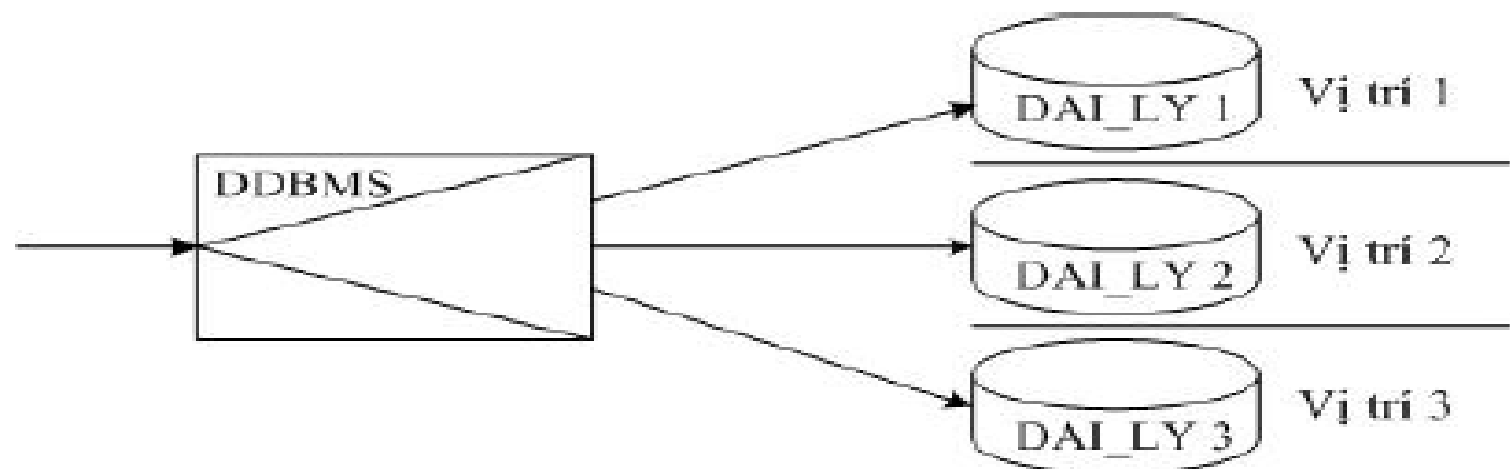
DAI_LY 2 (ID, Họ tên, Tài khoản, Điện thoại, Địa chỉ)

DAI_LY 3 (ID, Họ tên, Tài khoản, Điện thoại, Địa chỉ)

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 DAI_LY mà không cần biết quan hệ DAI_LY có phân tán hay không.

```
SELECT *  
FROM   DAI_LY  
WHERE  ID= "Id1"
```



Hình 1.10. Trong suốt phân đoạn

- ❑ Trong suốt trong cấp phát: là mức trong suốt thấp hơn và yêu cầu người sử dụng chỉ được làm việc trên các mảnh địa phương (thay cho làm việc trên quan hệ toàn cục) mà không biết mảnh đó ở đâu.

Người sử dụng không cần biết về vị trí vật lý của dl mà có quyền truy nhập tới csdl tại bất cứ vị trí nào

Ví dụ 1.2. 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= “Id1”

```
SELECT *  
FROM   DAI_LY 1  
WHERE  Id= “Id1”  
  
      IF NOT # FOUND THEN  
SELECT *  
FROM   DAI_LY 2  
WHERE  Id= “Id1”
```

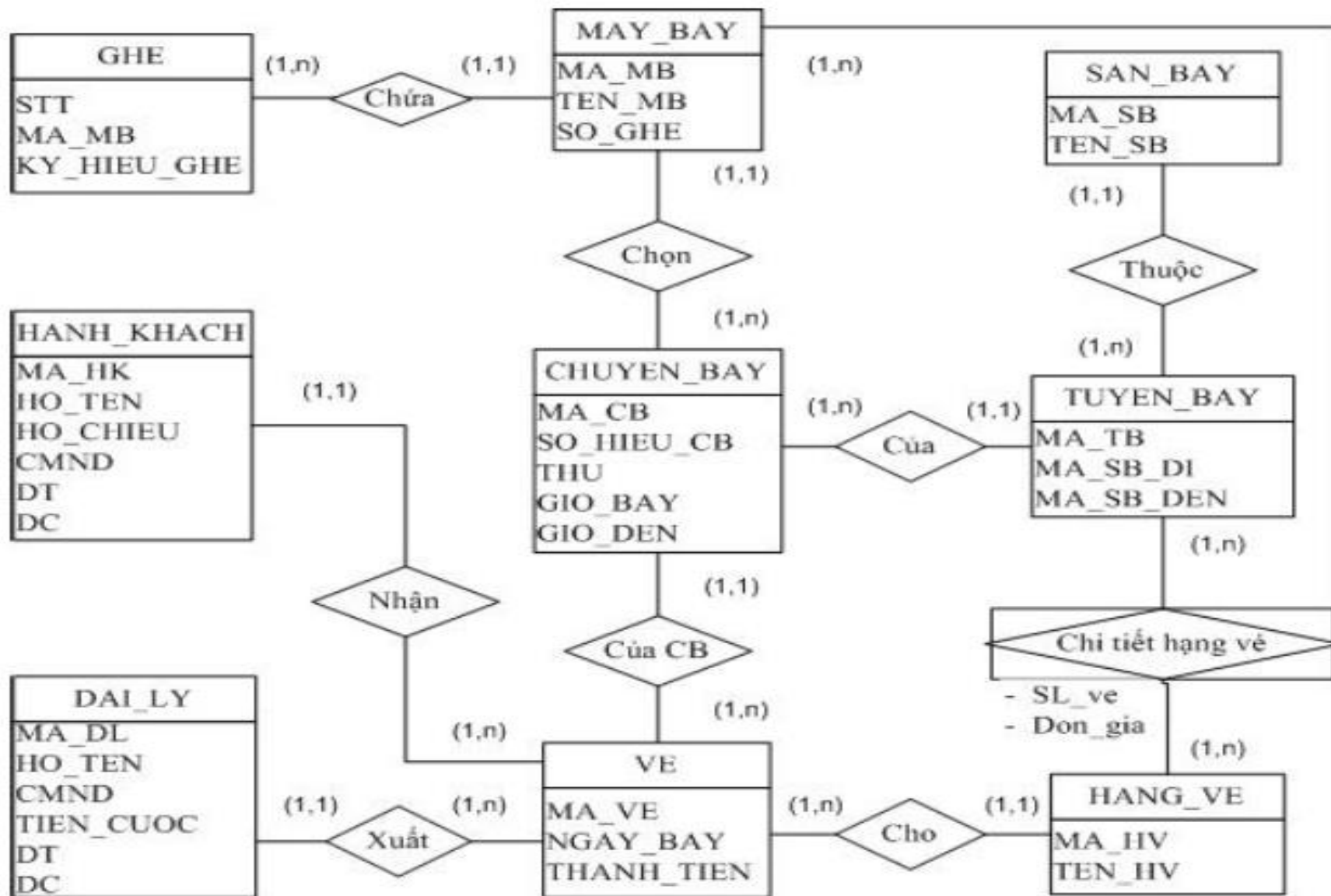
Đầu tiên hệ thống sẽ thực hiện tìm kiếm ở phân đoạn DAI_LY 1 và nếu DDBMS trả về biến điều khiển # FOUND thì một câu lệnh truy vấn tương tự được thực hiện trên phân đoạn DAI_LY 2 ,...

Ở đây quan hệ DAI_LY 2 được sao làm hai bản trên hai vị trí 2 và vị trí 3, ta chỉ cần tìm thông tin trên quan hệ DAI_LY 2 mà không cần quan tâm nó ở vị trí nào.

