

kafka偏移量重置

在Kafka Version为0.11.0.0之后，Consumer的Offset信息不再默认保存在Zookeeper上，而是选择用Topic的形式保存下来。

在命令行中可以使用kafka-consumer-groups的脚本实现Offset的相关操作。

更新Offset由三个维度决定：Topic的作用域、重置策略、执行方案。

Topic的作用域

- `--all-topics`：为consumer group下所有topic的所有分区调整位移
- `--topic t1 --topic t2`：为指定的若干个topic的所有分区调整位移
- `--topic t1:0,1,2`：为指定的topic分区调整位移

重置策略

- `--to-earliest`：把位移调整到分区当前最小位移
- `--to-latest`：把位移调整到分区当前最新位移
- `--to-current`：把位移调整到分区当前位移
- `--to-offset`：把位移调整到指定位移处
- `--shift-by N`：把位移调整到当前位移 + N处，注意N可以是负数，表示向前移动
- `--to-datetime`：把位移调整到大于给定时间的最早位移处，datetime格式是yyyy-MM-ddTHH

ss.xxx，比如2017-08-04T00:00:00.000

- `--by-duration`：把位移调整到距离当前时间指定间隔的位移处，duration格式是PnDTnHnMnS，比如PT0H5M0S
- `--from-file`：从CSV文件中读取调整策略

确定执行方案

- 什么参数都不加：只是打印出位移调整方案，不具体执行
- `--execute`：执行真正的位移调整
- `--export`：把位移调整方案按照CSV格式打印，方便用户成csv文件，供后续直接使用

注意事项

1. consumer group状态必须是inactive的，即不能是处于正在工作中的状态
2. 不加执行方案，默认是只做打印操作

常用示例

更新到当前group最初的offset位置

```
1. bin/kafka-consumer-groups.sh --bootstrap-server localhost:9092 --group test-group --reset-offsets --all-topics --to-earliest --execute
```

更新到指定的offset位置

```
1. bin/kafka-consumer-groups.sh --bootstrap-server localhost:9092 --group test-group --reset-offsets --all-topics --to-offset 500000 --execute
```

更新到当前offset位置（解决offset的异常）

```
1. bin/kafka-consumer-groups.sh --bootstrap-server localhost:9092 --group test-group --reset-offsets --all-topics --to-current --execute
```

offset位置按设置的值进行位移

```
1. bin/kafka-consumer-groups.sh --bootstrap-server localhost:9092 --group test-group --reset-offsets --all-topics --shift-by -100000 --execute
```

offset设置到指定时刻开始

```
1. bin/kafka-consumer-groups.sh --bootstrap-server localhost:9092 --group test-group --reset-offsets --all-topics --to-datetime 2017-08-04T14:30:00.000
```

查看消费组状态

```
1. /opt/kafka/bin/kafka-consumer-groups.sh --bootstrap-server 127.0.0.1:9092 --group source_nginx --describe
```

查看kafka topic消息数量

```
1. #time -1 当前topic 当前分区最大位移 总消息
2. kafka-run-class.sh kafka.tools.GetOffsetShell --broker-list localhost:9092 --topic topic_name --time -1
3. #time -2 当前topic 当前分区最早位移 历史消息
4. kafka-run-class.sh kafka.tools.GetOffsetShell --broker-list localhost:9092 --topic topic_name --time -2
5. #分区当前的消息总数 time-1] - [time-2]
```