

BÀI TẬP LỚN MÔN KHO DỮ LIỆU

Mô tả yêu cầu của bài toán: Thiết kế và cài đặt một kho dữ liệu cho một hệ thống xử lý đặt hàng của khách hàng cho một công ty sử dụng MS SQL Server và Oracle với những thông tin như sau.

Hệ thống kho dữ liệu được xây dựng cho một doanh nghiệp bao gồm nhiều cửa hàng nằm rải rác ở nhiều thành phố và bang khác nhau. Mỗi cửa hàng có một số lượng các mặt hàng với số lượng hàng khác nhau. Thêm nữa, doanh nghiệp cũng lưu trữ thông tin về các khách hàng bao gồm hai loại: loại thứ nhất là khách du lịch thường được dẫn đến bởi các hướng dẫn viên các đoàn du lịch, loại hai bao gồm những khách hàng đặt hàng qua đường bưu điện. Thành phố sinh sống của các khách hàng, cùng với dữ liệu về đơn đặt hàng đầu tiên của khách hàng được lưu lại trong hệ thống hiện tại. Mỗi khách hàng chỉ sống ở một thành phố và doanh nghiệp đó sẽ cố gắng đáp ứng các mặt hàng trong đơn hàng của khách bằng cách lấy từ kho ở thành phố mà khách hàng đó sinh sống. Mỗi đơn đặt hàng của khách có thể đặt thoải mái về số lượng và số các mặt hàng và mỗi đơn đặt hàng được xác định bởi một mã số đơn đặt hàng.

Vị trí của các cửa hàng cũng được lưu trữ. Mỗi cửa hàng nằm tại một thành phố và có nhiều cửa hàng trong một thành phố. Mỗi thành phố có một văn phòng đại diện để quản lý tất cả các cửa hàng của thành phố đó. Mục tiêu của doanh nghiệp này là đáp ứng đầy đủ các yêu cầu của khách hàng từ những cửa hàng nằm trong thành phố mà khách hàng sinh sống. Nếu những yêu cầu đó không thể đáp ứng thì công ty sẽ chuyển yêu cầu tới các thành phố đến khi mặt hàng có đủ để cung cấp.

Một số thông tin xử lý là rất quan trọng đối với doanh nghiệp này. Ví dụ, tổng số lượng mỗi mặt hàng được lưu trữ tại mỗi thành phố. Sau mỗi lần một mặt hàng được phân phối (được bán), doanh nghiệp này cần biết tổng số lượng mỗi mặt hàng còn lại tại tất cả các cửa hàng trong một thành phố.

Lược đồ quan hệ của hệ thống cơ sở dữ liệu quan hệ hiện tại của doanh nghiệp này như sau:

Cơ sở dữ liệu Văn phòng đại diện bao gồm những quan hệ với các lược đồ như sau:

Khách hàng (Mã KH, Tên KH, Mã Thành phố, Ngày đặt hàng đầu tiên)

Khách hàng du lịch (*Mã KH, Hướng dẫn viên du lịch, Thời gian)

Khách hàng bưu điện (* Mã KH, Địa chỉ bưu điện, Thời gian)

Cơ sở dữ liệu Bán hàng gồm những quan hệ với lược đồ như sau:

Văn phòng đại diện (Mã Thành phố, Tên Thành phố, Địa chỉ VP, Bang, Thời gian)

Cửa hàng (Mã cửa hàng, * Mã Thành phố, Số điện thoại, Thời gian)

Mặt hàng (Mã MH, Mô tả, Kích cỡ, Trọng lượng, Giá, Thời gian)

Mặt hàng_được lưu trữ (* Mã cửa hàng, * Mã mặt hàng, Số lượng trong kho, Thời gian)

Đơn đặt hàng (Mã đơn, Ngày đặt hàng, Mã Khách hàng)

Mặt hàng được đặt (* Mã đơn, * Mã mặt hàng, Số lượng đặt, Giá đặt, Thời gian)

Trong đó những thuộc tính được gạch chân là khóa chính và thuộc tính đánh dấu (*) là khóa ngoại.

Yêu cầu nghiệp vụ:

Để đáp ứng được nhu cầu của khách hàng, hệ thống kho dữ liệu trích lọc dữ liệu từ hai cơ sở dữ liệu hiện có để cho vào kho dữ liệu và cung cấp các xử lý phân tích trực tuyến với các thao tác cuộn lên, khoan xuống, chọn và chiếu dựa trên những yêu cầu chọn lựa của khách hàng và các bảng theo chiều để đáp ứng yêu cầu của người sử dụng. Thiết lập một chiều thời gian, hệ thống sẽ sinh một báo cáo Phân tích trực tuyến cho những yêu cầu sau:

1. Tìm tất cả các cửa hàng cùng với thành phố, bang, số điện thoại, mô tả, kích cỡ, trọng lượng và đơn giá của tất cả các mặt hàng được bán ở kho đó.
2. Tìm tất cả các đơn đặt hàng với tên khách hàng và ngày đặt hàng được thực hiện bởi khách hàng đó
3. Tìm tất cả các cửa hàng cùng với tên thành phố và số điện thoại mà có bán các mặt hàng được đặt bởi một khách hàng nào đó
4. Tìm địa chỉ văn phòng đại diện với tên thành phố, bang của tất cả các cửa hàng lưu kho một mặt hàng nào đó với số lượng trên mức cụ thể.
5. Với mỗi một đơn đặt hàng của khách, liệt kê các mặt hàng được đặt cùng với mô tả, mã cửa hàng, tên thành phố và các cửa hàng có bán mặt hàng đó.
6. Tìm thành phố và bang mà một khách hàng nào đó sinh sống
7. Tìm mức độ tồn kho của một mặt hàng cụ thể tại tất cả các cửa hàng ở một thành phố cụ thể nào đó
8. Tìm các mặt hàng, số lượng đặt, khách hàng, cửa hàng và thành phố của một đơn đặt hàng.
9. Tìm các khách hàng du lịch, khách hàng đặt theo đường bưu điện và khách hàng thuộc cả hai loại

Định dạng tài liệu báo cáo:

1. Giới thiệu – Mục tiêu và phạm vi của bài tập lớn

2. Yêu cầu nghiệp vụ - nêu các đặc tả ứng dụng của kho dữ liệu cho người sử dụng
3. Đặc tả chức năng – đặc tả đầu vào và đầu ra của kho dữ liệu
4. Thiết kế kho dữ liệu – phương pháp luận cho các bước thiết kế bao gồm cả lược đồ hình sao
5. Cài đặt các khối dữ liệu – tự động bằng máy tính cài đặt kho dữ liệu và tải dữ liệu vào các khối.
6. Các báo cáo phân tích trực tuyến – nhúng các lệnh hoặc thanh lệnh để sinh báo cáo OLAP
7. Kiểm tra tính đúng đắn của dữ liệu – kiểm tra các báo cáo OLAP với nguồn dữ liệu từ các bảng quan hệ
8. Kết luận – tổng kết những công việc đã hoàn thành và đạt được của bài tập lớn.

Cách tính điểm:

Tích hợp lược đồ thành một mô hình thực thể liên kết mở rộng (20%)

Thiết kế lược đồ hình sao (20%)

Các báo cáo xử lý phân tích trực tuyến (20%)

Chạy thử nghiệm công nghệ kho dữ liệu và xử lý phân tích trực tuyến (20%)

Tài liệu báo cáo (20%)