Университет ИТМО

Факультет программной инженерии и компьютерной техники

Распределённые системы хранения данных. Лабораторная работа №3.

Группа: P33131
Студент: Смирнов Виктор Игоревич
Преподаватель: Афанасьев Дмитрий Борисович
Вариант: 736

2024

Ключевые слова

База данных, конфигурация PostgreSQL.

Содержание

1 Цель работы и контекст

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.

2 Этап 0. Контекст работы

В предыдущей лабораторной работы была создана база данных. Приведу здесь скрипты для ее инициализации.

2.1 Переменные окружения

```
#!/bin/sh

export DDB_PG_CONF="."
export DDB_PG_USER="postgres0"
export DDB_PG_PASS="pleasehelp"
export DDB_PG_PASS_FILE="$DDB_PG_CONF/pgpass.txt"
export DDB_PG_PORT=9666
export DDB_PG_DATABASE=postgres

export DDB_TABLESPACE_NAME=yqy90
export DDB_TABLESPACE_LOCATION="$HOME/$DDB_TABLESPACE_NAME"
export DDB_NEW_DATABASE_NAME=lazyorangehair
export DDB_NEW_USER=root
export DDB_NEW_USER_PASSWORD=rootik

export PGDATA="$HOME/kop67"

export DDB_PG_BIN_DIR=/usr/lib/postgresql/14/bin
export DDB_INITDB=$DDB_PG_BIN_DIR/initdb
export DDB_PGBIN=$DDB_PG_BIN_DIR/postgres
export DDB_PGBASEBACKUP=$DDB_PG_BIN_DIR/pg_basebackup
export DDB_PGCTL=$DDB_PG_BIN_DIR/pg_ctl
export DDB_PGVERIFYBACKUP=$DDB_PG_BIN_DIR/pg_verifybackup
export DDB_PGDUMP=$DDB_PG_BIN_DIR/pg_dump
export DDB_PGRESTORE=$DDB_PG_BIN_DIR/pg_restore

export DDB_BACKUP_DIR="primary/backup"
export DDB_BACKUP_BASE_DIR="$DDB_BACKUP_DIR/base"
export DDB_BACKUP_WAL_DIR="$DDB_BACKUP_DIR/wal"
export DDB_BACKUP_DUMP_DIR="$DDB_BACKUP_DIR/dump"

export DDB_PRIMARY_USER=$DDB_PG_USER

export DDB_STANDBY_HOST=ddb-standby
export DDB_STANDBY_USER=$DDB_PG_USER
```

Листинг 1: Переменные окружения

2.2 Конфигурация базы данных

```
# $PGDATA/pg_hba.conf (Host-based authentication)

# TYPE  DATABASE    USER    ADDRESS        METHOD
host    all         all     127.0.0.1/32   scram-sha-256 # Permit only localhost
host    all         all     ::1/128        scram-sha-256 # Permit only localhost
host    replication all     localhost      scram-sha-256 # Permit base backup
```

Листинг 2: Конфигурационный файл pg_hba.conf

```
# $PGDATA/postgresql.conf (PostgreSQL configuration file)

# Note: Optimized for OLAP load:
# 5 users, packet r/w 128MB

## CONNECTIONS

listen_addresses  = '127.0.0.1' # Available only from localhost
port              = 9666        # For security
unix_socket_directories = ''    # Only TCP/IP

max_connections                 = 6 # 5 users + 1 extra
superuser_reserved_connections  = 3


```

```
## AUTHENTICATION

authentication_timeout  = 20s           # Type password faster
password_encryption     = scram-sha-256 # Strong password hashing

## RESOURCE USAGE

shared_buffers             = 1024MB  # 128MB * (5 + 3) users
temp_buffers               = 128MB   # 128MB
max_prepared_transactions = 0        # We don't use transactions
work_mem                   = 256MB   # Expected packet size
hash_mem_multiplier        = 1.5     # Smaller hash tables
maintenance_work_mem       = 64MB    # ?
autovacuum_work_mem        = -1
max_stack_depth            = 4MB     # Be prepared for complex queries

temp_file_limit            = 4GB     # Something is wrong if we reach this

## WRITE-AHEAD LOG

checkpoint_timeout    = 5min
fsync                 = off         # Lost data is not critical, as we can recreate
synchronous_commit    = off         # Same
wal_level             = replica     # Enable replication
wal_compression       = off         # WAL must not be so huge?
commit_delay          = 200         # Acceptable to lose 200mc of data
effective_cache_size  = 4GB         # OK?


## REPORTING AND LOGGING

log_destination   = 'stderr'
logging_collector = off
log_directory     = 'log'
log_filename      = 'postgresql-%Y-%m-%d_%H%M%S.log'
log_min_messages  = warning

log_connections    = on
log_disconnections = on

## Archiving

archive_mode    = on
archive_timeout = 16s
archive_command = 'ssh -q <STANDBY_HOST> "test ! -e <STANDBY_WAL_DIR>/%f" && scp %p <
    STANDBY_HOST>:~/<STANDBY_WAL_DIR>'
```

