

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4555140号
(P4555140)

(45) 発行日 平成22年9月29日 (2010. 9. 29)

(24) 登録日 平成22年7月23日 (2010. 7. 23)

(51) Int. Cl.

F I

G 0 6 F 1/26 (2006. 01)

G 0 6 F 1/00 3 3 0 C

G 0 6 F 1/32 (2006. 01)

G 0 6 F 1/00 3 3 2 Z

G 0 6 F 1/28 (2006. 01)

G 0 6 F 1/00 3 3 3 Z

請求項の数 5 (全 22 頁)

(21) 出願番号 特願2005-124517 (P2005-124517)
 (22) 出願日 平成17年4月22日 (2005. 4. 22)
 (65) 公開番号 特開2006-302059 (P2006-302059A)
 (43) 公開日 平成18年11月2日 (2006. 11. 2)
 審査請求日 平成20年2月4日 (2008. 2. 4)

(73) 特許権者 000005108
 株式会社日立製作所
 東京都千代田区丸の内一丁目6番6号
 (74) 代理人 110000350
 ポレール特許業務法人
 (74) 代理人 100068504
 弁理士 小川 勝男
 (74) 代理人 100086656
 弁理士 田中 恭助
 (74) 代理人 100094352
 弁理士 佐々木 孝
 (72) 発明者 馬場 貴成
 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】 複合型計算機装置およびその管理方法

(57) 【特許請求の範囲】

【請求項 1】

各装置に電力を供給する少なくとも一つのAC-DC電源を有するAC-DC電源モジュールと、
 少なくとも一つのプロセッサを搭載し、前記AC-DC電源モジュールから供給される電力
 をDC-DC変換するDC-DC電源を搭載し、前記プロセッサの非稼動時にも電源供給制御を行い
 、前記AC-DC電源モジュールと電源制御用インターフェースで接続されAC-DC電源モジュー
 ルの出力制御を行うCPUモジュールコントローラとを搭載したCPUモジュールと、

管理用のインターフェースで前記CPUモジュールコントローラに接続され前記CPUモジュー
 ルの電源制御を行う管理モジュールと、

前記CPUモジュールの電源ON時にはCPUモジュール毎に稼動するオペレーションシステム
 およびシステムファームウェアとを有し、

前記プロセッサがスリープ状態を有する場合に、

前記プロセッサのスリープ状態として少なくとも一つの状態、または前記CPUモジュー
 ル毎にシステムのスリープ状態として少なくとも一つの状態、のいずれかのスリープ状態
 を持ち、

前記CPUモジュールコントローラは、

前記プロセッサの状態が通常稼動状態からスリープ状態に遷移する場合には、

前記システムファームウェアから通知されるスリープ状態への遷移命令により、一定時
 間スリープするためのタイマ機能を持ち、

また前記管理モジュールへプロセッサ状態変更を通知する機能を持ち、

10

20

前記プロセッサの状態がスリープ状態から通常稼動状態に遷移する場合には、
前記タイマ機能復帰を契機に前記管理モジュールへのプロセッサ状態変更を通知する機能を持ち、

前記管理モジュールからの前記AC-DC電源モジュールの設定完了通知を受けてスリープから通常稼動状態に戻る割り込みをプロセッサに投入する機能を持ち、

前記管理モジュールは、システム情報として、CPUモジュールの電源ONまたはOFFと、前記CPUモジュールのシステムの通常稼動状態およびスリープ状態と、搭載されたプロセッサの種類の情報とを含むCPUモジュール稼動情報を持ち、

プロセッサ情報として、前記CPUモジュール内のプロセッサの通常稼動状態およびスリープ状態のプロセッサの状態情報と、前記プロセッサの種類毎に、それぞれの状態に必要なとなるプロセッサ消費電力情報を持ち、

10

前記プロセッサの状態が通常稼動状態からスリープ状態に遷移する場合または前記プロセッサの状態がスリープ状態から通常稼動状態に遷移する場合には、

前記CPUモジュールコントローラからのプロセッサ状態変更の通知を受けて、前記CPUモジュール稼動情報と、前記プロセッサの状態情報と、前記プロセッサ消費電力情報とから複合型計算機装置全体に必要な消費電力を計算し、その消費電力に基づいてAC-DC電源モジュールの稼動数を計算し、AC-DC電源モジュール稼動数の変更が必要な場合はAC-DC電源モジュールの稼動数を変更する制御を行う機能を持ち、

前記CPUモジュールコントローラへ前記AC-DC電源モジュールが安定して稼動していることを示す設定完了を通知する機能を持つことを特徴とする複合型計算機装置。

20

【請求項 2】

請求項 1 記載の複合型計算機装置において、

前記システムのスリープ状態を持つ場合、

前記CPUモジュールコントローラは、

前記システムの状態が通常稼動状態からスリープ状態に遷移する場合に、

前記システムファームウェアからの通知されるスリープ状態への遷移命令により、管理モジュールへのプロセッサ状態変更を通知する機能を持ち、

前記システムの状態がスリープ状態から通常稼動状態に遷移する場合に、

前記管理モジュールからの通知されるCPUモジュールの起動要求を契機にCPUモジュールを起動させる機能を持ち、

30

前記管理モジュールは、システムの状態が通常稼動状態からスリープ状態に遷移する場合に、

前記CPUモジュールコントローラからのシステム状態変更の通知を受けて、前記CPUモジュール稼動情報と、前記プロセッサの状態情報と、前記プロセッサ消費電力情報とから複合型計算機装置全体に必要な消費電力を計算し、その消費電力に基づいて前記AC-DC電源モジュールの稼動数を計算し、AC-DC電源モジュール稼動数の変更が必要な場合はAC-DC電源モジュールの稼動数を変更する制御を行う機能を持ち、

前記システムの状態がスリープ状態から通常稼動状態に遷移する場合に、

前記CPUモジュールの電源投入契機を受ける機能を持ち、

前記CPUモジュール稼動情報と、前記プロセッサの状態情報と、前記プロセッサ消費電力情報とから複合型計算機装置全体に必要な消費電力を計算し、その消費電力に基づいてAC-DC電源モジュールの稼動数を計算し、AC-DC電源モジュール稼動数の変更が必要な場合はAC-DC電源モジュールの稼動数を変更する制御を行う機能を持ち、

40

前記CPUモジュールコントローラへCPUモジュールの起動要求を通知する機能を持つことを特徴とする複合型計算機装置。

【請求項 3】

請求項 1 記載の複合型計算機装置で、

前記CPUモジュールに搭載されたプロセッサが、2つ以上のプロセッサコアを持つ場合、

前記管理モジュールは、前記プロセッサの状態情報として、

前記CPUモジュール内のプロセッサのコア毎の通常稼動状態およびスリープ状態と、

50

前記CPUモジュールのシステムの通常稼動状態およびスリープ状態とを持ち、この情報を基に、前記プロセッサの通常稼動状態とスリープ状態の遷移の際に、前記複合型計算機装置全体に必要な消費電力を計算し、その消費電力に基づいて前記AC-DC電源モジュールの稼動数を計算し、前記AC-DC電源モジュール稼動数の変更が必要な場合はAC-DC電源モジュールの稼動数を変更する制御を行う機能を持つことを特徴とする複合型計算機装置。

【請求項 4】

装置全体に電力を供給する少なくとも一つのAC-DC電源モジュールと、
少なくとも一つのプロセッサを搭載し、前記AC-DC電源モジュールから供給される電力をDC-DC変換するDC-DC電源を搭載し、前記プロセッサの非稼動時にも電源供給制御を行い、前記AC-DC電源モジュールと電源制御用インターフェースで接続されAC-DC電源モジュールの出力制御を行うCPUモジュールコントローラを搭載したCPUモジュールと、

管理用のインターフェースで接続され前記CPUモジュールの電源制御を行う管理モジュールと、

前記CPUモジュールの電源ON時にはCPUモジュール毎に稼動するオペレーションシステムおよびシステムファームウェアとを有し、

前記プロセッサがスリープ状態を有する複合型計算機装置の管理方法であって、

前記プロセッサの状態が通常稼動状態からスリープ状態に遷移する場合に、

前記CPUモジュール上で稼動する前記オペレーションシステムは、前記システムファームウェアにプロセッサの状態を通常稼動状態からスリープ状態へ遷移させる命令を通知し、この通知によりCPUモジュール上で稼動する前記システムファームウェアは、前記CPUモジュールコントローラにプロセッサ状態変更を通知し、

また前記プロセッサに対してスリープ制御を行い、前記CPUモジュールコントローラは前記プロセッサ管理モジュールにプロセッサ状態変更を通知し、この通知により前記プロセッサ管理モジュールは前記プロセッサCPUモジュールのプロセッサの通常稼動状態およびスリープ状態のプロセッサの状態情報を更新し、このプロセッサの状態情報と、前記プロセッサCPUモジュールの電源ONおよびOFF状態と、前記プロセッサCPUモジュールのシステムの通常稼動状態およびスリープ状態および搭載されたプロセッサの種類の情報を持つCPUモジュール稼動情報と、各プロセッサ種類に対して各状態に必要なプロセッサ消費電力情報と、を用いて複合型計算機装置全体に必要な消費電力を計算し、前記消費電力に基づき前記AC-DC電源モジュールの稼動数を計算し、前記AC-DC電源モジュール稼動数の変更が必要な場合はAC-DC電源モジュールの稼動数を変更する制御を行い、

前記AC-DC電源モジュールの稼動状態確認後、前記管理モジュールは前記CPUモジュールコントローラにAC-DC電源モジュール制御完了の通知し、この通知により前記CPUモジュールコントローラは前記プロセッサに対して通常稼動状態へ遷移するための割込み信号を出すことを特徴とする複合型計算機装置の管理方法。

【請求項 5】

装置全体に電力を供給する少なくとも一つのAC-DC電源モジュールと、

少なくとも一つのプロセッサを搭載し、前記AC-DC電源モジュールから供給される電力をDC-DC変換するDC-DC電源を搭載し、前記プロセッサの非稼動時にも電源供給制御を行い、前記AC-DC電源モジュールと電源制御用インターフェースで接続されAC-DC電源モジュールの出力制御を行うCPUモジュールコントローラを搭載したCPUモジュールと、

