



# AWS SERVICES

PROJECT

*Szabó Bence*

DATA  
ENGINEERING



# PROJECT OVERVIEW

## 01 BUSINESS CASE

Defining the business case that the project can serve.

## 02 DATA SOURCES

Identifying the integration of static historical datasets from S3 with real-time market feeds from an external API

## 03 PIPELINE

Architecting the automated serverless workflow that ingests, transforms, and warehouses data for analysis.

## 04 SERVICE COST

Calculating the projected monthly infrastructure expenditure to demonstrate economic viability and budget compliance.



01



Articulating the strategic imperative to merge historical market data with real-time sentiment for enhanced investment decision-making.



*Leverage NLP sentiment analysis to optimize trading strategies and maximize portfolio returns.*



*Quantitative Analysts and Portfolio Managers requiring automated, low-latency trading signals.*



*Scalable analytics engine capable of synthesizing historical and live data for real-time consumption.*

# HYBRID DATA INGESTION



## Internal source

- **Source:** S3 data lake (historical stock data from yfinance)
- **Nature:** Static, Structured, High Volume (5+ years)
- **Role:** Provides the "baseline" model for training and long-term trend analysis
- **Format:** csv

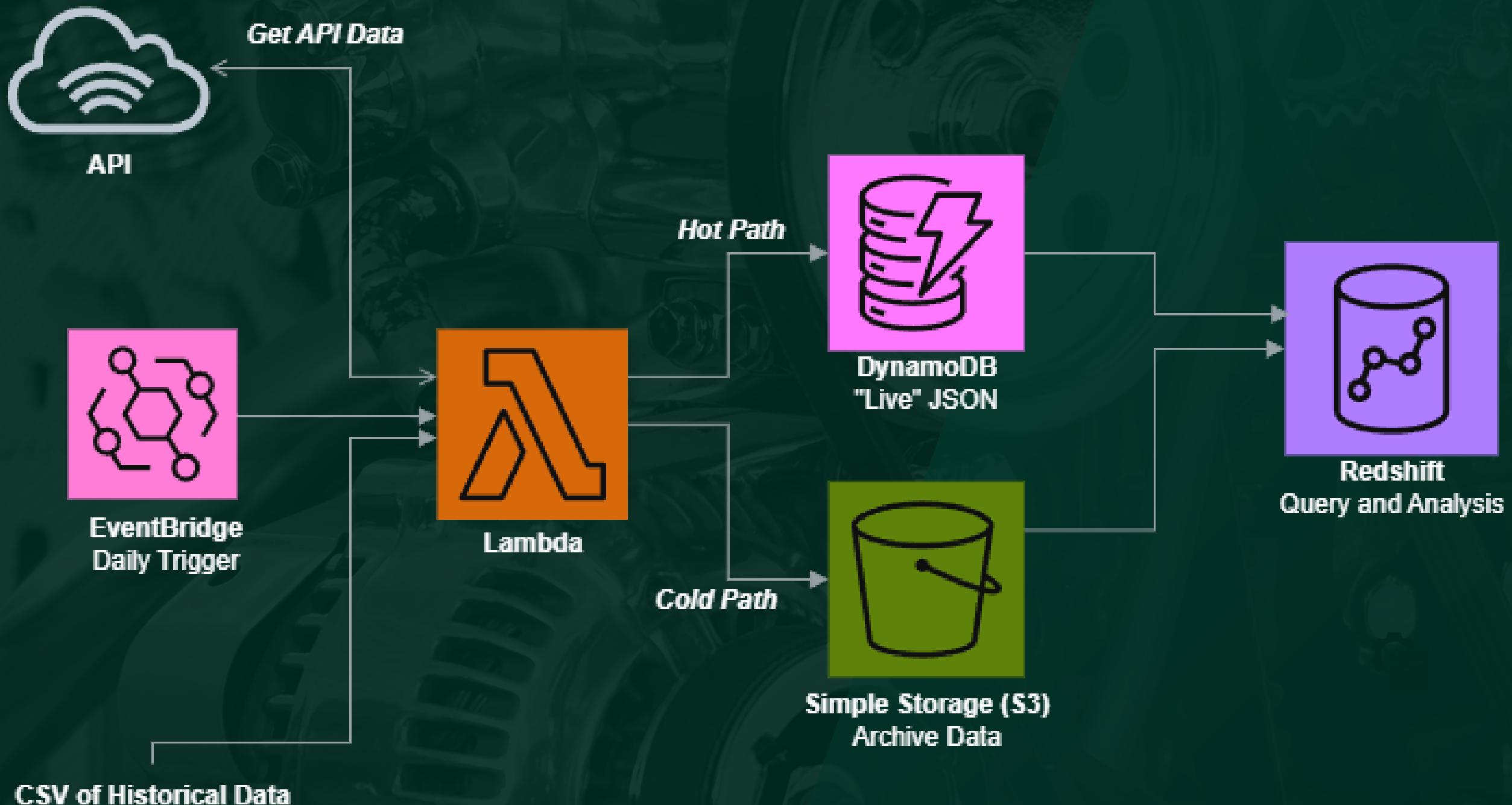


## External Source

- **Source:** Alpha Vantage API (REST Endpoint)
- **Nature:** Dynamic, Semi-structured, Low Latency (~Real-time)
- **Role:** Provides the "live pulse" to compare against historical baselines
- **Format:** JSON

03

## PIPELINE



04



# MONTHLY BILLING OF SERVICES

## AWS LAMBDA

Running the code to  
fetch stock prices  
from the API

0.00 USD

## SIMPLE STORAGE

Running the code to  
fetch stock prices  
from the API

0.01 USD

## DYNAMODB

Storing the daily  
"live" data for fast  
access

0.25 USD

## REDSHIFT

Running the code to  
fetch stock prices  
from the API

9.19 USD





THANK YOU  
FOR THE ATTENTION

THANK YOU