

POWER8 + PostgreSQL 性能検証報告



SRA OSS, INC.

2014/8/19

SRA OSS, Inc. 日本支社
〒170-0022 東京都豊島区南池袋 2-32-8 8F
Tel. 03-5979-2701 Fax. 03-5979-2702
<http://www.sraoss.co.jp/>

目次

1. 本文書について.....	2
2. 検証内容.....	2
2.1. 概要.....	2
2.2. 実施環境.....	2
3. pgbench 標準参照トランザクションによるテスト.....	3
3.1. テスト条件詳細.....	3
3.2. テスト結果と考察.....	4
3.3. POWER7 テストとの比較.....	5
3.4. 同様検証結果との比較.....	6
4. pgbench カスタム参照トランザクションによるテスト.....	7
4.1. テスト結果と考察.....	8
4.2. PGEcons テスト結果との比較.....	9
付録.....	10
POWER7 テストの検証条件について.....	10
参考資料.....	11
本文書についての注意事項.....	11

1. 本文書について

本文書は、POWER8 プロセッサを採用した IBM Power System における PostgreSQL の性能検証を行った結果を報告するものです。本検証にあたり、日本アイ・ビー・エム株式会社様より検証対象機器を提供いただきました。検証作業は 2014 年 6 月から 7 月にかけて実施しました。

2. 検証内容

2.1. 概要

本検証は POWER8 が PostgreSQL 最新版において効果的に動作するかを明らかにすることを目的としています。POWER8 は1コアあたり8スレッドの並行処理が可能です。この CPU 性能、特に並行処理性能を検証するため、参照マルチアクセスのテストを以下2種類実施しました。

カスタム参照トランザクションは、PostgreSQL エンタープライズコンソーシアムが他アーキテクチャのマシンに対して性能検証を行い結果を公表しているものに合わせた内容となっています。

1. pgbench 標準参照トランザクションによるテスト
2. pgbench カスタム参照トランザクションによるテスト

2.2. 実施環境

今回検証に使用した環境は次の通りです。テスト項目によって OS を変更しています。

マシン	IBM Power System S822L CPU: POWER8 3.42GHz 2CPU × 各 10 コア (Hardware Thread 有効、1 コアあたり 8 スレッド) メモリ: DDR3 128GB
OS	Red Hat Enterprise Linux Server release 6.5 64bit Red Hat Enterprise Linux Server release 7.0 64bit 特別なカーネルチューニング等なし
PostgreSQL	PostgreSQL 9.3.4 ソースコードから OS 付属 gcc を使用し、以下オプションでビルド RHEL 7.0 の場合: CFLAGS="-O3 -mcpu=power8 -mtune=power8" ./configure --without-readline RHEL 6.5 の場合: CFLAGS="-O3 -mcpu=power7 -mtune=power7" ./configure --without-readline
pgbench	PostgreSQL 9.3.4 付属のもの

また、PostgreSQL 設定 (postgresql.conf) の値は次の通りです。それ以外は無指定 (デフォルト値) としています。

```
shared_buffers = 32GB
work_mem = 10MB
checkpoint_segments = 16
```

```
checkpoint_timeout = 30min
max_connections = 510
logging_collector = on
```

3. pgbench 標準参照トランザクションによるテスト

pgbench に用意された参照系のテスト(pgbench -S)を行いました。このテストは pgbench コマンド内に組み込まれた形でデフォルトで用意されているものであるため、他所で実施され、公開されているデータが多く、結果を比較しやすいといえます。

3.1. テスト条件詳細

pgbench -S コマンドは以下のトランザクションを実行します。

scale には pgbench -s で与えるパラメータ 1000 が代入されます。setrandom 命令で第二引数を最小値、第三引数を最大値とする整数乱数値が第一引数に格納されます。bid、tid、delta は使われない値となります。

```
\set nbranches :scale
\set ntellers 10 * :scale
\set naccounts 100000 * :scale
\setrandom aid 1 :naccounts
\setrandom bid 1 :nbranches
\setrandom tid 1 :ntellers
\setrandom delta -5000 5000
BEGIN;
SELECT abalance FROM pgbench_accounts WHERE aid = :aid;
END;
```

