Comparative Analysis of Traditional vs Cloud Data Warehouse

Robbia Gulnar

Government College University, Faisalabad, Pakistan

Correspondence:
Robbia Gulnar: robiagulnar@gmail.com

Article Link: https://www.brainetwork.org/index.php/jcai/article/view/10

DOI: https://doi.org/10.69591/jcai.v1i1.10

Citation: Gulnar, R. (2023). Comparative Analysis of Traditional vs Cloud data warehouse, *Journal of Computing and Artificial Intelligence, 1*(1), 1-18

Conflict of Interest: Authors declared no Conflict of Interest

Acknowledgment: No administrative and technical support was taken for this research

Article History

Submitted: Mar 11, 2023
Last Revised: Apr 20, 2023
Accepted: May 15, 2023

Volume 1, Issue 1, 2023

Funding
No

Copyright
The Authors

Licensing

licensed under a Creative Commons Attribution 4.0 International License.

An official Publication of
Beyond Research Advancement & Innovation Network, Islamabad, Pakistan

Comparative Analysis of Traditional Vs Cloud Data Warehouse

Robbia Gulnar*

Government College University, Faisalabad, Pakistan

Abstract

In recent years due to the emerging of technologies, data amount is getting increase day by day. The mode and methods of data handling are getting upgraded. Prediction analysis seems to be very difficult but it generates interesting results. Various sectors like health, transport, education and economics are providing large amount of data. Recent developments in web technologies made it possible for operator and researcher to analyze and forecast that huge amount of data. The domain of BI (business intelligence) is core technology which help its users to extract meaningful information for decision making regarding any business. Data warehouse offers an insight view of business processes using previously provided data. Still, this traditional data warehouse system is not suitable for data analysis as it doesn't meet the requirements of evolving industries of this age. It cannot provide the up or down scale. Plus, it cannot handle the growing number of users. A new form of data warehouse has developed having better design and features, known as cloud data warehouse. This system has evolved in recent years, which directly affects the business and application domain as well. It has evolved in such a way to control large-scale data and provide up and down scale of business anytime. Moreover, it can handle a large number of users. In this review paper, comparison of traditional and cloud data warehouse is provided.

Keywords: business intelligence, warehouse, large data, geographic

Introduction

Nowadays, data warehousing system is the main element or component among the all areas of computing industry. For business personnel, it has provided the important enhancement to their business processes, while information system personnel consider it best quality method to eliminate the obstacle for providing enterprise records for executives and customers.

* Corresponding Author: robiagulnar@gmail.com

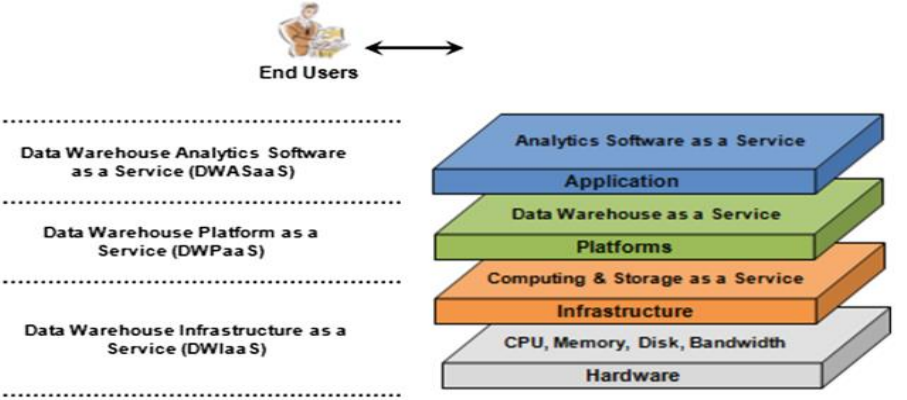
DWH is a vital feature of business intelligence as it used for reporting and provide better data analysis. DWHs provides a repository to save historical and sometimes current data from several sources. The current data warehouse has extra profitable and fascinating features than traditional data warehouse [1]. Bill Inmon said "A data warehouse is basically non-volatile, integrated, subject oriented and time varying collection of data, that is used in administrative decision making" [2]. DWH support online analytical processing (OLAP), which uses difficult questions to analyze aggregated data from OLTP (Online transaction processing). Traditional DWH is full supported to SQL but it is not scalable. It takes a lot of time to organize, optimize and operate this system. That is why, nowadays industries are moving to cloud data warehouse (CDWH). CDWH is completely different from traditional DWH. CDWH can provide data from various time-zones and geographic-zone to anywhere in the world.

