МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное бюджетное образовательное учреждение   
высшего образования

«УЛЬЯНОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Факультет «Информационных систем и технологий»

Кафедра «Вычислительная техника»

Дисциплина «Экспериментальные исследования в проектировании автоматизированных систем»

**КУРСОВОЙ ПРОЕКТ (РАБОТА)**

Тема «Экспериментальные исследования реализации учебных планов преподавателей вуза»

Выполнил студент \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ /П.С. Кондратьев/

подпись инициалы, фамилия

Курс 2 Группа ИВТАСмд-21

Направление/ специальность 09.04.01 «Информатика и вычислительная техника»

Руководитель Негода Виктор Николаевич

Дата сдачи:

«\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_20\_\_\_\_г.

Дата защиты:

«\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_20\_\_\_\_г.

Оценка: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Ульяновск, 2022 г.

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное бюджетное образовательное учреждение   
высшего образования

«УЛЬЯНОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Факультет «Информационных систем и технологий»

Кафедра «Вычислительная техника»

Дисциплина «Экспериментальные исследования в проектировании автоматизированных систем»

**ЗАДАНИЕ НА КУРСОВОЙ ПРОЕКТ (РАБОТУ)**

студенту ИВТАСмд-21 . Кондратьеву Павлу Сергеевичу у

Тема проекта (работы) «Экспериментальные исследования реализации учебных планов преподавателей вуза»

Срок сдачи законченного проекта (работы) «\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_20\_\_\_\_г.

Содержание пояснительной записки (перечень подлежащих разработке вопросов)

Введение, теоретическая часть, решение задачи, заключение, список литературы

Руководитель профессор кафедры ВТ \_\_\_\_\_\_\_\_\_\_\_\_\_ /В. Н. Негода/

подпись

«\_\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_20\_\_г

Студент \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ **/**П. С. Кондратьев/

подпись

«\_\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_20\_\_г.

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное бюджетное образовательное учреждение   
высшего образования

«УЛЬЯНОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

**ОТЗЫВ  
руководителя на курсовой проект (работу)**

студента Кондратьева Павла Сергеевича

Факультет ИСТ группа ИВТАСмд-21 курс 2 2

Дисциплина «Экспериментальные исследования в проектировании автоматизированных систем» .

Тема проекта (работы) «Экспериментальные исследования реализации учебных планов преподавателей вуза»

Отзыв Курсовая работа выполнена в полном объеме

Руководитель профессор кафедры «ВТ» \_\_\_\_\_\_\_\_\_\_\_\_\_ /В. Н. Негода/

подпись инициалы, фамилия

«\_\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_20\_\_г.

**Оглавление**

[Введение 5](#_Toc93860232)

[1. Цели экспериментирования, критериальные параметры и факторы, замыслы экспериментов 7](#_Toc93860233)

[1.1 Формулирование целей экспериментов 7](#_Toc93860234)

[1.2 Формулирование гипотез 8](#_Toc93860235)

[2. Разработка планов экспериментов 10](#_Toc93860236)

[2.1 Эксперимент 1: исследование зависимости TPS от рабочей нагрузки 10](#_Toc93860237)

[2.2 Эксперимент 2: исследование зависимости TPS от количества строк в базе данных 11](#_Toc93860238)

[2.3 Эксперимент 3: исследование зависимости TPS от наличия индекса 11](#_Toc93860239)

[3. Разработка средств поддержки проведения экспериментов 11](#_Toc93860240)

[3.1 Настройка тестового стенда 12](#_Toc93860241)

[4. Проведение экспериментов и анализ их результатов 13](#_Toc93860242)

[Заключение 18](#_Toc93860243)

[Список используемых источников 19](#_Toc93860244)

# Введение

Современный инженер имеет дело со сложными техническими системами. В рамках одного методологического подхода [1], к решению таких систем, могут существовать ряд разных теорий, взглядов, положений, имеющих одну и ту же концептуальную основу исследования.

