

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320-088X

IJCSMC, Vol. 5, Issue. 1, January 2016, pg.86 – 90



Efficient Reporting Tool for Business Intelligence with Increasing Volume of Data

B.Mounica

New Horizon College of Engineering, India

premkumarmounica@gmail.com

Abstract: Traditional databases are not sufficient to store and retrieve large volumes of data, read operations are time consuming through reporting tools because of I/O operations are restrictive in nature ,only fixed number of I/O operations with current technology is possible. so instead of that coming up with Netezza [1]; Sybase iq[2] databases which are highly tuned for read intensive queries. So this paper gives an idea on efficient use of reporting tool using in memory approach for business intelligence with increasing volume of data.

Keywords: OLTP: online transaction system; OLAP: online analysis Processing; I/O: input output; ETL: Extraction, Transformation, and Loading.

I. Introduction

Data warehouse is a system used for business reporting and analysis of data for making useful information from the database. Database store huge amount of current and historical data in structured, semi structured and unstructured manner. In data warehouse the data will be in structured data. But coming to big data the data can be of any type.

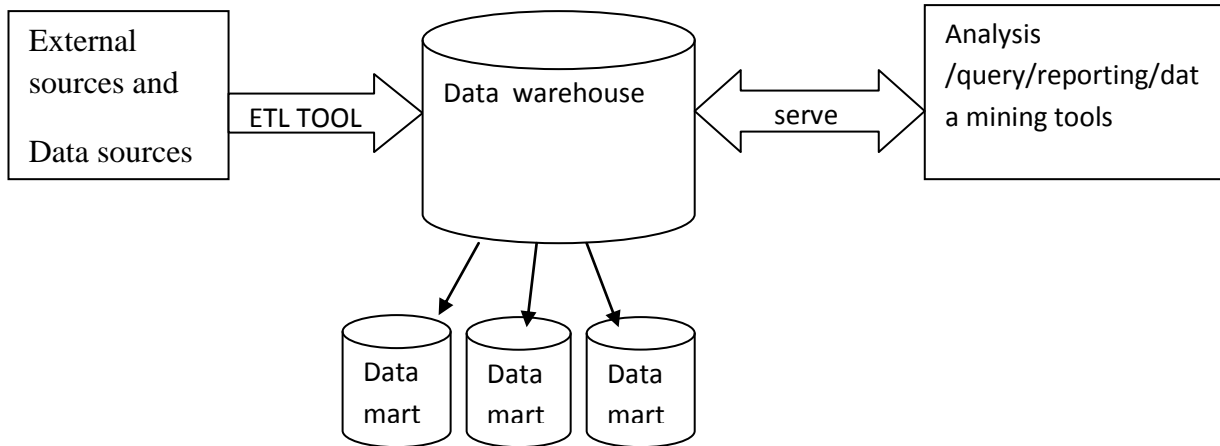


Fig. 1 Data warehouse architecture

In the above figure the architecture of data warehouse is showed ETL tools are applied on external sources and data sources and place the data in data warehouse and then data will be divided in to data marts and serve for analysis ,query ,reporting ,data mining tools. Data marts are the slices of data warehouse for example company details can be make in to different data marts depending on sales details ,salary details etc.

II. Types of Data

There are 3 different types of data structured, semi structured and unstructured data. In Structured data the data that can be fixed in the records. Unstructured data the data which cannot be fixed in to the records .In unstructured data it is a combination of photos, graphics, images, videos and text. Semi structured data is a combination of both structured data and unstructured data

III. Types of system

In data warehouse there are many types of system like data mart, OLTP, OLAP etc. In this paper OLTP, OLAP data warehouse is used which is shown in the figure.

OLTP: online transaction system many short online queries and updates are done example online train ticket booking. OLTP is customer-oriented and is used for transaction and query processing by clerks, clients and information technology professionals and OLAP: online analysis processing large amount of data and complex queries are done for example sales record annually. . An OLAP system is market-oriented and is used for data analysis by knowledge workers, including managers, executives and analysts

ETL is the Software tools used in this paper. ETL means Extraction, Transformation, and Loading which is used to convert OLTP in to OLAP shown in the figure.

The benefits of using data warehouse are to maintain history of data and to improve quality of data and in business applications to generate reports like annual and quarterly comparisons of sales and salary etc.

In traditional databases 10gb to 20gb of data can be stored, examples of traditional databases are oracle and db2; along with that simple reporting tools like crystal reports and VB is substantial. But now the data volume increased enormously so these kinds of systems are not sufficient. Data volume usage by the social media websites gone up to 600 TB per day; example face book [3]. Since traditional systems not workout to keep up the pace of the data volume generated. But oracle and DB2 are good enough for transactional database.

IV. Business Intelligence Reporting Tools

Business intelligence reporting tools are used to transfer the raw data which is stored in database in to meaningful and useful information, which will generate reports for business analysis.

Currently new customized database started emerging which is very specific for data warehousing purpose .databases like neteeza; Sybase iq very specific for data warehousing, business intelligence and these database are highly tuned for read intensive queries.

These data bases are used for bulk updates; 1/5th time differences can be noticed between traditional systems and these systems.

By using small write operations stack holder can enter data in to the OLTP database and this data is transferred to OLAP by using ETL tools like Informatica [4] we can extract different types of data sources and transform the data to store in a proper format and then load the data in final OLAP database for future analysis and quering. In OLAP read operations are efficient; by using reporting tools like qlikview, OBIEE, crystal report etc[5] data can be generated in form of reports for business applications.

The current issues in the reporting tool are I/O operations bandwidth is limited which is in between the reporting tool and OLAP database. The final report which is given to the user is slow and time consuming in the query which is fixed against large volumes of data.

