

# **CƠ SỞ DỮ LIỆU (DATABASE)**

# NỘI DUNG



- GIỚI THIỆU
- KHÁI NIỆM MỞ ĐẦU VỀ CƠ SỞ DỮ LIỆU
- LÝ DO TỔ CHỨC CƠ SỞ DỮ LIỆU
- CÁC VẤN ĐỀ CƠ BẢN
- MỘT SỐ VẤN ĐỀ NÂNG CAO
- NỘI DUNG MÔN HỌC
- YÊU CẦU ĐỐI VỚI SINH VIÊN

# KHÁI NIỆM CƠ SỞ DỮ LIỆU



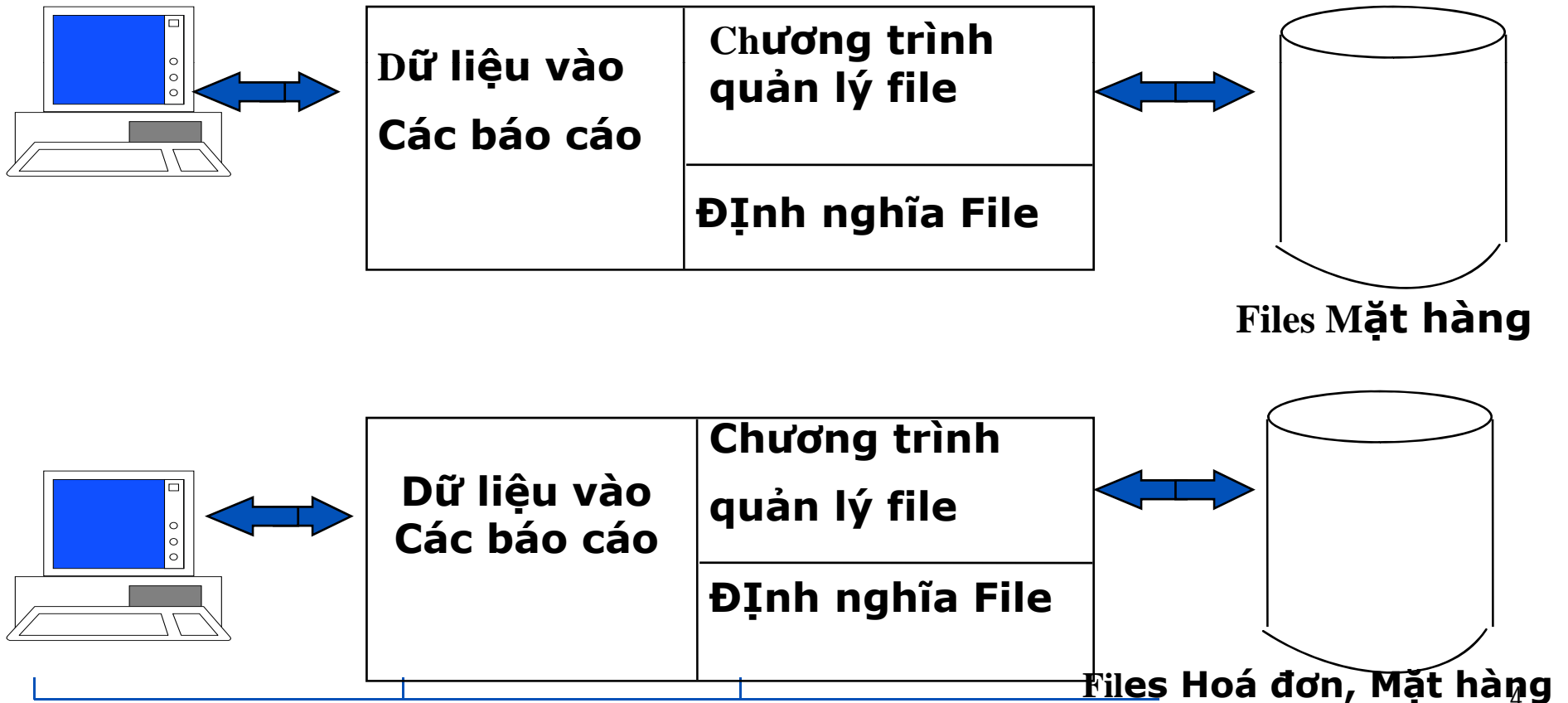
## ■ Theo Englels:

- ◆ CSDL là một bộ sưu tập những dữ liệu tác nghiệp, được lưu trữ lại và được các hệ ứng dụng của một xí nghiệp cụ thể nào đó sử dụng
- ◆ “xí nghiệp”: thuật ngữ chung chỉ những hoạt động thương mại, khoa học, kỹ thuật hoặc các hoạt động khác có quy mô đủ lớn
- ◆ “dữ liệu tác nghiệp”: không bao gồm dữ liệu vào ra, dãy thứ tự công việc hoặc những thông tin thuần túy là tạm thời

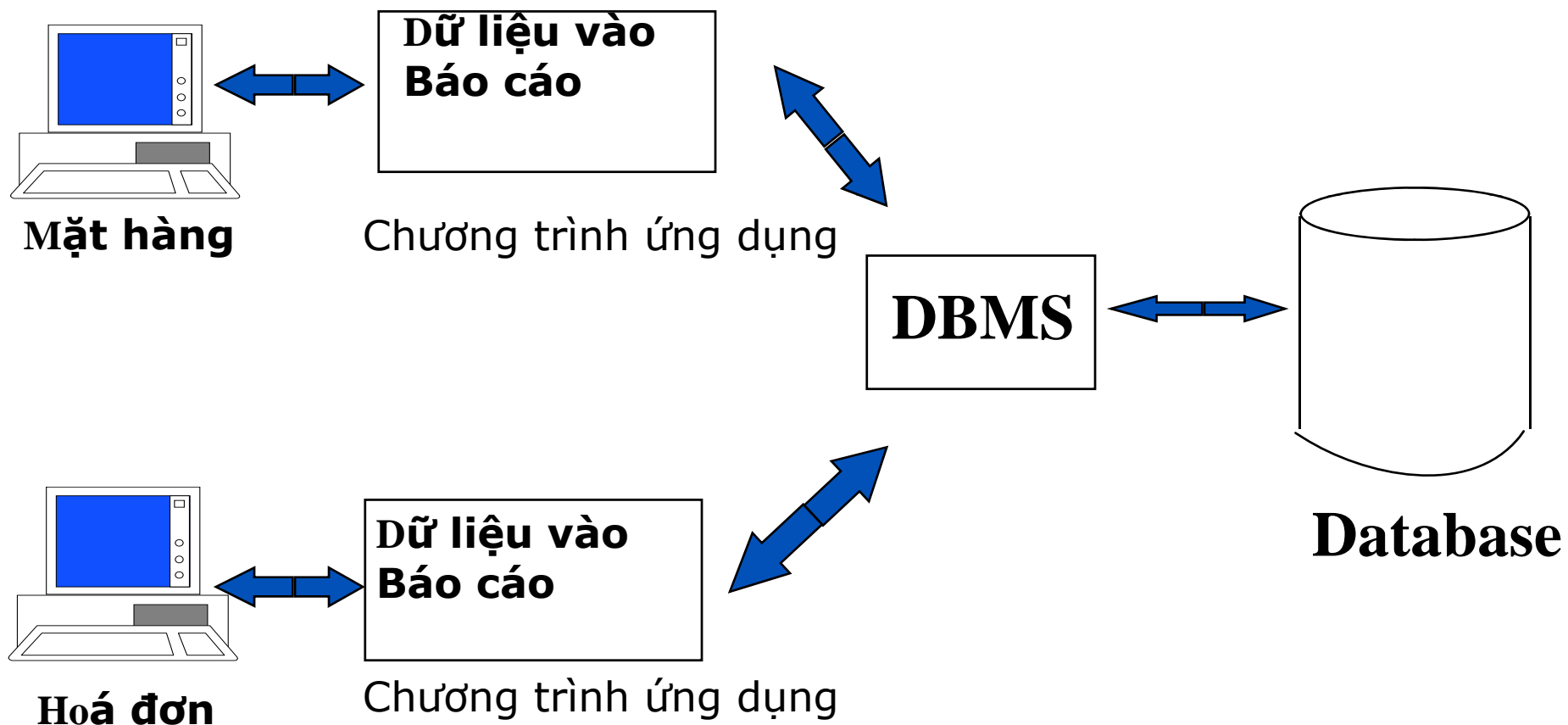
# LÝ DO TỔ CHỨC CƠ SỞ DỮ LIỆU

## ■ Cách xử lý tệp truyền thống

- ◆ Mỗi chương trình ứng dụng có riêng một số tệp dữ liệu
- ◆ viết bằng ngôn ngữ lập trình (C, Pascal,...)



## ■ Tiếp cận CSDL



## ■ Theo Meltzer:

- ◆ “CSDL đảm bảo cho xí nghiệp khả năng quản lý tập trung dữ liệu tác nghiệp của mình”

## ■ Ưu điểm của tổ chức CSDL

- ◆ Giảm dư thừa dữ liệu
- ◆ Tăng khả năng tránh không nhất quán dữ liệu
- ◆ Chia sẻ, dùng chung dữ liệu
- ◆ Tuân thủ các tiêu chuẩn thống nhất
- ◆ Áp dụng được các biện pháp an toàn dữ liệu
- ◆ Cân bằng các yêu cầu đối lập
- ◆ Đảm bảo tính độc lập dữ liệu

# CÁC VẤN ĐỀ CƠ BẢN



- Cơ sở dữ liệu – Hệ quản trị CSDL
- Mô hình dữ liệu
- Thao tác dữ liệu (Tạo lập, Cập nhật, Tìm kiếm)
- Thiết kế CSDL
- Tổ chức dữ liệu vật lý
- Quản lý giao tác
- Phục hồi dữ liệu
- Bảo mật và an toàn dữ liệu

# MỘT SỐ VẤN ĐỀ NÂNG CAO



- CSDL đa phương tiện (multi-media database)
