

Data Management Using Microsoft SQL Server

Are you registered with Onlinevarsity.com?

No

Did you download this book from Onlinevarsity.com?

Yes

Yes



No



Scores

For each YES you score 50

For each NO you score to

If you score less than 100 this book is illegal.

Register on www.onlinevarsity.com

Quản lý dữ liệu sử dụng Microsoft SQL Server Hướng dẫn người học © 2014 Aptech Limited

Mọi quyền được bảo lưu.

Không có phần nào của cuốn sách này có thể được sao chép hoặc sao chép dưới mọi hình thức hoặc bằng bất kỳ phương tiện nào – đồ họa, điện tử hoặc cơ khí, bao gồm cả photocopy, ghi âm, ghi hình, hoặc lưu trữ trong hệ thống thông tin hoặc gửi hoặc truyền đi mà không có sự cho phép trước bằng văn bản của chủ sở hữu quyền tác giả Aptech Limited.

Mọi nhãn hiệu thương mại đã được công nhận. APTECH LIMITED

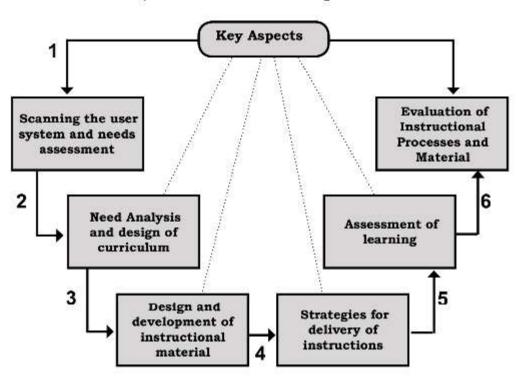
E-mail:ov-support@onlinevarsity.com

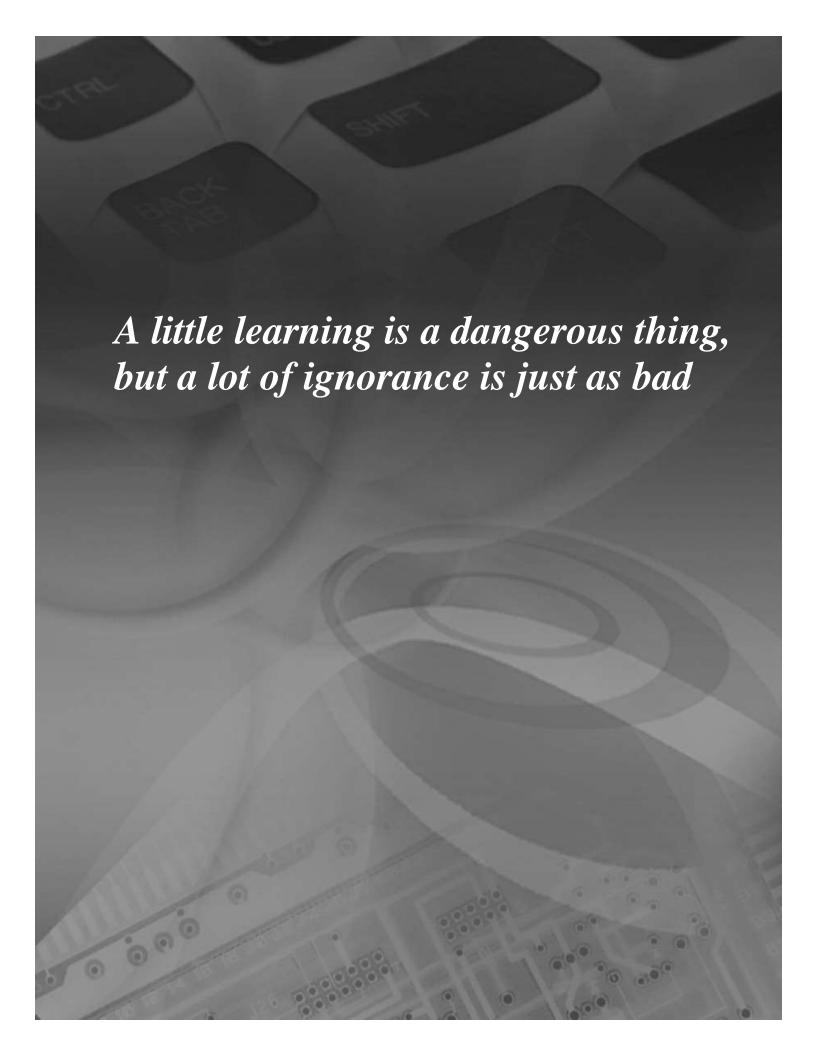






Aptech New Products Design Model







Lời nói đầu

SQL Server 2012 là Hệ thống quản lý cơ sở dữ liệu quan hệ (RDBMS) dựa trên máy khách-máy chủ mới nhất của Microsoft. Hệ thống cung cấp một nền quản lý dữ liệu mức doanh nghiệp cho tổ chức. SQL Server bao gồm rất nhiều tính năng và công cụ làm cho nó thành một cơ sở dữ liệu và nền tảng phân tích dữ liệu xuất sắc. Nó còn được nhắm tới Xử lý giao dịch trực tuyến (OLTP) quy mô lớn, lưu kho dữ liệu, và các ứng dụng thương mại điện tử. Một trong những tính năng chính của phiên bản SQL Server này là nó có sẵn trên nền tảng điện toán đám mây.

Cuốn sách này bắt đầu với phần giới thiệu về các khái niệm RDBMS và chuyển sang giới thiệu ngắn gọn về SQL Azure. Cuốn sách này sau đó đề cập đến nhiều chủ đề về SQL Server 2012 như là các loại dữ liệu, sử dụngTransact-SQL, và các đối tượng cơ sở dữ liệu như các chỉ mục, thủ tục đã lưu trữ, hàm, và vân vân. Cuốn sách này còn mô tả các giao dịch, các phần tử lập trình với Transact-SQL, và cuối cùng khắc phục sự cố các lỗi với các kỹ thuật xử lý lỗi.

Cuốn sách này là kết quả của một nỗ lực tập trung của Đội thiết kế, là sự phấn đấu liên tục để mang lại cho bạn điều tốt nhất và mới nhất trong Công nghệ thông tin. Quy trình thiết kế là một phần của chứng chỉ ISO 9001 cho Aptech-IT Division, Dịch vụ Hỗ trợ Giáo dục. Là một phần động lực chất lượng của Aptech, Nhóm thiết kế tiến hành nghiên cứu chuyên sâu và làm phong phú chương trình giảng dạy để bắt kịp các xu hướng của ngành.

Chúng tôi sẽ rất vui khi nhận được ý kiến đóng góp của bạn.

Nhóm thiết kế





Mục lục

Module

1. Các khái niệm RDBMS01
2. Mô hình và chuẩn hóa thực thể - mối quan hệ27
3. Giới thiệu về SQL 201255
4. SQL Azure
5. Transact – SQL85
6. Tạo và quản lý cơ sở dữ liệu105
7. Tạo các bảng
8. Truy cập dữ liệu
9. Truy vấn và phép nối nâng cao
10. Sử dụng giao diện, thủ tục đã lưu trữ và truy vấn siêu dữ liệu237
11. Chỉ mục
12. Khởi phát
13. Lập trình transact-SQL
14. Giao tác
15. Xử lý lỗi

Learning is not compulsory but neither is survival



Khái niệm RDBMS



1.1 Giới thiệu

Các tổ chức thường xuyên duy trì một lượng lớn dữ liệu, được tạo ra như là kết quả của hoạt động hàng ngày. Một cơ sở dữ liệu là một hình thức tổ chức các dữ liệu đó. Nó có thể bao gồm các hạng mục liên quan một hoặc nhiều dữ liệu được gọi là bản ghi dữ liệu. Hãy coi một cơ sở dữ liệu như là một tập hợp dữ liệu mà các câu hỏi khác nhau có thể được hỏi. Ví dụ, 'số điện thoại và địa chỉ của các năm Bưu điện gần nhất là gì?' hay 'Liệu chúng ta có bất kỳ cuốn sách nào trong thư viện nói về sức khỏe và thực phẩm? Nếu có, thì chúng nằm ở đâu? 'hoặc 'Cho tôi xem hồ sơ nhân viên và doanh số bán hàng của năm người bán hàng tốt nhất, trong quý hiện tại, nhưng các chi tiết về địa chỉ của họ không cần phải được đưa ra'

1.2 Dữ liệu và CSDL

Dữ liệu là thông tin, và nó là thành phần quan trọng nhất trong bất kỳ công việc nào được thực hiện. Trong hoạt động hàng ngày, các dữ liệu hiện có được sử dụng, hoặc nhiều dữ liệu nữa được tạo ra. Khi dữ liệu này được thu thập và phân tích, nó mang thông tin. Nó có thể là bất kỳ thông tin như thông tin về chiếc xe, thể thao, đường bay, và ...vv. Ví dụ, một nhà báo của tạp chí thể thao (là một người đam mê bóng đá) tập hợp các điểm (dữ liệu) thi đấu của Đức trong 10 trận đấu cúp thế giới. Các điểm này tạo thành dữ liệu. Khi dữ liệu này được so sánh với các dữ liệu của 10 trận đấu Cup thế giới được chơi tại Brazil, nhà báo có thể có được thông tin như quốc gia nào có một đội bóng chơi bóng đá tốt hơn.

