

Pgpool-IIの紹介

～誕生から17年目を迎えたPgpool-IIの過去、
現在、未来について～



2021/1/19

SRA OSS, Inc. 日本支社

石井 達夫

本日のアジェンダ

- Pgpool-IIの生い立ち
- Pgpool-IIの現在
- Pgpool-IIの将来



自己紹介

- 石井達夫
 - SRA OSS, Inc.日本支社で経営者兼OSS開発者をしています
 - Pgpool-IIの開発者、PostgreSQLコミッタ
 - 趣味は音楽鑑賞、SF、下手な料理



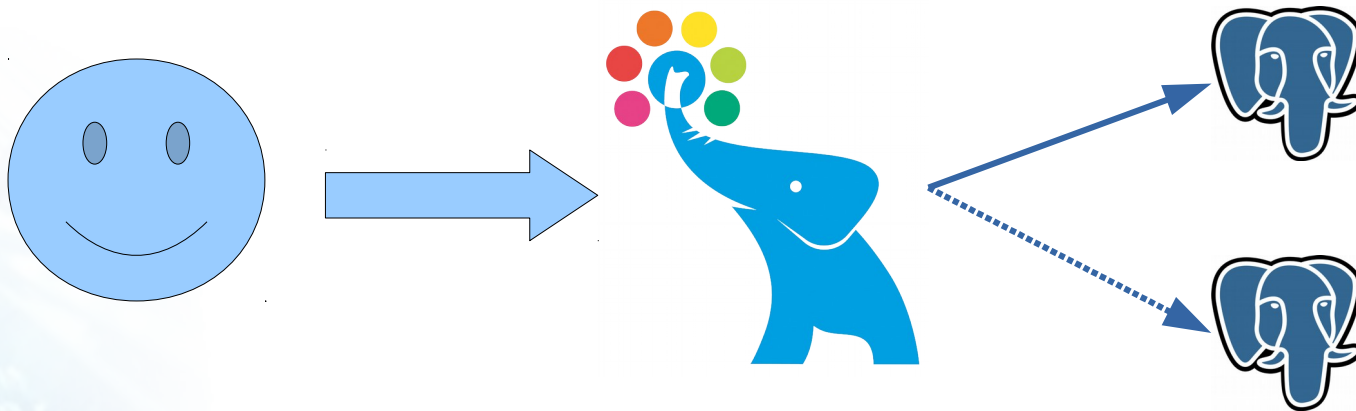
2018年PGConf.EU開催地ポルトガルにて

Pgpool-IIの生い立ち



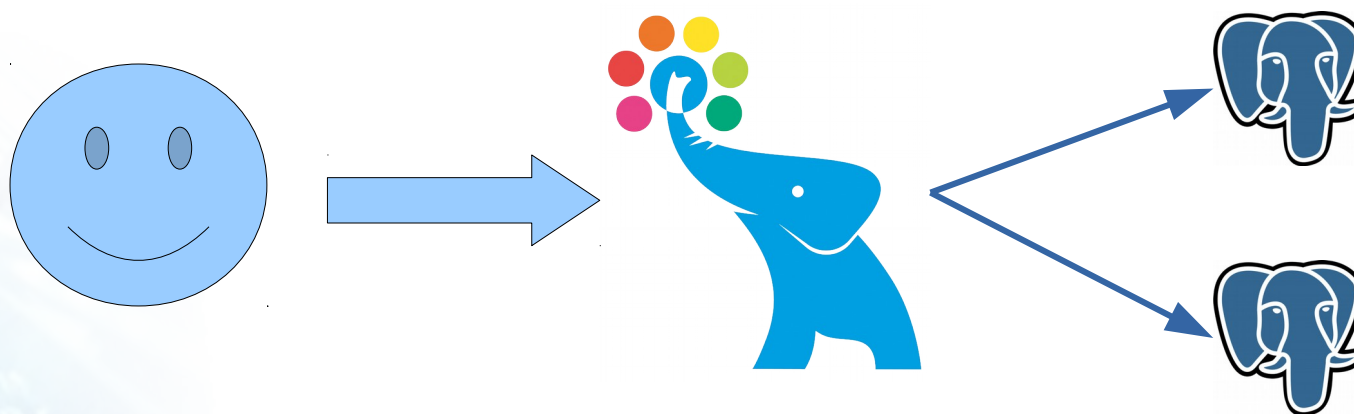
Pgpoolの誕生

- 2003年PostgreSQL用のコネクションプーリングソフトとして公開
- フェイルオーバー機能も同時に実装
- バックアップ用のPostgreSQLサーバの指定も可能



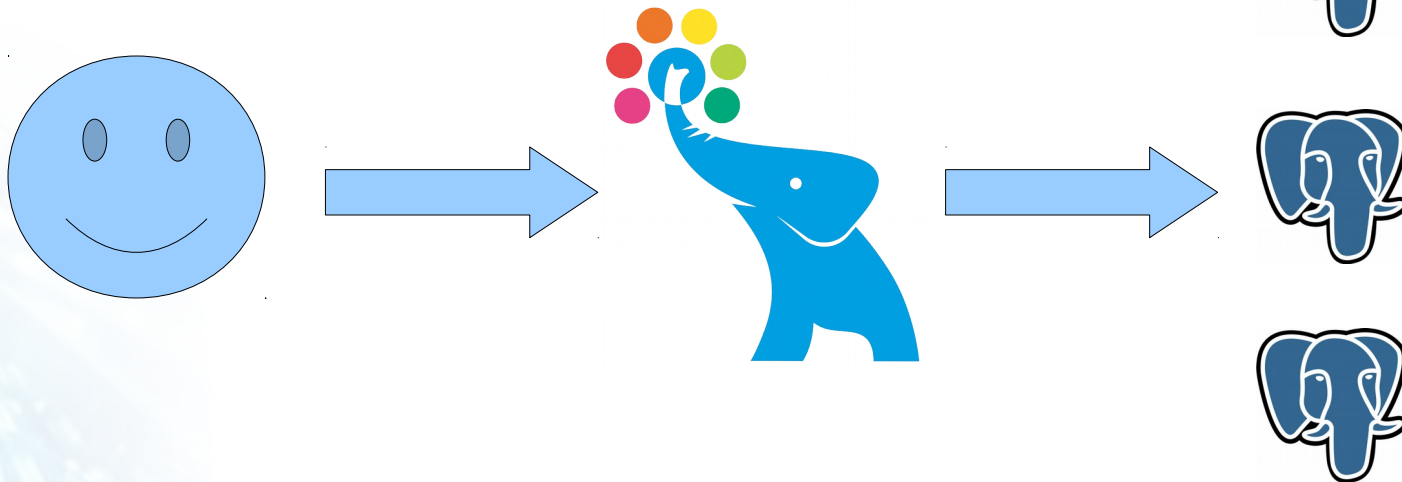
レプリケーションの実装

- 2004年現在の「ネイティブレプリケーションモード」に相当する機能を実装
- サーバは2台まで



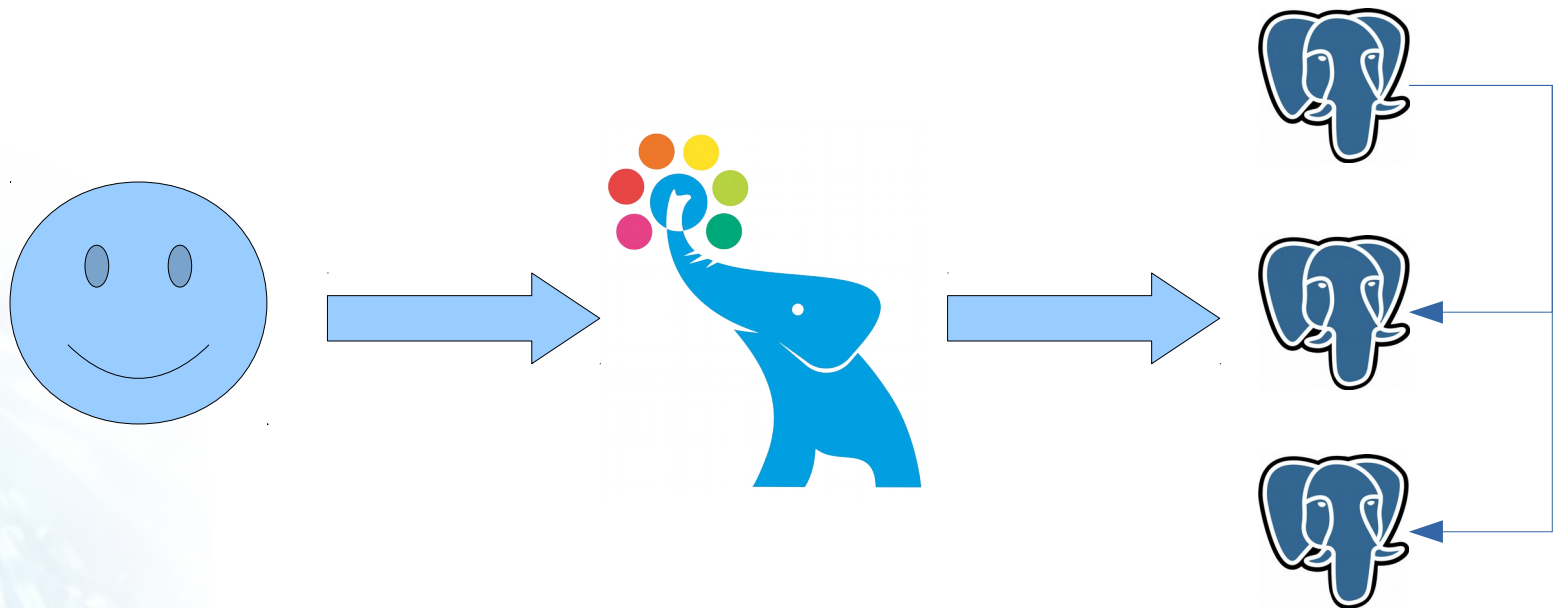
Pgpool-IIの誕生(2006)

