

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6079426号  
(P6079426)

(45) 発行日 平成29年2月15日(2017.2.15)

(24) 登録日 平成29年1月27日(2017.1.27)

(51) Int.Cl. F I  
H O 4 L 12/713 (2013.01) H O 4 L 12/713

請求項の数 4 (全 19 頁)

(21) 出願番号	特願2013-103968 (P2013-103968)	(73) 特許権者	000005223
(22) 出願日	平成25年5月16日(2013.5.16)		富士通株式会社
(65) 公開番号	特開2014-225774 (P2014-225774A)		神奈川県川崎市中原区上小田中4丁目1番1号
(43) 公開日	平成26年12月4日(2014.12.4)	(74) 代理人	100103528
審査請求日	平成28年2月26日(2016.2.26)		弁理士 原田 一男
		(72) 発明者	佐藤 昌浩
			神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
		審査官	大石 博見

最終頁に続く

(54) 【発明の名称】 情報処理システム、方法、装置及びプログラム

(57) 【特許請求の範囲】

【請求項1】

中継装置と、

前記中継装置に接続されている第1の情報処理装置及び第2の情報処理装置と、

を有し、

前記第2の情報処理装置は、

冗長構成を実現する複数の仮想マシンのうち現用の仮想マシンが前記第1の情報処理装置から前記第2の情報処理装置に移動してきたことを検出する第1処理部と、

前記現用の仮想マシンが第2の情報処理装置に移動してきたことが前記第1処理部により検出された場合、前記中継装置に、前記複数の仮想マシンが属するグループのアドレスを送信する第2処理部と、

を有し、

前記中継装置は、

アドレスとポートの識別情報とを対応付けて格納するデータ格納部と、

前記データ格納部における、受信した前記グループのアドレスに対応付けて格納されているポートの識別情報を、前記第1の情報処理装置が接続されているポートの識別情報から前記第2の情報処理装置が接続されているポートの識別情報に変更する第3処理部と、

を有する情報処理システム。

【請求項2】

前記第1処理部は、

10

20

前記現用の仮想マシンが送信する所定のメッセージの転送元が前記中継装置から前記第2の情報処理装置内の仮想マシンに変化したことを検出することによって、前記現用の仮想マシンが移動してきたことを検出する

ことを特徴とする請求項1記載の情報処理システム。

【請求項3】

冗長構成を実現する複数の仮想マシンのうち現用の仮想マシンが他の情報処理装置から移動してきたことを検出する第1処理部と、

前記現用の仮想マシンが移動してきたことが前記第1処理部により検出された場合、前記複数の仮想マシンが属するグループのアドレスを中継装置に送信する第2処理部と、  
を有する情報処理装置。

10

【請求項4】

冗長構成を実現する複数の仮想マシンのうち現用の仮想マシンが他の情報処理装置から移動してきたことを検出し、

前記現用の仮想マシンが移動してきたことが検出された場合、前記複数の仮想マシンが属するグループのアドレスを中継装置に送信する

処理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、冗長構成の制御技術に関する。

20

【背景技術】

【0002】

VRRP (Virtual Router Redundancy Protocol) は、ネットワークにおけるノードを冗長化するためのプロトコルである。VRRPにおいては、特定のVRRPグループに属する複数のノードのうち1つのノードをマスターノード(すなわち現用のノード)とし、残りのノードをバックアップノード(すなわち待機のノード)とする。通常はマスターノードのみが処理を行い、マスターノードに障害が発生するとバックアップノードのうちの一つが代わりに処理を行う。

【0003】

各VRRPグループには、1つの仮想MACアドレス及び仮想IPアドレスが割り当てられる。VRRPを利用するネットワークにおいては、この仮想MACアドレス及び仮想IPアドレス宛のデータがマスターノードに転送されるように、ネットワークにおけるスイッチ等の設定が行われる。

30

【0004】

ところで、VRRPは、物理マシンだけではなく仮想マシンに対しても適用される。そのため、VRRPグループにおけるマスターノードが仮想マシンである場合、マスターノードがライブマイグレーション(Live Migration)によって別のサーバ等に移動することがある。

【0005】

たとえマスターノードが別のサーバ等に移動したとしても、それに対する対処を行わない限りVRRPグループ宛のデータは元のサーバに転送される。しかし、マスターノードは元のサーバには存在しないため、通信障害が発生する。従来の技術においては、このような問題には着目されていない。

40

【先行技術文献】

【特許文献】

【0006】

【特許文献1】特表2010-518710号公報

【発明の概要】

【発明が解決しようとする課題】

【0007】

50

従って、本発明の目的は、1つの側面では、冗長構成を実現する複数の仮想マシンのうち現用の仮想マシンが移動する場合において、通信障害の発生を抑制するための技術を提供することである。

【課題を解決するための手段】

【0008】

本発明に係る情報処理システムは、中継装置と、中継装置に接続されている第1の情報処理装置及び第2の情報処理装置とを有する。そして、上で述べた第2の情報処理装置は、冗長構成を実現する複数の仮想マシンのうち現用の仮想マシンが第1の情報処理装置から第2の情報処理装置に移動してきたことを検出する第1処理部と、現用の仮想マシンが第2の情報処理装置に移動してきたことが第1処理部により検出された場合、中継装置に、複数の仮想マシンが属するグループのアドレスを送信する第2処理部とを有する。そして、上で述べた中継装置は、アドレスとポートの識別情報とを対応付けて格納するデータ格納部と、データ格納部における、受信したグループのアドレスに対応付けて格納されているポートの識別情報を、第1の情報処理装置が接続されているポートの識別情報から第2の情報処理装置が接続されているポートの識別情報に変更する第3処理部とを有する。

10

【発明の効果】

【0009】

冗長構成を実現する複数の仮想マシンのうち現用の仮想マシンが移動する場合において、通信障害の発生を抑制できるようになる。

【図面の簡単な説明】

20

【0010】

【図1】図1は、仮想環境におけるVRRPの設定について説明するための図である。

【図2】図2は、VMのライブマイグレーションについて説明するための図である。

【図3】図3は、VRRPグループの仮想MACアドレスが更新されないことによる問題について説明するための図である。

【図4】図4は、本実施の形態におけるシステム概要を示す図である。

【図5】図5は、仮想スイッチの機能ブロック図である。

【図6】図6は、レイヤ2スイッチの機能ブロック図である。

【図7】図7は、管理テーブル格納部に格納されるデータの一例を示す図である。

【図8】図8は、FDBに格納されるデータの一例を示す図である。

30

【図9】図9は、VRRP Advertisementメッセージの一例を示す図である。

【図10】図10は、VRRPセグメントを示す図である。

【図11】図11は、データ通信用セグメントを示す図である。

【図12】図12は、マスタノードの移動を検出する方法について説明するための図である。

【図13】図13は、マスタノードの移動を検出する方法について説明するための図である。

【図14】図14は、メインの処理フローを示す図である。

【図15】図15は、システムの動作の具体例を示す図である。

40

【図16】図16は、システムの動作の具体例を示す図である。

【図17】図17は、システムの動作の具体例を示す図である。

【図18】図18は、システムの動作の具体例を示す図である。

【図19】図19は、システムの動作の具体例を示す図である。

【図20】図20は、システムの動作の具体例を示す図である。

【図21】図21は、システムの動作の具体例を示す図である。

【図22】図22は、システムの動作の具体例を示す図である。

【図23】図23は、システムの動作の具体例を示す図である。

【図24】図24は、コンピュータの機能ブロック図である。

【発明を実施するための形態】

50

## 【 0 0 1 1 】

図 1 乃至図 3 を用いて、ライブマイグレーションによるマスタノードの移動について簡単に説明する。

## 【 0 0 1 2 】

まず、図 1 を用いて、仮想環境における VRRP の設定について説明する。図 1 の例においては、ホスト # 1 と、サーバ # 1 と、サーバ # 2 とが L (Layer) 2 スイッチに接続されている。サーバ # 1 及びサーバ # 2 においては、仮想スイッチが実行されている。サーバ # 1 においては、VM (Virtual Machine) # 1 及び VM # 2 が実行されている。VM # 1 及び VM # 2 は、サーバ # 1 における仮想スイッチに接続されている。ここで、VM # 1 及び VM # 2 は同じ VLAN (Virtual Local Area Network) (ここでは VLAN 1) に属し、VM # 1、VM # 2 及びホスト # 1 は同じ VLAN (ここでは VLAN 2) に属する。VLAN 1 は VRRP Advertisement メッセージを送受信するための VLAN であり、VLAN 2 はホスト # 1 と VM # 1 及び VM # 2 との間でデータを送受信するための VLAN である。

