

МИНИСТЕРСТВО НАУКИ И ОБРАЗОВАНИЯ РФ

Федеральное государственное автономное
образовательное учреждение высшего образования
«Национальный исследовательский университет
ИТМО»

ФАКУЛЬТЕТ ПРОГРАММНОЙ ИНЖЕНЕРИИ И КОМПЬЮТЕРНОЙ ТЕХНИКИ

ЛАБОРАТОРНАЯ РАБОТА №4

по дисциплине
БАЗЫ ДАННЫХ

Вариант - 25254

Выполнила:
Студент группы Р3132
Чмурова Мария
Владиславовна
Проверила:
Харитоновна
Анастасия Евгеньевна

г. Санкт-Петербург
2023 год

Задание

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:

Таблицы: Н_ЛЮДИ, Н_ВЕДОМОСТИ.

Вывести атрибуты: Н_ЛЮДИ.ОТЧЕСТВО, Н_ВЕДОМОСТИ.ИД.

Фильтры (AND):

- а) Н_ЛЮДИ.ИМЯ > Роман.
- б) Н_ВЕДОМОСТИ.ДАТА < 2010-06-18.

Вид соединения: INNER 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.НЗК,
Н_УЧЕНИКИ.НАЧАЛО.

Фильтры: (AND)

а) Н_ЛЮДИ.ОТЧЕСТВО > Владимирович.

б) Н_ОБУЧЕНИЯ.ЧЛВК_ИД < 113409.

с) Н_УЧЕНИКИ.НАЧАЛО < 2008-09-01.

Вид соединения: LEFT JOIN.

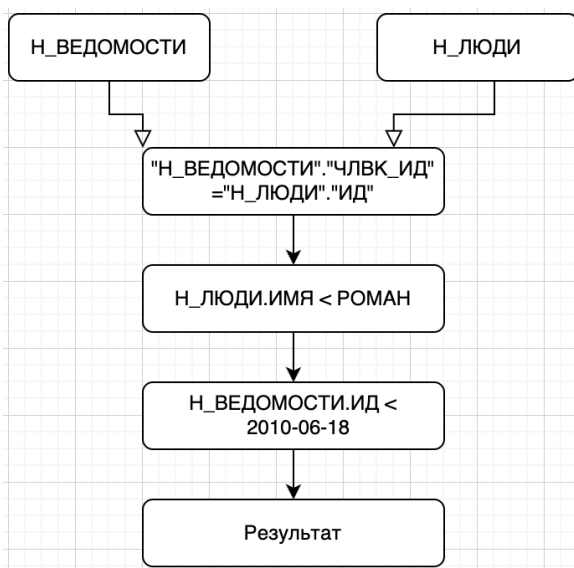
Ход работы

1. Запрос:

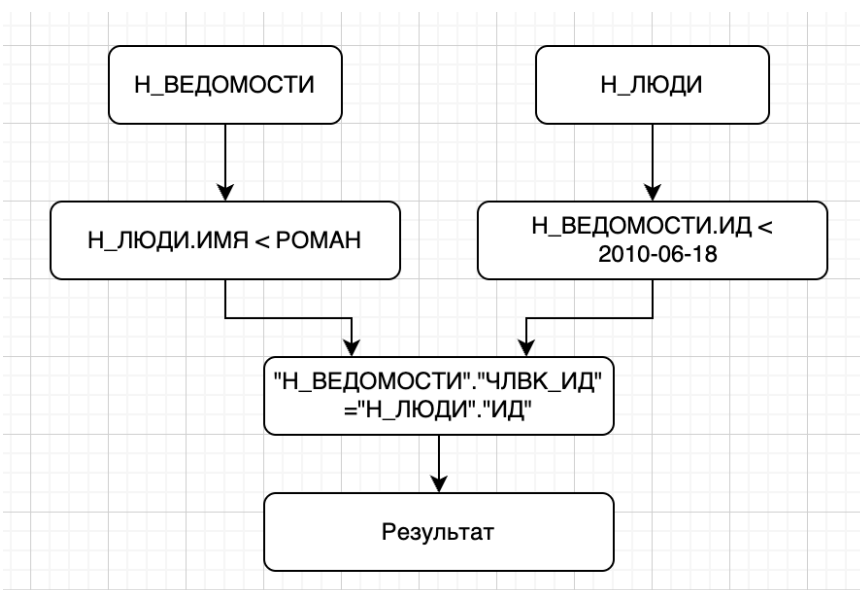
```
SELECT "Н_ЛЮДИ"."ОТЧЕСТВО", "Н_ВЕДОМОСТИ"."ИД"  
from "Н_ЛЮДИ"  
INNER JOIN "Н_ВЕДОМОСТИ" on "Н_ЛЮДИ"."ИД" = "Н_ВЕДОМОСТИ"."ЧЛВК_ИД"  
WHERE "Н_ЛЮДИ"."ИМЯ" > 'Роман' AND "Н_ВЕДОМОСТИ"."ДАТА" < '2010-06-  
18'::date;
```

Возможные планы выполнения запроса (без индексации)

1 План:



2 План:



На первом плане сначала происходит соединение таблиц, а затем выборка нужных данных. На втором плане сначала находятся нужные нам данным, а потом соединяются таблицы, это делает возможный план более быстрым и оптимальным.

Возможные индексы для оптимизации запроса:

```
CREATE INDEX person_name_index ON "Н_ЛЮДИ" USING btree ("ИМЯ");  
CREATE INDEX ved_date_index ON "Н_ВЕДОМОСТИ" USING btree ("ДАТА");
```

Для таблицы "Н_ЛЮДИ" создается индекс на поле "ИМЯ" с использованием В-tree, так как в запросе применяется операция сравнения по имени.

Поскольку в запросе применяется операция сравнения по полю "ДАТА", создание индекса на этом поле может улучшить производительность запроса. В-tree индексы обеспечивают эффективный поиск и сортировку данных для операций сравнения.

Результат EXPLAIN ANALYZE:

```
Hash Join (cost=173.36..7535.75 rows=28814 width=24) (actual time=2.603..67.593  
rows=28463 loops=1)  
" Hash Cond: ("Н_ВЕДОМОСТИ"."ЧЛВК_ИД" = "Н_ЛЮДИ"."ИД")"  
" -> Seq Scan on "Н_ВЕДОМОСТИ" (cost=0.00..6846.50 rows=196366 width=8) (actual  
time=0.015..41.954 rows=196568 loops=1)"  
" Filter: ("ДАТА" < '2010-06-18'::date)"  
Rows Removed by Filter: 25872  
-> Hash (cost=163.97..163.97 rows=751 width=24) (actual time=2.579..2.580 rows=753  
loops=1)  
Buckets: 1024 Batches: 1 Memory Usage: 52kB  
" -> Seq Scan on "Н_ЛЮДИ" (cost=0.00..163.97 rows=751 width=24) (actual  
time=0.007..2.435 rows=753 loops=1)"  
" Filter: (("ИМЯ")::text > 'Поман'::text)"  
Rows Removed by Filter: 4365  
Planning Time: 0.415 ms  
Execution Time: 68.972 ms
```

2. Запрос:

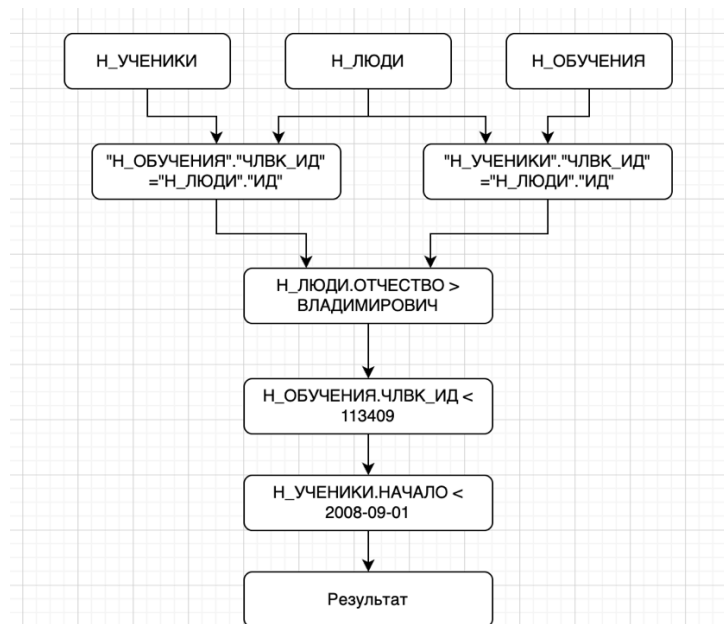
```

SELECT "Н_ЛЮДИ"."ОТЧЕСТВО", "Н_ОБУЧЕНИЯ"."НЗК", "Н_УЧЕНИКИ"."НАЧАЛО"
FROM "Н_ЛЮДИ"
LEFT JOIN "Н_ОБУЧЕНИЯ" ON "Н_ОБУЧЕНИЯ"."ЧЛВК_ИД" = "Н_ЛЮДИ"."ИД"
LEFT JOIN "Н_УЧЕНИКИ" ON "Н_УЧЕНИКИ"."ЧЛВК_ИД" = "Н_ЛЮДИ"."ИД"
WHERE ("Н_ЛЮДИ"."ОТЧЕСТВО" > 'Владимирович')
      AND ("Н_ОБУЧЕНИЯ"."ЧЛВК_ИД" < '113409')
      AND ("Н_УЧЕНИКИ"."НАЧАЛО" < '2008-09-01'::date);

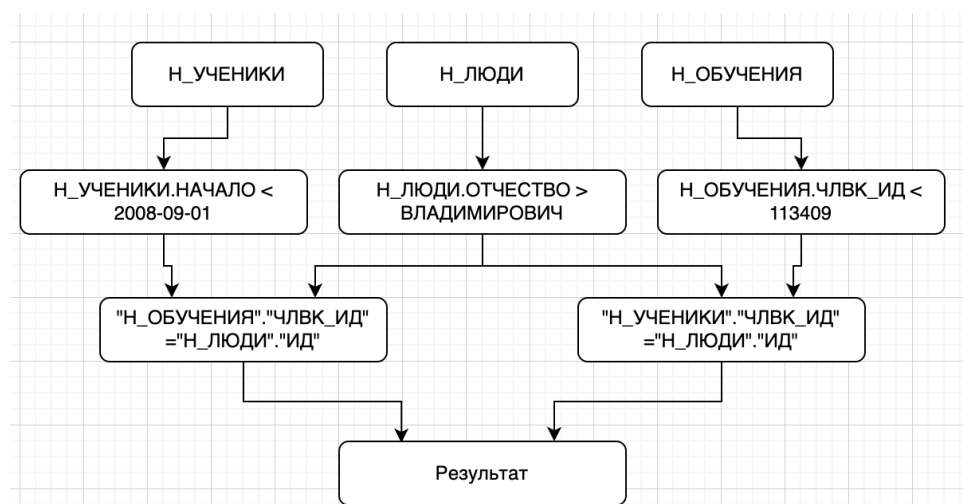
```