Thông tin giúp dự đoán, và tổ chức sự kiện. Hiểu biết dữ liệu một cách thông minh mang lại thông tin. Trong thế giới kinh doanh, để có thể dự đoán một sự kiện và kế hoạch cho nó, có thể tiết kiệm thời gian và tiền bạc. Hãy xem một ví dụ, khi một công ty sản xuất xe hơi đang có kế hoạch hàng năm để mua một số bộ phận của xe, được nhập khẩu vì nó không phải là có sẵn tại địa phương. Nếu dữ liệu của việc mua các bộ phận trong năm năm qua là có sẵn, người đứng đầu công ty thực sự có thể biết được thông tin về tổng số lượng linh kiện nhập khẩu. Dựa trên các phát hiện đó, một kế hoạch sản xuất có thể được chuẩn bị. Do đó, thông tin là một yếu tố quan trong

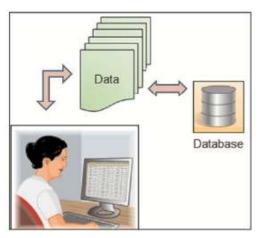
Một cơ sở dữ liệu là một tập hợp dữ liệu. Một số người thích nghĩ một CSDL là một cơ chế được tổ chức có khả năng lưu trữ thông tin. Các thông tin này có thể được lấy bởi người dùng trong một cách hiệu quả.

Một danh bạ điện thoại là một cơ sở dữ liệu. Dữ liệu được chứa bao gồm tên cá nhân, địa chỉ và số điện thoại. Danh sách được liệt kê theo thứ tự chữ cái, hoặc theo chỉ mục. Điều này cho phép người sử dụng tìm hiểu một cư dân địa phương cụ thể, một cách dễ dàng. Cuối cùng, dữ liệu này được lưu trữ trong cơ sở dữ liệu ở đâu đó trên một máy tính. Khi mọi người di chuyển đến các thành phố, hoặc các quốc gia khác nhau, các mục có thể được thêm vào, hoặc loại bỏ khỏi danh bạ điện thoại. Tương tự như vậy, các mục sẽ phải được sửa đổi cho người dân thay đổi tên, địa chỉ, số điện thoại, và ...vv.





Hình 11 minh họa khái niệm của một cơ sở dữ liệu.



Hình 1.1: Database

Như vậy, một cơ sở dữ liệu là một tập các dữ liệu được tổ chức để nội dung của nó có thể được truy cập, quản lý, và cập nhật một cách dễ dàng

1.3 Data Management (Quản lý dữ liệu)

Quản lý dữ liệu là việc quản lý một số lượng thông tin lớn, trong đó bao gồm cả việc lưu trữ các thông tin và cung cấp cơ chế cho các thao tác thông tin. Ngoài ra, hệ thống cũng nên cung cấp sự an toàn của thông tin được lưu trữ trong những hoàn cảnh khác nhau, chẳng hạn như người dùng truy cập nhiều lần, và...vv

Có hai phương pháp tiếp cận khác nhau trong việc quản lý dữ liệu là Hệ thống dựa trên tập tin (file-based systems), và hệ thống cơ sở dữ liệu (database systems).

1.3.1 File-based Systems (Hệ thống dựa trên tập tin)

Lưu trữ số lượng lớn dữ liệu luôn luôn là vấn đề lớn được quan tâm. Trong những ngày đầu, hệ thống dựa trên tập tin được sử dụng. Trong hệ thống này, dữ liệu được lưu trữ trong các tập tin rời rạc, và một tập các tập tin được lưu trữ trên một máy tính. Các dữ liệu đó có thể được truy cập bởi một nhà điều hành máy tính. Các tập tin của dữ liệu được lưu trữ được gọi là các bảng, bởi vì chúng trông giống các bảng được sử dụng trong việc lưu trữ tập tin truyền thống. Các hàng trong bảng được gọi là các bản ghi dữ liệu (records), và các cột được gọi là các trường (fields)

Thông thường, trước khi hệ thống cơ sở dữ liệu phát triển, thì dữ liệu trong hệ thống phần mềm được lưu trữ trong các tập tin phẳng.



Khái niệm RDBMS



Một ví dụ của hệ thống dựa trên tập tin được minh họa trong table 1.1.

First Name	Last Name	Address	Phone
Eric	David	ericd@eff.org	213-456-0987
Selena	Sol	selena@eff.org	987-765-4321
Jordan	Lim	nadroj@otherdomain.com	222-3456-123

Table 1.1: File-based System

Nhược điểm của hệ thống dựa trên tập tin

Trong một hệ thống dựa trên tập tin, các chương trình khác nhau trong cùng một ứng dụng, có thể được tương tác với các tập tin dữ liệu cá nhân khác nhau. Không có hệ thống thực thi bất kỳ điều khiển tiêu chuẩn hóa về tổ chức và cấu trúc của các tập tin dữ liệu.

Data redundancy and inconsistency (Dw thừa dữ liệu và tính không nhất quán)

Do các file và các trình ứng dụng được tạo ra bởi các nhà lập trình khác nhau, nên các file có định dạng khác nhau, các chương trình được viết trong các ngôn ngữ lập trình khác nhau, cùng một thông tin có thể được lưu trữ trong các file khác nhau. Tính không thống nhất và dư thừa này sẽ làm tăng chi phí truy xuất, và lưu trữ. Hơn nữa nó dẫn đến tính không nhất quán của dữ liệu: các bản sao của cùng một dữ liệu có thể không nhất quán

• Unanticipated queries (Khó khăn trong việc truy xuất dữ liệu)

Môi trường của việc xử lý file thông thường không cung cấp các công cụ cho phép truy xuất các thông tin một cách hiệu quả, và thuận lợi.

Data isolation (Sự cô lập dữ liệu)

Các dữ liệu được lưu trữ trong cơ sở dữ liệu phải thỏa mãn một số các ràng buộc về tính nhất quán của dữ liệu. Trong hệ thống xử lý file thông thường, rất khó khăn trong việc thay đổi chương trình để thỏa mãn các yêu cầu thay đổi ràng buộc. Vấn đề trở nên khó khăn hơn, khi các ràng buộc liên quan đến các hạng mục dữ liệu nằm trong các file khác nhau

Concurrent access anomalies (Tính bất thường trong truy xuất cạnh tranh)

Một hệ thống cho phép nhiều người sử dụng cập nhật dữ liệu đồng thời, có thể dẫn đến kết quả là dữ liệu không nhất quán. Điều này đòi hỏi một sự giám sát, hệ thống xử lý tập tin thông thường, không cung cấp chức năng này

Security problems (Vấn đề bảo mật)

Trong các ứng dụng dữ liệu chuyên sâu, bảo mật dữ liệu là một mối quan tâm lớn. Người sử dụng nên được chỉ truy cập vào dữ liệu cần thiết, và không phải cho toàn bộ cơ sở dữ liệu.

Concepts





Ví dụ, trong một hệ thống ngân hàng, nhân viên biên chế chỉ cần phải xem một phần của cơ sở dữ liệu có thông tin về các nhân viên ngân hàng khác nhau. Họ không cần phải truy cập vào các thông tin về tài khoản của khách hàng. Vì các chương trình ứng dụng được thêm vào hệ thống một cách tùy biến, nó rất khó để thực thi việc hạn chế bảo mật. Trong một hệ thống dựa trên tập tin, điều này có thể được xử lý chỉ bằng cách lập trình bổ sung thêm trong mỗi ứng dụng.