10

## 【 0 0 1 3 】

VM # 1 及び VM # 2 は同一の VRRP グループに属しており、VM # 1 の優先度が VM # 2 の優先度より高いため、VM # 1 がマスタノード (すなわち、現用のノード) であり、VM # 2 がバックアップノード (すなわち、待機のノード) である。VM # 1 及び VM # 2 が属する VRRP グループの仮想 MAC アドレスは「A」であり、仮想 IP アドレスは「1.1.1.1」である。このように、各 VRRP グループに対して 1 つの仮想 MAC アドレス及び仮想 IP アドレスが割り当てられているので、ホスト # 1 は、たとえマスタノードが変わったとしても宛先アドレスの変更をしなくてもよい。

20

## 【 0 0 1 4 】

図 2 を用いて、VM # 1 のライブマイグレーションについて説明する。図 2 において、VM # 1 はライブマイグレーションによってサーバ # 2 に移動している。VM # 1 が移動すると、サーバ # 2 の仮想スイッチは移動を検出し、VM # 1 の MAC アドレス「a」を含む ARP (Gratuitous Address Resolution Protocol) パケットを L 2 スイッチに送信する。L 2 スイッチは、送信元の MAC アドレスが「a」であるパケットをポート「P2」から受け取り、FDB (Forwarding DataBase) のデータを更新する。すなわち、FDB において MAC アドレス「a」に対応付けられている出力 IF (InterFace) (ここではポート) の識別子を「P1」から「P2」に変更する。なお、ライブマイグレーションとは、特定の物理サーバ等において稼働している VM を、OS (Operating System) 等のソフトウェアを停止させることなく別の物理サーバ等に移動させることである。

30

## 【 0 0 1 5 】

図 2 を用いて説明した処理は、従来より行われている処理である。このような処理を行えば、VM # 1 の MAC アドレス「a」については FDB に格納されている出力 IF を更新することができる。しかしながら、VM # 1 及び VM # 2 が属する VRRP グループの仮想 MAC アドレス「A」に対応付けられている出力 IF は更新されない。

## 【 0 0 1 6 】

図 3 を用いて、VRRP グループの仮想 MAC アドレスが更新されないことによる問題を説明する。図 3 においては、VRRP グループの仮想 MAC アドレス「A」及び仮想 IP アドレス「1.1.1.1」を宛先とするデータをホスト # 1 が L 2 スイッチに送信している。L 2 スイッチにおける FDB においては、仮想 MAC アドレス「A」を宛先とするデータを出力 IF 「P1」に出力することを示すデータが格納されているので、ホスト # 1 が送信したデータはポート「P1」を介してサーバ # 1 に送信される。しかしながら、VRRP グループのマスタノードである VM # 1 はサーバ # 1 には存在しないため、データに対する処理を行うことができない。

40

## 【 0 0 1 7 】

そこで、以下では、マスタノードがライブマイグレーションによって移動した場合にレイヤ 2 スイッチの設定を迅速に行うことで、通信障害が発生するのを抑制する方法につい

50

て説明する。

【0018】

図4に、本実施の形態におけるシステム概要を示す。レイヤ2スイッチ3には、物理サーバであるサーバ100と、物理サーバであるサーバ110と、物理サーバと通信を行うホスト5とが接続されている。サーバ100においては、VM101と、VM102と、仮想スイッチ103とが実行されている。仮想スイッチ103は、サーバ100内でデータの中継を行う。また、サーバ110においては、VM111と、VM112と、仮想スイッチ113とが実行されている。仮想スイッチ113は、サーバ110内でデータの中継を行う。図4において、サーバ100及びサーバ110において実行されるVMの数は2であるが、数に限定は無い。

10

【0019】

図5に、仮想スイッチ103及び113の機能ブロック図を示す。仮想スイッチ103及び113においては、I/O(Input/Output)インタフェース1001と、ブリッジング部1002と、監視部1003と、スヌーピング部1004と、GARP管理部1006とが実行される。また、例えばサーバ100及び110におけるハードディスク等の記憶装置には、管理テーブル格納部1005のための領域が確保される。

【0020】

I/Oインタフェース1001は、データの入出力を行うためのインタフェース(例えば仮想ポート)である。ブリッジング部1002は、仮想スイッチ内においてデータの中継を行う。監視部1003は、同じサーバ内にマスタノードであるVMが移動してきたことを検出する。また、VRRP Advertisementメッセージを受信したポートの識別子を管理テーブル格納部1005に格納する。スヌーピング部1004は、VRRP Advertisementメッセージから仮想MACアドレス及び仮想IPアドレスを取得し、管理テーブル格納部1005に格納する。GARP管理部1006は、マスタノードが同じサーバに移動してきた場合に、VRRPグループの仮想MACアドレス及び仮想IPアドレスを含むGARPの packets をレイヤ2スイッチ3に送信する。

20

【0021】

図6に、レイヤ2スイッチ3の機能ブロック図を示す。レイヤ2スイッチ3は、処理部31と、FDB32とを含む。処理部31は、FDB32に格納されているデータを管理する処理等を実行する。

30

【0022】

図7に、仮想スイッチ103及び113における管理テーブル格納部1005に格納されるデータの一例を示す。図7の例では、仮想MACアドレス及び仮想IPアドレスと、接続IFの識別子(ここでは、ポートの識別子)とが格納される。

【0023】

図8に、レイヤ2スイッチ3のFDB32に格納されるデータの一例を示す。図8の例では、宛先アドレスと、出力IFの識別子(ここでは、ポートの識別子)とが格納される。図8に示したようなデータは、VLAN毎に用意される。

【0024】

図9に、VRRPにおけるVRRP Advertisementメッセージの一例を示す。図9の例では、VRRP Advertisementメッセージには、宛先MACアドレス及び送信元MACアドレスを含むイーサネット(登録商標)ヘッダと、宛先IPアドレス及び送信元IPアドレスを含むIPヘッダと、優先度と仮想IPアドレスとを含むデータとが含まれる。図9に示すように、宛先には専用のマルチキャストアドレス(IPアドレスは「224.0.0.18」、MACアドレスは「01:00:5e:00:00:12」)が設定される。送信元MACアドレスは、VRRPグループに割り当てられた仮想MACアドレスである。送信元IPアドレスは、送信元のVMの送信ポートに割り当てられた物理IPアドレスである。データには、優先度及びVRRPグループに割り当てられた仮想IPアドレスが含まれる。

40

【0025】

50

図10及び図11に、本実施の形態におけるシステムの論理ネットワークを示す。まず、図10に、システムにおけるVRRPセグメント(本実施の形態においては、VLAN1とする)を示す。VRRPセグメントには、サーバ100と、サーバ110とが属する。VRRPセグメントは、VRRP Advertisementメッセージが通過するネットワークである。

【0026】

一方、図11に示すように、データ通信用セグメントには、サーバ100と、サーバ110と、ホスト5とが属する。データ通信用セグメントを、VLAN2とする。データ通信用セグメントは、ホストとサーバとの間における通信データが通過するネットワークである。

10

【0027】

このように、VRRPセグメントとデータ通信用セグメントとを分離して構築する理由は、VRRP Advertisementメッセージがホストに到達することを防ぐためである。

【0028】

図12及び図13を用いて、マスタノードの移動を検出する方法について説明する。まず、図12に、移動前の状態の一例を示す。マスタノードであるVM101はサーバ100に存在し、仮想スイッチのポート「p1」に接続されている。マスタノードがVRRP Advertisementメッセージをマルチキャストで送信すると、VRRP AdvertisementメッセージはバックアップノードであるVM102に到達すると共に、サーバ100内の仮想スイッチを経由してサーバ110内の仮想スイッチまで到達する。ここで、サーバ100内の仮想スイッチは、管理テーブル格納部1005に、VRRP Advertisementメッセージの受信ポートの識別子「p1」を登録する。また、VRRP Advertisementメッセージをスヌーピングし、仮想MACアドレス及び仮想IPアドレスを管理テーブル格納部1005に登録する。サーバ110内の仮想スイッチも、管理テーブル格納部1005への登録を行う。図12の例の場合、VRRP Advertisementメッセージの受信ポートの識別子は「p4」である。