テスト条件は以下の通りです。

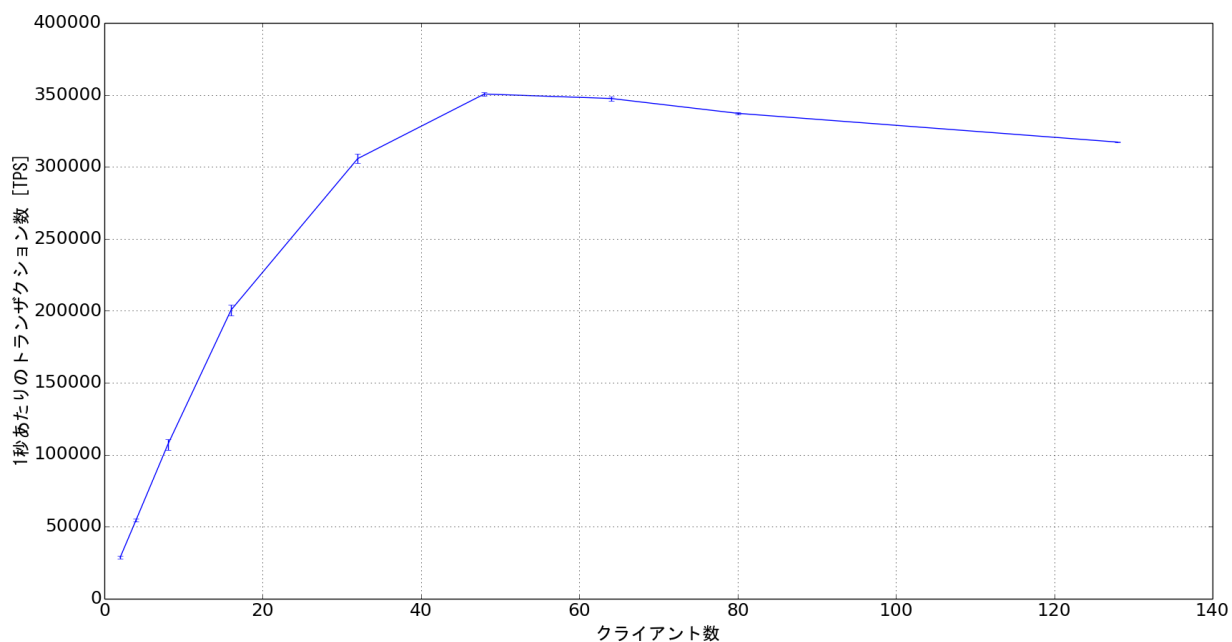
同時接続数(-c 指定)	2,4,8,16,32,48,64,80,128 の各値を指定
データ初期化パラメータ (test はデータベース名)	initdb --no-locale createdb test pgbench -i -s 1000 test
テストパラメータ	pgbench -n -S -c (同時接続数) -j (同時接続数/2) -T 300 test
各同時接続数の試行回数	5
その他	データ初期化は1回だけ行う。 最初に pgbench -S -T 600 を実行する。 その後に各同時接続数のテストをランダム順序で行う。 各テストの間に 60 秒 の休止を入れる。

また、OS に Redhat Enterprise Linux 6.5 を使用しました。

3.2. テスト結果と考察

結果を以下に示します。縦軸が1秒あたりのトランザクション数(グラフ内で「TPS」と記載)、横軸が同時接続数(グラフ内で「クライアント数」と記載)です。値は平均値です。

