



# HỆ CƠ SỞ DỮ LIỆU

GV: ThS.Trịnh Thị Ngọc Linh  
Email: [ngoclinhnl@yahoo.com](mailto:ngoclinhnl@yahoo.com)

# MỤC TIÊU

- Giới thiệu khái quát về lý thuyết cơ sở dữ liệu
- Cung cấp các kiến thức nền tảng về cơ sở dữ liệu để sinh viên có thể tiếp cận các môn học khác dễ dàng hơn

# NỘI DUNG

- Các hạn chế của hệ xử lý tập tin truyền thống và sự cần thiết của hệ cơ sở dữ liệu
- Các khái niệm chính về hệ cơ sở dữ liệu, kiến trúc 3 mức, các ngôn ngữ cơ sở dữ liệu
- Các mô hình dữ liệu, đặc biệt là mô hình quan hệ và mô hình thực thể kết hợp
- Các vấn đề khác như chuẩn hóa, an toàn dữ liệu, quản lý giao dịch, cơ sở dữ liệu phân tán

# TÀI LIỆU THAM KHẢO

- Trường CĐ CNTT Hữu nghị Việt – Hàn, Giáo trình Hệ cơ sở dữ liệu
- Lê Tiến Vương, Nhập môn cơ sở dữ liệu quan hệ
- Nguyễn Kim Anh, Nguyên lý các hệ cơ sở dữ liệu
- Hồ Thuấn - Hồ Cẩm Hà, Các hệ cơ sở dữ liệu - Lý thuyết và thực hành
- Phương Lan, Giáo trình nhập môn cơ sở dữ liệu
- Thomas Connolly - Carolyn Begg, Database systems - a practical approach to design, implementation, and management, Forth edition

# TIÊU CHUẨN ĐÁNH GIÁ SINH VIÊN

- Điểm kiểm tra thường kỳ: 10%
- Điểm kiểm tra giữa kỳ: 20 %
- Điểm thi kết thúc học phần: 70 %

# CHƯƠNG 1

## GIỚI THIỆU VỀ CƠ SỞ DỮ LIỆU

- 1.1. Hệ xử lý tập tin truyền thống
- 1.2. Hạn chế của hệ xử lý tập tin truyền thống
- 1.3. Hệ cơ sở dữ liệu
- 1.4. Đối tượng sử dụng
- 1.5. Ưu và nhược điểm của hệ CSDL

Hệ xử lý tập tin truyền thống

Hạn chế



Hệ cơ sở dữ liệu

Người sử dụng

Ưu điểm,  
nhược điểm

**Mục tiêu: Hiểu được hệ cơ sở dữ liệu là gì?**

**Tầm quan trọng của nó trong đời sống hiện nay**

# 1.1. Hệ xử lý tập tin truyền thống

- Hệ xử lý tập tin là một tập hợp các chương trình ứng dụng đáp ứng các yêu cầu của người sử dụng. Mỗi chương trình định nghĩa và quản lý dữ liệu của chính nó





# 1.1. Hệ xử lý tập tin truyền thống

- Mỗi chương trình ứng dụng đều có một tập tin dữ liệu tương ứng, mỗi khi chương trình ứng dụng cần được sửa đổi hoặc mở rộng thì tập tin dữ liệu tương ứng cũng thay đổi theo
- Đối với một công ty, tập trung vào các nhu cầu xử lý dữ liệu của các phòng riêng lẻ trong một công ty, mà không xem xét công ty này như một tổng thể

# Ví dụ 1.1

- **Phòng Kế toán**
  - Tính lương và in danh sách lương
  - Họ và Tên, Hệ số lương, Hệ số phụ cấp, Phụ cấp khác
  - Microsoft Excel
- **Phòng Tổ chức**
  - Quản lý thông tin lý lịch của CNVC chi tiết hơn
  - Họ CNVC, Tên CNVC, Bí danh, Giới tính, Ngày sinh, Ngày tuyển dụng, Hoàn cảnh gia đình...
  - FoxPro for Windows
- **Phòng Tổ chức cán bộ - Tổng công ty**
  - Quản lý CNVC gồm các cán bộ chủ chốt từ trưởng phó phòng, quản đốc và phó quản đốc xí nghiệp trở lên của các công ty con trực thuộc
  - Microsoft Access

# Ví dụ 1.1

- Ưu điểm
  - Ít tốn thời gian do khối lượng thông tin cần quản lý và khai thác nhỏ
  - Khả năng đáp ứng nhanh chóng, kịp thời
- Nhược điểm
  - Lãng phí công sức nhập tin và không gian lưu trữ thông tin
  - Không nhất quán dữ liệu
  - Do hệ thống được tổ chức thành các tập tin riêng lẻ nên việc nâng cấp ứng dụng sẽ là rất khó khăn

