

(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2008-33483
(P2008-33483A)

(43) 公開日 平成20年2月14日(2008.2.14)

(51) Int.Cl.

G06F 9/46 (2006.01)

F I

G06F 9/46 420Z

テーマコード (参考)

審査請求 有 請求項の数 22 O L (全 25 頁)

(21) 出願番号 特願2006-204286 (P2006-204286)
(22) 出願日 平成18年7月27日 (2006.7.27)

(71) 出願人 000004237
日本電気株式会社
東京都港区芝五丁目7番1号
(74) 代理人 100102864
弁理士 工藤 実
(72) 発明者 妹尾 賢一
東京都港区芝五丁目7番1号 日本電気株式会社社内

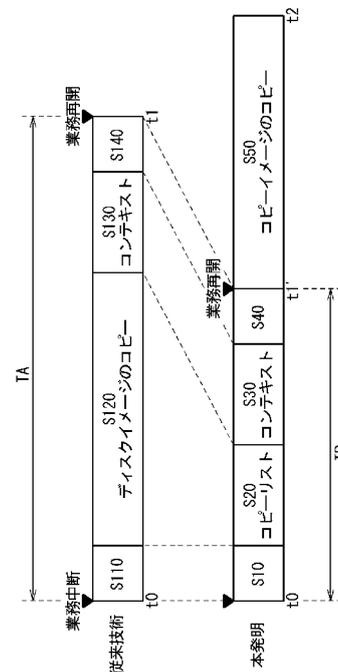
(54) 【発明の名称】 計算機システム、計算機および計算機動作環境の移動方法

(57) 【要約】

【課題】第1ディスクを使用する第1計算機の動作環境を、第2ディスクを使用する第2計算機に移動させる際に、業務の中断時間を短縮することができる技術を提供すること。

【解決手段】本発明に係る動作環境移動方法は、(A) 第1計算機100の動作を中断するステップと、(B) 上記(A)ステップの後、第1ディスク110上のコピーイメージCIに含まれるファイルのリストCLを作成するステップと、(C) 上記(A)ステップの後、第1計算機100の実行コンテキストECを、第2計算機200にコピーするステップと、(D) 上記(B)及び(C)ステップの後、第2計算機200において上記動作を再開させるステップと、(E) 上記(D)ステップの後、上記リストCLを参照して、コピーイメージCIを第1ディスク110から第2ディスク(210)にコピーするステップと、を有する。

【選択図】 図4



【特許請求の範囲】**【請求項 1】**

第 1 ディスクを使用する第 1 計算機の動作環境を、第 2 ディスクを使用する第 2 計算機に移動させる動作環境移動方法であって、

(A) 前記第 1 計算機の動作を中断するステップと、

(B) 前記 (A) ステップの後、前記第 1 ディスクのディスクイメージの少なくとも一部であるコピーイメージに含まれるファイルのリストを作成するステップと、

(C) 前記 (A) ステップの後、前記第 1 計算機の実行コンテキストを、前記第 2 計算機にコピーするステップと、

(D) 前記 (B)、(C) ステップの後、前記第 2 計算機において前記動作を再開させるステップと、

(E) 前記 (D) ステップの後、前記リストを参照して、前記コピーイメージを前記第 1 ディスクから前記第 2 ディスクにコピーするステップと

を有する

動作環境移動方法。

【請求項 2】

請求項 1 に記載の動作環境移動方法であって、

前記コピーイメージは、前記ディスクイメージの全てである

動作環境移動方法。

【請求項 3】

請求項 1 に記載の動作環境移動方法であって、

更に、(F) 前記 (A) ステップより前に、前記ディスクイメージの一部をベースイメージとして前記第 2 計算機に配布するステップを有し、

前記コピーイメージは、前記ディスクイメージと前記ベースイメージとの差分である

動作環境移動方法。

【請求項 4】

請求項 3 に記載の動作環境移動方法であって、

前記 (F) ステップは、

(F1) 前記 (A) ステップより前のあるタイミングで、前記ベースイメージを作成するステップと

(F2) 前記作成されたベースイメージを前記第 2 計算機に配布するステップとを含み、

前記 (B) ステップにおいて、前記リストは、前記 (F1) ステップから前記 (A) ステップまでの間のファイルの変更情報を示すログファイルに基づいて作成される

動作環境移動方法。

【請求項 5】

請求項 4 に記載の動作環境移動方法であって、

前記あるタイミングは、前記第 1 計算機における OS のインストール直後である

動作環境移動方法。

【請求項 6】

請求項 1 乃至 5 のいずれかに記載の動作環境移動方法であって、

前記 (B) ステップにおいて、前記作成されたリストは、前記第 2 ディスクに格納される

動作環境移動方法。

【請求項 7】

請求項 1 乃至 6 のいずれかに記載の動作環境移動方法であって、

前記 (E) ステップにおいて、前記コピーイメージは、前記第 2 計算機の動作のバックグラウンドでコピーされる

動作環境移動方法。

【請求項 8】

10

20

30

40

50

請求項 1 乃至 6 のいずれかに記載の動作環境移動方法であって、
前記 (E) ステップにおいて、前記コピーイメージは、前記第 2 計算機の要求に
応答してオンデマンド方式でコピーされる
動作環境移動方法。

【請求項 9】

請求項 7 又は 8 に記載の動作環境移動方法であって、
前記 (E) ステップは、
(E 1) 前記第 2 計算機が前記第 2 ディスクへのアクセスする際、アクセス対象である
対象データの実体が存在するか否かの判定を行うステップと、
(E 2) 前記対象データの実体が存在しない場合、前記リストを参照して、前記コピー
イメージ中の前記対象データを、前記第 1 ディスクから前記第 2 ディスクにコピーするス
テップと
を含む
動作環境移動方法。

10

【請求項 10】

請求項 9 に記載の動作環境移動方法であって、
前記リストは、前記コピーイメージに含まれる各ファイルがコピー済か否かを
示すコピー情報を有し、
前記 (E 1) ステップにおいて、前記リスト中の前記対象データに対応する前記
コピー情報に基づいて前記判定が行われ、
前記 (E 2) ステップにおいて、前記対象データに対応する前記コピー情報が
コピー済に変更される
動作環境移動方法。

20

【請求項 11】

ディスクへのアクセス時、アクセス対象である対象データの実体が存在するか
否かの判定を行うコピー判定モジュールと、
前記対象データの実体が存在しない場合、前記対象データが格納された移動元
ディスクにアクセス可能な装置に対して、前記ディスクへの前記対象データの
コピーを要求するコピーモジュールと
を備える
計算機。

30

【請求項 12】

請求項 11 に記載の計算機であって、
前記ディスクには、コピーが必要なファイルのリストが格納されており、前記
リストは、前記ファイルの各々がコピー済か否かを示すコピー情報を有し、
前記コピー判定モジュールは、前記リスト中の前記対象データに対応する前記
コピー情報に基づいて前記判定を行い、
前記コピーモジュールは、前記対象データのコピーが完了した後、前記対象
データに対応する前記コピー情報をコピー済に変更する
計算機。

40

【請求項 13】

請求項 12 に記載の計算機であって、
前記コピー判定モジュールは、前記リストに対応付けられたハッシュテーブル
を利用することによって、前記リストから前記対象データに関する情報の検索
を行う
計算機。

【請求項 14】

請求項 11 乃至 13 のいずれかに記載の計算機であって、
仮想マシンと、
前記仮想マシンを制御する仮想マシンモニタと
を具備し、

50

前記仮想マシンモニタが、前記コピー判定モジュールと前記コピーモジュールを有する計算機。

【請求項 15】

第1ディスクを使用する第1計算機の動作環境を、第2ディスクを使用する第2計算機に移動させる計算機システムであって、

前記第1計算機の動作中断後、前記第1ディスクのディスクイメージの少なくとも一部であるコピーイメージに含まれるファイルのリストを作成するコピーリスト作成モジュールと、

前記第1計算機の動作中断後、前記第1計算機の実行コンテキストを、前記第2計算機にコピーするコンテキストコピーモジュールと

を備え、

前記作成されたリストは前記第2ディスクに格納される計算機システム。

【請求項 16】

請求項 15 に記載の計算機システムであって、

前記コピーイメージは、前記ディスクイメージの全てである計算機システム。

【請求項 17】

請求項 15 に記載の計算機システムであって、

更に、

前記第1計算機の動作中断より前に、前記ディスクイメージの一部をベースイメージとして作成するベースイメージ作成モジュールと、

前記第1計算機の動作中断より前に、前記ベースイメージを前記第2計算機に配布するベースイメージ配布モジュールと

を備え、

前記コピーイメージは、前記ディスクイメージと前記ベースイメージとの差分である計算機システム。

【請求項 18】

請求項 17 に記載の計算機システムであって、

前記コピーリスト作成モジュールは、前記ベースイメージの作成から前記動作中断までの間のファイルの変更情報を示すログファイルに基づいて、前記コピーリストを作成する計算機システム。

【請求項 19】

請求項 15 乃至 18 のいずれかに記載の計算機システムであって、

前記第1ディスクにアクセス可能な第1実計算機と、

ネットワークを介して前記第1実計算機に接続され、前記第2ディスクにアクセス可能な第2実計算機と

を具備し、

前記第1計算機は、前記第1実計算機上で動作する第1仮想マシンモニタによって提供される仮想マシンであり、

前記第2計算機は、前記第2実計算機上で動作する第2仮想マシンモニタによって提供される仮想マシンであり、

前記第1仮想マシンモニタが、前記コピーリスト作成モジュール及び前記コンテキストコピーモジュールを有し、

前記第1仮想マシンモニタは、前記第2仮想マシンモニタを介して、前記作成されたリスト及び前記実行コンテキストを前記第2ディスク及び前記第2計算機にコピーする計算機システム。

