



**PGEC**cons  
PostgreSQL Enterprise Consortium

# 最新版PostgreSQLの パフォーマンスを引き出すためのポイント

**2015年06月10日**

**PostgreSQL エンタープライズ・コンソーシアム  
技術部会**

<http://www.pgecons.org/>



# 目次

1. はじめに (本日お伝えしたいこと)
2. PGEConsについて
3. 技術部会の活動内容
4. PostgreSQL技術紹介
5. さいごに



# 1. はじめに(本日お伝えしたいこと)

- ① 企業の基幹業務にPostgreSQLの導入が加速しており、ミッションクリティカル性の高い領域への普及促進、課題解決に奮闘している企業連携の団体(PGECons)があること
- ② PGEConsの活動内容と活動成果のご紹介

# PostgreSQLのエンタープライズにおける活用

## 【PGEConsで紹介してきた活用事例（1 / 2）】

### ■ 住友電気工業様

- 2005年より社内標準DBをPostgreSQLに

### ■ ヤマハモーターソリューション様

- 世界各地の拠点の基幹業務システムに採用

### ■ 株式会社キャム様

- クラウド・SaaS型 統合基幹業務システム「CAM MACS」

### ■ 株式会社スポットライト様

- スマートフォンを使った共通来店ポイントサービス「スマポ」

# PostgreSQLのエンタープライズにおける活用

## 【PGEEConsで紹介してきた活用事例（2/2）】

### ■ 日本電信電話株式会社様

- NTTグループ100システム以上でのPostgreSQLの利用、推進を実施

### ■ 株式会社中電シーティーアイ様

- PostgreSQLのサービス化を提供開始

### ■ 株式会社インテリジェンス ビジネスソリューションズ様

- 共通基盤の基幹システムへの導入

### ■ 株式会社Gengo様

- pgpool-IIで可用性を向上させたWebシステムでの利用

# PostgreSQLのエンタープライズにおける活用

## 【その他の導入事例】

### ■ 英気象庁

- 年間サポート費用と保守費用の負担が大きいため、商用DBMSへの依存度を低下させることを狙って、PostgreSQLを活用

- ITメディア様の記事より

<http://techtarget.itmedia.co.jp/tt/news/1407/24/news01.html>

### ■ 仏社会保障システム

- メインフレームからの移行

- Let's postgres の記事より

[http://lets.postgresql.jp/documents/case/Bull\\_Case\\_in\\_France/](http://lets.postgresql.jp/documents/case/Bull_Case_in_France/)



## **2. PGECons (PostgreSQL Enterprise Consortium) について**

## 2. PGEECons について

### ■ PGEEConsの発足と目的

- 2012年4月11日発足
- ミッションクリティカル性の高い**エンタープライズ領域へのPostgreSQLの普及を推進**するため、各種ツールやPostgreSQL本体に関する**利用技術情報**の収集と提供および、その整備などの**活動を企業ベースで展開**する

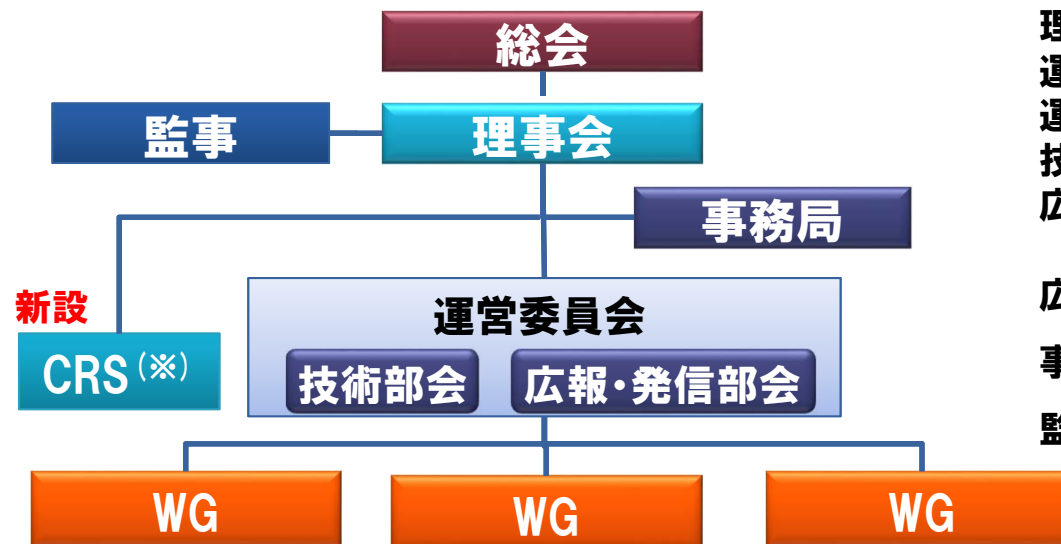
| 活動項目              | 概要                                                            |
|-------------------|---------------------------------------------------------------|
| 情報発信              | 会員の導入実績を基に、PostgreSQLおよび周辺ツールに関する情報を集約し、情報発信サイトや 세미나等を通じて提供する |
| 共同検証              | エンタープライズ領域への適用に向けて必要となる情報を、実証を通じて充実を図る                        |
| 開発コミュニティへのフィードバック | よりミッションクリティカル性の高い領域への適用に向けた技術的な課題を集約し、開発コミュニティに要望を発信する        |
| 開発プロジェクト支援        | 会員間での機能拡張に関する連携開発や、必要な周辺ツールの開発プロジェクト支援を進める                    |



# 平成27年度 体制・会員構成案

## ■ 会員は法人およびそれに準ずる団体で構成

平成27年度体制 (昨年から変更なし)



理事長 : 日本電信電話株式会社  
 運営委員長 : 日本電気株式会社  
 運営副委員長 : 株式会社 日立製作所  
 技術部会長 : 富士通株式会社  
 広報・発信部会長 : 日本ヒューレット・パッカード株式会社  
 広報・発信副部会長 : 株式会社アシスト  
 事務局長 : SRA OSS, Inc. 日本支社  
 監事 : 税理士法人ジャストスタッフ

(※) CRS: Community Relationship Sub-committee

| 種別   |      | 概要                              | 総会議決権 |
|------|------|---------------------------------|-------|
| 正会員  | 理事   | 理事会に参加、理事長および運営委員長は理事のうちから就任する  | あり    |
|      | 運営委員 | 運営委員会に参加、部会長およびWG長は運営委員から就任する   |       |
|      |      | ワーキンググループ(WG)やCRS等に参加し、活動に貢献    |       |
| 一般会員 |      | メーリングリストやWebなどから、活動情報を取得することが可能 | なし    |

# 平成27年度 参加会員一覧(1/2)

(平成27年6月8日 現在)

## ■ 49社(正会員 17社、一般会員32社)が活動

| 参加法人名 (順不同)  |                |                     |                       |
|--------------|----------------|---------------------|-----------------------|
| 正会員<br>(17社) | 運営委員           | (株)アシスト             | 日本電信電話(株)             |
|              |                | SRA OSS, Inc. 日本支社  | 日本ヒューレット・パッカー(株)      |
|              |                | NECソリューションイノベータ(株)  | (株)日立製作所              |
|              |                | 特定非営利活動法人 LPI-Japan | (株)日立ソリューションズ         |
|              |                | TIS(株)              | 富士通(株)                |
|              |                | 日本電気(株)             | (株)富士通ソーシャルサイエンスラボラトリ |
|              | NTTソフトウェア(株)   |                     | 大日本印刷(株)              |
|              | (株)HTKエンジニアリング |                     | 日本電子計算(株)             |
|              | サイオステクノロジー(株)  |                     |                       |
|              |                |                     |                       |

# 平成27年度参加会員一覧(2/2)

(平成27年6月8日 現在)

|               | 参加法人名（順不同）          |                     |                   |
|---------------|---------------------|---------------------|-------------------|
| 一般会員<br>(32社) | (株)アイ・アイ・エム         | キーウェアソリューションズ(株)    | 東芝ソリューションズ(株)     |
|               | (株)アイ・ティ・プロデュース     | (株)ギークフィード          | (株)ニイズ            |
|               | 岩通ソフトシステム(株)        | (株)キャム              | 日本アイ・ビー・エム(株)     |
|               | (株)インフォメーションクリエーティブ | クオリカ(株)             | (株)フィックスターズ       |
|               | (株)エクサ              | ジャパンシステム(株)         | フューチャーアーキテクト(株)   |
|               | SFKメディカル(株)         | 住友電気工業(株)           | (株)マインド           |
|               | (株)エニブラ             | 住友電工情報システム(株)       | ミュートック(株)         |
|               | (株)エム・オー・エム・テクノロジー  | ZEKKO INC.          | (株)メトロシステムズ       |
|               | エンタープライズDB(株)       | (株)seiwa            | ヤマハモーターソリューション(株) |
|               | (株)オンザマーク           | (株)中電シーティーアイ        | ローリーコンサルティング(株)   |
|               | 関電システムソリューションズ(株)   | (株)デジタル・ヒュージ・テクノロジー |                   |

