НИУ ИТМО

Факультет программной инженерии и компьютерных технологий

Отчет по лабораторной работе №3

по дисциплине Распределенные системы хранения данных

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

Санкт-Петербург

2024

# Задание

Цель работы – настроить процедуру периодического резервного копирования базы данных, сконфигурированной в ходе выполнения лабораторной работы №2, а также разработать и отладить сценарии восстановления в случае сбоев.

Узел из предыдущей лабораторной работы используется в качестве основного. Новый узел используется в качестве резервного. Учётные данные для подключения к новому узлу выдаёт преподаватель. В сценариях восстановления необходимо использовать копию данных, полученную на первом этапе данной лабораторной работы.

## Требования к отчёту

Отчет должен быть самостоятельным документом (без ссылок на внешние ресурсы), содержать всю последовательность команд и исходный код скриптов по каждому пункту задания. Для демонстрации результатов приводить команду вместе с выводом (самой наглядной частью вывода, при необходимости).

## Этап 1. Резервное копирование

* Настроить резервное копирование с основного узла на резервный следующим образом:
  + Периодические холодные полные копии.
  + Полная копия (rsync) по расписанию (cron) раз в сутки. СУБД на время копирования должна останавливаться. На резервном узле хранить 14 копий, после успешного создания пятнадцатой копии, самую старую автоматически уничтожать.
* Подсчитать, каков будет объем резервных копий спустя месяц работы системы, исходя из следующих условий:
  + Средний объем новых данных в БД за сутки: 750МБ.
  + Средний объем измененных данных за сутки: 300МБ.
* Проанализировать результаты.

## Этап 2. Потеря основного узла

Этот сценарий подразумевает полную недоступность основного узла. Необходимо восстановить работу СУБД на РЕЗЕРВНОМ узле, продемонстрировать успешный запуск СУБД и доступность данных.

## Этап 3. Повреждение файлов БД

Этот сценарий подразумевает потерю данных (например, в результате сбоя диска или файловой системы) при сохранении доступности основного узла. Необходимо выполнить полное восстановление данных из резервной копии и перезапустить СУБД на ОСНОВНОМ узле.

Ход работы:

* Симулировать сбой:
  + удалить с диска директорию любой таблицы со всем содержимым.
* Проверить работу СУБД, доступность данных, перезапустить СУБД, проанализировать результаты.
* Выполнить восстановление данных из резервной копии, учитывая следующее условие:
  + исходное расположение директории PGDATA недоступно - разместить данные в другой директории и скорректировать конфигурацию.
* Запустить СУБД, проверить работу и доступность данных, проанализировать результаты.

## Этап 4. Логическое повреждение данных

Этот сценарий подразумевает частичную потерю данных (в результате нежелательной или ошибочной операции) при сохранении доступности основного узла. Необходимо выполнить восстановление данных на ОСНОВНОМ узле следующим способом:

* Восстановление с использованием архивных WAL файлов. (СУБД должна работать в режиме архивирования WAL, потребуется задать параметры восстановления).

Ход работы:

* В каждую таблицу базы добавить 2-3 новые строки, зафиксировать результат.
* Зафиксировать время и симулировать ошибку:
  + удалить любые две таблицы (DROP TABLE)
* Продемонстрировать результат.
* Выполнить восстановление данных указанным способом.
* Продемонстрировать и проанализировать результат.

# Выполнение

## Этап 1. Резервное копирование

### Настройка резервного копирования

Файл cold\_copy.sh на основном узле

Изображение выглядит как текст, снимок экрана, Шрифт

Автоматически созданное описание

Файл remove\_old\_backups.sh на дополнительном узле

Изображение выглядит как текст, снимок экрана, число, Шрифт

Автоматически созданное описание

### Подсчет объема

Так как при создании резервной копии происходит копирование файлов БД целиком, то объем измененных данных не будет влиять на объем таких копий.

По условию:

* прирост объема данных в сутки ;
* максимальное количество копий равно ;
* копии создаются раз в сутки;
* требуется подсчитать объем за месяц.

Объем информации на хранилище (дни от до ) можно найти как сумму арифметической прогрессии:

Так как максимальное количество копий равно , то

Номер последней копии будет равен номеру дня за месяц, пусть будет . Значит

Объем данных резервных копий с номерами от 17 до 30 включительно:

### Анализ

При физическом копировании, обновление данных не влияет на размер копии, поэтому этот вид резервного копирования полезен, когда значение объема обновленных данных мало. Хотя даже при небольшом приросте данных в сутки, объем всех резервных копий может достигать слишком больших значений.

