

Bootcamp: Arquiteto Cloud Computing

Desafio

Módulo 3

Soluções de Dados, Big Data e Machine Learning

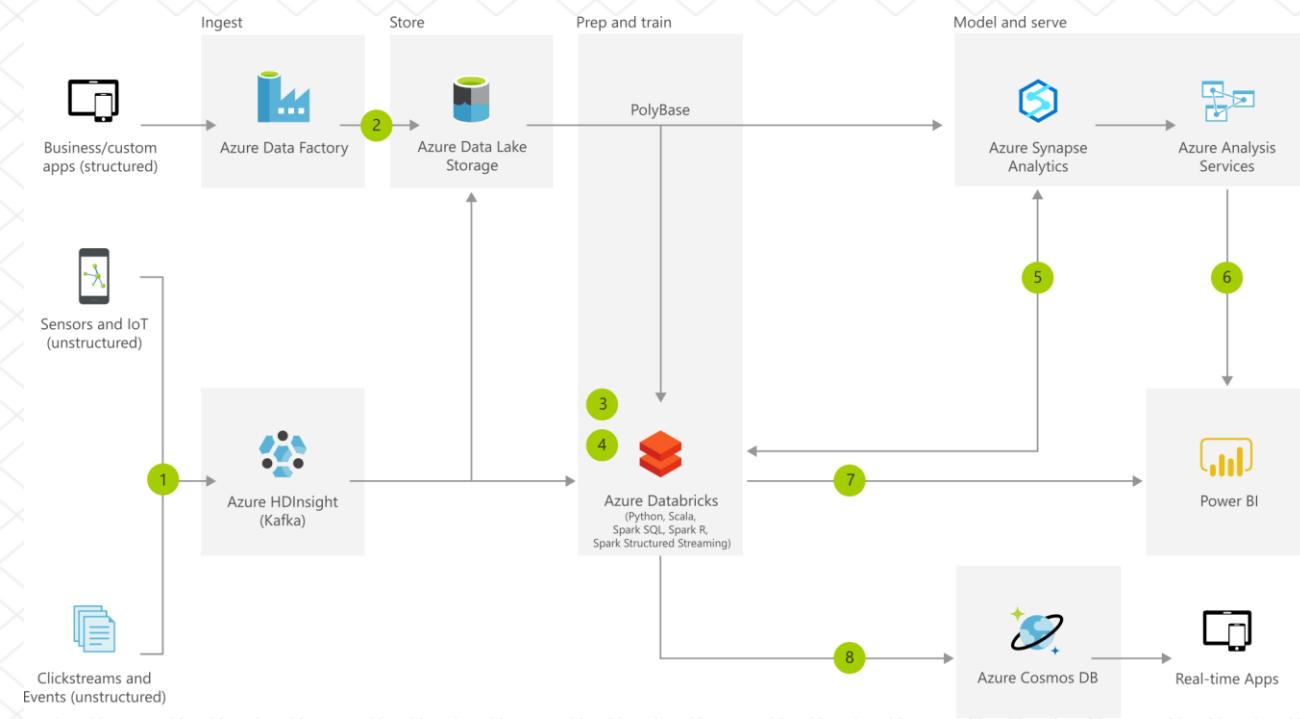
Objetivos

Exercitar os seguintes conceitos trabalhados no Módulo acerca das seguintes soluções do Microsoft Azure:

- ✓ Azure Data Factory.
- ✓ Azure Data Lake Storage.
- ✓ Azure HDInsight.
- ✓ Azure Databricks.
- ✓ Azure Cosmos DB.
- ✓ Azure Synapse Analytics.

Enunciado

Para a necessidade de uma grande multinacional, que consiste em realizar análises de dados em tempo real com Big Data, o arquiteto de solução optou por utilizar a arquitetura de referência da Microsoft, disponível no [site](#) e mostrada a seguir:



Atividades

A empresa lhe contratou para fazer a implantação da infraestrutura, no Azure, para a arquitetura de solução acima.

Você deverá desempenhar as seguintes atividades:

1. Fazer o deploy dos recursos abaixo, no Azure;
 - Azure Data Lake Storage para armazenar os dados que serão ingeridos de um banco de dados relacional, usando-se o Azure Data Factory;
 - Cluster Hadoop do Azure HDInsight.
 - Cluster Spark do Azure Databricks.
 - Ambiente do Azure Synapse Analytics.
 - Um banco de dados Azure Cosmos DB para cada API que ele suporta.
2. Construir um pipeline de exemplo para copiar os dados de um banco de dados relacional para o Azure Data Lake Storage.

Respostas Finais

Você deverá desenvolver a prática e, depois, responder às seguintes questões objetivas: