



OSDL-DBT-3 による PostgreSQLの評価

(株)SRA オープンソースカンパニー
松田 亮一

データベースのパフォーマンス(1)

■ インデックスを効果的に使う

- 検索対象行数、テーブル行数にもよるが、全行を読み取るシーケンシャルスキャンよりも、インデックススキャンの方が、圧倒的に速い。

■ キャッシュを効果的に使う

- PostgreSQLが管理するキャッシュ
- OSが管理するファイルキャッシュ

データベースのパフォーマンス(2)

■ インデックスが効果的に使えないケース

- 集約クエリー、全行検索クエリーなど

```
SELECT AVG(カラム) FROM テーブル;
```

```
SELECT * FROM テーブル WHERE カラム LIKE '%キーワード%';
```

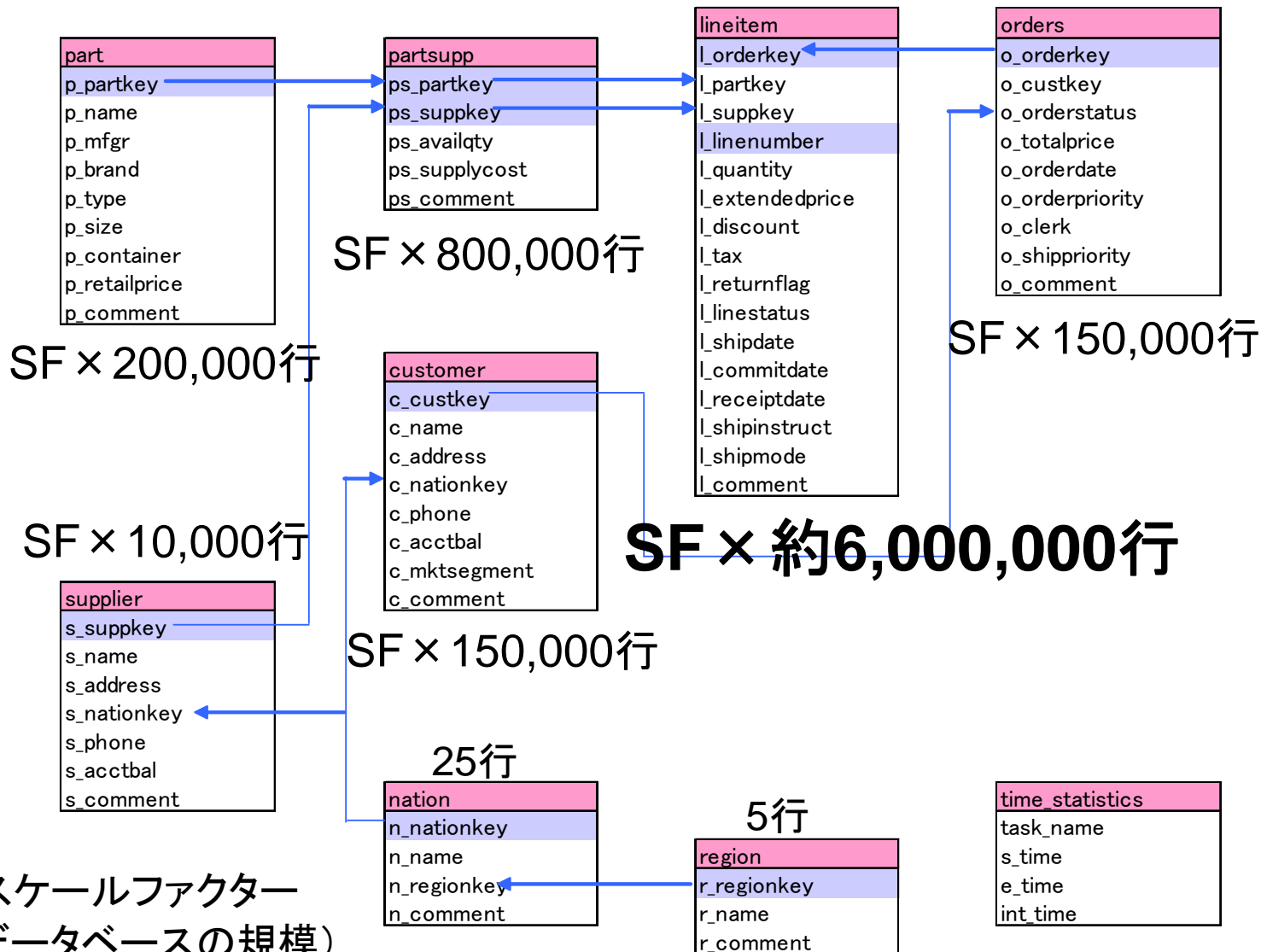
■ キャッシュが効果的に使えないケース

- データベースの規模が大きくて、メモリに入りきらない。キャッシュから、あふれてしまう。

OSDL-DBT-3とは、

- 意志決定支援環境のベンチマーク
 - 大規模なデータベース
(キャッシュが効果的に使えない)
 - 複雑なクエリー
(インデックスが効果的に使えない)
- TPC-Hを簡略化した物
 - DBT-3の実行結果は、TPC-Hとは無関係です。

大規模なデータベース



※SF=スケールファクター
(データベースの規模)

複雑なクエリー

- 22個の検索クエリー
(意志決定支援情報の検索)

Q7	<pre>select supp_nation, cust_nation, l_year, sum(volume) as revenue from (select n1.n_name as supp_nation, n2.n_name as cust_nation, extract(year from l_shipdate) as l_year, l_extendedprice * (1 - l_discount) as volume from supplier, lineitem, orders, customer, nation n1, nation n2 where s_suppkey = l_suppkey and o_orderkey = l_orderkey and c_custkey = o_custkey