

# Datacoral: Database replication using CDC

No-code data replication, with data quality guarantees

## OVERVIEW

Organizations today deeply understand the need for comprehensive data analytics. Data continues to be siloed in many different data sources (databases, APIs, file systems) and the volume and velocity of data continue to grow rapidly. The solution that data teams are searching for is a fast, scalable cloud warehouse combined with an easy-to-use ETL solution that helps them reliably centralize data from data sources such as databases.

## CHALLENGE

Centralizing an organization's data from different data sources, especially databases, into a warehouse is a frustrating experience. It involves piecing together multiple complex systems, while dealing with increasing variety, velocity and volumes of data. Change Data Capture (CDC) is the recommended way to replicate data from large databases but companies need to think about the unique challenges posed by CDC - replication lag, data quality, historical syncs and schema changes.

## SOLUTION

Snowflake combined with a cloud-native, ETL solution like Datacoral is the modern answer to today's analytics challenges.

- **End-to-end ETL** - Easy to deploy integrations, managed over their lifecycle. Data and schema reliably replicated into Snowflake.
- **Source system support** - Multiple types of sources (databases, APIs, etc) and complex CDC sources should work out-of-the-box.
- **Complete observability** - Data quality, monitoring and alerting are critical parts of ETL, and not an afterthought.

## BENEFITS



### No-code CDC Connectors

Get data flowing from databases such as PostgreSQL, MySQL, MongoDB into Snowflake with just a few clicks.



### Real-time Data Replication

Datacoral reads from database logs and writes to Snowflake with minimal replication lag so data teams always have complete and up-to-date data.



### Data Quality Checks

Out-of-the-box monitoring sends out alerts on failures. Regular, automated data quality checks make sure that data matches between source and destination at all times.



### Schema Change Handling

As tables and columns are added, removed or modified in your source database, Datacoral ensures that your tables in Snowflake remain in sync.



### Fast Historical Syncs

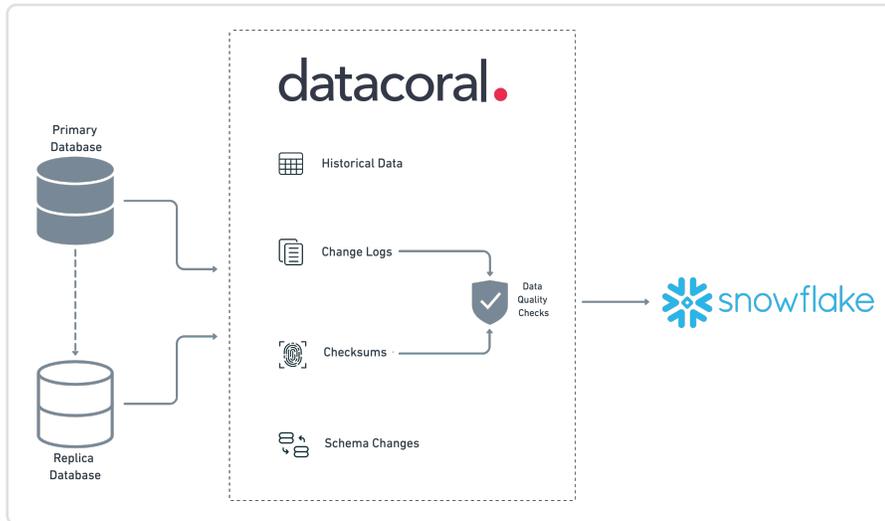
Historical syncs are performed by reading from replicas of your database. This leads to a faster initial sync and reduces the load on your production DB.



### Data Security

Datacoral is deployed in your cloud VPC which means that no data leaves your system. Data is encrypted at-rest and in-transit and all actions are audited.

## REFERENCE ARCHITECTURE



## How it works

1. Automated schema mapping from source DB to Snowflake
2. Automated historical syncs read from the replica database
3. Change logs read from database and applied to Snowflake tables
4. Automated data quality checks run regularly, comparing source to destination.
5. Schema changes are detected as they happen and applied to Snowflake

## “ flexport.

“We chose Datacoral's Change Data Capture (CDC) connector to replicate our PostgreSQL databases to Snowflake after evaluating multiple SaaS and open-source solutions.

Datacoral's fully managed pipeline reduces operational burden on our DevOps Team and our Security Team's stringent requirements were satisfied with the entire product deployed on-prem in our cloud.”

- Stephan Goergen, Data Engineer at Flexport

## GET STARTED TODAY

**Free 30 day trial:** Sign up at <https://app.datacoral.io/sign-up>

**Datacoral demo:** Fill the form at <https://datacoral.com/request-info/>

**Have questions?** Email us at [hello@datacoral.co](mailto:hello@datacoral.co)

## ABOUT DATACORAL

Datacoral offers a no-code data integration platform for analytics. It provides over 80 connectors for APIs, databases, file systems and event streams to ingest data into data warehouses. Datacoral's customers include data-driven companies such as Greenhouse, Frame.io, Betterment and Flexport who trust Datacoral for their data integration and data pipeline needs. Datacoral is based in San Francisco and is an AWS Advanced Technology Partner with AWS Data & Analytics Competency.