

## PT-16 Tantor: Настройка производительности PostgreSQL 16.

### Описание:

Знания, полученные при прохождении курса, позволяют настраивать работу экземпляра СУБД PostgreSQL любых производителей на операционных системах семейства Linux для достижения высокой производительности. В практических заданиях используется Tantor Postgres SE 16 на Astra Linux 1.8

### Аудитория:

- Администраторы;
- Разработчики;
- Инженеры технической поддержки, в задачи которых входит настройка производительности СУБД семейства PostgreSQL.

### Цель курса:

- Настройка операционных систем Linux;
- Настройка экземпляра, обслуживающего кластер баз данных PostgreSQL;
- Оптимизация структур хранения данных PostgreSQL.

### На курсе слушателями будут исследованы и так же закреплены в лабораторных работах следующие темы:

- Создание и использование тестов производительности, программ тестирования;
- Использование основных ресурсов: памяти, процессоров, дискового и сетевого ввода-вывода;
- Настройка операционной системы Linux для оптимальной работы СУБД PostgreSQL;
- Структуры хранения данных;
- Особенности использования индексов типа btree;
- Настройка работы автовакуума;
- Использование накопительной статистики.

### Предварительная подготовка:

- Знания в объеме курса "Администрирование PostgreSQL"
- Опыт администрирования СУБД PostgreSQL в операционных системах семейства Linux

Глава	Содержание курса
1-1	<b>Обзор</b>
	О компании Тантор
	СУБД Tantor
	Тантор xData
	Тантор PipelineDB
	Платформа Тантор
	О курсе
	Общие понятия настройки производительности
	Методология настройки производительности
	Последовательность действий по настройке производительности
	Пример исключения задержки на стороне сервера приложений
1-2	<b>Использование утилиты pgbench</b>
	Бенчмаркинг
	Результат бенчмаркинга
	pgbench - утилита бенчмаркинга PostgreSQL
	Три встроенных теста pgbench
	Параметры запуска pgbench
	Рекомендации по использованию pgbench
	Пример использования pgbench
1-3	<b>Использование утилит sysbench и fio</b>

	sysbench - утилита для тестирования производительности
	Использование sysbench для тестирования процессоров памяти дисков
	Тестирование аппаратных ресурсов
	Тестирование ввода-вывода утилитой Flexible IO Tester (fio)
1-4	<b>Тесты ТРС</b>
	Тесты ТРС
	Тесты ТРС-В и ТРС-С
	Тест ТРС-Е тестирование устойчивости к сетевым сбоям
	Реализация теста ТРС-С
	Приложение HammerDB
	Параметры для теста тип-С HammerDB
	Утилита Go-TPC
	Практика
2-1	<b>Память</b>
	Оперативная память
	Виртуальная адресация памяти
	Размер страниц памяти
	Размер буфера ассоциативной трансляции (TLB)
	Огромные страницы (Huge Pages)
	Использование Huge Pages
	Использование экземпляром Huge Pages
	Прозрачные огромные страницы (Transparent Huge Pages)
	Неравномерный доступ к памяти (NUMA)
2-2	<b>Нехватка памяти</b>
	Out Of Memory (OOM)
	Размер резидентного набора (RSS)
	Параметр oom_score_adj
	Параметр vm.overcommit_memory
	Установка значений overcommit и swap
	Параметр vm.swappiness
	Дедупликация страниц памяти (KSM)
	Выделение локальной памяти процессами экземпляра
	ОШИБКА: invalid memory alloc request size
	Параметр enable_large_allocations
2-3	<b>Страничный кэш</b>
	Страничный кэш linux
	Доля изменённых ("грязных") страниц в кэше
	Фрагментация памяти
	Дефрагментация памяти
	Длительность удержания грязных страниц в кэше
	Параметр backend_flush_after
	Практика
3-1	<b>Процессоры</b>
	Simultaneous Multi-Threading (SMT) и Hyper-Threading (HT)
	Привязка процесса к процессору (CPU affinity)
	Просмотр списка процессов утилитой ps
	Запись и просмотр метрик утилитой atop
	Переключения контекста выполнения
	Планировщик операционной системы
	Использование процессорного времени (пропорция USER/SYS)
	Источник времени (clock source)
	Сравнение источников времени
	Сравнение источников времени в PostgreSQL
	Замена источника времени
3-2	<b>Сеть</b>
	Основные параметры сети

