RocksDB 写入流程详解-博客-云栖社区-阿里云

笔记本: 存储
创建时间: 2018/8/25 20:40
标签: rocksDB
URL: https://yq.aliyun.com/articles/409102?spm=a2c4e.11153959.teamhomeleft.1.8WKxt7

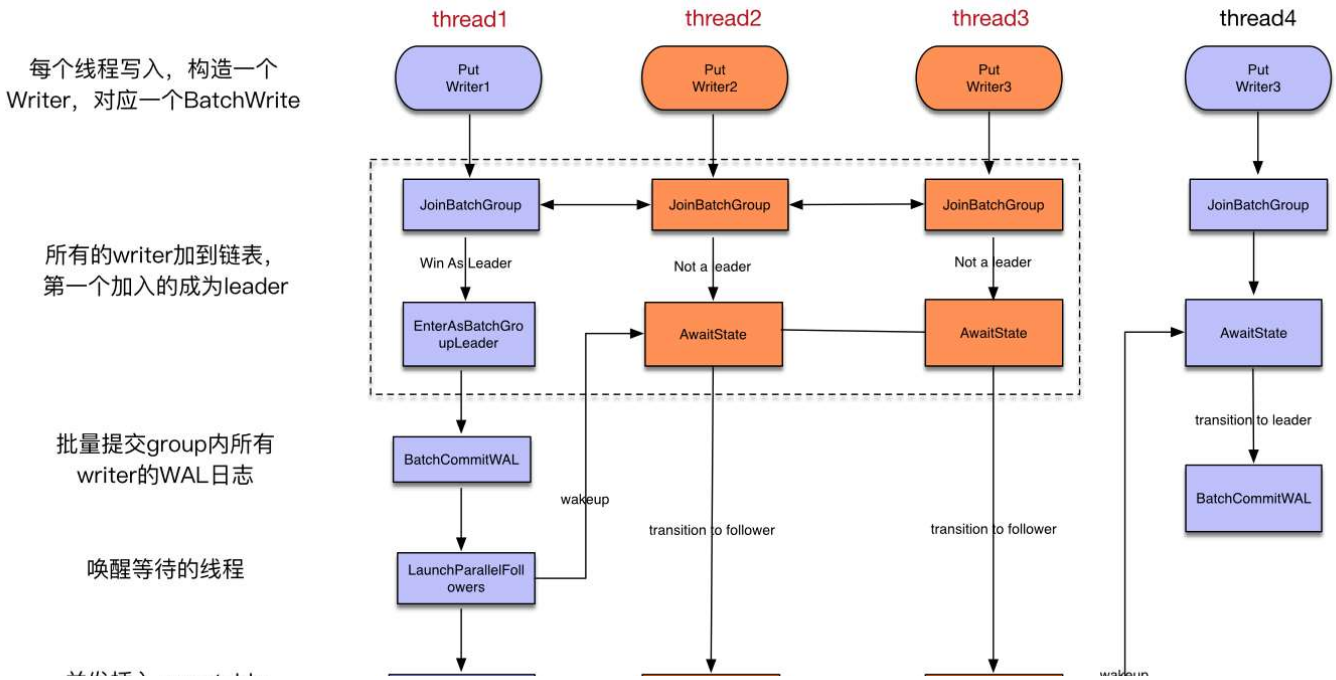
RocksDB 写入流程详解

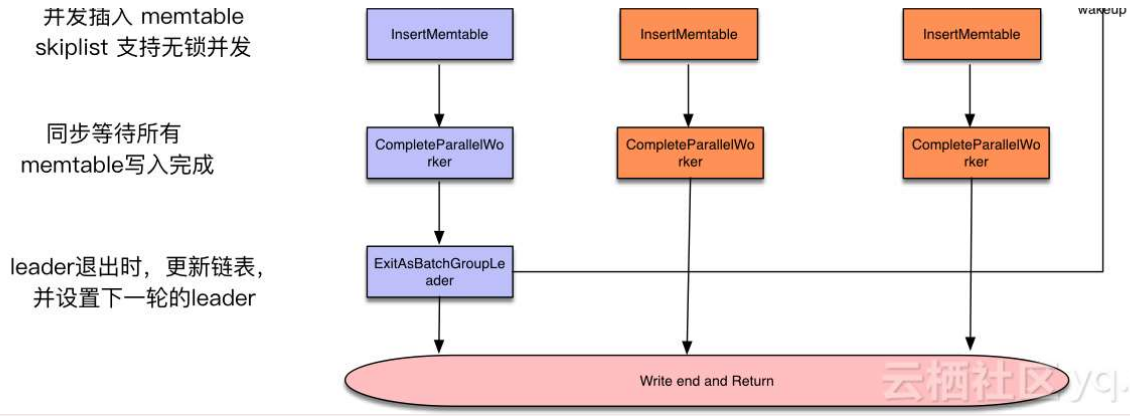
张友东 2018-01-29 11:35:40 浏览1690 评论0

性能 线程 levelDB pipeline html Group rocksdb 存储引擎 kv

摘要: 最初的写入流程，继承自 leveldb，多个 写线程组成一个 group， leader 负责 group 的 WAL 及 memtable 的提交，提交完后唤醒所有的 follwer，向上层返回。 支持 allow_concurrent_memtable_write 选项，在1的基础上，leader 提交完 WAL 后，group 里所有线程并发写 memtable。

- 最初的写入流程，继承自 leveldb，多个 写线程组成一个 group， leader 负责 group 的 WAL 及 memtable 的提交，提交完后唤醒所有的 follwer，向上层返回。
- 支持 allow_concurrent_memtable_write 选项，在1的基础上，leader 提交完 WAL 后，group 里所有线程并发写 memtable。原理如下图所示，这个改进在 sync=0的时候，有3倍写入性能提升，在 sync=1时，有2倍性能提升，参考Concurrent inserts and the RocksDB memtable
- 支持 enable_pipelined_write 选项，在2的基础上，引入流水线，第一个 group 的 WAL 提交后，在执行 memtable 写入时，下一个 group 同时开启，已到达 Pipeline 写入的效果。





上一篇：MongoDB 存储引擎 WiredTiger 原理解析

下一篇：In-place update in WiredTiger

相关文章

MyRocks写入分析

MySQL · myrocks · myrocks写入分…

MySQL · myrocks · myrocks写入分…

Flink原理与实现：详解Flink中的状态管理

RocksDB事务实现TransactionDB分析

RocksDB TransactionDB事务实现分析

【RocksDB】TransactionDB源码分析

RocksDB · 特性介绍 · HashLinkLis…

如何做到“恰好一次”地传递数十亿条消息

阿里云数据库MongoDB版正式支持3.4、RocksD…

网友评论

登录后可评论，请 登录 或 注册