

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4581962号
(P4581962)

(45) 発行日 平成22年11月17日(2010.11.17)

(24) 登録日 平成22年9月10日(2010.9.10)

(51) Int.Cl. F I
G06F 17/30 (2006.01) G O 6 F 17/30 4 1 4 B
G06F 12/00 (2006.01) G O 6 F 17/30 1 1 0 C
 G O 6 F 12/00 5 2 0 A

請求項の数 13 (全 15 頁)

(21) 出願番号	特願2005-312138 (P2005-312138)	(73) 特許権者	000005108
(22) 出願日	平成17年10月27日(2005.10.27)		株式会社日立製作所
(65) 公開番号	特開2007-122302 (P2007-122302A)		東京都千代田区丸の内一丁目6番6号
(43) 公開日	平成19年5月17日(2007.5.17)	(74) 代理人	100100310
審査請求日	平成19年3月15日(2007.3.15)		弁理士 井上 学
		(72) 発明者	八高 克志
			神奈川県横浜市戸塚区戸塚町5030番地
			株式会社日立製作所 ソフトウェア事業部内
		(72) 発明者	原 憲宏
			神奈川県横浜市戸塚区戸塚町5030番地
			株式会社日立製作所 ソフトウェア事業部内

最終頁に続く

(54) 【発明の名称】 情報検索システムとインデクス管理方法およびプログラム

(57) 【特許請求の範囲】

【請求項1】

情報検索要求を受付けるフロントエンドノードと、複数の情報項目で構成されえる情報をn個の情報項目群に分割し、該n個の情報項目群を検索対象範囲としてそれぞれに割当てたn個の検索ノードを有し、

該検索ノードは各検索対象に対応したインデクスを記憶手段に格納した情報検索システムであって、

n個の前記検索ノードで並列に前記インデクスを用いた情報検索を行なう情報検索システムにおけるインデクス管理方法において、

前記インデクスはインデクスキーと、該インデクスキーに対応する前記情報項目へのアドレスのリストである転置ファイルを有し、

前記フロントエンドノードは、

該情報検索システムの検索対象に前記情報項目を登録する際に

前記情報項目から前記インデクスキーを抽出し、前記インデクスキーを前記情報項目へのアドレスと対にすることでインデクス情報のエントリを作成し、前記情報項目毎のインデクス情報エントリを集積してインデクス情報を作成するステップと、

前記インデクス情報を前記インデクス情報エントリの部分である分割キーによってm(m>n)個のバケットに論理的に分割し、前記インデクス情報に対して前記バケットに対応した前処理を行って前処理済みインデクス情報を作成するステップと、

前処理済みインデクス情報を前処理済みインデクス情報格納部に格納するステップと、

前記ノードの検索対象範囲割当変更の際に、前記バケットのノード割当を変更するステップとを有し、

前記検索ノードは、

前記ノード割当変更によって前記ノードに割当てられた一つないし複数のバケットに該当する格納した前処理済みインデクス情報を用いてインデクスを作成するステップを有することを特徴とするインデクス管理方法。

【請求項2】

請求項1記載のインデクス管理方法であって、

前記インデクス情報に対して前記バケットに対応した前処理を行って前処理済みインデクス情報を作成するステップにおいて、

前記バケットに分割されたインデクス情報に対してインデクスキーとインデクスキーが合致する情報へのアドレスのリストであるアドレスリストを作成し、前記バケット内の複数のインデクスキーとアドレスリストの対からなる部分転置ファイルを作成することでインデクス前処理を行なうことを特徴とするインデクス管理方法。

【請求項3】

請求項1記載のインデクス管理方法において、

前記フロントエンドノードは、

検索ノードを追加する際に、新規に追加した検索ノードに検索担当範囲として割当てるバケットを決定するステップと、

前記新規検索ノードに割当てられたバケットに該当する一つないし複数の前処理済みインデクス情報を前記新規検索ノードに転送するステップとを有し、

前記新規検索ノードは、

転送された一つないし複数の前処理済みインデクス情報の部分転置ファイルにおける同一インデクスキーのアドレスリストを併合するステップと、

前記新規検索ノード内の複数のインデクスキーと前記併合したアドレスリストの複数の対からなる転置ファイルを作成するステップとを有することを特徴とするインデクス管理方法。

【請求項4】

請求項1記載のインデクス管理方法において、

前記フロントエンドノードは、

検索対象範囲割当変更にともない特定の前記検索ノードの検索対象割当範囲を拡大する際に、該当検索ノードに追加で割当てられるバケットを決定するステップと、

該当検索ノードに新規に割当てられたバケットに該当する一つないし複数の前記前処理済みインデクス情報を該当検索ノードに転送するステップとを有し、

前記検索ノードは、

転送された一つないし複数の前処理済みインデクス情報の部分転置ファイルにおける同一インデクスキーのアドレスリストを前記検索ノードに既存のインデクスの転置ファイルにおける同一インデクスキーのアドレスリストに合併するステップとを有することを特徴とするインデクス管理方法。

【請求項5】

請求項1記載のインデクス管理方法において、

前記フロントエンドノードは、

前記前処理インデクス作成時にバケットを識別するバケットの番号をアドレスと対にしてアドレスリストに格納するステップと、

検索対象割当変更に伴い特定の前記検索ノードの検索対象割当範囲を一部削減する際に、該当検索ノードから割当てが解除されるバケットを決定するステップとを有し、

前記検索ノードは、

インデクスの転置ファイル内のアドレスリストをサーチし、割当て解除対象の前記バケット番号が付与されたアドレスをアドレスリストから削除するステップとを有することを特徴とするインデクス管理方法。

10

20

30

40

50

【請求項6】

請求項1記載のインデクス管理方法において、

前記フロントエンドノードは、

インデクス作成時にバケットを識別するバケットの番号をアドレスと対にしてアドレスリストに格納するステップと、

検索対象割当変更に伴い特定の前記検索ノードの検索対象割当範囲を一部削減する際に、該当検索ノードから割当てが解除されるバケットを決定するステップとを有し、

前記検索ノードは、

インデクスの転置ファイル内のアドレスリストをサーチし、割当て解除対象の前記バケット番号が付与されたアドレスをアドレスリストから削除するステップでインデクス更新するとともに、割当てが解除される前記バケットを削除対象バケット番号一覧に記憶するステップと、前記情報検索要求に対してインデクスの転置ファイルを参照し、得られたアドレスとバケット番号の一覧から削除対象バケット番号一覧を参照して削除対象のバケットのアドレスを除外するステップを有し、論理的に検索対象範囲割当てを一部削除が完了した状態で情報検索処理を実行させながら、インデクス更新を行なうことを特徴とするインデクス管理方法。

