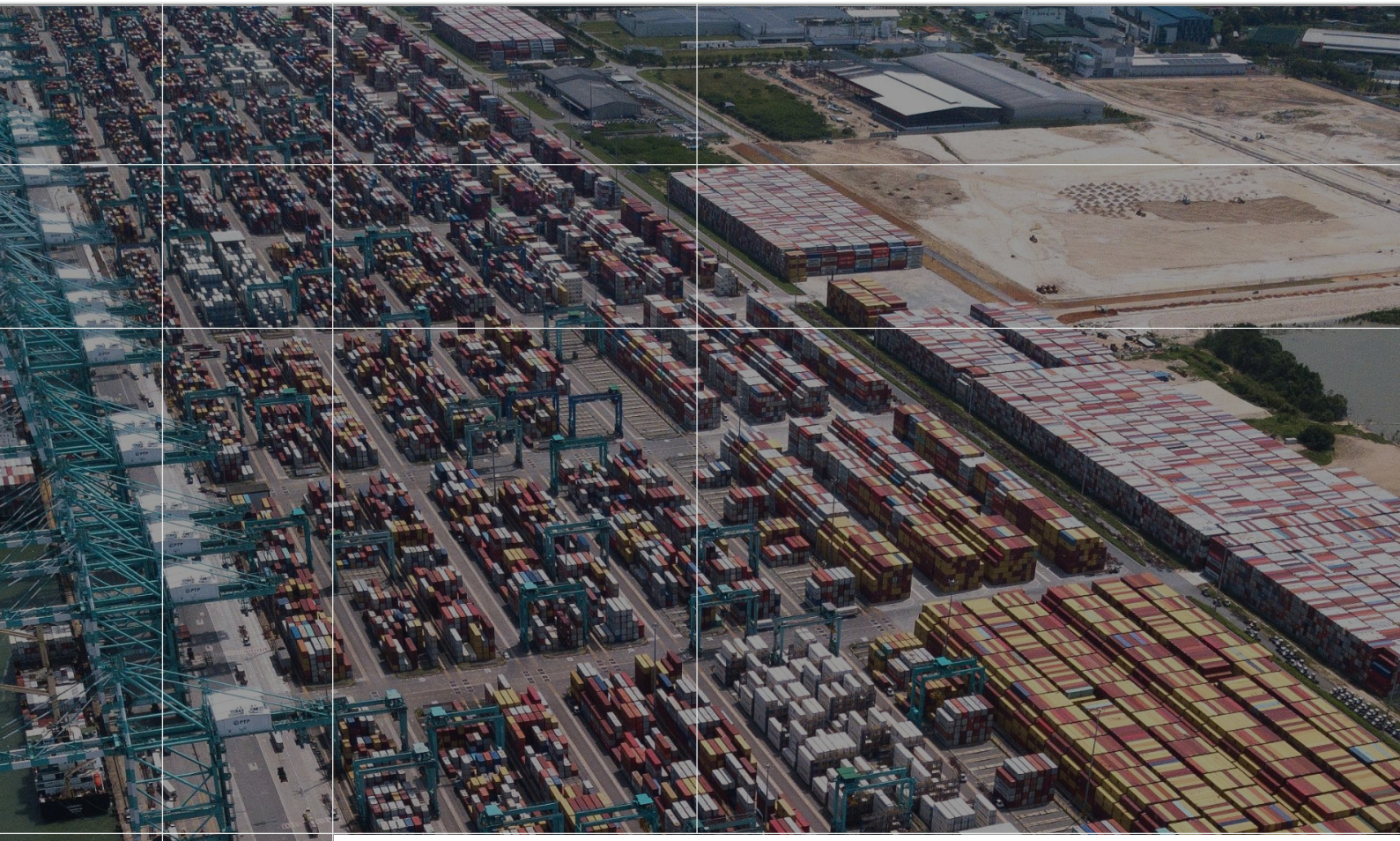
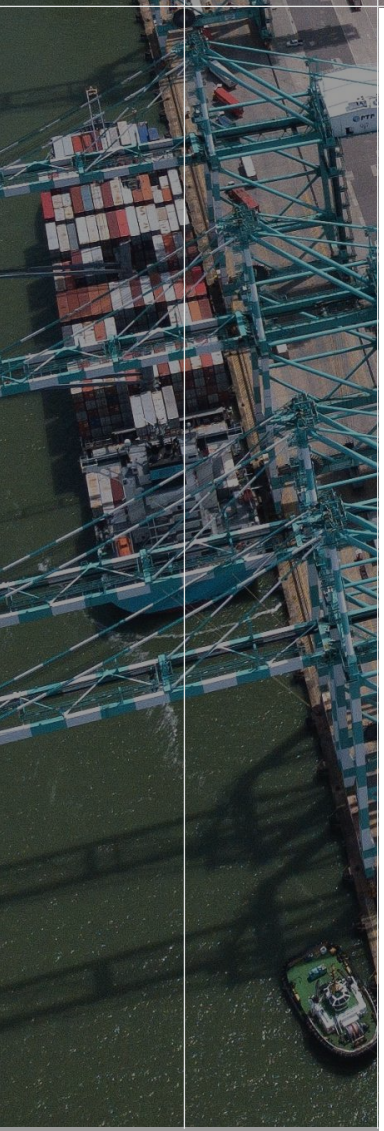
Tiger Analytics

