International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 05 | May 2021 www.irjet.net p-ISSN: 2395-0072

Overview of Data Warehouse

Saadhvi Hosmane[1], Rakshitha S[2]

[1] *Student, Department of Information Science and Engineering, Global Academy of Technology, Bangalore.*
[2]*Student, Department of Information Science and Engineering, JSS Academy of Technical Education, Bangalore.*

Abstract – *In latest years, the database community has seen major developments in a new technology: Data Warehouse. A data warehouse is a special repository that stores pre-processed queries on data which resides in many legacy, heterogenous, or operational sources. The foremost goal of this paper is to introduce the learners to the basic concepts and terminology associated with data warehouse.*

Key Words: **Data Warehouse, Repository, Data Marts, OLAP Tools, Unstructured Data, Metadata.**

1. INTRODUCTION

Have you ever had questions about the product recommendations you get when shopping online? For example, when you buy a mobile phone, it is recommended to buy other related products online (such as mobile phone cases, tempered glass, etc.). This is an example of data mining which is the process of discovering useful patterns in a huge data set. This huge data is created by integrating current and past data from different sources and store them centrally in a dedicated repository called *Data Warehousing (DW).*

A data warehouse means various things to various people. Some definitions are limited to data; other definitions apply to people, processes, software, and tools. One of the main definitions is: A data warehouse is a set of data that support decision making process. It provides the subsequent features: It is subject -oriented, integrated, consistent and it shows evolution over-time, and it is not volatile.

In other words, we can be said that a data warehouse is a place where the data is stored for archival, security and analysis reasons. Usually, one or more computers are linked together to create a huge computer system**.** There may be raw or formatted data related to various topics, including company sales, salary, operating data, summary data, reports, data copies, data from the human resources department, inventory data, and external data to provide simulations and analysis etc. Data warehouse system is also known by the following names:

- Decision Support System (DSS).
- Executive Information System.
- Management Information System.
- Business Intelligence Solution.
- Analytic Application.

The data in a warehouse never updated but used only to respond to queries from end users who usually make decision. Typically, data warehouses are huge and contain billions of records. In many cases, an organization may have multiple local or departmental data warehouses often called "data marts". A data mart is a data warehouse that is designed to meet the needs of a specific group of users. Depending on the subject area, it can be large or small. The primary advantage of a data warehouse is that it provides easy access to and analysis of vast stores of information on many subject areas.

The remaining part of this paper is organized as follows:
In Section 2, we discuss about the literature survey. We discuss the evolution of the data warehouse in Section 3 and then followed by the need and working of the data warehouse in the Section 4.

In Section 5, we highlight the characteristics of the data warehouse, following which we have discussed about the data warehouse architecture and types in the Section 6. In the Section 7, we highlighted the five main DW components. Section 8 provides a brief information about the three main types of DW.

Usually there are some misconceptions regarding data warehouse with other concepts such as data lakes, data marts and database which are highlighted in Section 9. Finally, we have discussed about the applications of the data warehouse, benefits & drawbacks, future aspects of the data warehouse in the Section 10,11,12, respectively.

2. LITERATURE SURVEY

Data warehouse is a relational database management system (RDBMS), which is specially developed for the requirements of transaction processing system. Once the relevant data is collected and stored in the data warehouse, the potential of the data warehouse can be enhanced.

The main advantage of a data warehouse is that it provides easy access and analysis of vast storage of information on various subject areas. In [1], the author W.H. Inmon describes the evolution of decision support systems, like how data warehouse environment is, and going further, he also describes the data warehouse design, granularity, advanced topics related to data warehouse and its future as well.

In [2], Rajan Tiwari gives detailed information about the Data Warehouse, where he describes its meaning, chara-

cteristics, architecture and all the related concepts for building a data warehouse. He also speaks about OLAP.

[4] Here, the author Jake Frankenfield mainly focused on working of data warehouse, he also discusses about data mining process. A data mining is a process which is used to convert raw data into useful information by business organizations.

The authors Matteo Golfarelli and Stefano Rizzi in [6], speaks about the basic definitions used in this data warehouse field. They also describe about lifecycle of data warehouse systems and suggested a methodological approach for designing them. They also focus on how data is extracted from sources, transformed, cleansed, and then used to populate DW.

