

ngx_http_groonga 全文検索nginx

須藤功平

クリアコード

nginx Tech Talks
2016-02-08





nginx

- HTTPサーバー
- HTTPリバースプロキシサーバー
- メールプロキシサーバー
- TCPプロキシサーバー
1.9.0から。まだmainline。

nginx = いろいろできるサーバー



全文検索nginx

- HTTPサーバー
- HTTPリバースプロキシサーバー
- メールプロキシサーバー
- TCPプロキシサーバー
- **[New!]** 全文検索エンジン！



使い方：登録

登録



HTTP
POST

全文検索
NGINX



使い方：検索

検索



HTTP GET
?query=nginx



全文検索
NGINX



速度：1リクエスト

- 対象：Wikipedia日本語版
- レコード数：約185万件
- データサイズ：約7GB
- 検索キーワード：nginx

0.6ms

Intel Core i7-6700 3.40GHz



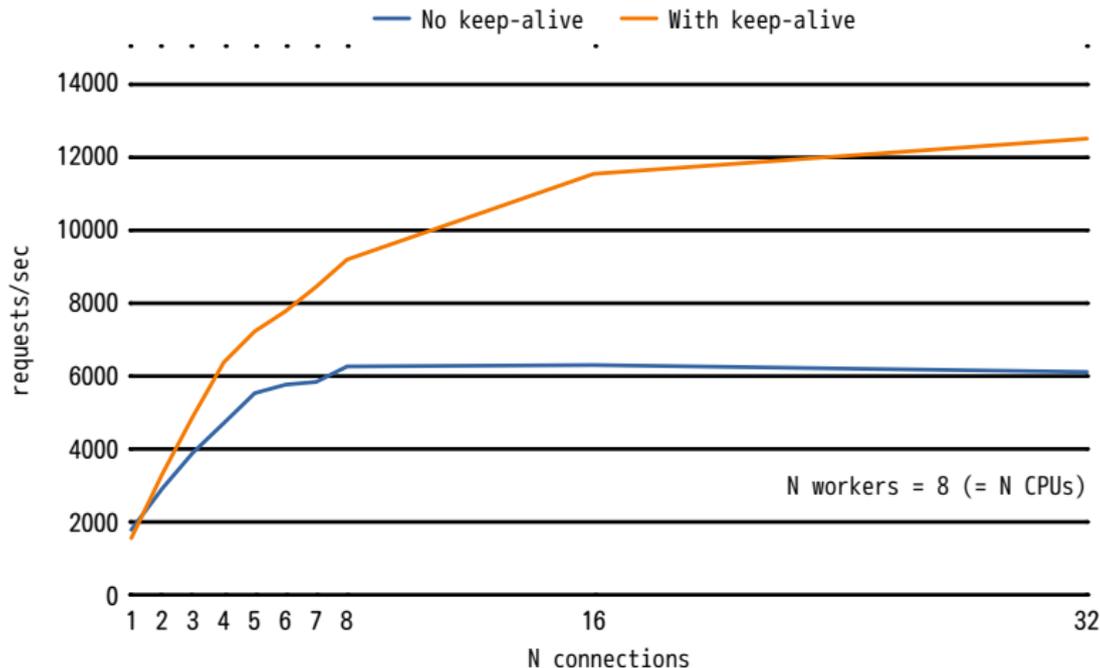
速度：スループット



ワーカー数を増やすとCPUコア数までスケール



速度：スループット



クライアントがkeep-aliveを使うとさらにスケール



速度

- 1リクエストの処理も速い
- スループットも出る
 - ワーカー数を増やすとスケール
 - keep-aliveでさらにスケール



速さの理由

- 本物の全文検索エンジンを組込
 - →1リクエストの処理が速い
(しかも豊富な機能)
- それもnginxのよさを殺さずに
 - 例：マルチプロセスでスケール
 - 例：性能向上用のHTTPの機能を利用
(keep-alive、レスポンス圧縮など)
 - 例：I/O多重化 (同時接続数の増加に対応)

 本物の全文検索エンジン

grnga

(ぐるんが)



Groonga

- 速い
- 日本語に強い（国産）
- ライブラリーとして使える
 - 組み込みやすい
- マルチプロセス対応
 - 複数のプロセスで同時にDBを使える



組込方法

モジュール



ngx_http_groonga



HTTP

HTTP, I/O, TLS, ...

ngx_http_groonga



普通の
← nginx

全文検索
NGINX

Groongaから見たメリット

- nginxと連携すると…
 - プロセス管理を任せられる
 - クライアントとのI/Oを任せられる
 - keep-alive・認証・TLS・圧縮対応
 - HTTP2対応 (まだmainline)

全文検索に集中できる！



困ったこと

- nginxのWindows用バイナリーをGNU/Linux上でビルドできない
 - Groongaはバイナリーを配布したい
 - Windowsが必要だとリリース作業がツライ



対策

パッチを送る



結果

取り込まれた
(1.7.7 released at 2014-10-28)

パッチが取り込まれるまで

- hgでパッチを作る
- メーリングリストに送る
- やりとりする
- 取り込まれる

結構反応してくれるから
なにかあったら送るといいよ！



全文検索nginx

■ nginx + Groonga

= groonga-httpd : Groongaパッケージに含まれている

- nginxが全文検索エンジンに！
- nginxのモジュール機能で実装

モジュールを作って
nginxと共存しよう！



お知らせ

- MySQL・PostgreSQLとも共存
- Groongaは肉の日リリース
2月9日にイベントあります！

MySQLとPostgreSQLと
日本語全文検索
(DMM.comラボにて)