Для проведения научного исследования необходимо методологией, исследовательскими приемами и методами. При проведении исследований на предприятиях перед исследователями ставятся цели, требующие не только анализа и выявления проблем, но и обоснования рекомендаций, предложенных для их разрешения. Оценка состояния, диагностика, профилактика негативных тенденций, поиск «узких мест» в системе управления, определение новых направлений деятельности требуют от менеджера системного видения, владения научно-категориальным аппаратом исследования. Для достижения цели исследования необходимо четко определить основные понятия: объект, предмет, новизну, практическую значимость, методы исследования; знать технологии управления, уметь определять проблему в исследовании процессов и систем управления, осуществлять системный анализ факторов прямого и косвенного воздействия, понимать эффективность, ограничения и условия использования различных методов.

Значительное место в организации исследования на предприятиях занимает работа с информацией, ее анализ, идентификация, классификация и обработка данных. Банк данных предприятий включает большую группу документов, которые могут являться информационной основой исследования. Это устав, баланс и другие документы финансовой отчетности, структура, приказы, отчеты функциональных служб, планы, протоколы, нормативные документы, рекламации и т.д. Получение данных экспертной оценки [2] и социологические опросы существенно расширят проблемное поле исследования, позволят восполнить недостаток информации, полученной из других источников.

При изучении закономерностей функционирования сложных систем приходится прибегать к различным способам познания. Основным методом эмпирического познания стал эксперимент [3], т.е. совокупность операций и в случае необходимости воздействий на изучаемый объект, выполняемых для получения информации о нем на основе результатов опытов.

# 1. Цели экспериментирования, критериальные параметры и факторы, замыслы экспериментов

## 1.1 Формулирование целей экспериментов

Для определения целей экспериментального анализа необходимо более подробно описать объект автоматизации. Данная курсовая работа выполнена в рамках работы над платформой для функциональной реализации учебных планов преподавателей вуза [4].

Система “Учебных планов” [5] не должна обеспечивать автоматизацию абсолютно всех задач, использующих на учебные планы. Однако эти системы могут использовать систему “Учебные планы” как источник достоверных данных для своих нужд. Содержание ученых планов является структурой образующей основой для многих процессов управлением обучением.

Т. о. прежде всего к основным задачам системы относится реализация возможности сбора и хранения всей необходимой информации по учебным планам.

Кроме всего прочего в системе должно поддерживаться хранение и редактирование дополнительной справочной информации, например, такой как перечень кафедр, факультетов, дисциплин, необходимой для корректной работы с учебными планами.

Учебный план можно описать в виде:

*A=<department, plans, discipline, teacher>, где*

department – множество кафедр

plans – множество планов

discipline – множество дисциплин

teacher – множество преподавателей

В свою очередь splans (множество учебных планов) описывается:

*splans = {plan1, plan2, … plann}, где*

plani = {sdepartment.disci | sdepartment.disci ∈ sdiscipline}

*sdiscipline = {disci | i ∈ N*

В рамках данной курсовой работы будет проводиться эксперимент над функцией выбора всех дисциплин, которые ведет кафедра. Эта функция имеет вид:

*fSelectDisc: splans x sdepartment => sdisc\*, где*

sdisc\* ∈ sdiscipline

Таким образом, не взирая на степень вовлечения содержания учебного плана в реализацию тех или иных функций организации учебного процесса, все они обслуживают так или иначе реализацию совокупности учебных планов.

## 1.2 Формулирование гипотез

Перед проведением тестирования сформулируем гипотезы [6], которые и будут проверяться при проведении экспериментов.

**Гипотеза 1:** Очевидно, что при росте количества одновременных соединений и количество потоков, используемых нагрузочной утилитой, TPS [7] будет линейно расти до определенного порога. Чтобы преодолеть данных порог можно использовать индексы.

**Гипотеза 2:** Метрика TPS имеет обратную зависимость от количества строк в таблице.

**Гипотеза 3:** По сколько в запросе выполняет join для двух таблиц, то можем сделать индексы на поля, к которым делается join. Это существенно ускорит select запрос.

