

如何手动更新Kafka中某个Topic的偏移量

本文将介绍如何手动更新Kafka存在Zookeeper中的偏移量。我们有时候需要手动将某个主题的偏移量设置成某个值，这时候我们就需要更新Zookeeper中的数据了。Kafka内置为我们提供了修改偏移量的类：kafka.tools.UpdateOffsetsInZK，我们可以通过它修改Zookeeper中某个主题的偏移量，具体操作如下：

```
[iteblog@www.iteblog.com ~]$ bin/kafka-run-class.sh kafka.tools.UpdateOffsetsInZK
USAGE: kafka.tools.UpdateOffsetsInZK$ [earliest | latest] consumer.properties topic
```

在不输入参数的情况下，我们可以得知kafka.tools.UpdateOffsetsInZK类需要输入的参数。我们的consumer.properties文件配置内容如下：

```
zookeeper.connect=www.iteblog.com:2181

# timeout in ms for connecting to zookeeper
zookeeper.connection.timeout.ms=6000

#consumer group id
group.id=group
```

这个工具只能把Zookeeper中偏移量设置成earliest或者latest，如下：

```
[iteblog@www.iteblog.com ~]$ bin/kafka-run-class.sh kafka.tools.UpdateOffsetsInZK \W
  earliest config/consumer.properties iteblog
updating partition 0 with new offset: 276022922
updating partition 1 with new offset: 234360148
updating partition 2 with new offset: 157237157
updating partition 3 with new offset: 106968019
updating partition 4 with new offset: 80696130
updating partition 5 with new offset: 317144986
updating partition 6 with new offset: 299182459
updating partition 7 with new offset: 197012246
updating partition 8 with new offset: 230433681
updating partition 9 with new offset: 120971431
updating partition 10 with new offset: 51200673
updated the offset for 11 partitions
```

在有些场景下，这个工具不满足我们的需求，我们需要的是能够手动设置分区的偏移量为任何有意义的值，而不仅仅是earliest或者latest。那咋办？

我们都知道，Kafka topic的偏移量一般都是存储在Zookeeper中，具体的路径为/consumers/[groupId]/offsets/[topic]/[partitionId]，比如iteblog主题分区10的偏移量获取如下：

```
[zk: www.iteblog.com(CONNECTED) 7] get /consumers/group/offsets/iteblog/10
70332526
cZxid = 0x1ec272a4c
ctime = Tue Apr 12 19:15:19 CST 2016
mZxid = 0x256b4306a
mtime = Tue Apr 19 18:55:34 CST 2016
pZxid = 0x1ec272a4c
cversion = 0
dataVersion = 1768
aclVersion = 0
ephemeralOwner = 0x0
dataLength = 8
numChildren = 0
```

所以，我们可以通过set命令来设置某个分区的偏移量，如下；

```
[zk: www.iteblog.com(CONNECTED) 11] set /consumers/group/offsets/iteblog/10 1024
cZxid = 0x1ec272a4c
ctime = Tue Apr 12 19:15:19 CST 2016
mZxid = 0x256ca2bd7
mtime = Tue Apr 19 19:03:39 CST 2016
pZxid = 0x1ec272a4c
cversion = 0
dataVersion = 1771
aclVersion = 0
ephemeralOwner = 0x0
dataLength = 8
numChildren = 0
```

这样我们就将iteblog主题的分区10的偏移量设置成1024了。

本文提供的两种方式用于更新Zookeeper中Topic的偏移量要么不能满足我们的需求（使用kafka.tools.UpdateOffsetsInZK），要么就是太麻烦了很容易出错（直接通过Zookeeper客户端更新），后期我将会介绍另外一种方式来更新Kafka中Topic的偏移量，欢迎关注。

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: [【】（）](#)