openGauss 安全体系创新

实践课



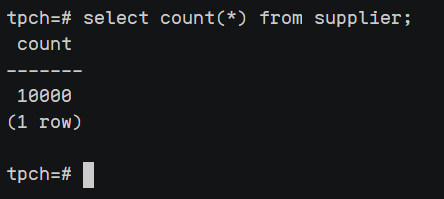
华为技术有限公司

# **关卡二、openGauss数据导入及行存列存**

任务一：数据初始化验证

1. 查询supplier表的行数，并将结果进行图：

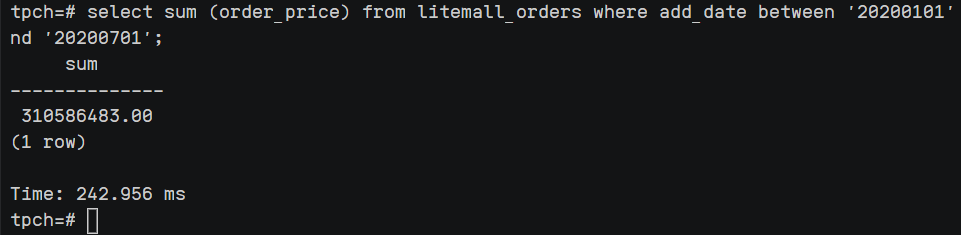
select count(\*) from supplier;;



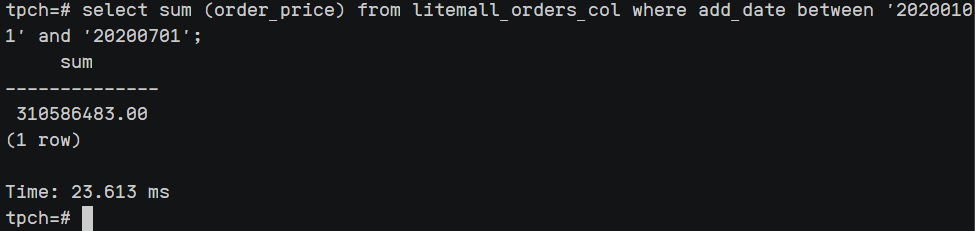
任务二：行存表与列存表执行效率对比

1. 2020年上半年litemall\_orders行存表与litemall\_orders\_col列存表中的order\_price的总和查询，并对比执行效率截图

select sum (order\_price) from litemall\_orders where add\_date between '20200101' and '20200701';

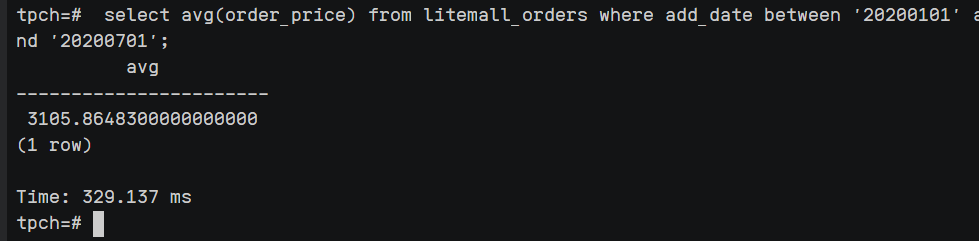


select sum (order\_price) from litemall\_orders\_col where add\_date between '20200101' and '20200701';

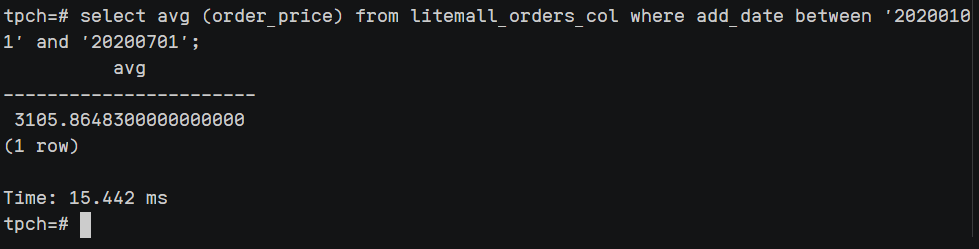


2. 2020年上半年litemall\_orders行存表与litemall\_orders\_col列存表中的order\_price的平均值查询，并对比执行效率截图

select avg (order\_price) from litemall\_orders where add\_date between '20200101' and '20200701';

