

Hệ quản trị cơ sở dữ liệu

Database Management System

# Giới thiệu

- Tên môn học: **Hệ quản trị cơ sở dữ liệu**  
(**D**atabase **M**anagement **S**ystem)
- Số tín chỉ: 3
  - Lý thuyết: 30 tiết
  - Thực hành: 30 tiết

# Nội dung

Giới thiệu DBMS

Lưu trữ dữ liệu

Tối ưu truy vấn

Quản lý giao tác

Khôi phục sự cố

# Đánh giá

- Điểm học phần = 40% điểm TX + 60% điểm thi
- **Đánh giá thường xuyên:**
  - Chuyên cần
  - Kiểm tra lý thuyết
  - Làm bài thực hành
- **Thi cuối kỳ:** viết (*không sử dụng tài liệu*)

# Tài liệu học tập

- Bài giảng và bài tập:

[sites.google.com/site/aguims301](https://sites.google.com/site/aguims301)

hoặc

[courses.agu.edu.vn](https://courses.agu.edu.vn)

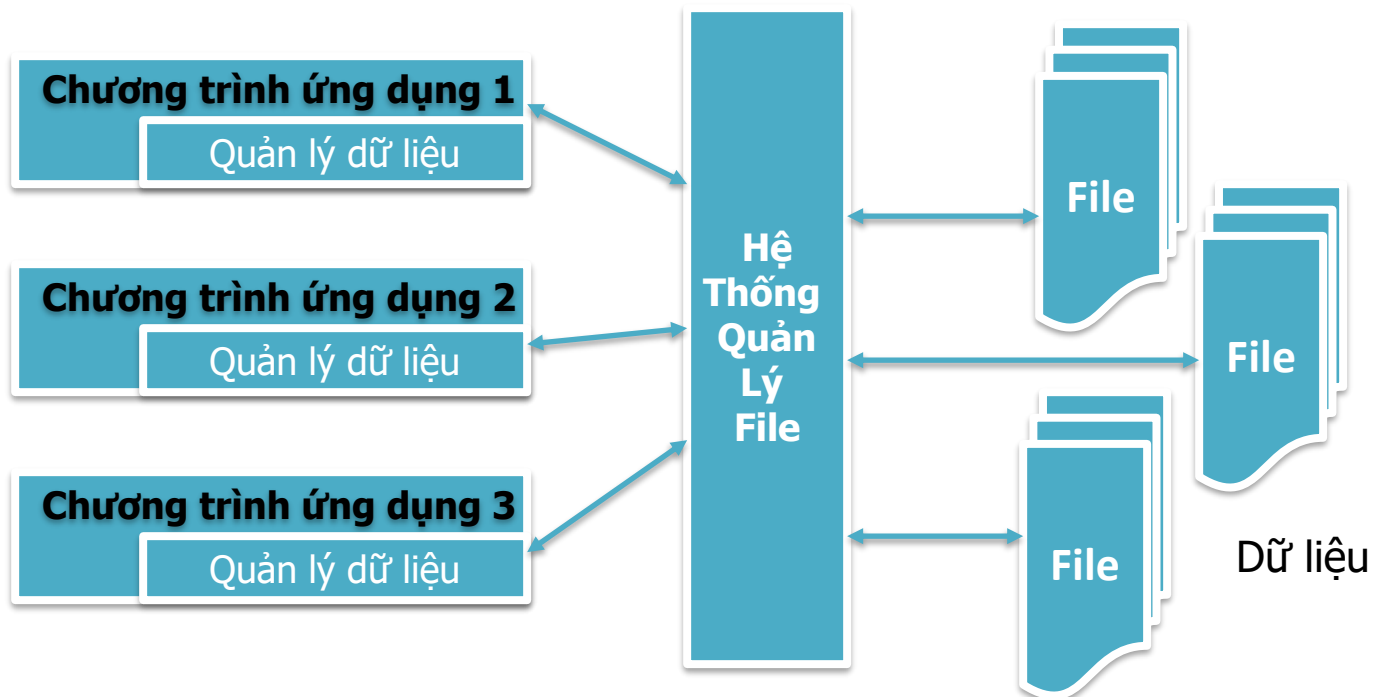
- Thực hành: SQL Server
- Khuyến khích SV đọc thêm các tài liệu khác có liên quan

