

(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2010-33404

(P2010-33404A)

(43) 公開日 平成22年2月12日(2010.2.12)

(51) Int.Cl. F I テーマコード (参考)
G 0 6 F 9/46 (2006.01) G O 6 F 9/46 3 5 0 5 B 0 1 4
G 0 6 F 13/10 (2006.01) G O 6 F 13/10 3 4 0 A

審査請求 未請求 請求項の数 16 O L (全 25 頁)

(21) 出願番号	特願2008-195936 (P2008-195936)	(71) 出願人	000005108 株式会社日立製作所 東京都千代田区丸の内一丁目6番6号
(22) 出願日	平成20年7月30日(2008.7.30)	(74) 代理人	100080001 弁理士 筒井 大和
		(72) 発明者	高野 光弘 神奈川県秦野市堀山下1番地 株式会社日立製作所エンタープライズサーバ事業部内
		(72) 発明者	上野 仁 神奈川県秦野市堀山下1番地 株式会社日立製作所エンタープライズサーバ事業部内
		Fターム(参考)	5B014 EB04 FB05

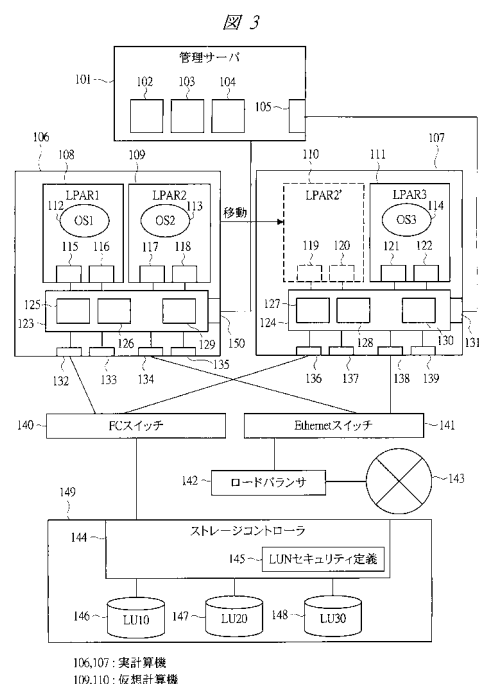
(54) 【発明の名称】 仮想計算機システムおよび仮想計算機システムの制御方法

(57) 【要約】

【課題】仮想計算機の移動を行う場合に、簡単な手順で、安全に移動処理を行うことができる仮想計算機システムの制御方法を提供する。

【解決手段】仮想計算機システムの制御方法において、実計算機(106)上で動作している仮想計算機(109)を別の実計算機(107)上に移動させる際、管理サーバ(101)により、移動元の仮想計算機(109)を停止させ、移動先の仮想計算機(110)を別の実計算機(107)上に定義し、その起動を抑止した状態で、移動元の仮想計算機(109)の仮想ネットワークアドレスおよび仮想ストレージインターフェイスアドレスを、移動先の仮想計算機(110)に定義された仮想ネットワークアドレスおよび仮想ストレージインターフェイスアドレスと交換した後、移動先の仮想計算機(110)の起動の抑止を解除する。

【選択図】図3



【特許請求の範囲】**【請求項 1】**

複数の実計算機と、システムディスクと、仮想計算機の制御を行う管理部とを備え、前記実計算機上の制御部により複数の論理的な仮想計算機を動作させ、前記仮想計算機のそれぞれは論理的な通信ネットワークインターフェイスおよび論理的なストレージ接続インターフェイスを有し、それぞれアドレス情報として、仮想ネットワークアドレスおよび仮想ストレージインターフェイスアドレスを有し、前記論理的なストレージ接続インターフェイスに前記仮想計算機を起動するシステムが格納された前記システムディスクが接続された仮想計算機システムの制御方法であって、

前記実計算機上で動作している前記仮想計算機を別の前記実計算機上に移動させる際、
前記管理部により、

移動元の前記仮想計算機を停止させ、

移動先の前記仮想計算機を別の前記実計算機上に定義し、

その起動を抑止した状態で、移動元の前記仮想計算機の構成情報及び移動先の前記仮想計算機の構成情報を変更し、

前記構成情報の変更後に、移動先の前記仮想計算機の起動を抑止を解除することを特徴とする仮想計算機システムの制御方法。

【請求項 2】

請求項 1 記載の仮想計算機システムの制御方法において、

前記構成情報は、仮想ネットワークアドレス、前記仮想ストレージインターフェイスアドレス及び時刻差分情報のうち、少なくとも 1 つであることを特徴とする仮想計算機システムの制御方法。

【請求項 3】

請求項 1 記載の仮想計算機システムの制御方法において、

移動元の前記仮想計算機の構成情報及び移動先の前記仮想計算機の構成情報を変更する際、

前記管理部により、移動元の前記仮想計算機の前記仮想ネットワークアドレスおよび前記仮想ストレージインターフェイスアドレスを、移動先の前記仮想計算機に上書きし、移動元の前記仮想計算機の前記仮想ネットワークアドレスおよび前記仮想ストレージインターフェイスアドレスを無効データとすることを特徴とする仮想計算機システムの制御方法。

【請求項 4】

請求項 1 記載の仮想計算機システムの制御方法において、

移動元の前記仮想計算機の構成情報及び移動先の前記仮想計算機の構成情報を変更する際、

前記管理部により、移動元の前記仮想計算機の前記仮想ネットワークアドレスおよび前記仮想ストレージインターフェイスアドレスを、移動先の前記仮想計算機に定義された前記仮想ネットワークアドレスおよび前記仮想ストレージインターフェイスアドレスと交換することを特徴とする仮想計算機システムの制御方法。

【請求項 5】

請求項 1 記載の仮想計算機システムの制御方法において、

前記管理部により、移動元の前記仮想計算機の起動を抑止することを特徴とする仮想計算機システムの制御方法。

【請求項 6】

請求項 1 記載の仮想計算機システムの制御方法において、

前記管理部により、同一ネットワークに接続されている前記実計算機に散在している前記仮想計算機を動作させている前記制御部のネットワークアドレスを探索することを特徴とする仮想計算機システムの制御方法。

【請求項 7】

請求項 1 記載の仮想計算機システムの制御方法において、

前記管理部により、同一ネットワークに接続されている前記実計算機に散在している前記仮想実計算機の仮想ネットワークアドレスを探索し、前記仮想ネットワークアドレスを持つ前記仮想実計算機が、現在どの前記実計算機上で動作しているかを探索することを特徴とする仮想計算機システムの制御方法。

【請求項 8】

請求項 1 記載の仮想計算機システムの制御方法において、

前記管理部は、前記複数の実計算機と接続され前記仮想計算機の管理を行う管理サーバに含まれることを特徴とする仮想計算機システムの制御方法。

【請求項 9】

複数の実計算機と、システムディスクと、仮想計算機の制御を行う管理部とを備え、前記実計算機上の制御部により複数の論理的な仮想計算機を動作させ、前記仮想計算機のそれぞれは論理的な通信ネットワークインターフェイスおよび論理的なストレージ接続インターフェイスを有し、それぞれアドレス情報として、仮想ネットワークアドレスおよび仮想ストレージインターフェイスアドレスを有し、前記論理的なストレージ接続インターフェイスに前記仮想計算機を起動するシステムが格納された前記システムディスクが接続された仮想計算機システムであって、

10

前記実計算機上で動作している前記仮想計算機を別の前記実計算機上に移動させる際、前記管理部は、

移動元の前記仮想計算機を停止させ、

移動先の前記仮想計算機を別の前記実計算機上に定義し、

20

その起動を抑止した状態で、移動元の前記仮想計算機の構成情報及び移動先の前記仮想計算機の構成情報を変更し、

前記構成情報の変更後に、移動先の前記仮想計算機の起動を抑止を解除することを特徴とする仮想計算機システム。

【請求項 10】

請求項 9 記載の仮想計算機システムにおいて、

前記構成情報は、仮想ネットワークアドレス、前記仮想ストレージインターフェイスアドレス及び時刻差分情報のうち、少なくとも 1 つであることを特徴とする仮想計算機システム。

【請求項 11】

30

請求項 9 記載の仮想計算機システムにおいて、

移動元の前記仮想計算機の構成情報及び移動先の前記仮想計算機の構成情報を変更する際、

前記管理部は、移動元の前記仮想計算機の前記仮想ネットワークアドレスおよび前記仮想ストレージインターフェイスアドレスを、移動先の前記仮想計算機に上書きし、移動元の前記仮想計算機の前記仮想ネットワークアドレスおよび前記仮想ストレージインターフェイスアドレスを無効データとすることを特徴とする仮想計算機システム。

【請求項 12】

請求項 9 記載の仮想計算機システムにおいて、

移動元の前記仮想計算機の構成情報及び移動先の前記仮想計算機の構成情報を変更する際、

40

前記管理部は、移動元の前記仮想計算機の前記仮想ネットワークアドレスおよび前記仮想ストレージインターフェイスアドレスを、移動先の前記仮想計算機に定義された前記仮想ネットワークアドレスおよび前記仮想ストレージインターフェイスアドレスと交換することを特徴とする仮想計算機システム。

【請求項 13】

請求項 9 記載の仮想計算機システムにおいて、

前記管理部は、移動元の前記仮想計算機の起動を抑止することを特徴とする仮想計算機システム。

【請求項 14】

50

請求項 9 記載の仮想計算機システムにおいて、

前記管理部は、同一ネットワークに接続されている前記実計算機に散在している前記仮想計算機を動作させている前記制御部のネットワークアドレスを探索することを特徴とする仮想計算機システム。

【請求項 15】

請求項 9 記載の仮想計算機システムにおいて、

前記管理部は、同一ネットワークに接続されている前記実計算機に散在している前記仮想実計算機の仮想ネットワークアドレスを探索し、前記仮想ネットワークアドレスを持つ前記仮想実計算機が、現在どの前記実計算機上で動作しているかを探索することを特徴とする仮想計算機システム。