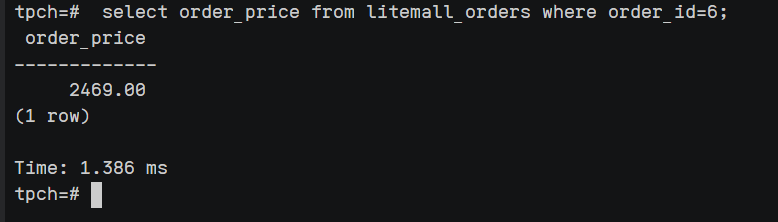


select avg (order\_price) from litemall\_orders\_col where add\_date between '20200101' and '20200701';

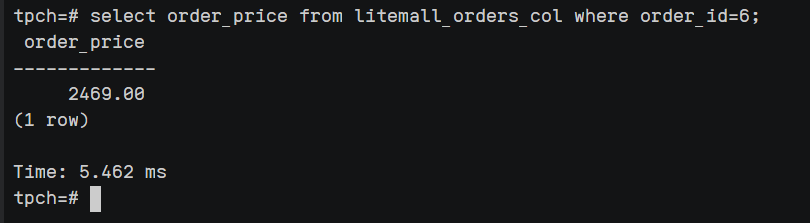


3. 查询litemall\_orders行存表与litemall\_orders\_col列存表中order\_id为6的order\_price的值，并对比执行效率截图。

select order\_price from litemall\_orders where order\_id=6;

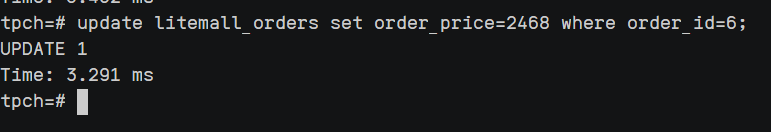


select order\_price from litemall\_orders\_col where order\_id=6;

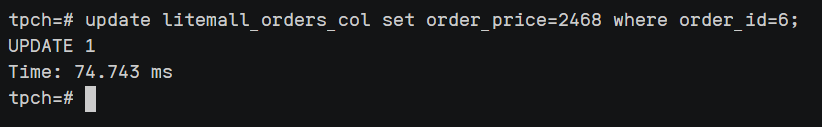


4. 将litemall\_orders行存表与litemall\_orders\_col列存表中order\_id为6的order\_price修改为2468，并对比执行效率截图。

update litemall\_orders set order\_price=2468 where order\_id=6;



update litemall\_orders\_col set order\_price=2468 where order\_id=6;



任务三：实践思考题

思考题1：

行存表与列存表在执行相同的SQL语句时，为何执行的时间不同？

行存表存储结构将数据以行为单位存储在磁盘上，而列存表则将数据以列为单位存储在磁盘上。这两种存储结构的差异决定了它们在执行查询时的执行情况。

思考题2：

在执行哪些类型SQL时，行存表效率更高？在执行哪些类型SQL时，列存表效率更高？

在执行哪些类型SQL时，行存表效率更高？在执行哪些类型SQL时，列存表效率更高？

在执行查询类 SQL 时，一般情况下行存表（Row-based Table）和列存表（Column-based Table）的效率取决于查询涉及的数据的类型以及查询操作本身。

行存表适合于如下类型 SQL：

对数据单行的增、删、改操作，例如：INSERT、UPDATE 和 DELETE。

需要使用聚合函数如 SUM、COUNT 和 AVG 的查询，由于聚合函数需要对每行数据进行计算。

在小规模数据操作时，由于每行数据内容较少，对于单条记录的增、删、改操作，行存表的 I/O 开销会更低。

而对于如下类型 SQL，一般来说对于较大数量的数据使用列存表效率会更高：

查询某个字段的值时，由于列存表将同一列数据进行整合存储在一起，因此从磁盘读取时可以减少 I/O 操作，效率更高。

需要跨多个表进行联合查询时，复杂 JOIN 操作可能需要读取大量的数据来完成关联，此时列存表会有更好的查询性能。

针对大型数据集的 OLAP（Online Analytical Processing）查询。这种查询通常需要在大数据集中调查数据的分布情况，而且通常不需要返回完整行数据。