УНИВЕРСИТЕТ ИТМО

Факультет программной инженерии и компьютерной техники

Направление подготовки 09.03.04 Программная инженерия

Дисциплина «Информационный системы и базы данных»

**Лабораторная работа №4**

*Вариант 3703*

Студент

*Митрофанов Е. Ю.*

*P33101*

Преподаватель

*Николаев В. В.*

Санкт-Петербург, 2021 г.

Описание 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]

Реализация первого запроса

Сделать запрос для получения атрибутов из указанных таблиц, применив фильтры по указанным условиям:  
Таблицы: Н\_ЛЮДИ, Н\_СЕССИЯ.  
Вывести атрибуты: Н\_ЛЮДИ.ИМЯ, Н\_СЕССИЯ.УЧГОД.  
Фильтры (AND):  
a) Н\_ЛЮДИ.ОТЧЕСТВО = Сергеевич.  
b) Н\_СЕССИЯ.ЧЛВК\_ИД = 105948.  
Вид соединения: LEFT JOIN.

SELECT Н\_ЛЮДИ.ИМЯ, Н\_СЕССИЯ.УЧГОД  
FROM Н\_ЛЮДИ  
 LEFT JOIN Н\_СЕССИЯ ON Н\_ЛЮДИ.ИД = Н\_СЕССИЯ.ЧЛВК\_ИД  
WHERE ( Н\_ЛЮДИ.ОТЧЕСТВО = 'Сергеевич' AND Н\_СЕССИЯ.ЧЛВК\_ИД = 105948);

*Планы выполнения:*

result

Н\_ЛЮДИ.ОТЧЕСТВО = Сергеевич

Н\_ЛЮДИ

Н\_СЕССИЯ

Н\_СЕССИЯ.ЧЛВК\_ИД = 105948

Н\_ЛЮДИ.ИМЯ, Н\_СЕССИЯ.УЧГОД

Н\_ЛЮДИ.ИД = Н\_СЕССИЯ.ЧЛВК\_ИД

Н\_СЕССИЯ.ЧЛВК\_ИД = 105948

Н\_ЛЮДИ.ОТЧЕСТВО = Сергеевич

result

Н\_ЛЮДИ.ИМЯ, Н\_СЕССИЯ.УЧГОД

Н\_ЛЮДИ.ИД = Н\_СЕССИЯ.ЧЛВК\_ИД

Н\_ЛЮДИ

Н\_СЕССИЯ

Во втором плане происходит объединение только нужной выборки, а не всех сущностей. Размер промежуточных данных меньше, значит этот план является оптимальным.

*Индексы:*

CREATE INDEX ON "Н\_ЛЮДИ" USING HASH ("ИД");  
CREATE INDEX ON "Н\_ЛЮДИ" USING BTREE ("ОТЧЕСТВО");  
CREATE INDEX ON "Н\_СЕССИЯ" USING BTREE ("ЧЛВК\_ИД");

Выборка происходит с использованием операторов сравнения, поэтому оптимально использование BTREE. Соединение таблиц с использованием ИД – оптимально использование HASH.

*Explain Analyse:*

Nested Loop (cost=4.62..34.20 rows=7 width=23) (actual time=0.024..0.024 rows=0 loops=1)

-> Index Scan using "ЧЛВК\_PK" on "Н\_ЛЮДИ" (cost=0.28..8.30 rows=1 width=17) (actual time=0.023..0.023 rows=0 loops=1)

Index Cond: ("ИД" = 105948)"

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

Rows Removed by Filter: 1

-> Bitmap Heap Scan on "Н\_СЕССИЯ" (cost=4.33..25.83 rows=7 width=14) (never executed)

Recheck Cond: ("ЧЛВК\_ИД" = 105948)

-> Bitmap Index Scan on "SYS\_C003500\_IFK" (cost=0.00..4.33 rows=7 width=0) (never executed)

Index Cond: ("ЧЛВК\_ИД" = 105948)

Planning time: 0.149 ms

Execution time: 0.057 ms

Реализация второго запроса

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):  
a) Н\_ЛЮДИ.ИМЯ = Владимир.  
b) Н\_ВЕДОМОСТИ.ДАТА > 2010-06-18.  
c) Н\_СЕССИЯ.ДАТА > 2012-01-25.  
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\_ЛЮДИ.ИД = Н\_СЕССИЯ.ЧЛВК\_ИД  
WHERE (Н\_ЛЮДИ.ИМЯ = 'Владимир' AND Н\_ВЕДОМОСТИ.ДАТА > '2010-06-18' AND Н\_СЕССИЯ.ДАТА > '2012-01-25');

