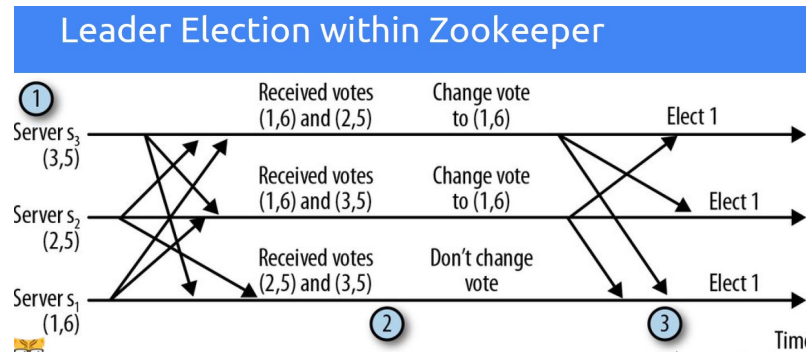


1. Q: Coordination和Scheduling有什么区别

- Coordination
 - Zookeeper, Consul, Etcd, Eureka, etc
- Scheduling
 - Mesos, Yarn, Oozie, etc

A: Coordination也就是协调，也就是帮助系统中的多个部件对某些信息达成一致，比如谁做master等。Scheduling指的是调度，也就是将任务按照比较优化的方法分出去。

2. Master 数据版本



Zookeeper中每次对data tree的操作都会导致zxid的增加。在上面这个图中的5, 6就好比是zxid。这个zxid很重要，因为它代表着谁有着更加新的数据。然后在首领选举的阶段就会根据这个信息来帮助决定推选谁为master。比如在上面的图中， s_2 和 s_3 就决定 s_1 的信息更加新，于是就改投票给 s_1 。

3. 一个partition对应一个log，是这样么？

在kafka中，每个partition其实对应的是硬盘上的一个文件夹。然后文件夹中会有日志文件。然后这个日志文件其实会分成多个小段。

4. Docker -d -p的参数意义

docker的文档中写出了这些参数的意义：

<https://docs.docker.com/engine/reference/run/>。

-d这个参数指示docker将容器跑在背景模式，如果不加上的话，启动的docker容器就会将日志输出到当前的terminal上，就类似于那个容器就跟你的terminal绑定了。一旦你的terminal退出，那个容器也会停止。

-p这个参数非常重要。在Docker中会对容器内部的端口和对外显示的端口进行一个映射，而且默认情况下是会随机挑一个端口。举个例子，比如你在容器内部运行着一个网络应用程序，需要通过8080端口访问，然后如果不加上-p这个约束的话，那个端口就可能会被随机映射到20000。-p这个参数可以帮助你固定住对外的端口。

5. 三个端口

```
docker run -d -p 2181:2181 -p 2888:2888 -p 3888:3888 --name zookeeper confluent/zookeeper
```

Zookeeper跟外界的连接主要有三个端口：2181，2888和3888。2181端口是客户端来链接用的，2888跟3888都是在集群环境下节点之间互相通信用的。

6. Zookeeper下quota的作用

```
get /zookeeper/quota
```

quota这个znode是zookeeper内部使用的一个znode。Zookeeper提供了一个给某个

znode的数据量设定一个上限的功能。这个上限也就是quota。对于各个znode的quota信息都存储在/zookeeper/quota这个znode中。

7. 学生

consume msg的时候有这个错误

```
YuchangtekiMacBook-Pro:bin yuchangchen$ ./kafka-console-consumer.sh
--zookeeper `docker-machine ip bigdata`:2181 --topic bigdata
[2016-08-23 16:57:05,294] WARN Fetching topic metadata with correlation id
0 for topics [Set(bigdata)] from broker
[BrokerEndPoint(0,9f9c8c44b562,9092)] failed (kafka.client.ClientUtils$)
java.nio.channels.ClosedChannelException
    at kafka.network.BlockingChannel.send(BlockingChannel.scala:110)
    at
```

这个问题往往是由于zookeeper没有启动或是没有提前创建topic引起的。Kafka内部将有关topic的一些元数据存储zookeeper中。然后当客户端试图链接的时候，客户端会先向zookeeper询问这个数据。这也是为什么大家必须要输入--zookeeper这个参数的原因。