­­­НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО

Факультет Нейротехнологии и программирование

Дисциплина

Лабораторная работа № 4

Выполнил студент

Доровский Егор Сергеевич

Группа № Р3123

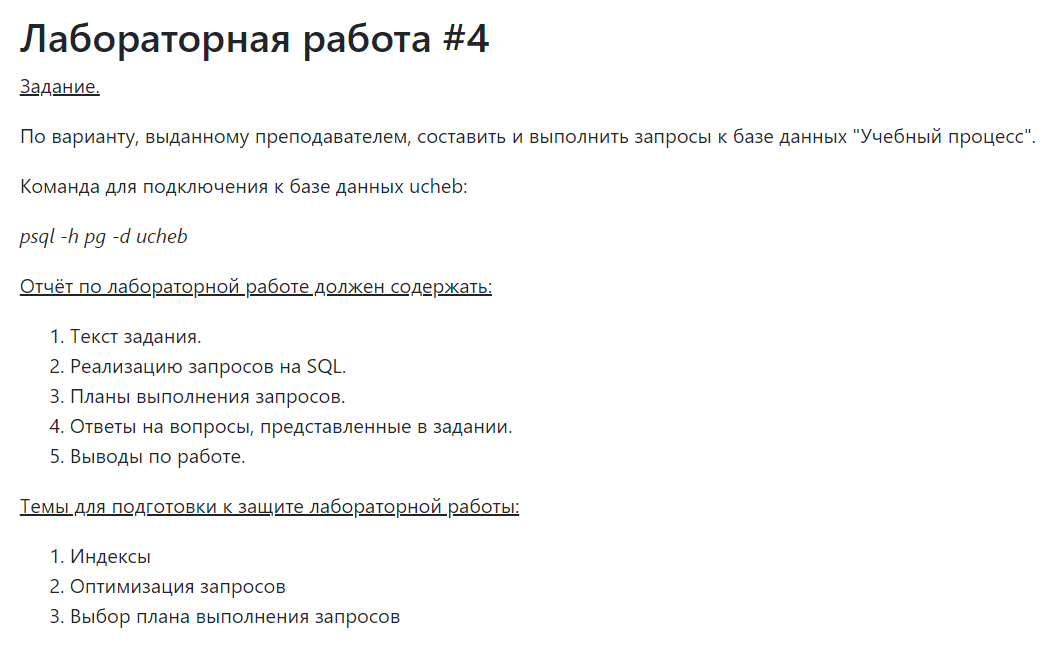
Преподаватель: Горбунов Михаил Витальевич

