

---

# PostgreSQL Update

## 9.0 のおさらいと、来るべき 9.1

---

オープンソースカンファレンス 2011 Tokyo/Spring  
2011-02-05 13:00-13:45 (61号館) 210室

特定非営利活動法人 日本 PostgreSQL ユーザ会  
高塚 遥



もともとの予定では・・・

- PostgreSQL バージョン 9.0 のおさらい
- 開発中の次バージョン 9.1 の紹介

だったのですが、

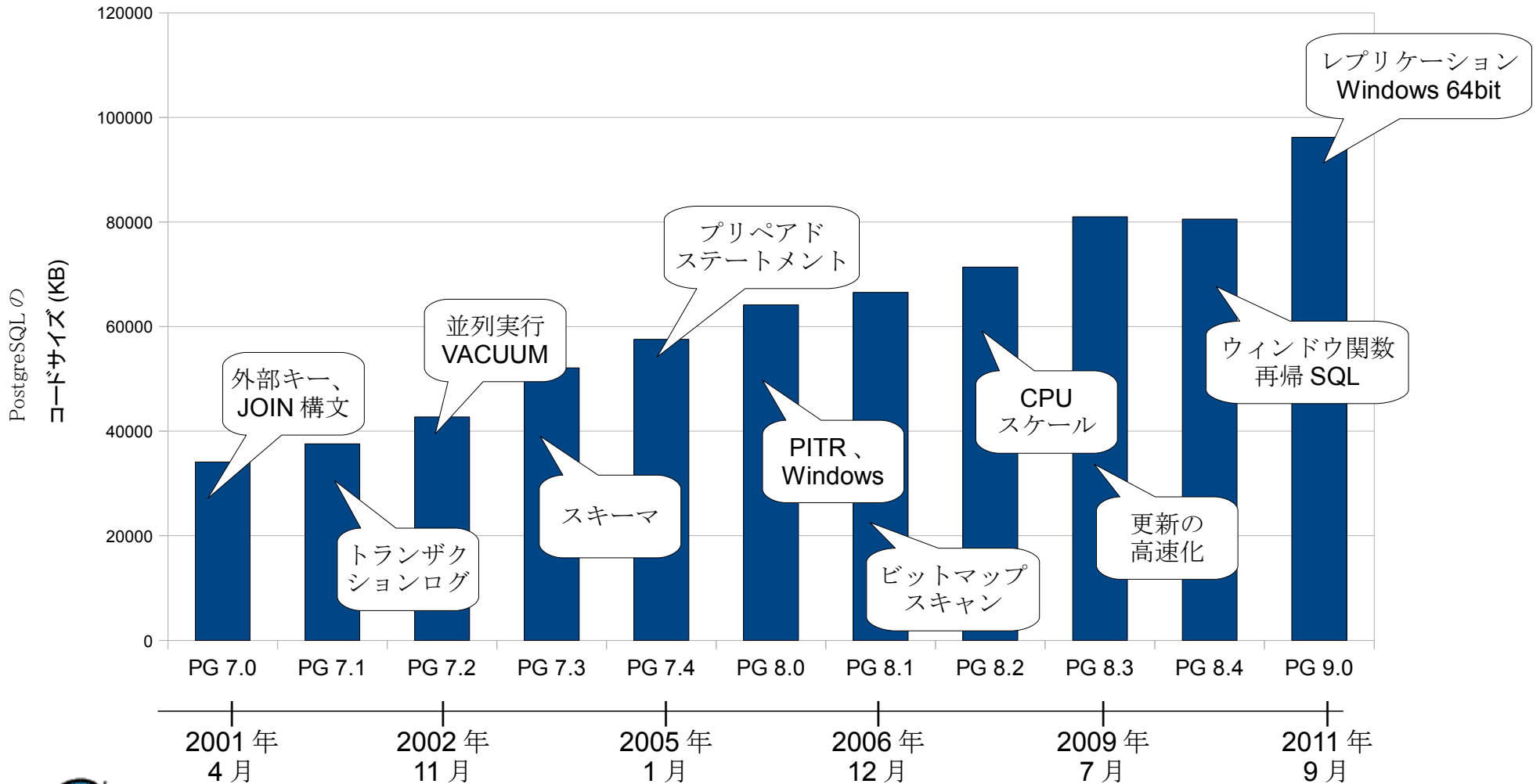
なんだか PostgreSQL、  
それも 9.0 をターゲットにした講演が多い

