Amazon RDS for PostgreSQL 検証報告

AWSプロダクトシリーズ|よくわかるクラウドデータベース
2014/01/17

SRA OSS, Inc. 日本支社
石井　達夫
正野　裕大

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

性能

- スペックの異なる RDS インスタンスでベンチマークを測定

技術的トピック

- 素の PostgreSQL と違う点

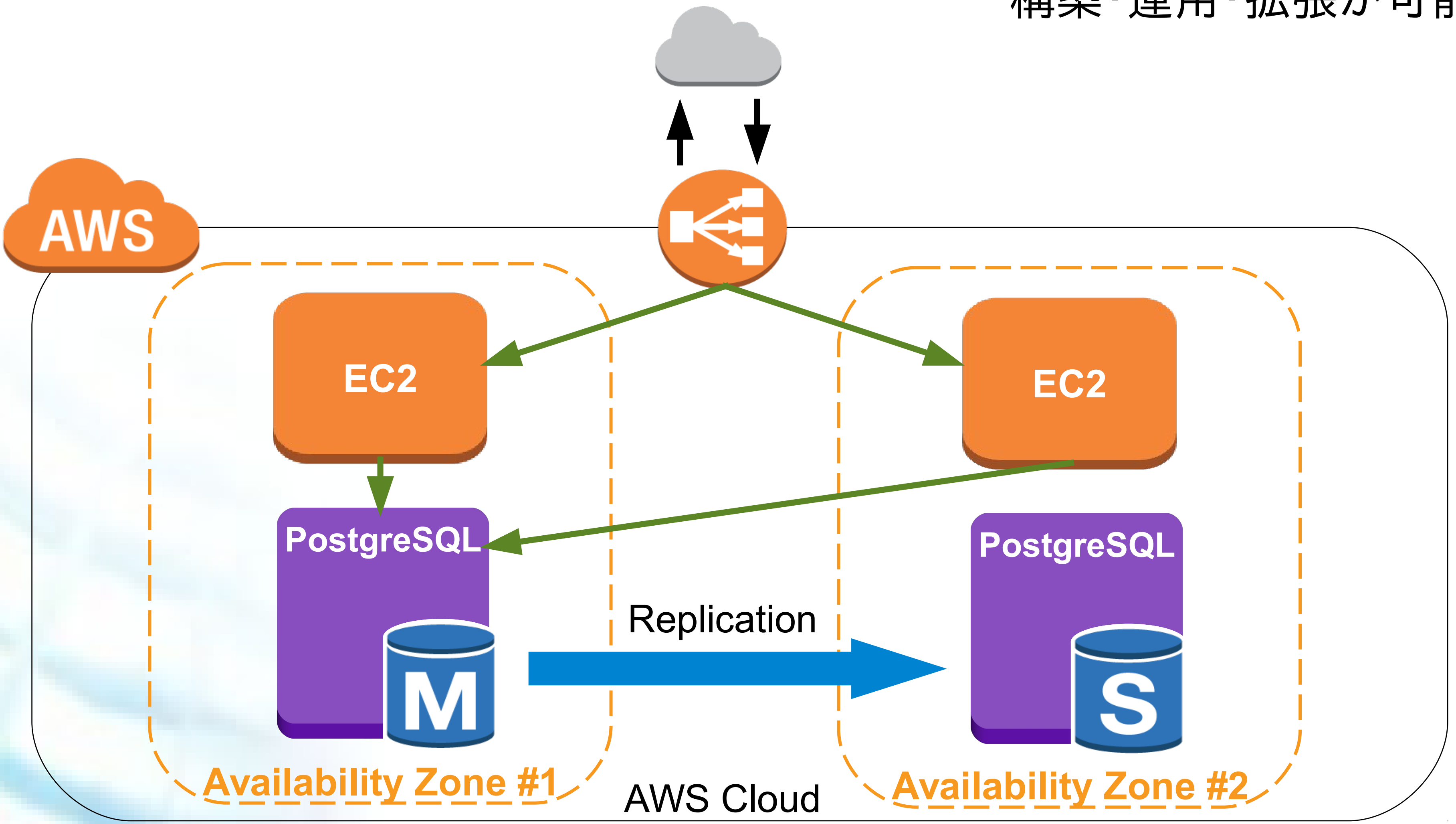
付録

- Amazon RDS と PostgreSQL の概要
- RDS for PostgreSQL を使ってみた
  RDS インスタンスの作成 / Multi-AZ / バックアップ / スケールアップ / モニタリング

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

Amazon RDS とは?

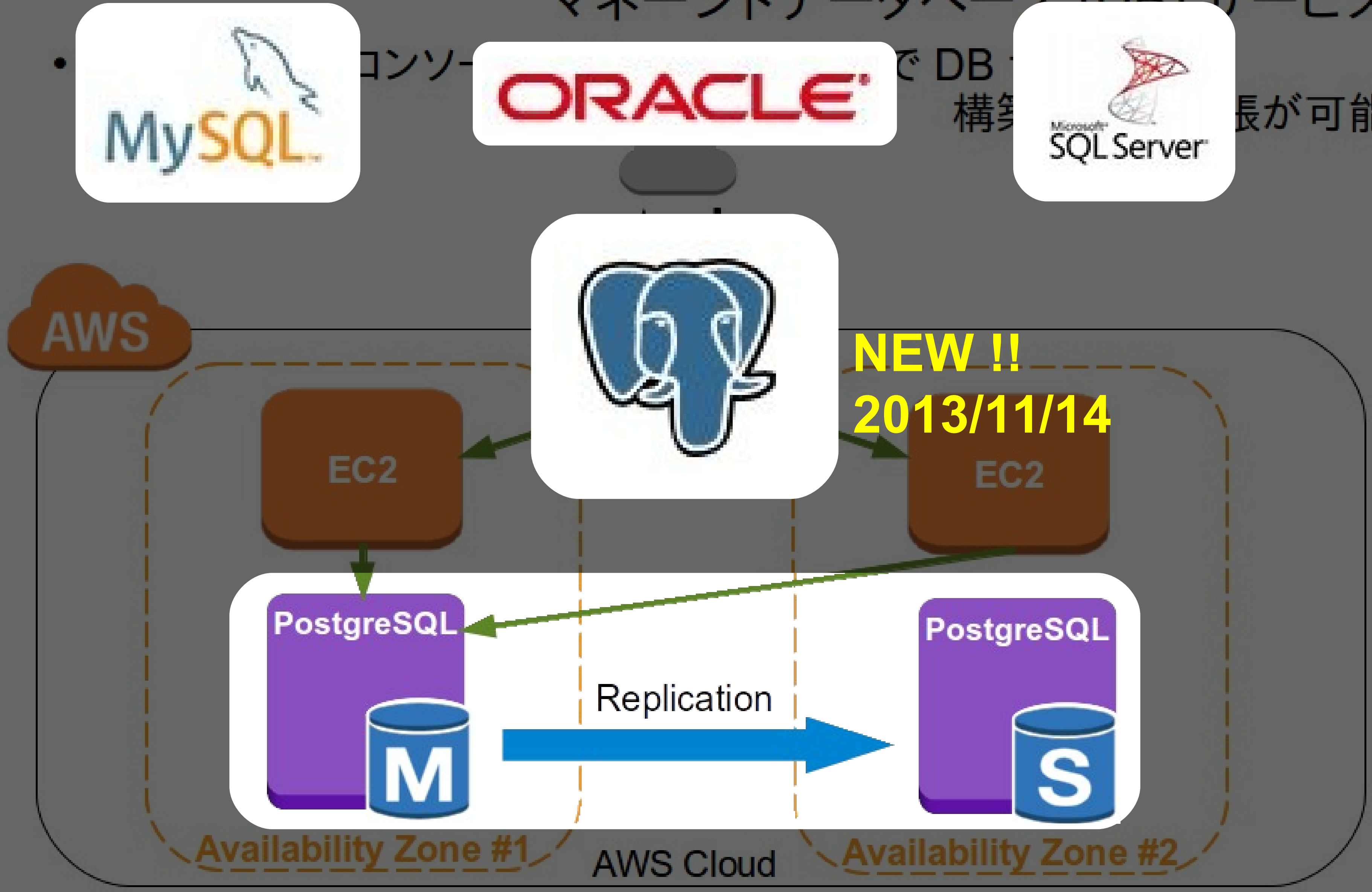
- Amazon Web Services が提供する マネージドデータベース (DB) サービス
  - 簡単に Web コンソールを通してクラウド上で DB サーバの構築・運用・拡張が可能



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

Amazon RDS とは?

- Amazon Web Services が提供する
  マネージドデータベース (DB) サービス
  - コンソー… で DB … 構築 … 張が可能



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

性能

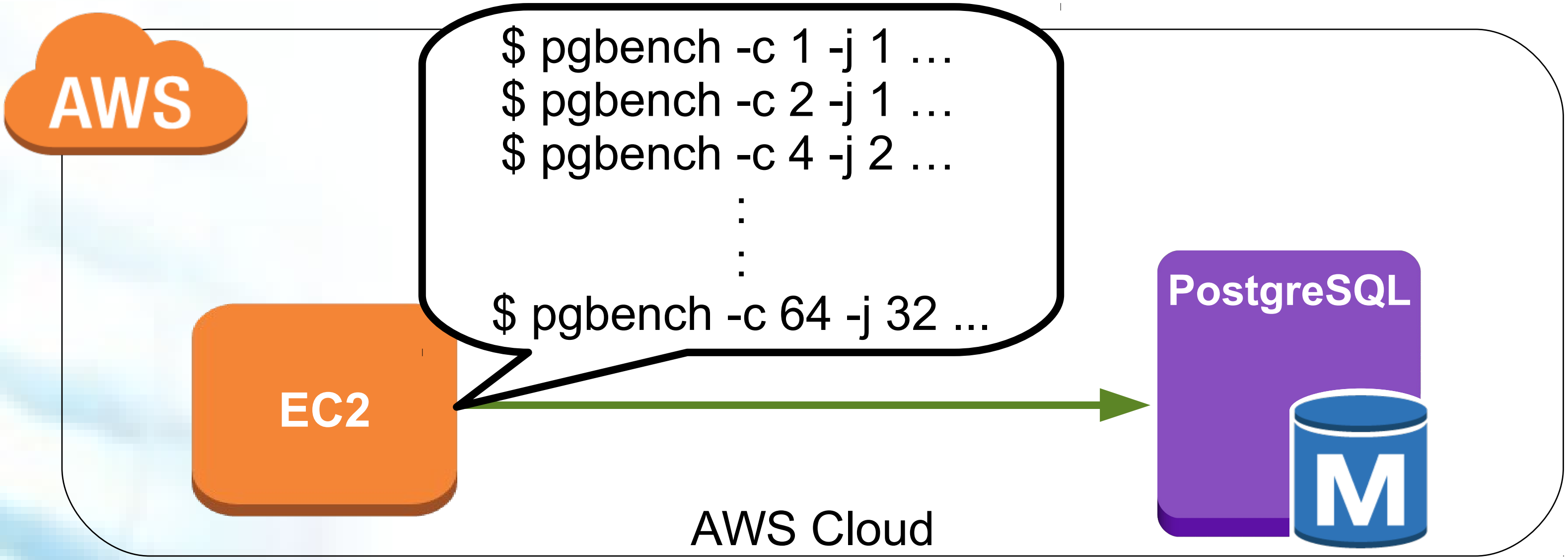
- PostgreSQL に付属する標準ベンチマークツール (pgbench) でベンチマークを実施

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

- 3 つの RDS インスタンスでベンチマークを測定

DB インスタンスクラス	CPU	メモリ
ハイメモリエクストララージ	3.25 ECU × 2 仮想コア	17.1 GB
ハイメモリダブルエクストララージ	3.25 ECU × 4 仮想コ	34 GB
ハイメモリクラスタエイトエクストララージ	88 ECU（32 ハイパースレッドコア）	244 GB

- EC2 インスタンスをクライアントとして pgbench を実行



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

ベンチマーク測定方法

- RDS インスタンスの情報 | DB サーバ（pgbench 対象）

DB インスタンスクラス	ストレージサイズ	IOPS	パラメータ	ベンチ用 DB
ハイメモリエクストララージ	100GB	1000	デフォルト	testdb
ハイメモリダブルエクストララージ				
ハイメモリクラスタエイトエクストララージ				

- EC2 インスタンス | クライアント（pgbench 実行側）

インスタンスクラス	ハイメモリクラスタエイトエクストララージ (cr1.8xlarge)
スペック	32 コア（仮想） / 88 ECU / RAM 244GB
OS	Red Hat Enterprise Linux 6.4
PostgresSQL 9.3.1	PostgreSQL コミュニティのリポジトリを使用

pgbench 実行側が
ボトルネックにならないように
高性能のインスタンスを
使用

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

ベンチマーク測定方法

- pgbench 用 DB testdb の初期化
  - スケールファクタ 1000（データ行 1 億件 / DB サイズ 15GB）

```
[ec2-user@ec2 ~]$ pgbench -i -s 1000 -U <DB USER> -h <RDS> testdb
```

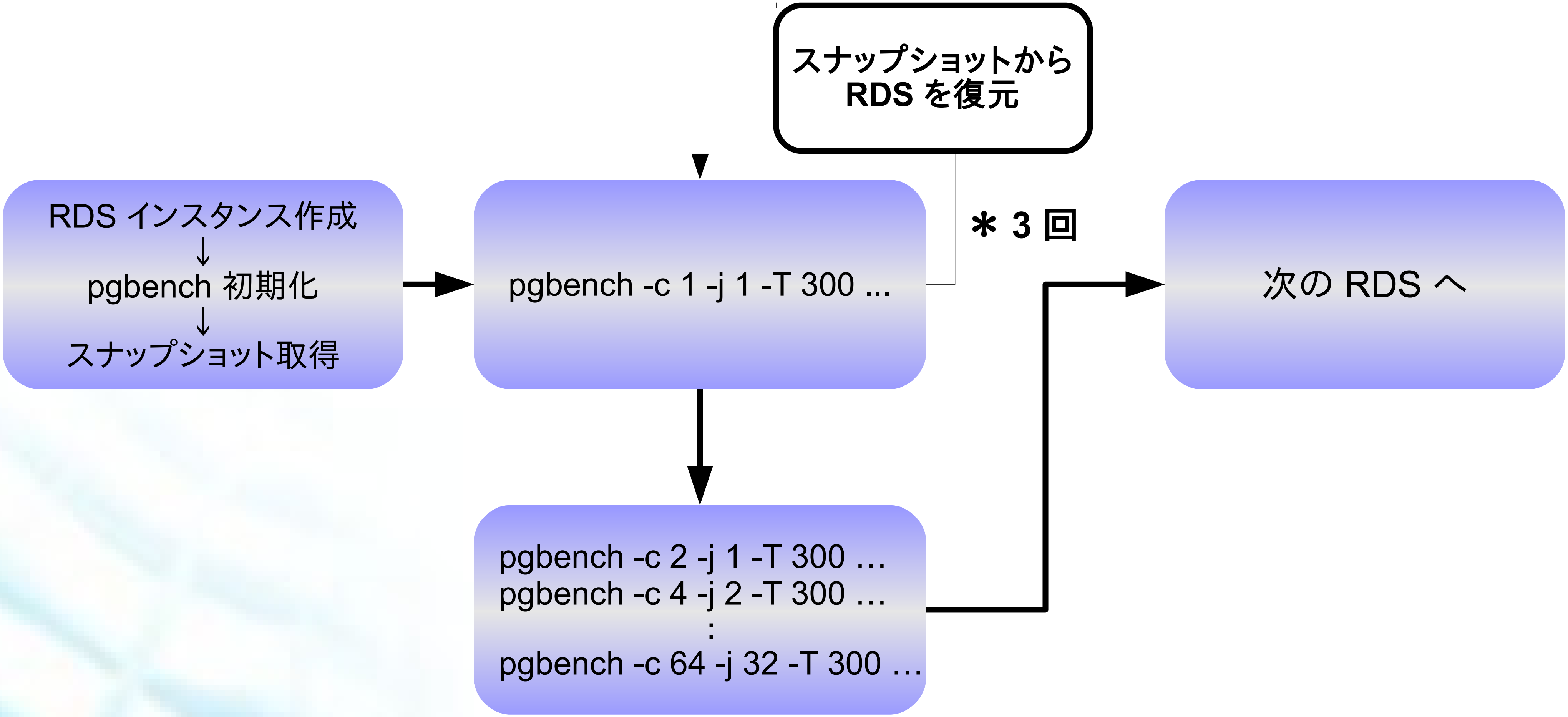
- pgbench の実行
  - 標準 (TCP-B) の pgbench
  - 同時接続数 1, 2, 4, 8, 16, 32, 64 を 3 回ずつ実行
  - 実行時間は 300 秒
  - 実行すると TPS（1 秒間に実行したトランザクション数）が出力される

```
[ec2-user@ec2 ~]$ pgbench -c 1   -j 1   -T 300 -U <DB USER> -h <RDS> testdb   * 3 回
[ec2-user@ec2 ~]$ pgbench -c 2   -j 1   -T 300 -U <DB USER> -h <RDS> testdb   * 3 回
[ec2-user@ec2 ~]$ pgbench -c 4   -j 2   -T 300 -U <DB USER> -h <RDS> testdb   * 3 回
[ec2-user@ec2 ~]$ pgbench -c 8   -j 4   -T 300 -U <DB USER> -h <RDS> testdb   * 3 回
[ec2-user@ec2 ~]$ pgbench -c 16 -j 8   -T 300 -U <DB USER> -h <RDS> testdb   * 3 回
[ec2-user@ec2 ~]$ pgbench -c 32 -j 16 -T 300 -U <DB USER> -h <RDS> testdb   * 3 回
```

