

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«Санкт-Петербургский национальный исследовательский университет
информационных технологий, механики и оптики»

ФАКУЛЬТЕТ ПРОГРАММНОЙ ИНЖЕНЕРИИ И
КОМПЬЮТЕРНОЙ ТЕХНИКИ

ЛАБОРАТОРНАЯ РАБОТА № 4

по дисциплине
«Базы Данных»

Вариант № 444

Выполнил:

Студент группы Р3125

Шадрухин Александр Сергеевич

Преподаватель:

Егошин Алексей Васильевич

г. Санкт-Петербург, 2023

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_СЕССИЯ.
Вывести атрибуты: Н_ЛЮДИ.ОТЧЕСТВО, Н_СЕССИЯ.ЧЛВК_ИД.
Фильтры (AND):
 - а) Н_ЛЮДИ.ИМЯ < Александр.
 - б) Н_СЕССИЯ.ЧЛВК_ИД = 106059.Вид соединения: LEFT JOIN.
2. Сделать запрос для получения атрибутов из указанных таблиц, применив фильтры по указанным условиям:
Таблицы: Н_ЛЮДИ, Н_ВЕДОМОСТИ, Н_СЕССИЯ.
Вывести атрибуты: Н_ЛЮДИ.ИД, Н_ВЕДОМОСТИ.ДАТА, Н_СЕССИЯ.ДАТА.
Фильтры (AND):
 - а) Н_ЛЮДИ.ИМЯ > Роман.
 - б) Н_ВЕДОМОСТИ.ЧЛВК_ИД = 153285.
 - с) Н_СЕССИЯ.ЧЛВК_ИД = 105948.Вид соединения: INNER JOIN.

