ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

An Extensive Study on Data Warehouse, OLAP used for Implementing ERP Systems

B.V. Baiju, R. Vardhini,

Assistant Professor, Dept. of Information Technology, Hindustan Institute of Technology and Science, Chennai, India

B.Tech Student, Dept. of Information Technology, Hindustan Institute of Technology and Science, Chennai, India

ABSTRACT: Data warehouse is a repository of data and related information which is an integrated, subject oriented, time variant and non-volatile collection of data used for decision making. It is a type of database for decision making which is separately maintained from the operational database. Data warehouse provides Online Analytical Processing (OLAP) tool, which is based on the multidimensional data model. It provides the capability for complex calculations, trend analysis and sophisticated data modeling. So, data warehouse and OLAP plays an important role in decision making and knowledge discovery process. PowerBI tool is a visualization tool thatis used for business analytics service which gives you a single view of your most critical business data. In this paper, we discussed about data warehouse and OLAP and also described about how these can be used for ERP implementation.

KEYWORDS: Data warehouse, OLAP, Multidimensional data model, Decision Making, PowerBI, Analytics

I. INTRODUCTION

A data warehouse environment stores detailed, integrated, clean and consistent data extracted from different data sources and properly processed by means of ETL tools. It also includes statistical analysis, reporting, data mining capabilities, client analysis tools, and other applications. William Inmon defines data warehouse in the sense of subject oriented, integrated, the Nonvolatile and Time variant which is also referred as the characteristics of the data warehouse. Instead of transaction processing, data warehouse is designed as a relational database that is used for query and analysis of data. Generally it contains historical data which are derived from transaction data, but the data can be included from different sources. Analysis workload can be separated from transaction workload for the purpose of enabling an organization to consolidate data from different sources.

ERP systems have exponentially grown in recent years, but the implementation of ERP systems has been problematic for many organizations.To deal with the interaction of people and technology the concept of Information Systems is used. An Enterprise Resource Planning system is included in information system,which is a user-interfaced and designedto provide information useful to support strategy, operations, management analysis, and decision-makingfunctions in an organization.

OLTP of ERP systems integrates all day-to-day transactions of an enterprise in the way which enables users to reduce operating cost, verify consistency of transactions, facilitate day-to-day management and provide detailed or summarized reports. When we convert the OLTP data sources into the data warehouse, it enables to apply OLAP and data mining techniques to provide more analytical reports and provide answers of miscellaneous questions to decision makers.

II. LITERATURE SURVEY

The implementation of ERP systems has been one of the first problems addressed by the literature on the subject. This area remains very active over the last 2 years, as it is stated in [1], where two distinct research streams are observed from the literature, the first one focuses on the fundamental corporate capabilities driving ERP as a strategic concept, the second on the details associated with implementing information systems and their relative success and cost. Based on this distinction, this section has been divided in sub-sections addressing the problems of sociological

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

and cultural factors influencing the implementation success, the implementation steps, the business process alignment phase and the factors of success/ reasons for failures.

An ERP system is a set of packaged application software modules, with an integrated architecture, that can be used by organizations as their primary engine for integrating data, processes and information technology, in real time, across internal and external value chains (Shang and Seddon 2002). It impounds deep knowledge of business practices that vendors have accumulated from implementations in a wide range of client organizations that can exert considerable influence on the design of processes within new client organizations (Shang and Seddon 2002).

ERP systems enable organizations to share common information and activities across the entire organization, automate and integrate the critical parts of business processes, and access information in a real-time environment. Since ERP systems can facilitate the productivity and efficiency of firms, the majority of organizations implement ERP systems to increase organizational competitiveness (Glover et al.1999; Davenport 1998; Pan et al. 2001). ERP systems touch on many aspects of a company's internal and external operations and provide organizations with an overall view of the business through multidimensional information (Gefen and Ragowsky 2005; Markus and Tanis 2000). Consequently, successful deployment and use of ERP systems are critical to organizational performance and survival (Markus and Tanis 2000). Nevertheless, the bulk of research on the implementation of ERP systems focuses on firms either prior to, during, or immediately after ERP software implementation (Krasner 2000; McNurlin 2001). Conventional wisdom saw "going live" as the end of ERP implementation and ignored the second wave, the post adoptive stage, which refers to the actions that are taken after going live that help organizations achieve the full capabilities and benefits (Deloitte Consulting 1999).

