ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Special Issue | Volume XIV, Issue XIII, October 2025

Exploring Cloud Solutions for Real-Time Big Data Processing

Vandana Nemane*, Prajakta Patil

Department of Computer Science, Dr. D. Y. Patil Arts, Commerce and Science College, Pimpri, Pune, Maharashtra, India

DOI: https://doi.org/10.51583/IJLTEMAS.2025.1413SP017

Received: 26 June 2025; Accepted: 30 June 2025; Published: 23 October 2025

Abstract: Cloud-based solutions provide flexibility in storage and processing capabilities, allowing for tailored adjustments as organizational needs evolve. Collaboration is fostered, enabling data sharing and teamwork among diverse users and teams. Accessibility becomes universal, harnessing the potential of big data analytics from any location with an internet connection. The period of big data has brought about unknown challenges and openings for data processing and analytics. The volume, haste, and variety of data generated bear scalable and effective processing results. Real- time, contextual, and secure data is now charge critical.data streaming platforms (DSPs) play a vital part in simplifying access to real- time data and making it easy to exercise whenever and wherever it's demanded. This paper explores the concept of big data processing in the cloud and its significance in enabling real- time data analytics. By using cloud infrastructure, we can process huge amounts of data assisting real-time data analysis.

The paper delves into the benefits of cloud based big data processing, including elastic resource provisioning, cost optimization, and simplified data operation. Through this comprehensive analysis, the paper aims to exfoliate light on the eventuality of cloud computing in scalable and real- time big data analytics, empowering associations to decide precious perceptivity and make data-driven opinions.

Keywords: Big Data Processing, Cloud Computing, Real-time Data Analytics

I. Introduction:

For organizations looking to gain insightful knowledge and make data-driven choices, the exponential growth of data in the digital age has brought both tremendous potential and tremendous obstacles. The enormous amount, rapid speed, and wide range of data generated from several sources, including social media, Internet of Things (IoT) devices, sensors, and more, define this phenomena, which is also referred to as big data. For businesses in a variety of sectors, efficiently handling and evaluating this enormous volume of data has become essential.

Cloud computing has become a revolutionary technology that offers scalable, adaptable, and affordable solutions to the problems associated with processing large amounts of data. In order to effectively manage large-scale data processing operations, the big data processing in the cloud idea makes use of distributed computing frameworks and cloud infrastructure. Organizations may process large datasets in parallel and extract real-time insights from the data thanks to cloud-based platforms' ability to dynamically scale computing resources on-demand.

Cloud computing's real-time data processing has become a key component of digital transformation, allowing businesses to handle and analyze data streams as they are created in real time. With the continued acceleration of global data generation—current estimates show that companies process—this skill has become more and more important. The applications and advantages of cloud solutions for real-time big data processing are examined in this research.

II. Literature Review:

A study was conducted by Uğur KEKEVİ, Ahmet Arif AYDIN (2022) on the Real-Time Big Data Processing and Analytics: Concepts, Technologies, and Domains emphasize the following elements: The terms and concepts pertinent to real-time analytics and data processing are given. The capabilities, traits, working principles, and uses of common real-time data processing technologies are covered. Real-time data processing application domains and popular NoSQL storage solutions—especially those utilized in real-time data processing—are described. In order to assist data processing and analytics researchers as well as developers of data-intensive systems, the main goal of this article is to present a comparative analysis of real-time data processing and NoSQL storage technologies by emphasizing the salient features, capabilities, and attributes of these technologies.

Another related study was carried out by Venkata Reddy Keesara (2025) Real-Time Data Processing in Cloud Computing: Enterprise Implementation Strategies shows cloud computing-enabled real-time data processing has become a disruptive force in contemporary enterprise design, radically altering how businesses manage, examine, and extract value from their data streams. It illustrated how real-time processing has a big impact on how businesses operate and make decisions. The limits of real-time data processing are being pushed further by the incorporation of cutting-edge technologies like edge computing, artificial intelligence, and quantum-inspired algorithms.