## 1.2. Hạn chế của hệ xử lý tập tin truyền thống

- Cô lập và hạn chế chia sẻ dữ liệu:
  - Dữ liệu nằm trong nhiều tập tin, các tập tin có khuôn dạng khác nhau
  - NSD khó khai thác các dữ liệu chứa trong các tập tin đang tồn tại
- Dư thừa dữ liệu và không nhất quán
  - Cùng một dữ liệu có thể được lưu trữ trong nhiều tập tin khác nhau
  - Khi tiến hành cập nhật có thể bỏ sót và dẫn đến không nhất quán
- Khó khăn trong việc truy xuất dữ liệu
  - Không cung cấp các công cụ cho phép truy xuất dữ liệu một cách hiệu quả và thuận lợi

## 1.2. Hạn chế của hệ xử lý tập tin truyền thống

- Các vấn đề về tính nguyên tử
  - Tính nguyên tử của một giao dịch: hoặc thực hiện hoàn toàn hoặc không thực hiện gì cả
  - Khó đảm bảo tính nguyên tử
- Tính dị thường trong truy cập tương tranh
  - Một hệ thống có nhiều NSD cập nhật dữ liệu đồng thời có thể dẫn đến dữ liệu không nhất quán
- Vấn đề an toàn
  - Hệ thống khó đảm bảo được tính phân quyền, chống truy xuất trái phép

## 1.3. Hệ cơ sở dữ liệu

1.3.1. Cơ sở dữ liệu (Database)

1.3.2. Hệ quản trị cơ sở dữ liệu (DataBase Management System - DBMS)

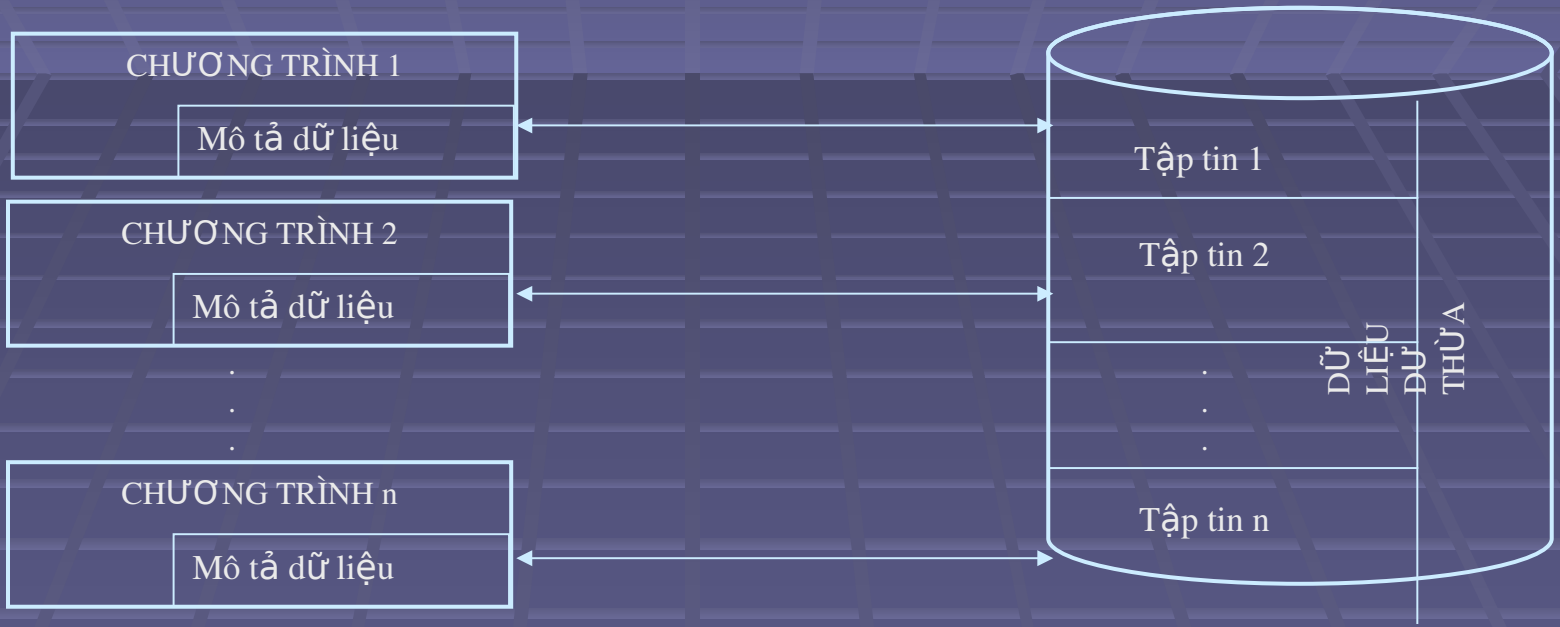
1.3.3. Hệ cơ sở dữ liệu (Database system)

