Cloudera Edge Management with IBM

Manage, control and monitor the edge for all your streaming and IoT initiatives

Why Cloudera Edge Management with IBM?

100 percent open source technology
Only vendor with this strategy; prevents vendor lock-in and encourages continuous innovation.

Agility of building edge apps
Build edge data flows visually and with no code for edge data collection and processing. This method also reduces the cost of developing Internet of Things (IoT) applications.

Operational ease of edge management
Deploy updates over the air (OTA) to thousands of edge agents at the same time.

Key enabler for IoT initiatives
Build successful IoT initiatives by collecting, curating and analyzing data from thousands of edge devices.

Enable edge intelligence
Process and react to data quickly at the edge.

Operational confidence
Gain complete operational confidence with your edge implementation by gaining visibility into all your deployed agents.

Built-in data provenance
Only product in the market to offer out-of-the-box data lineage tracking and provenance on data in motion.

Cloudera Edge Management with IBM

Cloudera Edge Management with IBM (CEM) is an edge management solution made up of edge agents and an edge management hub. It manages, controls and monitors edge agents to collect data from edge devices and push intelligence back to the edge. CEM allows you to develop, deploy, run and monitor edge flow apps on thousands of edge devices. CEM is a key part of Cloudera DataFlow with IBM (CDF) a comprehensive edge-to-enterprise streaming data platform that addresses the key data management challenges with streaming and IoT data for all types of enterprises. Apache MiNiFi is a lightweight edge agent that implements the core features of Apache NiFi, focusing on data collection and processing at the edge. Edge Flow Manager (EFM) is an agent management hub that supports a graphical flow-based programming model to develop, deploy and monitor edge flows on thousands of MiNiFi agents.

Edge data collection and management challenges

The key challenges that enterprises are facing today in IoT are about edge data collection and edge management. These challenges can be addressed easily with Cloudera Edge Management with IBM. Some of these challenges are:

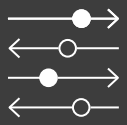
- Lack of tooling to collect and process data at the edge.
- Expensive to move data from edge to cloud.
- Building edge data collection / IoT apps require lots of coding that’s time-consuming.
- Managing apps on thousands of edge points is a complex problem.
- No tooling to monitor thousands of applications running on the edge.

Cloudera data-flow platform with IBM

Edge management

Edge data collection, routing and monitoring

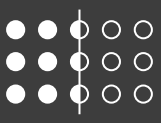
- MiNiFi
- Edge flow manager
- NiFi registry



Flow management

Enterprise data ingestion, transformation and enrichment

- Apache NiFi
- NiFi registry



Stream processing

Real-time stream processing at IoT scale

- Apache Kafka
- Schema registry

Stream management

- Streams messaging manager
- Streams replication manager

Stream analytics

Predictive analytics and real-time insights

- Kafka streams
- Apache flink
- Spark streaming

Enterprise services

Provisining, management and monitoring

- Unified security
- Edge-to-enterprise governance
- Single sign-on

MiNiFi

MiNiFi is a sub-project of Apache NiFi and was started a few years ago. Originally built out of a strong need to have the same kind of capabilities and strengths that NiFi has but at the edge, MiNiFi was created directly from NiFi by stripping out some core enterprise features like the user interface. The intent was to keep the edge agent as lightweight as possible. Today, with MiNiFi available in two flavors, Java and C++, it can be embedded inside any small edge device like a sensor or Raspberry Pi.

Key features of MiNiFi

Some of the key capabilities and features of MiNiFi are:

- Lightweight and portable, C++ and Java agents
- Guaranteed delivery
- Data buffering
- Prioritized queuing
- Flow Specific QoS (latency v throughput, loss tolerance, etc.)
- Data provenance
- Extensible architecture
- Site-to-site communication protocol
- TensorFlow support

Cloudera Edge Flow Manager (EFM)

EFM is an agent management hub that provides a graphical user interface (GUI) for designing, deploying and monitoring edge flow applications on thousands of MiNiFi agents. It also acts as the single management and monitoring layer for all the MiNiFi agents deployed in the field. EFM provides three key capabilities to the edge flow lifecycle:

Flow authorship
Cloudera Edge Flow Manager addresses the challenge of developing IoT applications by offering a code-free drag-and-drop development environment. This development environment offers a NiFi-like experience for capturing, filtering, transforming and transferring data from edge agents to upstream enterprise systems like CDH. It also comes with hundreds of pre-built processors to make the development much easier.

Flow deployment
Managing the deployment of IoT applications has been an industry challenge. EFM alleviates this challenge by offering a simple, yet powerful, model for deploying applications to agents. Agents registered with EFM are notified when a new or modified application is available. The agents themselves acquire the updated application from EFM to avoid networking complications. Agents will update their local runtime with the application and send verification checks back to EFM once the operation has been verified to have completed successfully.

Flow monitoring
Agents in CEM send regularly scheduled heartbeats to their EFM instance. The heartbeat represents the most recent snapshot of the agents' runtime. EFM stores, analyzes, and renders these heartbeats to end users. The heartbeats allow operators to visualize detail such as flow throughput, connection depths, processors running and overall agent health. Visualizing this information allows the operations administrator to take appropriate actions where needed.

For more information

To learn more about Cloudera Edge Management with IBM, visit the IBM and Cloudera webpage or contact an IBM data management expert.

IBM Corporation
New Orchard Road
Armonk, NY 10504

Produced in the United States of America
December 2019