

# Plataformas Programação Distribuída



Programação distribuída é uma abordagem para desenvolver sistemas em que componentes de software são executados em diferentes máquinas ou ambientes.

Essa abordagem é usada para criar sistemas escaláveis, resilientes e que podem lidar com grandes volumes de dados e alta demanda.

# Apache Hadoop



**Descrição:** Um framework open-source que permite o processamento e armazenamento distribuído de grandes volumes de dados usando um cluster de máquinas.

## Componentes Principais:

- **HDFS (Hadoop Distributed File System):** Sistema de arquivos distribuído que armazena dados em grandes blocos replicados.
- **MapReduce:** Modelo de programação para processamento paralelo e distribuído de grandes conjuntos de dados.
- **YARN (Yet Another Resource Negotiator):** Gerencia recursos e agendamento de tarefas no cluster.

**Aplicação:** Processamento de big data, análise de dados e armazenamento distribuído.

# Apache Spark



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

**Descrição:** Um framework de processamento de dados em memória que permite processamento rápido e distribuído.

## **Características:**

- **RDD (Resilient Distributed Dataset):** Estrutura de dados fundamental do Spark que fornece tolerância a falhas e processamento paralelo.
- **Spark SQL:** Integra SQL com o processamento distribuído.
- **MLlib:** Biblioteca de aprendizado de máquina distribuído.
- **GraphX:** Computação gráfica e análise de grafos distribuídos.

**Aplicação:** Processamento de big data, análise em tempo real, aprendizado de máquina e gráficos.

# Kubernetes



**Descrição:** Uma plataforma open-source para automação de implantação, escalabilidade e gerenciamento de aplicativos em contêineres.

## **Características:**

- **Orquestração de Contêineres:** Gerencia a execução de contêineres em um cluster de máquinas.
- **Escalabilidade Horizontal:** Escala aplicativos facilmente adicionando ou removendo contêineres.
- **Gerenciamento de Estado:** Garante que o estado desejado do sistema seja mantido.

**Aplicação:** Gerenciamento de aplicativos em contêineres, implantação de microserviços, e orquestração de serviços distribuídos.

# Apache Kafka



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

**Descrição:** Uma plataforma de streaming distribuído para processamento e transmissão de dados em tempo real.

## Componentes Principais:

- **Producer:** Envía dados para tópicos.
- **Consumer:** Lê dados dos tópicos.
- **Broker:** Armazena dados e os distribui para consumidores.

**Aplicação:** Processamento de eventos em tempo real, integração de sistemas, e pipelines de dados.

# RabbitMQ



**Descrição:** Um sistema de mensagens open-source baseado no protocolo AMQP (Advanced Message Queuing Protocol).

## **Características:**

- **Mensageria Assíncrona:** Permite comunicação assíncrona entre componentes distribuídos.
- **Suporte a Vários Protocolos:** AMQP, MQTT, STOMP.
- **Alta Disponibilidade:** Suporte a clustering e replicação.

**Aplicação:** Envio e recebimento de mensagens, integração de sistemas, e processamento de eventos.

# Amazon Web Services (AWS)



**Descrição:** Uma plataforma de serviços em nuvem que oferece uma ampla gama de serviços para computação, armazenamento e redes.

## Serviços Relacionados:

- **Amazon EC2:** Instâncias de computação escalável.
- **Amazon S3:** Armazenamento de objetos distribuído.
- **Amazon RDS:** Banco de dados relacional gerenciado.
- **Amazon ECS/EKS:** Orquestração de contêineres com Docker e Kubernetes.

**Aplicação:** Desenvolvimento de aplicativos escaláveis, armazenamento em nuvem, e análise de dados.

# Microsoft Azure



**Descrição:** Plataforma de nuvem da Microsoft que fornece serviços para computação, armazenamento e redes.

## Serviços Relacionados:

- **Azure Virtual Machines:** Máquinas virtuais escaláveis.
- **Azure Blob Storage:** Armazenamento de objetos.
- **Azure Kubernetes Service (AKS):** Orquestração de contêineres com Kubernetes.
- **Azure Service Bus:** Mensageria e integração de sistemas.

**Aplicação:** Aplicativos em nuvem, integração de sistemas, e análise de dados.



# Google Cloud Platform (GCP)



**Descrição:** Plataforma de serviços em nuvem do Google que oferece infraestrutura, análise de dados e serviços de IA.

## **Serviços Relacionados:**

- **Google Compute Engine:** Máquinas virtuais escaláveis.
- **Google Cloud Storage:** Armazenamento de objetos.
- **Google Kubernetes Engine (GKE):** Orquestração de contêineres com Kubernetes.
- **Pub/Sub:** Sistema de mensageria distribuído.

**Aplicação:** Desenvolvimento de aplicativos escaláveis, processamento de big data, e serviços de IA.