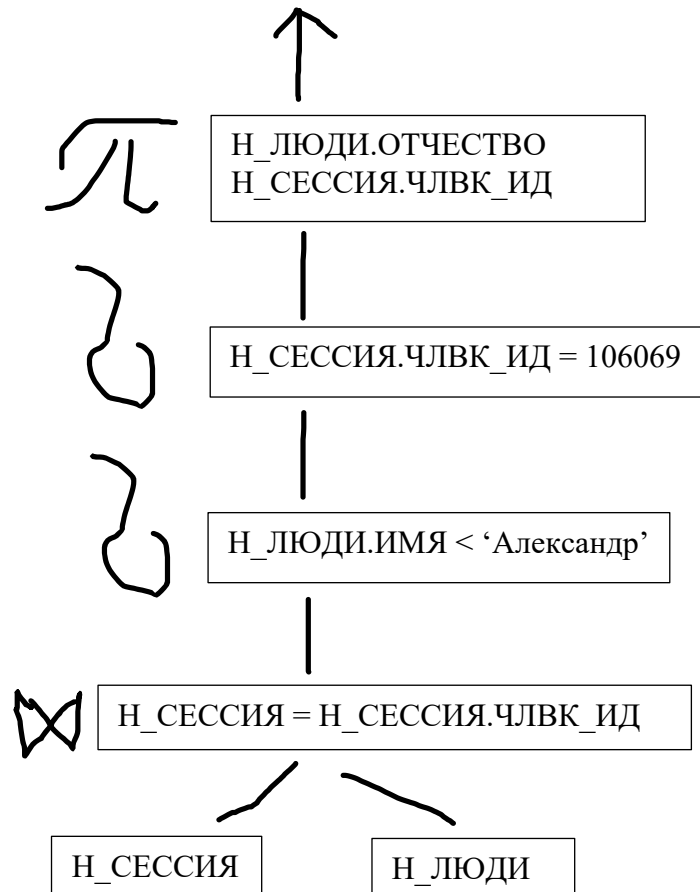
Запрос 1

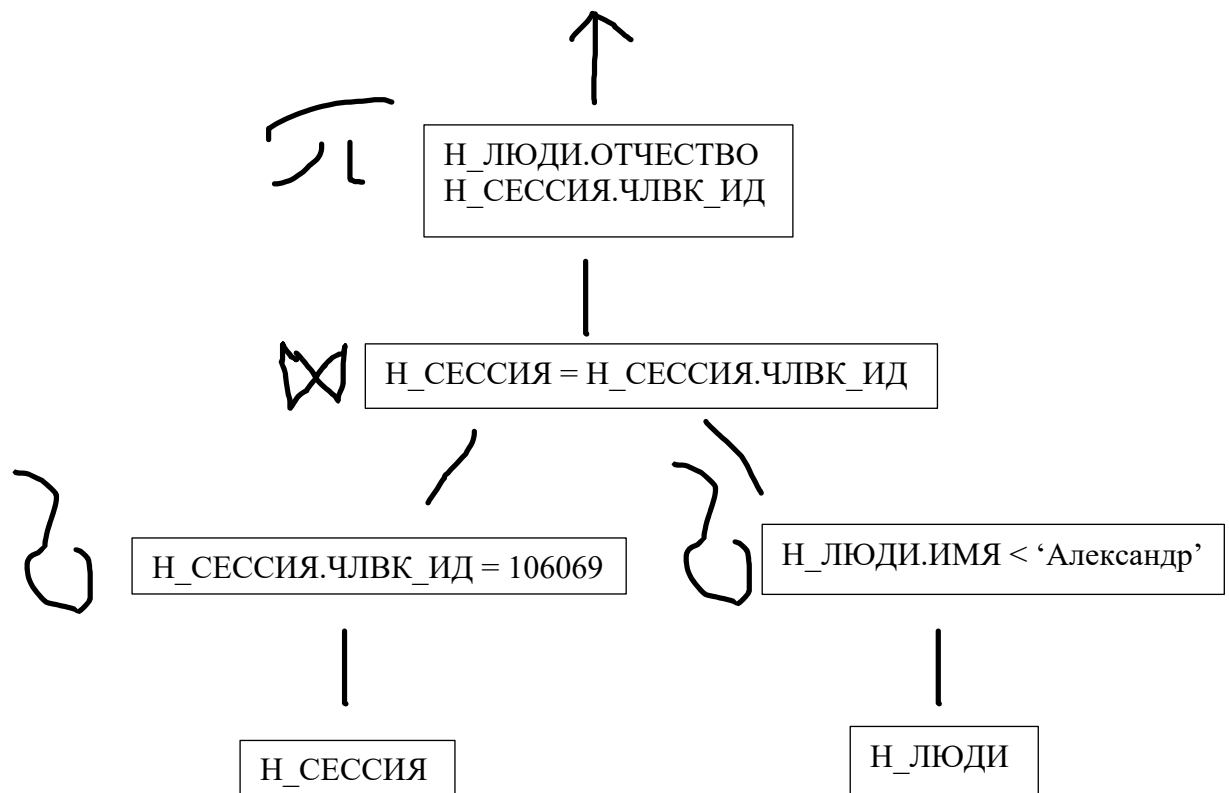
```
SELECT Н_ЛЮДИ.ОТЧЕСТВО, Н_СЕССИЯ.ЧЛВК_ИД
FROM Н_ЛЮДИ
LEFT JOIN Н_СЕССИЯ ON Н_ЛЮДИ.ИД = Н_СЕССИЯ.ЧЛВК_ИД
WHERE Н_ЛЮДИ.ИМЯ < 'Александр' AND Н_СЕССИЯ.ЧЛВК_ИД = 106059;
```

QUERY PLAN

```
Nested Loop (cost=0.56..12.83 rows=9 width=24) (actual time=0.039..0.040 rows=0 loops=1)
-> Index Scan using "ЧЛВК_PK" on "Н_ЛЮДИ" (cost=0.28..8.36 rows=1 width=21) (actual time=0.038..0.038 rows=0 loops=1)
    Index Cond: ("ИД" = 106059)
    Filter: ((("ИМЯ")::text < 'Александр')::text)
    Rows Removed by Filter: 1
-> Index Only Scan using "SYS_COPYING_IDX" on "Н_СЕССИЯ" (cost=0.28..4.44 rows=9 width=4) (never executed)
    Index Cond: ("ЧЛВК_ИД" = 106059)
    Heap Fetches: 0
Planning Time: 0.871 ms
Execution Time: 0.095 ms
(10 строк)
```

С точки зрения индексирования данных в базах данных, для индексации поля ЧЛВК_ИД, где проверка происходит на равенство, наиболее подходящим индексом будет hash-индекс. Hash-индекс использует хэш-функцию для быстрого поиска данных в индексе. Он позволяет быстро определить, есть ли в индексе запись с определенным ключом. Для индексации поля Н_ЛЮДИ.ИМЯ, где сравнение происходит по знаку "<", наиболее подходящим индексом будет B-tree. B-tree индекс используется для быстрого поиска данных в индексе, когда условие поиска базируется на сравнении данных. B-tree индекс упорядочивает данные в древовидной структуре, что облегчает поиск данных и определение их порядка.





2 план

План 2 будет самым эффективным, потому что сначала мы ограничиваем выборку, затем используем join. В `Н_СЕССИЯ.ЧЛВК_ИД = 106069` `ЧЛВК_ИД` – первичный ключ. Поэтому записей будет либо 1, либо 0 и запрос выполнится быстро.

