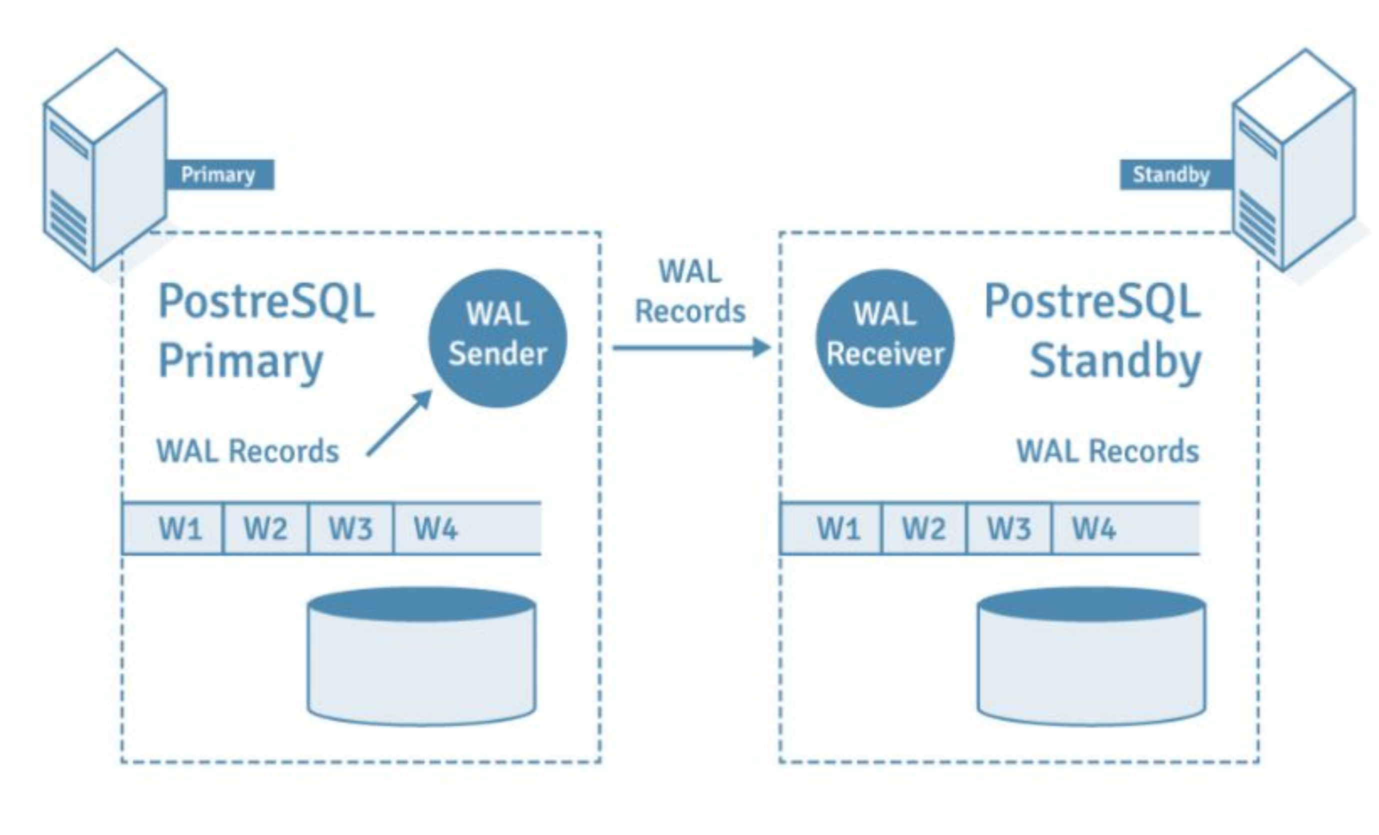
Отладка потоковой репликации в PostgreSQL

