

Yappoの思想すべて見せます

iYappo

大沢 和宏

本日の内容

- iYappoと検索ツール
- 私の採ったツール選定
- 将来について

iYappo とは



携帯サイト検索



iYappo



[GOODS ONLINE](#) -Frame Roll-



890円 1,365円 2,100円 1,890円 1,365円 3,990円

[スヌービー アクセサリーがいっぱい GOODS ONLINE](#)

最新ブログ検索



i.yappo.jp/

iYappo / YappoLabs

- 携帯向け検索サイト
- 1999年7月にサービス開始
 前身のYappoは1997年1月から
- 登録サイト数、約 8万4千
- YappoLabsは、iYappoの新機能を実験する牧場
- 個人運営

iYappoの問題

iYappoの問題

- 全文検索機能をMySQLに頼る
- 具体的にはwhere ~ like 文
grep検索と大差ない
- 効率的なスケールアップに難点
MySQLありきのスケールアップ
- 検索速度が遅い・・・

検索速度向上手法

- アプリ側のSQLクエリキャッシュ
- 動的なHTMLのキャッシュ

根本的な解決にならず
子供だまし

全文検索ソフト導入

選定基準

- N-gram であること
 - 携帯ユーザーには分かち書きが通用せず
- 検索結果のソートカスタマイズ
 - 独自スコアリングでソート
- めんどくさくない事