# Các hình thức lưu trữ dữ liệu



# Hệ thống tập tin

- Dữ liệu của ứng dụng được lưu trữ trên những tập tin dữ liệu riêng



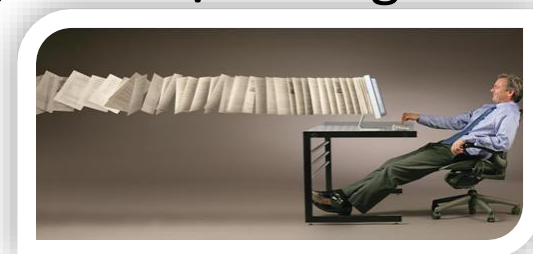
# Hệ thống tập tin

## ■ Ưu điểm:

- Gọn nhẹ, phù hợp thực tiễn. Ít tốn thời gian, chi phí thấp
- Khả năng đáp ứng khai thác nhanh chóng và kịp thời

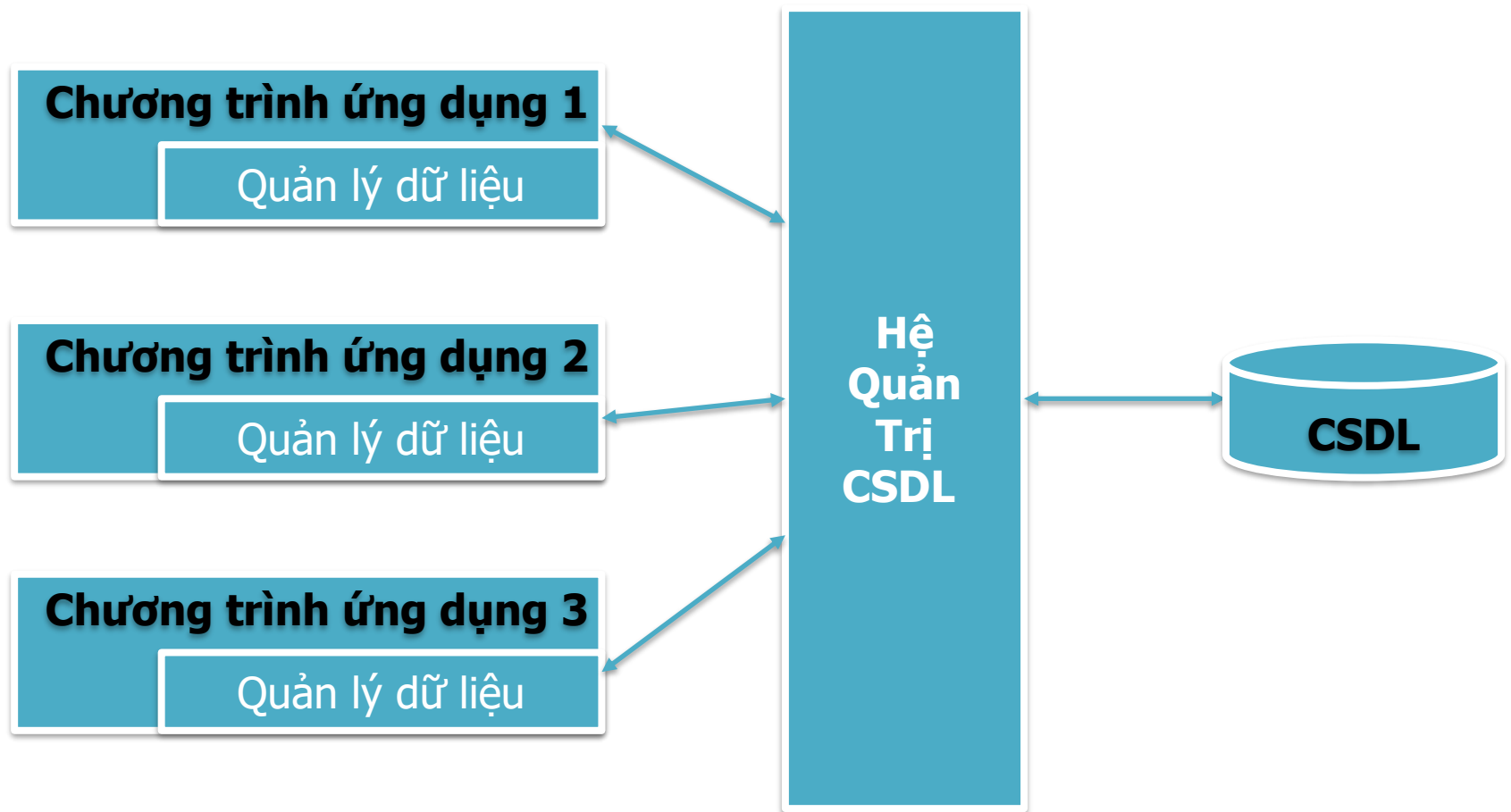
## ■ Nhược điểm:

