

# ROBUST END-TO-END DATA PIPELINES

## Auto-ingestion, change data capture, and in-stream data processing for up-to-the-moment insight



### CONTINUOUS DATA PIPELINES

Simplify ingesting and processing of data across a broad range of use cases. Move data within a Snowflake environment continuously, without the need for scripts or custom code.



### STREAM DATA FROM APACHE KAFKA™

Support event-driven architectures by ingesting JSON or Avro messages from Kafka using a Snowflake native connector.



### AUTO-INGEST USING SNOWPIPE

Automatically stream files from Amazon Web Services (AWS), Microsoft Azure, or Google Cloud Platform (GCP) object stores as new data arrives. Parse and insert files using Snowpipe without separate scheduling or orchestration.



### IMPLEMENT CHANGE DATA CAPTURE

Capture changes made to the underlying table via Snowflake table streams. Enable change data capture (CDC) use cases for continuously loading or changing tables, regardless of data scale.



### BUILD ROBUST DATA TRANSFORMATIONS

Schedule workloads to execute data transformations when needed. For use in more complex event-centric pipelines, trigger tasks that are based on outcomes of other tasks.



### SUPPORT ALL YOUR USERS

Easily scale and allocate resources to different workgroups without data or resource contention. Take advantage of Snowflake's instant and infinite scaling of independent workloads for any level of concurrency without impacting performance.

As data architectures evolve, so do the requirements for how data is loaded and processed within a data warehouse. Rapid and immediate access to live data, and the ability to analyze it in the moment, are key requirements so the data-driven enterprise can streamline operations, better serve customers, and uncover new market opportunities. Data pipelines in Snowflake can be batch or continuous, and processing can happen directly within Snowflake itself. Thanks to Snowflake's multi-cluster compute approach, these pipelines can handle complex transformations, without impacting the performance of other workloads.

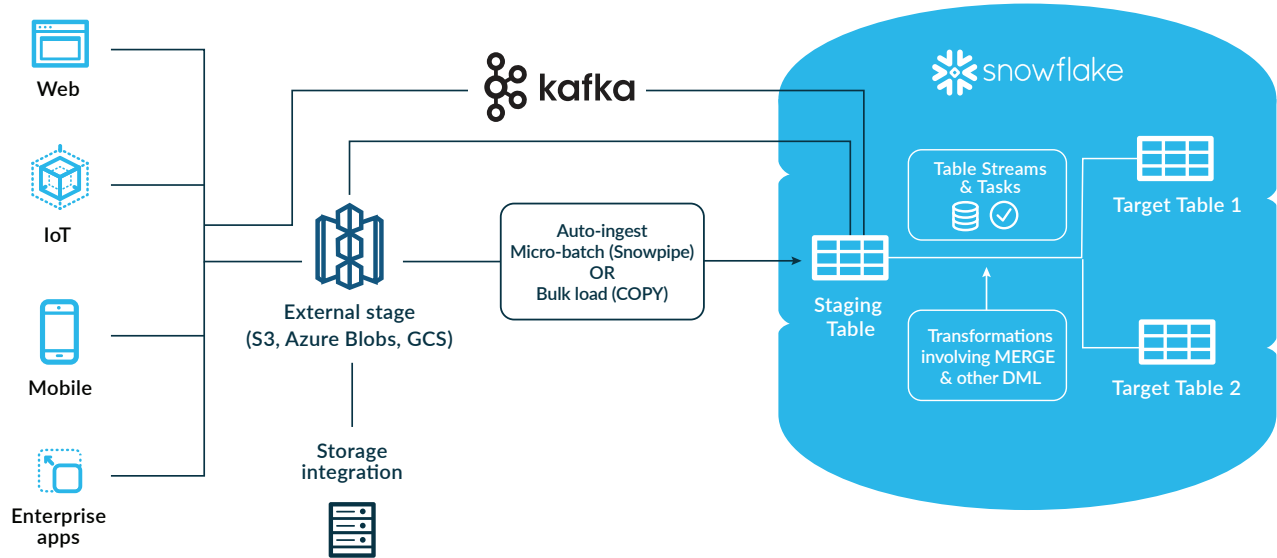
### AUTOMATIC INGESTION AND CONTINUOUS DATA PROCESSING

Snowflake first introduced Snowpipe to simplify the process of moving files from object stores into Snowflake. Now Snowpipe can automatically ingest data, so loading data happens continuously as new data arrives in an AWS, Azure, or Google object store, eliminating the need for scripts and scheduling tools. Snowflake also offers an Apache Kafka™ consumer connector to easily ingest event streams from Kafka.

As data lands in Snowflake, table streams capture changes that have occurred since the last load. Table streams support change data capture (CDC) use cases, which is particularly helpful as data warehouses grow. Tasks in Snowflake trigger processing without the need for complex scheduling routines.

Together, Snowpipe auto-ingest, table streams, and tasks enable the continuous movement of data within a Snowflake cloud-built data warehouse. Delivered as an easy-to-use data warehouse as a service, Snowflake enables you to process and analyze all your diverse data, build multiple databases, query with a common robust ANSI SQL environment, and execute ACID transactional capabilities.

## Streams and Tasks in Snowflake



*Ingesting structured and unstructured data from Kafka and a data lake into Snowflake.*

### CUSTOMER SUCCESS



Blackboard Inc. provides educational institutions with data about their academic programs and student performance. Blackboard uses Snowflake for its analytics products such as Blackboard Predict, which focuses on student retention by using data to accurately and proactively identify students at risk and help them succeed with targeted interventions. Blackboard also uses Snowflake to understand system usage in order to prioritize features and continuously improve its products.

Snowflake's Snowpipe, streams and tasks, and stored procedures enabled Blackboard to build an ingestion platform for most data pipelines hydrating Blackboard's data lake. These pipelines serve more than a thousand sites with hundreds of tables per site. This results in a significant reduction of Blackboard's infrastructure management and costs. It also delivers a streamlined architecture with less complexity and handoff points.

### ABOUT SNOWFLAKE

Snowflake is the only data warehouse built for the cloud, enabling the data-driven enterprise with instant elasticity, secure data sharing, and per-second pricing across multiple clouds. Snowflake combines the power of data warehousing, the flexibility of big data platforms, and the elasticity of the cloud at a fraction of the cost of traditional solutions. Snowflake: Your data, no limits. Find out more at [snowflake.com](https://www.snowflake.com).