• Integrity problems (Vấn đề toàn vẹn)

Trong bất kỳ ứng dụng nào, sẽ có các quy tắc toàn vẹn dữ liệu nhất định, mà cần phải được duy trì. Đây có thể là trong các hình thức điều kiện nhất định / ràng buộc trên các phần tử của các bản ghi dữ liệu. Trong ứng dụng gửi tiền tiết kiệm ngân hàng, một nguyên tắc toàn vẹn có thể là "Customer ID , đó là định danh duy nhất cho một bản ghi dữ liệu khách hàng, không được để trống. Có thể có một vài quy tắc toàn vẹn như vậy. Trong một hệ thống dựa trên tập tin, tất cả các quy tắc cần phải được lập trình một cách rõ ràng trong chương trình ứng dụng.

Mặc dù tất cả đều là những vấn đề chung của mối quan tâm đến bất kỳ ứng dụng dữ liệu chuyên sâu, mỗi ứng dụng đã phải xử lý tất cả những vấn đề riêng của mình. Các lập trình viên ứng dụng cần phải bận tâm không chỉ về việc thực hiện các quy tắc kinh doanh ứng dụng mà còn, về xử lý những vấn đề chung.

1.3.2 Database Systems (Hệ thống CSDL)

Hệ thống cơ sở dữ liệu phát triển vào cuối những năm 1960 để giải quyết các vấn đề phổ biến trong các ứng dụng xử lý khối lượng lớn dữ liệu, mà cũng là số liệu chuyên sâu. Một số vấn đề có thể là từ các nhược điểm của các hệ thống dựa trên tập tin (File-based systems)

Cơ sở dữ liệu được sử dụng để lưu trữ dữ liệu một cách hiệu quả, và có tổ chức. Một cơ sở dữ liệu cho phép quản lý dữ liệu nhanh chóng và dễ dàng. Ví dụ, một công ty có thể duy trì các thông tin chi tiết của nhân viên, trong các cơ sở dữ liệu khác nhau. Tại bất kỳ thời điểm nào, dữ liệu có thể được lấy từ, hoặc thêm vào cơ sở dữ liệu, và dữ liệu có thể được tìm kiếm dựa trên một số tiêu chí trong các cơ sở dữ liệu đó.

Việc lưu trữ dữ liệu có thể đạt được ngay cả khi sử dụng các tập tin thủ công đơn giản. Ví dụ, một trường đại học có để duy trì thông tin về các giáo viên, sinh viên, môn học, và các kỳ thi.

Chi tiết về các giáo viên có thể được duy trì trong một register (sổ ghi) Nhân viên, và thông tin chi tiết của các học sinh có thể được nhập vào một register (sổ ghi) Sinh viên, và...vv. Tuy nhiên, dữ liệu được lưu trữ ở dạng này không phải là vĩnh viễn. Các bản ghi trong các tập tin thủ công chỉ có thể được duy trì trong một vài tháng, hoặc vài năm. Các sổ ghi, hoặc các tập tin lớn, làm tiêu tốn rất nhiều không gian, và do đó, không thể được giữ trong nhiều năm.

Thay vào đó, nếu cùng một dữ liệu được lưu trữ bằng cách sử dụng hệ thống cơ sở dữ liệu, thì nó có thể được bền lâu hơn.

Ưu điểm của hệ thống cơ sở dữ liệu

Thông tin hay dữ liệu có thể được lưu trữ vĩnh viễn trong hình thức cơ sở dữ liệu trên máy tính. Một hệ thống CSDL là một điểm mạnh, bởi vì nó cung cấp một chế độ tập trung kiểm soát các dữ liệu





Một số lợi ích của việc sử dụng một hệ thống cơ sở dữ liệu tập trung là như sau:

Số lượng dư thừa trong dữ liệu được lưu trữ có thể được giảm

Trong một tổ chức, một số phòng ban thường lưu trữ các dữ liệu giống nhau. Việc duy trì một cơ sở dữ liệu tập trung, giúp các dữ liệu giống nhau được truy cập bởi nhiều bộ phận. Như vậy, sự trùng lặp dữ liệu hoặc 'thừa dữ liệu' có thể được giảm.

Không còn sự thiếu nhất quán trong dữ liệu

Khi dữ liệu được sao chép trên một số phòng ban, bất kỳ sửa đổi nào đối với các dữ liệu phải được phản ánh trên tất cả các phòng ban. Đôi khi, điều này có thể dẫn đến sự không thống nhất trong dữ liệu. Nếu có một trung tâm cơ sở dữ liệu , thì một người có thể được giao nhiệm vụ cập nhật dữ liệu theo định kỳ. Hãy xem trường hợp ông Larry Finner, một nhân viên của một tổ chức được thăng chức thành một Quản lý cấp cao, từ chức vụ quản lý. Trong một trường hợp như vậy, đó chỉ là một bản ghi trong cơ sở dữ liệu cần phải được thay đổi. Kết quả là, mâu thuẫn dữ liệu được giảm.

Dữ liệu được lưu trữ có thể được chia sẻ

Một trung tâm cơ sở dữ liệu có thể được đặt trên một máy chủ, có thể được chia sẻ bởi một số người dùng. Bằng cách này, tất cả các người dùng có thể truy cập, và cập nhật thông tin phổ biến ở mọi lúc.

Các tiêu chuẩn có thể được thiết lập, và tuân theo

Một trung tâm điều khiển đảm bảo một tiêu chuẩn nhất định trong việc trình bày dữ liệu có thể được thiết lập, và tuân theo. Ví dụ, tên của một nhân viên sẽ được trình bày như là 'Mr. Larry Finner'. Việc trình bày này có thể được chia thành các thành phần sau:

- Danh xưng (Mr.)
- Tên (Larry)
- Họ (Finner)

Chắc chắn rằng tất cả các tên được lưu trong cơ sở dữ liệu sẽ thực hiện theo các định dạng như trên, nếu các tiêu chuẩn được thiết lập theo cách này.

Tính toàn ven dữ liệu có thể được duy trì

Tính toàn vẹn dữ liệu là tính chính xác của dữ liệu trong cơ sở dữ liệu. Ví dụ, khi một nhân viên từ chức, và rời khỏi công ty, thì phòng Kế toán đã được cập nhật cơ sở dữ liệu của nó, nhưng bộ phận Hành chính nhân sự thì chưa cập nhật bản ghi của mình. Các dữ liệu trong các bản ghi của công ty không chính xác là do vậy. Sự kiểm soát tập trung của CSDL giúp tránh các lỗi trên. Chắc chắn rằng nếu một bản ghi bị xóa từ một bảng, thì bản ghi liên kết của nó trong bảng khác cũng bị xóa.

Bảo mật dữ liệu có thể được thực hiện

Trong một hệ thống cơ sở dữ liệu tập trung, đặc quyền của việc sửa đổi cơ sở dữ liệu không phải là cho tất cả mọi người. Quyền này được đưa ra chỉ để một người có toàn quyền kiểm soát các cơ sở dữ liệu. Người này được gọi là Database Administrator hoặc DBA. DBA có thể thực hiện việc bảo mật bằng cách đặt các hạn chế về dữ liệu. Căn cứ vào các quyền được cấp cho họ, người dùng có thể thêm, sửa đổi, hoặc truy vấn dữ liệu





1.4 Hệ thống quản trị Cơ sở dữ liệu (DBMS)

Một hệ thống quản trị CSDL (DBMS) có thể được định nghĩa là một tập hợp các bản ghi liên quan, và một tập hợp các chương trình truy cập, và thao tác các bản ghi. Một hệ thống quản trị CSDL cho phép người dùng nhập vào, lưu trữ và quản lý dữ liệu. Vấn đề chính với các gói hệ thống quản trị CSDL trước đó là các dữ liệu đã được lưu trữ trong các định dạng tập tin phẳng. Vì vậy, các thông tin về các đối tượng khác nhau được duy trì một cách riêng biệt trong các tập tin vật lý khác nhau. Do đó, quan hệ giữa các đối tượng này, nếu có, phải được duy trì trong một tập tin vật lý riêng biệt. Như vậy, một gói duy nhất sẽ bao gồm quá nhiều tập tin, và các chức năng rộng lớn để tích hợp chúng vào một hệ thống duy nhất

Một giải pháp cho những vấn đề này, đến dưới hình thức của một hệ thống cơ sở dữ liệu tập trung. Trong một hệ thống cơ sở dữ liệu tập trung, CSDL được lưu ở vị trí trung tâm. Mọi người đều có thể truy cập vào dữ liệu được lưu trữ trong một vị trí trung tâm từ máy tính của họ. Ví dụ, một hệ thống cơ sở dữ liệu trung tâm lớn sẽ chứa tất cả các dữ liệu liên quan đến các nhân viên. Các bộ phận Kế toán, và Hành chính nhân sự sẽ truy cập vào các dữ liệu cần thiết bằng cách sử dụng các chương trình phù hợp. Các chương trình đó, hoặc toàn bộ ứng dụng sẽ nằm trên các thiết bị đầu cuối là máy tính cá nhân.

