

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6257778号
(P6257778)

(45) 発行日 平成30年1月10日(2018.1.10)

(24) 登録日 平成29年12月15日(2017.12.15)

(51) Int.Cl.		F I			
G06F	9/48	(2006.01)	G06F	9/46	311B
G06F	9/46	(2006.01)	G06F	9/46	350

請求項の数 13 (全 17 頁)

(21) 出願番号	特願2016-541090 (P2016-541090)	(73) 特許権者	503433420
(86) (22) 出願日	平成26年12月12日 (2014.12.12)		華為技術有限公司
(65) 公表番号	特表2017-504109 (P2017-504109A)		HUAWEI TECHNOLOGIES
(43) 公表日	平成29年2月2日 (2017.2.2)		CO., LTD.
(86) 国際出願番号	PCT/CN2014/093686		中華人民共和国 518129 広東省深
(87) 国際公開番号	W02015/090158		▲チェン▼市龍崗区坂田 華為總部▲ベン
(87) 国際公開日	平成27年6月25日 (2015.6.25)		▼公樓
審査請求日	平成28年7月26日 (2016.7.26)		Huawei Administrati
(31) 優先権主張番号	201310716755.7		on Building, Bantia
(32) 優先日	平成25年12月20日 (2013.12.20)		n, Longgang Distric
(33) 優先権主張国	中国 (CN)		t, Shenzhen, Guangd
			ong 518129, P. R. Ch
			ina
		(74) 代理人	110001243
			特許業務法人 谷・阿部特許事務所
			最終頁に続く

(54) 【発明の名称】 仮想ネットワークインターフェースカードの割込みの親和性結合のための方法、およびコンピュータデバイス

(57) 【特許請求の範囲】

【請求項 1】

仮想ネットワークインターフェースカードの割込みの親和性結合のための方法であって、

IaaSリソース管理システムによって送信された要求メッセージを受信するステップであって、前記要求メッセージは、仮想ネットワークインターフェースカードの割込み親和性ポリシーパラメータを伝送する、ステップと、

複数の仮想中央処理装置VCPUと複数の物理中央処理装置PCPUとの間で1対1対応親和性結合を実行するステップと、

前記仮想ネットワークインターフェースカードの仮想割込みとVCPUとの間で親和性結合を実行するステップと、

前記親和性ポリシーパラメータに応じて、前記仮想ネットワークインターフェースカードの物理割込みと、対応するPCPUとの間で親和性結合を実行するステップと

を備えることを特徴とする方法。

【請求項 2】

前記親和性ポリシーパラメータに応じて前記仮想ネットワークインターフェースカードの物理割込みと対応するPCPUとの間で親和性結合を実行するステップは、

前記仮想ネットワークインターフェースカードの前記仮想割込みが前記複数のVCPUの第1のVCPUと親和性結合している場合、前記ポリシーパラメータに応じて、前記仮想ネットワークインターフェースカードの前記物理割込みと前記複数のPCPUの第1の

10

20

PCPUとの間で結合を実行するステップであって、前記第1のVCPUは、前記第1のPCPUとの親和性結合にあるVCPUである、ステップを特に備えることを特徴とする請求項1に記載の方法。

【請求項3】

前記仮想ネットワークインターフェースカードの前記仮想割込みが、第2のVCPUとの親和性結合にあるように変更する場合、前記仮想ネットワークインターフェースカードの前記物理割込みと第2のPCPUとの間で結合を実行するステップであって、前記第2のVCPUは、前記第2のPCPUとの親和性結合にあるVCPUである、ステップをさらに備えることを特徴とする請求項2に記載の方法。

【請求項4】

前記仮想ネットワークインターフェースカードの前記物理割込みと前記複数のPCPUの第1のPCPUとの間で結合を前記実行するステップは、前記仮想ネットワークインターフェースカードの仮想割込み番号に応じて、対応する物理割込みを決定するように対応する物理割込み番号を見出すステップと、前記VCPUと前記PCPUとの間の1対1対応親和性結合関係に応じて、対応する第1のPCPUを決定するステップと、前記仮想ネットワークインターフェースカードの前記物理割込みと前記対応する第1のPCPUとの間で結合を実行するステップとを特に備えることを特徴とする請求項2に記載の方法。

【請求項5】

前記仮想ネットワークインターフェースカードの前記物理割込みと第2のPCPUとの間で結合を前記実行するステップは、前記仮想ネットワークインターフェースカードの仮想割込み番号に応じて、対応する物理割込みを決定するように対応する物理割込み番号を見出すステップと、前記VCPUと前記PCPUとの間の1対1対応親和性結合関係に応じて、対応する第2のPCPUを決定するステップと、前記仮想ネットワークインターフェースカードの前記物理割込みと前記対応する第2のPCPUとの間で結合を実行するステップとを特に備えることを特徴とする請求項3に記載の方法。

【請求項6】

前記親和性ポリシーパラメータに応じて、前記仮想ネットワークインターフェースカードの物理割込みと、対応するPCPUとの間で親和性結合を前記実行するステップは、

前記ポリシーパラメータに応じて、前記仮想ネットワークインターフェースカードの前記物理割込みとアイドルPCPUとの間で親和性結合を実行するステップを特に備えることを特徴とする請求項1に記載の方法。

【請求項7】

前記アイドルPCPUの状態が非アイドルに変更された場合、前記仮想ネットワークインターフェースカードの前記物理割込みと別のアイドルPCPUとの間で親和性結合が実行されることを特徴とする請求項6に記載の方法。

【請求項8】

前記ポリシーパラメータに応じて、前記仮想ネットワークインターフェースカードの物理割込みと、対応するPCPUとの間で親和性結合を前記実行するステップは、

前記ポリシーパラメータに応じて、前記仮想ネットワークインターフェースカードの前記物理割込みと、前記仮想ネットワークインターフェースカードの前記物理割込みを処理するのに使用される予約されたPCPUとの間で親和性結合を実行するステップを特に備えることを特徴とする請求項1に記載の方法。

【請求項9】

IaaSリソース管理システムによって送信された要求メッセージを受信するように構成された受信モジュールであって、前記要求メッセージは、仮想ネットワークインターフェースカードの割込み親和性ポリシーパラメータを伝送する、受信モジュールと、

複数の仮想中央処理装置VCPUと複数の物理中央処理装置PCPUとの間で1対1対応親和性結合を実行し、かつ前記仮想ネットワークインターフェースカードの仮想割込みとVCPUとの間で親和性結合を実行するように構成された第1の結合モジュールと、

前記親和性ポリシーパラメータが第1の所定のものである場合に、前記第1の結合モジ

10

20

30

40

50

ジュールが、前記仮想ネットワークインターフェースカードの前記仮想割込みと前記V C P Uとの間で親和性結合を実行した後、前記親和性ポリシーパラメータに応じて、前記仮想ネットワークインターフェースカードの物理割込みと、対応するP C P Uとの間で親和性結合を実行するように構成された第2の結合モジュールと

