

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局



(10) 国際公開番号

WO 2011/043317 A1

(43) 国際公開日

2011年4月14日(14.04.2011)

PCT

- (51) 国際特許分類:
H04L 12/28 (2006.01) G06F 9/46 (2006.01)
G06F 1/32 (2006.01) G06F 9/50 (2006.01)
- (21) 国際出願番号: PCT/JP2010/067414
- (22) 国際出願日: 2010年10月5日(05.10.2010)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2009-233366 2009年10月7日(07.10.2009) JP
- (71) 出願人 (米国を除く全ての指定国について): 日本電気株式会社(NEC CORPORATION) [JP/JP]; 〒1088001 東京都港区芝五丁目7番1号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 村上 貴彦 (MURAKAMI Takahiko) [JP/JP]; 〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP). 大和 純一(YAMATO Junichi) [JP/JP]; 〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP).
- (74) 代理人: 工藤 実(KUDOH Minoru); 〒1400013 東京都品川区南大井六丁目24番10号カドヤビル6階 Tokyo (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, 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, 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, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR,

[続葉有]

(54) Title: POWER SAVING SYSTEM AND POWER SAVING METHOD

(54) 発明の名称: 省電力化システム、及び省電力化方法

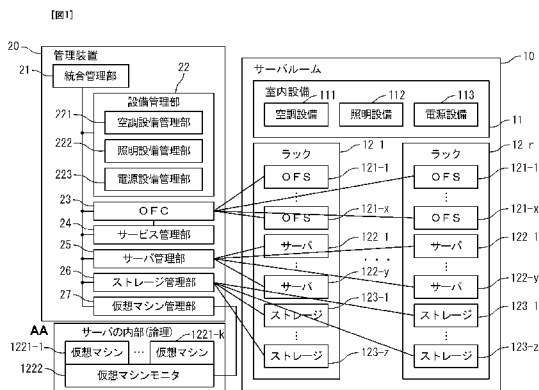


FIG. 1:
 10 SERVER ROOM
 11 INTERIOR EQUIPMENT
 111 AIR-CONDITIONING EQUIPMENT
 112 LIGHTING EQUIPMENT
 113 POWER SOURCE EQUIPMENT
 12-1, 12-n RACK
 122-1, 122-y SERVER
 123-1, 123-z STORAGE
 20 MANAGEMENT DEVICE
 21 INTEGRATED MANAGEMENT UNIT
 22 EQUIPMENT MANAGEMENT UNIT
 221 AIR-CONDITIONING EQUIPMENT MANAGEMENT UNIT
 222 LIGHTING EQUIPMENT MANAGEMENT UNIT
 223 POWER SOURCE EQUIPMENT MANAGEMENT UNIT
 24 SERVICE MANAGEMENT UNIT
 25 SERVER MANAGEMENT UNIT
 26 STORAGE MANAGEMENT UNIT
 27 VIRTUAL MACHINE MANAGEMENT UNIT
 1221-1, 1221-k VIRTUAL MACHINE
 1222 VIRTUAL MACHINE MONITOR
 AA INSIDE OF SERVER (LOGIC)

(57) Abstract: Provided are a power saving system and a power saving method in which power saving is achieved by stopping unnecessary devices in a server room and interior equipment. Specifically, consolidation of virtual machines (VM) providing services is performed, processes are concentrated on a server in a predetermined rack out of a group of racks in the server room, the racks having no servers in operation are stopped, processes of interior equipment operating for the racks and their surroundings are stopped or adjusted so as to achieve power saving.

