

**TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI**

## **Khoa Điện**

**Bộ môn**

**Kỹ thuật đo và Tin học Công nghiệp**

# **Cơ sở dữ liệu**

Người soạn: **ThS. Lê Văn Lợi**

**Viện NC Điện tử, Tin học, Tự động hóa  
Bộ Công nghiệp**

**HÀ NỘI – 2005**

## Nhập môn

Giáo trình này được soạn để phục vụ các bài giảng của môn “Cơ sở dữ liệu”. Đối tượng của giáo trình này là sinh viên Khoa Điện, Bộ môn Kỹ thuật đo và Tin học Công nghiệp. Chúng ta cần lưu ý rằng môn học “Cơ sở dữ liệu” luôn luôn là môn cơ bản của sinh viên ngành Công nghệ Thông tin. Môn này thường được giảng trong nhiều học kỳ, kết hợp với phần thực hành tại trường, thực tập tại các Viện, công ty chuyên ngành cơ sở dữ liệu. Điều này chứng tỏ rằng với thời lượng 30 tiết (2 học trình) giáo trình này chỉ có thể giới thiệu ở mức cơ bản, gợi mở để sinh viên tìm hiểu thêm, tự học là chính.

Tuy vậy, mục tiêu của môn học vẫn định hướng đem lại cho sinh viên khả năng ứng dụng được cơ sở dữ liệu vào các vấn đề thực tế ngay sau khi kết thúc môn học. Để đạt được mục tiêu này, giáo trình sẽ gắn liền với việc thực hành và vận hành hệ quản trị cơ sở dữ liệu MySQL. Hệ quản trị cơ sở dữ liệu MySQL được phát triển bởi công ty MySQL AB, là một công ty của Thụy Điển. Có người sẽ hỏi “Vậy tại sao lại chọn MySQL”? Các lý do chính là:

- ◆ Tác giả của giáo trình này hiểu rõ MySQL hơn các hệ quản trị cơ sở dữ liệu khác;
- ◆ MySQL là phần mềm mã nguồn mở, miễn phí, có thể dễ dàng cài đặt lên một máy PC, có đủ cả phần Server, Client, có thể tích hợp một cách dễ dàng với PHP, Java, VisualBasic, và các ứng dụng khác. MySQL có thể chạy trên một máy, trên mạng Intranet, trên mạng Internet.
- ◆ Qua kinh nghiệm thực tế của tác giả, sử dụng MySQL rất ít khi gặp lỗi, có các công cụ rất dễ sử dụng, giải quyết các vấn đề thực tế một cách dễ dàng.

(Tham khảo MySQL tại <http://www.mysql.com>)

Đó là về công cụ thực hành. Với mục tiêu là giúp sinh viên hiểu rõ hơn làm thế nào để ứng dụng các phép toán cơ sở dữ liệu vào thực tế, tác giả đặt hai bài toán giả định như sau:

- ◆ *Bài toán đo lường*: giả thiết là nhà máy sản xuất xi măng X có một hệ SCADA có thể đo được các đại lượng đầu vào là đá, clinke, phụ gia, công suất tiêu thụ điện và số lượng xi măng xuất xưởng (xi măng rời, xi măng đóng bao) một cách tự động. Dữ liệu đo được nhập theo ca, ngày, tuần, tháng, quý và năm. Giả thiết rằng các đại lượng đo sau đó được chuyển vào cơ sở dữ liệu trên máy chủ của nhà máy, máy chủ có kết nối với mạng Intranet (nội bộ) và Internet. Sinh viên được đặt vấn đề là hình dung, mô hình hóa, thiết kế, truy vấn cơ sở dữ liệu, lập trình từng phần hoặc toàn bộ sao cho:
  - Ban giám đốc nhà máy, với các quyền truy xuất được cấp, có thể biết được tình hình nguyên vật liệu tiêu thụ, sản phẩm xuất xưởng của nhà máy dù họ ở bất cứ đâu (tại nhà máy, đi công tác);
  - Bộ phận cung ứng vật tư, với các quyền truy xuất được cấp, có thể thống kê được tình hình tiêu thụ và lập kế hoạch mua sắm vật tư một cách thích hợp;

- ◆ *Bài toán điều khiển*: máy chủ gia đình (Home Server). Kịch bản là: trong một xã hội hiện đại, các gia đình đều có Home Server. Home Server là một máy tính nhỏ, tiêu thụ ít năng lượng và được gắn với các bộ phận khác như máy điều hòa nhiệt độ, quạt thông gió, máy pha cà phê, ghi và trả lời điện thoại, hệ thống an ninh gia đình bằng cách giám sát qua video, ghi và tự động phát hiện, cảnh báo đột nhập bất hợp pháp, ... Sinh viên được đặt vấn đề là hình dung, mô hình hóa, thiết kế, truy vấn cơ sở dữ liệu, lập trình từng phần hoặc toàn bộ sao cho:
  - Trước khi về nhà, thành viên gia đình có thể bật máy điều hòa, bật máy pha cà phê trước 10 phút thông qua điện thoại di động, máy PDA hoặc máy tính xách tay (với điều kiện phải có mã hóa bảo mật, tránh trường hợp bất cứ ai cũng có thể bật điều hòa hoặc hệ thống thông gió).
  - Từ xa, thành viên gia đình vẫn có thể tra danh bạ điện thoại được lưu trên Home Server.