- Hệ thống thông tin địa lý (GIS – Geographic Information Systems)
- Kho dữ liệu (Data Warehouse)
- CSDL thời gian thực (Real-time database)
- CSDL web
- CSDL phân tán (Distributed database)
- CSDL song song (Parallel Database)
- CSDL suy diễn (Deductive database)
- CSDL mờ (Fuzzy Database)



# NỘI DUNG MÔN HỌC



- Cơ sở dữ liệu – Hệ quản trị CSDL
- Mô hình dữ liệu
- Thao tác dữ liệu (Tạo lập, Cập nhật, Tìm kiếm)
- Thiết kế CSDL
- Tổ chức dữ liệu vật lý

# TÀI LIỆU THAM KHẢO

- Hồ Thuần, Hồ Cẩm Hà, *Các hệ CSDL: lí thuyết và thực hành*, Tập 1, 2, NXB GD, 2004
- Nguyễn Kim Anh, *Nguyên lý của các hệ CSDL*, NXB Đại học Quốc gia Hà Nội, 2004
- C.J. Date, *An introduction to data base system*, Addition-Wesley, 1986.
- Clement T. Yu & Weiyi Meng, *Principles of Database Query Processing for Advanced Applications*, Morgan Kaufmann Publisher, 1998
- J.D. Ullman, *Principles of Database and Knowledge-base systems*, Tập 1, Computer Science Press, 1988

(Bản dịch: Nguyên lý các hệ CSDL và cơ sở tri thức, Người dịch:  
Trần Đức Quang)

# TÀI LIỆU THAM KHẢO



- Elmasri, Navathe, *Fundamentals of Database*, New York : benjamin/Cummings Publishing Company.
- Molina, Hector Garcia,J.D. Ullman, *Database Systems: The complete book*, New Jersey, Prentice Hall, 2002.

# MỘT SỐ VẤN ĐỀ NÂNG CAO

## ■ CSDL đa phương tiện

- ◆ có thể lưu trữ ảnh, âm thanh, video,...

