

Кафедра вычислительной техники

Информационные системы и базы данных

Лабораторная работа №4

«EXPLAIN INDEX»

Вариант 92637

Преподаватель:

 Байрамова Хумай Бахруз Кызы

Выполнил:   
Горинов Даниил Андреевич

Р33151

Санкт-Петербург

2024

**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.  
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:
   * Таблицы:

Н\_ТИПЫ\_ВЕДОМОСТЕЙ, Н\_ВЕДОМОСТИ.

* + Вывести атрибуты:

Н\_ТИПЫ\_ВЕДОМОСТЕЙ.ИД, Н\_ВЕДОМОСТИ.ДАТА.

* + Фильтры (AND):

a) Н\_ТИПЫ\_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ > Перезачет.

b) Н\_ВЕДОМОСТИ.ЧЛВК\_ИД > 105590.

Вид соединения: RIGHT JOIN.

1. Сделать запрос для получения атрибутов из указанных таблиц, применив фильтры по указанным условиям:

* Таблицы:

Н\_ЛЮДИ, Н\_ОБУЧЕНИЯ, Н\_УЧЕНИКИ.

* Вывести атрибуты:

Н\_ЛЮДИ.ФАМИЛИЯ, Н\_ОБУЧЕНИЯ.НЗК, Н\_УЧЕНИКИ.ИД

* Фильтры: (AND)

a) Н\_ЛЮДИ.ОТЧЕСТВО < Александрович.  
b) Н\_ОБУЧЕНИЯ.НЗК > 933232.  
c) Н\_УЧЕНИКИ.НАЧАЛО = 2009-02-09.

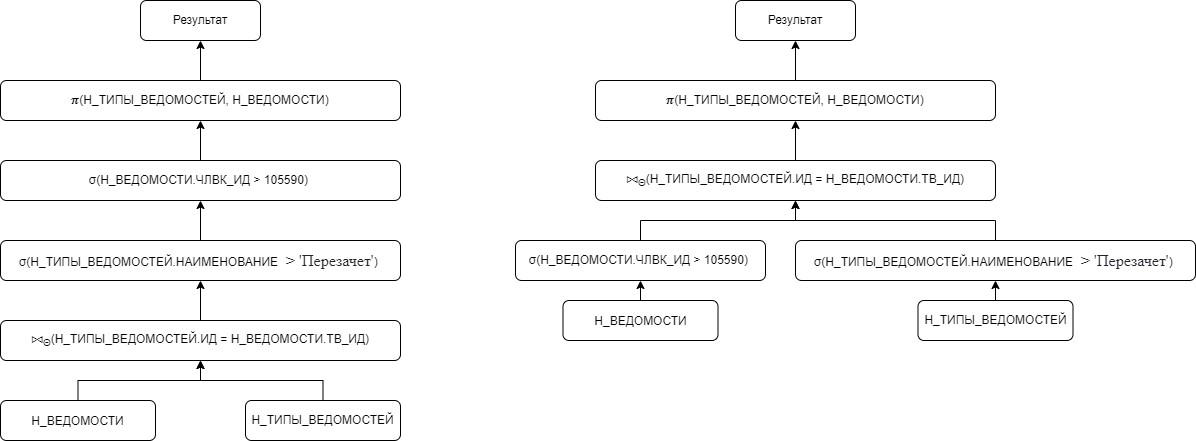
Вид соединения: LEFT JOIN.

**Запросы**

* №1:

|  |
| --- |
| SELECT  ТИПЫ\_ВЕДОМОСТЕЙ.ИД,  ВЕДОМОСТИ.ДАТА  FROM  Н\_ВЕДОМОСТИ AS ВЕДОМОСТИ  RIGHT JOIN Н\_ТИПЫ\_ВЕДОМОСТЕЙ AS ТИПЫ\_ВЕДОМОСТЕЙ ON (  ТИПЫ\_ВЕДОМОСТЕЙ.ИД = ВЕДОМОСТИ.ТВ\_ИД  )  WHERE (  ТИПЫ\_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ > 'Перезачет'  AND ВЕДОМОСТИ.ЧЛВК\_ИД > 105590  ); |

Возможные планы запроса:



Оптимальным планом выполнения запроса является правый, так как происходит объединение только необходимой выборки вместо полного объединения таблиц.