In [14] and [18], the author Sumit Thakur discusses about the applications and disadvantages of the data warehouse, respectively. The benefits of the data warehouse are discussed by respective authors in [15]-[17].

The future of the data warehouse is cloud based. According to author of [20], many believe that the data warehouse is alive. Many organizations are using data warehouse and there maybe more than one. Since the technology is improving new data sources, new database technologies and new users of data come into picture. There might be a question raised regarding how the data warehouse can fit into the modern data management?. This is an architectural question that challenges the original positioning of the data warehouse. The author speaks about his thoughts regarding positioning of the data warehouse in the modern data management architecture. Actian, Denodo, Slamdata, Wipro give their views on the future of data warehouse.[22]

3. EVOLUTION OF DATA WAREHOUSE

The Data warehouse helps the users to understand and improve the performance of their organization. As computer systems become more complex and necessary to process more and more information the demand for data warehouse. Here are few key occasions in evolution of Data Warehouse:

- In 1960, Dartmouth and General Mills together in a research project developed the terms dimensions and facts.
- In 1970, A Nielson And IRI introduced dimensional data marts for retail sales.
- Later in the year of 1983, Tera Data Corporation introduced a DBMS which was particularly designed for decision support.
- Finally, in late 1980s Data warehousing was started when IBM workers Paul Murphy and Barry Devlin developed the Business Data Warehouse.

However, the actual concept was given by *William H Inmon.* He is considered the *father of Data Warehouse*. He has written many articles on various topics related to the building, usage, and maintenance of the warehouse and the Corporate Information Factory.

Factors, which lead to data warehousing are as follows: Many factors have affected the rapid development of data warehouse discipline. The most important factor is the advancement of hardware and software technology. The price of software and hardware has dropped significantly. Higher capacity memory chips can be obtained at a very low price. Factors are: *Powerful Preprocessors, Inexpensive Disks, Desktop Powerful for analysis tools, Server Software.*

The dazzling performance of hardware and software and the availability of affordable and easy-to-use reporting and analysis tools played a vital role in the evolution of data warehouses.

4. NEED AND WORKING OF DATA WAREHOUSE

An ordinary database can store MBs to GBs of data, which is used for specific purposes. To store larger data like TBs of data, the storage must be shifted to data warehouse. In addition to this, a transactional database does not provide itself for analytics. So, in order to conduct effective analysis, a company needs to maintain a central data warehouse to fully understand its business by organizing, studying and using its historic data to make strategic decisions and analyze the upcoming trends.

By comparing consolidated data from multiple different sources, the data warehouse can be used to better understand a company's performance. The data warehouse is used to run query and historical data from transaction sources. When data enters the warehouse, it does not change, and it cannot because the data warehouse will analyze the events that have occurred and pay special attention to the changes in the data over time. Generally, data is never deleted from the data warehouse and updates are usually performed when the data warehouse is offline. This means that the data warehouse can be treated as a read-only database. The Warehoused data here is safe, reliable, easy to access and easy to manage.

To create a data warehouse, certain steps must be followed. The first step is data extraction, this involves collecting large amount of data from multiple sources. After the data is compiled, it will go through data cleaning, it is a process in which the data will be checked for errors, and the errors found will be corrected or eliminated. Then the cleaned-up data is converted from a database format to warehouse format. Once the data is stored in the warehouse, it is sorted, merged, summarized, etc. for more coordination and ease of use. Over time, as multiple data sources are updated, more data will be added to the warehouse.

5. CHARACTERISTICS OF DATA WAREHOUSE

W. H. Inmon characterized a data warehouse as *"a subject oriented, integrated, non-volatile, time-variant collection of data in support of management's decisions".* The data warehouse provides access to data for complex analysis, discovery, and decision-making. A data warehouse has the following characteristics:

- **Subject-Oriented**: Data are organized by subject rather than purpose e.g., Insurance companies using data warehouses organize data based on customers, premiums, and claims rather than different products. The data sorted by subject contains only the information needed to process decision support.
- **Integrated***:* In a data warehouse, integration means establishing a common unit of measurement for all similar data from different databases. The data should also be stored in the data warehouse in a common and generally acceptable way. This integration facilitates effective data analysis.
- **Time-Variant:** The time frame of the data warehouse is much longer than the time frame of the operational systems. Each key structure in the data warehouse contains a time element (explicitly or implicitly). The data collected in a data warehouse is recognized with a particular period and the information will be provided from a historical point of view.
- **Non-Volatile:** The data warehouse is also non-volatile, which means that the old data will not be deleted when new data is entered into it. The data is read-only and updated regularly. In data warehouse environment, operations such as delete, update, and insert that are performed while the application is running will be ignored. Only two types of data operations are performed in the data warehousing: *Data Loading and Data Access.*

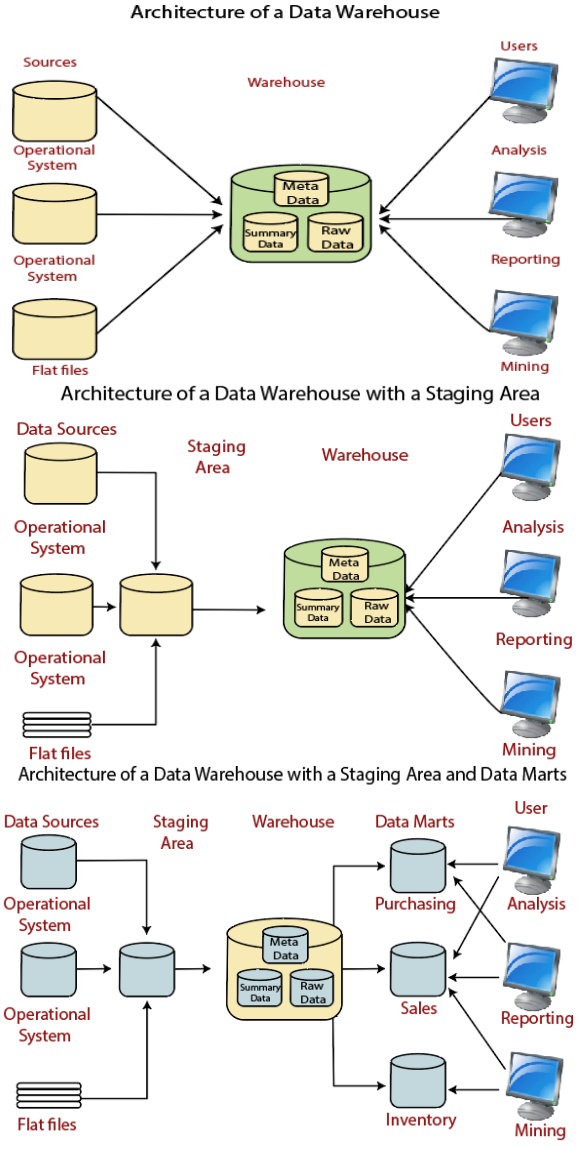
6. DATA WAREHOUSE ARCHITECTURE

Data warehouse architecture is a method to define all computing and representation architectures of end-user computing in an enterprise. Every data warehouse is different, but they all have important standard components. The data warehouse and its architecture largely depend on the elements of the business situation. The following architecture properties are very much necessary for a data warehouse system (Kelly,1997):

- **Separation:** Analysis and transaction processing should be carried out separately.
- **Scalability:** Hardware and software architectures should be simple to upgrade the data volume.
- **Extensibility:** The architecture should be able to handle new processes and technologies without having to redesign the entire system.
- **Security:** Monitoring accesses are necessary because of the strategic data stored in the data warehouses.
- **Administrability**: The Data warehouse management should not be difficult.

There are three common architectures namely,

- Basic Data Warehouse Architecture.
- Data Warehouse Architecture with Staging Area.
- Data Warehouse Architecture with Staging Area & Data Marts.



[24] Fig-1: Data Warehouse Architecture.

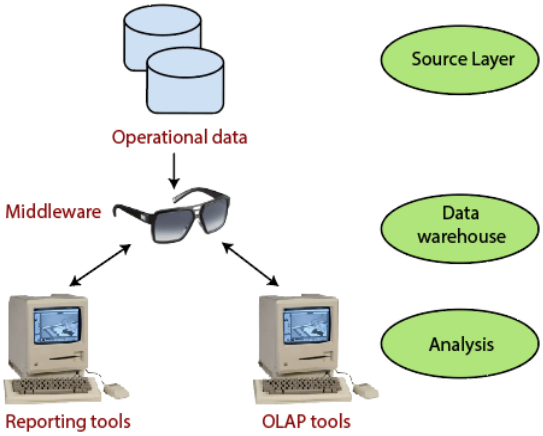
6.1 TYPES OF DATA WAREHOUSE ARCHITECTURE

Data warehouse Architectures are mainly of three types:

- Single-Tier Architecture.
- Two-Tier Architecture.
- Three-Tier Architecture.

