International Journal of Advances in Engineering and Management (IJAEM)
Volume 8, Issue 4 April. 2026, pp: 317-321 www.ijaem.net ISSN: 2395-5252

Methods to improve the performance of operational Data Warehousing

Dr G. Rajitha Devi
Associate professor in Computer Science

Date of Submission: 05-04-2026 Date of Acceptance: 16-04-2026

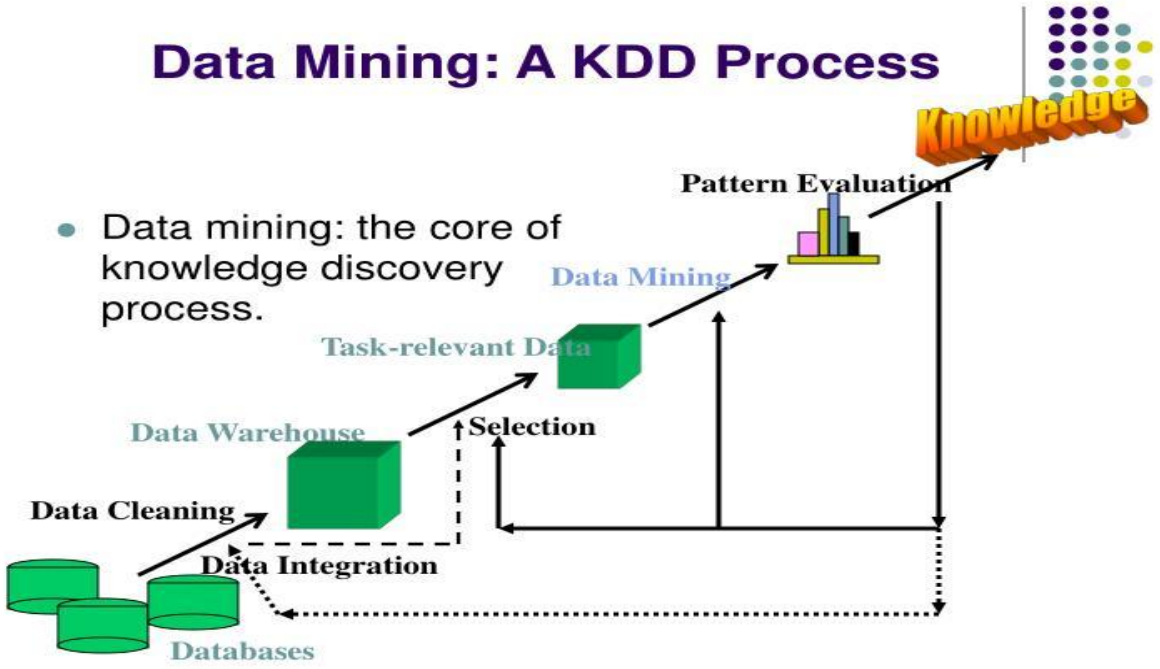
Abstract:
A Data Warehouse (DW) is a relational database designed for processing rather than fining transactions. It includes historical data obtain from transaction data from single and multiple origin. It is used not for daily operations and transaction processing but for decision making. "A data warehouse is a type of data management system designed to authorize business intelligence.It is designed to reinforce the function of data warehousing not only to store the data but it also provides the security of the dataThough, vast research has been done in the past on this technology, still more research is required to further improve its quality and application. In this paper we have tried to present outcome of thorough reference of details of research endeavor put up by various research groups in key areas of the technology like data warehouse modelling & design, ETL process designing, data warehouse maintenance, and OLAP design. The objective of our study is to determine past sequence of advancements, current state of the art, and future course of research needs in this domain.

