

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第6155861号
(P6155861)

(45) 発行日 平成29年7月5日 (2017.7.5)

(24) 登録日 平成29年6月16日 (2017.6.16)

(51) Int.Cl.

F I

G O 6 F 12/00 (2006.01)

G O 6 F 13/00 (2006.01)

G O 6 F 12/00 5 4 5 A

G O 6 F 13/00 3 5 7 Z

請求項の数 9 (全 28 頁)

(21) 出願番号	特願2013-120260 (P2013-120260)	(73) 特許権者	000005223
(22) 出願日	平成25年6月6日 (2013.6.6)		富士通株式会社
(65) 公開番号	特開2014-238678 (P2014-238678A)		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26年12月18日 (2014.12.18)	(74) 代理人	100089118
審査請求日	平成28年3月10日 (2016.3.10)		弁理士 酒井 宏明
		(72) 発明者	濱田 圭
			福岡県福岡市早良区百道浜2丁目2番1号
			富士通九州ネットワークテクノロジーズ株式会社内
		(72) 発明者	横山 亮平
			福岡県福岡市早良区百道浜2丁目2番1号
			富士通九州ネットワークテクノロジーズ株式会社内

最終頁に続く

(54) 【発明の名称】 データ管理方法、データ管理プログラム、データ管理システム及びデータ管理装置

(57) 【特許請求の範囲】

【請求項 1】

データを蓄積する複数の第1のノードと、各第1のノードを管理する第2のノードとを有するデータ管理システムが実行するデータ管理方法であって、

前記第1のノードが、

前記第1のノードがそれぞれ所属する複数のエリアが、それぞれ対応するデータを識別するIDの範囲を有し、前記複数のエリアの全エリアにおける前記IDの範囲を示す全エリアのIDの範囲と、前記データを更新する更新リクエストのデータから算出する第1のIDの導出可能範囲とを、前記第2のノードから受信し、

前記更新リクエストを検出した場合に、前記全エリアのIDの範囲と前記導出可能範囲とに基づいて、前記更新リクエストのデータから前記第1のIDを算出し、

第1のIDと第1のノードとを対応付けるノード情報を参照して、前記算出された第1のIDに対応する、前記更新リクエストのデータを記憶する第1のノードを決定し、

前記決定された第1のノードが他の前記エリアに所属する第1のノードである場合には、前記決定された第1のノードに前記更新リクエストを転送する処理を実行し、

前記第2のノードが、前記全エリアの各第1のノードから負荷量を収集し、当該収集された負荷量が所定負荷量を超えた第1のノードを検出した場合に、当該所定負荷量を超えた第1のノードを含むエリアに対応した第2のIDの範囲に、当該エリアに隣接するエリアの前記第2のIDの範囲を追加することで前記導出可能範囲を生成して前記第1のノードに送信し、

10

20

前記決定する処理は、前記第 1 の I D に対応する第 2 の I D と前記全エリアの各第 1 のノードとを対応付ける情報を前記ノード情報として参照して、前記算出された第 1 の I D に対応する、前記更新リクエストのデータを記憶する第 1 のノードを決定する、

ことを特徴とするデータ管理方法。

【請求項 2】

前記第 2 のノードが、エリアごとに、当該エリアごとに割り当てられる、前記第 1 の I D に対応する第 2 の I D の範囲と、該第 2 の I D の範囲に割り当てられる第 1 のノードとを対応付けて前記ノード情報を生成して、予め前記第 1 のノードに送信することを特徴とする請求項 1 に記載のデータ管理方法。

【請求項 3】

前記第 1 のノードが、さらに、

前記データを参照する参照リクエストを検出した場合に、前記参照リクエストのデータの一部を示す文字列と前記ノード情報とに基づいて前記第 2 の I D を算出し、

前記ノード情報を参照して、前記算出された第 2 の I D に対応する、前記参照リクエストのデータを参照する第 1 のノードを決定し、

前記決定された第 1 のノードが他の前記エリアに所属する第 1 のノードである場合には、前記決定された第 1 のノードに前記参照リクエストを転送する、

処理を実行することを特徴とする請求項 1 又は 2 に記載のデータ管理方法。

【請求項 4】

前記第 1 のノードを決定する処理は、

前記更新又は参照リクエストのデータの一部を示す文字列に基づいて、ハッシュ関数により前記更新又は参照リクエストのデータの一次 I D を算出し、

前記一次 I D と、前記ノード情報の第 2 の I D の範囲とに基づいて、前記更新又は参照リクエストのデータの第 2 の I D を算出し、

自ノードの前記データを記憶するデータ記憶部から、前記更新又は参照リクエストのデータの第 2 の I D 及び前記文字列に対応するデータを検索し、

当該データが前記データ記憶部に記憶されている場合は、前記更新又は参照リクエストのデータを記憶又は参照する第 1 のノードを自ノードに決定し、

前記データ記憶部の前記データを更新又は参照することを特徴とする請求項 3 に記載のデータ管理方法。

【請求項 5】

前記第 1 のノードを決定する処理は、

前記データが前記データ記憶部に記憶されていない場合に、転送先の第 1 のノードを記憶する位置記憶部から、前記更新又は参照リクエストのデータの第 2 の I D 及び前記文字列に対応する第 1 のノードを検索し、

当該第 1 のノードが前記位置記憶部に記憶されている場合は、前記更新又は参照リクエストのデータを記憶又は参照する第 1 のノードを、他の前記エリアに所属する第 1 のノードに決定し、

前記転送する処理は、

前記更新又は参照リクエストを決定された第 1 のノードに転送することを特徴とする請求項 4 に記載のデータ管理方法。

【請求項 6】

前記第 1 のノードを決定する処理は、

前記第 1 のノードが前記位置記憶部に記憶されていない場合に、前記一次 I D と前記導出可能範囲とに基づいて、前記更新リクエストのデータの第 1 の I D を算出し、

前記ノード情報を参照して、当該算出された前記更新リクエストのデータの第 1 の I D に対応する第 2 の I D が割り当てられた第 1 のノードを、前記更新リクエストのデータを記憶する第 1 のノードに決定し、

決定された第 1 のノードを前記位置記憶部に記憶し、

前記転送する処理は、

前記決定された第1のノードに対して、前記更新リクエストを転送することを特徴とする請求項5に記載のデータ管理方法。

【請求項7】

データを蓄積する複数の第1のノードと、各第1のノードを管理する第2のノードとを有するデータ管理システムの前記第1のノードであるコンピュータに実行させるデータ管理プログラムであって、

前記コンピュータに、

前記第1のノードがそれぞれ所属する複数のエリアが、それぞれ対応するデータを識別するIDの範囲を有し、前記複数のエリアの全エリアにおける前記IDの範囲を示す全エリアのIDの範囲と、前記データを更新する更新リクエストのデータから算出する第1のIDの導出可能範囲とを、前記第2のノードから受信し、

前記更新リクエストを検出した場合に、前記全エリアのIDの範囲と前記導出可能範囲とに基づいて、前記更新リクエストのデータから前記第1のIDを算出し、

第1のIDと第1のノードとを対応付けるノード情報を参照して、前記算出された第1のIDに対応する、前記更新リクエストのデータを記憶する第1のノードを決定し、

前記決定された第1のノードが他の前記エリアに所属する第1のノードである場合には、前記決定された第1のノードに前記更新リクエストを転送する処理を実行させ、

前記第2のノードでは、前記全エリアの各第1のノードから負荷量を収集し、当該収集された負荷量が所定負荷量を超えた第1のノードを検出した場合に、当該所定負荷量を超えた第1のノードを含むエリアに対応した第2のIDの範囲に、当該エリアに隣接するエリアの前記第2のIDの範囲を追加することで前記導出可能範囲が生成されて前記第1のノードに送信され、

前記決定する処理は、前記第1のIDに対応する第2のIDと前記全エリアの各第1のノードとを対応付ける情報を前記ノード情報として参照して、前記算出された第1のIDに対応する、前記更新リクエストのデータを記憶する第1のノードを決定する、

ことを特徴とするデータ管理プログラム。

【請求項8】

データを蓄積する複数の第1のノードと、各第1のノードを管理する第2のノードとを有するデータ管理システムであって、

前記第2のノードは、

前記第1のノードがそれぞれ所属する複数のエリアが、それぞれ対応するデータを識別するIDの範囲を有し、前記複数のエリアの全エリアにおける前記IDの範囲を示す全エリアのIDの範囲を記憶する記憶部と、

前記データを更新する更新リクエストのデータから算出する第1のIDの導出可能範囲を生成する生成部と、

前記全エリアのIDの範囲と前記導出可能範囲とを前記第1のノードに送信する送信部と、を有し、

前記第1のノードは、

前記全エリアのIDの範囲と前記導出可能範囲とを受信する受信部と、

前記更新リクエストを検出する検出部と、

前記更新リクエストを検出した場合に、前記全エリアのIDの範囲と前記導出可能範囲とに基づいて、前記更新リクエストのデータから前記第1のIDを算出する算出部と、

第1のIDと第1のノードとを対応付けるノード情報を参照して、前記算出された第1のIDに対応する、前記更新リクエストのデータを記憶する第1のノードを決定する決定部と、

前記決定された第1のノードが他の前記エリアに所属する第1のノードである場合には、前記決定された第1のノードに前記更新リクエストを転送する転送部と、を有し、

前記第2のノードの前記生成部は、前記全エリアの各第1のノードから負荷量を収集し、当該収集された負荷量が所定負荷量を超えた第1のノードを検出した場合に、当該所定負荷量を超えた第1のノードを含むエリアに対応した第2のIDの範囲に、当該エリアに

隣接するエリアの前記第 2 の I D の範囲を追加することで前記導出可能範囲を生成し、

前記第 1 のノードの前記決定部は、前記第 1 の I D に対応する第 2 の I D と前記全エリアの各第 1 のノードとを対応付ける情報を前記ノード情報として参照して、前記算出された第 1 の I D に対応する、前記更新リクエストのデータを記憶する第 1 のノードを決定する、

ことを特徴とするデータ管理システム。

【請求項 9】

第 1 のノードがそれぞれ所属する複数のエリアが、それぞれ対応するデータを識別する I D の範囲を有し、前記複数のエリアの全エリアにおける前記 I D の範囲を示す全エリアの I D の範囲と、蓄積したデータを更新する更新リクエストのデータから算出する第 1 の I D の導出可能範囲とを第 2 のノードから受信する受信部と、

前記更新リクエストを検出する検出部と、

前記更新リクエストを検出した場合に、前記全エリアの I D の範囲と前記導出可能範囲とに基づいて、前記更新リクエストのデータから前記第 1 の I D を算出する算出部と、

第 1 の I D と第 1 のノードとを対応付けるノード情報を参照して、前記算出された第 1 の I D に対応する、前記更新リクエストのデータを記憶する第 1 のノードを決定する決定部と、

前記決定された第 1 のノードが他の前記エリアに所属する第 1 のノードである場合には、前記決定された第 1 のノードに前記更新リクエストを転送する転送部と、を有し、

前記第 2 のノードでは、前記全エリアの各第 1 のノードから負荷量を収集し、当該収集された負荷量が所定負荷量を超えた第 1 のノードを検出した場合に、当該所定負荷量を超えた第 1 のノードを含むエリアに対応した第 2 の I D の範囲に、当該エリアに隣接するエリアの前記第 2 の I D の範囲を追加することで前記導出可能範囲が生成されて前記第 1 のノードに送信され、

前記決定部は、前記第 1 の I D に対応する第 2 の I D と前記全エリアの各第 1 のノードとを対応付ける情報を前記ノード情報として参照して、前記算出された第 1 の I D に対応する、前記更新リクエストのデータを記憶する第 1 のノードを決定する、

ことを特徴とするデータ管理装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、データ管理方法、データ管理プログラム、データ管理システム及びデータ管理装置に関する。

【背景技術】

【0002】

大量のデータを収集するデータ収集システムは、1箇所のデータセンタでデータを収集すると、データセンタ側のネットワークがボトルネックとなる。また、データを蓄積するデータベースが単一ノードで構成される場合、データベースの容量または処理性能にスケールビリティがないという問題がある。このため、本発明のデータ収集システムでは、データを収集して蓄積するデータベースを分散化した分散データベース（以下、分散 DB という）を採用したうえで、エリアごとに分散 DB ノード群を構成する。