## Этап 2. Потеря основного узла

### Восстановление базы данных на резервном узле

Файл run\_last\_backup.sh на резервном узле

Изображение выглядит как текст, снимок экрана, Шрифт, число

Автоматически созданное описание

Изображение выглядит как текст, снимок экрана, Шрифт, черный

Автоматически созданное описание

Так как в директории pg\_tblspc кластера БД хранятся символические ссылки на табличные пространства, а операция копирования cp не смогла скопировать эти файлы (нет доступа) – их необходимо создать заново для каждого табличного пространства.

### Демонстрация доступности данных

Изображение выглядит как текст, снимок экрана, Шрифт

Автоматически созданное описание

Изображение выглядит как текст, снимок экрана, черно-белый, монохромный

Автоматически созданное описание

## Этап 3. Повреждение файлов БД

### Симуляция сбоя

Удаление с диска директории таблицы “street” со всем содержимым

Изображение выглядит как текст, снимок экрана, Шрифт, черный

Автоматически созданное описание

### Проверка работы СУБД

Изображение выглядит как текст, снимок экрана, Шрифт

Автоматически созданное описание

### Восстановление

Файл load\_and\_run\_last\_backup.sh на основном узле

Изображение выглядит как текст, снимок экрана, Шрифт, число

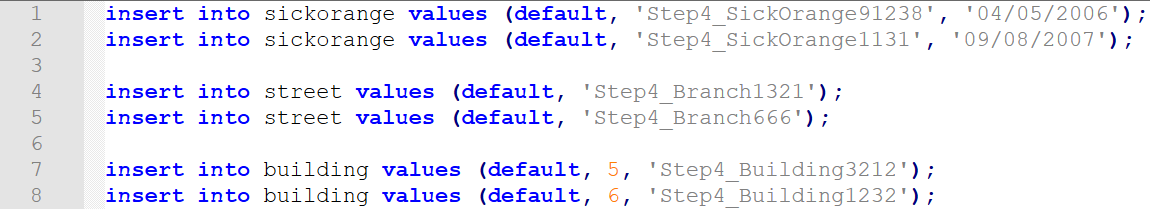
Автоматически созданное описание

Изображение выглядит как текст, снимок экрана, Шрифт, дизайн

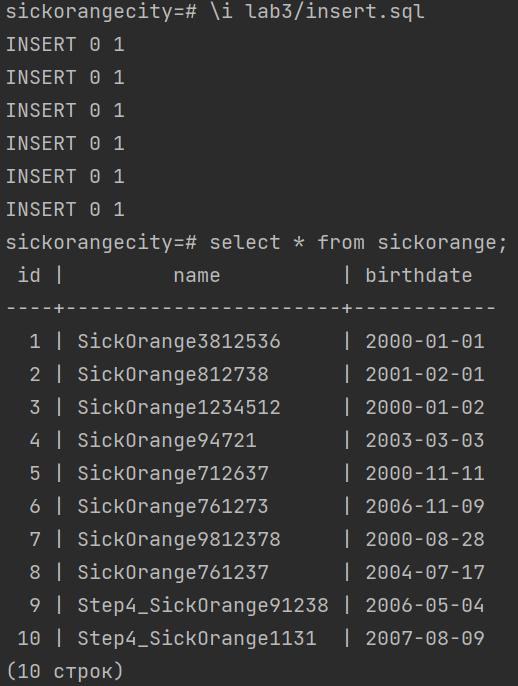
Автоматически созданное описание

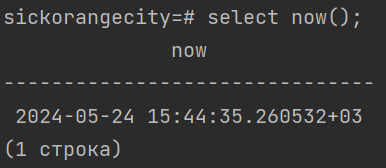
## Этап 4. Логическое повреждение данных

