

(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2014-41414

(P2014-41414A)

(43) 公開日 平成26年3月6日(2014.3.6)

(51) Int.Cl.

F I

テーマコード (参考)

G 0 6 F 9/46 (2006.01)

G 0 6 F 9/46 3 5 0

G 0 6 F 9/50 (2006.01)

G 0 6 F 9/46 4 6 2 Z

審査請求 未請求 請求項の数 7 O L (全 17 頁)

(21) 出願番号 特願2012-182340 (P2012-182340)
(22) 出願日 平成24年8月21日 (2012.8.21)

(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二丁目3番1号
(74) 代理人 100123788
弁理士 宮崎 昭夫
(74) 代理人 100148297
弁理士 ▲高▼森 俊夫
(72) 発明者 金子 斉
東京都千代田区大手町二丁目3番1号 日
本電信電話株式会社内
(72) 発明者 樋口 健一
東京都千代田区大手町二丁目3番1号 日
本電信電話株式会社内

最終頁に続く

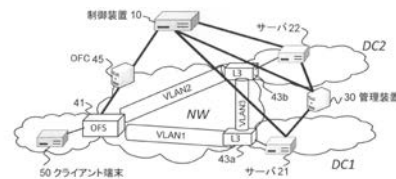
(54) 【発明の名称】 仮想マシンのマイグレーションシステム、制御装置およびマイグレーション方法

(57) 【要約】

【課題】 仮想マシンのライブマイグレーションを実施する際のサービス停止時間を短縮可能にしたマイグレーションシステムを提供する。

【解決手段】 第1 データセンタに設けられ、ネットワーク内の第1通信経路を介してクライアント端末と接続される第1サーバと、第2 データセンタに設けられ、ネットワーク内の第2通信経路を介してクライアント端末と接続される第2サーバと、クライアント端末と第1および第2のサーバとの通信経路となる第1および第2の通信経路を切り替え可能なネットワーク切替装置と、第1サーバが第1通信経路を介してクライアント端末に提供している仮想マシンを、第1サーバに対して第2サーバに移行する旨を指示した後、仮想マシンの第1サーバから第2サーバへの移行が完了する前に、クライアント端末に接続される通信経路を第1通信経路から第2通信経路に切り替える旨をネットワーク切替装置に指示する制御装置と、を有する。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

第 1 のデータセンタに設けられ、ネットワーク内の第 1 の通信経路を介してクライアント端末と接続される第 1 のサーバと、

前記第 1 のデータセンタとは異なる第 2 のデータセンタに設けられ、前記ネットワーク内の第 2 の通信経路を介して前記クライアント端末と接続される第 2 のサーバと、

前記クライアント端末と前記第 1 および第 2 のサーバとの通信経路となる、前記第 1 および第 2 の通信経路を切り替え可能なネットワーク切替装置と、

前記第 1 のサーバが前記第 1 の通信経路を介して前記クライアント端末に提供している仮想マシンを、該第 1 のサーバに対して前記第 2 のサーバに移行する旨を指示した後、該仮想マシンの前記第 1 のサーバから前記第 2 のサーバへの移行が完了する前に、前記クライアント端末に接続される通信経路を前記第 1 の通信経路から前記第 2 の通信経路に切り替えるように、前記ネットワーク切替装置に通信経路の切り替えを指示する制御装置と、を有する仮想マシンのマイグレーションシステム。

10

【請求項 2】

請求項 1 記載の仮想マシンのマイグレーションシステムにおいて、

前記制御装置は、

前記第 1 のサーバに関する所定のパラメタの値を取得し、取得したパラメタの値に基づいて、該第 1 のサーバに対して前記仮想マシンを前記第 2 のサーバに移行する旨を指示した時刻である移行指示時刻から前記ネットワーク切り替え装置に前記通信経路の切り替えを指示するまでの時間を算出し、前記移行指示時刻から該時間が経過したときに、前記ネットワーク切替装置に前記通信経路の切り替えを指示する、仮想マシンのマイグレーションシステム。

20

【請求項 3】

請求項 1 記載の仮想マシンのマイグレーションシステムにおいて、

前記第 2 のサーバは、

前記仮想マシンが前記第 1 のサーバで停止したことを検出すると、ダミーの G A R P 通知を前記制御装置に送信し、

前記制御装置は、

前記第 2 のサーバから前記ダミーの G A R P 通知を受信すると、前記ネットワーク切替装置に前記通信経路の切り替えを指示する、仮想マシンのマイグレーションシステム。

30

【請求項 4】

第 1 のデータセンタに設けられ、ネットワーク内の第 1 の通信経路を介してクライアント端末と接続される第 1 のサーバと、第 2 のデータセンタに設けられ、前記ネットワーク内の第 2 の通信経路を介して前記クライアント端末と接続される第 2 のサーバと、該クライアント端末と前記第 1 および第 2 のサーバとの通信経路となる、前記第 1 および第 2 の通信経路を切り替え可能なネットワーク切替装置と接続される制御装置であって、

前記第 1 のサーバが前記第 1 の通信経路を介して前記クライアント端末に提供している仮想マシンを、該第 1 のサーバに対して前記第 2 のサーバに移行する旨を指示する V M 切替制御部と、

40

前記仮想マシンの前記第 1 のサーバから前記第 2 のサーバへの移行が完了する前に、前記クライアント端末に接続される通信経路を前記第 1 の通信経路から前記第 2 の通信経路に切り替えるように、前記ネットワーク切替装置に通信経路の切り替えを指示する N W 切替制御部と、を有する制御装置。

【請求項 5】

請求項 4 記載の制御装置において、

前記第 1 のサーバに関する所定のパラメタの値を取得するパラメタ抽出部と、

前記パラメタ抽出部が取得した前記所定のパラメタの値に基づいて、前記第 1 のサーバに対して前記仮想マシンを前記第 2 のサーバに移行する旨を指示した時刻である移行指示

50

時刻から前記ネットワーク切り替え装置に前記通信経路の切り替えを指示するまでの時間を算出する計算部と、をさらに有し、

前記NW切替制御部は、

前記移行指示時刻から前記計算部が算出した時間が経過したときに、前記ネットワーク切替装置に前記通信経路の切り替えを指示する、制御装置。

【請求項6】

第1のデータセンタに設けられ、ネットワーク内の第1の通信経路を介してクライアント端末と接続される第1のサーバと、第2のデータセンタに設けられ、前記ネットワーク内の第2の通信経路を介して前記クライアント端末と接続される第2のサーバと、該クライアント端末と前記第1および第2のサーバとの通信経路となる、前記第1および第2の通信経路を切り替え可能なネットワーク切替装置と接続される制御装置による、仮想マシンのマイグレーション方法であって、

