

# Apache Kafka

Grupo:

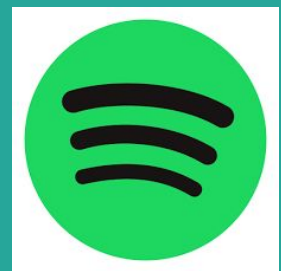
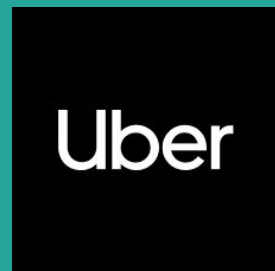
---

Ana Júlia Oliveira Lins

Lucas Rafael de Oliveira Rodrigues

Yohanna de Oliveira Cavalcanti

# Introdução



# Vantagens

- Rápido e versátil;
  - Utiliza o disco rígido ao invés da memória para armazenar os dados
  - Alta disponibilidade;
  - Transmissões assíncronas;
  - Sua funcionalidade de mensageria já possui filas integradas por padrão;
  - Garantia de entrega das mensagens ao consumer.
-

# Desvantagens

- As mensagens não são entregues de maneira ordenada;
  - Exige uma conexão de internet forte. Isso porque pequenas oscilações podem afetar a comunicação.
  - Curva de aprendizado e complexidade nas configurações.
-

# *Qual problema a tecnologia resolve?*

Exemplo:

A ingestão de dados em tempo real é um problema comum em análises em tempo real, pois em uma plataforma como o e-commerce, os usuários ativos em um determinado momento e a quantidade de eventos criados por cada usuário ativo são muitos. Portanto, as recomendações (ou seja, previsões) para cada evento ou grupos de eventos devem ser quase em tempo real.

# Arquitetura do Apache Kafka

- Mensagens
- Tópicos
- Producer
- Consumer
- Apache Zookeeper
- Kafka Brokers | Kafka Clusters

# Como o Apache Kafka cuida das falhas

- A proteção do Apache Kafka contra falhas ocorre por meio de replicação de dados;
- A unidade de replicação utilizada pelo Apache Kafka é a partição;
- Em um cluster, um broker sempre é designado como líder de uma ou mais partições;
- Os brokers secundários sempre são atualizados a partir do líder;
- Cada réplica de dados está armazenada em um broker diferente;
- O broker líder é o responsável por executar as escritas e leituras na lista de in-sync réplicas, que são replicadas para os brokers seguidores.

# Como o Apache Kafka funciona

**“ Não tenha medo do sistema de armazenamento “**

Podemos simplificar o fluxo básico do Kafka, que pode ser resumido em 5 ações bem simples:

1. O produtor produz uma mensagem.
2. Essa mensagem é colocada em um tópico.
3. Cada tópico terá uma ou várias partições
4. A mensagem mandada pelo produtor deve ter uma chave, que definirá em qual partição ficará
5. E o consumidor consome essa mensagem da partição.



# Alternativa ao Apache Kafka

## RabbitMQ

O RabbitMQ é um intermediário de mensagens, ele dará a sua aplicação uma plataforma comum para mandar e receber mensagens, além de mantê-las em um local até serem consumidas e descartadas.



# Casos de uso

Uber - reúne dados de usuários, táxis e viagens em tempo real para calcular e prever a demanda e calcular preços de pico em tempo real

LinkedIn - Previne spam e coleta interações do usuário para fazer melhores recomendações de conexão em tempo real

Empresas de Logística podem utilizar para acompanhar carros, caminhões, navios em tempo real

...

# Conclusão

- ★ Pode agir como um buffer para que os sistemas não travem
- ★ Redução da necessidade de múltiplas integrações

...

# Referências

<https://kafka.apache.org/>

<https://medium.com/@gabrielqueiroz/o-que-%C3%A9-esse-tal-de-apache-kafka-a8f447cac028>

<https://luby.com.br/infraestrutura/o-que-e-apache-kafka/>

<https://www.confluent.io/blog/how-kafka-is-used-by-netflix/>

<https://www.tibco.com/pt-br/reference-center/what-is-apache-kafka>

<https://www.infoq.com/br/articles/apache-kafka-licoes/>

<https://fullcycle.com.br/apache-kafka-trabalhando-com-mensageria-e-real-time/>