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

ベンチマーク測定方法

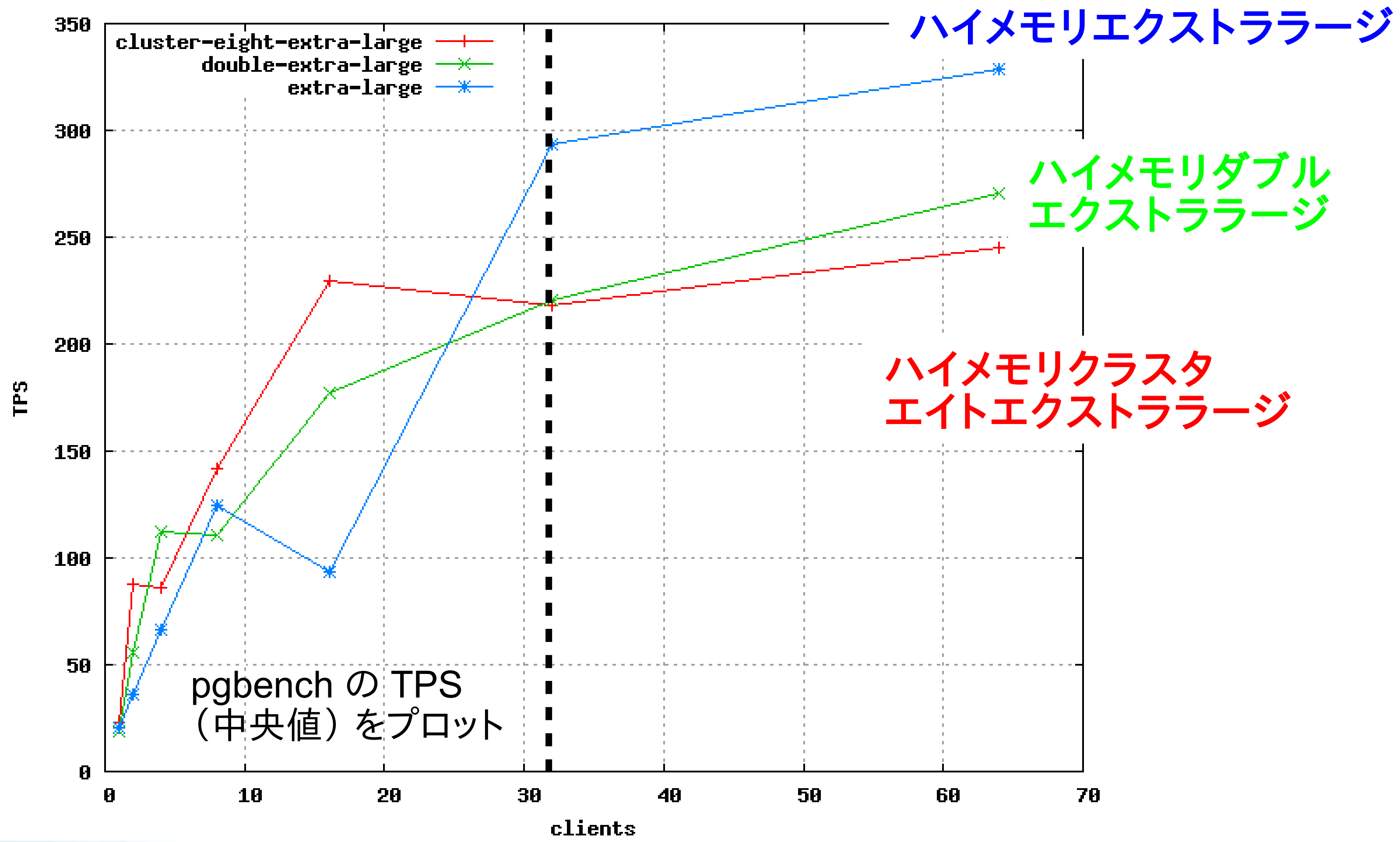
- 測定の流れ
  - 各インスタンスで以下を実行



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

ベンチマーク結果 | TPS

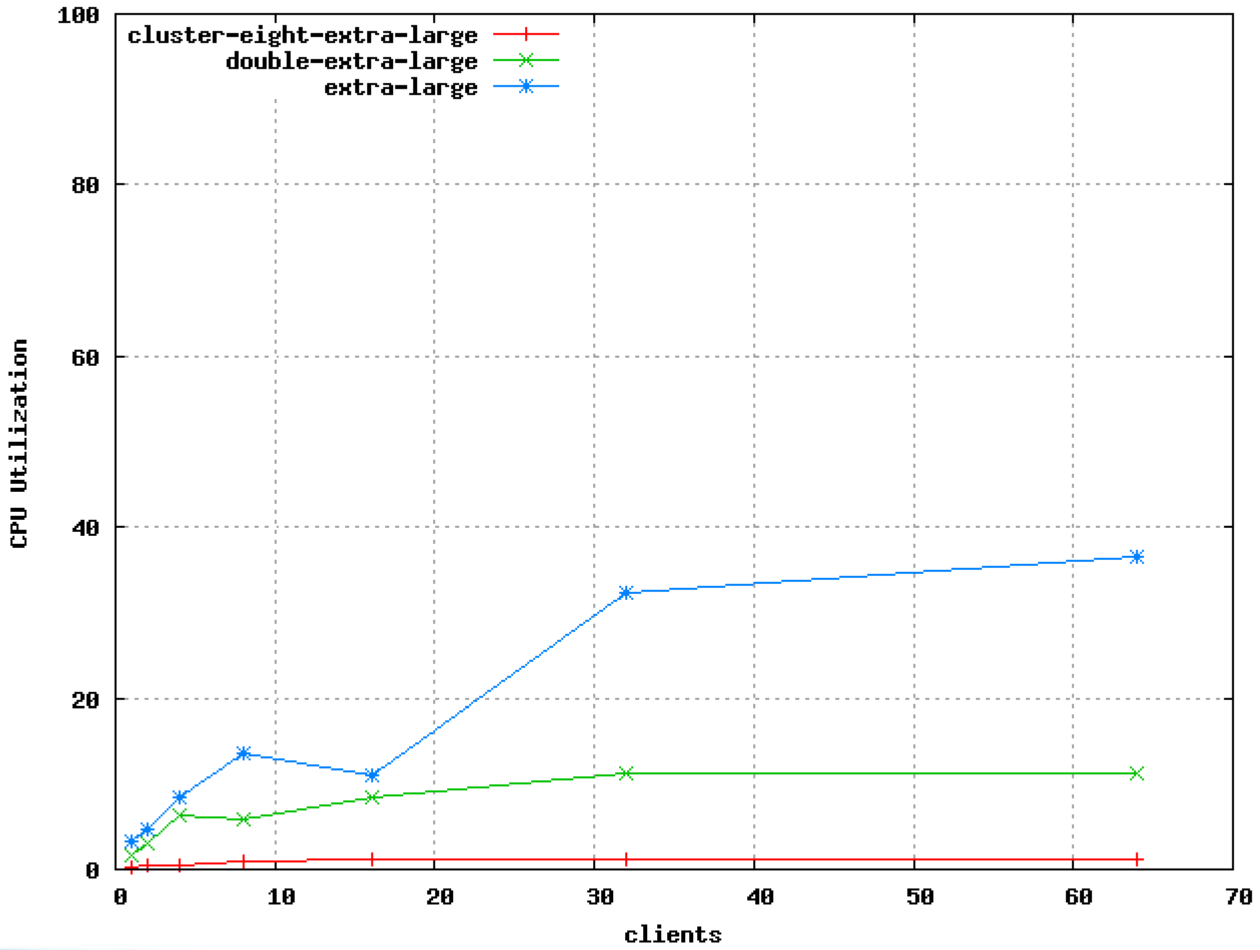
- 同時接続数が増えるにつれて性能 (TPS) が上昇している
- 同時接続数がコア数を越えても性能が上昇し続けた
- 最も性能の高い DB インスタンスクラス が一番 TPS が低い



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

ベンチマーク結果 | CPU

- 性能の高い DB インスタンスほど CPU 利用率が低い
- おそらく I/O がボトルネックになっている



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

ベンチマーク再測定

- 以下の条件で pgbench -c 64 -j 32 -T 300 … を実行
  - RDS インスタンスの情報 | DB サーバ（pgbench 対象）

DB インスタンスクラス	ストレージサイズ	IOPS	パラメータ	ベンチ用 DB
ハイメモリエクストララージ	**600GB**	**6000**	デフォルト	testdb

IOPS を 6 倍に

  - EC2 インスタンス | クライアント（pgbench 実行側）

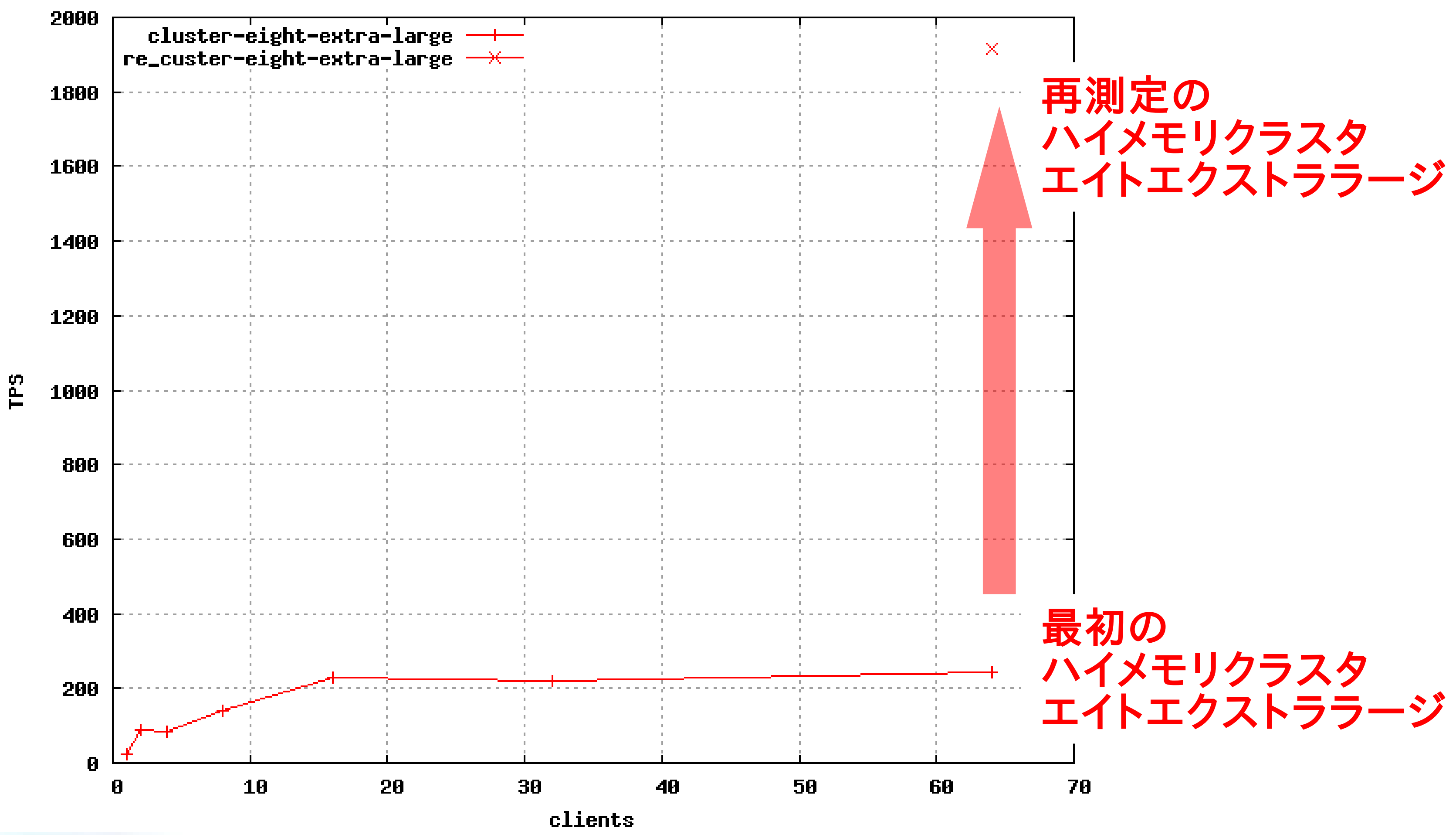
pgbench 実行側が
ボトルネックにならない程度に
高性能のインスタンスを
使用
（少しスペックダウン）

インスタンスクラス	M1 エクストララージ (m1.xlarge)
スペック	4 コア（仮想） / 8 ECU / RAM 15GB
OS	Red Hat Enterprise Linux 6.4
PostgresSQL 9.3.1	PostgreSQL コミュニティのリポジトリを使用

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

ベンチマーク結果(再)

- IOPS を 6 倍にしたら TPS がおよそ 8 倍に
  - Write IOPS は 5000 前後 に
  - CPU 使用率は低いまま



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

技術的トピック

- 素の PostgreSQL と違う点

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

素の PostgreSQL と違う点 1/3

- クライアント認証の設定を行う手段が提供されていない
  - PostgreSQL では pg_hba.conf ファイルを編集してクライアント認証を設定する
    - 接続方式、DB ユーザ、接続 DB、接続元、認証方式 を制御できる
  - RDS では AWS のセキュリティグループでしか設定できない
    - 接続元のみ
- DB のスーパーユーザになれない
  - 以下はできる（AWS が PostgreSQL に手を加えている?）
  - テーブルスペースの作成
  - 他のユーザのセッションの確認と切断
  - AWS が用意した拡張モジュールのロード

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

素の PostgreSQL と違う点 2/3

- PostgreSQL の拡張モジュールは以下が使える

btree_gin	btree_gist	chkpass	citext	cube
dblink	dict_int	dict_xsyn	earthdistance	fuzzystrmatch
hstore	intagg	intarray	isn	ltree
pgcrypto	pgrowlocks	pg_trgm	plperl	plpgsql
pltcl	postgis	postgis_tiger_ geocoder	postgis_topology	sslinfo
tablefunc	tsearch2	unaccent	uuid-ossp	

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

素の PostgreSQL と違う点 2/3

- PostgreSQL の拡張モジュールは以下が使える

btree_gin	btree_gist		citext	cube
dblink	dict_int	dict_xsyn	earthdistance	fuzzystrmatch
hstore	intagg	intarray		ltree
pgcrypto		pg_trgm	plperl	plpgsql
pltcl	postgis		ostgis_topology	sslinfo
tablefunc	tsearch2	unaccent	uuid-ossp	

大文字/小文字の区別がない文字列型を提供 (citext)

ユーザ定義関数の手続き言語（plperl, pltcl もある） (plpgsql)

データベースの暗号化と複合 (pgcrypto)

地理空間情報を扱う (postgis)

接続中クライアントの SSL 証明書の情報を提供 (sslinfo)

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

素の PostgreSQL と違う点 3/3

- デフォルトの設定では日本語を DB に格納するには不向き
  - 日本語のデータベースでは DB のロケールを C（ロケールなし）にしておく必要がある
    - ロケール：文字列照合順序、通貨単位、小数点記号などを規定
- postgresql.conf パラメータは RDS 側で設定されたもの、固定されたものが運用上は特に問題なし
  - 共有メモリのサイズはおよそ RDS インスタンスのメモリの 1/4
  - WAL バッファ（トランザクションログ書き込み時に利用されるバッファ）は共有メモリの 1/32（PostgreSQL のデフォルト）
- ストリーミングレプリケーションは行えない
- デフォルトで SSL 通信が有効
  - MySQL および SQL Server、PostgreSQL のみ対応
  - SSL 通信はそこそこ重い処理
  - クライアント側で無効にできる

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

Amazon RDS と PostgreSQL の概要

- Amazon RDS for PostgreSQL の特徴
- RDBMS における PostgreSQL の位置づけ

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

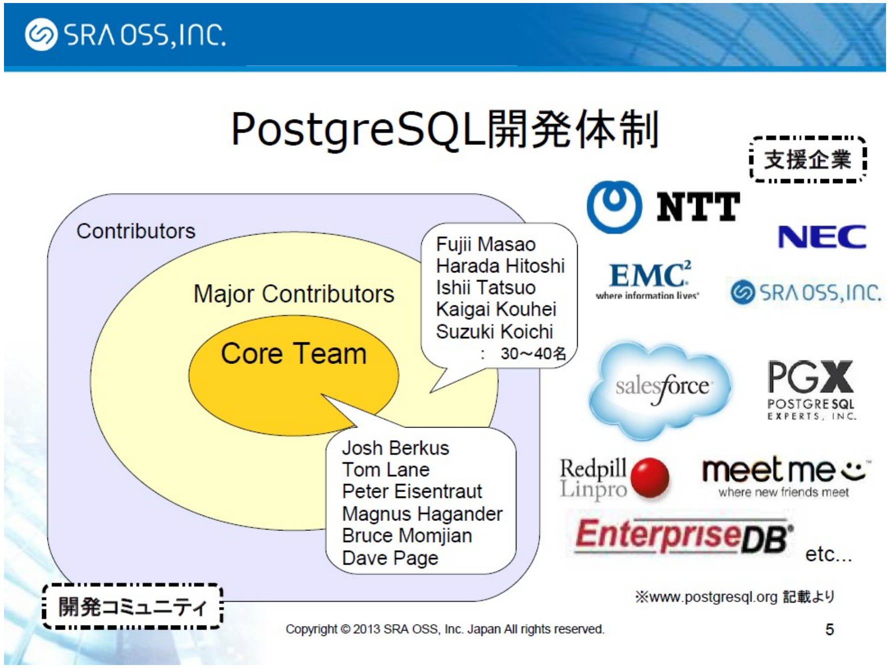
Amazon RDS for PostgreSQL の特徴

- DB サーバのマネージドサービス
  - 手軽にセットアップ･運用･スケーリング
    - パラーメータの最適化
    - CPU 使用率･メモリ･I/O･接続数などのモニタリング
    - PostgreSQL の自動アップデート
- HA 構成で可用性アップ
  - Multi-AZ で同期物理レプリケーションと自動フェイルオーバー
- バックアップ
  - 自動バックアップ
  - 任意のスナップショット

