

pg_rewind

Michael Paquier

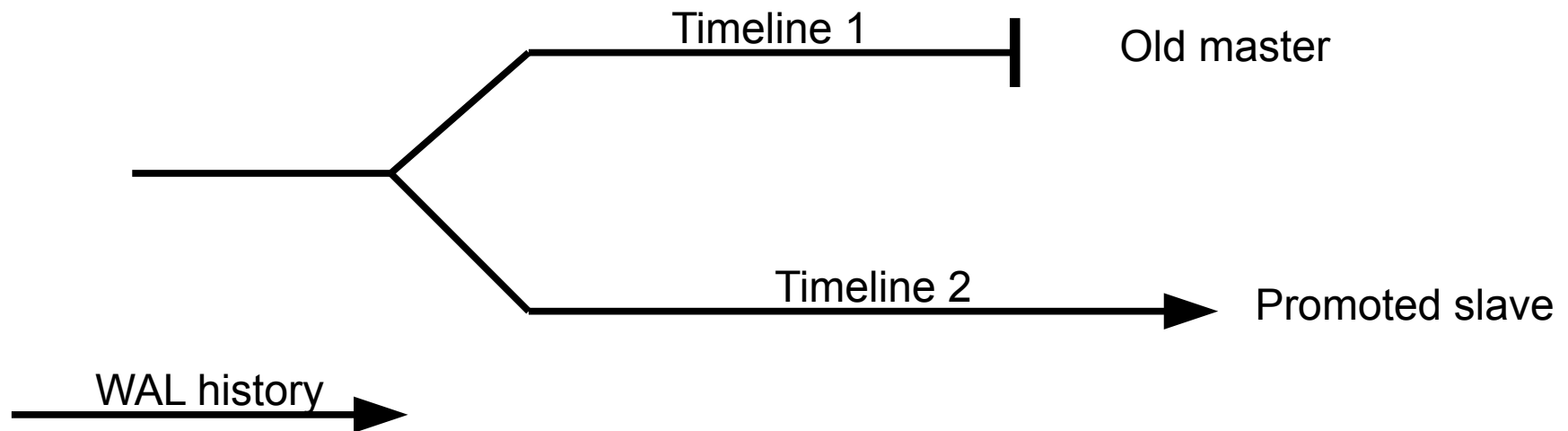
pgunconf#2 @ Tokyo2013/07/13

The lecturer

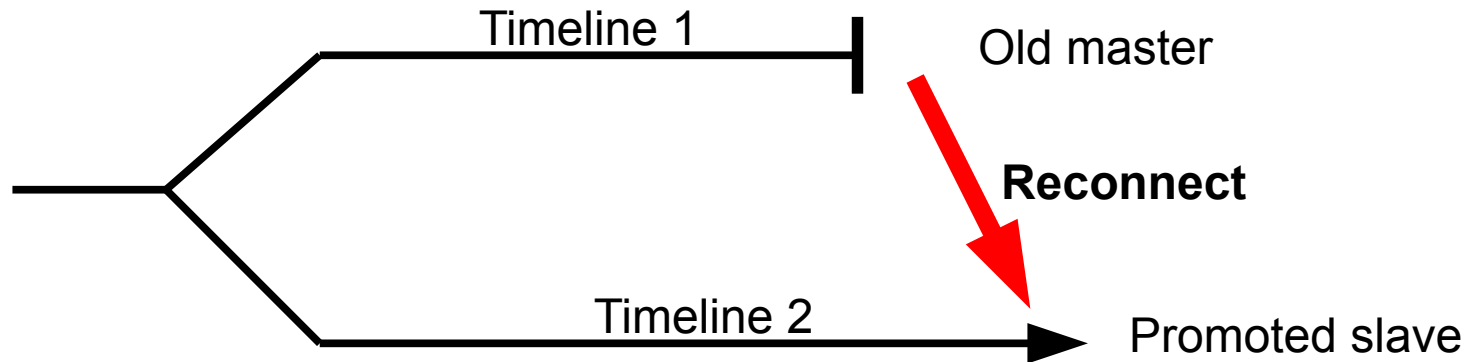
- Michael Paquier
- PostgreSQL hacker and blogger
- Working at VMware on PostgreSQL
- Contact
 - Twitter @michaelpq
 - Mail michael.paquier@gmail.com

About pg_rewind

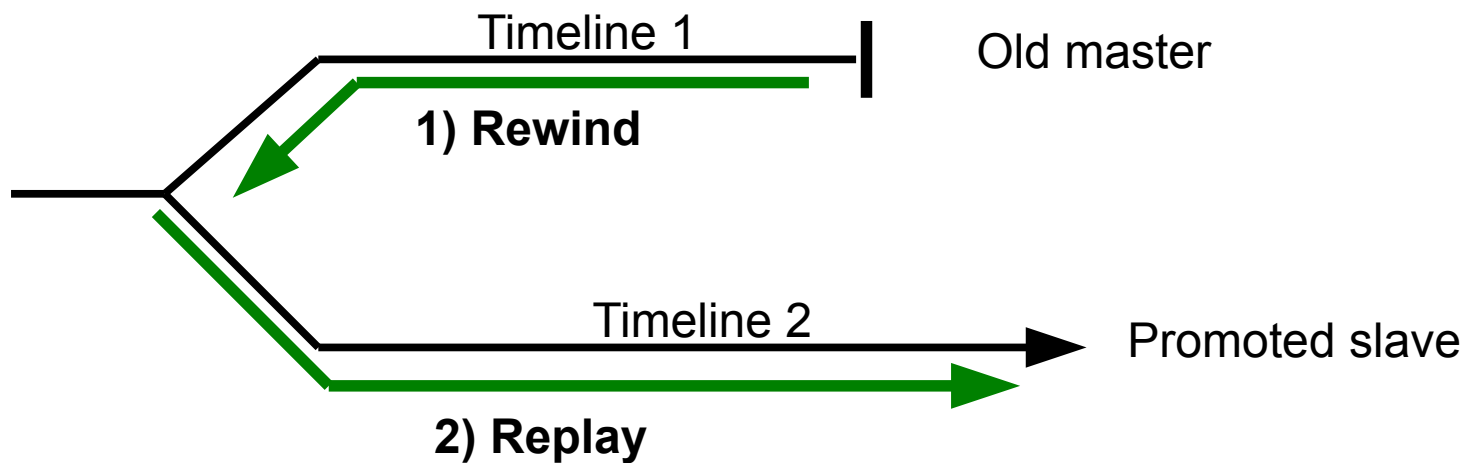
- Resync an existing data directory without new base backup
 - Replug old master to existing cluster => “rewind it”



About pg_rewind (2)



NO!



YES!

WAL history →

About pg_rewind (3)

- Code: https://github.com/vmware/pg_rewind.git
- Largely faster than new base backup
- Limitations
 - Hint bits operations not in WAL, so no track of modified blocks in this case
 - Need to use checksums with full page write to WAL => `initdb -k !`
 - Might not have all WAL for replay, can copy them though
 - No support for tablespace

How to rewind

- Scan old master data folder
 - Start from WAL fork point
 - Record data blocks touched
- Copy all changed blocks from promoted slave to old master
- Copy clog, conf files...
- Replay WAL from master, starting from failover checkpoint

Demonstration

- Master/slave configuration
- Operations
 - Promote slave
 - Stop master (Should be ahead of slave)
 - Apply `pg_rewind` to resync old master
 - `pg_rewind --target-data=$OLD_MASTER_DATA ¥`
`--source-pgdata=$NEW_PROMOTED_DATA`
 - Reconnect old master to promoted slave with `recovery.conf` update

Thanks!