14th Annual International Conference on Industrial Engineering and Operations Management Dubai United Arab Emirates (UAE), February 12-14, 2024

Publisher: IEOM Society International, USA DOI: 10.46254/AN14.20240541
Published: February 12, 2024

Into the Future with Cloud: A Comparison with On-premises Data Warehouse

Iman Noor and Saad Bin Tariq
Institute of Geographic Information Systems
National University of Sciences and Technology
Islamabad, Pakistan
imannoor.kr2016@gmail.com, saadbintariq01@gmail.com

Aisha Shabbir
NUST Institute of Civil Engineering, School of Civil and Environmental Engineering
National University of Sciences and Technology
Islamabad, 44000, Pakistan
aisha.shabbir@nice.nust.edu.pk

Mary Aksa
Department of Software Engineering
Foundation University
Islamabad,44000, Pakistan
Mary.aksa2015@gmail.com

Abstract

The need for data is growing at an extremely steep rate in the ever-digital realm, where terms like "big data" are becoming a thing of the past. All this development requires the use of modern and advanced data handling techniques, where users and researchers can analyze and predict vast amounts of data efficiently. Data warehouses are centralized repositories of data used for business intelligence activities such as analysis and reporting. In this paper, a comparative emphasis has been laid down on two different types of data warehouses, on-premises, and cloud data warehouses. The on-premises are known to be physically housed inside an organization's infrastructure. Cloud data warehouses are online-accessible repositories for data that is stored on cloud platforms. This paper provides a comparative analysis of both types in the context of deployment, scalability, flexibility, query management, cost analysis, access and integration, data security, data storage, data recovery, self-service capabilities and nonetheless, speed and performance. This article further highlights the evolution of data warehouses onto cloud and accentuates the growing demand for an efficient data warehouse, due to the amplification of volume, velocity, variety, value, and veracity of the incoming data in all realms. Furthermore, it provides an in-depth analysis of the advantages of the most suitable data warehouse and discusses the limitations of both.

Keywords

Cloud data warehouse, on-premises data warehouse, cloud computing, big data

1. Introduction

The concept of data warehousing dates to the 1970s, when various organizations, businesses, and enterprises started pondering the importance of handling, storing, and managing the growing amount of data. To address these arising problems, the father of data warehouse, also known as Bill Inmon, introduced the term “data warehouse” in the late

80s. Similarly, an IBM Systems Journal article 'An architecture for a business information system', published in 1988, concocted the term Business Data Warehouse. Such was the journey that opened gates to endless possibilities to handle the vastly escalating data.

As of now, Data Warehousing is known to be the most significant part of a business, an organization, or a company. The world of Business Intelligence has grasped the concept of data warehousing firmly, eventually using it for decision-making purposes, for managing the reports of the business world, and for data analytical needs. In the current times, where data is been sighted as the latest weapon, and the most precious resource, the worth of finding a suitable data warehouse solution is increasing rapidly (Vaisman and Zimányi 2014). Data Warehouse is known as a subject-oriented, integrated, time-variant, and non-volatile repository used for supporting the decision-making processes of an organization (Chen 2004) (Inmon 2005). Data Warehousing is stated as the notion of assembling and depositing big data from multiple sources in such a manner that they can be efficiently retrieved and analyzed. Data warehouse is stated to be a collection of databases, that are used for storing data, analyzing the data, and manipulating data as per the need. The analysts use this data to reproduce online analysis, reports, and grant help in the decision-making aspects of an organization (Giannakouris and Smihily 2014). Therefore, the data warehouse occupies the most important and centralized place in the business intelligence world. These data warehouses collect the data from various sources across various platforms and queries and henceforth, by the usage of different decision support queries they produce an analytic data report. (Abdelaziz and Mohamed 2015).

