

Kafka集群Leader均衡(Balancing leadership)

当一个broker停止或者crashes时，所有本来将它作为leader的分区将会把leader转移到其它broker上去。这意味着当这个broker重启时，它将不再担任任何分区的leader，kafka的client也不会从这个broker来读取消息，从而导致资源的浪费。比如下面的broker 7是挂掉重启的，我们可以发现Partition 1虽然在broker 7上有数据，但是由于它挂了，所以Kafka重新将broker 3当作该分区的Leader，然而broker 3已经是Partition 6的Leader了。

```
[iteblog@www.iteblog.com ~]$ kafka-topics.sh --topic iteblog  W  
--describe --zookeeper www.iteblog.com:2181
```

```
Topic:iteblog PartitionCount:7 ReplicationFactor:2 Configs:  
Topic: iteblog Partition: 0 Leader: 1 Replicas: 1,4 Isr: 1,4  
Topic: iteblog Partition: 1 Leader: 3 Replicas: 7,3 Isr: 3,7  
Topic: iteblog Partition: 2 Leader: 5 Replicas: 5,7 Isr: 5,7  
Topic: iteblog Partition: 3 Leader: 6 Replicas: 6,1 Isr: 1,6  
Topic: iteblog Partition: 4 Leader: 4 Replicas: 4,2 Isr: 4,2  
Topic: iteblog Partition: 5 Leader: 2 Replicas: 2,5 Isr: 5,2  
Topic: iteblog Partition: 6 Leader: 3 Replicas: 3,6 Isr: 3,6
```

幸运的是，Kafka中有一个被称为优先副本(preferred replicas)的概念。如果一个分区有3个副本，且这3个副本的优先级别分别为1，5，9，根据优先副本的概念，1会作为leader。为了使kafka集群恢复默认的leader，需要运行以下命令：

```
[iteblog@www.iteblog.com ~]$ kafka-preferred-replica-election.sh  W  
--zookeeper www.iteblog.com:2181
```

```
Successfully started preferred replica election for partitions Set([iteblog,1],  
[iteblog,5], [iteblog,4], [iteblog,6], [iteblog,2], [iteblog,0], [iteblog,3])
```

每次运行上面的命令是比较烦躁的，不过Kafka为我们提供了一个参数，可以使得Kafka集群自动平衡Leader，我们只需要在server.properties文件中配置如下设置：

```
auto.leader.rebalance.enable=true
```

这个值默认就是打开的。下面是Leader平衡的结果：

```
[iteblog@www.iteblog.com ~]$ kafka-topics.sh --topic iteblog --describe --zookeeper www.iteblog.com:2181
```

```
Topic:iteblog PartitionCount:7 ReplicationFactor:2 Configs:  
Topic: iteblog Partition: 0 Leader: 1 Replicas: 1,4 Isr: 1,4  
Topic: iteblog Partition: 1 Leader: 7 Replicas: 7,3 Isr: 3,7  
Topic: iteblog Partition: 2 Leader: 5 Replicas: 5,7 Isr: 5,7  
Topic: iteblog Partition: 3 Leader: 6 Replicas: 6,1 Isr: 1,6  
Topic: iteblog Partition: 4 Leader: 4 Replicas: 4,2 Isr: 4,2  
Topic: iteblog Partition: 5 Leader: 2 Replicas: 2,5 Isr: 5,2  
Topic: iteblog Partition: 6 Leader: 3 Replicas: 3,6 Isr: 3,6
```

可以看出broker 7重新变成Partition 1的Leader了。

本博客文章除特别声明，全部都是原创！
原创文章版权归过往记忆大数据（[过往记忆](#)）所有，未经许可不得转载。
本文链接: [【】](#) ()