databricks





Delivering Real-Time Operational Insights for PTP With a Modern Lakehouse Platform

Accelerating modernization by adopting future-proof solutions

Executive Summary

Pelabuhan Tanjung Pelepas Sdn Bhd (PTP) is Malaysia's premier transshipment port, strategically positioned to serve as a vital gateway for global trade. With state-of-the-art facilities, advanced equipment, and cutting-edge information technology systems, PTP seamlessly integrates all port users, ensuring efficient and reliable operations across the entire logistics chain.

PTP recognized the opportunity to enhance its data ecosystem by exploring strategies to enhance the integration of data across multiple systems. To scale the agility of its operations and respond quickly to market changes, a comprehensive view of all its different data systems would help its leadership make better, data-driven decisions. This called for a centralized repository and modern data systems to serve as a single source of truth across the entire logistics value chain.

To kick-start this transformation journey, PTP partnered with **Tiger Analytics** to develop a **Modern Data Lakehouse** Platform, powered by **Databricks**. The data lake would handle a wide range of data types, from large volumes to different speeds and sources, organizing them centrally, making it easier to extract insights. Tiger Analytics leveraged **Databricks' Unity Catalog**, Databricks' unified solution to improve data governance, security, and accessibility. This ensured that data was well-governed, secure, and easily discoverable, empowering data-driven decisions.

Implementation of this solution would help:

- Consolidate all structured and unstructured data into a unified repository.
- Enable real-time analytics and business intelligence.
- Break down data silos to strengthen trend analysis, pattern recognition, and insights.
- Enhance decision making, operational efficiency, and strategic growth across Operations, Productivity, Maintenance, and Finance.

Business Challenges

The fragmentation of data across multiple systems had a significant impact across key functions at PTP such as operations, productivity, maintenance, and finance. Some of these were:

Need for integration across 10+ operational source systems and a 17% year-over-year increase in data volume.

Capturing and managing intermediate data states during updates and inserts.

Generating real-time operational reports.

Developing machine learning (ML) and business intelligence (BI) use cases.

Delay in report generation due to reliance on macros-enabled Excel, where complex calculations take over an hour to process, leading to potential manual errors.

Reliance on manual judgment to estimate and adjust the number of prime movers required for Quay Cranes, resulting in discrepancies between reported and actual performance.