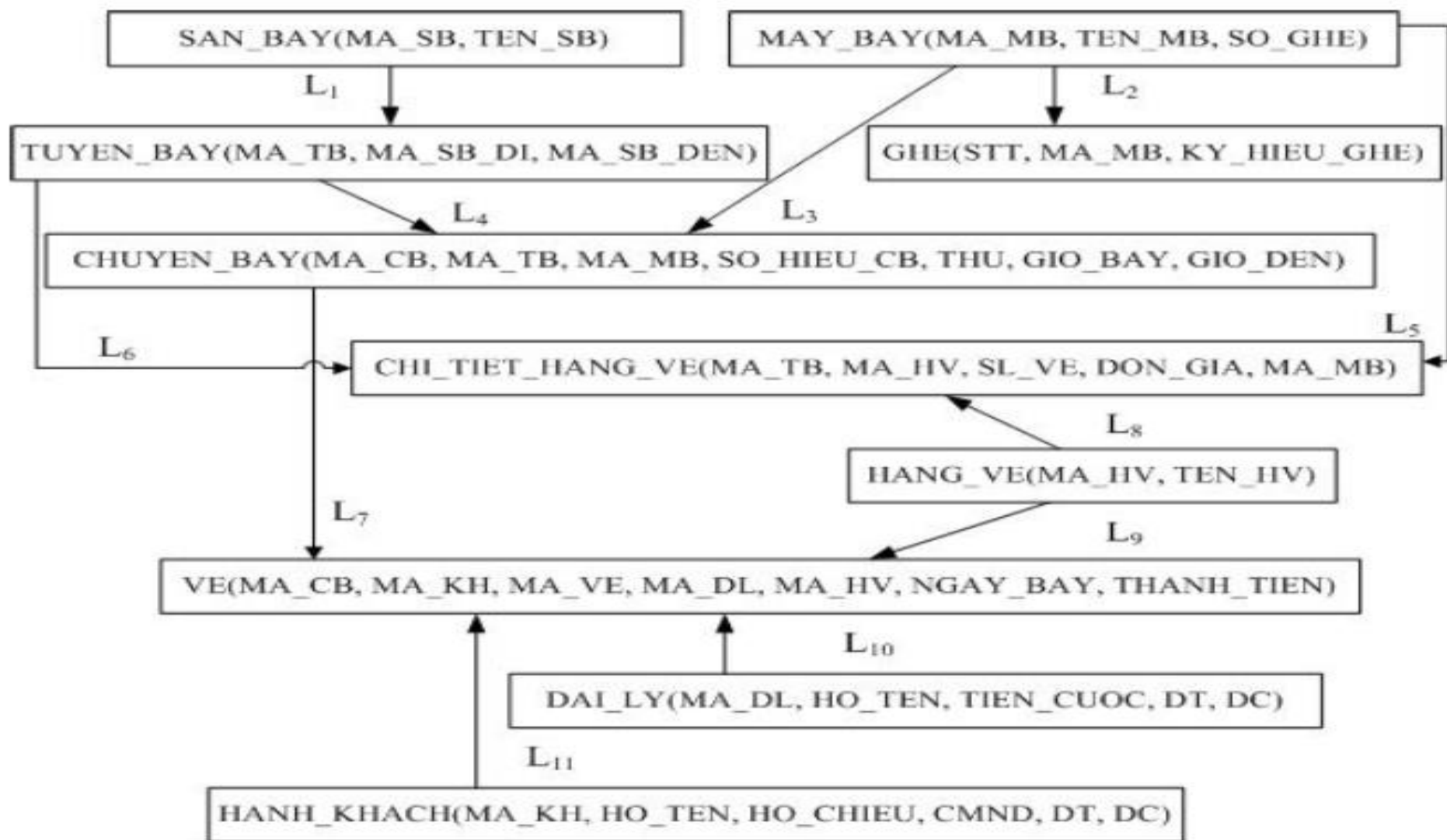
VD về hệ thống bán vé máy bay

- ✍ -Lược đồ tổng thể
- ✍ - Lược đồ phân mảnh
- ✍ - Lược đồ định vị
- ✍ - Lược đồ ánh xạ

Lược đồ tổng thể



Hình 3.1: Mô hình thực thể quan hệ
 Học Viện Công nghệ BGVT, Phan Thị
 Hà, hathiphan@yahoo.com



Hình 3.2. Sơ đồ tổng thể của hệ thống

Phân mảnh (tiếp vd)

Vị trí 1: dữ liệu đặt tại miền bắc (Hà Nội)

Vị trí 2: dữ liệu đặt tại miền trung (Đà Nẵng)

Vị trí 3: dữ liệu đặt tại miền nam (Sài Gòn)

Phân rã quan hệ **SAN_BAY** thành các mảnh ngang SAN_BAY 1, SAN_BAY 2, SAN_BAY 3 được định nghĩa như sau:

$SAN_BAY\ 1 = \sigma_{MA_SB} = \text{'Các mã sân bay mà vị trí 1 quản lý'} (SAN_BAY)$

$SAN_BAY\ 2 = \sigma_{MA_SB} = \text{'Các mã sân bay mà vị trí 2 quản lý'} (SAN_BAY)$

$SAN_BAY\ 3 = \sigma_{MA_SB} = \text{'Các mã sân bay mà vị trí 3 quản lý'} (SAN_BAY)$

Phân rã quan hệ **TUYEN_BAY** thành các mảnh ngang TUYEN_BAY 1, TUYEN_BAY 2, TUYEN_BAY 3 được định nghĩa như sau:

$TUYEN_BAY\ 1 = TUYEN_BAY \bowtie SAN_BAY\ 1$

$TUYEN_BAY\ 2 = TUYEN_BAY \bowtie SAN_BAY\ 2$

$$\text{TUYEN_BAY 3} = \text{TUYEN_BAY} \triangleright \triangleleft \text{SAN_BAY 3}$$

trong đó, thuộc tính mã sân bay đi của quan hệ TUYEN_BAY là trường để nối tự nhiên với mã sân bay của quan hệ SAN_BAY

còn mã sân bay đến của quan hệ TUYEN_BAY sẽ được chọn giá trị từ các mảnh SAN_BAY 1, SAN_BAY 2, SAN_BAY 3

Phân rã quan hệ **CHUYEN_BAY** thành các mảnh ngang CHUYEN_BAY 1, CHUYEN_BAY 2, CHUYEN_BAY 3 được định nghĩa như sau:

$$\text{CHUYEN_BAY 1} = \text{CHUYEN_BAY} \triangleright \triangleleft \text{TUYEN_BAY 1}$$

$$\text{CHUYEN_BAY 2} = \text{CHUYEN_BAY} \triangleright \triangleleft \text{TUYEN_BAY 2}$$

$$\text{CHUYEN_BAY 3} = \text{CHUYEN_BAY} \triangleright \triangleleft \text{TUYEN_BAY 3}$$

Phân rã quan hệ **CHI_TIET_HANG_VE** thành các mảnh ngang
CHI_TIET_HANG_VE 1, CHI_TIET_HANG_VE 2, CHI_TIET_HANG_VE 3,
được định nghĩa như sau:

CHI_TIET_HANG_VE 1 = CHI_TIET_HANG_VE ▷◁ TUYEN_BAY 1

CHI_TIET_HANG_VE 2 = CHI_TIET_HANG_VE ▷◁ TUYEN_BAY 2

CHI_TIET_HANG_VE 3 = CHI_TIET_HANG_VE ▷◁ TUYEN_BAY 3

Phân rã quan hệ **VE** thành các mảnh ngang VE 1, VE 2, VE 3 được định nghĩa
như sau:

VE 1 = VE ▷◁ CHUYEN_BAY 1

VE 2 = VE ▷◁ CHUYEN_BAY 2

VE 3 = VE ▷◁ CHUYEN_BAY 3

Phân rã quan hệ **DAI_LY** thành các mảnh ngang DAI_LY 1, DAI_LY 2,
DAI_LY 3 được định nghĩa như sau:

DAI_LY 1 là các quầy bán vé của CHUYEN_BAY 1

DAI_LY 2 là các quầy bán vé của CHUYEN_BAY 2

DAI_LY 3 là các quầy bán vé của CHUYEN_BAY 3

Phân rã quan hệ **HANH_KHACH** thành các mảnh ngang HANH_KHACH 1, HANH_KHACH 2, HANH_KHACH 3 được định nghĩa như sau:

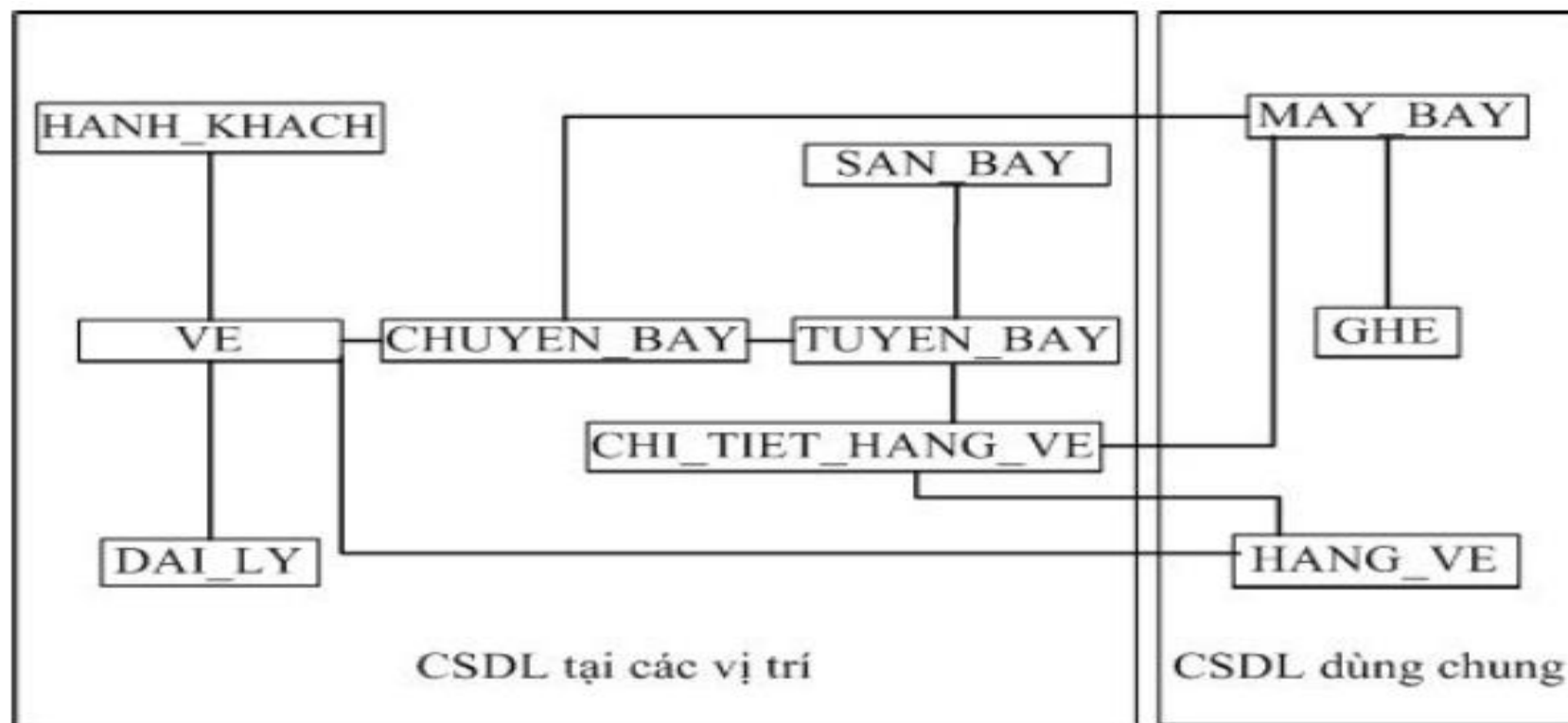
HANH_KHACH 1 là các hành khách mua vé CHUYEN_BAY 1 mà DAL_LY 1 bán