Modern businesses of this era are born on the cloud. Their systems are built with cloud- architecture, and their data teams work with cloud data systems rather than on-premises servers. DWH can also be taken as integrated software section of the cloud which provide support to timely and accurate response to queries that helps in OLAP tasks. A cloud setup is designed or arranged in such a way so that it can take advantage of maximum number of user applications so that it becomes easy for a DBA to provide services for that environment [3]. So, the cloud data warehouse with time must take and supply new technology, techniques and automate tons of those manual tasks. Extract, Transform and Load (ETL) method is generally applied to deal with a data warehouse. ETL is basically a back-end process that collects data from multiple resources, transform it according to business needs then stack this data to a data warehouse. Extraction is a process getting data from different resources. Later on, this gathered data is transformed in a well-managed form so that it can be used later on for useful purposes is transformation. ETL processes use almost 70% of computing resources of the whole data warehousing [4]. Business Intelligence (BI) that is a main technology of this age used for decision making. All that required information is pulled from the data warehouse. In fact, the transforming of the raw data (collected from multiple resources) into useful information (used for decision making) is called Business Intelligence.

Cloud based Data warehouse architecture is an architecture that meet the needs to sort big data and used in business Intelligence BI applications that is built upon the low-cost large number of computer Nodes. It is an emerging and different way to sort and get the intuition from big data. It forms a unique type of service focused on DNS (Decision support systems). Cloud bases architecture for data warehouse consists on different big data storage, analytic techniques into services, management of data and the concerns from users is just to use the services what they want, when and where they want to use and store data, search data and analyze the data and visualize the data [5]. End users are the users like administrators, developers and customers that uses the Cloud data warehouse. Service end user can be categorized as normal users and professional users. To Use decision support systems which is service oriented is one of the important improvements for various organizations in anticipations of becoming more agile. Comparable to the cloud architecture, the architecture of cloud data warehouse is majorly divided into three layers. These various layers classified practically from analytics software to infrastructure as planned in diagram. Cloud architecture of data warehouse provides numerous type of services, platform level, infrastructure level and analytics software as service level [6]. By using various layered services integration can be simply made with other system.

Figure 1

Layers classified from analytics software to infrastructure.



At the lowest layer, all cloud data warehouse architectures offer fundamental storage and computing capabilities for handling big data. The

infrastructure of data warehouse as a service combines two key components: first, infrastructure as a service in cloud computing, and second, computing and storage as services. This combination enables the storage and processing of big data, allowing for its analysis and evaluation.

Big data poses significant challenges to traditional infrastructures due to its unique functions and features. To effectively manage big data, a specialized infrastructure tailored to its needs is required. Such an infrastructure must deliver optimized performance to enable swift data access and processing, meeting users' requirements in real-time [7].

A critical aspect of designing data warehouse infrastructure for services is the provision for handling various data types, including both existing and emerging new types [8].

Distributed file systems and NoSQL Database Management Systems are part of this lower layer to maintain unstructured data and perform parallel processing on Big Data. Data warehouse infrastructure as a service is defined through two sub-layers, namely the

1. Hardware resources layer
2. Infrastructure resources layer.

The resources of hardware layer deals with the abstractions of low-level physical devices like a a simple machine or server and the virtualize management tools used to organize great number of virtual machines on physical devices/hardware. Every virtual machine will be used as a computing, storage node and processing. There is major two kinds of node present, the slave and the master node [9].

The layer of infrastructure resources is an application that are deployed in VM to provide Hadoop Services that are called as HDaaS. It is Further divide in two services computing as a service that is basically the Map-Reduce paradigm that is used to distributed processing and Storage for service [10].

In the architecture model of a cloud data warehouse, the upper layer represents the data warehouse platform as a service. Here, the service provider establishes infrastructure for data management and creates an execution environment for data applications and scripts. Typically, this layer consists of NoSQL data stores and distributed data accessed through

query languages, forming the backbone of the cloud data warehouse architecture [11].