and s_nationkey = n1.n_nationkey and c_nationkey = n2.n_nationkey and ((n1.n_name = 'JAPAN' and n2.n_name = 'VIETNAM') or (n1.n_name = 'VIETNAM' and n2.n_name = 'JAPAN')) and l_shipdate between date '1995-01-01' and date '1996-12-31') as shipping group by supp_nation, cust_nation, l_year order by supp_nation, cust_nation, l_year;</pre>
----	--

- 2個のリフレッシュクエリー
(システム運用中のデータの増減をエミュレート)

RF1	<pre>insert into lineitem ...; Inesrt into orders ...;</pre>
RF2	<pre>delete from lineitem ...; delete from orders ...;</pre>

DBT-3で測定する内容

- 以下の3つテストを連続実行する。

- ロードテスト

- データベースにデータをファイルからロードした時間を測定する。
 - データファイルのサイズは、「スケールファクター×1ギガバイト」。

- パワーテスト

- 同時に実行するトランザクション数=1
 - 22+2個のクエリーを実行した時間を測定する。

- スループットテスト

- 同時に実行するトランザクション数=任意指定(今回は4)
 - 22+2個のクエリーを実行した時間を測定する。

測定結果

■ 測定結果は、HTMLレポート

Compsite

下記2つの平均値。

Query Processing Power

同時トランザクション数=1で、
「1時間あたりに実行できる
クエリー数 × SF」

Throughput Numerical Quantity

同時トランザクション数=nで、
「1時間あたりに実行できる
クエリー数 × SF」

今回はトランザクション数=4

↓スクロールすると、
詳細な情報がある

DBT-3 Test Result

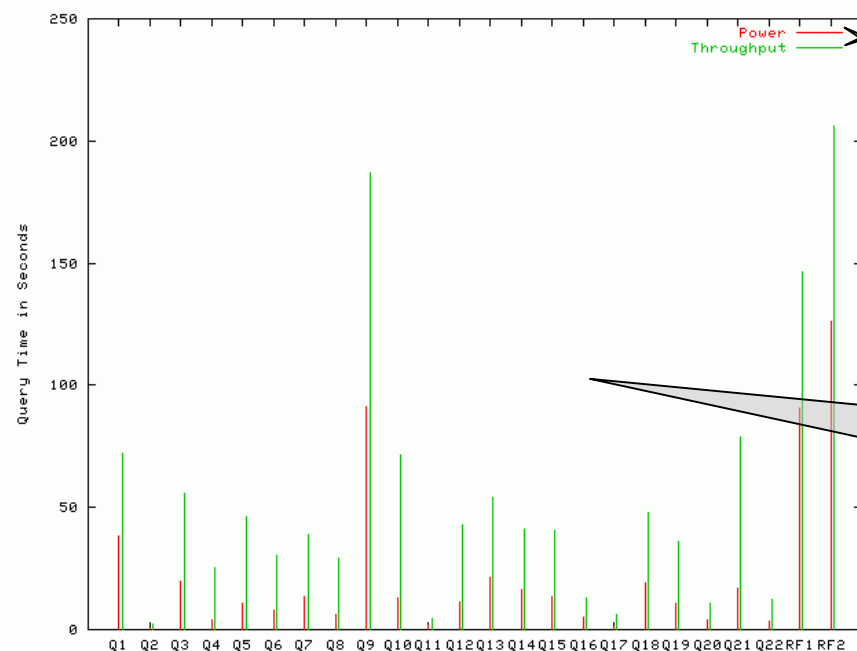
Configurations:

Software Version	Hardware Configuration	Run Parameters
Linux Kernel: 2.4.21-9.38AXsmp	4 CPU @ 2787.511 MHz	Database Scale Factor: 1
PostgreSQL: 8.0.0beta5	CPU model Intel(R) Xeon(TM) CPU 2.80GHz	Number of streams for throughput run: 4
sysstat: 4.0.7	2573188 kB Memory	shmmax: 2147483648
distribution: MIRACLE LINUX V3.0 (koumei)	Node: srpc2408.co.jp	database parameter:
procps: 2.0.13		Put postgres tmp on different drive: 0
Test Kit Version 1.5		Put WAL on different drive: 1

DBT-3 Metrics:

Composite	Query Processing Power	Throughput Numerical Quantity
275.16	336.99	224.68

Query Times



スケール
ファクター(SF)

ストリーム数
(同時接続数)

Throughput
ストリーム数=n
のベンチマーク

Power
ストリーム数=1
のベンチマーク

クエリーごとの
所要時間の
グラフ(秒)

結果の傾向 (ロード)

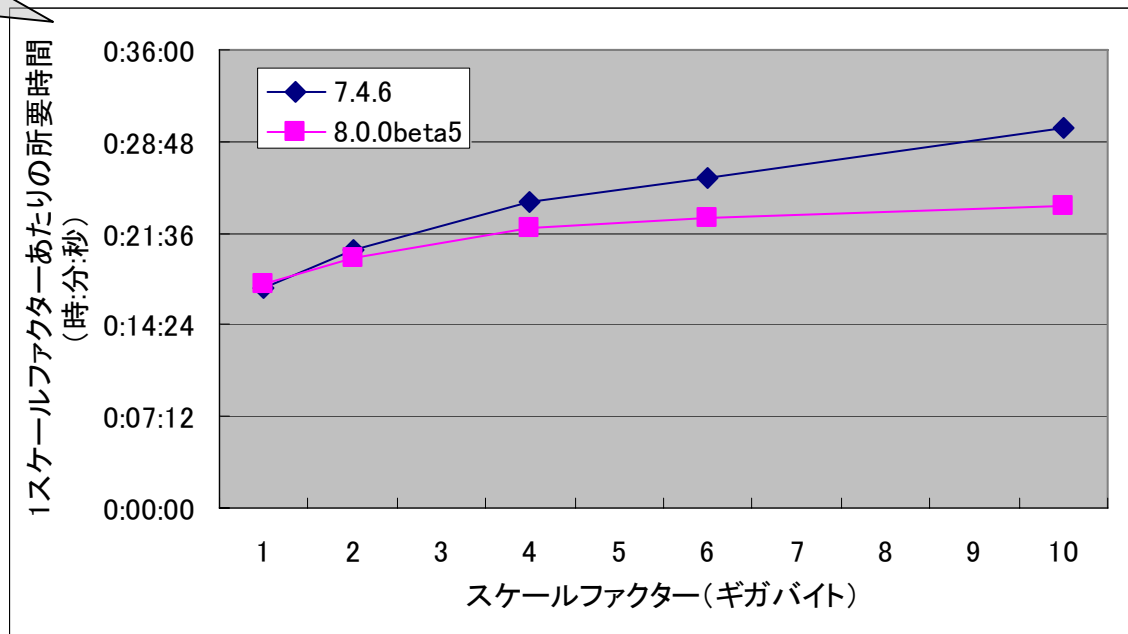
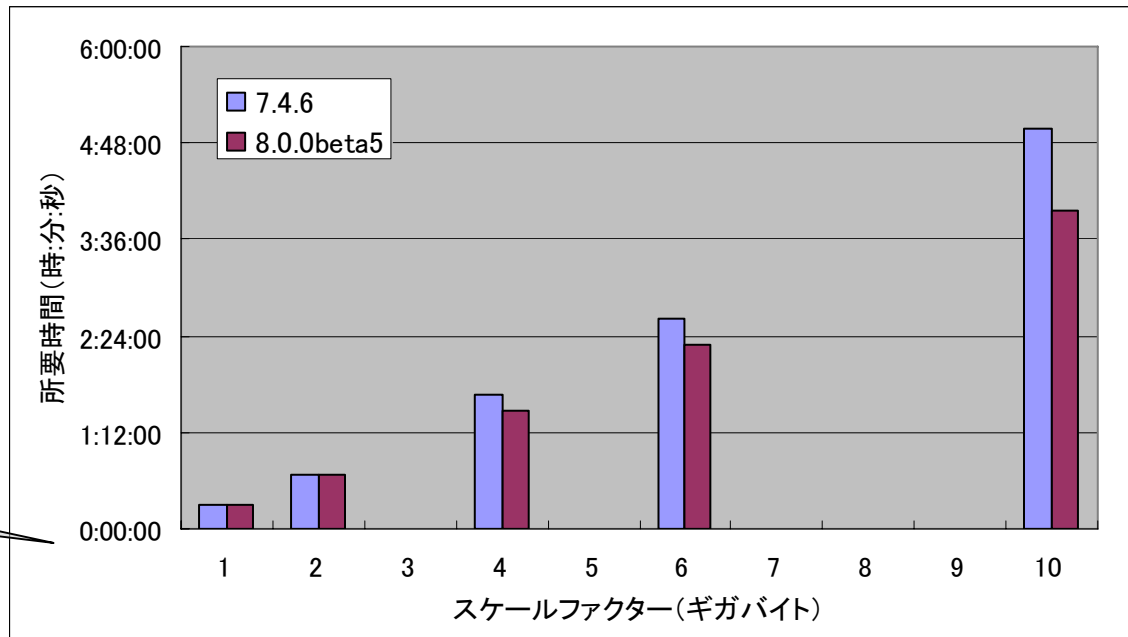
スケールファクターごとの
ロード時間のグラフ