Traditional data warehouses are known to be unscalable, despite the support that a traditional Data Warehouse provides, the time required for optimization, configuration, and management exceeds the ever-growing income of data (Rehman and Mahmood 2018). The load on the large queries and the traffic clustering has eventually led to a decline in the performance observed in a data warehouse (Abdelaziz and Mohamed 2015). The resolution to this problem had been provided by the emergence of Cloud-based solutions. Factors such as cost-effectiveness, scalability, and the services provided by the Cloud Data Warehouse solutions have left the world in awe. Data warehouses have continued to unstoppably grow in this era of complex data generation, coping with the expanding demands of business intelligence. This new paradigm of cloud data warehouse stores the data remotely and provides solutions to the business world. (Abdelaziz and Mohamed 2015). A vast number of organizations have transformed their data warehouses from the on-premises approach towards hosting them on the cloud. And many are yet in the phase of transformation.

In this paper, we have provided a comparative analysis between on-premises and cloud data warehouses. The paper is structured as follows; the literation review is provided in section 2 of the paper. The comparison and the approaches used to decide between on-premises and cloud data warehouses have been laid down in section 3. Section 4 of the paper is the Discussion. Finally, the conclusion has been drafted in the last section of this paper.

2. Literature Review

Cloud data warehouses have been surrounding and evolving the world of business intelligence lately. Recent studies have been emphasizing how to promote your data warehouse from a traditional approach to a cloud-based data warehouse. In a world where big data has created hype, a much larger emphasis is being placed on the approaches on handling the velocity, variety, volume, and veracity attributed to big data. In previous research, big data has been classified by these four attributes, and further detail has been laid out into these characteristics of big data. Patently, conclusions have been laid down stating big data plays the most significant role in various applications (Yadav et al. 2019). Parallelly, the spotlight has been drawn to all possible methods of tackling and handling the extensive and immense amounts of data produced daily in the business environment. Amongst these, cloud data warehousing is making its way through to being the spotlight of the future of data warehousing.

However, at the moment, all the attention is drifting towards cloud computing, cloud-based data warehouses, and their solutions; nonetheless, recent studies have been laying the foundation for a comparison between on-premises data warehouses and cloud data warehouses (Golec et al. 2021). In the study by Rehman and Mahmood (2018), authors have compared the traditional data warehouse and cloud data warehouse on certain parameters, which include the possibility of scaling up or down their data warehouses, the flexibility of data warehouses in perspective of both approaches, cost analysis, the ability to handle diverse data types (semi-structured and structured), and others. It states that cloud data warehouses outnumbered traditional data warehouses based on the parameters in the comparative analysis. Furthermore, the research clarifies that cloud data warehouses are free from limitations such as maintenance of indexes, cleaning of files, updating metadata files, etc., making firm grounds for their conclusion that cloud data

warehouses are taking over traditional warehouses with respect to decision support and business analytics. A much similar comparison analysis has been provided by another study (Golec et al. 2021). The study brought attention to several advantages of cloud data warehouses, which include improved access and integration, improved speed and performance, low cost of ownership, leveraged cloud and elasticity, and others. This paper provides the advantages in the form of a figure and table. The paper deals with providing a cloud strategy; a cloud strategy is stated to be the transmutation from a traditional data warehouse onto a cloud data warehouse.

Although a huge amount of attention has been drawn to the comparison between traditional and modern warehouses, studies have shifted the focus to the advantages and challenges of cloud data warehouses, among which data security is highlighted as a major issue (Boyko and Shakhovska 2018). In the midst of the transition to cloud data warehouses, organizations and businesses are reluctant due to the security issues of personal data in the cloud (Guermazi et al. 2001). According to other research organizations, as healthcare bodies, government departments, and financial institutions deal with sensitive information, organizations are skeptical about whether to trust cloud solutions with this data or not (Shaikh and Meshram 2021) (Onyebuchi et al. 2022). In counterpoint, another study has stated that cloud data warehouses are undertaking certain sets of policies to ensure data security and convince a larger community of businesses to entrust cloud data warehouses (Vaishnav and Prasad 2021). To ensure the security of personal as well as nonpersonal information, the European Union has contributed towards creating a platform on which the countries of the European Union have come together to sign a joint pact on cloud computing. This platform aims to create a European cloud, in which the member countries would be assured of a regulated free flow of non-personal data, cybersecurity, and data protection in the cloud. As a result, research has stated that 41% of the European Union enterprises had converted to cloud computing in the year 2021. A high turnout was observed in Sweden (75%), Finland (75%), the Netherlands (65%), and Denmark (65%) (Giannakouris and Smihily 2014).

