pg_paxos

The Paxos Algorithm

In a group of servers, the paxos(k,v) function reaches consensus on a value for key k.

All nodes in the group will see the same output value for the same k and it will be one of the input values.

Can tolerate a minority of nodes failing.

The Paxos Algorithm

paxos(k,v):

1. Ask the majority of servers to grant you the (preemptable) lock and return any existing value for key k.
   If a value exist, v becomes the most recent existing value.
2. Ask the servers from round 1 to accept value v for key k, return v

If in either round you cannot get a majority, restart.

The Multi-Paxos Algorithm

Each node maintains a log of changes, use Paxos to reach consensus on change.

To write a value to the distributed log at position i:

```
while(paxos(i,query) != query) i++;
```

To confirm consensus on a round i:

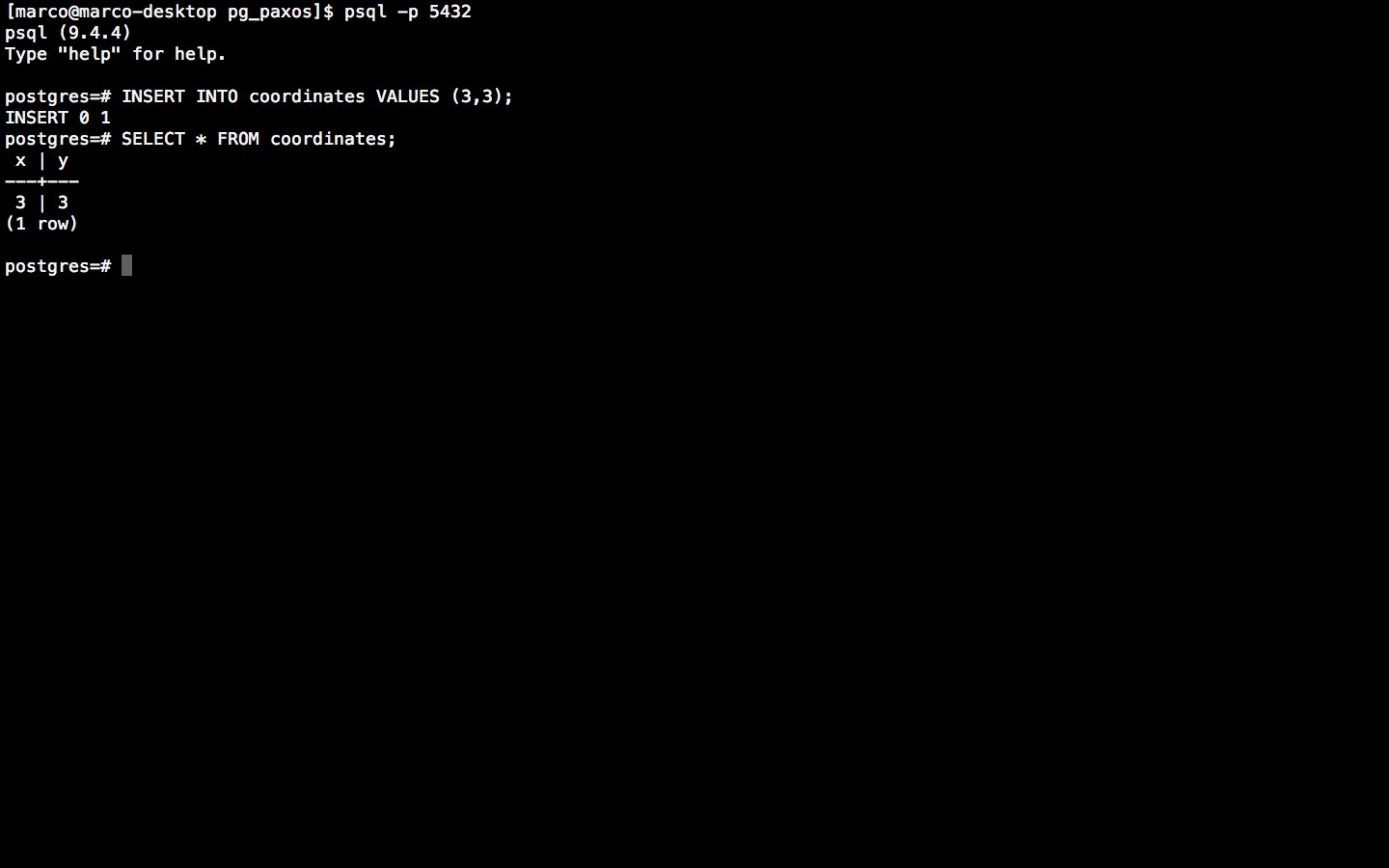
```
paxos(i,'');
```

pg_paxos

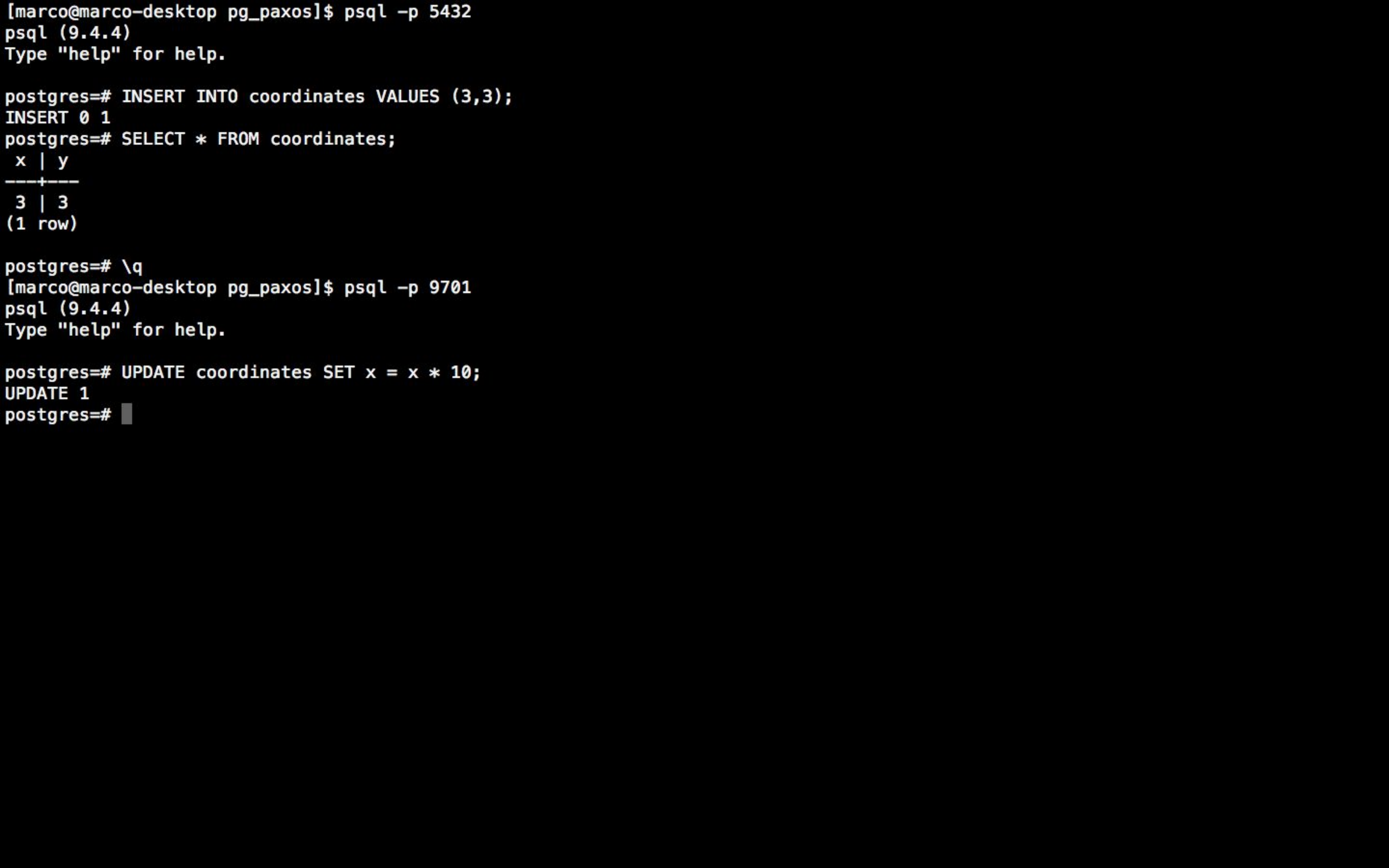
Extension for PostgreSQL

1. Basic implementation of Paxos and Multi-Paxos in PL/pgSQL (~1000 lines)
2. Consistent table replication implemented using Multi-Paxos