## 1.3.1. Cơ sở dữ liệu

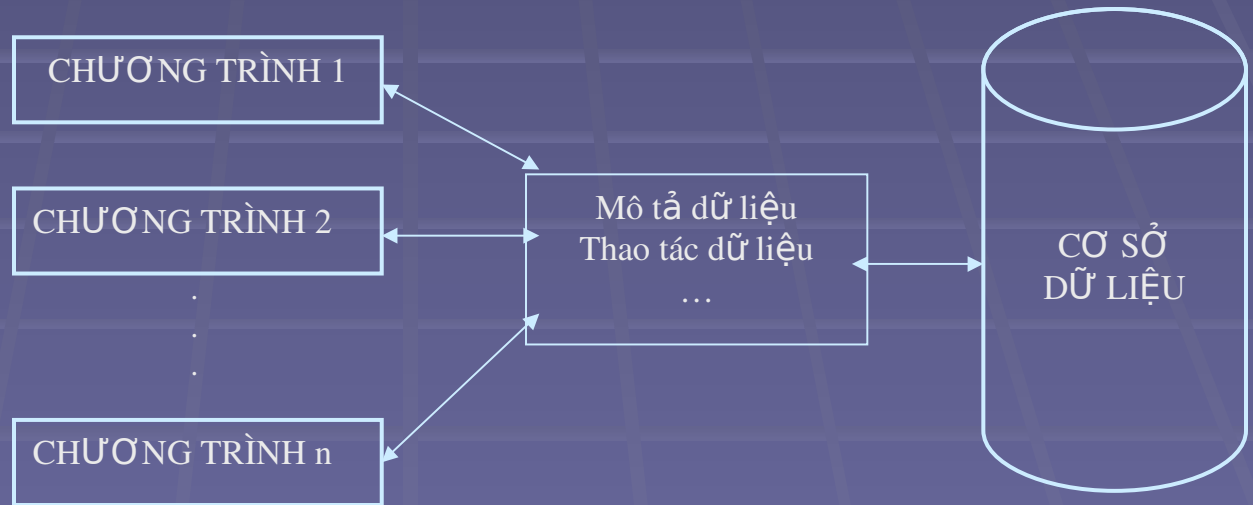
- Dữ liệu là các sự kiện, văn bản, hình ảnh, âm thanh, phim ảnh... có ý nghĩa được ghi lại và lưu trữ trên máy tính
- Thông tin là các dữ liệu đã được xử lý, hữu ích cho việc đưa ra các quyết định
- Cơ sở dữ liệu (CSDL)
  - là một hệ thống **các thông tin có cấu trúc**
  - được lưu trữ trên các thiết bị lưu trữ thông tin thứ cấp (như băng từ, đĩa từ...)
  - để có thể thỏa mãn yêu cầu khai thác thông tin **đồng thời** của nhiều NSD hay nhiều chương trình ứng dụng với nhiều mục đích khác nhau

# 1.3.1. Cơ sở dữ liệu (tt)

## Hệ xử lý tập tin truyền thống



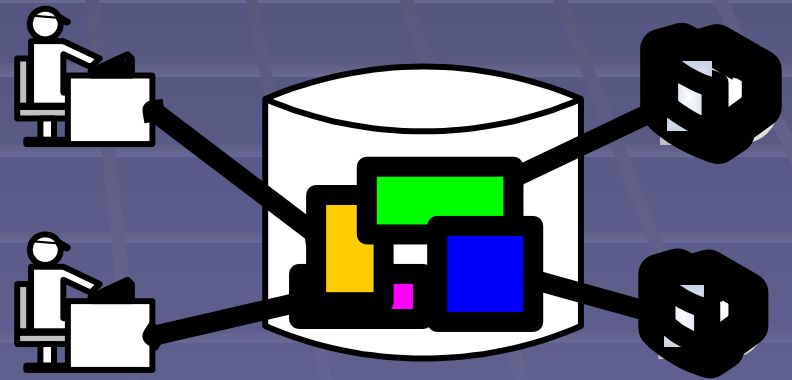
## Hệ cơ sở dữ liệu





## 1.3.1. Cơ sở dữ liệu (tt)

- Ví dụ về cơ sở dữ liệu
  - Trang niên giám điện thoại chẳng hạn
  - Danh sách khách hàng
  - Danh mục hay danh sách sinh viên
  - Danh sách hàng hóa
  - Trang Web