With all the talk revolving around the advantages of cloud-based data warehouses and the ease that these warehouses have provided to users and the business community, the challenges of cloud-based data warehouses are often overlooked. Whereas the study by Kurunji et al. (2012) has provided a thorough model for communication cost optimization. According to the author, in a cloud data warehouse, the relevant partitions are never guaranteed to be saved on the same physical machine, and during the execution of a query, inter-node communication requires to traverse through various nodes, thereby adding up to increased inter-node communication. The size of communication messages grows in proportion to the number of nodes and data amount, which leads to performance degradation. Thus, the study provides a PK-map-based storage structure alongside a query processing algorithm that has been proved to minimize the inter-node communication and decrease the workload of the joins upon query execution. Such studies have paved the path of transforming into a cloud data warehouse over the traditional. These works have indicated how the studies are being initiated to further move a step towards the cloud computing approach.

Moreover, the studies did not terminate at the optimization of communication costs; Abdelaziz and Mohamed (2015) laid out a plan for the optimization of queries in a cloud data warehouse. In the article, an approach to improving the cloud data warehouse's performance has been emphasized. The authors' work highlighted that storing in a cloud data warehouse is done via nodes, and optimizing inter-node communication is necessary to improve processing and response time. As a result, it proposes an approach to enhance the performance of cloud data warehouses based on a classification technique and an algorithm built on the MapReduce programming model, and this methodology has been claimed to minimize internode communication and hence query processing time. Similar to the optimization model stated before, research is being put into practice as to how cloud data warehouses can be moved a step further. All the advantages aside, these optimization approaches add a bonus point as to why the modification from on-premises data warehouses to cloud data warehouses should be done.

Moving further, studies have advanced from the advantages and optimization of cloud data warehouses to cloud data warehouse solutions. Abdelaziz and Mohamed (2015) have proposed up to ten cloud-based data warehouse solutions, which include Panoply, which tops the list, Teradata Integrated Data Warehouse, Yellowbrick Data, Oracle Autonomous Warehouse, IBM Db2 Warehouse, Google BigQuery (Ali et al. 2021), Amazon Redshift, a few others. Further study of these solutions has also been called attention to in the articles by Shaping Europe’s Digital Future. The authors Gupta et al. (2015) have presented Amazon Redshift as a cloud-based data warehousing solution. They in their work brought about a comparison of traditional data warehouses versus the Amazon Redshift. The arguments state that the traditional data warehouse systems are complex and costly, whereas Amazon Redshift turns out to be a cost-effective alternative. In terms of simplicity, flexibility, and scalability, the Amazon Redshift cloud-based data warehouse is said to be well-suited for a business environment and is very capable of handling big data.

In the world of big data, where the Amazon Redshift discussion is the hot topic, technologies such as Oracle Database Cloud Service are capturing the public's attention with their capabilities. Research has been performed on the Oracle Database Cloud Service, which states that only Oracle offers a complete platform as a service (PaaS) environment that allows a blended approach catering to the hardware and software together. Additionally, this advanced technology in the world of cloud computing offers users the option of using the Oracle Database Exadata Cloud Service for boosted performance (Oracle 2021).