# WGの活動風景

## ■ 検討会



## ■ 実機検証



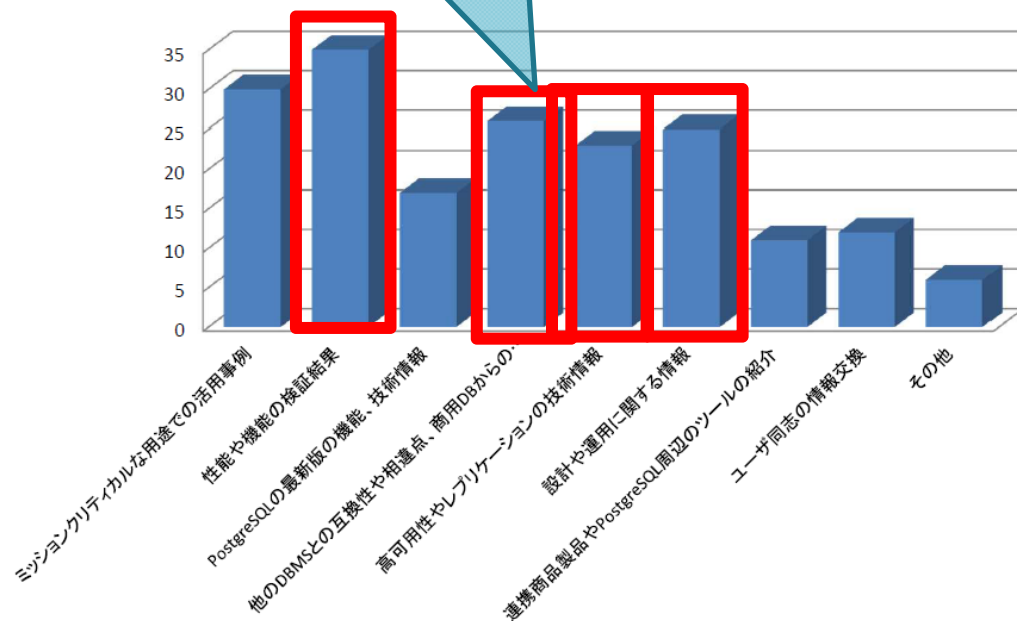


### 3. 技術部会の活動内容

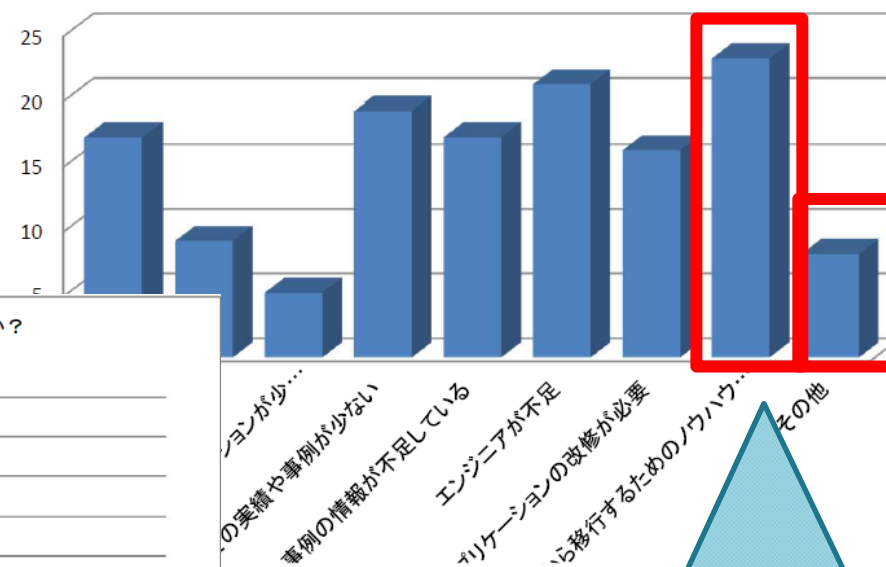
# 過去セミナーのアンケートよりテーマを選定

PGEConsへの**期待**  
「性能や移行、高可用性、  
設計・運用に関する  
情報発信」

Q6. PGEConsに何を期待していますか？またどんな情報を発信してほしいですか？  
(複数回答可)



Q5. PostgreSQLをよりミッションクリティカル性の高いエンタープライズ領域で採用するための課題は何だとお考えですか？(複数回答可)



PostgreSQL採用の**課題**  
「個々のコメントから具体的  
テーマを選定」



# 2014年度 PGECcons WG活動報告 (技術部会)

## 【2014年度 各WGの活動テーマ (PGECconsにおける課題領域より)】

※赤字:2014年度の活動テーマ

|                |        |                                                                   |                            |
|----------------|--------|-------------------------------------------------------------------|----------------------------|
| WG1<br>(2012～) | 性能     | 性能評価手法、 <u>スケールアップ</u> 、スケールアウト、性能向上手法、チューニング                     | 定点観測 + 仮想環境での性能検証          |
|                | 可用性    | 高可用クラスタ、 <u>BCP</u>                                               | 災害対策における可用性の運用検証           |
| WG3<br>(2013～) | 保守性    | 保守サポート、トレーサビリティ                                                   |                            |
|                | 運用性    | 監視運用、バックアップ運用、サイジング設計                                             |                            |
|                | セキュリティ | <u>監査</u>                                                         | PostgreSQL版セキュリティガイドラインの作成 |
| WG2<br>(2012～) | 互換性    | <u>データ</u> 、 <u>スキーマ</u> 、 <u>SQL</u> 、ストアドプロシージャ、チューニング、バージョンアップ | 移行時の試験項目を検討                |
|                | 接続性    | 他ソフトウェアとの連携                                                       |                            |

# 技術部会の活動内容と方針

## 【2015年度 各WGの活動テーマ(予定)】

※赤字:2015年度の活動テーマ

|     |        |                                                                                             |                                           |
|-----|--------|---------------------------------------------------------------------------------------------|-------------------------------------------|
| WG1 | 性能     | <u>性能評価手法</u> 、 <u>スケールアップ</u> 、 <u>スケールアウト</u> 、 <u>性能向上手法</u> 、 <u>チューニング</u>             | 定点観測 + OLAP性能検証やカーネルによる性能比較               |
|     | 可用性    | 高可用クラスタ、BCP                                                                                 |                                           |
| WG3 | 保守性    | 保守サポート、トレーサビリティ                                                                             |                                           |
|     | 運用性    | <u>監視運用</u> 、 <u>バックアップ運用</u> 、 <u>サイジング設計</u>                                              | 性能チューニングや運用面の課題解決に向けた検討                   |
|     | セキュリティ | <u>監査</u>                                                                                   | 2014年度実施のセキュリティ調査の実機検証                    |
| WG2 | 互換性    | <u>データ</u> 、 <u>スキーマ</u> 、 <u>SQL</u> 、 <u>ストアドプロシージャ</u> 、 <u>チューニング</u> 、 <u>バージョンアップ</u> | 高難易度のオブジェクト移行やセキュリティ、NoSQLなど新たな移行検証テーマの追加 |
|     | 接続性    | 他ソフトウェアとの連携                                                                                 |                                           |



# WG1 (性能WG)の2014年度テーマ

## 2013年度の参照系に加え、更新系のスケールアップ検証と仮想環境下 (KVM、Docker) での性能検証を実施

### ■ スケールアップ検証

- 参照系処理時の多コアCPU上でのPostgreSQL 9.4の到達点把握  
PostgreSQL 9.3との比較、2013年度実施page checksum性能影響の再検証
- 更新系処理時のWAL性能改善検証としてPostgreSQL 9.3とPostgreSQL 9.4との比較および多コアCPU下でのスケーラビリティの到達点把握

### ■ 仮想化環境下での性能検証

- KVM上の仮想環境と物理環境の参照系・更新系の性能検証  
複数仮想マシンでの性能影響有無の把握
- Dockerと非Docker、仮想ネットワークの有無、コンテナ内のファイルシステム利用時の参照系・更新系の性能検証

# WG1 2014年度活動内容

## 活動体制（企業名アイウエオ順、敬称略）

- SRA OSS, Inc.日本支社
- 日本電気株式会社
- 日本電信電話株式会社
- 日本ヒューレット・パッカート株式会社
- 株式会社日立製作所

対面による、年間11回の会合を実施

## 検証環境（企業名アイウエオ順、敬称略）

参加企業の検証環境・機器を活用

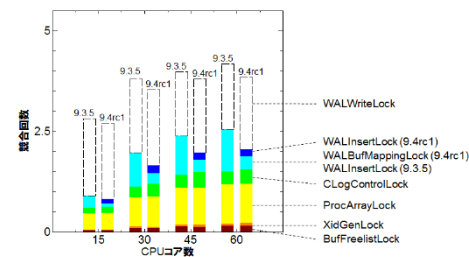
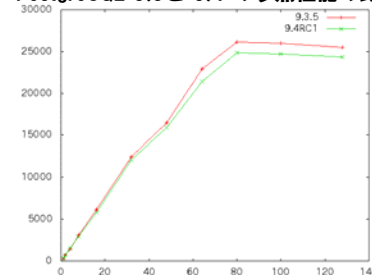
- 日本ヒューレット・パッカート株式会社提供  
HPソリューションセンター

本年度テーマで、合計28日の検証作業を実施

## 活動成果（63ページの検証レポートを作成）

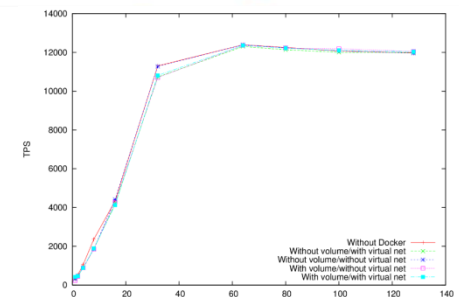
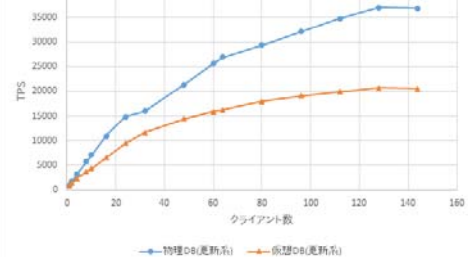
- スケールアップ検証・参照系（定点観測）  
CPUコア数60まで9.4も9.3と同様に良好なスケール性を確認
- スケールアップ検証・更新系（WAL性能改善）  
9.4の高負荷時の更新性能が9.3比で改善する点を確認
- 仮想環境（KVM）検証  
物理/仮想環境（10コア 96GB）を比較し各性能特性を調査
- Docker検証  
非コンテナとコンテナ環境（NW・ファイルシステム）の特性を把握

PostgreSQL 9.3と9.4の参照性能の比較



更新検証時のWALInsertLockを想定ロック競合回数

KVM利用時の物理DBと仮想DBのTPS（更新系）



Docker利用パターンと非Dockerの参照系比較

# 2014年度WG2活動概要

システム移行後に実施する試験項目および手順の検討を行います。

前提となるシステム移行は前年度までに作成した移行手法を参考に実施します。

調査を行った試験大項目は以下の通りです。

## ■ 移行結果確認試験

- ☐ スキーマ移行結果確認試験
- ☐ アプリケーション移行結果確認試験
- ☐ データ移行結果確認試験

活動体制（企業名アイウエオ順、敬称略）

- NECソリューションイノベータ株式会社
- 日本電信電話株式会社
- 富士通株式会社
- 株式会社富士通  
ソーシャルサイエンスラボトリ

# 「異種DBMSからPostgreSQLへの移行ガイド」の構成

- 移行作業の全体像を解説
  - DB移行フレームワーク編 (21ページ)
- 移行作業に含まれる作業内容、手順の調査
  - システム構成調査編 (29ページ)
  - 異種DB間連携調査編 (18ページ)
  - スキーマ移行調査編 (25ページ+別表)
  - データ移行・文字コード変換編 (49ページ)
  - ストアドプロシージャ移行調査編 (34ページ)
  - アプリケーション移行調査編 (10ページ)
  - SQL移行調査編 (18ページ+別表)
  - 組み込み関数移行調査編 (15ページ+別表)
  - チューニング編 (30ページ+別表)
  - バージョンアップ編 (39ページ+別表7)
  - **試験編<NEW!>**
- 移行作業を試行する検証
  - データ移行調査および実践編 (60ページ+別表3)
  - アプリケーション移行実践編 (25ページ+別表)



移行プロセス全体像

# WG3(設計・運用WG) 2014年度活動テーマ

## ■ 可用性（昨年度からの継続テーマ）

- 昨年度：PostgreSQLで可用性を担保するシステム構成(HA・レプリケーション)の特徴を整理し、実機検証まで実施
- 今年度：災害対策を想定し、検討範囲をDRサイト構成に発展

## ■ セキュリティ（今年度からの新設テーマ）

- 利用できる機能の整理と実用レベルのバランス調査
- 他DBMS向けに記述されたDBSCによる文書を参考に、PCIDSS準拠のためのPostgreSQLにおける対処策をまとめる
- ※DBSC: DataBase Security Consortium

