

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2017-146764

(P2017-146764A)

(43) 公開日 平成29年8月24日(2017.8.24)

(51) Int.Cl.

G06F 12/00 (2006.01)

F I

G06F 12/00 501B

G06F 12/00 545A

テーマコード (参考)

審査請求 未請求 請求項の数 5 O L (全 10 頁)

(21) 出願番号 特願2016-27779 (P2016-27779)

(22) 出願日 平成28年2月17日(2016.2.17)

(71) 出願人 000004226

日本電信電話株式会社

東京都千代田区大手町一丁目5番1号

(74) 代理人 110002147

特許業務法人酒井国際特許事務所

(72) 発明者 安藤 嵩

東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

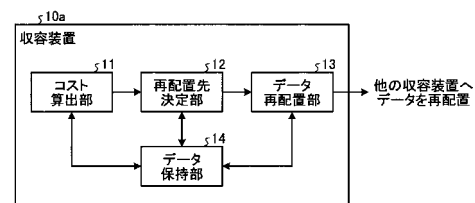
(54) 【発明の名称】 データ再配置システム、データ再配置装置、及びデータ再配置方法

(57) 【要約】

【課題】分散して保持されたデータに対し、最適なデータ再配置を自動的に行うことである。

【解決手段】収容装置10aは、再配置先決定部12とデータ再配置部13とを有する。再配置先決定部12は、複数の収容装置の内、該複数の収容装置の何れかに接続された複数の対向装置の各々からの通信コストの合計値が最小である収容装置を、何れかの収容装置に保持されたデータの再配置先となる収容装置に決定する。データ再配置部13は、再配置先決定部12により決定された収容装置に対し、上記データを再配置する。

【選択図】図2



【特許請求の範囲】**【請求項 1】**

複数の収容装置と、該複数の収容装置の何れかに接続された複数の対向装置とを有するデータ再配置システムであって、

前記複数の収容装置の各々は、

前記複数の収容装置の内、前記複数の対向装置の各々からのコストの合計値が最小である収容装置を、前記データ再配置システム内の何れかの収容装置に保持されたデータの再配置先となる収容装置に決定する決定部と、

前記決定部により決定された収容装置に対し、前記データを再配置する再配置部とを有することを特徴とするデータ再配置システム。

10

【請求項 2】

前記決定部は、前記コストの合計値が最小である収容装置が複数存在する場合、該複数の収容装置の内、前記データに対する呼量が最大である対向装置に最も近い収容装置を、前記データの再配置先となる収容装置に決定することを特徴とする請求項 1 に記載のデータ再配置システム。

【請求項 3】

前記複数の収容装置の各々は、

前記複数の対向装置からの呼量が大きいデータから順に、前記コストを算出する算出部を更に有し、

前記決定部は、前記算出部により算出されたコストの合計値が最小である収容装置を、前記データの再配置先となる収容装置に決定することを特徴とする請求項 1 に記載のデータ再配置システム。

20

【請求項 4】

複数の収容装置の内、該複数の収容装置の何れかに接続された複数の対向装置の各々からのコストの合計値が最小である収容装置を、何れかの収容装置に保持されたデータの再配置先となる収容装置に決定する決定部と、

前記決定部により決定された収容装置に対し、前記データを再配置する再配置部とを有することを特徴とするデータ再配置装置。

【請求項 5】

複数の収容装置と、該複数の収容装置の何れかに接続された複数の対向装置とを有するデータ再配置システムで実行されるデータ再配置方法であって、

30

前記複数の収容装置の内、少なくとも 1 つの収容装置が、

前記複数の収容装置の内、前記複数の対向装置の各々からのコストの合計値が最小である収容装置を、前記データ再配置システム内の何れかの収容装置に保持されたデータの再配置先となる収容装置に決定する決定工程と、

