

Online Analytical Processing (OLAP) for Disaster Report

Kadek Ary Budi Permana^{1*}, Dewa Ayu Putri Wulandari², and Putu Arya Mertasana³

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

³Department of Electrical and Computer Engineering, Udayana University

*Email: arybudi1@gmail.com

Abstract—Data analysis of disaster is very important for supporting in making decision to cope with the next disaster or prepare if the disaster will happen again. Online Analytical Processing (OLAP) becomes solution to get analysis by using certain analysis that supplied in OLAP. To supply the data analysis in which done by OLAP so the data as the representative in multidimensional model that manages the data in cube form. Warehouse data can be called as the place from source data that used to make easier in analysis information process in it because concept of data dimensional offered. The planning data warehouse is basically by source data that gotten and decided by the description of needs from the users. OLAP in this disaster report success in implementation with helping data warehouse and imply snowflake schema to dimensional data first.

Keywords—data warehouse; disaster; online analytical processing; snowflake schema.

I. INTRODUCTION

Disaster is the event that causes the hard to get, lose out or even threaten the life. Disaster can cause death, the environment damages, financial damages, and influence physiologist human being. The disaster can be caused by natural disasters, non-natural and human errors. Data disaster analysis is very important to get support the next making decision. The decisions can be some steps to represent of disaster that will happen, prepare it well if it will happen disaster again, and etc.

On-line Analytical Process (OLAP) can be the interesting solution to present disaster data in making report that easy to be analysis by the users. OLAP is multidimensional analysis used as support system decision to find the information in which take from the big data base [1]. OLAP is also used to see the fact from the data based on some dimension. To supply data analysis done by OLAP so the data is represented in multidimensional model that manages the data in cube form [2]. Based on the needs, data is implied into warehouse data with the certain schema implementation.

Warehouse data is the concept that make easy in cure the data history from the transactional process. This concept can guarantee information, it also guarantees access faster and make easy in analysis as supporters in running process [3]. Warehouse data can be called the place source data that can be used as to make easy in analysis information process inside because the concept of dimensional data that offered.

The implementation certain schema in data warehouse model is very important. The model warehouse has certain model such as star and snowflake schemas. Snowflake schema has advantages where this schema implements the sub dimension table beside dimension table that all be integrated

with the fact support the information that can be shown and get information more detail [4].

The aim of this paper is to show how the data warehouse to processing data the disaster in form data dimensional by implementation snowflake schema and present with OLAP technology. The choosing of snowflake schema based by the data collection and decided by description the needs from the users.

In this part of this paper is ordered as the following such as in part 2, we shortly explain some literatures reviews that have relation with this paper. In part 3 we describe the data and planning from warehouse data. in part 4 we show how OLAP is able to analyze the data with some basic operation. In part 5 is the conclusion from our results researches.

II. LITERATURE REVIEW

A. Online Analytical Processing (OLAP)

OLAP is able to solve the problem in mechanism the processing data to know some information from some view point [5]. Technology OLAP can be more flexible to analyze the collecting multidimensional data and fast find the data in certain rules [6]. OLAP has some basic operational to do analysis to the dimensional data, such as [7]:

- 1) Roll-Up: Navigation the data from the low data to the higher data.
- 2) Drill-down: Navigation the data from the high level to lower data.
- 3) Slicing: Describe the choice data as long as single dimation that show as the table.
- 4) Dicing: Describe the data choice based on the showing as sub cube.

B. Data Warehouse

Data warehouse is the system that taken and process the data from data source periods in forming dimensional data store (DDS) [8]. Data warehouse patch data that has characteristic history from the process business transaction that happens in source data.

There are some kinds of transactional data that processed into data warehouse as following [9]:

- 1) Record Data.
- 2) The data is being in environment operational department.
- 3) The last data changing is being in environment department since the last process into data warehouse.

