**Mục lục**

[**1.** **Môi trường CSDL Tập chung và Phân tán** 2](#_Toc83656677)

[**2.** **Database Management (Quản lý CSDL)** 2](#_Toc83656678)

[**3.** **Hệ thống phân tán** 3](#_Toc83656679)

[**4.** **CSDL Phân tán** 3](#_Toc83656680)

[**5.** **Các thành phần cơ bản cần thiết của hệ QTCSDLPT** 3](#_Toc83656681)

[**6.** **Mô hình kiến trúc tham chiếu hệ CSDL phân tán** 5](#_Toc83656682)

[**7.** **Kiếm trúc hệ CSDL phân tán** 7](#_Toc83656683)

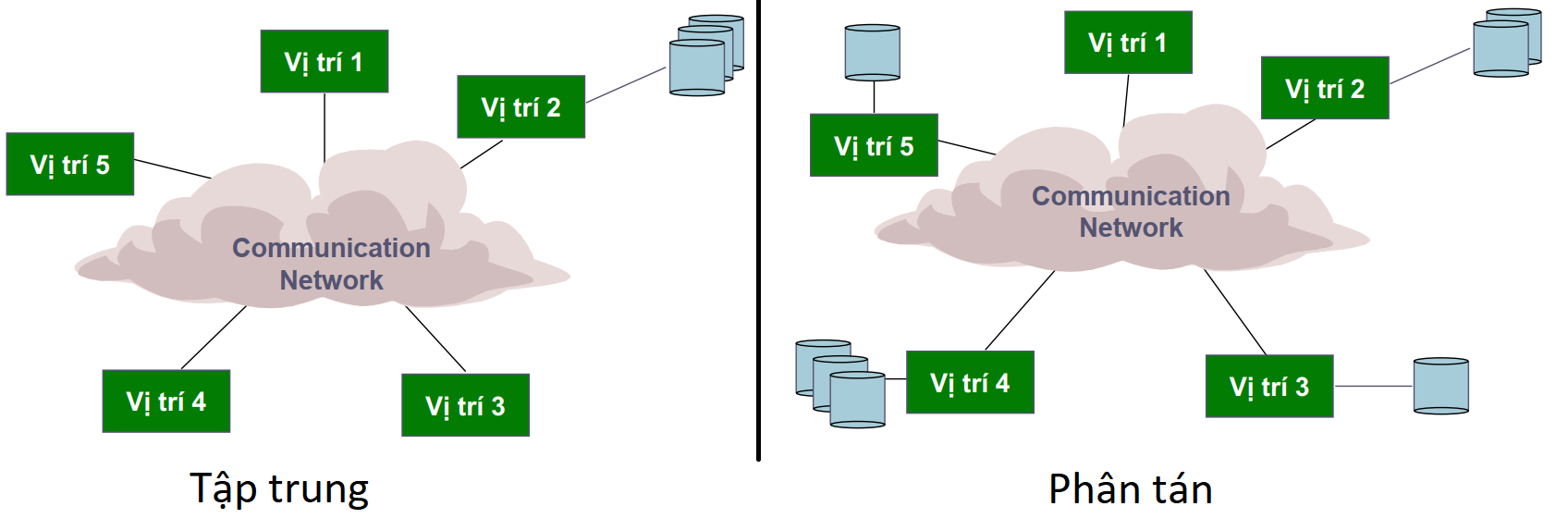
[*I.* *Kiến trúc ngang hàng (DDBS Peer-to-Peer Architecture)* 7](#_Toc83656684)

[*II. Kiến trúc Hệ thống quản trị CSDL phân tán (DDBMS Architecture)* 9](#_Toc83656685)

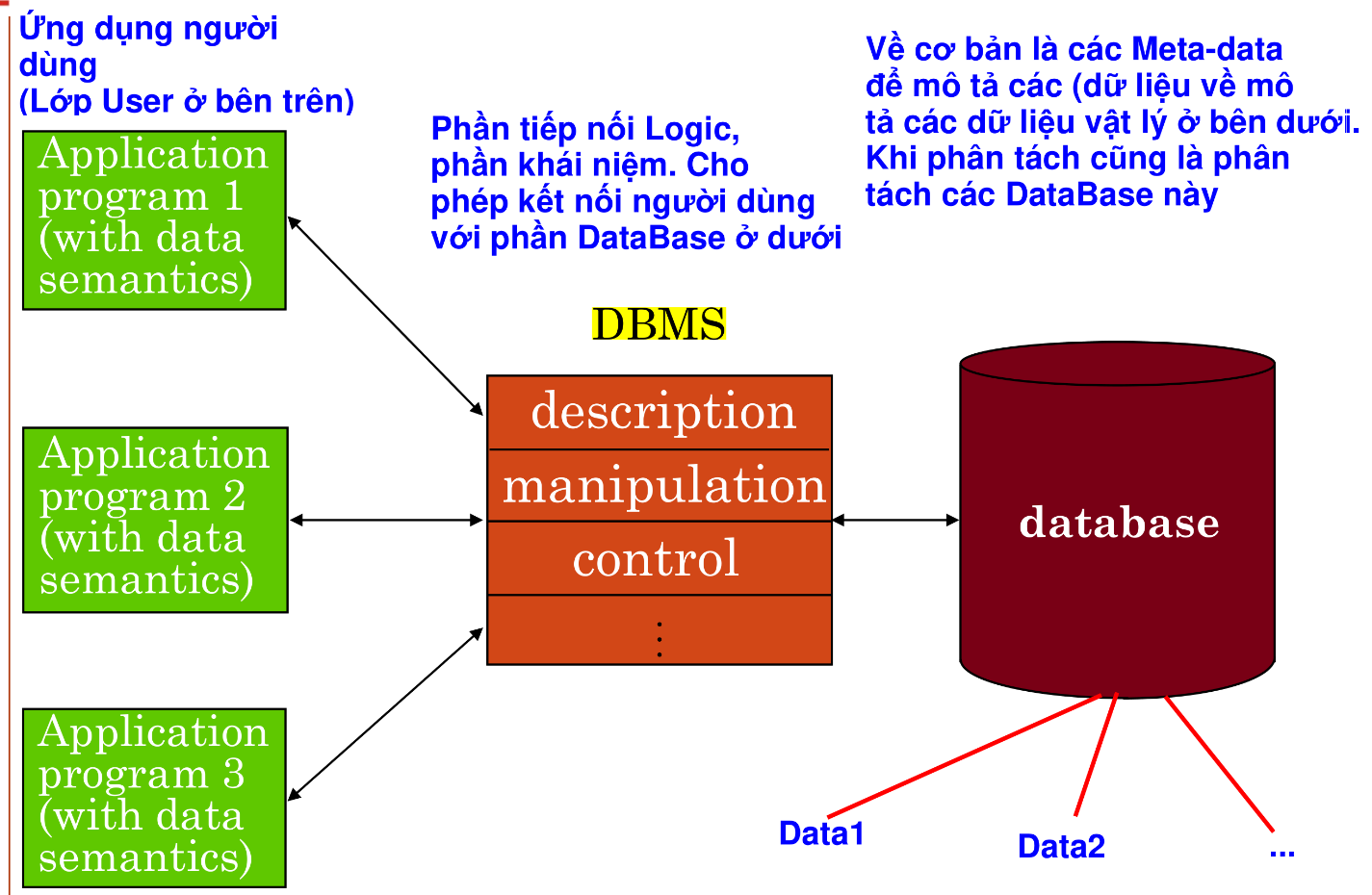
[*III. Kiến trúc Client/Server* 10](#_Toc83656686)