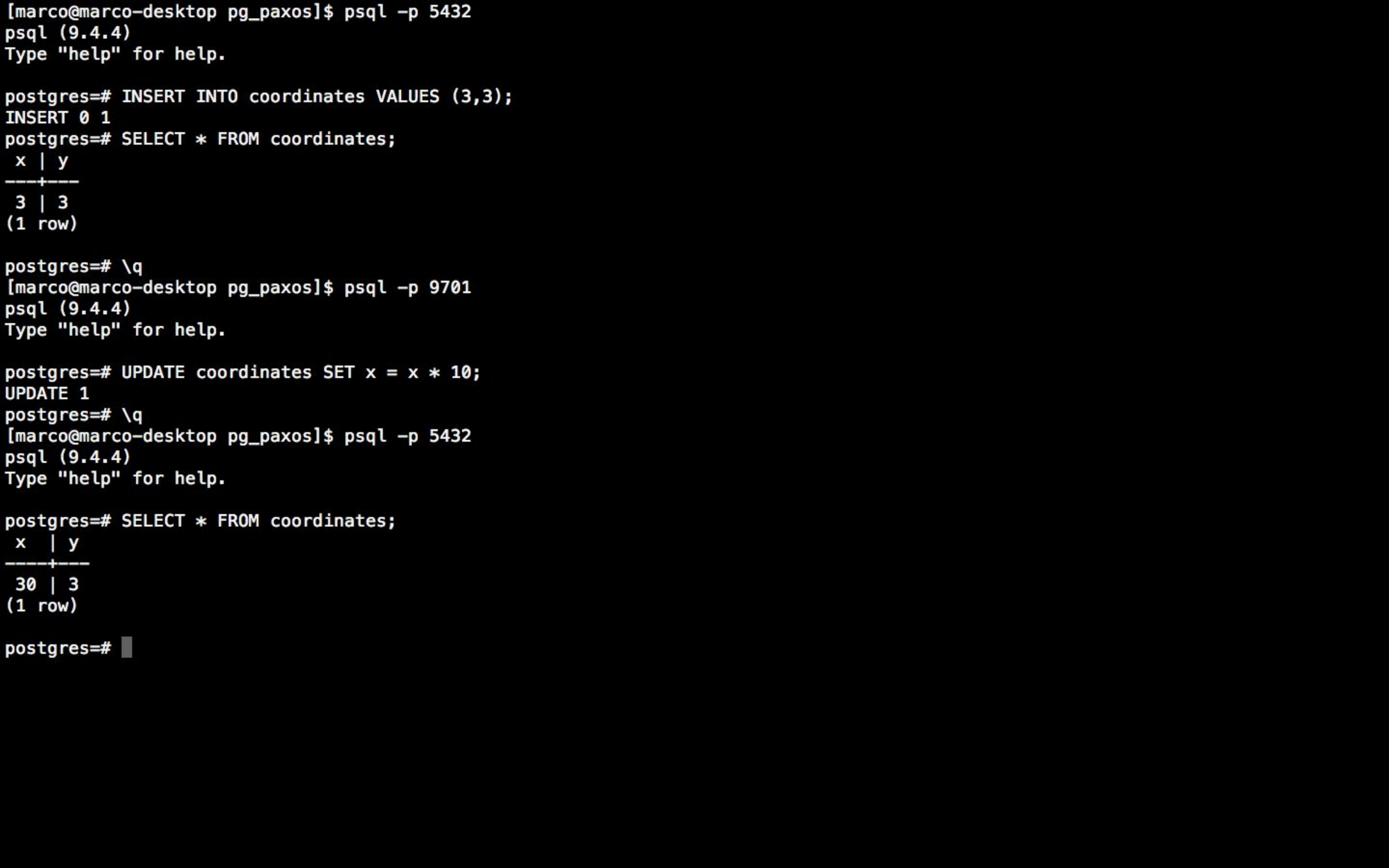
https://github.com/citusdata/pg_paxos/



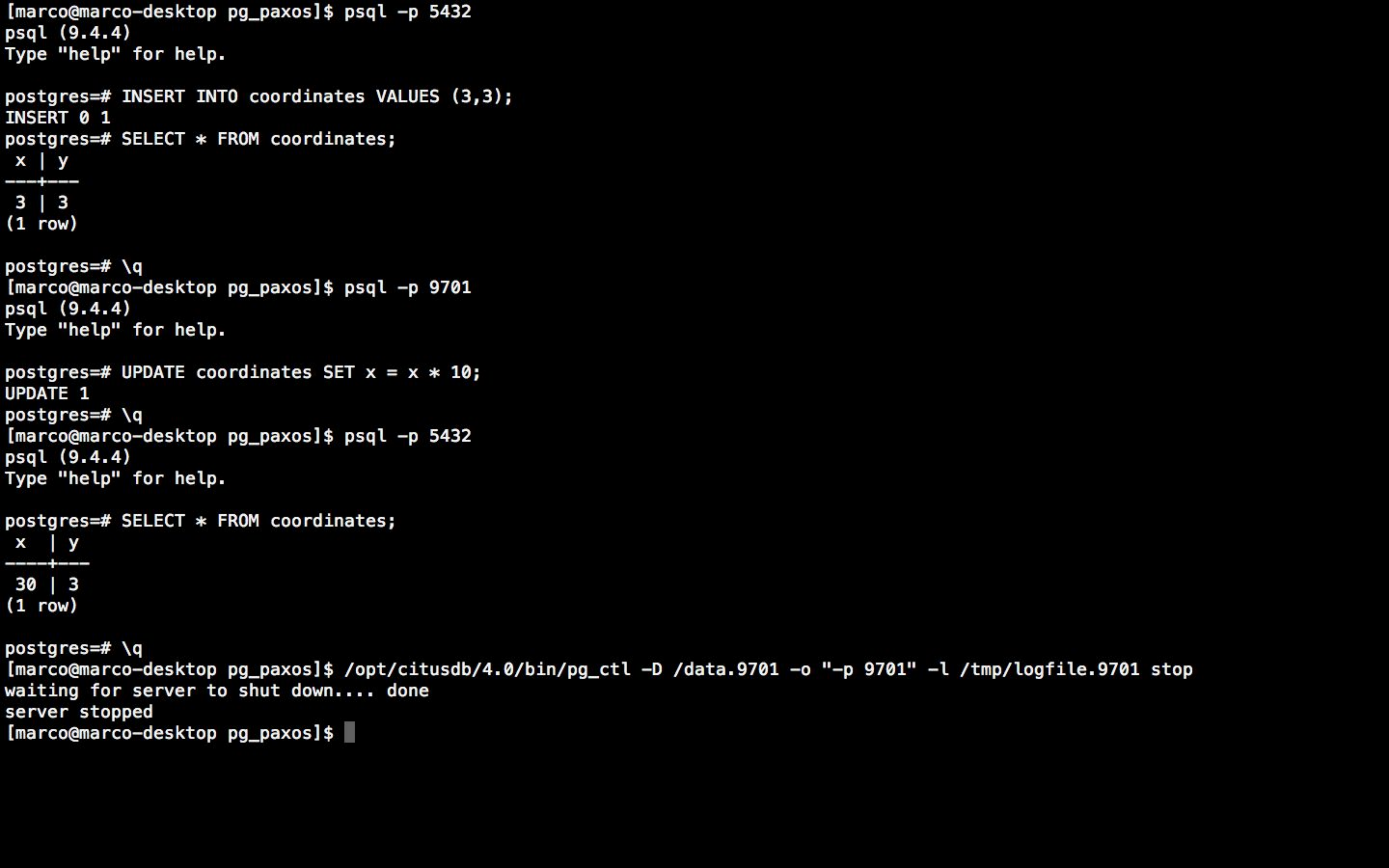
```
[marco@marco-desktop pg_paxos]$ psql -p 5432
psql (9.4.4)
Type "help" for help.

postgres=# INSERT INTO coordinates VALUES (3,3);
INSERT 0 1
postgres=# SELECT * FROM coordinates;
 x | y
---+---
 3 | 3
(1 row)

postgres=# 
```