前記第1のサーバが前記第1の通信経路を介して前記クライアント端末に提供している仮想マシンを、該第1のサーバに対して前記第2のサーバに移行する旨を指示し、

前記仮想マシンの前記第1のサーバから前記第2のサーバへの移行が完了する前に、前記クライアント端末に接続される通信経路を前記第1の通信経路から前記第2の通信経路に切り替えるように、前記ネットワーク切替装置に通信経路の切り替えを指示する、仮想マシンのマイグレーション方法。

【請求項7】

請求項6記載の仮想マシンのマイグレーション方法において、

前記第1のサーバに関する所定のパラメタの値を取得し、

取得したパラメタの値に基づいて、該第1のサーバに対して前記仮想マシンを前記第2のサーバに移行する旨を指示した時刻である移行指示時刻から前記ネットワーク切り替え装置に前記通信経路の切り替えを指示するまでの時間を算出し、

前記移行指示時刻から該時間が経過したときに、前記ネットワーク切替装置に前記通信経路の切り替えを指示する、仮想マシンのマイグレーション方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、異なるデータセンタ拠点間で仮想マシンのライブマイグレーションを実行するためのマイグレーションシステム、制御装置およびマイグレーション方法に関する。

【背景技術】

【0002】

仮想マシン（Virtual Machine：以下では、「VM」と表記する）のライブマイグレーションの一例が、特許文献1に開示されている。ライブマイグレーションとは、あるコンピュータにおけるVMで稼働中のOS（Operating System）およびソフトウェアプログラムを停止させることなく、他のコンピュータに「移行」させることを意味する。ライブマイグレーションでは、VMを動作させるコンピュータを単に変更するだけでなく、その過程でも、クライアントに対してVM使用のサービスを提供し続ける必要がある。以下に、ライブマイグレーションの一例を簡単に説明する。

【0003】

図8は従来のマイグレーションシステムの一構成例を示すブロック図である。

【0004】

図8に示すように、データセンタDC1内にサーバ121が設けられ、データセンタDC2内にサーバ122が設けられている。これらの2つのデータセンタDC1、DC2のそれぞれに設けられたLAN（Local Area Network）が1つのネットワークNWを介して接続されている。ネットワークNWは、例えば、WAN（Wide Area Network）である。ネットワークNWにおける通信では、通信規約として、例えば、TCP（Transmission Control Protocol）／IP（Internet Protocol）が使用される。

【0005】

サーバ 121、122 は、クライアントに提供する VM を動作させる情報処理装置である。サーバ 121 はネットワーク NW の L3 スイッチ 43a と接続され、サーバ 122 はネットワーク NW の L3 スイッチ 43b と接続されている。L3 スイッチ 43a と L3 スイッチ 43b は仮想ネットワーク VLAN3 を介して接続されている。

【0006】

クライアント端末 50 は、クライアントがサーバ 121 またはサーバ 122 で VM を使用するための端末であり、ネットワーク NW の OFS 41 と接続されている。クライアント端末 50 は、VM 宛にデータをパケットで送信する際、VM に予め割り当てられた IP アドレスをパケットのヘッダの宛先に書き込み、パケットを送信する。

【0007】

管理装置 30 は、サーバ 121、122 の負荷を監視し、サーバの負荷が予め決められた閾値以上になると、負荷が閾値以上になったサーバの情報を含む負荷情報を制御装置 100 に送信する。

【0008】

OFC (Open Flow Controller) 45 は、ネットワーク内に設けられた OFS (Open Flow Switch) 41 を制御する。OFC 45 は、制御装置 100 と接続され、IP アドレスに対応する MAC (Media Access Control) アドレスの変更情報を含む経路切替指示を制御装置 100 から受け取ると、その変更情報を含む経路切替情報を OFS 41 に通知する。

【0009】

OFS 41 はネットワーク切替装置の一種である。OFS 41 は、OFC 45 から受信する経路切替情報にしたがってフローテーブルを更新し、フローテーブルを参照してクライアント端末 50 とサーバ 121、122 との通信経路を切り替える。

【0010】

制御装置 100 は、管理装置 30 から負荷情報を受信すると、負荷の高いサーバで稼働している VM を負荷の低いサーバに移行させるために、負荷の高いサーバに、VM 移行の動作を指示する。また、VM の移動先のサーバから GARP (Gratuitous Address Resolution Protocol) 通知を受信すると、経路切替指示を OFC 45 に送信する。

【0011】

図 9 は図 8 に示すシステムの動作手順を示すシーケンス図である。

【0012】

ここでは、データセンタ DC1 内のサーバ 121 で稼働中の VM を、サーバ 121 の過負荷などの要因に因り、データセンタ DC2 のサーバ 122 にマイグレーションする場合で説明する。その際、遅延防止、および VM の負荷分散だけでなく、ネットワークの負荷分散も実現するために、VM のライブマイグレーション（以下では、VM マイグレーションと称する）に連動して、仮想ネットワーク VLAN による通信経路も切り替える。

【0013】

制御装置 100 は管理装置 30 を介してサーバ 121 の負荷を監視する。制御装置 100 は、サーバ 121 の過負荷を検出すると、VM 移行を指示する旨の VM 移行指示を、管理装置 30 を介してサーバ 121 に送信する（ステップ 301）。サーバ 121 は、稼働中の VM で使用される OS、ソフトウェアプログラムおよびデータを、仮想ネットワーク VLAN3 を介してサーバ 122 にコピーする VM 移行動作を行い（ステップ 303）、続いて、VM 切替動作を行う（ステップ 304）。この「VM 切替」とは、サーバ 121 で稼働している VM を一旦停止し、VM のコピー中に生じたデータ分をサーバ 122 に新たに稼働させる VM にコピーした後、新たに稼働させる VM 側でメモリ再開処理を行うことである。

【0014】

サーバ 122 は、VM マイグレーションが終了すると、GARP 通知を制御装置 100 に送信する（ステップ 305）。この GARP 通知は、IP アドレスで指定される VM に関して、「VM 移行」と「VM 切替」を含む VM マイグレーションが終了したことと、V

