SRA OSS, INC.

# Watchdog Internals

pgpool-II day 2015
2015/05/15, Tokyo

Yugo Nagata
pgpool-II Global Development Group

# Who am I?

- Yugo Nagata
  - pgpool-II developper
- Handling pgpool-II issues generally
  - Developing, bug fixes, analysis, documentation, release, yum repository, buildfarm, RPM, pgpoolAdmin, ...
- Mainly working on watchdog.
  - I started my work on pgpool-II work from testing and debugging of watchdog.

# What I'll talk about?

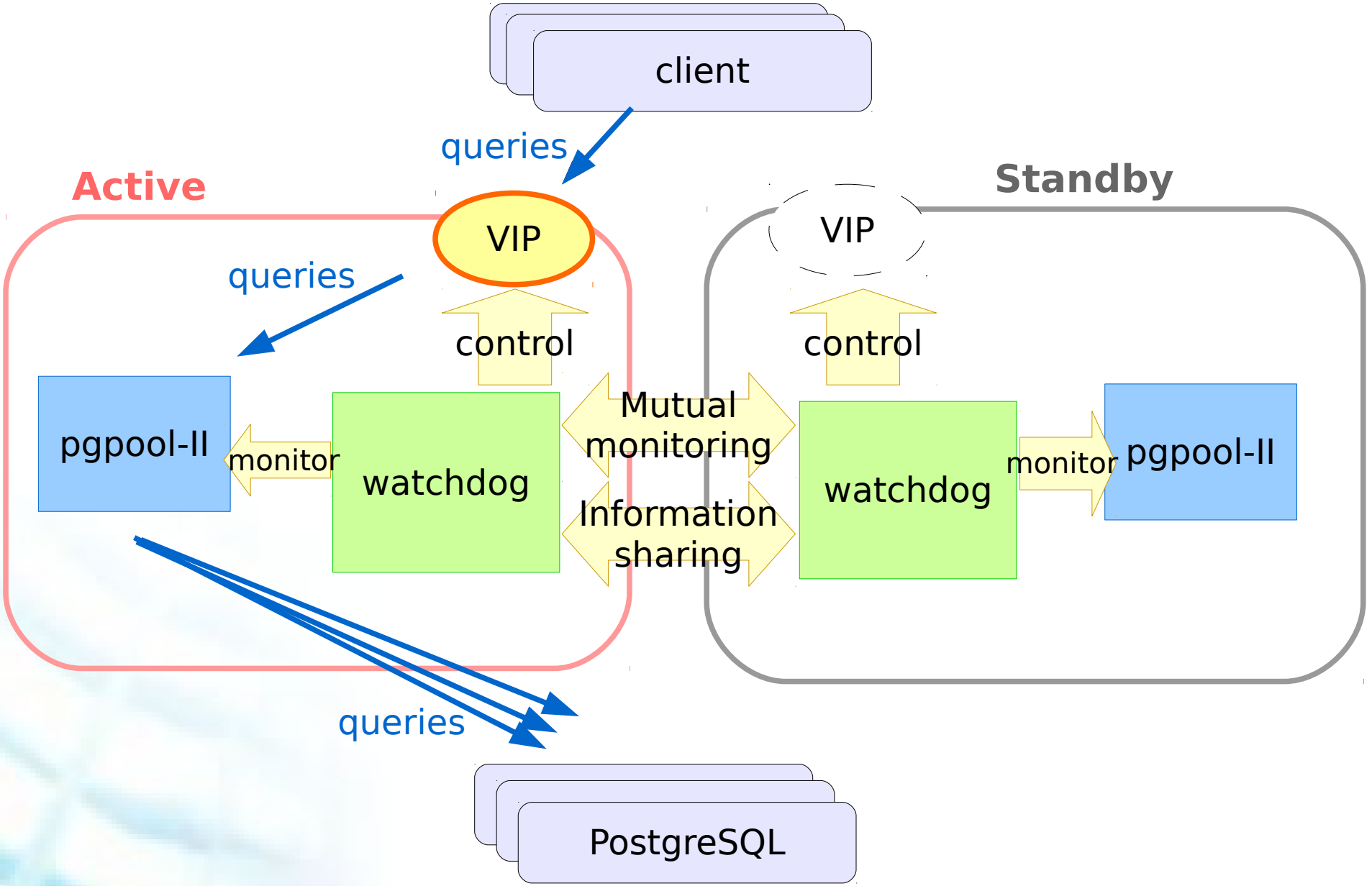
- Watchdog = pgpool-II built-in HA functionality
  - Avoids pgpool-II's Single Point of Failure(SPoF)

- Main features
  - Alive monitoring of pgpool-II
  - Virtual IP control
  - Information sharing between pgpool-IIs

- I'll talk a bit of details of the internal mechanism
  - Processes that constitutes watchdog
  - How the processes communicate between each other
  - How the processes work

