資料倉儲規劃與建置--以成衣業為例

Data Warehousing Design and Construction – the Case Study for a Garment Company

Shu-Fen Lin　　　　　　**Huei-huang Chen**
Department of Computer Science and Engineering, Tatung University
Lin.emmy@mas.hinet.net　　　hhchen@ttu.edu.tw

摘要

　　本研究嘗試建構傳統成衣產業之資料倉儲雛型，以提供決策者快速查詢相關統計資料為基本需求，整理出從企業資源規劃系統資料庫中萃取之資料，建構出相關多維度的資料模型。運用資料倉儲及線上分析處理技術，決策者可依其需求，彈性且快速地以不同角度查詢相關統計資料，以提昇決策之品質及時效。

　　配合企業資源規劃系統，基於重要性資料優先之考量，本研究採用之建置方法為由下而上(bottom-up)先建置個別之資料超市後，再整合各個資料超市成為資料倉儲。依使用者需求整理出所需要的資料型式，再自 IBM AS/400 企業資源規劃系統中將所需之資料擷取出來，轉換後儲存至建置於 SQL Server 資料庫中之儲存區後，再利用線上分析工具 Analysis services 提供決策者線上即時查詢功能。

關鍵詞：資料倉儲建置, 線上分析處理

ABSTRACT

　　This research attempts to design and construct a data warehouse fledgling model for a garment company based on the ERP (Enterprise Resource Planning) system to provide basic requirement for rapidly inquiring related statistic data for decision makers. By applying the technology of Data Warehouse and OLAP, and extracting data from ERP system's database, establishing a related multi-dimensional data model, decision makers can inquire related statistic data elastically and rapidly depending on their requirements and hence enhancing the quality and time effectiveness of decision making.

　　With the policy of implementing the most important data/information first consideration, we use bottom-up approach to implement the data warehousing system. We gathered the users'

requirements to define the needed metrics and the architecture of the data warehouse: data acquisition from the ERP system on IBM AS/400 DB2, data storage and information delivery on a SQL Server platform.

Key words：Data warehousing, OLAP

1. INTRODUCTION

1.1 Background

　　Rapid access to the right information is crucial for thriving in the complex and turbulent environment of the 21st century. A good measurement system is key to any competitive management system. Performance measurements are important for effective business planning and follow-up, improving business processes, motivating individuals and teams, organizational learning and improved decision-making. Trends and business characteristics, such as process orientation, supply chain collaboration, customer relationship management and e-business, drive the renewal of an organization's performance measurement system. In recent years, data warehouse has emerged as a powerful technology for integrating sparsely distributed operational data into a comprehensive analytical fashion to enable decision-making. The company that we study is the largest OEM jeans producer in the world with annual output of 40 million pairs, and is also the sixth largest denim producer at 60 million yards per year. The company has also expanded into casual wear, producing 10 million pairs of pants per year. Along with local production facilities, they have also expanded abroad include three garment factories in Africa, five garment factories and a dyeing house in Central America, and one denim mill and garment factory in Mexico. Global facilities make them the largest specialized denim fabric and garment in-one-stream.

　　To catch up with the market trend of "high quality, low price, quick response", this

company vertically integrates from spinning, dying, weaving till garment finishing to generate streamline benefit. They have been implementing ERP (Enterprise Resource Planning) since 2004 and centralizing logistic system to generate resource allocation efficiency and vendor gathering effect.

1.2 Research Motivation and Objective

For precisely capturing the future global trend of fashion and effectively improving the quality, shortening the delivery and saving the cost, the executives and managers who are responsible for keeping the enterprise competitive need information to make proper decisions. In order to make long-term managerial decisions, they have to exploit very large volumes of data, generally stored in their operational databases. The exploitation of these data is sometimes carried out in an empirical way using traditional techniques (SQL queries, views), and the decision-making process is often tiresome and does not properly respond to user requests. The data warehousing can overcome these insufficiencies.

This research aims to design and implement an online multidimensional analysis system that base on existed ERP system in the company context. Because of important data/information first consideration, we use bottom-up approach to implement system. We gathered the users' requirements to define the needed metrics the architecture of the data warehouse: data acquisition from IBM AS/400 DB2, data storage on SQL Server and information delivery for online inquiry.

2. Related Works

2.1 Basic Data Warehouse Concepts

A data warehouse is a subject-oriented, integrated, time-variant, and nonvolatile collection of data that supports managerial decision-making [1]. A data warehouse copies its data from existing systems like order entry, general ledger, human resources to form a physically separate store of data in contrast to the application data found in the operational environment.

Data warehouse is not a brand new technology, new platform, or new product. It is just a frame, strategy, or solution to construct an integrated information infrastructure based on the traditional database management system. The very essence of a data warehouse is to provide a solid platform of integrated, historical data from which informational, analytical processing and analysis over a long historical time perspective can be done. It generally provides the facility for

integration in a world of isolated application systems. There is indeed a world of promise in building and maintaining a data warehouse as the center of the architecture for information systems.

Data warehousing is the process to cleanse, filter, transform, summarize, and aggregate data generated in the operational environment and then put the data in a specially structured data warehouse for easy access and analysis by a disparate group of users [2].

2.2 Data Mart

Data mart is a subset of the overall enterprise data warehouse. It has all the characteristics of a data warehouse. The only difference between the data warehouse and the data mart is that the data in the data mart is organized according to a given department of the enterprise or a particular subject. Generally it is considered that the data of the data mart comes from the data warehouse [3].