10

20

30

40

50

Mの移動先となるサーバ122を制御装置100に通知するためのものである。そのため、GARP通知は、サーバ121で稼働していたVMを指定するためのIPアドレスとVMの移動先となるサーバ122のMACアドレスの情報を含む。

【0015】

制御装置100は、GARP通知をサーバ122から受信すると、VMマイグレーションが終了したことを認識し、VMおよびクライアント端末50間の通信経路を仮想ネットワークVLAN1から仮想ネットワークVLAN2に切り替えるための経路切替指示をOFC45に送信する(ステップ306)。OFC45は、制御装置100から経路切替指示を受け取ると、経路切替情報をOFS41に送信する(ステップ307)。

【0016】

OFS41は、OFC45から経路変更情報を受け取ると、フローテーブルにおいて、VMのIPアドレスに対応するMACアドレスをサーバ121からサーバ122に書き換える。そして、OFS41は、更新したフローテーブルにしたがって、経路切替処理を実行する。具体的には、OFS41は、クライアント端末-サーバ間のフローを仮想ネットワークVLAN1から仮想ネットワークVLAN2に切り替える。

【先行技術文献】

【特許文献】

【0017】

【特許文献1】特開2011-210032号公報(第7図)

【発明の概要】

【発明が解決しようとする課題】

【0018】

同じデータセンタ内でライブマイグレーションを行う場合、仮想ネットワークVLAN1から仮想ネットワークVLAN2への切り替えのようなネットワークの切り替えを行う必要がない。そのため、クライアント端末50からVM宛に送信されたパケットが行先不明で損失してしまうパケットロスが発生する時間は、図9に示すステップ304の「VM切替」の時間で済む。

【0019】

しかし、異なるデータセンタ間でライブマイグレーションを行う場合、つまり、ライブマイグレーションがデータセンタ間に跨る場合、ステップ304の「VM切替」からネットワーク切替終了まで、サービス停止時間が継続するため、パケットロスの量が多くなってしまう。

【0020】

この問題を、特許文献1に開示された方法で説明する。特許文献1の図7を参照すると、Manager151は、VM125からVM起動通知711を受け取ってから、ノード104に切替指示433を送信している。そのため、特許文献1に開示された方法では、上記の問題が発生するおそれがある。

【0021】

本発明は上述したような技術が有する問題点を解決するためになされたものであり、ライブマイグレーションを実施する際のサービス停止時間を短縮可能にした、仮想マシンのマイグレーションシステム、制御装置およびマイグレーション方法を提供することを目的とする。

【課題を解決するための手段】

【0022】

上記目的を達成するための本発明の仮想マシンのマイグレーションシステムは、

第1のデータセンタに設けられ、ネットワーク内の第1の通信経路を介してクライアント端末と接続される第1のサーバと、

前記第1のデータセンタとは異なる第2のデータセンタに設けられ、前記ネットワーク内の第2の通信経路を介して前記クライアント端末と接続される第2のサーバと、

前記クライアント端末と前記第1および第2のサーバとの通信経路となる、前記第1お

10

20

30

40

50

よび第 2 の通信経路を切り替え可能なネットワーク切替装置と、

前記第 1 のサーバが前記第 1 の通信経路を介して前記クライアント端末に提供している仮想マシンを、該第 1 のサーバに対して前記第 2 のサーバに移行する旨を指示した後、該仮想マシンの前記第 1 のサーバから前記第 2 のサーバへの移行が完了する前に、前記クライアント端末に接続される通信経路を前記第 1 の通信経路から前記第 2 の通信経路に切り替えるように、前記ネットワーク切替装置に通信経路の切り替えを指示する制御装置と、を有する構成である。

【0023】

また、本発明の制御装置は、第 1 のデータセンタに設けられ、ネットワーク内の第 1 の通信経路を介してクライアント端末と接続される第 1 のサーバと、第 2 のデータセンタに設けられ、前記ネットワーク内の第 2 の通信経路を介して前記クライアント端末と接続される第 2 のサーバと、該クライアント端末と前記第 1 および第 2 のサーバとの通信経路となる、前記第 1 および第 2 の通信経路を切り替え可能なネットワーク切替装置と接続される制御装置であって、

10

前記第 1 のサーバが前記第 1 の通信経路を介して前記クライアント端末に提供している仮想マシンを、該第 1 のサーバに対して前記第 2 のサーバに移行する旨を指示する VM 切替制御部と、

前記仮想マシンの前記第 1 のサーバから前記第 2 のサーバへの移行が完了する前に、前記クライアント端末に接続される通信経路を前記第 1 の通信経路から前記第 2 の通信経路に切り替えるように、前記ネットワーク切替装置に通信経路の切り替えを指示する NW 切替制御部と、を有する構成である。

20

【0024】

さらに、本発明の仮想マシンのマイグレーション方法は、第 1 のデータセンタに設けられ、ネットワーク内の第 1 の通信経路を介してクライアント端末と接続される第 1 のサーバと、第 2 のデータセンタに設けられ、前記ネットワーク内の第 2 の通信経路を介して前記クライアント端末と接続される第 2 のサーバと、該クライアント端末と前記第 1 および第 2 のサーバとの通信経路となる、前記第 1 および第 2 の通信経路を切り替え可能なネットワーク切替装置と接続される制御装置による、仮想マシンのマイグレーション方法であって、

30

前記第 1 のサーバが前記第 1 の通信経路を介して前記クライアント端末に提供している仮想マシンを、該第 1 のサーバに対して前記第 2 のサーバに移行する旨を指示し、

前記仮想マシンの前記第 1 のサーバから前記第 2 のサーバへの移行が完了する前に、前記クライアント端末に接続される通信経路を前記第 1 の通信経路から前記第 2 の通信経路に切り替えるように、前記ネットワーク切替装置に通信経路の切り替えを指示するものである。

【発明の効果】

【0025】

本発明によれば、異なるデータセンタ間で仮想マシンのライブマイグレーションを実施する際に発生するサービス停止時間を短縮できる。

40

【図面の簡単な説明】

【0026】

【図 1】第 1 の実施形態の VM マイグレーションの一構成例を示すブロック図である。

【図 2】図 1 に示した制御装置の一構成例を示すブロック図である。

【図 3】第 1 の実施形態の VM マイグレーションの手順を示すシーケンス図である。

【図 4】第 1 の実施形態の VM マイグレーションの効果を説明するための表である。

【図 5】第 2 の実施形態におけるサーバの一構成例を示すブロック図である。

【図 6】第 2 の実施形態における制御装置の一構成例を示すブロック図である。

【図 7】第 2 の実施形態の VM マイグレーションの手順を示すシーケンス図である。

【図 8】従来のマイグレーションシステムの構成例を示すブロック図である。

50

【図 9】図 8 に示したシステムの手順を示すシーケンス図である。

【発明を実施するための形態】

【0027】

従来技術では、制御装置は、VMマイグレーション終了（新たなVMの動作開始）に伴う「GARP通知」をトリガにしてOFCを介してOFSに「経路切替指示」を発したが、本発明では、VMマイグレーションが終了する前に、制御装置がOFCを介してOFSに経路切替指示を発することで、パケットロスの時間を低減するものである。これを実現可能にするための本発明の実施形態を、以下に説明する。