The engineering data warehouses are variations depend on the specific each organization. In deciding the engineering model, there are three main components, as following [10]:

- 1) Acquisition from the OLPT system and other resources;

- 2) Warehouse data its self and system data stock management.
- 3) System analysis and presentation from the warehouse data.

C. Snowflake Shema

There are some schemas in the modeling warehouse data. These schemas describe how fact table connected to dimension tables. In the star schema is described as the star where the fact table is in the middle and around by the dimension tables. While snowflake schema is probably dimension table has sub dimension table. The sub dimension tables can minimalist the data that overload. The implementation snowflake schema in some analysis can run better than the other schema model [8]. The advantage in this analysis is gotten because in the snowflake schema get more information that can be used the data in processing of making decision [4].

D. ETL

There are 3 steps in ETL process namely extract, transform, and load. The extract process is reasonable into interact the data from source data. Every source data has specific characters that difference that be managed. The specific characters are such as the management business system in difference data. In the transformation steps tend make some cleaners and arrangement to the data that enter in getting the right data, complete, consistent, and not ambiguity. The last steps namely load where the steps of this data written into structures dimension that the next can be accessed by the last users [11].

ETL process could not be separated in the planning of data warehouse. In this development there are some kinds of ETL, as following [12],

- 1) ETL Traditional with data based supposed as resource entry data and usually operate in batch processing mode
- 2) Near real time ETL that process the data with reduce refresh time to keep the quality data in warehouse data process.
- 3) Real time ETL does process data by real time from sources into target system that signed by high speed.

III. METHODOLOGY

A. Data Resource

Data resource is the Indonesian Disaster Data Information that be owned by the Safety Disaster National Foundation (BNPB). The resource data used year 2015 till 2018. The data resource contents information about kind of disaster, place, and the effect of disaster.

B. Need Information

In this step, describe about information that needed by the users. The information is used based on the collecting data, namely,

- 1) Disaster Report based on the period
- 2) Disaster Report based on the district
- 3) Disaster Report based on kind of disasters
- 4) Disaster Report based on the effect of caused disaster

C. Design Snowflake Schema

The data dimensional is formed based on the need information that has been described before. The data dimensional is formed by implementation the snowflake schema. The schema is implied by forming fact table and also the other tables as dimensional table and also sub dimensional table.

Table 1 Show the using dimension table.

TABLE I. DIMENSION TABLE

Dimension table	Explanation
dim_disaster	Information is about kind disaster that ever happened such as flood, abrasion, erosion, forest burning, estate burning, social conflict, and other.
dim_impact	Information is about effect that caused by disaster such as the victims death / lost, victims evacuation, damage houses, damage public facility and other.
dim_date	Information is about the time of disaster that can be time, date, day, month, year, and quarter. . .
dim_regency_city	Information is about region or city as the place of disaster happening.
dim_province	Information is about province as the place of disaster happening.

The table fact disaster is the table that the fact in which consisting information about the disaster based on the dimension. Figure 1 shows the snowflake schema from the fact and dimension table.

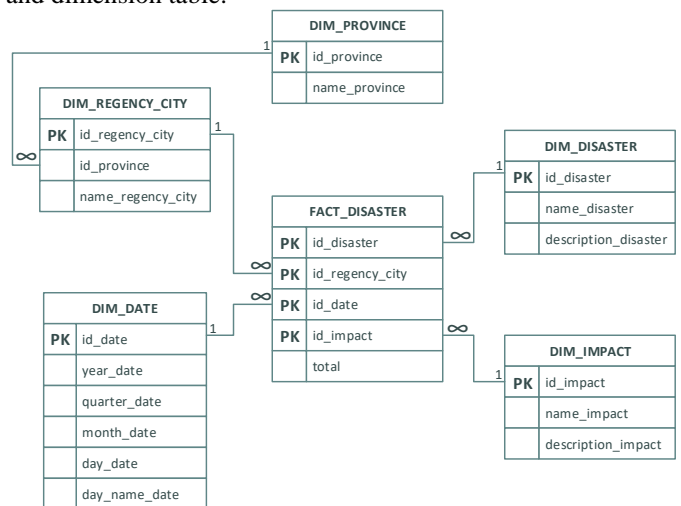


Fig 1. Snowflake Schema

IV. RESULT AND DISCUSSION

ETL process is from sources data to data warehouse into form dimensional data (snowflake schema) has succeeded to be done. Based on the need information that has definite in the beginning so the report which needed can be represented by using OLAP. The followings are some examples report analysis with use basic OLAP report.