【請求項 20】

請求項 15 乃至 18 のいずれかに記載の計算機システムであって、

ネットワークを介して前記第1計算機及び前記第2計算機に接続され、前記第1ディス

10

20

30

40

50

ク及び前記第 2 ディスクにアクセス可能な管理計算機を具備し、

前記管理計算機が、前記コピーリスト作成モジュール及び前記コピーモジュールを有し

、
前記管理計算機は、前記作成されたリスト及び前記実行コンテキストを前記第 2 ディスク及び前記第 2 計算機にコピーする

計算機システム。

【請求項 2 1】

(a) 第 1 ディスク上のファイルに対する操作の履歴を示すログファイルを解析するステップと、

(b) 前記履歴のうち最後に行われた前記操作が M o d i f y の場合、当該ファイルに固有 I D を割り当てるステップと、

(c) 前記割り当てられた固有 I D と、当該ファイルの名前と、当該ファイルが第 2 ディスクにコピー済か否かを示すコピー情報とを含むエントリからなるリストを作成するステップと

をコンピュータに実行させるためのプログラム。

【請求項 2 2】

請求項 2 1 に記載のプログラムであって、

更に、(d) 前記リスト中の固有 I D に基づいて、前記リストの検索に用いられるハッシュテーブルを作成するステップ

をコンピュータに実行させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、計算機システムにおいて、ある計算機の動作環境を他の計算機に移動させるための技術に関する。

【背景技術】

【0002】

ネットワークを介して様々なサービスを提供するサーバシステムが知られている。サーバシステムにおいて、パフォーマンス増強、負荷分散、メンテナンスなどを目的として、運用中のサーバによる業務を他のサーバに移行させる場合がある。その場合、運用中のサーバの業務が中断され、そのサーバの動作環境を、他のサーバにそっくり移動させることが必要である。特に、ディスクが共有されていない場合は、運用中のサーバのディスクイメージを、他のサーバのディスクにコピーすることが必要である。

【0003】

従来のサーバ動作環境の移動方法は、次の通りである。まず、動作環境の移動に先立って、移動元サーバと移動先サーバとの間でハードウェア構成（I/Oカードの-slot位置など）の比較が行われ、ディスクイメージの継承が可能であることが確認される。その後、動作環境の移動が開始する。動作環境の移動処理においては、図 1 に示されるフローが実行される。

【0004】

まず、移動元サーバの動作が中断される（ステップ S 1 1 0）。続いて、移動元サーバのディスクから移動先サーバのディスクに、移動元サーバのディスクイメージがコピーされる（ステップ S 1 2 0）。また、移動元サーバから移動先サーバに、「実行コンテキスト（execution contexts）」がコピーされる（ステップ S 1 3 0）。ここで、実行コンテキストとは、実行中のプログラムが利用しているレジスタセットやメインメモリの内容である。ディスクイメージ及び実行コンテキストのコピーが完了すると、移動先サーバの動作が開始させられる。これにより、中断していた業務が、移動先サーバにおいて再開する（ステップ S 1 4 0）。このようにして、移動元サーバの業務が、移動先サーバに引き継がれる。

【0005】

尚、関連する従来技術として、次のものが知られている。

【0006】

特許文献1には、複数の実計算機上に構築される仮想計算機システム間で、ある仮想計算機（VM:Virtual Machine）を移動させる方式が記載されている。この方式によれば、移動元計算機と移動先計算機は、磁気ディスク装置を共有している。移動元の仮想計算機システムには、移動させるVMの構成情報と対応する実資源の論理名称を通知する手段が設けられる。移動先の仮想計算機システムには、受信した論理名称に基づいて同一のVM構成を生成する手段が設けられる。

【0007】

【特許文献1】特開平10-283210号公報

10

【発明の開示】

【発明が解決しようとする課題】

【0008】

本願発明者は、次の点に着目した。ディスクイメージのサイズは数Gbyteと大きく、ディスクイメージのコピー（ステップS120）に要する時間は非常に長い。サーバ動作環境の移動時間、すなわち業務中断（ステップS110）から業務再開（ステップS140）までの時間は、ディスクイメージのコピー時間にほぼ律速されていると言える。サービス提供者の観点から、業務の中断時間はできる限り短いことが望ましい。

【0009】

従って、本発明の目的は、ある計算機の動作環境を他の計算機に移動させる際に、業務の中断時間を短縮することができる技術を提供することにある。

20

【課題を解決するための手段】

【0010】

以下に、[発明を実施するための最良の形態]で使用される番号・符号を用いて、[課題を解決するための手段]を説明する。これらの番号・符号は、[特許請求の範囲]の記載と[発明を実施するための最良の形態]との対応関係を明らかにするために括弧付きで付加されたものである。ただし、それらの番号・符号を、[特許請求の範囲]に記載されている発明の技術的範囲の解釈に用いてはならない。

【0011】

本発明の第1の観点において、第1ディスク（110）を使用する第1計算機（100）の動作環境を、第2ディスク（210）を使用する第2計算機（200）に移動させる動作環境移動方法が提供される。その動作環境移動方法は、（A）第1計算機（100）の動作を中断するステップと、（B）上記（A）ステップの後、第1ディスク（110）のディスクイメージ（10）の少なくとも一部であるコピーイメージ（CI）に含まれるファイルのリスト（CL）を作成するステップと、（C）上記（A）ステップの後、第1計算機（100）の実行コンテキスト（EC）を、第2計算機（200）にコピーするステップと、（D）上記（B）及び（C）ステップの後、第2計算機（200）において上記動作を再開させるステップと、（E）上記（D）ステップの後、上記リスト（CL）を参照して、コピーイメージ（CI）を第1ディスク（110）から第2ディスク（210）にコピーするステップと、を有する。

30

40

【0012】

本発明によれば、第1計算機（100）の動作中断から第2計算機（200）における動作再開までの間に、実行コンテキスト（EC）が第2計算機（200）にコピーされる。一方、第1ディスク（110）上のコピーイメージ（CI）に関しては、それに含まれるファイルのリスト（CL）だけが作成され、コピーイメージ（CI）の実体のコピーは行われない。コピーイメージ（CI）の実体のコピーは、第2計算機（200）における動作再開後、作成されたリスト（CL）を参照することにより行われる。従って、第1計算機（100）の動作中断から第2計算機（200）における動作再開までの時間が短縮される。すなわち、本発明によれば、業務の中断時間を短縮することが可能となる。

【0013】

50

本発明の第2の観点において、計算機が提供される。その計算機は、コピー判定モジュール(40)とコピーモジュール(50)を備えている。コピー判定モジュール(40)は、ディスク(210)へのアクセス時、アクセス対象である対象データの実体が存在するか否かの判定を行う。対象データの実体が存在しない場合、コピーモジュール(50)は、対象データが格納された移動元ディスク(110)にアクセス可能な装置(100, 120, 130, 300)に対して、自身のディスク(210)への対象データのコピーを要求する。

【0014】

このような構成を有する計算機は、業務の最中にコピー処理を行うことが可能である。第1計算機(100)の動作環境が第2計算機(200)に移動される際、第1計算機(100)の動作中断から第2計算機(200)における動作再開までの間、業務は中断する。その業務中断期間において、第2計算機(200)は、第1計算機(100)の実行コンテキスト(EC)と、コピーが必要なファイルのリスト(CL)だけを受け取ればよい。その業務中断期間に、第2計算機(200)が第1計算機(100)のディスク(110)上のコピーイメージ(CI)の実体を受け取る必要はない。それは、第2計算機(200)における動作再開後、受け取ったリスト(CL)を参照することによって、そのコピーイメージ(CI)の実体をコピーすることが可能だからである。従って、業務の中断時間が短縮される。

【0015】

本発明の第3の観点において、第1ディスク(110)を使用する第1計算機(100)の動作環境を、第2ディスク(210)を使用する第2計算機(200)に移動させる計算機システムが提供される。その計算機システムは、コピーリスト作成モジュール(20)と、コンテキストコピーモジュール(30)とを備えている。コピーリスト作成モジュール(20)は、第1計算機(100)の動作中断後、第1ディスク(110)のディスクイメージ(10)の少なくとも一部であるコピーイメージ(CI)に含まれるファイルのリスト(CL)を作成する。作成されたリスト(CL)は、第2ディスク(210)に格納される。コンテキストコピーモジュール(30)は、第1計算機(100)の動作中断後、第1計算機(100)の実行コンテキスト(EC)を、第2計算機(200)にコピーする。

【0016】

本発明によれば、第2計算機(200)は、上記実行コンテキスト(EC)と上記リスト(CL)を受け取ると直ぐに、業務を再開させることができる。それは、上記リスト(CL)を参照することによって、第1ディスク(110)上のコピーイメージ(CI)の実体を、第2計算機(200)の動作再開後にコピーすることが可能だからである。第1計算機(100)の動作中断から第2計算機(200)における動作再開までの間、第2計算機(200)がコピーイメージ(CI)の実体を受け取る必要はない。従って、動作中断から動作再開までの時間、すなわち、業務の中断時間を短縮することが可能となる。

【発明の効果】

【0017】

本発明によれば、ある計算機の動作環境を他の計算機に移動させる際に、業務の中断時間を短縮することが可能となる。

【発明を実施するための最良の形態】

【0018】

添付図面を参照して、本発明の実施の形態に係る計算機システムおよび計算機動作環境の移動方法を説明する。計算機システムとして、自律コンピューティングシステム、ユーティリティ・コンピューティングシステム、グリッドシステム、仮想化システムなどが挙げられる。本実施の形態において、計算機システムとして、様々なサービスを提供するサーバシステムが例示される。

【0019】

1. 概要

10

20

30

40

50

まず、本発明の概要を説明する。図2は、本発明の実施の形態に係るサーバシステム1を概念的に示している。本実施の形態において、移動元サーバ100の「動作環境」が、移動先サーバ200に移動させられる。移動元サーバ100は移動元ディスク110を使用し、移動先サーバ200は移動先ディスク210を使用する。移動元ディスク110は、移動先サーバ200によって共有されていないとする。

【0020】

移動元サーバ100の動作環境とは、移動元ディスク110のディスクイメージ10に加えて、移動元サーバ100の実行コンテキストECを含む。実行コンテキストECとは、実行中のプログラムが利用しているCPU中のレジスタの内容や、メインメモリ(RAM)の内容である。つまり、実行コンテキストECは、ある時点における計算機の動作状態(実行状態)を再現することができる情報の束である。

10

【0021】

図3は、本実施の形態に係るサーバ動作環境の移動方法を示すフローチャートである。図2及び図3を参照して、本実施の形態に係る動作環境移動方法を説明する。まず、移動元サーバ100の動作が中断される(ステップS10)。これにより、移動元サーバ100による業務が中断する。

【0022】

次に、移動元ディスク110から移動先ディスク210にコピーされるべきコピーイメージCIに含まれるファイルのリスト(インデックス)が作成される。ここで、コピーイメージCIは、移動元ディスク100のディスクイメージ10の全体の少なくとも一部である。また、作成されるリストは、以下、「コピーリストCL」と参照される。作成されたコピーリストCLは、移動先サーバ200がアクセス可能なディスクに格納される。例えば図2に示されるように、コピーリストCLは、移動先ディスク210に格納される(ステップS20)。更に、移動元サーバ100の実行コンテキストECが、移動先サーバ200にコピーされる。つまり、実行コンテキストECが、移動元サーバ100から移動先サーバ200に移動する(ステップS30)。

20

【0023】

次に、移動先サーバ200において動作が再開される(ステップS40)。つまり、移動先サーバ200は、実行コンテキストECを受け取った段階で、中断していた業務を再開する。この段階で、コピーイメージCIの実体は未だコピーされていない。その代わりに、コピーイメージCIのファイルリストであるコピーリストCLが、移動先サーバ200がアクセス可能なディスクに格納されている。

30

【0024】

その後、移動先サーバ200の業務の最中に、コピーイメージCIが、移動元ディスク110から移動先ディスク210にコピーされる(ステップS50)。コピーイメージCIのコピーは、上記コピーリストCLに基づいて行われる。例えば、コピーイメージCIのコピーは、移動先サーバ200の要求に回答してオンデマンド方式で行われる。また、コピーイメージCIは、移動先サーバ200の動作のバックグラウンドでコピーされてもよい。コピーイメージCIのコピーが完了すると、動作環境の移動処理は終了する。

【0025】

図4は、従来技術と本発明との比較を示している。従来技術によれば、時刻t0における業務中断の後、ディスクイメージ全体のコピー(ステップS120)及び実行コンテキストのコピー(ステップS130)が実行される。その後、時刻t1において、移動先サーバで業務が再開される。

40

【0026】

本発明によれば、時刻t0における業務中断から時刻t1'における業務再開までの間、実行コンテキストECはコピーされる(ステップS30)。しかし、移動元ディスク110上のコピーイメージCIに関しては、そのファイルリストであるコピーリストCLだけが作成され(ステップS20)、コピーイメージCIの実体のコピーは行われない。その段階で、移動先サーバ200において動作が再開される。コピーイメージCIの実体が

50

コピーされないため、業務中断から業務再開までの期間が T A から T B に短縮される。すなわち、業務の中断時間が短縮される。

【 0 0 2 7 】

コピーイメージ C I の実体に関しては、移動先サーバ 2 0 0 における動作再開後、業務の最中に、オンデマンド・コピーあるいはバックグラウンド・コピーを行うことが可能である。それは、移動先サーバ 2 0 0 がアクセス可能なようにコピーリスト C L が作成されているからである。コピーイメージ C I のコピーは、時刻 t 2 に完了する。動作環境 (= 実行コンテキスト E C + コピーイメージ C I) 全体のコピー時間は長くなる場合があるが、移動先サーバ 2 0 0 の業務開始までの時間は短縮される。このことは、サービス提供の継続性の観点から好ましい。

10

【 0 0 2 8 】

以下、本発明に係るサーバ動作環境の移動方法に関して詳細に説明する。

【 0 0 2 9 】

2 . 第 1 の実施の形態

第 1 の実施の形態において、移動元サーバ 1 0 0 及び移動先サーバ 2 0 0 は、サーバ仮想化技術によって実現される仮想マシン (V M : Virtual Machine) である。

【 0 0 3 0 】

2 - 1 . 構成

図 5 は、第 1 の実施の形態に係る構成を示している。図 5 において、実計算機 1 2 0 及び 2 2 0 が、LAN 等のネットワークを介して相互に接続されている。実計算機 1 2 0 は、移動元ディスク 1 1 0 にアクセス可能であり、移動元マシン 1 2 0 と参照される。一方、実計算機 2 2 0 は、移動先ディスク 2 1 0 にアクセス可能であり、移動先マシン 2 2 0 と参照される。

20

【 0 0 3 1 】

実計算機 1 2 0、2 2 0 の各々には、仮想マシンモニタ (V M M : Virtual Machine Monitor) が搭載されている。V M M は、実計算機上で動作するソフトウェアであり、複数の V M を構築、制御することができる。具体的には、移動元マシン 1 2 0 に搭載された移動元 V M M 1 3 0 が、仮想サーバである移動元サーバ 1 0 0 を提供し、またその動作を制御する。また、移動先マシン 2 2 0 に搭載された移動先 V M M 2 3 0 が、仮想サーバである移動先サーバ 2 0 0 を提供し、またその動作を制御する。移動元サーバ 1 0 0 は移動元ディスク 1 1 0 を使用し、移動先サーバ 2 0 0 は移動先ディスク 2 1 0 を使用する。

30

【 0 0 3 2 】

本実施の形態においては、移動元サーバ 1 0 0 の動作環境の移動先サーバ 2 0 0 への移動は、移動元 V M M 1 3 0 と移動先 V M M 2 3 0 の協働によって実現される。そのため、移動元 V M M 1 3 0 は、コピーリスト作成モジュール 2 0 とコンテキストコピーモジュール 3 0 を有している。また、移動先 V M M 2 3 0 は、コピー判定モジュール 4 0 とコピーモジュール 5 0 を有している。

【 0 0 3 3 】

2 - 2 . 動作

次に、本実施の形態におけるサーバシステムの動作を説明する。動作環境の移動前、移動元マシン 1 2 0 において、移動元 V M M 1 3 0 が移動元サーバ 1 0 0 を構築しており、移動元サーバ 1 0 0 のゲスト O S (G O S) が様々な処理を実行している。一方、移動先マシン 2 2 0 において、移動先 V M M 2 3 0 は移動先サーバ 2 0 0 の仮想ハードウェア構成を提供することができるが、ゲスト O S は未定義であり、その移動先サーバ 2 0 0 は未だ稼動していない。また、V M 環境の場合、動作環境の移動に先立って、移動元サーバ 1 0 0 と移動先サーバ 2 0 0 との間でハードウェア構成 (I O カードのスロット位置など) の比較を行う必要はない。但し、移動先ディスク 2 1 0 の容量は、移動元ディスク 1 1 0 の容量以上である必要はある。

40

【 0 0 3 4 】

図 3 及び図 5 を参照して、本実施の形態におけるサーバ動作環境の移動方法を説明する

50

。本実施の形態において、コピーイメージC Iは、移動元ディスク110のディスクイメージ10の全体である。

【0035】

ステップS10：

まず、ユーザによる指示に応答して、移動元VMM130は、移動元サーバ100の動作を中断させる。これにより、移動元サーバ100による業務が中断する。

【0036】

ステップS20：

移動元VMM130のコピーリスト作成モジュール20は、移動元ディスク110上のコピーイメージC Iを参照し、コピーリストC Lを作成する。コピーリストC Lは、コピーイメージC Iに含まれるファイルのリストであり、移動先ディスク210へのコピーが必要なファイルのリストである。コピーリスト作成モジュール20は、作成されたコピーリストC Lを移動先VMM230に転送する。移動先VMM230は、受け取ったコピーリストC Lを移動先ディスク210に格納する。

10

【0037】

図6は、コピーリストC Lの一例を示している。図6に示されるように、コピーリストC Lは、ヘッダ部C L - Hとボディ部C L - Bから構成されている。ヘッダ部C L - Hは、移動元サーバ100に関する情報を示しており、ネットワークアドレス、認証情報、ゲストOS情報などを含んでいる。例えば、ネットワークアドレスは、移動元サーバ100がネットワーク上のどこに存在するかを示し、ゲストOS情報は、移動されるゲストOSのIDを示している。

20

【0038】

一方、ボディ部C L - Bは、コピーイメージC Iのファイルリストを示している。例えば図6に示されるように、ボディ部C L - Bは、各ファイルに関する名前、格納されているディスクのID、そのディスク内の位置（開始ブロック）、ブロックサイズなどを示している。更に、ボディ部C L - Bは、各ファイルがコピー済みか否かを示す「フラグ（コピー情報）」を有している。コピーリストC Lの作成時、全てのフラグの状態は「未コピー」に設定されている。

【0039】

後述されるように、アクセス対象のファイルは、コピーリストC Lのボディ部C L - Bから検索される。その検索速度を高めるため、図7に示されるようなハッシュテーブルHTが利用されてもよい。ハッシュテーブルHTには、ハッシュキーとIDが対応付けられている。そのIDは、コピーリストC Lのボディ部C L - Bに示されているIDと対応している。このようなハッシュテーブルHTを利用することによって、コピーリストC L上に当該ファイルエントリが存在するか否かを高速に検索することができる。ハッシュテーブルHTが利用される場合、コピーリスト作成モジュール20は、コピーリストC Lの作成と同時に、コピーリストC L中のIDに基づいてハッシュテーブルHTを作成する。そして、コピーリスト作成モジュール20は、作成されたハッシュテーブルHTを移動先VMM230に転送する。移動先VMM230は、そのハッシュテーブルHTを移動先サーバ200のRAMに展開する。

30

40

【0040】

ステップS30：

移動元VMM130のコンテキストコピーモジュール30は、移動元サーバ100の実行コンテキストE Cを、移動先VMM230に送信する。移動先VMM230は、その実行コンテキストE Cを移動先サーバ200に転送する。このようにして、移動元サーバ100の実行コンテキストE Cが、移動先サーバ200にコピーされる。この時点で、移動先サーバ200において、ゲストOS（GOS）が定義される。

【0041】

ステップS40：

移動先VMM230は、移動先サーバ200の動作を再開させる。これにより、移動先

50

サーバ 200 において、中断されていた業務が再開する。

【 0042 】

ステップ S 50 :

その後、移動先サーバ 200 の業務の最中に、コピーイメージ C I が、移動元ディスク 110 から移動先ディスク 210 にコピーされる。コピーイメージ C I のコピーは、オンデマンドあるいは / 及びバックグラウンドで行われる。図 8 A 及び図 8 B は、ステップ S 50 における処理を詳しく示すフローチャートである。