【先行技術文献】

【特許文献】

【0003】

【特許文献 1】特開 2003 - 216474 号公報

【特許文献 2】特開 2009 - 230686 号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

しかしながら、分散 DB は、あるエリアの分散 DB ノード群に負荷が集中したとしても

10

20

30

40

50

、他のエリアの分散ＤＢノード群にその負荷を分散することができず、分散ＤＢ全体のリソースの利用効率が悪くなる。なお、分散ＤＢは、システム全体として１つの分散ＤＢノード群を有し、１つの管理サーバを設けることもできるが、管理サーバのネットワーク帯域の使用量が増大する。

【０００５】

一つの側面では、リソースの利用効率を向上できるデータ管理方法、データ管理プログラム、データ管理システム及びデータ管理装置を提供することにある。

【課題を解決するための手段】

【０００６】

１つの実施態様では、データ管理システムは、データを蓄積する複数の第１のノードと、各第１のノードを管理する第２のノードとを有する。第１のノードは、全エリアのＩＤの範囲と、前記データを更新する更新リクエストのデータから算出する第１のＩＤの導出可能範囲とを、前記第２のノードから受信する。第１のノードは、前記更新リクエストを検出した場合に、前記全エリアのＩＤの範囲と前記導出可能範囲とに基づいて、前記更新リクエストのデータから前記第１のＩＤを算出する。第１のノードは、第１のＩＤと第１のノードとを対応付けるノード情報を参照して、前記算出された第１のＩＤに対応する、前記更新リクエストのデータを記憶する第１のノードを決定する。第１のノードは、前記決定された第１のノードが他の前記エリアに所属する第１のノードである場合には、前記決定された第１のノードに前記更新リクエストを転送する。

10

【発明の効果】

20

【０００７】

リソースの利用効率を向上できる。

【図面の簡単な説明】

【０００８】

【図１】図１は、実施例のデータ管理システムの構成の一例を示す説明図である。

【図２】図２は、実施例の管理ノードの一例を示すブロック図である。

【図３】図３は、静的エリア情報記憶部の一例を示す説明図である。

【図４】図４は、ノード情報記憶部の一例を示す説明図である。

【図５】図５は、ＩＤノード情報記憶部の一例を示す説明図である。

【図６】図６は、動的エリア情報記憶部の一例を示す説明図である。

30

【図７】図７は、実施例のノードの一例を示すブロック図である。

【図８】図８は、データ記憶部の一例を示す説明図である。

【図９】図９は、位置記憶部の一例を示す説明図である。

【図１０】図１０は、再配置記憶部の一例を示す説明図である。

【図１１】図１１は、実施例の蓄積クライアントの一例を示すブロック図である。

【図１２】図１２は、実施例の解析サーバの一例を示すブロック図である。

【図１３】図１３は、実施例のノードのハードウェア構成の一例を示すブロック図である。

【図１４】図１４は、静的ＩＤとエリア及びノードとの関係の一例を示す説明図である。

【図１５】図１５は、動的ＩＤとエリア及びノードとの関係の一例を示す説明図である。

40

【図１６】図１６は、実施例の初期設定の動作の一例を示すシーケンス図である。

【図１７】図１７は、実施例の更新リクエスト及び参照リクエストの動作の一例を示すシーケンス図である。

【図１８】図１８は、実施例のノードの更新リクエスト受信時における動作の一例を示すフローチャートである。

【図１９】図１９は、実施例のノードの参照リクエスト受信時における動作の一例を示すフローチャートである。

【図２０】図２０は、データ管理プログラムを実行するデータ管理機器の一例を示す説明図である。

【発明を実施するための形態】

50

【 0 0 0 9 】

以下、図面に基づいて、本願の開示するデータ管理方法、データ管理プログラム、データ管理システム及びデータ管理装置の実施例を詳細に説明する。尚、本実施例により、開示技術が限定されるものではない。また、以下の実施例は、矛盾しない範囲で適宜組みあわせてもよい。

【 実施例 1 】

【 0 0 1 0 】

図 1 は、実施例のデータ管理システムの構成の一例を示す説明図である。図 1 に示すデータ管理システム 1 0 は、データセンタ 2 0 と、分散 D B ノード群を表すエリア 1 ~ エリア 4 とを有する。データセンタ 2 0 は、管理ノード 1 0 0 と、解析サーバ 4 0 0 とを有する。エリア 1 ~ 4 は、それぞれノード 2 0 0 と、蓄積クライアント 3 0 0 とを有する。管理ノード 1 0 0 と、エリア 1 ~ 4 の各蓄積クライアント 3 0 0 と、解析サーバ 4 0 0 とは、ネットワーク N を介して、有線により通信可能に接続される。エリア 1 ~ 4 の各ノード 2 0 0 は、所属するエリアの蓄積クライアント 3 0 0 に接続される。なお、例えば、管理ノード 1 0 0 は、第 2 のノードであり、例えば、ノード 2 0 0 は、第 1 のノードである。

10

【 0 0 1 1 】

[1 . 管理ノードの構成]

管理ノード 1 0 0 の構成の一例について説明する。図 2 は、実施例の管理ノードの一例を示すブロック図である。管理ノード 1 0 0 は、通信部 1 1 0 と、記憶部 1 2 0 と、制御部 1 3 0 とを有する。管理ノード 1 0 0 は、データを識別する I D とエリアを識別する情報とを用いて、リクエストに対応するノードを分散する情報を生成する。尚、管理ノード 1 0 0 は、管理ノード 1 0 0 の管理者等から各種操作を受け付ける入力部（例えば、キーボードやマウス等）や、各種情報を表示するための表示部（例えば、液晶ディスプレイ等）を有してもよい。

20

【 0 0 1 2 】

通信部 1 1 0 は、例えば、N I C (Network Interface Card) 等によって実現される。通信部 1 1 0 は、ネットワーク N と有線で接続され、ネットワーク N を介して、ノード 2 0 0 、蓄積クライアント 3 0 0 や解析サーバ 4 0 0 との間で情報の通信を司る通信インタフェースである。

【 0 0 1 3 】

記憶部 1 2 0 は、例えば、R A M (Random Access Memory) 、フラッシュメモリ (Flash Memory) 等の半導体メモリ素子、ハードディスクや光ディスク等の記憶装置によって実現される。記憶部 1 2 0 は、静的エリア情報記憶部 1 2 1 と、ノード情報記憶部 1 2 2 と、I D ノード情報記憶部 1 2 3 と、動的エリア情報記憶部 1 2 4 とを有する。

30

【 0 0 1 4 】

静的エリア情報記憶部 1 2 1 には、各エリアの位置関係を I D 空間に割り当てた情報である静的エリア情報が記憶されている。ここで、図 3 は、静的エリア情報記憶部の一例を示す説明図である。図 3 に示すように、静的エリア情報記憶部 1 2 1 は、エリア 1 2 1 A 、I D 範囲の始点 1 2 1 B 及び I D 範囲の終点 1 2 1 C 等の項目を対応付けて管理している。

40

【 0 0 1 5 】

エリア 1 2 1 A は、地域ごとの分散 D B ノード群を示す。例えば、図 3 では、エリア 1 2 1 A は、エリア 1 ~ エリア 4 の 4 つとなる。I D 範囲の始点 1 2 1 B は、I D 空間に割り当てられた各エリアの始点の静的 I D を示す。I D 範囲の終点 1 2 1 C は、I D 空間に割り当てられた各エリアの終点の静的 I D を示す。なお、静的 I D は、第 2 の I D である。例えば、図 3 では、I D 空間全体を「 1 ~ 1 0 0 0 (0 で表す) 」とし、エリア 1 に I D 「 1 ~ 2 5 0 」を、エリア 2 に I D 「 2 5 1 ~ 5 0 0 」を、エリア 3 に I D 「 5 0 1 ~ 7 5 0 」を、エリア 4 に I D 「 7 5 1 ~ 0 」を割り当てている。

【 0 0 1 6 】

ノード情報記憶部 1 2 2 には、各ノードと、各ノードのホスト名と、各ノードの所属エ

50

リアとを対応付けたノード情報が記憶されている。ここで、図4は、ノード情報記憶部の一例を示す説明図である。図4に示すように、ノード情報記憶部122は、ノード122A、ホスト名122B及び所属エリア122C等の項目を対応付けて管理している。

【0017】

ノード122Aは、各ノードを識別するための識別情報を示す。ホスト名122Bは、各ノードをネットワーク上で特定するための情報を示す。所属エリア122Cは、各ノードが所属するエリアを示す。図4の所属エリア122Cのエリア1には、例えば、“A1”、“A2”及び“A3”のノード200が所属し、エリア2には、例えば、“B1”、“B2”及び“B3”のノード200が所属している。

【0018】

IDノード情報記憶部123には、ID空間の静的ID範囲と各ノードとを対応付けたIDノード情報が記憶されている。なお、IDノード情報は、全エリアのIDの範囲の情報を有し、静的IDと全エリアの各ノード200とを対応付ける。ここで、図5は、IDノード情報記憶部の一例を示す説明図である。図5に示すように、IDノード情報記憶部123は、ID範囲の始点123A、ID範囲の終点123B及びノード123C等の項目を対応付けて管理している。

【0019】

ID範囲の始点123Aは、ID空間に割り当てられた各ノードの始点の静的IDを示す。ID範囲の終点123Bは、ID空間に割り当てられた各ノードの終点の静的IDを示す。ノード123Cは、ID範囲に対応するノードを示す。例えば、図5では、ID「1～80」が“A1”のノード200に、ID「81～160」が“A2”のノード200に、ID「161～250」が“A3”のノード200に対応する。また、ID「251～330」が“B1”のノード200に、ID「331～410」が“B2”のノード200に、ID「411～500」が“B3”のノード200に対応する。

【0020】

動的エリア情報記憶部124には、エリアごとにID空間に動的IDを対応付けて管理する動的エリア情報が記憶されている。なお、動的IDは、第1のIDであり、動的エリア情報は、動的IDの導出可能範囲を表す。ここで、図6は、動的エリア情報記憶部の一例を示す説明図である。図6に示すように、動的エリア情報記憶部124は、エリア124A、ID範囲の始点124B及びID範囲の終点124C等の項目を対応付けて管理している。

【0021】

エリア124Aは、地域ごとの分散DBノード群を示す。例えば、図6では、エリア124Aは、エリア1～エリア4の4つとなる。ID範囲の始点124Bは、ID空間に割り当てられた各エリアの始点の動的IDを示す。ID範囲の終点124Cは、ID空間に割り当てられた各エリアの終点の動的IDを示す。例えば、図6では、ID空間全体を「1～1000（0で表す）」とし、エリア1にID「1～500」を、エリア2にID「251～500」を、エリア3にID「501～750」を、エリア4にID「751～1000」を割り当てている。つまり、ID「251～500」は、エリア2だけでなく、エリア1の動的IDとしても使用できる。

【0022】

図2の説明に戻って、制御部130は、例えば、CPU（Central Processing Unit）やMPU（Micro Processing Unit）等によって、内部の記憶装置に記憶されているプログラムがRAMを作業領域として実行されることにより実現される。また、制御部130は、例えば、ASIC（Application Specific Integrated Circuit）やFPGA（Field Programmable Gate Array）等の集積回路により実現されるようにしても良い。

【0023】

制御部130は、図2に示すように、IDノード情報生成部131と、収集部132と、動的エリア情報生成部133とを有し、以下に説明する情報処理の機能や作用を実現又

10

20

30

40

50

は実行する。尚、制御部 130 の内部構成は、図 2 に示した構成に限られず、後述する情報処理を行う構成であれば他の構成であってもよい。

【0024】

ID ノード情報生成部 131 は、静的エリア情報記憶部 121 に記憶された静的エリア情報と、ノード情報記憶部 122 に記憶されたノード情報とに基づいて、ID ノード情報を生成する。ID ノード情報は、例えば、各ノードについて、ノード名からハッシュ関数を用いて一次 ID を算出して、以下の式 (1) によりノード ID を算出する。次に、全ノードをノード ID 順に並べて、対象ノードの範囲を前ノードのノード ID より大きく、自ノードのノード ID 以下の範囲として、ID ノード情報を生成する。

