Empowering Educational Institutes Using Data Warehouse

Vinay Nayak[1], Roopa E[2]

[1,2] *Department of Information Science & Engineering, R.V.College of Engineering,Bengaluru-59, Karnataka, India*

***Abstract*-The need to access data is a very common requirement now a days, be it a non-profit organization or full-fledged business organization. Decision makers have always needed data to make informed and accurate decisions. Data-driven decision support systems, such as data warehouses can serve the requirement of extraction of information from more than one subject area. Data warehouse can provide the information required by the decision makers at their fingertips. Usually data warehouses are built for an organization where revenue and profits are more important. The data warehousing technology need not be limited to usage by business organizations, its usage can be extended to non-profit areas like education and e-governance which can enhance the productivity.**

Current enrollments count in thousands in any reputed institute. The management challenges include meeting diverse student needs, increased complexity in academic processes. The complexity of these challenges requires continual improvements in operational strategies based on accurate, timely and consistent information. An information warehouse can deliver the required information to the decision makers and provide an insight into the overall situation. This greatly facilitates decision makers in taking micro level decisions in a timely manner without the need to depend on their IT staff. This scenario is applicable to education domain also. This paper provides the way in which productivity of education institute can be improved by using data warehousing technology.

***Keywords* ---Data analytics, Data warehouse, Data Modeling, ETL, Decision Making**

I. INTRODUCTION

Often the decision makers of educational institute take decisions based on assumption and rely on what they believe is true. They also depend on manual methods which needs resources like staff and time. In any educational Institute, data is distributed across the campus in different automation systems the institute has. When data resides on multiple systems distributed across the campus, it is difficult to increase access and make intelligent use of data .This leads to lack of ability for deeper analysis of data. Institutions are locked into manual, paper based methods for their decision making.

Data warehouse technology is the technology which holds huge information loaded from various sources. Usually, there will be multiple OLTP (Online Transaction Processing) systems in education institute for processing of various purposes. The data stored in these OLTP systems are in different platforms like SQL-Server, MySQL or any other type of data base platforms. Apart from this, there is also chances of important data being stored the form of excel file or text file. The information from all these sources needs to be consolidated and report is to be generated as and when required in the current setup. This is a complicated task which requires lot of manpower and services of IT staff. Building a data warehouse for education institute solves these problems. The decision makers can retrieve any information at any time without any dependency on IT staff.

II. RELATED WORK

Shaweta [1] discussed the issues related to building and maintenance of a data warehouse for the Institutions. Queries such as total number of professors in a span of years, number of students who have got the scholarship in recent years will be answered by institute's data warehouse. The researcher opines that the Database conversion and schema integration is most difficult task in building and maintaining a data warehouse for an Educational institute. The challenges also include heterogeneous data challenge and Complexities of the organization.

Azwa Abdul Aziza et al., [2] proposed an educational data warehouse architecture that employs the integration of proprietary and open source BI tools. The education data warehouse eduBI proposed in the research aims to adopt and adapt BI technologies in educational settings.

Umesh I M et al., [3] have built a data warehouse specific to engineering institution. The specific data marts which hold information about specific subject areas are employee information, student information and assets information. The queries to specific data marts provide required information for decision makers.

The outcome of the study of related works demands the more customized data warehouse for educational institute. This work attempts to build one such data warehouse.

III. PROPOSED MODEL

The first step to building data warehouse is to identify the source systems. The institute with student strength of more

than six thousand and faculty/non-teaching staff strength of five hundred was chosen for this study.

The various sources of information in an educational institute are result sheets, asset management (stock) registers, employee information. Once the sources are identified, the next step is to perform ETL i.e., Extract, Transform and Load. The proposed model is shown in the figure 1. The various OLTP systems have platforms like SQL server, MySQL and excel file or flat files. The data is to be extracted, transformed into required format and loaded into data warehouse.

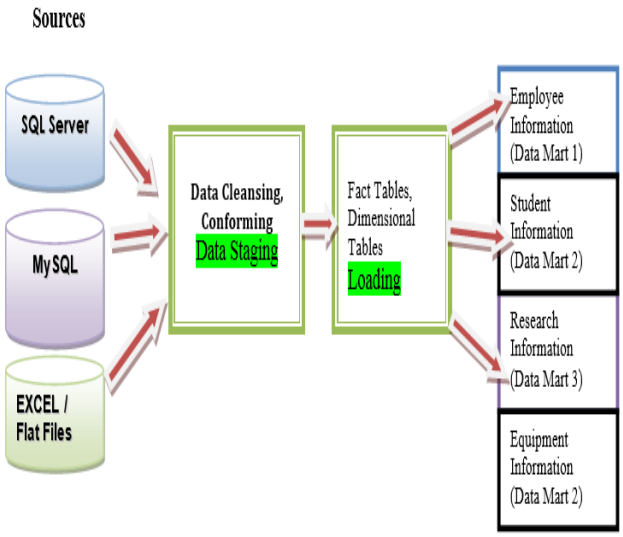

Figure 1. Education Data warehouse architecture

A. *Identifying Sources*

Usually, Institutes have different approaches for decision making processes. These approaches vary from using Campus management systems to department wise automation systems. The institute chosen for this study uses various methods for information storage. The back end information was stored in different platforms like SQL Server, MySQL in different automation systems. Student information such as results and non-academic activities related information was stored in Excel files and flat files. Research information was stored MySQL instance of WAMP server.

