The data warehouse for down payment administration in the Constitutional Court of Republic of Indonesia

Abba Suganda Girsang[1], Sani Muhamad Isa[1], Nanda Adytiansya[1], OeiKurniawanUtomo[1], Janner Simarmata[2]

[1]Computer Science Department, BINUS Graduate Program-Master of Computer Science, Bina Nusantara University,Jakarta, Indonesia 11480, Email : [2]Universitas Negeri Medan, Medan, Indonesia

agirsang@binus.edu, sani.m.isa@binus.ac.id, nanda.adytiansya@binus.ac.id, Oei.Kurniawan@binus.ac.id, jannersimarmata@unimed.ac.id

Abstract. The Constitutional Court of The Republic of Indonesia is an institution which has authority to judge and examine constitutions. In this institution, each component has responsibility to manage their own finance and budgeting. Based on this case, they need a solution to summarize the amount and absorption of the budget in every component. This paper proposes a solution to design a data warehouse model to process and analyse the down payment administration data. The nine steps design methodology for developing data warehouse is implemented in this work. This paper aims to bring some benefit for The Constitutional Court, especially in form of information about down payment absorption for every component each year. The result is presented in form of pdf report and informative dashboard using data warehouse tools.

1. Introduction

The Constitutional Court of the Republic of Indonesia is a state institution that performs an independent judicial authority to judge, examined, and review constitutionality under constitution. It implements a modern court and trusted principle since 2007. This institution was established based on 1945 constitutions, clause 24C.

Data Warehouse contains of various aspects that exist within the organization [1] [2]. This paper explains the implementation of data warehouse to store analytical data about down payment administration in every events and units in constitutional court. The data warehouse provides relevant information to management and executives to make decisions [3]. All of the payment detail has been recorded by existing web application that has been used by all departments. The employees who request approval for the invoice and approve it are recorded to the system. The payment divided to 9 departments and the amount of budget are not same, depend on every departments.

As the size of the data warehouse increases the value it brings to an organization tends to increase [4]. There is a business requirement to manage this down payment administration data so the annual amount and absorption for every component, output, activity, and program can be retrieved easily. The problem is, the initial budget for every component is not stored in a dedicated table. The budget table stores the remaining budget for every transaction after an invoice is approved. So, in order to calculate the absorption (in percentage), a complicated SQL query must be run to sum all transaction amount, while calculating the initial budget manually. The existing procedure to generate the required report is done using painful SQL

query, take a lot of time, and take a bunch of resources from the existing running web application server.

This paper proposes a data warehouse model of down payment administration in the Constitutional Court so that the report can be processed fast and all the information can be seen on an informative dashboard.

2. Related Work

In data warehouse (DW) scenarios, ETL (extraction, transformation, loading) processes are responsible for the extraction of data from heterogeneous operational data sources, their transformation (conversion, cleaning, normalization, etc.) and their loading into the DW [5]. A data warehouse is used a central repository of data of medium and large business organizations [6]. Data warehouse is a database that contains data from multiple operating systems that are integrated and structured, so it can be used to support analysis and decision-making process in business [7].

The category of tools which are responsible for extraction of data from several source systems, their cleansing, transformation and inserting them into a data warehouse are called ETL tools [8]. The most important component of the data warehouse is ETL process. ETL process extract and integrate data from diverse homogeneous and heterogeneous sources [9]. Extract is a process of choosing and taking some data from any source, then read it. Transform is a process of clearing and convert the data become something to be useful for any requirement of data warehouse. Load is the last process which has a function to input data with running a script periodically [10].

3. Proposed Method

There are nine steps to develop datawarehouse as follows.

Choosing the process. The process used for case study of the data warehouse model is the invoice approval administration flow. The invoice approval request flow is shown in Figure 1.
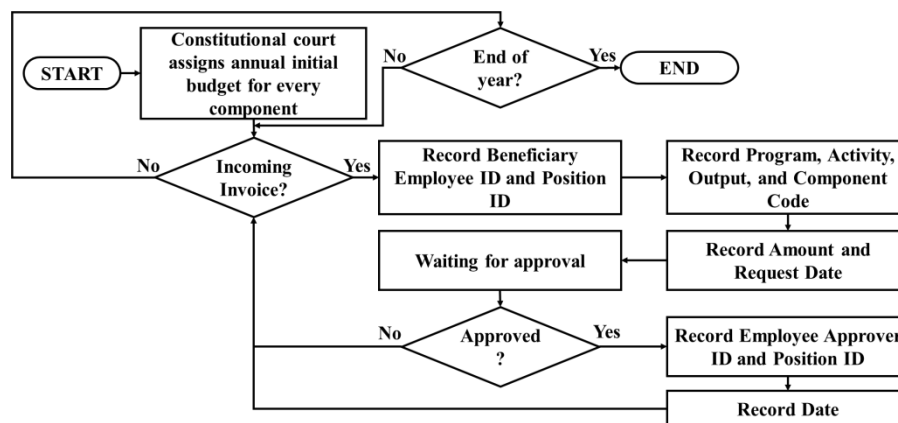