10

【請求項7】

情報検索要求を受付けるフロントエンドノードと、複数の情報項目で構成される情報をn個の情報項目群に分割し、該n個の情報項目群を検索対象範囲としてそれぞれに割当てたn個の検索ノードを有し、該検索ノードは各検索対象に対応したインデクスを記憶手段に格納した情報検索システムであって、

20

n個の前記検索ノードで並列にインデクスを用いた情報検索を行なう情報検索システムにおいて、

前記インデクスはインデクスキーと、インデクスキーに対応する情報項目へのアドレスのリストである転置ファイルを有し、

前記フロントエンドノードは、

該情報検索システムの検索対象に情報項目を登録する際に

前記情報項目から前記インデクスキーを抽出し、前記インデクスキーを前記情報項目へのアドレスと対にすることでインデクス情報のエントリを作成し、前記情報項目毎のインデクス情報エントリを集積してインデクス情報を作成する手段と、

30

前記インデクス情報を前記インデクス情報エントリの部分である分割キーによってm(m>n)個のバケットに論理的に分割する手段と、

前記インデクス情報に対してバケットに対応した前処理を行って前処理済みインデクス情報を作成する手段を有するインデクス前処理部を備え、

前処理済みインデクス情報を前処理済みインデクス情報格納部に格納し、前記ノードの検索対象範囲割当変更の際に、前記バケットのノード割当てを変更する構成管理の手段とを有し、

前記検索ノードは、

前記割当て変更によって前記ノードに割当てられた一つないし複数のバケットに該当する前処理済みインデクス情報を用いてインデクスを作成するインデクス作成手段を有することを特徴とする情報検索システム。

40

【請求項8】

請求項7記載の情報検索システムにおいて、

前記インデクス情報に対して前記バケットに対応した前処理を行って前処理済みインデクス情報を作成する手段は、

バケットに分割されたインデクス情報に対してインデクスキーとインデクスキーに合致する情報へのアドレスのリストであるアドレスリストを作成し、バケット内の複数のインデクスキーとアドレスリストの対からなる部分転置ファイルを作成してインデクス前処理を行なうことを有することを特徴とする情報検索システム。

【請求項9】

50

請求項7記載の情報検索システムにおいて、

前記フロントエンドサーバは、

検索ノードを追加する際に、新規に追加した検索ノードに検索担当範囲として割当てするバケットを決定する手段と、

前記新規検索ノードに割当てられたバケットに該当する一つないし複数の前処理済みインデクス情報を前記新規検索ノードに転送する手段とを有し、

前記新規検索ノードは、

転送された一つないし複数の前処理済みインデクス情報の部分転置ファイルにおける同一インデクスキーのアドレスリストを併合する手段と、

前記新規検索ノード内の複数のインデクスキーと前記併合したアドレスリストの複数の対からなる転置ファイルを作成してインデクスを作成する手段を有する

ことを特徴とする情報検索システム。

【請求項10】

請求項7記載の情報検索システムにおいて、

前記フロントエンドノードは、

検索対象範囲割当て変更にもとない特定の前記検索ノードの検索対象割当て範囲を拡大する際に、該当検索ノードに追加で割当てられるバケットを決定する手段と、

該当検索ノードに新規に割当てられたバケットに該当する一つないし複数の前記前処理済みインデクス情報を該当検索ノードに転送する手段とを有し、

前記検索ノードは、

転送された一つないし複数の前処理済みインデクス情報の部分転置ファイルにおける同一インデクスキーのアドレスリストを前記検索ノードに既存のインデクスの転置ファイルにおける同一インデクスキーのアドレスリストに併合してインデクス更新する手段を有することを特徴とする情報検索システム。

【請求項11】

請求項7記載の情報検索システムにおいて、

前記フロントエンドノードは、

前記前処理済みインデクス作成時にバケットを識別するバケットの番号をアドレスと対にしてアドレスリストに格納する手段と、

検索対象割当て変更に伴い特定の前記検索ノードの検索対象割当て範囲を一部削減する際に、該当検索ノードから割当てが解除されるバケットを決定する手段とを有し、

前記検索ノードは、

インデクスの転置ファイル内のアドレスリストをサーチし、割当て解除対象の前記バケット番号が付与されたアドレスをアドレスリストから削除してインデクス更新する手段を有することを特徴とする情報検索システム。

【請求項12】

請求項7記載の情報検索システムにおいて、

前記フロントエンドノードは、

インデクス作成時にバケットを識別するバケットの番号をアドレスと対にしてアドレスリストに格納する手段と、

検索対象割当て変更に伴い特定の前記検索ノードの検索対象割当て範囲を一部削減する際に、該当検索ノードから割当てが解除されるバケットを決定する手段とを有し、

前記検索ノードは、

インデクスの転置ファイル内のアドレスリストをサーチし、割当て解除対象の前記バケット番号が付与されたアドレスをアドレスリストから削除してインデクス更新するとともに、割当てが解除される前記バケットを削除対象バケット番号一覧に記憶し、前記情報検索要求に対してインデクスの転置ファイルを参照し、得られたアドレスと前記バケット番号の一覧から削除対象バケット番号一覧を参照して削除対象のバケットのアドレスを除外する手段を有し、論理的に検索対象範囲割当てを一部削除が完了した状態で情報検索処理を実行させながら、インデクス更新を行なうことを特徴とする情報検索システム。