を備えることを特徴とするコンピュータデバイス。

【請求項10】

前記第1の結合モジュールによって、前記仮想ネットワークインターフェースカードの前記仮想割込みと前記V C P Uとの間の親和性結合を実行する、または親和性結合を変更する動作をキャプチャし、かつ前記仮想ネットワークインターフェースカードの前記物理割込みと、前記対応するP C P Uとの間で親和性結合を実行するように前記第2の結合モジュールに命令するように構成されたキャプチャモジュールをさらに備え、

前記第2の結合モジュールは、前記キャプチャモジュールの通知に応じて、前記仮想ネットワークインターフェースカードの前記物理割込みと、前記対応するP C P Uとの間で親和性結合を実行するようにさらに構成され、前記対応するP C P Uは、前記V C P Uとの1対1対応親和性結合にあるP C P Uであることを特徴とする請求項9に記載のコンピュータデバイス。

【請求項11】

前記親和性ポリシーパラメータが第2の所定のものである場合に、前記第2の結合モジュールは、前記仮想ネットワークインターフェースカードの前記物理割込みとアイドルP C P Uとの間で親和性結合を実行するように特に構成されることを特徴とする請求項9に記載のコンピュータデバイス。

【請求項12】

前記親和性ポリシーパラメータが第3の所定のものである場合に、前記第2の結合モジュールは、前記仮想ネットワークインターフェースカードの前記物理割込みと、前記仮想ネットワークインターフェースカードの前記物理割込みを処理するのに使用される予約されたP C P Uとの間で親和性結合を実行するように特に構成されることを特徴とする請求項9に記載のコンピュータデバイス。

【請求項13】

請求項1ないし8のいずれかの方法をコンピュータに実行させるプログラムを記録したコンピュータ可読記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、通信技術の分野に関し、特に、仮想ネットワークインターフェースカードの割込みの親和性結合のための方法、およびコンピュータデバイスに関する。

【背景技術】

【0002】

遠隔通信サービスは、一般に、コンピューティングネットワークを多用するアプリケーションであり、ネットワークインターフェースカードのパケット送信および受信を処理するのにC P Uによって使用されるソフトウェア割込み処理およびハードウェア割込み処理の極めて多数のオーバーヘッドがある（例えば、512バイトUDPパケットは、300Mbpsの固定レートでギガビットネットワークインターフェースカード上で受信および送信され、ギガビットネットワークインターフェースカードのソフトウェア割込み処理およびハードウェア割込み処理のオーバーヘッドは、シングルコアC P U使用の30%に達する可能性がある）。

【0003】

物理マシンシナリオにおいて、マルチコアプロセッサに関して、いつ、いずれのC P U上でネットワークインターフェースカードの割込み処理が実行するかは、物理ネットワークインターフェースカードの割込みと物理C P Uの間で親和性結合を設定することによって制御されることが可能である。このようにして、遠隔通信サービスプログラムは、要求

10

20

30

40

50

されたとおりネットワークインターフェースカードの割込み処理ポリシーを正確に制御して、サービスプログラムの実行に対する予測不可能な干渉を回避するようにすることができる。

【0004】

クラウドコンピューティングの概念の絶え間ない推進とともに、仮想マシン上で遠隔通信サービスを実行することが、一般的な傾向となっている。仮想マシン上で遠隔通信サービスを実行することの信頼性および予測可能性の要件を満たすのに、仮想化プラットフォームは、ネットワークインターフェースカードの割込み処理ポリシーを正確に制御することができて、サービスプログラム、およびその仮想マシンと物理マシンを共有する隣接する仮想マシンのプログラムの実行に対する予測不可能な干渉を回避するようにする必要もある。

10

【0005】

図1に示されたとおり、仮想マシンが導入された後、仮想ネットワークインターフェースカード(VNIC、Virtual Network Interface Card)の割込みパケット送信および受信処理は、2段階に分割され、すなわち、第1に、仮想ネットワークインターフェースカードの物理割込みが、物理マシン上で実行するホストオペレーティングシステム(Host Operating System、Host OS)にパケットを送信し、次に、ホストOSが、そのパケットを、仮想ネットワークインターフェースカードの仮想割込みをトリガするように仮想マシン(Virtual Machine、VM)に配信し、次に、仮想マシン上で実行するゲストオペレーティングシステム(Guest Operating System、Guest OS)が、仮想ネットワークインターフェースカードの仮想割込みのパケット送信および受信を処理する。仮想ネットワークインターフェースカードの物理割込みから仮想ネットワークインターフェースカードの仮想割込みに対する転送のプロセスにおいて、仮想ネットワークインターフェースカードの物理割込みが、処理のためにパケットを送信する物理CPU、および仮想ネットワークインターフェースカードの仮想割込みが、処理のためにパケットを送信する仮想CPUに関して、デフォルトで固定の対応はなく、物理CPUと仮想CPUの間にも固定の対応はない。

20

【0006】

仮想ネットワークインターフェースカードの割込みパケット送信および受信処理に関する明確なスケジューリングポリシーがない場合、仮想ネットワークインターフェースカードの割込みパケット送信および受信の影響は、仮想マシン環境において予測されることが可能でない。図2に示されたとおり、VM1のアプリケーション3(Application 3、APP3)は、ネットワークパケット送信および受信を担うプロセスであり、VCPU3(Virtual Central Processing Unit、仮想中央処理装置)上で実行し、VCPU3は、PCPU3(Physical Central Processing Unit、物理中央処理装置3)上で実行するようにスケジュールされるものと想定される。明確なポリシー制御がない場合、仮想ネットワークインターフェースカードのものであり、APP3に対応する物理割込みのパケット送信および受信処理は、PCPU2上で生じることが可能であり、仮想ネットワークインターフェースカードの仮想割込みのパケット送信および受信処理は、VCPU1上で生じることが可能であり、その結果、VCPU2上で実行するAPP2、およびVCPU1上で実行するAPP1に対して干渉がそれぞれもたらされる。さらに、仮想ネットワークインターフェースカードの物理割込みおよび仮想割込みのパケット送信および受信に関する明確なポリシー制御がない場合、物理マシンを共有するVM間で隣接する干渉をもたらす可能性も高い。図2に示されたとおり、VM2のAPP5は、ネットワークパケット送信および受信を担い、VCPU5上で実行するプロセスであるものと想定され、VCPU5は、PCPU5上で実行するようにスケジュールされ、物理ネットワークインターフェースカードの対応する物理割込み処理は、PCPU4上であることが可能であり、VM1のAPP4に対して干渉がもたらされる。

30

40

50

【0007】

遠隔通信システムに関して、ネットワークインターフェースカードのソフトウェア割込み処理およびハードウェア割込み処理の多数のオーバヘッドがあるため、ネットワークインターフェースカードのソフトウェア割込み処理およびハードウェア割込み処理のオーバヘッドが、正確に制御される必要があり、したがって、物理マシンを共有する仮想マシン間のネットワークインターフェースカードの割込みパケット送信および受信処理が互いに干渉しないことが実施され、仮想マシン内で、その仮想マシンにおけるネットワークインターフェースカードの割込み処理がV C P U間で干渉を課さないことが実施される必要がある。