2.3 Technical Architecture

The architecture provides the overall framework for developing and deploying the data warehouse; it is a comprehensive blueprint. Figure2.2. [2] groups the components into the three major areas of data acquisition, data storage and information delivery. The technical architecture of a data warehouse is, therefore, the complete set of functions and services provided within its components. The technical architecture also includes the procedures and rules that are required to perform the functions and provide the services. The technical architecture also encompasses the data stores needed for each component to provide the services. The technical architecture in each of the three major areas of the data warehouse is described in the following subsections.
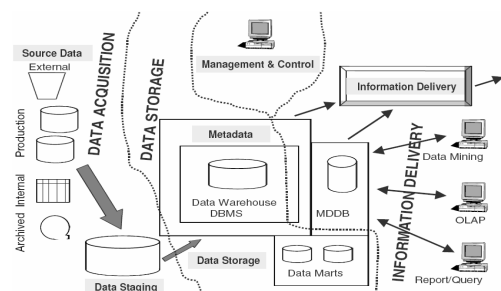

Figure 2.2 Architectural components in the three major areas

2.3.1 Data Acquisition

This area covers the entire process of extracting data from the data sources, moving all the extracted data to the staging area, and preparing the data for loading into the data warehouse repository. The two major

architectural components identified earlier as part of this area are source data and staging area.

2.3.2 Data Storage

This area covers the process of loading the data from the staging area into the data warehouse repository. All functions for transforming and integrating the data are completed in the data staging area. The prepared data in the data warehouse is like the finished product that is ready to be stacked in an industrial warehouse.

2.3.3 Information Delivery

This area spans a broad spectrum of many different methods of making information available to users. The information delivery component makes it easy for the user to access the information either directly from the enterprise-wide data warehouse, from the dependent data marts, or from the set of conformed data marts. Most of the information access in a data warehouse is through online queries and interactive analysis sessions.

2.4 Design Approaches

Writing in a leading trade magazine in 1998, Bill Inmon stated, "The single most important issue facing the IT manager this year is whether to build the data warehouse first or the data mart first." The statement is true even today. The two basic approaches are described in the following subsections [3].

2.4.1 Top-Down Approach

Top-down implementation means that a global data warehouse covering all the work-groups, departments, or lines of businesses is constructed first, and data marts according to different subjects will be built afterward. This approach is shown in Figure 2.3. The global data warehouse provides all the users through the enterprise with a consistent data definition and software environment. Then, if needed, data marts corresponding to separate departments and businesses can be built with their data populated from the global data warehouse rather than directly from the operational or external data sources.

2.4.2 Bottom-Up Approach

Bottom up implementation means that data marts are designed and constructed before, or in parallel with, a global data warehouse, as depicted in Figure 2.4. Data marts can extract their data either from a global data warehouse or directly from the operational or external data sources. While initial data mart implementations expand, a more global data warehouse will be
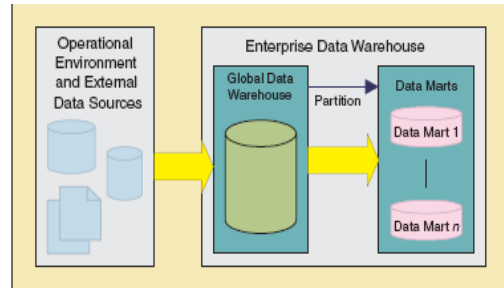

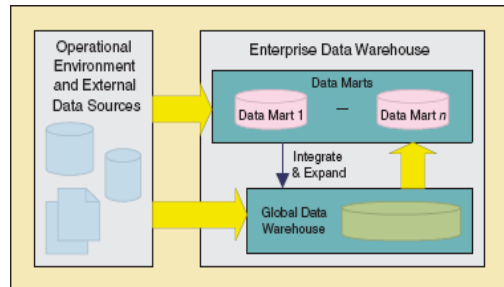Figure 2.3 Top-Down Implementation

Figure 2.4 Bottom-Up Implementation

built incrementally. This approach is more widely accepted today than the top-down approach because data marts have a less complex design and are more economical in terms of hardware and other resources than a global data warehouse, which is more easily evaluated and can result in much faster harvest and payback.

2.5 Online Analytical Processing (OLAP)

OLAP is designed to provide summary views of the basis for the aggregation of the transactions according to user-defined business dimensions from the static stores of the data warehouse. Data warehouses support OLAP applications by storing and maintaining data in multidimensional format. Data in an OLAP warehouse is extracted and loaded from multiple OLTP data sources using Extract, Transfer, and Load (ETL) tools. Three types of relational DBMS servers— relational, multidimensional, and hybrid online analytical processing — now support OLAP on data warehouses built with relational database systems [4].

➢ ROLAP servers: ROLAP middleware servers sit between the relational back-end server where the data warehouse is stored and the client front-end tools. They identify the views to be materialized, rephrase user queries in terms of the appropriate materialized views, and generate multi-statement SQL for the back-end server.

➢ MOLAP servers: MOLAP is a native server architecture that does not exploit the

functionality of a relational back end but directly supports a multidimensional data view through a multidimensional storage engine. MOLAP enables the implementation of multidimensional queries on the storage layer via direct mapping.

➢ HOLAP servers: The HOLAP architecture combines ROLAP and MOLAP technologies. HOLAP's selective view materialization, selective index building, and query and resource scheduling are similar to its MOLAP and ROLAP counter parts.

