

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5505925号
(P5505925)

(45) 発行日 平成26年5月28日 (2014.5.28)

(24) 登録日 平成26年3月28日 (2014.3.28)

(51) Int. Cl.		F I			
G06F	9/50	(2006.01)	G06F	9/46	462Z
G06F	1/20	(2006.01)	G06F	1/00	360B
G06F	1/32	(2006.01)	G06F	1/00	332Z

請求項の数 3 (全 18 頁)

(21) 出願番号	特願2009-188033 (P2009-188033)	(73) 特許権者	000208891
(22) 出願日	平成21年8月14日 (2009.8.14)		KDDI株式会社
(65) 公開番号	特開2011-39889 (P2011-39889A)		東京都新宿区西新宿二丁目3番2号
(43) 公開日	平成23年2月24日 (2011.2.24)	(74) 代理人	100084870
審査請求日	平成24年2月24日 (2012.2.24)		弁理士 田中 香樹
		(74) 代理人	100092772
			弁理士 阪本 清孝
		(74) 代理人	100119688
			弁理士 田邊 壽二
		(72) 発明者	荒井 大輔
			埼玉県ふじみ野市大原二丁目1番15号
			株式会社KDDI研究所内
		(72) 発明者	吉原 貴仁
			埼玉県ふじみ野市大原二丁目1番15号
			株式会社KDDI研究所内

最終頁に続く

(54) 【発明の名称】 ネットワーク運用管理方法及びネットワーク運用管理装置

(57) 【特許請求の範囲】

【請求項1】

データセンタ内に複数の物理サーバを設置し、1台の物理サーバを論理的に複数台の論理サーバとして動作させる仮想化技術が適用された複数の物理サーバがネットワークを介して接続され、前記各物理サーバのリソースの使用状況に応じて物理サーバのリソースの論理サーバへの割り当てを動的に変更する一方、データセンタ内に設置された空調機により前記各物理サーバを冷却するネットワーク運用管理方法において、

前記物理サーバのハードウェアスペック及び論理サーバに割り当てたリソース量を管理し、

前記物理サーバ及び前記空調機に関する電力消費量を予め管理するとともに、前記データセンタ内の全体の空気の流れ、及び、全体の空気の流れに対して前記各物理サーバから排気された空気が前記空調機に循環することなく物理サーバの吸気側に流れ込む空気の割合 について、各物理サーバにおいて、異なる時間帯、稼働状況の排気面及び給気面の温度をそれぞれ検出することで、前記各物理サーバにおける異なる温度の組を2つ以上収集し作成した方程式を解くことで算出し、予め若しくは動的に管理し、

前記各物理サーバの電力消費の総量を試算するとともに、前記各物理サーバを冷却するための空調機の電力消費量を試算することで、データセンタ全体の電力消費量を試算し、

各物理サーバの使用状況に応じ、物理サーバのリソースの論理サーバへの割り当てについて、複数の組み合わせを算出し、各組における電力消費量を前記データセンタ電力消費試算部により試算し、最も電力消費が小さいものを選択し、

その出力結果を基にデータセンタ内における物理サーバのリソースの論理サーバへの割り当てを動的に変更することを特徴とするネットワーク運用管理方法。

【請求項 2】

データセンタ内に複数の物理サーバを設置し、1台の物理サーバを論理的に複数台の論理サーバとして動作させる仮想化技術が適用された複数の物理サーバが、ネットワークを介して接続され、前記各サーバのリソースの使用状況に応じて、物理サーバのリソースの論理サーバへの割り当てを動的に変更する一方、データセンタ内に設置された空調機により前記各物理サーバを冷却するネットワーク運用管理装置であって、

前記物理サーバのハードウェアスペック及び論理サーバに割り当てたリソース量を管理する物理サーバ情報管理データベースと、

前記物理サーバ及び前記空調機に関する電力消費量、前記データセンタ内の全体の空気の流れ、及び、全体の空気の流れに対して前記各物理サーバから排気された空気が前記空調機に循環することなく物理サーバの吸気側に流れ込む空気の割合 について、各物理サーバにおいて、異なる時間帯、稼働状況の排気面及び給気面の温度をそれぞれ検出することで、前記各物理サーバにおける異なる温度の組を2つ以上収集し作成した方程式を解くことで算出して管理する電力消費モデル管理データベースと、

前記物理サーバや前記空調機の情報を前記物理サーバ情報管理データベース及び前記電力消費モデル管理データベースに登録する管理インタフェース提供部と、前記物理サーバのリソースの使用状況を定期的に取り得し前記物理サーバ情報管理データベースに保存するリソース利用状況取得部と、

前記各物理サーバの電力消費の総量を試算する物理サーバ電力消費試算部と、前記各物理サーバを冷却するための空調機の電力消費量を試算する空調機電力消費試算部からなり、各部の和によりデータセンタ全体の電力消費量を試算するデータセンタ電力消費試算部と、

各物理サーバの使用状況に応じ、物理サーバのリソースの論理サーバへの割り当てについて、複数の組み合わせを算出し、各組におけるデータセンタ全体の電力消費量を前記データセンタ電力消費試算部により試算し、最も電力消費が小さいものを選択する物理サーバ配置算出部と、

前記物理サーバ配置算出部の出力結果を基にデータセンタ内における物理サーバのリソースの論理サーバへの割り当てを変更し、物理サーバの再配置（物理サーバのリソースの論理サーバへの割り当て）結果を前記物理サーバ情報管理データベースに保存する物理サーバ構成管理部と、

前記各部を制御する中央制御部とを備えたことを特徴とするネットワーク運用管理装置。

【請求項 3】

前記電力消費モデル管理データベースは、複数の前記各物理サーバに対して、当該物理サーバを含めた複数の各物理サーバの吸気側に流れ込む空気の割合を予め登録したテーブルを備えることで、前記空気の割合を管理する請求項 2 に記載のネットワーク運用管理装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、複数のサーバが接続されたネットワークの運用管理に関し、詳しくは、データセンタ内に設置された複数のサーバ機器に対して、物理的に1台のサーバ機器（物理サーバ）を論理的に複数台のサーバ機器（論理サーバ）として動作させることにより、限られた物理リソースを有効活用する仮想化技術が導入されたサーバ群が接続されるネットワークのネットワーク運用管理方法及びネットワーク運用管理装置に関する。

