

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6167015号
(P6167015)

(45) 発行日 平成29年7月19日(2017.7.19)

(24) 登録日 平成29年6月30日(2017.6.30)

(51) Int. Cl.	F I				
G06F 17/30	(2006.01)	G06F	17/30	110C	
G06F 12/00	(2006.01)	G06F	17/30	340B	
		G06F	17/30	220C	
		G06F	12/00	520A	

請求項の数 5 (全 20 頁)

(21) 出願番号	特願2013-225881 (P2013-225881)	(73) 特許権者	000005223
(22) 出願日	平成25年10月30日(2013.10.30)		富士通株式会社
(65) 公開番号	特開2015-87953 (P2015-87953A)		神奈川県川崎市中原区上小田中4丁目1番1号
(43) 公開日	平成27年5月7日(2015.5.7)	(74) 代理人	100074099
審査請求日	平成28年7月5日(2016.7.5)		弁理士 大菅 義之
		(74) 代理人	100133570
			弁理士 ▲徳▼永 民雄
		(74) 復代理人	100141232
			弁理士 飯塚 達
		(72) 発明者	熊野 達夫
			神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
		審査官	成瀬 博之

最終頁に続く

(54) 【発明の名称】 情報処理システム、管理プログラム、及びインデックス管理方法

(57) 【特許請求の範囲】

【請求項1】

ネットワークを介して第1の情報処理装置と接続された複数台の第2の情報処理装置、及び管理装置を備え、
前記複数台の第2の情報処理装置は、
データ、及び前記データのメタデータを用いて検索用に構築されたインデックスを分散して管理し、
前記管理装置は、
前記第1の情報処理装置から送信された検索リクエストを監視する監視部と、
同一の第2の情報処理装置に前記インデックスを管理させるべき前記メタデータの種類の組み合わせを、前記検索リクエストに含まれる検索条件における前記メタデータの種類の組み合わせについての共起率に基づいて特定すると共に、該特定した前記組み合わせのメタデータの種類に基づいて前記同一の第2の情報処理装置を特定し、該特定した前記同一の第2の情報処理装置に、該特定した前記組み合わせのメタデータの種類のインデックスを管理させるインデックス管理部と、を有する、
ことを特徴とする情報処理システム。

【請求項2】

前記監視部が監視した前記検索リクエストで指定されているメタデータの種類の組み合わせ毎に、前記組み合わせでメタデータが同時に指定された結果を表す統計情報を生成する情報生成部、を有し、

前記インデックス管理部は、前記統計情報を基に、前記同一の第2の情報処理装置に前記インデックスを管理させるべき前記メタデータの種類の組み合わせを特定する、
ことを特徴とする請求項1記載の情報処理システム。

【請求項3】

前記情報生成部は、2種類のメタデータの組み合わせ毎に、前記2種類のメタデータが同時に指定された回数を計数して、前記統計情報を生成・更新し、

前記インデックス管理部は、前記2種類のメタデータのインデックスを管理単位として、前記同一の第2の情報処理装置に前記インデックスを管理させるべきか否かの判定を行う、

ことを特徴とする請求項2記載の情報処理システム。

10

【請求項4】

ネットワークと接続されたコンピュータに、

データ、及び前記データのメタデータを用いて検索用に構築されたインデックスを分散して管理する複数台の第2の情報処理装置に対し、前記ネットワークを介して第1の情報処理装置から送信された検索リクエストの監視を行い、

同一の第2の情報処理装置に前記インデックスを管理させるべき前記メタデータの種類の組み合わせを、前記検索リクエストに含まれる検索条件における前記メタデータの種類の組み合わせについての共起率に基づいて特定すると共に、該特定した前記組み合わせのメタデータの種類に基づいて前記同一の第2の情報処理装置を特定し、該特定した前記同一の第2の情報処理装置に、該特定した前記組み合わせのメタデータの種類のインデックスを管理させる、

20

処理を実行させる管理プログラム。

【請求項5】

ネットワークと接続されたコンピュータが、

データ、及び前記データのメタデータを用いて検索用に構築されたインデックスを分散して管理する複数台の第2の情報処理装置に対し、前記ネットワークを介して第1の情報処理装置から送信された検索リクエストの監視を行い、

同一の第2の情報処理装置に前記インデックスを管理させるべき前記メタデータの種類の組み合わせを、前記検索リクエストに含まれる検索条件における前記メタデータの種類の組み合わせについての共起率に基づいて特定すると共に、該特定した前記組み合わせのメタデータの種類に基づいて前記同一の第2の情報処理装置を特定し、該特定した前記同一の第2の情報処理装置に、該特定した前記組み合わせのメタデータの種類のインデックスを管理させる、

30

ことを特徴とするインデックス管理方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、データ、及びそのデータのメタデータを用いて検索用に構築されたインデックスを複数台の情報処理装置に分散させて管理させるための技術に関する。

【背景技術】

40

【0002】

コンピュータ等の情報処理装置に蓄積させるデータ量は増大傾向にある。そのため、現在では、大量のデータをより低コストで管理することが強く求められている。分散オブジェクトストレージは、大量のデータを低コストで管理可能な技術の1つとして着目されている。

【0003】

分散オブジェクトストレージは、複数台のサーバを用いた情報処理システムとして構築される。分散オブジェクトストレージでは、データはオブジェクトとして管理され、オブジェクトは、分散オブジェクトストレージを構成するサーバに分散して保存される。オブジェクトの分散により、そのオブジェクトの保存、及び計算（処理）に利用可能な各リソ

50

ース量は共に、サーバの台数に応じて向上することになる。そのため、分散オブジェクトストレージでは、サーバの追加により、記憶容量、及びパフォーマンスを共に向上させることができる。

【0004】

オブジェクトの分散により、処理の負荷も分散させることができる。そのため、サーバに要求される性能も抑えられる。このことから、分散オブジェクトストレージでは、サーバの台数を増やすスケールアウトに要するコストを抑制することができる。

【0005】

分散オブジェクトストレージでは、オブジェクトの管理にメタデータが用いられる。そのメタデータは、対応するオブジェクトに関連する情報であり、対応するオブジェクトの検索・抽出に用いられる。そのため、オブジェクトを検索・抽出するためのデータであるインデックスは、メタデータを用いて構築される。メタデータは、1つのオブジェクトに対し1つ以上、作成される。

10

【0006】

分散オブジェクトストレージにおける検索は、インデックスを管理するサーバによって行われる。分散オブジェクトストレージにおけるインデックスの管理には、以下の方式が用いられている。

(a) 特定の専用サーバ(例えば1台のサーバ)が全てのインデックスを集中管理する(以降「集中管理方式」と表記)

(b) インデックス毎に、異なるサーバが管理する(以降「個別管理方式」と表記)

(c) 自身が保存するオブジェクトのインデックスも併せて管理する(以降「並行管理方式」と表記)

20

【0007】

上記各管理方式には、それぞれ以下のような問題点がある。

集中管理方式では、検索、更にはインデックスの更新が専用サーバに集中する。インデックスを管理する専用サーバの台数は、通常、オブジェクトを管理するサーバの台数と比較して、非常に小さい。そのため、専用サーバの負荷が重くなり易い。重い負荷は検索性能を低下させる。

【0008】

分散オブジェクトストレージは、通常、複数のメタデータを指定したアンド(AND)検索をサポートする。個別管理方式では、アンド検索の場合、指定されたメタデータ毎に、そのメタデータに対応するインデックスを管理するサーバに問い合わせを送信しなければならない。また、フロントエンド等において、各問い合わせによって特定された対象オブジェクトを各サーバから取得し、取得した対象オブジェクトのなかからアンド検索で指定の検索条件を満たす対象オブジェクトを抽出するといった処理を行わなければならない。このようなことから、個別管理方式では、データの転送量が大きくなり易く、高い検索性能は得られ難い。

30

【0009】

並行管理方式でも、個別管理方式と同様に、アンド検索の場合、指定されたメタデータ毎に、そのメタデータに対応するインデックスを管理するサーバに問い合わせを送信しなければならない。そのため、高い検索性能は得られ難い。

