



10 Years of Postgres in Production

Madhav Khakhar
Endurance International Group

About Us

We help small and medium-sized business strive on the web by providing web presence solutions.

- Web hosting - Shared / Dedicated
- Domain names Registration
- Email Marketing



Our Brands



DevCloud

Experience so far

- It's been pretty good
- > 10 years
- Started with 8.7 -> 9.0 -> 9.3 -> ..
- A single outage in last 2 years

Configuration and stats

- OLTP Workload
- Size of PGDATA
 - Primary site ~**700 GB**
 - Archive site ~**1.8 TB**
- Stats
 - 20k queries per second (Average QPS)
 - 5.5k Select
 - 0.2k Insert/Update/Delete
 - 14k prepared queries
- Server Configuration - 48 core, 128 GB RAM, Centos 6.7
- PGDATA partition - 3 TB RAID1 array of SSDs

Postgres Cluster Configuration

The postgres cluster being used by the web application includes:

- 1 master - RW DB
- 3 slaves - RO DB
- 1 DR slave - RO DB
- Postgres physical streaming replication
- RBAC, HBAC (pg_hba.conf)
- Vacuum Analyze (Daily)

We also have an Archive Database Cluster containing 2 nodes. The archived transaction details are moved to the archive database periodically.

Backups and Recovery

Everything is backed up at 2 places:

- Within the DC
- Offshore (S3)

We take following backups:

- Physical backup - Basebackup
- Logical Backup - PG Dump
- WAL file backups

Verification of the backup on an off-shore instance for checking Data Corruption.

Monitoring

- Ganglia
- New Relic
- Pgbadger
- Nagios
 - Replication lag
 - High number of DB connections
 - Long Running Query
- Auxiliary Bash scripts / sysmon
 - Table and index sizes (hourly)
 - idle in transaction (every 5 minutes)
 - Pg_stat_activity (every minute)