To convince enterprises that the future lies in the hands of cloud computing and that there should be an adaptation of cloud data warehouses, a thorough comparison between on-premises and cloud data warehouses is necessary. The comparison can be made using an extended set of parameters. This paper's research attempts to bridge the comparability gap by delving further into both warehouses.

3. Comparative Analysis of On-Premises and Cloud-based Data Warehouse

The most prominent property that distinguishes the on-premises and the cloud data warehouse is where the data is situated, in the case of on-premises data warehouse, the data as well as the software, hardware, and applications required to handle the data are stored on-site, and henceforth, the management of this repository is solely handled by the organization itself. However, in the scenario of a cloud data warehouse, the data is hosted off premises and is handled by another organization. These data centers are run and managed by the organizations responsible for the cloud data warehouse, and yet the user of the business can uninterruptedly utilize the data in real-time, without any sort of hindrance. To further clarify the concept of a cloud data warehouse, it is stated to be a depository that stores the data remotely, however, the utilization of data is similar to that of any other data warehouse. Furthermore, it is asserted that the cloud data warehouse is available to the users in various formats, as per the need of the user. It provides a service to the users in which the applications virtually hosted on the cloud can be installed and run. This service is named as Infrastructure as a Service (IaaS) and amongst the notable provides Amazon Elastic Compute Cloud (EC2), and Cisco Vblock hold the spotlight (Bogdándy et al. 2020). Moreover, a cloud data warehouse provides Platform as a Service (PaaS), where the cloud provides platforms such as web servers, mail servers, databases etc. An example of this service is Microsoft Windows Azure (Sharma 2023). Lastly, the cloud provides Software as a Service (SaaS), where users are granted access to ready-to-use applications that can be used by multiple users simultaneously (Abdelaziz and Mohamed 2015). The comparison between the two data warehouses can be laid down on several parameters, the parameters that this study has utilized have been illustrated and described below in Figure 1. The result of the comparative analysis has been summarized in the form of a table, and it has been inserted at the end of this section under the heading of Table 1.

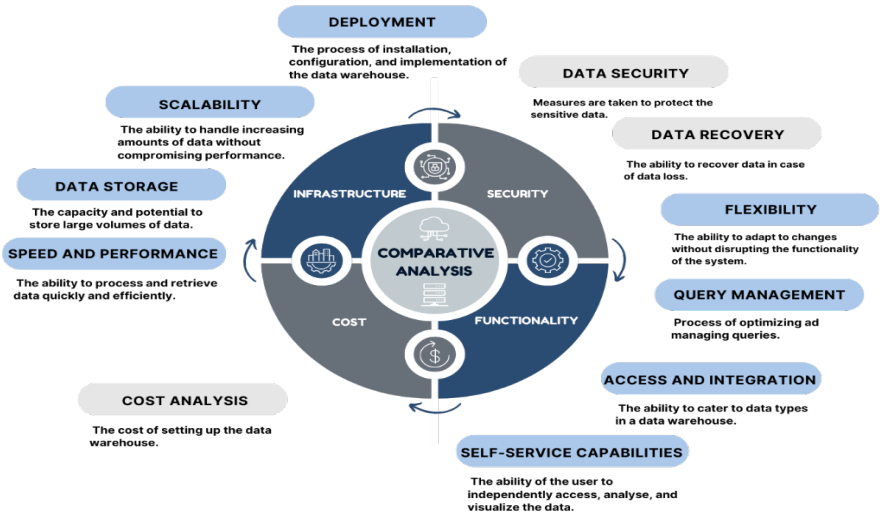

Figure 1. Comparative Analysis Parameters

3.1 Deployment
The deployment of an on-premises data warehouse is tedious as compared to the cloud data warehouse. As for the deployment of a traditional data warehouse the planning phase is solely handled by the business organization itself, thereof, it is stated to be an enormous task to be handled. Whereas, for the cloud data warehouse, the party responsible caters to and fulfils these requirements for an organization.