Further study was conducted by Mustapha Malami Idina (2023) The Concept of Big Data and Solutions of Cloud Computing where Big data is a vast topic that is changing quickly. Two of the main applications of big data platforms are the discovery of previously unnoticed patterns and the comprehension of previously unidentified behaviors. The rapid growth of AI, ML, and Big

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Special Issue | Volume XIV, Issue XIII, October 2025

Data demands the creation of effective techniques for examining the massive amounts of data produced every day. Our civilization depends heavily on big data and cloud computing, especially in sectors like the medical, commercial, and aviation sectors, among others, which produce vast volumes of data that must be cleaned up to solve the numerous issues they face. The efficacy of big data solutions is greatly increased by cloud settings, which give big data systems environments that are available, scalable, and fault-tolerant.

Another study was done by Ibrahim Abaker Targio Hashem et. al. (2014) The rise of "big data" on cloud computing: Review and open research issues observed that one of the most effective technologies for sophisticated and large-scale computing is cloud computing. It does away with the requirement to maintain costly computer gear, specialized space, and software. There has been a noticeable increase in the volume of data, or big data, produced by cloud computing. Additionally examined are Hadoop technology, big data storage systems, and the connection between big data and cloud computing. Scalability, availability, data integrity, data transformation, data quality, data heterogeneity, privacy, legal and regulatory concerns, and governance are among the other study topics that are examined.

One more study was conducted by Ahmed Hadi Ali AL-Jumaili (2023) Big Data Analytics Using Cloud Computing Based Frameworks for Power Management Systems: Status, Constraints, and Future Recommendations introduced the concepts and features of cloud computing are presented, together with a brief description of new parallel programming models like Hadoop, Spark, and Storm, and cloud computing solutions for the field of power equipment condition monitoring against the backdrop of big data. Lastly, a thorough cloud computing platform architecture is created to satisfy the multi-level real-time needs for various application scenarios in the field of power equipment condition monitoring by utilizing the Internet's experience.

III. Methodology:

Cloud-based platforms offer scalable and efficient solutions for real-time big data processing, with services like Amazon Kinesis, Google Cloud Dataflow, and Azure Stream Analytics. These platforms enable organizations to ingest, process, and analyze vast amounts of data as it streams, providing immediate insights and facilitating faster decision-making.

Real-Time Data Processing Technologies:

This section presents a list of prominent technologies used in real-time data processing that have been recognized in published research.

Apache Kafka:

Apache Kafka is a scalable software framework for managing and making judgments regarding massive volumes of streaming data utilizing in-memory analytics. Four prominent properties of Kafka are distributed processing, high throughput, persistent messaging, and disk structures. It combines online and offline processing of these datasets with real-time processing of activity data and operational data. Kafka distributes a partition across several nodes because it assigns sequential identities to messages.

Apache Spark Streaming:

Apache Spark is an open- source, distributed processing system designed for large- scale data processing and analytics. It provides a unified machine for different workloads like batch processing, interactive queries, real- time analytics, machine literacy, and graph processing. Spark utilizes in- memory hiding and optimized query execution to achieve rapid processing speeds for large datasets.

Apache Flink:

It is a stream processing framework that supports complex event processing and stateful computations. [11] Apache Flink is an open-source stream processing framework with powerful stream- and batch-processing capabilities. Apache Flink is a versatile alternative to MapReduce that supports both batch and real-time data processing.

[12] Flink aims to deliver more accessible and scalable real-time computing solutions, empowering organizations to fully embrace real-time capabilities across the entire spectrum of big data and AI applications.

Apache Storm:

A distributed real-time computation system for processing unbounded data streams. [13] Apache Storm is a free and open source distributed real-time computation system. Apache Storm makes it easy to reliably process unbounded streams of data, doing for real time processing what Hadoop did for batch processing. Apache Storm is simple, can be used with any programming language.

IV. Applications of real-time data processing:

A wide range of sectors and fields have embraced real-time data processing to handle important opportunities and difficulties. Among its main uses are:

- **Fraud detection:** Quickly spots questionable transactional trends to stop fraud and safeguard financial integrity.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Special Issue | Volume XIV, Issue XIII, October 2025

- **Trading and Finance:** Provides real-time market monitoring, algorithmic trading, and risk management, guaranteeing quick reaction to changing market conditions.
- **IoT and Smart Systems:** Handle massive amounts of sensor data, enabling predictive maintenance, smart home/city applications, and effective operational automation.
- **Healthcare Monitoring:** Provides the energy for real-time patient monitoring systems that recognize vital health indicators, facilitating prompt medical attention and better patient results.
- **Logistics and Transportation:** Evaluates real-time traffic and logistics data to optimize real-time scheduling, routing, and delivery procedures.
- **Cybersecurity Threat Detection:** Constantly assesses data feeds to identify and promptly address possible security events or breaches.
- **Customer Analytics:** Provides targeted communication, real-time recommendations, and instantaneous analysis of user activity to enable personalized customer experiences.