【0025】

$$\begin{aligned} \text{ノード ID} = & \text{一次 ID} \times \text{ノードの所属エリアの静的 ID 範囲} / \text{全体 ID 範囲} \\ & + \text{ノードの所属エリアの始点の静的 ID} \quad \dots (1) \end{aligned}$$

【0026】

例えば、図 3 の静的エリア情報と、図 4 のノード情報とから、ノード ID を算出する場合を説明する。例えば、“A2” のノード 200 の一次 ID が「636」であった場合のノード ID を算出する。“A2” のノード 200 は、ノード情報よりエリア 1 に所属する。“A2” のノード 200 の所属するエリア 1 の静的 ID 範囲は、静的エリア情報より「1 ~ 250」となる。また、エリア 1 の始点の静的 ID は「1」である。全体 ID 範囲は、静的エリア情報より「1000」である。これらのパラメータを式 (1) に代入すると、ノード ID = $636 \times 250 / 1000 + 1$ となり、ノード ID は、「160」となる。また、“A1” のノード 200 のノード ID が「80」とであるとすると、“A2” のノード 200 に対応する静的 ID 範囲は、「81 ~ 160」となる。

【0027】

ID ノード情報生成部 131 は、生成した ID ノード情報を ID ノード情報記憶部 123 に記憶する。また、ID ノード情報生成部 131 は、静的エリア情報と ID ノード情報とを通信部 110 を介して、各ノード 200、各蓄積クライアント 300 及び解析サーバ 400 に送信する。

【0028】

収集部 132 は、各ノード 200 から送信される負荷情報を、通信部 110 を介して受信して収集する。負荷情報は、各ノード 200 の負荷量を表す情報であり、例えば、各ノード 200 の CPU 使用率やディスク使用量等のリソース使用状況を含むものである。収集部 132 は、収集した負荷情報を動的エリア情報生成部 133 に出力する。

【0029】

動的エリア情報生成部 133 は、例えば、収集部 132 から負荷情報が入力されると、負荷量が所定負荷量を超えたノード 200 があるか否かを判定する。動的エリア情報生成部 133 は、負荷量が所定負荷量を超えたノード 200 を検出した場合、当該ノード 200 が所属するエリアの動的 ID 範囲を生成する。当該エリアの動的 ID 範囲は、所定負荷量を超えたノード 200 を含むエリアに対応した静的 ID 範囲に、当該エリアに隣接するエリアの静的 ID 範囲を追加して割り当てることで生成する。例えば、図 6 では、エリア 1 に所属するノード 200 の負荷量が所定負荷量を超えた場合に、エリア 1 の静的 ID 範囲である ID「1 ~ 250」に、隣接エリア (エリア 2) の静的 ID 範囲である ID「251 ~ 500」をエリア 1 に追加して割り当てる。つまり、エリア 1 は、動的 ID として ID「1 ~ 500」が割り当てられる。動的エリア情報生成部 133 は、エリアごとに動的 ID を対応付けて動的エリア情報を生成する。尚、負荷量が所定負荷量を超えたノード 200 を検出しない場合には、静的エリア情報と同様の ID 範囲を割り当てる。動的エリア情報生成部 133 は、生成した動的エリア情報を動的エリア情報記憶部 124 に記憶するとともに、通信部 110 を介して各ノード 200 に送信する。

【0030】

[2. ノードの構成]

ノード 200 の構成の一例について説明する。図 7 は、実施例のノードの一例を示すブ

10

20

30

40

50

ロック図である。ノード２００は、通信部２１０と、記憶部２２０と、制御部２３０とを有する。ノード２００は、当該ノード２００自体が所属するエリアを管轄する蓄積クライアント３００又は他のノード２００からデータを受信して蓄積する。尚、ノード２００は、ノード２００の管理者等から各種操作を受け付ける入力部（例えば、キーボードやマウス等）や、各種情報を表示するための表示部（例えば、液晶ディスプレイ等）を有してもよい。

【００３１】

通信部２１０は、例えば、ＮＩＣ等によって実現される。通信部２１０は、所属するエリアを管轄する蓄積クライアント３００と有線で接続され、蓄積クライアント３００、及び、蓄積クライアント３００を介して他のノード２００との間で情報の通信を司る通信インタフェースである。尚、通信部２１０は、直接ネットワークＮと有線で接続され、ネットワークＮを介して、蓄積クライアント３００や他のノード２００との間で情報の通信を司る通信インタフェースでとすることができる。

10

【００３２】

記憶部２２０は、例えば、ＲＡＭ、フラッシュメモリ等の半導体メモリ素子、ハードディスクや光ディスク等の記憶装置によって実現される。記憶部２２０は、静的エリア情報記憶部２２１と、ＩＤノード情報記憶部２２２と、動的エリア情報記憶部２２３と、データ記憶部２２４と、位置記憶部２２５と、再配置記憶部２２６とを有する。

【００３３】

静的エリア情報記憶部２２１には、静的エリア情報が記憶されている。静的エリア情報は、管理ノード１００から通信部２１０を経由して受信した、例えば、エリア１２１Ａ、ＩＤ範囲の始点１２１Ｂ及びＩＤ範囲の終点１２１Ｃ等の項目を対応付けて管理している情報である。尚、静的エリア情報は、管理ノード１００の静的エリア情報記憶部１２１に記憶された静的エリア情報の内容と同一である。

20

【００３４】

ＩＤノード情報記憶部２２２には、ＩＤノード情報が記憶されている。ＩＤノード情報は、管理ノード１００から通信部２１０を経由して受信した、例えば、ＩＤ範囲の始点１２３Ａ、ＩＤ範囲の終点１２３Ｂ及びノード１２３Ｃ等の項目を対応付けて管理している情報である。尚、ＩＤノード情報は、管理ノード１００のＩＤノード情報記憶部１２３に記憶されたＩＤノード情報の内容と同一である。

30

【００３５】

動的エリア情報記憶部２２３には、動的エリア情報が記憶されている。動的エリア情報は、管理ノード１００から通信部２１０を経由して受信した、例えば、エリア１２４Ａ、ＩＤ範囲の始点１２４Ｂ及びＩＤ範囲の終点１２４Ｃ等の項目を対応付けて管理している情報である。尚、動的エリア情報は、管理ノード１００の動的エリア情報記憶部１２４に記憶された動的エリア情報の内容と同一である。

【００３６】

データ記憶部２２４には、所属するエリアの蓄積クライアント３００から通信部２１０を介して受信したデータが記憶されている。ここで、図８は、データ記憶部の一例を示す説明図である。図８に示すように、データ記憶部２２４は、静的ＩＤ２２４Ａ及びデータ２２４Ｂ等を対応付けて管理している。また、データ２２４Ｂは、key ２２４Ｃ及びvalue ２２４Ｄを有する。なお、静的ＩＤ２２４Ａは、第２のＩＤである。

40

【００３７】

静的ＩＤ２２４Ａは、ノード２００に記憶するデータ、つまり、蓄積するデータを識別するための識別情報である。静的ＩＤ２２４Ａは、管理ノード１００から通信部２１０を介して受信した静的エリア情報のＩＤ範囲のうちの任意の静的ＩＤである。また、１つの静的ＩＤ２２４Ａには、複数のデータが対応付けて記憶される。データ２２４Ｂは、蓄積するデータを示す。key ２２４Ｃ（以下、キーとも表す）は、蓄積するデータの一部を示す文字列であり、静的ＩＤ２２４Ａと併せて各データを識別する。つまり、静的ＩＤ２２４Ａとkey ２２４Ｃとを用いることで、一意のデータを識別できる。value ２２

50

4 D は、蓄積するデータそのものである。value 2 2 4 D は、例えば、電話回線の通話データである。

【 0 0 3 8 】

位置記憶部 2 2 5 には、ノード 2 0 0 にデータを蓄積するリクエストのうち、リクエストを他のノード 2 0 0 に転送した場合の転送先のノード 2 0 0 が記憶されている。ここで、図 9 は、位置記憶部の一例を示す説明図である。図 9 に示すように、位置記憶部 2 2 5 は、静的 ID 2 2 5 A 及びデータ 2 2 5 B を対応付けて管理している。また、データ 2 2 5 B は、key 2 2 5 C 及び転送先ノード 2 2 5 D を有する。

【 0 0 3 9 】

静的 ID 2 2 5 A は、他のノード 2 0 0 に転送したリクエストに対応するデータを識別するための識別情報である。静的 ID 2 2 5 A は、管理ノード 1 0 0 から通信部 2 1 0 を介して受信した静的エリア情報の ID 範囲のうちの任意の静的 ID である。また、1 つの静的 ID 2 2 5 A には、複数のデータが対応付けて記憶される。データ 2 2 5 B は、リクエストの転送先を示す。key 2 2 5 C は、蓄積するデータの一部を示す文字列であり、データ記憶部 2 2 4 と同様に、静的 ID 2 2 5 A と併せて各データを識別する。転送先ノード 2 2 5 D は、リクエストを転送した先のノード 2 0 0 を示す。転送先ノード 2 2 5 D は、例えば、“B 2 ”等のノード名を用いることができる。

【 0 0 4 0 】

再配置記憶部 2 2 6 には、他のノード 2 0 0 から転送されたリクエストに対応するデータが記憶されている。ここで、図 1 0 は、再配置記憶部の一例を示す説明図である。図 1 0 に示すように、再配置記憶部 2 2 6 は、静的 ID 2 2 6 A 及びデータ 2 2 6 B を対応付けて管理している。また、データ 2 2 6 B は、key 2 2 6 C 及びvalue 2 2 6 D を有する。

【 0 0 4 1 】

静的 ID 2 2 6 A は、他のノード 2 0 0 から転送されたリクエストに対応するデータを識別するための識別情報である。静的 ID 2 2 6 A は、管理ノード 1 0 0 から通信部 2 1 0 を介して受信した静的エリア情報の ID 範囲のうちの任意の静的 ID である。また、1 つの静的 ID 2 2 6 A には、複数のデータが対応付けて記憶される。データ 2 2 6 B は、蓄積するデータを示す。key 2 2 6 C は、蓄積するデータの一部を示す文字列であり、データ記憶部 2 2 4 と同様に、静的 ID 2 2 6 A と併せて各データを識別する。value 2 2 6 D は、蓄積するデータそのものである。value 2 2 6 D は、例えば、電話回線の通話データである。

【 0 0 4 2 】

図 7 の説明に戻って、制御部 2 3 0 は、例えば、CPU や MPU 等によって、内部の記憶装置に記憶されているプログラムが RAM を作業領域として実行されることにより実現される。また、制御部 2 3 0 は、例えば、ASIC や FPGA 等の集積回路により実現されるようにしても良い。

【 0 0 4 3 】

制御部 2 3 0 は、図 7 に示すように、ID 算出部 2 3 1 と、ID 変換部 2 3 2 と、決定部 2 3 3 と、再配置部 2 3 4 と、負荷情報送信部 2 3 5 とを有し、以下に説明する情報処理の機能や作用を実現又は実行する。尚、制御部 2 3 0 の内部構成は、図 7 に示した構成に限られず、後述する情報処理を行う構成であれば他の構成であってもよい。

【 0 0 4 4 】

ID 算出部 2 3 1 は、蓄積クライアント 3 0 0 からデータの更新を行うリクエスト、又は、解析サーバ 4 0 0 からデータの参照を行うリクエストを、通信部 2 1 0 を介して受信すると、当該データの静的 ID を算出する。つまり、ID 算出部 2 3 1 は、例えば、検出部である。ID 算出部 2 3 1 は、リクエストに含まれる当該データのキーと、静的エリア情報記憶部 2 2 1 に記憶されている静的エリア情報とに基づいて、データの静的 ID を算出する。ID 算出部 2 3 1 は、例えば、データのキーからハッシュ関数を用いて一次 ID を算出して、以下の式 (2) によりデータの静的 ID を算出する。なお、データの静的 ID