【発明の概要】

10

【課題を解決するための手段】

【0008】

技術的問題を解決するのに、本発明の実施形態は、仮想ネットワークインターフェースカードの割込み親和性ポリシーの設定を実施するための方法、および仮想ネットワークインターフェースカードの割込みの処理が互いに干渉しないことを実施するコンピュータデバイスを提供する。

【0009】

本発明の実施形態における仮想ネットワークインターフェースカードの割込みの親和性結合のための方法は、

I a a Sリソース管理システムによって送信された要求メッセージを受信するステップであって、要求メッセージは、仮想ネットワークインターフェースカードの割込み親和性ポリシーパラメータを伝送する、ステップ、

20

複数の仮想中央処理装置V C P Uと複数の物理中央処理装置P C P Uとの間で1対1対応親和性結合を実行するステップ、

仮想ネットワークインターフェースカードの仮想割込みとV C P Uとの間で親和性結合を実行するステップ、および

親和性ポリシーパラメータに応じて、仮想ネットワークインターフェースカードの物理割込みと、対応するP C P Uとの間で親和性結合を実行するステップを含む。

【0010】

特に、仮想ネットワークインターフェースカードの仮想割込みが、複数のV C P Uの第1のV C P Uとの親和性結合にある場合、仮想ネットワークインターフェースカードの物理割込みと複数のP C P Uの第1のP C P Uとの間でポリシーパラメータに応じて結合が実行され、第1のV C P Uは、第1のP C P Uとの親和性結合にあるV C P Uである。

30

【0011】

本方法は、仮想ネットワークインターフェースカードの仮想割込みが、第2のV C P Uとの親和性結合にあるように変更する場合、仮想ネットワークインターフェースカードの物理割込みと第2のP C P Uとの間で結合を実行するステップであって、第2のV C P Uは、第2のP C P Uとの親和性結合にあるV C P Uである、ステップをさらに含む。

【0012】

仮想ネットワークインターフェースカードの物理割込みと、対応するP C P Uとの間でポリシーパラメータに応じて親和性結合を実行するステップは、

40

仮想ネットワークインターフェースカードの物理割込みとアイドルP C P Uとの間でポリシーパラメータに応じて親和性結合を実行するステップ、または

ポリシーパラメータに応じて、仮想ネットワークインターフェースカードの物理割込みと、仮想ネットワークインターフェースカードの物理割込みを処理するのに使用される予約されたP C P Uとの間で親和性結合を実行するステップを特に含む。

【0013】

本発明の実施形態におけるコンピュータデバイスは、

I a a Sリソース管理システムによって送信された要求メッセージを受信するように構成された受信モジュールであって、要求メッセージは、仮想ネットワークインターフェー

50

スカードの割込み親和性ポリシーパラメータを伝送する、受信モジュールと、

複数の仮想中央処理装置V C P Uと複数の物理中央処理装置P C P Uとの間で1対1対応親和性結合を実行し、かつ仮想ネットワークインターフェースカードの仮想割込みとV C P Uとの間で親和性結合を実行するように構成された第1の結合モジュールと、

第1の結合モジュールが、仮想ネットワークインターフェースカードの仮想割込みとV C P Uとの間で親和性結合を実行した後、親和性ポリシーパラメータに応じて仮想ネットワークインターフェースカードの物理割込みと、対応するP C P Uとの間で親和性結合を実行するように構成された第2の結合モジュールとを含む。

【0014】

コンピュータデバイスは、

第1の結合モジュールによって、仮想ネットワークインターフェースカードの仮想割込みとV C P Uとの間の親和性結合を実行する、または親和性結合を変更する動作をキャプチャし、かつ仮想ネットワークインターフェースカードの物理割込みと、対応するP C P Uとの間で親和性結合を実行するように第2の結合モジュールに命令するように構成されたキャプチャモジュールをさらに含み、

第2の結合モジュールは、キャプチャモジュールの通知に応じて、仮想ネットワークインターフェースカードの物理割込みと、対応するP C P Uとの間で親和性結合を実行するようにさらに構成され、対応するP C P Uは、V C P Uとの1対1対応親和性結合にあるP C P Uである。

【0015】

第2の結合モジュールが、仮想ネットワークインターフェースカードの物理割込みと対応するP C P Uの間で親和性結合を実行することは、

第2の結合モジュールによって、仮想ネットワークインターフェースカードの物理割込みとアイドルP C P Uとの間で親和性結合を実行すること、または

第2の結合モジュールによって、仮想ネットワークインターフェースカードの物理割込みと、仮想ネットワークインターフェースカードの物理割込みを処理するのに使用される予約されたP C P Uとの間で親和性結合を実行することを特に含む。

【0016】

本発明の実施形態における技術的解決策を使用することにより、仮想割込みとV C P Uとの間、物理割込みとP C P Uとの間、および対応するV C P Uと、同一の仮想ネットワークインターフェースカードのP C P Uとの間の3つの親和性結合関係が一貫しており、したがって、別のV C P U上で実行するA P Pにもたらされる干渉がなく、同一のコンピュータデバイス上の別の仮想マシン上で実行するA P Pにもたらされる干渉もない。さらに、仮想ネットワークインターフェースカードの物理割込みは、P C P Uの負荷ステータスに応じてホストO Sによって平衡された状態で自動的にスケジュールされ、したがって、アイドルP C P Uが、P C P Uリソースの浪費を回避するようにより柔軟に使用されることが可能である。さらに、仮想マシンのV C P Uが、専用のP C P U上でスケジュールされなくてもよく、仮想ネットワークインターフェースカードの物理割込みが専用のP C P Uに結合され、その結果、スケジューリング動作を単純化し、別のV C P U上で実行するA P Pに対する干渉を回避する。

【図面の簡単な説明】

【0017】

本発明の実施形態における技術的解決策をより明確に説明するため、以下では、それらの実施形態を説明するのに要求される添付の図面を簡単に概説する。明らかに、以下の説明における添付の図面は、本発明のいくつかの実施形態を示すに過ぎず、当業者は、それでも、独創的な取組みなしにこれらの添付の図面から他の図面を導き出すことが可能である。

