

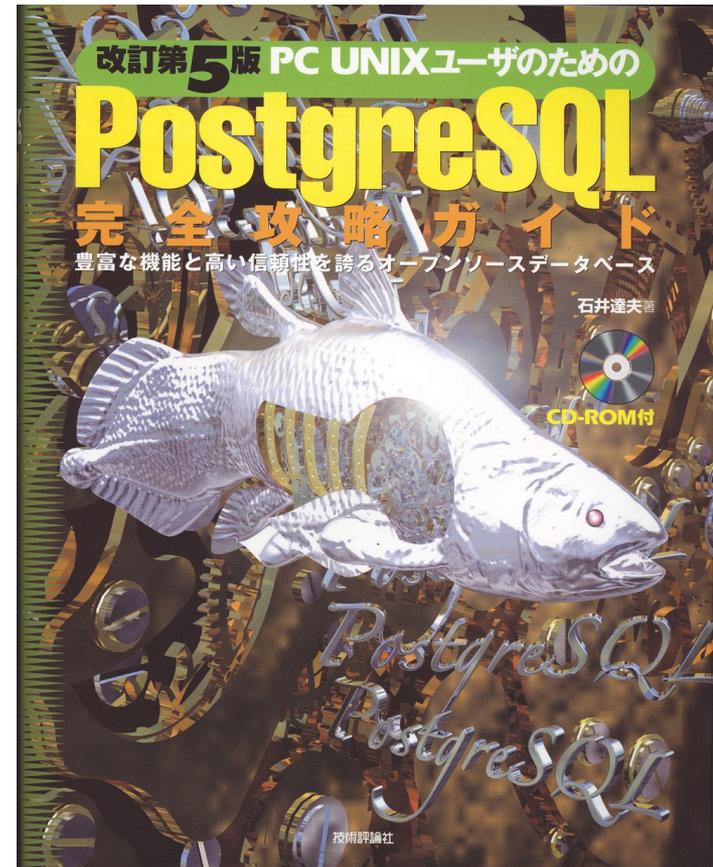
An introduction to pgpool-II Version 3

SRA OSS, Inc. Japan

Tatsuo Ishii

About me

- PostgreSQL & pgpool-II developer
- Working as a PostgreSQL evangelist(16 talks in 2010)
- Working for SRA OSS, Inc. Japan



More PostgreSQL books written by my colleagues



“PostgreSQL complete guide
Third edition”

