

Oracle Cloud Infrastructure In Lab

Arquitetura de Dados Episódio II

- As forças imperiais comandadas por Lakehouse lançam um ataque para tratar a atomicidade, consistência, isolamento e durabilidade de seus dados no DataLake. Venha conhecer como trabalhar com DeltaLake em OCI .

Início do evento 14:00



Heloisa Escobar

Senior Cloud Engineer

17/02/2023

Declaração de Porto Seguro (Safe Harbor Statement) - Geral

O texto a seguir tem como objetivo traçar a orientação dos nossos produtos em geral. É destinado somente a fins informativos e não pode ser incorporado a um contrato. Ele não representa um compromisso de entrega de qualquer tipo de material, código ou funcionalidade e não deve ser considerado em decisões de compra. O desenvolvimento, a liberação, a data de disponibilidade e a precificação de quaisquer funcionalidades ou recursos descritos para produtos da Oracle estão sujeitos a mudanças e são de critério exclusivo da Oracle Corporation.

Esta é a tradução de uma apresentação em inglês preparada para a sede da Oracle nos Estados Unidos. A tradução é realizada como cortesia e não está isenta de erros. Os recursos e funcionalidades podem não estar disponíveis em todos os países e idiomas. Caso tenha dúvidas, entre em contato com o representante de vendas da Oracle.

Declaração de Porto Seguro (Safe Harbor Statement) - específica para o Brasil e Portugal

As informações a seguir foram coletadas pela Oracle a partir de um analista terceirizado do setor e têm como objetivo apresentar destaques da orientação dos nossos produtos em geral. São destinadas somente a fins informativos e não podem ser consideradas para diferenciar produtos da Oracle dos produtos de outras empresas.

A Oracle conduziu a análise de “prova de conceito” com um alto nível de atendimento padrão do setor, mas as informações a seguir não estão isentas de erros. A Oracle não afirma ou garante que os resultados da análise possam ser reproduzidos e/ou demonstrados de maneira lógica. A Oracle não fornecerá nenhuma informação referente à análise, incluindo, mas não se limitando a, método, lógica, ideia, estratégia de análise e política de análise.





Data Warehouse



Data Lake



5



Características Data Warehouse



- Armazenamento de Forma estruturada dos dados Direcionado a finalidade Específica
- Disponibilização de dados relevantes com visão estratégica
- Confiabilidade das Informações
- Qualidade dos Dados



- Estruturas Complexas, dificultando escalabilidade
- Alto Custo de manutenção e armazenamento
- Lento para Implementação de novos projetos
- Não suporta dados semi e não estruturados
- Não suporta a Streaming
- Não suporta processos de Exploração, Machine Learning e AI





Características Data Lake



- Centralização dos Dados
- Agilidade e Disponibilização dos Dados
- Estrutura simplificada
- Segregação de armazenamento do processamento
- Suporta todo tipo de Dado e Workload
- Expansão para cases de Data Science, Machine Learning e AI
- Custo Benefício e Escalável



- Divisão das Tecnologias de Data Warehouse e Data Lake
- Governança e Qualidade dos Dados
- Versionamento e Gestão da Informação
- Disponibilização de Reports e Análises
- Não suporta transações ACID



Data Warehouse



Data Lake



Lakehouse



O que é ACID?

- Propriedades de Transação de um sistema de banco de dados;
- Sequência de Operações que devem ser executadas em conjunto
- Mudança de estado consistente



O que é ACID?

Atomicidade

- **Alteração do estado deve ser complete ou nada será realizado**
Ou o dinheiro é subtraído de uma conta e adicionado à outra ou nada acontece.

Consistência

- **Garante o banco de dados em um estado consistente**
O dinheiro não pode estar em ambas as contas ao mesmo tempo ou perdido completamente.

Isolamento

- **A transação é executada “sozinha” e não tem efeitos em outras atividades**
Uma transferência não interfere em outras realizadas ao mesmo tempo.

