



LUẬN VĂN TỐT NGHIỆP

**XÂY DỰNG VÀ QUẢN LÝ
DATAWAREHOUSE LĨNH VỰC
CHỨNG KHOÁN**

LỜI CẢM ƠN

Chúng em xin gửi lời cảm ơn chân thành đến Ban giám hiệu trường ĐH Kỹ Thuật Công Nghệ Thành Phố Hồ Chí Minh , đặc biệt là Cô, KS. Phạm Thị Lan Anh đã tận tình hướng dẫn và tạo điều kiện cho chúng em có được thành tích quý báu như ngày hôm nay.

Nhờ sự giúp đỡ và hướng dẫn của quý nhà trường và Cô mà chúng em đã có được những kiến thức bổ ích cũng như tinh thần, kinh nghiệm thực tế cho công việc của chúng em sau này.

Ngoài ra, xin được gửi lời cảm ơn sự giúp đỡ của:

- Anh Đặng Vương Anh, trưởng phòng IT, công ty Chứng khoán Sài Gòn Bảo Tín.
- Anh Hoàng Công Nguyên Vũ, trưởng phòng Lưu ký, công ty Chứng Khoán Sài Gòn Bảo Tín.
- Anh Phạm Công Khoa Hương, Trưởng phòng Giám sát thông tin, Sở Giao Dịch Chứng Khoán Thành Phố.

Và các anh chị thuộc Công ty chứng khoán Sài Gòn Bảo Tín đã tạo điều kiện và giúp đỡ chúng em về kiến thức chuyên môn cũng như nghiệp vụ chứng khoán.

Xin gửi lời cảm ơn chân thành đến Ba, Mẹ là những người đã có công nuôi dạy chúng con nên người và có được thành quả tốt đẹp như ngày nay. Cảm ơn tất cả những người bạn thân yêu đã luôn giúp đỡ, chia sẻ nhưng kinh nghiệm bổ ích và những lời động viên, đóng góp ý kiến quý báu của các bạn.

Chúng em xin hứa sẽ cố gắng nhiều hơn nữa để không phụ lòng cũng như niềm hi vọng từ phía những người thân, Ba, Mẹ, Thầy, Cô, Bạn Bè,... Để phần đầu trở

thành những công dân tốt, những kỹ sư giỏi với vốn kiến thức và những kinh nghiệm quý báu mà chúng em có được sau hơn 4 năm học đại học ở trường.

Cũng xin cảm ơn quý trường ĐH Kỹ Thuật Công Nghệ Thành Phố Hồ Chí Minh đã tạo mọi điều kiện cho chúng em có được thời gian thuận lợi và hợp lý nhất, để tập trung hoàn toàn cho bài luận văn tốt nghiệp, giúp chúng em hoàn thành tốt nhiệm vụ được giao.

LỜI NÓI ĐẦU

Để quán triệt nguyên tắc " **HỌC ĐI ĐÔI VỚI HÀNH** ", nhằm tạo điều kiện cho sinh viên làm quen với thực tế, có được cái nhìn tổng hợp, giúp cho sinh viên chúng em hạn chế được sự ngỡ ngàng khi ra trường, hàng năm nhà trường tổ chức cho sinh viên đi thực tập tại các cơ sở theo nội dung ngành nghề đào tạo. Đây là điều kiện thuận lợi giúp chúng em phát huy được năng lực bản thân cũng như khả năng áp dụng lý thuyết được trang bị ở trường vào thực tế.

Công nghệ thông tin đã có những bước phát triển mạnh mẽ. Máy tính điện tử không còn là phương tiện quý hiếm mà đang ngày một gần gũi với con người.

Đứng trước sự bùng nổ thông tin, các tổ chức và các doanh nghiệp đều tìm mọi biện pháp để xây dựng hoàn thiện hệ thống thông tin của mình nhằm tin học hoá các hoạt động tác nghiệp của đơn vị mình. Mức độ hoàn thiện tùy thuộc vào quá trình phân tích và thiết kế hệ thống.

Từ nhu cầu nêu trên, trong thời gian thực tập tốt nghiệp chúng em đã sử dụng vốn kiến thức ít ỏi của mình tìm hiểu và phân tích bài toán **Xây dựng và Quản lý DataWarehouse lĩnh vực chứng khoán**. Đề tài gồm các phần:

Chương I: Tổng quan

Chương II: Các kỹ thuật lập trình

Chương III: Giới thiệu bài toán và phân tích ứng dụng

Chương IV: Thiết kế ứng dụng

Chương V: Kết quả thực nghiệm.

Chương VI: Tổng kết

Tuy đã rất cố gắng học hỏi dựa trên kiến thức đã học và thực tế tại công ty nhưng do khả năng và thời gian có hạn nên báo cáo của chúng em không thể tránh khỏi những thiếu sót. Chúng em kính mong quý Thầy cô cùng bạn bè thông cảm và góp ý để chúng em kịp thời lấp kín những lỗ hổng kiến thức và chương trình đạt hiệu quả cao hơn.

Chúng em xin chân thành cảm ơn **Cô Phạm Thị Lan Anh** đã hết lòng chỉ bảo để chúng em hoàn thành đồ án này.

TP Hồ Chí Minh, tháng 12 / 2008

Sinh viên thực hiện

Bùi Quốc Vũ – Lê Hà Kha

MỤC LỤC

Tóm tắt nội dung luận văn	1
Chương I: Tổng quan	
I.1. Mô hình công ty chứng khoán	2
I.2. Các vấn đề tồn tại ở các công ty chứng khoán Việt Nam	5
I.3. Mục tiêu đề ra.....	5
Chương II: Các kỹ thuật lập trình	
II.1. Giới thiệu bộ quản trị database Microsoft SQL Server 2005.....	6
II.2. Giới thiệu DataWarehouse	10
II.2.1. Khái niệm.....	10
II.2.2. Lợi ích của kho dữ liệu	12
II.2.3. Thành phần Datawarehouse	13
II.3. Crystal Report Built-in Visual Studio 2005	15
Chương III: Giới thiệu bài toán và phân tích ứng dụng	
III.1. Phát biểu bài toán	17
III.2. Phân tích Use case.....	19
III.2.1. Chức năng Backup Database	19
III.2.2 Restore database.....	20
III.2.3 Dữ liệu từ hai sàn HASTC, HOSTC đổ về	
Stock_CurrentStockInfo_Full.....	22

III.2.4 Các flat file từ phantichcophieu.vn sẽ được import vào STSWEB.....	23
III.2.5 Trích xuất báo cáo, người dùng có thể tùy biến các báo cáo thông qua các điều kiện nhập vào.....	25
Chương IV. Thiết kế ứng dụng	
IV.1. Cơ sở dữ liệu STSWEB.....	26
IV.1.1. Lược đồ quan hệ	26
IV.1.2. Mô tả database	28
IV.1.3. Ràng buộc toàn vẹn	29
IV.1.4. Quy định	30
IV.2. STSWEB Report	31
IV.2.1 Mô tả chức năng.....	31
IV.2.2 Mô tả xử lý.....	32
IV.2.3 Data Flow.....	32
Chương V. Kết quả thực nghiệm	
V.1. Chức năng Backup Database.....	33
V.1.1 Cấu trúc chương trình	33
V.1.2 Các bước chạy chương trình	35
V.2. Chức năng Restore Database.....	37
V.2.1. Cấu trúc chương trình	37
V.2.2 Các bước chạy chương trình	38
V.3. Mô phỏng dữ liệu đổ từ hai sàn HOSTC và HASTC	40

V.3.1. Cấu trúc chương trình	40
V.3.2. Các bước chạy chương trình	41
V.4. Flat file từ phantichcophieu.vn import vào STSWEB	42
V.4.1 Cấu trúc chương trình	42
V.4.2 Các bước chạy chương trình	44
V.5. Report STSWEB	46
V.5.1. Cấu trúc chương trình	46
V.5.2 Các bước chạy chương trình	46
Chương VI: Tổng kết	
VI.1. Kết luận	53
VI.1.1 Về mặt lý thuyết.....	53
VI.1.2 Về mặt ứng dụng.....	53
VI.2. Hướng phát triển.....	54
Tài liệu tham khảo	55

TÓM TẮT NỘI DUNG LUẬN VĂN

Chương I. Tổng quan: mô hình công ty, các vấn đề tồn tại thường gặp trong database chứng khoán.

Chương II. Các kỹ thuật lập trình: Giới thiệu kỹ thuật lập trình, các công nghệ mới trên SQL Server 2005 và Crystal Report Built-in Visual Studio 2005, các khái niệm cơ bản của Data Warehouse.

Chương III. Giới thiệu bài toán và phân tích ứng dụng: : Mô tả bài toán, phân tích các use case. Trình bày các sơ đồ tuần tự và một số chức năng nổi bật của ứng dụng.

Chương IV. Thiết kế ứng dụng: Mô tả cơ sở dữ liệu của ứng dụng và trình bày sơ đồ màn hình trong ứng dụng.

Chương V. Kết quả thực nghiệm: Kết quả thực nghiệm trong quá trình chạy ứng dụng, màn hình làm việc cùng kết quả tương ứng.

Chương V. Tổng kết: tổng kết kết quả thu được về mặt lý thuyết và ứng dụng, hướng phát triển trong thời gian tới

Chương I

TỔNG QUAN

I.1 Mô hình công ty chứng khoán:

Công ty CP Chứng Khoán Sài Gòn Bảo Tín, được thành lập vào tháng 11 năm 2008, vốn điều lệ hơn 300 tỉ đồng, là công ty trực thuộc Công ty cổ phần liên hiệp vận chuyển GEMADEPT. STSI được thực hiện đầy đủ các nghiệp vụ kinh doanh, bao gồm:

- Môi giới chứng khoán
- Tư vấn đầu tư chứng khoán
- Bảo lãnh phát hành
- Tự doanh chứng khoán
- Lưu ký chứng khoán.

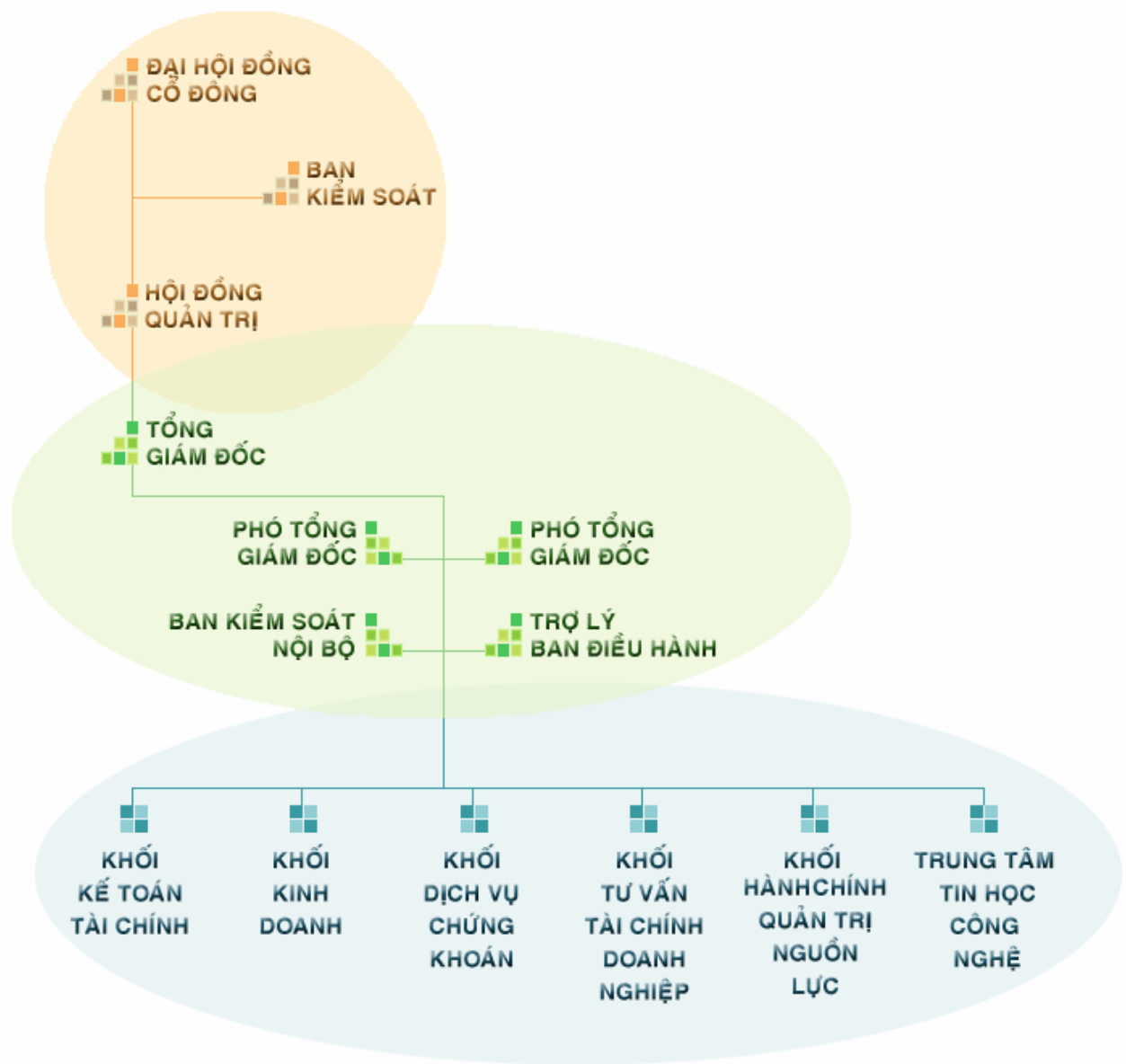
Khách hàng của STSI là các nhà đầu tư chứng khoán, bao gồm khách hàng cá nhân và khách hàng doanh nghiệp. Các khách hàng này có thể thực hiện các tác vụ như: đặt lệnh mua, bán, cầm cố, vay chứng khoán. Để thực hiện các tác vụ này, khách hàng phải đáp ứng các yêu cầu sau:

- Có tài khoản tiền gửi giao dịch chứng khoán tại các ngân hàng mà STSI liên kết.
- Có một tài khoản giao dịch tại STSI.

Khi đáp ứng tất cả các yêu cầu trên, khách hàng có thể lựa chọn một trong các cách giao dịch sau:

- Giao dịch dựa vào chức năng Giao dịch trực tuyến tại website của STSI
- Giao dịch tại sàn giao dịch của STSI
- Giao dịch qua tin nhắn SMS đến hệ thống tổng đài của STSI.
- Giao dịch qua điện thoại với các giao dịch viên của STSI.

Sơ đồ tổ chức:



Hình 1.1 Sơ đồ tổ chức công ty Chứng Khoán Sài Gòn Bảo Tín.

I.2 Các vấn đề tồn tại ở các công ty chứng khoán Việt Nam:

- Các công ty Chứng khoán tại Việt Nam đa phần không có nơi lưu trữ lại lịch sử giao dịch phục vụ nhu cầu tham khảo, đối chiếu sau này.
- Thông thường, các công ty chứng khoán sử dụng phần mềm backup và restore database của hãng thứ 3 và không tận dụng được các tính năng có sẵn của Hệ quản trị cơ sở dữ liệu. Chi phí cho việc sử dụng các phần mềm rất lớn. Cách sử dụng cũng hết sức phức tạp.
- Các báo cáo kết quả giao dịch đa phần chưa được tự động hóa, hầu như do đội ngũ kế toán tổng kết thông qua các phiếu lệnh và kết quả giao dịch từ sàn trả về.

I.3 Mục tiêu đề ra:

- Nắm bắt tất cả các nghiệp vụ chứng khoán liên quan đến các vấn đề tồn tại ở công ty.
- Giải quyết được các vấn đề tồn tại trong hệ thống database của công ty.