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

RDBMS における PostgreSQL の位置づけ

- 代表的なオープンソース RDBMS
  - PostgreSQL 6.0 (1996 ～) から 15 年以上の歴史
- 商用 DB と比較しても機能･性能面で引けをとらない
  - 業務基幹システムの商用 DB 製品の置き換え
  - Blog、SNS、ゲーム等の新しいオンラインサービス
- コミュニティベースで開発
  - 一つのオーナー企業、
    オーナー個人を持たない
  - 1 バージョン / 1 年 を
    10年以上



オープンソースデータベース PostgreSQL 最新動向のご紹介
(SRA OSS, Inc. 2013-06-27)

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

付録

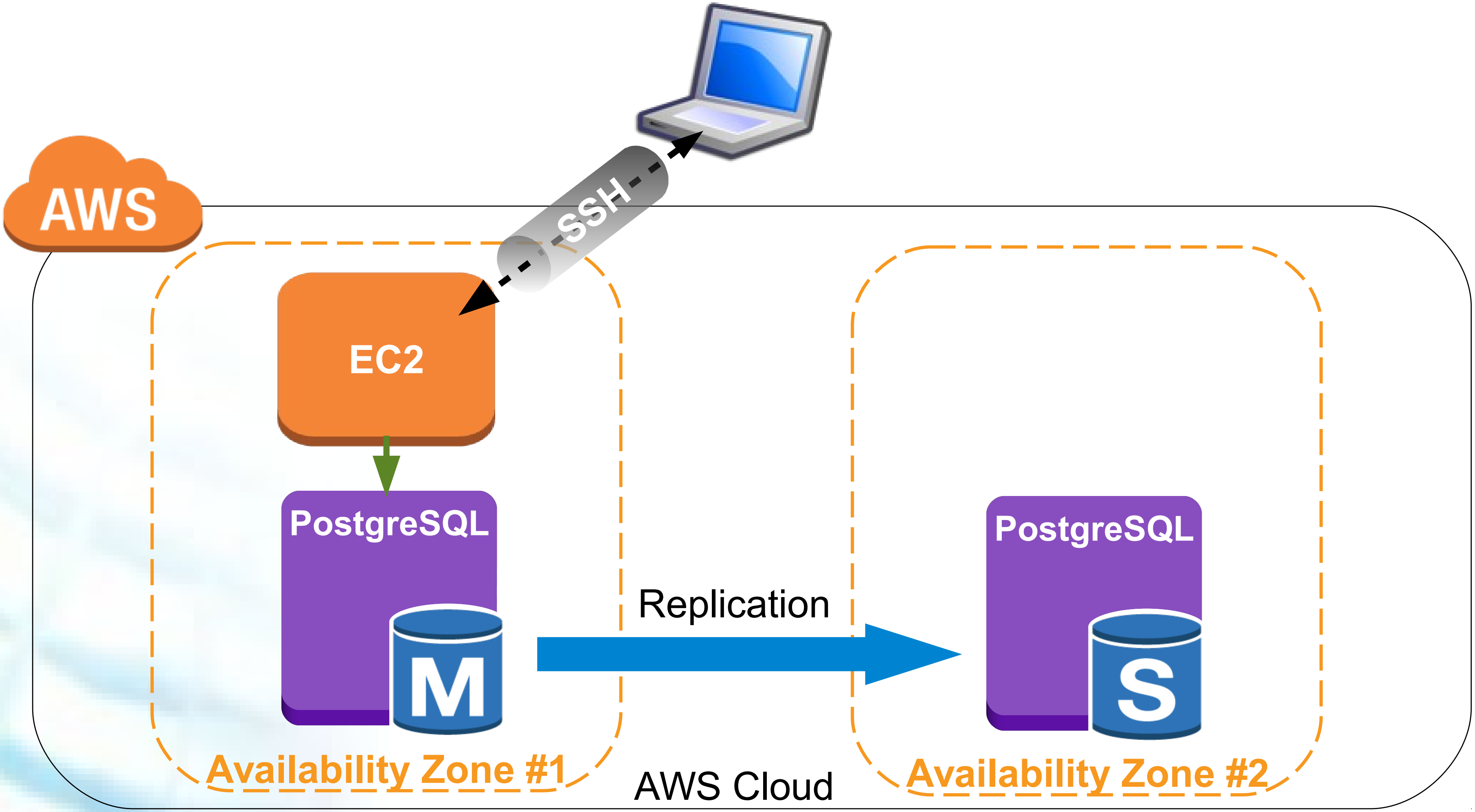
ユーザビリティ

- RDS インスタンスの作成
- バックアップ
- モニタリング
- Multi-AZ
- スケールアップ

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

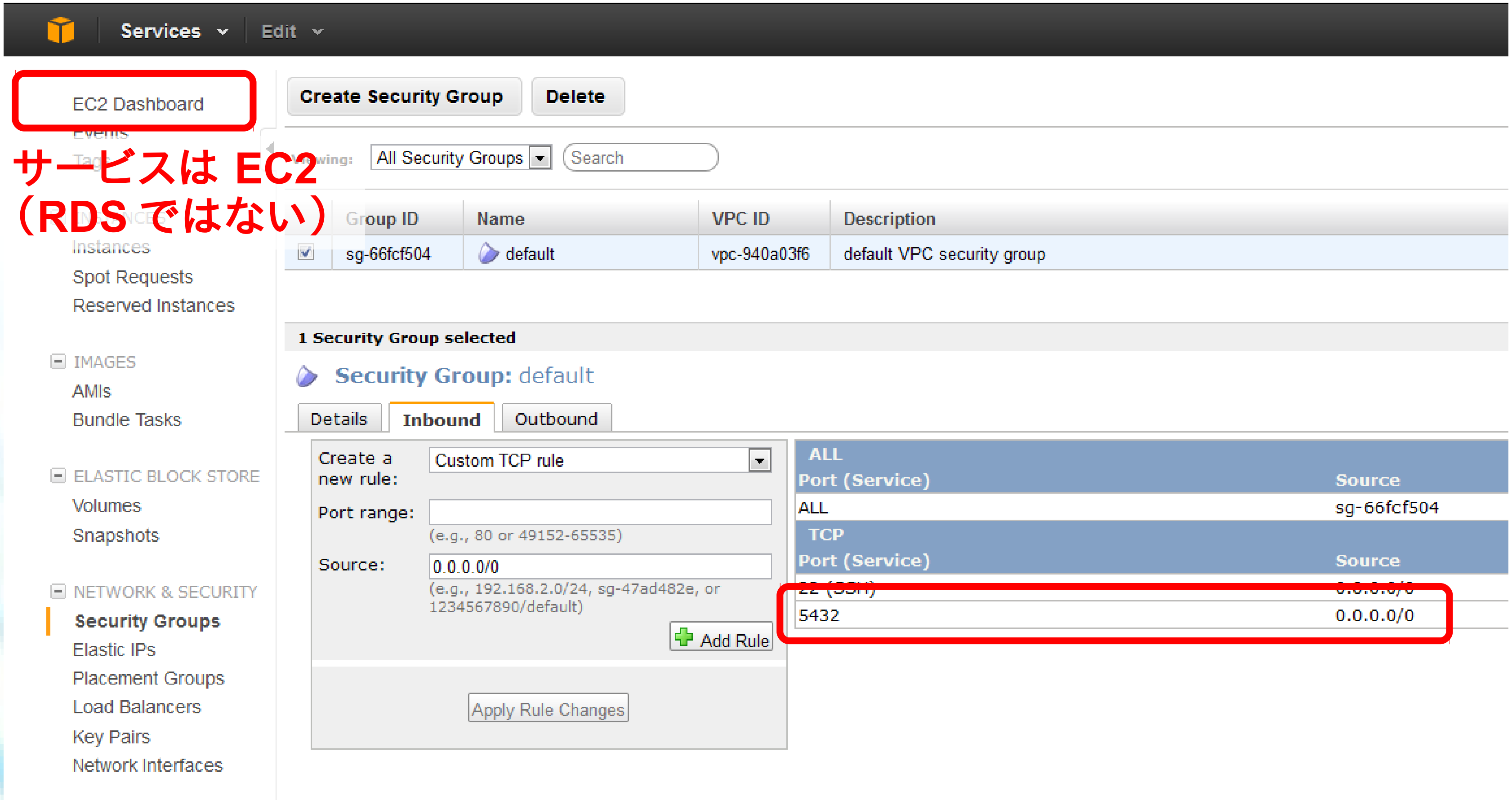
作成するサーバ構成

- クライアント用 EC2 インスタンスは作製済み
- PostgreSQL の RDS インスタンス | HA 構成 を作ってみる



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

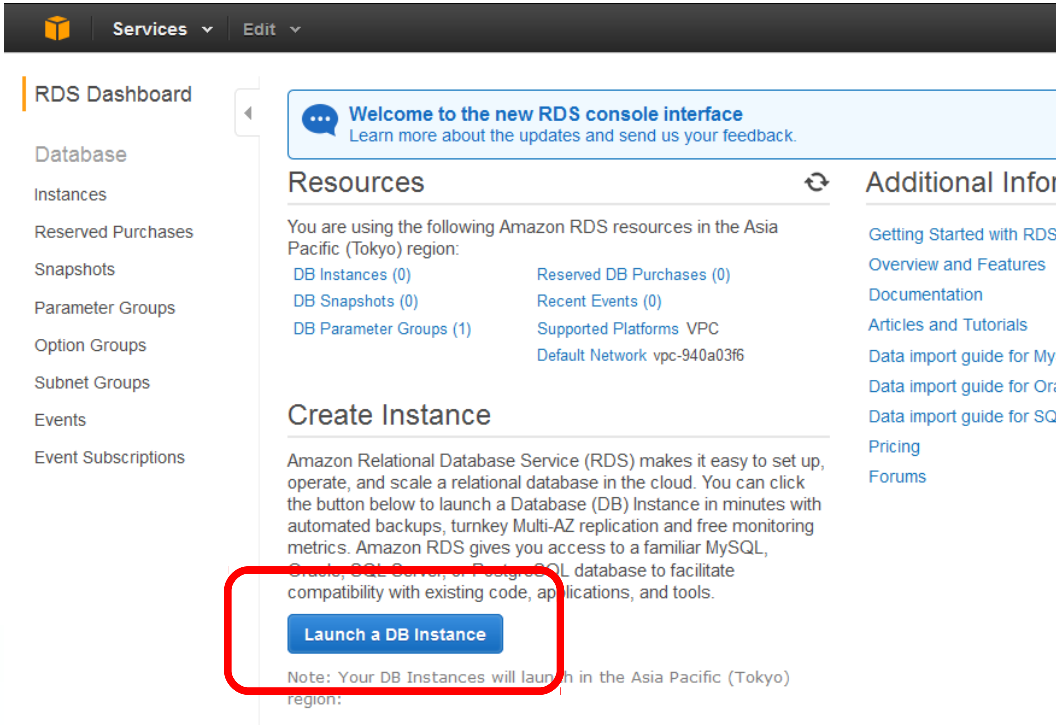
- 事前に**セキュリティグループ**を設定しておく
  - AWS の仮想ファイアウォール
  - インスタンスごとに任意のセキュリティグループを設定できる
  - 今回はデフォルトの設定 + 5432 ポートを開けておく
    - 5432 :PostgreSQL のデフォルトポート



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

RDS インスタンスの作成

- Web コンソールから　Launch a DB Instance !