管理用のインターフェースで接続され前記CPUモジュールの電源制御を行う管理モジュールと、

前記CPUモジュールの電源ON時にはCPUモジュール毎に稼動するオペレーションシステムおよびシステムファームウェアとを有し、

前記CPUモジュールがスリープ状態を有する複合型計算機装置の管理方法において、

システムの状態がスリープ状態から通常稼動状態に遷移する場合に、

前記管理モジュールは、CPUモジュールの電源投入契機を受け、前記プロセッサの状態情報および前記CPUモジュール稼動情報とを更新し、このプロセッサの状態情報と、CPUモ

10

20

30

40

50

ジュール稼動情報と、前記プロセッサ消費電力情報と、を用いて複合型計算機装置全体に必要な消費電力を計算し、前記消費電力に基づいてAC-DC電源モジュールの稼動数を計算し、前記AC-DC電源モジュール稼動数の変更が必要な場合はAC-DC電源モジュールの稼動数を変更する制御を行い、

前記AC-DC電源モジュールの稼動状態確認後、前記管理モジュールは、前記CPUモジュールコントローラにCPUモジュールの電源投入命令を通知し、この通知により前記CPUモジュールコントローラはCPUモジュールを起動させることを特徴とする複合型計算機装置の管理方法。

【発明の詳細な説明】

【技術分野】

10

【0001】

本発明は、計算機装置における管理技術に関して、特に1つの筐体に複数の計算機装置を搭載した複合型計算機装置における電源の制御に適用して有効な技術に関するものである。

【背景技術】

【0002】

企業で用いられる業務システム内のWEBサーバ、アプリケーションサーバ、データベースサーバには、プロセッサ、メモリ、IOデバイス、ストレージデバイス等で構成した計算機装置が用いられる。近年、計算機装置の省スペース性やシステム内のサーバを一箇所に配置し統合して管理をすること等を狙った複合型計算機装置が用いられるようになってい 20
る。複合型計算機装置としては、ブレードサーバが良く知られている。このブレードサーバは、CPUモジュールもしくはCPUブレードと呼ばれる薄いケース内にCPU、メモリ、HDD等の計算機装置の部品が搭載され、このCPUモジュールをシャーシもしくはエンクロージャと呼ばれるケース内に複数枚搭載することにより高密度な複合型計算機装置を実現し、従来の計算機装置に比べ省スペース性が可能となっている。また、ブレードサーバにはシャーシ内の計算機装置部品に対して電力を供給するためのAC-DC電源モジュールが搭載されていて、1個のAC-DC電源モジュールの障害時でも複合型計算機装置が稼動し続けることができるように、AC-DC電源モジュールを複数個搭載するのが一般的である。このような複数の計算機装置の統合管理を実現のために、ブレードサーバにはシャーシ内に搭載されたCPUモジュールやAC-DC電源モジュール等の計算機装置を構成するモジュールの制御および 30
管理を行う管理モジュールを搭載しているのが一般的である。管理モジュールでは、例えばCPUモジュールやAC-DC電源モジュール等のモジュールの電源ON/OFFの制御や、CPUモジュールやシャーシ内の温度のモニター、CPUモジュールやAC-DC電源モジュール等のモジュールの障害監視および障害処理を行う。ブレードサーバの管理者は、この管理モジュールに例えばtelnet等の手段によりアクセスし、ブレードサーバ内の計算機装置の電源のON/OFF等を管理することが可能となっている。さらに、計算機装置上で稼動するオペレーションシステムやアプリケーション等を含めたシステム管理を行うためには、ブレードサーバの外部の計算機装置上いわゆる管理サーバで稼動する管理ソフトウェアを用いる。この管理サーバは、上で述べた管理モジュールと管理ネットワーク、例えばローカルエリアネットワーク（LAN）で接続され、管理モジュールと通信を行うことによりブレードサーバ内 40
の計算機装置の電源のON/OFF等を管理することが可能となっている。また、この管理サーバはブレードサーバ内の個々の計算機装置と管理ネットワークで接続され、計算機装置上で稼動するオペレーションシステムやアプリケーション等の管理を行う。管理ソフトウェアと計算機装置の通信には、計算機装置上で動作するエージェントソフトウェアを用いることが一般的である。

【0003】

なお、この種のブレードサーバについては特許文献1（「カートリッジ型サーバユニットおよび該サーバユニット搭載用筐体」）に記載されている。

【0004】

【特許文献1】特開2002-32153号公報

50

【特許文献2】米国特許US 6、785、827号明細書

【発明の開示】

【発明が解決しようとする課題】

【0005】

ところが、上記のような複合型計算機装置におけるAC-DC電源モジュールの管理では、次のような問題点があることが本発明者により見い出された。

【0006】

すなわち、AC-DC電源モジュールに必要な出力電力の最大値は、シャーシ内に搭載される計算機装置の構成要素である全モジュールが稼動時に必要な最大の消費電力値で決められているため、シャーシ内の一部の計算機装置が電源OFFの状態である場合やスリープ状態にある場合や、計算機装置に搭載されたプロセッサがスリープ状態にある場合には、消費電力が少なくなるためシャーシ内の計算機装置の稼動のために過剰な出力電力を持つAC-DC電源モジュールにより電力が供給されることになる。この場合にはAC-DC電源モジュール一台あたりの出力電力がAC-DC電源モジュールの最大出力電力値よりも少ない低負荷状態で稼動することになる。一般的にAC-DC電源は、低負荷状態では効率が悪いいため、必要な消費電力に対して、無駄な消費電力を消費してしまう。図16は、典型的なAC-DC電源の負荷（横軸）に対する消費電力（縦軸）の関係を示したグラフである。図16の曲線1601は、AC-DC電源の出力に接続された負荷に対してAC-DC電源が消費する電力を示している。この図16から、負荷が0の状態でも消費電力としてW0が必要であることが分かる。このW0の値についてはAC-DC電源の性能に依存するが、計算機装置に用いられるAC-DC電源では、10ワット（W）～50ワット（W）程度ある。低負荷時LLでは、電源の効率が悪く無駄な消費電力が大きい。一方、通常負荷時SLでは、電源の効率が良く無駄な消費電力は小さい。ブレードサーバでは複数のAC-DC電源モジュールを搭載するため、 $(\text{無駄な消費電力 (W)}) \times (\text{AC-DC電源モジュールの個数}) \quad (\text{W})$ だけ無駄な消費電力が発生してしまうという問題がある。

【0007】

このような問題に対する公知例として、特許文献2（「System for determining server power supply requirement by sampling power usage values thereof at a rate based upon the criticality of its availability」）がある。この公知例では、システム内で個々の計算機装置に必要な消費電力をモニターして、その合計により必要な電力消費量を算出し、その電力消費量に応じて必要なAC-DC電源の個数を決定するという技術である。しかしながら、この公知例の方式では、急に多くの計算機装置を電源ON状態にする場合、システムがスリープ状態から通常稼動状態に復帰する場合、または、プロセッサがスリープ状態から通常稼動状態に復帰する場合において、消費電力のモニターによって最大消費電力を計算し必要な電源の個数を決定する方法では急な消費電力の増加に対応できないという問題がある。

【0008】

本発明の目的は、このような急に消費電力が変化するような場合でも、AC-DC電源モジュールの稼働台数の制御を行い、無駄な消費電力を削減することが可能となる技術を提供することにある。

【0009】

本発明の前記ならびにその他の目的と新規な特徴は、本明細書の記述および添付図面から明らかになるであろう。

【課題を解決するための手段】

【0010】

上記目的は、

（1）プロセッサと、前記プロセッサに電源を供給するDC-DC電源とを有する少なくとも一つのCPUモジュールと、前記CPUモジュールごとに接続され前記DC-DC電源に前記プロセッサの稼動状態に合わせて必要な電源量を供給するAC-DC電源を有するAC-DC電源モジュールと、前記プロセッサの非稼動時においても電源供給制御を行い、前記AC-DC電源モジュール

ールと電源制御用インターフェースで接続され前記AC-DC電源モジュールの出力制御を行う前記CPUモジュールに搭載されたCPUモジュールコントローラと、管理用インターフェースを介して前記CPUモジュールコントローラと接続され前記CPUモジュールの電源制御を行う管理モジュールと、を備え、前記管理モジュールは、前記プロセッサのステータス情報と、前記CPUモジュール稼動情報と、前記プロセッサ種毎の消費電力情報とを含むシステム情報を保持し、前記システム情報に基づいて、前記CPUモジュールごとに必要とする消費電力を計算し、その計算された消費電力を確保するのに必要なAC-DC電源モジュールの稼動数を計算し、その計算結果に応じて前記AC-DC電源モジュールの稼動台数を決定することを特徴とする複合型計算機装置により、