20

【0029】

図13に、移動後の状態を示す。サーバ100内のマスタノードはライブマイグレーションによってサーバ110に移動する。移動完了後、マスタノードはVRRP Advertisementメッセージを送信する。VRRP Advertisementメッセージはサーバ110内の仮想スイッチを通過し、サーバ100内の仮想スイッチに到達する。VRRP Advertisementメッセージはバックアップノードに転送される。各仮想スイッチが受信するVRRP Advertisementメッセージの受信ポートは、移動前と移動後とで異なるため、管理テーブル格納部1005に格納されているデータが更新される。サーバ100内の仮想スイッチは「p3」を登録し、サーバ110内の仮想スイッチは「p5」を登録する。また、更新後のポートがL2スイッチ側ではなくVM側であるか否かを確認することで、同じ物理サーバ内にマスタノードが移動してきたか否かを判断する。サーバ100においては、ポート「p3」はVMに接続されていないため、同じサーバ以外にマスタノードが移動したことを検出する。サーバ110においては、ポート「p5」はVMに接続されているため、同じサーバ内にマスタノードが移動してきたことを検出する。

30

40

【0030】

次に、図14を用いて、本実施の形態におけるシステムの動作を詳細に説明する。まず、仮想スイッチにおけるI/Oインタフェース1001は、パケットを受信する(図14:ステップS1)。I/Oインタフェース1001は、受信したパケットをブリッジング部1002に出力する。ブリッジング部1002は、I/Oインタフェース1001から受け取ったパケットを監視部1003に出力する。

【0031】

50

監視部1003は、宛先アドレスが所定のマルチキャストアドレスであるか判断する(ステップS3)。所定のマルチキャストアドレスとは、VRRP Advertisementメッセージの宛先アドレスに設定される専用のマルチキャストアドレスである。上で述べたように、IPアドレスは「224.0.0.18」であり、MACアドレスは「01:00:5e:00:00:12」である。

**【0032】**

宛先アドレスが所定のマルチキャストアドレスではない場合(ステップS3:Noルート)、VRRP Advertisementメッセージではないため、処理を終了する。一方、宛先アドレスが所定のマルチキャストアドレスである場合(ステップS3:Yesルート)、監視部1003は、パケットの受信ポートが前回の受信ポートと同じであるか判断する(ステップS5)。前回の受信ポートの識別子は、管理テーブル格納部1005に格納されている。

10

**【0033】**

前回の受信ポートと同じである場合(ステップS5:Yesルート)、VMは移動していないので、処理を終了する。一方、前回の受信ポートと同じではない場合(ステップS5:Noルート)、監視部1003は、受信ポートの識別子を管理テーブル格納部1005に格納する(ステップS7)。すなわち、ステップS7においては、既に格納されている識別子を、今回の受信ポートの識別子で更新する。監視部1003は、パケットをスヌーピング部1004に出力する。

**【0034】**

スヌーピング部1004は、パケットから仮想MACアドレス及び仮想IPアドレスを取得する(ステップS9)。そして、スヌーピング部1004は、取得された仮想MACアドレス及び仮想IPアドレスを管理テーブル格納部1005に格納する(ステップS11)。ステップS11においては、ステップS7において格納された受信ポートの識別子に対応付けて仮想MACアドレス及び仮想IPアドレスを格納する。

20

**【0035】**

監視部1003は、今回の受信ポートはVMが接続されるポートであるか判断する(ステップS13)。仮想スイッチは各ポートの接続先を管理する情報を保持しているため、監視部1003は、その情報を用いてステップS13の処理を行う。

**【0036】**

受信ポートはVMが接続されるポートではない場合(ステップS13:Noルート)、仮想スイッチが実行されているサーバと同じサーバにVMが移動してきたわけではないので、処理を終了する。一方、受信ポートはVMが接続されるポートである場合(ステップS13:Yesルート)、監視部1003は、GARP管理部1006に処理の実行を要求する。GARP管理部1006は、パケットから取得された仮想MACアドレス及び仮想IPアドレスを含むGARPパケットを生成し、レイヤ2スイッチ3に送信する。具体的には、生成されたGARPパケットをブリッジング部1002に出力し、ブリッジング部1002はそのGARPパケットをI/Oインタフェース1001に出力する。I/Oインタフェース1001は、ブリッジング部1002から受け取ったGARPパケットをレイヤ2スイッチ3に送信する(ステップS15)。そして処理を終了する。

30

40

**【0037】**

なお、レイヤ2スイッチ3における処理部31は、FDB32においてGARPパケットに含まれる仮想MACアドレスに対応付けて格納されているポートの識別子を、今回GARPパケットを受信したポートの識別子に変更する。

**【0038】**

以上のような処理を実行すれば、マスタノードであるVMが移動した後VRRP Advertisementメッセージを送信したタイミングにおいて、レイヤ2スイッチ3のFDB32が更新される。これによって、VRRPグループ宛のデータが元のサーバに送信されてしまうことを抑制できるので、通信障害の発生を抑制することにつながる。

**【0039】**

50

次に、図15乃至図23を用いて、本実施の形態におけるシステムの動作を具体的に説明する。図15に、ネットワークトポロジを示す。図15においては、ホスト#1と、サーバ#1と、サーバ#2とがL2スイッチに接続されている。サーバ#1及びサーバ#2は仮想化環境として使用される。サーバ#1においてはVM#1とVM#2とが実行されている。VM#1の仮想NIC(Network Interface Card)に割り当てられるMACアドレスを「a」、VM#2の仮想NICに割り当てられるMACアドレスを「b」とする。ホスト#1の物理NICのMACアドレスは「c」とする。

【0040】

VM#1とVM#2との間にVRRPを設定し、VM#1をマスタノードとする。VRRPで使用する仮想IPアドレスは「1.1.1.1」とし、仮想MACアドレスは「A」とする。VRRP Advertisementメッセージが通過するセグメントをVLAN1とし、ホスト#1とVM#1及びVM#2との間において通信データが通過するセグメントをVLAN2とする。この段階において、仮想スイッチにおける管理テーブル格納部1005及びL2スイッチにおけるFDB32にはデータが格納されていない。

【0041】

図16に、管理テーブル格納部1005におけるエントリを生成する動作を示す。VRRPが設定されると、マスタノードからVRRP Advertisementメッセージがマルチキャストで定期的送信される。VRRP Advertisementメッセージを受信した、サーバ#1の仮想スイッチは、受信ポートが前回の受信ポートと異なる場合に、管理テーブル格納部1005にエントリを生成する。具体的には、サーバ#1の仮想スイッチは、管理テーブル格納部1005における接続IFにVRRP Advertisementメッセージの受信ポートの識別子「p1」を登録する。また、VRRP Advertisementメッセージをスヌーピングし、「A」及び「1.1.1.1」を仮想MACアドレス及び仮想IPアドレスとして管理テーブル格納部1005に登録する。一方、サーバ#2の仮想スイッチは、管理テーブル格納部1005における接続IFにVRRP Advertisementメッセージの受信ポートの識別子「p4」を登録する。また、VRRP Advertisementメッセージをスヌーピングし、「A」及び「1.1.1.1」を仮想MACアドレス及び仮想IPアドレスとして管理テーブル格納部1005に登録する。

【0042】

L2スイッチは、VRRP Advertisementメッセージを受信すると、送信元MACアドレスである「A」と受信ポートの識別子である「P1」とを対応付けてFDB32に登録する。VRRP AdvertisementメッセージはVLAN1のセグメントを通過するため、FDB32におけるVLAN1用のデータを更新する。

【0043】

図17に、ホスト#1がARP要求を送信した場合における動作を示す。ホスト#1は、VRRPグループの仮想IPアドレスである「1.1.1.1」に対応するMACアドレスを調べるため、ARP要求をブロードキャストで送信する。ARP要求を受信したレイヤ2スイッチは、ARP要求をサーバ#1に転送し、サーバ#1における仮想スイッチは、VM#1及びVM#2に対してARP要求を送信する。ARP要求を受信したVMのうちマスタノードであるVM#1は、ARP応答を送信する。

【0044】

L2スイッチは、ARP要求については送信元MACアドレスである「c」と受信ポートの識別子である「P3」とを対応付けてFDB32に登録し、ARP応答については送信元MACアドレスである「A」と受信ポートの識別子である「P1」とを対応付けてFDB32に登録する。なお、ARP要求及びARP応答はVLAN2のセグメントを通過するため、FDB32におけるVLAN2用のデータを更新する。