【 0043 】

ステップ S 51 :

図 8 A に示されるオンデマンド・コピーの場合、まず、動作中のプログラム（例えば、ゲスト OS）から移動先ディスク 210 へのアクセス（リード要求）が発生する。この時、後述のバックグラウンド・コピーは、一時中断される。

10

【 0044 】

ステップ S 52 :

移動先 VMM 230 のコピー判定モジュール 40 は、アクセス対象である対象データの実体が移動先ディスク 210 に存在するか否かの判定を行う。本実施の形態において、コピー判定モジュール 40 は、上述のコピーリスト C L を参照し、そのコピーリスト C L に基づいて判定を行う。

【 0045 】

具体的には、コピー判定モジュール 40 は、対象データがコピーリスト C L に含まれているかを調べる。ここで、コピー判定モジュール 40 は、ハッシュテーブル H T を利用し、ハッシュテーブル H T において当該ファイルがヒットするか否かを調べてもよい。対象データがコピーリスト C L に含まれている場合、コピー判定モジュール 40 は、その対象データに対応するフラグを調べる（図 6 参照）。コピー判定モジュール 40 は、そのフラグが「コピー済」あるいは「未コピー」かによって、対象データの実体が移動先ディスク 210 に存在するか否かを判定することができる。

20

【 0046 】

ステップ S 53 :

フラグが「コピー済」の場合（ステップ S 53 ; N o）、対象データは既にコピーされている。従って、移動先ディスク 210 上の指定されたブロックから、対象データが読み出される（ステップ S 56）。読み出されたデータは、データの要求元に返される。一方、フラグが「未コピー」の場合（ステップ S 53 ; Y e s）、対象データは未だコピーされていない。つまり、未コピー領域へのアクセスが発生していることがわかる。その場合、処理は、ステップ S 54 に移行する。

30

【 0047 】

ステップ S 54 :

移動先 VMM 230 は、移動元 VMM 130 に対して通信を行い、対象データのコピーを要求する。具体的には、移動先 VMM 230 のコピーモジュール 50 は、コピーリスト C L のヘッダ部 C L - H 及びボディ部 C L - B を参照し、移動元や対象データの格納情報を認識する（図 6 参照）。そして、コピーモジュール 50 は、それら情報に基づいて、移動元 VMM 130 に対して対象データのコピーを要求する。そのコピー要求に応答して、移動元 VMM 130 は、コピーイメージ C I 中の対象データを移動元ディスク 110 から読み出し、読み出された対象データを移動先 VMM 230 に返す。移動先 VMM 230 は、対象データを、移動先ディスク 210 上の指定されたブロックに書き込む。このようにして、対象データの実体が、移動元ディスク 110 から移動先ディスク 210 にコピーされる。

40

【 0048 】

ステップ S 55 :

対象データが移動先ディスク 210 にコピーされると、コピーモジュール 50 は、コピーリスト C L を更新する。つまり、コピーモジュール 50 は、対象データに対応するフラ

50

グを、「未コピー」から「コピー済」に変更する。その後、移動先ディスク 2 1 0 上の指定されたブロックから、対象データが読み出される（ステップ S 5 6）。

【 0 0 4 9 】

バックグラウンド・コピーは、上記オンデマンド・コピーを補完する役割を果たす。コピーイメージ C I にはなかなかアクセスされないファイルも含まれるため、オンデマンドコピーだけでは、コピーイメージ C I 全体のコピーが終了するまでの時間が長くなる可能性がある。バックグラウンド・コピーを併用することによって、コピーイメージ C I 全体のコピー時間が短縮される。バックグラウンド・コピーは、移動先 V M M 2 3 0 によって同様に行われ得る。

【 0 0 5 0 】

ステップ S 5 7 :

図 8 B に示されるバックグラウンド・コピーの場合、移動先 V M M 2 3 0 が、開始要求を発行する。例えば移動先 V M M 2 3 0 が、システム負荷をモニタし、システム負荷が軽い場合にコピー判定モジュール 4 0 に対してバックグラウンド・コピーの開始を指示する。

10

【 0 0 5 1 】

ステップ S 5 8 :

コピー判定モジュール 4 0 は、コピーリスト C L を参照し、ファイルリストの先頭からファイルを選択する。

【 0 0 5 2 】

その後は、オンデマンド・コピーの場合と同様である。コピー判定モジュール 4 0 は、コピーリスト C L 中の選択されたファイルに対応するフラグを調べる（ステップ S 5 2）。フラグが「コピー済」の場合（ステップ S 5 3 ; N o）、処理はステップ S 5 8 に戻り、次のファイルが選択される。フラグが「未コピー」の場合（ステップ S 5 3 ; Y e s）、上述のステップ S 5 4、S 5 5 が実行される。ステップ S 5 5 が終わると、処理はステップ S 5 8 に戻り、次のファイルが選択される。オンデマンド・コピーが発生したり、システム負荷が重くなった場合、移動先 V M M 2 3 0 は、バックグラウンド・コピーを中断させる。

20

【 0 0 5 3 】

このようにして、移動先サーバ 2 0 0 の業務の途中に、コピーイメージ C I を移動元ディスク 1 1 0 から移動先ディスク 2 1 0 にコピーすることが可能となる。従って、移動元サーバ 1 0 0 の動作中断から移動先サーバ 2 0 0 における動作再開までの間に、コピーイメージ C I の実体をコピーする必要はない。その結果、業務の中断時間が短縮される。尚、本実施の形態において、移動元 V M M 1 3 0 や移動先 V M M 2 3 0 が行っていた処理は、所定の管理 V M（仮想管理サーバ）によって実行されてもよい。

30

【 0 0 5 4 】

3 . 第 2 の実施の形態

第 2 の実施の形態において、移動元サーバ 1 0 0 及び移動先サーバ 2 0 0 は、実計算機によって実現される。尚、第 1 の実施の形態と同じ構成には同じ符号が付され、重複する説明は適宜省略される。

40

【 0 0 5 5 】

3 - 1 . 構成

図 9 は、第 2 の実施の形態に係る構成を示している。本実施の形態に係るサーバシステムは、L A N 等のネットワークを介して相互に接続された移動元サーバ 1 0 0、移動先サーバ 2 0 0、及び管理サーバ 3 0 0 を備えている。管理サーバ 3 0 0 は、サーバ全体の管理を行うサーバである。移動元ディスク 1 1 0 は移動元サーバ 1 0 0 に直結されており、移動先ディスク 2 1 0 は移動先サーバ 2 0 0 に直結されている。管理サーバ 3 0 0 は、移動元ディスク 1 1 0 と移動先ディスク 2 1 0 の両方にアクセス可能である。

【 0 0 5 6 】

本実施の形態において、管理サーバ 3 0 0 が、コピーリスト作成モジュール 2 0 及びコ

50

ンテキストコピーモジュール 30 を有している。これらモジュール 20、30 は、演算処理装置とソフトウェアの協働によって実現される。

【0057】

また、本実施の形態において、移動先サーバ 200 が、コピー判定モジュール 40 及びコピーモジュール 50 を有している。例えば、移動先サーバ 200 には、ファームウェアが格納された ROM 60 が搭載されている。その ROM 60 に格納されたファームウェアによって、コピー判定モジュール 40 及びコピーモジュール 50 が実現される。

【0058】

3 - 2 . 動作

次に、本実施の形態におけるサーバシステムの動作を説明する。動作環境の移動に先立って、移動元サーバ 100 と移動先サーバ 200 との間でハードウェア構成（I/Oカードのスロット位置など）の比較が行われ、ディスクイメージの継承が可能であることが確認される。

10

【0059】

図 3 及び図 9 を参照して、本実施の形態におけるサーバ動作環境の移動方法を説明する。本実施の形態において、コピーイメージ C I は、移動元ディスク 110 のディスクイメージ 10 の全体である。

【0060】

ステップ S 10 :

まず、ユーザあるいは管理サーバ 300 によって、移動元サーバ 100 の動作が中断される。これにより、移動元サーバ 100 による業務が中断する。

20

【0061】

ステップ S 20 :

管理サーバ 300 のコピーリスト作成モジュール 20 は、移動元ディスク 110 上のコピーイメージ C I にアクセスし、コピーリスト C L を作成する。コピーリスト作成モジュール 20 は、作成されたコピーリスト C L を移動先ディスク 210 に格納する。また、コピーリスト作成モジュール 20 は、コピーリスト C L の作成と同時に、ハッシュテーブル H T を作成してもよい。その場合、コピーリスト作成モジュール 20 は、作成されたハッシュテーブル H T を移動先サーバ 200 の R A M に展開する。

30

【0062】

ステップ S 30 :

管理サーバ 300 のコンテキストコピーモジュール 30 は、移動元サーバ 100 にアクセスし、移動元サーバ 100 の実行コンテキスト E C を得る。そして、コンテキストコピーモジュール 30 は、その実行コンテキスト E C を移動先サーバ 200 に送る。このようにして、移動元サーバ 100 の実行コンテキスト E C が、移動先サーバ 200 にコピーされる。

【0063】

ステップ S 40 :

移動先サーバ 200 において、動作が開始し、中断されていた業務が再開する。ROM 60 中のコピー判定モジュール 40 及びコピーモジュール 50 も動作を開始する。

40

【0064】

ステップ S 50 :

その後、移動先サーバ 200 の業務の最中に、コピーイメージ C I が、移動元ディスク 110 から移動先ディスク 210 にコピーされる。そのコピーにおける処理フローは、既出の図 8 A 及び図 8 B に示されたフローと同様である。

【0065】

ステップ S 51 :

まず、動作中の O S あるいはプログラムから移動先ディスク 210 へのアクセス（リード要求）が発生する。この時、バックグラウンド・コピーは、一時中断される。

【0066】

50

ステップ S 5 2 :

アクセスが発生すると、コピー判定モジュール 4 0 は、アクセス対象である対象データの実体が移動先ディスク 2 1 0 に存在するか否かの判定を行う。具体的には、コピー判定モジュール 4 0 は、上述のコピーリスト C L を参照し、対象データに対応するフラグに基づいて判定を行う。この際、ハッシュテーブル H T が利用されてもよい。

【 0 0 6 7 】

ステップ S 5 3 :