**Гипотеза 4:** Добавление индексов к полям join, замедлит обновление и вставку строк в таблицу, поскольку каждый раз нужно будет перестаивать этот самый индекс.

Сейчас в качестве СУБД в проекте используется PostgresSQL [8]. В таком случае может быть несколько вариантов для оптимизации. Например, кеширование результатов некоторых запросов, добавление дополнительных индексов, нормализация таблиц, введение репликации и т. д. Но вероятность того, что проблема производительности будет именно в базе данных очень мала, так как в данный момент в прототипе совершается довольно мало транзакций и не выполняется никаких тяжелых запросов.

Так же есть вероятность, что мы упремся в оперативную память. Вероятность этого так же мала, но, если это произойдет можно либо нарастить память на сервере, либо не хранить данные в оперативной памяти, а вынести их в БД или другое хранилище.

Возможно проблема с производительностью будет связана с ограничениями пропускного канала сети. Если такая проблема действительно возникнет, то с ней нужно будет разбираться отдельно, так как на ширину пропускного канала мы повлиять не сможем. Но такая проблема так же скорее все не возникнет, так как сейчас в проекте происходит довольно мало сетевых взаимодействий.

Наиболее вероятно, что узким местом станет процессор, так как именно на него приходится большая часть нагрузки при обработке запросов. В таком случае можно масштабировать систему вертикально или горизонтально. При вертикальном масштабировании системы мы просто поставим на сервере процессор с большим количеством ядер. Но при вертикальном масштабировании мы рано или поздно упремся в потолок по производительности, так как очевидно нельзя до бесконечности наращивать мощности одного сервера. Тогда будем масштабироваться горизонтально и распределим нагрузку между несколькими серверами.

# 2. Разработка планов экспериментов

## 2.1 Эксперимент 1: исследование зависимости TPS от рабочей нагрузки

1. В данном эксперименте проверяются первую гипотезу, а именно: метрика TPS будет линейно расти до определенного порога, при возрастании нагрузки.
2. Факторами в данном эксперименте будут количество одновременных соединений и количество потоков, проверяемый параметр TPS – количество транзакций в секунду.
3. Теперь выберем запрос, который будет проверяться на нагрузке. Реально один из наиболее популярных запросов на кафедре является выборка дисциплин, которые ведутся на кафедре.
4. Распишем требования к проведению эксперимента:

* Необходимо запускать один и тот же скрипт несколько раз, так как при последующих запусках может отработать кэш, что может заметно повлиять на результаты тестирования;
* Для большей точности измерения TPS каждый нагрузочный тест должен продолжаться не менее 30 секунд;
* Для проведения экспериментов необходимо развернуть тестовое окружение. Тестировать производительность на локальной машине нельзя, так как большую часть ресурсов процессора и оперативной памяти будут занимать сторонние процессы, соответственно результат нагрузочного тестирования будет не репрезентативным. Для проведения стрессового тестирования имеет смысл развернуть отдельный сервер, на котором будет работать только тестируемое приложение. Проще всего для этих целей будет развернуть облачный сервер.

1. Напишем формулу для вычисления целевого параметра TPS. Из названия метрики очевидно, что она рассчитывается по формуле

TPS =

1. Теперь выразим формулой наши предположения относительно результатов эксперимента.

RPS = , где

1. То есть по нашим предположениям TPS линейно зависит от количества потоков и соединений, обрабатываемых в одном потоке.

## 2.2 Эксперимент 2: исследование зависимости TPS от количества строк в базе данных

В данном эксперименте проверяются вторую гипотезу, а именно: метрика TPS имеет обратную зависимость от количества строк в таблице.

Факторами в данном эксперименте будут количество строк в таблицах “Кафедры” и “ПланыСтроки”. Проверяемый параметр TPS – количество транзакций в секунду.

Запрос, который будет проверяться на нагрузке – выборка дисциплин, которые ведутся на кафедре.