【図1】従来技術における仮想ネットワークインターフェースカードの割込みを示す概略図である。

【図2】従来技術における仮想ネットワークインターフェースカードの仮想割込みおよび

10

20

30

40

50

物理割込みの処理を示す概略図である。

【図 3】本発明による仮想ネットワークインターフェースカードの割込みの親和性結合のためのシステムアーキテクチャを示す図である。

【図 4 A】本発明の実施形態 1 による仮想ネットワークインターフェースカードの割込みの親和性結合を示す流れ図である。

【図 4 B】本発明の実施形態 1 による仮想ネットワークインターフェースカードの割込みの親和性結合を示す流れ図である。

【図 5】本発明の実施形態 1 による仮想ネットワークインターフェースカードの割込みの親和性結合を示す概略ブロック図である。

【図 6 A】本発明の実施形態 2 による仮想ネットワークインターフェースカードの割込みの親和性結合を示す流れ図である。

10

【図 6 B】本発明の実施形態 2 による仮想ネットワークインターフェースカードの割込みの親和性結合を示す流れ図である。

【図 7】本発明の実施形態 2 による仮想ネットワークインターフェースカードの割込みの親和性結合を示す概略ブロック図である。

【図 8 A】本発明の実施形態 3 による仮想ネットワークインターフェースカードの割込みの親和性結合を示す流れ図である。

【図 8 B】本発明の実施形態 3 による仮想ネットワークインターフェースカードの割込みの親和性結合を示す流れ図である。

【図 9】本発明の実施形態 3 による仮想ネットワークインターフェースカードの割込みの親和性結合を示す概略ブロック図である。

20

【図 10】本発明の実施形態 4 によるコンピュータデバイスの内部構造を示す概略図である。

【発明を実施するための形態】

【0018】

本発明の目的、技術的解決策、および利点をより明確にするため、以下では添付の図面を参照して本発明の実施形態を詳細にさらに説明する。

【0019】

図 3 のシステムアーキテクチャ図に示されるように、システムアーキテクチャは、サービスリソース管理システムと、サービスとしてのインフラストラクチャ (Infrastructure as a Service、IaaS) リソース管理システムと、仮想マシンのための実行環境を提供するコンピュータデバイスとを主に含み、コンピュータデバイスは、例えば、物理マシンとすることが可能である。

30

【0020】

サービスリソース管理システムの主な機能は、特定のサービスの実行リソース要件に応じて仮想マシンのための IaaS リソース管理システムに適用し、仮想ネットワークインターフェースカードの割込み親和性ポリシーパラメータが伝送される。

【0021】

IaaS リソース管理システムの主な機能は、サービスリソース管理システムから送信された仮想マシンリソース要件に応じて、その要件を満たすアイドル(idle)コンピュータデバイスを選択し、対応する仮想マシンを作成するようにその選択されたコンピュータデバイスに命令し、仮想マシンとは、現実のマシンと同様にプログラムを実行することができるコンピュータのソフトウェア実装を指し、仮想マシンは、複数の仮想中央処理装置 VCPU と、対応するオペレーティングシステム (ゲスト OS) と、アプリケーション (App) とを含む。コンピュータデバイスは、複数の物理中央処理装置 PCPU を含む。ホスト OS は、物理マシン上で実行し、ゲスト OS は、仮想マシン上で実行する。

40

【0022】

特定の手順は、以下のとおりである。すなわち、コンピュータデバイスが、IaaS リソース管理システムによって送信された要求メッセージを受信し、要求メッセージは、仮想ネットワークインターフェースカードの割込み親和性ポリシーパラメータを伝送する、

50

コンピュータデバイスが、複数の仮想中央処理装置V C P Uと複数の物理中央処理装置P C P Uとの間で1対1対応親和性結合を実行し、かつ仮想ネットワークインターフェースカードの仮想割込みとV C P Uとの間で親和性結合を実行する、および

コンピュータデバイスが、親和性ポリシーパラメータに応じて、仮想ネットワークインターフェースカードの物理割込みと、対応するP C P Uとの間で親和性結合を実行する。

【0023】

V C P UとP C P Uとの間の親和性結合とは、V C P Uが実行するようにスケジュールされるP C P Uが決定されることに基づく結合関係を指し、例えば、V C P U 1とP C P U 1との間の1対1対応親和性結合は、V C P U 1が、別のP C P U上で実行するのに代えて、P C P U 1上で実行するようにスケジュールされることを示す。

10

【0024】

仮想ネットワークインターフェースカードの割込みの親和性結合は、仮想ネットワークインターフェースカードの物理割込みの親和性結合と、仮想ネットワークインターフェースカードの仮想割込みの親和性結合とを含む。仮想ネットワークインターフェースカードの物理割込みの親和性結合は、仮想ネットワークインターフェースカードの物理割込みが生成された後、コンピュータデバイス上のホストOSが、処理のために物理割込みとの親和性結合にあるP C P Uに物理割込みを送信することを指す。仮想ネットワークインターフェースカードの仮想割込みの親和性結合は、仮想ネットワークインターフェースカードの仮想割込みが生成された後、コンピュータデバイス上のゲストOSが、処理のために仮想割込みとの親和性結合にあるV C P Uに仮想割込みを配信することを指す。

20

【0025】

親和性ポリシーパラメータは、異なる特定のサービスシナリオに応じて、サービスリソース管理システムによって異なるパラメータとして設定され、これらの異なるパラメータは、自動追従ポリシー、自動平衡ポリシー、自動的に排他的なポリシーなどを特に含む異なる親和性結合ポリシーを実行するようにホストOSに別々に命令する。

【0026】

本発明の解決策によれば、サービスとしてのインフラストラクチャI a a Sリソース管理システムが、仮想ネットワークインターフェースカードの割込みパケット送信および受信処理のスケジューリングポリシーを構成する一般的な機能を提供し、したがって、仮想マシンを作成する際、コンピュータデバイスが、I a a Sリソース管理システムによって送信された割込み親和性ポリシーパラメータに応じて仮想ネットワークインターフェースカードの割込みに対して対応する結合およびスケジューリングを実行して、その結果、構成自動化を実施することができる。

30

【0027】

実施形態1

親和性ポリシーパラメータが自動追従ポリシーとして設定された場合、仮想ネットワークインターフェースカードの割込み親和性ポリシーパラメータは、物理割込みとP C P Uの間の親和性結合が、仮想割込みとV C P Uの間の親和性結合の変更に従うことを示す。図4 Aおよび図4 Bに示されるとおり、仮想ネットワークインターフェースカードの割込みの親和性結合のための方法は、特に以下のとおりであるステップ(1)、(2)、および(3)を主に含む。すなわち、

40

【0028】

ステップ(1)は、以下の3つのサブステップを主に含む。

【0029】

サービスリソース管理システムが、仮想マシンリソース要求を開始し、仮想マシンリソース要求の中の仮想ネットワークインターフェースカードの割込み親和性ポリシーパラメータは、自動的追従である。

【0030】

要求を受信した後、I a a Sリソース管理システムが、リソース要件を満たすコンピュータデバイスを見出し、コンピュータデバイス上で実行するホストOSに対して仮想マシ

50

ンを作成するための要求を開始し、要求は、仮想ネットワークインターフェースカードの自動追従割込み親和性ポリシーを伝送する。

【0031】

ホストOSは、対応する仮想マシンの仮想ネットワークインターフェースカードの自動追従親和性ポリシーを記録し、仮想マシンにおけるゲストOSによる、仮想ネットワークインターフェースカードの仮想割込みの親和性結合を設定する動作がキャプチャされることが可能な自動キャプチャ機能をイネーブル(enable)にする。