Durabilidade

- **As alterações a serem aplicadas por nossa transação são armazenadas de forma permanente e Segura.**
A transferência é mantida em disco em vez de ser mantida na memória e poder ser “reproduzida em caso de falha”

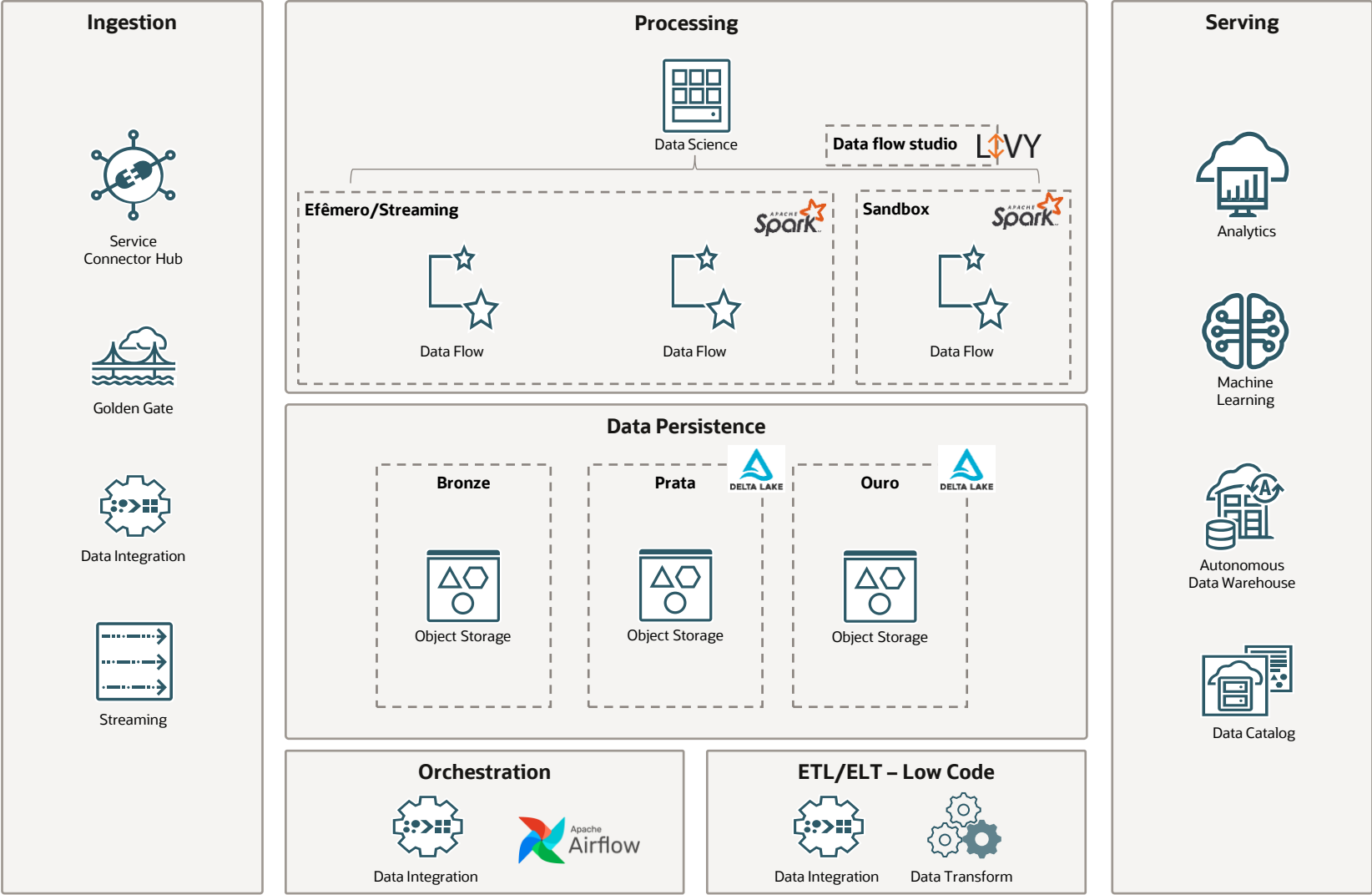
Arquitetura Conceitual



A laptop is shown from a front-facing perspective, open. The screen is white and displays the word "DEMO" in a bold, yellow, serif typeface, centered horizontally and vertically. The laptop's bezel is black, and its base is a dark grey or black. The entire device is set against a plain white background.

DEMO

OCI Lake House



OCI Data Flow

Seu cluster Spark Gerenciado

Spark on Demand

Implante clusters Spark em minutos sem nada para manter

Operação simplificada

Uis para monitoramento, alertas e diagnósticos.

Baixo Custo

Não existe custo pelo serviço,
Pague apenas pela Infraestrutura utilizada.