HANH_KHACH 2 là các hành khách mua vé CHUYEN_BAY 2 mà DAL_LY 2 bán

HANH_KHACH 3 là các hành khách mua vé CHUYEN_BAY 3 mà DAL_LY 3 bán

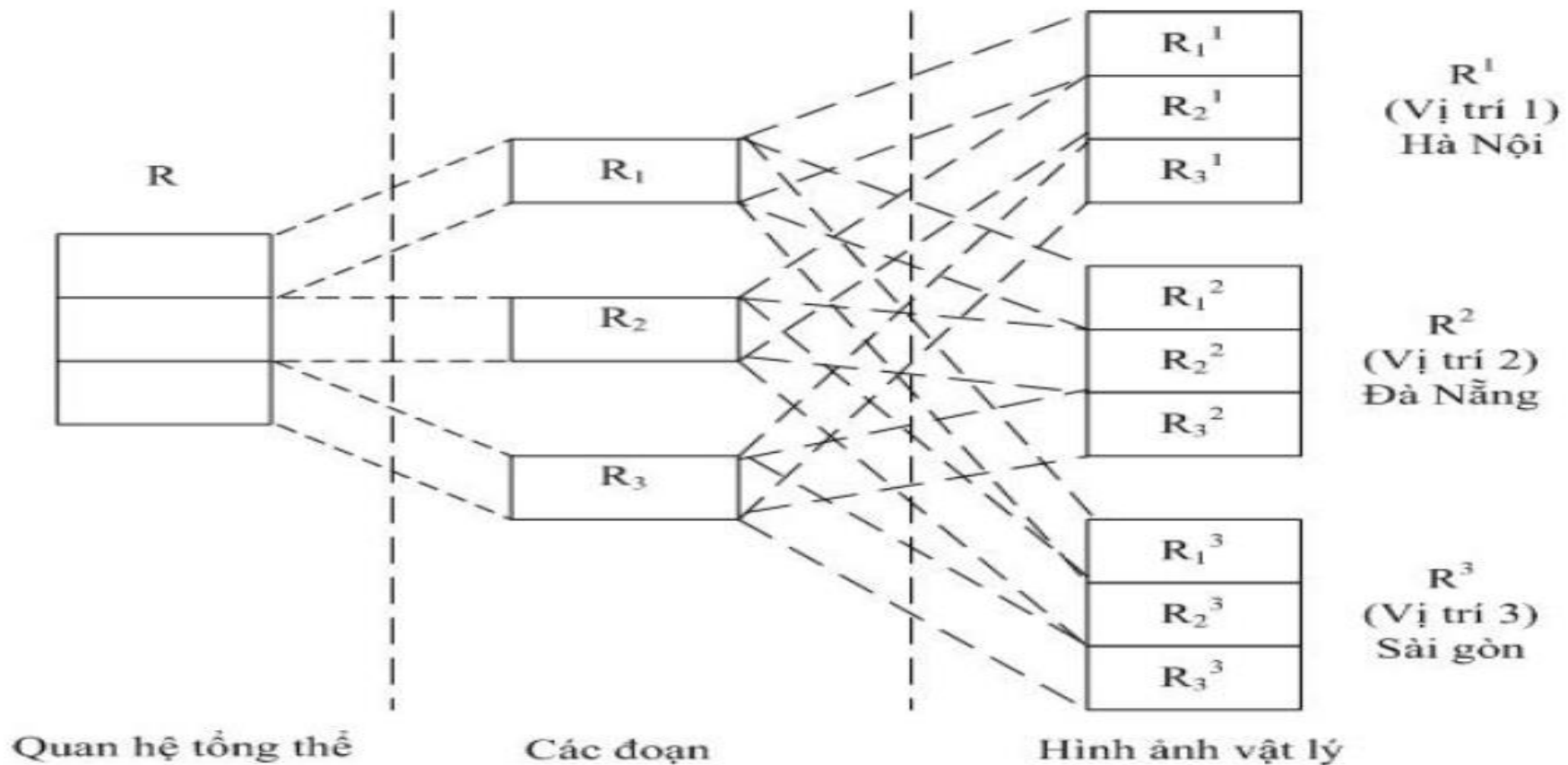
Đối với các quan hệ **MAY_BAY, GHE, HANG_VE** dùng chung cho 3 vị trí nên ta không phân mảnh nên ta thường xuyên thêm mới và cập nhật đồng bộ giữa 3 vị trí.

(tiếp vd)Thiết kế định vị (cấp phát)



Hình 3.3. Sơ đồ định vị của các mảnh tại các vị trí

(tiếp vd)



Hình 3.4. Các đoạn và hình ảnh vật lý của một quan hệ tổng thể

TABLE 3.1. Các đoạn và hình ảnh vật lý của một quan hệ tổng thể

R : Gồm các quan hệ tổng thể là SAN_BAY, TUYEN_BAY, CHUYEN_BAY, CHI_TIET_HANG_VE, VE, DAI_LY, HANH_KHACH.

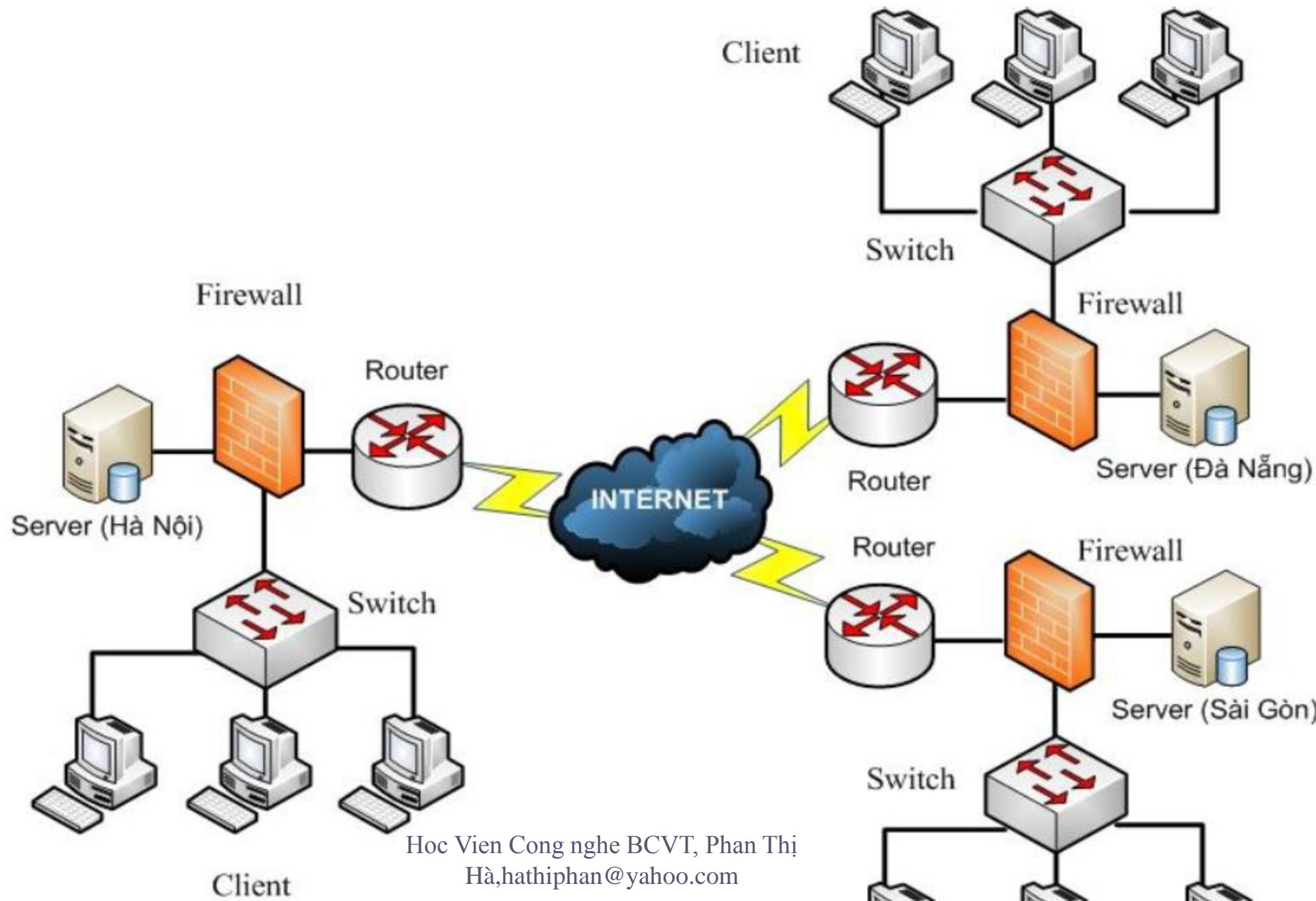
Còn quan hệ tổng thể MAY_BAY, GHE, HANG_VE dùng chung tại ba vị trí nên ta chỉ thêm mới và cập nhật đồng bộ giữa ba vị trí.

Thiết kế CSDL vật lý tại mọi trạm cho DDBMS (I ĐAX cục hồ)

Bảng 3.2. Danh mục các bảng dữ liệu của hệ thống bán vé máy bay

STT	Tên bảng	Ý nghĩa
1	SAN_BAY	Sân bay, chứa thông tin liên quan đến sân bay phục vụ cho công tác quản lý các sân bay.
2	MAY_BAY	Máy bay, chứa thông tin liên quan đến máy bay phục vụ cho công tác quản lý các máy bay.
3	GHE	Ghế, chứa thông tin liên quan đến ký hiệu ghế trong mỗi máy bay .
4	CHUYEN_BAY	Chuyến bay, chứa thông tin liên quan đến chuyến bay phục vụ cho công tác quản lý các chuyến bay.
5	TUYEN_BAY	Tuyến bay, chứa thông tin liên quan đến các tuyến bay của hệ thống.
6	HANG_VE	Hạng vé, chứa thông tin liên quan đến hạng vé chuyến bay mà hành khách lựa chọn.
7	CHI_TIET_HANG_VE	Chi tiết hạng vé, chứa thông tin liên quan đến số lượng vé, đơn giá theo từng hạng vé của mỗi máy bay ứng với mỗi tuyến bay.
8	VE	Vé, chứa thông tin liên quan đến thông tin của vé xuất cho hành khách.
9	DAI_LY	Đại lý, chứa thông tin liên quan đến đại lý phục vụ cho việc bán vé trong hệ thống.
10	HANH_KHACH	Hành khách, chứa thông tin liên quan đến hành khách mua vé trong hệ thống.