- 個人プロジェクトからOSSプロジェクトへ
- サーバ台数の制限を撤廃
- SQLパーサを搭載
- 現在のPgpool-IIの基礎



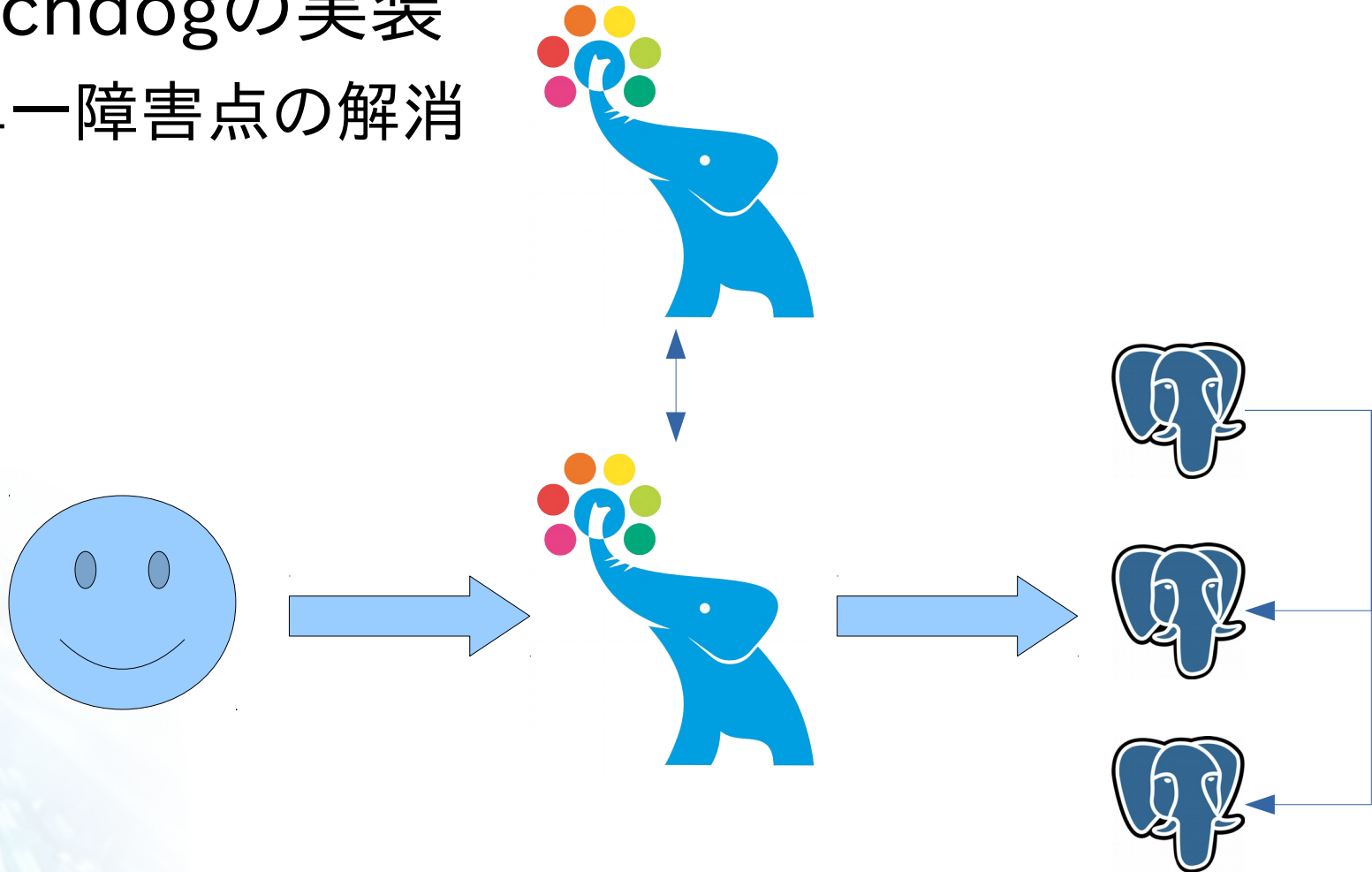
Pgpool-II 3.0 (2010)

- ストリーミングレプリケーションへの対応



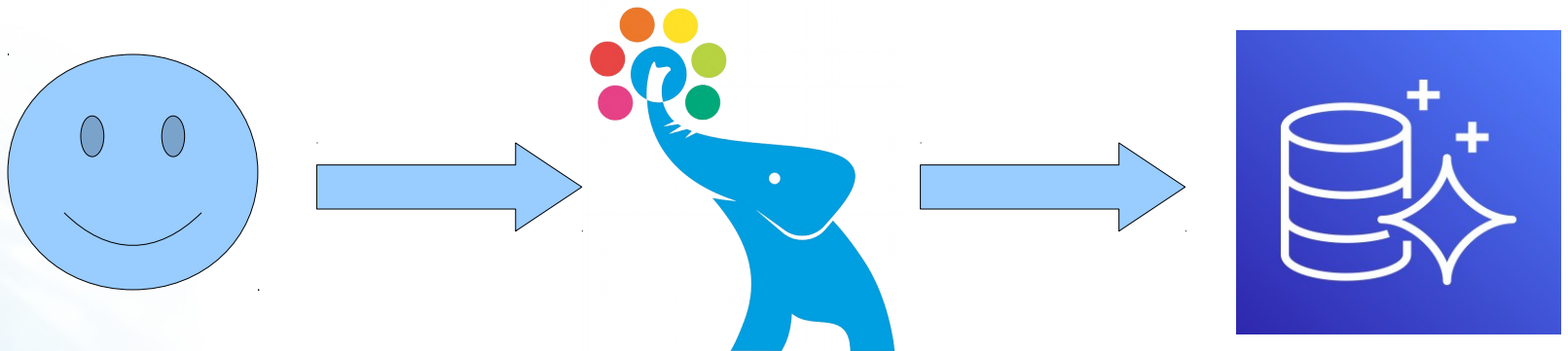
Pgpool-II 3.2 (2012)

- Watchdogの実装
 - 単一障害点の解消



Amazon Auroraへの対応(2017)