The platform layer of the cloud data warehouse architecture primarily consists of NoSQL data stores and distributed data accessible through structured query languages. This layer hosts the logical model of the data warehouse stored in NoSQL databases, empowering users to develop logical applications on extensive datasets using the data warehouse as a service. Integration of diverse data sources aims to offer end-users a unified view of these sources, a crucial aspect explored in business intelligence studies focusing on constructing efficient data warehouses as a service with the 4V features of big data.

The data warehouse as a service encompasses various main services, including Data Warehouse Design, Metadata Management, ETL (Extract, Transform, Load) Processes, Query Management, and Data Delivery.

In the final layer of the cloud data warehouse architecture, the analytics layer of the data warehouse as a service is dedicated to big data analytics services. This involves analyzing various forms of big data to uncover hidden patterns, unexpected relationships, and valuable insights. The complexity of algorithms designed for big data analytics often surpasses the capabilities of most organizations' IT departments, leading businesses to adopt analytics software for data warehouse as a service to harness the potential of their unstructured data [12].

By leveraging large volumes of organized and unstructured data, the cloud data warehouse architecture facilitates real-time and intelligent results for end-users. Conversely, Analytics Software as a Service (ASaaS) offerings enable users to interact with diverse analytics platforms at a higher level of abstraction, typically involving the analysis and execution of scripts crafted by programmers or data experts to create dashboards, reports, and visualizations. These services encapsulate a variety of big data analytics methods, including Visualization, Mining, OLAP (Online Analytical Processing), Dashboarding, and Reporting [13].

This information can be used for better strategic and decision-making jobs. In this review paper we have discussed the key differences between traditional and cloud data warehouse as shown in Table 1.

Literature Review

The advancements in data warehousing technology offer solutions across various industries by integrating real-time data with historical backups. Among the critical aspects of data warehousing is cloud computing, facilitated by cloud-based architectural frameworks and a multitude of accessible data sources. The emergence of web-based inquiries and the proliferation of distributed computing components, which encompass applications, software frameworks, and hardware infrastructure delivered as services over the internet, are noteworthy trends [14].

Server virtualization, utility computing, and distributed computing have paved the way for software as a service offering. Business intelligence, a distinct segment from traditional business realms, relies on data analytics sourced from multiple sources, predominantly historical data. Businesses leverage historical data to anticipate trends and patterns before they unfold [15].

Cloud data warehousing introduces new techniques and software distinct from the traditional practice of storing data in historical folders within conventional data warehouses. D.J Adabi and his colleagues consider the opportunities, restrictions and drawbacks of transportation of data into the cloud. They concluded that present database systems are unable to move data into cloud. Although DSS (Decision support system) can be used which can take the advantages of cloud. [16]

M. Brantner et al. Tried to build a database system on Amazon's S3, with the considerations of OLTP. They mainly focus on whether a simplest form of database can be build on S3. [11] D. Lomet and his fellows research propose a system for deployment of OLTP in cloud. This research in mainly based on the idea if database system needs to be split into data components and transactional components in order to make the database more flexible to operate in the cloud. [17]

S. Des and his colleagues developed Elastic Transactional relational database system. They focus on transactional database. The diverse nature of architecture makes it more flexible to accept parallel workloads [13].

Aboulnaga A. has faced some major challenges while deployment of a data warehouse onto cloud. This challenges are related to positioning of VM's across physical machines, handling of active workloads and segregation of resources against VM. He also provides the solution of these problems [18]. N. W. Paton and his fellows focus on a reliable method to transfer adaptive workload accomplishment in cloud [19].