フラグが「コピー済」の場合（ステップ S 5 3 ; N o ）、移動先ディスク 2 1 0 上の指定されたブロックから、対象データが読み出される（ステップ S 5 6 ）。読み出されたデータは、データの要求元に返される。一方、フラグが「未コピー」の場合（ステップ S 5 3 ; Y e s ）、処理は、ステップ S 5 4 に移行する。

10

【 0 0 6 8 】

ステップ S 5 4 :

移動先サーバ 2 0 0 のコピーモジュール 5 0 は、コピーリスト C L のヘッダ部 C L - H 及びボディ部 C L - B を参照し、移動元や対象データの格納情報を認識する（図 6 参照）。そして、コピーモジュール 5 0 は、それら情報に基づいて、対象データのコピーを指示する。

【 0 0 6 9 】

例えば、コピーモジュール 5 0 は、管理サーバ 3 0 0 に対して、対象データのコピーを要求する。管理サーバ 3 0 0 は、そのコピー要求に基づいて移動元ディスク 1 1 0 にアクセスし、コピーイメージ C I から対象データを読み出す。そして、管理サーバ 3 0 0 は、移動先ディスク 2 1 0 にアクセスし、指定されたブロックに対象データを書き込む。このようにして、対象データの実体が、移動元ディスク 1 1 0 から移動先ディスク 2 1 0 にコピーされる。

20

【 0 0 7 0 】

あるいは、対象データのコピーは、移動先サーバ 2 0 0 と移動元サーバ 1 0 0 との間の R P C (Remote Procedure Call) 通信によって実現されてもよい。コピーモジュール 5 0 は、移動元サーバ 1 0 0 に対して、対象データのコピーを要求する。そのコピー要求に回答して、移動元サーバ 1 0 0 の O S は一時動作再開し、移動元ディスク 1 1 0 上のコピーイメージ C I から対象データを読み出す。移動元サーバ 1 0 0 の O S は、読み出された対象データを、移動先サーバ 2 0 0 に送り返す。コピーモジュール 5 0 は、対象データを、移動先ディスク 2 1 0 上の指定されたブロックに書き込む。このようにして、対象データの実体が、移動元ディスク 1 1 0 から移動先ディスク 2 1 0 にコピーされる。

30

【 0 0 7 1 】

尚、この R P C 通信の場合、移動先サーバ 2 0 0 が、対象データの具体的な場所やサイズを認識する必要はない。それは、移動元サーバ 1 0 0 が対象データの具体的な場所やサイズを認識しているからである。よって、本実施の形態において作成されるコピーリスト C L は、図 1 0 に示されるように簡素化されてもよい。図 6 に示されたコピーリスト C L と比較して、ディスク I D、開始ブロック、サイズが省略されている。図 1 0 に示されるコピーリスト C L は、ファイル名とフラグ（コピー情報）との対応関係を示している。この場合、コピーリスト C L のサイズが縮小され、好適である。コピーモジュール 5 0 は、ファイル名を指定することによって、移動元サーバ 1 0 0 に対して対象データのコピーを要求する。

40

【 0 0 7 2 】

ステップ S 5 5 :

対象データコピーが完了すると、コピーモジュール 5 0 は、対象データに対応するフラグを、「未コピー」から「コピー済」に変更する。その後、移動先ディスク 2 1 0 上の指定されたブロックから、対象データが読み出される（ステップ S 5 6 ）。

【 0 0 7 3 】

ステップ S 5 7 :

50

バックグラウンド・コピーは、上記オンデマンド・コピーを補完する役割を果たす。移動先サーバ200のOSは、開始要求を発行する。例えばそのOSが、システム負荷をモニタし、システム負荷が軽い場合にコピー判定モジュール40に対してバックグラウンド・コピーの開始を指示する。

【0074】

ステップS58：

コピー判定モジュール40は、コピーリストCLを参照し、ファイルリストの先頭からファイルを選択する。

【0075】

その後は、オンデマンド・コピーの場合と同様である。コピー判定モジュール40は、コピーリストCL中の選択されたファイルに対応するフラグを調べる(ステップS52)。フラグが「コピー済」の場合(ステップS53; No)、処理はステップS58に戻り、次のファイルが選択される。フラグが「未コピー」の場合(ステップS53; Yes)、上述のステップS54、S55が実行される。ステップS55が終わると、処理はステップS58に戻り、次のファイルが選択される。オンデマンド・コピーが発生したり、システム負荷が重くなった場合、OSは、バックグラウンド・コピーを中断させる。

10

【0076】

このようにして、移動先サーバ200の業務の途中に、コピーイメージCIを移動元ディスク110から移動先ディスク210にコピーすることが可能となる。従って、移動元サーバ100の動作中断から移動先サーバ200における動作再開までの間に、コピーイメージCIの実体をコピーする必要はない。その結果、業務の中断時間が短縮される。

20

【0077】

4. 第3の実施の形態

4-1. 概念

既出の実施の形態では、ステップS50においてコピーされるコピーイメージCIは、ディスクイメージ10の全てであった。それに対し、第3の実施の形態では、ディスクイメージ10の一部が、業務中断前にあらかじめ移動先ディスク210に配布される。そして、業務再開後のステップS50においては、残りのイメージだけがコピーされる。例えば、移動元サーバ100におけるOSのインストール直後の汎用的なイメージが、移動先となり得る候補に予め配布される。その場合、OSインストール後の変更部分が、ステップS50において、コピーイメージCIとしてコピーされる。

30

【0078】

より一般化すると、次の通りである。図11は、本実施の形態におけるディスクイメージ10の分割を概念的に示している。ディスクイメージ10は、第1のイメージ11と第2のイメージ12に区分けされる。第1のイメージ11は、コピーイメージCIの作成の基準となるイメージであり、「ベースイメージBI」と参照される。このベースイメージBIが、業務中断前に、移動先となり得る候補に予め配布される。一方、第2のイメージ12は、ディスクイメージ10とベースイメージBIとの差分である。ベースイメージBIの作成後に追加された部分、及び、ベースイメージBI中で変更された部分が、第2のイメージ12に含まれる。この第2のイメージ12が、ステップS50においてコピーイメージCIとしてコピーされる。

40

【0079】

図12は、本実施の形態に係るサーバ動作環境の移動方法を示すフローチャートである。以下の説明において、既出の実施の形態と重複する部分は適宜省略される。

【0080】

本実施の形態において、上述のステップS10より前に、新たにステップS0が設けられている。つまり、ステップS10より前のあるタイミングで、ベースイメージBIが作成される。そのベースイメージBIは、移動元ディスク110のディスクイメージ10の一部である。作成されたベースイメージBIは、所定のディスクに保管される。そして、ステップS10より前の段階で、そのベースイメージBIは、移動先サーバ200の移動

50

先ディスク 210 に予め配布される。

【0081】

コピーイメージ C I もディスクイメージ 10 の一部である。このコピーイメージ C I は、ステップ S 10 の段階でのディスクイメージ 10 と上記ベースイメージ B I との差分である。ベースイメージ B I の作成後に追加された部分、及び、ベースイメージ B I 中で変更された部分が、コピーイメージ C I に含まれる。ステップ S 20 においては、そのコピーイメージ C I に関するコピーリスト C L が作成される。ステップ S 50 においては、そのコピーリスト C L を参照することにより、コピーイメージ C I がコピーされる。

【0082】

本実施の形態によれば、既出の実施の形態と比較して、コピーイメージ C I のサイズが縮小される。その結果、ステップ S 50 に要する時間が短縮される。言い換えれば、図 4 で示された期間 $t_0 \sim t_2$ が短縮され、動作環境全体のコピーが完了するまでの期間が短縮される。尚、本実施の形態は、既出の実施の形態のいずれとも組み合わせることが可能である。以下、具体的な例を挙げて、本実施の形態を更に詳しく説明する。

10

【0083】

4 - 2 . 構成

図 13 は、本実施の形態に係る構成の一例を示している。図 13 に示された構成は、第 1 の実施の形態で示された構成 (V M 環境) に対応している。移動元 V M M 130 は、更に、ベースイメージ作成モジュール 70 とベースイメージ配布モジュール 80 を有している。ベースイメージ作成モジュール 70 は、ベースイメージ B I を作成する。ベースイメージ配布モジュール 80 は、そのベースイメージ B I を移動先ディスク 210 に配布する。

20

【0084】

図 14 は、本実施の形態に係る構成の他の例を示している。図 14 に示された構成は、第 2 の実施の形態で示された構成 (実計算機環境) に対応している。図 14 においては、管理サーバ 300 が、ベースイメージ作成モジュール 70 とベースイメージ配布モジュール 80 を有している。これらモジュール 70、80 は、演算処理装置とソフトウェアの協働によって実現される。

【0085】

4 - 3 . 動作例

図 12 ~ 図 14 を参照して、本実施の形態におけるサーバ動作環境の移動方法を説明する。本実施の形態において、コピーイメージ C I は、移動元ディスク 110 のディスクイメージ 10 の一部である。

30

【0086】

ステップ S 0 :

ステップ S 10 より前に、ベースイメージ作成モジュール 70 は、移動元ディスク 110 にアクセスし、ベースイメージ B I を作成する。このベースイメージ B I は、現時点でのディスクイメージ 10 のバックアップであると言える。作成されたベースイメージ B I は、所定のディスクに格納される。例えば、作成されたベースイメージ B I は、移動元ディスク 110 に格納される。

40

【0087】

あるいは、ベースイメージ作成モジュール 70 は、バックアップとしてのベースイメージ B I を作成する代わりに、現時点でのディスクイメージ 10 のファイルのリストを作成してもよい。図 15 は、そのようなファイルリスト (以下、「ベースリスト B L」と参照される) を示している。ベースリスト B L は、ステップ S 0 の時点における移動元ディスク 110 のディスクイメージ 10 に含まれるファイルをリスト形式で示している。

