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 75752*

Студент

*Кочнев Р.Д.*

*P33081*

Преподаватель

*Шешуков Дмитрий Михайлович*

Санкт-Петербург, 2023 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]

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.ОТЧЕСТВО, Н\_СЕССИЯ.УЧГОД.

Фильтры (AND):

a) Н\_ЛЮДИ.ОТЧЕСТВО > Владимирович.

b) Н\_СЕССИЯ.УЧГОД = 2008/2009.

Вид соединения: LEFT JOIN.

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.ИД.

Фильтры: (AND)

a) Н\_ЛЮДИ.ИД = 163484.

b) Н\_ОБУЧЕНИЯ.ЧЛВК\_ИД = 163484.

c) Н\_УЧЕНИКИ.ГРУППА = 3100.

Вид соединения: INNER JOIN.

Первый запрос

SELECT Н\_ЛЮДИ.ОТЧЕСТВО, Н\_СЕССИЯ.УЧГОД

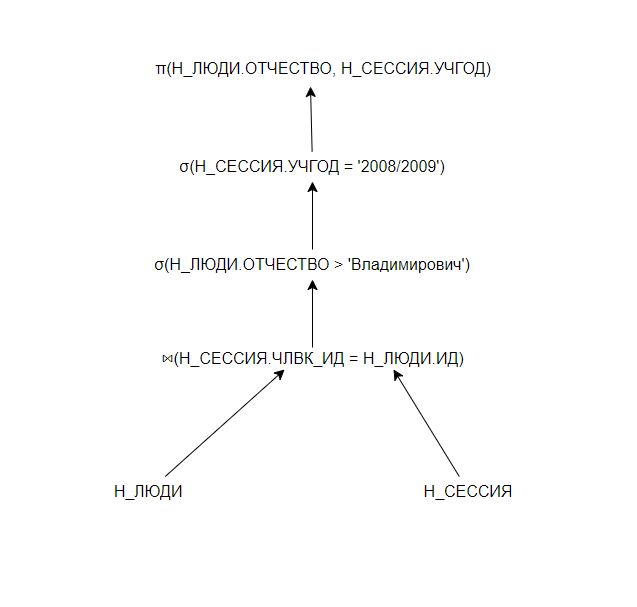
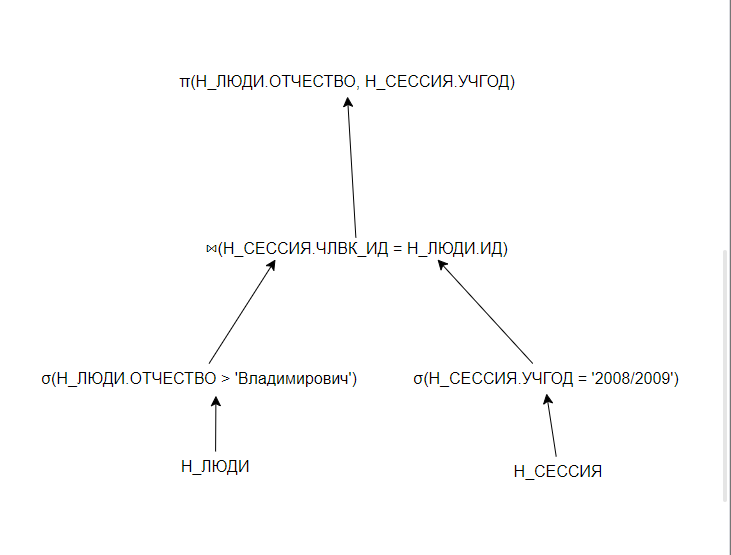
FROM Н\_ЛЮДИ

LEFT JOIN Н\_СЕССИЯ ON Н\_СЕССИЯ.ЧЛВК\_ИД = Н\_ЛЮДИ.ИД

WHERE Н\_ЛЮДИ.ОТЧЕСТВО > 'Владимирович'