【0032】

ステップ(2)は、以下の2つのサブステップを主に含む。

【0033】

サービスリソース管理システムが、仮想マシンを起動するための要求を開始し、要求を受信した後、IaaSリソース管理システムが、対応するコンピュータデバイス上で実行するホストOSに対して対応する仮想マシンを起動するための要求を開始する。

【0034】

仮想マシンが開始された後、ゲストOSが実行し、ゲストOSが実行した後、APPが実行し、仮想マシンが起動された後、ホストOSが、仮想マシンのVCPUと、対応するPCPUとの間の1対1対応親和性結合を自動的に設定する(図5に示されるように、1対1対応親和性結合は、VCPU1からVCPU4までと、PCPU1からPCPU4までとの間でそれぞれ実行される)。

【0035】

ステップ(3)は、以下の2つのサブステップを主に含む。

【0036】

仮想マシン上で実行するゲストOSが、仮想ネットワークインターフェースカードの仮想割込みとVCPUの間で親和性結合を設定する、または仮想ネットワークインターフェースカードの仮想割込みとVCPUとの間に親和性結合を変更する。

【0037】

ホストOSによってイネーブルにされた自動キャプチャ機能が、前述の仮想マシンにおけるゲストOSによる、仮想ネットワークインターフェースカードの仮想割込みの親和性結合を設定する、または変更する動作をキャプチャし、かつ自動追従ポリシーに応じて仮想ネットワークインターフェースカードの物理割込みと、対応するPCPUとの間で親和性結合関係を自動的に結合し、対応するPCPUは、VCPUとの1対1対応親和性結合にあるPCPUである。

【0038】

仮想ネットワークインターフェースカードの物理割込みと、対応するPCPUとの間で親和性結合関係を自動的に結合する特定の手順は、以下のとおりである。すなわち、ホストOSが、仮想ネットワークインターフェースカードの仮想割込みに応じて、仮想ネットワークインターフェースカードの対応する仮想割込み番号を見出し、仮想割込み番号に応じて対応する物理割込み番号を見出し(同一の仮想ネットワークインターフェースカードの仮想割込み番号と物理割込み番号は、1対1対応にある)、対応する物理割込みをさらに決定し、ホストOSが、VCPUとPCPUとの間の1対1対応親和性結合関係に応じて、対応するPCPUを決定し、仮想ネットワークインターフェースカードの物理割込みと、対応するPCPUとの間で結合を実行するようにして、その結果、仮想割込みとVCPUとの間、物理割込みとPCPUとの間、および対応するVCPUと、同一の仮想ネットワークインターフェースカードのPCPUとの間の3つの親和性結合関係の一貫性を実施する。

【0039】

仮想マシン上で実行するゲストOSが、仮想ネットワークインターフェースカードの仮想割込みとVCPUとの間で親和性結合関係を設定し、または変更し、ホストOSが、その動作をキャプチャし、仮想ネットワークインターフェースカードの物理割込みと対応するPCPUとを結合するように従う。この様態が、自動追従ポリシーである。

【0040】

ステップ（３）は、図５における概略ブロック図を使用することによって示されることが可能であり、図における（２）、（３）、および（４）は、前述のステップ（２）および（３）に対応し、右の図（図における破線の右側）は、仮想ネットワークインターフェースカードの仮想割込みとＶＣＰＵとの間の親和性結合関係が変更した（ＶＣＰＵ３からＶＣＰＵ４に）場合、ホストＯＳが、仮想ネットワークインターフェースカードの物理割込みと対応するＰＣＰＵとの間で親和性関係を結合するように自動的に従って、仮想ネットワークインターフェースカードの物理割込みとＰＣＰＵ４との間で親和性結合を実行することが可能であり、ＶＣＰＵ４は、ＰＣＰＵ４との親和性結合にあることを示す。

【0041】

10

前述の手順から、自動追従ポリシーが使用される場合、仮想割込みとＶＣＰＵとの間、物理割込みとＰＣＰＵとの間、および対応するＶＣＰＵと、同一の仮想ネットワークインターフェースカードのＰＣＰＵとの間の３つの親和性結合関係が一貫しており、したがって、別のＶＣＰＵ上で実行するＡＰＰにもたらされる干渉がなく、同一のコンピュータデバイス上の別の仮想マシン上で実行するＡＰＰにもたらされる干渉もないことが知られる。

【0042】

実施形態２

親和性ポリシーパラメータが自動平衡ポリシーとして設定された場合、仮想ネットワークインターフェースカードの割込み親和性ポリシーパラメータは、物理割込みを処理するためのＰＣＰＵの平衡されたスケジューリングを示す。仮想ネットワークインターフェースカードの割込みの親和性結合のための方法が、図６Ａおよび図６Ｂに示され、特に以下のとおりであるステップ（１）、（２）、および（３）を主に含む。

20

【0043】

ステップ（１）は、以下の３つのサブステップを主に含む。

【0044】

サービスリソース管理システムが、仮想マシンリソース要求を開始し、仮想マシンリソース要求の中の仮想ネットワークインターフェースカードの割込み親和性ポリシーパラメータは、自動平衡である。

【0045】

30

要求を受信した後、ＩaaSリソース管理システムが、リソース要件を満たすコンピュータデバイスを見出し、コンピュータデバイス上で実行するホストＯＳに対して仮想マシンを作成するための要求を開始し、要求は、仮想ネットワークインターフェースカードの自動平衡割込み親和性ポリシーを伝送する。

【0046】

ホストＯＳが、対応する仮想マシンの仮想ネットワークインターフェースカードの自動平衡親和性ポリシーを記録し、自動キャプチャ機能を無効(disable)にして、仮想マシンにおけるゲストＯＳによる、仮想ネットワークインターフェースカードの仮想割込みの親和性結合を設定する動作が、キャプチャされることが可能でないようにする。

【0047】

40

ステップ（２）は、以下の２つのサブステップを主に含む。

【0048】

サービスリソース管理システムが、仮想マシンを起動するための要求を開始し、要求を受信した後、ＩaaSリソース管理システムが、対応するコンピュータデバイス上で実行するホストＯＳに対して対応する仮想マシンを起動するための要求を開始する。

【0049】

仮想マシンが起動された後、ゲストＯＳが実行し、ゲストＯＳが実行した後、ＡＰＰが実行し、仮想マシンが起動された後、ホストＯＳが、仮想マシンのＶＣＰＵと対応するＰＣＰＵとの間で１対１対応親和性結合を自動的に設定する（図７に示されるとおり、１対１対応親和性結合は、ＶＣＰＵ１からＶＣＰＵ４までと、ＰＣＰＵ１からＰＣＰＵ４まで