(57) 要約: サーバルーム内の不用になった機器や室内設備を停止し、省電力化を実現する。具体的には、サービスを提供する仮想マシン (VM) の片寄せを行い、サーバルーム内のラック群のうち、所定のラック内のサーバに処理を集中させ、稼働中のサーバがなくなったラックを停止し、当該ラック又はその周辺を対象にした室内設備の処理も停止、或いは、調節し、省電力化を図る。

GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG). 添付公開書類:
— 国際調査報告 (条約第 21 条(3))

明 細 書

発明の名称：省電力化システム、及び省電力化方法

技術分野

[0001] 本発明は、省電力化システムに関し、特に仮想ネットワーク環境における省電力化システムに関する。

背景技術

[0002] 近年、ネットワークと仮想マシン（VM：Virtual Machine）を連動して柔軟なクラウド・コンピューティングを実現し、インフラ（Infrastructure）を構築したシステムが利用されている。仮想マシンは、アプリケーションプログラムを実行することで、ネットワークを介して様々なサービスをユーザに提供する。

[0003] しかし、このような仮想マシンの多くは、企業等の組織が有するサーバールーム内のラック（rack）群に搭載された複数の物理サーバ上で稼動している。そのため、少なくない数のサーバやスイッチ、及びサーバールームの室内設備を常に稼動し続ける必要があり、それに伴う消費電力量やリソース消費量も膨大なものとなっていた。

[0004] 現状では、サーバやスイッチ、及びサーバールームの室内設備のような仮想マシンの稼動に必要な機器に対する電力供給は、機器の使用状態に関係なく実施されている。また、電力供給を停止する方法としては、機器が不使用状態になったことを確認した上で、人間の手により物理的に機器を停止させるしかなかった。

[0005] そこで、効率的な省電力化を実現するために、コンピュータとネットワークを連携し、負荷に応じて仮想マシンを少数のサーバに集約し、機器の使用状態に応じて、機器に対する電力供給を制御する手法が求められる。

[0006] 関連する技術として、特開2009-169858号公報（特許文献1）にサーバシステム、及びサーバシステムの電力削減方法が開示されている。特許文献1には、仮想マシンのマイグレーション（migration）に

よって稼動するサーバ数を削減し、消費電力を削減するサーバシステムが記載されている。また、サーバの消費電力が最大電力に達した場合、仮想マシンを他のサーバに移動する技術が記載されている。なお、マイグレーションとは、プログラムやデータの移行・変換作業のことであり、特に、OS等の環境が異なるシステムへの移行を指すことが多い。

[0007] また、特開2009-176033号公報（特許文献2）にストレージシステム及びその消費電力削減方法が開示されている。特許文献2には、仮想マシンである仮想サーバをマイグレーションすることで稼動するサーバ数を削減し、消費電力を削減するストレージシステムが記載されている。また、仮想サーバによるネットワークインタフェース、ストレージインタフェース、制御装置に使用頻度を監視し、時間帯毎に統計情報を作成し、仮想サーバの稼動率を決定する技術が記載されている。

[0008] また、特開2003-281008号公報（特許文献3）にサーバ計算機負荷分配装置、サーバ計算機負荷分配方法、サーバ計算機負荷分配プログラム及びサーバ計算機システムが開示されている。特許文献3には、インフラ、データセンタ、ワークステーション、ラック、ブレード単位で電力制御し、不要な電力消費を抑えるための技術が記載されている。

[0009] また、特開2009-181571号公報（特許文献4）に大規模データセンターまたはITインフラにおけるエネルギー消費の確定及び動的制御方法が開示されている。特許文献4には、サーバへのデータ要求量に対して、サーバシステムのデータ供給可能量が最適な大きさとなるように、最適な数のサーバのみを稼動させるシステムが記載されている。

