



Data and Analytics Platform leveraging Lambda, Amazon Glue, Redshift,

A StepFunction for a Global Pharmaceutical Major

A leading US pharmaceutical firm aimed to shift from on-premises infrastructure to a centralized cloud data ecosystem. Discover how Tiger Analytics established a strong data foundation, automated provisioning with AWS Lambda, and modernized their data environment. Reduced provisioning time to a day, enabled data asset reuse, standardized structures for quicker projects, and achieved cost savings with AWS Lambda serverless model.

The Background

Our client, a prominent global pharmaceutical firm headquartered in the United States, specializes in pioneering research, development, and the commercialization of innovative medicines, vaccines, and animal health products.

Currently relying on an on-premises infrastructure, the client faces a prolonged process, often spanning weeks, for infrastructure setup. As part of their strategic vision, the client aspired to establish a centralized, resilient, and standardized cloud data ecosystem



Key Challenges

AWS cloud was used but use cases were running in Silos.



Lack of standardized data engineering practices.



Need to define the process from scratch



Our Solutions



Established Data Foundation

Evaluated basic requirements, reviewed existing data and analytics landscape – tools, frameworks, use cases, etc., selected appropriate tools and services, etc.

Leveraged Automated DataOps Framework

Using AWS Cloud Services Developed a solution for automated provisioning, Setup of the Unified Control Environment, Validation and testing of the solution.

Processed big data using AWS stack

Data from different sources such as BigQuery, Redshift, CSV files were uploaded to S3(the Landing zone), with the data processed using AWS Glue and then loaded into the Processed/Published layer. Serverless Glue was considered to provide fault tolerance during disaster recovery. Once the data was available in the Processed layer, multiple Glue jobs were used to apply the necessary business transformations such as Aggregation of the KPIs, Data Harmonization/Standardization, Generation of necessary Features etc., post which the transformed data was copied into the Published layer.

AWS Lambda

Lambda is primarily used for triggering the ETL data loads. The triggering of lambda is event driven. These events include the on demand triggers like S3 upload and scheduled Event Bridge notifications. Lambda is also used as an intermediary step inside a step function to ensure seamless integration between different components of step function

Automation

Deployment of code is being done by using Jenkins with Terraform as IAAS tool

Tech Stack

/ Lambda

/ Glue

/ AWS Redshift

/ AWS S3

/ Step Function

/ SNS

/ Event Bridge

Value Delivered



- **Reduced time** for provisioning environments, services, roles across various regions, markets, and verticals from 1 month to 1 day
- **Built a robust framework** with a centralized data lake using Amazon S3 & AWS Redshift to enable existing data assets to be repurposed for other projects.
- **Standardized data structures** along with role enabled access allowed faster onboarding and turnaround time for new projects.
- **Serverless architecture** of AWS Redshift led to cost savings by taking advantage of a “Pay as You Go” model for infrastructure.
- **AWS lambda** is a crucial component which is being used at many places to achieve this **end-to-end automated workflow**.

About Tiger Analytics

Tiger Analytics is a global leader in AI and analytics, helping Fortune 1000 companies solve their toughest challenges. We offer full-stack AI and analytics services & solutions to help businesses achieve real outcomes and value at scale. We are on a mission to push the boundaries of what AI and analytics can do to help enterprises navigate uncertainty and move forward decisively. Our purpose is to **provide certainty to shape a better tomorrow**.

Being a recipient of multiple industry awards and recognitions, we have 4500+ technologists and consultants, working from multiple cities in 5 continents.

www.tigeranalytics.com

US | UK | Canada | India | Singapore | Australia