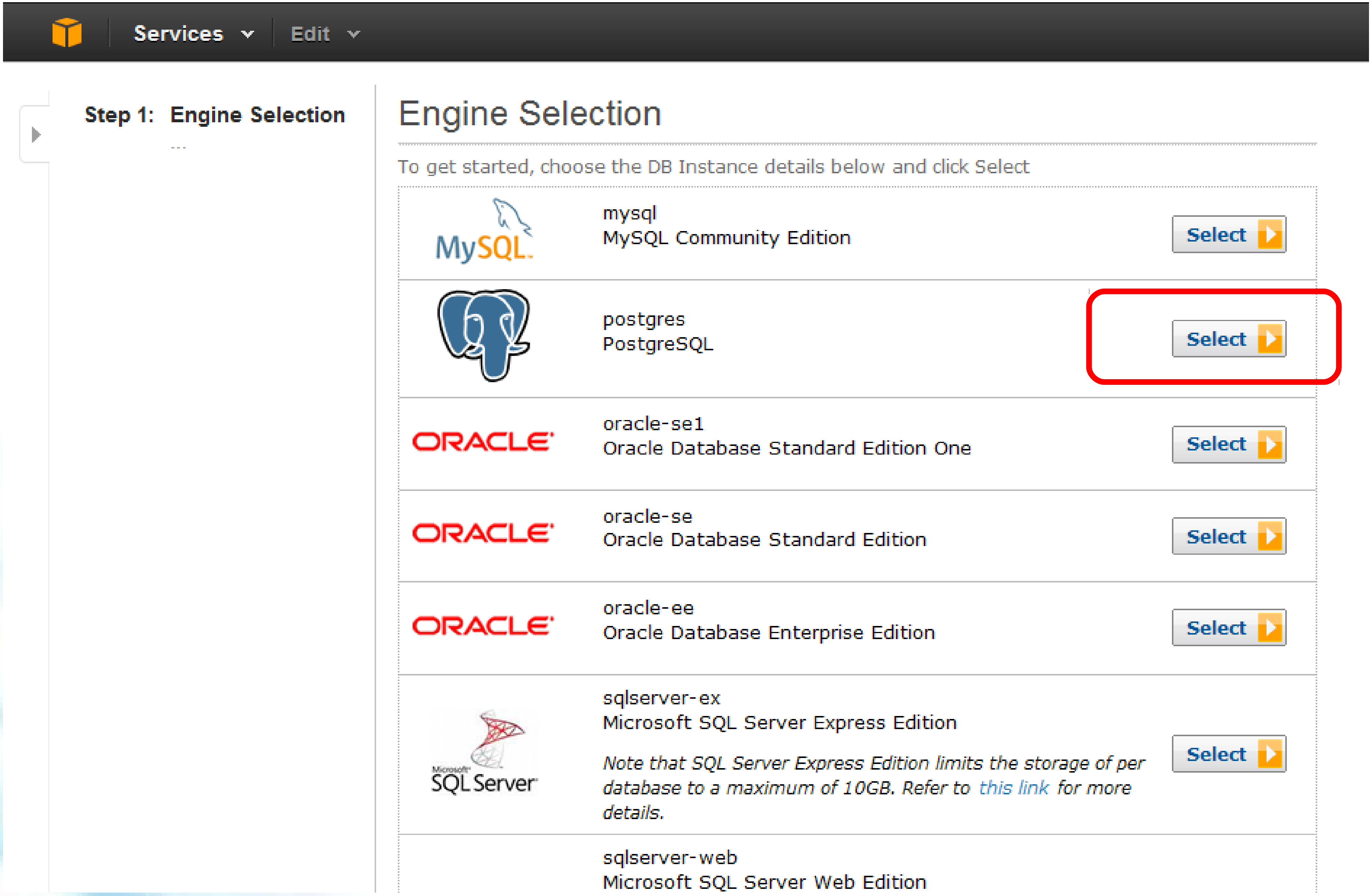
- ウィザードに従って項目を埋めていくだけ

DB エンジン → インスタンス スペック → セキュリティ ネットワーク構成 → DB パラメータ → バックアップ メンテナンス

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

DB エンジンの指定

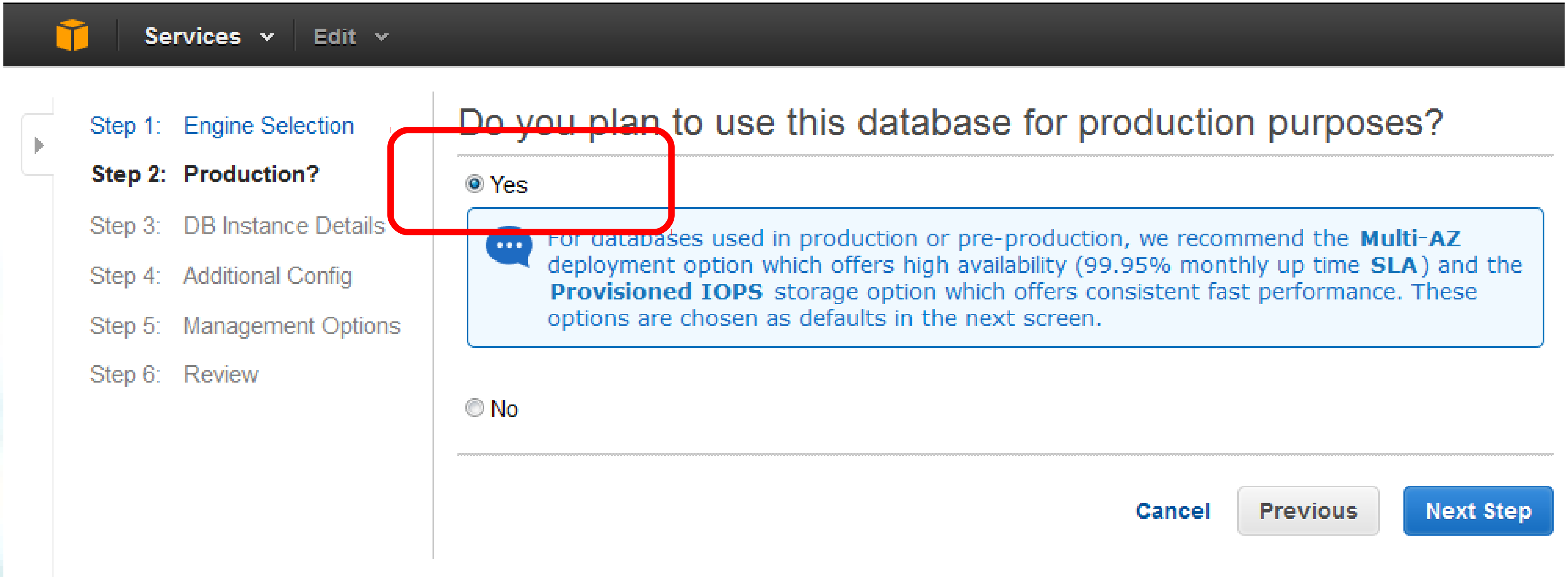
- PostgreSQL を選択



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

プロダクション用途か？

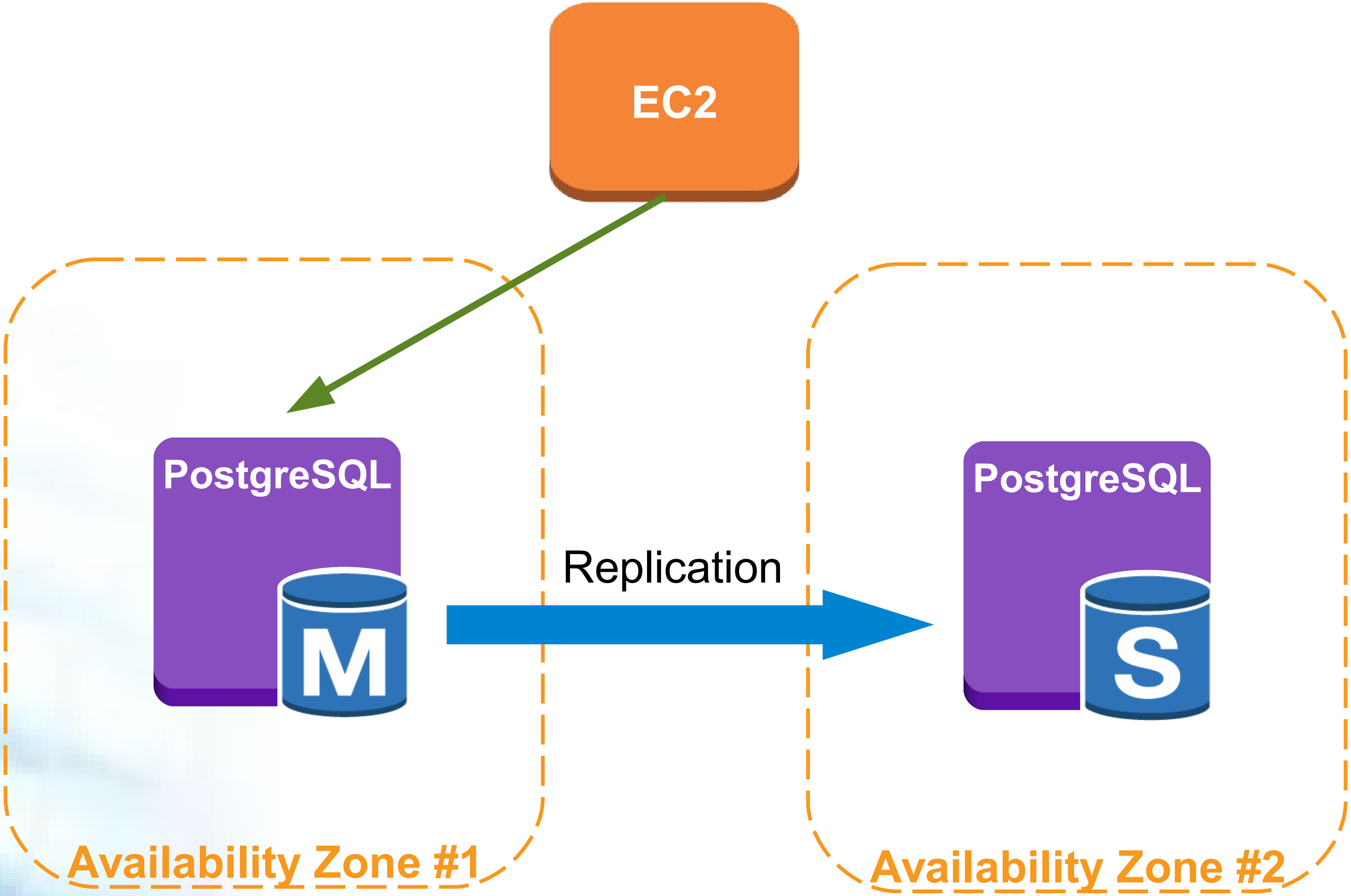
- Yes　Multi-AZ・Provisioned IOPS を利用する
- No　Multi-AZ・Provisioned IOPS を利用しない



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

Multi-AZ とは？

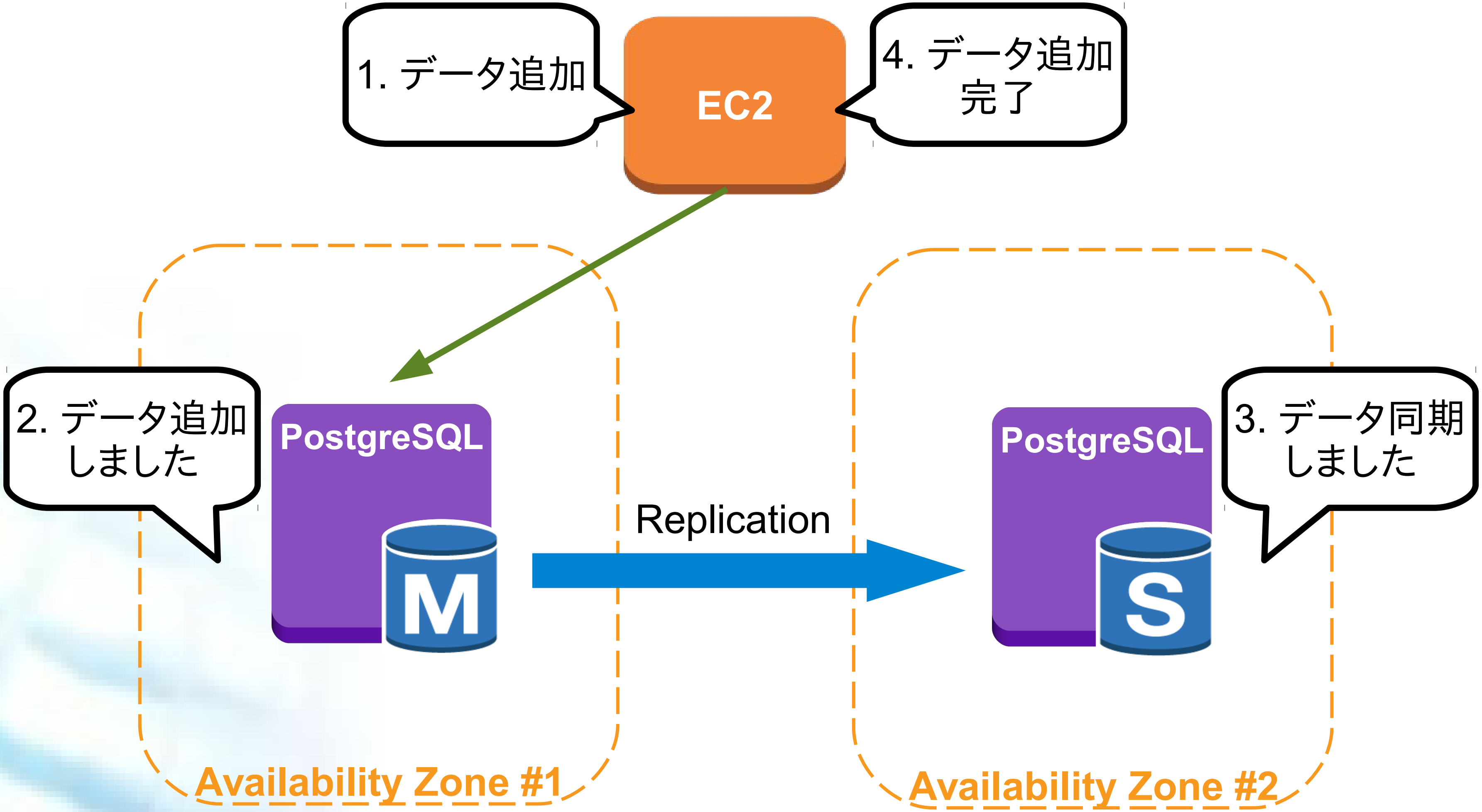
- 耐障害性・可用性を向上させる **AWS RDS の標準機能**
- 異なるアベイラリティゾーン（≒データセンタ）に待機系サーバを作成
- 同期物理レプリケーションと自動フェイルオーバ



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

Multi-AZ とは?

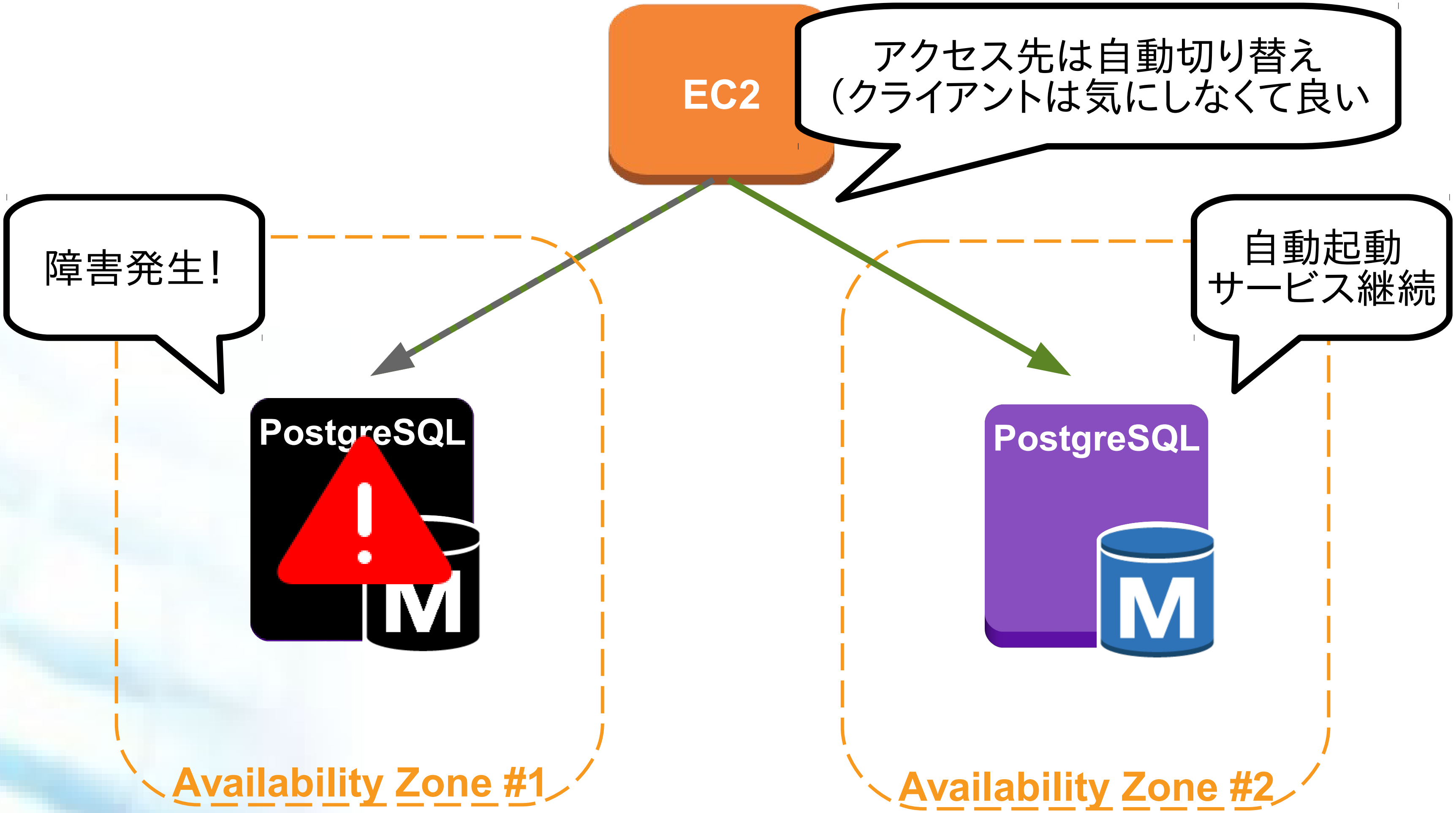
- 耐障害性・可用性を向上させる **AWS RDS の標準機能**
- 異なるアベイラリティゾーン(≒データセンタ)に待機系サーバを作成
- **同期物理レプリケーション**と自動フェイルオーバ



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

Multi-AZ とは?

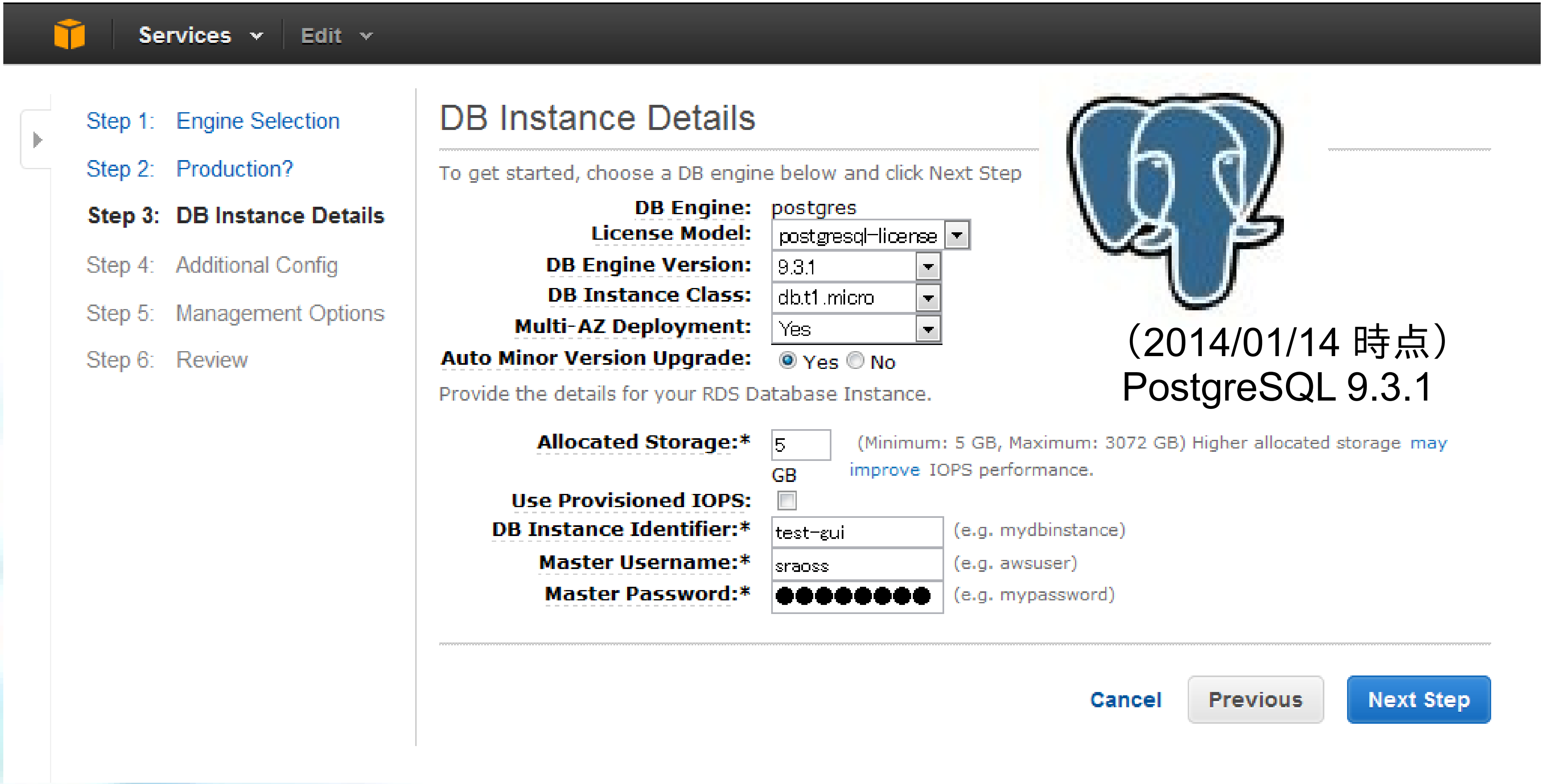
- 耐障害性・可用性を向上させる **AWS RDS の標準機能**
- 異なるアベイラリティゾーン(≒データセンタ)に待機系サーバを作成
- 同期物理レプリケーションと**自動フェイルオーバ**