先行技術文献

特許文献

[0010] 特許文献1：特開2009-169858号公報

特許文献2：特開2009-176033号公報

特許文献3：特開2003-281008号公報

特許文献4：特開2009-181571号公報

発明の概要

- [0011] 本発明の目的は、コンピュータとネットワークを連携させた負荷判断及び省電力化を行う省電力化システムを提供することである。
- [0012] 本発明の省電力化システムは、管理装置と、コントローラとを具備する。管理装置は、複数のサーバ上でそれぞれ稼動する仮想マシンにより提供されるサービスの通信情報を監視し、通信情報から求められた仮想マシン毎のサービス負荷を基に、仮想マシンを他のサーバ上へ移動させ、稼動する仮想マシンが存在しなくなったサーバを未稼動状態にする。コントローラは、サービスの通信パケットの転送を行うスイッチに対してフロー単位の経路情報を設定する。また、コントローラは、仮想マシンの移動に従ってスイッチに対して設定した経路情報を変更する。
- [0013] 本発明の省電力化方法は、コンピュータにより実施される省電力化方法である。この省電力化方法では、複数のサーバ上でそれぞれ稼動する仮想マシンにより提供されるサービスの通信情報を監視する。また、通信情報から求められた仮想マシン毎のサービス負荷を基に、仮想マシンを他のサーバ上へ移動させ、稼動する仮想マシンが存在しなくなったサーバを未稼動状態にする。また、サービスの通信パケットの転送を行うスイッチに対してフロー単位の経路情報を設定する。また、仮想マシンの移動に従ってスイッチに対して設定した経路情報を変更する。
- [0014] 本発明に係る記憶媒体は、以下のステップをコンピュータに実行させるための省電力化用プログラムを記憶した記憶媒体である。この記憶媒体から省電力化用プログラムを読み出して実行したコンピュータは、第1のステップとして、複数のサーバ上でそれぞれ稼動する仮想マシンにより提供されるサービスの通信情報を監視する。また、第2のステップとして、通信情報から求められた仮想マシン毎のサービス負荷を基に、仮想マシンを他のサーバ上へ移動させ、稼動する仮想マシンが存在しなくなったサーバを未稼動状態にする。また、第3のステップとして、サービスの通信パケットの転送を行うスイッチに対してフロー単位の経路情報を設定する。また、第4のステップ

として、仮想マシンの移動に従ってスイッチに対して設定した経路情報を変更する。

- [0015] これにより、サービスを提供する仮想マシン（VM）の片寄を行い、稼働中の仮想マシンが存在しなくなったサーバを未稼働状態にすることで、省電力化を図る。また、仮想マシンの移動により、ラック群のうち所定のラック内のサーバに仮想マシンを集中させることで、サービスの提供に必要な機器や設備を削減し、不用になった機器や設備を停止することで、省電力化を実現することができる。

図面の簡単な説明

- [0016] [図1]本発明の省電力化システムの構成例を示すブロック図である。
[図2]サービスの基本処理を示すシーケンス図である。
[図3]フローの定義情報を設定した後の情報の収集及び構成変更の契機を示すシーケンス図である。
[図4]稼動リソース最適化処理を示すフローチャートである。
[図5]移動先サーバの決定処理を示すフローチャートである。
[図6]稼動処理を示すフローチャートである。
[図7]停止処理を示すフローチャートである。

発明を実施するための形態

- [0017] 以下に、本発明の実施形態について添付図面を参照して説明する。

- [0018] <基本構成>

図1に示すように、本発明の省電力化システムは、サーバルーム10と、管理装置20とを含む。

- [0019] サーバルーム10は、企業等の組織内において多数のサーバを収容するための部屋である。なお、サーバルームは一例に過ぎない。実際には、本発明の要旨を逸脱しない範囲で、サーバルームと類似の環境をサーバルーム10としても良い。管理装置20は、サーバルーム10内の環境を管理するための装置であり、サーバルーム10内の各種の設備や機器とネットワークを介して接続されている。管理装置20は、1台で複数のサーバルーム10を管

理しても良い。反対に、管理装置 20 は、複数台で 1 つのサーバールーム 10 を管理しても良い。ここでは、説明の簡略化のため、1 台の管理装置 20 が、サーバールーム 10 を 1 部屋だけ管理する事例について説明する。

