

REAL-TIME DATA STREAMING SOLUTIONS IN DISTRIBUTED SYSTEMS

Samarth Shah¹ & Dr S P Singh²

¹*University at Albany, Washington Ave, Albany, NY 12222, United States*

²*Ex-Dean, Gurukul Kangri University, Haridwar, Uttarakhand, India*

ABSTRACT

Real-time data streaming has become a cornerstone for distributed systems, facilitating the rapid and continuous processing of massive data volumes generated by modern applications. This paper explores the core principles, architectural paradigms, and technologies underpinning real-time data streaming in distributed environments. Real-time streaming systems enable low-latency processing, making them essential for applications like fraud detection, predictive analytics, and dynamic resource management.

Key architectural components of these systems include data producers, stream processors, and storage systems. The challenges of achieving scalability, fault tolerance, and consistency are central to the design of these solutions. Strategies such as partitioning, replication, and distributed consensus algorithms are employed to ensure system resilience and reliability in the face of network failures and varying data loads.

Emerging trends focus on the integration of microservices-based architectures and event-driven designs, enhancing modularity and adaptability. Advances in stream processing frameworks and data serialization techniques have further improved throughput and reduced latency, enabling more efficient distributed workflows. Furthermore, the paper addresses the complexities of real-time analytics, including windowing operations and stateful computations, which are critical for deriving actionable insights from continuous data streams.

By synthesizing insights from various domains, this study highlights the evolution of real-time data streaming solutions in distributed systems. It emphasizes the importance of optimizing resource utilization and ensuring seamless integration with diverse data sources, paving the way for future innovations in this dynamic field.

KEYWORDS: *Real-Time Data Streaming, Distributed Systems, Stream Processing, Low-Latency Analytics, Fault Tolerance, Scalability, Partitioning, Event-Driven Architecture, Stateful Computations, Continuous Data Processing*

Article History

Received: 10 Nov 2024 | Revised: 13 Nov 2024 | Accepted: 16 Nov 2024