- Pgpool-IIは負荷分散を担当
- フェイルオーバーはAuroraに任せる



その後のPgpool-II

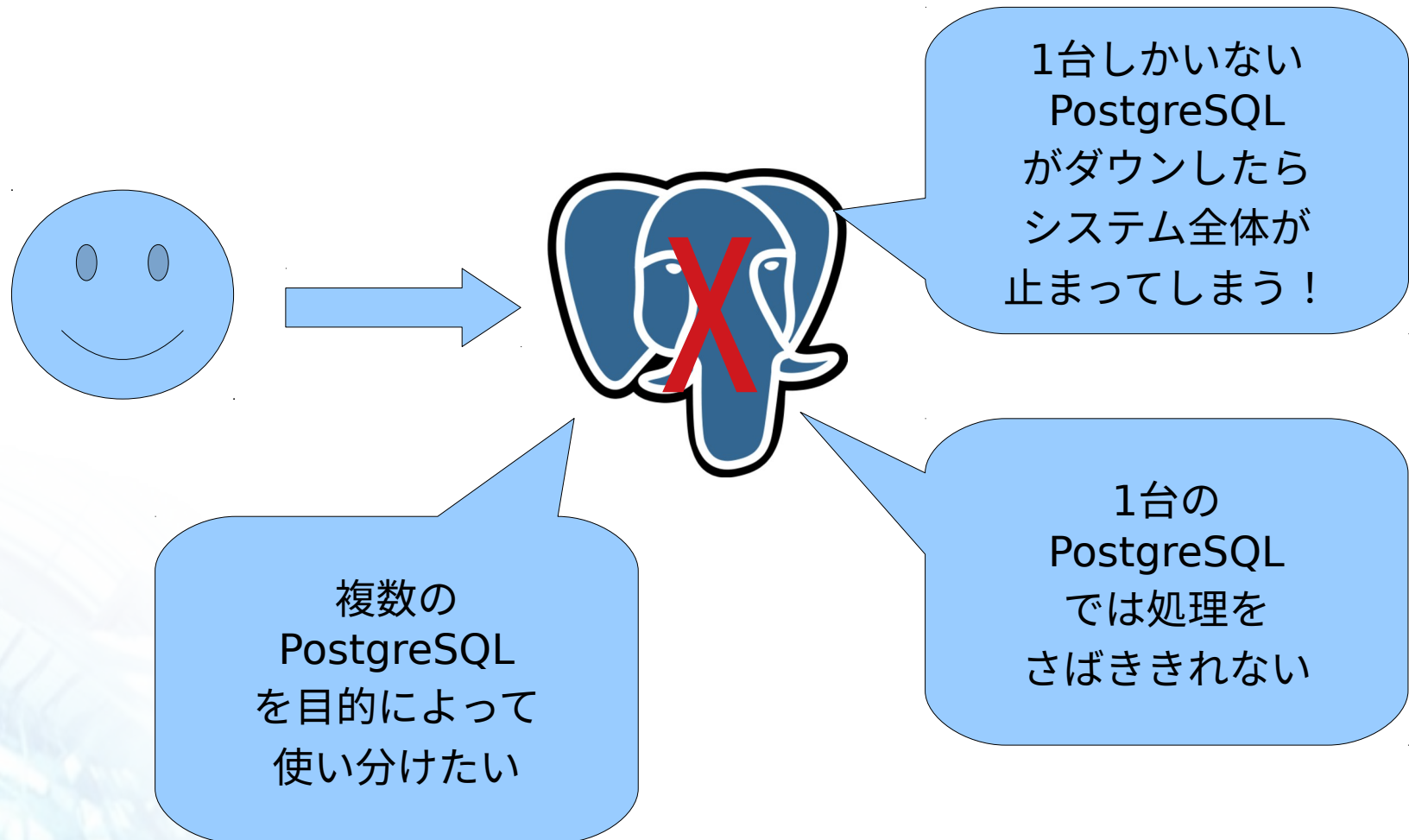
- 2018年 Pgpool-II 4.0リリース
 - SCRAM認証、証明書認証
 - AES256形式のパスワード
 - 負荷分散方式の改良
 - 不正プライマリの検出
- 2019年 Pgpool-II 4.1リリース
 - 文レベルの負荷分散
 - 自動フェイルバック
 - 共有リレーションキャッシュ

- 2020年 Pgpool-II 4.2リリース
 - LDAP認証
 - ログ収集プロセス
 - ヘルスチェックとSQLの統計情報
 - pgpool.confの記述のシンプル化
 - 設定ファイル、ドキュメント、ソースコードで使用する用語を適切化

