

Ingestion and management of real-time streaming data

CLUDERA FLOW MANAGEMENT

Why Cloudera Flow Management?

- **100% open source technology** - Only vendor with this strategy; prevents vendor lock-in and encourages continuous innovation
- **450+ pre-built processors** - Only product to offer such comprehensive connectivity and a no-code user experience from edge to enterprise
- **Built-in data provenance** - Only product in the market to offer out-of-the-box data lineage tracking and provenance on data-in-motion
- **DevOps model boosts agility** - NiFi Registry fosters collaboration amongst developers by enabling versioning, promotion and deployment of flow assets.
- **Monitor Flows** - Custom build NiFi monitoring of flow status and task completion.

Cloudera 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. Cloudera Flow Management (CFM) is a key part of the CDF platform and it addresses enterprise challenges around data ingestion and processing streaming data from a wide range of data sources.

Cloudera Flow Management

Cloudera Flow Management is a no-code data ingestion and management solution powered by Apache NiFi. With a slick user interface, 450+ processors and the NiFi Registry, CFM delivers highly scalable data management and DevOps capabilities to the enterprise. Apache NiFi is meant for large scale, high velocity enterprise data ingestion use cases. Primarily meant for real-time streaming sources such as clickstreams, social streams, log data etc., Apache NiFi can handle all types of data across any type of data source. NiFi Registry, which augments NiFi, enables DevOps teams with versioning, deployment and development of flow applications.

Flow Management cluster templates are also available on Cloudera Data Platform Data Hub to spin up NiFi clusters on your favorite public cloud in a matter of minutes.



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

Data Ingestion Challenges

The key challenges around data ingestion that enterprises are facing today can be addressed easily with Cloudera Flow Management. Some of these challenges are:

- Data movement - Unable to move extremely large amounts of data from one data center into another or from on-premises to cloud or vice-versa.
- Continuous data ingestion - Unable to handle the constant flow of high-velocity, high-volume streaming data from multiple sources.
- Log ingestion - Unable to ingest log data from thousands of applications or systems across the enterprise.
- IoT data processing - Unable to ingest the volume of IoT data coming into the enterprise from thousands of edge agents.

Apache NiFi

Originally built for a government intelligence agency, Apache NiFi has grown and matured over the last several years as a highly scalable and trustworthy data ingestion and integration engine for hundreds of companies using its mission-critical use cases. Backed by an ever-growing community of committers and developers, Apache NiFi has evolved to address some of the most complex problems in the industry.

Apache NiFi has a really intuitive user interface for designing data flow orchestrations for acquiring, processing, and routing data from any source to any target. This is accomplished with a no-code approach to designing these flows by dragging-and-dropping pre-built processors onto the canvas and connecting them up.

Apache NiFi has over 450+ processors that connect with a range of data sources to handle a wide variety of data formats including structured and unstructured data. It can integrate easily with hadoop and non-hadoop type data storage solutions. Apache NiFi can also integrate easily with cloud storage solutions with pre-built processors for all major cloud vendors such as Amazon Web Services, Azure and Google Cloud.

Key Features of Apache NiFi

Some of the key capabilities and features of Apache NiFi are -

- Intuitive visual design tool
- Flow templates
- Guaranteed delivery
- Prioritized queuing
- Flow Specific QoS (latency v throughput, loss tolerance, etc.)
- Data Provenance
- Comprehensive security (Authentication and Authorization)
- Extensible architecture
- Site-to-site communication protocol
- Flexible scaling model
- Parametrization for seamless deployment of flows
- Stateless NiFi execution mode for extremely high performance
- Create custom NiFi monitoring dashboard