【0028】

（第 1 の実施形態）

本実施形態のVMマイグレーションについて説明する。

【0029】

図 1 は本実施形態のVMマイグレーションの一構成例を示すブロック図である。なお、図 8 を参照して説明した構成と同様な構成についての詳細な説明を省略する。

【0030】

図 1 に示すように、データセンタDC1内にサーバ21が設けられ、データセンタDC2内にサーバ22が設けられている。サーバ21、22は、クライアントに提供するVMを動作させる情報処理装置である。サーバ21はネットワークNWのL3スイッチ43aと接続され、サーバ22はネットワークNWのL3スイッチ43bと接続されている。

【0031】

なお、説明を簡単にするために、本実施形態では、図 1 に示すように、サーバがサーバ21、22の2台の場合で説明するが、VMを稼働させることが可能なサーバの台数が3台以上であってもよい。データセンタの数も2つに限らない。また、OFC45に1つのOFS41が接続される場合を示しているが、複数のOFS41がOFC45に接続されていてもよい。

【0032】

次に、本実施形態における制御装置10の構成を説明する。図 2 は図 1 に示した制御装置の一構成例を示すブロック図である。

【0033】

図 2 に示すように、制御装置10は、制御部11、VM切替制御部12、およびNW切替制御部13を含む構成である。制御部11は、パラメタ抽出部14と、Config部15と、X値計算部16と、連動制御部17とを有する。

【0034】

Config部15には、経路切替指示を送信するタイミングを決定するために用いられるパラメタの値が予め格納されている。Config部15が記憶するパラメタについては後述する。

【0035】

パラメタ抽出部14は、管理装置30から負荷情報を受信すると、負荷情報から高負荷のサーバの情報を読み出し、そのサーバに予め決められたパラメタの値を要求する。そして、パラメタ抽出部14は、移行対象のVMを動作させているサーバからパラメタの値を取得すると、取得したパラメタの値をConfig部15から読み出したパラメタの値と一緒にしてX値計算部16に渡す。

【0036】

X値計算部16は、連動制御部17がVM移行指示を送信した時刻の情報を連動制御部17から受け取り、パラメタ抽出部14からパラメタの値を受け取ると、パラメタの値を所定の計算式に代入して、「VM移行指示を送信した時刻」から「経路切替指示を送信する時刻」までの時間Xを算出する。以下に、算出方法の一例を説明する。

【0037】

<パラメタ>

C： メモリコピー帯域 [Mb/s]

10

20

30

40

50

M： サーバ 2 1 のメモリ領域 [Mbyte]

R： サーバ 2 1 のメモリ使用率

V： VM切替時間 [sec]

P： 調整パラメタ

N： ネットワーク切り替え時間 [sec]

T： VM移行指示からメモリコピー開始までの時間 [sec]

これらのパラメタのうち、C、M、V、P、NおよびTはC o n f i g部 1 5に予め格納されている。

【0038】

X値計算部 1 6は、サーバ 2 1からパラメタRおよびTの値を取得し、C o n f i g部 1 5からパラメタC、M、V、P、NおよびTの値を読み出し、これらのパラメタの値を、下記の式（1）に代入して、X値を算出する。

【0039】

$$X = (M \times R \times 8 \times P / C) + (V / 2) - N + T \cdots \text{式 (1)}$$

ただし、式（1）において、Xが負の場合には、連動制御部 1 7が送信するVM移行指示と経路切替指示を逆転させて調整する。

【0040】

連動制御部 1 7は、管理対象のサーバ 2 1、2 2のMACアドレスの情報を記憶する。連動制御部 1 7は、管理装置 3 0から負荷情報を受信すると、負荷情報から高負荷のサーバの情報を読み出し、そのサーバ宛にVM移行指示を送信するようにVM切替制御部 1 2に通知する。また、連動制御部 1 7は、X値計算部 1 7からX値を受け取ると、VM移行指示を送信した時刻からXの時間経過したときに、経路切替指示を送信するようにNW切替制御部 1 3に通知する。このとき、連動制御部 1 7は、VM移行指示を送信した際にVMをどのサーバに移行させるかを認識しているので、移行前のVMが使用していたIPアドレスと移行先のサーバのMACアドレスの情報を含む経路情報を経路切替指示に添付する。

【0041】

VM切替制御部 1 2は、連動制御部 1 7からの通知にしたがって、指定されたサーバ宛にVM移行指示を送信する。このVM移行指示には、移行対象のVMを特定するための情報と、移行先のサーバを特定するための情報が含まれている。

【0042】

NW切替制御部 1 3は、連動制御部 1 7からの通知にしたがって、経路切替指示をO F C 4 5に送信する。本実施形態では、VM移行指示をトリガとし、VM移行指示からX秒後に「経路切替指示」を発するようにしている。Xの値は、「経路切替指示」に因るO F S 4 1での経路切替が、「VM切替」のちょうど中間点で終了するように、上記のパラメタから算出される。

【0043】

なお、制御部 1 1、VM切替制御部 1 2およびNW切替制御部 1 3の一部または全部は、各機能を実行するための専用回路で構成されてもよく、プログラムを記憶するメモリ（不図示）とCPU（Central Processing Unit）（不図示）とが制御部 1 0に予め設けられ、CPUがプログラムにしたがって処理を実行することで仮想的に構成されてもよい。

【0044】

また、サーバ 2 1で稼働するVMで使用されていたIPアドレスは、GARP通知よりも前に、管理装置 3 0から制御装置 1 0に予め通知されていてもよく、管理装置 3 0が負荷情報を制御装置 1 0に送信する際に制御装置 1 0に通知してもよく、その方法は限定されない。

【0045】

次に、本実施形態のVMマイグレーションにおける、マイグレーション方法の動作手順を説明する。図 3は本実施形態のVMマイグレーションの動作手順を示すシーケンス図である。

