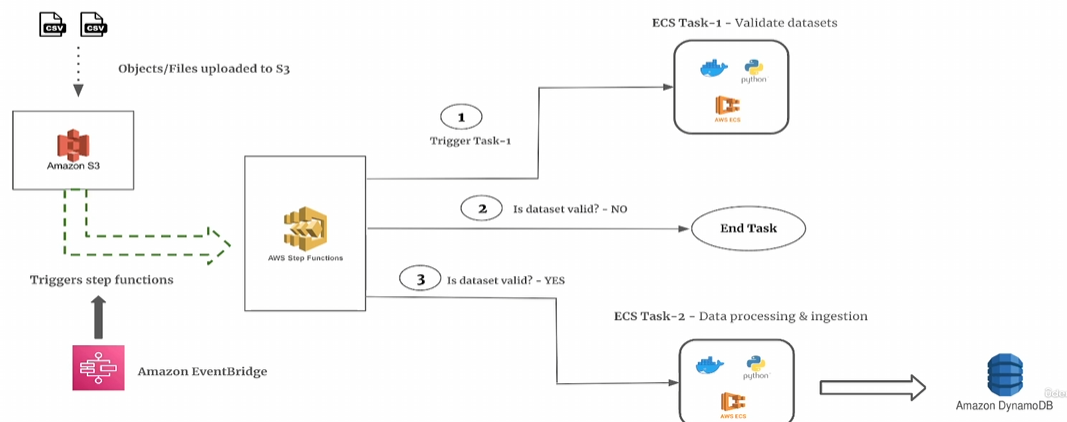
# Technical Report: E-commerce Data Processing & Lakehouse Implementation

This report provides a detailed breakdown of the architecture and implementation of our event-driven data processing pipeline using AWS ECS, Step Functions, and a Lakehouse architecture using Spark and S3.

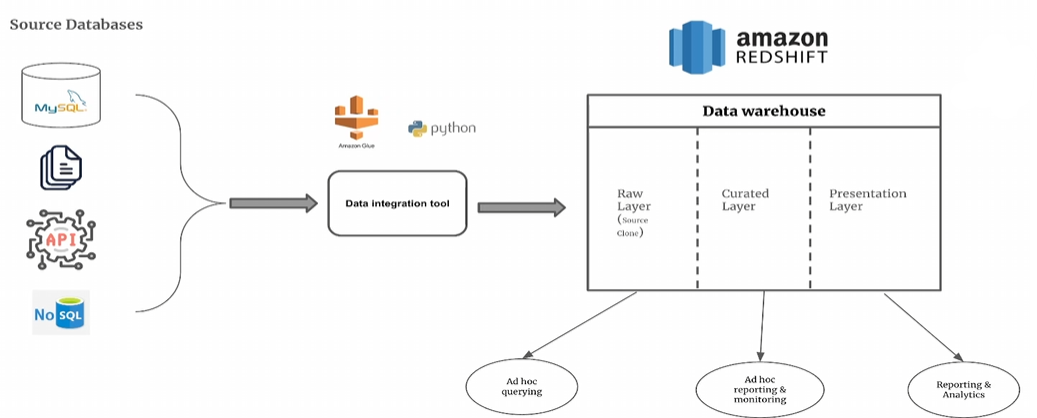
## 1. Event-Driven Data Processing with AWS ECS & Step Functions

We use AWS EventBridge to trigger workflows when new data arrives in Amazon S3. The workflow includes data validation, transformation, and storage in DynamoDB.



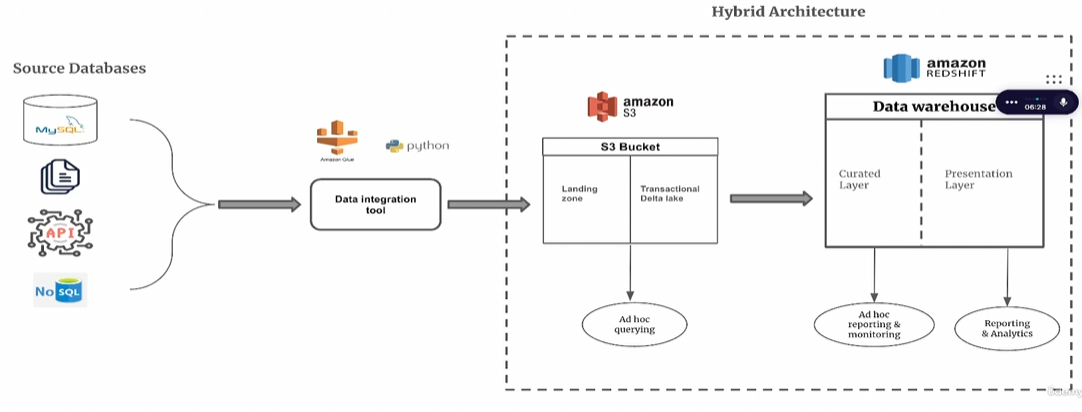
## 2. Lakehouse Architecture Using PySpark & S3

Our data lakehouse is built using Amazon S3, AWS Glue, and PySpark with Delta Tables. This enables schema evolution, time travel, and ACID transactions.



## 3. Querying Data with Athena & Redshift

Processed data is accessible via AWS Athena and Redshift for analytical workloads. We use Glue Data Catalog for seamless integration.



For more technical details, refer to our GitHub repository or internal documentation.