The roots of enterprise resource planning (ERP) systems may go back to half a century ago. With the development of information technology (IT) and the demands of organizations, ERP which originated from manufacturing cores had covered nearly all essential processes and functions of organizations two decades ago. Although some shifts happened during these years, according to Columbus (2013), in despite of the worldwide ERP software market share in 2012 shows that the SAP is still leading the worldwide market with 24.6 percent market share, new ERP vendors with tremendous growth indeed pose a potential threat. Meantime, the worldwide ERP market experienced slow growth of 2.2 percent, yet quoted from Columbus (2013) "Software-asa-Service (SaaS), financial management and Human Capital Management (HCM) applications showed potential for breakout growth." The ERP report of Panorama Consulting Solutions (2013) pointed out that the traditional ERP software was chosen by the majority of 61 percent with an increase of 3 percent over 2012, and 26 percent of respondents selected software as a SaaS and cloud ERP. To put it bluntly, the traditional ERP industry is shifting. The various demands of organizations are accelerating the diversity of ERP software. Apparently, the survival in natural selection will lead the next generation.

III. TECHNIQUES AND TOOLS FOR ERP SYSTEM

ONLINE TRANSACTION PROCESSING

OLTP applications are used in many applications in which the online users can have direct access to information. The OLTP is an application procedure unit of work, called transactions. An application can be simulated to run in parallel on a tightly coupled multiprocessor so that it cans concurrently procedure of multiple transactions. The quantity of parallelism depends on the system programmer, the transaction manager, and the hardware that the OLTP application runs on. To span a network online transaction processing increasingly requires support for transactions. To solve this issue, new online transaction processing software uses client or server processing and brokering software.It allows transactions to run on different computer platforms in a network.

ONLINE ANALYTICAL PROCESSING

OLAP is characterized by relatively low volumes of transaction. Queries are often very difficult and they involve aggregations. For OLAP systems a response time is a good organizational measure. OLAP applications are widely used by the Data Mining procedure. In the OLAP database there is aggregated, chronological data, stored in multi-dimensional schemas (usually star schema). OLAP summarizes data and makes forecasts.Component-based OLAP systems offer a number of benefits for both the user and the developer. The user can choose different suppliers and also combine different query optimization strategies and query evaluation algorithms. Extensible OLAP systems offer a

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

major benefit to developers, as well. The following will show how to decompose an OLAP system into functional units, which can communicate using a "software bus". It proposes a data model based on sets and vectors for the communication.The various multi-dimensional operations are:

- Rolling up (producing marginal)
- Drilling (going down levels of aggregation—the opposite of rolling up),
- Slicing (conditioning on one variable),
- Dicing (conditioning on many variables) and
- Pivoting (rotating the data axes to provide an alternative presentation of the data)

The enterprise data was split into OLTP and OLAP. OLTP is the necessary prerequisite for OLAP, however only with OLAP companies are able to understand their business and come to conclusions about how to steer and change course. Fig.1. Shows how data is provided by OLTP and alayzed by OLAP.

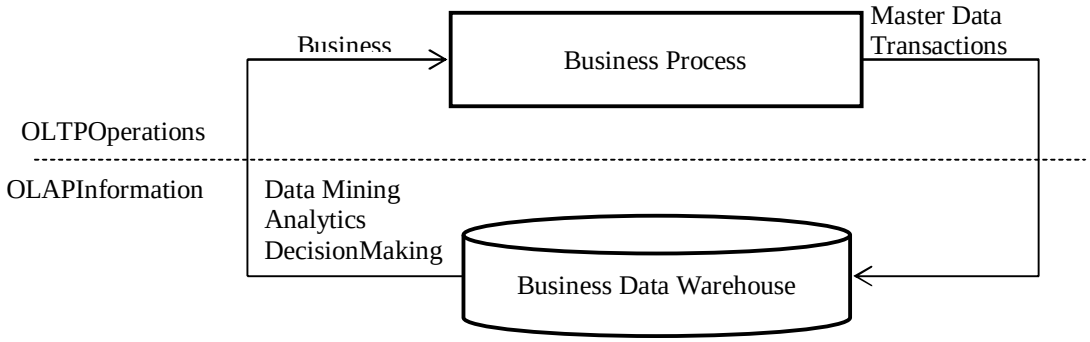

Fig.1. Data over OLAP and OLTP

Business becomes transparent and decisions can be made only when we match the planned data and actual data. The following table1 summarizes the major differences between OLTP and OLAP system design.

Factor	OLTP	OLAP
Source of data	Original operational data.	Consolidated data where data come from various OLTP databases.
Purpose of data	Used to control and run fundamental business processes.	Used to help with planning, problem solving and decision support.
Queries	Relatively standardized and simple queries returning few records.	Complex queries with aggregations.
Processing speed	Very fast.	The speed depends on the amount of data involved..
Space requirements	Can be relatively small if historical data is archived.	There is a larger requirement of space due to the existence of aggregation structures and history data.
Database design	The design is highly normalized with many tables.	De-normalized with fewer tables with the use of a star or snowflake schema.
Backup and recovery	Data loss is likely to entail significant monetary loss and legal liability.	Instead of regular backups, some environments may consider simply reloading the OLTP data as a recovery method.