Pgpool-IIの現在

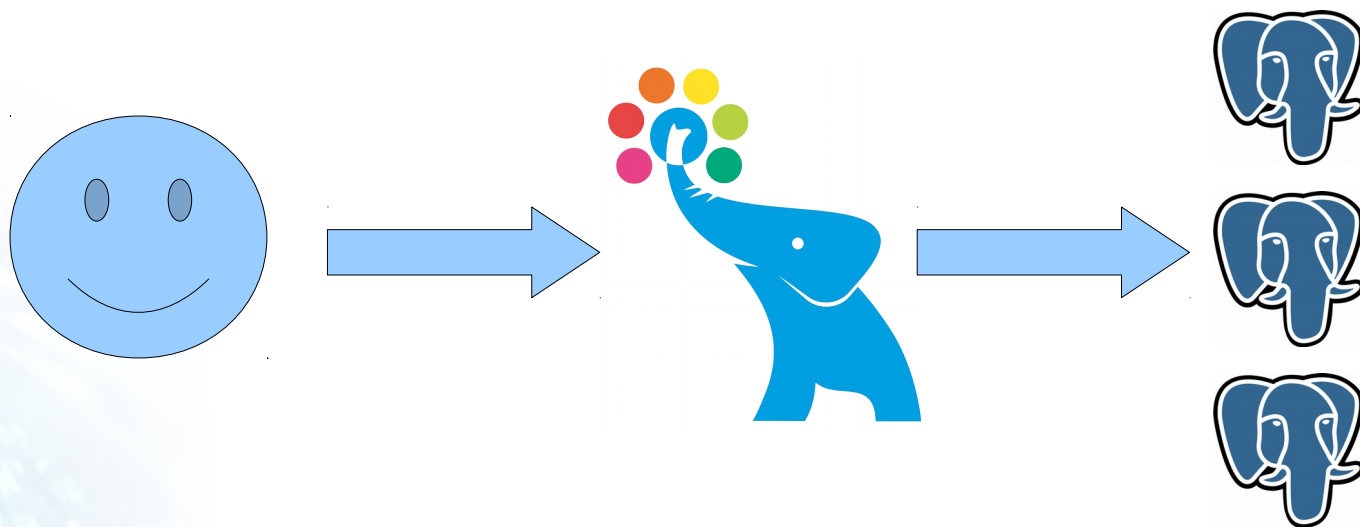


そもそもなぜクラスタは必要なの？

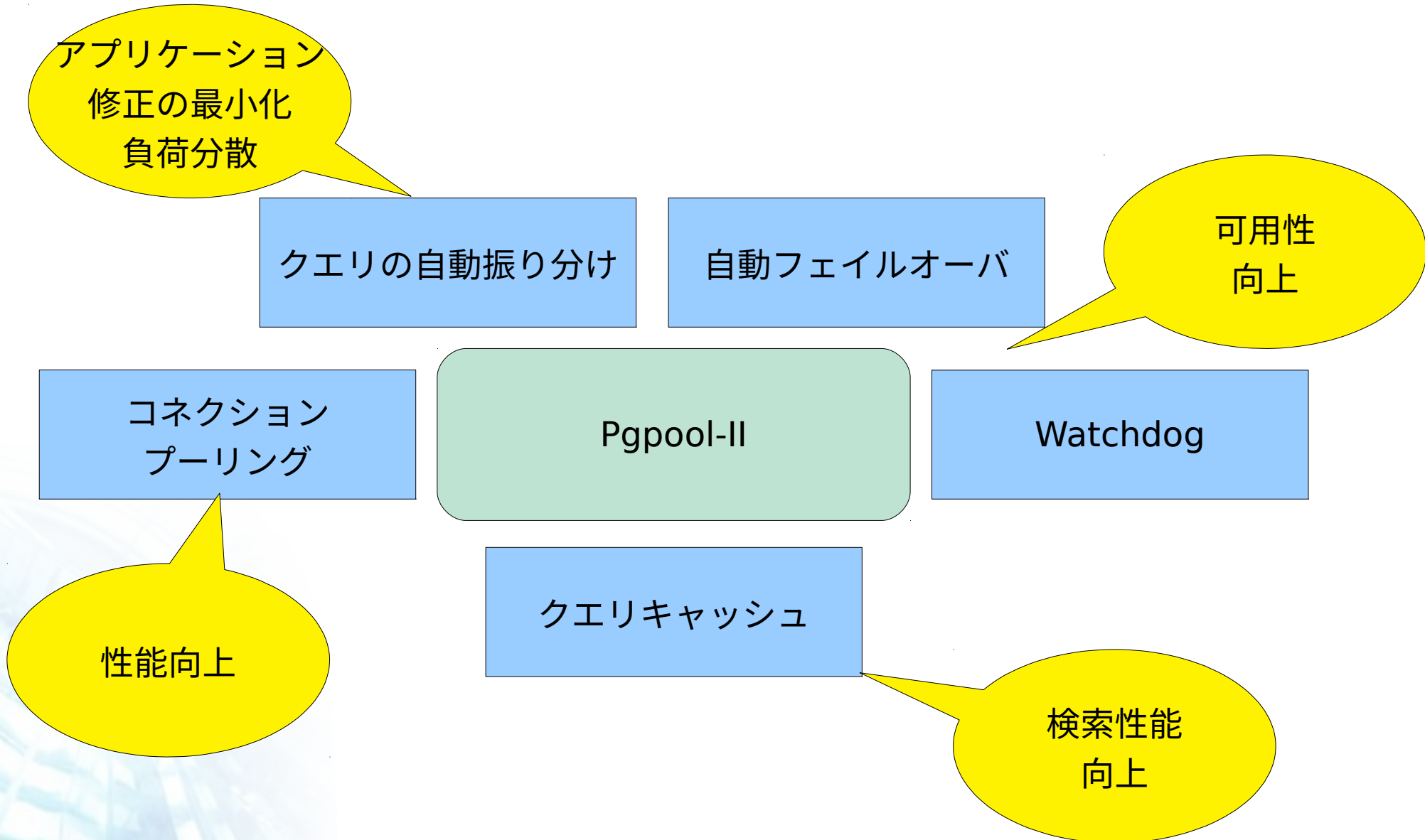


Pgpool-IIのコンセプト

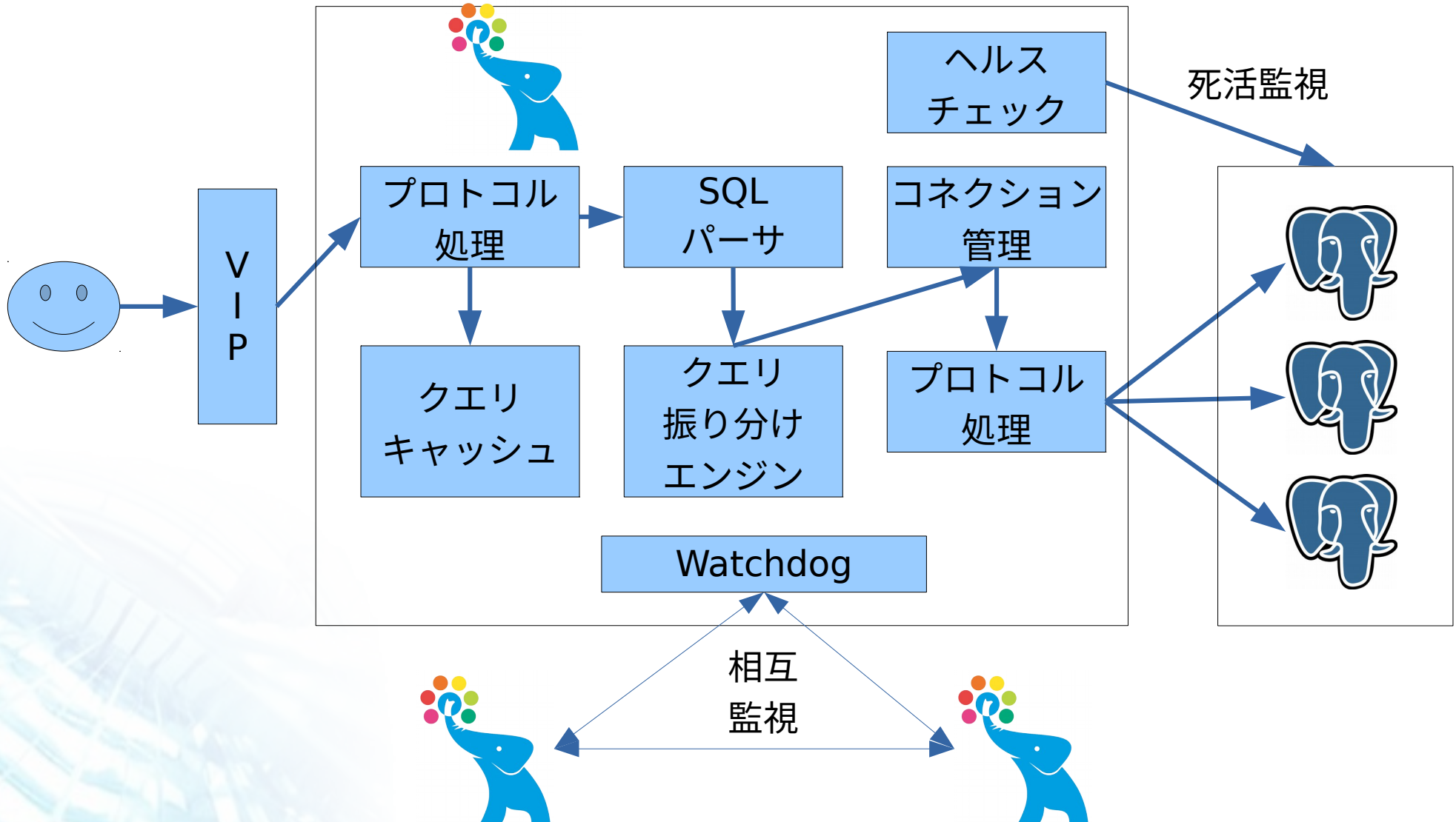
- ✓ ユーザから見て、なるべく複数PostgreSQLによるクラスタを使っていることを意識させたくない
- ✓ シングルPostgreSQL用のアプリケーションがなるべく変更なしに複数PostgreSQLによるクラスタを使う環境に移行できるようにしたい