Một CSDL là một tập hợp các dữ liệu liên quan đến nhau, và một hệ thống quản trị CSDL là một tập hợp các chương trình được sử dụng để thêm hoặc sửa đổi dữ liệu này. Như vậy, một hệ thống quản trị CSDL là một tập hợp các chương trình phần mềm cho phép cơ sở dữ liệu được xác định, xây dựng, và thao tác.

Một hệ thống quản trị CSDL cung cấp một môi trường thuận tiện, và hiệu quả để sử dụng khi có một khối lượng lớn dữ liệu và, nhiều giao dịch được xử lý. Các thể loại khác nhau của hệ thống quản trị CSDL có thể được sử dụng, từ các hệ thống nhỏ chạy trên máy tính cá nhân tới các hệ thống lớn có thể chạy trên máy tính lớn.

Ví dụ về các ứng dụng cơ sở dữ liệu bao gồm những điều sau đây:

- > Hệ thống thư viện trên máy tính
- Máy rút tiền tự động ATM
- Hệ thống đặt vé máy bay
- > Hệ thống thống kê hàng hóa trên máy tính

Từ một quan điểm kỹ thuật, các sản phẩm của hệ thống quản trị CSDL có thể khác nhau rất nhiều. Các hệ thống quản trị CSDL khác nhau hỗ trợ ngôn ngữ truy vấn khác nhau, mặc dù có một ngôn ngữ truy vấn bán tiêu chuẩn gọi là Structured Query Language (SQL). Các ngôn ngữ phức tạp để quản lý hệ thống cơ sở dữ liệu được gọi là ngôn ngữ thế hệ thứ tư – Fourth Generation Language (4GLs).

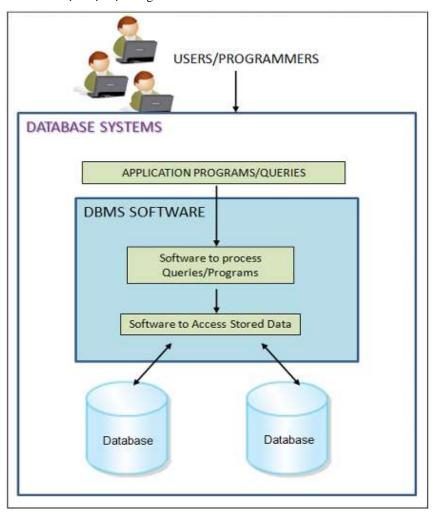
Thông tin từ một CSDL có thể được trình bày trong một loạt các định dạng. Hầu hết các hệ thống quản trị CSDL bao gồm một chương trình viết báo cáo cho phép người sử dụng xuất dữ liệu trong hình thức của một báo cáo. Nhiều hệ thống quản trị CSDL cũng bao gồm một thành phần đồ họa cho phép người sử dụng xuất các thông tin đầu ra trong các dạng đồ thị và sơ đồ.

Nó không phải là cần thiết để sử dụng mục đích chung của hệ thống quản trị CSDL cho việc thực hiện một CSDL trên máy tính. Các người dùng có thể viết các chương trình riêng của họ để tạo ra, và duy trì CSDL, có hiệu lực nhằm tạo ra phần mềm quản trị CSDL với mục đích đặc biệt của riêng mình. Các cơ sở dữ liệu và phần mềm hợp lại được gọi là một hệ thống CSDL, mà người dùng đầu cuối truy cập vào hệ thống CSDL thông qua các chương trình ứng dụng và truy vấn. Các phần mềm Hệ thống quản trị CSDL cho phép người sử dụng xử lý các truy vấn và các chương trình được đặt bởi người dùng đầu cuối. Phần mềm truy cập vào dữ liệu từ cơ sở dữ liêu





Hình 1.2 minh họa một hệ thống CSDL.



Hình 1.2: A Simplified Database System Environment

1.4.1 Lợi ích của Hệ thống quản trị CSDL

Một hệ thống quản trị CSDL có trách nhiệm xử lý dữ liệu, và chuyển đổi nó thành thông tin. Với mục đích này, các cơ sở dữ liệu phải được vận dụng trong việc truy vấn cơ sở dữ liệu, để lấy dữ liệu cụ thể, cập nhật cơ sở dữ liệu, và cuối cùng, tạo ra các báo cáo



Các báo cáo đó là nguồn của thông tin, là các dữ liệu được xử lý. Một hệ thống quản trị CSDL cũng chịu trách nhiệm đối với tính bảo mật và tính toàn vẹn của dữ liệu.

Các lợi ích của một hệ thống quản trị CSDL là như sau:

Lưu trữ dữ liệu

Các chương trình cần thiết cho việc lưu trữ dữ liệu vật lý, được xử lý bởi một hệ thống quản trị CSDL, được thực hiện bằng cách tạo ra các cấu trúc dữ liệu phức tạp, và quá trình này được gọi là quản trị lưu trữ dữ liệu.

> Đinh nghĩa dữ liệu

Một hệ thống quản trị CSDL cung cấp các chức năng để xác định cấu trúc của dữ liệu trong ứng dụng. Chúng bao gồm việc xác định và thay đổi cấu trúc bản ghi, các loại và kích thước của các trường, và những ràng buộc / điều kiện khác nhau sẽ được thỏa mãn bởi các dữ liệu trong mỗi trường.

> Thao tác dữ liệu

Một khi các cấu trúc dữ liệu được định nghĩa, dữ liệu cần phải được đưa vào, chỉnh sửa hoặc xóa bỏ. Các chức năng thực hiện các bài toán này, cũng là một phần của một hệ thống quản trị CSDL.Các chức năng đó có thể xử lý các nhu cầu thao tác dữ liệu trong, hoặc ngoài ý muốn. Các truy vấn trong dự kiến là một phần của ứng dụng. Các truy vấn ngoài dự kiến là các truy vấn ad-hoc, được thực hiện trên một nhu cầu cơ bản

Bảo mật và toàn vẹn dữ liệu

Bảo mật dữ liệu là điều vô cùng quan trọng khi có nhiều người dùng truy cập vào CSDL. Nó là cần thiết cho việc thường xuyên kiểm tra các dữ liệu truy cập bởi người sử dụng. Các quy tắc bảo mật xác định quyền truy cập vào CSDL của người dùng, và các thành phần dữ liệu mà người dùng có thể truy cập, các thao tác dữ liêu mà người dùng có thể thực hiện.

Dữ liệu trong CSDL chứa càng ít lỗi càng tốt. Ví dụ, nhập số lượng nhân viên không được để trống khi thêm một nhân viên mới. Trường chứa số điện thoại chỉ được chứa chữ số. Những việc kiểm tra như vậy được làm bởi một hệ thống quản trị CSDL

Như vậy, hệ thống quản trị CSDL chứa các chức năng xử lý an ninh và tính toàn vẹn của dữ liệu trong ứng dụng. Những chức năng đó có thể dễ dàng được gọi bởi các ứng dụng, và do đó các lập trình viên ứng dụng không cần viết code cho các chức năng đó trong chương trình.

> Phục hồi dữ liệu, và tính nhất quán

Việc phục hồi dữ liệu sau khi hệ thống bị lỗi, và việc truy cập các bản ghi cùng một lúc bởi nhiều người dùng, cũng được xử lý bởi một hệ thống quản trị CSDL

> Hiệu suất

Tối ưu hóa hiệu suất của các truy vấn là một trong những chức năng quan trọng của một hệ thống quản trị CSDL. Do đó, các hệ thống quản trị CSDL có một tập hợp các chương trình tạo thành Query Optimizer, để đánh giá việc triển khai khác nhau của một truy vấn, và chọn một lựa chọn tốt nhất trong số đó





Kiểm soát truy cập đa người dùng

Tại bất kỳ thời điểm nào, cũng có nhiều hơn một người sử dụng có thể truy cập vào cùng một dữ liệu. Một hệ thống quản trị CSDL quản lý việc chia sẻ dữ liệu giữa nhiều người dùng, và duy trì tính toàn ven dữ liêu.