Листинг 3: Конфигурационный файл postgresql.conf

2.3 Создание .pgpass

```
#!/bin/sh

set -e

cd "$(dirname "$0")"

echo "" > ~/.pgpass
echo "localhost:$DDB_PG_PORT:*:$DDB_PG_USER:$DDB_PG_PASS" >> ~/.pgpass
echo "localhost:$DDB_PG_PORT:*:$DDB_STANDBY_USER:$DDB_PG_PASS" >> ~/.pgpass
echo "localhost:$DDB_PG_PORT:*:$DDB_PRIMARY_USER:$DDB_PG_PASS" >> ~/.pgpass
echo "localhost:$DDB_PG_PORT:*:$DDB_NEW_USER:$DDB_NEW_USER_PASSWORD" >> ~/.pgpass
chmod 0600 ~/.pgpass
```

Листинг 4: Файл .pgpass

2.4 Инициализация базы данных

```
#!/bin/sh

set -e

cd "$(dirname "$0")"
```

```

mkdir "$PGDATA" 2> /dev/null

echo "$DDB_PG_PASS" > "$DDB_PG_PASS_FILE"

"$DDB_INITDB" \
  --pgdata="$PGDATA" \
  --locale="ru_RU.CP1251" \
  --encoding="WIN1251" \
  --pwfile="$DDB_PG_PASS_FILE"

cp "$DDB_PG_CONF/pg_hba.conf" "$PGDATA/pg_hba.conf"
cp "$DDB_PG_CONF/postgresql.conf" "$PGDATA/postgresql.conf"
```

Листинг 5: Инициализация базы данных

2.5 Запуск базы данных

```
#!/bin/sh

set -e

cd "$(dirname "$0")"

"$DDB_PGBIN" -D "$PGDATA"
```

Листинг 6: Запуск базы данных

2.6 Настройка базы данных

```
#!/bin/sh

set -e

cd "$(dirname "$0")"

sql() {
    psql -h localhost -p "$DDB_PG_PORT" -c "$1" "$DDB_PG_DATABASE"
}

mkdir "$DDB_TABLESPACE_LOCATION" 2>/dev/null

sql "CREATE TABLESPACE $DDB_TABLESPACE_NAME LOCATION '$DDB_TABLESPACE_LOCATION';"
sql "ALTER DATABASE template1 SET TABLESPACE $DDB_TABLESPACE_NAME;"
sql "CREATE DATABASE $DDB_NEW_DATABASE_NAME TEMPLATE template1;"
sql "CREATE ROLE tester;"
sql "CREATE USER $DDB_NEW_USER WITH LOGIN PASSWORD '$DDB_NEW_USER_PASSWORD';"
sql "GRANT tester TO $DDB_NEW_USER;"

sql() {
  psql -U "$DDB_NEW_USER" -h localhost -p $DDB_PG_PORT -c "$2" "$1"
}

PRV="$DDB_PG_DATABASE"
NEW="$DDB_NEW_DATABASE_NAME"

sql "$PRV" "CREATE TABLE note_prv (id serial PRIMARY KEY, content text NOT NULL);"
sql "$NEW" "CREATE TABLE note_new (id serial PRIMARY KEY, content text NOT NULL);"

sql "$PRV" "INSERT INTO note_prv (content) VALUES ('Note at postgres');"
sql "$NEW" "INSERT INTO note_new (content) VALUES ('Note at lazyorangehair');"

sql "$PRV" "SELECT * FROM note_prv;"
sql "$NEW" "SELECT * FROM note_new;"
```

Листинг 7: Настройка базы данных

3 Этап 1. Резервное копирование

3.1 Задача

1. Настроить резервное копирование с основного узла на резервный следующим образом:

(a) Первоначальная полная копия + непрерывное архивирование.