前記決定工程にて決定された収容装置に対し、前記データを再配置する再配置工程とを含むことを特徴とするデータ再配置方法。

【発明の詳細な説明】**【技術分野】****【0001】**

40

本発明は、データ再配置システム、データ再配置装置、及びデータ再配置方法に関する。

【背景技術】**【0002】**

従来、分散システムにおいて、大規模な通信システムを構築する技術が提案されている。例えば、非特許文献 1 には、分散システムを可能にする分散サービス基盤上で、インフラに関する変化を検出し、最適な配備先の計算と再配備を繰り返すことで、実行環境の変化に追従した最適な分散配備を実現する技術が提案されている。上記の様なシステムの構築では、一般的に、データ操作の処理時間に対する要件が非常に厳しく、データ更新にリアルタイム性が強く求められる。

50

【 0 0 0 3 】

図 4 は、従来技術に係る分散システム 1 0 2 を含むネットワーク 1 0 0 の構成例を示す図である。図 4 に示す様に、ネットワーク 1 0 0 は、保守端末 1 0 1 と、分散システム 1 0 2 と、サービス提供装置 1 0 3 a、1 0 3 b と、ユーザ端末 1 0 4 a、1 0 4 b、1 0 4 c、1 0 4 d とを有する。これら各構成部分は、一方向又は双方向に、各種信号やデータの送受信が可能な様に接続されている。また、分散システム 1 0 2 は、D B (Data Base) を有する複数のノード 1 0 2 a、1 0 2 b、1 0 2 c、1 0 2 d により構成される。情報通信技術の発展に伴い、D B においては、外部装置からの大量アクセスに対してリアルタイムな応答が要求される。このような要求に対応するため、D B は、分散・冗長化して運用することが一般的になりつつある。

10

【 0 0 0 4 】

以下、図 5 A 及び図 5 B を参照しながら、特にリアルタイム性の高いレスポンスが求められるサービス（例えば、通信呼処理やゲーム等の大規模ユーザに対するサービス）へデータを提供する場合を想定し、従来のデータ再配置方法について説明する。図 5 A は、従来技術に係る分散システム 1 0 2 におけるデータ再配置方法の前半部分を説明するための図である。図 5 A に示す様に、例えば、データ D 1 0 2 の再配置先となる収容装置 1 0 0 b は、呼量やデータの重さを考慮して決定される。また、図 5 B は、従来技術に係る分散システム 1 0 2 におけるデータ再配置方法の後半部分を説明するための図である。図 5 B に示す様に、例えば、実際にデータ D 1 0 1 を収容装置 1 0 0 a から収容装置 1 0 0 c に再配置する際には、データ更新のリアルタイム性を改善するために、保守管理者が、データ移行工事、所分割工事等を実施する。

20

【 先行技術文献 】

【 非特許文献 】

【 0 0 0 5 】

【 非特許文献 1 】 “クラウド環境を広域ネットワーク上で最適化する分散サービス基盤技術を開発”、[o n l i n e]、株式会社富士通研究所 システムソフトウェア研究所 分散プラットフォーム研究部、[平成28年2月1日検索]、インターネット<<http://pr.fujitsu.com/jp/news/2014/03/14.html>>

【 発明の概要 】

【 発明が解決しようとする課題 】

30

【 0 0 0 6 】

しかしながら、上述した従来の技術には、以下の様な問題点があった。例えば、呼量やデータの重さは、時間やデータの属性による変動が大きいため、ある時点の呼量やデータの重さを考慮しても、それ以降の運用やトラフィックの変化によって、最適な再配置先が変わってしまう。このため、データ再配置先の検討時に想定した効果が得られないことがある。また、データ移行工事、所分割工事等は、システム運用者にとって、保守稼働が非常に大きくなることがある。これらのことが、最適なデータ再配置を阻害し、データ更新速度を低下させる要因となっていた。

【 0 0 0 7 】

開示の技術は、上記に鑑みてなされたものであって、分散して保持されたデータに対し、最適なデータ再配置を自動的に行うデータ再配置システム、データ再配置装置、及びデータ再配置方法を提供することを目的とする。

40

【 課題を解決するための手段 】

【 0 0 0 8 】

上述した課題を解決し、目的を達成するために、本願の開示するデータ再配置システムは、一つの態様において、複数の収容装置と、該複数の収容装置の何れかに接続された複数の対向装置とを有するデータ再配置システムであって、前記複数の収容装置の各々は、前記複数の収容装置の内、前記複数の対向装置の各々からのコストの合計値が最小である収容装置を、前記データ再配置システム内の何れかの収容装置に保持されたデータの再配置先となる収容装置に決定する決定部と、前記決定部により決定された収容装置に対し、