(2) また、各装置に電力を供給する少なくとも一つのAC-DC電源を有するAC-DC電源モジュールと、少なくとも一つのプロセッサを搭載し、前記AC-DC電源モジュールから供給される電力をDC-DC変換するDC-DC電源を搭載し、前記プロセッサの非稼動時にも電源供給制御を行い、前記AC-DC電源モジュールと電源制御用インターフェースで接続されAC-DC電源モジュールの出力制御を行うCPUモジュールコントローラとを搭載したCPUモジュールと、管理用のインターフェースに接続され前記CPUモジュールの電源制御を行う管理モジュールと、前記CPUモジュールの電源ON時にはCPUモジュール毎に稼動するオペレーションシステムおよびシステムファームウェアとを有し、前記プロセッサがスリープ状態を有する場合に、前記プロセッサのスリープ状態として少なくとも一つの状態、または前記CPUモジュール毎にシステムのスリープ状態として少なくとも一つの状態、のいずれかのスリープ状態を持ち、前記CPUモジュールコントローラは、前記プロセッサの状態が通常稼動状態からスリープ状態に遷移する場合には、前記システムファームウェアから通知されるスリープ状態への遷移命令により、一定時間スリープするためのタイマ機能を持ち、また前記管理モジュールへプロセッサ状態変更を通知する機能を持ち、前記プロセッサの状態がスリープ状態から通常稼動状態に遷移する場合には、前記タイマ機能復帰を契機に前記管理モジュールへのプロセッサ状態変更を通知する機能を持ち、前記管理モジュールからの前記AC-DC電源モジュールの設定完了通知を受けてスリープから通常稼動状態に戻る割り込みをプロセッサに投入する機能を持ち、前記管理モジュールは、システム情報として、CPUモジュールの電源ONまたはOFFと、前記CPUモジュールのシステムの通常稼動状態およびスリープ状態と、搭載されたプロセッサの種類の情報とを含むCPUモジュール稼動情報を持ち、プロセッサ情報として、前記CPUモジュール内のプロセッサの通常稼動状態およびスリープ状態のプロセッサの状態情報と、前記プロセッサの種類毎に、それぞれの状態に必要なプロセッサ消費電力情報を持ち、前記プロセッサの状態が通常稼動状態からスリープ状態に遷移する場合または前記プロセッサの状態がスリープ状態から通常稼動状態に遷移する場合には、前記CPUモジュールコントローラからのプロセッサ状態変更の通知を受けて、前記CPUモジュール稼動情報と、前記プロセッサの状態情報と、前記プロセッサ消費電力情報とから複合型計算機装置全体に必要な消費電力を計算し、その消費電力に基づいてAC-DC電源モジュールの稼動数を計算し、AC-DC電源モジュール稼動数の変更が必要な場合はAC-DC電源モジュールの稼動数を変更する制御を行う機能を持ち、前記CPUモジュールコントローラへ前記AC-DC電源モジュールが安定して稼動していることを示す設定完了を通知する機能を持つことを特徴とする複合型計算機装置により、

(3) あるいは、装置全体に電力を供給する少なくとも一つのAC-DC電源モジュールと、少なくとも一つのプロセッサを搭載し、前記AC-DC電源モジュールから供給される電力をDC-DC変換するDC-DC電源を搭載し、前記プロセッサの非稼動時にも電源供給制御を行い、前記AC-DC電源モジュールと電源制御用インターフェースで接続されAC-DC電源モジュールの出力制御を行うCPUモジュールコントローラを搭載したCPUモジュールと、管理用のインターフェースで接続され前記CPUモジュールの電源制御を行う管理モジュールと、前記CPUモジュールの電源ON時にはCPUモジュール毎に稼動するオペレーションシステムおよびシステムファームウェアとを有し、前記プロセッサがスリープ状態を有する場合に、前記プロセッサのスリープ状態として少なくとも一つの状態、または前記CPUモジュール毎にシステムのスリープ状態として少なくとも一つの状態のいずれか一方のスリープ状態を持

10

20

30

40

50

ち、前記プロセッサの状態が通常稼動状態からスリープ状態に遷移する場合に、前記CPUモジュール上で稼動する前記オペレーションシステムは、前記システムファームウェアにプロセッサの状態を通常稼動状態からスリープ状態へ遷移させる命令を通知し、この通知によりCPUモジュール上で稼動する前記システムファームウェアは、前記CPUモジュールコントローラにプロセッサ状態変更を通知し、また前記プロセッサに対してスリープ制御を行い、前記CPUモジュールコントローラは前記プロセッサ管理モジュールにプロセッサ状態変更を通知し、この通知により前記プロセッサ管理モジュールは前記プロセッサCPUモジュールのプロセッサの通常稼動状態およびスリープ状態のプロセッサの状態情報を更新し、このプロセッサの状態情報と、前記プロセッサCPUモジュールの電源ONおよびOFF状態と、前記プロセッサCPUモジュールのシステムの通常稼動状態およびスリープ状態および搭載されたプロセッサの種類の情報を持つCPUモジュール稼動情報と、各プロセッサ種類に対して各状態に必要なとなるプロセッサ消費電力情報と、を用いて複合型計算機装置全体に必要な消費電力を計算し、前記消費電力に基づき前記AC-DC電源モジュールの稼動数を計算し、前記AC-DC電源モジュール稼動数の変更が必要な場合はAC-DC電源モジュールの稼動数を変更する制御を行うことを特徴とする複合型計算機装置の管理方法により達成できる。

10

【発明の効果】

【0011】

本願において開示される発明のうち、代表的なものによって得られる効果を簡単に説明すれば以下の通りである。

20

【0012】

複合型計算機装置において計算機装置の稼動状況が変わっても、プロセッサの状態やシステムの状態に合わせてAC-DC電源モジュールの稼働台数を制御することができるので、電力の効率が悪化しないで無駄な消費電量を削減することが可能となる。

【発明を実施するための最良の形態】

【0013】

以下、本発明の実施の形態を図面に基づいて詳細に説明する。なお、実施の形態を説明するために全図において、同一の部材には原則として同一の符号を付し、その繰り返しの説明は原則として省略する。

【0014】

30

先ず本発明の一実施の形態による複合型計算機装置の構成について説明する。図1は、本発明の一実施の形態による複合型計算機装置のブロック図、図2は図1の複合型計算機装置に設けられた管理モジュールが有するCPUモジュール稼動情報の一例を示す説明図、図3は図1の複合型計算機装置に設けられた管理モジュールが有するプロセッサの状態情報の一例を示す説明図、図4は図1の複合型計算機装置に設けられた管理モジュールが有するプロセッサ消費電力情報の一例を示す説明図、である。図1に示すように、本実施の形態に置いて複合型計算機装置100は、1以上のCPUモジュール1011、1012、1013、...、1以上のAC-DC電源モジュール1061、1062、1063、1064、1065、...、ならびに管理モジュール107から構成されている。CPUモジュール1011(、1012、1013、...)は、1枚もしくは相互に接続された複数枚の配線基板上にプロセッサ1021、1031、... (、1022、1032、...、1023、1033、...)、DC-DC電源1041(、1042、1043、...)、CPUモジュールコントローラ1051(、1052、1053、...)から構成されている。CPUモジュールコントローラ1051(、1052、1053、...)は、例えばノースブリッジ等のチップセットやベースボードマネジメントコントローラ等のマネジメント用チップセット等で構成される。次にCPUモジュール1011を例にとりCPUモジュール内の接続関係を示す。他のCPUモジュール1012、1013、...についても同様な接続関係を持つ。CPUモジュールコントローラ1051がDC-DC電源を制御するために制御線PDで接続される。CPUモジュールコントローラ1051がプロセッサの割り込みを発行等のプロセッサ1021、1031、...の制御を行うための制御線BCで接続される。また、DC-DC電源1041は、電源供給のためプロセッサ1021、1031、...と電源供給配線PCで接続される。また、CPUモジュール1011で稼動するソフトウェア1091は、オペレーションシステム1101、

40

50

システムファームウェア1111から構成される。オペレーションシステム1101とシステムファームウェア1111とはソフトウェアインタフェースBAで接続され、システムファームウェア1111はCPUモジュールコントローラ1051とソフトウェアインタフェースBBにより接続される。このソフトウェア1091の構成についても、他のCPUモジュール1012、1013、...についても同様な構成および接続関係を持つ。次に複合型計算機装置100内の各モジュール間の接続関係について説明する。管理モジュール107は、CPUモジュールコントローラ1051、1052、1053、...、と管理用のインターフェースMIで接続される。管理用のインターフェースMIには、例えばイーサネット（登録商標）やI2C（Inter-Integrated Circuit）等が用いられる。この管理用インターフェースMIを使用することにより管理モジュール107は、CPUモジュール1011、1012、1013、...、電源ON/OFF、障害監視、プレゼンス確認等の管理を行う。また、管理モジュール107は、AC-DC電源モジュール1061、1062、1063、1064、1065、...、とAC-DC電源モジュール制御用のインターフェースPAで接続される。AC-DC電源モジュール制御用のインターフェースPAには、例えばI2C（Inter-Integrated Circuit）等が用いられる。このAC-DC電源モジュール制御用のインターフェースPAを使用することにより管理モジュール107はAC-DC電源モジュール1061、1062、1063、1064、1065、...、の出力制御や障害監視等の管理を行う。AC-DC電源モジュール1061、1062、1063、1064、1065、...、は複合型計算機装置外部の電源、例えばブレーカ電源からAC電力を供給され、電源供給配線PBによりDC電力をCPUモジュール1011、1012、1013、...、に供給する。管理モジュール107は、複合型計算機装置100を管理するためにシステム情報108を保持している。このシステム情報108の詳細については図2、図3、図4を用いて詳細に説明する。

【0015】

システム情報108の情報は、少なくとも図2に示すCPUモジュール稼動情報FT2、図3に示すプロセッサの状態情報FT3、図4に示すプロセッサ消費電力情報FT4から構成される。CPUモジュール稼動情報FT2は、表形式で表すとCPUモジュールの番号K201、システムの状態情報K202、CPUモジュールの各スロットに搭載されているプロセッサの種類情報K2031、K2032、...、の列で構成される。CPUモジュールの番号K201の各行は、例えば図1のCPUモジュール1011、1012、1013、...に対応する。システムの状態K202は、システムの通常稼動状態、スリープ状態（システム電源OFF状態も含む）を示している。本実施例では、B0をシステム通常稼動状態、Bx（x=1、2、3、...）をシステムスリープ状態としている。このシステムの状態を定義している例としては、ACPI（Advanced Configuration and Power Interface）でのシステムスリープ状態S0、S1、...、S5がある。