【0045】

ARPの処理が完了すると、図18で示すように、ホスト#1とマスタノードであるVM#1との間ではユニキャストで通信が行われる。すなわち、ホスト#1が宛先MACア

10

20

30

40

50



ドレスを「A」とし宛先IPアドレスを「1.1.1.1」とするユニキャストのデータを送信する。L2スイッチは、ユニキャストのデータを受信すると、FDB32におけるエントリに従い、ユニキャストのデータをポート「P1」から出力する。サーバ#1は、ユニキャストのデータを受信すると、管理テーブル格納部1005に格納されているデータに従い、ユニキャストのデータをポート「p1」から出力する。そして、VM#1がユニキャストのデータを受信する。なお、VM#1がホスト#1にユニキャストのデータを送信する場合にも同様の中継が行われる。

**【0046】**

図19に、マスタノードがライブマイグレーションによってサーバ#2に移動した場合における動作を示す。マスタノードがサーバ#2に移動すると、サーバ#2の仮想スイッチは、VM#1の仮想NICのMACアドレスである「a」を含むGARPパケットを、VLAN1及びVLAN2に対して送信する。GARPパケットを受信したL2スイッチは、FDB32におけるVLAN1及びVLAN2用のデータの各々について「a」を含むエントリを更新する。具体的には、VLAN1については出力IFが「P1」から「P2」に変更され、VLAN2については新規にエントリが追加される。

10

**【0047】**

図20に、マスタノード移動後におけるVRRP Advertisementメッセージの転送を示す。ライブマイグレーション完了後、VRRP Advertisementメッセージはサーバ#2におけるマスタノードから送信され、サーバ#2における仮想スイッチを通過し、サーバ#1における仮想スイッチに到達する。サーバ#1の仮想スイッチに到達したVRRP Advertisementメッセージは、バックアップノードに転送される。

20

**【0048】**

サーバ#2の管理テーブル格納部1005における接続IFは「p4」から「p5」に変更され、サーバ#1の管理テーブル格納部1005における接続IFは「p1」から「p3」に変更される。L2スイッチは、VRRP Advertisementメッセージを受信した後、FDB32におけるVLAN1用のデータにおいて、「A」を含むエントリを更新する。具体的には、出力IFが「P1」から「P2」に変更される。

**【0049】**

サーバ#2の管理テーブル格納部1005における接続IFは「p5」に変更されるが、「p5」の接続先はVMであるため、仮想スイッチは、仮想MACアドレスである「A」を含むGARPパケットをL2スイッチに送信する。このGARPパケットは、VLAN2に対してのみ送信される。GARPパケットを受信したL2スイッチは、FDB32におけるVLAN2用のデータについて、「A」を含むエントリを更新する。具体的には、出力IFが「P1」から「P2」に変更される。

30

**【0050】**

図21に、ライブマイグレーション後におけるデータ通信を示す。ホスト#1が宛先MACアドレスを「A」とし宛先IPアドレスを「1.1.1.1」とするユニキャストのデータを送信する。L2スイッチは、ユニキャストのデータを受信すると、FDB32におけるエントリに従い、ユニキャストのデータをポート「P2」から出力する。サーバ#2における仮想スイッチは、ユニキャストのデータを受信すると、管理テーブル格納部1005に格納されているデータに従い、ユニキャストのデータをポート「p5」から出力する。そして、VM#1がユニキャストのデータを受信する。なお、VM#1がホスト#1にユニキャストのデータを送信する場合にも同様の中継が行われる。

40

**【0051】**

図22に、図18の状態においてバックアップノードがライブマイグレーションによってサーバ#2に移動した場合における動作を示す。バックアップノードがサーバ#2に移動すると、仮想スイッチは、VM#2の仮想NICのMACアドレス「b」を含むGARPパケットをVLAN1及びVLAN2に対して送信する。GARPパケットを受信したL2スイッチは、FDB32におけるVLAN1用のデータ及びVLAN2用のデータに

50

ついて「b」を含むエントリを更新する。具体的には、VLAN1及びVLAN2について新規のエントリが追加される。

【0052】

図23に、バックアップノード移動後におけるVRRP Advertisementメッセージの転送を示す。ライブマイグレーション完了後、VRRP Advertisementメッセージはサーバ#1におけるマスタノードから送信され、サーバ#1における仮想スイッチを通過し、サーバ#2における仮想スイッチに到達する。サーバ#2の仮想スイッチに到達したVRRP Advertisementメッセージは、バックアップノードに転送される。但し、マスタノードが移動したわけではないため、各仮想スイッチにおける管理テーブル格納部1005のデータは更新されない。

10

【0053】

以上のような処理を実行すれば、ライブマイグレーションによってマスタノードの配置が変更されたとしても迅速にFDB32を更新できるので、通信障害の発生を抑制できるようになる。

【0054】

以上本発明の一実施の形態を説明したが、本発明はこれに限定されるものではない。例えば、上で説明したレイヤ2スイッチ3並びにサーバ100及び110の機能ブロック構成は実際のプログラムモジュール構成に一致しない場合もある。

【0055】

また、上で説明した各テーブルの構成は一例であって、上記のような構成でなければならぬわけではない。さらに、処理フローにおいても、処理結果が変わらなければ処理の順番を入れ替えることも可能である。さらに、並列に実行させるようにしても良い。

20

【0056】

例えば、ステップS15においてはGARPのパケットを送信する例を説明したが、GARPのパケットに限られるわけではない。VRRPグループの仮想MACアドレスを含むパケットを送信し、レイヤ2スイッチ3がそのパケットの受信に応じてFDB32を更新してもよい。

【0057】

なお、上で述べたサーバ100及び110は、コンピュータ装置であって、図24に示すように、メモリ2501とCPU(Central Processing Unit)2503とハードディスク・ドライブ(HDD:Hard Disk Drive)2505と表示装置2509に接続される表示制御部2507とリムーバブル・ディスク2511用のドライブ装置2513と入力装置2515とネットワークに接続するための通信制御部2517とがバス2519で接続されている。オペレーティング・システム(OS:Operating System)及び本実施例における処理を実施するためのアプリケーション・プログラムは、HDD2505に格納されており、CPU2503により実行される際にはHDD2505からメモリ2501に読み出される。CPU2503は、アプリケーション・プログラムの処理内容に応じて表示制御部2507、通信制御部2517、ドライブ装置2513を制御して、所定の動作を行わせる。また、処理途中のデータについては、主としてメモリ2501に格納されるが、HDD2505に格納されるようにしてもよい。本発明の実施例では、上で述べた処理を実施するためのアプリケーション・プログラムはコンピュータ読み取り可能なリムーバブル・ディスク2511に格納されて頒布され、ドライブ装置2513からHDD2505にインストールされる。インターネットなどのネットワーク及び通信制御部2517を経由して、HDD2505にインストールされる場合もある。このようなコンピュータ装置は、上で述べたCPU2503、メモリ2501などのハードウェアとOS及びアプリケーション・プログラムなどのプログラムとが有機的に協働することにより、上で述べたような各種機能を実現する。

30

40

【0058】

以上述べた本発明の実施の形態をまとめると、以下ようになる。

【0059】

50

本実施の形態に係る情報処理システムは、(A)中継装置と、(B)中継装置に接続されている第1の情報処理装置及び第2の情報処理装置とを有する。そして、上で述べた第2の情報処理装置は、(b1)冗長構成を実現する複数の仮想マシンのうち現用の仮想マシンが第1の情報処理装置から第2の情報処理装置に移動してきたことを検出する第1処理部と、(b2)現用の仮想マシンが第2の情報処理装置に移動してきたことが第1処理部により検出された場合、中継装置に、複数の仮想マシンが属するグループのアドレスを送信する第2処理部とを有する。そして、上で述べた中継装置は、(a1)アドレスとポートの識別情報とを対応付けて格納するデータ格納部と、(a2)データ格納部における、受信したグループのアドレスに対応付けて格納されているポートの識別情報を、第1の情報処理装置が接続されているポートの識別情報から第2の情報処理装置が接続されているポートの識別情報に変更する第3処理部とを有する。

10

**【0060】**

このようにすれば、現用の仮想マシンが移動した場合において、グループ宛のデータが移動後の現用の仮想マシンに中継されるように迅速に中継装置を設定できるので、通信障害の発生を抑制できるようになる。