Figure 1. The Invoice Approval Request flow

The system will record the employee's ID (as beneficiary account), employee's position, program code, activity code, output code, component code, amount code, and date from an incoming invoice approval request to invoice table. If the invoice request is approved, the system will record the employee's ID (as source account), employee's position, date attributes to invoice table as update operation. The invoice approval request transaction is selected as

the case study for building the data warehouse to fulfill the business requirement. The invoice table is selected as the fact table for the data warehouse model.

Choosing the grain. Every approved transaction in invoice approval request is selected for the grain of this case study. The reason behind is, a detailed transaction with component code, beneficiary, and source account information will make the grouping process of aggregate function easy.

Identifying and conforming the dimension. The dimension tables of the invoice approval request model are (a) source account, which stores the employee ID and name information as approver, (b) source position, which stores the employee position information as approver, (c) beneficiary account, which stores the employee ID and name information as request maker, (d) beneficiary position, which stores the employee position information as request maker, (e) component, which stores the program, activity, output, and component, and (d) date, which stores the day, month, and year information. Those dimension tables are selected based on the business requirement and provide context to the analysis on top of the star schema.

Choosing the fact. Based on the business requirement and selected process, the invoice table is chosen to be the fact table. It contains all foreign key to the dimensional table and fulfilled the grain model that is every approved invoice approval request.

Storing pre-calculation in the fact table. The down payment absorption, in percentage, for every transaction can calculated by divided the amount to the initial budget for every component. As mentioned earlier, the initial budget for every component is not stored in a dedicated table. So, the initial budget can be retrieved by searching the oldest record for every component. The absorption calculation process is done in the transform step in the ETL process.

Rounding out the dimensions table. Every dimension's primary key is reflected to a surrogate key, which has integer data type.

Choosing the duration of database. The ETL process will be processed monthly. So, the data available for analysis is now minus 1 Month. For this case study, the 2017 data from January to December is used for analysis.

Tracking slowly changing dimension. The slowly changing dimension (SCD) type 2 is implemented in this work. The tables that possibly change is the component table. The old data will not be deleted and the new data can easily be added by creating new surrogate key.

Decide the physical design. The data warehouse server is designed to run on separated machine from the existing application's machine. Hence, the performance of the existing application will not be interfered by the data warehouse analytics process.

4. Result
4.1. Star Schema
After implementing the nine-step design methodology, the star schema of down payment administration system is shown in Figure 2. The facts contain analysis indicators, while a dimension includes analysis parameters organized in hierarchies from the lower granularity (most detailed) to the higher granularity (most general) [11].
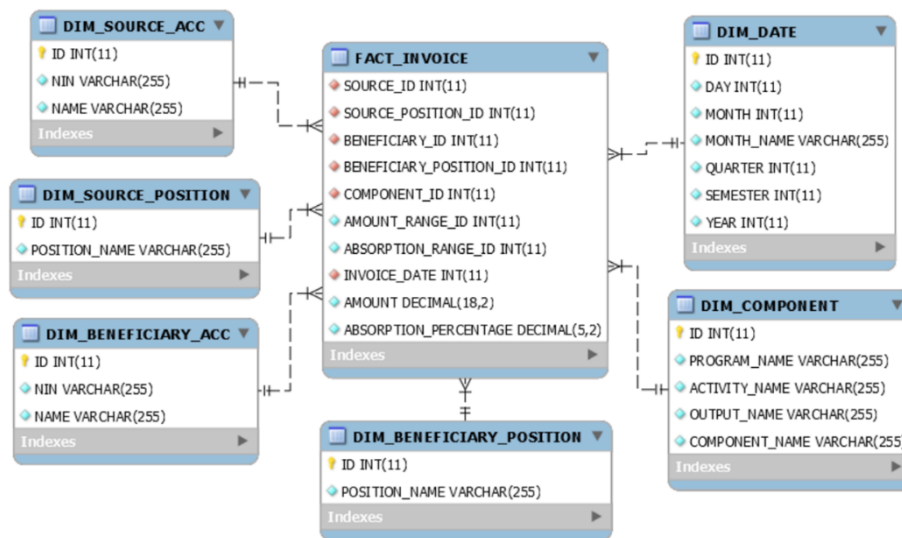
Figure 2. The Star Schema