10

【請求項 16】

請求項 9 記載の仮想計算機システムにおいて、

前記管理部は、前記複数の実計算機と接続され前記仮想計算機の管理を行う管理サーバに含まれることを特徴とする仮想計算機システム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、仮想計算機システムの制御方法に関し、特に、仮想計算機が持つ外部とのインターフェイスとしてネットワークインターフェイスとストレージインターフェイスを対象に、周囲の機器から透過的に扱い、仮想計算機群について情報を集約していない管理サーバを用いて、実計算機をまたいだ仮想計算機の移動を安全に行う制御方法に関する。

20

【背景技術】

【0002】

仮想計算機は実計算機の資源を有効に利用するための手段として実用化されている技術である。特に、利用率が低い複数の実計算機を仮想計算機として動作させる運用に変えることで資源を集約し、ストレージやネットワーク、電力などを節約する用途が注目されている。

【0003】

仮想計算機は、物理的な資源の集約や電力の節約など、得られる利点は多いが、近年では仮想計算機の管理が複雑化していることが問題となっている。多数の仮想計算機を扱う場合にこの傾向は顕著に現れる。

30

【0004】

ある実計算機上で動作していた 1 つの仮想計算機上のゲストシステムをいったん停止し、別の実計算機上の 1 つの仮想計算機上で動作させることを、仮想計算機を移動することであると定義すると、例えば、仮想計算機を実計算機をまたいで移動させるような運用や管理を行うときには、それぞれの仮想計算機がどのような状況にあり、仮想計算機を移動させるときには仮想計算機に接続されている機器にどのような影響があるのかを把握した管理が必要となり、管理サーバを含め、実計算機、仮想計算機の全てを考慮し、複雑なシステムとして扱うことになる。

【0005】

40

このようなシステムとして、ESX Server 3 i Configuration Guide (非特許文献 1)、Cisco MDS and Emulex Virtual HBA Solutions for VMware Infrastructure 3 (非特許文献 2) に記載された技術があった。

【非特許文献 1】ESX Server 3 i Configuration Guide、インターネット<URL:http://www.vmware.com/pdf/vi3_35/esx_3i_e/r35/vi3_35_25_3i_server_config.pdf>

【非特許文献 2】Cisco MDS and Emulex Virtual HBA Solutions for VMware Infrastructure 3、インターネット<URL:<http://www.emulex.com/white/hba/CiscoEmulexVirtualizationforVM>>

50

ware.pdf >

【発明の開示】

【発明が解決しようとする課題】

【0006】

しかしながら、非特許文献1、2に記載のように、実計算機をまたいで仮想計算機を移動させるときには、仮想計算機の周辺の環境に影響を及ぼさないように管理ソフトウェアがネットワークの制御やストレージの制御を行っているが、これらの管理ソフトウェアは仮想計算機を制御するソフトウェアとは別のものとして提供されており、それぞれの機器が仮想計算機の移動と干渉した不具合を起こさないように留意しつつ操作する必要がある、複雑な制御が必要である。

10

【0007】

例えば、ストレージについては移動元となる実計算機と移動先となる実計算機からのアクセスを許可しなければならない。また、仮想計算機が格納されているストレージが移動元と移動先の実計算機からのアクセスを行えなければ、実計算機をまたいだ仮想計算機の移動ができない。

【0008】

そのため、従来では、実計算機をまたいだ仮想計算機の移動を実現するためには、仮想計算機のストレージとして利用する媒体について、移動先となりうる複数の実計算機からのアクセスを許可する。この仕組みを利用することで、移動元となる実計算機と移動先となる実計算機から、同一のディスクにアクセスすることができ、実計算機をまたいだ仮想計算機の移動が可能となるが、これは複数の実計算機、さらには複数の仮想計算機から同時にアクセスすることを可能としていた。

20

【0009】

仮想計算機上で起動するOSを格納するような一般的なディスクでは、複数の計算機から共有された環境で利用することを前提にした仕組みを持たないため、複数の計算機から同時にアクセスしたディスクはファイルシステムの整合性が取れなくなり、破損してしまうという問題がある。

【0010】

また、ストレージコントローラにはこのような危険性を排除するために、ディスクと実計算機のストレージインターフェイスを対応付ける仕組みを持つものがあるが、この機能を利用したストレージは特定のストレージインターフェイスからのアクセスのみしか受け付けられないため、実計算機をまたいで仮想計算機を移動させるにはストレージコントローラの管理ソフトウェアを用いてストレージコントローラを制御することが必要となる。

30

【0011】

また、同時起動を防いだり、ネットワークやストレージの状況を把握するために集約的な管理サーバが必要となり、この集約的な管理サーバに障害が発生した場合に仮想計算機の運用についての情報が失われ、それぞれの仮想計算機がどの実計算機で稼働しているか、また、どのような状態であるかを把握することが困難である。

【0012】

例えば、非特許文献2に示す技術では、1台の仮想計算機を別の実計算機に移動させた場合に、仮想計算機が保持しているストレージ接続用のインターフェイスである仮想FC-HBAの持つ仮想WWNを変更する必要がない旨が記載されている。

40

【0013】

この方式を採用すればストレージコントローラ内のディスクと計算機のストレージインターフェイスを対応付ける仕組みに対して、仮想計算機を移動することによるWWNに関する設定変更をする必要がない。

【0014】

しかしながら、非特許文献2の技術によれば、仮想計算機を実計算機間で移動した際に、どの仮想WWNがどの実計算機上で使われているか、あるいは新たな仮想計算機に対して新たな仮想WWNを割り当てるとき、どのような値のWWNを割り当てれば重複が発生

50

しないかを管理する手段が必要となり、その管理手段が障害により失われると仮想 W W N の割り当て管理ができなくなるという課題がある。

【 0 0 1 5 】

ここで、図 1 0 および図 1 1 により、従来の仮想計算機システムが複雑な制御による具体的な問題点について説明する。図 1 0 および図 1 1 は従来の仮想計算機システムの制御における具体的な問題点を説明するための説明図であり、図 1 0 は実計算機を用いてスケールイン・スケールアウトを行う場合、図 1 1 は仮想サーバ技術で仮想計算機を移動する場合を示している。

【 0 0 1 6 】

図 1 0 において、システムは管理サーバ (3 0 1)、3 台の実計算機 (3 0 2 , 3 0 3 , 3 0 4)、F C スイッチ (3 3 1)、E t h e r n e t (登録商標) スイッチ (3 3 2)、ロードバランサ (3 3 3)、ストレージ装置 (3 4 1)、そしてロードバランサの外にある外部ネットワーク (3 3 5) からなる。実計算機は物理 F C H B A (3 1 1 , 3 1 2 , 3 1 5 , 3 1 6 , 3 1 9 , 3 2 0) と物理 N I C (3 1 3 , 3 1 4 , 3 1 7 , 3 1 8 , 3 2 1 , 3 2 2) を持つほか、管理用の N I C (3 0 8 , 3 0 9 , 3 1 0) も備えている。

10

【 0 0 1 7 】

また、L U 1 0 (3 4 4) に格納された業務システム 1 (3 2 1) と L U 3 0 (3 4 6) に格納された業務システム 3 (3 2 3) が同じ機能を持ち、ロードバランサを用いてこのふたつのシステムをロードバランサルール A により、負荷分散している。

20

【 0 0 1 8 】

L U 2 0 (3 4 5) に格納された業務システム 2 (3 2 2) と L U 4 0 (3 4 7) に格納された業務システム 4 (3 2 3) は同じ機能を持つシステムだが、業務システム 2 のみが起動されているとする。

【 0 0 1 9 】

このとき、業務システム 3 の処理能力が不要となり、業務システム 4 (3 2 4) を起動し、業務システム 2 との間でロードバランサルール B により負荷を分散する構成に変更する必要がある場合には以下のような作業が必要となる。

【 0 0 2 0 】

(1) 管理サーバ (3 0 1) が N I C (3 0 7) を用いて、実計算機 3 の N I C (3 1 0) を経由して実計算機 3 で動作する業務システム 3 に停止を指示する。

30

【 0 0 2 1 】

(2) 管理サーバ (3 0 1) が N I C を用いて、ストレージコントローラ (3 4 2) へ N I C (3 0 5) を経由して、L U N セキュリティ定義 (3 4 3) を変更するよう、指示する。

【 0 0 2 2 】

* L U 3 0 W W N 3 0

↓ 変更

* L U 4 0 W W N 3 0

なお、ストレージコントローラに指示を与えるため、管理サーバ (3 0 1) にはストレージコントローラの管理ソフトウェア (3 2 5) が必要である。

40

【 0 0 2 3 】

(3) 管理サーバ (3 0 1) が N I C を用いて、ロードバランサへ N I C (3 0 6) を経由してロードバランサ内で管理されているルールテーブル (3 3 4) に格納されているロードバランサルール A から業務システム 3 の登録を解除する。

【 0 0 2 4 】

* 業務システム 1 業務システム 3

↓ 変更

* 業務システム 1

なお、ロードバランサに指示を与えるため、管理サーバ (3 0 1) にはロードバランサ

50

の管理ソフトウェア（３２６）が必要である。

【００２５】

（４）管理サーバ（３０１）がNICを用いて、ロードバランサへNICを経由してロードバランサルールＢに業務システム４の登録を追加する。

【００２６】

* 業務システム 2

↓ 変更

* 業務システム 2 業務システム 4

上記と同様にこの操作ではロードバランサに指示を与えるため、管理サーバ（３０１）にはロードバランサに対応した管理ソフトウェアのインストールが必要である。

10

【００２７】

（５）管理サーバ（３０１）がNICを用いて、実計算機３のNICを経由して実計算機３の起動を指示する。

【００２８】

以上のような手順が発生し、実計算機の管理のみでは制御することができず、ストレージコントローラやロードバランサの管理が必要となるという問題がある。

【００２９】

また、図１１において、システムは管理サーバ（４０１）、２台の実計算機（４０６、４０７）、FCスイッチ（４４０）、Ethernet（登録商標）スイッチ（４４１）、ロードバランサ（４４２）、ストレージ（４４７）、外部ネットワーク（４４３）で構成されている。実計算機ではL P A R 1（４０８）、L P A R 2（４０９）、L P A R 3（４１１）の仮想計算機が動作し、その上でOS 1（４１２）、OS 2（４１３）、OS 3（４１４）のシステムが動作している。