(b) Включить для СУБД режим архивирования WAL;

(c) настроить копирование WAL (scp) на резервный узел;

(d) создать первоначальную резервную копию (pg_ basebackup),

(e) скопировать на резервный узел (rsync).

2. Подсчитать, каков будет объем резервных копий спустя месяц работы системы, исходя из следующих условий:

   (a) Средний объем новых данных в БД за сутки: 650МБ.

   (b) Средний объем измененных данных за сутки: 950МБ.

3. Проанализировать результаты.

3.2 Подготовка секретов

Нам необходимо будет отправлять базовую резервную копию, а так WAL файлы на резервный узел, так что сперва следует сгенерировать и распределить ключи шифрования для безопасной передачи данных между узлами.

```
#!/bin/sh

set -e

echo "[primary] Generating ssh key..."
ssh-keygen -t rsa -f ~/.ssh/id_rsa -N ""

echo "[primary] Generated ssh key:"
cat ~/.ssh/id_rsa.pub
```

Листинг 8: Генерация ключей

Далее я авторизовал публичный ключ primary узла на standby. Теперь можно проверить, что передача данных скорее всего будет работать.

```
#!/bin/sh

set -e

ssh -q $1 "echo Hello, World!"
```

Листинг 9: Проверка подключения

3.3 Конфигурация primary узла для резервного копирования

Включаем архивирование, будем отправлять WAL файлы на standby каждые 16 секунд.

```
## Archiving

archive_mode    = on
archive_timeout = 16s
archive_command = 'ssh -q <STANDBY_HOST> "test ! -e <STANDBY_WAL_DIR>/%f" && scp %p <
    STANDBY_HOST>:~/<STANDBY_WAL_DIR>'
```

Листинг 10: Ключевые строчки в конфигурационном файле

3.4 Создание базовой резервной копии

```
#!/bin/sh

set -e

echo "[primary] Creating base backup..."
"$DDB_PGBASEBACKUP" \
    --host="localhost" \
    --port="$DDB_PG_PORT" \
```

```
    --pgdata="$HOME/$DDB_BACKUP_BASE_DIR" \
    --format="tar" \
    --wal-method="fetch" \
    --no-password

echo "[primary] Sending to '$DDB_STANDBY_USER@$DDB_STANDBY_HOST'..."
rsync -ave ssh \
    "$HOME/$DDB_BACKUP_BASE_DIR" \
    $DDB_STANDBY_USER@$DDB_STANDBY_HOST:~/$DDB_BACKUP_DIR
```

Листинг 11: Создание базовой резервной копии

3.5 Подготовка standby

```
#!/bin/sh

set -e

echo "[standby] Prepairing..."

mkdir -p "$HOME/$DDB_BACKUP_WAL_DIR"
mkdir -p "$HOME/$DDB_BACKUP_BASE_DIR"
mkdir -p "$HOME/$DDB_BACKUP_DUMP_DIR"
```

Листинг 12: Подготовка standby

3.6 Полная настройка primary узла

```
#!/bin/sh

set -e

echo "[primary] Creating '.pgpass' file..."
sh common/pgpass.sh

echo "[primary] Editing 'postgresql.conf' file..."
sh primary/config.sh

echo "[primary] Initializing the database..."
sh primary/init.sh

echo "[primary] Starting the database..."
sh common/start.sh &

echo "[primary] Waiting the database startup..."
sleep 2

echo "[primary] Settings up the database..."
sh primary/setup.sh

echo "[primary] All right!"
```

Листинг 13: Полная настройка primary узла

3.7 Действия системного администратора по первоначальной настройке системы

1. Получить конфигурационные файлы системы из репозитория `https://github.com/vityaman-edu/ddb-homework/tree/trunk/lab-3`
2. Доставить директории `db/common` и `db/primary` на `primary` узел, разместив их в домашней директории пользователя, от лица которого будет запущена система
3. Доставить директории `db/common` и `db/standby` на `standby` узел, разместив их в домашней директории пользователя, от лица которого будет запущена система
4. На `primary` узле сгенерировать ключи для `primary` узла при помощи скрипта `common/ssh-keygen.sh` и авторизовать публичный ключ на узле standby
5. Проверить на `primary` узле возможность ssh соединения с `standby` узлом при помощи `common/ssh-test.sh`