Results & Discussion

Comparative Analysis

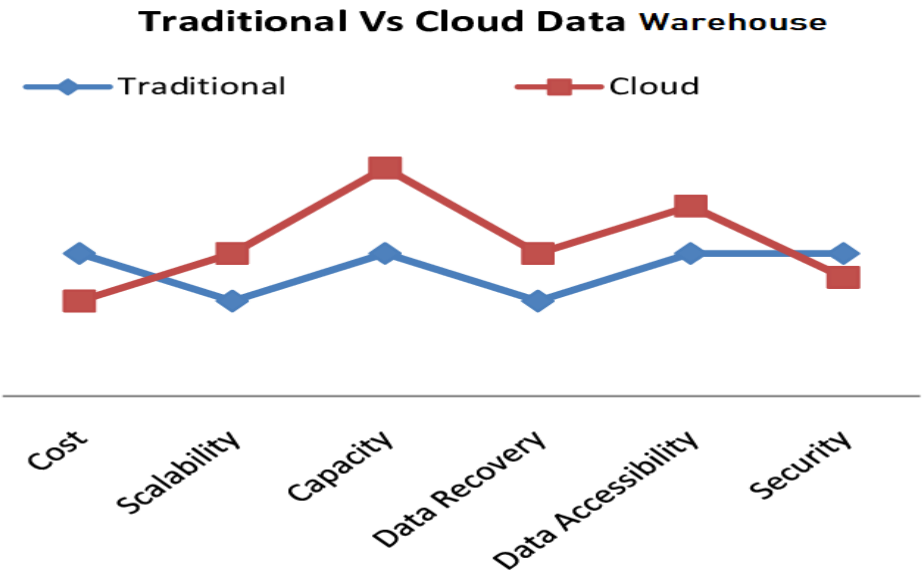
Cloud data warehousing offers storage services, contrasting with traditional data warehousing, which entails significant time for software, hardware, and infrastructure configuration, optimization, and management. Cloud data warehousing capitalizes on a large user and application base, making scaling easier and more efficient compared to the tedious and resource-intensive scaling process of traditional data warehousing [20]. With cloud solutions, data scaling occurs instantly and seamlessly, accommodating business fluctuations, whereas traditional systems struggle to provide immediate scaling adjustments [21].

Obtaining a cloud data warehouse is relatively straightforward and swift, unlike the time-consuming setup required for traditional data warehousing. Cloud data warehouses excel in analytics by utilizing columnar storage and massively parallel processing (MPP), which significantly boost the efficiency of intricate queries. This optimization empowers business users to redirect their attention from system management to data analysis, diverging from the traditional data warehousing model, which prioritizes data management [22].

Cloud data warehousing offers cost savings by eliminating hefty upfront expenses associated with on-premises systems. With no need for server rooms, hardware procurement, IT staffing, or operational expenses, cloud solutions streamline the complexity and cost of maintaining infrastructure, enabling users to concentrate on deriving value from their data rather than managing software and hardware [23]. Figure 1 illustrates the significant advantages of cloud data warehousing over traditional data warehousing.

Figure 2

Cloud vs Traditional data warehouse [17]



CDW also delivers new techniques, technology and automates various manual duties whether that can be related to:

1. Managing encryption
2. Managing the data
3. Automatic determination of data distribution
4. Affordability
5. Volume and pace of data
6. Delivering automatic performance optimizations
7. Limitation of users
8. Deployment

The main differences in traditional DWH and cloud DWH are explained well in Table 1 given below:

Table 1

Differences in traditional DWH and cloud DWH

Traditional data warehouse	Cloud data warehouse
It is a huge task for planning while using traditional data warehouse.	There is no need to plan your data warehouse while using cloud.
A traditional DWH is obstinate and it may cause the extra provisioning and paying.	It is flexible and routinely grows when it needs increases. It can also cause the cost to reduce.
It directly affects the queries when the data raises.	It does not affect the queries at all when the data raises.
It does not shrink when it is under-utilized and hence increase cost.	It automatically gets shrink when it is under-utilized to save costs.
The standards for choosing a traditional DWH is large data volume but not growing figure of users.	The standards for choosing a cloud DWH is dynamic and flexible data volume and growing figure of users
It cannot scale up or down data at any time. It takes time to organize software, hardware and infrastructure. Plus, it cannot handle the growing number of customers.	It can cause the data to scale up and scale down at any time. It will not take time to organize software, hardware and infrastructure. It can also handle growing number of customers.
It is expensive and difficult to scale when the data increase.	A Data warehouse constructed for the cloud is easy and low-cost to size and scale.
It cannot handle varied data forms. It can only deal with structured data.	It is capable to handle varied data forms i.e., semi structured and structured.
It forces you to purchase storage.	It allows you to resize compute collections (only pay for your need and when you need it).

