

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

Факультет программной инженерии и компьютерной техники

Лабораторная работа №4

Вариант 42144

Студент:
Кулагин Вячеслав, Р3109
Преподаватель:
Райла Мартин

Санкт-Петербург
2024

Оглавление

Задание	3
Реализация запросов на SQL	4
Ответы на вопросы, представленные в задании	4
Планы выполнения запросов	7
Вывод	8

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.

Изменяться 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).

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.УЧГОД.

Фильтры (AND):

а) Н_ЛЮДИ.ФАМИЛИЯ > Иванов.

б) Н_СЕССИЯ.ДАТА = 2002-01-04.

Вид соединения: RIGHT JOIN.

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.ИД, Н_СЕССИЯ.ДАТА.

Фильтры (AND):

а) Н_ЛЮДИ.ОТЧЕСТВО > Сергеевич.

б) Н_ВЕДОМОСТИ.ДАТА = 2010-06-18.

с) Н_СЕССИЯ.ЧЛВК_ИД < 126631.

Вид соединения: INNER JOIN.

Реализация запросов на SQL

Запрос №1

```
SELECT "Н_ЛЮДИ"."ИД", "Н_СЕССИЯ"."УЧГОД"  
FROM "Н_ЛЮДИ" RIGHT JOIN "Н_СЕССИЯ" ON "Н_ЛЮДИ"."ИД" = "Н_СЕССИЯ"."ЧЛВК_ИД"  
WHERE "ФАМИЛИЯ" > 'Иванов' AND "ДАТА" = '2002-01-04';
```

Запрос №2

```
SELECT "Н_ЛЮДИ"."ИД", "Н_ВЕДОМОСТИ"."ИД", "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"."ОТЧЕСТВО" > 'Сергеевич'  
AND "Н_ВЕДОМОСТИ"."ДАТА" = '2010-06-18'  
AND "Н_СЕССИЯ"."ЧЛВК_ИД" < 126631;
```

Ответы на вопросы, представленные в задании

Предлагаемые индексы для запроса №1

- В таблицу Н_ЛЮДИ можно добавить индекс на столбец ФАМИЛИЯ типа B-Tree. Это отношение достаточно большое, при этом мы используем сравнение, значит будет эффективно использовать именно B-Tree
- В таблицу Н_СЕССИЯ можно добавить индекс на столбец ДАТА типа Hash. Это отношение достаточно большое, при этом активно используется равенство, что эффективно при использовании Hash-индексов.

Предлагаемые индексы для запроса №2

- В таблицу Н_ЛЮДИ на столбец ОТЧЕСТВО можно добавить индекс типа B-Tree, потому что отношение достаточно большое, но при этом используется оператор сравнения
- В таблицу Н_ВЕДОМОСТИ на столбец ДАТА можно добавить индекс Hash, потому что отношение достаточно большое, при этом используется оператор равенства
- В таблицу Н_СЕССИЯ на столбец ЧЛВК_ИД можно добавить индекс типа B-Tree, потому что отношение достаточно большое, но при этом используется оператор сравнения

Вывод EXPLAIN ANALYZE для запроса №1

Nested Loop (cost=0.28..159.42 rows=3 width=14) (actual time=0.076..0.453 rows=7 loops=1)

" -> Seq Scan on ""H_СЕССИЯ"" (cost=0.00..117.90 rows=5 width=14) (actual time=0.065..0.421 rows=7 loops=1)"

" Filter: (""ДАТА"" = '2002-01-04 00:00:00'::timestamp without time zone)"

Rows Removed by Filter: 3745

" -> Index Scan using ""ЧЛВК_РК"" on ""H_ЛЮДИ"" (cost=0.28..8.30 rows=1 width=4) (actual time=0.004..0.004 rows=1 loops=7)"

" Index Cond: (""ИД"" = ""H_СЕССИЯ"".""ЧЛВК_ИД"")"

" Filter: ((""ФАМИЛИЯ"")::text > 'Иванов'::text)"

Planning Time: 0.363 ms

Execution Time: 0.481 ms

Вывод EXPLAIN ANALYZE для запроса №2

Nested Loop (cost=169.92..394.86 rows=3 width=16) (actual time=2.950..2.952 rows=0 loops=1)

-> Hash Join (cost=169.64..386.36 rows=6 width=12) (actual time=2.834..2.937 rows=17 loops=1)

" Hash Cond: (""H_ВЕДОМОСТИ"".""ЧЛВК_ИД"" = ""H_ЛЮДИ"".""ИД"")"

" -> Index Scan using ""ВЕД_ДАТА_I"" on ""H_ВЕДОМОСТИ"" (cost=0.29..216.82 rows=73 width=8) (actual time=0.034..0.119 rows=141 loops=1)"

" Index Cond: (""ДАТА"" = '2010-06-18 00:00:00'::timestamp without time zone)"