Single-Tier Architecture: Single-Tier Architecture is not often used in practice. The goal is to minimize the amount of stored data to achieve this goal, it eliminates data redundancy. The Fig-2 shows that the only layer which is physically available is the source layer. In this architecture, the data warehouses are *virtual*, which means that data warehouse is enforced as a multi-dimensional view of operational data generated by specific middleware or process middleware.

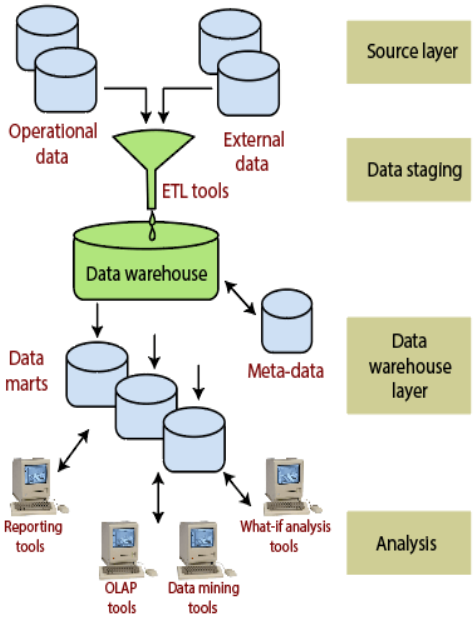
The weakness of this architecture is that it cannot meet the requirements for separation between analytical and transaction processing. After the middleware interprets it, it sends the analysis queries to the operational data. Thus, queries will affect normal transactional workloads.



[24] Fig-2: Single-Tier Architecture.

Two-Tier Architecture: The requirement for separation plays an important role in defining the two-tier architecture for a data warehouse system, as shown in Fig-3. Although it is usually referred to as a two-tier architecture to highlight the separation between physically available sources and data warehouses, it consists of four main sequential data flow stages: Source Layer, Data Staging, Data Warehouse Layer and Analysis.

- **Source Layer:** The data warehouse system uses heterogeneous data sources. The data is stored locally in the company's relational database or old database, or it can come from an information system outside the company.
- **Data Staging:** The data stored in the source must be extracted, cleaned up to eliminate inconsistencies and fill gaps, and integrated to combine the different sources into one standard architecture. *ETL (Extraction, Transformation and Loading)* Tools can combine heterogeneous patterns, extract, transform, clean, verify, filter and loads source data into a data warehouse.
- **Data Warehouse Layer:** Information is stored in a single, logically centralized repository: a data warehouse. The data warehouse may be accessed directly, however they can also be used as a source for creating data marts that replicate the contents of the data warehouse and target specific business units. Meta-data repositories store information on sources, access procedures, data staging users, data mart patterns, etc.
- **Analysis:** Here, integrated data can be accessed efficiently and flexibly to issue reports, dynamically analyse information, and model business scenarios. It should include an aggregated browser, complex query optimizers, and an easy-to-use graphical user Interfaces.



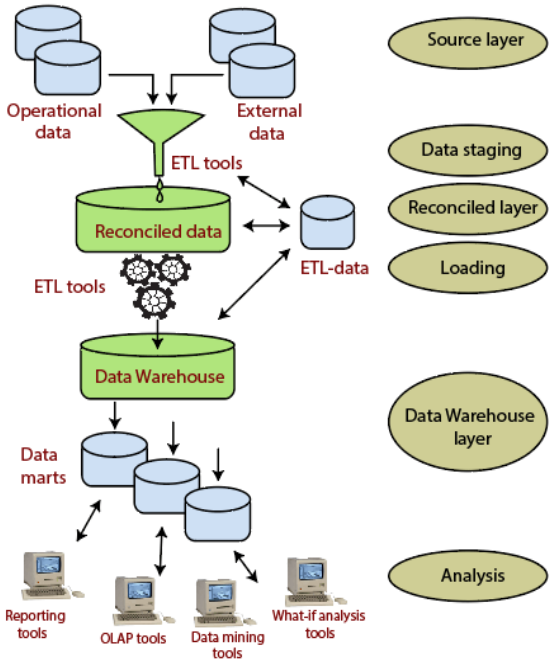
[24] Fig-3: Two-Tier Architecture.