3.2 Scalability
It is notably known that a data warehouse set on-premises cannot be scaled up or down. Moreover, the on-premises data warehouse lacks the capabilities of catering for the growing number of users. Moreover, studies have stated that it often takes up to days for the configuration of hardware, software, and infrastructure of an on-premises data warehouse. In contrast, a cloud data warehouse is renowned for its functionality of scaling up and down, as per the need. And is capable of handling the increasing usage. It is also proclaimed that an on-premises data warehouse lacks the capacity to down-sizing the warehouse upon the reduction in usage, but a cloud data warehouse automatically downsizes upon underutilization. (Boyko and Shakhovska 2018).

3.3 Flexibility
A cloud data warehouse is known to be flexible to the growing amount of usage, the incoming data, and it is known to automatically alter as the data warehouse's requisites and requirements change. The feasibility of drastically adapting to the changes ultimately reduces the cost of implementing a cloud data warehouse. On the other hand, an on-premises data warehouse is not flexible to the requirements of the user, thus leading to being more expensive as compared to a cloud data warehouse (Rehman and Mahmood 2018).

3.4 Query Management
An on-premises data warehouse is known to affect the queries upon data expansion, whereas a cloud data warehouse does not affect the queries in the stated scenario. (Rehman and Mahmood 2018).

3.5 Cost Analysis
To own and set up an on-premises data warehouse is already more expensive as compared to a cloud data warehouse. As every component required to set up the data warehouse, such as the server, the hardware elements, the software, and storage devices, arouses the cost of set up. Moreover, the maintenance and regular check-up on the data warehouse further incurs the cost of implantation of an on-premises data warehouse. Whereas, the cloud data warehouse works on the principles of renting, and the services that are utilized by the users are only billed (Golec et al. 2021). Furthermore, the maintenance, the upgradation, and the regular checkups are handled by service providers which eventually cuts down the cost to a great extent.

3.6 Access and Integration
Access and Integration are widely known for the potential to support structured and unstructured data types. An on-premises traditional data warehouse is inadequate to handle the massive types of data available, it is only capable of handling, storing, and managing the structured data type. However, the cloud data warehouse is recognized to be capable of managing and storing diverse data types, such as unstructured, semi-structured, and structured (Golec et al. 2021).

3.7 Data Security
The users of cloud data warehouses have been skeptical about the implementation of these data warehouses in regard to the security provided. However, over time, cloud solutions have become more secure. It is stated that cloud data warehouses provide multi-factor authentication facilities which have paved the path for more secure communication and data handling (Vaishnav and Prasad 2021).

3.8 Data Storage
As it has already been stated that the on-premises data warehouse is not able to scale the data warehouse as per the need, whereas the cloud data warehouse is. Similarly, a cloud data warehouse in need of more data storage easily provides this service to its users, but for the traditional data warehouse, this upper limit of data storage is a certain figure, which cannot be crossed.

3.9 Data Recovery

In case of a disaster occurrence, the recovery of an on-premises data warehouse is nearly impossible. In the scenario of a cloud data warehouse, the data is stored on multiple nodes, which asynchronously back up the data, due to which data is available uninterruptedly and continuously. Moreover, the requirement of setting up a secondary data storage center of cloud data warehouses is smaller than compared to an on-premises data warehouse (Golec et al. 2021) (Table 1).

Table 2. Summary of comparisons of Data Warehouses

Parameters	On-Premises Data Warehouse	Cloud Data Warehouse
Deployment		✓
Scalability		✓
Flexibility		✓
Query Management		✓
Cost Analysis		✓
Access & Integration		✓
Data Security	✓	
Data Storage		✓
Data Recovery		✓
Self-Service Capabilities		✓
Speed & Performance		✓

4. Discussion

Cloud data warehouses are emerging and replacing traditional data warehouse methods numerously. Cloud data warehouses are being vitally utilized based upon the potential of their powerful computing mechanisms, the high-speed retrieval of data for complex queries, and brisk communication network. Moreover, the world of business intelligence is drastically adopting cloud-based solutions since they can handle data up to terabytes yet are a cost-effective approach.

