International Journal of Information Technology and Knowledge Management
December 2012, Volume 6, No. 1, pp. 61-64

DESIGN AND IMPLEMENTATION OF ACCESS THE CONTENTS IN THE DATA WAREHOUSE

Nirmal Sharma[1] and S.K. Gupta[2]

Abstract: This paper describes the design & implementation of data analysis tools. Many tools & database systems use the decision making process. The increasing processing techniques are more powerful analytical tool [4]. There are a number of data analysis tools that any organization should consider a data warehouse, which can be the critical analytical tool for maximizing the organization's investment in the information it has collected and stored throughout the monitor. In this work we developed various tools, which can be used as the analytical applications. Software managers need to understand the rationale and importance of data warehouses because they may need to design and implement [2].

More generally, data warehouse is a collection of decision support technologies, aimed at enabling the knowledge worker, such as executive, manager, and analyst, to arrive at better and faster decisions. Business metadata may also serves as documentation for the data warehouse/business intelligent system. Users who browse the data warehouses are primarily viewing the business metadata.

The data warehouses are provided for storage, functionality and responsiveness to queries beyond the capabilities transaction-oriented databases. Data warehousing are using many analysis tool. In this paper we use online analytical tools with excel. It's provides functionality, reusability of these tools enhanced the application's reliability, analysis techniques & reduced the cost/time of application [1].

In this research paper use Online Analysis Tools in Excel Spreadsheets. Because it tools have unlimited techniques. Whatever you can generate report any manner. This tool chooses it because we can do online and offline. If any organizations use this tool easily activates reports. The data warehouse must hold all our accounting data. Restate the problem in quantifiable terms, like "The data warehouse must handle the current 30 GB of accounting data including all metadata and replicated data with an expected 30% growth per year".

Keywords: Meta data, Business Intelligence, Data Mining, Data Warehouse, Information Processing, Online Analytical Processing & Tools.

1. INTRODUCTION

A data warehouse is a subject-oriented, integrated, non-volatile and time-variant collection of data in support of management's decisions.

Business metadata is content from the data warehouse described in more user-friendly terms. The business metadata tells you, what data you have, where it comes from, what it means and what its relationship is to other data in the data warehouse.

We are solving the modern business problems like market analysis and financial forecasting requires query-centric databases schemas that are array-oriented and multidimensional. These business problems are specified by the need to manipulate large numbers of records from very large data sets.

Once the selected cuboids have been materialized, it is important to take benefit of them during query processing. Finally during clean, load and refresh, the materialized cuboids should be updated efficiently.

[1] Aryan Institute of Technology, Ghaziabad (UP). *E-mail: nirmal1709@rediffmail.com*

[2] B.I.E.T. Jhansi (UP), *E-mail: guptask_biet@rediffmail.com*

2. LITERATURE BACKGROUND

Data is the information system. The departments create, store and provide information data in its business contexts- what is the business need?

Company is betting on technology to provide the tools it needs to leverage reams of information tucked away throughout the giant retail chain. A data stores for a large amount of corporate data.

When we survey the market so that analyze the business questions are reflected in the fact- like "How much revenue did the new product generate and by period & region wise, compared with the plan?"- One way to set the multidimensional data model is to view it as cuboids & table form.

Table1
Base and Summary of Data Store

PID	*Product*	*Specification*	*Regions*	*Year*	*Sales*
1001	Suit	Standard	Noida	2009	39999.00

Table Cont'd

Table 1 Cont'd

1002	Coat	Large	Ghaziabad	2009	46789.00
1003	Trouser	Standard	Gr. Noida	2009	69999.00
….	….	….	….	….	….
4001	Garments	All	All City	2010	1,26,999.00

This tuple is at a more general level of abstraction than the tuples 1001, 1002 & 1003. The specification value has been generalized to All, so that the corresponding period value is March 2010, rather than just March 15 or 30, 2010. The special value All is used to represent subtotals in summarized data [8].

As an example, a Garment company may need to analyze their profits for the last three years. Looking at the data from the last three years will allow the user to view the transformations the company has gone through. Unfortunately, only having a current view will not allow the user to get the information they need.

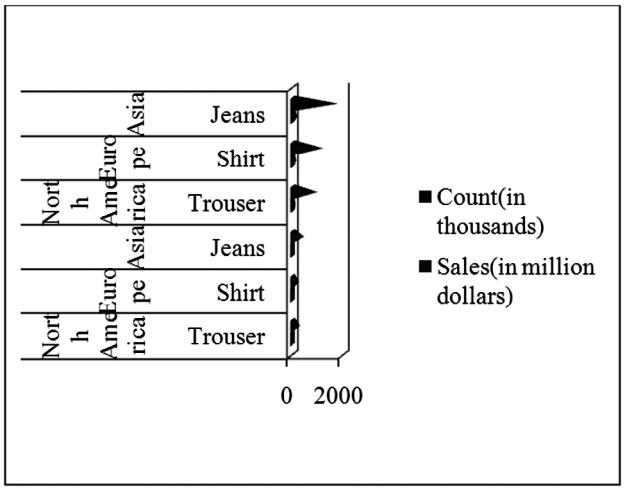
One of the biggest challenges that data warehouse managers face today is the issue of how to manage dimensional tables over a given period of time. Not only must they be able to do this, but they must also be able to manage this information with current data. When you manage your data warehouse, you will want to do more than simply place an emphasis on maintaining the data. You will want to maintain the data warehouse in a way which is directly related to caring for your customers.