50

との間でそれぞれ実行される)。

【0050】

ステップ(3)は、以下の2つのサブステップを主に含む。

【0051】

仮想マシン上で実行するゲストOSが、仮想ネットワークインターフェースカードの仮想割込みとVCPUとの間で親和性結合を設定する、または仮想ネットワークインターフェースカードの仮想割込みとVCPUとの間で親和性結合を変更する。

【0052】

ホストOSが、自動平衡ポリシーに応じて、自動キャプチャ機能を無効にして、仮想マシンにおけるゲストOSによって、仮想ネットワークインターフェースカードの仮想割込みの親和性結合を設定する、または変更する動作がキャプチャされないようにする。ホストOSが、仮想マシンが配置されたPCPUの負荷ステータスに応じて、仮想ネットワークインターフェースカードの物理割込みを平衡された状態で自動的にスケジュールし、かつ仮想ネットワークインターフェースカードの物理割込みを、それほど負荷がかけられていない、またはアイドルPCPUに動的に結合する。

【0053】

ステップ(3)は、図7における概略ブロック図を使用することによって示されることが可能であり、図における(2)および(3)は、前述のステップ(2)および(3)にそれぞれ対応し、APP2が配置されたPCPU2に負荷がそれほどかけられていない場合、ホストOSが、仮想ネットワークインターフェースカードのものであり、かつAPP3に対応する物理割込みとPCPU2との間で親和性結合を実行し、図の右の図(図における破線の右側)に示されるとおり、期間の後、APP2が配置されたPCPU2がビジーになり、APP4が配置されたPCPU4がアイドルになった場合、ホストOSが、仮想ネットワークインターフェースカードのものであり、かつAPP3に対応する物理割込みとPCPU4との間で親和性結合を実行する。

【0054】

自動追従ポリシーとは対照的に、ホストOSは、仮想ネットワークインターフェースカードの仮想割込みとVCPUとの間の親和性結合の動作をキャプチャせず、仮想ネットワークインターフェースカードの物理割込みが、PCPUの負荷ステータスに応じてホストOSによって平衡された状態で自動的にスケジュールされて、アイドルPCPUが、PCPUリソースの浪費を回避するようにより柔軟に利用されることが可能であるようになる。

【0055】

実施形態3

親和性ポリシーパラメータが自動的に排他的なポリシーとして設定された場合、仮想ネットワークインターフェースカードの割込み親和性ポリシーパラメータは、物理割込みを処理する予約されたPCPUがあることを示す。仮想ネットワークインターフェースカードの割込みの親和性結合のための方法が図8Aおよび図8Bに示され、特に以下のとおりであるステップ(1)、(2)、および(3)を主に含む。

【0056】

ステップ(1)は、以下の3つのサブステップを主に含む。

【0057】

サービスリソース管理システムが、仮想マシンリソース要求を開始し、仮想マシンリソース要求の中の仮想ネットワークインターフェースカードの割込み親和性ポリシーパラメータは、自動的に排他的な使用である。

【0058】

要求を受信した後、IaaSリソース管理システムが、リソース要件を満たすコンピュータデバイスを見出し、コンピュータデバイス上で実行するホストOSに対して仮想マシンを作成するための要求を開始し、要求は、仮想ネットワークインターフェースカードの自動的に排他的な使用割込み親和性ポリシーを伝送する。

【0059】

ホストOSは、対応する仮想マシンの仮想ネットワークインターフェースカードの自動的に排他的な使用親和性ポリシーを記録し、自動キャプチャ機能を無効にして、仮想マシンにおけるゲストOSによる、仮想ネットワークインターフェースカードの仮想割込みの親和性結合を設定する動作がキャプチャされることが可能でないようにする。

【0060】

ステップ(2)は、以下の2つのサブステップを主に含む。

【0061】

サービスリソース管理システムが、仮想マシンを起動するための要求を開始し、要求を受信した後、IaaSリソース管理システムが、対応するコンピュータデバイス上で実行するホストOSに対して対応する仮想マシンを起動するための要求を開始する。

10

【0062】

仮想マシンが起動された後、ゲストOSが実行し、ゲストOSが実行した後、APPが実行し、仮想マシンが起動された後、ホストOSが、仮想マシンのVCPUと対応するPCPUとの間で1対1対応親和性結合を自動的に設定する(図9に示されるとおり、1対1対応親和性結合は、VCPU1からVCPU4までと、PCPU1からPCPU4までとの間でそれぞれ実行される)。

【0063】

ステップ(3)は、以下の2つのサブステップを主に含む。

【0064】

20

仮想マシン上で実行するゲストOSが、仮想ネットワークインターフェースカードの仮想割込みとVCPUとの間で親和性結合を設定する、または仮想ネットワークインターフェースカードの仮想割込みとVCPUとの間で親和性結合を変更する。

【0065】

ホストOSが、自動的に排他的なポリシーに応じて、自動キャプチャ機能を無効にして、前述の仮想マシンにおけるゲストOSによって、仮想ネットワークインターフェースカードの仮想割込みの親和性結合を設定する、または変更する動作がキャプチャされないようにし、ホストOSが、仮想ネットワークインターフェースカードの物理割込みと、仮想ネットワークインターフェースカードの物理割込みを処理するのに使用される予約されたPCPUとの間で親和性結合を実行する。

30

【0066】

ステップ(3)は、図9における概略ブロック図を使用することによって示されることが可能であり、PCPU9およびPCPU10は、仮想ネットワークインターフェースカードの物理割込みのために専用に予約されたPCPUであり、仮想マシン上のVCPUは、PCPU9およびPCPU10の上でスケジュールされず、仮想ネットワークインターフェースカードに対応する物理割込みは、排他的にこれらの専用のプロセッサ上でスケジュールされる。

【0067】

実施形態2における解決策と同様に、ホストOSは、仮想ネットワークインターフェースカードの仮想割込みとVCPUとの間の親和性結合の動作をキャプチャせず、仮想マシンのVCPUは、専用のPCPU上でスケジュールされなくてもよく、仮想ネットワークインターフェースカードの物理割込みは、専用のPCPUに結合され、その結果、スケジューリング動作を単純化し、別のVCPU上で実行するAPPに対する干渉を回避する。

40

【0068】

実施形態4

図10は、

IaaSリソース管理システム00によって送信された要求メッセージを受信するように構成された受信モジュール101であって、要求メッセージは、仮想ネットワークインターフェースカードの割込み親和性ポリシーパラメータを伝送する受信モジュール101と、

50

複数の仮想中央処理装置 V C P U と複数の物理中央処理装置 P C P U の間で 1 対 1 対応親和性結合を実行し、かつ仮想ネットワークインターフェースカードの仮想割込みと V C P U との間で親和性結合を実行するように構成された第 1 の結合モジュール 1 0 2 と、

第 1 の結合モジュール 1 0 2 が、仮想ネットワークインターフェースカードの仮想割込みと V C P U の間で親和性結合を実行した後、親和性ポリシーパラメータに応じて、仮想ネットワークインターフェースカードの物理割込みと、対応する P C P U との間で親和性結合を実行するように構成された第 2 の結合モジュール 1 0 3 とを含むコンピュータデバイス 1 0 の内部構造の図を示す。