2.6 Cube

'Cube' is a metaphor used in an OLAP context and it is a storage format that enables high performance multidimensional analysis. A cube is n-dimensional, where n is the number of dimensions. Each cell in the cube contains the value of a certain measure for specified values of the n dimensions. Some of the aggregations in the dimension hierarchies are pre-calculated, which substantially reduces query times.

The star schema design is optimal for creating OLAP cubes. An OLAP cube can be accessed from a wide range of analytical products supporting OLE db for OLAP, in which the cube can be viewed as a report or in graphic format. OLAP has become a standard protocol for analytical processing.

The measures in an OLAP cube originate from a fact table, and dimension tables are the main source of OLAP dimensions. A dimension table can be used as the base for several different dimensions in a cube.

3. Requirement Definition

3.1 General Requirement Descriptions

We gather the different point of view from users' requirement as follow:

➢ Customer Centric Analysis. User want performed per customer and/or on a grouped level of customers. Of course the other dimensions like product may be analyzed as well.

◇ Does any customer group perform better than all the others?

◇ Profitability, as compared to an average for all customers.

◇ Discounts, costs and credits explaining the profitability.

◇ Credits and price adjustments — expressing bad quality and decrease in customer satisfaction.

◇ Sales and sales growth — and sales compared to the average customer.

➢ Product Analysis. User need performed per product and/or on a grouped level of products. Of course other dimensions like for example customer may be analyzed as well.

◇ Products with exceptional performance and growth

◇ Products selling below expected levels for profitability

◇ Which are the most sold products? And how do they perform from a profitability point of view?

◇ Which products have decreased most in sales? Should we plan for a phase out?

◇ Do we have products with exceptional levels of discounts and credits that erode our margins?

➢ Sales organizational analysis: Users' requirement performed from an internal organizational point of view in order to find out:

◇ Who are the top sales people performers?

◇ Which sales people improved the most, compared to last year?

◇ Which sales people reduced the profitability the most by manually reducing prices or generating credit notes?

◇ Which sales organization controls discount the best?

◇ Which sales organization is the most efficient by means of processed order volumes?

◇ Which sales organization is the most effective by means of correct deliveries, minimized credits and returns?

3.2 Business Metrics

The numbers the users analyze are the metrics that measure the profitability. These are the facts that indicate to the users how much profit is doing in fulfilling their objectives. In this case of the profitability, these metrics relate to the profit. These are the numbers that tell the users about their performance in profit. These are numbers about the profit of each individual item. The set of meaningful and useful metrics for analyzing profitability is gross line amount, total discount amount, net line amount, total cost and net margin. Main sources for the fact table

are the Sales Statistic Detail OSBSTD, Customer Order head OOHEAD, Customer Order lines OOLINE, Delivery Customer order head ODHEAD, Route Master file DROUTE, Item Master MITMAS and Customer Order types OOTYPE tables.

3.3 Business Dimensions

The dimension tables usually contain hierarchical data that are used in the OLAP (Online Analytical Processing) Cube Builder in Microsoft Analysis Manager to build hierarchies. The Analysis Manager is installed on an SQL server. The dimension hierarchy enables us to drill up and down in the OLAP cube. That is, we can follow a path from the highest analyze level down to its smallest components. Depending on which purpose the analysis serves, we probably want to be able to examine our business information on different levels. For every dimension a unique identifier must be included. This identifier is called a dimension key, to which the fact table is connected. The dimension key is automatically created when we create the dimension table in BPW Designer [10].

➢ The first obvious dimension is the calendar dimension the purpose of this dimension is to be able to analyze the measures in many models over a time period. The name of the dimension table is Calendar. The table is based on the following ERP table: CSYCAL.

➢ The customer dimension again for the profitability, analysis of sales must include analysis by breaking the sales down by items. The purpose of this dimension is to be able to analyze data for different customer-related aspects. The name of the dimension table is Customer. The table is based on the following ERP tables: OCUSMA and CSYTAB.

➢ The Delivery address dimension: Each customer can have several addresses for delivery and invoicing. This dimension includes only delivery addresses. The data in the dimension table is stored on the site level, where site is a unique combination of customer and address number. Compared with the dimension customers, this dimension includes more transactions since a customer can have several delivery addresses. The name of the dimension table is delivery addresses. The table is based on the following ERP tables: OCUSAD, OCUSMA and CSYTAB.

➢ The Sales organization: The purpose of this dimension is to be able to analyze data for different salespeople, sales departments and business areas. The name of the

dimension table is Sales organization. The table is based on the following ERP table: CSYTAB.

➢ The Item dimension: The purpose of this dimension is to be able to analyze data based on different item related aspects. The name of the dimension table is Item. The table is based on the following ERP tables: MITMAS, CSYTAB and SITCOM.

➢ The warehouse dimension: The purpose of this dimension is to be able to study data that is split (divided) between different warehouses. The name of the dimension table is Warehouse. The table is based on the following ERP tables: MITWHL, CFACIL, CMNDIV and CSYTAB.

➢ The Facility dimension: The purpose of this dimension is to be able to analyze data for different facilities. Facilities are used for activities such as sales, production or distribution. A facility always belongs to a division, which in turn belongs to a company. A facility can contain one or more warehouses. The name of the dimension table is Facility. The table is based on the following ERP table: CFACIL.