6. Подготовить `standby` узел к резервированию, выполнив на нем `source common/env.sh && sh standby/prepare.sh`
7. Запустить `primary` узел, выполнив на нем `source common/env.sh && sh common/pgpass.sh && sh primary/full.sh`
8. Создать базовую резервную копию на `primary` узле и отправить ее на `standby` узел: `sh primary/backup.sh`
9. Наполнить данными базу данных на `primary` узле: `sh primary/fill.sh`
10. Убедиться, что базовая резервная копия и WAL файлы доставлены на `standby` узел

3.8 Объем резервных копий спустя месяц работы

1. Средний объем новых данных в БД за сутки: N = 650МБ
2. Средний объем измененных данных за сутки: M = 950МБ

Базовая резервная копия у меня весит $B = 50$ `MB`.
Сперва оценим величину грубо:

$$B + (N + M) \cdot 30 = 46 \text{ GB}$$

Оценка является скорее всего не самой достоверной, ведь возможно использование алгоритмов сжатия для понижения объемов хранимых данных. Не берусь оценивать исследуемую величину, ведь это лучше доказывать экспериментально.

4 Этап 2. Потеря основного узла

4.1 Задача

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

4.2 Восстановление СУБД на резервном узле

Для этого необходимо просто выполнить `source common/env.sh && sh common/restore.sh standby anon-tblspc` на `standby` узле, предварительно убедившись, что в директории `primary/backup/base` находятся файлы базовой резервной копии, а в директории `primary/backup/wal` есть WAL сегменты.

```
#!/bin/sh

# set -e

INSTANCE=$1
echo "[restore] INSTANCE: $INSTANCE"

MODE=$2
echo "[restore] MODE: $MODE"

TARGET_TIME=$3
echo "[restore] TARGET_TIME: $TARGET_TIME"

echo "[restore] Removing existing database installation..."
sh common/clear.sh

mkdir -p $PGDATA
chmod 0700 $PGDATA

echo "[restore] Extracting base backup..."
tar \
    --extract \
    -f "$HOME/$DDB_BACKUP_BASE_DIR/base.tar" \
```

```
    --directory=$PGDATA

echo "[restore] Restoring tablespaces..."
while read -r line; do
    TABLESPACE_OID=$(echo $line | awk '{print $1}')

    if [ "$MODE" = "anon-tblspc" ]; then
        TABLESPACE_DIR="$HOME/tablespace/$TABLESPACE_OID"
    else
        TABLESPACE_DIR=$(echo $line | awk '{print $2}')
    fi

    echo "[restore][$TABLESPACE_OID] Restoring tablespace..."

    echo "[restore][$TABLESPACE_OID] Directory: $TABLESPACE_DIR"
    mkdir -p $TABLESPACE_DIR

    echo "[restore][$TABLESPACE_OID] Extracting..."
    tar --extract \
        -f "$HOME/$DDB_BACKUP_BASE_DIR/$TABLESPACE_OID.tar" \
        --directory=$TABLESPACE_DIR

    echo "[restore][$TABLESPACE_OID] Creating symbolic link..."
    ln -s $TABLESPACE_DIR $PGDATA/pg_tblspc/$TABLESPACE_OID
done < $PGDATA/tablespace_map

echo "[restore] Checking base backup integrity..."
"$DDB_PGVERIFYBACKUP" \
    --manifest-path="$HOME/$DDB_BACKUP_BASE_DIR/backup_manifest" \
    --wal-directory="$HOME/$DDB_BACKUP_WAL_DIR" \
    $PGDATA

echo "[restore] Removing 'tablespace_map'..."
rm $PGDATA/tablespace_map

# Patching is used as we need to change values depending on env variables
echo "[restore] Patching postgresql.conf: add 'restore_command'..."
RESTORE_CMD="restore_command = 'cp ~/$DDB_BACKUP_WAL_DIR/%f %p'"
echo "$RESTORE_CMD" >> $PGDATA/postgresql.conf

if [ "$INSTANCE" = "standby" ]; then
    echo "[restore] Patching postgresql.conf: disable archive..."
    sed -i -e "s+archive_mode+#archive_mode+g" $PGDATA/postgresql.conf
fi

if [ "$TARGET_TIME" != "" ]; then
    echo "[restore] Patching postgresql.conf: setting target time..."
    echo "recovery_target_time = '$TARGET_TIME'" >> $PGDATA/postgresql.conf
    echo "recovery_target_inclusive = false"     >> $PGDATA/postgresql.conf
fi

echo "[restore] Signalling of recovery..."
touch $PGDATA/recovery.signal

echo "[restore] Is ready for startup!"
```

