

知 MPP1数据分布默认是Hash分布还是随机分布?

厉梦如 2019-05-19 发表

问题描述

Q: MPP1数据分布默认是Hash分布还是随机分布?

解决方法

A: 默认是Hash分布。Hash分布就是按分布键对数据列进行hash取模存放对应的segment, 而随机分布是指数据随机分布在数据库, 每次查询都会查询所有的 segment, 因为数据随机分散在每一个节点中, 可以保证数据平均分布, 但是在执行 SQL 的过程中, 关联等操作都需要将数据重分布, 性能较差。一般很少使用随机分布。建议使用hash分布。