B. *Data Modeling*

For an educational institute, the information from different platforms includes employee information, Student information, Research Information and Equipment information. The proposed data warehouse architecture has different data marts for each type of information. A star schema is a modeling paradigm in which the data warehouse contains a large, single central Fact Table and set of smaller Dimension tables, one for each dimension. The fact table contains the detailed summary data. Its primary key has one key per dimension. Each dimension is a single, highly de-normalized table.

Every tuple in the fact table consists of fact or subject of interest, and dimension that provide the fact. The dimension table consists of columns that correspond to the attributes of the dimension.

C. *Data Staging and Loading*

Data staging is the process of moving data from the operational data base to the data warehouse. The main tasks in this process are extract, transform and load. This part is a very technical part of the warehouse development. The data from various sources are extracted, transformed in to required format using various transformations available in ETL tools. The target database used here is Oracle.

Figure 2 and figure 3 indicate screen shots of data warehouse building.

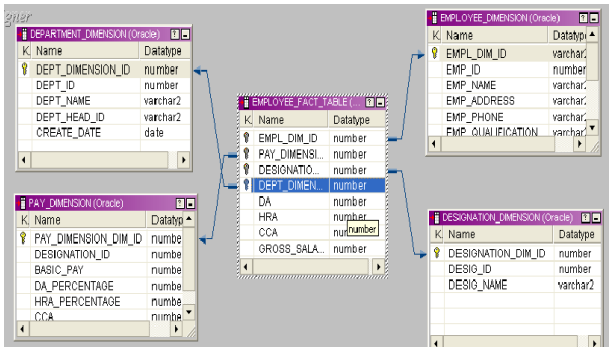

Figure 2. Star Schema designed for EMPLOYEE MART

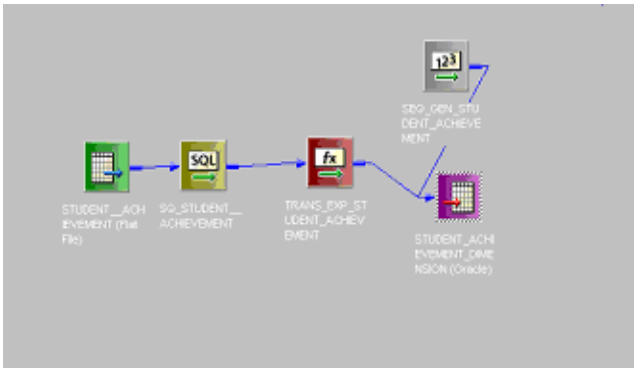

Figure 3. Mapping showing the data flow

IV. CONCLUSION

Most of the educational institutes are nonprofit organizations and the usefulness of the data warehouse can be done on the basis of its ability to meet user's requirements. The decision makers can extract information regarding three main components of the Institute, namely Employees, Students and the Infrastructure. The data warehouse can provide the

answers to questions like "Who are the toppers in each department?", "How much amount has been spent towards salary of employees?", "List of Equipments whose warranty has been expired?" The beneficiaries of the data warehouse technology include decision makers like the Management, the Principal, and the Heads of the Departments. Students and employees will be benefited as they will be in the receiving end of better decisions taken.

REFERENCES

[1]. Shaweta, Critical Need of the Data Warehouse for an Educational Institution and Its Challenges, international Journal of Computer Science and Information Technologies, Vol. 5 (3), 2014, 4556-4559

[2]. Azwa Abdul Aziza et al., Framework for educational data warehouse (edw) architecture using business intelligence (bi) technologies, Journal of Theoretical and Applied Information Technology, Vol. 69 No.1

[3]. Manjunath T. N., Ravindra S. Hegadi, Umesh I. M., and Ravi Kumar G. K. "Realistic Analysis of Data Warehousing and Data Mining Application in Education Domain'' in International Journal of Machine Learning and Computing, Vol. 2, No. 4, August 2012.

[4]. W. H. Inmon. "'Building the Data Warehouse." John Wiley & Sons, Third edition, USA, 2002.

[5]. R. Wrembel, C. Koncilia. "Data Warehouses and OLAP: Concepts, Architectures, and Solutions". Idea Group Inc (IGI), 2007.