When you maintain your data warehouse, it is important to place an emphasis on measurements. If you don't take the time to make measurements, your information and views will be subjective. Measuring your data warehouse will allow you to determine if you are improving as a company or organization. However, there are specific areas of a data warehouse that need to be measured.

With the Kimball approach, the data structures that must be obtained before the dimensional presentation will be dependent on the source data and transformation. In most cases, the duplicate storage of data is not required in both dimensional and online data analysis tools foundations.

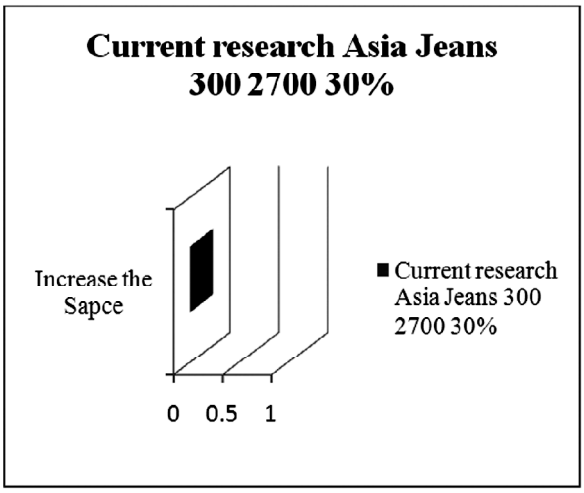
Table 2 and Chart 1
Dimension Relation for Last Three Year's Sale [5]

Location	*Product*	*Sales (in million dollars)*	*Count (in housands)*
Asia	Trouser	15	300
Europe	Shirt	12	250
North America	Jeans	28	450
Asia	Trouser	120	1000
Europe	Shirt	150	1200
North America	Jeans	200	1800



Many of the people who choose to use online data analysis tools structure believe that it is faster than the dimensional structure but we consider above techniques.

Chart 2 and Table 3
Comparison between Existing and Current research Analysis Tools



Location	*Product*	*Sales (in million dollars)*	*Count (in thousands)*	*Growth per year (2010-Till)*	*Increase the Space*
North America	Trouser	15	300	10%	10 GB
Europe	Shirt	12	250	10%	10 GB
Asia	Jeans	28	450	10%	10 GB
North America	Trouser	120	1000	20%	20 GB
Europe	Shirt	150	1200	20%	20GB
Asia	Jeans	200	1800	20%	20 GB
Current research Asia	Jeans	300	2700	30%	30 GB

[13] One example of these processes. You will want to spend time measuring the activities that you are carried out. It is also important for you to measure these activities over a given period of time. When it comes to data warehouses, measuring a process is distinct from measuring a product. It will cost your company less to develop a measurement program than you will have to pay if you measure data improperly. Your company must be able to deal with the many problems it will run into. To successfully maintain your data warehouse, it is important to understand where the true value lies. "The data warehouse must handle the current 30 GB of accounting data including all metadata and replicated data with an expected 30% growth per year.

3. METHODOLOGY

Data Implementation contains huge volume of data. OLAP servers demand that decision support quires be answered in the order of seconds. The data warehouse systems to support highly efficient cube computation techniques, access methods, and query processing techniques.

3.1 Efficient Computation of Data Cubes: [10]

The compute cube operator computes aggregates over all subsets of the dimensions specified in the operations. What is involved in the efficient computation of data cube? Suppose that you would like to create a data cube for All Garment sales that contains the following- city, product, year and sales. You analyze the data, with queries as following:

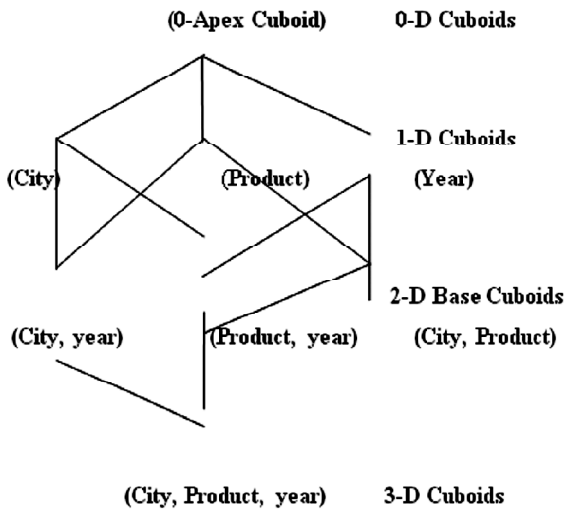


Figure 1: Relation Between Sales & Multidimensional Cuboids

Calculate the sum of sales, group by city and product.