Note that I'll talk about pgpool-II 3.4

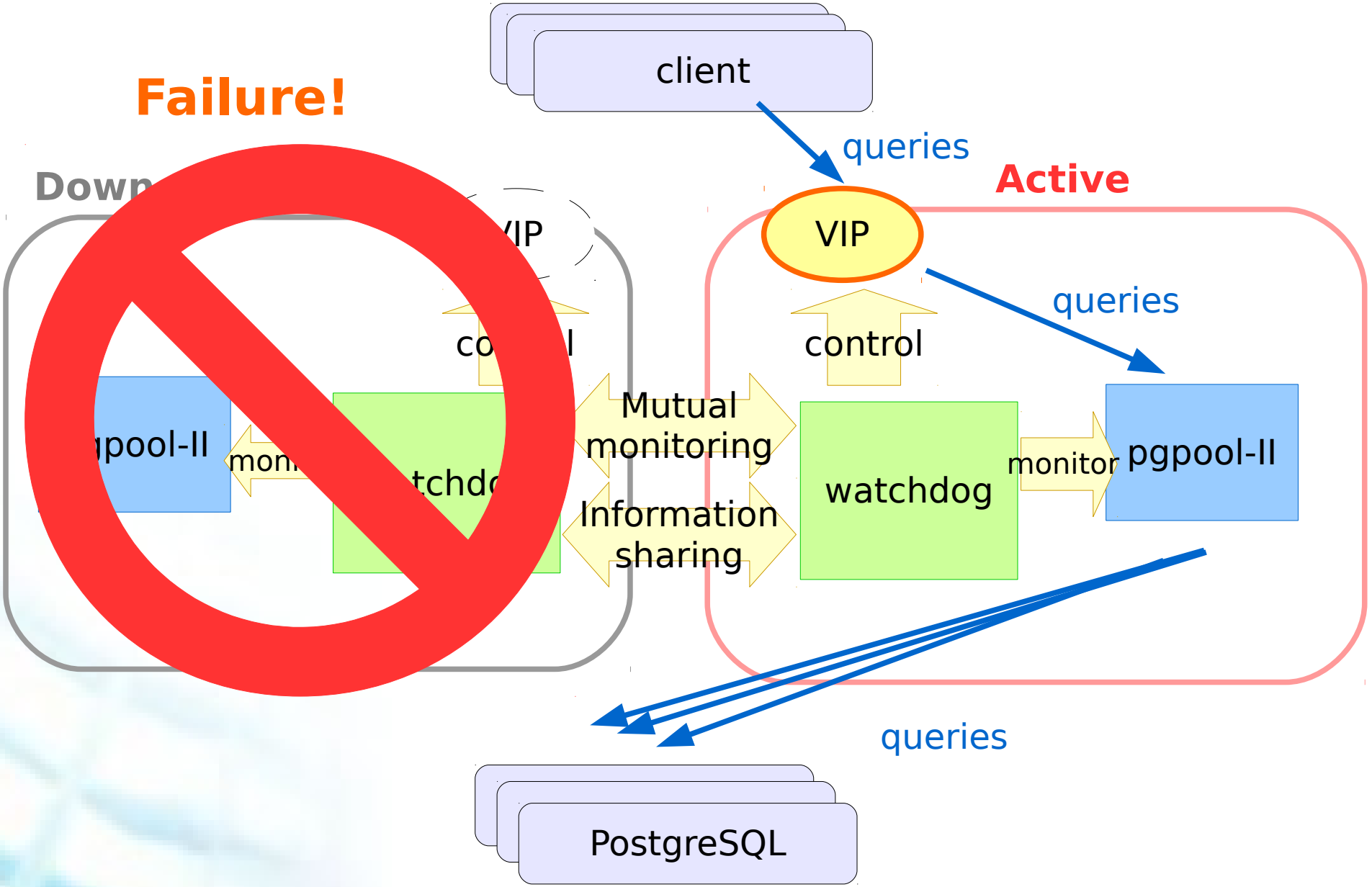
# Overview of watchdog functionality





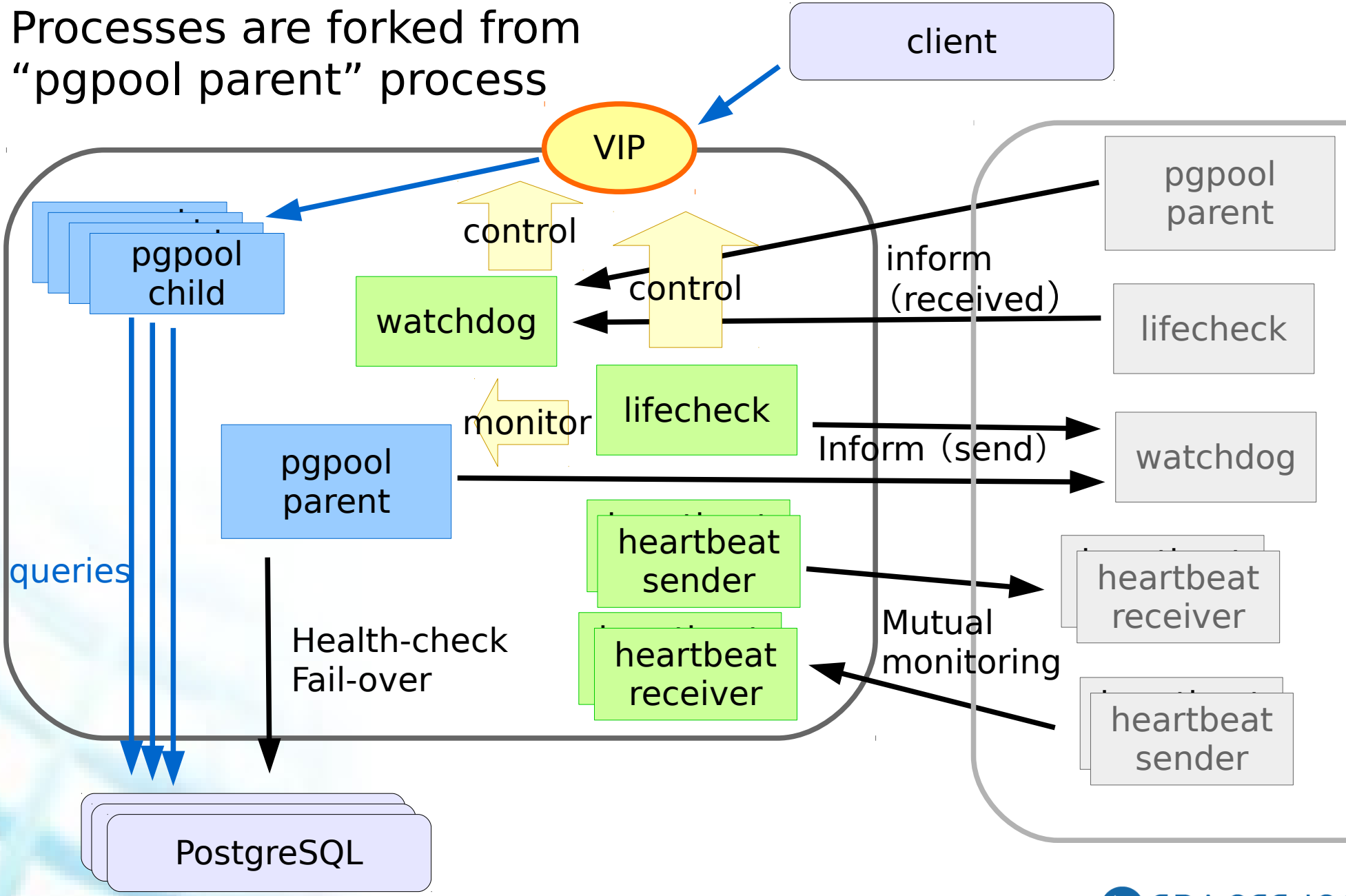

SRA OSS, INC.

# Overview of watchdog functionality



# Overview of process and communication

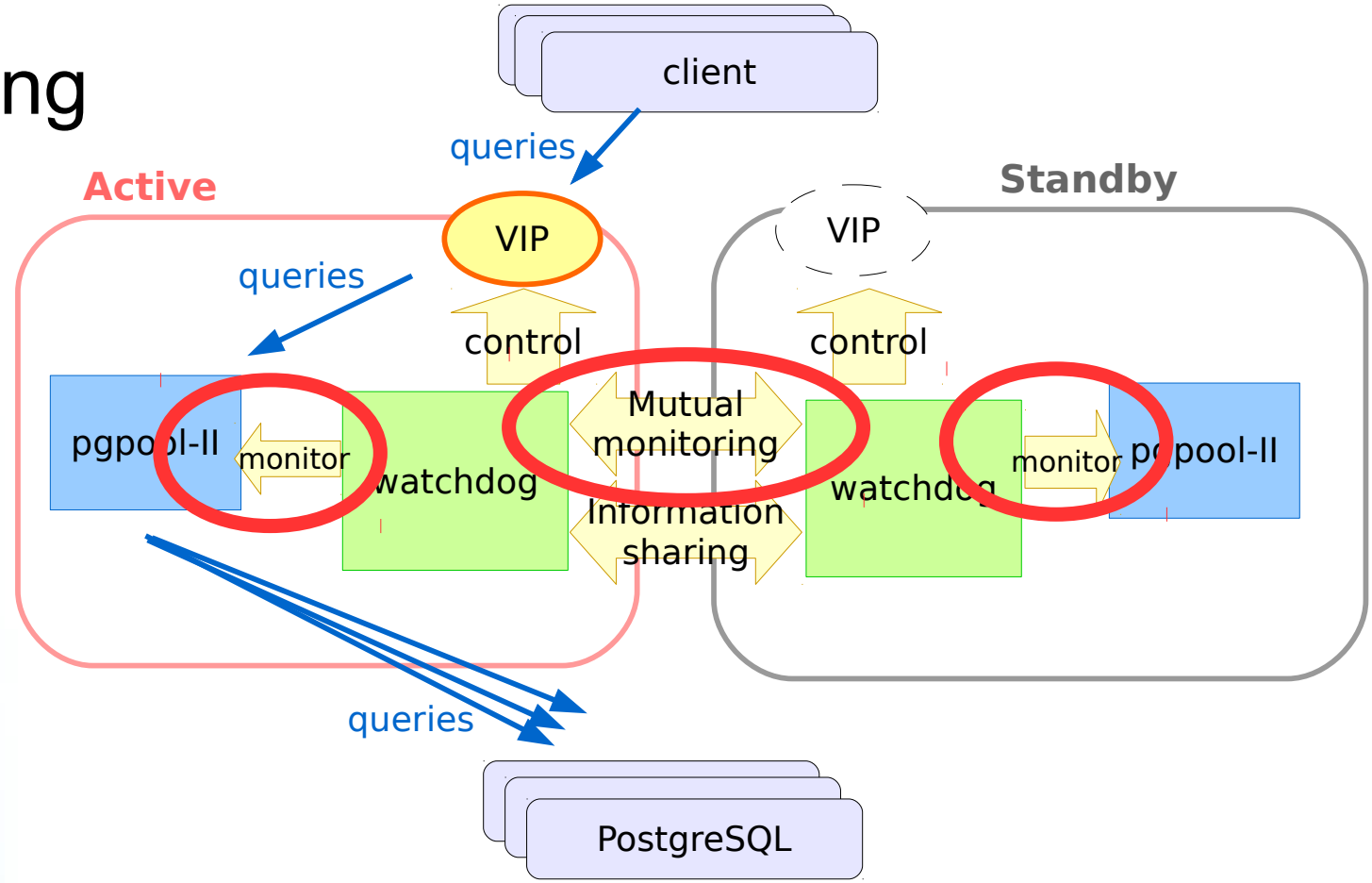
- Processes are forked from “pgpool parent” process





