

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

Факультет Программной Инженерии и Компьютерной Техники



Лабораторная работа № 4

Вариант № 313104

Выполнил:

Карташев Владимир Р3131

Преподаватель:

Наумова Надежда Александровна

г. Санкт-Петербург, 2023 г.

Для выполнения лабораторной работы №4 необходимо:

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

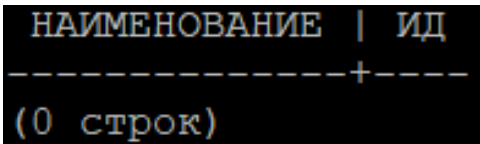
Для каждого запроса предложить индексы, добавление которых уменьшит время выполнения запроса (указать таблицы/атрибуты, для которых нужно добавить индексы, написать тип индекса; объяснить, почему добавление индекса будет полезным для данного запроса).

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.

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.	<p>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.ДАТА. Фильтры (AND): а) Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ > Перезачет. б) Н_ВЕДОМОСТИ.ЧЛВК_ИД > 153285. с) Н_ВЕДОМОСТИ.ЧЛВК_ИД = 142390. Вид соединения: RIGHT JOIN.</p>
	<pre>SELECT Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ, Н_ВЕДОМОСТИ.ИД FROM Н_ТИПЫ_ВЕДОМОСТЕЙ RIGHT JOIN Н_ВЕДОМОСТИ ON Н_ТИПЫ_ВЕДОМОСТЕЙ.ИД = Н_ВЕДОМОСТИ.ТВ_ИД WHERE Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ > 'Перезачет' AND Н_ВЕДОМОСТИ.ЧЛВК_ИД > 153285 AND Н_ВЕДОМОСТИ.ЧЛВК_ИД = 142390;</pre>
	

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.ЧЛВК_ИД, Н_СЕССИЯ.УЧГОД.
Фильтры (AND):
а) Н_ЛЮДИ.ФАМИЛИЯ = Петров.
б) Н_ВЕДОМОСТИ.ЧЛВК_ИД > 142390.
Вид соединения: INNER JOIN.

```
SELECT Н_ЛЮДИ.ИД, Н_ВЕДОМОСТИ.ЧЛВК_ИД, Н_СЕССИЯ.УЧГОД
FROM Н_ЛЮДИ
INNER JOIN Н_ВЕДОМОСТИ
ON Н_ЛЮДИ.ИД = Н_ВЕДОМОСТИ.ЧЛВК_ИД
INNER JOIN Н_СЕССИЯ ON Н_ВЕДОМОСТИ.СЭС_ИД = Н_СЕССИЯ.СЭС_ИД

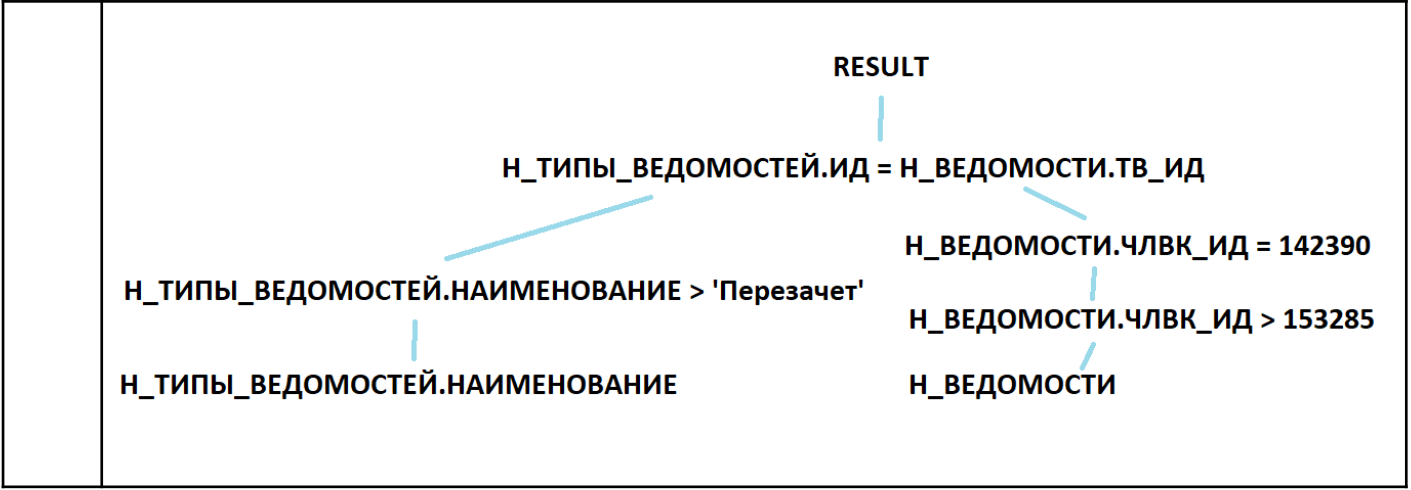
WHERE Н_ЛЮДИ.ФАМИЛИЯ = 'Петров'
AND Н_ВЕДОМОСТИ.ЧЛВК_ИД > 142390;
```

