

Amazon RDS+EC2+ PGroonga+ ロジカルレプリケーション を使った 低コスト高速全文検索

堀本泰弘

クリアコード

PostgreSQL Conference Japan 2019
2019-11-15





目次

- PGroongaとは？
- AmazonRDS と PGroonga



目次

- RDS + EC2 + PGroonga構成
 - 構成図
 - 負荷分散
 - 復旧

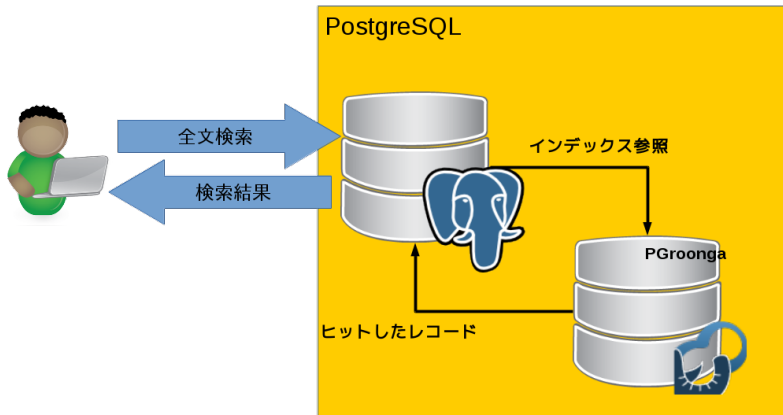


PGroongaとは？

- PostgreSQL経由で使える超高速な全文検索エンジン

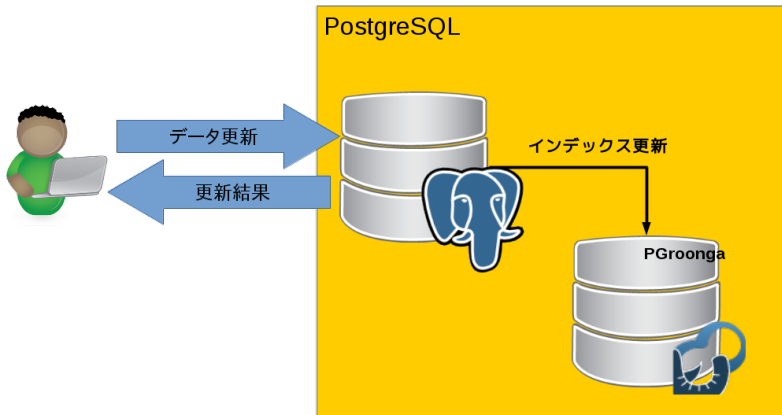


PGroonga : 検索





PGroonga : 更新





PGroonga : メリット

- PostgreSQL経由で使える
(=SQLで検索できる)
 - 実装コスト：小
 - メンテナンスコスト：小
 - LIKEも速くなる



PGroonga : 実行例

テーブル定義

```
-- PostgreSQLのテーブル作成と同じ  
CREATE TABLE entries (  
  title text,  
  content text  
);
```




PGroonga : 実行例

インデックス定義

```
-- 全文検索用インデックス
CREATE INDEX entries_full_text_search
ON entries
-- 「USING PGroonga」 = 「PGroongaを使う」
USING PGroonga (title, content);
```



PGroonga : 実行例

データ挿入

```
-- 普通に挿入するだけでよい  
INSERT INTO entries  
VALUES ('PGroongaで高速全文検索！',  
        '高速に全文検索したいですね！');
```



PGroonga : 実行例

全文検索

```
SELECT title FROM entries
WHERE
-- &@~で全文検索
-- 「検索」と「高速」をAND検索
title &@~ '検索 高速' OR
content &@~ '検索 高速';
```