10

20

30

40

50

【請求項 13】

情報検索要求を受付けるフロントエンドノードと、複数の情報項目で構成されえる情報をn個の情報項目群に分割し、該n個の情報項目群を検索対象範囲としてそれぞれに割当てたn個の検索ノードを有し、該検索ノードは各検索対象に対応したインデクスを記憶手段に格納した情報検索システムであって、

n個の前記検索ノードで並列にインデクスを用いた情報検索を行なう前記情報検索システムを実行させる情報検索プログラムにおいて、

前記インデクスはインデクスキーと、インデクスキーに対応する情報項目へのアドレスのリストである転置ファイルを有しており、

前記情報検索システムの検索対象に情報項目を登録する際に、

前記フロントエンドノードに、

前記情報項目から前記インデクスキーを抽出し、前記インデクスキーを前記情報項目へのアドレスと対にすることでインデクス情報のエントリを作成し、前記情報項目毎のインデクス情報を集積してインデクス情報を作成する機能と、

前記インデクス情報を前記インデクス情報のエントリの部分である分割キーによってm(m>n)個のバケットに論理的に分割する機能と、

前記インデクス情報に対してバケットに対応した前処理を行って前処理済みインデクス情報を作成し、前処理済みインデクス情報を格納する機能と、

前記検索ノードの検索対象範囲割当変更の際に、前記バケットのノード割当を変更する機能を実行させ、

前記検索ノードに、

前記割当変更によって前記ノードに割当てられた一つないし複数のバケットに該当する前処理済みインデクス情報を用いてインデクスを作成する機能を実行させる

ための情報検索プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、インデクスを用いた情報検索システムに係わり、特に、インデクスを分割して複数のノードで並列して検索を行なう検索システムの検索ノードの追加、削除または検索ノード間の処理量の変更に関する。

【背景技術】

【0002】

データベース管理システム(以下、DBMS)など情報検索を行なう情報処理システムの処理性能を向上する方法として、非特許文献1にデータベース処理負荷を複数のプロセッサに分散させ実行するアーキテクチャに関する技術が開示されている。上記従来技術の記載の Shared Everything, Shared Disk(共有型)では検索を行なう全てのノードまたはプロセッサから全てのディスクにアクセス可能であるが、Shared Nothing(無共有型)アーキテクチャでは各ノードに独立して属するディスクにのみアクセス可能である。無共有型アーキテクチャは共有型アーキテクチャに比べてアクセスが競合するリソースが少なく、スケーラビリティの点で優れている。

【0003】

無共有型アーキテクチャの情報システムではノードの追加、ノードの削除、または特定のノードへのアクセス集中により、それぞれのノードにおける処理量の変更が必要となった場合、それぞれが処理するデータ配分を変更する必要がある。ノード間のデータ配分を変更する最も単純な方法としては、データベースの内容を一旦バックアップし、データの配置定義を変更した後にバックアップしたデータを再ロードするものである。ただし、この方法では扱うデータ量が大规模となった場合、バックアップ、再ロードにかかる処理時間が膨大になるという問題がある。

【0004】

この問題に対して、特許文献1に、データをハッシュ関数等で予め複数のバケットに分

10

20

30

40

50

割しいくつかのバケットをプロセッサに対応させて管理する技術が開示されている。

【0005】

また、特許文献2ではデータを予めバケットに分割し、バケットを複数のディスクと対応表で管理し、検索ノードの追加等にもなうディスクの追加の際に移動データ量が最小になるようにバケットとディスクの対応を変更してデータを再配置する技術が開示されている。

【0006】

また、特許文献3ではデータを予めハッシュ関数により論理的にバケットに相当する単位に分割して複数のディスクと対応させて管理し、ディスク追加の際にバケットに相当する単位でデータの再配置中を行なうとともに、データの再配置中に検索、更新、挿入などの処理を実行する技術が開示されている。

【0007】

また、特許文献4ではバケットに相当する物理ディスクと仮想ディスクとのマッピングを変更することで、物理的なデータ移動なくプロセッサ間のデータ配分を変更する技術が開示されている。これにより、プロセッサ間のデータ配分変更の時間を飛躍的に向上し、ノード負荷の上昇に追従して動的にノード台数を増やすことが可能となっている。

【0008】

これらの技術はデータ一般に関する技術であり、転置ファイルからなるインデクスに対する特別な考慮はない。

【0009】

【特許文献1】米特許4,412,285

【特許文献2】特開2001-142752号公報

【特許文献3】特開2003-6021号公報

【特許文献4】特開2005-56077号公報

【非特許文献1】"Parallel Database Systems: The Future of High Performance Database Systems", COMMUNICATIONS OF THE ACM, Vol.35, NO6, 1992, P.85-P.98

【発明の開示】

【発明が解決しようとする課題】

【0010】

無共有型のインデクスを持つ情報検索システムにおいてノードの追加、削除およびノード間の負荷バランス変更のために各ノードのインデクスの検索対象範囲割当てを変更する必要がある。この要求を実現する基本的なアイデアは、インデクスがない場合の一般のデータと同様に予めバケットに分割しておき、各ノードに割当てする検索対象範囲割当てをバケット単位で変更する。

【0011】

ここで、情報処理システムのノード数の柔軟な変更や、ノードの負荷バランスの柔軟な変更のためには、データの配置の最小単位であるバケットのデータサイズが各ノードに割当てられるデータ量に比べて十分小さい必要がある。バケットサイズを小さくすると各ノードあたりのバケット数は必然的に多くなる傾向がある。

【0012】

また、情報の検索を高速化するために用いられる一般的なインデクスは検索で用いられるインデクスキーとそのインデクスキーに合致する情報項目のアドレスのリストからなる転置ファイルからなっている。

【0013】

ここで、もっとも単純にインデクスを再配分する方法として、バケット単位にインデクスを作成し、バケット単位のインデクスを再配置することが考えられる。しかし、先に述べた通りバケットは各ノードに多数存在するため、バケット単位にインデクスを作成した場合、各ノードにおいて多数のインデクスが存在することになり、情報検索時に同一検索キーに対して多数回のインデクス参照が発生する。多数のインデクスの参照は転置ファイルから目的のインデクスキーを探索する処理が多数回発生する点、目的のインデクスキー

10

20

30

40

50

に対応するアドレスリストがバケット毎に多数個に分割されているため、アドレスリストの一括読み込みが出来ない点でノード単位の大きなインデクスを1回参照する場合に比べて非効率である。検索性能の劣化は情報検索を主たる業務とする情報処理システムにとって重大な問題である。

