

PostgreSQLによる クラスタ構成の可能性

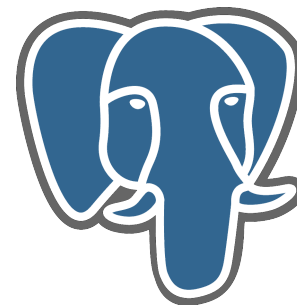
SRA OSS, Inc.日本支社

取締役支社長

石井 達夫

SRA OSS, Inc.のご紹介

- PostgreSQLを中心とした、OSSへの様々なサービスを提供
 - サポートサービス、コンサルティング
 - パッケージ製品
 - PowerGres, libTextConv, Sylpheed Pro
 - 教育サービス
 - トレーニング
 - 技術者認定制度(PostgreSQL CE)
 - [Http://www.sraoss.co.jp](http://www.sraoss.co.jp)
- 私自身は経営の立場とエンジニアの二足のわらじ
- Twitter: @tatsuo_ishii



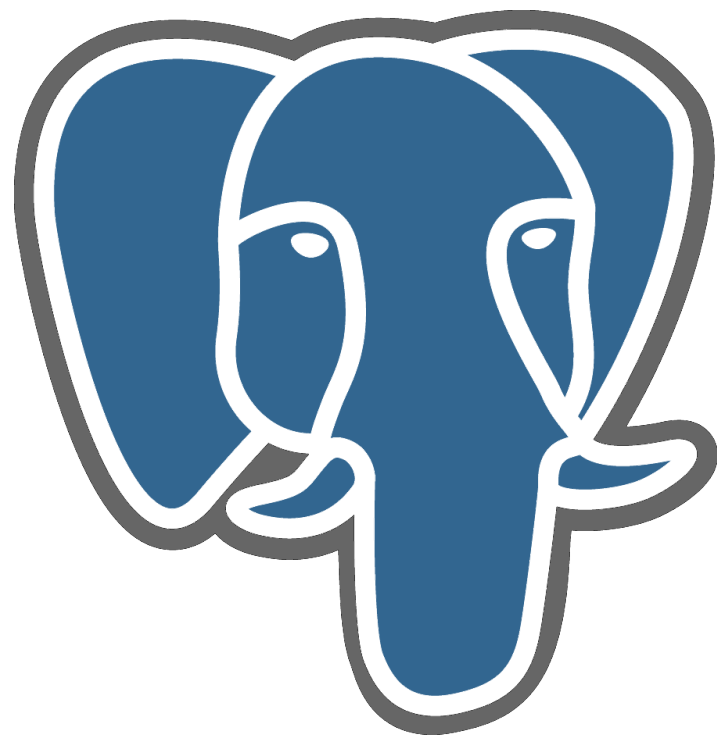
