# Университет ИТМО Факультет программной инженерии и компьютерной техники

# Учебно-исследовательская работа №4 (УИР 4)

По дисциплине «Распределённые системы хранения данных»

Выполнили: Нодири Хисравхон Бахилин Никита Витальевич Вариант: 3023

Преподаватель: Максимов Андрей Николаевич

# Содержание

1	Усл	овие	
	1.1	Цель работы	
	1.2	Требования к выполнению работы	
	1.3	Этап 1. Конфигурация	
	1.4	Этап 2. Симуляция и обработка сбоя	•
	1.5	Восстановление	
<b>2</b>	Эта	ы 1 — Конфигурация кластера и балансировщика	4
	2.1	/etc/hosts для упрощения адресации	4
	2.2	PostgreSQL 16: настройка master (pg1)	
	2.3	Поднимаем standby (pg2)	
	2.4	Заполняем тестовыми данными	
	2.5	Установка и настройка PgBouncer (pg1)	
	2.6	Проверка балансировки	
3	Этап 2. Симуляция и обработка сбоя		
	3.1	Client	7
	3.2	Crash & Recover	
	3.3	Демонстрация работы	
4	Вы	вод	14

### 1 Условие

#### 1.1 Цель работы

**Цель работы.** Ознакомление с методами и средствами построения отказоустойчивых решений на базе СУБД PostgreSQL; получение практических навыков восстановления работы системы после отказа. Работа рассчитана на двух человек и включает три этапа: настройка, симуляция и обработка сбоя, восстановление.

## 1.2 Требования к выполнению работы

- В качестве хостов использовать одинаковые виртуальные машины.
- В первую очередь необходимо обеспечить сетевую связность между ВМ.
- Для подключения к СУБД (например, через psql), использовать отдельную виртуальную или физическую машину.
- Демонстрировать наполнение базы и доступ на запись на примере не менее, чем двух таблиц, столбцов, строк, транзакций и клиентских сессий.

## 1.3 Этап 1. Конфигурация

Pазвернуть postgres на двух узлах в режиме балансировки нагрузки. Использовать pgbouncer. Продемонстрировать обработку транзакций обоими серверами.

#### 1.4 Этап 2. Симуляция и обработка сбоя

#### 2.1 Подготовка:

- Установить несколько клиентских подключений к СУБД.
- Продемонстрировать состояние данных и работу клиентов в режиме чтение/запись.

#### 2.2 Сбой:

• Симулировать ошибку диска на основном узле — удалить директорию PGDATA со всем содержимым.

#### 2.3 Обработка:

- Найти и продемонстрировать в логах релевантные сообщения об ошибках.
- Выполнить переключение (failover) на резервный сервер.
- Продемонстрировать состояние данных и работу клиентов в режиме чтение/запись.

#### 1.5 Восстановление

- Восстановить работу основного узла откатить действие, выполненное с виртуальной машиной на этапе 2.2.
- Актуализировать состояние базы на основном узле накатить все изменения данных, выполненные на этапе 2.3.
- Восстановить исправную работу узлов в исходной конфигурации (в соответствии с этапом 1).
- Продемонстрировать состояние данных и работу клиентов в режиме чтение/запись.

# 2 Этап 1 — Конфигурация кластера и балансировщика

#### Инфраструктура

```
• pg1 — 147.45.40.78
```

- pg2 87.120.36.97
- client Отдельная машина/ПК с psql

## 2.1 /etc/hosts для упрощения адресации

```
echo "147.45.40.78 pg1" | sudo tee -a /etc/hosts
echo "87.120.36.97 pg2" | sudo tee -a /etc/hosts
```

## 2.2 PostgreSQL 16: настройка master (pg1)

```
# /etc/postgresql/16/main/postgresql.conf (fragment)
listen_addresses = '*'
wal_level = replica
max_wal_senders = 10
wal_keep_size = 256MB
archive_mode = on
archive_command = 'test ! -f /var/lib/postgresql/archive/%f && cp %p /var/lib/
postgresql/archive/%f'
hot_standby = on
```

Создаём архив каталог и перезапускаем:

```
sudo mkdir -p /var/lib/postgresql/archive
sudo chown postgres:postgres /var/lib/postgresql/archive
sudo systemctl restart postgresql
```