【0014】

一方で、検索性能の劣化問題に対し、ノード間の検索対象範囲割当て変更の際に情報項目の原本からノード単位のインデクスを再作成することが考えられるが、インデクスの作成は転置ファイルにおけるインデクスキー毎にアドレスリストを作成するためのインデクスキーの比較が多数回発生することから計算量が多い。また、この他にも情報項目の原本を解析してインデクスキーとなる部分を抽出するなど計算量の多い処理がある。検索対象範囲割当て変更の要因がノードの負荷上昇によるノード追加等である場合、インデクス作成のために更にノード負荷を上昇させることとなるので、情報項目の原本からインデクスを再作成するのは適切でない。

10

【0015】

本発明は、情報検索効率を考えたノード単位のインデクスをもつ検索装置の検索対象範囲割当てを高速に変更することを目的とする。

【課題を解決するための手段】

【0016】

本発明では、 n 個のノードに割当てて検索対象情報項目を登録する際に、情報項目からインデクスキーおよび情報項目のアドレスの対の集合であるインデクス情報を抽出し、インデクス情報を m ($m > n$) 個のバケットに分割し、バケット内に閉じた部分転置ファイルを作成しておく。ここで、 m 、 n は1以上の整数とする。

20

【0017】

ノードの検索対象範囲割当ての変更の際には、各ノードに対するバケットの割当てを変更し、バケット単位の部分転置ファイルおよび既存インデクスにおける転置ファイルを併合することで新たなインデクスを作成することで、高速なインデクスの作成、更新を実現する。

【発明の効果】

【0018】

検索対象情報項目の登録時にインデクスキーと検索対象情報項目のアドレスの抽出を予め行っておくことで、インデクスの検索対象範囲割当て変更時にこの作業を不要とし、インデクス作成を高速化できる。

30

【0019】

また、バケット単位の部分転置ファイルを作成しておくことで、転置ファイルにおけるインデクスキーを情報項目毎に探索することなく、部分転置ファイルのアドレスリスト毎に探索し、併合することによって最終的なインデクスを作成でき、インデクスの作成を高速化できる。

【発明を実施するための最良の形態】

【0020】

以下発明の一実施例を、添付図面を用いて具体的に説明する。

40

【0021】

図1は実施形態の情報検索システムのハードウェアおよびソフトウェアの構成図である。クライアント100は本発明の実施例における情報検索システムに検索要求を発行するコンポーネントである。本発明の実施例における処理検索システムはフロントエンドノード110と複数の検索ノード160a、160b、160cおよび160dとこれらを結合する通信路150から成る。フロントエンド部110はクライアントからの検索要求を受け付ける受付部111、検索対象の情報項目の格納および取出し操作を行なう原本管理部121、情報検索対象の情報項目を保持する原本格納部122、検索対象登録時に情報項目の原本とその位置情報であるアドレスからインデクス作成のための前処理を行なうインデクス前処理部131、前処理済みインデクス情報を格納する前処理済みインデクス情報格納部132、検索ノードの数および配

50

置、前処理済みインデクス情報のバケット分割情報およびバケットの検索ノード割当て情報からなる構成情報を管理する構成管理部141および構成情報を保持する構成情報格納部142から構成される。検索ノード160a~160dは、インデクスを検索するインデクス検索部161a~161d、インデクスを作成するインデクス作成部162a~162d、インデクスを格納するインデクス格納部163a~163dから構成される。

【0022】

本発明の実施例では複数のフロントエンド110、受付部111は一つであるが、問合せ処理の負荷分散のために複数にすることも考えられる。同様に、本発明の実施例では原本管理部121、原本格納部122も格納する情報項目のデータ量、情報項目原本へのアクセス負荷に応じて複数にすることも考えられる。また、本発明の実施例では原本管理部121、原本格納部122はフロントエンドノード100内にあるが、フロントエンドノード100と独立したノードとすることも可能であり、また情報検索システムの外部コンポーネントであっても良い。さらに、本発明の実施例ではインデクス検索部161a~161dとインデクス作成部162a~162dはインデクス格納部163a~163dとともに組になり検索ノード160a~160dに存在するが、インデクス格納部163a~163dを共有型の外部記憶装置として検索ノード160a~160dから独立させた上で、インデクス検索部161a~161dとインデクス作成部162a~162dをそれぞれ独立したノードに配置することも可能である。これにより、情報検索処理に対するインデクス更新処理による負荷の影響を最小限にすることが出来る。

【0023】

次に、図1~図7を用いて検索対象である情報項目登録時の情報検索システムの動作を示す。情報項目登録の説明の際には検索ノードは160a~160cの3台(n=3)と仮定する。

【0024】

クライアント100から検索対象項目登録要求は(ステップ700)登録する情報項目とともにフロントエンドノード110の受付部111にて受信される(ステップ701)。次に、原本管理部121が原本格納部122に受付部111が受信した情報項目を格納する(ステップ702)。インデクス前処理部131は情報項目とその格納アドレスから図2に示すインデクス情報エントリを抽出する(ステップ703)。次に、インデクス前処理部131はインデクス情報エントリの分割キーに対してハッシュ関数を作用することでインデクス情報エントリを12(m=12)個のバケットに分類する(ステップ704)。ここで、本実施例では分割キーは情報項目のアドレスと仮定するが、情報項目から抽出したインデクスキーとは異なる値であっても良いし、インデクスキーでも良い。また、本発明の実施例ではバケット数を12としたが、一般的にはノード分割を柔軟にするために、より大きいバケット数、例えば256バケットとするのが良い。図3が0~11の各バケット番号320をもつ12個のバケット300~311に0~47の分割キーを持つ情報項目を分類した様子を論理的に示したものである。次に、インデクス前処理部131は各バケット毎に作成されている前処理済みインデクス情報132の部分転置ファイル520から、登録情報項目のインデクスキーと一致する部分転置ファイルエントリをサーチする(ステップ705)。ここで、前処理済みインデクス情報の構造を図5に示す。前処理済みインデクス情報500は情報項目から抽出したインデクスキー530とインデクスキー毎に作成された情報項目へのアドレス一覧であるアドレスリスト540の集合である部分転置ファイル520からなる。アドレスリストの各エントリは本発明の実施例ではアドレス541とバケット番号521の対から成り、これをアドレスリストエントリと呼ぶ。さらに前処理済みインデクス情報500はオプションとして、部分転置ファイル520に情報項目を登録する際や最終的な検索ノード160単位に分割したインデクスを作成する際に転置ファイル520から目的のインデクスキーの探索を高速化する目的の補助情報としてのインデクスキー高速探索用構造化データ510を持つことが出来る。前処理済みインデクス情報132の部分転置ファイル520に登録情報項目のインデクスキーと一致する部分転置ファイルエントリの有無を判断し(ステップ706)、部分転置ファイルエントリが無い場合は部分転置ファイル520に部分転置ファイルエントリを追加する(ステップ707)。次に部分転置ファイルの該当エントリのアドレスリスト540に登録情報項目へのアドレスを追加する(ステップ708)。部分転置ファイル520に登録情報項目のインデクスキーと一致する部分転置ファイルエントリがある場合は