- Thông tin lưu nhiều nơi, dư thừa, không nhất quán
- Lãng phí thời gian cập nhật dữ liệu và lưu trữ
- Phối hợp tổ chức và khai thác là khó khăn
- Thiếu sự chia sẻ thông tin giữa các đơn vị và bộ phận
- Khó khăn khi nâng cấp ứng dụng
- Không có người quản trị dữ liệu, mọi người có quyền sử dụng thêm, xóa, sửa không an toàn, không bảo mật thông tin





# Cơ sở dữ liệu

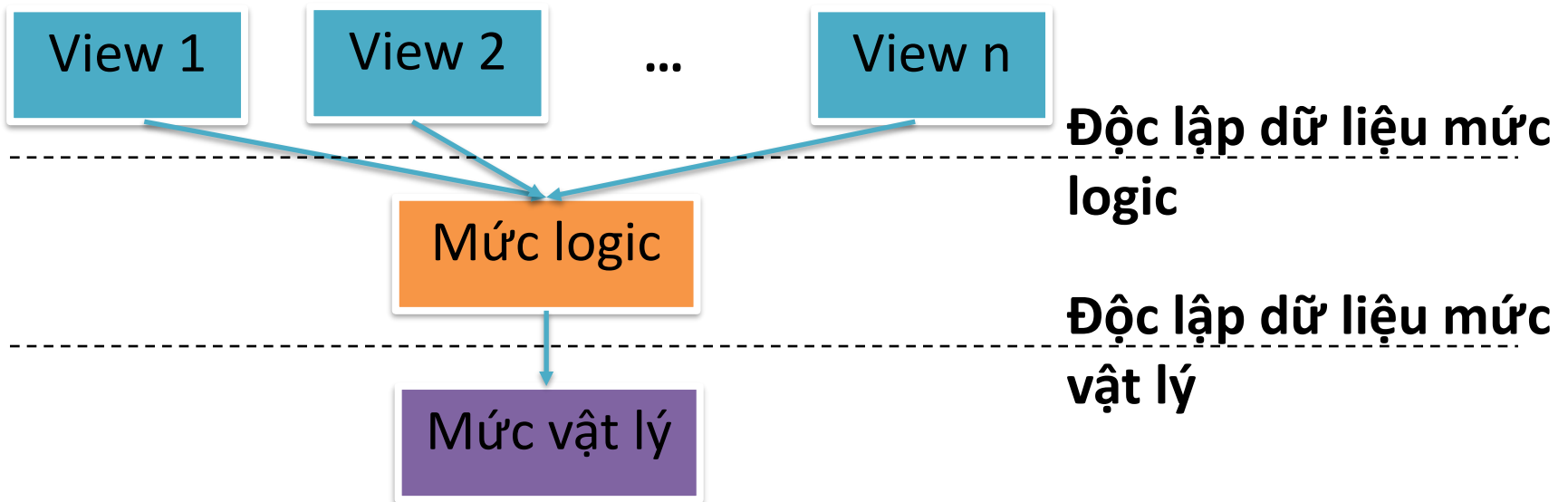


# Cơ sở dữ liệu

Các đặc trưng:

- Ít dư thừa dữ liệu
- Chia sẻ cho nhiều người sử dụng
- Có tính an toàn, bảo mật
- Cho phép truy xuất đồng thời
- Có thể khôi phục khi có sự cố
- Độc lập vật lý và logic

# Các mức trừu tượng dữ liệu



# Các mức trừu tượng dữ liệu

- Mức khung nhìn (View Level)
  - Mô tả một phần nào đó trong hệ thống
- Mức logic (Logical View)
  - Những thông tin gì được lưu trữ trong CSDL
  - Có những mối quan hệ nào giữa những thông tin đó
- Mức vật lý (Physical Level)
  - Dữ liệu được lưu trữ như thế nào

# Tổ chức, quản lý dữ liệu

TỔ CHỨC DỮ LIỆU	HỆ QUẢN LÝ DỮ LIỆU
Hệ thống tập tin (File system)	Hệ điều hành (Operating system)
Cơ sở dữ liệu (Database)	Hệ quản trị cơ sở dữ liệu (Database Management System)

# Tổng quan về HỆ QUẢN TRỊ CƠ SỞ DỮ LIỆU

# Tổng quan HQT CSDL

- Định nghĩa HQT CSDL
- Các thành phần của HQT CSDL
- Phân loại HQT CSDL
- Lợi ích từ HQT CSDL

# Tổng quan HQT CSDL

- Định nghĩa HQT CSDL
- Các thành phần của HQT CSDL
- Phân loại HQT CSDL
- Lợi ích từ HQT CSDL



# Hệ quản trị cơ sở dữ liệu

- Là hệ thống phần mềm cung cấp các công cụ để xây dựng và quản lý CSDL
  - Định nghĩa cấu trúc dữ liệu
  - Cung cấp khả năng thao tác trên CSDL
  - Hỗ trợ lưu trữ dữ liệu
  - Điều khiển truy xuất dữ liệu giữa nhiều người dùng