【背景技術】

【0002】

10

20

30

40

50

ネットワーク人口の増加や動画配信など各種通信サービスの普及により、インターネット内のトラフィック量が将来的に急激に増大することが予想される。図17は、経済産業大臣主催で開催された「グリーンITイニシアティブ会議(第1回)」で示されたトラフィック量及び消費電力の推計であり、2025年には2006年との比較でトラフィック量が190倍になると推計される。また、トラフィック量の増大に伴い、それを処理するサーバやルータなどのIT機器(ネットワーク機器)も増加する傾向にあり、その消費電力量は2025年には2006年との比較で5倍になると推計される。

【0003】

これらサーバやネットワーク機器の電力消費量を削減することは、トラフィック量の増大を支えるうえで、重要な課題である。特に、複数のサーバやネットワーク機器が集中設置されるデータセンタにおける省電力化は急務である。

10

【0004】

通常、データセンタには、複数のサーバ(サーバ群)を冷却するための空調機が設置され、データセンタ全体の省電力化を図るためには、空調機を駆動するための消費電力についても考慮する必要がある。

データセンタの電力消費量の内訳の一例として、空調機(Chiller(冷却機)+CRAC(Computer room air conditioning))に関するものが全体の42%、サーバ及びネットワーク機器(IT Equipment)に関するものが30%、電力系設備(PDU(Power Distributed Unit)、UPS(Uninterruptible PowerSupply))に関するものが23%、照明や静電気の発生を防ぐための加湿器が5%であることが知られている。

20

【0005】

サーバの仮想化技術が適用されたデータセンタを想定する。ここでは、簡単に仮想化技術について触れる。

非特許文献1及び非特許文献2に示された仮想化技術は、物理的な計算機資源を複数の論理的な資源として見せるための技術である。仮想化技術の適用により、単一の物理サーバ上で2つの(実際は2つ以上の複数の論理サーバを単一の物理サーバ上で実行することが可能)論理サーバを動作させることが可能となる。また、論理サーバは、物理サーバのハードウェアリソース(CPU、メモリ、ネットワーク、ハードディスク)を分割して利用する(分割したハードウェアリソースが論理サーバに割り当てられる)。

特許文献1、非特許文献3及び非特許文献4には、仮想化技術を応用し、仮想化されたサーバ(論理サーバ)への物理リソース(CPUやメモリ、ネットワーク帯域、ハードディスク)の割り当てを、リソースの利用状況に応じて動的に変更し、不要となった機器の電源を休止することで、更なる省電力化を図る技術が提案されている。

30

【0006】

従来技術(特許文献1、非特許文献3及び非特許文献4)の動作概要について、図15を参照しながら説明する。

従来技術においては、物理リソースの使用状況を定期監視し、リソースの使用状況が少ない場合(図15における左の例)、物理リソースの論理サーバへの割り当てを動的に再構成し、一部の物理サーバで全ての論理サーバを実行し、さらに不要となった物理サーバを休止することが行われる(図15における右の例)。

40

図15の例においては、5台の物理サーバが設置されたサーバラック内でのみ論理サーバの物理サーバへの再割り当てを行っているが、実際にはデータセンタ内に設置された任意の物理サーバに論理サーバを割り当てることが可能である。

【先行技術文献】

【特許文献】

【0007】

【特許文献1】特開2007-310791

【非特許文献】

【0008】

【非特許文献1】VMware: <http://www.vmware.com/>

50

【非特許文献2】Xen: <http://www.xen.org/>

【非特許文献3】畑崎 恵介, 高本 良史, "サーバ仮想化を用いたシステムの省電力ポリシー運用技術," 電子情報通信学会技術研究報告 コンピュータシステム研究会CPSY2006-44 Vol.106, No.436, pp. 37-42, Dec. 2006.

【非特許文献4】VMware Distributed Power Management(DPM)

【発明の概要】

【発明が解決しようとする課題】

【0009】

上述した従来技術(特許文献1、非特許文献3及び非特許文献4)においては、稼働中の物理サーバを削減することにより、サーバ群の電力消費量を削減することが行われていた。

10

すなわち、従来技術によれば、図16に示すように、サーバ1、ネットワーク機器2及び空調機5が設置されたデータセンタにおいて、サーバ群の使用状況に応じ、その都度最小のサーバ数でユーザからのサービス要求を処理するようサーバ群の構成を動的に変更することで、サーバ群の電力消費量を最小化することが行われていた。

その一方、一部のサーバに処理を集中させた場合、サーバの電力消費量は使用率に比例して増加し、また、そのサーバによる排熱は電力消費量に応じて増加するため、一部のサーバの排熱が大きくなり、それらを冷却するために(図16中、斜線で表示された矢印が冷氣)空調機5の出力を上げる必要が発生し、空調機5の電力消費量を含むデータセンタ全体の電力消費量が大きくなる現象が生じることが懸念される。

20

【0010】

すなわち、上述した従来技術(特許文献1、非特許文献3及び非特許文献4)のように、物理リソースの使用状況にのみ注目し、物理リソースの論理サーバへの割り当て、さらに不要となった物理サーバを休止し、サーバ群の電力消費量を少なくする技術だけでは、場合により、空調機(CRAC)の電力消費量の増加につながり、結果としてデータセンタDC全体の電力消費量も大きくなる可能性がある。

【0011】

本発明は上記事情に鑑みて提案されたもので、データセンタに電力消費量を削減するため、物理リソースの論理サーバへの割り当てを動的に変更する場合に、サーバの電力消費量に加え、空調機の電力消費量も考慮することにより、データセンタ全体の電力消費量を削減することができるネットワーク運用管理方法及びネットワーク運用管理装置を提供することを目的としている。

30

【課題を解決するための手段】

【0012】

上記目的を達成するため本発明は、仮想化技術が適用されたサーバと、サーバを冷却するための空調機を主な構成要素とするデータセンタにおいて、データセンタの構成(物理リソースの論理サーバへの割り当て、空調機の設定)を管理する管理サーバを新たに設置する。そして、この管理サーバが、データセンタの電力消費量のうち、サーバ群および空調機による電力消費量の合計を最小化するため、物理リソースの使用状況に応じて、空調機による電力消費量を考慮した物理リソースの論理サーバへの動的割り当て変更を行う。

40

【0013】

すなわち、請求項1の発明は、データセンタ内に複数の物理サーバを設置し、1台の物理サーバを論理的に複数台の論理サーバとして動作させる仮想化技術が適用された複数の物理サーバがネットワークを介して接続され、前記各物理サーバのリソースの使用状況に応じて、物理リソースの論理サーバへの割り当てを動的に変更する一方、データセンタ内に設置された空調機により前記各物理サーバを冷却するネットワーク運用管理方法において、

CPUの製品型式などハードウェアスペックより算出可能な物理サーバの電力消費モデルと、空調機の製品型式から算出可能な空調機の電力消費モデルと、物理サーバや空調機の物理的な位置から算出可能な、データセンタ内の空気の流れのモデルをあらかじめ管理し