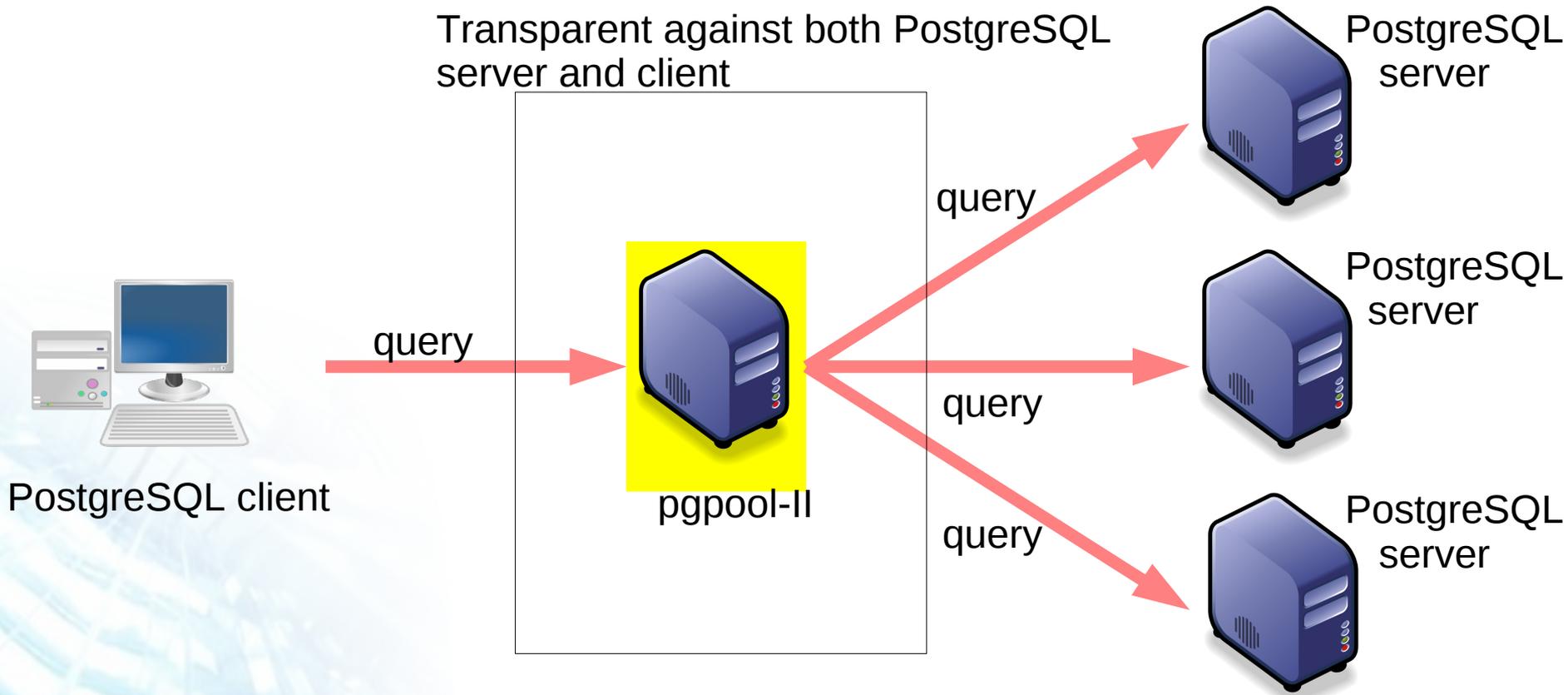


“Road to the database expert”

