

Analysis and Design of Data Mart Decision Support Systems at PT Marlindo Tirta Nusantara

Budi Sudrajat

Universitas Bina Sarana Informatika
Jl. Kamal Raya No. 18 Outer Ring Road Barat
Cengkareng, Jakarta Barat
e-mail : budi.bst@bsi.ac.id

Abstract— The role of information technology in each company is very influential in providing complete and accurate information. With a data warehouse or data mart the company can utilize important asset data to present the information needed.. Shipping firms nowadays are in operation with extremely competitive and challenging environment. Vast volume of knowledge is generated from various operational systems and these are used for determination of several business issues that needed imperative handling. The results obtained from this study are data marts that simplify and accelerate the provision of data and information to support decision making, so as to provide a basis for developing DSS and EIS applications. Conclusions obtained by the data warehouse or data mart can provide complete, accurate and integrated information as a basis for consideration for the executive in making decisions so that decisions taken are based on real facts owned by the company.

Keywords: Decision Support System, data mart, data warehouse, shipping.

I. INTRODUCTION

1.1 Introduction

Increasingly tight business competition now makes the role of information technology very important for many companies in various fields. Since information technology is growing, it encourages companies to develop existing information technology to improve company performance. Information plays an important role in winning business competition, therefore the existing data must be managed properly so as to produce quality, complete, precise and accurate information that will be useful for management. Data warehouse is a technology that is capable of managing large amounts of data. This technology can integrate data from various divisions that can facilitate management in conducting analysis, query and report processes. The truth value of the stored data will be maintained in a data warehouse system because it is only readable so that data changes cannot be made. The concept of a data warehouse is a technology that is very useful for companies to display information in the database quickly and accurately to meet the needs of the

company. The information generated is a summary that can help the executive in making decisions.

PT. Marlindo Tirta Nusantara is a company engaged in loading and unloading freight services.. PT. Marlindo Tirta Nusantara continues to strive to keep abreast of information technology to support its business activities in competing with similar companies.

1.2 Formulation of The Problem

How to analyze the ongoing On Line Transaction Processing (OLTP) database and design transport data marts to meet the needs of decision making in transport strategies and create on line Analytical Processing (OLAP) data marts to display company history data to meet executive needs.

II. RELATED STUDIES

2.1. Literature review

According to Connolly (2015), a database is a logically related data collection that is used together and a description of the data that is designed to meet the needs of an organization. Connolly further stated that the database approach is to separate the data

structure from the application program and store it in the database. To analyze the information needs of an organization, we first determine the entities, attributes and relationships.

OLTP is designed to allow simultaneous access by multiple OLTP is designed to allow simultaneous access by several users to the same data source and to arrange the necessary processes (Viera, 2010). Transactions carried out include insert, update, and delete operations. OLTP databases are usually relational and in the third normal form.

Data warehouse is a development of a database concept that provides better data resources for users, allowing users to manipulate and use the data intuitively (Mc Leod: 2011). In the data warehouse there are only 2 activities on the data, namely loading data and access data, there is no data update activity. Data from database operations and external sources are extracted, filtered, summarized and then loaded into the data warehouse. Because the data warehouse contains historical data, it often involves retrieving data that has been completely summarized.

It is important to note that there are huge differences between these two tools though they may serve same purpose. Firstly, data mart contains programs, data, software and hardware of a specific department of a company. There can be separate data marts for finance, sales, production or marketing. All these data marts are different but they can be coordinated. Data mart of one department is different from data mart of another department, and though indexed, this system is not suitable for a huge data base as it is designed to meet the requirements of a particular department. Data Warehousing is not limited to a particular department and it represents the database of a complete organization. The data stored in data warehouse is more detailed though indexing is light as it has to store huge amounts of information. It is also difficult to manage and takes a long time to process. It implies then that data marts are quick and easy to use, as they make use of small amounts of data. Data warehousing is also more expensive because of the same reason. (Neha Marhotra : 2015). The motivations behind the creation of these two types of data marts are also typically different. Dependent data marts are usually built to achieve improved performance and availability, better control, and lower telecommunication costs resulting from local access of data relevant to a specific department. The creation of independent data marts is often driven by the need to have a solution within a shorter time. (Pregibon : 2012).

