

hbasehw10

描述HBase集群主要进程和后台进程的主要功能

1、HBase Master

HBase Master管理集群中很多关键的功能

- * 协调RegionServers, 管理故障恢复
 - 处理region重新分配到其他RegionServers
- * 管理表和column family的新增和修改
 - 更新hbase:meta
- * 管理region变化, 如region的切分和重新分配
- * 多个Master可以同时运行
 - 仅一个Master处于活动状态
 - 如果主master失败, 其他master 通过竞争, 其中一个变成活动状态, 接管集群

2、HBase Master 后台进程

- Hbase通过LoadBalancer进程均衡region在集群中的分布
- Catalog Janitor进程, 检查不用的region并交给垃圾回收
- Log Cleaner进程删除旧的WAL文件

3、RegionServer

- RegionServer处理数据移动
 - 为gets 和scans 命令读取所有数据并返回给客户端
 - 为puts命令保存所有数据
 - 记录所有删除数据
- 处理所有compactions
 - Major和minor compactions
- 处理所有region splitting
- 维护Block Cache
 - Catalog tables, indexes, bloom filters, 和data blocks都在Block Cache中维护
- 管理Memstore和WAL
 - Puts和Deletes先写入WAL 然后添加到Memstore
 - Memstore 产生flushed
- 从Master接收新的regions
 - 需要重写WAL

4、RegionServer 后台进程

- RegionServer有CompactSplitThread进程
 - 监测哪些regions需要split

- 超过region的最大值
- 处理minor compactions
- MajorCompactionChecker 检测是否需要执行major compaction
- MemstoreFlusher检测是否Memstore 已满, 需要执行flush 把数据写入磁盘
- LogRoller 关闭WAL创建新文件