

Advanced messaging and streams management

CLUSTERA STREAMS MESSAGING

Why Clustera Streams Messaging?

- **100% open source** - Invest in your architecture and scale with confidence knowing that there is no vendor lock-in
- **Most comprehensive Kafka platform** - Clustera offers the best Streams Messaging solution with all necessary ecosystem components for a complete Kafka implementation
- **Data lineage and provenance** - Only platform to offer end-to-end data lineage tracking and provenance across Kafka, MiNiFi, NiFi and Flink
- **Choice of multiple streaming analytic engines** - Extend your Kafka implementation into Apache Flink, Kafka Streams, or Spark Streaming for real-time insights and predictive analytics
- **Hundreds of Kafka customers** - Clustera has hundreds of happy customers getting excellent support on their advanced Kafka deployments
- **Edge-to-cloud platform** - Kafka tightly integrated with NiFi, MiNiFi, Flink and the rest of the SDX components offering unified security and governance across hybrid and multi-cloud environments.

Clustera DataFlow (CDF) is a comprehensive edge-to-enterprise streaming data platform. It addresses the key data management challenges with streaming and IoT data for all types of enterprises. The Streams Messaging capabilities of CDF enable enterprises to ingest and buffer millions of streaming data inputs, allowing them to take on IoT-scale use cases.

Clustera Streams Messaging

Clustera delivers the most comprehensive streams messaging and management capabilities in the industry. It includes -

- Latest certified, secure, and governed Apache Kafka that provides the messaging backbone
- Schema Registry for centralized schema management
- Kafka Streams for real-time analytics
- Kafka Connect for native connectivity with key data sources
- Cruise Control for cluster management and monitoring
- Apache Ranger for rich access control and security
- Streams Messaging Manager for monitoring and management of enterprise Kafka
- Streams Replication Manager for disaster recovery and replication of enterprise Kafka clusters



About Cloudera

At Cloudera, we believe that data can make what is impossible today, possible tomorrow. We empower people to transform complex data into clear and actionable insights. Cloudera delivers an enterprise data cloud for any data, anywhere, from the Edge to AI. Powered by the relentless innovation of the open source community, Cloudera advances digital transformation for the world's largest enterprises.

Learn more at: cloudera.com

Streams Messaging Manager

Cure Kafka blindness with a single monitoring and management dashboard that lets you:

- Troubleshoot your Kafka environment to identify bottlenecks, throughputs, consumer patterns, traffic flow, and so on
- Visualize end-to-end flows and complete data lineage of message streams from producers to topics to consumers
- Optimize your Kafka environment based on the key performance insights gathered from various brokers and topics
- Get alerted on various metrics of critical components of a Kafka cluster including brokers, topics, consumers and producers
- Manage your schemas from within the user interface thanks to the tight integration with the Schema Registry

Streams Replication Manager

Next-gen Kafka replication powered by Mirrormaker 2.0 allows you to:

- Leverage significantly improved Kafka replication capabilities with Mirrormaker 2.0
- Extend your Kafka architecture to handle active-active clusters and disaster recovery scenarios with Streams Replication Manager
- Install an MM2 cluster easily, provide lifecycle management actions on the cluster (start, stop), and manage/monitor the replication flows across clusters
- Monitor all Kafka cluster replications from within Streams Messaging Manager

Cloudera Data Platform (CDP)

Streams Messaging cluster templates are also available on CDP Data Hub to spin up Kafka clusters on your favorite public cloud in a matter of minutes. This will give you access to Kafka, Schema Registry and Streams Messaging Manager. The entire Streams Messaging capability is available within CDP-DC edition. This has the huge benefit of integration with CDP's Shared Data Experience (SDX) for unified security and governance across all components of your streaming architecture.

Key Benefits

- Process millions of messages per second with Apache Kafka
- Adopt a hybrid cloud architecture for your streaming needs across any public cloud
- Re-use schemas, define relationship between schemas, and manage schema versions with Schema Registry
- Leverage the integration of Schema Registry across Kafka and Apache NiFi by using the same schemas from end-to-end
- Optimize and auto-scale your clusters with Cruise Control
- Cure "Kafka Blindness" by getting visibility into all your Kafka clusters with Streams Messaging Manager
- Manage enterprise Kafka data effectively for active-active cluster replication and disaster recovery use cases