

SISTEMAS DISTRIBUIDOS



KAFKA

João Vitor, Júlia Miranda e Myllene Silva

TÓPICOS



1.
APACHE
KAFKA

2.
CASOS DE
USO

3.
DESVANTAGENS

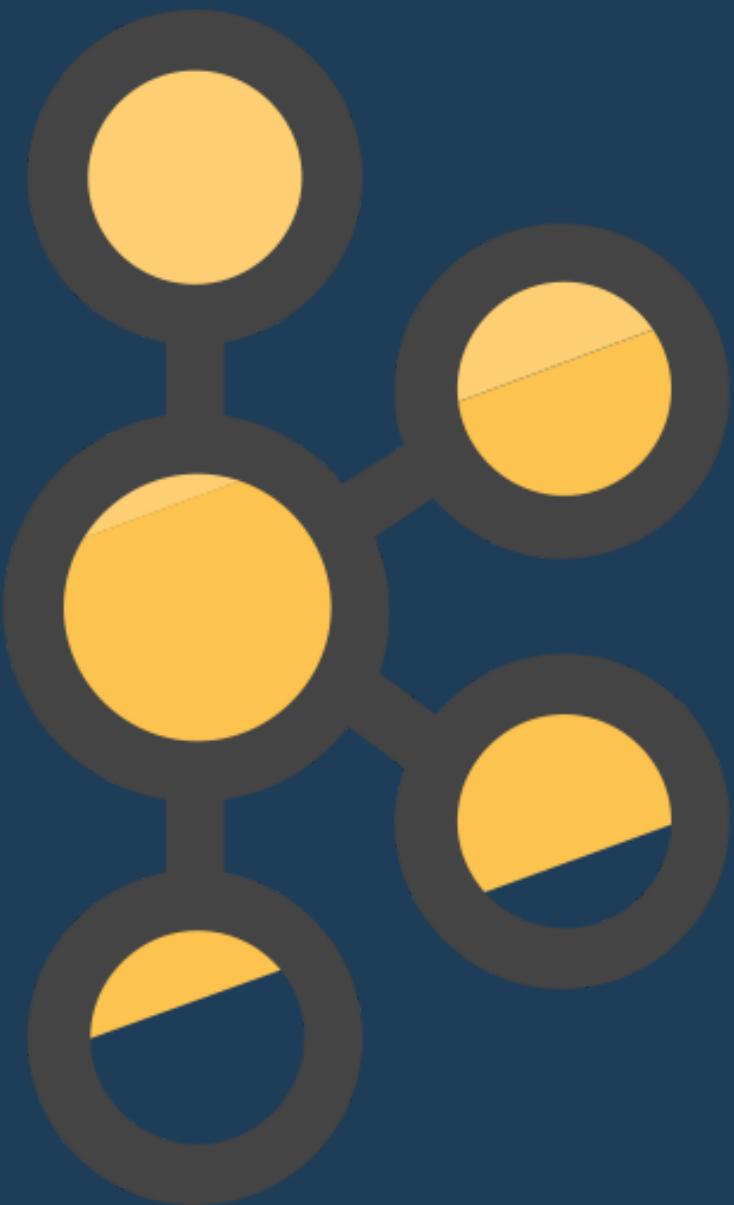
4.
CURIOSIDADES

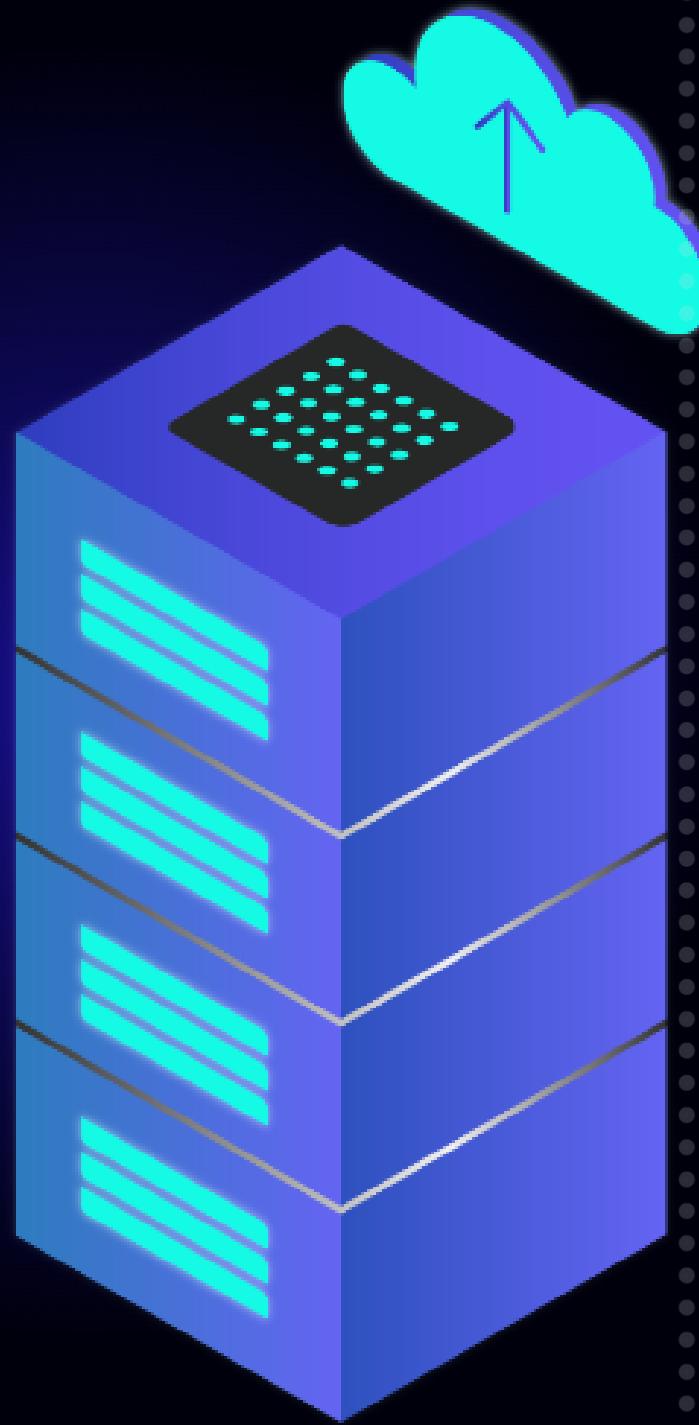
5.
ARQUITETURA

6.
IMPLEMENTAÇÃO

INTRODUÇÃO

O Apache Kafka é uma plataforma distribuída de transmissão de dados que é capaz de publicar, subscrever, armazenar e processar fluxos de registro em tempo real.





Apache Kafka

- Open source de Streaming de Eventos;
- Solução assíncrona, tolerante a falhas e escalável;
- Integração com diversas tecnologias de BD (Postgres, AWS, etc.);
- Mais de 1.000 casos de uso.

CASOS DE USO



GLOBO

Kafka é usado na Globo para enfileirar e transmitir dados em tempo real e em lote para nossa ingestão de eventos de Big Data.



O NEW YORK TIMES

Apache Kafka e a API Kafka Streams são usados para armazenar e distribuir, em tempo real, o conteúdo publicado para os diversos aplicativos e sistemas que o disponibilizam aos leitores.



NETFLIX

Monitoramento em tempo real e pipeline de processamento de eventos .

CASO DE USO

Hoje, o Kafka é usado por milhares de empresas, incluindo mais de 80% da Fortune 100.



DESVANTAGENS

MENSAGENS FORA DE ORDEM

EXIGE UMA CONEXÃO COM INTERNET
FORTE

COMPLEXIDADE NAS CONFIGURAÇÕES

CURIOSIDADES



1

Inicialmente, era um sistema interno do LinkedIn para processar 1,4 trilhões de mensagens.

2

Em novembro de 2014, Jun Rao, Jay Kreps, e Neha Narkhede, que trabalharam com o Kafka no LinkedIn, criaram uma nova empresa chamada Confluent com foco em Kafka

3

De acordo com um post no Quora de 2014, Kreps escolheu o nome do software, em referência ao autor Franz Kafka, porque ele é "um sistema otimizado para escrever", e ele gostava do trabalho de Kafka.