全体的なテーマとしては：

**PostgreSQL運用に必要な技術ノウハウ提供を目指す**

# WG3 2014年度活動内容

## 【可用性】

### 活動成果（63ページの報告書を作成）

- DRサイトをPostgreSQLで構築する場合の構成と特徴を机上調査で整理
  - 遠隔地バックアップ
  - ストレージレベルでのデータ同期
  - PostgreSQLストリーミング・レプリケーション（SR）
- SRを使ったDRサイトの実機検証
  - DRサイトの初期構築、障害時の運用手順を整備
  - Amazon Web Service環境で遠隔拠点間のレプリケーションを検証

### 活動体制（企業名アイウエオ順、敬称略）

- TIS株式会社
- 日本電信電話株式会社
- 株式会社日立ソリューションズ

## 【セキュリティ】

### 活動成果（83ページの報告書を作成）

- PostgreSQL版『データベースセキュリティガイドライン』を作成
  - DBSCによる『データベースセキュリティガイドライン 第2.0版』の「製品別機能対応表」をチェック
  - PCI DSSv2.0での要求項目に対し、PostgreSQLでの実現手法を整理
  - PostgreSQLの標準機能で実現できない箇所は、外部ソフトの利用やスクリプト記述などを使った対策方法を例示

### 活動体制（企業名アイウエオ順、敬称略）

- 株式会社アシスト
- SRA OSS, Inc. 日本支社
- NTTソフトウェア株式会社
- サイオステクノロジー株式会社
- 大日本印刷株式会社
- 日本電気株式会社



## **4. PostgreSQL技術紹介 (WG1 (性能WG) 成果物のご紹介)**





## おことわり

- 諸事情により商用DBMSとの比較については、PGEConsからの公開を行うことはありません。
- セミナー等で必ずと言っていいほどご要望をいただいておりますが、実施する予定はございません。ご了承ください。



活動報告1

# 定点観測/ スケールアップ検証 [参照系]

# 定点観測(スケールアップ検証)概要

## ■2012年度から参照系ベンチマークを継続的に実施

## ■計測内容は以下2項目

### □ 9.3と9.4の参照性能の比較

9.3と9.4で同傾向・同程度の参照性能が出るか？

### □ 9.4のpage checksum機能有無での参照性能の比較

page checksum機能のオーバーヘッドは？(2013年度の検証の再試)