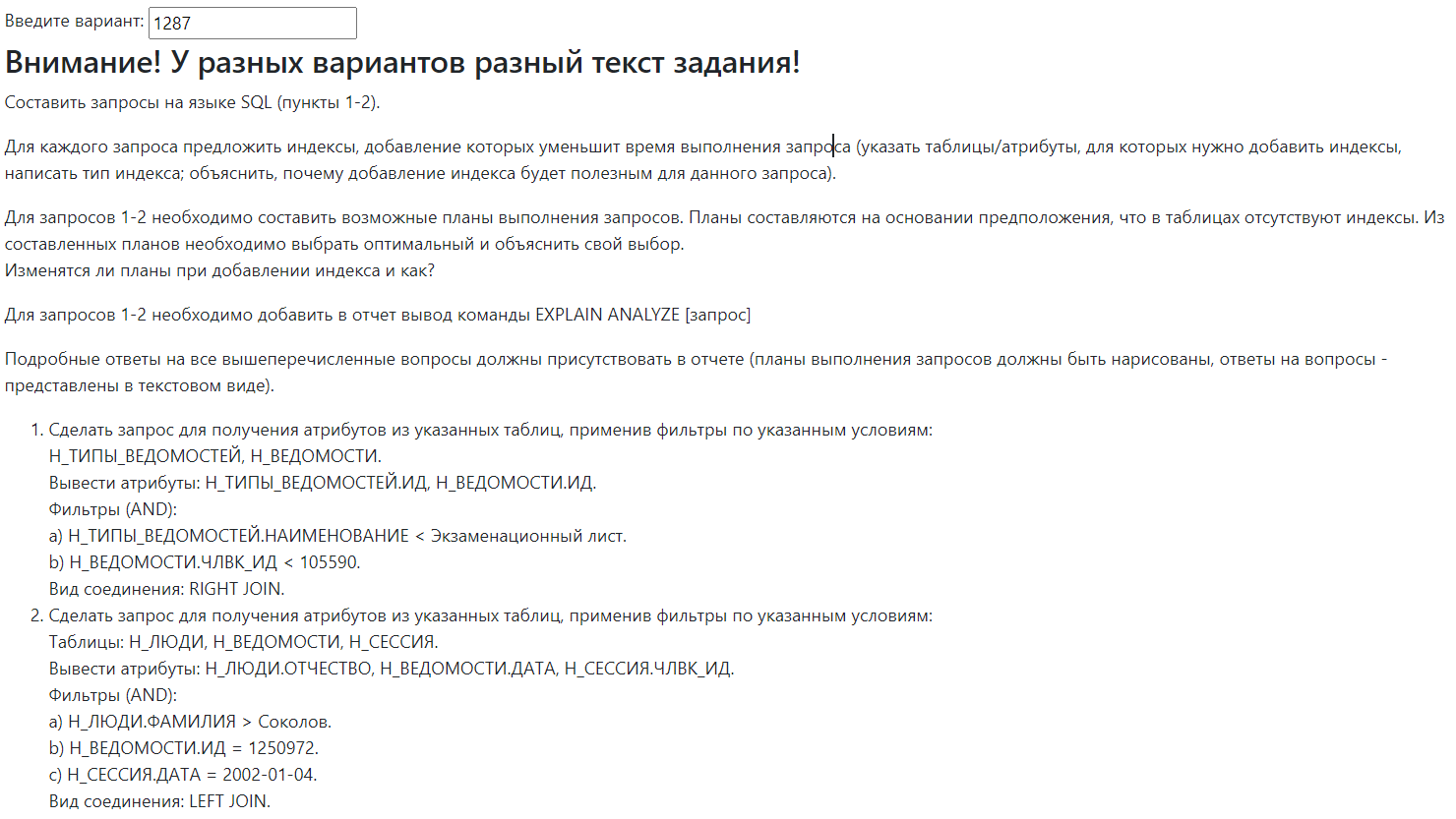
г. Санкт-Петербург

2024

**Вариант:** 1287

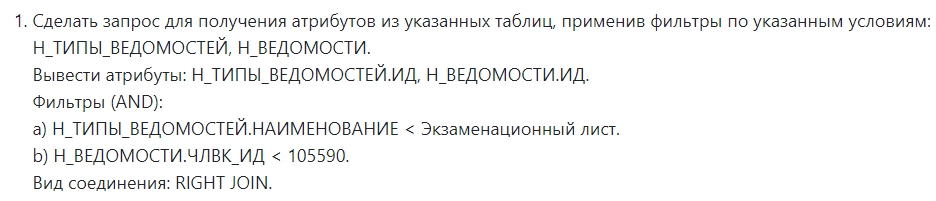
Задание:





Отчет

Первый запрос

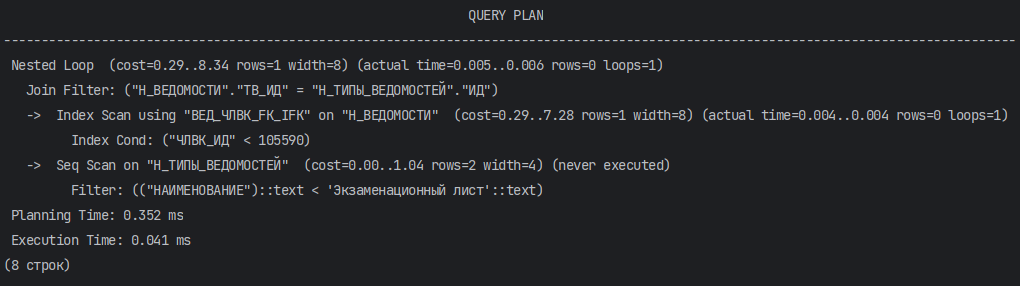


SELECT Н\_ТИПЫ\_ВЕДОМОСТЕЙ.ИД, Н\_ВЕДОМОСТИ.ИД

FROM Н\_ВЕДОМОСТИ

RIGHT JOIN Н\_ТИПЫ\_ВЕДОМОСТЕЙ ON Н\_ВЕДОМОСТИ.ТВ\_ИД = Н\_ТИПЫ\_ВЕДОМОСТЕЙ.ИД

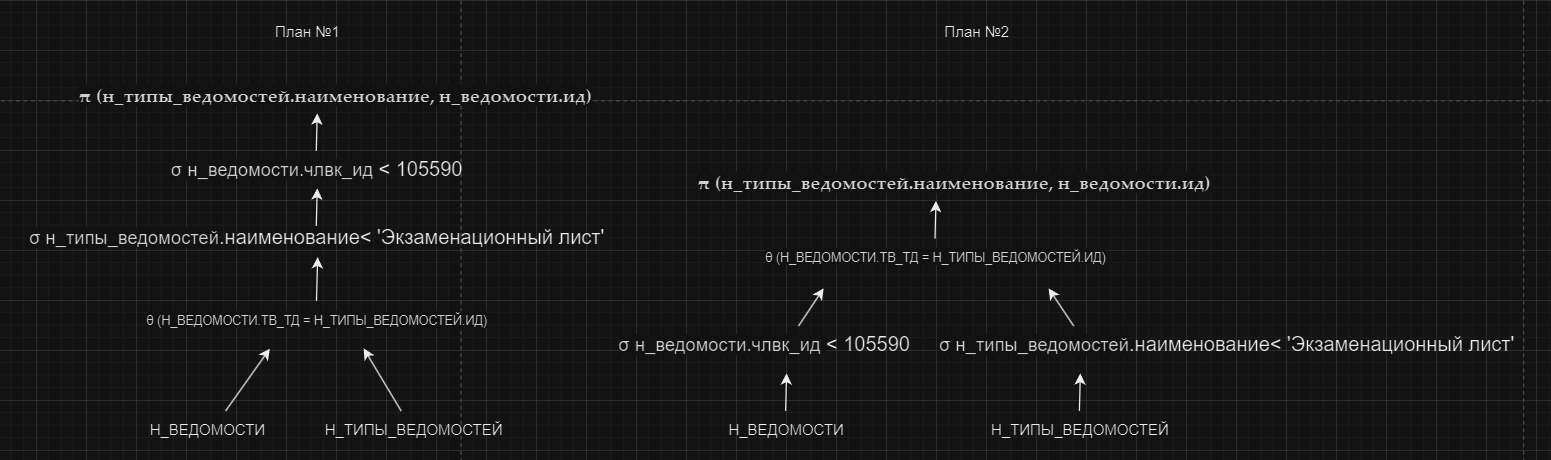
WHERE Н\_ТИПЫ\_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ < 'Экзаменационный лист' AND Н\_ВЕДОМОСТИ.ЧЛВК\_ИД<105590;



CREATE INDEX ON Н\_ВЕДОМОСТИ USING HASH (ТВ\_ИД);

CREATE INDEX ON Н\_ТИПЫ\_ВЕДОМОСТЕЙ USING HASH (ИД);

Планы запроса

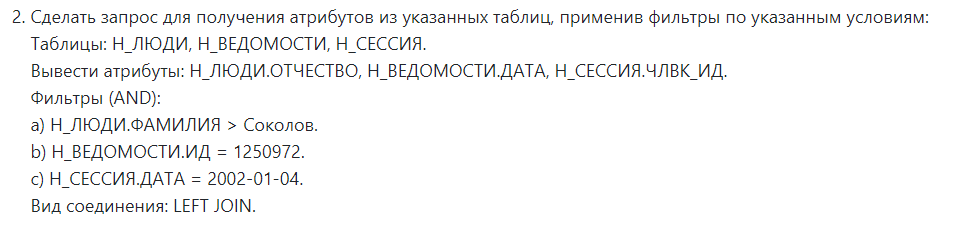


План №1. Производится полный скан таблиц Н\_ВЕДОМОСТИ и Н\_ТИПЫ\_ВЕДОМОСТЕЙ. Затем происходит правое их соединение (RIGHT JOIN) на основе соответствия Н\_ВЕДОМОСТИ.ТВ\_ИД = Н\_ТИПЫ\_ВЕДОМОСТЕЙ.ИД. Потом два раза фильтруем строки: Н\_ТИПЫ\_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ < ‘Экзаменационный лист’ и Н\_ВЕДОМОСТИ.ЧЛВК\_ИД < 105590.

План №2. Производится полный скан таблиц Н\_ВЕДОМОСТИ и Н\_ТИПЫ\_ВЕДОМОСТЕЙ. Строки в каждой из них фильтруются по условиям Н\_ВЕДОМОСТИ.ЧЛВК\_ИД < 105590 и Н\_ТИПЫ\_ВЕДОМОСТЕЙ.НАИМЕНОВАНИЕ < ‘Экзаменационный лист’ соответственно. Потом полученные строки объединяются с помощью RIGHT JOIN.

