

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2010-512600

(P2010-512600A)

(43) 公表日 平成22年4月22日(2010.4.22)

(51) Int.Cl.	F I	テーマコード (参考)
G 0 6 F 9/50 (2006.01)	G O 6 F 9/46 4 6 5 D	
G 0 6 F 9/46 (2006.01)	G O 6 F 9/46 3 5 O	

審査請求 未請求 予備審査請求 未請求 (全 17 頁)

(21) 出願番号	特願2009-541433 (P2009-541433)	(71) 出願人	508243639 エルエスアイ コーポレーション アメリカ合衆国 95035 カリフォル ニア, ミルピタス, バーバー レーン 1 621
(86) (22) 出願日	平成19年9月21日 (2007.9.21)	(74) 代理人	100064447 弁理士 岡部 正夫
(85) 翻訳文提出日	平成21年6月3日 (2009.6.3)	(74) 代理人	100094112 弁理士 岡部 譲
(86) 国際出願番号	PCT/US2007/079145	(74) 代理人	100085176 弁理士 加藤 伸晃
(87) 国際公開番号	W02008/073553	(74) 代理人	100104352 弁理士 朝日 伸光
(87) 国際公開日	平成20年6月19日 (2008.6.19)	(74) 代理人	100128657 弁理士 三山 勝巳
(31) 優先権主張番号	11/609,408		
(32) 優先日	平成18年12月12日 (2006.12.12)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 記憶装置関連負荷情報を使用するクラスタ化されたプロセッサでの仮想機械の負荷バランスのための方法およびシステム

(57) 【要約】

(ブレード・センタなどの) 複数の物理プロセッサのクラスタ内で動作可能な複数の仮想機械の負荷バランスのための記憶装置関連負荷率情報を生成するための方法およびシステム。全体としての記憶装置システムの動作に関する、かつ記憶装置システム内の複数の記憶装置コントローラの各々に関する負荷率情報が記憶装置システム内で生成される。記憶装置システム内でそのように生成された情報は、複数の仮想機械と関連する負荷バランス要素に通信される。次いでこの負荷バランス要素は、複数の仮想機械の動作を複数の物理プロセッサにわたって分散または再分散するために、任意選択で他の負荷率情報と組み合わせて、この記憶装置関連負荷率情報を利用する。

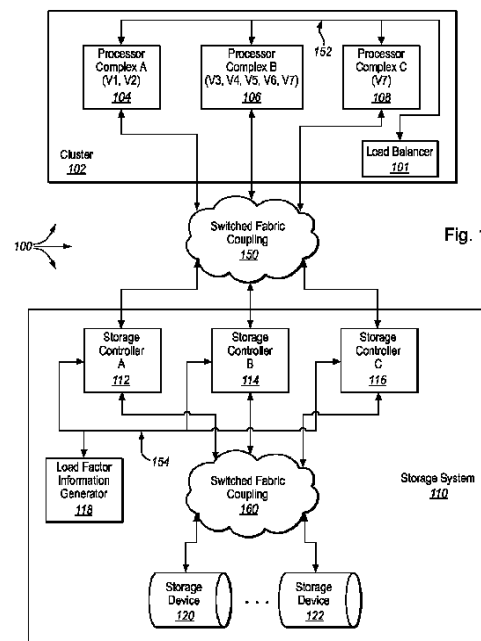


Fig. 1

【特許請求の範囲】**【請求項 1】**

物理プロセッサの各々が前記クラスタの他の各物理プロセッサに通信可能に結合される、物理プロセッサのクラスタと、

各仮想機械が前記物理プロセッサの任意のサブセット上で動作可能であり、複数の仮想機械が複数の前記物理プロセッサにわたって動作するように分散される、複数の仮想機械と、

前記複数の物理プロセッサに通信可能に結合され、負荷率情報に基づいて前記物理プロセッサの第 1 のサブセットから前記物理プロセッサの異なるサブセットに仮想機械を再分散するようになされた負荷バランスと、

複数の記憶装置コントローラを有し、前記物理プロセッサに結合され、記憶装置システムの前記記憶装置コントローラに関する記憶装置システム負荷率情報を生成するようになされた記憶装置負荷率情報ジェネレータを含む記憶装置システムであって、前記負荷率情報ジェネレータが、前記仮想機械を再分散するために中で使用するために前記負荷バランスに前記記憶装置システム負荷率情報を与えるように、前記負荷バランスに通信可能に結合される記憶装置システムとを備える、システム。

【請求項 2】

各物理プロセッサが、前記記憶装置システムの前記複数の記憶装置コントローラの対応するサブセットに結合される、請求項 1 に記載のシステム。

【請求項 3】

各物理プロセッサが、前記記憶装置システムの前記複数の記憶装置コントローラのうちのいずれかに結合され得る、請求項 1 に記載のシステム。

【請求項 4】

前記物理プロセッサが、前記記憶装置コントローラにスイッチ型ファブリックを介して結合される、請求項 3 に記載のシステム。

【請求項 5】

前記物理プロセッサが、前記記憶装置コントローラにスイッチ型シリアル付加 SCSI (「SAS」) ネットワークを介して結合される、請求項 3 に記載のシステム。

【請求項 6】

前記複数の記憶装置コントローラが互いに結合され、1 つの記憶装置コントローラが監督コントローラとして指定され、前記記憶装置負荷率情報ジェネレータが、前記複数のコントローラの各々に関する記憶装置負荷率情報を生成するように前記監督コントローラ内で動作可能である、請求項 1 に記載のシステム。

【請求項 7】

前記記憶装置システムが、前記複数の記憶装置コントローラの対応するコントローラ内で、前記対応するコントローラに対する記憶装置負荷率情報を生成するように、各々が動作可能な複数の記憶装置負荷率情報ジェネレータをさらに備える、請求項 1 に記載のシステム。

【請求項 8】

記憶装置システム内の複数の記憶装置コントローラの負荷に関する記憶装置負荷率情報を前記記憶装置システム内で生成することと、

前記記憶装置負荷率情報を複数の物理プロセッサと関連する負荷バランスに通信することと、

前記記憶装置コントローラの使用と関連する前記物理プロセッサ上の負荷をバランスさせるように、複数の仮想機械のうちの 1 つまたは複数のクラスタの異なる物理プロセッサ上で動作するように再分散することを含む、記憶装置システムに結合される物理プロセッサのクラスタ内で動作する複数の仮想機械の負荷をバランスさせる方法。

【請求項 9】

生成する前記ステップがさらに、

前記負荷率情報内の変化の検知に応答して前記負荷バランスを更新させるように、前記

10

20

30

40

50

記憶装置負荷率情報を生成することをさらに含む、請求項 8 に記載の方法。

【請求項 10】

生成する前記ステップが、

I / O 要求の処理のための、返されるステータス情報の一部分として、前記記憶装置負荷率情報を生成することをさらに含む、かつ、

通信する前記ステップが、

完了した I / O 要求に応答して、ステータス情報として、前記生成された記憶装置負荷率情報を返すことをさらに含む、請求項 8 に記載の方法。

【請求項 11】

前記記憶装置負荷率情報内の情報に関する 1 つまたは複数の閾値を設けることをさらに含む、

通信する前記ステップが

前記記憶装置負荷率情報内の値が対応する閾値を超えると、前記負荷バランサに前記記憶装置負荷率情報を通信することをさらに含む、請求項 8 に記載の方法。

【請求項 12】

設ける前記ステップが、

前記 1 つまたは複数の閾値を動的に更新することをさらに含む、請求項 11 に記載の方法。

【請求項 13】

設ける前記ステップが、

前記記憶装置コントローラの各々と現在通信している動作中の仮想機械の数と、

前記仮想機械の任意のセットの I / O 要求処理率と、

コントローラと仮想機械の任意のセットの間のデータ処理量と、

記憶装置コントローラ内の未処理の I / O 要求の数と、

仮想機械の 1 セットからの未処理の I / O 要求の数と、

記憶装置コントローラの任意のセットと仮想機械の任意のセットの間の可能な通信パスの数と、

仮想機械の任意のセットに接続される記憶装置コントローラの合計数と、

物理プロセッサの数、とからなる群から選択される 1 つまたは複数の尺度から導出される 1 つまたは複数の閾値を設けることをさらに含む、請求項 11 に記載の方法。

【請求項 14】

スイッチ型ファブリック接続を介して記憶装置システムに結合されるブレード・センタを有するシステム内で動作可能な複数の仮想機械用の負荷バランシング装置であって、前記ブレード・センタが複数のプロセッサ / メモリ複合体を備え、前記記憶装置システムが複数の記憶装置コントローラを備える負荷バランシング装置において、

前記複数の記憶装置コントローラのうちの 1 つまたは複数に関する記憶装置負荷率情報を生成するための、前記記憶装置システム内で動作可能な記憶装置システム負荷率情報ジェネレータと、

前記ブレード・センタ内で動作可能な、かつ前記記憶装置システム負荷率情報ジェネレータに通信可能に結合される負荷バランサであって、前記負荷バランサが前記ジェネレータから記憶装置負荷率情報を受け取るようになされ、かつ前記負荷バランサが、前記複数の仮想機械のうちの 1 つまたは複数の動作を前記ブレード・センタの前記複数のプロセッサ / メモリ複合体にわたって分散するのに前記記憶装置負荷率情報を利用するようになされている負荷バランサとを備える、負荷バランシング装置。

【請求項 15】

前記ジェネレータが、前記負荷率情報内で変化を検知するのに応答して前記記憶装置負荷率情報を生成するようになされた、請求項 14 に記載の装置。

【請求項 16】

前記ジェネレータが、記憶装置コントローラによる I / O 要求の処理に対して返されたステータス情報の一部分として、前記記憶装置負荷率情報を生成するようになされ、かつ

10

20

30

40

50

完了した I / O 要求に応答して、前記記憶装置負荷率情報をステータス情報として通信するようになされている、請求項 14 に記載の装置。

【請求項 17】

前記ジェネレータが、

前記記憶装置負荷率情報内の情報に関する閾値をさらに備え、

前記ジェネレータが、前記記憶装置負荷率情報内の値が前記閾値を超えると、前記記憶装置負荷率情報を前記負荷バランスに通信するようにさらになされている、請求項 14 に記載の装置。

【請求項 18】

前記ジェネレータが、前記閾値を動的に更新するようにさらになされている、請求項 17 に記載の装置。

【請求項 19】

前記閾値が、

前記記憶装置コントローラの各々と現在通信している動作中の仮想機械の数と、

前記仮想機械の任意のセットの I / O 要求処理率と、

コントローラと仮想機械の任意のセットの間のデータ処理量と、

記憶装置コントローラ内の未処理の I / O 要求の数と、

仮想機械の 1 セットからの未処理の I / O 要求の数と、

記憶装置コントローラの任意のセットと仮想機械の任意のセットの間の可能な通信パスの数と、

仮想機械の任意のセットに接続される記憶装置コントローラの合計数と、

物理プロセッサの数とからなる群から選択される 1 つまたは複数の尺度から導出される、請求項 17 に記載の装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は一般にクラスタ化されたコンピュータ環境で動作する複数の仮想機械のための負荷バランスに関し、より具体的には、いわゆる「ブレード・センタ」などクラスタの複数の物理プロセッサ上で動作可能な複数の仮想機械の負荷バランスに際し、記憶装置システム負荷情報を生成しかつ利用するためのシステムおよび方法に関する。

【背景技術】

【0002】

コンピュータ業界では、物理コンピュータ環境の「頂上」で動作する仮想機械環境を用意することは一般に知られている。ゲスト・オペレーティング・システムとも呼ばれる仮想機械環境は、ユーザおよび / またはアプリケーション・プログラムが、関連する専用のメモリ、記憶装置、および / または I / O 周辺装置を有する専用の物理プロセッサを、全て仮想オペレーティング・システム（「ゲスト」オペレーティング・システム）の制御下で利用しているかのように動作することができるように、仮想レイヤを提供する。したがって複数のそのような仮想機械は、複数のユーザおよび / またはアプリケーション・プログラムが、それぞれが専用の物理コンピュータ環境を使用しているかのように同時に / 並行して動作することができるように、単一のコンピュータ環境上で同時に（かつ / または並行して）動作することができる。例えば、IBM MVS (multiple virtual system) コンピュータ環境を含む、かつ仮想機械ソフトウェア・コンポーネントの VMware ファミリなどのより最近の例を含むそのような仮想機械コンピュータ技術は、かなり以前からコンピュータ分野で良く知られている。

【0003】

上記で言及したように、物理コンピュータ環境の頂上で仮想機械を動作させることの主要な利点は、複数のそのような仮想機械が同時に / 並行して単一の物理コンピュータ環境上で動作可能になることである。各仮想機械は、ユーザの対応するセットの代わりにかつ / またはアプリケーション・プログラムの対応するセットの代わりに動作可能である。そ

10

20

30

40

50

の上クラスタ化されたコンピュータは、複数の物理プロセッサが「クラスタ」内で密結合するように進化しているので、仮想機械コンピュータ技術は、そのようなクラスタでの複数のプロセッサを利用できるようになされている。特に、複数の仮想機械コンピュータ環境は並行して／同時に動作可能であり、かつクラスタ化されたコンピュータ環境内で構成される複数の物理プロセッサにわたって分散することができる。本明細書で使用されるとき、クラスタ化されたコンピュータは、短距離バス構造を介して通信する密結合された物理プロセッサを含むのみならず、複数のプロセッサが良く知られたコンピュータ・ネットワーク通信プロトコルおよび媒体を利用して「クラスタ」として論理的に結合される、より広く分散されるネットワーク・コンピュータも含むことを広く理解すべきである。特に本明細書で使用されるとき、「クラスタ」は、（各々がオペレーティング・システムの制御下で完全に動作するプロセッサ／複数のプロセッサおよびメモリを含む）複数のサーバが共有ファブリックを介してかつ／または共通のバックプレーン／ミッドプレーン・バス・アーキテクチャを介して密結合された「ブレード・センタ」と一般に呼ばれる構成を含む。したがってクラスタは、密結合された複数の物理プロセッサおよび／またはプロセス相互間通信のために短距離バス構造またはファブリック接続を共有する密接に物理的に接近するプロセッサ／メモリ複合体を意味することができ、あるいはコンピュータ・ネットワーク・プロトコルを利用する複数のプロセッサを結合するより広い物理的分布を含むことができる。さらに本明細書で使用されるとき、「複数のプロセッサ」は、単一のブレード／複合体の上に複数のプロセッサを有するブレードまたは複合体の単一のプロセッサからはっきりと区別される、ブレードなどのプロセッサ／メモリ複合体を指すことを意図している。換言すれば、本明細書で使用されるとき、「複数のプロセッサ」の各「プロセッサ」は、一般に単一のオペレーティング・システムの制御下のプロセッサ／メモリ複合体を含む。

10

20

30

40

50

【 0 0 0 4 】

複数の仮想機械を動作させるそのようなクラスタ・コンピュータ環境では、複数のプロセッサのうちの異なるプロセッサにわたって複数の仮想機械（すなわちゲスト・オペレーティング・システム）の動的な分散および再分散を可能にするように、クラスタ化された環境内に負荷バランシング方法および／またはコンポーネントを設けることは一般に知られている。換言すれば、物理プロセッサのうちの１つ（例えば、１つのブレードまたは複合体）が、その上で動作している複数の仮想機械に対する計算要求および／またはメモリ要求によって過負荷になる場合、過負荷になったブレード／複合体上で動作可能な複数の仮想機械のうちの１つまたは複数を１つまたは複数のより少ない負荷の掛かった物理プロセッサに再分散することができる。したがって、クラスタ化された環境のコンポーネントまたは方法が特定する負荷バランシングは、クラスタのブレード／複合体の起こりうる過負荷を避けるまたは緩和するために、仮想機械環境を、クラスタ化された環境内の複数の物理プロセッサの異なるプロセッサ（例えば、異なるブレードまたは複合体）に適切に分散または再分散することに責任を負っている。

【 0 0 0 5 】

クラスタ化されたコンピュータ環境で動作する複数の仮想機械に関連する現在の負荷バランシング・システムおよび技術は、クラスタ化されたコンピュータ物理プロセッサそれ自体に由来する負荷率情報のみ対処する。換言すれば、現行の負荷バランシング・システムおよび技術は、物理プロセッサの特定のＣＰＵまたはメモリが、関連する物理プロセッサ上の複数の仮想機械の動作によって計算的に過負荷になる可能性があることを判定することができる。しかしながら、そのような負荷率情報が仮想機械動作によって課される負荷の他の態様に対処できないことが一般に問題点である。特に、クラスタ化されたコンピュータ環境で動作する複数の仮想機械に対する現在の負荷バランシング技術は、付加された記憶装置システムに関連する負荷率情報に対処できない。しばしば、クラスタ化されたコンピュータ環境では、複数の物理プロセッサが記憶装置システムへのアクセスを共有する。この記憶装置システムは、例えばＲＡＩＤ記憶装置管理などの記憶装置管理技術を含む付属の物理プロセッサの代わりに、Ｉ／Ｏ要求処理を行うために全てが複数の記憶装置

に結合される複数の記憶装置コントローラを含む場合がある。しばしば、複数の仮想機械が何らかの特定の物理プロセッサのCPUまたは主メモリを著しく過負荷にしない可能性もある。しかしながら、この複数の仮想機械は、物理プロセッサを記憶装置システムの特定の記憶装置コントローラに結合する共通のインターフェイス構造を共有することによって、クラスタ化されたシステムの全体性能に対しボトルネックを生み出す場合がある。したがって、現行の負荷バランシング技術は一般にそのような負荷バランシング・ボトルネックを検出する能力がなく、したがって、1つまたは複数の仮想機械の他の物理プロセッサへの再分散を介してそのようなボトルネックを矯正する能力がない。

【発明の概要】

【発明が解決しようとする課題】

10

【0006】

上記の論議から、クラスタ化されたコンピュータ環境で動作する複数の仮想機械の負荷バランシングのための改善されたシステムおよび方法が求められていることは明らかである。

【課題を解決するための手段】

【0007】

本発明は、記憶装置システム関連の負荷率情報を生成するためのシステムおよび方法を提供し、かつそのような情報をクラスタ化されたコンピュータ環境で動作する複数の仮想機械の負荷を再分散するために適用される負荷バランシング技術に利用することによって、上記のかつ他の問題点を解決し、それによってこの有効な最先端技術を進歩させる。

20

【0008】

したがって本発明の第1の特徴は、物理プロセッサの各々がクラスタの他の各物理プロセッサに通信可能に結合される物理プロセッサのクラスタを含むシステムを提供することである。このシステムは、各仮想機械が物理プロセッサの任意のサブセット上で動作可能であり、かつ複数の仮想機械が複数の物理プロセッサにわたって動作するように分散される複数の仮想機械も含む。さらにこのシステムは、複数の物理プロセッサに通信可能に結合される負荷バランサを含み、この負荷バランサは負荷率情報に基づいて物理プロセッサの第1のサブセットから物理プロセッサの異なるサブセットに仮想機械を再分散するようになされている。最後にこのシステムは、複数の記憶装置コントローラを有する記憶装置システムを含む。この記憶装置システムは物理プロセッサに結合され、かつ記憶装置システムの記憶装置コントローラに関する記憶装置システム負荷率情報を生成するようになされた記憶装置負荷率情報ジェネレータを含む。この負荷率情報ジェネレータは、仮想機械を再分散するために中で使用するために負荷バランサに記憶装置システム負荷率情報を与えるように、負荷バランサに通信可能に結合される。

30

【0009】

本発明の別の別の特徴は、記憶装置システムに結合される物理プロセッサのクラスタ内で動作する複数の仮想機械の負荷をバランスさせるための方法を提供する。この方法は、記憶装置システム内の複数の記憶装置コントローラの負荷に関する記憶装置負荷率情報を記憶装置システム内で生成することを含む。この方法は、この記憶装置負荷率情報を複数の物理プロセッサと関連する負荷バランサに通信することを含む。最後にこの方法は、記憶装置コントローラの使用に伴う物理プロセッサ上の負荷をバランスさせるように、複数の仮想機械のうちの1つまたは複数のクラスタの異なる物理プロセッサ上で動作するように再分散することを含む。

40

【0010】

本発明のさらに別の特徴は、複数の仮想機械のための負荷バランシング装置を提供する。この装置および仮想機械は、スイッチ型ファブリック接続を介して記憶装置システムに結合されるブレード・センタを有するシステム内で動作可能である。このブレード・センタは、複数のプロセッサ/メモリ複合体、および複数の記憶装置コントローラを備える記憶装置システムを含む。この負荷バランシング装置は、複数の記憶装置コントローラのうちの1つまたは複数に関する記憶装置負荷率情報を生成するための、記憶装置システム内

50

で動作可能な記憶装置システム負荷率情報ジェネレータを含む。この装置は、ブレード・センタ内で動作可能な、かつ記憶装置システム負荷率情報ジェネレータに通信可能に結合される負荷バランサも含む。この負荷バランサは、ジェネレータから記憶装置負荷率情報を受け取るようになされ、かつさらにこの記憶装置負荷率情報を、複数の仮想機械のうちの１つまたは複数の動作をブレード・センタの複数のプロセッサ／メモリ複合体にわたって分散するのに利用するようになされている。

【図面の簡単な説明】

【 0 0 1 1 】

【図 1】記憶装置関連負荷率情報を生成し、かつそれを複数の物理プロセッサ上で動作する複数の仮想機械と関連する負荷バランサに通信するための、本発明の特徴および態様により高められた例示的なシステムのブロック図である。

10

【図 2】本発明の特徴および態様による、図 1 のシステムの 1 つまたは複数の記憶装置コントローラ内に一体化される記憶装置関連負荷率情報ジェネレータの例示的な実施形態を説明するブロック図である。

【図 3】本発明の特徴および態様による、図 1 のシステムの 1 つまたは複数の記憶装置コントローラ内に一体化される記憶装置関連負荷率情報ジェネレータの例示的な実施形態を説明するブロック図である。

【図 4】記憶装置システム内に記憶装置関連負荷率情報を生成し、かつ生成された情報を、複数の物理プロセッサ上で動作可能な複数の仮想機械の負荷をバランスさせるのに利用するために負荷バランサに通信するための、本発明の特徴および態様による例示的な方法を説明する流れ図である。

20

【図 5】複数の物理プロセッサ上で動作可能な複数の仮想機械用の負荷バランサで使用するために、記憶装置関連負荷率情報を生成し、かつそれを通信するための、図 4 の方法の実施形態の例示的な追加の詳細を提供する流れ図である。

【図 6】複数の物理プロセッサ上で動作可能な複数の仮想機械用の負荷バランサで使用するために、記憶装置関連負荷率情報を生成し、かつそれを通信するための、図 4 の方法の実施形態の例示的な追加の詳細を提供する流れ図である。

【図 7】複数の物理プロセッサ上で動作可能な複数の仮想機械用の負荷バランサで使用するために、記憶装置関連負荷率情報を生成し、かつそれを通信するための、図 4 の方法の実施形態の例示的な追加の詳細を提供する流れ図である。

30

【発明を実施するための形態】

【 0 0 1 2 】

図 1 は、システム 1 0 0 内で記憶装置関連負荷率情報を生成しかつ利用するための、本発明の特徴および態様を実施する例示的なシステム 1 0 0 のブロック図である。システム 1 0 0 は、複数の物理プロセッサ、すなわちプロセッサ A 1 0 4、プロセッサ B 1 0 6、およびプロセッサ C 1 0 8 を備えるクラスタ化されたコンピュータ環境 1 0 2 を含むことができる。この分野で一般に知られているように、そのようなクラスタ化された環境は、複数のプロセッサを結合するためのバックプレーンまたはミッドプレーン相互接続 1 5 2（または他の適切な相互接続構造）を含むことができる。そのような相互接続 1 5 2 の例には、様々なプロセッサ 1 0 4 から 1 0 8 が互いに通信するのを可能にするための、様々な P C I バス構造、A M B / A H B バス構造、および様々な高速シリアル相互接続構造およびプロトコルが含まれる。

40

【 0 0 1 3 】

複数の仮想機械は、クラスタ 1 0 2 のプロセッサ上で動作することができる。さらにこの複数の仮想機械は、クラスタ 1 0 2 内の複数の物理プロセッサのうちの 1 つまたは複数にわたって分散することができる。例えば、プロセッサ A 1 0 4 は、仮想機械 V 1 および仮想機械 V 2 が現在その上で動作可能であることを示す。プロセッサ B 1 0 6 は、仮想機械 V 3 から V 7 がそのプロセッサを利用して動作可能であることを示す。同様に、プロセッサ C 1 0 8 は、それがやはり仮想機械 V 7（例えば、V 7 は 2 台のプロセッサ 1 0 6 および 1 0 8 にわたって分散されている）によって利用されていることを示す。

50

【 0 0 1 4 】

システム 1 0 0 は、複数の物理プロセッサ上で動作する仮想機械の負荷を等化するまたはバランスさせるのを助けるために、相互接続 1 5 2 を介して複数の物理プロセッサ 1 0 4 から 1 0 8 と通信可能に結合される負荷バランサ 1 0 1 も含む。クラスタ 1 0 2 の負荷バランサ 1 0 1 は、例示として、クラスタ 1 0 2 の様々なプロセッサ 1 0 4 から 1 0 8 から区別される別体として図 1 に描かれている。当業者は、この負荷バランサ 1 0 1 は、クラスタ 1 0 2 内のそのような別個の、区別される処理要素として実施することができ、またはクラスタの特定のプロセッサ内に一体化することもでき、あるいは複数の物理プロセッサの各々内に存在する協働通信プロセスとして分散することもできることを容易に理解するであろう。そのような設計選択は、当業者に良く知られている。

10

【 0 0 1 5 】

上記で言及したように、現在実施される負荷バランシング技術および構造は一般に、各特定のプロセッサに関連しかつ特定のプロセッサによって求めることが可能な利用率メトリックのみ対処する。例えば、現在知られている負荷バランシング技術は、特定の物理プロセッサの CPU 利用率、その特定のプロセッサのメモリ利用率、そのプロセッサによって発生する I / O 要求割合、等（すなわち、各プロセッサによってそのプロセッサに対し求められるメトリック）のみ対処する傾向がある。現在実施されているように、複数の物理プロセッサ上で動作可能な仮想機械用の負荷バランシング技術は、複数の記憶装置コントローラ 1 1 2 から 1 1 6 が動作に利用可能である可能性がある記憶装置システム 1 1 0 などの共通の記憶装置に対する共有のアクセスを考慮に入れていない。例えば、現在の負荷バランシング技術は、複数の物理プロセッサ 1 0 4 から 1 0 8 のうちの 1 つまたは複数の上で動作可能ないくつかの仮想機械が、共通の記憶装置システム 1 1 0 の複数の記憶装置コントローラ 1 1 2 から 1 1 6 のうちのただ 1 つのみに対するアクセスを共有していることを理解できない可能性がある。その特定の共有の記憶装置コントローラが過負荷になる場合、その記憶装置コントローラに対するアクセスを共有する対応する仮想機械は、最適とは言い難く機能している可能性がある。例えば、特定の記憶装置コントローラに対するアクセスを共有する 1 つまたは複数の仮想機械が記憶装置システム 1 1 0 の複数の記憶装置コントローラ 1 1 2 から 1 1 6 のうちの異なる記憶装置コントローラと相互作用できるように、仮想機械をクラスタ 1 0 2 の異なるプロセッサに移動させることも可能である可能性がある。しかしながら、現在の負荷バランシング技術および構造は、そのような負荷バランシングを可能にせず、実際、そのような決定を可能にするための記憶装置システム動作に関連する負荷率に関する有用な情報を一般に欠いている。

20

30

【 0 0 1 6 】

本発明の特徴および態様によれば、記憶装置システム 1 1 0 は、記憶装置システム 1 1 0 の動作を全体として、かつ記憶装置システム 1 1 0 の複数の記憶装置コントローラ 1 1 2 から 1 1 6 の動作を解析するように動作可能な負荷率情報ジェネレータ要素 1 1 8 を含む。次いでジェネレータ 1 1 8 は、様々な記憶装置 1 1 2 から 1 1 6 の、かつ全体としてのシステム 1 1 0 の処理能力に関する負荷率情報を生成することができる。次いで要素 1 1 8 によって発生されたそのような負荷率情報は、クラスタ 1 0 2 の複数の物理プロセッサ 1 0 4 から 1 0 8 上で動作する複数の仮想機械の負荷をどのように分散または再分散するか決定するとき、負荷バランサ 1 0 1 が記憶装置コントローラおよび記憶装置システムの負荷率情報を考慮することができるように、クラスタ 1 0 2 内で動作可能な負荷バランサ 1 0 1 に通信される。

40

【 0 0 1 7 】

この分野で一般的に知られているように、クラスタ 1 0 2 の複数の物理プロセッサ 1 0 4 から 1 0 8 は、スイッチ型ファブリック結合 1 5 0 を介して記憶装置システム 1 1 0 の様々な記憶装置コントローラ 1 1 2 から 1 1 6 に通信することができる。そのようなスイッチ型ファブリックは、シリアル付加 SCSI (「 SAS 」) スwitch型ファブリック、 Fiber Channel スwitch型ファブリック、または任意の他のいくつかの良く知られた、市販のコンピュータおよび / または記憶装置ネットワーク構造およびプロトコ

50

ルとして実施することができる。負荷率情報ジェネレータ 118 は一般に、記憶装置関連負荷率情報を収集し、そのような情報を記憶装置コントローラおよびスイッチ型ファブリック 150 を介してクラスタ 102 内の負荷バランサ 101 に通信するように、記憶装置コントローラ 112 から 116 の各々に通信可能に結合される。

【0018】

通信パス 154 は、負荷率情報ジェネレータ 118 が記憶装置システム 110 の記憶装置コントローラ 112 から 116 の各々から情報を収集するのを可能にする、任意の適切なコントローラ間通信構造およびプロトコルを一般に意味する。パス 154 も、負荷率情報ジェネレータ 118 が、記憶装置コントローラ 112 から 116 のうちの 1 つまたは複数を介して、ファブリック 150 を経由して、1 つまたは複数のプロセッサ 104 から 108 を介して、生成された情報をクラスタ 102 (具体的には負荷バランサ 101) に通信するのを可能にする。より具体的には、負荷率情報ジェネレータ 118 は、記憶装置コントローラ 112 から 116 から情報を収集し、そのような収集された情報を負荷バランサ 101 に送るように負荷率情報としてフォーマット化するための任意の適切な構造および/またはプロセス間通信技術を意味することができる。したがってジェネレータ 118 は、複数の記憶装置コントローラ 112 から 116 とは別個のかつ区別される記憶装置システム 110 内のコンポーネントであることができ、あるいは 1 つまたは複数の記憶装置コントローラ内に一体化されかつ動作可能であることもできる。

【0019】

記憶装置システム 110 の記憶装置コントローラ 112 から 116 は、スイッチ型ファブリック結合 160 を介して記憶装置システム 110 の複数の記憶装置 120 から 122 と通信することもできる。上述のように、シリアル付加 SCSI (「SAS」)、パラレル SCSI、Fiber Channel、および他の良く知られた、市販の通信構造およびプロトコルを、記憶装置コントローラ 112 から 116 の各々を記憶装置 120 と 122 の各々に結合するために使用することができる。コントローラ間通信チャンネル 154 は、スイッチ型ファブリック結合 160 を介したコントローラからコントローラへの通信として単に実施することもできる。別個のかつ区別される通信チャンネル 154 が、コントローラ間通信の目的で様々な記憶装置コントローラ 112 から 116 とジェネレータ 118 を結合し、スイッチ型ファブリック結合 160 の全バンド幅を記憶装置コントローラ 112 から 116 と記憶装置 120 および 122 との間のデータ交換に使用可能に残すのが好ましい。

【0020】

図 1 に例示するように、システム 100 の様々なコンポーネントを結合するスイッチ型ファブリック 150 は、例えば、複数のプロセッサ (104 から 108) のうちのいずれかが、記憶装置システム 110 内の複数の記憶装置コントローラ (112 から 116) のいずれかに制御可能に結合されるのを可能にする。シリアル付加 SCSI (「SAS」) の使用は、現在の技術を適用する一般的な設計選択の例示である。しかしながら当業者は、市販の製品を利用する様々な他のスイッチ型ファブリック接続性アーキテクチャがそのようなフレキシブルな接続性を可能にすることを容易に理解するであろう。その上、本発明の特徴および態様は、プロセッサと記憶装置コントローラ間の接続性がスイッチ型ではなく、オペレータの構成およびインストールによって静的に決められる場合にも実施することができる。したがって図 1 に示されていないが、本発明の特徴および態様は、プロセッサと記憶装置コントローラ間のスイッチ型ファブリック接続性を欠いているシステムにも適用することができる。そのような場合、さらに本明細書で論じられるような負荷率情報は、複数のプロセッサのうちの特定の 1 つに、それに静的に結合される記憶装置コントローラのうちの対応する特定の 1 つから通信することができる。次いで受信プロセッサによる記憶装置関連負荷率情報の利用は、様々なプロセッサ上で動作可能な 1 つまたは複数の仮想機械を複数のプロセッサの全ての上にその負荷をより良くバランスさせるように再分散できるように、最初にこの情報を他のプロセッサとかがつ/または負荷バランサ要素 101 と共有することを含むことができる。

10

20

30

40

50

【 0 0 2 1 】

上記で言及しかつさらに本明細書で以下に論じるように、記憶装置システム 1 1 0 の記憶装置コントローラ 1 1 2 から 1 1 6 の動作に関するこの生成された負荷率情報は、記憶装置システム 1 1 0 内のジェネレータ 1 1 8 からクラスタ 1 0 2 内の負荷バランサ 1 0 1 に任意のいくつかの技術によって通信することができる。例えば、記憶装置システム 1 1 0 に関連する負荷率情報は、任意の I / O 動作に対して非同期的に - 例えば、周期的に生成するまたは特定の変化の検出に応答して生成するのいずれかで - 生成することができ、記憶装置コントローラ 1 1 2 から 1 1 6 を介してスイッチ型ファブリック結合 1 5 0 を経由して負荷バランサ 1 0 1 に、任意のプロセッサ 1 0 4 から 1 0 8 を介して転送することができる。したがって、負荷バランサ 1 0 1 は、複数のプロセッサ 1 0 4 から 1 0 8 のうちの 1 つまたは複数上で動作し、記憶装置システム 1 1 0 の現行の負荷情報を反映するような周期的なまたは間歇的な伝達を受け取るようになされた専門の管理アプリケーションを含むことができる。その上、負荷率情報ジェネレータ 1 1 8 の動作は、記憶装置システム 1 1 0 の記憶装置コントローラ 1 1 2 から 1 1 6 の各々の中の I / O 処理と緊密に一体化することができる。したがって、記憶装置システム 1 1 0 ならびにその中の各コントローラ (1 1 2 から 1 1 6) の動作に関連する現行の負荷率情報は、クラスタ 1 0 2 からの I / O 要求の処理と関連する返されたステータス情報の一部分として提供することができる。次いで記憶装置システム 1 1 0 に関するそのような負荷率情報は、複数の物理プロセッサ 1 0 4 から 1 0 8 とその上で動作可能な仮想機械との協働的な統合化を介して、返されたステータス情報から負荷バランサ 1 0 1 によって取り出すことができる。記憶装置システム 1 1 0 内で発生する負荷率情報がクラスタ 1 0 2 に返される方法に関わらず、負荷バランサ 1 0 1 は、記憶装置システム 1 1 0 に関連するそのような負荷率情報をクラスタ 1 0 2 の複数の物理プロセッサ上で動作可能な複数の仮想機械の負荷を分散または再分散するのに利用するように動作可能である。

【 0 0 2 2 】

当業者は、完全に機能的なシステム 1 0 0、クラスタ 1 0 2 または記憶装置システム 1 1 0 に対して、図 1 に描かれたものを超える多数の均等なかつ追加の要素を容易に理解するであろう。したがって図 1 は、複数の仮想機械 (または他のアプリケーション) の負荷を、クラスタに結合される記憶装置システム内の複数の記憶装置コントローラに関する負荷率情報に基づいて、クラスタ内の複数のプロセッサにわたって分散または再分散するためのクラスタ・コンピュータ環境内に実施されるような、本発明の特徴および態様の単なる例示として意図としている。そのような追加の特徴は、この論議を単に簡単かつ簡潔にするために省略されている。

【 0 0 2 3 】

図 1 に示すように、負荷率情報ジェネレータ 1 1 8 は、記憶装置システム 1 1 0 内の複数の記憶装置コントローラ 1 1 2 から 1 1 6 に対して別個かつ区別されるとして描かれている。しかし、上記で言及したように、この負荷率情報発生機能は、記憶装置コントローラの処理内に密接に一体化することもできる。したがって図 2 は、負荷率情報ジェネレータ 2 1 8 が複数の記憶装置コントローラ (2 1 2 から 2 1 6) の指定される監督記憶装置コントローラ 2 1 2 内に一体化される、記憶装置システム 1 1 0 の例示的な実施形態を示す。記憶装置システム 1 1 0 の他の記憶装置コントローラ 2 1 4 から 2 1 6 は、負荷率情報を監督記憶装置コントローラ 2 1 2 内で動作可能な負荷率情報ジェネレータ 2 1 8 と交換するように、監督記憶装置コントローラ 2 1 2 と協働的に動作可能である。

【 0 0 2 4 】

同様に、図 3 は、負荷率情報発生が複数の記憶装置コントローラの各々を介して分散される別の例示的な実施形態を示す。したがって、図 3 の記憶装置システム 3 1 0 は、それぞれ対応する一体化された負荷率情報ジェネレータコンポーネント 3 1 8 (A)、3 1 8 (B)、および 3 1 8 (C) をその各々が有する、3 つの記憶装置コントローラ 3 1 2 から 3 1 6 を示す。

【 0 0 2 5 】

10

20

30

40

50

図 2 および 3 では、複数の記憶装置コントローラ 2 1 2 から 2 1 6 および 3 1 2 から 3 1 6 は、任意のいくつかの良く知られた、市販の、コントローラ間通信媒体およびプロトコルを利用して通信することができる。したがって負荷率情報は、複数の記憶装置コントローラのいずれかまたは全てに関連する複数の記憶装置コントローラのうちの任意の 1 つまたは複数によって収集することができる。そのようなコントローラ間通信媒体およびプロトコルには、例えば、InfiniBand、Fibre Channel、シリアル付加 SCSI (SAS)、(PCI 標準品のいくつかの変形形態を含む) PCI バス構造、等が含まれ得る。したがって当業者は、全体としてこの記憶装置システムに関連するかつ/または記憶装置システム内の個々の記憶装置コントローラの各々に関連する負荷率情報を収集し、かつ付加されるクラスタ・コンピュータ環境内の適切な要素に通信する、様々な均等な構造を容易に理解するであろう。さらに当業者は、完全に機能的な記憶装置システム 1 1 0 内に多数の追加の要素が通常存在することを容易に理解するであろう。そのような追加の要素は、この論議の単純化および簡潔化のために、図 2 および 3 から削除されている。

10

20

30

40

50

【0026】

図 4 は、図 1 に対して上記で示しかつ説明したようなシステム内で動作可能な、本発明の特徴および態様による方法を大ざっぱに説明する流れ図である。要素 4 0 0 は、記憶装置システム内の任意の適切な手段によって最初に記憶装置関連負荷率情報を生成するように動作可能である。上記で言及したように、負荷率情報ジェネレータ要素は、記憶装置システムの記憶装置コントローラのうちのいずれかまたは全ての中で動作可能であり、あるいは記憶装置コントローラに対して別個のかつ区別される、かつそれに結合されるコンポーネントとして動作可能である。そのような負荷率情報は、関連する記憶装置システムおよび/または記憶装置コントローラの各々の動作の任意の数の態様を含むことができる。例えば、記憶装置システムおよび/または記憶装置コントローラの動作に関連する様々な種類の負荷率情報の中には、

- ・記憶装置システムの各記憶装置コントローラと現在通信している仮想機械の数；
 - ・複数の記憶装置コントローラのうちのいずれかと結合される仮想機械の任意の特定されるセットの I/O 要求処理率；
 - ・記憶装置コントローラの任意の特定されるセットと複数のプロセッサ上で動作可能な仮想機械の任意のセットの間の測定されるデータ・スループット率；
 - ・記憶装置コントローラの任意の特定されるセット内に係属中の、待機または未処理の I/O 要求の数；
 - ・仮想機械の任意の特定されるセットから発生する、待機または未処理の I/O 要求の数；
 - ・記憶装置コントローラの任意の特定されるセットと複数のプロセッサ上で動作可能な仮想機械の任意の特定されるセットの間の可能な通信バスの最大数；
 - ・仮想機械の任意の特定されるセットに現在結合される記憶装置コントローラの総計；
- および
- ・物理プロセッサの総数、がある。

【0027】

記憶装置システム内の負荷率情報ジェネレータ要素の動作によるそのような記憶装置関連負荷率情報を生成した後、要素 4 0 0 は、複数の物理プロセッサ上で動作可能な複数の仮想機械と関連する負荷バランス要素へのこの生成された負荷情報の通信を達成するように動作可能でもある。次いでこの負荷バランスは、複数の物理プロセッサ上で動作可能な複数の仮想機械を再分散するために、この記憶装置関連負荷率情報を(負荷バランス決定に対し有用な任意の他の情報と共に)利用することができる。この記憶装置関連負荷率情報を利用することによって、仮想機械の再分散は、記憶装置システムの処理および通信バンド幅ならびに物理プロセッサ活用の他の態様をより良好に利用できる。

【0028】

上記で言及しかつさらに本明細書で以下に論じるように、生成された記憶装置関連負荷

率情報の通信は、任意のいくつかの技術および構造を利用することができる。その上、この発生する情報は、複数の仮想機械と関連する負荷バランスを更新するために周期的に生成することができ、記憶装置システム内で検出される事象または変化した状態に基づいて間歇的に生成することができ、あるいは記憶装置システムによって受信されかつ処理される、複数の物理プロセッサ上で動作可能な1つまたは複数の仮想機械から記憶装置システムによって受信されかつ処理されるI/O要求の処理の一部として生成しかつ返すことができる。

【0029】

したがって、図4の要素402は、必要に応じて複数の仮想機械のうちの1つまたは複数を複数の物理プロセッサにわたって再分散するための、複数の仮想機械と関連する負荷バランス要素の処理を意味する。この再分散は、もしあれば、複数の物理プロセッサ上の負荷のより良好なバランスを実現するのを助け、したがって、複数の仮想機械のより良好な全体性能をもたらす。

【0030】

当業者は、図4の処理の多数の追加のステップおよび例示的な詳細な実施形態を容易に理解するであろう。詳細な処理のいくつかの例示的な実施形態はさらに本明細書で以下に図5から7に関して論じられる。他の実施形態および追加のステップは、当業者には容易に明らかであろうので、単にこの論議の簡潔化のために本明細書では省略されている。

【0031】

図5は、上記で論じた図4の要素400の例示的な詳細処理を提供する流れ図である。一般に図5の処理は、内在するステータスの変化が記憶装置システム内で検出されるときのみ、記憶装置システムと関連する負荷率情報を生成する。したがって、負荷率情報に影響を与える可能性のある記憶装置システムのステータス内で（おそらく、記憶装置システムによって検出される付加システムおよび仮想機械内の変化を含む）変化を検出することによって必要になったとき、負荷率情報が間歇的に生成される。したがって要素500は、負荷率情報が生成された最後のとき以来、いずれかの記憶装置システムのステータスが変化したかどうか判定するための処理を表す。有意な変更が全く検出されない場合は、処理は関連の変更が検出されるなどのそのようなときまで要素500上でループを続ける。記憶装置システム・ステータス内でそのような変更が検出される際、次いで要素502は、記憶装置システム内の負荷率関連情報の現行のステータスを反映する更新される負荷率情報を生成するように動作可能になる。次いで要素504は、生成された、更新された負荷率情報を複数の物理プロセッサ上で動作する複数の仮想機械と関連する負荷バランス要素に伝達するように動作可能になる。したがって、図5に詳細に示す要素400は、記憶装置システムのステータスに何らかの有意な変更が検出される際、記憶装置システムから負荷率情報を生成し、それを負荷バランス要素に伝達するように動作可能である。

【0032】

図6は、上記で論じた図4の要素400の例示的な詳細処理を提供する流れ図である。一般に図6の処理は、仮想機械および/または物理プロセッサから受け取った記憶装置システムによるI/O要求の受信および処理に応答して、現行の負荷率情報を生成しかつ返すように動作可能である。特に、要素600は、記憶装置システム内のI/O要求の受信と通常の処理を意味する。1つまたは複数の仮想機械が上で動作可能な複数の物理プロセッサのうちの1つまたは複数から受け取ったそのようなI/O要求は、記憶装置システム内の情報を検索するまたはシステム内に情報を記憶させる典型的なリードまたはライト要求を意味することができる。そのような通常の処理は、要求されるデータを記憶しまたは検索し、かつその要求が発せられた要求仮想機械および/または物理プロセッサに返すべきステータス応答を通常発生する。次いで要素602は、記憶装置システムおよび/または記憶装置システムの記憶装置コントローラに関する現行の負荷率情報を生成するように動作可能になる。この生成された情報は、要素600内に示されるようなI/O要求の処理によって発生する結果/ステータス情報に追加または一体化することができる。例えば、完了したI/O要求に対する返された/ステータス情報は、正常完了コードおよびエラ

10

20

30

40

50

ー完了コードを示すビット・フィールドおよび／または値をしばしば含む。そのようなステータス完了情報内のビットまたは値は、現行の記憶装置関連負荷率情報がＩ／Ｏ動作に対して返されたステータス・コードで追加または一体化された物理プロセッサおよび／または仮想機械に示すようにコード化することができる。最後に要素６０４は、要求している物理プロセッサに関連する負荷バランサ要素がその負荷バランシング決定に記憶装置関連負荷率情報を含むことができるように、この追加／一体化負荷率情報と一緒にＩ／Ｏ要求ステータス情報を伝達するように動作可能になる。したがって、図６に示すような要素４００は、付加クラスタ内の複数の物理プロセッサのうちのいずれかの上で動作可能な任意の複数の仮想機械から記憶装置システムに向けられるＩ／Ｏ要求の処理の際に、記憶装置システムから負荷率情報を生成し、かつそれを負荷バランサ要素に伝達するように動作可能であることができる。

10

【００３３】

図７は、図４の要素４００の処理の例示的な詳細を提供する流れ図である。一般に図７の処理は、負荷率情報の値が関連する閾値の範囲外に該当するとき、記憶装置関連負荷率情報を生成しかつ伝達するように動作可能である。したがって要素７００は、現行の負荷率情報を生成するように最初に動作可能である。この動作は、周期的にまたは様々な特定の事象の検出に基づいて間歇的に実施させることができる。次いで要素７０２は、負荷率情報内の値に対して比較するための現行の更新された閾値を生成するための任意選択の処理を表す。上記で言及したように、この閾値は記憶装置システムの製造時に静的に、または記憶装置システムのセットアップおよび設定での機器構成ステップとして決めることができ、あるいは記憶装置システムの変化する動作に基づいて動的に決めることもできる。次いで要素７０４は、要素７００によって生成された負荷率情報内のどのような値も対応する閾値によって画定される範囲外に該当するかどうか判定するために、（以前に静的に生成されていようと、または要素７０２の処理によって現行更新されていようとに関わらず）現行の閾値情報を利用するように動作可能である。範囲外に該当しない場合は、要素４００の処理は完了する。そうでない場合は、現行の、更新された、記憶装置関連負荷率情報が、付加クラスタ内の複数の仮想機械および複数の物理プロセッサと関連する負荷バランサ要素に伝達される。次いでそのような情報の受信に応答して、この負荷バランサ要素は、複数の仮想機械のうちの１つまたは複数を複数の物理プロセッサのうちの他のものに再分散することを選択することができる。したがって、図７に詳細に示される要素４００は、負荷率情報内の任意の値が対応する閾値によって画定される値の範囲外に該当したことを検出することに応答して、更新した負荷率情報を生成しかつ伝達するように動作可能である。さらに、言及したように、この閾値は静的に決めることができ、あるいは記憶装置システム内の現行の動作に基づいて動的に更新することができる。例示的な閾値は、負荷率情報の代表として上記で説明した例示的な値に対応することができる。

20

30

【００３４】

当業者は、完全に機能的な実施で、図４から７までの方法に組み込むことができる均等なまたは追加のステップを容易に理解するであろう。そのような追加の特徴およびステップは、この論議の簡単化および簡潔化のために本明細書では省略されている。さらに当業者は、例示的に詳細化された本方法の特徴を組み合わせることによって、上記の方法の特徴を実施するための他の詳細な方法を理解するであろう。例えば、負荷率情報はある値が画定される閾値範囲外に該当するときのみ生成することができ、そのように生成された場合、次いで完了したＩ／Ｏ要求の完了ステータスと共に返すことができる。多数の他の均等なステップおよび方法は、当業者に容易に明らかであろう。