20

【００３０】

仮想計算機を移動させるときには、移動元と移動先で同一の情報を用いるなどのことがないよう、管理サーバ（４０１）が仮想計算機についての構成情報を集約的に管理している。L P A R 2を移動元実計算機（４０６）から移動先実計算機（４０７）へ移動させるときには管理サーバ（４０１）が管理している情報を用い、以下のような処理を行う。

【００３１】

（１）管理サーバ（４０１）がNIC（４０５）を用いて、移動元の仮想計算機のNIC（４６０）を経由し、移動元の仮想計算機のハイパバイザエージェント（４２９）へ移動元仮想計算機の停止を要求する。

30

【００３２】

（２）管理サーバ（４０１）がNIC（４０５）を用いて、移動先の仮想計算機のNIC（４３１）を経由し、移動先の仮想計算機のハイパバイザエージェント（４３０）へ移動先仮想計算機（４１０）の起動を要求する。

【００３３】

管理サーバ（４０１）ではこれらの段階について、移動元の仮想計算機と移動先の仮想計算機で重複する情報を保持しないかをCPU（４０２）、主記憶（４０３）、ストレージ（４０４）を用いて管理する。

40

【００３４】

具体的には物理FC H B A（４３２，４３３，４３６，４３７）を仮想化した仮想FC H B A（４１５，４１７，４１９，４２１）や物理NIC（４３４，４３５，４３８，４３９）を仮想化した仮想NIC（４１６，４１８，４２０，４２２）の情報が重複していないかをハイパバイザ（４２３，４２４）が保持する仮想FC H B A管理テーブル（４２５，４２７）や仮想NIC管理テーブル（４２６，４２８）を参照する。

【００３５】

なお、従来のシステムではWWNやMACアドレスの仮想化は行わないことが一般的で、物理的なWWNやMACアドレスが仮想計算機によって利用される。

【００３６】

50

また、仮想サーバのシステムでは特殊なファイルシステムにより、１つのＬＵに複数のＯＳが格納されることがある。このシステムではＯＳ１の区画（４５０）とＯＳ２の区画（４５１）がＬＵ２０（４４８）に格納されている。

【００３７】

また、ＯＳ３の区画（４５２）はＬＵ３０（４４９）に格納されている。このシステムにおいて、以下のような短所がある。

【００３８】

（１）実計算機がハイパバイザを用いずにＬＵにアクセスしたとき、予期しないファイルシステムにアクセスしてしまい、ディスクを破損するおそれがある。

【００３９】

（２）管理サーバ（４０１）が集約的に仮想計算機についての構成情報を保持しているため、他の計算機では代替することができない。

【００４０】

（３）ＷＷＮの仮想化に対応していないため、管理者はストレージコントローラ（４４４）のＬＵＮセキュリティ定義（４４５）について以下のように定義し、ＬＵ２０を複数のＷＷＮに開放する。複数のサーバからアクセス可能であるため、サーバが同時に動作したときにディスク内容を破損してしまうおそれがある。

【００４１】

* ＬＵ１０（４４６） ＷＷＮ１０

* ＬＵ２０（４６１） ＷＷＮ２０

* ＬＵ２０（４６１） ＷＷＮ３０

* ＬＵ３０（４６２） ＷＷＮ４０

そこで、本発明の目的は、仮想計算機の移動を行う場合に、簡単な手順で、安全に移動処理を行うことができる仮想計算機システムの制御方法を提供することにある。

【００４２】

本発明の前記ならびにその他の目的と新規な特徴は、本明細書の記述および添付図面から明らかになるであろう。

【課題を解決するための手段】

【００４３】

本願において開示される発明のうち、代表的なものの概要を簡単に説明すれば、次のとおりである。

【００４４】

すなわち、代表的なものの概要は、実計算機上で動作している仮想計算機を別の実計算機上に移動させる際、管理部により、移動元の仮想計算機を停止させ、移動先の仮想計算機を別の実計算機上に定義し、その起動を抑止した状態で、移動元の仮想計算機の構成情報及び移動先の仮想計算機の構成情報を変更し、構成情報の変更後に、移動先の仮想計算機の起動の抑止を解除するものである。

【発明の効果】

【００４５】

本願において開示される発明のうち、代表的なものによって得られる効果を簡単に説明すれば以下のとおりである。

【００４６】

すなわち、代表的なものによって得られる効果は、仮想計算機のＭＡＣアドレスと仮想ＷＷＮが移動の前後で変化しないため、ルータやロードバランサなどのネットワーク機器やＳＡＮストレージ装置の設定を変更する必要がなく、仮想計算機の移動のための制御手順を簡単にできる。

【００４７】

また、仮想計算機１台とストレージ装置内のＬＵを１対１に対応付けたまま移動ができるので、誤って別のサーバがそのＬＵを重複してアクセスすることによるディスクデータの破壊が発生することがない、という信頼性上の効果がある。

10

20

30

40

50

【発明を実施するための最良の形態】**【0048】**

まず、本発明の概要について説明する。

【0049】

本発明では、仮想計算機を移動するときにネットワークとストレージについて、別の管理ソフトウェアに影響を及ぼさないために、移動元の仮想計算機と移動先の仮想計算機のネットワークインターフェイスとストレージインターフェイスが仮想計算機の周辺から同じ見え方にする。

【0050】

このようにネットワークインターフェイスとストレージインターフェイスを透過的に扱うことで、仮想計算機が抽象化され、仮想計算機がどのような実計算機で動作していてもネットワークやストレージの管理ソフトウェアに影響することなく仮想計算機を運用することができる。

【0051】

また、仮想計算機の周辺から透過的なネットワークインターフェイスとストレージインターフェイスを構成するために、仮想計算機の移動の際には、移動元仮想計算機が持つ仮想MACアドレスと仮想WWNを、移動先仮想計算機が持つ仮想MACアドレスと仮想WWNを交換して設定する手段を設ける。

【0052】

物理的なネットワークインターフェイスの識別子であるMACアドレスと物理的なストレージインターフェイスの識別子であるWWNについて仮想化を行うことで、管理サーバは仮想計算機の移動において、移動元の仮想計算機と移動先の計算機で仮想MACアドレスと仮想WWNを交換し、仮想計算機がどのような実計算機で動作しているものであっても、ネットワークの管理ソフトウェアやストレージの管理ソフトウェアでは同じ仮想計算機として扱う。

【0053】

これにより、ネットワークの管理ソフトウェアやストレージの管理ソフトウェアでは仮想計算機の移動にともなう操作が不要となる。しかし、仮想MACアドレスと仮想WWNを交換する手段を提供したときに、移動元の仮想計算機と移動先の仮想計算機で仮想MACアドレスおよび仮想WWNが重複する可能性がある。例えば、これらが重複した複数の仮想計算機が同時に起動すると、ネットワークやストレージに対して、物理的な識別子が重複してしまうため、予期しないネットワークアドレスが付与されることや複数の計算機からの同時のアクセスによりストレージが破壊される可能性がある。

【0054】

本発明ではインターフェイスの識別子が重複した複数の仮想計算機が同時に起動することを防ぐため、仮想計算機に起動を抑止する機能を用いる。また、仮想計算機の移動作業途中でハードウェア障害やソフトウェア障害などが発生したり、移動作業の管理が不十分であったりするとどの仮想計算機がどのような実計算機で動作中なのかが分からなくなる場合があるので、このような仮想計算機の配置情報を分析して表示する手段を設ける。

【0055】

以下、本発明の実施の形態を図面に基づいて詳細に説明する。なお、実施の形態を説明するための全図において、同一の部材には原則として同一の符号を付し、その繰り返しの説明は省略する。

【0056】

図1により、本発明の一実施の形態に係る制御方法が適用された仮想計算機システムの構成について説明する。図1は本発明の一実施の形態に係る制御方法が適用された仮想計算機システムの構成を示す構成図である。

【0057】

図1において、仮想計算機システムは、CPU(102)、主記憶(103)、ストレージ(104)、ネットワークインターフェイス(105)を持つ管理部である管理サー

10

20

30

40

50

バ(101)、2つの実計算機(106, 107)、FCスイッチ(140)、Ethernet(登録商標)スイッチ(141)、ロードバランサ(142)、ストレージ装置(149)から構成され、外部ネットワーク(143)に通じている。

【0058】

このとき、実計算機(106, 107)で、制御部となるハイパバイザ(123, 124)が動作し、ハイパバイザ(123, 124)はハイパバイザエージェント(129, 130)からの指示で仮想計算機を管理する。

【0059】

ハイパバイザエージェント(129, 130)はネットワークインターフェイス(150, 131)を用い、管理サーバ(101)から指示を受けることでハイパバイザ(123, 124)の挙動を変化させることができる。

10

【0060】

ハイパバイザ(123, 124)は、物理FC HBA(132, 133, 136, 137)を仮想計算機(108, 109, 111)で共有して利用するために、物理FC HBAと仮想FC HBA(115, 117, 121)を関連付けるための仮想FC HBA管理テーブル(125, 127)をもっている。

【0061】

同様に物理NIC(134, 135, 138, 139)を仮想計算機で共有して利用するために、物理NICと仮想NIC(116, 118, 122)を関連付けるための仮想NIC管理テーブル(126, 128)をもっている。

20

【0062】

このとき、仮想計算機(108)ではLU10(146)に格納されているOS1(112)が動作し、仮想計算機(109)ではLU20(147)に格納されているOS2(113)が動作し、仮想計算機(111)ではLU30(148)に格納されているOS3(114)がそれぞれ並列に動作している。

【0063】

なお、ストレージコントローラ(144)ではLUNセキュリティ定義(145)を操作し、仮想FC HBA(115)からはLU10(146)へのアクセスのみを許可し、仮想FC HBA(117)からはLU20(147)へのアクセスのみを許可し、仮想FC HBA(121)からはLU30(148)へのアクセスのみを許可している。

30

【0064】

次に、図2～図4により、本発明の一実施の形態に係る制御方法による仮想計算機の移動動作について説明する。図2～図4は本発明の一実施の形態に係る制御方法による仮想計算機の移動動作を説明するための説明図であり、仮想計算機の同時起動を防止し、安全に移動させるために管理サーバ(101)からハイパバイザ(123, 124)へ定められた順で指示を出すことでハイパバイザ(123, 124)が処理を行う際の処理時の仮想計算機システムの状態を示している。