50

物理リソースの論理サーバへの割り当てを変更する際、前記各物理サーバの電力消費モデルから、割り当て変更後の各物理サーバの電力消費量と、データセンタ内の全ての物理サーバによる電力消費量の総量を試算するとともに、前記物理サーバの電力消費モデルと、空調機の電力消費モデル、データセンタ内の空気の流れのモデル（データセンタ内の全体の空気の流れ、及び、全体の空気の流れに対して前記各物理サーバから排気された空気が前記空調機に循環することなく物理サーバの吸気側に流れ込む空気の割合 について、各物理サーバにおいて、異なる時間帯、稼働状況の排気面及び給気面の温度をそれぞれ検出することで、前記各物理サーバにおける異なる温度の組を2つ以上収集し作成した方程式を解くことで算出し、空気の割合を考慮したモデル）から、割り当て変更後の各物理サーバを冷却するための空調機の電力消費量を試算することで、データセンタ全体の電力消費量を試算し、

10

物理リソースの使用状況に応じ、物理リソースの論理サーバへの割り当てについて、論理サーバをデータセンタ内のいずれの物理リソースに割り当てるかの組み合わせを複数算出し、各組み合わせにおける、物理リソースの論理サーバへの割り当て変更後の、データセンタ全体の電力消費量（物理サーバによる電力消費量の総量と空調機の電力消費量の和）を、前述の手順により試算し、データセンタ全体の電力消費量が最も小さい組み合わせを選択し、

その選択結果を基に、物理リソースの論理サーバへの割り当てを動的に変更することを特徴としている。

20

【0014】

請求項2は、データセンタ内に複数の物理サーバを設置し、1台の物理サーバを論理的に複数台の論理サーバとして動作させる仮想化技術が適用された複数の物理サーバがネットワークを介して接続され、前記各物理サーバのリソースの使用状況に応じて、物理リソースの論理サーバへの割り当てを動的に変更する一方、データセンタ内に設置された空調機により前記各物理サーバを冷却するネットワーク運用管理装置であって、次の構成を含むことを特徴としている。

【0015】

前記物理サーバのCPUの製品型式などのハードウェアスペック及び論理サーバに割り当てた場合のリソース量を管理する物理サーバ情報管理データベース。

30

CPUの製品型式などハードウェアスペックより算出可能な前記物理サーバの電力消費モデルと、空調機の製品型式から算出可能な前記空調機の電力消費モデルと、物理サーバや空調機の物理的な位置から算出可能な、前記データセンタ内の全体の空気の流れのモデル、及び、全体の空気の流れに対して前記各物理サーバから排気された空気が前記空調機に循環することなく物理サーバの吸気側に流れ込む空気の割合 について、各物理サーバにおいて、異なる時間帯、稼働状況の排気面及び給気面の温度をそれぞれ検出することで、前記各物理サーバにおける異なる温度の組を2つ以上収集し作成した方程式を解くことで算出して管理する電力消費モデル管理データベース。

前記物理サーバや前記物理サーバの電力消費モデルと前記空調機の電力消費モデル、前記データセンタ内の空気の流れのモデルを前記物理サーバ情報管理データベース及び前記電力消費モデル管理データベースに登録する管理インタフェース提供部。

40

前記物理サーバのリソースの使用状況を定期的に取得し前記物理サーバ情報管理データベースに保存するリソース利用状況取得部。

前記各物理サーバの電力消費の総量を試算する物理サーバ電力消費試算部と、前記空調機の電力消費量を試算する空調機電力消費試算部からなり、各部の和によりデータセンタ全体の電力消費量を試算するデータセンタ電力消費試算部。

各物理サーバの使用状況に応じ、物理リソースの論理サーバへの割り当てについて、複数の組み合わせを算出し、各組における電力消費量を前記データセンタ電力消費試算部により試算し、最も電力消費が小さいものを選択する物理サーバ配置算出部。

前記物理サーバ配置算出部の出力結果を基にデータセンタ内における物理リソースの論

50

理サーバへの割り当てを変更し、物理サーバの再配置結果を前記物理サーバ情報管理データベースに保存する物理サーバ構成管理部。

前記各部を制御する中央制御部。

【0016】

請求項3は、請求項2のネットワーク運用管理装置において、前記電力消費モデル管理データベースは、複数の前記各物理サーバに対して、当該物理サーバを含めた複数の各物理サーバの吸気側に流れ込む空気の割合を予め登録したテーブルを備えることで、前記空気の割合を管理することを特徴としている。

【発明の効果】

【0017】

本発明によれば、物理サーバの使用状況に応じて、物理リソースの論理サーバへの割り当てを動的に変更する処理を行うに際して、物理サーバが消費する電力消費量及びそれらを冷却する空調機の電力消費量の両方を考慮して、データセンタ全体の電力消費量が最小となるように、複数の物理リソースの論理サーバへの割り当ての組合せの中から、データセンタ全体の電力消費量が最も小さくなる組み合わせを選び、物理リソースの論理サーバへの割り当てを変更するとともに、データセンタ内の全体の空気の流れ、及び、全体の空気の流れに対して前記各物理サーバから排気された空気が前記空調機に循環することなく物理サーバの吸気側に流れ込む空気の割合 について、各物理サーバにおいて、異なる時間帯、稼働状況の排気面及び給気面の温度をそれぞれ検出することで、前記各物理サーバにおける異なる温度の組を2つ以上収集し作成した方程式を解くことで算出した空気の割合 を考慮することで、物理サーバの使用状況にのみ注目し割り当てを変更する従来技術と比較し、データセンタの省電力化を実現することができる。

【図面の簡単な説明】

【0018】

【図1】本発明のネットワーク運用管理装置（管理サーバ）の一例を示すブロック図である。

【図2】ネットワーク運用管理装置の物理サーバ情報管理データベースが管理する情報の例を示すテーブル図である。

【図3】ネットワーク運用管理装置の電力消費モデル管理データベースが管理する情報の例を示すテーブル図である。

【図4】ネットワーク運用管理装置が設置されるデータセンタ内の想定環境及び動作概要を示すモデル図である。

【図5】ネットワーク運用管理装置が設置されるデータセンタ内の一つのサーバラックに対する物理サーバ配置算出部とデータセンタ電力消費試算部、物理サーバ構成管理部の動作概要を示すモデル図である。