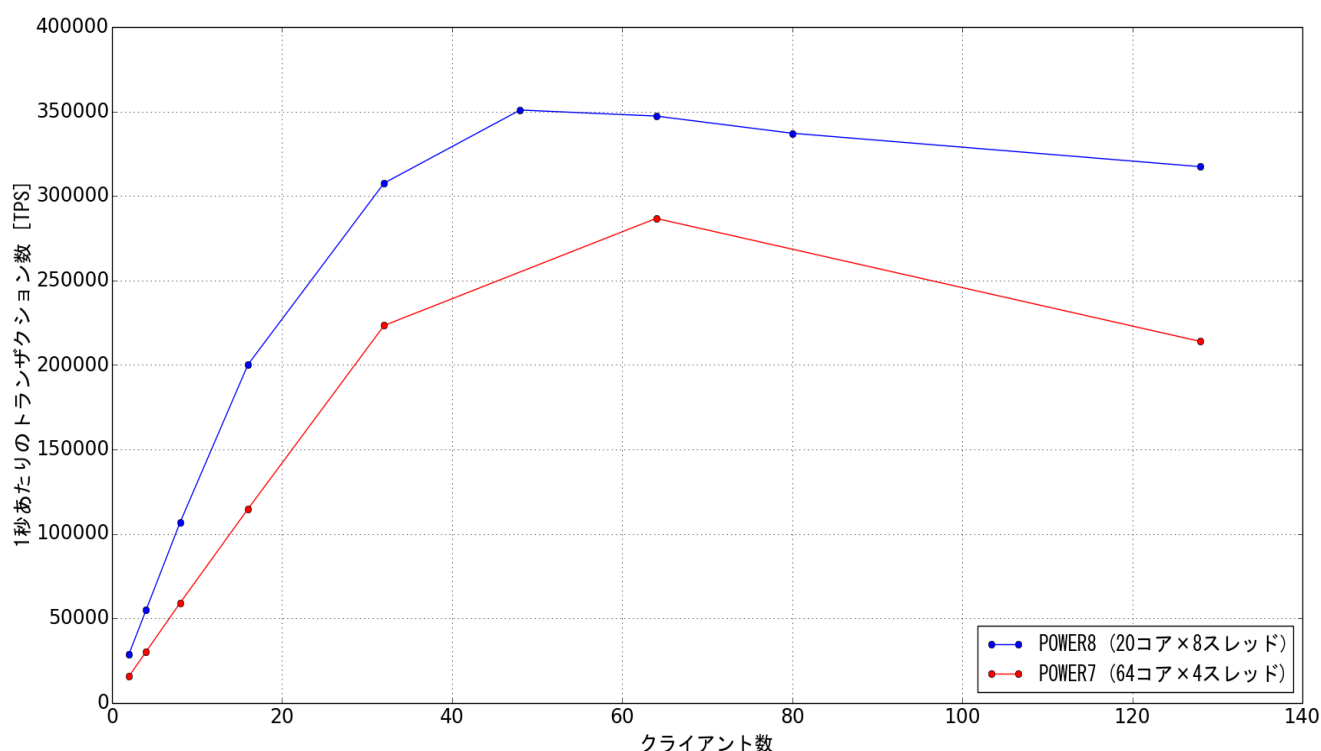
エラーバーは標準偏差を上下に取っています。線の長さが標準偏差の2倍となります。本テスト結果のバラつきは小さく、すべての水準(同時接続数)で標準偏差は値の5%以下、40 同時接続以上の結果だけでいえば1%以下でした。



事前想定としては、物理コア数 20 であることから 20 接続まで比例して TPS が伸び、それ以降は同時接続数を増やしても TPS は下がっていくことが予想されましたが、実際には、50~60 接続くらいにピークがあるという結果となりました。これは物理コア数の2~3倍です。

3.3. POWER7 テストとの比較

比較のため、前回 POWER7 マシンで行った検証結果を合わせてグラフにしますと以下ようになります。前回テストに合わせて値は中央値としています。



POWER7 テストの方が対象マシンの物理コア数(20 vs 64)、総スレッド数(160 vs 256)ともに多いにもかかわらず、POWER8 テストの方が全体的に良好な結果がでています。POWER7 テストでは物理コア数 64 を超えた同時接続数では性能ダウンが起きています。POWER8 テストでは物理コア数の2~3倍くらいまでの同時接続数において右肩上がりとなり、また、同時接続数 128 に至っても性能ダウンの度合は緩やかです。

