



Факультет программной инженерии и компьютерной техники
Информационные системы и базы данных

Лабораторная работа № 4
Вариант № 1156

Преподаватель: Харитоновна Анастасия Евгеньевна
Выполнил: Геллер Леонид Александрович
Группа: P33301

Задание

По варианту, выданному преподавателем, составить и выполнить запросы к базе данных "Учебный процесс".

Команда для подключения к базе данных `ucheb`:

```
psql -h pg -d ucheb
```

Составить запросы на языке SQL (пункты 1-7).

Для каждого запроса предложить индексы, добавление которых уменьшит время выполнения запроса (указать таблицы/атрибуты, для которых нужно добавить индексы, написать тип индекса; объяснить, почему добавление индекса будет полезным для данного запроса).

Для запросов 1- 2 необходимо составить возможные планы выполнения запросов. Планы составляются на основании предположения, что в таблицах отсутствуют индексы. Из составленных планов необходимо выбрать оптимальный и объяснить свой выбор.

Изменяются ли планы при добавлении индекса и как?

Для запросов 1- 2 необходимо добавить в отчет вывод команды EXPLAIN ANALYZE [запрос]

Подробные ответы на все вышеперечисленные вопросы должны присутствовать в отчете (планы выполнения запросов должны быть нарисованы, ответы на вопросы - представлены в текстовом виде).

1. Сделать запрос для получения атрибутов из указанных таблиц, применив фильтры по указанным условиям:
Н_ТИПЫ_ВЕДОМОСТЕЙ, Н_ВЕДОМОСТИ.
Вывести атрибуты: Н_ТИПЫ_ВЕДОМОСТЕЙ.ИД, Н_ВЕДОМОСТИ.ЧЛВК_ИД.
Фильтры (AND):
 - а) Н_ТИПЫ_ВЕДОМОСТЕЙ.ИД > 3.
 - б) Н_ВЕДОМОСТИ.ИД = 1457443.
 - в) Н_ВЕДОМОСТИ.ИД < 1490007.Вид соединения: RIGHT JOIN.
2. Сделать запрос для получения атрибутов из указанных таблиц, применив фильтры по указанным условиям:
Таблицы: Н_ЛЮДИ, Н_ОБУЧЕНИЯ, Н_УЧЕНИКИ.
Вывести атрибуты: Н_ЛЮДИ.ИД, Н_ОБУЧЕНИЯ.ЧЛВК_ИД, Н_УЧЕНИКИ.ИД.
Фильтры: (AND)
 - а) Н_ЛЮДИ.ИД > 163484.

б) Н_ОБУЧЕНИЯ.ЧЛВК_ИД < 105590.

Вид соединения: LEFT JOIN.

Отчёт по лабораторной работе должен содержать:

1. Текст задания.
2. Реализацию запросов на SQL.
3. Планы выполнения запросов.
4. Ответы на вопросы, представленные в задании.
5. Выводы по работе.

Темы для подготовки к защите лабораторной работы:

1. Индексы
2. Оптимизация запросов
3. Выбор плана выполнения запросов

Реализация запросов на SQL

1. Сделать запрос для получения атрибутов из указанных таблиц, применив фильтры по указанным условиям:

Н_ТИПЫ_ВЕДОМОСТЕЙ, Н_ВЕДОМОСТИ.

Вывести атрибуты: Н_ТИПЫ_ВЕДОМОСТЕЙ.ИД, Н_ВЕДОМОСТИ.ЧЛВК_ИД.

Фильтры (AND):

а) Н_ТИПЫ_ВЕДОМОСТЕЙ.ИД > 3.

б) Н_ВЕДОМОСТИ.ИД = 1457443.