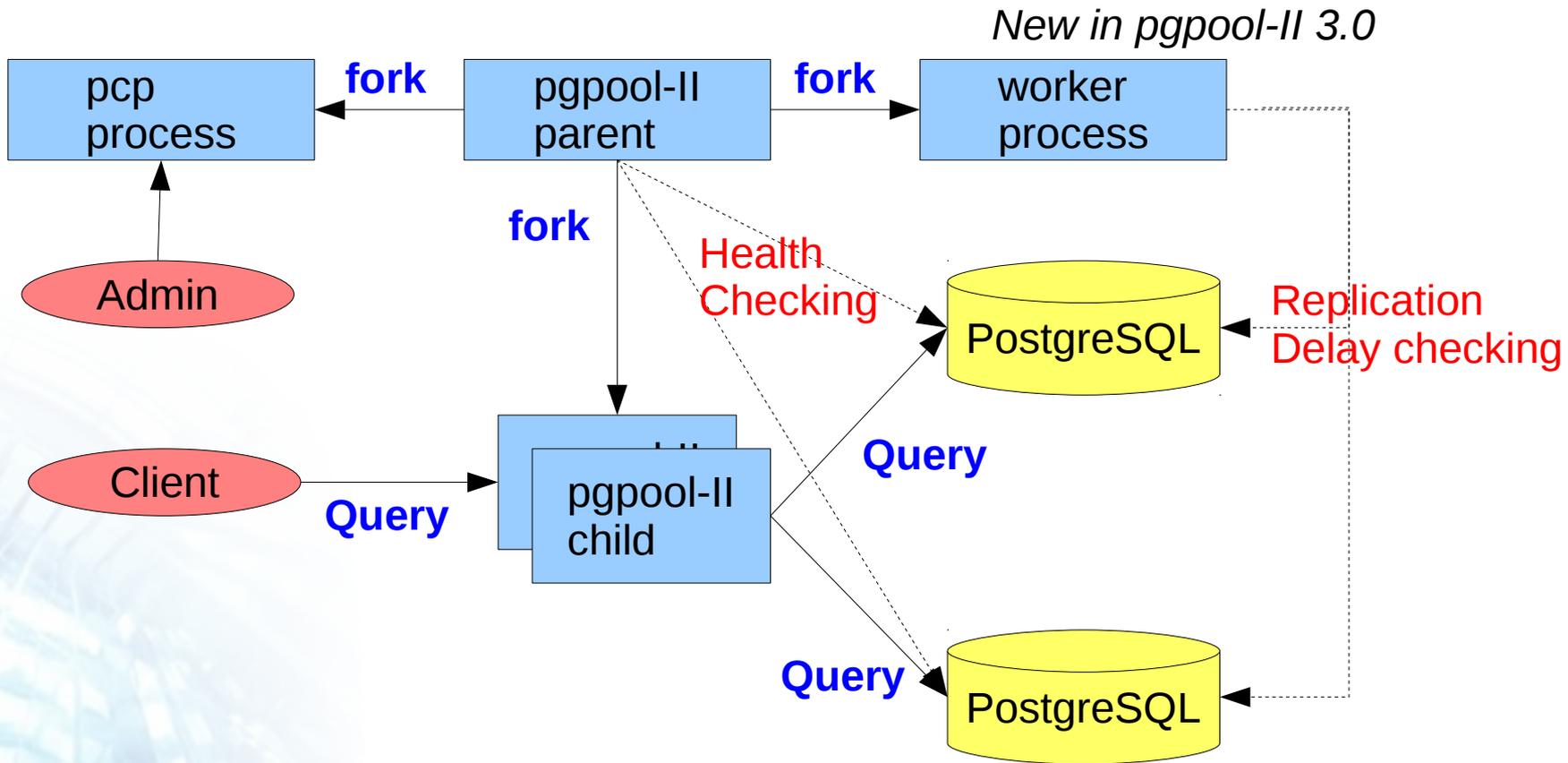
What is pgpool-II?