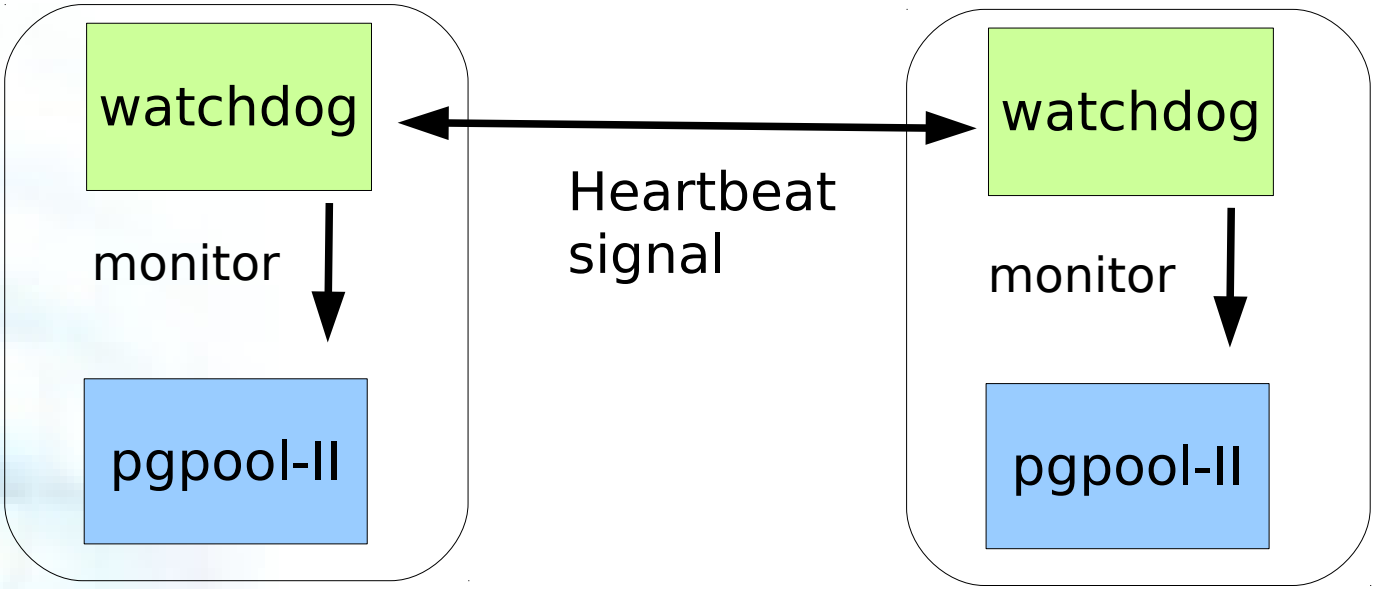

SRA OSS, INC.

# Alive monitoring



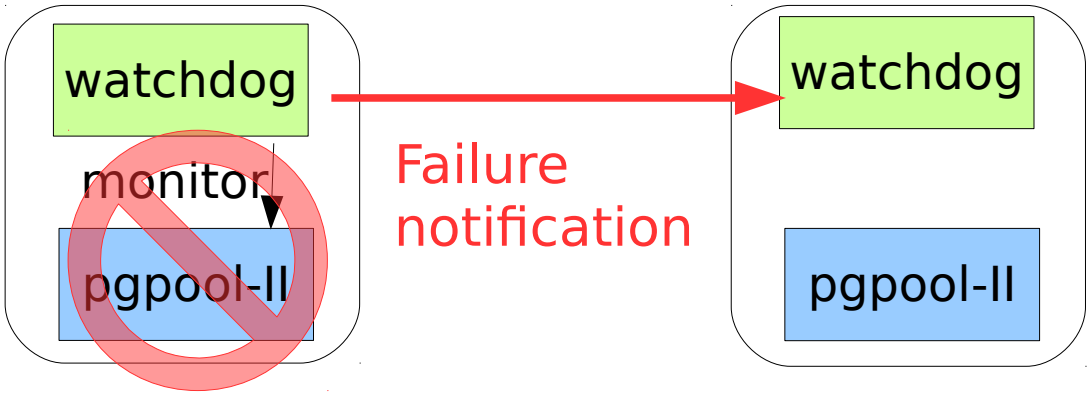
# Alive monitoring (life check) : overview

- Feature to monitor and detect failure of pgpool-II
  - Using heartbeat communication.
    - Heartbeat signal is sent between pgpool-IIs at a regular interval.
    - If a signal isn't received from an other pgpool-II for a time , the pgpool-II is regarded as failed.
  - Also, pgpool-II main process is checked if it's alive.

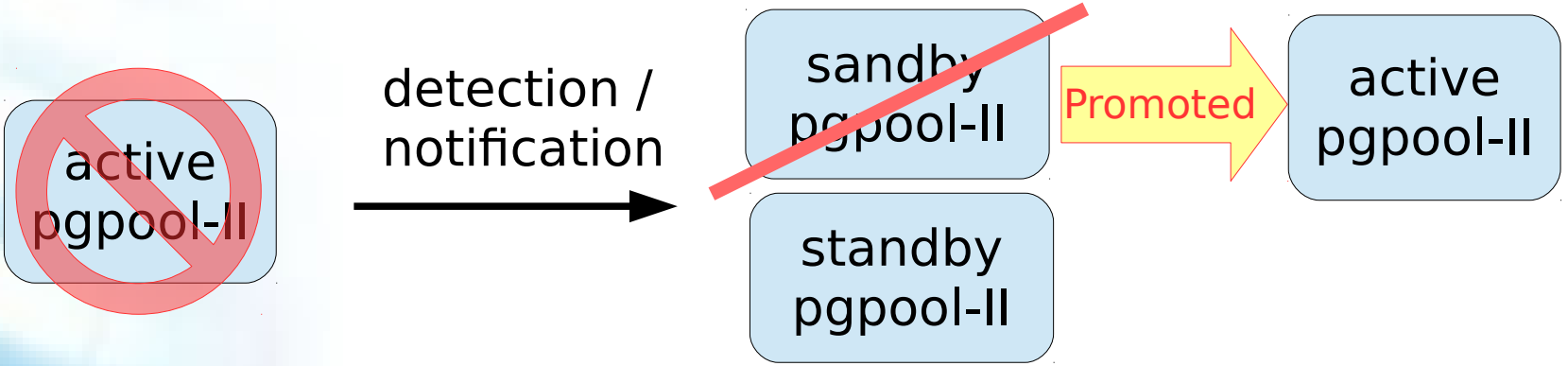


# Alive monitoring : actions after failure is detected

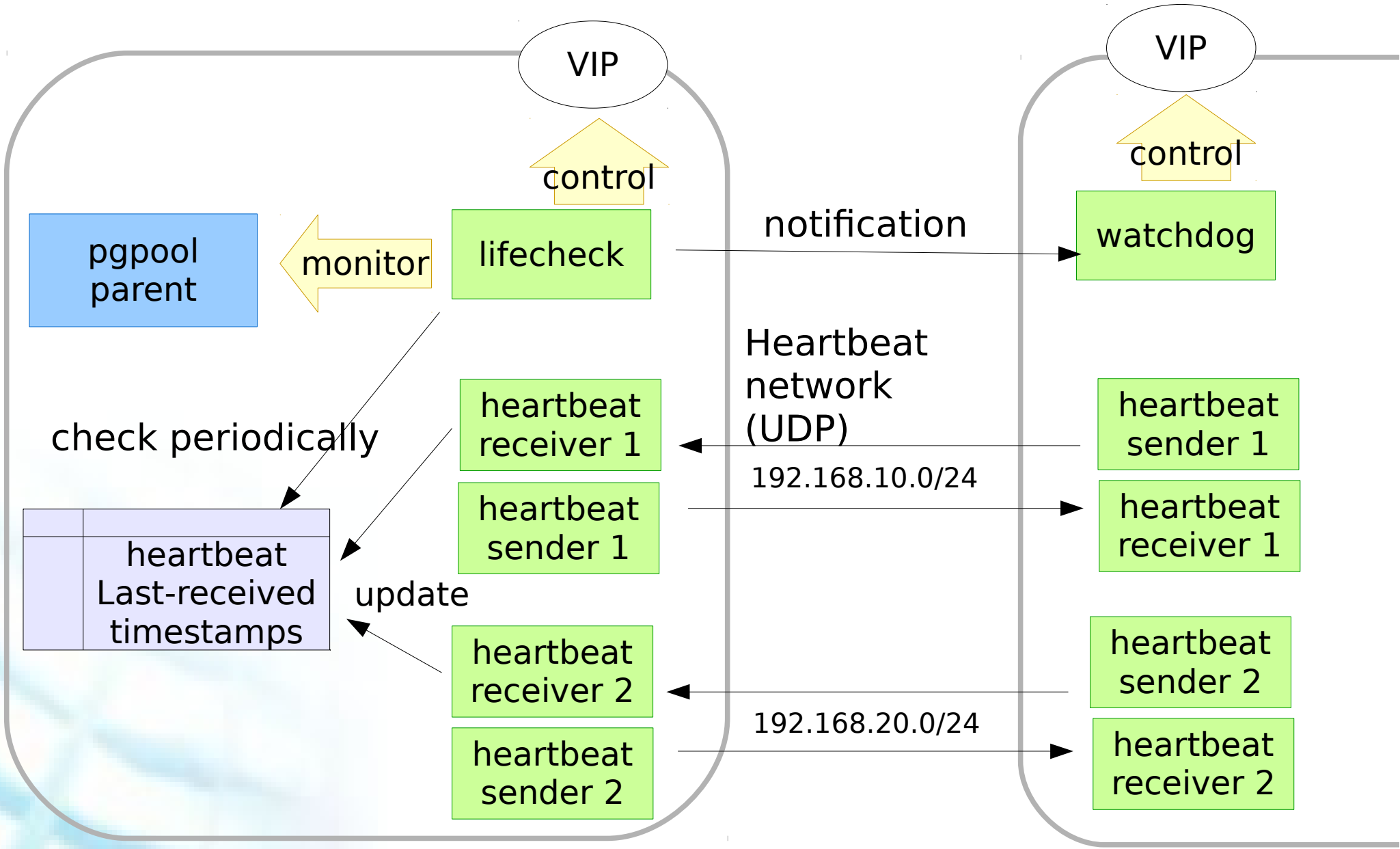
- When detecting a failure of the local pgpool-II
  - Notify other pgpool-IIs of the failure.