Data warehouses are useful to make it easier to create applications for the Decision Support System (DSS) and Executive Information System (EIS). In fact the data warehouse is a form of database, but the

data warehouse has an interactive nature in accessing information with high response times. Thus the data warehouse is different from the database because the database is used in the operational processes of business transactions and the data warehouse focuses on the process of analysis and decision making for executives. The database provided in the data warehouse environment is different from the database in the operational environment that supports OLTP. The difference is that the database in the data warehouse is an analysis database, while the database in the operational environment that supports OLTP is an operational database. Although sometimes used is confused, the term data warehousing and OLAP are different system components that are often referred to as decision support components. OLAP (Online Analytical Processing) is a technology that processes data in a data warehouse in a multidimensional structure, providing fast answers to complex query analysis. The purpose of OLAP is to organize large amounts of data so that it can be analyzed and evaluated quickly using online tools such as Microsoft Excel (Pivot Table), as well as graphical representation. Answers to querying historical data often require the next query, because researchers must answer and explore all possibilities. The OLAP system provides speed and flexibility to support analysis in real time.

2.2 Datawarehouse concept

The data warehouse contains combined data from many sources, plus summary information and covers a long period of time. General workloads that involve ad hoc queries, which are quite complex and fast response times are important. These characters distinguish warehouse applications from OLTP applications, and must use different DBMS design and implementation techniques to achieve satisfactory results. Distributed DBMS with good scalability and high availability (achieved by storing tables redundantly on more than one site) is needed for very large warehouses.

2.2.1 Data Warehouse Architecture

According to Connolly (2015), the processes, tools, and technologies related to the data warehouse architecture are:

1. Operational data

Data sources for the data warehouse can come from :

- a. Operational Data Mainframe
- b. Data on departmental parts that are stored in a variety of systems; proprietary file system, for example; RMS, relational DBMS, such as: Informix and Oracle.

c. Internal data stored from workstations and workstation servers.

d. Data originating from external systems, such as: the internet, commercial databases or company databases related to customers or suppliers.

2. Load Manager

Serves to perform all operations related to extracting and loading data into the data warehouse. The data can be obtained directly from the data source or usually comes from the operational data storage. This is done by the load manager, including simple data transformation to prepare data before entering into the data warehouse. The size and complexity of these components differ between data warehouses.

3. Warehouse manager

Serves to perform all operations related to data management in the data warehouse. Operations carried out by the warehouse manager include the following :

- a. Analyze data to ensure consistency.
- b. Transform and merge data sources from temporary data storage into data warehouse tables.
- c. Making indexes and views in the table.
- d. Generate denormalization (if needed).
- e. Generate aggregation (if needed).
- f. Back-up dan data archive.

4. Query manager

Serves to perform all operations related to management from user queries. The complexity of the query manager is determined from the facilities provided by the access device from the end user and database. The thing that is done by the query manager is to direct the query to the table that matches the scheduling of query execution.

5. Detail data

This section is where the data warehouse stores all detailed data with a database scheme, which aims to complete the data set for the data warehouse.

6. Lightly and highly summarized data

This section is where the data warehouse stores a collection of lightly and highly summarized data generated by the warehouse manager. The purpose of this information recapitulation (both lightly and

highly) is to improve query performance. This data recapitulation will be continuously updated every time when new data enters the data warehouse.

7. Archive or back up data

This section is where the data warehouse stores all detailed data and recapitulation data for documentation and backing up purposes.

8. Metadata

This section is where the data warehouse stores all information from metadata (data about data) that is used in each process in the data warehouse, for example the origin of data, data definition, data location, quality, and age of data in the data warehouse. These data are used as supporting data or supporting information generated by the data warehouse to support the decision making process.