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

DB インスタンスの詳細設定 1/2

- RDS インスタンス作成時に必要な項目を設定



（2014/01/14 時点）
PostgreSQL 9.3.1

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

DB Instance Class の種類

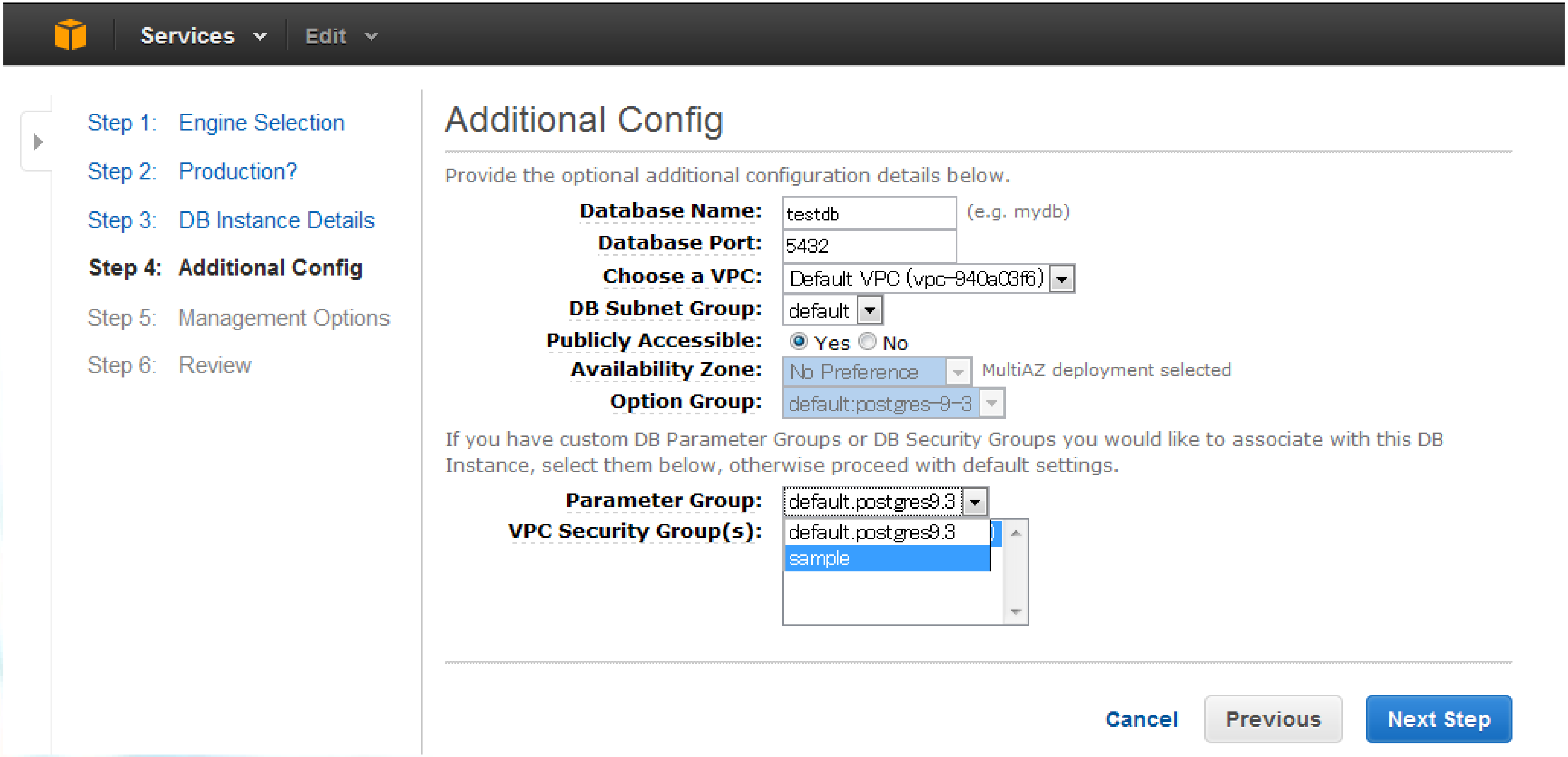
- RDS インスタンスのスペックを選択できる
- 後からインスタンスクラスは簡単に変更可能

DB インスタンスクラス	CPU	メモリ
マイクロ	最大 2 ECU	630 MB
スモール	1 ECU × 1 仮想コア	1.7 GB
ミディアム	2 ECU × 1 仮想コア	3.75 GB
ラージ	2 ECU × 2 仮想コア	7.5 GB
エクストララージ	2 ECU × 4 仮想コア	15 GB
ハイメモリエクストララージ	3.25 ECU × 2 仮想コア	17.1 GB
ハイメモリダブルエクストララージ	3.25 ECU × 4 仮想コ	34 GB
ハイメモリクアドラプルエクストララージ	3.25 ECU × 8 仮想コア	68 GB
ハイメモリクラスタエイトエクストララージ	88 ECU（32 ハイパースレッドコア）	244 GB

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

DB インスタンスの詳細設定 2/2

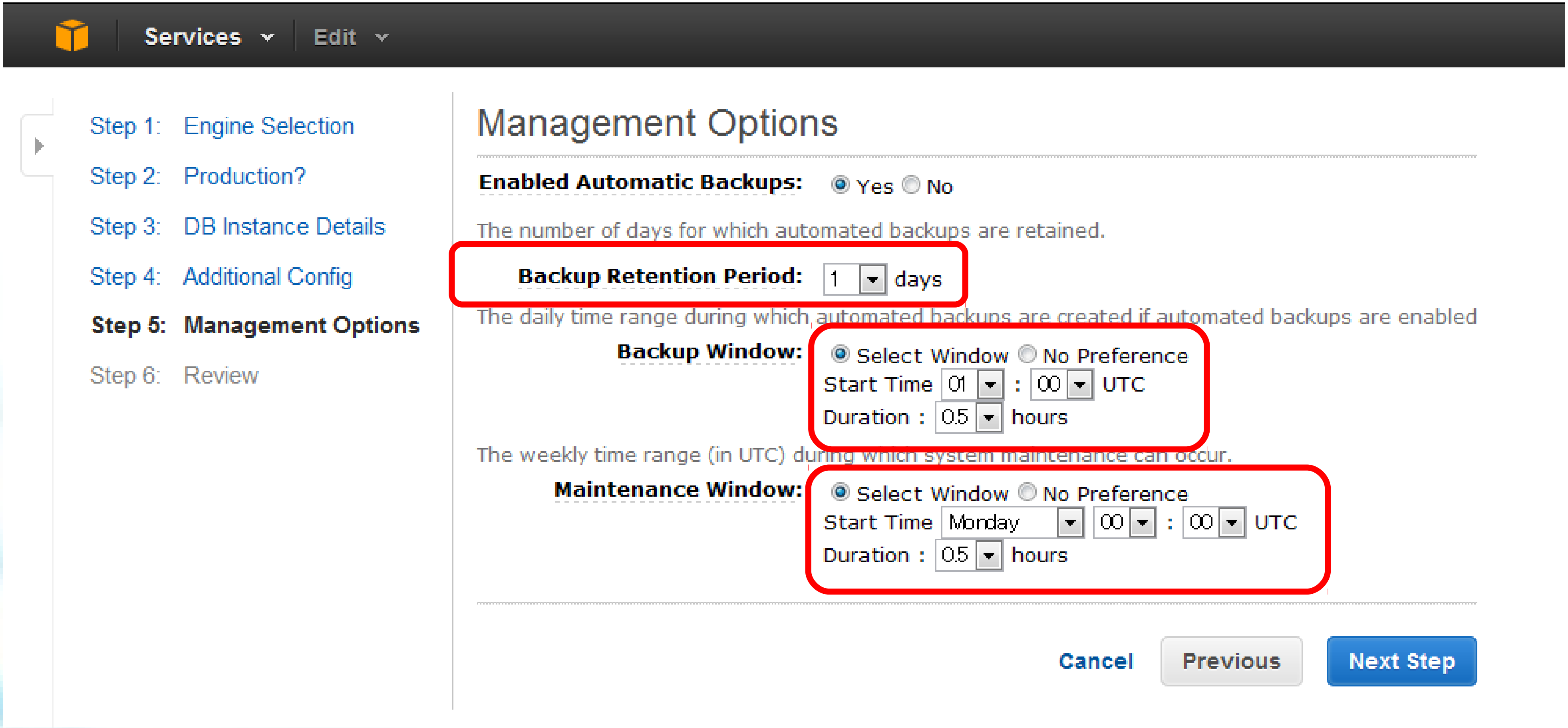
- DB 名・ポート・ネットワーク構成・パラメータ・セキュリティグループ etc.



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

バックアップ・メンテナンスの設定

- 自動バックアップを行う時間帯
- 自動バックアップデータの保存期間は？
- メンテナンスを行う時間帯 (0.5h, 1.0h, 1.5h, … 3.0h)

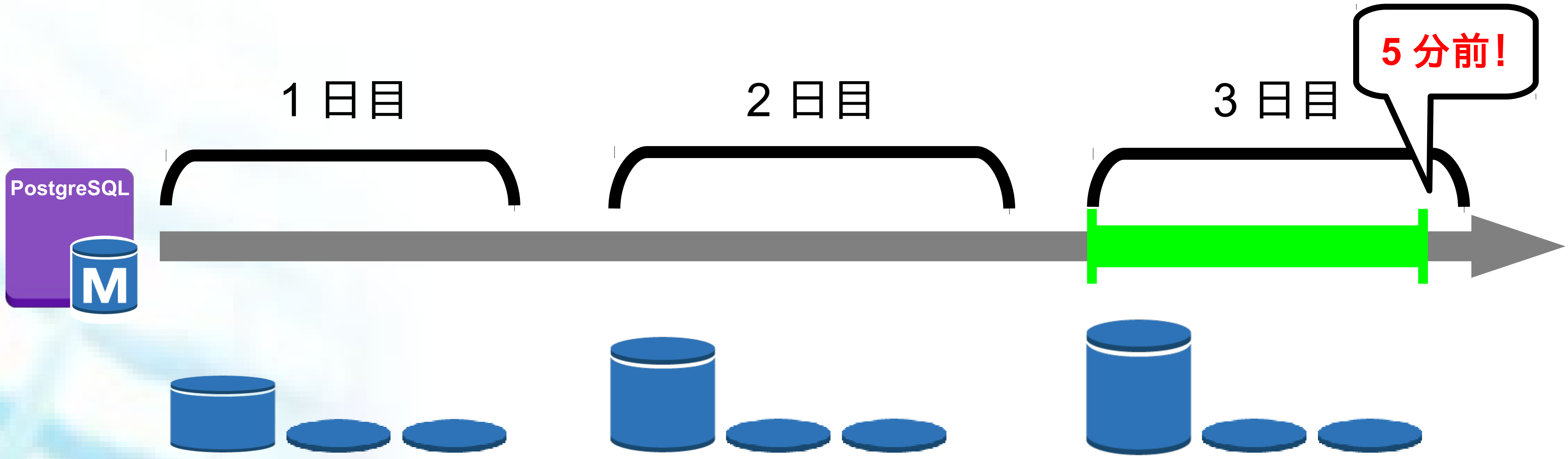


Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

自動バックアップ

- 1 日 1 回の RDS インスタンスのスナップショット + その日の差分情報
- 取得したバックアップデータを 0 ～ 35 日間保存できる
- 保存期間内の任意の時点に RDS をリカバリできる
- ただし最新の復元時点は現在時刻の **5 分前**
  - 差分情報は 5 分置きに取得されるため

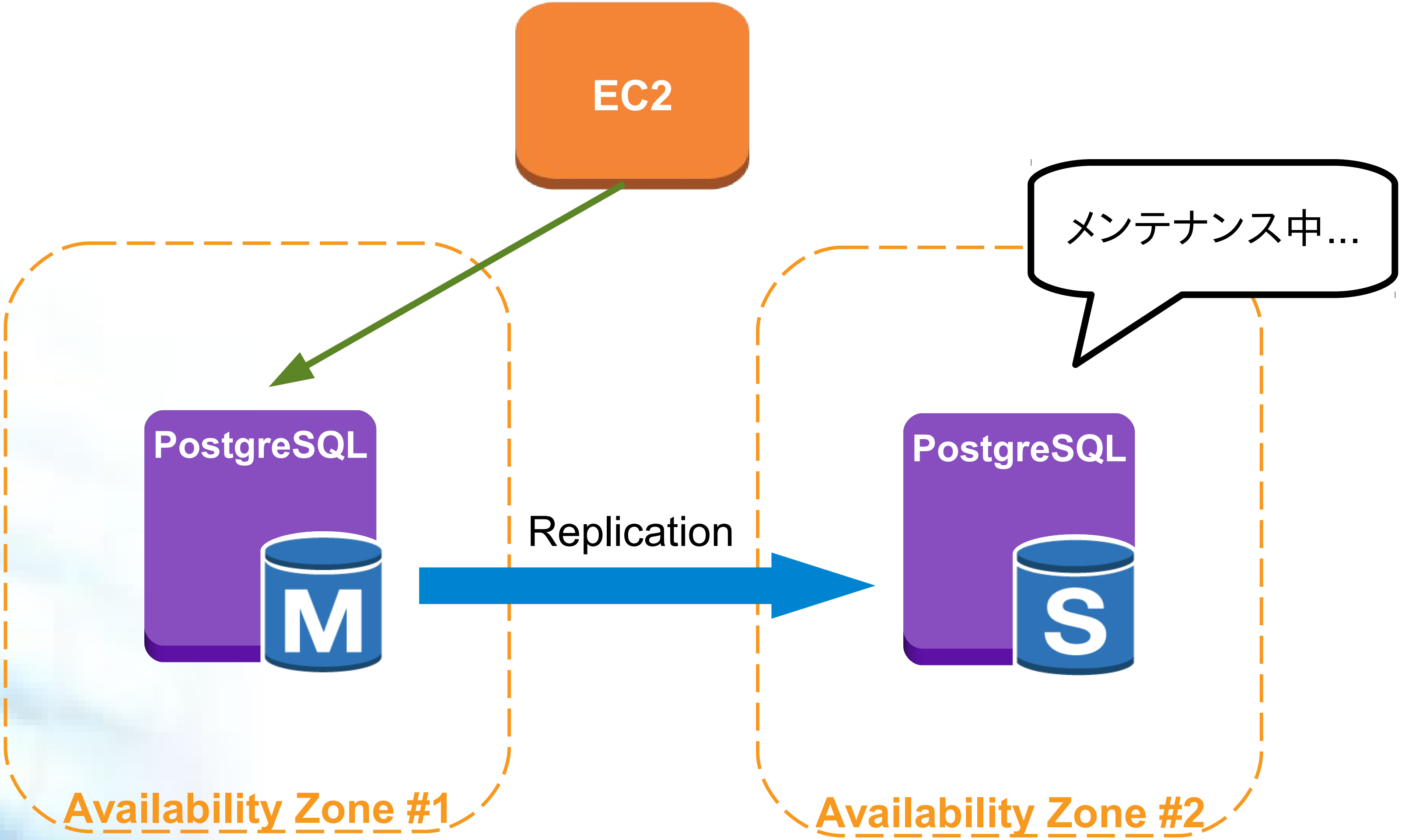
例: Backup Retention Period = 1 days



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

メンテナンスウィンドウ

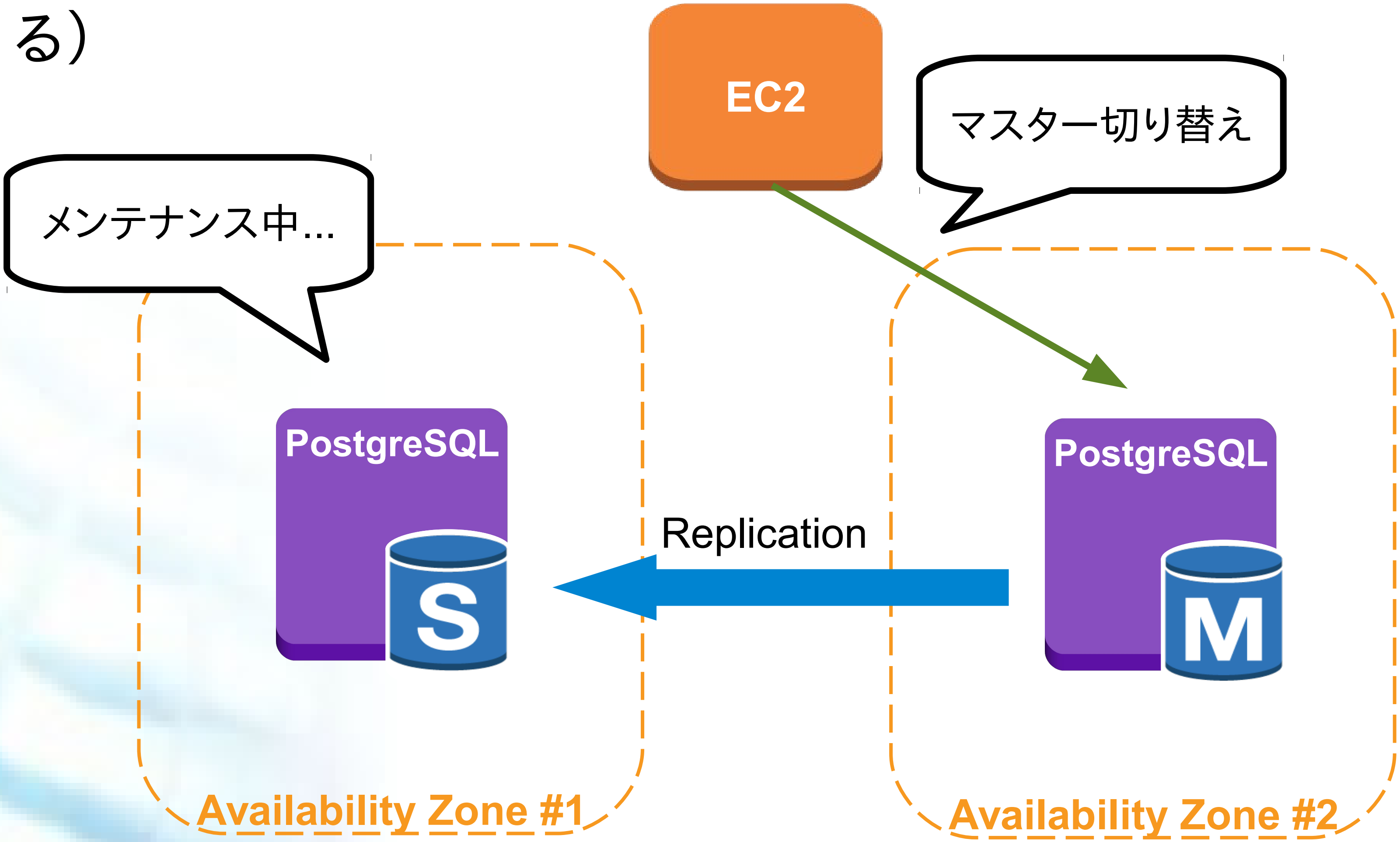
- AWS 側がソフトウェアパッチを当ててもよい時間を指定
- 数ヶ月に 1 度あるかないか
- RDS インスタンスが再起動することがある
- 数分で完了（Multi-AZ の場合はダウンタイムをさらに短縮できる）