[0020] サーバールーム 10 は、室内設備 11 と、ラック 12 (12-i、i=1~n : n は任意) を備える。

[0021] 室内設備 11 は、空調設備 111 と、照明設備 112 と、電源設備 113 を備える。

[0022] 空調設備 111 は、サーバールーム内の空調機全般である。この空調設備 111 は、ネットワークを介して受信した信号やコマンドに応じて、空調制御を実施したり、空調制御を停止したり、設定温度/風力/対象範囲を調節したりする。ここでは、空調設備 111 は、室内全体の空調を担当する空調機、ラック (rack) の近傍あるいはラック自体に設置された空調機を含む。なお、空調機 1 つにつき冷却されるラックの集合を、その空調機の「空調グループ」と呼ぶ。すなわち、1 つの空調機は、その空調機の空調グループに属するラックを冷却する。空調設備 111 の例として、AHU (Air Handling Unit : エアハンドリングユニット)、局所冷却装置 (High heat density cooling system)、熱交換器、エアコン、送風ファン等が考えられる。また、空調設備 111 は、水冷式の冷却機構 (水冷システム) でも良い。本質的に同じだからである。但し、実際には、これらの例に限定されない。

[0023] 照明設備 112 は、サーバールーム内の照明器具全般である。この照明設備 112 は、ネットワークを介して受信した信号やコマンドに応じて、特定の場所を明るくしたり、暗くしたりする。ここでは、照明設備 112 は、室内全体を照らす照明器具、ラックの近傍あるいはラック自体に設置された照明器具を含む。なお、照明器具 1 つにつき照らされるラックの集合を、その照明器具の「照明グループ」と呼ぶ。すなわち、1 つの照明器具は、その照明器具の照明グループに属するラックを照らす。照明設備 112 の例として、蛍光灯、蛍光ランプ、LED (Light Emitting Diode

: 発光ダイオード) 照明、電球等が考えられる。但し、実際には、これらの例に限定されない。

[0024] 電源設備 113 は、サーバールーム内の電源装置全般である。この電源設備 113 は、ネットワークを介して受信した信号やコマンドに応じて、電力供給を実施したり、電力供給を停止したり、供給される電力量を調節したりする。ここでは、電源設備 113 は、サーバや各種ネットワーク機器に電力を供給 (配電) するための電源装置、空調設備に電力を供給 (配電) するための電源装置、照明設備に電力を供給 (配電) するための電源装置を含む。なお、電源装置 1 つにより配電されるラックの集合を、その電源装置の「電源グループ」と呼ぶ。すなわち、1 つの電源装置は、その電源装置の電源グループに属するラックに配電する。電源設備 113 の例として、インテリジェント PDU (Power Distribution Unit: ラック用電源タップ)、UPS (Uninterruptible Power Supply: 無停電電源装置) 等が考えられる。但し、実際には、これらの例に限定されない。

[0025] ラック 12 (12-i、i=1~n) は、IDC (Internet Data Center: インターネットデータセンター) 等に設置されている、ラックマウント型サーバ (rack-mountable server) を効率的に設置できる形状の電子機器収納専用ラックである。なお、1 つのラックに収納される電子機器の集合を、そのラックの「ラックグループ」と呼ぶ。すなわち、1 つのラックは、そのラックのラックグループに属する電子機器を収納する。

[0026] ラック 12 (12-i、i=1~n) の各々は、OFS (OpenFlow Switch) 121 (121-i、i=1~x: x は任意) と、サーバ 122 (122-i、i=1~y: y は任意) と、ストレージ 123 (123-i、i=1~z: z は任意) を備える。なお、ラック内の構成 (例えば、サーバの台数) は、ラック毎に異なっても良い。

[0027] OFS 121 (121-i、i=1~x) は、物理スイッチであり、ラッ

ク内に少なくとも1台は存在する。同一ラック内のサーバ122 (122-i、i=1~y) は、同一ラック内のOFS121 (121-i、i=1~x) に接続される。ラック間の接続は、ラック内のOFSと、他のラック内のOFSを接続することで行う。

[0028] OFS121 (121-i、i=1~x) は、冗長経路を作る場合には2台以上存在する。この場合、サーバ122 (122-i、i=1~y) の各々は、異なる2台以上のOFS121 (121-i、i=1~x) と接続される。OFS121 (121-i、i=1~x) の例として、ラックに搭載されるエッジルータ (Edge Router) を想定している。なお、ラック内に限らず、サーバルーム10と管理装置20の間にも、OFSが存在している場合がある。但し、実際には、これらの例に限定されない。