Ta có mô hình hệ thống mạng mô phỏng như sau:



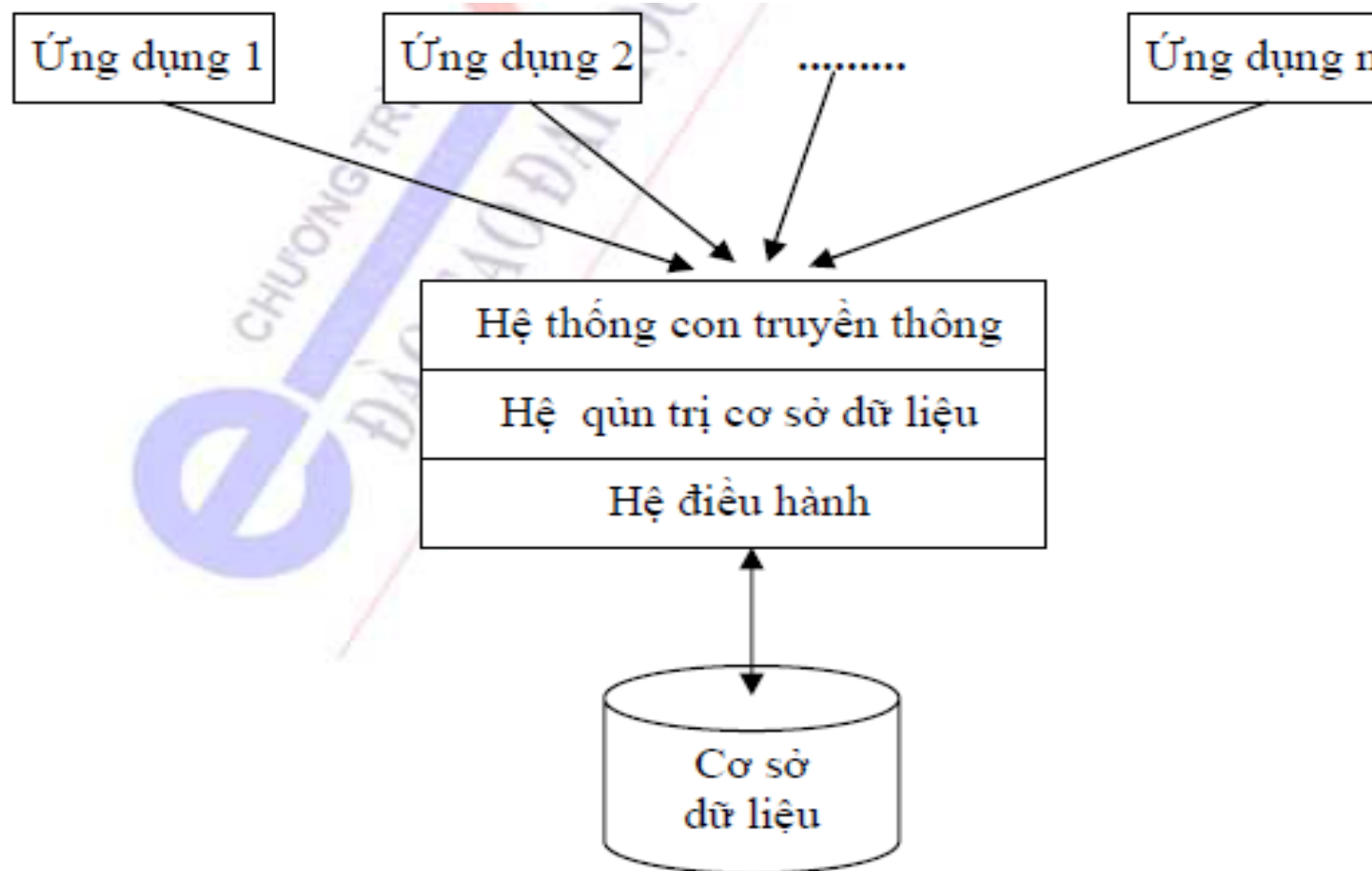
Các kiểu thiết kế CSDLPT trên mạng

- ✎ Các bản sao : CSDL được sao chép Lưu trữ trên các site khác nhau
- ✎ Phân mảnh: CSDL được phân ra thành các mảnh nhỏ lưu trên các site khác nhau
- ✎ Kết hợp các bản sao và phân mảnh: Trên 1 số site chứa bản sao, 1 số site khác chứa phân mảnh

8. Hệ quản trị CSDLPT

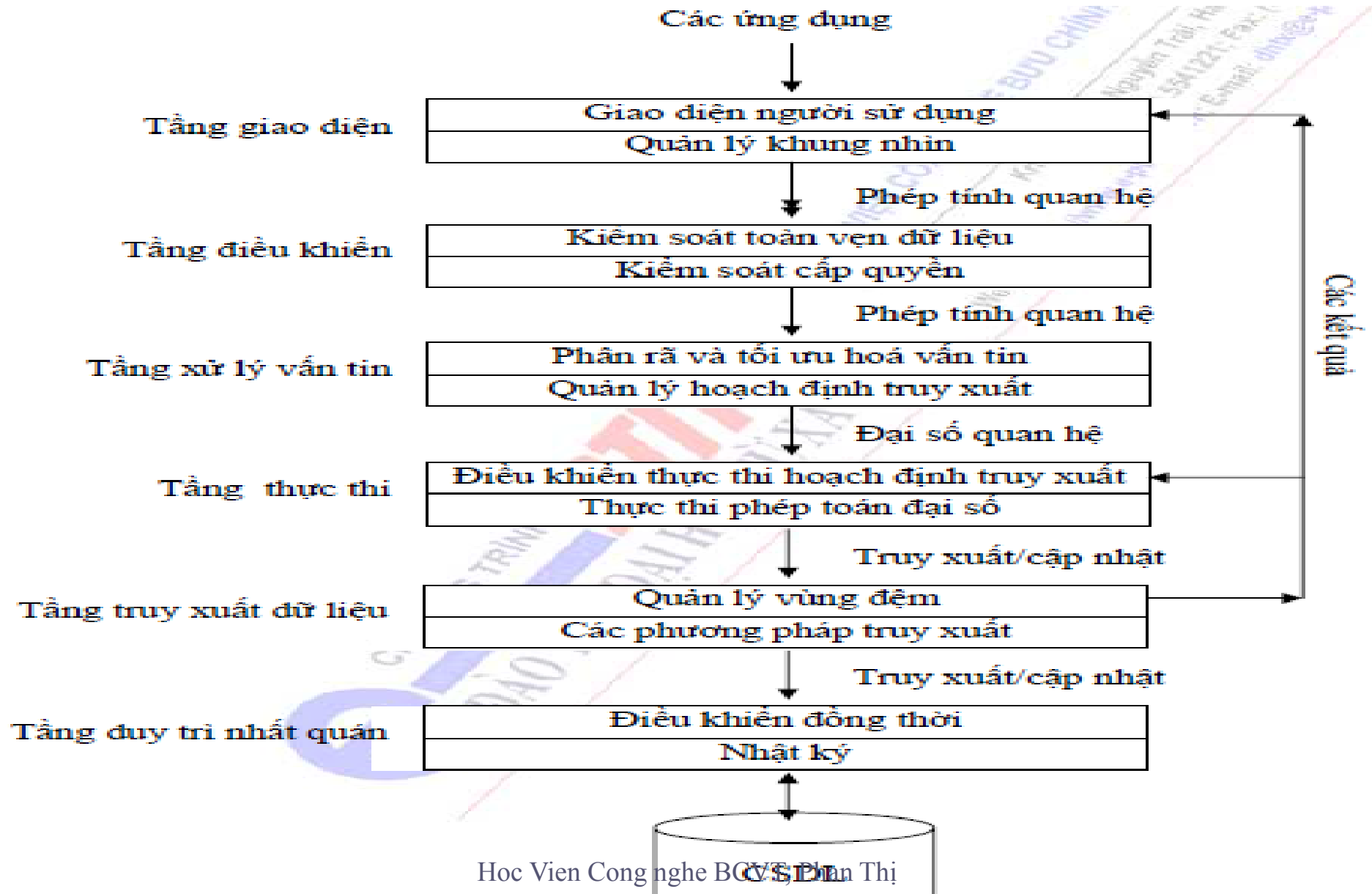
- ❑ Kiến trúc tổng quát về mô hình và chức năng của hệ QT CSDL (DBMS)
- ❑ Hệ QT CSDLPT (DDBMS)
- ❑ Các thành phần quản lý trên mỗi site của DDBMS
- ❑ Phương thức truy nhập từ xa của DDBMS
- ❑ DDBMS Thuần nhất và không thuần nhất
- ❑ Hệ QTCSDLPT kiểu Client/server và ngang hàng
- ❑ Các thành phần xử lý dữ liệu của DDBMS