***- by Team 8 CSDL Phân tán PTIT***

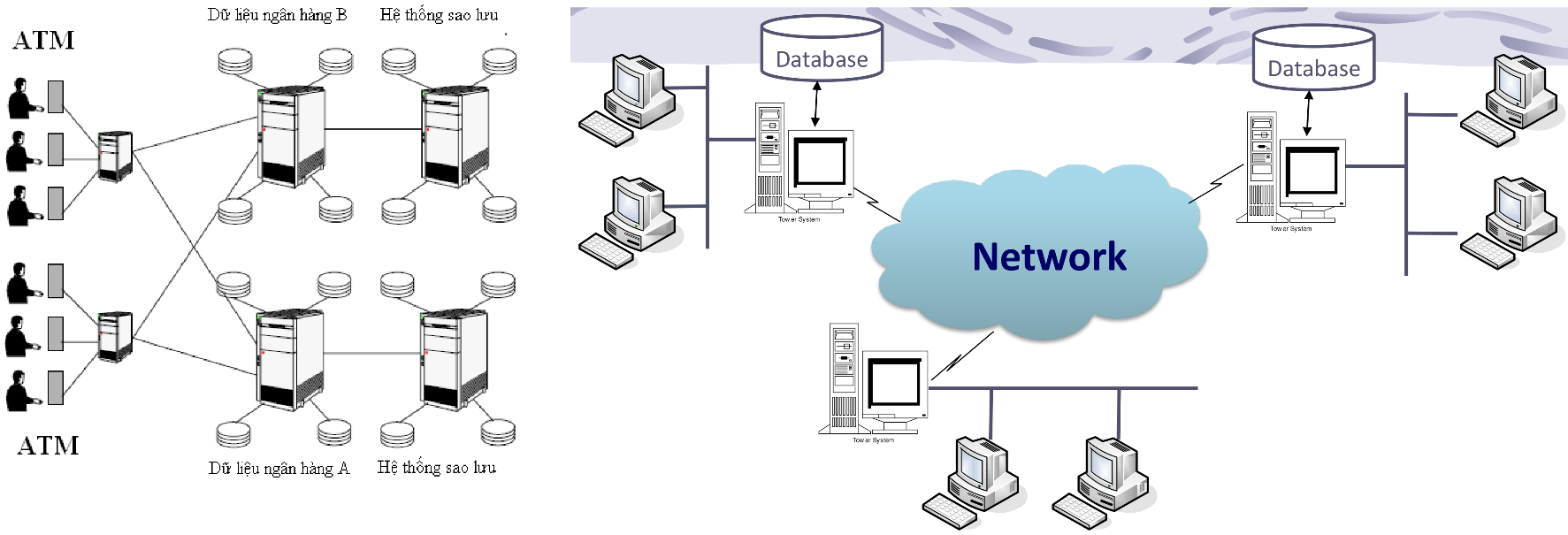
# **Môi trường CSDL Tập chung và Phân tán**