Choosing the Right Partner

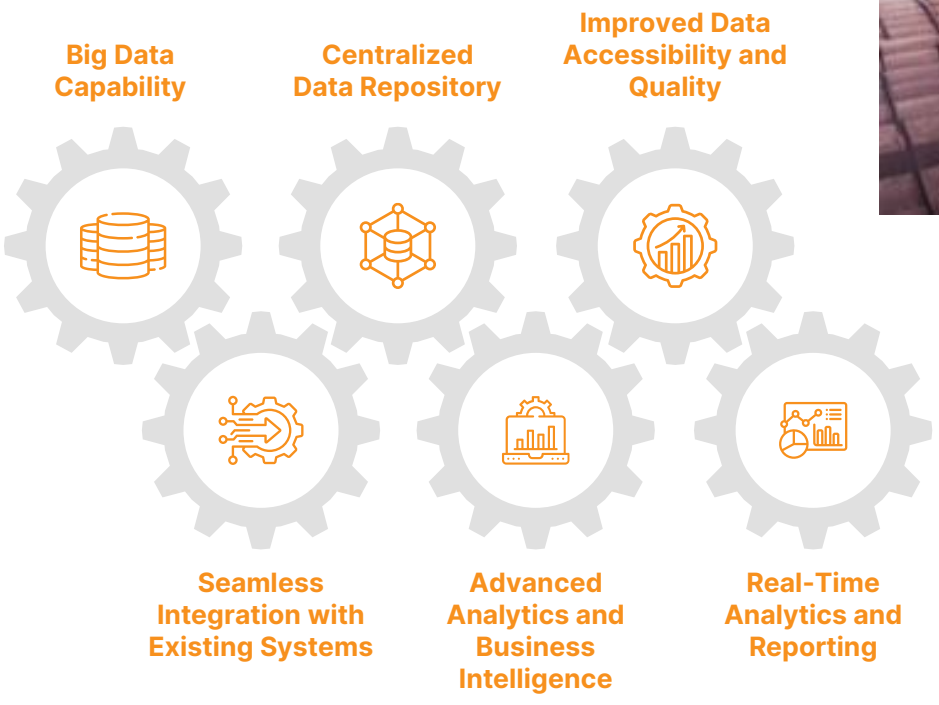
To set up a comprehensive data system that would reinforce their leadership in the global maritime industry, PTP needed a partner with the necessary technical expertise and the required execution rigor including:

- Expertise in modern data lake implementation
- Experience in executing substantial data engineering projects spanning multiple functions
- Advanced machine learning capabilities for potential future initiatives
- Deep understanding of BI use cases for streamlined reporting

Given our deep data engineering expertise across these building and implementing data lakes, advanced ML capabilities, domain expertise and experience working on BI projects, **Tiger Analytics** seamlessly aligned with the project requirements.

Proposed Strategy

To address the business need, a Data Lake and Analytics platform was chosen as the best solution. This next-generation solution would centralize PTP's data - both structured and unstructured - into a single, unified interface. The core features and benefits of the solution are:



Solution Approach

Tiger Analytics adopted a six-step approach in implementing this Data Lake platform at PTP:

Event Driven Processing:
Continuous event ingestion/processing from Terminal OS using Azure Event Hub and Databricks Workflows