1スケールファクターあたりの
スケールファクターごとの
ロード時間のグラフ

理想:
SF10の時間 = SF1の時間 × 10

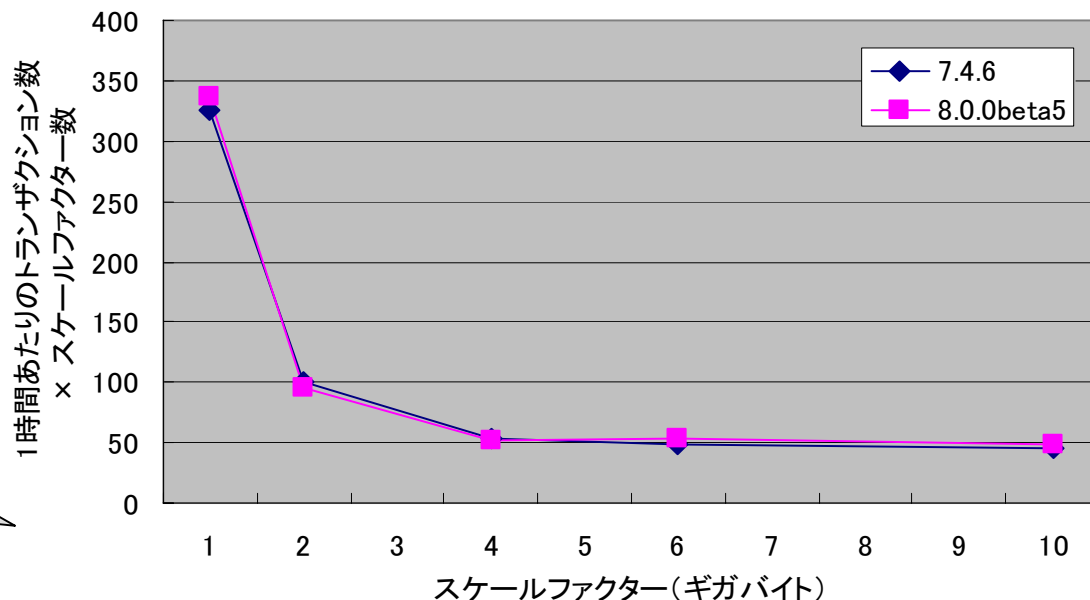
現実:
規模が大きいと管理コストも大きいので、それだけ時間も長くなる

結果の傾向:
8.0は7.4よりも、ロード時間の
増加率は改善されていた。



結果の傾向 (パワー)