【0016】

プロセッサの種類情報K2031、K2032、...、は、搭載されているプロセッサの種類を示している。複合型計算機装置に搭載可能なプロセッサについて固有の識別子、例えばCP1、CP2、...を付ける。本実施例では、例えばCP1はXeon 2.8GHz、CP2はItanium2 900MHzを示している。図3に示すプロセッサの状態情報FT3は、表形式で表すとCPUモジュールの番号K301、CPUモジュールの各スロットに搭載されているプロセッサの状態情報K3021、K3022、...、の列で構成される。CPUモジュールの番号K301の各行は、例えば図1のCPUモジュール1011、1012、1013、...に対応する。プロセッサの状態情報K3021、K3022、...は、CPUモジュールに搭載されている各プロセッサの通常稼動状態、スリープ状態（システムスリープ状態も含む）を示している。本実施例では、A0をプロセッサ通常稼動状態、Ax（x=1、2、3、...）をプロセッサスリープ状態、SSをシステムスリープ状態（システム電源OFF状態も含む）としている。このプロセッサの状態を定義している例としては、ACPI（Advanced Configuration and Power Interface Specification Revision 3.0（2004.9.2 発行）参照）でのプロセッサスリープ状態P0、P1、P2、P3がある。図4に示すプロセッサ消費電力情報FT4は、表形式で表すとプロセッサ種類識別子K401、プロセッサの状態A0、A1、A2、A3、...の消費電力情報K4021、K4022、K4023、K4024、...、システムの状態B0、B1、B2、B3、...の消費電力情報K4031、K4032、K4033、K4034、...、で構成される。また、プロセッサ名K404の情報を含んでも良い。プロセッサ種類識別子K401は、図3での複合型計算機装置に搭載可能なプロセッサについて固有の識別子CP1、CP2、...に対応している。消費電力情報

K4021、K4022、K4023、K4024、...は、プロセッサ種類毎のプロセッサ状態A0、A1、A2、A3、...における必要なプロセッサの消費電力を示す。一方、消費電力情報K4031、K4032、K4033、K4034、...は、プロセッサ種類毎のシステム状態B0、B1、B2、B3、...における必要なプロセッサの消費電力を示す。次にこれらのシステム情報FT2、FT3、FT4を用いて、複合型計算機装置100に必要な消費電力を算出する例を示す。CPUモジュール#0の場合は、プロセッサ#0のプロセッサの種類がCP1でプロセッサの状態がA0状態、プロセッサ#1のプロセッサの種類がCP1でプロセッサの状態がA3状態であるので、プロセッサ消費電力情報FT4を読み取って、CPUモジュール#0： $80(W) + 0(W) = 80(W)$ となる。システム状態がB0のときはシステムが通常稼動状態になるので、B0状態に対応する消費電力情報K4031から消費電力を読み取るのではなく、プロセッサ状態A0、A1、A2、A3、...に対応する消費電力情報K4021、K4022、K4023、K4024、...から読み取る。

10

同様に他のCPUモジュールについても計算すると、

CPUモジュール#1： $100(W) + 100(W) = 200(W)$

CPUモジュール#2： $80(W) + 80(W) = 160(W)$

CPUモジュール#3： $80(W) + 80(W) = 160(W)$

CPUモジュール#4： $0(W) + 0(W) = 0(W)$

必要な消費電力は合計で $600(W)$ となる。当然とことながら計算機装置では、プロセッサ以外のメモリ、HDD、I/Oカード等の構成要素でも電力を消費するため、プロセッサ消費電力情報FT4にプロセッサ以外の構成要素の消費電力が加味されていない場合は、上記計算した必要な消費電力にプロセッサ以外の構成要素の消費電力を付加する必要がある。管理モジュール107は、必要な消費電力をAC-DC電源モジュール1個あたりの最大出力電力で除することで必要なAC-DC電源モジュールの個数を計算し、稼動させるAC-DC電源モジュールの個数を制御する。なお、AC-DC電源モジュールを $N+1$ 冗長で稼動させる場合は、上記で計算した必要なAC-DC電源モジュールの個数に対して1を足した個数を稼動させるAC-DC電源モジュールの個数とする。ここでの $N+1$ 冗長とは、AC-DC電源モジュールを必要台数よりも1台多く稼動させておくことにより、1個のAC-DC電源モジュールの障害が発生しても残りのAC-DC電源モジュールが稼動しているので計算機装置が稼動し続けることが可能となる機能のことを指す。

20

【0017】

次に上記で説明した本発明の一実施の形態による複合型計算機装置でのAC-DC電源モジュールの管理方法についてフローチャートを用いて説明する。最初に、プロセッサの状態がスリープ状態に遷移する場合のAC-DC電源モジュールの管理方法について説明する。図5は図1の複合型計算機装置のプロセッサの状態遷移時のAC-DC電源モジュールの管理方法のフローチャート、図6は図1の複合型計算機装置のプロセッサの状態が通常稼動状態からスリープ状態へ遷移する場合のAC-DC電源モジュールの管理方法の一例を示す説明図、図7は図1の複合型計算機装置のプロセッサの状態がスリープ状態から通常稼動状態へ遷移する場合のAC-DC電源モジュールの管理方法の一例を示す説明図である。図5では、例えば図6、図7のCPUモジュール1011に搭載されたプロセッサ1021のプロセッサ状態がA0の通常稼動状態から Ax ($x=1, 2, \dots$)のスリープ状態を経てA0の通常稼動状態に戻る場合の図6、図7のAC-DC電源モジュール1061、1062、1063、1064、1065、...の管理方法のフローチャートを示している。プロセッサ状態がA0の通常稼動状態から Ax ($x=1, 2, \dots$)のスリープ状態についてのフローをstart1からend1に示し、その説明図を図6に示す。また、プロセッサ状態が Ax ($x=1, 2, \dots$)のスリープ状態からA0の通常稼動状態についてのフローをstart2からend2に示し、その説明図を図7に示す。なお、図6、図7中に記載されているS501~S516は、図5に記載のフローチャートのステップS501~S516に対応している。図5のCPUモジュール501は図6、図7のCPUモジュール1011の動作フローを示し、図5の管理モジュール502は図6、図7の管理モジュール107の動作フローを示す。プロセッサ状態503は、動作フローに対するプロセッサ1021のプロセッサ状態を示している。AC-DC電源モジュール稼動状態504は、AC-DC電源モジュールの稼働台数の変化を示している。本実施例の説明では、例えばPM1が5台稼動、PM2が4台稼動、PM3が5台稼動を示している

30

40

50

。また、T1はプロセッサがスリープ状態から復帰するまでの時間を定めたものであり、スリープ状態Ax (x = 1、2、...) により時間が異なる。

【 0 0 1 8 】

プロセッサ状態がA0の通常稼動状態からAx (x = 1、2、...) のスリープ状態に遷移するときのフローについて説明する。オペレーションシステム1101は、プロセッサの負荷やオペレーションシステムの省電力の設定に従って遷移するプロセッサのスリープ状態Ax (x = 1、2、...) を決定し、ソフトウェアインターフェースを用いて遷移するスリープ状態Ax (x = 1、2、...) をシステムファームウェア1111に通知する (ステップS501)。システムファームウェア1111は、ソフトウェアインターフェースを用いてCPUモジュールコントローラ1051に対して遷移するスリープ状態Ax (x = 1、2、...) を通知する (ステップS502) 10。CPUモジュールコントローラ1051は遷移するスリープ状態Ax (x = 1、2、...) に対応する時間T1をCPUモジュールコントローラ1051のタイマに設定し始動する (ステップS503)。CPUモジュールコントローラ1051は、制御経路BCを用いてプロセッサ1021に対してスリープ状態Ax (x = 1、2、...) へ遷移する命令を発行する (ステップS504)。これによりプロセッサ1021は、A0の通常稼動状態からAx (x = 1、2、...) のスリープ状態に遷移する。CPUモジュールコントローラ1051は、通信経路MIを用いて管理モジュール107にプロセッサ1021がスリープ状態Ax (x = 1、2、...) に遷移したことを通知する (ステップS505)。管理モジュール107は、保持するプロセッサの状態情報FT3のスリープ状態に遷移したプロセッサ 20に対応する状態をスリープ状態Ax (x = 1、2、...) に変更する (ステップS506)。システム情報108が更新されたため管理モジュール107は、システム情報108 (FT2、FT3、FT4の情報) を基に、必要な消費電力を計算し、さらに必要なAC-DC電源モジュールの個数を計算する (ステップS507)。管理モジュール107はAC-DC電源モジュールの稼動個数に変更があるかを判定し (ステップS508)、稼動台数に変更がある場合には、管理モジュール107は、制御用の通信経路PAを用いてAC-DC電源モジュール1061、1062、1063、1064、1065、... の出力を制御して稼働台数を変更する (ステップS509)。本実施例では、例えば5台のAC-DC電源モジュール1061、1062、1063、1064、1065の稼動から、AC-DC電源モジュール1065の出力をOFFにして、4台の稼動に変更している。一方、稼動台数に変更がない場合には、動作フローを終了する。

【 0 0 1 9 】

次に、プロセッサ状態がAx (x = 1、2、...) のスリープ状態からA0の通常稼動状態に遷移するときのフローについて説明する。ステップS503で設定したタイマのタイムアウトを契機に、CPUモジュールコントローラ1051は、通信経路MIを用いて管理モジュール107にプロセッサ1021が通常稼動状態A0に遷移することを通知する (ステップS510)。管理モジュール107は、保持するプロセッサの状態情報FT3のスリープ状態に遷移したプロセッサに対応する状態を通常稼動状態A0に変更する (ステップS511)。システム情報108が更新されたため管理モジュール107は、システム情報108 (FT2、FT3、FT4の情報) を基に、必要な消費電力を計算し、さらに必要なAC-DC電源モジュールの個数を計算する (ステップS512) 30。管理モジュール107はAC-DC電源モジュールの稼動個数に変更があるかを判定し (ステップS513)、稼動台数に変更がある場合には、管理モジュール107は、制御用の通信経路PAを用いてAC-DC電源モジュール1061、1062、1063、1064、1065、... の出力を制御して稼働 40台数を変更する (ステップS514)。本実施例では、例えば4台のAC-DC電源モジュール1061、1062、1063、1064の稼動から、AC-DC電源モジュール1065の出力をONにして、5台の稼動に変更している。一方、稼動台数に変更がない場合には、ステップS515の動作フローに移る。管理モジュール107は、制御用の通信経路PAを用いてAC-DC電源モジュール1061、1062、1063、1064、1065、... の正常に出力されていることを確認し、CPUモジュールコントローラ1051にAC-DC電源モジュールの設定完了を通知する (ステップS515)。CPUモジュールコントローラ1051は、制御経路BCを用いてプロセッサ1021に対して通常稼動状態A0へ遷移するため割り込み命令を発行する (ステップS516)。これによりプロセッサ1021は、Ax (x = 1、2、...) のスリープ状態からA0の通常稼動状態に遷移する。

