



Rroongaの全文 検索性能

Naoya(@naoa_y)

*61st Ruby/Rails勉強会@関西 LT
2014/05/31*

自己紹介



- ✓ Naoya (@naoa_y)
- ✓ 大学は情報系
- ✓ 新卒で3年半ほど金融系のユーザSIでインフラSE
- ✓ 現在は3年半ほどITと無縁の仕事
- ✓ Ruby歴は実働数日ほど



国産の全文検索 エンジン

Groongaって 知ってますか？

Groongaって？



- ✓ C言語で書かれた高速な全文検索ライブラリ/サーバ
 - ✓ MySQLのストレージエンジンとして実装されたMroonga
 - ✓ Rubyらしい記法でRubyから全文検索できるRroonga



個人でMroonga
でそこそこ大規模
な全文検索Web
サービスを作りま
した。



Groongaは少ないハードウェア資産でもかなり高速な全文検索ができる！



デスクトップ向けの全文検索アプリをRubyで作りたい！

Rroongaって？



- ✓ Rroongaのほかにサーバがなくともローカルのみで高速な全文検索機能が実現可能
- ✓ アプリに組み込みやすい。イメージとしてはSQLiteみたいなもの

Rroongaの利用例



- ✓ るりまサーチ
 - ✓ Rubyのリファレンスマニュアルを検索できるWebアプリ <http://docs.ruby-lang.org/ja/search/>
- ✓ Milkode
 - ✓ ソースコード検索エンジン、検索アプリ <http://milkode.ongaeshi.me/> (るびま0045号に記事掲載)



実際のところ
Roongaっ
てどのくらい
速いの？



Rubyだし
言うほどじ
や・・・？



試してみよう！

検証環境



- ✓ MacBook Air
 - ✓ CPU 1.4GHzデュアルコアIntel Core i5
 - ✓ Memory 4GB
 - ✓ SSD 128GB

検証方法

- ✓ 日本語Wikipediaのカテゴリ上位100件を検索
- ✓ カテゴリ上位100件で日本語Wikipediaのタイトルと本文を全文検索
- ✓ 検証プログラム <http://github.com/naoa/rroonga-wikipedia-bench>

検証1

日本語Wikipedia 1万件 XML
168MB

	SQLite3	Rroonga
カテゴリ	2314 msec	47 msec
本文	103 msec	6 msec

検証2

日本語Wikipedia 5万件 XML
512MB

	SQLite3	Rroonga
カテゴリ	9339 msec	88 msec
本文	341 msec	30 msec

検証3

日本語Wikipedia 183万件(全件)
XML 17GB

	Rroonga
カテゴリ	1601 msec
本文	56 msec

検証4



日本語Wikipedia 366万件(全件
*2) XML 33GB

	Rroonga
カテゴリ	3087 msec
本文	113 msec

検証5





**MacBook
Airのディスク
が逼迫したの
でここまで！**



**Mroongaでの経験
則から言うところの数
倍ぐらいなら高速に
全文検索できると思
います。(日本語で
Bigramの場合。)**



Mroongaを使った全文検索のWebサービスでは、最大で400GB超、1000万レコード超でもサーバ1台でそこそこ高速に検索できています。



**Rroongaで
は、高速な全文
検索のほかに位
置情報検索機能
もあります。**



先日、Heroku
でRroongaが
使えるビルドパ
ックがリリース
されています。



Rroongaを
使ってみた
くなりました
たよね？

おわりに



- ✓ 2014/6/27(金)19:00から神戸でGroongaに関する勉強会をします。
- ✓ Groongaに興味があれば是非参加してください！

おわりに

- ✓ Rroonga,Rabbit等の作者、Ruby/Groongaコミッタの須藤さん(@ktou)が東京から来てくれます！
- ✓ 2014/6/28(土)に大阪でGroongaに関するイベントができればと考えています。
- ✓ 会場提供等に協力していただける方がいたら嬉しいです。

ありがとうございます
うございました
ました