# **Database Management (Quản lý CSDL)**



# **Hệ thống phân tán**



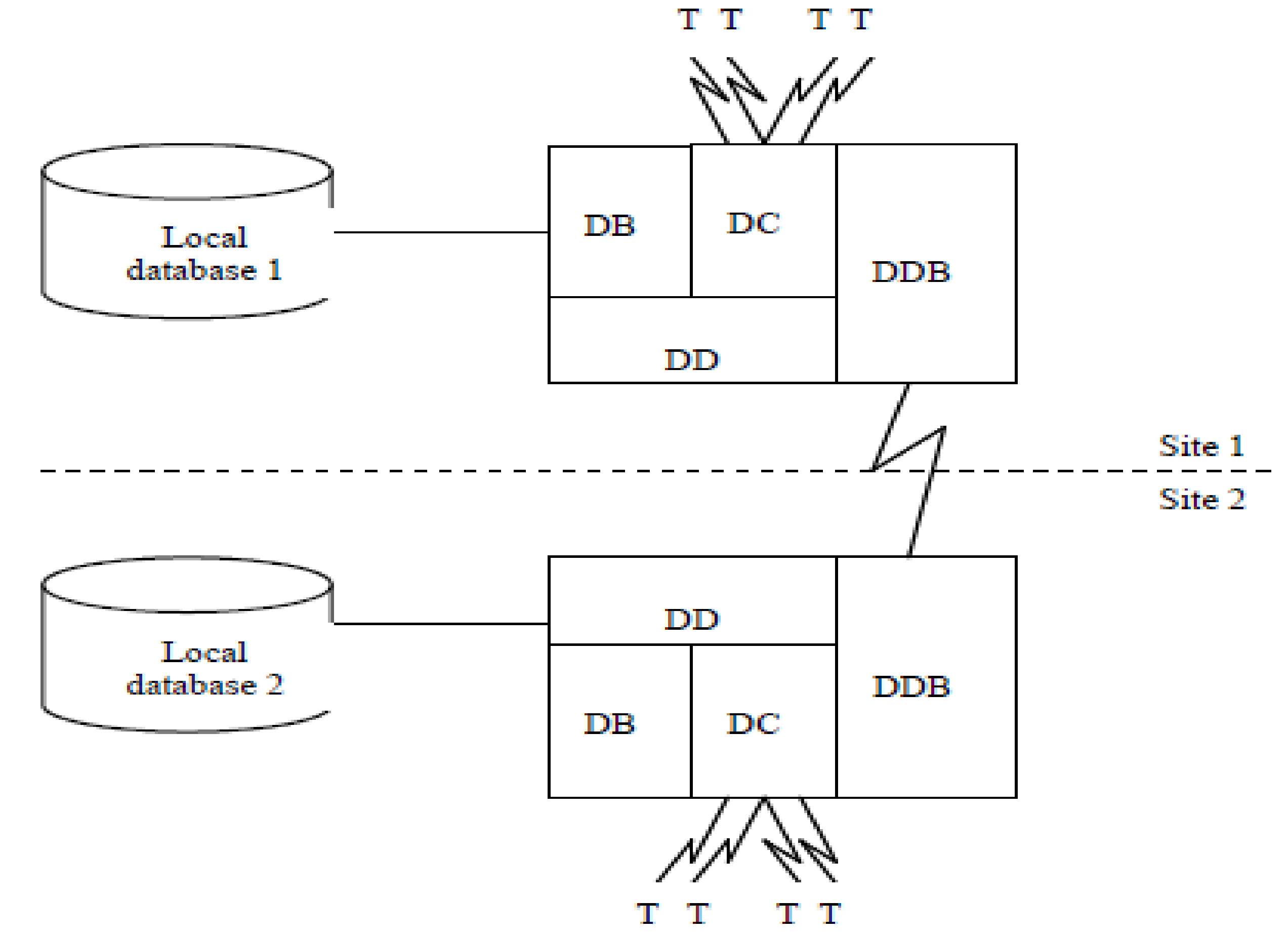
# **CSDL Phân tán**

**DDBS = DDB + DDBMS**

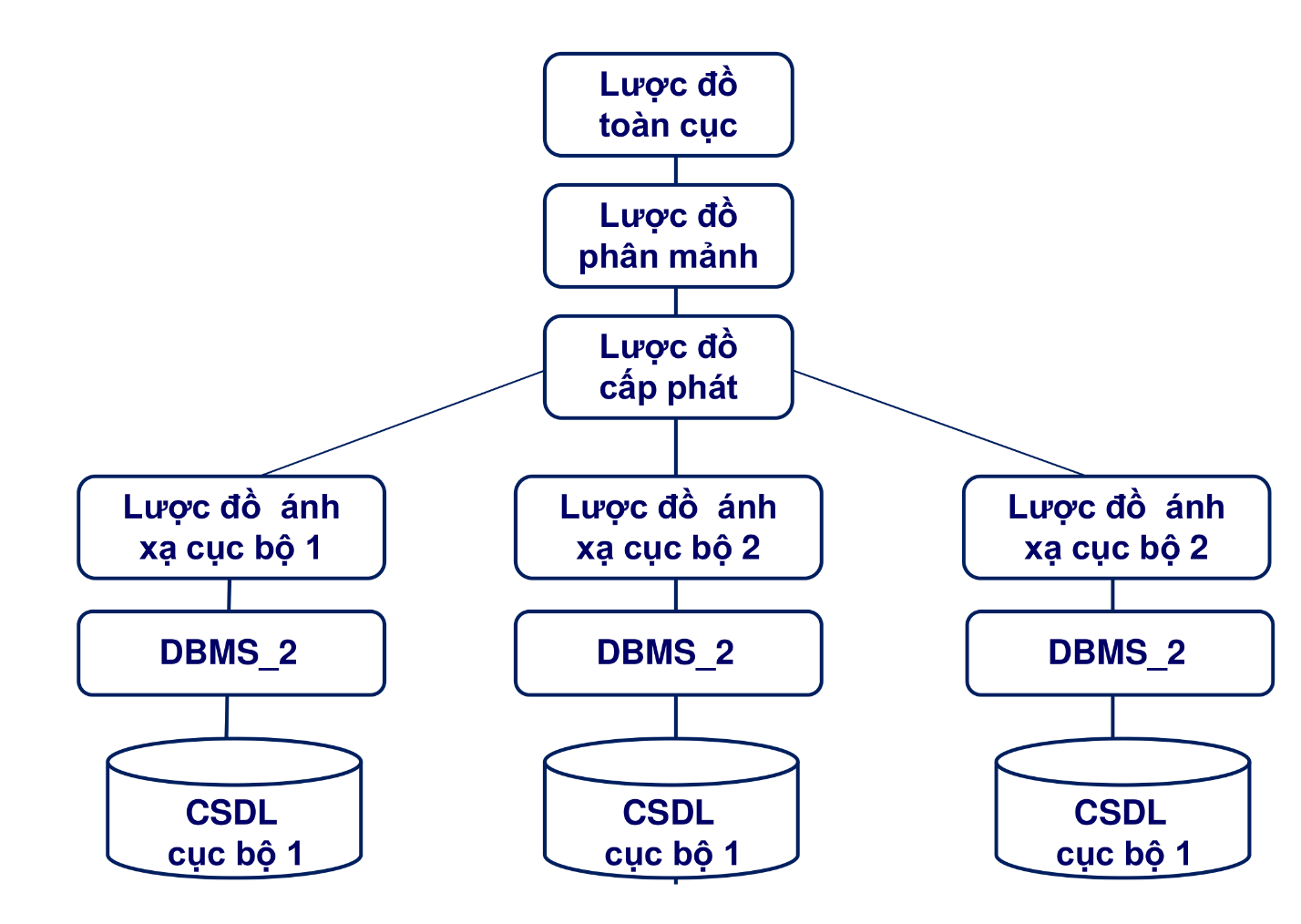
* **DDBS: Distributed Database Systems** (Hệ thống cơ sở dữ liệu phân tán)
* **DDB: Distributed Database** ( Cơ sở dữ liệu phân tán ) là cơ sở dữ liệu trong đó dữ liệu được lưu trữ trên các vị trí thực tế khác nhau. Nó có thể được lưu trữ trong nhiều máy tính ở cùng một vị trí thực tế; hoặc có thể phân tán trên một mạng lưới các máy tính được kết nối với nhau
* **DDBMS: Distributed Database Management System** (Hệ thống quản lý phân phối Cơ sở dữ liệu): là một ứng dụng tập trung quản lý cơ sở dữ liệu phân tán. hệ thống cơ sở dữ liệu này đồng bộ hóa dữ liệu theo định kỳ và đảm bảo rằng bất kỳ sự thay đổi trong dữ liệu do người dùng được phổ cập nhật trong cơ sở dữ liệu. DDBMS được sử dụng rộng rãi trong kho dữ liệu, nơi mà một lượng lớn dữ liệu được xử lý và truy cập bởi nhiều người dùng hoặc khách hàng cơ sở dữ liệu cùng một lúc. Hệ thống cơ sở dữ liệu này được sử dụng để quản lý dữ liệu trong mạng, duy trì bảo mật và xử lý toàn vẹn dữ liệu. Một hệ thống quản lý cơ sở dữ liệu phân tán được thiết kế cho nền tảng cơ sở dữ liệu không đồng nhất mà tập trung vào hệ thống quản lý cơ sở dữ liệu không đồng nhất.

