

Отчет об исследовании и оптимизации производительности СУБД в виртуализированной среде с использованием средств нагрузочного тестирования HammerDB

1. Введение и контекст

Данный отчет документирует практическую работу, направленную на исследование и оптимизацию производительности Системы Управления Базами Данных (СУБД) PostgreSQL в среде виртуальных машин (ВМ), управляемых VMware Workstation. Основная цель – понять влияние конфигурации ресурсов, выделяемых ВМ (оперативная память и виртуальные процессоры), на производительность СУБД под нагрузкой, используя инструмент нагрузочного тестирования HammerDB.

2. Цели работы

В соответствии с методическими указаниями к заданию, были поставлены следующие цели:

- Освоение средств виртуализации: Получение практического опыта установки, настройки и управления ВМ с использованием VMware Workstation.
- Работа с СУБД: Приобретение навыков развертывания и базовой конфигурации PostgreSQL на различных операционных системах.
- Навыки нагрузочного тестирования: Освоение инструмента HammerDB для генерации нагрузки, сбора и анализа метрик производительности СУБД.
- Идентификация оптимальной конфигурации: Определение идеальной конфигурации ресурсов ВМ для СУБД под заданной нагрузкой.

3. Подготовка тестовой среды

3.1. Описание аппаратного и программного обеспечения хост-среды

- Гипервизор: VMware Workstation.
- Инструмент нагрузочного тестирования: HammerDB версии 4.8, запущенный с другой машины (машины товарища).

3.2. Описание виртуальных машин (ВМ)

Были созданы и настроены две виртуальные машины в VMware Workstation:

1. ВМ Windows Server:
 - a. Операционная система: Windows Server.
 - b. Оперативная память (RAM): 4 ГБ.
 - c. Виртуальные процессоры (vCPU): 2.

- d. Выделенное дисковое пространство: 50 ГБ.
 - e. Сетевой режим: NAT.
2. VM Linux (CentOS 8):
- a. Операционная система: CentOS 8 64-бит.
 - b. Оперативная память (RAM): Варьировалась для тестов оптимизации (протестированные конфигурации: 2 ГБ, 3 ГБ, 4 ГБ).
 - c. Виртуальные процессоры (vCPU): Варьировались для тестов оптимизации (протестированные конфигурации: 2, 4).
 - d. Выделенное дисковое пространство: 50 ГБ.
 - e. Сетевой режим: Bridged (Мост).

3.3. Развертывание и настройка СУБД (PostgreSQL)

PostgreSQL версии 15.6 был развернут на обеих VM. Базовая конфигурация включала:

- `postgresql.conf`: `listen_addresses = '*'` для разрешения внешних подключений.
- `pg_hba.conf`: Правила доступа настроены для аутентификации md5 с IP-адреса машины HammerDB.
- Брандмауэр операционной системы: Порт 5432 был открыт.

4. Проведение базовых тестов

В этом разделе представлены результаты первоначальных тестов, проведенных для оценки базовой производительности PostgreSQL на каждой операционной системе со стандартной конфигурацией.

4.1. Методика тестирования

Для всех тестов применялась следующая методика:

- СУБД: PostgreSQL 15.6
- Инструмент тестирования: HammerDB 4.8
- Бенчмарк: TPC-C
- Количество складов (Warehouses): 100 (схема создана через HammerDB)
- Виртуальные пользователи: 3
- Продолжительность теста: 600 секунд (10 минут)

4.2. Результаты базовых тестов

4.2.1. VM Windows Server (4 ГБ ОЗУ, 2 vCPU, 50 ГБ диска)

- Время создания схемы (Build Schema): 563 секунды.
- Результаты теста производительности:

- TPM (Транзакций в минуту): 1152.02
- NOPM (Новых заказов в минуту): 579.52

4.2.2. VM CentOS 8 (3 ГБ ОЗУ, 2 vCPU, 50 ГБ диска)

- Время создания схемы (Build Schema): 501 секунда.
- Результаты теста производительности:
 - TPM: 1335.79
 - NOPM: 671.95

4.3. Первые наблюдения и обоснование выбора VM для оптимизации

Сравнение первоначальных показателей производительности показывает, что VM CentOS 8 превзошла VM Windows Server, демонстрируя более высокие значения TPM и NOPM, а также более короткое время создания схемы, несмотря на меньший объем оперативной памяти (на 1 ГБ).

- TPM: CentOS 8 (1335.79) > Windows Server (1152.02)
- NOPM: CentOS 8 (671.95) > Windows Server (579.52)

Следовательно, VM CentOS 8 была выбрана в качестве платформы для последующих этапов оптимизации, в соответствии с требованием задания выбрать наиболее производительную операционную систему для данного этапа.

5. Исследование производительности СУБД в различных конфигурациях VM и выявление оптимальной конфигурации

В этом разделе подробно описаны эксперименты, проведенные на VM CentOS 8 для определения влияния оперативной памяти и виртуальных процессоров на производительность СУБД PostgreSQL и выявления оптимальной конфигурации.

5.1. Описание проведенных экспериментов

На этом этапе VM CentOS 8 подвергалась серии нагрузочных тестов HammerDB с изменением параметров оперативной памяти и виртуальных процессоров. Выделенное дисковое пространство оставалось постоянным – 50 ГБ.