50

前記データを再配置する再配置部とを有する。

【 0 0 0 9 】

また、本願の開示するデータ再配置装置は、一つの態様において、複数の収容装置の内、該複数の収容装置の何れかに接続された複数の対向装置の各々からのコストの合計値が最小である収容装置を、何れかの収容装置に保持されたデータの再配置先となる収容装置に決定する決定部と、前記決定部により決定された収容装置に対し、前記データを再配置する再配置部とを有する。

【 0 0 1 0 】

更に、本願の開示するデータ再配置方法は、一つの態様において、複数の収容装置と、該複数の収容装置の何れかに接続された複数の対向装置とを有するデータ再配置システムで実行されるデータ再配置方法であって、前記複数の収容装置の内、少なくとも1つの収容装置が、前記複数の収容装置の内、前記複数の対向装置の各々からのコストの合計値が最小である収容装置を、前記データ再配置システム内の何れかの収容装置に保持されたデータの再配置先となる収容装置に決定する決定工程と、前記決定工程にて決定された収容装置に対し、前記データを再配置する再配置工程とを含む。

10

【 発 明 の 効 果 】

【 0 0 1 1 】

本願の開示するデータ再配置システム、データ再配置装置、及びデータ再配置方法は、分散して保持されたデータに対し、最適なデータ再配置を自動的に行うことができるという効果を奏する。

20

【 図 面 の 簡 単 な 説 明 】

【 0 0 1 2 】

【 図 1 A 】 図 1 A は、データの再配置前における分散システムの一例を示す図である。

【 図 1 B 】 図 1 B は、データの再配置後における分散システムの一例を示す図である。

【 図 2 】 図 2 は、本実施例に係る収容装置の構成例を示す図である。

【 図 3 】 図 3 は、データ再配置プログラムによる情報処理がコンピュータを用いて具体的に実現されることを示す図である。

【 図 4 】 図 4 は、従来技術に係る分散システムを含むネットワークの構成例を示す図である。

【 図 5 A 】 図 5 A は、従来技術に係る分散システムにおけるデータ再配置方法の前半部分を説明するための図である。

30

【 図 5 B 】 図 5 B は、従来技術に係る分散システムにおけるデータ再配置方法の後半部分を説明するための図である。

【 発 明 を 実 施 す る た め の 形 態 】

【 0 0 1 3 】

以下に、本願の開示するデータ再配置システム、データ再配置装置、及びデータ再配置方法の実施例を、図面を参照しながら詳細に説明する。なお、以下の実施例により本願の開示するデータ再配置システム、データ再配置装置、及びデータ再配置方法が限定されるものではない。

【 0 0 1 4 】

40

まず、本願の開示する一実施例に係るデータ再配置システムとしての分散システム 1 について説明する。図 1 A は、データ D 1 の再配置前における分散システム 1 の一例を示す図である。図 1 A に示す様に、分散システム 1 は、それぞれ “ 0 5 0 - A A A A ”、“ 0 5 0 - B B B B ”、“ 0 5 0 - C C C C ”により識別されるデータ D 1、D 2、D 3 を、収容装置 1 0 a、1 0 b、1 0 c に分散して保持する。図 1 A において、収容装置 1 0 a ~ 1 0 f 及び対向装置 1 0 g ~ 1 0 i 間に記載された数値は、各経路間の通信コストを表す。

【 0 0 1 5 】

例えば、データ I D “ 0 5 0 - A A A A ”により識別されるデータ D 1 は、現時点では、収容装置 1 0 a により保持され、対向装置 1 0 g からの呼量は「 1」、対向装置 1 0 h

50

からの呼量は「2」、対向装置10iからの呼量は「3」である。また、データID「050-BBBBB」により識別されるデータD2は、収容装置10bにより保持され、対向装置10g、対向装置10h、対向装置10iからの呼量は、それぞれ「2」、「2」、「1」である。同様に、データID「050-CCCCC」により識別されるデータD3は、収容装置10cにより保持され、対向装置10g、対向装置10h、対向装置10iからの呼量は、それぞれ「3」、「1」、「0」である。