【0088】

その後、移動元サーバ 100 の動作は継続し、移動元ディスク 110 上でファイルの追加・変更が行われる。

【0089】

50

また、ステップ S 1 0 以前のあるタイミングで、ベースイメージ配布モジュール 8 0 は、ベースイメージ B I を読み出し、そのベースイメージ B I を移動先となり得る候補に配布する。バックアップとしてのベースイメージ B I が存在する場合、ベースイメージ配布モジュール 8 0 は、そのベースイメージ B I をそのまま読み出す。一方、図 1 5 に示されたベースリスト B L が作成されている場合、ベースイメージ配布モジュール 8 0 は、そのベースリスト B L を参照して、移動元ディスク 1 1 0 からベースイメージ B I に相当する部分を読み出す。図 1 3 に示された V M 環境の場合、移動元 V M M 1 3 0 のベースイメージ配布モジュール 8 0 は、ベースイメージ B I を移動先 V M M 2 3 0 に送る。移動先 V M M 2 3 0 は、そのベースイメージ B I を移動先ディスク 2 1 0 に格納する。図 1 4 に示された実計算機環境の場合、管理サーバ 3 0 0 のベースイメージ配布モジュール 8 0 は、ベースイメージ B I を移動先ディスク 2 1 0 に直接書き込むことができる。

10

【 0 0 9 0 】

ステップ S 1 0 :

その後、ステップ S 1 0 において、移動元サーバ 1 0 0 の動作が中断する。この時点でのディスクイメージ 1 0 と上記ベースイメージ B I との差が、コピーイメージ C I となる。ステップ S 0 後に追加された部分、及び、ベースイメージ B I 中で変更された部分が、コピーイメージ C I に含まれる。

【 0 0 9 1 】

図 1 6 は、ベースイメージ B I とコピーイメージ C I に関する様々なパターンを示している。例えばパターン (a) の場合、移動元サーバ 1 0 0 における O S インストール直後のディスクイメージが、ベースイメージ B I として利用される。つまり、O S インストール直後に上記ステップ S 0 が実行される。この場合のベースイメージ B I は汎用的であり、移動先候補への配布が容易となる。一方、O S インストール後の追加・変更部分が、コピーイメージ C I となる。その追加・変更部分は、移動元ディスク 1 1 0 上のログファイル L O G に記録される。コピーイメージ C I のファイルリストであるコピーリスト C L は、そのログファイル L O G から作成することができる。

20

【 0 0 9 2 】

パターン (b) の場合、アプリケーションのインストール、業務・ユーザ設定が完了した時点におけるディスクイメージが、ベースイメージ B I として利用される。つまり、アプリケーションのインストール、業務・ユーザ設定が完了した時点で、上記ステップ S 0 が実行される。そして、その後の追加・変更部分がコピーイメージ C I となる。その追加・変更部分は、移動元ディスク 1 1 0 上のログファイル L O G に記録される。この場合のベースイメージ B I は、移動元サーバ 1 0 0 の業務に特化している。しかし、パターン (a) の場合と比較して、ステップ S 5 0 でコピーされるコピーイメージ C I のサイズが、より縮小されている。

30

【 0 0 9 3 】

パターン (c) の場合、移動元サーバ 1 0 0 の運用中の所定のタイミングが、“ 起点 (ステップ S 0) ” となる。その起点におけるディスクイメージが、スプリットミラー機能等によって瞬時に取り込まれ、ベースイメージ B I として利用される。そして、その起点以降の追加・変更部分がコピーイメージ C I となる。同様に、その追加・変更部分は、移動元ディスク 1 1 0 上のログファイル L O G に記録される。このパターン (c) の場合、ステップ S 5 0 でコピーされるコピーイメージ C I のサイズが、最も縮小されている。

40

【 0 0 9 4 】

ステップ S 0 後、ログファイル L O G を作成するための方法の一例を、以下に説明する。図 1 7 は、ステップ S 0 以降のディスクイメージの変更の一例を、概念的に示している。あるファイル / a / f i l e 1 が削除される一方、他のファイル / b / c 及び / b / c / f i l e 3 が作成されている。図 1 8 は、図 1 7 で示された変更に対応する動作シーケンスの一例を示している。

【 0 0 9 5 】

1 . / a / f i l e 1 に対する書き込み

50

2. / a / f i l e 1 の削除
3. / a / f i l e 2 に対する読み込み
4. / b / c の作成
5. / b / c / f i l e 3 の作成
6. / b / c / f i l e 3 に対する書き込み

【0096】

図18には、上記動作シーケンスに応じて作成されるログファイルLOGも示されている。ログファイルLOGの各エントリには、操作コードopとファイル名が対応付けられて格納される。操作コードopは、「M(modify)」あるいは「D(delete)」に設定される。書き込みの場合、ファイルの内容が変更されるため、「M」が選択される。読み込みの場合、ファイルの内容は変更されないため、ログファイルLOGに記録を残す必要はない。ファイルの削除の場合、「D」が選択される。ファイルの作成の場合、「M」が選択される。ファイルの移動の場合、移動元のファイルに関しては「D」が選択され、移動先のファイルに関しては「M」が選択される。このようにして、移動元ディスク110上のファイルに対する操作の履歴を示すログファイルLOGが作成される。

10

【0097】

ステップS20:

コピーリスト作成モジュール20は、コピーイメージCIに含まれるファイルのリストであるコピーリストCLを作成する。ここで、上述のログファイルLOGは、ベースイメージBIの作成(ステップS0)から業務中断(ステップS10)までのファイルの変更情報を示している。従って、コピーリスト作成モジュール20は、ログファイルLOGを参照し解析することによって、コピーリストCLを作成することができる。

20

【0098】

図19は、図18に示されたログファイルLOGからコピーリストCLが作成される例を示している。コピーリストCLにおいて、1つのファイルに1つのエントリが割り当てられている。ファイル名は、ログファイルLOGに現れるファイル名がそのまま使用されている。ログファイルLOGに同じファイル名で複数のエントリが存在する場合、ログファイルLOG中で最後に現れるエントリが参照される。

【0099】

例えば、ファイル/a/f i l e 1に関しては、ログファイルLOG中のエントリ「D, / a / f i l e 1」が参照される。操作コードopが「D」の場合、コピーリストCL中のIDとして「-1」が割り振られる。ファイル/b/cに関しては、ログファイルLOG中のエントリ「M, / b / c」が参照される。操作コードopが「M」の場合、コピーリストCL中のIDとして「固有ID」が割り振られる。この場合、固有IDは「1」である。ファイル/b/c/f i l e 3に関しては、ログファイルLOG中のエントリ「M, / b / c / f i l e 3」が参照される。この場合、固有IDとして「2」が割り振られる。全てのフラグは、初期状態として「未コピー」に設定される。

30

【0100】

このようにして、コピーリストCLが作成される。図19において、コピーリストCLの各エントリは、割り当てられたIDと、ファイル名と、フラグとを有している。更に、コピーリストCL中のIDに基づいて、上述のハッシュテーブルHTを作成することも可能である。

40

【0101】

ステップS50におけるバックグラウンドコピーにおいて、IDが「-1」であるファイルは、移動先ディスク210から削除される。他の処理に関しては、既出の実施の形態と同様である。

【0102】

尚、VM環境では、ある時点のイメージとそれ以降の変更部分を分けて管理する技術として、「VMware nonpersistent mode」や「UnionFS」といった機能も知られている。それら機能を利用することによって、ベースイメージBI

50

とコピーイメージC Iの管理が行われてもよい。

【0103】

他の例として、図20は、ベースイメージB IとコピーイメージC Iの分割が、移動元ディスク110のファイル管理によって実現される例を示している。図20に示されるように、移動元ディスク110は、第1の領域91と第2の領域92に分けて管理されている。第1の領域91には、プログラム等の変更されないファイルだけが格納される。一方、第2の領域92には、アプリケーションデータ等の変更され得るファイルが格納される。ファイルの書き換えは、第2の領域92に対してだけ実行される。この場合、第1の領域91のディスクイメージが、ベースイメージB Iとして利用されるとよい。また、第2の領域92のディスクイメージが、コピーイメージC Iとしてコピーされるとよい。この場合、変更部分を追跡するためのログファイルLOGは不要となる。また、コピーリスト作成モジュール20は、第2の領域92にアクセスすることによって、コピーリストCLを作成することができる。

10

【0104】

本実施の形態によれば、既出の実施の形態と同じ効果が得られる。すなわち、業務の中断時間が短縮される。更に、ステップS50においてコピーすべき量が減少するため、図4における期間t0~t2が短縮されるという追加的な効果が得られる。言い換えれば、動作環境全体のコピーが完了するまでの期間が短縮される。

【図面の簡単な説明】

【0105】

20

【図1】図1は、従来技術に係るサーバ動作環境の移動方法を示すフローチャートである。

【図2】図2は、本発明に係るサーバシステムの構成を概略的に示すブロック図である。

【図3】図3は、本発明に係るサーバ動作環境の移動方法を示すフローチャートである。

【図4】図4は、本発明の効果を説明するための図である。

【図5】図5は、本発明の第1の実施の形態に係るサーバシステムの構成を示すブロック図である。

【図6】図6は、コピーリストの一例を示す概念図である。

【図7】図7は、ハッシュテーブルの一例を示す概念図である。

【図8A】図8Aは、本実施の形態におけるコピーイメージのコピー方法を示すフローチャートである。

30

【図8B】図8Bは、本実施の形態におけるコピーイメージのコピー方法を示すフローチャートである。

【図9】図9は、本発明の第2の実施の形態に係るサーバシステムの構成を示すブロック図である。

【図10】図10は、コピーリストの一例を示す概念図である。

【図11】図11は、本発明の第3の実施の形態におけるディスクイメージを示す概念図である。

【図12】図12は、第3の実施の形態に係るサーバ動作環境の移動方法を示すフローチャートである。

40

【図13】図13は、第3の実施の形態に係るサーバシステムの構成の一例を示すブロック図である。

【図14】図14は、第3の実施の形態に係るサーバシステムの構成の他の例を示すブロック図である。

【図15】図15は、第3の実施の形態におけるベースリストの一例を示す概念図である。

【図16】図16は、第3の実施の形態におけるベースイメージとコピーイメージの例を示す概念図である。

【図17】図17は、ディスクイメージの追加・変更の例を示す概念図である。

【図18】図18は、ログファイルの作成例を示す概念図である。

50

【図19】図19は、ログファイルからコピーリストを作成する例を示す概念図である。

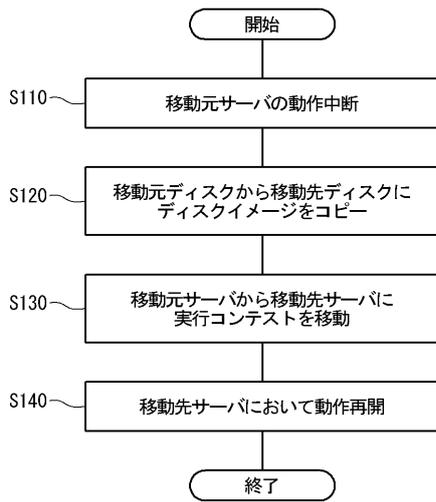
【図20】図20は、第3の実施の形態におけるベースイメージをコピーイメージの他の例を示す概念図である。