Keywords –Data warehousing,Decision making, OLAP, Data mining, ETL

I. Introduction

Data warehousing is a process of customizing, design and implementation of a data repository capable of servingcomposite business queries for varied decision making task of the organization. Data warehousing systems arecapable of serving all types of queries ranging from simple business reporting, online analytical processing (OLAP), or a complex data mining query. These systems store and continue data for past several years which makes size oftables very large, thus results in slow query execution. Improvement in query execution time requires use of varioustechniques out of which transpire view maintenance is most effective and useful in which pre-computedcollection information is stored in the data repository to save run time processing loads on the server and producethe response of query faster. Data warehousing systems get data from operational system of the organization andfrom some external data sources also. Design of both types of systems is different. While the operational systems areplan for operational productivity of the organization's business process, the data warehousing systems aredesigned to support decision making activities of the organization. Differences in objectives of these systems, causesdifferent design approaches. While operational system schema is designed to hold atomic data of all businesstransactions, data warehouse schema is designed to hold non atomic data at different levels of grains depending onorganization's data warehousing needs. Collecting data from multiple sources requires cleaning for namingdiscord, data type mismatches, size differences, treatment to missing data values, removal of irrelevant data fromthe point of view of data analysis, and selection of applicable attributes for data analysis needs. These steps are keycomponents of extraction transformation and loading (ETL) process. Figure 1 gives idea about data warehousingprocess and data mining application to extract knowledge useful in decision making of an organization.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 8, Issue 4 April. 2026, pp: 317-321 www.ijaem.net ISSN: 2395-5252



A data warehouse can be viewed as a data system with the following characteristics

- It is a database that can be used for inspecting tasks using data from various applications.
- It supports a like small number of customers with relatively long interactions. Contains current and historical data to provide a historical perspective of the information. Its use is readable.
- It has some big tables.
- "A subject-oriented integrated and timely source of information in support of data warehouse management decisionsdifferent collection."

Subject-Oriented

A data warehouse aims at the pattern and analysis of data for decision-makers; therefore, a data warehouse is usually a brief and direct overview of a particular topic, such as a customer product or sales, rather than the global organization's ongoing operations

Integrated

A data warehouse with various heterogeneous data sources such as RDBMS, flat files and online Integrates transaction records. Data cleaning and integration is required during data warehousing to ensure consistency in naming conventions, types of attributes, etc. among different data sources.

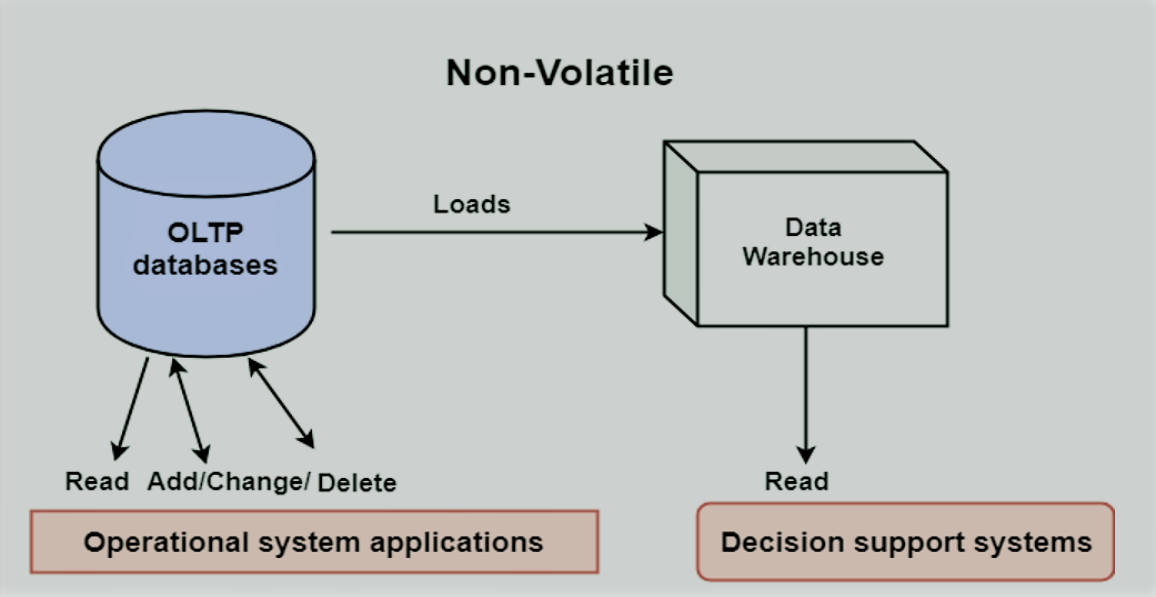
Time-Variant

Historical information is kept in a data warehouse. For example one can retrieve files of 3 months 6 months 12 months or even past data from a data warehouse. These change with the transactional system, where often only the most current file is kept.