➢ The Division dimension: The purpose of this dimension is to be able to analyze data for different companies and divisions. This is an identity for a legal unit within a company group. Division is a key value in the financial system in ERP. One division may consist of several facilities. Facilities are used in material and production management and for purchasing purposes. The name of the dimension table is Division. The table is based on the following ERP tables: CMNDIV and CMNCMP.

➢ The Currency dimension: The purpose of this dimension is to be able to have currency as an analysis dimension. Currencies can be maintained either on the division level or on the company level, or on a combination of both. The measurement model assumes that currencies are maintained on the company level. If they are not, the relation between the division columns in the composers must be added. The name of the dimension table is Currency. The table is based on the following ERP table: CSYTAB.

➢ The Order type dimension: The purpose of this dimension is to be able to analyze data for different customer order types. The order type indicates an order type which is a combined ID for the settings that

determines how the order is processed during order entry and in the processing flow. The name of the dimension table is order type. The table is based on the following ERP table: OOTYPE.

3.4 Dimension Hierarchies

The dimension hierarchies are the paths for drilling down or rolling up in user analysis. In our case, the hierarchy of the calendar dimension consists of the levels of year, quarter, and month. The user can first sees the net margins for the entire year. Then the user moves down further to the level of quarters and looks at the net margins by individual quarters. After this, the user moves down further to the level of individual months to look at monthly numbers. In the same way, the hierarchy of the item dimension is consists of the levels of product

group, item group, style and item. The user can drill down or roll up to look at which style's net margins is better or what kind of item group's total cost is overspend, and which product group sales amount is the best, and so on.

3.5 Information Package Diagram

As shown in Figure 3.1, the information subject here is profitability. The metrics that are of interest for analysis are shown in the bottom section of the package diagram. The measurements are gross line amount, total discount amount, net line amount, total cost and net margin. The business dimensions along which these measurements are to be analyzed are shown at the top of diagram as column headings. These dimensions are calendar, order type, delivery address, sales organization, item, warehouse, facility, division, currency and

Information Package Diagram: Profitability

Dimensions

Calendar	Customer	Delivery Address	Sales Organization	Item	Warehouse	Facility	Division	Currency	Order Type
Year	Customer Group	Area	Sales Department	Product Group	Country	Division	Company	Division	Category
Quarter	Customer	Country	Sales Person	Item Group	Facility	Facility	Division	Currency	Order Type
Month				Style	Warehouse				
				Item					

Metrics: Gross Line Amount, Total Discount Amount, Net Line Amount, Total Cost, Net Margin
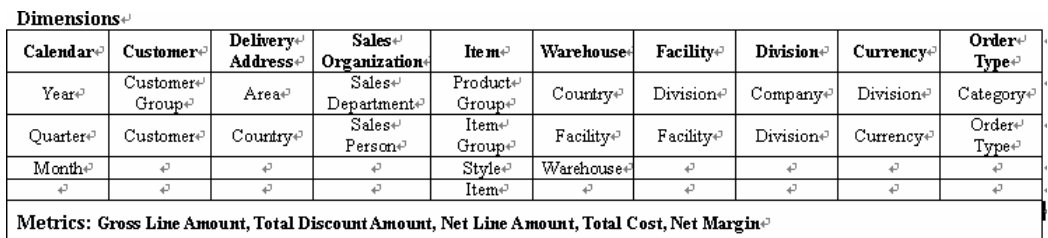
Figure 3.1 Information Package Diagram

customer. Each of these business dimensions contains a hierarchy or levels. For example, the sales organization dimension has the hierarchy going from sales department down to the level of individual sales person. The item dimension has the hierarchy going from product group down to the level of individual item.

4. Design

The requirements definition completely drives the data design for the data warehouse. Data design consists of putting together the data structures. A group of data elements form a data structure. The information package diagrams form the basis for the data warehouse. The data design process results in a dimensional data model.

4.1 Data Design

We designed the data warehouse using star schemas. Unlike the design of databases in online transaction processing (OLTP) systems, the star schema is a database design optimized for analytical needs, not least for multidimensional analysis. A star schema is organized around a fact table that contains data for different measures for a certain area, such as data for different order fulfillment indicators. This fact table is related to several dimension

tables containing information about the different structures that may be of interest.
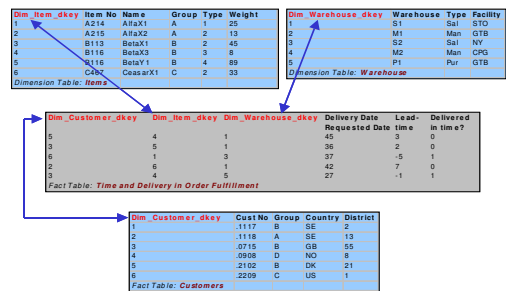

Figure 4.1 how an information model is constructed as a star schema

Figure4.1.[5] is an example of how an information model is constructed as a star schema with a fact table in the middle surrounded by dimension tables that each contains a dimension key that is connected to the fact table.

We will determine one of the data structures to be included in the dimensional model for profitability and derive the fact table from the information package diagram. Look at the calendar business dimension in Figure 3.1.

The customer business dimension is used when we want to analyze the facts by customers.

Sometimes our analysis could be a breakdown by individual customer. Another analysis could be at a higher level by customer groups. The list of data items relating to the customer dimension is customer group and customer. All of these relate of the customer in some way. We can, therefore, group all of these data items in one data structure of one relational table. We call this table the customer dimension table. The data items in the above list would all be attributes in this table.

