

Cenário Atual

Estamos passando por um processo de transformação digital, onde o seu papel como Arquiteto de Dados será definir a arquitetura de referência para plataforma de dados do Grupo Boticário e ser uma referência técnica para engenheiros e analistas de dados.

No cenário atual, utilizamos SAP Hana como nosso repositório principal de data warehouse. Existem processos de ETL que fazem ingestão de dados de 50 transacionais. Mais de 90% das bases são de origem transacionais de diferentes DBMS's (DB2, MS SQL etc) e estão alocados em ambiente on-premises.

Além do SAP Hana, a empresa possui algumas aplicações hospedadas em nuvens públicas como Microsoft Azure e Amazon Web Services.

Dentro da empresa, o tratamento e o consumo dos dados são tratados em silos, onde diferentes unidades de negócios acabam utilizando diferentes ferramentas para processar, analisar dados e apresentar dados. Algumas ferramentas que podemos citar como exemplo são Jupyter Notebook, Qlick, Qlick Sense.

Outro aspecto importante está ligado a governança de dados, onde aspectos como acesso a dados sensíveis, catalogação e permissionamento carecem de melhorias.

O que esperamos?

1. Que você defina uma arquitetura de referência com tecnologias de alguma nuvem pública, preferencialmente AWS ou GCP. Você deve considerar os seguintes requisitos:
 - Permeiar as camadas de ingestão, processamento, armazenamento, consumo, análise, segurança e governança;
 - Substituição gradativa do cenário on-premises atual;
 - Incorporação de componentes e tecnologias que permitam a analisarmos dados em tempo real;
 - Que a arquitetura considere componentes que a habilitem a empresa organizar e fornecer dados para diferentes fins, tais como: Analytics, Data Science, API's e serviços para integrações com aplicações. Ressaltando que necessariamente precisaremos manter a comunicação on-premises x cloud para diversas finalidades.
2. Que você prepare uma apresentação para discutir a arquitetura definida, detalhando a motivação da escolha dos componentes, especificando as vantagens, desvantagens e riscos do modelo definido.
3. Que você apresente de forma presencial ou videoconferência para nosso time de arquitetura. Você terá 30 minutos para apresentação.
4. Que você traga exemplos práticos, onde tenha aplicado anteriormente a utilização dos componentes selecionados para esta arquitetura.

Observação: Ao definir a arquitetura, considere que somos uma das maiores empresas de varejo do Brasil.

O que não esperamos?

Para este case, avaliaremos seu conhecimento de arquitetura e capacidade de atuar de forma consultiva junto a engenheiros, cientistas e analistas de dados. Não é esperado que você implemente a arquitetura ou que realize qualquer tipo de codificação.