

## openGauss 数据库维护管理实验

姓名: 蒋薇 学号: 2110957

### 实验步骤:

## ➤ openGauss 数据库安装

```

root@139.9.142.100 ~
Welcome to 4.19.90-2110.8.0.el19.oel.aarch64

System information as of time: Wed Mar 29 23:21:13 CST 2023

System load: 0.70
Processes: 144
Memory used: 10.3%
Swap used: 0.0%
Usage On: 14%
IP address: 192.168.0.235
Users online: 1

[root@ecs-cbbf ~]# su - oam
Last login: Wed Mar 29 14:08:13 CST 2023 on pts/0

Welcome to 4.19.90-2110.8.0.el19.oel.aarch64

System information as of time: Wed Mar 29 23:24:57 CST 2023

System load: 0.06
Processes: 140
Memory used: 10.3%
Swap used: 0.0%
Usage On: 14%
IP address: 192.168.0.235
Users online: 1

```

## ➤ 数据库性能检查实验

[illegible]

## ➤ 最大连接数设置实验

```

00112) is larger.
=====
Successfully started.
[omm@ecs-cb0f ~]$ gsql -d postgres -p 26888 -r
gsql ((openGauss 2.8.0 build /8688da9) compiled at 2021-05-31 21:05:52 comm
last mtr )
Non-SSL connection (SSL connection is recommended when requiring high-secu
type "help" for help.

postgres=# SHOW max_connections;
max_connections
-----
8088
(1 row)

postgres=# \q
[omm@ecs-cb0f ~]$ gsql -d postgres -p 26888 -r
gsql ((openGauss 2.8.0 build /8688da9) compiled at 2021-05-31 21:05:52 comm
last mtr )
Non-SSL connection (SSL connection is recommended when requiring high-secu
type "help" for help.

postgres=# alter system set max_connections=8088;
NOTICE: please restart the database for the POSTMASTER level parameter to
effect.
ALTER SYSTEM SET
postgres=# alter system set max_connections=8088;
NOTICE: please restart the database for the POSTMASTER level parameter to
effect.
ALTER SYSTEM SET
postgres=# \q
[omm@ecs-cb0f ~]$ gs_om -t stop;
Stopping cluster.
=====
Successfully stopped cluster.
=====
[omm@ecs-cb0f ~]$ gs_om -t start;

```

## ➤ 例行表、索引维护实验

```

postgres=#
postgres=# \d student
          Table "public.student"
   Column   |          Type          | Modifiers
-----|-----|-----
 std_id     | integer                | not null
 std_name   | character varying(20)  | not null
 std_sex    | character varying(8)   |
 std_birth  | timestamp(0) without time zone |
 std_in     | timestamp(0) without time zone | not null
 std_address | character varying(100) |

postgres=# vacuum student;
VACUUM
postgres=# vacuum student;
VACUUM
postgres=# delete from student where std_id>30;
DELETE 20
postgres=# vacuum full student;
VACUUM
postgres=# analyze student;
ANALYZE
postgres=# analyze verbose student;
INFO:  analyzing "public.student" (dp 8081 pid=40165)

```

## 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=# 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=# alter table student add primary key (std_id);
NOTICE:  ALTER TABLE / ADD PRIMARY KEY will create implicit index "student_pkey"

```

步骤 8 查看语句的执行计划。（[查询结果截图](#)）

```
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)
```

Seq Scan on student 表示使用的是全表扫描。

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=# \q
[omm@ecs-c55f ~]$
```

步骤 11 通过 hint 来优化语句扫描方式。

通过加 hint 来使查询语句进行索引扫描。（[查询结果截图](#)）

```
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=#
```

Index Scan using student\_pkey on student 表示语句通过 student 表上的主键索引 student\_pkey 进行了索引扫描。

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？）

```
(1 row)

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

postgres=# \q
[omm@ecs-c55f ~]$
```

Successfully started.

### 步骤 8 验证参数设置是否成功。(截图验证)

使用 `gsqsl` 客户端以管理员用户身份连接 `postgres` 数据库，然后查看参数值。

```
[omm@ecs-e1b3 ~]$ gsql -d postgres -6000 -rp 2
gsqsl ((openGauss 2.0.0 build 38a9312a) compiled at 2021-05-27 14:57:08 commit 472 last mr 549)
Non-SSL connection (SSL connection is recommended when requiring high-security)
Type "help" for help.
postgres=# SHOW max_connections;
max_connections
-----
6000
(1 row)
```

这里显示 `max_connections` 为 6000，说明前面参数的修改已经生效。

```
DELETE 20
postgres=# vacuum full student;
VACUUM
postgres=# analyze student;
ANALYZE
postgres=# analyze verbose student;
INFO:  analyzing "public.student"(dn_6001 pid=40166)
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=40166)
ANALYZE
```

ANALYZE

### 步骤 11 使用 `ANALYZE VERBOSE` 语句更新统计信息，并输出表的相关信息。(该步骤截图)

```
postgres=# analyze verbose student;
INFO:  analyzing " public.student " (dn_6001 pid=37195)
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=37195)
ANALYZE
```

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？

全表扫描和索引扫描的区别在于扫描的数据范围不同。全表扫描是对整张表进行扫描，而索引扫描只扫描索引所对应的数据。

全表扫描：遍历整张表，逐行进行判断和处理。

索引扫描：通过索引树找到符合条件的行，然后再通过主键或聚簇索引找到对应的数据行。

全表扫描 cost > 索引扫描 cost, 因为全表扫描需要遍历整张表，而索引扫描只需要扫描索引所对应的数据行。

因为主键是唯一的，而全表扫描需要遍历整张表，如果数据量很大，就会造成很大的开销。因此，在需要查询某个唯一值或者需要查询的数据量很小的情况下，通过主键进行查找会更加高效。

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。

数据库空间变化：当数据库的空间发生变化，如数据量增加、删除等，可能会导致索引的分布发生变化，需要重建索引。（索引碎片化：当索引中的数据被频繁地更新、删除或插入时，会导致索引碎片化，使得查询效率下降，需要重建索引。

索引过度填满：当索引的填充因子过高，导致索引树的高度变高，会导致查询效率下降，需要重建索引。）

DROP+CREATE 方式：首先使用 DROP INDEX 语句删除原有的索引，然后使用 CREATE INDEX 语句重新创建索引。这种方式效率较低，会造成索引不可用的时间。

REINDEX 方式：使用 REINDEX 语句，系统会自动重建索引，不会删除原有的索引，而是在后台创建一个新的索引，等新索引创建完成后，再将原有的索引删除，将新索引重命名为原有索引的名称。这种方式能够避免删除索引的时间，但是需要额外的磁盘空间。

CLUSTER 方式：使用 CLUSTER 语句，会重新对表进行物理排序，同时会对表上的所有索引重建。这种方式能够提高查询效率，但是需要额外的磁盘空间和时间。

VACUUM FULL 方式：使用 VACUUM FULL 语句，会对整张表进行重建，同时会对表上的索引进行重建。这种方式能够提高查询效率，但是需要额外的磁盘空间和时间，同时会锁定整张表，影响其他操作的执行。