1. Теперь выразим формулой наши предположения относительно результатов эксперимента.
2. TPS =
3. То есть по нашим предположениям TPS линейно зависит от количества ядер процессора.

## 2.3 Эксперимент 3: исследование зависимости TPS от наличия индекса

1. В данном эксперименте проверяются третью и четверную гипотезу, а именно: ускорение выборки и замедление вставки при наличии индекса.
2. Факторами в данном эксперименте будет индекс.
3. Запрос, который будет проверяться на нагрузке – выборка дисциплин, которые ведутся на кафедре.
4. Проверяемый параметр TPS – количество транзакций в секунду.

# 3. Разработка средств поддержки проведения экспериментов

## 3.1 Настройка тестового стенда

1. Для проведения экспериментов была развернута тестовая бд, в среде Postgres, с инструментом стресс-теста – pgbench [9]. Таблицы в базе данных были заполнены данными из файлов plx, содержимое которого состоит из спецификаций учебных планов.
2. Как тестирующая утилита была взята pgbench – это простая программа для запуска тестов производительности Postgres. Она многократно выполняет одну последовательность команд, возможно в параллельных сеансах базы данных, а затем вычисляет среднюю скорость транзакций (число транзакций в секунду). По умолчанию pgbench тестирует сценарий, примерно соответствующий TPC-B, который состоит из пяти команд SELECT, UPDATE и INSERT в одной транзакции. Однако можно легко протестировать и другие сценарии, написав собственные скрипты транзакций.
3. В экспериментах использовался компьютер на базе процессора I5 – 5700U (2,5 ГГц, 2 ядра, 4 потока), имеющий оперативную память 8 GB.

# 4. Проведение экспериментов и анализ их результатов

1. Теперь можно приступить непосредственно к тестированию. Команда для запуска такого теста будет выглядеть следующим образом:
2. *sudo -u postgres pgbench -f script.sql postgres*
3. В файле script.sql храниться наш запрос, а postgres – это название бд с которой мы работает в данный момент.
4. Будем запускать один и тот же скрипт несколько раз, так как при последующих запусках может отработать кэш, что заметно повлияет на результаты тестирования.
5. Перейдем к **перовому** экспериментальному тестированию.

*sudo -u postgres pgbench -c -j -f script.sql postgres , где*

-с – количество смоделированных клиентов, то есть количество параллельных сессий базы данных. По умолчанию 1.

-j – количество рабочих нитей в pgbench. Использование более чем одного потока может быть полезно на многопроцессорных машинах. Клиенты распределены между доступными потоками как можно более равномерно. По умолчанию 1.

После проведения первого эксперимента зафиксируем количество соединений (50) и потоков (4).

*sudo -u postgres pgbench -P 10 -T 30 -c 50 -j 4 -f script.sql postgres, где*

-P – Выводить отчёт о прогрессе через заданное число секунд (сек). Выдаваемый отчёт включает время, прошедшее с момента запуска, скорость (в TPS) с момента предыдущего отчёта, а также среднее время ожидания транзакций и стандартное отклонение. В режиме ограничения скорости (-R) время ожидания вычисляется относительно назначенного времени запуска транзакции, а не фактического времени её начала, так что оно включает и среднее время отставания от графика.

-T – Выполнять тест с ограничением по времени (в секундах), а не по числу транзакций для каждого клиента. Параметры -t и -T являются взаимоисключающими.