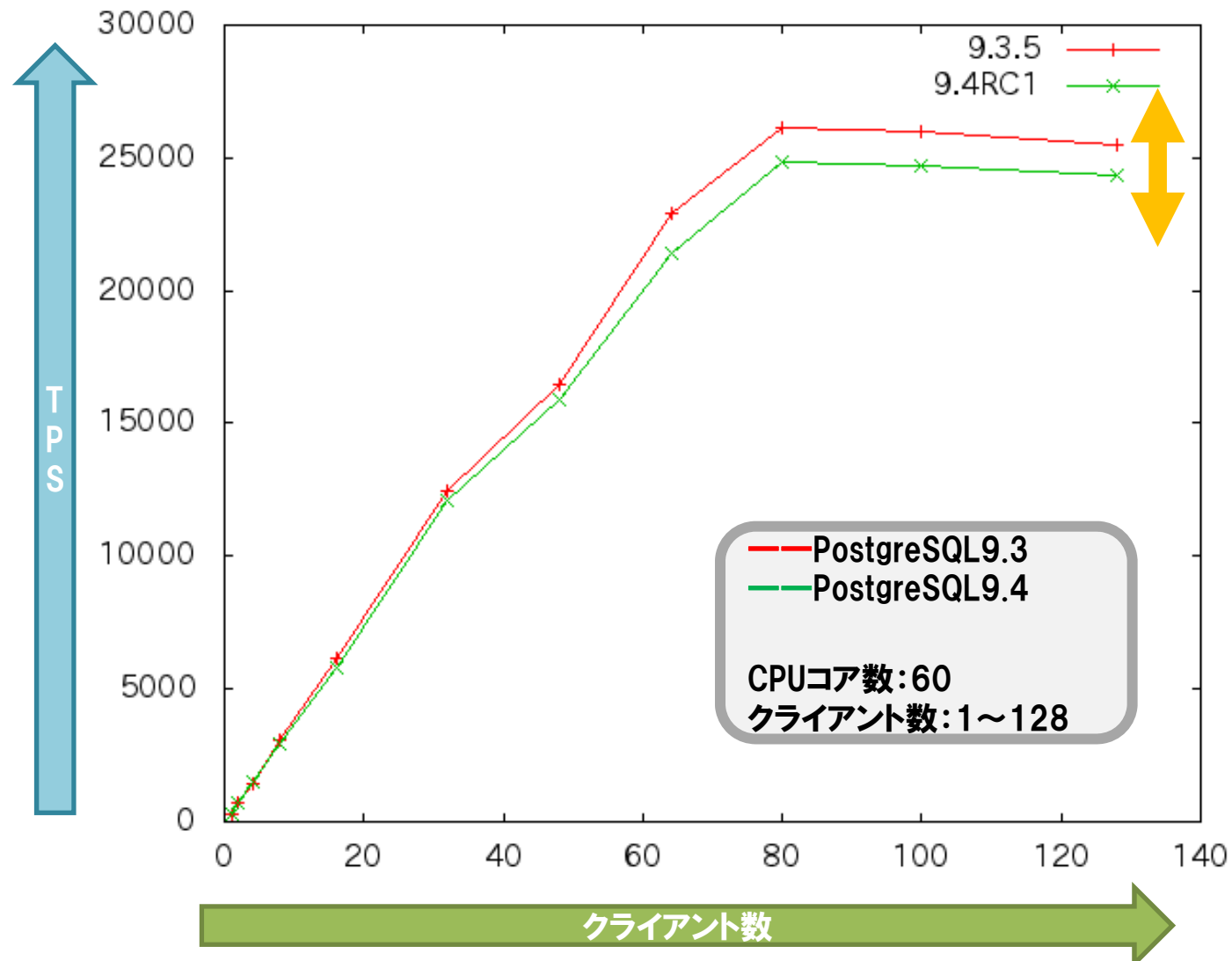
## ■pgbenchで計測

### □ スケールファクタ1,000で初期化(DBサイズは15GB)

### □ ランダムに10,000行取得するカスタムスクリプトを300秒ずつ実行(pgbench -S では十分な負荷がかからないため)

### □ 3回の計測結果の中央値を結果とした

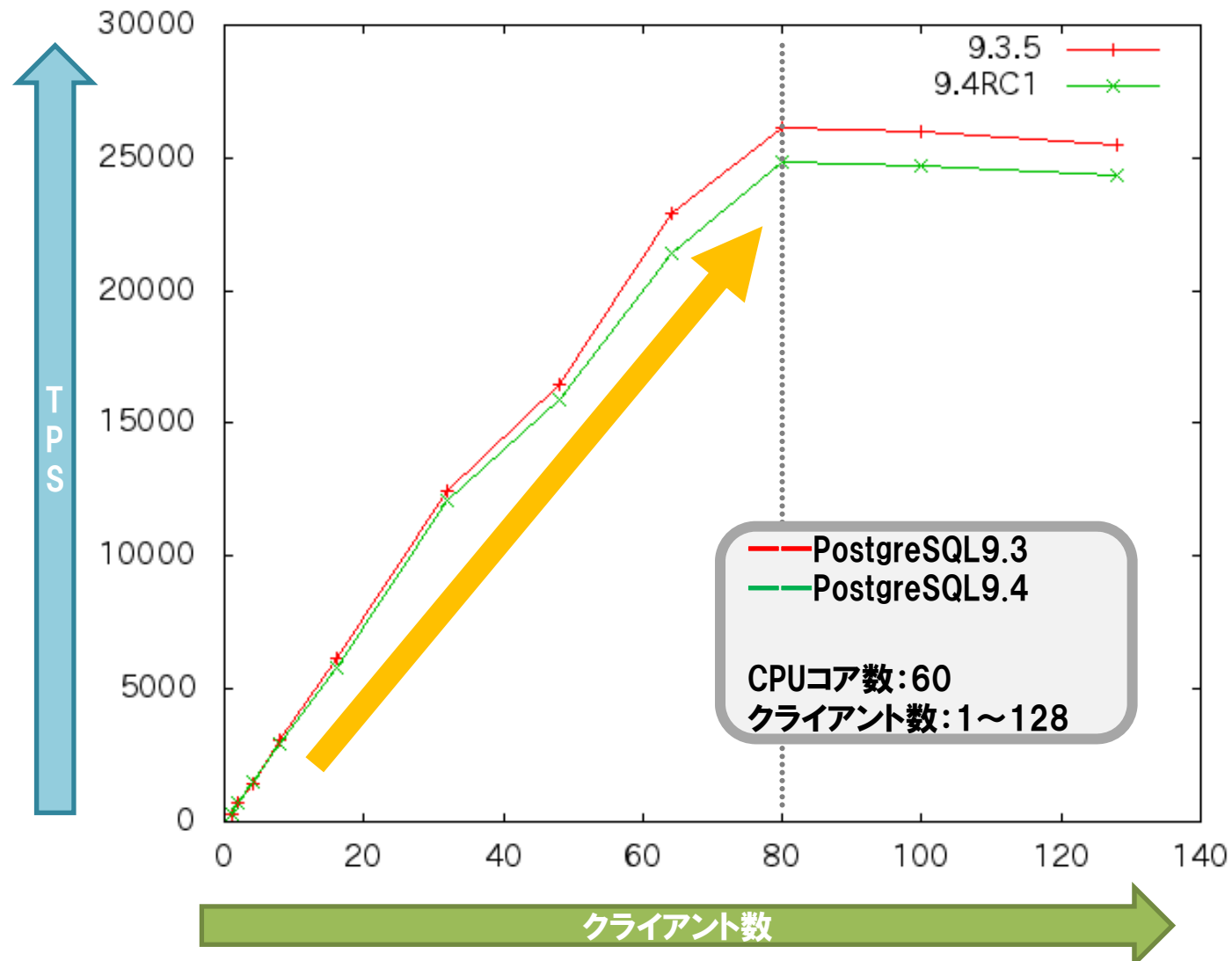
## 9.3 と 9.4 の参照性能の比較 <TPS値>



9.3 より 9.4 の方がわずかにTPSが低く、2～7%程度低下

9.4で行われた何らかの変更が影響？

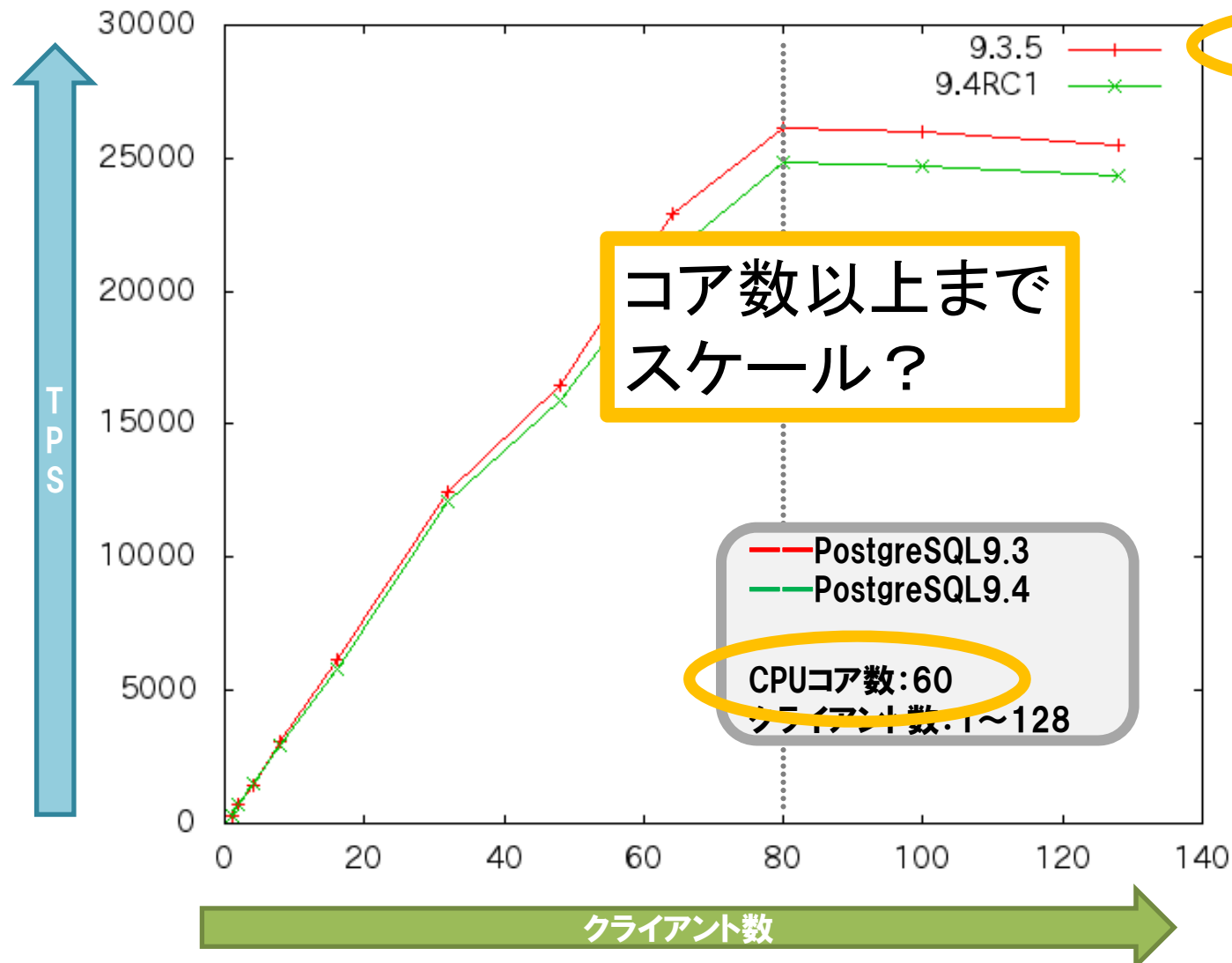
## 9.3 と 9.4 の参照性能の比較 <スケーラビリティ>



80 クライアント  
まで、ほぼリニア  
にスケール

9.4 と 9.3で  
同傾向のスケール  
ラビリティ特性  
を持っていること  
を確認

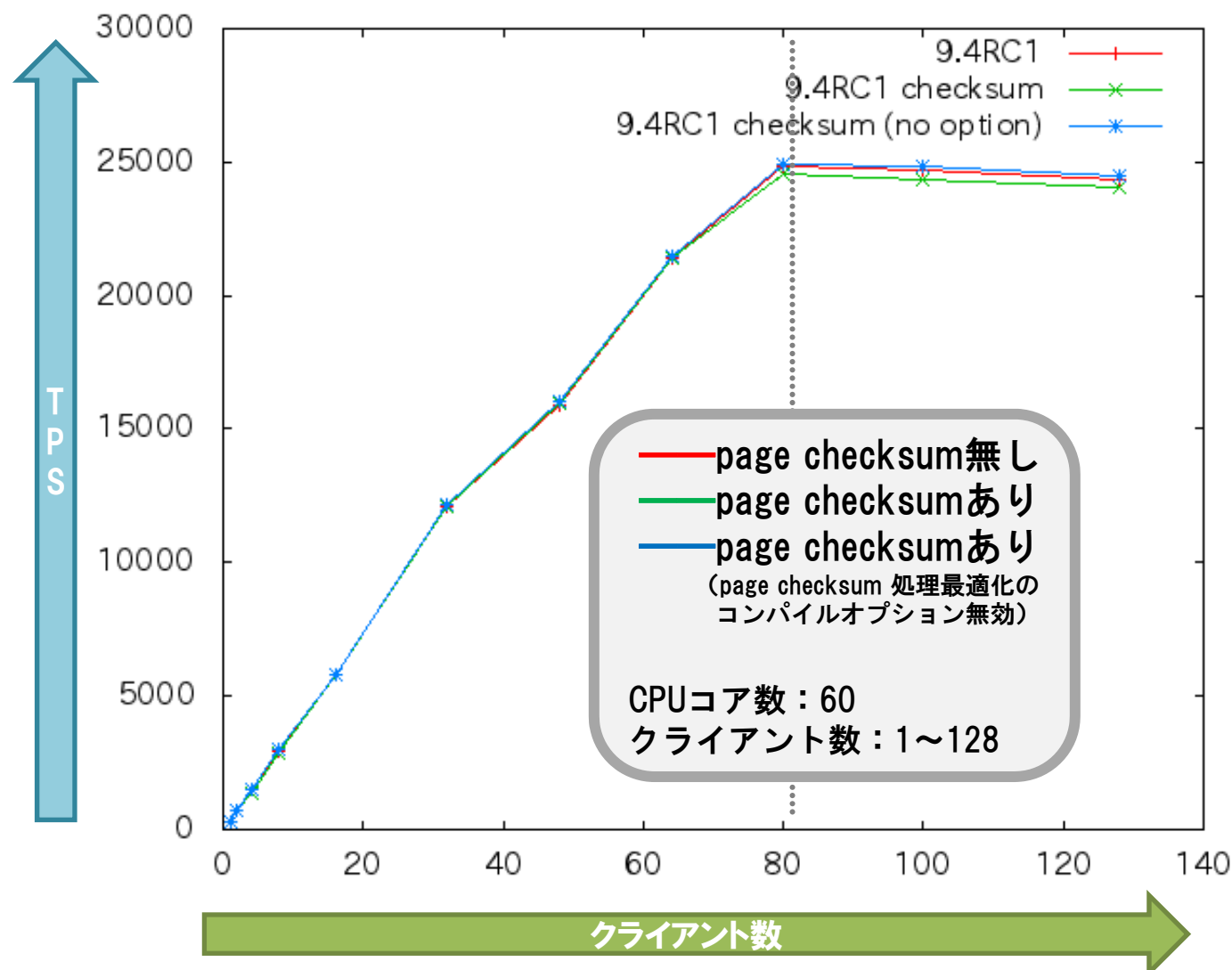
## 9.3 と 9.4 の参照性能の比較 <スケーラビリティ>



80 クライアント  
まで、ほぼリニア  
にスケール

クエリの負荷が  
不十分もしくは  
ネットワークレイ  
テンシの影響で、  
サーバのCPUを  
使い切れていな  
いのが原因と考  
えられる

## 9.4 page checksum有無での参照性能の比較



全条件でTPSに  
ほとんど差がな  
かった

9.4ではpage  
checksum が参  
照性能に与える  
影響は無視でき  
るほど小さい





活動報告2

# WAL改善/ スケールアップ検証 [更新系]

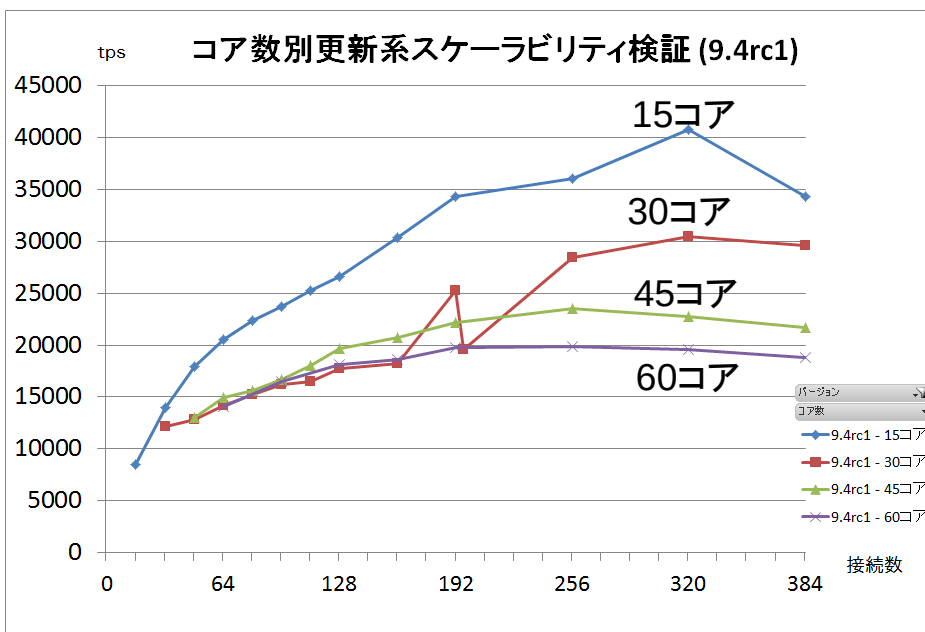
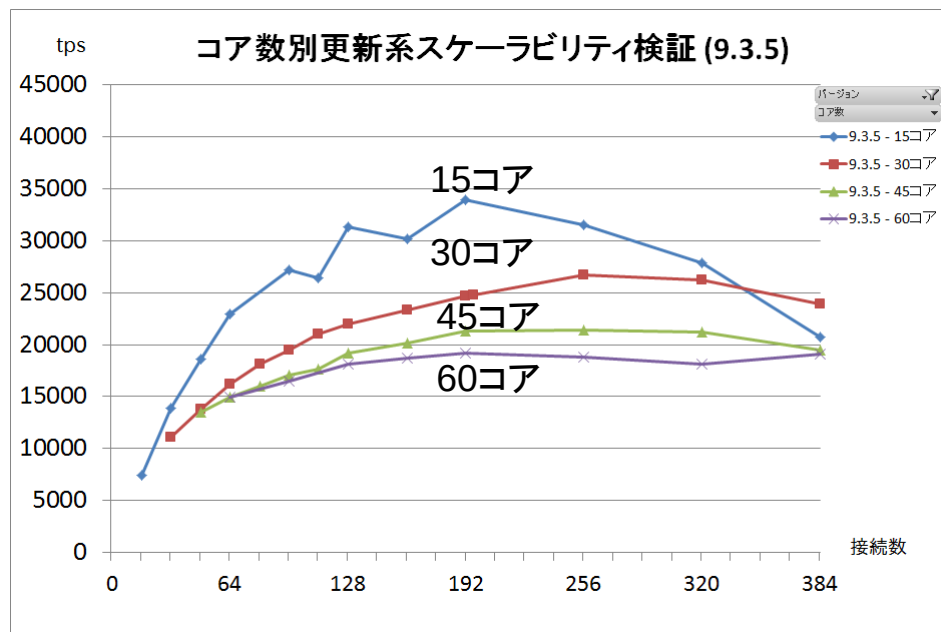
# WAL改善検証の背景・目的