The metadata structure is different in each process because each one has a different purpose. The purpose of using metadata is as follows :

- a. Extraction and loading processes to map data sources with data in the data warehouse.
- b. The process of managing a data warehouse for automation that produces a recapitulation table
- c. As part of the query management process to point the query to the most appropriate data source.
- d. Access tool for end users

9. End-User Access Tools

The principle of purpose of the data warehouse is to present information to the user to support decision making. Users will interact with the data warehouse through end user access devices. The following are some categories of end user access devices :

- a. Reporting and query tools
- b. Application development tools
- c. Executive Information System Device
- d. Online Analytical Processing Device
- e. Data Mining Device

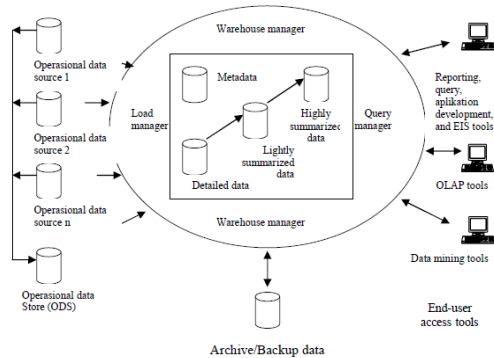


Figure 2.1
Data Warehouse Architecture (Connolly, 2015)

2.3 Data Mart

Data Mart is a smaller collection of data from a data warehouse that is used to analyze business in one division, while a data warehouse is a collection of all company data that is used to analyze the overall business. Data Marts are a small part of a data warehouse that supports requests from a department or a specific purpose of a business (Connolly, 2015).

A data mart contains a small portion of a data warehouse usually in the form of a brief data report that has a relationship with a department or a specific purpose of a business. Data marts can stand alone or be centrally tied to the company's data warehouse. When a data warehouse grows wider, the ability to fulfill all kinds of organizational needs can be met. The popularity of data marts gives the fact that a vast data warehouse of a company proved difficult to make and use. Characteristics that distinguish data marts and data warehouses :

1. Data mart focuses only on user requests relating to one department or specific purpose of a business.
2. Data mart usually do not contain operational data in detail, unlike a data warehouse.
3. Data mart when compared to data warehouses contain smaller amounts of data and data marts are easier to understand and process.

Gain data mart before building a data warehouse :

- a. Simple data mart to be implemented
- b. Users find business needs faster.
- c. Queries always run fast, because the database is smaller than the full warehouse.
- d. The data mart holds out unexpected data that should be contained in a data warehouse.
- e. The data mart is very fast to be built
- f. The hardware requirements of each department are usually smaller than for a full data warehouse.
- g. Data mart help in supporting competitive advantage.

Data mart losses before building a data warehouse :

- a. If a warehouse is built first, the addition of hardware is needed to support the marts in addition to the warehouse.
- b. Decentralized of data

III. RESEARCH METHOD

3.1 Stage of Research

Data from PT Marlindo Tirta Nusantara, it can be seen that the data provided is not well structured and is still seen in the condition of normalization. From here it can be analyzed and seen that the data cannot yet be processed and is not appropriate to be used as a data mart and has not been able to determine the measure, dimension, or fact table because there is no many-to-many relationship or many relationships in the current table, therefore the ERD design is done again which is intended to find the relationship of many to many or many