4.2. The ETL Process

The next step of the data warehouse design is implementing the ETL process based on the nine-step design methodology using Pentaho. In this work, the ETL process is always began with either load/read data from existing application's database or by generate certain rows (e.g. to generate date dimension).
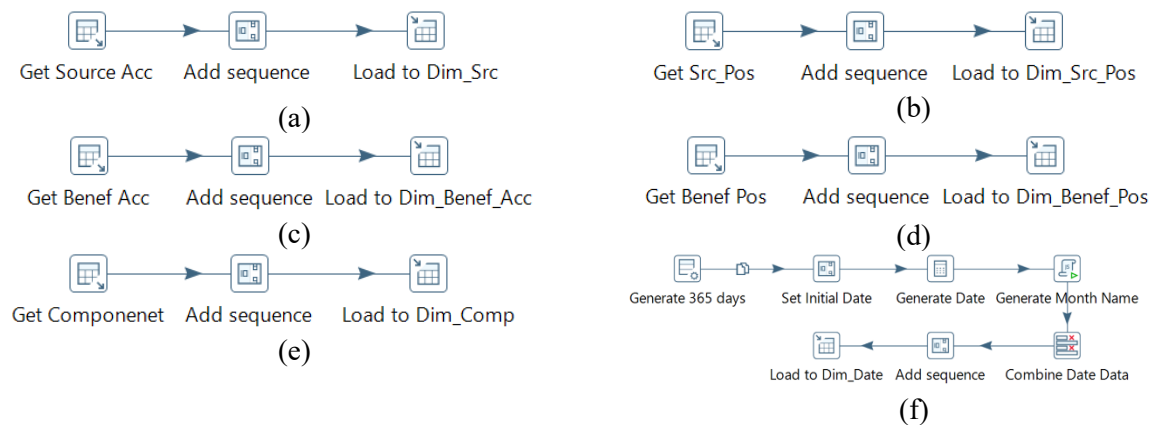

Figure 3. ETL Process of (a) employee source account, (b) employee source position, (c) employee beneficiary account, (d) employee beneficiary account, (e) program, activity, output, and component, (f) invoice date
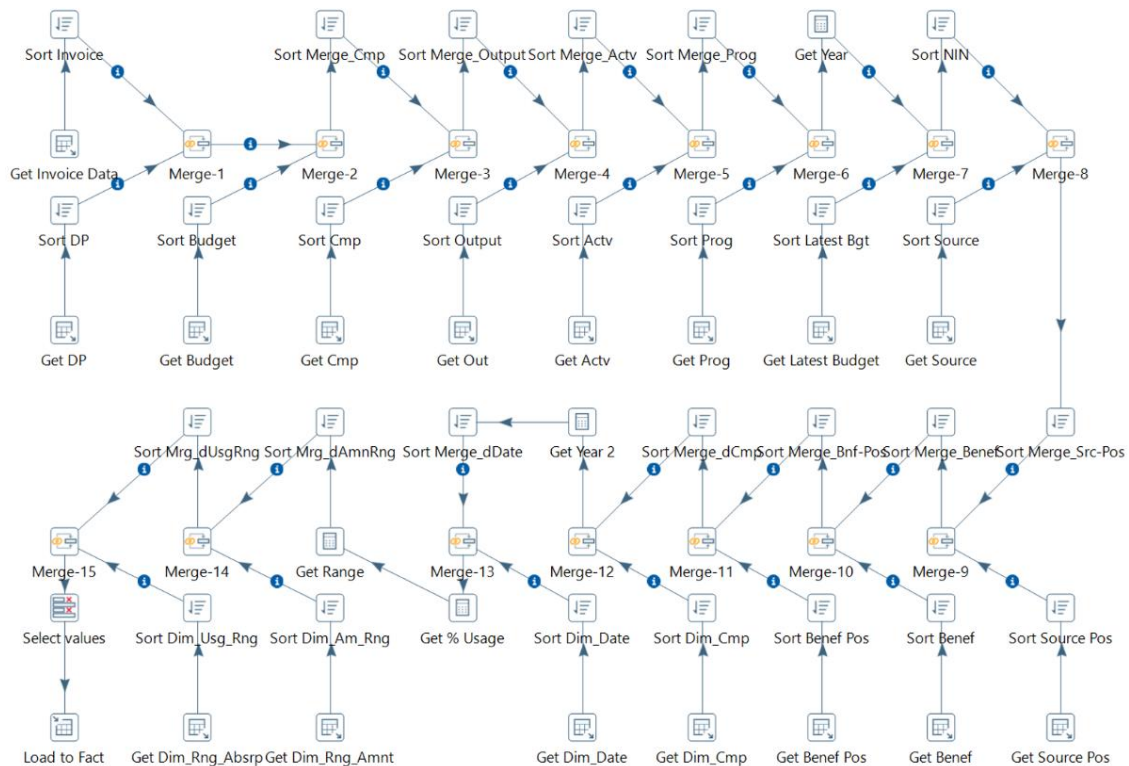
Figure 4. ETL process of invoice fact

