



# HỆ CƠ SỞ DỮ LIỆU

GV: ThS.Trịnh Thị Ngọc Linh





# CHƯƠNG 9. HỆ CƠ SỞ DỮ LIỆU PHÂN TÁN

**1**

**Khái niệm và kiến trúc**

**2**

**Ưu và nhược điểm**

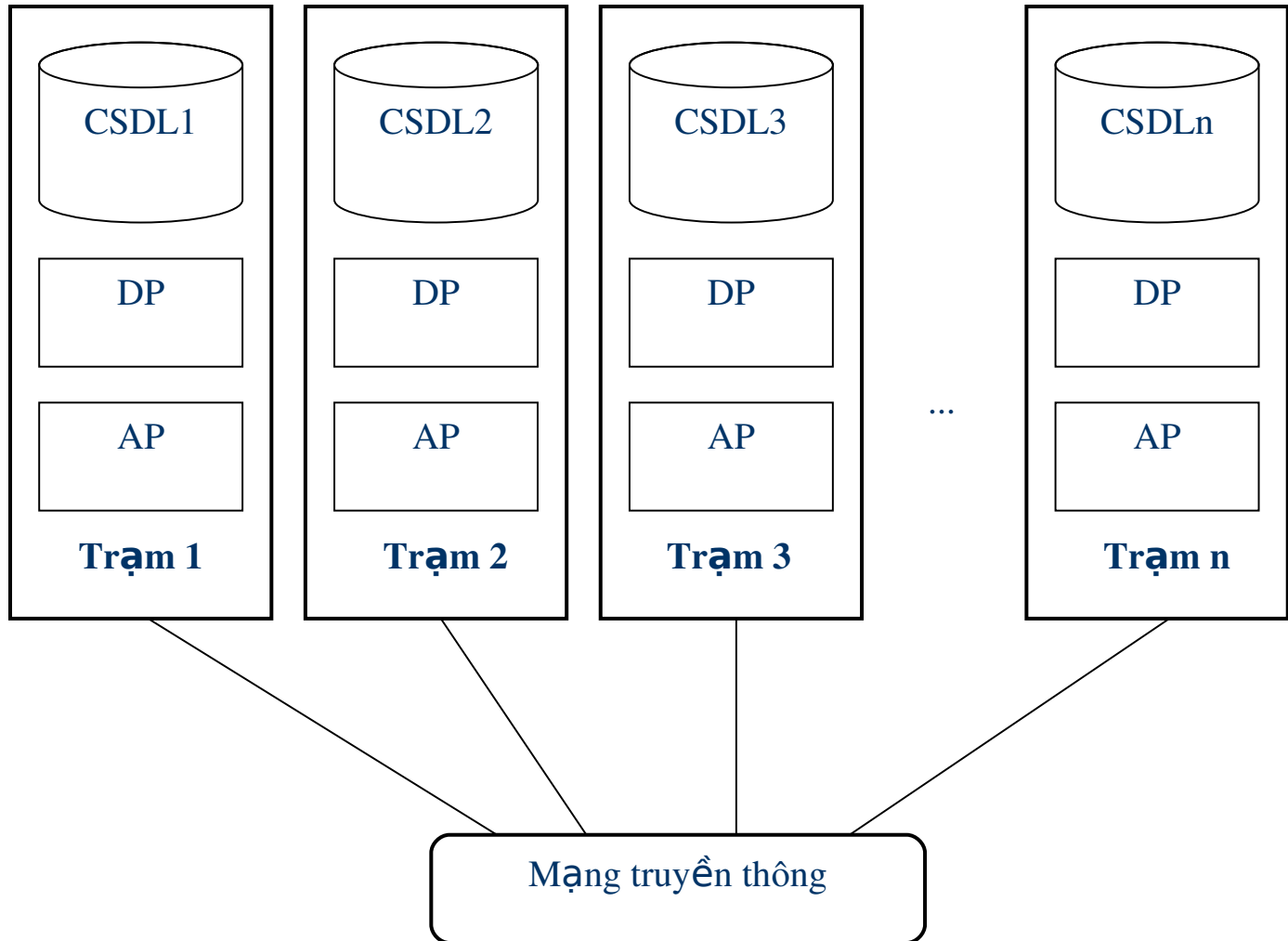
**3**

**Quản lý giao dịch phân tán**

- ❖ Một CSDL phân tán là một tập các CSDL cùng hợp tác hoạt động, được lưu trữ trên các máy tính khác nhau, gọi là các trạm (sites), và được nối với nhau bằng một mạng truyền thông. NSD tại một trạm bất kỳ có thể truy cập và thao tác dữ liệu trên các trạm khác.
- ❖ Hệ quản trị CSDL phân tán là một tập các phần mềm hệ thống hỗ trợ việc tạo lập và bảo trì các CSDL phân tán
- ❖ Phân loại
  - CSDL phân tán thuần nhất
  - CSDL phân tán không thuần nhất