【符号の説明】

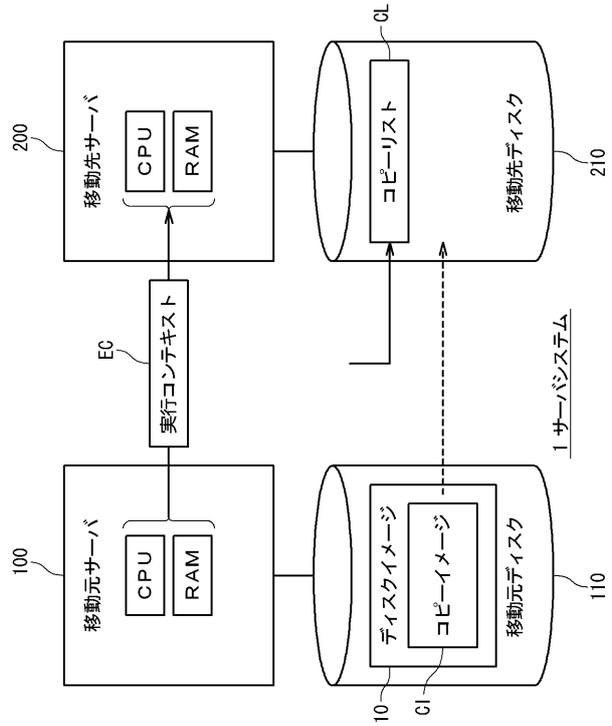
【0106】

1	サーバシステム	
10	ディスクイメージ	
11	第1のイメージ、ベースイメージ	
12	第2のイメージ、コピーイメージ	
20	コピーリスト作成モジュール	10
30	コンテキストコピーモジュール	
40	コピー判定モジュール	
50	コピーモジュール	
60	ROM	
70	ベースイメージ作成モジュール	
80	ベースイメージ配布モジュール	
100	移動元サーバ	
110	移動元ディスク	
120	移動元マシン	
130	移動元VMM	20
200	移動先サーバ	
210	移動先ディスク	
220	移動先マシン	
230	移動先VMM	
300	管理サーバ	
BI	ベースイメージ	
CI	コピーイメージ	
CL	コピーリスト	
CL - H	ヘッダ部	
CL - B	ボディ部	30
EC	実行コンテキスト	
GOS	ゲストOS	
HT	ハッシュテーブル	
LOG	ログファイル	

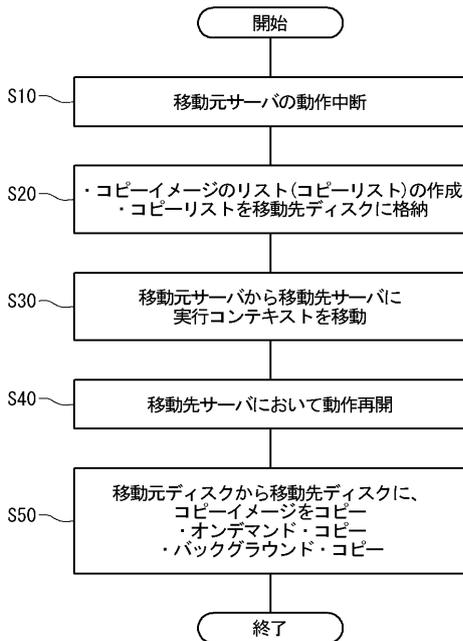
【 図 1 】



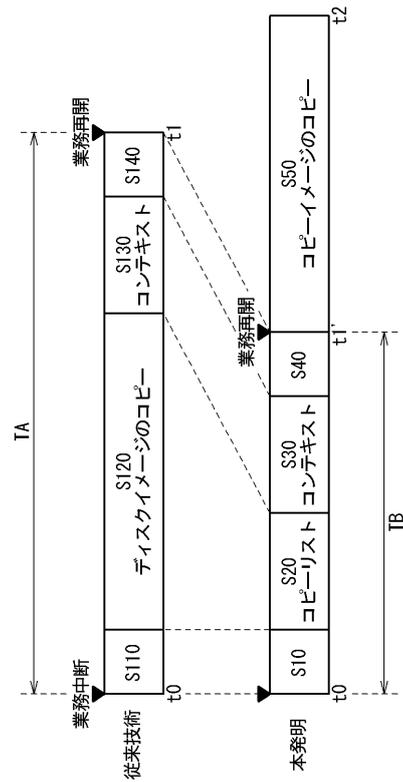
【 図 2 】



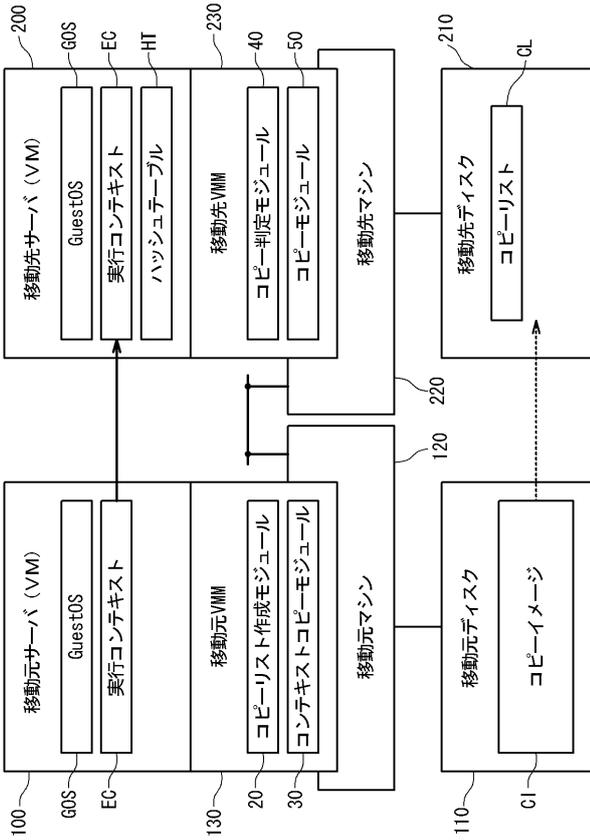
【 図 3 】



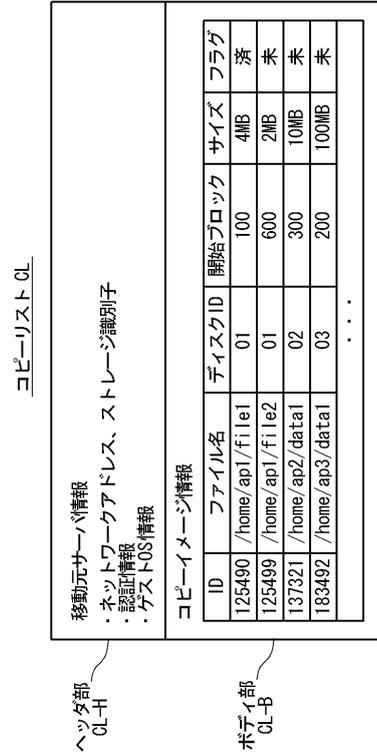
【 図 4 】



【 図 5 】



【 図 6 】

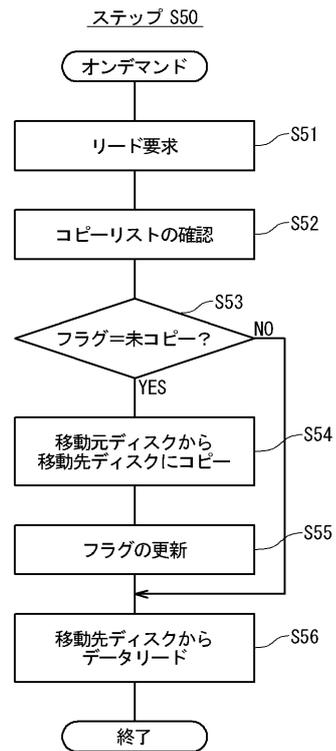


【 図 7 】

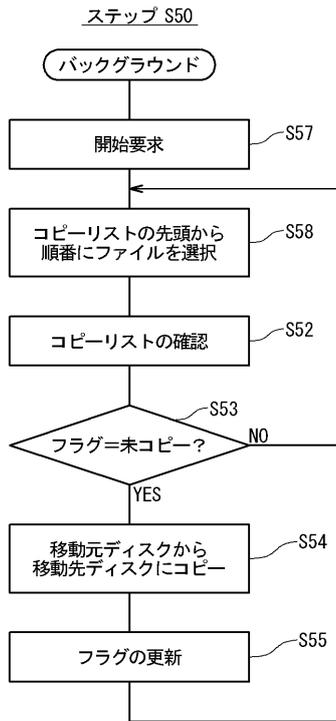
ハッシュテーブル HT

ハッシュキー	ID
..	..
321	137321
..	..
490	125490
..	..
492	183492, 254492, 383492
..	..
499	125499, 421499
..	..