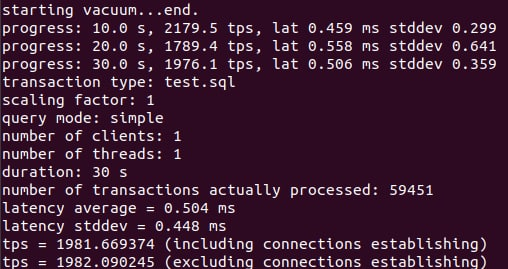


Рис. 1. Первый эксперимент, зависимость TPS от рабочей нагрузки

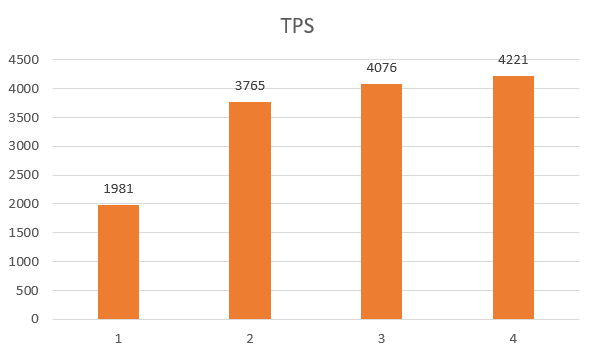


Рис. 2. Первый эксперимент, зависимость TPS от рабочей нагрузки

Из графика мы видим, что первые два столбца имеют линейную зависимость от рабочей нагрузки. При 3-х и более соединений, мы видим, что возрастания TPS не происходит, это говорит нам о том, что мы достигли порогового значения.

1. Перейдем ко **второму** экспериментальному тестированию.

*sudo -u postgres pgbench -P 10 -T 30 -c 50 -j 4 -f test.sql postgres*

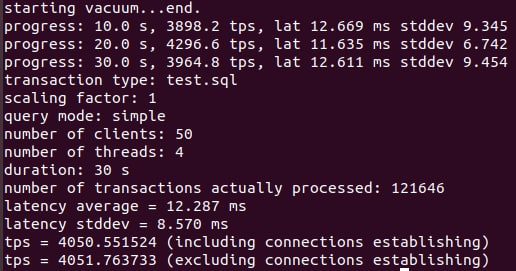


Рис. 3. Второй эксперимент

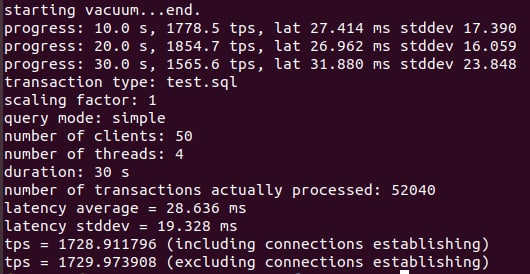


Рис. 4. Второй эксперимент, удвоили кол-во данных в таблицах

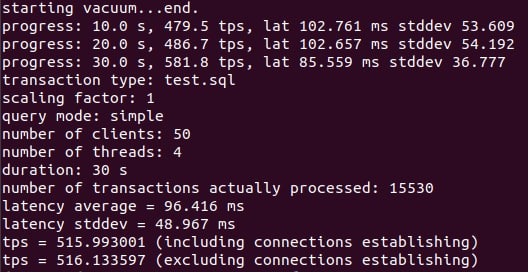


Рис.5. Второй эксперимент, учетверили кол-во данных

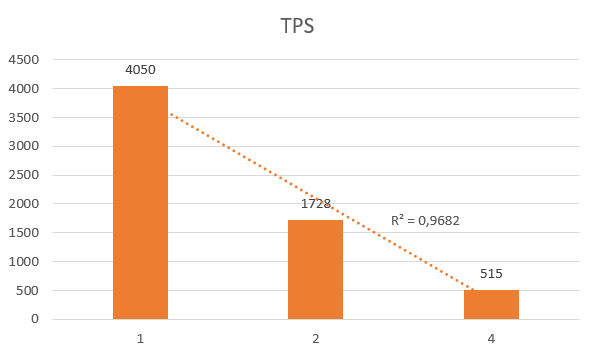


Рис. 6. Второй эксперимент – исследование зависимости TPS от количества строк в базе данных

Из гистограммы видно, что с увеличением объема данных линейно падает TPS. Построив линию тренда, мы убедились то, что стремиться к 1, а это значит, что зависимость действительно линейна.