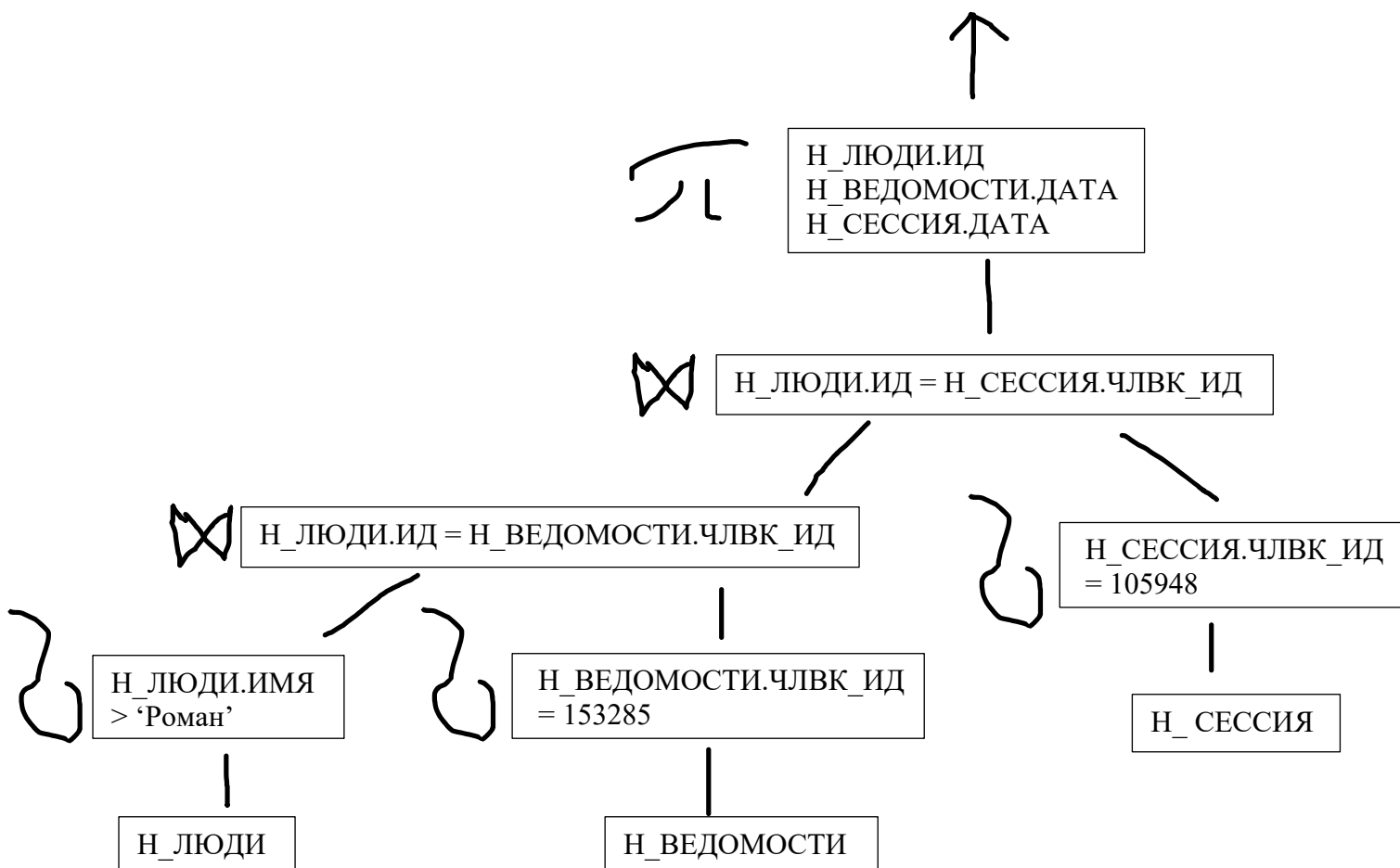
Запрос 2

```
SELECT Н_ЛЮДИ.ИД, Н_ВЕДОМОСТИ.ДАТА, Н_СЕССИЯ.ДАТА
FROM Н_ЛЮДИ
  INNER JOIN Н_СЕССИЯ ON Н_ЛЮДИ.ИД = Н_СЕССИЯ.ЧЛВК_ИД
  INNER JOIN Н_ВЕДОМОСТИ ON Н_ЛЮДИ.ИД = Н_ВЕДОМОСТИ.ЧЛВК_ИД
WHERE Н_ЛЮДИ.ИМЯ > 'Роман' AND Н_ВЕДОМОСТИ.ЧЛВК_ИД = 153285 AND
Н_СЕССИЯ.ЧЛВК_ИД = 105948;
```