PowerGres



PostgreSQL CE

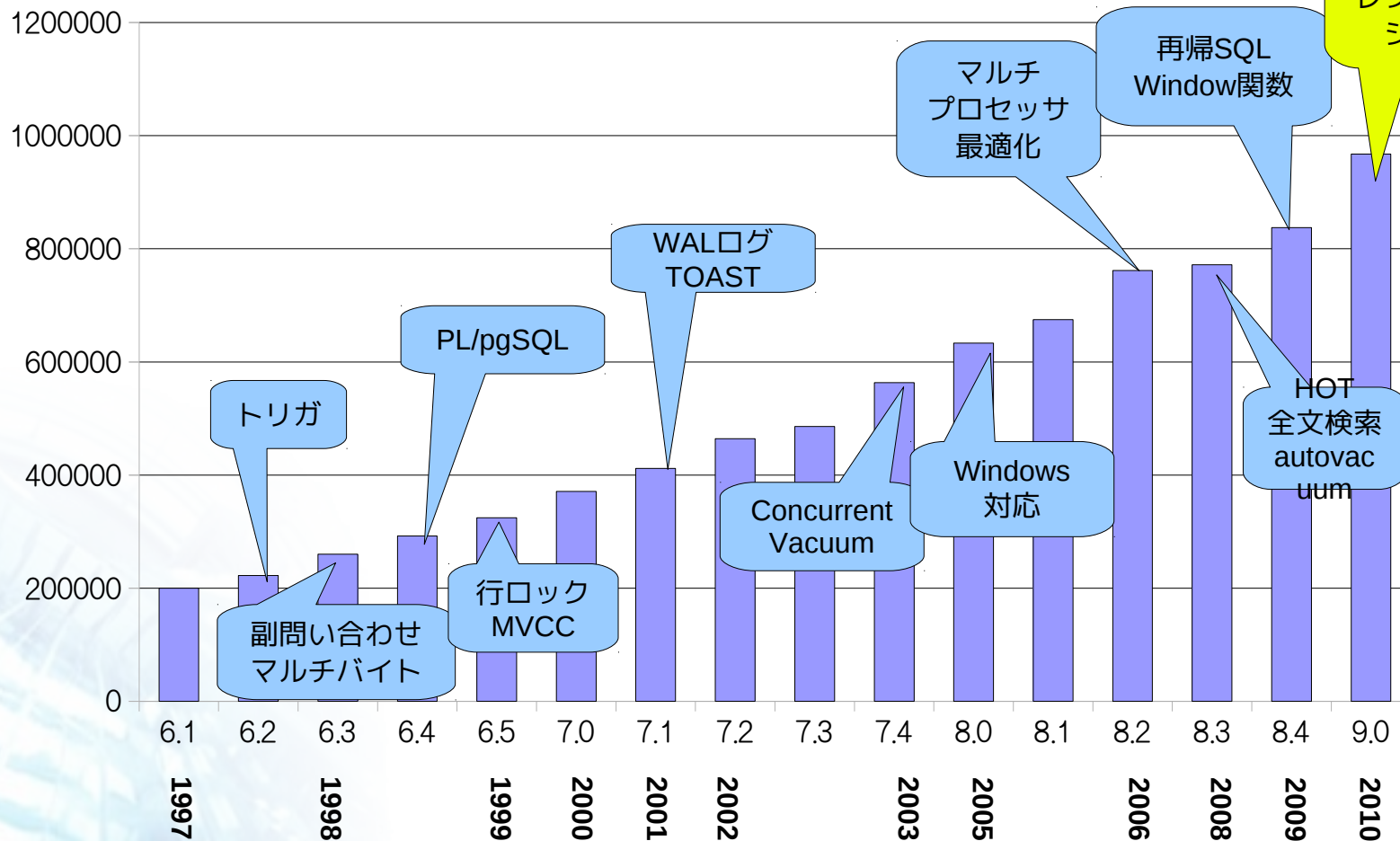
PostgreSQLとは

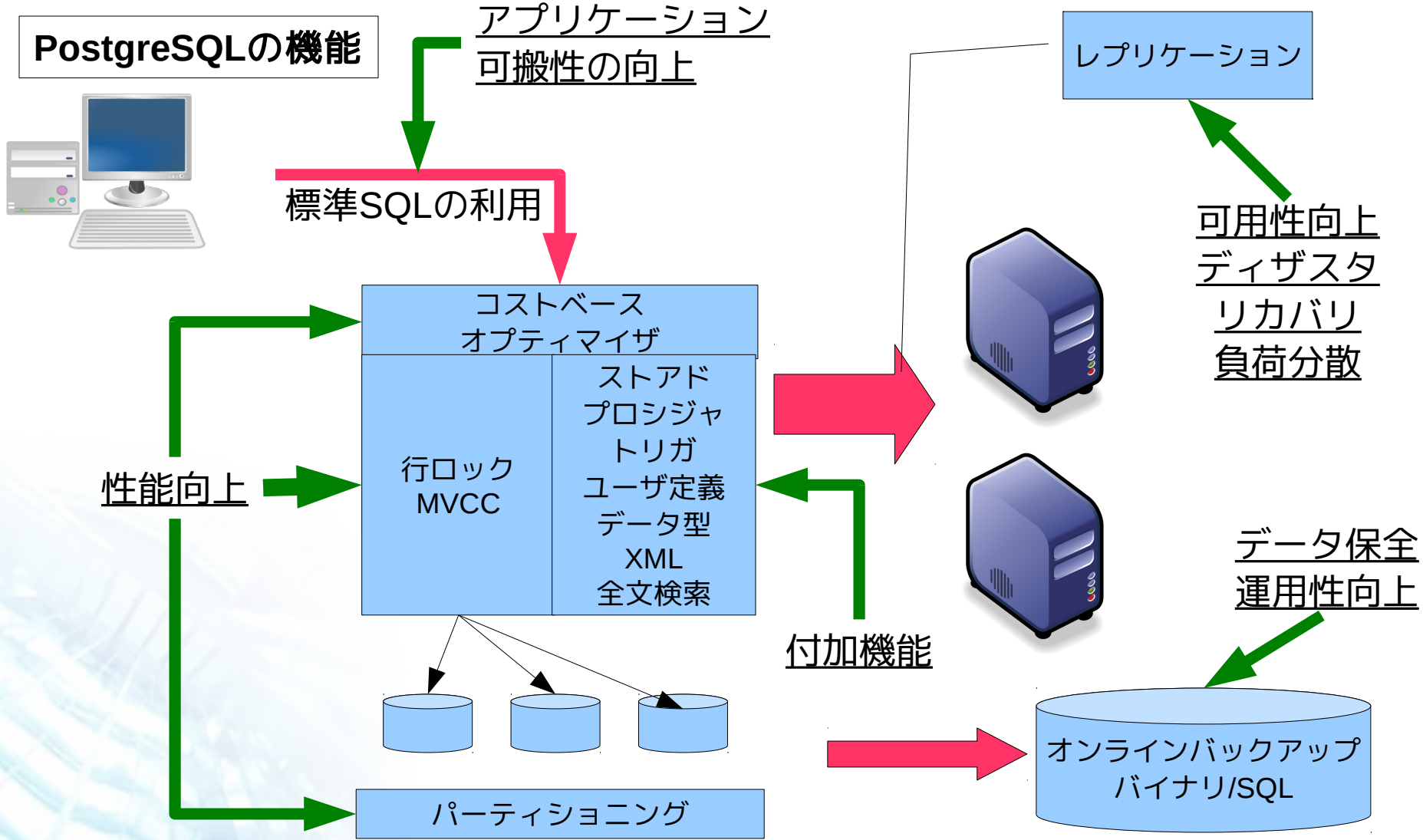
- 10年以上の歴史を持ち機能、信頼性、安定性に定評のあるデータベース
- CPUスケールビリティにも定評(12 CPU程度までスケール)
- TB級の大規模データベースの事例もある