Kiến trúc tổng quát ủa hệ QTCSDL



Hình 1.7: Kiến trúc tổng quát của mô hình hệ quản trị cơ sở dữ liệu quan hệ

Chức năng của hệ QTCSDL

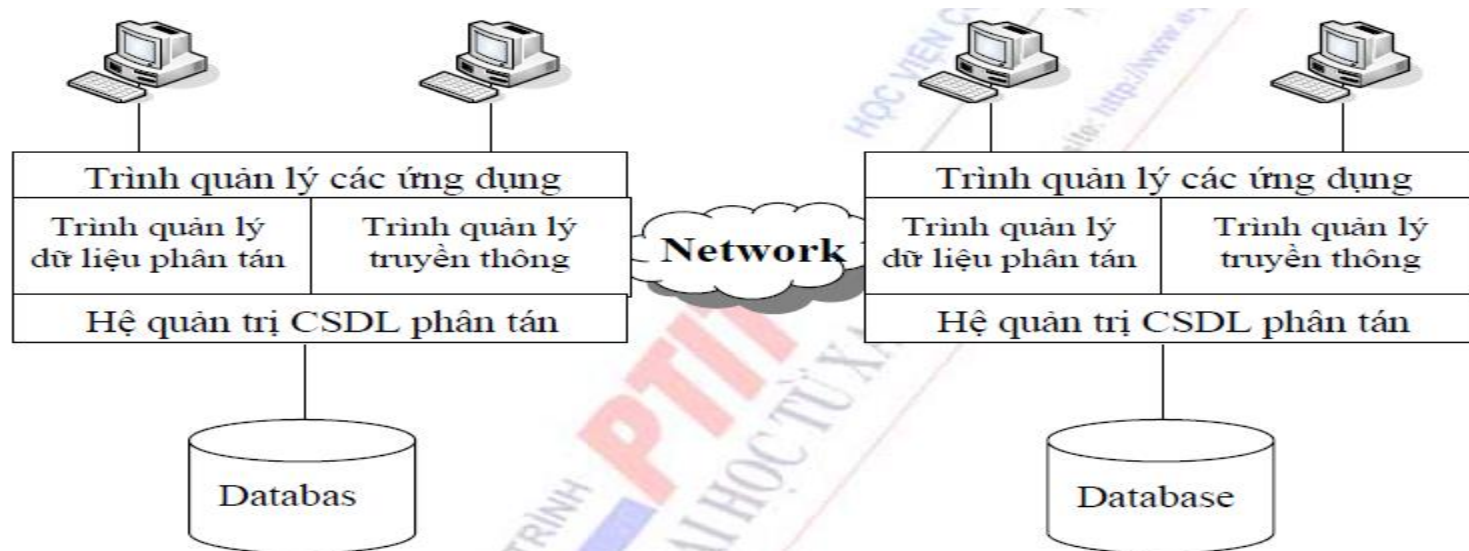


Hệ QTCSDLPT

- ☛ Là hệ thống pm quản lý CSDLPT và làm cho sự phân tán dl là trong suốt với ng sử dụng (csdl đc phân tán vật lý nhưng đc thống nhất, tổ chức như 1 csdl duy nhất)

Để quản lý dữ liệu phân tán tại môi trường cài đặt các thành phần sau

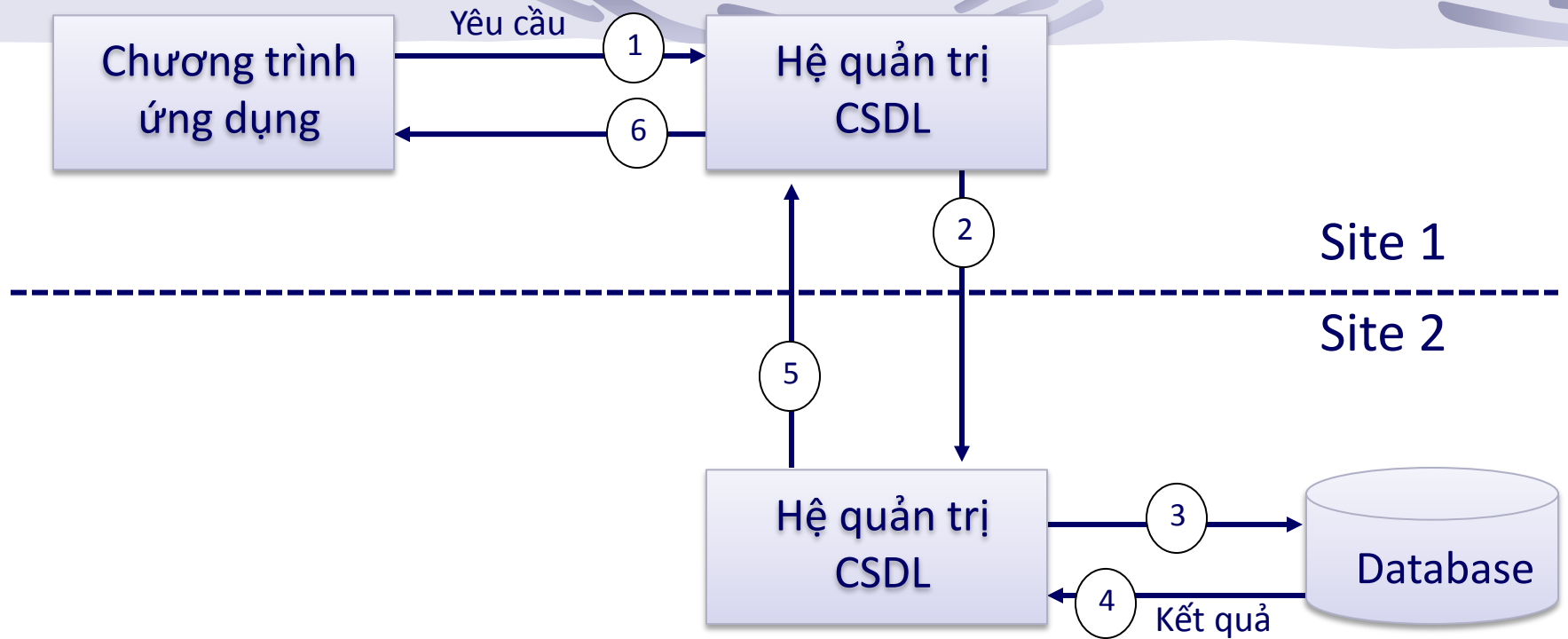
- ❑ TP quản trị CSDL (DM)
- ❑ Tp truyền thông (DC)
- ❑ Từ điển dữ liệu (DD): Thông tin dữ liệu phân tán trên mạng (phân mảnh, định vị, cấp phát,..)
- ❑ TP Cơ sở dữ liệu phân tán



DBMS hỗ trợ Truy nhập dữ liệu từ xa trực tiếp

- ❑ Các trình ứng dụng trên Client yêu cầu truy nhập trực tiếp vào các hệ cơ sở dữ liệu từ xa. Yêu cầu này được máy chủ CSDL thực hiện và gửi kết quả về Client.
- ❑ Tính trong suốt phân tán cao. Có thể thực hiện được bằng việc cung cấp các file chung và có thể truy nhập tự động theo địa chỉ các truy nhập trước đó.
- ❑ Tuy nhiên việc điều khiển tương tranh và khôi phục các tiến trình phân tán khi số người sử dụng tăng lên là khó khăn và phức tạp.

DBMS hỗ trợ Truy nhập dữ liệu từ xa trực tiếp



Hình 2.6 Phương thức truy nhập từ xa trực tiếp

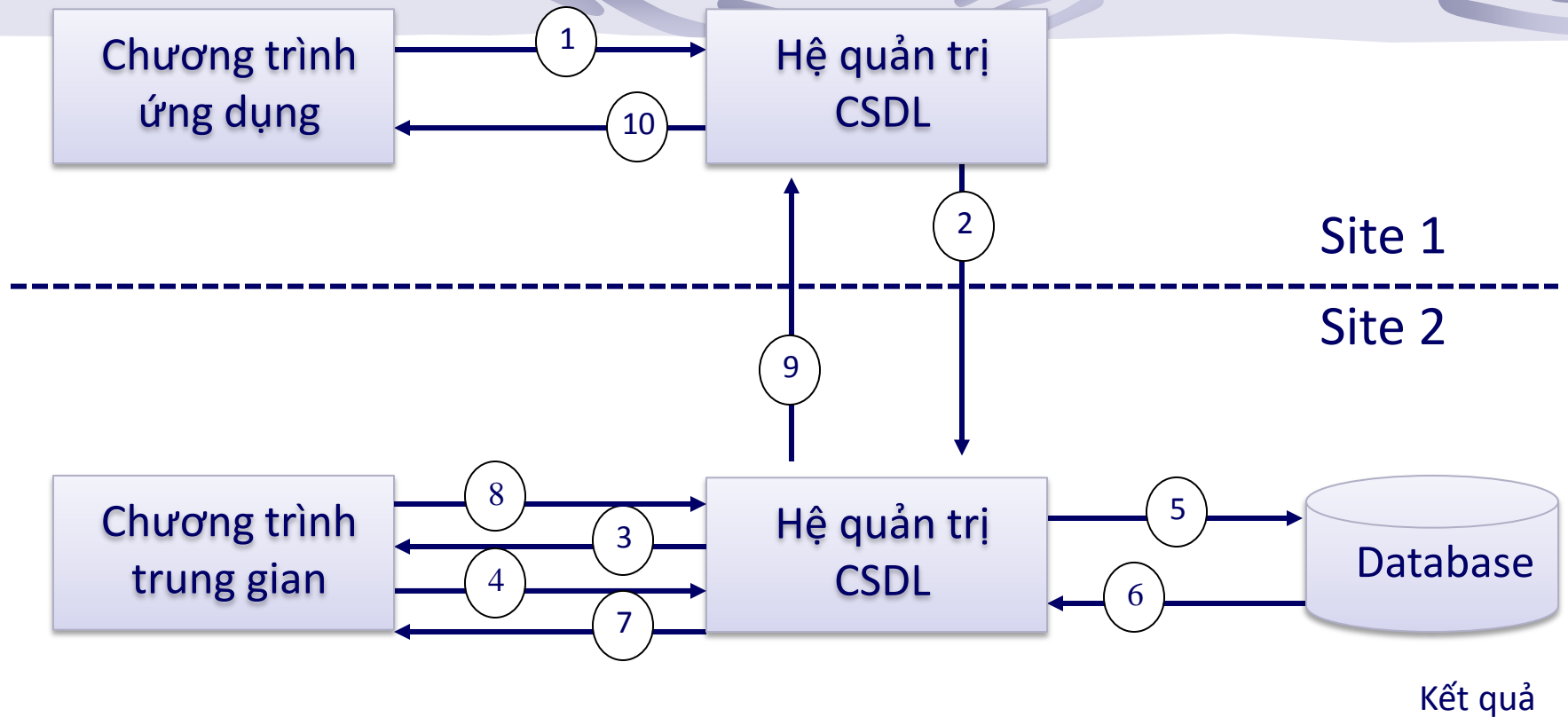
DBMS hỗ trợ Truy nhập dữ liệu từ xa qua trung

gian Middleware

- ❑ Middleware là bộ phần mềm trung gian, thực hiện việc điều khiển các tiến trình truyền thông và điều khiển cấp phát tài nguyên cho các tiến trình theo yêu cầu của các trình ứng dụng Client.

