

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC DÂN LẬP HẢI PHÒNG



ISO 9001:2008

PHẠM VĂN HẬU

LUẬN VĂN THẠC SĨ
NGÀNH HỆ THỐNG THÔNG TIN

Hải Phòng - 2016

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC DÂN LẬP HẢI PHÒNG

PHẠM VĂN HẬU

ỨNG DỤNG KHO DỮ LIỆU VÀ WEBSERVICE
ĐỀ TÍCH HỢP DỮ LIỆU
XÂY DỰNG HỆ THỐNG BÁO CÁO THỐNG KÊ
TẠI TRƯỜNG CAO ĐẲNG NGHỀ SỐ 3 BQP

LUẬN VĂN THẠC SĨ
NGÀNH CÔNG NGHỆ THÔNG TIN

CHUYÊN NGÀNH: HỆ THỐNG THÔNG TIN
MÃ SỐ: 60 48 01 04

NGƯỜI HƯỚNG DẪN KHOA HỌC:
PGS.TS. NGUYỄN VĂN VỸ

LỜI CẢM ƠN

Trân trọng cảm ơn tất cả các Giáo sư, Phó giáo sư, tiến sĩ, các thầy giáo cô giáo của Khoa CNTT trường Đại Học Dân Lập Hải Phòng đã nhiệt tình giảng dạy, tạo điều kiện thuận lợi cho tác giả trong quá trình học tập, nghiên cứu, hoàn thành chương trình học tập của khóa học.

Tác giả xin trân trọng cảm ơn PGS.TS. Nguyễn Văn Vy, các thầy cô trong hội đồng khoa học đã giành thời gian chỉ bảo tận tình giúp em hoàn thành luận văn.

Tác giả xin chân thành cảm ơn Ban giám hiệu trường Cao đẳng nghề số 3-BQP đã quan tâm giúp đỡ tạo mọi điều kiện thuận lợi cho tác giả trong suốt quá trình học tập, nghiên cứu và hoàn thành luận văn.

Tác giả xin cảm ơn gia đình, bạn, đồng nghiệp, đã động viên tiếp thêm nghị lực để tác giả hoàn thành khóa học và luận văn.

Mặc dù đã có nhiều cố gắng, song luận văn khó tránh khỏi những thiếu sót. Tác giả rất mong sự chỉ bảo, góp ý của các nhà khoa học, các thầy cô giáo và đồng nghiệp.

Xin chân trọng cảm ơn!

Hải Phòng, tháng 10 năm 2016

Tác giả

Phạm Văn Hậu

LỜI CAM ĐOAN

Tôi xin cam đoan rằng, đây là công trình nghiên cứu của tôi trong đó có sự giúp đỡ rất lớn của thầy *PGS.TS. Nguyễn Văn Vy*. Các nội dung nghiên cứu và kết quả trong đề tài này là hoàn toàn trung thực.

Trong luận văn, tôi có tham khảo đến một số tài liệu của một số tác giả đã được liệt kê tại phần Tài liệu tham khảo ở cuối luận văn.

Hải phòng, ngày.....tháng.....năm 2016

Tác giả

Phạm Văn Hậu

MỤC LỤC

LỜI CẢM ƠN	1
LỜI CAM ĐOAN.....	4
BẢNG CÁC CHỮ VIẾT TẮT.....	8
DANH MỤC BẢNG.....	9
DANH MỤC HÌNH	10
Chương 1: GIỚI THIỆU VÀ MÔ TẢ BÀI TOÁN.....	12
1.1 Giới thiệu.....	14
1.2 Mô hình tổ chức và quản lý của nhà trường.....	15
1.2.1 Sơ đồ cơ cấu tổ chức của Nhà trường	15
1.2.2 Các ngành nghề đào tạo của Nhà trường.....	15
1.2.3 Quy mô đào tạo của nhà trường	16
1.2.4 Đội ngũ cán bộ quản lý, giáo viên của Nhà trường.....	17
1.3 Hoạt động quản lý và các hệ thống tin học hóa hiện có.....	17
1.4 Bài toán lập các báo cáo trong quá trình quản lý	18
1.5 Đề xuất giải pháp cho các vấn đề đặt ra.....	22
Chương 2: LÝ THUYẾT KHO DỮ LIỆU VÀ TÍCH HỢP DỮ LIỆU	24
2.1 Lý thuyết về kho dữ liệu (DW)	24
2.1.1 Định nghĩa DW.....	24
2.1.2 Các tính chất của DW	25
2.1.3 Các loại DW thường gặp	29
2.1.4 Cơ sở dữ liệu tác nghiệp và DW	30
2.1.5 Các giải pháp kiến trúc DW	32
2.1.6 Các thành phần cơ bản của kiến trúc DW	36
2.1.7 Các lược đồ dữ liệu của DW	37
2.2 Tích hợp dữ liệu	39
2.2.1 Khái niệm về tích hợp dữ liệu	39
2.2.3 Các phương pháp tích hợp dữ liệu	41
2.2.4 Môi trường tích hợp đồng nhất.....	42
2.2.5 Môi trường tích hợp không đồng nhất.....	46

2.3 Web service	49
2.3.1 Định nghĩa Web service	49
2.3.2 Đặc điểm của Web service	50
2.3.3 Nền tảng của Web service	52
2.3.4 Các công nghệ của Web service	53
2.4 Mô hình tích hợp cho ứng dụng của nhà trường	58
2.4.1 Xác định nguồn dữ liệu	58
2.4.2 Các yêu cầu về báo cáo thông kê	60
Chương 3: XÂY DỰNG KHO DỮ LIỆU BÁO CÁO THỐNG KÊ TRƯỜNG CĐ NGHỀ SỐ 3 - BQP.....	62
3.1 Thiết kế hệ thống kho dữ liệu của trường Cao đẳng nghề số 3.....	62
3.1.1 Xây dựng cơ sở dữ liệu theo chuẩn DW	62
3.1.2 Xây dựng bảng các tiêu chí	69
3.1.3 Xây dựng quy trình tích hợp dữ liệu cho DW	72
3.1.4 Thu thập và tạo lập dữ liệu	74
3.1.5 Kiến trúc luồng dữ liệu.....	81
3.2 Thiết kế Webservice.....	82
3.2.1 Cách thức hoạt động.....	82
3.2.2 Cấu trúc một message theo dạng SOAP.....	84
3.2.3 Cách truyền thông SOAP xây dựng WS	85
3.2.4 Một số Service được lập từ 3 hệ CSDL để đưa vào DW	86
3.3 Lập báo cáo	88
3.3.1 Yêu cầu nghiệp vụ của hệ thống khai thác báo cáo	88
3.3.2 Yêu cầu cụ thể của hệ thống báo cáo	89
3.3.3 Giải pháp xây dựng hệ thống báo cáo	89
3.3.4 Quá trình xây dựng báo cáo gồm các bước	91
3.3.5 Một số báo cáo dự kiến thiết kế.....	91
Chương 4: TRIỂN KHAI THỬ NGHIỆM HỆ THỐNG	96
4.1 Kho dữ liệu thử nghiệm	96
4.2 Cơ sở dữ liệu của một số đơn vị và công cụ trích rút tương ứng.....	96
4.3 Một số báo cáo thử nghiệm thực hiện với kho dữ liệu	97
4.3.1 Một số báo cáo trong quản lý nhân sự.....	97
4.3.2 Một số báo cáo trong quản lý đào tạo	98

4.3.3 Một số báo cáo trong quản lý trang thiết bị.....	98
TÀI LIỆU THAM KHẢO.....	100
A. Phụ lục 1: Các bảng dữ liệu nguồn.....	101
a. Các bảng dữ liệu về nhân sự.....	101
b. Các bảng dữ liệu quản lý đào tạo	104
c. Các bảng dữ liệu quản lý vật tư, trang bị kỹ thuật, phương tiện	107
B. Phụ lục 2: Các đoạn chương trình	109
a. Đoạn COD 1	109
b. Đoạn COD 2	110
c. Đoạn COD 3	111

BẢNG CÁC CHỮ VIẾT TẮT

Chữ viết tắt	Tên đầy đủ	Ý nghĩa
CNTT	<i>Information Technology</i>	Công nghệ thông tin
COM	<i>Common Object Model</i>	Mô hình đối tượng chung
CORBA	<i>Common Object Request Broker Architecture</i>	Kiến trúc môi giới yêu cầu đối tượng chung
CSDL	<i>Database</i>	CSDL
DBMS	<i>Database Management Systems</i>	Hệ quản trị CSDL
DCOM	<i>Distributed Component Object Model</i>	Mô hình đối tượng thành phần phân tán
DW	<i>Data Warehouse</i>	Kho dữ liệu
EDW	<i>Enterprise Data Warehouse</i>	DW doanh nghiệp
HTTT	<i>Information System</i>	Hệ thống thông tin
IBIS	<i>Internet Base Information System</i>	Hệ thống thông tin trên nền Internet
JDK	<i>Java Development Kit</i>	Bộ công cụ phát triển với Java
JRMI	<i>Java Remote Method Invocation</i>	Triệu gọi từ xa với Java
JVM	<i>Java Virtual Machine</i>	Máy ảo java
ODS	<i>Operational Data Store</i>	Kho dữ liệu tác nghiệp
OLAP	<i>Online Analytical Processing</i>	Xử lý phân tích trực tuyến
OLTP	<i>Online Transaction Processing</i>	Xử lý giao dịch trực tuyến
RMI	<i>Remote Method Invocation</i>	Triệu gọi từ xa
SOAP	<i>Simple Object Access Protocol</i>	Giao thức truy nhập đối tượng đơn giản
UDDI	<i>Universal Description, Discovery, and Integration</i>	Mô tả chung, phát hiện và tích hợp

DANH MỤC BẢNG

Bảng 1.1: Các ngành nghề đào tạo của Nhà trường.....	16
Bảng 1.2: Quy mô đào tạo của Nhà trường từ (2015 – 2020).....	17
Bảng 1.3: Đội ngũ cán bộ quản lý, giáo viên năm 2011-2015.....	17
Bảng 1.4: Hiện trạng việc ứng dụng CNTT tại trường	18
Bảng 2.1: Dữ liệu trong DW gắn với thời gian.....	28
Bảng 2.2: So sánh những đặc trưng của hệ thống tác nghiệp và DW.....	31
Bảng 3.1: Các hệ thống phần mềm đang sử dụng	62

DANH MỤC HÌNH

Hình 1.1: Sơ đồ cơ cấu tổ chức của Nhà trường	15
Hình 2.1: Tính hướng chủ đề	26
Hình 2.2: Tính tích hợp dữ liệu	27
Hình 2.3: Tính không thay đổi của dữ liệu trong DW	28
Hình 2.4: Kiến trúc cơ bản của DW	32
Hình 2.5: Kiến trúc DW	33
Hình 2.6: Kiến trúc của DW với vùng gắn nhãn (Staging area)	34
Hình 2.7: Kiến trúc Datamart độc lập	35
Hình 2.8: Kiến trúc Datamart phụ thuộc	35
Hình 2.9: Lược đồ hình sao của DW	37
Hình 2.10: Lược đồ tuyết rơi của DW	38
Hình 2.11: Minh họa đồng bộ dữ liệu (EIA)	40
Hình 2.12: Giao tiếp giữa Máy trạm và Server qua RMI	44
Hình 2.13: Kiến trúc công nghệ DCOM [11]	45
Hình 2.14: Giao tiếp giữa bên sử dụng dịch vụ và cung cấp dịch vụ	47
Hình 2.15: Kiến trúc mô hình hệ thống Web service	50
Hình 2.16: Kiến trúc của dịch vụ Web service	52
Hình 2.17: Mô hình truy nhập CSDL web bằng Java Socket	53
Hình 2.18: Mô hình truy nhập CSDL bằng Servlet	55
Hình 2.19: Mô hình truy nhập CSDL web bằng RMI	56
Hình 2.20: Mô hình truy nhập cơ sở dữ web bằng Java CORBA	57
Hình 3.1: Kiến trúc DW BCTK	67
Hình 3.2: Cơ sở dữ liệu DW	69
Hình 3.3: Sơ đồ quan hệ chủ đề Quản Lý Nhân Sự	70
Hình 3.4: Sơ đồ quan hệ chủ đề Quản Lý Đào Tạo	71
Hình 3.5: Sơ đồ quan hệ chủ đề Quản lý thiết bị	72
Hình 3.6: Các bước của quá trình tích hợp dữ liệu vào DW	74
Hình 3.7: Quy trình tạo lập dữ liệu	75
Hình 3.8: Biểu đồ hoạt động (UML) lọc tinh chế dữ liệu	77
Hình 3.9: UML Sequence lọc tinh chế dữ liệu	78
Hình 3.10: Quy trình thu thập vào chuyển đổi dữ liệu	78
Hình 3.11: Kiến trúc tích hợp dữ liệu	80
Hình 3.12: Kiến trúc luồng dữ liệu trường Cao đẳng nghề số 3 BQP	81

Hình 3.13: Cấu trúc thông điệp SOAP	84
Hình 3.14: Cấu trúc một message theo dạng SOAP	84
Hình 3.15: Modul Service Quản Lý Đào Tạo	86
Hình 3.16: Modul Service Quản Lý Nhân Sự.....	87
Hình 3.17: Modul Service Quản Lý Thiết Bị.....	88
Hình 3.18: Quy trình giải pháp Business Intelligence (BI).....	90
Hình 3.19: Tổ chức báo cáo hệ thống	92
Hình 3.20: Báo cáo Thống Kê Sinh Viên Theo Ngành Học.....	92
Hình 3.21: Báo cáo Thống Kê Sinh Viên Tại Trường	93
Hình 3.22: Báo cáo Thống Kê Số Liệu Tuyển Sinh Qua Các Năm Đào Tạo.....	93
Hình 3.23: Báo cáo Thống Kê Danh Mục Vật Tư	94
Hình 3.24: Báo cáo Thống Kê Vật Tư Nhập Theo Năm	94
Hình 3.25: Báo cáo Thống Kê Danh Sách Cán Bộ Giảng Viên Trong Trường	95
Hình 3.26: Báo cáo Thống Kê Trình Độ Nhân Sự Trong Trường.....	95
Hình 4.1: Báo cáo danh sách cán bộ giáo viên	97
Hình 4.2: Báo cáo kết quả học sinh tốt nghiệp	98
Hình 4.3: Báo cáo cơ sở vật chất trang thiết bị dạy nghề	98

MỞ ĐẦU

1 Tính cấp thiết của đề tài

Trường cao đẳng nghề số 3-BQP có nhiệm vụ đào tạo nguồn lực cho toàn xã hội, đặc biệt cho lực lượng vũ trang và bán vũ trang. Hàng tháng, quý, năm, Thủ trưởng, Ban giám hiệu nhà trường yêu cầu các cơ quan, đơn vị tổng hợp và làm các báo cáo về các lĩnh vực như *cơ cấu tổ chức cán bộ, quản lý đào tạo, cơ sở vật chất trang thiết bị phục vụ cho công tác dạy nghề, công tác thu chi và quyết toán tài chính* v.v... Từ đó nhà trường tiến hành tổng hợp để tạo ra các báo cáo hỗ trợ việc ra các quyết định của nhà trường và gửi đến các cơ quan liên quan.

Khi tiến hành làm báo cáo, nhà trường phải cập nhật lại các dữ liệu thu thập được, tiến hành các điều chỉnh cần thiết về nội dung, định dạng dữ liệu, thực hiện các xử lý để đưa ra các báo cáo khác nhau theo từng yêu cầu cụ thể.

Xuất phát từ hiện trạng trên, khi nghiên cứu về các giải pháp công nghệ hiện có và cũng như khả năng thực thi của nó ở cơ sở, đề tài “*Ứng dụng kho dữ liệu và webservice để tích hợp dữ liệu xây dựng hệ thống báo cáo thống kê tại trường cao đẳng nghề số 3 BQP*” được chọn làm đề tài luận văn cao học của tôi.

2. Mục tiêu và nhiệm vụ nghiên cứu

Mục tiêu của đề tài tìm hiểu lý thuyết về kho dữ liệu Data Warehouse và tích hợp dữ liệu bằng công cụ webservice để rút trích dữ liệu để xây dựng kho dữ liệu báo cáo thống kê và hỗ trợ ra quyết định.

3. Đối tượng và phạm vi nghiên cứu

Đối tượng nghiên cứu: Tìm hiểu hệ thống báo cáo thông kê tại Trường Cao đẳng nghề số 3- BQP. Khảo sát các hệ thống nguồn xây dựng kho dữ liệu báo cáo thống kê

Phạm vi nghiên cứu: Đề tài được nghiên cứu và thực hiện tại Trường Cao đẳng nghề số 3- BQP

4. Phương pháp nghiên cứu

- Phương pháp nghiên cứu lý thuyết: kho dữ liệu và web service
- Phương pháp phân tích và thiết kế kho dữ liệu
- Phương pháp thực nghiệm

5. Ý nghĩa khoa học của đề tài

- Đề xuất một cách trích rút dữ liệu trong việc xây dựng kho dữ liệu từ nhiều kho dữ liệu đa dạng và phân tán
- Tạo một kho dữ liệu và bộ công cụ đáp ứng đầy đủ và nhanh chóng yêu cầu lập các báo cáo thống kê và quản lý đào tạo tại trường cao đẳng nghề số 3 BQP đạt hiệu quả cả về chi phí và triển khai

6. Bố cục của luận văn

Luận văn được chia làm 4 chương

Chương 1: Giới thiệu và mô tả bài toán

Chương 2: Lý thuyết kho dữ liệu và tích hợp dữ liệu

Chương 3: Xây dựng kho dữ liệu lập báo cáo thống kê tại trường cao đẳng nghề số 3- BQP

Chương 4: Triển khai thử nghiệm hệ thống được xây dựng

Chương 1: GIỚI THIỆU VÀ MÔ TẢ BÀI TOÁN

1.1 Giới thiệu

Trường Cao đẳng nghề số 3 – BQP tiền thân là trường Lái xe Quân khu 3, được thành lập ngày 17 tháng 3 năm 1960. Ngày 11 tháng 10 năm 1999, Bộ trưởng Bộ Quốc phòng có Quyết định số 1702/1999/QĐ-QP về việc nâng cấp trường thành Trường đào tạo nghề khu vực miền Bắc. Ngày 21 tháng 11 năm 2003, Bộ trưởng Bộ Quốc phòng có Quyết định số 276/2003/QĐ-QP đổi tên trường thành Trường Dạy nghề số 3 – BQP. Ngày 13 tháng 7 năm 2007, Bộ trưởng Bộ Quốc phòng có Quyết định số 112/2007/QĐ-BQP đổi tên trường thành trường Cao đẳng nghề số 3 – BQP.

Qua hơn 50 năm xây dựng và trưởng thành, trường Cao đẳng nghề số 3 – BQP đã trở thành một cơ sở lớn đào tạo nguồn nhân lực cho toàn quân và cho công cuộc công nghiệp hóa, hiện đại hóa đất nước đất nước. Để đảm bảo và nâng cao chất lượng đào tạo, Nhà trường đặc biệt coi trọng các biện pháp nâng cao chất lượng đội ngũ giáo viên và cán bộ quản lý giáo dục, chủ động lựa chọn, tạo nguồn, tích cực bồi dưỡng đào tạo bằng nhiều loại hình thức như cử đi học nâng cao nghiệp vụ tại các cơ sở giáo dục có uy tín ở trong nước cũng như ở nước ngoài. Nhờ sự đóng góp lớn lao của cán bộ, giáo viên, nhân viên và học viên trong toàn trường qua các thời kỳ đã được Đảng, Nhà nước và Quân đội tặng thưởng nhiều phần thưởng cao quý đó là Huân chương Lao động hạng nhất, Huân chương Lao động hạng nhì và nhiều phần thưởng cao quý của các Bộ, ngành liên quan.

Tên giao dịch: TRƯỜNG CAO ĐẲNG NGHỀ SỐ 3 – BỘ QUỐC PHÒNG

Tên giao dịch quốc tế: Vocational College N. 3 Ministry of Defence

Địa chỉ trụ sở chính:

Địa chỉ: Số 248 Lê Duẩn - Bắc Sơn, quận Kiến An, thành phố Hải Phòng

Điện thoại : 0313 877 987 – 0313 876 250

Fax : 0313 790 218

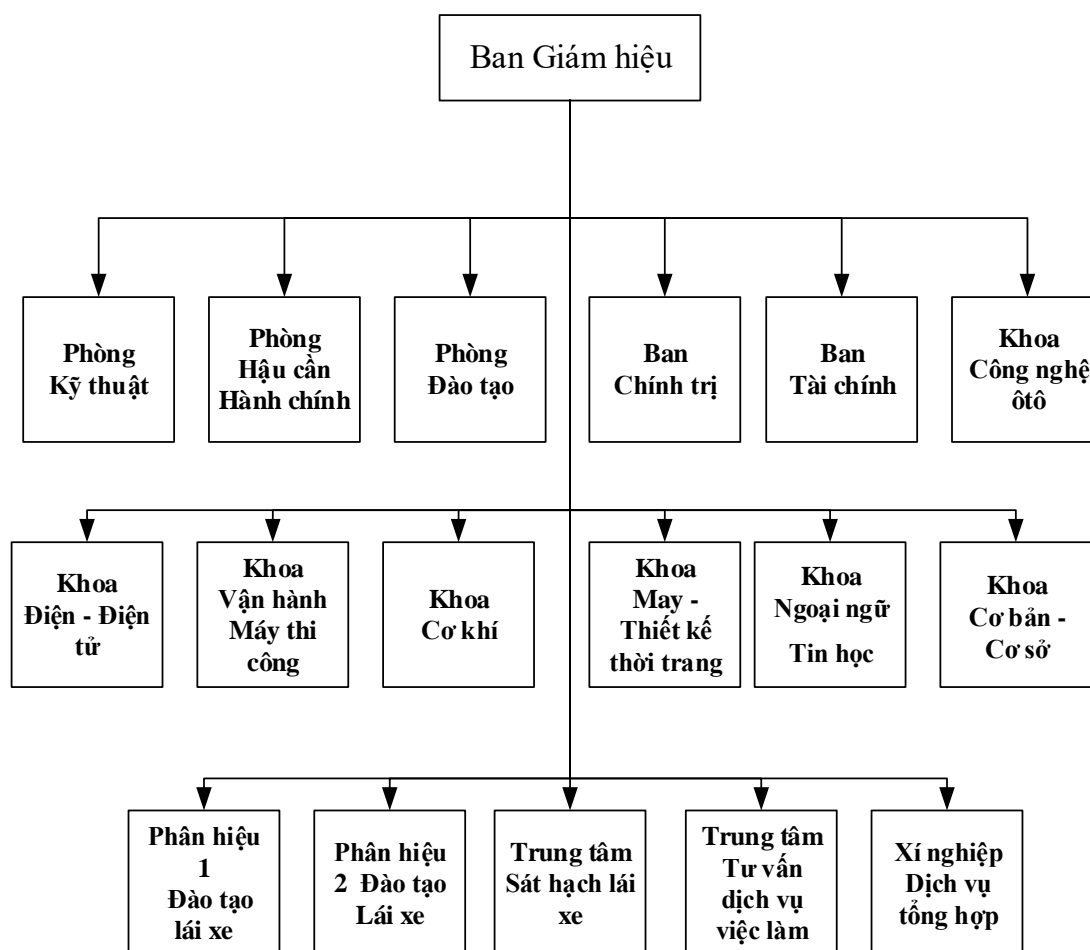
Website : <http://www.cdns3.edu.vn>

Email: truongcdns3@gmail.com

1.2 Mô hình tổ chức và quản lý của nhà trường

1.2.1 Sơ đồ cơ cấu tổ chức của Nhà trường

Cơ cấu tổ chức của nhà trường được mô tả ở hình 1.1. bao gồm Ban giám hiệu, bên dưới là các phong ban, các khoa và các trung tâm



Hình 1.1: Sơ đồ cơ cấu tổ chức của Nhà trường

1.2.2 Các ngành nghề đào tạo của Nhà trường

Trường Cao đẳng nghề số 3 so với các trường nghề khác trong và ngoài Quân đội là một trường có bề dày với hơn 50 năm thực hiện công tác đào tạo nghề (Nhà trường được thành lập từ năm 1960). Ngoài nhiệm vụ tổ chức đào tạo nghề, Nhà trường còn mở các lớp bồi dưỡng thi nâng bậc thợ cho các đơn vị trong và ngoài quân đội, đồng thời còn tổ chức liên kết đào tạo với Học viện kỹ thuật Quân sự, ĐHSPKT

Nam Định, ĐHSPTK Hưng Yên và một số cơ sở đào tạo nghề khác để thực hiện tốt nhiệm vụ được giao.

Bảng 1.1: Các ngành nghề đào tạo của Nhà trường

TT	Nghề đào tạo	Cao đẳng nghề	Trung cấp nghề	Sơ cấp nghề
1	Công nghệ ô tô	30 tháng	20 tháng	09 tháng
2	Hàn	30 tháng	20 tháng	09 tháng
3	Điện công nghiệp	30 tháng	20 tháng	06 tháng
4	KT máy lạnh và điều hoà không khí	30 tháng	20 tháng	06 tháng
5	Cơ điện tử			
6	Kế toán doanh nghiệp	30 tháng	20 tháng	06 tháng
7	Vận hành, SC thiết bị lạnh		20 tháng	06 tháng
8	Vận hành máy xúc		20 tháng	06 tháng
9	Vận hành cần trục		20 tháng	06 tháng
10	Kỹ thuật xây dựng		20 tháng	06 tháng
11	May và thiết kế thời trang		20 tháng	
12	Kỹ thuật SC, lắp ráp M. tính		20 tháng	
13	Điện dân dụng		20 tháng	06 tháng
14	Sửa chữa xe máy		20 tháng	06 tháng
15	Bảo vệ			03 tháng
16	Vận hành máy nâng chuyển			03 tháng

1.2.3 Quy mô đào tạo của nhà trường

Trong mấy năm gần đây, hàng năm, tổng số học sinh trường đào tạo đều trên ba nghìn em, bao gồm cao đẳng, trung cấp và sơ cấp (bảng 1.2)

Bảng 1.2: Quy mô đào tạo của Nhà trường từ (2010 – 2020)

Năm học	Đào tạo nghề			Tổng số
	Cao đẳng	Trung cấp	Sơ cấp	
2010- 2011	400	450	2200	3100
2012 – 2013	450	450	2500	3400
2014 - 2015	600	500	2700	3800
2016 - 2017	800	600	3000	4400

1.2.4 Đội ngũ cán bộ quản lý, giáo viên của Nhà trường

Tổng số cán bộ công nhân viên nhà trường hiện nay là 237 người, trong đó 125 là giáo viên. Số còn lại là cán bộ quản lý và nhân viên (Bảng 1.3). Trong đó chỉ có 28 người có trình độ sau đại học, chiếm 11,8% tổng số công nhân viên.

Bảng 1.3: Đội ngũ cán bộ quản lý, giáo viên năm 2011-2015

Đối tượng	Số lượng	Giới tính		Độ tuổi		Thâm niên		Trình độ chuyên môn			
		Tổng	Nam	Nữ	≤35	≥35	≤ 15	≥15	TH	Sau ĐH	ĐH CD
Cán bộ quản lý	54	38	16	14	40	14	40	4	6	44	0
Giáo viên	125	84	41	38	87	38	87	15	22	76	12
Nhân viên phục vụ	58	27	31	42	16	42	16	0	0	12	46
Tổng	237	149	88	94	143	94	143	19	28	132	58

1.3 Hoạt động quản lý và các hệ thống tin học hóa hiện có

Hiện nay nhà trường đã ứng CNTT dụng cho những hoạt động chính, như quản lý đào tạo, quản lý bảo hiểm quân đội, quản lý nhân viên kỹ thuật và trang thiết bị,

quản lý kế toán. Các phần mềm này được các đơn vị phát triển phần mềm khác nhau xây dựng và trên nền tảng công nghệ khác nhau, như Hệ quản trị CSDL SQL Server, Net FrameWork SQL Server, Net FrameWork MS_ACCESS và FOXPRO For WINDOWS (bảng 1.4). Chính vì vậy, chúng hoạt động như những hệ thống riêng rẽ, nên các kết quả của từng hệ thống chỉ đáp ứng được yêu cầu cơ bản của mỗi hoạt động quản lý. Tuy nhiên, việc tổng dữ liệu cho các báo cáo theo định kỳ hoặc làm những báo cáo phân tích phục vụ cho việc lập kế hoạch ngắn hay dài hạn gặp rất nhiều khó khăn. Nhiều thao tác lấy dữ liệu từ các hệ thống riêng lẻ phải thực hiện thủ công, khi tạo báo cáo phải làm đồng bộ dữ liệu và thực hiện các xử lý phải viết chương trình hay thao tác trực tiếp nên tốn nhiều thời gian. Vấn đề trở nên gay go và phức tạp khi kết quả không đúng đắn phải làm lại, thường hay chậm trễ, không đáp ứng được yêu cầu về thời gian.

Bảng 1.4: Hiện trạng việc ứng dụng CNTT tại trường

TT	Đơn vị	Dữ liệu sử dụng	Chương trình xử lý dữ liệu	Phần mềm đã sử dụng
1	Phòng đào tạo	Giáo viên, học sinh, kế hoạch đào tạo, quản lý điểm, quản lý tốt nghiệp, báo cáo thống kê, tổng hợp	Hệ quản trị CSDL SQL Server	Quản lý đào tạo
2	Ban Quân lực	Hồ sơ cán bộ, bảo hiểm xã hội, bảo hiểm y tế	Nền tảng công nghệ Net FrameWork SQL Server	Phần mềm Bảo hiểm quân đội ASMS
3	Phòng Kỹ thuật	Nhân viên kỹ thuật và trang bị vật tư kỹ thuật	Nền tảng công nghệ Net FrameWork SQL Server	Phần mềm quản lý nhân viên KT và trang bị
4	Ban Tài chính	Quản lý tài chính, thu học phí, cân đối thu chi và tiền lương	Sử dụng nền FOXPRO For WINDOWS	Phần mềm quản lý kế toán

1.4 Bài toán lập các báo cáo trong quá trình quản lý

Trường cao đẳng nghề số 3-BQP có nhiệm vụ đào tạo nguồn lực cho toàn xã hội, đặc biệt cho lực lượng vũ trang và bán vũ trang. Vì vậy tổ chức biên chế của Nhà

trường trực thuộc Bộ quốc phòng quản lý, còn lĩnh vực chuyên môn đào tạo nghề cho các đối tượng chính sách xã hội trực thuộc Tổng cục dạy nghề Bộ lao động thương binh và xã hội. Chính vì vậy theo định kỳ và đột xuất Nhà trường thường xuyên lập các báo cáo thống kê gửi các cơ quan chức năng chuyên môn của Bộ quốc phòng và Bộ lao động thương binh xã hội. Ngoài ra Nhà trường cần những số liệu thống kê chính xác để hỗ trợ cho việc ra quyết định trong lĩnh vực quản lý và dạy nghề.

Thực tế hiện nay tại Nhà trường bộ phận báo cáo thống kê hàng tháng, quý phải xây dựng các mẫu biểu báo cáo thống kê về cơ cấu tổ chức quản lý nhân sự như danh sách đội ngũ cán bộ giáo viên và nhân viên phục vụ. Trong lĩnh vực đào tạo thống kê được số lượng kết quả tuyển sinh theo đợt, theo năm, theo đối tượng tuyển sinh, kết quả đào tạo theo ngành nghề, danh sách học sinh sinh viên ra trường và việc làm sau đào tạo. Trong lĩnh vực quản lý vật tư trang thiết bị dạy nghề báo cáo được số lượng, chủng loại, chất lượng, năm sản xuất, năm sử dụng, mã ngành nghề của các loại vật tư trang thiết bị, phương tiện dạy nghề. Các loại mẫu biểu sổ sách thu chi tài chính v.v.

Tất cả các công việc đó Nhà trường đang làm thủ công, điều đó dẫn đến những bất cập như khi có yêu cầu bộ phận được giao nhiệm vụ thống kê báo cáo yêu cầu các cơ quan đơn vị trong Nhà trường cung cấp số liệu của cơ quan chuyên môn lĩnh vực bộ phận mình phụ trách gửi cho bộ phận báo cáo thống kê. Chính vì vậy xảy ra một số hạn chế như sau:

- Sự trùng lặp và dư thừa dữ liệu ví dụ như cùng một đối tượng quản lý con người bộ phận tổ chức nhân sự báo cáo danh sách cán bộ sĩ quan, quân nhân chuyên nghiệp, công nhân viên quốc phòng và lao động hợp đồng, mặt khác bộ phận quản lý đào tạo lại báo cáo danh sách cán bộ giáo viên và nhân viên phục vụ hay trong lĩnh vực quản lý trang thiết bị phòng kỹ thuật lại báo cáo vật tư trang bị phương tiện phục vụ huấn luyện còn trong quản lý đào tạo lại báo cáo vật tư trang thiết bị dạy nghề v.v
- Dữ liệu và tên gọi không thống nhất ví dụ mỗi cơ quan đơn vị có thể sử dụng những phần mềm riêng rẽ do vậy dữ liệu định dạng là khác nhau không thống nhất chẳng hạn như bộ phận quản lý nhân sự sử dụng Mã sĩ quan, Mã quân nhân

chuyên nghiệp, Mã nhân viên kỹ thuật, còn trong quản lý Cán bộ giáo viên lại sử dụng Mã giáo viên, Mã nhân viên phục vụ hay trong quản lý trang thiết bị trong quân đội gọi là nhóm phương tiện vật tư huấn luyện, còn trong quản lý đào tạo gọi là mã vật tư trang thiết bị dạy nghề.

- Tính thời gian do mọi thao tác bằng phương pháp thủ công chính vì vậy khi các báo cáo lập ra không đảm bảo tính kịp thời, ví dụ trong các hệ thống tác nghiệp của các cơ quan đơn vị được cập nhập thường xuyên theo từng ngày giờ vì vậy dữ liệu khi gửi cho bộ phận báo cáo thống kê ngày hôm nay khác với ngày hôm qua.
- Tính không đầy đủ: Chẳng hạn khi tiến hành báo cáo có những nguồn thông tin đầy đủ và chi tiết, có những nguồn thông tin còn thiếu so với yêu cầu chính vì vậy người lập báo cáo phải bổ xung , chỉnh sửa một cách thủ công.
- Không có tính kế thừa dữ liệu ví dụ mỗi khi thực hiện các yêu cầu báo cáo bộ phận báo cáo thống kê lại cập nhập dữ liệu từ đầu chứ không sử dụng lại dữ liệu đã có và bổ xung thêm phần dữ liệu mới phát sinh.