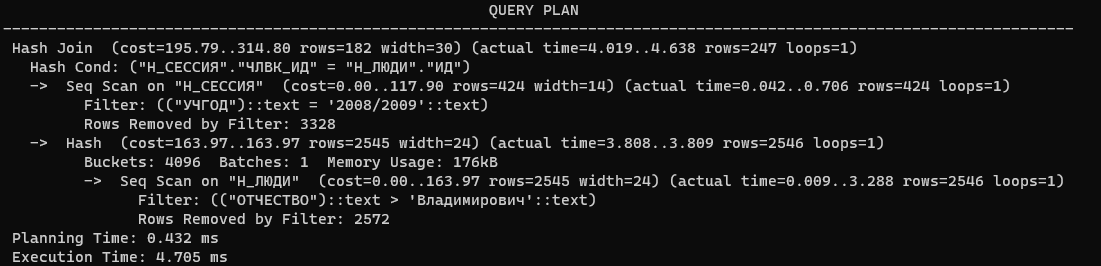
AND Н\_СЕССИЯ.УЧГОД = '2008/2009';

В втором случае выборка происходит раньше чем во втором, поэтому размер промежуточных данных будет меньше => будет меньше число операций чтения/записи во внешнюю память.

Добавим HASH индексы на Н\_ЛЮДИ.ИД и Н\_СЕССИЯ.ЧЛВК\_ИД, т.к. используется прямое сравнения на ИД. Ещё один индекс Н\_СЕССИЯ.УЧГОД бесполезен в случае конвейерной обработки. Тогда можно будет построить следующий план:

# EXPLAIN ANALYZE



CREATE INDEX ON "Н\_ЛЮДИ" USING HASH("ИД");

CREATE INDEX ON "Н\_СЕССИЯ" USING HASH("ЧЛВК\_ИД");

