

Designing a Virtual Data Warehouse in Supporting Sales Information Needs

(Case Study: National Scale Building Material Store X)

Andrew Sumichan^{1*}, I Made Gede Yudiana², Muhammad Ridwan Satrio³, I Made Sudarma⁴

^{1,2,3}Departement Electrical and Computer Engineering, Post Graduate Program, Udayana University
Department of Electrical Engineering, Udayana University
*andrewsumichan@gmail.com

Abstract - National-scale building material stores that have branches in various regions certainly want the business to run smoothly. This must be supported by good sales activities at the central store and also branch stores spread in various places. However, there is no solution to collect sales data from branch shops quickly, causing a slow response to the analysis of the increasing trend in sales of goods and goods whose sales trend has dropped. This can reduce sales competitiveness in this national scale business with similar competitors. Virtual data warehouses can help collect scattered operational data, in this case sales data located at branch stores in various locations are then collected into a data warehouse that is stored in the cloud server so that it can be processed for strategic decisions. Virtual data warehouse is a data warehouse that connects operational databases regardless of where the database is located and regardless of the format, looks as if somewhere and in a consistent format. This study produces a virtual data warehouse structure that applies the kimball nine step design method so that it produces a data warehouse schema model with the star scheme. This study explains the design of a virtual data warehouse that can facilitate analysis of sales data at national scale building material stores that have branches in various places.

Index Terms— Virtual Data Warehouse, Sales, Building Material Stores, Data Collection

I. INTRODUCTION

THE use of sales data at this time is very important. Because by knowing these things, the sales trend is obtained so that the trading business will be able to determine the number and type of goods to be sold to get more profits. National scale building material stores have branches in various regions. The absence of sales data processing that can provide information on trends in sales of goods can leave companies behind their competitors. Data warehouse can provide a solution for processing sales data seen in research [1], [2], [3], [4], [5]. In a study by A. Saputra and M. Sudarma (2018), mentioning the existence of a data warehouse on culinary sales, the information obtained is more effective and useful because it has been filtered out of the dimensions of data and sales reports from the data warehouse to produce accurate data and faster [1]. Then there are problems about how to implement a data warehouse if data sources are scattered in various places. These problems

can be overcome by other forms of data warehouses, namely virtual data warehouses that aim to collect information data if operational databases are scattered in various places in the research [6], [7], [8], [9].

In research T. R. Ross et al (2014), designing a virtual data warehouse for a collection of health research data from data sources spread across various health research data sites from various different website projects [6]. Then in the research of S. S. Kondabolu and J. Nasina (2010), this study succeeded in designing a virtual data warehouse for manufacturing industries that is useful in supporting data processing in the industry [7]. However, in the virtual data warehouse studies and as far as the excavation of the journal literature we have taken, there is no research to use a virtual data warehouse in supporting the sale of goods to companies that have branches in various places.

Therefore, this study contributes are introducing a virtual data warehouse structure that functions to facilitate the analysis of sales data in building material stores that have several branches. Virtual data warehouse designed using the design method. Nine step kimball then produces a data warehouse schema model with the star scheme. The overall

structure of the study uses the five-part form, an explanation of the methods and analysis of the virtual data warehouse that is applied in sections III and IV until concluding with the conclusions of this study in section V.

II. LITERATURE

A. Virtual Data Warehouse

Virtual data warehouse refers to layers of databases that sit on top of existing databases to permit users in the query of the whole databases as if they are a single entity although these databases are logically and physically separated [9].

B. Difference Between Virtual Data Warehouse And Traditional Data Warehouse

The data warehouse is the strategy of any business intelligence (BI) strategy and architecture. The warehouse is the repository of subject oriented, nonvolatile, historical enterprise data snapshots used in Acrossa's analytical and reporting purposes range. Securely storing data needed for analytical purposes, the data warehouse provides stability, reliability, quality and consolidation. It is a collection of Complete, Correct, Comprehensible and Consistent data. [10]

The functionality of the Traditional Data Warehouse are : [10]

- provides a central repository for information
- requires the development of metadata definitions
- provides data cleansing facilities

While, The functionality of the Virtual Data Warehouse are : [10]

- uses middleware as data hubs, allowing access to the corporate data stored in heterogeneous datasources
- uses middleware to build direct connections among disparate applications
- relies on the creation of an independent metadata definition of the corporate data
- accesses only data in its raw form unless the data access middleware is used to populate OLAP databases directly

III. RESEARCH METHODS

A. Data Source

The data source in the design of the virtual data warehouse is data on sales of materials that can be obtained from the Building Materials Store X. Data obtained include data on goods sales, purchase of goods, suppliers of data and customer data.