Mặt khác các thông tin dữ liệu về giáo viên, học viên, cơ sở vật chất, trang thiết bị, nguồn lực tài chính, các dữ liệu về chính sách xã hội, các chế độ ưu đãi đối với học viên là quân nhân xuất ngũ được tích lũy qua rất nhiều năm và được lưu trữ ở nhiều nơi, nhiều cơ quan, đơn vị khác nhau, với nhiều định dạng khác nhau trên nhiều phương tiện lưu trữ khác nhau và có chế độ bảo mật khác nhau. Mỗi cơ quan đơn vị đều xây dựng các ứng dụng công nghệ thông tin độc lập (trong đó gồm cả phần mềm và cơ sở dữ liệu)

Đặc biệt khi các tiêu chí báo cáo nhiều thì việc làm này mất rất nhiều thời gian và công sức đôi khi không chính xác và kịp thời dẫn đến việc hỗ trợ ra quyết định và định hướng trong lĩnh vực quản lý đào tạo và dạy nghề không chính xác cũng như gây ra sự lãng phí trong đầu tư cơ sở vật chất trang thiết bị dạy nghề v.v... Chính điều này, đã tạo ra những khó khăn nhất định trong việc thống kê báo cáo khi dữ liệu ngày càng lớn và các tiêu chí báo cáo ngày càng nhiều.

Xuất phát từ những bất cập ở trên bài toán đặt ra là làm thế nào để xây dựng được một hệ thống báo cáo thống kê một cách đơn giản, có thể kết nối đến tất cả những người sử dụng, những người tham gia vào các hoạt động báo cáo của Nhà trường, dữ liệu của các hệ thống tác nghiệp có thể tích hợp thành một kho dữ liệu. Các thông tin trong kho dữ liệu lớn, đa dạng và là nền tảng giúp cho việc báo cáo thống kê một cách đầy đủ và chi tiết các tiêu chí từ tổng hợp đến chi tiết một cách thống nhất và kịp thời và giúp cho nhà lãnh đạo quản lý đưa ra những quyết định và chiến lược phát triển trong tương lai qua đó tôi đưa ra giải pháp ***xây dựng kho dữ liệu Data Warehouse phục vụ cho việc báo cáo thống kê tại trường Cao đẳng nghề số 3 – BQP***, việc tạo ra kho dữ liệu với đầy đủ các dữ liệu cần thiết và có các công cụ lập báo cáo sẵn sàng sẽ đáp ứng dễ dàng mọi yêu cầu quản lý của lãnh đạo. Hơn nữa, với kho dữ liệu tại trường, Nhà trường có thể đưa ra ngay các báo cáo mỗi khi cần thiết mà không phải chờ đợi hay mất nhiều thời gian xử lý, mà chúng là nguyên nhân gây ra sự chậm chễ hiện nay.

Có nhiều cách tạo kho dữ liệu trong điều kiện hiện trạng của nhà trường, như lập cơ sở dữ liệu phân tán, tích hợp dữ liệu theo cách truy nhập cơ sở dữ liệu từ xa. Tuy nhiên, lập kho dữ liệu với cách trợ giúp của webservice để trích rút chỉ dữ liệu cần thiết từ các cơ sở dữ liệu độc lập là tiện lợi cả về chi phí và cách tổ chức thực hiện. trên cơ sở tích hợp và sử dụng lại các nguồn thông tin hiện có của các cơ sở dữ liệu tác nghiệp trong Nhà trường.

Với sự hỗ trợ tổ chức và khai thác dữ liệu hướng đối tượng, từ đó cung cấp khả năng khai thác dữ liệu theo nhiều mức khác nhau: tổng thể hoặc chi tiết; biến động theo thời gian; so sánh giữa các đối tượng báo cáo và theo nhiều chiều khác nhau như theo chiều thời gian, theo các đơn vị báo cáo, ...

- *Cung cấp mô hình phân tích thực trạng hệ thống dạy nghề bao gồm các chính sách của Chính phủ, hành lang pháp lý, phân tích các xu hướng phát triển nghề nghiệp trong nước và quốc tế.*
- *Cung cấp thông tin về cơ sở vật chất trang thiết bị, hạ tầng cơ sở thống kê được số diện tích, phòng học, nhà xưởng hiện có, tình hình tiềm năng của từng đơn*

vị, khoa cho từng chuyên ngành và đưa ra được nhu cầu sử dụng và phát triển trong giai đoạn tới.

- *Cung cấp thông tin về năng lực đội ngũ cán bộ giáo viên thống kê được số lượng, chất lượng, trình độ đội ngũ cán bộ giáo viên, nhu cầu còn thiếu bao nhiêu từ đó đề ra các phương án tuyển mới hay đào tạo lại bằng các chương trình đào tạo bồi dưỡng cán bộ giáo viên chính sách đãi ngộ như lương, thưởng, v.v. Từ đó xây dựng chiến lược phát triển nguồn nhân lực, đội ngũ giảng viên trong những năm tới như thế nào ?*
- *Hỗ trợ cải thiện chất lượng dạy nghề, lựa chọn phương pháp và công nghệ giảng dạy, lựa chọn hệ thống đánh giá kiểm tra đảm bảo cơ hội việc làm cho sinh viên sau khi tốt nghiệp.*
- *Hỗ trợ cho việc xây dựng các quan hệ đối tác căn cứ vào nhu cầu lao động theo từng giai đoạn (trong nước và ngoài nước), số lượng các doanh nghiệp và các khu công nghiệp.*
- *Cung cấp thông tin về nhu cầu học sinh đã đăng ký học nghề, số lượng đối tượng chính sách được ưu tiên, số lượng và chất lượng có việc làm sau đào tạo để hỗ trợ cho việc xây dựng chiến lược thu hút học sinh.*

1.5 Đề xuất giải pháp cho các vấn đề đặt ra

Nghiên cứu cơ sở lý thuyết về Data Warehouse và Webservices

Nghiên cứu các bước triển khai Data Warehouse bằng công cụ xây dựng và quản lý DW SQL Server 2008.

Xây dựng quy trình tích hợp dữ liệu từ các nguồn CSDL có sẵn và các các hệ thống tác nghiệp hiện có của Nhà trường vào DW bằng công cụ Web services, các nguồn dữ liệu này phải được xử lý, làm sạch, đồng bộ nhất quán, chính xác và tích lũy theo thời gian trước khi đưa vào kho phục vụ cho nhu cầu báo cáo thống kê của Nhà trường.

Xây dựng các giải pháp để khai thác các báo cáo trên cơ sở đã có Data Warehouse.

Hệ thống báo cáo phải phục vụ cho việc lập báo cáo theo mẫu có sẵn như báo cáo thống kê theo từng định kỳ báo cáo, báo cáo biến động theo thời gian, theo các tiêu chí báo cáo, theo các đầu mỗi đơn vị báo cáo.

Hệ thống phải phục vụ cho việc phân tích số liệu theo hệ thống tiêu chí như: tiêu chí về Cán bộ giáo viên, nhu cầu học nghề của học sinh sinh viên, nhu cầu về trang thiết bị vật chất phục vụ cho việc dạy nghề, nguồn lực tài chính, chế độ ưu đãi v.v...

Hệ thống báo cáo phải có yếu tố thời gian, kịp thời hỗ trợ cho việc ra quyết định điều hành và quản lý, xây dựng chiến lược về công tác dạy nghề của Nhà trường.

Cung cấp giao diện đơn giản, hỗ trợ khai thác hiệu quả báo cáo đã có trong CSDL với yêu cầu kiến thức CNTT người sử dụng ở mức thông thường.

Cung cấp khả năng truy xuất thông tin đột xuất và tức thời của báo cáo được cung cấp bởi dữ liệu có sẵn trong kho dữ DW lưu trữ hoặc được lấy từ các CSDL nghiệp vụ trên hệ thống một cách nhanh chóng người sử dụng chỉ cần đăng nhập vào hệ thống , Chọn chức năng lấy dữ theo liệu yêu cầu nghiệp vụ, chọn chức năng báo cáo.

Các mẫu báo cáo có thể chỉnh sửa số liệu theo yêu cầu nghiệp vụ trước khi xuất và in ấn và báo cáo có thể xuất ra dưới nhiều định dạng như Word, Excel, PDF. Dữ liệu trên báo cáo được thể hiện dưới các dạng báo cáo theo bảng biểu hoặc đồ thị , cột.... .

Tùy vào yêu cầu nghiệp vụ các báo cáo được trích rút theo các tiêu chí đặc trưng khác nhau như về khoảng thời gian, năm học, loại thiết bị, ngành nghề, cấp bậc, chức vụ công tác....Nhằm giúp quá trình quản lý tại trường được chính xác nhanh chóng hiệu quả đúng theo yêu cầu quản lý.

Có khả năng quản lý người sử dụng linh động, an toàn, bảo mật dữ liệu báo cáo theo thẩm quyền khai thác. Với mỗi người dùng trên hệ thống tùy từng vai trò chức năng được cấp trên phân công các cán bộ nghiệp vụ chỉ được thao tác quyền trên các nghiệp vụ được phân quyền tương ứng trên hệ thống và không có quyền thao tác lựa chọn, trích xuất các báo cáo nghiệp vụ không được phân công phân quyền.

Chương 2: LÝ THUYẾT KHO DỮ LIỆU VÀ TÍCH HỢP DỮ LIỆU

2.1 Lý thuyết về kho dữ liệu (DW)

2.1.1 Định nghĩa DW

Kho dữ liệu (Data Warehouse – DW) là hướng tiếp cận phổ biến cho các hệ thống dữ liệu lớn và có tính tích lũy tri thức cao. Nó đã được định nghĩa theo nhiều cách. Cho đến nay, chưa có một cách định nghĩa nào được xem là hoàn toàn chuẩn xác. Từ cách nhìn lưu trữ, DW được xem như là một Cơ sở dữ liệu (CSDL) lớn, được tổng hợp từ nhiều nguồn dữ liệu khác nhau, được tổ chức theo những cách riêng biệt ở trong hay ngoài doanh nghiệp (như từ các CSDL tác nghiệp khác nhau, từ các dự báo hay dữ liệu của khách hàng) để hỗ trợ phân tích hay tổng hợp trên CSDL lịch sử và các công cụ truy vấn dữ liệu mạnh. Việc sử dụng DW sẽ tạo ra một sự đồng nhất về thông tin doanh nghiệp, và từ đó doanh nghiệp có thể nhận được các chỉ số phân tích, hoặc dùng các công cụ hỗ trợ để theo dõi các chỉ số cần quan tâm. DW có thể xem là môi trường có cấu trúc các Hệ thống thông tin (HTTT), cung cấp cho người dùng những thông tin khó có thể truy vấn hoặc biểu diễn trong CSDL tác nghiệp truyền thống, nhằm mục đích hỗ trợ việc ra quyết định mang tính lịch sử hoặc hiện tại [7], người được xem là kiến trúc sư đầu tiên xây dựng DW đã định nghĩa: “*Kho dữ liệu là tập hợp dữ liệu hướng chủ đề, được tích hợp, gắn với thời gian và ổn định, được thiết kế cho việc hỗ trợ ra quyết định của người quản lý*” [9].

Về bản chất, DW hướng vào việc cung cấp một kiến trúc, công cụ để phát triển dữ liệu của các hệ thống tác nghiệp thành dữ liệu của hệ thống hỗ trợ ra quyết định. Dữ liệu được phát sinh từ các hoạt động hàng ngày và được thu thập xử lý để phục vụ công việc, nghiệp vụ cụ thể của một tổ chức thường được gọi là *dữ liệu tác nghiệp* (Operational Data). Các hoạt động thu thập, xử lý dữ liệu tác nghiệp được gọi là *xử lý giao dịch trực tuyến* (Online Transaction Processing – OLTP). Trái lại DW phục vụ cho việc phân tích với kết quả có hàm lượng thông tin cao. Các hệ thống thu thập xử lý dữ liệu loại này còn gọi là *xử lý phân tích trực tuyến* (Online Analytical Processing – OLAP). DW thường rất lớn tới hàng trăm GB hay thậm chí hàng Terabyte.

DW được xây dựng nhằm mục đích làm tiện lợi cho việc truy cập theo nhiều nguồn, nhiều kiểu dữ liệu khác nhau, sao cho có thể kết hợp được cả những ứng dụng của các hệ thống hiện đại và kế thừa được từ những hệ thống đã có sẵn từ trước. Nó được thiết kế để quản lý dữ liệu, cung cấp cho những ứng dụng yêu cầu một khối lượng dữ liệu lớn. DW là thể hiện mô hình hỗ trợ quyết định dạng quan hệ hoặc đa chiều được thiết kế cho việc truy vấn dữ liệu và phân tích hơn là cho việc xử lý giao dịch. Chúng thường chứa dữ liệu lịch sử và cho phép lấy dữ liệu từ nhiều nguồn cung cấp cho doanh nghiệp những thông tin mà họ cần để đưa ra các quyết định chiến lược.

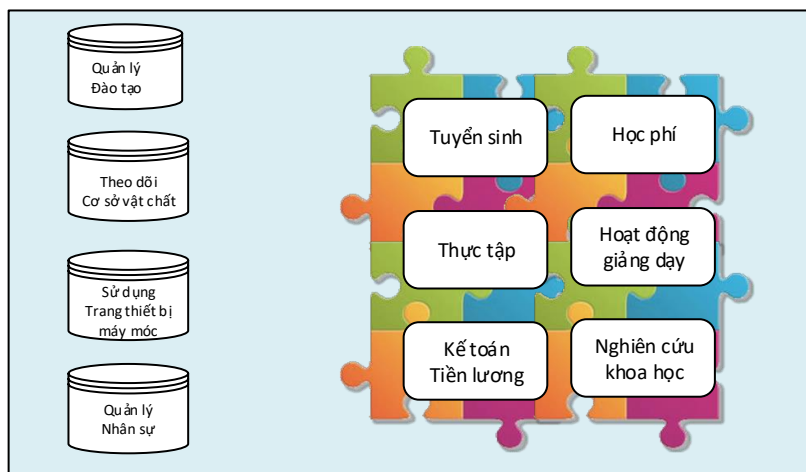
2.1.2 Các tính chất của DW

Một DW được xác định là một CSDL, trong đó có chứa các đặc trưng cơ bản nhất như hướng chủ đề, tính ổn định, được tích hợp, gắn với thời gian.

2.1.2.1 Hướng chủ đề

Hướng chủ đề có nghĩa là tất cả các dữ liệu được tổ chức xung quanh một chủ đề và được tập trung lại, lưu trữ trong cùng một tập có cấu trúc đơn giản dễ sử dụng. Thông thường, các ứng dụng tác nghiệp được thiết kế nhằm giải quyết một số giao dịch, công việc hằng ngày và dữ liệu được thiết kế nhằm mục đích phục vụ công việc đó. Trong khi đó, dữ liệu trong DW lại được dùng để mục đích phân tích và hỗ trợ ra quyết định. Đây là những chức năng phức tạp nên cần được cân nhắc nhiều khía cạnh, tiêu chuẩn và sự lựa chọn khác nhau. Dữ liệu trong DW cần được tích hợp theo cách để có thể thực hiện công việc này một cách tốt nhất, đó là phương thức tập hợp dữ liệu hướng chủ đề. Vì vậy, không phải toàn bộ các thông tin từ cơ sở tác nghiệp đều được đưa vào DW mà phải chọn lọc theo những chủ đề thích hợp, chứa thông tin cần thiết phục vụ cho việc thống kê, phân tích, báo cáo và hỗ trợ ra quyết định. Ví dụ, trong hệ thống quản lý tài chính cũ (có tính tác nghiệp) có thể dữ liệu được tổ chức theo chức năng: *cho vay, quản lý tín dụng, quản lý ngân sách*, v. v. Sự khác nhau hai cách tiếp cận trên dẫn đến sự khác nhau về nội dung dữ liệu được lưu trữ trong hệ thống. DW không lưu trữ dữ liệu chi tiết, chỉ lưu trữ dữ liệu có tính tổng hợp phục vụ chủ yếu cho quá trình phân tích để trợ giúp quyết định.

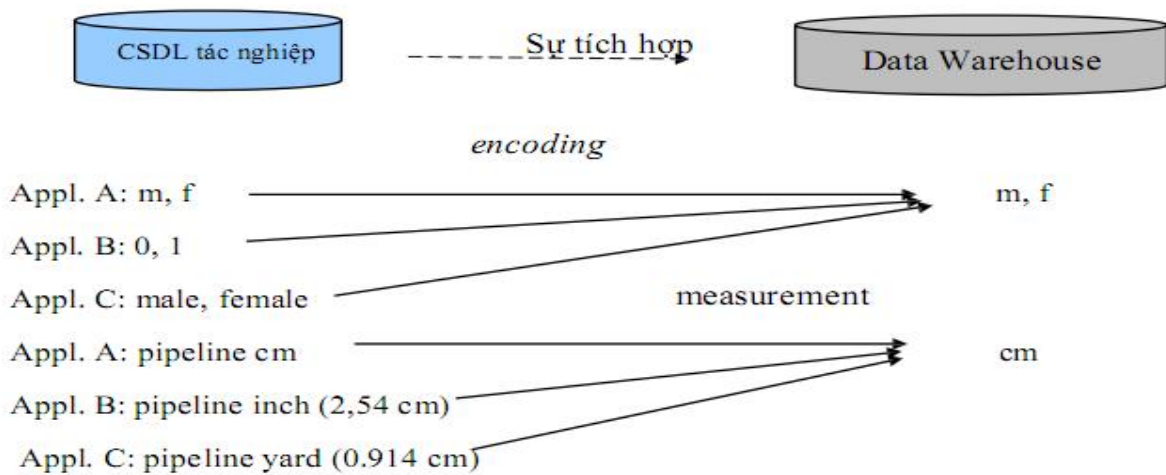
Hệ tác nghiệp DW



Hình 2.1: Tính hướng chủ đề

2.1.2.2 Tích tích hợp

Một DW được xây dựng bằng việc tích hợp nhiều nguồn dữ liệu, lưu trữ trong một cấu trúc thống nhất, toàn vẹn về việc đặt tên, đơn vị đo, cấu trúc mã hóa, các thuộc tính vật lý... Tích hợp dữ liệu là một vấn đề luôn được đặt ra của hầu hết các tổ chức, đặc biệt tích hợp dữ liệu phải đòi hỏi phải sử dụng đến nhiều công nghệ chuyên dụng đang được biết đến. Thông thường, dữ liệu trong DW được tích hợp từ nhiều nền tảng hệ quản trị dữ liệu khác nhau thành một hệ hợp nhất. Với một hệ thống tác nghiệp, điều này khó có thể thực hiện được. Ngoài ra, còn có một số nguyên nhân không dễ thấy được, như với cùng một kiểu dữ liệu, nhưng trong các hệ quản trị CSDL khác nhau là khác nhau. Vì vậy, việc truy cập theo phương thức tuần tự là không hiệu quả. Hơn nữa, trước khi dữ liệu được đưa vào kho, nó phải được tích hợp. Sự tích hợp này là một tiến trình xuyên suốt trong thời gian khi dữ liệu được chuyển vào kho và sau đó được đưa vào CSDL ứng dụng. Ngoài ra, với đặc điểm tích hợp tổng thể làm cho DW cung cấp thông tin nhất quán và có ý nghĩa. Ví dụ, dữ liệu từ những chương trình ứng dụng thực hiện trên các CSDL tác nghiệp được tích hợp lại theo một cách mã hóa và độ đo thống nhất như sau:



Hình 2.2: Tích tích hợp dữ liệu

Dữ liệu từ những chương trình ứng dụng thực hiện trên các CSDL tác nghiệp được tích hợp lại theo một cách mã hóa (encoding) và độ đo (measurement) thống nhất

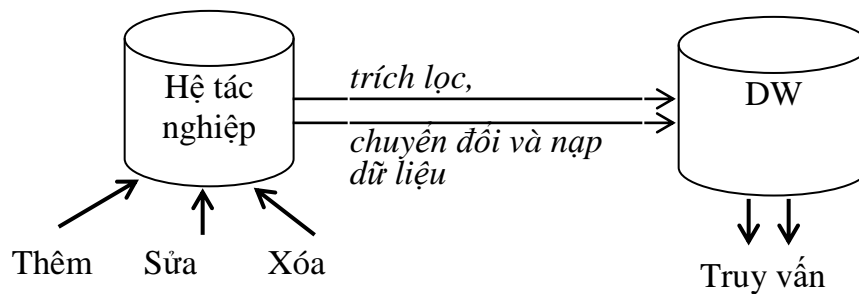
2.1.2.3 Gán nhãn thời gian

Đối với DW, yếu tố thời gian được gắn liền với các dữ liệu. Điều này có nghĩa là, dữ liệu được ghi lại theo thời gian lưu trữ các dữ liệu. Tuy nhiên, hầu hết các truy vấn đều được xử lý ngược lại với việc có một vài thuộc tính về thời gian được lưu trữ trong DW. Chúng ta đã xác định rằng, hầu hết các hệ tác nghiệp đều không chứa các thông tin về lịch sử của nó. Sẽ không thể đoán trước được chuyện gì sẽ xảy ra trong tương lai nếu chúng ta không quan sát được chuyện gì đã xảy ra trong quá khứ. Một DW sẽ giúp chúng ta lấy ra các thông tin cơ bản này bằng cách thêm trường lịch sử thời gian vào dữ liệu được lấy từ hệ tác nghiệp. Như vậy, đối với một doanh nghiệp thì bản thân DW là một “ảnh chụp nhanh” trạng thái của doanh nghiệp tại các thời điểm khác nhau. Nó hỗ trợ việc định ra phương hướng và công tác dự báo cho tương lai. Vì thế, thuộc tính thời gian được ngầm hiểu trong DW như là một thuộc tính bắt buộc. Thời gian đóng vai trò như một phần để đảm bảo tính đơn nhất của các bản ghi và cung cấp đặc trưng về thời gian cho dữ liệu. Dữ liệu trong kho được lưu trữ trong thời gian rất lâu, cỡ vài năm đến vài chục năm và được cập nhật định kỳ.

Bảng 2.1: Dữ liệu trong DW gắn với thời gian

Đặc điểm	Hệ tác nghiệp	DW
Thời gian	60 đến 90 ngày	2 đến 10 năm
Nội dung khóa	Không chứa yếu tố thời gian	Chứa yếu tố thời gian
Dữ liệu	Có thể cập nhật được	Không cập nhật

2.1.2.4 Tính ổn định



Hình 2.3: Tính không thay đổi của dữ liệu trong DW

Dữ liệu trong DW là dữ liệu hướng chủ đề, chỉ đọc và chỉ có thể được kiểm tra. Người sử dụng đầu cuối không được sửa đổi. Nó chỉ cho phép thực hiện hai thao tác cơ bản là nạp dữ liệu vào kho và truy cập vào các vùng trong DW. Nhờ yếu tố này mà tốc độ tổng hợp dữ liệu tăng lên một cách đáng kể. Hơn nữa, kết quả trả lại khi thực hiện công việc phân tích trong những thời gian gần nhau thì hầu như là giống nhau. Dữ liệu được nạp vào một lần từ các CSDL tác nghiệp nhằm mục đích phục vụ truy vấn thông tin. Việc cập nhật sửa đổi dữ liệu chỉ được thực hiện trên CSDL tác nghiệp thông thường; Việc thay đổi dữ liệu đã có chỉ được thực hiện hạn chế khi có lỗi. Nhờ vậy tốc độ tổng hợp dữ liệu tăng lên một cách đáng kể, kết quả trả lại khi thực hiện công việc phân tích của tuần này cũng giống như của tuần trước.

2.1.2.5 Dữ liệu không biến động

Thông tin trong DW được tải vào sau khi dữ liệu trong hệ thống điều hành được cho là quá cũ. Tính không biến động thể hiện ở chỗ, dữ liệu được lưu trữ lâu dài trong kho mặc dù có thêm dữ liệu mới nạp vào nhưng dữ liệu cũ trong kho vẫn không bị

xóa. Điều đó cho phép cung cấp thông tin về một khoảng thời gian dài, cung cấp đủ số liệu cần thiết cho các mô hình nghiệp vụ phân tích, dự báo. Từ đó, có những quyết định hợp lý. Các kỹ thuật liên quan tới phục hồi dữ liệu, đảm bảo toàn vẹn tham chiếu, phát hiện và giải quyết tắc nghẽn... thường không cần thiết.

2.1.2.6 Dữ liệu tổng hợp

Dữ liệu tác nghiệp thuần túy không được lưu trữ trong DW. DW chỉ lưu trữ những dữ liệu tổng hợp được tích lại qua nhiều giai đoạn khác nhau theo các chủ đề như đã nêu ở trên.

2.1.3 Các loại DW thường gặp

Có ba kiểu DW thường gặp:

- *DW doanh nghiệp (Enterprise Data Warehouse - EDW)*: DW loại này cung cấp một tập dữ liệu trung tâm được tổ chức hỗ trợ cho việc ra quyết định của cả công ty (xí nghiệp).
- *DW tác nghiệp (Operational Data Store - ODS)*: DW loại này dùng cho các công ty có phạm vi rộng, nhưng nó không giống với DW doanh nghiệp. Ở đây dữ liệu được làm tươi ngay tức thì và được sử dụng cho những hoạt động thường nhật. Theo định nghĩa của Inmon, ODS khác EDW ở chỗ, nó bị giới hạn về tính lịch sử của dữ liệu và được cập nhật dữ liệu thường xuyên hơn.
- *DW chuyên đề (Data Mart - DM)*: Là CSDL có những đặc điểm giống với DW nhưng quy mô nhỏ hơn và lưu trữ dữ liệu về một lĩnh vực, một chuyên đề, một chuyên ngành cụ thể. DM là DW thứ cấp chứa các dữ liệu tích hợp của DW, hướng tới phần dữ liệu thường được gọi là một vùng chủ đề được tạo ra dành cho một nhóm người sử dụng.

Dữ liệu trong DM cho thông tin về một chủ đề xác định, không phải của toàn bộ hoạt động nghiệp vụ đang diễn ra trong một tổ chức. Thông thường DM là một DW riêng được lưu trữ trên một máy chủ riêng, trong một mạng cục bộ phục vụ cho một nhóm người nhất định

Trong ba loại DW trên, thì *ODS* là sự nâng cấp từ CSDL vốn tồn tại trong doanh nghiệp nếu doanh nghiệp có tổ chức các CSDL phục vụ cho hoạt động của mình. DW và DM cần được phát triển dựa trên việc mô hình hoá dữ liệu theo các chiều, trong đó các *bảng sự kiện* (Fact table) được kết nối với các *bảng chiều* (Dimension table). Mục tiêu của việc xây dựng các DW này là nhằm phục vụ cho các hoạt động nghiệp vụ ở mức cao hơn với hiệu quả cao.

Nói đến DM người ta lại chia ra làm hai loại DM phụ thuộc và DM độc lập. Trong đó, DM phụ thuộc chứa dữ liệu được lấy từ DW và dữ liệu này sẽ được trích lọc tinh chế, tích hợp lại ở mức cao hơn cho một chủ đề nhất định và loại DW này được hình thành sau khi đã có DW. Chính vì vậy DM phụ thuộc vào DW

Còn DM độc lập được xây dựng trước DW và dữ liệu được lấy trực tiếp từ các nguồn tác nghiệp. Phương pháp này đơn giản hơn và chi phí thấp hơn nhưng mỗi DM độc lập có cách tích hợp riêng, do đó dữ liệu từ nhiều nguồn DM khó đồng nhất với nhau. Chính vì vậy DM loại này có trước DW và không phụ thuộc vào DW

2.1.4 Cơ sở dữ liệu tác nghiệp và DW

Thông thường, khi tin học hóa hoạt động quản lý một tổ chức, người ta thường xây dựng CSDL cho các hoạt động quản lý nghiệp vụ thường xuyên của tổ chức và ta có các CSDL tác nghiệp. HTTT tác nghiệp với các dữ liệu tác nghiệp có các đặc điểm sau:

- Trợ giúp công việc hàng ngày, chứa các dữ liệu hiện thời, phản ánh trạng thái của công việc, hoạt động của hệ thống thường đơn giản, giới hạn trong một phạm vi nghiệp vụ đã xác định, và hoạt động chính yếu là cập nhật dữ liệu. Xử lý thông tin hướng đến việc xử lý nhanh các tác vụ đã được định trước
- Người dùng là người làm công việc cụ thể, ở mức độ chi tiết như nhân viên bán hàng, thư ký, thủ kho, ...Thiết kế thường khó hiểu (các bảng dữ liệu phải đạt chuẩn 3 trở lên) đối với người dùng và che dấu đi những quan hệ trực quan của đời thường.

Trong khi đó, DW hướng vào trợ giúp quá trình phân tích và ra quyết định cần có các đặc điểm như trợ giúp quá trình quản lý và điều hành công việc. Nó chứa các dữ liệu mang tính lịch sử, thể hiện cách nhìn ổn định của công việc trong một giai đoạn hay những thời điểm trong quá khứ. Nó được tối ưu hóa cho việc truy vấn, với câu hỏi đã xác định trước hay được thiết lập theo yêu cầu người dùng. Người dùng là những nhà quản lý, phân tích, dự báo hay đánh giá công việc và ra quyết định, các yêu cầu thường đa dạng và có tính nghiệp vụ chuyên ngành. Dữ liệu được thiết kế dễ hiểu và dễ sử dụng đối với người dùng. DW tích hợp và kết nối thông tin từ nhiều nguồn khác nhau trên nhiều loại phương tiện lưu trữ và xử lý thông tin nhằm phục vụ cho những ứng dụng xử lý tác nghiệp trực tuyến. DW có thể lưu giữ thông tin tổng hợp theo một chủ đề nghiệp vụ nào đó nhằm tạo ra các thông tin phục vụ hiệu quả cho việc phân tích của người sử dụng.

Bảng 2.2: So sánh những đặc trưng của hệ thống tác nghiệp và DW

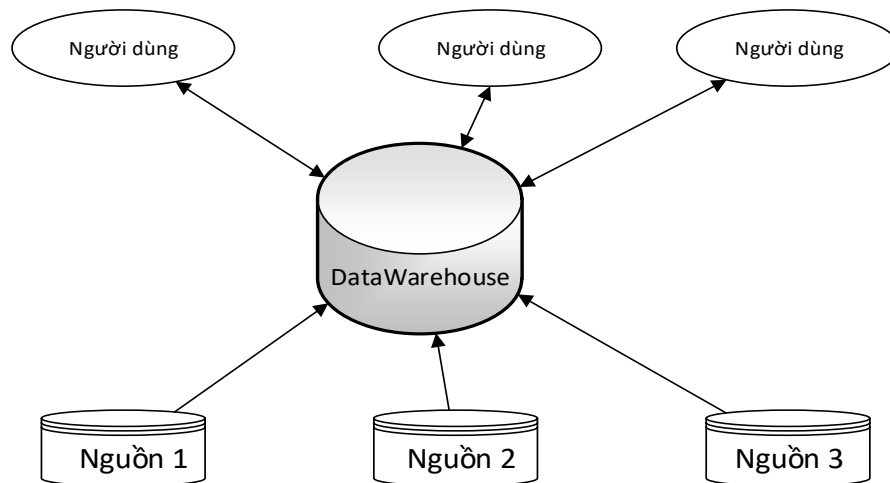
Đặc trưng	Hệ tác nghiệp	DW
Định hướng	Giao dịch tác nghiệp	Phân tích
Kích thước CSDL lớn	Cỡ Gigabyte	Cỡ Terabyte
Nguồn dữ liệu	Nội bộ	Bên trong và bên ngoài
Khả năng cập nhật	Có thể cập nhật được	Không cập nhật được
Nội dung khóa	Không chứa yếu tố thời gian	Chứa yếu tố thời gian
Dữ liệu lưu trữ	Mang tính hiện thời, cập nhật hàng ngày	Dữ liệu có tính lịch sử, duy trì theo khoảng thời gian
Khối lượng công việc	Có thể dự đoán được	Không dự đoán được
Hỗ trợ người dùng	Thao tác báo cáo	Phân tích dự báo, khai phá
Kết nối	Cần nhiều kết nối	Cần ít kết nối
Chức năng	Phục vụ hoạt động	Hỗ trợ ra quyết định

	hang ngày	
Tính ưu tiên	Hiệu suất cao, có tính sẵn sàng	Tính mềm dẻo cao, người dùng tự khai thác
Độ đo	Thông lượng giao dịch	Thông lượng truy vấn
Người dùng	Người dùng cuối	Quản lý, điều hành, phân tích

2.1.5 Các giải pháp kiến trúc DW

2.1.5.1 Kiến trúc cơ bản của DW

- **Lớp nguồn:** Dữ liệu được đưa vào từ rất nhiều nguồn khác nhau. Đây là một trong những đặc điểm chính của DW, dữ liệu nguồn có thể có sẵn trong hệ thống tác nghiệp của doanh nghiệp hoặc lấy từ nguồn bên ngoài.