Non-Volatile

A data warehouse is a physically separate data storage that is transformed from a source operational RDBMS. The operation of the data is not done in the updated data warehouse i.e. update, insert and delete operations are not done. Expressing data usually requires only two processes, the initial loading of the data and the access to the data. Therefore DW does not require transaction processing retrieval and concurrency capabilities which allows for substantial speed of data retrieval. Nonlexical defines that no more data should change once it has entered the warehouse.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 8, Issue 4 April. 2026, pp: 317-321 www.ijaem.net ISSN: 2395-5252



The idea of data warehousing originated in the late 1980s when IBM researchers Bury Devlin and Paul Murphy founded Business Data Warehouse. Briefly, the data warehousing idea was planned to support an architectural model for the flow of information from the operational system to the decision support environment. The concept tries to address various problems associated with flow. Mainly the high cost associated with it.The absence of a data warehousing architecture requires a large amount of space to support multiple decision support environments. It was common in large corporations to operate independently of various decision support environments.

Advantages

1. It increases Business Intelligence (M). Because there is a lot in this due to which business decision making has to be done.
2. It saves time. Its support can be done very fast, due to which time and money is spent.
3. It enhances the quality and consistency of the data
4. It Generates High Roll
5. It improves decision-making because it has high quality

Types of Data Warehouse It has three main types which are as follows

1- Enterprise Data Warehouse (EDW)

It is a centralized. Warehouse provides decision support service. It provides a unified approach (integrated approach) to organize and replicate data

2 Operational Data Store (ODS)

It is a data store and itsSupport system (DSS) is used as an alternative.

3 Data Mart

Data Mart is a subset of data warehouse; it is designed for specific purposes such as sales finance etc. In independent data mart, data is collected directly from sources.

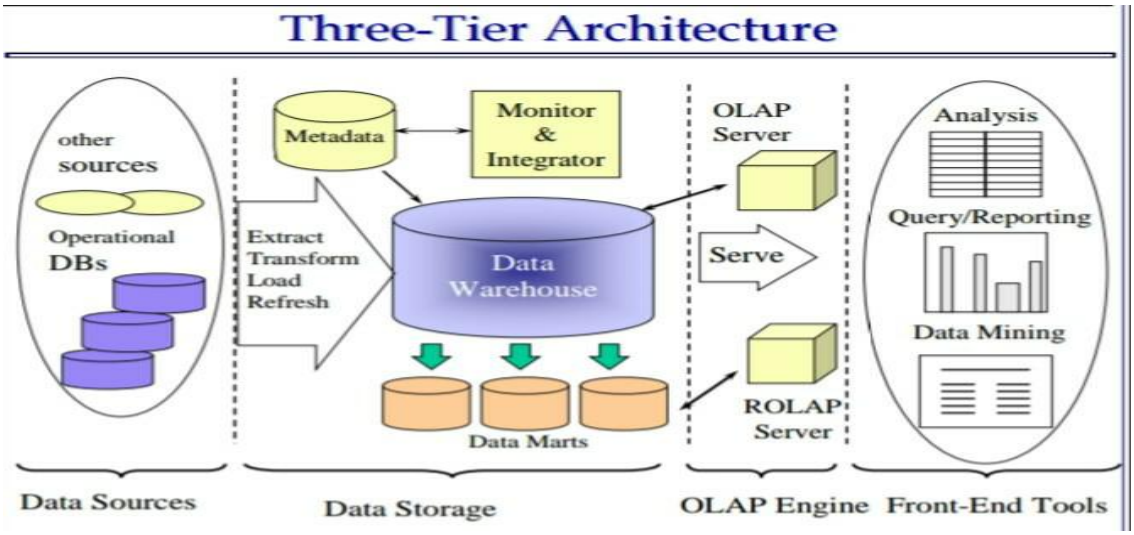
There are three types of data marts.

1. Dependent
2 Independent
3. Hybrid

Data warehouse architecture

A data warehouse architecture is a design that is maintained .We can maintain any purpose through the tractor, which the architecture is keeping in mind the need of the light and the goal of the employee.There are many different types of data warehouses.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 8, Issue 4 April. 2026, pp: 317-321 www.ijaem.net ISSN: 2395-5252



1. Bottom tier: The lowest layer contains the database server which is mostly always a relational one, in which the back end devices are used to load the data in the bottom tier and this data is taken from the operational database and external sources. Handles the ETL process of Mahavegas ETL process is the process in which the datais extracted, transformed and loaded.