40

【0010】

上記のように、何れの管理方式も、アンド検索時には検索性能が低下し易いという問題点がある。オブジェクトを必要とするユーザにとっての利便性を考慮するならば、アンド検索時における検索性能の低下を抑えることが重要と思われる。

【先行技術文献】

【特許文献】

【0011】

【特許文献1】特開2012-168781号公報

【特許文献2】特開昭62-121576号公報

50

【特許文献 3】特開 2002 - 7641 号公報

【特許文献 4】特開 2012 - 48592 号公報

【発明の概要】

【発明が解決しようとする課題】

【0012】

一側面では、本発明は、分散オブジェクトストレージのような複数台の情報処理装置（サーバ）を用いて構築された情報処理システムにおけるアンド検索時の検索性能の低下を抑制するための技術を提供することを目的とする。

【課題を解決するための手段】

【0013】

本発明を適用した 1 システムは、ネットワークを介して第 1 の情報処理装置と接続された複数台の第 2 の情報処理装置、及び管理装置を備え、複数台の第 2 の情報処理装置は、データ、及びデータのメタデータを用いて検索用に構築されたインデックスを分散して管理し、管理装置は、第 1 の情報処理装置から送信された検索リクエストを監視する監視部と、同一の第 2 の情報処理装置にインデックスを管理させるべきメタデータの種類の組み合わせを、検索リクエストに含まれる検索条件におけるメタデータの種類の組み合わせについての共起率に基づいて特定すると共に、特定した組み合わせのメタデータの種類に基づいて同一の第 2 の情報処理装置を特定し、該特定した同一の第 2 の情報処理装置に、該特定した組み合わせのメタデータの種類のインデックスを管理させるインデックス管理部と、を有する。

【発明の効果】

【0014】

本発明を適用した場合には、分散オブジェクトストレージのような複数台の情報処理装置（サーバ）を用いて構築された情報処理システムにおけるアンド検索時の検索性能の低下を抑制することができる。

【図面の簡単な説明】

【0015】

【図 1】本実施形態による管理装置を用いて構築された情報処理システムの構成例を説明する図である。

【図 2】統計情報の構成例を説明する図である。

【図 3】統計情報の更新例を説明する図である。

【図 4】インデックスを管理させるストレージサーバ兼インデックス管理サーバの変更例を説明する図である。

【図 5】統計情報更新処理のフローチャートである。

【図 6】第 1 の計数処理のフローチャートである。

【図 7】第 2 の計数処理のフローチャートである。

【図 8】第 2 の計数処理の変形例のフローチャートである。

【図 9】インデックス管理サーバ決定処理のフローチャートである。

【図 10】サーバ決定処理のフローチャートである。

【図 11】サーバ連表の構成例を説明する図である。

【図 12】本実施形態によるストレージサーバ兼インデックス管理サーバとして使用可能な情報処理装置の構成例を表す図である。

【発明を実施するための形態】

【0016】

以下、本発明の実施形態について、図面を参照しながら詳細に説明する。

図 1 は、本実施形態による管理装置を用いて構築された情報処理システムの構成例を説明する図である。

【0017】

図 1 に表すように、情報処理システムは、ネットワーク 1 に、複数台のアプリケーションサーバ（図 1 中「アプリサーバ」と表記。以降この表記を用いる）2、及び複数台のス

10

20

30

40

50

ストレージサーバ兼インデックス管理サーバ（以降「ストレージ・インデックス管理サーバ」と表記）3を接続させた構成となっている。

【0018】

各ストレージ・インデックス管理サーバ3には、管理対象とするデータ、及びそのデータの検索用のデータであるインデックスを保存可能なディスク装置4が接続されている。このディスク装置4を管理する全てのストレージ・インデックス管理サーバ3をネットワーク1と接続することにより、分散オブジェクトストレージが構築されている。

【0019】

ディスク装置4は、通常、複数台、各ストレージ・インデックス管理サーバ3に接続されるか、或いは1台のストレージである。ここでは、説明上、便宜的に、各ストレージ・インデックス管理サーバ3に接続されているディスク装置4は1台と想定する。

10

【0020】

各アプリサーバ2は、不図示の別のネットワークと接続された端末装置を使用するユーザに対してサービスを提供するサーバである。各アプリサーバ2は、分散オブジェクトストレージが管理するデータ（オブジェクト）のなかで端末装置のユーザが要求するデータを提供するフロントエンドとして機能する。また、各アプリサーバ2は、分散オブジェクトストレージの保守、オブジェクトの追加、削除、メタデータの登録、削除、等にも用いることができる。

【0021】

各ストレージ・インデックス管理サーバ3は、図1に表すように、機能構成として、ネットワーク入出力部31、データ読み書き部32、メタデータ読み書き部33、ディスク入出力部34、検索部35、統計情報取得部36、記憶部37、及びインデックス管理サーバ決定部38を備えている。

20

【0022】

ネットワーク入出力部31は、ネットワーク1を介した通信を実現させる。他のストレージ・インデックス管理サーバ3、及び各アプリサーバ2との通信は、ネットワーク入出力部31を用いて行われる。

【0023】

データ読み書き部32は、ディスク装置4からのオブジェクトの読み出し、ディスク装置4へのオブジェクトの書き込み、及びディスク装置4に格納されたオブジェクトの削除、等の操作のための機能である。実際の操作は、ディスク入出力部34によって行われる。

30

【0024】

メタデータ読み書き部33は、ディスク装置4からのメタデータの読み出し、ディスク装置4へのメタデータの書き込み、及びディスク装置4に格納されたメタデータの削除、等の操作のための機能である。ディスク装置4には、メタデータはインデックスを構成する一部のデータとして格納されている。そのため、実際の操作は、ディスク入出力部34を用いたインデックスの更新のための操作として行われる。

【0025】

このインデックスは、メタデータの種類毎に、そのメタデータを用いて構築される検索用のデータである。このインデックスは、メタデータの値（内容）毎に、その値をメタデータとして割り当てられたオブジェクトの保存先を表す保存先情報を備えている。そのインデックスでは、保存先情報からオブジェクトに割り当てられたメタデータの値を特定可能になっている。

40

【0026】

各アプリサーバ2は、端末装置のユーザからの要求により、そのユーザがメタデータの値を入力して指定した検索条件に従って、その検索条件を満たすオブジェクトの保存先を確認するための検索リクエストを送信する。ネットワーク入出力部31は、自ストレージ・インデックス管理サーバ3宛に送信された検索リクエストを受信し、受信した検索リクエストを検索部35に渡す。検索部35は、渡された検索リクエスト中に格納されたメタ

50

データの値を抽出し、インデックスに対する検索を行い、その検索結果をネットワーク入出力部 31 から送信させる。

【0027】

記憶部 37 は、統計情報 37 a の保存に用いられる。その統計情報 37 a は、メタデータの種類（以降「メタデータ名」と表記）の組み合わせ毎に、その組み合わせで各値が指定された検索条件での検索を行った回数を表す情報である。

【0028】

図 2 は、統計情報の構成例を説明する図である。図 2 中に表記の「メタデータ 1」「メタデータ 2」は、異なるメタデータ名を表している。「メタデータ」に付した「1」「2」の数値は、辞書に登録された順序、或いは識別用に割り当てられた番号を表している。ここでは、数値は辞書に登録された順序を表していると想定する。

10

【0029】

図 2 では、統計情報の構成例をテーブル形式で表している。それにより、各欄は、回数を計数する 1 種類のメタデータ名、或いは 2 種類のメタデータ名の組み合わせを表している。

【0030】

例えば行、列ともに「メタデータ 1」の欄には、メタデータ名としてメタデータ 1 のみ値が指定された検索条件で検索を行った回数が格納される。例えば行が「メタデータ 2」、列が「メタデータ 1」の欄に表記されている「未使用」は、回数が格納されないことを表している。これは、2 種類のメタデータ名の組み合わせで各値が指定された検索条件での検索では、辞書に登録された順序で 2 種類のメタデータ名の組み合わせを特定し、特定した組み合わせのみで回数を計数するからである。