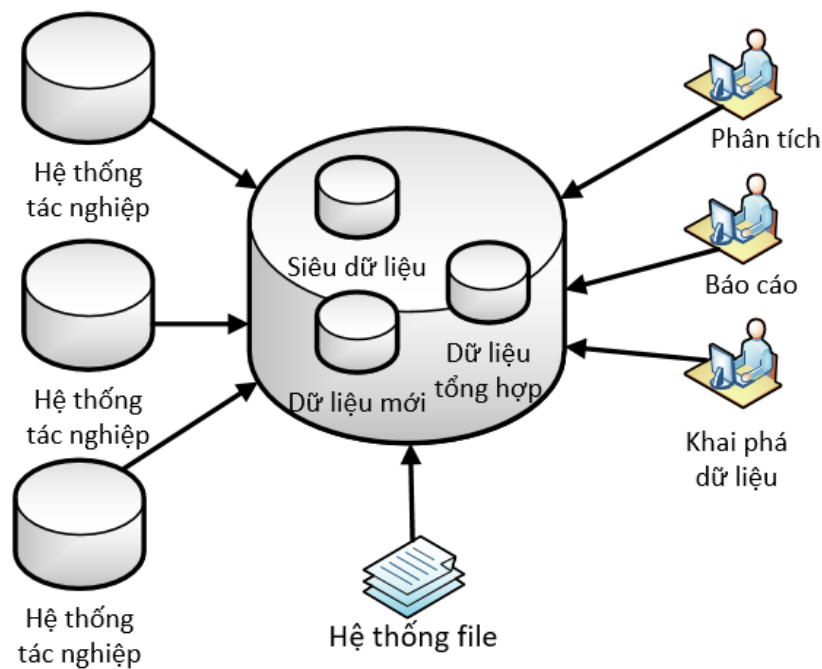


Hình 2.4: Kiến trúc cơ bản của DW

- **Lớp DataWarehouse:** Lớp này chứa đựng dữ liệu đã được tổng hợp cùng siêu dữ liệu mô tả chúng cũng như các tiến trình tổng hợp, phân bổ dữ liệu. Hai thành phần quan trọng nhất của DW là dữ liệu và siêu dữ liệu. Chúng được lưu trữ và mô tả một cách nhất quán về nội dung, ý nghĩa dữ liệu, cũng như hình thức khuôn dạng dữ liệu.

- **Lớp ứng dụng:** Làm nhiệm vụ tương tác với người dùng cuối. Một trong những đặc điểm quan trọng của DW là cung cấp thông tin cho nhiều người sử dụng những yêu cầu không thể dự đoán trước. Vì vậy, lớp người dùng thường có cấu trúc phức tạp, người dùng có thể sử dụng nhiều công nghệ, công cụ khác nhau để khai thác và truy xuất dữ liệu. Vì vậy, giao diện giữa hai lớp này cũng rất đa dạng.

2.1.5.2 Kiến trúc tổng thể đơn giản của DW mức doanh nghiệp

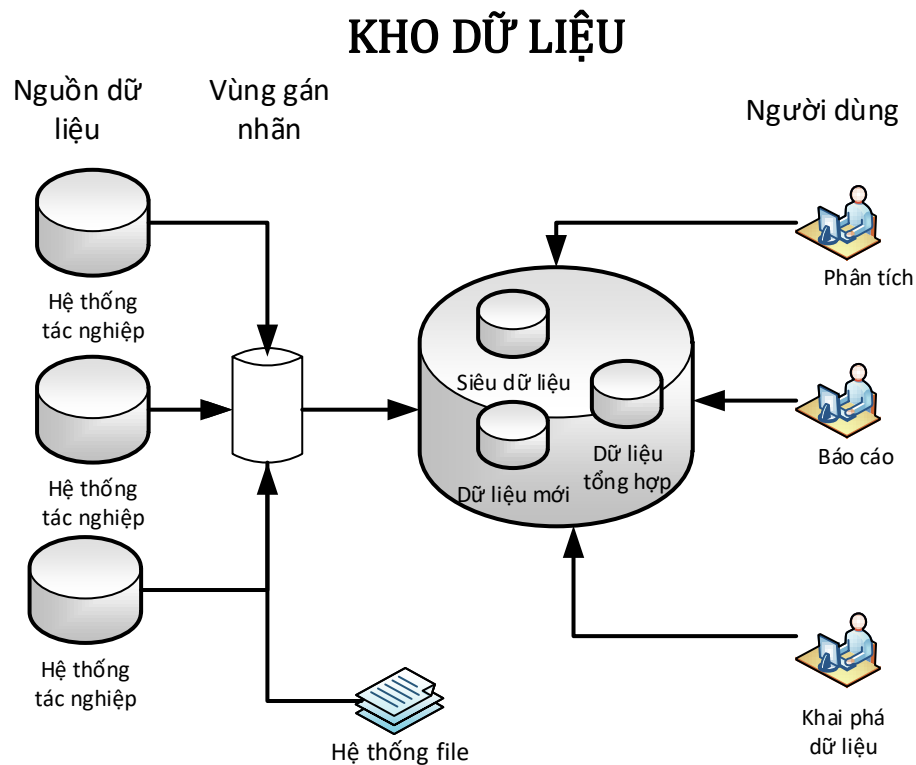


Hình 2.5: Kiến trúc DW

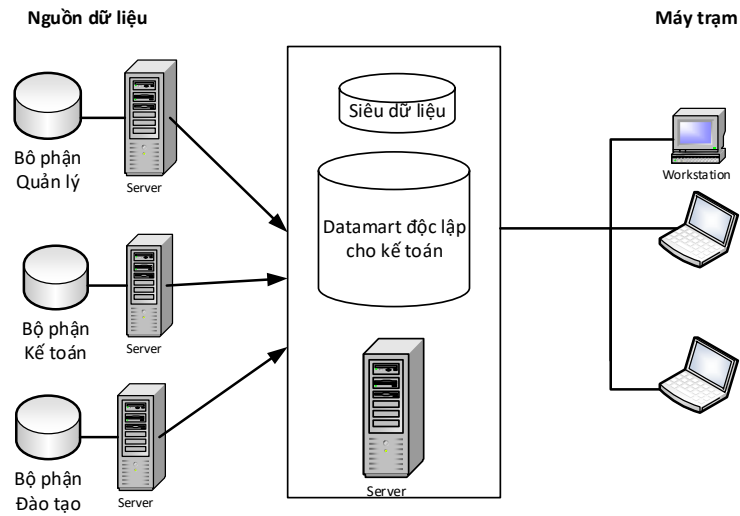
Kiến trúc DW tập trung có một số ưu điểm như: dữ liệu được lưu trữ tại một nơi nên có độ tin cậy và chính xác cao. Mọi chủ đề đều được xem xét và đưa vào sử dụng phục vụ cho nhu cầu toàn doanh nghiệp. Tuy nhiên, hạn chế của kiến trúc này là cấu trúc CSDL phải được điều chỉnh phù hợp cho tất cả các hoạt động, vì vậy thiết kế có thể không tối ưu cho một chức năng nào đó. Tính sẵn sàng bị giảm do rủi ro, sự cố dữ liệu hoặc bị ngưng trệ khi cập nhật dữ liệu. Các thay đổi hoặc cải thiện dữ liệu buộc phải có sự đồng ý của toàn bộ doanh nghiệp.

2.1.5.3 Kiến trúc DW với vùng gắn nhãn

Kiến trúc DW với vùng dữ liệu gắn nhãn được biểu diễn như hình 2.6 Ở đây siêu dữ liệu và các dữ liệu chưa chế biến của hệ thống OLTP hiện tại được đưa vào cùng dữ liệu tổng hợp. Các dữ liệu tổng hợp rất có giá trị trong DW vì nó được tính trước trong quá trình nạp dữ liệu vào DW, nên người dùng dễ dàng truy xuất dữ liệu. Như vậy, để có được các dữ liệu trong DW cần có một quá trình làm sạch và những thao tác chế biến trước khi đẩy dữ liệu vào DW.

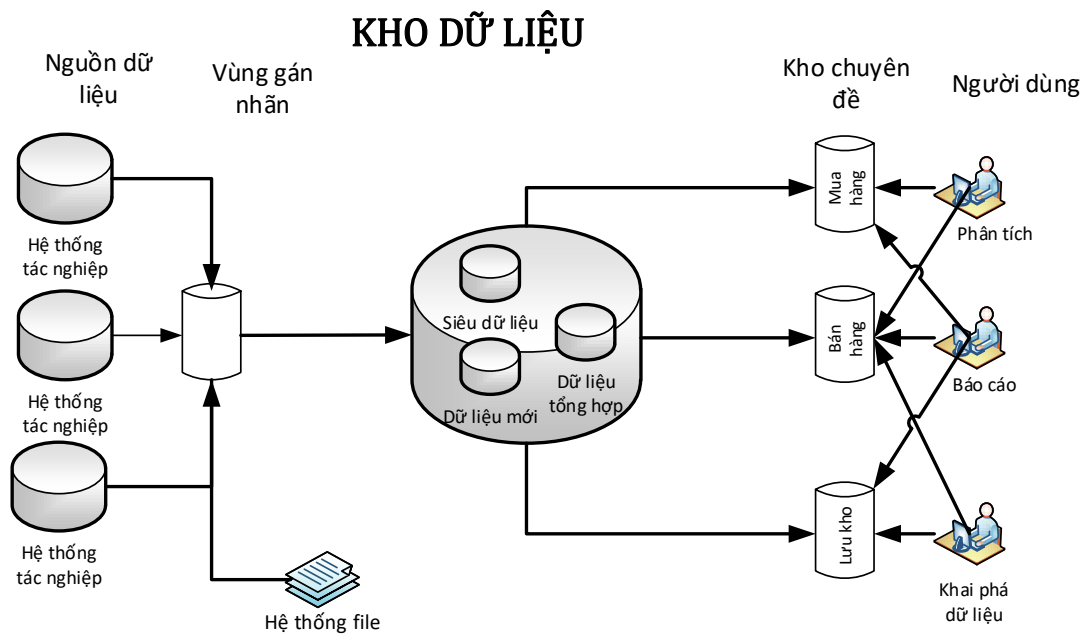


Hình 2.6: Kiến trúc của DW với vùng gắn nhãn (Staging area)



Hình 2.7: Kiến trúc Datamart độc lập

2.1.5.4 Kiến trúc DW chủ đề Datamart phụ thuộc



Hình 2.8: Kiến trúc Datamart phụ thuộc

2.1.6 Các thành phần cơ bản của kiến trúc DW

Kiến trúc tổng quát của DW bao gồm từ nhiều thành phần. Nguồn dữ liệu của nó lấy từ rất nhiều nguồn khác nhau và có cấu trúc dữ liệu khác nhau:

- *Dữ liệu từ hệ thống tác nghiệp*: Các dữ liệu chi tiết nhận được từ các hoạt động nghiệp vụ hàng ngày. Đây là nguồn dữ liệu chính để xây dựng DW, chứa các dữ liệu chi tiết hiện tại của hệ thống tác nghiệp.
- *Dữ liệu từ hệ thống phân tích*: Đây là dữ liệu được tổng hợp từ dữ liệu nguồn đã cũ và tổ chức lại theo nhiều phương pháp khác nhau.
- *Dữ liệu từ bên ngoài*: đây là các dữ liệu từ các nguồn ngoài hệ thống có thể do các tổ chức khác thu thập và tạo ra. Nó được sử dụng cho các yêu cầu phân tích dữ liệu. Ví dụ như các CSDL từ các tệp excel, các tệp thô, hay dạng XML,... Vì thế, trước khi đưa vào DW cần phải chuyển đổi và tích hợp dữ liệu.

Các công cụ trích lọc, chuyển đổi và nạp dữ liệu thực hiện các nhiệm vụ *thu thập, chuẩn hóa, làm sạch, chất lọc, tích hợp dữ liệu và nạp* vào DW.

DW là nơi lưu trữ các loại dữ liệu. Nó gồm ba khối lớn: *các dữ liệu thô, các dữ liệu tổng hợp và siêu dữ liệu*.

Các chức năng chính của siêu dữ liệu bao gồm:

- *Định nghĩa các dữ liệu* lưu trữ trong kho, mô tả các báo cáo và các truy vấn, mô tả các quy định sao chép, cập nhật và nạp dữ liệu, mô tả cách truy nhập dữ liệu.
- *Các DW chuyên đề* lưu các dữ liệu về một lĩnh vực, một chuyên ngành. Các công cụ truy vấn giúp tạo báo cáo, phân tích trực tiếp và khai phá dữ liệu.
- *Hệ quản trị DW*. Nó có chức năng nạp vào, nạp lại dữ liệu, trích lọc dữ liệu, đảm bảo an toàn truy nhập, sao lưu và phục hồi dữ liệu.
- *Hệ thống phân phối thông tin* bao gồm hai lớp. Lớp thứ nhất chuyên tải dữ liệu giữa các thành phần của kho và các bộ phận liên quan. Lớp thứ hai là lớp kết cấu hạ tầng cung cấp các công cụ cho tìm kiếm, quản lý, xác định các phần mềm ứng

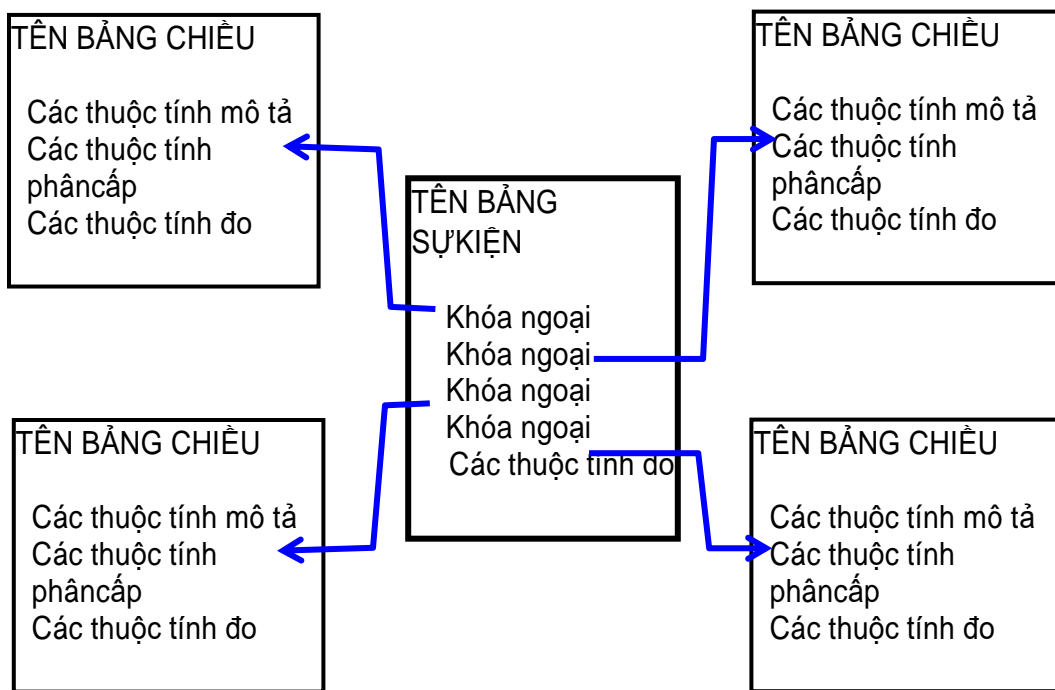
dụng cho người dùng, cho việc sao chép, cập nhật, kết nối, tổng hợp dữ liệu, cung cấp các dịch vụ quản lý các khối dữ liệu, và các xử lý yêu cầu.

2.1.7 Các lược đồ dữ liệu của DW

Khi phát triển một DW, người ta thường sử dụng hai lược đồ để tổ chức lưu trữ dữ liệu, đó là *lược đồ hình sao (Star Flake)* và *lược đồ hình tuyết rơi (Snow Flake)*.

2.1.7.1 Lược đồ hình sao

Lược đồ hình sao được Dr. Ralph Kimball đề xuất dùng trong thiết kế CSDL cho DW, là giản đồ dữ liệu quan hệ, diễn tả dữ liệu nhiều chiều. Nó là dạng lược đồ đơn giản nhất của DW, chứa một hoặc nhiều chiều và một bảng sự kiện. Nó được gọi là lược đồ hình sao vì sơ đồ quan hệ thực thể giữa các bảng chiều và bảng sự kiện giống như hình sao. Các sự kiện được nằm ở trung tâm của mô hình và được bao quanh bởi các chiều liên quan mà trong đó bảng sự kiện được kết nối với nhiều bảng chiều (hình vẽ 2.9). Lợi ích của lược đồ hình sao là sự tách được các dữ liệu thành các lớp nhỏ theo các bảng chiều, tăng tốc độ của việc trình diễn dữ liệu và dễ dàng hiểu được về dữ liệu.

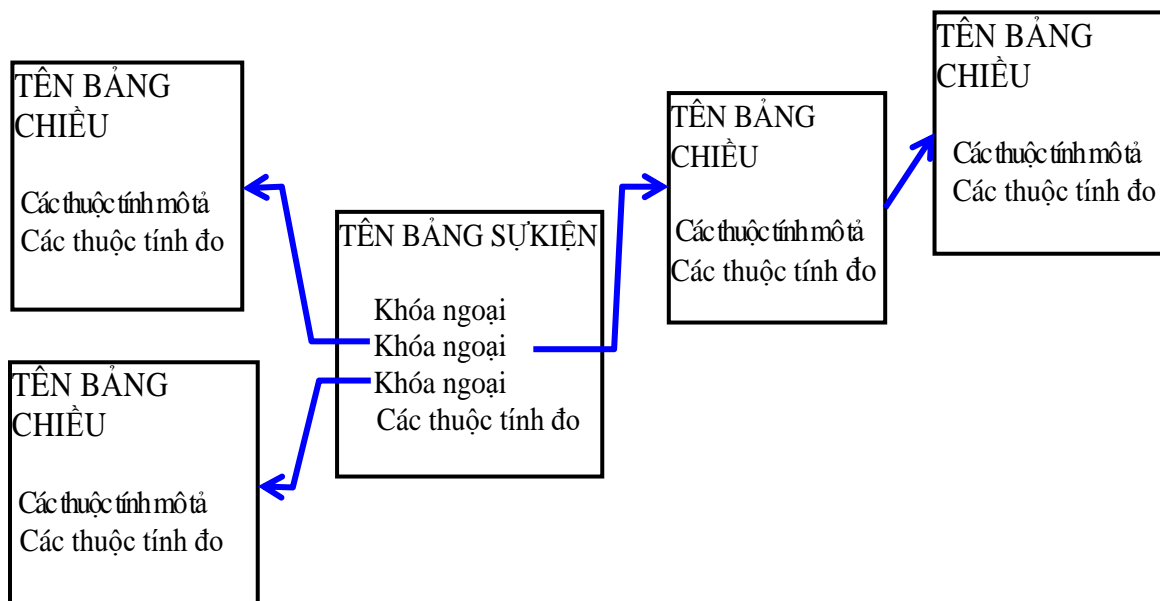


Hình 2.9: Lược đồ hình sao của DW

Trong CSDL thiết kế theo lược đồ hình sao, các truy vấn với những câu hỏi phức tạp liên quan đến nhiều bảng và số liệu tổng trở lên đơn giản hơn, số lượng công việc cần làm ít hơn so với một mô hình quan hệ chuẩn. Giản đồ hình sao rất trực quan, dễ sử dụng thể hiện khung nhìn đa chiều của dữ liệu, cải thiện đáng kể thời gian truy vấn và cho phép thực hiện một số tính năng đa phạm vi. Khóa của bản sự kiện được tạo bởi khóa của các bảng chứa thông tin theo chiều.

2.1.7.2 Lược đồ tuyết rơi

Lược đồ tuyết rơi là một cải tiến mở rộng của lược đồ hình sao, tại mỗi cánh sao không phải một bảng chiều mà có nhiều bảng. Trong đó một số chiều được phân cấp để thể hiện rõ ràng dạng chuẩn của bảng chiều. Điều này dẫn đến nhiều lợi ích trong việc duy trì các bảng chiều. Tuy nhiên, cấu trúc phi chuẩn của các bảng chiều trong lược đồ hình sao có thể thích hợp hơn cho việc duyệt các chiều. Trong lược đồ tuyết rơi (hình 2.10), các bảng chiều được chuẩn hóa. Trong một vài trường hợp, lược đồ này cải thiện việc truy xuất dữ liệu vì các bảng nhỏ hơn được liên kết với nhau nên dễ bảo trì, tăng tính mềm dẻo. Tuy nhiên, lúc này số lượng bảng sử dụng sẽ bị tăng lên, việc thực hiện những câu truy vấn khó khăn hơn nhiều vì có nhiều bảng cần liên kết với nhau.



Hình 2.20: Lược đồ tuyết rơi của DW

2.2 Tích hợp dữ liệu

2.2.1 Khái niệm về tích hợp dữ liệu

Tích hợp dữ liệu là một khái niệm mà có rất nhiều nhà khoa học khác nhau định nghĩa theo quan điểm riêng của mình. Chính vì vậy, thực chất tích hợp dữ liệu là vấn đề kết nối nhiều nguồn dữ liệu khác nhau và cung cấp cho người dùng một khung nhìn chung thống nhất trên tất cả các nguồn dữ liệu đó. Đặc điểm của hệ thống tích hợp dữ liệu tổng quát là bao gồm các nguồn dữ liệu phân tán, không nằm tập trung tại một địa điểm và cũng không đồng nhất về mặt kiến trúc cũng như về ngữ nghĩa của thông tin. Các nguồn dữ liệu này có thể là các CSDL trong các hệ thống khác nhau, cũng có thể là các trang Web ở các địa chỉ khác nhau, cũng có thể là của cá nhân con người với các quan điểm khác nhau về một vấn đề nào đó. Sự không đồng nhất này thể hiện ở các ngôn ngữ biểu diễn và từ vựng biểu diễn dữ liệu. Các nguồn dữ liệu có thể có ngôn ngữ biểu diễn khác nhau, ví dụ CSDL của một nguồn được biểu diễn theo dạng XML, nhưng một nguồn dữ liệu khác lại được biểu diễn theo CSDL quan hệ.

Qua hệ thống tích hợp dữ liệu, chúng ta có thể lấy thông tin từ tất cả các nguồn dữ liệu mà chúng ta mong muốn. Điều đó có nghĩa là, chúng ta thao tác với các nguồn dữ liệu được lấy thông tin từ các nguồn khác nhau với những ràng buộc giữa các nguồn dữ liệu đó. Khi kết nối từ các nguồn dữ liệu lại với nhau để đưa ra câu trả lời thống nhất phù hợp với yêu cầu người sử dụng. Tuy nhiên, cần có giải pháp xử lý dữ liệu trước đó, nhằm đảm bảo sự đồng bộ dữ liệu. Giải pháp hữu hiệu nhất trong trường hợp này là tích hợp dữ liệu. Tùy vào quy mô và cấu trúc CSDL mà chi phí cho dự án tích hợp phù hợp khi sử dụng một trong hai mô hình tích hợp sau:

2.2.1.1 Tích hợp dữ liệu trực tiếp

Với mô hình tích hợp này, dữ liệu trong các CSDL có cùng định dạng, hoặc xây dựng một CSDL mới bao quát các CSDL có cấu trúc khác nhau. Mô hình này thường áp dụng cho các ứng dụng đơn giản, có chi phí thấp.

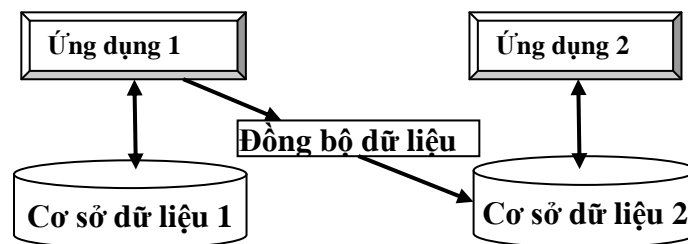
2.2.1.2 Tích hợp dữ liệu qua thành phần trung gian

Đây là mô hình tích hợp dùng cho các CSDL có cấu trúc và định dạng dữ liệu khác nhau. Thành phần trung gian đóng vai trò trung chuyển, có nhiệm vụ nhận dữ liệu từ CSDL nguồn, sau đó biến đổi và chuyển tới CSDL đích. Do đó cấu trúc nguyên thủy của các CSDL ban đầu không bị thay đổi. Mô hình này đang được sử dụng phổ biến.

Một vấn đề quan trọng của tích hợp dữ liệu là tính toàn vẹn dữ liệu, phải đảm bảo không có mâu thuẫn dữ liệu giữa các CSDL khi một ứng dụng được tích hợp làm thay đổi dữ liệu. Trường hợp đơn giản đối với một ứng dụng Máy trạm/Server: khi Máy trạm yêu cầu dữ liệu, Server cung cấp cho nó một bản sao dữ liệu. Máy trạm thực hiện việc đọc/ghi trên bản sao dữ liệu đó, sau khi kết thúc giao dịch, dữ liệu trong bản sao sẽ phải được cập nhật vào bản chính trên Server. Vì vậy dữ liệu luôn đảm bảo toàn vẹn. Trong trường hợp nhiều bản sao dữ liệu cùng tồn tại trong các CSDL, khi dữ liệu ở một bản sao nào đó thay đổi thì dữ liệu trên các bản sao ở các CSDL khác cũng phải thay đổi.

Chẳng hạn: Thông tin người lao động của một công ty được lưu trong nhiều CSDL. CSDL nhân sự chứa thông tin cá nhân, CSDL tài chính chứa thông tin tiền lương. Cả hai CSDL này đều chứa thông tin về người lao động. Do đó, khi có một nhân viên thêm vào công ty thì dữ liệu phải được cập nhật trên cả hai CSDL.

Để đảm bảo tính toàn vẹn dữ liệu, một trong những phương pháp được ứng dụng là **đồng bộ dữ liệu**. Đồng bộ dữ liệu là sự làm phù hợp các nội dung trong hai hay nhiều CSDL nhằm đảm bảo tính nhất quán của dữ liệu. Với phương thức này, dữ liệu được thay đổi ở một CSDL và sau đó được sao chép tới các CSDL liên quan khác. Tùy thuộc vào các yêu cầu sử dụng, dữ liệu có thể được đồng bộ ngay lập tức khi có sự kiện xảy ra hoặc định kỳ theo một khoảng thời gian nhất định.



Hình 2.11: Minh họa đồng bộ dữ liệu (EIA)

Trong hình 2.11, CSDL 2 chứa một bản sao của CSDL 1. Khi ứng dụng 1 thay đổi dữ liệu trong bản chính ở CSDL 1 thì dữ liệu của bản sao trong CSDL 2 được cập nhật lại, do đó không có mâu thuẫn dữ liệu.

2.2.3 Các phương pháp tích hợp dữ liệu

Các phương pháp tích hợp dữ liệu dựa trên cơ sở của kỹ thuật tích hợp.

2.2.3.1 Tích hợp dữ liệu dựa trên ước lượng không chắc chắn

Tích hợp dữ liệu dựa trên ước lượng không chắc chắn là phương pháp tính toán độ phù hợp của các dữ liệu thu thập được với yêu cầu của người dùng hoặc ứng dụng cụ thể, sau đó chọn ra dữ liệu có độ phù hợp cao nhất.

Ví dụ, trong các ứng dụng tìm kiếm truy xuất thông tin dữ liệu trên Web, như Yahoo, Google, Alta Vista... độ phù hợp của một thông tin dữ liệu được tính qua hai tham số là độ chính xác và khả năng chấp nhận dữ liệu.

2.2.3.2 Tích hợp dữ liệu dựa trên các ràng buộc dữ liệu

Các phương pháp thuộc về dạng này được áp dụng cho hệ thống bao gồm các nguồn biểu diễn dữ liệu dưới dạng các hệ CSDL và cấu trúc, ràng buộc trong các hệ CSDL này là có thể biết được. Mục đích của các hệ thống này là trả lời các truy vấn của người dùng về thông tin dữ liệu trong nhiều nguồn khác nhau mà không cần truy nhập trực tiếp vào các nguồn thông tin này. Tiêu biểu cho phương pháp tích hợp dữ liệu thuộc loại này là phương pháp dùng cho *HTTT dựa trên Internet* (Internet Base Information System - IBIS). Phương pháp tích hợp dữ liệu này đưa ra lược đồ bộ ba (G, S, M) được xây dựng từ các nguồn thông tin dữ liệu cần tích hợp:

- *Lược đồ toàn cục (Global schema – G)*: giống như lược đồ quan hệ trong lý thuyết về CSDL, mô tả các ràng buộc nhất quán, các ràng buộc khóa và các yêu cầu về tính độc lập giữa các nguồn thông tin dữ liệu.
- *Lược đồ dữ liệu (Source schema – S)*: Mô tả cấu trúc của tập các nguồn dữ liệu cần tích hợp trong hệ thống.

- Các ánh xạ (*Map schema –M*): bao gồm các ánh xạ được thiết lập giữa lược đồ toàn cục và các lược đồ nguồn dữ liệu.

Trên cơ sở xem xét các ràng buộc được định nghĩa trong G và cấu trúc biểu diễn trong S. Người thiết kế hệ thống sẽ xác định các ánh xạ tương ứng giữa các thực thể dữ liệu trong các nguồn dữ liệu.

2.2.3.3 Tích hợp dữ liệu tự động dựa trên ontology

Phương pháp tích hợp dữ liệu dựa trên ontology sử dụng một cấu trúc phân lớp các khái niệm, thuật ngữ và các quan hệ giữa các khái niệm để biểu diễn các nguồn dữ liệu cần tích hợp. Cả nội dung và ngữ nghĩa thông tin dữ liệu được biểu diễn thông qua tương tác giữa các thành phần dựa trên ontology của các dữ liệu nguồn được tích hợp. Chính vì vậy, ontology giải quyết được vấn đề không đồng nhất về ngữ nghĩa thông tin dữ liệu. Quá trình tích hợp dữ liệu này sẽ diễn ra một cách tự động thông qua việc xác định các ánh xạ tương đương hoặc không tương đương giữa các khái niệm trong ontology khác nhau.

2.2.4 Môi trường tích hợp đồng nhất

2.2.4.1 Công nghệ Java RMI

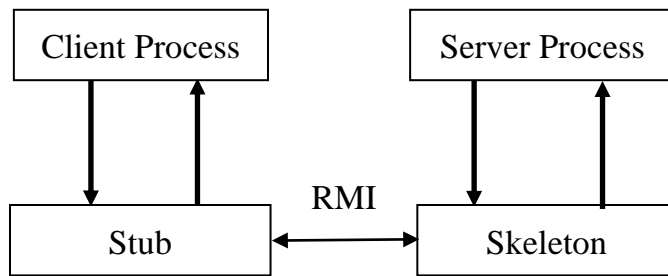
Lập trình đối tượng phân tán là một trong những vấn đề nóng bỏng của công nghệ phân tán phần mềm ngày nay. Java là ngôn ngữ đi tiên phong với *Phương pháp triệu gọi từ xa (Remote Method Invocation - RMI)* - một kỹ thuật cài đặt các đối tượng phân tán vô cùng hiệu quả và linh hoạt. Thông thường, các chương trình Java được viết dưới dạng thủ tục hoặc hàm, và việc các hàm gọi lẫn nhau, truyền tham số hay kết quả cho nhau chỉ xảy ra ở máy cục bộ. Kỹ thuật RMI - triệu gọi phương thức từ xa, là cách thức giao tiếp giữa các đối tượng trong Java có mã lệnh cài đặt nằm trên các máy khác nhau có thể triệu gọi lẫn nhau. Công nghệ Java RMI [10] sử dụng để tích hợp các ứng dụng được xây dựng cùng một ngôn ngữ lập trình. Vì vậy, các ứng dụng có thể kết nối và trao đổi dữ liệu trực tiếp với nhau không cần thành phần chuyển đổi ngôn ngữ trung gian. Công nghệ cho môi trường tích hợp đồng nhất chủ yếu là JRMI (*Java RMI*) của IBM và *Mô hình đối tượng thành phần phân tán (Distributed Component*

Object – DCOM) của Microsoft [11]. Tuy nhiên, ngôn ngữ Java chiếm ưu thế cho lập trình phân tán RMI hỗ trợ các đối tượng từ xa thông qua giao thức JRMI. Với công nghệ này, phía yêu cầu thông tin và phía cung cấp thông tin đều được xây dựng trên ngôn ngữ Java. Mỗi đối tượng RMI Server định nghĩa một giao diện để có thể truy cập đối tượng Server từ bên ngoài *máy ảo Java (Java Virtual Machine - JVM)* hiện hành và trên JVM của máy tính khác. Giao diện là một tập các cách thức để biểu diễn các dịch vụ và triệu gọi đối tượng phía server. Nó được xây dựng và biên dịch bằng chương trình RMI. Một RMIRegistry trên máy chủ chứa thông tin về các đối tượng server và cung cấp tên dịch vụ cho RMI. Khi máy khách yêu cầu một đối tượng tham chiếu tới đối tượng máy chủ thì máy khách sẽ tìm kiếm trong RMIRegistry. Khi một RMI máy trạm yêu cầu một dịch vụ từ RMI server, nó thực hiện các bước sau:

- Kết nối với JVM có chứa các đối tượng ở xa.
- Truyền các thông số cho JVM ở xa.
- Chờ các kết quả của các phương thức yêu cầu.
- Nhận về giá trị cần gọi.

Bằng việc sử dụng các đối tượng, các dữ liệu và mã lệnh có thể được trao đổi trực tiếp giữa máy chủ và máy khách. Vì thế, một đối tượng có thể chạy được trên cả máy chủ lẫn máy khách.

Để giải quyết một số vấn đề trong việc truyền thông giữa Máy trạm/Server. RMI không cho phép gọi trực tiếp mà thông qua lớp trung gian. Lớp này tồn tại ở cả hai phía Máy trạm và Server. Lớp ở máy Máy trạm gọi là **Stub**, lớp ở máy Server gọi là **Skeleton**. Máy trạm gọi sẽ gọi đến đối tượng đại diện Stub. Stub chuyển lời gọi đến đối tượng đại diện trên máy chủ là Skeleton thông qua kiến trúc RMI. Sau đó, Skeleton gọi đối tượng thật trên máy chủ và kết quả trả về cho máy khách. Quá trình trả kết quả về cho máy khách được thực hiện tuần tự ngược lại với quá trình gọi.