【0046】

ここでは、データセンタDC1内のサーバ21で稼働中のVMを、サーバ21の過負荷などの要因に因り、データセンタDC2のサーバ22にマイグレーションする場合で説明する。

【0047】

制御装置10は管理装置30を介してサーバ21の負荷を監視する。制御装置10は、サーバ21の過負荷を検出すると、メモリ使用率等のパラメタの値をサーバ21から取得する(ステップ201)。続いて、制御装置10は、VM移行指示を管理装置30を介してサーバ21に送信する(ステップ202、203)。サーバ21は、稼働中のVMで使
10

用されるOS、ソフトウェアプログラムおよびデータを、仮想ネットワークVLAN3を介してサーバ22にコピーするVM移行動作を行う(ステップ204)。

【0048】

一方、制御装置10は、ステップ202の後、X値を算出し、VM移行指示を送信してからX秒後に経路切替指示をOFC45に送信する(ステップ205)。サーバ21、22は、ステップ204の後、VM切替動作を行う(ステップ206)。OFC45は、制御装置10から経路切替指示を受信すると、経路切替情報を含む経路切替指示をOFC41に送信する(ステップ207)。OFC45は、OFC41から受信した経路切替情報にしたがってフローテーブルを更新し、VMの通信経路として、仮想ネットワークVLAN1から仮想ネットワークVLAN2に切り替える。その結果、ステップ206のVM切替
20

の途中で、ネットワーク切替が終了する。

【0049】

サーバ22は、VM切替動作が終了すると、GARP通知を制御装置10に送信する(ステップ208)。制御装置10は、GARP通知をサーバ22から受信すると、VMマイグレーションが終了したことを認識する。

【0050】

図3を参照して説明したように、VM切替の途中でネットワーク切替が終了しているため、VM切替の後にネットワーク切替にかかる時間が不要となり、クライアントに対するサービスの停止がVM切替の時間で済む。

【0051】

次に、本実施形態のマイグレーション方法による効果を具体例で説明する。図4は本実施形態のVMマイグレーションによるシミュレーションの結果を示す表である。
30

【0052】

図4は、パターン1およびパターン2の2つのパターンの場合を示す。パターン1は、メモリコピー帯域Cが100[Mb/sec]の場合であり、パターン2は、メモリコピー帯域Cが1000[Mb/sec]の場合である。

【0053】

VM移行指示の時刻をT0とすると、パターン1の場合、時刻T0からネットワーク切替時刻(NM切替時刻)までの時間が49.05秒であり、時刻T0からVM切替時刻までの時間が48.3~49.8である。また、パターン2の場合、時刻T0からNM切替時刻までの時間が5.85秒であり、時刻T0からVM切替時刻までの時間が5.3~6.8である。パターン1、2のいずれの場合でも、ネットワーク切替がVM切替中に行われていることがわかる。時刻T0は、制御装置10がサーバ21に対してVMをサーバ21に移行する旨を指示した時刻に相当する。
40

【0054】

従来では、パケットロスの発生する時間が「VM切替時間V+ネットワーク切替時間N」であったのに対し、本実施形態のマイグレーション方法では、パケットロスの発生する時間を「VM切替時間V」に短縮できる。図4に示す例では、パケットロスの発生する時間が4.5秒から1.5秒に短縮されている。

【0055】

本実施形態によれば、制御装置が、VM移行前のサーバから抽出したパラメタの値に基
50

づいて、VM移行指示を送信した時刻から経路切替指示を送信するまでの時間を算出し、VMマイグレーションの終了前に経路切替指示を送信している。その結果、VM切替の動作の途中で、ネットワーク切り替えが行われる。そのため、制御装置がGARP通知を受信してから経路切替指示を送信する従来技術に比べて、「ネットワーク切替」の時間分、クライアントに対するサービスの停止時間を短縮することができる。

【0056】

また、クライアント端末50とVMとが仮想ネットワークVLANで通信していない場合、サーバ22は、VMマイグレーションが終了したとき、ネットワークNWに設けられた全てのネットワーク機器に対して、MACアドレスの変更を通知するためのARP通知をブロードキャストで通知する必要がある。これに対して、本実施形態では、VMマイグレーションに連動して仮想ネットワークVLANによる通信経路も切り替えているので、ネットワークの負荷分散を実現することができる。

10

【0057】

(第2の実施形態)

第1の実施形態では、経路切替指示を発するタイミングを制御装置が所定のパラメタの値に基づいて決定するものであったが、本実施形態は、VMの移行先のサーバが経路切替指示の発生トリガとなる「ダミーのGARP通知」を制御装置に送信するものである。

【0058】

本実施形態のVMマイグレーションを説明する。なお、本実施形態のシステムの全体は、図1に示したシステムと同様であり、図1または図8を参照して説明した構成と同様な構成についての詳細な説明を省略し、第1の実施形態と異なる点を詳しく説明する。

20

【0059】

はじめに、本実施形態におけるサーバ22の構成を説明する。図5は本実施形態におけるサーバの一構成例を示すブロック図である。

【0060】

本実施形態では、VMがサーバ21からサーバ22にライブマイグレーションする場合とし、サーバ22の構成を説明するが、サーバ21もサーバ22と同様な構成であってもよい。また、本実施形態では、本発明のライブマイグレーション方法に関連する構成について説明し、VMを稼働するための機能についての詳細な説明を省略する。

【0061】

30

図5に示すように、サーバ22は、VMで使用されるOS、ソフトウェアプログラムおよびデータを格納するための記憶部25と、ネットワークNWを介してクライアント端末50とデータを送受信する通信部26と、各部を制御する制御部27とを有する。

【0062】

制御部27は、メモリコピー等のVM移行の動作が終了し、サーバ21でのVM停止を検出すると、ダミーのGARP通知を制御装置10に送信する。ダミーのGARP通知は、制御装置10に経路切替指示を発生させるためのトリガとなるものである。ダミーのGARP通知には、少なくとも制御装置10に経路切替指示を送信可能にするための情報が含まれていればよい。例えば、サーバ21で稼働していたVMを指定するためのIPアドレスとVMの移動先となるサーバ22のMACアドレスの情報がダミーのGARP通知に含まれている。そして、制御部27は、VMマイグレーションが終了すると、GARP通知を制御装置10に送信する。このGARP通知には、VMマイグレーションが終了したことを示す情報が含まれている。