A. Roll-Up Operation

This operation increases the level hierarchy from dimension. This operation is perhaps to result information shorter. As the examples can be seen in figure 2 shows that the number disaster happened based on the kind of disaster and month. By using roll-up operation is so the level hierarchy can be increased by seen the number kind of disaster and year as shown in figure 3.

Disaster	Year	Month	Totals		
ABRASION OR TIDAL WAVE	2015	April	1		
		January	4		
		June	1		
		March	1		
EARTHQUAKE	2015	December	2		
		February	1		
		July	3		
		June	11		
		March	1		
		May	2		
		November	5		
		September	1		
		FLOOD	2015	April	75
				August	7
December	56				
February	107				
January	125				
July	5				
June	35				
March	32				
May	35				
November	38				
September	16				
Totals					564

Fig 2. Disaster Report based on kind of disaster, year, and month

Disaster	Year	Totals
ABRASION OR TIDAL WAVE	2015	7
EARTHQUAKE	2015	26
FLOOD	2015	531
Totals		564

Fig 3. Disaster Report based on kind of disaster and year

B. Drill Down Operation

Drill down operation is the opposite from roll-up operation. This operation reduces level hierarchy from dimension. This operation is perhaps to result information more detail. As the example it can be seen in figure 4 show that the number disaster that happened based on the kind of disaster and province. By using drill down operation so the level hierarchy can be reduced

based on the kind disaster, province, and regency / city as shown in figure 5.

Disaster	Province	Totals
ABRASION OR TIDAL WAVE	BALI	2
EARTHQUAKE	BALI	1
FLOOD	BALI	20
Totals		23

Fig 4. Disaster Report based on kind of Disaster and Province

Disaster	Province	Regency / City	Totals
ABRASION OR TIDAL WAVE	BALI	JEMBRANA	2
EARTHQUAKE	BALI	KOTA DENPASAR	1
FLOOD	BALI	BADUNG	7
		BANGLI	1
		BULELENG	5
		JEMBRANA	4
		KARANGASEM	2
		TABANAN	1
		Totals	

Fig 5. Disaster Report based on kind of Disaster, Province, and Regency / City

C. Slicing and Dicing Operation

Slicing does the data choice as long as certain dimension. It is perhaps the users to do data selection to certain dimension. As the example the users do selection in dim impact to show the victim death as shown in figure 6.

Dicing does the operation by cutting cube so information that result more specific. Dicing divinites sub cube with choose two or more dimensions cube. As the example, the users use dim disaster, dim date, and dim impact to know the disaster that had happened and the effect that caused from that disaster. For more detail can be seen in figure 6.

Disaster	Year	Victims Death	Totals
ABRASION OR TIDAL WAVE	2015	0	
		1	1
EARTHQUAKE	2015	0	
		0	
FLOOD	2015	1	20
		2	14
		3	6
		4	8
		Totals	

Fig 6. Disaster Report based on kind of Disaster, Year, and Number of Victims Death

D. Graph View

The results of operations on OLAP can be displayed in graphs to make it easier to read data and find trend data. Figure 7 shows the graph view of disaster report based on kind of disaster and year.