20

【0031】

図 3 (a) ~ 図 3 (c) は、統計情報の更新例を説明する図である。図 3 (a) ~ 図 3 (c) に表す例は、メタデータ 1 として「名前」、メタデータ 2 として「年齢」をそれぞれ想定した場合の例である。図 3 (b) は、図 3 (a) の状態の統計情報 37 a から 1 回の検索を行った後の状態、図 3 (c) は、図 3 (b) の状態の統計情報 37 a から 1 回の検索を行った後の状態、をそれぞれ表している。

【0032】

図 3 (a) に表す状態は、統計情報 37 a がクリアされた状態で、例えば「名前 = 田中 AND 年齢 = 20」のような 2 種類のメタデータ名の値が指定された検索条件による検索を行った後の状態である。その検索条件では、1 種類のメタデータ名として「名前」及び「年齢」が指定され、2 種類のメタデータ名の組み合わせが「名前」及び「年齢」である。そのため、行、及び列が共に「名前」の欄、行、及び列が共に「年齢」の欄、並びに行が「名前」で列が「年齢」の欄の各回数は、それぞれ 0 から 1 に更新されている。

30

【0033】

図 3 (b) に表す状態は、統計情報 37 a が図 3 (a) に表す状態時に、例えば「名前 = 山田」のような 1 種類のメタデータ名の値が指定された検索条件による検索を行った後の状態である。その検索条件では、行、及び列が共に「名前」の欄の回数が 1 から 2 に更新されている。

40

【0034】

図 3 (c) に表す状態は、統計情報 37 a が図 3 (b) に表す状態時に、例えば「名前 = 上田 AND 年齢 < 30」のような 2 種類のメタデータ名の値が指定された検索条件による検索を行った後の状態である。その検索条件により、行、及び列が共に「名前」の欄では回数が 2 から 3 に更新されている。行、及び列が共に「年齢」の欄では回数が 1 から 2 に更新されている。行が「名前」で列が「年齢」の欄では、回数が 1 から 2 に更新されている。

【0035】

本実施形態では、上記のように、1 種類のメタデータ名の値（メタデータ）のみが検索条件に指定された検索、及び 2 種類のメタデータ名の組み合わせで各値が検索条件に指定

50

された検索の両方で回数を計数し、各計数結果を統計情報 37a として保存している。3 種類以上のメタデータ名の組み合わせは、2 種類のメタデータ名の組み合わせに分解し、分解した組み合わせ毎に回数を計数するようにしている。この統計情報 37a は、各ストレージ・インデックス管理サーバ 3 に管理させるべきメタデータ名のインデックスの決定に用いられる。

【0036】

検索条件 (AND 検索) で値が指定された 2 種類のメタデータ名のインデックスを同じストレージ・インデックス管理サーバ 3 が管理している場合、そのストレージ・インデックス管理サーバ 3 は、その検索条件を満たすオブジェクトの保存先情報を全て抽出することができる。そのため、アンド検索の検索条件を格納した検索リクエストが送信されたアプリサーバ 2 は、その検索リクエストを送信したストレージ・インデックス管理サーバ 3 からのレスポンスにより、その検索条件を満たすオブジェクトを全て取得することができる。

10

【0037】

1 台のストレージ・インデックス管理サーバ 3 からのレスポンスにより、検索条件を満たすオブジェクトを全て取得できることは、ネットワーク 1 を介して送受信するオブジェクトの数を抑えられることを意味する。1 台のストレージ・インデックス管理サーバ 3 は、他のストレージ・インデックス管理サーバ 3 との通信を行うことなく、送信すべきレスポンスを作成することができる。このようなことから、アンド検索時、検索条件で値が指定された複数のメタデータ名のインデックスを同じストレージ・インデックス管理サーバ 3 が管理している場合、データ送受信による遅延を最小限に抑えることが可能になる。

20

【0038】

また、レスポンスは、ストレージ・インデックス管理サーバ 3 内の処理により作成されることから、ストレージ・インデックス管理サーバ 3 の計算リソースがより有効に活用することができる。これは、フロントエンドであるアプリサーバ 2 の負荷をより抑制できることを意味する。負荷の抑制により、アプリサーバ 2 はより迅速に検索結果としてのオブジェクトを端末装置のユーザに提示することができる。このようなことから、検索条件で値が指定された複数のメタデータ名のインデックスが同じストレージ・インデックス管理サーバ 3 に管理させている場合、単位時間あたりに実行可能な検索の件数はより多くなり、検索性能は事実上、向上する。

30

【0039】

このように、アンド検索時の検索条件で値が指定される複数のメタデータ名のインデックスを同じストレージ・インデックス管理サーバ 3 に管理させることは、検索性能を向上させるうえで有効である。そのため、統計情報 37a は、同じストレージ・インデックス管理サーバ 3 にインデックスを管理させるべきメタデータ名の組み合わせの特定に用いられる。

【0040】

本実施形態では、図 2 に表すように、組み合わせるメタデータ名の数は 2 つとしている。これは、組み合わせるメタデータ名の数が増えるほど、同じメタデータ名のインデックスを管理させるストレージ・インデックス管理サーバ 3 の台数も多くなり、各ストレージ・インデックス管理サーバ 3 に管理させるインデックスの数も多くなるからである。言い換えれば、ネットワーク 1 を介したデータの転送量の増大、及びストレージ・インデックス管理サーバ 3 における計算リソースの浪費の増大、等の影響が大きくなり易くなるためである。このことから、本実施形態では、2 種類のメタデータ名の組み合わせのみに着目している。

40

【0041】

上記検索部 35 は、ネットワーク入出力部 31 から渡された検索リクエストを統計情報取得部 36 に渡す。統計情報取得部 36 は、渡された検索リクエストから検索条件を抽出し、抽出した検索条件に該当する回数をインクリメントすることにより、統計情報 37a を更新、つまり新しい統計情報 37a を生成する。

50

【 0 0 4 2 】

インデックス管理サーバ決定部 3 8 は、統計情報 3 7 a を参照し、2 種類のメタデータ名の組み合わせ毎に、その 2 種類のメタデータ名のインデックスを管理させるべきストレージ・インデックス管理サーバ 3 を決定する。インデックス管理サーバ決定部 3 8 は、その決定結果に従って、インデックスを管理させるストレージ・インデックス管理サーバ 3 を変更させる制御を行う。

【 0 0 4 3 】

本実施形態による管理装置は、各ストレージ・インデックス管理サーバ 3 に管理させるインデックスを動的に変更する制御を行う装置である。その管理装置は、インデックス管理サーバ決定部 3 8 によって実現されている。つまり、本実施形態による管理装置は、各

10

【 0 0 4 4 】

なお、管理装置は、各アプリサーバ 2 に搭載させても良い。また、ネットワーク 1 に接続させた別の情報処理装置（コンピュータ）を管理装置として用いても良い。その別の情報処理装置は、1 台であっても良く、2 台以上であっても良い。

【 0 0 4 5 】

図 4 は、インデックスを管理させるストレージサーバ兼インデックス管理サーバの変更例を説明する図である。

【 0 0 4 6 】

図 4 では、便宜的に、3 台のストレージ・インデックス管理サーバ 3（3 - 1 ~ 3 - 3）のみがネットワーク 1 と接続されているシステム構成を想定している。図 4 の上方に表すシステム構成は、インデックスを管理させるストレージ・インデックス管理サーバ 3 を変更する前の状態である。図 4 の下方に表すシステム構成は、インデックスを管理させるストレージ・インデックス管理サーバ 3 を変更した後の状態である。各システム構成において、保存されているオブジェクトとしてオブジェクト A ~ C のみを表し、インデックスとしては、メタデータ名が「名前」「年齢」の 2 つのみを表している。