- 2012年度、2013年度は参照系のスケーラビリティを確認してきたが、更新系は未実施であった
- PostgreSQL9.4では2つの更新系性能改善が実装された
  - WAL (Write Ahead Log)データ圧縮によるファイル出力量の軽減
  - WAL出力のロック競合の軽減による並列書込み性能の改善
- これら性能改善を機に、目的を以下の2つを確認することに設定
  1. PostgreSQLの更新系のCPUスケーラビリティの達成状況
  2. PostgreSQL9.3.5 とPostgreSQL9.4RC1 を比較し、更新性能の改善状況

## 測定条件・環境

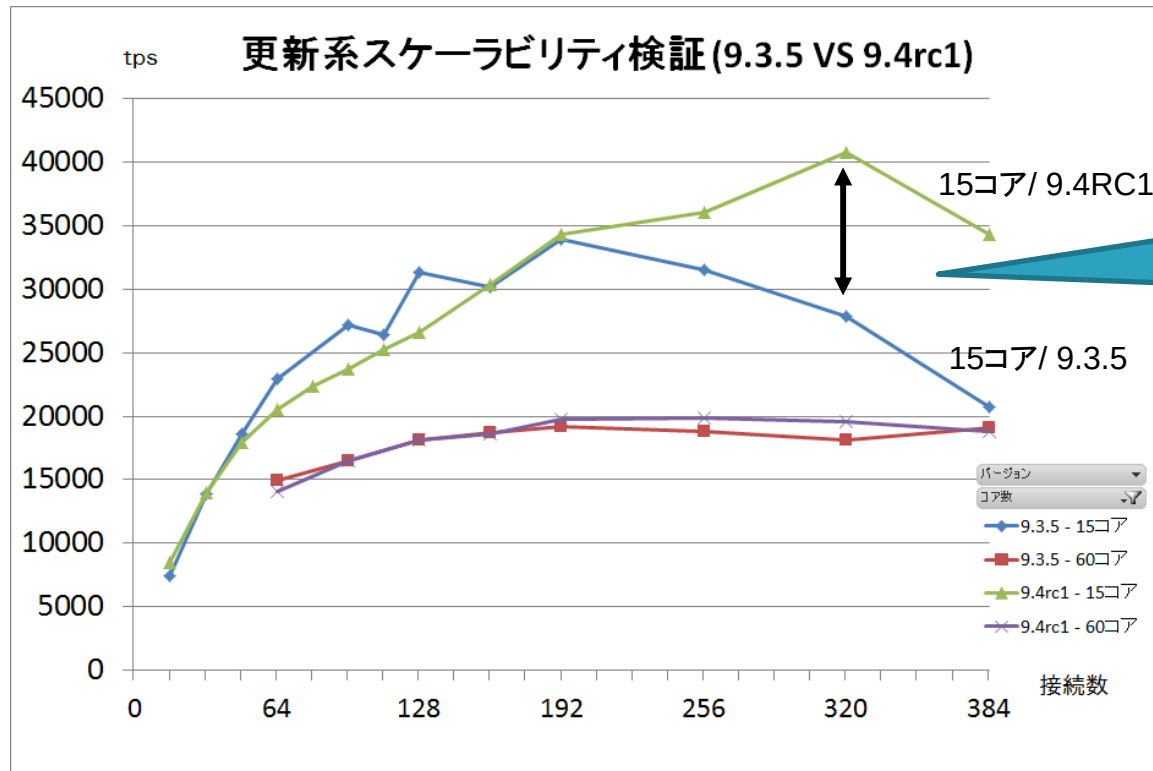
- マシンは後述の検証環境を使用
- 測定ツールはpgbenchを利用し、SF=7,000 (約100GB) でDBを作成
  - トランザクションは独自に1つの大きいテーブル (pgbench\_accounts) に対するUPDATEのみと設定し、ブロック排他の集中を避ける
  - FILLFACTOR=80で定義
- データベースクラスタ (/data) とWAL (/pg\_xlog) を分割した2ボリュームを使用
  - 各ボリュームは36台RAID10相当を4つに論理分割したもの
- サーバ側のコア数 (15,30,45,60コア) とpgbenchの接続数を変更して測定
- 測定中にCHECKPOINTは発生させないように設定
- 3回の計測結果の中央値を結果とする

## 9.3.5/9.4RC1コア数別スケールアップ検証結果



- 参照系とは異なり/0が発生し、コア数よりも多くの接続数でピーク値となる
  - 例: 9.3.5/15コア: 192接続      9.4RC1/ 15コア 320接続
- PostgreSQL 9.3.5、9.4RC1共に検証の中で最小の15コアが最大性能に
- 全体の傾向としてコア数が増加すると、スループットが低下している
  - 9.3.5では負荷が高いところで15コアと30コアが逆転しているところも
  - 9.4RC1では負荷が低いところで30コアが45コアよりも下がっている、原因は不明

# 15コア/60コアの9.3.5 VS 9.4RC1



9.3.5比で9.4RC1  
が最大46%性能改  
善している

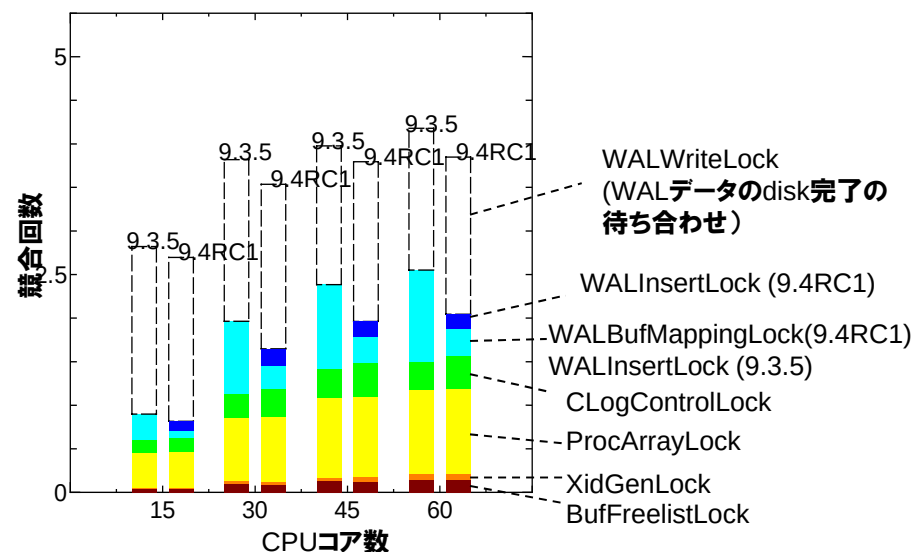
- コア数が少ない (15コア) 場合
  - 負荷(接続数)が高いところでは9.4RC1が高い性能が出ている
  - ピーク値も9.4RC1のほうが高い負荷に対応できている
- コア数が多い (60コア) 場合
  - 9.3.5と9.4RC1の傾向に差が出ていない

# 測定結果の分析

## ■ 9.4RC1の改善について

- ピーク時のストレージやCPU負荷がボトルネックでない可能性が高くロックの動作状況を確認
- WALInsertLockの負荷が実質的に下がっていることが有効であったと想定

1トランザクションあたりの  
PostgreSQL9.4RC1と9.3.5の  
実質的なロック競合想定回数



## ■ CPUスケーラビリティについて

- 同様にピーク時のストレージI/O、CPU負荷のボトルネックは確認できず
- コア数増加に伴い、ProcArrayLockを中心にロック競合の回数が増加しており、性能が伸びない一因であると推定

詳細は、2014年度WG1活動報告をご参照ください

## 結果まとめ

- (今回では特に15コアで) 9.4RC1が9.3.5と比較して更新性能が改善されることが確認できた
- 15/30/45/60コアの測定単位では、単純な更新のみのワークロードにおいて9.3.5、9.4RC1共にCPUスケーラビリティについて確認できなかった
  - ただし、参照系でのCPUスケーラビリティが確認できているため、採用システムのワークロード(更新・参照の割合等)に応じたCPU設計が必要であると判断
- 来年度以降の課題: 計測時のコア数粒度をもっと少なくし、コア数のピーク値と傾向を判断する
  - ピーク値は10コア? それとも20コア?



活動報告3

# 仮想化環境(KVM)検証



# 検証概要

## ■ 検証目的

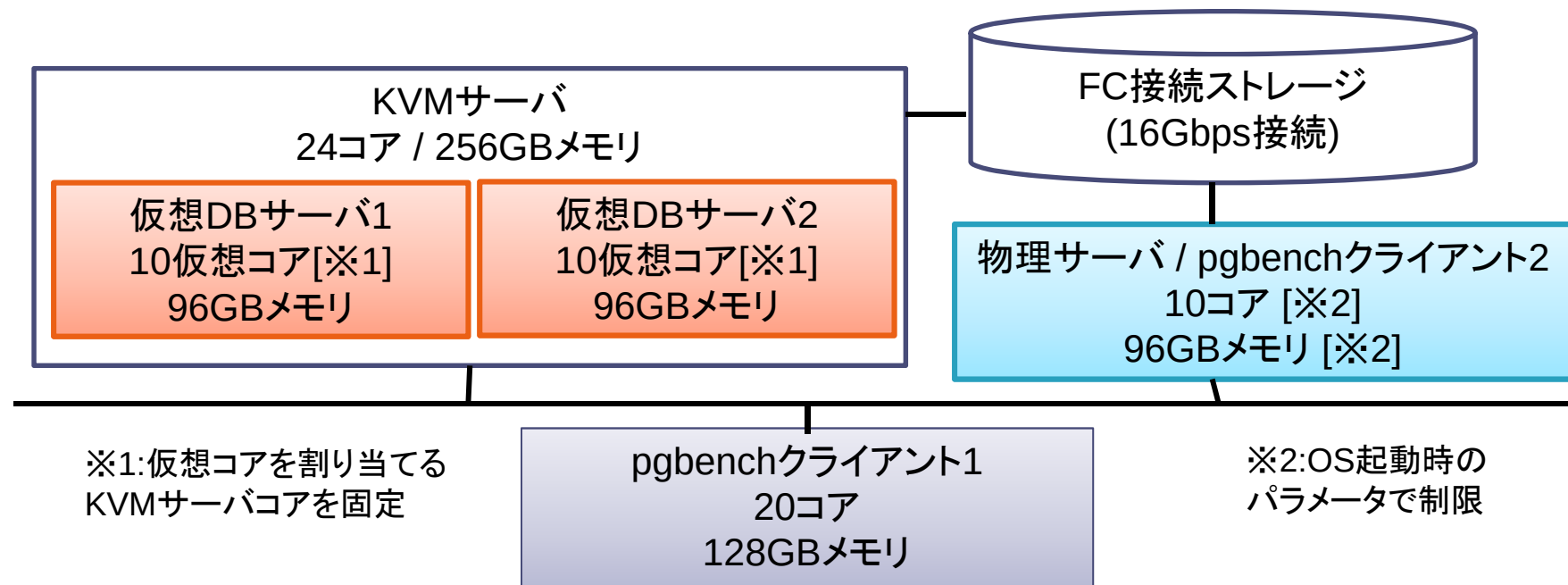
- 仮想環境におけるデータベース性能への影響調査
- ベンチマーク結果公開のため仮想化技術はKVMを使用

