

📁 - Kafka 📁 📁 📁 📁

- 1. Kafka 📁 📁
 - Kafka 📁 📁 📁
- 2. Kafka 📁
 - Kafka, Zookeeper 📁 📁 📁

1. Kafka □ □

1. Kafka 简介

Kafka 是什么

Kafka 简介

Kafka 是一个分布式流处理平台，它允许你以高吞吐量接收、存储和处理大量数据。

Kafka 是一个开源项目。

1. Kafka 的架构

1.1. 代理(Broker)

- Kafka 由多个代理组成，每个代理负责存储和处理数据。
- 每个代理都是一个 Kafka 实例。

1.2. 主题(Topic)

- 主题是 Kafka 中数据的逻辑集合。
- 每个主题由一个或多个分区组成。

1.3. 分区(Partition)

- 分区是 Kafka 中的最小数据单元。
- 每个分区由一个或多个副本组成。

1.4. 复制(Replication)

- 复制是指将数据副本存储在多个代理上。
- `replication-factor` 参数指定了副本的数量。







1.5. 生产者(Producer)

- 生产者负责将数据发送到 Kafka。
- 每个生产者都是一个 Kafka 实例。

1.6. 消费者(Consumer)

- Kafka
-

1.7. ZooKeeper (Kafka 依赖 ZooKeeper)

- Kafka   
- Kafka      

2. Kafka

2. Kafka 安装

Kafka, Zookeeper 安装与配置

1. 环境准备

1.1. Java 安装

Kafka 需要 Java 环境。Java 8 及以上版本，64 位系统。

```
java -version
```

- 推荐使用 OpenJDK 或 Oracle JDK。

```
brew update
```

```
brew install openjdk@11
```

1.2. 安装 Kafka

Kafka 需要安装 `/tmp/kafka-logs` 目录。创建如下：

2. ZooKeeper 安装与配置

2.1. ZooKeeper 安装

ZooKeeper 是 Kafka 的依赖。需要安装。

```
brew install zookeeper
```

2.2. ZooKeeper 配置

ZooKeeper 需要配置 `zoo.cfg`。

```
brew services start zookeeper
```

2.3. ZooKeeper 启动

ZooKeeper 是分布式系统协调服务。

```
brew services list
```

2.4. ZooKeeper 配置文件

ZooKeeper 的配置文件位于 `/etc/zookeeper/conf/zoo.cfg`。

```
dataDir=/var/lib/zookeeper
clientPort=2181
maxClientCnxns=60
initLimit=10
syncLimit=5
server.1=localhost:2888:3888
```

- **dataDir** :
 - ZooKeeper 数据目录。
 - ZooKeeper 数据目录。
 - `/usr/local/var/run/zookeeper`
- **clientPort** :
 - ZooKeeper 客户端连接端口。
 - `2181` 端口，ZooKeeper 客户端连接端口。
- **maxClientCnxns** :
 - ZooKeeper 客户端连接数限制。
 - 每个客户端（IP 地址）最多连接数。
- **initLimit** :
 - ZooKeeper 初始化（follower）与（leader）之间的连接数限制。
 - 初始化连接数限制。
- **syncLimit** :
 - ZooKeeper 同步（leader）与（follower）之间的连接数限制。
 - 同步连接数限制。
- **server.x** :
 - ZooKeeper 服务器配置，`x` 为服务器编号。
 - `server.<ID>=<IP>:<follower-to-leader>:<leader-to-follower>`
 - `<ID>` : ZooKeeper 服务器 ID
 - `<IP>` : ZooKeeper 服务器 IP
 - `<follower-to-leader>` : follower 到 leader 的连接端口
 - `<leader-to-follower>` : leader 到 follower 的连接端口

3. Kafka 简介

3.1. Kafka 简介

```
brew install kafka
```

3.2. Kafka 安装

Kafka 安装需要 ZooKeeper，先安装 ZooKeeper。

```
brew services start kafka
```

4. 环境变量配置

4.1. 设置环境变量

```
~/.bashrc 和 ~/.zshrc 中添加以下配置。
```

```
export KAFKA_HOME=/usr/local/opt/kafka
export PATH=$PATH:$KAFKA_HOME/bin
```

4.2. 加载环境变量

```
source ~/.bashrc
```

5. Kafka 配置

5.1. 配置 Kafka

Kafka 配置文件位于 `/usr/local/etc/kafka/server.properties`。

5.2. 配置参数

```
broker.id=0
log.dirs=/tmp/kafka-logs
zookeeper.connect=localhost:2181
num.partitions=1
default.replication.factor=1
listeners=PLAINTEXT://0.0.0.0:9092
```


advertised.listeners=PLAINTEXT://localhost:9092

- **broker.id** : 唯一 ID。 集群中每个 broker 必须不同。
- **log.dirs** : Kafka 日志存储目录。
- **zookeeper.connect** : ZooKeeper 连接字符串。
- **num.partitions** : 每个主题默认分区数。
- **default.replication.factor** : 每个分区默认副本数。
- **advertised.listeners** :
 - **Kafka** : 监听器名称。 必须与 `listeners` 中的名称一致。
 - **PLAINTEXT://localhost:9092** : 本地监听地址。
 - **localhost** : 本地 IP 地址。
 - **9092** : 本地端口。
 - **PLAINTEXT** : 监听协议。
- **listeners** :
 - **Kafka** : 监听器名称。 必须与 `advertised.listeners` 中的名称一致。
 - **PLAINTEXT://0.0.0.0:9092** : 监听地址。
 - **PLAINTEXT** : 监听协议。
 - **0.0.0.0** : 监听 IP 地址。
 - **9092** : 监听端口。