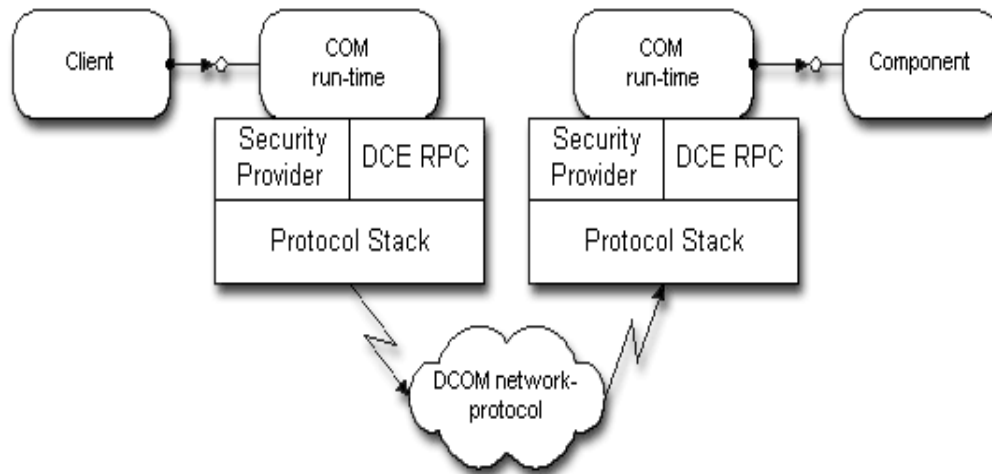
Hình 2.32: Giao tiếp giữa Máy trạm và Server qua RMI

2.2.4.2 Công nghệ DCOM

Ngoài công nghệ Java RMI còn có một công nghệ sử dụng cho môi trường đồng nhất là *Mô hình đối tượng thành phần phân tán* (Distributed Component Object Model – DCOM). DCOM là một chuẩn do tập đoàn Microsoft phát triển, nó là mở rộng của chuẩn *Mô hình đối tượng thành phần* (Component Object Model - COM). DCOM cung cấp các đối tượng từ xa thông qua một giao thức được gọi là ORPC. ORPC được xây dựng trên DCE/RPC và tương tác với các dịch vụ thực thi trong COM. Công nghệ DCOM có một số đặc điểm sau [12]

- DCOM là một mô hình phân tán dễ triển khai với chi phí thấp, hỗ trợ kết nối chặt giữa các ứng dụng và hệ điều hành. COM định nghĩa cách thức các thành phần và máy trạm liên lạc trao đổi với nhau trên cùng một máy. DCOM mở rộng COM bằng cách sử dụng các giao thức mạng chuẩn khi cần trao đổi dữ liệu với máy khác trên mạng. DCOM hỗ trợ kết nối giữa các đối tượng và những kết nối này có thể được thay đổi lúc đang chạy. Các đối tượng DCOM được triển khai bên trong các gói nhị phân chứa các mã lệnh quản lý chu kỳ sống của đối tượng và việc đăng ký đối tượng.
- DCOM mang đến nhiều ưu điểm như tính ổn định, không phụ thuộc vị trí địa lý, quản lý kết nối hiệu quả và dễ dàng mở rộng. Nó là một lựa chọn tốt cho các doanh nghiệp có các ứng dụng chạy trên nền Windows. Tuy nhiên, đây cũng

chính là nhược điểm lớn của các công nghệ của Microsoft vì chúng bị giới hạn trên nền tảng Windows.



Nguồn: [4]

Hình 2.43: Kiến trúc công nghệ DCOM [11]

Các kiến trúc trên đều hướng đến việc xây dựng một hệ thống hướng dịch vụ, tuy nhiên chúng vẫn còn gặp phải một số vấn đề như sau:

- Chúng có kết nối chặt, nghĩa là *kiến trúc triển khai cài đặt bên phía nhà cung cấp dịch vụ và phía sử dụng dịch vụ phải giống nhau*. Điều này đồng nghĩa với khó khăn, mỗi khi có sự thay đổi từ một trong hai phía, bởi vì mỗi thay đổi cần được đánh giá, lên kế hoạch và sửa chữa ở cả hai phía.
- Những chuẩn trên đa phần là chuẩn đóng, chúng hầu như không thể kết hợp, hoạt động với chuẩn khác. Ví dụ như bắt đối tượng Java trao đổi dữ liệu trực tiếp với một đối tượng DCOM là không thể.
- Các đối tượng của các mô hình trên là fine-grained, nghĩa là lượng thông tin trong mỗi lần thực hiện giao dịch là ít, và được thực hiện nhiều lần dẫn đến chiếm dụng băng thông sử dụng và tăng thời lượng đáp trả dữ liệu. .

2.2.5 Môi trường tích hợp không đồng nhất

Một thực tế của các doanh nghiệp hiện nay là, không phải tất cả các ứng dụng đều được xây dựng trên cùng một ngôn ngữ, mà trên nhiều ngôn ngữ lập trình khác nhau như Java, Visual Basic, Visual C++, ... Do đó, khi tích hợp các ứng dụng không cùng ngôn ngữ thì các công nghệ tích hợp trong môi trường đồng nhất như Java RMI hay DCOM không thực hiện được. Điều này dẫn tới yêu cầu phải có công nghệ có thể hiểu được các ngôn ngữ của các ứng dụng tích hợp. Để giải quyết vấn đề này một số công nghệ cho môi trường tích hợp không đồng nhất ra đời.

2.2.5.1 Công nghệ CORBA

Kiến trúc môi giới yêu cầu đối tượng chung (Common Object Request Broker Architecture – CORBA) là công nghệ tích hợp các ứng dụng phân tán được chuẩn hóa bởi tổ chức OMG, sử dụng được trên các nền hệ điều hành và tương thích được với nhiều ngôn ngữ lập trình [13]. Trong CORBA việc triệu gọi các đối tượng từ xa thông qua giao thức IIOP, tất cả đều dựa vào thành phần *Môi giới yêu cầu đối tượng (Object Request Broker - ORB)* hoạt động như một kênh đối tượng trung tâm. Mỗi đối tượng CORBA có thể tương tác trong suốt với đối tượng CORBA khác trong cùng một máy hoặc ở máy khác trong mạng. Mỗi đối tượng có một giao diện và giao diện này mô tả một tập các phương thức. Máy trạm có thể sử dụng các phương thức gọi trên các đối tượng tham chiếu giống như các đối tượng CORBA server đang nằm trong không gian địa chỉ của CORBA máy trạm. ORB đảm nhận tìm một sự thực thi của đối tượng CORBA so sánh, nhận gửi đi các yêu cầu, nhận lại kết quả cho máy trạm. Đối tượng CORBA tương tác với ORB thông qua giao diện của ORB.

Trong CORBA gồm có thành phần trung gian, đối tượng đại diện cho đối tượng cần gọi phía máy khách là Stub, đối tượng đại diện cần triệu gọi phía máy chủ là Skeleton. Cả Stub và Skeleton được mô tả bằng *Ngôn ngữ định nghĩa mô tả giao diện (Interface Discription Definition Language – IDDL)*. Sau khi xây dựng một giao diện chung, các ứng dụng có thể giao tiếp với nhau thông qua Stub và Skeleton, thành phần môi giới trung gian ORB và môi trường mạng với giao thức IIOP.

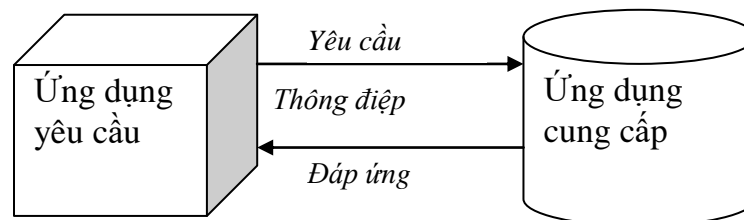
Một điểm giống với công nghệ Java RMI của CORBA đó là cách triệu gọi phương thức của đối tượng trên ứng dụng khác thông qua các đối tượng đại diện Stub và Skeleton. Tuy nhiên, các đối tượng Stub và Skeleton của CORBA sinh ra từ IDDL, không giống như của Java RMI được sinh ra từ định nghĩa đối tượng Java.

2.2.5.2 Công nghệ Web Services (Dịch vụ Web)

Web Services cũng là một công nghệ tích hợp ứng dụng sử dụng trong môi trường tích hợp không đồng nhất và các ứng dụng luôn có sự thay đổi.

Một dịch vụ Web là một ứng dụng có khả năng xử lý một công việc độc lập, được mô đun hóa, tự mô tả và dễ dàng kết nối với các ứng dụng khác để thực hiện xử lý những công việc phức tạp hơn thông qua môi trường mạng. Giống như các công nghệ DCOM, RMI, CORBA, Web Services cũng là một công nghệ dùng cho tích hợp các ứng dụng phân tán. Với Web Services, người sử dụng có thể biết được cách thức triệu gọi các dịch vụ thông qua giao diện mô tả dịch vụ. Các ứng dụng trao đổi dữ liệu với nhau thông qua các thông điệp được xây dựng theo một cấu trúc chung và sử dụng giao thức truyền thông HTTP. Với cách thức giao dịch bằng thông điệp, các ứng dụng gửi và nhận dữ liệu bằng các thông điệp [14].

Web Services được xây dựng từ các chuẩn SOAP, WSDL, UDDI trên nền tảng XML. Người sử dụng có thể tìm thấy dịch vụ ở mục UDDI, biết được giao diện của dịch vụ bằng mô tả WSDL và sử dụng dịch vụ thông qua giao thức SOAP.



Hình 2.54: Giao tiếp giữa bên sử dụng dịch vụ và cung cấp dịch vụ

Ưu điểm của Web Services là các chuẩn được xây dựng trên nền tảng ngôn ngữ XML. Đây là sự thuận lợi cho việc mở rộng ứng dụng mà không ảnh hưởng đến mối liên kết nội tại với các ứng dụng khác. Một ưu điểm nữa của Web Services là không

cần quan tâm các ứng dụng được viết bằng ngôn ngữ gì, thực thi trên hệ điều hành nào, cấu hình phần cứng ra sao và định vị ở đâu. Qua đó cho thấy, các ứng dụng được tích hợp có liên kết lỏng lẻo và tính trong suốt khi sử dụng công nghệ Web Services.

Tính liên kết lỏng lẻo: Trong môi trường liên kết chặt, mỗi ứng dụng kết nối với ứng dụng khác qua một tổ hợp các giao diện đóng, giao thức mạng đóng, nên việc tích hợp các ứng dụng cần nhiều thời gian và chi phí cao để mở và đóng liên kết giữa các máy tính trong hệ thống. Với tính liên kết lỏng của Web Services, mỗi một phần của phần mềm như là một dịch vụ Web, dễ dàng di chuyển đến các vị trí khác nhau và có thể liên kết lại một cách đơn giản.

Tính trong suốt: Tính trong suốt của Web Services cho phép các ứng dụng được tích hợp ở bất kỳ chỗ nào trong mạng mà không ảnh hưởng đến hoạt động tích hợp. Bởi vì, mỗi dịch vụ có một địa chỉ URL riêng trên mạng Internet, chỉ cần triệu gọi nó thông qua địa chỉ URL.

Trên đây giới thiệu một số công nghệ tích hợp ứng dụng gồm Java RMI, DCOM, CORBA và Web Service. Chúng có điểm chung là: được sử dụng để xây dựng các ứng dụng phân tán trên môi trường mạng. Tuy nhiên, mỗi công nghệ đáp ứng được các yêu cầu khác nhau với mức độ và điều kiện khác nhau cho một giải pháp tích hợp.

Môi trường tích hợp: Công nghệ JavaRMI và DCOM hiệu quả trong môi trường tích hợp đồng nhất, nhưng trong môi trường tích hợp không đồng nhất thì chúng không sử dụng được. Để khắc phục vấn đề này, công nghệ CORBA ra đời để thay thế cho Java RMI và DCOM. Nhưng hạn chế của công nghệ này là các ứng dụng tích hợp phụ thuộc vào giao diện IDL. Vì vậy Web Services là công nghệ được lựa chọn cho cả giải pháp tích hợp đồng nhất và không đồng nhất.

Các lợi thế của Web Services

- **Hỗ trợ ngôn ngữ lập trình:** Các công nghệ khác chỉ được hỗ trợ bởi một ngôn ngữ lập trình, thì Web Services được hỗ trợ bởi nhiều ngôn ngữ lập trình.
- **Chuẩn mở:** Web Services là một công nghệ mới được xây dựng trên các chuẩn WSDL, UDDI, SOAP mà nền tảng là ngôn ngữ XML. Với đặc điểm này, nó có nhiều ưu điểm hơn các công nghệ khác khi tích hợp ứng dụng trên nền web.

- **Đơn giản:** Web Services dễ thiết kế, phát triển và bảo trì. Một khung phát triển, sử dụng là có sẵn và dễ dàng đưa ra qui trình nghiệp vụ mới từ các ứng dụng.
- **Mềm dẻo:** Các giải pháp tích hợp ứng dụng thường theo kiểu điểm nối điểm. Khi có sự thay đổi ở một điểm nào đó sẽ dẫn tới sự thay đổi của điểm kia làm cho chúng không linh hoạt. Với Web Services thì hoàn toàn mềm dẻo, nó được xây dựng từ kết nối lỏng lẻo giữa ứng dụng công bố dịch vụ và ứng dụng sử dụng dịch vụ đó.
- **Chi phí:** Web Services có chi phí rẻ hơn, nhanh hơn các các công nghệ tích hợp ứng dụng khác.
- **Hiệu quả:** Web Services cho phép các ứng dụng phân chia thành các phần nhỏ hơn, làm cho việc tích hợp các ứng dụng dễ dàng, chiếm ít không gian mạng. Đây là một công nghệ tích hợp ứng dụng mang lại nhiều hiệu quả hơn so với các giải pháp truyền thống khác.

Với các ưu điểm trên, công nghệ Web Services thực sự nổi trội hơn cả và là một sự lựa chọn hàng đầu về giải pháp công nghệ cho sự tích hợp các ứng dụng.

2.3 Web service

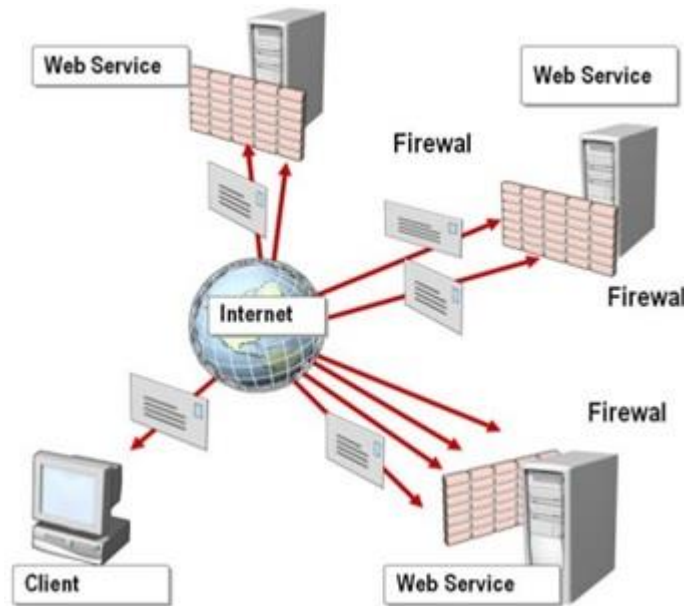
2.3.1 Định nghĩa Web service

Theo định nghĩa của W3C (World Wide Web Consortium). *Webservice là một hệ thống phần mềm được thiết kế để hỗ trợ cho khả năng tương tác giữa các ứng dụng trên các máy tính khác nhau trong môi trường Internet thông qua các giao diện (Interface) chung và sự gắn kết được mô tả bằng XML.*

Web service là tài nguyên phần mềm có thể xác định bằng địa chỉ URL để thực hiện các chức năng và đưa thông tin ra cho người dùng.

Web service được tạo ra bằng cách lấy các chức năng và đóng gói chúng sao cho các ứng dụng khác dễ dàng tìm thấy chúng và truy cập tới sử dụng, đồng thời vẫn có thể yêu cầu thông tin từ các dịch vụ khác. Web service bao gồm các mô đun độc lập để thực hiện yêu cầu nghiệp vụ của doanh nghiệp được thực thi trên Server.

Ứng dụng cơ bản của Web service là tích hợp các hệ thống và là một trong những hoạt động chính khi phát triển hệ thống. Trong hệ thống này, các ứng dụng cần được tích hợp với CSDL và các ứng dụng khác, người sử dụng sẽ giao tiếp với CSDL để tiến hành phân tích và lấy dữ liệu.



Hình 2.15: Kiến trúc mô hình hệ thống Web service

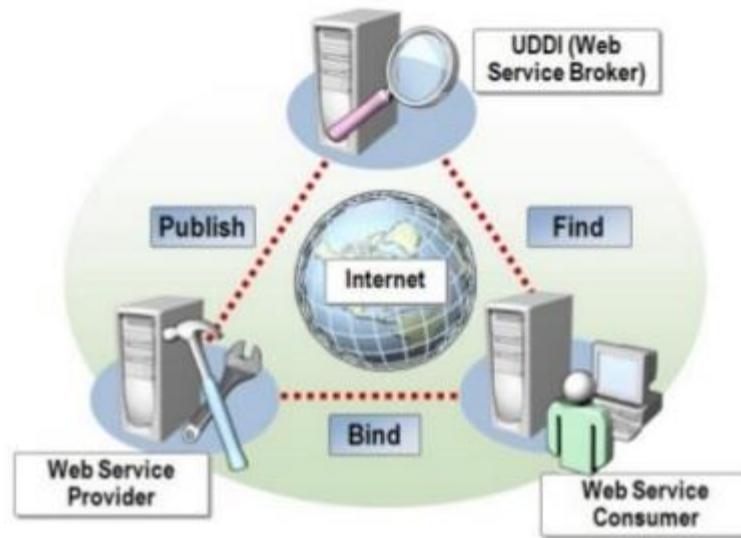
2.3.2 Đặc điểm của Web service

Web service cho phép các ứng dụng khác nhau từ các nguồn khác nhau có thể giao tiếp với các ứng dụng khác mà không đòi hỏi nhiều thời gian lập trình, do tất cả các quá trình giao tiếp cùng tuân theo định dạng XML. Cho nên Web service không bị phụ thuộc vào bất kỳ hệ điều hành hay ngôn ngữ lập trình nào. Web service cho phép máy trạm và server có thể tương tác được với nhau trên các nền tảng khác nhau mà không cần bất cứ thay đổi hay yêu cầu đặc biệt nào. Ví dụ, chương trình viết bằng ngôn ngữ Java cũng có thể trao đổi dữ liệu với các chương trình viết bằng Perl. Các ứng dụng chạy trên nền Windows cũng có thể trao đổi dữ liệu với các ứng dụng chạy trên nền Linux. Công nghệ Web service không yêu cầu phải sử dụng trình duyệt và ngôn ngữ HTML.

Phần lớn kỹ thuật của Web Service được xây dựng trên mã nguồn mở và được phát triển từ các chuẩn đã được công nhận. Nó tích hợp các ứng dụng trên nền Web lại với nhau bằng cách sử dụng các công nghệ XML, SOAP, WSDL, và UDDI trên nền tảng giao thức internet với mục tiêu tích hợp ứng dụng và truyền thông điệp. XML được sử dụng để đánh dấu dữ liệu, SOAP được dùng để truyền dữ liệu, WSDL được sử dụng để mô tả các dịch vụ có sẵn và UDDI được sử dụng để liệt kê những dịch vụ nào hiện tại đang có sẵn để sử dụng. Web service cho phép các tổ chức có thể trao đổi dữ liệu với nhau mà không cần phải có kiến thức hiểu biết về HTTP đứng sau Firewall. Web service có thể gồm nhiều mô đun được công bố trên internet. Là sự kết hợp của việc phát triển theo hướng từng thành phần với những lĩnh vực cụ thể và cơ sở hạ tầng Web, đưa ra lợi ích cho cả doanh nghiệp, khách hàng, những nhà cung cấp dịch vụ khác và các cá nhân thông qua mạng internet.

Web service khi được triển khai sẽ được hoạt động theo mô hình máy trạm-server. Nó có thể được triển khai bởi một phần mềm ứng dụng phía Server như PHP, JSP, ASP.NET, ... Không giống như mô hình máy trạm-server truyền thống, chẳng hạn như hệ thống Web server-web page, Web service không cung cấp cho người dùng một giao diện đồ họa nào. Web service đơn thuần chỉ là việc chia sẻ các dữ liệu logic và xử lý các dữ liệu đó thông qua một giao diện chương trình ứng dụng được cài đặt xuyên suốt trên mạng máy tính.

Tính tương thích (Interoperability) là một lợi thế vô cùng mạnh mẽ của Web service. Thông thường, các công nghệ Java và công nghệ của Microsoft rất khó có thể tích hợp được với nhau. Nhưng với Web service thì các ứng dụng và máy khách sử dụng 2 công nghệ trên hoàn toàn có khả năng tương tác với nhau thông qua Web service.



Hình 2.66: Kiến trúc của dịch vụ Web service

2.3.3 Nền tảng của Web service

Dịch vụ Web cũng có thể xem như các khối cơ bản được xây dựng để di chuyển trong hệ thống máy tính phân tán trên internet. Các chuẩn mở và việc tập trung vào giao tiếp và làm việc cộng tác giữa con người và các ứng dụng đã tạo nên một môi trường Web service làm nền tảng cho việc tích hợp ứng dụng. Các ứng dụng được sử dụng các Web service đủ loại từ nhiều nguồn khác nhau để làm việc cùng với nhau bất kể chúng ở đâu và được triển khai như thế nào. Có thể có các định nghĩa khác nhau về Web service khi các công ty xây dựng chúng, nhưng hầu hết các định nghĩa đều có chung các điểm sau:

- Web service đưa ra chức năng hữu dụng cho người sử dụng Web thông qua một giao thức chuẩn Web. Trong hầu hết các trường hợp, giao thức được sử dụng đó là SOAP.
- Web service đưa ra cách mô tả các giao diện của chúng một cách đủ chi tiết nhằm giúp cho người sử dụng xây dựng một ứng dụng máy trạm để giao tiếp được với chúng. Mô tả này thường được cung cấp ở dạng một tài liệu XML gọi là một tài liệu về ngôn ngữ mô tả dịch vụ web (*Web service Description Language - WSDL*)

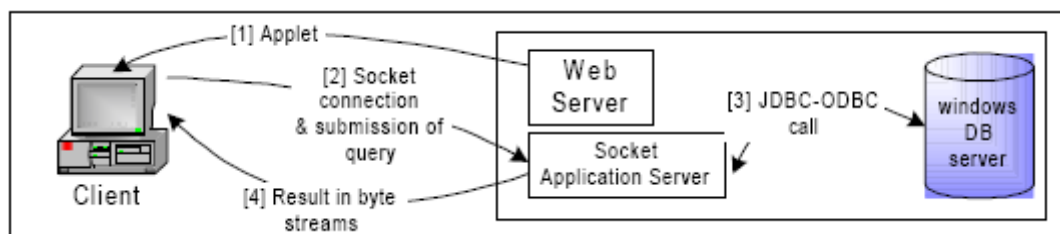
- Web service được đăng ký sao cho khách hàng tiềm năng là người sử dụng có thể tìm thấy chúng một cách dễ dàng. Điều này được thực hiện với UDDI (Universal Discovery Description and Integration). Web service như là một dịch vụ phần mềm được trình bày trên Web thông qua giao thức SOAP, được mô tả bằng một tệp WSDL và được đăng ký trong UDDI. Các dịch vụ - Web service là nguồn thông tin mà ta có thể dễ dàng kết hợp vào các ứng dụng. Dễ dàng nhận ra toàn bộ lớp ứng dụng có thể được xây dựng để phân tích và tích hợp thông tin quan tâm và trình bày nó theo nhiều cách khác nhau.

Việc trình bày các ứng dụng đang có như các dịch vụ Web service cho phép người sử dụng xây dựng các ứng dụng có các tính năng mạnh hơn thông qua việc sử dụng Web service như những khối được xây dựng sẵn. Ví dụ, người sử dụng có thể phát triển một ứng dụng mua bán để tự động lấy các thông tin về giá cả từ nhiều nhà cung cấp khác nhau, cho phép người dùng chọn nhà cung cấp, chuyển đơn hàng và sau đó theo dõi việc chuyển hàng cho tới khi nhận được hàng. Ứng dụng của nhà cung cấp, khi trình bày các dịch vụ của họ trên Web, có thể quay ra sử dụng các dịch vụ Web service để kiểm tra tín dụng của khách hàng, lấy tiền từ tài khoản của khách hàng và thiết lập việc chuyển hàng với một công ty vận tải.

2.3.4 Các công nghệ của Web service

Hiện nay có rất nhiều giải pháp cho bài toán như trên, tuy nhiên lựa chọn phương pháp nào thì tùy thuộc vào hạ tầng cơ sở viễn thông, quy định về mặt pháp lý và định hướng của cơ quan đó. Để sử dụng Web service xây dựng các ứng dụng, chúng ta cần xem xét Web service dựa trên nền tảng công nghệ nào?

2.3.4.1 Phương pháp Java Socket



Hình 2.77: Mô hình truy nhập CSDL web bằng Java Socket

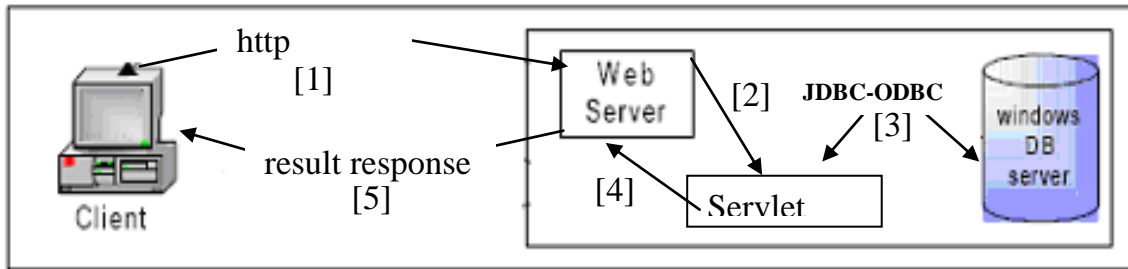
Ngôn ngữ lập trình Java hỗ trợ hai dạng chương trình ứng dụng chính là ứng dụng độc lập (Java application) và ứng dụng nhúng (Java applet). Các Java applet có thể được máy trạm tải xuống từ một máy ở xa thông qua trình duyệt Web và thực thi tại máy đó. Do tính bảo mật của ngôn ngữ Java nên máy ảo Java sẽ không cho phép các Java applet được quyền truy nhập tài nguyên cục bộ như CSDL Web đặt trên máy server. Vì vậy, để bảo đảm được hai yếu tố của phương pháp Java socket là truy nhập CSDL từ xa thông qua trình duyệt Web và nhận được kết quả trả về cần có thêm thành phần trung gian đứng giữa máy trạm và CSDL Web. Thành phần trung gian trong phương pháp Java socket là một chương trình ứng dụng độc lập.

Hoạt động của mô hình truy nhập CSDL Web bằng phương pháp Java socket thực hiện qua những bước sau :

- Máy trạm truy nhập vào trình chủ Web trên máy server thông qua trình duyệt Web. Một ứng dụng Java applet có chức năng truy nhập CSDL từ trên máy chủ Web được tải về máy trạm cùng với trang chủ Web.
- Ứng dụng Java applet truy nhập CSDL được người dùng khởi động tại máy trạm và kết nối tới thành phần trung gian trên máy chủ Web. Khi kết nối thành công thì máy trạm gửi yêu cầu truy nhập dữ liệu cho thành phần trung gian.
- Kết nối được chấp nhận thì chương trình trung gian sẽ truy nhập vào CSDL đặt trên máy chủ Web lấy dữ liệu theo yêu cầu của máy trạm.
- Thành phần trung gian trả dữ liệu kết quả về cho ứng dụng Java applet ở phía máy trạm, sau đó applet chuyển dữ liệu kết quả cho trình duyệt Web để nó hiển thị dữ liệu kết quả lên cho người dùng

2.3.4.2 Phương pháp servlet

Phương pháp servlet thường được dùng để tạo ra các trang Web động. Mọi thao tác xử lý theo yêu cầu của máy trạm được thực hiện tại server như viết mã lệnh để tạo ra trang Web, truy nhập CSDL. Điều này rất có ý nghĩa đối với các máy trạm có năng lực xử lý hạn chế. Một ưu điểm nổi bật của phương pháp Servlet là giúp giảm tải mạng, do không cần phải duy trì một kết nối mạng thường xuyên giữa máy trạm và server trong quá trình máy trạm truy nhập CSDL.



Hình 2.88: Mô hình truy nhập CSDL bằng Servlet

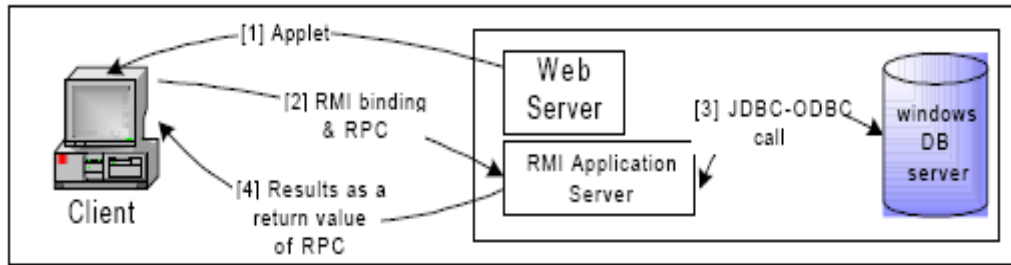
Thành phần trung gian trong phương pháp này là một servlet. Nó là một chương trình Java được thực hiện như là một tiến trình con trong môi trường của một trình máy chủ Web có hỗ trợ Java. Trình chủ Web có nhiệm vụ định tuyến cho các yêu cầu từ phía máy trạm đến được servlet có nhiệm vụ thực thi yêu cầu đó. Ngoài ra trình chủ Web còn đảm nhiệm các công việc: nạp, khởi động, chạy và kết thúc các servlet.

Hoạt động của mô hình truy nhập CSDL bằng Servlet thực hiện theo các bước như sau:

- Máy trạm truy nhập trình chủ Web trên máy server bằng trình duyệt Web.
- Trình chủ Web gọi servlet tương ứng thực thi yêu cầu từ phía máy trạm.
- Chương trình servlet truy nhập vào CSDL cục bộ lấy dữ liệu theo yêu cầu của máy trạm.
- Chương trình servlet chuyển dữ liệu kết quả cho trình chủ Web
- Trình chủ Web trả dữ liệu kết quả cho máy trạm. Trình duyệt Web tại máy trạm sẽ hiển thị dữ liệu đã yêu cầu lên cho người dùng.

2.3.4.2 Phương pháp RMI

RMI là một giao diện ứng dụng cho phép thực thi các lời gọi phương thức từ xa giữa các đối tượng Java phân tán.



Hình 2.19: Mô hình truy nhập CSDL web bằng RMI

Thành phần trung gian trong phương pháp RMI bao gồm hai đối tượng:

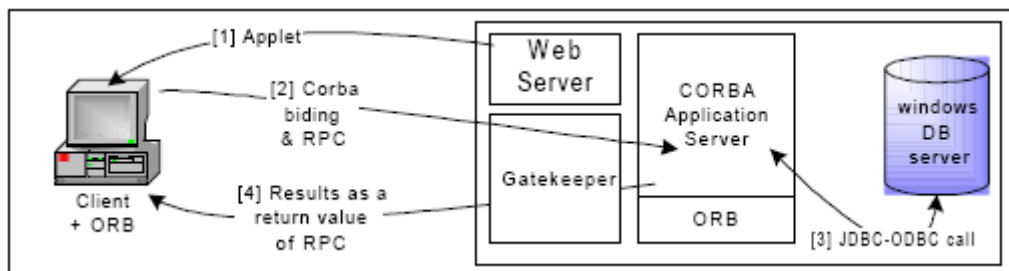
- Chương trình ứng dụng độc lập Java, làm nhiệm vụ cài đặt và thực hiện các phương thức được máy trạm triệu gọi từ xa.
- Ứng dụng chạy nền `rmiregistry.exe` đi kèm trong bộ JDK từ phiên bản 1.3 trở lên làm hai nhiệm vụ: Khởi động ứng dụng server và đăng ký tên duy nhất cho ứng dụng server với máy ảo Java chạy trên trình chủ Web.

Hoạt động của mô hình truy nhập CSDL Web bằng phương pháp RMI thực hiện qua những bước sau:

- Máy trạm truy nhập vào trình chủ Web trên máy server bằng trình duyệt Web. Java applet có nhiệm vụ truy nhập CSDL Web bằng lời gọi phương thức từ xa được tải từ server về máy trạm cùng với trang chủ Web.
- Applet truy nhập CSDL Web được người dùng kích hoạt sẽ thực hiện tìm kiếm đối tượng từ xa trên máy chủ Web dựa vào trình đăng ký tên dịch vụ duy nhất `rmiregistry.exe` chạy trên máy chủ Web. Nếu tìm thấy, applet thực hiện lời gọi phương thức từ xa để lấy dữ liệu.
- Ứng dụng server đáp ứng yêu cầu được trình đăng ký tên dịch vụ duy nhất chạy trên máy chủ Web khởi động và thực hiện truy nhập CSDL để lấy dữ liệu theo yêu cầu của máy trạm.
- Ứng dụng server trả dữ liệu kết quả về cho máy trạm bằng phương thức được gọi từ xa của nó.

2.3.4.3 Phương pháp Java CORBA

CORBA là một chuẩn đối tượng phân tán, định nghĩa các mối quan hệ máy trạm/server giữa các đối tượng trong một ngôn ngữ giao diện chung [13]. Chương trình RMI chỉ có cài đặt thể thực thi bằng ngôn ngữ lập trình Java nhưng chương trình CORBA có thể được cài đặt và thực thi bằng một ngôn ngữ lập trình bất kỳ [16]



Hình 2.90: Mô hình truy nhập cơ sở dữ web bằng Java CORBA

Đối tượng ứng dụng máy trạm CORBA muốn gọi đúng được đối tượng ứng dụng server CORBA cần có một đối tượng thứ ba có thể cung cấp phương tiện giao tiếp giữa các ứng dụng, dịch vụ và các tiện ích mạng gọi là ORB. ORB được quan niệm như là một loại bus mềm hay đường trục sống, cung cấp các giao diện chung giữa nhiều loại đối tượng khác nhau để có thể giao tiếp được với nhau theo mô hình bình đẳng.

Đối tượng máy trạm gửi yêu cầu đến ORB, nhiệm vụ của ORB là tìm đối tượng server hay tìm đối tượng có thể biết các server, sau đó thiết lập quá trình truyền thông giữa máy trạm và server này. Đối tượng server gửi đáp ứng cho ORB, nó định dạng lại và chuyển tiếp đáp ứng về cho nơi phát ra yêu cầu. ORB phải được nạp trên cả máy trạm và server, do tính bảo mật nên CORBA không cho phép một sử dụng Java truy nhập trực tiếp vào một đối tượng server CORBA ở xa, mà phải đi qua tường lửa IIOP. IIOP là một phần của CORBA. Nó cung cấp phương tiện để các đối tượng CORBA có thể tương tác với mạng TCP/IP, bao gồm cả mạng Internet. IIOP làm việc kết hợp hoặc thay thế cho HTTP, một giao thức cơ bản tương tác giữa trình duyệt Web và server.