10

20

30

40

50

ステップ707をスキップしてステップ708に進む。次に、前処理済みインデクス情報500のインデクスキー高速探索用構造化データ510を更新する(ステップ709)。次に、構成管理部141の構成管理情報におけるバケット検索ノード対応表を参照し、登録情報エントリに対応する検索ノード160a、160bまたは160cを決定する(ステップ710)。バケット検索ノード対応表は図4に示す通り、バケット番号を表す部分420とそのバケットを担当する検索ノード番号430の対がバケット毎に並んでいる400~411。決定した検索ノードにインデクス情報エントリとバケット番号を転送し(ステップ711)、フロントエンドノードにおけるクライアントからのデータ挿入処理を終了する。なお、ステップ710、ステップ711はステップ704より後ろであれば、より前段階で行なうことができる。

【0025】

フロントエンドノードでのクライアントからの情報項目登録手順と平行して、ステップ710で決定した検索ノードではステップ720からステップ726にしたがってインデクスの更新を行なう。まず、ステップ711でフロントエンドノードから転送されたインデクス情報エントリ、バケット番号を受信し、登録情報項目のインデクスキーでインデクス600の転置ファイル620から転置ファイルエントリ630を探索する(ステップ721)。ここで、インデクスの構造を図6に示す。インデクス600は情報項目から抽出したインデクスキー630とインデクスキー毎に作成された情報項目へのアドレス一覧であるアドレスリスト640の集合である転置ファイル620からなる。アドレスリストの各エントリは本発明の実施例ではアドレス641とバケット番号621の対から成り、これをアドレスリストエントリと呼ぶ。さらにインデクス600はオプションとして、転置ファイル620に情報項目を登録または参照する際に転置ファイル620から目的の転置ファイルエントリの探索を高速化するためにインデクスキー高速探索用構造化データ610を持つのが一般的である。インデクス600の転置ファイル620に登録情報項目のインデクスキーと一致する転置ファイルエントリの有無を判断し(ステップ722)、転置ファイルエントリが無い場合は転置ファイル620に転置ファイルエントリを追加する(ステップ723)。次に転置ファイルの該当エントリのアドレスリスト640に登録情報項目へのアドレスを追加する(ステップ724)。転置ファイル620に登録情報項目のインデクスキーと一致する転置ファイルエントリがある場合は、ステップ723をスキップしてステップ724に進む。次に、インデクス620のインデクスキー高速探索用構造化データを更新して(ステップ725)、検索ノードにおけるインデクス更新を終了する(ステップ726)。

【0026】

次に、図1~図6と図8を用いて情報検索時の情報検索システムの動作を説明する。情報検索の説明の際には検索ノードは160a~160cの3台(n=3)と仮定する。

【0027】

クライアント100から検索キーとともに発行された情報検索要求は(ステップ800)、フロントエンドノード110の受付部111にて受信される(ステップ801)。次に、受付部が問合せを解析し検索キーを含む検索手順を生成する(ステップ802)。次に、全ての検索ノード160a、160bおよび160cに対して検索手順を送信してインデクス検索を指示する(ステップ803、804、805)。各検索ノードにインデクス検索指示を出し終わると、受付部111は検索結果としての情報項目原本へのアドレス1件ずつに対してステップ807からステップ811までを繰り返す。検索結果としての各情報項目原本へのアドレスを、検索ノード160a、160bまたは160cから受信すると(ステップ808)、このアドレスをもって原本管理部121が原本格納部122から情報項目原本を取り出し(ステップ809)、検索結果としてクライアントに返却する(ステップ810)。全ての検索ノードにおける全ての検索結果について処理を終えると(ステップ811)、情報検索処理を終了する(ステップ812)。

【0028】

また、検索ノード160a、160bおよび160cでは、並列してステップ820からステップ829の処理を行なう。検索キーを含む検索手順をフロントエンドノードから受け取ると(ステップ821)、インデクス600のインデクスキー高速探索用構造化データ610を参照して転置ファイル620の該当転置ファイルエントリ630を探索する(ステップ823)。検索キーに該当する

10

20

30

40

50

転置ファイルエントリ630が存在した場合は(ステップ824)、該当転置ファイルエントリのアドレスリスト640からアドレスの一覧を取得し(ステップ825)、取得したアドレスを順次受付ノードに返却する(ステップ826、827、828)。取得したアドレスを全て返却すると(ステップ828)、検索ノードでのインデクス検索を終了する(ステップ829)。

【 0 0 2 9 】

次に、図1～図6、図9および図12を用いて、検索ノード追加時のインデクスの再配置における情報検索システムの動作を説明する。検索ノード追加の説明の際には検索ノードは160a～160cの3台(n=3)から160a～160cに160dを加えた4台(n=4)に変更するものと仮定する。