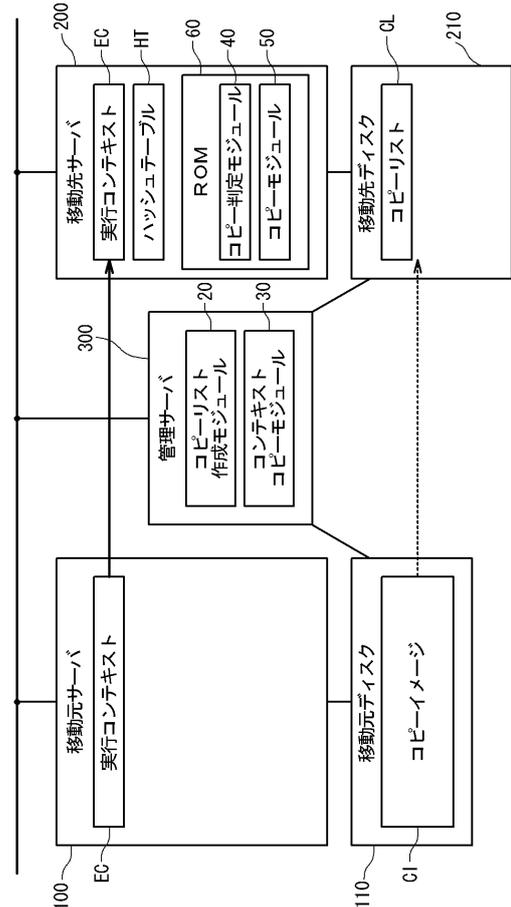
【 図 8 A 】



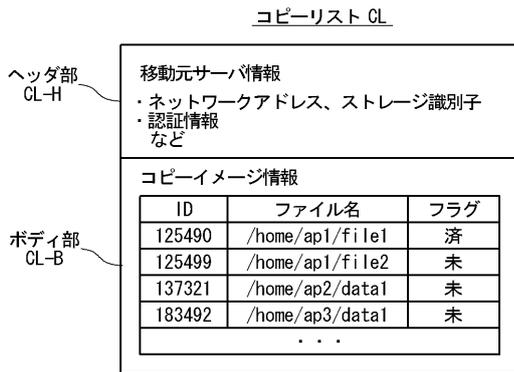
【 図 8 B 】



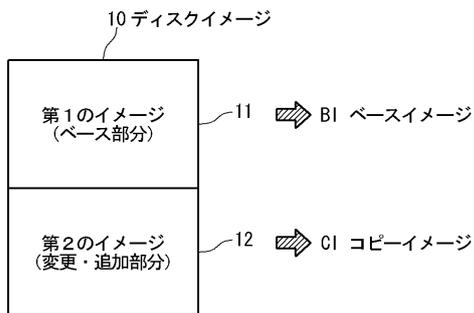
【 図 9 】



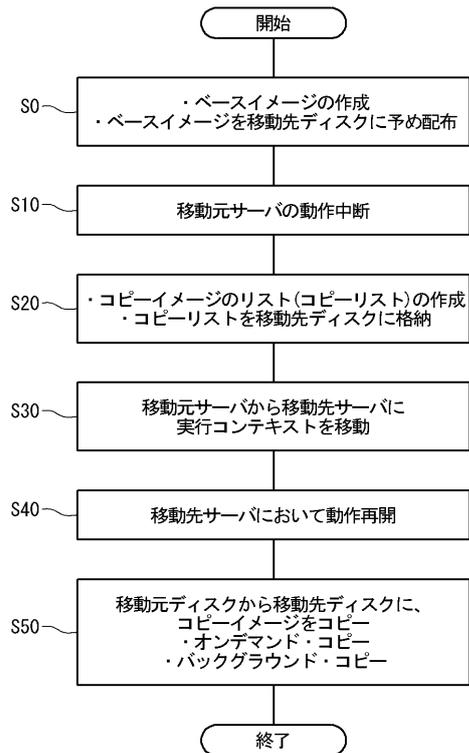
【 図 1 0 】



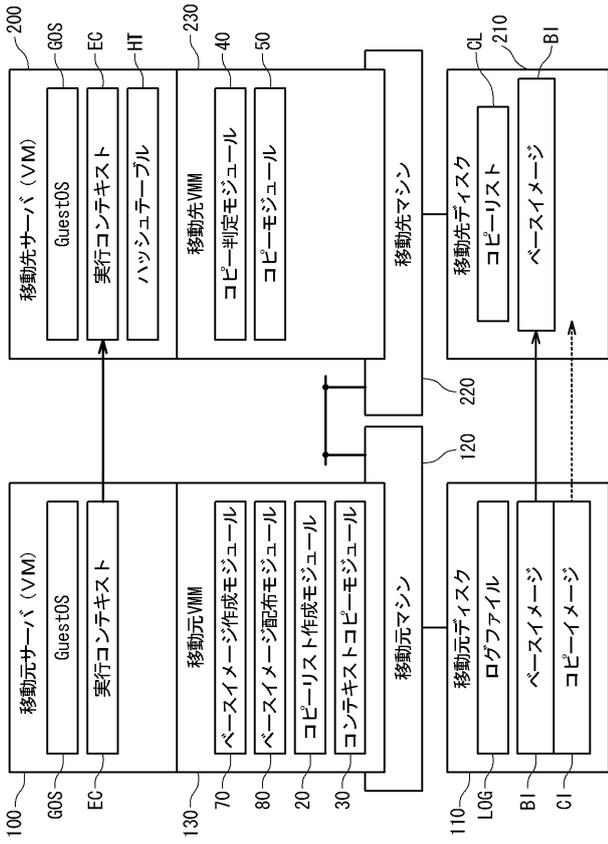
【 図 1 1 】



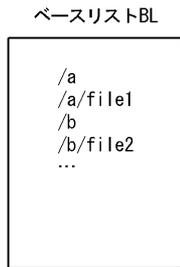
【 図 1 2 】



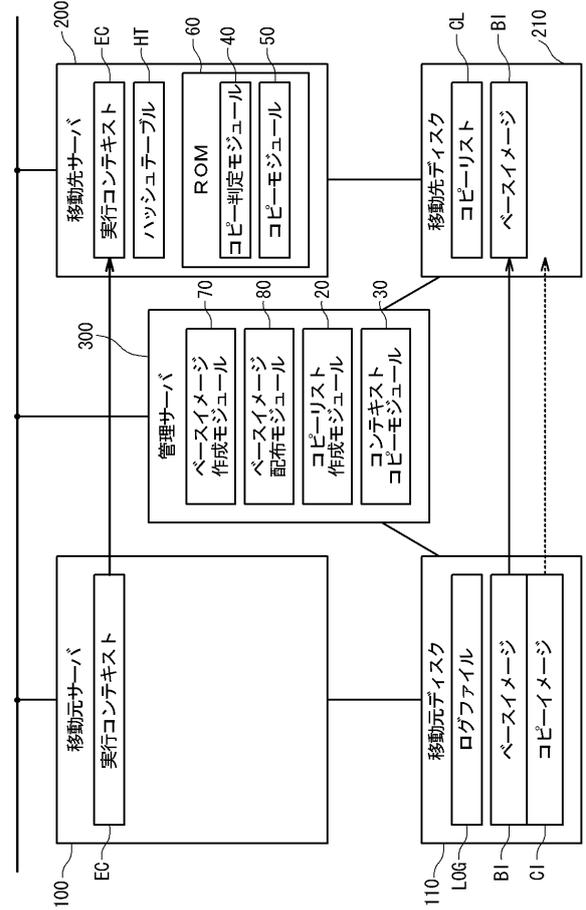
【 図 1 3 】



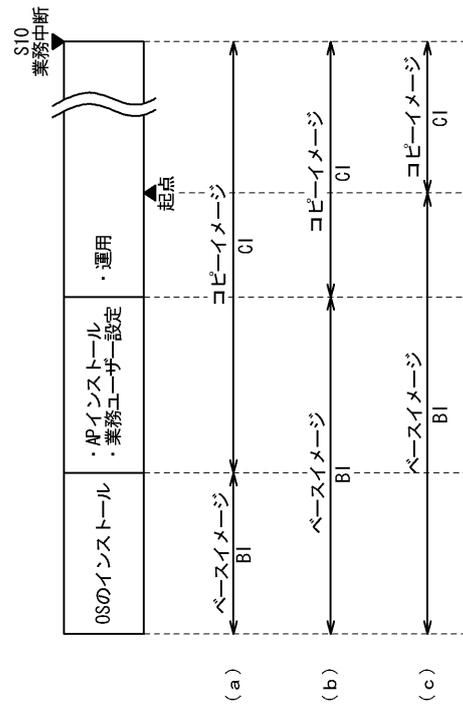
【 図 1 5 】



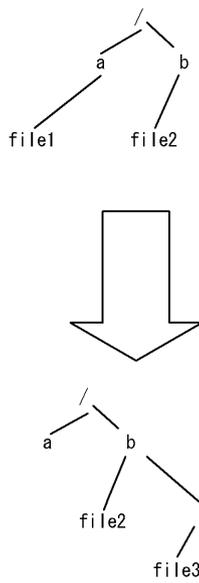
【 図 1 4 】



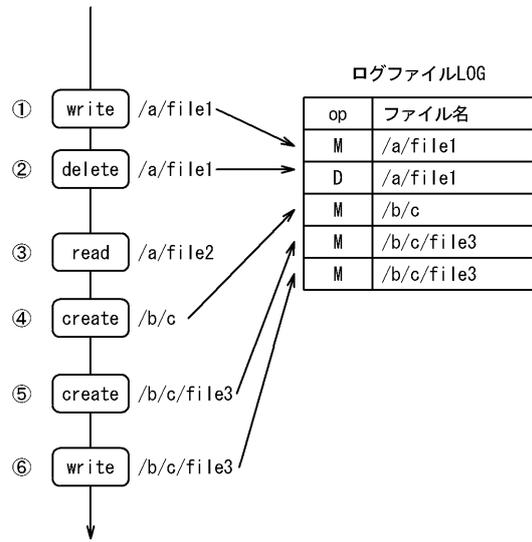
【 図 1 6 】



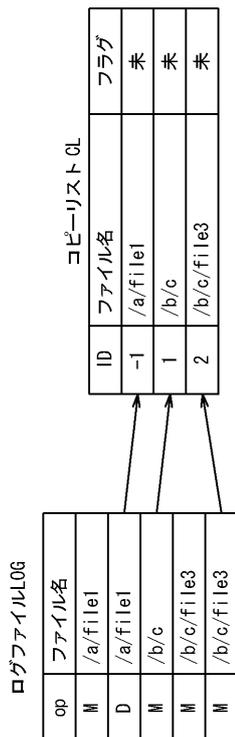
【 図 1 7 】



【 図 1 8 】



【 図 1 9 】



【 図 2 0 】