Looking further into the information package diagram, we note that business dimensions shown as column headings. In our case of the profitability information package diagram, these other business dimensions are delivery address, sales organization, item, warehouse, facility, division, currency, order type and calendar. Just as we formed the customer dimension table, we can form the remaining dimension tables of delivery address, sales organization, item, warehouse, facility, division, currency, order type and calendar. The data items shown within each column would then be the attributes for each corresponding dimension table.

The dimensional model should primarily facilitate queries and analyses. The metrics inside the fact table are analyzed across one or more dimensions using the dimension table attributes. We are analyzing facts along attributes in the various dimension tables. The attributes in the dimension table are act as constraints and filters in our queries. We also find that any or all of the attributes of each dimension table can participate in a query. Further, each dimension table has an equal chance to be part of a query.

4.2 The Star Schema

The star schema for profitability is shown in Figure 4.2. It consists of the profitability fact table in the middle of schema diagram. Surrounding the fact table are ten dimension tables of calendar, customer, sales organization, currency, item, facility, order type, warehouse, division and delivery address. All these dimensions along which the users will analyze are found in the structure. Each dimension table is related to the fact table in a one-to-many relationship.

We have seen that a key component of the star schema is the set of dimension tables. This model that we created is used for analyzing sales statistics generated by SQL programs. The fact table's master file is (OSBSTD), but several other files are also used. Each transaction contains values for a delivery index, belonging
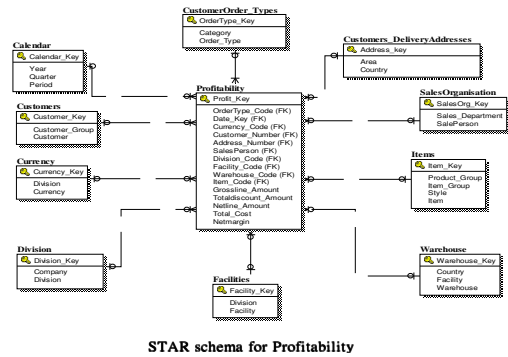

STAR schema for Profitability

Figure 4.2 The Star Schema for Profitability

to a customer order, an order line, and a line suffix. The measures are including gross line amount, net line amount, total discount amount, total cost and net margin. Each measurement goes into the fact table as an attribute for profitability.

4.3 Star Schema Keys

Refer to Figure 4.2 to see how the keys are formed for the dimension and fact tables. Each row in a dimension table is identified by a unique value of an attribute designated as the surrogate key of the dimension. The data warehouse contains historic data. Assume that the customer primary key gets changed in the middle of a year, because the customer is now stored in a different warehouse of the company. So we have to change the customer primary key in the data warehouse. If the customer key is the primary key of the customer dimension table, then the newer data for the same customer will reside in the data warehouse with different key values. This could cause problems if we need to aggregate the data from before the change with the data from after the change to the customer.

The surrogate keys are simply system-generated sequence numbers. They do not have any built-in meanings. Of course, the surrogate keys will be mapped to the customer system keys. Nevertheless, they are different. The customer key is the surrogate primary key for the store dimension table. The operational system primary key for the store reference table may be kept as just another non-key attribute in the store dimension table.

Each dimension table is in a one-to-many relationship with the central fact table. So the primary key of each dimension table must be a foreign key in the fact table. In our case, there are ten dimension tables of calendar, customer, sales organization, currency, item, facility, order type, warehouse, division and delivery address representative, then the primary key of each of these ten tables must be present in the profitability fact table as foreign keys.

4.4 Data Warehouse Tools

Business Performance Warehouse (BPW) [9] that we use tool in our study case is a system that easy-to-use data warehouse built on leading-edge data warehouse technology. Figure 4.3 provides an overview of the BPW architecture [6].

The BPW databases are stored in the repository. BPW repository stores all metadata and is the center of a data warehouse. The repository is located in a database on an SQL server designed to give users the ability to carry out high-performance data warehouse operations. Tables, views, columns and relationships between tables are examples of technical metadata; the business-oriented metadata is always mapped to technical metadata. All objects (also called entities) and columns have descriptions that we can modify. Business logic and application logic is also stored in the repository. In Figure 4.4 we can see the databases in BPW repository enlarged with connections between them and the tree view tabs in BPW Designer.
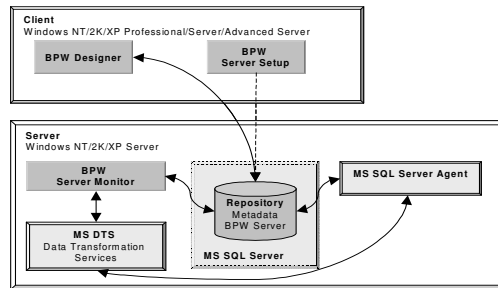
Figure 4.3 BPW Architecture
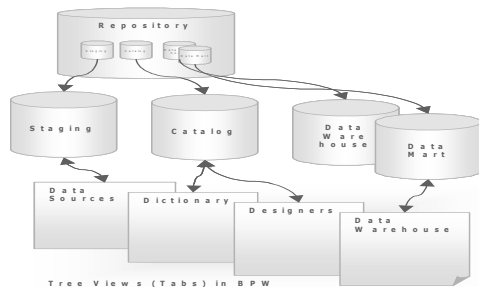
Figure 4.4 BPW Repository

4.5 Data Extraction, Transformation and Loading (ETL)

The goal of the data extraction step is to bring data from disparate sources into a database where it can be modified and incorporated into the data warehouse.