Hence, overall, we analyzed that cloud data warehouse has more potential over traditional warehouse because of many above discussed reasons [24].

In cloud, data warehouse upgrading and maintenance is easy and it is comparable to traditional data warehouse because there is no need to clean files and maintain any kind of index etc. The cloud warehouse has no cost of complexity and it has cheap storage [25]. It provides independent scaling to compute and adopted new ways to adopt traditional warehouse system. Consequently, cloud data warehouse has replaced the traditional data warehouse because it supports business analytics smoothly [26]. In the next section some of the characteristics of cloud and requirements for a DW in cloud environment and its benefits are discussed.

Cloud's Characteristics

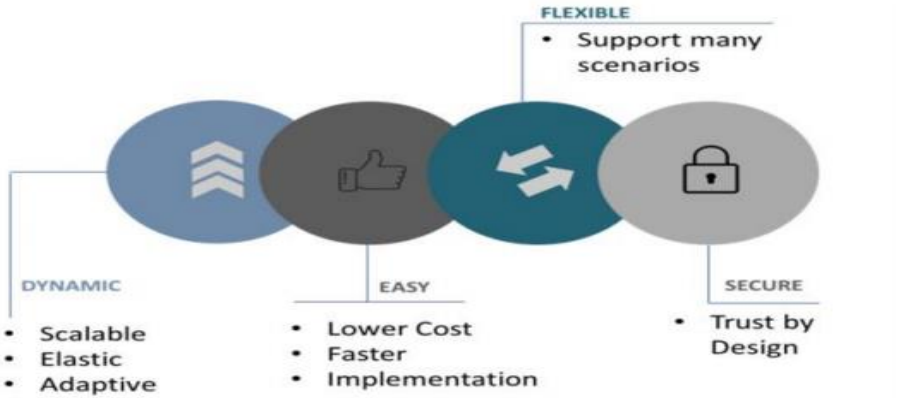
There are several reasons why cloud computing is better than traditional techniques of computing. Some reasons are given below:

1. Infinite scalability: It provides the limitless volume for computing.
2. Deployment speed: The on-time delivery of complete services will minimize the deployment duration.
3. Elasticity: The common payment method is used for this cloud computing is pay-per-use. It means you need to pay for the services you actually use. It saves cost of start-up and over-p provisioning.
4. Reliability: A cloud computing provides high reliability by having several data centers.

In figure 2 some of the characteristics of a cloud are shown.

Figure 3

Cloud Characteristics



Requirements for Data Warehousing Systems (Cloud-Based)

Data warehousing systems must meet specific conditions to adapt and thrive in a cloud environment. Although there are numerous technical criteria for data warehousing systems, we'll focus on the most critical ones directly pertinent to transitioning to the cloud [27].

1. **Performance**: Data transferred swiftly between permanent cloud storage and computing nodes, or among computing nodes themselves. There are occasions where large volumes of data, even up to terabytes in size, need to be transferred as new nodes are engaged, such as to accommodate evolving workloads or during query processing.
2. **Flexibility:** Whenever the number of users or workloads increase or decrease the data warehousing systems should be able to automatically balance up and down in order to achieve flexibility or elasticity.
3. **Multi-latency:** Data warehouse networks in the cloud should be able to handle multiple users while using the same storage server. In the cloud systems, data storage systems the database schemes should be strictly isolated.
4. **Privacy:** Data warehouse systems in the cloud must encrypt data locally to ensure privacy. A useful addition would be the possibility to work on encrypted data.
5. **Security:** Data warehouse should be secure and safe so that legitimate personnel will be able to access cloud data warehouse, lack of security can lead to unauthorized access and breach of Confidentiality of data.
6. **Monitoring:** Managers must have monitoring capabilities in order to evaluate the processes of the systems. Potential bottlenecks must also be detected and cancelled if necessary (e.g. 'killer-queries'). [19, 20]

Benefits of Cloud Data Warehousing

1. Cost reduced due to sharable resources i.e. server and network
2. Improve storage speed and performance as data analyst try to access application database it slows down the performance because analytical queries executed to analyze number of rows at a time