【 0 0 2 0 】

10

20

30

40

50

次に、システムの状態が通常稼動状態からスリープ状態に遷移する場合のAC-DC電源モジュールの管理方法について説明する。図8は図1の複合型計算機装置のシステムの状態が通常稼動状態からスリープ状態へ遷移する場合のAC-DC電源モジュールの管理方法のフローチャート、図10は図1の複合型計算機装置のシステムの状態が通常稼動状態からスリープ状態へ遷移する場合のAC-DC電源モジュールの管理方法の一例を示す説明図である。

【0021】

図10のブロック図で説明されていないものについて以下で説明する。複合型計算機100の外部の管理サーバ10012は、複合型計算機装置100やその他の計算機システムの構成要素を管理するためのサーバであり、管理ソフトウェア10013を搭載している。管理サーバ10012は、管理用のネットワークML1で複合型計算機装置100内のCPUモジュール1011、1012、1013、...と接続される。また、管理サーバ10012は、管理用のネットワークML2で複合型計算機装置100内の管理モジュール107と接続される。管理用のネットワークML1およびML2には、例えばイーサネットが用いられる。管理ソフトウェア10013と通信するためのソフトウェアとしてCPUモジュール1011、1012、1013、...上にエージェントソフトウェア100141、100142、100143、...を搭載する場合もある。管理ソフトウェア10013は、エージェントソフトウェア100141、100142、100143、...もしくはオペレーションシステム1101、1102、1103、...と通信を行い、例えばCPUモジュール1011、1012、1013、...のオペレーションシステム1101、1102、1103、...のシャットダウン、アプリケーションソフトウェアの情報取得または設定等の管理を行う。また、管理ソフトウェア10013は管理モジュールと通信を行い、例えばCPUモジュール1011、1012、1013、...の電源のON/OFF、ハードウェア構成情報の取得等の管理を行う。

【0022】

図8では、例えば図10のCPUモジュール1011で構成される計算機装置のシステムの状態がB0の通常稼動状態からBx (x=1、2、...)のスリープ状態に遷移する場合の図10のAC-DC電源モジュール1061、1062、1063、1064、1065、...の管理方法のフローチャートを示している。また、その説明図を図10に示す。なお、図10中に記載されているS801~S810は、図8に記載のフローチャートのステップS801~S810に対応している。図8の管理サーバ801は図10の管理サーバ10012の動作フローを示し、図8のCPUモジュール802は図10のCPUモジュール1011の動作フローを示し、図8の管理モジュール803は図10の管理モジュール107の動作フローを示す。図8のシステムの状態804は、動作フローに対するCPUモジュール1011のシステムの状態を示している。AC-DC電源モジュール稼動状態805は、AC-DC電源モジュールの稼働台数の変化を示している。本実施例の説明では、例えばPM1が5台稼動、PM2が4台稼動を示している。

【0023】

CPUモジュール1011で構成される計算機装置のシステムの状態がB0の通常稼動状態からBx (x=1、2、...)のスリープ状態に遷移するときのフローについて説明する。管理サーバ10012上の管理ソフトウェア10013は、業務サービスの負荷状況等により稼動する計算機装置の台数を変更するために計算機装置の遷移するシステムのスリープ状態Bx (x=1、2、...)を決定する。管理ソフトウェア10013は、管理用のネットワークML1を用いてCPUモジュール1011で動作するエージェントソフトウェア100141に対してシステムのスリープ状態Bx (x=1、2、...)への移行を要求する(ステップS801)。エージェントソフトウェア100141は、ソフトウェアインターフェースを用いて遷移するスリープ状態Bx (x=1、2、...)をオペレーションシステム1101に要求する(ステップS802)。オペレーションシステム1101は、ソフトウェアインターフェースを用いて遷移するスリープ状態Bx (x=1、2、...)をシステムファームウェア1111に命令する(ステップS803)。システムファームウェア1111は、ソフトウェアインターフェースを用いてCPUモジュールコントローラ1051に対して遷移するスリープ状態Bx (x=1、2、...)を通知する(ステップS804)。CPUモジュールコントローラ1051は遷移するスリープ状態Bx (x=1、2、...)への移行命令を実行する(ステップS805)。これによりCPUモジュール1011で構成される計算機装置のシステムの状態

は、B0の通常稼働状態からBx (x = 1、2、...)のスリープ状態に遷移する。CPUモジュールコントローラ1051は、通信経路MIを用いて管理モジュール107にシステムがスリープ状態Bx (x = 1、2、...)に遷移したことを通知する(ステップS807)。管理モジュール107は、保持するCPUモジュール稼働情報FT2のスリープ状態に遷移したシステムに対応する状態をスリープ状態Bx (x = 1、2、...)に変更し、保持するプロセッサの状態情報FT3のスリープ状態に遷移したプロセッサに対応する状態をシステムのスリープ状態SSに変更する。(ステップS807)。システム情報108が更新されたため管理モジュール107は、システム情報108 (FT2、FT3、FT4の情報)を基に、必要な消費電力を計算し、さらに必要なAC-DC電源モジュールの個数を計算する(ステップS808)。管理モジュール107はAC-DC電源モジュールの稼働個数に変更があるかを判定し(ステップS809)、稼働台数に変更がある場合には、管理モジュール107は、制御用の通信経路PAを用いてAC-DC電源モジュール1061、1062、1063、1064、1065、...の出力を制御して稼働台数を変更する(ステップS810)。本実施例では、例えば5台のAC-DC電源モジュール1061、1062、1063、1064、1065の稼働から、AC-DC電源モジュール1065の出力をOFFにして、4台の稼働に変更している。一方、稼働台数に変更がない場合には、動作フローを終了する。

【0024】

次に、システムの状態がスリープ状態から通常稼働状態に遷移する場合のAC-DC電源モジュールの管理方法について説明する。図9は図1の複合型計算機装置のシステムの状態がスリープ状態から通常稼働状態へ遷移する場合のAC-DC電源モジュールの管理方法のフローチャート、図11は図1の複合型計算機装置のシステムの状態がスリープ状態から通常稼働状態へ遷移する場合のAC-DC電源モジュールの管理方法の一例を示す説明図である。

【0025】

図9では、例えば図11のCPUモジュール1011で構成される計算機装置のシステムの状態がBx (x = 1、2、...)のスリープ状態からB0の通常稼働状態に遷移する場合の図11のAC-DC電源モジュール1061、1062、1063、1064、1065、...の管理方法のフローチャートを示している。また、その説明図を図11に示す。なお、図11中に記載されているS901~S907は、図9に記載のフローチャートのステップS901~S907に対応している。図9の管理サーバ901は図11の管理サーバ10012の動作フローを示し、図9のCPUモジュール902は図11のCPUモジュール1011の動作フローを示し、図9の管理モジュール903は図11の管理モジュール107の動作フローを示す。図9のシステムの状態904は、動作フローに対するCPUモジュール1011のシステムの状態を示している。AC-DC電源モジュール稼働状態905は、AC-DC電源モジュールの稼働台数の変化を示している。本実施例の説明では、例えばPM2が4台稼働、PM3が5台稼働を示している。

【0026】

CPUモジュール1011で構成される計算機装置のシステムの状態がBx (x = 1、2、...)のスリープ状態からB0の通常稼働状態に遷移するときのフローについて説明する。管理サーバ10012上の管理ソフトウェア10013は、業務サービスの負荷状況等により稼働する計算機装置の台数を変更するために計算機装置のスリープ状態から通常稼働状態への遷移を決定する。管理ソフトウェア10013は、管理用のネットワークML2を用いて管理モジュール107にCPUモジュール1011のシステムの通常稼働状態B0への移行、すなわち起動を要求する(ステップS901)。管理モジュール107は、保持するCPUモジュール稼働情報FT2の通常稼働状態に遷移するシステムに対応する状態を通常稼働状態B0に変更し、保持するプロセッサの状態情報FT3のシステムの通常稼働状態B0に遷移したプロセッサに対応するプロセッサの通常稼働状態A0に更する。(ステップS902)。システム情報108が更新されたため管理モジュール107は、システム情報108 (FT2、FT3、FT4の情報)を基に、必要な消費電力を計算し、さらに必要なAC-DC電源モジュールの個数を計算する(ステップS903)。管理モジュール107はAC-DC電源モジュールの稼働個数に変更があるかを判定し(ステップS904)、稼働台数に変更がある場合には、管理モジュール107は、制御用の通信経路PAを用いてAC-DC電源モジュール1061、1062、1063、1064、1065、...の出力を制御して稼働台数を変更する

(ステップS905)。本実施例では、例えば4台のAC-DC電源モジュール1061、1062、1063、1064の稼働から、AC-DC電源モジュール1065の出力をONにして、5台の稼働に変更している。一方、稼働台数に変更がない場合には、ステップS906の動作フローに移る。管理モジュール107は、制御用の通信経路PAを用いてAC-DC電源モジュール1061、1062、1063、1064、1065、...の正常に出力されていることを確認し、CPUモジュールコントローラ1051にAC-DC電源モジュールの設定完了を通知する(ステップS906)。CPUモジュールコントローラ1051は、CPUモジュール1011で構成される計算機装置のシステム起動命令を実行する(ステップS907)。これによりCPUモジュール1011で構成される計算機装置は、システムのBx(x=1、2、...)のスリープ状態からB0の通常稼働状態に遷移する。