Namazu / Estraier

- 2004年秋での有力候補
 選択肢がなかった
- 組み込み作業が大変
- **そもそも分かち書き**

Hyper Estraier

- 十分興味があった
P2Pが1998年ごろの思想
- でも、公開直前だった
- 今も勉強不足・・・

Senna

- Senna + MySQLはとても強力
- 検証作業に時間がかかりそう
- 文書属性の面倒を見てくれない
テーブル設計面倒

Rast

- N-gram
- BDB + APRなので、信頼できそう
Subversionモデル
- 文書属性に対する柔軟さ
- なんか「分散」って書いてある

結論

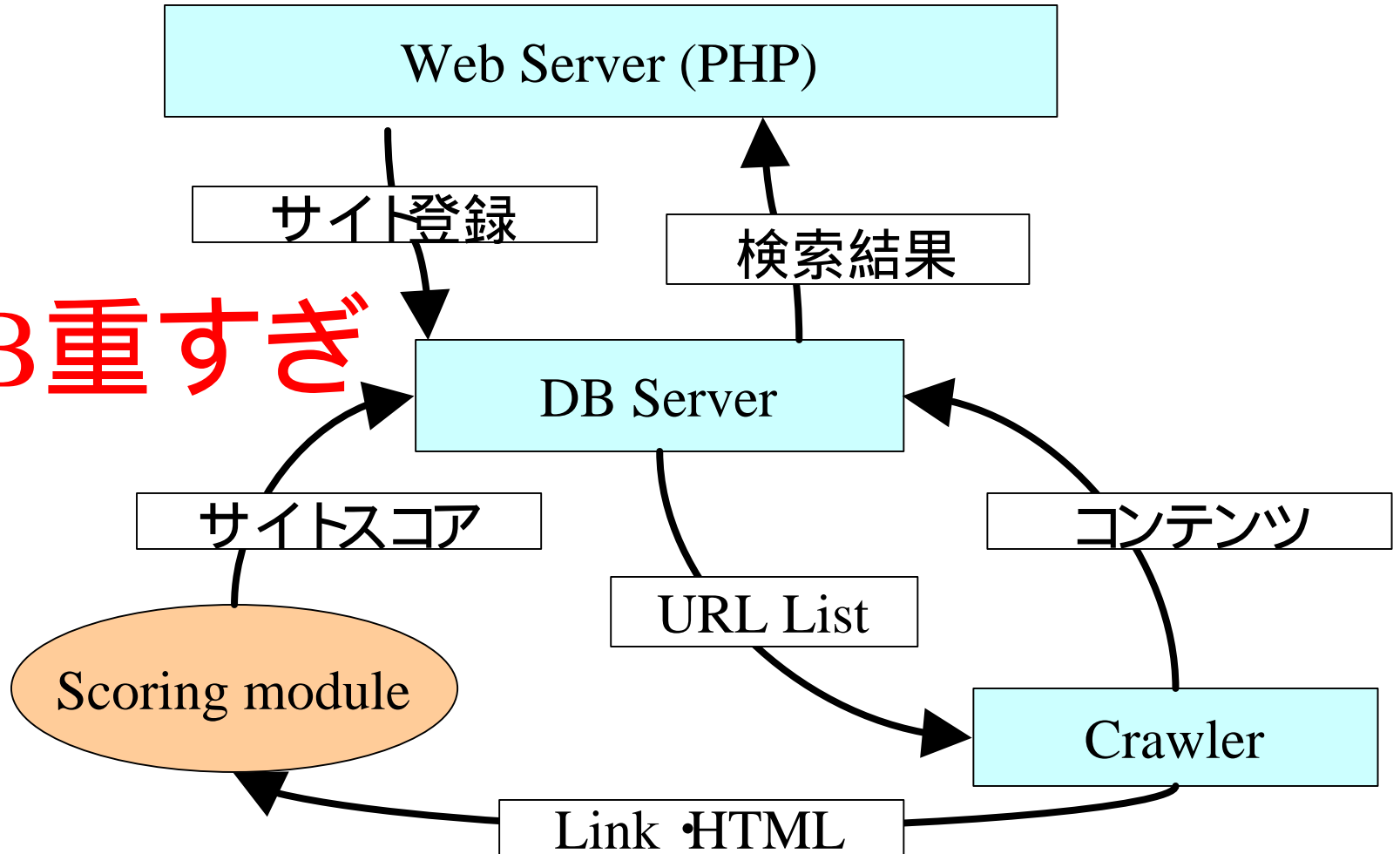
一目惚れ重要

Rastの利用方法

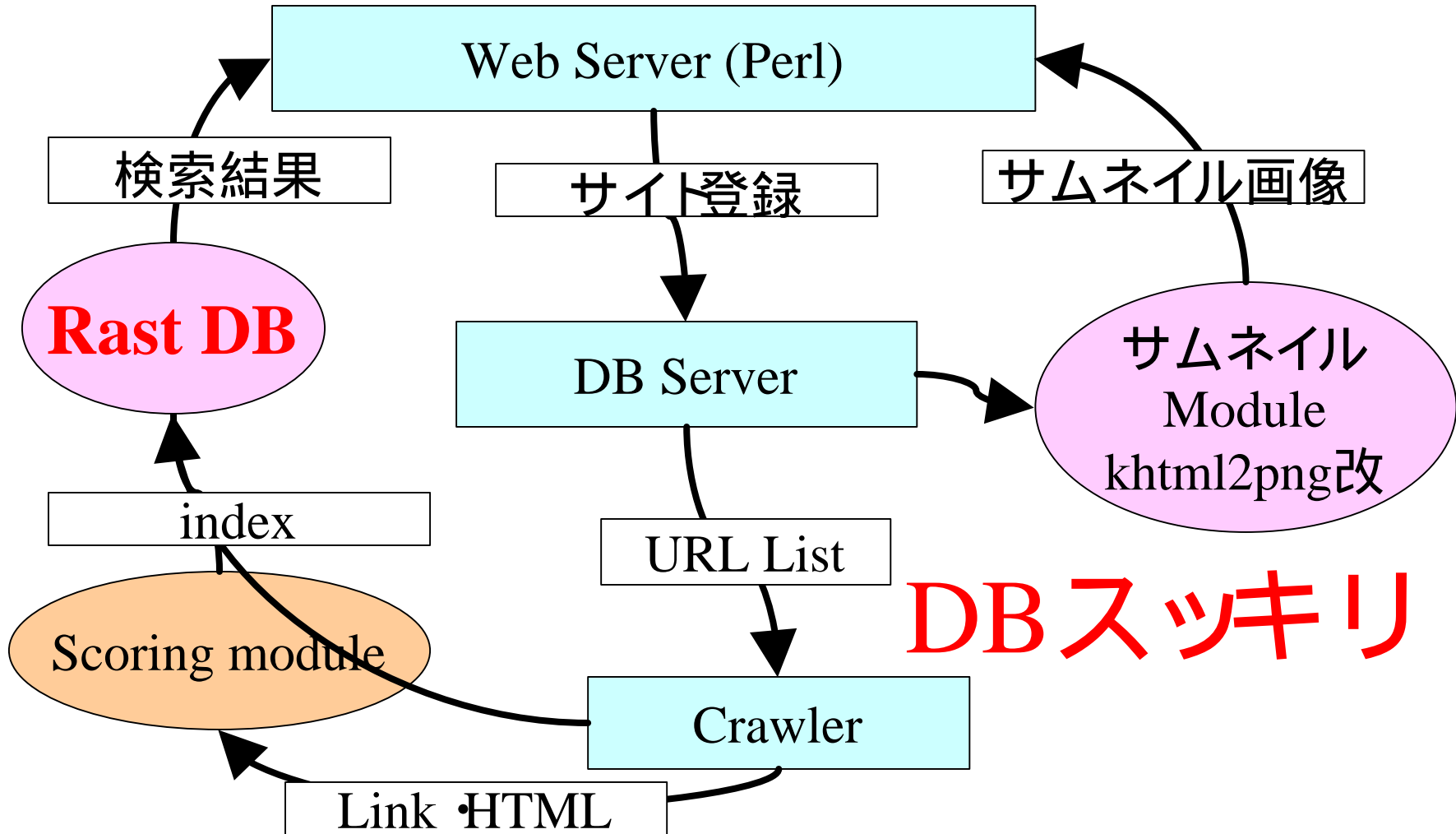
- 現在はSearchのみの評価フェーズ
- 文書属性に独自スコアを入れる
リンク関係やコンテンツのパターンで
スコア決定。ソートに利用
- 今後は、カテゴリ検索もRastで
属性の前方一致検索を利用

iYappo (Before)

DB重すぎ



YappoLabs (After)



新たな問題

Rastの問題点

- index速度が気になる
 - sync処理で引っかかり
- 検索とindexが同時実行できない
 - Namazu手法やXML-RPCで回避

index速度改善案

- BDB自体のチューニング
- BDBをQDBMと入れ替えてみる
- indexを分割してXML-RPCする
もっとも正しい手法？めんどい

これ有ったら嬉しい

ちょっとした要望

- 洗練されたコンポーネント化
機能別に取り替えしてみたい
- プラグイン機能
- High Availabilityの選択権
- 開発ポリシーを知っておきたい

今後の取り組み

iYappoでのやりたい事

- Blog Searchとの密接な連携
- Folksonomyを携帯ユーザに
- Personal Search

何故？

- 話題性のある内容を的確に提供
- 寂しがりやにFolksonomy
- キーワード検索の概念を捨てる

結論

検索窓とかは不要
(超高度な検索システム必須)

アクセスしただけでの的確な
情報を提供すべき。

ありがとう
ございました