20

【 0 0 4 7 】

図 4 の上方に表すシステム構成では、3 つのオブジェクト A ~ C はそれぞれストレージ・インデックス管理サーバ 3 - 1 ~ 3 - 3 に保存されている。メタデータ名が「名前」のインデックスはストレージ・インデックス管理サーバ 3 - 1 で管理され、メタデータ名が「年齢」のインデックスはストレージ・インデックス管理サーバ 3 - 3 で管理されている。各インデックスは、共に、オブジェクト A ~ C の保存先の検索に用いられる。メタデータ名が「名前」のインデックスで内容例として表記の「田中：オブジェクト A」等、メタデータ名が「年齢」のインデックスで内容例として表記の「20：オブジェクト A」等は、そのことを表している。

30

【 0 0 4 8 】

図 4 の下方に表すシステム構成では、メタデータ名が「名前」及び「年齢」の 2 つのインデックスは、ストレージ・インデックス管理サーバ 3 - 2 に管理されている。それにより、ストレージ・インデックス管理サーバ 3 - 1 及び 3 - 3 は共に、管理するインデックスが存在しない状態となっている。

40

【 0 0 4 9 】

「名前」及び「年齢」は、検索条件でその値が同時に指定されたメタデータ名である。本実施形態では、設定された変更条件を満たす場合に、検索条件で値が同時に指定された 2 種類のメタデータ名のインデックスを同じストレージ・インデックス管理サーバ 3 に管理させるようにしている。その変更条件を設けることにより、本実施形態では、インデックスを管理させるストレージ・インデックス管理サーバ 3 の変更を制限している。

【 0 0 5 0 】

本実施形態では、この変更条件は、2 種類のメタデータ名の値が検索条件で同時に指定される頻度に着目して設定している。より具体的には、変更条件は、閾値 r を超える共起率 r が得られることとしている。

50

【0051】

共起率 r は、2種類のメタデータ名の値が検索条件で同時に指定される割合のことである。2種類のメタデータ名の値のうち的一方、及び他方がそれぞれ検索条件で指定された回数を c_A 、及び c_B 、2種類のメタデータ名の値が検索条件で同時に指定された回数を c_{AB} と表す場合、共起率 r は例えば以下の式を用いて計算される。

$$r = c_{AB} / (c_A + c_B - c_{AB}) \quad \dots (1)$$

【0052】

この共起率 r は、以下の式を用いて計算しても良い。

$$r = c_{AB} / C \quad \dots (2)$$

ここで、 C は検索が行われた総回数である。

10

【0053】

共起率 r が低いことは、2種類のメタデータ名のインデックスを同じストレージ・インデックス管理サーバ3に管理させても、その2つのインデックスを対象にした検索を同時に行う頻度が小さいことを意味する。その2つのインデックスへの検索を同時に行わないのであれば、各インデックスをそれぞれ異なるストレージ・インデックス管理サーバ3に管理させることにより、2台のストレージ・インデックス管理サーバ3の計算リソースをより有効に利用することができる。このようなことから、本実施形態では、閾値 r を設け、変更条件 ($r >$ の関係) が満たされた2種類のメタデータ名のインデックスを同じストレージ・インデックス管理サーバ3に管理させるようにしている。閾値 r としては、例えば0.5の付近の値を設定することが考えられる。

20

【0054】

インデックス管理サーバ決定部38は、統計情報37aを参照して、2種類のメタデータ名の組み合わせ毎に共起率 r を算出する。インデックス管理サーバ決定部38は、2種類のメタデータ名の組み合わせ毎に、算出した共起率 r からその2種類のメタデータ名のインデックスを同じストレージ・インデックス管理サーバ3に管理させるべきか否かを判定する。それにより、インデックス管理サーバ決定部38は、算出した共起率 r が閾値 r を超えている2種類のメタデータ名のインデックスを同じストレージ・インデックス管理サーバ3に管理させる。

【0055】

インデックス管理サーバ決定部38が共起率 r を算出する2種類のメタデータ名の組み合わせは、統計情報37aによってサポートされる、管理単位となる組み合わせである。統計情報37aを用いて回数が計数可能なメタデータ名は、自ストレージ・インデックス管理サーバ3に送信される検索リクエスト中の検索条件で値が指定されるメタデータ名に限定される。

30

【0056】

本実施形態では、何れかのストレージ・インデックス管理サーバ3のインデックス管理サーバ決定部38の決定結果に従い、対象となるインデックスを管理させるストレージ・インデックス管理サーバ3を直ちに更新させるようにしている。インデックスを管理させるストレージ・インデックス管理サーバ3を更新させる場合、その更新が終了するまで、検索リクエストへの適切なレスポンスが行えない可能性がある。そのため、本実施形態では、インデックス管理サーバ決定部38は、外部からの指示により起動させ、インデックス管理サーバ決定部38を動作させている間、検索サービスを停止させるようにしている。

40

【0057】

インデックス管理サーバ決定部38によって同じストレージ・インデックス管理サーバ3にインデックスを管理させるべきと判定された2種類のメタデータ名の組み合わせ結果(以降「インデックス管理情報」と表記)は、各ストレージ・インデックス管理サーバ3で共有される。そのため、インデックスを管理させるストレージ・インデックス管理サーバ3の更新は、インデックス管理情報に登録されていない組み合わせのみを対象に行われる。

50

【 0 0 5 8 】

図 1 2 は、本実施形態によるストレージサーバ兼インデックス管理サーバとして使用可能な情報処理装置の構成例を表す図である。ここで、図 1 2 を参照し、本実施形態によるストレージ・インデックス管理サーバ 3 として使用可能な情報処理装置例について具体的に説明する。

【 0 0 5 9 】

この情報処理装置は、図 1 2 に表すように、CPU (Central Processing Unit) 5 1、FWH (Firm-Ware Hub) 5 2、メモリ (メモリモジュール) 5 3、NIC (Network Interface Card) 5 4、ハードディスク装置 (HD) 5 5、ファン 5 6、コントローラ 5 7、及び BMC (Baseboard Management Controller) 5 8 を備えている。この構成は、1 例であり、ストレージ・インデックス管理サーバ 3 として使用可能な情報処理装置の構成は、図 1 2 に表すようなものに限定されない。

10

【 0 0 6 0 】

FWH 5 2 は、ファームウェアを格納したメモリである。このファームウェアは、CPU 5 1 によってメモリ 5 3 に読み出され実行される。ハードディスク装置 5 5 には、OS (Operating System)、及び各種アプリケーション・プログラム (以降「アプリケーション」と略記) が格納されている。アプリケーションとしては、分散オブジェクトストレージを構成するノード (ストレージ・インデックス管理サーバ 3) として機能させるためのもの (以降「分散オブジェクトストレージシステム」と表記) が含まれる。CPU 5 1 は、ファームウェアの起動が完了した後、コントローラ 5 7 を介してハードディスク装置 5 5 から OS、及び分散オブジェクトストレージシステムを読み出して実行することができる。NIC 5 4 を介した通信は、OS の起動によって可能となる。

20

【 0 0 6 1 】

この分散オブジェクトストレージシステムは、ハードディスク装置 5 5 以外のストレージ、或いは記録媒体に格納しても良い。ストレージ、或いは記録媒体は、NIC 5 4 がネットワークを介して通信可能な外部装置がアクセス可能なものであっても良い。このことから、分散オブジェクトストレージシステムは、外部装置から受信しても良い。

【 0 0 6 2 】

BMC 5 8 は、情報処理装置を管理するための専用の管理装置である。BMC 5 8 は、CPU 5 1 のオン/オフ、各構成要素に発生するエラーの監視、等を行う。コントローラ 5 6 は、図 1 に表すディスク装置 4 が接続可能なインターフェースである。