【 0 0 3 0 】

先ず、検索ノード追加の構成変更指示を受けたフロントエンドノード110の構成管理部141は図4に示したバケット検索ノード対応表を参照し、新規追加検索ノード160dに割当てられる検索対象範囲に該当するバケットを決定する(ステップ901)。この際、既存の検索ノード160a、160bおよび160cの検索対象範囲割当ての変更が最小限になるようにする。本発明の実施例における本ノード追加例では検索ノード160a、160bおよび160cからそれぞれバケット番号3、7、11のバケットを新規追加検索ノード160d用に割当てを変更する。この結果、図4に示したバケット検索ノード対応関係は図12に示したバケット検索ノード対応関係に変更される。図12のバケット検索ノード対応関係のデータ構造に基づいて、バケットの検索ノード割り当てを変更し、インデクスを作成する。このバケット再配置方法は特開2001-142752に詳細に記載されている。次に、インデクス前処理部141が新規追加検索ノード160dに割当てられるバケットに該当する前処理済みインデクス情報500を前処理済みインデクス情報格納部142から読み出し新規追加検索ノード160dに転送し、新規追加検索ノード160dにインデクスの作成を指示する(ステップ902)。また、新規追加検索ノード160dに割当てが変更になり各既設検索ノード160a～160cで検索不要となったバケット番号を各既設検索ノード160a～160cに転送し、検索対象範囲の削除を指示する(ステップ903、904、905)。次に、バケット検索ノード対応表の検索ノード番号430をステップ901で決定したバケット検索ノード対応に従って更新し(ステップ906)、フロントエンドノード110におけるインデクスの再配置を終了する(ステップ907)。

【 0 0 3 1 】

新規追加ノード160dにおいては、ステップ902で転送された前処理済みインデクス情報500を全て受信し(ステップ911)、前処理済みインデクス情報500のインデクスキー高速探索用構造化情報510を参照し、それぞれの部分転置ファイルエントリ530をインデクスキー毎にグルーピングする(ステップ912)。次に、インデクスキー毎にグルーピングされた部分転置ファイルエントリ530をマージしてインデクスの転置ファイル620を作成する(ステップ913)。次に、各前処理済み各インデクス情報500のインデクスキー高速探索用構造化データ510をマージして、インデクス600のインデクスキー高速探索用構造化データ610を作成する(ステップ914)。最後に、インデクスキー高速探索用構造化データ610と転置ファイル620からなるインデクス600をインデクス格納部163dに格納して(ステップ915)インデクス作成を終了する(ステップ916)。

【 0 0 3 2 】

既設検索ノードにおいては、ステップ905で割当てが変更になり検索不要となった削除対象バケット番号の一覧を受信する(ステップ921)。次に、各検索ノードのインデクス600のインデクスキー毎(ステップ922)、各インデクスキーのアドレスリストにおけるアドレスリストエントリ毎(ステップ923)にステップ924からステップ926までの処理を行なう。各アドレスリストエントリのバケット番号を参照し(ステップ924)、バケット番号が不要となったバケット番号一覧に存在するか確認し(ステップ925)、不要となったバケット番号一覧に存在する場合、転置ファイルエントリ630のアドレスリスト640におけるアドレスリストエントリを削除する(ステップ926)。アドレスリストエントリのバケット番号が不要となったバケット番号一覧に存在しない場合はステップ927に進む。全てのインデクスキー、アドレスリストエントリに対して処理を終了すると(ステップ928)、既存検索ノード

10

20

30

40

50

ドでの不要検索範囲の削除を終了する(ステップ929)。

【0033】

以上で検索ノード追加の基本手順の説明は完了であるが、以下に既存検索ノードにおける不要となった検索対象範囲の削除と情報検索手順の改良案を開示する。

【0034】

既存検索ノードにおける不要となった検索対象範囲を削除するインデクス更新手順は大きな計算量を必要とする処理であるが、図10を用いて下記に説明するように情報検索手順を一部変更することで、不要となった検索対象範囲を削除するインデクス更新手順を情報検索要求の処理と平行して実行できるようになる。

【0035】

図10のフロントエンドノードにおける情報検索手順のステップ800から812は図8と同様である。図10の検索ノードにおける情報検索手順のステップ820から825は図8と同様である。検索ノードでアドレスリストから取得したアドレス一覧をフロントエンドノード110に返却する際に、インデクス600の転置ファイル620の検索キーと合致する転置ファイルエントリ630におけるアドレスリスト640の各アドレスリストエントリについて(ステップ826)、アドレスリストエントリのバケット番号842が削除対象バケット番号一覧に存在するかチェックし(ステップ1001)、削除対象バケット番号一覧に存在しなければ、受付ノードにアドレスを返却する(ステップ827)。アドレスリストエントリの削除対象バケット番号一覧に存在した場合はステップ828にスキップする。アドレスリストエントリを全て処理すると(ステップ828)、検索ノードでのインデクス検索を終了する(ステップ829)。

【0036】

次に、図1～図6と図11を用いて、検索ノード削除時のインデクスの再配置における情報検索システムの動作を説明する。検索ノード追加の説明の際には検索ノードは160a～160dの4台(n=4)から160dを除いた160a～160cの3台(n=3)に変更するものと仮定する。

【0037】

まず、検索ノード削除の構成変更指示を受けたフロントエンドノード110の構成管理部141はバケット検索ノード対応表を参照し、削除検索ノード160dに割当ててあった検索対象範囲に該当するバケットの再割当て先を決定する(ステップ1101)。この際、残存する検索ノード160a、160bおよび160cの割当て済み検索対象範囲は変更せず、検索ノード160dの検索対象範囲割当ての一部が残存ノードの検索対象範囲割当てに追加になるようにする。本発明の実施例における本ノード削除例では検索ノード160a、160bおよび160cのそれぞれにバケット番号3、7、11のバケットの割当てを変更する。次に、インデクス前処理部141が残存検索ノード160aから160c毎にループして(ステップ1102)、各残存検索ノードに割当てするバケットに該当する前処理済みインデクス情報500を前処理済みインデクス情報格納部142から読み出して各残存検索ノードに転送し、各残存検索ノードにインデクスの更新を指示する(ステップ1103)。全ての残存検索ノードにインデクス更新指示を終えると(ステップ1104)、バケット検索ノード対応表の検索ノード番号430をステップ1101で決定したバケット検索ノード対応に従って更新し(ステップ1105)、フロントエンドノード110におけるインデクスの再配置を終了する(ステップ1106)。