- Yet another cluster management tool dedicated for PostgreSQL
- OSS project(pgfoundry, BSD License)
- Rich features
 - Synchronous replication
 - Load balancing, automatic failover, connection pooling etc.
 - Collaborating with other replication tools
 - Streaming replication, Slony-I

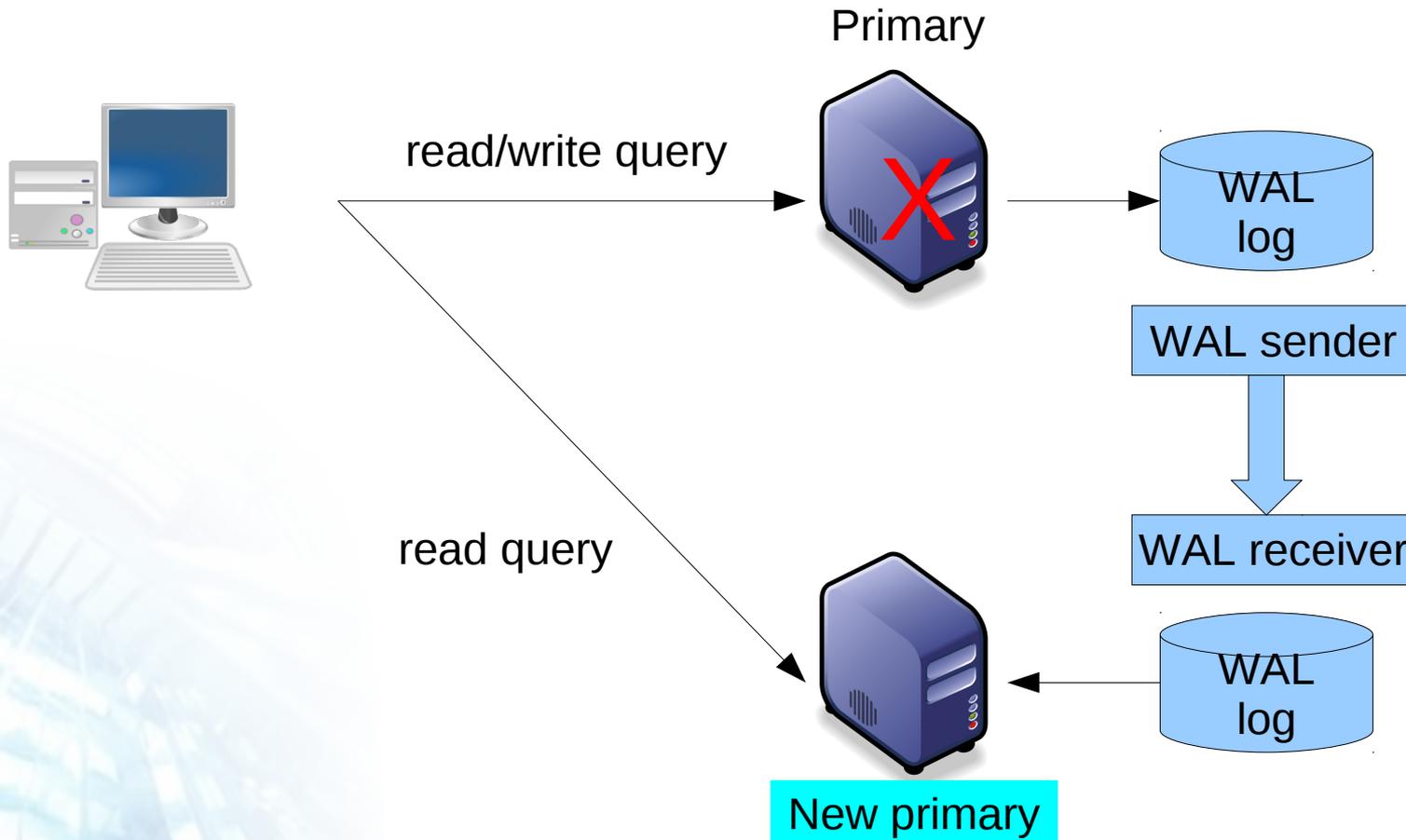
Basic idea of pgpool-II



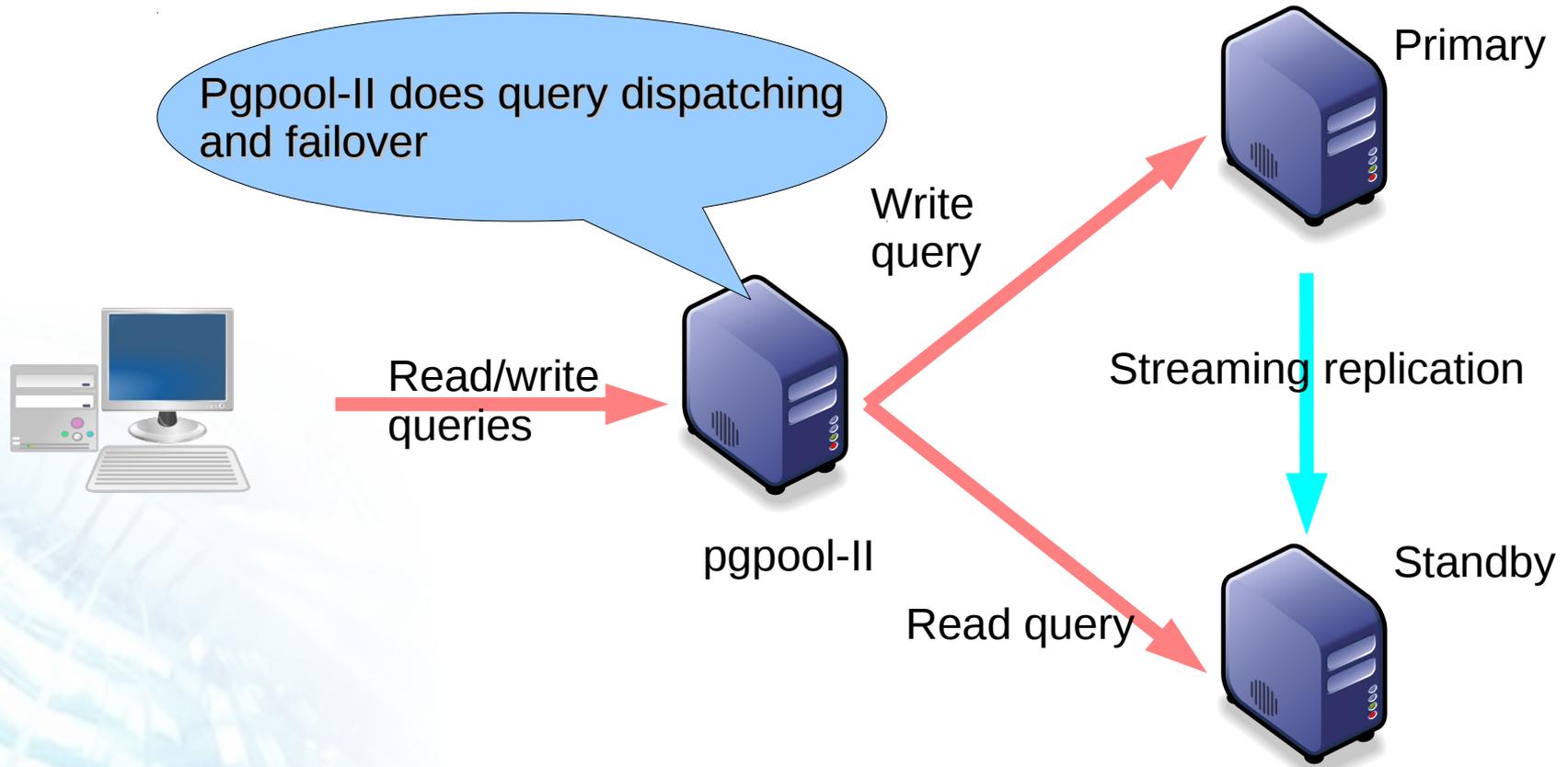
The architecture of pgpool-II