Ngoại trừ giao thức IIOP, thành phần trung gian trong mô hình CORBA giống như RMI: bao gồm hai đối tượng là đối tượng phục vụ ứng dụng (CORBA Application Server) và một đối tượng cài đặt. Đối tượng cài đặt ở đây là một ứng dụng chạy ngầm (`tnameserv.exe`) được tích hợp trong bộ JDK từ phiên bản 1.3 trở lên. Đối tượng cài đặt được sử dụng để nạp ORB và đăng ký CORBA Application Server dưới một tên dịch vụ duy nhất với ORB.

Hoạt động của mô hình truy nhập CSDL Web bằng phương pháp Java CORBA thực hiện theo các bước sau :

- Máy trạm truy nhập vào web server trên máy server qua trình duyệt Web, applet có chức năng truy nhập CSDL Web được tải về máy máy trạm từ máy server.
- Applet được khởi động từ phía máy trạm. Sau khi nạp xong ORB nó kết nối với ứng dụng CORBA server trên máy server thông qua Gatekeeper bằng cách gọi một phương thức đặc biệt và chuyển tên dịch vụ duy nhất của ứng dụng server đi giống như tham số của phương thức.
- Ứng dụng CORBA Server thực hiện truy nhập CSDL Web cục bộ, lấy dữ liệu theo yêu cầu của phía máy trạm.
- Ứng dụng CORBA Server gửi dữ liệu kết quả về cho phía máy trạm giống như giá trị trả về của lời gọi phương thức.

2.4 Mô hình tích hợp cho ứng dụng của nhà trường

Mục tiêu của nhà trường là xây dựng nhà trường thành trường cao đẳng nghề chất lượng cao khu vực Asian và quốc tế theo quyết định 761 của Thủ tướng Chính phủ theo đề án đến năm 2020 nhằm đáp ứng nguồn nhân lực chất lượng cao cho xã hội. Để thực hiện được mục tiêu này, và có thể cạnh tranh được với các trường nghề khác trong khu vực, trường Cao đẳng nghề số 3 BQP cần phải có HTTT trợ giúp trường trong công tác phân tích, thống kê, dự báo, đánh giá chất lượng đào tạo, v. v... thường xuyên nhằm giúp các nhà quản lý, các nhà lãnh đạo trong việc điều hành và ra các quyết định có tính chiến lược.

2.4.1 Xác định nguồn dữ liệu

Do đặc thù trường Cao đẳng nghề số 3-BQP đang trên đà xây dựng phát triển, chính vì vậy việc ứng dụng CNTT vào việc quản lý đào tạo và giảng dạy còn đơn lẻ, riêng rẽ từng bộ phận. Mỗi bộ phận hiện đang sử dụng các phần mềm riêng rẽ để phục vụ cho nhiệm vụ quản lý của mình. Ví dụ, bộ phận quản lý nhân sự hiện đang sử dụng phần mềm Phần mềm Bảo hiểm quân đội ASMS để quản lý đội ngũ Cán bộ, giáo viên và nhân viên phục vụ trong toàn trường. Bộ phận quản lý trang thiết bị và phương tiện ô tô xe máy phục vụ việc dạy nghề hiện đang sử dụng phần mềm quản lý nhân viên KT và trang bị. Bộ phận quản lý đào tạo sử dụng phần mềm quản lý đào tạo. Bộ phận tài chính sử dụng phần mềm quản lý kế toán quân đội. Các phần phần mềm này được xây dựng độc lập và cài đặt trên từng máy tính đơn lẻ, trong môi trường hệ điều hành và nền tảng kiến trúc khác nhau. Để kế thừa các nguồn dữ liệu có sẵn của từng bộ phận, đơn vị, mô hình tích hợp dữ liệu được xây dựng trên cơ sở sử dụng các nguồn dữ liệu từ các đơn vị thành viên, tạo thành DW phục vụ cho mục đích phân tích, báo cáo và hỗ trợ ra quyết định của nhà trường.

Căn cứ vào các hệ thống nguồn cũng như nhu cầu quản lý, việc xây dựng DW ở trường Cao đẳng nghề số 3- BQP nhằm tạo ra các báo cáo thông kê, phân tích được thực hiện bằng cách xây dựng DW. DW có các DM là các chủ đề chủ đề tương ứng sau:

- Dữ liệu nhân sự về Cán bộ giáo viên, nhân viên
- Dữ liệu về cơ sở vật chất trang thiết bị dạy nghề
- Dữ liệu về quá trình đào tạo

Trên cơ sở khảo sát các nguồn dữ liệu tại trường Cao đẳng nghề số 3 BQP, bằng cách xây dựng các DW theo hướng chủ đề như nhắc đến ở trên, sau đó tích hợp lại với nhau tạo thành DW dùng chung của Nhà trường. Tuy nhiên, do hiện trạng hệ thống dữ liệu phân tán hiện tại, nên có nhiều khó khăn trong việc tích hợp, chuyển đổi dữ liệu cũng như công tác quản trị dữ liệu, vì mỗi DBMS có cấu trúc dữ liệu, ràng buộc, cách thức truy vấn, bảo mật khác nhau. Chính vì, trong luận văn này tôi sử dụng công nghệ Web service làm công cụ chính để tích hợp dữ liệu từ các nguồn khác nhau vào một DW dùng chung. Vì Web service cho phép các ứng dụng khác nhau từ các nguồn khác

nhau có thể giao tiếp với mà không đòi hỏi nhiều thời gian mã hóa. Vì tất cả các quá trình giao tiếp giữa các ứng dụng thông qua Web service đều sử dụng định dạng XML, cho nên chúng không bị phụ thuộc vào bất kỳ hệ điều hành hay ngôn ngữ lập trình nào mà ứng dụng sử dụng.

2.4.2 Các yêu cầu về báo cáo thông kê

Hệ thống được xây dựng cần hỗ trợ tổ chức và khai thác dữ liệu hướng đối tượng, từ đó cung cấp khả năng khai thác dữ liệu theo nhiều mức khác nhau: tổng thể hoặc chi tiết; biến động theo thời gian; so sánh giữa các đối tượng báo cáo và theo nhiều chiều khác nhau như theo chiều thời gian, theo các đơn vị báo cáo, các thông tin này có thể là:

- Thông tin về số lượng, chất lượng, trình độ chuyên môn của đội ngũ Cán bộ giáo viên, nhân viên trong nhà trường.
- Thông tin về cơ sở vật chất trang thiết bị phục vụ công tác dạy nghề
- Thông tin về học sinh, sinh viên, đang theo học tại nhà trường như số lượng, chất lượng, trình độ, đối tượng và kết quả đào tạo.
- Thông tin về nhu cầu học nghề của các đối tượng chính sách xã hội cũng như nhu cầu về việc làm sau đào tạo.
- Đánh giá chất lượng đào tạo bằng một loạt các báo cáo tiêu chí, yếu tố khác nhau, biểu đồ theo dõi kết quả tốt nghiệp. v. v.

Ngoài ra, với công nghệ DW, có thể tích hợp dữ liệu của nhà trường với dữ liệu của các sở ngành, bộ liên quan. Đặc biệt là kết nối với CSDL quốc gia về dạy nghề của Tổng cục dạy nghề Bộ lao động thương binh và xã hội. Từ đó Ban giám hiệu Nhà trường có thể dự báo được nhu cầu về nguồn nhân lực của từng địa phương, từng ngành nghề phát triển của xã hội và có những chiến lược hợp lý hơn về công tác dạy nghề và việc làm cho bộ đội xuất ngũ và các đối tượng chính sách xã hội.

Cung cấp giao diện đơn giản, hỗ trợ khai thác hiệu quả báo cáo đã có trong CSDL với yêu cầu kiến thức CNTT người sử dụng ở mức thông thường.

- Cung cấp khả năng truy xuất thông tin đột xuất; các dạng biểu diễn báo cáo theo bảng biểu hoặc đồ thị.
- Khai thác trực tiếp dữ liệu trên mạng cục bộ hoặc truy nhập từ xa qua mạng WAN của nhà trường dựa trên giao diện WEB.
- Có khả năng quản lý người sử dụng, an toàn, bảo mật dữ liệu báo cáo theo thẩm quyền khai thác, có thể đưa ra các câu truy vấn bất thường
- Cung cấp mô hình phân tích, dự báo, hỗ trợ giao diện với các phần mềm văn phòng của Microsoft .

Chương 3: XÂY DỰNG KHO DỮ LIỆU BÁO CÁO THỐNG KÊ TRƯỜNG CĐ NGHỀ SỐ 3 - BQP

Chúng ta cần xây dựng kho dữ liệu DW phục vụ cho việc lập các báo cáo thống kê của trường Cao đẳng nghề số 3- BQP theo các tiêu chí yêu cầu của các bộ phận chức năng khác nhau. Quá trình xây dựng phải trải qua các giai đoạn sau:

- Xây dựng CSDL theo chuẩn DW phục vụ:
 - Báo cáo thống kê dựa trên các tiêu chí
 - Báo cáo thống kê mang tính tổng hợp, chi tiết
- Xây dựng quy trình tích hợp cho DW
- Yêu cầu dữ liệu lấy từ hệ thống CSDL của các phần mềm quản lý nhân sự, quản lý đào tạo, quản lý trang bị vật tư đã có
- Xây dựng giải pháp để khai thác báo cáo trên cơ sở đã có DW. Trong đó DW đã được tổ chức đa chiều nhất quán, chính xác và tích lũy theo thời gian

3.1 Thiết kế hệ thống kho dữ liệu của trường Cao đẳng nghề số 3

3.1.1 Xây dựng cơ sở dữ liệu theo chuẩn DW

3.1.1.1 Khảo sát hệ thống nguồn

Sau khi tiến hành khảo sát và tìm hiểu tại trường Cao đẳng nghề số 3 BQP và các cơ quan đơn vị, khoa giáo viên trực thuộc đã xác định được một số nguồn thông tin như sau:

Bảng 3.1: Các hệ thống phần mềm đang sử dụng

STT	Tên hệ thống, đơn vị sử dụng	Mô tả chức năng
1	Website của nhà trường : http://www.cdns3.edu.vn	Website giới thiệu tin tức, sự kiện và hoạt động của nhà trường
2	File báo cáo hàng tháng, quý, năm	Báo cáo số liệu hoạt động của nhà trường

3	Bộ phận quản lý nhân sự thuộc phòng đào tạo sử dụng phần mềm BHXH và BHYT quân đội	Quản lý hồ sơ cán bộ, sĩ quan, QNCN, CNVQP, Lao động hợp đồng
4	Bộ phận kế hoạch thuộc phòng đào tạo sử dụng phần mềm quản lý đào tạo	Quản lý các hoạt động đào tạo, thời khóa biểu, tiến độ đào tạo....
5	Bộ phận quản lý điểm và văn bằng chứng chỉ bằng phần mềm quản lý đào tạo thuộc phòng đào tạo	Quản lý kết quả học tập và văn bằng chứng chỉ
6	Bộ phận tuyển sinh phòng đào tạo	Quản lý hồ sơ học sinh, sinh viên
7	Bộ phận quản lý vật chất trang thiết bị dạy nghề thuộc phòng kỹ thuật sử dụng phần mềm Phần mềm quản lý nhân viên Kỹ thuật và trang bị	Quản lý vật chất trang thiết bị phục vụ công tác dạy nghề và phương tiện ô tô xe máy
8	Bộ phận tài chính thuộc ban tài chính sử dụng phần mềm quản lý kế toán	Quản lý tài chính, thu học phí, cân đối thu chi và tiền lương
9	Các khoa	Quản lý chuyên môn giảng dạy

a. Cơ sở dữ liệu quản lý nhân sự

Căn cứ vào hệ thống quản lý nhân sự, chúng ta xác định các bảng sự kiện và bảng chiều. Mỗi đơn vị sẽ có nhiều nhân viên với các thông tin khác nhau. Khi chọn một nhân viên thì chúng ta có thể xem được các thông tin liên quan đến nhân viên đó như lý lịch các nhân, quá trình công tác,...

Sự kiện nhập thông tin nhân viên được thực hiện khi nhân viên được nhận vào làm việc tại trường. Người quản lý nhân sự sẽ tham gia vào các sự kiện quản lý thông tin, đối tượng quản lý như quản lý cán bộ sĩ quan, quân nhân chuyên nghiệp, công nhân viên quốc phòng, lao động hợp đồng, quá trình công tác, khen thưởng kỷ luật, quan hệ gia đình, học hàm, học vị, chức vụ, quân hàm... dưới đây là các bảng mô tả chi tiết trong kho dữ liệu chủ đề nhân sự và để đảm bảo yếu tố thời gian trên mỗi bảng thông tin nhân viên ta thêm vào các bảng trường thời gian như ngày cập nhật (CreateDate), ngày sửa chữa (ModifiedDate) gồm các bảng sau (xem phụ lục 1, phần a.):

- Bảng Hồ sơ nhân viên: qlns_hoso_nhanvien

- Bảng thâm niên: qlns_thamnien
- Bảng chức vụ: qlns_chucvu
- Bảng cấp bậc: qlns_capbac
- Bảng phòng ban: qlns_phongban
- Bảng trình độ chuyên môn: qlns_trinhdo_chuyenmon
- Bảng trình độ ngoại ngữ: qlns_trinhdo_ngoaingu
- Bảng trình độ tin học: qlns_trinhdo_tinhoc
- Bảng lương: qlns_bangluong
- Bảng hợp đồng lao động: qlns_hopdong_laodong
- Bảng khen thưởng: qlns_khenthuong
- Bảng kỷ luật: qlns_kyluat

b. Cơ sở dữ liệu quản lý đào tạo

Căn cứ vào hệ thống quản lý đào tạo chúng ta xác định các bảng sự kiện và bảng chiều. Ở đây chúng ta có những đối tượng cần quản lý như hồ sơ sinh viên và kết quả học tập .v.v. Khi chọn một đối tượng thì ta có thể xem được các thông tin liên quan đến đối tượng đó, như lý lịch, hồ sơ sinh viên, quá trình học tập và rèn luyện v.v

Bộ phận kế hoạch giáo vụ, bộ phận tuyển sinh, bộ phận quản lý điểm kết quả học tập và văn bằng chứng chỉ sẽ tham gia vào các sự kiện quản lý đào tạo. Dưới đây là các bảng mô tả chi tiết trong kho dữ liệu chủ đề đào tạo và để đảm bảo yếu tố thời gian trên mỗi bảng thông tin nhân viên ta thêm vào các bảng trường thời gian như ngày cập nhật (CreateDate), ngày sửa chữa (ModifiedDate) gồm các bảng sau (xem phụ lục 1, phần b.):

- Bảng Hồ sơ sinh viên: qltd_hoso_sinhvien
- Bảng hệ đào tạo: qltd_he
- Bảng ngành đào tạo: qltd_nganh
- Bảng lớp học: qltd_lop

- Bảng điểm học kỳ: qltd_diem_hocky
- Bảng học kỳ: qltd_hocky
- Bảng Năm học: qltd_namhoc
- Bảng giáo viên chủ nhiệm: qltd_giaovien_chunhiem
- Bảng kết quả thi tốt nghiệp: qltd_ketqua_thitotnghiep
- Bảng môn thi tốt nghiệp: qltd_monthi_totnghiep

c. Cơ sở dữ liệu quản lý vật tư trang bị kỹ thuật, phương tiện

Căn cứ vào hệ thống quản lý vật tư trang bị thiết bị, phương tiện phục vụ công tác dạy nghề, chúng ta xác định các bảng sự kiện và bảng chiều. ở đây chúng ta có những đối tượng cần quản lý tên trang thiết bị phương tiện, chủng loại, đối tượng sử dụng,... Khi chọn một chủng loại vật tư trang bị chúng ta có thể xem được các thông tin liên quan đến chủng loại vật tư trang bị đó như tên vật tư trang bị, chủng loại, nhóm,...

Bộ phận nhân viên chuyên môn kỹ thuật sẽ tham gia vào các sự kiện quản lý trang thiết bị vật tư phương tiện kỹ thuật. Dưới đây là các bảng mô tả chi tiết trong kho dữ liệu chủ đề trang bị vật tư phương tiện phục vụ công tác dạy nghề và để đảm bảo yếu tố thời gian trên mỗi bảng thông tin nhân viên ta thêm vào các bảng trường thời gian như ngày cập nhật (CreateDate), ngày sửa chữa (ModifiedDate) gồm các bảng sau (xem phụ lục 1, phần c.):

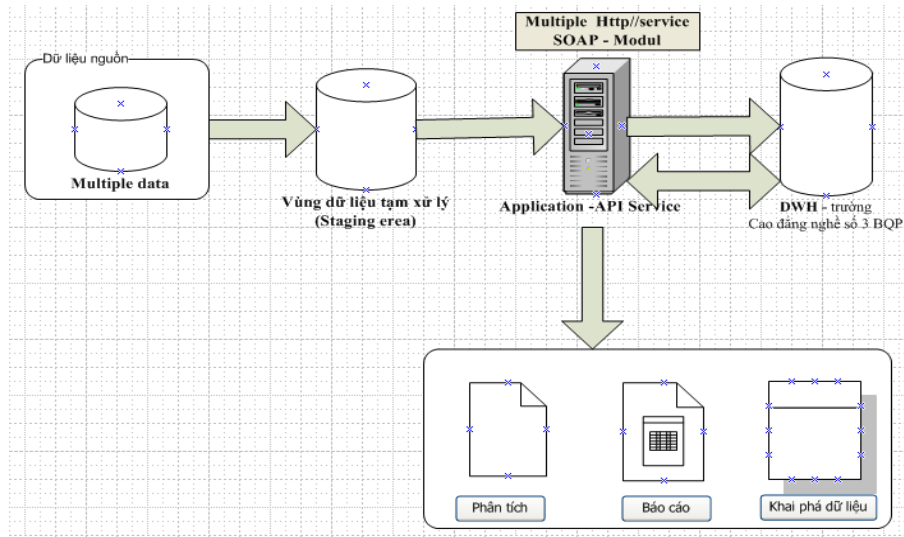
- Bảng trang thiết bị vật tư: qltb_thietbi_vattu
- Bảng mã nhóm trang thiết bị :qltb_nhom_trangthietbi
- Bảng phòng ban đơn vị quản lý: qltb_phongban_quanly
- Bảng chủng loại trang thiết bị: qltb_chungloai_thietbi
- Bảng phiếu nhập trang thiết bị: qltb_phieunhap_thietbi
- Bảng phiếu xuất trang thiết bị: qltb_phieuxuat_thietbi
- Bảng phương tiện ô tô xe máy: qltb_phuongtien

d. Cơ sở dữ liệu quản lý thu chi tài chính

CSDL về quản lý thu chi tài chính được tập hợp từ ban tài chính của nhà trường bao gồm các công việc Quản lý tài chính, thu học phí, cân đối thu chi và tiền lương. Hiện đang sử dụng phần mềm Phần mềm quản lý kế toán của quân đội sử dụng nền FOXPRO For WINDOWS.

3.1.1.2 Kiến trúc của DW báo cáo thống kê

Việc thiết kế và tổ chức một DW là rất quan trọng vì nó ảnh hưởng đến việc tổ chức và khai thác báo cáo sau này. Do vậy quá trình này đòi hỏi những người thiết kế phải hiểu biết về các kiến trúc, các thành phần và các lược đồ của DW cũng như các ưu nhược điểm của mỗi loại. Việc chọn kiến trúc, lược đồ để xây dựng DW phải dựa trên những yêu cầu và đặc thù của bài toán lập các báo cáo thống kê và tận dụng nguồn dữ liệu từ những hệ thống quản lý dữ liệu cơ sở sẵn có. Từ yêu cầu của việc khai thác các báo cáo như mô tả ở trên, chúng tôi chọn xây dựng một DW BCTK, không có các DM. Sở dĩ có quyết định này vì các đơn vị khai thác báo cáo độc lập nhau (có những mẫu biểu báo cáo khác nhau và nhu cầu về các thông tin cũng khác nhau) nhưng có thể dùng chung nguồn dữ liệu. Ví dụ như cơ quan chủ quản BQP cũng có nhu cầu báo cáo thống kê tháng, quý, năm về quân số, trang thiết bị vật chất, phương tiện huấn luyện, quy mô tuyển sinh, quy mô đào tạo và kết quả tuyển sinh, rèn luyện. v.v ..., BLĐT BXH cũng quan tâm đến thông tin này, ngoài ra còn thông tin về số lượng, chất lượng đội ngũ giáo viên, thông tin về việc làm sau đào tạo ở các mức độ khác nhau và mẫu biểu khác nhau. Do vậy nếu chúng ta chia các DM theo các lĩnh vực của các nhóm chỉ tiêu thì việc tổ chức khai thác báo cáo sẽ rất khó khăn. Một đơn vị có thể sẽ lấy số liệu từ nhiều các DM. Để phù hợp với thực tế, chúng tôi xây dựng một DW với nhiều bảng Fact và các bảng Dimension. Trong đó mỗi bảng Fact sẽ tương ứng với một nhóm chỉ tiêu, các Dimension đại diện cho các chiều.



Hình 3.1: Kiến trúc DW BCTK

Kiến trúc DW đầy đủ gồm các lớp sau:

- **Dữ liệu nguồn**: được tạo bởi một tập các hệ CSDL 3 hệ cơ sở dữ liệu CSDL Quản lý Nhân sự, Quản lý Vật tư, Quản lý đào tạo
- **Vùng dữ liệu tạm** (Staging area): lưu dữ liệu tạm thời chưa kiểm tra để phục vụ tính toán hợp lệ trước khi đưa vào DW. Phục vụ cho quá trình làm sạch và trích chọn dữ liệu
- **API Service**: là ứng dụng được phát triển trên nền Web application có chức năng tiếp nhận các dữ liệu đã được modul hóa theo nghiệp vụ của chủ đề theo chuẩn SOAP đưa dữ liệu vào DW. Khai thác dữ liệu trong DW để đưa ra phân tích, báo cáo, khai phá dữ liệu...
- **DW**: chứa toàn bộ dữ liệu bao gồm các bảng Dimension, Facts, Atomic table phục vụ nghiệp vụ. Hệ quản trị kho dữ liệu là MySQL
- **Dimension**: là các table chứa dữ liệu đại diện cho một chiều, chính là các điều kiện dùng khi query
- **Fact**: là các table chứa dữ liệu theo nhiều chiều khác nhau
- **Atomic**: là các bảng dữ liệu được chuẩn hóa đưa vào DW

3.1.1.3 Đặc tả cấu trúc DW của trường Cao đẳng nghề số 3-BQP

Căn cứ vào hệ thống nguồn hiện có, nhu cầu quản lý và đối tượng báo cáo thống kê, việc xây dựng kho dữ liệu báo cáo thống kê của Nhà trường được xác định bằng việc xây dựng các kho dữ liệu chủ đề tương ứng sau:

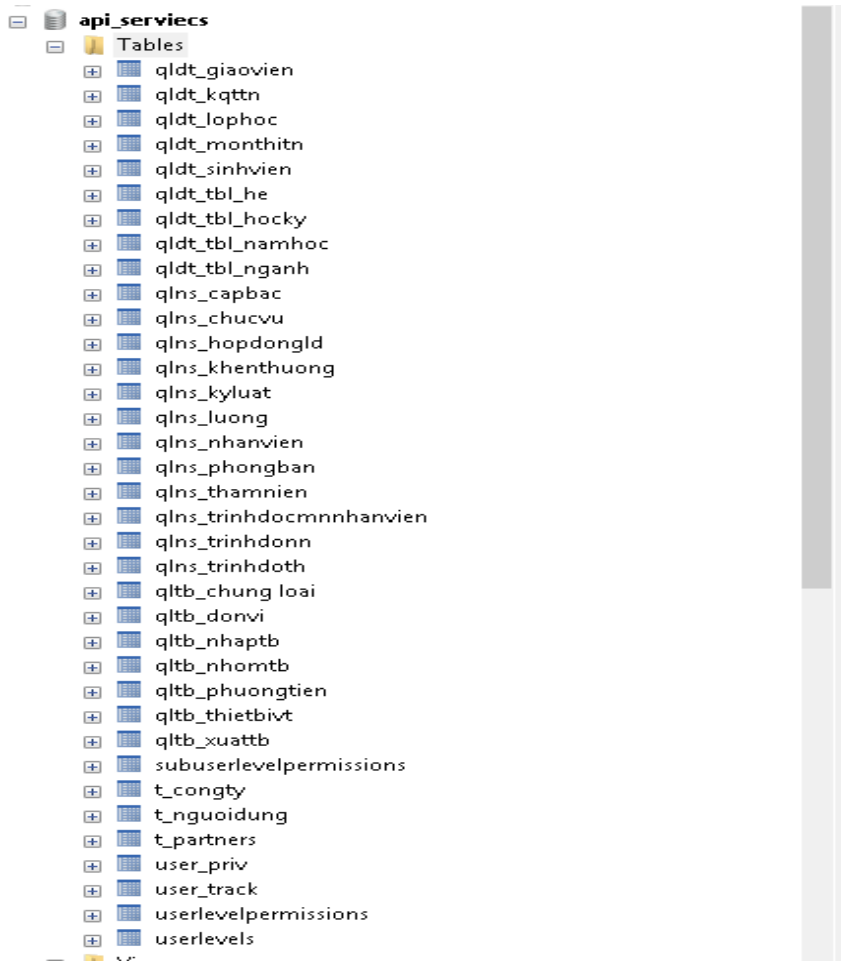
- Dữ liệu nhân sự
- Dữ liệu đào tạo
- Dữ liệu vật tư trang thiết bị

Chúng ta chọn sơ đồ hình sao để xây dựng các kho dữ liệu chủ đề, bởi vì giản đồ hình sao rất thuận lợi và dễ dàng thực hiện quá trình ETL để nạp dữ liệu vào chiều lưu trữ và tăng tốc độ truy xuất trong kho. Để xây dựng kho dữ liệu chủ đề, chúng ta cần xác định bảng sự kiện cùng các bảng chiều liên quan.

Quy ước đặt tên cho các bảng

Với kho DW tại trường Cao đẳng nghề số 3 – BQP, các bảng được phân biệt bởi các tiền tố theo nghiệp vụ.

- **qldt:** là tiền tố đặt tên với các bảng liên quan đến nghiệp vụ quản lý đào tạo. Các bảng được đặt tên theo quy ước **qldt_nametable**.
- **qlns:** là tiền tố đặt tên với các bảng liên quan đến nghiệp vụ quản lý nhân sự. Các bảng được đặt tên theo quy ước **qlns_nametable**.
- **qttb:** là tiền tố đặt tên với các bảng liên quan đến nghiệp vụ quản lý vật tư. Các bảng được đặt tên theo quy ước **qttb_nametable**.



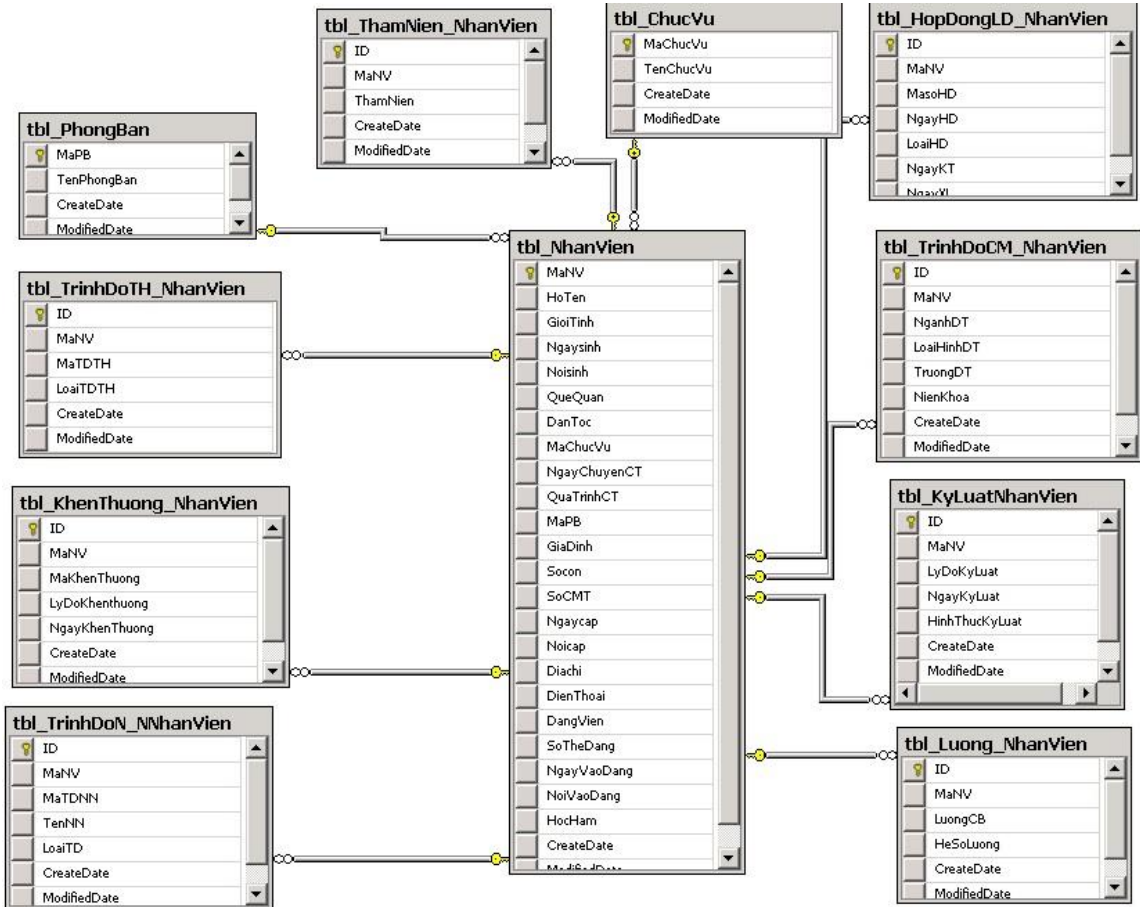
Hình 3.2: Cơ sở dữ liệu DW

3.1.2 Xây dựng bảng các tiêu chí

3.1.2.1 Bảng dữ kiện chuyên đề nhân sự

Bảng dữ kiện (Dimension Table - Fact Table) cho kho dữ liệu chuyên đề nhân sự:

- Bảng Dimension: tbl_PhongBan, tbl_ChucVu
- Bảng Fact: tbl_TrinhDoTH_NhanVien, tbl_NhanVien, tbl_HopDongLD_NhanVien, tbl_TrinhDoTH_NhanVien, tbl_NhanVien, tbl_KhenThuong_NhanVien, tbl_KhenThuong_NhanVien, tbl_Luong_NhanVien...

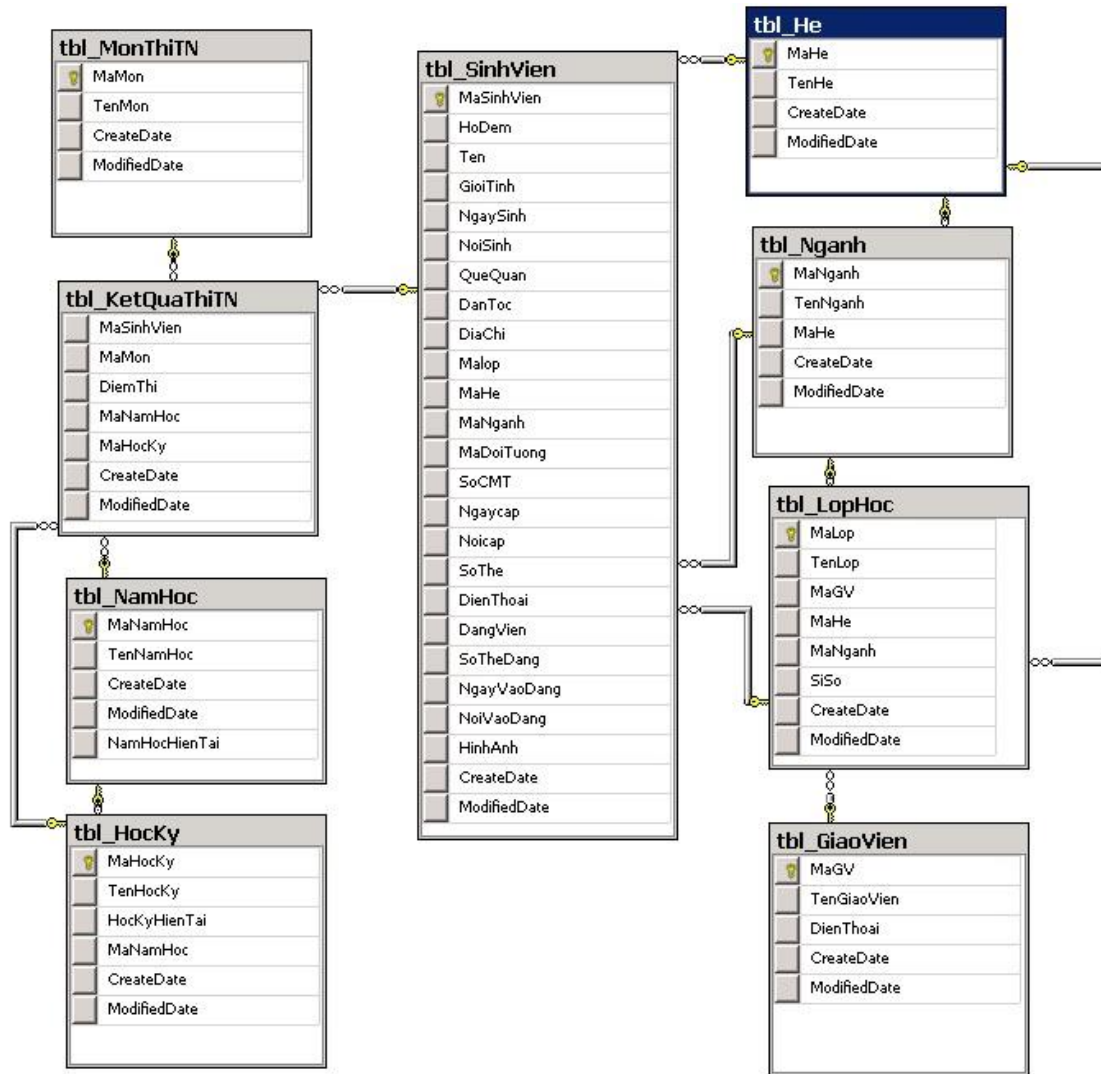


Hình 3.3: Sơ đồ quan hệ chủ đề Quản Lý Nhân Sự

3.1.2.2 Bảng dữ kiện chuyên đề đào tạo

Bảng dữ kiện (Dimension Table - Fact Table) cho kho dữ liệu chuyên đề đào tạo:

- Bảng Dimension: tbl_MonThi, tbl_He, tbl_HocKy, tbl_NamHoc, tbl_Nganh, tbl_Khoa ..
- Bảng Fact: tbl_SinhVien, tbl_KetQuaThiTN, tbl_LopHoc...