A data mart can be seen as a small local data warehouses that can replicate and aggregate the part of a primary data warehouse required for a specific application domain. A data mart is a subset or collection of data stored in a primary data warehouse and is a collection of information related to a specific area of a company, department, or category of users.

Three-Tier Architecture: The three-tier architecture of the source layer (containing multiple source systems), the reconciled layer and the data warehouse layer (containing both data warehouse and data marts). The reconciled layer lies between the source data and data warehouse.

The main advantage of the reconciled layer is that it establishes a standard reference data model for the entire company. At the same time, it separates the problem of extracting and integrating source data from those of data warehouse population. In some cases, the consistency layer can also be directly used to better perform certain operational tasks, such as: for example, generating daily reports that cannot be correctly created using company applications, or generating data flows to provide external processes and to benefit from cleaning an integration. This architecture is particularly useful for large-scale enterprise systems.

The disadvantage of this structure is the use of additional file storage space used through the extra redundant reconciled layer. In addition to this, the analytical tools are still a little away from being real-time.



[24] Fig-4: Three-Tier Architecture.

7. DATA WAREHOUSE COMPONENTS

The data warehouse is based on the RDBMS server (a central information repository), which is surrounded by some important data warehousing components. A typical data warehouse consists of five main components: *a central data warehouse database, ETL tools, metadata, access/query tools and data marts.*

All these components are designed to run quickly, so you can get results and analyze data quickly anytime, anywhere.

Data Warehouse Database: The central data warehouse database is the foundation the data warehouse environment. The database is implemented with RDBMS technology. Traditionally, these are standard relational databases running locally or in the cloud.

Sourcing, Acquisition, Clean-up & Transformation Tools (ETL): The data sourcing, transformation, and migration tools are used for performing all necessary conversions, summarizations, and all the changes needed to transform data into a unified format in the data warehouse. They are also known as Extract, Transform and Load (ETL) Tools. Their functionalities include:

- Eliminating unnecessary data in the operational databases from loading into data warehouse.
- Search and replace common names and definitions for data arriving from various sources.
- In case of any missing data, they populate with the default. And calculate summaries and derived data.
- De-duplicated repeated data arriving from multiple data sources.

These Extract, Transform, and Load tools may generate cron jobs, background jobs, Cobol programs, shell scripts, etc. that regularly update data in data warehouse. These tools are helpful to maintain the Metadata. These tools must deal with challenges of the database and data heterogeneity. These tools can save a large amount of time and effort.

Meta Data: Metadata is data about the data that defines the data warehouse, and is used for creating, managing, and maintaining the data warehouse. In the data warehouse architecture, metadata plays an important role in defining the source, purpose, value, and attributes of the data in the data warehouse, as well as in defining how to modify and manipulate the data.

Metadata is mainly classified into two types:

Technical Metadata- This type of metadata contains information about the warehouse used by data warehouse designers and administrators.
Business Metadata- This type of metadata contains information that enables the end user to easily understand the information stored in the data warehouse.

Query Tools (Access Tools): One of the main tasks of the data warehousing is to provide the company with the information needed to make strategic decisions. Users use front-end tools to interact with the data warehouse. The query tool allows users to interact with the data warehouse system. These tools can be divided into four categories: *query and reporting tools, application development tools, online analytical processing tools (OLAP), and data mining tools.*

Query and reporting tools can be divided into two categories: reporting tools and managed query tools. Further the reporting tools are divided into production reporting tools and report writers.

Companies use these *production reporting tools* to generate regular operational reports or support large amount of work, such as calculating and printing paychecks. On the other hand, report writer is an affordable desktop tool front end user.

The *managed query tools* protect the end users from the complexities of SQL and database structure by adding a meta-layer between the user and the database.

