

**Trabalho:** Construir um desenho de arquitetura e descrever os fluxos de processamentos de dados entre zonas e aplicar os recursos demonstrados em aula até o momento.

**Data de entrega:** 27/11/2019

**Peso da avaliação:** 10

**Enunciado:**

A empresa E-commerce-XYZ deseja criar um data lake para tornar a sua empresa orientada a dados com base centralizada para estudos. Podendo cruzar informações de múltiplos sistemas em um único lugar. Atualmente a empresa dispõe dos seguintes recursos nos seus sistemas transacionais:

- Base de dados SQL Server.
- Base de dados Oracle.
- Sistemas cliente-server.
- Aplicações Java para micro serviços.
- Aplicações Mobile.
- Aplicações .NET.
- Utiliza Apache Kafka para integrações entre aplicações.
- Processos ETL para integrações pontuais entre sistemas.
- Recebe arquivos de dados de fontes externas no formato CSV.

**Requisitos do projeto:**

- Armazenados de forma centralizada no data lake organizado por zonas e com fluxo de processamento de dados entre as zonas.
- Deve conter particionamento.
- Utilizar armazenamento com compressão e colunar.
- Ilustrar e descrever camada de ingestão.
  - Transacional
  - Mensageria
- Camada de catalogação de dados.
- Camada de self-service.
  - SQL
  - Notebook

**Sugestão:**

Utilizar draw.io para o desenho.

Link: <https://www.draw.io>

Desenho inicial:

