

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5433837号
(P5433837)

(45) 発行日 平成26年3月5日(2014.3.5)

(24) 登録日 平成25年12月20日(2013.12.20)

(51) Int. Cl.		F I			
G06F	9/50	(2006.01)	G06F	9/46	465D
G06F	9/46	(2006.01)	G06F	9/46	350
G06F	1/32	(2006.01)	G06F	1/00	332Z

請求項の数 15 (全 27 頁)

(21) 出願番号	特願2007-314761 (P2007-314761)	(73) 特許権者	000005108
(22) 出願日	平成19年12月5日(2007.12.5)		株式会社日立製作所
(65) 公開番号	特開2009-140157 (P2009-140157A)		東京都千代田区丸の内一丁目6番6号
(43) 公開日	平成21年6月25日(2009.6.25)	(74) 代理人	100114236
審査請求日	平成22年10月6日(2010.10.6)		弁理士 藤井 正弘
審判番号	不服2012-20841 (P2012-20841/J1)	(74) 代理人	100075513
審判請求日	平成24年10月23日(2012.10.23)		弁理士 後藤 政喜
		(72) 発明者	對馬 雄次
			東京都国分寺市東恋ヶ窪一丁目280番地
			株式会社日立製作所 中央研究所内
		(72) 発明者	上原 敬太郎
			東京都国分寺市東恋ヶ窪一丁目280番地
			株式会社日立製作所 中央研究所内

最終頁に続く

(54) 【発明の名称】 仮想計算機システム、仮想計算機の制御方法及びプログラム

(57) 【特許請求の範囲】

【請求項1】

スリープ状態と通常の動作状態とを切り替え可能な物理CPUを複数備えた物理計算機と、

前記物理計算機を複数の論理区画に分割し、各論理区画上でそれぞれゲストOSを動作させて、各論理区画に対する物理計算機の資源の割当を制御する仮想化制御部と、を備えた仮想計算機システムであって、

前記仮想化制御部は、

前記論理区画に対する操作指令を受け付ける論理区画制御部と、

前記操作指令に基づいて前記論理区画に前記ゲストOSを実行する仮想CPUを割り当てる仮想CPU制御部と、

前記操作指令に基づいて前記仮想CPUを前記物理CPUに割り当て、前記物理CPUの動作状態を制御する仮想-物理CPU割当制御部と、

前記仮想-物理CPU割当制御部からの操作指令に基づいて前記物理CPUの動作状態を制御する物理CPU制御部と、を含み、

前記物理CPUは、動作クロックの周波数を変更するクロック制御部を有し、

前記操作指令は、前記論理区画に割り当てられた当該仮想CPUが前記物理CPUを使用する比率としての割り当て率を含み、

前記仮想-物理CPU割当制御部は、

前記論理区画制御部から仮想CPUを生成する操作指令を受け付けると、前記複数の物

10

20

理CPUのうち、前記通常動作状態の物理CPUに前記仮想CPUを割り当て可能か否かを判定し、割り当て可能な場合には当該仮想CPUを前記通常動作状態の物理CPUに割り当て、割り当て不能な場合にはスリープ状態の物理CPUを起動する指令を前記物理CPU制御部に発行して、前記仮想CPUを当該起動した物理CPUに割り当て、仮想CPUと物理CPUの割り当て状態と物理CPUの動作状態を管理するテーブルに前記割り当てた仮想CPUを追加して更新し、

前記論理区画制御部から仮想CPUを削除する操作指令を受け付けると、当該仮想CPUを前記テーブルから削除し、当該仮想CPUを削除した物理CPUに割り当てられている仮想CPUがなくなった場合には、当該物理CPUをスリープ状態に移行するよう前記物理CPU制御部に指令し、

前記物理CPUに対する仮想CPUの割り当てを変更したときには、当該物理CPUに割り当てられている仮想CPUの割り当て率の和に基づいて前記物理CPUの動作クロックの周波数を決定し、当該動作クロックの周波数を前記物理CPU制御部へ指令し、

前記物理CPU制御部は前記動作クロックの周波数を前記クロック制御部へ指令することを特徴とする仮想計算機システム。

【請求項2】

前記仮想 - 物理CPU割当制御部は、

前記論理区画制御部から仮想CPUを削除する操作指令を受け付けると、当該仮想CPUを前記テーブルから削除した後に、当該テーブルを参照して当該仮想CPUを削除した第1の物理CPUに割り当てられた第2の仮想CPUが存在し、かつ、通常動作状態にある第2の物理CPUに前記第2の仮想CPUを割り当て可能な場合には、前記第2の仮想CPUを前記第2の物理CPUに再割り当てを行い、前記第1の物理CPUに割り当てられている仮想CPUがなくなると、当該第1の物理CPUをスリープ状態に移行させることを特徴とする請求項1に記載の仮想計算機システム。

【請求項3】

前記論理区画は、ゲストOSが利用する仮想CPUの実使用率を検知して前記仮想化制御部へ通知する稼動状態検出部を含み、

前記仮想 - 物理CPU割当制御部は、

前記物理CPUに対する仮想CPUの割り当てを変更したときには、当該物理CPUに割り当てられている仮想CPUの実使用率の和に基づいて物理CPUに対する仮想CPUの割り当てを変更することを特徴とする請求項1に記載の仮想計算機システム。

【請求項4】

前記仮想 - 物理CPU割当制御部は、

前記論理区画制御部から仮想CPUを削除する操作指令を受け付けると、当該仮想CPUを前記テーブルから削除した後に、当該テーブルを参照して当該仮想CPUを削除した第1の物理CPUに割り当てられた第2の仮想CPUが存在するときには、前記稼動状態検出部からの実使用率を取得して、通常動作状態にある第2の物理CPUに割り当てられている仮想CPUの実使用率の和に、前記第2の仮想CPUの実使用率を加算した値が所定の閾値未満であれば、前記第2の仮想CPUを前記第2の物理CPUに再割り当てを行い、前記第1の物理CPUに割り当てられている仮想CPUがなくなると、当該第1の物理CPUをスリープ状態に移行させることを特徴とする請求項3に記載の仮想計算機システム。

【請求項5】

前記仮想 - 物理CPU割当制御部は、

前記第2の物理CPUに割り当てられている仮想CPUの実使用率の和に、前記第2の仮想CPUの実使用率を加算した値が所定の閾値未満であれば、前記第2の物理CPUに割り当てられている仮想CPUの割り当て率の和に前記第2の仮想CPUの割り当て率を加算した値が100%を超えることを許容することを特徴とする請求項4に記載の仮想計算機システム。

【請求項6】

10

20

30

40

50

スリープ状態と通常の動作状態とを切り替え可能な物理CPUを複数備えた物理計算機を複数の論理区画に分割し、各論理区画上でそれぞれゲストOSを動作させて、各論理区画に対する物理計算機の資源の割当を制御する仮想化制御部を備えた仮想計算機システムの制御方法であって、

前記物理CPUは、動作クロックの周波数を変更するクロック制御部を有し、

前記仮想化制御部が前記論理区画に対する操作指令を受け付けるステップと、

前記仮想化制御部は、前記受け付けた操作指令が前記論理区画に前記ゲストOSを実行する仮想CPUを生成する操作指令のときには、前記複数の物理CPUのうち、前記通常動作状態の物理CPUに前記仮想CPUを割り当て可能か否かを判定し、割り当て可能な場合には当該仮想CPUを前記通常動作状態の物理CPUに割り当て、前記判定結果が割り当て不能な場合にはスリープ状態の物理CPUを起動して、前記仮想CPUを当該起動した物理CPUに割り当てるステップと、

前記仮想化制御部が、仮想CPUと物理CPUの割り当て状態と物理CPUの動作状態を管理するテーブルに前記割り当てた仮想CPUを追加して更新するステップと、

前記仮想化制御部は、前記受け付けた操作指令が前記論理区画制御部から仮想CPUを削除する操作指令のときには、当該仮想CPUを前記テーブルから削除し、当該仮想CPUを削除した物理CPUに割り当てられている仮想CPUがなくなった場合には、当該物理CPUをスリープ状態に移行させるステップと、

を含み、

前記操作指令は、前記論理区画に割り当てられた当該仮想CPUが前記物理CPUを使用する比率としての割り当て率を含み、

前記仮想CPUを当該起動した物理CPUに割り当てるステップは、

新たに仮想CPUを割り当てた前記物理CPUに割り当てられている仮想CPUの割り当て率の和に基づいて前記物理CPUの動作クロックの周波数を変更し、

前記仮想CPUを前記テーブルから削除し、当該仮想CPUを削除した物理CPUに割り当てられている仮想CPUがなくなった場合には、当該物理CPUをスリープ状態に移行させるステップは、

当該物理CPUに他の仮想CPUが割り当てられているときには、割り当てられている仮想CPUの割り当て率の和に基づいて前記物理CPUの動作クロックの周波数を変更することを特徴とする仮想計算機の制御方法。

【請求項7】

前記仮想CPUを前記テーブルから削除し、当該仮想CPUを削除した物理CPUに割り当てられている仮想CPUがなくなった場合には、当該物理CPUをスリープ状態に移行させるステップは、

当該仮想CPUを前記テーブルから削除するステップと、

前記テーブルを参照して当該仮想CPUを削除した第1の物理CPUに割り当てられた第2の仮想CPUが存在し、かつ、通常動作状態にある第2の物理CPUに前記第2の仮想CPUを割り当て可能か否かを判定するステップと、

前記判定結果が割り当て可能な場合には、前記第2の仮想CPUを前記第2の物理CPUに再割り当てを行い、前記第1の物理CPUに割り当てられている仮想CPUがなくなると、当該第1の物理CPUをスリープ状態に移行させるステップと、

を含むことを特徴とする請求項6に記載の仮想計算機の制御方法。

【請求項8】

前記ゲストOSが利用する仮想CPUの実使用率を検知して前記仮想化制御部へ通知するステップをさらに含み、

前記仮想CPUを前記テーブルから削除し、当該仮想CPUを削除した物理CPUに割り当てられている仮想CPUがなくなった場合には、当該物理CPUをスリープ状態に移行させるステップは、

当該物理CPUに他の仮想CPUが割り当てられているときには、割り当てられている仮想CPUの前記実使用率の和に基づいて当該物理CPUに対する仮想CPUの割り当て

10

20

30

40

50

を変更することを特徴とする請求項 6 に記載の仮想計算機の制御方法。

【請求項 9】

前記仮想 CPU を前記テーブルから削除し、当該仮想 CPU を削除した物理 CPU に割り当てられている仮想 CPU がなくなった場合には、当該物理 CPU をスリープ状態に移行させるステップは、

当該仮想 CPU を前記テーブルから削除した後に、当該テーブルを参照して当該仮想 CPU を削除した第 1 の物理 CPU に割り当てられた第 2 の仮想 CPU が存在するときには、前記実使用率を取得して、通常動作状態にある第 2 の物理 CPU に割り当てられている仮想 CPU の実使用率の和に、前記第 2 の仮想 CPU の実使用率を加算した値が所定の閾値未満であれば、前記第 2 の仮想 CPU を前記第 2 の物理 CPU に再割り当てを行い、前記第 1 の物理 CPU に割り当てられている仮想 CPU がなくなると、当該第 1 の物理 CPU をスリープ状態に移行させることを特徴とする請求項 8 に記載の仮想計算機の制御方法。

10

【請求項 10】

前記仮想 CPU を前記テーブルから削除し、当該仮想 CPU を削除した物理 CPU に割り当てられている仮想 CPU がなくなった場合には、当該物理 CPU をスリープ状態に移行させるステップは、

前記第 2 の物理 CPU に割り当てられている仮想 CPU の実使用率の和に、前記第 2 の仮想 CPU の実使用率を加算した値が所定の閾値未満であれば、前記第 2 の物理 CPU に割り当てられている仮想 CPU の割り当て率の和に前記第 2 の仮想 CPU の割り当て率を加算した値が 100% を超えることを許容することを特徴とする請求項 9 に記載の仮想計算機の制御方法。

20

【請求項 11】

スリープ状態と通常の動作状態とを切り替え可能な物理 CPU を複数備えた物理計算機を複数の論理区画に分割し、各論理区画上でそれぞれゲスト OS を動作させて、各論理区画に対する物理計算機の資源の割当を制御するプログラムであって、

前記物理 CPU は、動作クロックの周波数を変更するクロック制御部を有し、

前記論理区画に対する操作指令を受け付けるステップと、

前記受け付けた操作指令が前記論理区画に前記ゲスト OS を実行する仮想 CPU を生成する操作指令のときには、前記複数の物理 CPU のうち、前記通常動作状態の物理 CPU に前記仮想 CPU を割り当て可能か否かを判定し、割り当て可能な場合には当該仮想 CPU を前記通常動作状態の物理 CPU に割り当て、前記判定結果が割り当て不能な場合にはスリープ状態の物理 CPU を起動して、前記仮想 CPU を当該起動した物理 CPU に割り当てるステップと、

30

前記仮想 CPU と物理 CPU の割り当て状態と物理 CPU の動作状態を管理するテーブルに前記割り当てた仮想 CPU を追加して更新するステップと、

前記受け付けた操作指令が前記論理区画制御部から仮想 CPU を削除する操作指令のときには、当該仮想 CPU を前記テーブルから削除し、当該仮想 CPU を削除した物理 CPU に割り当てられている仮想 CPU がなくなった場合には、当該物理 CPU をスリープ状態に移行させるステップと、を前記物理計算機に実行させ、

40

前記操作指令は、前記論理区画に割り当てられた当該仮想 CPU が前記物理 CPU を使用する比率としての割り当て率を含み、

前記仮想 CPU を当該起動した物理 CPU に割り当てるステップは、

新たに仮想 CPU を割り当てた前記物理 CPU に割り当てられている仮想 CPU の割り当て率の和に基づいて前記物理 CPU の動作クロックの周波数を変更し、

前記仮想 CPU を前記テーブルから削除し、当該仮想 CPU を削除した物理 CPU に割り当てられている仮想 CPU がなくなった場合には、当該物理 CPU をスリープ状態に移行させるステップは、

当該物理 CPU に他の仮想 CPU が割り当てられているときには、割り当てられている仮想 CPU の割り当て率の和に基づいて前記物理 CPU の動作クロックの周波数を変更す

50

ることを特徴とするプログラム。

【請求項 1 2】

前記仮想 CPU を前記テーブルから削除し、当該仮想 CPU を削除した物理 CPU に割り当てられている仮想 CPU がなくなった場合には、当該物理 CPU をスリープ状態に移行させるステップは、

当該仮想 CPU を前記テーブルから削除するステップと、

前記テーブルを参照して当該仮想 CPU を削除した第 1 の物理 CPU に割り当てられた第 2 の仮想 CPU が存在し、かつ、通常動作状態にある第 2 の物理 CPU に前記第 2 の仮想 CPU を割り当て可能か否かを判定するステップと、

前記判定結果が割り当て可能な場合には、前記第 2 の仮想 CPU を前記第 2 の物理 CPU に再割り当てを行い、前記第 1 の物理 CPU に割り当てられている仮想 CPU がなくなると、当該第 1 の物理 CPU をスリープ状態に移行させるステップと、

を含むことを特徴とする請求項 1 1 に記載のプログラム。

10

【請求項 1 3】

前記ゲスト OS が利用する仮想 CPU の実使用率を検知して前記仮想化制御部へ通知するステップをさらに含み、

前記仮想 CPU を前記テーブルから削除し、当該仮想 CPU を削除した物理 CPU に割り当てられている仮想 CPU がなくなった場合には、当該物理 CPU をスリープ状態に移行させるステップは、

当該物理 CPU に他の仮想 CPU が割り当てられているときには、割り当てられている仮想 CPU の前記実使用率の和に基づいて当該物理 CPU に対する仮想 CPU の割り当てを変更することを特徴とする請求項 1 1 に記載のプログラム。

20

【請求項 1 4】

前記仮想 CPU を前記テーブルから削除し、当該仮想 CPU を削除した物理 CPU に割り当てられている仮想 CPU がなくなった場合には、当該物理 CPU をスリープ状態に移行させるステップは、

当該仮想 CPU を前記テーブルから削除した後に、当該テーブルを参照して当該仮想 CPU を削除した第 1 の物理 CPU に割り当てられた第 2 の仮想 CPU が存在するときには、前記実使用率を取得して、通常動作状態にある第 2 の物理 CPU に割り当てられている仮想 CPU の実使用率の和に、前記第 2 の仮想 CPU の実使用率を加算した値が所定の閾値未満であれば、前記第 2 の仮想 CPU を前記第 2 の物理 CPU に再割り当てを行い、前記第 1 の物理 CPU に割り当てられている仮想 CPU がなくなると、当該第 1 の物理 CPU をスリープ状態に移行させることを特徴とする請求項 1 3 に記載のプログラム。

30

【請求項 1 5】

前記仮想 CPU を前記テーブルから削除し、当該仮想 CPU を削除した物理 CPU に割り当てられている仮想 CPU がなくなった場合には、当該物理 CPU をスリープ状態に移行させるステップは、

前記第 2 の物理 CPU に割り当てられている仮想 CPU の実使用率の和に、前記第 2 の仮想 CPU の実使用率を加算した値が所定の閾値未満であれば、前記第 2 の物理 CPU に割り当てられている仮想 CPU の割り当て率の和に前記第 2 の仮想 CPU の割り当て率を加算した値が 100% を超えることを許容することを特徴とする請求項 1 4 に記載のプログラム。

40

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、複数の物理計算機上でゲスト OS を稼働させる仮想計算機システムの改良に関するものである。

【背景技術】

【0002】

近年、プロセッサ (CPU) の高性能化に伴って計算機で消費される電力も増加してき

50

た。特にサーバ用途で用いられるCPUでは大容量のキャッシュメモリを搭載しているため性能も高いが消費電力も大きい。さらに、サーバでは複数のCPUを搭載しているため、消費電力がさらに大きくなる。このため、多数のサーバを運用するデータセンタ等では、サーバの消費電力が大きいとサーバの設置条件や冷却等への影響が大きくなる。

【0003】

また、サーバを用いたITシステムでは、一般に、CPUのピーク性能に応じたシステム設計を行う。このため、多数のサーバを運用するデータセンタ等では、CPUの稼働率が高くないサーバが多数存在している。ここで、近年のCPUでは、動作クロックを変更したりCPUの駆動電圧を変更して消費電力を低減する技術が採用されているが、負荷が低い状態（アイドル状態や待機状態など）であっても電力消費はゼロにはならない。このため、複数のサーバを1台のサーバに集約してCPUの稼働率を向上させるサーバ仮想化（仮想計算機）の技術に注目が集まっている。

10

【0004】

サーバ仮想化の技術としては、ホストOS上で複数のゲストOSを稼働させるVMware（登録商標）のESX Server（非特許文献1）や特許文献1に開示される技術が知られており、ハイパバイザを用いる仮想化技術としては、XenSource社のXen（非特許文献2）等が知られている。

【0005】

非特許文献1では、仮想CPUを物理CPUへ割り付ける処理は、ホストOS（VMKernel）のスケジューラに依存している。また、特許文献1では、管理OS上でゲストOSを稼働させ、ゲストOSからの指示に基づき、管理OSが物理CPUの電源制御やクロック制御を実施して仮想計算機の低消費電力を抑制しようとするものである。

20

【0006】

また、非特許文献2では、ハイパバイザがSEDF（Simple Early Deadline First）/BVT（Borrowed Virtual Timer Scheduler）のようなりアルタイムスケジューリング方法によって仮想CPUを物理CPUに割り当てるものである。

【特許文献1】特開2006-113767号公報

【非特許文献1】「VMware ESX Server」、[online]、VMware corp. 発行、[平成19年11月1日検索]、インターネット<URL: http://www.vmware.com/files/jp/pdf/esx_datasheet.pdf>

30

【非特許文献2】「Xen 3.0 Datasheet」、[online]、XenSource Inc. 発行、[平成19年11月1日検索]、インターネット<URL: http://www.vmware.com/files/jp/pdf/esx_datasheet.pdf>

【発明の開示】

【発明が解決しようとする課題】

【0007】

上記非特許文献1の従来例では、仮想CPUと物理CPUの管理がホストOS側で実施されるため、ホストOS上の仮想化ソフトウェアによって消費電力を考慮した物理CPUの制御が行えない。さらに、仮想サーバが利用する仮想CPUが、どの物理CPUに割り当てられるかについても、ホストOSのスケジューラに依存することになる。このため、仮想化ソフトウェアによる、仮想CPUと物理CPUの割当て制御による消費電力の低減は困難であるという問題があった。

40

【0008】

また、上記非特許文献2の従来例では、ハイパバイザは物理CPUに対する仮想サーバが利用する仮想CPUの割当てを制御しているため、消費電力を意識した割当ての実施は可能である。しかし、非特許文献2の技術で利用されているリアルタイムスケジューリング方法では、仮想サーバからの要求に応じて何時でも物理CPUに仮想CPUを割り当てることを可能にしている。このため、物理CPUは常に稼働状態となっており、実行する命令がない、つまり、仮想CPUが割り当てられていなくても電力を消費しているという問題があった。

50

【 0 0 0 9 】

さらに、上記特許文献 1 の従来例では、管理 OS によって物理 CPU の消費電力を低減する処理を行っているが、本処理の動作の契機はゲスト OS からの指示に基づき管理 OS が物理 CPU の電源およびクロックの制御を行っている。このため、ゲスト OS からの指示があるまでは、管理 OS は物理 CPU に対して消費電力を低減する処理を実行しない。このため、仮想計算機システム全体で消費電力を低減することができないという問題があった。

【 0 0 1 0 】

そこで本発明は、上記問題点に鑑みてなされたもので、仮想計算機システムにおいて、仮想 CPU と物理 CPU の対応付けを制御することにより、仮想 CPU が割り当てられていない物理 CPU を計算機システム全体で最適化することで消費電力の低減を図ることを目的とする。

【課題を解決するための手段】

【 0 0 1 1 】

本発明は、スリープ状態と通常の動作状態とを切り替え可能な物理 CPU を複数備えた物理計算機と、前記物理計算機を複数の論理区画に分割し、各論理区画上でそれぞれゲスト OS を動作させて、各論理区画に対する物理計算機の資源の割当を制御する仮想化制御部と、を備えた仮想計算機システムであって、前記仮想化制御部は、前記論理区画に対する操作指令を受け付ける論理区画制御部と、前記操作指令に基づいて前記論理区画に前記ゲスト OS を実行する仮想 CPU を割り当てる仮想 CPU 制御部と、前記操作指令に基づいて前記仮想 CPU を前記物理 CPU に割り当て、前記物理 CPU の動作状態を制御する仮想 - 物理 CPU 割当制御部と、前記仮想 - 物理 CPU 割当制御部からの操作指令に基づいて前記物理 CPU の動作状態を制御する物理 CPU 制御部と、を含み、前記物理 CPU は、動作クロックの周波数を変更するクロック制御部を有し、前記操作指令は、前記論理区画に割り当てられた当該仮想 CPU が前記物理 CPU を使用する比率としての割り当て率を含み、前記仮想 - 物理 CPU 割当制御部は、前記論理区画制御部から仮想 CPU を生成する操作指令を受け付けると、前記複数の物理 CPU のうち、前記通常動作状態の物理 CPU に前記仮想 CPU を割り当て可能か否かを判定し、割り当て可能な場合には当該仮想 CPU を前記通常動作状態の物理 CPU に割り当て、割り当て不可能な場合にはスリープ状態の物理 CPU を起動する指令を前記物理 CPU 制御部に発行して、前記仮想 CPU を当該起動した物理 CPU に割り当て、仮想 CPU と物理 CPU の割り当て状態と物理 CPU の動作状態を管理するテーブルに前記割り当てた仮想 CPU を追加して更新し、前記論理区画制御部から仮想 CPU を削除する操作指令を受け付けると、当該仮想 CPU を前記テーブルから削除し、当該仮想 CPU を削除した物理 CPU に割り当てられている仮想 CPU がなくなった場合には、当該物理 CPU をスリープ状態に移行するよう前記物理 CPU 制御部に指令し、前記物理 CPU に対する仮想 CPU の割り当てを変更したときには、当該物理 CPU に割り当てられている仮想 CPU の割り当て率の和に基づいて前記物理 CPU の動作クロックの周波数を決定し、当該動作クロックの周波数を前記物理 CPU 制御部へ指令し、前記物理 CPU 制御部は前記動作クロックの周波数を前記クロック制御部へ指令する。

【 0 0 1 2 】

また、前記仮想 - 物理 CPU 割当制御部は、前記論理区画制御部から仮想 CPU を削除する操作指令を受け付けると、当該仮想 CPU を前記テーブルから削除した後に、当該テーブルを参照して当該仮想 CPU を削除した第 1 の物理 CPU に割り当てられた第 2 の仮想 CPU が存在し、かつ、通常動作状態にある第 2 の物理 CPU に前記第 2 の仮想 CPU を割り当て可能な場合には、前記第 2 の仮想 CPU を前記第 2 の物理 CPU に再割り当てを行い、前記第 1 の物理 CPU に割り当てられている仮想 CPU がなくなると、当該第 1 の物理 CPU をスリープ状態に移行させる。

【 0 0 1 3 】

また、前記仮想 - 物理 CPU 割当制御部は、前記論理区画制御部から仮想 CPU を削除

10

20

30

40

50

する操作指令を受け付けると、当該仮想CPUを前記テーブルから削除した後に、当該テーブルを参照して当該仮想CPUを削除した第1の物理CPUに割り当てられた第2の仮想CPUが存在するときには、前記稼働状態検出部からの稼働率を取得して、通常動作状態にある第2の物理CPUに割り当てられている仮想CPUの稼働率の和に、前記第2の仮想CPUの稼働率を加算した値が所定の閾値未満であれば、前記第2の仮想CPUを前記第2の物理CPUに再割り当てを行い、前記第1の物理CPUに割り当てられている仮想CPUがなくなると、当該第1の物理CPUをスリープ状態に移行させる。さらに、前記第2の物理CPUに割り当てられている仮想CPUの稼働率の和に、前記第2の仮想CPUの稼働率を加算した値が所定の閾値未満であれば、前記第2の物理CPUに割り当てられている仮想CPUの割り当て率の和に前記第2の仮想CPUの割り当て率を加算した値が100%を超えることを許容する。

10

【発明の効果】

【0014】

したがって、本発明は、論理区画に仮想CPUを割り当てた後に、仮想CPUを削除する際には、仮想CPUを削除した物理CPUに他の仮想CPUが割り当てられていない場合にはこの物理CPUをスリープ状態にして、物理計算機の消費電力を積極的に抑制することが可能となる。

【0015】

また、仮想CPUを削除した第1の物理CPUに割り当てられている第2の仮想CPUが存在し、かつ、通常動作状態にある第2の物理CPUに前記第2の仮想CPUを割り当て可能な場合には、前記第2の仮想CPUを前記第2の物理CPUに再割り当てを行い、前記第1の物理CPUに割り当てられている仮想CPUがなくなると、当該第1の物理CPUをスリープ状態に移行させることで、積極的にスリープ状態の物理CPUを生成することで、仮想CPUを割り当て最適化して物理計算機の消費電力を低減することができる。

20

【0016】

さらに、仮想CPUの稼働率を取得して、前記第2の仮想CPUを前記第2の物理CPUに再割り当てを行う際には、第2の物理CPUに割り当てる仮想CPUの稼働率が閾値未満であれば仮想CPUの割り当て率が100%を超えることを許容するので、さらに効率よく仮想CPUの再配置を行い、スリープ状態の物理CPUの数を最大にして物理計算機の消費電力を低減することができる。

30

【発明を実施するための最良の形態】

【0017】

以下、本発明の一実施形態を添付図面に基づいて説明する。

【0018】

図1は、第1の実施形態を示し、本発明を適用する仮想計算機システムの一例を示すブロック図である。

【0019】

仮想計算機システムは、ゲストOS（またはアプリケーション：AP）100、110を実行する仮想CPU201、202、211、212が割り当てられた論理区画（Logical PARTition）200、210と、これら複数の論理区画200、210及び仮想CPU201、202、211、212を提供する仮想化ソフトウェア300と、仮想化ソフトウェア300を実行する物理計算機400と、物理計算機400に接続されて仮想化ソフトウェア300に操作指令を送信する管理コンソール500と、を備える。なお、論理区画200、210で稼働するゲストOS100、110が仮想サーバを提供する。

40

【0020】

<物理計算機の構成>

物理計算機400は、演算処理を行う物理CPU410、411、412と、データやプログラムを格納するメモリ420と、物理CPU410～412及びメモリ420へ電力を供給する電源供給部430と、物理CPU410～412及びメモリ420へクロック

50

クを供給するクロック供給部を備える。また、物理計算機 400 は、図示しないインターフェースを介して接続された管理コンソール 500 を有し、この管理コンソール 500 は、表示装置と入力装置を備える。メモリ 420 には仮想化ソフトウェア 300 がロードされ、物理 CPU 410 ~ 412 により実行される。

【0021】

なお、物理計算機 400 は、図示はしないがストレージ装置やネットワークインターフェースなどの I/O 装置を備えることができる。

【0022】

ここで、物理 CPU 410 ~ 412 は、仮想化ソフトウェア 300 からの操作指令に応じてクロックと電力を変更可能な省電力機構を備える。例えば、物理 CPU 410 は、図 2 で示すように、複数のプロセッサコア 4101, 4102 と、データや命令を一時的に格納するキャッシュ 4103 と、電源供給部 430 から入力された電力と、クロック供給部 440 から入力されたクロックをプロセッサコア 4101, 4102 及びキャッシュに供給するクロック・電力制御部 4104 を備え、このクロック・電力制御部 4104 が後述するように省電力機構を含む。なお、物理 CPU 411, 412 も図 2 の物理 CPU 410 と同一の構成である。

【0023】

クロック・電力制御部 4104 は、仮想化ソフトウェア 300 から受け付けた指令に基づいて、プロセッサコア 4101, 4102 に供給する電圧とクロックを変更し、プロセッサコア 4101, 4102 の動作モードを動的に変更する。クロック・電力制御部 4104 の電圧とクロックの変更は、プロセッサコア 4101, 4102 のみに電力とクロックを遮断し、その他の部分に電源供給を継続して再起動可能な状態を維持する動作状態を含む。また、クロック・電力制御部 4104 は、プロセッサコア 4101, 4102 が通常の稼働状態のときに、プロセッサコア 4101, 4102 の処理能力が最大となる定格クロック及び定格電圧から、プロセッサコア 4101, 4102 の処理能力が最小となる最低クロック及び最低電圧の間で電圧及びクロックを変更し、物理 CPU 410 の処理能力を優先させる動作状態から消費電力を低減させる動作状態の間で動作状態を変更させることができる。なお、クロック・電力制御部 4104 は、プロセッサ動作状態コア 4101, 4102 毎にクロック周波数と電圧を変更可能である。

【0024】

次に、仮想化ソフトウェア 300 は、物理 CPU 410 ~ 412 のクロック・電力制御部 4104 に動作モードを設定することで、物理 CPU 410 ~ 412 の動作モードを変更する。ここで、物理 CPU 410 ~ 412 の動作モードは、ACPI (Advanced Configuration and Power Interface Specification) で規定されるプロセッサの動作モード C0 ~ C3 を用いることができる。なお、プロセッサの動作モード C0 ~ C3 は、ACPI の Rev 3.0 で規定されるものである (<http://www.acpi.info/spec30a.htm>、ADVANCED CONFIGURATION AND POWER INTERFACE SPECIFICATION、Revision 3.0a, December 30, 2005)。プロセッサ (物理 CPU 410 ~ 412) の動作モード C0 ~ C3 は、次のとおりである。

- ・ C0 : 動作状態 (通常稼働状態)
- ・ C1 : アイドル状態 (スリープ状態) : 物理 CPU の待機状態で、動作クロックの停止や、HLT 命令などによる命令実行の停止状態を示す
- ・ C2 : C1 よりも低消費電力の状態
- ・ C3 : スタンバイ状態 (プロセッサコア 4101, 4102 への電源遮断状態。再起動に必要な電力は継続して供給される)

なお、物理 CPU 410 ~ 412 の動作モードは、上記 ACPI の C0 ~ C3 に代わって、システムの動作状態を示す S0 ~ S3 を用いても良い。S0 ~ S3 の動作モードは、次のとおりである。

- ・ S0 : 動作状態 (通常稼働状態)
- ・ S1 : スリープ状態。VGA 信号はオフ、物理 CPU はパワーダウン。

- ・ S 2 : 物理 CPU の電源遮断 (再起動に必要な電力は継続して供給される)。
- ・ S 3 スタンバイ状態。作業内容をメモリに書き込み、メモリへの通電と再起動に必要な電流のみ流れる。

【 0 0 2 5 】

仮想化ソフトウェア 3 0 0 は、物理 CPU 4 1 0 ~ 4 1 2 のクロック・電力制御部 4 1 0 4 に指令を送ることで、物理 CPU 4 1 0 ~ 4 1 2 の動作モードを、定格の動作クロックの周波数で処理を行う通常モードと、処理の実行を停止するスリープ状態のスリープモードと、処理を実行するが動作クロックの周波数を低下させる低消費電力モードの何れかに変更することができる。

【 0 0 2 6 】

< 仮想化ソフトウェアの構成 >

図 1 において、仮想化ソフトウェア 3 0 0 は、例えば、ハイパバイザなどホスト OS を利用しないソフトウェアで構成されており、物理計算機 4 0 0 の計算機資源を複数の論理区画 2 0 0 (L P A R # 1)、2 1 0 (L P A R # 2) に分割する L P A R 制御部 3 1 0 と、論理区画 2 0 0 , 2 1 0 に物理 CPU 4 1 0 ~ 4 1 2 を仮想化した仮想 CPU 2 0 1 , 2 0 2 , 2 1 1 , 2 1 2 を割り当てる仮想 CPU 制御部 3 2 0 と、物理 CPU 4 1 0 ~ 4 1 2 を制御する物理 CPU 制御部 3 3 0 と、仮想 CPU と物理 CPU の割り当て状況を管理する仮想 - 物理 CPU 割当制御部 3 4 0 とを含む。この仮想化ソフトウェア 3 0 0 は、メモリ 4 2 0 にロードされて物理 CPU 4 1 0 ~ 4 1 2 により実行される。また、仮想化ソフトウェア 3 0 0 は、図示しないストレージ装置や不揮発性メモリ等の記憶媒体に格納される。

【 0 0 2 7 】

L P A R 制御部 3 1 0 は、システム管理者等が操作する管理コンソール 5 0 0 からの操作指令に応じて論理区画 2 0 0 , 2 1 0 の生成や不要になった論理区画の削除を行う。L P A R 制御部 3 1 0 は、管理コンソール 5 0 0 からの操作指令に応じて論理区画に対して物理計算機 4 0 0 の計算機資源を割り当てる。管理コンソール 5 0 0 から設定する計算機資源としては、メモリ 4 2 0 の割当量 (アドレス範囲など) と仮想 CPU 及び I / O 装置を各論理区画 2 0 0 , 2 1 0 毎に設定する。なお、L P A R 制御部 3 1 0 が、物理計算機 4 0 0 の計算機資源を割り当てる手法については、公知の手法を用いることができる。

【 0 0 2 8 】

L P A R 制御部 3 1 0 では、生成した論理区画 2 0 0 , 2 1 0 に仮想 CPU を割り当てて、論理区画 2 0 0 , 2 1 0 のそれぞれでゲスト OS 1 0 0 , 1 1 0 を動作させ、各ゲスト OS 1 0 0 , 1 1 0 上でそれぞれアプリケーション (A P) を動作させる。

【 0 0 2 9 】

仮想 CPU 制御部 3 2 0 は、L P A R 制御部 3 1 0 から要求された仮想 CPU の生成と削除 (論理区画の生成と削除に対応) と、仮想 CPU の稼動状態の管理を行う。このため、仮想 CPU 制御部 3 2 0 は、論理区画 2 0 0 , 2 1 0 と仮想 CPU 2 0 1 ~ 2 1 2 の対応関係を保持する論理区画対応表 3 4 2 を備えている。

【 0 0 3 0 】

仮想 CPU 制御部 3 2 0 では、L P A R 制御部 3 1 0 から仮想 CPU の生成を受け付けた場合は、当該仮想 CPU の割り当て先となる物理 CPU の決定を仮想 - 物理 CPU 割当制御部 3 4 0 に仮想 CPU の生成を指令し、仮想 CPU の生成が完了した通知を受けると論理区画対応表 3 4 2 に仮想 CPU を割り当てた論理区画の情報を更新する。

【 0 0 3 1 】

また、仮想 CPU 制御部 3 2 0 が L P A R 制御部 3 1 0 から仮想 CPU の削除を受け付けた場合は、当該仮想 CPU の割り当て先である物理 CPU と、当該消滅対象の仮想 CPU の対応関係を削除するよう仮想 - 物理 CPU 割当制御部 3 4 0 に指令し、このとき、仮想 - 物理 CPU 割当制御部 3 4 0 は、当該物理 CPU に割り当てられている仮想 CPU がなくなる場合には、物理 CPU 制御部 3 3 0 へ当該物理 CPU の動作モードを省電力モードへ移行するように指令する。さらに、仮想 CPU 制御部は、削除した仮想 CPU を論理

10

20

30

40

50

区画対応表 3 4 2 から削除する。なお、物理 CPU の動作モード及び省電力モードについては後述する。

【 0 0 3 2 】

さらに、仮想 CPU 制御部 3 2 0 は、各論理区画 2 0 0 , 2 1 0 に割り当てられた仮想 CPU 2 0 1 , 2 0 2 , 2 1 1 , 2 1 2 の稼働状態 (割当状態) を監視して論理区画対応表 3 4 2 を更新し、取得した稼働状態を仮想 - 物理 CPU 割当制御部 3 4 0 へ通知する。

【 0 0 3 3 】

物理 CPU 制御部 3 3 0 は、物理 CPU 4 1 0 ~ 4 1 2 の稼働状態 (動作モード) を監視して、仮想 - 物理 CPU 割当制御部 3 4 0 からの要求に応じて物理 CPU 4 1 0 ~ 4 1 2 の動作モードを制御する。物理 CPU 制御部 3 3 0 は、物理 CPU 4 1 0 ~ 4 1 2 に対して仮想 CPU が割り当てられていない場合には、該当する物理 CPU の動作モードを省電力モードに変更する。また、仮想 - 物理 CPU 割当制御部 3 4 0 から物理 CPU 4 1 0 ~ 4 1 2 に新たに仮想 CPU を割り当ててる場合、該当する物理 CPU 4 1 0 ~ 4 1 2 が省電力モードの場合には、当該物理 CPU の動作モードを稼働状態へ移行させる。

【 0 0 3 4 】

仮想 - 物理 CPU 割当制御部 3 4 0 は、仮想 CPU 制御部 3 2 0 からの要求に応じて物理 CPU 4 1 0 ~ 4 1 2 に対する仮想 CPU の割当状況を動的に管理する。このため、仮想 - 物理 CPU 割当制御部 3 4 0 は、仮想 CPU 2 0 1 , 2 0 2 , 2 1 1 , 2 1 2 と物理 CPU 4 1 0 ~ 4 1 2 の対応関係を設定した仮想 - 物理 CPU 割当表 4 3 1 を保持する。

【 0 0 3 5 】

仮想 - 物理 CPU 割当制御部 3 4 0 は、仮想 CPU 制御部 3 2 0 から仮想 CPU の稼働状態を取得し、物理 CPU 制御部 3 3 0 から物理 CPU 4 1 0 ~ 4 1 2 の稼働状態を取得し、仮想 CPU 制御部 3 2 0 からの要求と、仮想 CPU の稼働状態の変化に応じて、物理 CPU 4 1 0 ~ 4 1 2 に対する仮想 CPU の割り当てを動的に変更する。このため、仮想 - 物理 CPU 割当制御部 3 4 0 は、物理 CPU 4 1 0 ~ 4 1 2 に割り当てた仮想 CPU を管理する仮想 - 物理 CPU 割当表 3 4 1 を備える。

【 0 0 3 6 】

本第 1 の実施形態においては、仮想 - 物理 CPU 割当制御部 3 4 0 は、仮想 CPU 制御部 3 2 0 からの要求と、仮想 CPU 2 0 1 , 2 0 2 , 2 1 1 , 2 1 2 の割当率に基づいて物理 CPU 4 1 0 ~ 4 1 2 に対する仮想 CPU の割り当てを動的に変更し、物理 CPU 4 1 0 ~ 4 1 2 の動作モードを動的に変更する例を示す。

【 0 0 3 7 】

< 仮想 CPU の割り当て >

次に、仮想化ソフトウェア 3 0 0 が行う仮想 CPU の割り当てについて説明する。仮想化ソフトウェア 3 0 0 が物理計算機 4 0 0 の計算機資源を各論理区画 2 0 0 , 2 1 0 に割当てる方法は、専用割当と共用割当の二種類ある。

【 0 0 3 8 】

専用割当は、特定の計算機資源を特定の論理区画に専用的に割当てる方法である。計算機資源のうち、メモリ 4 2 0 (主記憶) と I / O 装置は共用割当される。

【 0 0 3 9 】

なお、物理 CPU 4 1 0 ~ 4 1 2 を専用割当にすることもできる。物理 CPU 4 1 0 ~ 4 1 2 の専用割当の場合、ある論理区画に対して専用割当する物理 CPU 4 1 0 ~ 4 1 2 の数をその L P A R に対する CPU の割当量と呼ぶ。

【 0 0 4 0 】

一方、共用割当は、計算機資源を各論理区画に少しづつ時分割によって割り当てる。共有割当では、ある論理区画の仮想 CPU に対して物理 CPU 4 1 0 ~ 4 1 2 を割当てている時間の、全 L P A R に CPU を割当てている時間に対する割合を、仮想 CPU の割当率と呼ぶ (% で表す。値は 0 ~ 1 0 0 の間) 。

【 0 0 4 1 】

仮想化ソフトウェア 3 0 0 の仮想 - 物理 CPU 割当制御部 3 4 0 が管理する仮想 - 物理

10

20

30

40

50

CPU割当表341は、図3で示すように、物理CPU410～412のそれぞれについて割り当てられている仮想CPUの識別子と割り当て率が格納される。

【0042】

図3において、仮想-物理CPU割当表341は、物理CPU410～412の識別子を格納する物理CPU番号3410と、物理CPU番号3410で特定される物理CPUの動作状態を格納する動作状態3411と、複数の仮想CPUの情報を格納する仮想CPUリストから1つのエントリが構成される。仮想CPUリストは、物理CPU番号3410で特定される物理CPUに割り当てられている仮想CPUの識別子を格納する仮想CPU番号3412と、仮想CPU番号3412で特定される仮想CPUの割り当て率を格納する割り当て率3413で構成され、仮想CPUの数に応じて仮想CPU番号3412と割り当て率3413が繰り返して格納される。

10

【0043】

例えば、図3の物理CPU番号3410=1のエントリには、物理CPUの動作状態3411が通常の稼働状態（例えば、クロック周波数及び供給電圧が定格）であることが示されており、この物理CPUには、仮想CPU番号が1の仮想CPUが割り当て率40%で割り当てられ、さらに、仮想CPU番号が3の仮想CPUが割り当て率20%で割り当てられていることを示す。この物理CPU番号=1の物理CPUは、残りの40%の割り当て率で仮想CPUを割り当てることが可能であることを示している。仮想CPUリストには、当該物理CPU番号3410に割り当てられた仮想CPUの数に応じた仮想CPU番号3412と割り当て率3413がペアで格納される。

20

【0044】

なお、図3と図1との対応関係は、物理CPU410～412が物理CPU番号=1～3に対応し、仮想CPU201, 202, 211, 212が、仮想CPU番号=1～4に対応する。

【0045】

また、本第1実施形態では、物理CPUの動作状態3411は、定格で動作する「通常」と、低消費電力状態である「スリープ」の何れかに設定され、「スリープ」は上記ACPIで定義されるC1以上(C1～C3)またはS1以上(S1～S3)の状態であればよい。なお、以下の説明においても、スリープ状態は上記ACPIのプロセッサの動作モードC1～C3(またはS1～S3)を指すものである。

30

【0046】

仮想CPU制御部320が管理する論理区画対応表342は、図4で示すように構成される。図4において、論理区画対応表342は、論理区画の識別子を格納する論理区画番号3420と、論理区画番号3420で特定される論理区画に割り当てられている仮想CPUの識別子を格納する仮想CPUリストから構成される。仮想CPUリストは、論理区画番号3420で特定され論理区画に割り当てられている仮想CPUの識別子を格納する仮想CPU番号3421で構成され、ひとつの論理区画に複数の仮想CPUが割り当てられている場合には、仮想CPU番号3421も複数設定される。図示の例では、論理区画番号3402=1の論理区画201には、2つの仮想CPU#1と#2が割り当てられ、論理区画番号3402=2の論理区画210には、2つの仮想CPU#3と#4が割り当てられている状態を示している。

40

【0047】

<仮想化ソフトウェアの動作>

次に、仮想化ソフトウェア300で行われる処理の一例について、以下に説明する。図5は、仮想化ソフトウェア300で行われる仮想サーバ(または仮想CPU)を生成する処理の一例を示すフローチャートである。このフローチャートは、仮想化ソフトウェア300が管理コンソール500から指定された論理区画200, 210に仮想サーバ(または仮想CPU)を生成する操作指令を受け付けたときに実行される処理である。なお、この処理を実行する以前に、管理コンソール500からの操作指令によって既に仮想化ソフトウェア300が論理区画201, 202を設定したものとする。

50

【 0 0 4 8 】

仮想化ソフトウェア 3 0 0 は、管理コンソール 5 0 0 から仮想サーバ（または仮想 CPU）を生成する論理区画と仮想 CPU の数と割り当て率を受け付けると、ステップ S 1 で仮想 - 物理 CPU 割当表 3 4 1 を参照し、仮想 CPU を実行可能な動作状態の物理 CPU 4 1 0 ~ 4 1 2 を参照する。ステップ S 2 では、上記ステップ S 1 で取得した動作状態 3 4 1 1 が「通常」となっている物理 CPU 4 1 0 ~ 4 1 2 が存在するか否かを判定する。動作状態 3 4 1 1 が「通常」の物理 CPU が存在すればステップ S 3 へ進み、存在しなければステップ S 6 に進む。

【 0 0 4 9 】

次に、ステップ S 3 では、動作状態 3 4 1 1 が「通常」の物理 CPU の中から、管理コンソール 5 0 0 から要求された仮想 CPU を割り当て可能であるか否かを判定する。この判定は、動作状態 3 4 1 1 が「通常」となっている物理 CPU のうち、ひとつの物理 CPU に割り当てられている仮想 CPU の割り当て率の総和を 1 0 0 % から差し引いた値が、生成する仮想 CPU の割り当て率以上であれば、管理コンソール 5 0 0 から要求された仮想 CPU を当該物理 CPU で割り当て可能と判定してステップ S 4 に進む。一方、当該物理 CPU に割り当てられている仮想 CPU の割り当て率の総和を 1 0 0 % から差し引いた値が、生成する仮想 CPU の割り当て率未満であれば、管理コンソール 5 0 0 から要求された仮想 CPU を当該物理 CPU に割り当て不能と判定してステップ S 5 に進む。ステップ S 5 では、要求された仮想 CPU を割り当てる対象の物理 CPU を他の物理 CPU に切り替えてからステップ S 2 に戻り、再度仮想 CPU を割り当て可能な物理 CPU を判定する。なお、ステップ S 5 ~ S 2 のループは、仮想 - 物理 CPU 割当表 3 4 1 のエントリ毎に順次判定を行うようにすればよい。

【 0 0 5 0 】

ステップ S 4 では、ステップ S 3 で選択した物理 CPU に管理コンソール 5 0 0 から要求された仮想 CPU を割指定された論理区画に割り当てて、論理区画対応表 3 4 2 の該当する論理区画に仮想 CPU の識別子を書き込む。そして、仮想化ソフトウェア 3 0 0 は、仮想 - 物理 CPU 割当表 3 4 1 の該当物理 CPU のエントリに、新たに割り当てる仮想 CPU の番号と割り当て率を書き込んで更新してから、新たな仮想 CPU を起動する。

【 0 0 5 1 】

一方、上記ステップ S 2 の判定で、動作状態 3 4 1 1 が「通常」で、かつ新たな仮想 CPU を割り当て可能な物理 CPU がない場合のステップ S 6 では、仮想 - 物理 CPU 割当表 3 4 1 を参照して動作状態 3 4 1 1 が「スリープ」となっている物理 CPU があるか否かを判定する。動作状態 3 4 1 1 が「スリープ」の物理 CPU がある場合には、ステップ S 7 へ進んで、当該物理 CPU の動作状態 3 4 1 1 を「通常」に変更して起動する。そして、当該起動した物理 CPU を仮想 CPU を割り当てる対象としてステップ S 4 に進み、上記と同様に割り当てを実施する。

【 0 0 5 2 】

上記ステップ S 6 の判定で、動作状態 3 4 1 1 が「スリープ」の物理 CPU がない場合には、仮想 CPU を割り当てる物理 CPU が存在しないため、仮想化ソフトウェア 3 0 0 は管理コンソール 5 0 0 に対して新たな仮想 CPU を割り当てる計算機資源が不足したことを示す警告を通知する。

【 0 0 5 3 】

以上の処理により、管理コンソール 5 0 0 から新たな仮想 CPU の生成要求があると、仮想化ソフトウェア 3 0 0 は、まず、動作状態 3 4 1 1 が「通常」の稼働中の物理 CPU から割当先を探索し、稼働中の物理 CPU に仮想 CPU を割り当てることができないときにのみ動作状態 3 4 1 1 が「スリープ」の物理 CPU を起動して新たな仮想 CPU を割り当てる。これにより、仮想 CPU を割り当てるのに必要最小限の物理 CPU のみを「通常」の動作状態 3 4 1 1 で稼働させて、動作状態 3 4 1 1 が「スリープ」の物理 CPU を再起動させるのを極力抑制することで、物理計算機 4 0 0 の消費電力が増大するのを抑制できる。

10

20

30

40

50

【 0 0 5 4 】

図 6 は、仮想化ソフトウェア 3 0 0 が管理コンソール 5 0 0 から、指定された仮想サーバ（または仮想 CPU）を削除する操作指令を受け付けたときに実行される処理の一例を示すフローチャートである。

【 0 0 5 5 】

仮想化ソフトウェア 3 0 0 は、管理コンソール 5 0 0 から仮想サーバ（または仮想 CPU）を削除する仮想 CPU の識別子を受け付けると、ステップ S 1 1 で仮想 - 物理 CPU 割当表 3 4 1 を参照し、指定された仮想 CPU を実行している物理 CPU のエントリから該当する仮想 CPU の識別子と割当率を削除する。また、仮想化ソフトウェア 3 0 0 は、論理区画対応表 3 4 2 を参照し、管理コンソール 5 0 0 から指定された仮想 CPU の識別子があれば当該識別子を削除する。

10

【 0 0 5 6 】

ステップ S 1 2 では、上記ステップ S 1 1 で仮想 CPU の識別子と割当率を削除した物理 CPU のエントリで、割り当てられている仮想 CPU があるか否かを判定する。仮想 CPU が存在する場合には処理を終了し、割り当てられている仮想 CPU が存在しなければステップ S 1 3 に進む。

【 0 0 5 7 】

ステップ S 1 3 では、ステップ S 1 1 で仮想 CPU の識別子を削除した物理 CPU の動作状態をスリープ状態とするよう物理 CPU 制御部 3 3 0 に指令してから処理を終了する。

20

【 0 0 5 8 】

以上の処理により、仮想 CPU の削除時には、仮想 CPU の割り当てがなくなった物理 CPU をスリープ状態にして物理計算機 4 0 0 の消費電力を低減することができる。

【 0 0 5 9 】

このように、本発明の第 1 の実施形態によれば、サーバの仮想化を行う仮想計算機システムにおいて、仮想 CPU と物理 CPU の対応付けを制御して、仮想 CPU が割り当てられていない物理 CPU をスリープ状態へ移行させることで、仮想計算機システム全体の消費電力の低減を図ることができる。

【 0 0 6 0 】

なお、上記第 1 実施形態では、マルチコアの物理 CPU 4 1 0 ~ 4 1 2 に仮想 CPU を割り当てる際に、物理 CPU 毎に仮想 CPU を割り当てる例を示したが、プロセッサコア毎に仮想 CPU を割り当ててもよい。

30

【 0 0 6 1 】

< 第 2 実施形態 >

図 7 ~ 図 9 は第 2 の実施形態を示し、前記第 1 実施形態に示した物理 CPU 制御部 3 3 0 が物理 CPU 4 1 0 ~ 4 1 2 の動作クロックを制御し、さらに、仮想 - 物理 CPU 割当制御部 3 4 0 が、仮想 CPU を削除する際に、スリープ状態の物理 CPU の数を最大にするよう仮想 CPU の割り当てを制御するものである。物理計算機 4 0 0 の構成は前記第 1 実施形態と同様であり、仮想化ソフトウェア 3 0 0 は、仮想 - 物理 CPU 割当制御部 3 4 0 と物理 CPU 制御部 3 3 0 の機能の一部が前記第 1 実施形態と異なり、仮想 - 物理 CPU 割当表 3 4 1 A の構成の一部が前記第 1 実施形態と異なる。

40

【 0 0 6 2 】

図 7 は、仮想 - 物理 CPU 割当制御部 3 4 0 が管理する仮想 - 物理 CPU 割当表 3 4 1 A の構成を示す説明図である。本第 2 実施形態の仮想 - 物理 CPU 割当表 3 4 1 A は、前記第 1 実施形態の図 3 に示した仮想 - 物理 CPU 割当表 3 4 1 に物理 CPU 4 1 0 ~ 4 1 2 の動作クロックの周波数を格納する動作周波数 3 4 1 4 を加えたもので、その他は前記第 1 実施形態と同様である。動作周波数 3 4 1 4 は、物理 CPU 4 1 0 ~ 4 1 2 の定格動作時のクロック周波数に対する実際に設定された動作クロックの周波数の比率を百分率で表した値である。

【 0 0 6 3 】

50

物理CPU 410～412は、前記第1実施形態の図2で示したように、クロック供給部440から供給されたクロックを制御するクロック・電力制御部4104を備えて、動作クロックの周波数を変更可能に構成される。物理CPU 410～412の動作クロック周波数の変更は、物理CPU割当制御部340が物理CPU 410～412のクロック・電力制御部4104に指示する周波数の値を変更することで実現される。

【0064】

また、図2において、プロセッサコア4101, 4102毎に動作クロックの周波数を独立して制御する場合には、仮想-物理CPU割当表341Aのエントリを物理CPU 410～412のプロセッサコア毎に設けてもよい。

【0065】

次に、仮想-物理CPU割当制御部340は、物理CPU 410～412に仮想-物理CPUを割り当てる際に、仮想CPUの割り当て率に応じて動作クロックの周波数を設定し、物理CPU制御部330に動作クロックの周波数を指令する。この動作クロックの周波数の設定は、例えば、物理CPU 410～412に割り当てられている仮想CPUの割り当て率の合計値を動作周波数3414として設定することができる。つまり、図7において、物理CPU番号3410が「1」のエントリでは、仮想CPU#1の割り当て率3413が40%で、仮想CPU#3の割り当て率3413が20%の場合、割り当て率3413の合計値である60%を物理CPU#1の動作周波数3414とすればよい。さらに、物理CPU 410～412の動作クロックの周波数を変更する条件として、当該物理CPUに割り当てられている仮想CPUの割り当て率3413の合計値が所定の第1の閾値（例えば、90%）未満であれば、動作クロックの周波数を変更することができる。なお、第1の閾値の値は、ゲストOS 100, 110や、アプリケーションの特性などに応じて適宜設定すればよい。

【0066】

また、仮想-物理CPU割当制御部340は、後述するように、仮想CPUを削除する際には、スリープ状態となる物理CPU 410～412の数が最大となるよう、仮想CPUの割り当てを変更し、物理計算機400の消費電力が最小となるよう仮想CPUの割り当てを最適化する。

【0067】

図8は、仮想化ソフトウェア300で行われる仮想サーバ（または仮想CPU）を生成する処理の一例を示すフローチャートである。このフローチャートは、前記第1実施形態の図5と同様に、仮想化ソフトウェア300が管理コンソール500から指定された論理区画200, 210に仮想サーバ（または仮想CPU）を生成する操作指令を受け付けたときに実行される処理である。なお、この処理を実行する以前に、管理コンソール500からの操作指令によって既に仮想化ソフトウェア300が論理区画201, 202を設定したものとする。なお、前記第1実施形態の図5と同様の処理については、同一の符号を付した。

【0068】

本第2実施形態の図8は、前記第1実施形態の図5に示したステップS3とS4の間に、物理CPU 410～412の動作クロックの周波数を変更するステップS21とステップS22を加えたものであり、その他の構成は前記図5と同様である。

【0069】

管理コンソール500から新たな仮想CPUの生成要求があると、仮想化ソフトウェア300は、まず、ステップS1～S3で、仮想-物理CPU割当表341Aを参照して動作状態3411が「通常」の稼働中の物理CPUから割当先を探索し、稼働中の物理CPUに仮想CPUを割り当てることができないときのみステップS6, S7で動作状態3411が「スリープ」の物理CPUを起動して新たな仮想CPUを割り当てる。また、スリープ状態の物理CPUがない場合にはステップS8で警告を通知する。

【0070】

ステップS3またはステップS7で指定された仮想CPUを割り当てる物理CPU 41

10

20

30

40

50

0 ~ 4 1 2 が決定すると、ステップ S 2 1 に進んで当該物理 CPU 4 1 0 ~ 4 1 2 の動作クロックの周波数を変更するか否かを判定する。すなわち、仮想化ソフトウェア 3 0 0 は、仮想 - 物理 CPU 割当表 3 4 1 A を参照し、選択した物理 CPU に既に割り当てられている仮想 CPU の割り当て率 3 4 1 3 の合計と、管理コンソール 5 0 0 から指令された新たな仮想 CPU の割り当て率との和が所定の第 1 の閾値（例えば、90%）未満であるか否かを判定する。前記和の値が第 1 の閾値未満であれば物理 CPU 4 1 0 ~ 4 1 2 の動作クロックの周波数の変更を実施すると判定してステップ S 2 2 へ進む。一方、前記和の値が第 1 の閾値以上であれば物理 CPU 4 1 0 ~ 4 1 2 の動作クロックの周波数を変更しないと判定してステップ S 4 へ進む。

【 0 0 7 1 】

ステップ S 2 2 では、上述したように、対象の物理 CPU に割り当てる仮想 CPU の割り当て率の合計値を当該物理 CPU の動作クロックの周波数の定格時に対する比率として設定する。そして、仮想化ソフトウェア 3 0 0 は、この周波数の比率に対応する周波数となるよう物理 CPU に動作クロックの周波数を変更するよう指令する。これにより、物理 CPU へ新たに仮想 CPU を割り当てると、割り当て率の和の増加に応じて動作クロックの周波数を増加させることができ、仮想 CPU の実行に必要な最小限の消費電力で物理 CPU を動作させることが可能となって、物理計算機 4 0 0 の消費電力を抑制することができる。

【 0 0 7 2 】

そして、ステップ S 4 では、上記ステップ S 3 または S 7 で選択した物理 CPU に対応する仮想 - 物理 CPU 割当表 3 4 1 A のエントリに、新たに割り当てた仮想 CPU の識別子と割り当て率を仮想 CPU 番号 3 4 1 2 と割り当て率 3 4 1 3 に書き込み、ステップ S 2 2 で設定した動作クロックの周波数の比率を動作周波数 3 4 1 4 に格納して処理を終了する。

【 0 0 7 3 】

以上の処理により、仮想 CPU の生成時には、物理 CPU 4 1 0 ~ 4 1 2 にそれぞれ割り当てられている仮想 CPU の割り当て率等に応じて物理 CPU 4 1 0 ~ 4 1 2 の動作クロックの周波数を変更することができ、物理計算機 4 0 0 の消費電力をさらに抑制することが可能となる。

【 0 0 7 4 】

次に、図 9 は、仮想化ソフトウェア 3 0 0 が管理コンソール 5 0 0 から、指定された仮想サーバ（または仮想 CPU）を削除する操作指令を受け付けたときに実行される処理の一例を示すフローチャートである。この処理は、前記第 1 実施形態の図 6 に示したフローチャートにステップ S 3 1 ~ S 3 3 の処理を追加したもので、ステップ S 1 1 ~ S 1 3 は前記第 1 実施形態の図 6 と同様である。

【 0 0 7 5 】

ステップ S 1 1 ~ S 1 3 の処理では、前記第 1 実施形態の図 6 と同様にして、仮想 CPU の削除時には、仮想 CPU の割り当てがなくなった物理 CPU をスリープ状態にして物理計算機 4 0 0 の消費電力を低減する。

【 0 0 7 6 】

さらに、ステップ S 1 2 の判定で管理コンソール 5 0 0 から指令された仮想 CPU を削除した物理 CPU に、他の仮想 CPU が割り当てられている場合にはステップ S 3 1 へ進む。

【 0 0 7 7 】

ステップ S 3 1 では、仮想 - 物理 CPU 割当表 3 4 1 A を参照して、動作状態 3 4 1 1 が「通常」となっている稼働中の物理 CPU が他にあるか否かを判定する。他の物理 CPU が通常の動作状態で稼働中であればステップ S 3 2 へ進み、他に通常の動作状態の物理 CPU 4 1 0 ~ 4 1 2 がなければそのまま処理を終了する。なお、動作状態 3 4 1 1 が「通常」とは、前記第 1 実施形態では物理 CPU 4 1 0 ~ 4 1 2 が定格で動作している状態を示したが、本第 2 実施形態では、動作クロックの周波数を変更して仮想 CPU を実行し

10

20

30

40

50

ている状態も「通常」に含む。

【0078】

次に、ステップS32では、仮想CPUを削除した現在の物理CPUに割り当てられている残りの全ての仮想CPUを割り当て可能な物理CPUがあるか否かを、仮想-物理CPU割当表341を参照して判定する。つまり、仮想CPUを削除した物理CPUに残っている全ての仮想CPUの割り当て率の和と、他の物理CPUに割り当てられている仮想CPUの割り当て率の和を加算した値が100%以下である物理CPUがあれば、この物理CPUを、仮想CPUの移動対象として選択し、ステップS33の処理へ進む。一方、前記加算した値が100%を超えた場合には仮想CPUの移動ができないので、そのまま処理を終了する。

10

【0079】

ステップS33では、仮想CPUを削除した物理CPUの仮想CPUを、ステップS32で選択した物理CPUに割り当て先を移動して、仮想-物理CPU割当表341を更新する。さらに、仮想CPUの移動によって、割り当てられていた仮想CPUがなくなった物理CPUをスリープ状態に設定し、仮想-物理CPU割当表341を更新してから処理を終了する。

【0080】

以上の処理により、仮想CPUの削除によって割り当てられている仮想CPUがなくなった物理CPUがスリープ状態に移行するのに加え、仮想CPUを削除した物理CPUに割り当てられている残りの仮想CPUを他の物理CPUで割り当て可能な場合には、これらの残っている仮想CPUを、他の物理CPUへ移動させて割り当て先を変更する。そして、仮想CPUの移動によって割り当てられている仮想CPUがなくなった物理CPUをスリープ状態に移行させることで、物理計算機400で全体の仮想CPUを実行するのに必要な最小限の物理CPUのみを通常の動作状態で稼働させ、他の物理CPUをスリープ状態に移行させることで、スリープ状態の物理CPUの数を最大にして物理計算機400の消費電力を大幅に低減できるように仮想CPUの割り当て状態を最適化することができるのである。

20

【0081】

例えば、図7に示した仮想-物理CPU割当表341を、物理CPUと仮想CPUの割り当て率で表現すると図10のようになる。図10では、物理CPU#3(412)に割り当てられている仮想CPU#4(212)を削除する場合を示しており、図中上段が削除前を示し、下段が削除後の状態を示す。

30

【0082】

物理CPU#1に割り当てられている仮想CPU#1、#3の割り当て率はそれぞれ40%、20%であり、物理CPU#3に割り当てられている仮想CPU#4、#2の割り当て率はそれぞれ20%、40%である。また、物理CPU#2は、スリープ状態となっている。

【0083】

ここで、物理CPU#3の仮想CPU#2を削除すると、物理CPU#3に残った仮想CPUは割り当て率が20%の仮想CPU#4のみとなる。ここで、割り当て率=20%の仮想CPU#4を割り当て可能な物理CPUを仮想-物理CPU割当表341から検索すると、全仮想CPUの割り当て率=60%の物理CPU#1に、仮想CPU#3の割り当て率=20%を加算した値は80%となり、所定の閾値未満である90%以下であるので、仮想CPU#3を物理CPU#1へ移動することができる。なお、本実施形態では、動作クロックの周波数を変更する閾値を、仮想CPUの割り当て率3413の合計値を判定する閾値として代用したが、これらの閾値を独立して設定してもよい。

40

【0084】

仮想化ソフトウェア300は、物理CPU#3の仮想CPU#3を物理CPU#1に移動して割り当て先を変更する。そして、割り当てられている仮想CPUのなくなった物理CPU#3をスリープ状態にする。

50

【 0 0 8 5 】

以上のように、仮想CPUの削除時には、仮想CPUを削除した物理CPU上に残った仮想CPUの割り当て率の和と、通常の動作状態の他の物理CPUの割り当て率の和とを加算した値が、所定の閾値未満となる物理CPUを順次検索し、2つの物理CPU上の仮想CPUの割り当て率の和が所定の閾値未満であれば、ひとつの物理CPUに仮想CPUを再配置することで、仮想CPUを割り当てない物理CPUを生成してスリープ状態に移行させ、物理計算機400の消費電力を低減させることができる。

【 0 0 8 6 】

さらに、物理CPUに割り当てる仮想CPUの割り当て状態が変更されたときには、物理CPUに割り当てられている仮想CPUの割り当て率の和に応じて動作クロックの周波数を増減するので、物理計算機400の消費電力を最小限に抑制することが可能となる。

10

【 0 0 8 7 】

なお、上記第2実施形態において、仮想化ソフトウェア300は仮想CPUの削除の指示を受け付けたときに、物理CPUに対する仮想CPUの割り当てを変更する例を示したが、仮想CPUの作成時や所定の周期で物理CPU毎の割り当て状態を監視して、上記図9のステップS12～S33を実行するようにしてもよい。

【 0 0 8 8 】

なお、上記では物理CPU410～412の動作クロックの周波数を変更する例を示したが、物理CPU410～412のクロック・電力制御部4104に電圧を変更するように指令しても良く、例えば、動作クロックの周波数を低下させる場合には、プロセッサコア4101、4102に供給する電圧も低下するように指令することができる。なお、動作クロックの周波数と供給電圧を変更する手法は、公知または周知の手法を用いればよい。

20

【 0 0 8 9 】

< 第3実施形態 >

図11は第3の実施形態を示し、前記第2実施形態の図9に示した仮想化ソフトウェア300が行う仮想CPUの削除処理フローチャートの一部を変更したもので、その他の構成は、前記第2実施形態と同様である。

【 0 0 9 0 】

図11のフローチャートは、図9のフローチャートに示したステップS32とステップS33の間に、物理CPU410～412の動作クロックの周波数を変更するステップS40を加えたものである。

30

【 0 0 9 1 】

ステップS40では、前記第2実施形態の図8のステップS21及びS22で示したように、現在着目している物理CPU410～412の動作クロックを変更できるか否かを判定し、変更可能な場合には新たな割り当て時に応じて動作クロックの周波数を変更するものである。

【 0 0 9 2 】

すなわち、図10において、物理CPU#3に割り当てられている仮想CPU#4を削除し、この物理CPU#3に割り当てられている仮想CPU#3を物理CPU#1へ移動した場合、物理CPU#1の動作クロックの周波数を、新たな割り当て率の和に応じて変更するものである。そして、前記第2実施形態と同様に、仮想CPUの割り当て率の和と閾値（例えば、90%）を比較して、割り当て率の和が閾値未満であれば和の値に応じた動作クロックの周波数に設定することができる。

40

【 0 0 9 3 】

これにより、仮想CPUを纏めた物理CPUの動作クロックの周波数を割り当て率の和に応じて低減することで、物理計算機400の処理能力を確保しながら消費電力の増大を抑制することができる。

【 0 0 9 4 】

また、上記物理CPU410～412の動作クロックの周波数の変更は、前記第1実施

50

形態の図 6 に示した仮想 CPU の削除時に適用しても良い。この場合、図 12 のように、ステップ S 12 の判定処理が Yes の場合にステップ S 41 を実行させる。

【 0095 】

ステップ S 41 では、上記図 11 のステップ S 40 と同様に、現在着目している物理 CPU 410 ~ 412 の動作クロックを変更できるか否かを判定し、変更可能な場合には新たな割り当て時に応じて動作クロックの周波数を変更するものである。すなわち、仮想 CPU を削除した物理 CPU に残っている仮想 CPU の割り当て率の和に応じて動作クロックの周波数を変更する。このとき、仮想 CPU の割り当て率の和と所定の閾値を比較し、割り当て率の和が閾値未満のときに動作クロックの周波数の変更を行うようにしても良い。

10

【 0096 】

これにより、仮想 CPU を実行するのに必要な動作クロックの周波数を設定でき、物理計算機 400 の消費電力をさらに低減できるのである。

【 0097 】

< 第 4 実施形態 >

図 13 は、第 4 の実施形態を示し、前記第 1 実施形態のゲスト OS 100、110 上で、ゲスト OS 毎の仮想 CPU の使用率（以下、稼働率とする）を監視して仮想化ソフトウェア 300 へ通知する稼働率監視ミドルウェア 500、510 を実行し、仮想化ソフトウェア 300 の仮想 - 物理 CPU 割当制御部 340 では、仮想 CPU の稼働率に応じて物理 CPU に対する仮想 CPU の割り当てを制御する。

20

【 0098 】

仮想 - 物理 CPU 割当制御部 340 は、図 14 で示すように、前記第 1 実施形態の図 3 に示した仮想 - 物理 CPU 割当表に稼働率 3415 と実使用率 3416 を加えた仮想 - 物理 CPU 割当表 341B を管理する。

【 0099 】

ここで、仮想 - 物理 CPU 割当制御部 340 は、仮想 - 物理 CPU 割当表 341B の稼働率 3415 に、各ゲスト OS 100、110 が利用している各仮想 CPU の稼働率を格納する。また、仮想 - 物理 CPU 割当制御部 340 は、仮想 - 物理 CPU 割当表 341B の実使用率 3416 には、各仮想 CPU の割り当て率に稼働率を乗じた値を、各仮想 CPU が実際に利用している物理 CPU の利用率を格納する。

30

【 0100 】

仮想 CPU を割り当てる際には、前記第 1 実施形態と同様に、管理コンソール 500 からの操作指令に応じた論理区画 200、210 に仮想 CPU を割り当てる。仮想 CPU を割り当てた論理区画 200、210 ではゲスト OS 100、110 がそれぞれ起動し、さらに、稼働率監視ミドルウェア 500、510 が起動して所定の周期などでゲスト OS 100、110 が使用する各仮想 CPU の稼働率を仮想化ソフトウェア 300 へ通知する。

【 0101 】

仮想化ソフトウェア 300 は、管理コンソール 500 から指定された仮想 CPU を削除の操作指令を受け付けると、図 15 に示すフローチャートを実行し、不要な物理 CPU 410 ~ 412 をスリープ状態に移行させ、さらに、各仮想 CPU の稼働率に応じて仮想 CPU を再配置する。

40

【 0102 】

図 15 のフローチャートは、前記第 2 実施形態の図 9 に示した処理のうち、ステップ S 32 の処理を新たなステップ S 50、S 51 に置き換えたものであり、その他の構成は前記第 2 実施形態と同様である。ステップ S 11、S 12 では、上述のとおり指定された仮想 CPU を削除し、仮想 - 物理 CPU 割当表 341B を更新する。そして、仮想 CPU を削除した物理 CPU に他の仮想 CPU が残っていれば、ステップ S 31 で当該物理 CPU の他に通常の動作状態の物理 CPU があるかを判定する。他の物理 CPU が通常の動作状態であれば、ステップ S 50 に進んで、物理計算機 400 で稼働中の仮想 CPU の稼働率を取得する。

50

【 0 1 0 3 】

ステップ S 5 0 では、稼働率監視ミドルウェア 5 0 0、5 1 0 が通知した仮想 CPU の稼働率のうち最新の値を取得するか、または各仮想 CPU 毎の稼働率の平均値を演算する。そして、仮想化ソフトウェア 3 0 0 は、取得した各仮想 CPU の割り当て率を仮想 - 物理 CPU 割当表 3 4 1 B の稼働率 3 4 1 5 に書き込んで更新する。さらに、仮想化ソフトウェア 3 0 0 は、仮想 - 物理 CPU 割当表 3 4 1 B の割り当て率 3 4 1 3 に稼働率 3 4 1 5 を乗じた値を各仮想 CPU が実際に利用している物理 CPU の使用率 3 4 1 6 として求め、この値を実使用率 3 4 1 6 に書き込んで更新する。

【 0 1 0 4 】

次に、ステップ S 5 1 では、仮想 CPU を削除した現在の物理 CPU に割り当てられている残りの全ての仮想 CPU を割り当て可能な物理 CPU があるか否かを、仮想 - 物理 CPU 割当表 3 4 1 B を参照して判定する。つまり、仮想 CPU を削除した物理 CPU に残っている全ての仮想 CPU の実使用率 3 4 1 6 の和と、他の物理 CPU に割り当てられている仮想 CPU の実使用率 3 4 1 6 の和を加算した値が所定の第 2 の閾値（例えば、1 0 0 %）以下である物理 CPU があれば、この物理 CPU を、仮想 CPU の移動対象として選択し、ステップ S 3 3 の処理へ進む。一方、前記加算した値が第 2 の閾値を超えた場合には仮想 CPU の移動ができないので、そのまま処理を終了する。

10

【 0 1 0 5 】

ステップ S 3 3 では、仮想 CPU を削除した物理 CPU の仮想 CPU を、ステップ S 5 1 で選択した物理 CPU に割り当て先を移動して、仮想 - 物理 CPU 割当表 3 4 1 B を更新する。さらに、仮想 CPU の移動によって、割り当てられていた仮想 CPU がなくなった物理 CPU をスリープ状態に設定し、仮想 - 物理 CPU 割当表 3 4 1 B を更新してから処理を終了する。

20

【 0 1 0 6 】

以上の処理を図 1 4 の仮想 - 物理 CPU 割当表 3 4 1 B に示した仮想 CPU # 4 を削除する例を、図 1 6 に示す。

【 0 1 0 7 】

図 1 6 は、図 1 4 の仮想 - 物理 CPU 割当表 3 4 1 B から仮想 CPU # 2 を削除した状態を図中上部の割当変更前に示し、仮想 CPU の割り当て変更後の状態を図中下部に示し、図中斜線の部分が仮想 CPU の実使用率 3 4 1 6 を示す。

30

【 0 1 0 8 】

物理 CPU # 1 には、仮想 CPU # 1、# 3 がそれぞれ 5 0 % と 3 0 % の割り当て率で割り当てられており、物理 CPU # 3 には仮想 CPU # 4 が 4 0 % の割り当て率で割り当てられている。なお、物理 CPU # 2 はスリープ状態である。

【 0 1 0 9 】

仮想化ソフトウェア 3 0 0 は、各仮想 CPU の稼働率 3 4 1 5 を稼働率監視ミドルウェア 5 0 0 から取得し、仮想 - 物理 CPU 割当表 3 4 1 B の割り当て率 3 4 1 3 に稼働率 3 4 1 5 を乗じた値の百分立を物理 CPU の実使用率 3 4 1 6 として求める。この結果、実使用率 3 4 1 6 は、仮想 CPU # 1 が 2 5 %、仮想 CPU # 3 が 9 %、仮想 CPU # 4 が 8 % となる。仮想化ソフトウェア 3 0 0 は、今回、仮想 CPU を削除した物理 CPU # 3 の仮想 CPU の実使用率 3 4 1 6 の合計値と、他の物理 CPU のうち通常の動作状態のものをひとつ選択して、この物理 CPU（この例では、物理 CPU # 1）の実使用率 3 4 1 6 の合計値を求める。ここで、物理 CPU # 1 を選択すると、仮想 CPU # 1 と仮想 CPU # 3 の実使用率 3 4 1 6 の和 = 3 4 % に、仮想 CPU を削除した物理 CPU # 3 の実使用率 3 4 1 6 の和 = 8 % を加算した値 = 4 2 % が第 2 の閾値（例えば、1 0 0 %）未満であるか否かを判定する。

40

【 0 1 1 0 】

この例では、選択した物理 CPU # 1 と物理 CPU # 3 の実使用率 3 4 1 6 の和は、4 2 % となって第 2 の閾値未満であるので、仮想化ソフトウェア 3 0 0 は物理 CPU # 3 の仮想 CPU # 4 を物理 CPU # 1 へ移動可能と判定する。そして、仮想化ソフトウェア 3

50

00は、仮想-物理CPU割当表341Bを図17で示すように更新し、物理CPU#3のエントリを仮想CPUの割り当てなしに更新し、物理CPU#1のエントリに仮想CPU#4の識別子と割り当て率3413、稼働率3415、実使用率3416を書き込んで更新する。この後、仮想化ソフトウェア300は、物理CPU#3をスリープ状態に移行させて処理を終了する。

【0111】

上記仮想CPUの再配置の結果、図16の下部に示すように、物理CPU#1には3つの仮想CPU#1、#3、#4が割り当てられて、割り当て率3413の和は120%となって、物理CPU#1のリソースの最大値を示す100%を超えているが、各仮想CPUの実使用率3416の和は100%未満の42%であるので、物理CPU#1は各仮想CPUを円滑に実行できる。

10

【0112】

このように、仮想CPUを削除して仮想CPUの再配置を行う際に、仮想CPUが実際に利用している物理CPUの実使用率3416を、稼働率監視ミドルウェア500から取得した値に基づいて求め、2つの物理CPU上の仮想CPUの実使用率3416の和が第2の閾値未満であれば、ひとつの物理CPUに仮想CPUを再配置することで、仮想CPUを割り当てない物理CPUを生成してスリープ状態に移行させ、物理計算機400の消費電力を低減させることができる。

【0113】

さらに、論理区画200、210では各ゲストOS100、110上で稼働率監視ミドルウェア500、510を稼働させて仮想化ソフトウェア300へ各ゲストOS100、110の仮想CPUの使用率を通知するようにしたので、既存のOSをそのままゲストOS100、110として利用しながら仮想化ソフトウェア300では仮想CPUの稼働率に応じた割り当て制御が可能となる。

20

【0114】

なお、上記では第2の閾値を100%とした例を示したが、各ゲストOS100、110の負荷変動を考慮した場合、第2の閾値を80%などの100%未満の値に設定することで、負荷変動による実使用率の急増に対応することができる。

【0115】

また、通常の動作状態の物理CPUが複数ある場合に、仮想化ソフトウェア300は、上記ステップS51の処理で、仮想CPUを削除した物理CPUの実使用率を、他の物理CPUの実使用率と順次比較し、第2の閾値未満となる物理CPUを検索すればよい。

30

【0116】

なお、上記各実施形態では、物理CPU410~412として複数のプロセッサコアを備えたホモジニアスのマルチコア・マルチプロセッサに本発明を適用した例を示したが、異種のプロセッサコアを備えたヘテロジニアスのマルチコア・マルチプロセッサに本発明を適用することができる。

【産業上の利用可能性】

【0117】

以上のように、本発明では、複数の物理CPUを備えて複数の仮想CPUを実行する仮想計算機システム及び仮想化ソフトウェアに適用することができる。

40

【図面の簡単な説明】

【0118】

【図1】第1の実施形態を示し、本発明を適用する仮想計算機システムの構成を示すブロック図。

【図2】第1の実施形態を示し、物理CPUの構成を示すブロック図。

【図3】第1の実施形態を示し、仮想化ソフトウェアが管理する仮想-物理CPU割当表の一例を示す説明図。

【図4】第1の実施形態を示し、仮想化ソフトウェアが管理する論理区画対応表の一例を示す説明図。

50

【図5】第1の実施形態を示し、仮想化ソフトウェアが行う仮想CPUの生成処理の一例を示すフローチャート。

【図6】第1の実施形態を示し、仮想化ソフトウェアが行う仮想CPUの削除処理の一例を示すフローチャート。

【図7】第2の実施形態を示し、仮想化ソフトウェアが管理する仮想 - 物理CPU割当表の一例を示す説明図。

【図8】第2の実施形態を示し、仮想化ソフトウェアが行う仮想CPUの生成処理の一例を示すフローチャート。

【図9】第2の実施形態を示し、仮想化ソフトウェアが行う仮想CPUの削除処理の一例を示すフローチャート。

【図10】第2の実施形態を示し、仮想CPUの削除処理による仮想CPUの再配置の様子を示す説明図で、図中上部が割り当て変更前の状態を示し、図中下部が割り当て変更後の状態を示す。

【図11】第3の実施形態を示し、仮想化ソフトウェアが行う仮想CPUの削除処理の一例を示すフローチャート。

【図12】第3の実施形態を示し、仮想化ソフトウェアが行う仮想CPUの削除処理の他の例を示すフローチャート。

【図13】第4の実施形態を示し、仮想計算機システムの構成を示すブロック図。

【図14】第4の実施形態を示し、仮想化ソフトウェアが管理する仮想 - 物理CPU割当表の一例を示す説明図。

【図15】第4の実施形態を示し、仮想化ソフトウェアが行う仮想CPUの削除処理の一例を示すフローチャート。

【図16】第4の実施形態を示し、仮想CPUの削除処理による仮想CPUの再配置の様子を示す説明図で、図中上部が割り当て変更前の状態を示し、図中下部が割り当て変更後の状態を示す。

【図17】第4の実施形態を示し、仮想化ソフトウェアが管理する仮想 - 物理CPU割当表で、仮想CPUの削除処理後の一例を示す説明図。

【符号の説明】

【0119】

100、110 ゲストOS

200、210 論理区画

201、202、211、212 仮想CPU

300 仮想化ソフトウェア

310 LPAR制御部

320 仮想CPU制御部

330 物理CPU制御部

340 仮想 - 物理CPU割当制御部

400 物理計算機

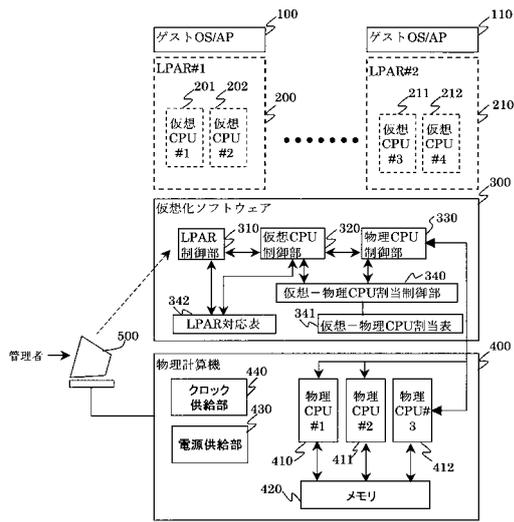
410 物理CPU

10

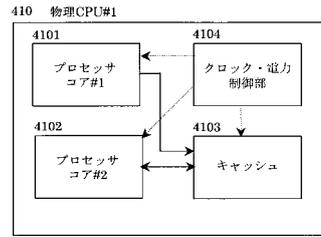
20

30

【図1】



【図2】



【図3】

341 仮想-物理CPU割当表

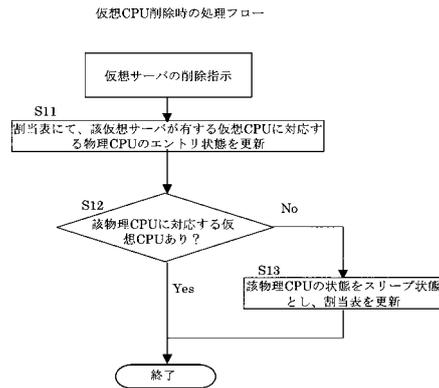
物理CPU#	状態	対応する仮想CPUリスト			
		仮想CPU#	割当率	仮想CPU#	割当率
1	通常	1	40%	3	20%
2	スリープ	なし	0%	なし	0%
3	通常	4	20%	2	40%
Σ	Σ	Σ	Σ	Σ	Σ
3410	3411	3412	3413	3412	3413

【図4】

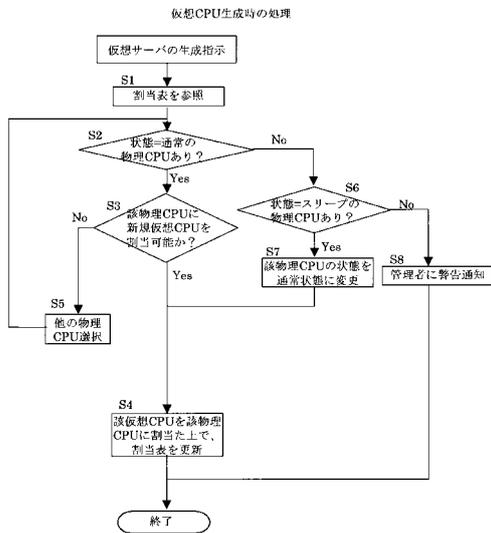
342 LPAR対応表

LPAR#	仮想CPUリスト		
	仮想CPU#	仮想CPU#	仮想CPU#
1	1	2	
2	3	4	
...

【図6】



【図5】



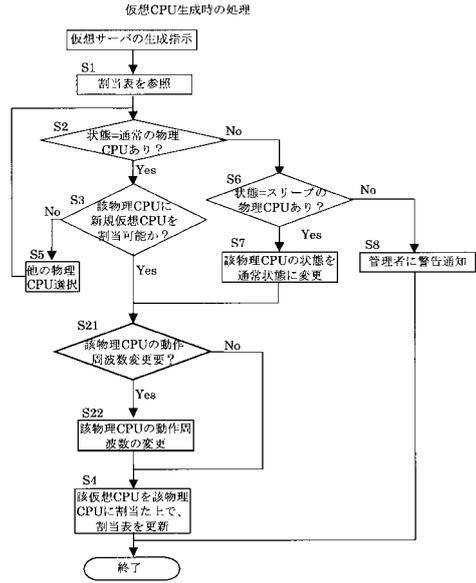
【図7】

341A 仮想-物理CPU割当表

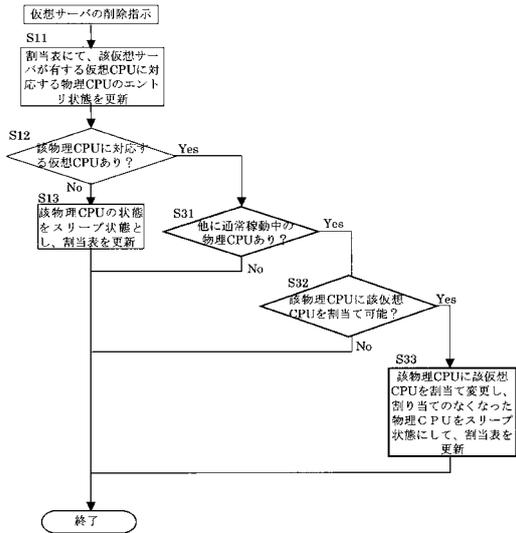
物理CPU#	状態	動作周波数	対応する仮想CPUリスト			
			仮想CPU#	割当率	仮想CPU#	割当率
1	通常	60%	1	40%	3	20%
2	スリープ	0%	なし	0%	なし	0%
3	通常	60%	4	20%	2	40%

3410 3411 3414 3412 3413 3412 3413

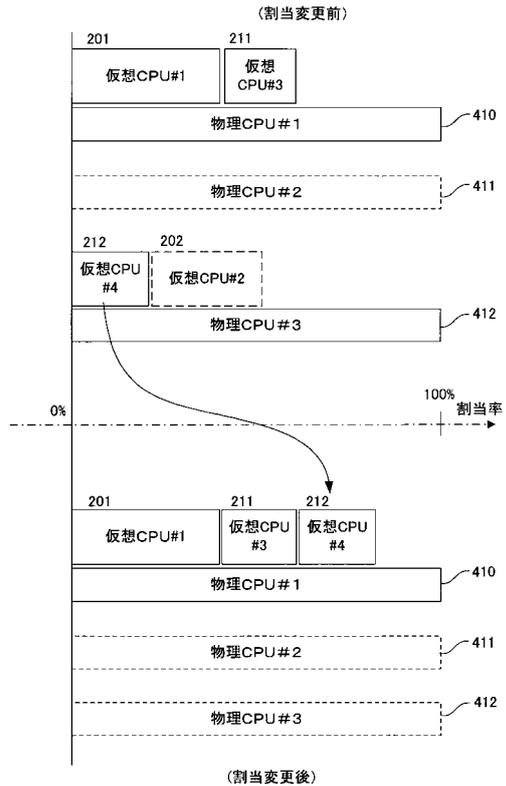
【図8】



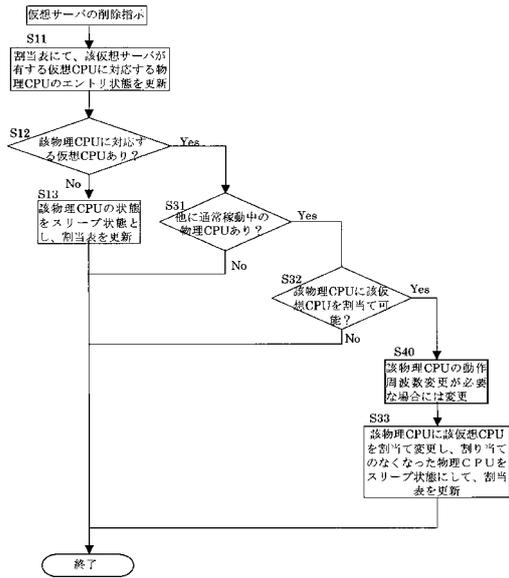
【図9】



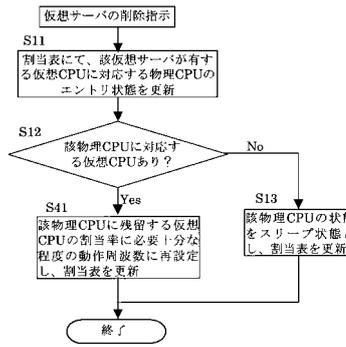
【図10】



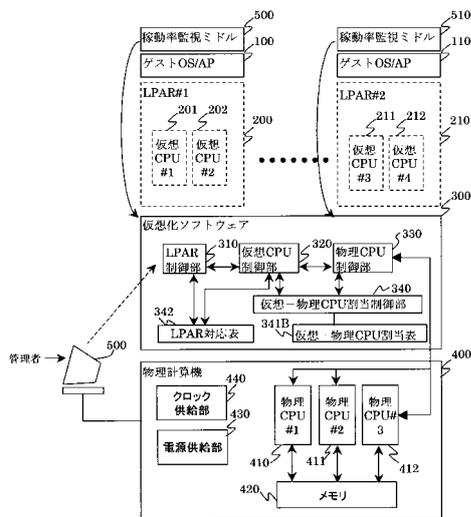
【図 1 1】



【図 1 2】



【図 1 3】

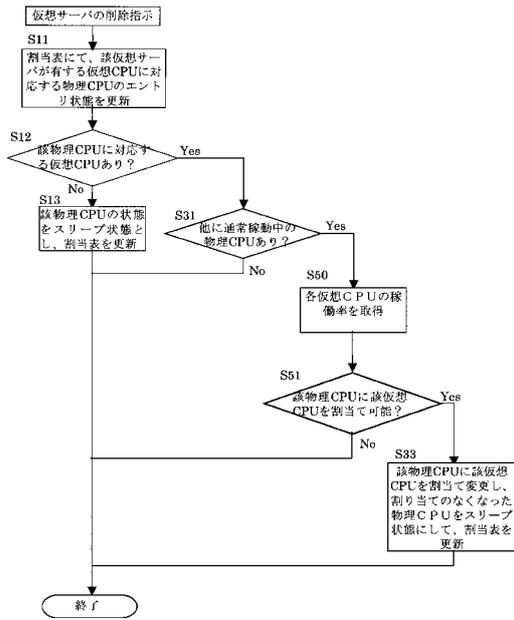


【図 1 4】

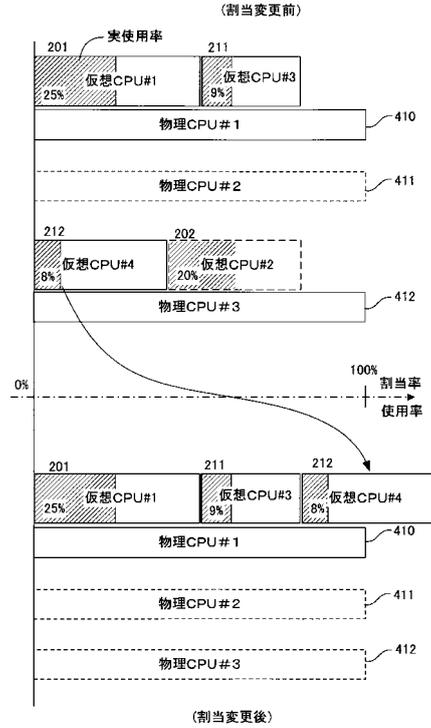
341B 仮想-物理CPU割当表

物理CPU#	状態	対応する仮想CPUリスト							
		仮想CPU#	割当率	稼働率	実使用率	仮想CPU#	割当率	稼働率	実使用率
1	通常	1	50%	50%	25%	3	30%	30%	9%
2	スリープ	なし	0%	0%	0%	なし	0%	0%	0%
3	通常	4	40%	20%	8%	2	40%	50%	20%
3410	ス	ス	ス	ス	ス	ス	ス	ス	ス
3411	ス	3412	3413	3415	3416	3412	3413	3415	3416

【図15】



【図16】



【図17】

341B 仮想-物理CPU割当表

物理CPU#	状態	仮想サーバの仮想CPUリスト							
		仮想CPU#	割当率	稼働率	実使用率	仮想CPU#	割当率	稼働率	実使用率
1	通常	3	50%	50%	50%	3	40%	20%	20%
2	スリープ	なし	0%	0%	0%	なし	0%	0%	0%
3	スリープ	なし	0%	0%	0%	なし	0%	0%	0%
3410	3411	3412	3413	3415	3416	3412	3413	3415	3416

フロントページの続き

- (72)発明者 森木 俊臣
東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所 中央研究所内
- (72)発明者 服部 直也
東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所 中央研究所内

合議体

- 審判長 辻本 泰隆
審判官 田中 秀人
審判官 仲間 晃

- (56)参考文献 特開2004-192612(JP,A)
特開平9-179667(JP,A)
特開2006-113767(JP,A)
特開2001-256067(JP,A)

(58)調査した分野(Int.Cl., DB名)

- G06F 9/46
G06F 1/32