Создаём роли и БД:

```
CREATE ROLE replicator WITH REPLICATION LOGIN PASSWORD 'replpass';
CREATE ROLE labuser WITH LOGIN PASSWORD 'labpass';
CREATE DATABASE labdb OWNER labuser;

GRANT ALL ON TABLE streams TO labuser;
GRANT ALL ON TABLE idols TO labuser;
GRANT USAGE, SELECT ON ALL SEQUENCES IN SCHEMA public TO labuser;
```

Правила доступа (pg\_hba.conf):

```
# Clients
host labdb labuser 0.0.0.0/0 md5
# Replication from pg2
host replication replicator 87.120.36.97/32 md5
```

## 2.3 Поднимаем standby (pg2)

```
# Off old cluster
sudo systemctl stop postgresql
sudo -u postgres rm -rf /var/lib/postgresql/16/main/*

# Base backup + slot
sudo -u postgres pg_basebackup \
-h pg1 -D /var/lib/postgresql/16/main \
-U replicator -Fp -Xs -P -R -C -S pg2_slot

sudo systemctl start postgresql
```

pg2 pg hba.conf:

```
host replication replicator 147.45.40.78/32 md5
host all all 0.0.0.0/0 md5
```

 $\Phi$ айл postgresql.auto.conf на pg2 содержит:

```
primary_conninfo = 'user=replicator password=replpass host=pg1 port=5432'
primary_slot_name = 'pg2_slot'
```

#### Проверяем:

```
-- pg1
SELECT client_addr,state,sync_state FROM pg_stat_replication;
-- pg2
SELECT pg_is_in_recovery();
```

#### 2.4 Заполняем тестовыми данными

```
CREATE TABLE IF NOT EXISTS idols (
               serial PRIMARY KEY,
               text NOT NULL,
    name_en
    name_jp
               text NOT NULL,
                text NOT NULL,
    gen
    debut_date date
);
CREATE TABLE IF NOT EXISTS streams (
               serial PRIMARY KEY,
                       REFERENCES idols(id),
    talent_id
               int
    title
                text,
    viewers_k int,
    stream_date date
);
BEGIN;
                                                     'O gen', '201/-00 (2018-06-01'), '2018-06-01'),
INSERT INTO idols (name_en, name_jp, gen, debut_date) VALUES
                                        ' ,
 ('Tokino Sora',
                                                                      <sup>'2017-09-07'</sup>),
                                                     '1 gen ',
 ('Shirakami Fubuki',
                                                    'GAMERS',
 ('Minato Aqua',
                                                     '3 gen ',
 ('Usada Pekora',
                                                                     <sup>'2019-07-17')</sup>,
                                                     '3 gen ',
                                                                     <sup>'2019-07-17')</sup>,
 ('Houshou Marine',
                                                      'E N Myth',
 ('Gawr Gura',
                                                                      '2020-09-13'),
  ('Mori Calliope',
                                                     'E N Myth',
                                                                     '2020-09-12');
WITH ids AS (
 SELECT id FROM idols WHERE name_en IN ('Usada Pekora', 'Gawr Gura')
INSERT INTO streams (talent_id, title, viewers_k, stream_date) VALUES
 ((SELECT id FROM ids OFFSET O LIMIT 1),
   'Pekora peko-peko', 180, '2025-05-17'),
  ((SELECT id FROM ids OFFSET 1 LIMIT 1),
   'Guraaaaa', 194, '2025-07-1');
COMMIT;
```

```
postgres=# SELECT client_addr , state , sync_state FROM pg_stat_replication ; client_addr | state | sync_state | sync_stat
```

# 2.5 Установка и настройка PgBouncer (pg1)

```
sudo apt-get install pgbouncer -y
   /etc/pgbouncer/pgbouncer.ini
labdb = host=147.45.40.78,87.120.36.97 port=5432 dbname=labdb
[pgbouncer]
                   = 0.0.0.0
listen_addr
listen_port
                   = 6432
auth_type
                   = md5
auth_file
                   = /etc/pgbouncer/userlist.txt
pool_mode
                   = transaction
server_round_robin = 1
   /etc/pgbouncer/userlist.txt
"labuser" "labpass"
```