30

【 0 0 6 3 】

上記のような構成の情報処理装置では、ネットワーク入出力部 3 1 は NIC 5 4 である。ディスク入出力部 3 4 は、コントローラ 5 7 である。記憶部 3 7 は、例えばメモリ 5 3、ハードディスク装置 5 5、及びコントローラ 5 7 によって実現される。データ読み書き部 3 2、メタデータ読み書き部 3 3、検索部 3 5、統計情報取得部 3 6、及びインデックス管理サーバ決定部 3 8 は、共に、CPU 5 1 が OS 上で分散オブジェクトシステムを実行することで実現される機能である。このことから、データ読み書き部 3 2、メタデータ読み書き部 3 3、検索部 3 5、統計情報取得部 3 6、及びインデックス管理サーバ決定部 3 8 は、共に、CPU 5 1、FWH 5 2、メモリ 5 3、ハードディスク装置 5 5、及びコントローラ 5 7 によって実現される。

40

【 0 0 6 4 】

上記分散オブジェクトストレージシステムは、CPU 5 1 に各種処理を実行させる。以降、図 5 ~ 図 1 0 に表す各種フローチャートを参照し、分散オブジェクトストレージシステムを実行することによる CPU 5 1 の動作について詳細に説明する。

【 0 0 6 5 】

図 5 は、統計情報更新処理のフローチャートである。この統計情報更新処理は、何れかのアプリサーバ 2 から検索を要求する検索リクエストを NIC 5 4 が受信することを契機に実行される処理であり、統計情報取得部 3 6 の機能の一部を実現させる。更新情報 3 7 a は、この統計情報更新処理の実行により、受信された検索リクエスト中の検索条件で値

50

が指定されたメタデータ名に応じて更新される。始めに図5を参照し、統計情報更新処理について詳細に説明する。

【0066】

先ず、CPU51は、検索リクエスト中の検索条件で値が指定されたメタデータ名毎に、その回数を更新するための第1の計数処理を実行する(S1)。その第1の計数処理の実行後、CPU51は、検索条件(図5中「検索式」と表記)に2種類以上のメタデータ名の値(メタデータ)が含まれているか否か判定する(S2)。検索条件で値が指定されたメタデータ名が2種類以上であった場合、S2の判定はyesとなってS5に移行する。検索条件で値が指定されたメタデータ名が1種類であった場合、S2の判定はnoとなり、ここで統計情報更新処理が終了する。

10

【0067】

S5は処理ループであり、この処理ループS5内では、CPU51は、検索リクエスト中の検索条件で値が指定された2種類のメタデータ名の組み合わせ毎に、その回数を更新するための第2の計数処理を実行する(S6)。2種類のメタデータ名の組み合わせの全てで回数を更新した場合、処理ループS5が終了し、統計情報更新処理も終了する。

【0068】

図6は、上記第1の計数処理のフローチャートであり、図7は、上記第2の計数処理のフローチャートである。ここで、図6、及び図7を参照し、上記統計情報更新処理内でS1、及びS6として実行される各サブルーチン処理について詳細に説明する。

【0069】

20

S1として実行される第1の計数処理では、図6に表すように、検索リクエスト中の検索条件で値が指定されたメタデータ名毎に、その回数を更新するための処理ループS10が実行される。その処理ループS10では、先ず、CPU51は、検索条件で値が指定されたメタデータ名を取得する(S11)。次に、CPU51は、取得したメタデータ名に対応する回数を統計情報37aから抽出する(S12)。図6では、統計情報37aを「表」、回数を「行列」とそれぞれ表記している。

【0070】

回数を抽出した後、CPU51は、その回数をインクリメントする(S13)。その回数のインクリメントにより、1メタデータ名に対する一連の処理が終了する。検索条件で値が指定された全てのメタデータ名を対象にしていた場合、一連の処理の終了後、処理ループS10が終了し、更に第1の計数処理が終了する。

30

【0071】

S6として実行される第2の計数処理では、図7に表すように、先ず、CPU51は、検索条件で値が指定された2種類のメタデータ名を取得する(S21)。次に、CPU51は、取得した2種類のメタデータ名を辞書順に比較し、辞書順で小さい方、つまり前の方を第1のメタデータ名、残りを第2のメタデータ名とする(S22)。

【0072】

続いてCPU51は、第1及び第2のメタデータ名の組み合わせに対応する回数を統計情報37a中から抽出する(S23)。その回数とは、図2において、第1のメタデータ名を行、第2のメタデータ名を列とする欄に格納されている回数である。回数を抽出した後、CPU51は、その回数をインクリメントする(S24)。その回数のインクリメントを行った後、第2の計数処理が終了する。

40

【0073】

上記第1及び第2の計数処理が統計情報更新処理内で実行されることにより、検索リクエストが受信される度に、図3に例を表すように統計情報37aが更新される。

【0074】

なお、本実施形態では、統計情報37a中からの2種類のメタデータ名の組み合わせに対応する回数の抽出を、メタデータ名が辞書に出現する順序の組で特定するようにしているが、回数の抽出は別の方法で行えるようにしても良い。具体的には、例えば統計情報37aは、ハッシュ値と回数を含む組の集合体としても良い。統計情報37aがハッシュ値

50

と回数を含む組の集合体であった場合、第2の計数処理は、図8にフローチャートを表すような内容としても良い。

【0075】

図8にフローチャートを表す第2の計数処理では、まず、CPU51は、検索条件で値が指定された2種類のメタデータ名を取得する(S31)。次に、CPU51は、取得した2種類のメタデータ名を辞書順に比較し、辞書順で小さい方を第1のメタデータ名、残りを第2のメタデータ名とし、第1のメタデータ名、第2のメタデータ名の順に並べて結合したものを最終的なメタデータ名とする(S32)。

【0076】

続いてCPU51は、最終的なメタデータ名を用いてハッシュ値を計算し、計算したハッシュ値を用いた統計情報37aの検索により、対応する回数を抽出する(S33)。回数を抽出した後、CPU51は、その回数をインクリメントする(S34)。その回数のインクリメントを行った後、変形例の第2の計数処理が終了する。

10

【0077】

図9は、インデックス管理サーバ決定処理のフローチャートである。このインデックス管理サーバ決定処理では、同じストレージ・インデックス管理サーバ3にインデックスを管理させるべき2種類のメタデータ名の特定が行われる。また、このインデックス管理サーバ決定処理では、特定した2種類のメタデータ名のインデックスを管理させるべきストレージ・インデックス管理サーバ3に保存させるための処理が行われる。インデックス管理サーバ決定部38は、このインデックス管理サーバ決定処理をCPU51が実行することで実現される。このインデックス管理サーバ決定処理は、外部、例えば何れかのアプリサーバ2からインデックスの管理を要求するリクエストの受信によって実行される。そのリクエストは、例えば情報処理システムを管理する管理者の指示、或いは設定によって送信される。

20

【0078】

このインデックス管理サーバ決定処理では、2つの処理ループS40、及びS50がその順序で実行される。処理ループS40は、2種類のメタデータ名の組み合わせ毎に、その2つのインデックスを管理させるべきストレージ・インデックス管理サーバ3に管理させるための処理ループである。処理ループS50は、1種類のメタデータ名毎に、そのインデックスを管理させるべきストレージ・インデックス管理サーバ3に管理させるための

30

【0079】

各ストレージ・インデックス管理サーバ3で回数の計数が可能なのは、そのストレージ・インデックス管理サーバ3が受信する検索リクエスト中の検索条件で値が指定されている1種類のメタデータ名、及び2種類のメタデータ名の組み合わせである。各アプリサーバ2は、通常、端末装置のユーザが値を指定したメタデータ名から、検索リクエストを送信すべきストレージ・インデックス管理サーバ3を特定し、特定したストレージ・インデックス管理サーバ3宛に検索リクエストを送信する。このことから、本実施形態では、各ストレージ・インデックス管理サーバ3で対象にさせるメタデータ名、及び2種類のメタデータ名の組み合わせは、検索条件で指定されたことのあるもの、つまり回数が1以上のものとしている。