急速に成長を続けるPostgreSQL

ソースコード行数





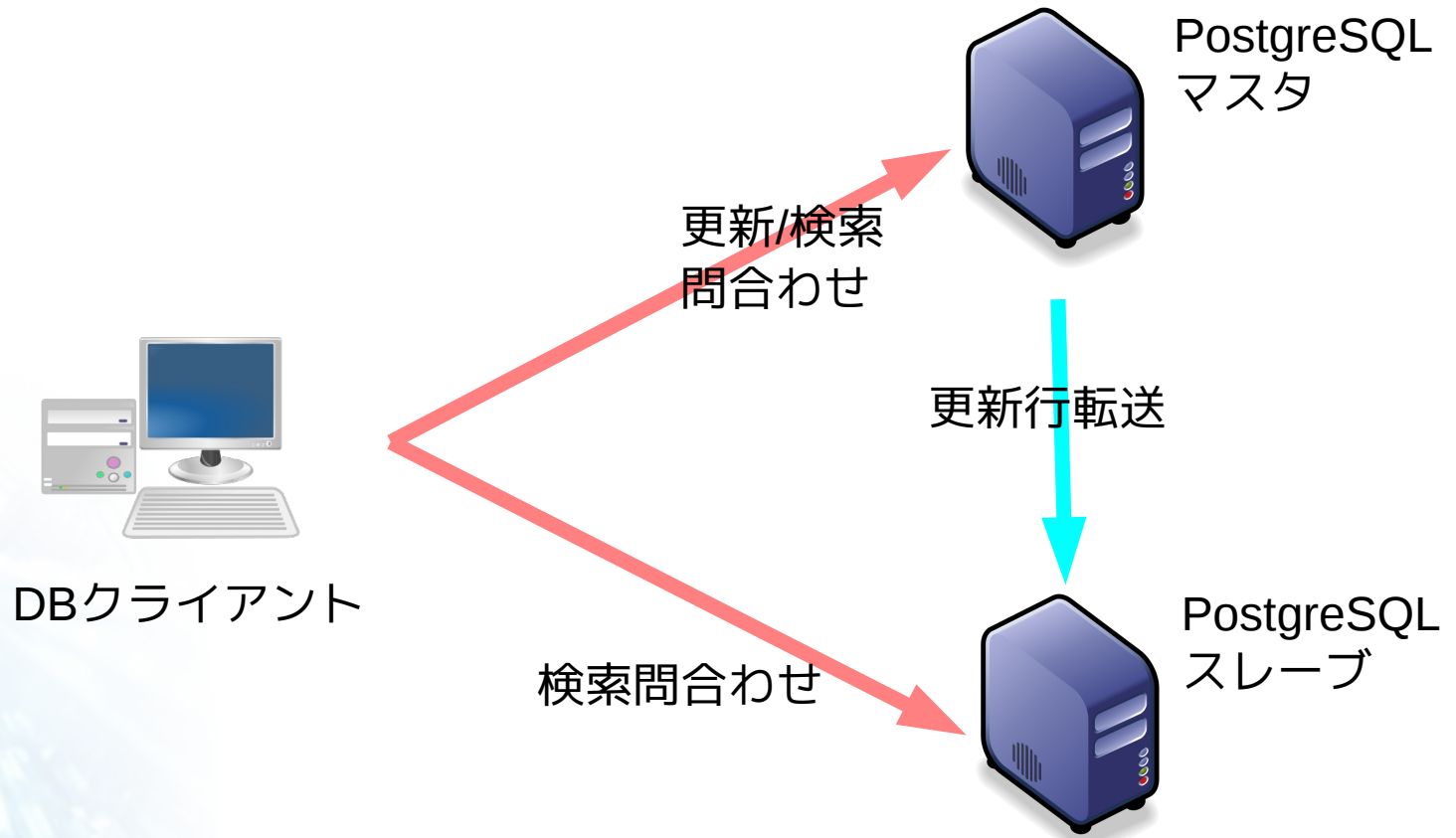
クラスタ構成で注目される OSS DBの利用

- コスト節約効果が顕著
- 製品グレードによってクラスタ機能が使えないなどの心配がない
- 必要な分だけ有償保守を利用可能
- ライセンス管理の煩わしさから解放
- ビジネスの成長に合わせて手軽にクラスタ規模を拡大可能