[0029] サーバ122 (122-i、i=1~y) は、物理サーバであり、ラック内に少なくとも1台は存在する。サーバ122 (122-i、i=1~y) の例として、ラックマウント型サーバを想定している。但し、実際には、これらの例に限定されない。

[0030] サーバ122 (122-i、i=1~y) の各々は、仮想マシン (VM) 1221 (1221-i、i=1~k : kは任意) と、仮想マシンモニタ (VMM: Virtual Machine Monitor) 1222を備える。

[0031] 仮想マシン1221 (1221-i、i=1~k) は、論理サーバであり、サーバ122 (122-i、i=1~y) 上で稼動し、サービスを提供する。サービスとは、Webサービス等、ネットワークを介してユーザマシンから要求され、アプリケーションプログラムにより実行される処理である。複数のサーバ (論理サーバを含む) で同一のサービスを提供する場合もある。

[0032] 仮想マシン1221 (1221-i、i=1~k) が使用するディスクイメージは、共有ディスク (Shared Disk) に格納される。この共有ディスクは、サーバ122 (122-i、i=1~y) の各々が備えてい

ても良いし、ストレージ123（123-i、i=1~z）やOFS121（121-i、i=1~x）、又はサーバルーム10内の全ラックに共通の記憶装置、或いは管理装置20が備えていても良い。

[0033] 仮想マシン1221（1221-i、i=1~k）には、それぞれMACアドレス（Media Access Control Address）やIPアドレス（Internet Protocol Address）が割り当てられている。このMACアドレスは、仮想MACアドレスであり、任意に設定することができる。同様に、このIPアドレスは、仮想IPアドレスであり、任意に設定することができる。なお、同一のサーバ122（122-i、i=1~y）内や、各サーバ122（122-i、i=1~y）間を、仮想マシン1221（1221-i、i=1~k）が移動しても、仮想マシン1221（1221-i、i=1~k）のMACアドレスやIPアドレスは変わらない。

[0034] 仮想マシンモニタ1222は、仮想マシン1221（1221-i、i=1~k）の生成・稼動・移動・停止・削除を監視し管理する。また、仮想マシンモニタ1222は、仮想マシン1221（1221-i、i=1~k）のメモリ管理を行う。すなわち、仮想マシンモニタ1222は、仮想マシン1221（1221-i、i=1~k）で使用されるデータやプログラムをストレージ123（123-i、i=1~z）に格納したり取り出したりする。

[0035] 更に、仮想マシンモニタ1222は、他の仮想マシンモニタとの間で通信することが可能である。仮想マシンモニタ間の通信とは、互いに異なる仮想マシンモニタに属する仮想マシン1221（1221-i、i=1~k）同士の通信である。仮想マシンモニタ1222上の仮想マシン1221（1221-i、i=1~k）に対しては、仮想マシン1221（1221-i、i=1~k）のMACアドレスやIPアドレスを宛先として、外部から直接通信できる。なお、仮想マシンモニタ1222は、サーバ内外の仮想マシン間をブリッジ接続する仮想スイッチとしての機能を持つ。

- [0036] ストレージ123 (123-i、i=1~z) は、データやプログラムを格納する補助記憶装置である。ストレージ123 (123-i、i=1~z) の例として、HDD (Hard Disk Drive) やSSD (Solid State Drive) 等が考えられる。
- [0037] また、ストレージ123 (123-i、i=1~z) は、コンピュータ本体に内蔵された記憶装置に限らず、周辺機器 (外付けHDD等) や外部のサーバ (ストレージサーバ等) に設置された記憶装置でも良い。例えば、ストレージ123 (123-i、i=1~z) は、DAS (Direct Attached Storage)、FC-SAN (Fibre Channel - Storage Area Network)、NAS (Network Attached Storage)、IP-SAN (IP - Storage Area Network) 等を用いたストレージ装置でも良い。すなわち、ストレージ123 (123-i、i=1~z) は、必ずしもラック内になくても良い。但し、実際には、これらの例に限定されない。
- [0038] 管理装置20は、統合管理部21と、設備管理部22と、OFC (OpenFlow Controller) 23と、サービス管理部24と、サーバ管理部25と、ストレージ管理部26と、仮想マシン管理部27を備える。
- [0039] 統合管理部21は、設備管理部22、OFC23、サービス管理部24、サーバ管理部25、ストレージ管理部26、及び仮想マシン管理部27を総合的に管理する。統合管理部21は、これらの管理のために、予め必要な設定情報を保持している。
- [0040] また、統合管理部21は、必要に応じて、設備管理部22、OFC23、サービス管理部24、サーバ管理部25、ストレージ管理部26、及び仮想マシン管理部27を介して、室内設備11、OFS121 (121-i、i=1~x)、サーバ122 (122-i、i=1~y)、及びストレージ123 (123-i、i=1~z) に関する情報を収集し、収集された情報を

