Volume 9, Issue 4, pp. 81-85, 2025.
International Journal of Scientific Engineering and Science
ISSN (Online): 2456-7361

Modern Models of Real-Time Data Processing in Cloud Infrastructures

Prashant Dathwal
Senior Principal Engineer, Oracle
Santa Clara, California, USA

***Abstract*— This article explores modern models of real-time data processing in cloud and hybrid infrastructures. The relevance of the topic is driven by the rapid growth in data volumes and the growing demand for instant analytics across various domains—from industrial manufacturing to smart city intelligence. The novelty of the study lies in examining the capabilities of distributed architectures, security mechanisms, and hybrid approaches that integrate cloud and edge computing. The paper reviews existing solutions and analyzes the advantages of micro service architectures and containerized services, which have enabled flexibility and rapid scalability. It also investigates approaches to ensuring security, with particular attention given to authentication methods, encryption, and anomaly detection. The aim of the study is to iden tify key trends and determine the factors influencing the successful implementation of streaming analytics. To achieve this, the article employs comparative analysis, a historical-analytical review of sources, and content analysis. The conclusion presents the results of the architectural comparison, assessing the effectiveness of various models and outlining promising directions for further development. This work will be useful for IT p rofessionals, system architects, and researchers interested in distributed computing systems.*

***Keywords*— Real-time data processing, cloud computing, hybrid infrastructure, edge computing, microservices, containerization, scalability, security, authentication, streaming analytics.*

I. Introduction

The relevance of this topic is driven by the rapidly increasing volume of information generated by sensors, industrial devices, and various online services. In many sectors, real-time analytics is critical for timely decision-making, requiring the use of distributed cloud-based solutions capable of processing large volumes of streaming data in a flexible and secure manner. At the same time, cloud–edge integration helps reduce latency and optimize system load while maintaining high fault tolerance.

The aim of this study is to identify the key characteristics and most effective approaches for organizing high-performance and secure real-time data processing based on cloud and hybrid architectures.

To achieve this, the following objectives were set:

1. To analyze current solutions in microservice architecture and containerization that enable flexibility and scalability.
2. To examine existing cloud security mechanisms, including authentication, encryption, and anomaly detection systems.
3. To assess the effectiveness and prospects of hybrid approaches that integrate cloud resources with edge computing.

The novelty of this research lies in the simultaneous exploration of several aspects of streaming analytics, including scalability, security, and distributed architecture.

II. Materials And Methods

The study is based on the work of a number of authors who have examined modern approaches to real-time data processing in cloud environments and related issues of security and scalability. A.K. Keesara [4] presented a concept for using cloud platforms to address continuous analytics in large enterprise systems, showing how optimized stream processing can improve performance and reduce infrastructure costs. A.H. Mohammed [5] focused on the combined use of cloud and edge resources, emphasizing the potential for load balancing and reduced network latency by bringing computation closer to the data source. A.H.A. Al-Jumaili, R.C. Muniyandi, M.K. Hasan, J.K.S. Paw, M.J. Singh [2] examined the advantages and limitations of using big data cloud frameworks in energy consumption management, identifying key obstacles to stream analytics and suggesting optimization strategies.

N. Gomez Larrakoetxea, B. Sanz Uquijo, I.P. Lopez, J.G. Barruetabena, P.G. Bringas [3] analyzed the Edge Computing paradigm in the context of Industry 4.0, demonstrating how distributed architecture enhances real-time efficiency when collecting and processing sensor data. N.F. Prangon, J. Wu [7] addressed the convergence of cloud and edge computing, particularly the application of artificial intelligence and machine learning within a unified ecosystem. M. Pum [8] focused on AI-driven monitoring and troubleshooting mechanisms, detailing how to respond promptly to failures in large-scale distributed infrastructures.

F. van der Vlist, A. Helmond, F. Ferrari [9] introduced the concept of "Big AI," highlighting the growing interdependence between cloud providers and AI-based applications, as well as how this dynamic influences the industrialization of AI technologies. F. Paul, J. Bauer [6] studied hybrid deployment models where some functions are processed locally while others are offloaded to public cloud platforms, emphasizing the need for scalable architectures and high responsiveness in stream data scenarios. Y.W. Weldegeorgise, Z. Samira, O.S. Osundare, H.O. Ekpobimi, R.C. Kandekere [10] explored the role of cloud technologies in supporting small and medium-sized enterprises, outlining mechanisms for rapid scalability and advantages for real-time operations. R. Alghofaili, A. Albattah, N. Alrajeh, M.A. Rassam, B.A.S. Al-rimy [1] addressed cybersecurity concerns and described methods for ensuring infrastructure protection, including distributed authentication systems and anomaly detection tools.