【0027】

次に図1に示す本発明の一実施の形態による複合型計算機装置において、CPUモジュール1011、1012、1013、...に搭載されるプロセッサがマルチコアを持つ場合の複合型計算機装置の構成とAC-DC電源モジュールの管理方法について述べる。図12は図1の複合型計算機装置に搭載されるCPUモジュールのプロセッサがマルチコアを有する場合のCPUモジュールのブロック図、図13は図1の複合型計算機装置に搭載されるCPUモジュールのプロセッサがマルチコアを有する場合の管理モジュールが有するプロセッサの状態情報の一例を示す説明図である。図12のCPUモジュール1011は、例えば図1、6、7、10、11に記載されたCPUモジュール1011に置き換えて考えることができる。図12では、プロセッサ1021、1031、...が内部に2つ以上のプロセッサコア121511、121512、121513、121514、...、121521、121522、121523、121524、...、を持つ構成となる。また、図13に示すように、管理モジュール107が保持する図3のプロセッサの稼働情報FT3が上記のプロセッサコア121511、121512、121513、121514、...、121521、121522、121523、121524、...、の個々のプロセッサコア状態を持つプロセッサの稼働情報FT13ように変更される。すなわち図13に示すプロセッサの状態情報FT13は、表形式で表すとCPUモジュールの番号K1301、CPUモジュールの各スロットに搭載されているプロセッサの状態情報K13021、K13022、...、の列で構成され、プロセッサの状態情報は、プロセッサコアの状態情報K130311、K130312、K130313、K130314、...、K130321、K130322、K130323、K130324、...、の列で構成される。CPUモジュールの番号K301の各行は、例えば図1のCPUモジュール1011、1012、1013、...に対応する。プロセッサコアの状態情報K130311、K130312、K130313、K130314、...、K130321、K130322、K130323、K130324、...は、CPUモジュールに搭載されているプロセッサの各プロセッサコアの通常稼働状態、スリープ状態(システムスリープ状態も含む)を示している。プロセッサの状態情報FT3と同様に本実施例では、A0をプロセッサ通常稼働状態、Ax(x=1、2、3、...)をプロセッサスリープ状態、SSをシステムスリープ状態(システム電源OFF状態も含む)としている。CPUモジュール1011、1012、1013、...に搭載されているプロセッサがマルチコア持つ場合は、上記プロセッサの状態情報FT13を用いて各マルチコアの状態毎に必要な消費電力を計算する。計算方法については、上記プロセッサの状態情報FT3の説明で示した方法と同じ方法を用いる。また、AC-DC電源モジュールの管理方法については、消費電力の計算方法が各マルチコアの状態毎に必要な消費電力を計算する点と、プロセッサ単位のプロセッサの状態遷移がプロセッサコア単位のプロセッサコアの状態遷移に変更される点と、を変更すれば良い。

【0028】

次に、プロセッサのスリープ状態Ax(x=1、2、3、...)に対するスリープ時間をカウントするタイマのタイマ設定時間T1の決定方法について述べる。図14は図1の複合型計算機装置のプロセッサのスリープ状態のためのタイマ設定時間が典型的なスリープ時間をベースに決定する方法を示す説明図、図15は図1の複合型計算機装置のプロセッサのスリープ状態のためのタイマ設定時間が電源の制御時間を考慮した典型的なスリープ時間をベースに決定する方法を示す説明図である。図14および図15で、時刻順序1401は時間の進行方向を示し図中では左から右に時間が進行する。プロセッサの状態1402は、例えばプロセッサ1021のプロセッサの状態を示し、図14、図15に示す例では通常稼働状態A0からスリープ状態Ax(x=1、2、3、...)に遷移し、さらに通常稼働状態A0に戻る場合を記載

10

20

30

40

50

している。プロセッサのスリープ時間1403は、実際にプロセッサがスリープしている実スリープ時間TSPを示しており、スリープ状態Ax (x=1、2、3、...)の期間と一致する。スリープ用タイマ時間1404は、CPUモジュールコントローラ1051がタイマに設定するタイマ設定時間T1を示す。電源制御用時間1405は、管理モジュール107が、プロセッサの状態変更通知を受けてから、AC-DC電源モジュール1061、1062、1063、1064、1065、...の稼働台数を変更し、AC-DC電源モジュールの出力を確認するまでのAC-DC電源モジュール制御時間TCP1、TCP2を示す。具体的には、AC-DC電源モジュール制御時間TCP1は図5のステップS505からステップ509までの時間を示し、AC-DC電源モジュール制御時間TCP2は図5のステップS510からステップ515までの時間を示す。図14に示すタイマ設定時間T1の決定方法では、本発明の適用前のプロセッサのスリープ時間(典型的なスリープ時間)TSAx (x=1、2、3、...)をT1として採用する。この場合は、実スリープ時間TSPは、 $TSP = T1 + TCP2 = TSAx + TCP2$ となる。このため典型的なスリープ時間TSAx (x=1、2、3、...)よりもAC-DC電源モジュール制御時間TCP2だけ長い期間スリープすることになる。一方、図15に示すタイマ設定時間T1の決定方法では、本発明の適用前のプロセッサのスリープ時間TSAx (x=1、2、3、...)からAC-DC電源モジュール制御時間TCP2を引いた時間をT1として採用する。この場合は、実スリープ時間TSPは、 $TSP = T1 + TCP2 = TSAx - TCP2 + TCP2 = TSAx$ となる。このため本発明を適用するときと同じ典型的なスリープ時間TSAx (x=1、2、3、...)だけスリープする。ただし、タイマ設定時間T1は、計算機装置の固有の値として装置ベンダーが自由に決めることができるため、この決定方法が本発明の全ての場合に適用されるものではない。

10

20

【0029】

次にプロセッサ消費電力情報FT4が持つプロセッサの状態A0、A1、A2、A3、...の消費電力情報K4021、K4022、K4023、K4024、...に定める消費電力の決定方法について示す。あるプロセッサ例えばCP1に対応するプロセッサのプロセッサの状態A0、A1、A2、A3、...に対応するスリープ時間TAx (x=1、2、3、...)、消費電力値をPAx (x=0、1、2、3、...)、プロセッサ消費電力情報FT4に記載される消費電力値MPAx (x=0、1、2、3、...)とする。一般的に、スリープ時間TAx (x=1、2、3、...)および消費電力値をPAx (x=0、1、2、3、...)は次のような関係がある。

$TA1 < TA2 < TA3 < \dots$

$PA0 > PA1 > PA2 > PA3 > \dots$

30

つまり、スリープ時間TAx (x=1、2、3、...)は、xが大きくなるほど長くなり、消費電力値をPAx (x=0、1、2、3、...)は、xが大きくなるほど小さくなる。このときプロセッサ消費電力情報FT4に記載される消費電力値MPAx (x=0、1、2、3、...)の決定方法として、例えば全AC-DC電源モジュール制御時間TCP (=TCP1 + TCP2)に対して、スリープ時間TAx (x=1、2、3、...)が小さい場合(TAx < TCP)の場合は、プロセッサ消費電力情報FT4に記載される消費電力値MPAx (x=0、1、2、3、...)をMPAx = PA0と定め、スリープ時間TAx (x=1、2、3、...)が大きい場合(TAx > TCP)の場合は、プロセッサ消費電力情報FT4に記載される消費電力値MPAx (x=0、1、2、3、...)をMPAx = PAxと定める。この方法によりプロセッサ消費電力情報FT4に記載される消費電力値MPAx (x=0、1、2、3、...)を定めることにより、スリープ時間TAx (x=1、2、3、...)が全AC-DC電源モジュール制御時間TCPよりも短い場合は、AC-DC電源モジュール制御が行われないのでAC-DC電源モジュール制御のためにスリープ時間が長くなることを防ぐことができる。

40

【0030】

一方、システムの状態B0、B1、B2、B3、...の消費電力情報K4031、K4032、K4033、K4034、...に定める消費電力の決定方法についても、上記プロセッサの状態A0、A1、A2、A3、...に対する消費電力情報の定め方と同じ考え方で決定することができる。ただし、上でも述べたように各スリープ状態のスリープ時間は、計算機装置の固有の値として装置ベンダーが自由に決めることができるため、この決定方法が本発明の全ての場合に適用されるものではない。

【0031】

50

以上、本発明者によってなされた発明を実施の形態に基づき具体的に説明したが、本発明は前記実施の形態に限定されるものではなく、その要旨を逸脱しない範囲で種々変更可能であることはいうまでもない。

【0032】

例えば、前記実施の形態においては、管理モジュール107が持つ機能が複合型計算機装置100の外部の計算機装置が持っても良い。また、AD-DC電源モジュール1061、1062、1063、1064、1065、...が複合型計算機装置100の外部にあっても良い。また、管理サーバ10012が、複合型計算機装置100内部のCPUモジュール1011、1012、1013、...が代替しても良い。

【図面の簡単な説明】

【0033】

【図1】本発明の一実施の形態による複合型計算機装置のブロック図である。

【図2】図1の複合型計算機装置に設けられた管理モジュールが有するCPUモジュール稼動情報の一例を示す説明図である。

【図3】図1の複合型計算機装置に設けられた管理モジュールが有するプロセッサの状態情報の一例を示す説明図である。

【図4】図1に示す複合型計算機装置に設けられた管理モジュールが有するプロセッサ消費電力情報の一例を示す説明図である。

【図5】図1に示す複合型計算機装置のプロセッサの状態遷移時のAC-DC電源モジュールの管理方法のフローチャートである。

【図6】図1に示す複合型計算機装置のプロセッサの状態が通常稼動状態からスリープ状態へ遷移する場合のAC-DC電源モジュールの管理方法の一例を示す説明図である。

【図7】図1に示す複合型計算機装置のプロセッサの状態がスリープ状態から通常稼動状態へ遷移する場合のAC-DC電源モジュールの管理方法の一例を示す説明図である。

【図8】図1に示す複合型計算機装置のシステムの状態が通常稼動状態からスリープ状態へ遷移する場合のAC-DC電源モジュールの管理方法のフローチャートである。

【図9】図1に示す複合型計算機装置のシステムの状態がスリープ状態から通常稼動状態へ遷移する場合のAC-DC電源モジュールの管理方法のフローチャートである。

【図10】図1に示す複合型計算機装置のシステムの状態が通常稼動状態からスリープ状態へ遷移する場合のAC-DC電源モジュールの管理方法の一例を示す説明図である。

【図11】図1に示す複合型計算機装置のシステムの状態がスリープ状態から通常稼動状態へ遷移する場合のAC-DC電源モジュールの管理方法の一例を示す説明図である。

【図12】図1に示す複合型計算機装置に搭載されるCPUモジュールのプロセッサがマルチコアを有する場合のCPUモジュールのブロック図である。

【図13】図1に示す複合型計算機装置に搭載されるCPUモジュールのプロセッサがマルチコアを有する場合の管理モジュールが有するプロセッサの状態情報の一例を示す説明図である。