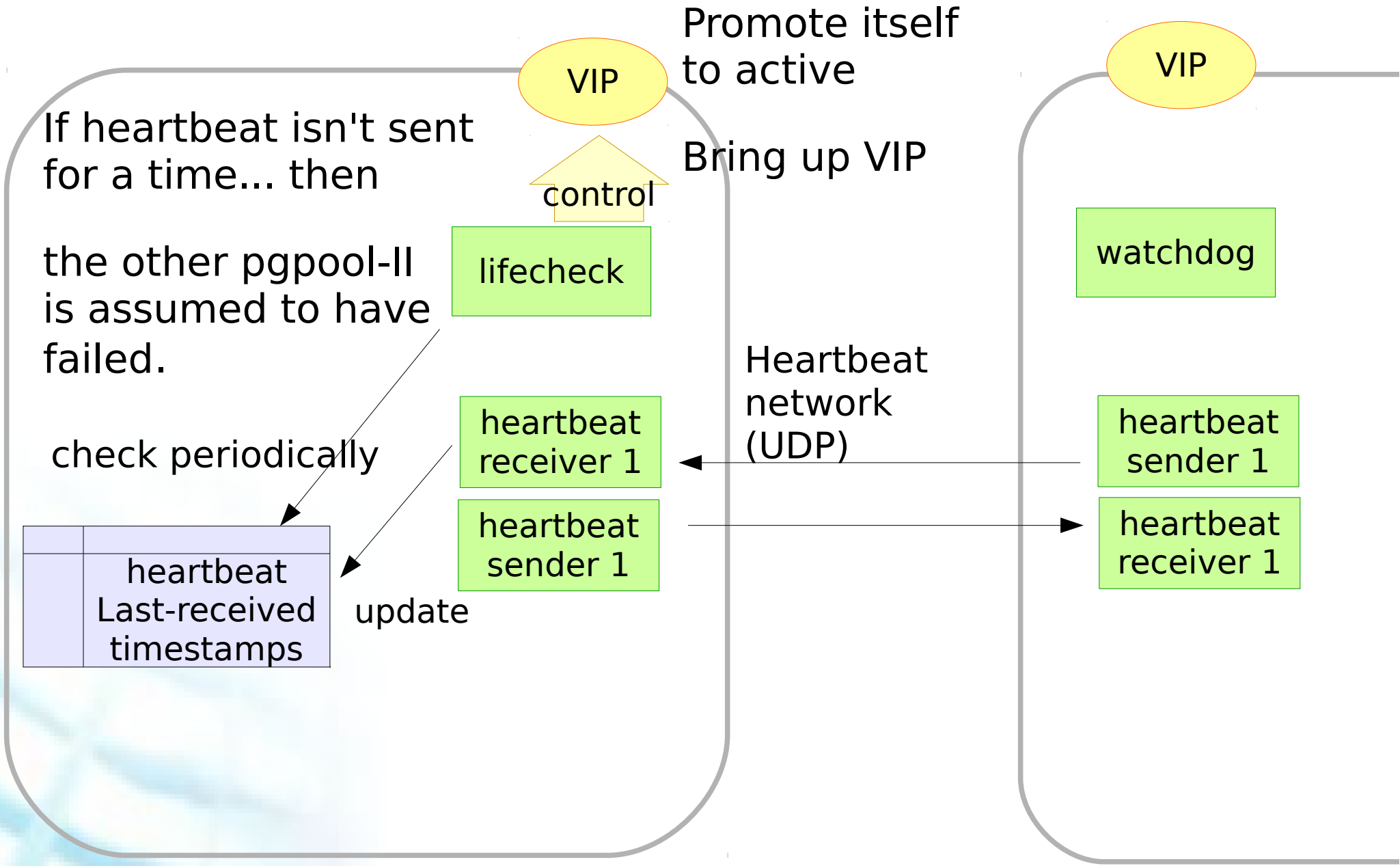

- When a of failure of an other active pgpool-II is detected
  - One of the standby pgpool-IIs is promoted to the active.



