



A Conceptual Architecture of Faster Data Retrieval from Multimedia Data Warehouse

Devendra Singh Shekhawat

MBA-IT, SICSR, affiliated to Symbiosis International University (SIU), Pune, Maharashtra, India

Dr. Pravin S. Metkewar

Assoc. Professor, SICSR, affiliated to Symbiosis International University (SIU), Pune, Maharashtra, India

ABSTRACT

Since the term data warehousing was coined it became an essential part of an enterprise. The growing usage of multimedia data for decision making in the past few years has led to generation of Multimedia Data Warehouse. Multimedia data warehousing is a concept through which useful information is fetched from the data like images, graphics, videos and audios. These kind of data needs to be managed rightly so that it does not affect the overall performance of the system. It should be studied in order to provide an accurate environment to the enterprise so that it can make efficient use of its multimedia data and it gets stored, accessed and analyzed in a proper manner. In this paper, we propose a faster data retrieval technique from the data warehouse.

KEYWORDS : Multimedia Data Warehouse, Hashing Technique, Hash Function, Hash Key, Data Warehouse, Data Processing, Architecture, Models

INTRODUCTION

Data Warehousing is used for storing huge amount of data into a central place which is further used for decision making and analysis. A Data warehouse built by integrating large amount of data from multiple heterogeneous sources and it is organized in such a way to provide business with vital strategic information. Data warehouse supports ad hoc reporting, decision making and analytical reports. Multimedia data warehouse is as the name suggests a data warehouse which stores multimedia data like videos, audios, images and graphics and other data sets which are converted into digital format from some other data format. In the fields like medical or bioinformatics, multimedia data is very essential and is used as valuable information to produce results. Multimedia data warehouse uses star and snow flake schema for storing of the data in warehouse.

At the beginning the data warehouses had to deal with only numerical and textual data but with the change in time a demand for storing



multimedia data rose ergo led to a demand of such kind of data warehouse which is capable of storing multimedia data like images, videos, audios in to it.

As we all know the data is not limited to numeric or textual form in today's business scenario. It has spread his wings and now also includes maps, images, audio, videos tapes etc. Everyday new architectures and frameworks are being

proposed for multimedia data warehouse just to achieve more efficiency and better processing of the multimedia in the warehouse.

The focus of this paper is to address the retrieval speed of data from the data warehouse. What is generally used and what we can use instead of that to speed up the process.

CREATING MULTIMEDIA DATA WAREHOUSE

Data warehousing is a simple concept originated because of the need of strategic information grew day by day. It is a very result of

this need. Traditional approaches as we all know we not efficient, time consuming, confusing and provided a high level on dissatisfaction. Something new was needed which could perform operations on existing data and produce fresh data for analysis. With the growing world the internet became the new hub. Organizations and People started sharing multimedia data over the internet and with this rose the need of storing the multimedia data which led to the birth of multimedia data warehouse.

Following are few steps to be considered while creating multimedia data warehouse:

1.1- Clear Business Objectives

Prior to the setting up a data warehouse the company should clearly define what is has to achieve, where it is now and where it wants to be. With a clear vision, mission and objective it will be able to filter out what is necessary for the organization and what is not.

1.2- Data Extraction

The most important function of a data warehouse is to pull out data from various source and sometimes it could be tricky because the organization should have a clear idea about what data is to be extracted. Data is extracted from various sources which can be internal or external. Internal sources can be CRM, HRM etc. and external could be the business partners, social sites etc.

1.3- Data Transformation

After the data is extracted another important part is to filter the data and only relevant data should be stored in the multimedia data warehouse which is done under the process called data transformation. Organization should clearly define what data is relevant to them and the data which is irrelevant to should be ignored.

- Data Model

data model comes dimension models which further consists of facts and dimensions. The dimensional approach provides improved query processing without affecting the integrity of the data. It is much faster and efficient than the relational approach. Fact tables have foreign keys of each dimension table and it also have measures. The measures are nothing but a factual representation of how well an organization is doing. Dimension are what the users want in the reports they want to generate.

