**大数据实训环境安装**

**Zookeeper**

|  |  |  |
| --- | --- | --- |
|  | 文件标识： |  |
| 当前版本： | 1.0 |
| 作 者： | 马卫花 |
| 完成日期： | 2022-10 |

**目 录**

[1.Zookeeper的主要作用 3](#_Toc28541)

[2. Zookeeper与HBase的关系 3](#_Toc18881)

[3.Zookeeper的主要作用 4](#_Toc5338)

## 1.Zookeeper的主要作用

Zookeeper是一个分布式协调服务，其主要作用是为分布式系统提供可靠的协调和一致性功能。它具有以下主要作用：

* 分布式锁：Zookeeper提供了分布式锁的实现，可以帮助解决多个进程或线程之间的并发访问问题。
* 命名服务：Zookeeper可以作为命名服务，用于将节点及其相关信息命名并注册到Zookeeper的目录结构中。
* 分布式队列：Zookeeper可以实现分布式队列，用于协调多个节点之间的任务分配和处理。
* 数据发布与订阅：Zookeeper可以实现发布/订阅模式，使多个节点能够实时获取数据的变化。
* 分布式协调：Zookeeper提供了一些原语（如条件变量、屏障等）来帮助开发者实现更复杂的分布式协调算法。

总的来说，Zookeeper通过提供一致性和可靠的分布式协调服务，帮助解决了分布式系统中的各种问题，包括数据一致性、节点故障处理、分布式任务调度等。

## Zookeeper与HBase的关系

Zookeeper与HBase之间有密切的关系，可以说Zookeeper是HBase的基础设施之一。下面是它们之间关系的几个方面：

* 分布式协调：HBase是一个分布式的列式数据库，它的各个节点需要进行协调和同步才能保持一致性。Zookeeper作为一个分布式协调服务，提供了数据一致性和可靠的协调功能，帮助HBase进行节点管理、选举主节点、故障检测等。
* 主节点选举：HBase的每个RegionServer节点中有一个HMaster节点负责整个集群的管理。Zookeeper在这个过程中起到了重要的作用，它记录了集群中哪个节点是HMaster，并在发生故障时协助进行主节点的选举。
* 配置管理：HBase的配置信息也是由Zookeeper来管理的。每个HBase节点在启动时，会从Zookeeper获取最新的配置信息，确保集群中各个节点的配置保持一致。
* 故障检测与恢复：Zookeeper可以监测HBase集群中节点的存活状态，当某个节点发生故障或宕机时，Zookeeper会通知其他节点，并帮助进行故障恢复。

总的来说，Zookeeper提供了HBase所需的分布式协调和一致性服务，帮助HBase实现高可用性、数据一致性和故障恢复等功能。它是HBase集群的重要组成部分，保证了HBase的稳定运行。

## 3.Zookeeper的主要作用

1.从Zookeeper官网下载Zookeeper的安装文件，界面如下所示,下载ZooKeeper：

zookeeper-3.4.12.tar.gz

2.下载后解压下载的文件如下图所示:

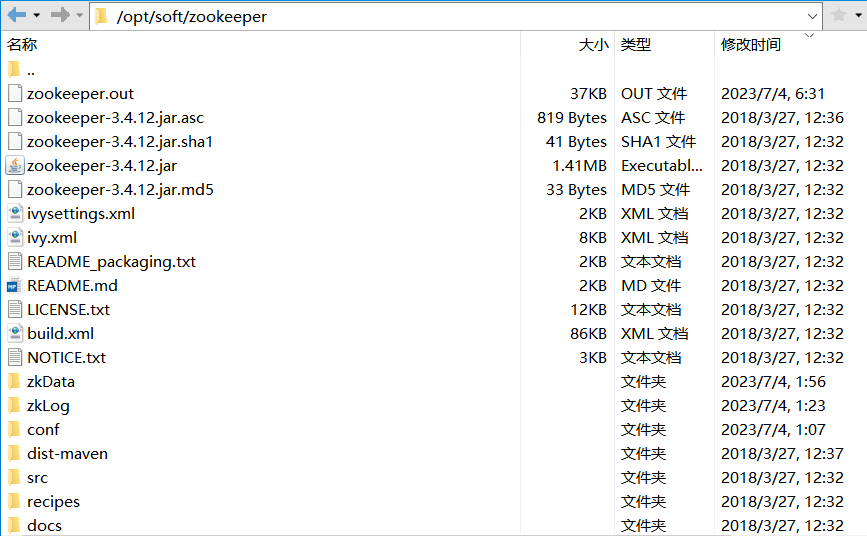
tar zxvf zookeeper-3.4.12.tar.gz

解压后的目录路径可按如下路径设置。

[root@master opt]# cd spark

[root@master soft]# ls

Zookeeper解压后的目录如下所示:



1. 配置ZooKeeper

进入ZooKeeper解压后的目录，找到conf子目录。复制zoo\_sample.cfg文件并将复制的文件重命名为zoo.cfg。打开zoo.cfg文件，根据需要进行适当的配置更改。主要的配置项包括：dataDir（ZooKeeper数据目录）、clientPort（客户端连接端口）等。可参考如下配置：

|  |
| --- |
| # The number of milliseconds of each tick  tickTime=2000  # The number of ticks that the initial  # synchronization phase can take  initLimit=10  # The number of ticks that can pass between  # sending a request and getting an acknowledgement  syncLimit=5  # the directory where the snapshot is stored.  # do not use /tmp for storage, /tmp here is just  # example sakes.  dataDir=/opt/soft/zookeeper/zkData  dataLogDir=/opt/soft/zookeeper/zkLog  # the port at which the clients will connect  clientPort=2181  server.1=hadoop:2888:3888  # the maximum number of client connections.  # increase this if you need to handle more clients  #maxClientCnxns=60  #  # Be sure to read the maintenance section of the  # administrator guide before turning on autopurge.  #  # http://zookeeper.apache.org/doc/current/zookeeperAdmin.html#sc\_maintenance  #  # The number of snapshots to retain in dataDir  #autopurge.snapRetainCount=3  # Purge task interval in hours  # Set to "0" to disable auto purge feature  #autopurge.purgeInterval=1 |

1. 启动ZooKeeper

在ZooKeeper的安装目录下，打开终端或命令提示符。使用以下命令启动ZooKeeper服务器bin/zkServer.sh start。

【over】