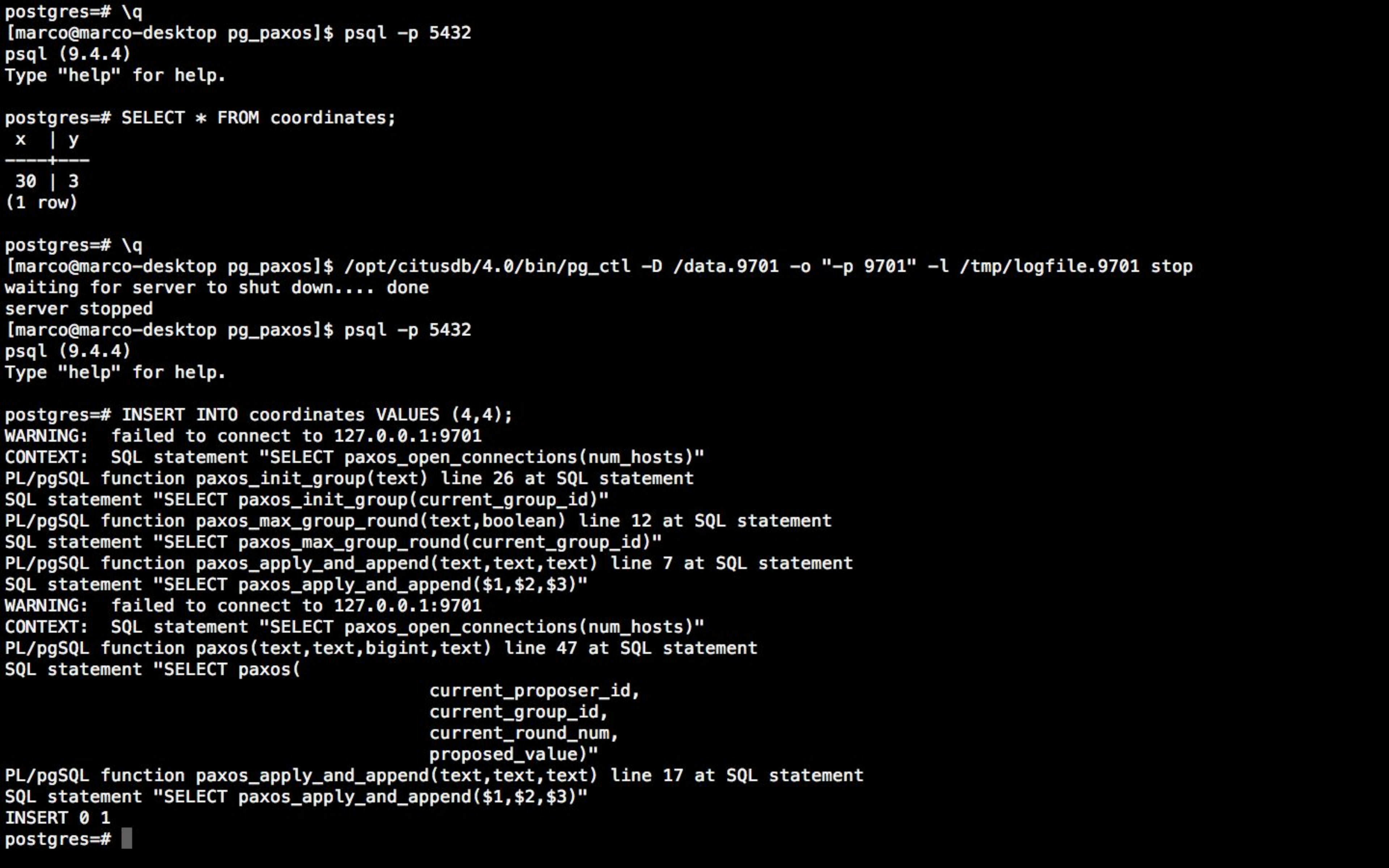
```
[marco@marco-desktop pg_paxos]$ psql -p 5432
psql (9.4.4)
Type "help" for help.

postgres=# INSERT INTO coordinates VALUES (3,3);
INSERT 0 1
postgres=# SELECT * FROM coordinates;
 x | y
---+---
 3 | 3
(1 row)

postgres=# \q
[marco@marco-desktop pg_paxos]$ psql -p 9701
psql (9.4.4)
Type "help" for help.

postgres=# UPDATE coordinates SET x = x * 10;
UPDATE 1
postgres=#
```