# Hệ quản trị cơ sở dữ liệu

- Một số HQT CSDL nổi bật
  - MS-SQL Server
  - ORACLE
  - DB2



ORACLE®



SYBASE®  
An SAP Company



INGRES®



HyperSQL



Informix®



PROGRESS



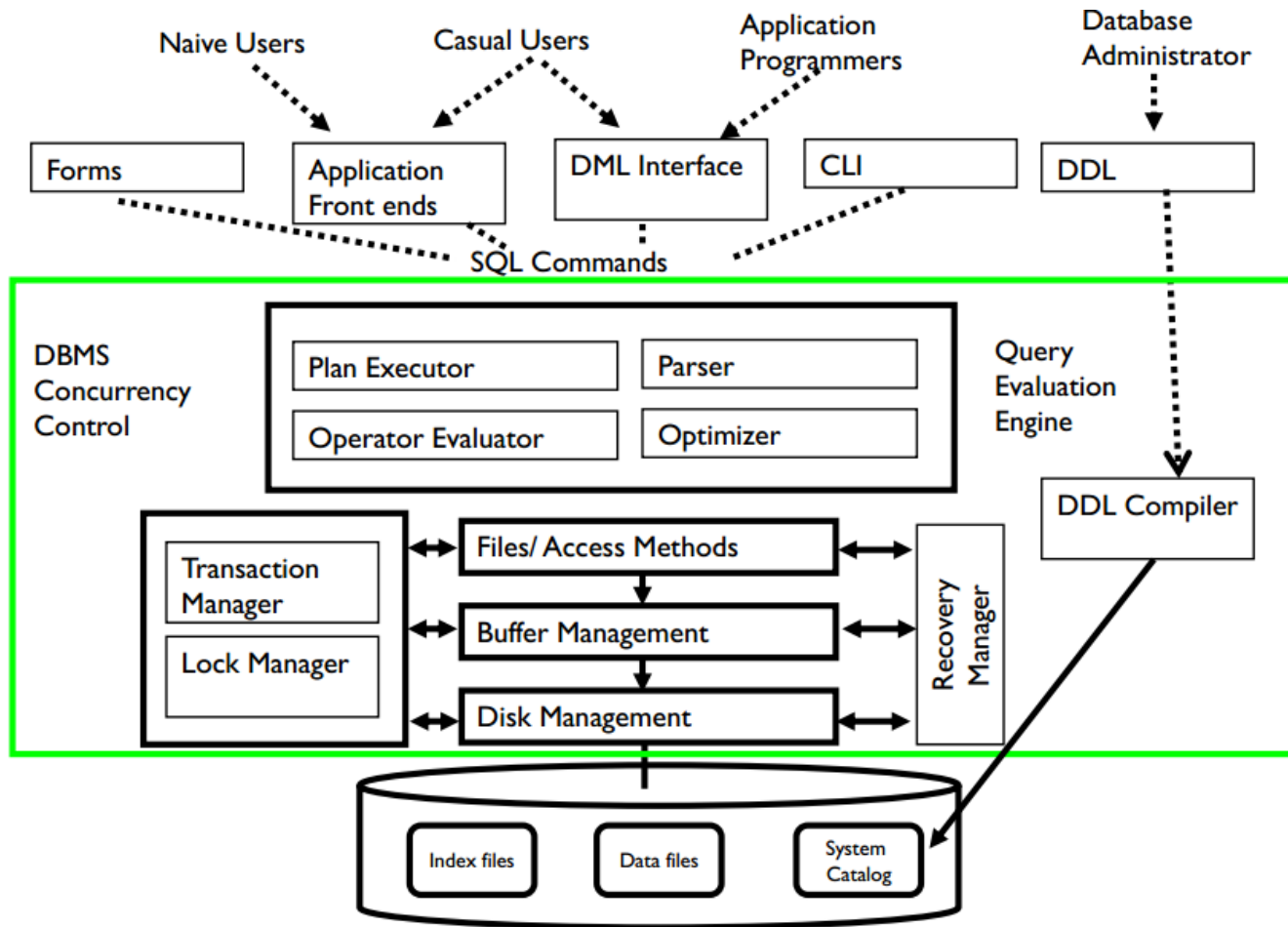
# Lịch sử phát triển

- Từ giữa 1960s: mô hình mạng và phân cấp, dùng file
  - CODASYL
- Cuối 1970s: mô hình dữ liệu quan hệ
  - SYSTEM-R, DB2, Informix, Sysbase, Oracle
- 1980s: mô hình dữ liệu hướng đối tượng
  - ObjectStone, O2
- 1990s: XML

# Tổng quan HQT CSDL

- Định nghĩa HQT CSDL
- Các thành phần của HQT CSDL
- Phân loại HQT CSDL
- Lợi ích từ HQT CSDL

# Kiến trúc DBMS



# Kiến trúc DBMS

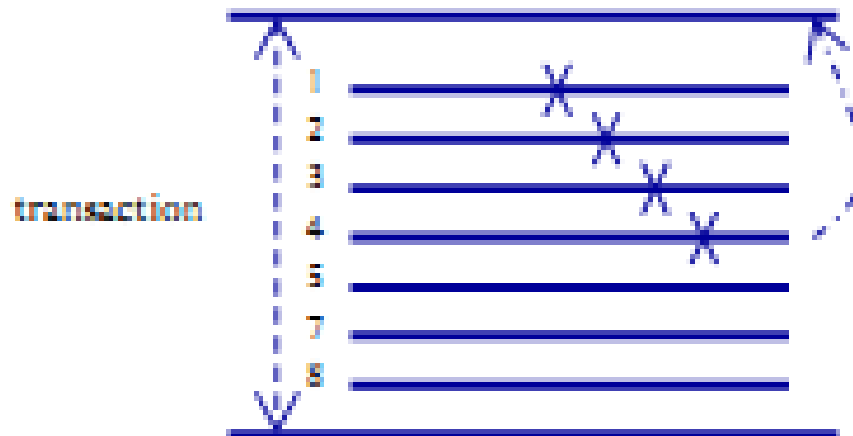
- Các thành phần chính của DBMS:
  - Ngôn ngữ giao tiếp
  - Quản lý giao tác
  - Xử lý truy vấn
  - Quản lý lưu trữ
  - Quản lý khôi phục