Ngôn ngữ truy cập CSDL, và Giao diện lập trình ứng dụng (APis)

Ngôn ngữ truy vấn của một hệ thống quản trị CSDL thực hiện việc truy cập dữ liệu. SQL là ngôn ngữ truy vấn thông dụng nhất. Một ngôn ngữ truy vấn là một ngôn ngữ phi thủ tục, nơi mà người sử dụng cần để yêu cầu những dữ liệu cần thiết, mà không cần phải xác định nó phải làm như thế nào. Một số ngôn ngữ thủ tục như C, Visual Basic, Pascal, và ..vv, cung cấp việc truy cập dữ liệu cho các nhà lập trình.

1.5 Các mô hình CSDL

Cơ sở dữ liệu có thể được phân biệt dựa trên các chức năng và mô hình dữ liệu. Một mô hình dữ liệu mô tả một container cho việc lưu trữ dữ liệu, và các quá trình lưu trữ và lấy dữ liệu từ nó. Các phân tích và thiết kế của các mô hình dữ liệu là nền tảng của sự phát triển của CSDL.

Các mô hình CSDL sẽ được thảo luận thêm trong các chương sau

1.5.1 Mô hình dữ liệu tập tin phẳng

Trong mô hình này, CSDL chỉ bao gồm một bảng, hoặc một tập tin. Mô hình này được sử dụng cho các CSDL đơn giản.

-Ví dụ, để lưu trữ Mã sinh viên, tên, môn học, và điểm của một nhóm Sinh viên. Mô hình này không thể xử lý các dữ liệu thật phức tạp. Nó có thể gây ra sự dư thừa khi dữ liệu bị lặp đi lặp lại nhiều lần. Table 1.2 mô tả cấu trúc của CSDL một tập tin phẳng

Roll Number	FirstName	LastName	Subject	Marks
45	Jones	Bill	Maths	84
45	Jones	Bill	Science	75
50	Mary	Mathew	Science	80

Table 1.2: Structure of Flat File Data Model

1.5.2 Mô hình dữ liệu phân cấp

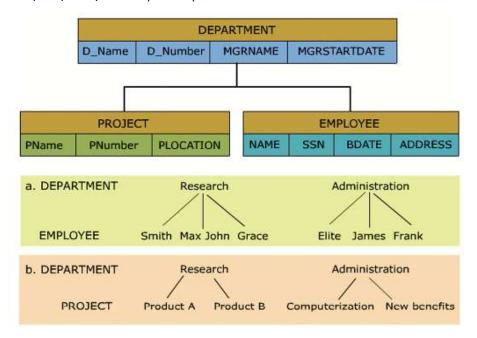
Trong mô hình phân cấp, các nút khác nhau được liên kết với nhau thông qua cấu trúc phân cấp, hoặc cấu trúc cây. Trong mô hình này, các mối liên hệ được xem như các nút cha và nút con. Một nút cha có thể có nhiều nút con, nhưng một nút con chỉ có thể có một nút cha. Để tìm dữ liệu được lưu trữ trong mô hình này, người dùng cần phải biết cấu trúc của cây

Windows Registry là một ví dụ về một cơ sở dữ liệu phân cấp lưu trữ các thiết lập cấu hình và các tùy chon trên hê điều hành Microsoft Windows.





Hình 13 minh họa một ví dụ về kiểu phân cấp.



Hình 1.3: Example of a Hierarchical Model

Trong mô hình phân cấp, Department được coi là nút cha của các phân đoạn. Các bảng, Project và Employee, là các nút con. Một nhánh từ các phân đoạn của nút cha bắt đầu từ bên trái, xác định cây. Điều này xác định trình tự sắp xếp của các phân đoạn theo các cấu trúc phân cấp được gọi là phân cấp nhánh

Có thể thấy rõ ràng từ hình ảnh minh họa, thì trong một Department duy nhất, có thể có nhiều employees (nhân viên), và một Department có thể có nhiều project (dự án)

Ưu điểm của mô hình phân cấp

Ưu điểm của mô hình phân cấp là như sau:

- Dữ liệu được lưu trữ tại một cơ sở dữ liệu chung, để việc chia sẻ dữ liệu trở nên dễ dàng hơn,
 và an ninh được cung cấp, và thực thi bởi một hệ thống quản trị CSDL.
- Dữ liệu độc lập được cung cấp bởi một hệ thống quản trị CSDL, làm giảm bớt các nỗ lực, và chi phí trong việc duy trì chương trình.

Mô hình này rất hiệu quả khi một cơ sở dữ liệu chứa một khối lượng lớn dữ liệu. Ví dụ, hệ thống tài khoản khách hàng của một ngân hàng phù hợp với mô hình phân cấp bởi vì mỗi tài khoản của khách hàng có thể có nhiều giao dịch





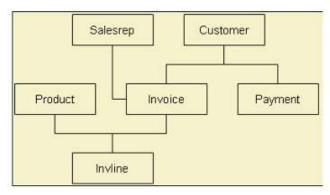
1.5.3 Mô hình dữ liệu mạng

Mô hình này cũng tương tự như mô hình dữ liệu phân cấp. Mô hình phân cấp thực sự là một tập con của mô hình mạng. Tuy nhiên, thay vì sử dụng một hệ thống phân cấp cây chỉ có duy nhất một nút cha, mô hình mạng sử dụng lý thuyết tập hợp để cung cấp một hệ thống phân cấp cây với các ngoại lệ mà các nút con được phép có nhiều hơn một nút cha.

Trong mô hình mạng, dữ liệu được lưu trữ trong các tập, thay vì trong định dạng phân cấp cây. Điều này giải quyết vấn đề dư thừa dữ liệu. Các lý thuyết tập hợp của mô hình mạng không sử dụng một hệ thống phân cấp cây chỉ có một nút cha duy nhất. Nó cho phép một nút con có nhiều hơn một nút cha. Do đó, các bản ghi có thể được liên kết thông qua danh sách được liên kết. Hệ thống quản trị CSDL tích hợp (IDMS) của Associates International Computer Inc, và Server quản trị CSDL Raima (RDM) của Raima Inc là các ví du của một hệ thống quản trị CSDL mang.

Mô hình mạng cùng với mô hình dữ liệu phân cấp là một mô hình dữ liệu quan trọng cho việc thực hiện nhiều hệ thống quản trị CSDL thương mại. Các cấu trúc mô hình mạng, và cấu trúc ngôn ngữ đã được định nghĩa tại Hội nghị Data Systems Language (CODASYL).

Đối với mỗi CSDL, một định nghĩa của tên CSDL, loại bản ghi cho từng bản ghi, và các thành phần tạo nên những bản ghi đó được lưu trữ. Điều này được gọi là sơ đồ mạng. Một phần của CSDL theo cách nhìn của các chương trình ứng dụng đó, thực sự tạo ra các thông tin như mong muốn từ các dữ liệu được chứa trong CSDL được gọi là sub-schema (sơ đồ con). Nó cho phép các chương trình ứng dụng truy cập các dữ liêu cần thiết từ cơ sở dữ liêu.



Hình 1.4: Network Model

Mô hình mạng được cho thấy trong Hình 1.4 minh họa một loạt các quan hệ một - nhiều (1-n) như sau :

- 1. Một đại diện bán hàng có thể đã viết nhiều hóa đơn, nhưng mỗi hóa đơn được viết bởi một đại diện bán hàng duy nhất (Salesrep).
- 2. Một khách hàng (Customer) có thể mua hàng vào nhiều dịp khác nhau. Một khách hàng có thể có nhiều hóa đơn, nhưng mỗi hóa đơn chỉ thuộc về một khách hàng duy nhất
- Một hóa đơn (Invoice) có thể có nhiều dòng đơn hàng (Invline), nhưng mỗi dòng đơn hàng chỉ được tìm thấy trên một hóa đơn duy nhất.
- 4. Một sản phẩm (Product) có thể xuất hiện trên nhiều dòng hóa đơn khác nhau, nhưng mỗi dòng hóa đơn chỉ chứa một sản phẩm duy nhất.