Chương II

CÁC KỸ THUẬT LẬP TRÌNH

II.1 Giới thiệu bộ quản trị database Microsoft SQL Server 2005:

SQL Server 2005 là một hệ thống quản lý cơ sở dữ liệu (Relational Database Management System (RDBMS)) sử dụng Transact-SQL để trao đổi dữ liệu giữa Client computer và SQL Server computer. Một RDBMS bao gồm databases, database engine và các ứng dụng dùng để quản lý dữ liệu và các bộ phận khác nhau trong RDBMS.

SQL Server 2005 được tối ưu để có thể chạy trên môi trường cơ sở dữ liệu rất lớn (Very Large Database Environment) lên đến Tera-Byte và có thể phục vụ cùng lúc cho hàng ngàn user. SQL Server 2005 có thể kết hợp "ăn ý" với các server khác như Microsoft Internet Information Server (IIS), E-Commerce Server, Proxy Server...



Hình II.1 Màn hình đăng nhập SQL Server.

Khi SQL Server 2005 ra đời, thì những kiến thức về SQL 2000 thực sự không đủ để làm chủ được SQL 2005. Khả năng và độ linh hoạt của SQL Server 2005 được phản ánh rất rõ trong công cụ Management Studio mới và BIDS. Các tính năng nổi bật trong bộ Microsoft SQL Server 2005:

a. Nâng cao bảo mật: Bảo mật là trọng tâm chính cho những tính năng mới trong SQL Server 2005. Điều này phản ánh sự phản ứng lại của Microsoft với sâu máy tính Slammer đã tấn công SQL Server 2000. Nó cũng cho thấy một thế giới ngày càng có nhiều dữ liệu kinh doanh có nguy cơ bị lộ ra ngoài Internet.

- **Bảo mật nhóm thư mục hệ thống:** Nhóm mục hệ thống bao gồm các View bên dưới cấu trúc dữ liệu hệ thống. Người sử dụng không thấy được bất cứ bảng bên dưới nào, vì thế những người dùng không có kỹ năng hoặc có ý phá hoại không thể thay đổi hoặc làm hư hỏng các bảng này được. Điều này ngăn bạn hoặc bất kỳ ai khác làm hỏng cấu trúc chính mà SQL Server phụ thuộc vào.

- **Bắt buộc chính sách mật khẩu:** Khi bạn cài Window Server 2003, bạn có thể áp dụng chính sách mật khẩu của Window (bạn đang áp dụng) cho SQL Server 2005. Bạn có thể thi hành chính sách về mức độ và ngày hết hạn của mật khẩu trên SQL Server 2005 giống hệt như cho tài khoản đăng nhập vào Windows mà trong 2000 không hỗ trợ tính năng này. Bạn có thể tắt hoặc mở việc bắt buộc chính sách mật khẩu cho từng đăng nhập riêng.

b. Mở rộng T-SQL:

- **Transact – SQL:** là một phiên bản của Structured Query Language (SQL), được dùng bởi SQL Server 2005. Transact-SQL thường được gọi là T-SQL. T-SQL có nhiều tính năng do Microsoft phát triển không có trong ANSI SQL (SQL chuẩn).

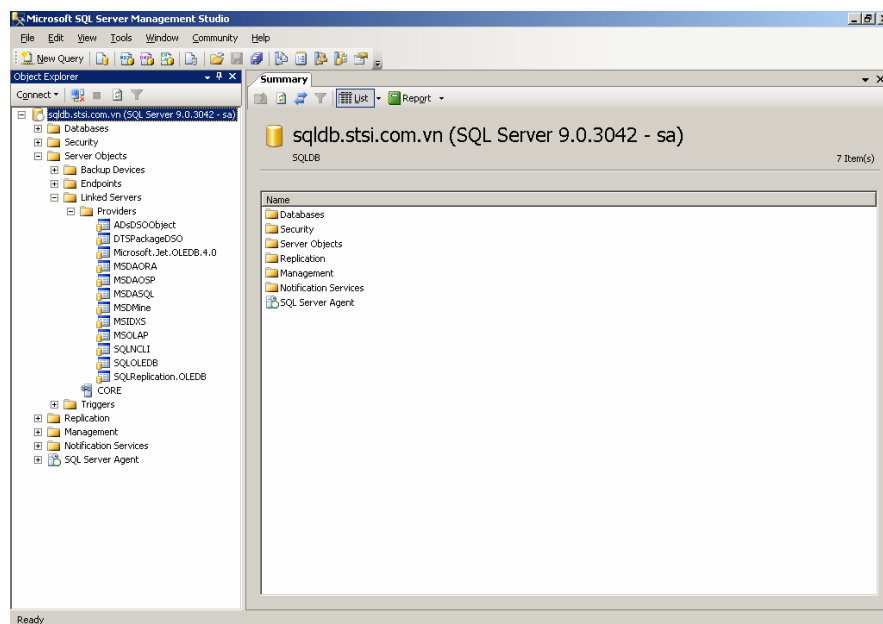
- **Cải tiến khả năng hỗ trợ XML:** SQL Server 2000 cho phép bạn nhận dữ liệu quan hệ ở dạng XML với mệnh đề FOR XML, hoặc lưu trữ XML như dữ liệu quan hệ trong SQL Server sử dụng mệnh đề OPEN XML. SQL Server 2005 có thêm một kiểu dữ liệu mới là XML cho phép bạn viết mã nhận dữ liệu XML như là XML, tránh việc biến đổi từ XML thành dữ liệu quan hệ khi dùng OPEN XML. Bạn cũng có thể dùng tài liệu giản đồ biểu diễn trong ngôn ngữ W3C XML Schema Definition (đôi khi gọi là giản đồ XSD) để chỉ ra cấu trúc hợp lệ trong XML.
- **Sử dụng khối Try... Catch trong mã T-SQL:** cho phép chỉ ra điều gì phải làm khi lỗi xảy ra.
- Trong SQL Server Management Studio, có thể tìm thấy nhiều đoạn mã mẫu giúp bạn thực hiện những tác vụ thường gặp với T-SQL.

c. Tăng cường hỗ trợ người phát triển:

- **Hỗ trợ cho Common Language Runtime (CLR):** CLR được dùng bởi mã .NET, được nhúng vào trong cỗ máy CSDL SQL Server 2005, giúp developer có thể viết các thủ tục lưu sẵn, trigger, hàm, tính toán tập hợp và các kiểu dữ liệu do người dùng định nghĩa bằng cách sử dụng các ngôn ngữ như VB.NET hoặc C#.
- **Các kiểu dữ liệu mới:** Varchar(max) cho phép bạn dùng chuỗi kí tự lớn hơn 8000 byte (8000 kí tự), tối đa là 2 GB. Nvarchar(max) cho phép dùng chuỗi kí tự Unicode lớn hơn 8000 byte (4000 kí tự), tối đa là 2 GB. Varbinary(max) cho phép dùng dữ liệu nhị phân lớn hơn 8000 byte....
- **SQL Management Object (SMO):** SMO thay thế cho Distributed Management Objects (DMO) được dùng trong SQL Server 2000. SMO nhanh hơn DMO ở nhiều thiết lập bởi vì mỗi đối tượng chỉ được thực hiện từng phần.
- **Tự động thực thi mã kịch bản:** Nếu bạn đã dùng các chương trình của Microsoft như Microsoft Access, Excel, bạn biết rằng có thể tạo các macro (mã

thực thi) cho phép bạn thực hiện tự động một số tác vụ nào đó. SQL Server 2005 bây giờ có tính năng tự động tạo mã kịch bản T-SQL từ những hành động mà bạn dùng giao diện hình ảnh trong SQL Server Management Studio.

- **Truy cập HTTP:** Dùng giao thức HTTP để truy cập vào SQL Server 2005 là tính năng mới cho phép người lập trình truy cập vào SQL Server mà không phụ thuộc vào việc IIS có đang chạy trên cùng máy hay không. SQL Server có thể cùng tồn tại với IIS nhưng không giống với SQL Server 2000, IIS không còn là yêu cầu bắt buộc với SQL Server 2005. Truy cập HTTP cho phép phát triển dung XML Web Service với SQL Server 2005. Truy cập HTTP có thể thực thi nhóm lệnh T-SQL hoặc thủ tục lưu sẵn. Tuy nhiên, vì lí do bảo mật truy cập, HTTP mặc định sẽ bị vô hiệu hóa. Để sử dụng truy cập HTTP bạn phải chỉ rõ người dùng, thủ tục lưu sẵn và CSDL được phép hỗ trợ nó.



Hình II.2 Giao diện SQL Server 2005

- **Hỗ trợ đầy đủ Linked Server:** Nhờ Linked Server, việc thao tác dữ liệu trên các Data Source khác trở nên hết sức dễ dàng và thuận tiện.

II.2 Giới thiệu DataWarehouse:

II.2.1: Khái niệm:

Kho dữ liệu về bản chất cũng là một database bình thường, các hệ Quản trị cơ sở dữ liệu quản lý và lưu trữ nó như các database thông thường (tuy nhiên có hỗ trợ thêm về quản lý dữ liệu lớn và truy vấn). Thực ra nét khác biệt của datawarehouse so với database là ở quan niệm, cách nhìn vấn đề:

- Trước tiên Datawarehouse là database rất lớn, gọi một cách chính xác hơn là kho thông tin (information warehouse), là một cơ sở dữ liệu hướng đối tượng được thiết kế với việc tiếp cận các ý kiến trong mọi lĩnh vực kinh doanh. Nó cung cấp các công cụ để đáp ứng thông tin cần thiết cho các nhà quản trị kinh doanh tại mọi cấp độ tổ chức - không những chỉ là những yêu cầu dữ liệu phức hợp, mà còn là điều kiện thuận tiện nhất để đạt được việc lấy thông tin nhanh, chính xác. Một kho dữ liệu được thiết kế để người sử dụng có thể nhận ra thông tin mà họ muốn có và truy cập đến bằng những công cụ đơn giản.
- Datawarehouse là một tập hợp thông tin cơ bản trên máy tính mà chúng có tính quyết định đến việc thực hiện thành công bước đầu trong công việc kinh doanh, là một sự pha trộn của nhiều công nghệ, bao gồm các cơ sở dữ liệu đa chiều và mối quan hệ giữa chúng, kiến trúc chủ khách, giao diện người dùng đồ họa và nhiều nữa. Dữ liệu trong kho dữ liệu không giống dữ liệu của hệ điều hành là loại chỉ có thể đọc nhưng không chỉnh sửa được. Hệ điều hành tạo ra, chỉnh sửa và xóa những dữ liệu sản xuất mà những dữ liệu này cung cấp cho kho dữ liệu. Nguyên nhân chính cho sự phát triển một kho dữ liệu là hoạt động tích hợp dữ liệu từ nhiều nguồn khác nhau vào một kho dữ liệu đơn lẻ và dày đặc mà kho này cung cấp cho việc phân tích và ra quyết định trong công việc kinh doanh. Đối với một số công việc kinh doanh tin rằng thông tin là nguồn tài nguyên có giá trị rất lớn thì một kho dữ liệu tương đối giống như một nhà kho chứa hàng. Hệ điều hành tạo ra những phần dữ liệu và nạp chúng vào

kho. Một số phần được tóm tắt trong thành phần thông tin và được cất vào kho. Người sử dụng kho dữ liệu đưa ra những yêu cầu và được cung cấp sản phẩm được tạo ra từ các thành phần và các phân đoạn được lưu trong kho.

- Database hướng về xử lý thời gian thực, datawarehouse hướng về tính ổn định.
- Database phục vụ xử lý transaction, cập nhật. Datawarehouse thường chỉ đọc, phục vụ cho những nhu cầu báo cáo. VD: Chúng ta sẽ yêu cầu hãy cho biết trong 5 năm, bộ phận phần mềm đã làm được những dự án nào, từ đó chúng ta sẽ có quyết định về hiệu năng của nhóm này .
- Datawarehouse sẽ lấy thông tin có thể từ nhiều nguồn khác nhau: DB2, Oracle, SQL Server thậm chí cả File thông thường, làm sạch chúng và đưa vào cấu trúc của nó. Datawarehouse rất lớn, nên muốn cho từng bộ phận chuyên biệt người sử dụng cuối cùng có thể khai thác thông tin dễ dàng thì bản thân datawarehouse phải được chuyên hoá, phân ra thành những chủ đề, do đó những chủ đề chuyên môn hóa đó tạo thành một Database chuyên biệt-đó là Data mart. Ví dụ: Datawarehouse của Microsoft là rất lớn, trong một núi thông tin đó, rất khó để khai thác. Vì thế có rất nhiều Data mart về kinh doanh, tiếp thị, kỹ thuật, testing,... Có một điểm lưu ý ở đây là có một công cụ hay đúng hơn là một chuẩn công cụ mà mọi hệ quản trị Database hỗ trợ cho việc truy vấn thông tin trong Datamart rồi đưa ra những quyết định, nhận định những thông tin trong Datamart - Đó là OLAP, bộ phân tích trực tuyến (Online Analyze Proceesing).
- Một điểm quan trọng là Database thường được chuẩn hóa (dạng chuẩn 1, 2, 3, BCK) để khai thác. Datawarehouse phải phi chuẩn hoá rồi sau đó có thể chuẩn hoá theo Data mart, điều này đồng nghĩa với việc Datawarehouse sẽ trùng lặp thông tin. Thật ra điều là hiển nhiên, vì việc chuẩn hoá nhằm tránh sự trùng lặp thông tin, do đó sẽ nhất quán trong việc cập nhật, thêm, xoá, sửa. Tuy nhiên Datawarehouse là Database rất lớn phục vụ cho báo cáo, truy vấn chỉ đọc nên

- việc trùng lặp thông tin sẽ giúp thao tác tìm kiếm sẽ nhanh hơn. Đây cũng là một quy luật: Càng Trùng lặp thông tin thì tìm kiếm càng dễ dàng và ngược lại.
- Datawarehouse là một hướng công nghệ nóng nhất. Một kho dữ liệu được xác định đúng hướng, hoạt động hiệu quả có thể trở thành một công cụ cạnh tranh có giá trị cao trong kinh doanh.