【0016】

図1Bは、データD1の再配置後における分散システム1の一例を示す図である。本実施例では、分散して保持されたデータD1、D2、D3の内、データD1を再配置の対象データとする。図1Bに示す様に、データID「050-AAAA」を有するデータD1
10
に関し、収容装置10a～10fのトータルコストは、それぞれ「13」、「9」、「10」、「11」、「10」、「10」である。このため、収容装置10aに配置されていたデータD1は、矢印Y1に示す様に、トータルコストが最小値「9」をとる収容装置10bに再配置される。これにより、データD1を用いたデータ更新に際してのトータルコストが低減される。また、データID「050-AAAA」を有するデータD1は、呼量が最大値「3」をとる対向装置10iの近傍に再配置されることになり、データ配置が最適化される。

【0017】

図2は、本実施例に係る収容装置10aの構成例を示す図である。図2に示す様に、収容装置10aは、コスト算出部11と再配置先決定部12とデータ再配置部13とデータ
20
保持部14とを有する。これら各構成部分は、一方向又は双方向に、各種信号やデータの入出力が可能な様に接続されている。コスト算出部11は、複数の対向装置10g～10iからの呼量が大きいデータから順に通信コストを算出する。再配置先決定部12は、複数の収容装置10a～10fの内、複数の対向装置10g～10iの各々からの通信コストの合計値（トータルコスト）が最小である収容装置を、データの再配置先となる収容装置に決定する。データ再配置部13は、決定された収容装置に対し、上記データを再配置する。データ保持部14は、再配置対象のデータを保持する。

【0018】

以上、収容装置10aの構成及び動作を代表的に説明したが、収容装置10a以外の他の収容装置10b～10fの構成及び動作は、収容装置10aの構成及び動作と同様であるので、その図示及び詳細な説明は省略する。
30

【0019】

以上説明した様に、分散システム1は、複数の収容装置10a～10fと、該複数の収容装置10a～10fの何れかに接続された複数の対向装置10g～10iとを有する。複数の収容装置10a～10fの各々は、再配置先決定部12とデータ再配置部13とを有する。再配置先決定部12は、複数の収容装置10a～10fの内、複数の対向装置10g～10iの各々からの通信コストの合計値（トータルコスト）が最小である収容装置10bを、分散システム1内の何れかの収容装置10aのデータ保持部14に保持されたデータの再配置先となる収容装置に決定する。データ再配置部13は、再配置先決定部12により決定された収容装置10bに対し、上記データを再配置する。
40

【0020】

分散システム1によれば、時間経過によるトラフィックやデータ収容の変動の影響が抑制され、リアルタイム性の高いデータ再配置が実現される。すなわち、分散システム1は、時間経過による変動の少ないネットワーク構成を基に通信コストを評価し、該評価の結果を用いて、最適なデータ配置先を自動で算出し、データの再配置を実行する。これにより、最適なデータ再配置によるデータ更新速度（例えば、対向装置に保持されたデータのアップグレード速度等）の向上が可能となる。また、分散システム1は、自動的なデータ再配置を行うことにより、リアルタイム性を向上させるためのデータ移行においても、保守稼働が増大しないため、保守稼働に伴う保守管理者の負担を軽減することが可能となる。

【 0 0 2 1 】

また、分散システム 1 において、収容装置 1 0 a の再配置先決定部 1 2 は、上記通信コストの合計値が最小である収容装置が複数存在する場合、該複数の収容装置 1 0 a ~ 1 0 f の内、上記データに対する呼量が最大である対向装置 1 0 i に最も近い収容装置を、上記データの再配置先となる収容装置に決定するものとしてもよい。これにより、あるデータに関し、通信コストが最小の収容装置が複数存在する場合でも、該データの参照回数の高い対向装置の近傍に優先的にデータを再配置することができるため、より効率的なデータの再配置ひいてはデータ更新速度の向上が実現可能となる。

【 0 0 2 2 】

ここで、最も近い収容装置とは、必ずしも物理的な距離が最小の収容装置に限らず、例えば、最少のホップ数、最短のデータ送受信時間、最良のデータ通信品質、あるいは最少の通信費用等をとる収容装置であってもよい。