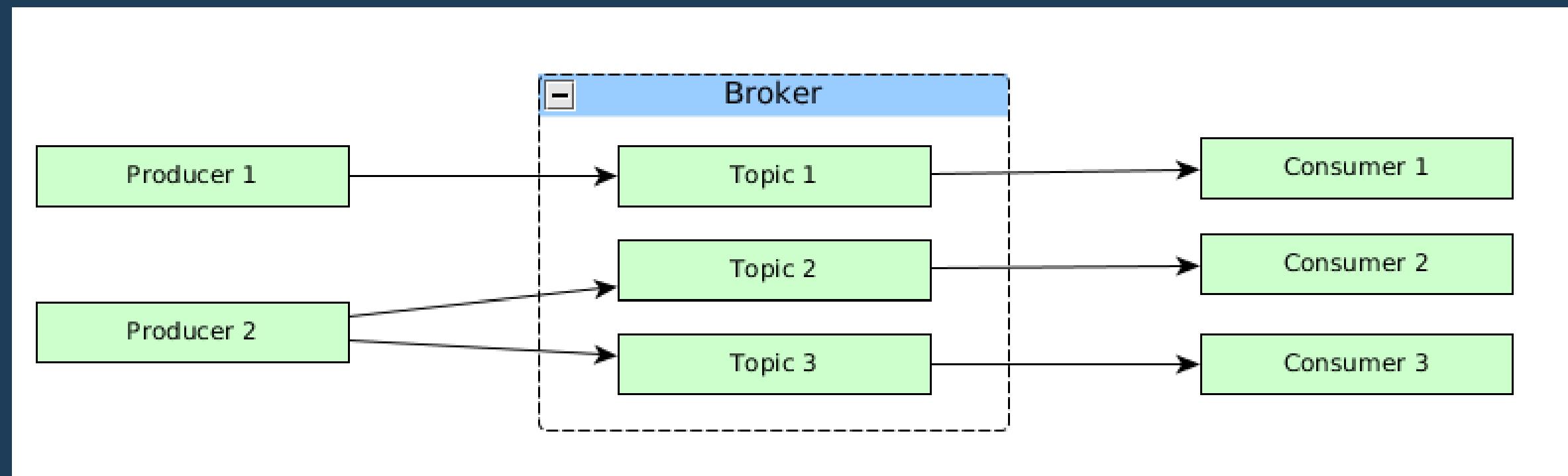


ARQUITETURA

A arquitetura do Kafka é composta por consumers, producers, topics e o cluster (chamado de broker).



ARQUITETURA



PRODUCERS

É quem envia mensagens para o cluster.

BROKERS

Servidor do Kafka que recebe as mensagens e gerencia os tópicos.

CONSUMERS

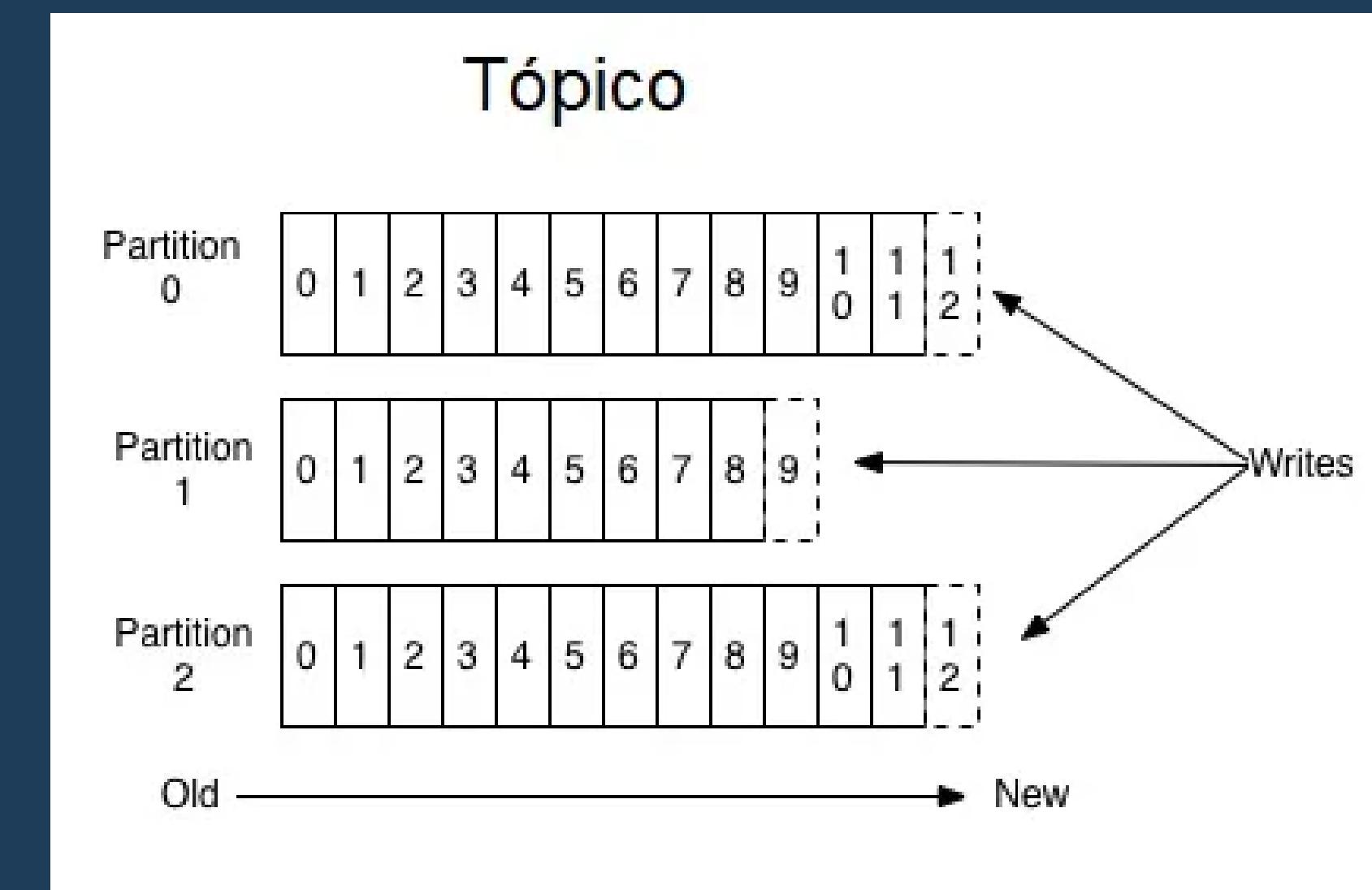
Recebe e le as mensagens do cluster.