40

【0063】

次に、本実施形態における制御装置10の構成を説明する。図6は本実施形態における制御装置の一構成例を示すブロック図である。

【0064】

図6に示すように、本実施形態の制御装置10は、連動制御部17を含む制御部11と、VM切替制御部12と、NW切替制御部13とを有する構成である。

【0065】

50

連動制御部 17 は、管理対象のサーバ 21、22 の MAC アドレスの情報を記憶する。連動制御部 17 は、管理装置 30 から負荷情報を受信すると、負荷情報から高負荷のサーバの情報を読み出し、そのサーバ宛に VM 移行指示を送信するように VM 切替制御部 12 に通知する。また、連動制御部 17 は、サーバからダミーの GARP 通知を受け取ると、経路切替指示を送信するように NW 切替制御部 13 に通知する。このとき、連動制御部 17 は、VM をどのサーバに移行させるかを認識しているので、移行前の VM が使用していた IP アドレスと移行先のサーバの MAC アドレスの情報を含む経路情報を経路切替指示に添付する。

【0066】

VM 切替制御部 12 は、連動制御部 17 からの通知にしたがって、指定されたサーバ宛に VM 移行指示を送信する。この VM 移行指示には、移行対象の VM を特定するための情報と、移行先のサーバを特定するための情報が含まれている。NW 切替制御部 13 は、連動制御部 17 からの通知にしたがって、経路切替指示を OFC 45 に送信する。

【0067】

次に、本実施形態の VM マイグレーションにおける、マイグレーション方法の動作手順を説明する。図 7 は本実施形態の VM マイグレーションの動作手順を示すシーケンス図である。

【0068】

本実施形態においても、データセンタ DC 1 内のサーバ 21 で稼働中の VM を、サーバ 21 の過負荷などの要因に因り、データセンタ DC 2 のサーバ 22 にマイグレーションする場合で説明する。

【0069】

制御装置 10 は管理装置 30 を介してサーバ 21 の負荷を監視する。制御装置 10 は、サーバ 21 の過負荷を検出すると、VM 移行指示を管理装置 30 を介してサーバ 21 に送信する（ステップ 251、252）。サーバ 21 は、稼働中の VM で使用される OS、ソフトウェアプログラムおよびデータを、仮想ネットワーク VLAN 3 を介してサーバ 22 にコピーする VM 移行動作を行う（ステップ 253）。

【0070】

サーバ 22 は、ステップ 253 の VM 移行動作が終了したときに、ダミーの GARP 通知を制御装置 10 に送信し（ステップ 254）、続いて、サーバ 21、22 は、VM 切替動作を開始する（ステップ 255）。制御装置 10 は、ダミーの GARP 通知をサーバ 22 から受信すると、VM の IP アドレスとサーバ 22 の MAC アドレスの情報を含む経路切替指示を OFC 45 に送信する（ステップ 256）。サーバ 22 は、VM 切替動作が終了すると、GARP 通知を制御装置 10 に送信する（ステップ 257）。制御装置 10 は、GARP 通知をサーバ 22 から受信すると、この時点で、VM マイグレーションが終了したことを認識する。

【0071】

一方、OFC 45 は、制御装置 10 から経路切替指示を受信すると、経路切替情報を含む経路切替指示を OFC 41 に送信する（ステップ 258）。OFC 45 は、OFC 41 から受信した経路切替情報にしたがってフローテーブルを更新し、VM の通信経路として、仮想ネットワーク VLAN 1 から仮想ネットワーク VLAN 2 に切り替える。

【0072】

図 7 を参照して説明したように、VM 移行動作が終了したときに、サーバ 22 から送出されるダミーの GARP 通知をトリガとして、制御装置 10 が経路切替指示を発しているため、「VM 切替」の時間分、クライアントに対するサービスの停止時間が短縮される。

【0073】

本実施形態によれば、制御装置が、VM マイグレーションの終了前に、サーバから受信した、ダミーの GARP 通知をトリガとして、経路切替指示を送信している。その結果、制御装置が GARP 通知を受信してから経路切替指示を送信する従来技術に比べて、クライアントに対するサービスの停止時間を短縮することができる。サーバがダミーの GARP

10

20

30

40

50

P通知をVM切替動作の前に送信すれば、「VM切替」動作の時間分、サービス停止時間を短縮できる。

【0074】

なお、上述の実施形態では、サーバ21からサーバ22にVMを移行する場合で説明したが、はじめにサーバ22でVMを稼働させ、サーバ22からサーバ21にVMを移行してもよい。また、上述の実施形態では、サーバ21がサーバ22に比べて高負荷の場合で説明したので、管理装置30がサーバ21を監視する場合で説明したが、管理装置30は、サーバ21、22を均等に監視してもよく、また、サーバ22が低負荷または予備装置であることを予め認識している場合には、サーバ22を監視していなくてもよい。

【0075】

また、上述の実施形態において、サーバ22はGARP通知を制御装置10に送信する際、仮想ネットワークVLAN2にブロードキャストで送信してもよい。

【0076】

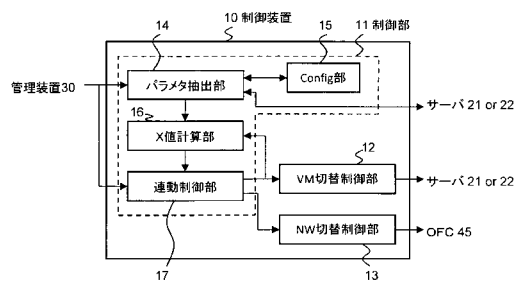
また、上述の実施形態では、サーバ21が高負荷になったときにVMマイグレーションを実施する場合で説明したが、サーバ21に故障が発生した場合など他の要因によってVMマイグレーションを実施する場合にも、本発明を適用してもよい。

