

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

Факультет программной инженерии и компьютерной техники

Лабораторная работа №4

По базам данных

Вариант №696

Выполнил:

Ступин Тимур Русланович

Группа № Р3108

Преподаватель:

Афанасьев Дмитрий Борисович

Санкт-Петербург 2024

Содержание

Текст задания	3
Уменьшение времени выполнения 1 запроса	4
Уменьшение времени выполнения 2 запроса	8
Вывод.....	12

Текст задания

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:
Н_ТИПЫ_ВЕДОМОСТЕЙ, Н_ВЕДОМОСТИ.
Вывести атрибуты: Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ,
Н_ВЕДОМОСТИ.ЧЛВК_ИД.
Фильтры (AND):
 - a) Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ = Ведомость.
 - b) Н_ВЕДОМОСТИ.ИД = 39921.
 - c) Н_ВЕДОМОСТИ.ИД < 1250981.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.ИД, Н_ОБУЧЕНИЯ.НЗК,
Н_УЧЕНИКИ.ГРУППА.
Фильтры: (AND)
 - a) Н_ЛЮДИ.ИД > 152862.
 - b) Н_ОБУЧЕНИЯ.ЧЛВК_ИД = 112514.Вид соединения: INNER JOIN.

Уменьшение времени выполнения 1 запроса

```
-- Сделать запрос для получения атрибутов из указанных таблиц,  
применив фильтры по указанным условиям:  
-- Н_ТИПЫ_ВЕДОМОСТЕЙ, Н_ВЕДОМОСТИ.  
-- Вывести атрибуты: Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ,  
Н_ВЕДОМОСТИ.ЧЛВК_ИД.  
-- Фильтры (AND):  
-- а) Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ = Ведомость.  
-- б) Н_ВЕДОМОСТИ.ИД = 39921.  
-- в) Н_ВЕДОМОСТИ.ИД < 1250981.  
-- Вид соединения: INNER JOIN.  
  
SELECT Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ, Н_ВЕДОМОСТИ.ЧЛВК_ИД  
FROM Н_ТИПЫ_ВЕДОМОСТЕЙ  
INNER JOIN Н_ВЕДОМОСТИ ON Н_ВЕДОМОСТИ.ТВ_ИД = Н_ТИПЫ_ВЕДОМОСТЕЙ.ИД  
WHERE Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ = 'Ведомость' AND  
Н_ВЕДОМОСТИ.ИД = 39921 AND  
Н_ВЕДОМОСТИ.ИД < 1250981;
```

Индексы, добавление которых уменьшит время выполнения запроса:

1. В таблице Н_ТИПЫ_ВЕДОМОСТЕЙ

- Индекс на атрибуте НАИМЕНОВАНИЕ (B-tree). Это ускорит выполнение фильтрации строк по условию “Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ = Ведомость”

```
CREATE INDEX Н_ТИПЫ_ВЕДОМОСТЕЙ_НАИМЕНОВАНИЕ ON Н_ТИПЫ_ВЕДОМОСТЕЙ  
USING btree(НАИМЕНОВАНИЕ);
```

2. В таблице Н_ВЕДОМОСТИ

- Индекс на атрибуте ИД (B-tree). Это ускорит выполнение фильтрации строк по условиям “Н_ВЕДОМОСТИ.ИД = 39921” и “Н_ВЕДОМОСТИ.ИД < 1250981”

```
CREATE INDEX Н_ВЕДОМОСТИ_ИД ON Н_ВЕДОМОСТИ USING btree(ИД);
```

- Индекс на атрибуте ТВ_ИД (B-tree). Это ускорит выполнение соединения таблиц по атрибуту ТВ_ИД