ИД	ЧЛВК_ИД	УЧГОД
146111	146111	2009/2010
146111	146111	2009/2010
146111	146111	2010/2011
146111	146111	2010/2011
146111	146111	2010/2011
146111	146111	2010/2011
146111	146111	2010/2011
146111	146111	2010/2011
157307	157307	2010/2011
157307	157307	2010/2011
157307	157307	2010/2011
157307	157307	2010/2011
157307	157307	2010/2011
157307	157307	2010/2011
(141 строка)		

3.	Для запроса 1 необходимо добавить в отчет вывод команды EXPLAIN ANALYZE [запрос]
	<p>EXPLAIN ANALYZE</p> <p>SELECT Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ, Н_ВЕДОМОСТИ.ИД FROM Н_ТИПЫ_ВЕДОМОСТЕЙ RIGHT JOIN Н_ВЕДОМОСТИ ON Н_ТИПЫ_ВЕДОМОСТЕЙ.ИД = Н_ВЕДОМОСТИ.ТВ_ИД</p> <p>WHERE Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ > 'Перезачет' AND Н_ВЕДОМОСТИ.ЧЛВК_ИД > 153285 AND Н_ВЕДОМОСТИ.ЧЛВК_ИД = 142390;</p>
<pre>учеб=> \i analyze1request.sql QUERY PLAN ----- Nested Loop (cost=0.29..15.34 rows=1 width=422) (actual time=0.031..0.031 rows=0 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=422) (actual time=0.023..0.025 rows=1 loops=1) Filter: (("НАИМЕНОВАНИЕ")::text > 'Перезачет')::text Rows Removed by Filter: 2 -> Index Scan using "ВЕД_ЧЛВК_ФК_ИФК" on "Н_ВЕДОМОСТИ" (cost=0.29..14.26 rows=3 width=8) (actual time=0.004..0.004 rows=0 loops=1) Index Cond: (("ЧЛВК_ИД" > 153285) AND ("ЧЛВК_ИД" = 142390)) Planning Time: 0.223 ms Execution Time: 0.060 ms (9 строк)</pre>	

4.	Для запроса 2 необходимо добавить в отчет вывод команды EXPLAIN ANALYZE [запрос]
	<p>EXPLAIN ANALYZE</p> <p>SELECT Н_ЛЮДИ.ИД, Н_ВЕДОМОСТИ.ЧЛВК_ИД, Н_СЕССИЯ.УЧГОД FROM Н_ЛЮДИ INNER JOIN Н_ВЕДОМОСТИ ON Н_ЛЮДИ.ИД = Н_ВЕДОМОСТИ.ЧЛВК_ИД INNER JOIN Н_СЕССИЯ ON Н_ВЕДОМОСТИ.СЭС_ИД = Н_СЕССИЯ.СЭС_ИД</p> <p>WHERE Н_ЛЮДИ.ФАМИЛИЯ = 'Петров' AND Н_ВЕДОМОСТИ.ЧЛВК_ИД > 142390;</p>
<pre>ucheб=> \i analyze2request.sql QUERY PLAN ----- Nested Loop (cost=4.99..884.06 rows=91 width=18) (actual time=0.072..0.837 rows=141 loops=1) -> Nested Loop (cost=4.71..832.72 rows=150 width=12) (actual time=0.057..0.385 rows=267 loops=1) -> Bitmap Heap Scan on "Н_ЛЮДИ" (cost=4.41..49.43 rows=17 width=4) (actual time=0.036..0.055 rows=17 loops=1) Recheck Cond: (("ФАМИЛИЯ")::text = 'Петров')::text) Heap Blocks: exact=12 -> Bitmap Index Scan on "ФАМ_ЛЮД" (cost=0.00..4.41 rows=17 width=0) (actual time=0.027..0.027 rows=17 loops=1) Index Cond: (("ФАМИЛИЯ")::text = 'Петров')::text) -> Index Scan using "ВЕД_ЧЛВК_ФК_ИФК" on "Н_ВЕДОМОСТИ" (cost=0.29..45.94 rows=14 width=8) (actual time=0.001..0.016 rows=16 loops=1) Index Cond: (("ЧЛВК_ИД" = "Н_ЛЮДИ"."ИД") AND ("ЧЛВК_ИД" > 142390)) -> Index Scan using "СЕС_СЭС_ФК" on "Н_СЕССИЯ" (cost=0.28..0.32 rows=2 width=14) (actual time=0.001..0.001 rows=1 loops=267) Index Cond: ("СЭС_ИД" = "Н_ВЕДОМОСТИ"."СЭС_ИД") Planning Time: 0.580 ms Execution Time: 0.893 ms (13 строк)</pre>	