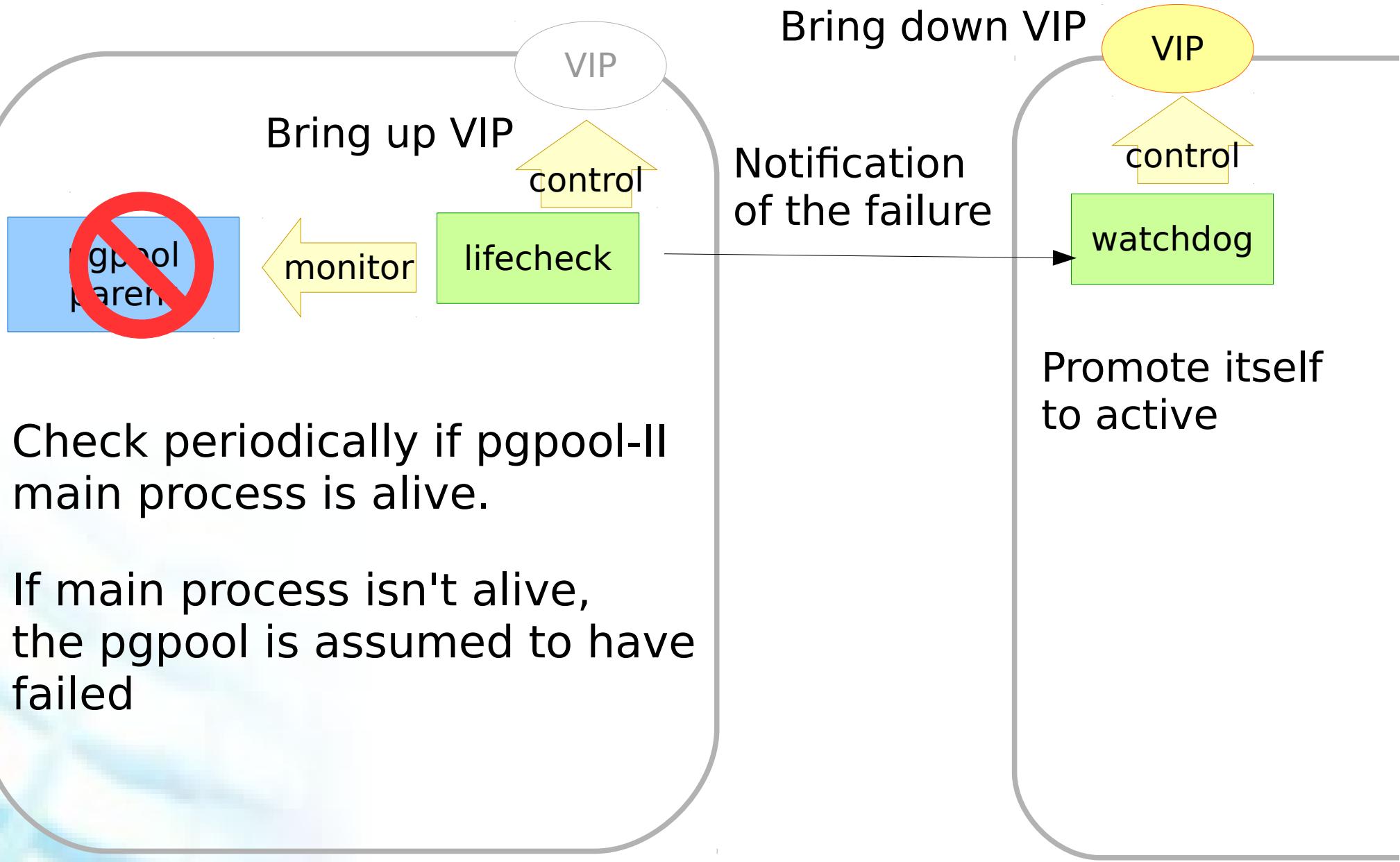
# Alive monitoring : overview



# Alive monitoring : heartbeat (mutual monitoring)



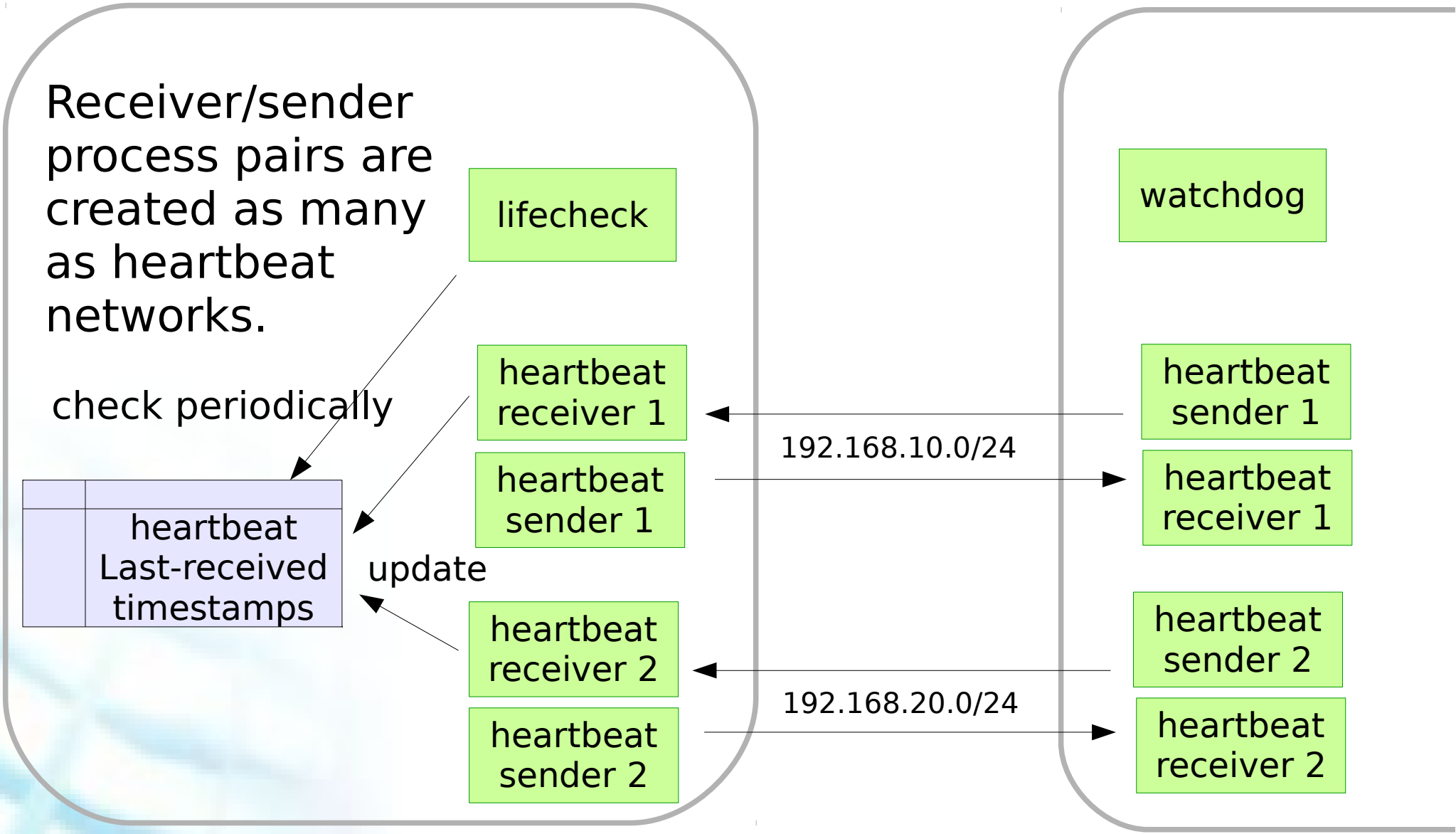
# Alive monitoring : process monitoring (self-monitoring)




Check periodically if pgpool-II main process is alive.