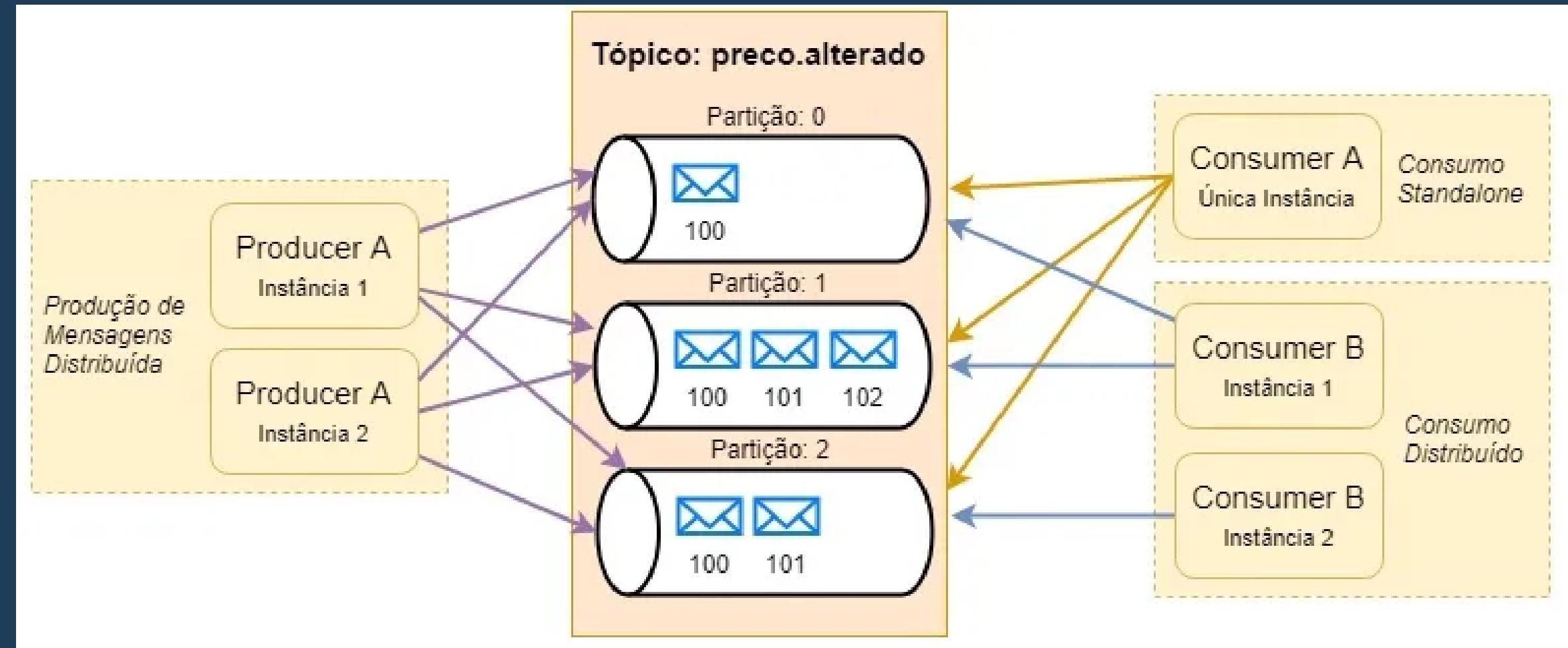
ARQUITETURA



TOPICS

Um tópico é constituído de no mínimo uma partição e zero ou mais réplicas seguidoras.





Visão Geral

Tópicos são particionados, permitindo leitura simultânea de um mesmo tópico, resultando em uma alta taxa de processamento. Os producers enviam as mensagens para o tópico e os consumers acessam a informação de cada tópico.

IMPLEMENTAÇÃO



REFERÊNCIAS

- <https://kafka.apache.org/>
- https://pt.wikipedia.org/wiki/Apache_Kafka
- https://pt.wikipedia.org/wiki/Franz_Kafka



OBRIGADA!

Alguma dúvida?