【図14】図1に示す複合型計算機装置のプロセッサのスリープ状態のためのタイマ設定時間が典型的なスリープ時間をベースに決定する方法を示す説明図である。

【図15】図1に示す複合型計算機装置のプロセッサのスリープ状態のためのタイマ設定時間が電源の制御時間を考慮した典型的なスリープ時間をベースに決定する方法を示す説明図である。

【図16】典型的なAC-DC電源の負荷と消費電力の関係を示すグラフである。

【符号の説明】

【0034】

100...複合型計算機装置、
1011,1012,1013...CPUモジュール、
1021,1022,1023...プロセッサ、
1031,1032,1033...プロセッサ、
1041,1042,1043...DC-DC電源、
1051,1052,105...CPUモジュールコントローラ、

10

20

30

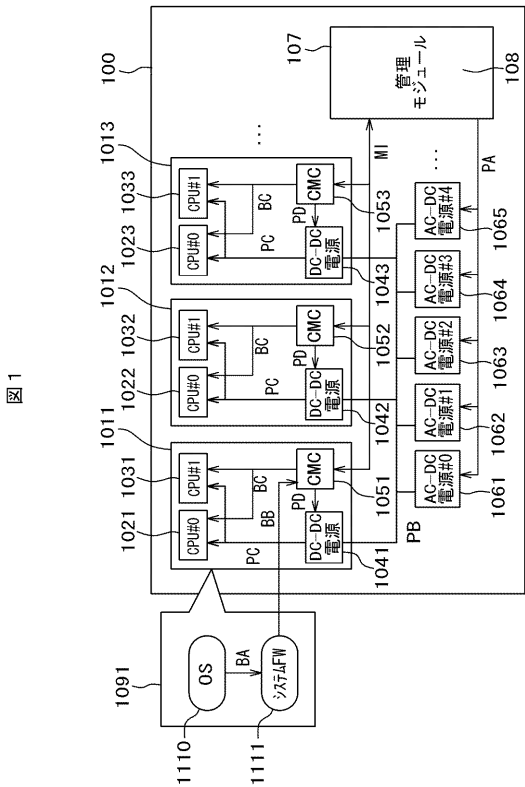
40

50

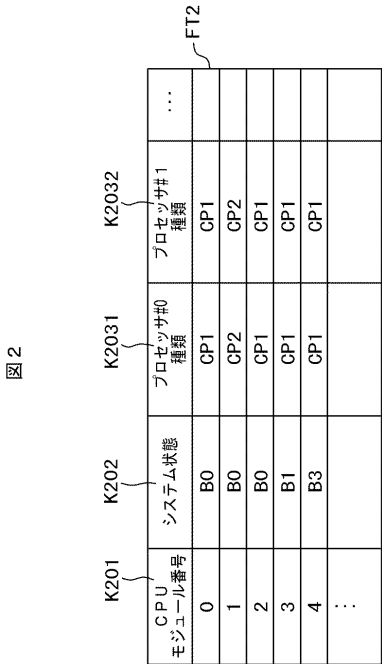
1061, 1062, 1063, 1064, 1065...AC-DC電源モジュール、	
107... 管理モジュール、	
108... システム情報、	
1091, 1092, 1093...CPUモジュールで稼動するソフトウェア群、	
1091, 1092, 1093...CPUモジュールで稼動するソフトウェア群、	
1101, 1102, 1103... オペレーションシステム、	
1111, 1112, 1113... システムファームウェア、	
10012... 管理サーバ、	
10013... 管理ソフトウェア、	
100141, 100142, 100143... エージェントソフトウェア、	10
121511, 121512, 121513, 121514... プロセッサコア、	
121521, 121522, 121523, 121524... プロセッサコア、	
1401... 時間順序、	
1402... プロセッサの状態、	
1403... プロセッサのスリープ時間、	
1404... スリープ用のタイマ時間、	
1405... 電源制御用時間、	
1601... 典型的なAC-DC電源の負荷-消費電力の関係を示す曲線、	
BA... オペレーションシステムとシステムファームウェア間のソフトウェアインターフェース、	20
BB... システムファームウェアとCPUモジュールコントローラ間のソフトウェアインターフェース、	
BC... プロセッサとCPUモジュールコントローラ間の制御用インターフェース、	
PA... 管理モジュールとAC-DC電源モジュール間の制御用のインターフェース、	
PB... AC-DC電源モジュールとDC-DC電源間の電源配線、	
PC... DC-DC電源とプロセッサ間の電源配線、	
PD... CPUモジュールコントローラとDC-DC電源間の制御用のインターフェース、	
MI ... 管理モジュールとCPUモジュールコントローラ間の管理用のインターフェース、	
ML1... 管理サーバとCPUモジュール間の管理用のインターフェース、	
ML2... 管理サーバと管理モジュール間の管理用のインターフェース、	30
FT2... CPUモジュール稼動情報、	
FT3... プロセッサの状態情報、	
FT4... プロセッサ消費電力情報、	
FT13... プロセッサがマルチコアの場合のプロセッサの情報、	
K201, K301, K1301... CPUモジュールの番号を示す列、	
K202... システムの状態を示す列、	
K2031, K2032... CPUモジュールに搭載されるプロセッサの種類を示す列、	
K3021, K3022... CPUモジュールに搭載されるプロセッサの状態を示す列、	
K401... プロセッサ種類を示す列、	
K4021, K4022, K4023, K4024... プロセッサの状態A0、A1、A2、A3、... に対応する必要な消費電力を示す列、	40
K4031, K4032, K4033, K4034... プロセッサの状態B0、B1、B2、B3、... に対応する必要な消費電力を示す列、	
K404... プロセッサの名称を示す列、	
K13021, K13022... CPUモジュールに搭載されるプロセッサを示す列、	
K130311, K130312, K130313, K130314... CPUモジュールに搭載されるプロセッサ内のプロセッサコアの状態を示す列、	
CP1, CP... プロセッサの種類を識別するための識別子、	
A0, A1, A2, A3... プロセッサの状態を表す識別子、	
B0, B1, B2, B3... システムの状態を表す識別子、	50

SS... システムの状態がスリープ状態であることを表す識別子、
TSP... 実スリープ時間、
TSAx... スリープ状態Axの（典型的な）スリープ時間、
TCP1 v ...AC-DC電源モジュール制御時間（スリープ移行時）、
TCP2...AC-DC電源モジュール制御時間（通常稼動復帰時）、
TCP... 全AC-DC電源モジュール制御時間。

【 図 1 】



【 図 2 】



【 図 3 】

図 3

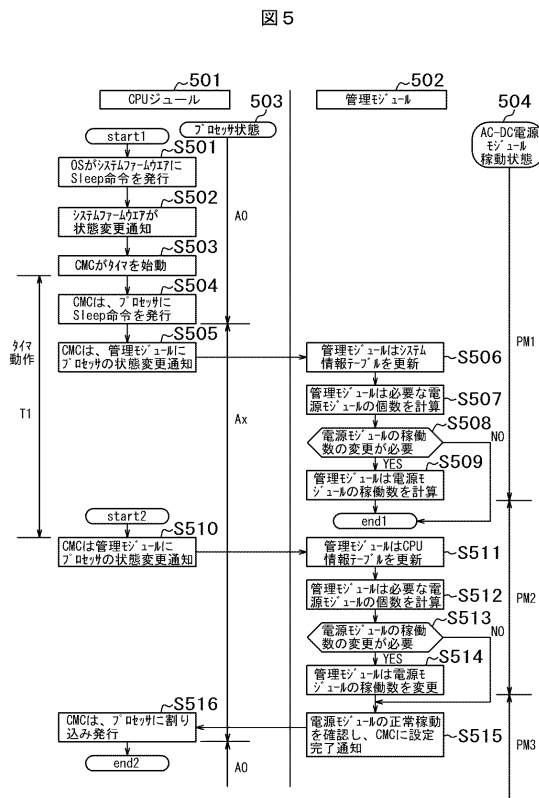
CPU モジュール番号	K301	K3031	K3032	プロセッサ#1	...	FT3
0		A0		A3		
1		A0		A1		
2		A0		A1		
3		SS		SS		
4		SS		SS		
⋮						

【 図 4 】

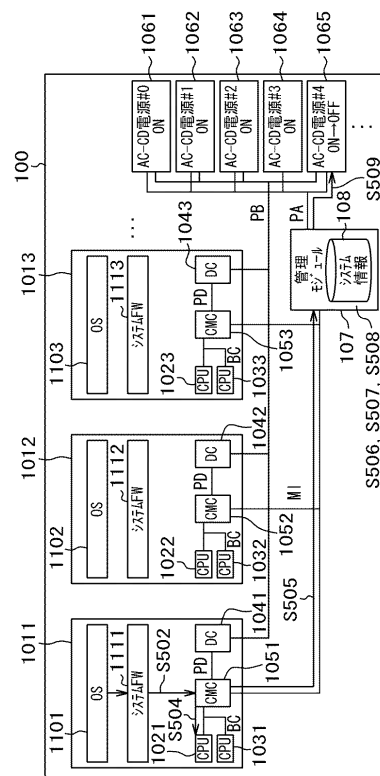
	K401	K4021	K4022	K4023	K4024	K4031	K4032	K4033	K404	
	A0	A1	A2	A3	...	B0	B1	B2	B3	...
プロセッサ種類	プロセッサ名									
CP1	80	80	0	0		80	80	0	0	Xeon 2.8GHz
CP2	100	100	0	0		100	100	0	0	Itanium2 900MHz
⋮										

単位(W)