# **Các thành phần cơ bản cần thiết của hệ QTCSDLPT**

* Thành phần quản trị cơ sở dữ liệu (DB: Database Management)
* Thành phần truyền dữ liệu (DC: Data Communication)
* Tự điển dữ liệu (DD : Data Dictionnary) mở rộng để biểu diễn thông tin về sự phân tán dữ liệu trên mạng.
* Thành phần cơ sở dữ liệu phân tán (DDB:Distributed Database)



# **Mô hình kiến trúc tham chiếu hệ CSDL phân tán**



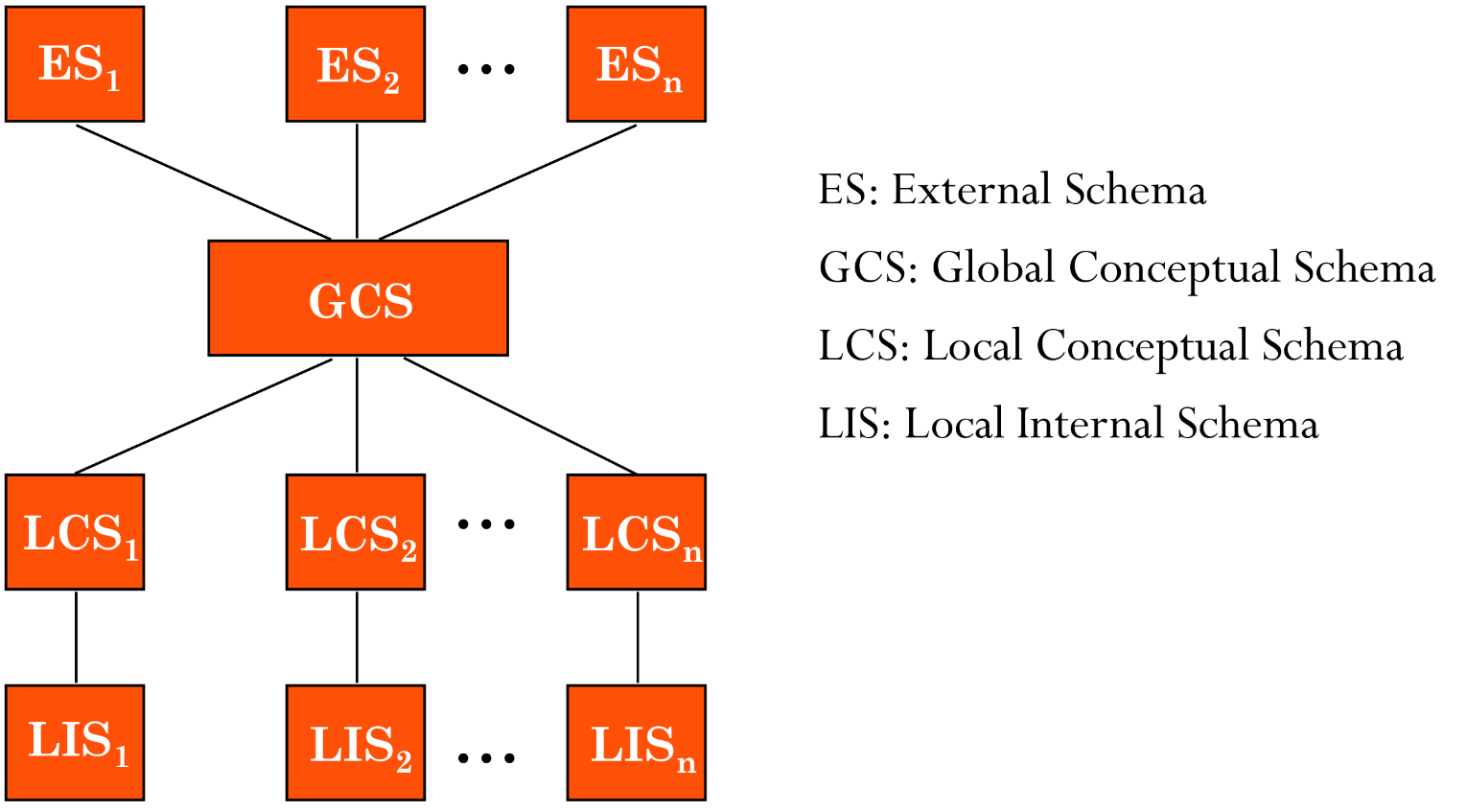
* **Lược đồ toàn cục:** tất cả dữ liệu được chứa trong CSDL phân tán như trong CSDL tập trung. Bao gồm thông tin về các thực thể, tập các phụ thuộc hàm và mối quan hệ giữa các thực thể,
* **Phân mảnh:** Quan hệ toàn cục có thể chia thành nhiều quan hệ con nhỏ hơn được gọi là các mảnh, tách biệt với nhau nhau. (Các phân mảnh là các thành phần của các quan hệ toàn cục mà được lưu trữ vật lý tại một hay một số địa điểm).
* **Lược đồ phân mảnh:** Là ánh xạ giữa quan hệ toàn cục và các phân mảnh. Ánh xạ này là mối quan hệ một-nhiều.
  + **Tên các mảnh =** tên của quan hệ toàn cục + một chỉ số (chỉ số phân mảnh)
  + Các mảnh có thể được cài đặt tại một hay nhiều vị trí khác nhau trên mạng
  + Các kiểu phân mảnh:
    - Phân mảnh ngang
    - Phân mảnh dọc
    - Phân mảnh hỗn hợp (cả 2)
  + Một mảnh được định nghĩa bằng một biểu thức đại số quan hệ trong đó các toán hạng là các quan hệ toàn cục.
* **Lược đồ cấp phát:** chứa thông tin các mảnh được chứa trên những địa điểm (site) nào – vì mỗi mảnh có thể được chứa ở một hay nhiều nơi (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 / địa điểm. Tất cả các mảnh của một quan hệ tổng thể trên cùng một trạm tạo ra ảnh vật lý
  + Tất cả các mảnh được liên kết với cùng một quan hệ tổng thể R và được định vị (đặt) tại cùng một trạm j 🡪 ảnh vật lý quan hệ tổng thể R tại trạm j.

🡺 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 đặt tên = tên của quan hệ tổng thể và một chỉ mục trạm

