

SETOR: MINERAÇÃO E SIDERURGIA · SAP · KAFKA · DATABRICKS · AZURE

Integração SAP em tempo real para acelerar a disponibilidade de dados no Databricks

Introdução

Uma empresa do setor de mineração e siderurgia precisava evoluir a forma como extraía, integrava e disponibilizava dados do SAP para sua plataforma analítica. O desafio era transformar dados antes restritos ao sistema transacional em uma base mais acessível, confiável e preparada para análises, dashboards, APIs, indicadores e novas iniciativas orientadas por dados.

Big Numbers & Destaques



Tempo Real

Dados do SAP disponíveis em tempo real no Databricks



~30 min

Para configurar uma nova tabela no framework



3 meses

Da implementação batch até a migração para streaming



Menor esforço

Técnico para novas integrações SAP



Redução de riscos

Com padronização e automação



Framework ativo

Frente de ingestões ativa para novas demandas

Contexto e desafio da empresa

A empresa precisava extrair dados de negócio e operação do SAP sem depender exclusivamente dos frameworks e ferramentas do próprio sistema. Enquanto essas informações permaneciam restritas ao ambiente transacional, as análises ficavam limitadas e o acesso aos dados poderia se tornar um gargalo para a evolução data-driven da companhia. O objetivo era disponibilizar esses dados em uma plataforma analítica moderna, permitindo cruzamento com outras fontes, mais autonomia para o time técnico e maior capacidade de resposta às demandas das áreas de negócio.

Solução

A Objective implementou um framework de ingestão de dados estruturados do SAP em tempo real, conectando a origem transacional ao Databricks com uma arquitetura modular, escalável e aderente à stack homologada pelo cliente. A solução evoluiu de um modelo batch para uma abordagem em streaming com CDC, combinando SAP, Kafka, Databricks, Azure, Python, Spark, Delta Lake, Spark Streaming, Databricks Workflows e Jobs Databricks. Além da implementação técnica, o projeto incluiu definição de padrões, documentação, automação e governança para facilitar novas ingestões com mais qualidade e menor dependência de configurações manuais.

SAP

Kafka

Databricks

Azure

Python

Spark

Delta Lake

Spark Streaming

Resultados

Com o novo framework, a empresa passou a ter uma esteira mais fluida e confiável para disponibilizar dados do SAP no Databricks. A principal evolução foi sair de um processo com janelas fixas e maior esforço manual para uma ingestão contínua, padronizada e escalável. Na prática, isso reduziu retrabalho técnico, acelerou novas integrações e fortaleceu a base de dados para relatórios, dashboards, APIs, indicadores e iniciativas de IA com mais governança, autonomia e consistência.