

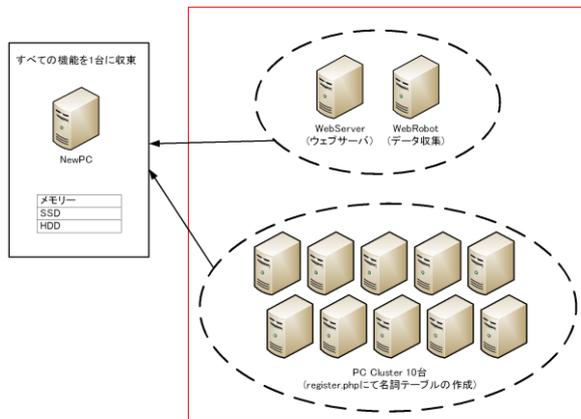
1.目的

現代において高速で情報を検索できるようになった背景にはハードウェアの低価格、高性能化による恩恵がある。本研究は最新のハードウェア環境に過去に作成された検索エンジンを適応させ高速化、検証、新機能の追加をすることを目標とした。

前年度まで研究が行われていた複数台のPCを利用した旧システムは環境構築が非常に難しく、さらに保守管理や機能拡張において様々な問題があったため旧システムを元に再構築し、環境構築方法をマニュアル化してシステム全体の簡素化を行った。

2.ハードウェア構成

旧システムはPCclusterを利用して複数台のPCに処理を分担し高速化されていた(右赤枠内)。



[新システム構成図]

新システムは簡素化のため最新のハードウェアを利用し、すべての動作を1台のPC上で行わせることにした(左図)。ハードウェアを選定しCore i7-2600 3.4GHz、メモリ(DDR3 1333 16GB)、SSD(Crucial m4 64GB)を導入。ボトルネックとなる部分を無くするために1台のPCの中にあるストレージのベンチマークを計測し、実用性を考えながら処理に適した場所にデータベースやプログラムを配置し実行した。

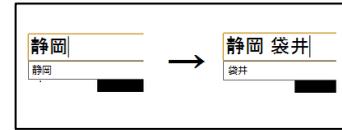
3.ソフトウェアシステムの再構築

開発環境は今後研究室でのデモや成果物として環境構築がしやすいようWindows上での動作を基本とし、XAMPPという開発環境上での安定動作を目標とした。

開発言語のPHPは、バージョン6以降無くなる予定の関数を利用している場合エラーとして警告するようになった。これまでバージョン5を基準として構築されていた旧システム内では数多くのエラーが発生してしまっ。研究室の成果物として展示できる形として残すことも目標としており、今後のアップデートで完全に動かなくなってしまうという事態を避けなければならない。そのため、再構築上で発生したエラーと改善点を文章としてまとめた。

4.追加された新機能

「autocomplete_db」



これは入力した文字と似ているワードを自動的に検索時に表示する機能である。

「インスタントプレビュー」



検索で表示されたURLリンクを画像としてプレビュー表示する。これによってリンク先の情報が有益か視覚的に確認できるようになった。

5.検索速度比較のためのベンチマークテスト

検索エンジンを動作させるにあたってどのストレージ上で動作させればバランスよくシステムを動作させることができるのか、ベンチマークソフトを利用して洗い出した。結果、SSD上ですべてのプログラムを動作させることとし、「北海道 観光」と検索したところ以下の結果となった。

旧システム	新システム
登録 URL 数 : 1000 件	登録 URL 数 : 1000 件
ヒット数 : 記述なし	ヒット数 : 10 件
検索時間 : 1.16 秒 (1.160sec)	検索時間 : 0.77 秒 (0.775sec)

検証結果を見るとPCを10台利用した旧システムに比べて、改良を加えた新システムは高速に応答していることがベンチマーク結果からわかる。

1秒未満で結果を応答することができているため、ユーザーにとっては待ち時間を意識すること無く利用可能になっていると言える。複数台のPCを使用して高速化した旧システムと比較して、1台のPCでそれを上回る応答速度を出すことができている、十分に実用可能と考える。