

Data Center Data Warehouse Development at Z Bali Clinic Using the Kimball Nine-Step Method

I Gusti Ngurah Wira Partha^{1*}, Philipus Novenando Mamang Weking², and Putu Arya Mertasana³

^{1,2} Department of Electrical and Computer Engineering, Post Graduate Program, Udayana University

³ Department of Electrical and Computer Engineering, Udayana University

*Email : kobayakawa.sena62@gmail.com

Abstract— Z Bali Clinic has implemented a clinical information system as an operational system. However, the system applied is only limited to carrying out registration and payment functions. Over time, the number of patients from the Z Bali Clinic is increasing and causes more and more complex patient visit data to be managed.

The data is managed to be used in making reports. However, the process of making the report is still done through manual calculations, causing it to become an obstacle in the data processing and reporting process.

In this study, the process of developing a Data center data warehouse at Z Bali Clinic Using the Kimball Nine-Step Method, where the data warehouse can be used as a center for processing and analyzing outpatient visit data, and to produce reports that can facilitate the executive in analyzing data on outpatient visits as consideration in making decisions at the Z Bali Clinic.

Keywords—Data warehouse, Data Center, Kimball Nine-Step.

I. INTRODUCTION

Z Bali Clinic is one of the health service facilities in the city of Denpasar. Z Bali Clinic has implemented a clinical information system as an operational system. However, the system applied is only limited to carrying out registration and payment functions. Over time, the number of patients from the Z Bali Clinic is increasing and causes more and more complex patient visit data to be managed.

The data needs to be managed and reported to the executive to be analyzed, so that the executive

can find out the progress of the clinic at a certain time period. However, the process of making the report is still done through manual calculations, as well as the need for various report formats that support multidimensional analysis, causing it to become an obstacle in the data processing and reporting process. Therefore, there is a need for a data processing tool that can support the analysis process of outpatient visit data, one of the tools that can be used is a data warehouse.

Data warehouse is a collection of data obtained from various sources used to support management decision-making processes within the company [1]. The data warehouse design method is carried out by applying the 9 steps (Nine-Step Methodology) used by Kimball. The 9 stages are process selection, selection, grain, identification and adjustment, fact selection, pre-calculation storage in fact tables, ensuring dimension tables, selecting database duration, tracking changes from dimensions slowly, determining priorities and query models [2].

Data warehouse based on OLAP provides solutions to problems of data integration, information access, and data analysis that enable users to access information quickly and accurately so executives can make decisions efficiently [3]. OLAP and data warehouses have complementary functions, data warehouses are used to store and manage data, while OLAP is used to convert data in the data warehouse into information [4].

Research related to the development of the data warehouse uses the kimball method the previous steps have been carried out by Nur Ardista, Taufik, Purbandini (2017) [5]. The study was conducted to design and build an OLAP-based data warehouse so that it can be used by the RSUD medical record section in preparing

reports. The data warehouse is built through seven stages, namely analysis, design, ETL process (Extraction, Transformation, and Loading), applying OLAP, testing, exploration for the results of reports and analysis, and evaluation. The design of the data warehouse uses the Nine Step Methodology with modeling in the form of fact constellation schema. The results of the implementation of the data warehouse are OLAP applications that can be used to help the performance of the RSUA medical record section in making reports, both in the form of pivot tables and graphs.

Similar research has also been carried out by Suzana (2013) [6]. The research was conducted to build a data warehouse at the Palembang Bari Hospital in order to meet the information needs of the executive so that it can be used to support decision making so that it can improve performance in handling patients. In this case, the data warehouse is designed to manage outpatient and inpatient data. The design of the data warehouse uses nine step methodology by modeling fact constellation schema consisting of two fact tables, namely outpatient fact tables and fact hospitalization tables. The dimensions used in the outpatient facts are time, clinic, diagnosis, and insurance dimensions, while for the fact the dimensions of hospitalization used are time, clinic, insurance, and room dimensions.