This can be overcome if reporting tool is with in memory enabled. In memory means OLAP data is divided in to data marts and stored in reporting tool .data can be fetch directly from the reporting tool. Example a company profits can be in different data marts like sales, marketing etc. which can be stored in reporting tool from the OLAP database to the reporting tool when a query rises against this information business analyst can search directly in reporting tool and reports can be generated within time and star schema is implemented.

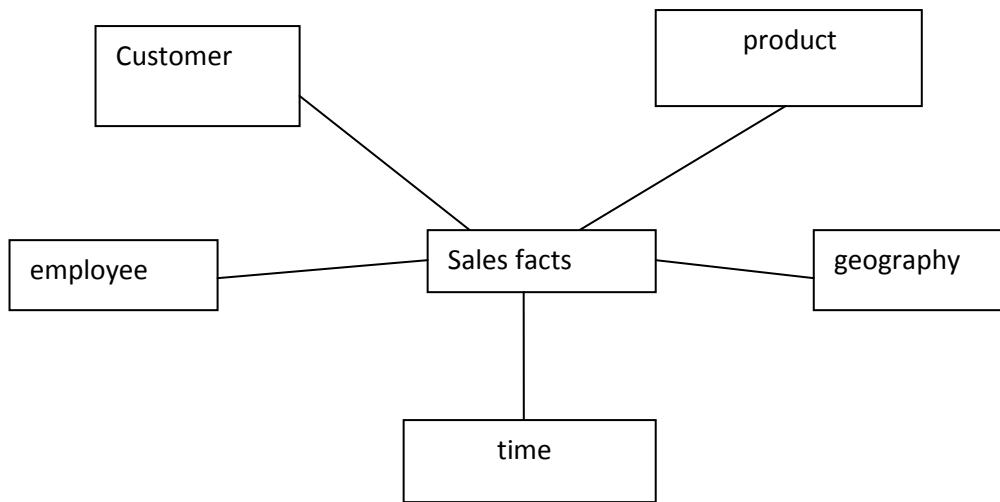


Fig. 2 Star schema

In the above example the sales facts of a company with the help of star schema is showed and the in memory concept is explained in detail in the following below diagram.

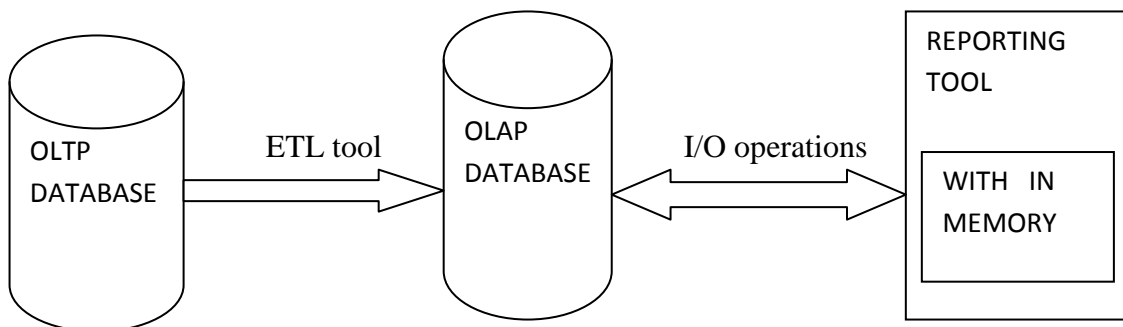


Fig. 3 Efficient reporting tool for business intelligence

V. Future Enhancement

Databases read operations are time consuming because of large volume of data is stored and I/O operations band width is limited and The final report which is given to the user is slow and time consuming in the query which is fixed against large volumes of data . By using in memory reporting tool the inefficient I/O operation bandwidth between OLAP database and reporting tool can be overcome and efficient reporting tool for business intelligence can be generated with increasing volume of data.

This way it can provide new insights in to the efficiency and performance for business intelligence tool with increasing volume of data.

The same can be applied with big data with unstructured and semi structured data for generating reports which is used for business analysis.

VI. Conclusion

By using in memory concepts in reporting tool the current issues in the reporting tool with limited I/O operations bandwidth in between the reporting tool and OLAP database reporting tool can be overcome. Thus the inefficient I/O operation bandwidth between OLAP database and reporting tool can be overcome and efficient reporting tool for business intelligence can be generated with increasing volume of data.

References

- [1] Netezza available <https://en.wikipedia.org/wiki/Netezza>
- [2] Sybase iq available <http://www.sap.com/pc/tech/database/software/sybase-iq-big-data-management/index.html>
- [3] <https://code.facebook.com/posts/229861827208629/scaling-the-facebook-data-warehouse-to-300-pb/>
- [4] <http://datawarehouse4u.info/ETL-tools.html>
- [5] <http://dwhlaureate.blogspot.in/2012/08/list-of-popular-reporting-tools-in.html>
- [6] http://www.zenithresearch.org.in/images/stories/pdf/2012/Feb/ZIJMR/17_ZEN_VOL2ISSUE2_FEB12
- [7] http://www.csjournals.com/IJITKM/PDF%206-1/Article_15
- [8] *Enhancing Business Intelligence using Data Warehousing: A Multi Case Analysis* by Sonal sharmal and Rajini jain ISSN: **2321-7782 (Online)**
- [9] "The Study on Data Warehouse Design and Usage" International Journal of Scientific and Research Publications, Volume 3, Issue 3, March 2013 1 ISSN 2250-3153
- [10] "An Overview of Data Warehousing and OLAP Technology" <http://www.cs.sfu.ca/CourseCentral/459/han/papers/chaudhuri97>