40

【0080】

処理ループS40では、まず、CPU51は、統計情報37aを参照して、回数が1以上の2種類のメタデータ名の組み合わせの1つを抽出し、抽出した組み合わせの共起率 r を計算して、計算した共起率 r が閾値より大きいか否か判定する(S41)。その組み合わせが同時に指定された検索が頻繁に要求されているような場合、S41の判定はyesとなってS42に移行する。その組み合わせが同時に指定された検索が頻繁に要求されていない場合、S41の判定はnoとなり、処理ループS40内の一連の処理が終了する。その結果、処理ループS40では、回数が1以上の2種類のメタデータ名の組み合わせのなかで抽出していないものが有るか否かの判定が行われる。抽出対象となる組み合わせ

50

が存在しない場合、処理ループ S 4 0 が終了し、抽出対象となる組み合わせが残っている場合、S 4 1 が再度、実行されることとなる。

【 0 0 8 1 】

S 4 1 の判定処理には、実際には、上記インデックス管理情報が用いられる。そのインデックス管理情報には、例えば S 4 1 の判定が初めて y e s となった 2 種類のメタデータ名の組み合わせが登録される。それにより、インデックス管理情報に既に登録されている組み合わせは、共起率 r の値に係わらず、S 4 1 で n o と判定される。

【 0 0 8 2 】

S 4 2 では、C P U 5 1 は、抽出した 2 種類のメタデータ名を辞書順に比較し、辞書順で小さい方を第 1 のメタデータ名、残りを第 2 のメタデータ名とし、第 1 のメタデータ名、第 2 のメタデータ名の順に並べて結合する。次に C P U 5 1 は、2 種類のメタデータ名のインデックスを管理させるストレージ・インデックス管理サーバ 3 を決定し、決定したストレージ・インデックス管理サーバ 3 にその 2 つのインデックスを移動させるサーバ決定処理を実行する (S 4 3)。そのサーバ決定処理を実行した後、処理ループ S 4 0 内の一連の処理が終了する。S 4 2 で結合したメタデータ名は、2 つのインデックスを管理させるストレージ・インデックス管理サーバ 3 の決定に用いられる。

【 0 0 8 3 】

図 1 0 は、サーバ決定処理のフローチャートである。ここで図 1 0 を参照し、上記 S 4 3、及び後述する S 5 1 で実行されるサーバ決定処理について詳細に説明する。インデックス管理サーバ決定処理は、このサーバ決定処理に対し、メタデータ名を引数として渡す。インデックスを管理させるストレージ・インデックス管理サーバ 3 の決定には、図 1 1 に表すようなサーバ一覧表が用いられる。

【 0 0 8 4 】

このサーバ一覧表は、例えば各アプリサーバ 2、及び各ストレージ・インデックス管理サーバ 3 が保持する共有情報である。図 1 1 に表すように、サーバ一覧表には、ストレージ・インデックス管理サーバ 3 毎に、そのストレージ・インデックス管理サーバ 3 に割り当てられたハッシュ値の範囲が定義されている。それにより、インデックス管理サーバ決定処理から渡されるメタデータ名は、ハッシュ値の計算に用いられる。

【 0 0 8 5 】

先ず、C P U 5 1 は、インデックス管理サーバ決定処理から渡されたメタデータ名を用いてハッシュ値を計算する (S 6 1)。次に、C P U 5 1 は、計算したハッシュ値を用いて、インデックスを管理させるストレージ・インデックス管理サーバ 3 を決定し、決定したストレージ・インデックス管理サーバ 3 にインデックスを移動させるための処理ループ S 7 0 を実行する。

【 0 0 8 6 】

この処理ループ S 7 0 では、先ず、C P U 5 1 は、サーバ一覧表中で対象とする 1 台のストレージ・インデックス管理サーバ 3 を選択し、計算したハッシュ値が、選択したストレージ・インデックス管理サーバ 3 に割り当てられたハッシュ値の範囲内か否か判定する (S 7 1)。計算したハッシュ値がその範囲内であった場合、S 7 1 の判定は y e s となって S 7 2 に移行する。計算したハッシュ値がその範囲外であった場合、S 7 1 の判定は n o となり、ここで処理ループ S 7 0 内の一連の処理が終了する。このとき、選択の対象となるストレージ・インデックス管理サーバ 3 が存在しない場合、処理ループ S 7 0 が終了し、その後にサーバ決定処理が終了する。

【 0 0 8 7 】

S 7 2 では、C P U 5 1 は、渡されたメタデータ名から特定されるインデックスを、現在、選択しているストレージ・インデックス管理サーバ 3 に移動させる処理を行う。その処理の実行後、処理ループ S 7 0 内の一連の処理が終了する。

【 0 0 8 8 】

インデックス管理サーバ決定処理から渡されたメタデータ名は、1 種類のメタデータ名か、或いは 2 種類のメタデータ名の結合体である。そのため、移動の対象となるインデッ

10

20

30

40

50

クスは渡されたメタデータ名から特定することができる。

【 0 0 8 9 】

移動対象となるインデックスを管理しているストレージ・インデックス管理サーバ3は、サーバー一覧表から特定される。そのストレージ・インデックス管理サーバ3は、移動対象となるインデックスのメタデータ名を用いて計算されるハッシュ値が、割り当てられたハッシュ値の範囲内となっているストレージ・インデックス管理サーバ3である。それにより、CPU51は、移動対象となるインデックスを管理しているストレージ・インデックス管理サーバ3（以降「移動元サーバ3」と表記）が、計算したハッシュ値を範囲内とするストレージ・インデックス管理サーバ3（以降「移動先サーバ3」と表記）と異なる場合、インデックスを移動させなければならない。そのため、CPU51は、例えば移動元サーバ3に対し、移動先サーバ3、及び移動対象とするインデックスのメタデータ名等を送信して、移動対象とするインデックスの移動先サーバ3への送信を指示する。そのようにして、移動先サーバ3に、移動元サーバ3から送信されるインデックスを保存させる。

10

【 0 0 9 0 】

移動先サーバ3では、インデックスが増えたことにより、それまで送信されなかった検索リクエストがアプリサーバ3から送信される可能性が生じる。そのため、移動先サーバ3では、統計情報37aをクリアしている。

【 0 0 9 1 】

図9のインデックス管理サーバ決定処理の説明に戻る。

20

処理ループS40の終了により、処理ループS50が実行される。処理ループS40の実行により、同じストレージ・インデックス管理サーバ3に管理させるべき2種類のメタデータ名のインデックスは、管理させるべきストレージ・インデックス管理サーバ3に管理される。そのため、処理ループS50でインデックスを移動させる可能性のあるメタデータ名は、インデックス管理情報に登録されていないメタデータ名のみとなる。

【 0 0 9 2 】

その処理ループS50では、まず、CPU51は、統計情報37aを参照して、回数が1以上の1種類のメタデータ名を抽出し、抽出したメタデータ名がインデックス管理情報に登録されていないか否かが判定する（S51）。抽出したメタデータ名がインデックス管理情報に登録されていない場合、S51の判定はyesとなってS52のサーバ決定処理に移行する。このとき、抽出したメタデータ名は引数としてサーバ決定処理に渡される。抽出したメタデータ名がインデックス管理情報に登録されている場合、S51の判定はnoとなり、ここで処理ループS50内の一連の処理が終了する。

30

【 0 0 9 3 】

この一連の処理の終了時、抽出の対象となるメタデータ名が存在しない場合、処理ループS50が終了し、その後にインデックス管理サーバ決定処理が終了する。

【 0 0 9 4 】

1種類のメタデータ名のインデックスは、何れかのストレージ・インデックス管理サーバ3に保存させた後には、サーバー一覧表が更新されたといった特別の理由がない限り、移動させる必要はない。そのため、処理ループS50は、通常、実行されない。しかし、ここでは、1種類のメタデータ名のインデックスが最初に保存されるストレージ・インデックス管理サーバ3を明確にするために、処理ループS50を処理ループS40と共にフローチャートに表している。各アプリサーバ2は、図11に表すようなサーバー連表、及びインデックス管理情報を参照することにより、検索リクエストを送信すべきストレージ・インデックス管理サーバ3を特定することができる。

