



MODERNIZED BIG DATA PLATFORM FOR AMERICAS LEADING BANKING CLIENT

The situation: Decommissioning of on premise Data Lake

One of the Americas leading bank had an on-premise Data Lake built on Big Insights which was nearing its end of life and support. On premise data lake need to be migrated and decommissioned at the earliest before the expire of the license and ensure all data and processes are migrated along with enablement of next gen capabilities.

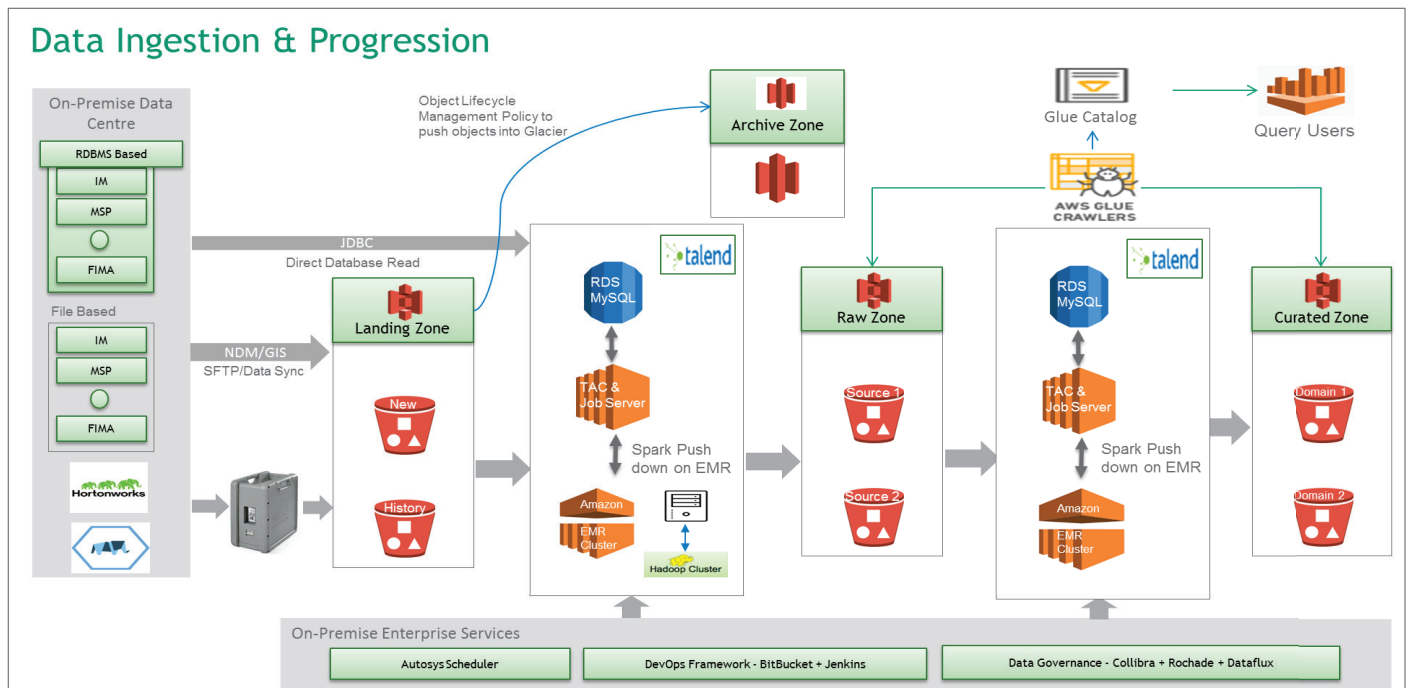
The need: Migrate Data Lake and enable next gen capabilities

- Scalability/Stability of the platform: on premise infrastructure had problems in processing huge volumes of data resulting in missed opportunities
- High cost of maintenance with dependency on IT teams to on board new applications and data feeds
- Limited capability for processing semi structured and unstructured data
- Data silos with quality issues resulted in challenges on getting actionable insights

The Objective: Scalable Data lake with AI, ML Capabilities

To harness and monetize growing volumes of customer and transaction data with more advanced AI and ML processing, the company decided to modernize their architecture by migrating their data from on premise Hadoop platform to AWS platform to leverage the agility and scalability also have the capability to use different AI/ML tools available in cloud platform

The Solution: Next gen data lake in AWS



Details of Solution:

- Created a robust Big data foundation with Talend based metadata driven data movement framework
- Delivered end to end cloud migration solution, starting from environment setup, architecture, source analysis, ingestion and history data loads
- Incremental Business value delivered by agile methodology.
- Infrastructure design, which is high available, fault tolerant and cost effective
- Data is ingested using metadata driven ingestion framework for heterogeneous sources
- Push down of Talend Data processing in to EMR Spark
- Performed comprehensive analysis and assessment of each data source and processing in existing on premise data lake to establish accuracy and coverage of migration
- History data was migrated using combination of AWS Snowball and custom ingestion frameworks
- Cost effective storage layers using Amazon S3 for Landing, Raw and Curated data
- Glue Data Catalog was used as centralized metadata repository and integrated with existing Data Governance tools
- Automated deployment of AWS Infrastructure using Cloud Formation
- Leverage capabilities offered by Target platform
 - Managed services to simplify & save on administration cost
 - Usage of temporary, on-demand storage & processing cluster (ephemeral model) compared to persistence
 - Make impact on pricing due to storage/compute design as regular task and not once-a-while
- Fine grained and Secure data access solution implemented using AWS lake formation

The outcome: Secure and Future ready Data Lake with Enhanced capabilities

- Secured and Future Ready Data Lake on AWS
- Trusted Data Source to drive Consumption
- Agile delivery with DevOps/DevSecOps capability
- Migrated 400 TB of data from On premise data lake to AWS Cloud using AWS Snowball
- Accelerated on boarding of new data feeds by 100X by leveraging metadata driven data movement using Talend and AWS native services
- Enabled data ingestion from ~86 source systems (4000 data feeds from various source like Files, Mainframe, Databases) using metadata driven framework. Migrated Users ~ 200 data Users and 13 live consumer applications to consume data from AWS Data Lake
- Enhanced Self Service and Exploration capabilities using Athena
- Enabled Easier access to data using s3
- 30% cost saving through on demand data processing
- Realized Platform architecture that can be quickly scalable to take care of diverse data processing patterns on demand



Infosys Cobalt is a set of services, solutions and platforms for enterprises to accelerate their cloud journey. It offers over 14,000 cloud assets, over 200 industry cloud solution blueprints and a thriving community of cloud business and technology practitioners to drive increased business value. With Infosys Cobalt, regulatory and security compliance, along with technical and financial governance comes baked into every solution delivered.

For more information, contact askus@infosys.com



© 2021 Infosys Limited, Bengaluru, India. All Rights Reserved. Infosys believes the information in this document is accurate as of its publication date; such information is subject to change without notice. Infosys acknowledges the proprietary rights of other companies to the trademarks, product names and such other intellectual property rights mentioned in this document. Except as expressly permitted, neither this documentation nor any part of it may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, printing, photocopying, recording or otherwise, without the prior permission of Infosys Limited and/ or any named intellectual property rights holders under this document.

[Infosys.com](https://www.infosys.com) | NYSE: INFY

Stay Connected   