【0038】

残存検索ノードにおいては、ステップ1103で転送された前処理済みインデクス情報500を全て受信し(ステップ1111)、前処理済みインデクス情報500のインデクスキー高速探索用構造化情報510を参照し、それぞれの部分転置ファイルエントリ530をインデクスキー毎にグルーピングする(ステップ1112)。次に、インデクスキー毎にグルーピングされた部分転置ファイルエントリ530を各残存ノードのインデクスにおける転置ファイルにマージしてインデクスの転置ファイル620を作成する(ステップ1113)。次に、各前処理済み各インデクス情報500のインデクスキー高速探索用構造化データ510をインデクス600のインデクスキー高速探索用構造化データ610にマージして更新する(ステップ1114)。最後に、インデクスキー高速探索用構造化データ610と転置ファイル620からなるインデクス600をインデクス格納部163dに格納して(ステップ1115)インデクス更新を終了する(ステップ1116)。

10

20

30

40

50

【図面の簡単な説明】

【0039】

【図1】本発明を実施するための情報検索システムの構成図を示す。

【図2】本発明の実施例における、情報項目から抽出しインデクスの元情報となるインデクス情報エントリのデータ構造を示す。

【図3】本発明の実施例における、インデクス情報をバケット単位に論理的分割した概念図を示す。

【図4】本発明の実施例における、バケットと検索ノードの対応を表すデータ構造と、3検索ノードの場合のバケット配置例を示す。

【図5】本発明の実施例における前処理済みインデクス情報のデータ構造を示す。 10

【図6】本発明の特徴であるインデクス一般における共通格納構造を示す。

【図7】本発明の実施例における検索対象情報項目の登録時のインデクス更新概略処理フローを示す。

【図8】本発明の実施例における情報検索の概略処理フローを示す。

【図9】本発明の実施例における検索ノード追加時のインデクス作成および更新概略処理フローを示す。

【図10】本発明の実施例における検索ノード追加に伴う既設検索ノードの検索対象範囲割当て削除中の情報検索概略処理フローを示す。

【図11】本発明の実施例における検索ノード削除時のインデクス更新概略処理フローを示す。 20

【図12】本発明の実施例における、バケットと検索ノードの対応を表すデータ構造と、4検索ノードの場合のバケット配置例を示す。

【符号の説明】

【0040】

100：情報検索クライアント

110：フロントエンドエンドノード

111：受付部

121：原本管理部

122：原本格納部

131：インデクス前処理部 30

132：前処理済みインデクス情報格納部

141：構成管理部

142：構成情報格納部

150：通信路

160a～160d：検索ノード

161a～161d：インデクス検索部

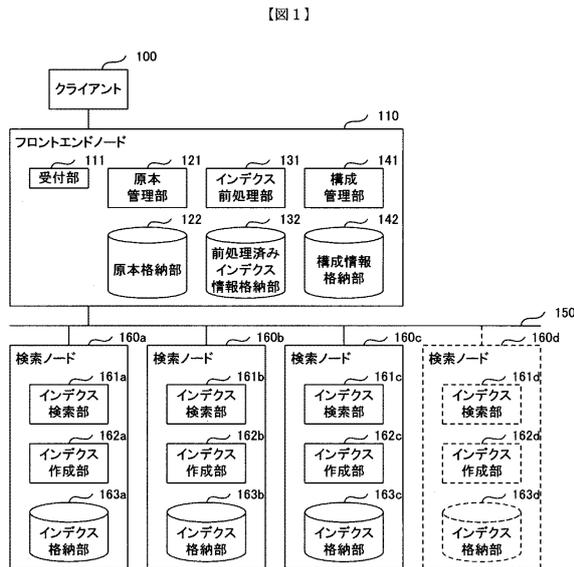
162a～162d：インデクス更新部

163a～163d：インデクス格納部

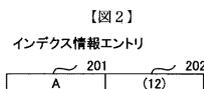
500：前処理済みインデクス情報

600：インデクス 40

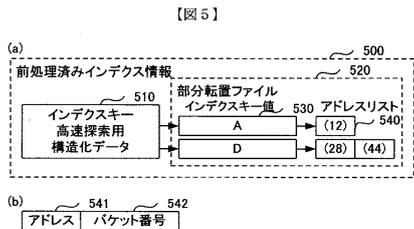
【図1】



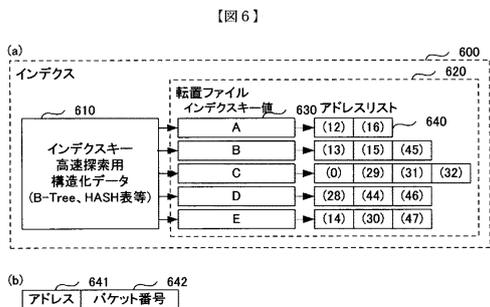
【図2】



【図5】



【図6】



【図3】

【図3】

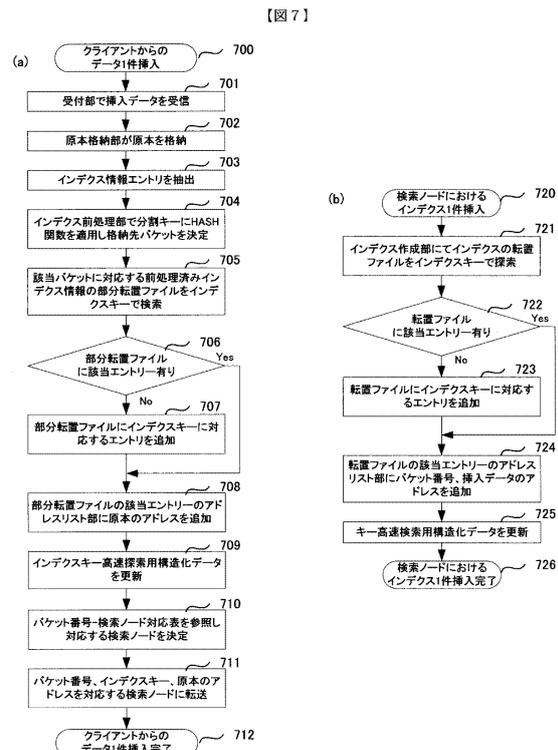
バケット番号	分割キー値
0	12, 28, 44
1	13, 14, 29, 30, 45, 46
2	15, 31, 47
3	0, 16, 32
4	1, 2, 17, 18, 33, 34
5	3, 19, 35
6	4, 20, 36
7	5, 6, 21, 22, 37, 38
8	7, 23, 39
9	1, 13, 25, 37
10	8, 24, 40
11	11, 27, 43

