Data lake setup aiding rapid insights with regulatory compliance

Customer: A large US based medical equipment company

Summary

The customer is a leading US-based medical equipment company catering mainly to cloud-connected medical devices that transform care for people with sleep apnea, COPD and other chronic diseases. They are looking at integrating their MyApp application's data to MosaIQ Data Lake platform on AWS cloud. MyApp is a self-monitoring sleep therapy progress application used extensively by medical representatives and caregivers.

About Customer

The customer is one of the top medical equipment companies based in San Diego, California. They primarily provide cloud-connectable medical devices for the treatment of sleep apnea, chronic obstructive pulmonary disease (COPD) and other respiratory conditions. It employs more than 7,500 employees worldwide with a presence in more than 120 countries globally that have manufacturing facilities in Australia, France, Singapore and the United States.

Problem Statement

MyApp is the customer's patient self-monitoring application that helps track patient's sleep therapy progress both online as well as on smartphones. MyApp facilitates tailored coaching and handy tips to make therapy more comfortable. The Customer wanted to,

- To integrate MyApp application data to MosaIQ Data Lake platform on AWS.
- Reuse and replicate data flow of AirView, inclusive of policy, pseudo rules, de-identification, Protected Health Information (PHI) and non-PHI.
- Build code for data staging, data transformations for regulatory adherence and storage on AWS Simple Storage Service (S3).

Proposed Solution

Powerup to analyze and define the scope of integration. Obtain complete access to AWS development, system integration test and production setups and create AWS services catering to Virtual Private Network (VPC)s, subnets, route tables and Internet gateways. Define fixed and incremental S3 buckets for PHI as well as non-PHI accounts.

Ensure that a detailed definition of MyApp S3 policies including source connections and scheduling is made available before coding in the development environment. Also, freeze all policies and pseudo rules for PHI and non-PHI data encryption until coding completion and migration to test environment.

Implement Data Migration Service (DMS) to migrate data from on-prem to AWS cloud storage S3. Data with all the files to be pushed inside a single folder per table in the S3 bucket via lambda functions. CDC to be implemented for incremental data transfer to S3 event which in turn will trigger and push the requests to Amazon Simple Queue Service (SQS).

Leverage Fargate containers to run scripts in order to check data against the IDs. Run Electronic Medical Records (EMR) cluster by applying masking logic to this data which is sent for further analytics. Identify and save the same in S3 buckets. The next step is to create a test strategy for unit and integration tests.

Powerup DevOps to configure Complement Fixation Test (CFT) and implement continuous integration and continuous deployment (CI/CD) process for MyApp migration. Create integration test scripts, test CI/CD process before the actual system integration migration (SIT), prepare migration to development and UAT environments and devise automation.

The next task is to migrate to SIT through Ci/CD to validate all the resources and execute full load and schedule trigger for CDC load before moving to production deployment. Repeat the process in the production environment and perform UAT.