If main process isn't alive, the pgpool is assumed to have failed

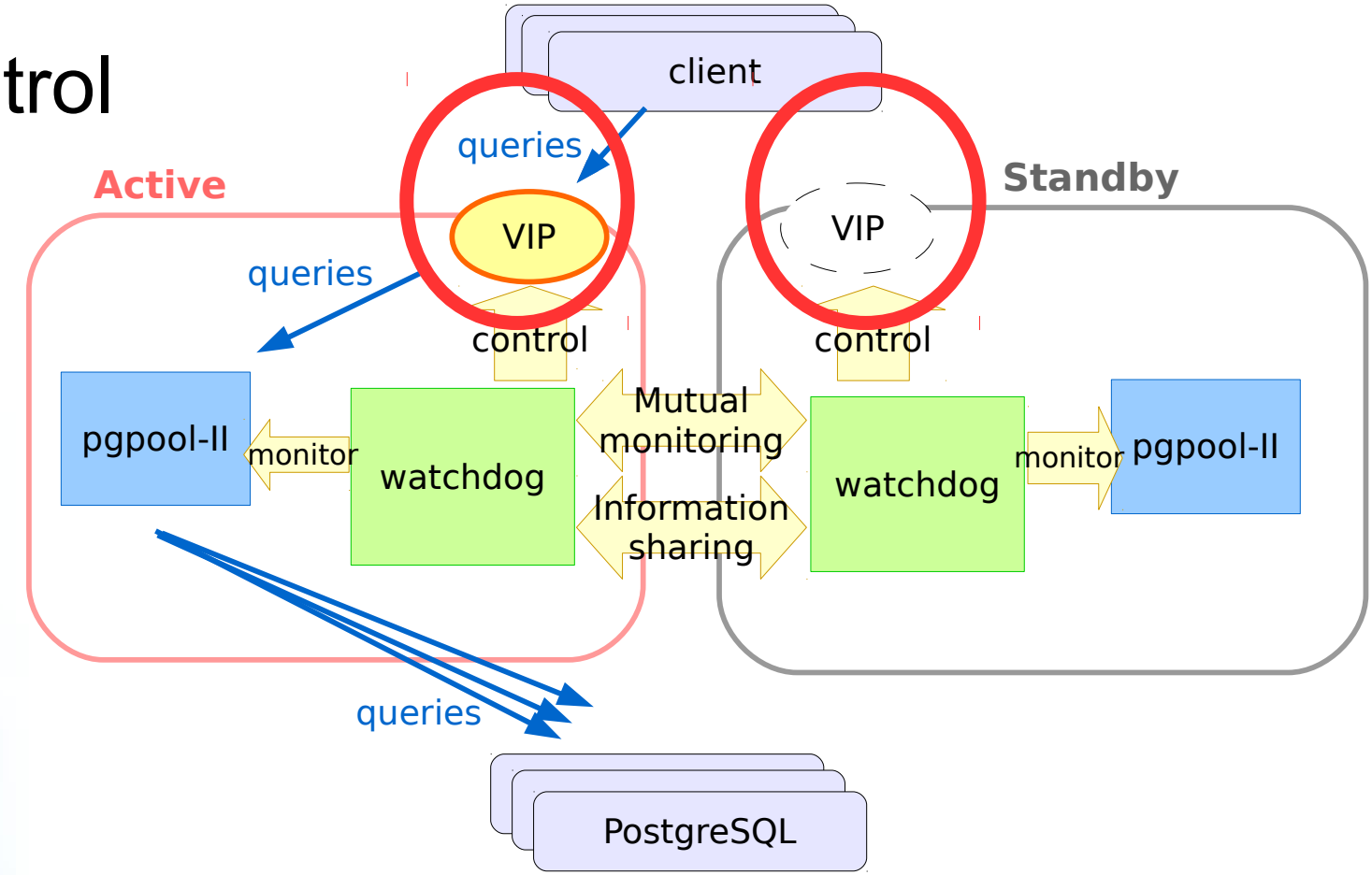
# Alive monitroing: redundant heartbeat





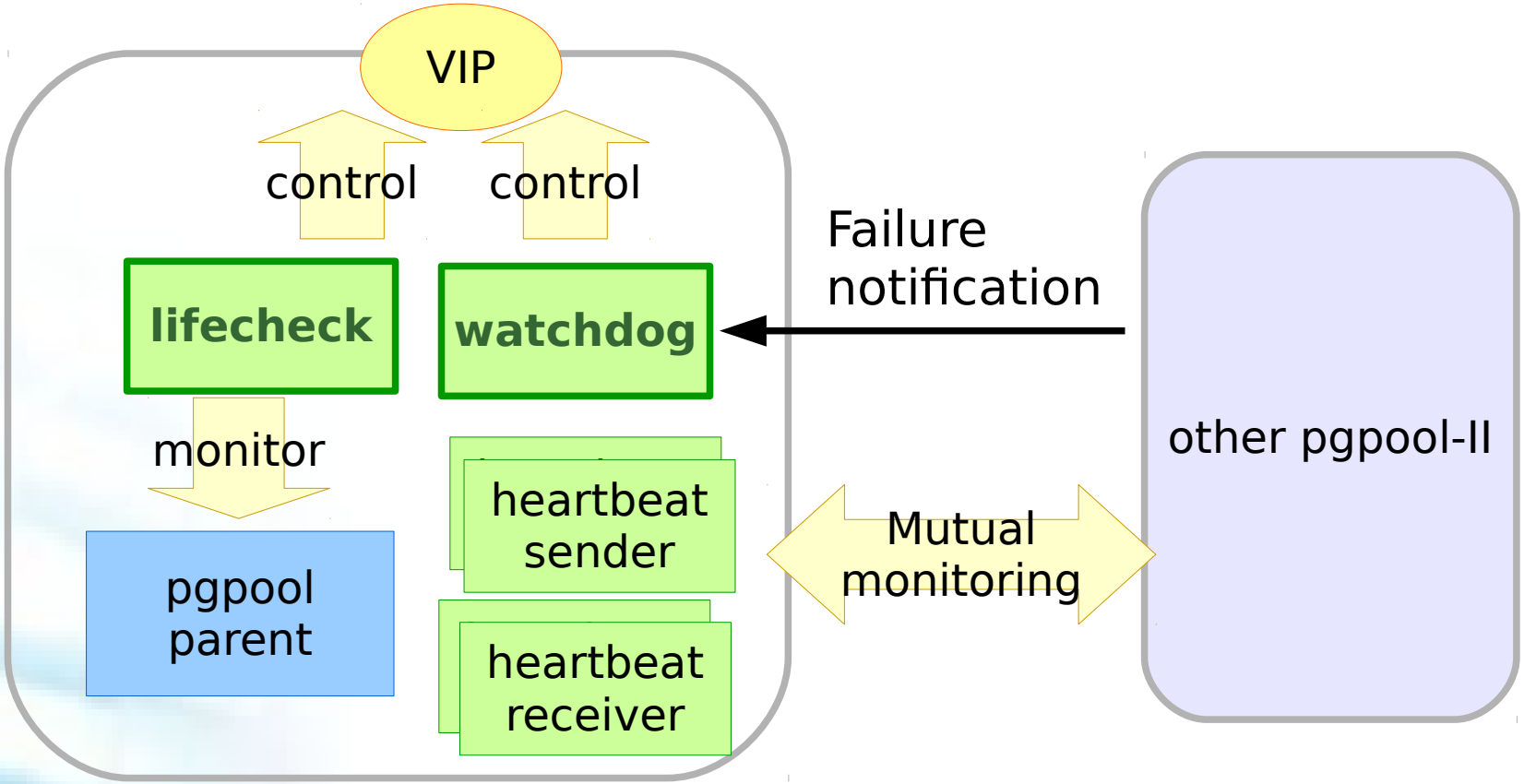

SRA OSS, INC.

# Virtual IP control



# Virtual IP control

- VIP is brought up:
  - When standby pgpool-II is promoted to active
- VIP is brought down:
  - When active pgpool-II goes to down



# Virtual IP control : commands

- Commands executed for bringing up/down VIP
  - ifconfig command is used at default.

    (ip command will be used in the next version.)

- Bringing up the virtual IP

```
if_up_cmd = 'ifconfig eth0:0 inet $_IP_$ netmask 255.255.255.0'
```

- Bringing down the virtual IP

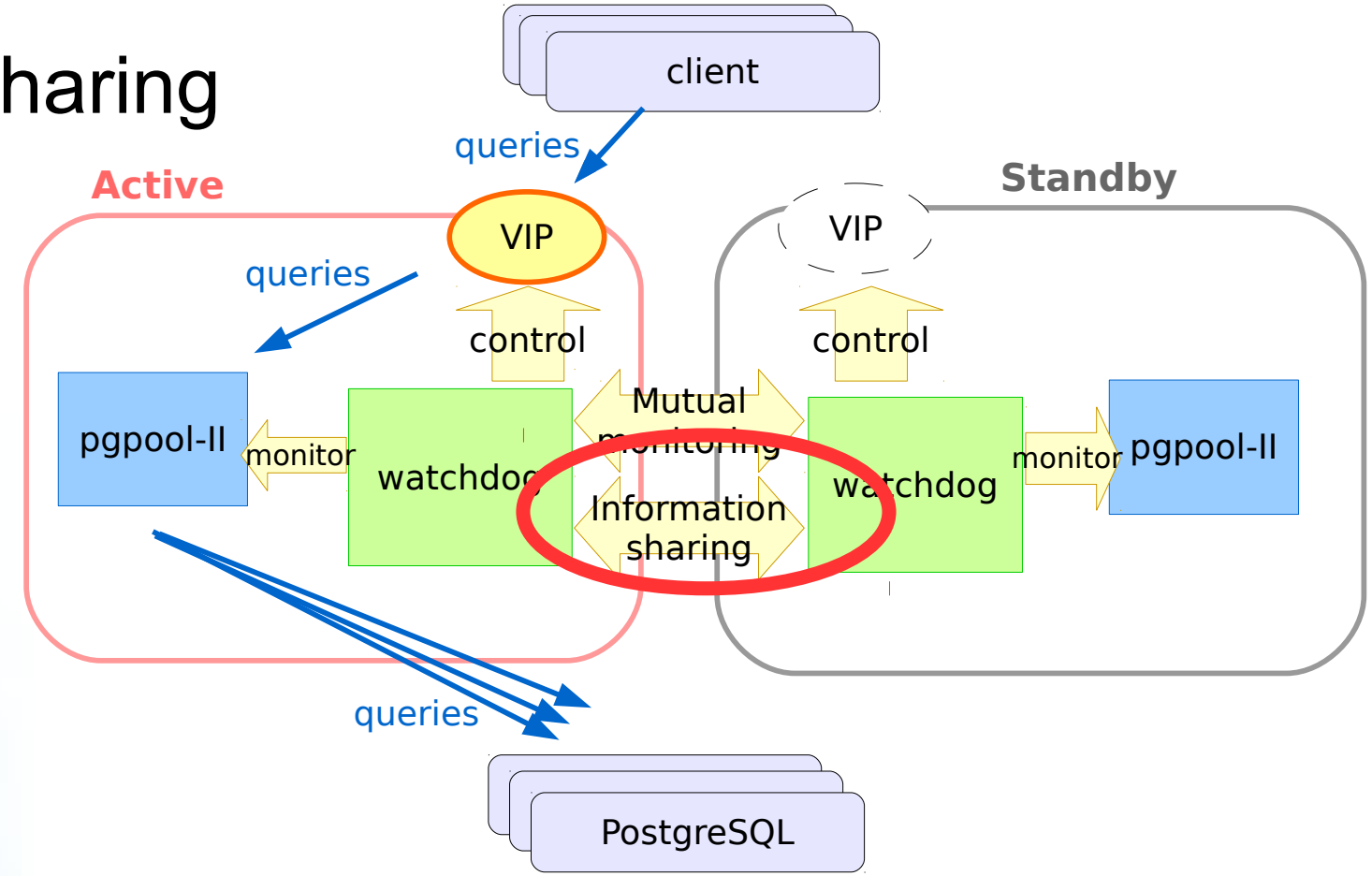
```
if_down_cmd = 'ifconfig eth0:0 down'
```

- Updating neighborhood's ARP cache after bringing up the virtual IP

```
arping_cmd = 'arping -U $_IP_$ -w 1'
```

- You can use your custom command executing AWS CLI etc.
  - More researches and enhancements are needed for controlling a floating IP in cloud network environments.