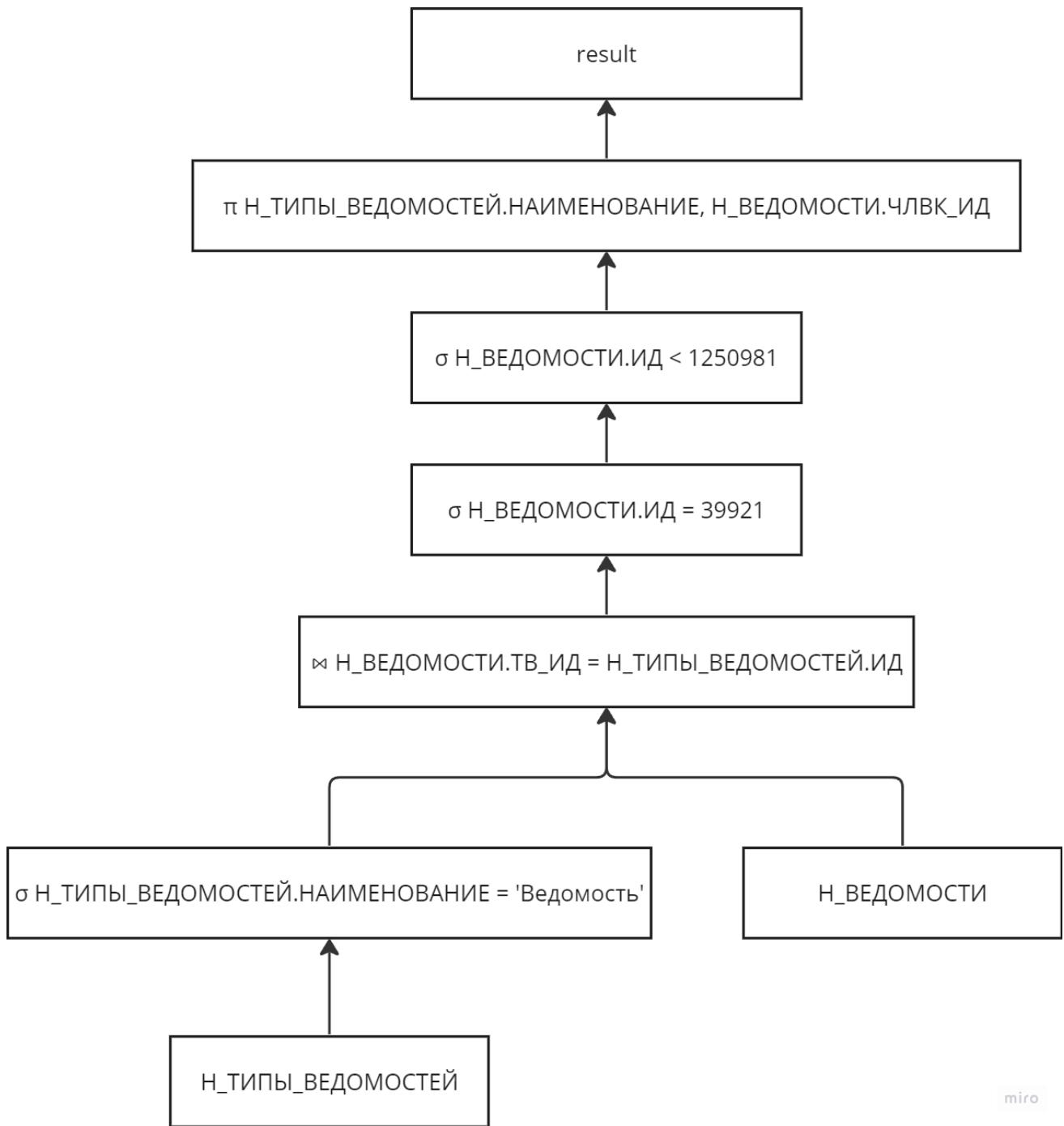
```
CREATE INDEX Н_ВЕДОМОСТИ_ТВ_ИД ON Н_ВЕДОМОСТИ USING btree(ТВ_ИД);
```

Добавление индексов на указанные столбцы позволит ускорить поиск нужных записей по фильтрам в запросе, а также выполнение соединений.

