StreamSets

BENEFITS

Low footprint agent at <5MB in size.

On-edge transformations/ analytics and bi-directional data movement.

Automated provisioning at scale.

Continuous performance management.

COMMON USE CASES

Internet of Things

Leverage SDC Edge to ingest data directly from sensor driven machinery and componentry. Scale to any number of connected devices without an IoT gateway.

Cybersecurity

Analyze network system and endpoint logs for improved threat detection by deploying SDC Edge as a lightweight agent (small attack surface).

StreamSets Data Collector Edge™

Build and operate dataflow pipelines on lightweight and nomadic systems

With the advent of huge numbers of digital systems enterprises need a new approach to how they ingest and transform data for initiatives such as the Internet of Things (IoT) and cybersecurity.
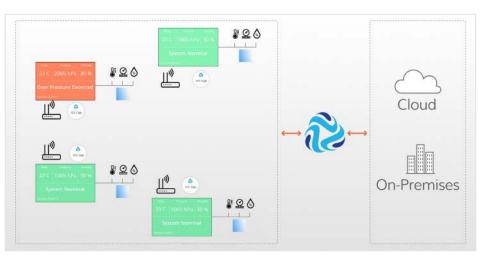
Lightweight sensor-enabled systems and nomadic personal devices present new challenges that are very different from those for systems that sit close to the data center. For one, the number of systems out on the edge is voluminous and comprises a variety of brands and versions. Also, these systems have limited memory and compute and can't host heavy, resource-intensive software. Lastly, they often have periods where they are disconnected yet still need to send data once back online.

These unique characteristics make it difficult to design, provision and manage edge data collection. Numerous device brands and versions means frequent changes that can break pipelines. Resource constraints eliminate most data integration software solutions from consideration and intermittent connections put data fidelity at risk.

StreamSets Data Collector Edge (SDC Edge) brings the edge within reach, delivering an ultralight yet powerful data ingestion solution for constrained systems. Developers can build robust ingest pipelines that are deployed in a small footprint, low resource utilization package. A drag-and-drop GUI makes it easy to prototype and deploy edge pipelines to huge populations of systems. Written in Go, SDC Edge provides a single solution across a broad range of edge hardware platforms.

SDC Edge pipelines address a range of functional requirements. They can act as a simple data forwarder, sending data to multiple destinations using pre-built connectors. Or they can be configured to perform transformations and analytics on edge. Data movement can even be bidirectional, for instance to flow control data or contextual data back to the edge device.

As part of the StreamSets Data Operations Platform, SDC Edge pipelines get automated provisioning and performance management at scale. With full lifecycle management, health monitoring and automated pipeline deployment you can efficiently support as many edge pipelines as you need.

IoT Dataflows using StreamSets Data Collector Edge

StreamSets

FEATURES

Lightweight agent runs anywhere

- Less than 5MB installation footprint.
- Low memory and CPU (1-2%) utilization.
- Single Go binary (no external dependencies).
- Apache licensed open source.

On-edge transformations and bi-directional dataflows

- Run multiple dataflow pipelines on any edge agent.
- At-least once and at-most once delivery guarantees.
- Structured and semi-structured data supported.

Complete operational control

- Drag-and-drop build of edge pipelines.
- Automate deployment and maintenance of pipelines at scale.
- Manage real-time pipeline and dataflow topology performance with StreamSets Dataflow Performance Manager™.

Platform and protocol agnostic

- Support for a broad range of communications protocols, including HTTP, MQTT, CoAP, and WebSockets.
- Runs on leading device platforms including OS X, Windows, Android and Linux.

The StreamSets Data Operations Platform is designed to simplify the entire dataflow lifecycle, including how to build, execute and operate enterprise dataflows at scale. Developers can design batch and streaming pipelines with a minimum of code, while operators can aggregate dataflows into topologies for centralized provisioning and performance management.

ABOUT

StreamSets is headquartered in San Francisco. Our mission is to help enterprises harness their data in motion. StreamSets software is in use at hundreds of organizations and we're backed by top-tier Silicon Valley venture capital firms, including Accel Partners, Battery Ventures, Ignition Partners, and New Enterprise Associates.

LEARN MORE

Get up and running with StreamSets in minutes. Visit us at:

www.streamsets.com