Streaming replication+Hot Standby



Pgpool-II helps Streaming replication

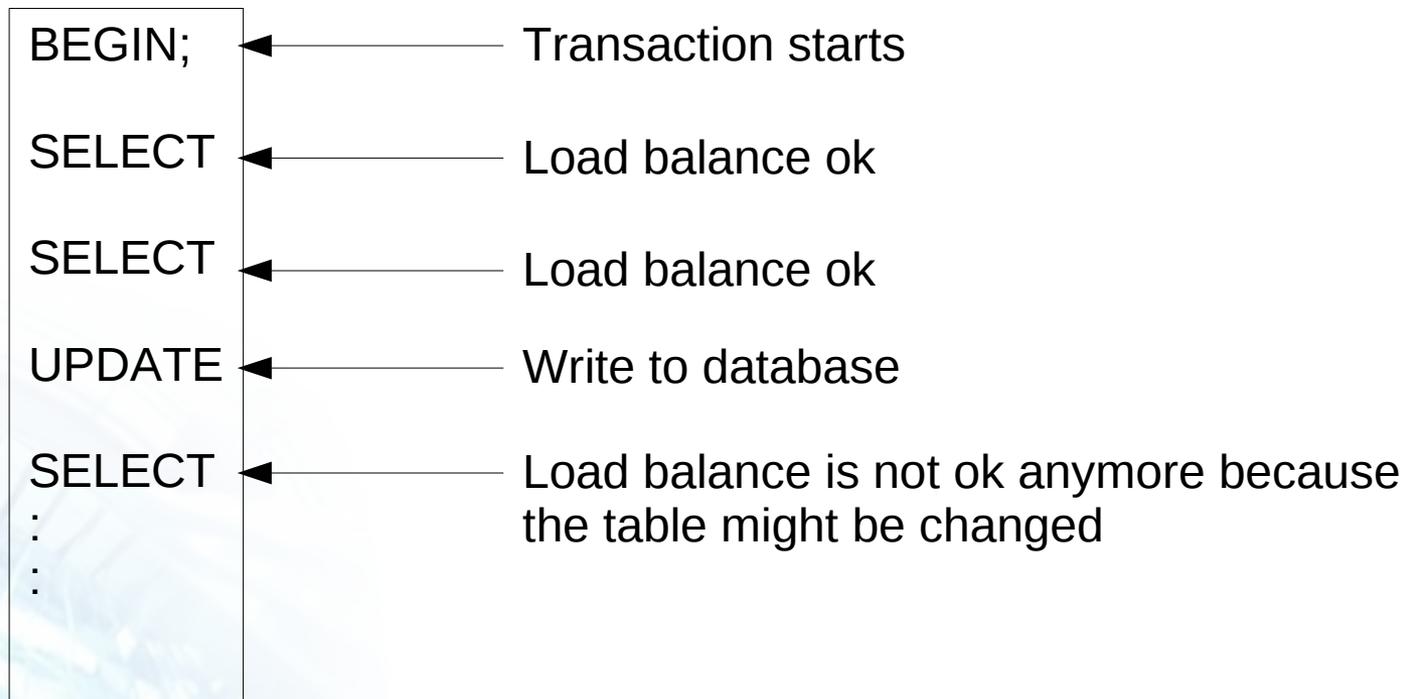


Pgpool-II V3 helps PostgreSQL

- New “stream” mode dedicated for SR/HS
- Automatic query dispatching
- Allow to load balance SELECTs in an explicit transaction
- Allow to specify functions which are writing to DB(they need to be executed on the primary)
- Adding standby servers without stopping pgpool-II