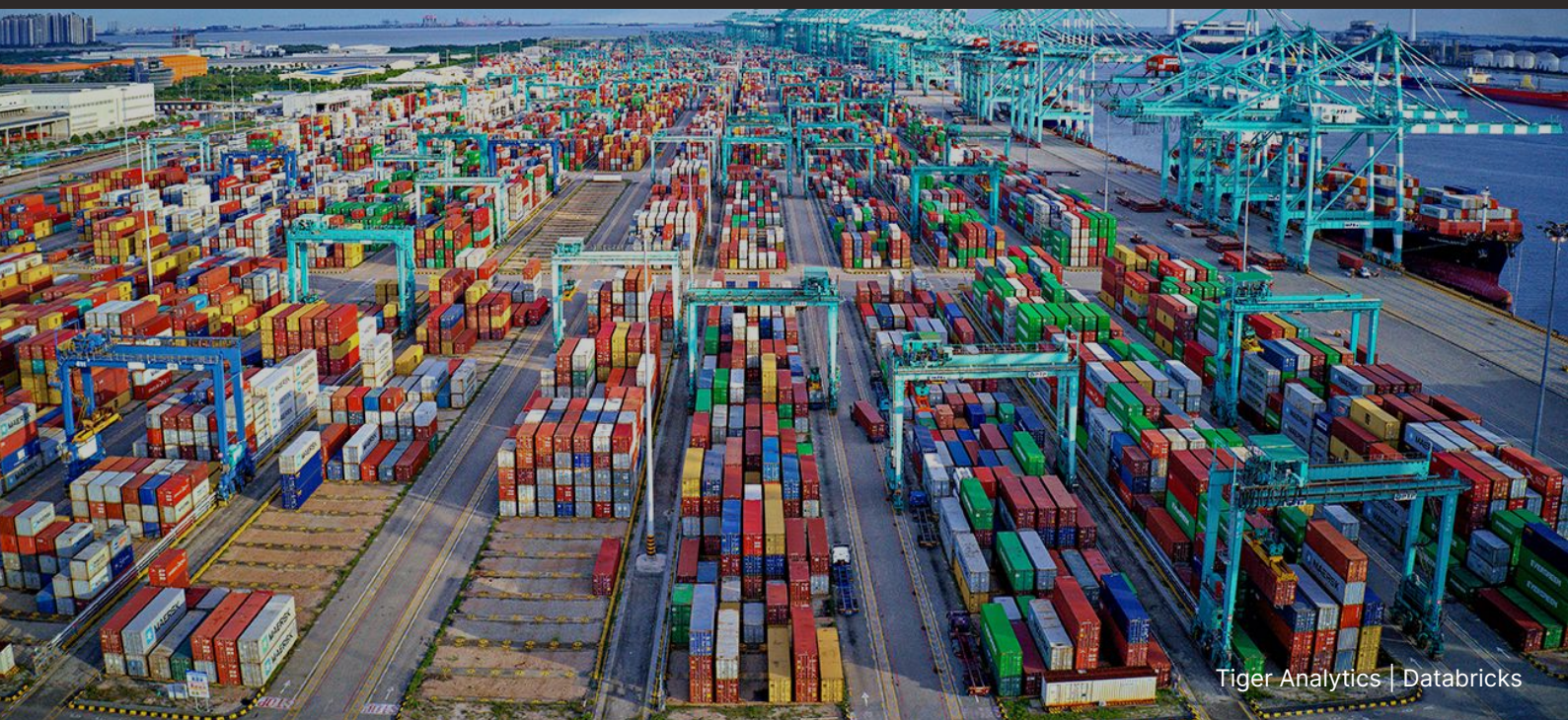
Data Capture and Transmission:
Employed Debezium connectors, APIs, file systems, SharePoint to capture and send both real-time and batch data

Data Organization:
Applied a Medallion architecture to organize data into raw, cleaned, and business-level formats

Scheduling and Scalability:
Utilized Azure Databricks Workflows for near real-time job scheduling and execution, providing a scalable and reliable solution with quick data updates.

Data Orchestration and Transformation:
Utilized Azure Data Factory to manage data flows and Azure Databricks to apply business logic and prepare data for Power BI

ML Pipelines-based Performance Forecasting:
Past metrics, scenarios captured to train models in Databricks for real time performance analytics



Business Outcomes

Delivering this data lake resulted in significant advantages for PTP, a few of which are outlined below:

Unified Data Platform for Integration and Scalability	Centralized data from across 10+ operational sources into a single platform
	The data lake's scalable storage and processing capabilities can handle increasing data volumes without affecting performance, ensuring the organization can continue to ingest and process data as it grows
Enabled Real-Time Operational Reporting	Real-time insights provide a strategic advantage, allowing the business to make data-driven decisions faster and with greater accuracy
Facilitated Development of ML and BI Use Cases	Centralized repository of both structured and unstructured data for ML and BI initiatives
	More sophisticated analytics and predictive models, helping the business to uncover new insights and optimize operations
Streamlined Report Generation and Reduction in Manual Errors	Complex calculations are handled by the data lake's processing power, reducing report generation time from hours to minutes
	Eliminates the risk of manual errors and provides a more reliable, automated, and faster reporting process
	Enhanced decision-making for resource allocation