10

20

30

40

50

Dは、データの第2のIDである。

【0045】

データの静的ID = 一次ID × ノードの所属エリアの静的ID範囲 / 全体ID範囲
+ ノードの所属エリアの始点の静的ID ... (2)

【0046】

例えば、エリア1に配置されるノード200において、図3に示す静的エリア情報と、一次IDとして算出された「800」とに基づいて、データの静的IDを算出する場合を説明する。当該ノード200の所属するエリアの静的ID範囲は、静的エリア情報より「250」となる。また、エリア1の始点の静的IDは「1」である。全体ID範囲は、静的エリア情報より「1000」である。これらのパラメータを式(2)に代入すると、データの静的ID = $800 \times 250 / 1000 + 1$ となり、データの静的IDは、「201」となる。ID算出部231は、生成したデータの静的IDを、ID変換部232及び決定部233に出力する。

10

【0047】

ID変換部232は、決定部233から後述する動的ID生成情報が入力されると、該当するデータの静的IDを動的IDに変換する。ID変換部232は、データの静的IDを算出した一次IDと、動的エリア情報記憶部223に記憶されている動的エリア情報とに基づいて、当該データの動的IDを算出する。つまり、ID変換部232は、例えば、算出部である。ID変換部232は、以下の式(3)によりデータの動的IDを算出する。なお、データの動的IDは、データの第1のIDである。

20

【0048】

データの動的ID = 一次ID × ノードの所属エリアの動的ID範囲 / 全体ID範囲
+ ノードの所属エリアの始点の動的ID ... (3)

【0049】

例えば、エリア1に配置されるノード200において、図6に示す動的エリア情報と、一次IDとして算出された「800」とに基づいて、データの動的IDを算出する場合を説明する。ノード200の所属エリアの動的ID範囲は、動的エリア情報より「500」となる。また、エリア1の始点の動的IDは「1」である。全体ID範囲は、動的エリア情報より「1000」である。これらのパラメータを式(3)に代入すると、データの動的ID = $800 \times 500 / 1000 + 1$ となり、データの動的IDは、「401」となる。ID変換部232は、生成したデータの動的IDを決定部233に出力する。

30

【0050】

決定部233は、蓄積クライアント300からデータの更新を行うリクエスト、又は、解析サーバ400からデータの参照を行うリクエストを、通信部210を介して受信する。また、決定部233は、ID算出部231からデータの静的IDが入力される。決定部233は、リクエストに対応するデータがデータ記憶部224に記憶されているか否かを、リクエストのデータの静的ID及びキーに基づいて検索する。決定部233は、リクエストに対応するデータがデータ記憶部224に記憶されている場合は、データを参照又は更新する。決定部233は、データの参照が完了すると、解析サーバ400に参照レスポンスを送信する。また、決定部233は、データの更新が完了すると、蓄積クライアント300に更新レスポンスを送信する。

40

【0051】

決定部233は、リクエストに対応するデータがデータ記憶部224に記憶されていない場合は、位置記憶部225に転送先のノード200が記憶されているか否かを、リクエストのデータの静的ID及びキーに基づいて検索する。決定部233は、リクエストに対応する転送先のノード200が位置記憶部225に記憶されている場合は、リクエストを検索にヒットした転送先のノード200に通信部210を介して転送する。決定部233は、参照レスポンスを転送先のノード200から受信すると、解析サーバ400に参照レスポンスを転送する。また、決定部233は、更新レスポンスを転送先のノード200から受信すると、蓄積クライアント300に更新レスポンスを転送する。

50

【 0 0 5 2 】

決定部 2 3 3 は、リクエストに対応する転送先のノード 2 0 0 が位置記憶部 2 2 5 に記憶されていない場合は、ID 変換部 2 3 2 に対して動的 ID 生成情報を出力する。決定部 2 3 3 は、ID 変換部 2 3 2 からデータの動的 ID が入力されると、ID ノード情報記憶部 2 2 2 に記憶されている ID ノード情報を参照（検索）する。決定部 2 3 3 は、ID ノード情報の検索の結果に基づいて、データの動的 ID に対応する静的 ID が割り当てられたノード 2 0 0 を、転送先のノード 2 0 0 として決定する。決定部 2 3 3 は、決定したノード 2 0 0 を転送先ノード 2 0 0 として、位置記憶部 2 2 5 に記憶する。また、決定部 2 3 3 は、リクエストを決定した転送先のノード 2 0 0 に通信部 2 1 0 を介して転送する。さらに、決定部 2 3 3 は、転送先のノード 2 0 0 から更新レスポンスを受信すると、更新レスポンスを蓄積クライアント 3 0 0 に転送する。つまり、決定部 2 3 3 は、例えば、検出部、決定部及び転送部である。

10

【 0 0 5 3 】

再配置部 2 3 4 は、他のノード 2 0 0 から通信部 2 1 0 を介して転送されたリクエストを受信する。再配置部 2 3 4 は、転送された更新リクエストを受信すると、リクエストの送信元の蓄積クライアント 3 0 0 に対してデータの送信を許可し、当該データを受信して再配置記憶部 2 2 6 に記憶する。再配置部 2 3 4 は、データの再配置記憶部 2 2 6 への記憶が完了すると、更新レスポンスを転送元のノード 2 0 0 に通信部 2 1 0 を介して送信する。また、再配置部 2 3 4 は、転送された参照リクエストを受信すると、リクエストの送信元の解析サーバ 4 0 0 に対してデータを送信する。再配置部 2 3 4 は、データの送信が完了すると、参照レスポンスを転送元のノード 2 0 0 に通信部 2 1 0 を介して送信する。

20

【 0 0 5 4 】

負荷情報送信部 2 3 5 は、自らのノード 2 0 0 の負荷情報を収集し、通信部 2 1 0 を介して管理ノード 1 0 0 に送信する。負荷情報は、例えば、ノード 2 0 0 の CPU 使用率やディスク使用量等のリソース使用状況である。CPU 使用率及びディスク使用量は、例えば、百分率で表すことができる。他にも、ディスク使用量としては、残り容量を負荷情報として採用してもよい。

【 0 0 5 5 】

[3 . 蓄積クライアントの構成]

蓄積クライアント 3 0 0 の構成の一例について説明する。図 1 1 は、実施例の蓄積クライアントの一例を示すブロック図である。蓄積クライアント 3 0 0 は、通信部 3 1 0 と、記憶部 3 2 0 と、制御部 3 3 0 とを有する。蓄積クライアント 3 0 0 は、所属するエリアのデータを取得すると、取得されたデータの更新リクエストに対応するノード 2 0 0 に送信する。尚、蓄積クライアント 3 0 0 は、蓄積クライアント 3 0 0 の管理者等から各種操作を受け付ける入力部（例えば、キーボードやマウス等）や、各種情報を表示するための表示部（例えば、液晶ディスプレイ等）を有してもよい。

30

【 0 0 5 6 】

通信部 3 1 0 は、例えば、NIC 等によって実現される。通信部 3 1 0 は、ネットワーク N と有線で接続され、ネットワーク N を介して、管理ノード 1 0 0、他のエリア内の蓄積クライアント 3 0 0 や解析サーバ 4 0 0 との間で情報の通信を司る通信インタフェースである。また、通信部 3 1 0 は、同じエリア内のノード 2 0 0 と通信接続し、各ノード 2 0 0 との間で情報の通信を司る通信インタフェースである。

40

【 0 0 5 7 】

記憶部 3 2 0 は、例えば、RAM、フラッシュメモリ等の半導体メモリ素子、ハードディスクや光ディスク等の記憶装置によって実現される。記憶部 3 2 0 は、静的エリア情報記憶部 3 2 1 と、ID ノード情報記憶部 3 2 2 とを有する。

【 0 0 5 8 】

静的エリア情報記憶部 3 2 1 には、静的エリア情報が記憶されている。静的エリア情報は、管理ノード 1 0 0 から通信部 3 1 0 を経由して受信した、例えば、エリア 1 2 1 A、ID 範囲の始点 1 2 1 B 及び ID 範囲の終点 1 2 1 C 等の項目を対応付けて管理している

50

情報である。尚、静的エリア情報は、管理ノード１００の静的エリア情報記憶部１２１に記憶された静的エリア情報の内容と同一である。

【００５９】

ＩＤノード情報記憶部３２２には、ＩＤノード情報が記憶されている。ＩＤノード情報は、管理ノード１００から通信部３１０を経由して受信した、例えば、ＩＤ範囲の始点１２３Ａ、ＩＤ範囲の終点１２３Ｂ及びノード１２３Ｃ等の項目を対応付けて管理している情報である。尚、ＩＤノード情報は、管理ノード１００のＩＤノード情報記憶部１２３に記憶されたＩＤノード情報の内容と同一である。

【００６０】

制御部３３０は、例えば、ＣＰＵやＭＰＵ等によって、内部の記憶装置に記憶されているプログラムがＲＡＭを作業領域として実行されることにより実現される。また、制御部３３０は、例えば、ＡＳＩＣやＦＰＧＡ等の集積回路により実現されるようにしても良い。

10

【００６１】

制御部３３０は、図１１に示すように、ＩＤ算出部３３１と、ノード決定部３３２とを有し、以下に説明する情報処理の機能や作用を実現又は実行する。尚、制御部３３０の内部構成は、図１１に示した構成に限られず、後述する情報処理を行う構成であれば他の構成であってもよい。

【００６２】

ＩＤ算出部３３１は、データの更新を行う場合に、当該データの静的ＩＤを算出する。ＩＤ算出部３３１は、データのキーと、静的エリア情報記憶部３２１に記憶されている静的エリア情報とに基づいて、データの静的ＩＤを算出する。ＩＤ算出部３３１は、ノード２００のＩＤ算出部２３１と同様に、データの静的ＩＤを算出する。ＩＤ算出部３３１は、算出したデータの静的ＩＤをノード決定部３３２に出力する。

20

【００６３】

ノード決定部３３２は、ＩＤ算出部３３１からデータの静的ＩＤが入力されると、データの更新を行う更新リクエストの宛先を決定する。ノード決定部３３２は、データの静的ＩＤと、ＩＤノード情報記憶部３２２に記憶されているＩＤノード情報とに基づいて、更新リクエストの宛先となるノード２００を決定する。ノード決定部３３２は、決定されたノード２００に対して、通信部３１０を介して更新リクエストを送信する。ここで、更新リクエストには、データの一部を示す文字列であるキーも含まれる。また、ノード決定部３３２は、更新リクエストを送信したノード２００から更新レスポンスを受信すると、データの蓄積処理が完了したことを検出する。

30

【００６４】

[４．解析サーバの構成]

解析サーバ４００の構成の一例について説明する。図１２は、実施例の解析サーバの一例を示すブロック図である。解析サーバ４００は、通信部４１０と、記憶部４２０と、制御部４３０とを有する。解析サーバ４００は、各ノード２００に蓄積されたデータを参照して解析を行う。尚、解析サーバ４００は、解析サーバ４００の管理者等から各種操作を受け付ける入力部（例えば、キーボードやマウス等）や、各種情報を表示するための表示部（例えば、液晶ディスプレイ等）を有してもよい。

40

【００６５】

通信部４１０は、例えば、ＮＩＣ等によって実現される。通信部４１０は、ネットワークＮと有線で接続され、ネットワークＮを介して、各エリア内の蓄積クライアント３００や蓄積クライアント３００に接続されている各ノード２００との間で情報の通信を司る通信インタフェースである。また、通信部４１０は、同じデータセンタ２０内の管理ノード１００と通信接続し、管理ノード１００との間で情報の通信を司る通信インタフェースである。

【００６６】

記憶部４２０は、例えば、ＲＡＭ、フラッシュメモリ等の半導体メモリ素子、ハードデ

50

ィスクや光ディスク等の記憶装置によって実現される。記憶部４２０は、静的エリア情報記憶部４２１と、ＩＤノード情報記憶部４２２とを有する。

【００６７】

静的エリア情報記憶部４２１には、静的エリア情報が記憶されている。静的エリア情報は、管理ノード１００から通信部４１０を経由して受信した、例えば、エリア１２１Ａ、ＩＤ範囲の始点１２１Ｂ及びＩＤ範囲の終点１２１Ｃ等の項目を対応付けて管理している情報である。尚、静的エリア情報は、管理ノード１００の静的エリア情報記憶部１２１に記憶された静的エリア情報の内容と同一である。