Database Failover Process

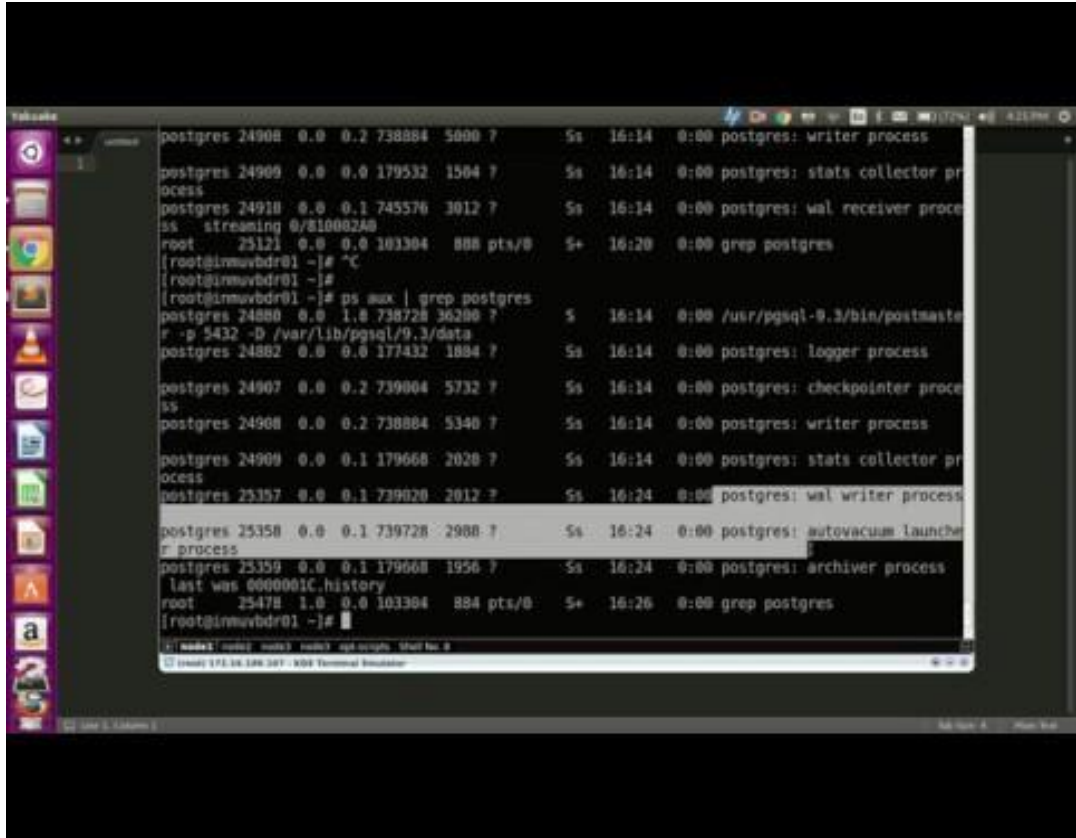
Configuration

- VIPs
- Connection strings
- `recovery_target_timeline = 'latest'`

Steps

- Disable traffic to the RW, RO databases
- Flush WAL to disk on master database
- Check replication lag before going ahead with promotion of slave
- Stop Postgres on master
- Promote Slave DB
- Swap VIPs, Enable traffic

Database Failover Demo



A terminal window showing the output of the 'ps' command, listing PostgreSQL processes. The processes include writer, stats collector, wal receiver, checkpointer, and archiver. The terminal also shows the execution of 'grep postgres' and 'ps aux | grep postgres' commands.

```
postgres 24908 0.0 0.2 738884 5090 ? Ss 16:14 0:00 postgres: writer process
postgres 24909 0.0 0.0 179532 1564 ? Ss 16:14 0:00 postgres: stats collector process
postgres 24910 0.0 0.1 745576 3012 ? Ss 16:14 0:00 postgres: wal receiver process
ss streaming @/810002A0
root 25121 0.0 0.0 103304 888 pts/0 S+ 16:20 0:00 grep postgres
[root@irmuvbdr01 ~]# ^C
[root@irmuvbdr01 ~]#
[root@irmuvbdr01 ~]# ps aux | grep postgres
postgres 24880 0.0 1.0 730728 36200 ? S 16:14 0:00 /usr/pgsql-9.3/bin/postmaster -p 5432 -D /var/lib/pgsql/9.3/data
postgres 24882 0.0 0.0 177432 1894 ? Ss 16:14 0:00 postgres: logger process
postgres 24907 0.0 0.2 739004 5732 ? Ss 16:14 0:00 postgres: checkpointer process
ss
postgres 24908 0.0 0.2 738884 5340 ? Ss 16:14 0:00 postgres: writer process
postgres 24909 0.0 0.1 179668 2020 ? Ss 16:14 0:00 postgres: stats collector process
postgres 25357 0.0 0.1 739020 2012 ? Ss 16:24 0:00 postgres: wal writer process
postgres 25358 0.0 0.1 739728 2988 ? Ss 16:24 0:00 postgres: autovacuum launcher process
postgres 25359 0.0 0.1 179668 1956 ? Ss 16:24 0:00 postgres: archiver process
last was 0000001C.history
root 25478 1.0 0.0 103304 894 pts/0 S+ 16:26 0:00 grep postgres
[root@irmuvbdr01 ~]#
```

Database Failover Process (Cont..)

- Why perform a database failover?
- Rationale behind keeping the process the way it is
- Fast Failover

Other Postgres related Tools

- Barman
- Repmgr
- BDR
- Postgres XL

What's next?

- Upgrade to Postgres 9.6
 - Pg_rewind
 - Pglogical
 - Replicate a subset of tables rather than whole database
 - Replicate across versions - Easier postgres upgrades
 - Push real-time updates to downstream application by integrating with Kafka
- Microservices based architecture with separate postgres instances

THANKS!

Any questions?

Email: madhav.k@endurance.com

LinkedIn: <https://in.linkedin.com/in/madhavkhakhar08>

Twitter: @mkmaddy08