

openGauss 数据库维护管理实验

姓名： 汤清云 学号： 2013536

实验步骤：

- openGauss 数据库安装
- 数据库性能检查实验
- 最大连接数设置实验
- 例行表、索引维护实验

实验报告

实验步骤截图：

截图 1：指导手册 25 页顺序扫描执行计划截图

```
postgres=# ANALYZE VERBOSE student;
INFO: analyzing "public.student"(dn_6001 pid=11115)
INFO: ANALYZE INFO : "student": scanned 1 of 1 pages, containing 50 live rows and 0 dead rows; 50 rows in sample, 50 estimated
total rows(dn_6001 pid=11115)
ANALYZE
postgres=# explain select * from student where std_id=30;
               QUERY PLAN
-----
Seq Scan on student  (cost=0.00..1.62 rows=1 width=62)
  Filter: (std_id = 30)
(2 rows)

postgres=#
```

截图 2：指导手册 26 页索引扫描执行计划截图

```
postgres=# explain select /*+indexscan(student student_pkey)*/ * from student where std_id=30;
               QUERY PLAN
-----
[Bypass]
Index Scan using student_pkey on student  (cost=0.00..8.27 rows=1 width=62)
  Index Cond: (std_id = 30)
(3 rows)

postgres=#
```

截图 3：将最大连接数设置为 8000 并验证设置是否成功截图（注意，指导手册中将最大连接数设置为 6000，怎样重新设置为 8000 呢？）

```
ze the memory protect for g_instance.attr.attr_storage.cstore_buffers (16 Mbytes) or shared memory (2363 Mbytes) is
=====
Successfully started.
[omm@ecs-6002 ~]$ gsql -d postgres -p 26000 -r
gsql ((openGauss 2.0.0 build 78689da9) compiled at 2021-03-31 21:03:52 commit 0 last mr )
Non-SSL connection (SSL connection is recommended when requiring high-security)
Type "help" for help.

postgres=# SHOW max_connections;
max_connections
-----
8000
(1 row)

postgres=#
```

截图 4：使用 ANALYZE VERBOSE 语句更新统计信息，并输出表的相关信息。

```
DELETE 20
postgres=# vacuum full student;
VACUUM
postgres=# analyze student;
ANALYZE
postgres=# analyze verbose student;
INFO:  analyzing "public.student"(dn_6001 pid=50714)
INFO:  ANALYZE INFO : "student": scanned 1 of 1 pages, containing 30 live rows and 20 dead rows; 30 rows in sample, 30 estimated
total rows(dn_6001 pid=50714)
ANALYZE
postgres=#
```

实验思考题：

1. 全表扫描和索引扫描的区别是什么？具体是如何实现的？比较两种扫描方式的 cost（提供查询结果截图），为什么全表扫描比索引扫描 cost 更小？在什么情况下通过主键进行查找会比全表扫描更节省时间？

全表扫描：在数据库中，对没有主键索引的表进行查询一般称为全表扫描。

全表扫描是数据库服务器用来搜寻表的每一条记录的过程，Oracle 读取表中所有的行，并检查每一行是否满足语句的 WHERE 限制条件。直到所有符合给定条件的记录返回为止。实现方式：数据库通过调用 I/O 每次读取多个数据块，

索引扫描：在执行计划中为 index scan，适合扫描整个索引的数据。类似于全表扫描(只扫描索引)，索引访问方法负责把它拿到的那些据说匹配扫描键字的所有元组之 TID 的回流。实现方法：我们先通过 index 查找到数据对应的 rowid 值(对于非唯一索引可能返回多个 rowid 值),然后根据 rowid 直接从表中得到具体的数据，一个 rowid 唯一的表示一行数据，该行对应的数据块是通过一次 I/O 得到的,在此情况下该次 I/O 只会读取一个数据库块。

全表扫描的 cost:

```
-----  
Seq Scan on student (cost=0.00..1.62 rows=1 width=62)
```

索引扫描的 cost:

```
[Bypass]  
Index Scan using student_pkey on student (cost=0.00..8.27 rows=1 width=62)  
Index Cond: (std_id = 30)
```

本题中只建立了一张表，数据容量小，所以全表扫描的 cost 小于索引扫描。

当有多张表、存储了大量数据时，索引扫描会优于全表扫描。

2. 请列举一种需要重建索引的情况和原因，并说明 openGauss 中重建索引的方式有哪些。

原因：索引不断增加，但删除的空间没有重复利用；索引的群集因子不同步；索引碎片不断增加；数据发生大量删除后，索引页面上的索引键将被删除，导致索引页面数量的减少，造成索引膨胀。重建索引可回收浪费的空间。

重建索引的方式：1. 用 REINDEX 语句重建索引 (REINDEX INTERNAL TABLE name) 2. 先删除索引 (DROP INDEX)，再创建索引 (CREATE

INDEX)。