II.2.2 Lợi ích của kho dữ liệu:

- Tạo ra những quyết định có ảnh hưởng lớn. Một kho dữ liệu cho phép trích rút tài nguyên nhân lực và máy tính theo yêu cầu để cung cấp các câu truy vấn và các báo cáo dựa vào cơ sở dữ liệu hoạt động và sản xuất. Điều này tạo ra sự tiết kiệm đáng kể. Có kho dữ liệu cũng trích rút tài nguyên khan hiếm của hệ thống sản xuất khi thực thi một chương trình quá lâu hoặc các báo cáo và các câu truy vấn phức hợp.
- Công việc kinh doanh trở nên thông minh hơn. Tăng thêm chất lượng và tính linh hoạt của việc phân tích kinh doanh do phát sinh từ cấu trúc dữ liệu đa tầng của kho dữ liệu, đó là nơi cung cấp dữ liệu được sắp xếp từ mức độ chi tiết của công việc kinh doanh cho đến mức độ cao hơn - mức độ tổng quát. Đảm bảo được dữ liệu chính xác và đáng tin cậy do đảm bảo được là trong kho dữ liệu chỉ chứa duy nhất dữ liệu có chất lượng cao và ổn định (trusted data).
- Dịch vụ khách hàng được nâng cao. Một doanh nghiệp có thể giữ gìn mối quan hệ với khách hàng tốt hơn do có mối tương quan với dữ liệu của tất cả khách hàng qua một kho dữ liệu riêng.
- Tái sáng tạo những tiến trình kinh doanh. Sự cho phép phân tích không ngừng thông tin kinh doanh thường cung cấp sự hiểu biết mọi mặt của phương thức kinh doanh do đó có thể làm nảy sinh ra những ý kiến cho sự sáng tạo ra những tiến trình này lại. Chỉ khi xác định chính xác các nhu cầu từ kho dữ liệu thì mới

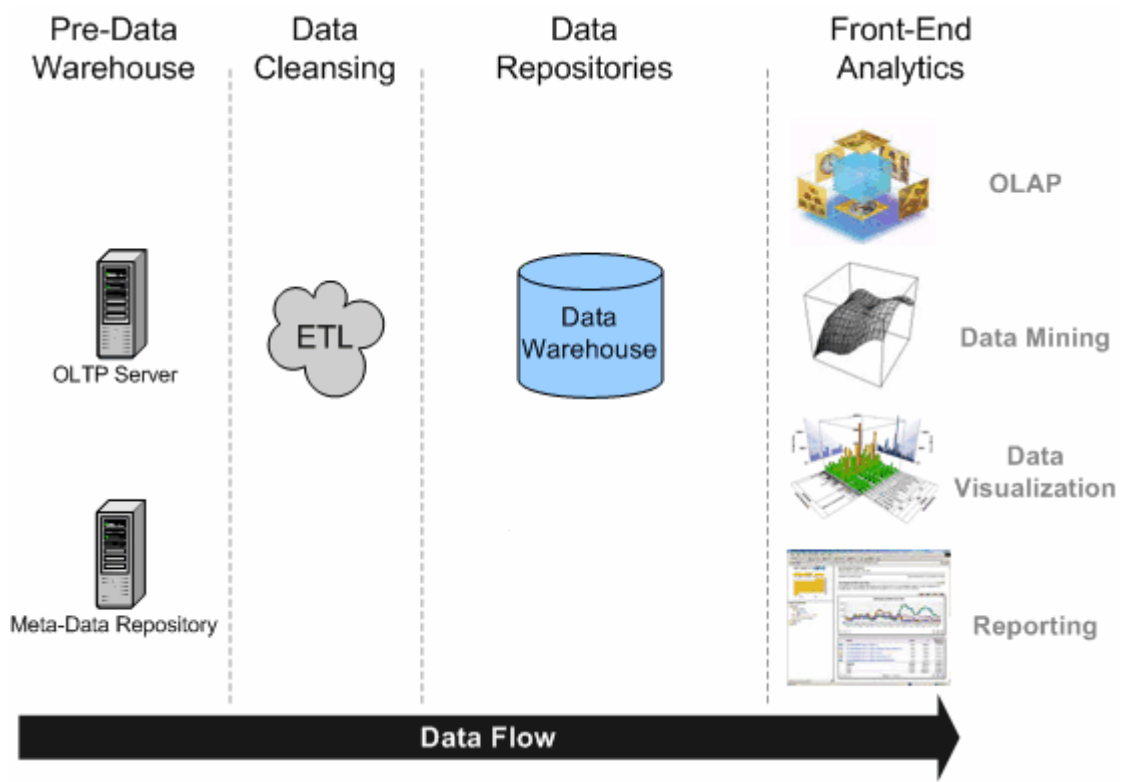
giúp ta đánh giá được những hạn chế và mục tiêu kinh doanh một cách chính xác hơn.

- Tái sáng tạo hệ thống thông tin. Một kho dữ liệu là nền tảng cho các yêu cầu dữ liệu trong mọi lĩnh vực kinh doanh, nó cung cấp một chi phí ảnh hưởng nghĩa là đưa ra thói quen cho cả hai sự chuẩn hóa dữ liệu và sự chuẩn hóa hoạt động của hệ điều hành theo chuẩn quốc tế.

II.2.3 Thành phần Datawarehouse:

Kiến trúc DataWarehouse có thể được chia làm 4 phần chính:

- **Pre- Data Warehouse:** Đây là các nguồn dữ liệu bên ngoài, sẽ được sàng lọc và đưa vào lưu trữ bên trong Data Warehouse.
- **Data Cleansing - ETL:** là viết tắt của cụm từ Extraction (phân rã), Transformation (chuyển đổi), Loading (dung nạp). Đây là thành phần chịu trách nhiệm sàng lọc, chuyển đổi thành các định dạng phù hợp với kiến trúc bên trong của Data Warehouse.
- **Data Repositories:** bao gồm các Metadata. Metadata là dữ liệu để mô tả dữ liệu. Khi dữ liệu được cung cấp cho người dùng cuối, thông tin metadata sẽ cung cấp những thông tin cho phép họ hiểu rõ hơn bản chất về dữ liệu mà họ đang có. Những thông tin này sẽ giúp cho người dùng có được những quyết định sử dụng đúng đắn và phù hợp về dữ liệu mà họ có. Bản thân Metadata cũng có thể quan niệm là nơi để chứa chính dữ liệu, nhưng số lượng phải là cực lớn.
- **Front-End Analytics:** là các ứng dụng sử dụng dữ liệu có trong Data Warehouse như OLAP (On-Line Analysis Processing - Xử lý phân tích trực tuyến), Data Mining (Khai thác dữ liệu), Data Visualization (mô hình hóa dữ liệu), Reporting (trích xuất báo cáo).



Hình II.3 Mô hình luồng dữ liệu Data Warehouse.

II.3 Crystal Report Built-in Visual Studio 2005:

Crystal Report for VisualStudio .NET là công cụ báo cáo chuẩn dành cho VisualStudio.NET. Bạn có thể tiếp quản những báo cáo này trên nền web và windows và phân bổ chúng ở dạng dịch vụ web hoặc win-form trên server.



Hình II.4 Một mẫu báo cáo tạo bằng Crystal Report.

Crystal Report được mệnh danh là chuyên viên phân tích tình hình hoạt động của doanh nghiệp, mang tính định lượng cho tất cả các lĩnh vực từ Tài Chính, Kinh doanh, Marketing cho đến Tổ Chức hoạch định. Hầu hết các doanh nghiệp luôn phải đầu đầu với những con số từ các báo cáo của nhân viên trong công ty. Crystal Report với đầy đủ tính năng và điều kiện để làm các báo cáo, kèm theo những phân tích về nguồn vốn, tình hình chi phí, thu nhập, kết quả hoạt động sản xuất kinh doanh, diễn

giải tình hình tăng giảm tài sản cố định trong thư mục tài chính; quản lý chặt chẽ tình hình doanh số, sản phẩm, thu chi và phần trăm kế hoạch kinh doanh của nhân viên kinh doanh, đánh giá tiềm năng thị trường dựa trên các số liệu của các nhà phân phối, những khách hàng chủ lực để giúp doanh nghiệp có định lượng chính xác cho kế hoạch đầu tư kinh doanh dài hạn tại từng thị trường; không những vậy Crystal Report cũng là 1 chuyên viên Marketing cừ khôi cho kế hoạch triển khai sản phẩm, cân đối chi phí cho hoạt động Marketing của doanh nghiệp theo tiêu chuẩn của quốc tế.v.v...

Crystal Report là công cụ thiết kế cho phép chúng ta tạo ra những bản báo cáo mà kết quả trả về được rút trích từ một Database hoặc Datasource khác nhau. Thay vào đó để cho việc xem data trong một datasource trở nên đơn giản hơn, Crystal Report tạo ra những bản báo cáo một cách đa dạng.

Những dạng báo cáo mà doanh nghiệp thường hay dùng hiện nay bằng Excel, Word, hay Power point, PDF (Acrobat..), sẽ không dùng nữa mà thay vào đó là Crystal Report. Chỉ cần chuyển tải các số liệu vào Crystal Report, doanh nghiệp sẽ có những báo cáo hoàn hảo.

Chương III

GIỚI THIỆU BÀI TOÁN VÀ PHÂN TÍCH ỨNG DỤNG

III.1 Phát biểu bài toán:

Mô tả: Mọi giao dịch của khách hàng tại STSI, sẽ ngay lập tức được chuyển về hai sàn chứng khoán trung tâm là sàn Thành phố Hồ Chí Minh (HOSTC) và sàn Hà Nội (HASTC). Đồng thời, trong khoảng thời gian 10 giây một lần, hệ thống STSI sẽ liên tục nhận thông tin trả về từ hai sàn trung tâm.

Hệ thống lưu trữ dữ liệu tại STSI được chia làm hai phần, Core và Web.

Core là hệ thống lưu trữ các thông tin về tài khoản khách hàng, các thông tin về hợp đồng, các lệnh đặt mua, bán, cầm cố, vay chứng khoán, thông tin lịch sử giao dịch, số dư tài khoản tiền mặt, số dư tài khoản chứng khoán. Core được lưu trữ bằng hệ quản trị CSDL Oracle 10G-R2.

Web là hệ thống website bao gồm mô tả về công ty, tin tức thị trường, các biểu đồ phân tích cổ phiếu, biến động giá, các sự kiện chứng khoán và giao dịch trực tuyến. Website có hai database, STSWEB và WebHavest, sử dụng Hệ quản trị SQL Server 2005. Tin tức thị trường được lấy thông qua chương trình lấy tin tự động, dữ liệu sẽ được chuyển đổi thành dạng xml, lưu trữ tạm trong database WebHavest. Phòng kinh doanh sẽ chịu trách nhiệm kiểm duyệt tin tức và lưu trữ lại vào trong STSWEB. Các biểu đồ được xây dựng bằng webpart, lấy thông tin lưu trữ trong database thông qua các store procedure. Khách hàng cần đăng nhập mới có thể sử dụng chức năng Giao dịch trực tuyến. Mọi giao dịch tại đây sẽ được chuyển qua Core để giao dịch viên xử lý. Website còn có các bảng điện tử thể hiện biến động giá theo thời gian thực của thị trường. Các biến động giá trên thị trường, sẽ được chuyển về, lưu trữ trong database STSWEB. Bảng điện tử sẽ lấy thông tin từ STSWEB, sau đó hiển thị lên cho nhà đầu tư xem.

Qua việc phân tích hệ thống ở các công ty bạn, cũng như góp ý từ phía khách hàng, STSI dự định sẽ xây dựng thêm một số tính năng sau:

- Lúc 17h, mọi thông tin trong STSWEB sẽ được backup lại, chuyển qua lưu trữ trên File Server.
- Thông tin biến động giá từ hai sàn HASTC và HOSTC sẽ được lưu trữ trong table stock_CurrentStockInfo_Full, phục vụ nhu cầu xem lại các biến động giá với tần số chuyển đổi 5 phút/5s.
- Các flat file từ phantichcophieu.vn sẽ được import vào STSWEB tự động vào lúc 14h30.
- Tự động trích xuất các báo cáo hằng ngày, người dùng có thể tùy biến các báo cáo thông qua các điều kiện nhập vào.
- Các quá trình backup, restore sẽ ngắt các kết nối đến hệ thống, các thay đổi trên database từ các kết nối này sẽ được tự động rollback lại.
- Kiểm tra thường xuyên các kết nối đang trong trạng thái Idle. Nếu quá thời gian quy định sẽ drop.

Yêu cầu hệ thống: Ứng dụng cần phát triển và hiện thực trên hệ thống sau:

- Hệ điều hành Win 2003 server
- Microsoft SQL Server 2005
- Crystal Report Built in MS Visual Studio 2005
- Mạng LAN
- Hệ thống lưu trữ SAN.

III.2 Phân tích Use case:

III.2.1 Chức năng Backup Database:

Nhận xét:

- Việc backup database phải được thực hiện thường xuyên, khi kết thúc giờ làm việc vào lúc 17h30. Việc backup này cần được thực hiện tự động, đồng thời phải có thể thực hiện vào bất kỳ lúc nào, phục vụ việc chạy thử, kiểm tra của đội ngũ phát triển database.
- Phải cho phép người sử dụng tùy chọn giữa các loại backup khác nhau:
 - Backup Full: Dùng để backup toàn bộ dữ liệu database
 - Backup Differential: Dùng để backup những dữ liệu thay đổi từ lần Backup Full gần đó nhất
- Tùy vào chính sách của cty mà có thể tùy chọn loại backup phù hợp.
- Chỉ nhân viên của đội ngũ phát triển database mới có thể thực hiện tác vụ backup.
- Toàn bộ các backup file phải đặt tên theo một quy luật thống nhất, gọi nhớ thời điểm backup, nội dung backup và kiểu backup.
- Trước khi thực hiện backup, ngắt toàn bộ các kết nối đang có đến database. Mọi tác vụ đang thực hiện sẽ được rollback lại.
- Các backup file sau khi được tạo ra phải nhanh chóng chuyển qua lưu trữ trên file server.
- Có thể backup chỉ database log.

- Khi backup, đóng gói data file và log file thành một file, giúp tiện lợi cho việc di chuyển, lưu trữ và đồng bộ khi restore.
- Cấu hình tác vụ backup phải rõ ràng, dễ hiểu, dễ làm với người không chuyên database.
- Việc cấu hình tác vụ backup không cần phải đăng nhập SQL Server, không sửa chữa cấu hình lại database.

Hướng phát triển:

- Sử dụng một store procedure đặt trên database master, với đối số truyền vào là tên database cần backup và loại backup.
- Sử dụng ngày giờ lúc thực hiện, làm thành phần của file backup, giúp tránh trùng tên file và gợi nhớ thời điểm backup.
- Sử dụng chức năng Scheduled Task để lập thời biểu backup và theo dõi quá trình hoạt động backup.
- Dùng công cụ SQLCMD để thiết lập các cấu hình backup và gọi store procedure thực hiện.

III.2.2 Restore database:

Nhận xét:

- Chức năng Restore phải luôn đi kèm với chức năng backup database.
- Chức năng Restore cần tự phát hiện file backup thuộc dạng nào, Full, Differential hay là file log.

- Chỉ nhân viên của đội ngũ phát triển database mới có thể thực hiện tác vụ Restore.
- Trước khi thực hiện Restore, ngắt toàn bộ các kết nối đang có đến database. Mọi tác vụ đang thực hiện sẽ được rollback lại.
- Cần cho phép tùy chọn nơi đặt file data và log file.
- Cấu hình tác vụ backup phải rõ ràng, dễ hiểu, dễ làm với người không chuyên database.
- Việc cấu hình tác vụ backup không cần phải đăng nhập SQL Server, không sửa chữa cấu hình lại database.

Hướng phát triển:

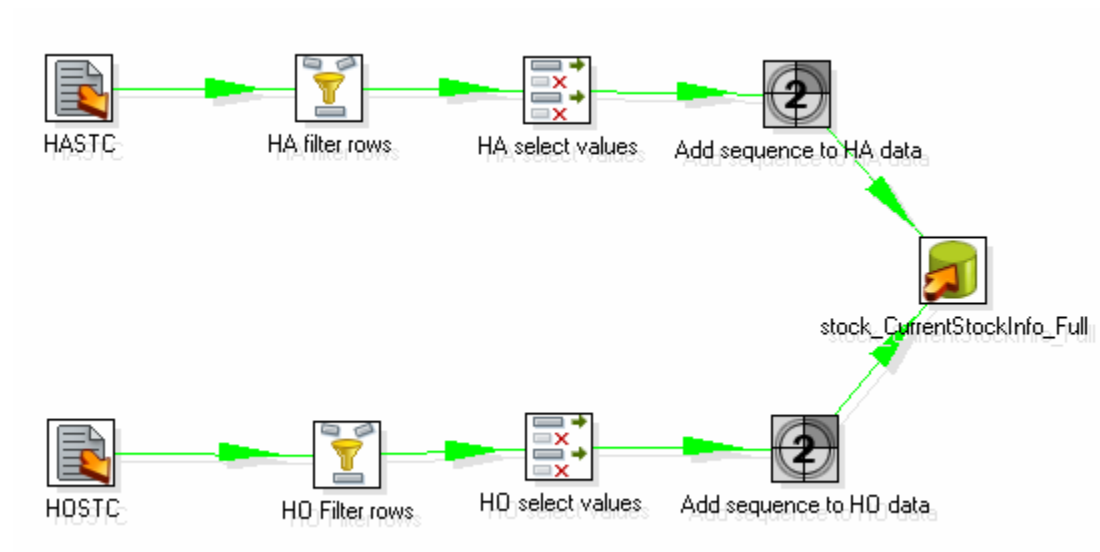
- Sử dụng 1 file SQL script để lưu trữ cấu hình và các thông số cần thiết để thực hiện restore như: tên backup file, restore vào database, data file và log file nào.
- Dùng công cụ SQLCMD để thiết lập các cấu hình restore và gọi script thực hiện.