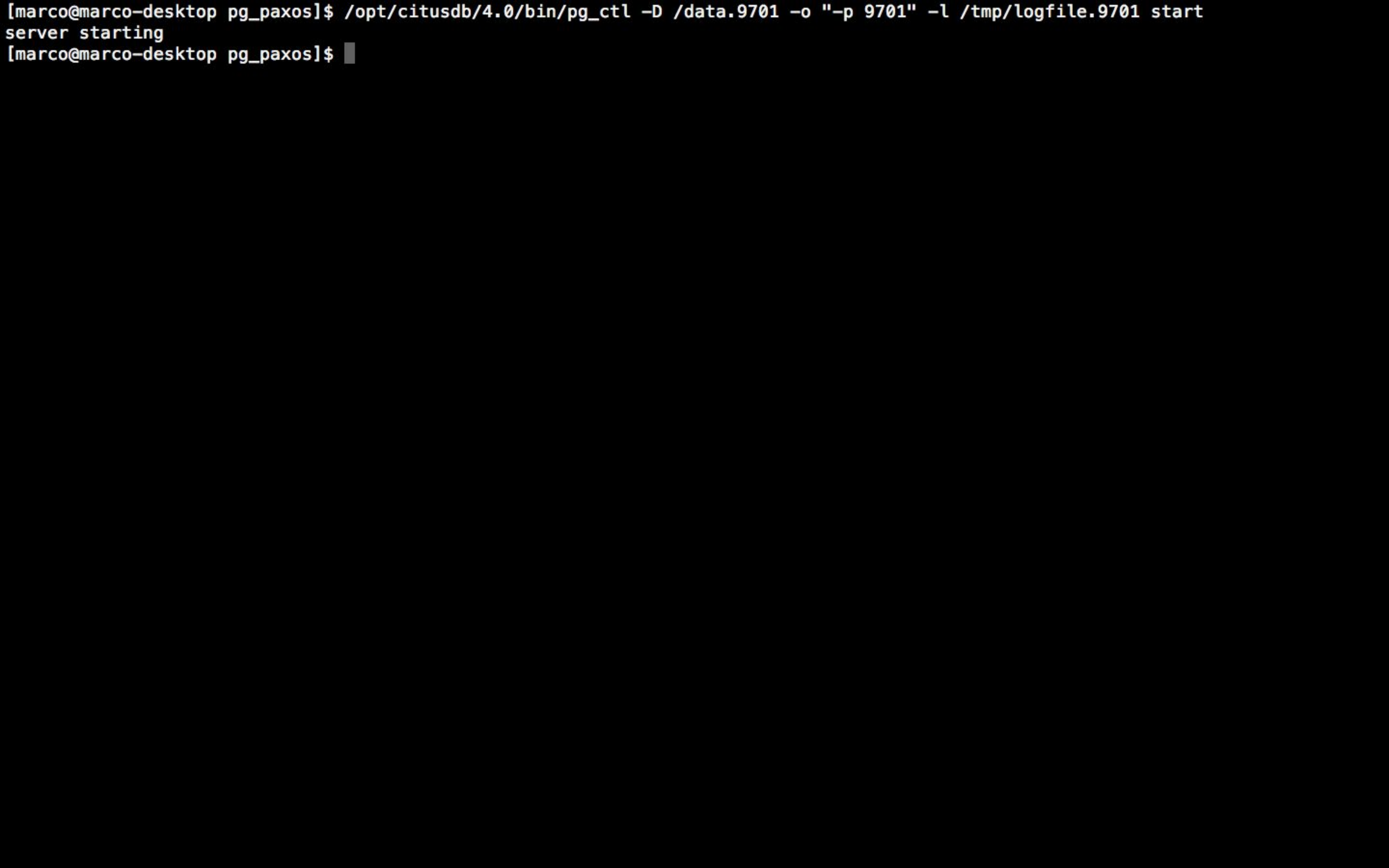
```
[marco@marco-desktop pg_paxos]$ psql -p 5432
psql (9.4.4)
Type "help" for help.

postgres=# INSERT INTO coordinates VALUES (3,3);
INSERT 0 1
postgres=# SELECT * FROM coordinates;
 x | y
---+---
 3 | 3
(1 row)

postgres=# \q
[marco@marco-desktop pg_paxos]$ psql -p 9701
psql (9.4.4)
Type "help" for help.

postgres=# UPDATE coordinates SET x = x * 10;
UPDATE 1
postgres=# \q
[marco@marco-desktop pg_paxos]$ psql -p 5432
psql (9.4.4)
Type "help" for help.

postgres=# SELECT * FROM coordinates;
 x  | y
----+---
 30 | 3
(1 row)

postgres=# 
```



