**TECHNICAL REPORT WRITING FOR CA2 EXAMINATION**



**NAME**- SAPNADARSHI SINGHA

**STERAM-** COMPUTER SCIENCE AND BUSINESS SYSTEM

**ROLL NO** – 18731122014

**SUBJECT NAME**- DATABASE MANAGEMENT SYSTEMS

**SUBJECT CODE**- PCC CSBS 502

**YEAR**- THIRD

**SEMESTER**- FIFTH

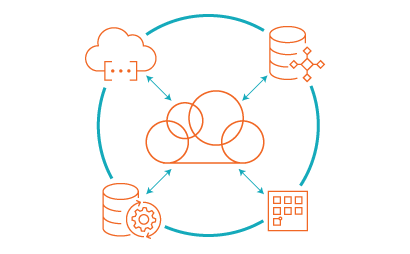
**TOPIC- RELATION BETWEEN CLOUD AND DATA WAREHOUSE**

**INTRODUCTION**

The rapid growth of data in recent years has transformed how organizations manage, store, and analyze information, resulting in a shift to cloud-based solutions. Among these, cloud data warehouses have emerged as an important component of modern data architecture, combining traditional data warehousing capabilities with the flexibility and scalability of the cloud. A data warehouse is a centralized repository that consolidates data from multiple sources, allowing organizations to run complex queries and analyses that are critical for business intelligence, reporting, and decision-making.

Cloud data warehouses like Amazon Redshift, Google Big Query, Snowflake, and Microsoft Azure Synapse Analytics utilize cloud resources to provide scalable, cost-effective, and high-performance data handling solutions. Cloud data warehouses differ from traditional on-premises data warehouses by operating on a pay-as-you-go model, which decreases costs and resource requirements, opposed to requiring extensive capital investment and continual maintenance. They allow companies to manage large quantities of structured and semi-structured data, conduct real-time analysis, and effortlessly connect with other cloud-based tools and services.

The merging of cloud services and data warehousing offers many benefits such as the ability to expand resources flexibly, improve teamwork, and utilize advanced analytics like machine learning and artificial intelligence. Nevertheless, it also brings in new factors to think about, like data security and governance.



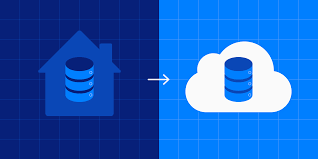
* **CONCEPT OF CLOUD AND DATA WAREHOUSING**

A data warehouse in the cloud is a system for storing and processing data on cloud infrastructure, enabling organizations to store, handle, and analyse extensive amounts of structured and semi-structured data.

Amazon Redshift, Google BigQuery, Snowflake, and Microsoft Azure Synapse Analytics are all instances of cloud data warehouses.

Old-fashioned data warehouses are located onsite, receiving data from relational databases, transactional systems, business applications, and other source systems. Nevertheless, they are usually created to collect a portion of information at once and save it according to strict structures, which prevents them from being used for immediate inquiries or live analysis. Companies are also required to buy their own hardware and software when using an on-premises data warehouse, which can be costly to expand and upkeep. In a conventional storage facility, there is usually less space for storage in comparison to processing power, leading to rapid data transformation and deletion to maintain free storage capacity.  
  
Today, data analytics now play a central role in all key business operations, such as generating revenue, minimizing costs, enhancing operations, and improving customer satisfaction. With the evolution and diversification of data, organizations require stronger data warehouse solutions and advanced analytic tools to store, manage, and analyse vast amounts of data throughout their entities.

These systems need to be able to grow in size, be dependable, have strong security features suitable for industries with regulations, and be adaptable to handle a range of data formats and big data scenarios. They must also offer flexible pricing and computing options, allowing you to pay only for what you use rather than estimating your needs. Most legacy data warehouses do not meet the requirements. Consequently, numerous businesses are shifting towards data warehouse solutions that are based in the cloud.



* **HOW DATA WAREHOUSING WORKS**

Similar to a classic data warehouse, cloud data warehouses gather, consolidate, and store information from both internal and external sources. Data is commonly moved from one system to another through a data pipeline. Information is taken from the source system, altered, and then placed into the data warehouse using the ETL process. Information can be transmitted straight to a main database and then transformed through ELT (extract, load, transform) methods. From that point, users are able to utilize various business intelligence (BI) tools for accessing, analyzing, and reporting data. Cloud data warehouses must also be able to accommodate streaming scenarios in order to take action on data immediately or almost immediately.  
  
Within a public cloud environment, cloud data warehouses provide storage, processing, integration, cleansing, and loading of structured and semi-structured data. You can also utilize them alongside a cloud data lake in gathering and storing unstructured data. With certain providers, you can merge your data warehouse and data lake to have one central location for managing your enterprise data.  
  
Various cloud providers may have different strategies in providing cloud data warehouse services. Some cloud data warehouses, for instance, might utilize a cluster-based design that resembles that of a traditional data warehouse. On the other hand, some choose to utilize a modern serverless structure, effectively reducing data management duties even more. Nevertheless, the majority of cloud data warehouses come equipped with integrated data storage and capacity management functionalities as well as automatic updates.

* **WHY USE CLOUD BASED DATA WAREHOUSING**

Cloud data warehouses have numerous advantages over traditional on-premises data warehouses, making them a popular option for modern data management. Here are the main benefits of cloud data warehouses.

1. Scalability and Flexibility.
2. Cost Efficiency.
3. Performance and Speed.
4. Accessibility and Collaboration.
5. Enhanced Security and Compliance
6. Managed Services and Automation

**CONCLUSION**

Cloud data warehouses have transformed how businesses manage and analyse data by offering a scalable, cost-effective, and adaptable solution that addresses the challenges of traditional on-premises data warehousing. These modern data warehouses provide unrivalled benefits in terms of scalability, performance, and ease of integration with other cloud-based tools and services. They enable businesses to handle massive amounts of data with agility, perform real-time analytics, and gain advanced insights through seamless integration with machine learning and AI platforms.

The pay-as-you-go pricing model and managed services reduce the financial and operational burdens associated with maintaining hardware and software, allowing organizations to focus more on deriving actionable insights from their data. Additionally, the enhanced security features and compliance capabilities of cloud data warehouses provide robust protection for sensitive information, ensuring data integrity and resilience in a rapidly evolving digital landscape.

**REFERENCES**

[**https://cloud.google.com/learn/what-is-a-data-warehouse**](https://cloud.google.com/learn/what-is-a-data-warehouse)

[**https://www.informatica.com/in/resources/articles/what-is-a-cloud-data-warehouse.html**](https://www.informatica.com/in/resources/articles/what-is-a-cloud-data-warehouse.html)

[**https://www.matillion.com/learn/blog/cloud-data-warehouse**](https://www.matillion.com/learn/blog/cloud-data-warehouse)