Apart from these advantages, the platform also enabled multiple use cases across various functions within the organization. These include:

01 Prime Mover (PM) Deployment Planning

- **Accurate PM Estimation:** Provides precise forecasts of the number of PMs needed for each shift, improving planning accuracy and resource allocation.
- **Enhanced Resource Utilization:** Optimizes the deployment of PMs by considering multiple variables, leading to better management of equipment and personnel.
- **Optimized Shift Coverage:** Ensures that the right number of PMs is available for each hour, aligning with the specific needs of the shift and vessel operations.
- **Simulation Models:** Simulates various scenarios (e.g., twin container handling, deck levels) to understand their effect on PM deployment needs.
- **Informed Planning:** Helps the execution team anticipate and plan for variations in move complexity and distance, improving overall efficiency.
- **Strategic Insights:** Offers valuable data on past performance, helping to refine future deployment strategies and optimize resource management over time.

02 Upcoming ML Based Solutions
Likelihood of Delinquency Prediction

Proactive Management	Allows the OLM team to identify Terminal Equipment Operator (TEO) at higher risk of delinquency early, enabling targeted interventions and preventive measures.
Improved Resource Allocation	Helps in allocating additional support and training resources to those identified as higher risk, improving overall workforce efficiency and safety.
Targeted Training	Identifies specific areas where behavioral improvements are needed, allowing for tailored training programs to address these issues.
Real-Time Insights	Enables the OLM team to monitor and respond to performance issues as they arise, ensuring timely and effective interventions.
Behavioral Scoring Systems	Develop scoring systems based on a combination of non-performance metrics to assess overall TEO behaviour and risk profiles.
Holistic Assessment	Offers a comprehensive view of TEO performance by incorporating various non-performance factors, providing a more complete picture of operator behaviour.
Informed Decision-Making	Enhances the ability of the OLM team to make data-driven decisions regarding training, resource allocation, and performance management.

03 Power Consumption Anomaly Detection and Equipment Efficiency using IoT

Early Detection
Identifies abnormal power consumption early, allowing for timely investigation and resolution of potential issues.

Focused Diagnostics
Helps pinpoint specific equipment or containers experiencing unusual power consumption, facilitating targeted troubleshooting.

Detailed Diagnostics
Breaks down power consumption into various influencing factors, enabling a comprehensive RCA for identifying specific causes of power spikes or inefficiencies.

Targeted Interventions
Facilitates precise action plans based on the identified root causes, improving overall power efficiency and reducing operational costs.