【 図 5 】

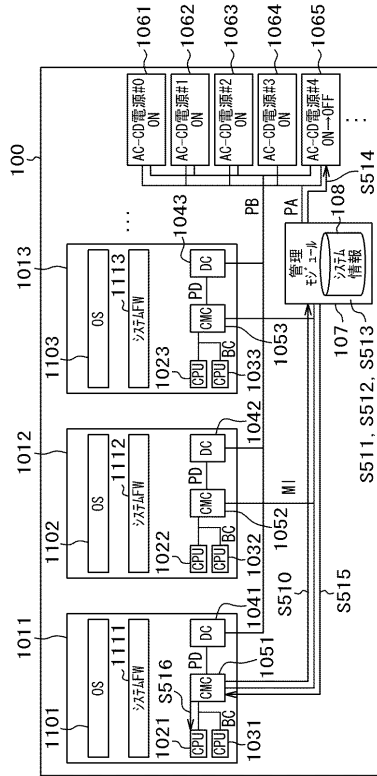


【 図 6 】



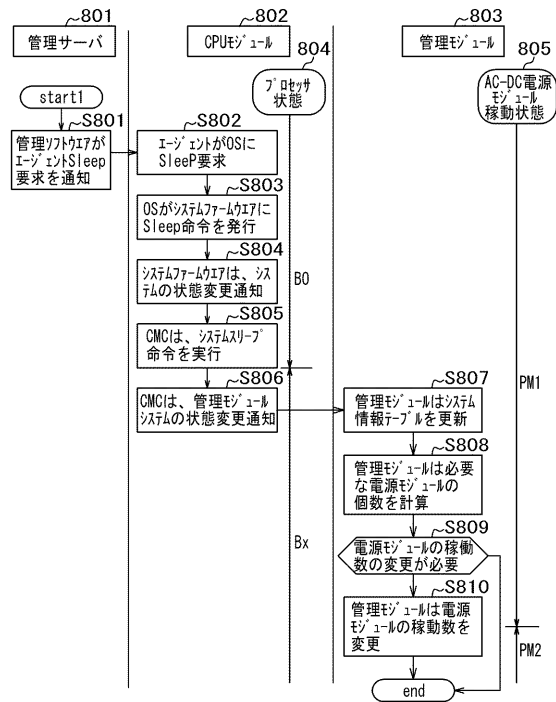
【図 7】

図 7



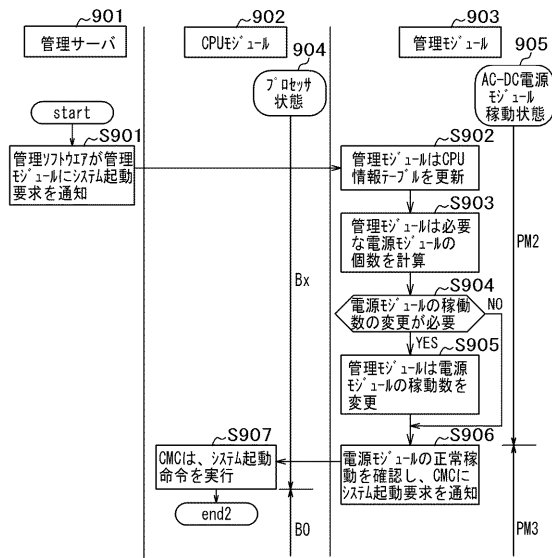
【図 8】

図 8



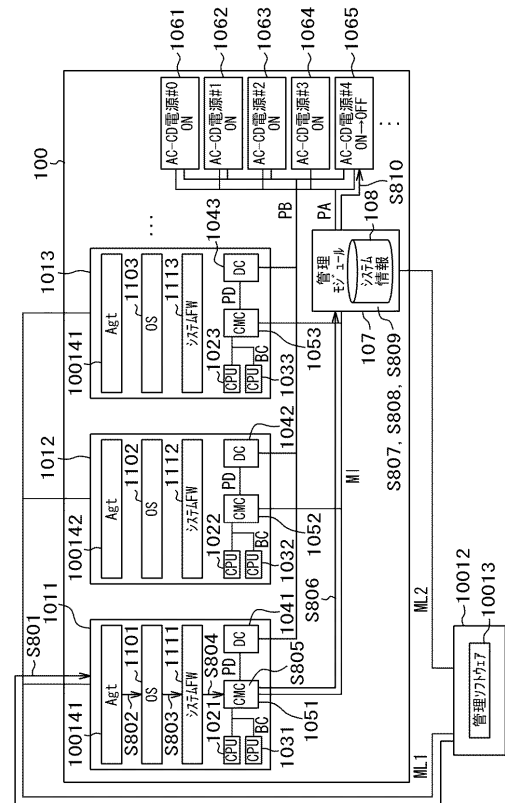
【図 9】

図 9



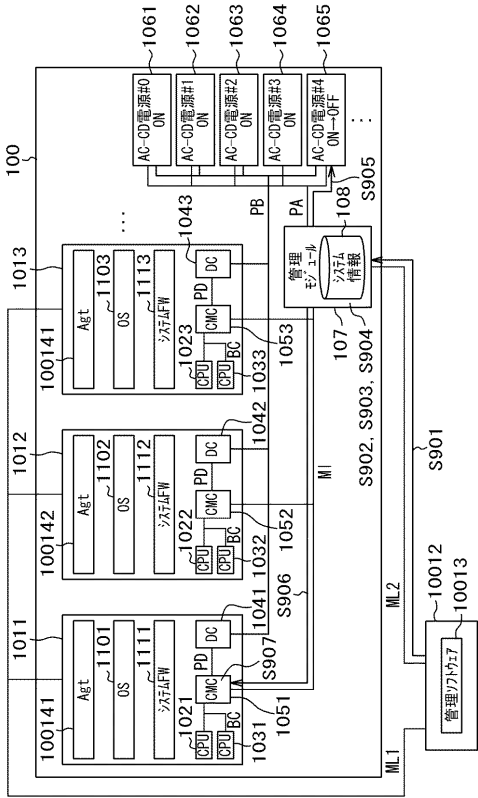
【図 10】

図 10



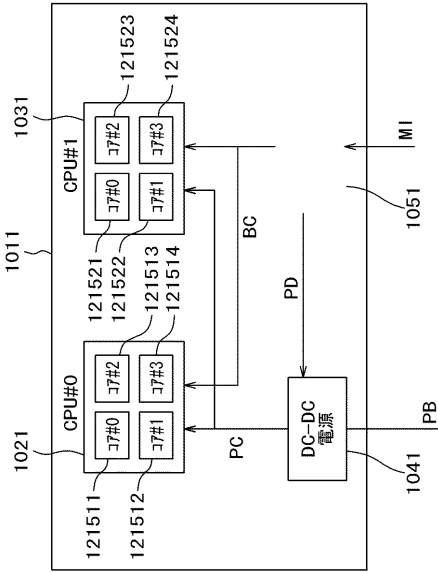
【図 1 1】

図 1 1



【図 1 2】

図 1 2



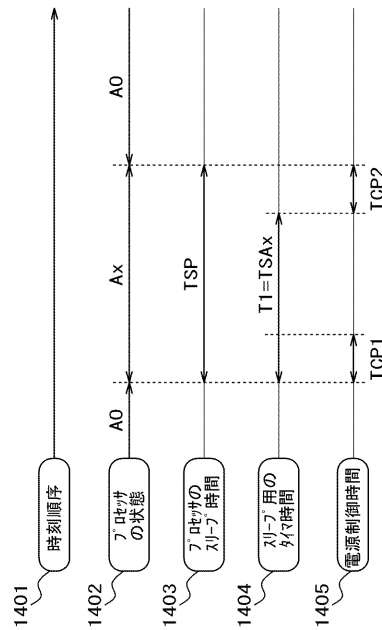
【図 1 3】

図 1 3

CPU 番号	プロセッサ#0				プロセッサ#1				...
	コア#0	コア#1	コア#2	コア#3	コア#0	コア#1	コア#2	コア#3	
0	A0	A0	A1	A2	A3	A3	A3	A0	...
1	A0	A0	A0	A0	A1	A1	A3	A3	...
2	A0	A2	A1	A0	A1	A1	A3	A3	...
3	SS	SS	SS	SS	SS	SS	SS	SS	...
4	SS	SS	SS	SS	SS	SS	SS	SS	...
...

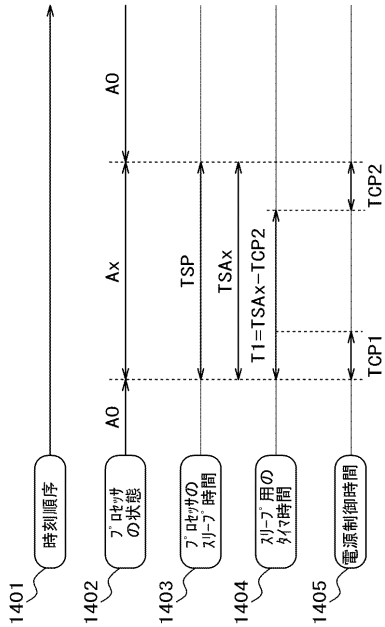
【図 1 4】

図 1 4



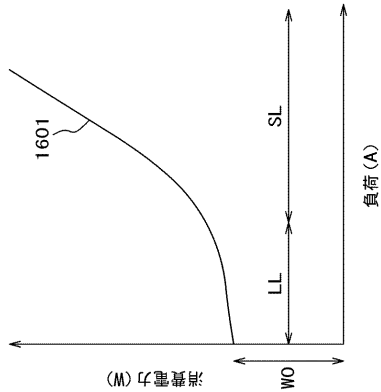
【図 1 5】

図 1 5



【図 1 6】

図 1 6



フロントページの続き

(72)発明者 堀本 和秀

東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社日立製作所 中央研究所内

審査官 三浦 みちる

(56)参考文献 特開平 0 4 - 3 3 2 0 1 3 (J P , A)
特開平 0 9 - 2 0 4 2 4 0 (J P , A)
特開 2 0 0 0 - 2 1 6 3 1 5 (J P , A)
特開 2 0 0 1 - 1 4 2 5 7 9 (J P , A)
米国特許第 0 6 7 8 5 8 2 7 (U S , B 1)
特開 2 0 0 4 - 3 0 3 2 0 6 (J P , A)
特開 2 0 0 5 - 0 8 5 1 6 4 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)

G 0 6 F 1 / 2 6

G 0 6 F 1 / 2 8

G 0 6 F 1 / 3 2