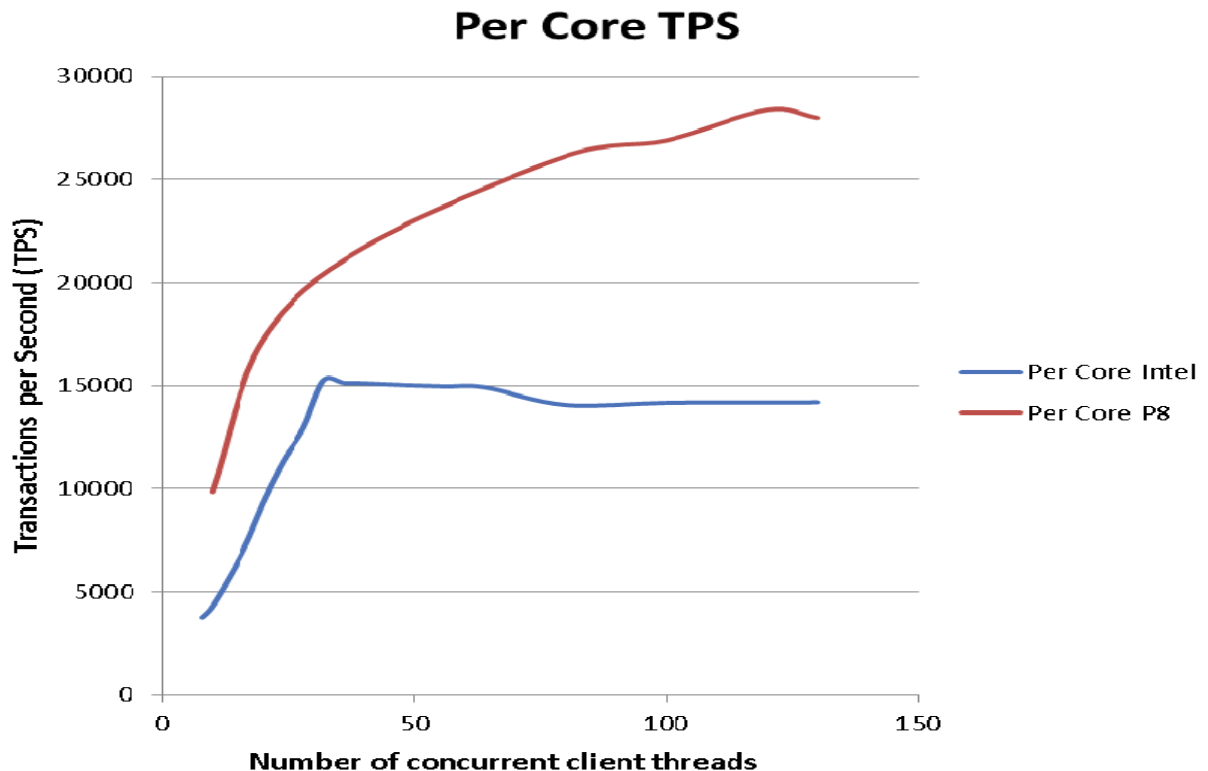
ハードウェアスレッドがより有効に働いて処理性能に貢献しているものと考えられます。

前回の POWER7 テストはいくつかの点で実施条件が異なります。詳細は付録「POWER7 テストの検証条件について」節に記載しています。

3.4. 同様検証結果との比較

Postgres Plus Advanced Server 9.3 (PostgreSQL 9.3 をベースとした商用製品のひとつ)にて、POWER8 のマシン、インテル CPU (Ivy Bridge アーキテクチャ) のマシンとで比較した検証結果がありますので、本節で紹介します。

以下のような結果となっています。



グラフは pgbench の参照テスト (-S オプションによる SELECT only テスト) の結果をコア数あたりの時間当たりトランザクション数でプロットしたものです。コア数あたりで比較するとインテル CPU を大きく上回っていることがわかります。

本テストは POWER8 が 10 コア、Intel CPU が 16 コアで試験を行っています。上記グラフの検証の方が、本検証よりも同時接続数が多い領域において良好な成績となっているのは、カーネルおよびデータベースに対するチューニングを加えた状態での結果であるためと考えられます。