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

メンテナンスウィンドウ

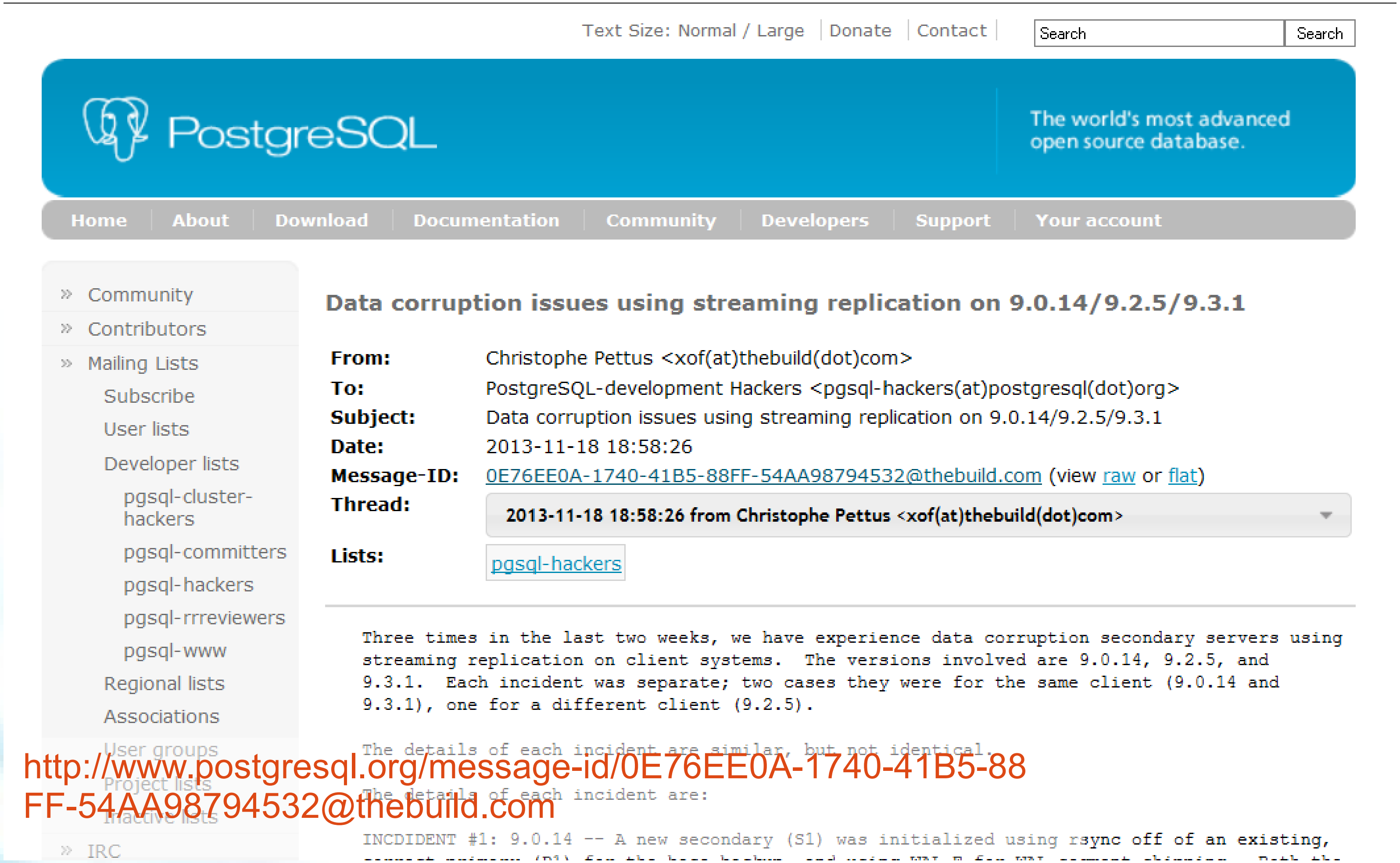
- AWS 側がソフトウェアパッチを当ててもよい時間を指定
- 数ヶ月に 1 度あるかないか
- RDS インスタンスが再起動することがある
- 数分で完了（Multi-AZ の場合はダウンタイムにさらに短縮できる）



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

ちなみに最近の PostgreSQL では

- PostgreSQL 9.0.14/9.2.5/**9.3.1** でバグ
  - マイナーアップデート (2013/12/05) があった
- RDS for PostgreSQL のメンテナンスあるかも



Text Size: Normal / Large | Donate | Contact | Search

PostgreSQL — The world's most advanced open source database.

Home | About | Download | Documentation | Community | Developers | Support | Your account

» Community
» Contributors
» Mailing Lists
Subscribe
User lists
Developer lists
pgsql-cluster-hackers
pgsql-committers
pgsql-hackers
pgsql-rrreviewers
pgsql-www
Regional lists
Associations
User groups
Project lists
Inactive lists
» IRC

Data corruption issues using streaming replication on 9.0.14/9.2.5/9.3.1

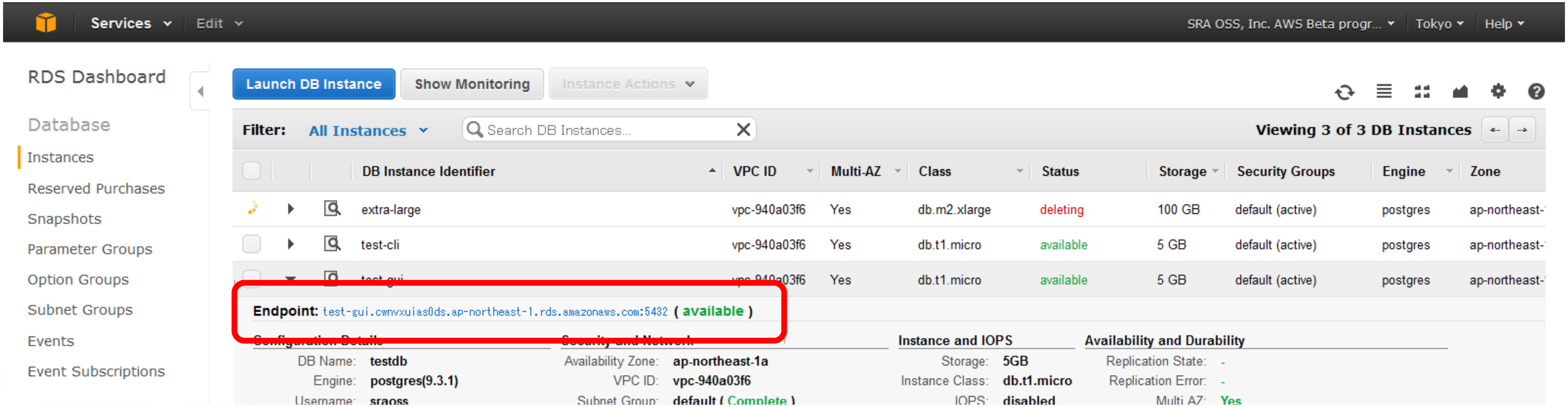
From:	Christophe Pettus <xof(at)thebuild(dot)com>
To:	PostgreSQL-development Hackers <pgsql-hackers(at)postgresql(dot)org>
Subject:	Data corruption issues using streaming replication on 9.0.14/9.2.5/9.3.1
Date:	2013-11-18 18:58:26
Message-ID:	0E76EE0A-1740-41B5-88FF-54AA98794532@thebuild.com (view raw or flat)
Thread:	2013-11-18 18:58:26 from Christophe Pettus <xof(at)thebuild(dot)com>
Lists:	pgsql-hackers

```
Three times in the last two weeks, we have experience data corruption secondary servers using
streaming replication on client systems.  The versions involved are 9.0.14, 9.2.5, and
9.3.1.  Each incident was separate; two cases they were for the same client (9.0.14 and
9.3.1), one for a different client (9.2.5).

The details of each incident are similar, but not identical.

The details of each incident are:

INCDIDENT #1: 9.0.14 -- A new secondary (S1) was initialized using rsync off of an existing,
```

http://www.postgresql.org/message-id/0E76EE0A-1740-41B5-88FF-54AA98794532@thebuild.com

作成完了

- 簡単に作成できる
- 作成した DB のアクセスは Endopoint から

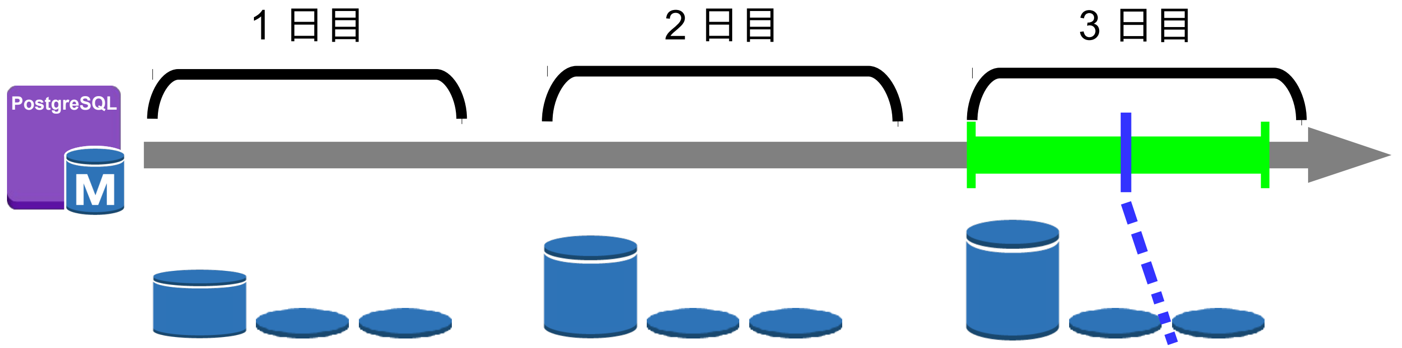


- 100 GB / 1000 IPOS で 15 分程度

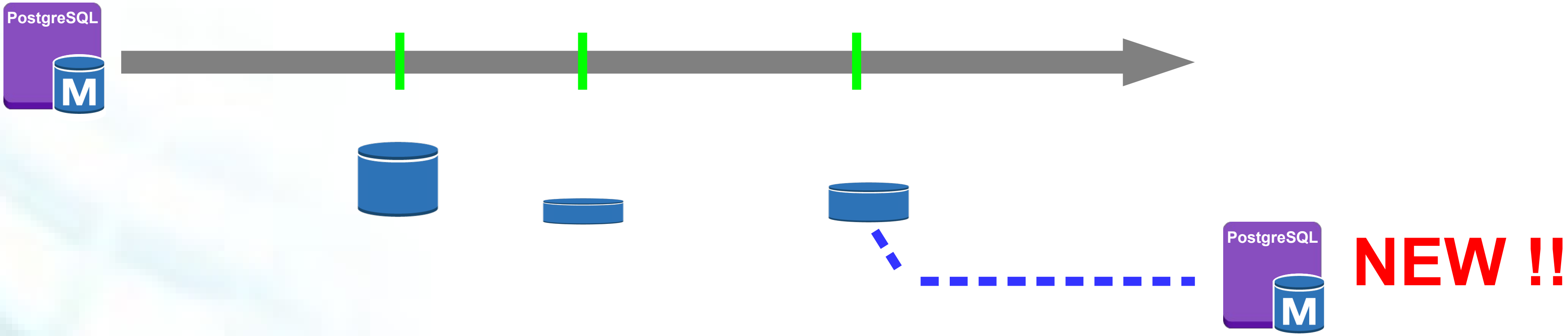
Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

RDS for PostgreSQL のバックアップ戦略

- 自動バックアップ
  - 指定期間内の任意の時点にリカバリ



- スナップショットの取得
  - 取得時点に RDS インスタンスをリストア（複数取得＝差分取得）

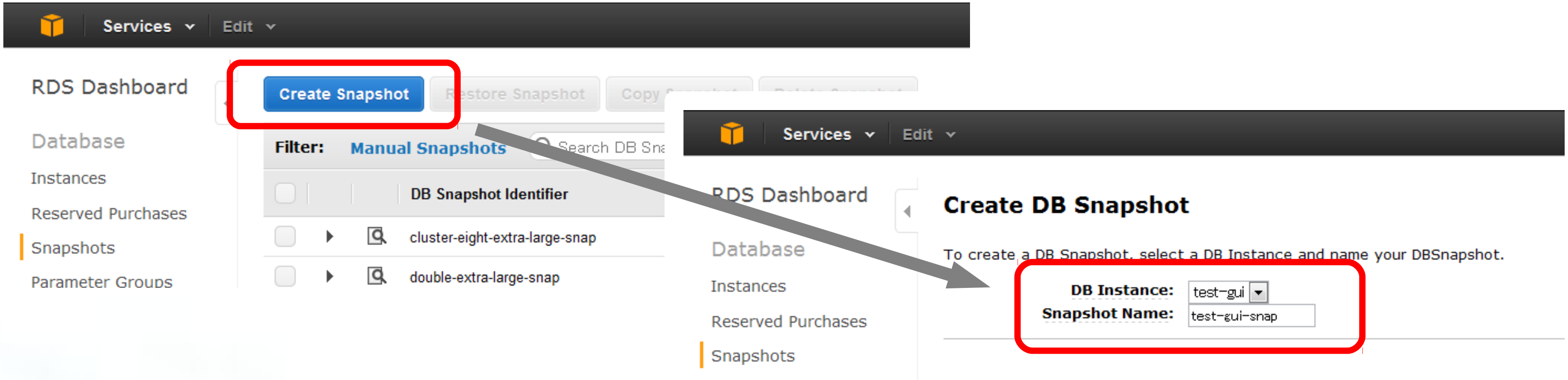


- どちらも**新しいインスタンスを作成する**ことになる
- リカバリ/リストア時に**インスタンススペックを変更できる**

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

スナップショットの取得方法

- Create Snapshot
- 取得したい RDS インスタンスを選択して
  スナップショット名を指定



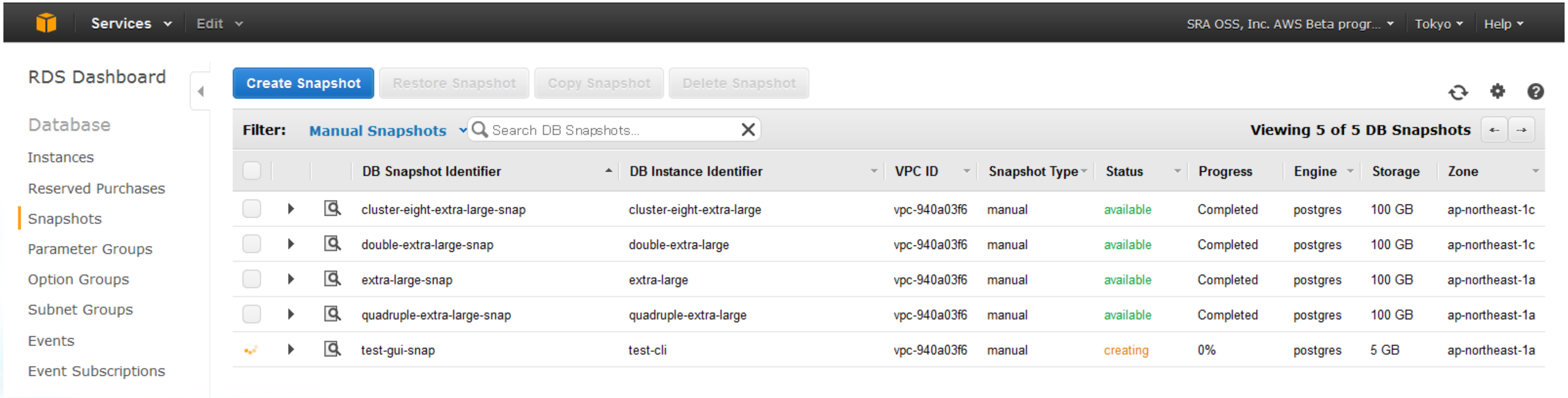
- スナップショットの取得時間は 7 分程度

DB インスタンスクラス	使用量/ストレージサイズ	取得時間
ハイメモリエクストララージ	15 GB / 100 GB	7 分
ハイメモリダブルエクストララージ	15 GB / 100 GB	9 分
ハイメモリクラスタエイトエクストララージ	15 GB / 100 GB	6 分

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

スナップショットの取得方法

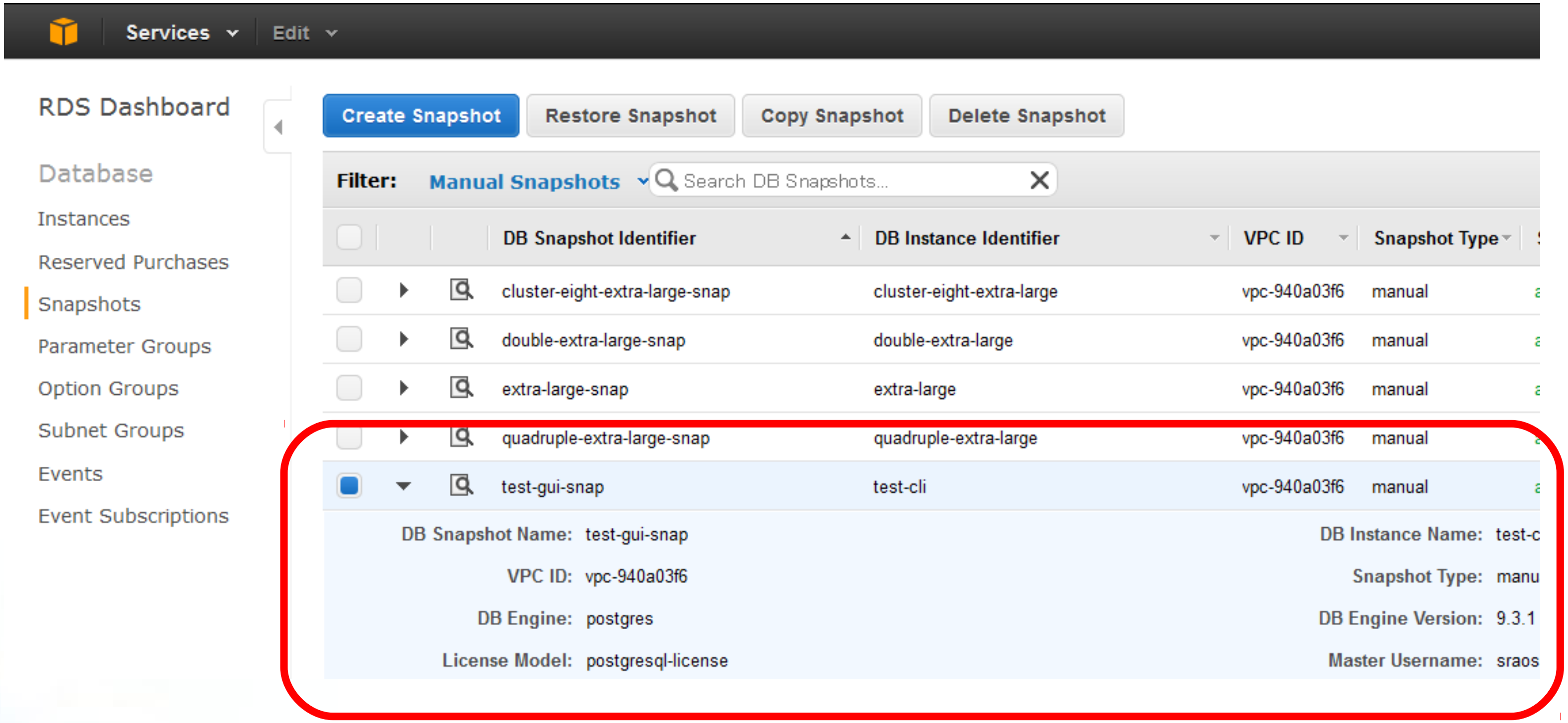
- スナップショットは複数取得可能
- スナップショットのコピーもできる
  - コピー先を異なる AZ に配置してさらに安心
- Multi-AZ を有効にしていないと数分間 DB に接続できなくなる