**【0061】**

また、上で述べた第1処理部は、(b11)現用の仮想マシンが送信する所定のメッセージの転送元が中継装置から第2の情報処理装置内の仮想マシンに変化したことを検出することによって、現用の仮想マシンが移動してきたことを検出してもよい。このようにすれば、上記のグループが、現用系の仮想マシンが所定のメッセージを送ることを特徴とするグループ(例えばVRRPグループ)である場合に、現用の仮想マシンの移動を効率的に検出できるようになる。

20

**【0062】**

また、上で述べた第2処理部は、(b21)複数の仮想マシンが属するグループのアドレスを含むアドレス解決プロトコルのパケットを中継装置に送信してもよい。上記のような情報処理システムにおいてアドレス解決プロトコルは通常使用されるプロトコルであるので、アドレス解決プロトコルを利用すれば、情報処理システムに対する改変を少なくできるようになる。

**【0063】**

また、上で述べたグループのアドレスは、グループの仮想MACアドレスを含んでもよい。このようにすれば、中継装置が例えばレイヤ2のスイッチである場合においても、設定を適切に行えるようになる。

30

**【0064】**

また、上で述べた所定のメッセージは、VRRP(Virtual Router Redundancy Protocol)におけるAdvertisementメッセージであってもよい。VRRPにおけるAdvertisementメッセージは定期的には送信されるので、Advertisementメッセージを利用すれば、上記のグループがVRRPグループである場合において確実に設定の変更を行えるようになる。

**【0065】**

なお、上記方法による処理をコンピュータに行わせるためのプログラムを作成することができ、当該プログラムは、例えばフレキシブルディスク、CD-ROM、光磁気ディスク、半導体メモリ、ハードディスク等のコンピュータ読み取り可能な記憶媒体又は記憶装置に格納される。尚、中間的な処理結果はメインメモリ等の記憶装置に一時保管される。

40

**【0066】**

以上の実施例を含む実施形態に関し、さらに以下の付記を開示する。

**【0067】**

(付記1)

中継装置と、

前記中継装置に接続されている第1の情報処理装置及び第2の情報処理装置と、

を有し、

50

前記第 2 の情報処理装置は、  
 冗長構成を実現する複数の仮想マシンのうち現用の仮想マシンが前記第 1 の情報処理装置から前記第 2 の情報処理装置に移動してきたことを検出する第 1 処理部と、  
 前記現用の仮想マシンが第 2 の情報処理装置に移動してきたことが前記第 1 処理部により検出された場合、前記中継装置に、前記複数の仮想マシンが属するグループのアドレスを送信する第 2 処理部と、  
 を有し、  
 前記中継装置は、  
 アドレスとポートの識別情報とを対応付けて格納するデータ格納部と、  
 前記データ格納部における、受信した前記グループのアドレスに対応付けて格納されているポートの識別情報を、前記第 1 の情報処理装置が接続されているポートの識別情報から前記第 2 の情報処理装置が接続されているポートの識別情報に変更する第 3 処理部と、  
 を有する情報処理システム。

10

【 0 0 6 8 】

(付記 2)

前記第 1 処理部は、  
 前記現用の仮想マシンが送信する所定のメッセージの転送元が前記中継装置から前記第 2 の情報処理装置内の仮想マシンに変化したことを検出することによって、前記現用の仮想マシンが移動してきたことを検出することを特徴とする付記 1 記載の情報処理システム。

20

【 0 0 6 9 】

(付記 3)

前記第 2 処理部は、  
 前記複数の仮想マシンが属するグループのアドレスを含むアドレス解決プロトコルのパケットを前記中継装置に送信することを特徴とする付記 1 又は 2 記載の情報処理システム。

【 0 0 7 0 】

(付記 4)

前記グループのアドレスは、前記グループの仮想 MAC アドレスを含むことを特徴とする付記 1 乃至 3 のいずれか 1 つ記載の情報処理システム。

30

【 0 0 7 1 】

(付記 5)

前記所定のメッセージは、VRRP (Virtual Router Redundancy Protocol) における Advertisement メッセージであることを特徴とする付記 2 記載の情報処理システム。

【 0 0 7 2 】

(付記 6)

冗長構成を実現する複数の仮想マシンのうち現用の仮想マシンが他の情報処理装置から移動してきたことを検出する第 1 処理部と、  
 前記現用の仮想マシンが移動してきたことが前記第 1 処理部により検出された場合、前記複数の仮想マシンが属するグループのアドレスを中継装置に送信する第 2 処理部と、  
 を有する情報処理装置。

40

【 0 0 7 3 】

(付記 7)

冗長構成を実現する複数の仮想マシンのうち現用の仮想マシンが他の情報処理装置から移動してきたことを検出し、  
 前記現用の仮想マシンが移動してきたことが検出された場合、前記複数の仮想マシンが属するグループのアドレスを中継装置に送信する処理をコンピュータに実行させるためのプログラム。

【 0 0 7 4 】

50

(付記 8)

冗長構成を実現する複数の仮想マシンのうち現用の仮想マシンが他の情報処理装置から移動してきたことを検出し、

前記現用の仮想マシンが移動してきたことが検出された場合、前記複数の仮想マシンが属するグループのアドレスを中継装置に送信する

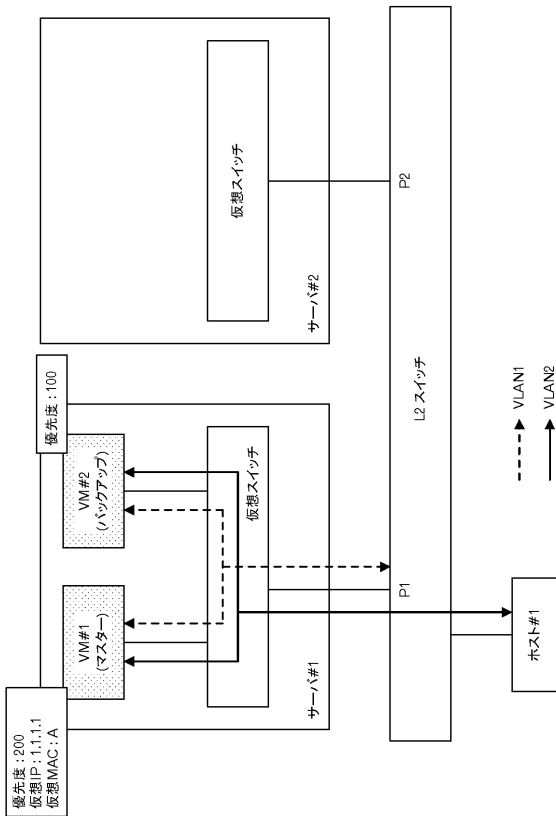
処理をコンピュータが実行する情報処理方法。

【符号の説明】

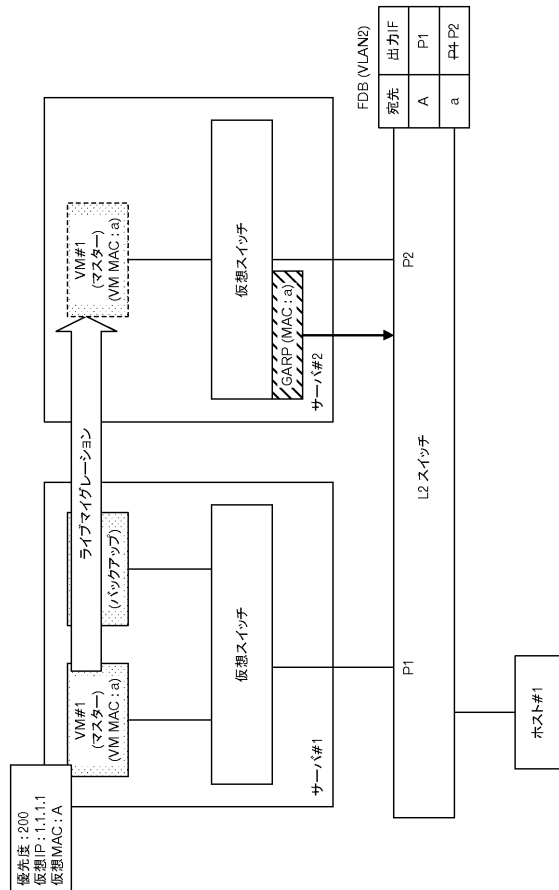
【0075】

- 100, 110 サーバ 101, 102, 111, 112 VM
- 103, 113 仮想スイッチ 3 レイヤ2スイッチ
- 5 ホスト 1001 I/Oインタフェース
- 1002 ブリッジング部 1003 監視部
- 1004 スヌーピング部 1005 管理テーブル格納部
- 1006 GARP管理部 31 処理部
- 32 FDB