Листинг 14: Восстановление СУБД

4.3 Действия на primary узле

```
postgres0@2d09031f584d:~$ history
1  ls
2  source common/env.sh
3  sh common/ssh-keygen.sh
4  sh common/ssh-test.sh ddb-standby
5  sh common/pgpass.sh
6  sh primary/full.sh
7  sh primary/backup.sh
8  sh primary/fill.sh
9  history
```

Листинг 15: Действия на primary узле

4.4 Действия на standby узле

```
postgres0@c33884c20d42:~$ history
2  ls
3  source common/env.sh
4  vim .ssh/authorized_keys
5  sh standby/prepare.sh
6  ls
7  ls primary/backup/base/
8  ls primary/backup/wal/
9  ls primary/backup/base
10  ls primary/backup/wal/
11  sh standby/restore.sh
12  sh common/start.sh
13  bg
14  history
```

Листинг 16: Действия на standby узле

```
    $ sh common/start.sh
2024-05-28 17:29:22.989 GMT [61131] СООБЩЕНИЕ: запускается PostgreSQL 14.2 on amd64-portbld-freebsd1
compiled by FreeBSD clang version 11.0.1 (git@github.com:llvm/llvm-project.git llvmorg-11.0.1-0-g43f
64-bit
2024-05-28 17:29:22.989 GMT [61131] СООБЩЕНИЕ: для приёма подключений по адресу IPv4 "127.0.0.1"откр
порт 9666
2024-05-28 17:29:22.991 GMT [61132] СООБЩЕНИЕ: работа системы БД была прервана; последний
момент работы: 2024-05-28 17:19:51 GMT
cp: /var/db/postgres1/primary/backup/wal/00000002.history: No such file or directory
2024-05-28 17:29:23.012 GMT [61132] СООБЩЕНИЕ: начинается восстановление архива
2024-05-28 17:29:23.014 GMT [61132] СООБЩЕНИЕ: файл журнала "000000010000000000000002"восстановлен
из архива
2024-05-28 17:29:23.017 GMT [61132] СООБЩЕНИЕ: запись REDO начинается со смещения 0/2000028
2024-05-28 17:29:23.017 GMT [61132] СООБЩЕНИЕ: согласованное состояние восстановления достигнуто
по смещению 0/2000138
2024-05-28 17:29:23.018 GMT [61131] СООБЩЕНИЕ: система БД готова принимать подключения в
режиме "только чтение"
2024-05-28 17:29:23.019 GMT [61132] СООБЩЕНИЕ: файл журнала "000000010000000000000003"восстановлен
из архива
2024-05-28 17:29:23.029 GMT [61132] СООБЩЕНИЕ: файл журнала "000000010000000000000004"восстановлен
из архива
cp: /var/db/postgres1/primary/backup/wal/000000010000000000000005: No such file or directory
2024-05-28 17:29:23.031 GMT [61132] СООБЩЕНИЕ: записи REDO обработаны до смещения 0/4000110,
нагрузка системы: CPU: пользов.: 0.00 с, система: 0.00 с, прошло: 0.01 с
2024-05-28 17:29:23.031 GMT [61132] СООБЩЕНИЕ: последняя завершённая транзакция была выполнена
в 2024-05-28 17:20:13.468706+00
2024-05-28 17:29:23.034 GMT [61132] СООБЩЕНИЕ: файл журнала "000000010000000000000004"восстановлен
из архива
cp: /var/db/postgres1/primary/backup/wal/00000002.history: No such file or directory
2024-05-28 17:29:23.035 GMT [61132] СООБЩЕНИЕ: выбранный ID новой линии времени: 2
2024-05-28 17:29:23.043 GMT [61132] СООБЩЕНИЕ: восстановление архива завершено
cp: /var/db/postgres1/primary/backup/wal/00000001.history: No such file or directory
2024-05-28 17:29:23.054 GMT [61131] СООБЩЕНИЕ: система БД готова принимать подключения
```

```
$ sh common/connect.sh postges
postgres=# select * from note_prv;
 id |          content
----+---------------------------
  1 | Note at postgres
  2 | Another note at postgres
(2 rows)

$ sh common/connect.sh lazyorangehair
lazyorangehair=# select * from note_new;
 id |             content
```

```
----+--------------------------------
  1 | Note at lazyorangehair
  2 | Another note at lazyorangehair
(2 rows)
```

Листинг 17: Состояние СУБД на standby узле

Как мы видим, в базе данных сохранились не только данные с базовой копии, но и подтянулись изменения из WAL сегментов.

