

Федеральное государственное автономное
образовательное учреждение высшего образования
«Научно-образовательная корпорация ИТМО»

Факультет программной инженерии и компьютерной техники
Направление подготовки 09.03.04 Программная инженерия

Отчёт по лабораторной работе №3

По дисциплине «Распределённые системы хранения данных» (семестр 6)

Студент:

Дениченко Александр Р3312

Практик:

Осипов Святослав

Санкт-Петербург
2025 г.

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.

Этап 1. Резервное копирование

Настроить резервное копирование с основного узла на резервный следующим образом: Периодические полные копии с помощью SQL Dump. По расписанию (cron) раз в сутки, методом SQL Dump с сжатием. Созданные архивы должны сразу перемещаться на резервный хост, они не должны храниться на основной системе. Срок хранения архивов на резервной системе - 4 недели. По истечении срока хранения, старые архивы должны автоматически уничтожаться.

Изначально сделаем донастройку конфигураций прежней базы данных.

Листинг 1: kitty

```
1 docker create --name postgres-cont-1 -e POSTGRES_PASSWORD=root -p 9193:9193 postgres
```

Листинг 2: kitty

```
1 docker start postgres-cont-1
2 docker exec -it postgres-cont-1 /bin/bash
```

Изменение некоторых настроек разрешений

Листинг 3: postgresql.conf

```
1 listen_addresses = '*'
```

Листинг 4: pg_hba.conf

```
1 host      all      all      0.0.0.0/0      scram-sha-256
```

Подключение к первому узлу теперь выглядит следующим образом.

Листинг 5: kitty

```
1 psql -h 127.0.0.1 -p 9193 -U postgres -d postgres
2 psql -h 127.0.0.1 -p 9193 -U postgres -d fatrednews
3 psql -h 127.0.0.1 -p 9193 -U fatreduser -d fatrednews
```

postgres - root
fatreduser - changeMe

Табличные пространства из прошлой лабы

Листинг 6: kitty

| List of tablespaces | | | | | | |
|---------------------|----------|---------------------------|-------------------|---------|-------|--|
| Name | Owner | Location | Access privileges | Options | Size | |
| Description | | | | | | |
| pg_default | postgres | | | | 22 MB | |
| pg_global | postgres | | | | 565 | |
| sgk31 | postgres | /var/lib/postgresql/sgk31 | | | 0 | |
| bytes | | | | | | |

```

7   yrp30      | postgres | /var/lib/postgresql/yrp30 |      | 0
   bytes |
8   yva58      | postgres | /var/lib/postgresql/yva58 |      | 0
   bytes |
9   (5 rows)

```

Создан узел для хранения бэкапов

Листинг 7: kitty

```

1   docker create --name postgres-backup -p 9194:9193 ubuntu:latest tail -f /dev/null

```

Добавлены утилиты на резервном узле и на основном

Листинг 8: kitty

```

1   apt-get install -y cron openssh-client openssh-server gzip

```

Была сделана сеть докер для обмена

Листинг 9: kitty

```

1   docker network create --driver bridge postgres-backup-net
2   docker network connect postgres-backup-net postgres-cont-1 postgres-backup
3   docker restart postgres-cont-1 postgres-backup

```

Добавим конфиг в сервер

Листинг 10: kitty

```

1   echo "PermitRootLogin yes" >> /etc/ssh/sshd_config
2   chmod 700 /root/.ssh
3   mkdir -p /run/sshd
4   chmod 755 /run/sshd
5   service ssh start

```

Добавлены ключи для упрощения обмена данными на основном сервере

Листинг 11: kitty

```

1   ssh-keygen -t rsa -b 4096
2   cat ~/.ssh/id_rsa.pub

```

И добавили ключи в резервный узел

Листинг 12: kitty

```

1   vim ~/.ssh/authorized_keys ....

```

Настройка скрипта для копирования

Листинг 13: kitty

```

1  #!/bin/bash
2
3  mkdir -p /tmp/backups
4  mkdir -p "/tmp/backups/backup_$(date +%Y%m%d_%H%M%S)"
5
6  pg_dumpall -p 9193 -U postgres > "/tmp/backups/backup_$(date
   +%Y%m%d_%H%M%S)/full_backup.sql"
7
8  cp "/var/lib/postgresql/data/postgresql.conf" "/tmp/backups/backup_$(date +%Y%m%d_%H%M%S)/"
9  cp "/var/lib/postgresql/data/pg_hba.conf" "/tmp/backups/backup_$(date +%Y%m%d_%H%M%S)/"
10
11 mkdir -p "/tmp/backups/backup_$(date +%Y%m%d_%H%M%S)/pg_tblspc"

```

```

12 cp -r "/var/lib/postgresql/data/pg_tblspc/"* "/tmp/backups/backup_$(date
    +%Y%m%d_%H%M%S)/pg_tblspc/"
13
14 for tblspc in "/var/lib/postgresql/data/pg_tblspc/"*; do
15     if [ -L "$tblspc" ]; then
16         mkdir -p "/tmp/backups/backup_$(date +%Y%m%d_%H%M%S)/tblspc_data/$(basename
            "$tblspc")"
17         cp -r "$(readlink -f "$tblspc")/"* "/tmp/backups/backup_$(date
            +%Y%m%d_%H%M%S)/tblspc_data/$(basename "$tblspc")/"
18     fi
19 done
20
21 tar -czf "/tmp/backups/full_backup-$(date +%Y%m%d_%H%M%S).tar.gz" -C
    "/tmp/backups/backup_$(date +%Y%m%d_%H%M%S)" .
22
23 scp "/tmp/backups/full_backup-$(date +%Y%m%d_%H%M%S).tar.gz" "postgres-backup:/backups/"
24
25 rm -rf "/tmp/backups/backup_$(date +%Y%m%d_%H%M%S)"
26 rm "/tmp/backups/full_backup-$(date +%Y%m%d_%H%M%S).tar.gz"

```

Сама настройка для автоматизации на основном хосте

Листинг 14: kitty

```

1 crontab -e
2 */1 * * * * /backup.sh
3 cron

```

Настройка для автоматизации на доп хосте

Листинг 15: kitty

```

1 crontab -e
2 */2 * * * * /cleanup_dumps.sh
3 cron

```

Расчет объема резервных копий через месяц

Исходные данные:

- Средний объем новых данных в БД за сутки: 950 МБ.
- Средний объем измененных данных за сутки: 150 МБ.
- Период хранения резервных копий: 4 недели (28 дней).

Предположения:

1. Бэкап делается полностью (включает все данные, а не только новые или измененные).
2. Измененные данные не увеличивают общий объем бэкапа, так как они уже входят в полную копию.
3. Объем БД увеличивается со временем за счет новых данных.

Каждый день создается новая полная резервная копия. Объем бэкапа на n-й день равен всему объему базы данных на тот момент.

Объем базы через n дней можно выразить как:

$$V(n) = V_0 + 950 \times n$$

Где:

- V_0 — начальный объем базы (пусть 0 для расчета за 1 месяц),
- 950 — рост базы в день.

Общий объем всех бэкапов за 28 дней:

$$V_{\text{total}} = \sum_{n=1}^{28} (950 \times n)$$

—
Рассчитаем сумму:

$$V_{\text{total}} = 950 \times (1 + 2 + \dots + 28)$$

Сумма арифметической прогрессии:

$$S = \frac{n(n+1)}{2}$$

Где $n = 28$:

$$S = \frac{28 \times 29}{2} = 406$$

Подставляем:

$$V_{\text{total}} = 950 \times 406 = 385700 \text{ МБ} = 385.7 \text{ ГБ}$$

Потеря основного узла

Листинг 16: kitty

```
1 #!/bin/bash
2
3 PG_USER="postgres"
4 BACKUP_DIR="/backups"
5 RESTORE_TEMP_DIR="/tmp/restore_temp"
6 PGDATA="/var/lib/postgresql/data"
7 LOG_FILE="/tmp/restore_log.txt"
8 BACKUP_FILE="$1"
9
10 if [ -z "$BACKUP_FILE" ]; then
11     echo "Error: Backup file not specified. Usage: $0
12         /backups/full_backup-YYYYMMDD_HHMMSS.tar.gz" >> $LOG_FILE
13     exit 1
14 fi
15
16 echo "Starting restore: $(date +%Y%m%d_%H%M%S)" >> $LOG_FILE
17
18 pg_ctl stop -D "$PGDATA/"
19
20 rm -rf "$PGDATA"/* || {
21     echo "Failed to clean PGDATA" >> $LOG_FILE
22     exit 1
23 }
24
25 initdb -D "$PGDATA"
26 pg_ctl start
27
28 mkdir -p "$RESTORE_TEMP_DIR"
29 gunzip -c "$BACKUP_FILE" > "$RESTORE_TEMP_DIR/${basename "$BACKUP_FILE" .gz}" || {
30     echo "Failed to extract backup" >> "$LOG_FILE"
31     exit 1
32 }
33
34 for tblspc in "$RESTORE_TEMP_DIR/pg_tblspc"/*; do
35     if [ -f "$tblspc" ]; then
36         TBL_ID=$(basename "$tblspc")
37         TBL_PATH=$(readlink "$tblspc")
```

```

37     mkdir -p "$TBL_PATH" || echo "Failed to create tablespace dir $TBL_PATH" >>
        $LOG_FILE
38     ln -s "$TBL_PATH" "$PGDATA/pg_tblspc/$TBL_ID" || echo "Failed to link tablespace
        $TBL_ID" >> $LOG_FILE
39 fi
40 done
41
42 if [ -d "$RESTORE_TEMP_DIR/tblspc_data" ]; then
43     for tblspc in "$RESTORE_TEMP_DIR/tblspc_data/"*; do
44         TBL_ID=$(basename "$tblspc")
45         TBL_PATH=$(readlink "$RESTORE_TEMP_DIR/pg_tblspc/$TBL_ID")
46         cp -r "$tblspc/"* "$TBL_PATH/" || echo "Failed to restore tablespace data for
            $TBL_ID" >> $LOG_FILE
47     done
48 fi
49
50 cp "$RESTORE_TEMP_DIR/postgresql.conf" "$PGDATA/" || echo "Failed to restore
    postgresql.conf" >> $LOG_FILE
51 cp "$RESTORE_TEMP_DIR/pg_hba.conf" "$PGDATA/" || echo "Failed to restore pg_hba.conf" >>
    $LOG_FILE
52
53
54 psql -U $PG_USER -f "$RESTORE_TEMP_DIR/full_backup.sql" || {
55     echo "Failed to restore database" >> $LOG_FILE
56     exit 1
57 }
58
59 pg_ctl restart
60
61 rm -rf "$RESTORE_TEMP_DIR"
62
63 echo "Restore completed successfully" >> $LOG_FILE

```

Пример восстановления базы данных

Запуск резервного узла

Листинг 17: kitty

```

1 docker exec -it postgres-backup /bin/bash

```

Должно быть включено ssh и общая есть у двух контейнеров.

Листинг 18: kitty

```

1 service ssh start

```

Делаем типом последний бэкап с основного узла перед его выходом из строя

Листинг 19: kitty

```

1 root@7f898d3520a8:/# ./backup.sh
2 full_backup-20250401_190056.tar.gz          100%   14KB   30.0MB/s   00:00

```