B. Research Method

The method and research flow in this study can be described as follows:

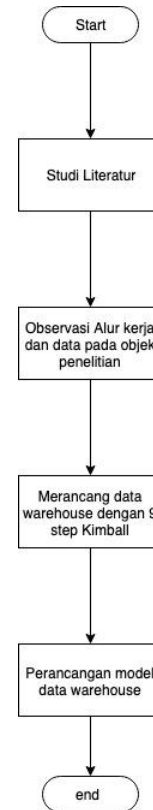


Figure 1. Research Method Flowchart

The study began by conducting a literature study of the supporting theories of the Virtual Data Warehouse and 9 Steps Methodology by Kimball and Roses. The research then continued with observing workflows and data on research objects, namely building materials store X. After getting enough data and knowing how the workflow was done, data warehouse was processed with 9 Steps Methodology by Kimball and Roses and finally the datawarehouse model was obtained as the final result .

IV. ANALYSIS AND RESULTS

A. Virtual Data Warehouse Structure

The Virtual Data Warehouse architecture in this study is as follows:

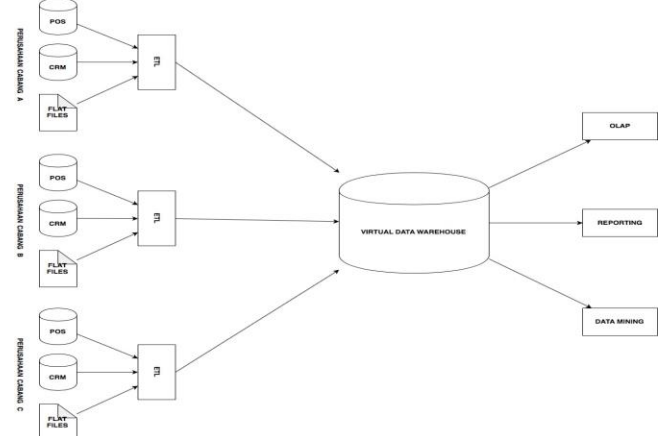


Figure 2. Virtual Data Warehouse Structure

The concept of Virtual Data Warehouse is a platform that becomes a bridge between end users and several data sources. In this study the data source came from various branches of building stores that have different information systems and databases so that the concept of virtual data warehouse greatly assists top management in analyzing the condition of the company and making decisions.

Seen in Figure 2, the structure of the Virtual Data Warehouse serves as a bridge to various data sources which, after going through ETL processes, store their data in a virtual data warehouse. This platform will then provide access to the end user to query so that all data in all branches can be accessed in a data warehouse design.

B. Steps to Design a Virtual Data Warehouse

This study produced a data warehouse design that was built through the process of collecting data and analysis on research objects, namely building material store X. The framework used was 9 Steps Methodology by Kimball & Roses.

1. Choose The Process

At this stage, the process of selecting the data will be entered into the data warehouse. The selected process for this research is

- Process of Selling Goods
- Process of Purchasing Goods
- Process of Paying Customer Receivables

2. Choosing The Grain

The next step is to choose a representation from the fact table. Of the three processes selected three grains were obtained, namely:

- The best-selling item
- Customers shop the most
- The most current customer payments receivables

3. Identifying and confirming the dimensions

This process identifies and connects fact tables with dimension tables. The dimensions that can be identified are:

- Customer Dimensions
- Dimensions of Merchandise

4. Choosing the Facts

This stage determines what information we need in the fact table.