【図6】空調機とサーバ及びネットワーク機器の関係を示すモデル図である。

【図7】データセンタ内における空気の流れを説明するためのモデル図である。

【図8】データセンタ内におけるサーバラックの配置例を示すモデル図である。

【図9】図8に対応する排気効率(物理サーバからの排気熱がどの程度、空調機までとどくか)の例を示すモデル図である。

【図10】図8に対応する給気効率(他の物理サーバからの排気が空調機からの冷気にどの程度混ざるか)の例を示すモデル図である。

【図11】データセンタ内の一つのサーバに着目した空気の流れを示すモデル図である。

【図12】ある空調機の電力効率モデルを示すグラフ図である。

【図13】サーバの電力消費のモデルを示すグラフ図である。

【図14】ネットワーク運用管理装置が設置されるデータセンタ内の想定環境を示すモデル図である。

【図15】従来の物理サーバの使用率に応じて、物理リソースの論理サーバへの割り当てを変更する技術において、その動作概要を示すモデル図である。

【図16】従来技術において空調機とサーバ及びネットワーク機器の各電力消費量の関係

10

20

30

40

50

を示すモデル図である。

【図17】トラヒック量及び消費電力量の推計を示したグラフ図である。

【発明を実施するための形態】

【0019】

以下、本発明のネットワーク運用管理装置の実施形態の一例について、図1乃至図5を参照しながら説明する。

本発明のネットワーク運用管理装置は、複数のサーバ及びネットワーク機器と、それらの安定稼働のための空調機が設置されるデータセンタにおいて、データセンタの電力消費量を最小化することを目的に、時々刻々と変化するユーザのデータセンタの利用量に応じて、サーバの電力消費量と空調機の電力消費量の両方を考慮し、データセンタの構成を動的に変更する処理を行うものである。

10

【0020】

ネットワーク運用管理装置は、図4に示すように、仮想化技術が適用された物理サーバ1（論理的に複数台の論理サーバ1'として動作させる）と、物理サーバ1の安定稼働のため、物理サーバ1を冷却する空調機5を主な構成要素とするデータセンタDCにおいて、データセンタの構成（物理リソースの論理サーバへの割り当て）を管理する管理サーバ10を新たに設置することで実現される。各物理サーバ1には、物理サーバ1から排気する空気の温度を測定する温度センサ8と、各物理サーバ1が給気する空気の温度を測定する温度センサ9が設けられている。

【0021】

20

各物理サーバ1には、その物理サーバ1が安定稼働するための動作保障温度が、物理サーバを製造する製造会社等により設定される。データセンタDC内においては、各物理サーバ1が給気する空気の温度が、動作保障温度以下となるように、空調機5が出力する空気の温度が決まる。物理サーバ1の給気側に到達した空調機5からの空気は、各物理サーバ1の電力消費が熱エネルギーとなるため、各物理サーバ1内で温度が高くなり、排気側より排熱される。排気された空気は、空調機5に循環され、空調機5で冷却され再度、各物理サーバ1に到達するが、循環の際、各物理サーバ1から排気される電力消費により温められた空気の一部（全体の空気の流れに対する割合）は空調機5に循環することなく、一部の物理サーバ1（図4の例では上側の3つの物理サーバ1）の給気側に流れ込む。各物理サーバ1から排気された空気が、空調機5に循環することなく、そのまま物理サーバの吸気側に流れ込む空気の割合は、データセンタDCの広さや物理サーバ1及び空調機5の数や配置状態により異なる。

30

運用者は、本発明のネットワーク運用管理装置を使用する際、データセンタDC内の全ての物理サーバ1の組み合わせに対する空気の流れの割合を管理サーバ10に登録する。更には、CPUの製品型式などハードウェアスペックより算出可能な物理サーバの電力消費モデルと、空調機の製品型式から算出可能な空調機の電力消費モデルを空気の流れの割合に加えて管理サーバ10に登録する。

【0022】

空気の流れの割合（他のノードもしくはノード自身の排気が含まれる割合）は、データセンタDCの設計時点で運用者がシミュレーションにより求めたものを管理サーバ10に登録してもよいし、データセンタDC内の各サーバに対して設けた温度センサ8, 9等を用いて得た環境情報を基に管理サーバ10が算出してもよい。

40

例えば、各ノードの稼働状況において、温度センサ8及び温度センサ9から排気面及び給気面の温度を検出し、各ノードの稼働状況が異なる時点の温度センサ8および温度センサ9（各センサの値は、稼働状況が異なるため、異なる温度を示す）から、を求める方程式を作成し、解くことでを求める。

【0023】

単純のため、2つのノードのみで構成されたデータセンタDCを例に説明する。空調機が供給する空気の温度を T_{sup} 、ノード1の温度センサ8の温度を T_{1_out} 、ノード2の温度センサ8の温度を T_{2_out} 、同様にノード1の温度センサ9の温度を T_{1_in} 、ノード2

50

の温度を $T2_in$ とする。さらに、ノード i の給気に含まれるノード j からの排気の割合を ji と標記した場合、 $T1_in = (1 - (c11 + c21)) \times Tsup + c11 \times T1_out + c21 \times T2_out$ となる。異なる温度の組を2つ以上収集することで、上記式を解き、を求めることができる。

【0024】

管理サーバ10は、論理サーバ1'のリソースの使用状況(各論理サーバに割り当てられたCPU、メモリ、ネットワーク、ディスク等の物理リソースがどの程度使用されているか)を定期的、もしくは運用者からの要求により収集する。

そして、管理サーバ10は、物理サーバ1による電力消費量と空調機(CRAC)5による電力消費量の和が最小となるよう、どの物理サーバ1にどの論理サーバ1'を割り当てるのか、また各論理サーバ1'にどの程度の物理リソースを割り当てるのかを算出し、データセンタDC内における物理リソースの論理サーバへの割り当ての動的変更を行う。

10

【0025】

管理サーバ10ではこれらの処理が繰り返し実行されることで、データセンタDCの電力消費量が最小化されるように、時々刻々と変化するユーザからのサービス要求量に応じて、都度最適な(電力消費量を最小化する)データセンタDC内の構成を算出し、さらに構成変更を行う。

すなわち、管理サーバ10が、データセンタの電力消費量のうち、サーバ群、および、空調機による電力消費量を最小化するため、論理サーバの使用状況に応じて、物理リソースの論理サーバへの割り当てを動的に変更する処理を行うように構成されている。