Moreover, the most important reason that organizations are adopting the cloud strategy is that the service providers perform all the tedious tasks for the organization. Whether an organization wants to handle the purchase and sale orders, whether if an organization is willing to apply the sale upon the entire data, or if the data is to generate a report or likewise operations, all of these are ultimately handled by the service providers, hence, leaving a business company with least worried over such matters. The users have immensely consumed the services provided by cloud data warehouses, the ease of scalability, the proficiency of cloud solutions to provide flexibility in their approach, and above all, the affordability of cloud solutions compared against all parameters eventually led the organizations to transform their data warehouses onto the cloud.

The dire need to manage big data organizations has started the endless journey to the deployment of their warehouses on the cloud. The business world has been contemplating the restrictions that traditional data warehouses come about with. Above all, the fact that the on-premises data warehouses store the data in relational tables, however, the drastically evolving big data now demands a modern data warehouse that has the capacity of storing data from multiple data sources such as web pages, social media feeds, indexed, and sensors (Karkouda et al. 2019).

The implementations of cloud-based solutions have been advancing considerably, with the features greatly being utilized including data warehouses, sandbox development environments, data marts, and data back-ups, widely used in the scenario of disaster recovery (Oracle 2021).

Many service providers have started delivering cloud services whether they may be needed to host a data warehouse platform, to set up a cloud database, or to enhance the performance of the already hosted data warehouse on the cloud. Services such as Oracle or Amazon web services have been implemented by users due to their agile functionality.
The Amazon web services (AWS) provides its users with economical and cost-effective approaches to deploy and host their data warehouses on the cloud. AWS provides the world of business intelligence with cost-effective data lakes that are used to perform complex queries on petabytes of structured as well as semi-structured data such that the result can reproduce highly efficient reports and dashboards. Moreover, AWS provides access to computing SQL and analytical queries on various data types in the warehouse and data lakes without the requirement of relocating the data. Lastly, AWS is highly recognized for its efficient automation of setting and managing the data lakes for its users, thus reducing the workload of months into days.

Similarly, the Oracle cloud-based solution also provides users with inimitable features. Oracle solutions have been divided into three prime components, improved agility, cost-effective approach, and co-location features. As for agility, the data in Oracle cloud solutions can be moved efficiently. Furthermore, Oracle's cloud services provide data security at default, thus reducing the cost to a great extent. It has further been stated that the business world has been significantly transforming their applications onto Oracle Cloud Solutions, all due to the time available data.

There are further several reasons why the provision of Oracle data warehouse cloud services should be implemented. The unique functionality of Oracle to innovate on hardware and software has led to a much faster approach to deploying a data warehouse on Oracle than any other service provider as Amazon, Microsoft, etc. Secondly, the Oracle solution is known for its big data enablement. As the demand and recommendation of implementing a hybrid platform for the extraction of data is escalating rapidly. To cater for this, Oracle provides customers with cloud services that enable the creation of hybrid platforms, alongside it offers an inferior option of leveraging Hadoop as the engine for new data sources. Thirdly, Oracle is known for its capabilities of maintaining a data warehouse in the cloud which ultimately leads to ease of distributing the data and maintaining a geographically disbursed system (Oracle 2021).