5 Этап 3. Повреждение файлов БД

5.1 Задача

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:

1. Симулировать сбой: удалить с диска директорию любой таблицы со всем содержимым.
2. Проверить работу СУБД, доступность данных, перезапустить СУБД, проанализировать результаты.
3. Выполнить восстановление данных из резервной копии, учитывая следующее условие: исходное расположение дополнительных табличных пространств недоступно - разместить в другой директории и скорректировать конфигурацию.
4. Запустить СУБД, проверить работу и доступность данных, проанализировать результаты.

5.2 Решение

Попробуем удалить таблицу `note_prv`. Для этого узнаем, где представлены ее данные в файловой системе.

```
$ psql -h localhost -p "$DDB_PG_PORT" "$DDB_PG_DATABASE"
postgres=# SELECT pg_relation_filepath('note_prv');
 pg_relation_filepath
----------------------
 base/14115/16389
(1 stroka)
```

Листинг 18: Получаем физическую локацию таблицы

Начинаем уничтожение.

```
$ rm $PGDATA/base/14374/16389
```

Листинг 19: Удаляем файл таблицы

Наблюдаем эффекты.

```
 $ psql -h localhost -p "$DDB_PG_PORT$DDB_PG_DATABASE"
postgres=# select * from note_prv;
ERROR: could not open file "base/14374/16389": No such file or directory
```

Попробуем перезапустить базу данных.

```
$ sh common/stop.sh
waiting for server to shut down.... done
server stopped

$ sh common/start.sh
2024-05-28 10:52:21.905 GMT [9200] LOG:  starting PostgreSQL 14.12 (Ubuntu 14.12-1.
    pgdg20.04+1) on x86_64-pc-linux-gnu, compiled by gcc (Ubuntu 9.4.0-1ubuntu1~20.04.2)
     9.4.0, 64-bit
2024-05-28 10:52:21.905 GMT [9200] LOG:  listening on IPv4 address "127.0.0.1", port
    9666
2024-05-28 10:52:21.908 GMT [9206] LOG:  database system was shut down at 2024-05-28
    10:52:16 GMT
```

```
2024-05-28 10:52:21.916 GMT [9200] LOG:  database system is ready to accept connections

$ psql -h localhost -p "$DDB_PG_PORT" "$DDB_PG_DATABASE"
postgres=# select * from note_prv;
2024-05-28 10:54:18.117 GMT [9230] ERROR:  could not open file "base/14374/16389": No
    such file or directory
2024-05-28 10:54:18.117 GMT [9230] STATEMENT:  select * from note_prv;
ERROR:  could not open file "base/14374/16389": No such file or directory
```

Листинг 20: Перезапускаем датабазу

Видим, что postgres не заметил пропажи, значит не проверил целостность файлов. Наверное, должен существовать способ осуществить проверку целостность его файлов. `pg_checksums` не обнаружил проблему. По идее, можно использовать для проверки утилиту `pg_verifybackup`, которая проверит, соответствует ли содержимое `PGDATA` заданному бэкапу.

Скачиваем бэкап с `standby`, сносим нашу больную датабазу, запускаем восстановление и проверяем результат.

```
$ sh primary/download.sh

$ ls primary/backup/base primary/backup/wal/
primary/backup/base:
16384.tar  backup_manifest  base.tar

primary/backup/wal/:
000000010000000000000001
000000010000000000000002
000000010000000000000002.00000028.backup
000000010000000000000003
000000010000000000000004
000000010000000000000005

$ sh common/restore.sh
```