Các thành phần của ngôn ngữ được sử dụng với các mô hình mạng là như sau:

- Data Definition Language (DDL) ngôn ngữ định nghĩa dữ liệu được sử dụng để tạo, và xóa các cơ sở dữ liệu và các đối tượng của CSDL. Nó cho phép người quản trị CSDL xác định các thành phần của sơ đồ
- 2. Môt sơ đồ con DDL cho phép người quản tri CSDL xác định các thành phần của CSDL
- Data Manipulation Language (DML) Ngôn ngữ thao tác dữ liệu được sử dụng để chèn, lấy, và chỉnh sửa thông tin CSDL. Tất cả người dùng của CSDL sử dụng các lệnh đó trong các hoạt động thường ngày của CSDL.
- 4. Data Control Language (DCL) Ngôn ngữ điều khiển dữ liệu được sử dụng để quản lý các quyền trên các cơ sở dữ liệu và các đối tượng của CSDL.

Ưu điểm của mô hình mạng

Các ưu điểm của một cấu trúc như trên được xác định như sau:

- Trong mô hình mạng, các mối quan hệ nhiều nhiều có thể dễ dàng biểu diễn hơn trong mô hình phân cấp
- Mô hình này nâng cao tính toàn vẹn của CSDL
- Mô hình này bảo đảm tính độc lập của CSDL.

> Nhược điểm của mô hình mạng

Các nhược điểm được xác định như sau:

- Các CSDL trong mô hình này khó thiết kế và sử dung
- Lập trình viên phải nắm rất vững về cấu trúc CSDL, thì mới có thể làm tốt được công việc của họ
- Mô hình này cung cấp một môi trường truy cập dữ liệu theo một hướng nhất định. Do đó, để di chuyển từ A đến E trong thứ tự A-B-C-D-E, người dùng phải di chuyển qua B,C và D để tới E.

Nói chung mô hình này là khó thực hiện và duy trì. Nó chỉ thích hợp với các lập trình viên, và các nhà quản trị, hơn là cho các người dùng đầu cuối.

1.5.4 Mô hình dữ liệu quan hệ

Khi nhu cầu thông tin trở nên lớn và các CSDL phức tạp hơn, các ứng dụng đã được yêu cầu, việc thiết kế cơ sở dữ liệu, quản lý, sử dụng và trở nên quá cồng kềnh. Việc thiếu cơ sở truy vấn làm mất rất nhiều thời gian của các lập trình viên để tạo ra các báo cáo, dù là đơn giản nhất. Điều này dẫn đến sự phát triển của mô hình dữ liệu quan hệ (Relational Data Model)

Thuật ngữ 'Quan hệ' được bắt nguồn từ lý thuyết của toán học. Trong mô hình quan hệ không có các liên kết vật lý giống như trong các mô hình phân cấp, và mạng. Tất cả các dữ liệu được duy trì trong các hình thức của bảng gồm các hàng và cột. Dữ liệu trong hai bảng có liên quan thông qua các cột chung, và không liên kết vật lý. Các toán tử được cung cấp để tính toán trên các hàng trong bảng.





Các hệ thống quản trị CSDL quan hệ phổ biến là Oracle, Sybase, DB2, Microsoft SQL Server, và ...vv

Mô hình này biểu diễn CSDL như là một tập các mối quan hệ. Trong thuật ngữ của mô hình này, một hàng được gọi là một bộ dữ liệu, một cột là một thuộc tính, và bảng được gọi là một quan hệ. Danh sách của các giá trị có thể áp dụng cho một khu vực cụ thể được gọi là miền. Nó có thể cho một số thuộc tính có cùng một miền. Số lượng các thuộc tính của một mối quan hệ được gọi là degree (bậc) của mối quan hệ. Số lượng các bộ dữ liệu xác định lượng số của mối quan hệ.

Để hiểu được mô hình quan hệ, xem các table 1.3 và 1.4.

Roll Number	Student Name
1	Sam Reiner
2	John Parkinson
3	Jenny Smith
4	Lisa Hayes
5	Penny Walker
6	Peter Jordan
7	Joe Wong

Table 1.3: Students Table

Roll Number	Marks Obtained
1	34
2	87
3	45
4	90
5	36
6	65
7	89

Table 1.4: Marks Table

Bảng Students hiển thị Roll Number (Mã sinh viên) và Student Name (Tên Sinh viên), và bảng Marks hiển thị Roll Number và Marks (điểm) có được bởi các Sinh viên. Bây giờ, có hai bước cần phải được thực hiện đối với các Sinh viên đã đạt trên 50 điểm. Đầu tiên, xác định vị trí các Mã sinh viên của những người đã đạt trên 50 điểm từ bảng Marks. Thứ hai, tên của họ phải được đặt trong bảng Students bằng cách chọn các Mã sinh viên phù hợp. Kết quả sẽ được trình bày trong Table 1.5.

Roll Number	Student Name	Marks Obtained
2	John	87
4	Lisa	90
6	Peter	65
7	Joe	89

Table 1.5: Displaying Student Names and Marks





Nó có thể để lấy được thông tin này, bởi vì hai nguyên nhân. Thứ nhất là cả hai bảng đều có một cột chung – **Roll Number**. Thứ hai, dựa trên cột này, các bản ghi từ hai bảng khác nhau có thể được so sánh, và có thể có được các thông tin cần thiết.

Trong một mô hình quan hệ, dữ liệu được lưu trữ trong các bảng. Một bảng trong một cơ sở dữ liệu có một cái tên duy nhất để xác định nội dung của nó. Mỗi bảng có thể được định nghĩa như là một giao điểm của các hàng và cột.

Ưu điểm của mô hình quan hệ

Mô hình CSDL quan hệ cho lập trình viên thời gian để tập trung vào các điểm hợp lý của CSDL, chứ không phải bị làm phiền về ngoại hình. Một trong những lý do cho sự phổ biến của các CSDL quan hệ là sự linh hoạt truy vấn. Hầu hết các CSDL quan hệ sử dụng Structured Query Language (SQL), hay còn được gọi là Ngôn ngữ truy vấn có cấu trúc. Một hệ thống quản trị CSDL sử dụng SQL để biên dịch các truy vấn của người dùng vào các mã kỹ thuật cần thiết, để lấy dữ liệu theo yêu cầu. Mô hình quan hệ là rất dễ dàng sử dụng, ngay cả những người chưa qua đào tạo cũng có thể sử dụng nó để tạo ra các báo cáo, và truy vấn mà không phải tốn công suy nghĩ khi thiết kế một CSDL hợp lý.

Nhược điểm của mô hình quan hệ

Mặc dù mô hình dữ liệu quan hệ ẩn đi tất cả các sự phức tạp của hệ thống, nhưng nó có xu hướng chậm hơn so với các hệ thống cơ sở dữ liệu khác.

So với tất cả các mô hình khác, thì mô hình dữ liệu quan hệ là phổ biến nhất, và được sử dụng rộng rãi.

1.6 Hệ thống quản trị CSDL quan hệ (RDBMS)

Mô hình dữ liệu quan hệ là một nỗ lực để đơn giản hóa cấu trúc cơ sở dữ liệu. Nó biểu diễn tất cả các dữ liệu trong CSDL như các bảng đơn giản (hàng & cột) của các giá trị dữ liệu. Một hệ thống quản trị CSDL quan hệ là một chương trình phần mềm giúp bạn tạo ra, duy trì, và thao tác một cơ sở dữ liệu quan hệ. Một cơ sở dữ liệu quan hệ là một CSDL được chia thành các đơn vị logic gọi là các bảng, nơi mà các bảng có liên quan đến nhau trong cơ sở dữ liệu

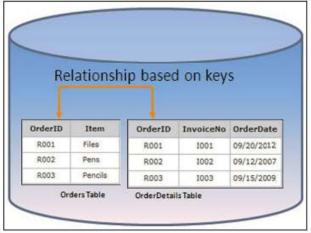
Các bảng có liên quan trong một CSDL quan hệ, cho phép lấy được đầy đủ dữ liệu trong một truy vấn duy nhất (mặc dù các dữ liệu mong muốn có thể tồn tại, trên nhiều bảng). Do có các khóa, hoặc các lĩnh vực chung trong các bảng CSDL quan hệ, nên dữ liệu từ nhiều bảng có thể được ghép lại để tạo thành một tập kết quả lớn.

Concepts





Hình 1.5 cho thấy hai bảng có liên quan với nhau bằng một khóa chung (giá trị dữ liệu) trong một CSDL quan hệ.



