distributede table

分布式表不存储数据,仅负责分发和查询数据,实际数据存储 依赖本地表.

本地表仅mergetree引擎家族拥有复制能力,可以完成本地表之间的数据同步

分布式表以分片未单位, 一个完整的表数据包含在不同分片中

一个分片可以包含多个物理节点,不同物理节点可以用来做数据备份,当一个分片接受到一个写入请求时,根据internal_replicated的值判断时写入当前分片下所有物理节点,还是写入任意物理节点,也就是分片的复制

replicated(没有主从之分,所有表都是副本)

所有mergetree family 都可以增加此前缀用于复制表,复制表依赖于clickhouse keeper进行元数据的管理,核心依赖于集群的分片设置,也就是分片包含了多少个物理节点,那么也就可以包含多少个副本

复制表的复制是基于表级别的,分片有自己的复制机制。

mergetree family

mergetree 核心引擎

repliacingmergetree 可使用insert into 进行更新的引擎

summingmergetree 预聚合的引擎

aggregatemergetree 扩展了smming的,常用于物化视图

collapsingmergetree 由于repliacing的合并删除的不定期性,此引擎增加sign符号字段,可通过sql合并统计

versionedcollapsingmergetree collapsing 基础上增加了版本字段用于控制多进程写入的情况

other

未完

join 用于更新部分表,无法使用select,但是需要注意where条件 set 用于in查询,无法单独使用select ,写入无法使用select memory 内存表引擎,用于临时数据处理 buffer 用于写入另一张表的临时缓存

distributede table internal_replicated= false internal_replicated= false internal_replicated= true internal_replicated= true shard shard shard shard 写入所有节点 写入所有节点 写入任意节点数据 写入任意节点数据 replicated replicated replicated replicated mergetree mergetree mergetree mergetree mergetree mergetree mergetree mergetree 错误配置,数据在任意时刻 写入任意本地表,但是本地 由于网络原因,可能出现 数据不会丢失, 但是在 官方推荐,分配到本分片 某个表未被成功写入的情 表之间没有任何同步机制, 反复同步过程中, 浪费 数据写入任意复制表, 使 况,但不会进行后续同 而读取时候会返回分片任意 资源 用本地表数据同步 步,可能导致数据不一致 本地表数据, 导致分布式表 数据错误