# Ngôn ngữ giao tiếp

- DBMS cung cấp giao diện lập trình để sử dụng với một ngôn ngữ lập trình CSDL
  - SQL Server: Transaction-SQL (T-SQL)
  - Oracle: PL/SQL
- Ngôn ngữ bao gồm
  - Định nghĩa dữ liệu
  - Thao tác dữ liệu

# Quản lý giao tác

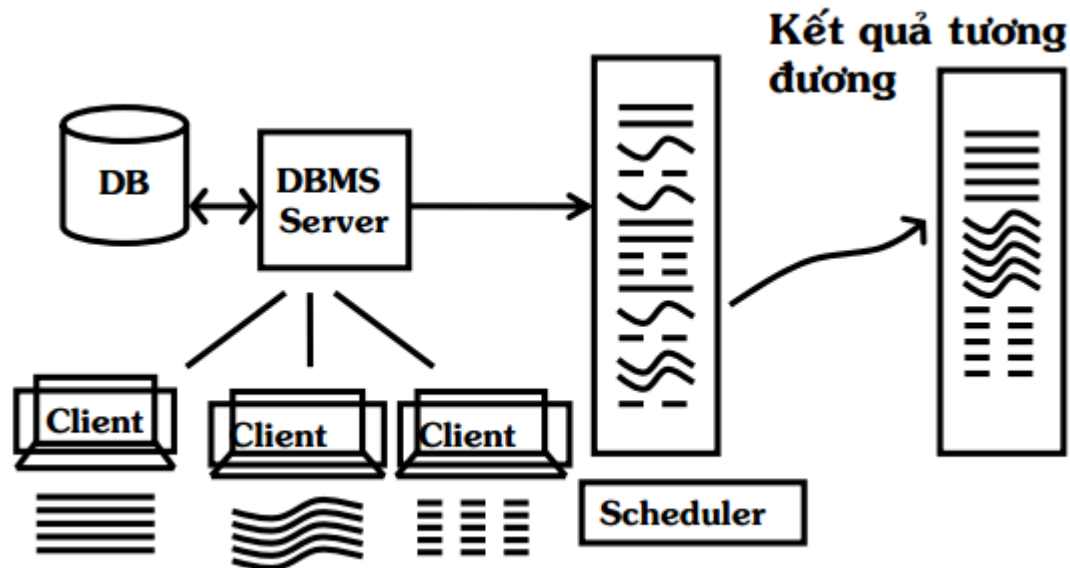
- Thành phần quản lý các giao tác có ảnh hưởng đến CSDL
  - Giao tác là một nhóm các hành động mà nếu thực hiện được thì phải thực hiện hết tất cả các hành động trong giao tác đó, ngược lại xem như không thực hiện hành động nào





# Quản lý giao tác

- Điều khiển đồng thời
  - Bộ lập lịch (scheduler) – có nhiệm vụ lập 1 lịch thực hiện từ n giao tác được kích hoạt đồng thời
  - Cơ chế khóa (lock) – ngăn 2 giao tác cùng thao tác lên cùng 1 đơn vị dữ liệu tại một thời điểm



# Quản lý khôi phục

- Để CSDL được bền vững, mọi thay đổi lên CSDL phải được ghi nhận lại
- Log manager – ghi chép nhật ký
  - Đảm bảo CSDL vẫn nguyên vẹn khi có sự cố xảy ra.
- Recovery manager – khôi phục
  - Dựa vào nhật ký để phục hồi lại CSDL về trạng thái nhất quán trước đó (đảm bảo tất cả các RBTV của CSDL đó).