-> Hash (cost=163.97..163.97 rows=430 width=4) (actual time=2.770..2.771 rows=430 loops=1)

Buckets: 1024 Batches: 1 Memory Usage: 24kB

" -> Seq Scan on ""H_ЛЮДИ"" (cost=0.00..163.97 rows=430 width=4) (actual time=0.011..2.687 rows=430 loops=1)"

" Filter: ((""ОТЧЕСТВО"")::text > 'Сергеевич'::text)"

Rows Removed by Filter: 4688

" -> Index Scan using ""SYS_C003500_IFK"" on ""H_СЕРВИС"" (cost=0.28..1.28 rows=14 width=12) (actual time=0.000..0.000 rows=0 loops=1)"

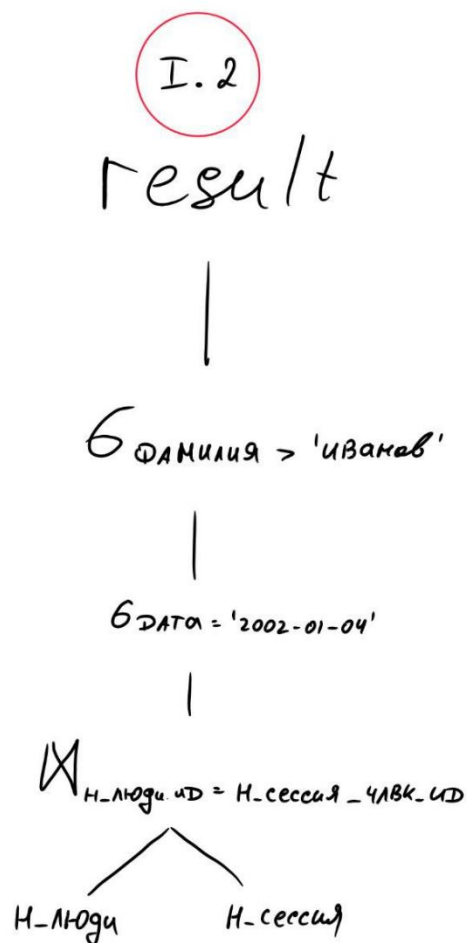
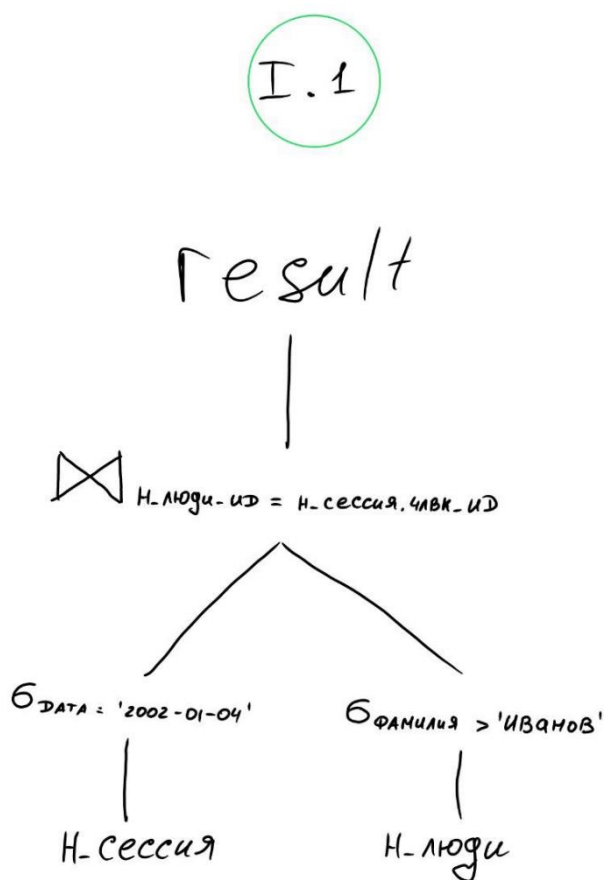
" Index Cond: ((""ЧЛВК_ИД"" = ""H_ЛЮДИ"".""ИД"") AND (""ЧЛВК_ИД"" < 126631))"