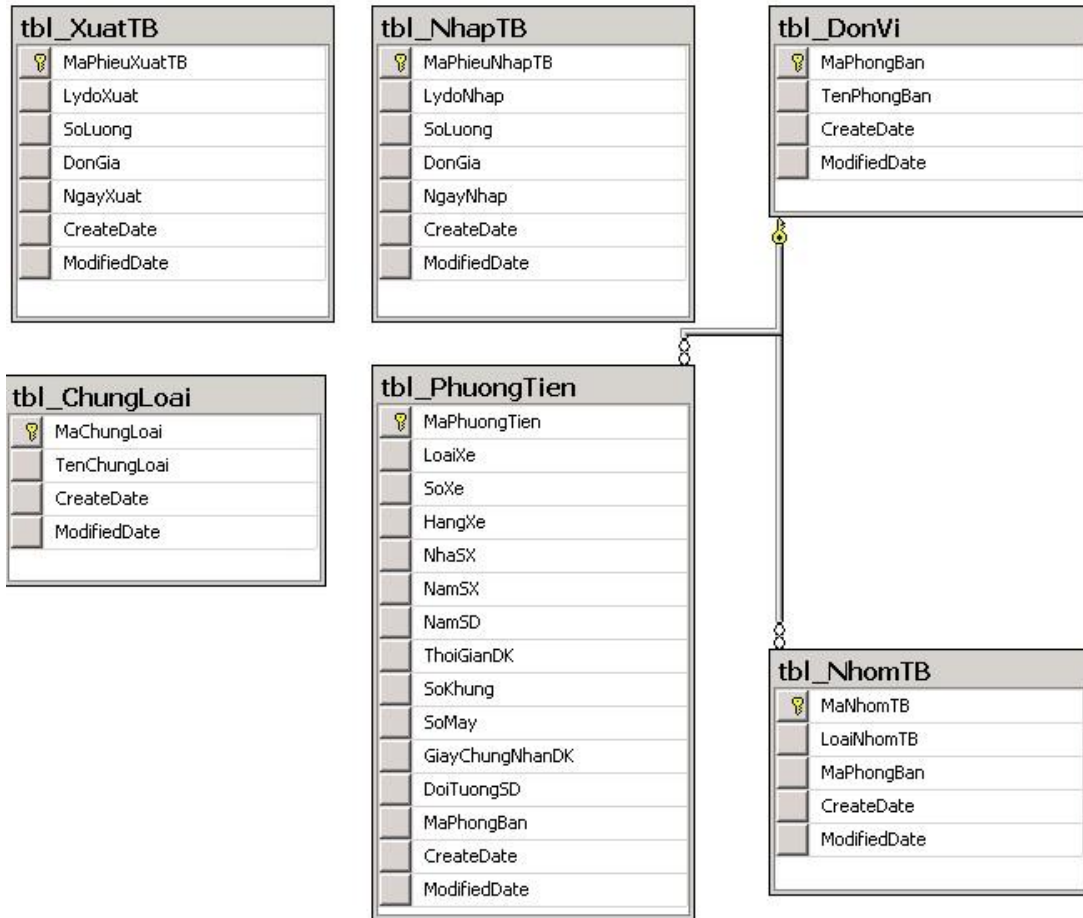


Hình 3.4: Sơ đồ quan hệ chủ đề Quản Lý Đào Tạo

3.1.2.3 Bảng dữ kiện chuyên đề vật tư trang thiết bị dạy nghề

Bảng dữ kiện (Dimension Table - Fact Table) cho kho dữ liệu chuyên đề trang thiết bị dạy nghề:

- Bảng Dimension: tbl_DonVi, tbl_NhomTB, , tbl_ChungLoai ..
- Bảng Fact: tbl_PhuongTien, tbl_XuatTB, tbl_NhapTB ...
- Hình 3.5. Sơ đồ quan hệ chủ đề Quản lý thiết bị



Hình 3.5: Sơ đồ quan hệ chủ đề Quản lý thiết bị

3.1.3 Xây dựng quy trình tích hợp dữ liệu cho DW

3.1.3.1 Giải pháp tích hợp dữ liệu vào kho dữ liệu BCTK

Trên thực tế có 2 giải pháp giúp cho việc tích hợp dữ liệu Full replacement và Incremental.

- Full replacement

Thực hiện cập nhật lại toàn bộ dữ liệu mỗi khi khi tích hợp. Phương pháp này mất nhiều thời gian và chi phí, do đó chỉ được áp dụng cho kho dữ liệu nhỏ.

-Incremental

Giải pháp này thực hiện cập nhật dữ liệu mới phát sinh cho DW, do đó thích hợp cho các kho dữ liệu lớn. Có 2 cách thức áp dụng:

+ **Thực hiện tích hợp dựa trên dữ liệu nguồn**

- Phụ thuộc vào dữ liệu nguồn
- Giảm được thời gian tích hợp (do lọc được khối lượng dữ liệu nguồn)

+ **Dựa vào dữ liệu đích khi tích hợp**

- Không phụ thuộc vào dữ liệu nguồn
- Sao chép toàn bộ dữ liệu nguồn và đích vào vùng đệm, sau đó thực hiện đồng bộ dữ liệu.
- Mất nhiều thời gian để thực hiện

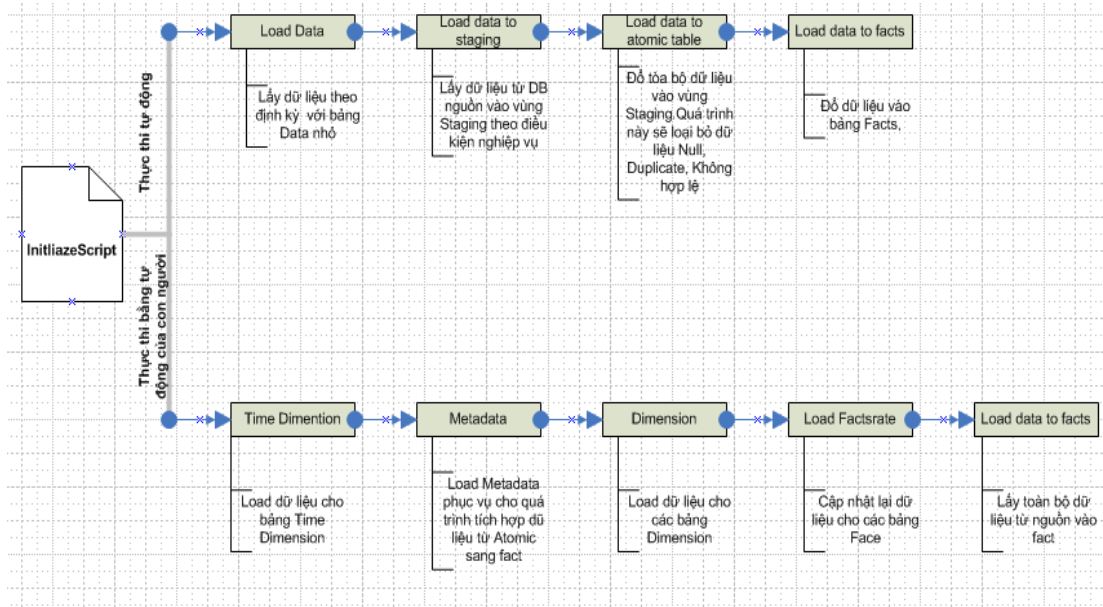
Trên cơ sở những mô tả của bài toán nghiệp vụ và căn cứ vào yêu cầu tích hợp dữ liệu ở trên, ta nhận thấy việc chọn tích hợp dữ liệu cho DW BCTK theo giải pháp *Incremental* và cách thực hiện tích hợp dựa trên dữ liệu nguồn là tối ưu nhất vì kho dữ liệu BCTK rất lớn. Với việc chọn giải pháp này, các bảng dữ liệu nguồn phải có 1 trường *timestamp* để lưu vết thời gian khi dữ liệu được cập nhật. Khi thực hiện, công cụ tích hợp sẽ dựa vào trường này để lọc dữ liệu mới phát sinh như sau:

Thực hiện theo lịch lập sẵn: sử dụng tích hợp dữ liệu phát sinh hàng ngày, dữ liệu của các chỉ tiêu...

Tác động trực tiếp của con người:

- Thực hiện cập nhật lại dữ liệu theo các điều kiện lọc khi có nhu cầu, nghiệp vụ
- Cập nhật lại dữ liệu lỗi (do yếu tố khách quan: mạng lỗi..) của các lần thực hiện trước
- Cập nhật lại dữ liệu sai logic
- Các nhu cầu khác : metadata tích hợp dữ liệu

3.1.3.2 Các bước thực hiện quá trình tích hợp dữ liệu



Hình 3.6: Các bước của quá trình tích hợp dữ liệu vào DW

3.1.4 Thu thập và tạo lập dữ liệu

Một phần quan trọng của việc cài đặt kho dữ liệu là sử dụng những dữ liệu đã được tinh chế từ 3 hệ thống tác nghiệp và đưa chúng vào một khuôn dạng thích hợp Staging cho ứng dụng

Giai đoạn này thực hiện các công việc chuyển đổi tóm tắt những thay đổi quan trọng, những thay đổi về cấu trúc và những cô đọng cần thiết cho sự chuyển đổi dữ liệu riêng rẽ thành thông tin có thể được dùng trong các hệ ra quyết định

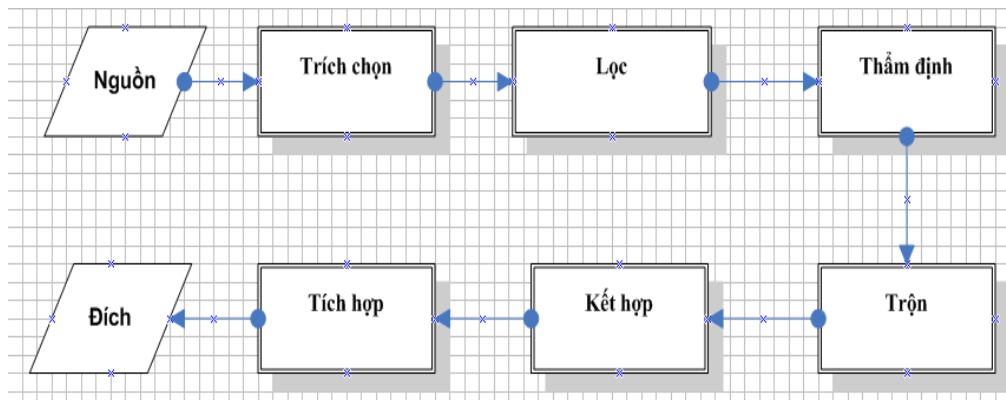
Các chức năng chính bao gồm:

- Lựa chọn, loại bỏ những dữ liệu không mong muốn lấy từ 3 hệ cơ sở Quản lý Vật tư, Quản lý Đào tạo, Quản lý Nhân sự
- Chuyển đổi thành những tên gọi và những định nghĩa dữ liệu chung, tổng quát
- Thực hiện , tính toán và tổng hợp dữ liệu được chuẩn hóa
- Thiết lập mặc định cho các dữ liệu bị mất
- Làm cho những thay đổi về định nghĩa dữ liệu nguồn trở lên thích hợp

Các chức năng này có thể tiết kiệm được một cách đáng kể thời gian và sức lực. Do vậy việc phát triển những thủ tục tinh chế cho một số lĩnh vực ứng dụng là cần thiết cho việc tinh chế dữ liệu.

Quá trình tạo lập dữ liệu của DW thực hiện các công đoạn bao gồm

- Trích chọn dữ liệu
- Lọc và tinh chế dữ liệu
- Thẩm định dữ liệu
- Gộp, kết hợp dữ liệu
- Lưu trữ và phát triển, phân phối dữ liệu



Hình 3.7: Quy trình tạo lập dữ liệu

3.1.4.1 Trích chọn dữ liệu

Trích dữ liệu là phép xử lý lấy các dữ liệu đã được xác định trước từ các hệ thống tác nghiệp và các nguồn dữ liệu bên ngoài. Các dữ liệu thường di chuyển từ các hệ thống OLTP tới kho dữ liệu. Khi trích chọn dữ liệu có thể gặp một số vấn đề như sau:

- Không có đủ thông tin chi tiết về hệ thống.
- Hoặc người sử dụng đầu cuối yêu cầu thông tin ở mức thấp hơn mức thông tin của hệ thống lưu trữ.

Một nhiệm vụ quan trọng của bước này là phải lập ra kế hoạch và tần suất tiến trình trích chọn dữ liệu, tối ưu hóa các tác động các hệ thống và thi hành các tác vụ

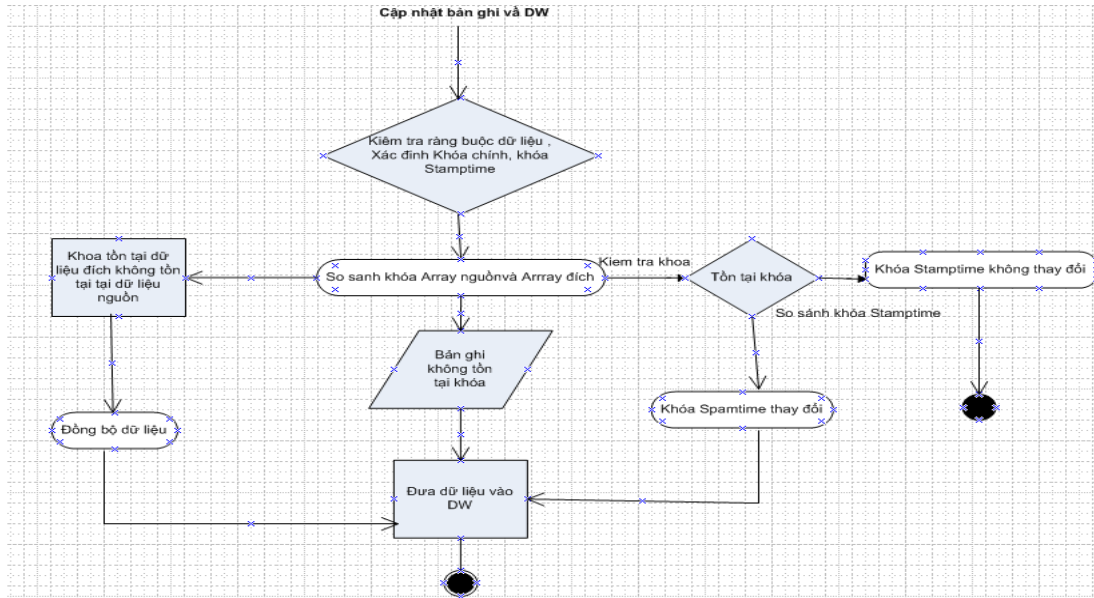
này trong một cửa sổ xử lý theo lô. Đối với các bảng dữ liệu khác nhau thì tần suất trích dữ liệu sẽ khác nhau. Việc trích dữ liệu còn phụ thuộc vào sự ảnh hưởng từ hệ thống nguồn và loại dữ liệu sẽ được trích.

Ví dụ với kho dữ liệu về chủ đề Đào tạo, dữ liệu về sinh viên được phát sinh theo từng học kỳ, từng năm nên có thể trích dữ liệu theo từng kỳ hoặc từng năm. Các dữ liệu trích chọn đầu tiên được chuyển tải và kho dữ liệu tạm thời, Trong khuôn khổ luận văn này quá trình tích hợp dữ liệu được thực hiện thông qua Web Service với các hệ CSDL tác vụ SQL Server Procedure.

3.1.4.2 Lọc và tinh chế dữ liệu

Sau khi dữ liệu được trích chọn, nó được tinh chế thông qua các công việc chọn lọc, làm sạch để thu được dữ liệu không bị thay đổi và đúng với các dữ liệu về nghiệp vụ.

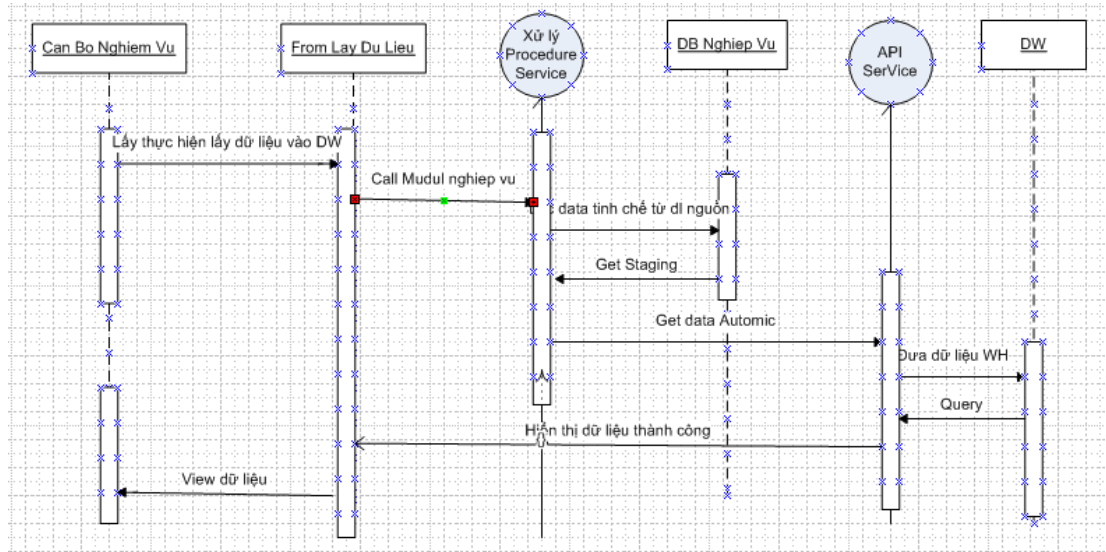
Quá trình lọc, làm sạch dữ liệu kiểm tra và sửa chữa các lỗi của dữ liệu để đảm bảo tính đúng đắn của dữ liệu. Quá trình này luôn tồn tại hai khóa: một khóa chính và một khóa phụ (Dùng time creat; time edit và cờ trạng thái) thời gian xác định sự thay đổi. Chọn giải pháp sử dụng cơ chế Slowly Changing Dimension (Cách thức cập nhật dữ liệu cho các bảng dimension) theo cách tạo một bản ghi mới dùng để phân tách rõ ràng dữ liệu lịch sử và dữ liệu hiện tại.



Hình 3.8: Biểu đồ hoạt động (UML) lọc tinh chế dữ liệu

Công việc này bao gồm các thao tác sau:

- Kiểm tra các trường dữ liệu đơn lẻ hoặc các trường có liên kết chéo
- Kiểm tra bản ghi được thiết lập :thu nhật, sắp sếp so sánh đối chiếu
 - TH1: Bản ghi không tồn tại Khóa → cập nhật vào DW
 - TH2: Bản ghi tồn tại Khóa nhưng giá trị Khóa phụ không thay đổi → không cập nhật DW
 - TH3: Bản ghi tồn tại Khóa nhưng giá trị Khóa phụ thay đổi → cập nhật vào DW
 - TH4: Bản ghi tồn tại khóa ở dữ liệu đích không tồn tại ở dữ liệu nguồn → đồng nhất dữ liệu trên DW

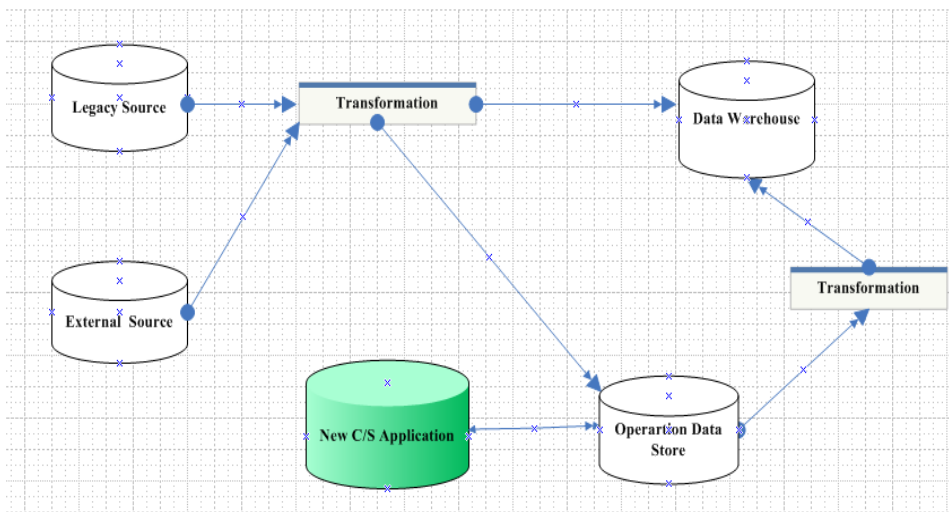


Hình 3.9: UML Sequence lọc tinh chế dữ liệu

3.1.4.3 Thẩm định và chuyển đổi dữ liệu

Mục đích của các bước này là chuyển dữ liệu thành thông tin và làm cho chúng dễ hiểu và dễ sử dụng hơn đối với người sử dụng đảm bảo hiệu năng tối đa trên hệ thống đưa việc tính toán thông tin về DW. Việc chuyển đổi có thể bao gồm các bước sau:

- Chuyển trực tiếp dữ liệu đồng nhất về trường lưu trữ và dữ liệu
- Xây dựng lại định dạng dữ liệu, có thể chuyển một phần hoặc phải gộp cả các trường lại với nhau để tạo lên một trường mới

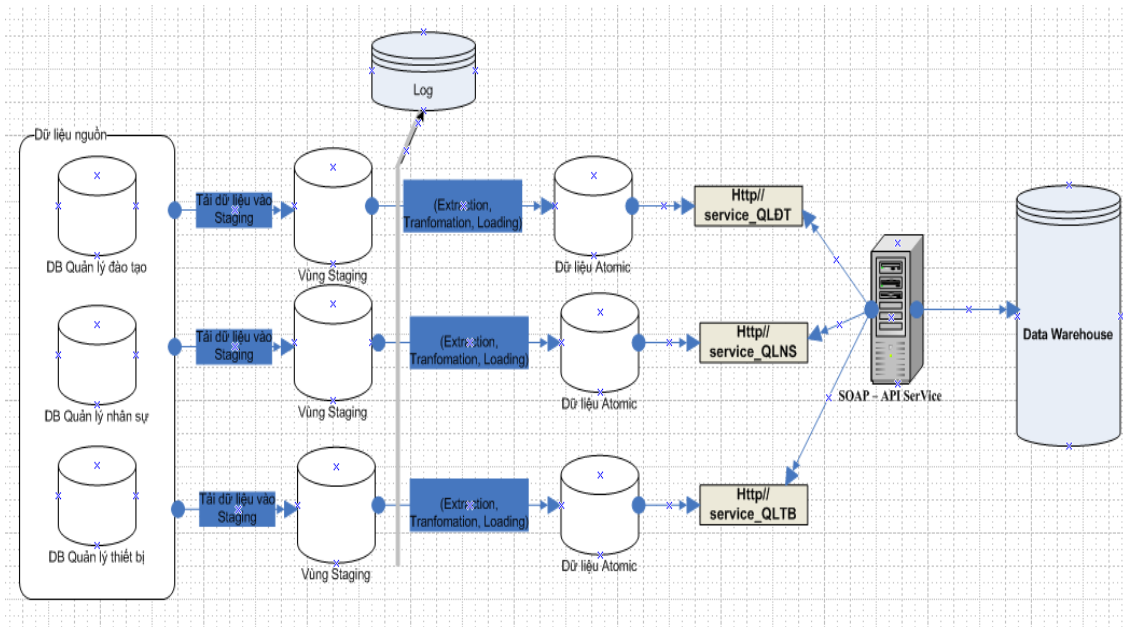


Hình 3.10: Quy trình thu thập vào chuyển đổi dữ liệu

3.1.4.4 Tích hợp dữ liệu

Khi có nhiều nguồn dữ liệu thì chúng ta cần thiết phải tích hợp lại để hợp nhất và tổ chức lại dữ liệu cho phù hợp với kiến trúc DW và nhu cầu sử dụng tại trường cao đẳng nghề số 3- BQP. Tích hợp thông tin từ nhiều nguồn dữ liệu hay từ hệ thống này sang hệ thống khác bao gồm:

- Sắp xếp hợp nhất khi một bảng dữ liệu đích được tạo lên từ nhiều nguồn dữ liệu. Khi đó dữ liệu phải được sắp xếp lại và loại bỏ đi các bản ghi giống nhau. Sau giai đoạn sắp xếp thì dữ liệu mới được hợp nhất thành một file dữ liệu duy nhất.
- Chia cắt dữ liệu nếu như từ một dữ liệu nguồn cần tạo ra nhiều dữ liệu đích.
- Đưa ra và giải quyết các vi phạm về tính nguyên vẹn của dữ liệu. Một giải pháp là lưu trữ dữ liệu đệm tiền xử lý.
- Tạo ra các khóa tổng hợp có khả năng tách biệt và bảo vệ khỏi các thay đổi trong hệ thống nguồn.
- DW cần phải được xây dựng để khỏi bị tác động của các thay đổi và sửa đổi đặc biệt là khóa nguồn.



Hình 3.11: Kiến trúc tích hợp dữ liệu

Các bước thực hiện

- Tải toàn bộ dữ liệu vào vùng Staging theo điều kiện chức năng đã được thiết lập
- Thực hiện trích lọc, chuyển đổi và tải dữ liệu vào các bảng tạm (Vùng Staging)
- Ghi log dữ liệu lỗi nếu xảy ra.
- Trích tách, chuyển đổi và tải dữ liệu sạch từ các bảng dữ liệu tạm sang các Atomic data.
- Sử dụng Webservice xây dựng trên chuẩn SOAP trên nền tảng là XML modul hóa các nghiệp vụ đưa vào DW kho dữ liệu trường Cao đẳng nghề số 3 BQP.

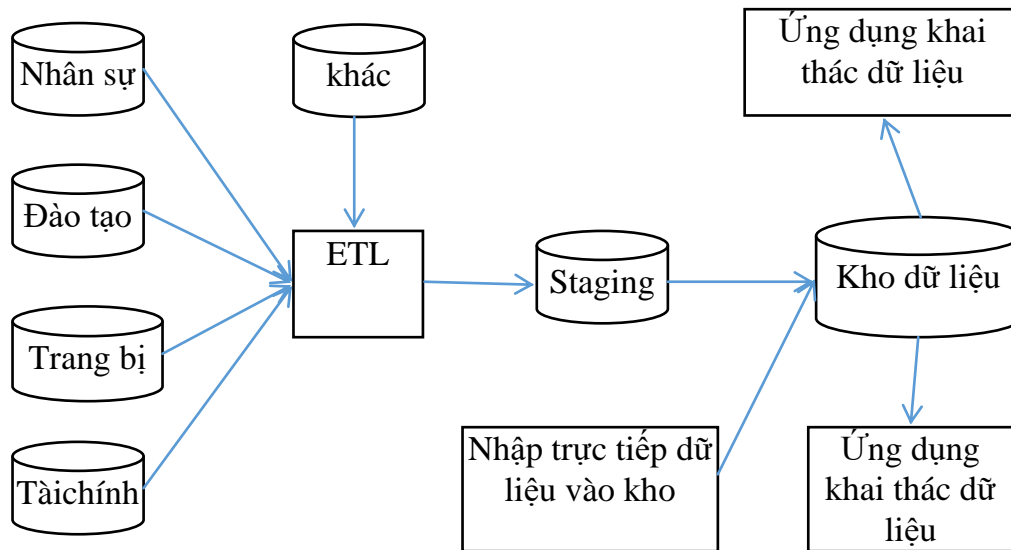
Tải dữ liệu là quá trình chuyển dữ liệu từ nơi lưu trữ nguồn vào CSDL đích. Trước hết cần xác định khoảng thời gian và định kỳ cần thiết để chuyển dữ liệu vào kho .không những cần xem lần nạp dữ liệu đầu tiên mà còn phải quan tâm đến việc làm tươi dữ liệu. tiến trình tải đầu tiên dùng để chuyển dữ liệu vào DW mới được xây dựng, tiến trình này chỉ được thực hiện một lần và không cần quan tâm đến vấn đề thời gian. Các tiến trình nạp dữ liệu tiếp theo là công việc hàng ngày của DW, tiến trình này được kiểm soát chặt chẽ về mặt thời gian. Thời gian thực hiện tiến trình tải dữ liệu luôn là một vấn đề cần quan tâm của các nhà quản trị DW do DW sử dụng dữ liệu từ

rất nhiều nguồn khác nhau mà mỗi loại lại có một khoảng thời gian tổng hợp dữ liệu riêng. Để có thể đồng bộ hóa được dữ liệu từ nhiều nguồn dữ liệu khác nhau cần tính toán kỹ trước khi thực hiện.

Để giải quyết vấn đề này chúng ta đi xây dựng kho tạm để toàn bộ dữ liệu được tập trung vào kho tạm sau đó chuyển vào DW tại thời điểm thích hợp trong ngày, thời gian tải dữ liệu được lựa chọn sao cho ít ảnh hưởng đến công việc hiện tại của DW.

3.1.5 Kiến trúc luồng dữ liệu

Qua kết quả khảo sát nhận thấy có rất nhiều nguồn dữ liệu khác nhau nên kho dữ liệu trường Cao đẳng nghề số 3 BQP cũng được tích hợp từ những nguồn dữ liệu đó, đồng thời cần bổ sung thêm một số dữ liệu cần thiết còn thiếu. Kiến trúc luồng dữ liệu của kho dữ liệu nhà trường được thiết kế như sau:



Hình 3.12: Kiến trúc luồng dữ liệu trường Cao đẳng nghề số 3 BQP

Quá trình ETL nhận dữ liệu từ CSDL quản lý nhân sự, quản lý đào tạo, quản lý trang bị vật tư thiết bị, quản lý tài chính và các nguồn dữ liệu khác nạp về vùng chứa tạm Staging. CSDL tạm (Staging) là nơi lưu trữ tạm thời dữ liệu trích được từ hệ thống nguồn, trước khi thực hiện các bước tiếp theo. Một Staging rất cần thiết đối với việc chuyển đổi phức tạp, dữ liệu nguồn lớn và giảm tối đa thời gian trích dữ liệu từ hệ

thông nguồn rồi nạp về kho dữ liệu của nhà trường. Ngoài ra những dữ liệu trong kho còn bị thiếu cần phải cập nhập trực tiếp vào kho để phục vụ các ứng dụng khai thác dữ liệu trong kho.

3.2 Thiết kế Webservice

Lựa chọn hướng tiếp cận (bottom-up) từ 3 hệ CSDL Qlvt, Qlns, Qldt để xây dựng một dịch vụ Web, đảm bảo hệ thống đáp ứng các nhu cầu về nghiệp vụ quản lý tại trường Cao đẳng nghề số 3 – BQP. Trong đó tối đa việc sử dụng lại các chức năng, các thành phần, modul đã được xây dựng.

Quy trình xây dựng một dịch vụ Web service bao gồm các bước sau:

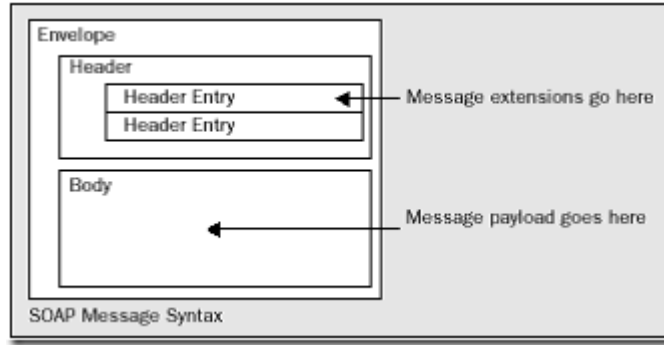
1. Định nghĩa và xây dựng các chức năng, các dịch vụ mà dịch vụ sẽ cung cấp theo nghiệp vụ
2. Tạo WSDL cho dịch vụ
3. Xây dựng SOAP server (*Sử dụng ngôn ngữ lập trình C#, SQL Procedure , SQL Server*)
4. Đăng ký WSDL với UDDI registry để cho phép các client có thể tìm thấy và truy xuất
5. Client nhận file WSDL và từ đó xây dựng SOAP client để có thể kết nối với SOAP server
6. Xây dựng ứng dụng phía client và sau đó gọi thực hiện dịch vụ thông qua việc kết nối tới SOAP server. (*Sử dụng ngôn ngữ lập trình PHP, library NuSOAP , MySQL Server*)

3.2.1 Cách thức hoạt động

- Dựa trên Nền tảng cơ bản của WS là XML + HTTP.XML cung cấp một ngôn ngữ mà có thể được sử dụng giữa ngôn ngữ lập trình và các nền tảng khác. Đồng thời, nó còn có thể được dùng để mô tả những thông điệp và chức năng phức tạp. Do web service là sự kết hợp của nhiều thành phần khác nhau, do đó web services sử dụng các tính năng và đặc trưng của các thành phần này để giao tiếp với nhau.

Vì vậy XML là một công cụ chính yếu để giải quyết vấn đề này. Web service tận dụng khả năng giải quyết vấn đề của các ứng dụng lớn trên các hệ điều hành khác nhau cho chúng giao tiếp với nhau.