【 0 0 6 9 】

さらに、コンピュータデバイスは、第 1 の結合モジュール 1 0 2 によって、仮想ネットワークインターフェースカードの仮想割込みと V C P U との間で親和性結合を実行する、または親和性結合を変更する動作をキャプチャし、かつ仮想ネットワークインターフェースカードの物理割込みと、対応する P C P U との間で親和性結合を実行するように第 2 の結合モジュール 1 0 3 に命令するように構成されたキャプチャモジュール 1 0 4 をさらに含む。

【 0 0 7 0 】

特に、第 2 の結合モジュール 1 0 3 は、キャプチャモジュール 1 0 4 の通知に応じて、仮想ネットワークインターフェースカードの物理割込みと、対応する P C P U との間で親和性結合を実行するように構成され、対応する P C P U は、V C P U との 1 対 1 対応親和性結合にある P C P U であり、または

第 2 の結合モジュール 1 0 3 は、仮想ネットワークインターフェースカードの物理割込みとアイドル P C P U との間で親和性結合を実行し、または

第 2 の結合モジュール 1 0 3 は、仮想ネットワークインターフェースカードの物理割込みと、仮想ネットワークインターフェースカードの物理割込みを処理するのに使用される予約された P C P U との間で親和性結合を実行する。

【 0 0 7 1 】

当業者は、実施形態のステップのすべて、またはいくつかは、ハードウェア、または関連付けられたハードウェアに命令するプログラムによって実施されることが可能であることを理解することができる。そのプログラムは、コンピュータ可読記憶媒体の中に記憶されることが可能である。記憶媒体は、読取り専用メモリ、磁気ディスク、または光ディスクを含むことが可能である。

【 0 0 7 2 】

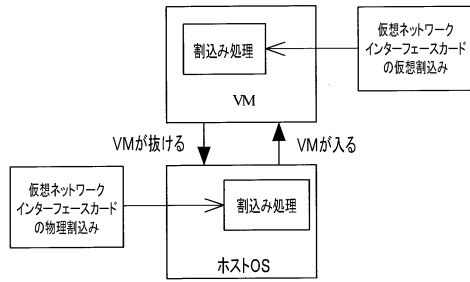
前述の説明は、本発明の例示的な実施形態に過ぎないが、本発明を限定するように意図されない。本発明の趣旨および原理を逸脱することなく行われるいずれの変形、均等の置換、および改良も、本発明の保護範囲内に入るものとする。