Sometimes the built-in analytical and graphical tools cannot meet the company's analytical needs, in this case, *application development tools* will be used to develop custom reports.

In data mining, new and meaningful correlations, patterns, and trends are discovered by extracting large amounts of data. *Data mining tools* are used to automate this process.

OLAP tools are based on the concept of multidimensional databases, allowing users to use complex multidimensional views to analyze data.

Data Marts: The *data mart* is an access layer which is used for sending data to the users. It is an option for large data warehouses because it takes less time and money to build. In simple words, we can say a data mart is a *branch* of a data warehouse. Data marts can be created in the same database as the DW or a physically separate database.

8. TYPES OF DATA WAREHOUSE

Enterprise Data Warehouse (EDW): Enterprise Data Warehouse (EDW) is essentially a centralized warehouse. It guarantees choice assistance carrier at some point of the enterprise. It presents a coalesce technique for organizing and representing records. It additionally dispenses the potentiality to categorize records in line with the situation and provide get entry to in line with the one's divisions.

Operational Data Store: Operational records shop, usually known as ODS that is a records shop commonly required while neither records warehouse nor OLTP structures guide businesses reporting needs. In ODS, the records warehouse is refreshed right here and now. Due to this nature of ODS its especially favored for recurring sports like storing information of the employees.

Data Mart: A records warehouse subset is a records mart. The records may be without delay gathered from the assets in an unbiased records mart. It is designed for commercial enterprises together with income and finance.

9. WHAT A DATA WAREHOUSE IS NOT?

It is quite easy to be confused with the concepts of *data warehouse, database, data marts and data lakes*, but they are not the same.

DATA WAREHOUSE IS NOT A DATABASE

Data warehouse is not a database, but they do share some similarities. However, the main difference arises when companies need to *analyse large data sets*. The DW is designed for this type of problem, whereas the database is not.

DATA WAREHOUSE IS NOT A DATA LAKE

Both are used for business analytics purpose, but the main difference between a data lake and a data warehouse is that the data lake stores all types of raw, structured, and unstructured data from all data sources in a native format until needed. Whereas data warehouses store data in files or folders in a more organized way, which can be used for reporting and data analysis.

DATA WAREHOUSE IS NOT A DATA MARTS

Data warehouses are sometimes confused with data marts, but data warehouses are usually larger and contain more data, and data marts have limited uses. Data marts are usually subset of warehouse and can be used to easily deliver certain data to specific users of specific applications. In short, we can say that data marts can cover single-subject, while a data warehouse multiple subjects.

10. APPLICATIONS OF DATA WAREHOUSE

Here are the few applications of the data warehouses:

Banking Industry: In the banking enterprise, awareness is given to threat control and coverage reversal as nicely studying patron information, marketplace developments, authorized policies and reviews, and greater importantly monetary selection making. Most banks additionally use warehouses to control the assets to be had on deck powerfully. Certain banking sectors make use of them for marketplace studies, overall performance evaluation of every product, interchange and change rates, and to broaden advertising programs.

Government and Education: The federal authorities make use of the warehouses for studies in compliance, while the national authorities make use of it for offerings associated with human assets like recruitment, and accounting like payroll control.

The authorities make use of data warehouses to hold and examine tax data, fitness coverage data, and their respective providers, and additionally their whole crook regulation database is hooked up to the nation's information warehouse.

Universities use warehouses for extracting statistics used for the idea of studies grants, knowledge of their pupil demographics, and human useful resource control. The whole monetary branch of maximum universities relies upon data warehouses, along with the Financial Aid branch.

Healthcare: Healthcare is one of the most important sectors using data warehouses. All their financial, medical and employee data are sent into the warehouse for strategy formulation and predict the outcomes, track, and analyze their service reviews, helps to generate patient reports, share important data with the tied-up insurance companies.

The Retailers: The retailer acts as an intermediary between the producer and the consumer. Keeping records of both parties is important to ensure their presence in the market. They often use warehouses to track products, to advertise their promotions. They also analyze sales to determine fast selling and slow selling product lines and define their shelf space through a process of elimination.

Services Sector: Data warehouse are very useful in the service sector for maintaining financial reports, revenue structure, customer profiles and for resource management.