Discover your Data Lake

Consultar dados estruturados e não estruturados
No Object Storage com SQL

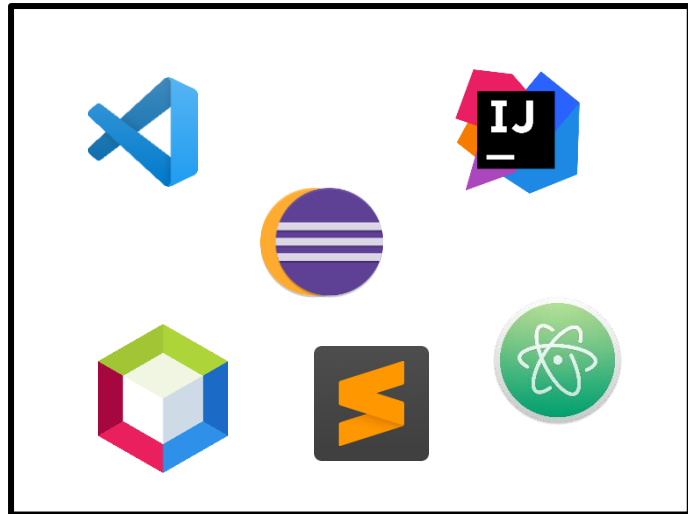
Integrated

Consulta com o Oracle Analytics Cloud ou traga sua ferramenta de BI favorita

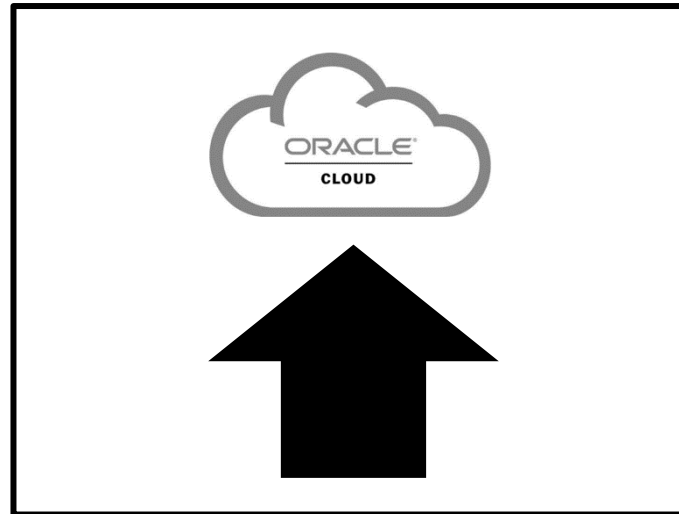


OCI Data Flow

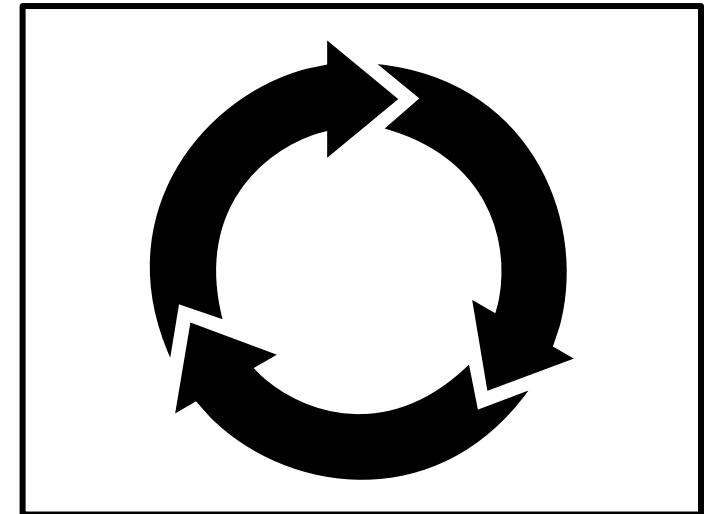
Data Flow torna a operação mais simples



Desenvolva e teste aplicativos Spark localmente usando o IDE de sua escolha.



Carregue aplicativos como estão, sem modificações.



Execute na nuvem em qualquer escala.