### Добавление новых данных



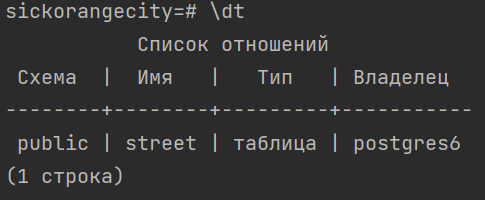
### Фиксация времени и симуляция ошибки



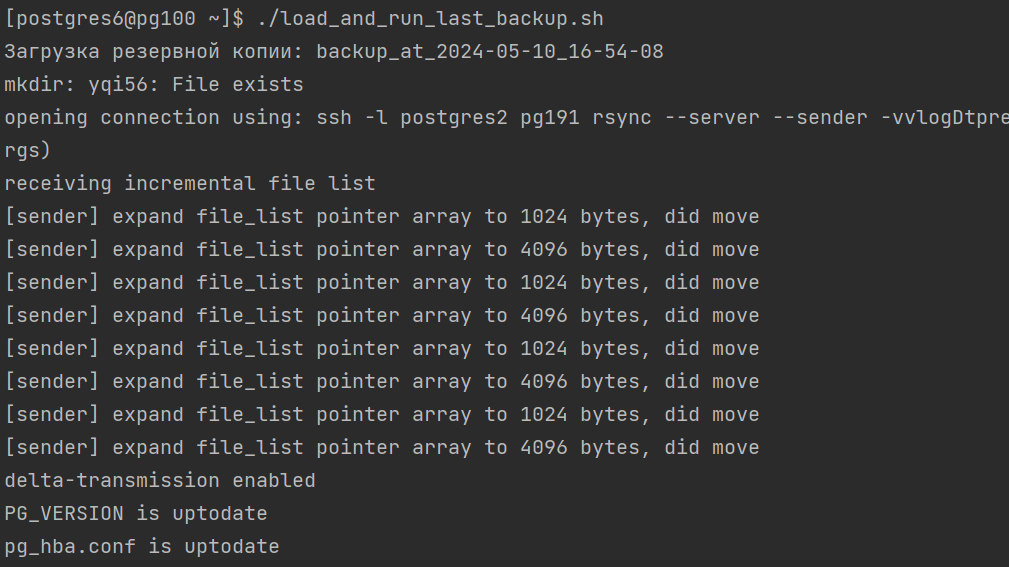




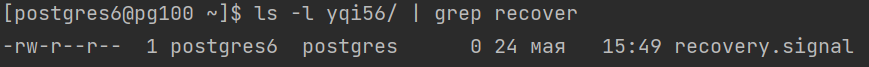
### Результат выполнения ошибки



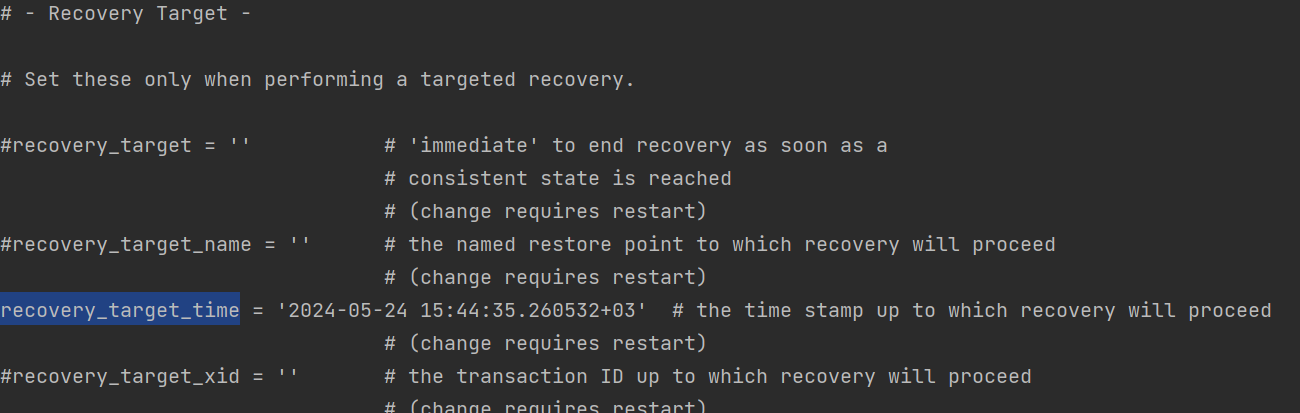
### Восстановление



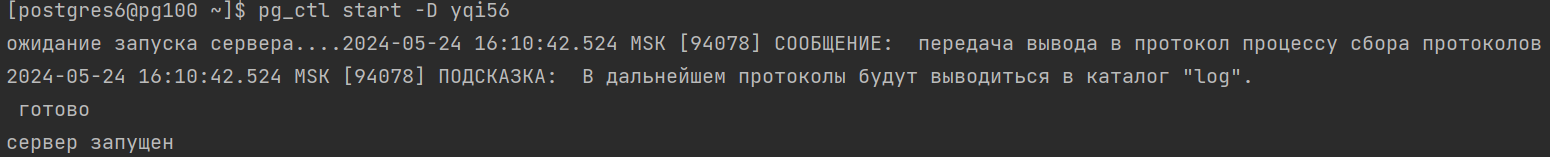
Загрузка последней копии с резервного узла