ETL (extract transform and load process is performed by back end tools

2: Middle tier

This tuer is also called application tier, this layer handles the request of all the users. In this layer B (Business Intelligence) This OLAP server is implemented in the following two ways.

1- ROLAP (Relational OLAP)

These servers are intermediate servers which are relational back end servers and the lights are located between the front end dowels

2. MOLAP (multidimensional OLAP) This server supports multi-dimensional view of data.

3. Top Tier This tier is also called presentation tier, it is the front end client layer. This

The layer contains query tools, reporting tools, analysis tools, and eta mining tools. This layer interacts with the end users, that is, the data reaches the end users through this layer.

II. Conclusion:

Data warehousing applications provide these features to business users of the organizations. To have highly efficient data warehousing solutions, many aspects of this technology needed research initiatives that have been put up by the research community. Total forty research papers have been referred. Major contributions are seen in the area of data warehouse modeling & design, data warehouse accessing technologies and data warehouse maintenance. From the review of literature it is clear that in the beginning major focus was on modeling and design

and by now this aspect of data warehousing has matured enough. Though, the work in data accessing techniques has also been done but with the ever increasing volume of data, we need to work further in the direction of data accessing and visualization.

REFERENCES

[1]. K.Y. Lee and M.H. Kim. “Optimizing the Incremental Maintenance of Multiple Join Views”, Eighth ACM International Workshop onData Warehousing and OLAP, ISBN: 1-59593-162-7, ACM Press, NY, 2005.

[2]. F. Dehneet al. “Parallel Querying of ROLAP Cubes in the Presence of Hierarchies”, Eighth ACM International Workshop on DataWarehousing and OLAP, ISBN: 1-59593-162-7, ACM Press, NY, 2005.

[3]. Jens Lechtenborger. “Issues in Data Warehouse Modeling”, Eighth ACM International Workshop on Data Warehousing and OLAP, ISBN:1-59593-162-7, ACM Press, NY, 2005.

[4]. Fan, H., &Poulovassilis, A., “Using AutoMed Metadata in Data Warehousing Environments DOLAP’03”, November 7, 2003, NewOrleans, Louisiana, USA. Copyright 2003 ACM 1-58113-727-3/03/0011 ...$5.00.P86-93.

[5]. Efficient incremental view maintenance in data warehouses. Ki Yong Lee, Jin HyunSon,

Myoung Ho Kim. Korea Advanced Institute of Science and Technology Document.

[6]. Vishal Gour, Dr.S.S.Sarangdevot, Govind Singh Tanwar "Performance Tunning Mechanisms for Data Warehouse: Query cache"in International Journal in Computer Application Volume 2 – No.2, May 2010 pp. 70-75

[7]. R. Winter and B. Strauch.A method for demand-driven information requirements analysis in data warehousing Projects. In Proc. HICSS, Hawaii, 2003 pp.1359–1365,.

[8]. Albrecht, J.; Günzel, H.; Lehner, W.: Set-Derivability of Multidimensional Aggregates, in: Proceedings of the First International Conference on Data Warehousing and KnowledgeDiscovery (DAWAK'99, Florence, Italy, August 30 - September 1), 1999

[9]. Albrecht, J.; Bauer, A.; Deyerling, O.; Günzel, H.; Hümmer, W.; Lehner, W.; Schlesinger, L.: Management of multidimensional Aggregates for efficient Online Analytical Processing, in: International Database Engineering and Applications Symposium (IDEAS'99, Montreal, Canada, August 1-3), 1999

[10]. Cohen, S.; Nutt, W.; Serebrenik, A.: Rewriting Aggregate Queries Using Views, in: 18th Symposium on Principles of Database Systems (PODS'99, Philadelphia, Pennsylvania, USA, May 31 - June 2), 1999

[11]. Yu, C.T.; Sun, W.: Automatic Knowledge Acquisition and Maintenance for Semantic Query Optimization, in: IEEE Transactions on Knowledge and Data Engineering (TKDE),1989

[12]. V. Markl and R. Bayer. Processing Relational OLAP Querieswith UB-Trees and Multidimensional ierarchical Clustering. InProceedings of DMDW 2000, June 5-6, 2000.