

kudu性能调优

1、Kudu Tablet Server Maintenance Threads

解释：Kudu后台对数据进行维护操作，如写入数据时的并发线程数，一般设置为4，官网建议的是数据目录的3倍

参数：maintenance_manager_num_threads

1

2、Kudu Tablet Server Block Cache Capacity Tablet

解释：分配给Kudu Tablet Server块缓存的最大内存量，建议是2-4G

参数：block_cache_capacity_mb

1

3、Kudu Tablet Server Hard Memory Limit Kudu

解释：Tablet Server能使用的最大内存量，有多大，设置多大，tablet Server在批量写入数据时并非实时写入磁盘，而是先Cache在内存中，在flush到磁盘。这个值设置过小时，会造成Kudu数据写入性能显著下降。对于写入性能要求比较高的集群，建议设置更大的值（一般是机器内存的百分之80）

参数：memory_limit_hard_bytes

1

补充：

Cgroup 内存软限制，这个限制并不会阻止进程使用超过限额的内存，只是在系统内存不足时，会优先回收超过限额的进程占用的内存，使之向限定值靠拢，当进程试图占用的内存超过了cgroups的限制，会触发out of memory，导致进程被kill掉

```
memory.soft_limit_in_bytes:
```

```
## Cgroup 内存硬限制，限制该组中的进程使用的物理内存总量不超过设定值
```

```
memory.limit_in_bytes:
```

```
1
```

```
2
```

```
3
```

```
4
```

```
报错: Service unavailable: Soft memory limit exceeded (at 96.35% of capacity)
```

4、建议每个表50columns左右，不能超过300个

5、hash分区数量*range分区数量不能超过60个（1.7.0版本之后没限制了）

6、设置block的管理器为文件管理器（默认是日志服务器）

解释：并非所有文件系统格式都需要设置该选项。ext4、xfs格式支持hole punching（打孔），所以不需要设置block_manager=file，但是ext3 格式需要。可以通过df -Th命令来查看文件系统的格式。

```
参数: --block_manager=file
```

```
1
```

7、设置ntp服务器的时间误差不超过20s（默认是10s）

```
参数: max_clock_sync_error_usec=20000000
```

```
1
```

8、设置rpc的连接时长（默认是3s，建议不要设置）

参数: `--rpc_negotiation_timeout_ms=300000`

1

9、设置rpc一致性选择的连接时长（默认为1s，建议不要设置）

参数: `--consensus_rpc_timeout_ms=1000`

1

10、记录kudu的crash的信息

解释:

Kudu在Kudu遇到崩溃时，使用Google

Breakpad库来生成minidump。这些minidumps的大小通常只有几MB，

即使禁用了核心转储生成，也会生成，生成minidumps只能在Linux上建立。

minidump文件包含有关崩溃的进程的重要调试信息，包括加载的共享库及其版本，崩溃时运行的线程列表，处理器寄存器的状态和每个线程的堆栈内存副本，以及CPU和操作系统版本信息。

Minidump可以通过电子邮件发送给Kudu开发人员或附加到JIRA，以帮助Kudu开发人员调试崩溃。为了使其有用，开发人员将需要知道Kudu的确切版本和发生崩溃的操作系统。请注意，虽然minidump不包含堆内存转储，但它确实包含堆栈内存，因此可以将应用程序数据显示在minidump中。如果机密或个人信息存储在群集上，请不要共享minidump文件。

参数:

`--minidump_path=minidumps`

`--max_minidumps=9`

1

2

3

（默认是在设置的log目录下生成minidumps目录，里边包含最多9个以dmp结尾的文件，无法设置为空值，需要注意的是如果自定义minidump文件，在master不能启动的情况下，需要将该目录中的文件删除）

11、Stack WatchLog

解释：每个Kudu服务器进程都有一个称为Stack Watchdog的后台线程，它监视服务器中的其他线程，以防它们被阻塞超过预期的时间段。这些跟踪可以指示操作系统问题或瓶颈存储。通过WARN日志信息的跟踪（Trace）可以用于诊断由于Kudu以下的系统（如磁盘控制器或文件系统）引起的根本原因延迟问题。

12、kudu表如果不新建的情况下，在表中增加字段，对数据是没有影响的

参考：

http://blog.csdn.net/weixin_39478115/article/details/79036390

13、cdh设置多master

参数：--

master_addresses=hadoop4:7051,hadoop5:7051,hadoop6:7051

1

14、kudu出现启动速度特别慢

解决办法：

1、取消所有配置参数（除了资源、时间同步）

2、升级版本到kudu1.6.0

3、client必须停止（client不占用io的情况，3台机器，每台机器60G，127分区数量，启动速度3分钟）

4、查看io使用情况 `iostat -d -x -k 1 200`

1

2

3

4

参考：

http://blog.csdn.net/weixin_39478115/article/details/

15、单hash分区最大是60

16、安装kudu过程中，会要求CPU支持ssc4.2指令集，但是我们的虚拟机cpu没有这个执行集，所以无法安装

17、设置client长连接过期时间

参数：`--authn_token_validity_seconds=12960000`（150天）

1

注意：设置到tserver的配置文件中

18、tserver和master的wal和data目录要分隔（或者是目录设置为lvm卷轴）

原因：wal目录只能设置为1个

参数：`--fs_wal_dir_reserved_bytes`

1

解释：

Number of bytes to reserve on the log directory filesystem for non-Kudu usage. The default, which is represented by -1, is that 1% of the disk space on each disk will be reserved. Any other value specified represents the number of bytes reserved and must be greater than or equal to 0. Explicit percentages to reserve are not currently supported

用于非kudu都使用的日志目录文件系统的字节数，默认情况下是-1，每个磁盘上的磁盘空间的1%将被保留，指定的任何其他值表示保留的字节数，必须大于或等于0。

19、设置用户权限，能移动tablet

参数: `--superuser_acl=*`

1

20、说明文档

<http://kudu.apache.org/releases/1.5.0/doc>

21、tserver宕掉后，5分钟后没有恢复的情况下，该机器上的tablet会移动到其他机器

参数: `--follower_unavailable_considered_failed_sec=300`

1

22、超过参数时间的历史数据会被清理，如果是base数据不会被清理。而真实运行时数据大小持续累加，没有被清理。

参数: `--tablet_history_max_age_sec=900`

版权声明：本文为CSDN博主「@black」的原创文章，遵循CC 4.0 BY-SA版权协议，转载请附上原文出处链接及本声明。

原文链接：

https://blog.csdn.net/weixin_39478115/java/article/details/78