Возможные планы выполнения запроса без индексов:

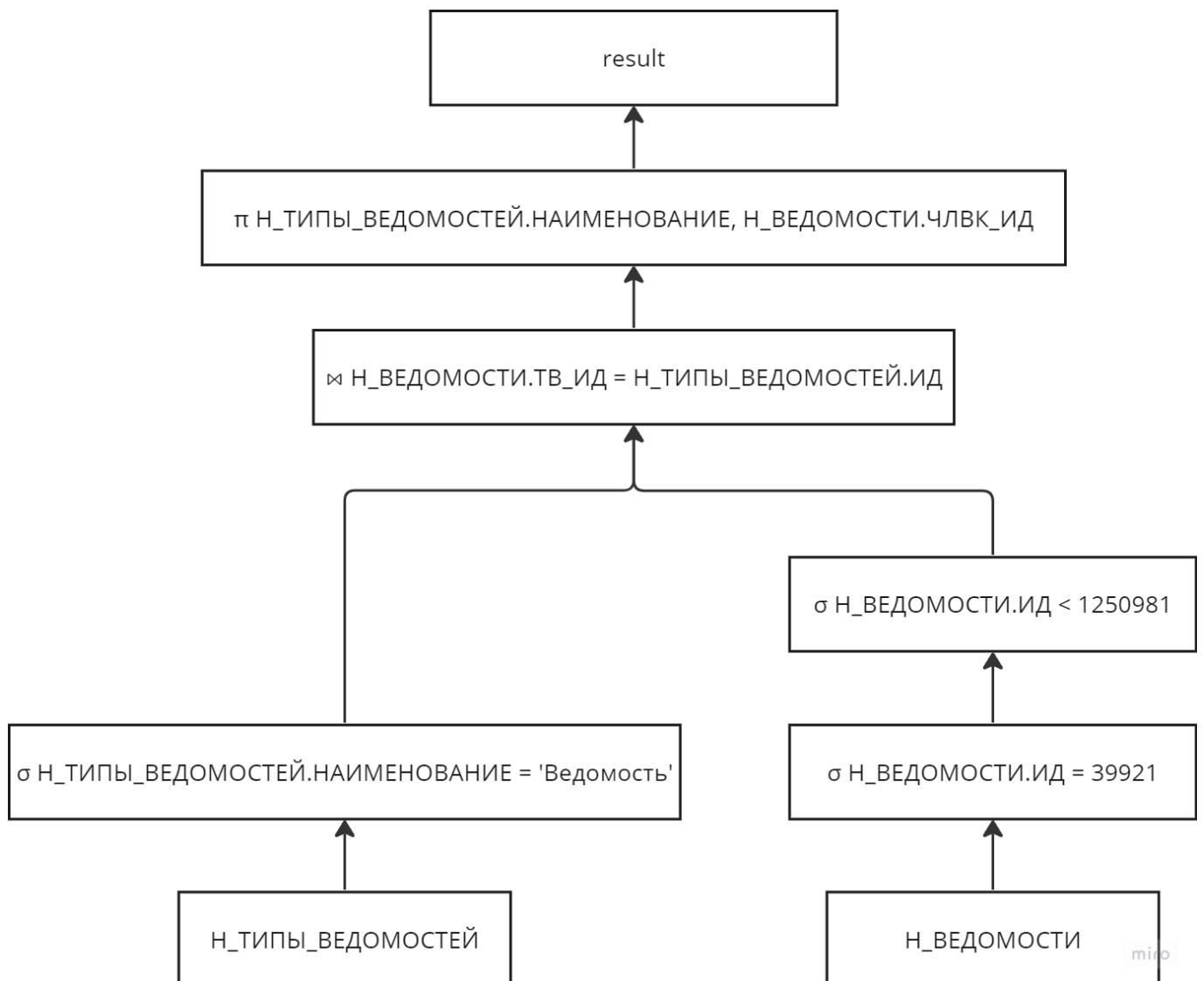
План 1:

- Полный скан таблицы Н_ТИПЫ_ВЕДОМОСТЕЙ с применением фильтра по условию “Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ = Ведомость”
- Полный скан таблицы Н_ВЕДОМОСТИ
- Соединение таблиц с использованием Nested Loop Join по атрибуту ТВ_ИД
- Фильтрация результата соединения по условию “Н_ВЕДОМОСТИ.ИД = 39921”
- Фильтрация результата соединения по условию “Н_ВЕДОМОСТИ.ИД < 1250981”



План 2:

- Полный скан таблицы Н_ТИПЫ_ВЕДОМОСТЕЙ с применением фильтра по условию “Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ = Ведомость”
- Полный скан таблицы Н_ВЕДОМОСТИ с применением фильтра по условию “Н_ВЕДОМОСТИ.ИД = 39921”
- Фильтрация результата сканирования таблицы Н_ВЕДОМОСТИ по условию “Н_ВЕДОМОСТИ.ИД < 1250981”
- Соединение таблиц с использованием Nested Loop Join по атрибуту ТВ_ИД



Оптимальный план:

- План 2, так как фильтрация данных выполняется до соединения таблиц, что уменьшает количество строк для обработки

При добавлении индексов планы выполнения запросов изменятся:

- Вместо полного скана таблиц будет использован индексный скан
- Nested Loop Join станет быстрее благодаря индексу на атрибуте ТВ_ИД

QUERY PLAN

```
Nested Loop (cost=0.42..9.49 rows=1 width=30) (actual time=0.061..0.063 rows=1 loops=1)
  Join Filter: ("Н_ТИПЫ_ВЕДОМОСТЕЙ"."ИД" = "Н_ВЕДОМОСТИ"."ТВ_ИД")
    -> Seq Scan on "Н_ТИПЫ_ВЕДОМОСТЕЙ" (cost=0.00..1.04 rows=1 width=30) (actual
        time=0.010..0.011 rows=1 loops=1)
        Filter: (("НАИМЕНОВАНИЕ")::text = 'Ведомость'::text)
        Rows Removed by Filter: 2
    -> Index Scan using "ВЕД_PK" on "Н_ВЕДОМОСТИ" (cost=0.42..8.44 rows=1 width=8)
        (actual time=0.046..0.046 rows=1 loops=1)
        Index Cond: (("ИД" < 1250981) AND ("ИД" = 39921))
Planning Time: 1.063 ms
Execution Time: 0.121 ms
```

Уменьшение времени выполнения 2 запроса

```
-- 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.НЗК, Н_УЧЕНИКИ.ГРУППА.  
-- Фильтры: (AND)  
-- а) Н_ЛЮДИ.ИД > 152862.  
-- б) Н_ОБУЧЕНИЯ.ЧЛВК_ИД = 112514.  
-- Вид соединения: INNER JOIN.  
  
SELECT Н_ЛЮДИ.ИД, Н_ОБУЧЕНИЯ.НЗК, Н_УЧЕНИКИ.ГРУППА  
FROM Н_ЛЮДИ  
INNER JOIN Н_ОБУЧЕНИЯ ON Н_ОБУЧЕНИЯ.ЧЛВК_ИД = Н_ЛЮДИ.ИД  
INNER JOIN Н_УЧЕНИКИ ON Н_УЧЕНИКИ.ЧЛВК_ИД = Н_ЛЮДИ.ИД  
WHERE Н_ЛЮДИ.ИД > 152862 AND  
    Н_ОБУЧЕНИЯ.ЧЛВК_ИД = 112514;
```

Индексы, добавление которых уменьшит время выполнения запроса:

1. В таблице Н_ЛЮДИ

- Индекс на атрибуте ИД (B-tree). Это ускорит выполнение фильтрации строк по условию “Н_ЛЮДИ.ИД > 152862”

```
CREATE INDEX Н_ЛЮДИ_ИД ON Н_ЛЮДИ USING btree(ИД);
```