上記グラフの検証レポートは IBM Partner World サイトにて公開されています。テスト条件の詳細につきましては、当該レポートを参照ください。参考資料の節にて URL を示します。

4. pgbench カスタム参照トランザクションによるテスト

PostgreSQL エンタープライズコンソーシアム (以下 PGECons と表記) にて、PostgreSQL 9.3 系バージョンの性能を検証するために pgbench にカスタムトランザクションを与える負荷テストをインテル CPU のマシンを対象に行っています。この結果と比較できるようにするため、同じ内容のカスタムトランザクションにてテスト実施しました。

本テストは標準参照系テストよりも負荷が高いものとなっております。トランザクション内容は以下の通りです。ランダムな範囲の値の件数を応答するというものです。ここでも bid、tid、delta は使われない値となります。

```
¥set nbranches :scale
¥set ntellers 10 * :scale
¥set naccounts 100000 * :scale
¥set row_count 10000
¥set aid_max :naccounts - :row_count
¥setrandom aid 1 :aid_max
¥setrandom bid 1 :nbranches
¥setrandom tid 1 :ntellers
¥setrandom delta -5000 5000

SELECT count(abalance) FROM pgbench_accounts WHERE aid BETWEEN :aid and :aid +
:row_count
```

設定パラメータは以下の通りです。

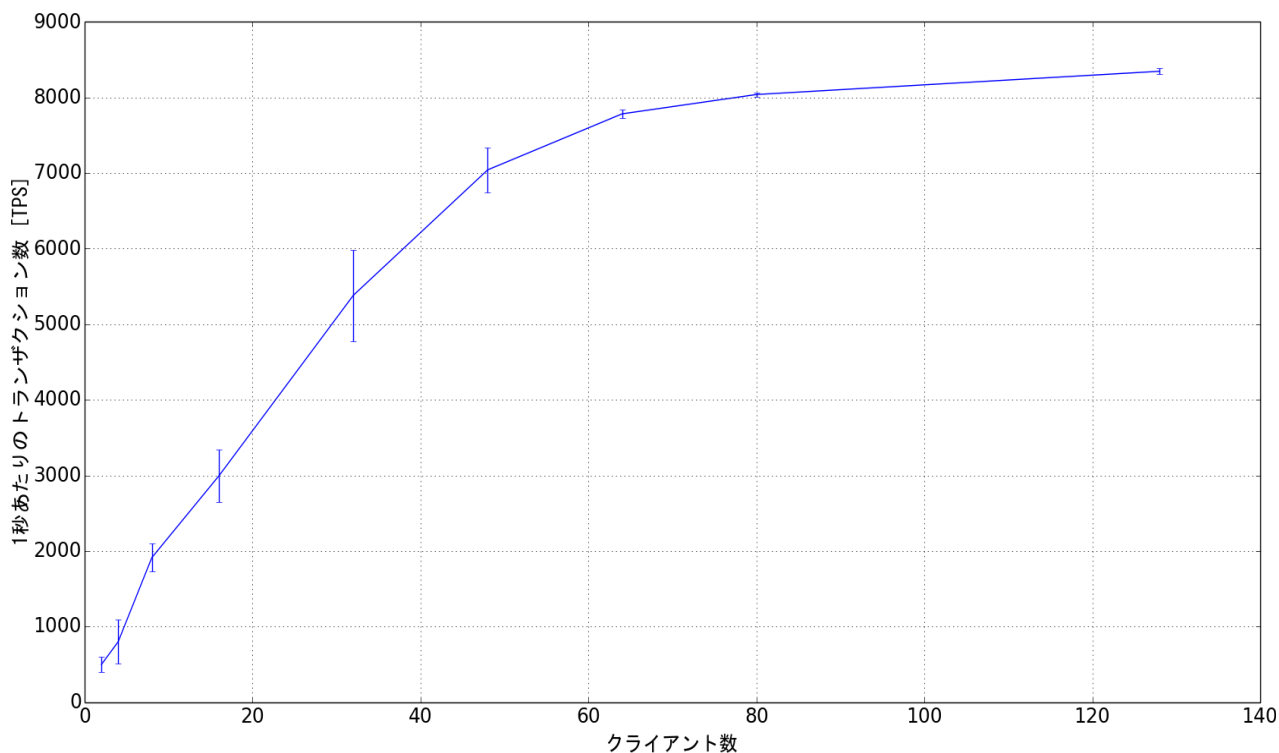
物理コア数	20 (論理 160)
同時接続数	変動 (2,4,8,16,32,48,64,80,128)
データ初期化パラメータ (test はデータベース名)	initdb --no-locale createdb test pgbench -i -s 1000 test
負荷パラメータ	pgbench -n -f custom.sql -c (同時接続数) -j (同時接続数/2) -T 300 test
その他	データ初期化は1回だけ行う。 最初に pgbench -S -T 600 を実行する。 その後に各同時接続数のテストをランダム順序で行う。 各テストの間に 60 秒 の休止を入れる。

