International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 4 Issue: 1 123 - 125

Using Network Model Represent Metadata in Data Warehouse

Sampa Das, Assistant Prof., CSE Dept.

Abstract **-** Using network model metadata representation have becomes a necessity not only for better knowledge, but also to handle the overall database of a huge numbers of information This work is aimed at represent metadata using network model. A model network is prepared using the entity which is present in the database.

Keywords:-*Data Warehouse, Metadata, Online Transaction Processing, Online Analytical Processing.*

*****

I. INTRODUCTION

Metadata plays a key role in converting raw data into knowledge as it helps to provide valuable description about the data so data it can be understood and converted into meaningful information. Metadata is simply defined as data about data. The data that is used to represent other data is known as metadata. For example, the index of a book serves as a metadata for the contents in the book. In other words, we can say that metadata is the summarized data that leads us to detailed data. In terms of data warehouse, we can define metadata as follows.

- Metadata is the road-map to a data warehouse.
- Metadata in a data warehouse defines the warehouse objects.
- Metadata acts as a directory. This directory helps the decision support system to locate the contents of a data warehouse.

A metadata does not give you just the description of the entity but also gives other details explaining the syntax and semantic of data elements. Metadata describes all the pertinent aspects of the data in the data warehouse fully and precisely to help the users and the developers of the data warehouse. A typical metadata contains information about the following:

- Structure of data from the programmer's perspective.
- Structure of data from the end-user's perspective
- Source systems that feed the data warehouse.
- Transformation process that was applied before the data from the source system could pass into the data warehouse.
- Data model.
- History of data extraction process.[1]

Metadata describes the details about the data in a data warehouse. Such a description may be in terms of the contents and source of data flows into the data warehouse. Metadata repository is an integral part of a data warehouse system. It contains the following metadata:

Business metadata - It contains the data ownership information, business definition, and changing policies.

Operational metadata - It includes currency of data and data lineage. Currency of data refers to the data being active, archived, or purged. Lineage of data means history of data migrated and transformation applied on it.

Data for mapping from operational environment to data warehouse - It metadata includes source databases and their contents, data extraction, data partition, cleaning, transformation rules, data refresh and purging rules.

The current paper proposes a network model inspired approach toward the evaluation of metadata. The flexibility that the system allows would be limited to the domain of knowledge supported by a tree model. Presently the work is restricted only in warehouse database.

Data warehouse

Data warehouse is a collection of data; these data are copied from other systems and assembled into one place. Once assembled it is made available to end users, who can use it to support a plethora of different kinds of business decision support and information collection activities [15]. DWs are central repositories of integrated data from one or more disparate sources. They store current and historical data and are used for creating analytical reports for knowledge workers throughout the enterprise. Data warehouse contains critical metrics of the business processes store along business dimensions.[10]

II. LITERATURE SURVEY

Various methods of managing collections of data (e.g., databases) have been developed since data was first stored in electronic form. From initial systems and applications that simply collected data in one or more flat database files to

present sophisticated database management systems (DBMS), different solutions have been developed to meet different requirements. A database may be considered distinct from a DBMS, in that a DBMS, as the name implies, includes utilities for accessing, updating and otherwise managing or operating a database. As the amount and complexity of data stored in databases has increased, DBMS design and development efforts have increasingly focused upon the ability to organize, store and access data quickly and efficiently. As a result, today's database management systems can be very effective in managing collections of linear information such as inventory, customer lists, etc [14].

The information professional has spent a life dedicated to process and functional analysis, user requirements, maintenance archit- ectures, and the like. Simply from the standpoint of who needs help the most in terms of finding one's way around data and systems, It is expected that the information technology community is computer literate, and able to find his/her way around systems. it is assumed the DSS analysis community requires a much more formal and intensive level of support than the information technology community. For this reason alone, the formal establishment of ongoing support of metadata becomes important in the data warehouse environment [3].

Huynh et al. [4] propose the use of metadata to map between object oriented and relational environment within the metadata layer of an object-oriented data warehouse.

Eder et al. [5] propose the COMET model that registers all changes to the schema and structure of data warehouses. They consider the COMET model as the basis for OLAP tools and transformation operations with the goal to reduce incorrect OLAP results.

Stohr et al. [6] have introduced a model which uses a uniform representation approach based on the Uniform Modeling Language (UML) to integrate technical and semantic meta data and their inter depen-dencies.

Katic et al. [7] propose a model that covers the security-relevant aspects of existing OLAP/ data warehouse solutions. They assert that this particular aspect of metadata has seen rather little interest from product developers and is only beginning to be discussed in the research community.

Shankaranarayanan & Even and Foshay et al. [8] provide a good description of business metadata and associated data quality. They argue that managerial decision-making stands to benefit from business metadata.

Kim et al. [9] provide a general overview of a metadata-oriented methodology for building data warehouses that includes legacy, extraction, operational data store, data warehouse, data mart, application, and metadata.

Our work covers a broad range of metadata aspects. We provide a means to manage data warehouse refresh observation timestamps, capturing message logs to detect any load or data issues. We also discuss in detail how to control individual job run behavior of subsequent batch cycles runs.