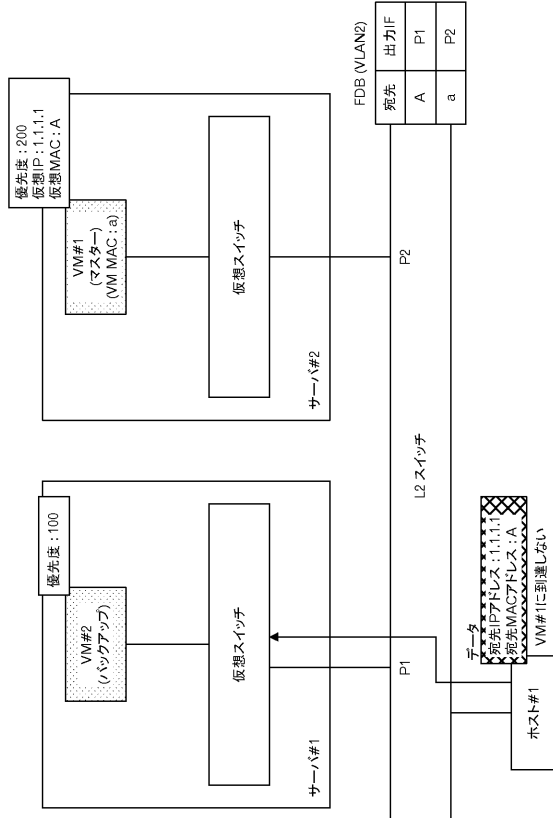
【図 1】



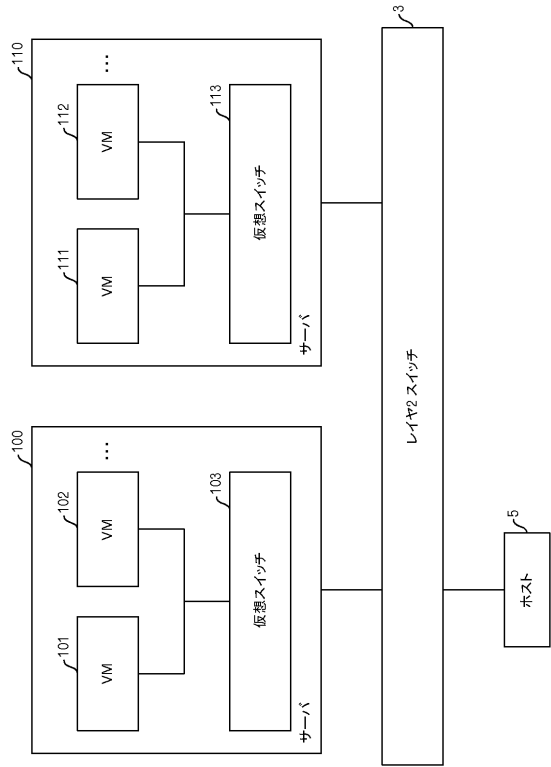
【図 2】



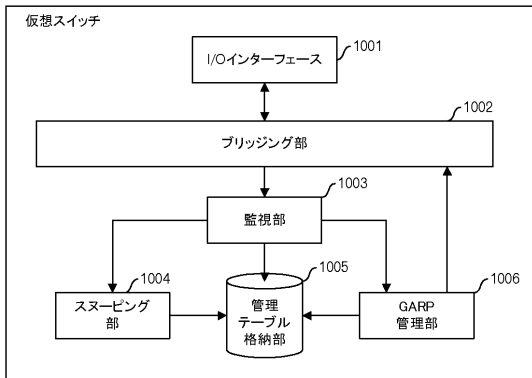
【図3】



【図4】



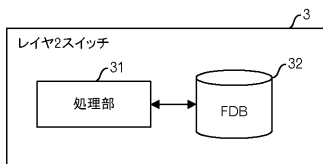
【図5】



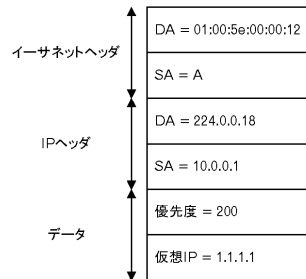
【図8】

宛先アドレス	出力IF
B	P10
b	P12

【図6】



【図9】

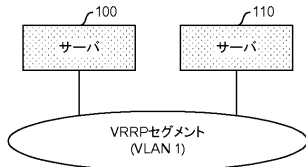


【図7】

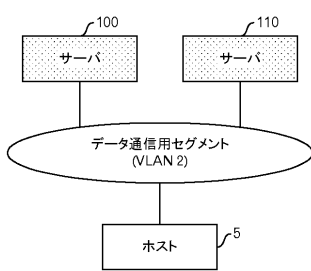
仮想MACアドレス/仮想IPアドレス	接続IF
B / 1.1.1.2	p10

- イーサネットヘッダ
  - DA : 宛先MACアドレス
  - SA : 送信元MACアドレス
- IPヘッダ
  - DA : 宛先IPアドレス
  - SA : 送信元IPアドレス