## ■ 検証の観点とシナリオ

- 参照系処理への影響の大きさを知る  
⇒シナリオ1:参照系ワークロードにて、仮想・物理の性能を比較
- 更新系処理への影響の大きさを知る  
⇒シナリオ2:更新系ワークロードにて、仮想・物理の性能を比較
- 仮想DBを同時に実行した場合の性能への影響を知る  
⇒シナリオ3:仮想DBの単独実行と仮想DBを同時に2つ実行した場合で性能を比較

# 測定環境

- KVMサーバに仮想サーバを二つ作成しそれぞれにDBを作成
- 物理環境と仮想環境のコア数とメモリ量を統一して比較
  - コア数: 10コア、メモリサイズ: 96GB
- WAL領域は仮想サーバ内から直接外部のストレージ領域に格納
- ベンチマークプログラムは定点観測 (参照系・更新系) に同じ



# 測定条件

## ■ 参照系：(シナリオ1)

**仮想化によるCPU処理のオーバヘッドの影響を確認する**

- スケールアップ検証 (参照系) と同じ pgbench のカスタムスクリプトを使用
- DBサイズは約15GBで、検索はオンメモリで処理される

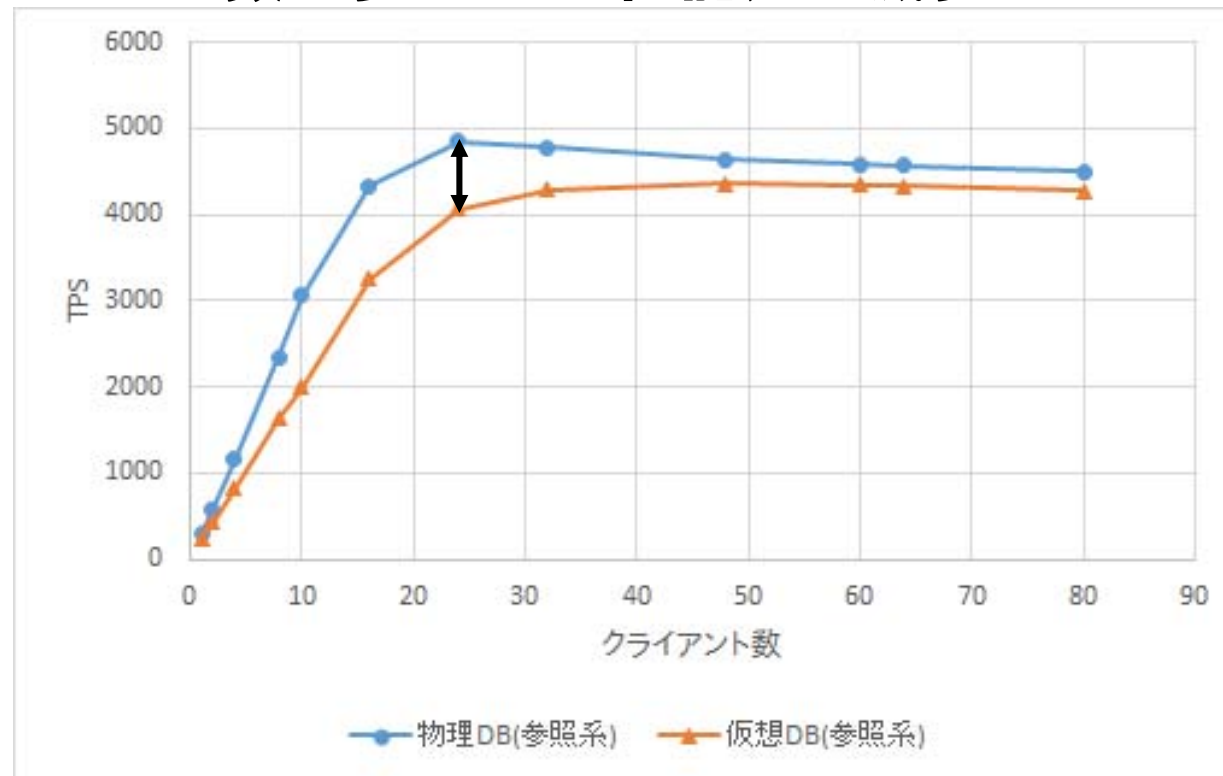
## ■ 更新系：(シナリオ2,3)

**仮想化によるストレージの書き込みへの影響を確認する**

- スケールアップ検証 (更新系) と同じくpgbenchを使用
- DBサイズは約100GBで、検索はほぼオンメモリで処理される
- ストレージへの書き込みはWAL (トランザクションログ) 出力のみとなるように、CHECKPOINTが実行されない設定をとする
  - DBクラスタは仮想マシン上のストレージプール、WALファイルは直接外部のストレージ領域に格納 (詳細は付録を参照)

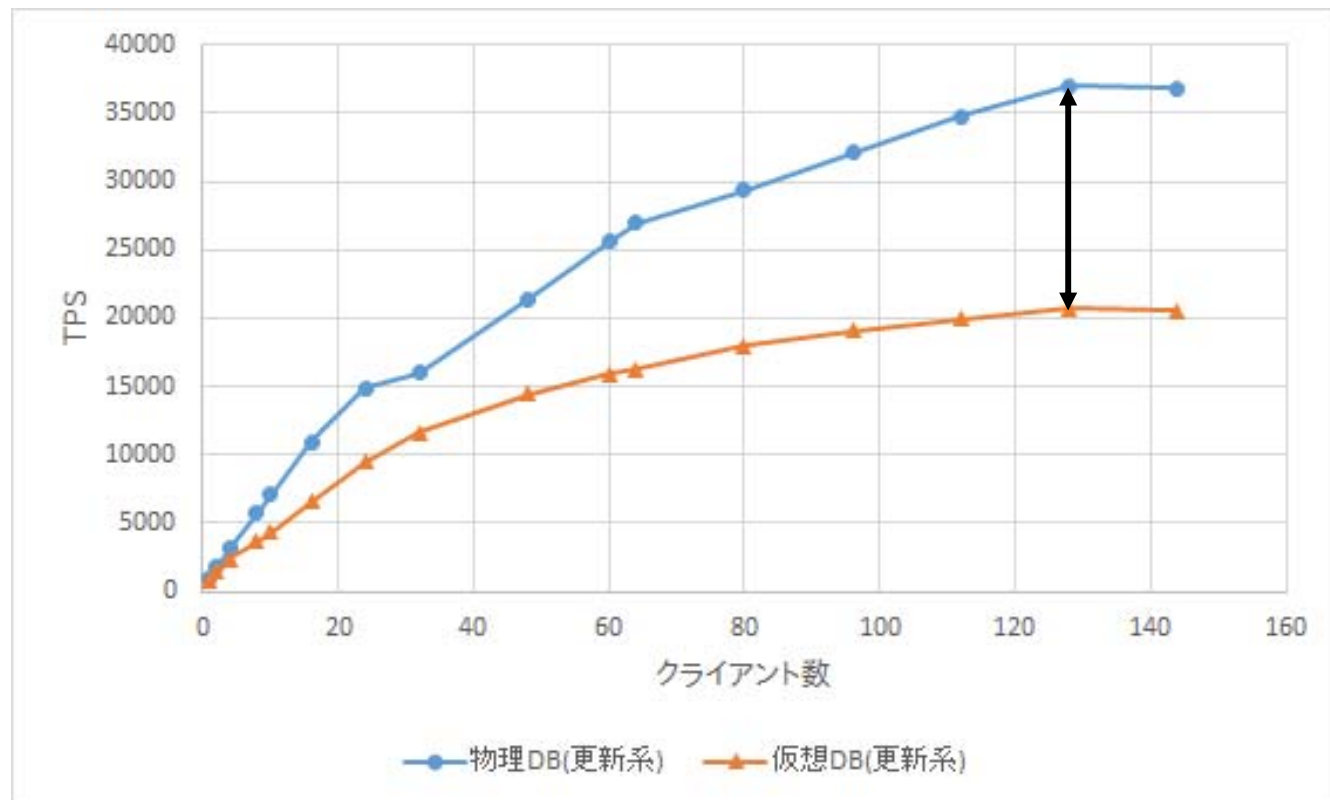
# シナリオ1：参照系での物理と仮想の性能比較

- ピーク時性能の差は約10%
  - 物理は24クライアント、仮想は48クライアントでピーク
  - 24クライアントで比較すると差は約16%
- クライアント数が多くなると性能差は減少



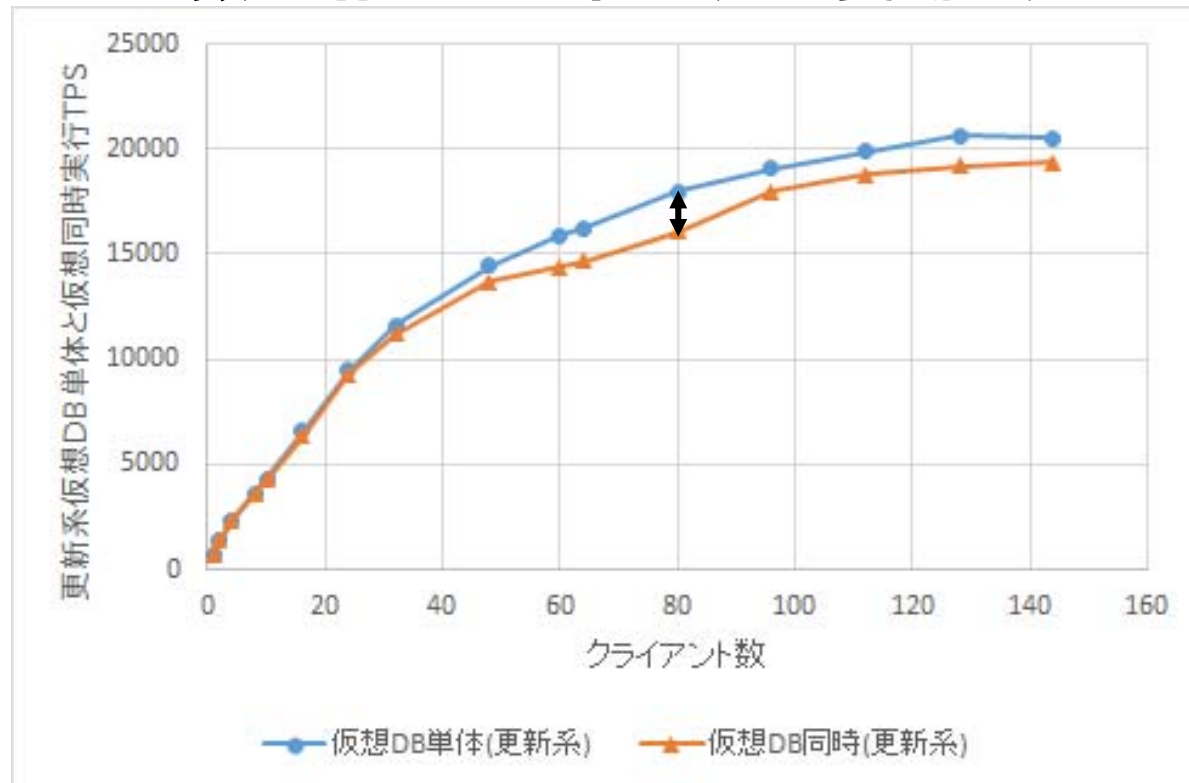
## シナリオ2: 更新系での物理と仮想の性能比較

- ピーク時性能の差は約40% (128クライアント時)
- クライアント数が増えると性能差が増大



## シナリオ3: 更新系仮想単体と仮想複数同時性能比較

- シナリオ2 (更新系) と同じベンチマークを使用
- ピーク性能の差は約7% (128クライアント時)
  - 最大で約10%の差が見られた (80クライアント時)
- クライアント数が増えても、性能差の変動は見られない



