How to calculate how much memory needs Kafka Broker?

0

How to calculate how much memory needs Kafka Broker having given troughput, average message size and retention time.

I am asking, as I see in our monitoring system, that after starting, Kafka broker lineary increases used JVM heap. At some point we got OutOfMemoryError.

We extend JVM Heap couple of times and the problem remains.

So it would be better not to guess, what Kafka needs by experimenting, but to calculate the value :-).

java memory-management configuration apache-kafka out-of-memory

1 Answer

2

It's hard to calculate a precise value but you can use the rule of thumb from the Kafka documentation:

> You need sufficient memory to buffer active readers and writers. You can do a back-of-the-envelope estimate of memory needs by assuming you want to be able to buffer for 30 seconds and compute your memory need as write_throughput*30.

Also since 1.0.0, you can use `queued.max.request.bytes` to limit how much memory brokers can use for incoming requests. That can help you avoid OOM and if you can monitor its usage with:

```
kafka.network:type=SocketServer,name=MemoryPoolAvailable
kafka.network:type=SocketServer,name=MemoryPoolUsed
```

That should allow you to determine how much memory your workload requires without too many issues.

I have quite low throughput, but the rule does not work for me - somehow. Just numbers from 3 installations I have (two Kafka nodes each), which I had measured by monitoring JMX information about used heap. 1. Trougput 3,8 MB/s. Expected memory usage: 114MB, Min used heap in 1300 MB, 4200MB in peak. 2. Troughput 1,0 MB/s. Expected memory usage: 30MB, Min used heap 400MB, 1000MB in peak. 3 Throughput 0,3 MB/s. Expected memory usage: 9 MB, Min used heap 330 MB, 1000MB in peak. So there is bias ~330 MB to be added, and also regression lead through min-s gives coeff: 320*throughput and not 30 – Seweryn Habdank-Wojewódzki Jan 19 '18 at 7:57