DBMS Truy nhập dữ liệu từ xa qua trung gian

Middleware



Hình 2.7 Phương thức truy nhập xa qua trung gian

DBMS hỗ trợ Truy nhập dữ liệu từ xa qua trung

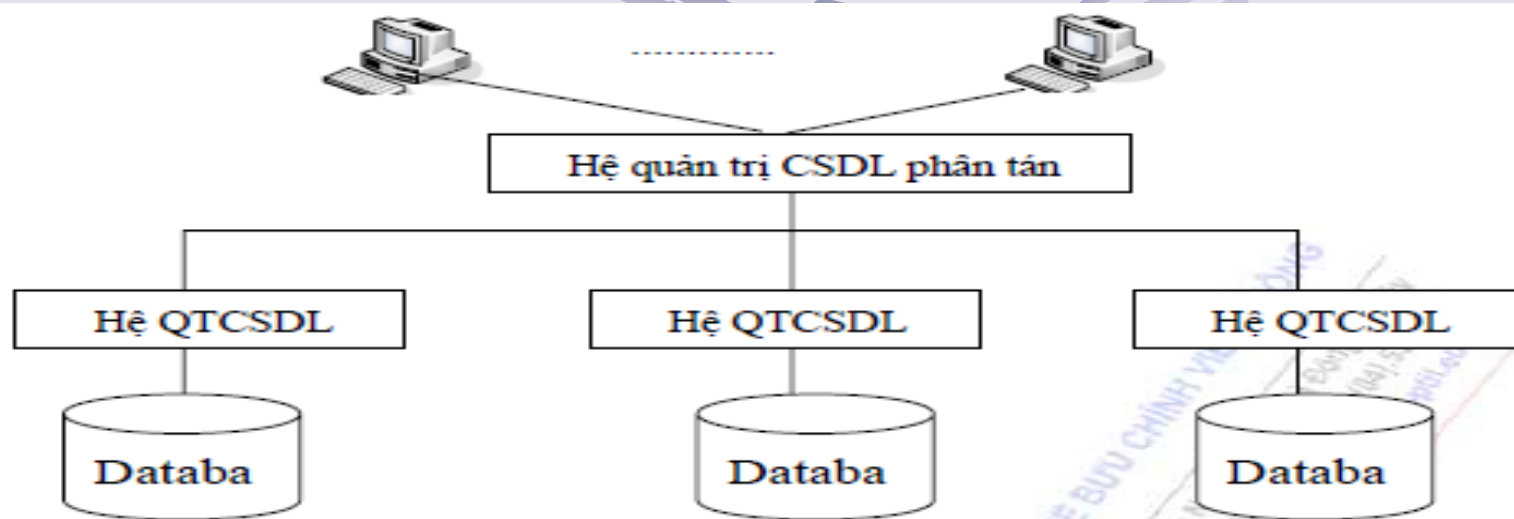
gian Middleware

- ❑ Middleware là bộ phần mềm trung gian, thực hiện việc điều khiển các tiến trình truyền thông và điều khiển cấp phát tài nguyên cho các tiến trình theo yêu cầu của các trình ứng dụng Client.
- ❑ Phương thức này hiệu quả cao. Khi nhiều trình ứng dụng yêu cầu truy nhập phân tán từ xa, Middleware sẽ thực hiện và điều khiển các tiến trình, hỗ trợ điều khiển tương tranh và khôi phục nhanh các tiến trình phân tán.

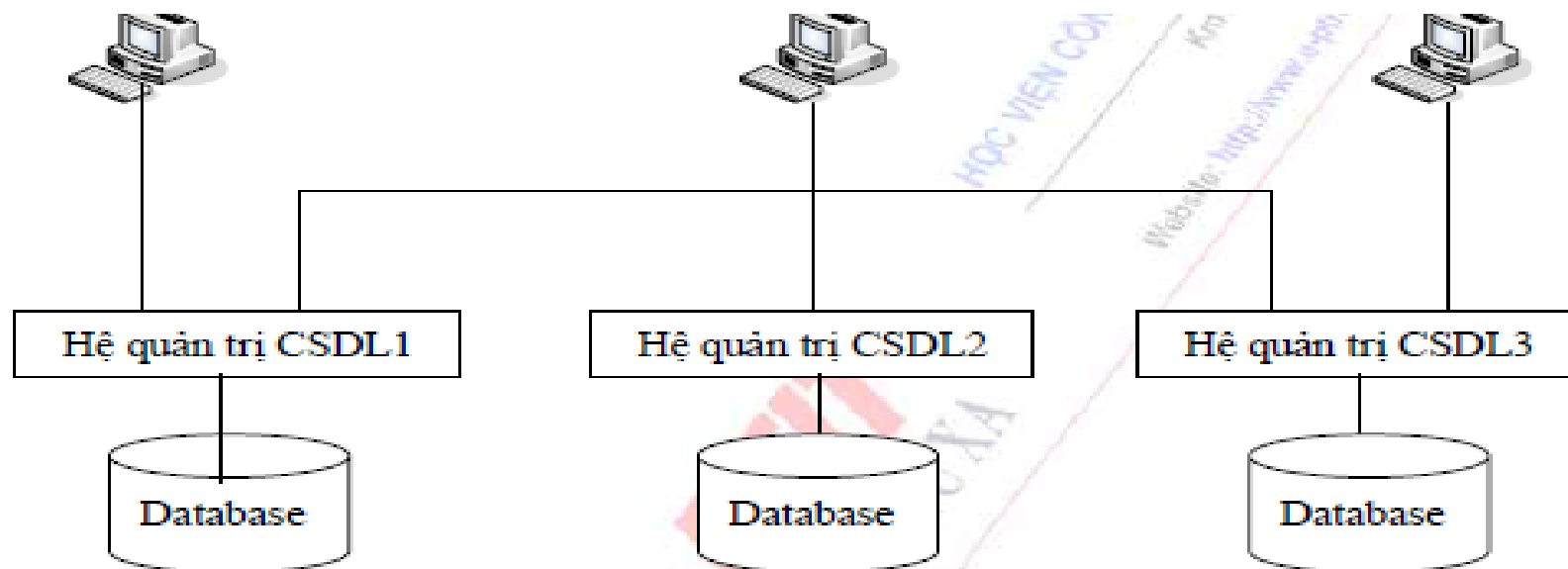
Hệ QTCSDLPT thuần nhất và không thuần nhất

Đặc tính trong suốt trong ax cục bộ cho phép xd ht CSDLPT đồng nhất và không đồng nhất

- Đồng nhất: Lược đồ độc lập tại các site đc đn sử dụng cùng 1 mô hình như DBMS cục bộ
- Không đồng nhất: Lược đồ ánh xạ cục bộ dùng để phối hợp các kiểu khác nhau của DBMS



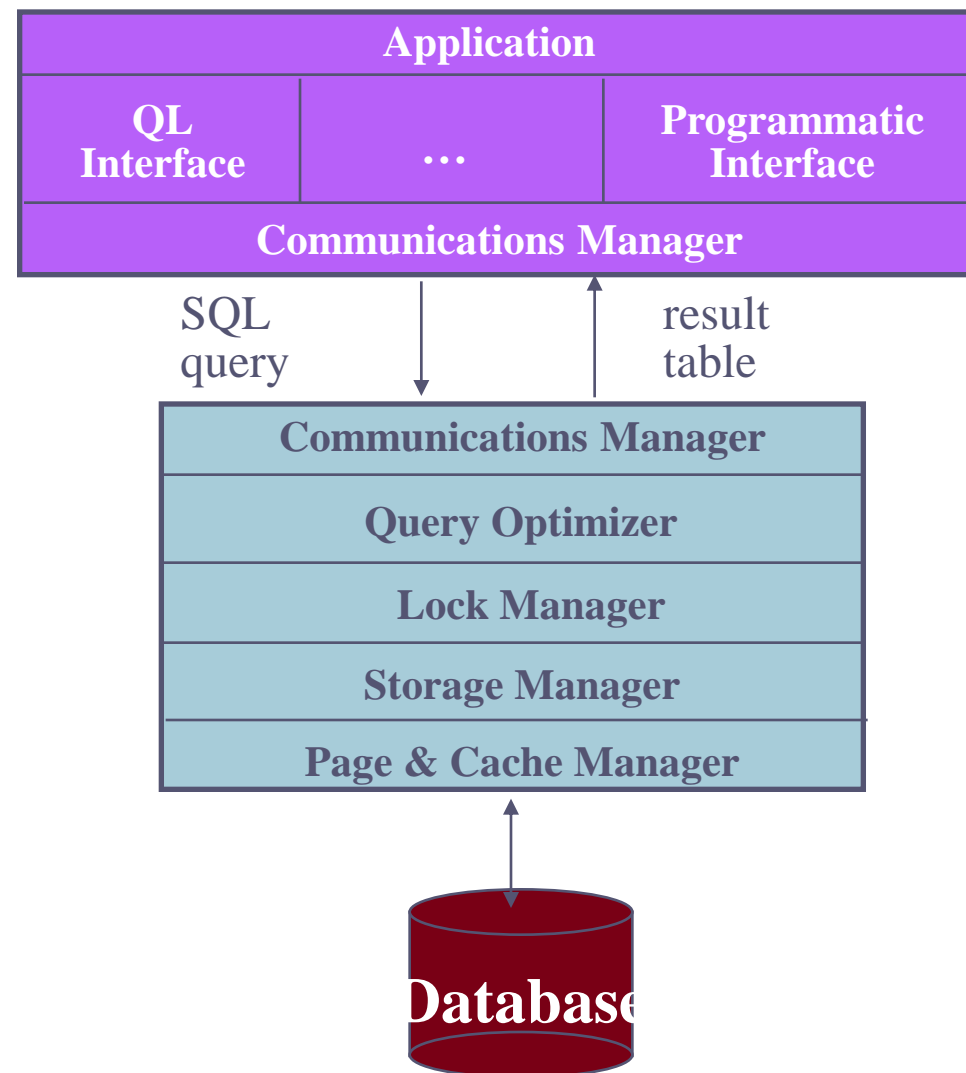
Hình 1.11 Kiến trúc mô hình hệ QTCSDLPT thuần nhất



