Министерство науки и высшего образования Российской Федерации

Федеральное государственное автономное образовательное

учреждение высшего образования

"Национальный исследовательский университет ИТМО»

Факультет программной инженерии и компьютерных наук

09.03.04 «Программная инженерия»

Системное и прикладное программное обеспечение

**Отчёт по лабораторной работе №4.**

**По предмету "Базы данных»**

**Вариант: 657489314513**

Выполнил:

Шубин Егор Вячеславович

Группа Р3109

Проверил:

Воронина Дарья Сергеевна

Оглавление

[1 Задание 3](#_Toc198131422)

[2 Ход работы 4](#_Toc198131423)

[2.1 Запросы 4](#_Toc198131424)

[2.2 Добавление индексов 5](#_Toc198131425)

[2.3 Возможные планы выполнения запроса 1: 6](#_Toc198131426)

[2.4 Возможные планы выполнения запроса 2: 8](#_Toc198131427)

[2.5 Фактические планы запросов, полученные с помощью EXPLAIN ANALYSE 10](#_Toc198131428)

[Запрос 1 10](#_Toc198131429)

[Запрос2 10](#_Toc198131430)

[Вывод 11](#_Toc198131431)

# 1 Задание

Составить запросы на языке SQL (пункты 1-2).

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.  
Изменятся 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:  
   Таблицы: Н\_ЛЮДИ, Н\_СЕССИЯ.  
   Вывести атрибуты: Н\_ЛЮДИ.ИД, Н\_СЕССИЯ.ИД.  
   Фильтры (AND):  
   a) Н\_ЛЮДИ.ОТЧЕСТВО > Владимирович.  
   b) Н\_СЕССИЯ.ЧЛВК\_ИД = 151200.  
   c) Н\_СЕССИЯ.ЧЛВК\_ИД < 105948q.  
   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.ОТЧЕСТВО, Н\_ОБУЧЕНИЯ.ЧЛВК\_ИД, Н\_УЧЕНИКИ.ИД.  
   Фильтры: (AND)  
   a) Н\_ЛЮДИ.ОТЧЕСТВО = Владимирович.  
   b) Н\_ОБУЧЕНИЯ.ЧЛВК\_ИД < 163276.  
   Вид соединения: INNER JOIN.

# 2 Ход работы

## 2.1 Запросы

1 запрос:

SELECT  
 "Н\_ЛЮДИ"."ИД",  
 "Н\_СЕССИЯ"."ИД"  
FROM "Н\_ЛЮДИ"  
RIGHT JOIN "Н\_СЕССИЯ" ON "Н\_СЕССИЯ"."ЧЛВК\_ИД" = "Н\_ЛЮДИ"."ИД"  
WHERE "Н\_ЛЮДИ"."ОТЧЕСТВО" > 'Владимирович'  
 AND "Н\_СЕССИЯ"."ЧЛВК\_ИД" = 151200  
 AND "Н\_СЕССИЯ"."ЧЛВК\_ИД" < 105948;

2 запрос:  
SELECT "Н\_ЛЮДИ"."ОТЧЕСТВО", "Н\_ОБУЧЕНИЯ"."ЧЛВК\_ИД", "Н\_УЧЕНИКИ"."ИД"  
FROM "Н\_ЛЮДИ"  
INNER JOIN "Н\_ОБУЧЕНИЯ" on "Н\_ЛЮДИ"."ИД" = "Н\_ОБУЧЕНИЯ"."ЧЛВК\_ИД"  
INNER JOIN "Н\_УЧЕНИКИ" on "Н\_ЛЮДИ"."ИД" = "Н\_УЧЕНИКИ"."ЧЛВК\_ИД"  
WHERE "Н\_ЛЮДИ"."ОТЧЕСТВО" = 'Владимирович'  
 AND "Н\_ОБУЧЕНИЯ"."ЧЛВК\_ИД" < 163276;

2.2 Добавление индексов

1 запрос: Н\_ЛЮДИ.ИД и Н\_СЕССИЯ.ИД являются первичными ключами, значит в postgres им соответствуют B-tree индексы по умолчанию. Можно установить индекс на Н\_ЛЮДИ.ИД для ускорения соединения таблиц

2 запрос: Можно установить B-tree индексы на Н\_ЛЮДИ. ОТЧЕСТВО, Н\_ОБУЧЕНИЯ.ЧЛВК\_ИД и на Н\_УЧЕНИКИ.ИД для ускорения поиска подходящих строк

## 2.3 Возможные планы выполнения запроса 1:

Условия заведомо неверные, но при верности условий можно было составить следующие планы:

1 план:

• Полный скан таблицы Н\_ЛЮДИ, Н\_СЕССИЯ с фильтрацией по условию

• Соединение таблиц с помощью Hash Join

2 план:

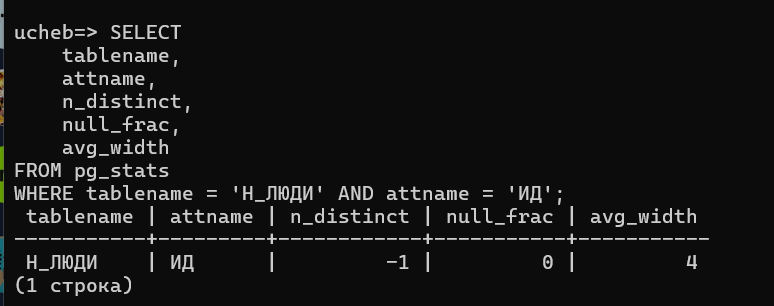
• Полный скан таблицы Н\_ЛЮДИ

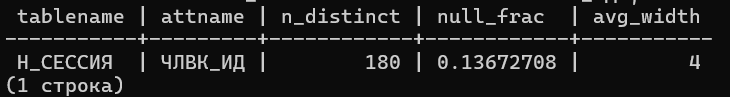
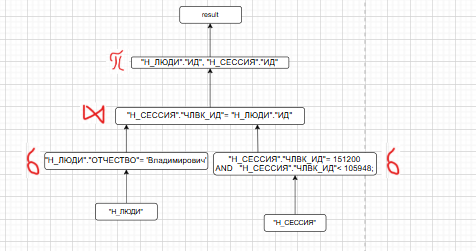
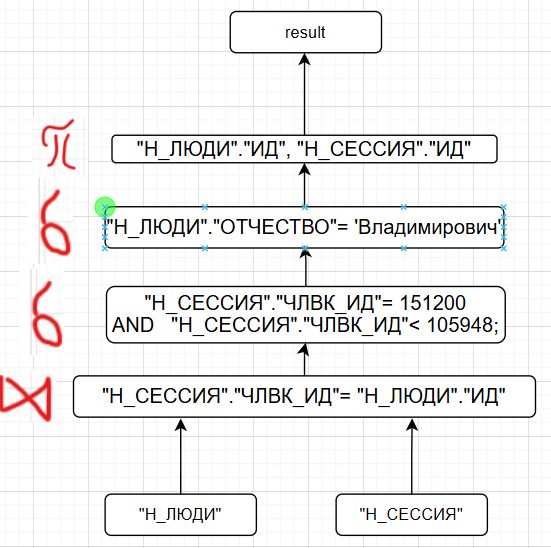
• Полный скан таблицы Н\_СЕССИЯ

• Соединение таблиц с использованием Nested Loop Join

• Фильтрация результата соединения по условия

План 1 оптимальный, так как фильтры к таблицам применяются до соединения таблиц, что уменьшает время выполнения запроса Если установить индекс на Н\_ЛЮДИ.ИД, т.к. значения уникальны то таблицы могут соединиться быстрее, создание индекса для Н\_СЕССИЯ.ЧЛВК\_ИД не нужно, т.к. всего 180 уникальных значений из 3752.



## 2.4 Возможные планы выполнения запроса 2:

1 план:

• Полный скан таблицы Н\_ЛЮДИ с фильтрацией по условию ОТЧЕСТВО = Владимирович

• Скан и фильтрация таблицы Н\_ОБУЧЕНИЯ по условию ) ЧЛВК\_ИД < 163276.

• Соединение отфильтрованных таблиц Н\_ЛЮДИ и Н\_ОБУЧЕНИЯ

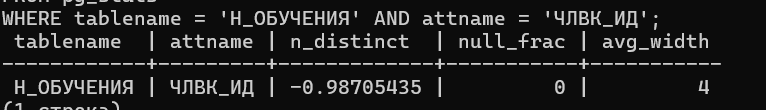
• Полный скан таблицы Н\_УЧЕНИКИ

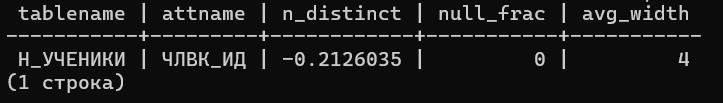
• Присоединение таблицы Н\_УЧЕНИКИ

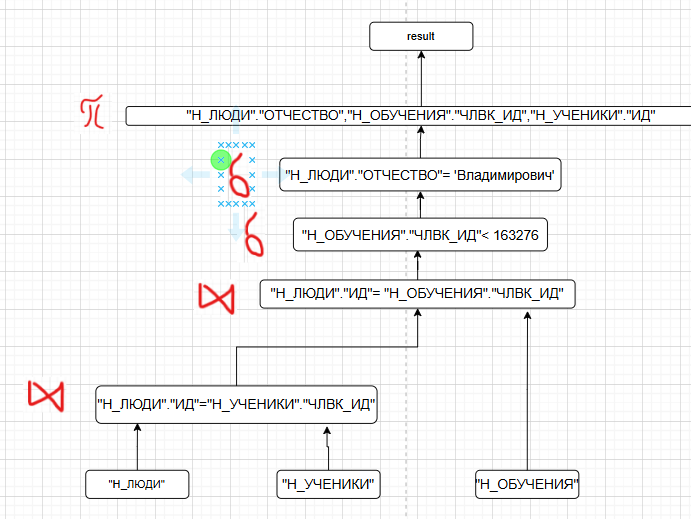
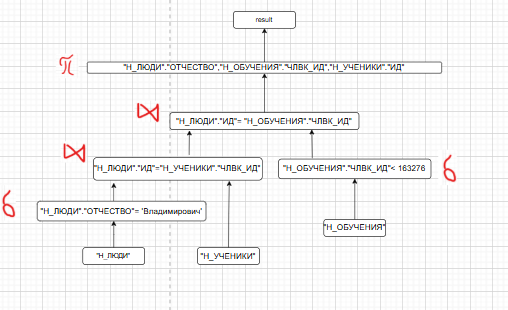
2 план:

* Сканирование и соединение таблиц с помощью hash join
* Фильтрация таблицы Н\_ОБУЧЕНИЯ по условию ) ЧЛВК\_ИД < 163276.
* Фильтрация таблицы по условию ОТЧЕСТВО = Владимирович

Для Н\_ОБУЧЕНИЯ.ЧЛВК\_ИД создаем индекс, в связи с высокой уникальностью. Всего – 5021 значений, уникальных 4956

Для Н\_УЧЕНИКИ\_.ЧЛВК\_ИД индекс нет смысла создавать в связи с низкой уникальностью значений Всего – 23311 значений, уникальных 4956.



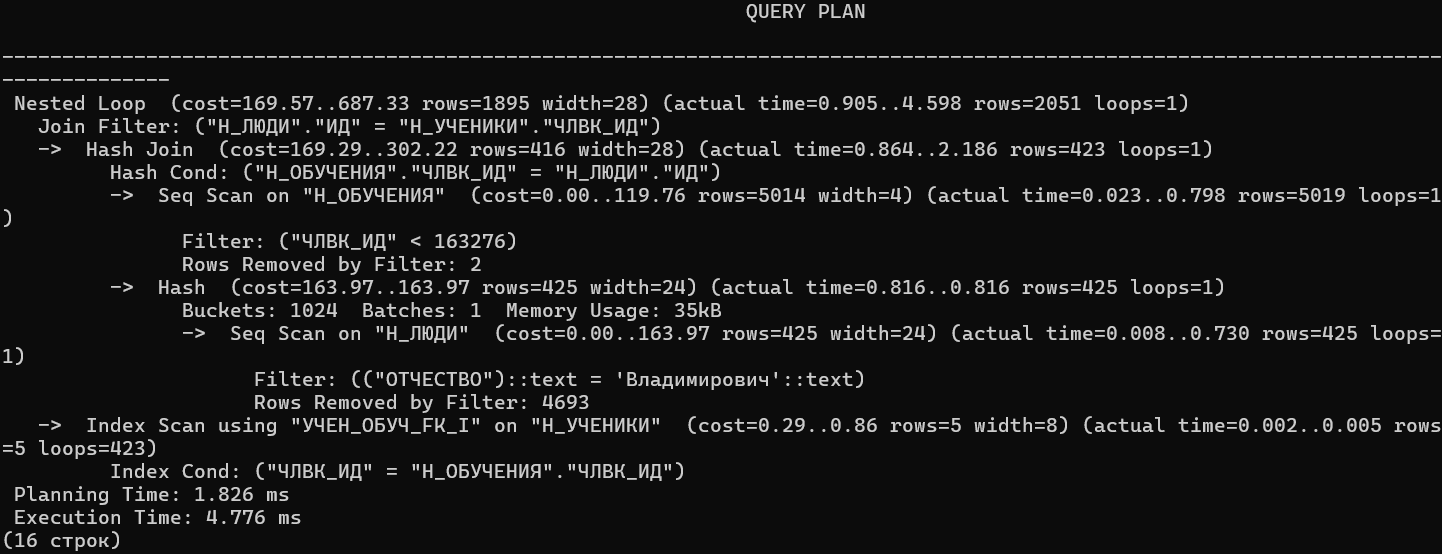


## 2.5 Фактические планы запросов, полученные с помощью EXPLAIN ANALYSE

### Запрос 1

Таблицы соединяются с помощью nested loop, в индексе, созданного после сканирования Н\_ЛЮДИ.ИД фактически возвращено 0 строк, хотя ожидалось 1 строка, т.к. данная строка удаляется при выполнении фильтра ОТЧЕСТВО > 'Владимирович'. Затем выполняется сканирование индекса Н\_СЕССИЯ.ЧЛВК\_ИД и фильтруются по двум противоречащим условиям.

### Запрос2

Сканируются таблицы Н\_ЛЮДИ И Н\_ОБУЧЕНИЯ и фильтруются по условиям "Н\_ЛЮДИ"."ОТЧЕСТВО" = 'Владимирович' и "Н\_ОБУЧЕНИЯ"."ЧЛВК\_ИД" < 163276;. Затем соединяются с помощью hash\_join. После этого с помощью nested loop присоединяется индекс Н\_УЧЕНИКИ.ЧЛВК\_ИД.

# Вывод

Во время выполнения данной лабораторной работы я научился работать с индексами и с pg\_stats.