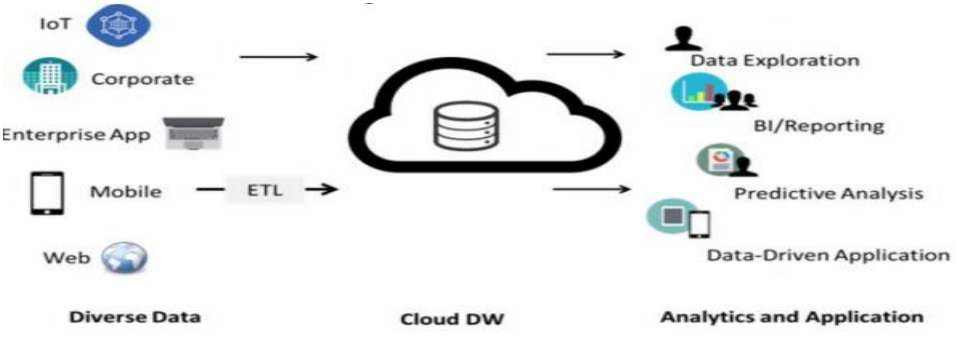
3. Data warehouse system must be able to encrypt the data to ensure privacy of data.
4. Cloud should be able to load the data rapidly which permits the vendor to work with enormous datasets within the shorter time constraint specified by the clients, furthermore servicing additional customers at a given time.
5. Optimization performed manually or automatically by using database management system, it enables SaaS BI vendors to enhance the classification of stored data. Execution of queries in optimal way
6. It enables quick or fast analytical reporting and execution of queries become faster with the help of cloud infrastructure.

By using cloud's services, we are able to get that knowledge that our organization might not have. There are many benefits like cost, operational and managerial benefits while using cloud's services provided by cloud suppliers rather than doing it by self [28].

Cloud data warehouse is also used scale up the systems by many organizations when needed. CWD allows automatic scaling for analytics, storage or resources accordingly. In contrast to traditional data warehouse CWD services can be scaled easily up and down according to needs of business. Cloud services can easily assign different tasks to even thousands of CPU's concurrently. Google cloud, IBM cloud, oracle, SAP and many other such cloud service providers is catching a reasonable amount of recognition and features in these days, but on the other hand, still the traditional DW is being used by many corporations for analysis of historical or recent data. There's no need to worry about maintenance and cost of DW while using CWD because maintenance becomes very easy and effective the cost also reduces in cloud services [21].

A cloud DW also make it easy to clone our precious data. Cloning is basically a process to make a snapshots and time data sets and this is done even without duplicating the storage, by using the consistent model of data. In Cloning we can also create different copies of data that can be used in future for testing different predictive models simultaneously. And most important of all, there will be no interference in the created models. A generic model of a cloud DW is shown Figure-2.

Figure 4
Cloud Data Warehouse



A cloud DW provides concurrency, better performance and more and implicitly while doing data analytics tasks. Importance of a cloud for data warehousing is generally linked with following to three main factors:

A) Enhanced agility

B) Effective cost control

C) Co-location

A. Enhanced agility

Movement of data becomes quick and easy while using cloud services.

B. Effective cost control

It becomes much easier to manage and reduce costs of predictive analysis tasks while using cloud services.

C. Co-location for faster loading: Data loading becomes even more speedier due to Co-Location due to which the consumers get access to data timelier. That's why E-business clients are now trying to move their data to cloud. [22]

Data Warehouse on Cloud Challenges

Potential disadvantages and bottleneck to deploy Data warehouse system in the cloud, there are few challenges which are discussed below:

1. Communication: when dealing with huge data (i.e. terabyte or petabyte) there might be risk of bottleneck. Data transferred over WAN link it is generally not faster than data transfer in local system. WAN link can lead to latency issue.
2. Performance: As many customers are active on the cloud at the same time, it may slow down the speed and performance might affect. specially I/Os to write to traditional disk
3. Loss of Control: when an organization use cloud as a service or platform, it loses some of control as it previously has over information assets.
4. High Cost: Amount for transfer of terabyte data to cloud Amazon charge 100$ to transfer 1TB to their cloud. Data transfer to cloud become more expensive process. It was claimed that shipping data on the disk to different location is a cheapest way of transferring data

Conclusion

The modern trend is being achieved by powerful and effective plan that are following modern data warehouse that has changed the big data and other business intelligences and moved them to cloud. The working data can be processed and maintained by speed and storage capability that has amazed. In cloud data warehouse upgrading and maintenance is easy and it is comparable to traditional data warehouse because there is no need to clean files and maintain any kind of index etc. The cloud warehouse has no cost of complexity and it has cheap storage. It provides independent scaling to compute and adopted new ways to adopt traditional warehouse system. Consequently, cloud data warehouse has replaced the traditional data warehouse because it supports business analytics smoothly.