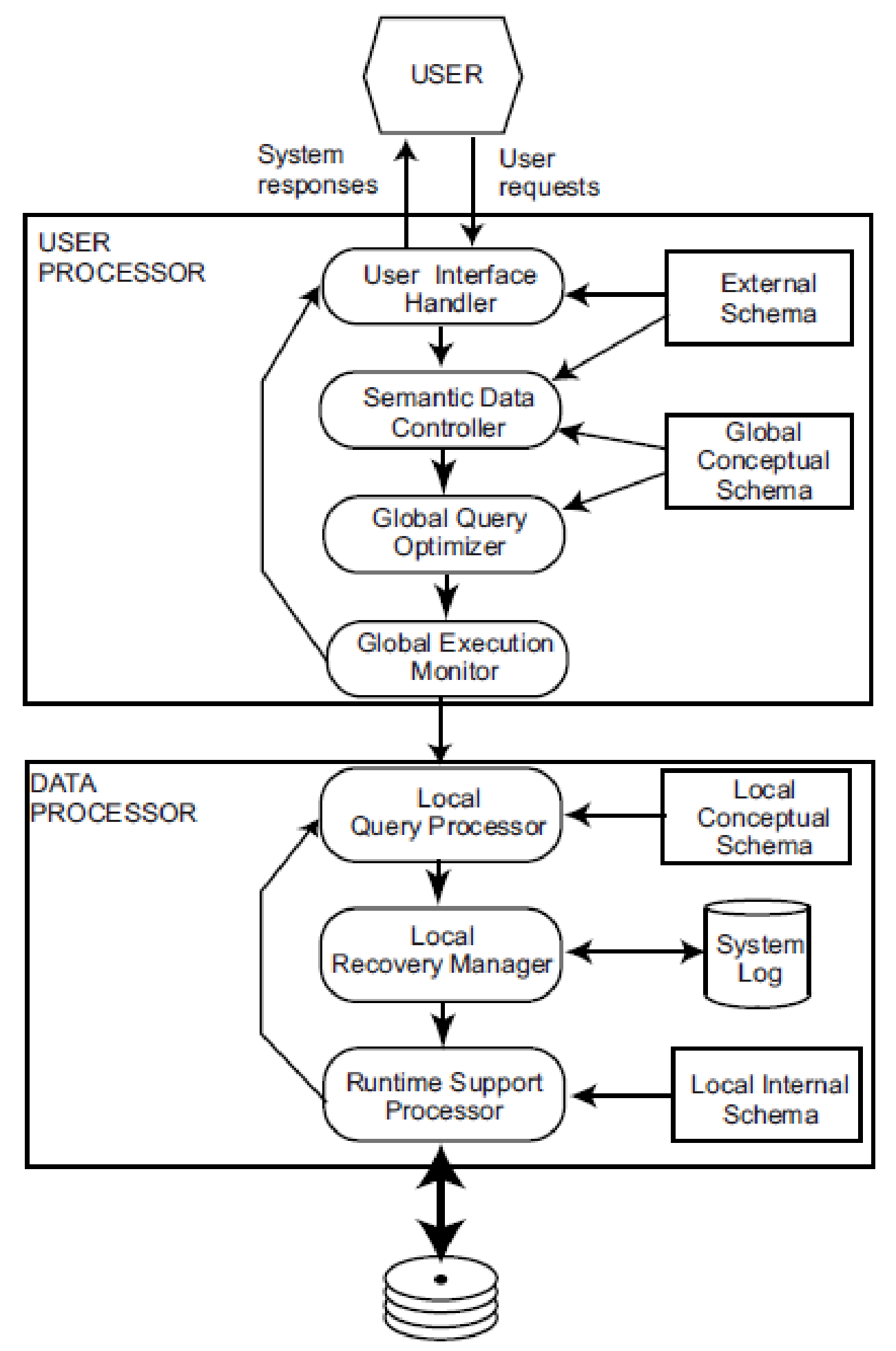
* + Ký hiệu là mảnh thứ i của quan hệ R trên địa điểm (site) j.  
    Ví dụ quan hệ toàn cục R được phân làm 3 mảnh con và được cấp phát cho 4 địa điểm:
    - Tại địa điểm 1 (site 1): Có 1 bản sao và . Kí hiệu: và
    - Địa điểm 2 và 3 tương tự.
* **Lược đồ ánh xạ cục bộ:** ánh xạ các ảnh vật lý đến các đối tượng được thao tác bởi các hệ quản trị cơ sở dữ liệu cục bộ/ địa phương. Ánh xạ này được gọi là lược đồ ánh xạ cục bộ

# **Kiến trúc hệ CSDL phân tán**

## ***Kiến trúc ngang hàng (DDBS Peer-to-Peer Architecture)***



* Cách thành phần kiến trúc ngang hàng (Component Peer-to-Peer Architecture):