Запуск и автозапуск:

```
sudo systemctl enable --now pgbouncer
```

#### 2.6 Проверка балансировки

Client:

# 3 Этап 2. Симуляция и обработка сбоя

### 3.1 Client

```
#!/usr/bin/env bash
2
   # set -euo pipefail
   PG1_IP=147.45.40.78
4
   PG2_IP=87.120.36.97
5
   PGHOST = $PG1_IP
6
   PGPORT = 6432
   DB=labdb
   USER=labuser
   PASS=labpass
10
   DSN="host=$PGHOST port=$PGPORT dbname=$DB user=$USER password=$PASS"
11
   c_reset="\033[0m"
13
   c_red="\033[31m"]
14
   c_green="\033[32m"
15
   c_cyan="\033[36m"
16
   c_yellow="\033[33m"
17
18
              { printf '%(%F %T)T \033[32m[+]\033[0m %s\n' -1 "$*"; }
   log_ok()
19
   log_fail() { printf '%(%F %T)T \033[31m[ ]\033[0m %s\n' -1 "$*"; }
20
21
22
   color_ip() {
     case "$1" in
23
       $PG1_IP) printf "${c_cyan}%s${c_reset}" "$1" ;;
24
       $PG2_IP) printf "${c_yellow}%s${c_reset}" "$1" ;;
25
                printf "%s" "$1" ;;
       *)
26
     esac
27
   }
28
29
   get_primary_ip() {
30
     for ip in "$PG1_IP" "$PG2_IP"; do
31
       if [[ $(PGPASSWORD=$PASS psql -h "$ip" -U "$USER" -d "$DB" -At -c "SELECT
           pg_is_in_recovery();") == "f" ]]; then
         echo "$ip"
33
34
         return
       fi
35
     done
36
```

```
log_fail "No primary detected"
     exit 1
   }
39
40
   writer() {
41
     while true; do
42
       primary_ip=$(get_primary_ip)
43
       dsn="host=$primary_ip port=5432 dbname=$DB user=$USER password=$PASS"
44
45
       if out=$(psql "$dsn" -AtF',' \
46
            -c "INSERT INTO streams(talent_id,title,viewers_k,stream_date)
47
                   VALUES (1,'live workload',123,now())
48
                RETURNING id, inet_server_addr();" 2>&1); then
49
          IFS=',' read -r id ip <<<"$out"</pre>
50
         log_ok "writer : inserted id=$id backend=$(color_ip "$ip")"
51
       else
52
         log_fail "writer : $out"
53
       fi
54
       sleep 5
55
     done
56
   }
57
58
   reader() {
60
     tag=$1
     while true; do
61
       if out=$(psql "$DSN" -AtF',' \
62
             -c "SELECT count(*), inet_server_addr() FROM streams;" 2>&1); then
63
          IFS=',' read -r cnt ip <<<"$out"</pre>
64
         log_ok "$tag: rows=$cnt backend=$(color_ip "$ip")"
65
       else
66
         log_fail "$tag: $out"
67
       fi
68
69
       sleep 1
70
     done
71
   }
72
   schema_inspector() {
73
     \log_{-}ok "inspector : inspecting schema and privileges "
74
     psql "$DSN" -c "\z" 2>&1 | sed 's/^/inspector : /'
75
76
77
   schema_inspector
78
79
   writer &
80
81
   sleep 1
82
   for n in {1..10}; do
83
     reader "r$n" &
     sleep 1
84
   done
85
86
   wait
87
```

Workload.sh

#### 3.2 Crash & Recover

```
#!/usr/bin/env bash
set -euo pipefail

PG1=147.45.40.78
PG2=87.120.36.97
PGPORT=5432
BOUNCER_PORT=6432
BB=labdb
APPUSER=labuser
```