Future Scope

Cloud characteristic i.e. reliability, elasticity, scalability, reduce cost, performance, queries execution time or deployment time etc. due to which data warehouse on cloud have latent potential.

References

[1] S. Chaudhuri and U. Dayal, "An overview of data warehousing and OLAP technology," *ACM SIGMOD Rec.*, vol. 26, no. 1, pp. 65–74, Mar. 1997, doi: 10.1145/248603.248616.

[2] W. H. Inmon, *Building the data warehouse*. John wiley & sons, 2005.

[3] S. H. A. El-Sappagh, A. M. A. Hendawi, and A. H. El Bastawissy, "A proposed model for data warehouse ETL processes," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 23, no. 2, pp. 91–104, 2011.

[4] H. Dabbèchi, A. Nabli, and L. Bouzguenda, "Towards Cloud-Based Data Warehouse as a Service for Big Data Analytics," in *Computational Collective Intelligence*, vol. 9876, N. T. Nguyen, L. Iliadis, Y. Manolopoulos, and B. Trawiński, Eds., in Lecture Notes in Computer Science, vol. 9876. , Cham: Springer International Publishing, 2016, pp. 180–189. doi: 10.1007/978-3-319-45246-3_17.

[5] U. Aftab and G. F. Siddiqui, "Big data augmentation with data warehouse: A survey," in *2018 IEEE International Conference on Big Data (Big Data)*, IEEE, 2018, pp. 2785–2794. Available: https://ieeexplore.ieee.org/abstract/document/8622206/

[6] S. Chaudhuri, U. Dayal, and V. Narasayya, "An overview of business intelligence technology," *Commun. ACM*, vol. 54, no. 8, pp. 88–98, Aug. 2011, doi: 10.1145/1978542.1978562.

[7] D. Agrawal, S. Das, and A. El Abbadi, "Big data and cloud computing: current state and future opportunities," in *Proceedings of the 14th International Conference on Extending Database Technology*, Uppsala Sweden: ACM, Mar. 2011, pp. 530–533. doi: 10.1145/1951365.1951432.

[8] L. Duan and L. Da Xu, "Business intelligence for enterprise systems: a survey," *IEEE Trans. Ind. Inform.*, vol. 8, no. 3, pp. 679–687, 2012.

[9] H. Chen, R. H. Chiang, and V. C. Storey, "Business intelligence and analytics: From big data to big impact," *MIS Q.*, pp. 1165–1188, 2012.

[10] A. Cuzzocrea, L. Bellatreche, and I.-Y. Song, "Data warehousing and OLAP over big data: current challenges and future research

directions," in *Proceedings of the sixteenth international workshop on Data warehousing and OLAP*, San Francisco California USA: ACM, Oct. 2013, pp. 67–70. doi: 10.1145/2513190.2517828.

[11] E. Abdelaziz and O. Mohamed, "Optimisation of the queries execution plan in cloud data warehouses," in *2015 5th World Congress on Information and Communication Technologies (WICT)*, IEEE, 2015, pp. 219–133. Available: https://ieeexplore.ieee.org/abstract/document/7489659/

[12] D. J. Abadi, "Data management in the cloud: Limitations and opportunities.," *IEEE Data Eng Bull*, vol. 32, no. 1, pp. 3–12, 2009.

[13] M. Brantner, D. Florescu, D. Graf, D. Kossmann, and T. Kraska, "Building a database on S3," in *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, Vancouver Canada: ACM, Jun. 2008, pp. 251–264. doi: 10.1145/1376616.1376645.

