

COMPARISON OF DATA WAREHOUSE ARCHITECTURE BASED ON DATA MODEL

Bikramjit Pal

Data warehouse is a repository of an organization's electronically stored data. Data warehouses are designed to facilitate reporting and analysis. A data warehouse houses a standardized, consistent, clean and integrated form of data sourced from various operational systems in use in the organization, structured in a way to specifically address the reporting and analytic requirements. This definition of the data warehouse focuses on data storage. However, the means to retrieve and analyze data, to extract, transform and load data, and to manage the data dictionary are also considered essential components of a data warehousing system. Many references to data warehousing use this broader context. Thus, an expanded definition for data warehousing includes business intelligence tools, tools to extract, transform, and load data into the repository, and tools to manage and retrieve metadata. There are mainly three classifications of data warehouse architecture based on data model given by Inmon and Ralph. There is a fourth classification that is discussed in this paper and it has been compared with the other three.

Keywords: Data Warehouse, Star Schema, Business Intelligence and Data Mart.

1. INTRODUCTION

Data warehouse is a repository of an organization's electronically stored data. Data warehouses are designed to facilitate reporting and analysis. A data warehouse houses a standardized, consistent, clean and integrated form of data sourced from various operational systems in use in the organization, structured in a way to specifically address the reporting and analytic requirements. This definition of the data warehouse focuses on data storage.

However, the means to retrieve and analyze data, to extract, transform and load data, and to manage the data dictionary are also considered essential components of a data warehousing system. Many references to data warehousing use this broader context. Thus, an expanded definition for data warehousing includes business intelligence tools, tools to extract, transform, and load data into the repository, and tools to manage and retrieve metadata. There are mainly three classifications of data warehouse architecture based on data model given by Inmon and Ralph.

2. MODEL

The three architectural classifications are:

Assistant Prof., Department of Computer Application, JIS College of Engineering

Block 'A' Phase III, Kalyani, Nadia-741235, West Bengal, INDIA

Email: biku_paul@rediffmail.com

2.1. Top Down Architecture, Top Down Implementation

In this architecture data marts are created from data warehouse and user can then directly use the data warehouse. Data marts provide faster access and very specific analysis. Architecture is given below:

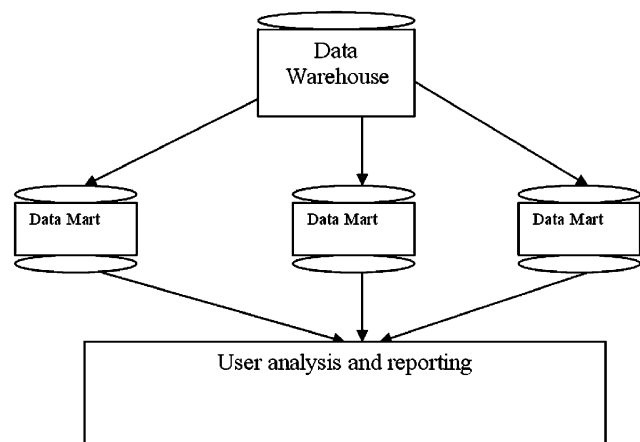


Fig. Top Down Architecture, Top Down Implementation

2.2. Bottom up Architecture, Bottom up Implementation

Based on star schema in this architecture data warehouse is created from the union of all data marts and data warehouse is the used by users for analysis purpose. Here specified data for use can be available if data warehouse can go through staging process. Architecture is given below:

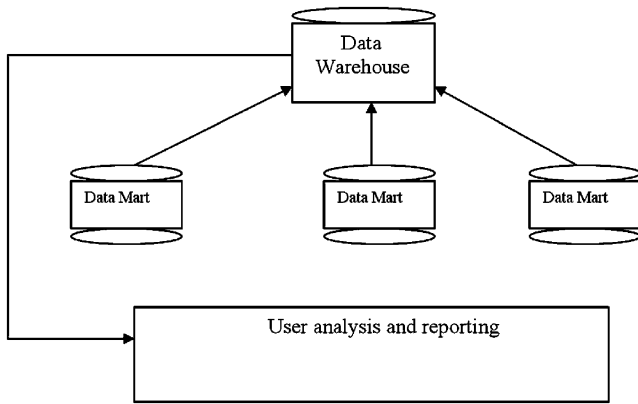


Fig. Bottom up Architecture, Bottom up Implementation

2.3. Top Down Architecture, Bottom up Implementation

In this architectural approach data marts are created from data warehouse but when implementing, the users have to follow the above approach. This architecture is not commonly used in industry because data mart creation in this case what I think is not optimal in use.

3. IMPLEMENTATION

I thought of having one more combination if possible of this architectural concept, i. e. Bottom Up Architecture, Top Down Implementation. The possible architecture could be User analysis and reporting

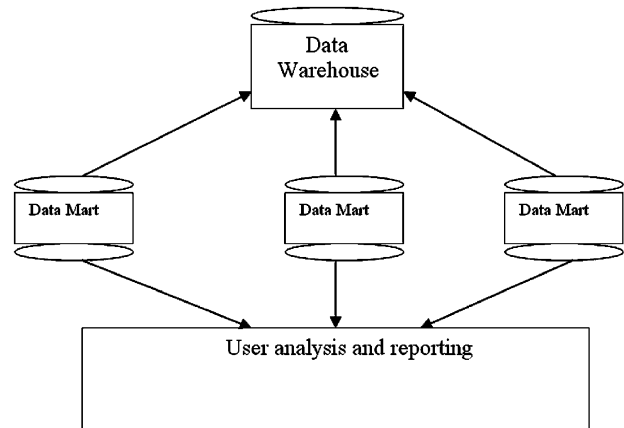


Fig. Bottom Up Architecture, Top Down Implementation

Here, if data marts are directly available then users can use them for analysis and reporting without the need of even creating a data warehouse. If companies can improve the performance of data marts separately then it will be quite beneficial for the users.

4. CONCLUSION AND RESULT

The architecture given in 2.1 is very common in industry. The one discussed in 3 also can be used for various applications which require fast processing. The architecture of 2.3 is not of much use.

REFERENCES

- [1] en.wikipedia.org/wiki/Data_warehouse
- [2] Data Warehousing and OLAP: a Research-oriented Bibliography by A. Mendelzon, C. Hurtado, and D. Lemire.
- [3] <http://db.stanford.edu/pub/papers/warehouse-research.ps>