5. Conclusion

Precisely, the world is in an endless race towards the best-suited approach for big data, business analytics, intelligence, and so on. Studies are consistently providing seamless approaches to handling the vast amount of data and are persuading the business community to shift from traditional methodologies to cloud data warehouses. Similarly, this paper has used an approach of thoroughly comparing the on-premises data warehouse and cloud data warehouse. The paper has reviewed the alternatives through a certain set of parameters and summarized the stances in a table. The modernization of the data warehouses has eventually paved the path to cloud strategy. Nonetheless, the case studies, the comparisons, and the optimization approaches tested on cloud data warehouses as compared to traditional data warehouses have certainly given an upper hand to cloud solutions. To further elucidate the dilemma between the data warehouse approaches services such as Amazon Redshift, Amazon DynamoDB, Cloudera, Snowflake, Teradata, Apache Hive, and Microsoft Azure are growing their customers rapidly. It may be stated that the on-premises data warehouse provides its users with a greater sense of data security, however, the services that the cloud has been offering nullify the capabilities of an on-premises data warehouse. The functionalities of the cloud such as pay-as-you-go service, scalability, flexibility, backup services and cost-effectiveness have led the world with a very clear motive for what has to be the future of data warehouses. Thus, business intelligence is meticulously delving into the capacity of cloud data warehouses and making the required shift towards it. Henceforth, it can be stated that between these two alternates, the cloud data warehouse is mounting the cliff of the data warehouse future with immensely densified services.

References

Abdelaziz, E. and Mohamed O., Optimisation of the queries execution plan in cloud data warehouses. *Proceedings of the 5th World Congress on Information and Communication Technologies* (WICT), pp. 219-133, 2015. IEEE.

Ali, M.H., Hosain, M.S. and Hossain, M.A., Big Data analysis using BigQuery on cloud computing platform. *Australian Journal of Engineering Innovation and Technology,* vol. 3, no. 1, pp. 1-9, 2021.

Bogdándy, B., Kovács, Á. and Tóth, Z., Case Study of an On-premise Data Warehouse Configuration. *Proceedings of the 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom),* pp. 000179-000184, September 23, 2020.

Boyko, N. and Shakhovska, N., Prospects for using cloud data warehouses in information systems. *Proceedings of the IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT)*, vol. 2, pp. 136-139, September 11, 2018.

Chen, X., E-Business Data Warehouse Design and Implementation. *Proceedings of the International Conference on Industrial Engineering and Operations Management,* 2004.

Cloud computing. (n.d.). Shaping Europe's Digital Future. Digital Strategy. Available: https://digital-strategy.ec.europa.eu/en/policies/cloud-computing, Accessed on August 19, 2023

Cloud Data Warehouse | Solutions Providers | EM360. (n.d.). Cloud Data Warehouse | Solutions Providers | EM360. Available: https://em360tech.com/data_management/tech-features-featuredtech-news/top-10-cloud-data-warehouse-solution-providers, Accessed on June 22, 2023.

Free Analytics Services - Free Datalakes - AWS. (n.d.). Amazon Web Services, Inc. *AWS Free Analytics Services.* Available: https://aws.amazon.com/free/analytics/, Accessed on August 20, 2023.

Giannakouris, K. and M. Smihily, Cloud computing statistics on the use by enterprises, *Eurostat Statistics Explained,* 2014.

Golec, D., Strugar, I. and Belak, D., The benefits of enterprise data warehouse implementation in cloud vs. on-premises. *Entrenova-Enterprise Research Innovation,* vol. 7, no. 1, pp. 66-74, 2021.

Guermazi, E., Ben Ayed, M. and Ben-Abdallah, H., A survey of data warehouse security. *Ingénierie des Systèmes d'Information*, vol. 19, no. 5, pp. 75-96, 2014.

Gupta, A., Agarwal, D., Tan, D., Kulesza, J., Pathak, R., Stefani, S. and Srinivasan, V., Amazon Redshift and the case for simpler data warehouses. *Proceedings of the 2015 ACM SIGMOD International Conference on- Management of Data*, pp. 1917-1923, May 27, 2015.

Inmon, W.H., Building the data warehouse. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, 2005.

Karkouda, K., Nabli, A. and Gargouri, F., SecuredDW: A Homomorphic Schema to Securely Hosting Data Warehouse in the Cloud. *Research in Computer Science*, vol. 148, no. 4, pp. 99-118, 2019.