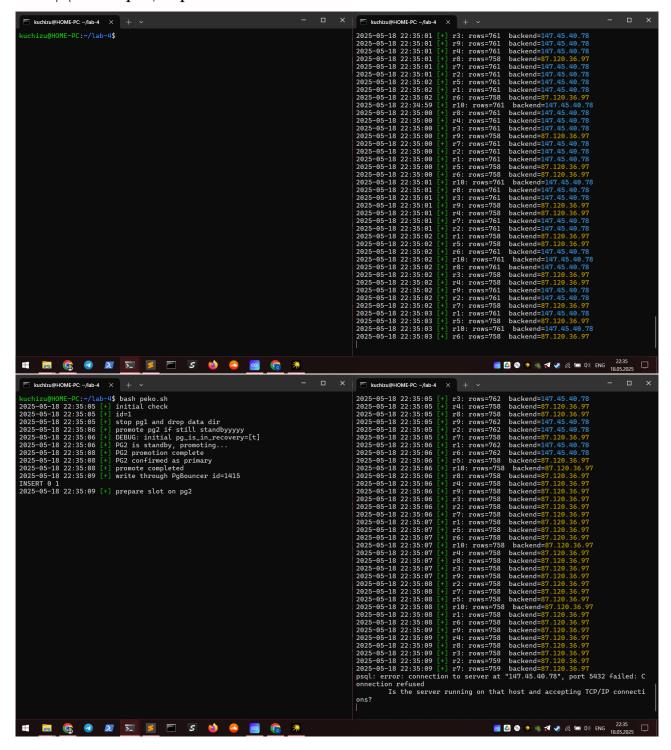
```
APPPASS=labpass
   REPLUSER=replicator
   REPLPASS=replpass
12
   DATA_DIR=/var/lib/postgresql/16/main
13
   PATH_ADDITION="/usr/bin:/usr/lib/postgresql/16/bin"
14
   slot="pg1_slot_$(date +%s)"
15
16
   g="\033[32m[+]\033[0m"
17
   r="\033[31m[ ]\033[0m"
18
        { printf '%(%F %T)T %b %s\n' -1 "$g" "$*"; }
19
   fail() { printf '%(%F %T)T %b %s\n' -1 "$r" "$*"; }
20
   # psql_b() {
22
       PGPASSWORD = "$APPPASS" psql "host = $PG1 port = $BOUNCER_PORT dbname = $DB user = $APPUSER
23
       " - qAt -c "$1"
24
25
   # psql_p2() {
26
      PGPASSWORD="$APPPASS" psql -h "$PG2" -p "$PGPORT" -U "$APPUSER" -d "$DB" -qAt -c
27
       "$1"
   # }
28
   psql_b(){ PGPASSWORD=$APPPASS psql "host=$PG1 port=$BOUNCER_PORT dbname=$DB user=
       $APPUSER" -At -c "$1"; }
31
32
   psq1_p2(){
     out=$(PGPASSWORD=$APPPASS psq1 -h "$PG2" -p "$PGPORT" -U "$APPUSER" -d "$DB" -At -c
33
          "$1") || {
       fail "psql_p2 error while running: $1"
34
       exit 1
35
     }
36
     ok "$out"
37
38
39
   psql_p2_bool() {
40
     PGPASSWORD=$APPPASS psql -h "$PG2" -p "$PGPORT" \
41
       -U "$APPUSER" -d "$DB" -At -c "SELECT pg_is_in_recovery();" \
42
     | awk '{print $1}'
43
44
45
   ssh1(){ ssh root@"$PG1" "export PATH=$PATH_ADDITION:\$PATH; $*"; }
46
   ssh2(){ ssh root@"$PG2" "export PATH=$PATH_ADDITION:\$PATH; $*"; }
47
48
   ok "initial check"
   row1=$(psql_b "SELECT id from streams limit 1") || { fail "write failed"; exit 1; }
50
   ok "id=$row1"
51
52
   # exit 0;
53
54
   ok "stop pg1 and drop data dir"
55
   ssh1 "systemctl stop postgresql && rm -rf $DATA_DIR"
56
57
   ok "promote pg2 if still standbyyyyy"
58
59
   # if [[ $(psql_p2 "SELECT pg_is_in_recovery();") == "t" ]]; then
60
       ssh2 \ "rm - f \ \$DATA\_DIR/standby.signal \ \&\& sudo - u \ postgres \ pg\_ctl \ promote - D
61
       $DATA_DIR"
       sleep 3
62
   # fi
63
64
   # [[ $(psql_p2 "SELECT pg_is_in_recovery()") == "f" ]] // { fail "pq2 promotion
65
       failed"; exit 1; }
66
   status=$(psql_p2_bool)
67
   ok "DEBUG: initial pg_is_in_recovery=[$status]" >&2
69
```

