Cloudera Data in Motion

Capture, process, and distribute any data anywhere for real-time insights

Choosing the right Data in Motion solution is critical to respond to urgent business events in every major industry. Some of the most transformative uses of data such as network monitoring, cybersecurity, fraud prevention, predictive maintenance, and personalized e-commerce require real-time intelligence. As organizations seek to build applications that deliver greater customer value and operational efficiency across the enterprise, demand for real-time intelligence grows. It is a real-time intelligence network that organizations must deliver, but the traditional analytics architecture was built for business intelligence and reporting on historical data, not real-time data.

To maximize the business and operational impact of investments in real-time data from sensors, modernized applications and systems,change data capture tools, third-party API integrations, networks and infrastructure, organizations need to capture, process, and distribute any data anywhere in real-time and at massive scale:

- **Scalable data movement.** Connect to any data source and make that data available to data streams and any target system.
- **Enterprise Stream Management.** Operate a real-time data store shared between one or many publishing and subscribing applications or systems.
- **Democratized Stream Processing.** Empower non-coder domain experts to process real-time data for automated responses to critical events and analytic insights.

Together, these capabilities enable organizations to maximize the value of their investments, deliver intelligence with near-zero latency, and unlock new use cases with reliable and easily modifiable data pipelines at scale.

Cloudera Data in Motion: Key Features

Data Flow Designer

Powered by Kafka, Kconnect, etc)

Intuitive GUI Self-service
450+ processors
Centralized management

Scalable data movement

Streams Messaging Manager

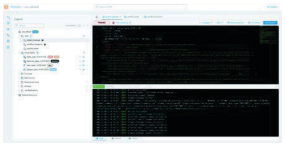
(Powered by Kafka, Kconnect, etc)

Cluster management
Schema goveranance
Replication service

Enterprise stream management

SQL Stream Builder

(Powered by Flnk and Flink SQL)



Instand data access
unified SQL processing
recomposable asset repository

Democratized stream processing

Data in Motion: At a Glance

- As enterprise organizations seek to deliver greater customer value and operational efficiency with data, data teams need a core set of data in motion capabilities to capture, process, and distribute any data anywhere in real-time and at massive scale.
- Cloudera data in motion provides composable services based on open source that make it easy to build, run, and monitor streaming pipelines and deliver actionable insights at the time of an event.
- Data in motion powers the most transformative business and operational use cases, including cybersecurity, fraud detection and prevention, predictive maintenance, and more.

The Business Impact of Data in Motion

PT Bank Rakyat Indonesia, one of the largest and oldest banks in Indonesia, built and deployed a real-time fraud prevention service with Cloudera and **reduced fraud by**

40%

Cloudera Data Flow

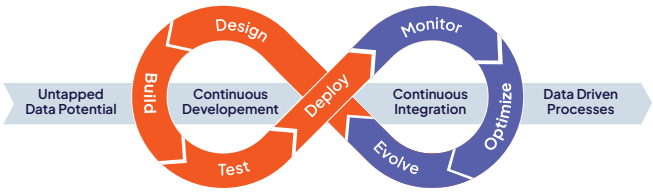
Built on Apache NiFi and MiNiFi, Data Flow provides a drag-and-drop GUI to build and reconfigure pipelines as a series of "processors" that connect to any source and perform a variety of processing functions like filtering, cleansing, enrichment and normalization to deliver processed data to any destination. Data Flow features more than 450 processors and a ReadyFlow Gallery of pre-built templates to jumpstart development. Built-in lineage and provenance makes data governance, security, auditability, and troubleshooting easier than ever.

Streams Messaging Manager

Built on Apache Kafka, Streams Messaging Manager provides a comprehensive management console for data streams operations, including cluster management, schema governance, security, and performance monitoring. Streams Messaging Manager includes features for high-availability and protection from data loss, including intelligent load balancing and data replication.

SQL Stream Builder

Built on Apache Flink, SQL Stream Builder is a comprehensive console that enables customers to build, run, and monitor stream processing jobs. SQL Stream Builder uses SQL to make streaming data accessible to non-coder domain experts so that they can query streaming data and data at rest. It includes a composable asset repository that supports streaming DevOps by making data assets available for collaboration and versioning, so domain experts are even more effective and efficient.



Real-time data pipelines must evolve over time. They require continuous development and integration to embed data deeper into processes and adapt to changing requirements as businesses seek to leverage data to make rapid, incremental improvements. Cloudera data in motion delivers sustained innovation speed over the long term. By building on open source technology, Cloudera data in motion delivers continuous innovation to vibrant developer communities with enterprise-grade support, tooling, and integration. The combined power of Apache NiFi, Apache Kafka, and Apache Flink are brought into a complete user-friendly development experience to make developers maximally productive, delivering business agility.

Benefits of Cloudera Data in Motion

- **Real-time intelligence at scale.** Maximize operational awareness and business responsiveness by identifying and taking action on critical business events as they are happening. Use cases such as fraud monitoring, cyber security, real time customer offers, and process monitoring demand real-time intelligence.
- **Any data anywhere.** Easily capture structured and unstructured data from any source, process data in flight according to the needs of the target system or user, and distribute data to any destination with a universal toolset for data movement and streaming.
- **Efficient scaling.** Data in motion is built for efficiency at massive scale, powered by distributed engines optimized to scale up in any environment. The holistic approach to building, running, and monitoring pipelines eliminates redundant tools and "data hops" between sources and destinations.
- **Deployment Flexibility.** Deploy data-in-motion tools in the cloud to take advantage of elasticity, accessibility, and integration with data at rest, or on premises for predictable costs, ultra-low latency, and advanced monitoring and control. Quickly deploy and manage data-in-motion components independently with Kubernetes operators for lightweight and efficient deployments independent from other Cloudera data services.

Visit us on our website and learn more. cloudera.com/products/dataflow | cloudera.com/products/stream-processing

CLOUDERA

Cloudera, Inc. | 5470 Great America Pkwy, Santa Clara, CA 95054 USA | cloudera.com

Cloudera is the only true hybrid platform for data, analytics, and AI. With 100x more data under management than other cloud-only vendors, Cloudera empowers global enterprises to transform data of all types, on any public or private cloud, into valuable, trusted insights. Our open data lakehouse delivers scalable and secure data management with portable cloud-native analytics, enabling customers to bring GenAI models to their data while maintaining privacy and ensuring responsible, reliable AI deployments. The world's largest brands in financial services, insurance, media, manufacturing, and government rely on Cloudera to be able to use their data to solve the impossible—today and in the future.

To learn more, visit Cloudera.com and follow us on LinkedIn and X.