

# ZooKeeper手册

Apache ZooKeeper 是一个开源的分布式协调服务，用于管理大型分布式系统中的配置信息、命名、同步和分布式服务，它通过提供简单的原语来帮助构建更可靠的分布式系统。

## 原理

核心概念：

- **节点 (Znode)**：ZooKeeper 中的数据单元，类似于文件系统中的文件和目录
- **会话 (Session)**：客户端和 ZooKeeper 服务器之间的连接
- **数据一致性**：ZooKeeper 保证在集群中数据的一致性
- **顺序一致性**：所有更新按顺序应用
- **原子性**：所有更新要么成功要么失败，没有部分完成的状态
- **单一视图**：无论客户端连接到哪个服务器，都能看到相同的视图
- **可靠性**：一旦应用了更新，它将一直存在，直到被覆盖

## 核心组件

- **Leader**：负责处理所有写请求和管理集群的同步
- **Follower**：处理读取请求，并在 Leader 失败时参与选举新的 Leader
- **Observer**：不参与选举，只处理读取请求，通常用于扩展读取能力

## 重要配置参数

- **tickTime**：基本时间单位 (ms)，用于心跳和超时计算
- **initLimit**：Follower 初始化连接到 Leader 的时间限制
- **syncLimit**：Follower 同步数据到 Leader 的时间限制
- **dataDir**：数据存储目录（快照文件）
- **dataLogDir**：日志存储目录（事务日志）
- **clientPort**：客户端连接 ZooKeeper 服务器的端口
- **server.x=hostname:peerPort:leaderElectionPort**：集群中的服务器节点 (x - myid)

## 部署方案

每台机器（最好是奇数台）：

1. 安装 Java 和 ZooKeeper
2. 创建数据目录并创建 myid 文件
3. 复制并编辑示例配置文件 (zoo\_sample.cfg)，命名为 zoo.cfg
4. 启动实例
5. 检查实例状态

## 常用命令

连接到 ZooKeeper CLI: `bin/zkCli.sh -server <hostname>:<port>`

启动 ZooKeeper: `zkServer.sh start`

停止 ZooKeeper: `zkServer.sh stop`

查看状态: `zkServer.sh status`

查看日志: `tail -f zookeeper.out`

以下命令常于 ZooKeeper 客户端中使用:

创建节点: `create /path "myData"`

获取节点数据: `get /path`

设置节点数据: `set /path "newData"`

删除节点: `delete /path`

检查节点状态: `stat /path`