➢ Specify Operational Database: When we first logon BPW Designer we will get an attention message that informs us that we have not yet set up the proper system connection to the specific data source.

Click yes to be prompted with the system properties dialog where we can set up a correct destination connection for the data source. The system name can be changed. Select source/operational from the drop-down menu in the system type field. We can select a connection to local operational database that contains data from the drop-down menu. In Figure 4.6 we see all the connections and their properties.
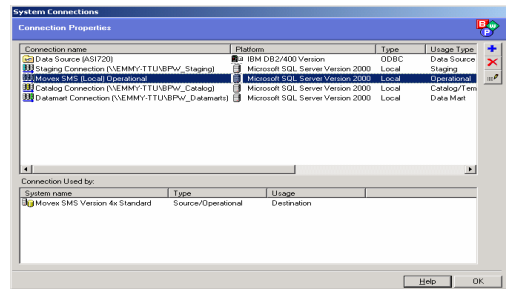
Figure 4.6 All connections

➢ Specify ODBC Connections: For any data source system we are going to use in BPW we must specify the system connection from which the data will be copied and extracted, and the destination at which the data should be stored. In this case, we extract business data from an IBM AS/400 platform/data source with an ODBC connection [7] to a staging database located on the data warehouse server.

➢ Import metadata: Metadata in BPW can be defined as the foundation that keeps the data warehouse organized, logical and together. It brings necessary meaning to the source system tables and components that are stored. Metadata describes and characterizes a specific resource. It can be compared to the alphabetizing and classification that we use to find what we are looking for in a library or archive. BPW is delivered together with ERP metadata. Depending on which ERP version we use [8]. Normally a metadata representation has the following characteristics: it includes a limited number of elements (such as title, type, format and relation) and it includes the name of each element together with the element's meaning. All metadata is stored in the BPW Repository, thus providing a so-called electronic map of the enterprise. We always see a status report when the metadata import is complete. After it appears we can continue with our design in BPW Designer. Figure4.7 illustrates the import metadata from source system IBM AS/400, then select an ODBC data source
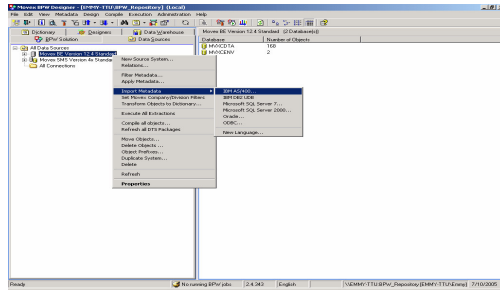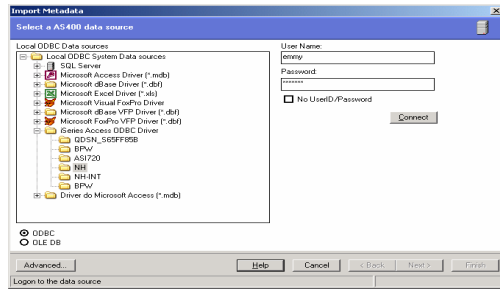
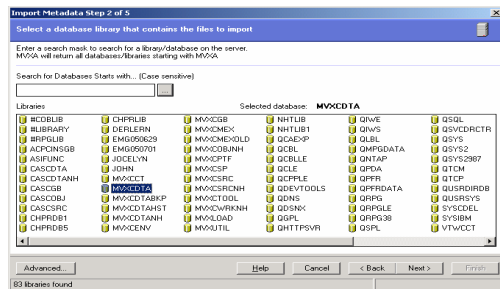Figure 4.7 Import metadata

Figure 4.8 Logon the data source

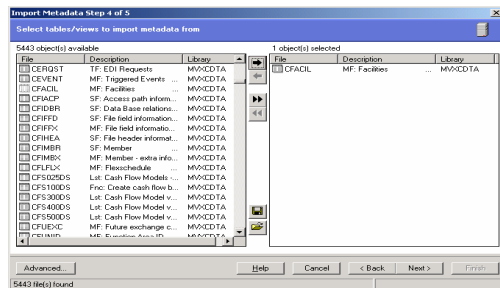Figure 4.9 Select database library

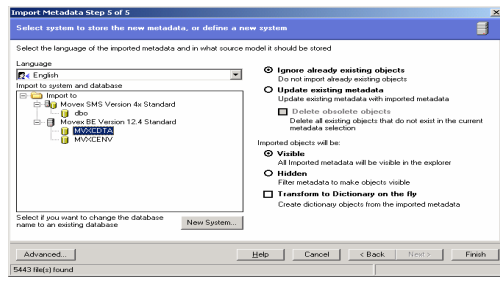Figure 4.10 Select tables/views to
import metadata

Figure 4.11 Select system to store
the new metadata

and then enter the user name and the password. click connect to continue shown in Figure4.8. Figure 4.9 starts with a database libraries search. Type the library name MVXCDTA to import metadata from. In Figure 4.10 on the left side the column we mark the objects we want to import. Use the arrow buttons to transfer objects from the available column to the selected column. When we are finished, click next. Figure 4.11 selects a system and database in BPW to import the metadata to. Ignore already existing objects is selected by default, change only when required. Click finish to start the metadata import. When the import is finished, a status report is displayed.