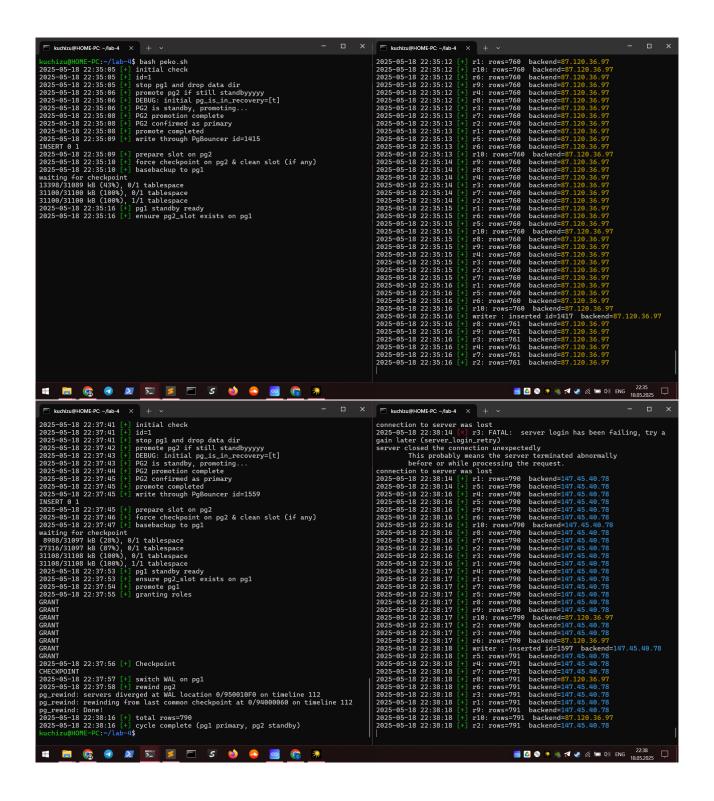
```
if [[ "$status" == "t" ]]; then
70
      ok "PG2 is standby, promoting..."
71
      ssh2 "pg_ctlcluster 16 main promote"
72
      until [[ "$(psql_p2_bool)" == "f" ]]; do
73
74
        sleep 1
75
      done
      ok "PG2 promotion complete"
76
77
78
    status=$(psql_p2_bool)
79
    if [[ "$status" != "f" ]]; then
80
      fail "PG2 promotion failed (still in recovery)"
81
82
    fi
    ok "PG2 confirmed as primary"
84
85
    ok "promote completed"
86
87
    row1=$(psql_b "INSERT INTO streams(talent_id,title,viewers_k,stream_date) VALUES (1,'
88
       row1',111,now()) RETURNING id;")
    ok "write through PgBouncer id=$row1"
89
    ok "prepare slot on pg2"
91
    ssh2 \ \ "sudo -u \ postgres \ psql - qAt \ 2 > /dev/null - c \ \ "SET \ client\_min\_messages = error;
       SELECT pg_drop_replication_slot('pg1_slot') WHERE EXISTS (SELECT 1 FROM
       pg_replication_slots WHERE slot_name='pg1_slot'); \""
93
    ok "force checkpoint on pg2 & clean slot (if any)"
94
    ssh2
95
      sudo -u postgres psql -d postgres -qAt -c \"
96
          SELECT pg_drop_replication_slot('$slot')
97
          WHERE EXISTS (SELECT 1 FROM pg_replication_slots WHERE slot_name = '$slot');
98
99
          CHECKPOINT;
100
102
    ok "basebackup to pg1"
103
    ssh1 "
104
      rm -rf $DATA_DIR &&
105
      mkdir -p $(dirname $DATA_DIR) &&
106
      chown -R postgres:postgres $(dirname $DATA_DIR)
107
108
      sudo -u postgres env PGPASSWORD=$REPLPASS pg_basebackup \
109
           -h $PG2 \
110
           -D $DATA_DIR \
111
           -U $REPLUSER \
112
           -Fp -Xs -P -R \
113
           -C -S $slot -w
114
115
      systemctl start postgresql &&
116
      for i in {1..60}; do sudo -u postgres pg_isready -q && exit 0; sleep 1; done; exit
117
          111
118
119
    [ $? -eq 0 ] || { fail "pg1 failed to start"; exit 1; }
120
    ok "pg1 standby ready"
122
123
    ok "ensure pg2_slot exists on pg1"
124
    ssh1 'sudo -u postgres psql -qAt -c "
125
      DO \$\$
126
127
        IF NOT EXISTS (SELECT 1 FROM pg_replication_slots WHERE slot_name = '\'', pg2_slot'
128
            \'') THEN
          PERFORM pg_create_physical_replication_slot('\''pg2_slot'\'');
129
        END IF;
```

```
END:
131
      \$\$;
132
133
134
    ok "promote pg1"
135
    ssh1 "pg_ctlcluster 16 main promote"
136
    # sleep 3
137
138
    ok "granting roles"
139
140
      sudo -u postgres psql -d postgres -c \"
141
        GRANT EXECUTE ON FUNCTION pg_read_binary_file(text)
                                                                                             TO
142
             $REPLUSER;
        GRANT EXECUTE ON FUNCTION pg_read_binary_file(text, boolean)
                                                                                             TΠ
