authentication with field-level encryption to protect patient data in automated workflows. Full audit logs ensure HIPAA compliance.

Future Trends

With AI and cloud-native advancements, workflow orchestration is evolving toward intelligence and ubiquity:

Intelligence: Large language models redefine human-machine interaction. Tests show GPT-4 can generate executable Airflow code in 10 seconds from prompts like "Monitor server CPU usage; scale resources and alert staff if exceeding 80 %," achieving 92 % accuracy. This lowers design barriers and enables "conversational automation."

Edge Computing: Extends workflows to endpoints. Industrial devices with WebAssembly runtimes execute local quality checks: edge AI analyzes camera images to trigger defect sorting in under 50ms latency, enhancing real-time performance and privacy.

Cross-Ecosystem Interoperability: Leading vendors build open platforms blending blockchain smart contracts with traditional systems. A logistics firm uses Chainlink oracles to optimize routes based on real-time fuel prices, saving $3M annually.

Conclusion

Automated workflow orchestration has evolved from an efficiency tool to a cornerstone of digital transformation. Its value extends beyond replacing repetitive tasks to unlocking systemic synergies through intelligent scheduling. With AI, future systems will exhibit greater adaptability, serving as intelligent bridges between physical and digital ecosystems.

References
1. Apache Airflow in Action. [Electronic resource]. – Access mode : https://airflow.apache.org. – Access date : 20.04.2025.
2. Intelligent Automation: From Tool to Strategy [Electronic resource]. – Access mode : https://www.gartner.com. – Access date : 21.04.2025.
3. Zapier Enterprise Automation Whitepaper [Electronic resource]. – Access mode : https://zapier.com/customer-stories. – Access date : 22.04.2025.
4. Kubeflow [Electronic resource]. – Access mode : https://www.kubeflow.org/docs/ components/ pipelines/. – Access date : 3.05.2025.

UDC 004.415

A REAL-TIME DATA PROCESSING SYSTEM BASED ON APACHE FLINK

Guo Rujie, master's degree student,
Dunina E., PhD in Physics and Mathematics, associate professor,
Dzerkachenka P., senior lecturer
Vitebsk State Technological University, Vitebsk, Republic of Belarus

Abstract. *Real-time data processing has become a cornerstone of modern data-driven industries, enabling organizations to extract actionable insights from high-velocity data streams. Apache Flink, an open-source stream processing framework, addresses the challenges of real-time analytics through its robust architecture and low-latency capabilities. This paper explores the core features of Flink, its applications across diverse domains, and the technical challenges in its deployment.*

Keywords: stream processing, dynamic resource allocation, Apache Flink.

The exponential growth of data in industries such as finance, manufacturing, and healthcare has necessitated advanced tools for real-time analytics. Traditional batch processing systems fall short in addressing the latency and scalability demands of streaming workloads. Apache Flink, a state-of-the-art stream processing framework, bridges this gap by providing low-latency, high-throughput, and fault-tolerant capabilities. This paper examines Flink's architecture, its applications, and the

challenges in leveraging its potential for real-time data processing.

Flink is designed for stream computing, where data is processed as it arrives without requiring intermediate storage. Unlike batch systems that process data in fixed intervals, Flink enables event-driven computation, ensuring responses in milliseconds. For instance, in financial risk control, he analyzes transaction streams in real time to detect fraudulent patterns, achieving sub-second latency.

He supports elastic scaling to adapt to fluctuating workloads. During peak traffic, clusters can scale horizontally to handle millions of events per second. Conversely, during off-peak hours, resources are dynamically reduced to optimize costs.

Flink employs checkpointing and state backends (e.g., RocksDB) to ensure fault tolerance. By persisting intermediate states and leveraging event-time processing, Flink guarantees exactly-once semantics even during node failures. In logistics tracking systems, this ensures uninterrupted analysis of vehicle trajectories despite network disruptions.

Flink's versatility spans multiple industries:
- real-time recommendations: adjusting user recommendations based on clickstream data,
- monitoring sensor data for predictive maintenance.
- financial services: detecting anomalies in transaction streams.

Flink delivers significant business benefits:
- reduced latency: from minutes to milliseconds,
- high throughput: processes millions of records per second.
- complex event processing (CEP): identifies patterns in vehicle telemetry data (e.g., speeding, illegal lane changes),
- cost efficiency: minimizes redundant data storage and optimizes resource usage.
- flexibility: supports dynamic rule updates via SQL or DataStream APIs.

Flink's performance depends on operator chaining and backpressure mechanisms to prevent downstream overloads.

Future advancements may integrate machine learning for automated resource allocation.

While Flink excels in real-time processing, it complements batch systems like Hadoop MapReduce. ERP systems, for example, rely on Flink for dynamic scheduling while managing static resource planning.

Table 1 – Comparison Between Flink and Hadoop MapReduce

Feature	Apache Flink (Stream Processing)	Hadoop MapReduce (Batch Processing)
Latency	Millisecond-level	Minute-level
Resource utilization	Dynamic allocation	Static allocation
Fault Tolerance	Checkpointing (exactly-once semantics)	Data retransmission (at-least-once)
Applicable Scenarios	Real-time analytics, IoT	Log analysis, data warehousing

Leveraging cloud-native architectures (e.g., Kubernetes) and edge computing will enhance Flink's scalability and reduce latency in distributed environments.

Apache Flink has emerged as a transformative tool for real-time data processing, enabling enterprises to derive immediate value from streaming data. Its strengths in low-latency computation, dynamic scalability, and fault tolerance address critical challenges in modern data ecosystems. While challenges remain in algorithmic optimization and system integration, Flink's evolution alongside cloud and edge technologies promises to further unlock its potential in diverse industries.

Reference
1. Website « Apache Flink Official Documentation» [Electronic resource] – Access mode: https:// flink.apache.org – Access date: 10.02.2025.
2. Zhang, L., Wang, J. Flink in Action / Manning Publications, 2023.