[14] D. Lomet, A. Fekete, G. Weikum, and M. Zwilling, "Unbundling Transaction Services in the Cloud." arXiv, Sep. 09, 2009.. Available: http://arxiv.org/abs/0909.1768

[15] S. Das, D. Agrawal, and A. El Abbadi, "ElasTraS: An elastic, scalable, and self-managing transactional database for the cloud," *ACM Trans. Database Syst.*, vol. 38, no. 1, pp. 1–45, Apr. 2013, doi: 10.1145/2445583.2445588.

[16] A. Aboulnaga, K. Salem, A. A. Soror, U. F. Minhas, P. Kokosielis, and S. Kamath, "Deploying database appliances in the cloud.," *IEEE Data Eng Bull*, vol. 32, no. 1, pp. 13–20, 2009.

[17] N. Paton, M. A. De Aragão, K. Lee, A. A. Fernandes, and R. Sakellariou, "Optimizing utility in cloud computing through autonomic workload execution," *Bull. Tech. Comm. Data Eng.*, vol. 32, no. 1, pp. 51–58, 2009.

[18] J. Widom, "Research problems in data warehousing," in *Proceedings of the fourth international conference on Information and knowledge management - CIKM '95*, Baltimore, Maryland, United States: ACM Press, 1995, pp. 25–30. doi: 10.1145/221270.221319.

[19] E. Guermazi, M. B. Ayed, and H. Ben-Abdallah, "Adaptive security for Cloud data warehouse as a service," in *2015 IEEE/ACIS 14th International Conference on Computer and Information Science*

(ICIS), IEEE, 2015, pp. 647–650. Available: https://ieeexplore.ieee.org/abstract/document/7166672/

[20] A. R. Nimje, "Data Analytics as a Service (DAaaS): An Arriving Technology in Cloud Computing," *Int. J. Emerg. Trend Eng. Basic Sci.*, vol. 2, no. 1, pp. 181–186, 2015.

[21] S. Kurunji, T. Ge, B. Liu, and C. X. Chen, "Communication cost optimization for cloud Data Warehouse queries," in *4th IEEE International Conference on Cloud Computing Technology and Science Proceedings*, IEEE, 2012, pp. 512–519. Available: https://ieeexplore.ieee.org/abstract/document/6427580/

[22] A. M. Van, H. L. Lv, V. L. Cheng, and F. V. Wang, "Design of cloud data warehouse and its application in smart grid," in *International Conference on Automatic Control and Artificial Intelligence (ACAI 2012)*, Xiamen, China: Institution of Engineering and Technology, 2012, pp. 849–852. doi: 10.1049/cp.2012.1108.

[23] A. Abid *et al.*, "A survey on search results diversification techniques," *Neural Comput. Appl.*, vol. 27, no. 5, pp. 1207–1229, Jul. 2016, doi: 10.1007/s00521-015-1945-5.

[24] M. G. Kahn *et al.*, "Migrating a research data warehouse to a public cloud: challenges and opportunities," *J. Am. Med. Inform. Assoc.*, vol. 29, no. 4, pp. 592–600, 2022.

[25] L. Dinesh and K. G. Devi, "An efficient hybrid optimization of ETL process in data warehouse of cloud architecture," *J. Cloud Comput.*, vol. 13, no. 1, p. 12, Jan. 2024, doi: 10.1186/s13677-023-00571-y.

[26] A. Nambiar and D. Mundra, "An overview of data warehouse and data lake in modern enterprise data management," *Big Data Cogn. Comput.*, vol. 6, no. 4, p. 132, 2022.

[27] S. Rohajawati, "Effectiveness of Digital Transformation in Data Warehouse Technology," *Indones. J. Comput. Sci.*, vol. 13, no. 1, 2024. Available: http://ijcs.stmikindonesia.ac.id/ijcs/index.php/ijcs/article/view/3759

[28] M. Patel and D. B. Patel, "Data Warehouse Modernization Using Document-Oriented ETL Framework for Real Time Analytics," in *Rising Threats in Expert Applications and Solutions*, vol. 434, V. S. Rathore, S. C. Sharma, J. M. R. S. Tavares, C. Moreira, and B. Surendiran, Eds., in Lecture Notes in Networks and Systems, vol.

434. , Singapore: Springer Nature Singapore, 2022, pp. 33–41. doi: 10.1007/978-981-19-1122-4_5.