本テストでは Redhat Enterprise Linux 7.0 を使用しました。

4.1. テスト結果と考察

結果を以下に示します。縦軸が1秒あたりのトランザクション数(グラフ内で「TPS」と記載)、横軸が同時接続数(グラフ内で「クライアント数」と記載)です。値は平均値です。

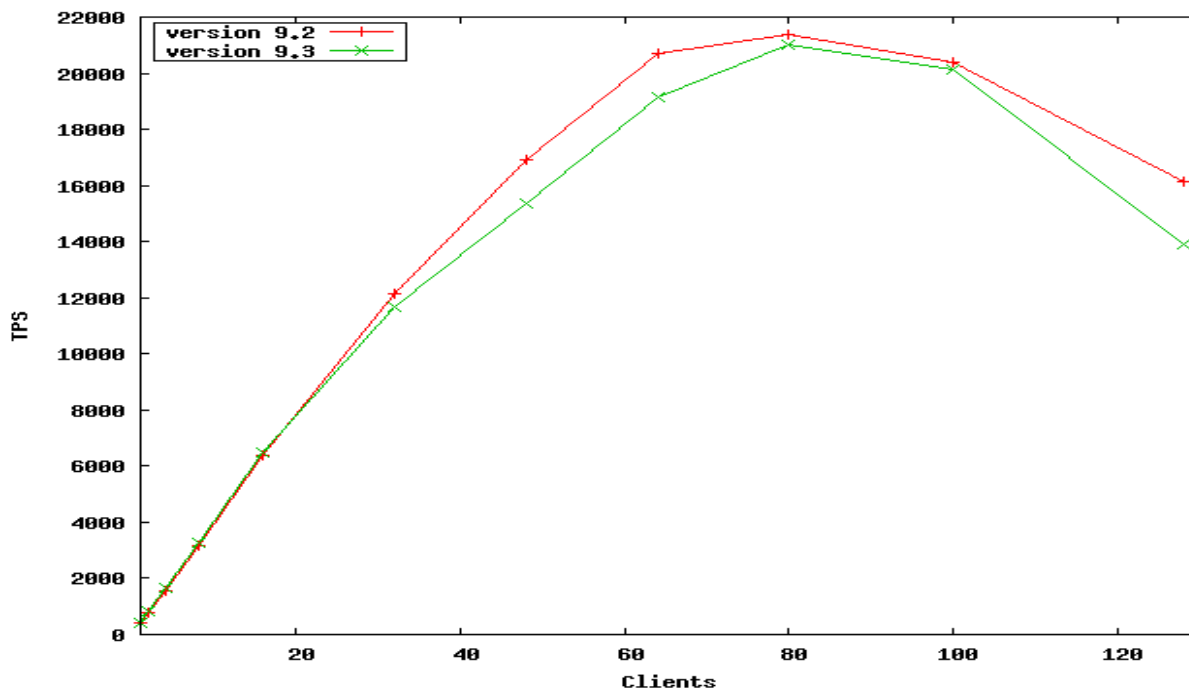
エラーバーは標準偏差を上下に取っています。線の長さが標準偏差の2倍となります。



40 接続くらいまで接続数に対して直線的に性能が向上し、80 接続程度から緩やかな傾きになるものの、同時 128 接続まで、時間あたり処理トランザクション数が右肩上がりとなっています。pgbench 標準参照テストの結果と同様に POWER8 では、物理コア数を超えて有効な並列処理をこなせることがわかります。