III.2.3 Dữ liệu từ hai sàn HASTC, HOSTC đổ về Stock_CurrentStockInfo_Full:

Nhận xét:

- Cấu trúc của stock_CurrentStockInfo_Full phải có cấu trúc tương tự như table stock_CurrentStockInfo
- Trong stock_CurrentStockInfo_Full cần thêm field xác định thời điểm nhận dữ liệu và thứ tự của dữ liệu đó trong ngày.
- Cần 1 timer điều khiển việc lấy dữ liệu 5 phút/lần.
- Dữ liệu trả về từ hai sàn chứng khoán HoSTC và HoSTC là dữ liệu liên tục, tần số truyền về là 5 giây. Dữ liệu được truyền dưới dạng flat file.
- Dữ liệu trả về từ sàn được cập nhật liên tục vào table stock_CurrentStockInfo, stock_CurrentMarketInfo và chuyển sang lưu trữ trong stock_CurrentStockInfo_Full.
- Vào đầu phiên giao dịch thứ nhất, cần kiểm tra dữ liệu đổ về từ hai sàn, so khớp số lượng mã chứng khoán có trên hai sàn với lượng mã chứng khoán có trong hệ thống. Nếu có mã chứng khoán mới lên sàn, lập tức phải đưa ra thông báo về mã mới, thuộc sàn nào để Phòng Kinh doanh cập nhật thông tin hệ thống.

Data Flow:



Hình III.1 Mô hình luồng dữ liệu từ hai sàn HA và HO đổ về database

Hướng phát triển:

- Sử dụng store procedure để thực hiện việc đọc, chọn lọc và lấy dữ liệu đổ vào stock_CurrentStockInfo_Full
- Sử dụng các hàm data-time để thực hiện việc chọn lọc dữ liệu.

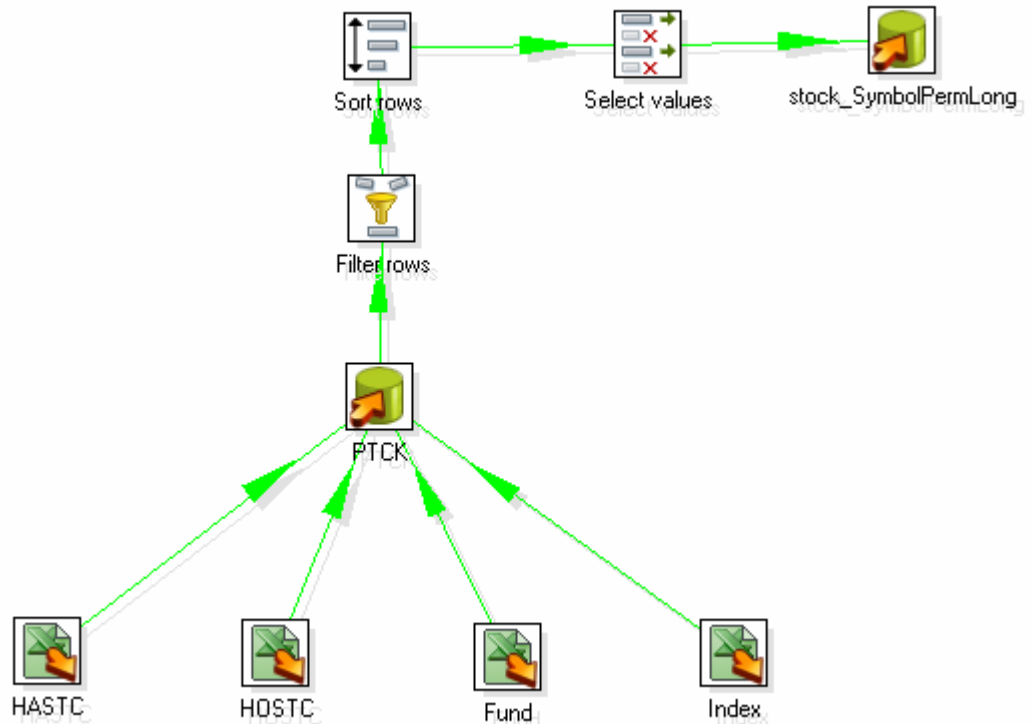
III.2.4 Các flat file từ phantichcophieu.vn sẽ được import vào STSWEB:

Nhận xét:

- Dữ liệu kết quả giao dịch cuối ngày được trả về từ trang web phantichcophieu.vn, gồm 4 file Excel: HaSTC.xls, HoSTC.xls, Index.xls, Fund.xls.
- Dữ liệu kết quả giao dịch cuối ngày được cập nhật vào table stock_SymbolPermLong. Table stock_SymbolPermLong là nơi chứa toàn bộ lịch sử kết quả giao dịch cuối ngày, từ 28/07/2000 đến nay.

- Việc cập nhật dữ liệu cuối phiên giao dịch được giao cho phòng Phân tích đảm trách nên chương trình cập nhật đòi hỏi phải dễ dàng, không qua nhiều bước thực hiện.

Data Flow:



Hình III.2 Mô hình luồng dữ liệu import từ phantichcophieu.vn vào database.

Hướng phát triển:

- Dữ liệu được đẩy vào table PTCK và làm sạch trước khi chuyển vào lưu trữ trên stock_SymbolPermLong.
- Sử dụng SQLCMD để đơn giản hóa công việc cập nhật.
- Cần kiểm tra độ mới của dữ liệu.

III.2.5 Trích xuất báo cáo, người dùng có thể tùy biến các báo cáo thông qua các điều kiện nhập vào:

Nhận xét:

- Các báo cáo phải là báo cáo động, dựa vào điều kiện người dùng đưa ra.
- Báo cáo phải dựa trên mẫu quy định của công ty. Công ty có thể có rất nhiều mẫu báo cáo trong tương lai.
- Dữ liệu để tạo báo cáo được lấy trong database với các table tương ứng.

Hướng phát triển:

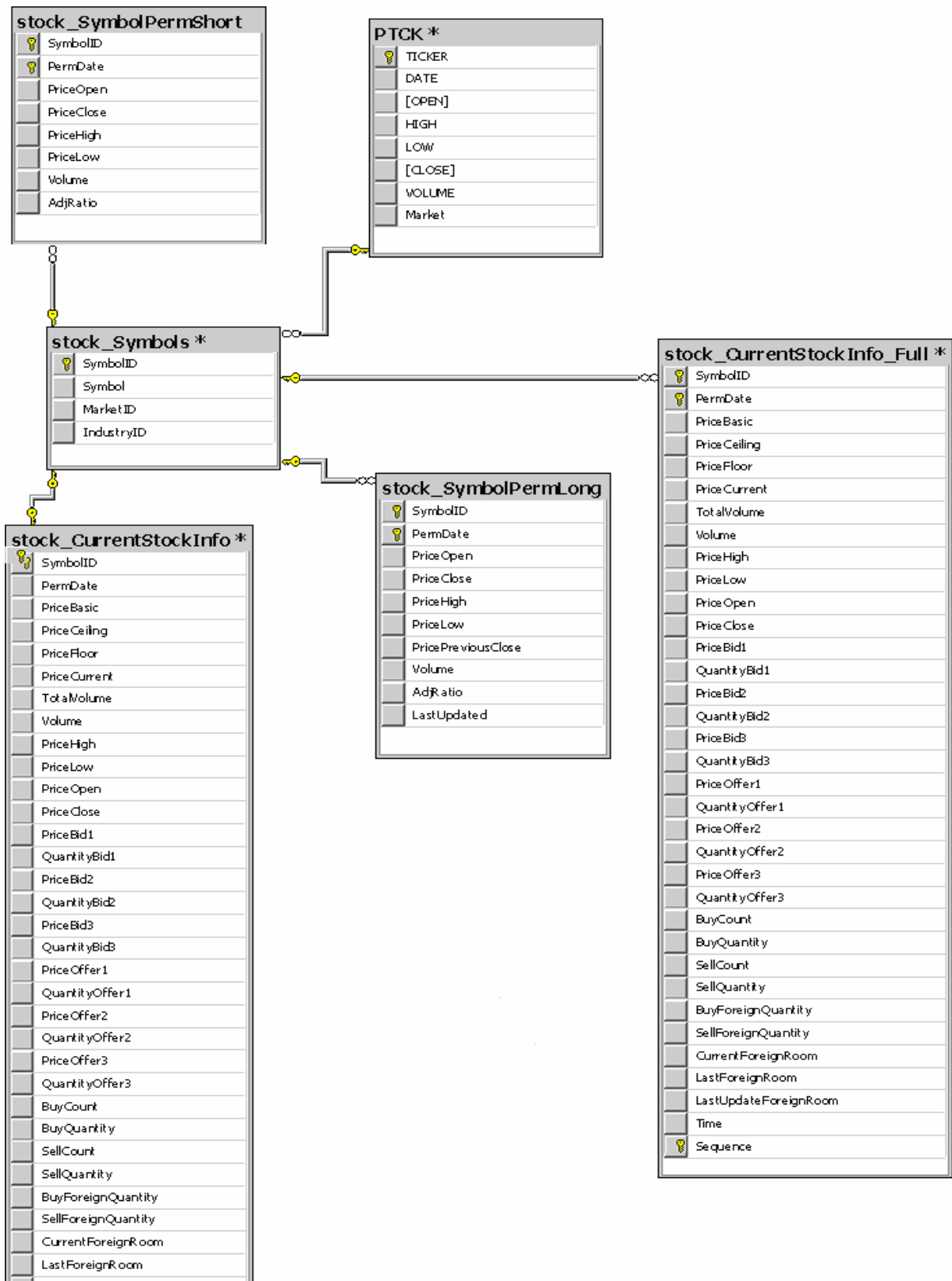
- Mẫu báo cáo được tạo ra trên form của C#. Điều kiện của các mẫu báo cáo được rút trích ra từ store tương ứng được tạo ra trong database.
- Report (mẫu báo cáo) được kết nối với store thông qua dataset.
- Khi muốn xem báo cáo người dùng chỉ việc nhập vào những điều kiện tương ứng với từng mẫu báo cáo được đưa ra.

Chương IV

THIẾT KẾ ỨNG DỤNG

IV.1. Cơ sở dữ liệu STSWEB:

IV.1.1. Lược đồ quan hệ:



Hình IV.1 Lược đồ quan hệ Data Warehouse STSWEB.

IV.1.2. Mô tả database:

Database STSWEB có chức năng lưu trữ toàn bộ thông tin website www.stsi.com.vn và các dữ liệu giao dịch. Trong đó, các table `stock_SymbolPermShort`, `stock_SymbolPermLong`, `stock_CurrentStockInfo`, `stock_CurrentStockInfo_Full`, `stock_Symbols`, PTCK là các table thuộc kiến trúc Data Warehouse.

Tên table	Mô tả
stock_SymbolPermShort	Lưu trữ thông tin kết quả giao dịch cuối ngày của một năm.
stock_SymbolPermLong	Lưu trữ toàn bộ thông tin giao dịch cuối ngày của các năm.
stock_CurrentStockInfo	Lưu trữ biến động giá hiện tại, phục vụ nhu cầu hiển thị giá trên bảng điện tử.
stock_CurrentStockInfo_Full	Lưu trữ lịch sử biến động giá từng ngày của quý.
stock_Symbols	Lưu trữ thông tin các mã chứng khoán trên sàn Thành phố Hồ Chí Minh (HOSTC) và sàn Hà Nội (HASTC).
PTCK	Lưu trữ thông tin của phantichcophieu.vn trước khi làm sạch.

IV.1.3. Ràng buộc toàn vẹn:

Tên ràng buộc	Mô tả
RB-1	Symbol là duy nhất.
RB-2	$\text{PriceCeiling} = \text{PriceBasic} + x\% * \text{PriceBasic}$ <p>x = 5 với Symbol thuộc HOSTC</p> <p>x = 7 với Symbol thuộc HASTC</p>
RB-3	$\text{PriceFloor} = \text{PriceBasic} - x\% * \text{PriceBasic}$ <p>x = 5 với Symbol thuộc HOSTC</p> <p>x = 7 với Symbol thuộc HASTC</p>
RB-4	$\text{PriceClose} - x\% * \text{PriceClose} \leq \text{PriceOpen} \leq \text{PriceClose} + x\% * \text{PriceClose}$ <p>x = 5 với Symbol thuộc HOSTC</p> <p>x = 7 với Symbol thuộc HASTC</p>
RB-5	$\text{PriceFloor} \leq \text{PriceCurrent}, \text{PriceHigh}, \text{PriceLow}, \text{PriceBid}, \text{PriceOffer} \leq \text{PriceCeiling}.$
RB-6	Symbol chỉ có 1 Market duy nhất.

IV.1.4. Quy định:

Tên Quy định	Mô tả
QD-1	SymbolID là số thứ tự mã chứng khoán có trong hệ thống. Khi thêm mới phải kiểm tra có những số thứ tự bị xóa thì bổ sung vào nhưng ký tự còn trống này. Nếu không có thì thêm vào số thứ tự tiếp theo.
QD-2	MarketID quy định sàn mà mã chứng khoán đó thuộc về. 4 là HASTC, 6 là HOSTC.
QD-3	Time lưu lại thời điểm lấy dữ liệu
QD-4	Sequence lưu lại thứ tự của lần lấy dữ liệu đó trong ngày.

IV.2. STSWEB Report:

IV.2.1 Mô tả chức năng:

STSWEB Report là chương trình, được xây dựng để trích xuất các báo cáo dựa vào các dữ liệu có trong DataWarehouse. Để tạo báo cáo, người dùng cần tiến hành các bước sau:

- Mở chương trình, chọn mẫu báo cáo cần tạo.
- Sau khi chọn, sẽ mở hộp thoại cho phép người dùng nhập vào các điều kiện cần thiết để tạo báo cáo.
- Kiểm tra điều kiện nhập vào. Nếu đúng, sẽ tạo báo cáo dựa trên điều kiện nhập vào. Nếu sai, xuất hiện thông báo lỗi.

Báo cáo có thể dành cho nhiều phòng ban. Ví dụ, phòng Kinh Doanh, Phân Tích có thể tham khảo các báo cáo tài chính, phòng Lưu Ký có thể xem các báo cáo về khách hàng, tình trạng hợp đồng... Ngoài ra, còn một số báo cáo đặc biệt, dành cho phòng IT, dùng để theo dõi tình trạng, độ phình, cũng như theo dõi việc chỉnh sửa, thay đổi cấu trúc database.

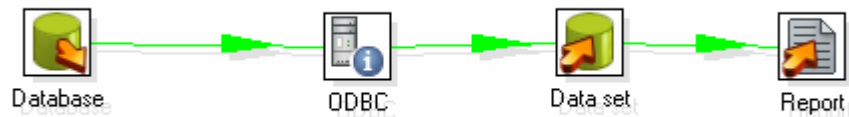
Các mẫu báo cáo mới có thể tạo mới, thêm vào, hoàn thiện trong tương lai.

IV.2.2 Mô tả xử lý:

Quy trình xử lý bao gồm các công việc sau:

- Kiểm tra kết nối đến database.
- Hiện thị form chọn báo cáo.
- Nhận điều kiện nhập vào tương ứng với báo cáo đã chọn.
- Kiểm tra điều kiện.
- Truyền điều kiện vào store procedure.
- Tạo dataset để nhận kết quả trả về từ store procedure.
- Trình bày dữ liệu ra báo cáo.