# Information sharing

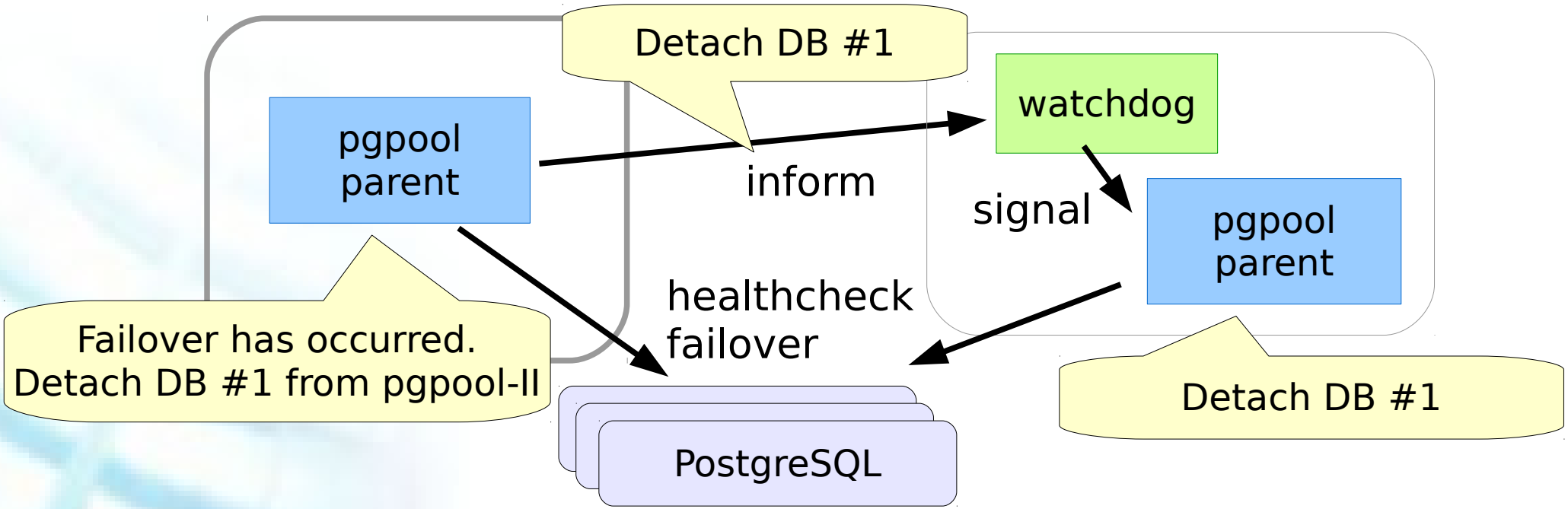


# Information sharing: pgpool-II server info.

- pgpool-IIs share their server information between each other
  - Host name and port number
    - Used for communication and as pgpool-II identifier
  - Status
    - active / standby / down / and so on
  - VIP configuration
    - Required to be the same between all pgpool-II
  - Start-up timestamp
    - Used to determine a new active
- When starting up, each server information is sent to other pgpool-IIs.
  - When restarting, the information is updates.

# Information sharing : DB node info.

- DB node information
  - is managed by pgpool-II
  - Is primary DB or standby DB, and so on.
- When DB failover or failback etc. occurs
  - Inform other pgpool-IIs of the event
    - Event kinds: detach, attach, promote, online-recovery



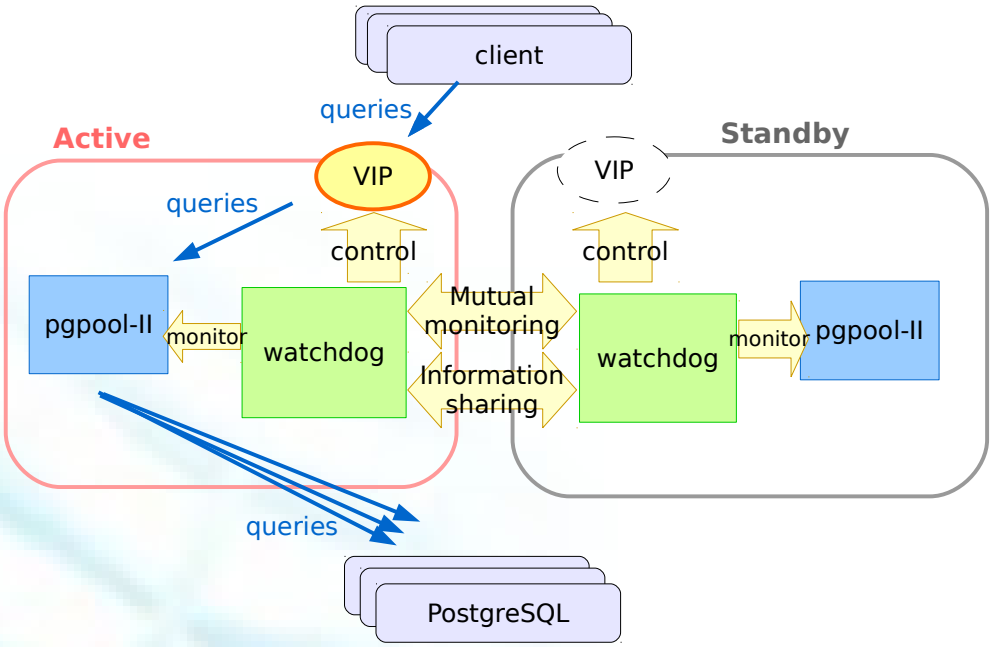
# Reference

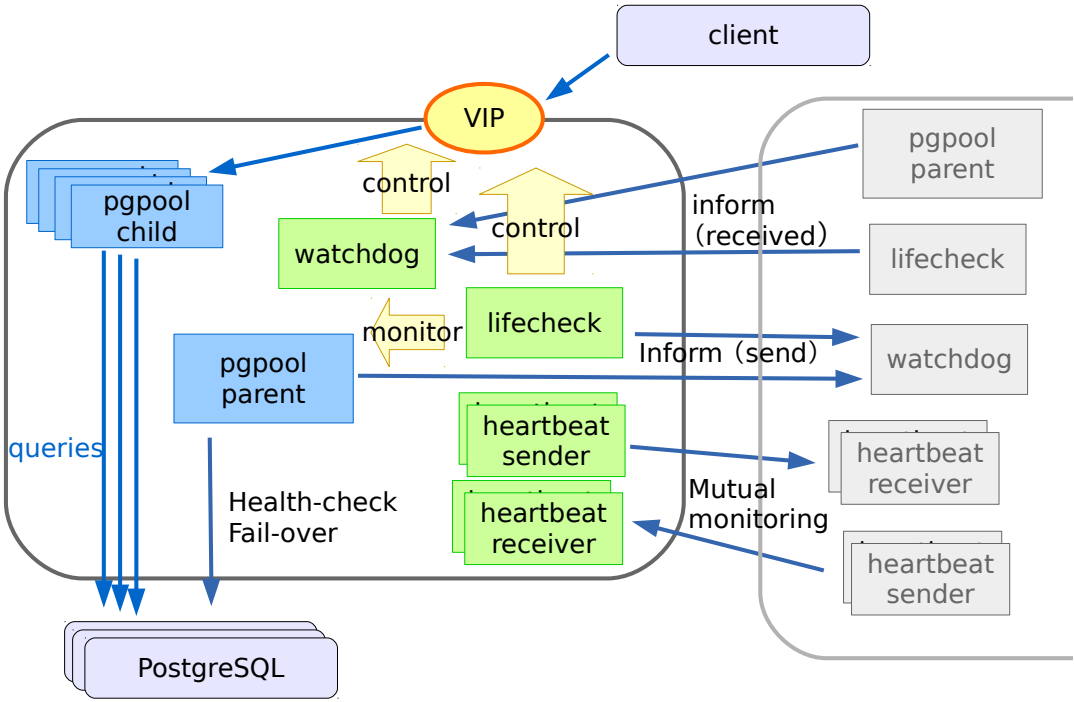
- Pgpool-II official wiki
  - http://www.pgpool.net/ (English)
  - http://www.pgpool.net/jp/ (Japanese)
- SRA OSS, Inc. Japan
  - Seminar slides, Case studies, Technical information (Japanese)

    http://www.sraoss.co.jp/
- Let's Postgres
  - PostgreSQL information web portal (Japanese)
  - http://lets.postgresql.jp/
- Malling Lists
  - pgpool-general-jp@sraoss.jp (Japanese)
  - pgpool-general@pgpool.net (English)

# Summary

- Watchdog's internals
  - Alive monitoring
  - Virtual IP control
  - Information sharing




I you have any questions please feel free to contact us!