編集して保持する。

- [0041] 更に、統合管理部 2 1 は、サーバ管理部 2 5 及び仮想マシン管理部 2 7 に指示（命令を通知）して、サービスを提供する仮想マシン 1 2 2 1（1 2 2 1 - i、i = 1 ~ k）をマイグレーションすることで、所定のサーバ 1 2 2（1 2 2 - i、i = 1 ~ y）に片寄せするようにしても良い。
- [0042] 設備管理部 2 2 は、空調管理部 2 2 1 と、照明管理部 2 2 2 と、電源管理部 2 2 3 を備える。
- [0043] 空調管理部 2 2 1 は、指定された空調 I D の空調設備 1 1 1 の起動、停止、又は調節を行う。具体的には、空調管理部 2 2 1 は、ネットワークを介して空調設備 1 1 1 に信号やコマンドを送信し、空調設備 1 1 1 による冷却を開始／停止したり、冷却の程度（設定温度、風力、対象範囲、稼動設備の数、等）を調節したりする。
- [0044] 照明管理部 2 2 2 は、指定された照明 I D の照明設備 1 1 2 の起動、停止、又は調節を行う。具体的には、照明管理部 2 2 2 は、ネットワークを介して照明設備 1 1 2 に信号やコマンドを送信し、照明設備 1 1 2 を点灯／消灯したり、照明の程度（輝度、光度、対象範囲、稼動設備の数、等）を調節したりする。
- [0045] 電源管理部 2 2 3 は、指定された電源 I D の電源設備 1 1 3 の起動、停止、又は調節を行う。具体的には、電源管理部 2 2 3 は、ネットワークを介して電源設備 1 1 3 に信号やコマンドを送信し、電源設備 1 1 3 による電力の供給（配電）を開始／停止したり、電力の供給（配電）の程度（供給電力、対象範囲、稼動設備の数、等）を調節したりする。
- [0046] O F C 2 3 は、オープンフロー（Open Flow）技術により、システム内の通信を制御するコントローラである。
- [0047] オープンフロー技術とは、コントローラ（ここでは O F C 2 3）が、ルーティングポリシーとして自身に設定されたフロー定義情報に従い、マルチレイヤ及びフロー単位の経路情報（フローテーブル）をスイッチに設定し、経路制御やノード制御を行う技術を示す。

- [0048] オープンフローの詳細については、以下の文献に記載されている。 “OpenFlow Switch Specification Version 0.9.0 (Wire Protocol 0x98) July 20, 2009 Current Maintainer: Brandon Heller (brandonh@stanford.edu)” <<http://www.openflowswitch.org/documents/openflow-spec-v0.9.0.pdf>>
- [0049] これにより、経路制御機能がルータやスイッチから分離され、コントローラによる集中制御によって最適なルーティング、トラフィック管理が可能となる。オープンフロー技術が適用されるスイッチは、従来のルータやスイッチのようにパケットやフレームの単位ではなく、END2END (End to End) のフローの単位で通信を取り扱う。
- [0050] 詳細には、OFC23は、スイッチやノード毎にフローを設定することで当該スイッチやノードの動作（例えばパケットデータの中継動作）を制御する。
- [0051] [フローの例]
- フローとは、所定の規則（ルール）に適合するパケットに対して行うべき所定の処理（アクション）を定義したものである。フローのルールは、MACフレームのヘッダ領域に含まれる宛先アドレス (Destination Address)、送信元アドレス (Source Address)、宛先ポート (Destination Port)、送信元ポート (Source Port) のいずれか又は全てを用いた様々な組み合わせにより定義され、区別可能である。なお、上記のアドレスには、MACアドレス (Media Access Control Address) やIPアドレス (Internet Protocol Address) を含むものとする。また、上記に加えて、入口ポート (Ingress Port) の情報も、フローのルールとして使用可能である。
- [0052] ここで、OFC23によって制御されるスイッチは、OFS121 (12