## 1.3.2. Hệ quản trị cơ sở dữ liệu

- Phần mềm cho phép người dùng giao tiếp với CSDL, cung cấp một môi trường thuận lợi và hiệu quả để tìm kiếm và lưu trữ thông tin của CSDL

# Quá trình phát triển các hệ quản trị CSDL

- Đầu những năm 1960: các hệ QT CSDL đầu tiên ra đời dựa trên **mô hình dữ liệu phân cấp và mạng**
  - **IMS** của hãng IBM dựa trên mô hình dữ liệu phân cấp
- Năm 1976: hệ quản trị CSDL đầu tiên dựa trên **mô hình dữ liệu quan hệ**: mục tiêu tổ chức dữ liệu, truy cập và cập nhật những khối lượng lớn dữ liệu một cách thuận lợi, an toàn và hiệu quả
  - **System-R** của hãng IBM
- Từ năm 1980: các hệ quản trị CSDL quan hệ khác
  - **DB2, Dbase, Sybase, Oracle, Informix, SQL-Server, Access...**

# Quá trình phát triển các hệ quản trị CSDL (tt)

- Từ những năm 1990 người ta bắt đầu cố gắng xây dựng các **hệ quản trị CSDL hướng đối tượng**. Tuy nhiên hầu hết các hệ này đều vẫn là quan hệ - hướng đối tượng
  - **Orion, Illustra, Itasca...**
- Hệ quản trị CSDL hướng đối tượng thuần nhất ra đời vào năm 1996: hỗ trợ các ứng dụng đa phương tiện (văn bản, âm thanh, hình ảnh) và động (các chương trình, mô phỏng)
  - **ODMG**

# Quá trình phát triển các hệ quản trị CSDL (tt)

- Hiện nay có khá nhiều hệ quản trị CSDL mạnh:
  - Visual FoxPro
  - Microsoft Access
  - SQL-Server
  - DB2
  - Sybase
  - Paradox
  - Informix
  - Oracle
  - ...

# Các chức năng của hệ quản trị CSDL

- Cung cấp cho người dùng khả năng lưu giữ, truy xuất và cập nhật dữ liệu
- Cung cấp cho người dùng một từ điển dữ liệu, đó là mô tả về dữ liệu được lưu trữ
- Hỗ trợ các giao dịch bằng cách cung cấp một cơ chế đảm bảo rằng: hoặc tất cả các cập nhật trong một giao dịch làm việc được thực hiện, hoặc không thao tác cập nhật nào trong giao dịch này được thực hiện
- Cung cấp các dịch vụ điều khiển tương tranh để đảm bảo tính nhất quán dữ liệu khi có nhiều người đồng thời truy cập vào CSDL

# Các chức năng của hệ quản trị CSDL (tt)

- Cung cấp một cơ chế để khôi phục dữ liệu
- Cung cấp các dịch vụ bản quyền: những người có quyền mới được truy cập CSDL ở những mức độ khác nhau
- Hỗ trợ cho truyền thông dữ liệu: Hệ quản trị CSDL phải có khả năng tích hợp được với các phần mềm truyền thông
- Cung cấp các dịch vụ đảm bảo tính toàn vẹn dữ liệu: Điều này nghĩa là đảm bảo dữ liệu trong CSDL và những thay đổi dữ liệu tuân theo những luật đã xác định

## 1.3.3. Hệ cơ sở dữ liệu

- Hệ CSDL là một hệ thống gồm có bốn thành phần:
  - Cơ sở dữ liệu hợp nhất: CSDL của hệ có hai tính chất là tối thiểu hoá dư thừa và tính chia sẻ
  - Người sử dụng
  - Phần mềm hệ quản trị CSDL
  - Phần cứng: Phần cứng của hệ bao gồm các thiết bị nhớ thứ cấp được sử dụng để lưu trữ dữ liệu
- Có thể chia hệ CSDL thành hai loại:
  - Hệ CSDL tập trung
  - Hệ CSDL phân tán