Hình 1.12 Kiến trúc mô hình hệ QTCSDLPT không thuần nhất

Ý tưởng chung: Chia các chức năng thành 2 lớp

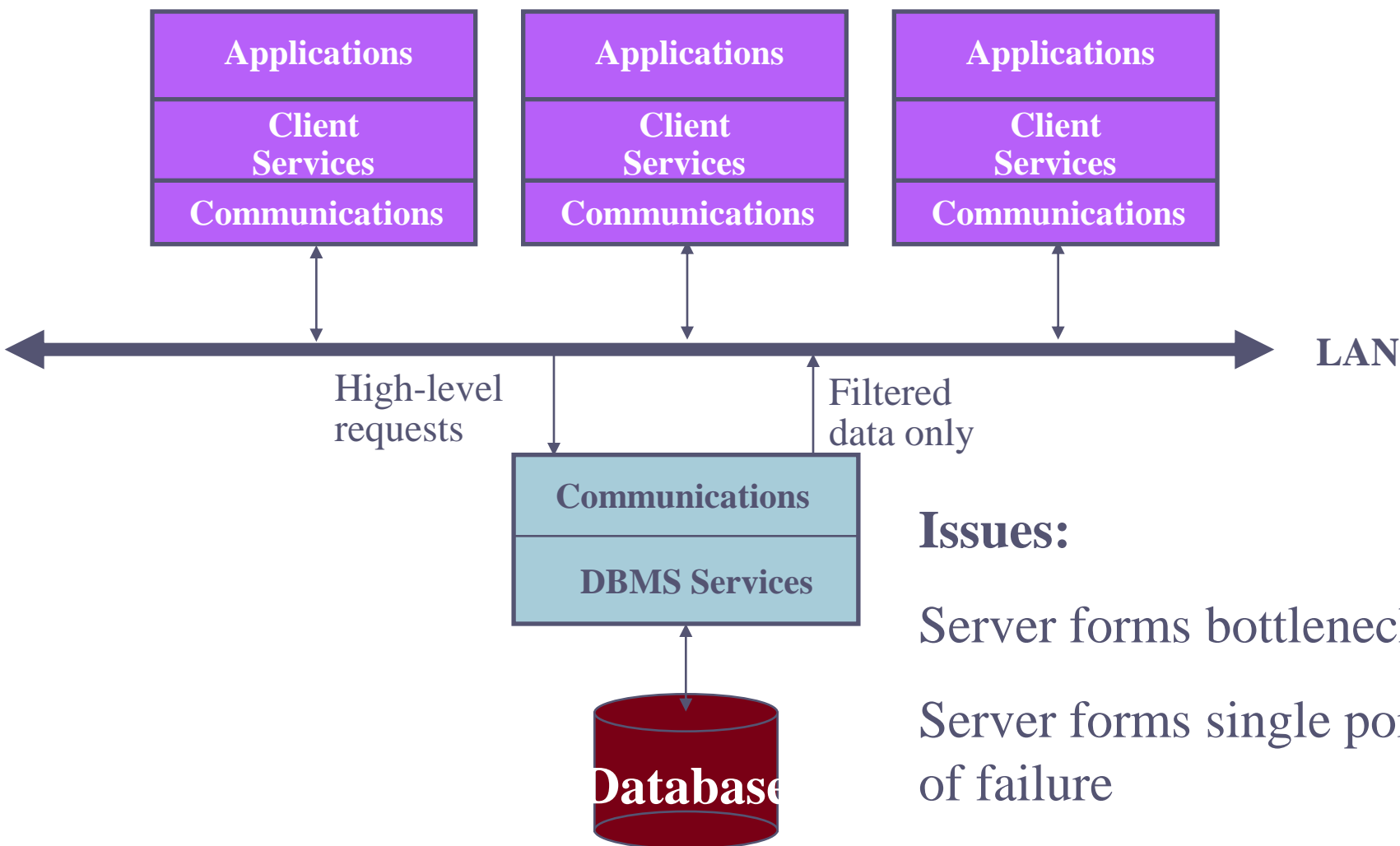
- Các chức năng phía chủ
 - Quản lý dữ liệu, tối ưu xử lý truy vấn, quản lý giao dịch v.v
- Các chức năng phía khách
 - Có thể bao gồm một phần chức năng quản lý dữ liệu, không chỉ UI
- Cung cấp kiến trúc 2 mức
- Phân chia hiệu quả hơn công việc
- Các loại kiến trúc C/S:
 - Multiple client/single server
 - Multiple client/multiple server



Multiclients/Single Server

www.ptit.edu.vn

n



Issues:

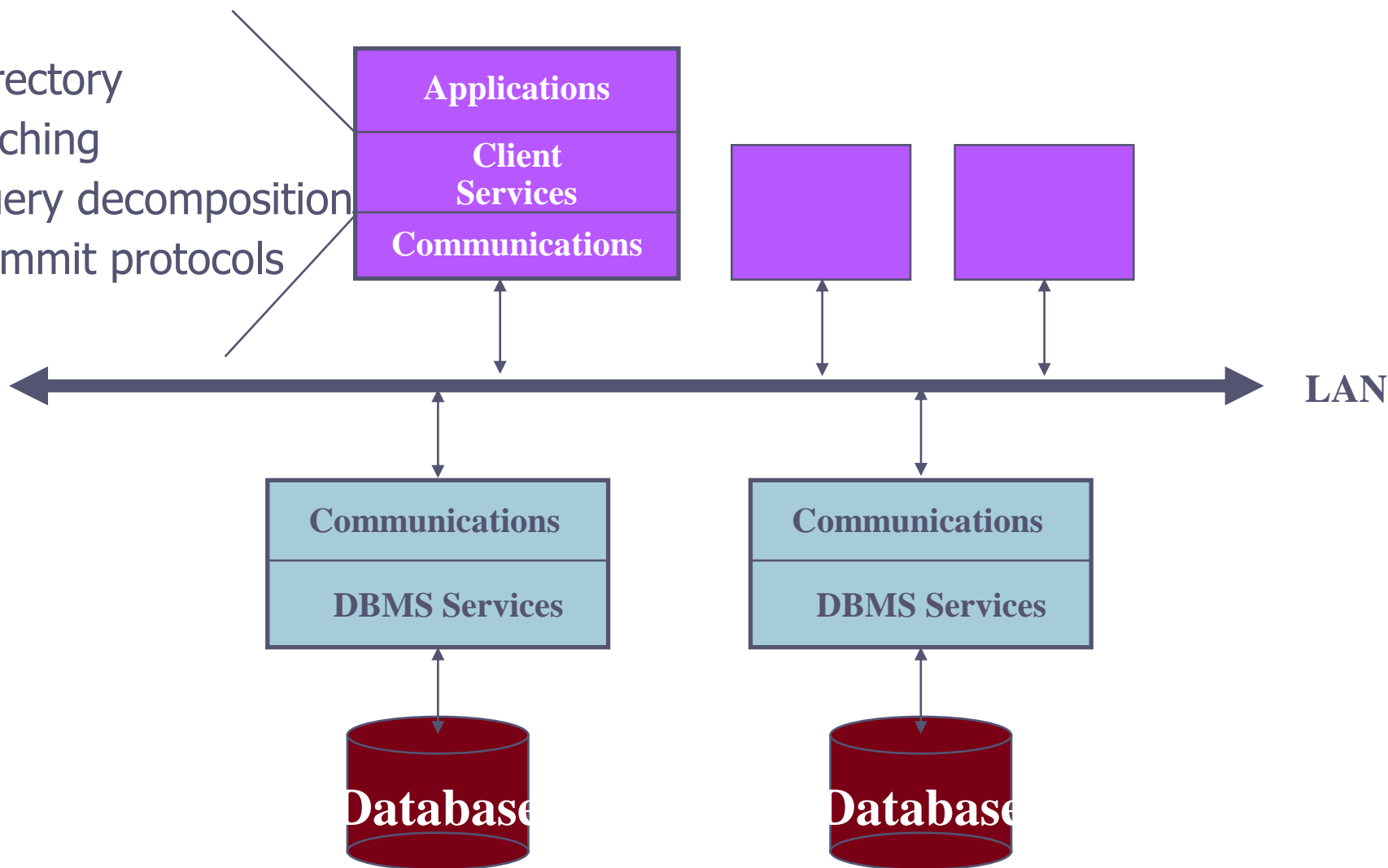
Server forms bottleneck

Server forms single point of failure

Database scaling difficult

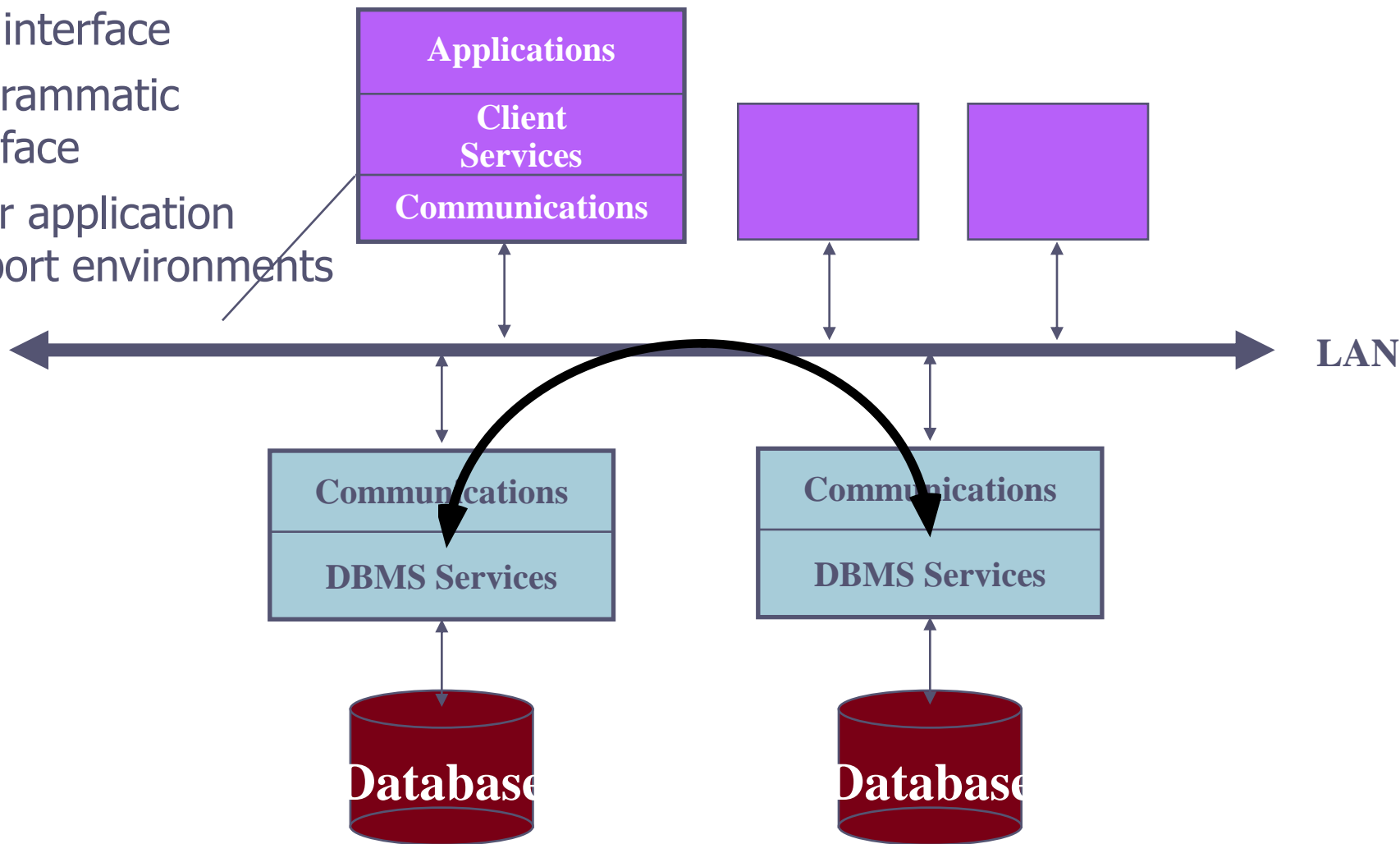
Multi Clients/ Multi Servers (1)

- directory
- caching
- query decomposition
- commit protocols

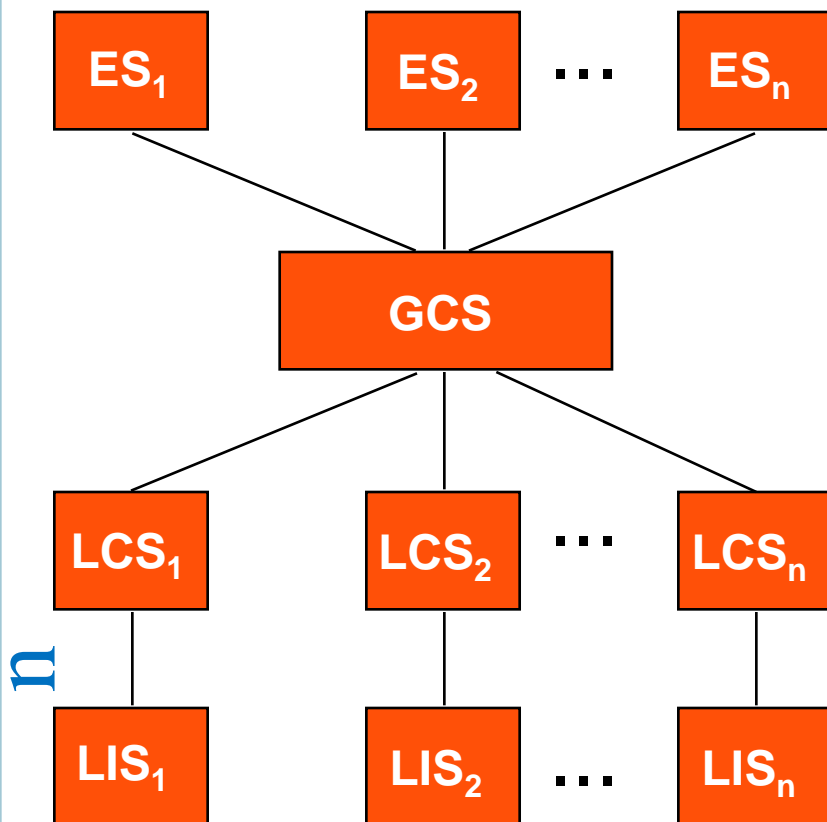


Multi Clients/ Multi Servers (2)

- SQL interface
- programmatic interface
- other application support environments



Kiến trúc ngang hàng của DBMS

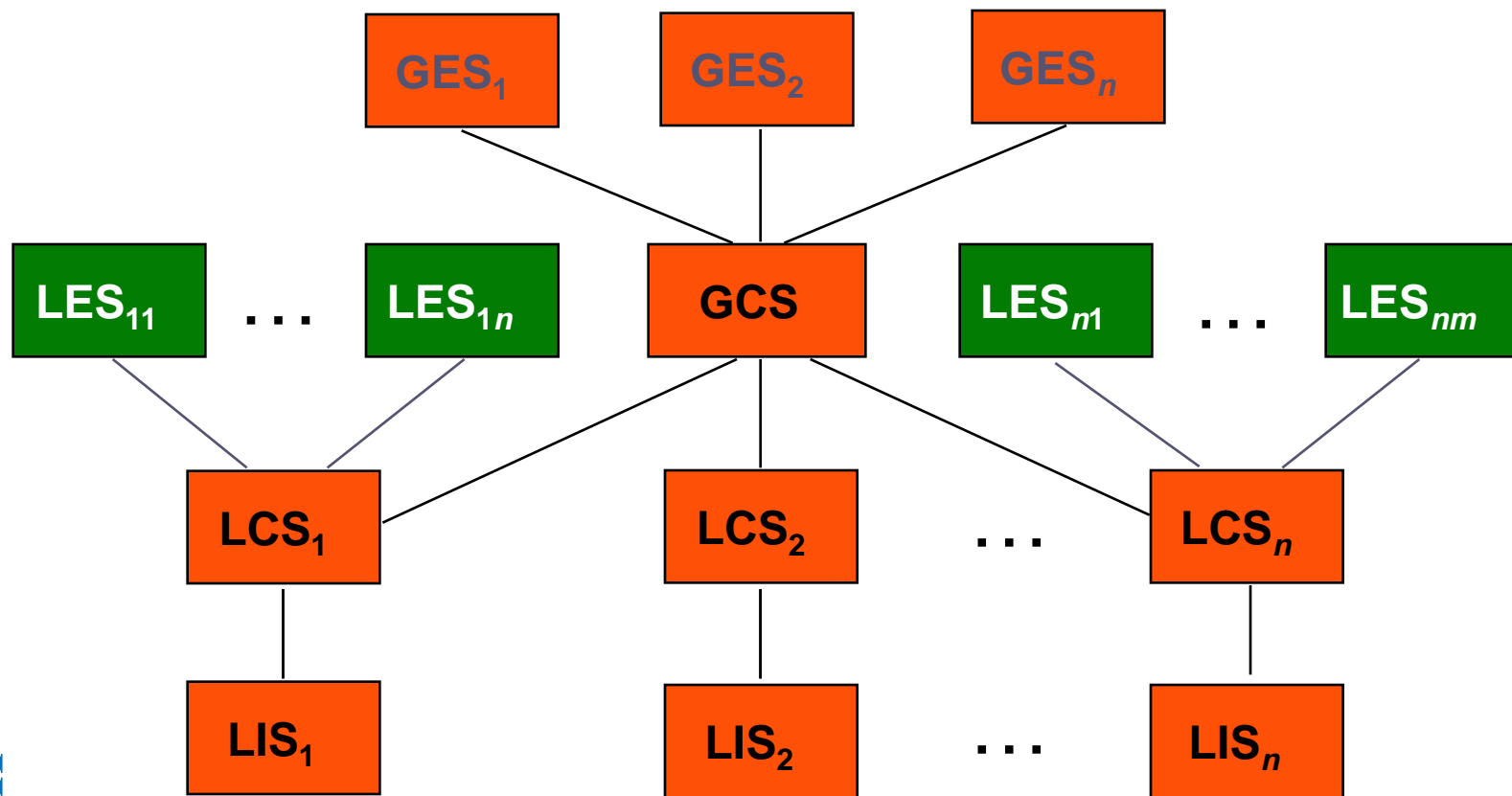


ES: External Schema

GCS: Global Conceptual Schema

LCS: Local Conceptual Schema

LIS: Local Internal Schema



- GES: Global External Schema
- LES: Local External Schema

- LCS: Local Conceptual Schema
- LIS: Local Internal Schema

1. Introduction