A range of scientific methods were used to develop a comprehensive understanding of the topic. Comparative analysis enabled the juxtaposition of different authors' conclusions, revealing common trends and differences in their approaches to infrastructure organization and security. A historical-analytical review traced the evolution of cloud and edge computing concepts and illustrated how priorities in scaling and data protection have shifted over time. Content analysis of scientific and applied sources helped structure the key solutions presented by researchers and identify the factors that influence the successful implementation of streaming analytics. The generalization and synthesis of these findings contributed to the development of an integrated model that reflects current trends, challenges, and future directions for cloud and hybrid infrastructures in real-time data processing.

III. Results

The evolution of modern approaches to real-time data processing in cloud infrastructures is grounded in the integration of distributed computing resources, scalable containerization platforms, and optimized machine learning techniques, as confirmed by numerous experimental studies and practical use cases [4]. Experts note that when dealing with large volumes of telemetry data from industrial sensors, systems must deliver high performance, reliability, and adaptability under fluctuating workloads [5].

One of the key trends is the integration of cloud computing with edge nodes, which significantly reduces data exchange latency and streamlines analytics, particularly in critical real-time scenarios [3]. This is achieved by relocating part of the computational workload as close to the data source as possible, which is especially relevant for intelligent manufacturing systems and Internet of Things (IoT) environments [7].

Implementing high-performance systems for real-time (or near-real-time) data processing requires strict adherence to technical specifications across all levels of the distributed infrastructure. Research by Mohammed A. H. [5] demonstrated that the intelligent optimization of co-flow scheduling in large-scale data exchanges can increase throughput by up to 3.66 times compared to traditional per-flow planning. In distributed architectures processing numerous parallel streams, such optimization maintains high efficiency—over 95%—even under heavy load and competition for network resources. Additional experiments confirm that well-designed co-flow strategies enable near-optimal stream execution times, reducing average data transmission completion times by factors of 2.5 to 4.7.

In the same study, Mohammed A. H. [5] presented the advantages of dynamic packet management in streaming analytics systems. The results show that adaptive packet scaling reduces average latency by 45% while maintaining or even improving throughput (by up to 50% in some scenarios compared to static configurations). These mechanisms prove highly effective across a wide range of workloads (from 100 to 10,000 events per second), sustaining stable performance in over 90% of cases despite varying traffic profiles. The performance gains across four real-time system metrics—efficiency, load stability, end-to-end latency reduction, and overall throughput—are illustrated in Figure 1.

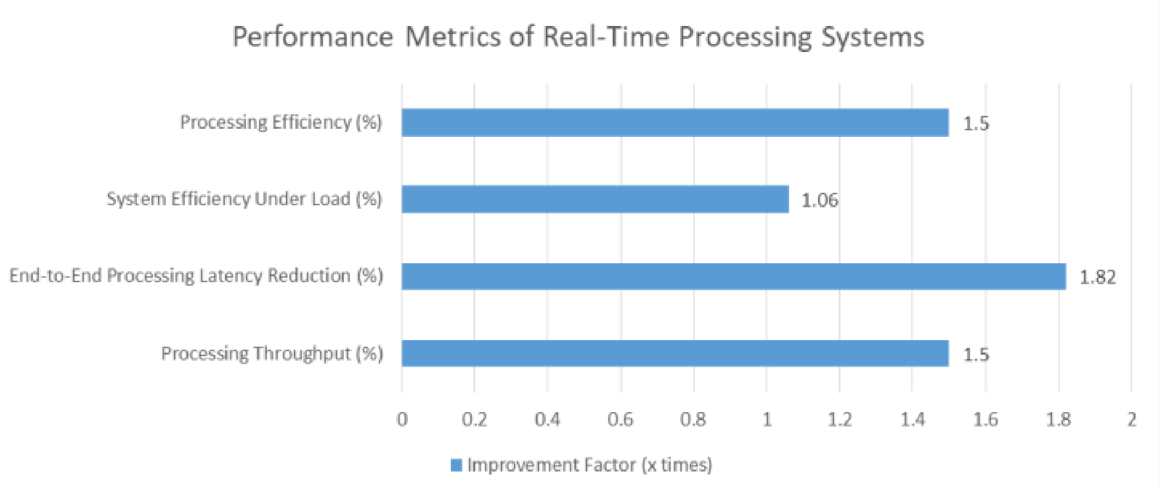

Figure 1. Performance metrics of real-time data processing systems [5]

These findings underscore the need for combining intelligent network interaction planning with adaptive packet-level scaling to improve responsiveness and throughput amid constant changes in workload profiles. This is especially critical in hybrid or cloud–edge systems, where high performance can be sustained through flexible stream orchestration and dynamic allocation of computing resources.