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

スナップショットからのリストア

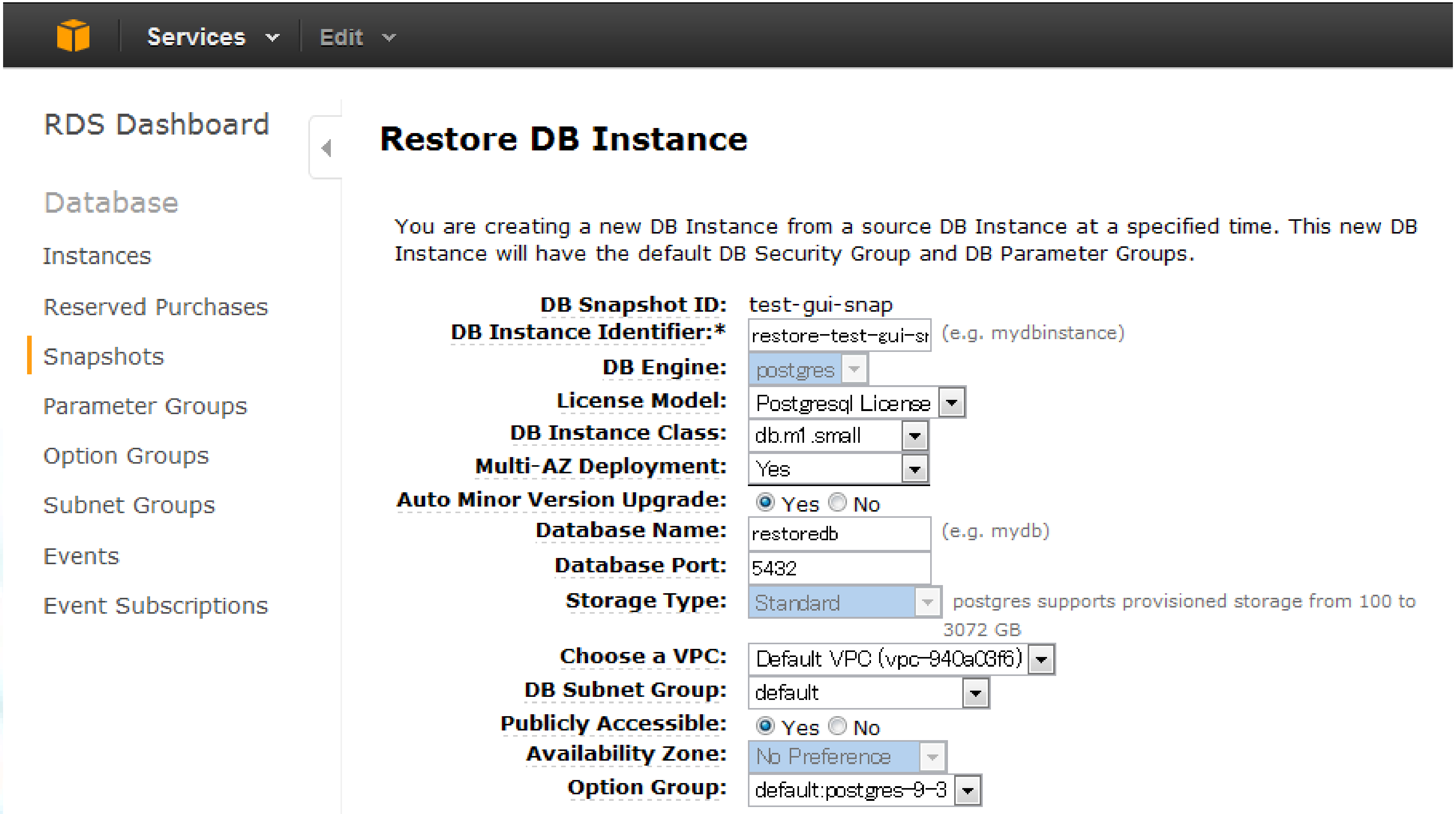
- リストアしたいスナップショットを選択



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

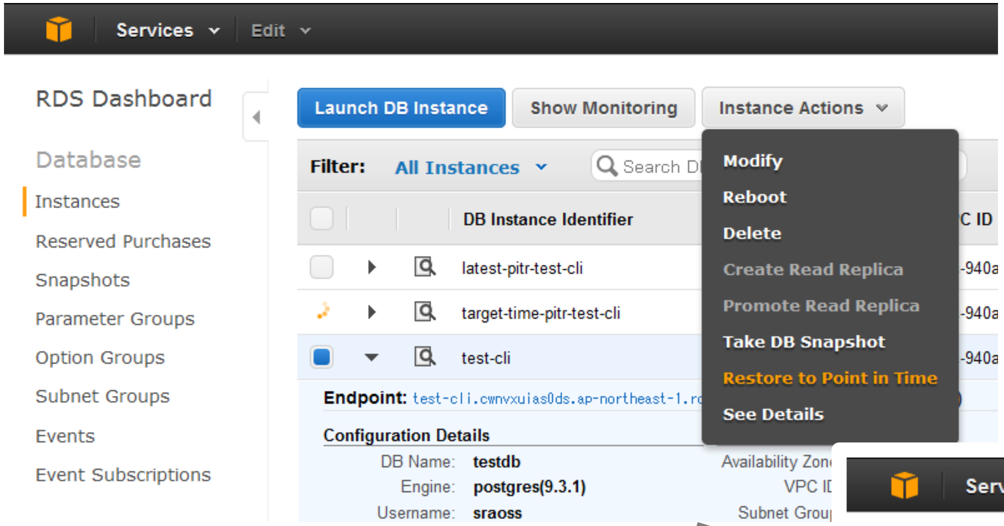
スナップショットからのリストア

- リストア時に**インスタンスのスペックを変更できる!**

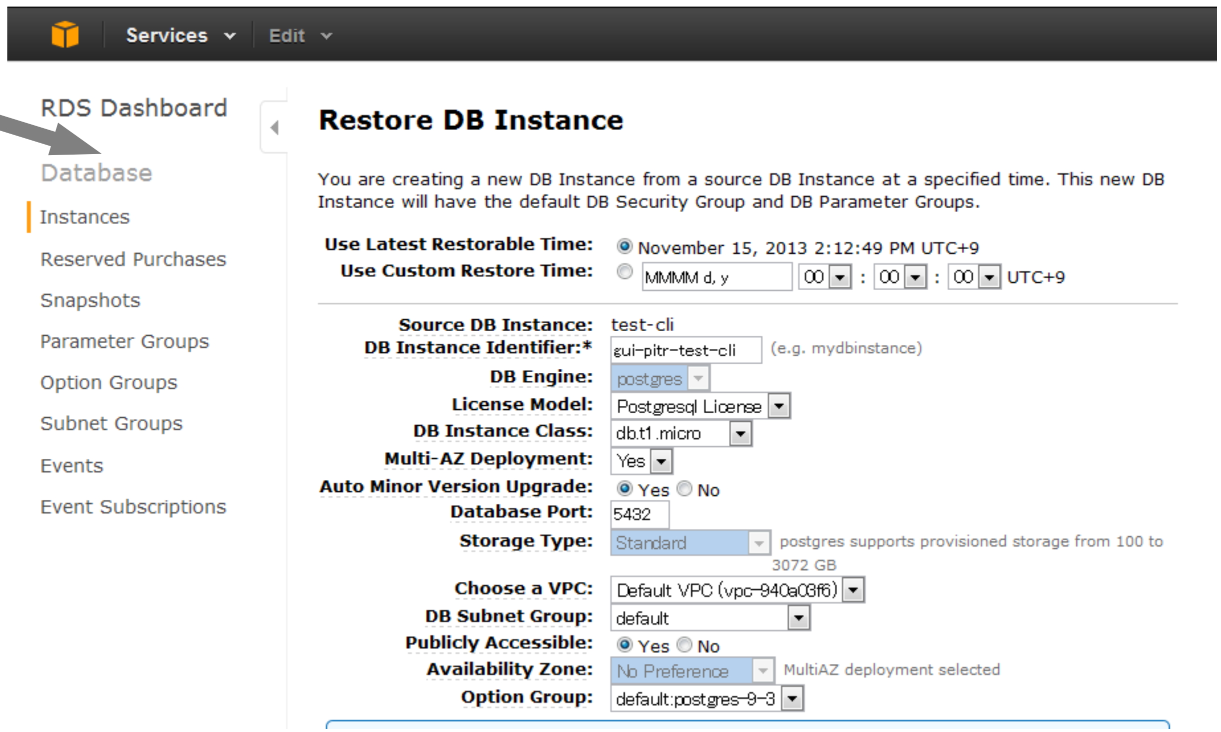


Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

自動バックアップからのリカバリ



- リストアしたい
  RDS インスタンスを選択



- リカバリしたい時刻を選択
- リカバリ時に
  インスタンスのスペックを変更できる!

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

RDS インスタンスのスケールアップ

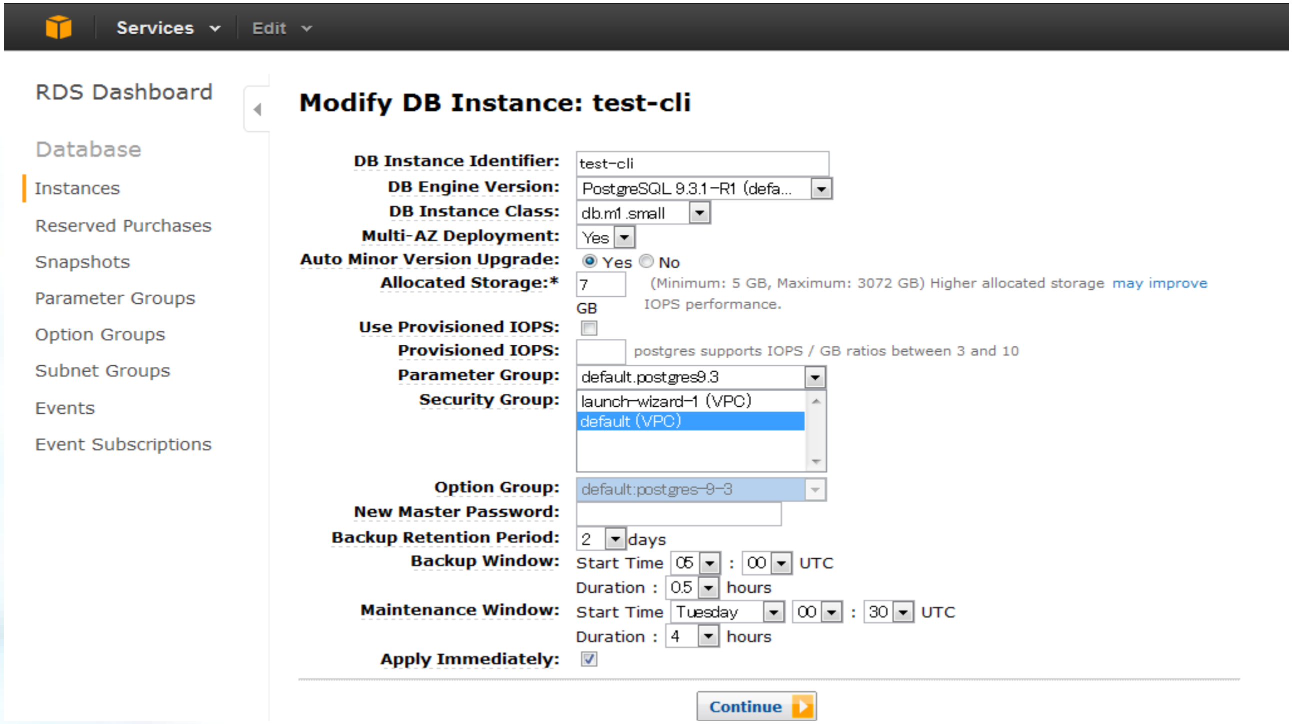
- スケールアップのタイミング
  - 自動バックアップからのリカバリ
  - スナップショットからのリストア
  - **任意のタイミングで**

- 任意のタイミングでスケールアップ（**RDS インスタンスの変更**）ができる

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

RDS インスタンスの変更 | スケールアップ

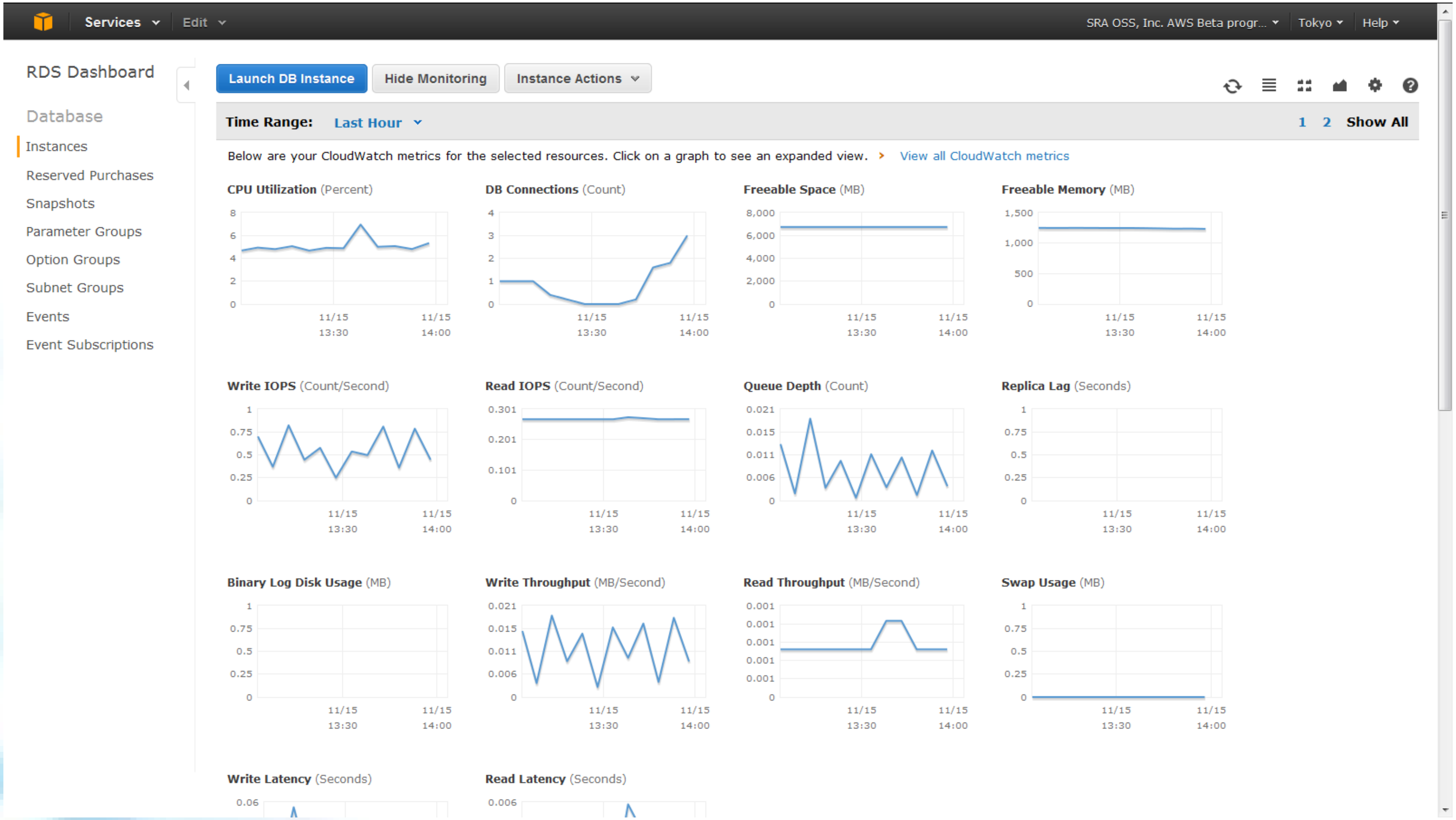
- 指定の RDS インスタンスの内容を変更できる
- 新たにインスタンスが作成されるわけではない
- 即時変更反映 OR メンテナンスウィンドウで指定した時間帯
- 数分ほど RDS インスタンスは停止する
- **ストレージサイズ**の増加はできても**減少はできない**



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

RDS インスタンスのモニタリング

- Web コンソールから 2 週間分の RDS インスタンスの統計情報を確認可能
- この統計情報は、Amazon CloudWatch（AWS クラウドリソースのモニタリング機能）を使っている