20

【0026】

ネットワーク運用管理装置を実現する管理サーバ10の詳細構成について、図1を参照して説明する。

ネットワーク運用管理装置(管理サーバ)10は、各物理サーバ1に関する各種情報を管理する物理サーバ情報管理データベース11と、物理サーバ1及び空調機5の電源消費モデル等を管理する電力消費モデル管理データベース12と、運用者が情報を入力するための管理インタフェース提供部13と、電力消費量が最小となるデータセンタDCの構成候補を求める物理サーバ配置算出部14と、物理リソースの論理サーバへの割り当て変更(物理サーバの再配置)と再配置結果を管理する物理サーバ構成管理部15と、データセンタDCの構成候補から電力消費を試算するデータセンタ電力消費試算部16と、リソース利用状況取得部17と、各部を制御する中央制御部18を備えて構成されている。

30

【0027】

物理サーバ情報管理データベース11は、図2に示すように、複数の物理サーバ1の各ハードウェアスペック、各物理サーバを複数の論理サーバに割り当てた場合の割り当りリソース量(各論理サーバの使用率)等の物理サーバに関する各情報が管理されている。物理サーバのハードウェアスペックとは、例えば、CPUの製品型式のような情報である。これらの情報は、運用者が管理インタフェース提供部13から入力して記憶させる。

【0028】

電力消費モデル管理データベース12は、図3に示すように、物理サーバ1の電源消費モデル、空調機5の電源消費モデル、空気の流れ(空気の流れの割合)のモデルが登録されて管理されている。物理サーバ1の電源消費モデルとは、各物理サーバにおける物理リソースの使用量に対する電力消費量の値であり、各サーバに数式として登録されている。

40

物理サーバ1の電力消費モデルは、例えばCPUの製品型式から一意に求めることができるCPU使用率(物理リソースの使用量)と、その時の電力消費量とすることができる。空調機5の電源消費モデルとは、空調機の冷却能力に対する電力消費量である。

空調機5の電力消費モデルは、例えば空調機の製品型式から一意に求めることができる。空気の流れ(空気の流れの割合)のモデルとは、各サーバの排熱が他のサーバの給気側に流れ込む割合等であり、各サーバに対するテーブルにより管理される。空気の流れのモデルは、例えばデータセンタDCの設計時点で運用者がシミュレーションにより求

50

めることができる。管理されるテーブルの具体例については後述する。

【 0 0 2 9 】

電力消費モデルデータベース 1 2 で管理される内容は、データセンタ D C の設計時点で運用者がシミュレーションにより求めたものを、予め運用者により管理インタフェース部 1 3 を介して手動で入力して登録される。また、必要に応じて、データセンタ D C 内に配置された温度センサ 8 , 9 等を用いて、ネットワーク運用管理装置（管理サーバ）1 0 がセンサより収集した情報を基に算出するものであってもよい。

【 0 0 3 0 】

管理インタフェース提供部 1 3 は、運用者がデータセンタ D C 内の物理サーバやその構成、上記各電力消費モデルなどを登録するためのユーザインタフェースとしての機能を有する。例えば、運用者は、自身の P C にインストールされたブラウザ等により、管理サーバ 1 0 に接続し、各物理サーバ 1 に関する情報を登録する。登録された情報は物理サーバ情報管理データベース 1 1 に保存される。

10

また、電力消費モデル管理データベース 1 2 に登録されるデータセンタ D C 内の空気の流れは、温度センサ 8 , 9 等を用いて、管理サーバ 1 0 がセンサより収集した情報を基に算出して登録してもよいし、データセンタ D C の設計時点で運用者がシミュレーションにより求めた空気の流れを直接管理インタフェース提供部 1 3 を介して電力消費モデル管理データベース 1 2 に手動登録してもよい。

運用者がデータセンタ D C の電力消費を試算するために手動で登録する情報としては、各物理サーバ 1 の電力消費モデル、空調機 5 の電力消費モデル、データセンタ D C 内の空気の流れのモデルが考えられる。

20

【 0 0 3 1 】

物理サーバ配置算出部 1 4 は、データセンタ D C の電力消費量が最小となるデータセンタ D C の構成（どのように物理リソースを論理サーバに割り当てるか）を求めるものであり、管理インタフェース提供部 1 3 を介しての運用者からの指示、又は、中央制御部 1 8 からの定期的な呼び出しにより実行される。

電力消費量が最小となるデータセンタ D C の構成を求めることは、例えば、図 5 に示すように、各論理サーバの使用状況（論理サーバに割り当てた物理リソースがどの程度の割合で使用され、物理リソースがどの程度の割合で使用されているか）に応じ、物理リソースの論理サーバへの割り当てについて複数の組み合わせを求め、さらに、各組の構成とした場合の電力消費量を、データセンタ電力消費試算部 1 6 により組み合わせ毎に試算し、その結果、最も電力消費が最も小さい組み合わせを選択する手順により実現される。

30

【 0 0 3 2 】

物理サーバ構成管理部 1 5 は、物理サーバ配置算出部 1 4 の出力結果を基に、データセンタ D C の構成を変更する。図 5 の例では、物理リソースの論理サーバへの割り当てについて、物理サーバ配置算出部 1 4 が 4 種の構成について、それぞれデータセンタ D C の電力消費量を試算し、結果、データセンタ D C の電力消費量が一番少ない図 5 中最右の構成（電力消費量：1465.7W）を選択する。物理サーバ構成管理部 1 5 は、物理サーバ配置算出部 1 4 の結果に従い、物理リソースの論理サーバへの割り当てを変更する。さらに、物理サーバ情報管理データベース 1 1 では、物理サーバ毎に、実行する論理サーバ等を管理するため、物理サーバ構成管理部 1 5 が割り当てを変更した結果を物理サーバ情報管理データベース 1 1 に保存する。

40

【 0 0 3 3 】

データセンタ電力消費試算部 1 6 は、物理サーバ再配置算出部より呼び出され、データセンタ D C の構成候補から、その際のデータセンタ D C の電力消費を試算する機能を提供する。データセンタ電力消費試算部 1 6 は、物理サーバ電力消費試算部 1 6 a と空調機電力消費試算部 1 6 b からなり、それら各部が試算する電力消費量の和により、データセンタ D C の電力消費量が求められる。各部の試算のため、電力消費モデル管理データベース 1 2 より、物理サーバや空調機、さらにデータセンタ D C 内の空気の流れのモデルを取得する。

50

【 0 0 3 4 】

リソース利用状況取得部 17 は、運用者からの指示、または中央制御部からの定期呼び出しにより、論理サーバのリソースの使用状況を定期的に取得する。各論理サーバのリソースの使用状況は、前述の通り、利用者からのサービス要求量に応じ、時々刻々と変化する。また、使用状況の取得には、仮想化ソフトウェアが提供する管理インタフェース等を用いることで実現することもできる。取得した使用状況に関する情報は、物理サーバ情報管理データベース 11 に保存され、時々刻々と変化する各論理サーバによる物理リソースの使用状況が、物理サーバ情報管理データベース 11 に反映される。

