



## CASE STUDY

# Serverless ETL in a Cloud Data Warehouse

### Just the Headlines

Short on time? Here are the key facts.

- An agricultural technology client needed to automatically schedule ETL jobs to run from AWS through the layers of their data warehouse without allocating resources to manually trigger the load.
- 2nd Watch leveraged their ETL framework and best practices to run ETL jobs using Google Cloud Composer.
- The resulting serverless ETL in a cloud data warehouse has automated previously manual processes, saving the company time and money.

### The Challenge

2nd Watch's client wanted to move their data, housed in AWS, into a data warehouse on Google Cloud Platform. In doing this, they needed a way to automatically schedule the ETL jobs to run from AWS through the layers of data warehouse without allocating resources to manually trigger the load.

#### Industry

Agricultural  
Technology



#### Featured Technologies

Airflow

AWS

Cloud Composer

Google Cloud Platform

## The Solution

To alleviate this problem, 2nd Watch leveraged their ETL framework and best practices to run ETL jobs using Google Cloud Composer, a managed version of Apache Airflow. 2nd Watch created Python scripts that are run daily at scheduled times, using Composer, to transform tables from AWS and load them into the cloud.

## The Outcome

The serverless ETL solution with Google Cloud Composer manages the daily raw loads and business logic loads for the client, so employees do not need to spend their time manually triggering the jobs. This ensures that the correct data will be available in a cloud-based data warehouse, saves the company time and money, and allows them to expand their Google Cloud platform capabilities.