The ETL process of the fact table is shown in Figure 3. The process begins with loading the data from down payment and invoice table to get the source data, beneficiary data, amount, and invoice date. Then, it will get the program, activity, output, and component name from program, activity, output, and component table. Next, the latest budget is retrieved by sorting the budget table descending by its date. After loading from all these data from existing application's database, the process continues with joining all data with the dimension table to get the surrogate key. Then, the absorption is calculated in the end.

The OLAP Cube design can be used to make the analytics in form of report and dashboard [12]. The report and dashboard show the data of total amount and absorption for every component based on the selected program, activity, and output as report parameters. The x-axis shows the component, while y-axis shows the amount or the absorption.
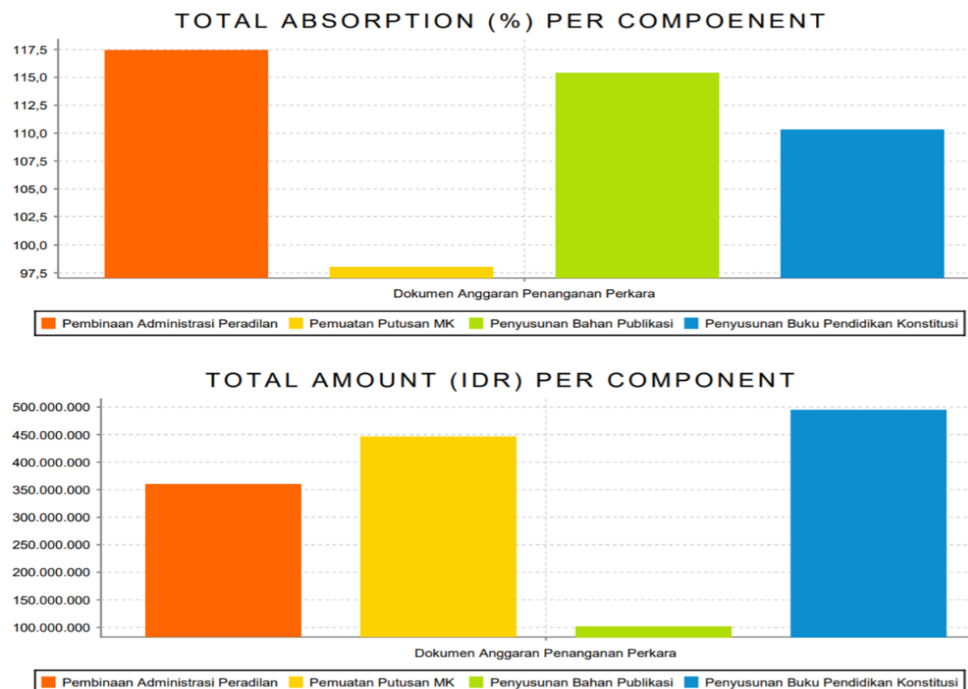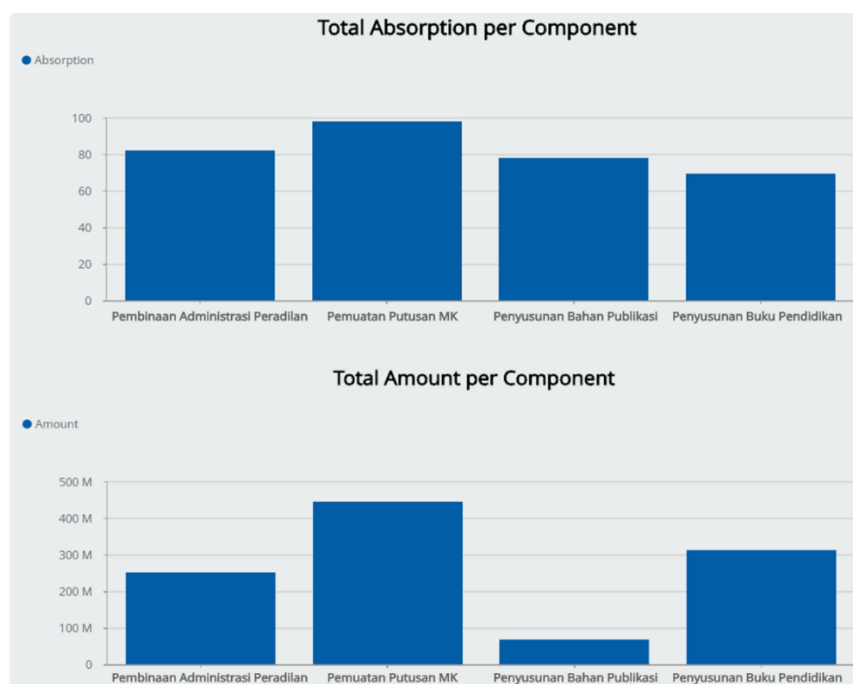
Figure 5. Pdf report