Hình 1.5: Relationship between Tables

Như vậy, một cơ sở dữ liệu quan hệ là một cơ sở dữ liệu được cấu trúc trên mô hình quan hệ. Các đặc trưng cơ bản của một mô hình quan hệ là trong một mô hình quan hệ, dữ liệu được lưu trữ trong các mối quan hệ. Để hiểu được các mối quan hệ, cần xem các ví dụ sau đây.

Bảng Capitals được thấy trong table 1.6 hiển thị một danh sách tên các quốc gia, và thủ đô của họ, và bảng Currency được thấy trong table 1.7 hiển thị tên các quốc gia, và tiền tệ được sử dụng.

Country	Capital
Greece	Athens
Italy	Rome
USA	Washington
China	Beijing
Japan	Tokyo
Australia	Sydney
France	Paris

Table 1.6: Capitals





Country	Currency
Greece	Drachma
Italy	Lira
USA	Dollar
China	Renminbi (Yuan)
Japan	Yen
Australia	Australian Dollar
France	Francs

Table 1.7: Currency

Cả hai bảng đều có một cột chung, đó là cột Country. Bây giờ, nếu người dùng muốn hiển thị thông tin về tiền tệ được sử dụng ở Rome, đầu tiên phải tìm tên của quốc gia của Rome. Thông tin này có thể được lấy từ Table 1.6. Tiếp theo, tên quốc gia đó nên được tìm trong Table 1.7 để tìm ra tiền tệ

Nó là khả thi để có được thông tin này, bởi vì nó có thể thiết lập một mối quan hệ giữa hai bảng thông qua một cột chung được gọi là **Country** (Quốc gia)

1.6.1 Các quy tắc liên quan đến Hệ thống quản trị CSDL quan hệ

Có những quy tắc nhất định, chủ yếu được sử dụng trong một hệ thống quản trị CSDL quan hệ. Những quy tắc đó được miêu tả như sau:

- Dữ liệu được biểu diễn như một tập các quan hệ.
- > Mỗi quan hệ được mô tả như một bảng
- Các cột là các thuộc tính
- Các hàng (bộ dữ liệu) đại diện cho các thực thể
- Mỗi một bảng có một tập các thuộc tính được gọi chung là một 'key' (khóa) duy nhất, để xác định mỗi thực thể

Ví dụ, một công ty có thể có một bảng Employee (nhân viên), với một hàng cho từng nhân viên. Có những thuộc tính nào là phù hợp cho một bảng như vậy? Điều này sẽ phụ thuộc vào các ứng dụng và các cách sử dụng dữ liệu sẽ được đưa vào, và được xác định tại thời điểm thiết kế cơ sở dữ liệu.

Xem xét kịch bản của một công ty duy trì thông tin khách hàng, và đơn đặt hàng cho các sản phẩm đang được bán, và thông tin chi tiết đặt hàng của khách hàng cho một tháng cụ thể, chẳng hạn như, tháng Tám.





Tables 1.8, 1.9, 1.10, và 1.11 được sử dụng để minh họa kịch bản này. Các bảng đó mô tả các bộ dữ liệu, và các thuộc tính trong hình thức các hàng, các cột. Các quy tắc khác nhau liên quan đến các bảng trên được đưa ra trong table 1.12

Cust_No	Cust_Name	Phone No
002	David Gordon	0231-5466356
003	Prince Fernandes	0221-5762382
003	Charles Yale	0321-8734723
002	Ryan Ford	0241-2343444
005	Bruce Smith	0241-8472198

Table 1.8: Customer

Item_No	Description	Price
HW1	Power Supply	4000
HW2	Keyboard	2000
HW3	Mouse	800
SW1	Office Suite	15000
SW2	Payroll Software	8000

Table 1.9: Items

Ord_No	ltem_No	Qty
101	HW3	50
101	SW1	150
102	HW2	10
103	HW3	50
104	HW2	25
104	HW3	100
105	SW1	100

Table 1.10 Order_Details

Ord_No	Ord_Date	Cust_No
101	02-08-12	002
102	11-08-12	003
103	21-08-12	003
104	28-08-12	002
105	30-08-12	005

Table 1.11 Order_August





Thuật ngữ	Định nghĩa	Ví dụ từ các kịch bản
Quan hệ	Một bảng	Order_August, Order_Details, Customer và Items
Bộ dữ liệu	Một hàng hoặc một bảng ghi trong một quan hệ	Một hàng từ quan hệ Customer là một bộ dữ liệu của Customer
Thuộc tính	Một trường hoặc một cột trong một quan hệ	Ord_Date, ltem_No, Cust_Name, vàvv
Cardinality (Lượng số) của một quan hệ	Số lượng của các bộ dữ liệu trong một quan hệ	Lượng số của mối quan hệ Order_Details là 7
Degree (bậc) của một quan hệ	Số lượng các thuộc tính trong một quan hệ	Bậc của quan hệ Customer là 3
Miền của một thuộc tính	Một tập tất cả các giá trị có thể được truy xuất bởi thuộc tính	Miền của Qty trong Order_Details là 1 tập của tất cả các giá trị biểu diễn số lượng của một mục được đặt mua
Primary Key (Khóa chính) của một quan hệ	Một thuộc tính hoặc, một sự kết hợp của các thuộc tính duy nhất, xác định 1bộ dữ liệu trong một mối quan hệ	Primary Key của quan hệ Customer là Cust_No Sự kết hợp của Ord_No và ltem_No tạo thành primary key của Order Details
Foreign Key (Khóa ngoài)	Một thuộc tính hoặc một sự kết hợp của các thuộc tính trong một mối quan hệ R1, để chỉ ra các mối quan hệ của R1 với một mối quan hệ R2 Các thuộc tính khóa ngoài trong R1 phải chứa các giá trị phù hợp với những giá trị trong R2	Cust_ No trong quan hệ Order_August là một foreign key (khóa ngoài) tạo ra tham chiếu từ bảng Order_August đến bảng Customer. Điều này là cần thiết để chỉ ra mối quan hệ giữa các đơn đặt hàng trong bảng Order_August và Customer

Table 1.12: Terms Related to Tables

1.6.2 Người sử dụng của hệ thống quản trị CSDL quan hệ

Mục tiêu chính của một hệ thống CSDL là cung cấp một môi trường để lấy thông tin từ CSDL, và lưu trữ thông tin mới vào đó.

Đối với một cơ sở dữ liệu cá nhân cỡ nhỏ, một người thường xác định các cấu trúc, và thao tác các cơ sở dữ liệu. Tuy nhiên, cần có nhiều người tham gia vào việc thiết kế, sử dụng, và duy trì một cơ sở dữ liệu lớn với hàng trăm người dùng.

> Database Administrator (DBA) - Người quản trị CSDL

Một DBA là một người thu thập các thông tin sẽ được lưu trữ trong cơ sở dữ liệu. Một cơ sở dữ liệu được thiết kế để cung cấp các thông tin đúng, vào đúng thời điểm cho đúng người.





Quản lý các nguồn tài nguyên này là trách nhiệm của Database Administrator (Người quản trị CSDL). DBA chịu trách nhiệm cho phép truy cập vào CSDL, điều phối và giám sát việc sử dụng của nó, và cho việc mua phần mềm, và tài nguyên phần cứng khi cần thiết. DBA có trách nhiệm đối với các vấn đề như vi phạm an ninh, hay thời gian đáp ứng của hệ thống bị chậm trễ.

> Database Designer (Người thiết kế CSDL)

Người thiết kế CSDL có trách nhiệm xác định các dữ liệu được lưu trữ trong CSDL, và cho việc lựa chọn các cấu trúc phù hợp để trình bày, và lưu trữ dữ liệu này. Đó là trách nhiệm của các nhà thiết kế CSDL để giao tiếp với tất cả các người dùng cơ sở dữ liệu tương lai, để hiểu được yêu cầu của họ, và để có được một thiết kế đáp ứng các yêu cầu.

> Các nhà phân tích hệ thống, và lập trình ứng dụng

Các nhà phân tích hệ thống xác định các yêu cầu của người dùng đầu cuối, và phát triển các thông số kỹ thuật để đáp ứng các yêu cầu này. Các nhà lập trình viên ứng dụng thực hiện các thông số kỹ thuật như các chương trình; sau đó, họ kiểm tra, gỡ lỗi, dẫn chứng, và bảo trì các giao dịch đã được xác định trước đó.