1スケールファクターあたりの
スケールファクターごとの
パワーテスト結果のグラフ



理想:

SF10の、1SFあたりのトランザクション数 = SF1の、1SFあたりのトランザクション数 × 10

現実:

規模が大きいと管理コストも大きいので、それだけ性能が低下する。

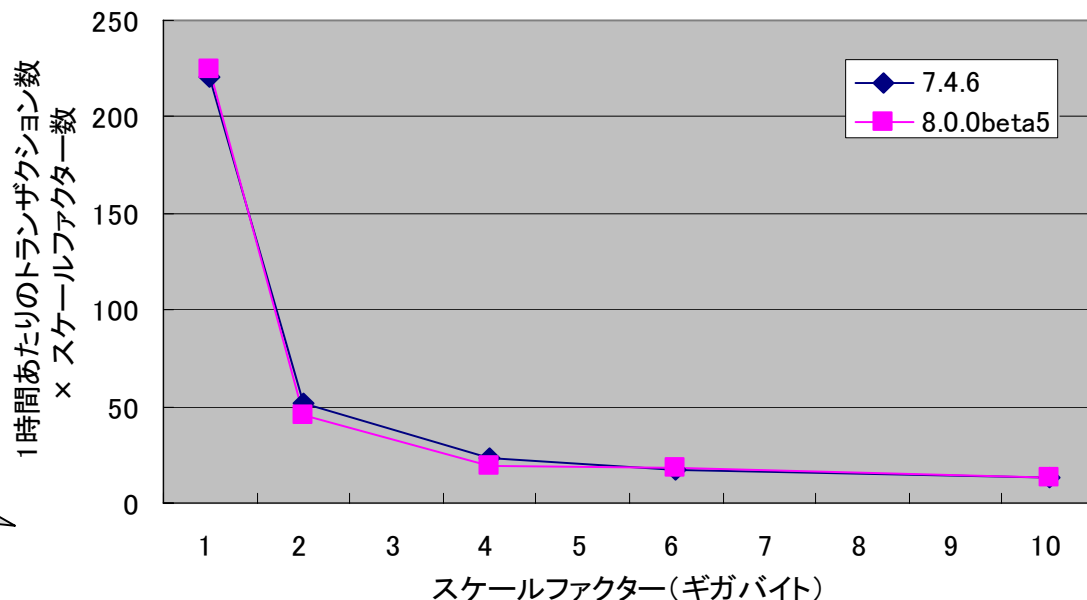
規模が小さいとキャッシュが効果的に使われるので、それだけ性能が向上する。

結果の傾向:

マシンのメモリ(2.5GB)までの規模なら、キャッシュの効果で性能が向上したと考えられる。
SF4以上はキャッシュの効果は無いが、管理コストによる性能低下は少なかった。

結果の傾向 (スループット)

1スケールファクターあたりの
スケールファクターごとの
スループットテスト結果のグラフ



理想と現実：
前ページのパワーと同じ。

結果の傾向：
マシンのメモリ(2.5GB)までの規模なら、キャッシュの効果で性能が向上したと考えられる。
SF4以上はキャッシュの効果は無いが、管理コストによる性能低下は若干あった。

HDDは150,000回転のSCSIディスクを使用した。
SF4以上は、ディスク性能の影響が大きいだろう。

今後の課題と謝辞

■ SF1～10を測定した

- HDDの使用量にして、およそ4GB～40GBになる。
- 規模が大きいと管理コストも大きくなるが、その増加率が一定なのが望ましい。作りが悪いと、ある時点から急に性能が劣化する場合がある。



- SF1～10では問題なかったが、より大きな規模での測定が望まれる
 - SF10は、およそ3日かかった。
 - TPC-Hでは、テラバイト級で測定が行われている。

■ 謝辞

- 本評価は、「独立行政法人情報処理推進機構 オープンソースソフトウェア活用基盤整備事業」に係る委託業務の一環として実施しました。
- 本評価では、OSDL-Jの協力を得ました。
- DBT-3の手順と評価結果の詳細は、IPAのWebサイトで公開しました。