2. В таблице Н_ОБУЧЕНИЯ

- Индекс на атрибуте ЧЛВК_ИД (B-tree). Это ускорит выполнение фильтрации строк по условию “Н_ОБУЧЕНИЯ.ЧЛВК_ИД = 112514” а также соединения таблиц по атрибуту ЧЛВК_ИД

```
CREATE INDEX Н_ОБУЧЕНИЯ_ЧЛВК_ИД ON Н_ОБУЧЕНИЯ USING btree(ЧЛВК_ИД);
```

3. В таблице Н_УЧЕНИКИ

- Индекс на атрибуте ЧЛВК_ИД (B-tree). Это ускорит выполнение соединения таблиц по атрибуту ЧЛВК_ИД

```
CREATE INDEX Н_УЧЕНИКИ_ЧЛВК_ИД ON Н_УЧЕНИКИ USING btree(ЧЛВК_ИД);
```

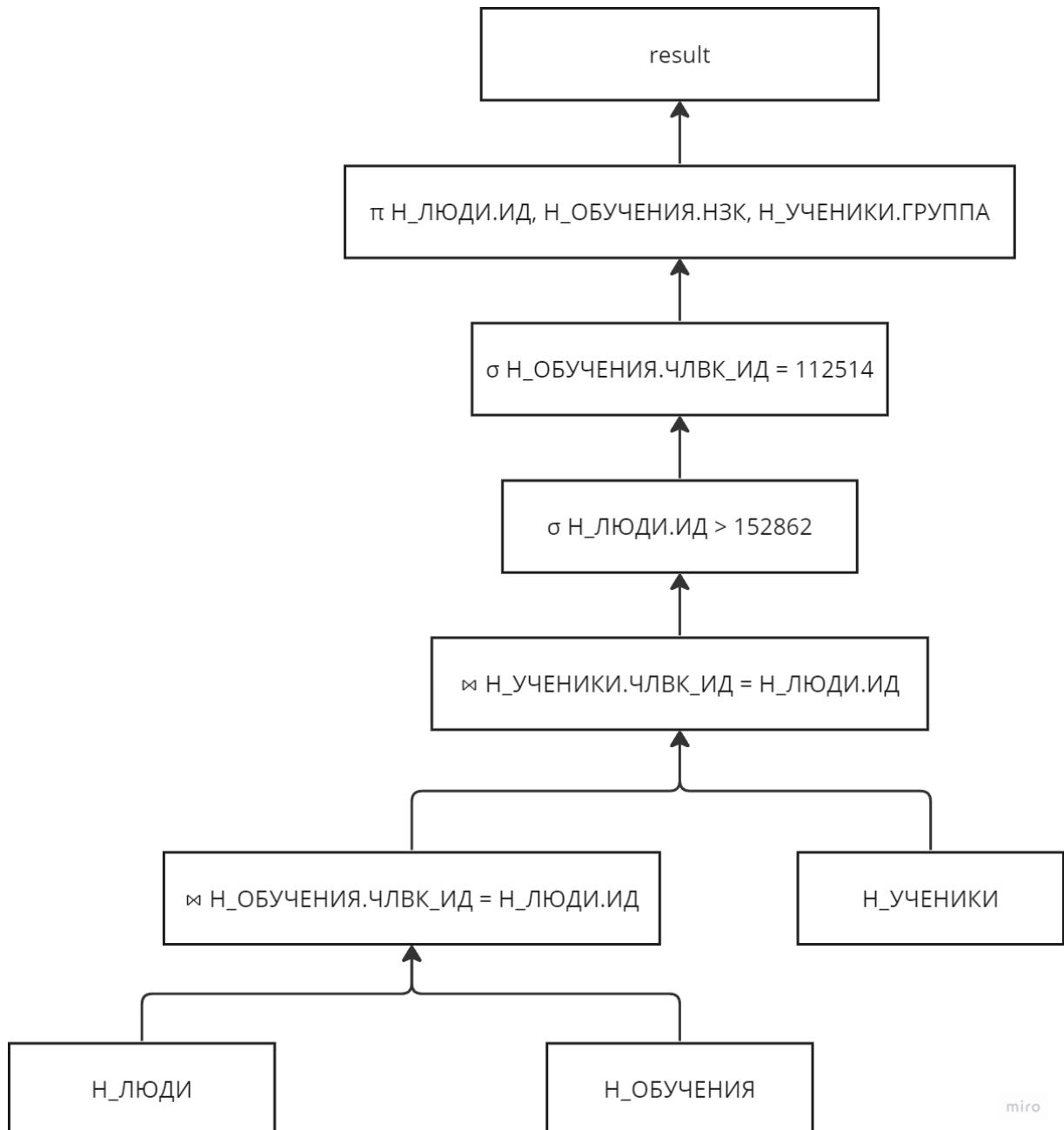
Добавление индексов на указанные столбцы позволит ускорить поиск нужных записей по фильтрам в запросе, а также выполнение соединений.