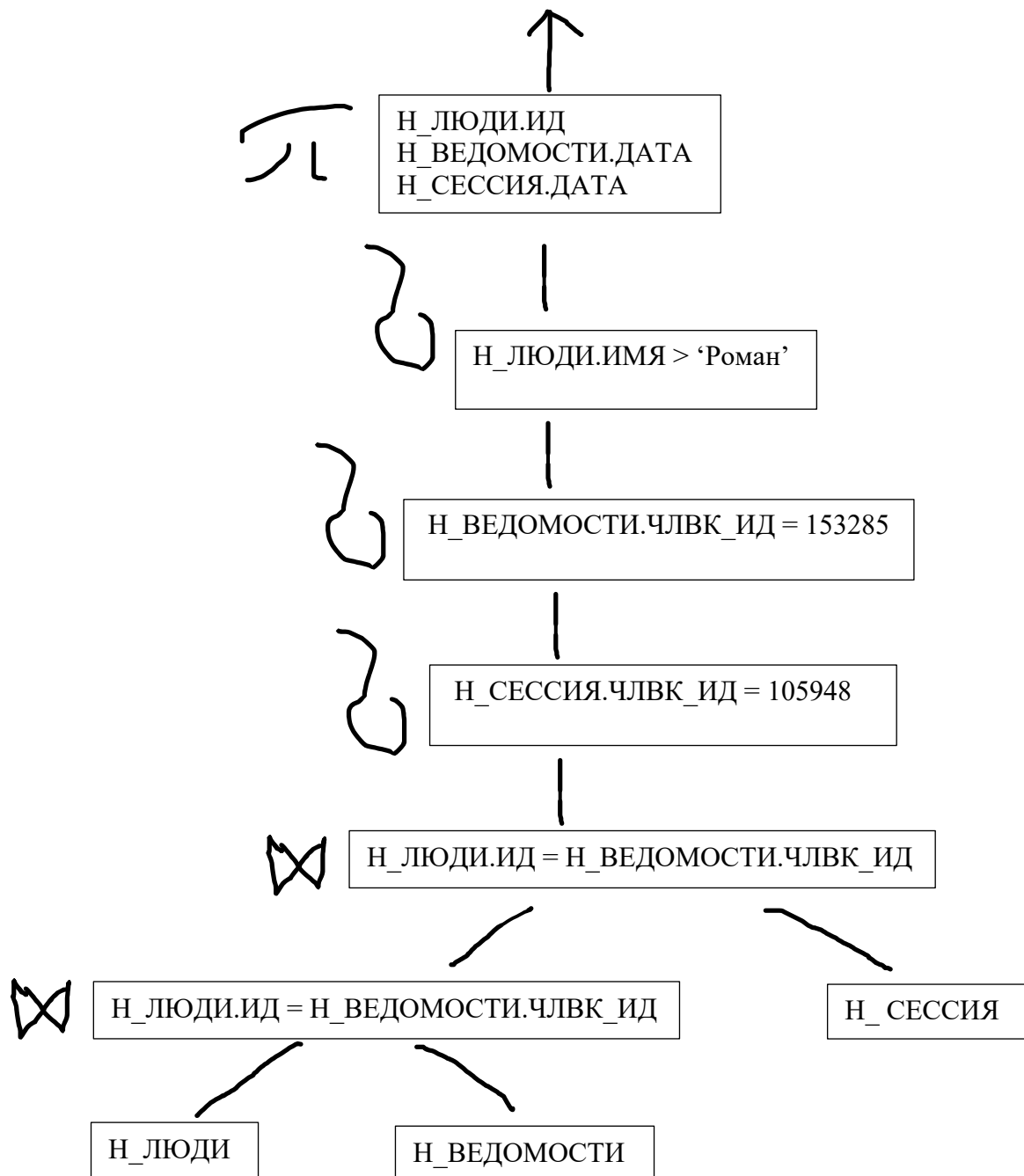
QUERY PLAN

Result (cost=0.00..0.80 rows=0 width=20) (actual time=0.001..0.001 rows=0 loops=1)
One-Time Filter: false
Planning Time: 1.322 ms
Execution Time: 0.018 ms
(4 строки)

Для оптимизации данного запроса можно добавить индекс B-tree на атрибуты Н_ЛЮДИ.ФАМИЛИЯ и Н_ЛЮДИ.ИМЯ, так как производится операция сравнения с этими полями. Также можно добавить индекс Hash на Н_ВЕДОМОСТИ.ИД, так как производится сравнение по знаку “=”. Однако стоит отметить, что добавление индексов нужно оценивать не только по одному запросу, но и в контексте всей системы, чтобы знать, какие индексы в долгосрочной перспективе будут полезны.



1 план



2 план

План 1 будет выгоднее, потому что сначала я выбираю нужные значения из таблицы, а потом использую join. Это уменьшает количество следующих записей в несколько раз.

Вывод

В процессе выполнения лабораторной работы я усвоил методы оптимизации выполнения запросов, в том числе через добавление индексов к атрибутам таблиц. Также научился определять условия, при которых необходимо добавлять атрибуты, и использовать команду EXPLAIN ANALYZE для получения информации о выполнении запроса.