IV.2.3 Data Flow:



Hình IV.2 Mô hình luồng dữ liệu khi tạo báo cáo.

Chương V

KẾT QUẢ THỰC NGHIỆM

V.1. Chức năng Backup Database:

V.1.1 Cấu trúc chương trình:

Chương trình gồm hai thành phần: store procedure **sp_BackupDatabase** và một file **run.bat**.

a. sp_BackupDatabase:

```
Create PROCEDURE [dbo].[sp_BackupDatabase]
@databaseName sysname, @backupType CHAR(1)
AS
BEGIN
SET NOCOUNT ON;

DECLARE @sqlCommand NVARCHAR(1000)
DECLARE @dateTime NVARCHAR(20)

SELECT @dateTime = REPLACE(CONVERT(VARCHAR, GETDATE(),105), '/', '')
+ '--'+
REPLACE(CONVERT(VARCHAR, GETDATE(),108), ':', '')

IF @backupType = 'F'
SET @sqlCommand = 'BACKUP DATABASE ' + @databaseName +
' TO DISK = 'C:\Backup STSWEB\' + @databaseName + '_Full_' +
@dateTime + '.BAK''

IF @backupType = 'D'
SET @sqlCommand = 'BACKUP DATABASE ' + @databaseName +
' TO DISK = 'C:\Backup STSWEB\' + @databaseName + '_Diff_' +
@dateTime + '.BAK' WITH DIFFERENTIAL'

IF @backupType = 'L'
SET @sqlCommand = 'BACKUP LOG ' + @databaseName +
' TO DISK = 'C:\Backup STSWEB\' + @databaseName + '_Log_' +
@dateTime + '.TRN''

EXECUTE sp_executesql @sqlCommand

END
```

sp_BackupDatabase nhận hai tham số:

- *@databaseName*: tên database cần backup.
- *@backupType*: Kiểu backup.

@backupType nhận các tham số sau:

- F: backup toàn bộ database.
- D: backup chỉ những thông tin thay đổi so với lần backup Full gần nhất.
- L: backup chỉ file database_log.

b. run.bat : Đây là chương trình dùng để gọi thực hiện *sp_BackupDatabase* :

- Tạo ổ đĩa mạng từ File Server 192.168.33.4:

```
net use T: \\192.168.33.4\Backup_STSWEB_79 /user:administrator  
baotin@CK
```

- Dùng *sqlcmd* để gọi thực hiện *sp_BackupDatabase* backup STSWEB và WebHavest database:

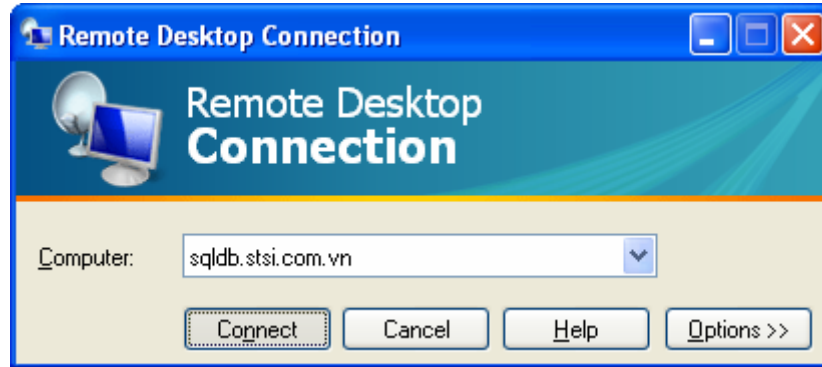
```
SQLCMD.EXE -S 192.168.33.79 -d "master" -Q "exec  
sp_BackupDatabase 'STSWEB', 'F'"  
SQLCMD.EXE -S 192.168.33.79 -d "master" -Q "exec  
sp_BackupDatabase 'WebHavest', 'F'"
```

- Chuyển các file vừa backup thành công qua lưu trữ trên File Server:

```
cd Backup  
echo Chuyen qua File Server.....  
copy *.bak T:\  
del *.bak
```

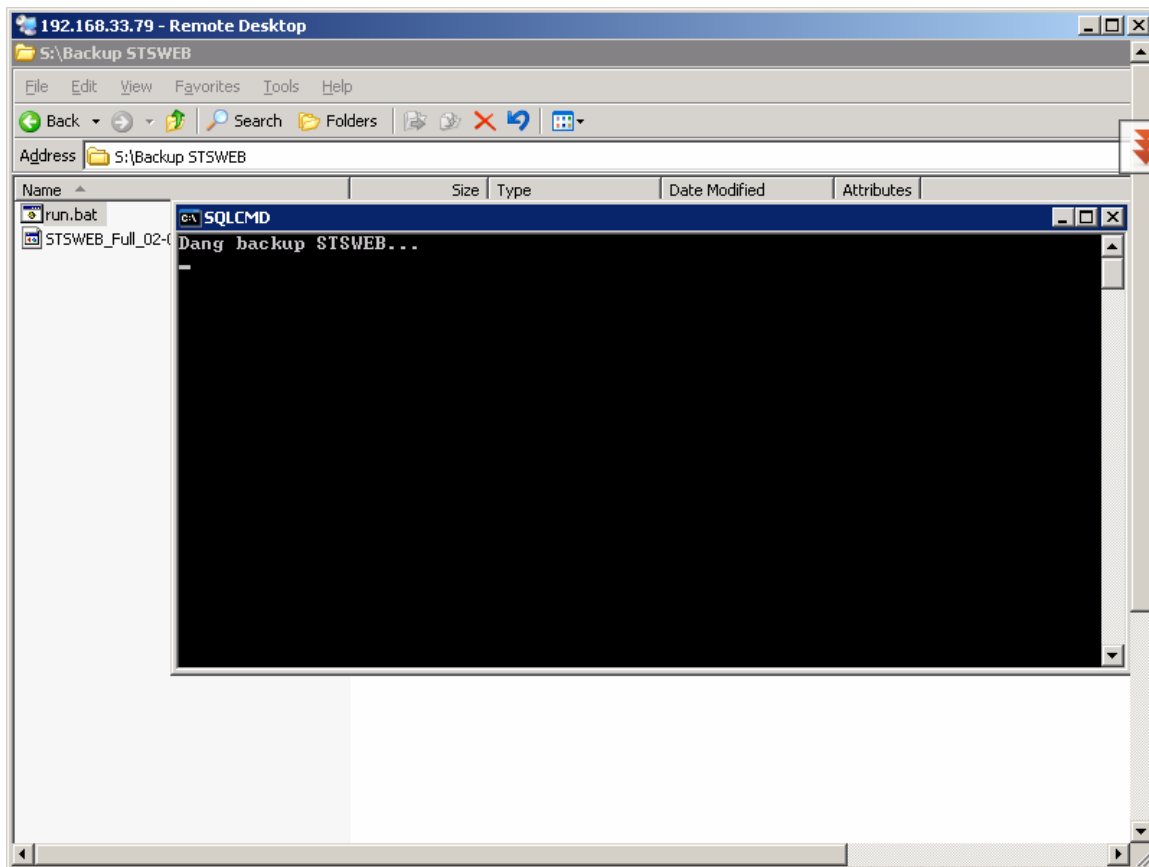
V.1.2 Các bước chạy chương trình:

- Đăng nhập vào database server sqldb.stsi.com.vn:



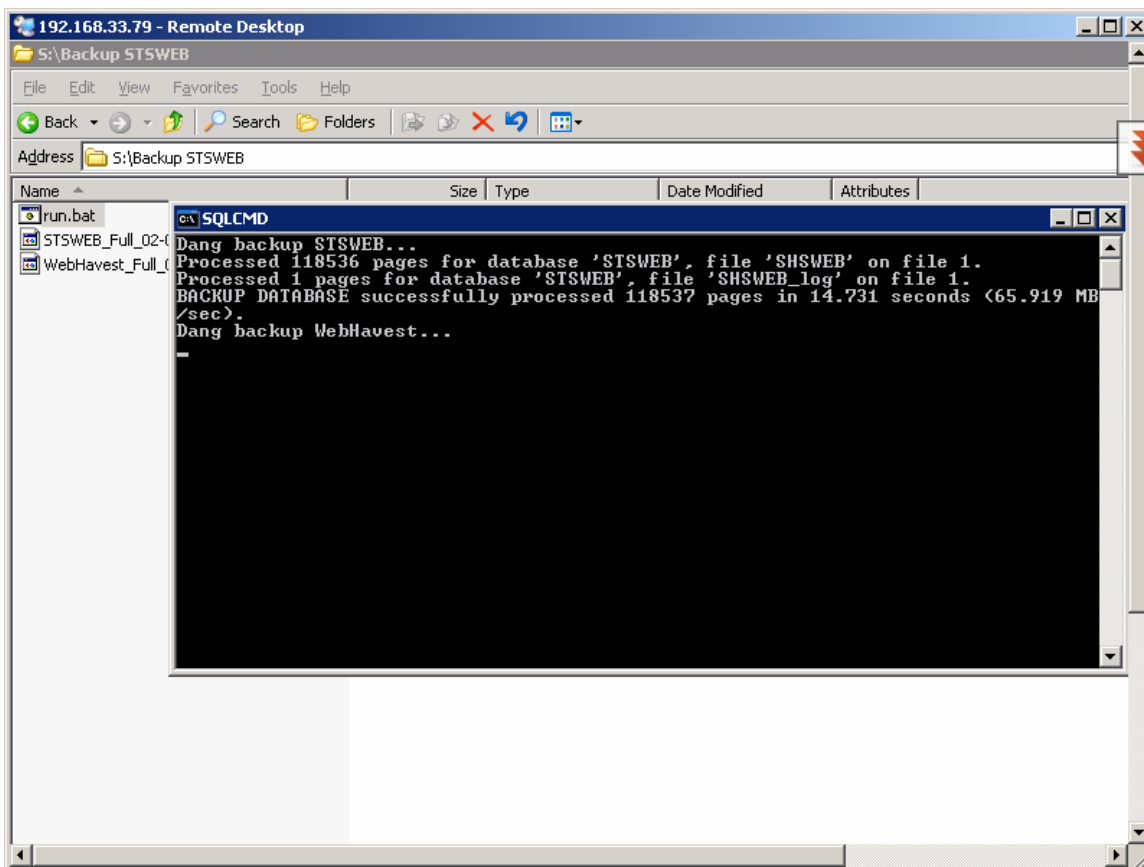
Hình V.1 Kết nối đến Database Server.

- Tìm đến thư mục **D:\Backup**, chạy file **run.bat** :



Hình V.2 Thực hiện backup STSWEB

- Chương trình sẽ backup các database chỉ định:



Hình V.3 Backup thành công STSWEB, chuẩn bị backup WebHavest

- Các file đã backup thành công, sẽ chuyển qua File Server.

V.2. Chức năng Restore Database:

V.2.1. Cấu trúc chương trình:

Chương trình gồm hai module: **restore.bat** và **STSWEB.sql**.

a. STSWEB.sql:

Chứa thông tin cấu hình cần thiết cho tác vụ restore database. Cấu trúc lệnh restore:

```
restore database <database_name>  
  
from disk = '<dia_chi_file_backup>'  
  
<option>
```

Nội dung file STSWEB.sql:

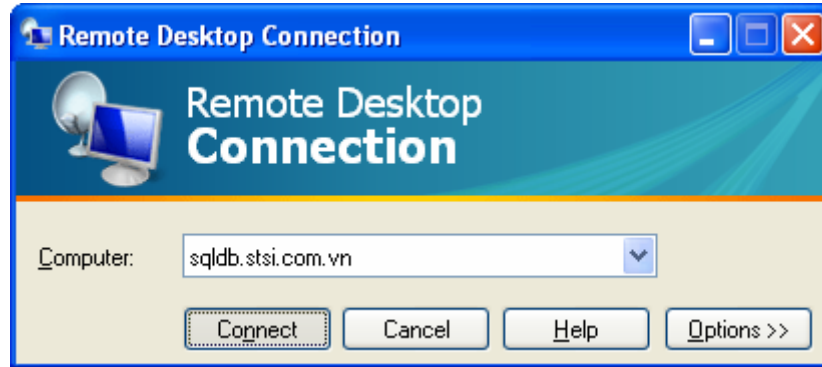
```
print "Restore STSWEB"  
restore database STSWEB  
from disk = 'D:\Quoc Vu\DATN\Restore Database\STSWEB_Full_31-  
12-2008--170000.BAK'  
WITH  
MOVE 'SHSWEB' TO 'C:\Program Files\Microsoft SQL  
Server\MSSQL.1\MSSQL\Data\STSWEB.mdf',  
MOVE 'SHSWEB_log' TO 'C:\Program Files\Microsoft SQL  
Server\MSSQL.1\MSSQL\Data\STSWEB_log.ldf',  
Replace
```

b. Restore.bat: Đây là chương trình để gọi thực hiện script STSWEB.sql:

```
sqlcmd -S 192.168.1.34 -d "master" -i "D:\Quoc Vu\DATN\Restore  
Database\STSWEB.sql"
```

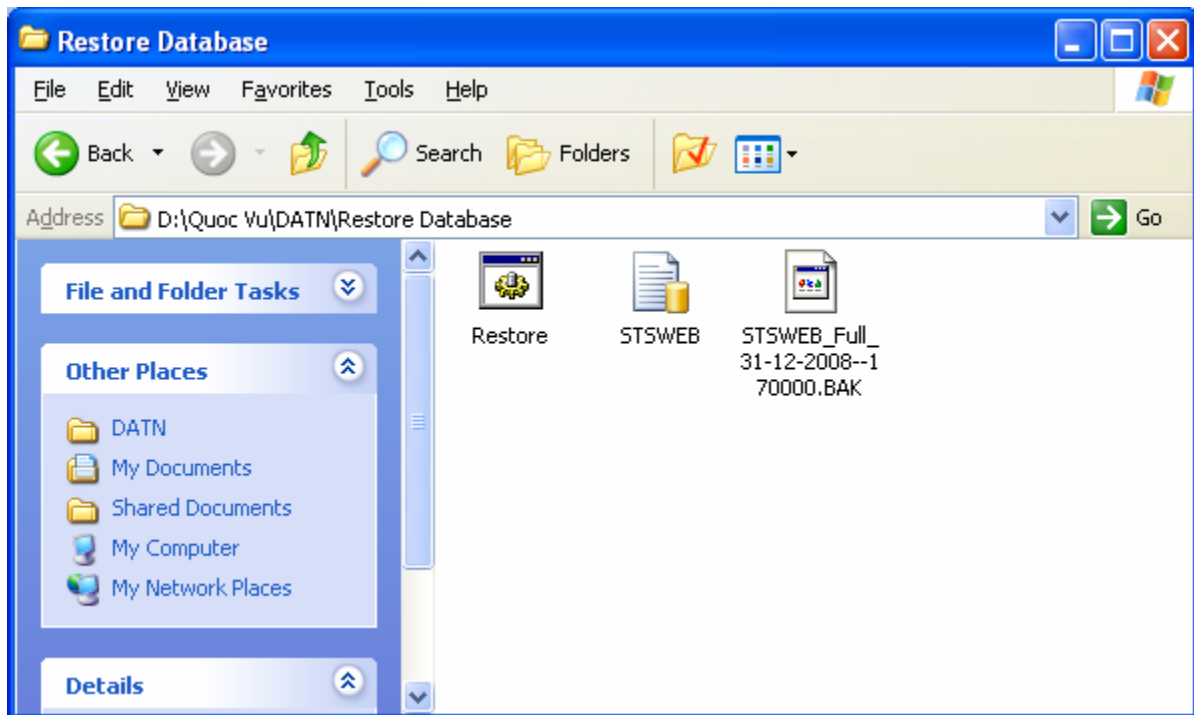
V.2.2 Các bước chạy chương trình:

- Đăng nhập vào database server sqldb.stsi.com.vn:



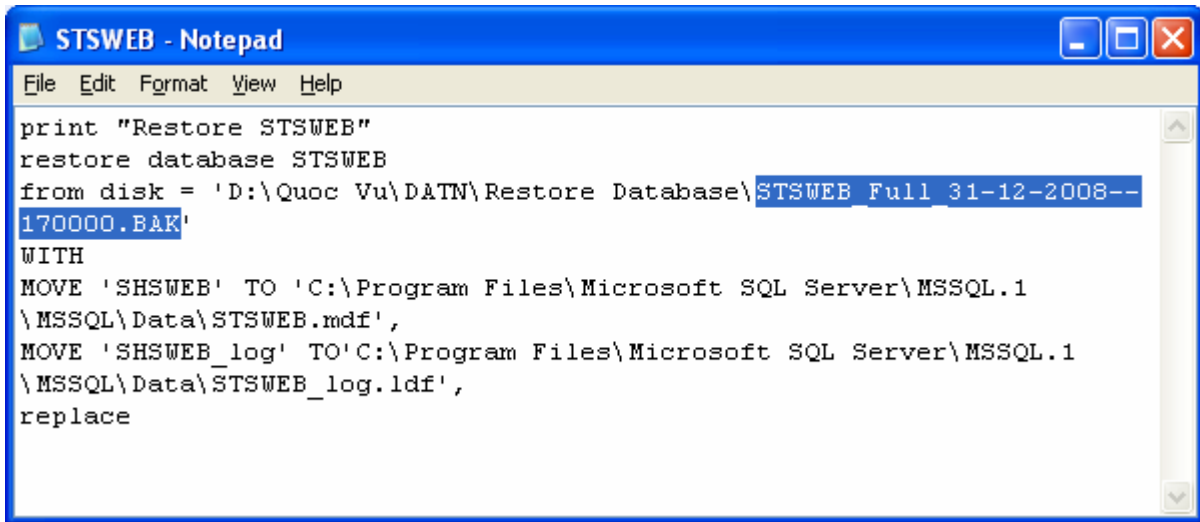
Hình V.4 Kết nối đến Database Server

- Chép file backup vào cùng thư mục chứa **restore.bat** và **STSWEB.sql**:



Hình V.5 Chuẩn bị restore đến thời điểm 17 giờ ngày 31/12/2008

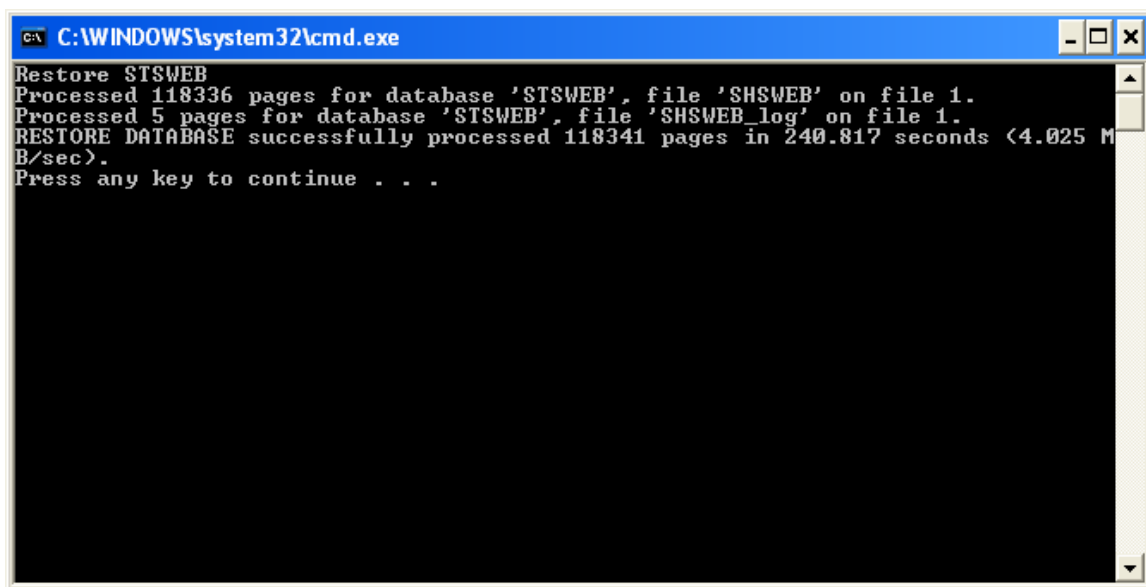
- Chỉnh sửa tên file backup trong STSWEB.sql:



```
STSWEB - Notepad
File Edit Format View Help
print "Restore STSWEB"
restore database STSWEB
from disk = 'D:\Quoc Vu\DATN\Restore Database\STSWEB_Full 31-12-2008--
170000.BAK'
WITH
MOVE 'SHSWEB' TO 'C:\Program Files\Microsoft SQL Server\MSSQL.1
\MSSQL\Data\STSWEB.mdf',
MOVE 'SHSWEB_log' TO 'C:\Program Files\Microsoft SQL Server\MSSQL.1
\MSSQL\Data\STSWEB_log.ldf',
replace
```

Hình V.6 Sửa đổi nội dung file STSWEB.sql

- Chạy file Restore.bat:



```
C:\WINDOWS\system32\cmd.exe
Restore STSWEB
Processed 118336 pages for database 'STSWEB', file 'SHSWEB' on file 1.
Processed 5 pages for database 'STSWEB', file 'SHSWEB_log' on file 1.
RESTORE DATABASE successfully processed 118341 pages in 240.817 seconds (4.025 M
B/sec).
Press any key to continue . . .
```

Hình V.7 Restore thành công

V.3. Mô phỏng dữ liệu đồ từ hai sàn HOSTC và HASTC:

V.3.1. Cấu trúc chương trình: bao gồm một đoạn script lặp vô tận, liên tục đưa dữ liệu vào table stock_CurrentStockInfo_Full.

- Module lấy giá ngẫu nhiên, mô phỏng biến động giá, nằm trong khoảng giá trần và giá sàn của phiên giao dịch gốc.

```
round((PriceCeiling+PriceFloor)/2 +  
(PriceCeiling+PriceFloor)/2*sin(rand()*1000)*0.02,-2 )
```

- Module lấy khối lượng giao dịch ngẫu nhiên, bằng $\pm 10\%$ giá trị phiên giao dịch gốc.

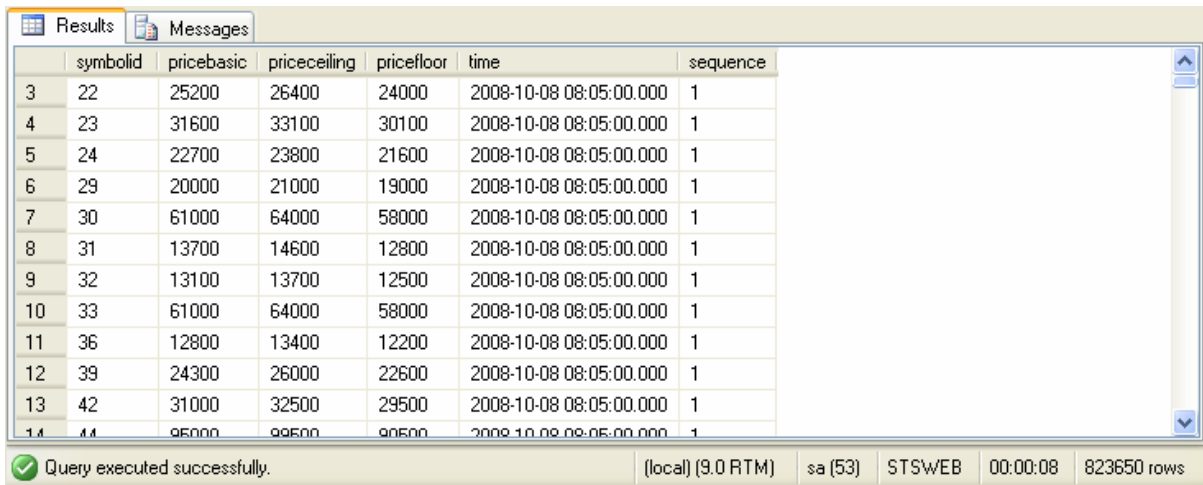
```
round(Volume + Volume*sin(rand()*1000)*0.02,-1 )
```

- Mô phỏng thời gian đưa dữ liệu vào lưu trữ.

```
declare @temp datetime, @d datetime, @t datetime  
declare @i int  
set @d = (select max(permdate) from stock_CurrentStockInfo)  
set @temp = (select datepart(hh,getdate()))  
set @i=0  
while 1=1  
begin  
    begin  
        set @i=@i+1  
        set @t=dateadd(se,8*60+@i*5,@d)  
        if @i=3600/5  
            begin  
                set @i=0  
                set @d = dateadd(dd,1,@d)  
            end  
    end  
    set @temp = (select datepart(hh,getdate()))  
end
```


V.3.2. Các bước chạy chương trình:

Sau khi chạy đoạn script giả lập trên, ta sẽ có dữ liệu trong table `stock_CurrentStockInfo_Full` như sau:

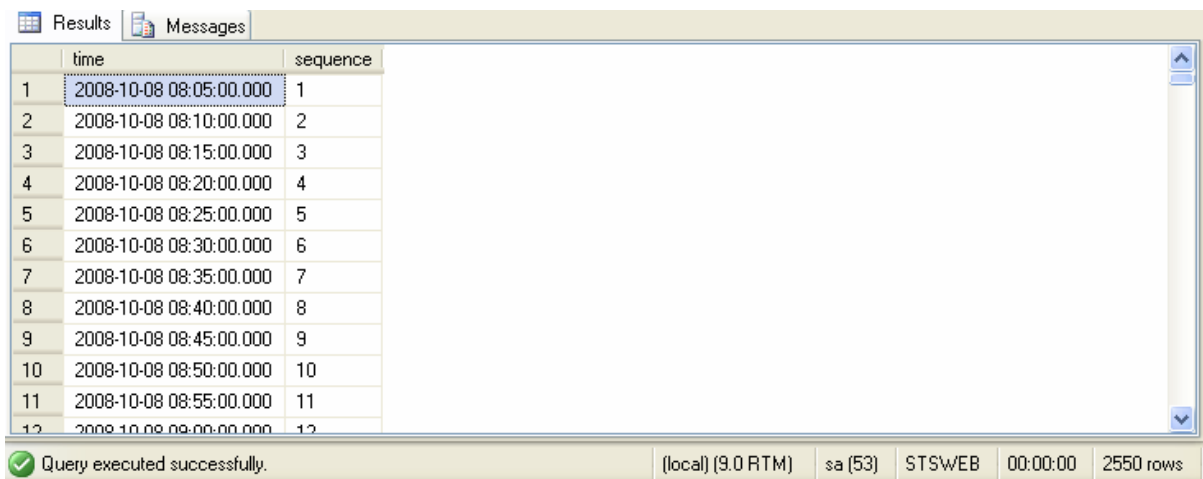


	symbolid	pricebasic	priceceiling	pricefloor	time	sequence
3	22	25200	26400	24000	2008-10-08 08:05:00.000	1
4	23	31600	33100	30100	2008-10-08 08:05:00.000	1
5	24	22700	23800	21600	2008-10-08 08:05:00.000	1
6	29	20000	21000	19000	2008-10-08 08:05:00.000	1
7	30	61000	64000	58000	2008-10-08 08:05:00.000	1
8	31	13700	14600	12800	2008-10-08 08:05:00.000	1
9	32	13100	13700	12500	2008-10-08 08:05:00.000	1
10	33	61000	64000	58000	2008-10-08 08:05:00.000	1
11	36	12800	13400	12200	2008-10-08 08:05:00.000	1
12	39	24300	26000	22600	2008-10-08 08:05:00.000	1
13	42	31000	32500	29500	2008-10-08 08:05:00.000	1
14	44	95000	99500	90500	2008-10-08 08:05:00.000	1

Query executed successfully. (local) (9.0 RTM) sa (53) STSWEB 00:00:08 823650 rows

Hình V. 8 Dữ liệu có trong `stock_CurrentStockInfo_Full`

Thời gian giãn cách đang cài đặt là 5 phút, và số sequence, lưu giữ lần lấy dữ liệu trong ngày.



	time	sequence
1	2008-10-08 08:05:00.000	1
2	2008-10-08 08:10:00.000	2
3	2008-10-08 08:15:00.000	3
4	2008-10-08 08:20:00.000	4
5	2008-10-08 08:25:00.000	5
6	2008-10-08 08:30:00.000	6
7	2008-10-08 08:35:00.000	7
8	2008-10-08 08:40:00.000	8
9	2008-10-08 08:45:00.000	9
10	2008-10-08 08:50:00.000	10
11	2008-10-08 08:55:00.000	11
12	2008-10-08 09:00:00.000	12

Query executed successfully. (local) (9.0 RTM) sa (53) STSWEB 00:00:00 2550 rows

Hình V.9 Thời gian và thứ tự lấy dữ liệu trong ngày.

V.4. Flat file từ phantichcophieu.vn import vào STSWEB:

V.4.1 Cấu trúc chương trình:

Chương trình gồm có 2 module chính: store procedure **Update_CurrentStockInfo** và file **PTCK.bat** .

a. Update_CurrentStockInfo:

- Module lấy thông tin từ 4 file Excel: HoSTC.xls, HaSTC.xls, Index.xls, Fund.xls sử dụng hàm OpenRowset của SQL Server.

```
Print 'Insert cac dong co trong file Index.xls'
Insert into
PTCK(TICKER,DATE,[OPEN],HIGH,LOW,[CLOSE],VOLUME)
select TICKER,DATE,[OPEN],HIGH,LOW,[CLOSE],VOLUME from
OPENROWSET('Microsoft.Jet.OLEDB.4.0','Excel
8.0;Database=S:\PTCK\index.xls;HDR=YES','SELECT * FROM
[Price$]')
Print 'Insert cac dong co trong file Fund.xls'
Insert into
PTCK(TICKER,DATE,[OPEN],HIGH,LOW,[CLOSE],VOLUME)
select TICKER,DATE,[OPEN],HIGH,LOW,[CLOSE],VOLUME from
OPENROWSET('Microsoft.Jet.OLEDB.4.0','Excel
8.0;Database=S:\PTCK\Fund.xls;HDR=YES','SELECT * FROM
[Price$]')
Print 'Insert cac dong co trong file HaSTC.xls'
Insert into
PTCK(TICKER,DATE,[OPEN],HIGH,LOW,[CLOSE],VOLUME,Market)
select
TICKER,DATE,[OPEN],HIGH,LOW,[CLOSE],VOLUME,'HASTC' from
OPENROWSET('Microsoft.Jet.OLEDB.4.0','Excel
8.0;Database=S:\PTCK\HASTC.xls;HDR=YES','SELECT * FROM
[Price$]')
Print 'Insert cac dong co trong file HoSTC.xls'
Insert into
PTCK(TICKER,DATE,[OPEN],HIGH,LOW,[CLOSE],VOLUME,Market)
select
TICKER,DATE,[OPEN],HIGH,LOW,[CLOSE],VOLUME,'HOSTC' from
OPENROWSET('Microsoft.Jet.OLEDB.4.0','Excel
8.0;Database=S:\PTCK\HOSE.xls;HDR=YES','SELECT * FROM
[Price$]')
```

- Module kiểm tra có mã chứng khoán mới lên sàn để phòng Phân tích cập nhật thông tin vào hệ thống.

```
insert into Temp_Stock
select ticker, Market from PTCK where ticker not in (select
symbol from stock_Symbols union select symbol from
temp_Stock)
delete from Temp_Stock
where symbol in (select symbol from stock_Symbols)
declare @count int
set @count = (select count(*) from temp_stock)
print ''
print 'Kiem tra Symbol moi...'
print ''
if @count>0
begin
    print 'Co '+str(@count,1,1)+' Symbol moi!..'
    select Symbol,Market from temp_stock
    print ''
    print 'Huy bo cap nhat, cho insert symbol moi'
    return 0
end
```

b. PTCK.bat:

Gọi thực hiện store **Update_CurrentStockInfo**:

```
SQLCMD.EXE -S 192.168.33.79 -d "STSWEB" -Q "exec
Update_CurrentStockInfo"
```

V.4.2 Các bước chạy chương trình:

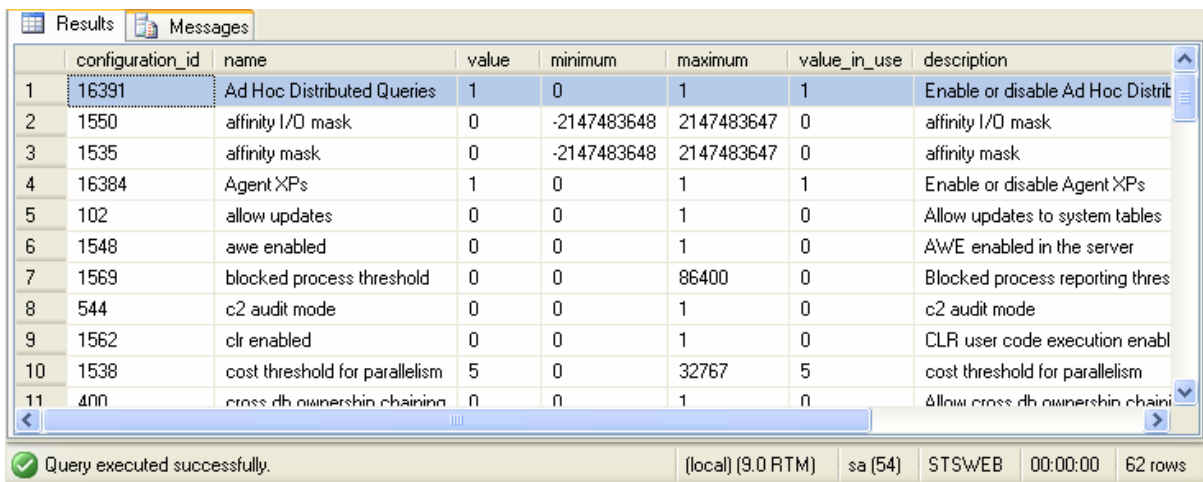
- Lần đầu tiên chạy chương trình, cần cài đặt SQL Server cho phép sử dụng chức năng “Ad Hoc Distributed Queries”. Đây là chức năng cho phép người dùng sử dụng câu lệnh OPENROWSET.

```
sp_configure 'show advanced options', 1  
reconfigure
```

```
sp_configure 'Ad Hoc Distributed Queries', 1  
reconfigure
```

- Có thể kiểm tra SQL Server đã cấu hình chưa bằng cách dùng lệnh:

```
SELECT * FROM sys.configurations  
order by name
```



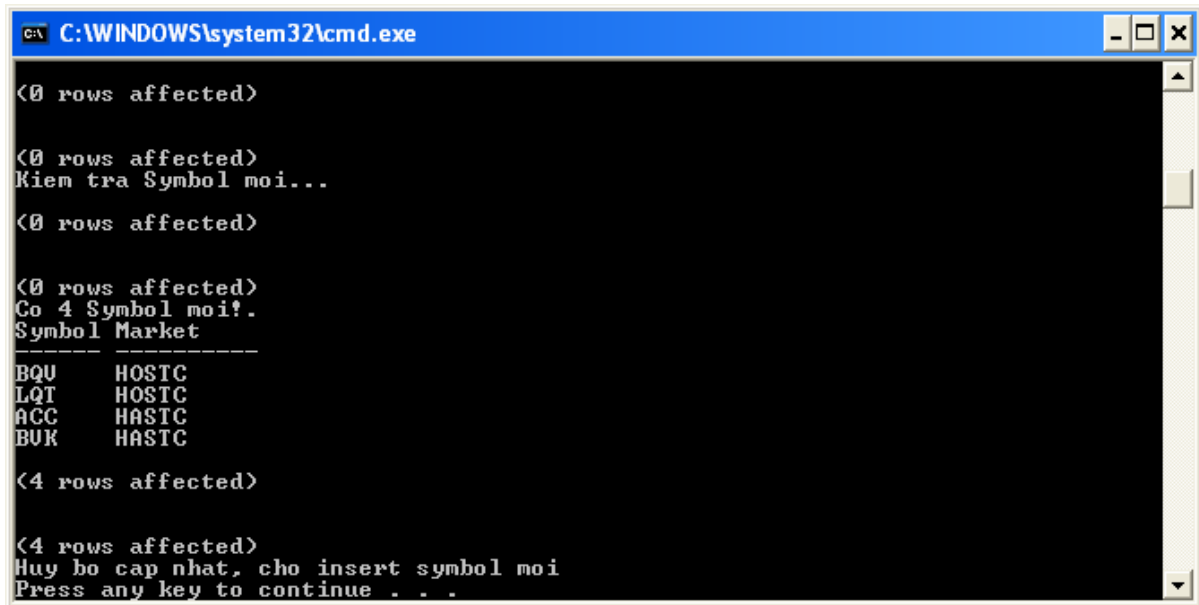
	configuration_id	name	value	minimum	maximum	value_in_use	description
1	16391	Ad Hoc Distributed Queries	1	0	1	1	Enable or disable Ad Hoc Distrib
2	1550	affinity I/O mask	0	-2147483648	2147483647	0	affinity I/O mask
3	1535	affinity mask	0	-2147483648	2147483647	0	affinity mask
4	16384	Agent XPs	1	0	1	1	Enable or disable Agent XPs
5	102	allow updates	0	0	1	0	Allow updates to system tables
6	1548	awe enabled	0	0	1	0	AWE enabled in the server
7	1569	blocked process threshold	0	0	86400	0	Blocked process reporting thres
8	544	c2 audit mode	0	0	1	0	c2 audit mode
9	1562	clr enabled	0	0	1	0	CLR user code execution enabl
10	1538	cost threshold for parallelism	5	0	32767	5	cost threshold for parallelism
11	400	cross db ownership chaining	0	0	1	0	Allow cross db ownership chaini

Query executed successfully. (local) (9.0 RTM) sa (54) STSWEB 00:00:00 62 rows

Hình V.10 Các thiết lập cấu hình SQL Server

- Copy 4 file: HoSTC.xls, HaSTC.xls, Index.xls, Fund.xls vào thư mục **D:\Quoc Vu\DATN\PTCK**.
- Chạy file PTCK.bat

- Nếu phát hiện có mã chứng khoán mới, chương trình đưa thông tin về mã đó và tạm dừng cập nhật để đảm bảo không mất dữ liệu.



```
C:\WINDOWS\system32\cmd.exe

<0 rows affected>

<0 rows affected>
Kiem tra Symbol moi...

<0 rows affected>

<0 rows affected>
Co 4 Symbol moi!.
Symbol Market
-----
BQU    HOSTC
LQT    HOSTC
ACC    HASTC
BUK    HASTC

<4 rows affected>

<4 rows affected>
Huy bo cap nhat, cho insert symbol moi
Press any key to continue . . .
```

Hình V.11 Các mã chứng khoán mới lên sàn.

- Nếu không có mã chứng khoán mới, chương trình sẽ tiếp tục cập nhật vào hệ thống.
- Nếu dữ liệu đã được cập nhật trước đó, sẽ không thể cập nhật lần hai trong cùng một ngày.

```
C:\WINDOWS\system32\cmd.exe
<0 rows affected>
Kiem tra du lieu co san trong CurrentStockInfo.....
<0 rows affected>
<0 rows affected>
<0 rows affected>
Du lieu ngay 31/12 da duoc cap nhat truoc do!. Khong tien hanh cap nhat!
<0 rows affected>
<0 rows affected>
<0 rows affected>
Ket thuc cap nhat, xoa tat ca du lieu tam, dong ung dung...
<0 rows affected>
Press any key to continue . . .
```

Hình V.12 Không cập nhật dữ liệu trùng.

V.5. Report STSWEB:

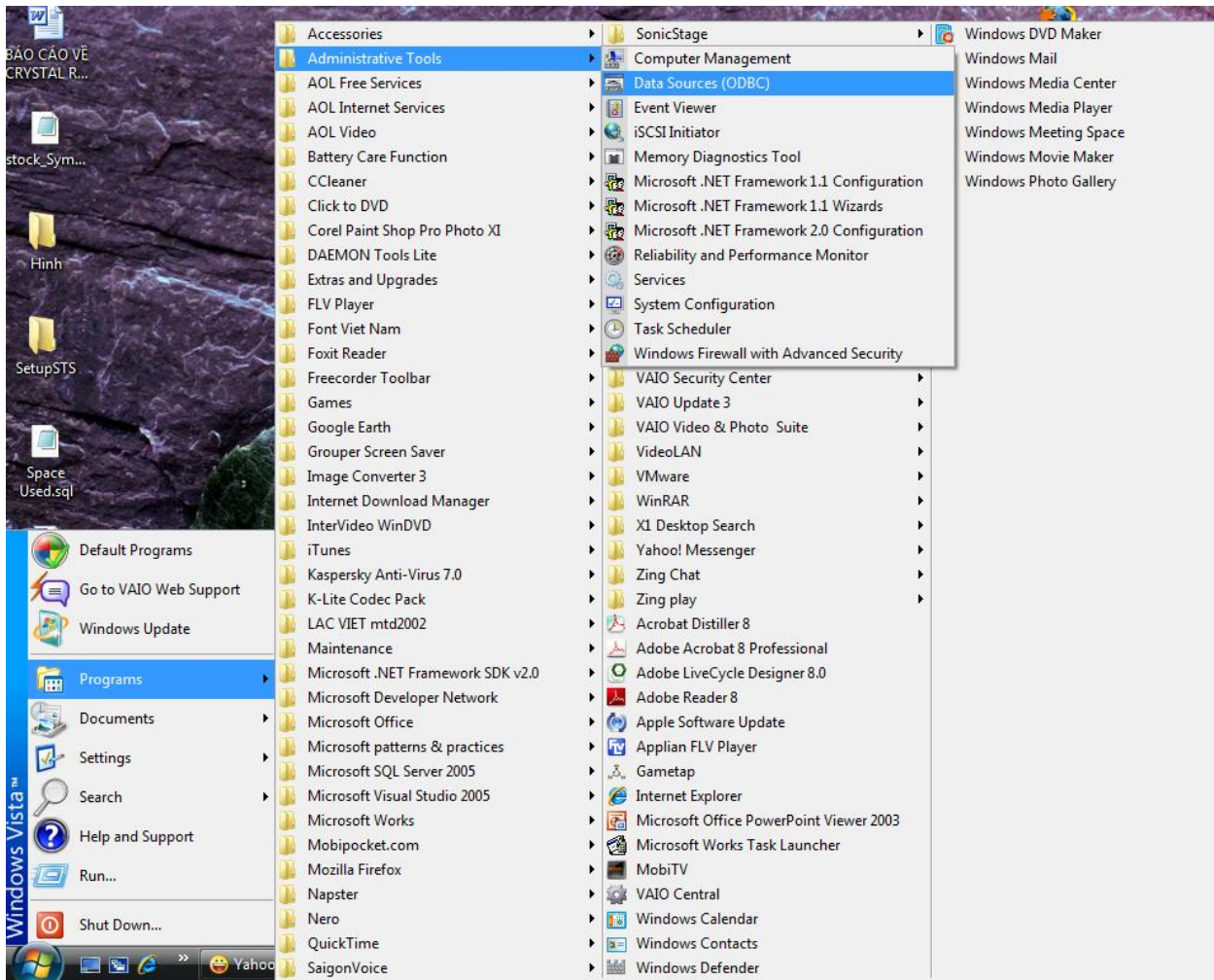
V.5.1. Cấu trúc chương trình:

Chương trình gồm có 3 phần:

- Form hiển thị danh sách các báo cáo.
- Form cho phép nhập điều kiện để tạo báo cáo.
- Form xuất báo cáo

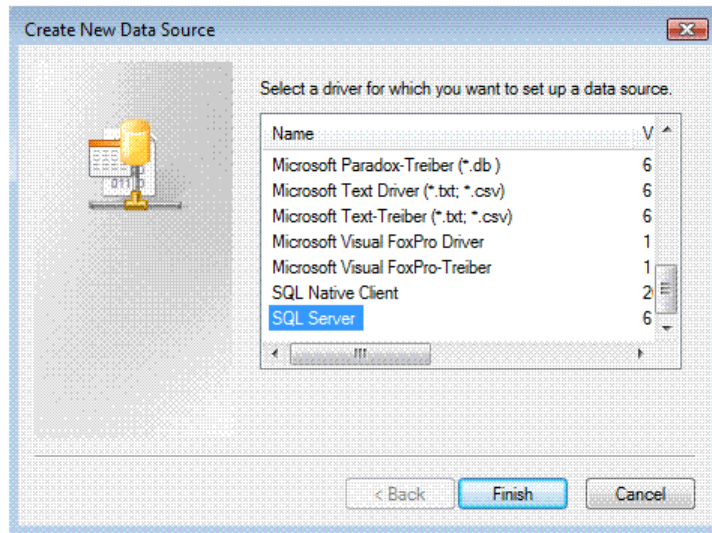
V.5.2 Các bước chạy chương trình:

- Để kết nối được với database, máy tính cần cài đặt ODBC Data Source.