【 0 0 3 5 】

中央制御部 18 は、各部を制御しデータセンタ DC の省電力オペレーションシステムの機能を実現するものである。

10

【 0 0 3 6 】

上述したネットワーク運用管理装置においては、データセンタ電力消費試算部 16 でデータセンタ DC 全体の電力消費量を算出するため、空調機 5 の消費電力についても考慮する必要がある。そのため、データセンタ内の各サーバに対する空調機 5 からの空気の流れを電力消費モデル管理データベース 12 にデータとして入力しておく必要がある。以下、各サーバに対する空調機 5 からの空気の流れ等の把握の仕方について、図 6 ~ 図 14 を参照しながら説明する。

【 0 0 3 7 】

図 6 において、ユーザからのサービス要求は、インターネットなどのネットワークを介し、データセンタ DC に到達する。データセンタ DC 内部には、ルータやスイッチなどのネットワーク機器 2 が存在し、ユーザからのサービス要求を適切なサーバ 1 に転送する。サービス要求を受け取ったサーバ 1 は、ユーザにサービス応答を返す。ユーザからのサービス要求量は、例えば平日昼間が多く休日夜間は少ないなど、時々刻々と変化する。

20

【 0 0 3 8 】

また、サービス要求量に応じ、ネットワーク機器群及びサーバ群の処理負荷も時々刻々と変化する。各サーバ 1 は、電力により稼働される。また、各サーバ 1 で消費された電力は、熱としてデータセンタ DC 内に放出される。各サーバ 1 には、そのサーバ 1 が正常に動作するための温度（動作保障温度）がサーバ 1 毎に仕様として決まっており、サーバ 1 を正常に動作させるためには、データセンタ DC 内において動作保障温度以下の空気をサーバ 1 に供給する必要がある。そのため、各サーバからの排熱は、データセンタ内に設置された空調機 (CRAC) 5 に取り込まれ、空調機 (CRAC) 5 にて取り込んだ各サーバの排熱を含む空気を冷却し、再度サーバ 1 に供給することが行われる。

30

【 0 0 3 9 】

この場合、図 6 に示すように、サーバ 1 からの排熱の 100% が空調機 (CRAC) 5 に取り込まれるわけではなく、その一部は、サーバ 1 に供給される (空調機 (CRAC) 5 により冷却された) 空気と混ざり、サーバ 1 に取り込まれることとなる。上述したように、各サーバ 1 に対しては、正常動作する温度（動作保障温度）以下の空気を供給する必要があるため、サーバ 1 からの排熱がサーバに供給される空気と多量に混ざる場合には、空調機 (CRAC) 5 の出力を大きくする必要がある (電力消費量も大きくなる)。また、Chiller は、データセンタ DC の屋外に設置される冷却のための熱源を生成する、一般家庭のエアコンに例えると室外機にあたる機器であり、空調機 (CRAC) 5 がデータセンタ DC 屋内の空気を冷却するための熱源 (例えば冷却水など) を空調機 (CRAC) 5 に提供する。

40

【 0 0 4 0 】

データセンタ DC の冷却方法は、多種多様なベンダより様々な方式が提案されている。多くのデータセンタ DC で一般的に採用されている構造としては、例えば図 7 に示すように、底上げした床面 4 より空調機 (CRAC) 5 からの空気を、サーバラック 3 に収納されたサーバに分配し、排熱を天井面より収集する。

その他、データセンタの天井より空調機 (CRAC) 5 からの空気をサーバに供給し、さらに排気も天井から行う天井方式冷却ソリューション方式や、サーバラック 3 ごとに空調機 (C

50

RAC) 5 を配置しサーバラック内で給気と排気を行う方式などが提案されている。いずれの方式であっても、サーバの電力消費により排熱が発生し、空調機 (CRAC) 5 が排熱を取り込み冷却し再度サーバに提供すること、及び、排熱の一部が空調機 (CRAC) 5 からの冷却された空気と混ざること (図 6 における空気の流れ) は変わらない。

【 0 0 4 1 】

一般的なデータセンタ DC においては、図 7 に示すように、複数の縦長のサーバラック 3 が設置され、各サーバラック 3 に複数のサーバが納められている。空調機 (CRAC) 5 により冷却された空気は、床下を通り、サーバラック 3 の給気面 (多くの場合ラック前面) から給気される。また、サーバの電力消費により発生した熱はサーバラック 3 の排気面 (多くの場合ラック背面) から排気される。排気は、天井付近を通り、空調機 (CRAC) 5 に取り込まれ、取り込まれた空気は空調機 (CRAC) 5 で冷却されて再度サーバラック 3 側に供給される。

10

【 0 0 4 2 】

各サーバラック 3 は、サーバの排熱が空調機 5 により冷却された給気用の空気と混合することを防ぐため、サーバラック 3 の給気面は一般的に他のサーバラック 3 の給気面と向かい合うように、また、排気面は他のサーバラック 3 の排気面と向かい合うように設置される。

サーバラック 3 に納められた各サーバからの排熱の一部は、空調機 (CRAC) 5 から供給される冷却された空気と混合するが、その混ざり方は、サーバラック 3 や冷却された空気を供給する位置、排熱が空調機 (CRAC) 5 に取り込まれる経路によりデータセンタ DC 毎に不均一となる。そのため、データセンタ DC 内の空気の流れは、データセンタ DC 毎に異なるものとなる。

20

【 0 0 4 3 】

図 7 に示したデータセンタ DC の構造のように、床面より冷気を供給し、天井面より排気を取り込む場合を例に、データセンタ DC 内で給気と排気の流れがどのようになるかの一例について、図 8 ~ 図 10 を参照しながら説明する。

データセンタ DC 内には、図 8 に示すように、10 個のサーバラックにそれぞれ 5 つのサーバ (サーバ A ~ E) を設置した場合を想定する。図 6 で述べたように、サーバからの排熱の一部は、自身もしくは他のサーバの給気に混ざり、その割合を α 、排熱を T_{out} とすると " $\alpha \times T_{out}$ " で表すことができる。また、残りは空調機 (CRAC) 5 に取り込まれ、空調機に取り込まれる熱を T_{ACin} とすると、" $T_{ACin} = T_{out} - (\alpha \times T_{out})$ " と表すことができる。

30

図 9 に、各サーバの排気がどの程度空調機 (CRAC) 5 に取り込まれるかの一例を示す ($\beta = (T_{out} - (\alpha \times T_{out})) / T_{out}$)。縦軸の 1.0 は、排気の全てが (CRAC) 5 に取り込まれることを示す。