- Data Loading

In the loading part the data converted, merged and brought on the same ground is injected or loaded into the data warehouse. At this part we can separate hash keys to each record using hashing function and these hash keys will make the users to retrieve the data in a much faster manner while fetching the data out of the data warehouse

- Generating Analysis

Here using various tools and techniques we fetch data according to our requirement from the data warehouse. Various business intelligence tools are used by the end users and queries are fired to generate analysis. Queries result in the data fetched from the data warehouse.

2. PROBLEM STATEMENT

With this paper, we try to propose an alternative for fetching data from data warehouse. Commonly indexing is used for the purpose of getting the data as result.

Let's first understand what is indexing and how it works?

A good example to make everyone understand indexing is the index of books. You can refer to the index to go to particular topic.

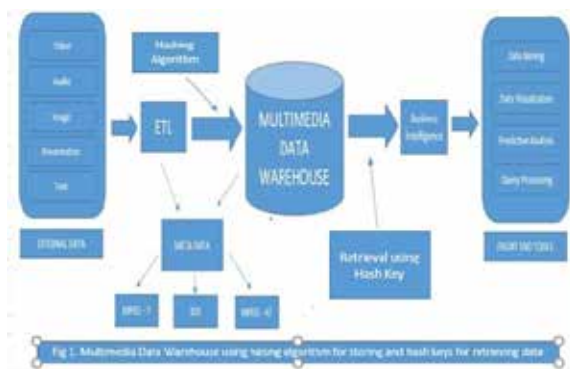
Indexing are basically lookup tables which makes retrieval of data possible in a fast manner. The problem with using indexing is that searches every record for the searched query and then retrieves the desired record.

3. PROPOSED SOLUTION

The proposed architecture we have suggests replaces indexing with hashing techniques to retrieve data. Hashing techniques uses hash functions, it basically transforms a string input value into a shorter fixed-value key which represents the original input value and is later on used for retrieving the data.

Hash function works like associative arrays, they put key value for each of the record and when called the database search engine looks for the key instead of searching in each record as is done in indexing this makes the retrieval time very less in comparison to the indexing approach.

For instance, you have to search the term "Matthew" in a table. In indexing approach, the database search engine will go on searching in every record for the term and when found it will return the values. This is time consuming if we compare it to hashing technique. When we use hashing technique it allots a key to every record and for the above example it will allot a key like "023201" to the record "Matthew" and when called upon it will just search for the key to retrieve the data.



CONCLUSIONS

The paper has tried to suggest a framework for multimedia data warehouse using hashing technique to retrieve data. At the users end the time taken for extraction could be more than expected while using indexing and to reduce this it should be replaced with hashing as it takes lesser time to fetch resulted data. With the proposed approach we can curb this problem to a certain extent.

REFERENCES:

[1] Bhavani Thuraisingham, "Managing and Mining Multimedia Databases," The MITRE Corporation, Bedford, MA
 [2] Dr. S.Vijayarani and Ms. A.Sakila, "Multimedia Mining Research – An Overview", Department of Computer Science, Bharathiar University, Coimbatore.
 [3] Mital Vora, Jelam Vora Dr. N. N. Jani, "Modelling Architecture for Multimedia Data Warehouse, Assistant Professor, Department of Computer Science, T. N. Rao College of I.T.,

Rajkot, Gujarat, India¹ Assistant Professor, Department of Computer Science, T. N. Rao College of I.T., Rajkot, Gujarat, India² Director, SKPIMCS - MCA, Gandhinagar, Gujarat, India³
 [4] ¹Meenakshi Srivastava, ²Dr. S.K.Singh, ³Dr. S.Q.Abbas, "An Architecture for Creation of Multimedia Data Warehouse"¹ Assistant Professor, Amity University , Lucknow Campus, India, ² Professor, Amity University Lucknow Campus, India, ³ Director, Professor, Ambalika Institute of Management and Technology, Lucknow, India