1. Перейдем ко **третьему** экспериментальному тестированию.
2. *sudo -u postgres pgbench -P 10 -T 30 -c 50 -j 4 -f test.sql postgres*

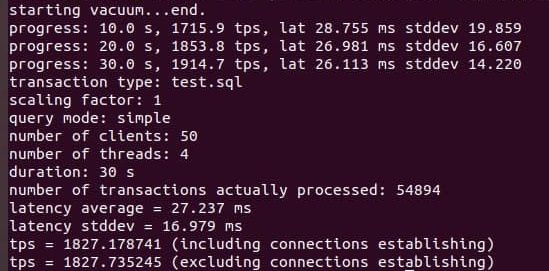


Рис. 7. Третий эксперимент, добавили индексы

При добавлении индексов на поля, которые используется в тестовом запросе, мы видим значительное увеличение значения TPS. Проверить то, что используются индексы, можно при использовании explain, который показывает план запроса.

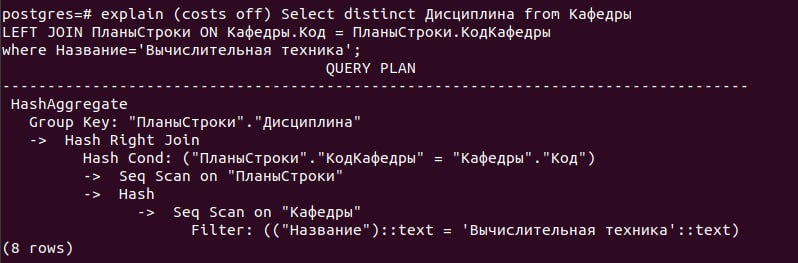


Рис.8. План запроса без индексов

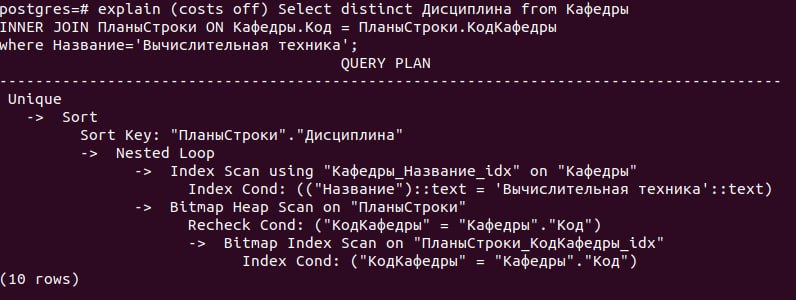


Рис.8. План запроса с индексами

# Заключение

1. В ходе выполнения данной курсовой работы были сформулированы и проверены гипотезы о производительности платформы и влиянии различных факторов на порог максимальной нагрузки, которую может выдержать прототип платформы по функциональной реализации учебных планов преподавателей вуза.
2. Были проведены эксперименты по определению производительности платформы и способах ее масштабирования с помощью нагрузочной утилиты pgbench. А так же построены графики зависимости производительности платформы от различных факторов.

# Список используемых источников

1. Основной методологический способ – это подход. URL: <https://molotokrus.ru/osnovnoy-metodologicheskiy-sposob-eto-podhod/>
2. Экспертное оценивание. URL: <https://ru.wikipedia.org/wiki/Экспертное_оценивание>
3. Эксперимент как важнейший способ эмпирического познания. URL: <https://studref.com/483556/filosofiya/eksperiment_vazhneyshiy_sposob_empiricheskogo_poznaniya>
4. Учебный план в формате .plx
5. Положение об индивидуальных планах работы преподавателей. URL: <http://www.chgpu.edu.ru/uploads/files/1556256339_pol.-ob-ind.plane-raboty-prepodavatelya-na-sayt.pdf>
6. Экспериментальная гипотеза. URL: <https://studme.org/144040/psihologiya/eksperimentalnaya_gipoteza>
7. Пропускная способность системы, TPS (QPS), параллелизм пользователей, концепции и формулы тестирования производительности. URL: <https://russianblogs.com/article/9416794810/>
8. PostgreSQL: The World's Most Advanced Open Source Relational Database. URL: <https://www.postgresql.org/>
9. Pgbench. URL: <https://runebook.dev/ru/docs/postgresql/pgbench>