V. Benefits of Real-Time Data Processing:

Because it allows for instantaneous, actionable insights, real-time data processing has clear advantages over conventional techniques. Some significant benefits are:

- **Better Decision-Making**: Quick, well-informed decisions are made possible by instant insights.
- **Enhanced System Dependability:** Facilitates real-time monitoring and notifications, leading to prompt problem detection and proactive fixing.
- **Competitive Advantage:** Businesses that use real-time data are able to quickly modify their plans in response to early market insights.
- **Improved Customer Experience:** Businesses can provide real-time interactions, personalized recommendations, and prompt customer care thanks to immediate data processing.
- **Increased Operational Efficiency:** Resource usage is optimized, procedures are made simpler, and operational waste is decreased with real-time visibility into operations.
- **Quick Threat Identification:** Constant real-time analytics enable prompt detection of irregularities, security risks, or fraudulent activity, enabling prompt preventive action.

VI. Conclusion:

The digital age demands companies to be visionary rather than reactive when it comes to exercising data. They must concentrate on making real- time data work for prophetic perceptivity that drive strategic opinions. The line of real- time data processing is clear further integration, briskly processing, and sophisticated logical capabilities. In the future, businesses will continue to evolve their structure to handle larger volumes of data more efficiently, integrating pall technologies and advanced algorithms that make real- time perceptivity indeed more practicable. Real- time data processing through pall computing has surfaced as a transformative force in ultramodern enterprise armature, unnaturally changing how associations handle, dissect, and decide value from their data aqueducts. The elaboration from traditional batch processing to sophisticated sluice processing infrastructures has enabled associations to achieve unknown situations of functional effectiveness and client responsiveness. The integration of arising technologies similar as edge computing, artificial intelligence, and amount- inspired algorithms continues to push the boundaries of what is possible in real- time data processing. As associations continue to borrow and optimize these technologies, the future of real- time processing looks promising, with new infrastructures and methodologies arising to address decreasingly complex business conditions. The challenges of maintaining data thickness, managing peak loads, and optimizing costs remain important considerations, but innovative results and stylish practices are continuously evolving to address these enterprises.

References:

1. Mustapha Malami Idina (2023) “The Concept of Big Data and Solutions of Cloud Computing”, International Journal of Advanced Engineering and Management Research, Vol. 8, No. 02, ISSN: 2456-3676
2. Venkata Reddy Keesara (2025) “Real-Time Data Processing in Cloud Computing: Enterprise Implementation Strategies”, International Journal of Scientific Research in Computer Science, Engineering and Information Technology, Volume 11, Issue 1, ISSN: 2456-3307
3. Uğur KEKEVİ et.al. (2022) “Real-Time Big Data Processing and Analytics: Concepts, Technologies, and Domains”, Journal of Computer Science, Volume 7, issue 2, e-ISSN –2548-1304.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Special Issue | Volume XIV, Issue XIII, October 2025

4. Ahmed Hadi Ali AL-Jumaili et. al. (2023) "Big Data Analytics Using Cloud Computing Based Frameworks for Power Management Systems: Status, Constraints, and Future Recommendations", Sensors 2023, 23, 2952. https://doi.org/10.3390/s23062952
5. Ibrahim Abaker Targio Hashem et. al. (2014) "The rise of "big data" on cloud computing: Review and open research issues", Information Systems, Elsevier, ISSN - 0306-4379
6. Apache Software Foundation. (2022g). Kafka. https://kafka.apache.org/
7. Philip Chen, C. L., & Zhang, C. Y. (2014).
8. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. Information Sciences, 275, 314–347. https://doi.org/10.1016/j.ins.2014.01.015
9. https://www.researchgate.net/publication/366071082_Real-Time_Big_Data_Processing_and_Analytics_Concepts_Technologies_and_Domains
10. Apache Flink 2.0.0: A new Era of Real-Time Data Processing March 24, 2025 - Xintong Song
11. https://www.instaclustr.com/education/real-time-streaming/real-time-data-processing-benefits-challenges-and-best-practices/
12. https://aws.amazon.com/what-is/apache-flink
13. https://storm.apache.org/