10

【 0 0 2 3 】

また、分散システム 1 において、複数の収容装置 1 0 a ~ 1 0 f の各々は、複数の対向装置 1 0 g ~ 1 0 i からの呼量が大きいデータから順に、所定の最適経路探索（例えば、ダイクストラ法、ベルマンフォード法）を実行し、上記通信コストを算出するコスト算出部 1 1 を更に有するものとしてもよい。かかる態様では、再配置先決定部 1 2 は、コスト算出部 1 1 により算出された通信コストの合計値が最小である収容装置を、上記データの再配置先となる収容装置に決定する。これにより、例えば、参照回数の高いデータから優先的に最適な収容装置に移行させることができるため、より効率的なデータの再配置ひいてはデータ更新速度の向上が実現可能となる。

20

【 0 0 2 4 】

なお、上記実施例では、上述したデータ再配置の機能を、分散システム 1 内の各収容装置 1 0 a ~ 1 0 f が個別に具備するものとしたが、上記機能は、収容装置 1 0 a ~ 1 0 f とは別体に構成された中央集中型装置が具備するものとしてもよい。かかる態様では、中央集中型装置は、分散システム 1 の内部または外部の何れに配置されてもよい。

【 0 0 2 5 】

また、上記実施例では、データの再配置先となる収容装置を決定する際の評価対象として、通信コストを例示したが、呼量やデータの重さ（収容効率）と比較して、時間経過による変動の少ない指標であれば、他の指標（例えば、収容装置、対向装置等の装置間の遅延時間）を、評価対象として用いてもよい。また、通信コストについても、物理的な距離に限らず、ホップ数、回線の速度や状態、通信費用等を用いてもよい。更に、通信コストの合計値が最小である収容装置が複数存在する場合に考慮される呼量は、例えば、データの参照回数であるが、データの使用頻度等であってもよい。

30

【 0 0 2 6 】

また、データの再配置処理を実行する契機は、例えば、該データに関する通信コストの合計値の変動（例えば、あるデータに関し、トータルコストが最小となる収容装置の変化）であるが、これに限らず、上記データに対する呼量が最大である対向装置の変化あるいは、該対向装置に最も近い収容装置の変化等であってもよい。

【 0 0 2 7 】

40

（データ再配置プログラム）

図 3 は、データ再配置プログラムによる情報処理がコンピュータ 2 0 0 を用いて具体的に実現されることを示す図である。図 3 に示す様に、コンピュータ 2 0 0 は、例えば、メモリ 2 0 1 と、CPU（Central Processing Unit）2 0 2 と、ハードディスクドライブインタフェース 2 0 3 と、ディスクドライブインタフェース 2 0 4 と、シリアルポートインタフェース 2 0 5 と、ビデオアダプタ 2 0 6 と、ネットワークインタフェース 2 0 7 とを有し、これらの各部はバス C によって接続される。

【 0 0 2 8 】

メモリ 2 0 1 は、図 3 に示す様に、ROM（Read Only Memory）2 0 1 a 及び RAM（Random Access Memory）2 0 1 b を含む。ROM 2 0 1 a は、例えば、BIOS（Ba

50

sic Input Output System)等のブートプログラムを記憶する。ハードディスクドライブインタフェース203は、図3に示す様に、ハードディスクドライブ208に接続される。ディスクドライブインタフェース204は、図3に示す様に、ディスクドライブ209に接続される。例えば磁気ディスクや光ディスク等の着脱可能な記憶媒体が、ディスクドライブ209に挿入される。シリアルポートインタフェース205は、図3に示す様に、例えばマウス210、キーボード211に接続される。ビデオアダプタ206は、図3に示す様に、例えばディスプレイ212に接続される。