Research on cloud services highlights the importance of resource orchestration and adaptive load balancing, where the synergy between microservice architecture and containerized infrastructure simplifies the deployment of streaming analytics applications and improves responsiveness to data spikes [2]. At the same time, advanced security mechanisms are emerging to protect infrastructure and stored data, as the multi-tenant nature

International Journal of Scientific Engineering and Science
Volume 9, Issue 4, pp. 81-85, 2025. ISSN (Online): 2456-7361

of the cloud introduces risks of data leakage and unauthorized access to sensitive information [1].

To detect suspicious activity and automatically counter threats in real time, distributed monitoring systems powered by machine learning are being implemented. These systems adjust trust levels for network nodes based on user behavior analysis and identifiable intrusion patterns [8]. Researchers emphasize that deploying such security mechanisms requires efficient access management, especially in environments where the infrastructure frequently scales and the composition of computing resources is constantly evolving [9].

In addition to classical public cloud models, the hybrid approach is being actively explored, wherein part of the workload is processed on local clusters while the rest is offloaded to powerful external resources. This configuration enables both system stability and minimal latency, while also allowing for cost-efficiency during peak loads [6]. Experiments show that adaptive function distribution is particularly effective in hybrid scenarios: priority and mission-critical data streams are directed to edge nodes or on-premises servers, while less time-sensitive analytical tasks are executed in the public cloud, which provides scalable storage and deep analytics services [10]. This architecture enables agile responsiveness to market fluctuations or shifts in industrial conditions without compromising performance indicators: response times remain stable even under high input volumes, and overall long-term resource consumption is optimized due to the distributed nature of the system [4].

Collectively, the research confirms that well-designed distributed data processing technologies—tightly integrated with machine learning techniques and robust security mechanisms—form a universal platform for agile and efficient information stream management [5]. Such a platform can scale virtually without limitation, ensuring minimal processing latency and robust security across multiple industries, including manufacturing, energy, logistics, and smart city infrastructure management [2].

Moreover, implementation statistics reveal reductions in operational costs and accelerated decision-making processes, as high-performance streaming analytics systems not only detect anomalies in real time but also predict their likelihood, which is especially valuable when processing large volumes of heterogeneous data [7].

The adoption of hybrid and distributed cloud solutions brings the protection of edge nodes to the forefront, since a substantial share of computational activities takes place outside traditional data centers and requires robust authentication and encryption protocols [1]. Researchers emphasize that a multi-layered security architecture, integrated throughout every phase of the data lifecycle, ensures resilience, reliable access control, and rapid incident mitigation [8]. Additionally, the scale of distributed services enables the deployment of anomaly detection systems trained on diverse datasets sourced from thousands of sensors [9].

Below is a summary table (Table 1) that consolidates the main approaches to real-time and near-real-time data processing in cloud and hybrid infrastructures. It highlights architectural features, key research outcomes, and references to the respective studies where experimental prototypes and methodologies are described in detail.

Below is the second table (Table 2), which outlines the most commonly used mechanisms and solutions for ensuring security in cloud systems that process real-time streaming data. The table summarizes key threat mitigation methods, highlights the strengths and limitations of each approach, and provides references to sources where detailed studies and experiments are discussed.

TABLE 1. Key approaches to real-time and near-real-time data processing in cloud systems

Approach / Model	Description	Outcomes and Benefits	Reference
Microservice architecture with containers	Use of lightweight containers (e.g., Docker, Kubernetes) for flexible orchestration of streaming apps	Simplified scaling, rapid deployment, high fault tolerance	[4]
Distributed streaming analytics	Application of distributed frameworks (e.g., Spark Streaming, Flink) for parallel processing of telemetry	Reduced latency in data ingestion; near-real-time processing achieved	[5]
Cloud–edge integration (Edge Computing)	Offloading computation to the data source to minimize network delays and optimize resource usage	Improved system responsiveness; better handling of large volumes of sensor data	[3]
End-to-end streaming pipeline	Unified “collection–processing–storage” architecture with adaptive load balancers	Reduced response times via dynamic computation allocation; reliable delivery at peak	[7]

(*Source: systematized by the author based on [3; 4; 5; 7]*)