Figure 6. Dashboard

Figure 5 shows the report generated in pdf, while Figure 6 shows the dashboard manager. From the report and the dashboard, the information shown can be used to make a summary of total amount and absorption per component. The month name can be selected to make a month filter in the amount and absorption aggregation process.

5. Conclusions

By implementing the data warehouse design for down payment administration in The Constitutional Court, the amount and absorption for every component can be monitored. The dashboard can give a visualization of the data gathered from the invoice table from existing web application. The report can be generated faster, easy, and more resource-friendly.

For the future research, the data can be processed to show the amount and absorption for every component every month, so the management can predict the future budget for specific component.

Reference

[1]   Abba Suganda Girsang, Candrauji Wira Prakoso. "Data Warehouse Development for Customer WIFI Access Service at a Telecommunication Company." *International Journal on Communications Antenna and Propagation* 7 (2017): 114-124.

[2]   Andersen, Ove, Christian Thomsen and Kristian Torp. "SimpleETL: ETL Processing by Simple Specifications." *CEUR* (2018).

[3]   Kimball, Ralph and Margy Ross. *The Data Warehouse Toolkit Second Edition The Complete Guide to Dimensional Modelling*. Toronto: Robert Ipsen, 2002.

[4]   Lohiya, Akshay S., et al. "Optimize ETL For Banking DDS : Data Refinement Using ETL Process For Banking Detail Data Store(DDS)." *Imperial Journal of Interdisciplinary Research (IJIR)* 3.3 (2017): 1839.

[5]   Luján-Mora, Sergio, Panos Vassiliadis and and Juan Trujillo. "Data Mapping Diagrams for Data Warehouse Design with UML." *Springer* (2004): 191-204.

[6]   Pandya, Bhadresh and Dr. Sanjay Shah. "Query Optimizer for the ETL Process in Data Warehouses." *IJSRSET* (2015): 329-333.

[7]   Prakash, Gajare Harikishan and Prof. S. P. Rangdale. "ETL Data conversion: Extraction, transformation and loading Data conversion." *International Journal Of Engineering And Computer Science* (2017): 22545-22550.

[8]   Rahman, Nayem. "A Simulation Model for Application Development in Data Warehouses." *International Journal of Operations Research and Information Systems* 9.1 (2018): 66-80.

[9]   —. "Enterprise Data Warehouse Governance Best Practices." *International Journal of Knowledge-Based Organizations* 6.2 (2016): 21-37.

[10] Ravat, Franck, Jiefu Song and Olivier Teste. "OLAP Analysis Operators for Multi-State Data Warehouses." *International Journal of Data Warehousing and Mining* (2016): 20-53.

[11] Talib, Ramzan, et al. "A Multi-Agent Framework for Data Extraction,Transformation and Loading in Data Warehouse." *International Journal of Advanced Computer Science and Applications* 7 (2016): 351-354.

[12] Vassiliadis, Panos, Alkis Simitsis and Spiros Skiadopoulos. "Conceptual Modelling for ETL Processes." *Dolap* 14 (2002): 17.