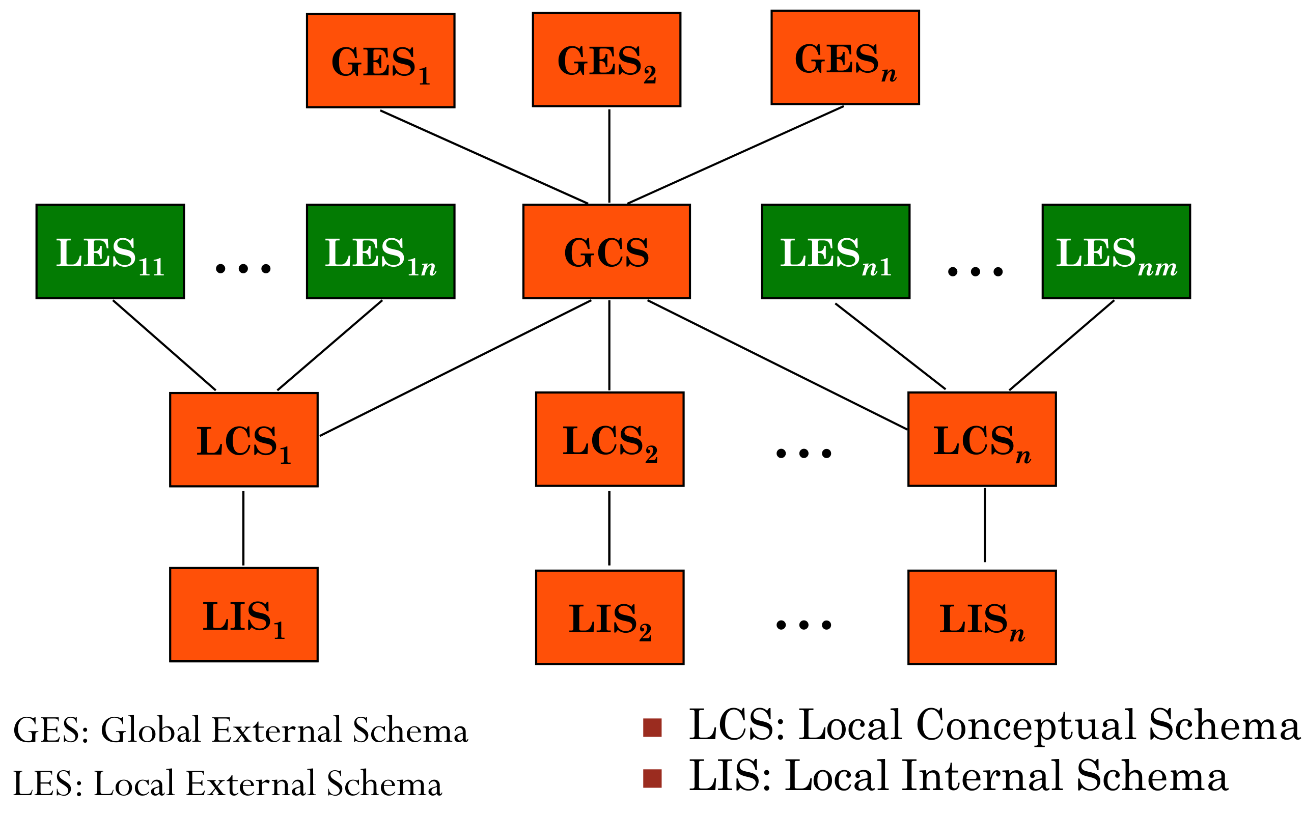


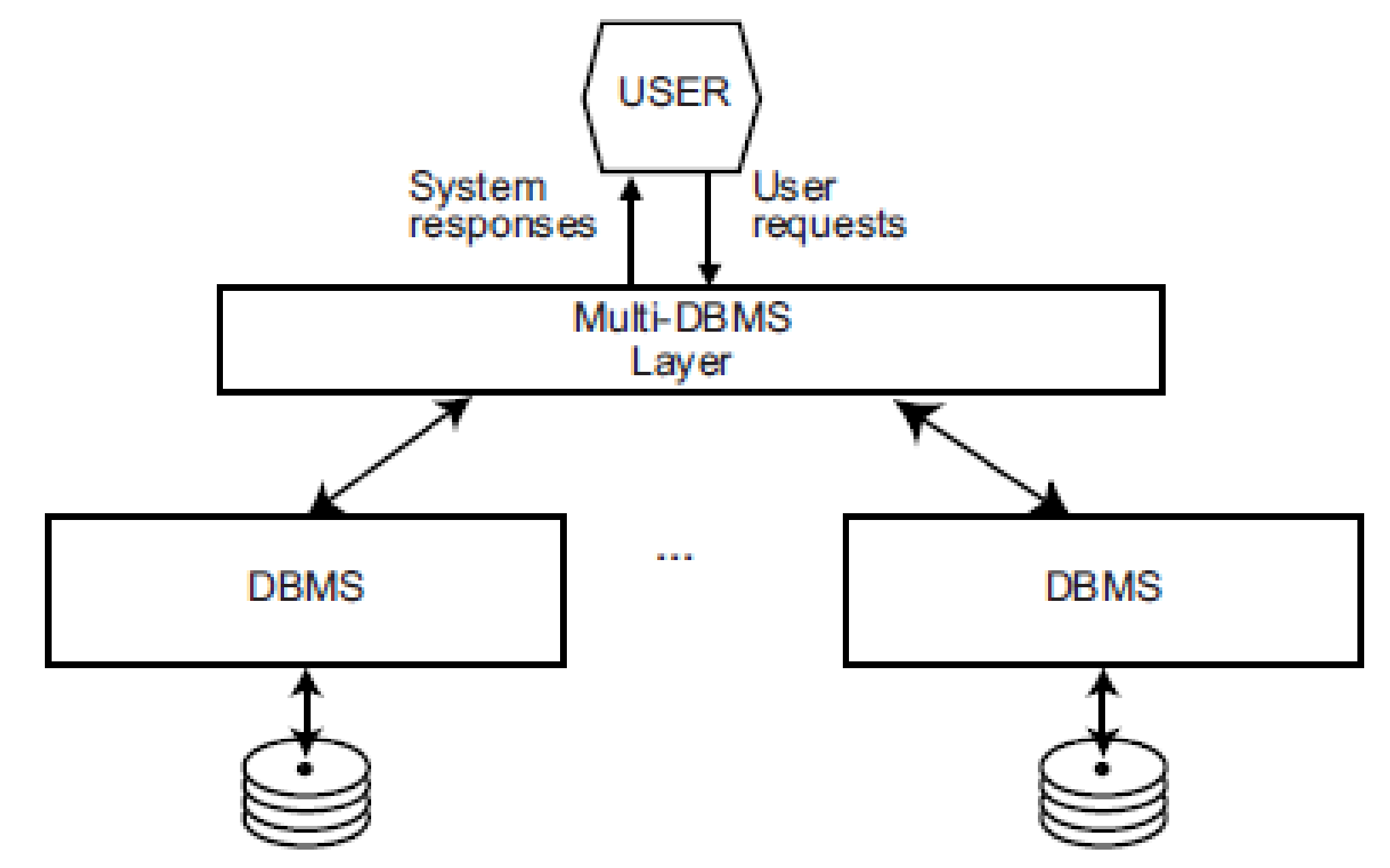
* + Bộ phận giao tiếp (user interface handler): dịch các câu lệnh của người sử dụng và định dạng dữ liệu kết quả để chuyển cho người sử dụng.
  + Bộ phận kiểm soát dữ liệu ngữ nghĩa (semantic data controller): sử dụng các ràng buộc toàn vẹn và thông tin quyền hạn, được định nghĩa như thành phần của lược đồ quan niệm toàn cục để kiểm tra xem các câu truy vấn có thể xử lý được hay không.
  + Bộ phận phân rã và tối ưu hoá vấn tin toàn cục (global query optimizer and decomposer): xác định như một chiến lược hoạt động nhằm giảm thiểu chi phí, phiên dịch các câu vấn tin toàn cục thành các câu vấn tin cục bộ bằng cách sử dụng các lược đồ quan niệm toàn cục, lược đồ quan niệm cục bộ và thư mục toàn cục. Bộ phận tối ưu vấn tin toàn cục còn chịu trách nhiệm tạo ra một chiến lược thực thi tốt nhất cho phép nối phân tán.
  + Bộ phận giám sát hoạt động phân tán (distributed execution monitor): điều phối việc thực hiện phân tán các yêu cầu người sử dụng và cũng được gọi là bộ quản lý giao tác phân tán (distributed transaction manager).

\*\*\*

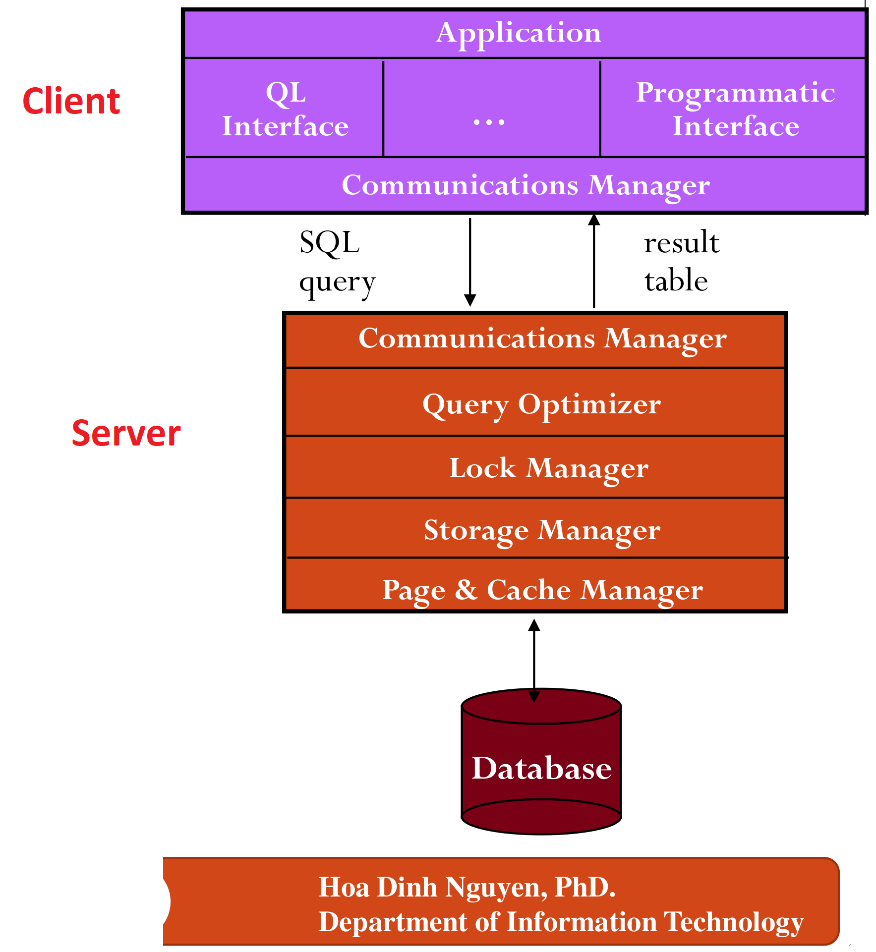
* + Bộ phận tối ưu hoá vấn tin cục bộ (local query optimizer): thường hoạt động như bộ chọn đường truy xuất, chịu trách nhiệm chọn ra một đường truy xuất thích hợp nhất để truy xuất các mục dữ liệu.
  + Bộ phận khôi phục cục bộ (local recovery manager) bảo đảm cho các cơ sở dữ liệu cục bộ vẫn duy trì được tính nhất quán ngay cả khi có sự cố xảy ra.
  + Bộ phận hỗ trợ lúc thực thi (run-time support processor): truy xuất cơ sở dữ liệu tùy thuộc vào các lệnh trong lịch biểu do bộ phận tối ưu vấn tin tạo ra. Nó chính là bộ giao tiếp với hệ điều hành và chứa bộ quản lý vùng đệm cơ sở dữ liệu, chịu trách nhiệm quản lý vùng đệm và quản lý việc truy xuất dữ liệu