# 結果まとめ（参照系一検証1）

## ■ ピーク値

- 物理で4,900TPS、仮想で4,400TPS程度と約10%の違い
- ピーク値状態でのCPU利用率はともに、ほぼ100%
  - 仮想環境でのピーク値低下は仮想化によるオーバーヘッドによるものと思われる

## ■ ピーク値に達するクライアント数

- 物理で24、仮想で48と差がある
- 24クライアント（仮想）でのCPU利用率を見るとidleが見られることから、仮想化によるレイテンシの増加があると思われる



# 結果まとめ（更新系一検証2）

## ■ ピーク値

- 物理で37,000TPS、仮想で20,700TPS程度となり仮想環境では約半分の性能しか出ない
  - 物理・仮想で約20,000TPSの時のCPU利用率はほぼ同じ
    - CPU処理のオーバーヘッドが原因ではない
  - 仮想環境では、ピーク値が出ている時でも、CPU利用率のidleが20%を超えておりCPUを使いきれていない
- 検証3にて2DBを同時走行させたケースのピーク値は19,400TPS程度であり、検証2の仮想環境と大差はない
  - ホストOS側やその先のストレージ装置がボトルネックになっているのではなく、ゲスト-ホスト間のレイヤのどこかにボトルネックがあることが推測される。

## ■ ピーク値に達するクライアント数

- 物理・仮想ともに128で同じ



## 結果まとめ（単体と同時実行の比較―検証3）

### ■ ピーク値

- 単体走行で約20,700TPS、同時走行で約19,400TPSであり、単体の方が約7%優れる
- 同時走行した2つのDBで、ほぼ同じ
  - クライアント数を変えても、ほぼ同じスループットであった

### ■ ピーク値に達するクライアント数

- ピーク値に達するクライアント数は単体、同時走行ともに128であった



活動報告4

# コンテナ環境(Docker)検証

# コンテナ環境(Docker)での性能検証

## ■ 背景

- アプリケーションの可搬性の向上、運用コストの削減、デプロイの効率化などの理由でコンテナ実装のひとつであるDockerが注目されている
- しかし、PostgreSQLなどのOSSデータベースをコンテナ環境で動作させたときの性能を含めた検証はあまり行われていない

## ■ 目的

Dockerを上でPostgreSQLを動作させた際の性能を実機検証によって確認する

- オーバヘッドの有無とオーバヘッドを軽減する運用手法の調査
- 更新系・参照系での性能傾向の把握

## 測定条件・環境

- 測定ツールはpgbenchを利用し、SF=1,000 (約15GB) でDBを作成
  - 負荷かけサーバは20コア、メモリ128GB、Red Hat Enterprise Linux (以下RHEL) 6.5、PostgreSQL 9.4RC1
  - コンテナ検証サーバは24コア、メモリ256GB、RHEL6.6、PostgreSQL 9.4RC1  
(RHELが6.6なのは、6.5だとDockerをインストールできないため)
  - DockerコンテナのOSはRHEL6.5の公式イメージからインストール
- 実際に使用したマシンは付録をご覧ください

# 測定条件・環境(続き)

## ■ テスト共通

- Dockerを使わない場合と、Dockerを使う場合(4通り)を比較

| Dockerを使う場合のテストケース |       | ネットワーク |       |
|--------------------|-------|--------|-------|
|                    |       | コンテナ仮想 | ホスト直接 |
| ファイルシステム           | ホスト直接 | ケース1   | ケース2  |
|                    | コンテナ内 | ケース3   | ケース4  |

- 3回の計測結果の中央値を結果とする(各回5分)