【0029】

ここで、図3に示す様に、ハードディスクドライブ208は、例えば、OS(Operating System)208a、アプリケーションプログラム208b、プログラムモジュール208c、プログラムデータ208d、データD1~D3を記憶する。すなわち、開示の技術に係るデータ再配置プログラムは、コンピュータ200によって実行される指令が記述されたプログラムモジュール208cとして、例えばハードディスクドライブ208に記憶される。具体的には、上記実施例で説明したコスト算出部11、再配置先決定部12、データ再配置部13の各々と同様の情報処理を実行する各種手順が記述されたプログラムモジュール208cが、ハードディスクドライブ208に記憶される。また、データ再配置プログラムによる情報処理に用いられるデータは、プログラムデータ208dとして、例えばハードディスクドライブ208に記憶される。そして、CPU202が、ハードディスクドライブ208に記憶されたプログラムモジュール208cやプログラムデータ208dを必要に応じてRAM201bに読み出し、上記各種手順を実行する。

【0030】

なお、データ再配置プログラムに係るプログラムモジュール208cやプログラムデータ208dは、ハードディスクドライブ208に記憶される場合に限られず、例えば着脱可能な記憶媒体に記憶され、ディスクドライブ209等を介してCPU202によって読み出されてもよい。あるいは、データ再配置プログラムに係るプログラムモジュール208cやプログラムデータ208dは、ネットワーク(LAN(Local Area Network)、WAN(Wide Area Network)等)を介して接続された他のコンピュータに記憶され、ネットワークインタフェース207を介してCPU202によって読み出されてもよい。

【0031】

また、上述した分散システム1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、各装置の分散・統合の具体的態様は、図示のものに限らず、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もできる。例えば、収容装置10aに関し、再配置先決定部12とデータ再配置部13とを1つの構成要素として統合してもよい。反対に、再配置先決定部12に関し、通信コストの合計値から再配置先を決定する部分と、呼量の最大値から再配置先を決定する部分とに分散してもよい。更に、ハードディスクドライブ208を、分散システム1の外部装置として、ネットワークやケーブル経由で接続する様にしてもよい。

【符号の説明】

【0032】

1 分散システム

10a、10b、10c、10d、10e、10f 収容装置

10g、10h、10i 対向装置

11 コスト算出部

12 再配置先決定部

13 データ再配置部

14 データ保持部

100 ネットワーク

100a、100b、100c、100d、100e、100f 収容装置

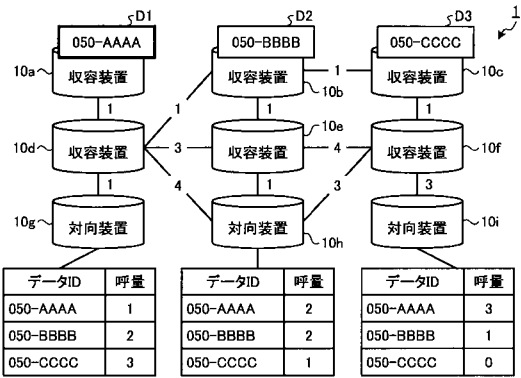
100g、100h、100i 対向装置

1 0 1 保守端末
1 0 2 分散システム
1 0 2 a、1 0 2 b、1 0 2 c、1 0 2 d ノード
1 0 3 a、1 0 3 b サービス提供装置
1 0 4 a、1 0 4 b、1 0 4 c、1 0 4 d ユーザ端末
2 0 0 コンピュータ
2 0 1 メモリ
2 0 1 a R O M
2 0 1 b R A M
2 0 2 C P U
2 0 3 ハードディスクドライブインタフェース
2 0 4 ディスクドライブインタフェース
2 0 5 シリアルポートインタフェース
2 0 6 ビデオアダプタ
2 0 7 ネットワークインタフェース
2 0 8 ハードディスクドライブ
2 0 8 a O S
2 0 8 b アプリケーションプログラム
2 0 8 c プログラムモジュール
2 0 8 d プログラムデータ
2 0 9 ディスクドライブ
2 1 0 マウス
2 1 1 キーボード
2 1 2 ディスプレイ
C バス
D 1、D 2、D 3、D 1 0 1、D 1 0 2、D 1 0 3 データ