Đó là các bài toán lớn của thực tế. Để đạt được điều đó, sinh viên cần thực hành các bước nhỏ hơn như:

- ◆ Tạo cơ sở dữ liệu
- ◆ Kiến tạo và cập nhật kiến trúc các quan hệ (bảng)
- ◆ Lập mối quan hệ giữa các bảng
- ◆ Truy vấn
- ◆ Kết nối với một ngôn ngữ lập trình,
- ◆ ...

Tất cả các thao tác đó đều được thực tập trực tiếp trên MySQL.

Vậy sinh viên cần làm gì?

1. Trước hết sinh viên cần nắm các khái niệm cơ bản thông qua các bài giảng
2. Sinh viên cần củng cố các kiến thức bằng cách thực tập trên máy
3. Dựa vào việc đặt vấn đề của các bài toán giả định, sinh viên được khuyến khích đọc thêm các tài liệu khác nhằm giải quyết các bài toán đó.

Sinh viên cần lưu ý rằng, trong phần kiểm tra cuối cùng, họ cần đưa ra các lời giải chính xác, chi tiết, cụ thể. Đáp án của các câu hỏi kiểm tra cuối kỳ thường có tập hợp hữu hạn các giải pháp và có thể có các câu hỏi trắc nghiệm. Tác giả khuyến khích sinh viên nên suy nghĩ rộng nhưng thực hành thì nhỏ, chi tiết, cụ thể.

Mặt khác, sinh viên cần lưu ý, giáo trình được soạn ra chủ yếu để định hướng, bài giảng trên lớp có thể khác và thường được cập nhật hơn giáo trình.

## **Phần I. Các khái niệm cơ bản (4 tiết)**

- Cơ sở dữ liệu là gì?
- Tại sao lại cần cơ sở dữ liệu?
- Kiến trúc tổng thể của một hệ CSDL
- Kiến trúc 3 mức
- Khái niệm về quản trị CSDL và hệ quản trị CSDL
- Khái niệm cơ bản về Client/Server
- Khái niệm về xử lý phân tán
- Các khái niệm cơ bản về mô hình quan hệ

## **Phần II. Cơ sở dữ liệu quan hệ (6 tiết)**

- Miền và quan hệ
- Tính toàn vẹn của dữ liệu quan hệ
- Các chuẩn quan hệ
- Khóa, khóa chính, khóa phụ, khóa ngoại
- Khái niệm về giá trị Null trong CSDL quan hệ
- Sơ lược về các phép toán quan hệ: union, intersection, difference, product, restriction, projection, join, natural join, division
- Các phép toán cập nhật: Insert, Update, Delete
- Cơ bản về ngôn ngữ SQL
- Các lệnh DDL
- Các lệnh DML
- Cài đặt MySQL
- Các bài tập cơ bản về CSDL

## **Phần III. Mô hình hóa cơ sở dữ liệu (6 tiết)**

- Mô hình hóa là gì?
- Tại sao lại cần mô hình hóa
- Các đặc điểm chính của UML
- Sơ lược về XML
- Chi tiết về ERD
- Các bước biến đổi ERD về các quan hệ
- Thiết kế các bài toán về xử lý trong công nghiệp
- Giới thiệu một số phương pháp và công cụ thiết kế thông dụng
- Bài tập về thiết kế bài toán công nghiệp (kết hợp với các môn khác)

## **Phần IV. Khảo sát CSDL mã nguồn mở (8 tiết)**

- Khái niệm về mã nguồn mở
- Các phương pháp chung về sử dụng phần mềm mã nguồn mở
- Chi tiết về MySQL
- Đặc điểm thiết kế ứng dụng trong công nghiệp
- Đặc tính thời gian thực của các bài toán trong công nghiệp
- Cơ sở để lựa chọn HQT CSDL cho ứng dụng trong công nghiệp
- Chọn ngôn ngữ phát triển ứng dụng: C, C++, Java, PHP, ....

- Bài tập về thiết lập MySQL lên PC và Server

**Phần V. CSDL trên mạng Internet/Intranet (6 tiết)**

- Sơ lược về TCP/IP
- Sơ lược về HTML
- Tích hợp CSDL với Web Server – các nguyên tắc cơ bản
- Khảo sát một số Web Server thông dụng: IIS, Apache
- Giới thiệu một số ngôn ngữ lập trình CGI: Perl, PHP, ASP, JSP, javascript,  
...
- Tích hợp mạng công nghiệp với Internet/Intranet
- Bài tập: tạo Website công nghiệp

Do hạn chế về mặt thời gian, tôi không soạn ngay được một giáo trình hoàn chỉnh. Toàn bộ đề cương nêu trên phần lớn sẽ được đưa vào các bài giảng. Để dễ theo dõi tôi soạn thành các bài đánh số từ bài 1 đến bài 10.

Các bài từ 1-5 sẽ phủ phần lý thuyết và các bài 6-10 sẽ tập trung vào việc ứng dụng thực tế.

Các bài giảng có thể tải về từ <http://www.vielina.com/uni/db/>

Sinh viên có thể thực hành một số câu lệnh SQL từ <http://www.vielina.com/uni/sql/>