PGroonga : 機能

- 本格的な全文検索エンジンをバックエンドで使用
 - 全文検索に十分な機能
 - 同義語検索
 - 類似文書検索
 - 読みがな検索 などなど



PGroonga : 速度

- 検索が安定して速い
 - ヒット数が多くても、あまり速度が変わらない
 - 更新中でも検索が遅くならない
 - 更新されたデータは即検索できる



Amazon RDS と PGroonga

- Amazon RDSは管理コストを低減できる
- PGroongaは速くて機能が豊富



Amazon RDS と PGroonga

- RDSとPGroongaを組み合わせれば、運用が楽で高機能、高速な全文検索が実現！



Amazon RDS と PGroonga

- ただ、RDSは自由に拡張機能をインストールできない
- PGroongaもRDSでは使えない



Amazon RDS と PGroonga

- RDSにPGroongaをインストールしなくても、RDSのメリットを活かして全文検索したい！



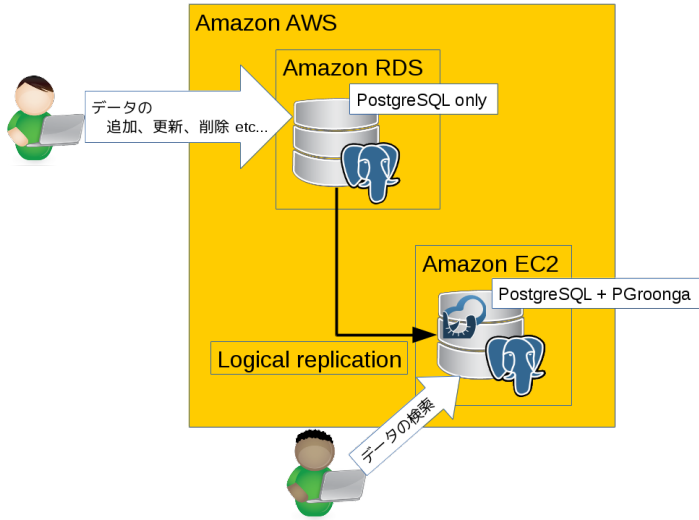
システム構成

以下の4つを組み合わせで実現

- Amazon RDS
- Amazon EC2
- PGroonga
- ロジカルレプリケーション



構成図





Amazon RDSと Amazon EC2の役割

- RDSは更新専用
- EC2は検索専用



PGroongaの インデックス設定

- ロジカルレプリケーション
 - DDLをレプリケーションしない
- この特徴を利用



PGroongaの インデックス設定

- EC2だけにPGroonga拡張をインストール
- EC2だけに全文検索用のインデックスを設定



データの管理

- データは、RDSからロジカルレプリケーションで取得
 - RDS上にデータがあれば、EC2での検索が可能



負荷分散

- 検索リクエストが増えてくると
EC2 1台ではリクエストが処理
できない



負荷分散

- データはRDS上にある
- EC2へはロジカルレプリケーションでデータを同期



負荷分散

- 新規にEC2インスタンスを作成
→サブスクライバとしてRDS
に接続
→検索用のサーバーが完成!



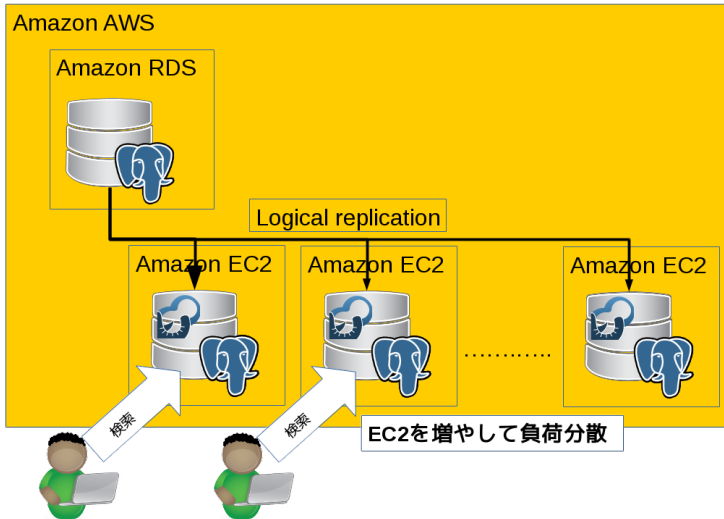
負荷分散

つまり

- EC2を増やすことで負荷分散ができる



負荷分散：構成





復旧

- EC2へアクセス不能なときは？



復旧

- データはRDSにある
 - RDSにアクセスできるEC2を新しく作ればOK
- アクセス不能なEC2は破棄
 - 復旧をがんばらなくてよい



まとめ

- データの管理はRDSで実施
 - 運用負荷低減
- 検索はPGroongaで高速に検索
 - EC2は検索に専念できる
- 運用負荷を低減しつつ、高速、高機能な全文検索を実現！



最後に

高速で高機能な全文検索をお求め
の方は↓↓

問い合わせ先：

[https://www.clear-code.com/contact/?
type=groonga](https://www.clear-code.com/contact/?type=groonga)