## ■ GIS

- ◆ Dữ liệu địa lý: dữ liệu không gian + dữ liệu thuộc tính
- ◆ Nhận và lưu trữ, phân tích, quản lý, biểu diễn dữ liệu liên quan đến 1 vị trí nào đó.
- ◆ Ứng dụng:
  - Lập bản đồ, phân tích các sự vật, hiện tượng trên trái đất
  - Phân tích sự kiện, dự đoán tác động, hoạch định chiến lược
  - Ở Việt Nam: quy hoạch lâm nghiệp, quản lý rừng, tư liệu địa chất, đo đạc bản đồ, địa chính, quản lý đô thị,...

## ■ CSDL web:

- ◆ Chú trọng: Khả năng duyệt và tìm thông tin

## ■ CSDL thời gian thực

- ◆ dữ liệu thay đổi liên tục
- ◆ “thời gian thực”: 1 giao tác được xử lý đủ nhanh để kết quả trả về và thực thi ngay tức thì
- ◆ Ứng dụng trong thanh toán, ngân hàng, luật pháp, điều khiển tiến trình, phân tích dữ liệu khoa học, ....

## ■ Kho dữ liệu

- ◆ Đặc tính:
  - Hướng chủ đề
  - Tích hợp
  - Biến thái đối với thời gian
  - “Tĩnh”
- ◆ Ứng dụng: Phân tích, hỗ trợ ra quyết định

## ■ Cơ sở dữ liệu phân tán

- ◆ Dữ liệu được duy trì trên một mạng máy tính
- ◆ Các vấn đề:
  - Kỹ thuật phân tán dữ liệu?
  - Truy nhập của người dùng? (ứng dụng cục bộ, ứng dụng toàn cục)
- ◆ Đặc tính quan trọng: Tính “trong suốt” (transparency)

## ■ Cơ sở dữ liệu song song

- ◆ Nâng cao khả năng biểu diễn song song hoá các phép toán khác nhau: tải dữ liệu, xây dựng chỉ mục, định giá truy vấn

## ■ Cơ sở dữ liệu suy diễn

- ◆ Kỹ thuật CSDL + Toán logic + Trí tuệ nhân tạo
- ◆ CSDL gồm: Các sự kiện; Các luật suy diễn

# YÊU CẦU ĐỐI VỚI SINH VIÊN

- Làm đầy đủ bài tập theo yêu cầu
- Báo cáo BTL:
  - ◆ Thiết kế CSDL cho một ứng dụng cụ thể. Nộp báo cáo vào tuần học thứ 12
  - ◆ Nội dung báo cáo:
    - Trình bày bài toán thực tế
    - Xác định các chức năng trong chương trình ứng dụng
    - Mô hình thực thể liên kết
    - Thiết kế CSDL quan hệ, ánh xạ từ mô hình thực thể liên kết
    - Lấy ví dụ 10 câu hỏi biểu diễn bằng ngôn ngữ đại số và 10 câu hỏi biểu diễn bằng ngôn ngữ SQL (kèm theo hình ảnh minh họa kết quả thực hiện câu truy vấn trong một hệ quản trị csdl cụ thể)
  - ◆ GỢI Ý MỘT SỐ ĐỀ TÀI
    - Quản lý đoàn viên
    - Quản lý hồ sơ sinh viên
    - Quản lý kí túc xá
    - Quản lý hộ khẩu thường trú
    - Quản lý thư viện, .....

## ■ Phương án đánh giá kết quả:

- ◆ Điểm giữa kì = Điểm báo cáo bài tập lớn + Điểm tích cực – Điểm phạt
- ◆ Điểm tích cực: Đánh giá qua mức độ tích cực tham gia đóng góp xây dựng bài giảng trên lớp và bài tập về nhà (nếu có)
- ◆ Điểm phạt: Nếu phát hiện sao chép bài tập (trên lớp hoặc về nhà) (-1 đ/1 bài) và bài tập lớn (-100% số điểm báo cáo BTL).
- ◆ Điểm cuối kì = Điểm bài thi cuối kì (Thi viết 90 phút, thang điểm 7) + Điểm báo cáo chương trình (thang điểm 3) + Điểm tích cực (nếu còn)