



## PARTNER BRIEF

# Accelerate the Pace of Innovation With Snowflake Data Cloud and Snowpark

StreamSets provides secure and reliable enterprise data integration for companies who need to integrate corporate data assets with Snowflake for modern data & analytics initiatives.

**Traditional data pipelines break when the unexpected happens, holding up business reporting and analytics initiatives while teams isolate, refactor, and redeploy traditional data pipelines.**

StreamSets intelligent pipelines are resilient and adaptable to change, ensuring continuous, secure data ingestion into Snowflake from on-premise and cloud environments. Organizations can do more with their enterprise data and easily leverage StreamSets across reporting, operational analytics, data science, data sharing initiatives and more.

## Accelerate Snowflake Use Cases



### Migrate to Data Cloud

With StreamSets Data Collector, you can bridge the gap between new and legacy environments easily and securely. With a data “mission control” across all your environments, you can easily move between the Data Cloud and on-premises. StreamSets intelligent pipelines are unique for hybrid data integration environments and automatically adapt to and continue operating even when data formats and schemas change (Data Drift).



### Operational/Advanced Analytics & Reporting with Snowflake

StreamSets helps companies do more with Data Cloud: financial regulatory reporting and ESG, enhance customer experiences, generate greater marketing and advertising

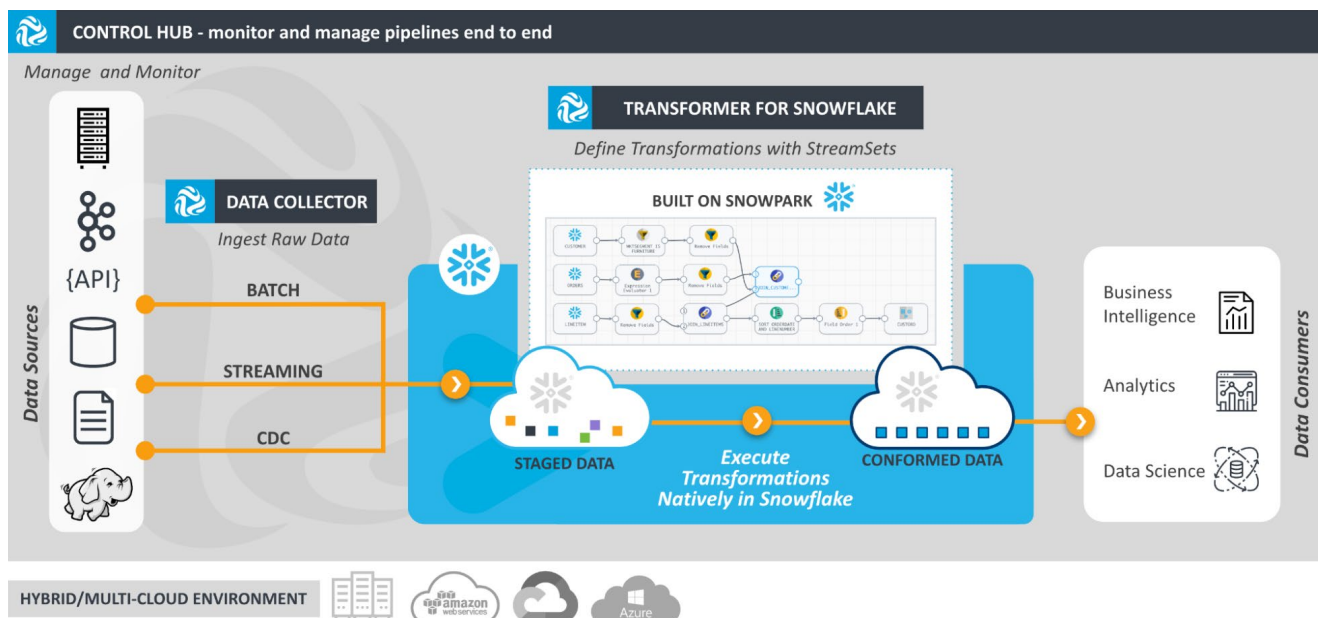
## StreamSets and Snowflake are Partnered to Accelerate Reliable Data Integration to the Snowflake Data Cloud

StreamSets is proud to be a Premier Technology Partner with Snowflake and has completed a 3rd party technical validation to confirm integrations with Snowflake are optimized with an emphasis on functional and performance best practices. StreamSets Transformer for Snowflake is built on Snowpark to enable both the expressiveness and flexibility of Snowpark's multi-language support, as well as the simplicity of data cloud operations.



## Reference Architecture

Integrating data from various sources to Snowflake Data Cloud



ROI, fuel machine learning, operational/advanced analytics, and drive overall long-term business growth.

With StreamSets Transformer for Snowflake, data analysts and scientists can easily design complex data transformation jobs which can also incorporate custom functions. Because Transformer for Snowflake is built on Snowpark, all the data processing is done natively and securely in Snowflake eliminating the need for an external, complex Spark environment.



### Change Data Capture (CDC) to Snowflake

It's important to keep data fresh and updated through ongoing change data capture operations. With StreamSets, this is as easy as leveraging one of our prebuilt CDC origins. You can ensure that data from your core systems is continuously replicated into your Data Cloud platform. StreamSets provides out-of-the-box CDC enabled sources to easily develop and automate Change Data Capture (CDC) operations.

## StreamSets Reference Architecture for Snowflake:

### Overview

StreamSets data integration platform has 3 core capabilities: Control Hub monitors and manages StreamSets data pipelines end to end. Data Collector provides secure connectivity to enterprise data sources (on-premise or in the cloud) and data ingestion into Snowflake. Transformer for Snowflake transforms, joins, and cleanses data in Snowflake from raw staging tables to a final conformed state so it can be used for reporting, analytics and data science initiatives.

### Capabilities & Benefits

#### Build All Your Enterprise Data Pipelines in One Place

Get started quickly with an intuitive, easy to learn interface. StreamSets has out-of-the box connectivity for all your databases, messaging queues, legacy data stores and operational systems (including mainframes). Create batch, CDC and streaming pipelines with all data types (structured, semistructured and unstructured). StreamSets engines are secure and designed for hybrid environments. Build all your pipelines in one place. Just hit run.

### **Intelligent, Resilient and Reusable Data Ingestion Pipelines**

Drag and drop over 50 pre-built processors into your pipelines to intelligently format, shape and enrich your data on its way into Snowflake. StreamSets will automatically identify and adapt to Data Drift (reducing pipeline breakages by 80%). Create templates and pipeline fragments easily deploy common logic and pipelines 100's of times over with just a few lines of code. You can trust that your enterprise data is analytics ready as it is ingested into Snowflake

### **Design & Execute Complex Data Transformations Natively in Snowflake**

Build data transformation jobs of any complexity. Incorporate your custom Python, Java or Scala libraries and have it all executed natively in Snowflake through Snowpark. Save time and money - no need to extract your data for additional processing in complex Spark environments.

## **About StreamSets**

StreamSets, a Software AG company, eliminates data integration friction in complex hybrid and multi-cloud environments to keep pace with need-it-now business data demands. Our platform lets data teams unlock data—without ceding control—to enable a data-driven enterprise. Resilient and repeatable pipelines deliver analytics-ready data that improve real-time decision-making and reduce the costs and risks associated with data flow across an organization. That's why the largest companies in the world trust StreamSets to power millions of data pipelines for modern analytics, smart applications, and hybrid integration.

To learn more, visit [www.streamsets.com](https://www.streamsets.com) and follow us on [LinkedIn](#).