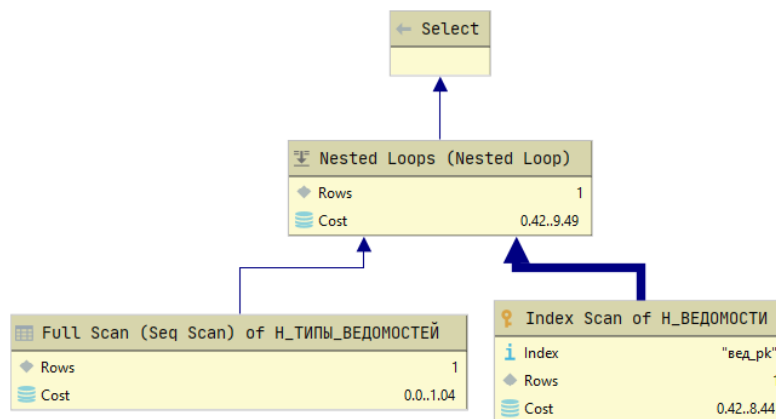
с) Н_ВЕДОМОСТИ.ИД < 1490007.

Вид соединения: RIGHT JOIN.

```
select type.ИД, ved.ЧЛВК_ИД from Н_ТИПЫ_ВЕДОМОСТЕЙ as type
right join Н_ВЕДОМОСТИ as ved on type.ИД = ved.ТВ_ИД where
type.ИД > 3 and ved.ИД = 1457443 and ved.ИД < 1490007;
```

План выполнения запроса:

```
uscheb=> explain analyze select type.ИД, ved.ЧЛВК_ИД from Н_ТИПЫ_ВЕДОМОСТЕЙ as type right join Н_ВЕДОМОСТИ as ved on type.ИД = ved.ТВ_ИД where type.ИД > 3
ved.ИД = 1457443 and ved.ИД < 1490007;
QUERY PLAN
-----
Nested Loop (cost=0.42..9.49 rows=1 width=8) (actual time=0.011..0.011 rows=0 loops=1)
  Join Filter: (type."ИД" = ved."ТВ_ИД")
  -> Seq Scan on "Н_ТИПЫ_ВЕДОМОСТЕЙ" type (cost=0.00..1.04 rows=1 width=4) (actual time=0.009..0.009 rows=0 loops=1)
    Filter: ("ИД" > 3)
    Rows Removed by Filter: 3
  -> Index Scan using "ВЕД_ПК" on "Н_ВЕДОМОСТИ" ved (cost=0.42..8.44 rows=1 width=8) (never executed)
    Index Cond: (("ИД" < 1490007) AND ("ИД" = 1457443))
Planning time: 0.214 ms
Execution time: 0.071 ms
(9 rows)
```



“Н_ТИПЫ_ВЕДОМОСТЕЙ” сканируется полностью (Seq Scan), однако в добавлении индекса нет смысла т. к. таблица содержит всего 3 записи и маловероятно, что когда-либо будет расширена.

“Н_ВЕДОМОСТИ” сканируется по первичному ключу - ИД (Index Scan), и по тому же ИД выполняется проверка (правда фактически выполнение туда не заходит из-за результатов проверки по Н_ТИПЫ_ВЕДОМОСТЕЙ). Таким образом, во введении дополнительных индексов нет необходимости

Во введении индексов нет необходимости

2. Сделать запрос для получения атрибутов из указанных таблиц, применив фильтры по указанным условиям:
Таблицы: Н_ЛЮДИ, Н_ОБУЧЕНИЯ, Н_УЧЕНИКИ.
Вывести атрибуты: Н_ЛЮДИ.ИД, Н_ОБУЧЕНИЯ.ЧЛВК_ИД, Н_УЧЕНИКИ.ИД.
Фильтры: (AND)
а) Н_ЛЮДИ.ИД > 163484.
б) Н_ОБУЧЕНИЯ.ЧЛВК_ИД < 105590.
Вид соединения: LEFT JOIN.

```

select people.ИД, studies.ЧЛВК_ИД, students.ИД from Н_ЛЮДИ as people
left join Н_ОБУЧЕНИЯ as studies on people.ИД = studies.ЧЛВК_ИД
left join Н_УЧЕНИКИ as students on students.ЧЛВК_ИД = people.ИД
where people.ИД > 163484 AND studies.ЧЛВК_ИД < 105590;
  
```

План выполнения запроса:

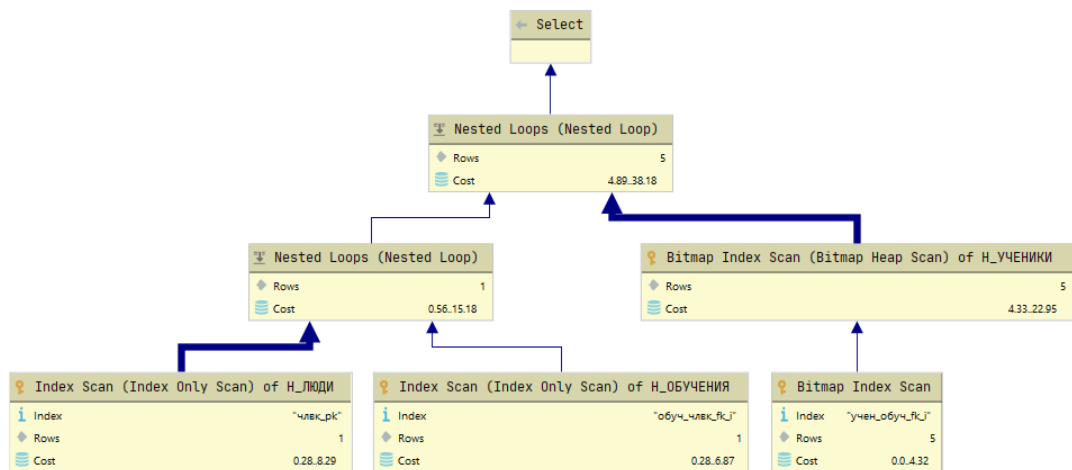
```

учеб=> explain analyse select people.ИД, studies.ЧЛВК_ИД, students.ИД from Н_ЛЮДИ as people left join Н_ОБУЧЕНИЯ as studies on people.ИД = stu
dies.ЧЛВК_ИД left join Н_УЧЕНИКИ as students on students.ЧЛВК_ИД = people.ИД where people.ИД > 163484 AND studies.ЧЛВК_ИД < 105590;
QUERY PLAN

Nested Loop Left Join (cost=4.89..38.19 rows=5 width=12) (actual time=0.007..0.007 rows=0 loops=1)
-> Nested Loop (cost=0.56..15.18 rows=1 width=8) (actual time=0.005..0.005 rows=0 loops=1)
  Join Filter: (people."ИД" = studies."ЧЛВК_ИД")
  -> Index Only Scan using "ЧЛВК_ИД" on "Н_ЛЮДИ" people (cost=0.28..8.29 rows=1 width=4) (actual time=0.005..0.005 rows=0 loops=1)
    Index Cond: ("ИД" > 163484)
    Heap Fetches: 0
  -> Index Only Scan using "ОБУЧ_ЧЛВК_ИД" on "Н_ОБУЧЕНИЯ" studies (cost=0.28..6.87 rows=1 width=4) (never executed)
    Index Cond: ("ЧЛВК_ИД" < 105590)
    Heap Fetches: 0
-> Bitmap Heap Scan on "Н_УЧЕНИКИ" students (cost=4.33..22.95 rows=5 width=8) (never executed)
  Recheck Cond: ("ЧЛВК_ИД" = people."ИД")
  -> Bitmap Index Scan on "УЧЕН_ОБУЧ_ИД" (cost=0.00..4.32 rows=5 width=0) (never executed)
    Index Cond: ("ЧЛВК_ИД" = people."ИД")

Planning time: 0.589 ms
Execution time: 0.091 ms
(15 rows)

учеб=>
  
```



Поиск по "Н_ОБУЧЕНИЯ" производится по столбцу ЧЛВК_ИД, так как из таблицы необходим только этот столбец и его же требуется вывести, Index Only Scan по нему – оптимальное решение

Поиск по "Н_ЛЮДИ" производится по первичному ключу, из таблицы нужен он же, фильтр по нему же – Index Only Scan снова оптимальное решение

Затем результат будет слит в один столбец с помощью JOIN (столбцы равны по странному условию варианта)

Затем производится Bitmap Index Scan по ЧЛВК_ИД в "Н_УЧЕНИКИ", планировщик снова прав.

Во введении индексов нет необходимости

Вывод

В ходе выполнения лабораторной работы я ознакомился с основами оптимизации запросов SQL, индексами в PostgreSQL и DataGrip.