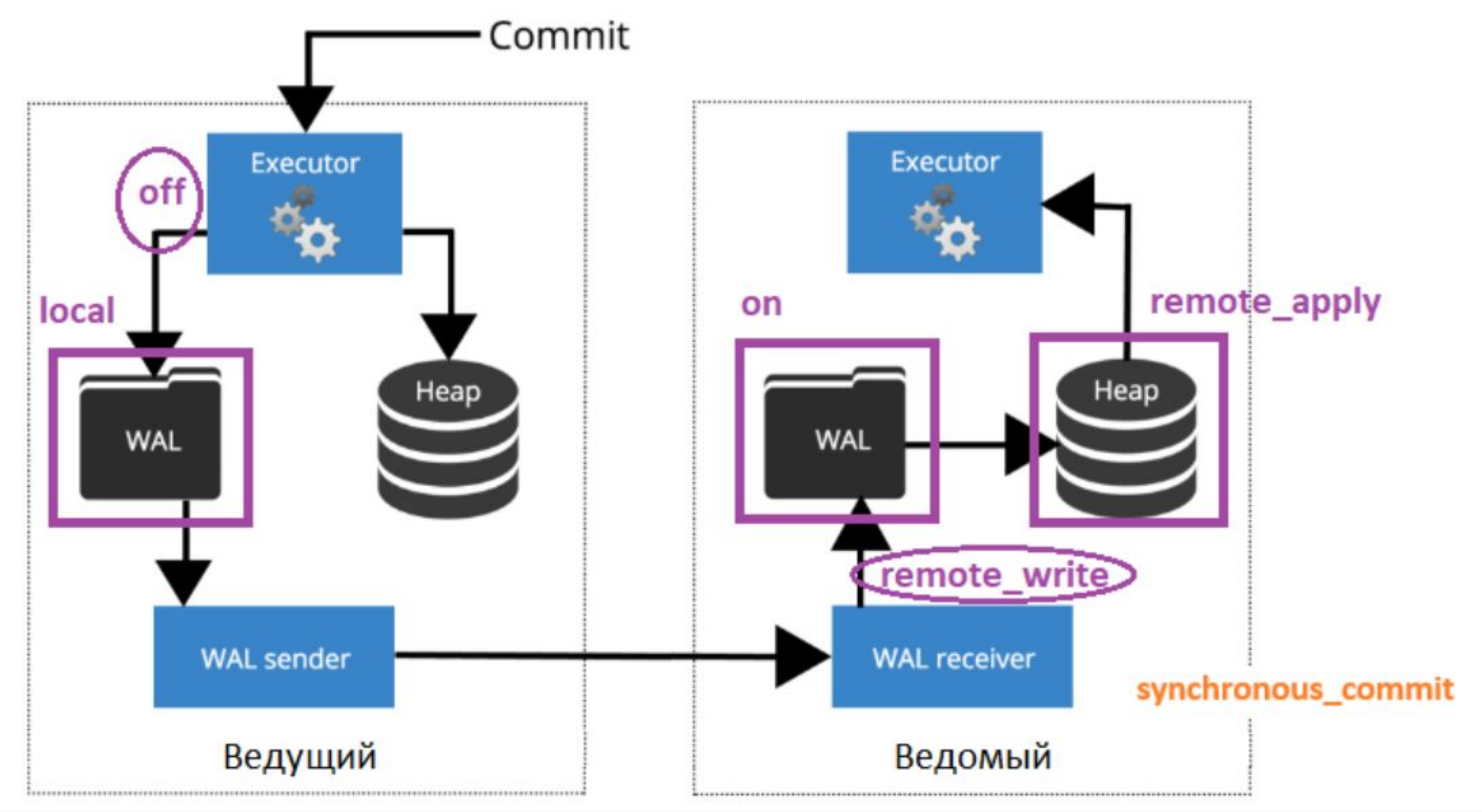
Илья Евдокимов, «Тантор Лабс»

Мастер-класс

Потоковая репликация PostgreSQL



Потоковая репликация PostgreSQL



pg_waldump

Содержимое файлов WAL:

```
/usr/local/pgsql/bin/pg_waldump -f
                              pg_data/pg_wal/000000010000000000000001
>>>
    rmgr: XLOG len (rec/tot): 30/ 30, tx: 0, lsn: 0/01569360, prev
    0/01569328, desc: INSERT
    ...
```

pg_current_wal_lsn(), pg_last_wal_receive_lsn()

Текущее позиция на ведущем:

```
SELECT pg_current_wal_lsn();
>>>
    pg_current_wal_lsn
   --------------------
     0/15694A8
   (1 row)
```

Последняя позиция WAL, которую получила реплика:

```
SELECT pg_last_wal_receive_lsn();
>>>
    pg_last_wal_receive_lsn()
   --------------------
     0/15694A0
   (1 row)
```

pg_wal_lsn_diff()

Количество байт, которые не были доставлены на реплику:

```
SELECT pg_wal_lsn_diff(pg_current_wal_lsn(),
                       pg_last_wal_receive_lsn());
>>>
    pg_wal_lsn_diff
   -----------------
     8
    (1 row)
```

Параметры postgresql.conf

```
SELECT name, short_desc
FROM   pg_settings
WHERE  name = 'synchronous_commit' OR name = 'wal_level' OR
name = 'wal_keep_size' OR name = 'wal_buffers' OR name = 'max_wal_size';
>>>
        name        |                    short_desc
--------------------+--------------------------------------------------
 max_wal_size       | Sets the WAL size that triggers a checkpoint.
 synchronous_commit | Sets the current transaction's synchronization
                           level.
 wal_buffers        | Sets the number of disk-page buffers in shared
                           memory for WAL.
 wal_keep_size      | Sets the size of WAL files held for standby
                           servers.
 wal_level          | Sets the level of information written to the WAL.
```

walsender и walreceiver

```
SELECT name, short_desc FROM pg_settings
WHERE  name = 'wal_receiver_timeout'
OR     name = 'wal_sender_timeout';
>>>
         name          |                short_desc
-----------------------+----------------------------------------
 wal_receiver_timeout | Sets the maximum wait time to receive
                           data from the sending server.
 wal_sender_timeout   | Sets the maximum time to wait for WAL
                           replication.
```

Сборка PostgreSQL

```
git clone https://github.com/postgres/postgres
cd postgres/

./configure --prefix=/usr/local/pgsql \
     --with-python \
     --with-icu \
     --with-lz4 \
     --with-zstd \
     --enable-tap-tests \
     --enable-debug \
     --enable-cassert

make -j 4
sudo make install
make -C src/tests/modules/tap_test/ check
```

Primary

Создаем кластер для primary

```
use PostgreSQL::Test::Cluster;

my $primary = PostgreSQL::Test::Cluster->new('main');
$primary = init(allows_streaming => 1);
```

Конфигурация primary

Все параметры запишутся в конец postgresql.conf

```
$primary->append_conf(
     wal_writer_delay = 10000ms
     log_min_messages = debug1
     synchronous_commit = on
     …
);
```

Standby

Создаем standby

```
my $standby = PostgreSQL::Test::Cluster->new('standby');
$standby->init_from_backup(
$primary, 'backup',
standby => 1, has_streaming => 1);
$standby->start;
```

TAP-тесты

```
$result == 'walcrc'
```

```
is($result, 'walcrc', 'Comment of test');
```

Демонстрация!!!

Спасибо за внимание!

https://tantorlabs.ru/

info@tantorlabs.ru

+7 495 787 51 78