PostgreSQL failover and bring an old primary server as a standby:

OS: RHEL 9

PostgreSQL version: 15.7

Promote the Standby Server as a Primary:

Current Primary Server IP: 172.31.28.62

Current Standby server IP: 172.31.30.125

On Current Standby server:

Step1) Verify Standby Synchronization with below query

#Recovery Control Functions

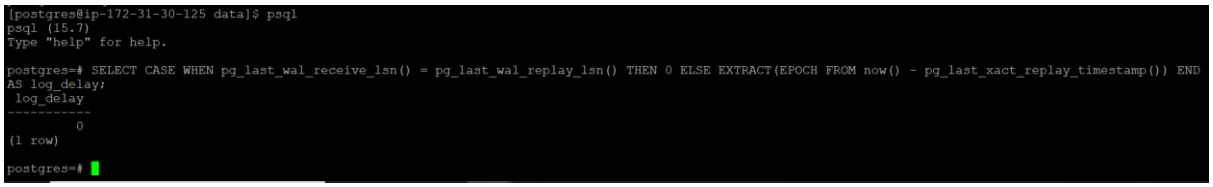
pg_last_wal_receive_lsn () --> Returns the last write-ahead log location that has been received and synced to disk by streaming replication.

pg_last_wal_replay_lsn () --> Returns the last write-ahead log location that has been replayed during recovery.

pg_last_xact_replay_timestamp () --> Returns the time stamp of the last transaction replayed during recovery

#Query

SELECT CASE WHEN pg_last_wal_receive_lsn() = pg_last_wal_replay_lsn() THEN 0 ELSE EXTRACT(EPOCH FROM now() - pg_last_xact_replay_timestamp()) END AS log_delay;



Step2) Check and create archive directory on standby server if doesn't exist

cat $PGDATA/postgresql.conf |grep archive_command

ls -ltrh /mnt/server/archivedir/

On Current Primary server:

Step3) Shutdown the PostgreSQL service on current primary server

sudo systemctl stop postgresql-15

sudo systemctl status postgresql-15

```
[root@ip-172-31-28-62 ec2-user]# sudo systemctl stop postgresql-15
[root@ip-172-31-28-62 ec2-user]# sudo systemctl status postgresql-15
● postgresql-15.service - PostgreSQL 15 database server
     Loaded: loaded (/usr/lib/systemd/system/postgresql-15.service; enabled; preset: disabled)
     Active: inactive (dead) since Fri 2024-06-14 10:59:49 UTC; 6min ago
   Duration: 27min 43.118s
       Docs: https://www.postgresql.org/docs/15/static/
    Process: 1803 ExecStartPre=/usr/pgsql-15/bin/postgresql-15-check-db-dir ${PGDATA} (code=exited, status=0/SUCCESS)
    Process: 1809 ExecStart=/usr/pgsql-15/bin/postmaster -D ${PGDATA} (code=exited, status=0/SUCCESS)
   Main PID: 1809 (code=exited, status=0/SUCCESS)
        CPU: 529ms

Jun 14 10:32:06 ip-172-31-28-62.ec2.internal systemd[1]: Starting PostgreSQL 15 database server...
Jun 14 10:32:06 ip-172-31-28-62.ec2.internal postmaster[1809]: 2024-06-14 10:32:06.212 UTC [1809] LOG:  redirecting log output to logging collector process
Jun 14 10:32:06 ip-172-31-28-62.ec2.internal postmaster[1809]: 2024-06-14 10:32:06.212 UTC [1809] HINT:  Future log output will appear in directory "log".
Jun 14 10:32:06 ip-172-31-28-62.ec2.internal systemd[1]: Started PostgreSQL 15 database server.
Jun 14 10:59:49 ip-172-31-28-62.ec2.internal systemd[1]: Stopping PostgreSQL 15 database server...
Jun 14 10:59:49 ip-172-31-28-62.ec2.internal systemd[1]: postgresql-15.service: Deactivated successfully.
Jun 14 10:59:49 ip-172-31-28-62.ec2.internal systemd[1]: Stopped PostgreSQL 15 database server.
[root@ip-172-31-28-62 ec2-user]#
```

On Current Standby server:

Step4) Promote the Standby to Read-Write mode

pg_promote() --> end standby mode and begin read-write operations.

pg_is_in_recovery() --> Returns true if recovery is still in progress.

SELECT pg_promote();

SELECT pg_is_in_recovery();

```
[postgres@ip-172-31-30-125 data]$ psql
psql (15.7)
Type "help" for help.

postgres=# SELECT pg_is_in_recovery();
 pg_is_in_recovery
-------------------
 t
(1 row)

postgres=# SELECT pg_promote();
 pg_promote
------------
 t
(1 row)

postgres=# SELECT pg_is_in_recovery();
 pg_is_in_recovery
-------------------
 f
(1 row)

postgres=#
```

#Check the PostgreSQL logs

tail -f $PGDATA/log/postgresql-Sat.log

```
[postgres@ip-172-31-30-125 ~]$ tail -f $PGDATA/log/postgresql-Sat.log
2024-06-15 06:14:55.507 UTC [1673] FATAL:  could not connect to the primary server: connection to server at "172.31.28.62", port 5432 failed: Connection refu
sed
        Is the server running on that host and accepting TCP/IP connections?
2024-06-15 06:14:55.507 UTC [1592] LOG:  waiting for WAL to become available at 0/4000018
2024-06-15 06:14:55.989 UTC [1592] LOG:  received promote request
2024-06-15 06:14:55.989 UTC [1592] LOG:  redo done at 0/3000148 system usage: CPU: user: 0.00 s, system: 0.00 s, elapsed: 387.69 s
2024-06-15 06:14:55.990 UTC [1592] LOG:  selected new timeline ID: 2
2024-06-15 06:14:56.054 UTC [1592] LOG:  archive recovery complete
2024-06-15 06:14:56.060 UTC [1590] LOG:  checkpoint starting: force
2024-06-15 06:14:56.063 UTC [1588] LOG:  database system is ready to accept connections
2024-06-15 06:14:56.069 UTC [1590] LOG:  checkpoint complete: wrote 2 buffers (0.0%); 0 WAL file(s) added, 0 removed, 0 recycled; write=0.001 s, sync=0.003 s
, total=0.009 s; sync files=2, longest=0.003 s, average=0.002 s; distance=16383 kB, estimate=16384 kB
```

#To verify insert some data into the table

INSERT INTO categories (category_name, description)

VALUES

('Produce', 'Dried fruit and bean curd'),

('Seafood', 'Seaweed and fish');

```
[postgres@ip-172-31-30-125 ~]$ psql
psql (15.7)
Type "help" for help.

postgres=# SELECT * FROM categories;
 category_id | category_name  |                        description
-------------+----------------+-------------------------------------------------------------
           1 | Beverages      | Soft drinks, coffees, teas, beers, and ales
           2 | Condiments     | Sweet and savory sauces, relishes, spreads, and seasonings
           3 | Confections    | Desserts, candies, and sweet breads
           4 | Dairy Products | Cheeses
           5 | Grains/Cereals | Breads, crackers, pasta, and cereal
           6 | Meat/Poultry   | Prepared meats
(6 rows)

postgres=# INSERT INTO categories (category_name, description)
VALUES
  ('Produce', 'Dried fruit and bean curd'),
  ('Seafood', 'Seaweed and fish');
INSERT 0 2
postgres=# SELECT * FROM categories;
 category_id | category_name  |                        description
-------------+----------------+-------------------------------------------------------------
           1 | Beverages      | Soft drinks, coffees, teas, beers, and ales
           2 | Condiments     | Sweet and savory sauces, relishes, spreads, and seasonings
           3 | Confections    | Desserts, candies, and sweet breads
           4 | Dairy Products | Cheeses
           5 | Grains/Cereals | Breads, crackers, pasta, and cereal
           6 | Meat/Poultry   | Prepared meats
          38 | Produce        | Dried fruit and bean curd
          39 | Seafood        | Seaweed and fish
(8 rows)

postgres=#
```

Bring an old primary server back online as a standby:

New Primary Server IP: 172.31.30.125

Old Primary server IP: 172.31.28.62

On New Primary server:

Step5) Allow the remote connections between the servers

pg_reload_conf() --> Cause server processes to reload their configuration files

#Edit $PGDATA/pg_hba.conf file

#normal user connection

host all postgres 172.31.28.62/32 scram-sha-256

#replication user connection

host replication rep_user 172.31.28.62/32 scram-sha-256

save&exit

```
[postgres@ip-172-31-30-125 ~]$ vim $PGDATA/pg_hba.conf
[postgres@ip-172-31-30-125 ~]$ cat $PGDATA/pg_hba.conf |grep "172.31.28.62"
host    all             postgres            172.31.28.62/32      scram-sha-256
host    replication     rep_user            172.31.28.62/32      scram-sha-256
[postgres@ip-172-31-30-125 ~]$ psql
psql (15.7)
Type "help" for help.

postgres=# SELECT pg_reload_conf();
 pg_reload_conf
----------------
 t
(1 row)

postgres=#
```

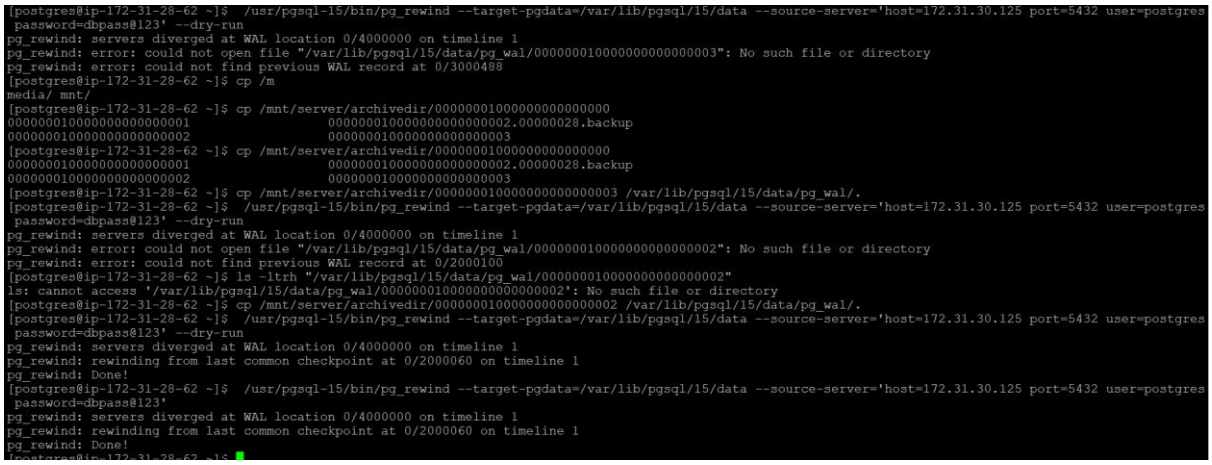
Step6) Use pg_rewind command to sync the New and old primary server clusters

pg_rewind is a tool for synchronizing a PostgreSQL cluster with another copy of the same cluster, after the clusters' timelines have diverged. A typical scenario is to bring an old primary server back online after failover as a standby that follows the new primary

/usr/pgsql-15/bin/pg_rewind --target-pgdata=/var/lib/pgsql/15/data --source-server='host=172.31.30.125 port=5432 user=postgres password=dbpass@123' --dry-run

/usr/pgsql-15/bin/pg_rewind --target-pgdata=/var/lib/pgsql/15/data --source-server='host=172.31.30.125 port=5432 user=postgres password=dbpass@123'

Note: If you get an error about could not found WAL records, you can copy manually wal file from archive logs directory to old primary database's $PGDATA/pg_wal directory.

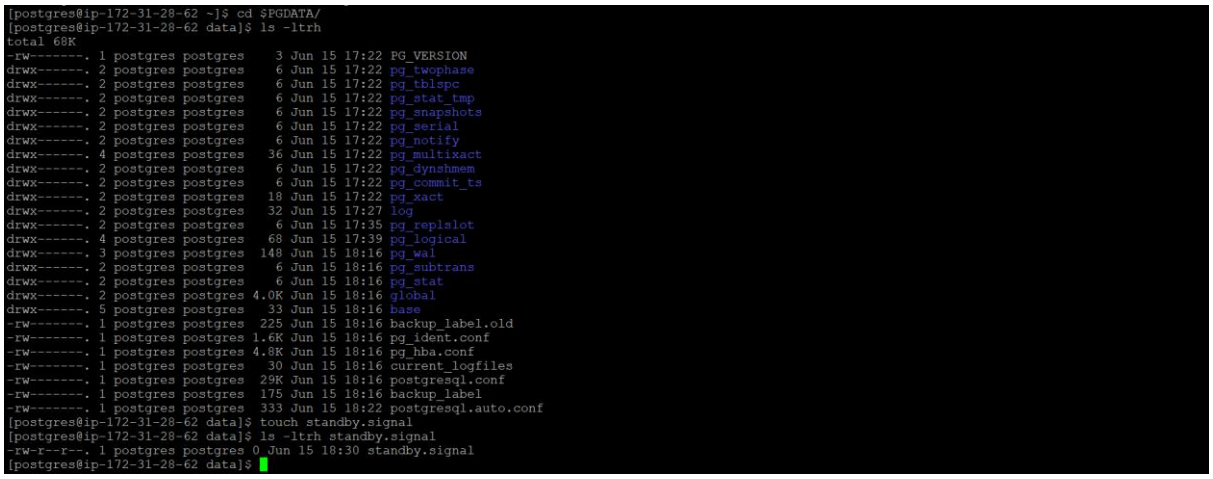


Step7) Create standby.signal file in old primary server data directory

A server enters standby mode if a "**standby.signal**" file exists in the data directory when the server is started.

cd $PGDATA/

touch standby.signal



Step8) Add new primary server IP address in a $PGDATA/postgresql.auto.conf file

vim $PGDATA/postgresql.auto.conf

primary_conninfo = 'user=rep_user password=''rep@123'' channel_binding=prefer host=172.31.30.125 port=5432 sslmode=prefer sslcompression=0 sslsni=1 ssl_min_protocol_version=TLSv1.2 gssencmode=prefer krbsrvname=postgres target_session_attrs=any'

save&exit

```
[postgres@ip-172-31-28-62 ~]$ vim $PGDATA/postgresql.auto.conf
[postgres@ip-172-31-28-62 ~]$ cat $PGDATA/postgresql.auto.conf
# Do not edit this file manually!
# It will be overwritten by the ALTER SYSTEM command.
primary_conninfo = 'user=rep_user password=''rep@123'' channel_binding=prefer host=172.31.30.125 port=5432 sslmode=prefer sslcompression=0 sslsni=1 ssl_min_p
rotocol_version=TLSv1.2 gssencmode=prefer krbsrvname=postgres target_session_attrs=any'
[postgres@ip-172-31-28-62 ~]$
```

Step9) Configure below parameters in a $PGDATA/postgresql.conf

vim $PGDATA/postgresql.conf

restore_command = 'cp /mnt/server/archivedir/%f %p'

recovery_target_timeline = 'latest'

save&exit

```
[postgres@ip-172-31-28-62 ~]$ vim $PGDATA/postgresql.conf
[postgres@ip-172-31-28-62 ~]$ cat $PGDATA/postgresql.conf |grep -iaE "restore_command|recovery_target_timeline"
restore_command = 'cp /mnt/server/archivedir/%f %p'           # command to use to restore an archived logfile segment
recovery_target_timeline = 'latest'     # 'current', 'latest', or timeline ID
[postgres@ip-172-31-28-62 ~]$
```

Step10) Start and check status of the postgresql service

sudo systemctl start postgresql-15

sudo systemctl status postgresql-15

```
[root@ip-172-31-28-62 ec2-user]# sudo systemctl status postgresql-15
● postgresql-15.service - PostgreSQL 15 database server
     Loaded: loaded (/usr/lib/systemd/system/postgresql-15.service; enabled; preset: disabled)
     Active: active (running) since Sat 2024-06-15 18:35:30 UTC; 8s ago
       Docs: https://www.postgresql.org/docs/15/static/
    Process: 2438 ExecStartPre=/usr/pgsql-15/bin/postgresql-15-check-db-dir ${PGDATA} (code=exited, status=0/SUCCESS)
   Main PID: 2443 (postmaster)
      Tasks: 6 (limit: 4400)
     Memory: 15.7M
        CPU: 62ms
     CGroup: /system.slice/postgresql-15.service
             ├─2443 /usr/pgsql-15/bin/postmaster -D /var/lib/pgsql/15/data/
             ├─2444 "postgres: logger "
             ├─2445 "postgres: checkpointer "
             ├─2446 "postgres: background writer "
             ├─2447 "postgres: startup recovering 000000020000000000000004"
             └─2453 "postgres: walreceiver "

Jun 15 18:35:30 ip-172-31-28-62.ec2.internal systemd[1]: Starting PostgreSQL 15 database server...
Jun 15 18:35:30 ip-172-31-28-62.ec2.internal postmaster[2443]: 2024-06-15 18:35:30.956 UTC [2443] LOG:  redirecting log output to logging collector process
Jun 15 18:35:30 ip-172-31-28-62.ec2.internal postmaster[2443]: 2024-06-15 18:35:30.956 UTC [2443] HINT:  Future log output will appear in directory "log".
Jun 15 18:35:30 ip-172-31-28-62.ec2.internal systemd[1]: Started PostgreSQL 15 database server.
```

#Check the logs

tail -f $PGDATA/log/postgresql-Sat.log

```
2024-06-15 18:35:30.995 UTC [2447] LOG:  completed backup recovery with redo LSN 0/2000028 and end LSN 0/2000100
2024-06-15 18:35:30.995 UTC [2447] LOG:  consistent recovery state reached at 0/4001140
2024-06-15 18:35:30.995 UTC [2447] LOG:  invalid record length at 0/4001140: wanted 24, got 0
2024-06-15 18:35:30.995 UTC [2443] LOG:  database system is ready to accept read-only connections
2024-06-15 18:35:31.020 UTC [2453] LOG:  started streaming WAL from primary at 0/4000000 on timeline 2
```

Check the replication status:

Current Primary Server IP: 172.31.30.125

Current Standby server IP: 172.31.28.62

On Current Primary:

SELECT * FROM pg_stat_replication;

```
[postgres@ip-172-31-30-125 pg_wal]$ psql
psql (15.7)
Type "help" for help.

postgres=# \x
Expanded display is on.
postgres=# SELECT * FROM pg_stat_replication;
-[ RECORD 1 ]----+------------------------------
pid              | 2222
usesysid         | 16397
usename          | rep_user
application_name | walreceiver
client_addr      | 172.31.28.62
client_hostname  |
client_port      | 53440
backend_start    | 2024-06-15 18:35:31.00628+00
backend_xmin     |
state            | streaming
sent_lsn         | 0/4001140
write_lsn        | 0/4001140
flush_lsn        | 0/4001140
replay_lsn       | 0/4001140
write_lag        |
flush_lag        |
replay_lag       |
sync_priority    | 0
sync_state       | async
reply_time       | 2024-06-15 18:48:12.756185+00

postgres=#
```

CREATE TABLE customers (

customer_id SERIAL NOT NULL PRIMARY KEY,

customer_name VARCHAR(255),

contact_name VARCHAR(255),

address VARCHAR(255),

city VARCHAR(255),

postal_code VARCHAR(255),

country VARCHAR(255)

);

INSERT INTO customers (customer_name, contact_name, address, city, postal_code, country)

VALUES

('Alfreds Futterkiste', 'Maria Anders', 'Obere Str. 57', 'Berlin', '12209', 'Germany'),

('Ana Trujillo Emparedados y helados', 'Ana Trujillo', 'Avda. de la Constitucion 2222', 'Mexico D.F.', '05021', 'Mexico'),

('Antonio Moreno Taquera', 'Antonio Moreno', 'Mataderos 2312', 'Mexico D.F.', '05023', 'Mexico'),

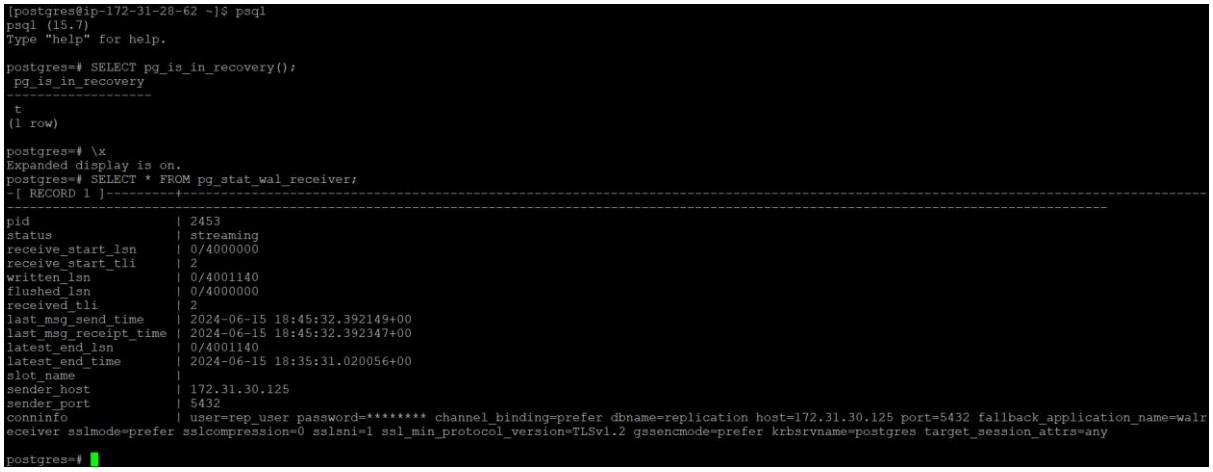
('Around the Horn', 'Thomas Hardy', '120 Hanover Sq.', 'London', 'WA1 1DP', 'UK')

```
[postgres@ip-172-31-30-125 pg_wal]$ psql
psql (15.7)
Type "help" for help.

postgres=# CREATE TABLE customers (
  customer_id SERIAL NOT NULL PRIMARY KEY,
  customer_name VARCHAR(255),
  contact_name VARCHAR(255),
  address VARCHAR(255),
  city VARCHAR(255),
  postal_code VARCHAR(255),
  country VARCHAR(255)
);
CREATE TABLE
postgres=# INSERT INTO customers (customer_name, contact_name, address, city, postal_code, country)
VALUES
  ('Alfreds Futterkiste', 'Maria Anders', 'Obere Str. 57', 'Berlin', '12209', 'Germany'),
  ('Ana Trujillo Emparedados y helados', 'Ana Trujillo', 'Avda. de la Constitucion 2222', 'Mexico D.F.', '05021', 'Mexico'),
  ('Antonio Moreno Taquera', 'Antonio Moreno', 'Mataderos 2312', 'Mexico D.F.', '05023', 'Mexico'),
  ('Around the Horn', 'Thomas Hardy', '120 Hanover Sq.', 'London', 'WA1 1DP', 'UK');
INSERT 0 4
postgres=# SELECT * FROM customers;
 customer_id |           customer_name            |  contact_name  |            address            |    city     | postal_code | country
-------------+------------------------------------+----------------+-------------------------------+-------------+-------------+---------
           1 | Alfreds Futterkiste                | Maria Anders   | Obere Str. 57                 | Berlin      | 12209       | Germany
           2 | Ana Trujillo Emparedados y helados | Ana Trujillo   | Avda. de la Constitucion 2222 | Mexico D.F. | 05021       | Mexico
           3 | Antonio Moreno Taquera             | Antonio Moreno | Mataderos 2312                | Mexico D.F. | 05023       | Mexico
           4 | Around the Horn                    | Thomas Hardy   | 120 Hanover Sq.               | London      | WA1 1DP     | UK
(4 rows)

postgres=#
```

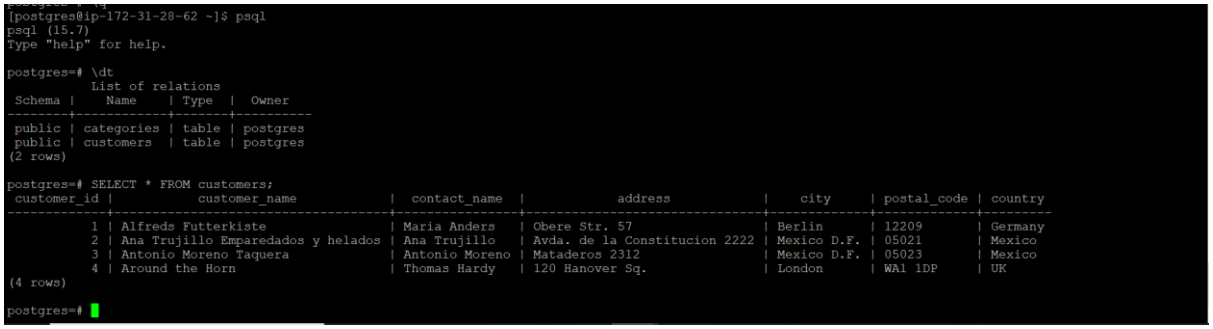
On current standby:

SELECT pg_is_in_recovery();

SELECT * FROM pg_stat_wal_receiver;



SELECT * FROM customers;



Ref:

https://www.postgresql.org/docs/current/app-pgrewind.html#:~:text=pg_rewind%20is%20a%20tool%20for,that%20follows%20the%20new%20primary.

https://www.2ndquadrant.com/en/blog/introduction-to-pgrewind/

https://www.fatihacar.com/blog/how-to-use-pg_rewind-in-postgresql-12/

https://www.postgresql.org/docs/current/warm-standby-failover.html