【0065】

(1)まず、図2に示すように管理サーバ(101)が移動先の仮想計算機(110)を定義する。この仮想計算機(110)は他の仮想計算機(108, 109, 111)と同様に仮想FC HBA(119)と仮想NIC(120)をもつ。

40

【0066】

(2)移動先の実計算機(107)で動作するハイパバイザエージェント(130)が管理サーバ(101)からの指示により、定義した移動先の仮想計算機(110)の起動を抑止する。起動の抑止をされた仮想計算機は停止した状態から起動することができない。

【0067】

(3)移動元の実計算機(106)で動作するハイパバイザエージェント(129)が管理サーバ(101)からの指示により、移動元の仮想計算機(109)のOS2を停止する。OS2が停止されたことにより、LU20(147)へのディスクアクセスがなく

50

なる。

【 0 0 6 8 】

(4) 管理サーバ (1 0 1) は移動元の実計算機 (1 0 6) で動作するハイパバイザエージェント (1 2 9) に仮想計算機の状態を問い合わせ、移動元仮想計算機 (1 0 9) が停止していることを確認する。

【 0 0 6 9 】

(5) 移動元の実計算機 (1 0 6) で動作しているハイパバイザエージェント (1 2 9) が管理サーバ (1 0 1) からの指示により移動元の仮想計算機 (1 0 9) の起動を抑止する。

【 0 0 7 0 】

(6) 管理サーバ (1 0 1) が移動元の仮想計算機 (1 0 9) と移動先の仮想計算機 (1 1 0) の仮想 W W N を交換する。交換は以下の (a) ~ (d) の手順で行われる。

【 0 0 7 1 】

(a) 管理サーバ (1 0 1) が移動元の実計算機 (1 0 6) で動作するハイパバイザエージェント (1 2 9) から移動元の仮想計算機 (1 0 9) の仮想 F C H B A (1 1 7) に設定されている仮想 W W N を取得する。取得した仮想 W W N は管理サーバ (1 0 1) の主記憶 (1 0 3) に記憶する。

【 0 0 7 2 】

(b) 管理サーバ (1 0 1) が移動先の実計算機 (1 0 7) で動作するハイパバイザエージェント (1 3 0) から移動先の仮想計算機 (1 1 0) の仮想 F C H B A (1 1 9) に設定されている仮想 W W N を取得する。取得した仮想 W W N は管理サーバ (1 0 1) の主記憶 (1 0 3) に記憶する。

【 0 0 7 3 】

(c) 管理サーバ (1 0 1) が移動先の実計算機 (1 0 7) で動作するハイパバイザエージェント (1 3 0) に指示を出し、移動先の仮想計算機 (1 1 0) の仮想 F C H B A (1 1 9) に取得した移動元の仮想計算機 (1 0 9) の仮想 W W N の設定を要求する。管理サーバ (1 0 1) は応答で設定の完了を確認する。

【 0 0 7 4 】

(d) 管理サーバ (1 0 1) が移動元の実計算機 (1 0 6) で動作するハイパバイザエージェント (1 2 9) に指示を出し、移動元の仮想計算機 (1 0 9) の仮想 F C H B A (1 1 7) に取得した移動先の仮想計算機 (1 1 0) の仮想 W W N の設定を要求する。管理サーバ (1 0 1) は応答で設定の完了を確認する。

【 0 0 7 5 】

(7) 移動元の仮想計算機 (1 0 9) と移動先の仮想計算機 (1 1 0) の仮想 N I C (1 1 8 , 1 2 0) について仮想 M A C アドレスの交換を行う。この交換の手順は上記 (a) ~ (d) の手順と同様とする。

【 0 0 7 6 】

(8) 移動元の仮想計算機 (1 0 9) と移動先の仮想計算機 (1 1 0) の時刻差分情報を交換する。交換の手順は上記 (a) ~ (d) の手順と同様とする。なお、時刻差分情報はハイパバイザ (1 2 3 , 1 2 4) が管理している仮想計算機の情報に含まれている。

【 0 0 7 7 】

なお、この時点の状態を示したものが図 3 である。

【 0 0 7 8 】

(9) 移動先の実計算機 (1 0 7) で動作するハイパバイザエージェント (1 3 0) が管理サーバ (1 0 1) からの指示により、移動先の仮想計算機 (1 1 0) の起動の抑止を解除する。

【 0 0 7 9 】

(1 0) 移動先の実計算機 (1 0 7) で動作するハイパバイザエージェント (1 3 0) が管理サーバ (1 0 1) からの指示により、移動先の仮想計算機 (1 1 0) を起動する。

【 0 0 8 0 】

10

20

30

40

50

これにより、図 4 に示すような状態となり、移動先の仮想計算機（１１０）で OS 2 が起動する。

【００８１】

以上の手順で安全に仮想計算機を実計算機をまたいで移動させることができる。ただし、移動元の仮想計算機（１０９）と移動先の仮想計算機（１１０）の間での構成情報（仮想 W W N , 仮想 M A C アドレス, 時刻差分情報）の交換についてはそれぞれ順不同で行うことができる。

【００８２】

また、同じ移動の手順を同一の組の仮想計算機の間で行うときには、移動元であった仮想計算機が移動の完了後に定義済みの仮想計算機として残るため、移動先の仮想計算機を定義する手順については不要である。

10

【００８３】

次に、図 5 により、本発明の一実施の形態に係る制御方法による仮想計算機の移動シーケンスについて説明する。図 5 は本発明の一実施の形態に係る制御方法による仮想計算機の移動シーケンスを示す図であり、移動元の実計算機（１０６）で動作するハイパバイザエージェント（１２９）、管理サーバ（１０１）、移動先の実計算機（１０７）で動作するハイパバイザエージェント（１３０）に注目して表している。図 5 に示した制御手順を L P A R 移動シーケンスとする。

【００８４】

（１）管理サーバ（１０１）が、移動元計算機で動作するハイパバイザエージェント（１２９）に起動抑止処理への対応を確認する（ステップ 2 0 1）。これは、後ほど移動元の仮想計算機（１０９）と移動先の仮想計算機（１１０）との間で仮想 W W N や仮想 M A C アドレスなどの構成情報を交換するときに双方の仮想計算機（１０９, １１０）が起動抑止処理を行わなければ、安全な状態で構成情報を交換できないためである。

20

【００８５】

（２）管理サーバ（１０１）が、移動先計算機で動作するハイパバイザエージェント（１３０）に移動先の仮想計算機（１１０）について自動起動の無効を確認する（ステップ 2 0 2）。自動起動が有効であると移動先の実計算機（１０７）で予期しない再起動が発生したときに、仮想計算機（１１０）が予期しない構成情報を保持して起動する可能性があり、安全ではないためである。

30

【００８６】

（３）管理サーバ（１０１）が、移動元の実計算機（１０６）で動作するハイパバイザエージェント（１２９）に移動元の仮想計算機（１０９）の起動状態を確認する（ステップ 2 0 3）。管理サーバ（１０１）は起動状態の確認によって、移動元の仮想計算機（１０９）を停止する必要があるかを判断する。

【００８７】

（４）管理サーバ（１０１）が、移動先の実計算機（１０６）で動作するハイパバイザエージェント（１２９）に移動先の仮想計算機（１１０）に起動抑止処理を行うように指示する（ステップ 2 0 4）。移動先の仮想計算機（１１０）はこの起動抑止処理によって、起動を抑止され、構成情報を交換する間に起動することを防ぐことができ、安全に構成情報を交換することができる。

40

【００８８】

（５）管理サーバ（１０１）は、ステップ 2 0 3 において、移動元の実計算機（１０６）で動作するハイパバイザエージェント（１２９）に移動元の仮想計算機（１０９）の起動状態を確認したときに、起動されている状態であった場合に、移動元計算機で動作するハイパバイザエージェントに移動元仮想計算機の停止処理を要求する（ステップ 2 0 5）。仮想計算機が起動している状態では安全に仮想計算機を移動させることができないためである。

【００８９】

（６）管理サーバ（１０１）は、移動元の実計算機（１０６）で動作するハイパバイザ

50

エージェント(129)はステップ205による停止処理の要求を受け付け、正常に処理されたことを確認することで、仮想計算機の停止を確認する(ステップ206)。

【0090】

(7)管理サーバ(101)は、移動元の実計算機(106)で動作するハイパバイザエージェント(129)に対し、移動元の仮想計算機(109)の起動抑止処理を行う(ステップ207)。この段階で移動元の仮想計算機(109)の起動を抑止することで、移動元の仮想計算機(109)と移動先の仮想計算機(110)が起動抑止状態となり、構成情報の交換をしている間も双方の仮想計算機(109, 110)が起動することなく、安全に構成情報を交換できる。

【0091】

(8)管理サーバ(101)は、移動元計算機で動作するハイパバイザエージェントへ移動元仮想計算機のFC情報の取得を要求する(ステップ208)。この要求で仮想計算機がどの物理FC HBAを利用しているのか、また仮想FC HBAに設定されている仮想WWNを取得する。

【0092】

(9)管理サーバ(101)は、移動先の実計算機(107)で動作するハイパバイザエージェント(130)へ移動先の仮想計算機(110)のFC情報の取得を要求する(ステップ209)。

【0093】

(10)管理サーバ(101)は、移動先の実計算機(107)で動作するハイパバイザエージェント(130)に移動先の仮想計算機(110)の仮想WWNの設定を要求する(ステップ210)。設定を要求する仮想WWNの値はステップ208でのFC情報の取得による情報に含まれるものである。

【0094】

(11)管理サーバ(101)は、移動元の実計算機(106)で動作するハイパバイザエージェント(129)に移動元の仮想計算機(109)の仮想WWNの設定を要求する(ステップ211)。設定を要求する仮想WWNの値は、ステップ209でのFC情報の取得による情報に含まれるものである。

【0095】

(12)管理サーバ(101)は、移動元の実計算機(106)で動作するハイパバイザエージェント(129)へ移動元の仮想計算機(109)の仮想MACアドレスの取得を要求する(ステップ212)。