Planning Time: 1.105 ms

Execution Time: 3.016 ms

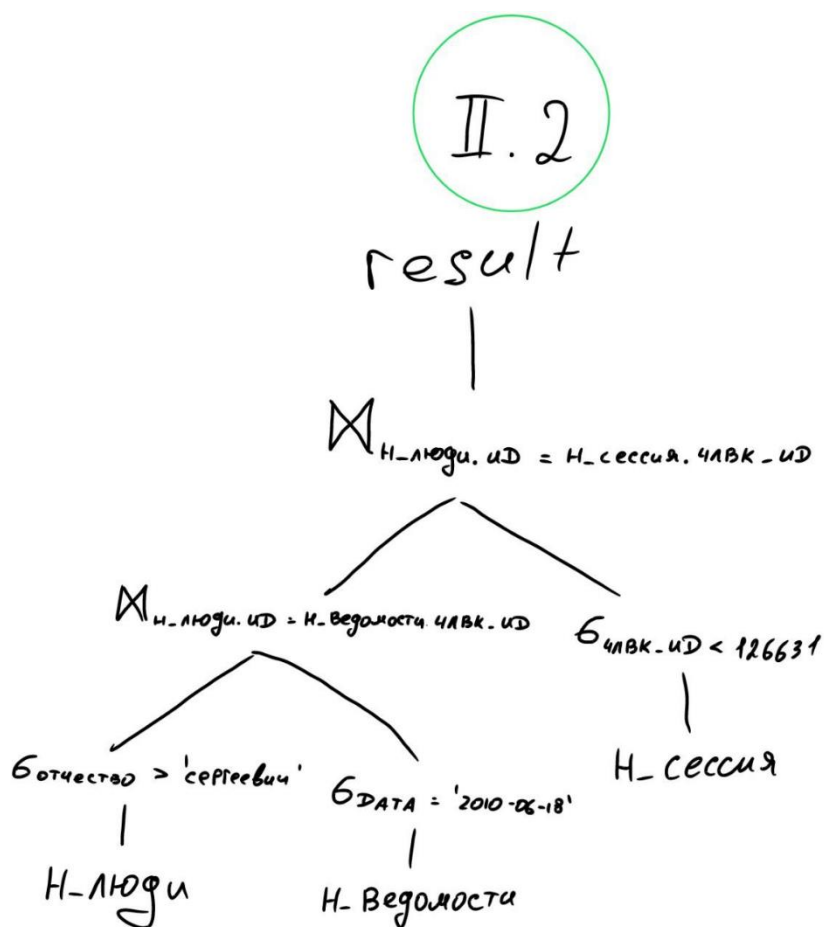
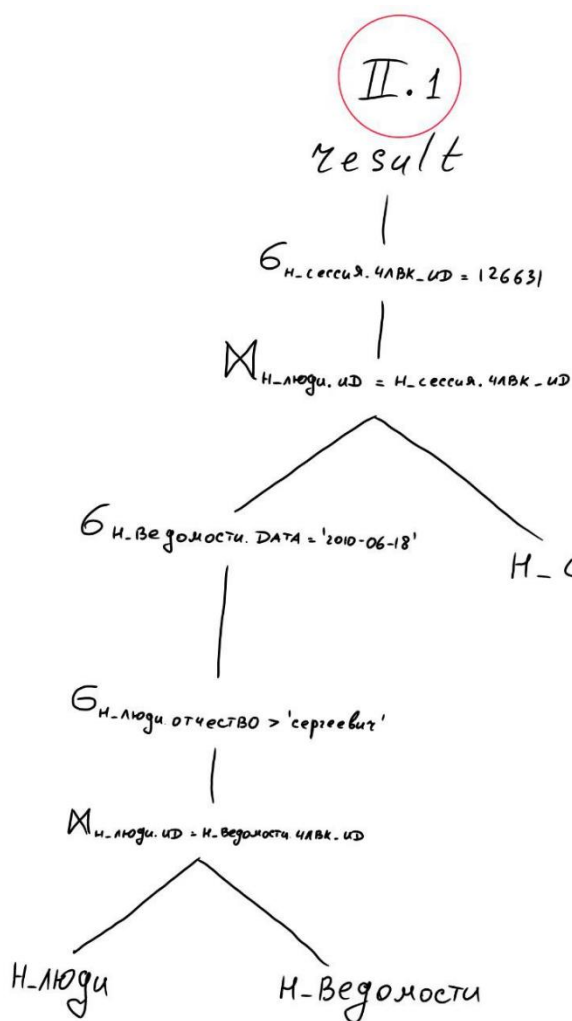
Планы выполнения запросов

Предлагаемые планы для выполнения запроса №1



Из предложенных вариантов план I.1 более эффективен, потому что в нем фильтрация происходит до объединения таблиц, а значит, она будет занимать меньше времени, ведь количество данных до объединения меньше, чем после. При добавлении индексов в план, глобально он не изменится, изменится лишь Seq Scan на Index Scan.

Предлагаемые планы для выполнения запроса №2



Из предложенных вариантов план II.2 более эффективен, потому что в нем фильтрация происходит до объединения таблиц, а значит, она будет занимать меньше времени, ведь количество данных до объединения меньше, чем после. При добавлении индексов в план, глобально он не изменится, изменится лишь Seq Scan на Index Scan.

Вывод

Проведя эту работу, я узнал, какие существуют индексы, чем они различаются, а также что такое и как составляется план выполнения запроса.