Вывод EXPLAIN ANALYZE:

|  |
| --- |
| QUERY PLAN  --------------------------------------------------------------------------------------------------------------------  Hash Join (cost=1.06..8255.94 rows=148250 width=12) (actual time=0.044..66.770 rows=21468 loops=1)  Hash Cond: ("ВЕДОМОСТИ"."ТВ\_ИД" = "ТИПЫ\_ВЕДОМОСТЕЙ"."ИД")  -> Seq Scan on "Н\_ВЕДОМОСТИ" "ВЕДОМОСТИ" (cost=0.00..6846.50 rows=222375 width=12) (actual time=0.011..35.656 rows=222413 loops=1)  Filter: ("ЧЛВК\_ИД" > 105590)  Rows Removed by Filter: 27  -> Hash (cost=1.04..1.04 rows=2 width=4) (actual time=0.011..0.013 rows=1 loops=1)  Buckets: 1024 Batches: 1 Memory Usage: 9kB  -> Seq Scan on "Н\_ТИПЫ\_ВЕДОМОСТЕЙ" "ТИПЫ\_ВЕДОМОСТЕЙ" (cost=0.00..1.04 rows=2 width=4) (actual time=0.006..0.007 rows=1 loops=1)  Filter: (("НАИМЕНОВАНИЕ")::text > 'Перезачет'::text)  Rows Removed by Filter: 2  Planning Time: 0.516 ms  Execution Time: 67.824 ms |

Индексы:

|  |
| --- |
| CREATE INDEX ВЕДОМОСТИ\_ЧЛВК\_ИД \_ИНДЕКС ON Н\_ ВЕДОМОСТИ USING BTREE(ЧЛВК\_ИД); |

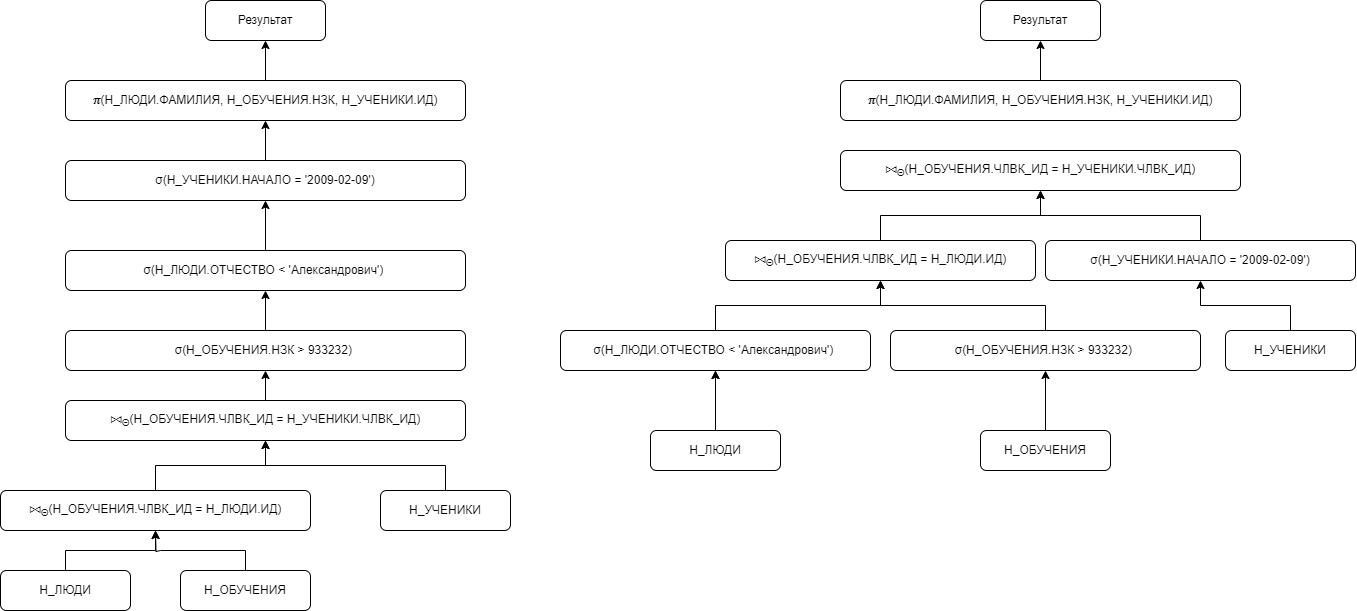
Тип индекса BTREE выбран из-за наличия оператора ‘>’.  
  
Строить индекс по ТИПЫ\_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ не релевантно, поскольку НАИМЕНОВАНИЕ имеет низкую селективность (там всего 3 значения).

В Postgres при создании таблицы создается BTREE индекс по её PK. Поэтому индекс по ТИПЫ\_ВЕДОМОСТЕЙ.ИД мы не создаем.