*Планы выполнения:*

Н\_ВЕДОМОСТИ.ДАТА = '2010-06-18'

Н\_ЛЮДИ.ИМЯ = Владимир

result

Н\_ЛЮДИ.ИД, Н\_СЕССИЯ. ЧЛВК\_ИД, Н\_СЕССИЯ. УЧГОД

Н\_ЛЮДИ.ИД = Н\_ВЕДОМОСТИ.ЧЛВК\_ИД

Н\_ЛЮДИ

Н\_ВЕДОМОСТИ

Н\_СЕССИЯ

Н\_ЛЮДИ.ИД = Н\_СЕССИЯ.ЧЛВК\_ИД

Н\_СЕССИЯ.ДАТА = '2012-01-25'

result

Н\_ЛЮДИ

Н\_СЕССИЯ

Н\_ЛЮДИ.ИД, Н\_ЛЮДИ.ИМЯ

Н\_ВЕДОМОСТИ

Н\_ВЕДОМОСТИ.ЧЛВК\_ИД, Н\_ВЕДОМОСТИ.ДАТА

Н\_ВЕДОМОСТИ.ДАТА = '2010-06-18'

Н\_ЛЮДИ.ИМЯ = Владимир

Н\_ЛЮДИ.ИД = Н\_СЕССИЯ.ЧЛВК\_ИД

Н\_СЕССИЯ.ЧЛВК\_ИД, Н\_СЕССИЯ.ДАТА

Н\_СЕССИЯ.ДАТА = '2012-01-25'

Н\_ЛЮДИ.ИД = Н\_СЕССИЯ.ЧЛВК\_ИД

Н\_ЛЮДИ.ИД, Н\_СЕССИЯ. ЧЛВК\_ИД, Н\_СЕССИЯ. УЧГОД

Второй план является оптимальным. Из-за того, что выборка происходит на более ранних этапах, идет соединение только нужных атрибутов, и размер промежуточных данных меньше.

*Индексы:*

CREATE INDEX ON "Н\_ЛЮДИ" USING HASH ("ИД");  
CREATE INDEX ON "Н\_ЛЮДИ" USING BTREE ("ИМЯ");  
CREATE INDEX ON "Н\_СЕССИЯ" USING BTREE ("ДАТА");  
CREATE INDEX ON "Н\_ВЕДОМОСТИ" USING BTREE ("ДАТА");

Выборка происходит с использованием операторов сравнения, поэтому оптимально использование BTREE. Соединение таблиц с использованием ИД – оптимально использование HASH.

Nested Loop (cost=0.70..238.55 rows=1 width=18) (actual time=1.341..1.341 rows=0 loops=1)

-> Nested Loop (cost=0.28..126.21 rows=1 width=18) (actual time=1.341..1.341 rows=0 loops=1)

-> Seq Scan on "Н\_СЕССИЯ" (cost=0.00..117.90 rows=1 width=14) (actual time=1.339..1.339 rows=0 loops=1)

Filter: ("ДАТА" > '2012-01-25 00:00:00'::timestamp without time zone)

Rows Removed by Filter: 3752

-> Index Scan using "ЧЛВК\_PK" on "Н\_ЛЮДИ" (cost=0.28..8.30 rows=1 width=4) (never executed)

Index Cond: ("ИД" = ""Н\_СЕССИЯ"".""ЧЛВК\_ИД"")

Filter: (("ИМЯ")::text = 'Владимир'::text)

-> Index Scan using "ВЕД\_ЧЛВК\_FK\_IFK" on ""Н\_ВЕДОМОСТИ" (cost=0.42..112.26 rows=8 width=4) (never executed)

Index Cond: ("ЧЛВК\_ИД" = "Н\_ЛЮДИ"."ИД")

Filter: ("ДАТА" > '2010-06-18 00:00:00'::timestamp without time zone)

Выводы

При выполнении лабораторной работы я познакомился с использованием индексов для ускорения обработки запросов в SQL. Изучил, как поэтапно происходит выполнение запроса, какие планы могут использоваться и как индексы влияют на процесс.