## ***II. Kiến trúc Hệ thống quản trị CSDL phân tán (DDBMS Architecture)***

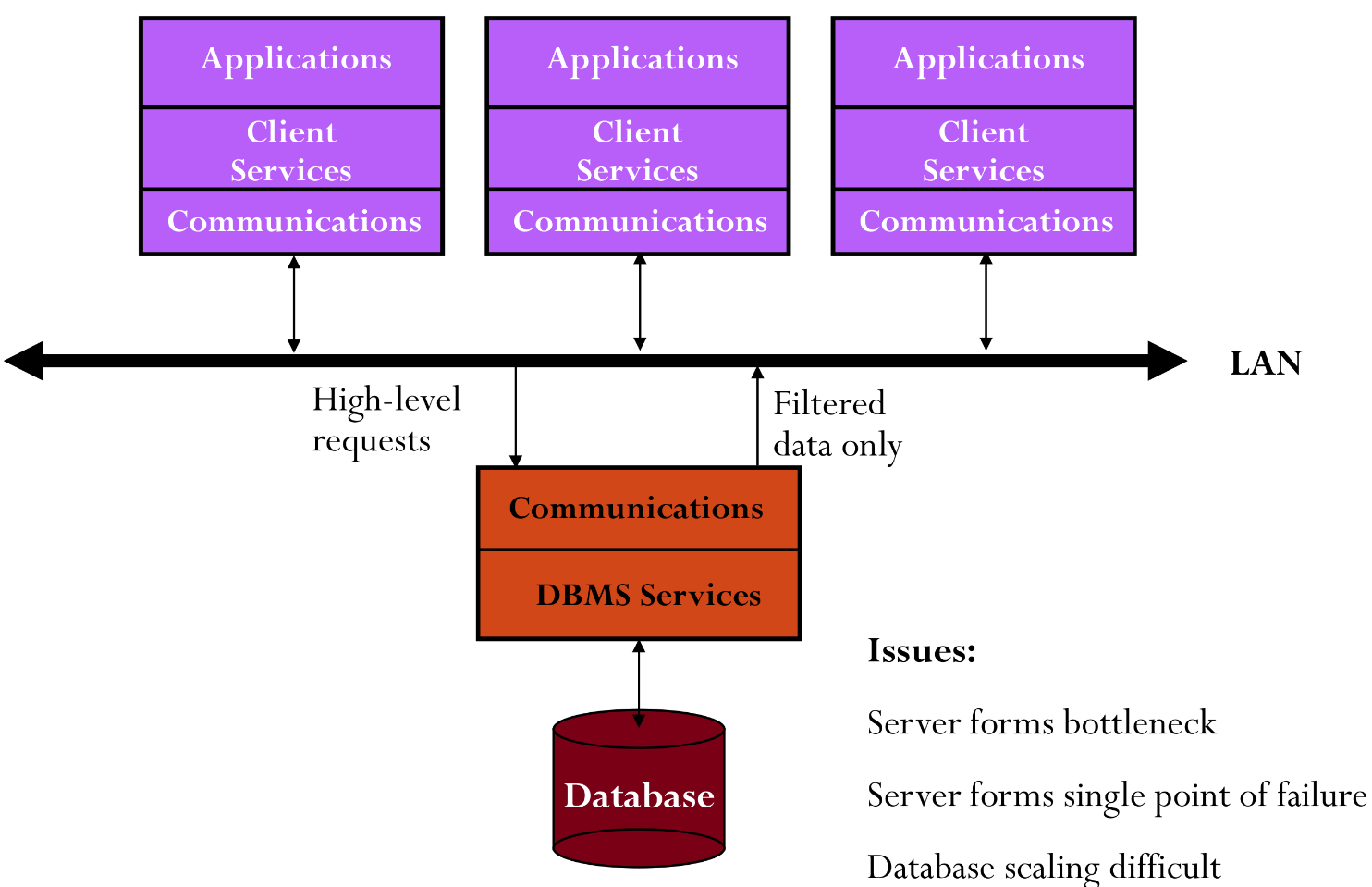


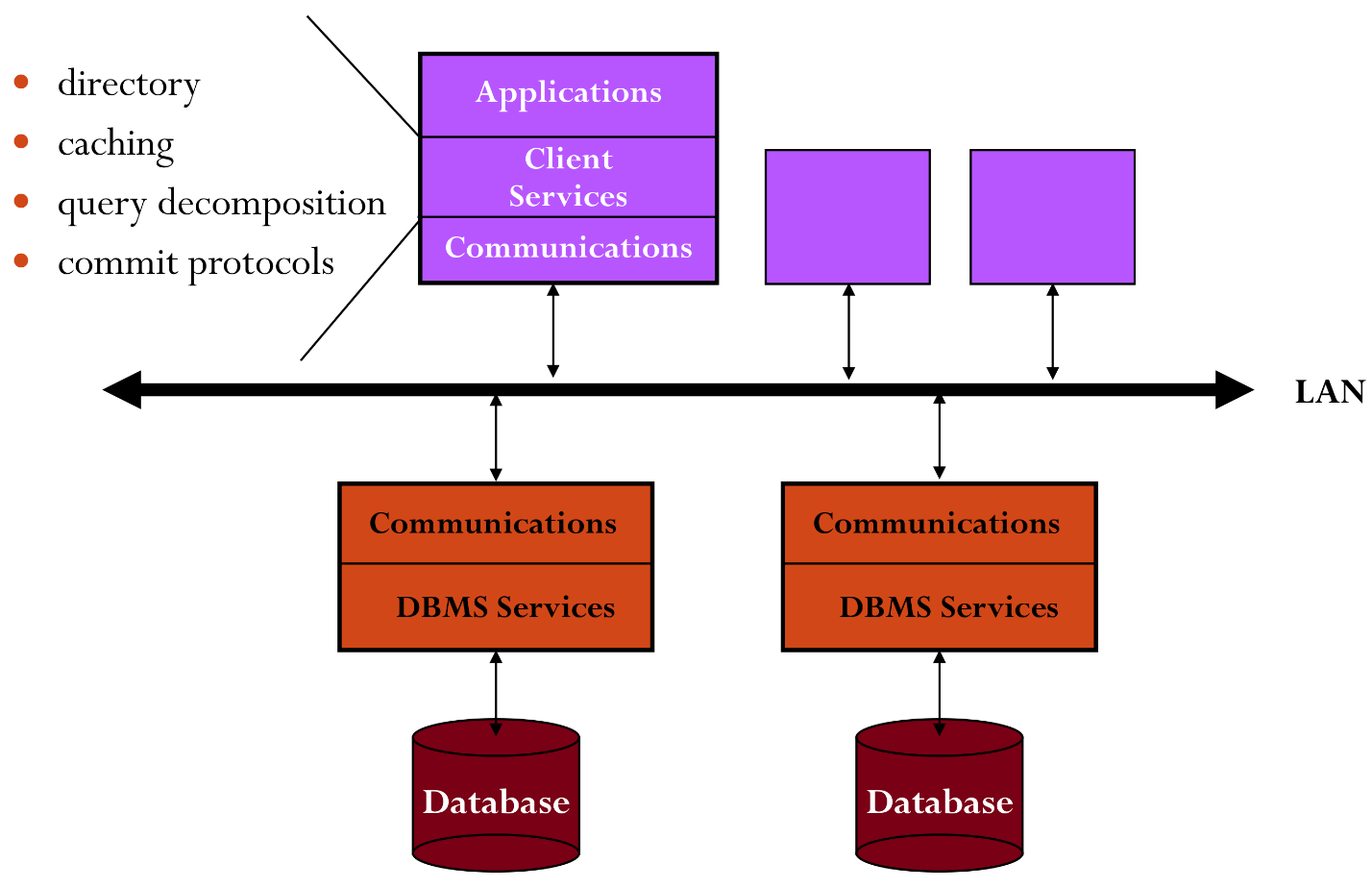


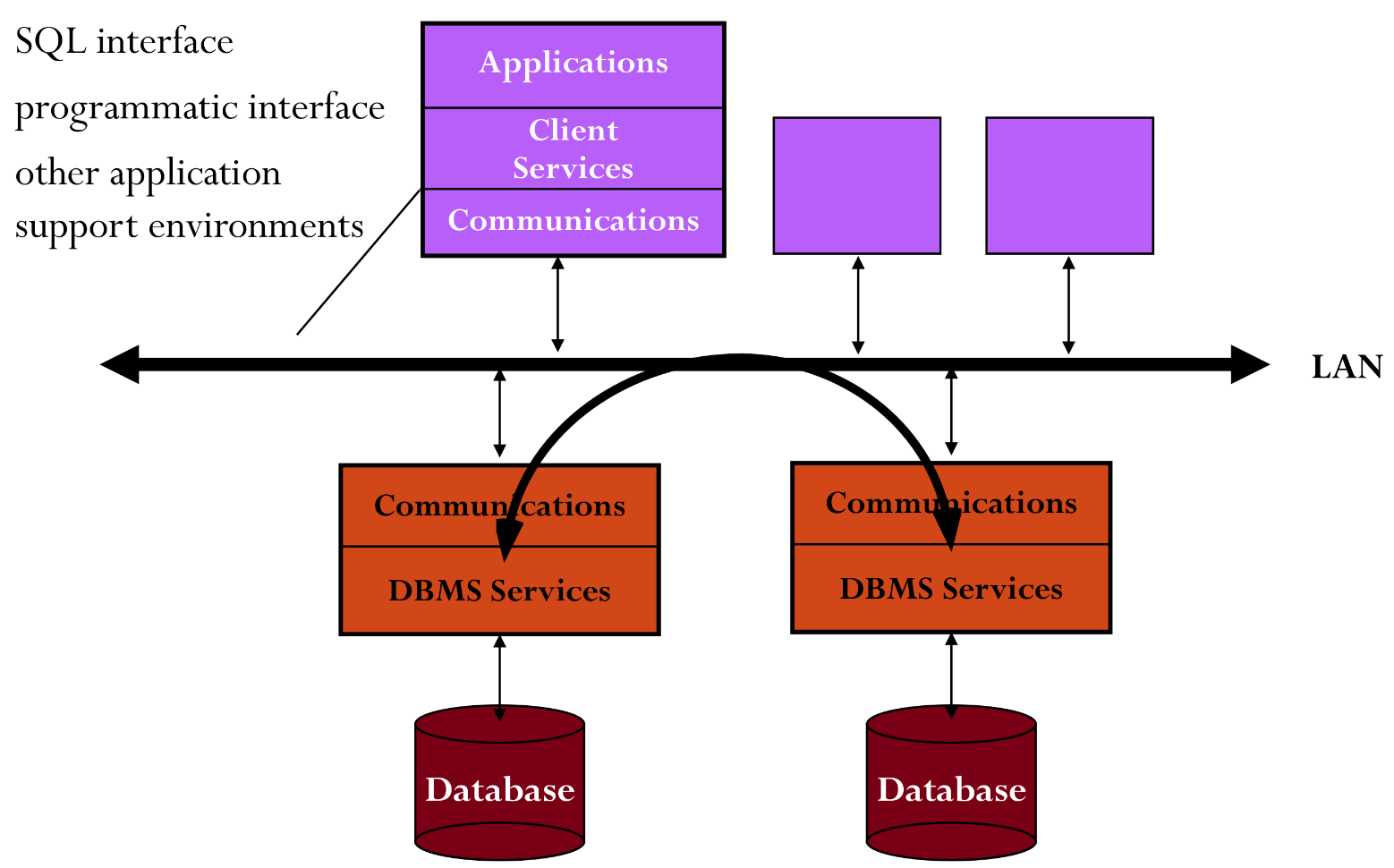
## ***III. Kiến trúc Client/Server***



* Cung cấp một kiến trúc hai mức, dễ dàng cho việc quản lý mức độ phức tạp của các hệ quản trị cơ sở dữ liệu hiện đại và độ phức tạp của việc phân tán dữ liệu.
* Máy chủ thực hiện phần lớn các công việc quản lý dữ liệu. Điều này có nghĩa là mọi việc xử lý và tối ưu hóa vấn tin, quản lý giao tác và quản lý thiết bị lưu trữ đều được thực hiện tại máy chủ.
* Khách, ngoài giao diện và ứng dụng, sẽ có một module quản trị cơ sở dữ liệu, khách chịu trách nhiệm quản lý dữ liệu được gửi đến.
* **Đa người dùng – đơn server:**

****

****

****