【0096】

(13)管理サーバ(101)は、移動先の実計算機(107)で動作するハイパバイザエージェント(130)へ移動先の仮想計算機(110)の仮想MACアドレスの取得を要求する(ステップ213)。

【0097】

(14)管理サーバ(101)は、移動先の実計算機(107)で動作するハイパバイザエージェント(130)へ移動先の仮想計算機(110)の仮想MACアドレスの設定を要求する(ステップ214)。設定を要求する仮想MACアドレスの値は、ステップ212での移動元の実計算機(106)で動作するハイパバイザエージェント(129)への仮想MACアドレスの取得に含まれるものである。

【0098】

(15)管理サーバ(101)は、移動元の実計算機(106)で動作するハイパバイザエージェント(129)へ移動元の仮想計算機(109)の仮想MACアドレスの設定を要求する(ステップ215)。設定を要求する仮想MACアドレスの値は、ステップ213での移動先の実計算機(107)で動作するハイパバイザエージェント(130)への仮想MACアドレスの取得に含まれるものである。

【0099】

(16)管理サーバ(101)は、移動元の実計算機(106)で動作するハイパバイ

10

20

30

40

50

ザエージェント(129)へ移動元の仮想計算機(109)についての時刻差分情報の取得を要求する(ステップ216)。この時刻差分情報には移動元の仮想計算機(109)と移動元の実計算機(106)のハイパバイザ(123)の間の時刻差分情報が含まれている。

【0100】

(17)管理サーバ(101)は、移動先の実計算機(107)で動作するハイパバイザエージェント(130)へ移動先の仮想計算機(110)についての時刻差分情報の取得を要求する(ステップ217)。この時刻差分情報には移動先の仮想計算機(110)と移動先の実計算機(107)のハイパバイザ(124)の間の時刻差分情報が含まれている。

10

【0101】

(18)管理サーバ(101)は、移動先の実計算機(107)で動作するハイパバイザエージェント(130)へ移動先の仮想計算機(110)の時刻差分情報の設定を要求する(ステップ218)。設定を要求する時刻差分情報は、ステップ216での移動元の実計算機(106)で動作するハイパバイザエージェント(129)へ要求した時刻差分情報の取得に含まれているものである。

【0102】

(19)管理サーバ(101)は、移動元の実計算機(106)で動作するハイパバイザエージェント(129)へ移動元の仮想計算機(109)の時刻差分情報の設定を要求する(ステップ219)。設定を要求する時刻差分情報は、ステップ217での移動先の実計算機(107)で動作するハイパバイザエージェント(130)へ要求した時刻差分情報の取得に含まれているものである。

20

【0103】

(20)管理サーバ(101)は、移動元の実計算機(106)で動作するハイパバイザエージェント(129)へ構成情報の保存を要求する(ステップ220)。構成情報を保存することで、以後、移動元の実計算機(106)が再起動しても、仮想計算機(109)が移動された状態で起動する。

【0104】

(21)管理サーバ(101)は、移動先の実計算機(107)で動作するハイパバイザエージェント(130)へ構成情報の保存を要求する(ステップ221)。構成情報を保存することで、以後、移動先の実計算機(107)が再起動しても、仮想計算機(110)が移動された状態で起動する。

30

【0105】

(22)管理サーバ(101)は、移動先の実計算機(107)で動作するハイパバイザエージェント(130)へ移動先の仮想計算機(110)の起動抑止解除を要求する(ステップ222)。これにより、LPAR移動フローの開始直後に行ったステップ204での起動抑止処理が解除され、移動先の仮想計算機(110)が起動できる状態となる。

【0106】

(23)管理サーバ(101)は、移動先の実計算機(107)で動作するハイパバイザエージェント(130)へ移動先の仮想計算機(110)の起動処理を要求する(ステップ223)。起動処理の要求を受け取った仮想計算機(110)は移動元の仮想計算機(109)がアクセスしていたディスクから起動する。

40

【0107】

(24)管理サーバ(101)は、移動先の実計算機(107)で動作するハイパバイザエージェント(130)へ移動先の仮想計算機(110)の起動状態を確認する(ステップ224)。移動先の仮想計算機(110)の起動状態を確認することで、正常に仮想計算機の移動が行われたと判断し、LPAR移動フローは終了となる。

【0108】

次に、図6により、本発明の一実施の形態に係る制御方法による仮想計算機の移動処理を行うスクリプトプログラムの処理について説明する。図6は本発明の一実施の形態に係

50

る制御方法による仮想計算機の移動処理を行うスクリプトプログラムの処理を示すフローチャートであり、このスクリプトプログラムは、管理サーバ(101)上で実行されている。なお、図6において、同様の処理が続く部分については省略した。省略部分を含む各処理の動作は以下で説明するとおりである。

【0109】

なお、それぞれの処理は、管理サーバ(101)から、移動元と移動先それぞれのLPARが構成されている実計算機(106, 107)で動作するハイババイザエージェント(129, 130)に指示を送ることで実現される。

【0110】

(1) スクリプトプログラムが起動の指示を管理サーバ(101)から受けたときに開始する(ステップ500)。

10

【0111】

(2) ConfirmSrcLparSupportActInhibitAction [移動元LPARが起動抑止に対応していることを確認する] (ステップ501)。

【0112】

(3) 直前の処理に成功したか否かを判断し(ステップ502)、ステップ502で直前の処理に成功していない場合は、処理を続行するか否かを判断する(ステップ503)。続行するか否かの判断材料は管理サーバ(101)の入力から与えられる。ステップ502で直前の処理に成功した場合、またはステップ503で処理を続行する場合、次のステップに進み、ステップ503で処理を続行しない場合は、異常終了として、スクリプトが異常に終了した状態を管理サーバ(101)に応答する(ステップ511)。

20

【0113】

(4) ConfirmSrcLparAutoActNoAction [移動元LPARの自動起動が無効であることを確認する] (ステップ504)。

【0114】

(5) 直前の処理に成功したか否かを判断し(ステップ505)、ステップ505で直前の処理に成功していない場合は、処理を続行するか否かを判断する(ステップ506)。続行するか否かの判断材料は管理サーバ(101)の入力から与えられる。ステップ505で直前の処理に成功した場合、またはステップ506で処理を続行する場合、次のステップに進み、ステップ506で処理を続行しない場合は、異常終了として、スクリプトが異常に終了した状態を管理サーバ(101)に応答する(ステップ511)。

30

【0115】

以下、各処理の後に、ステップ502およびステップ503(ステップ505およびステップ506)の処理が行われ、直前の処理に成功した場合、または処理を続行する場合、次のステップに進み、処理を続行しない場合は、異常終了として処理される。

【0116】

(6) ConfirmSrcLparActAction [移動元LPARが起動状態であることを確認する] (ステップ507)。

【0117】

以下、図示しないが、各処理を行う。

40

【0118】

(7) RequestDestLparActInhibitYesAction [移動先LPARに起動抑止状態を設定する]。

【0119】

(8) RequestSrcLparDeactAction [移動元LPARの停止を要求する]。

【0120】

(9) ConfirmSrcLparDeactAction [移動元LPARの停止を確認する]。

【0121】

50

- (10) RequestSrcLparActInhibitYesAction [移動元LPARに起動抑止状態を設定する]。
- 【0122】
- (11) SetSrcLparVfcNumberInhibitYesAction [移動元LPARの仮想FCの分割識別子を固定する]。
- 【0123】
- (12) SetDestLparVfcNumberInhibitYesAction [移動先LPARの仮想FCの分割識別子を固定する]。
- 【0124】
- (13) GetSrcLparAutoVnicAction [移動元LPARの仮想NICの情報を取得する]。 10
- 【0125】
- (14) GetDestLparAutoVnicAction [移動先LPARの仮想NICの情報を取得する]。
- 【0126】
- (15) SetDestLparAutoVnicAction [移動先LPARの仮想NICの情報を設定する]。
- 【0127】
- (16) SetSrcLparAutoVnicAction [移動元LPARの仮想NICの情報を設定する]。 20
- 【0128】
- (17) GetSrcLparPhysicalFCAction [移動元LPARの実FCの情報を取得する]。
- 【0129】
- (18) GetSrcLparLogicalFCAction [移動元LPARの仮想FCの情報を取得する]。
- 【0130】
- (19) GetDestLparPhysicalFCAction [移動先LPARの実FCの情報を取得する]。 30
- 【0131】
- (20) GetDestLparLogicalFCAction [移動先LPARの仮想FCの情報を取得する]。
- 【0132】
- (21) SetDestLparWWNAction [移動先LPARの仮想WWNの情報を設定する]。
- 【0133】
- (22) SetSrcLparWWNAction [移動元LPARの仮想WWNの情報を設定する]。
- 【0134】
- (23) GetSrcLparRTCDifferenceAction [移動元LPARの時刻差分情報を取得する]。 40
- 【0135】
- (24) GetDestLparRTCDifferenceAction [移動先LPARの時刻差分情報を取得する]。
- 【0136】
- (25) SetDestLparRTCDifferenceAction [移動先LPARの時刻差分情報を設定する]。
- 【0137】
- (26) SetSrcLparRTCDifferenceAction [移動元LPARの時刻差分情報を設定する]。 50

【0138】

(27) RequestSrcLparSaveConfigAction [移動元 LPAR の構成情報の保存を要求する]。

【0139】

(28) ConfirmSrcLparSaveConfigAction [移動元 LPAR の構成情報の保存を確認する]。

【0140】

(29) RequestDestLparSaveConfigAction [移動先 LPAR の構成情報の保存を要求する]。

【0141】

(30) ConfirmDestLparSaveConfigAction [移動先 LPAR の構成情報の保存を確認する]。

【0142】

(31) RequestDestLparActInhibitNoAction [移動先 LPAR の起動抑止状態を解除する]。

【0143】

(32) RequestDestLparActAction [移動先 LPAR の起動を要求する]。

【0144】

(33) ConfirmDestLparActAction [移動先 LPAR の起動を確認する]。

【0145】

(34) 全ての処理が終了すると正常終了として、スクリプトが正常に終了した状態を管理サーバ(101)に応答する(ステップ510)。

【0146】

次に、図7～図9により、本発明の一実施の形態に係る制御方法による仮想計算機の配置情報の分析処理について説明する。図7は本発明の一実施の形態に係る制御方法による仮想WWNを解析するスクリプトプログラムの動作を示すフローチャート、図8は本発明の一実施の形態に係る制御方法による仮想MACアドレスを解析するスクリプトプログラムの動作を示すフローチャート、図9は本発明の一実施の形態に係る制御方法によるハイパバイザエージェントを探索するスクリプトプログラムの動作を示すフローチャートであり、この各スクリプトプログラムは、管理サーバ(101)上で実行されている。