III. METADATA IN DATAWAREHOUSE

Metadata is a data about data in the data warehouse. Metadata in a data warehouse is similar to the data dictionary. In the data dictionary hold the information about files and addresses, information about the indexes, information about the logical data structure. Similarly metadata components are the same as data dictionary.
Metadata in a data warehouse fall into three major categories:

- Operational metadata
- Extraction and transformation metadata
- End-user metadata

Operational metadata contains all of the information about the operational data source.

Extraction and Transformation metadata contains information about all the data transformations that take place in the data staging area.

The end-user metadata allows the end-users to use their own business terminology. [10]

Basically, a metadata dictionary describes the following:

1. common metadata meanings (*semantics*),
2. common grammar and rules for expressing data (*syntax*),
3. commonly defined metadata dictionary element properties (*attributes*)[1]

Some metadata models or schemas are based on a logical, hierarchical arrangement of their metadata elements, not only in the way they are conceptually presented, but also in how they are applied in actual metadata and asset management systems.

IV. NETWORK MODEL

The network model is a database model conceived as a flexible way of representing objects and their relationships. Its distinguishing feature is that the schema, viewed as a graph in which object types are nodes and relationship types are arcs, is not restricted to being a hierarchy or lattice.
The network model organizes data using two fundamental concepts, called *records* and *sets*. Records contain fields (which may be organized hierarchically). Sets define one-to-many relationships between records: one owner, many members. A record may be an owner in any number of sets, and a member in any number of sets.[11]

A set consists of circular linked lists where one record type, the set owner or parent, appears once in each circle, and a second record type, the subordinate or child, may appear multiple times in each circle. In this way a hierarchy may be established between any two record types, e.g., type A is the owner of B. At the same time another set may be defined where B is the owner of A. Thus all the sets comprise a general directed graph (ownership defines a direction), or *network* construct. Access to records is either sequential (usually in each record type) or by navigation in the circular linked lists.[11]

The operations of the network model are navigational in style: a program maintains a current position, and navigates from one record to another by following the relationships in which the record participates. Records can also be located by supplying key values.

Although it is not an essential feature of the model, network databases generally implement the set relationships by means of pointers that directly address the location of a record on disk. This gives excellent retrieval performance, at the expense of operations such as database loading and reorganization.

V. CONCLUSION

In this paper provides the knowledge about metadata, data quality is increase if metadata is properly managed in data warehouse. We also define metadata using network model, so that reason it is easy to understand how to manage all the data and how to all the data are related each other in the data warehouse. Users easily can find the answer to the question regarding the data. It gives the internal view of the data warehouse.

REFERENCES

[1] Reema Thereja, "Data warehouse", Published by Oxford University Press, ISBN- 13:978-0-19-5699616
[2] International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.3, No.3, May 2013 'A CONCEPTUAL METADATA FRAMEWORK FOR SPATIAL DATA WAREHOUSE'
[3] Journal of Computing and Information Technology - CIT 20, 2012, 2, 95–111 doi:10.2498 /cit .1002046 'An ETL Metadata Model for Data Warehousing'
[4] T. N. HUYNH, O. MANGISENGI AND A. M. TJOA,Metadata for Object-Relational Data Warehouse, in Proceedings of the International Workshop on Design and Management of Data Warehouses(DMDW'2000), Stockholm, Sweden, June 5–6, 2000.
[5] J. EDER, C.KONCILIA AND T. MORZY, The COMET Metamodel for Temporal Data Warehouses, in Proceedings of the 14th International Conference on Advanced Information System Engineering (Caise 2002), Toronto, Canada, LNCS 2348, pp. 83–99.
[6] T. STOHR, R. MULLER AND E. RAHM, An Integrative and Uniform Model for Metadata Management in Data Warehousing Environments, in Proceedings for the International Workshop on Design and Management of Data Warehouses (DMDW'99), Heidelberg, Germany, June 14–16, 1999.
[7] T. CHENOWETH, K. CORRAL AND H. DEMIRKAN, Seven Key Interventions for Data Warehouses Success, *Communications of the ACM*, January 2006, Vol. 49, No. 1.
[8] N. FOSHAY, A. MUKHERJEE AND A. TAYLOR, Does Data Warehouse End-User Metadata Add Value?, Communications of the ACM, November 2007, Vol. 50, No. 2.
[9] T. KIM, J. KIM AND H. LEE, A Metadata-Oriented Methodology for Building Data Warehouse: A Medical Center Case, Informs & Korms – 928, Seoul 2000 (Korea).
[10] Paulraj Ponniah ,"Data Warehouse",Wiley Publication,second Edition.
[11] "http://www.ustudy.in/node/12070"
[12] C. WHITE, Data Integration: Using ETL, EAI, and EII Tools to Create an Integrated Enterprise, The Data Warehousing Institute, October, 2005.
[13] N. RAHMAN, Refreshing Data Warehouses with Near Real-Time Updates, Journal of Computer Information Systems, Vol. 47, Part 3, 2007, pp. 71–80.
[14] Ravi Kothuri, Siva Ravada, Jayant Sharma, Jayanta Banerjee, "Relational database system for storing nodes of a hierarchical index of multi-dimensional data in a first module and metadata regarding the index in a second module" US6505205 B1.
[15] Jossen, Claudio, Dittrich, Klaus R,(2007) "The process of Metadata Modeling in industrial Data Warehouse Environments", BTW Workshops 2007, Aachen, Germany, pp16-27.