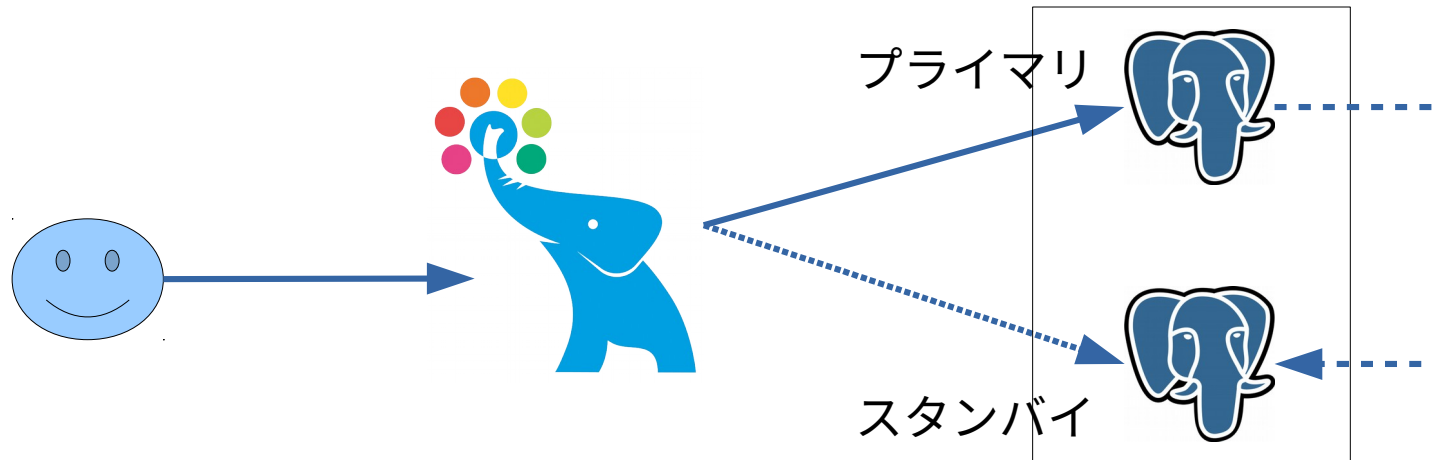
Pgpool-IIの機能



Pgpool-IIの構造

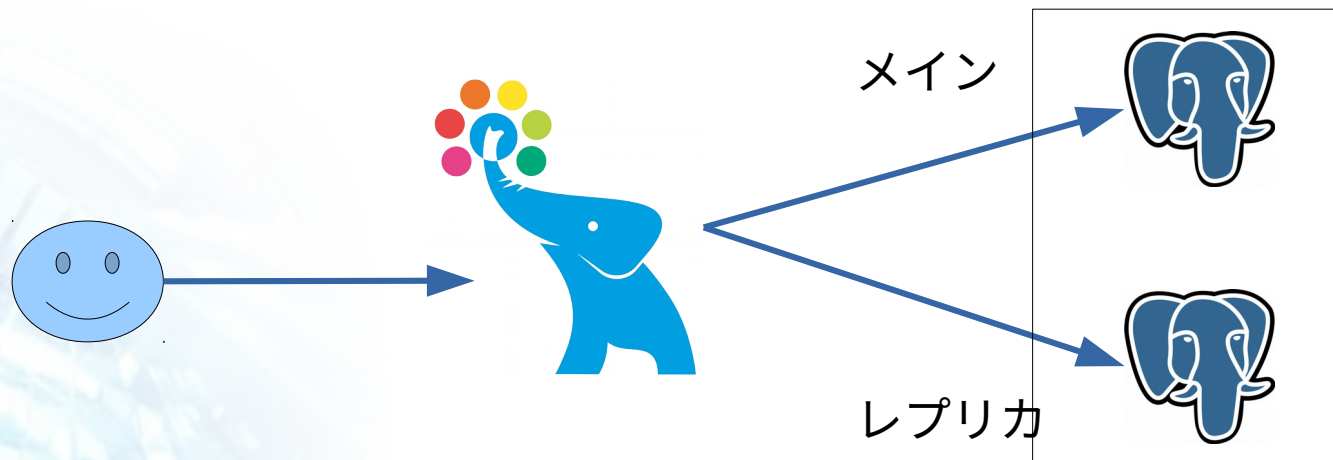


Pgpool-IIの機能: クエリの振り分け



本日主に
取り上げるのは
こちら

ストリーミング
レプリケーション
モード
更新クエリを
プライマリにのみ
投げる



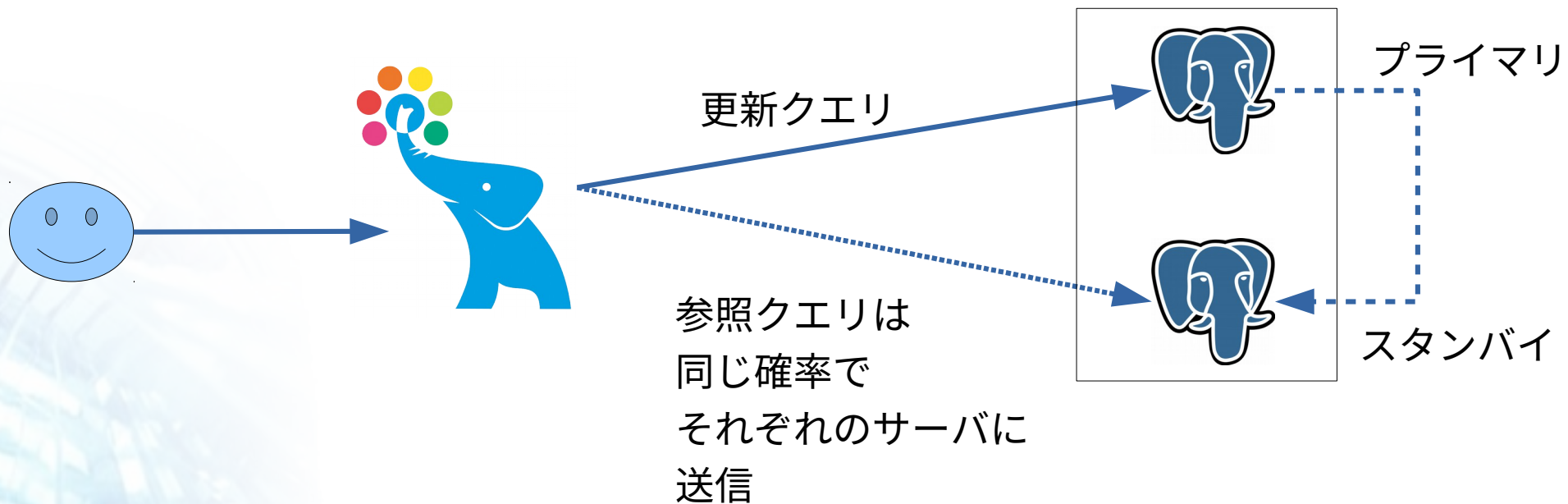
ネイティブ
レプリケーション
モード、SIモード
更新クエリを
両方に投げる

Pgpool-IIの機能:クエリの振り分け

- Pgpool-IIを使いたい最大の理由がクエリの振り分け
 - 他のソフトでは実現できないPgpool-IIだけの強み
 - スタンバイに投げてはいけないクエリをプライマリにのみ投げる
 - 更新クエリ
 - 強いロック
 - シリアライザブルトランザクション
 - その他
 - それ以外のクエリはスタンバイに投げる
 - クエリ振り分けのメリット
 - シングルPostgreSQL用のアプリケーションをストリーミングレプリケーション環境にそのまま移行できる
 - スタンバイのリソースを有効活用できる
 - プライマリサーバの負担を軽減する

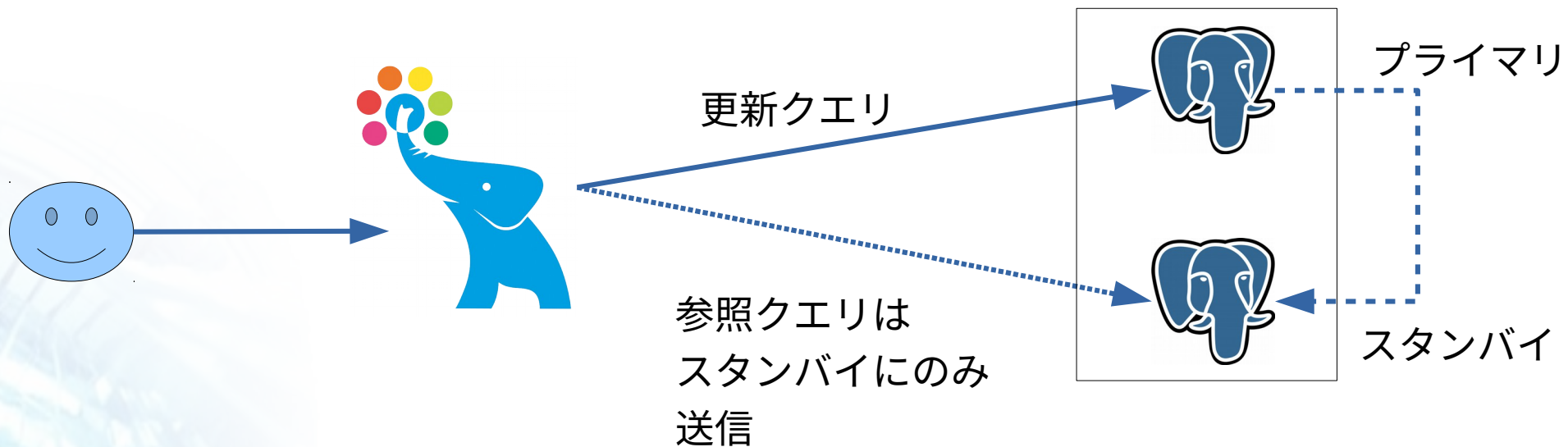
クエリの振り分けの設定

- 負荷分散の重み付けが設定可能
 - `backend_weight0 = 0.5`
 - `backend_weight1 = 0.5`



クエリの振り分けの設定

- プライマリサーバを更新処理に専念させ、負荷を下げる設定の例
 - `backend_weight0 = 0.0`
 - `backend_weight1 = 1.0`

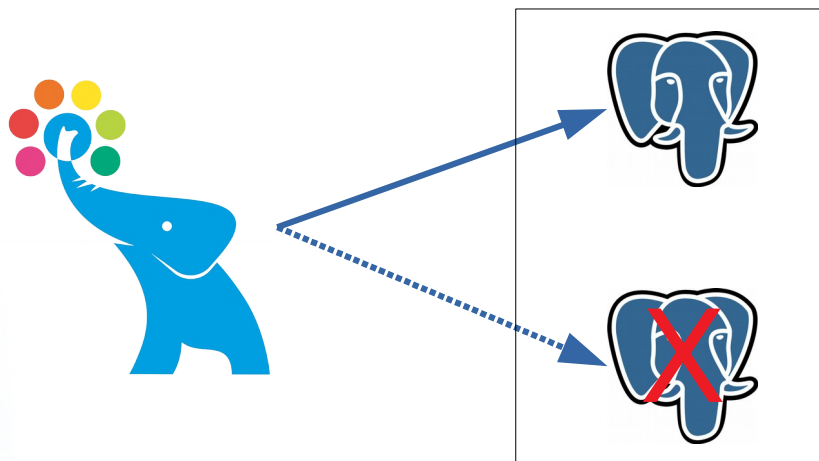


その他のクエリ振り分けの設定

- `statement_level_load_balance`
 - クエリの振り分けをセッション単位ではなく、SQL文単位で決定
- `database_redirect_preference_list`
 - データベース名によってクエリの振り分け先を決定
- `database_redirect_preference_list`
 - アプリケーション名によってクエリの振り分け先を決定
- `disable_load_balance_on_write`
 - トランザクション内のクエリの振り分けの挙動を制御