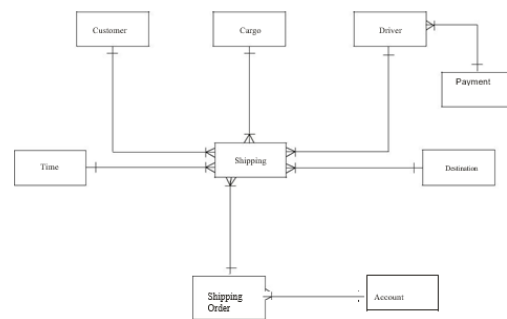


Figure 3.1 Entity Relationship Diagram

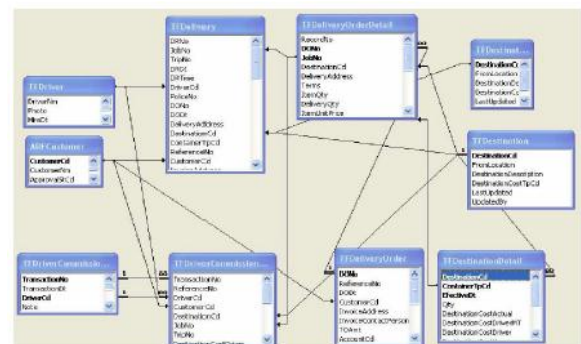


Figure 3.2 Entity Relationship Diagram

Data model proposed for PT Marlindo Tirta Nusantara :

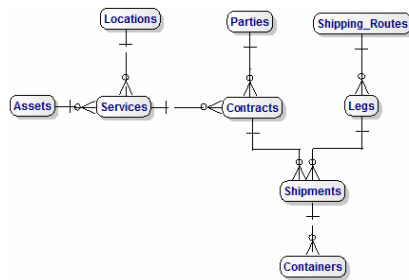


Figure 3.3 Data Model

The stages of research are systematic sequences in conducting research. In the preparation of this final assignment the stages of analysis carried out in general can be seen in the framework of the following research stages :

1. Analysis OLTP

This method of analysis is done to tidy up or reconstruct data that is not yet structured and flat (not yet in its normal form) and then made into a structured (normalization). This method is designed with the aim of maximizing the transaction processing capacity of the company.

2. Data Warehouse Analysis

From structured data, analysis is then carried out using the data warehouse method. This analytical method is carried out to support ad hoc query processes and support relatively small numbers of unpredictable transactions and to be able to answer whether the existing queries are ad hoc, structured and heuristic. This analysis a measure, dimension and directive key will be determined which will then be used as a fact table. After the fact table is determined, the next step is to enter into OLAP Implementation by using OLAP manager.

3. OLAP Implementation

At this stage, steps are taken to make OLAP from connection to data source to the results of browsing cube that provides information from fact and dimension tables.

4. Integration

This analysis is useful for combining data marts into the data warehouse. The data mart is part of the data warehouse but flat flats can still be seen using this analysis process.

5. Reporting

This process is a process for making excel based front end users because the application does not have this application. Reporting is done to facilitate or streamline an executive in seeing information that has been provided through a pivot table .

IV. RESULTS of ANALYSIS

4.1 Analysis

At the analysis stage, identify the things needed in business analysis to support the company's business objectives, by making the priority level based on the results of business analysis.

ANALYSIS NEEDED	LOW	MEDIUM	HIGH
Sales			x
Customers			x
Competitors		x	
Product			x
Location			x
Promotion	x		
Labour Resource	x		

Table 4.1 Analysis Needed Priority

4.2 Role and Process

By identifying all potential players who have responsibilities in the transport area to be analyzed later, it means that they have the same perspective from an executive. For each role, identifying individuals who will represent the needs of these roles include :

1. Identify the role of each player who has potential involvement in the analysis
2. Identify the analysis process that occurs in the player.
3. Understand the information requirements needed from each analysis process.
 - a. Transportation Commission Analysis
Amount of shipment : "What is the quantity of order delivery sent based on destination, driver, commission payment, charge, customer, time ".
 - b. Analysis of the Financial Commission
Amount of turnover : "What is the amount of turnover obtained based on destination, driver, commission payment, charge, customer, time".
Amount of transportation costs: "What is the amount of gross and clean transportation costs incurred based on destination, driver, payment of commission, cargo, customer, time".

Commission amount: "What is the commission amount based on destination, driver, commission payment, charge, customer, time".

Amount of gross profit: "What is the amount of gross profit that can be obtained based on purpose, driver, commission payment, charge, customer, time".