【0147】

まず、仮想WWNを解析するスクリプトプログラムの動作は、図7に示すように以下に説明するとおりである。

【0148】

(1) スクリプトプログラムは管理サーバ(101)から起動の指示を受けたときに開始する(ステップ600)。

【0149】

(2) WWPN(World Wide Port Name)かWWNN(World Wide Node Name)かを解析する(ステップ601)。仮想WWNの最下位1ビット目からの1ビットはWWPNかWWNNかを示している。スクリプトプログラムはこの情報を取り出す。

【0150】

(3) シャーシ固有のシリアル番号を解析する(ステップ602)。仮想WWNの最下位2ビット目からの15ビットはシャーシに固有なシリアル番号を示している。スクリプトプログラムはこの情報を取り出す。

【0151】

(4) パーティション番号を解析する(ステップ603)。仮想WWNの最下位17ビット目からの3ビットはパーティション番号を示している。スクリプトプログラムはこの

10

20

30

40

50

情報を取り出す。

【 0 1 5 2 】

(5) 物理 F C H B A の相対スロット番号を解析する (ステップ 6 0 4) 。 仮想 W W N の最下位 2 0 ビット目からの 4 ビットは物理 F C H B A の相対スロット番号を示している。スクリプトプログラムはこの情報を取り出す。なお、相対スロット番号とは物理的なブレード位置と物理 F C H B A が取り付けられている物理的な位置の差分を取り、相対的な表現にしたものである。

【 0 1 5 3 】

(6) 物理 F C H B A のポート番号を解析する (ステップ 6 0 5) 。 仮想 W W N の最下位 2 4 ビット目からの 1 ビットは物理 F C H B A 上でどのポート番号にあたるかを示している。スクリプトプログラムはこの情報を取り出す。

10

【 0 1 5 4 】

(7) ベンダ固有の番号と照合する (ステップ 6 0 6) 。 仮想 W W N の最下位 2 5 ビット目から 2 4 ビットはベンダに固有な番号を示している。スクリプトプログラムはこの情報を取り出し、予期したベンダ固有の値と一致していることを確認する。

【 0 1 5 5 】

(8) 仮想 F C H B A 識別番号を解析する (ステップ 6 0 7) 。 仮想 W W N の最下位 4 9 ビット目から 3 ビットは仮想 F C の識別番号を示している。スクリプトプログラムはこの情報を取り出す。なお、仮想 F C H B A の識別番号とは、物理 F C H B A を複数の仮想 F C H B A として仮想化したときに、仮想 F C H B A のそれぞれを識別するために付与される仮想 F C H B A に固有な値である。

20

【 0 1 5 6 】

(9) プラットフォームを解析する (ステップ 6 0 8) 。 仮想 W W N の最下位 5 2 ビット目から 3 ビットはプラットフォームを示している。スクリプトプログラムはこの情報を取り出す。なお、プラットフォームとはハードウェア製品につけられた名前のことである。

【 0 1 5 7 】

(1 0) 予約された値に一致するかを照合する (ステップ 6 0 9) 。 仮想 W W N の最下位 5 5 ビット目から 2 ビットは予約された値が格納されている。スクリプトプログラムはこの情報を取り出し、予約された値と一致しているかを確認する。

30

【 0 1 5 8 】

(1 1) ベンダ管理番号を解析する (ステップ 6 1 0) 。 仮想 W W N の最下位 5 7 ビット目から 4 ビットはベンダ管理番号を示している。スクリプトプログラムはこの情報を取り出す。ベンダ管理番号とは、ベンダに与えられた W W N 空間について、重複しないように、ベンダが用途別に割り当てた値である。

【 0 1 5 9 】

(1 2) 既定の値に一致するかを照合する (ステップ 6 1 1) 。 仮想 W W N の最下位 6 1 ビット目から 4 ビットは既定されている。スクリプトプログラムはこの情報を取り出し、既定された値と一致することを確認する。

【 0 1 6 0 】

(1 3) 全ての情報を取り出し、解析を終えたスクリプトプログラムは終了する (ステップ 6 1 2) 。

40

【 0 1 6 1 】

また、仮想 M A C アドレスを解析するスクリプトプログラムの動作は、図 8 に示すように以下に説明するとおりである。

【 0 1 6 2 】

(1) スクリプトプログラムは管理サーバ (1 0 1) から開始の指示を受けることで仮想 M A C アドレスの解析を開始する (ステップ 6 2 0) 。

【 0 1 6 3 】

(2) ベンダ識別番号と照合する (ステップ 6 2 1) 。 仮想 M A C アドレスの上位 1 バ

50

イト目からの3バイトはベンダ識別番号を示している。スクリプトプログラムはこの情報を取り出し、期待した値と一致することを確認する。

【0164】

(3) 仮想NICシステム識別番号を解析する(ステップ622)。仮想MACアドレスの上位4バイト目からの2バイトは仮想NICシステム識別番号を基に生成されている。この情報は仮想NICシステム識別番号と1対1で対応している。スクリプトプログラムはこの情報を取り出し、解析し、仮想NICシステム識別番号を取り出す。

【0165】

(4) LPAR番号を解析する(ステップ623)。仮想MACアドレスの上位6バイト目からの1バイトはLPAR番号を基に生成されている。この情報はLPAR番号と1対1で対応している。スクリプトプログラムはこの情報を取り出し、解析し、LPAR番号を取り出す。

【0166】

(5) 全ての情報を取り出し、解析を終えたスクリプトプログラムは終了する(ステップ624)。

【0167】

また、ハイパバイザエージェントを探索するスクリプトプログラムの動作は、図9に示すように以下に説明するとおりである。

【0168】

ハイパバイザエージェント(129, 130)は、LPARの特別な一種として構成されており、MACアドレスを解析し、特定の条件を満足するとき、具体的にはLPARの管理番号が17番であるときにハイパバイザエージェントとして一覧表示する。

【0169】

(1) スクリプトプログラムは管理サーバ(101)からの起動の指示を受けたときに実行を開始する(ステップ700)。

【0170】

(2) ICMPエコー要求をブロードキャストする(ステップ701)。スクリプトプログラムは管理サーバ(101)を通し、同じネットワークの全ての計算機に対してICMPエコー要求をブロードキャストする。このとき、ICMPエコー要求の送信を行うために、IPアドレスからMACアドレスを引き出すために、ARP要求を送信する。ARP要求によって解決されたIPアドレスとMACアドレスはARPテーブルに格納される。

【0171】

(3) ICMPエコー応答を受信する(ステップ702)。ICMPエコー要求を受信した計算機はICMPエコー要求の送信元へICMPエコー応答を送信する。スクリプトプログラムはこのICMPエコー応答を受信する。

【0172】

(4) テーブルにMACアドレスとIPアドレスを登録する(ステップ703)。ICMPエコー要求のブロードキャストとICMPエコー応答の受信を終えたシステムのARPテーブルには同じネットワークのIPアドレスとMACアドレスの対が登録されている。スクリプトプログラムはテーブルにこのMACアドレスとIPアドレスの対を登録する。

【0173】

(5) テーブルからMACアドレスを取り出し解析する(ステップ704)。スクリプトプログラムはテーブルに登録したMACアドレスとIPアドレスの対から1つを取り出し、MACアドレスについてそのMACアドレスが仮想MACアドレスであると仮定して解析する。

【0174】

(6) テーブルに格納されたMACアドレス全てについての解析が終わったかを判断する(ステップ705)。MACアドレスの解析は、テーブルに格納された数だけ行う。

10

20

30

40

50

【 0 1 7 5 】

(7) ステップ 7 0 5 でテーブルに格納された M A C アドレス全てについての解析が終わっていない場合は、繰り返しステップ 7 0 4 での解析結果がハイパバイザエージェントを示しているかを判断する (ステップ 7 0 6)。スクリプトプログラムはステップ 7 0 4 での解析結果について M A C アドレスがハイパバイザエージェントのものと期待されるかを判断する。

【 0 1 7 6 】

(8) ステップ 7 0 6 で解析結果がハイパバイザエージェントを示していない場合は、ステップ 7 0 4 に戻り、ステップ 7 0 6 で解析結果がハイパバイザエージェントを示している場合は、テーブルから解析した M A C アドレスに対応する I P アドレスを取り出し、その I P アドレスと解析した M A C アドレスの対を表示する (ステップ 7 0 7)。

【 0 1 7 7 】

(9) ステップ 7 0 5 で、テーブルに格納された M A C アドレス全てについての解析が終わり、ハイパバイザエージェントの I P アドレスと M A C アドレスの対と期待される情報を全て出力したスクリプトプログラムは終了する (ステップ 7 0 8)。

【 0 1 7 8 】

以上の仮想 W W N の解析、仮想 M A C アドレスの解析、ハイパバイザエージェントの探索により、仮想計算機の配置情報を分析することができ、仮想計算機を移動した結果として、現在どの実計算機上で仮想計算機が動作しているかを容易に判断することができる。

【 0 1 7 9 】

以上、本発明者によってなされた発明を実施の形態に基づき具体的に説明したが、本発明は前記実施の形態に限定されるものではなく、その要旨を逸脱しない範囲で種々変更可能であることはいうまでもない。

【 0 1 8 0 】

例えば、本実施の形態では、仮想計算機を移動する際に、移動元の仮想計算機と、移動先の仮想計算機の各情報を入れ替えることにより、移動を行うようにしているが、移動先の仮想計算機の情報に、移動元の仮想計算機の情報を上書きするようにしてもよい。その場合は、移動元の仮想計算機の情報を意味のない情報にするなどして、移動先の仮想計算機とは異なる情報とすればよい。