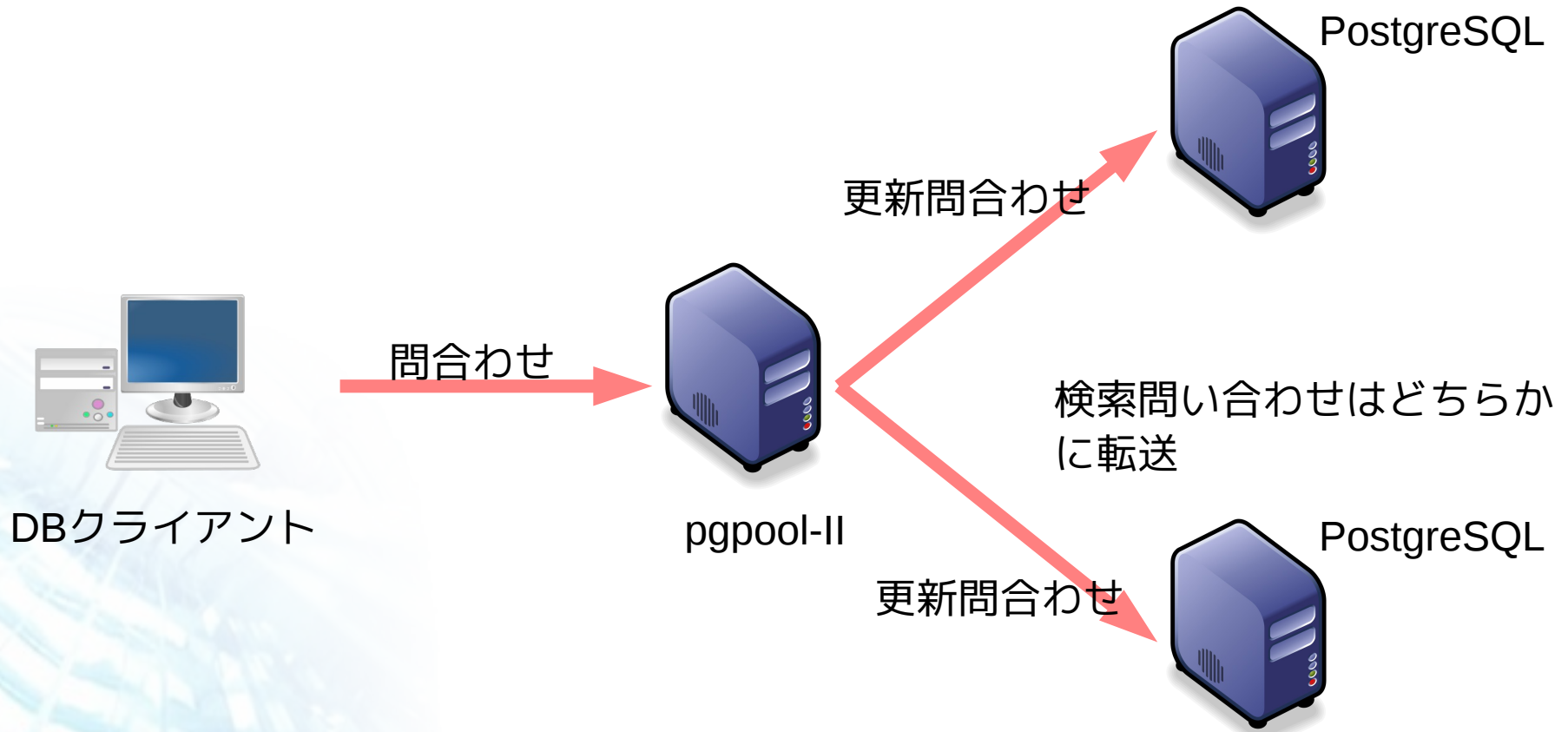
PostgreSQL 9.0以前に利用可能だった 代表的なレプリケーション

- Slony-I
 - トリガを使って更新データを転送してレプリケーション(トリガベースレプリケーション)
 - 非同期レプリケーション/更新性能の劣化が少ない
- Pgpool-II
 - SQL文をコピーして転送してレプリケーション(クエリベースレプリケーション)
 - 同期レプリケーション/更新性能の劣化が多い