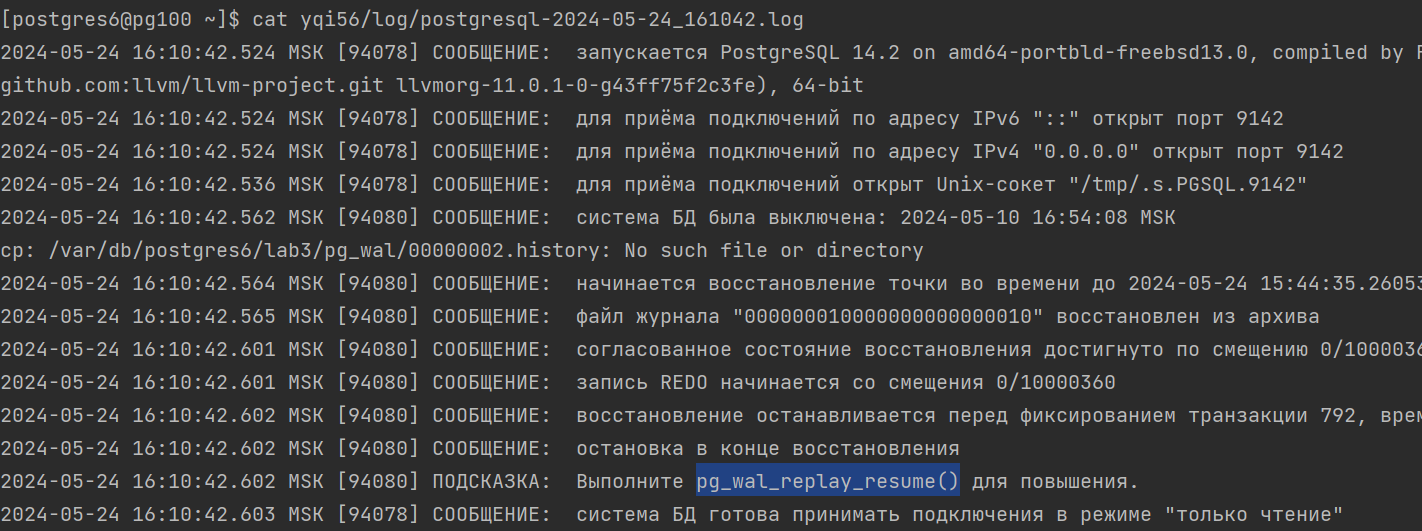


Создание файла recovery.signal в PGDATA

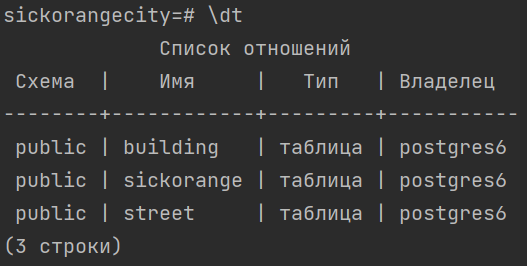


Изменение параметра recovery\_target\_time в postgresql.conf.

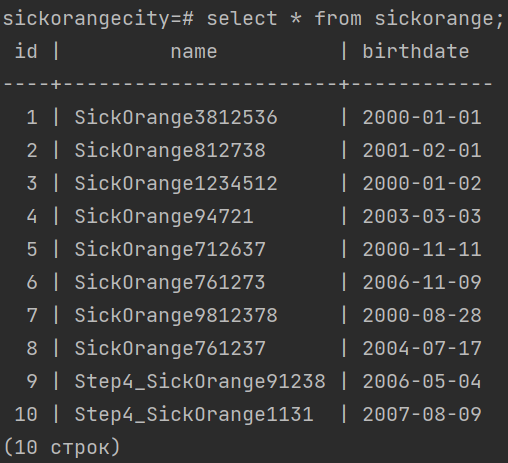




### Результат восстановления



Таблицы восстановлены.



Данные в таблицах соответствуют тому, что было перед их удалением.