- Sales: ID Sales, Customers, IDs, Total Sales, Total Goods, Customer Payments
- Purchases: id Purchases, idSuplayer, idBars, totalBuy, totalBags
- Item Inventory: Item, Item Enter, Item Out, Time Stock Out

5. Store pre-calculations in the fact tables

Store information that has been calculated in advance to reduce calculation errors for various data.

- Sales Calculation Facts: Amount of Sales

(Number of items x price of goods)

- Customer Payments (period <45 days, 60 days, 90 days)
- Purchase Facts: Amount of Purchases (Amount of Goods x price of Goods)

6. Rounding out the dimensions Table

The identified dimension is made a description that contains structured information about the attributes in the dimension table.

- Time => contains information on the transaction time
- Item => contains merchandise information
- Customer => contains customer information
- Suplayer => contains supplier information

7. Choosing DB duration

Select the duration of the database to be entered into the data warehouse. In this study the data entered into the data warehouse is data for 3 (three) years from the company.

8. Tracking slowly changing Dimensions

It is important to store historical data in data warehouse for analysis needs. Data that changes slowly will be entered into a new record in this study.

9. Deciding the Query priorities

At this stage physical planning is carried out which results in a database scheme.

C. Results of Fact Tables formed

After passing through the processes of 9 Kimball and Roses steps, the results are in the form of a database scheme that has a Star Scheme.



Figure 3. Fact table from the virtual data warehouse

Obtained two fact tables, namely sales facts and purchase facts with four tables of dimensions, namely suppliers of goods, dimensions of customers, time and suppliers.

V. CONCLUSION

In the research design of virtual data warehouse in supporting the information needs of sales, it produces a virtual data warehouse design that functions to facilitate the analysis of sales data in building material stores that have several branches. This study succeeded in designing a virtual data warehouse that implemented the nine step kimball method and then produced a data warehouse schema model

in the form of a star scheme.

ACKNOWLEDGMENT

We, as the authors, would like to express our first gratitude to the Almighty God who made this research complete smoothly. then to the lecturer who guides us and to all parties and colleagues who support us in doing this research.

REFERENCES

- [1] A. Saputra and M. Sudarma, "Designing Data Warehouse for Analysis of Culinary Sales with Multidimensional Modeling – Star Schema Design," *International Journal of Engineering and Emerging Technology*, vol. 3, no. 1, p. 4, 2018.
- [2] C. Ju and M. Han, "Effectiveness of OLAP-Based Sales Analysis in Retail Enterprises," in *2008 ISECS International Colloquium on Computing, Communication, Control, and Management*, Guangzhou, China, 2008, pp. 240–244.
- [3] C. Yu, "Sale Decision Analysis System Design Based on Data Warehouse," in *2016 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, Changsha, China, 2016, pp. 468–471.
- [4] A. H. Kridalaksana, H. Hamdani, and N. Yanti, "Perancangan Data Warehouse Penjualan Uuntuk Data Warehouse Penjualan Uuntuk Perhitungan Laba Rugi (Studi Kasus di Era 5000 Lembuswana Samarinda)," *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, vol. 6, no. 3, pp. 119–125, Jun. 2016.
- [5] D. A. Savitri, F. Yasin, and Y. S. Nugroho, "Implementasi Data Warehouse Sistem Penjualan Batik Di Kampung Batik Laweyan (Studi Kasus Batik Mahkota Laweyan)," p. 12.
- [6] T. R. Ross *et al.*, "The HMO Research Network Virtual Data Warehouse: A Public Data Model to Support Collaboration," *eGEMs*, vol. 2, no. 1, p. 2, Mar. 2014.
- [7] S. S. Kondabolu and J. Nasina, "A Virtual Data Warehouse for Manufacturing Industry," *SSRN Journal*, 2010.
- [8] J. P. McGlothlin, A. Madugula, and I. Stojic, "The Virtual Enterprise Data Warehouse for Healthcare:," in *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies*, Porto, Portugal, 2017, pp. 469–476.
- [9] F. Ahmedy, "Developing Virtual Data Warehouse For Rehabilitation Registry In Sabah, Borneo: Towards Big Data Analytics And Geomapping," p. 8.
- [10] R. Parabia, "Trade and Tricks: Traditional vs. Virtual Data Warehouse," *steps-india.com*.