【 0 0 4 4 】

図 10 に、各サーバ給気に、どの程度、他もしくは自身の排熱が混ざることか、その割合 ($1 - \beta$) の一例を示す。本例では、サーバラックにおいて物理的に低い位置に設置されたサーバ (サーバ A, B, C) の給気には、ほとんど他もしくは自身からの排熱が混ざることとはなく、物理的に高い位置に設置されたサーバ (サーバ D, E) の給気には、他もしくは自身の排気が多く混ざることが分かる。

40

【 0 0 4 5 】

図 9 の排気効率及び図 10 の給気効率は、データセンタの想定環境により分布は異なるが、一般的に空気の流れは図 11 のように表現することができる。図 11 において、サーバ (S_i) ($i = 1 \sim 3$) からの排気が隣接するサーバ (S_j) ($j = 1 \sim 3$) の給気に混ざり割合を γ_{ij} と表す。

空調機 (CRAC) 5 の電力消費量は、供給する冷気の温度 (T_{sup}) と冷却する熱量から求まる。図 12 に実験的に求めたある空調機 (CRAC) 5 の供給する空気の温度 (T_{sup}) とその電力効率 (COP, Coefficient of Performance) を示す。

【 0 0 4 6 】

図 12 より、供給する空気の温度が高いほど、電力効率が優れていることが分かる。例

50

えば、10kwの熱を冷却し15 としたい場合、COP = 2 = 0.0068*15*15+0.0008*15 + 0.458 であることから、空調機(CRAC) 5の電力消費量は5kw = 10kw / 2となる。

また、10kwの熱を冷却し20 としたい場合、COP=3.338 であることから、空調機(CRAC) 5の電力消費量は2.996kw (電力消費量が40%削減される)となる。一方で、最終的に各サーバに到達する空気の温度(Tin)は、各サーバが正常に動作するための温度(動作保障温度)以下となる必要がある。

Cillerの電力消費量は、図6に示した通り、空調機(CRAC) 5の動作と密接に関係しており、空調機(CRAC) 5の電力消費量に単純に比例するものと考えることができる。

【0047】

サーバの電力消費量は、図13中の実線または点線で示されるような、サーバのCPU等の物理リソースの使用率に応じて変化する特性を有することが知られている。図13中の実線及び点線のいずれの特性となるかは、CPUの製品型式により決まる。実線及び点線のいずれの特性においても、「使用率の増加に比例して、電力消費量が増加する」と想定できる。図13における実線を数式で示すと下記の式1となる。このような電力消費のモデルは、一般的に全てのサーバに適用可能なものである。

P = Pidle + (Pmax - Pidle) * n (式1)

Pidle:サーバのCPU使用率が0%の時の電力消費量

Pmax:サーバのCPU使用率が100%の時の電力消費量

n:サーバの使用率(0~1)

【0048】

データセンタDCの電力消費量の具体例について、図14を参照して説明する。

図14では、5つのサーバを設置したデータセンタDCを仮定する。データセンタDC内の空気の流れは、表1に示したようなサーバ同士のテーブルを仮定する。表1に示されるデータセンタDC内の空気の流れのモデルは、電力消費モデル管理データベース12に記憶され、これらの情報は、前記したデータセンタ電力消費試算部16でデータセンタDCの電力消費を試算する際に用いられる。

表1では、サーバAの排熱のうち、5%が自身に、2.5%がサーバBに、10%がサーバCに、2.5%がサーバDに、2.5%がサーバEの給気と混ざり、残りが空調機(CRAC)5に取り込まれるとしている。

【0049】

【表1】

	A	B	C	D	E
サーバA	5	2.5	10	25	25
サーバB	1	5	10	25	25
サーバC	1	1	5	10	10
サーバD	1	1	1	5	10
サーバE	1	1	1	1	5

【0050】

また、サーバの電力消費は、前記した式1で計算し、Pidle = 100W(サーバの使用率が0%の時の電力消費量)、Pmax = 300W(サーバの使用率が100%の時の電力消費量)と仮定した(実際の高密度サーバではさらに大きな電力を消費するが、ここではあくまで例示するための値とする)。

【0051】

また、各サーバの排気温度は、給気温度が15 だとすると、300W(サーバの使用率100%)消費した場合は30、100W(Pidle)消費した場合、20 の排気温度になるとする。

排気温度 = 給気温度 + K * サーバの電力消費量

(注) Kは、空気の密度とサーバを流れる空気の流量、空気の比熱から求めることができる係数。ここでは簡単のため0.05(定数)を用いる。

さらに、空調機(CRAC)の電力消費は、図12に前述の電力消費モデルを用いる。また、冷却する熱量は、簡単のため、サーバの電力消費量の和とした。

サーバが安定的に動作するための温度は、25とし、空調機は、起動中のサーバの給気温度が25以下となるよう0.5刻みで出力を変更するものとする。

【0052】

上述の各値を使用し、各サーバの使用率を図15における左の例とした場合、データセンタ全体DCの電力消費量(サーバと空調機(CRAC)の電力消費量の和)は、

空調機(CRAC)の設定温度: 16.5度

サーバの電力消費量の和: 1090W

空調機(CRAC)の電力消費量の和: 469.3W

データセンタDC全体の電力消費量: 1559.3W

となる。

【0053】

また、各サーバの使用率を図15における右の例とした場合、データセンタDC全体の電力消費量(サーバと空調機(CRAC)の電力消費量の和)は、

空調機(CRAC)の設定温度: 16度

サーバの電力消費量の和: 1080W

空調機(CRAC)の電力消費量の和: 488.3W

データセンタDC全体の電力消費量: 1568.3W

となる。

【0054】

上述の2例によれば、図15における左の例から右の例とすることで、サーバ1台を休止状態とすることでサーバの電力消費量の和を1090Wから1080Wに削減したにも関わらず、空調機(CRAC)の電力消費量が増加し、データセンタDC全体の電力消費量が1559.3Wから1568.3Wに増加している。

すなわち、サーバの使用率にのみ注目し、サーバ群の電力消費量を少なくしただけでは、データセンタDC全体の電力消費量が大きくなる可能性がある。

【0055】

これに対して上述した本発明の実施例によれば、データセンタ電力消費試算部16を設け、物理サーバ電力消費試算部16aと空調機電力消費試算部16bが試算する電力消費量の和によりデータセンタDCの電力消費量を求めるようにしたので、物理サーバ配置算出部14によりデータセンタDC全体の電力消費量が最小となる論理サーバの構成(物理リソースの論理サーバへの割り当ての組み合わせ)を選択することができる。