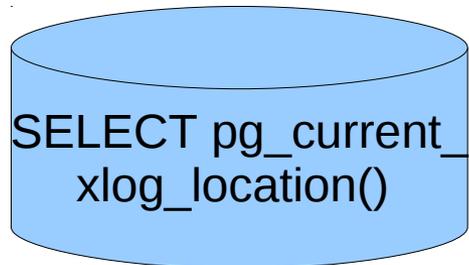
Load balance SELECTs in an explicit transaction

transaction

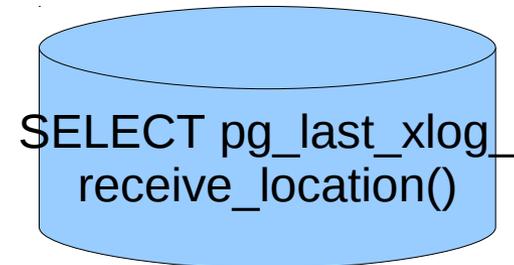


Replication delay detection: how it works

Primary PostgreSQL



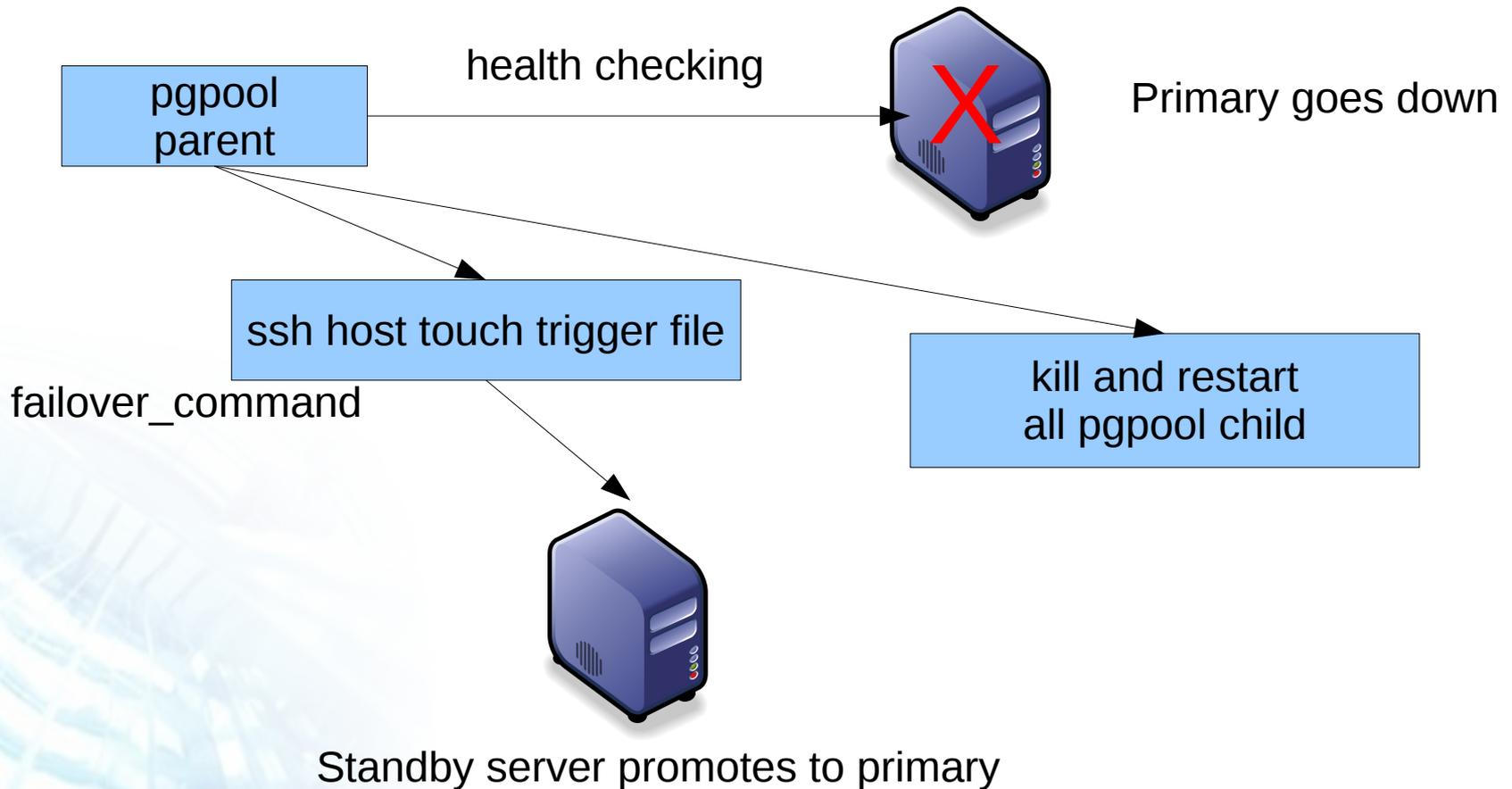
Standby PostgreSQL



call the functions during each “health check period” second and calculate the difference.
if it exceeds the delay threshold, sends read queries to the primary sever only.

worker process

Automatic promoting standby



Logging replication delay

- `log_standby_delay`
 - 'none': no logging delay
 - 'if_over_threshold': log only when delay exceeds `delay_threshold`
 - 'always': always log delay

```
2010-06-28 15:51:32 LOG: pid 13223: Replication of node:1 is behind 1228800 bytes from the primary server (node:0)
2010-06-28 15:51:42 LOG: pid 13223: Replication of node:1 is behind 3325952 bytes from the primary server (node:0)
2010-06-28 15:51:52 LOG: pid 13223: Replication of node:1 is behind 974848 bytes from the primary server (node:0)
2010-06-28 15:52:02 LOG: pid 13223: Replication of node:1 is behind 2990080 bytes from the primary server (node:0)
2010-06-28 15:52:12 LOG: pid 13223: Replication of node:1 is behind 901120 bytes from the primary server (node:0)