Второй запрос

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

FROM Н\_ЛЮДИ

JOIN Н\_ОБУЧЕНИЯ ON Н\_ОБУЧЕНИЯ.ЧЛВК\_ИД = Н\_ЛЮДИ.ИД

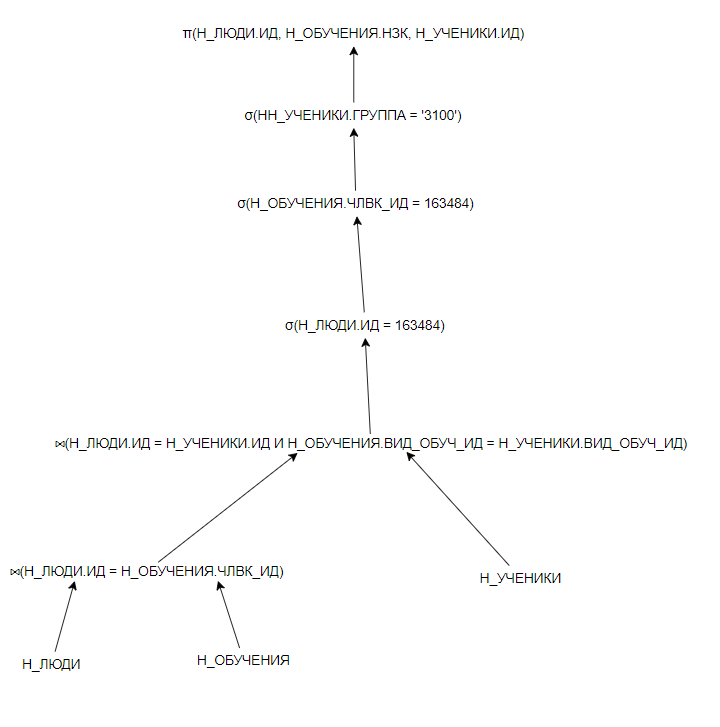
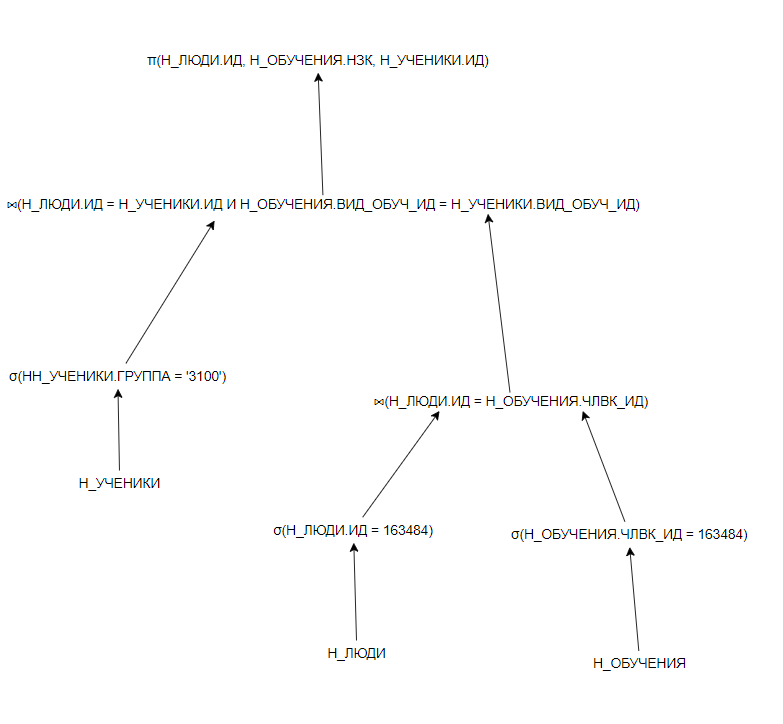
JOIN Н\_УЧЕНИКИ ON Н\_УЧЕНИКИ.ЧЛВК\_ИД = Н\_ЛЮДИ.ИД

AND Н\_УЧЕНИКИ.ВИД\_ОБУЧ\_ИД = Н\_ОБУЧЕНИЯ.ВИД\_ОБУЧ\_ИД

WHERE Н\_ЛЮДИ.ИД = 163484

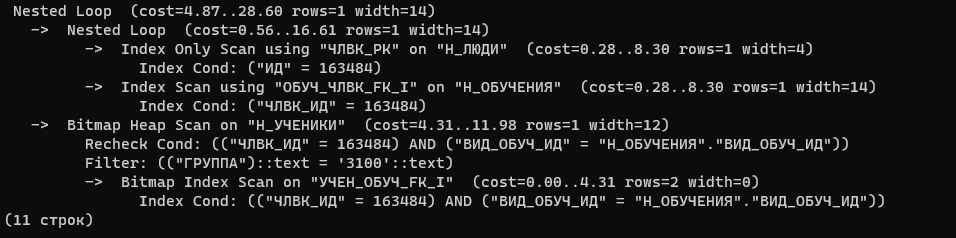
AND Н\_ОБУЧЕНИЯ.ЧЛВК\_ИД = 163484

AND Н\_УЧЕНИКИ.ГРУППА = '3100';

Размер промежуточных данных в втором случае будет меньше, потому что выборка происходит на более раннем этапе и фильтрует большое количество данных, также оба плана используют конвейерную обработку (левосторонние планы)

# EXPLAIN ANALYZE



CREATE INDEX ON "Н\_ЛЮДИ" USING HASH("ИД");

CREATE INDEX ON "Н\_ОБУЧЕНИЯ" USING HASH("ЧЛВК\_ИД");

CREATE INDEX ON "Н\_УЧЕНИКИ" USING HASH("ЧЛВК\_ИД");

CREATE INDEX ON "Н\_УЧЕНИКИ" USING HASH("ГРУППА");

Вывод

В ходе работы я узнал, что такое индексы и как их лучше применять.

Рассмотрел различные планы выполнения запроса и сам проанализировал работу индексов.