Post the integration, Powerup took up the responsibility of architectural assessment and went ahead with the Well-Architected Review (WAR) framework. WAR is an architectural assessment based on AWS framework that is built on five pillars – operational efficiency, reliability, security, performance efficiency and cost optimization.
Powerup identified the workload to be reviewed and once relevant data were identified, reviews were arranged with the stakeholders at the company. Review could be conducted onsite or remotely. A report aligning with AWS best practices, categorized as critical, needs improvement or meets best practices were generated for the selected workload. The report highlights the priority with which remediation should be carried out.
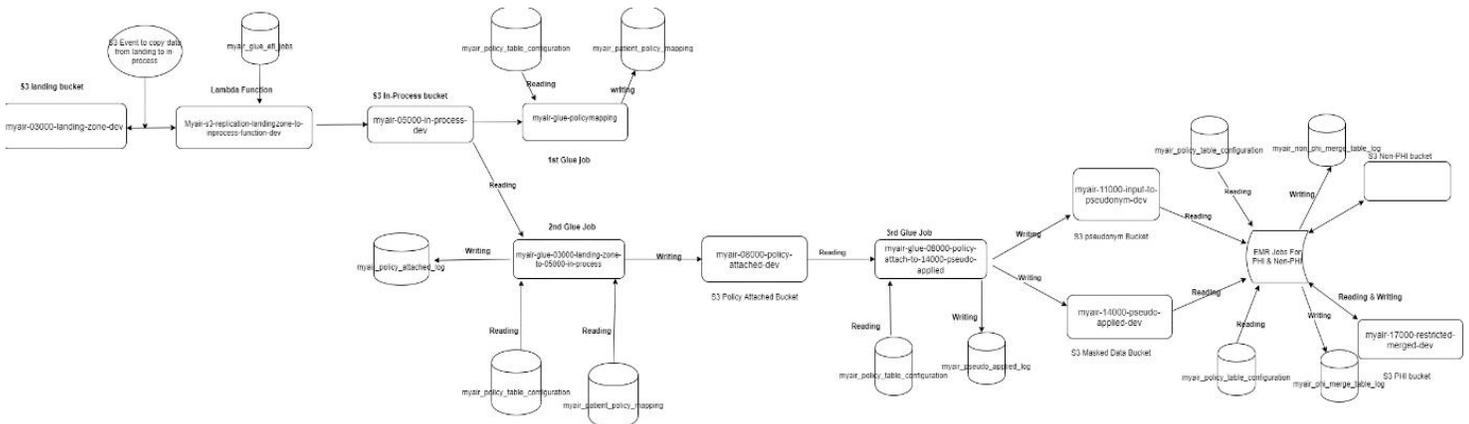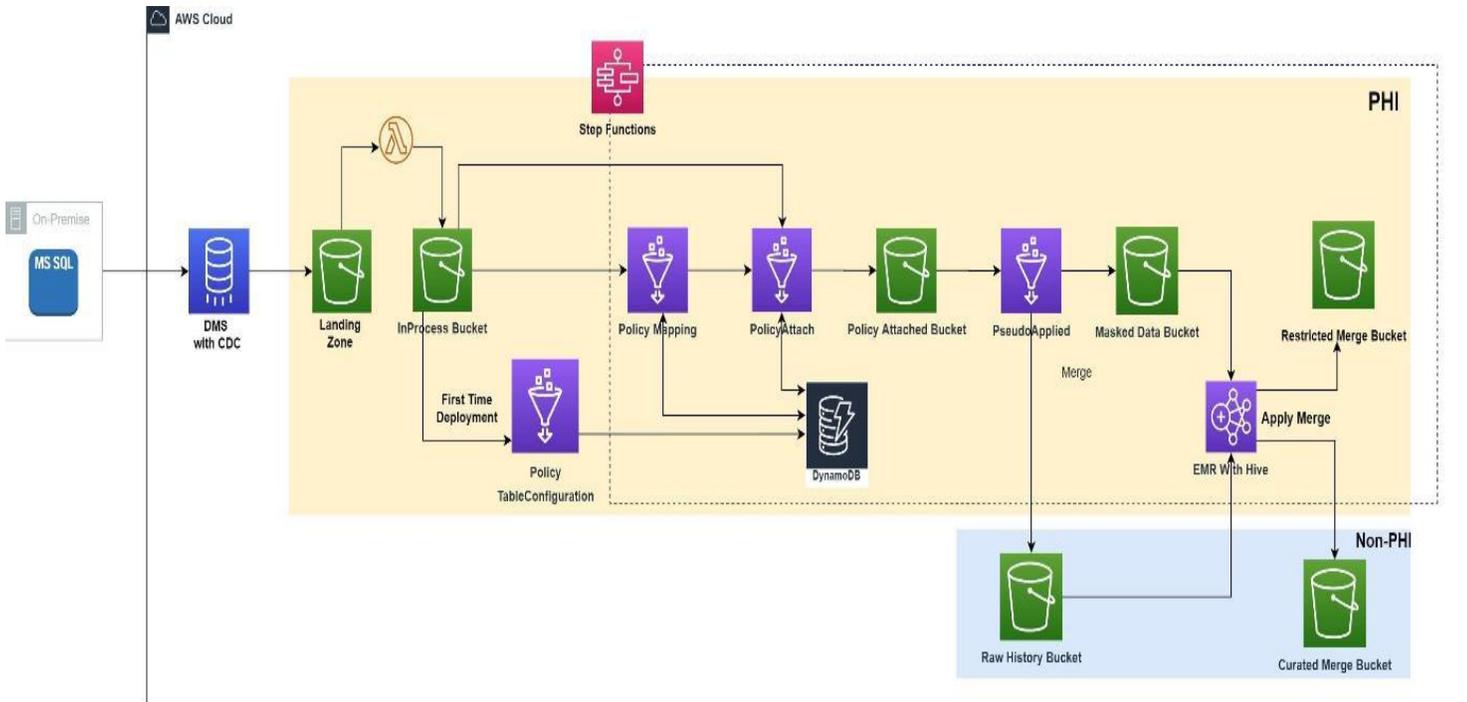
Benefits

MyApp application data has been integrated to MosaIQ on AWS cloud successfully. This platform can now provide capabilities to wider business team communities as MosaIQ is a data lake platform built on top of AWS stack and stores structure and unstructured data in raw format. It assists in the rapid discovery of actionable insights to improve patient care and business outcomes while maintaining security and regulatory compliance.

MosaIQ platform allows analytics, engineers, and data scientists to respond more efficiently and provide timely information to support better business decisions. This is mainly because data segregation is more organized and bifurcated for PHI and non-PHI data.

Reusable design from MyApp integration can be utilized for similar use cases across the company. A significant improvement in performance was noticed due to features like scalability and reduction of in-memory processing.

Cloud platform

AWS.

Technologies used

AWS S3, Lambda, AWS Glue, AWS EMR, AWS DynamoDB, AWS Step Function, AWS CloudFormation, AWS DMS + CDC.