TABLE 2. Security and data protection methods in cloud infrastructures

Method / Technology	Description	Advantages and Limitations	Reference
Distributed monitoring systems	Use of agent networks and machine learning to detect anomalies and intrusions across nodes and channels	High accuracy in threat detection, but depends on training data quality; requires substantial processing power	[8]
Multi-level authentication	Combination of identity verification and access control mechanisms (e.g., tokens, certificates, RBAC)	Flexible permission management and scalability; complex to configure in dynamic node environments	[9]
Data encryption at collection	Encryption applied at the source (edge) to ensure confidentiality during cloud transmission	Prevents data leakage in transit, but may overload edge devices and requires well-defined key management	[1]
Automated access rights management	Dynamic reassignment of roles and policies based on the principle of least privilege	Reduces risk of privilege escalation and unauthorized access, but needs a centralized platform for change tracking	[2]
Hybrid incident response model	Combines local analysis capabilities with large-scale cloud services for rapid threat containment	Ensures fast incident response and system recovery, but can be complex to administer and requires advanced orchestration	[6]

(*Source: systematized by the author based on [1; 2; 6; 8; 9]*)

IJSES
International Journal of Scientific Engineering and Science
Volume 9, Issue 4, pp. 81-85, 2025. ISSN (Online): 2456-7361

The two tables presented above provide a structured overview of the approaches to implementing high-performance streaming systems, along with the principal methods for ensuring data security in real-time and near-real-time environments. This comparative analysis confirms the relevance of distributed and hybrid architectures that balance performance requirements with the complexity of security implementation across all layers of infrastructure.

Thus, real-time analytical solutions based on cloud or hybrid infrastructures combine the computational power of data centers with the low-latency benefits of edge computing, while also enabling continuous monitoring and protection at every level [3]. Together, these factors offer a strong foundation for the continued development of distributed service ecosystems, suitable for a wide range of applications that demand rapid processing of large data volumes, high resilience, and advanced analytics capabilities [10].

IV. Discussion

The analysis of the reviewed studies shows that the effectiveness of modern cloud and hybrid solutions for real-time data processing depends not only on the use of specific technological tools—such as distributed frameworks, containerization, and orchestrators—but also on the degree of coherence in developing and maintaining a unified architectural approach [3; 4]. In environments with rapidly growing data volumes, the distributed model allows for changes to be made in individual modules without interrupting the entire system, enabling flexible update cycles and reducing the risk of large-scale failures [5; 7].

Practical examples [2; 8] demonstrate that the proper design of monitoring and orchestration systems is a key factor in the successful implementation of streaming analytics. A large number of microservices, each with its own database and asynchronous communication channels, can significantly increase the load on communication infrastructure and complicate debugging and configuration management. However, these challenges are offset by the ability to precisely scale individual services—such as machine learning modules or business logic components—that experience the highest demand [1; 10].

A critical aspect of distributed architecture remains secure interservice communication. Establishing a unified gateway for request control, securing APIs, and encrypting communication channels requires a systemic approach and continuous vulnerability monitoring [6; 9]. Researchers emphasize that hybrid scenarios, involving edge components and public cloud environments, must take into account the complexity of access management and permission segregation. As the number of nodes changes dynamically, the system becomes more vulnerable unless clearly defined and unified standards are in place [5].

Industry practice [3; 8] confirms that deep automation of CI/CD processes (continuous integration and delivery), combined with containerization tools such as Docker and Kubernetes, accelerates version rotation and improves resource management. Distributed logging systems and configuration management tools enable the isolation and resolution of issues within specific microservices without compromising the overall stability of the application [4; 10].

Authors [2; 7] also point out that adopting modern technologies in a multi-team development environment requires alignment in design principles, testing practices, and security protocols. While high autonomy among teams facilitates faster feature delivery, it also necessitates maintaining a unified architectural vision to prevent service duplication and excessive functional redundancy [1].

Thus, accumulated experience suggests that with proper planning, standardization, and a mature DevOps culture, a distributed architecture combined with a hybrid deployment model can ensure high fault tolerance, scalable performance, and effective data protection. This, however, requires not only technical expertise in orchestration, containerization, and monitoring, but also a well-thought-out organizational structure that supports consistent rules and standards throughout development and operations.

V. Conclusion

The findings of this study indicate that modern cloud and hybrid infrastructures designed for real-time data analytics deliver high performance, flexibility, and reliability—provided that architectural decisions are well-founded.