40

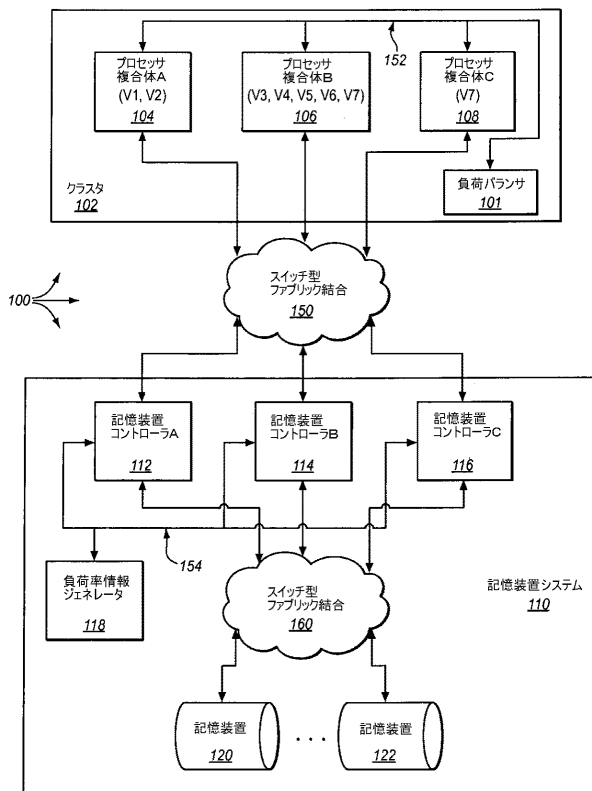
【００３５】

本発明は図面に図示され前述の説明で説明されているが、そのような図示および説明は、性格的に限定的ではなく例示として理解されたい。本発明の１つの実施形態およびその小さな変形形態が示され、説明されてきた。本発明の趣旨内に入る全ての変更および改変に対して保護が所望される。当業者は、本発明の範囲内に該当する上記で説明した実施形態の変形形態を理解するであろう。特に、当業者は、本発明の特徴および態様は、電子回

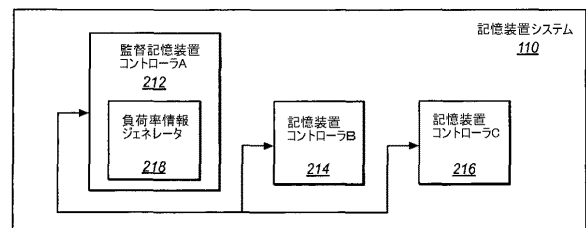
50

路内で均等にあるいは一般的なまたは特殊な目的のプロセッサの適切にプログラムされる命令として実施できることを容易に理解するであろう。回路またはプログラミング設計のそのような均等性は、設計選択の問題として当業者に良く知られている。結果として、本発明は上記で論じた特定の例および図示に限定されず、添付の特許請求の範囲およびそれらの均等物によってのみ限定される。

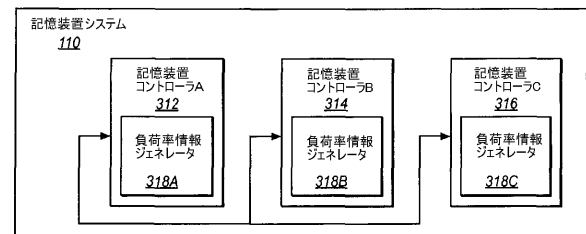
【 図 1 】



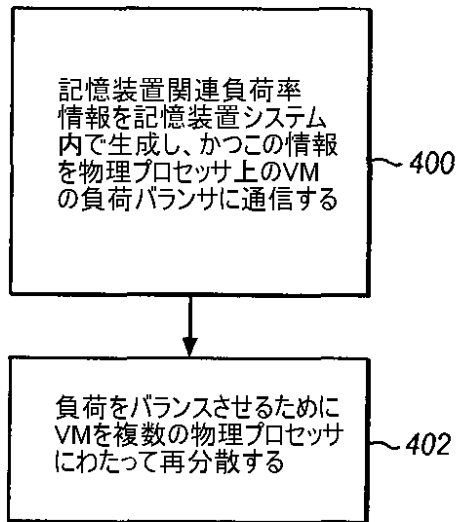
【 図 2 】



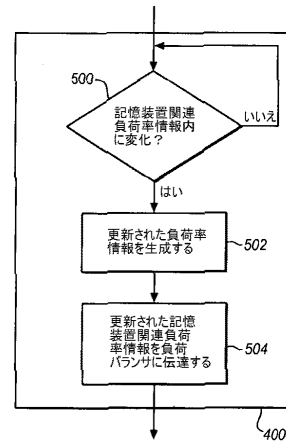
【 図 3 】



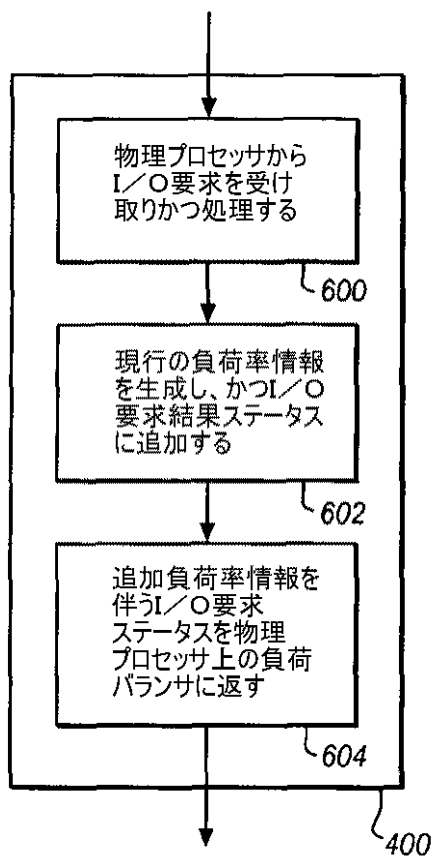
【図 4】



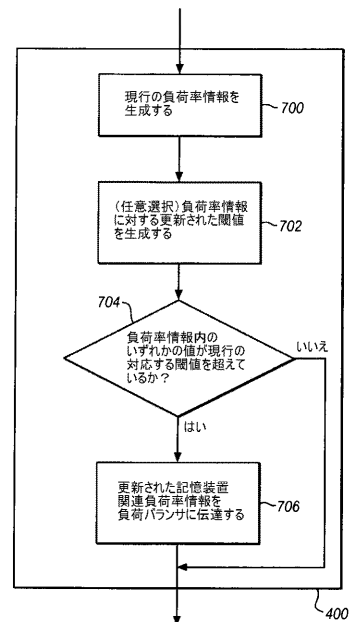
【図 5】



【図 6】



【図 7】



【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US07/79145
A. CLASSIFICATION OF SUBJECT MATTER IPC: G06F 13/00(2006.01) USPC: 711/6,154,165;718/105 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) U.S. : 711/6, 154, 165; 718/105		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched IEEE NPL		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EAST		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 7,203,944 B1 (van Rietschote et al), col. 3, lines 20-43, col. 4, line 59 - col. 5, line 21.	1-19
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 23 June 2008 (23.06.2008)		Date of mailing of the international search report 07 JUL 2008
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (571) 273-3201		Authorized officer STEPHEN C. ELMORE Telephone No. (571) 272-2100

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(74)代理人 100160967

弁理士 濱 口 岳久

(72)発明者 ジョンソン, スチーヴン ビー.

アメリカ合衆国 80909 コロラド, コロラド スプリングス, ロッチ ロモンド レーン
4225