【００６８】

ＩＤノード情報記憶部４２２には、ＩＤノード情報が記憶されている。ＩＤノード情報は、管理ノード１００から通信部４１０を経由して受信した、例えば、ＩＤ範囲の始点１２３Ａ、ＩＤ範囲の終点１２３Ｂ及びノード１２３Ｃ等の項目を対応付けて管理している情報である。尚、ＩＤノード情報は、管理ノード１００のＩＤノード情報記憶部１２３に記憶されたＩＤノード情報の内容と同一である。

【００６９】

制御部４３０は、例えば、ＣＰＵやＭＰＵ等によって、内部の記憶装置に記憶されているプログラムがＲＡＭを作業領域として実行されることにより実現される。また、制御部４３０は、例えば、ＡＳＩＣやＦＰＧＡ等の集積回路により実現されるようにしても良い。

【００７０】

制御部４３０は、図１２に示すように、ＩＤ算出部４３１と、ノード決定部４３２とを有し、以下に説明する情報処理の機能や作用を実現又は実行する。尚、制御部４３０の内部構成は、図１２に示した構成に限られず、後述する情報処理を行う構成であれば他の構成であってもよい。

【００７１】

ＩＤ算出部４３１は、データの参照を行う場合に、当該データの静的ＩＤを算出する。ＩＤ算出部４３１は、データのキーと、静的エリア情報記憶部４２１に記憶されている静的エリア情報とに基づいて、データの静的ＩＤを算出する。ＩＤ算出部４３１は、ノード２００のＩＤ算出部２３１と同様に、データの静的ＩＤを算出する。ＩＤ算出部４３１は、算出したデータの静的ＩＤをノード決定部４３２に出力する。

【００７２】

ノード決定部４３２は、ＩＤ算出部４３１からデータの静的ＩＤが入力されると、データの参照を行う参照リクエストの宛先を決定する。ノード決定部４３２は、データの静的ＩＤと、ＩＤノード情報記憶部４２２に記憶されているＩＤノード情報に基づいて、参照リクエストの宛先となるノード２００を決定し、決定したノード２００に対して参照リクエストを送信する。

【００７３】

[５．ノード等のハードウェアの構成]

続いて、ノード２００のハードウェア構成の一例について説明する。図１３は、実施例のノードのハードウェア構成の一例を示すブロック図である。

【００７４】

ノード２００は、通信インタフェース２０１と、ＨＤＤ(Hard Disk Drive)２０２と、ドライブ装置２０３と、ＣＰＵ２０４と、メモリ２０５と、入出力装置２０６とを有し、それぞれバス２０７に接続されている。通信インタフェース２０１は、通信部２１０である。ＨＤＤ２０２は、記憶部２２０であり、例えば、ＲＡＩＤ(Redundant Arrays of Inexpensive Disks)を構築することによって、信頼性及び速度等を向上させることができる。

【００７５】

ドライブ装置２０３は、記憶部２２０であり、例えば、光ディスク等を用いることができる。ＣＰＵ２０４は、制御部２３０である。メモリ２０５は、記憶部２２０であり、例

10

20

30

40

50

えば、RAM、フラッシュメモリ等の半導体メモリ素子を用いることができる。入出力装置206は、例えば、入力部（例えば、キーボードやマウス等）や、表示部（液晶ディスプレイ等）である。バス207は、通信インタフェース201と、HDD202と、ドライブ装置203と、CPU204と、メモリ205と、入出力装置206との間で情報の送受信を行うためのものである。尚、説明の便宜上、図13では、ノード200のハードウェア構成について説明したが、例えば、管理ノード100、蓄積クライアント300及び解析サーバ400のハードウェア構成についても同一の構成であるため、その構成及び動作の説明については省略する。

【0076】

[6 . 静的IDとエリア及びノードとの関係]

続いて、静的IDとエリア及びノードとの関係の一例を説明する。図14は、静的IDとエリア及びノードとの関係の一例を示す説明図である。エリア1～エリア4は、例えば、地理的に円環状に位置するものとする。エリア1は、エリア2とエリア4とに隣接する。また、エリア2は、エリア3とエリア1とに隣接する。さらに、エリア3は、エリア4とエリア2とに隣接する。また、エリア4は、エリア1とエリア3とに隣接する。

【0077】

エリア1は、ノード200として、“A1”と、“A2”と、“A3”とを有する。エリア2は、ノード200として、“B1”と、“B2”と、“B3”とを有する。エリア3は、ノード200として、“C1”と、“C2”と、“C3”とを有する。エリア4は、ノード200として、“D1”と、“D2”と、“D3”とを有する。このとき、静的ID範囲は、ID空間全体を「1～1000（0で表す）」とし、エリア1にID「1～250」を、エリア2にID「251～500」を、エリア3にID「501～750」を、エリア4にID「751～0」を割り当てる。さらに、各エリアの静的ID範囲は、各エリア内で各ノード200に割り当てる。エリア1を例にとると、ID「1～80」を“A1”に、ID「81～160」を“A2”に、ID「161～250」を“A3”に割り当てる。

【0078】

[7 . 動的IDとエリア及びノードとの関係]

次に、動的IDとエリア及びノードとの関係の一例を説明する。図15は、動的IDとエリア及びノードとの関係の一例を示す説明図である。エリア1～エリア4は、例えば、静的IDの場合と同様に、地理的に円環状に位置するものとする。また、各エリアに所属するノード200も静的IDの場合と同様であるとする。このとき、動的ID範囲は、ID空間全体を「1～1000（0で表す）」とし、エリア1にID「1～500」を、エリア2にID「251～500」を、エリア3にID「501～750」を、エリア4にID「751～0」を割り当てる。つまり、エリア1には、隣接エリアであるエリア2の静的ID範囲であるID「251～500」を追加して割り当てる。これにより、エリア1の蓄積クライアント300は、エリア2のノード200も使用できる。

【0079】

[8 . データ管理システムの動作]

次に、実施例のデータ管理システム10の動作について説明する。図16は、実施例の初期設定の動作の一例を示すシーケンス図である。まず、管理ノード100は、管理者から静的エリア情報を入力し（ステップS1）、入力された静的エリア情報を静的エリア情報記憶部121に記憶する。続いて、管理ノード100は、管理者からノード情報を入力し（ステップS2）、入力されたノード情報をノード情報記憶部122に記憶する。

【0080】

管理ノード100のIDノード情報生成部131は、静的エリア情報とノード情報とに基づいて、IDノード情報を生成する。IDノード情報生成部131は、静的エリア情報とIDノード情報とを通信部110を介して、各ノード200、各蓄積クライアント300及び解析サーバ400に送信する（ステップS3）。各ノード200、各蓄積クライアント300及び解析サーバ400は、受信した静的エリア情報とノード情報とを、それぞ

10

20

30

40

50

れ静的エリア情報記憶部 2 2 1 , 3 2 1 , 4 2 1、及び、ノード情報記憶部 2 2 2 , 3 2 2 , 4 2 2 に記憶する。

【 0 0 8 1 】

管理ノード 1 0 0 の動的エリア情報生成部 1 3 3 は、負荷量が所定負荷量を超えたノード 2 0 0 を検出した場合、当該ノード 2 0 0 が所属するエリアの動的 ID を生成する。動的エリア情報生成部 1 3 3 は、負荷量が所定負荷量を超えたノード 2 0 0 を検出しない場合は、各エリアに静的エリア情報と同様の ID 範囲を割り当てて、各エリアの動的 ID 範囲を生成する。動的エリア情報生成部 1 3 3 は、生成した動的エリア情報を動的エリア情報記憶部 1 2 4 に記憶するとともに、通信部 1 1 0 を介して各ノード 2 0 0 に送信する（ステップ S 4）。各ノード 2 0 0 は、受信した動的エリア情報を動的エリア情報記憶部 2 2 3 に記憶する。尚、動的エリア情報は、動的エリア情報生成部 1 3 3 が、負荷量が所定負荷量を超えたノード 2 0 0 を検出した場合に、負荷量を反映して再生成され、各ノード 2 0 0 に送信される。

10

【 0 0 8 2 】

続いて、更新リクエスト及び参照リクエスト時のデータ管理システム 1 0 の動作について説明する。図 1 7 は、実施例の更新リクエスト及び参照リクエストの動作の一例を示すシーケンス図である。ここでは、更新リクエストに対応するデータは、“ A 3 ” のノード 2 0 0 及び“ B 2 ” のノード 2 0 0 のどちらにも記憶されておらず、更新リクエストが、“ A 3 ” のノード 2 0 0 から“ B 2 ” のノード 2 0 0 に転送される場合のシーケンスについて説明する。また、参照リクエストに対応するデータは、“ B 2 ” のノード 2 0 0 に記憶されており、参照リクエストが、“ A 3 ” のノード 2 0 0 から“ B 2 ” のノード 2 0 0 に転送される場合のシーケンスについて説明する。

20

【 0 0 8 3 】

まず、更新リクエスト時の動作の一例について説明する。蓄積クライアント 3 0 0 の ID 算出部 3 3 1 は、更新するデータが発生すると、当該データのキーと、静的エリア情報記憶部 3 2 1 に記憶されている静的エリア情報とに基づいて、データの静的 ID を算出する。ID 算出部 3 3 1 は、算出したデータの静的 ID をノード決定部 3 3 2 に出力する。

【 0 0 8 4 】

ノード決定部 3 3 2 は、ID 算出部 3 3 1 からデータの静的 ID が入力されると、データの静的 ID と、ID ノード情報記憶部 3 2 2 に記憶されている ID ノード情報に基づいて、更新リクエストの宛先となる“ A 3 ” のノード 2 0 0 を決定する。ノード決定部 3 3 2 は、決定した“ A 3 ” のノード 2 0 0 に対して更新リクエストを送信する（ステップ S 1 0）。

30

【 0 0 8 5 】

“ A 3 ” のノード 2 0 0 の ID 算出部 2 3 1 は、蓄積クライアント 3 0 0 からデータの更新リクエストを受信すると、データの静的 ID を算出する。ID 算出部 2 3 1 は、リクエストに含まれる当該データのキーと、静的エリア情報記憶部 2 2 1 に記憶されている静的エリア情報とに基づいて、データの静的 ID を算出する。ID 算出部 2 3 1 は、生成したデータの静的 ID を、ID 変換部 2 3 2 及び決定部 2 3 3 に出力する。

【 0 0 8 6 】

40

“ A 3 ” のノード 2 0 0 の決定部 2 3 3 は、蓄積クライアント 3 0 0 からデータの更新リクエストを受信する。また、決定部 2 3 3 は、ID 算出部 2 3 1 からデータの静的 ID が入力される。決定部 2 3 3 は、データ記憶部 2 2 4 を参照して更新リクエストのデータの静的 ID 及びキーを検索し、その検索結果に基づき、更新リクエストに対応するデータがデータ記憶部 2 2 4 に記憶されているか否かを判定する。決定部 2 3 3 は、更新リクエストに対応するデータがデータ記憶部 2 2 4 に記憶されていないため、位置記憶部 2 2 5 を参照して更新リクエストのデータの静的 ID 及びキーに基づいて検索する。決定部 2 3 3 は、その検索結果に基づき、転送先のノード 2 0 0 が位置記憶部 2 2 5 に記憶されているか否かを判定する。

【 0 0 8 7 】

50

決定部 233 は、更新リクエストに対応する転送先のノード 200 が位置記憶部 225 に記憶されていないため、ID変換部 232 に対して動的 ID 生成情報を入力する。ID 変換部 232 は、決定部 233 から動的 ID 生成情報が入力されると、該当するデータの静的 ID を動的 ID に変換する。ID 変換部 232 は、データの静的 ID を算出した一次 ID と、動的エリア情報記憶部 223 に記憶されている動的エリア情報とに基づいて、データの動的 ID を算出する。ID 変換部 232 は、算出したデータの動的 ID を決定部 233 に出力する。