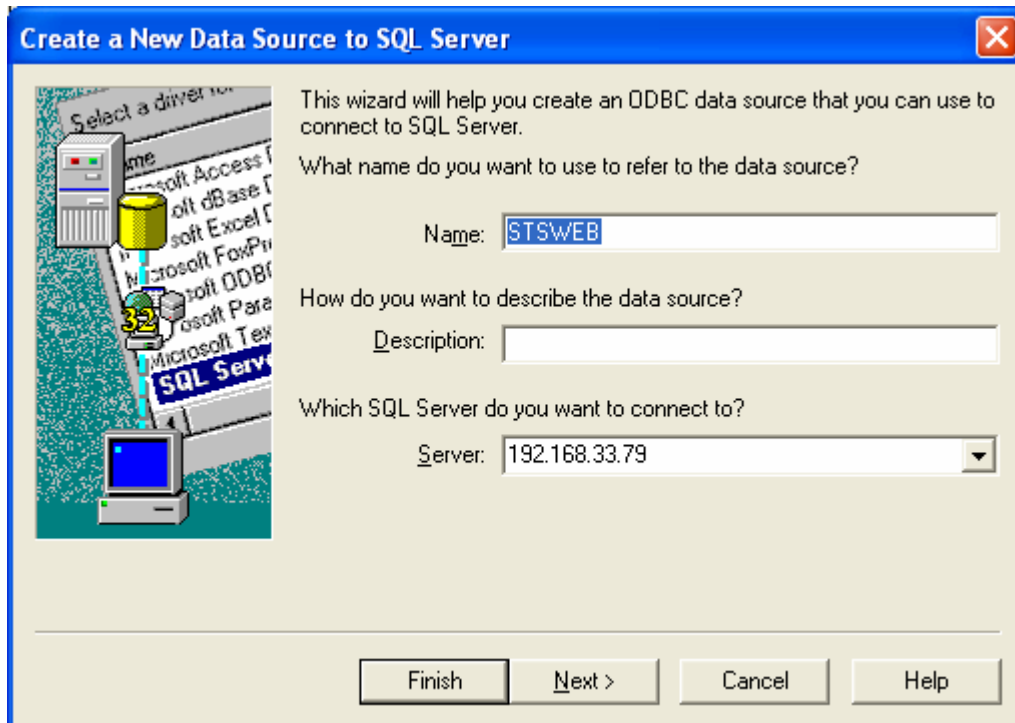
Hình V.13 Chạy ODBC

- Thêm Data Source SQL Server:



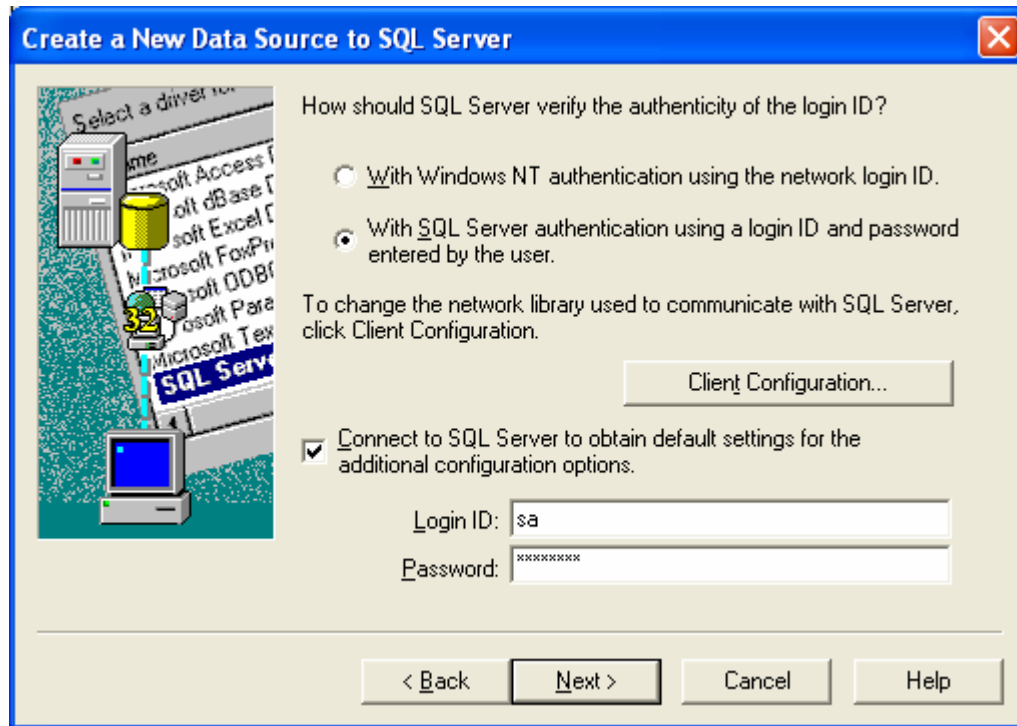
Hình V.14 Tạo data source

- Đặt tên kết nối và khai báo địa chỉ database server.



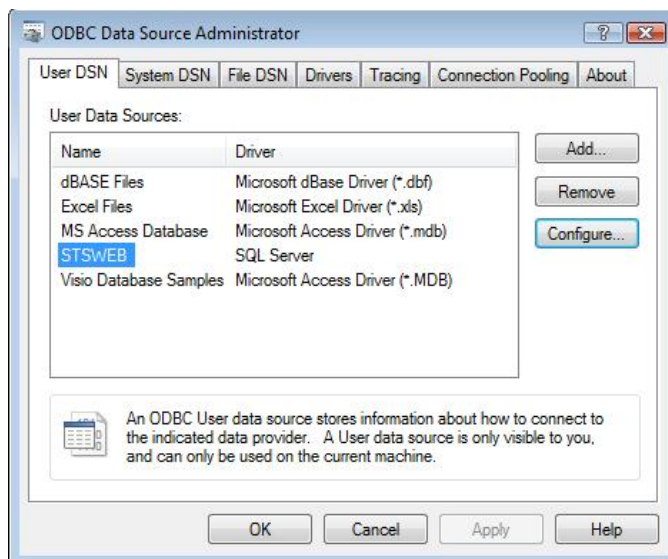
Hình V.15 Tên data source và địa chỉ

- Gỡ mật khẩu SQL Server.



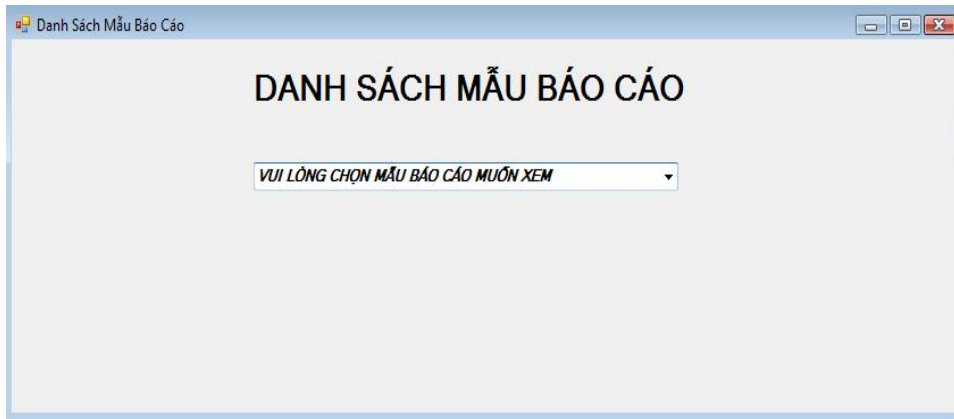
Hình V.16 Cung cấp User và Password để kết nối.

- Kết nối thành công đến SQL Server.



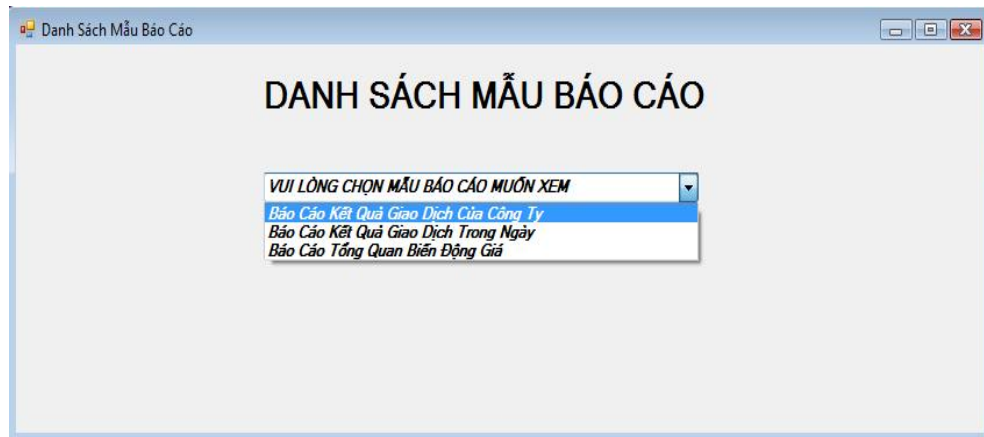
Hình V.17 Tạo Data Source thành công

- Chạy chương trình STSWEB Report.



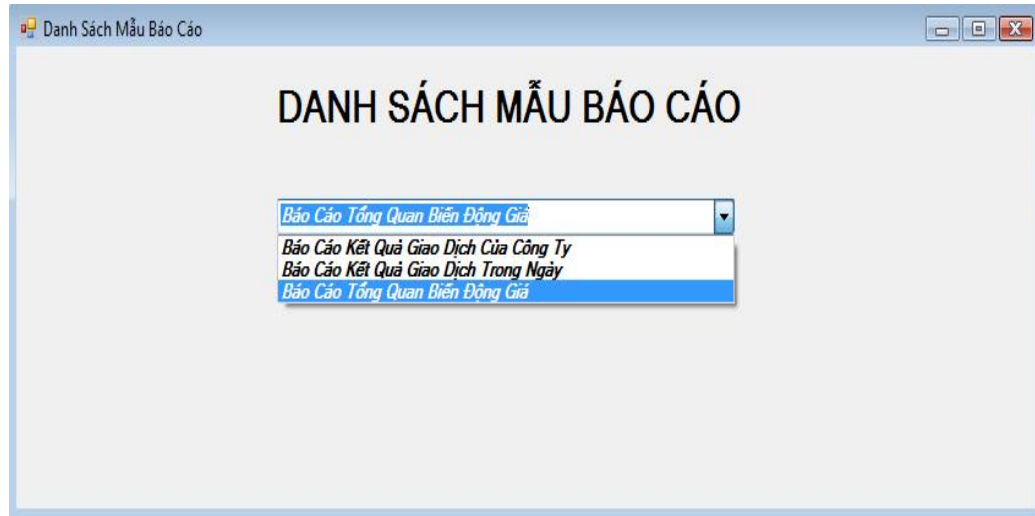
Hình V.18 Danh sách mẫu báo cáo

- Nhấp chuột vào list box, xuất hiện các mẫu báo cáo.

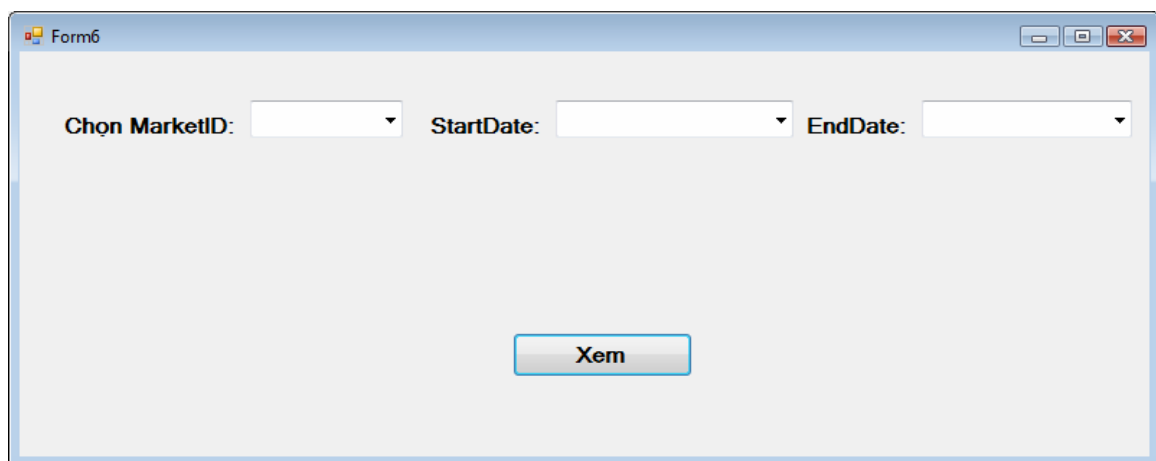


Hình V.19 Chọn mẫu báo cáo

Giả sử chọn Báo cáo Tổng Quan Biến Động Giá.



- Xuất hiện hộp thoại nhập liệu, cho phép người dùng đặt điều kiện để lấy kết quả báo cáo.



Hình V.20 Form cho phép người dùng nhập báo cáo.

- MarketID có thể nhận 3 giá trị. 1 là biến động giá từ cả hai sàn, 4 là biến động giá sàn Hà Nội, 6 là biến động giá sàn Thành phố Hồ Chí Minh.
- Giả sử cần lập báo cáo biến động giá của sàn Hà Nội từ ngày 11/11/2008 đến ngày 11/12/2008.

Form6

Chọn MarketID: 4 StartDate: 11/11/2008 EndDate: 11/12/2008

Xem

Hình V.21 Người dùng nhập điều kiện.

- Xuất báo cáo

Form7

Main Report

Bảo Tín Securities Head Office
 Add: 121 Lê Lợi, Quận 1, TP.HCM
 Phone: {+84-8}9148148 - Fax: {+84-8}9148147 - Website: <http://www.stsi.com.vn>

1/4/2009 **BÁO CÁO TỔNG QUAN BIẾN ĐỘNG GIÁ**
 Từ ngày: 11/11/2008 Đến ngày: 11/12/2008

Symbol	PriceHigh	PriceLow	Volume	PriceCurrent	PriceBasic
ACB	44000	40000	2951800	43500	42000
B82	13500	12200	9600	13100	12400
BCC	13000	11600	375400	12400	12400
BLF	10300	9200	168200	10200	9800
BTS	11300	10500	22000	11300	11200
BVS	32500	28500	499000	28500	30500
CDC	26000	23800	82000	24900	25400
CIC	10300	9700	9000	9800	10000
CMC	11900	11000	56100	11900	11200
CSG	10500	9600	102500	10400	10400
CTN	12200	11300	19200	12200	12000
DAC	30400	27500	13000	28000	29400
DAE	12800	11400	14800	12800	12100
DBC	25300	22500	1096600	24400	23600
DCS	8600	7300	47800	8000	7700
DTC	46000	41400	12700	46000	44400
EBS	14900	13400	85700	14200	14000
GHA	23000	21500	4000	23000	22500
HAI	22000	20000	11300	21400	21400
HEV	14900	13900	2000	13900	14900
HHC	20000	17700	8000	20000	18300
HJS	12200	10700	9000	12200	11300
HLV	38900	35900	15700	36200	38400
HNM	11800	10300	238900	11500	11000
HPC	19500	17400	370800	17500	18700
HTD	40600	32000	44600	40600	36000

Current Page No.: 1 Total Page No.: 1+ Zoom Factor: 100%

Hình V.22 Xuất báo cáo.

Chương VI

TỔNG KẾT

VI.1. Kết luận

VI.1.1 Về mặt lý thuyết:

Chúng em đã tìm hiểu được những vấn đề sau:

- Bản chất, đặc điểm của hệ thống Data Warehouse.
- Một số khái niệm và nghiệp vụ chứng khoán.
- Tiếp xúc và làm việc với một số công nghệ mới như SQL Server 2005, Crystal Report....

VI.1.2 Về mặt ứng dụng:

- Các chương trình Backup database, Restore Database, hệ thống lưu trữ Data Warehouse đã được ứng dụng và chạy tốt trong vòng 4 tháng qua.
- Ứng dụng Report có thể cho phép người dùng tạo báo cáo một cách dễ dàng vào bất kỳ thời điểm nào một cách tổng quan và đầy đủ nhất.
- Số lượng báo cáo có thể được thêm vào trong tương lai.

Ngoài ra, chúng em cũng đã tinh giản độ phức tạp của việc sử dụng ứng dụng, giúp cho nhân viên dễ cấu hình và quản lý hệ thống.

VI.2. Hướng phát triển:

Do thời gian xây dựng để kịp hạn định của đồ án, không thể tránh khỏi những thiếu hụt, chưa hợp lý trong ứng dụng. Trong tương lai, chúng em dự định sẽ hoàn chỉnh những chức năng sau cho hệ thống:

- Hệ thống Data Warehouse cần hỗ trợ thêm nhiều data source, tạo sự uyển chuyển trong việc thu thập dữ liệu.
- Ứng dụng Report có thể import mẫu báo cáo thông qua các template bằng XML.
- Ứng dụng Report cần có cơ chế xác thực quyền để có thể xác định quyền tạo báo cáo.
- Phát triển Report theo nhiều định dạng như: biểu đồ, đồ thị trực quan...

TÀI LIỆU THAM KHẢO

- Professional Crystal Reports for Visual Studio .NET (Second Edition). David McAmis.
- Giải pháp lập trình C#. Tổng hợp và biên dịch (Nguyễn Ngọc Bình Phương – Thái Thanh Phong), cùng sự cộng tác của (Nguyễn Thanh Nhân – Trần Lê Vĩnh Phong – Nguyễn Quang Nam – Đinh Phan Chí Tâm – Bùi Minh Khoa – Lê Ngọc Sơn – Thái Kim Phụng – Lê Trần Nhật Quỳnh). Chịu trách nhiệm xuất bản: TS. Nguyễn Xuân Thủy. Nhà xuất bản Giao Thông Vận Tải.
- Pro C# 2005 and the .NET 2.0 Platform, Third Edition – ANDREW TROELSEN.
- Website: http://blogs.msdn.com/sql_protocols/
- Website: <http://www.codeproject.com>.
- Website: Google.com
- Website: <http://vnoug.org/>
- Website: www.sqlcentral.com