Kurunji, S., Ge, T., Liu, B. and Chen, C.X., Communication cost optimization for cloud Data Warehouse queries. *Proceedings of the 4th IEEE International Conference on Cloud Computing Technology and Science*, pp. 512-519, December 3, 2012.

Onyebuchi, A., Matthew, U.O., Kazaure, J.S., Okafor, N.U., Okey, O.D., Okochi, P.I., Taiwo, J.F. and Matthew, A.O., Business demand for a cloud enterprise data warehouse in electronic healthcare computing: Issues and developments in e-healthcare cloud computing. *International Journal of Cloud Applications and Computing (IJCAC),* vol. 12, no. 1, pp. 1-22, 2022.

Oracle (2021), "Data Warehouse In The Cloud, Bringing Decades of Data Management Innovations to the Cloud", Available: https://www.oracle.com/a/ocom/docs/data-warehouse-in-the-cloud-brief-3157435.pdf, Accessed on June 15, 2023

Rehman, K.U.U., Ahmad, U. and Mahmood, S., A Comparative Analysis of Traditional and Cloud Data Warehouse. *Proceedings of Vawkum Transactions on Computer Science,* vol. 6, no. 1, pp. 34-40, 2018, https://doi.org/10.21015/vtcs.v15i1.487

Shaikh, A.H. and Meshram, B.B., Security issues in cloud computing. *Proceedings of Intelligent Computing and Networking: IC-ICN 2020,* pp. 63-77, Springer Singapore, 2021.

Sharma, P., Cloud Computing for Supply Chain Management and Warehouse Automation: A Case Study of Azure Cloud. *Proceedings of the International Conference on Industrial Engineering and Operations Management,* 2023.

Vaishnav, J. and Prasad, N.H., Security aspects in cloud tools and its analysis—a study. *Proceedings of Inventive Systems and Control: ICISC,* pp. 927-937, Springer Singapore, 2021.

Vaisman, A. and Zimányi, E., Data warehouse systems. *Data-Centric Systems and Applications*, 2014.

Yadav, P.K., Sharma, S. and Singh, A., Big Data and Cloud Computing: An Emerging Perspective and Future Trends. *Proceedings of the International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT),* vol. 1, pp. 1-4, September 27, 2019.

Biographies

Iman Noor, a dedicated and ambitious individual, is a final-year student pursuing a degree in Geoinformatics at the National University of Sciences and Technology (NUST). With her expertise and experience lying in the fields of spatial data analytics, machine learning, data warehousing, and web. Her hands-on experience in diverse projects showcases her adaptability and forward-thinking mindset.

Saad Bin Tariq is an aspiring Geoinformatics student with vast experience in web development and spatial data-analytics. He is currently in final year of bachelors engineering degree from National University of Sciences and Technology. He has worked on various projects in different organizations to polish his technical skills and cope up in the digital world covering subjects like Machine Learning, Data Mining and warehousing, Spatial System Development.

Aisha Shabbir is currently serving as an Assistant Professor at School of Civil & Environmental Engineering (SCEE), National University of Sciences and Technology (NUST), Islamabad. She holds a Ph.D. degree in Computer Science from University Technology Malaysia. She has completed her MS from Ghulam Ishaq Khan Institute (GIKI). She has presented her research work in international conferences and got the best paper award. She published her research work in many reputed journals. Her research interests include Data mining, machine learning, and Big Data processing techniques.

Mary Aksa is serving as Visiting Lecturer in Foundation University Islamabad, Pakistan. Mary holds a Masters of Science degree in Computer Science from COMSATS Institute of Information Technology, Wah and a Bachelors of Science degree in Bioinformatics from Mohammad Ali Jinnah University, Islamabad. She has taught courses in Artificial Intelligence, Analysis of Algorithm, Data Structures, Professional Practices in IT and Introduction to Bioinformatics passionately. Her research area is in Data Analytics.