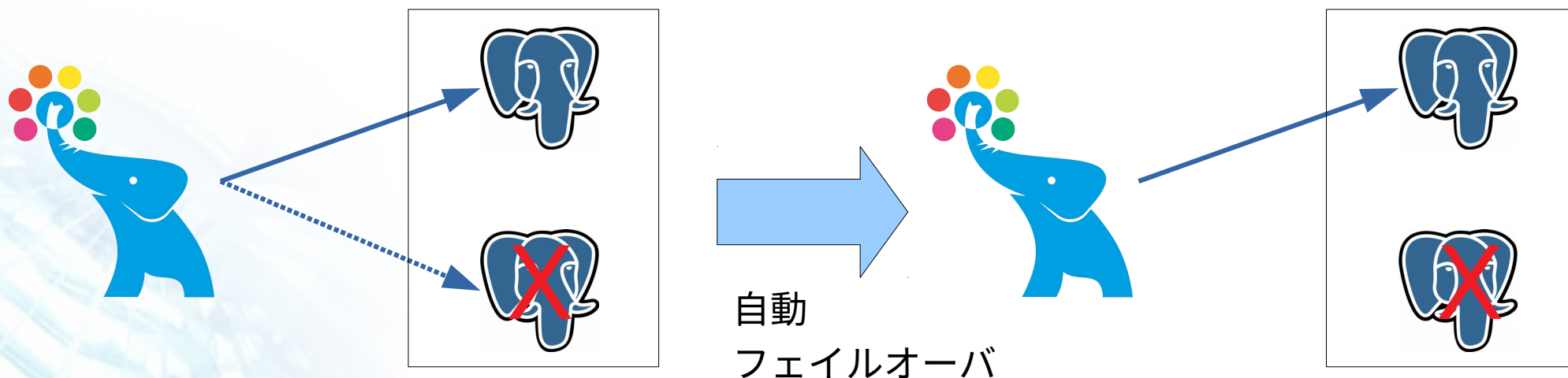
Pgpool-IIの機能:ヘルスチェック

- 定期的にPostgreSQLに接続し、PostgreSQLが正常に稼働しているかどうかを確認
- 接続ができなければ自動的にそのサーバを切り離すリクエストを出す(次のスライド)



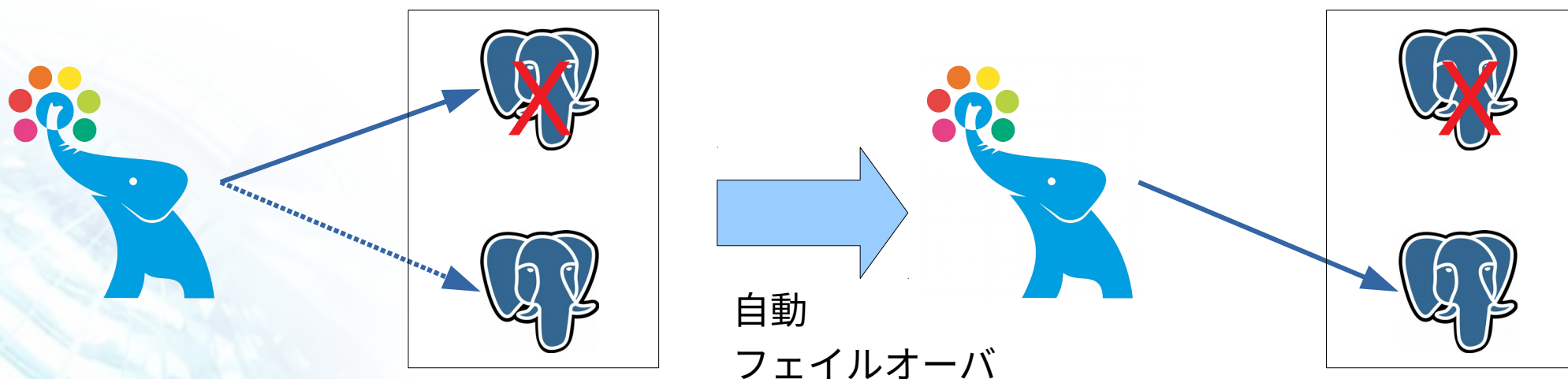
Pgpool-IIの機能：自動フェイルオーバ

- PostgreSQLサーバがダウンしたらそのサーバを切り離し、残りのサーバで処理を継続
 - 定期的にヘルスチェックプロセスがPostgreSQLサーバと通信ができるかどうかを確認
 - 通信ができなければ自動的にそのサーバを切り離す



