



CASE STUDY



Title: Modernization of Big Data Platform

Client: Infosys

Location: Providence, RI

Industry: Banking and Finance

Duration: 12 Months

Challenges

One of Americas leading bank had an op-premises Data Lake built on Data Insights which was nearing its end-of-life support. On premises data lake need to be migrated and decommissioned quickly before its license expires and ensure all the data and processes are migrated along with enabling next-gen capabilities.

Some of the key challenges associated with this transformation were:

- Scalability / Stability – On premises data lake had issues with processing huge volume of data
- High cost of maintenance with dependency on IT systems
- Data in silos results in quality issues and getting actionable insights

Company has decided to modernize and migrate the data lake and Hadoop cluster to AWS platform and leverage agility and scalability

Key Specs

Migration Type: Rearchitect / Modernize

Technology Stack: Data Lake, Data insights, Hadoop

Modernization Method: Rearchitect and utilize cloud native tools

AWS Services Utilized: EMR, Glue, RDS, EC2, Elastic Load Balancer (ELB), EBS, S3 Buckets, CloudWatch, VPC, IAM

Timeline: Dec 2019 to Nov 2020

How did Matilda help?

Matilda executed the complete migration journey in following phases

- **Discovery:** Matilda Discovery helped to completely assess existing infrastructure and provide infrastructure, services and application dependency
- **Assessment:** Matilda provided detailed assessment at infrastructure and application level with the cost and risk involved in migration process including target architecture
 - Assessment includes working with dependent packages
- **Migration Planning:** Migration planning involves
 - Automation engine recommends for target architecture design including security configurations and optimized resource allocation

- Comprehensive analysis of data source and processing in existing on premise data lake to establish accuracy and coverage of migration.
- **Migration Execution:** Migration process includes mix of automated and manual steps.
 - **Provisioning:** With the recommended design or finalized application design, Matilda automatically provided the target infrastructure (Infrastructure as Code using Terraform) and included necessary packages and services as required by application demands
 - **Application Rearchitect:** to optimize data processing and analytics using EMR Spark, RDS MySQL and AWS S3 storage.
 - **Data Migration:** On prem data was migrated using AWS Snowball, AWS Data Sync and Custom ingestion solution.
 - **Deployment:** end to end deployment flow is created using Matilda workflow.
- **Monitoring Setup:**
 - Matilda was able to setup automatically the required monitoring configurations for applications like CloudWatch setup and routing application logs to S3 buckets
 - Installing required agents at host, service, and database layers

Solution Summary

- Created robust big data platform with Talend based metadata driven framework
- Delivered end to end migration including environment setup, architecture, solution analysis and data ingestion
- Design and Provision highly scalable and fault tolerant infrastructure
- Data ingestion from Heterogeneous sources
- Push down Talend data processing to EMR
- Cost effective storage layers using S3 Buckets for raw and clustered data
- Glue catalog was used for centralized repository

Business Value

- Secure and future ready data lake on AWS
- Agile delivery with DevOps/DevSecOps capability
- Accelerated on boarding of new data feeds by leveraging AWS native services
- Streamlined data feeds from different platforms
- 30% cost savings over onprem data processing
- Overall manual effort during migration is reduced by 70%

Contact us for more information

972.525.2300

Sales@matildacloud.com