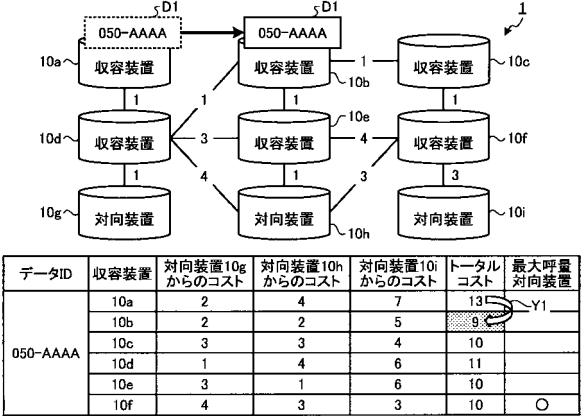
10

20

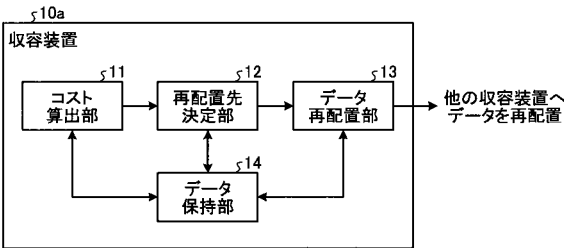
【 図 1 A 】



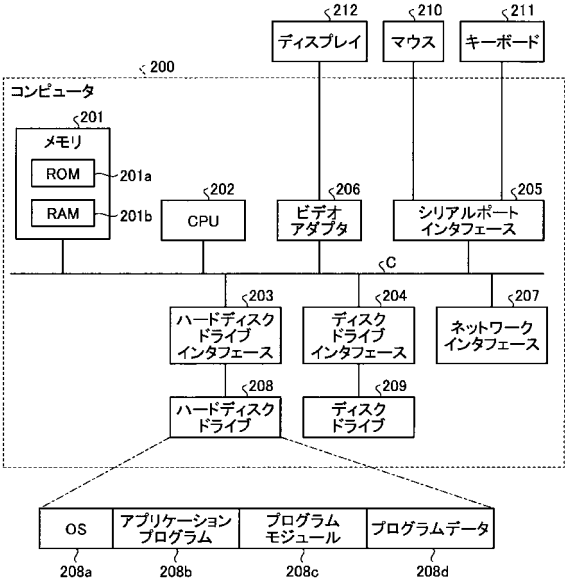
【 図 1 B 】



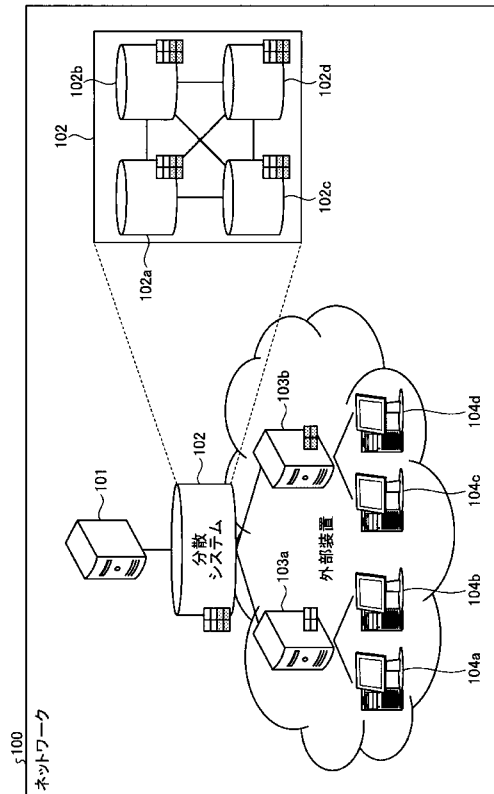
【 図 2 】



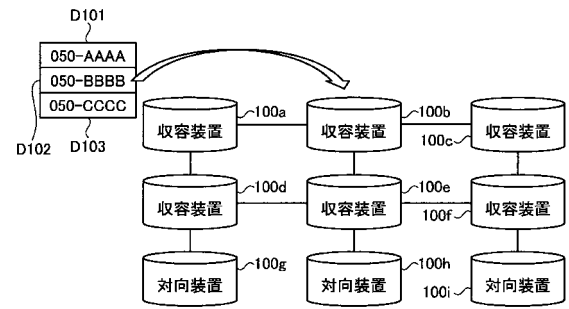
【 図 3 】



【図 4】



【図 5 A】



【図 5 B】