Возможные планы выполнения запроса без индексов:

План 1:

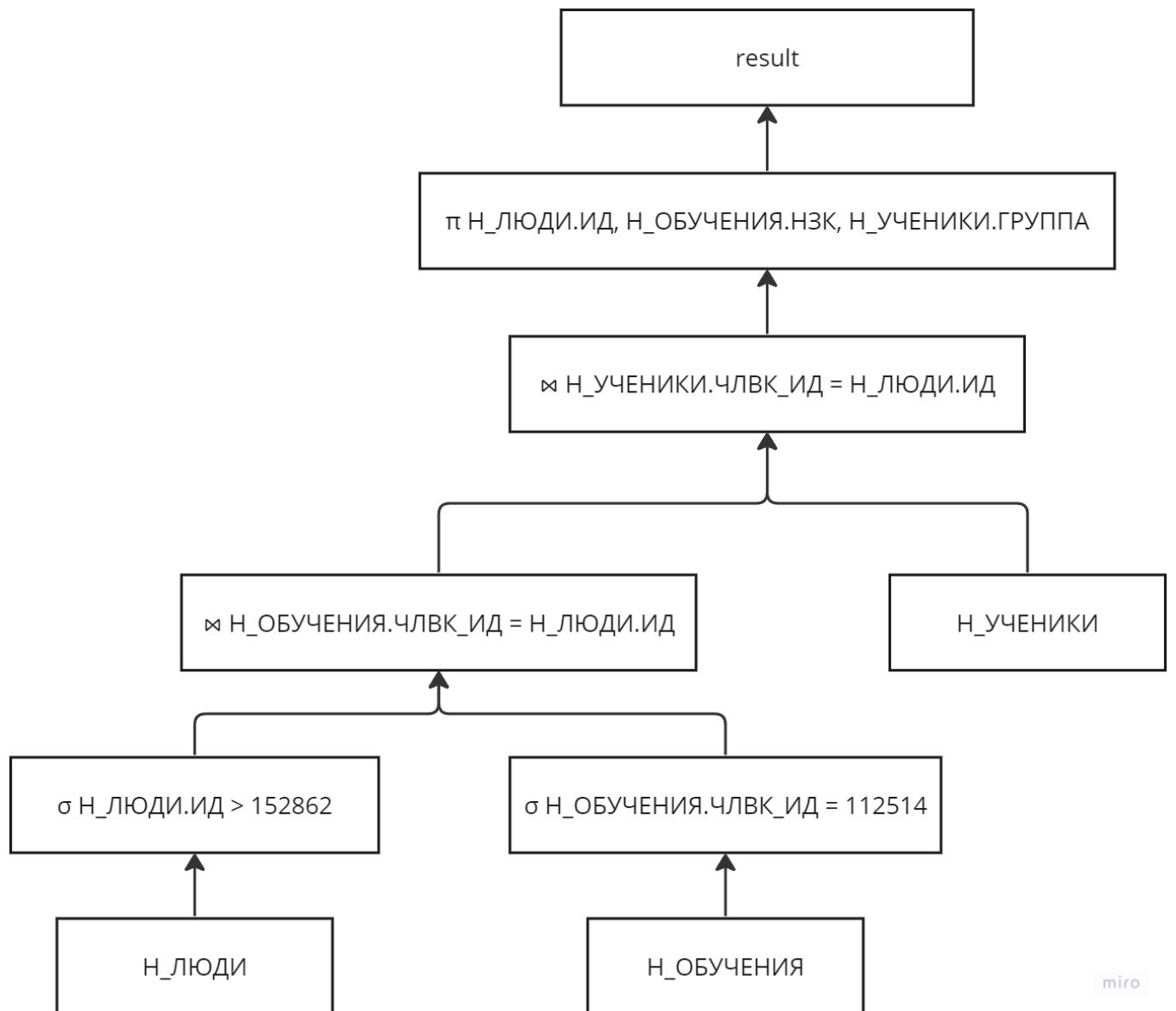
- Полный скан таблицы Н_ЛЮДИ

- Полный скан таблицы Н_ОБУЧЕНИЯ
- Соединение таблиц с использованием Nested Loop Join по атрибуту ЧЛВК_ИД
- Полный скан таблицы Н_УЧЕНИКИ
- Соединение результатов предыдущего соединения с таблицей Н_УЧЕНИКИ с использованием Nested Loop Join по атрибуту ЧЛВК_ИД
- Фильтрация результата соединения по условию “Н_ЛЮДИ.ИД > 152862”
- Фильтрация результата соединения по условию “Н_ОБУЧЕНИЯ.ЧЛВК_ИД = 112514”



План 2:

- Полный скан таблицы Н_ЛЮДИ с применением фильтра по условию “Н_ЛЮДИ.ИД > 152862”
- Полный скан таблицы Н_ОБУЧЕНИЯ с применением фильтра по условию “Н_ОБУЧЕНИЯ.ЧЛВК_ИД = 112514”
- Соединение таблиц с использованием Nested Loop Join по атрибуту ЧЛВК_ИД
- Полный скан таблицы Н_УЧЕНИКИ
- Соединение результатов предыдущего соединения с таблицей Н_УЧЕНИКИ с использованием Nested Loop Join по атрибуту ЧЛВК_ИД



Оптимальный план:

- План 2, так как фильтрация данных выполняется до соединения таблиц, что уменьшает количество строк для обработки

При добавлении индексов планы выполнения запросов изменятся:

- Вместо полного скана таблиц будет использован индексный скан