The first objective—analyzing microservice architecture and containerization—demonstrated that these approaches significantly simplify scaling and automate deployment processes in rapidly changing load conditions. The second objective—examining security mechanisms—confirmed the importance of multi-level authentication, encryption at all stages of data transmission and storage, and the application of machine learning for anomaly detection. The third objective—assessing the prospects of a hybrid model—revealed that combining cloud resources with edge computing enables fast response times, resource efficiency, and continuous monitoring.

In summary, a distributed architecture that incorporates modern requirements for scalability and security provides an effective solution for managing the growing flow of data and offers ample opportunities for the continued evolution of real-time technologies.

References

1. Alghofaili Y., Albattah A., Alrajeh N., Rassam M. A., Al-rimy B. A. S. Secure Cloud Infrastructure: A Survey on Issues, Current Solutions, and Open Challenges // Applied Sciences. – 2021. – Vol. 11. – Article ID: 9005. – DOI: 10.3390/app11199005. – URL: https://doi.org/10.3390/app11199005 (accessed: April 1, 2025).
2. Al-Jumaili A. H. A., Muniyandi R. C., Hasan M. K., Paw J. K. S., Singh M. J. Big Data Analytics Using Cloud Computing Based Frameworks for Power Management Systems: Status, Constraints, and Future Recommendations // Sensors. – 2023. – Vol. 23. – Article ID: 2952. – DOI: https://doi.org/10.3390/s23062952.

[2]. Gómez Larrakoetxea N., Sánz Uquijo B., López I. P., Barruetabeña J. G., Bringas P. G. Enhancing Real-Time Processing in Industry 4.0 Through the Paradigm of Edge Computing // Mathematics. – 2025. – Vol. 13. – Article ID: 29. – DOI: 10.3390/math13010029. – URL: https://doi.org/10.3390/math13010029 (accessed: April 3, 2025).

[3]. Keesara V. R. Real-Time Data Processing in Cloud Computing: Enterprise Implementation Strategies // International Journal of Scientific Research in Computer Science Engineering and Information Technology. – 2025. – Vol. 11, No. 1. – P. 1295–1304. – DOI:

Volume 9, Issue 4, pp. 81-85, 2025. ISSN (Online): 2456-7361

10.32628/CSEIT251112134. – URL: https://www.researchgate.net/publication/388655648_Real-Time_Data_Processing_in_Cloud_Computing_Enterprise_Implementation_Strategies (accessed: April 3, 2025).
[4]. Mohammed A. H. Real-Time Data Processing in Cloud and Edge Computing // International Journal of Computer Engineering & Technology. – 2024. – Vol. 15, No. 6. – P. 1940–1951. – DOI: 10.34218/IJCET_15_06_166. – URL: https://www.researchgate.net/publication/387555297_REAL-TIME_DATA_PROCESSING_IN_CLOUD_AND_EDGE_COMPUTING (accessed: April 3, 2025).
[5]. Paul F., Bauer J. Real-Time Data Processing in Hybrid Cloud Systems: Architecting for Speed and Scalability. – 2025. – March. – URL: https://www.researchgate.net/publication/389986786_Real-Time_Data_Processing_in_Hybrid_Cloud_Systems_Architecting_for_Speed_and_Scalability (accessed: April 3, 2025).
[6]. Prangon N. F., Wu J. AI and Computing Horizons: Cloud and Edge in the Modern Era // Journal of Sensor and Actuator Networks. – 2024. – Vol. 13. – Article ID: 44. – DOI: 10.3390/jsan13040044. – URL: https://doi.org/10.3390/jsan13040044 (accessed: April 2, 2025).
[7]. Pum M. Artificial Intelligence for Real-Time Cloud Monitoring and Troubleshooting. – 2024. – December. – URL: https://www.researchgate.net/publication/387140941_Artificial_Intelligence_for_Real-Time_Cloud_Monitoring_and_Troubleshooting (accessed: April 3, 2025).
[8]. Van der Vlist F., Helmond A., Ferrari F. Big AI: Cloud Infrastructure Dependence and the Industrialisation of Artificial Intelligence // Big Data & Society. – 2024. – Vol. 11, No. 1. – DOI: 10.1177/20539517241232630. – URL: https://doi.org/10.1177/20539517241232630 (accessed: April 3, 2025).
[9]. Weldegeorgise Y. W., Samira Z., Osundare O. S., Ekpobimi H. Framework for Real-Time Data Analytics Using Cloud Technologies. – 2024. – September. – URL: https://www.researchgate.net/publication/384935988_Framework_for_Real-Time_Data_Analytics_Using_Cloud_Technologies (accessed: April 3, 2025).