

Cloudera Edge Management Overview

Date published: 2019-04-15

Date modified: 2023-10-27

The Cloudera logo is displayed in a bold, orange, sans-serif font. The word "CLOUDERA" is written in all caps, with the letter 'E' stylized as a horizontal bar with a small triangle in the center.

Legal Notice

© Cloudera Inc. 2024. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

Cloudera Edge Management overview.....	4
---	----------

Cloudera Edge Management overview

Cloudera Edge Management (CEM) enables edge device managers, data engineers, and IoT administrators to gain control of the data from edge devices with realtime edge data collection and management. Get familiar with CEM, its components, key features, and use cases.

CEM is an edge management solution made up of edge agents and a management hub for the agents. The edge management hub manages, controls, and monitors agents to collect data from the edge devices and push intelligence back to those devices. CEM offers out-of-the-box data lineage tracking and provenance of data-in-motion.

For additional information about the main functionalities of CEM (like flow creation, authentication, authorization, monitoring, agent management and much more), check out the videos on the Cloudera Edge Management YouTube playlist:

Components

CEM consists of two components:

- 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.
- Apache MiNiFi is a light-weight edge agent that implements the core features of Apache NiFi, focusing on data collection, processing, and distribution at the edge. It can be embedded inside any small edge device like a sensor or Raspberry Pi. It is available in two flavors: Java and C++.

Capabilities

CEM provides three main capabilities to the edge flow lifecycle:

Flow creation

EFM 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 CDP Private Cloud Base or CDP Public Cloud.

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 flows to agents. Agents registered with EFM are notified when a new or modified flow is available. The agents access the flow and apply it locally.

Flow monitoring

Agents in CEM send periodic heartbeats to their EFM instance. The heartbeat contains information about the deployment and runtime metrics. EFM stores, analyzes, and renders these heartbeats to end users. The heartbeats enable operators to visualize details such as flow throughput, connection depths, processors running, and overall agent health.

Key features

- No-code drag-and-drop UI

Hundreds of pre-built processors are available to connect with a range of data sources, devices, and protocols to build sophisticated data flow pipelines.

- Edge management hub

You can ingest, capture, and deliver data in real-time from any streaming source, including clickstreams, social media, mobile, or IoT devices.

- Flow designer for edge flows

You can build edge dataflows visually through a NiFi-like user interface for edge data collection and processing.

- MiNiFi edge agents

Lightweight and portable C++ and Java agents with fine-grained data-lineage information generated constantly.

- Enterprise-grade security and data provenance

Robust options for authentication and authorization and out-of-the-box data lineage tracking and provenance on data-in-motion.

- Edge management and data collection

A custom-built dashboard enables edge management at scale with command, control, and monitoring of hundreds of thousands of agents with minimal footprint to collect, filter, and process data. You can deploy updates to thousands of edge agents at the same time.

Use cases

Predictive maintenance

A data-driven approach to analyze IoT and sensor data from connected equipment to effectively predict when and how an asset might fail, detect variances, understand warning signals, and quickly identify patterns that might indicate a potential breakdown in, for example, manufacturing or fleet management. CEM offers a simple and small footprint solution for data ingestion from connected assets to enhance predictive maintenance.

Patient monitoring

Biometric and telemetric devices are used in healthcare organizations to monitor post-surgery or high-risk patients. Ingesting sensor data from these devices about various patient vitals helps to detect abnormalities or concerning patterns. CEM helps to capture patient-monitoring data and deliver them to stream-processing engines for insights.

Data movement

Traditional ETL processes are for use cases where data must move from one database to another. Modern enterprises transfer data from on-premises to cloud or cloud-to-cloud, moving petabytes of information in a matter of just hours.

Log and metric collection

The agents can be used to collect any type of logs or metrics from the hosts where the agents are installed. They can also be used as log aggregators. This data can be filtered, enriched, processed locally before being sent to any destination. This type of deployment is very common for cybersecurity use cases.

Cloud native applications

The MiNiFi C++ agent is commonly used as a side-car container or as a DaemonSet to collect data, logs, metrics, and so on, from cloud native applications running on Kubernetes. It can also be used as a way to get data to the running applications. Besides, with EFM, it is possible to update the processing on the agents without redeploying the application.