Based on the description of the above problems and several references, this research will develop a Data center data warehouse at Z Bali Clinic Using the Kimball Nine-Step Method, where the data warehouse can be used as a center for processing and analyzing outpatient visit data, as well as for produce reports that can facilitate the executive in analyzing outpatient visit data as a consideration in making decisions at the Z Bali Clinic.

II. THEORETICAL BASIS

A. Data warehouse

The data warehouse is in the form of a database that stores a collection of current data and past data from various sources. Data and information are collected from internal sources and external sources into a collection of data placed on a large storage medium. Data storage is processed into a form of multidimensional

storage and designed for querying and reporting and data mining.

According to W.H. Inmon and Richard D.H, the data warehouse is a collection of data that has the nature of a subject oriented, integrated, time variant, and still on data collection to support the management decision-making process. Based on the understanding expressed by Inmon and Richard, the data warehouse has four characteristics, namely Subject Oriented, Integrated, Time Variant, Non-Volatile [1].

B. Prinsip *Data warehouse*

In principle, the data warehouse is built by collecting data and information from internal sources in the form of operational data and external sources. Furthermore, the data warehouse manager performs grouping and processing it first to then store it into the data warehouse center. This data warehouse manager also regulates data warehouse utilization by users or users. Inmon and Richard also confirmed, "the managers know all too well that integration is needed to deliver on the promises of data warehousing." [7].

In principle, the data warehouse is used to meet the needs of reporting and developing applications, On-Line Analytical Processing (OLAP), and data mining [8].

C. Benefits of a Data warehouse

A number of benefits can be obtained by building a data warehouse in the form of strategic benefits to exceed competitors in the business [9].

1. Ability to access large data.
2. The ability to have consistent data.
3. The ability of fast analysis performance.
4. Knowing that there are repeated results.
5. Finding a defect in the business knowledge or business process.
6. Reducing administrative costs.
7. Give authority to all members of the company by providing information needed so that performance can be more effective.

D. OLAP (Online Analytical Processing)

OLAP is a device that describes technology using multidimensional visualization of a number of data for the purpose of speeding up analysis. In the OLAP data model, information is conceptually described as a cube (cube), which

consists of descriptive categories (dimensions) and quantitative values (measures) [10].

E. The Kimball Nine-Step Method

Kimball's 9-step method or Kimball's steps were introduced by a data warehouse and business intelligence expert named Ralph Kimball (1944) who formulated steps to build and develop a data warehouse. [7]. The nine steps are as follows :

1. Choosing the process.
2. Choosing the grain.
3. Identifying and conforming the dimensions.
4. Choosing the fact.
5. Storing pre-calculations in the fact table.
6. Rounding-out the dimension tables.
7. Choosing the duration of the database.
8. Tracking slowly changing dimension.
9. Deciding the query priorities and the query modes.

III. METHODOLOGY

A. Research Location

The author conducted research at the Z Bali Clinic located in Denpasar, Bali. The author conducts research using observational methods on work processes that occur at the Z Clinic Bali, besides that during the observation process, the author also collects the data needed during the research at that location using the documentation method.

B. Data Sources

Data collection method is a method used to collect data needed in research. The methods used in this study are :

1. Primary Data

The primary data used in this study are observational data from the system and work processes that occur at the Z Bali Clinic.

2. Secondary Data

Secondary data used in this research are books, journals, documentation data from research locations both in the form of documents, images, photographs and electronics, as well as data from various sources that support the basic concept of a data warehouse.

C. Data Collection Techniques

1. Observation

The author made observations at the research location, namely the Z Bali Clinic. The author observes the work process at the research site in making reports of outpatient visits.

The results of several observations made by the author regarding the process of making reports of outpatient visits at the Z Bali Clinic, namely the process carried out is still manual and the need for various report formats that support multidimensional analysis, causing it to become an obstacle in data processing and manufacturing report.