	Алгоритмы Congestion и slow start
	Алгоритм BBR (Bottleneck Bandwidth и Round Trip Time)
	Параметры сетевых соединений
	Параметры энергосбережения
	Практика
4	<b>Система хранения</b>
	Дисковая подсистема
	HDD SSD NVMe
	Блочные устройства
	Планировщик ввода/вывода (I/O Scheduler)
	Изменение I/O Scheduler
	Физический сектор диска
	Взаимодействие процессов экземпляра с диском
	Синхронизация файлов данных с диском
	Размер блока файловой системы
	Параметр wal_sync_method
	Гарантия записи в WAL
	Быстрые фиксации изменений в журнале файловой системы ext4 (fast_commit)
	Включение быстрой фиксации (fast_commit)
	Утилита pg_test_fsync
	Групповая фиксация транзакций
	Параметры commit_delay и commit_siblings
	Команды шин ввода-вывода discard/trim
	Поддержка discard/trim
	Рекомендации по использованию SSD
	Параметр max_files_per_process
	Увеличение значения max_files_per_process
	Временная файловая система (tmpfs)
	RAID
	LVM
	Практика
5	<b>Начальная настройка СУБД</b>
	Конфигураторы
	Параметры shared_buffers temp_buffers effective_cache_size
	Параметры work_mem hash_mem_multiplier maintenance_work_mem
	Параметр autovacuum_work_mem
	Параметры temp_file_limit и temp_tablespace
	Параметры max_slot_wal_keep_size и transaction_timeout
	Параметры max_connections и client_connection_check_interval
	Параметр max_locks_per_transaction
	Фоновые рабочие процессы
	Параметры max_worker_processes и max_parallel_workers
	Параметр max_parallel_workers_per_gather
	Параметры системы хранения
	Параметры контрольных точек
	Параметры процесса фоновой записи bgwriter
	Практика
6	<b>Структуры хранения</b>
	Таблицы
	Служебные столбцы
	Расширение pageinspect
	Padding и aligning
	Aligning (выравнивание)
	cache line
	Структура блока данных
	Число строк в блоке

	Порядок следования столбцов в таблице
	Порядок следования столбцов и производительность
	Практика
7	<b>Индексы типа btree</b>
	Методы доступа к строкам
	Класс операторов для индекса
	Семейства и классы операторов
	Поддерживаемые функции для индекса
	Индексы для ограничений целостности
	Индекс типа btree
	Функции расширения pageinspect для btree
	Индексы с дедупликацией в листовых блоках
	Проверка поддерживается ли дедупликация
	Параметры создания индекса и их влияние на производительность
	Частичные (partial) индексы
	Эволюция индексов: создание удаление перестройка
	Структура индекса типа btree
	High key в структуре индекса
	Изменение структуры индекса при добавлении строк
	Пример роста индекса при вставке строк
	Структура индекса после его перестроения
	FILLFACTOR в индексах типа btree
	Быстрый путь (fastpath) вставки в индексы
	Внутристраничная очистка в индексах
	Влияние удаления строк на индексы
	Исключение блоков из структуры индекса
	Число исключённых блоков из структуры индекса
	Практика
8-1	<b>TOAST</b>
	TOAST (The Oversized-Attribute Storage Technique)
	Поля переменной длинны
	Вытеснение полей в TOAST
	Алгоритм вытеснения полей в TOAST
	TOAST chunk
	Ограничения TOAST
	Выравнивание строк с вытесненными в TOAST полями
	Параметры toast_tuple_target и default_toast_compression
	Оптимизация Heap Only Tuple
	Мониторинг HOT update
	Влияние FILLFACTOR на HOT cleanup
	Внутристраничная очистка в таблицах
	Внутристраничная очистка в индексах
8-2	<b>Типы данных</b>
	Типы данных наименьшего размера: boolean char smallint
	Типы данных переменной длинны
	Целочисленные типы данных
	Выбор типов данных для первичного ключа
	Параметр cache у последовательностей
	Хранение дат, времени, их интервалов
	Функции проверки типа данных и размера поля
	Типы данных для вещественных чисел
	Параметр конфигурации extra_float_digits
	Хранение вещественных чисел
	Разрядность результата деления numeric
	Практика
9-1	<b>Архитектура</b>