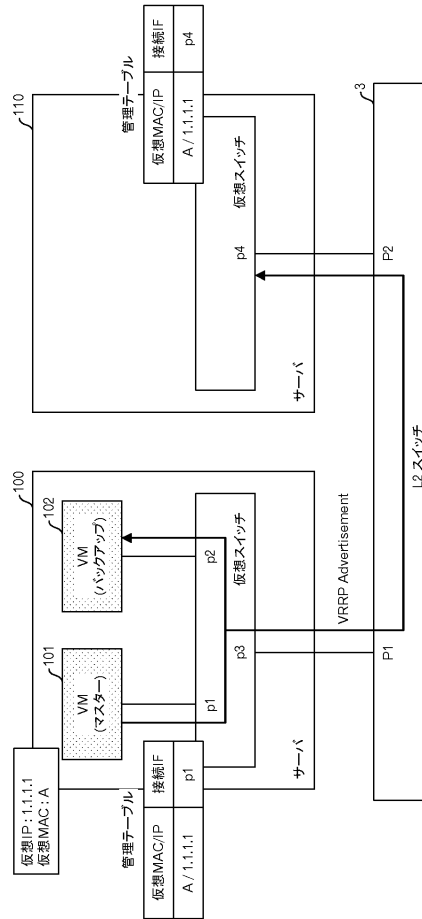
【図10】



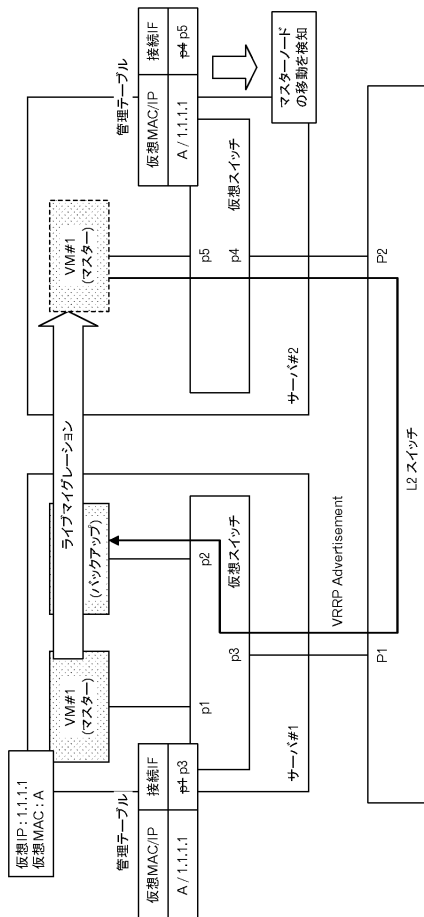
【図11】



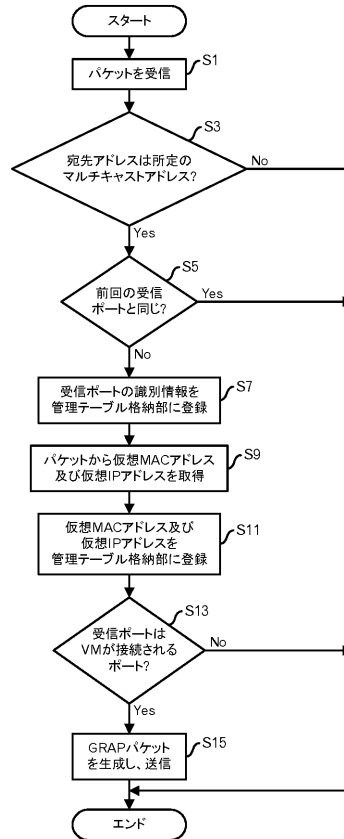
【図12】



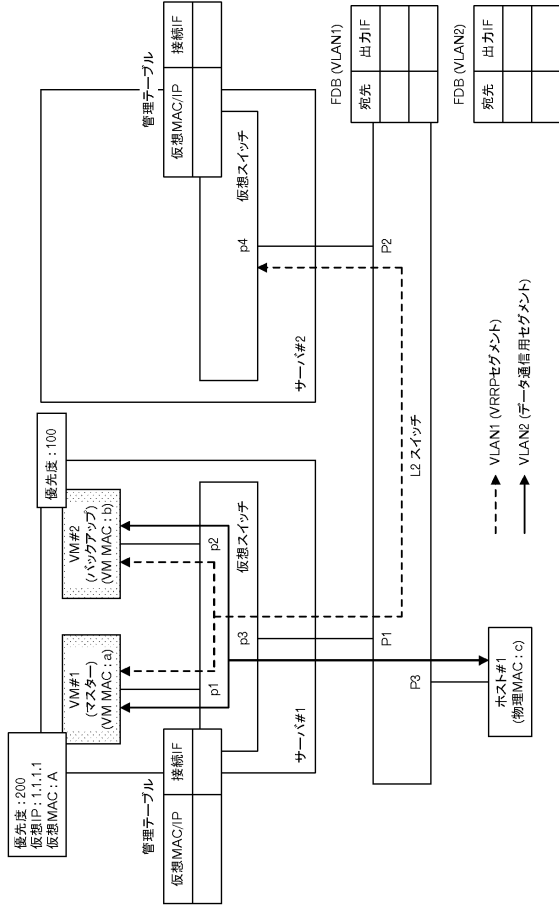
【図13】



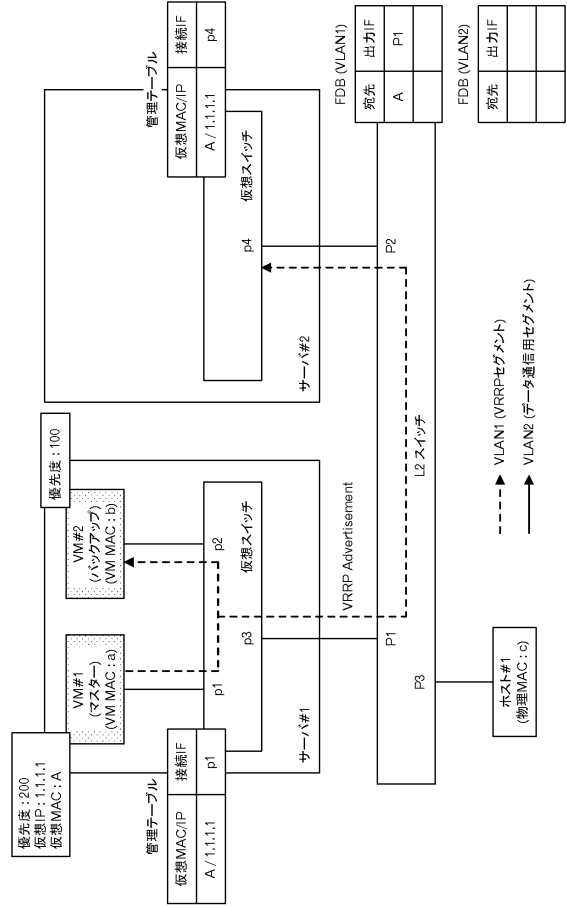
【図14】



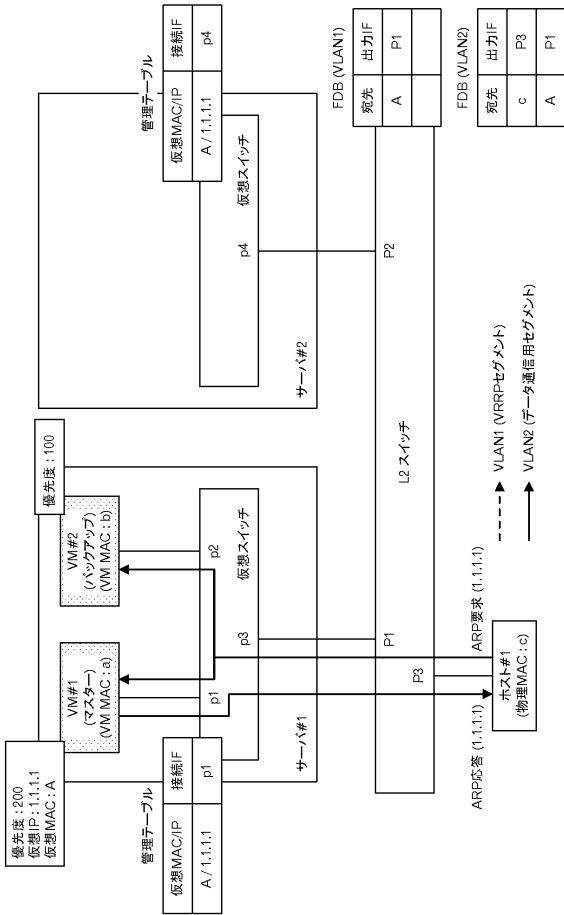
【 図 15 】



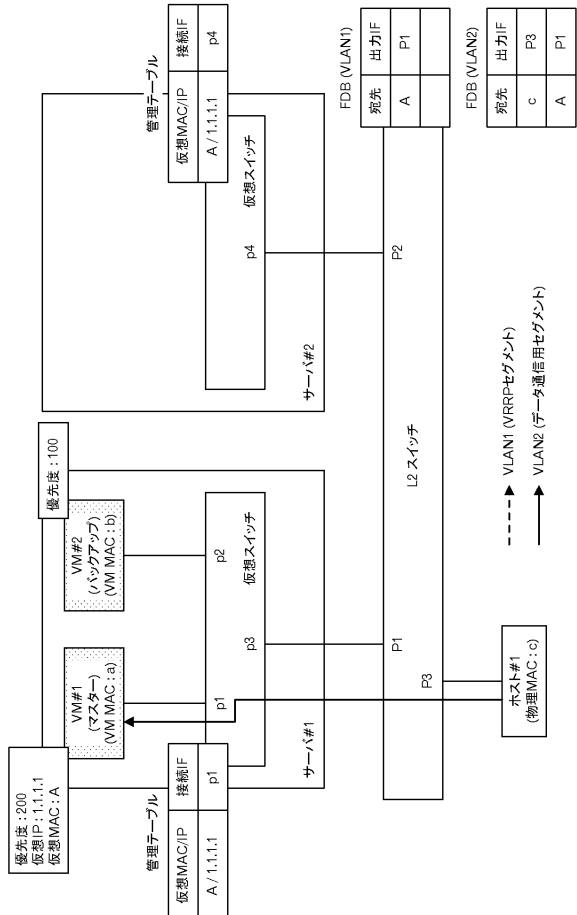
【 図 16 】



【 図 17 】

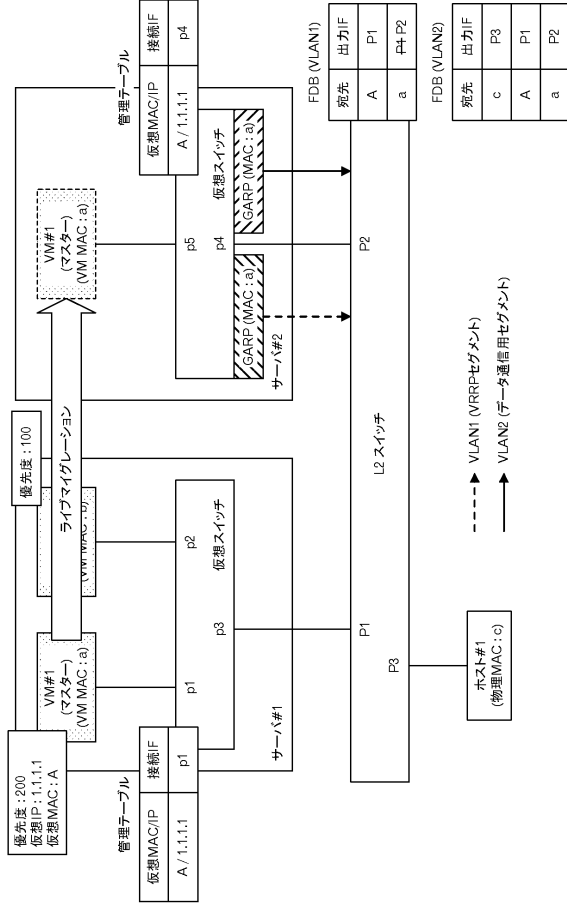