4.2. PGEcons テスト結果との比較

PGEcons (PostgreSQL エンタープライズコンソーシアム) の 2013 年度 WG1 (ワークグループ1) 成果資料から、同様のカスタムテストの結果を紹介します。



上記グラフのテストは PostgreSQL 9.2 と 9.3 を比較するために実施されたものです。総物理コア数 80 のインテル CPU (Ivy Bridge アーキテクチャ) マシンを使っています。物理コア数分までの同時接続数に対して TPS が向上し、それ以降は比較的急速に性能ダウンすることがわかります。

ただし、TPS の絶対値としては、上記グラフのテスト結果の方が高い値となっています。上記のインテル CPU マシンの結果ではピークが 21000 TPS 程度であり、本検証の POWER8 での結果ではピークは 8400 TPS 程度となります。物理コア数が 1/4 (25%) であるのに対して 40% 程度の性能を実現できていることになります。

PGEcons によるテストについての詳細は、当該レポートを参照ください。PGEcons の Web サイトで公開されています。参考資料の節にて URL を示します。

付録

POWER7 テストの検証条件について

3 節で 2012 年 6 月に実施した POWER7 検証結果と比較するグラフを掲載しましたが、テスト条件が下表のように異なります。

	POWER7 テスト	POWER8 テスト
検証マシン	IBM Power 780 server CPU: POWER7 3.86GHz 64 コア (Hardware Thread 有効、 1 コアあたり 4) メモリ: DDR3 256GB	IBM Power System S822L CPU: POWER8 3.42GHz 20 コア (Hardware Thread 有効、 1 コアあたり 8) メモリ: DDR3 128GB
PostgreSQL バージョン	9.2 beta 2	9.3.4
各水準の試行数	3 (グラフの値は中央値)	5 (グラフの値は中央値)
スケールファクター	100	1000
PostgreSQL 設定	shared_buffers = 16GB work_mem = 1GB checkpoint_segments = 16 checkpoint_timeout = 30min	shared_buffers = 32GB work_mem = 10MB checkpoint_segments = 16 checkpoint_timeout = 30min max_connections = 510 logging_collector = on

しかしながら、以下の観点から比較可能であるものと判断しています。

- 全てのデータが共有バッファ上に載るように設定されている
- PostgreSQL 9.2.x から 9.3.x へのバージョンでは、単純な SQL による参照の多重アクセスにおける性能向上要素は無い(他の各種の試験で性能向上が無いことが確認されている)

参考資料

[1] IBM Partner World

<https://www-304.ibm.com/partnerworld/wps/servlet/ContentHandler/partnerworld-public>

"IBM Power Systems solution for EnterpriseDB Postgres Plus Advanced Server "

[https://www-](https://www-304.ibm.com/partnerworld/wps/servlet/ContentHandler/stg_ast_sys_wp_power-systems-solution-enterprise-db-postgres)

[304.ibm.com/partnerworld/wps/servlet/ContentHandler/stg_ast_sys_wp_power-systems-solution-enterprise-db-postgres](https://www-304.ibm.com/partnerworld/wps/servlet/ContentHandler/stg_ast_sys_wp_power-systems-solution-enterprise-db-postgres)

[2] PostgreSQL エンタープライズコンソーシアム

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

"2013 年度 WG1 活動報告"

<https://www.pgecons.org/downloads/43>

本文書についての注意事項

本文書では TM マーク等を明記していませんが、文中の製品名は一般に各社の登録商標または商標となります。

本文書は、SRA OSS, Inc. 日本支社が作成したものです。SRA OSS, Inc. 日本支社は、本文書に対して、正確性、有用性、その他いかなる保証をするものではありません。本文書の内容を利用する場合、利用者の責任において行っていただくものとなります。