10

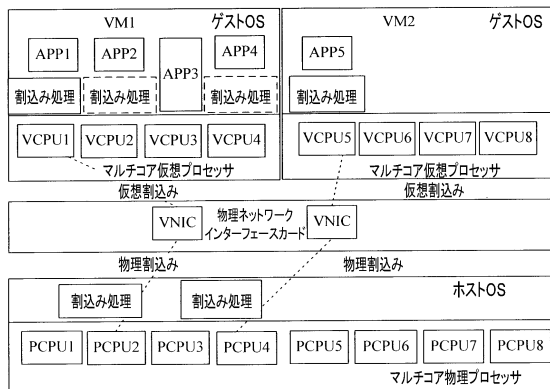
20

30

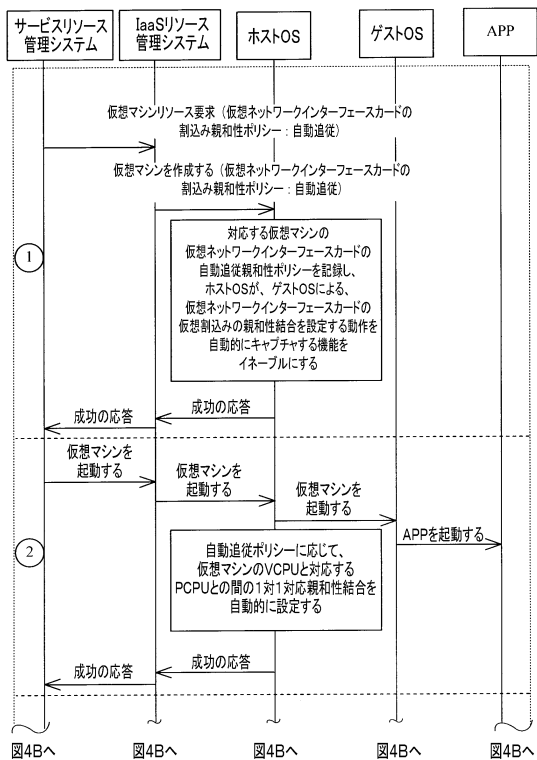
【図 1】



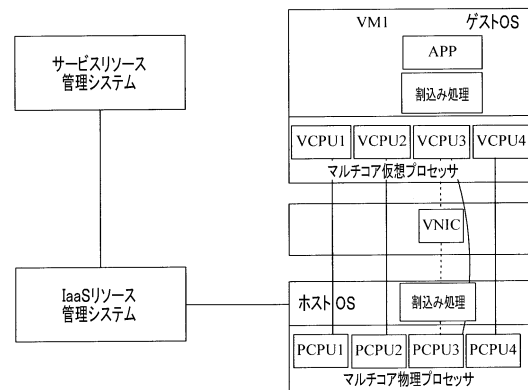
【図 2】



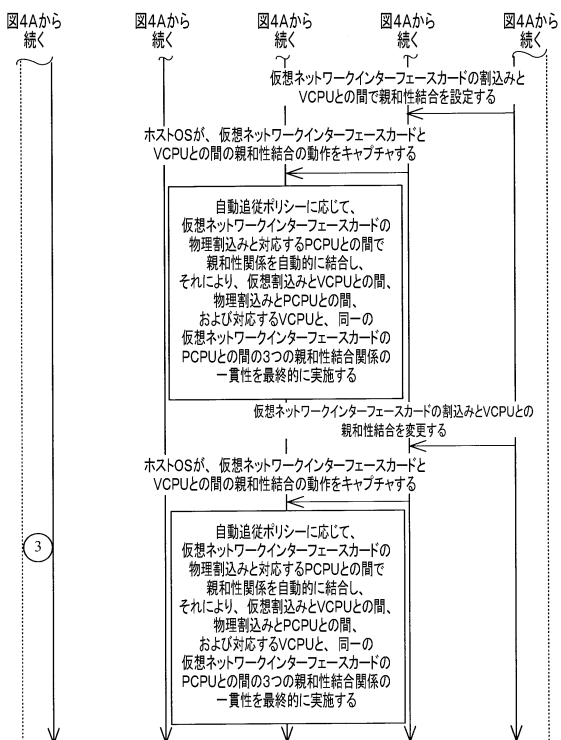
【図 4 A】



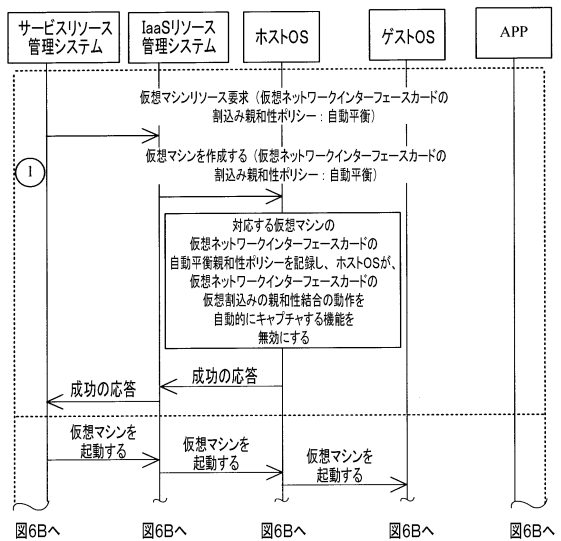
【図 3】



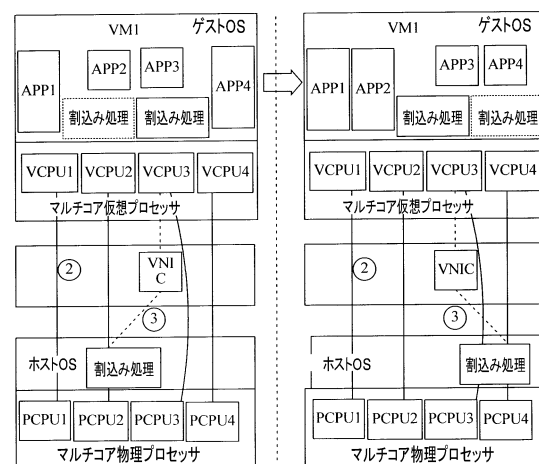
【図 4 B】



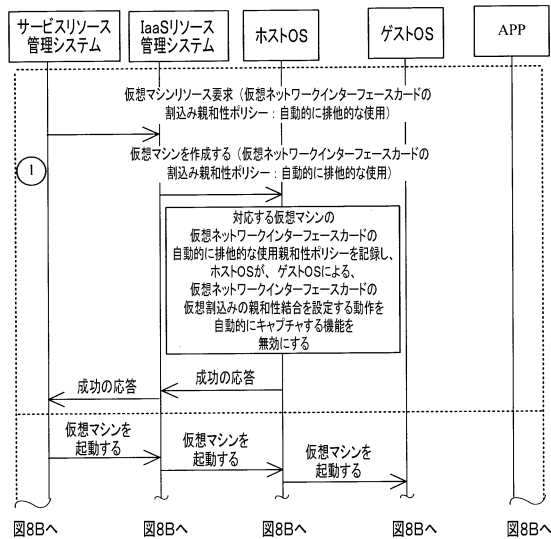
【図 6 A】



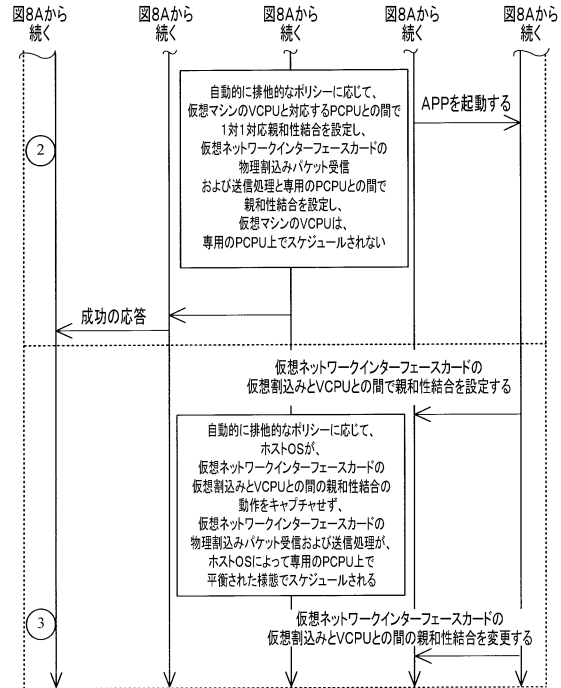
【図 7】



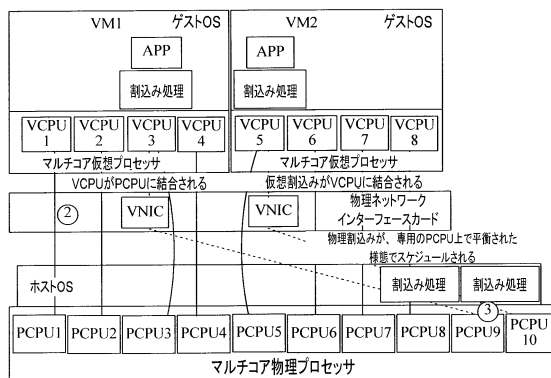
【図 8 A】



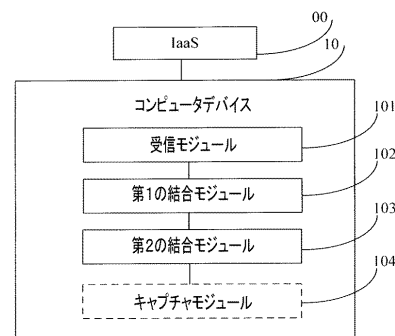
【図 8 B】



【图 9】



【図 10】



フロントページの続き

(72)発明者 夏 ▲海▼

中華人民共和国 518129 広東省深▲チェン▼市龍崗区坂田 華為總部▲ベン▼公樓

審査官 清木 泰

(56)参考文献 特表2012-529092 (JP, A)

特表2012-515995 (JP, A)

特表2010-538373 (JP, A)

特開2010-128911 (JP, A)

特開2004-303237 (JP, A)

米国特許出願公開第2013/0104127 (US, A1)

米国特許出願公開第2011/0202699 (US, A1)

米国特許出願公開第2009/0183153 (US, A1)

高橋浩和, 仮想マシンモニターXen3.0解説室 第6回仮想割り込み, オープンソースマガジン, 日本, ソフトバンククリエイティブ株式会社, 2006年 9月 1日, 第15巻, 第9号, Pages:147~154

Yaozu Dong, Dongxiao Xu, Yang Zhang, Guangdeng Liao, Optimizing Network I/O Virtualization with Efficient Interrupt Coalescing and Virtual Receive Side Scaling, Proceedings of 2011 IEEE International Conference on Cluster Computing (CLUSTER), IEEE, 2011年 9月26日, Pages:26-34

Jiuxing Liu, Bulent Abali, Virtualization Polling Engine(VPE): Using Dedicated CPU Cores to Accelerate I/O Virtualization, Proceedings of the 23rd International Conference on Supercomputing (ICS'09), ACM, 2009年 6月 8日, Pages:225-234

(58)調査した分野(Int.Cl., DB名)

G06F 9/46- 9/54