Сравнение. Второй план более оптимален, потому что в нем мы фильтруем строки, прежде чем соединять таблицы. Чем больше объем данных, тем больше будет заметна разница в производительности.

Второй запрос

****

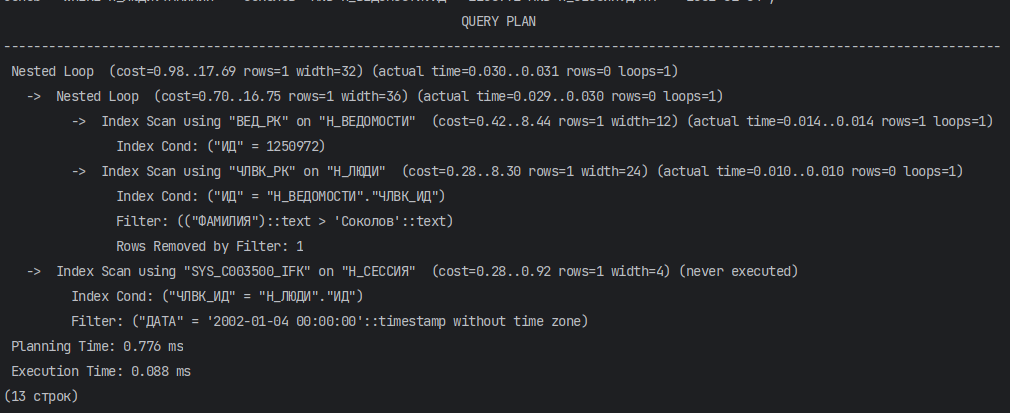
SELECT Н\_ЛЮДИ.ОТЧЕСТВО, Н\_ВЕДОМОСТИ.ДАТА, Н\_СЕССИЯ.ЧЛВК\_ИД

FROM Н\_ЛЮДИ

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

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

WHERE Н\_ЛЮДИ.ФАМИЛИЯ > 'Соколов' AND Н\_ВЕДОМОСТИ.ИД = 1250972 AND Н\_СЕССИЯ.ДАТА = '2002-01-04';

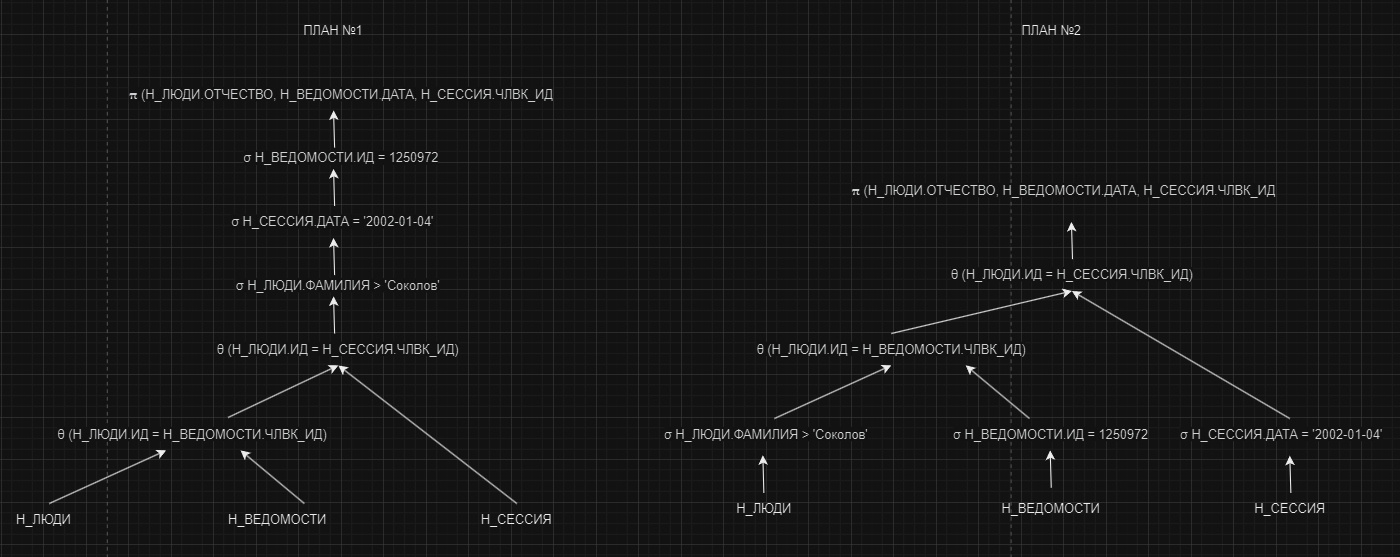


CREATE INDEX ON Н\_ЛЮДИ USING BTREE(ФАМИЛИЯ);

