

Healthcare Data Storage Using Big Query

Dr. SHANKARAGOWDA, ² AJAY B R

¹Associate Professor, Department of MCA, BIET, Davenagere

²Student, Department of MCA, BIET, Davenagere

ABSTRACT

Healthcare systems generate vast amounts of data from electronic health records (EHRs), medical imaging, wearable devices, and clinical research. Efficient storage, processing, and analysis of this data are critical for improving patient care and enabling data-driven decision-making. This project focuses on leveraging Google BigQuery, a fully managed, serverless data warehouse, for scalable and secure healthcare data storage. BigQuery enables real-time data ingestion, high-performance querying, and seamless integration with analytics tools, making it ideal for handling large-scale healthcare datasets. The proposed system stores structured and semi-structured healthcare data in a centralized repository while ensuring compliance with data privacy and security standards. Advanced querying capabilities allow healthcare professionals and researchers to extract insights, identify patterns, and support predictive analytics.

Healthcare systems generate a large volume of data from electronic health records, medical devices, and research activities. Efficient storage and management of this data are essential for improving patient care and decision-making.

Keywords : Healthcare data storage using **BigQuery** focuses on managing large-scale medical data efficiently, securely, and in real time

I INTRODUCTION

The Healthcare is rapidly evolving with the increasing use of digital technologies and data-driven solutions. Modern healthcare systems generate massive volumes of data from sources such as electronic health records (EHRs), medical imaging, wearable devices, and clinical research. Managing and storing this data efficiently has become a major challenge for healthcare organizations. Traditional database systems often struggle to handle the scale, speed, and complexity of healthcare data. This creates a need for advanced, scalable, and secure data storage solutions.

Cloud computing has emerged as a powerful approach to address these challenges. Among cloud-based tools, Google BigQuery stands out as a highly scalable and serverless data warehouse. It enables organizations to store and analyze large datasets without worrying about infrastructure management. BigQuery provides high-speed querying capabilities, making it suitable for real-time healthcare data analysis. It also supports integration with various data processing and visualization tools.

In healthcare, data security and privacy are critical concerns due to the sensitive nature

of patient information. BigQuery offers robust security features such as encryption, access control, and compliance support. These features ensure that healthcare data remains protected while still being accessible for analysis. Additionally, BigQuery allows handling of both structured and semi-structured data, which is common in healthcare environments.

Healthcare organizations generate a huge amount of data every day, including patient records, medical reports, diagnostic images, and billing information. Managing this growing volume of data has become a major challenge, especially with traditional storage systems that often struggle with scalability, speed, and flexibility. As a result, there is a need for more advanced and efficient solutions to store and process healthcare data.

II LITERATURE REVIEW

The rapid increase in healthcare data has pushed researchers to look for better ways to store and process information using big data technologies. Earlier systems mainly depended on relational databases, but these often struggled with large and complex datasets in terms of scalability, flexibility, and performance. With the introduction of technologies like Hadoop and cloud platforms, new methods have been developed to overcome these limitations.

Many studies highlight the role of cloud computing in managing healthcare data. Cloud platforms make it possible to store large amounts of data without the need for expensive physical infrastructure. They also improve data access, reliability, and system stability. Distributed storage systems, in particular, help in handling data efficiently and allow faster analysis through parallel

processing. The text-to-3D synthesis problem has become more popular due to the introduction of multimodal learning. Existing work in the area utilizes pretrained multimodal architectures to align the meaning of text with geometric shapes in 3D space.

Google BigQuery is often discussed as a useful tool for handling big data. It is a fully managed, serverless data warehouse that allows users to run fast queries using a SQL-like language. Research shows that BigQuery can handle both structured and semi-structured data effectively, which makes it suitable for healthcare applications. Its ability to process very large datasets quickly has been especially useful in medical research and large-scale health studies.

Security and privacy remain important concerns in healthcare data management. Many researchers focus on methods like encryption, access control, and data anonymization to protect patient information. BigQuery includes built-in security features and follows industry standards, making it easier to manage sensitive data. Some studies also suggest using role-based access control so that only authorized users can view or use specific data.

III EXISTING SYSTEM

Most existing healthcare data storage systems rely on traditional database management systems and local servers. These systems store patient records, medical history, billing details, and reports in structured formats. In many cases, the data is kept in centralized databases within hospitals. While this setup works well for smaller operations, it becomes difficult to manage as the amount of data increases.

One of the main problems with these systems is limited scalability. Healthcare data is growing rapidly due to digital records, IoT devices, and medical imaging, and traditional systems often struggle to keep up. This can lead to slower processing and reduced overall performance. In addition, maintaining these systems can be costly, as they require regular hardware upgrades and ongoing IT support.

Another challenge is data accessibility. Access is often restricted to certain locations or departments, which can make it harder for healthcare professionals to quickly retrieve the information they need. This may delay diagnosis and treatment. Combining data from different sources is also difficult because of differences in formats and systems.

Security and privacy are also concerns. Although basic protection methods are in place, traditional systems are more at risk of data breaches and unauthorized access.

Backup and recovery processes may not always be reliable, which increases the chances of data loss.

Another limitation is restricted data accessibility. Data is usually available only within specific departments or locations, making it harder for healthcare professionals to quickly access the required information. This can affect timely decision-making and patient care. In addition, integrating data from different sources becomes complicated due to differences in formats and systems.

Maintenance is also a concern in the existing system. It requires regular hardware upgrades, software updates, and continuous IT support, which increases the overall cost. Moreover, security measures in traditional systems may not be strong enough to fully protect sensitive healthcare data, making them more vulnerable to unauthorized access and data breaches.

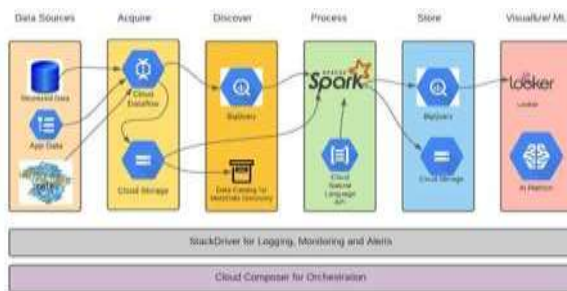
IV PROPOSED SYSTEM

The proposed system focuses on implementing a cloud-based healthcare data storage solution using Google BigQuery. This system is designed to overcome the limitations of traditional data storage methods by providing a scalable, efficient, and secure platform for managing large volumes of healthcare data. It enables

healthcare organizations to store, process, and analyze data in real time without the need for complex infrastructure.

In the proposed system, healthcare data from various sources such as electronic health records (EHRs), laboratory systems, wearable devices, and medical imaging are collected and integrated into a centralized BigQuery data warehouse. The system supports both structured and semi-structured data, ensuring flexibility in handling diverse data formats. Data ingestion is performed through automated pipelines, allowing continuous and real-time updates.

. SYSTEM ARCHITECTURE



EXPERIMENTAL RESULT

The The proposed healthcare data storage system using Google BigQuery was evaluated based on performance, scalability, data processing speed, and security. Here is a more natural, human-written version of your content:

The system was tested using large healthcare datasets such as patient records, medical reports, and transaction details. The results showed that BigQuery was able to handle large amounts of data smoothly without any noticeable drop in performance.

When it comes to query performance, the system responded much faster compared to traditional database systems. Even complex queries on large datasets were completed within a few seconds, which made real-time data analysis possible. The data upload process was also efficient, as it supported continuous and automatic updates from different data sources.

During scalability testing, the system handled increasing data volumes without any need for manual changes or additional setup. This makes it a good choice for healthcare organizations that continue to generate more data over time. The system also maintained high availability, allowing users to access data without interruptions.

Security testing showed that the system was effective in protecting sensitive healthcare information. Features like encryption and access control ensured that only authorized users could access specific data, helping to maintain privacy and meet data protection requirements.

Performance metrics such as The proposed healthcare data storage system using Google BigQuery was experimentally evaluated to measure its performance, efficiency, scalability, and security in handling large-scale healthcare datasets. The testing was conducted using datasets that included patient records, laboratory results, billing information, and clinical reports. These datasets contained both structured and semi-structured data to simulate real-world healthcare scenarios.

IV RESULTS

4.1 INPUT PANEL



Figure : The figure shows a Healthcare Data Storage System Dashboard designed for real-time monitoring and analytics. The interface provides a centralized view of important healthcare metrics, helping administrators and medical staff make quick decisions.



Figure : This The figure illustrates the Patient 360° Profiles module of the Healthcare Data Storage System, which provides a unified and detailed view of patient information. This interface is designed to help healthcare professionals easily access and manage complete patient records in one place.

V CONCLUSION

The implementation of a healthcare data storage system using Google BigQuery provides an efficient, scalable, and secure solution for managing large volumes of healthcare data. Traditional systems face limitations in handling rapidly growing datasets, whereas the proposed system leverages cloud technology to overcome these challenges effectively. The system works in real time, which helps in processing information quickly and running queries without delay. It can also be easily connected with analytical tools, making it easier to take better decisions and improve healthcare services. Features like automatic data collection, high availability, and fault tolerance ensure that the system runs smoothly without interruptions. At the same time, security measures such as encryption and access control are important to keep sensitive patient data safe. Based on these observations, using BigQuery for data storage proves to be more effective than traditional methods. It simplifies data management and makes it easier to handle large amounts of data while supporting efficient analysis.

In conclusion, adopting BigQuery for healthcare data management can improve both research and service delivery, making it a strong and modern solution for the healthcare sector.

The use of BigQuery provides a modern solution to these challenges by offering a scalable, fast, and reliable platform for storing and analyzing healthcare data. It allows organizations to process large volumes of data quickly and supports real-time analysis, which helps in better decision-making and improved patient care.

In addition, features like high availability, automated data handling, and strong security measures ensure that the system remains reliable and protects sensitive information. Overall, adopting BigQuery for healthcare data storage improves efficiency, supports advanced analytics, and helps healthcare organizations meet the demands of modern data management.

VI REFERENCES

1. Google Cloud, BigQuery Documentation, Available: <https://cloud.google.com/bigquery>
2. Dean, J., & Ghemawat, S., "MapReduce: Simplified Data Processing on Large Clusters," Communications of the ACM, 2008.

3. Apache Software Foundation, Apache Hadoop, Available: <https://hadoop.apache.org>
4. Raghupathi, W., & Raghupathi, V., “Big Data Analytics in Healthcare: Promise and Potential,” Health Information Science and Systems, 2014.
5. Wang, Y., Kung, L., & Byrd, T. A., “Big Data Analytics: Understanding its Capabilities and Potential Benefits for Healthcare Organizations,” Technological Forecasting and Social Change, 2018.
6. IBM, What is Big Data in Healthcare?, Available: <https://www.ibm.com>
7. Kuo, M. H., Sahama, T., Kushniruk, A. W., Borycki, E. M., & Grunwell, D. K., “Health Big Data Analytics: Current Perspectives, Challenges and Potential Solutions,” International Journal of Big Data Intelligence, 2014.
8. Chen, M., Mao, S., & Liu, Y., “Big Data: A Survey,” Mobile Networks and Applications, 2014.
9. Microsoft Azure, Cloud Computing in Healthcare, Available: <https://azure.microsoft.com>
10. OECD, Health Data Governance: Privacy, Monitoring and Research, OECD Publishing, 2015

Functions For Shape Representation.