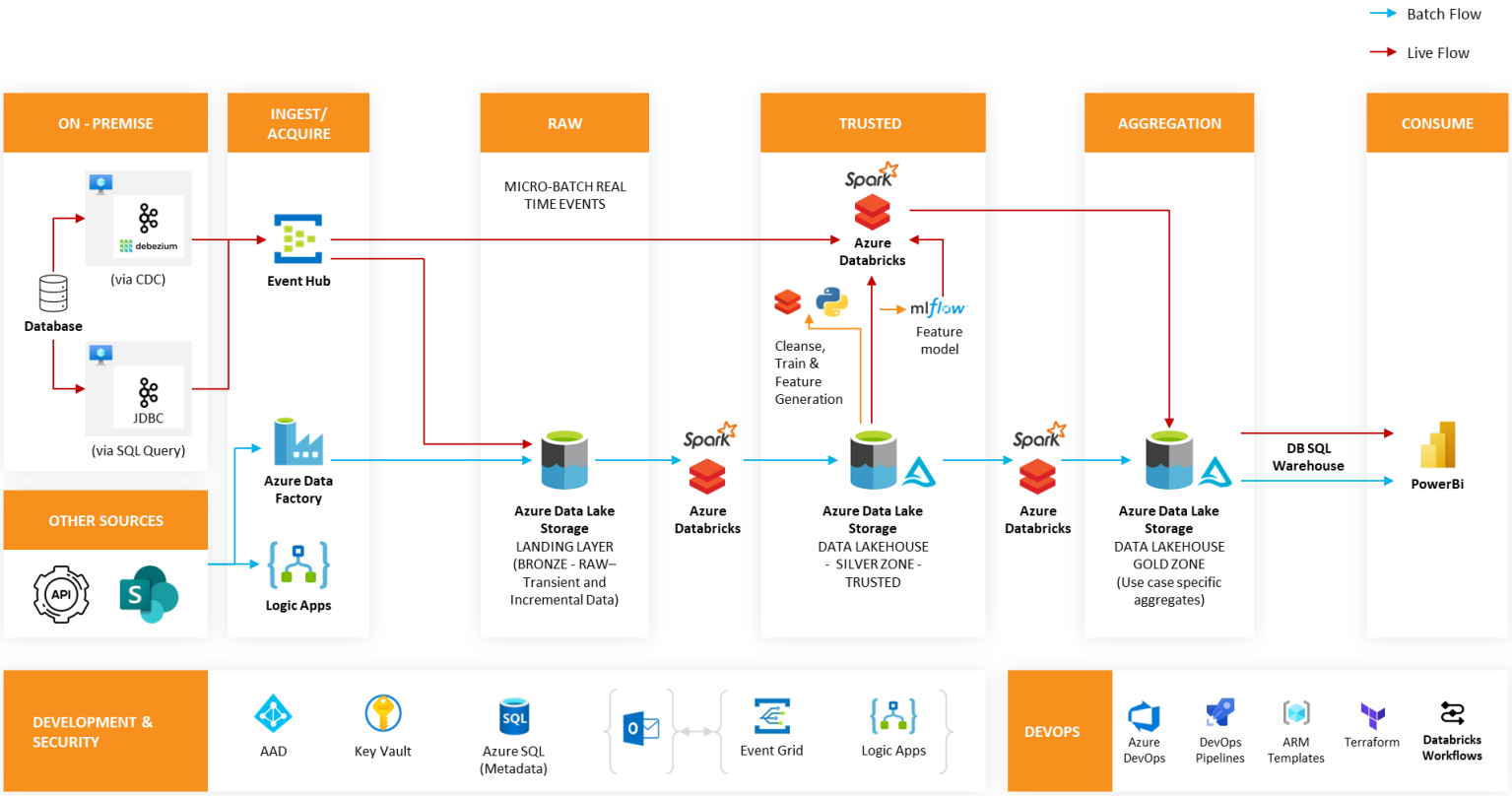
Streamlined Reporting
Automates the generation of monthly electricity consumption reports, reducing manual effort and ensuring timely and accurate reporting.

Real-Time Monitoring
Provides daily updates on power consumption and cost, enabling ongoing monitoring and prompt action based on the latest data.