➢ Filter and extraction source data: All source objects are always available in the metadata repository, but to make our easier we do not need to have all objects visible. We can set a filter on an entire data source system or on just a small part of it, such as an individual information model or a fact table. Figure 4.12 to set up general filter criteria for extracting data from ERP system. Right-click a data source system name and click execute all extractions on the shortcut menu show on Figure 4.13.
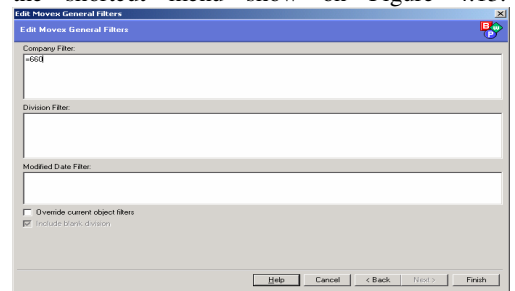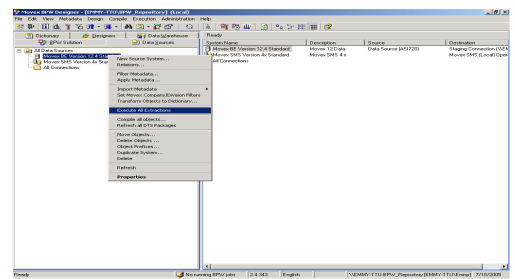

Figure 4.12 Set filters
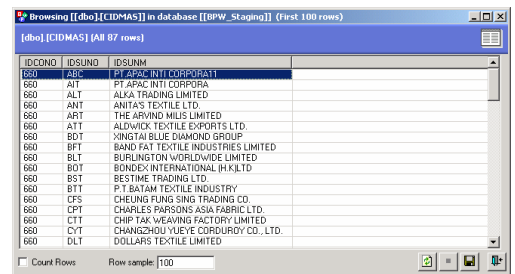

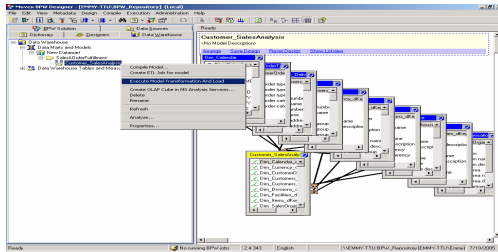Figure 4.13 Execute all extractions

Figure 4.14 Browsing data

Figure 4.15 Model transformation and load

It creates a background job that automatically extracts data from all compiled tables. Figure 4.14 browse extraction data. Figure 4.15 Right-click an information model name customer Sales Analysis on the Data Warehouse tab. Dimension tables and fact table belong to the model are list. Detail progress and status are display on Figure 4.16.

4.6 Transformation to OLAP Cubes

We transform dimension into cubes by an OLAP tool, such as Analysis Services, which is included in MS SQL server 2000. Analysis Services is also integrated with BPW, which means that cubes can be created directly from BPW Designer. We created a new data mart by right-clicking the top node of the tree structure and then clicking New Data Mart on the shortcut menu. When we create a model group we need to have a data mart to place the model group in.

Figure 4.17 illustrates the data warehouse region including all the compiled information models and sorted in model group(s). A model group belongs to a data mart. We can have more than one (parallel) data marts in the data warehouse. Right-click or double-click the model or model group for more information. Right-click to create an OLAP cube in Microsoft analysis services based on a selected model. Figure 4.18 shows the beginning of storage design wizard.
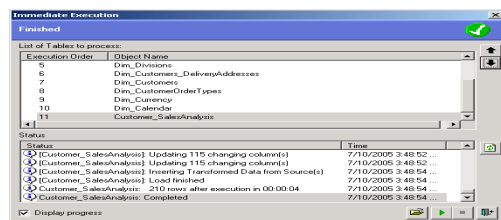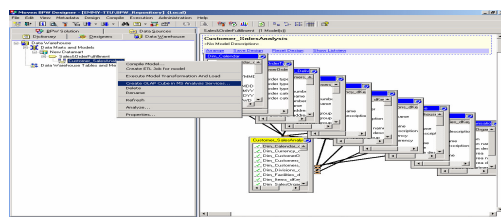

Figure 4.16 Display progress

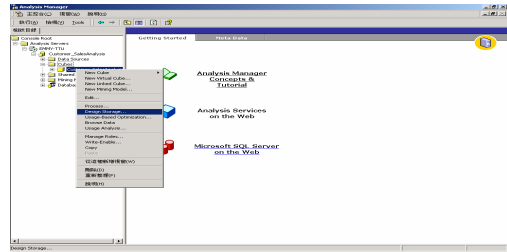Figure 4.17 The data warehouse regions

Figure 4.18 Execute Analysis Manager

5. Information Delivery

After performing all of these tasks that identified pertinent source systems, extracted and transformed the source data, design the data warehouse repository and loading the transformed and cleansed data into the data warehouse database effectively. We have to provide the mechanism for information delivery to our users.

Our information access in data warehouse is through online queries and interactive analysis sessions. The Microsoft analysis service is easy for the user to access the information either directly from the enterprise-wide data warehouse, from the dependent data marts, or from the set of conformed data marts.

In Figure 5.1 cube browser is to present the dimension tables on the top of area, and calendar dimension on the left of area, metrics values are shown on the right side. We can select dimension value. When the dimension is changed, metrics are also changed. As shown in Figure 5.2, we change the calendar dimension, we can see the data is spread on quarter 1, and the net margin is minus. We can drill down to find out each month's performance. Figure 5.3 shows we add customer dimension to analysis the casual loss is because of customer group VIP.