* №2:

|  |
| --- |
| SELECT  Н\_ЛЮДИ.ФАМИЛИЯ,  Н\_ОБУЧЕНИЯ.НЗК,  Н\_УЧЕНИКИ.НАЧАЛО  FROM  Н\_УЧЕНИКИ  LEFT JOIN Н\_ОБУЧЕНИЯ ON Н\_ОБУЧЕНИЯ.ЧЛВК\_ИД = Н\_УЧЕНИКИ.ЧЛВК\_ИД  LEFT JOIN Н\_ЛЮДИ ON Н\_ОБУЧЕНИЯ.ЧЛВК\_ИД = Н\_ЛЮДИ.ИД  WHERE  Н\_ОБУЧЕНИЯ.НЗК::integer > 933232  AND Н\_ЛЮДИ.ОТЧЕСТВО < 'Александрович'  AND Н\_УЧЕНИКИ.НАЧАЛО = '2009-02-09'; |

Возможные планы запроса:



Оптимальным планом выполнения запроса является правый, так как за счет раннего использования выборки происходит соединение не целых сущностей, а только нужных нам выборок.

Вывод EXPLAIN ANALYZE:

|  |
| --- |
| QUERY PLAN  --------------------------------------------------------------------------------------------------------------------------------  Nested Loop (cost=171.69..278.13 rows=1 width=30) (actual time=3.307..3.309 rows=0 loops=1)  Join Filter: ("Н\_УЧЕНИКИ"."ЧЛВК\_ИД" = "Н\_ОБУЧЕНИЯ"."ЧЛВК\_ИД")  -> Hash Join (cost=171.41..276.55 rows=1 width=32) (actual time=3.268..3.294 rows=1 loops=1)  Hash Cond: ("Н\_УЧЕНИКИ"."ЧЛВК\_ИД" = "Н\_ЛЮДИ"."ИД")  -> Bitmap Heap Scan on "Н\_УЧЕНИКИ" (cost=4.54..109.59 rows=32 width=12) (actual time=0.020..0.045 rows=32 loops=1)  Recheck Cond: ("НАЧАЛО" = '2009-02-09 00:00:00'::timestamp without time zone)  Heap Blocks: exact=15  -> Bitmap Index Scan on "УЧЕН\_НАЧ\_I" (cost=0.00..4.53 rows=32 width=0) (actual time=0.012..0.013 rows=32 loops=1)  Index Cond: ("НАЧАЛО" = '2009-02-09 00:00:00'::timestamp without time zone)  -> Hash (cost=163.97..163.97 rows=232 width=20) (actual time=3.234..3.235 rows=233 loops=1)  Buckets: 1024 Batches: 1 Memory Usage: 21kB  -> Seq Scan on "Н\_ЛЮДИ" (cost=0.00..163.97 rows=232 width=20) (actual time=0.016..3.164 rows=233 loops=1)  Filter: (("ОТЧЕСТВО")::text < 'Александрович'::text)  Rows Removed by Filter: 4885  -> Index Scan using "ОБУЧ\_ЧЛВК\_FK\_I" on "Н\_ОБУЧЕНИЯ" (cost=0.28..1.57 rows=1 width=10) (actual time=0.011..0.011 rows=0 loops=1)  Index Cond: ("ЧЛВК\_ИД" = "Н\_ЛЮДИ"."ИД")  Filter: (("НЗК")::integer > 933232)  Rows Removed by Filter: 1  Planning Time: 0.788 ms  Execution Time: 3.365 ms |

Индексы:

|  |
| --- |
| CREATE INDEX Н\_УЧЕНИКИ\_НАЧАЛО\_ИНДЕКС ON Н\_УЧЕНИКИ USING HASH (НАЧАЛО);  CREATE INDEX Н\_ОБУЧЕНИЯ\_НЗК\_ИНДЕКС ON Н\_ОБУЧЕНИЯ USING BTREE (НЗК);  CREATE INDEX Н\_ЛЮДИ\_ОТЧЕСТВО\_ИНДЕКС ON Н\_ЛЮДИ USING BTREE (ОТЧЕСТВО); |

Добавление этих индексов должно сильно ускорить выполнение запросов, так как в них присутствует выборка с использованием операторов сравнения и соединение таблиц.

**Вывод**

В ходе данной лабораторной работы мы изучили индексы и их влияние на нагрузку системы, ознакомились с планировщиком и его системой выбора оптимального плана выполнения запроса.