143
             $REPLUSER:
        GRANT EXECUTE ON FUNCTION pg_read_binary_file(text, bigint, bigint)
                                                                                             TΩ
144
             $REPLUSER:
        GRANT EXECUTE ON FUNCTION pg_read_binary_file(text, bigint, bigint, boolean)
                                                                                             TO
145
             $REPLUSER;
        GRANT EXECUTE ON FUNCTION pg_ls_dir(text)
                                                                                             TO
146
             $REPLUSER:
        GRANT EXECUTE ON FUNCTION pg_ls_dir(text, boolean, boolean)
                                                                                             TO
             $REPLUSER;
                                                                                             ΤN
        GRANT EXECUTE ON FUNCTION pg_stat_file(text)
             $REPLUSER;
        GRANT EXECUTE ON FUNCTION pg_stat_file(text, boolean)
                                                                                             TΠ
149
             $REPLUSER;
      \"
150
151
152
    ok "Checkpoint"
153
    ssh1 "sudo -u postgres psql -c 'CHECKPOINT;'"
154
155
    ok "switch WAL on pg1"
156
    ssh1 "sudo -u postgres psql -q -c 'CHECKPOINT; SELECT pg_switch_wal();' > /dev/null
157
       2>&1"
158
    ok "rewind pg2"
159
    ssh2 "
160
      systemctl stop postgresql
161
      sudo -u postgres env PGPASSWORD=$REPLPASS /usr/lib/postgresql/16/bin/pg_rewind \
162
        --target-pgdata=$DATA_DIR \
163
        --source-server=\"host=$PG1 port=$PGPORT user=$REPLUSER password=$REPLPASS dbname
164
            =postgres\"
      echo \"primary_conninfo = 'host=$PG1 port=$PGPORT user=$REPLUSER password=$REPLPASS
          '\" > $DATA_DIR/postgresql.auto.conf
      echo \"primary_slot_name = 'pg2_slot'\" >> $DATA_DIR/postgresql.auto.conf
166
      touch $DATA_DIR/standby.signal
167
      systemctl start postgresql
168
169
170
    sleep 7
171
172
    \# row3=\$(psql_b \ "INSERT \ INTO \ streams(talent_id,title,viewers_k,stream_date) \ VALUES
173
       (1, 'row3', 3, now()) RETURNING id") | { fail "final write failed"; exit 1; }
    total=$(psql_b "SELECT count(*) FROM streams")
174
    # ok "final write id=$row3"
176
    ok "total rows=$total"
177
    ok "cycle complete (pg1 primary, pg2 standby)"
178
```

Peko.sh

### 3.3 Демонстрация работы





# 4 Вывод

В ходе выполнения работы была реализована автоматизированная система отказоустойчивости PostgreSQL с использованием двух узлов (primary/standby), PgBouncer и скрипта, обеспечивающего полный цикл переключения ролей и восстановления репликации.(((((