# Xử lý truy vấn

- Biểu diễn câu truy vấn ở dạng ngôn ngữ cấp cao (SQL) và thực hiện câu truy vấn có hiệu quả
- Query compiler – biên dịch
  - Query parser
    - Xây dựng cấu trúc hình cây từ câu truy vấn
  - Query preprocessor
    - Kiểm tra ngữ nghĩa của câu truy vấn
    - Chuyển đổi cấu trúc cây sang ngôn ngữ đại số quan hệ
  - Query optimizer
    - Sắp xếp các phép toán nhằm mục đích tối ưu hóa câu truy vấn

# Tổng quan HQT CSDL

- Định nghĩa HQT CSDL
- Các thành phần của HQT CSDL
- Phân loại HQT CSDL
- Lợi ích từ HQT CSDL

# Phân loại DBMS

- Theo mô hình dữ liệu
  - Mạng (network)
  - Phân cấp (Hierarchical)
  - Quan hệ (Relation)
  - Hướng đối tượng (Oriented-Object)
  - XML
- Theo kiến trúc
  - Đơn người dùng /Client – Server
  - Tập trung / Phân tán
- Khác
  - Multimedia
  - Hypertext

# Tổng quan HQT CSDL

- Định nghĩa HQT CSDL
- Các thành phần của HQT CSDL
- Phân loại HQT CSDL
- Lợi ích từ HQT CSDL

# Thuận lợi khi sử dụng DBMS

- Chống được việc dư thừa dữ liệu trong lưu trữ
- Kiểm soát được việc truy nhập dữ liệu trái phép
- Cung cấp những cách chứa dữ liệu hiệu quả tăng tốc độ truy xuất
- Cung cấp hệ thống lưu trữ và phục hồi dữ liệu
- Cung cấp nhiều giao diện
- Giúp biểu diễn những quan hệ phức tạp của dữ liệu
- Bảo đảm được tính thống nhất và không mâu thuẫn của CSDL thông qua các ràng buộc
- Chia sẻ dữ liệu

# Thuận lợi khi sử dụng DBMS

- Tăng tính chuẩn hóa của cả hệ thống
- Giảm thời gian phát triển các ứng dụng
- Dễ dàng thay đổi cấu trúc dữ liệu
- Cập nhật thông tin nhanh
- Lợi ích về kinh tế:
  - Ngăn chặn được sự chồng chéo giữa các nguồn nhân lực và tài nguyên bằng việc dùng một CSDL chung giữa các bộ phận



# Tuy nhiên

- DBMS đòi hỏi cấu hình và giá cả nhất định:
  - Đầu tư ban đầu và các phần cứng bổ sung.
  - Để có một DBMS ưng ý tốn khá nhiều tiền.
- DBMS có thể không cần thiết trong các trường hợp:
  - Dữ liệu đơn giản, định nghĩa rõ ràng, ít thay đổi.
  - Nếu yếu tố real-time là vô cùng quan trọng (Game: Counter Strike).
  - Nếu việc truy xuất đồng thời là không cần thiết.
- DBMS có thể không đáp ứng được trong các trường hợp:
  - Mô hình dữ liệu quá phức tạp.
  - Có những bước thực hiện đặc biệt mà DBMS không đáp ứng được.

# Ai sử dụng HQT CSDL?

- **Thiết kế viên cơ sở dữ liệu** (Database Designer)
- **Quản trị viên cơ sở dữ liệu** (Database Administrators – DBA)
- **Người dùng cuối** (End Users)
  - ND thường xuyên
  - ND không thường xuyên
  - ND chuyên gia

# Câu hỏi

1. So sánh ưu điểm và nhược điểm của hệ thống quản lý tập tin truyền thống và hệ quản trị cơ sở dữ liệu.
2. Hệ quản trị cơ sở dữ liệu là gì? Nêu tên một số DBMS phổ biến hiện nay.
3. Trình bày tóm tắt chức năng các thành phần chính của DBMS.