2010-06-28 15:52:22 LOG: pid 13223: Replication of node:1 is behind 2433024 bytes from the primary server (node:0)
```

New show commands

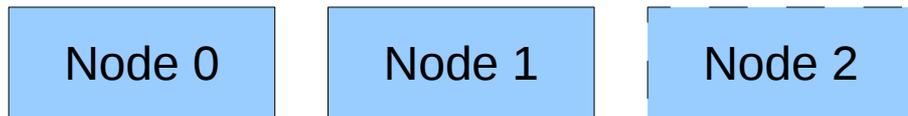
- show pool_nodes
 - show DB node information
- show pool_processes
 - show pgpool child process information
- show pool_pools
 - show connection pool information
- show pool_version
 - show pgpool-II version

Plans for pgpool-II 3.1

- Import PostgreSQL 9.0 SQL parser
- Allow to switch back/force primary node and standby node without restarting pgpool-II
- Allow to use regular expression in black/white_function_list
- Allow to use syslog
- Better handling for more than 1 standby
- French translation

Switch back/force failover

Pgpool-II 3.0



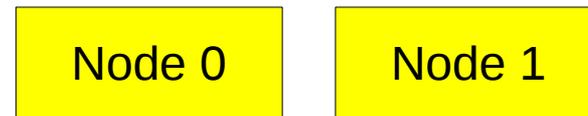
Primary
Down

Standby
Primary
Primary

Standby

Cannot make node 0 as primary.
Because primary needs to be
"younger" than standby node.

Pgpool-II 3.1



Primary
Down
Standby
Primary
Primary

Standby
Primary
Primary
Down
Standby

Switch back/force failover!

More than 1 standby servers. What's the problem anyway?

~~Primary~~

Standby A

Standby B

Which standby should be promoted?
Current rule is “the least position in the config file”. This is not too intelligent.