4.3 Key Performance Indikator (KPI)

Strategic information needed by the executive at PT Marlindo Tirta Nusantara, among others :

1. Executives want a system that is able to provide historical information.
2. Executives want a new system that can support it decision-making.
3. Executives want a new system that can display information on commission payments per driver and crew.
4. Executives want a new system that can display turnover statistics and gross profit per period time in graphical form.
5. Executives want a system that just provides information regarding the amount, turnover, net shipping costs, fees dirty shipments, driver and crew commission information one of the vital parts of a company's performance, and gross profits.

Look at the analysis that has been done before and interview what is done with the executive can be determined by KPI and measurement for making data mart transportation at PT Marlindo Tirta Nusantara.

1. Identifying KPI

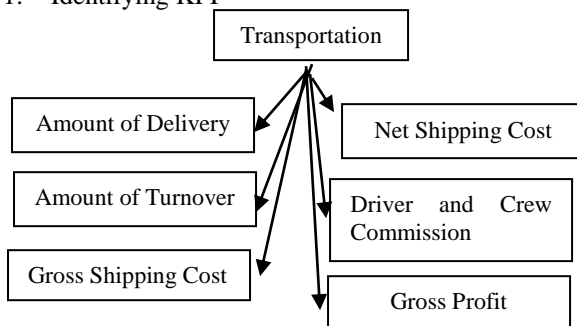


Figure 4.1 Key Performance Indicator

2. Identifies information objects and events

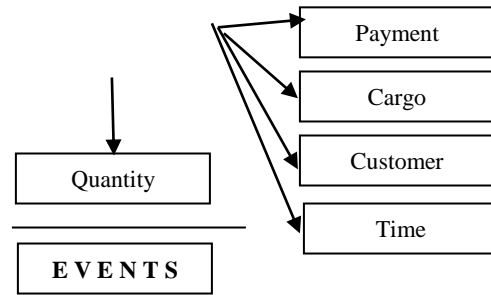
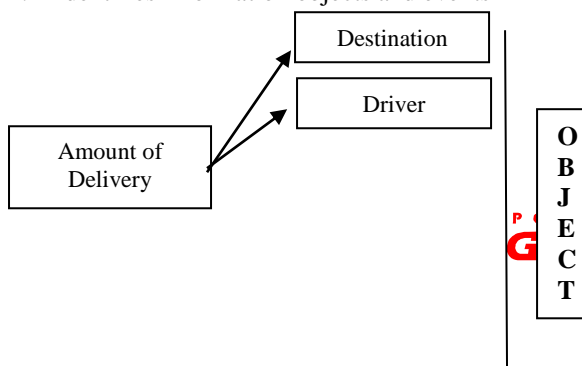


Figure 4.2 Objects and events of the number of shipments

4.4 Identifying Facts

Facts represent the measures that result from events. Based on the analysis above, measures are obtained between others:

- a. Total Quantity
- b. Total Turnover
- c. Total gross shipping costs = Total Delivery
- d. Net Shipping Costs = Destination Cost
- e. Total Driver Commission
- f. Total Crew Commission
- g. Total Commission
- h. Gross Profit

4.5 Dimension

The next area that will be identified from the information is dimension. Dimension is an entity that interacts with an event.

Dimension can be identified by looking at the relationships found in the fact table. Entities that have relations with the fact table can be used as candidates for dimension.

4.6 Source and Transformation

Source and transformation is one of the key components in the data warehouse life cycle Source and transforming applications are carried out using DTS (Data Transformation Services).

DTS is a very powerful application to consolidate data from various data sources and move it into a warehouse. Data in the Transport division system before (data source) is transferred to the data mart using DTS.

There are some data that are not found in the data source, but at this stage of analysis the data appears like the dimension of time, obtained by looking at the Transport data in the delivery date field. So data can be created for dimension time and can integrated with existing fact tables.