【0088】

決定部 233 は、ID 変換部 232 からデータの動的 ID が入力されると、ID ノード情報記憶部 222 に記憶されている ID ノード情報を参照して、データの動的 ID に対応する静的 ID が割り当てられた“B2”のノード 200 を決定する。決定部 233 は、決定した“B2”のノード 200 を転送先ノード“B2”として、位置記憶部 225 に記憶する。また、決定部 233 は、更新リクエストを決定した転送先の“B2”のノード 200 に転送する（ステップ S11）。

10

【0089】

“B2”のノード 200 の再配置部 234 は、“A3”のノード 200 から転送された更新リクエストを受信すると、更新リクエストの送信元の蓄積クライアント 300 に対してデータの送信を許可し、データを受信して再配置記憶部 226 に記憶する。再配置部 234 は、データの再配置記憶部 226 への記憶が完了すると、更新レスポンスを転送元の“A3”のノード 200 に送信する（ステップ S12）。

20

【0090】

“A3”のノード 200 の決定部 233 は、転送先の“B2”のノード 200 から更新レスポンスを受信すると、更新レスポンスを蓄積クライアント 300 に転送する（ステップ S13）。蓄積クライアント 300 は、更新レスポンスを受信すると、データの蓄積処理が完了したことを検出する。

【0091】

次に、参照リクエスト時の動作の一例について説明する。解析サーバ 400 の ID 算出部 431 は、データの参照要求が発生すると、当該データのキーと、静的エリア情報記憶部 421 に記憶されている静的エリア情報とに基づいて、データの静的 ID を算出する。ID 算出部 431 は、算出したデータの静的 ID をノード決定部 432 に出力する。

30

【0092】

ノード決定部 432 は、ID 算出部 431 からデータの静的 ID が入力されると、データの静的 ID と、ID ノード情報記憶部 422 に記憶されている ID ノード情報とに基づいて、参照リクエストの宛先となる“A3”のノード 200 を決定する。ノード決定部 432 は、決定された“A3”のノード 200 に対して参照リクエストを送信する（ステップ S14）。

【0093】

“A3”のノード 200 の ID 算出部 231 は、解析サーバ 400 からデータの参照リクエストを受信すると、データの静的 ID を算出する。ID 算出部 231 は、参照リクエストに含まれる当該データのキーと、静的エリア情報記憶部 221 に記憶されている静的エリア情報とに基づいて、データの静的 ID を算出する。ID 算出部 231 は、算出されたデータの静的 ID を ID 変換部 232 及び決定部 233 に出力する。

40

【0094】

“A3”のノード 200 の決定部 233 は、解析サーバ 400 からデータの参照リクエストを受信する。また、決定部 233 は、ID 算出部 231 からデータの静的 ID を入力する。決定部 233 は、データ記憶部 224 を参照して参照リクエストのデータの静的 ID 及びキーを検索し、その検索結果に基づき、参照リクエストに対応するデータがデータ記憶部 224 に記憶されているか否かを判定する。決定部 233 は、参照リクエストに対応するデータがデータ記憶部 224 に記憶されていないため、位置記憶部 225 を参照して参照リクエストのデータの静的 ID 及びキーに基づいて検索する。決定部 233 は、そ

50

の検索結果に基づき、転送先のノード２００が位置記憶部２２５に記憶されているか否かを判定する。

【００９５】

決定部２３３は、参照リクエストに対応する転送先の“Ｂ２”のノード２００が位置記憶部２２５に記憶されている場合、参照リクエストを転送先の“Ｂ２”のノード２００に転送する（ステップＳ１５）。

【００９６】

“Ｂ２”のノード２００の再配置部２３４は、“Ａ３”のノード２００から転送された参照リクエストを受信すると、参照リクエストに対応するデータを再配置記憶部２２６から読み出して、解析サーバ４００に送信する。再配置部２３４は、データの解析サーバ４００への送信が完了すると、参照レスポンスを転送元の“Ａ３”のノード２００に送信する（ステップＳ１６）。

10

【００９７】

“Ａ３”のノード２００の決定部２３３は、転送先の“Ｂ２”のノード２００から参照レスポンスを受信すると、参照レスポンスを解析サーバ４００に転送する（ステップＳ１７）。解析サーバ４００は、参照レスポンスを受信すると、データの読み出し処理が完了したことを検出する。

【００９８】

続いて、更新リクエスト受信時のノード２００の動作の詳細について説明する。図１８は、実施例のノードの更新リクエスト受信時における動作の一例を示すフローチャートである。

20

【００９９】

ノード２００のＩＤ算出部２３１は、蓄積クライアント３００からデータの更新を行うリクエストを受信する（ステップＳ１０１）。ＩＤ算出部２３１は、リクエストに含まれる当該データのキーと、静的エリア情報記憶部２２１に記憶されている静的エリア情報とに基づいて、データの静的ＩＤを算出する。ＩＤ算出部２３１は、算出したデータの静的ＩＤを、ＩＤ変換部２３２及び決定部２３３に出力する。

【０１００】

決定部２３３は、蓄積クライアント３００からデータの更新を行うリクエストを受信する。また、決定部２３３は、ＩＤ算出部２３１からデータの静的ＩＤが入力される。決定部２３３は、データ記憶部２２４を参照して更新リクエストのデータの静的ＩＤ及びキーを検索する（ステップＳ１０２）。決定部２３３は、その検索結果に基づき、更新リクエストに対応するデータがヒットしたか否かを判定する（ステップＳ１０３）。決定部２３３は、更新リクエストに対応するデータがヒットした場合には（ステップＳ１０３：肯定）、データ記憶部２２４に記憶されているデータを更新する（ステップＳ１０４）。決定部２３３は、データの更新が完了すると、蓄積クライアント３００に更新レスポンスを送信する（ステップＳ１１６）。

30

【０１０１】

決定部２３３は、更新リクエストに対応するデータがヒットしない場合には（ステップＳ１０３：否定）、位置記憶部２２５を参照して参照リクエストのデータの静的ＩＤ及びキーに基づいて検索する（ステップＳ１０５）。決定部２３３は、位置記憶部２２５の検索の結果に基づいて、更新リクエストに対応する転送先のノード２００がヒットしたか否かを判定する（ステップＳ１０６）。決定部２３３は、更新リクエストに対応する転送先のノード２００がヒットした場合には（ステップＳ１０６：肯定）、更新リクエストを転送先のノード２００に転送する（ステップＳ１０７）。決定部２３３は、更新レスポンスを転送先のノード２００から受信すると（ステップＳ１０８）、蓄積クライアント３００に更新レスポンスを転送すべく、ステップＳ１１６に移行する。

40

【０１０２】

決定部２３３は、更新リクエストに対応する転送先のノード２００がヒットしない場合には（ステップＳ１０６：否定）、ＩＤ変換部２３２に対して動的ＩＤ生成情報を出力す

50

る。ID変換部232は、決定部233から動的ID生成情報が入力されると、データの静的IDを算出した一次IDと、動的エリア情報記憶部223に記憶されている動的エリア情報とに基づいて、データの動的IDを算出する(ステップS109)。ID変換部232は、算出したデータの動的IDを決定部233に出力する。

【0103】

決定部233は、ID変換部232からデータの動的IDが入力されると、IDノード情報記憶部222内のIDノード情報を検索する(ステップS110)。決定部233は、IDノード情報の検索の結果に基づいて、データの動的IDに対応する静的IDがヒットしたか否かを判定する(ステップS111)。決定部233は、データの動的IDに対応する静的IDがヒットした場合には(ステップS111：肯定)、データの動的IDに対応する静的IDが割り当てられたノード200を、転送先のノード200として決定する。また、決定部233は、決定したノード200を転送先ノード200として、位置記憶部225にエントリを追加して記憶する(ステップS112)。さらに、決定部233は、更新リクエストを転送先のノード200に転送する(ステップS113)。決定部233は、更新レスポンスを転送先のノード200から受信すると(ステップS114)、蓄積クライアント300に更新レスポンスを転送すべく、ステップS116に移行する。

10

【0104】

決定部233は、データの動的IDに対応する静的IDがヒットしない場合には(ステップS111：否定)、エラー処理を実行する(ステップS115)。決定部233は、蓄積クライアント300に対して、エラー処理の結果を更新レスポンスとして送信すべく、ステップS116に移行する。

20

【0105】

ノード200は、データ記憶部224を参照して更新リクエストに対応するデータがヒットした場合には、当該データ記憶部224に記憶されているデータを更新する。その結果、エリアを跨ぐリクエストの転送トラヒックの発生を防止することができる。

【0106】

また、ノード200は、位置記憶部225を参照して更新リクエストに対応する転送先のノード200がヒットした場合には、更新リクエストを転送先のノード200に転送する。その結果、蓄積クライアント300は、更新リクエストの転送の有無に関わらず、転送元のノード200にリクエストを送信することで、転送先のノード200のデータを更新することができる。

30

【0107】

また、ノード200は、IDノード情報記憶部222を参照してデータの動的IDに対応する静的IDがヒットした場合には、データの動的IDに対応する静的IDが割り当てられたノード200を、転送先のノード200として決定する。その結果、各エリアで使用するノード200を、動的エリア情報を用いてエリアごとの動的IDの範囲で管理するので、分散DBノード群の管理コストを低減することができる。

【0108】

次に、参照リクエスト受信時のノード200の動作の詳細について説明する。図19は、実施例のノードの参照リクエスト受信時における動作の一例を示すフローチャートである。

40

【0109】

ノード200のID算出部231は、解析サーバ400からデータの参照を行うリクエストを受信する(ステップS201)。ID算出部231は、参照リクエストに含まれる当該データのキーと、静的エリア情報記憶部221に記憶されている静的エリア情報とに基づいて、データの静的IDを算出する。ID算出部231は、算出したデータの静的IDを、ID変換部232及び決定部233に出力する。

【0110】

決定部233は、解析サーバ400からデータの参照を行うリクエストを受信する。また、決定部233は、ID算出部231からデータの静的IDが入力される。決定部23

50

3は、データ記憶部224を参照して参照リクエストのデータの静的ID及びキーを検索する(ステップS202)。決定部233は、その検索結果に基づき、参照リクエストに対応するデータがヒットしたか否かを判定する(ステップS203)。決定部233は、参照リクエストに対応するデータがヒットした場合には(ステップS203：肯定)、データ記憶部224に記憶されているデータを読み出して、解析サーバ400に送信する(ステップS204)。決定部233は、データの送信が完了すると、解析サーバ400に更新レスポンスを送信する(ステップS210)。

【0111】

決定部233は、参照リクエストに対応するデータがヒットしない場合には(ステップS203：否定)、位置記憶部225を参照して参照リクエストのデータの静的ID及びキーに基づいて検索する(ステップS205)。決定部233は、その検索結果に基づき、転送先のノード200がヒットしたか否かを判定する(ステップS206)。決定部233は、参照リクエストに対応する転送先のノード200がヒットした場合には(ステップS206：肯定)、参照リクエストを転送先のノード200に転送する(ステップS207)。決定部233は、参照レスポンスを転送先のノード200から受信すると(ステップS208)、解析サーバ400に参照レスポンスを転送すべく、ステップS210に移行する。

【0112】

決定部233は、参照リクエストに対応する転送先のノード200がヒットしない場合には(ステップS206：否定)、エラー処理を実行する(ステップS209)。決定部233は、解析サーバ400に対して、エラー処理の結果を参照レスポンスとして送信すべく、ステップS210に移行する。

【0113】

ノード200は、データ記憶部224を参照して参照リクエストに対応するデータがヒットした場合には、当該データ記憶部224に記憶されているデータを参照する。その結果、エリアを跨ぐリクエストの転送トラヒックの発生を防止することができる。

【0114】

また、ノード200は、位置記憶部225を参照して参照リクエストに対応する転送先のノード200がヒットした場合には、参照リクエストを転送先のノード200に転送する。その結果、解析サーバ400は、参照リクエストの転送の有無に関わらず、転送元のノード200にリクエストを送信することで、転送先のノード200のデータを更新することができる。

【0115】

[9. 効果及びその他]

