

# Leveraging Modern Data Lake Architecture to accelerate growth for a leading Oil & Gas Company

## Challenge

The client, a leading oil and gas company, sought to migrate their data processing and analytics workloads from EMR to Databricks to leverage its advanced capabilities, such as optimized performance, better cost efficiency, and enhanced data governance.

## Technology stack

- **AWS S3:** For data storage
- **Databricks:** For data processing, transformation and unified data analysis
- **Medallion Architecture:** For data architecture and organization
- **Config-driven Frameworks:** For data integration, data transformation and data quality
- **Azure DevOps:** For CI/CD pipelines
- **Terraform:** For infrastructure as code (IaC) and environment setup

## Business Challenges

- Client was facing performance and scaling issues on AWS EMR cluster
- Cost management was out of control for every increasing demand
- Lack of search and discovery of datasets was causing time to market challenges for new analytics products
- Data and business logic was duplicated with pointed solutions in data marts

## Outcomes

- **40% improvement** in data processing performance
- **30% reduction** in operational cost through optimized resource utilization and auto scaling
- **Enhanced Scalability:** enabled seamless scaling of data processing workloads to accommodate growing data volumes
- **Better Data Governance:** implemented robust data governance practices, ensuring data accuracy, security and compliance
- **Operational Excellence:** improved reliability and maintainability of data pipelines, resulting in fewer operational issues and faster resolution times
- **Self-Service** reporting providing full transparency into operations and billing

## About Straive

As a data analytics and AI operationalization company, we don't just build top-tier data analytics and AI solutions—we integrate them seamlessly into your core workflows. This approach drives enhanced efficiency, improves user experience, and boosts revenue, setting you apart from the competition.

Serving a diverse range of industries— including Pharmaceutical & Life Sciences, Logistics, Supply Chain & Manufacturing, Research & Publishing, Information Services, EdTech, Banking & Financial services, and Retail Tech, Entertainment & Media— we have a global client base spanning over 30 countries.

Our strategically positioned resource pool operates across seven countries, including the Philippines, India, the United States, Nicaragua, Vietnam, the United Kingdom, and Singapore, where the company is headquartered.