**Data Management Layer Design**

Một nhóm dự án thiết kế lớp quản lý dữ liệu của một hệ thống sử dụng một quy trình gồm bốn bước:

* Chọn định dạng lưu trữ,
* Ánh xạ các lớp miền vấn đề sang định dạng đã chọn.
* Tối ưu hóa bộ nhớ để thực hiện hiệu quả.
* Thiết kế truy cập dữ liệu cần thiết và các lớp thao tác.

Mục tiêu của chương:

* Làm quen với một vài định dạng tồn tại đối tượng.
* Có thể ánh xạ các đối tượng miền vấn đề với các định dạng tồn tại đối tượng khác nhau.
* Có thể áp dụng các bước bình thường hóa vào một cơ sở dữ liệu quan hệ.
* Có thể tối ưu hóa cơ sở dữ liệu quan hệ để lưu trữ và truy cập đối tượng.
* Làm quen với các chỉ mục cho cơ sở dữ liệu quan hệ.
* Có thể ước lượng kích thước của một cơ sở dữ liệu quan hệ.
* Hiểu được ảnh hưởng của các yêu cầu phi chức năng đối với lớp quản lý dữ liệu.
* Có thể thiết kế các lớp truy cập dữ liệu và các thao tác.

**I. Cơ sở dữ liệu quan hệ**

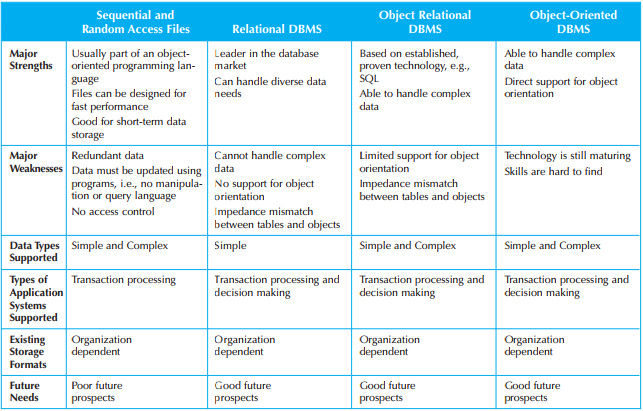
- Cơ sở dữ liệu quan hệ là loại cơ sở dữ liệu phổ biến nhất để phát triển ứng dụng hiện nay. Một cơ sở dữ liệu quan hệ được dựa trên sự kết hợp của các bảng với mỗi bảng có một khóa chính-một lĩnh vực hoặc các lĩnh vực có giá trị là duy nhất cho mỗi hàng của bảng. Hầu hết các hệ thống quản lý cơ sở dữ liệu quan hệ (RDBMS) hỗ trợ toàn vẹn referential, hoặc ý tưởng đảm bảo rằng các giá trị liên kết các bảng với nhau thông qua các khóa chính và khóa ngoại được hợp lệ hóa và chính xác. Các bảng có một số cột đã được thiết lập và một số lượng các hàng có thể thay đổi sự xuất hiện của dữ liệu. Ngôn ngữ truy vấn có cấu trúc (SQL) là ngôn ngữ chuẩn để truy cập dữ liệu trong bảng, và nó hoạt động trên các bảng hoàn chỉnh. Do đó, một truy vấn được viết bằng SQL được áp dụng cho tất cả các hàng trong một bảng tất cả tại một lần, khác với nhiều ngôn ngữ lập trình, thao tác dữ liệu hàng từng hàng. Khi truy vấn phải bao gồm thông tin từ nhiều bảng, các bảng đầu tiên được kết hợp với nhau dựa trên các khóa chính và khóa ngoại. Các ví dụ về phần mềm RDBMS là Microsoft Access, Oracle, DB2, MySQL và Microsoft SQL Server…

- Các hệ thống quản lý cơ sở dữ liệu quan hệ đối tượng (ORDBMS)

- Các hệ thống quản lý cơ sở dữ liệu quan hệ hướng đối tượng (OODBMS)

Mỗi tệp tin và các định dạng lưu trữ cơ sở dữ liệu đã được trình bày có những điểm mạnh và điểm yếu, và không có một định dạng nào là tốt hơn so với những định dạng khác. Trong thực tế, đôi khi nhóm dự án sẽ chọn nhiều định dạng. Vì vậy, điều quan trọng là phải hiểu những điểm mạnh và điểm yếu của mỗi định dạng và khi nào sử dụng.

- So sánh các định dạng đối tượng:



=> Lựa chọn thiết kế cơ sở dữ liệu cho BTL của nhóm BK\_FS là Cơ sở dữ liệu quan hệ.

**II. Các dạng chuẩn cơ bản trong thiết kế Cơ sở dữ liệu**

- Dạng chuẩn 1 (1NF): Một bảng có dạng chuẩn 1NF nếu và chỉ nếu mỗi hàng có cùng một số lượng các cột (nghĩa là không có cột lặp).

- Dạng chuẩn 2 (2NF): Một bảng có dạng 2NF nếu và chỉ nếu bảng này có dạng chuẩn 1NF và mỗi cột không phải là thành phần của khoá sẽ phụ thuộc hàm vào toàn bộ khoá (các cột không là khoá phải phụ thuộc hàm vào khoá và các cột không phải khoá không được phụ thuộc hàm vào một phần của khoá).

- Dạng chuẩn 3(3NF): Một bảng có dạng 3NF nếu và chỉ nếu bảng thuộc 2NF và mỗi cột không phải là thành phần của khoá chỉ phụ thuộc vào khoá(nghĩa là không phụ thuộc hàm vào những cột không phải là khoá).

**III. Thiết kế lưu trữ các dữ liệu hệ thống Website Tin tức**

Ảnh có chứa văn bản, biên lai, ảnh chụp màn hình

Mô tả được tạo với mức tin cậy rất cao

- Việc thiết kế lưu trữ các dữ liệu của hệ thống Website tin tức được thực hiện trên hệ quản trị cơ sở dữ liệu MySQL. Lý do chọn MySQL bởi:

* MySQL là phần mềm mã nguồn mở phổ biến nhất, có sự ổn định và cơ chế xử lý nhanh, chi phí thấp.
* MySQL dễ cài đặt và cài đặt nhanh chóng đồng thời còn có thể chạy trên nhiều nền tảng (Linux, Apache,…) nên rất linh hoạt.
* MySQL hỗ trợ giao dịch mạnh 1 cách tự động, thống nhất, độc lập và bền vững; sở hữu engine xử lý tốc độ cao và khả năng chèn dữ liệu nhanh.
* MySQL có cơ chế bảo mật tốt.

=> Phù hợp để xây dựng một trang Website tin tức.

- Việc sử dụng dạng chuẩn 3 (3NF) trong thiết kế lưu trữ các dữ liệu nhằm:

* Giảm lượng dữ liệu dư thừa
* Đảm bảo sự độc lập dữ liệu

=> Tối ưu hóa hiệu quả lưu trữ

- Về tối ưu hóa tốc độ truy cập dữ liệu:

* Ở trong MySQL khi lập Index cho trường dữ liệu thì dữ liệu sẽ được sắp xếp trên một file riêng, khi chúng ta thực hiện truy vấn dữ liệu thông qua các trường index này, các giải thuật tìm kiếm được cài đặt sẵn trên hệ thống sẽ phát huy hiệu quả tối đa của nó, đặc biệt là ở các trường index dạng số.

- Ước lượng Dung lượng lưu trữ dữ liệu: (cái này là số ở trong thiết kế VD int (11))