- Sử dụng Giao thức HTTP là giao thức được sử dụng nhiều nhất trong các giao thức trên internet.
- Web Service lựa chọn sử dụng giao thức SOAP (Simple Object Access Protocol)
- SOAP là một giao thức dựa trên nền XML cho phép ứng dụng trao đổi các thông tin qua HTTP. Nói cụ thể hơn thì SOAP là một giao thức dùng để truy cập các dịch vụ web.
- SOAP là một giao thức truyền thông.
- SOAP là một định dạng dùng để gửi đi các thông điệp.
- SOAP được thiết kế để giao tiếp thông qua internet.
- SOAP là một nền tảng XML.
- SOAP rất đơn giản và có thể mở rộng.
- SOAP cho phép chúng ta vượt qua bức tường lửa (firewall)
- SOAP là một tiêu chuẩn W3C.
- SOAP là giao thức sử dụng XML để định nghĩa dữ liệu dạng thuần văn bản (plain text) thông qua HTTP. SOAP là cách mà web service sử dụng để truyền tải dữ liệu. Vì dựa trên XML nên SOAP là một giao thức không phụ thuộc platform cũng như bất kỳ ngôn ngữ lập trình nào.
- Một thông điệp SOAP được chia thành hai phần là header và body. Phần header chỉ ra địa chỉ web service, host, Content-Type, Content-Length tương tự như một thông điệp HTTP.

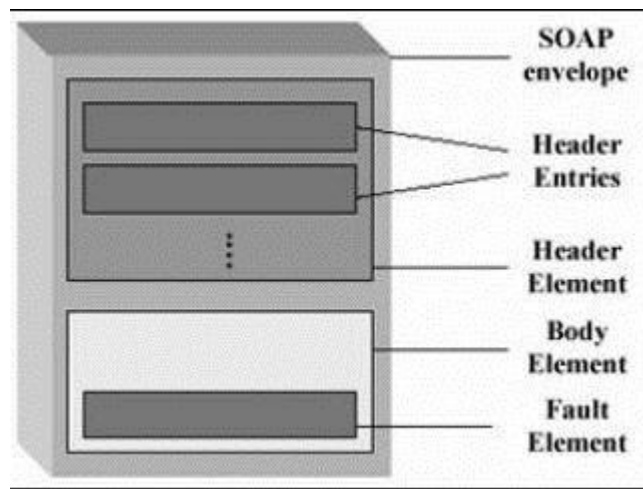


Hình 3.13: Cấu trúc thông điệp SOAP

Envelop là thành phần gốc của một thông điệp SOAP, nó chứa các thành phần Header và Body. Thành phần Header là một cơ chế mở cho phép thêm các tính năng vào bên trong một thông điệp SOAP. Mỗi thành phần con của Header gọi là một Header Entry. Các Header Entry dùng để diễn giải, quy định một số ngữ nghĩa của thông điệp SOAP. Các ứng dụng có thể xử lý và định tuyến các thông điệp dựa trên thông tin header và thông tin bên trong thông điệp đó. Đây là ưu điểm mà các mô hình kiến trúc như DCOM, CORBA và RMI không có được, vì các protocol header của chúng phải được chỉ định chi tiết cho mỗi ứng dụng.

3.2.2 Cấu trúc một message theo dạng SOAP

Cấu trúc một message theo dạng SOAP được mô tả như hình dưới đây:



Hình 3.14: Cấu trúc một message theo dạng SOAP

Message theo dạng SOAP là một văn bản XML bao gồm các phần tử sau:

- *Phần tử gốc – envelop*: Phần tử bao trùm nội dung message, khai báo văn bản XML như là một thông điệp SOAP.
- *Phần tử đầu trang – header*: Chứa các thông tin tiêu đề cho trang, phần tử này không bắt buộc khai báo trong văn bản. Những đầu mục còn có thể mang những dữ liệu chứng thực, những chữ ký số hóa, và thông tin mã hóa, hoặc những cài đặt cho giao tác.
- *Phần tử khai báo nội dung chính trong thông điệp – body*, chứa các thông tin yêu cầu và phản hồi.

Phần tử phát sinh lỗi (Fault) cung cấp thông tin lỗi xảy ra trong quá trình xử lý thông điệp.

Trong trường hợp đơn giản nhất, phần thân của SOAP message gồm có:

- *Tên của message*

Một tham khảo tới một thể hiện service.

- Một hoặc nhiều tham số mang các giá trị và mang các tham chiếu. Có 3 kiểu thông báo
 - *Request messages*: Với các tham số gọi thực thi một service
 - *Response messages*: Với các tham số trả về, được sử dụng khi đáp ứng yêu cầu.
 - *Fault messages*: báo tình trạng lỗi.

3.2.3 Cách truyền thông SOAP xây dựng WS

Sử dụng hai kiểu truyền thông SOAP hỗ trợ để xây dựng WS:

- Remote procedure call (RPC): Cho phép gọi hàm hoặc thủ tục qua mạng. Kiểu này được khai thác bởi nhiều web service và có nhiều trợ giúp.
- Document: Kiểu này cung cấp một lớp thấp của sự trừu tượng hóa và yêu cầu người lập trình nhiều hơn khi làm việc.

Các định dạng message, tham số và lời gọi đến các API thì tương ứng trong RPC và document là khác nhau. Nên việc quyết định chọn cái nào tùy thuộc vào thời gian xây dựng và sự phù hợp của service cần xây dựng.

Mục đích của mô hình dữ liệu SOAP là cung cấp sự trừu tượng hóa độc lập ngôn ngữ cho kiểu ngôn ngữ lập trình. Nó gồm có:

- Những kiểu XSD đơn giản như những kiểu dữ liệu cơ bản trong đa số các ngôn ngữ lập trình như int, string, date...
- Những kiểu phức tạp, có hai loại là struct và array.

Tất cả các phần tử và những định danh có trong mô hình dữ liệu SOAP thì được định nghĩa bằng namespace SOAP-ENC

3.2.4 Một số Service được lập từ 3 hệ CSDL để đưa vào DW

3.2.4.1 Hệ Quản lý đào tạo được chạy trên địa chỉ: <http://10.1.0.236:8087>

← → 10.1.0.236:8087

QuanLyDaoTao

The following operations are supported. For a formal definition, please review the [Service Description](#).

- [Sinhvien](#)
Danh sách sinh viên: DataTable Sinhvien(); Output(MaSinhVien, HoDem, Ten, GioiTinh, NgaySinh, NoiSinh, QueQuan, DanToc, DiaChi, Malop, MaHe, TenNganh, MaDoiTuong, SoCMT, CreateDate)
- [SinhvienTN](#)
Danh sách sinh viên tốt nghiệp: DataTable SinhvienTN(); Output(MaSinhVien, HoDem, Ten, GioiTinh, NgaySinh, QueQuan, Malop, MaHe, MaNganh, MaNamHoc, MaHocky, DiemTBTN)

This web service is using <http://tempuri.org/> as its default namespace.

Recommendation: Change the default namespace before the XML Web service is made public.

Each XML Web service needs a unique namespace in order for client applications to distinguish it from other services on the Web. <http://tempuri.org/> is available for XML Web services that are under development, but published XML Web services should use a more permanent namespace.

Your XML Web service should be identified by a namespace that you control. For example, you can use your company's Internet domain name as part of the namespace. Although many XML Web service namespaces look like URLs, they need not point to actual resources on the Web. (XML Web service namespaces are URIs.)

For XML Web services creating using ASP.NET, the default namespace can be changed using the WebService attribute's Namespace property. The WebService attribute is an attribute applied to the class that contains the XML Web service methods. Below is a code example that sets the namespace to "<http://microsoft.com/webservices/>":

```
C#
[WebService(Namespace="http://microsoft.com/webservices/")]
public class MyWebService {
    // implementation
}

Visual Basic
<WebService(Namespace="http://microsoft.com/webservices/")> Public Class MyWebService
    ' implementation
End Class

C++
[WebService(Namespace="http://microsoft.com/webservices/")]
public ref class MyWebService {
    // implementation
};
```

For more details on XML namespaces, see the W3C recommendation on [Namespaces in XML](#).

For more details on WSDL, see the [WSDL Specification](#).

10.1.0.236

Hình 3.15: Modul Service Quản Lý Đào Tạo

(Xem B. Phụ lục 2, a. Đoạn COD 1)

3.2.4.2 Hệ Quản lý nhân sự được chạy trên địa chỉ: <http://10.1.0.236:8086>

The following operations are supported. For a formal definition, please review the [Service Description](#).

- **Nhanvien**
Danh sách nhân viên: DataTable NhanVien(); Output(MaNV, HoTen, GioiTinh, Ngaysinh, TenPhongBan, TenChucVu, LuongCB, HeSoLuong, ThamNien, CreateDate, ModifiedDate)

This web service is using <http://tempuri.org/> as its default namespace.

Recommendation: Change the default namespace before the XML Web service is made public.

Each XML Web service needs a unique namespace in order for client applications to distinguish it from other services on the Web. <http://tempuri.org/> is available for XML Web services that are under development, but published XML Web services should use a more permanent namespace.

Your XML Web service should be identified by a namespace that you control. For example, you can use your company's Internet domain name as part of the namespace. Although many XML Web service namespaces look like URLs, they need not point to actual resources on the Web. (XML Web service namespaces are URIs.)

For XML Web services creating using ASP.NET, the default namespace can be changed using the WebService attribute's Namespace property. The WebService attribute is an attribute applied to the class that contains the XML Web service methods. Below is a code example that sets the namespace to "<http://microsoft.com/webservices/>":

C#

```
[WebService(Namespace="http://microsoft.com/webservices/")]
public class MyWebService {
    // implementation
}
```

Visual Basic

```
<WebService(Namespace="http://microsoft.com/webservices/")> Public Class MyWebService
    ' implementation
End Class
```

C++

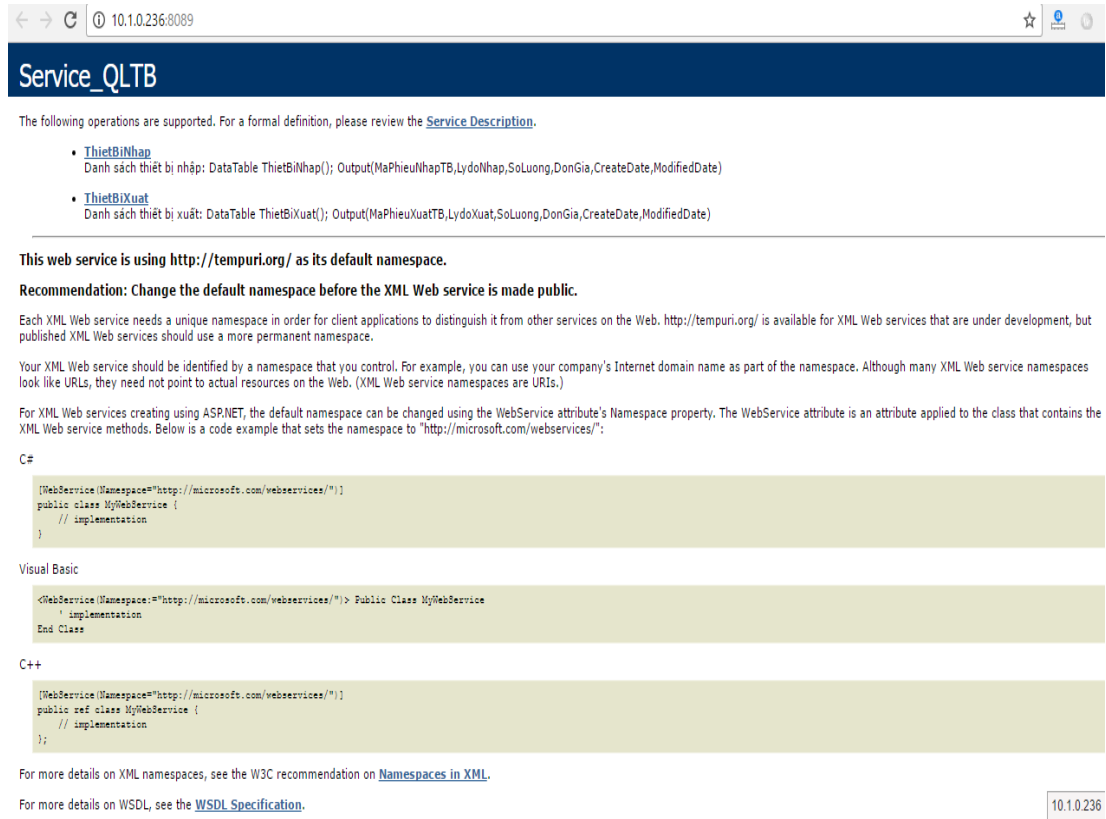
```
[WebService(Namespace="http://microsoft.com/webservices/")]
public ref class MyWebService {
    // implementation
};
```

For more details on XML namespaces, see the W3C recommendation on [Namespaces in XML](#).

For more details on WSDL, see the [WSDL Specification](#).

Hình 3.16: Modul Service Quản Lý Nhân Sự
(xem B. Phụ lục 2, b. Đoạn COD 2)

3.2.4.3 Hệ Quản lý thiết bị được chạy trên địa chỉ: <http://10.1.0.236:8089>



The following operations are supported. For a formal definition, please review the [Service Description](#).

- [ThietBiNhap](#)
Danh sách thiết bị nhập: DataTable ThietBiNhap(); Output(MaPhieuNhapTB, LydoNhap, SoLuong, DonGia, CreateDate, ModifiedDate)
- [ThietBiXuat](#)
Danh sách thiết bị xuất: DataTable ThietBiXuat(); Output(MaPhieuXuatTB, LydoXuat, SoLuong, DonGia, CreateDate, ModifiedDate)

This web service is using <http://tempuri.org/> as its default namespace.

Recommendation: Change the default namespace before the XML Web service is made public.

Each XML Web service needs a unique namespace in order for client applications to distinguish it from other services on the Web. <http://tempuri.org/> is available for XML Web services that are under development, but published XML Web services should use a more permanent namespace.

Your XML Web service should be identified by a namespace that you control. For example, you can use your company's Internet domain name as part of the namespace. Although many XML Web service namespaces look like URLs, they need not point to actual resources on the Web. (XML Web service namespaces are URIs.)

For XML Web services creating using ASP.NET, the default namespace can be changed using the WebService attribute's Namespace property. The WebService attribute is an attribute applied to the class that contains the XML Web service methods. Below is a code example that sets the namespace to "<http://microsoft.com/webservices/>":

C#

```
[WebService(Namespace="http://microsoft.com/webservices/")]  
public class MyWebService {  
    // implementation  
}
```

Visual Basic

```
<WebService(Namespace="http://microsoft.com/webservices/")> Public Class MyWebService  
    ' implementation  
End Class
```

C++

```
[WebService(Namespace="http://microsoft.com/webservices/")]  
public ref class MyWebService {  
    // implementation  
};
```

For more details on XML namespaces, see the W3C recommendation on [Namespaces in XML](#).

For more details on WSDL, see the [WSDL Specification](#).

10.1.0.236

Hình 3.17: Modul Service Quản Lý Thiết Bị

(xem B. Phụ lục 2, c. Đoạn COD 3)

3.3 Lập báo cáo

3.3.1 Yêu cầu nghiệp vụ của hệ thống khai thác báo cáo

Hệ thống được xây dựng cần:

- Hỗ trợ tổ chức và khai thác dữ liệu hướng đối tượng, từ đó cung cấp khả năng khai thác dữ liệu theo nhiều mức khác nhau: tổng thể hoặc chi tiết; biến động theo thời gian; so sánh giữa các đối tượng báo cáo và theo nhiều chiều khác nhau như theo chiều thời gian, theo các đơn vị báo cáo,
- Cung cấp giao diện đơn giản, hỗ trợ khai thác hiệu quả báo cáo đã có trong CSDL với yêu cầu kiến thức CNTT người sử dụng ở mức thông thường.

- Cung cấp khả năng truy xuất thông tin đột xuất; các dạng biểu diễn báo cáo theo bảng biểu hoặc đồ thị.
- Khai thác trực tiếp dữ liệu trên mạng cục bộ hoặc truy nhập từ xa qua mạng WAN dựa trên giao diện Web.
- Có khả năng quản lý người sử dụng, an toàn, bảo mật dữ liệu báo cáo theo thẩm quyền khai thác.
- Cung cấp mô hình phân tích, dự báo.
- Hỗ trợ giao diện với các phần mềm văn phòng của Microsoft

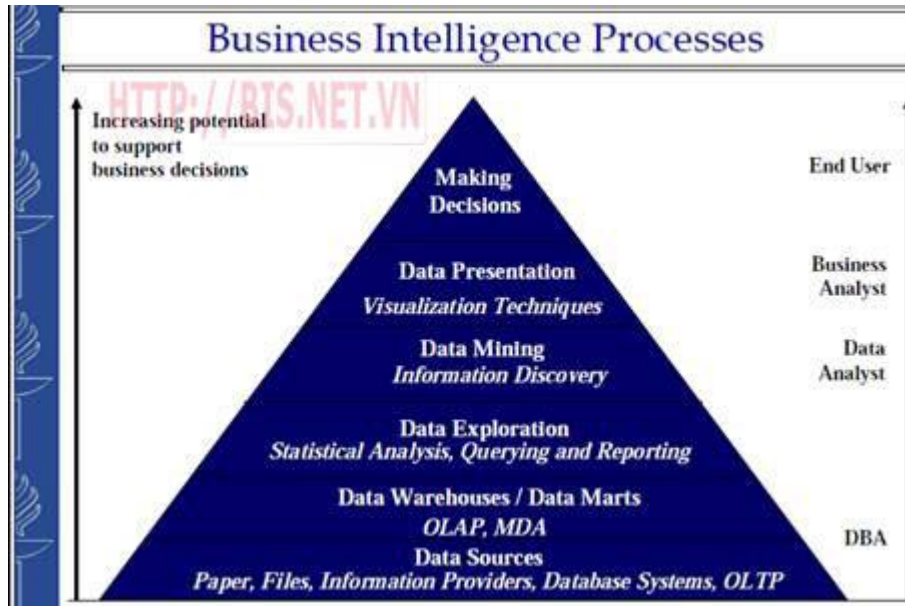
3.3.2 Yêu cầu cụ thể của hệ thống báo cáo

Hệ thống báo cáo cụ thể phải phục vụ cho việc lập các loại báo cáo theo mẫu có sẵn:

- Báo cáo thống kê theo định kỳ báo cáo.
- Báo cáo biến động đột xuất
- Báo cáo tổng hợp và báo cáo chi tiết

3.3.3 Giải pháp xây dựng hệ thống báo cáo

Sau khi đã có kho dữ liệu DW với các bảng Fact chứa dữ liệu theo các chỉ tiêu và các bảng Dimension về các chiều mà hệ thống khai thác báo cáo sẽ sử dụng. Việc tổ chức báo cáo nhằm hướng tới nghiệp vụ tại trường Cao đẳng nghề số 3 BQP phục vụ trong công tác quản lý và hệ ra quyết định tại trường. Tại đơn vị khai thác báo cáo người sử dụng có thể truy cập qua mạng Lan tại đơn vị truy cập được vào hệ thống báo cáo theo từng chủ đề. Báo cáo được thể hiện dựa trên các mẫu biểu có sẵn, hoặc dạng bảng biểu đồ thị.... Chính vì yêu cầu này cần phải tổ chức kho dữ liệu DW thành các thành phần (Các lớp) dữ liệu cho từng chủ đề khai thác (vật tư, nhân sự, đào tạo) theo quy trình giải pháp Business Intelligence (BI)



Hình 3.18: Quy trình giải pháp Business Intelligence (BI)

Data Sources: Là cơ sở dữ liệu thô (thường là cơ sở dữ liệu quan hệ) đến từ nhiều nguồn khác nhau như các ứng dụng business như Human Resource Management(HRM), Customer relationship management(CRM), phần mềm bán hàng, website thương mại điện tử... Có thể là bất cứ hệ quản trị cơ sở dữ liệu nào như MySQL, Oracle, MSSQL, DB2, ...Thường được thiết kế theo mô hình cơ sở dữ liệu quan hệ (vì dạng mô hình này đang rất phổ biến trong thực tế).

Data Warehouse: Là cơ sở dữ liệu được thiết kế theo mô hình khác với CSDL quan hệ và là nơi lưu trữ dữ liệu lâu dài của tổ chức. Dữ liệu của DW chỉ có thể đọc, không ghi hay update được và chỉ được update bởi gói ETL chuyển đổi dữ liệu từ Data Sources vào Data Warehouse.

Integrating Server: Chịu trách nhiệm trung gian vận hành gói ETL để chuyển đổi dữ liệu từ Data Sources vào Data Warehouse.

Analysis Server: Chịu trách nhiệm thực thi các Cube được thiết kế dựa trên các Dimension dữ liệu và tri thức nghiệp vụ Cube chịu trách nhiệm nhận input data từ DW và thực thi theo nghiệp vụ định nghĩa sẵn để trả về output.

Reporting Server: Thực thi các report với output nhận được từ Analysis Server. Nơi quản trị tập trung các report trên nền web, các report này có thể được attach vào ứng dụng web, hay application

Data Mining: Là quá trình trích xuất thông tin dữ liệu đã qua xử lý (phù hợp với yêu cầu riêng của doanh nghiệp) từ Data Warehouse rồi kết hợp với các thuật toán để đưa ra (hoặc dự đoán) các quyết định có lợi cho việc kinh doanh của doanh nghiệp. Đây là một quá trình quan trọng trong BI, thông thường một doanh nghiệp muốn sử dụng giải pháp BI thường kèm theo về Data Mining.

Data Presentation: Tạo ra các báo cáo, biểu đồ từ quá trình data mining để phục vụ cho nhu cầu của người dùng cuối.

3.3.4 Quá trình xây dựng báo cáo gồm các bước

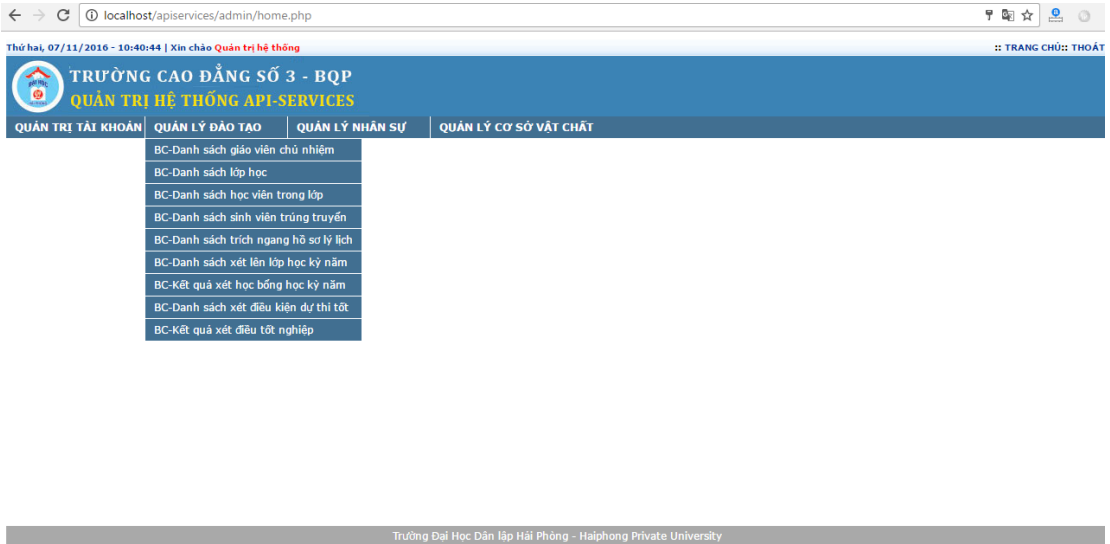
Tổng hợp và lưu trữ dữ liệu thô: Là việc tổng hợp dữ liệu thô từ nhiều nguồn vào một nơi lưu trữ chung phục vụ cho việc khai thác dữ liệu sau này

Sàng lọc dữ liệu: Dữ liệu thô chưa xử lý thường lớn và có thể trùng lặp, không cần thiết, sàng lọc dữ liệu khiến cho chất lượng dữ liệu đầu vào chính xác và phù hợp với nhu cầu sử dụng hơn

Lưu trữ dữ liệu: Thông qua việc nghiên cứu các hoạt động tại trường và nhu cầu của nghiệp vụ quản lý, dữ liệu được sàng lọc và lưu trữ vào Data Warehouse (DW)

Phân tích dữ liệu: Sử dụng các nguồn dữ liệu từ DW, xây dựng các cấu trúc dữ liệu tạo thành các Data Marts (DM). Data Marts là lớp đưa các dữ liệu từ DW tới người dùng. Reporting: Các báo cáo, thống kê, phân tích được xây dựng sử dụng cấu trúc từ Data Marts, phục vụ cho việc hiển thị dữ liệu báo cáo tới người sử dụng

3.3.5 Một số báo cáo dự kiến thiết kế



Hình 3.19: Tổ chức báo cáo hệ thống



Hình 3.20: Báo cáo Thống kê Sinh Viên Theo Ngành Học

TRƯỜNG CAO ĐẲNG SỐ 3 - BQP
QUẢN TRỊ HỆ THỐNG API-SERVICES

QUẢN TRỊ TÀI KHOẢN | QUẢN LÝ ĐÀO TẠO | QUẢN LÝ NHÂN SỰ | QUẢN LÝ CƠ SỞ VẬT CHẤT

BỘ QUỐC PHÒNG
Trường Cao đẳng nghề số 3

DANH SÁCH SINH VIÊN

STT	Mã SV	Họ đệm	Tên	Lớp	Giới tính	Ngày sinh	Dân tộc	Trạng thái
1	1215C-65001	Bùi Đức	Anh	CTCL201	Nam	30/9/1992		Đang học
2	1308C-65015	Nguyễn Mạnh	Cường	CTC1001	Nam	1/8/1994	1	Đang học
3	500094	Bùi Văn	Đức	CTC501	Nam	26/6/1990		Đang học
4	1208C-65019	Hoàng Tú Trường	Giảng	CTC901	Nam	3/10/1994		Đang học
5	1008C-65014	Hoàng Trọng	Hữu	CTC701	Nam	14/4/1990		Đang học
6	1108C-65013	Nguyễn Quang	Hiếu	CTC801	Nam	25/6/1993		Đang học
7	1208C-65023	Nguyễn Thái	Hoàng	CTC901	Nam	9/6/1993		Đang học
8	1408C-65004	Châu Quốc	Huy	CTC1101	Nam	10/6/1996	1	Đang học
9	1208C-65025	Nguyễn Trung	Kiến	CTC901	Nam	2/9/1994		Đang học
10	1008C-65036	Mai Trọng	Linh	CTC701	Nam	3/4/1992		Đang học
11	1208C-65001	Đoàn Xuân	Lộc	CTC901	Nam	1/3/1994		Đang học
12	1408C-65010	Lê Hoàng Phương	Nam	CTC1101	Nam	27/6/1996	1	Đang học
13	1008C-65021	Đào Nhật	Ninh	CTC701	Nam	25/1/1991		Đang học
14	1008C-65004	Nguyễn Quốc	Phong	CTC701	Nam	22/9/1992		Đang học
15	1208C-65016	Lương Duy	Phượng	CTC901	Nam	13/7/1993		Đang học
16	061C-65015	Nguyễn Ngọc	Sơn	CTC601	Nam	5/2/1991		Đang học
17	1608C-65001	Đào Trường	Sơn	CTC1301	Nam	25/6/1995		Đang học
18	1208C-65021	Trần Ngọc	Tấn	CTC901	Nam	2/1/1994		Đang học
19	1308C-65010	Phạm Văn	Thu	CTC1001	Nam	21/6/1994	1	Đang học
20	1008C-65012	Bùi Trung	Tiến	CTC701	Nam	9/12/1992		Đang học
21	1208C-65020	Lưu Phú	Toàn	CTC901	Nam	11/12/1993		Đang học
22	1008C-65007	Đỗ Văn	Trung	CTC701	Nam	30/3/1992		Đang học
23	1508C-65002	Đào Quang	Trung	CTC1201	Nam	5/3/1997	1	Đang học
24	1108C-65031	Nguyễn Ngọc	Tuấn	CTC801	Nam	15/4/1992		Đang học
25	1008C-65035	Nguyễn Văn	Tuấn	CTC701	Nam	28/9/1991		Đang học
26	1208C-65027	Bùi Duy	Tường	CTC901	Nam	5/11/1993		Đang học
27	1108C-65019	Hoàng Quốc	Việt	CTC801	Nam	15/1/1992		Đang học
28	1408C-65007	Phạm Văn	Vũ	CTC1101	Nam	9/2/1993	1	Đang học

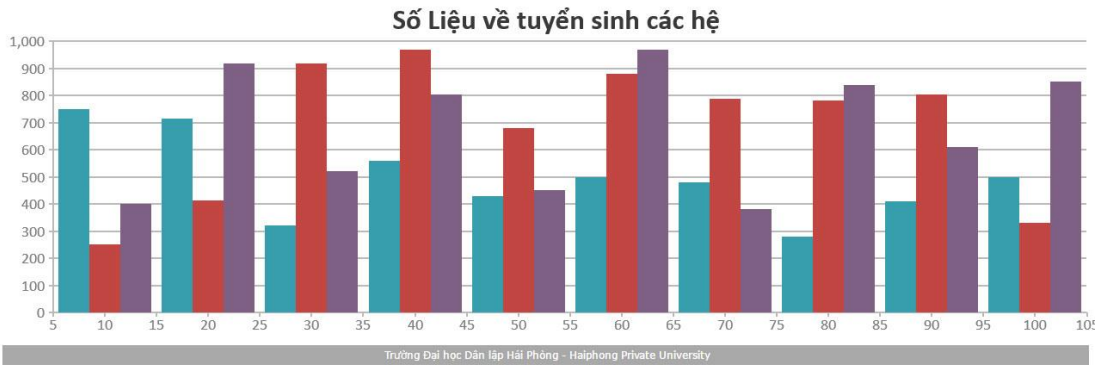
Trường Đại học Dân lập Hải Phòng - Haiphong Private University

Hình 3.21: Báo cáo Thống Kê Sinh Viên Tại Trường

Thứ hai, 24/10/2016 - 17:00:49 | Xin chào **Quản trị hệ thống**

TRƯỜNG CAO ĐẲNG SỐ 3 - BQP
QUẢN TRỊ HỆ THỐNG API-SERVICES

QUẢN TRỊ TÀI KHOẢN | QUẢN LÝ ĐÀO TẠO | QUẢN LÝ NHÂN SỰ | QUẢN LÝ CƠ SỞ VẬT CHẤT



Hình 3.22: Báo cáo Thống Kê Số Liệu Tuyển Sinh Qua Các Năm Đào Tạo

TRƯỜNG CAO ĐẲNG SỐ 3 - BQP
QUẢN TRỊ HỆ THỐNG API-SERVICES

QUẢN TRỊ TÀI KHOẢN
QUẢN LÝ ĐÀO TẠO
QUẢN LÝ NHÂN SỰ
QUẢN LÝ CƠ SỞ VẬT CHẤT

BẢNG TỔNG HỢP DANH MỤC VẬT TƯ

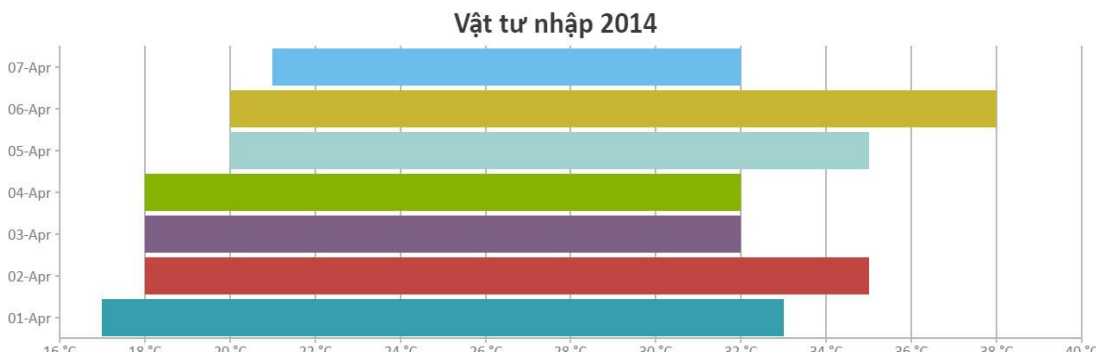
TT	Tên tài sản	Hiện trạng tài sản	Năm nhập, thu hồi	ĐVT	Số lượng	Đề xuất
1	Dây cứu hỏa	Hồng	2012	Cuộn	2	Tái sử dụng
2	H.thống bảng chuyển chiếu sáng máy	Hồng	1995	Bộ	1	Tái sử dụng, POTVT cho tháo xếp gọn (QĐ thanh lý số 7809/QĐ-BTC ngày 19/12/2012)
3	Đục nguội	Hồng	2012	Cái	271	Tái sử dụng TT
4	Chạm chặt	Hồng	2012	Cái	105	Tái sử dụng TT
5	Tấm rào sắt hồng các loại	Phế liệu	2012	kg	Cán thực tế	Tái sử dụng làm nhà xe SV
6	Thép cũ chữ U nhà ĐH 7 tầng	Phế liệu	2012	Kg	297	Tái sử dụng làm nhà xe SV
7	áo veston nữ		2012	Cái	14	Giao cho CB/Đoàn TNCSHCM làm từ thiện
8	áo đồng phục		2012	Cái	1237	Giao cho CB/Đoàn TNCSHCM làm từ thiện
9	áo BHLĐ	58 áo mẫu cũ	2012	Cái	316	Giao cho các xưởng, khoa phát cho HSSV
10	Quần BHLĐ		2012	Cái	382	Giao cho các xưởng, khoa phát HSSV

Trường Đại học Dân lập Hải Phòng - Haiphong Private University

Hình 3.23: Báo cáo Thống Kê Danh Mục Vật Tư

TRƯỜNG CAO ĐẲNG SỐ 3 - BQP
QUẢN TRỊ HỆ THỐNG API-SERVICES

QUẢN TRỊ TÀI KHOẢN
QUẢN LÝ ĐÀO TẠO
QUẢN LÝ NHÂN SỰ
QUẢN LÝ CƠ SỞ VẬT CHẤT



Hình 3.24: Báo cáo Thống Kê Vật Tư Nhập Theo Năm

TRƯỜNG CAO ĐẲNG SỐ 3 - BQP
QUẢN TRỊ HỆ THỐNG API-SERVICES

QUẢN TRỊ TÀI KHOẢN | QUẢN LÝ ĐÀO TẠO | QUẢN LÝ NHÂN SỰ | QUẢN LÝ CƠ SỞ VẬT CHẤT

Danh Sách Cán Bộ Giảng Viên Trường Cao Đẳng Nghề Số 3 - BQP

STT	Họ và tên	Chức vụ	Đơn vị	Ghi chú
69	Vũ Văn Hiệp	Giảng viên	BM Đầu máy toa xe	
70	Hoàng Thị Hồng Lê	Phó Trưởng khoa	Khoa KTVT	
71	Lê Thu Sao	Phó trưởng khoa	Khoa KTVT	
72	Phạm Trung Nghĩa	Quản lý HSSV	Khoa KTVT	
73	Hoàng Văn Lâm	P.Trưởng Bộ môn	BM VT sắt bộ	
74	Nguyễn Việt Thắng	P.Trưởng Bộ môn	BM Quản trị DN	
75	Phạm Đức Tấn	P.Trưởng Bộ môn	BM Kinh tế XD	
76	Trần Trung Kiên	P.Trưởng Bộ môn	BM Kinh tế XD	
77	Đỗ Thị Hồng Vân	Trưởng Bộ môn	BM CSN KT	
78	Lâm Phạm Thị Hải Hà	P.Trưởng Bộ môn	BM CSN KT	
79	Chu Thị Bích Hạnh	P.Trưởng Bộ môn	BM Kế toán KT	
80	Nguyễn Minh Nguyệt	P.Trưởng Bộ môn	BM Tài chính NH	
81	Trần Thái Minh	P.Trưởng khoa	Khoa KHCB	
82	Lưu Thị Thu Hà	Trưởng Bộ môn	BM Hóa học	
83	Trần Quốc Tuấn	Trưởng Bộ môn	BM Vật lý	

Trường Đại học Dân lập Hải Phòng - Haiphong Private University

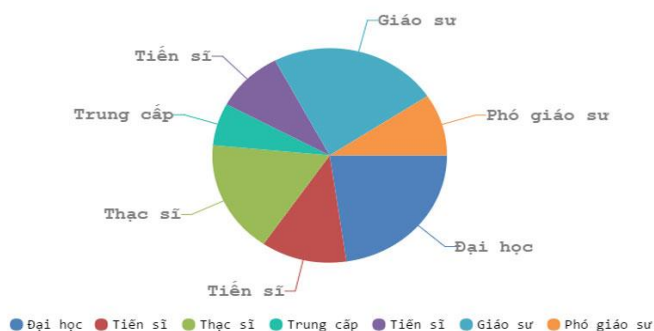
Hình 3.25: Báo cáo Thống Kê Danh Sách Cán Bộ Giảng Viên Trong Trường

Thứ hai, 24/10/2016 - 17:00:49 | Xin chào Quản trị hệ thống

TRƯỜNG CAO ĐẲNG SỐ 3 - BQP
QUẢN TRỊ HỆ THỐNG API-SERVICES

QUẢN TRỊ TÀI KHOẢN | QUẢN LÝ ĐÀO TẠO | QUẢN LÝ NHÂN SỰ | QUẢN LÝ CƠ SỞ VẬT CHẤT

Danh sách trình độ nhân sự



Trường Đại học Dân lập Hải Phòng - Haiphong Private University

Hình 3.26: Báo cáo Thống Kê Trình Độ Nhân Sự Trong Trường

Chương 4: TRIỂN KHAI THỬ NGHIỆM HỆ THỐNG

4.1 Kho dữ liệu thử nghiệm

Cài đặt hệ quản trị SQL Server Express. SQL Server là viết tắt của Structure Query Language, nó là một công cụ quản lý dữ liệu được sử dụng phổ biến ở nhiều lĩnh vực. Hầu hết các ngôn ngữ bậc cao đều có trình hỗ trợ SQL như Visual BASIC, Oracle, Visual C...