Листинг 21: Восстанавливаем базу данных

```
    $ sh common/start.sh
2024-05-28 17:42:12.835 GMT [61895] СООБЩЕНИЕ: запускается PostgreSQL 14.2 on amd64-portbld-freebsd1
compiled by FreeBSD clang version 11.0.1 (git@github.com:llvm/llvm-project.git llvmorg-11.0.1-0-g43f
64-bit
2024-05-28 17:42:12.835 GMT [61895] СООБЩЕНИЕ: для приёма подключений по адресу IPv4 "127.0.0.1"откр
порт 9666
2024-05-28 17:42:12.838 GMT [61896] СООБЩЕНИЕ: работа системы БД была прервана; последний
момент работы: 2024-05-28 17:19:51 GMT
cp: /var/db/postgres0/primary/backup/wal/00000002.history: No such file or directory
2024-05-28 17:42:12.840 GMT [61896] СООБЩЕНИЕ: начинается восстановление архива
2024-05-28 17:42:12.842 GMT [61896] СООБЩЕНИЕ: файл журнала "000000010000000000000002"восстановлен
из архива
2024-05-28 17:42:12.843 GMT [61896] СООБЩЕНИЕ: запись REDO начинается со смещения 0/2000028
2024-05-28 17:42:12.843 GMT [61896] СООБЩЕНИЕ: согласованное состояние восстановления достигнуто
по смещению 0/2000138
2024-05-28 17:42:12.844 GMT [61895] СООБЩЕНИЕ: система БД готова принимать подключения в
режиме "только чтение"
2024-05-28 17:42:12.845 GMT [61896] СООБЩЕНИЕ: файл журнала "000000010000000000000003"восстановлен
из архива
2024-05-28 17:42:12.852 GMT [61896] СООБЩЕНИЕ: файл журнала "000000010000000000000004"восстановлен
из архива
2024-05-28 17:42:12.854 GMT [61896] СООБЩЕНИЕ: файл журнала "000000010000000000000005"восстановлен
из архива
2024-05-28 17:42:12.856 GMT [61896] СООБЩЕНИЕ: файл журнала "000000010000000000000006"восстановлен
из архива
cp: /var/db/postgres0/primary/backup/wal/000000010000000000000007: No such file or directory
2024-05-28 17:42:12.858 GMT [61896] СООБЩЕНИЕ: записи REDO обработаны до смещения 0/60000D8,
нагрузка системы: CPU: пользов.: 0.00 с, система: 0.00 с, прошло: 0.01 с
2024-05-28 17:42:12.858 GMT [61896] СООБЩЕНИЕ: последняя завершённая транзакция была выполнена
в 2024-05-28 17:20:13.468706+00
```

```
2024-05-28 17:42:12.860 GMT [61896] СООБЩЕНИЕ: файл журнала "000000010000000000000006"восстановлен
из архива
cp: /var/db/postgres0/primary/backup/wal/00000002.history: No such file or directory
2024-05-28 17:42:12.862 GMT [61896] СООБЩЕНИЕ: выбранный ID новой линии времени: 2
2024-05-28 17:42:12.865 GMT [61896] СООБЩЕНИЕ: восстановление архива завершено
cp: /var/db/postgres0/primary/backup/wal/00000001.history: No such file or directory
2024-05-28 17:42:12.876 GMT [61895] СООБЩЕНИЕ: система БД готова принимать подключения
$ psql -h localhost -p "$DDB_PG_PORT$DDB_PG_DATABASE"
postgres=# select * from note_prv;
id | content
--+-------------
1 | Note at postgres
2 | Another note at postgres
(2 rows)
```

Починили.

Вторая часть выполняется аналогично, но нужно вызывать восстановление через `sh common/restore.bash primary anon-tblspc`.

6 Этап 4. Логическое повреждение данных

6.1 Задача

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:

Генерация файла на резервном узле с помощью pg_dump и последующее применение файла на основном узле.

Ход работы:

1. В каждую таблицу базы добавить 2-3 новые строки, зафиксировать результат.
2. Зафиксировать время и симулировать ошибку: в любой таблице с внешними ключами подменить значения ключей на случайные (INSERT, UPDATE)
3. Продемонстрировать результат.
4. Выполнить восстановление данных указанным способом.
5. Продемонстрировать и проанализировать результат.

6.2 Фиксируем состояние БД

Сперва запустим вставку нескольких строк на основном узле через скрипт, после чего фиксируем текущее время 2024-05-28 17:46:57.818.

```
#!/bin/sh

set -e

cd "$(dirname "$0")"

sql() {
    psql -U "$DDB_PG_USER" -h localhost -p $DDB_PG_PORT -c "$1" $DDB_PG_DATABASE
}

sql "INSERT INTO note_prv (content) VALUES ('The first testing note');"
sql "INSERT INTO note_prv (content) VALUES ('The second testing note');"
sql "INSERT INTO note_prv (content) VALUES ('The third testing note');"
```

Листинг 22: Вставка дополнительных строк

Теперь испортим. Не очень понятно, как я должен портить данные, поэтому просто вставлю строку-индикатор: `insert into note_prv (content) values ('aaaaaaaaaaa');`