【 図 18 】

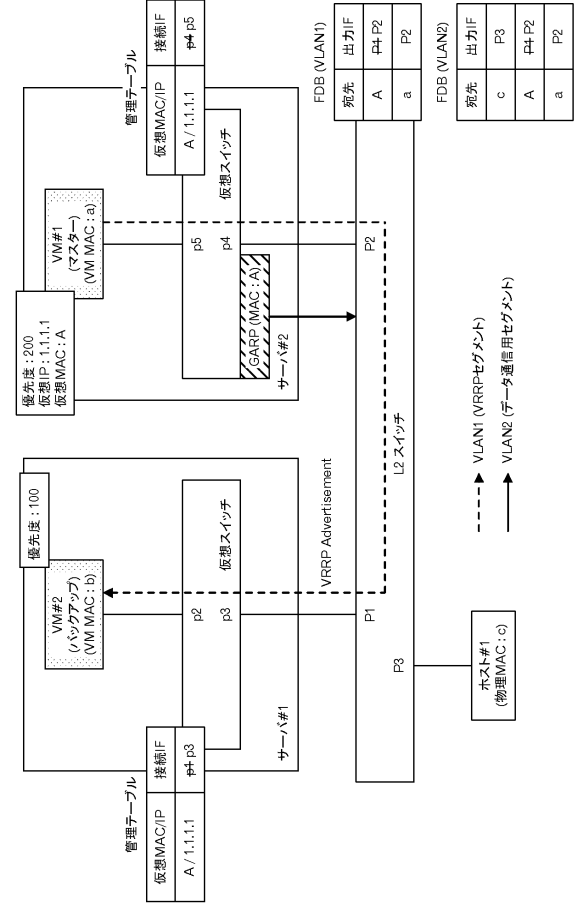




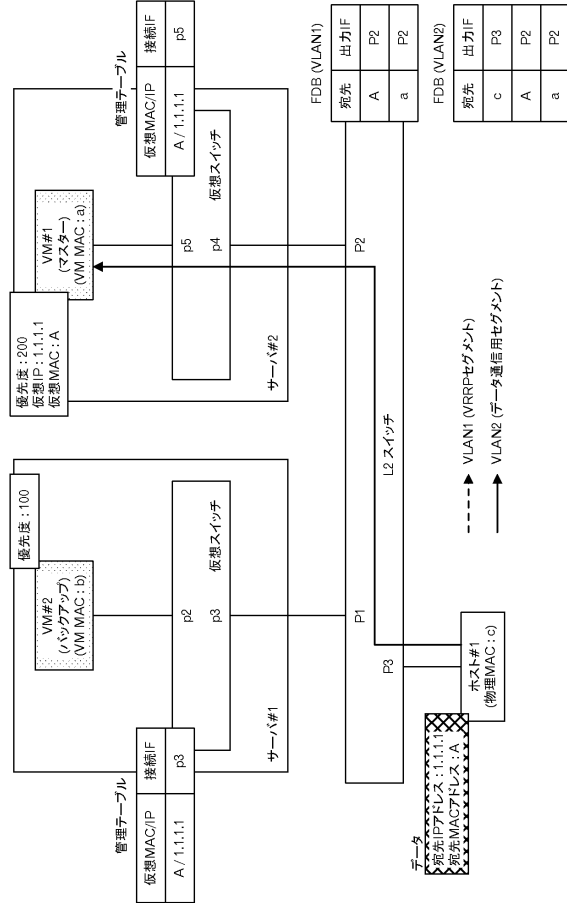
【 図 19 】



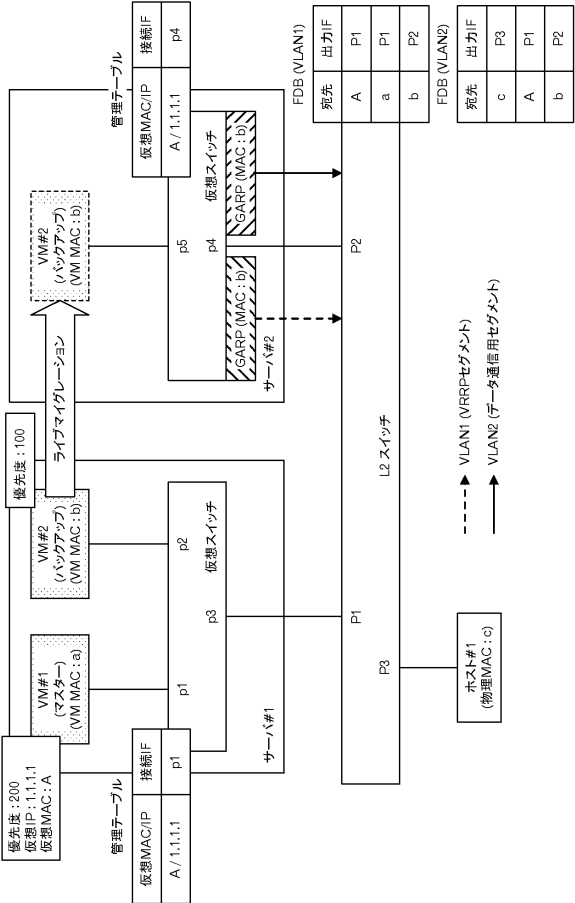
【 図 20 】



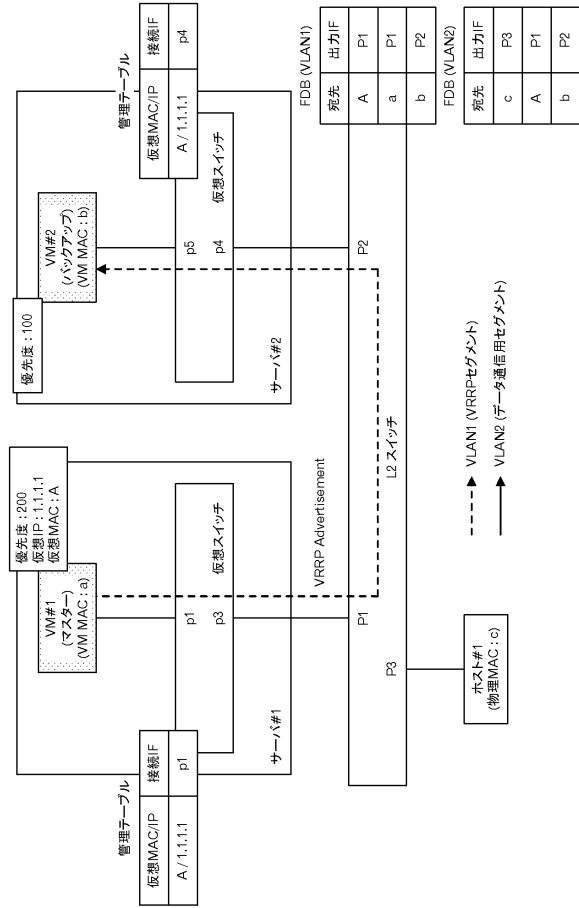
【 図 21 】



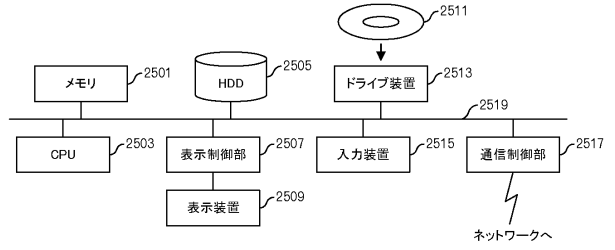
【 図 22 】



【図23】



【図24】



---

フロントページの続き

(56)参考文献 特開2011-070549(JP,A)  
特開2010-124129(JP,A)  
特開2011-198299(JP,A)

(58)調査した分野(Int.Cl., DB名)  
H04L 12/713