1-i、i=1~x)や、仮想マシン1221(1221-i、i=1~k)が機能として持つ仮想スイッチ等である。また、OFC23によって制御されるノードは、サーバ122(122-i、i=1~y)、仮想マシン1221(1221-i、i=1~k)、仮想マシンモニタ1222、ストレージ123(123-i、i=1~z)等である。

[0053] すなわち、OFC23は、OFS121(121-i、i=1~x)、サーバ122(122-i、i=1~y)、ストレージ123(123-i、i=1~z)の各々をノードとみなしてノード管理を行う。

[0054] 例えば、OFC23は、各ノードの負荷監視や、指定されたスイッチIDのOFS121(121-i、i=1~x)の起動及び停止を行う。なお、停止には、ハードウェアのシャットダウンに限らず、スタンバイやスリープ、休止状態等の省電力モードへの移行を含むものとする。

[0055] サービス管理部24は、サービスの論理構成・物理構成の管理を行う。サービスは、当該サービスを提供するためのアプリケーションプログラムを実行している仮想マシン1221(1221-i、i=1~k)により提供される。また、サービス管理部24は、サービス毎の負荷監視を行う。例えば、サービス管理部24は、OFC23からフロー数を収集する。

[0056] サーバ管理部25は、サーバ122(122-i、i=1~y)の動作や負荷の監視を行う。また、サーバ管理部25は、指定されたサーバIDのサーバ122(122-i、i=1~y)の起動及び停止を行う。なお、停止には、ハードウェアのシャットダウンに限らず、スタンバイやスリープ、休止状態等の省電力モードへの移行を含むものとする。

[0057] ストレージ管理部26は、ストレージ123(123-i、i=1~z)のアクセスコントロールを行う。アクセスコントロールには、FC(Fibre Channel)スイッチ等によるゾーニング(zoning)を含む。ゾーニングとは、FCスイッチが提供する応用的な機能の一つで、デバイスやポート毎に「Read/Write可」、「Read only」、「不可視」といったアクセス制御を提供する機能である。

- [0058] 仮想マシン管理部 27 は、各サーバ 122 (122-i、i=1~y) 上の仮想マシンモニタ 1222 の管理やマイグレーションを行う。仮想マシン管理部 27 は、仮想マシンモニタ 1222 の情報を保持する。また、仮想マシン管理部 27 は、仮想マシンモニタ 1222 上で動作する仮想マシン 1221 (1221-i、i=1~k) 及び仮想スイッチの情報を保持したり、仮想マシンモニタ 1222 から適宜入手したりする。
- [0059] ここでは、管理装置 20 の例として、PC (パソコン)、シンククライアントサーバ、ワークステーション、メインフレーム、スーパーコンピュータ等の計算機を想定している。
- [0060] この場合、統合管理部 21、設備管理部 22、OFC 22、サービス管理部 24、サーバ管理部 25、ストレージ管理部 26、及び仮想マシン管理部 27 は、プログラムに基づいて駆動し所定の処理を実行するプロセッサと、当該プログラムや各種データを記憶するメモリと、ネットワークを介して外部と通信するためのネットワークインタフェースによって実現される。
- [0061] 上記のプロセッサの例として、CPU (Central Processing Unit)、マイクロプロセッサ (microprocessor)、マイクロコントローラ、或いは、同様の機能を有する半導体集積回路 (Integrated Circuit (IC)) 等が考えられる。
- [0062] 上記のメモリの例として、RAM (Random Access Memory)、ROM (Read Only Memory)、EEPROM (Electrically Erasable and Programmable Read Only Memory) やフラッシュメモリ等の半導体記憶装置、HDD (Hard Disk Drive) や SSD (Solid State Drive) 等の補助記憶装置、又は、DVD (Digital Versatile Disk) 等のリムーバブルディスクや、SDメモリカード (Secure Digital memory card) 等の記憶媒体 (メディア) 等が考えられる。
- [0063] 上記のネットワークインタフェースの例として、ネットワーク通信に対応