【 0 1 8 1 】

また、本実施の形態では、管理サーバにより、仮想計算機の移動処理を行っているが、管理サーバ以外の管理部により、仮想計算機の移動処理を行うようにしてもよい。

【 産業上の利用可能性 】

【 0 1 8 2 】

本発明は、計算機を多数運用するデータセンタなどにおいて、省エネルギーを目的に、あるいはトータルの計算機台数を無駄に増加させないようにするためシステムの負荷を監視しながら仮想計算機を特定の実計算機に移動して、空いた実計算機の電源を切断し、別用途に用いるなど効率的な計算機利用を図るシステムなどの広く適用可能である。

【 0 1 8 3 】

また、実計算機のハードウェア保守のとき、仮想計算機を別の実計算機に移動させることにより、保守のためのシステム停止時間を最小限にするシステムなどにも広く適用可能である。

【 図面の簡単な説明 】

【 0 1 8 4 】

【 図 1 】 本発明の一実施の形態に係る制御方法が適用された仮想計算機システムの構成を示す構成図である。

【 図 2 】 本発明の一実施の形態に係る制御方法による仮想計算機の移動動作を説明するための説明図である。

【 図 3 】 本発明の一実施の形態に係る制御方法による仮想計算機の移動動作を説明するための説明図である。

10

20

30

40

50

【図４】本発明の一実施の形態に係る制御方法による仮想計算機の移動動作を説明するための説明図である。

【図５】本発明の一実施の形態に係る制御方法による仮想計算機の移動シーケンスを示す図である。

【図６】本発明の一実施の形態に係る制御方法による仮想計算機の移動処理を行うスクリプトプログラムの処理を示すフローチャートである。

【図７】本発明の一実施の形態に係る制御方法による仮想W W Nを解析するスクリプトプログラムの動作を示すフローチャートである。

【図８】本発明の一実施の形態に係る制御方法による仮想M A Cアドレスを解析するスクリプトプログラムの動作を示すフローチャートである。

【図９】本発明の一実施の形態に係る制御方法によるハイパバイゼエージェントを探索するスクリプトプログラムの動作を示すフローチャートである。

【図１０】従来の仮想計算機システムの制御における具体的な問題点を説明するための説明図である。

【図１１】従来の仮想計算機システムの制御における具体的な問題点を説明するための説明図である。

【符号の説明】

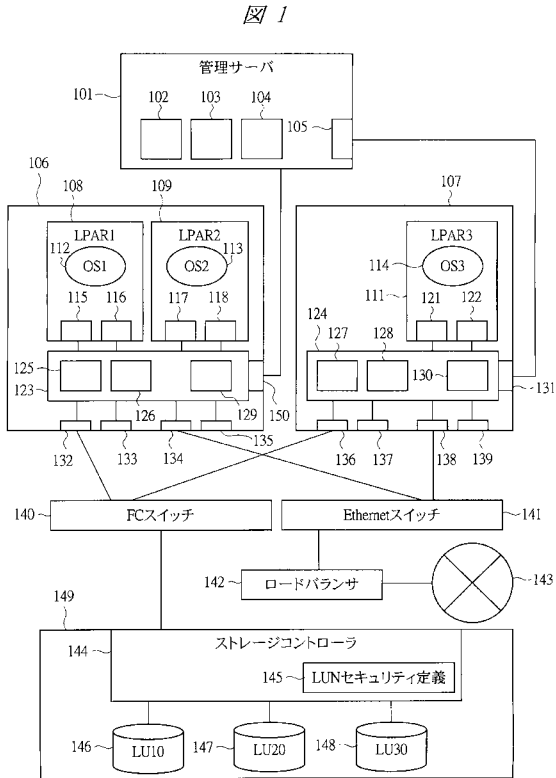
【 0 1 8 5 】

1 0 1 ... 管理サーバ、1 0 2 ... C P U、1 0 3 ... 主記憶、1 0 4 ... ストレージ、1 0 5 ... ネットワークインターフェイス、1 0 6 , 1 0 7 ... 実計算機、1 0 8 , 1 0 9 , 1 1 0 , 1 1 1 ... 仮想計算機、1 1 5 , 1 1 7 , 1 2 1 ... 仮想F C H B A、1 1 6 , 1 1 8 , 1 2 2 ... 仮想N I C、1 2 3 , 1 2 4 ... ハイパバイザ、1 2 5 , 1 2 7 ... 仮想F C H B A管理テーブル、1 2 6 , 1 2 8 ... 仮想N I C管理テーブル、1 2 9 , 1 3 0 ... ハイパバイゼエージェント、1 3 1 , 1 5 0 ... ネットワークインターフェイス、1 3 2 , 1 3 3 , 1 3 6 , 1 3 7 ... 物理F C H B A、1 3 4 , 1 3 5 , 1 3 8 , 1 3 9 ... 物理N I C、1 4 0 ... F Cスイッチ、1 4 1 ... E t h e r n e t (登録商標)スイッチ、1 4 2 ... ロードバランサ、1 4 3 ... 外部ネットワーク、1 4 4 ... ストレージコントローラ、1 4 5 ... L U Nセキュリティ定義、1 4 9 ... ストレージ装置、1 4 6 , 1 4 7 , 1 4 8 ... L U。