Calculate the sum of sales, group by city.

Calculate the sum of sales, group by product.

The methodology concept is based on business process and measure performance. It means application of business process [12]. It leads the steps to analysis tools:

- Process analysis and key process indicators setup
- System of evaluation and process execution.

The design and construction of a data warehouse may consist of the following steps:

Figure 2: Design and Construction of a Data Warehouse Steps

The data warehouse analysis tools order to provide an easy information access to the end users. In order to start this phase, it is essential that the data warehouse online analysis tools have undergone system testing.

The data online analysis tools will move from development environment to test environment to production environment.

4. CONCLUSIONS

Among the greatest benefits of a data warehouse is the ability to analyze and execute business decisions based on data from multiple sources. For example, an organization has collected valuable data and stored it in 30 databases. A data warehouse is not only a convenient way to analyze and compare data in all the databases, but it can also give historical data and perspective. Thus data warehouse is a one-stop shop, but it is also a one-stop shop from an historical perspective as well. Using data warehouse, one can look at past trends, whether they be product sales or customers or whatever and may be do some predictions of what is going to happen in the future [11].

One of the primary differences between these two techniques is the online data analysis tools foundation. Many of the people who choose to use online data analysis tools structure believe that it is faster than the dimensional structure. The primary purpose of these tools is to provide analysts with an awareness of the application's components, structure, and interaction. Therefore, it is important to carefully evaluate the types of analysis that can be conducted and the presentation of analysis results.

Most companies can benefit from a data warehouse when the proper tools are in place and users are trained in analysis of results.

However, data warehouses are still an expensive solution and typically found in large firms. The development of a central warehouse is a huge undertaking and capital intensive with large, potentially unmanageable

risks [5]. Only for Online Analysis Tools in Excel Spreadsheets provide a report writer through which users can specify their own reports. Micro focus stated that it provides consulting support for revolve users to assist in customizing reports [15].

REFERENCES

[1] Alex Guazzelli, Wen-Ching Lin, Tridivesh Jena Author, "PMML in Action: Unleashing the Power of Open Standards for Data Mining and Predictive Analytics", *Create Space*, 2010.

[2] Ralph Kimball, "The Data Warehouse Lifecycle Toolkit", ISBN 978-0-470-14977-5, page 10, 115-117, 131-132, 140, 154-155, Second Edition. New York, Wiley, 2007.

[3] Kimball Et Al., "The Data Warehouse Lifecycle Toolkit", ISBN 978-0-470-14977-5, 116-117, Third Edition. New York, Wiley,2008.

[4] Inmon, W.H. "Topic: What is a Data Warehouse? Prism Solutions", **1**. 1995.

[5] Alex Guazzelli, Michael Zeller, Wen-Ching Lin, Graham Williams. Author, Title: "PMML: An Open Standard for Sharing Models", *The R Journal*, **1/1**, May 2009.

[6] Griffin, J,. Author, Title: "Fundamental Pitfalls to Avoid in Your Data Warehouse". Revised Edition April 2003.

[7] Dunham M.H., Author, Title "Data Mining, Introductory and Advanced Topics", Prentice Hall, 2002.

[8] Han J., M. Kamber, Author, Title: "Data Mining: Concepts and Techniques", Academic Press, 2001.

[9] Efrem G. Mallach, Author, Title: "Decision Support & Data Warehouse Systems", ISBN-13: 978-007-48684-3, ISBN-10 0-07-048684-0, Tata MacGraw-Hill Edition 2002, Seventh (Reprint) Edition 2011

[10] Jiawei Han, Micheline Kamber, Jian Pei, Author, Title "Data Mining: Concepts and Techniques", Microsoft Corporation December 2001, Microsoft® SQL Server™ 2000.

[11] Alex Berson Stephen J. Smith Author, Title: "Data Warehousing, Data Mining & OLAP", ISBN 0-07-058741-8, Tata MacGraw-Hill Edition 2004, Tenth (Reprint) Edition 2007

[12] Shen, W.; Tan, W.; Zhao, J., "A Methodology for Dynamic Enterprise Process Performance Evaluation. Computers in Industry", **58(5)**, June, 2007, pp. 474-485, ISSN 0166-3615, 2007.

[13] Aalst V., et al. (2007). "Business Process Mining: An Industrial Application". *Information System*, **32(5)**, (July, 2007), pp. 713-732, ISSN 0306-4379.

[14] Dunham M.H., Author, Title "Data Mining, Introductory and Advanced Topics", Prentice Hall, 2002.

[15] Jiawei Han, Micheline Kamber, Jian Pei, Author, Title "Data Mining: Concepts and Techniques", Microsoft Corporation December 2001, Microsoft® SQL Server™ 2000.