SRA OSS, INC.

# Appendix

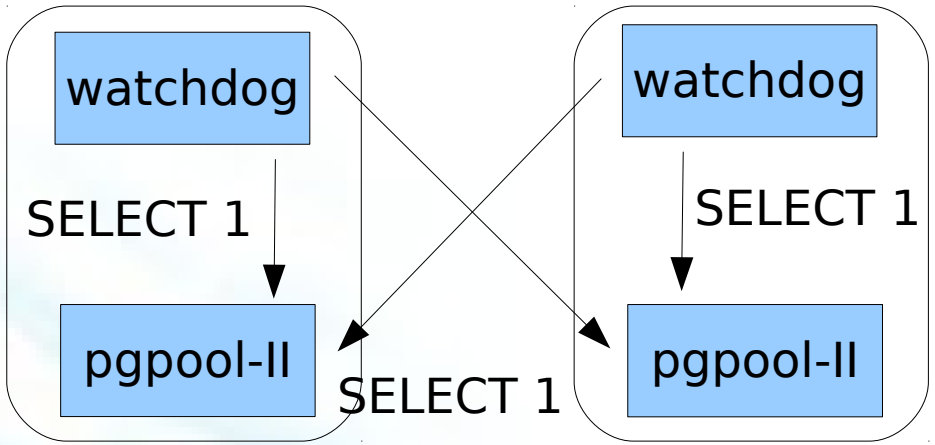
# watchdog Status

- Status:
  - Active
    - pgpool-II is holding Virtual IP.
    - There must be only one active pgpool-II.
  - Standby
    - pgpool-II is not holding Virtual IP and
    - Standby pgpool-II may be promoted to active.
  - Down
    - pgpool-II is regarded as out of service.
- When pgpool-II starts up:
  - The first pgpool-II becomes the active.
  - pgpool-II which starts afterwards sends add request to the active;
  - If this is acceped, the pgpool-II becomes a standby.

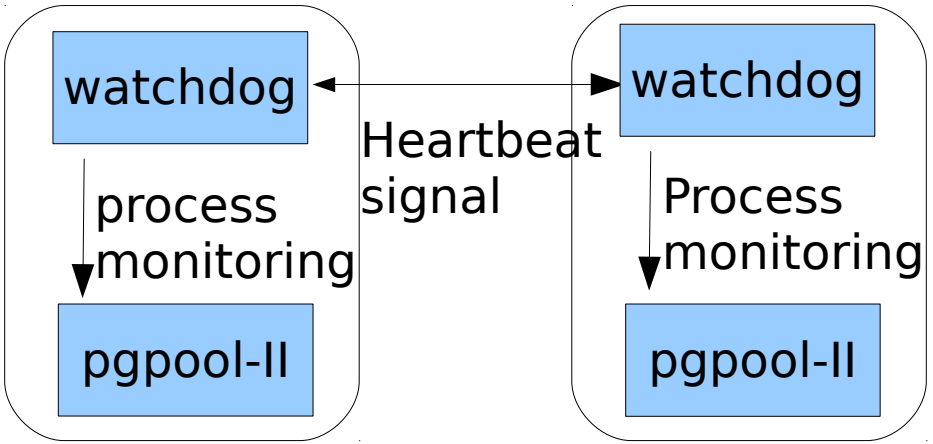
# Alive monitoring: overview

- Feature to monitor and detect failure of pgpool-II
  - There are two modes for life-checking

- Query mode (legacy)
  - Issue a query ('SELECT 1 ' etc.) to pgpool-II and check the response.



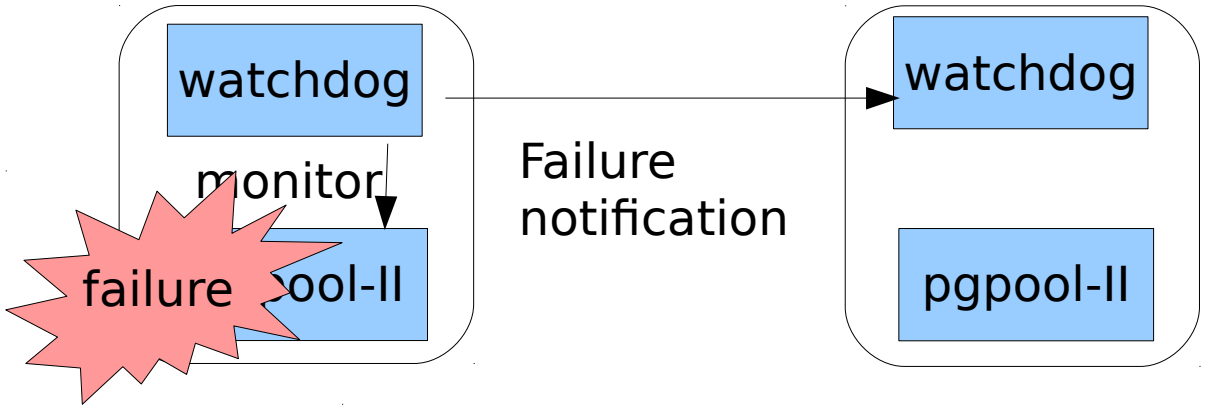

- Heartbeat mode (recommended)
  - Check other pgpool-II's heartbeat signals.



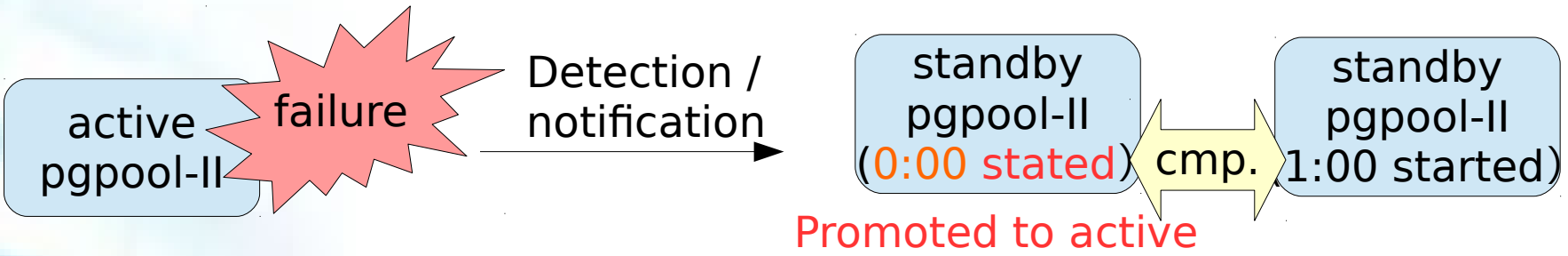

This talk assumes heartbeat mode is used.

# Alive monitoring : actions after failure is detected

- When detecting a failure of the local pgpool-II
  - Notify other pgpool-IIs of the failure




- When detecting a failure of an other active pgpool-II of receiving failure notification from the active pgpool-II
  - One of the standby pgpool-IIs is promoted to the active.
    - pgpool-II which started earliest will be selected.



# Virtual IP control : root privilege

- root privilege is required for controlling virtual IP

- Use one of the following options:

1. Run pgpool-II by a user with root permission.
2. Run pgpool-II by a user with sudo permission.
   - Configure VIP control commands as “sudo ifconfig ...” etc.
3. Set setuid to ifconfig command etc. (recommended)

```
# chmod 4755 /usr/sbin/ifconfig
```

   - This allows normal users to execute the command by root privilege.
   - In practice, it is better to prepare a special ifconfig command which only pgpool-II user can run.

# Information: Secure communication

- Security issue
  - Anyone who knows watchdog protocol can affect pgpool-IIs by spoofing.

- Solution
  - Use authentication key shared by all pgpool-IIs
    - Packets from pgpool-II with invalid key are rejected.