5.	Для запроса 1 необходимо составить возможный план выполнения запроса.
	<p>Индексы, добавление которых уменьшит время выполнения запроса:</p> <ul style="list-style-type: none"> • Н_ТИПЫ_ВЕДОМОСТЕЙ: <ol style="list-style-type: none"> 1. Индекс на атрибуте НАИМЕНОВАНИЕ (B-tree). Это ускорит фильтрацию строк с условием Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ > 'Перезачет'. • Н_ВЕДОМОСТИ: <ol style="list-style-type: none"> 1. Индекс на атрибуте ЧЛВК_ИД (B-tree). Это ускорит фильтрацию строк с условием ЧЛВК_ИД > 153285 и ЧЛВК_ИД = 142390. 2. Индекс на атрибуте ТВ_ИД (B-tree). Это ускорит соединение строк по атрибуту ТВ_ИД. <p>Возможные планы выполнения запроса без индексов:</p> <ol style="list-style-type: none"> 1. -Полное сканирование таблицы Н_ТИПЫ_ВЕДОМОСТЕЙ <ul style="list-style-type: none"> -Полное сканирование таблицы Н_ВЕДОМОСТИ по условиям: Н_ВЕДОМОСТИ.ЧЛВК_ИД > 153285 и Н_ВЕДОМОСТИ.ЧЛВК_ИД = 142390 -Соединение таблиц через Nested Loop Join по атрибуту ТВ_ИД -Фильтрация результата соединений по условию: Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ > 'Перезачет' 2. -Полное сканирование таблицы Н_ТИПЫ_ВЕДОМОСТЕЙ по условию: Н_ТИПЫ_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ > 'Перезачет' <ul style="list-style-type: none"> -Полное сканирование таблицы Н_ВЕДОМОСТИ по условиям: Н_ВЕДОМОСТИ.ЧЛВК_ИД > 153285 и Н_ВЕДОМОСТИ.ЧЛВК_ИД = 142390 -Соединение таблиц через Nested Loop Join по атрибуту ТВ_ИД <p>Оптимальным планом выполнения данного запроса является план #2, потому что фильтрация происходит до соединения, что уменьшает объем обработки строк.</p> <p>При добавлении индексов план выполнения запросов изменится:</p> <ul style="list-style-type: none"> -Вместо полного сканирования таблицы будет использоваться индексный скан -Nested Loop Join станет быстрее благодаря индексу на атрибуте ТВ_ИД



6.	Для запроса 2 необходимо составить возможный план выполнения запроса.
	<p>Индексы, добавление которых уменьшит время выполнения запроса:</p> <ul style="list-style-type: none"> • Н_ЛЮДИ: <ol style="list-style-type: none"> 1. Индекс на атрибуте ФАМИЛИЯ (B-tree). Это ускорит фильтрацию строк с условием Н_ЛЮДИ.ФАМИЛИЯ > 'Петров'. • Н_ВЕДОМОСТИ: <ol style="list-style-type: none"> 1. Индекс на атрибуте ЧЛВК_ИД (B-tree). Это ускорит фильтрацию строк с условием ЧЛВК_ИД > 142390. • Н_СЕССИЯ: <ol style="list-style-type: none"> 1. Индекс на атрибуте СЭС_ИД (B-tree). Это ускорит соединение строк по атрибуту СЭС_ИД. <p>Возможные планы выполнения запроса без индексов:</p> <ol style="list-style-type: none"> 1. -Полное сканирование таблицы Н_ЛЮДИ по условию: Н_ЛЮДИ.ФАМИЛИЯ = 'Петров'. <ul style="list-style-type: none"> -Полное сканирование таблицы Н_ВЕДОМОСТИ по условию: Н_ВЕДОМОСТИ.ЧЛВК_ИД > 142390. -Соединение таблиц Н_ЛЮДИ И Н_ВЕДОМОСТИ через Nested Loop Join по атрибуту ЧЛВК_ИД. -Полное сканирование таблицы Н_СЕССИЯ. -Соединение соединения Н_ЛЮДИ и Н_ВЕДОМОСТИ с таблицей Н_СЕССИЯ через Nested Loop Join. 2. -Полное сканирование таблицы Н_ЛЮДИ по условию: Н_ЛЮДИ.ФАМИЛИЯ = 'Петров'. <ul style="list-style-type: none"> -Полное сканирование таблицы Н_ВЕДОМОСТИ по условию: Н_ВЕДОМОСТИ.ЧЛВК_ИД > 142390. -Соединение таблиц Н_ЛЮДИ И Н_ВЕДОМОСТИ через Hash Join по атрибуту ЧЛВК_ИД. -Полное сканирование таблицы Н_СЕССИЯ. -Соединение соединения Н_ЛЮДИ и Н_ВЕДОМОСТИ с таблицей Н_СЕССИЯ через Hash Join. <p>Оптимальным планом выполнения данного запроса является план #1, потому что Nested Loop Join будет работать лучше, чем Hash Join в случаях, когда у нас небольшой обрабатываемый набор данных и выводимое количество строк.</p>

При добавлении индексов план выполнения запросов изменится:

-Вместо полного сканирования таблицы будет использоваться индексный скан

-Nested Loop Join станет быстрее благодаря индексу на атрибуте **ЧЛВК_ИД** и **СЭС_ИД**.