CREATE INDEX ON Н\_ВЕДОМОСТИ USING HASH (ИД);

CREATE INDEX ON Н\_СЕССИЯ USING HASH (ДАТА);

Планы запроса



План №1. Полностью сканируем таблицы Н\_ЛЮДИ и Н\_ВЕДОМОСТИ. Объединяем их с условием Н\_ЛЮДИ.ИД = Н\_ВЕДОМОСТИ.ЧЛВК\_ИД. Сканируем таблицу Н\_СЕССИЯ. Присоединяем её с условием Н\_ЛЮДИ.ИД = Н\_СЕССИЯ.ЧЛВК\_ИД. Применяем фильтры Н\_ЛЮДИ.ФАМИЛИЯ > 'Соколов' AND Н\_ВЕДОМОСТИ.ИД = 1250972 AND Н\_СЕССИЯ.ДАТА = '2002-01-04'.

План №2. Последовательно сканируем таблицу Н\_ЛЮДИ. Фильтруем строки в ней: Н\_ЛЮДИ.ФАМИЛИЯ > ‘Соколов’. Последовательно сканируем Н\_ВЕДОМОСТИ. Фильтруем строки в ней: Н\_ВЕДОМОСТИ.ИД = 1250972. Последовательно сканируем Н\_СЕССИЯ. Фильтруем строки в ней: Н\_СЕССИЯ.ДАТА = '2002-01-04'. Объединяем Н\_ЛЮДИ и Н\_ВЕДОМОСТИ: Н\_ЛЮДИ.ИД = Н\_ВЕДОМОСТИ.ЧЛВК\_ИД. Присоединяем это с Н\_СЕССИЯ с условием Н\_ЛЮДИ.ИД = Н\_СЕССИЯ.ЧЛВК\_ИД.

Сравнение. План №2 определенно быстрее, потому что в нем мы фильтруем строки, прежде чем соединять таблицы. Чем больше объем данных, тем больше будет заметна разница в производительности.

Вывод

Я научился технике эксплейн-анализа, разработке эффективных индексов. Понял, в чем разница фильтрации в WHERE и в JOIN. Научился строить планы запросов и выбирать эффективные среди них.

Список литературы

[Инфологическая модель БД "Учебный Процесс"](https://se.ifmo.ru/documents/10180/733702/%D0%91%D0%94+%D0%A3%D1%87%D0%B5%D0%B1%D0%BD%D1%8B%D0%B9+%D0%9F%D1%80%D0%BE%D1%86%D0%B5%D1%81%D1%81.pdf/2eae3fcd-ea34-4496-924b-6ee4e889a9e5)

[Документация к СУБД PostgreSQL](https://www.postgresql.org/docs/9.6/static/)

[Лекция 1](https://se.ifmo.ru/documents/10180/733702/isbd-2021-1.5.pdf/72944648-dbe7-ced4-38ac-35efe9f87e6f), [Лекция 2](https://se.ifmo.ru/documents/10180/733702/isbd-2021-2.6.pdf/e47a2e01-d445-e017-070f-a300ecdb71a8), [Лекция 3](https://se.ifmo.ru/documents/10180/733702/isbd-2021-3.3.pdf/4d439555-a3cb-883c-705a-bab0fa37687b), [Лекция 4](https://se.ifmo.ru/documents/10180/733702/isbd-2021-4.3.pdf/67de83a2-571f-527c-1ec7-7c9853c86576), [Лекция 5](https://se.ifmo.ru/documents/10180/733702/isbd-2021-5.2.pdf/beadf2a7-db1d-d0c9-c0aa-6abf923c4473), [Лекция 6](https://se.ifmo.ru/documents/10180/733702/isbd-2021-6.2.pdf/e31693c5-499d-3d2d-067c-3e6462d20d57), [Лекция 7](https://se.ifmo.ru/documents/10180/733702/isbd-2021-7.2.pdf/897de389-c846-3483-cd8e-98348c71547b)