【符号の説明】

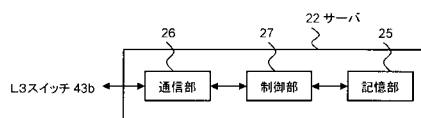
【0077】

- 10 制御装置
- 21、22 サーバ
- 30 管理装置
- 41 O F S
- 45 O F C
- 50 クライアント端末
- NW ネットワーク

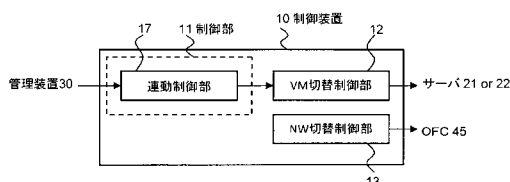
【図2】



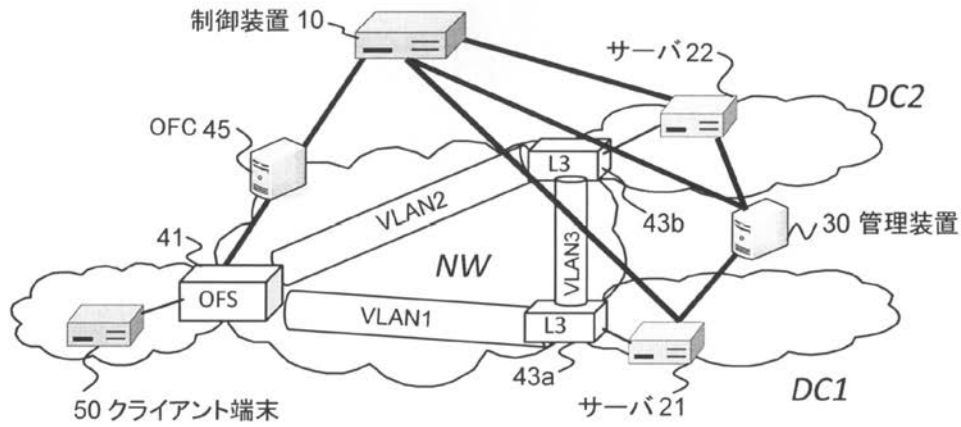
【図5】



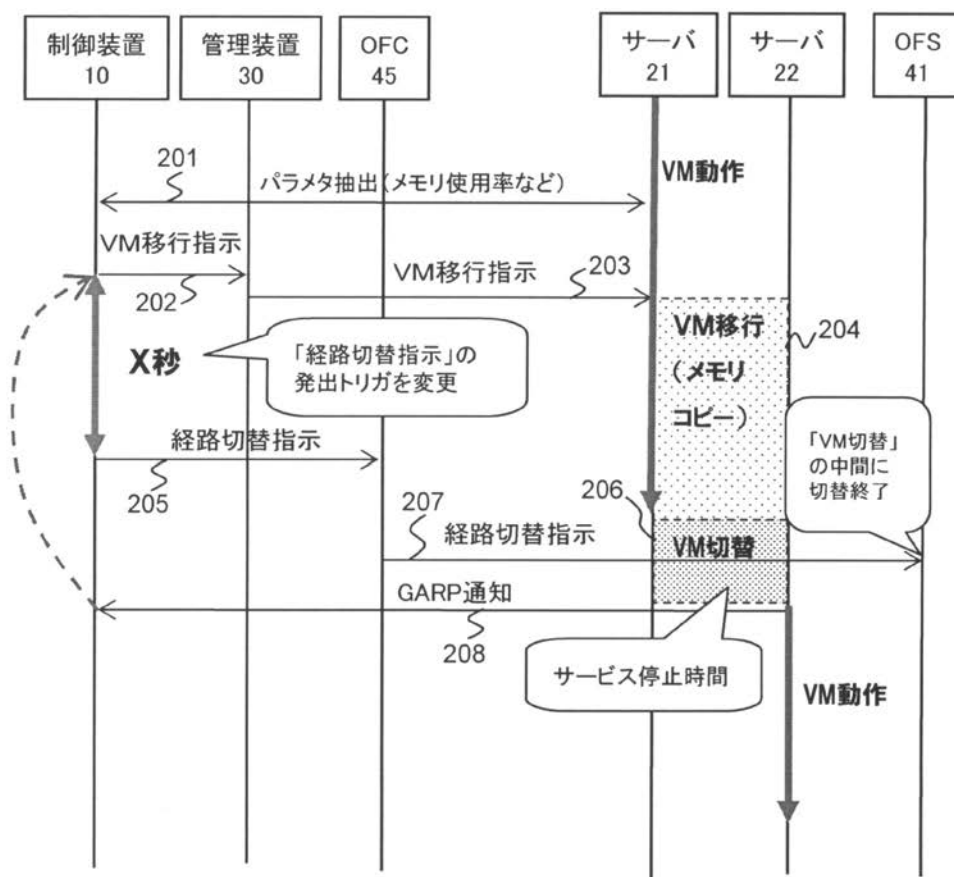
【図6】



【図 1】



【図 3】



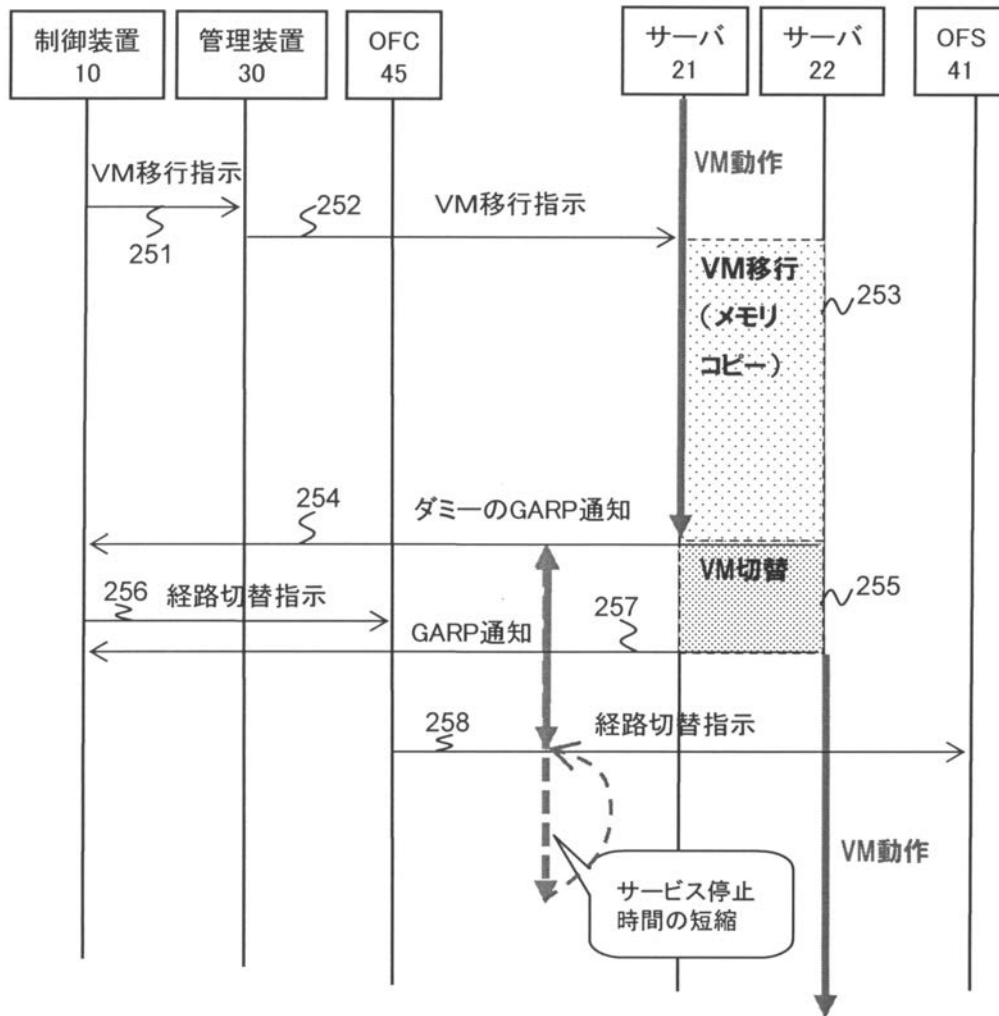
【図 4】

パラメタ等		パターン1	パターン2
パラメタ	C: メモリコピー帯域(Mb/sec)	100	1000
	M: サーバ21のメモリ容量(Mbyte)	2000	2000
	R: サーバ21のメモリ使用率	0.3	0.3
	V: VM切替時間(sec)	1.5	1.5
	P: 調整パラメタ	1	1
	N: ネットワーク切替時間	3	3