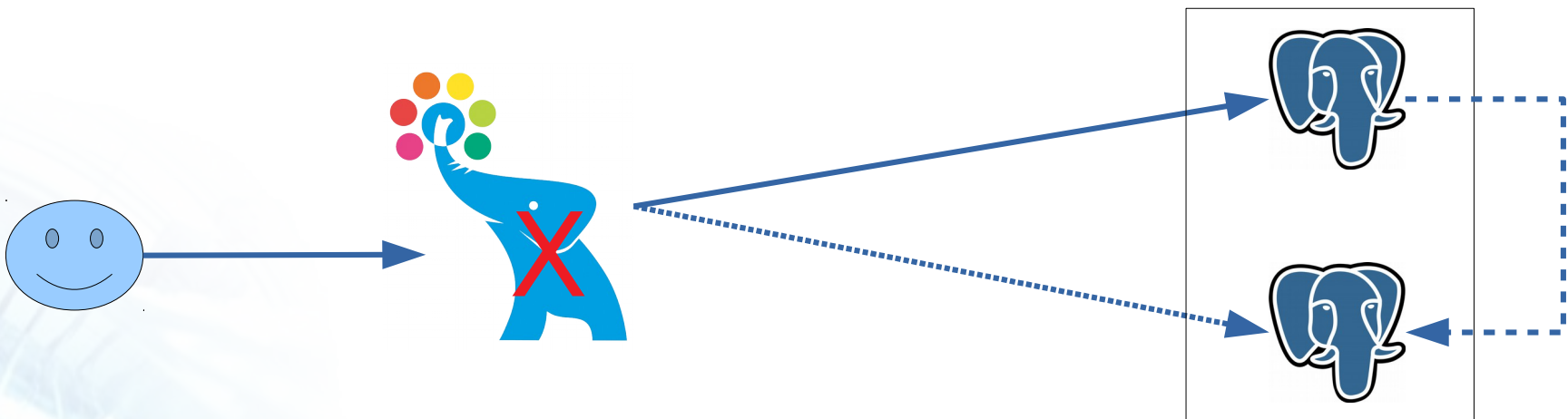
Pgpool-IIの機能：自動フェイルオーバー

- プライマリサーバがダウンした時は、スタンバイサーバを自動昇格して新しいプライマリサーバで処理を続行する
 - 昇格処理を行うカスタムスクリプトを指定できる



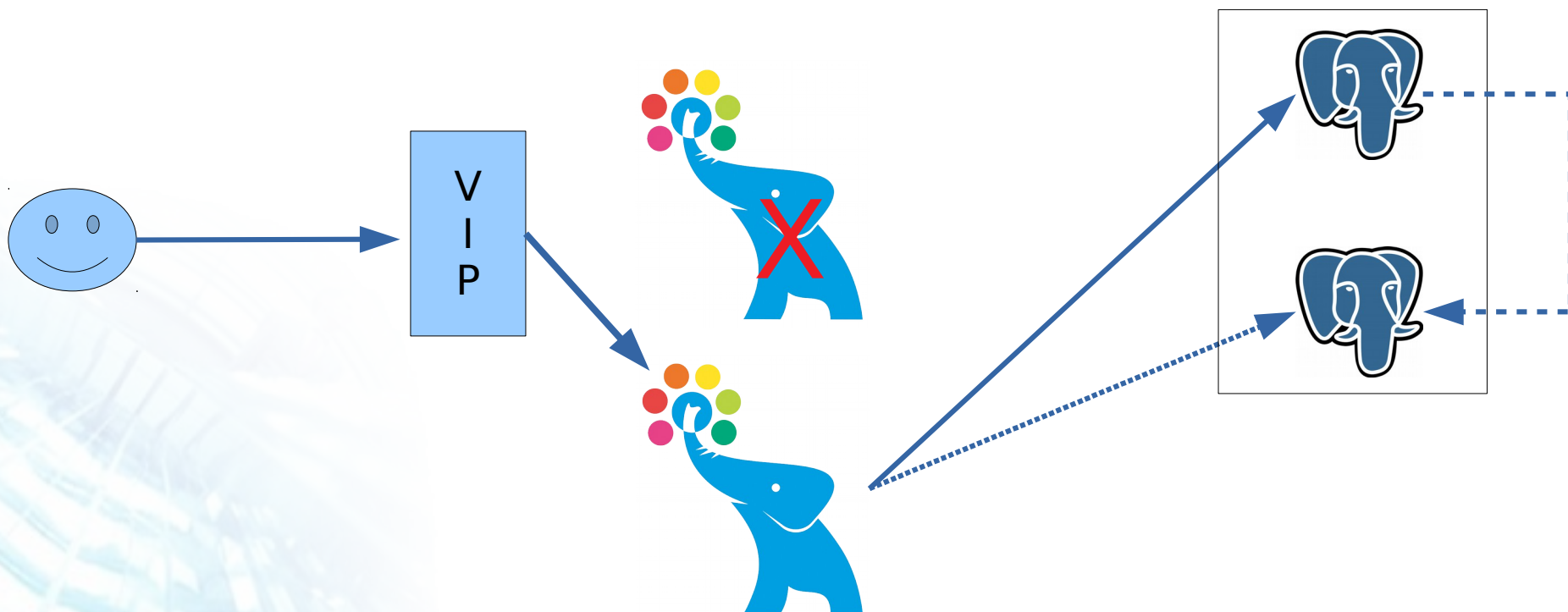
Pgpool-IIの機能: watchdog

- Pgpool-IIがダウンすると、PostgreSQLにアクセスできなくなってしまう(単一障害点、SPOF: Single point of failure)



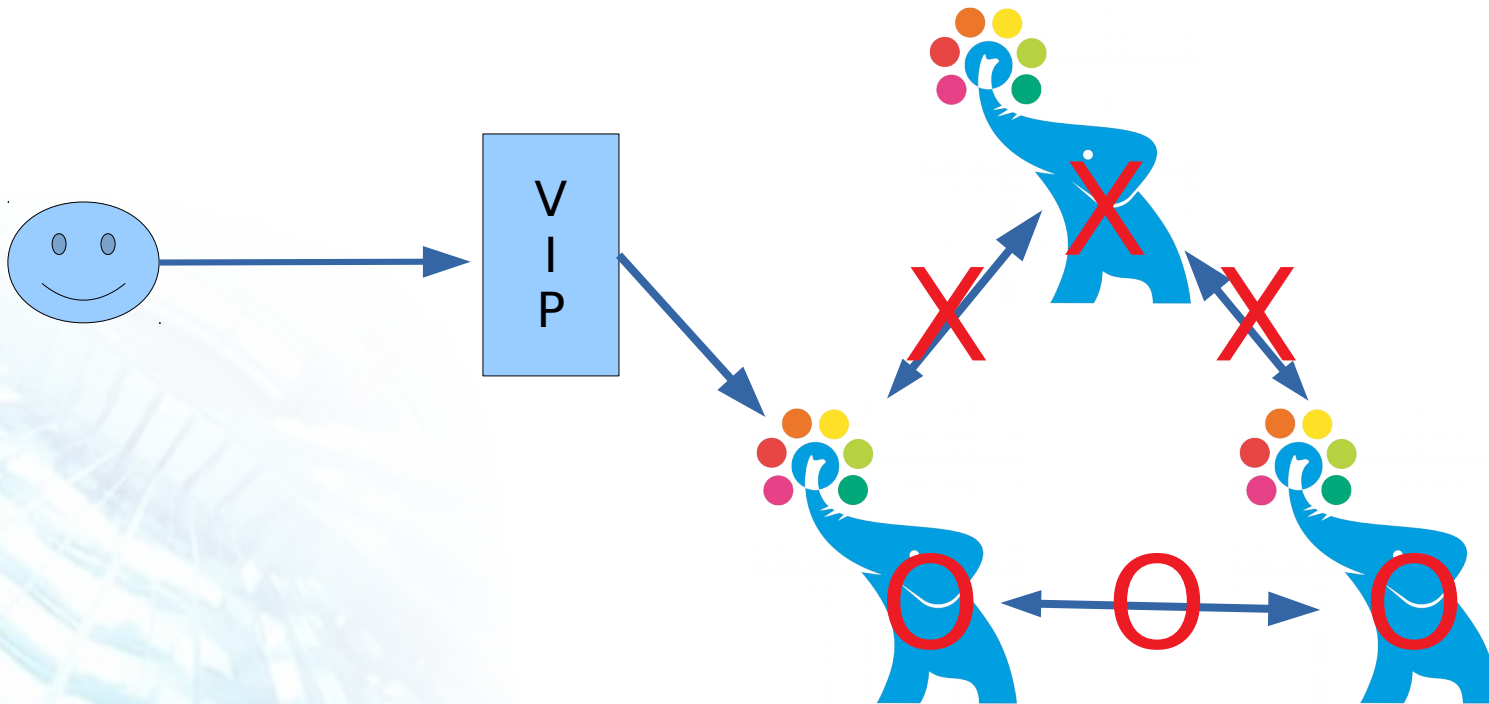
Pgpool-IIの機能: watchdog

- 複数のPgpool-IIを用意してお互いに監視し、相手がダウンしたときには待機系のPgpool-IIを昇格して運用を継続する
- Pgpool-IIへのアクセスはVIP経由にして、クライアント側では同一IPでアクセスし続けられるようにする



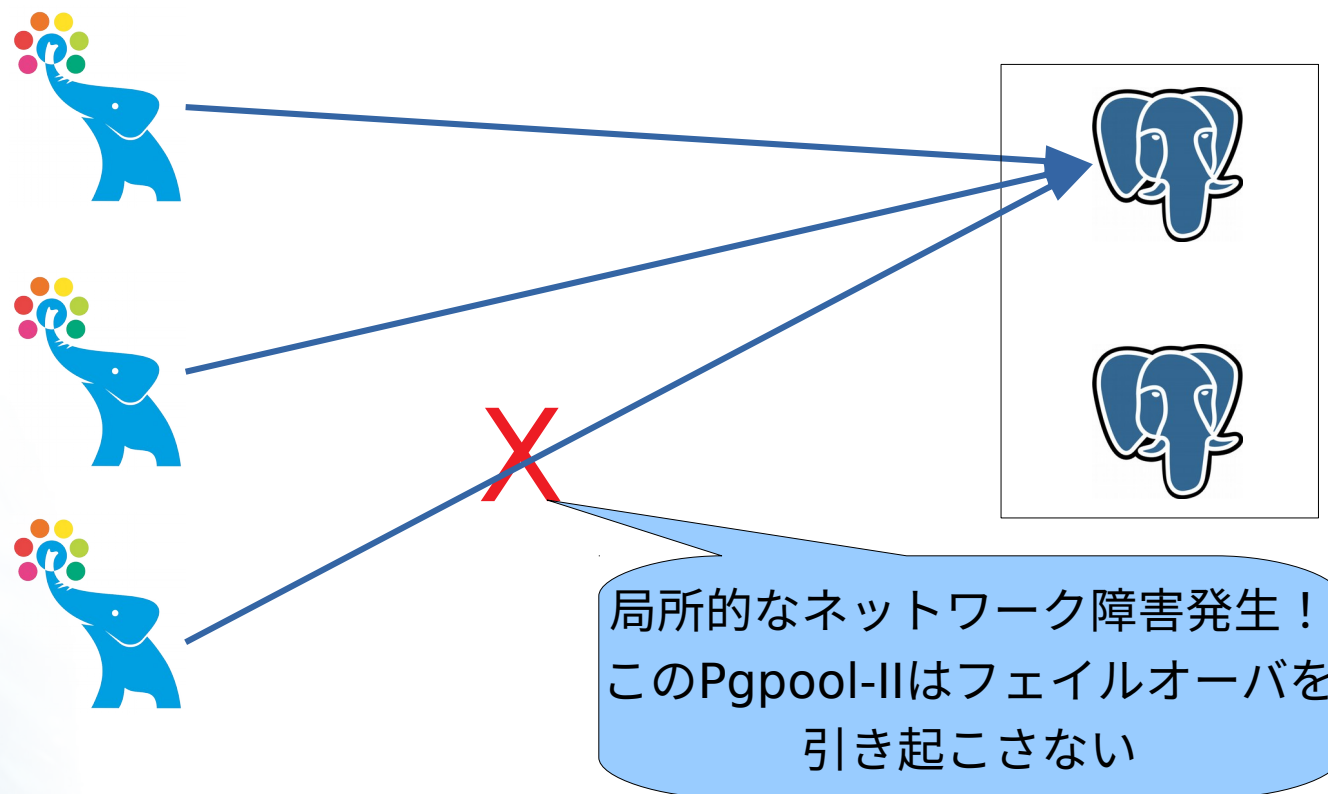
Pgpool-IIの機能: watchdog

- Pgpool-IIがお互いに相手がダウンしていると思ってしまう問題 (Split brain problem)の解決
- Pgpool-IIを奇数台用意することによって、多数決でどれが本当にダウンしているのかを正しく判断できる