このように、データ管理システム10は、ノード200が、IDノード情報と動的エリア情報とを管理ノード100から受信する。ノード200は、更新リクエストを検出した場合に、IDノード情報と動的エリア情報とに基づいて、更新リクエストのデータから動的IDを算出する。ノード200は、IDノード情報を参照して、算出された動的IDに対応する、更新リクエストのデータを記憶するノード200を決定する。ノード200は、決定されたノード200が他のエリアに所属するノード200である場合には、決定されたノード200に更新リクエストを転送する。その結果、データ管理システム10は、リソースの利用効率を向上できる。

【0116】

また、管理ノード100が、全エリアの各ノード200から負荷量を収集し、当該収集された負荷量が所定負荷量を超えたノード200を検出した場合に、当該所定負荷量を超えたノード200を含むエリアに対応した静的IDの範囲に、当該エリアに隣接するエリアの静的IDの範囲を追加することで動的エリア情報を生成する。その結果、データ管理システム10は、隣接するエリアのリソースを利用できる。

【0117】

また、管理ノード100が、エリアごとに、当該エリアごとに割り当てられる、動的I

10

20

30

40

50

Dに対応する静的IDの範囲と、該静的IDの範囲に割り当てられるノード200とを対応付けてIDノード情報を生成する。その結果、データ管理システム10は、静的IDの範囲に応じてエリアと所属するノード200が決定できる。

【0118】

また、ノード200は、データを参照する参照リクエストを検出した場合に、参照リクエストのデータの一部を示す文字列とIDノード情報とに基づいて静的IDを算出する。ノード200は、IDノード情報を参照して、算出された静的IDに対応する、参照リクエストのデータを参照するノード200を決定する。ノード200は、更新リクエストを検出した場合と同様に、決定されたノード200に参照リクエストを転送する。その結果、データ管理システム10は、リソースの利用効率を向上できる。

10

【0119】

また、ノード200は、更新又は参照リクエストのデータの一部を示す文字列に基づいて、ハッシュ関数により更新又は参照リクエストのデータの一次IDを算出する。ノード200は、一次IDとIDノード情報の静的IDの範囲とに基づいて、更新又は参照リクエストのデータの静的IDを算出する。ノード200は、データ記憶部224から、更新又は参照リクエストのデータの静的ID及び文字列に対応するデータを検索する。ノード200は、当該データがデータ記憶部224に記憶されている場合は、更新又は参照リクエストのデータを記憶又は参照するノード200を自ノードに決定する。ノード200は、データ記憶部224のデータを更新又は参照する。その結果、データ管理システム10は、全てのエリアの分散DBノード群の負荷が低い場合に、各エリア内のノード200のみでデータを蓄積することができる。このため、データ管理システム10は、エリアを跨ぐリクエストの転送トラヒックの発生を防止することができる。

20

【0120】

また、ノード200は、更新又は参照リクエストのデータがデータ記憶部224に記憶されていない場合に、転送先のノード200を記憶する位置記憶部225から、更新又は参照リクエストのデータの静的ID及び文字列に対応するノード200を検索する。ノード200は、検索したノード200が位置記憶部225に記憶されている場合は、更新又は参照リクエストのデータを記憶又は参照するノード200を、他のエリアに所属するノード200に決定する。ノード200は、更新又は参照リクエストを決定されたノード200に転送する。その結果、データ管理システム10は、負荷に応じて他のエリアのノード200に転送したデータを、転送元のノード200に更新又は参照リクエストを送信することで更新又は参照することができる。

30

【0121】

また、ノード200は、転送先のノード200が位置記憶部225に記憶されていない場合に、一次IDと動的エリア情報とに基づいて、更新リクエストのデータの動的IDを算出する。ノード200は、IDノード情報を参照して、当該算出された更新リクエストのデータの動的IDに対応する静的IDが割り当てられたノード200を、更新リクエストのデータを記憶するノード200に決定する。ノード200は、決定されたノード200を位置記憶部225に記憶すると共に、決定されたノード200に対して、更新リクエストを転送する。その結果、データ管理システム10は、転送元のノード200が転送したデータを管理できるので、データの位置管理をスケーラブルとすることができる。また、データ管理システム10は、各エリアで使用するノード200を、動的エリア情報を用いてエリアごとの動的IDの範囲で管理するので、分散DBノード群の管理コストを低減することができる。

40

【0122】

また、上記実施例では、エリアの数を4個、各エリアに所属するノード200の数を3個としたが、これに限定されない。エリア及び各エリアのノード200の数は、例えば、蓄積するデータの量に応じて適宜増減することができる。

【0123】

また、上記実施例では、静的ID及び動的IDの範囲を「1～1000」としたが、こ

50

れに限定されない。静的ID及び動的IDの範囲は、例えば、各エリアのノード200の個数に応じて、適宜増減することができる。

【0124】

また、図示した各部の各構成要素は、必ずしも物理的に図示の如く構成されていることを要しない。すなわち、各部の分散・統合の具体的な形態は図示のものに限られず、その全部又は一部を、各種の負荷や使用状況等に応じて、任意の単位で機能的又は物理的に分散・統合して構成することができる。

【0125】

更に、各部で行われる各種処理機能は、CPU（又はMPU（Micro Processing Unit）、MCU（Micro Controller Unit）等のマイクロ・コンピュータ）上で、その全部又は任意の一部を実行するようにしてもよい。また、各種処理機能は、CPU（又はMPU、MCU等のマイクロ・コンピュータ）で解析実行されるプログラム上、又はワイヤードロジックによるハードウェア上で、その全部又は任意の一部を実行するようにしてもよいことは言うまでもない。

【0126】

ところで、上記の実施例で説明した各種の処理は、予め用意されたプログラムをデータ管理機器で実行することで実現できる。そこで、以下では、上記の実施例と同様の機能を有するプログラムを実行するデータ管理機器の一例を説明する。図20は、データ管理プログラムを実行するデータ管理機器の一例を示す説明図である。

【0127】

図20に示すデータ管理プログラムを実行するデータ管理機器500は、インタフェース部511と、RAM512と、ROM（Read Only Memory）513と、プロセッサ514とを有する。インタフェース部511は、管理機器、蓄積機器、解析機器及び他のデータ管理機器と通信する。プロセッサ514は、データ管理機器500全体を制御する。

【0128】

そして、ROM513には、上記実施例と同様の機能を発揮するデータ管理プログラムが予め記憶されている。尚、ROM513ではなく、図示せぬドライブで読取可能な記録媒体にデータ管理プログラムが記録されていてもよい。また、記録媒体としては、例えば、CD-ROM、DVDディスク、USBメモリ等の可搬型記録媒体、フラッシュメモリ等の半導体メモリ等でもよい。データ管理プログラムとしては、図20に示すように、検出プログラム513A、算出プログラム513B、決定プログラム513C及び転送プログラム513Dである。尚、プログラム513A～513Dについては、適宜統合又は分散してもよい。また、RAM512には、静的エリア情報、IDノード情報、動的エリア情報、蓄積データ、転送した蓄積データの位置及び転送された蓄積データ等を格納するデータベースが記憶してある。

【0129】

そして、プロセッサ514が、これらのプログラム513A～513DをROM513から読み出し、これらの読み出された各プログラムを実行する。そして、プロセッサ514は、図20に示すように、各プログラム513A～513Dを、検出プロセス514A、算出プロセス514B、決定プロセス514C及び転送プロセス514Dとして機能することになる。

【0130】

データ管理機器500は、全エリアのIDの範囲と、蓄積したデータを更新する更新リクエストのデータから算出する第1のIDの導出可能範囲とを受信する。プロセッサ514は、更新リクエストを検出する。プロセッサ514は、更新リクエストを検出した場合に、全エリアのIDの範囲と導出可能範囲とに基づいて、更新リクエストのデータから第1のIDを算出する。プロセッサ514は、第1のIDと第1のノードとを対応付けるノード情報を参照して、算出された第1のIDに対応する、更新リクエストのデータを記憶する第1のノードを決定する。プロセッサ514は、決定されたノードが他のエリアに所属するノードである場合には、決定されたノードに更新リクエストを転送する。その結果

10

20

30

40

50

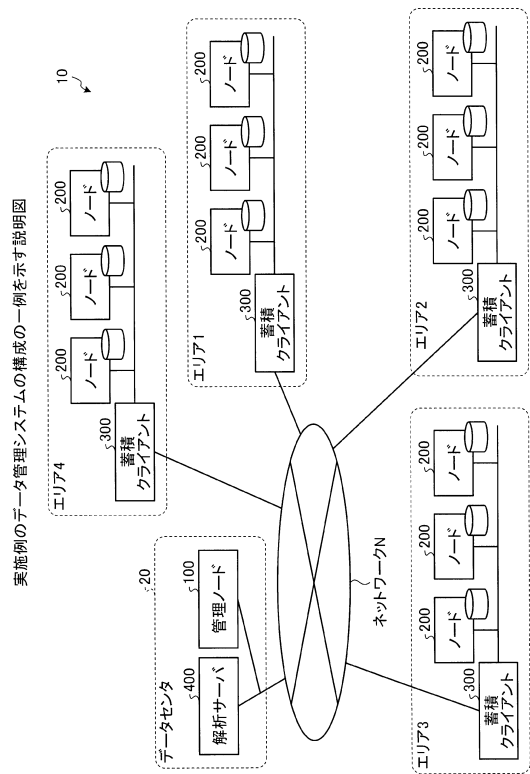
、リソースの利用効率を向上できる。

【符号の説明】

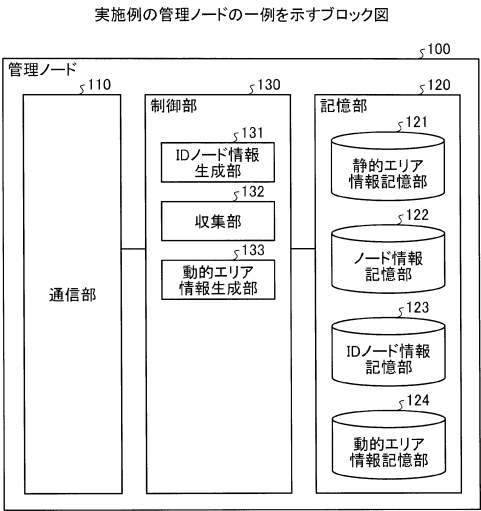
【0131】

10	データ管理システム	
20	データセンタ	
100	管理ノード	
110, 210, 310, 410	通信部	
120, 220, 320, 420	記憶部	
121, 221, 321, 421	静的エリア情報記憶部	
122	ノード情報記憶部	10
123, 222, 322, 422	IDノード情報記憶部	
124, 223	動的エリア情報記憶部	
130, 230, 330, 430	制御部	
131	IDノード情報生成部	
132	収集部	
133	動的エリア情報生成部	
200	ノード	
201	通信インタフェース	
202	HDD	
203	ドライブ装置	20
204	CPU	
205	メモリ	
206	入出力装置	
207	バス	
224	データ記憶部	
225	位置記憶部	
226	再配置記憶部	
231, 331, 431	ID算出部	
232	ID変換部	
233	決定部	30
234	再配置部	
235	負荷情報送信部	
300	蓄積クライアント	
332, 432	ノード決定部	
400	解析サーバ	

【図 1】



【図 2】



【図 3】

静的エリア情報記憶部の一例を示す説明図

エリア	ID範囲	
	始点	終点
エリア1	1	250
エリア2	251	500
エリア3	501	750
エリア4	751	0(999+1)

【図 4】

ノード情報記憶部の一例を示す説明図

ノード	ホスト名	所属エリア
A1	node1.area1.net.	エリア1
A2	node2.area1.net.	エリア1
A3	node3.area1.net.	エリア1
B1	node1.area2.net.	エリア2
B2	node2.area2.net.	エリア2
B3	node3.area2.net.	エリア2
...