⇒ 9.1 の話メインでいきます



# PostgreSQL の歩み

PostgreSQL のコードサイズ



PostgreSQL

特定非営利活動法人  
日本 PostgreSQL ユーザ会

# PostgreSQL 9.1



例によって1年毎メジャーバージョンリリース

- 9.0 に入りそびれた大物機能
- 運用、開発を便利にする使いたくなる機能



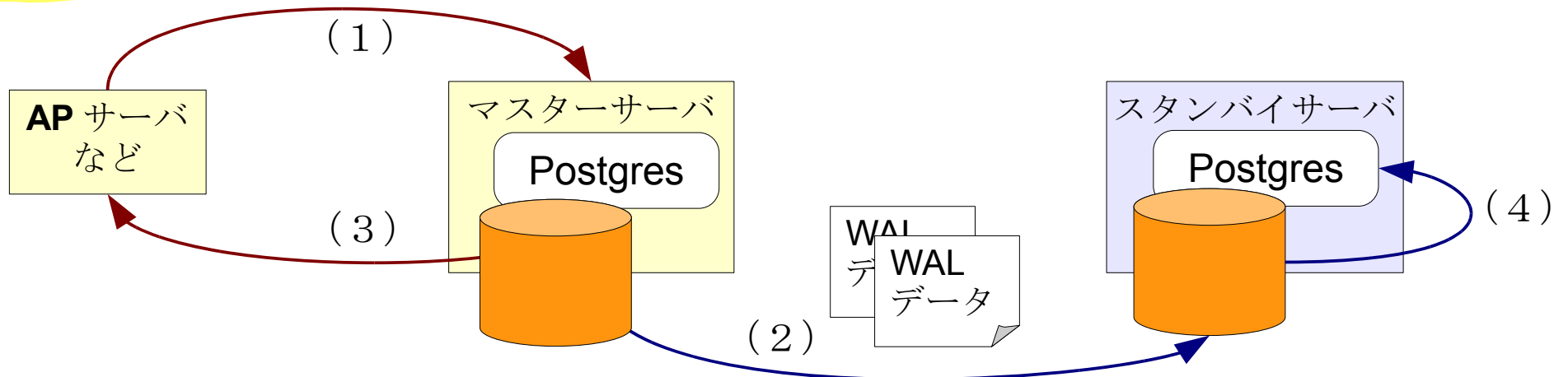
---

# レプリケーション関連

---



# 同期レプリケーション



- 9.0 では非同期＝障害直近でデータ損失あり
- WAL データを転送完了後にコミット完了
- 「スタンバイに問い合わせた内容がマスタと一致」を保障するわけではない



# pg\_stat\_replication ビュー

- マスターでレプリケーション状況を一覧できる
  - トランザクションのどこまで送信済み、どこまで適用済み、が見れる
    - 何時何分時点とかでは見れない
    - WAL ファイル位置「0/16D11F0」で
  - 9.0 では各スタンバイに問い合わせる必要があった

```
=# SELECT * FROM pg_stat_replication;
```



# pg\_basebackup コマンド

## pg\_ctl promote コマンド

- ベースバックアップがコマンドひとつで完了
  - `SELECT pg_start_backup('hoge');` ⇒ `tar` とか ⇒ `SELECT pg_stop_backup();` の手順が不要に
  - テーブルスペースにも対応
  - PostgreSQL の通常接続を使うので `scp` や `NFS` を用意しなくてよい
- スタンバイのレプリケーションを終了させて通常起動 (=マスタ昇格) させるコマンド
  - 以前はトリガーファイルを置いていた



# 拡張



**PostgreSQL**

特定非営利活動法人  
日本 PostgreSQL ユーザ会

# SQL/MED FOREIGN TABLE

- 外部データをテーブルとして読み込む仕組み
  - contrib/file\_fdw は CSV 形式などのファイルをテーブルとして読み込む外部データラッパーの実装例
  - さまざまな外部データ連携のモジュールが作れる
- 9.1 では参照だけ、更新は未対応



# CREATE EXTENSION

- 拡張モジュールをインストールする命令
- 所定の作法を守って書かれた拡張モジュールを、SQLコマンドで管理できる
  - ビルドと `.dll` / `.so` のインストールは予め必要です
- 拡張モジュールのバージョンアップに対応

```
=# CREATE EXTENSION hstore;  
=# DROP EXTENSION hstore;
```



# 性能



# UNLOGGED テーブル

- トランザクションログ（WAL）に記録しないテーブルを作れる
  - 書き込みが速い
  - 複数接続共有の一時テーブルとして
  - クラッシュしたらデータ消えてしまうかも
  - レプリケーションされない

```
=# CREATE UNLOGGED TABLE tb101 ( ... )
```



# ORDER BY pushdown

- プランナが改善
- Merge Append プランタイプ
- テーブルパーティショニングされたテーブルでソートを行うとき各々の子テーブルにあるインデックスが使われる

```
=> SELECT id, v1 FROM t_partitioned_oya  
ORDER BY v1 LIMIT 10;
```



# 賢い CLUSTER コマンド

- テーブル統計情報を見て、最適な再編成の方式を自動的に選択
  - 新方式 **Seq Scan + Sort** は速い、一時ファイルサイズを大きく使う

テーブル状態	処理方式
断片化が少ない	Index Full Scan (従来方式)
断片化が多い	Seq Scan + Sort (新方式)



---

# 機能

---



# KNN GiST インデックス

- k-NN(k Nearest Neighbor) 検索：空間上に指定された地点に近接するオブジェクトを、空間データベースの中から k 個求める
- GiST インデックスを距離が近いもののトップリストを出す検索に使える
  - POINT 型、 Geometry 型など

```
SELECT coordinates,  
       (coordinates <-> '5.0,5.0'::point) AS dist  
FROM spots ORDER BY dist ASC LIMIT 10;
```