Plataforma de Desenvolvimento – OCI Data Science

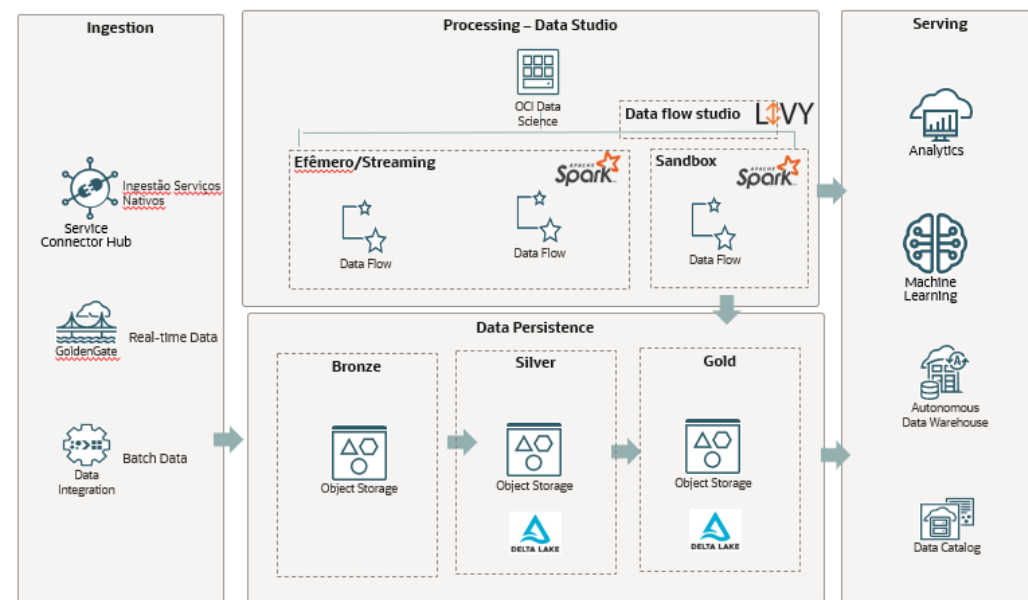
Serviço baseado em JupyterLab para Ciência de Dados e Engenharia de Dados

- Serviço PaaS gerenciado pela OCI
- Amplo Conjunto de Algoritmos e Estrutura de Código Aberto Python
- Infraestrutura Escalável para atender as necessidades específicas
- AutoML, Automates Python Algorithm selection, tuning



Data Flow Studio

Agilizar o tratamento, transformação e extração de valor de dados é crucial para estabelecer uma abordagem Orientada à Dados de sucesso



Obrigado!
Até a próxima sessão