A. Sources consist of :

1. OLTP (Online Transaction Processing), This stage is the operational system needed for run a business analysis of the company.
 2. Data Warehouse, This is an OLAP repository when the analysis will run or used as a basis for moving data to a data mart department. In some data warehouses, the data mart will become a source and will be integrated into the enterprise data warehouse.
- B. Transformations, Transformation can replace any place for data can be moved between two points:
1. From the OLTP system source into the OLAP data warehouse, the source will be extracted, cleaned, and loaded back into structured data with new sources.
 2. Data warehouse allows preparing to push for extraction into the data mart department.

4.7. Implementation OLAP

The Olap implementation is the creation of a cube that queries the database in question. In this thesis the cube is created using OLAP Manager, along with the steps to implement it :



Figure 4.3 Create New Data Base

The analysis carried out is for information systems data mart and not a data warehouse and because there is no data warehouse that can be used to be combined or integrated, there is no need for integration to merge data marts into the data warehouse.

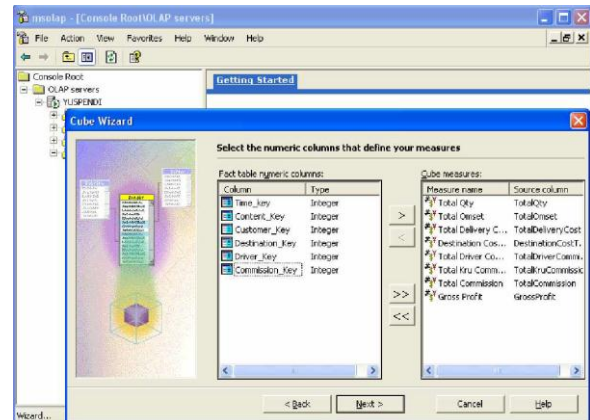


Figure 4.4 Measure Definition

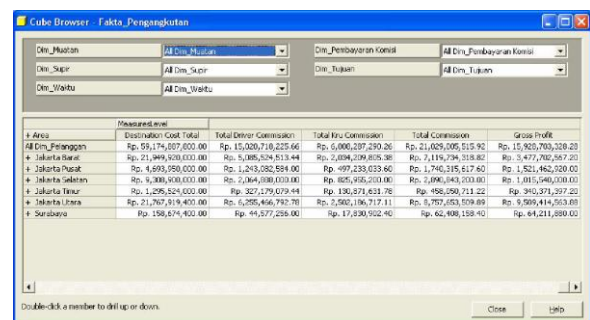


Figure 4.5 Browsing Cube

V. CONCLUSION

Based on the results of data mart analysis and design at PT Marlindo Tirta Nusantara, conclusions can be drawn :

1. With datamart, operational data and historical data from PT Marlindo Tirta Nusantara which was initially not optimally utilized, can be utilized and processed further as an analysis material by executive management.
2. Data mart is able to provide the information needed for shipping commission
3. The data mart stores all the information on the commission PT Marlindo Tirta Nusantara's transportation with the aim of helping or make it easier for executives to look for or analyze transport data that has existed for the past few years.

REFERENCES

- [1] Abbas. Application of Transportation. Surabaya: Pustaka Agung Harapan, 2011.
- [2] Connolly Thomas M and Carolyn E Begg. Database System A Practical Approach to Design, Implementation, and Management Edition. England: Pearson Education Limited. 2015.

- [3] Inmon W.H. Building the Data Warehouse Third Edition, USA: John Wiley, 2001
- [4] Neha Malhotra, Implementation of Data Marts in Data Ware House: International Journal of Advance Research, ideas and innovations in Technology, Vol 1 issue 2 (2015)
- [5] Peterson, Timothy and James Pinkelman. Microsoft OLAP Unleashed. USA: Sams Publishing. 2000
- [6] Poe, Vidette. Building a Data Warehouse for Decision Support. USA: Prentice-Hall, Inc, New Jersey. 1996
- [7] Pregibon, D. (2012). Data Mining. Statistical Computing and Graphics