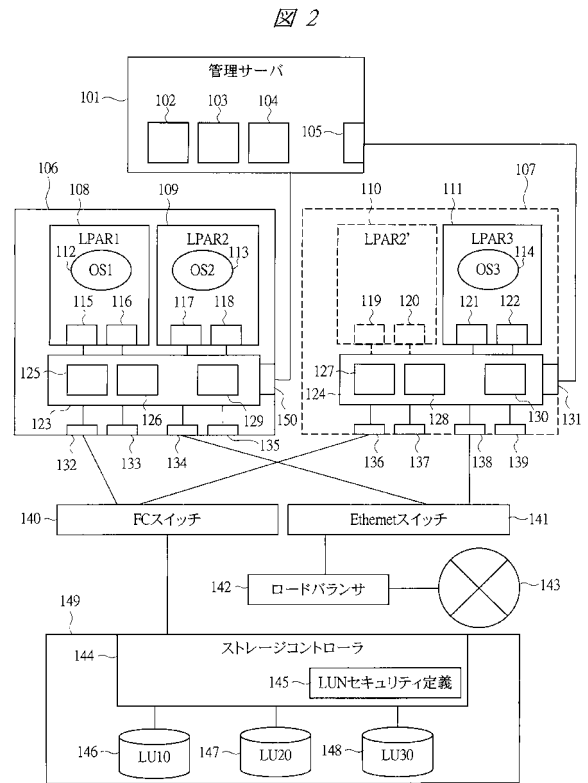
10

20

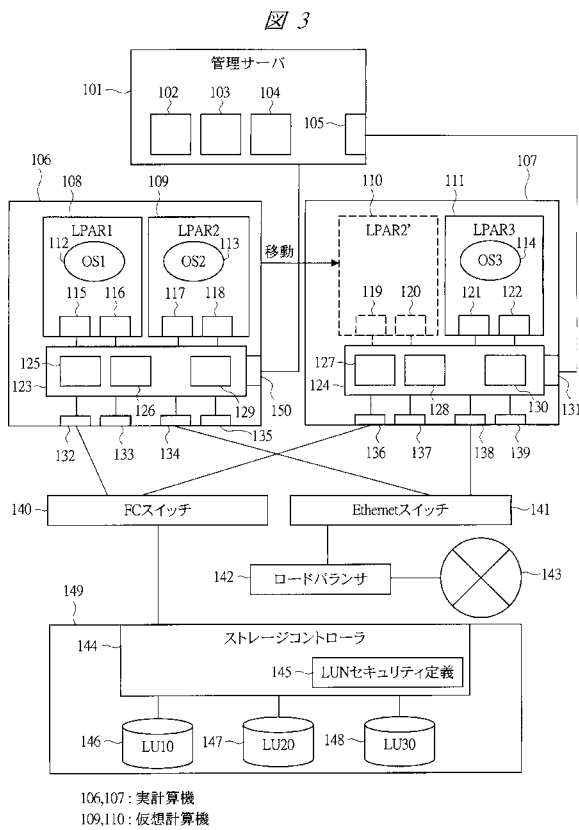
【図 1】



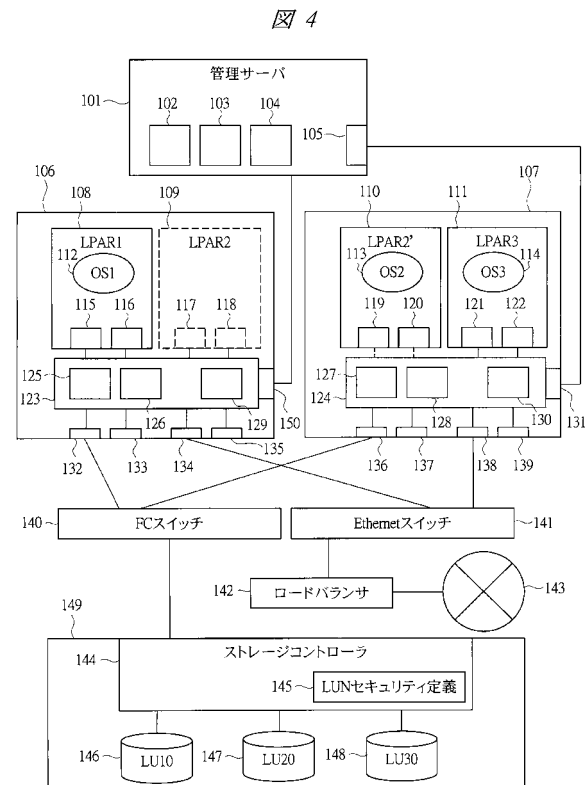
【図 2】



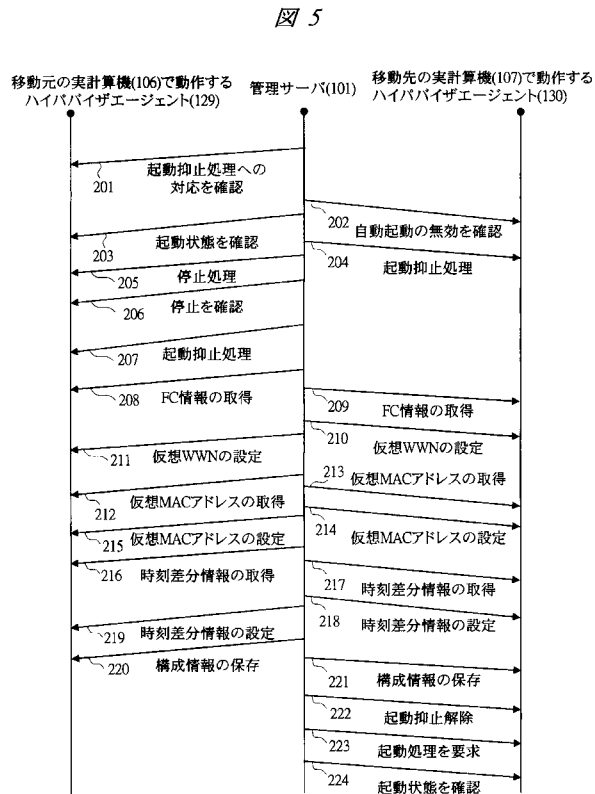
【図 3】



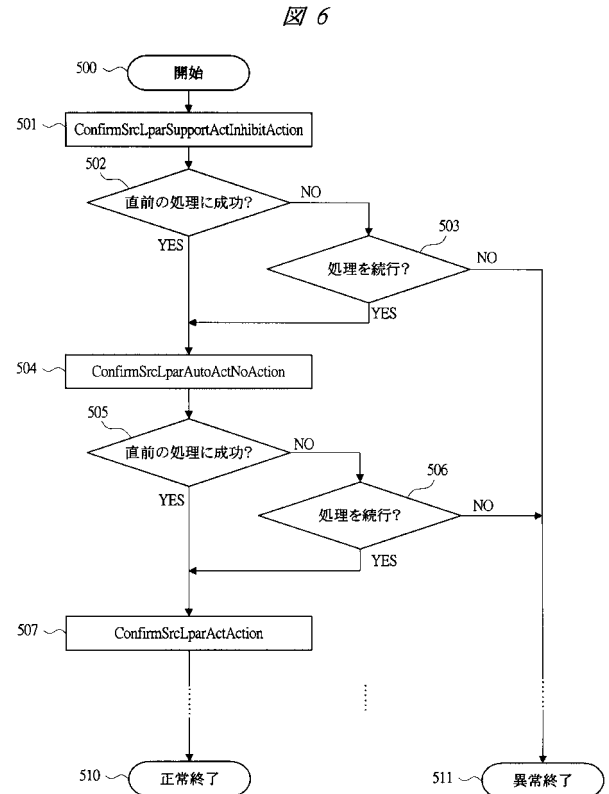
【図 4】



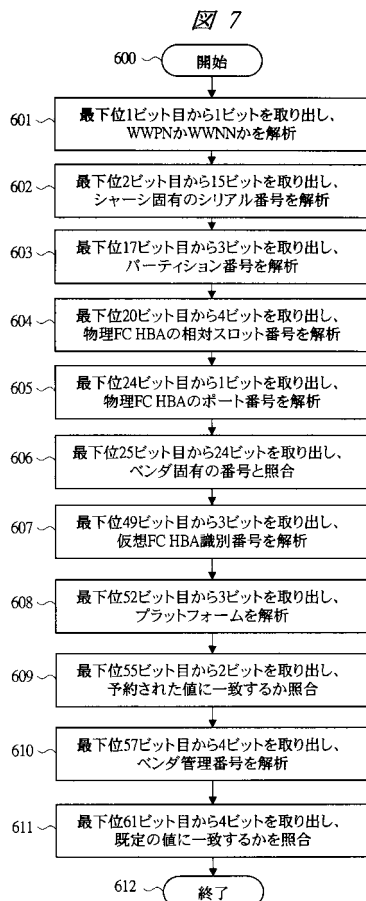
【図5】



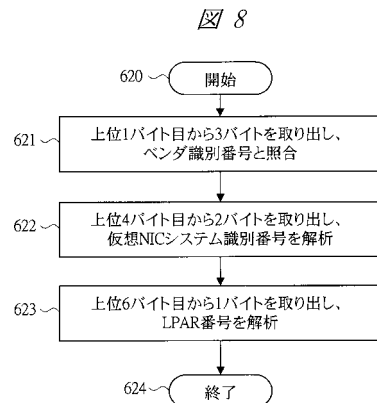
【図6】



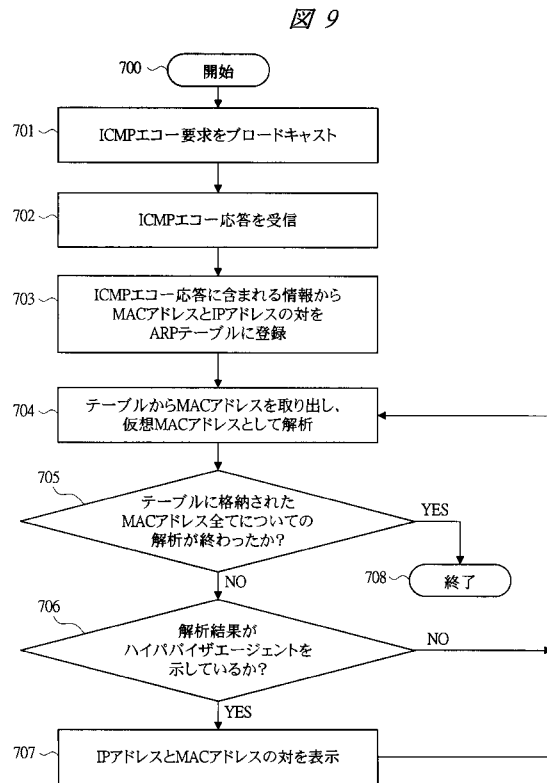
【図7】



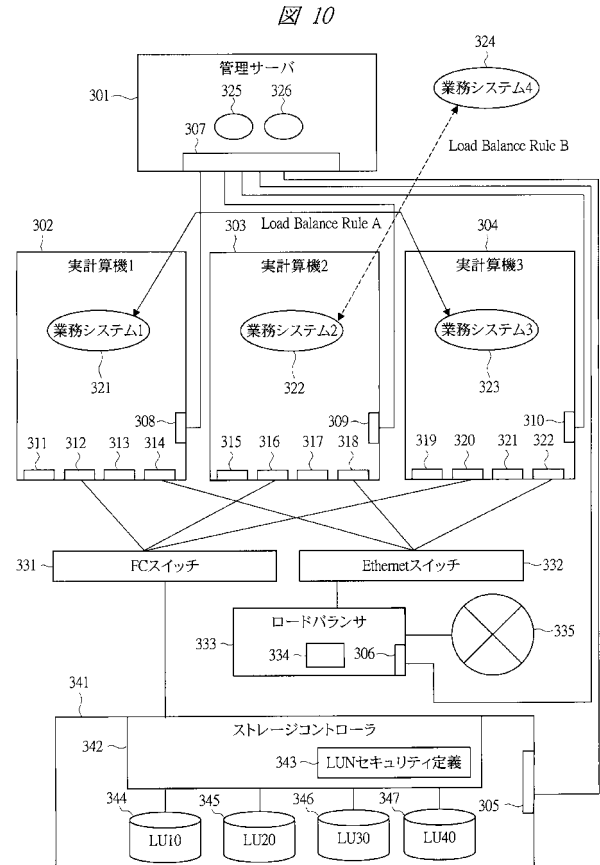
【図8】



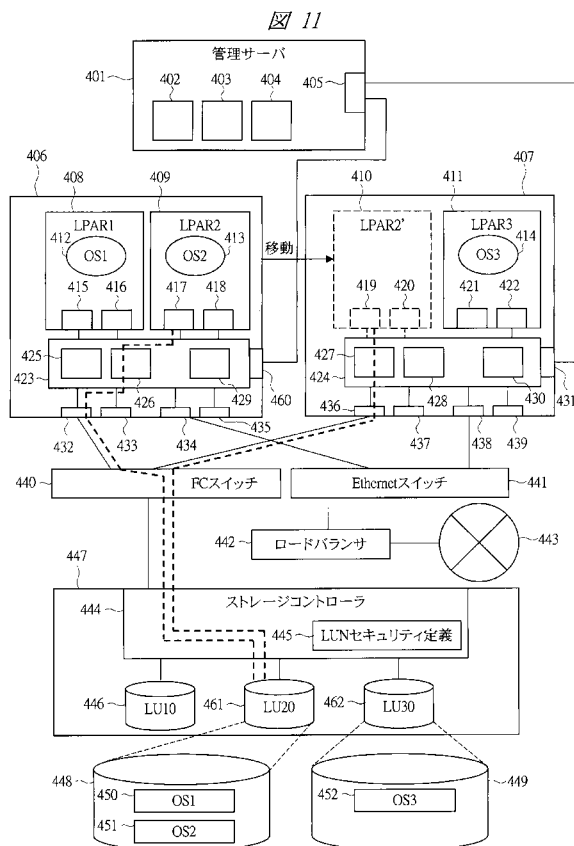
【図 9】



【図 10】



【図 11】



【手続補正書】

【提出日】平成21年6月22日(2009.6.22)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】0005

【補正方法】変更

【補正の内容】

【0005】

このようなシステムとして、ESX Server 3i Configuration Guide (非特許文献1)、Cisco MDS and Emulex Virtual HBA Solutions for VMware Infrastructure 3 (非特許文献2)に記載された技術があった。

【非特許文献1】ESX Server 3i Configuration Guide, Storage, 7 - Raw Device Mapping - Benefits of Raw device Mapping, (p113 - 115)、インターネット <URL:http://www.vmware.com/pdf/vi3_35/esx_3i_e/r35/vi3_35_25_3i_server_config.pdf>

【非特許文献2】Cisco MDS and Emulex Virtual HBA Solutions for VMware Infrastructure 3, 6 . Array level LUN Masking, Mapping and VMotion, (p12 - 14)、インターネット <URL:<http://www.emulex.com/white/hba/CiscoEmulexVirtualizationforVMware.pdf>>