40

【 0 0 9 5 】

なお、本実施形態では、2種類のメタデータ名のインデックスを同じストレージ・インデックス管理サーバ3に管理させるか否かの判定に共起率rを用いている。その判定には別の指標を用いても良い。或いは別の指標と組み合わせて用いても良い。例えばストレージ・インデックス管理サーバ3全体の負荷の重さと、検索時に同時に指定される2種類以

50

上のメタデータ名との間にパターンが観測される場合、そのパターンをインデックスの管理に反映させても良い。共起率 r の代わりにの指標としては、例えば単位時間当たりの検索件数、が考えられる。共起率 r の代わりにの指標、或いは λ 及び、組み合わせが可能な別の指標としては、オブジェクト、或いはインデックスのデータ量、1 件の検索に要する平均処理時間、1 件の検索の平均的な負荷の重さ、等が考えられる。

【0096】

共起率 r は、サンプル数が少ない状況では、大きく変化し易い。サンプル数が少ない状況時に、共起率 r が閾値より大きくなって、実際には不要な2つのインデックスを同じストレージ・インデックス管理サーバ3に管理させることが有り得る。このことから、共起率 r を計算する対象を制限するための条件を設けても良い。或いはインデックス管理情報に登録された2種類のメタデータ名の組み合わせのなかで、別のストレージ・インデックス管理サーバ3に管理させるべき組み合わせを特定し、インデックスの移動を行うようにしても良い。インデックスの移動の対象とする組み合わせは、閾値より小さい別の閾値を設定し、 $r <$ の関係を満たすことを条件に、特定させることが考えられる。

10

【0097】

また、本実施形態では、各ストレージ・インデックス管理サーバ3は、統計情報37aの生成・更新、及びインデックスを管理させるストレージ・インデックス管理サーバ3の決定を行えるようになっている。統計情報37aの生成・更新、及びインデックスを管理させるストレージ・インデックス管理サーバ3の決定は、それぞれ別の装置に行わせても良い。このことから、管理装置は、複数台の情報処理装置(コンピュータ)を用いて実現させても良い。

20

【0098】

インデックスを移動させるストレージ・インデックス管理サーバ3は、各ストレージ・インデックス管理サーバ3の負荷の重さに着目して決定しても良い。また、インデックスの移動対象とさせるストレージ・インデックス管理サーバ3の優先順位を設定し、設定された優先順位に従って、インデックスを移動させるストレージ・インデックス管理サーバ3を決定するようにしても良い。

上記以外にも、様々な変形を行うことができる。

【0099】

以上の変形例を含む実施形態に関し、更に以下の付記を開示する。

30

(付記1)

ネットワークを介して第1の情報処理装置と接続された複数台の第2の情報処理装置、及び管理装置を備え、

前記複数台の第2の情報処理装置は、

データ、及び前記データのメタデータを用いて検索用に構築されたインデックスを分散して管理し、

前記管理装置は、

前記第1の情報処理装置から送信された検索リクエストを監視する監視部と、

前記監視部による監視結果を基に、同一の第2の情報処理装置に前記インデックスを管理させるべき前記メタデータの種類の組み合わせ、及び前記同一の第2の情報処理装置を特定し、該特定した前記同一の第2の情報処理装置に、該特定した前記組み合わせのメタデータの種類のインデックスを管理させるインデックス管理部と、を有する、

40

ことを特徴とする情報処理システム。

(付記2)

前記監視部が監視した前記検索リクエストで指定されているメタデータの種類の組み合わせ毎に、前記組み合わせでメタデータが同時に指定された結果を表す統計情報を生成する情報生成部、を有し、

前記インデックス管理部は、前記統計情報を基に、前記同一の第2の情報処理装置に前記インデックスを管理させるべき前記メタデータの種類の組み合わせを特定する、

ことを特徴とする付記1記載の情報処理システム。

50

(付記3)

前記情報生成部は、2種類のメタデータの組み合わせ毎に、前記2種類のメタデータが同時に指定された回数を計数して、前記統計情報を生成・更新し、

前記インデックス管理部は、前記2種類のメタデータのインデックスを管理単位として、前記同一の第2の情報処理装置に前記インデックスを管理させるべきか否かの判定を行う、

ことを特徴とする付記2記載の情報処理システム。

(付記4)

ネットワークを介して第1の情報処理装置と接続され、データ、及び前記データのメタデータを用いて検索用に構築されたインデックスを分散して管理する複数台の第2の情報処理装置に対し、前記第1の情報処理装置から送信された検索リクエストを監視する監視部と、

前記監視部による監視結果を基に、同一の第2の情報処理装置に前記インデックスを管理させるべき前記メタデータの種類の組み合わせ、及び前記同一の第2の情報処理装置を特定し、該特定した前記同一の第2の情報処理装置に、該特定した前記組み合わせのメタデータの種類のインデックスを管理させるインデックス管理部と、

を有することを特徴とする管理装置。

(付記5)

ネットワークと接続されたコンピュータに、

データ、及び前記データのメタデータを用いて検索用に構築されたインデックスを分散して管理する複数台の第2の情報処理装置に対し、前記ネットワークを介して第1の情報処理装置から送信された検索リクエストの監視を行い、

前記監視の結果を基に、同一の第2の情報処理装置に前記インデックスを管理させるべき前記メタデータの種類の組み合わせ、及び前記同一の第2の情報処理装置を特定し、該特定した前記同一の第2の情報処理装置に、該特定した前記組み合わせのメタデータの種類のインデックスを管理させる、

処理を実行させる管理プログラム。

(付記6)

ネットワークと接続されたコンピュータを用いて、

データ、及び前記データのメタデータを用いて検索用に構築されたインデックスを分散して管理する複数台の第2の情報処理装置に対し、前記ネットワークを介して第1の情報処理装置から送信された検索リクエストの監視を行い、

前記監視の結果を基に、同一の第2の情報処理装置に前記インデックスを管理させるべき前記メタデータの種類の組み合わせ、及び前記同一の第2の情報処理装置を特定し、該特定した前記同一の第2の情報処理装置に、該特定した前記組み合わせのメタデータの種類のインデックスを管理させる、

ことを特徴とするインデックス管理方法。

【符号の説明】

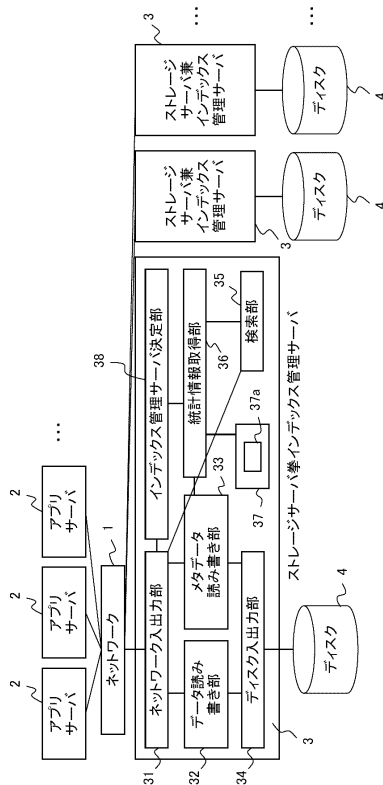
【0100】

- | | | |
|----|---|----|
| 1 | ネットワーク | 40 |
| 2 | アプリケーションサーバ(アプリサーバ) | |
| 3 | ストレージサーバ兼インデックス管理サーバ(ストレージ・インデックス管理サーバ) | |
| 4 | ディスク装置 | |
| 31 | ネットワーク入出力部 | |
| 32 | データ読み書き部 | |
| 33 | メタデータ読み書き部 | |
| 34 | ディスク入出力部 | |
| 35 | 検索部 | |
| 36 | 統計情報取得部 | 50 |

- 3 7 記憶部
- 3 7 a 統計情報
- 3 8 インデックス管理サーバ決定部

【 図 1 】

本実施形態による管理装置を用いて構築された情報処理システムの構成例を説明する図



【 図 2 】

統計情報の構成例を説明する図

	メタデータ1	メタデータ2	...
メタデータ1	メタデータ1が検索された回数	メタデータ1とメタデータ2が同時に検索された回数	...
メタデータ2	未使用	メタデータ2が検索された回数	...
...