2. Literature

The literature method is a method of collecting data by analyzing literature review as a theoretical basis. The data and information obtained from the literature method come from books, journals and data from various sources that support the basic concept of a data warehouse.

IV. RESULTS AND DISCUSSION

A. Design of Data warehouse

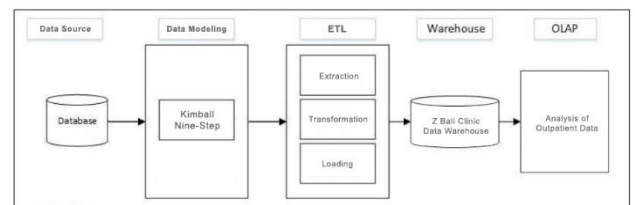


Figure 1. Design of Data center data warehouse at the Z Bali Clinic

B. Data Sources

The data source used is in the form of a database of clinical information systems owned by the Z Bali Clinic.

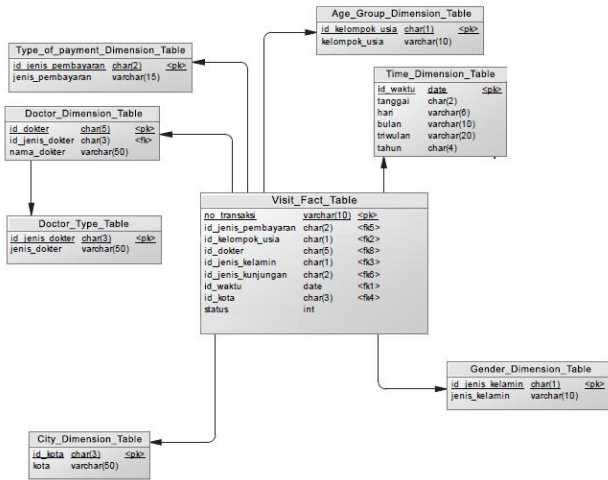


Figure 2. Physical Data Model (PDM) Clinical Information System Database Z Bali Clinic

C. Modeling the Data warehouse

In the data warehouse modeling this method will use the nine steps with the following steps.

1. Choosing the process.

Based on the scope of the research, the process chosen in building this data warehouse is the process of recording data on outpatient visits and the reporting process of outpatient visits at the Z Bali Clinic.

2. Choosing the grain.

Grain is a process for determining what records will represent in the fact table. Grain chosen in designing this data warehouse is an outpatient process. The grain was chosen based on the needs of the report, namely a report containing information on the number of outpatients that can be analyzed based on time, gender, age, city, type of payment, and doctor.

3. Identifying and conforming the dimensions.

There are 6 dimensions used in the data warehouse, namely the dimensions of time, gender dimensions, age group dimensions, city dimensions, payment type dimensions, and doctor dimensions, where this dimension can later be used for the management and reporting process.

4. Choosing the fact.

The fact chosen is the number of outpatient visits at the Z Bali Clinic. Following are the attributes in the fact table of outpatient visits at the Z Bali Clinic.

Table 1. Table of Facts on Visits of Outpatients at the Z Clinic Bali

Attribute	Data Type	Data Length
no_transaksi	Varchar	10
id_jenis_pembayaran	Varchar	10
id_kelompok_usia	Varchar	10
id_dokter	Varchar	10
id_jenis_kelamin	Varchar	10
id_kota	Varchar	10
id_waktu	Varchar	10
status	Interger	

5. Storing pre-calculations in the fact table.

The calculation in the fact table of outpatient visits at Z Bali Clinic is the calculation of the number of outpatient visits, where this calculation shows the number of patients who get services in outpatient units at the Z Bali Clinic.

6. Rounding-out the dimension tables.

This stage is the stage of adding information to each dimension to make it easier to use and understand by the user. The following are the attributes and descriptions of each predetermined dimension table.

