

Plataforma Big Data para a IIoT

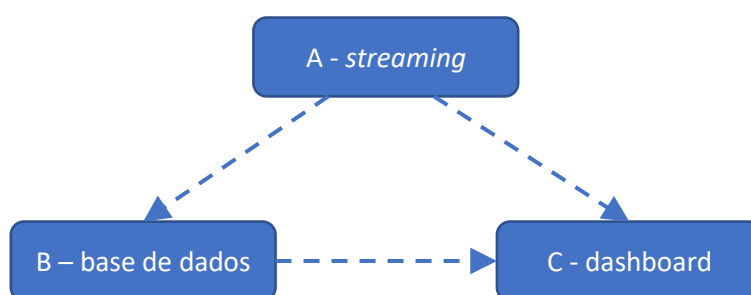
Nº de alunos: 4 a 5

Orientador: Hélder Troca Zagalo (htz@ua.pt)

Colaboradores: José Duarte e Rogério Costa

Descrição do projeto

O objetivo deste projeto é desenvolver uma plataforma para armazenar e processar dados provenientes de máquinas, dispositivos, sensores, pessoas ou sistemas de informação, em aplicações industriais (IIoT – Industrial IoT). A arquitetura recomendada é a seguinte:



- O módulo A recebe os dados *raw* de sensores, máquinas ou outros, e prepara-os para inserção numa base de dados (módulo B), ou para serem usados diretamente por uma aplicação cliente (módulo C). Recomenda-se a utilização de um modelo *Publisher-Subscriber*, usando ferramentas como o Apache Kafka, Apache Flink, ou outra.
- O módulo B representa um *cluster* com um número variável de *nodes* (3 a 8) para armazenar os dados numa base de dados Apache Impala. Deve ser possível adicionar *nodes* de forma transparente para adaptar a capacidade de processamento do sistema ao volume de dados a processar.
- O módulo C é uma aplicação cliente (*dashboard*) para visualização dos dados com origem no módulo A (em “quase” tempo real) ou no módulo B (base de dados).

Vão ser disponibilizados dados do projeto “PRODUTECH-SIF: Soluções para a Indústria de Futuro” e um simulador de dados para testar o sistema. Opcionalmente, os alunos podem desenvolver aplicações (por exemplo, uma aplicação móvel) à sua escolha para capturar e enviar dados em tempo real para o sistema.

Desafio suplementar:

Fazer testes de desempenho e definir regras de dimensionamento do sistema, por exemplo, para processar x operações por segundo recomendam-se n nodes.

Competências técnicas:

Conhecimentos em bases de dados e linguagens de programação (Java, Python ou Scala).