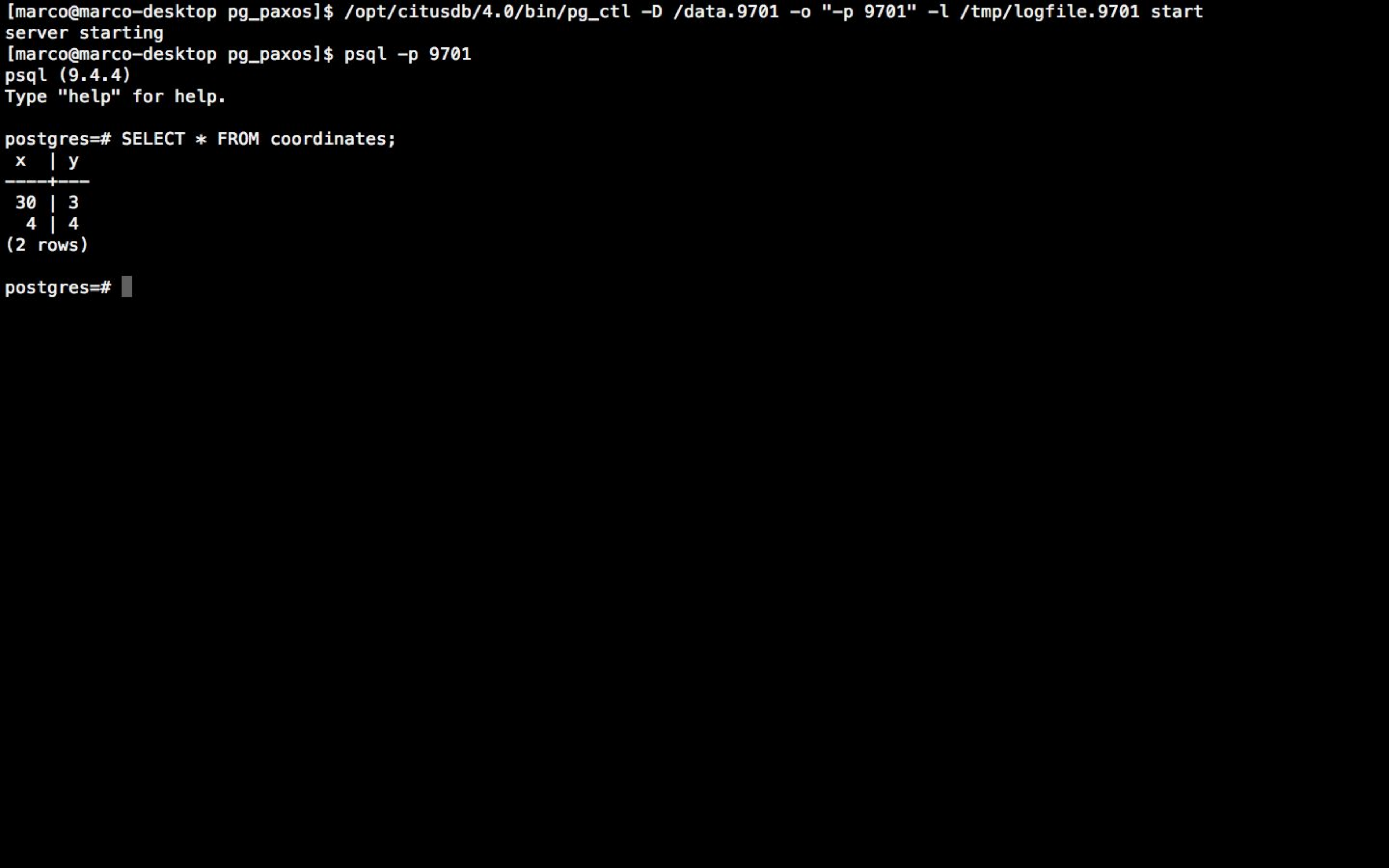
```
[marco@marco-desktop pg_paxos]$ psql -p 5432
psql (9.4.4)
Type "help" for help.

postgres=# INSERT INTO coordinates VALUES (3,3);
INSERT 0 1
postgres=# SELECT * FROM coordinates;
 x | y
---+---
 3 | 3
(1 row)

postgres=# \q
[marco@marco-desktop pg_paxos]$ psql -p 9701
psql (9.4.4)
Type "help" for help.

postgres=# UPDATE coordinates SET x = x * 10;
UPDATE 1
postgres=# \q
[marco@marco-desktop pg_paxos]$ psql -p 5432
psql (9.4.4)
Type "help" for help.

postgres=# SELECT * FROM coordinates;
 x  | y
----+---
 30 | 3
(1 row)

postgres=# \q
[marco@marco-desktop pg_paxos]$ /opt/citusdb/4.0/bin/pg_ctl -D /data.9701 -o "-p 9701" -l /tmp/logfile.9701 stop
waiting for server to shut down.... done
server stopped
[marco@marco-desktop pg_paxos]$ 
```



```
postgres=# \q
[marco@marco-desktop pg_paxos]$ psql -p 5432
psql (9.4.4)
Type "help" for help.

postgres=# SELECT * FROM coordinates;
 x  | y
----+---
 30 | 3
(1 row)

postgres=# \q
[marco@marco-desktop pg_paxos]$ /opt/citusdb/4.0/bin/pg_ctl -D /data.9701 -o "-p 9701" -l /tmp/logfile.9701 stop
waiting for server to shut down.... done
server stopped
[marco@marco-desktop pg_paxos]$ psql -p 5432
psql (9.4.4)
Type "help" for help.

postgres=# INSERT INTO coordinates VALUES (4,4);
WARNING:  failed to connect to 127.0.0.1:9701
CONTEXT:  SQL statement "SELECT paxos_open_connections(num_hosts)"
PL/pgSQL function paxos_init_group(text) line 26 at SQL statement
SQL statement "SELECT paxos_init_group(current_group_id)"