Возможные планы выполнения запроса (без индексации):

1 План:



2 План:



Аналогично плану из запроса 1, во втором плане находятся нужные данные, а потом соединяются таблицы, это позволяет плану быть более быстрым и оптимальным.

Возможные индексы для оптимизации запроса:

```
CREATE INDEX person_third_name_index ON "Н_ЛЮДИ" ("ОТЧЕСТВО");  
CREATE INDEX person_id_index ON "Н_ОБУЧЕНИЯ" ("ЧЛВК_ИД");  
CREATE INDEX student_person_id_index ON "Н_УЧЕНИКИ" ("ЧЛВК_ИД");  
CREATE INDEX student_start_date_index ON "Н_УЧЕНИКИ" ("НАЧАЛО");
```

Выбран тип индекса B-tree по умолчанию, поскольку он хорошо подходит для поиска диапазона значений. Он обеспечивает эффективный поиск и сортировку данных для операций сравнения.

Результат EXPLAIN ANALYZE:

```
Nested Loop  (cost=79.32..536.19 rows=702 width=34) (actual  
time=0.291..5.121 rows=766 loops=1)  
"  Join Filter: ("Н_ОБУЧЕНИЯ"."ЧЛВК_ИД" =  
"Н_УЧЕНИКИ"."ЧЛВК_ИД") "  
-> Hash Join  (cost=79.04..257.83 rows=209 width=34) (actual  
time=0.253..3.711 rows=198 loops=1)  
"      Hash Cond: ("Н_ЛЮДИ"."ИД" = "Н_ОБУЧЕНИЯ"."ЧЛВК_ИД") "  
"      -> Seq Scan on "Н_ЛЮДИ"  (cost=0.00..163.97 rows=2545  
width=24) (actual time=0.011..3.166 rows=2546 loops=1) "  
"          Filter: (("ОТЧЕСТВО")::text > 'Владимирович'::text) "  
          Rows Removed by Filter: 2572  
      -> Hash  (cost=73.79..73.79 rows=420 width=10) (actual  
time=0.222..0.222 rows=412 loops=1)  
          Buckets: 1024  Batches: 1  Memory Usage: 25kB  
"      -> Bitmap Heap Scan on "Н_ОБУЧЕНИЯ"  
(cost=11.54..73.79 rows=420 width=10) (actual time=0.058..0.141  
rows=412 loops=1) "  
"          Recheck Cond: ("ЧЛВК_ИД" < 113409) "  
          Heap Blocks: exact=10  
"      -> Bitmap Index Scan on "ОБУЧ_ЧЛВК_ФК_I"  
(cost=0.00..11.43 rows=420 width=0) (actual time=0.047..0.047 rows=412  
loops=1) "  
"          Index Cond: ("ЧЛВК_ИД" < 113409) "  
"  -> Index Scan using "УЧЕН_ОБУЧ_ФК_I" on "Н_УЧЕНИКИ"  
(cost=0.29..1.29 rows=3 width=12) (actual time=0.003..0.006 rows=4  
loops=198) "
```

```
"      Index Cond: ("ЧЛВК_ИД" = "Н_ЛЮДИ"."ИД") "  
"      Filter: ("НАЧАЛО" < '2008-09-01'::date) "  
Planning Time: 1.565 ms  
Execution Time: 5.245 ms
```


Вывод

В ходе данной лабораторной работы я познакомилась с индексами, узнала для чего они используются, а также научилась создавать возможные планы выполнения запросов