Figure 5.4 we drill down customer dimension from customer group to individual customer. Figure 5.5 shows that after we change the sale organization dimension, the metrics value changed. We can see the C1 sales department is loss. Figure 5.6 we can find the sale person Brian Lain made a good margin of profit. And we also can find out the loss is because of other sale person Simon Yang.
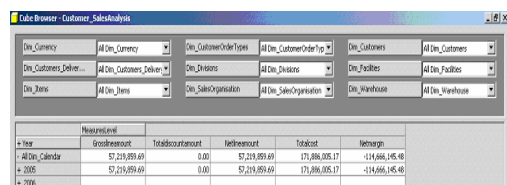

Figure 5.1 Cube browser present dimension tables and metrics
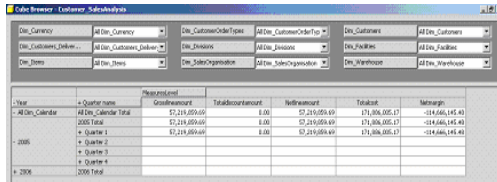
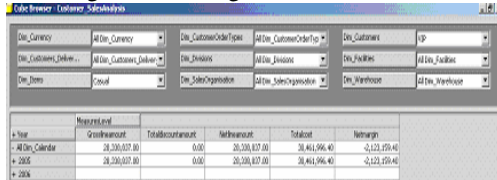Figure 5.2 Change calendar dimension
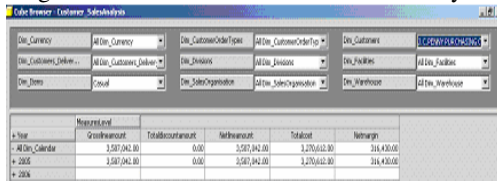

Figure 5.3 Add customer dimension to analysis
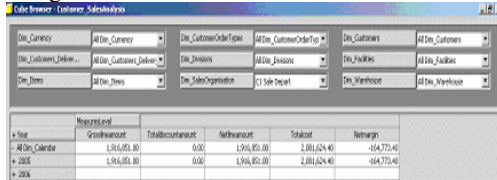

Figure 5.4 Drill down to individual customer

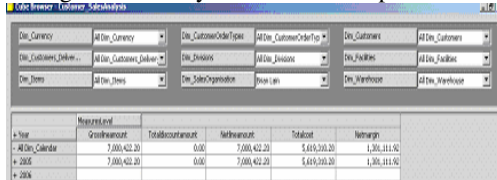Figure 5.5 Analysis casual sale department

Figure 5.6 Find out top saleperson

6. Conclusion and Future Works

In this paper, we design and implement an online multidimensional analysis system based on the existing ERP system for a garment company. First, we summarize the requirements of sales that users want to analyze profitability in a number of ways into an information package diagram. Secondly, we design the star schema according to the BPW [9] data warehouse architecture, and design the data storage in a SQL Server platform. Finally, we present the information delivery in various points of view with BPW's Analysis Services, to provide online inquiry for effectively improving the quality, shortening the delivery and saving the cost. The executives and managers who are responsible for keeping the enterprise competitive need information to make proper decisions can use such a system to improve their efficiency and effectiveness. Further research works are required to extend the scope of other models to integrate a full-fledged version and to consider needed security control functions.

The data warehouse exists for one reason and one reason alone. It is there just for providing strategic information to our users. This research offers a data warehouse for a garment company to plan the prototype with the construction, let enterprises understand that the data warehouse can reach those functions? Whether it accords with the demand that enterprises make policy and support and assess, and it can be as the garment industry or the reference of planning or construction data warehouse of other manufacturing industries to expect.

References

[1] Inmon, W.H., Building the Data Warehouse, 3rd edition, John Wiley & Sons, New York, 2002.

[2] Ponniah, Paulraj, Data Warehousing Fundamentals, John Wiley & Sons, New York, 2001.

[3] Shi, Dongyuan, Yinhong Lee, Xianzhong Duan, and Q.H. Wu, "Power System Data Warehouses," IEEE Computer Applications in Power, July 2001, pp. 49-55.

[4] Chaudhuri, Surajit, Umeshwar Dayal, and Venkatesh Ganti, "Database technology for decision support systems," IEEE Computer Magazine, Dec. 2001, pp. 48-55.

[5] BPM Team, "Overview Movex Business Measurement Models, "
http://wire.intentia.com/intentia/wire2000.nsf/Index2?Openframeset&login

[6] Melander, Per and Ingela Skoog, "Concept Document BPW Architecture, "
http://wire.intentia.com/intentia/wire2000.nsf/Index2?Openframeset&login

[7] Schyldt, Hans, "Technical Document Setup Client Access for ODBC Connections, "
http://wire.intentia.com/intentia/wire2000.nsf/Index2?Openframeset&login

[8] Skoog, Ingela, "Intentia Installation Guide Movex BPW, v2, "
http://wire.intentia.com/intentia/wire2000.nsf/Index2?Openframeset&login

[9] BPM Team, "78_Business Performance Warehousing BPW, "
http://wire.intentia.com/intentia/wire2000.nsf/Index2?Openframeset&login

[10] BPM Team, "Software Download for Movex BPW 2, "
http://wire.intentia.com/intentia/wire2000.nsf/Index2?Openframeset&login