Table 2. Dimension Table Attributes and Descriptions

Dimension Table	Attribute	Description
Time	id_waktu	Reports can be made based on day, date, month, quarter, and year
	tanggal	
	hari	
	bulan	
	triwulan	
Gender	id_jenis_kelamin	Reports can be made based on the patient's gender
	jenis_kelamin	

Age_Group	id_kelompok_usia	Reports and analysis can be made based on the patient's age group
	kelompok_usia	
City	id_kota	Reports can be made based on the city of origin of the patient
	kota	
Type_of_payment	id_jenis_pembayaran	Reports can be made based on the type of payment
	jenis_pembayaran	
Doctor	id_dokter	Reports can be made based on the doctor
	nama_dokter	

7. Choosing the duration of the database.

The duration of the database that will be entered into this data warehouse is for two years starting from the beginning of 2017 to the end of 2018.

8. Tracking slowly changing dimension.

Attributes from dimension tables do not always have a fixed value or are relatively static. Changes in attribute values can occur for a long time. The three basic types of changing attributes in dimensions are rewriting attributes, creating new records and creating new columns. Dimensional attributes that may change in the Data center data warehouse at Z Bali Clinic are the dimensions of payment types.

9. Deciding the query priorities and the query modes.

This stage focuses more on the problem of the physical data warehouse associated with index ordering. At this stage the provisions of the queries or reports (reporting) are made to be able to display the data desired by the user. Data center data warehouse at Z Bali Clinic has been sorted by primary key in each table, and this data warehouse was developed to display reports containing information on the number of outpatients that can be analyzed based on time, gender, age, city, type of payment and doctor.

Query results or reporting like this can be used by policy makers or leaders in making strategic decisions.

D. Designing a Star Schema

In the development of Data center data warehouses at the Z Clinic Bali uses a star schema, where the fact table is placed in the center, surrounded by dimension tables. The use of this star scheme is chosen because the form of this scheme is easy to understand and use, making it easier to carry out the process of forming queries. The star scheme produced in this study is the scheme of outpatient visits at the Z Bali Clinic.

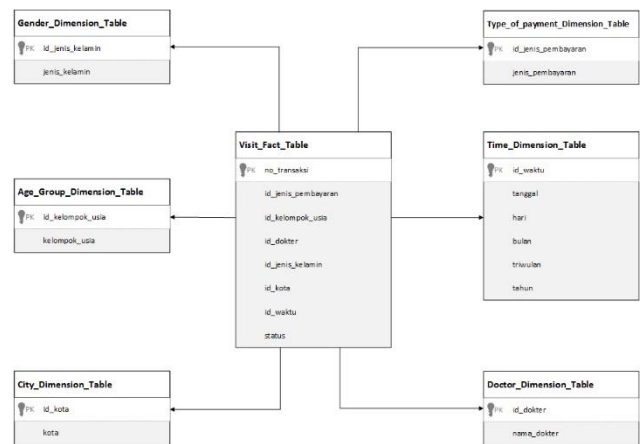


Figure 3. Star Schema Data center data warehouse at Z Bali Clinic

E. ETL Process (Extraction, Transformation, Loading)

After the data warehouse design is completed, the next process is doing the ETL process.

1. Extraction

Data on outpatient visits at the Z Bali Clinic are taken from data that has been stored in the clinical information system and the data taken is in accordance with the requirements in the data warehouse.

2. Transformation

Transformation is done to process outpatient visit data so that it has the format as needed.

3. Loading

Data on transformed outpatient visits are ready to be loaded into the data warehouse.

F. OLAP Process

From the results of the design of the data warehouse, we obtain information on the facts of inpatient visits at the Z Bali Clinic, where data from the ETL process (Transform and Load Extract), are displayed in the form of pivot tables. Pivot tables are designed based on the OLAP data model. OLAP is built on Microsoft Excel with the aim that users can perform multidimensional analysis using pivot tables to produce reports as needed.

G. Application of OLAP

Here are some reports that can be generated from the process of applying OLAP, namely the report on the number of patient visits based on the month and the report on the number of patient visits based on the year.