Để download cài đặt SQL Server Expresss truy cập qua địa chỉ:

<https://www.microsoft.com/en-us/sql-server/sql-server-editions-express>

Cài đặt IIS (IIS là viết tắt của từ Internet Information Services) là các dịch vụ dành cho máy chủ chạy trên nền Hệ điều hành Window nhằm cung cấp và phân tán các thông tin lên mạng, nó bao gồm nhiều dịch vụ khác nhau như Web Server, FTP Server... Trong đề tài tôi sử dụng phiên bản IIS 7.0

Được được download và cài đặt :

<https://www.microsoft.com/en-us/download/details.aspx?id=2299>

4.2 Cơ sở dữ liệu của một số đơn vị và công cụ trích rút tương ứng

Sử dụng ngôn ngữ lập trình PHP, library **NuSOAP** , MySQL SerVer: Xây dựng ứng dụng phía client và sau đó gọi thực hiện dịch vụ thông qua việc kết nối tới SOAP server.

Trong khuôn khổ của đề tài với 3 chương trình trên tôi dùng gói dịch vụ Xampp là chương trình tạo máy chủ Web (WebServer) được tích hợp sẵn Apache, PHP, MySQL, FTP Server, Mail Server và các công cụ như phpMyAdmin Phiên bản được dùng là XAMPP for Windows 5.2.38, 5.6.28 & 7.0.13

Được download và cài đặt tại: <https://www.apachefriends.org/download.html>

SOAP (viết tắt của Simple Object Access Protocol) là một trong những giao thức "chuẩn" dựa trên nền tảng XML để làm việc với Web Services. Với các dịch vụ được tích hợp

- SOAP 1.1 (<http://www.w3.org/TR/2000/NOTE-SOAP-20000508/>)

- WSDL 1.1 (<http://www.w3.org/TR/2001/NOTE-wsdl-20010315>)
 - SOAP Messages With Attachments (<http://www.w3.org/TR/SOAP-attachments>)
 - XML 1.0 (<http://www.w3.org/TR/2006/REC-xml-20060816/>)
 - Namespaces in XML 1.0 (<http://www.w3.org/TR/2006/REC-xml-names-20060816/>)
 - XML Schema 1.0 (<http://www.w3.org/TR/xmlschema-0/>)
 - RFC 2045 Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies
 - RFC 2068 Hypertext Transfer Protocol -- HTTP/1.1
 - RFC 2617 HTTP Authentication: Basic and Digest Access Authentication
- Thư viện được dowload và cài đặt: <http://sourceforge.net/projects/nusoap/>*

4.3 Một số báo cáo thử nghiệm thực hiện với kho dữ liệu

4.3.1 Một số báo cáo trong quản lý nhân sự

The screenshot shows a web application interface for HR management. The page title is "TRƯỜNG CAO ĐẲNG SỐ 3 - BQP" and "QUẢN TRỊ HỆ THỐNG API-SERVICES". The interface includes a search form with fields for "Phong Ban", "Chức Vụ", "Hệ Số Lương", "Create Date", "Modified Date", and "Trạng thái". Below the search form is a table of staff members with columns for "Mã Nhân Viên (*)", "Họ Tên (*)", "Giới Tính (*)", "Ngày sinh", "Phong Ban", "Chức Vụ", "Lương cơ bản", "Hệ Số Lương", "Thâm Niên", "Create Date", "Modified Date", "Trạng thái", "User Add", and "Add Time".

	Mã Nhân Viên (*)	Họ Tên (*)	Giới Tính (*)	Ngày sinh	Phong Ban	Chức Vụ	Lương cơ bản	Hệ Số Lương	Thâm Niên	Create Date	Modified Date	Trạng thái	User Add	Add Time	
Xem	000001	Nguyễn Văn An	Nam	01/01/1976	Ban Giám đốc	Giám đốc	1210000	5	10	12/10/2016	12/10/2016	Active	1	15/12/20	
Xem	Sửa	000002	Trần Hồng Quân	Nam	01/01/1977	Tổ chức nhân sự	Trưởng phòng	1210000	4	5	12/10/2016	12/10/2016	Active	1	15/12/20
Xem	Sửa	000003	Bùi Sỹ Thành	Nam	01/01/1977	Hành chính tổng hợp	Trưởng phòng	1210000	4	7	12/10/2016	12/10/2016	Active	1	15/12/20
Xem	Sửa	000004	Hồ Sỹ Đàm	Nam	01/01/1977	Bảo vệ	Trưởng phòng	1210000	3	9	12/10/2016	12/10/2016	Active	1	15/12/20

Hình 4.1: Báo cáo danh sách cán bộ giáo viên

4.3.2 Một số báo cáo trong quản lý đào tạo

Table: Báo Cáo Danh Sách Sinh Viên Tốt Nghiệp

Ngành	Lớp	Năm Học	Học Kỳ	Mã sinh viên (*)	Họ Tên (*)	Ngày sinh	Điểm TB TN	Create Date	Modified Date	User Add	Add Time	User Edit	Edit Time	Trạng thái
Xem	Sửa	101	CT2001	2015-2016	1516HK2	11101003	Hoàng Văn Hạnh	02/03/1997	6	22/08/2016 15:11:21	22/08/2016 15:11:21	1	20/12/2016	Active
Xem	Sửa	101	CT2001	2015-2016	1516HK2	11101004	Đoàn Hoa Vinh	26/09/1995	7	22/08/2016 15:11:22	22/08/2016 15:11:22	1	20/12/2016	Active
Xem	Sửa	101	CT2001	2015-2016	1516HK2	11101005	Cù Thế Vinh	07/11/1997	8	22/08/2016 15:11:22	22/08/2016 15:11:22	1	20/12/2016	Active

Hình 4.2: Báo cáo kết quả học sinh tốt nghiệp

4.3.3 Một số báo cáo trong quản lý trang thiết bị

Table: Báo Cáo Nhập Thiết Bị

Nhập TB id	Ngày Nhập	Mã phiếu	Lý do nhập (*)	Số Lượng	Đơn giá	User Add	Add Time	User Edit	Edit Time	Create Date	Modified Date	Trạng thái
Sửa	14	12/10/2016	N001	Xe tải 5 tấn theo yêu cầu mua sắm	10	500000000	1	20/12/2016		12/10/2016 14:36:53	12/10/2016 14:36:53	Active
Sửa	15	12/10/2016	N002	Xe tải 10 tấn theo yêu cầu mua sắm	10	700000000	1	20/12/2016		12/10/2016 14:36:53	12/10/2016 14:36:53	Active
Sửa	16	12/10/2016	N003	Xe con 4 chỗ theo yêu cầu mua sắm	10	500000000	1	20/12/2016		12/10/2016 14:36:53	12/10/2016 14:36:53	Active
Sửa	19	12/10/2016	N004	Xe con 7 chỗ theo yêu cầu mua sắm	50	800000000	1	20/12/2016	1	20/12/2016 00:00:00	19/10/2016 00:00:00	Lock

Hình 4.3: Báo cáo cơ sở vật chất trang thiết bị dạy nghề

KẾT LUẬN

Sau một thời gian nghiên cứu và hoàn thành luận văn, tôi đã thực hiện được mục tiêu đề ra. Các kết quả đạt được bao gồm:

1. Kết quả đạt được:

Đề tài đã đạt được những yêu cầu đề ra về mặt lý thuyết cũng như ứng dụng trong thực tiễn.

Về mặt lý thuyết: đề tài đã trình bày những khái niệm về kho dữ liệu Data Warehouse và Web service, cách tổ chức khai thác kho dữ liệu phục vụ cho việc báo cáo thống kê của Nhà trường.

Về mặt thực tiễn: đề tài đã xây dựng được kho dữ liệu Báo cáo thống kê tại trường Cao đẳng nghề số 3 –BQP.

2. Phạm vi áp dụng:

Mặc dù đối tượng nghiên cứu là trường Cao đẳng nghề số 3-BQP. Nhưng đề tài có thể áp dụng cho các trường cao đẳng và trung cấp nghề khác.

3. Hướng phát triển

Xây dựng kho dữ liệu phân tán

Hệ thống hỗ trợ cho việc quản lý, điều hành ra quyết định

TÀI LIỆU THAM KHẢO

Tiếng Việt:

1. Vũ Đức Thi, Lê Hải Khôi (1999). “*Một số nguyên lý hoạt động của kho dữ liệu*”, Tạp chí: Tin học và điều khiển học, 2:15, tr 27-32.
2. Nguyễn Văn Vy (2010), “*Phân tích và thiết kế các hệ thống thông tin hiện đại. Hướng cấu trúc và hướng đối tượng*”, Nhà xuất bản Thống kê, Hà Nội.
3. Lê Văn Phùng (2014), “*Hệ Thống Thông Tin Quản Lý*”, Nhà xuất bản Thông Tin Truyền Thông
4. Đoàn Văn Ban. “*Phương pháp thiết kế và khai thác kho dữ liệu.*” Đề tài cấp trung tâm KHTN & CNQG 1997, n.d”.
5. Hà, Hồ Cẩm. “*Thiết kế kho dữ liệu phục vụ công tác đào tạo ở trường Sư Phạm.*” Báo cáo đề tài khoa học cấp bộ 2007, n.d”.
6. Lê Văn Phùng (2014), “*Các Mô Hình Cơ Bản Trong Phân Tích Và Thiết Kế Hướng Đối Tượng*” . Nhà xuất bản Thông Tin Truyền Thông

Tiếng Anh:

7. W. H. Inmon (2002), *Building the Data Warehouse*, Wiley Computer Publishing.
8. Thomas Erl. (2008), *SOA: principles of service design*. Pearson Education, Inc.
9. Barry. “Commercial Data Mining (1997). Processing, Analysis and Modeling for Predictive.” n.d.
10. B Gray, N. A. (2004). Công nghệ Java RMI. n.d.
11. Dcom: Microsoft Distributed Component Object Model Paperback – September, 1997
12. Benson a, Smith S.J. “Data Warehousing, Data Mining Application For CRM, 1997.” n.d.
13. “CORBA Technology and the Java™ Platform Standard Edition.” (Oracle). n.d.
14. ERL, Thomas. “Service Orientecture Architecture, 2015.” n.d.
15. V.Poe, Prentice. *Buiding a Data Warehouse for Decision Support, 1996*, n.d.
16. Carl-Fredrik, (2001). “*DIF8901 Object-Oriented Systems A Comparison of Distributed Object Technologies.*” n.d.

A. Phụ lục 1: Các bảng dữ liệu nguồn

a. Các bảng dữ liệu về nhân sự

Bảng 3.a1. Hồ sơ nhân viên

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaNV	nvarchar	5	Mã nhân viên-khóa chính
HoTen	nvarchar	30	Họ tên
Gioitinh	nvarchar	5	Giới tính
Ngaysinh	Date/Time		Ngày sinh
Noisinh	nvarchar	50	Nơi sinh
QueQuan	nvarchar	50	Quê quán
DanToc	nvarchar	20	Dân tộc
<u>MaTN</u>	nvarchar	5	Mã thâm niên
<u>MaCV</u>	nvarchar	5	Mã chức vụ
<u>MaCB</u>	nvarchar	5	Mã cấp bậc
NgayChuyenCT	Date/Time		Ngày chuyển công tác
QuaTrinhCT	Memo	10	Quá trình công tác
<u>MaPB</u>	nvarchar	5	Mã phòng ban
<u>MaTDCM</u>	nvarchar	5	Mã trình độ chuyên môn
GiaDinh	Yes/No		Gia đình
Socon	int	2	Số con
SoCMT	nvarchar	11	Số chứng minh thư
Ngaycap	Date/Time		Ngày cấp
Noicap	nvarchar	50	Nơi cấp
Diachi	nvarchar	50	Địa chỉ
SoDT	int	13	Số điện thoại
DangVien	Yes/No		Đảng viên
SoTheDang	nvarchar	10	Số thẻ đảng
NgayVaoDang	Date/Time		Ngày vào đảng
NoiVaoDang	nvarchar	50	Nơi vào đảng
<u>MaTDNN</u>	nvarchar	5	Mã trình độ ngoại ngữ
<u>MaTDTH</u>	nvarchar	5	Mã trình độ tin học
HocHam	nvarchar	15	Học hàm
NgayNhanHH	Date/Time		Ngày nhận học hàm
<u>MaLuong</u>	nvarchar	5	Mã lương
<u>MasoHD</u>	nvarchar	5	Mã số hợp đồng
LoaiHD	nvarchar	50	Loại hợp đồng
NgayHD	Date/Time		Ngày hợp đồng
NgayBD	Date/Time		Ngày bắt đầu

NgàyKT	Date/Time		Ngày kết thúc
NgàyXL	Date/Time		Ngày xếp lương
SoBHXH	nvarchar	9	Số bảo hiểm xã hội
SoBHYT	nvarchar	9	Số bảo hiểm y tế
LuongTL	int	10	Lương tháng lĩnh
<u>MaKhenThuong</u>	nvarchar	5	Mã khen thưởng
<u>MaKyLuat</u>	nvarchar	5	Mã kỷ luật
HinhAnh	Picture		Hình ảnh
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

Bảng 3.a2. Thâm niên

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaTN	nvarchar	5	Mã thâm niên
Thamnien	nvarchar	50	Thâm niên
CreateDate	Date/time		Ngày khởi tạo
ModifiedDate	Date/time		Ngày cập nhật

Bảng 3.a3. Chức vụ

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaChucVu	nvarchar	5	Mã chức vụ
TenChucVu	nvarchar	50	Tên chức vụ
CreateDate	Date/time		Ngày khởi tạo
ModifiedDate	Date/time		Ngày cập nhật

Bảng 3.a4. Cấp bậc

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaCB	nvarchar	5	Mã cấp bậc
LoaiCB	nvarchar	50	Loại cấp bậc
CreateDate	Date/time		Ngày khởi tạo
ModifiedDate	Date/time		Ngày cập nhật

Bảng 3.a5. Phòng ban

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaPB	nvarchar	5	Mã phòng ban
TenPhongBan	nvarchar	50	Tên phòng ban
CreateDate	Date/time		Ngày khởi tạo
ModifiedDate	Date/time		Ngày cập nhật

Bảng 3.a6. Trình độ chuyên môn

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaTDCM	nvarchar	5	Mã trình độ chuyên môn
NganhDT	nvarchar	50	Ngành đào tạo
LoaiHinhDT	nvarchar	50	Loại hình đào tạo
TruongDT	nvarchar	50	Trường đào tạo
NienKhoa	nvarchar	10	Niên khóa
CreateDate	Date/time		Ngày khởi tạo
ModifiedDate	Date/time		Ngày cập nhật

Bảng 3.a7. Trình độ ngoại ngữ

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaTDNN	nvarchar	5	Mã trình độ ngoại ngữ
TenNN	nvarchar	20	Tên ngoại ngữ
LoaiTD	nvarchar	50	Loại trình độ ngoại ngữ
CreateDate	Date/time		Ngày khởi tạo
ModifiedDate	Date/time		Ngày cập nhật

Bảng 3.a8. Trình độ tin học

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaTDTH	nvarchar	5	Mã trình độ tin học
LoaiTDTH	nvarchar	50	Loại trình độ tin học
CreateDate	Date/time		Ngày khởi tạo
ModifiedDate	Date/time		Ngày cập nhật

Bảng 3.a9. Lương

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaLuong	nvarchar	5	Mã lương
LuongCB	int	10	Lương cơ bản
HesoLuong	int	2	Hệ số lương
CreateDate	Date/time		Ngày khởi tạo
ModifiedDate	Date/time		Ngày cập nhật

Bảng 3.a10. Hợp đồng lao động

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MasoHD	nvarchar	5	Mã số hợp đồng
NgayHD	Date/time		Ngày hợp đồng
LoaiHD	nvarchar	50	Loại hợp đồng
NgayKT	Date/time		Ngày kết thúc
NgayXL	Date/time		Ngày xếp lương

SoBHXH	nvarchar	9	Số bảo hiểm xã hội
SoBHYT	nvarchar	9	Số bảo hiểm y tế
CreateDate	Date/time		Ngày khởi tạo
ModifiedDate	Date/time		Ngày cập nhật

Bảng 3.a11. Khen thưởng

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaKhenThuong	nvarchar	5	Mã khen thưởng
LyDoKhenthuong	nvarchar	150	Lý do khen thưởng
NgayKhenThuong	Date/time		Ngày khen thưởng
CreateDate	Date/time		Ngày khởi tạo
ModifiedDate	Date/time		Ngày cập nhật

Bảng 3.a12. Kỷ luật

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaKyLuat	nvarchar	5	Mã kỷ luật
LyDoKyLuat	nvarchar	150	Lý do kỷ luật
NgayKyLuat	Date/time		Ngày kỷ luật
HinhThucKT	nvarchar	150	Hình thức kỷ luật
CreateDate	Date/time		Ngày khởi tạo
ModifiedDate	Date/time		Ngày cập nhật

b. Các bảng dữ liệu quản lý đào tạo

Bảng 3.b1. Hồ sơ sinh viên

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaSinhvien	nvarchar	8	Mã sinh viên
HoDem	nvarchar	30	Họ tên
Ten	nvarchar	10	Tên
Gioitinh	nvarchar	5	Giới tính
Ngaysinh	Date/Time		Ngày sinh
Noisinh	nvarchar	50	Nơi sinh
QueQuan	nvarchar	50	Quê quán
DanToc	nvarchar	20	Dân tộc
Diachi	nvarchar	50	Địa chỉ
<u>MaHe</u>	nvarchar	2	Mã hệ (CD,TC,SC)
<u>MaNganh</u>	nvarchar	10	Mã ngành học
<u>MaLop</u>	nvarchar	9	Mã lớp
<u>MaDoiTuong</u>	nvarchar	9	Mã đối tượng(BDXN,Dân sự)

SoCMT	nvarchar	11	Số chứng minh thư
Ngaycap	Date/Time		Ngày cấp
Noicap	nvarchar	50	Nơi cấp
SoThe	nvarchar	10	Số thẻ học nghề
Ngaycap	Date/Time		Ngày cấp thẻ học nghề
Noicap	nvarchar	50	Nơi
SoDT	int	13	Số điện thoại
DangVien	Yes/No		Đảng viên
SoTheDang	nvarchar	10	Số thẻ đảng
NgayVaoDang	Date/Time		Ngày vào đảng
NoiVaoDang	nvarchar	50	Nơi vào đảng
HinhAnh	Picture		Hình ảnh
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

Bảng 3.b2. Hệ đào tạo

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaHe	nvarchar	2	Mã hệ
TenHe	nvarchar	30	Tên hệ
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

Bảng 3.b3. Ngành đào tạo

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaNganh	nvarchar	10	Mã ngành học
TenNganh	nvarchar	30	Tên ngành học
MaHe	nvarchar	2	Mã hệ
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

Bảng 3.b4. Lớp học

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaLop	nvarchar	9	Mã lớp
TenLop	nvarchar	20	Tên lớp
MaGV	nvarchar	9	Mã giáo viên
MaNganh	nvarchar	10	Mã ngành học
SiSo	int	2	Sĩ số lớp
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

Bảng 3.b5. Điểm học kỳ

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
DiemTBC	Float	8	Điểm trung bình cộng
PhanLoaiRL	nvarchar	1	Phân loại rèn luyện
MaHocKy	nvarchar	2	Mã học kỳ(1,2..)
MaSinhvien	nvarchar	8	Mã sinh viên
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

Bảng 3.b6. Học kỳ

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaHocKy	nvarchar	2	Mã học kỳ(1,2..)
TenHocKy	nvarchar	20	Tên học kỳ
MaNamHoc	nvarchar	2	Mã năm học
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

Bảng 3.b7. Năm học

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaNamHoc	nvarchar	2	Mã năm học
TenNamHoc	nvarchar	20	Tên năm học
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

Bảng 3.b8. Giáo viên chủ nhiệm

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaGV	nvarchar	9	Mã giáo viên
TenGiaovien	nvarchar	50	Tên giáo viên
DienThoai	int	12	Điện thoại liên hệ
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

Bảng 3.b9. Kết quả thi tốt nghiệp

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaSinhvien	nvarchar	8	Mã sinh viên
DiemThiTN	Float	8	Điểm thi tốt nghiệp
MaMon	nvarchar	1	Mã môn thi tốt nghiệp
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

Bảng 3.b10. Môn thi tốt nghiệp

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
------------	--------------	--------	---------

MaMon	nvarchar	1	Mã môn thi tốt nghiệp
TenMon	nvarchar	20	Tên môn thi tốt nghiệp
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

c. Các bảng dữ liệu quản lý vật tư, trang bị kỹ thuật, phương tiện

Bảng 3.c1. Trang thiết bị vật tư

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaThietBi	nvarchar	10	Mã thiết bị
TenThietBi	nvarchar	50	Tên thiết bị
DonVi	nvarchar	10	Đơn vị
NhaSX	nvarchar	50	Nhà sản xuất
NamSX	Date/Time		Năm sản xuất
NamSD	Date/Time		Năm sử dụng
<u>MaNhomTB</u>	nvarchar	10	Mã nhóm trang bị
<u>MaPhongBan</u>	nvarchar	10	Mã đơn vị quản lý
<u>MaPhuongTien</u>	nvarchar	10	Mã phương tiện
MaChungLoai	nvarchar	10	Mã chủng loại thiết bị
SoLuongTT	int	8	Số lượng thực tế
SoLuongTang	int	8	Số lượng tăng
LyDotang	nvarchar	50	Lý do tăng
SoLuongGiam	int	8	Số lượng giảm
LyDoGiam	nvarchar	50	Lý do giảm
PhanCapCL	nvarchar	1	Phân cấp chất lượng
MaPhieuNhapTB	nvarchar	10	Mã phiếu nhập thiết bị
MaPhieuXuatTB	nvarchar	10	Mã phiếu xuất thiết bị
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

Bảng 3.c2. Mã nhóm trang thiết bị

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaNhomTB	nvarchar	10	Mã nhóm trang bị
LoaiNhomTB	int	1	Tên thiết bị
MaPhongBan	nvarchar	10	Mã đơn vị quản lý
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

Bảng 3.c3: Phòng ban đơn vị quản lý

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
------------	--------------	--------	---------

MaPhongBan	nvarchar	10	Mã đơn vị quản lý
TenPhongBan	nvarchar	50	Tên phòng ban đơn vị quản lý
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

Bảng 3.c4. Chủng loại trang thiết bị

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaChungLoai	nvarchar	10	Mã chủng loại thiết bị
TenChungLoai	nvarchar	50	Tên chủng loại thiết bị
MaPhongBan	nvarchar	10	Mã đơn vị quản lý
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

Bảng 3.c5. Phiếu nhập trang thiết bị

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaPhieuNhapTB	nvarchar	10	Mã phiếu nhập thiết bị
LydoNhap	nvarchar	150	Lý do nhập
SoLuong	int	8	Số lượng nhập
DonGia	Float	18	Đơn giá
TongTien	Float	18	Tổng tiền
NgayNhap	Date/Time		Ngày nhập

Bảng 3.c6. Phiếu xuất trang thiết bị

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaPhieuXuatTB	nvarchar	10	Mã phiếu xuất thiết bị
LydoXuat	nvarchar	50	Lý do xuất
SoLuong	int	8	Số lượng xuất
DonGia	Float	18	Đơn giá
TongTien	Float	18	Tổng tiền
NgayXuat	Date/Time		Ngày xuất

Bảng 3.c7. Phương tiện ô tô xe máy

Tên trường	Kiểu dữ liệu	Độ lớn	Miêu tả
MaPhuongTien	nvarchar	10	Mã phương tiện
LoaiXe	nvarchar	50	Loại xe
SoXe	int	10	Số xe
HangXe	nvarchar	50	Hãng xe
NhaSX	nvarchar	50	Nhà sản xuất
NamSX	Date/Time		Năm sản xuất
NamSD	Date/Time		Năm sử dụng
ThoiGianDK	Date/Time		Thời gian đăng ký
SoKhung	int	10	Số khung
SoMay	int	10	Số máy

GiayChungNhanDK	nvarchar	50	Giấy chung nhận đăng ký
DoiTuongSD	nvarchar	50	Mã nhóm trang bị
MaPhongBan	nvarchar	10	Mã đơn vị quản lý
CreateDate	Date/Time		Ngày khởi tạo
ModifiedDate	Date/Time		Ngày cập nhật

B. Phụ lục 2: Các đoạn chương trình

a. Đoạn COD 1

```

using System;
    using System.Data;
    using System.Data.SqlClient;
    using System.Web;
    using System.Web.Services;
    using System.Web.Services.Protocols;
    namespace Service_QLDT
    {
        /// <summary>
        /// Summary description for QuanLyDaoTao
        /// </summary>
        [WebService(Namespace = "http://tempuri.org/")]
        [WebServiceBinding(ConformsTo = WsiProfiles.BasicProfile1_1)]
        [System.ComponentModel.ToolboxItem(false)]
        // To allow this Web Service to be called from script, using ASP.NET AJAX, uncomment the
        // following line.
        // [System.Web.Script.Services.ScriptService]
        public class QuanLyDaoTao : System.Web.Services.WebService
        {
            private SqlConnection getConnection()
            {
                SqlConnection conn = new
                SqlConnection("Server=10.1.0.234;Database=QLDT;UID=qldt;PWD=123654");
                return conn;
            }
            //Lấy danh sách sinh viên
            [WebMethod(Description = "Danh sách sinh viên: DataTable Sinhvien()");
            Output(MaSinhVien, HoDem, Ten, GioiTinh, NgaySinh, NoiSinh, QueQuan, DanToc, DiaChi, Malo
            p, MaHe, TenNganh, MaDoiTuong, SoCMT, CreateDate")]
            public DataTable Sinhvien()
            {
                SqlConnection conn = getConnection();
                DataTable dt = new DataTable("Sinhvien");
                //string sql;
                SqlCommand cmd = new SqlCommand("uspSinhVien", conn);
                SqlDataAdapter da = new SqlDataAdapter(cmd);
                da.Fill(dt);
                conn.Close();
                return dt;
            }
            // Lấy danh sách sinh viên tốt nghiệp
            [WebMethod(Description = "Danh sách sinh viên tốt nghiệp: DataTable SinhvienTN()");
            Output(MaSinhVien, HoDem, Ten, GioiTinh, NgaySinh, QueQuan, Malop, MaHe,
            MaNganh, MaNamHoc, MaHocKy, DiemTBTN")]
            public DataTable SinhvienTN()
            {

```

```

        SqlConnection conn = getConnection();
        DataTable dt = new DataTable("SinhvienTN");
//string sql;
SqlCommand cmd = new SqlCommand("uspSinhVienTN", conn);
SqlDataAdapter da = new SqlDataAdapter(cmd);
da.Fill(dt);
conn.Close();
return dt;
}
}
}

```

b. Đoạn COD 2

```

using System;
using System.Data;
using System.Data.SqlClient;
using System.Web;
using System.Web.Services;
using System.Web.Services.Protocols;
namespace QLNS
{
    /// <summary>
    /// Summary description for Service_QLNS
    /// </summary>
    [WebService(Namespace = "http://tempuri.org/")]
    [WebServiceBinding(ConformsTo = WsiProfiles.BasicProfile1_1)]
    [System.ComponentModel.ToolboxItem(false)]
    // To allow this Web Service to be called from script, using ASP.NET AJAX, uncomment the
    // following line.
    // [System.Web.Script.Services.ScriptService]
    public class Service_QLNS : System.Web.Services.WebService
    {
        [WebMethod]
        private SqlConnection getConnection()
        {
            SqlConnection conn = new
                SqlConnection("Server=10.1.0.234;Database=QLNS;UID=qlns;PWD=123654");
            return conn;
        }
        //Lấy danh sách nhân viên
        [WebMethod(Description = "Danh sách nhân viên: DataTable NhanVien(); Output(MaNV, HoTen,
            GioiTinh, Ngaysinh, TenPhongBan, TenChucVu, LuongCB, HeSoLuong, ThamNien,
            CreateDate, ModifiedDate)")]
        public DataTable Nhanvien()
        {
            SqlConnection conn = getConnection();
            DataTable dt = new DataTable("Nhanvien");
//string sql;
SqlCommand cmd = new SqlCommand("uspNhanvien", conn);
SqlDataAdapter da = new SqlDataAdapter(cmd);
da.Fill(dt);
conn.Close();
return dt;
        }
    }
}

```

c. Đoạn COD 3

```
using System;
using System.Data;
using System.Data.SqlClient;
using System.Web;
using System.Web.Services;
using System.Web.Services.Protocols;
namespace QLTB
{
    /// <summary>
    /// Summary description for Service_QLTB
    /// </summary>
    [WebService(Namespace = "http://tempuri.org/")]
    [WebServiceBinding(ConformsTo = WsiProfiles.BasicProfile1_1)]
    [System.ComponentModel.ToolboxItem(false)]
    // To allow this Web Service to be called from script, using ASP.NET AJAX, uncomment the
    following line.
    // [System.Web.Script.Services.ScriptService]
    public class Service_QLTB : System.Web.Services.WebService
    {
        private SqlConnection getConnection()
        {
            SqlConnection conn = new
            SqlConnection("Server=10.1.0.234;Database=QLTB;UID=qltb;PWD=123654");
            return conn;
        }
        //Lấy danh sách thiết bị nhập
        [WebMethod(Description = "Danh sách thiết bị nhập: DataTable ThietBiNhap();
        Output(MaPhieuNhapTB, LydoNhap, SoLuong, DonGia, CreateDate, ModifiedDate)")]
        public DataTable ThietBiNhap()
        {
            SqlConnection conn = getConnection();
            DataTable dt = new DataTable("ThietBiNhap");
            //string sql;
            SqlCommand cmd = new SqlCommand("uspThietBiNhap", conn);
            SqlDataAdapter da = new SqlDataAdapter(cmd);
            da.Fill(dt);
            conn.Close();
            return dt;
        }
        //Lấy danh sách thiết bị xuất
        [WebMethod(Description = "Danh sách thiết bị xuất: DataTable ThietBiXuat();
        Output(MaPhieuXuatTB, LydoXuat, SoLuong, DonGia, CreateDate, ModifiedDate)")]
        public DataTable ThietBiXuat()
        {
            SqlConnection conn = getConnection();
            DataTable dt = new DataTable("ThietBiXuat");
            //string sql;
            SqlCommand cmd = new SqlCommand("uspThietBiXuat", conn);
            SqlDataAdapter da = new SqlDataAdapter(cmd);
            da.Fill(dt);
            conn.Close();
            return dt;
        }
    }
}
```