## ■ 参照系のテスト

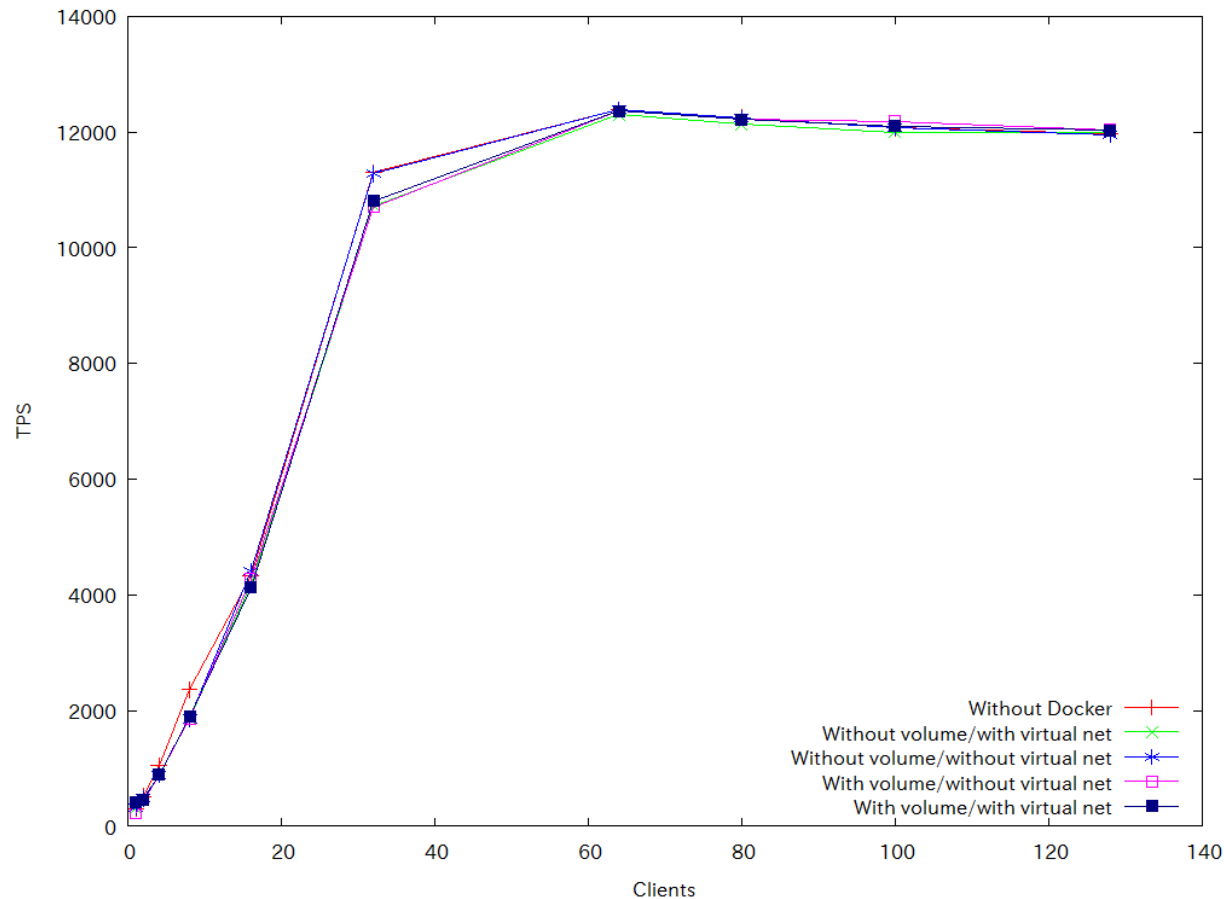
- クライアントからの同時接続数を1から128まで変化

## ■ 更新系のテスト

- 同時32ユーザのみ測定
- 小さなテーブルへ更新が集中してボトルネックになるのを避けるためにpgbench -Nを使用

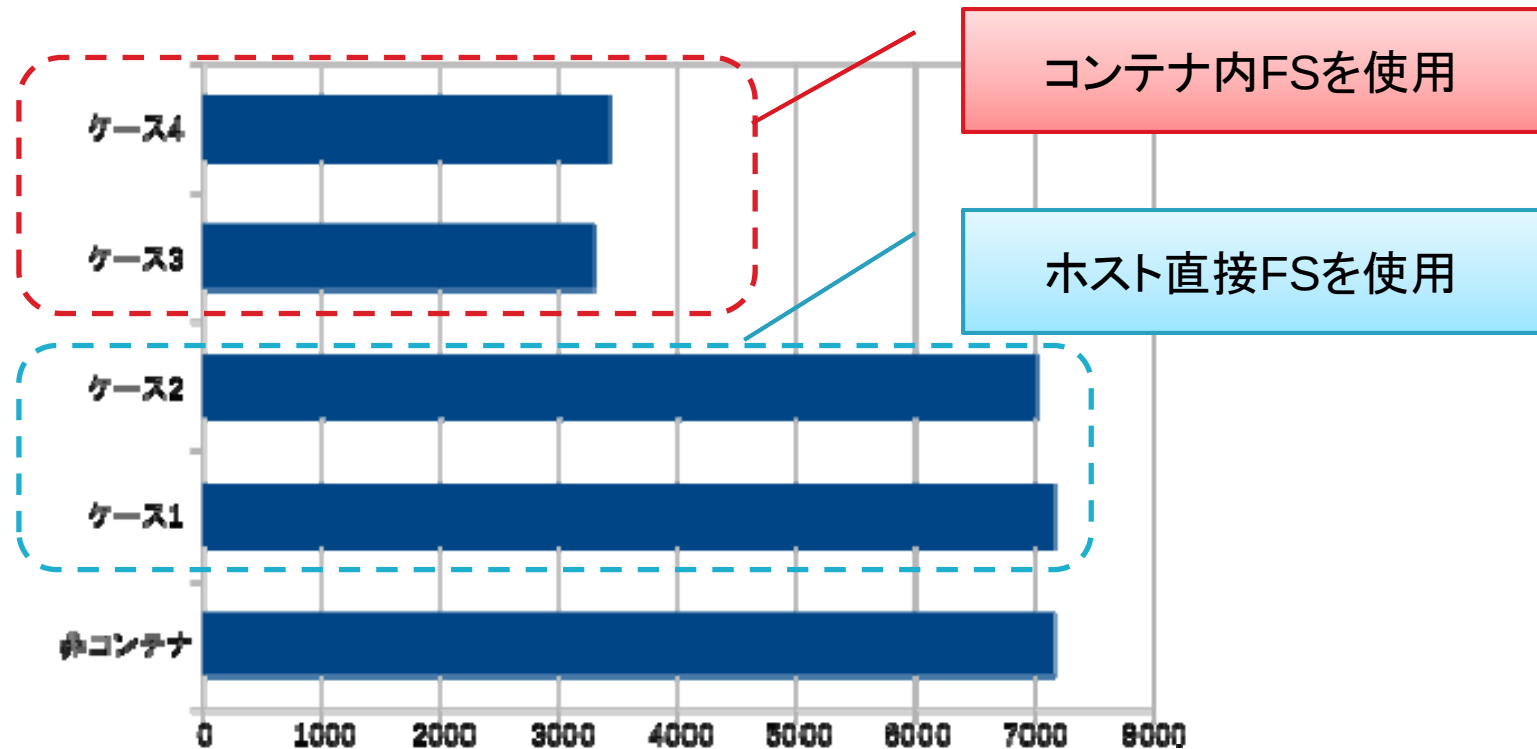
# 参照系性能測定結果

- すべての測定パターンではほぼ結果が同じ（誤差の範囲）
- 参照性能へのDockerの影響は最小限



# 更新系性能測定結果

- コンテナ内のファイルシステムを使うケースでは大きな性能低下が見られる



# 結果まとめ

## ■ 参照系

- オーバヘッドによる性能低下は見られない

## ■ 更新系

- ホストOSのファイルシステムにDBを格納するケースではオーバヘッドによる性能低下は見られない

このようなファイル配置とすることで、オーバヘッドによる性能低下を回避できる

- コンテナ内のファイルシステムにDBを格納するケースで大きく低下する

ただし、コンテナ内にDBを置くような運用はプロダクション環境ではあまり考えられず、実運用の障害にはならないと思われる

## ■ 全体的な性能の傾向

- オーバヘッドを除いて、物理環境と大きな違いはなく、Dockerにも不安定な挙動などは見られなかった



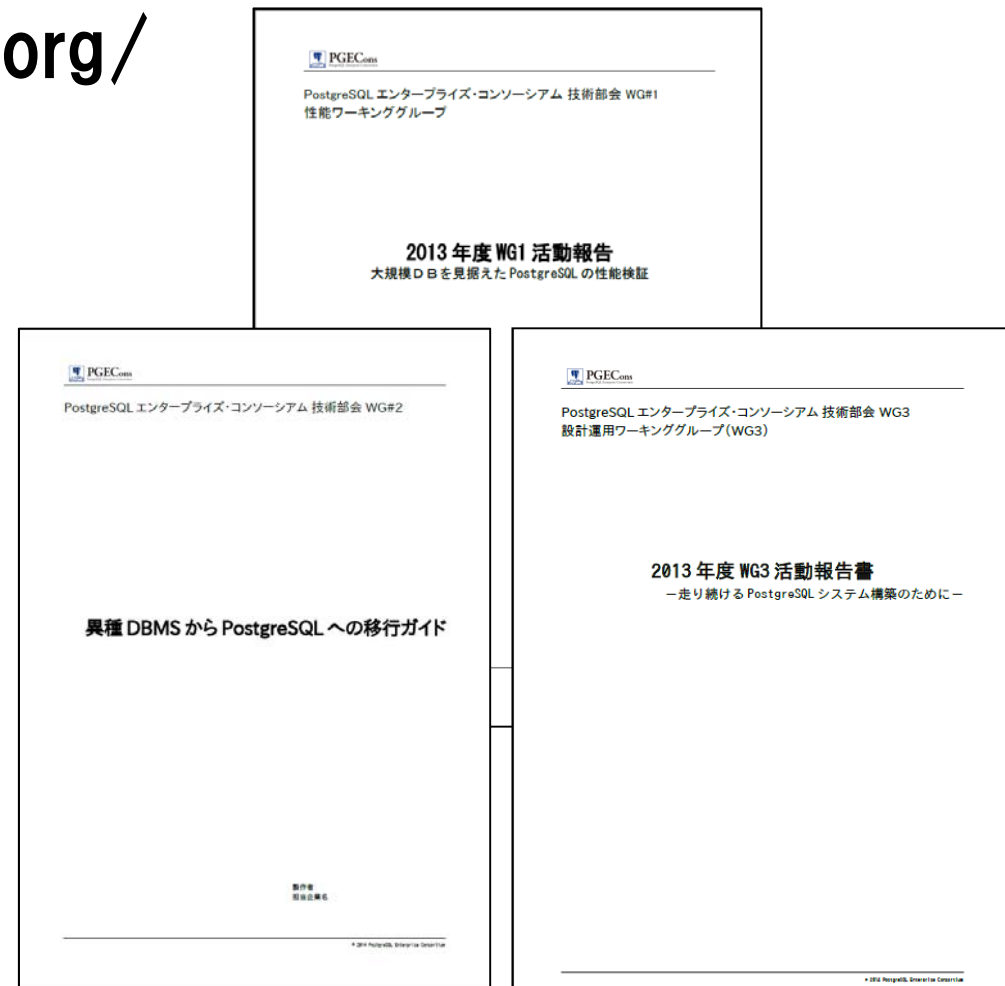


## 5. さいごに

# 成果物の公開

ダウンロード可能

- PGEConsサイトから、無料でダウンロードいただけます
- <http://www.pgecons.org/>



# 会員募集

- **正会員・一般会員を広く募集いたします**
  - WG・CRSと一緒に活動を行っていただける団体様  
⇒ 正会員
  - PostgreSQLのエンタープライズ領域に興味を持っている団体様 ⇒ 一般会員

**お問い合わせ先:**

**PostgreSQLエンタープライズ・コンソーシアム事務局**

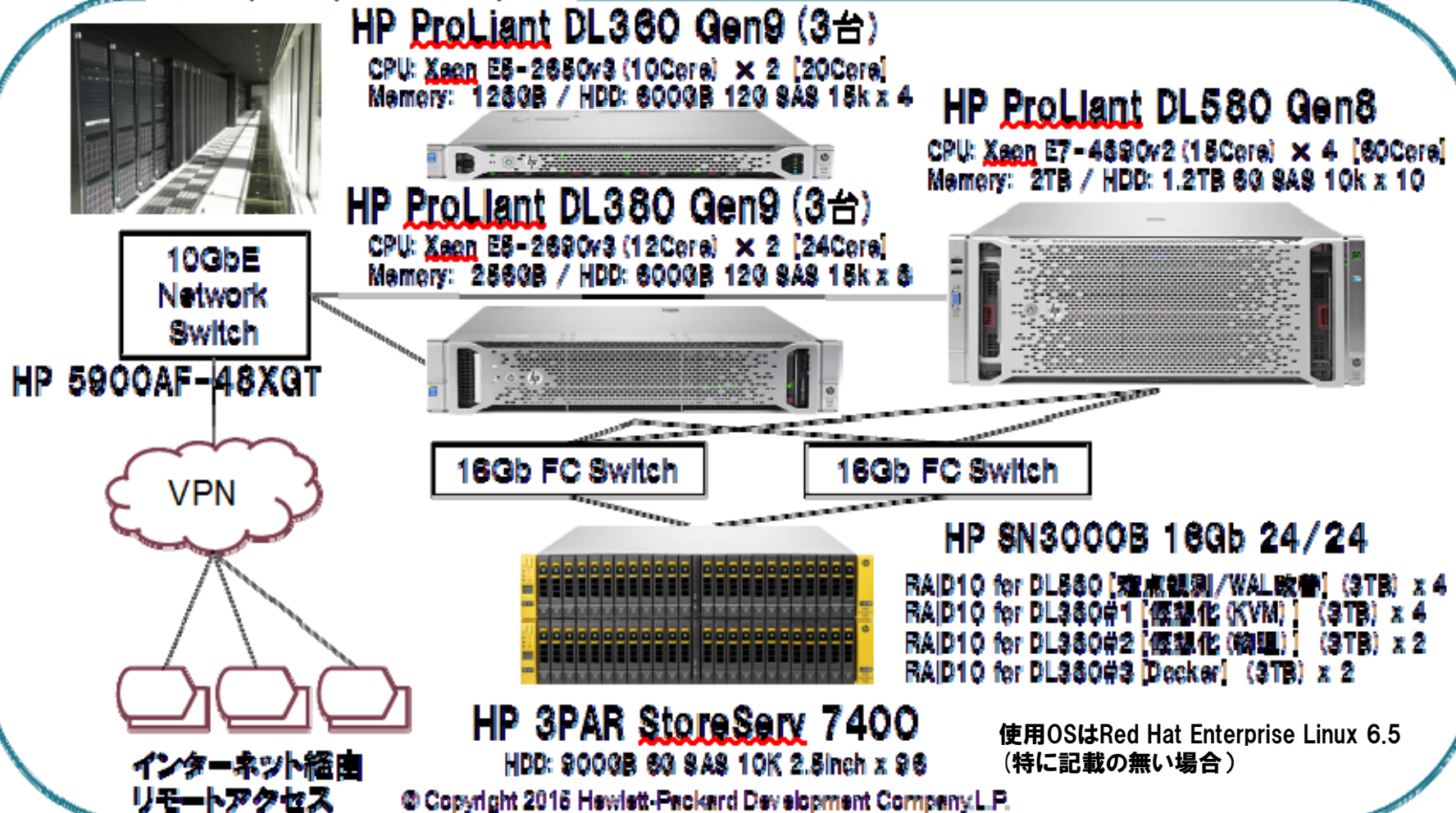
メール : [jimukyoku@pgecons.org](mailto:jimukyoku@pgecons.org)

Web : <http://www.pgecons.org/>

# 付録：検証環境 （提供：日本ヒューレット・パカード株式会社）

## ■ 2014年度WG1性能検証での使用機器/設備

### HPソリューションセンター



# 付録：仮想環境でのストレージ構成について

作成コマンド(ストレージプール)

```
#virsh pool-define-as vv13 fs - - /dev/mapper/mpathdp1 - /var/lib/libvirt/vv13/
```

- 仮想環境で使用するストレージは、仮想マシンを作成したストレージプールと、仮想マシンから直接mountするホスト上の外部ブロックデバイスの双方が利用可能。

```
#virsh attach-device vmpgsql01 ~/NewStorage.xml
```

作成コマンド(外部ストレージ)

```
<disk type='block' device='disk'>
  <driver name='qemu' type='raw' cache='none'/>
  <source dev='/dev/mapper/mpathbp1'/>
  <target dev='vdb' bus='virtio'/>
</disk>
```

NewStorage.xml

- 今回の更新系ベンチマークはCHECKPOINTが発生しないようにしてWALファイルへの書き込みが性能の上限となって、外部ブロックデバイスへのWALファイル配置が性能の決め手となった。

| パターン    | ファイルのマウント先 |            | 性能 |
|---------|------------|------------|----|
|         | データベースクラスタ | WALファイル    |    |
| 1       | ストレージプール   | ストレージプール   | 低  |
| 2(今回採用) | ストレージプール   | 外部ブロックデバイス | 高  |
| 3       | 外部ブロックデバイス | 外部ブロックデバイス | 高  |



**ご清聴ありがとうございました！**

**POSTGRESQL ENTERPRISE  
CONSORTIUM**



**PGEC**ons  
PostgreSQL Enterprise Consortium