НИУ ИТМО

Факультет программной инженерии и компьютерных технологий

Отчет по лабораторной работе №2

по дисциплине Распределенные системы хранения данных

|  |  |
| --- | --- |
| Студент группы № P33151 | Шипулин Павел Андреевич |
| Преподаватель | Перцев Тимофей Сергеевич |

Санкт-Петербург

2024

# Задание

## Этап 1. Инициализация кластера БД

* Директория кластера: $HOME/yqi56
* Кодировка: ANSI1251
* Локаль: русская
* Параметры инициализации задать через аргументы команды

## Этап 2. Конфигурация и запуск сервера БД

* Способы подключения: 1) Unix-domain сокет в режиме peer; 2) сокет TCP/IP, только localhost
* Номер порта: 9142
* Способ аутентификации TCP/IP клиентов: по паролю MD5
* Остальные способы подключений запретить.
* Настроить следующие параметры сервера БД:
  + max\_connections
  + shared\_buffers
  + temp\_buffers
  + work\_mem
  + checkpoint\_timeout
  + effective\_cache\_size
  + fsync
  + commit\_delay

Параметры должны быть подобраны в соответствии со сценарием OLAP: 5 одновременных пользователей, пакетная запись/чтение данных по 192МБ.

* Директория WAL файлов: $PGDATA/pg\_wal
* Формат лог-файлов: .log
* Уровень сообщений лога: WARNING
* Дополнительно логировать: попытки подключения и завершение сессий

## Этап 3. Дополнительные табличные пространства и наполнение базы

* Создать новые табличные пространства для различных таблиц: $HOME/mwd84, $HOME/orw97, $HOME/uzb16
* На основе template0 создать новую базу: sickorangecity
* Создать новую роль, предоставить необходимые права, разрешить подключение к базе.
* От имени новой роли (не администратора) произвести наполнение ВСЕХ созданных баз тестовыми наборами данных. ВСЕ табличные пространства должны использоваться по назначению.
* Вывести список всех табличных пространств кластера и содержащиеся в них объекты.

# Выполнение

## Этап 1. Инициализация кластера БД

Изображение выглядит как текст, снимок экрана, Шрифт

Автоматически созданное описание

Замечание преподавателя: нет смысла указывать файл лога в команде запуска, так как по заданию далее идет настройка файла лога в postgresql.conf.

## Этап 2. Конфигурация и запуск сервера БД

### 1. Настройка способов подключения

Изменения в файле pg\_hba.conf:

Изображение выглядит как текст, снимок экрана, Шрифт, число

Автоматически созданное описание

Изменения в файле postgresql.conf:





### 2. Параметры сервера в соответствии со сценарием OLAP

* max\_connections – отвечает за максимальное количество одновременных подключений.

Значение: 5

Замечание преподавателя: хотя бы одно соединение должно быть всегда доступно администратору, поэтому значение параметра должно быть хотя бы 6.

* shared\_buffers – определяет количество памяти, которое сервер использует для буферов общей памяти (доступной различным процессам). Рекомендуется устанавливать примерно 25% от доступной памяти ОЗУ (устанавливать много больше нет смысла так как postgres активно использует кэш). Изменение (более 25% ОЗУ) требует соответствующего изменения max\_wal\_size.

Значение: 2GB.

* temp\_buffers – определяет максимальное количество памяти, которое сервер использует для буферов отдельных сессий при доступе к временным таблицам.

Значение: 8MB (по-умолчанию).

* work\_mem – устанавливает количество памяти, которое используется в каждой отдельной операции запроса (например, order by / merge join в запросе может быть несколько) перед записью во временные файлы диска.

Значение: 4MB (по-умолчанию).

Замечание преподавателя: так как по условию объемы чтения/записи велики, а количество пользователей ограничено малым числом – этот параметр можно значительно увеличить.

* checkpoint\_timeout – максимальное время между автоматической установкой контрольных точек для WAL файлов. Чем меньше – тем чаще создаются контрольные точки, тем больше нагрузки.

Значение: 5min (по-умолчанию).

* effective\_cache\_size – устанавливает предположение планировщика об эффективном размере дискового кэша, доступного для одного запроса. Чем больше – тем более вероятней, что планировщик будет использовать сканирования индекса таблицы.

Значение: 2 GB.

* fsync – если включен, то сервер будет пытаться проверять, записаны ли изменения данных из основной памяти на диск.

Значение: on.

* commit\_delay – устанавливает задержку перед записью WAL файлов на диск. Чем больше – тем больше (успевших стать готовыми) транзакций будут записаны одной операцией WAL flush.

Значение: 1s.

Замечание преподавателя: задержка перед записью такие коммитов не имеет особого смысла в данной системе, так как упор идет на большой объем коммитов при их малом количестве.

### 3. Настройка WAL

Изображение выглядит как текст, Шрифт, линия, снимок экрана

Автоматически созданное описание

Замечание преподавателя: архивация по заданию не нужна.

### 4. Настройка логов

Изменения в файле `postgresql.conf`:







## Этап 3. Дополнительные табличные пространства и наполнение базы

### 1. Создание табличных пространств



### 2. Создание новой базы данных



Изображение выглядит как текст, снимок экрана, Шрифт

Автоматически созданное описание

### 3. Создание и настройка роли

Изображение выглядит как текст, Шрифт, снимок экрана, линия

Автоматически созданное описание

### 4. Наполнение базы данных

Изображение выглядит как текст, снимок экрана, Шрифт, документ

Автоматически созданное описание

### 5. Список табличных пространств и объектов в кластере

Изображение выглядит как текст, снимок экрана, Шрифт, черный

Автоматически созданное описание

Изображение выглядит как текст, снимок экрана, Шрифт

Автоматически созданное описание

Изображение выглядит как текст, снимок экрана, Шрифт, меню

Автоматически созданное описание

Изображение выглядит как текст, снимок экрана, Шрифт

Автоматически созданное описание

Изображение выглядит как текст, Шрифт, снимок экрана

Автоматически созданное описание

# Вывод

Научился инициализировать и настраивать кластеры БД postgresql. Создал новую базу данных на основе шаблона template0 и наполнил её тестовыми объектами. Научился создавать роли пользователей и устанавливать им права.