- Nested Loop Join станет быстрее благодаря индексу на атрибуте ЧЛВК_ИД

QUERY PLAN

Nested Loop (cost=4.89..35.61 rows=5 width=14) (actual time=0.014..0.015 rows=0 loops=1)

-> Nested Loop (cost=0.56..12.61 rows=1 width=10) (actual time=0.013..0.014 rows=0 loops=1)

-> Index Only Scan using "ЧЛВК_PK" on "Н_ЛЮДИ" (cost=0.28..4.30 rows=1 width=4) (actual time=0.013..0.013 rows=0 loops=1)

Index Cond: ((ИД > 152862) AND (ИД = 112514))

Heap Fetches: 0

-> Index Scan using "ОБУЧ_ЧЛВК_FK_I" on "Н_ОБУЧЕНИЯ" (cost=0.28..8.30 rows=1 width=10) (never executed)

Index Cond: ("ЧЛВК_ИД" = 112514)

-> Bitmap Heap Scan on "Н_УЧЕНИКИ" (cost=4.33..22.95 rows=5 width=8) (never executed)

Recheck Cond: ("ЧЛВК_ИД" = 112514)

-> Bitmap Index Scan on "УЧЕН_ОБУЧ_FK_I" (cost=0.00..4.32 rows=5 width=0) (never executed)

Index Cond: ("ЧЛВК_ИД" = 112514)

Planning Time: 1.198 ms

Execution Time: 0.105 ms

Вывод

В ходе работы я изучил как выполняются запросы в базе данных, научился строить план выполнения запроса с диаграммой, а также проводить его оптимизацию. Я изучил различные типы индексов и их влияние на скорость выполнения запроса. Также я познакомился с конструкциями SQL позволяющими определить план выполнения запроса.