PL/pgSQL function paxos_max_group_round(text,boolean) line 12 at SQL statement
SQL statement "SELECT paxos_max_group_round(current_group_id)"
PL/pgSQL function paxos_apply_and_append(text,text,text) line 7 at SQL statement
SQL statement "SELECT paxos_apply_and_append($1,$2,$3)"
WARNING:  failed to connect to 127.0.0.1:9701
CONTEXT:  SQL statement "SELECT paxos_open_connections(num_hosts)"
PL/pgSQL function paxos(text,text,bigint,text) line 47 at SQL statement
SQL statement "SELECT paxos(
                                        current_proposer_id,
                                        current_group_id,
                                        current_round_num,
                                        proposed_value)"
PL/pgSQL function paxos_apply_and_append(text,text,text) line 17 at SQL statement
SQL statement "SELECT paxos_apply_and_append($1,$2,$3)"
INSERT 0 1
postgres=#
```



```
[marco@marco-desktop pg_paxos]$ /opt/citusdb/4.0/bin/pg_ctl -D /data.9701 -o "-p 9701" -l /tmp/logfile.9701 start
server starting
[marco@marco-desktop pg_paxos]$ 
```



```
[marco@marco-desktop pg_paxos]$ /opt/citusdb/4.0/bin/pg_ctl -D /data.9701 -o "-p 9701" -l /tmp/logfile.9701 start
server starting
[marco@marco-desktop pg_paxos]$ psql -p 9701
psql (9.4.4)
Type "help" for help.

postgres=# SELECT * FROM coordinates;
 x  | y
----+---
 30 | 3
  4 | 4
(2 rows)

postgres=# 
```

pg_paxos

Extension for PostgreSQL

1. Basic implementation of Paxos and Multi-Paxos in PL/pgSQL (~1000 lines)
2. Consistent table replication implemented using Multi-Paxos

https://github.com/citusdata/pg_paxos/