【 図 3 】

統計情報の更新例を説明する図

検索条件: 名前=田中 AND 年齢=20

	名前	年齢
(a) 名前	1	1
年齢	-	1

検索条件: 名前=山田

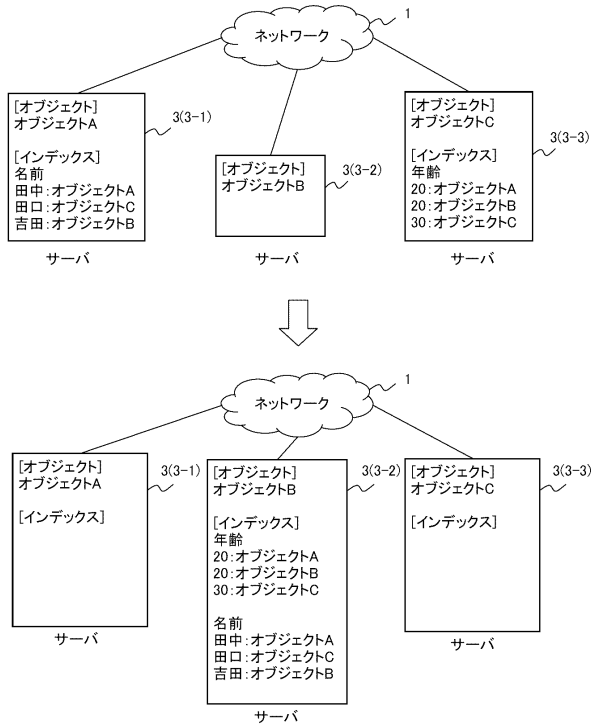
	名前	年齢
(b) 名前	2	1
年齢	-	1

検索条件: 名前=上田 AND 年齢<30

	名前	年齢
(c) 名前	3	2
年齢	-	2

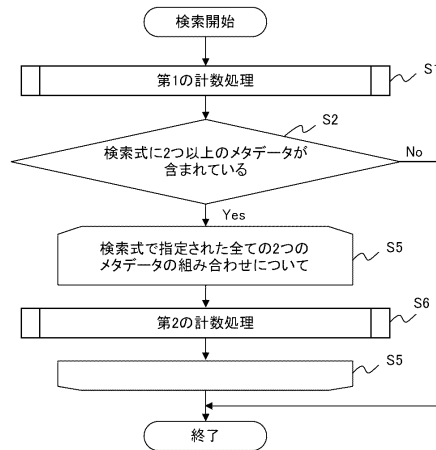
【図4】

インデックスを管理させるストレージサーバ
兼インデックス管理サーバの変更例を説明する図



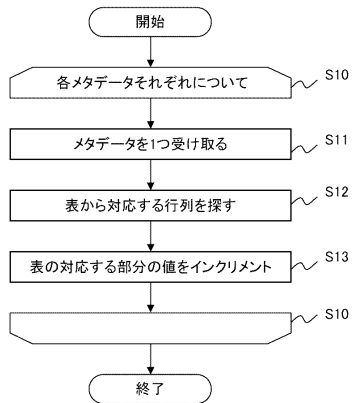
【図5】

統計情報更新処理のフローチャート



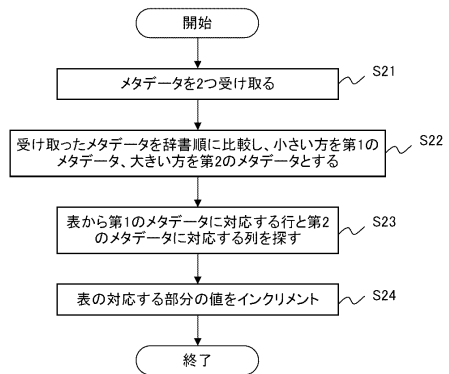
【図6】

第1の計数処理のフローチャート



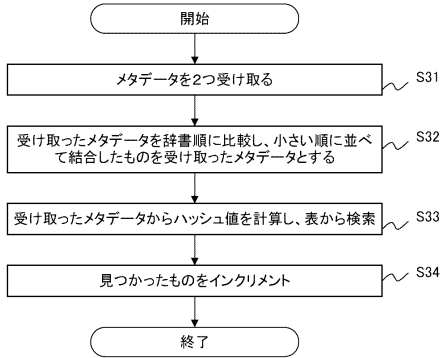
【図7】

第2の計数処理のフローチャート



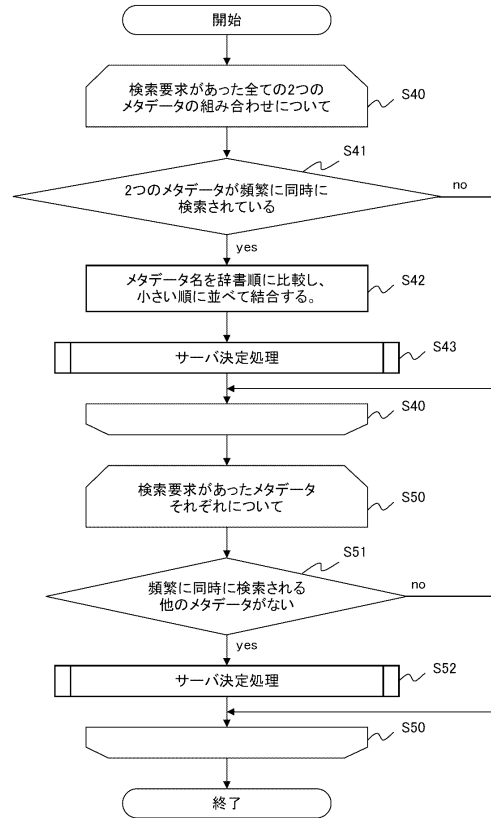
【図8】

第2の計数処理の変形例のフローチャート



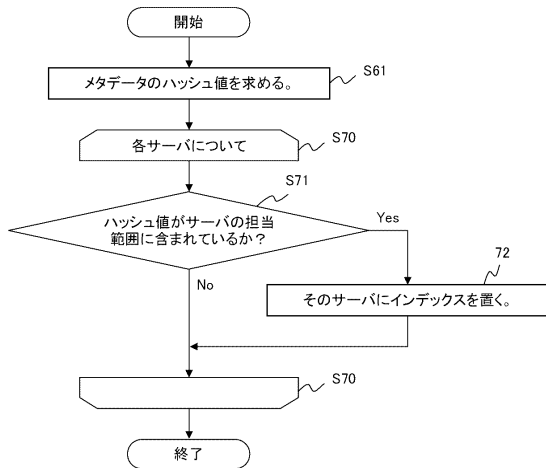
【図9】

インデックス管理サーバ決定処理のフローチャート



【図10】

サーバ決定処理のフローチャート



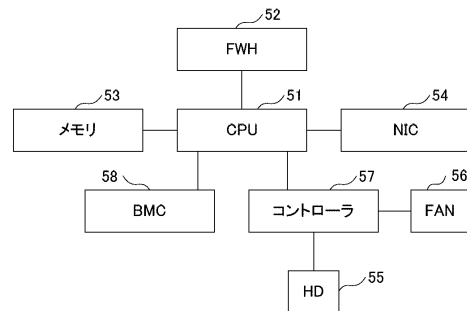
【図11】

サーバ連表の構成例を説明する図

サーバ名	担当範囲
サーバ1	サーバ1が担当するハッシュ値の範囲
サーバ2	サーバ2が担当するハッシュ値の範囲
⋮	⋮

【図12】

本実施形態によるストレージサーバ兼インデックス管理サーバとして使用可能な情報処理装置の構成例を表す図



フロントページの続き

(56)参考文献 特開2011-053868(JP,A)
特開2006-185019(JP,A)
国際公開第2010/116608(WO,A1)

(58)調査した分野(Int.Cl., DB名)
G06F 17/30
G06F 12/00