【符号の説明】

【0056】

1...物理サーバ(サーバ)、 1'...論理サーバ 2...ネットワーク機器、 3...サーバラック、 4...床面、 5...空調機、 6...室外機、 8, 9...温度センサ、 10...管理サーバ(ネットワーク運用管理装置)、 11...物理サーバ情報管理データベース、 12...電力消費モデル管理データベース、 13...管理インタフェース提供部、 14...物理サーバ配置算出部、 15...物理サーバ構成管理部、 16...データセンタ電力消費試算部、 17...リソース利用状況取得部、 18...中央制御部、

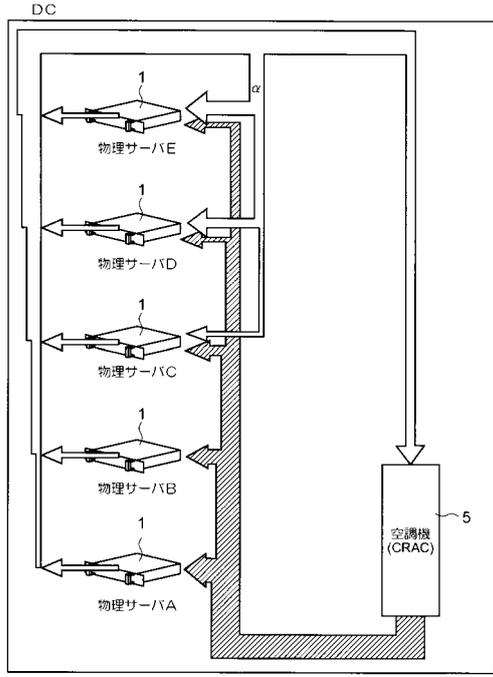
10

20

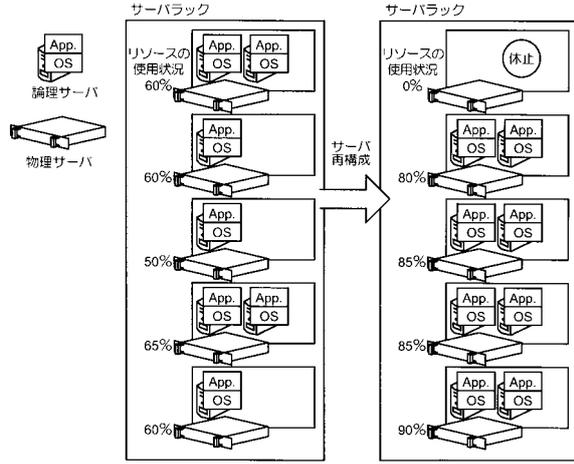
30

40

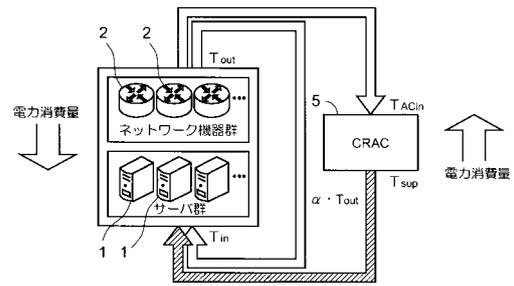
【図14】



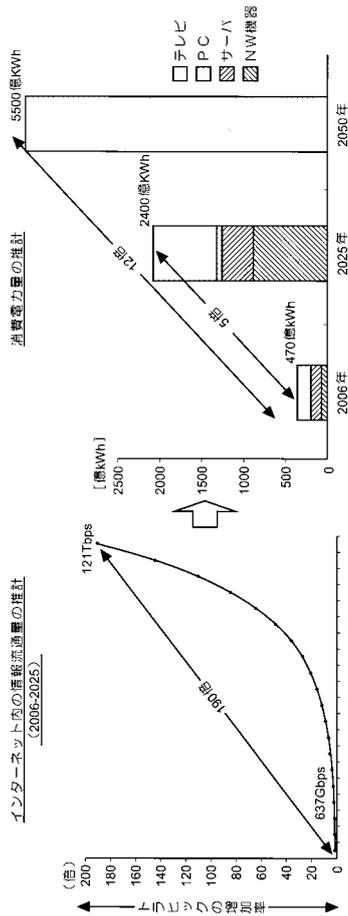
【図15】



【図16】

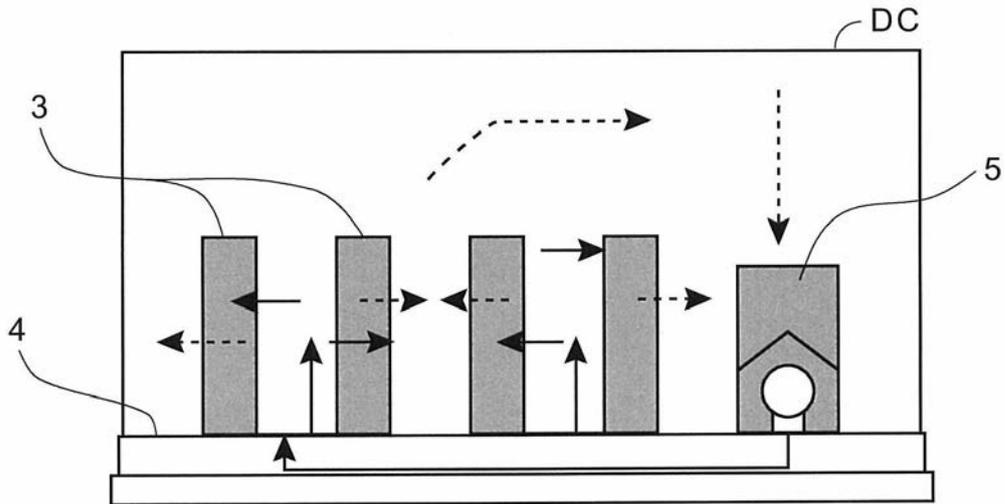


【図17】



経済産業省 グリーンITイニシアチブ会議 (第1回)
 (http://www.meti.go.jp/press/2007/12/07/06/20071207005.html)

【図7】



フロントページの続き

審査官 漆原 孝治

(56)参考文献 国際公開第2008/092001(WO, A1)

特開2007-179437(JP, A)

荒井 大輔 Daisuke ARAI, データセンタ省電力運用管理手法の提案 Proposal on ECO-friendly Data Center Operation Scheme, 電子情報通信学会技術研究報告 Vol. 108 No. 458 IEICE Technical Report, 2009年 2月24日, 第108巻, 第458号, p.469-474

(58)調査した分野(Int.Cl., DB名)

G06F 9/50

G06F 1/20

G06F 1/32