Telephone Industry: The telephone industry processes data loaded with large amounts of historical data both online and offline, and these data need to be merged and integrated. In addition to these operations, DW are needed to analyze fixed assets, analyze customer's calling patterns to the sales representatives regarding the advertising campaigns and track customer queries.

These were the few applications of data warehouse, which was briefly discussed, there are many more applications as such (e.g., In hospitality industry, Finance industry, Consumer Goods Industry, Manufacturing and distribution Industry, Transportation Industry.).

11. BENEFITS & DRAWBACKS OF DATA WAREHOUSE

A success implementation of a data warehouse can convey out many advantages to an organization such as:

- The data warehouse allows end users to easily access various data in it, it allows them to make informed decisions on the basics of key aspects.
- The decision-makers should not only be reliable on the limited data because they have access to information from multiple sources on a same platform. Effective decision can be made in less amount of time, it eliminates the need for organizations to waste time collecting data from various sources. In addition, data warehouses can be applied in the management of finance, processing businesses, segmentation of the market, can be included in inventory.
- Data warehouses can work together to increase the operational value of business applications, mainly in customer relationship management.
- Data warehouses can help users analyze different time periods, and usually store large amounts of historical data to make predictions about the future. The data warehouse supports the integration of multiple data sources, thereby reducing the load on the production system.
- The data warehouse ensures backup to the resources by enabling complete backup, a partial backup, cold backup, hot backup, and online backup.
- The data warehouse transforms data from multiple sources into a consistent format. Because the entire company's data is standardized, each department can provide consistent results to get more accurate data and make the right decisions.
- The data warehouse has a wide range of historical data storage libraries, which can display different trends and condition analysis in different time periods, to provide more accurate forecasts and results.

Even with countless advantages, implementing a data warehouse model do have some disadvantages:

- Datawarehouse is not an absolute option for *unstructured data*, and it requires large amount of data.
- The *initial cost of building* a warehouse is huge, and small businesses cannot afford the cost of building a warehouse, so they can only rely on public warehouses or rented warehouse to store products. So, we can say that high installation costs are one of the biggest obstacles for companies that have their own warehouses to store their products.
- Although the main objective of the data warehouse is to concentrate on the data, which is available at the same place for easy analysis and access, because different departments are unwilling to share their personal data in a central repository**,** it sometimes causes problems for each department, which also means security and increases *Ownership* issues in certain sectors. In this case, the company needs to ensure that data analysis is sent to trusted people in the company.
- In data warehouse, it is unable to capture the required amount of data which might be required for evaluation by a particular organization. The data may not be integrated into the warehouse which leads to *loss of information*.
- There is the *low flexibility* of quotation because it is difficult for the company to meet certain requirements since wages and salary costs are also rising.
- Sometimes, the internal sources that power the data warehouse continue to consolidate issues that have not been noticed for years. For instance, when entering customer data, some value fields use null

values. Even if the data is available, the customer data may be incomplete in the future.

12. FUTURE OF DATA WAREHOUSE

[22] "Data warehouses are poised to play a leading role in next-generation initiatives, from Artificial Intelligence to Machine Learning to the Internet of Things."

—Joe McKendrick

The future of the data warehouse is going to be cloud-based. To have a good performance, security, flexibility, and ease of use, organization are shifting to Cloud data warehousing. Many companies have used hybrid, public and private clouds to start their journey from data to decision making. Few of the cloud databases are: *Snowflake, BigQuery, RedShift, Azure SQL Data Warehouse*. With the advantages of hybrid and cloud platform, the next generation data warehouse tends to be more beneficial in 3 main aspects: *Storage, computing infrastructure and services*.

The data warehouse also assists AI and machine learning to obtain results. The data warehouse will not only be the foundation for AI data sets, but AI will also help to improve the operation and functions of the data warehouse. The next generation data warehouse market would help in the infrastructure that would support *Big Data.*

As a system, a concept, and a method of providing information about customers, markets, and companies, the data warehouse will not disappear anytime soon. Data warehouses are becoming a progressively important part of the digital world.

13. CONCLUSION

In this paper, we have seen what Data Warehouse is and some salient features and concepts of Data Warehouse. Data Warehouse is the paradigm of future as most of the business organizations use Data Warehouse for decision making process. Data Warehouse is advantageous for a fact that, all the users who want to access the data can get it from a single place rather than having to go to various operational applications directly. Thus, we have concentrated on basic concepts related to data warehouse to help the beginners to get a fair idea about the same.