	A	B	C	D
1	Sum of status	Column Labels		
2	Row Labels	2017	2018	Grand Total
3	Januari	1128	3707	4835
4	Februari	1145	3684	4829
5	Maret	1266	4616	5882
6	April	1395	4886	6281
7	Mei	1030	4580	5610
8	Juni	1274	5099	6373
9	Juli	1189	5129	6318
10	Agustus	1691	4911	6602
11	September	2323	5092	7415
12	Oktober	2909	5288	8197
13	Nopember	3083	5315	8398
14	Desember	3476	5424	8900
15	Grand Total	21909	57731	79640

Figure 4. Report on Number of Patient Visits by Month

	A	B	C	D
1	Row Labels	Sum of status		
2	2017	21909		
3	2018	57731		
4	Grand Total	79640		
5				
6				

Figure 5. Report on Number of Patient Visits by Year

V. CONCLUSION AND RECOMMENDATION

A. Conclusion

Based on the research conducted, conclusions can be drawn are :

1. A data center data warehouse was developed at the Z Bali Clinic using the Kimball Nine-Step method in the form of a star schema. The scheme is designed to use data relating to outpatient visits at the Z Bali Clinic.
2. OLAP has been able to produce several required report formats, and OLAP can display reports relating to outpatient visits at the Z Bali Clinic in the form of pivot tables..

B. Suggestions

Suggestions that can be conveyed to be used in further research are :

1. The scope of this research can be expanded so that it covers various fields in the Z Bali Clinic for example by displaying information about diagnoses and medical records.
2. This research can be further developed by implementing data mining so that the analysis process can be carried out more deeply and patterned.

REFERENCES

- [1] Inmon, W.H. 2005. **Building the Data warehouse 4th Edition**. Indianapolis: Wiley Publishing, Inc.
- [2] Rianto, Cucu Hadis,. 2017. **Perancangan Data warehouse Pada Rumah Sakit (Studi Kasus: BLUD RSUD Kota Banjar)**. Jurnal Siliwangi Vol.3. No.2, 2017 ISSN 2477-3891 Seri Sains dan Teknologi.
- [3] Sheta, O. E., & Eldeen, A. N., 2013. **The Technology of Using A Data warehouse to Support Decision-Making in Health Care**. International Journal of Database Management Systems , 5 (3), 75-86.
- [4] Reddy, G. S., Srinivasu, R., Rao, M. P., & Rikkula, S. R., 2010. **Data Warehousing, Data Mining, OLAP and OLTP Technologies are Essential Elements to Support Decision-Making Process in Industries**. International Journal on Computer Science and Engineering , 2 (9), 2865-2873.
- [5] Nur Ardista, Taufik, Purbandini. 2017. **Rancang Bangun Data warehouse Untuk Pembuatan Laporan dan Analisis pada Data Kunjungan Pasien Rawat Jalan Rumah Sakit Universitas Airlangga Berbasis Online Analytical Processing (OLAP)**. Journal of Information Systems

Engineering and Business Intelligence Vol. 3, No. 1, April 2017.

- [6] Suzana, M., 2013. **Analisis dan Perancangan Data warehouse Rumah Sakit Umum Daerah Palembang Bari.** Skripsi, Universitas Bina Darma, Program Studi Teknik Informatika.
- [7] Kimball, Ralph and Ross, Margy. 2002. **The Data warehouse Toolkit Second Edition.** John Wiley and Sons, Inc.
- [8] Mulyana JRP. 2014. **Pentaho: Solusi Open Source Untuk Membangun Data warehouse.** Yogyakarta: C.V. Andi Offset.
- [9] Nolan, Sean And Huguélet, Tom. 2000. **Microsoft SQL Server 7.0 Data Warehousing Training Kit.** Microsoft Prees.USA.
- [10] M.A.B.J. Dewa, Muhammad. N. 2013. **Penerapan OLAP untuk Monitoring Kinerja Perusahaan.** Seminar Nasional Sistem Informasi Indonesia.