

Data Lake implementation Using Open Source Tech Stack & AWS Services

Built one of the largest Data lakes for Al Taysar Group, a renowned Travel & Tourism Company in Middle East to serve as single source of truth for showcasing the enterprise's performance statistics at corporate and business unit level

Industry: Travel & Tourism

Tech Stack

- Hadoop
- Hive
- Nifi
- Spark
- Elastic Search
- Kibana
- Grafana
- Hue
- Airflow
- Redis
- Mysql RDS
- Oracle
- D3Js
- HTML/CSS
- React
- Node Js
- Python
- Scoop
- AWS Cloud services
- Kafka

The Client

Al Taysar Travel Group Holding Company is one of the largest travel and tourism companies in Saudi Arabia with a prominent local, regional and international presence. Al Taysar Travel Group Holding Company is a Saudi joint-stock company. Boasting over 3,000 employees, the company serves customers in over 430 branches in Saudi Arabia and abroad. The Group understands the value of having a presence in its client's favourite tourist destinations, and thus has invested in branches across KSA, the GCC, Egypt, Sudan, Lebanon, Malaysia, the United Kingdom, Canada and other popular tourist location.

Business Pain & Challenges

Al Taysar Travel Group has multiple businesses and subsidiaries working in travel and hospitality domain that store data in their respective enterprise systems. ATG wants to create one of the biggest Data lakes in the middle east which would serve as a single source of truth for the leadership. They required the data lake to be integrated with various sources of data and fetch data near real time. They wanted the data to be cleansed and processed based on rules and logic to provide valuable information in form of reports and dashboards to leadership to allow them take investment related decisions to achieve high returns. The scope of work involved:

- Setup of automated infrastructure for the Data Lake
- Implementation of right tools and processes to monitor and manage complete Cloud environment in real time.
- Provisioning for code/data movement from staging to production environment using CI/CD
- Control user's access to dashboards/reports based on their role
- Integration with multiple complex data sources to build state of art Data Management & Analytics Platform (DMAP), a Data Lake with data processing with ultimate goal of data orchestration, advanced analytics and reporting.
- Creation of multi-dimensional dashboards across various metrics and providing flexibility to change those dimensions as per business needs.

Business Solution

TO THE NEW worked in a distributed agile environment with teams located in 3 different geographies and handled end-to-end development of the platform including front-end, back-end, automated testing, cloud infrastructure and DevOps. Below is a brief overview of the solution we provided.

- Designed & implemented infrastructure & instances as per the agreed architecture for Staging and Production
- Big data development using Nifi, HDP and Spark for ingesting and transforming the data

Know more about our Product Engineering offerings



www.tothenew.com



Talk to our Experts

- Used Zeppelin, Grafana, Kibana and HUE JS monitoring and visualization
- Used D3JS and HTML at visualization layer to create dashboards
- DevOps implementation for the project which includes below
 - AWS Cloud Infrastructure setup and management for DMAP Staging & Production implementation.
 - Automation setup for AWS cloud infrastructure services and management
 - CI / CD Pipeline to setup complete DMAP Staging & Production environments.
 - Centralized logging of different components and stacks
 - Real Time Monitoring and Dashboards
- Followed a combination of Kanban and SCRUM agile approaches for application development

Business Benefits / Outcome

- Developed Data warehouse for ETL.
- Presented real time data transformed and extracted from Google Analytics using Apache Spark and Hadoop.
- Dashboards for presenting business reports with accuracy up to 95%.
- Enabled Corporate and Business Unit heads to view meaningful reports and take decision that could add value to the company

Know more about our Product Engineering offerings



www.tothenew.com



Talk to our Experts