# 述語ロック対応 SERIALIZABLE

- Serializable Snapshot Isolation
- 9.0 までの SERIALIZABLE は、不完全
  - マニュアルにも但し書きがついている

```
SELECT SUM(val) FROM tbl WHERE grp = 'a';
```

```
SELECT SUM(val) FROM tbl WHERE grp = 'b';
```

- 旧 SERIALIZABLE ⇒ (9.1) REPEATABLE READ
- 新 SERIALIZABLE ⇒ (9.1) SERIALIZABLE



# contrib/sepgsql

- データベースオブジェクトに強制アクセス制御
  - アクセス可否のチェックを **SE-Linux** と統合
    - 各 SQL について **SE Linux** に問い合わせるので  
**SELinux Enabled** が大前提
    - **SELinux** のポリシーとして設定を行う
  - **SE-Linux** 同様のラベルベース強制アクセス制御

```
SECURITY LABEL FOR selinux
ON TABLE mytable
IS 'system_u:object_r:sepgsql_table_t:s0';
```



# さまざまな文字列関数が追加

- `format('%s and %s', 'abc', '123')`
  - `printf` タイプのフォーマッタ
  - SQL リテラル、SQL 識別子のエスケープに対応
- `concat(a, b, c, ...)`、`concat_ws()`
  - `||` 演算子と同じだけど、某DBとの互換性的に
- `reverse()`
- `left()`、`right()`
  - 右から何文字、左から何文字



# 更新を行える WITH 句

- Writable CTE (Common Table Expressions)
- 処理行データを返す「RETURNING」と組み合わせて WITH クエリーで INSERT、UPDATE、DELETE が記述できる
  - DELETE した内容を別テーブルに INSERT

```
WITH t (id, v) AS (DELETE FROM t_a RETURNING *)  
INSERT INTO t_b SELECT * FROM t;
```

- 再帰 + Writable CTE の応用例は？



# ビューに対するトリガー

- 更新可能ビューを作るのに **RULE** でなくトリガーも使えるようになった

```
CREATE TRIGGER mytrig
  INSTEAD OF UPDATE ON myview
  FOR EACH ROW
  EXECUTE PROCEDURE myupdate ();
```



# カラム単位ロケール指定

- 検索毎、カラム毎で **COLLATE**（言語を考慮した文字比較）指定ができる
  - 8.3 まで データベースクラスター単位
  - 8.4 データベース単位
- **glibc** の日本語ロケールは例によって使えない

```
CREATE COLLATION c_en (LOCALE = 'en_US.UTF-8');
```

```
SELECT t FROM t_coll ORDER BY t COLLATE c_en;
```

```
SELECT t FROM t_coll ORDER BY t;
```

ABC  
XYZ  
abc  
xyz

abc  
ABC  
xyz  
XYZ



# レンジデータ型

- 範囲を表現するデータ型
  - 重なり検出する演算子 **&&**
  - 9.0 で導入された排他制約と組み合わせ
    - これまでは **POINT** 型くらいしか使える対象がなかった

データ型	要素データ型
int4range	int
int8range	bigint
numrange	numeric
tsrange	timestamp without timezone
tstzrange	timestamp with timezone
daterange	date

```
SELECT range(11.1, 22.2) && range(20.0, 30.0);
```

```
ALTER TABLE reservation  
ADD EXCLUDE USING gist (during WITH &&);
```



# contrib/pg\_tram の拡張

- LIKE 検索で中間一致にもインデックス検索
  - もともとは類似度一致の演算子やインデックス機能を提供するもの
  - 9.1 で LIKE、ILIKE に対応

```
CREATE INDEX ON docs  
  USING gin (doc gin_trgm_ops);
```

gist も可

```
SELECT * FROM docs WHERE doc LIKE '%foo%';
```



# 以上。そして注意点

- **Committed** 状態でないものも紹介しています  
(2011/3/3 時点)
  - つまり 9.1 に載るとも限らない
    - 先送りを繰り返してついに搭載という機能も多い
- 今回割愛した内容
  - PL/Tcl、PL/Python、PL/Perl、PL/pgsql 関連
  - ほか、いろいろ
    - パラレル pg\_dump、MERGE コマンド



# 新バージョン情報は

- コミットフェスト管理 Web サイト  
<https://commitfest.postgresql.org/>
- アルファ版が数ヶ月おきに  
<http://www.postgresql.org/developer/alpha>
- HEAD のマニュアルとか
- JPUG が各種イベントで講演など
  - 本講演で紹介した 9.1 新機能スライドは 2/25 PostgreSQL Conference 2011 Tokyo のライトニングトーク（板垣 貴裕氏）原稿がベース



# JPUG の Web サイト

- 日本 PostgreSQL ユーザ会 (JPUG)

<http://postgresql.jp>

- Let's postgres

( PostgreSQL の情報ポータルサイト )

<http://lets.postgresql.jp>



**PostgreSQL**

特定非営利活動法人  
日本 PostgreSQL ユーザ会