Table 1.OLTP vs OLAP

OLAP draws data sets that contain the information in multiple dimensions. These data sets are in turn drawn from relational sources.The OLAP Survey in terms of the areas of greatest interest to companies considering who should be on their OLAP shortlist:

- Functionality and performance
- Data sources for OLAP products

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

- Implementation
- Common customer issues
- Buyer influences.

The survey is done and the percentage of usageof OLAP for selected industriesis mentioned in table 2.

Industry	Average % OLAP usage	% IBM Cognos PowerPlay usage
Healthcare/Pharma	8.5%	16%
Information Technology	8.0%	9%
Insurance/Div. financial	7.1%	9%
Consumer packaged goods	6.1%	8%
Businessconsulting/ Professional services	5.4%	3%
Manufacturing – Process	3.4%	4%
Transportation	3.4%	4%

Table2: Industries using OLAP and Tools

ENTERPRISE RESOURCE PLANNING

An ERP system is a set of managerial tools that provide a balance between supply and demand within a business. It also provides a complete supply chain from suppliers to end users. ERP is used for providing well established business processes for decision making and a high degree of integration of all business processes within the enterprise. ERP systems are used to collect, store, manage and interpret data from many business activities, including finance, marketing and sales, inventory management, HR management, etc. The real value of an ERP system can be revealed only after the deployment of a BI system.Enterprise Resource Planning (ERP) is a set of applications for core business operations and back-end management that was originally developed for manufacturing and commercial companies. ERP permits organizations to integrate business processes and optimize the resources available.

The benefits of ERP systems are

- Business process automation
- Timely access to management information
- Increased flexibility
- Reduced quality cost
- Improved vendor performance
- Improved decision making capability

The reasons for a purchasing ERP system may include

- Efficiency of the current system
- Failure in the distributed system
- Maintenance overhead in the current system
- Competition
- Company growth
- E-commerce

ERP AND DATA MINING

Business Data Mining is the process of exploration and analysis. This is used to discover meaningful correlations, patterns and trend by examining through large amounts of data that are stored in repositories. Businesses may be able to compare customer feedback, retain highly valuable customers and also used to perform effective market analysis by identifying similar products, which are used to make smart business decisions. Several predictive and statistical methods are used in order to explore and analyze data. Association rule, linear regression, neural networks, regression trees, cluster analysis and classification trees are some of the methods included in Data Mining.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

INTEGRATING OLAP AND DATA MINING IN ERP

Integrating OLAP and data mining can benefit for the ultimate objective of efficiently providing a flexible answer to data mining queries addressed either to a relational or a multi-dimensionaldatabase. The implementation of OLAP relies on selection, projection and assembly operations on lattices.The four major steps in integrating multi-dimension module in ERP are

1. Data mart of ERP

A subset of the data warehouse is known as Data mart. It is the access layer that is used to get data out to the users.Data mart defines the overall methods and scopes during the evaluation stage of the design and determined using Schema or Snowflake in accordance with the requirements. It highlights the concepts of importing, purchasing or ordering of good on the single business activity of the enterprise.

2. Choices of fact:

The Cube is built and completed using Fact. A Fact table consists of measurements and metrics of a business process. A fact table is found at the cente of a star schema. Therefore Fact must be able to answer all the possible questions that may occur during the process of decision-making.

3. Establishment of Dimension:

Dimensions are a common way of analysing data. Dimensions store the textual descriptions of the business.The textual description should contain simple and useful message and avoid codes, abbreviations and Null. The explicated time, names or addresses allow more flexibility in inquiries. For processing the Fact Table each and every item of the Dimension Table carries multiple feedback capabilities. When there are changes over the data, the new data will be added to the “newly added data row”. The time is used to tell the track record at certain points. This method allows unlimited times for tracking the changes over data. The deficiency is that it must use time to identify the updated data row and increase the data rows of the dimension table. There is no need to build additional data row or to change the values architecture of dimension tables.

4. The design of aggregation

The purpose of Aggregation is to get more information about particular groups based on specific variables. It can also be made from different data occurs within the same data subjects, business transactions and de-normalized database. When facing complicated enquiries aggregation is used to increase the analysis speed by calculating the total amount of data.

POWERBI TOOL

Power BI is positioned by Microsoft as the ‘self-service BI solution’ that can integrate with many other Business Intelligence solutions. Microsoft provides thesolution for bridging the gap and reaching the Business Intelligence leaders Tableau and Qlik. The spearhead of the Microsoft BI stack is obviously Power BI that gives to the Redmond Company, together with other BI software SSRS and DataZen, the most complete suite of any vendor in the BI market today. PowerBI gave on free preview the capability to share on public websites the dashboards created also it is possible to share data, reports and stories to everyone, embedding in a web page one row of simple HTML code.