REFERENCES

[1] W. H. Inmon: "Building the data warehouse, Fourth Edition" Published by Wiley Publishing, Inc.

[2] Rajan Tiwari, "Data Warehousing", Jharkhand Rai University. http://www.jru.edu.in/wp-content/uploads/moocs/e-books/computer-science-and-IT/DATA_WAREHOUSING.pdf

[3] Surajit Chaudhuri, Umeshwar Dayal: "An Overview of Data Warehousing and OLAP Technology", (Appears in ACM Sigmod Record, March 1997) https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/sigrecord.pdf

[4] Jake Frankenfield, "Data Warehousing", Investopedia. https://www.investopedia.com/terms/d/data-warehousing.asp

[5] Ali Radhi Al Essa, Bach, Christian, "Data Mining and Warehousing". ASEE 2014 Zone I Conference, April 3-5, 2014, University of Bridgeport, Bridgeport, CT, USA.

[6] Matteo Golfarelli, Stefano Rizzi, University of Bologna, Italy: "Data Warehouse Design Modern principles and Methodologies". book-data-warehouse-design-golfarelli-_-rizzi.pdf

[7] The Data Administration Newsletter, "Components of a Data Warehouse". https://tdan.com/components-of-a-data-warehouse/4213

[8] Mehmed Kantardzic, University of Louisville: "Data Mining Concepts, Models, Methods, and Algorithms Second Edition". Data Mining_ Concepts, Models, Methods, and Algorithms (2nd ed.) [Kantardzic 2011-08-16].pdf

[9] Talend, "What is a Data Warehouse and why does it matter to your business?". https://www.talend.com/resources/what-is-data-warehouse/

[10] Sunil Samtani, Mukesh Mohania, Vijay Kumar, Yahiko Kambayashi: "Recent Advances and Research Problems in Data Warehousing".

[11] Oracle: "Data Warehousing and Business Intelligence" Oracle Database Online Documentation 12cRelease 1 (12.1). https://docs.oracle.com/database/121/DWHSG/concept.htm#DWHSG-GUID-452FBA23-6976-4590-AA41-1369647AD14D

[12] A G P Kujur1, Ajay Oraon2, "A Data Warehouse Design for Institutions Information System." International Journal of Science and Research (IJSR), Volume 5 Issue 11, November 2016.

[13] Irfan Ahmed, "What is Data Warehousing? Concepts, Features, and Examples". https://www.astera.com/type/blog/what-is-data-warehousing/

[14] Sumit Thakur, "12 Applications of Data Warehouse", June 14 ,2016. https://whatisdbms.com/12-applications-of-data-warehouse/

[15] Rahul Dubey, "Benefits of Data Warehouse",Educba. https://www.educba.com/benefits-of-data-warehouse/

[16] Data Warehouse information center, "Benefits of a Data Warehouse". https://datawarehouseinfo.com/data-warehouse/benefits-of-a-data-warehouse/

[17] Tauqeer Khurram, "Data Warehouse benefits and its challenges", Tech Funnel. https://www.techfunnel.com/information-technology/benefits-of-data-warehouse/

[18] Sumit Thakur, "9 Disadvantages and limitations of data warehouse."

[19] Prakash Tiwari, "What is Data Warehousing? How it works, Types, and General Stages", DataChannel, April 7,2021. https://datachannel.co/blogs/introduction-to-data-warehousing/

[20] Dave Wells, "Future of the data warehouse",Eckerson Group, October 10, 2017. https://www.eckerson.com/articles/the-future-of-the-data-warehouse

[21] Amber Lee Dennis, "The Data Warehouse: From the Past to the Present", Dataversity, January 4,2017. https://www.dataversity.net/data-warehouse-past-present/

[22] Actian, Denodo, SlamData, Wipro, "The Future of Data Warehousing." Rethinking-the-Future-of-Data-Warehousing.pdf

[23] Guru99, "What is Data Warehouse? Types, Definition & Example." https://www.guru99.com/data-warehousing.html#4

[24] Javatpoint, "Data Warehouse Architecture." https://www.javatpoint.com/data-warehouse-architecture