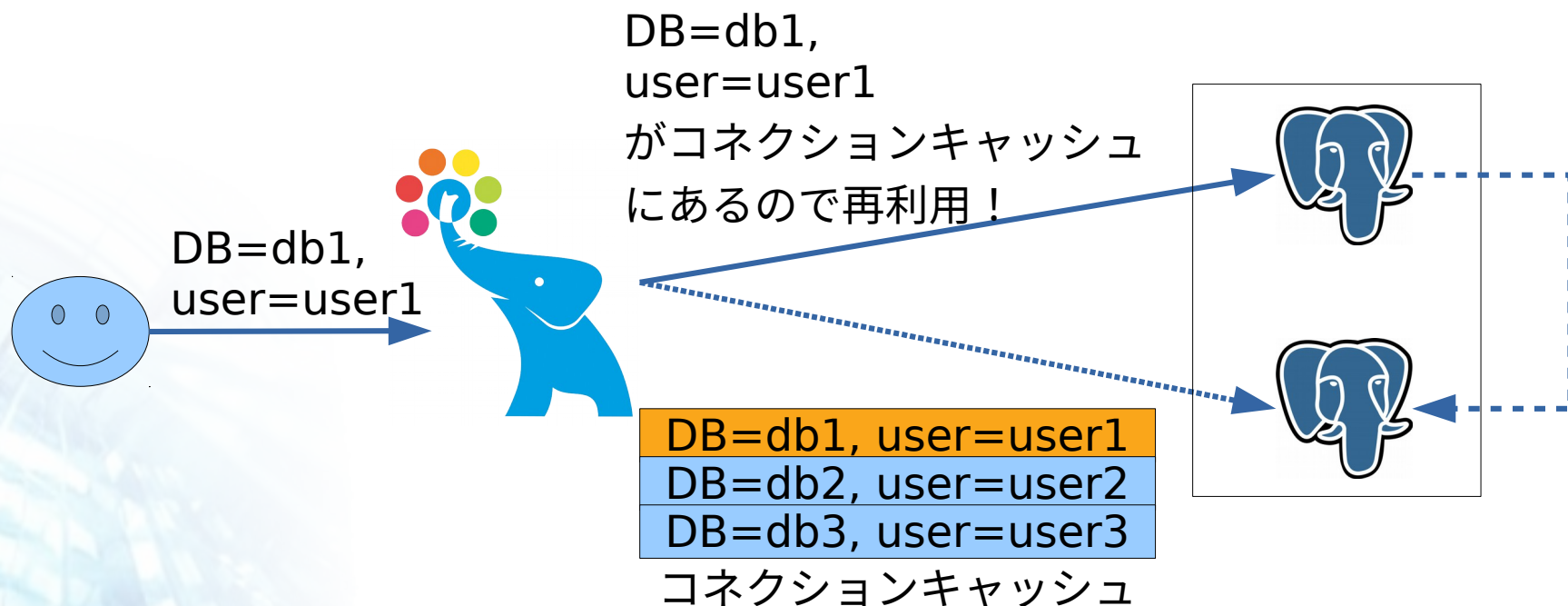
Pgpool-IIの機能: watchdog

- ヘルスチェックでも多数決の原理を利用
- 局所的なネットワーク障害による障害誤検知を防止



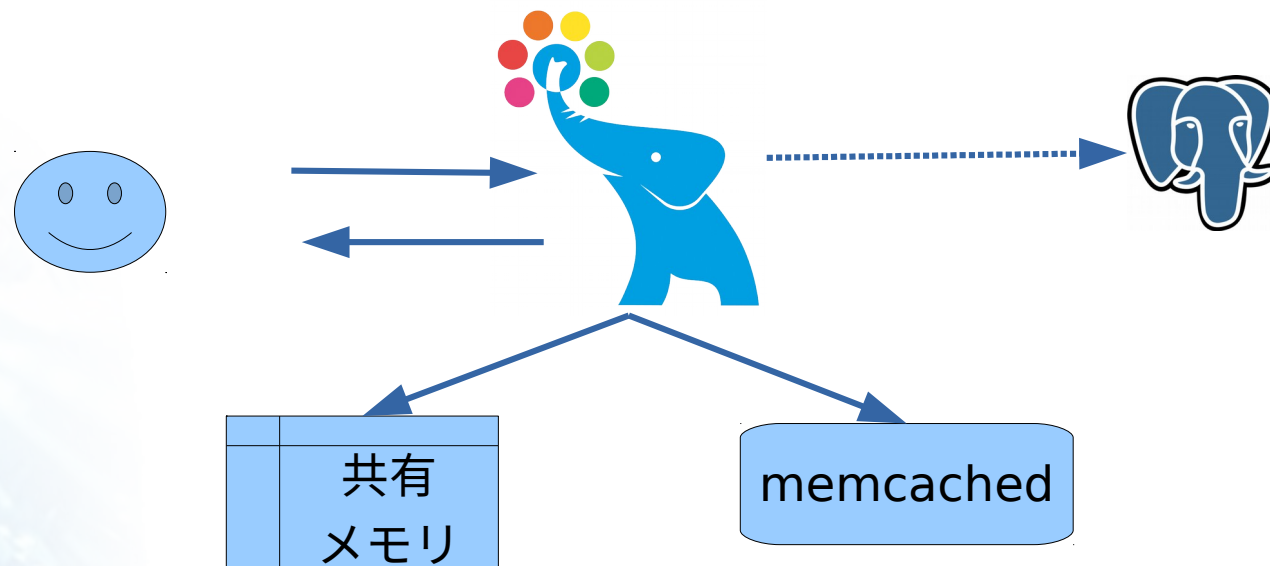
Pgpool-IIの機能： コネクションプーリング

- PostgreSQLに限らず、DBへの接続が時間がかかる
- そこで一度接続したDBへのコネクションを取っておき、条件が合えば再利用する
 - DB名、ユーザ名が同じであること



Pgpool-IIの機能:クエリキャッシュ

- 検索結果を共有メモリ(またはmemcached)に保存し、同じクエリ文字列が送られてきたら2回目以降はキャッシュから結果を返す
 - SQLパーサを経由せず、DBアクセスも発生しないので超高速
- キャッシュはテーブルに変更があったらクリアされる
- 更新の多いシステムには向かない(キャッシュヒット率70%以上での利用を推奨)



クエリキャッシュヒット率の確認

```
show pool_cache;
-[ RECORD 1 ]-----+-----
num_cache_hits      | 891703
num_selects         | 99995
cache_hit_ratio     | 0.90
num_hash_entries    | 131072
used_hash_entries   | 99992
num_cache_entries   | 99992
used_cache_entries_size | 12482600
free_cache_entries_size | 54626264
fragment_cache_entries_size | 0
```

Pgpool-IIの将来



今後の計画

- 統合PostgreSQLクラスタ管理ツールを目指す
 - より使いやすく
 - 設定ファイルを簡素化してインストールを容易に
 - より管理しやすく
 - 統計データの取得をより充実
 - ログを見なくても管理者が取るべき次のアクションを判断できるのが理想
 - クラウド・コンテナ対応
 - Kubernetes対応
 - 次世代高信頼性クラスタ
 - スナップショットアイソレーションモードの改良
 - 高速フェイルオーバ
 - 高度なデータ一貫性

参考URLなど

- Pgpool-II公式サイト
 - <https://pgpool.net>
- Pgpool-II on Kubernetes
 - https://github.com/pgpool/pgpool2_on_k8s
- Pgpool-II Prometheus exporter
 - https://github.com/pgpool/pgpool2_exporter
- SRA OSS技術ブログ
 - <https://www.sraoss.co.jp/tech-blog/category/pgpool/>

まとめ



- Pgpool-IIの生い立ち、現在そして将来について説明しました
- Pgpool-IIは総合PostgreSQLクラスタ管理ツールを目指して改良を続けていきます
- 他にも豊富な機能が盛り沢山なので、是非参考リンクからマニュアルやブログ記事をご覧ください