~~Primary~~

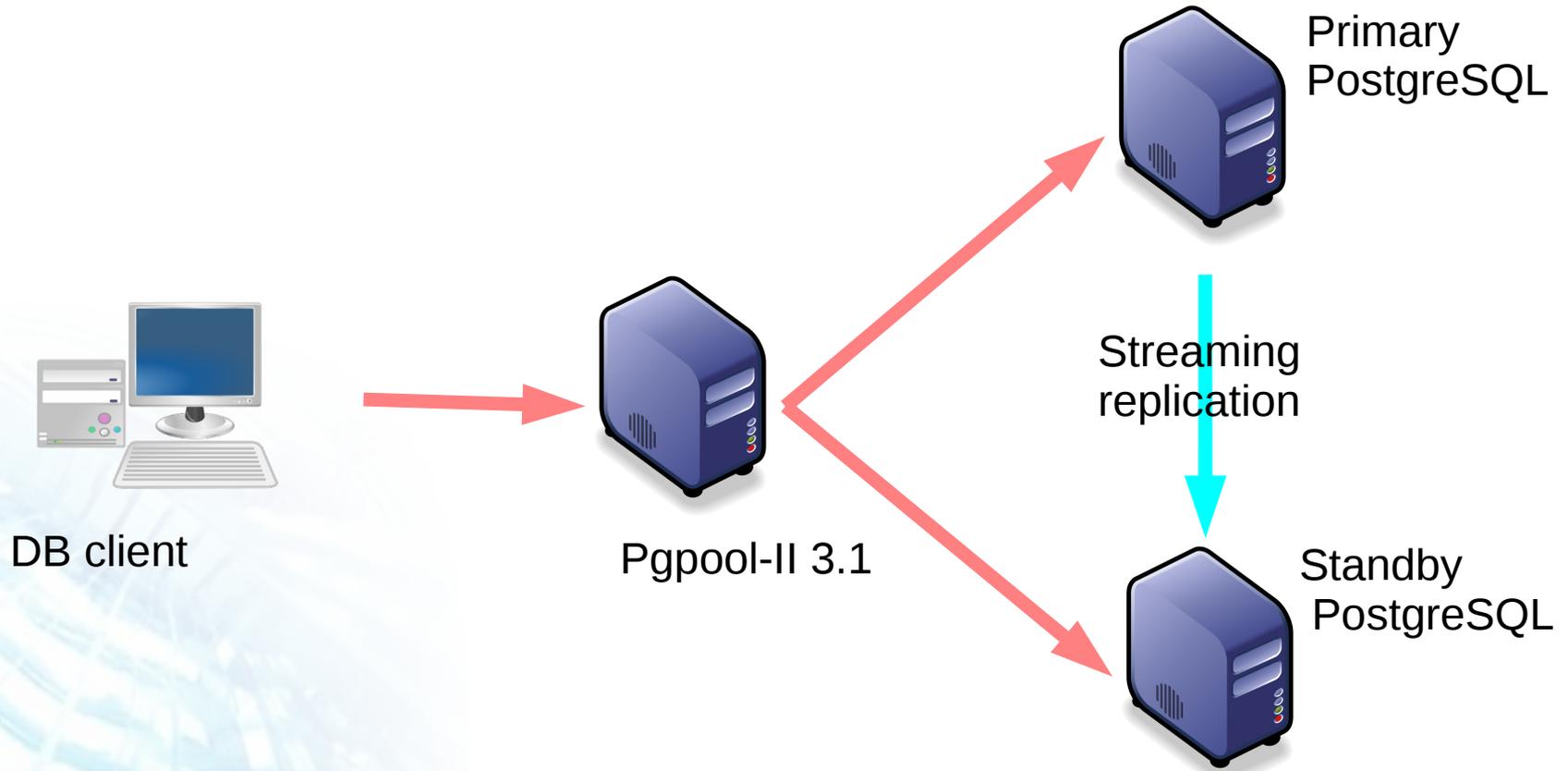
Standby A

New primary



The old standby cannot connect to new primary. So there's no standby server which can backup the system. Moreover, if the client connects to the old standby, it will read outdated data.

Demonstration



URL etc.

- Pgpool-II can be downloaded here:
 - <http://pgfoundry.org/projects/pgpool/>
- Twitter
 - @pgpool2
- SRA OSS's website
 - http://www.srasoss.co.jp/index_en.php