# Các hệ CSDL tập trung

- Tất cả các dữ liệu được định vị tại một trạm đơn lẻ, NSD tại các trạm từ xa có thể truy cập CSDL thông qua các công cụ truyền thông dữ liệu
- Có 3 kiểu:
  - Hệ CSDL cá nhân
    - một NSD đơn lẻ, vừa thiết kế, bảo trì, cập nhật CSDL trên cùng một máy tính
  - Hệ CSDL trung tâm
    - Dữ liệu lưu trữ trên một máy tính trung tâm
    - NSD từ xa có thể truy cập CSDL này thông qua các thiết bị đầu cuối và các móc nối truyền thông dữ liệu
  - Hệ CSDL khách/chủ
    - Dữ liệu được quản lý bởi máy chủ
    - Các máy khách truy nhập vào dữ liệu khi có yêu cầu

# Các hệ CSDL phân tán

- CSDL phân tán là tập hợp dữ liệu logic thuộc về cùng một hệ thống nhưng trải rộng ra nhiều điểm trên mạng máy tính
- Có 2 kiểu:
  - Hệ CSDL phân tán thuần nhất: Công nghệ CSDL (hệ điều hành, hệ quản trị CSDL) và dữ liệu tại các vị trí khác nhau là như nhau hoặc có thể tương thích với nhau
  - Hệ CSDL phân tán không thuần nhất: Công nghệ CSDL hoặc dữ liệu tại các vị trí khác nhau không tương thích với nhau

## 1.4. Đối tượng sử dụng

- Những người sử dụng CSDL không chuyên về lĩnh vực tin học và CSDL
- Các chuyên viên tin học biết khai thác CSDL (thiết kế cơ sở dữ liệu, phát triển các ứng dụng)
- Những người quản trị CSDL

# 1.5. Ưu và nhược điểm của hệ CSDL

1.5.1. Ưu điểm

1.5.2. Nhược điểm



## 1.5.1. Ưu điểm

- Giảm sự trùng lặp thông tin xuống mức thấp nhất và do đó tránh dư thừa dữ liệu, bảo đảm được tính nhất quán và toàn vẹn dữ liệu
- Quản lý được khối lượng dữ liệu lớn
- Có khả năng chia sẻ thông tin cho nhiều NSD và nhiều ứng dụng khác nhau
- Tăng cường tính bảo mật
- Các CSDL tuân theo các tiêu chuẩn nên dễ tích hợp
- Việc kết hợp các dữ liệu vào cùng một CSDL có thể tiết kiệm không gian lưu trữ, thời gian và công sức nhập dữ liệu

## 1.5.1. Ưu điểm (tt)

- Đảm bảo dữ liệu có thể được truy xuất theo nhiều cách khác nhau
- Hệ quản trị CSDL cung cấp một giao diện dễ sử dụng, hỗ trợ sẵn nhiều chức năng nên thuận tiện hơn cho việc phát triển các ứng dụng
- Tăng cường tính độc lập giữa dữ liệu và chương trình
- Hạn chế tranh chấp dữ liệu khi có nhiều người dùng đồng thời
- Có cơ chế sao lưu và phục hồi dữ liệu khi có sự cố

## 1.5.2. Nhược điểm

- Sử dụng CSDL sẽ phức tạp hơn so với sử dụng tập tin thông thường
- Kích thước lớn vì quản lý dữ liệu tập trung
- Phải chi phí cho hệ quản trị CSDL và các phần cứng
- CSDL lưu trữ một khối lượng dữ liệu lớn có thể làm cho các ứng dụng hoạt động chậm

# Câu hỏi

- 1.1. Cho một ví dụ và phân tích các ưu điểm, nhược điểm của một hệ xử lý tập tin truyền thống
- 1.2. Tại sao người ta lại sử dụng hệ CSDL để thay thế hệ xử lý tập tin truyền thống?
- 1.3. Hãy nêu ra các điểm khác nhau chính giữa một hệ xử lý tập tin truyền thống và một hệ CSDL
- 1.4. Phân biệt các thuật ngữ sau: dữ liệu, thông tin, cơ sở dữ liệu, hệ quản trị cơ sở dữ liệu, hệ cơ sở dữ liệu
- 1.5. Hãy nêu các nhóm đối tượng sử dụng cơ sở dữ liệu



## Câu hỏi (tt)

- 1.6. Hãy phân tích các nhóm đối tượng sử dụng CSDL tại một trường học. Giả sử trường học bao gồm cán bộ quản lý, giáo viên, các chuyên viên quản trị CSDL, các chuyên viên phát triển phần mềm ứng dụng cho trường, học sinh
- 1.7. Nêu các ưu và nhược điểm của hệ CSDL
- 1.8. CSDL được sử dụng rộng rãi trong cuộc sống, hãy nêu ví dụ về một số lĩnh vực ứng dụng CSDL, nêu rõ ứng dụng CSDL vào việc gì?