Kafka—Accelerating!

Quick tips and insights on how to make *Apache Kafka* work faster!

Hardware

- CPU doesn't matter that much.
- Memory helps a lot (a lot) in performance.
- SSDs are not required, since most operations are sequential read and writes.
- If possible run in *bare metal*.

Linux

- Configure to maximize memory usage (tweak until you feel comfortable):

```
vm.dirty_background_ratio = 5

vm.dirty_ratio = 80

vm.swappiness = 1
```

- Assuming you are using *ext4*, don't waste space with reserved blocks:

```
tune2fs -m 0 -i 0 -c -1 /dev/device
```

- Mount with *noatime*:

```
/dev/device       /mountpoint       ext4    defaults,noatime
```

- Keep an eye on the number of free *inodes*:

```
tune2fs -l /dev/device | grep -i inode
```

- Increase limits, for example, using *systemd*:

```
$ cat /etc/systemd/system/kafka.service.d/limits.conf

[Service]

LimitNOFILE=10000
```

- Tweak your network settings, for example:

```
net.core.somaxconn = 1024

net.core.rmem_max = 67108864

net.core.wmem_max = 67108864

net.ipv4.tcp_rmem = 4096 87380 33554432

net.ipv4.tcp_wmem = 4096 65536 33554432

net.ipv4.tcp_max_syn_backlog = 4096

net.ipv4.tcp_syncookies = 1
```

Kafka

- *log.dirs* accepts a comma separated list of disks and will distribute partitions across them, however:

1. Doesn't rebalance, some disks could be full and others empty.
2. Doesn't tolerate any disk failure, more info in KIP-18.
3. *Raid 10* is probably the best middle ground between performance and reliability.

- *num.io.threads,* number of I/O threads that the server uses for executing requests. You should have at least as many threads as you have disks.
- *num.network.threads*, number of network threads that the server uses for handling network requests. Increase based on number of producers/consumers and replication factor.
- Use Java 1.8 and *G1 Garbage collector*:

```
-XX:MetaspaceSize=96m

-XX:+UseG1            # use G1

-XX:MaxGCPauseMillis=20 # gc deadline

-XX:InitiatingHeapOccupancyPercent=35

-XX:G1HeapRegionSize=16M

-XX:MinMetaspaceFreeRatio=50 -XX:MaxMetaspaceFreeRatio=80
```

- *KAFKA_HEAP_OPTS,* 5–8Gb heap should be enough for most deployments, file system cache is way more important. *Linkedin* runs 5Gb heap in 32Gb RAM servers.
- *pcstat* can help understand how well the system is caching:

```
./pcstat /kafka/data/*
```

. . .

Any comments or suggestions are welcome!