Slony-Iによるレプリケーション

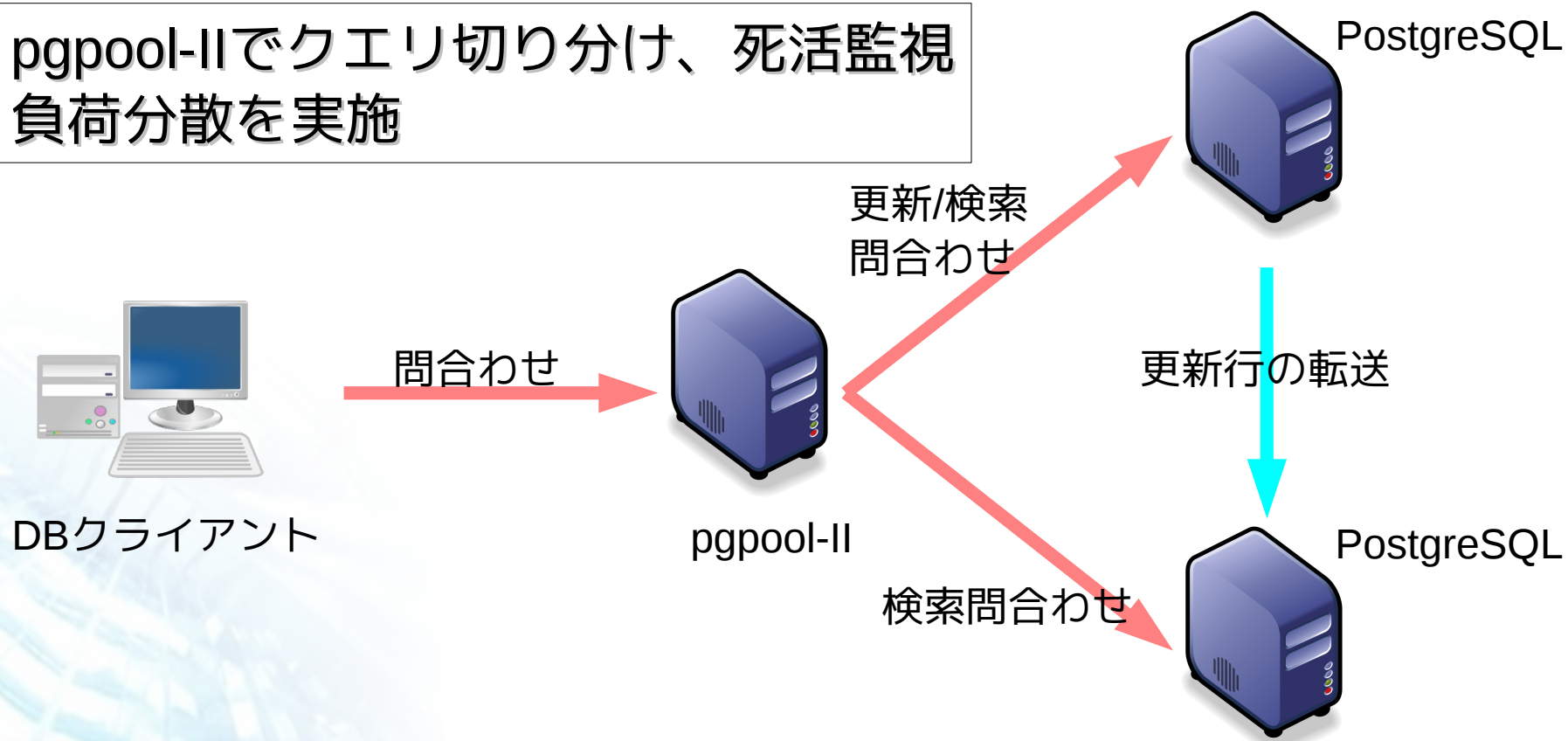


pgpool-IIによるレプリケーション



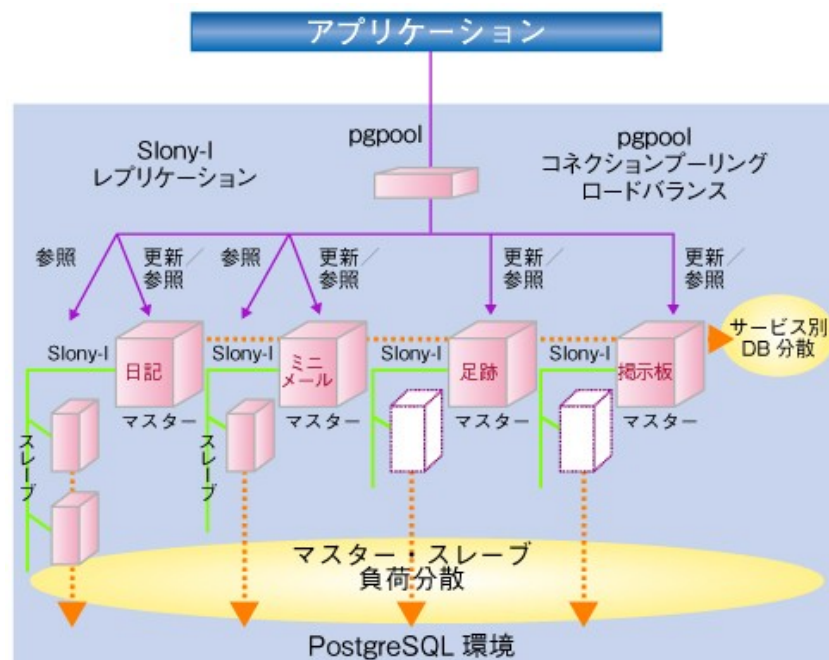
Slony-Iとpgpool-IIの組み合わせ

pgpool-IIでクエリ切り分け、死活監視
負荷分散を実施



大規模クラスタシステム事例

- オープンドア様事例
- 国内最大級規模の携帯電話向けSNS/ゲームサイト
- 月間4億PV, トランザクションの2割が更新系
- pgpool-II + Slony-Iで20台以上のPostgreSQLを管理. スケールアウトする大規模システムを構築