Ngoài những người thiết kế, sử dụng và quản lý một cơ sở dữ liệu, còn có những người khác có liên quan đến việc thiết kế, phát triển và hoạt động của các phần mềm hệ thống quản trị CSDL và môi trường hệ

Các nhà thiết kế và thực hiện hệ thống quản trị CSDL

Các nhà thiết kế, thiết kế và thực hiện các mô-đun, và giao diện DBMS như là một gói phần mềm. Một DBMS là một hệ thống phần mềm phức tạp bao gồm nhiều thành phần hoặc mô-đun, gồm các module để thực hiện các danh mục, ngôn ngữ truy vấn, xử lý giao diện, truy cập dữ liệu, và bảo mật. Một DBMS phải giao tiếp với các hệ thống phần mềm khác như các hệ điều hành, và các trình biên dịch cho các ngôn ngữ lập trình khác nhau.

End User (Người dùng đầu cuối)

Người dùng đầu cuối sẽ gọi một ứng dụng để tương tác với hệ thống, hoặc viết một truy vấn để dễ dàng truy xuất, sửa đổi, hoặc xóa dữ liệu.

1.7 Entities and Tables (Các thực thể & bảng)

Các thành phần của một hệ thống CSDL quan hệ là các thực thể, và các bảng, chúng sẽ được giải thích trong phần này

1.7.1 Entity (Thực thể)

Một thực thể là một người, địa điểm, điều, đối tượng, sự kiện, hoặc thậm chí là một khái niệm, mà có thế được xác định rõ ràng. Ví dụ, các thực thể trong một trường đại học là các sinh viên, giảng viên, và các khóa học.

Mỗi thực thể có một số đặc điểm được biết đến như là các thuộc tính. Ví dụ, các thực thể sinh viên có thể bao gồm các thuộc tính như số sinh viên, tên, và lớp. Mỗi thuộc tính nên được đặt tên một cách thích hợp.

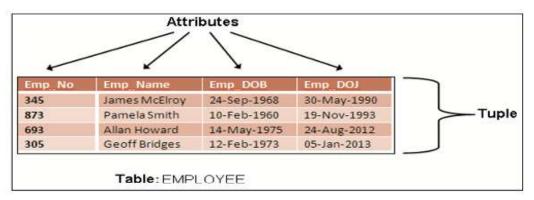




Một nhóm các thực thể có liên quan sẽ trở thành một tập thực thể. Mỗi tập thực thể được đặt một cái tên. Tên của tập thực thể phản ánh các nội dung. Do đó, các thuộc tính của tất cả các sinh viên của trường đại học sẽ được lưu trữ trong một tập thực thể gọi là Sinh viên.

1.7.2 Các bảng, và đặc điểm của chúng.

Việc truy cập và thao tác dữ liệu được hỗ trợ bởi việc tạo ra các mối quan hệ dữ liệu dựa trên một cấu trúc được gọi là một bảng. Một bảng chứa một nhóm các thực thể có liên quan đó là một tập thực thể. Thuật ngữ về tập thực thể, và bảng thường được sử dụng thay thế cho nhau. Một bảng cũng được gọi là một mối quan hệ. Các hàng được biết đến như là các bộ dữ liệu (tuples). Các cột được gọi là các thuộc tính. Hình 1.6 chỉ ra các đặc tính của một bảng.



Hình 1.6: Characteristics of a Table

Các đặc tính của một bảng là như sau:

- Một cấu trúc hai chiều gồm các hàng và cột được coi là một bảng.
- Mỗi tuple (bộ dữ liệu) đại diện cho một thực thể duy nhất trong tập thực thể
- Mỗi cột có một cái tên riêng
- Mỗi giao điểm của hàng/cột đại diện cho một giá trị dữ liệu duy nhất
- Mỗi bảng phải có một khóa chính (Primary key) duy nhất, để xác định mỗi hàng
- Tất cả các giá trị trong một cột phải phù hợp với các định dạng dữ liệu tương tự. Ví dụ, nếu thuộc tính được gán một định dạng dữ liệu số thập phân, thì tất cả các giá trị trong cột đại diện cho thuộc tính đó phải ở trong số thập phân.
- Mỗi cột có một phạm vi cụ thể của các giá trị, được biết đến như là miền thuộc tính.
- Mỗi hàng, mang thông tin mô tả một thực thể xuất hiện.
- Thứ tự của các hàng, và cột trong một hệ thống quản trị CSDL, là không quan trọng.





1.8 Sự khác biệt giữa hệ thống quản trị CSDL, và hệ thống quản trị CSDL quan hệ

Các sự khác biệt giữa một DBMS và một RDBMS được liệt kê trong table 1.13.

DBMS	RDBMS
Nó không cần phải có dữ liệu trong cấu trúc bảng, và nó cũng không thực thi các mối bảng hệ bảng giữa các mục dữ liệu.	Trong một RDBMS, cấu trúc bảng là phải có, và các mối quan hệ bảng được thực thi bởi hệ thống. Các mối quan hệ này cho phép người dùng áp dụng và quản lý các quy tắc kinh doanh với mã hóa tối thiểu
Số lượng nhỏ dữ liệu có thể được lưu trữ, và truy xuất	RDBMS có thể lưu trữ và truy xuất số lượng lớn dữ liệu.
DBMS kém an toàn hơn, so với RDBMS.	RDBMS an toàn hơn so với DBMS.
Nó là hệ thống một người dùng duy nhất	Nó là hệ thống đa người dùng
Most DBMSs do not support client/server architecture.	Hỗ trợ cấu trúc client/server .

Table 1.13: Difference between DBMS and RDBMS

Trong một RDBMS (hệ thống quản trị CSDL quan hệ), một mối quan hệ được xem là quan trọng hơn. Như vậy, các bảng trong một RDBMS là phụ thuộc, và người dùng có thể thiết lập các ràng buộc toàn vẹn khác nhau trên các bảng đó, để các dữ liệu cuối cùng được sử dụng bởi người dùng vẫn chính xác. Trong trường hợp của một DBMS (hệ thống quản trị CSDL), các thực thể được coi là quan trọng hơn, và không có mối quan hệ nào được thiết lập giữa các thực thể.





Network model

Flat File model

(A)

(B)

(A)	Flat File – Tệp tin phẳng	(C)	Network – Mang
(B)	Hierarchical – Phân cấp	(D)	Relational – Quan hệ
của CS		các qu	yền trên các CSDL và các đối t
(A)	Data Definition Language (DDL)	(C)	Sub-schema – Lược đồ con
_	Data Manipulation Language (DML) thuật ngữ của mô hình quan hệ, mộ ột được gọi là, và mộ		g được gọi là một
Trong	thuật ngữ của mô hình quan hệ, mớ	t hàng	g được gọi là một
Trong một c	thuật ngữ của mô hình quan hệ, mớ ột được gọi là, và một	ot hàng bảng	được gọi là
Trong một c (A) (B)	thuật ngữ của mô hình quan hệ, mớ ột được gọi là, và một attribute,tuple,relation	t hàng bảng (C) (D)	g được gọi là một
Trong một c (A) (B)	thuật ngữ của mô hình quan hệ, mớ ột được gọi là, và mộ attribute,tuple,relation tuple, attribute,relation có thể được định nghĩ	t hàng bảng (C) (D)	g được gọi là một

(C)

(D)

Data model

Relational model





1.9.1 Đáp án

1.	С
2.	D
3.	В
4.	В
5.	С

Concepts



RDBMS Concepts





- Một cơ sở dữ liệu là một tập dữ liệu có liên quan trong hình thức của một bảng
- Một mô hình dữ liệu mô tả một container cho việc lưu trữ và quá trình lưu trữ, và truy xuất dữ liệu từ container đó.
- Một DBMS (hệ thống quản trị CSDL) là một tập các chương trình cho phép người dùng lưu trữ, chỉnh sửa, và lấy thông tin từ một CSDL.
- Một hệ thống quản trị CSLD quan hệ (RDBMS) là một bộ các chương trình phần mềm để tạo, duy trì, sửa đổi, và thao tác một CSDL quan hệ.
- Một CSDL quan hệ được phân thành các đơn vị logic gọi là các bảng. Các đơn vị logic đó liên kết với nhau bên trong CSDL
- Các thành phần chính của một hệ thống CSDL quan hệ là các thực thể và các bảng
- Trong một RDBMS, một quan hệ là có tầm quan trọng hơn. Còn trong trường hợp của một DBMS, thì các thực thể được coi là quan trọng hơn, và không có bất kỳ mối quan hệ nào được thiết lập giữa chúng