Tech Stack

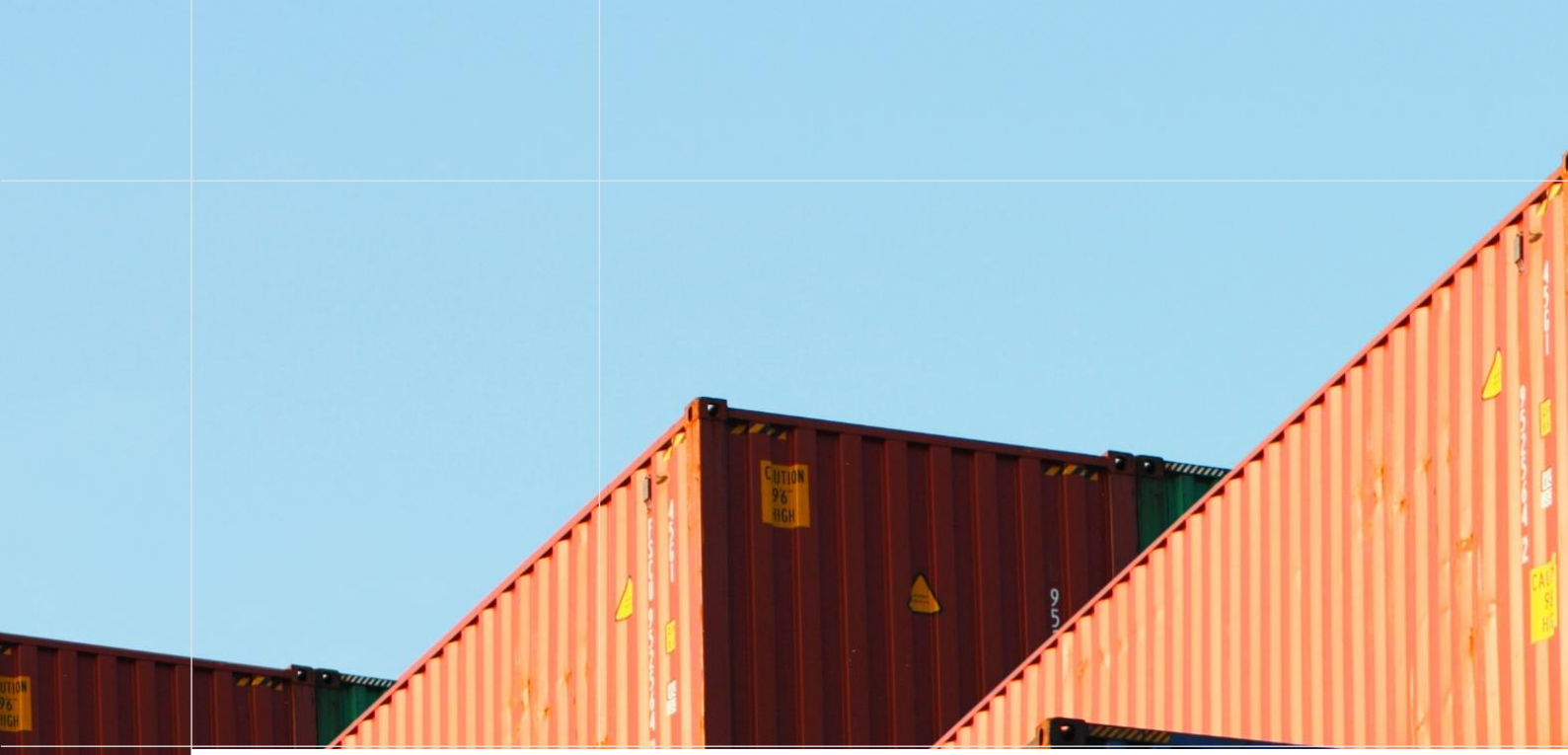
Kafka, Databricks for ETL + Processing & Streaming

Technical Data-lake Architecture



Conclusion

The data lake platform developed by Tiger Analytics and powered by Databricks has had a profound impact on PTP by improving operational efficiency, enabling smarter decision-making, optimizing resource management, and driving business growth. Additionally, it will empower PTP to fully capitalize on its data for strategic advantage, positioning the port as a more agile, efficient, and innovative leader in the global maritime sector.



About Tiger Analytics

Tiger Analytics is a global leader in AI and analytics, helping Fortune 1000 companies solve their toughest challenges. We offer full-stack AI and analytics services & solutions to help businesses achieve real outcomes and value at scale. We are on a mission to push the boundaries of what AI and analytics can do to help enterprises navigate uncertainty and move forward decisively. Our purpose is to **provide certainty to shape a better tomorrow.**

Being a recipient of multiple industry awards and recognitions, we have 5000+ technologists and consultants, working from multiple cities in 5 continents.

www.tigeranalytics.com