	T: VM移行指示からメモリコピー開始までの時間(sec)	0.3	0.3
時刻(*)	$X = (M \times R \times 8 \times P / C) + (V / 2) - N + T$	46.05	2.85
	NW切替時刻	49.05	5.85
	VM切替時刻	48.3~49.8	5.3~6.8

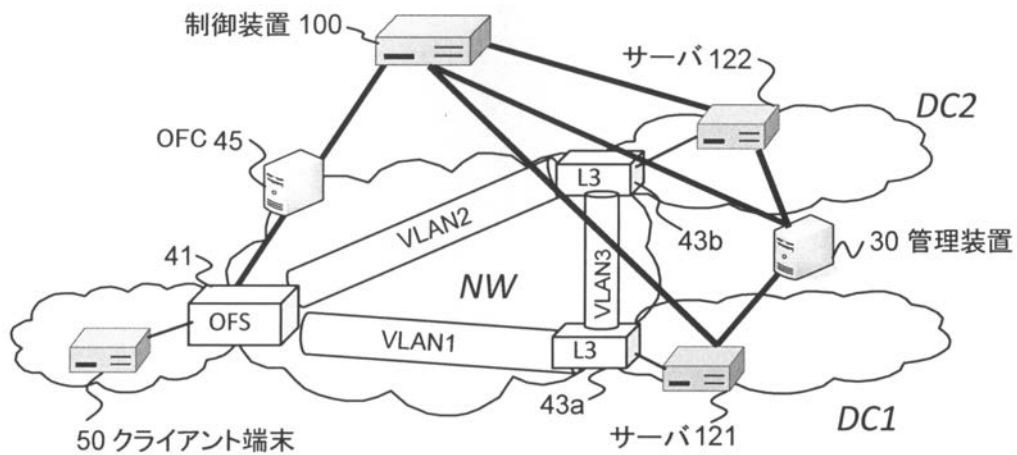
NW切替が、
VM切替中に
行われている。

(*)時刻: 「VM移行指示」を始点とした場合の時刻(sec)

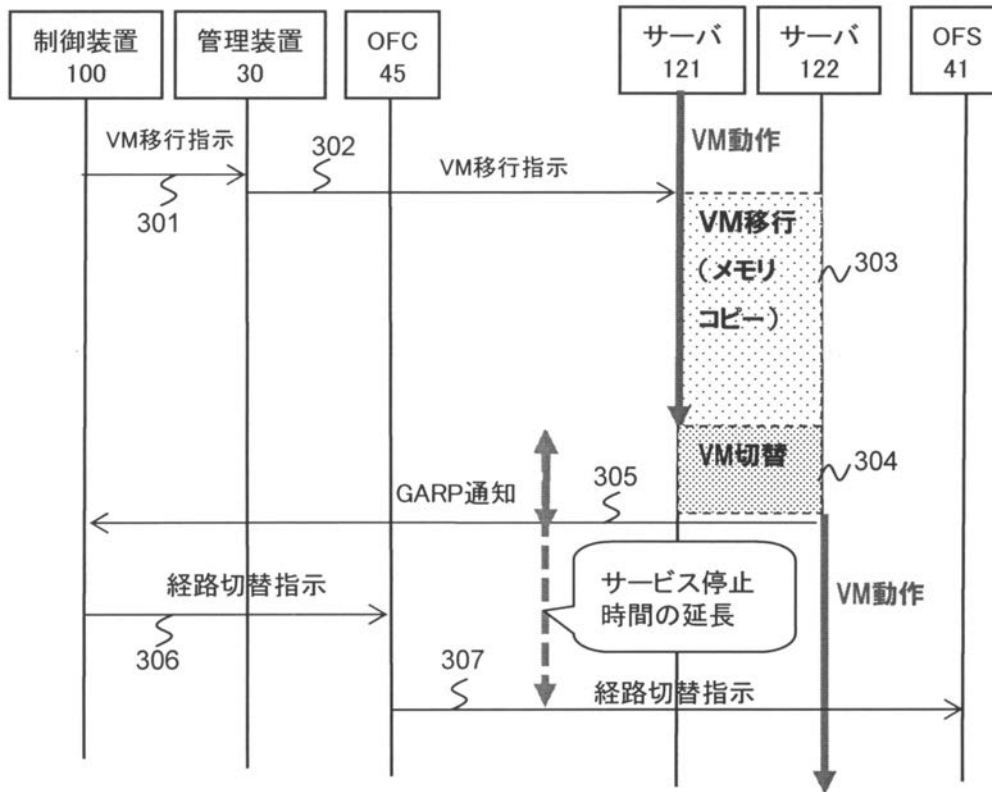
【図 7】



【图 8】



【図 9】



フロントページの続き

- (72)発明者 山崎 裕史
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
- (72)発明者 福岡 亜希
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
- (72)発明者 望月 このみ
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内