写程序的少年，走在市井之间

Moving Kafka Replica Between Disk

Posted on 2016-12-27 | In [infra](#) |

As we know, one Kafka partition replica resides under one directory. When the replica size grows, the directory maybe reach the limit of disk. And Kafka can't handle this automatically, when disk is full, IOException occurs.

Fundamentals

So we have to do this by hand. Before we start, let's take a look at some fundamentals about kafka's log partiton

There are four files under each kafka data directory(path config in `log.dirs`), in addition to the replica log data directory.

- `cleaner-offset-checkpoint`
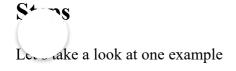
- `recovery-point-offset-checkpoint`

  As its name suggests, this file is used when broker restart, initialize the internal state of `LogManager`. Each time broker flush log, current `LogEndOffset` will be written to the file.

- `replication-offset-checkpoint`

  This file records the high watermark of partitions. The high watermark is the offset of the last message that was successfully copied to all of the log's replicas

- `meta.properties` holds the version and id of broker

Difference between Log end offset(LEO) and High Watermark(HW) is shown below. More details can be found [here](#).

Steps

Let's take a look at one example

```
~$ du -hd 1 /disk12/kafka-logs/
176G    /disk12/kafka-logs/midsource-fsrcproxy1-9
90G     /disk12/kafka-logs/dns-formated-log-48
270G    /disk12/kafka-logs/metrics-15
...
```

now we want to move partition `metrics-15` to another disk, for example `/disk7/kafka-logs`

Step 0: Stop the broker

Step 1: Migrating meta in recovery file

```
~$ cat /disk12/kafka-logs/recovery-point-offset-checkpoint
0
41
...
metrics 15 185980984
...
```

Main format of this file is `<topic> <partition id> <LogEndOffset>`, find our target partition, cut the line, and paste the line into `/disk7/kafka-logs/recovery-point-offset-checkpoint`

The second line shows how many partition under this log data dir. After we move partition, it should be reduced by one, so we change it from `41` to `40`.

At the same time the second line number in `/disk7/kafka-logs/recovery-point-offset-checkpoint` should be added one

Step 2: Migrating meta in replica offset file

Similar to step 1, except that the file is changed to `replication-offset-checkpoint`

Step 3: Move partition data

```
mv /disk12/kafka-logs/metrics-15 /disk7/kafka-logs/metrics-15
```

Step 4: Restart the broker

[# kafka](#)
[记一次写磁盘的缓存优化](#)
[利用 Shapley Value 计算带宽成本](#)

- Table of Contents
- Overview

Simon Xia

Simon Xia

[20 posts](#)
[4 categories](#)
[31 tags](#)
[RSS](#)
[GitHub](#) [StackOverflow](#) [Zhihu](#) [E-Mail](#)
Creative Commons

. Fundamentals

Powered by [Hexo](#)
|
Theme — [NexT.Muse](#)
0%