Протестированные конфигурации на CentOS 8:

- Конфигурация 1 (Базовая): 3 ГБ ОЗУ, 2 vCPU
- Конфигурация 2: 2 ГБ ОЗУ, 2 vCPU
- Конфигурация 3: 4 ГБ ОЗУ, 2 vCPU
- Конфигурация 4: 4 ГБ ОЗУ, 4 vCPU

5.2. Методика тестирования

Параметры HammerDB оставались идентичными тем, что были на этапе 4.1 (3 виртуальных пользователя, 100 складов, 10 минут теста). Для каждой новой конфигурации ВМ, ВМ перезагружалась после корректировки ресурсов в VMware Workstation.

5.3. Результаты тестов (зависимость ТРМ/NOPM от vCPU и RAM)

Ниже представлена сводная таблица полученных результатов для ВМ CentOS 8:

Конфигурация теста (CentOS 8)	ОЗУ (ГБ)	vCPU	ТРМ	NOPM
Базовый тест	3	2	1335.79	671.95
Тест 2	2	2	2829.17	1422.37
Тест 3	4	2	50944.00	22163.00
Тест 4	4	4	52084.75	26194.27

5.4. Анализ результатов: выявление зависимостей и определение оптимальной конфигурации ВМ

Анализ результатов показывает очень сильную зависимость производительности PostgreSQL от объема выделенной оперативной памяти, особенно для 100 складов и 3 виртуальных пользователей.

- Влияние ОЗУ: Производительность значительно возросла при переходе с 3 ГБ на 4 ГБ ОЗУ (с 2 vCPU): ТРМ увеличился с 1335.79 до 50944.00, а NOPM – с 671.95 до 22163.00. Это впечатляющее улучшение предполагает, что при объеме ОЗУ менее 4 ГБ база данных не могла полностью помещаться в память, что приводило к частым операциям ввода-вывода с диска и сильно ограничивало производительность. Переход на 4 ГБ ОЗУ, вероятно, позволил значительно большей части (если не всей) активной рабочей нагрузки PostgreSQL кэшироваться в памяти, резко сократив обращения к диску. Производительность при 2 ГБ ОЗУ была еще ниже, чем при 3 ГБ, что подтверждает эту тенденцию.
- Влияние vCPU: После того как объем ОЗУ стал достаточным (4 ГБ), увеличение vCPU с 2 до 4 привело к дополнительному приросту производительности. ТРМ незначительно вырос с 50944.00 до 52084.75 (+2.2%), в то время как NOPM показал более заметный прирост с 22163.00 до 26194.27 (+18.2%). Этот прирост, хотя и менее значительный, чем

оптимизация ОЗУ, указывает на то, что увеличение параллелизма может быть полезным после устранения узкого места по памяти.

Определение оптимальной конфигурации:

На основе проведенных тестов, наиболее производительной конфигурацией для VM CentOS 8 с HammerDB при 3 виртуальных пользователях и схеме на 100 складов является:

- ОЗУ: 4 ГБ
- vCPU: 4

Эта конфигурация показала наилучшие результаты по TPM и NOPM.

5.5. Выводы по этапу

Этот этап оптимизации продемонстрировал, что ОЗУ является наиболее критичным фактором производительности для PostgreSQL под нагрузкой в данной среде, особенно когда размер базы данных превышает доступную память. Порог в 4 ГБ ОЗУ был определен как необходимый для достижения значительно более высокой производительности. Увеличение vCPU (с 2 до 4) привело к дополнительным приростам, хотя и менее значительным, чем оптимизация ОЗУ, что предполагает полезность параллелизма после оптимизации памяти.

6. Общие выводы

6.1. Краткое резюме полученных результатов по всем этапам

Данная практическая работа позволила:

- Успешно развернуть и настроить виртуальные машины в VMware Workstation, а также СУБД PostgreSQL 15.6 на Windows Server и CentOS 8.
- Использовать HammerDB для проведения нагрузочных тестов TPC-C и сбора ключевых метрик производительности.
- Сравнить первоначальную производительность двух операционных систем, выявив CentOS 8 как наиболее производительную платформу для данного исследования.
- Провести серию тестов оптимизации на CentOS 8, выявив критическую зависимость производительности от ОЗУ и конфигурации vCPU, а также определив оптимальную конфигурацию.

6.2. Ответы на поставленные в начале работы цели и задачи

Все цели, поставленные в начале работы, были достигнуты:

- Компетенции в области виртуализации (VMware Workstation) были приобретены и применены в процессе создания и управления ВМ.
- Развертывание и базовая настройка PostgreSQL на системах Windows и Linux были освоены.
- Инструмент нагрузочного тестирования HammerDB был успешно использован для моделирования нагрузок и анализа метрик производительности.
- Была определена оптимальная конфигурация ВМ (4 ГБ ОЗУ, 4 vCPU на CentOS 8) для СУБД под тестовой нагрузкой.

6.3. Итоговые рекомендации по выбору ОС и конфигурации ВМ для развертывания выбранной СУБД под нагрузкой

На основе результатов данного исследования:

- Выбор операционной системы: Для развертывания PostgreSQL под нагрузкой в виртуализированной среде рекомендуется CentOS 8 по сравнению с Windows Server, поскольку она продемонстрировала превосходную производительность в проведенных тестах.
- Конфигурация ВМ: Для высокой производительности PostgreSQL при аналогичной рабочей нагрузке (3 виртуальных пользователя, 100 складов) оптимальной конфигурацией, выявленной на CentOS 8, является выделение 4 ГБ ОЗУ и 4 vCPU. Крайне важно выделить достаточный объем ОЗУ для эффективной работы базы данных в памяти, что является наиболее определяющим фактором производительности.

7. Список литературы

(Внешние ссылки, кроме документа к заданию и документации к инструментам, не использовались.)