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

RDS インスタンスのモニタリング

- Web コンソールから 2 週間分の RDS インスタンスの統計情報を確認可能
- この統計情報は、Amazon CloudWatch（AWS クラウドリソースのモニタリング機能）を使っている

統計情報	内容
CPUUtilization (%)	CPU 使用率
FreeableMemory (Bytes)	利用可能なメモリ量
FreeStorageSpace (Bytes)	利用可能なディスク量
DatabaseConnections (Count)	データベース接続数
ReadIOPS (Count/Second)	1秒辺りの平均ディスク読み込み回数
ReadLatency (Second)	上記の操作の平均所要時間
WriteIOPS (Count/Second)	1秒辺りの平均ディスク書き込み回数
WriteLatency (Second)	上記の操作の平均所要時間

…… など 13 種類

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

RDS インスタンスからイベント通知

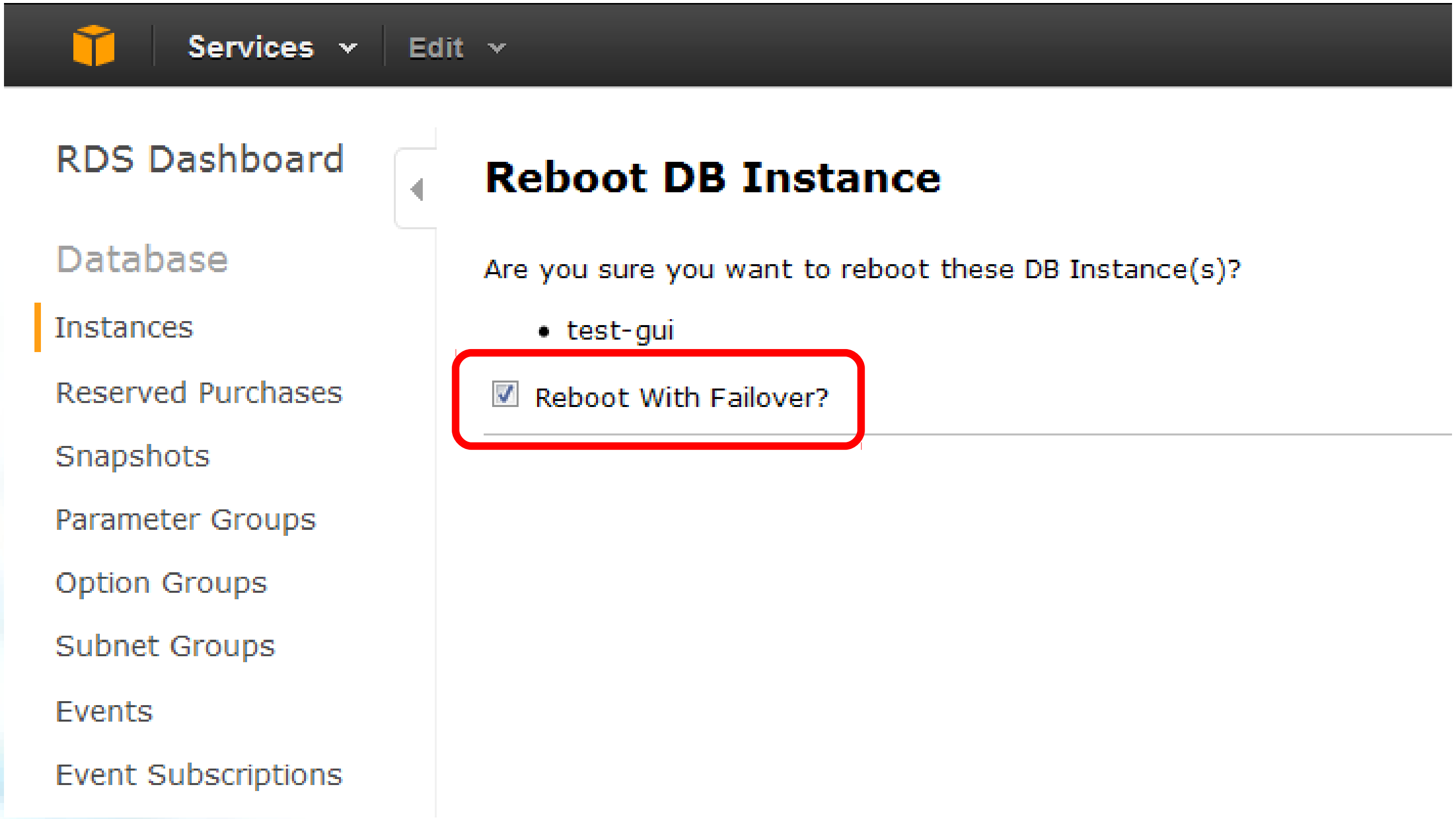
- Amazon Simple Notification Service を通して RDS インスタンスの状態変化をデバイス･アプリにプッシュ通知

Availability	RDS インスタンスのシャットダウン･再起動
Backup	自動バックアップの開始･終了
Configuration Change	セキュリティグループ、インスタンスのスケーリング、パスワード変更など
Creation	RDS インスタンス、スナップショットの作成･削除
Failover	フェイルオーバーの開始･完了
Low Storage	ストレージの容量不足
Maintenance	パッチインストールによるオフラインへの移行･オンラインへの復帰
Recovery	データベースインスタンスの復旧
Restoration	ポイント･イン･タイムまたはスナップショットからの RDS インスタンスのリストア

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

RDS インスタンスのフェイルオーバー

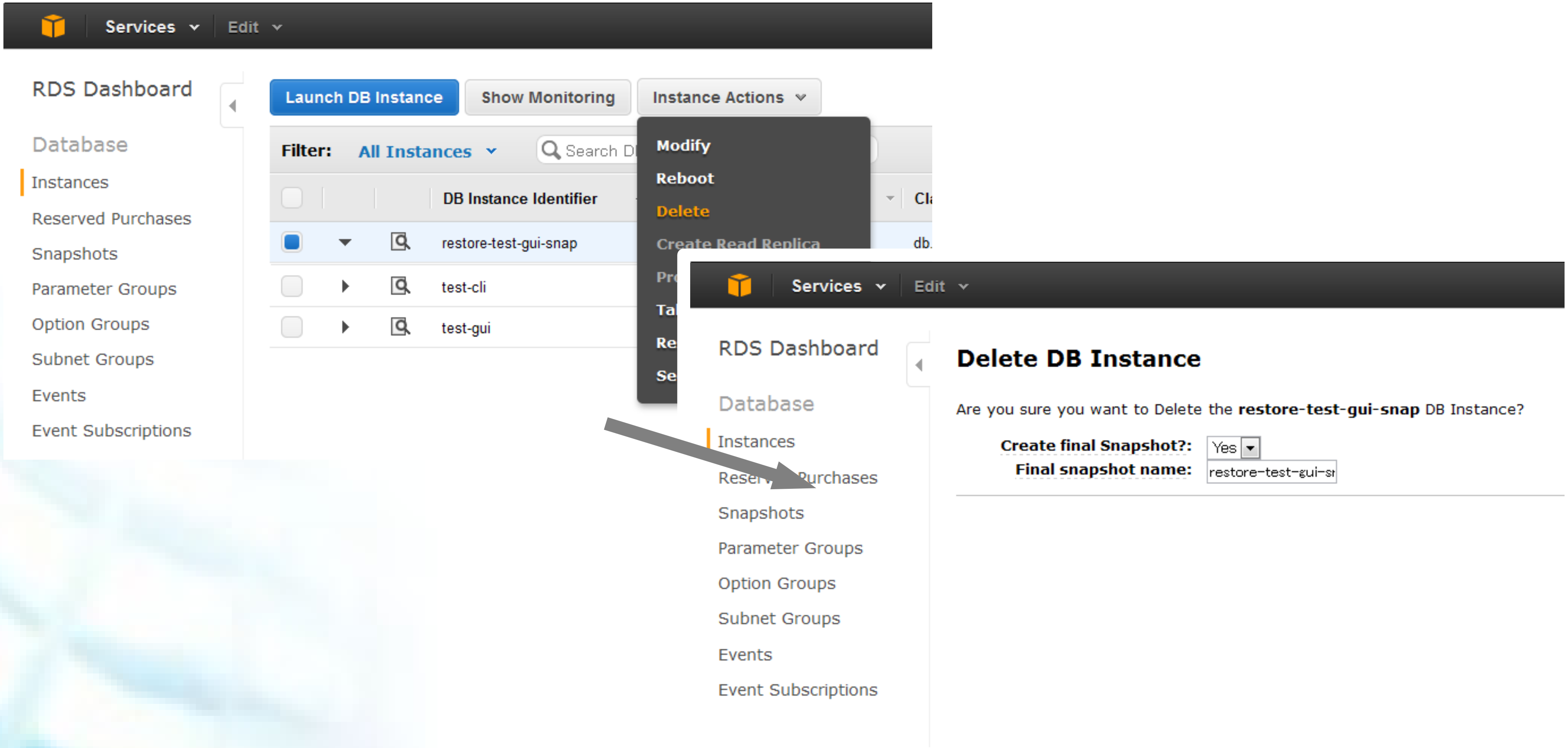
- Multi-AZ を有効にしていると、障害発生時に自動的にスタンバイ側に切り替わりサービス継続
- RDS インスタンスの再起動時にもフェイルオーバーを任意に選択できる



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

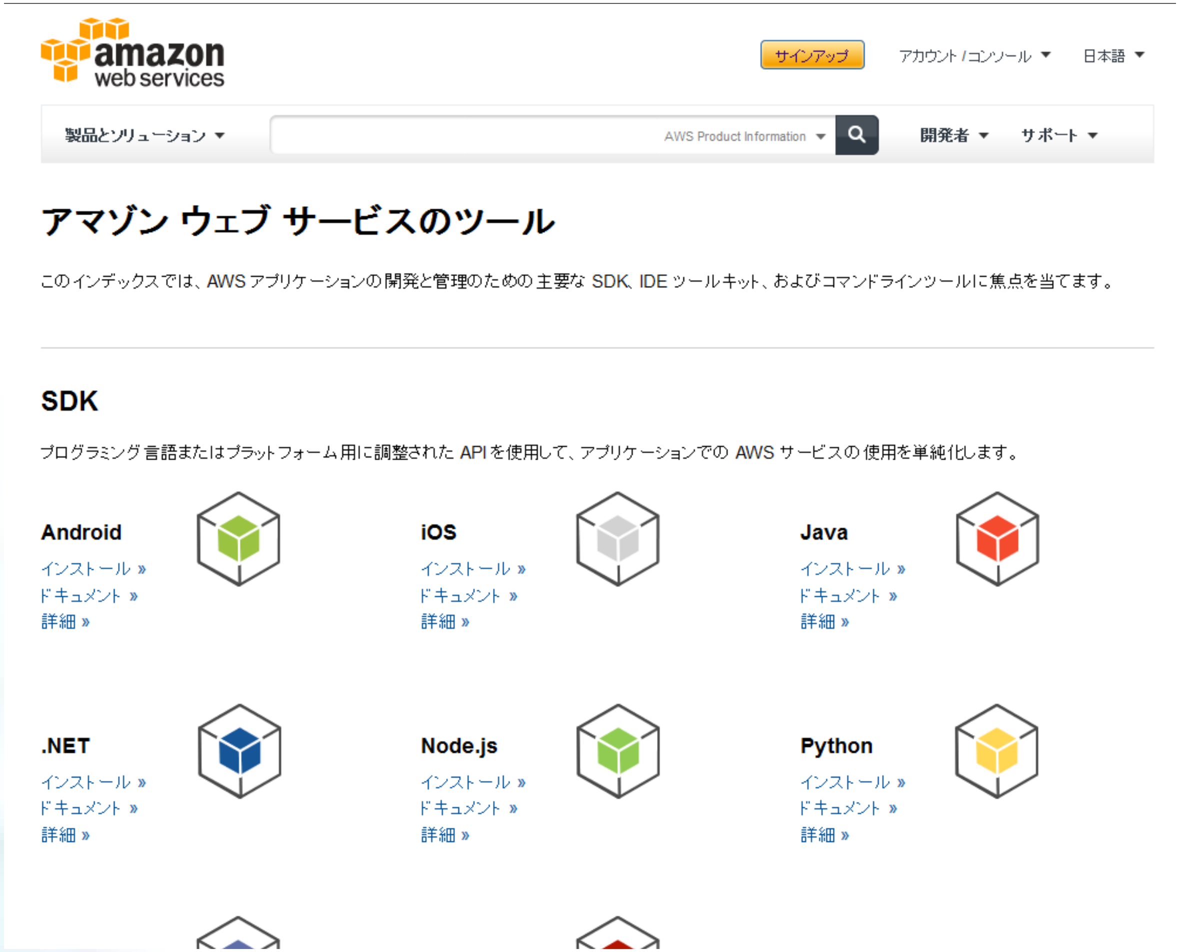
RDS インスタンスの削除

- インスタンスを選択して削除
- 削除時の最終スナップショットを取っておくことも可能
- DB のデータを手元に持ってくる機能は AWS にはないので PostgreSQL のダンプコマンドを使用する



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

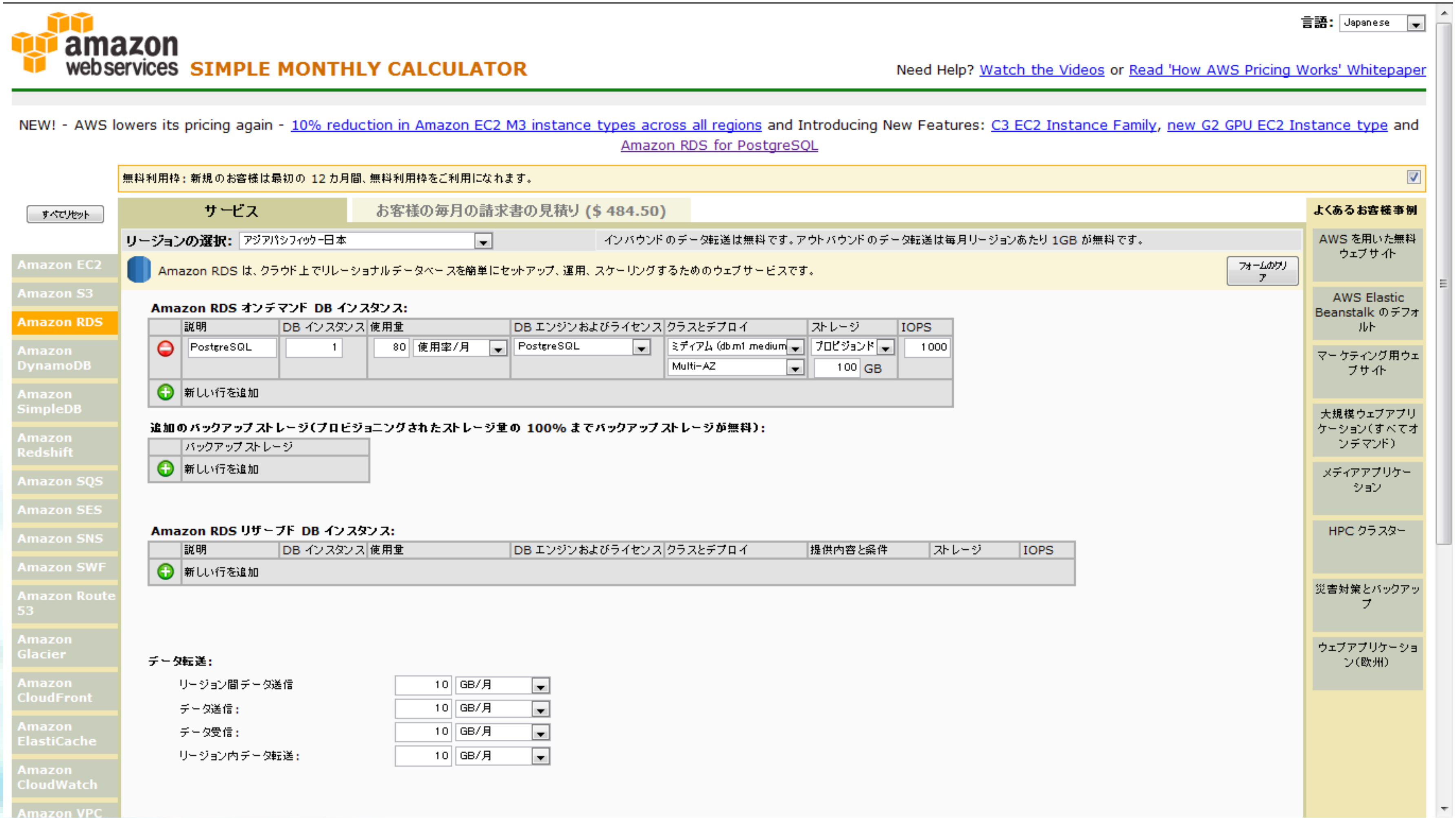
- 以上の操作は Web コンソールだけでなく、各種 API 、スクリプトでも実行可能。
- [AWS SDKs & Tools]
  - https://aws.amazon.com/jp/tools/



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

費用について

- 以下のサイトで試算可能
- [Amazon Web Services Simple Monthly Calculator]
  - http://calculator.s3.amazonaws.com/calc5.html



Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

AWS が提供しているドキュメント

- [ドキュメント]
  - http://aws.amazon.com/jp/documentation/
  - 各サービスのドキュメント
  - 未翻訳のものも多い

- [AWS SDKs & Tools]
  - http://aws.amazon.com/jp/tools/
  - 各種ツールのインストール・使い方

- [AWS クラウドサービス活用資料集]
  - http://aws.amazon.com/jp/aws-jp-introduction/
  - スライド資料
  - わかりやすい概要、事例紹介など

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.

オープンソースとともに

URL：http://www.sraoss.co.jp/
E-mail：sales@sraoss.co.jp
Tel：03-5979-2701

Copyright © 2014 SRA OSS, Inc. Japan All rights reserved.