6.3 Восстановление дампа из бэкапа на резервном узле

Теперь летим на резервный узел и там восстанавливаем бд до заданного момента времени: `sh standby/backup2dump.sh "2024-05-28 17:46:57"`.

```
#!/bin/sh

set -e

TARGET_TIME=$1
if [ "$TARGET_TIME" = "" ]; then
    echo "[backup2dump] Expected TARGET_TIME parameter!"
    exit 1
fi

echo "[backup2dump] Starting converting backup to sql dump..."

echo "[backup2dump] Starting restoration..."
sh common/restore.sh standby anon-tblspc "$TARGET_TIME"

echo "[backup2dump] Starting the database..."
sh common/start.sh &

echo "[backup2dump] Waiting the database..."
sleep 2

echo "[backup2dump] Starting dumping..."
sh common/sql-dump.sh

echo "[backup2dump] Shutting down the database..."
sh common/stop.sh

echo "[backup2dump] Done!"
```

Листинг 23: Скрипт backup2dump

```
#!/bin/sh

set -e

dump() {
    echo "[dump] Dumping '$1'..."
    pg_dump \
        --host=localhost \
        --port=$DDB_PG_PORT \
        --username=$DDB_PG_USER \
        --format=tar \
        --clean \
        --if-exists \
        --blobs \
        --file=$DDB_BACKUP_DUMP_DIR/$1.tar \
        $1
}

dump $DDB_PG_DATABASE
dump $DDB_NEW_DATABASE_NAME
```

Листинг 24: Скрипт sql-dump

После этого в директории с бэкапами появились нужные нам архивы.

Теперь летим на главный узел и запускаем наш кошмар там: `sh primary/donwload.sh && sh common/sql-restore.sh`.

```
#!/bin/sh

set -e

echo "[sql-restore] Restoring the database from an SQL dump..."

restore() {
    echo "[sql-restore] Restoring '$1'..."
    "$DDB_PGRESTORE" \
        --dbname=$1 \
        --host=localhost \
```

```
        --port=$DDB_PG_PORT \
        --clean \
        --if-exists \
        $DDB_BACKUP_DUMP_DIR/$1.tar
}

restore $DDB_PG_DATABASE
restore $DDB_NEW_DATABASE_NAME

echo "[sql-restore] Done!"
```

Листинг 25: Скрипт sql-restore

Проверяем, что в таблицах.

```
 postgres=# select * from note_prv;
id | content
1 | Note at postgres
2 | Another note at postgres
35 | The first testing note
36 | The second testing note
37 | The third testing note
```

7 Вывод

Данная лабораторная работа помогла мне по-настоящему ощутить, какого это - быть администратором PostgreSQL. Я научился создавать резервные копии и накатывать их. В предложенной реализации есть проблема, связанная с тем, что я не накатываю валы переодически, из-за чего через 20 лет я не смогу восстановить БД за разумное время. Также я научился писать переносимые шел скрипты для автоматизации развертывания СУБД.

8 Бонус

```
FROM ubuntu:20.04

RUN apt-get update
RUN apt-get -y install wget gpg gnupg2 lsb-release

RUN wget --quiet -O - https://www.postgresql.org/media/keys/ACCC4CF8.asc | apt-key add -
RUN sh -c 'echo "deb https://apt.postgresql.org/pub/repos/apt $(lsb_release -cs)-pgdg
    main" > /etc/apt/sources.list.d/pgdg.list'
RUN apt-get update

RUN DEBIAN_FRONTEND=noninteractive apt-get -y install postgresql-14

RUN apt-get -y install locales locales-all

RUN apt-get -y install vim rsync

RUN apt-get -y install openssh-client openssh-server
RUN mkdir -p /run/sshd && chmod 755 /run/sshd
RUN echo "PasswordAuthentication yes" >> /etc/ssh/sshd_config
RUN echo "PermitRootLogin no" >> /etc/ssh/sshd_config

RUN useradd -ms /bin/bash postgres0
RUN echo "postgres0:changeme" | chpasswd
USER postgres0
WORKDIR /home/postgres0
RUN mkdir ~/.ssh
RUN touch ~/.ssh/authorized_keys

USER root
RUN cd /etc/ssh && ssh-keygen -A

CMD service ssh start && bash
EXPOSE 22
```

Список литературы

[1] PostgreSQL Documentation: сайт. - 2024. - URL: https://www.postgresql.org/docs/14/index.html (дата обращения: 06.04.2024) - Текст : электронный.