gslab | GAVS

CASE STUDY

Highly Flexible Compliance Reporting using Big Data Analytics for U.S. based Commodities Exchange

Client Overview

The client is an Exchange House in the U.S. operating in commodities. They deal with end-to-end operations right from trading platforms to clearing and settlement to reporting.

The Business Situation

The client wanted to comply with the ever-changing regulatory compliance report requirements. Failing to do so would result in severe penalties. They were faced with the challenge of having disparate sources of data for the report. Moreover, the data, already in the order of exabytes, kept increasing as there was no specified data retention period. Also, the report generation algorithm was rule-based and rigid.

The Solution

GS Lab | GAVS, with deep experience in delivering value through Big Data Analytics solutions, architected a solution around data lakes that was highly flexible and reusable for multiple use cases. Several connectors were written to extract data from various source platforms. The ingested structured and unstructured data was pushed into the data lake solution. The crawler logic was kept configurable. Now, with some configurations, different queries could be set to extract the output in the format needed. The solution enabled them to take input data in different formats.

Challenges

- Difficulty complying with constantly changing regulatory compliance report requirements
- Severe penalties for non-compliance
- Disparate sources of data for report
- Increasing volumes of data, in exabytes
- Rule-based and rigid report generation

Solution Highlights

- Architected around data lakes
- Highly flexible and reusable architecture
- Connectors developed to extract data from various platforms
- Configurable crawler logic
- Queries to extract output in any required format

Solution Outcomes

- Drastically reduced report generation time - from 6-9 months to few weeks
- Increased flexibility due to highly configurable reporting rules
- Extremely low-cost storage for exabytes of data
- Ready availability of data for analytics regardless of data volume or formats