	Запуск экземпляра, процесс postgres
	Процесс startup
	Синхронизация PGDATA, параметр recovery_init_sync_method
	Синхронизация бэкапа, параметр pg_basebackup --sync-method
	Параметр restart_after_crash
	Особенности работы экземпляра в контейнере docker
	Что происходит при запуске серверного процесса
	Общая память процессов экземпляра
	Кэш таблиц системного каталога
	Представление pg_stat_slru
	Локальная память процесса
	Представление pg_backend_memory_contexts
	Функция pg_log_backend_memory_contexts(PID)
9-2	<b>Блокировки</b>
	Типы блокировок
	Параметры deadlock_timeout и log_lock_waits
	Параметр lock_timeout
	Подтранзакции
	Мультитранзакции
	Быстрый путь блокирования (fastpath)
	Сильные и слабые блокировки таблиц
	Справочник устанавливаемых командами блокировок
	Секции таблицы блокировок
	Транши блокировок (tranches)
	Легковесные блокировки
	Блокирование по быстрому пути и 16 блокировок
	Индексы соединения секции и быстрый путь
	Параметр join_collapse_limit
	Представление pg_locks
	Параметр track_commit_timestamp
	Практика
10	<b>Буферный кэш</b>
	Структуры памяти обслуживающие буферный кэш
	Структуры памяти обслуживающие буферный кэш (продолжение)
	Поиск свободного буфера
	Алгоритм вытеснения грязных буферов
	Стратегии замены буферов
	Поиск блока в буферном кэше
	Закрепление буфера (pin) и блокировка content_lock
	Освобождение буферов при удалении файлов
	Оптимизированное расширение файлов
	Изменение размера файлов и буферный кэш
	Предварительное чтение блоков (prefetch)
	Представление pg_stat_recovery_prefetch
	Расширение pg_prewarm
	Процесс фоновой записи bgwriter
	Алгоритм очистки кэша буферов процессом bgwriter
	Представление pg_stat_bgwriter
	Расширение pg_buffers
	Настройка размера кэша буферов
	Параметр synchronize_seqscans
	Практика
11	<b>Контрольная точка</b>
	Контрольная точка
	Шаги выполнения контрольной точки
	Шаги выполнения контрольной точки (продолжение)

	Параметры конфигурации процесса checkpointer
	Статистика для настройки параметров checkpointer
	Пример настройки параметров checkpointer
	Пример настройки параметров checkpointer (продолжение)
	Практика
12	<b>Автовакуум</b>
	Алгоритм вакуумирования
	Первая фаза вакуумирования
	Расчёт памяти под TID для вакуумирования
	Вторая и третья фазы вакуумирования
	Четвертая и пятая фазы вакуумирования
	Агрессивный (Aggressive) режим вакуумирования
	Заморозка строк (FREEZE)
	Вакуум в 17 версии PostgreSQL
	Сравнительное тестирование вакуума 16 и 17 версий PostgreSQL
	Контрольные суммы и WAL
	Параметры команды VACUUM
	Параметры команды VACUUM (продолжение)
	Расширение pg_visibility
	Мониторинг автовакуума
	Представление pg_stat_progress_vacuum
	Параметр log_autovacuum_min_duration
	Параметры конфигурации автовакуума
	Настройка автовакуума
	Параметр autovacuum_naptime
	Выбор таблиц автовакуумом
	Рекомендации по настройке автовакуума
	Важность наблюдения за горизонтом баз данных
	Мониторинг горизонта баз данных
	Параметры автовакуума на уровне таблиц
	Параметр default_statistics_target
	Раздувание (bloat) таблиц и индексов
	Практика
13	<b>Использование диагностического журнала</b>
	Диагностический журнал
	Параметры диагностики
	Отслеживание использования временных файлов
	Отслеживание работы автовакуума и автоанализа
	Наблюдение за контрольными точками
	Описание записей log_checkpoints
	Описание записей log_checkpoints (продолжение)
	Утилита pg_waldump и записи log_checkpoints
	Утилита pg_waldump и записи log_checkpoints
	Диагностика частоты соединений
	Диагностика блокирующих ситуаций
	Практика
14	<b>Накопительная статистика</b>
	Накопительная статистика
	Утилита pg_test_timing
	Просмотр статистики работы процессов
	Представление pg_stat_database
	Прогресс выполнения команд
	Представление pg_stat_io
	Статистики buffers_backend_fsync и fsyncs
	Строки представления pg_stat_io
	Характеристики pg_stat_io

	Статистики представления pg_stat_io
	Представления pg_statio_all_tables и pg_statio_all_indexes
	Представление pg_stat_all_tables
	Представление pg_stat_all_indexes
	Длительность удержания горизонта баз данных
	Представление pg_stat_wal
	Расширение pg_walinspect
	Использование расширения pg_walinspect
	Представление pg_stat_activity
	Блокирующие процессы и функция pg_blocking_pids()
	pg_cancel_backend() и pg_terminate_backend()
	Практика
15	<b>Расширения pg_stat_statements и pg_stat_kcache</b>
	Расширение pg_stat_statements
	Конфигурация pg_stat_statements
	Параметры конфигурации pg_stat_statements
	Представление pg_stat_statements
	Запросы к представлению pg_stat_statements
	Примеры запросов к представлению pg_stat_statements
	Метрики pg_stat_statements
	Примеры метрик представления pg_stat_statements
	Расширение pg_stat_kcache
	Статистики собираемые pg_stat_kcache
	Просмотр статистик pg_stat_kcache
	Практика
16	<b>Расширение pg_wait_sampling</b>
	Расширение pg_wait_sampling
	История событий ожидания
	История событий ожидания (продолжение)
	Параметры расширения pg_wait_sampling
	Профиль pg_wait_sampling
	Запросы к профилю pg_wait_sampling
	Запросы к профилю pg_wait_sampling (продолжение)
	Сброс статистик
	Практика