

赤峯享* 河合英紀* 小西弘一** 菊地賢太郎** 福島俊一*

NEC ヒューマンメディア研究所* NEC C&C メディア研究所**

1. はじめに

インターネット/イントラネットの普及により、一般ユーザがアクセス可能な電子化テキストが飛躍的に増加している。特に、WEB ページは、日本国内に限っても数千万ページに達し、さらに急激な増加を続けている。このテキストの激増により、ユーザが必要なテキストを見つけるための手段として、全文検索エンジンは必要不可欠なものになっている。

大規模テキストの検索を 1 台のマシンで実現しようとした場合、大容量のディスクとメモリを搭載した高価な大型マシンが必要[5]となり、さらに、検索対象テキストの増加に伴って、検索レスポンスが悪化するという問題が生じる。この問題を解決するために、テキストを複数の断片に分割し、PC 等の安価なマシンで並列検索する方式が有望視されている[2,3,4]。

筆者らは、大規模テキストを高速に検索するために、PC クラスタを用いた並列検索エンジンを構築した。本エンジンは、PC クラスタを用いることで、テキスト量の増加に対してスケラブルな検索が可能であり、さらに、各 PC の検索性能の違いや障害時に対応するために、冗長インデックスを用いた負荷分散方式を導入している。まず、本稿で、基本方式とスケラビリティに対する評価結果を述べ、次に、[1]で冗長インデックスを用いた負荷分散方式について述べる。

2. 設計方針

日本国内の WEB ページを対象とした実用的な検索エンジンには、最低限、数百万～数千万ページのテキストを 1 秒間に数十回検索できる性能が要求される。このため、1 台のマシンで実用的な検索性能

を保って全ての検索を行なうことは、もはや現実的ではなく、複数のマシンを用いて並列検索する並列検索エンジンが必要である。筆者らは、並列検索エンジンを以下の方針で設計した。(1)テキスト量の増加に対して、一定の検索性能を維持する[スケラビリティ]、(2)高価な特殊なマシンを使わずに、安価で性能の高いエンジンを構築する[価格性能比]、(3)システム中の 1 台のマシンが故障した場合でも、検索サービスを維持する[耐障害性]、(4)マシン毎の検索性能の違いによって生じる負荷のアンバランス吸収する[負荷分散]。

3. PC クラスタを用いた並列検索エンジン

スケラビリティと価格性能比に優れたアーキテクチャとして、PC クラスタが注目されている。PC クラスタは、複数の PC を高速ネットワークで接続した並列マシンであり、各 PC が協調して処理を行なうことにより仮想的な単一マシンとして動作する。

筆者らは並列検索エンジンの基盤として PC クラスタを用い、検索対象テキストを分割して、各 PC で部分テキストを検索する方式で並列検索エンジンを構築した。さらに、方針(3)と方針(4)に対応するために、冗長インデックスによる負荷分散方式を取り入れた[1]。

4. システム構成

図 1 に、PC クラスタを用いた並列検索エンジンの概要図を示す。並列検索エンジンは、ユーザからのクエリを受け付ける受付ノードの PC、クエリとマッチするテキストを検索する検索ノードの PC、PC 間を接続する高速な SAN(System Area Network)からなる。また、各検索ノードのローカルディスクには、検索対象テキストを分割して作成した部分インデックスを予め格納しておく。検索は、以下の手順で行う。(1)受付ノードは入力クエリを全ての検索ノードに送る。(2)個々の検索ノードは、部分イン

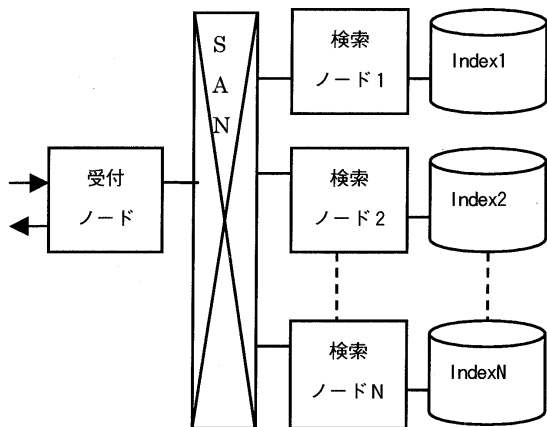


図1 並列検索エンジン概要図

デックスに対して検索を行い、検索結果を受付ノードに返す。(3) 受付ノードは全ての検索結果をマージして出力する。これにより、1 セットの PC クラスタを用いて、仮想的な 1 台のマシンで検索を行なうことができる。

5. 検索性能の評価

5.1 評価システム

4 章のシステム構成で、実際に並列検索エンジンを構築し、検索性能を評価した。各ノードは、512M バイトのメモリと PentiumII(450Mhz)の 1CPU を搭載した PC で構成した。PC の OS は LINUX を用い、SAN は Myrinet [6] を用いた。また、全文検索用のインデックスは、キー情報として n 文字 (n=1~5) のキーをもち、位置情報としてテキスト内の出現位置をもたずにテキスト ID のみをもつタイプを用いた。

5.2 評価方法

スケーラビリティの評価を行なうために、検索ノードを 1 台から 19 台まで順に増やし、スループットと応答時間を測定した。テキストは、検索ノード 1 台当たり 43 万ページ (テキスト 800M バイト、インデックス 1.1G バイト)、19 台構成の場合は全部で 817 万ページ (テキスト 15.2G バイト、インデックス 20.9G バイト) を持たせた。なお、クエリは実際の WEB 検索システム NETPLAZA [7] の検索ログを使用し、ディスクキャッシュの影響を考慮して 5 万クエリ程度検索した後のスループットが安定した 1 万クエリを性能評価の対象とした。

表1 スケーラビリティの評価結果

台数	テキスト量 [ページ]	スループット [クエリ/秒]	平均応答時間 [ms]	最大応答時間 [ms]
1	43 万	95.2	200	481
2	86 万	94.5	200	492
4	172 万	93.3	204	570
8	344 万	92.6	205	491
12	516 万	92.0	206	483
16	688 万	91.1	208	490
19	817 万	89.4	212	489

5.3 評価結果

検索ノードを 1 台から 19 台まで増やした場合のスループットと応答時間を表 1 に示す。スループット及び応答時間も 6% 程度の性能劣化はあるものの、検索ノードの台数 (= 検索対象ページ数) に依存せずに、スケーラブルな検索ができています。また、スループットは約 90 クエリ/秒、応答時間は平均で約 210ms、最大で約 500ms であり、実用的な性能を達成している。

6. おわりに

本稿では、PC クラスタを用いた並列検索エンジンを提案し、基本方式の評価結果を報告した。19 台構成の PC クラスタで、817 万ページの Web ページを対象とした検索でも、検索性能は飽和しておらず、高速でスケーラビリティに優れた検索が可能であることを示した。今後、本エンジンを基盤として、運用系を加え、実用レベルの検索システムを構築する予定である。

参考文献

- [1] 河合ほか、PC クラスタを用いた並列全文検索エンジン—負荷分散方式—、情処第 60 回全国大会、2000 年
- [2] 小西ほか、冗長分散格納による PC クラスタ上の動的負荷均衡化、SwoPP、1999 年
- [3] 森ほか、大規模並列検索エンジンの実装と高度化の検討、情処第 59 回全国大会、1999 年
- [4] FAST、<http://www.alltheweb.com/>
- [5] AltaVista、<http://www.altavista.com/>
- [6] Myrinet、<http://www.mri.com/>
- [7] NETPLAZA、<http://netplaza.biglobe.ne.jp/>