<http://members.techtarget.itmedia.co.jp/tt/members/0802/28/news01.html>

ITMedia techTarget記事より引用

Slony-Iによるレプリケーションの問題点

- 運用が煩雑
- レプリケーションできないデータがある
- レプリケーションの遅延が多い

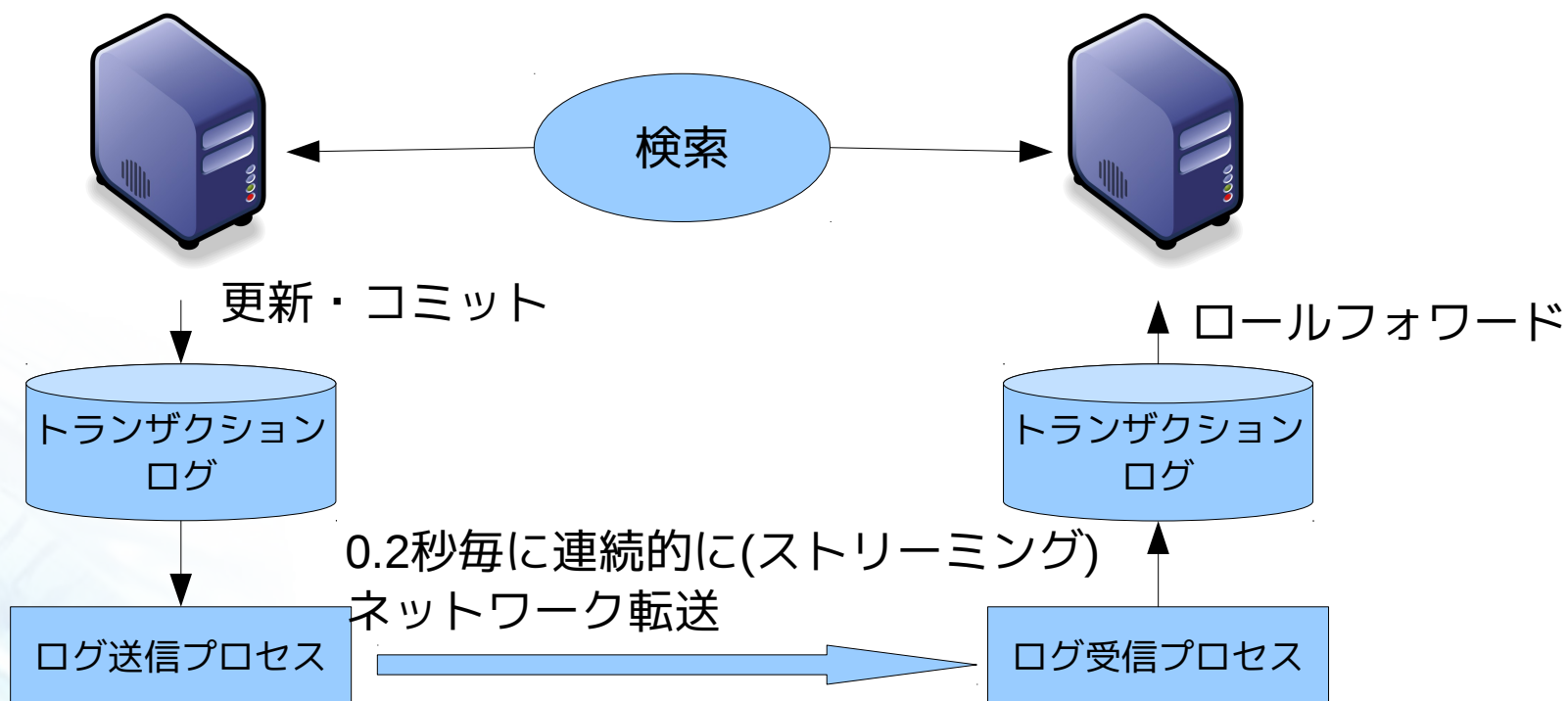
PostgreSQL 9.0のストリーミング レプリケーション

- 運用が容易
- すべてのデータをレプリケーション可能
- レプリケーションの遅延が少ない
- レプリケーションが本体に組み込みなので導入しやすい

ストリーミングレプリケーションの仕組み

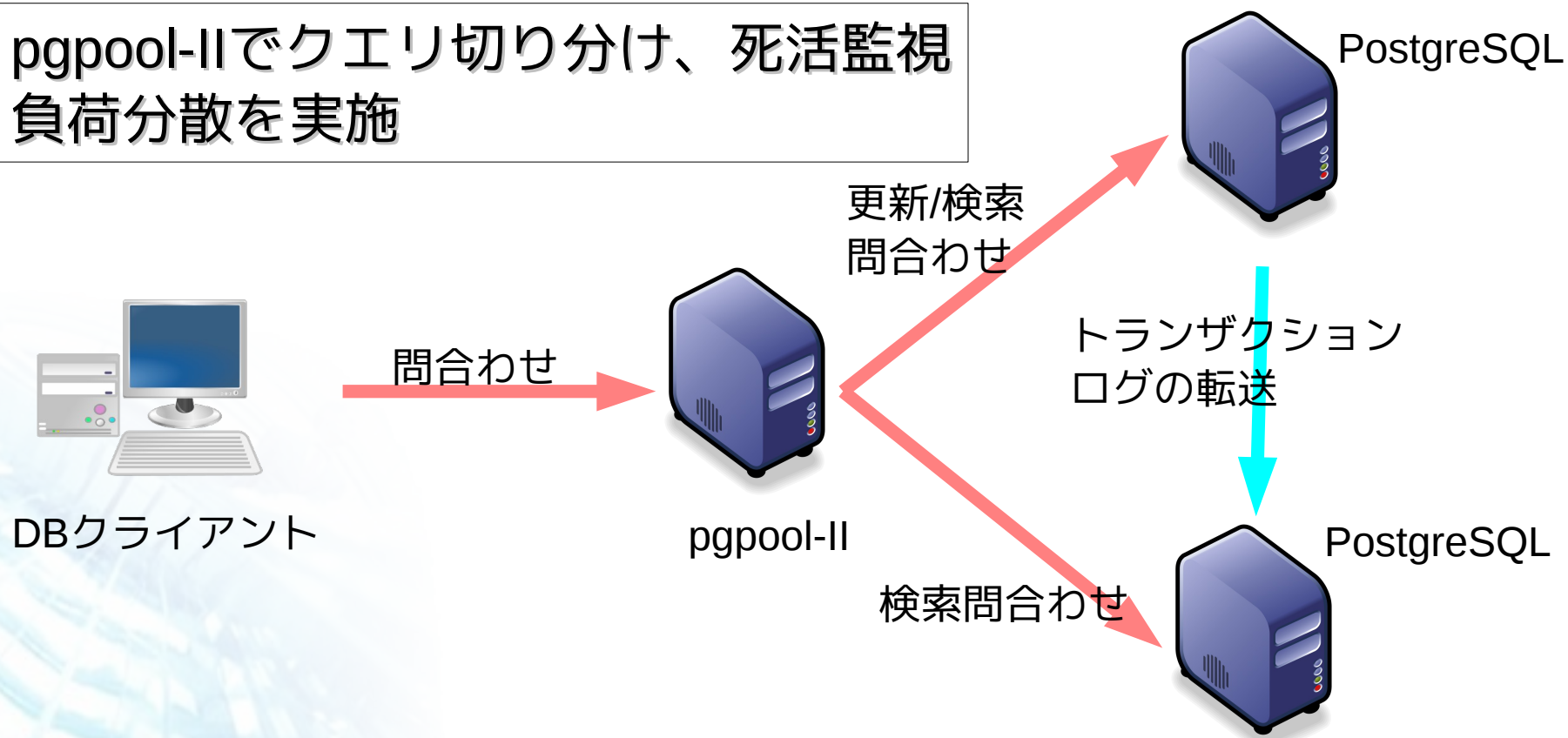
プライマリサーバ

スタンバイサーバ



Streaming Replication+ Hot Standbyとpgpool-IIの組み合わせ

pgpool-IIでクエリ切り分け、死活監視
負荷分散を実施



PostgreSQLによるクラスタリングの 現在の到達点

- 可用性は充分確保できている
- 検索性能はスケールアウト可能
- 更新性能の向上は対価が必要
 - 非同期レプリケーション
 - データ分割戦略
- 妥協の必要がない更新性能の向上は、PostgreSQLに限らずすべてのDBMSの課題

PostgreSQLにおけるレプリケーション クラスタソリューションの今後

- ストリーミングレプリケーションの改良
 - 同期レプリケーションの実装
- PostgreSQL-XCプロジェクト
 - NTTが中心になって実施しているOSSプロジェクト
 - 検索だけでなく、更新もスケールするクラスタシステム

Postgres-XCのアーキテクチャ



SQLの解析

グローバル
トランザクション
マネージャ



コーディネータ

コーディネータ

コーディネータ

SQLの実行

データノード

データノード

データノード

参考URL

- PostgreSQL
 - <http://www.postgresql.org>
- Slony-I
 - <http://www.slony.info/>
- Pgpool-II
 - <http://pgfoundry.org/projects/pgpool/>
- Postgres-XC
 - <http://postgres-xc.sourceforge.net/>

ご清聴ありがとうございました