IV. PROPOSED SYSTEM

The proposed system uses the PowerBI to visualize the generated report. The required data of all employees are stored in a database. These data are in the form of tables (SQL server) and they are created using a stored procedure, as a stored procedure has various advantages to its credit. It reduces the amount of information sent to the database and the compilation needs to be done only once. With all the benefits, the backend database is ready. In big organisations, data will be updated on a regular basis and due to it; the older data should not lost. In order to have the historical data, the data warehousing concept is being implemented. The data are subjected to a concept called as schedules. In scheduling the data will get updated on a regular timely interval and hence all the records will be available and updated regularly. This will make it more convenient to fetch them during use.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

The backend tabled are connected to the frontend of the PowerBI and all the data are now imported to the frame. With this a relationship can be formed with a central table connecting to all the sub tables. Now according to the visual need, the graphs can be imported by drag and drop methods. Different form of graphs like bar graph, pie chart can be used for each dimension like the type of employee, grade, gender, level, type of leave, reason and so on. It also uses various multi dimension operations like drill down, drill up to get into a clearer picture. Now the outputs are clearly reported and hence the analysis is easier to make.

V. CONCLUSION

The ERP system will definitely change in accordance to the way people work and also it provides a new platform with different interfaces, data entry is changed and report formats are different. Users often find these changes unnecessary and therefore refuse to accept them. The best way to address and reduce the impact of these changes is to encourage user participation in the implementation of ERP systems. ERP implementations are expensive and complex undertakings, but once they are successfully implemented, significant improvements can be achieved, such as easier access to reliable information, elimination of redundant data and operations, reduction of cycle times, increased efficiency hence reducing costs.

REFERENCES

1. Arnold.V,Hunton. J.E, Sutton, S.G.,"On the Death and Dying of Originalityin the Workplace: A Critical View of Enterprise Resource Planning Systems Impacton Workers and the Work Environment', Working Paper, University of South Florida, 2000.
2. T.H. Chang et al., ' Measuring the success possibility of implementing ERP by utilizing the Incomplete Linguistic Preference Relations', Applied Soft Computing 12 , 2012.
3. Wen Hsien Tsai et al.,'An empirical investigation of the impacts of internal/external facilitators on the project success of ERP : A structural equation model', Decision Support Systems 50 , 480-490, 2011.
4. Amin Amid et al.,'Identification and classification of ERP critical failure factors in Iranian industries , Information Systems'D. O. Leary, Enterprise, 2012.
5. Ahituv Niv, Neumann Seev, Zviran Moshe, 'A system development methodology for ERP systems', Journal of Computer Information Systems, 2002.
6. Berry.M. J., Linoff.G. S., 'Mastering Data Mining: The Art and Science of Customer Relationship Management',Wiley Computer Publishing, New York, 2000.
7. Goebel. M., Gruenwald, L., 'A Survey of Data Mining and Knowledge Discovery Software Tools', ACM SIGKDD(Special Interest Group on Knowledge Discovery and Data Mining) Explorations, June, (1:1), 20-28,1999.
8. T. Ariyachandra, H. J. Watson, 'Key organizational factors in data warehouse architecture selection', Decision Support Systems 200–212, 2010.
9. Wided Oueslati and Jalel Akaichi, 'A survey on data warehouse Evolution', International Journal of Database Management Systems (IJDMS), Vol.2, No.4, 2010.
10. A. Abdullah, Zahid Ullah, 'A Framework of an Automated Data Mining Systems Using ERP Model', International Journal of Computer and Electrical Engineering, Vol.1, No.5, pp. 651-655, 2009.
11. Mohammed K. Kolkas, 'Integrated Data Mining Techniques inEnterprise Resource Planning (ERP) Systems', International Journal of Information Science and Intelligent System, Vol.2, Issue 2, pp. 131-152, 2014.
12. Tamer S. Abdellatif, 'Comparing online analytical processing and data mining tasks in enterprise resource planning systems', International Journal of Computer Science Issues,Vol. 8, Issue 6(2), pp. 161-174 , 2011.
13. Mawilmada, Pubudika Kumari, 'Impact of a data warehouse model for improved decision-making process in healthcare', Masters by Researchthesis, Queensland University of Technology, 2011.

BIOGRAPHY

Mr. B.V. Baiju received M.E degree in Computer Science and Engineering from Hindustan College of Engineering. Chennai. Currently he is working as Assistant Professor at Hindustan Institute of Technology and Science, Chennai. His interested research domains are Bigdata and Software Engineering.

R.Vardhini is a student of the Department of Information technology, Hindustan Institute of Technology and Science, Chennai. Her research interests are Bigdata, Networking and Programming languages.