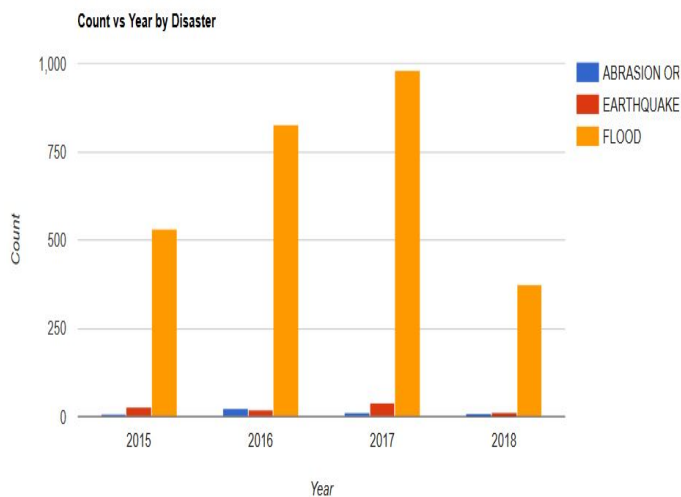


Fig 7. Graph View

V. CONCLUSION

Online Analytical Processing (OLAP) of the disaster report succeed to imply with building data warehouse and imply snowflake to dimensional data first. By using basic operation to OLAP, the needs the users to do analysis disaster data can be fulfilled and become more flexible. OLAP operation results can also be displayed in the graph. Graph view makes it easier for users to find trend data. OLAP will supply supporting in making decision by the users to prevent the next disaster.

REFERENCES

- [1] S. Tohir, K. Kusriani, and S. Sudarmawan, "On-Line Analytic Processing (OLAP) Modeling for Graduation Data Presentation," in 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), 2017, pp. 132–135.
- [2] G. Hayardisi, I. S. Sitanggang, and L. Syaufina, "Data warehouse and web-based OLAP for hotspot distribution in Indonesia," in 2nd Conference on Data Mining and Optimization, 2009, pp. 1–4.
- [3] N. Kumaladewi and Y. Sugiarti, "Design analysis of data warehouse for lecturer performance evaluation (Case study: Faculty of science and technology UIN Jakarta)," in 4th International Conference on Cyber and IT Service Management, 2016, pp. 1–6.
- [4] A. Dahlan and F. W. Wibowo, "Design of Library Data Warehouse Using Snowflake Scheme Method: Case Study: Library Database of Campus XYZ," in 7th International Conference on Intelligent Systems, Modelling and Simulation Design, 2016, pp. 318–322.
- [5] P. Widiadnyana, M. A. Maricar, I. N. Arnawan, and S. Ariyani, "Analysis of Data Warehouse for Transactional Analysis Methods Online Analytical Processing (OLAP) at Company XYZ Silver," International Journal of Engineering and Emerging Technology, vol. 2, no. 1, pp. 72–75, 2017.
- [6] J. Shi, J. Yao, and M. Liu, "Research on Database Audit Scheme Design of Life Insurance Industry Based on OLAP Technology," 2016.
- [7] K. Dhanasree and C. Shobabindu, "A survey on OLAP," in IEEE International Conference on Computational Intelligence and Computing Research, 2016.
- [8] V. Rainardi, Building a Data Warehouse with Examples in SQL Server. 2008.
- [9] W. H. Inmon, Building the Data Warehouse, Fourth Edition. 2005.
- [10] C. E. Poenaru, D. Merezeanu, R. Dobrescu, and E. Posdarascu, "Advanced Solutions for Medical Information Storing: Clinical Data Warehouse," in the 6th IEEE International Conference on E-Health and Bioengineering, 2017, pp. 37–40.
- [11] S. H. A. El-Sappagh, A. M. A. Hendawi, and A. H. El Bastawissy, "A proposed model for data warehouse ETL processes," J. King Saud Univ. - Comput. Inf. Sci., vol. 23, pp. 91–104, 2011.
- [12] K. V. Phanikanth and S. D. Sudarsan, "A Big Data Perspective of Current ETL Techniques K," pp. 330–334, 2016.