【図 6】

動的エリア情報記憶部の一例を示す説明図

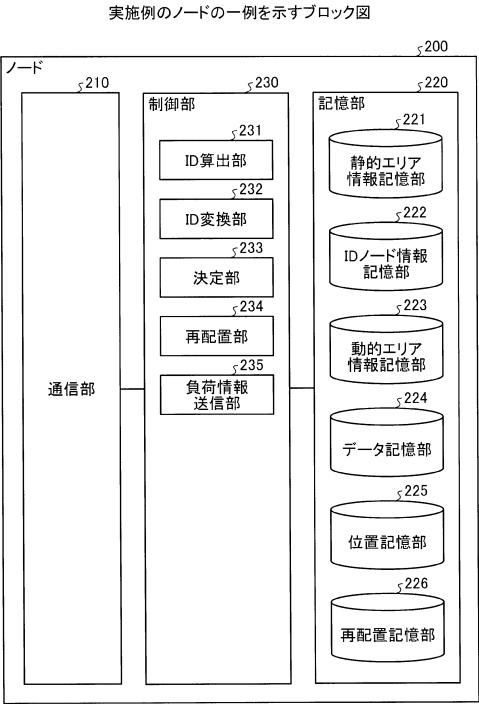
エリア	ID範囲	
	始点	終点
エリア1	1	500
エリア2	251	500
エリア3	501	750
エリア4	751	0(999+1)

【図 5】

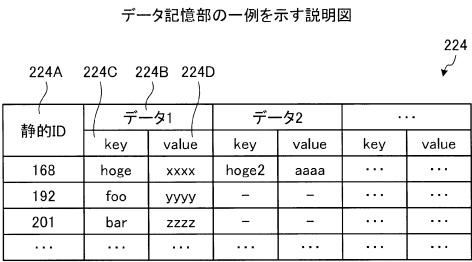
IDノード情報記憶部の一例を示す説明図

ID範囲		ノード
始点	終点	
1	80	A1
81	160	A2
161	250	A3
251	330	B1
331	410	B2
411	500	B3
...

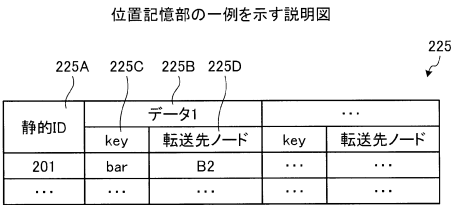
【図 7】



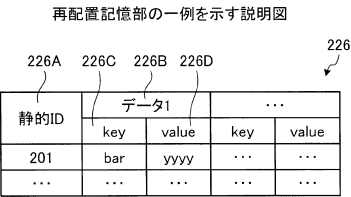
【図 8】



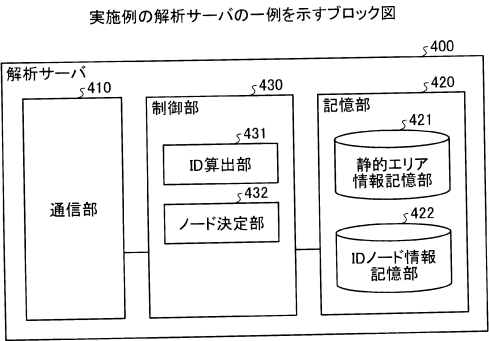
【図 9】



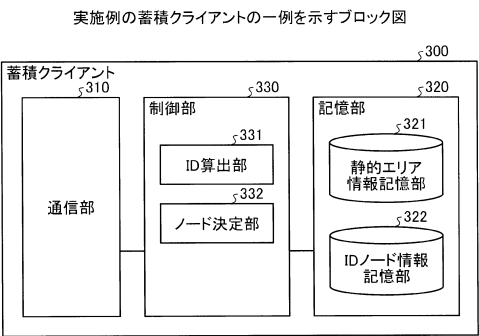
【図 10】



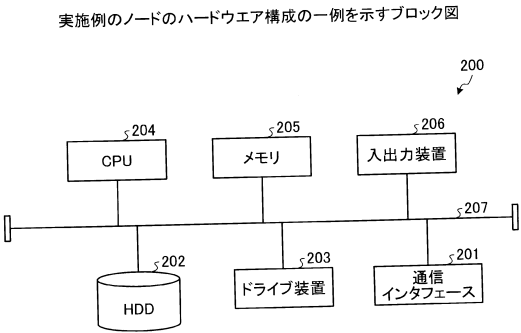
【図 12】



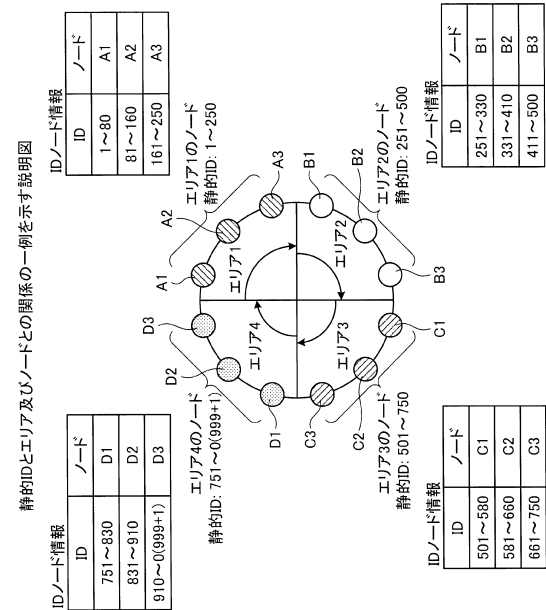
【図 11】



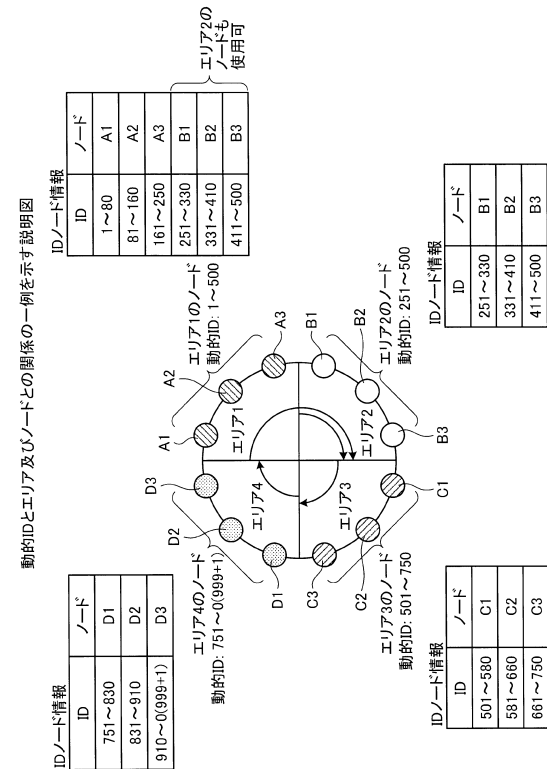
【図 13】



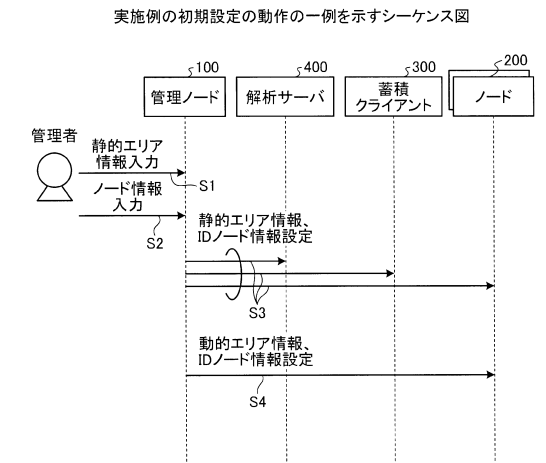
【図 1 4】



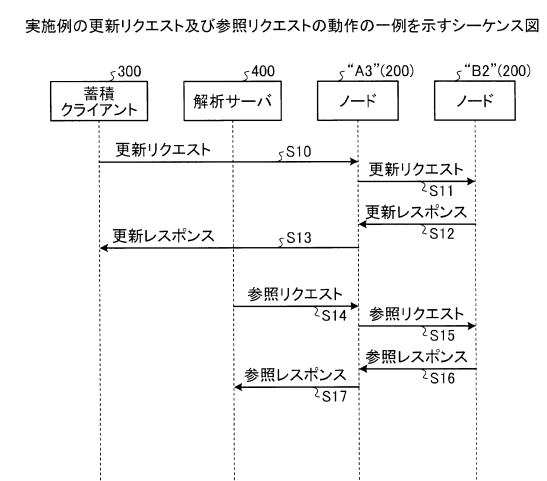
【図 1 5】



【図 1 6】

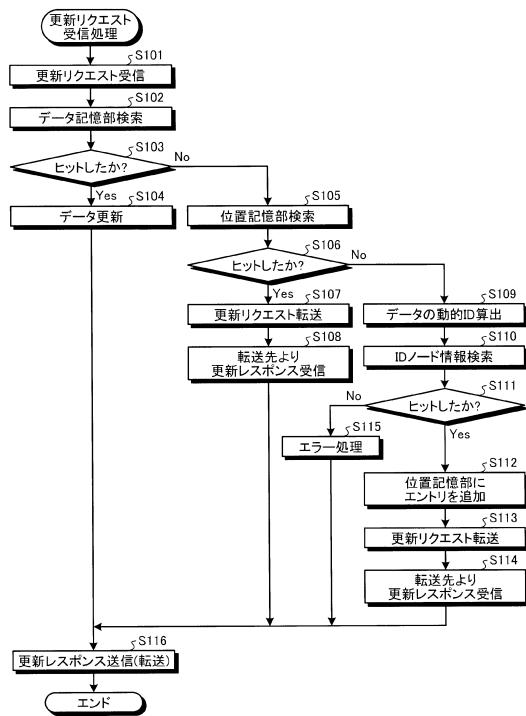


【図 1 7】



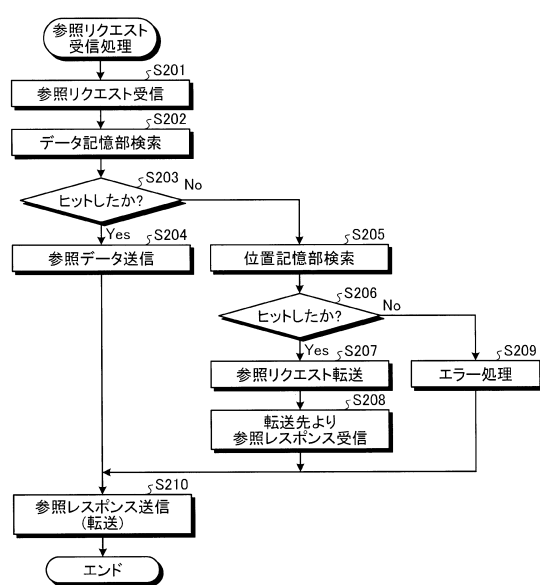
【図 18】

実施例のノードの更新リクエスト受信時における動作の一例を示すフローチャート



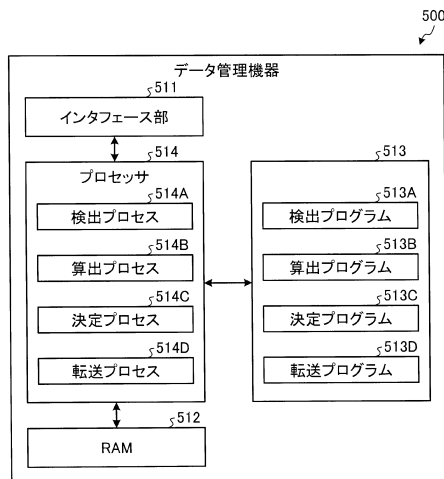
【図 19】

実施例のノードの参照リクエスト受信時における動作の一例を示すフローチャート



【図 20】

データ管理プログラムを実行するデータ管理機器の一例を示す説明図



フロントページの続き

審査官 井上 宏一

(56)参考文献 特開 2 0 1 0 - 2 7 1 7 9 8 (J P , A)
特開 2 0 1 3 - 6 1 7 3 9 (J P , A)
国際公開第 2 0 1 3 / 0 4 6 6 6 4 (W O , A 1)

(58)調査した分野(Int.Cl. , D B 名)
G 0 6 F 1 2 / 0 0
G 0 6 F 1 3 / 0 0