【図4】

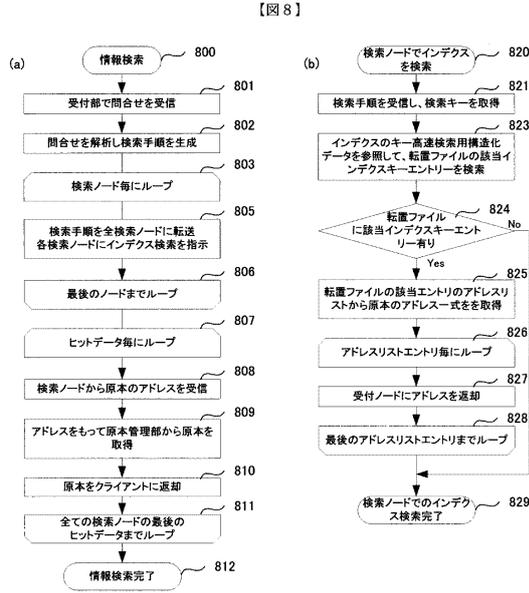
【図4】

バケット番号	検索ノード番号
0	0
1	1
2	2
3	0
4	1
5	2
6	0
7	1
8	2
9	0
10	1
11	2

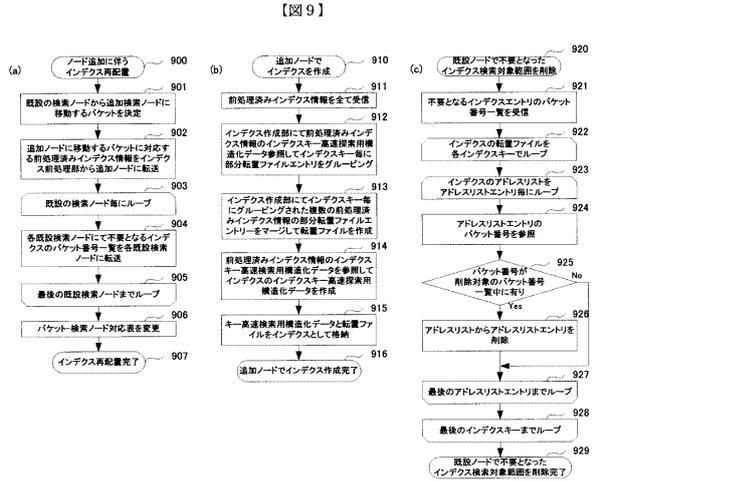
【図7】



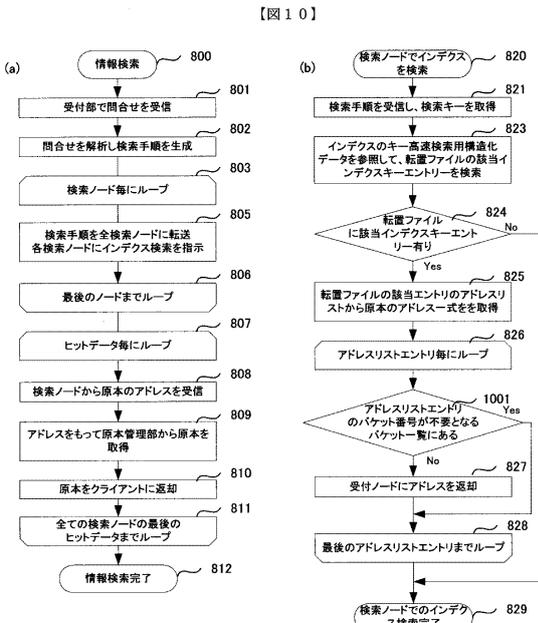
【図 8】



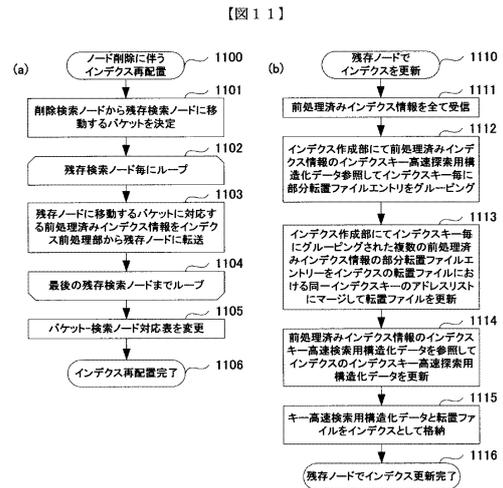
【図 9】



【図 10】



【図 11】



【図 12】

【図 12】

バケット番号	検索ノード番号	検索ノード番号
0	420	0
1		1
2		2
3		3
4		1
5		2
6		0
7		3
8		2
9		0
10		1
11		3

フロントページの続き

(72)発明者 松林 忠孝

神奈川県横浜市戸塚区戸塚町5030番地 株式会社日立製作所 ソフトウェア事業部内

審査官 吉田 誠

(56)参考文献 吉原 朋宏, 並列Btree構造における負荷分散処理の並行性制御への影響, 電子情報通信学会技術研究報告, 日本, 社団法人電子情報通信学会, 2005年 7月 6日, Vol. 105 No. 171, 185 - 190ページ

渡邊 明嗣, 値域分割された分散ストレージにおける効率的なアクセス負荷の記録と管理, 電子情報通信学会技術研究報告, 日本, 社団法人電子情報通信学会, 2003年10月 2日, Vol. 103 No. 358, 31 - 36ページ

宮崎 純, 無共有並列計算機向けディレクトリ構造Fat-Btreeの実装とその評価, 情報処理学会研究報告, 日本, 社団法人情報処理学会, 1999年 7月23日, Vol. 99 No. 61, 407 - 412ページ

(58)調査した分野(Int.Cl., DB名)

G06F 17/30

G06F 12/00