# Kiến trúc





- ❖ **Để quản lý được tính phức tạp của một Hệ CSDL phân tán, phần mềm tại các trạm được chia làm ba mô đun chính:**
  - Phần mềm xử lý dữ liệu (DP - Data Processor): Chịu trách nhiệm quản lý dữ liệu cục bộ (địa phương) tại một trạm, giống với phần mềm CSDL tập trung
  - Phần mềm xử lý ứng dụng (AP - Application Processor): Chịu trách nhiệm với phần lớn chức năng phân tán, nó truy cập thông tin phân tán từ thư mục CSDL phân tán và chịu trách nhiệm xử lý mọi yêu cầu đòi hỏi truy cập đến nhiều hơn một trạm
  - Phần mềm truyền thông: Cung cấp các nguyên thủy truyền thông được AP dùng để truyền lệnh và dữ liệu giữa các trạm khác nhau khi cần



## Các khái niệm

- ❖ **CSDL từ xa:** Là một CSDL ở trên một máy tính khác với máy tính của người dùng và được truy cập nhờ vào các lệnh truyền thông được xác định bởi người dùng
- ❖ **Một CSDL liên hiệp:** Không nhất thiết phải phân tán trên nhiều máy và cách tiếp cận này nhằm ưu tiên người dùng thao tác dữ liệu trên CSDL riêng của người đó và xem việc truy cập đến các CSDL khác chỉ là đột xuất
- ❖ **Một CSDL song song:** có thể hiểu là một CSDL phân tán thuần nhất, trong đó các trạm là các nút của một máy tính song song và trao đổi/liên lạc với nhau bằng các thông báo



# Các ưu điểm của việc phân tán dữ liệu

- ❖ **Lợi điểm về tổ chức và tính kinh tế:** Việc dùng CSDL phân tán phù hợp đối với tổ chức phân tán thành nhiều chi nhánh
- ❖ **Tận dụng những cơ sở dữ liệu sẵn có:** Có thể hình thành CSDL phân tán từ các CSDL tập trung có sẵn
- ❖ **Thuận lợi cho nhu cầu phát triển:** Xu hướng dùng cơ sở dữ liệu phân tán sẽ cung cấp khả năng phát triển thuận lợi hơn, giảm được xung đột về chức năng giữa các đơn vị đã tồn tại, và giảm xung đột giữa các chương trình ứng dụng khi truy cập đến CSDL
- ❖ **Giảm chi phí truyền thông:** Trong CSDL phân tán, chương trình ứng dụng có thể giảm bớt được chi phí truyền thông khi thực hiện bằng cách khai thác CSDL tại chỗ
- ❖ **Tăng số công việc thực hiện:** Hệ CSDL phân tán có thể tăng số lượng công việc thực hiện qua áp dụng nguyên lý xử lý song song với hệ thống xử lý đa nhiệm.
- ❖ **Tính dễ hiểu và sẵn sàng:** Hướng phát triển CSDL phân tán cũng nhằm đạt được tính dễ hiểu và tính sẵn sàng cao hơn. Tuy nhiên để đạt được mục tiêu này không phải là dễ làm và đòi hỏi sử dụng kỹ thuật phức tạp. Khả năng xử lý tự trị của các điểm làm việc khác nhau không đảm bảo tính dễ sử dụng



# Các nhược điểm của việc phân tán dữ liệu

- ❖ **Giá phát triển phần mềm:** Do các khó khăn khi cài đặt một hệ thống phân tán, giá thành sẽ tăng lên
- ❖ **Dễ mắc lỗi hơn:** Vì các trạm trong hệ phân tán làm việc song song, khó có thể đảm bảo thuật toán được thực hiện đúng trên tất cả các trạm. Do vậy mà số lỗi sẽ tăng lên. Các thuật toán phù hợp với CSDL phân tán còn đang được nghiên cứu
- ❖ **Khối lượng các xử lý tăng:** Hệ thống phân tán cần truyền nhiều thông báo, nhiều tính toán phụ. Do vậy khối lượng xử lý tăng lên so với hệ thống tập trung. Khi chọn thiết kế hệ thống CSDL, người thiết kế cần phải cân đối các ưu điểm và nhược điểm của việc phân tán các dữ liệu. Có dải rộng các thiết kế rất phân tán đến các hệ thống tập trung





# Quản lý giao dịch phân tán

## ❖ Điều khiển tương tranh

- Thời dấu: Thời dấu bao gồm việc sắp xếp các giao dịch phân tán khi cho thực hiện chúng và áp đặt các thao tác truy cập tới dữ liệu tôn trọng thứ tự được xác định trước
- Sự đặt khóa: Việc truy cập tới các hạt dữ liệu được chi phối bởi một giao thức nói chung bao gồm hai thao tác nguyên thủy:
  - LOCK: Yêu cầu thao tác trên các hạt dữ liệu theo một cách thức nhất định.
  - UNLOCK: báo hiệu sự kết thúc của thao tác trên hạt

## ❖ Sự hợp thức hoá giao dịch: tích hợp thực sự các cập nhật của nó vào CSDL