した基板（マザーボードやI/Oボード）等の半導体集積回路、NIC（Network Interface Card）等のネットワークアダプタや同様の拡張カード、アンテナ等の通信装置、接続口（コネクタ）等の通信ポート等が考えられる。

[0064] また、上記のネットワークの例として、インターネット、LAN（Local Area Network）、無線LAN（Wireless LAN）、WAN（Wide Area Network）、バックボーン（Backbone）、ケーブルテレビ（CATV）回線、固定電話網、携帯電話網、WiMAX、3G（第3世代携帯電話）、専用線（lease line）、IrDA（Infrared Data Association）、Bluetooth（登録商標）、シリアル通信回線、データバス等が考えられる。

[0065] なお、統合管理部21、設備管理部22、OFC22、サービス管理部24、サーバ管理部25、ストレージ管理部26、及び仮想マシン管理部27は、それぞれ独立した計算機でも良い。

[0066] 但し、実際には、これらの例に限定されない。

[0067] <サービスとフローの関係>

仮想マシン1221（1221-i、i=1~k）にサービスを要求するために、ユーザマシンがサーバに接続する毎にフローが発生する。すなわち、フローは、コネクションに相当する。

[0068] ここでは、フローは、TCP/IP（Transmission Control Protocol/Internet Protocol）パケットのアドレス領域のうち、デスティネーションアドレス（Destination Address：宛先アドレス）、ソースアドレス（Source Address：送信元アドレス）、デスティネーションポート番号、ソースポート番号のいずれか又は全てを用いた様々な組み合わせにより定義され、区別可能である。なお、上記のアドレスには、MACアドレスやIPアドレスを含むものとする。

[0069] 複数のサーバで同一のサービスを提供している場合、デスティネーションIPアドレス（宛先IPアドレス）でサービスとフローの対応付けができる。この場合、デスティネーションIPアドレスを基にフローを集計することで、フロー数からサービス数を算出することが可能になる。サービス数とは、同時に存在するフロー数であり、サービスによる負荷（サービス負荷）を示す値である。

[0070] また、サービス毎に使用するポート番号が異なる場合、デスティネーションポート番号でサービスとフローの対応付けができる。この場合、デスティネーションポート番号が同一であれば、同一のサービスであると判断することで、デスティネーションポート番号でそのサービスのフロー数をカウントし、サービス数を算出できる。

[0071] <データ構成>

次に、本発明の省電力化システムにおいて使用されるデータの所在及び詳細について説明する。

[0072] 以降、次のように記載を簡略化する。

- ・「ラック」は、ラック12（12-i、i=1~n）を示す。
- ・「スイッチ」は、OFS121（121-i、i=1~x）、又は仮想マシンモニタ1222の仮想スイッチを示す。
- ・「OFS」は、OFS121（121-i、i=1~x）を示す。
- ・「サーバ」は、サーバ122（122-i、i=1~y）を示す。
- ・「VM」は、仮想マシン1221（1221-i、i=1~k）を示す。
- ・「ストレージ」は、ストレージ123（123-i、i=1~z）を示す。

[0073] 統合管理部21は、室内設備情報と、ラック情報と、ストレージ接続情報と、サービス論理構成情報を保持する。

[0074] 室内設備情報は、空調情報と、照明情報と、電源情報とを含む。

[0075] 空調情報は、空調IDと、電力情報と、稼動状態情報と、関連ラック情報

を含む。空調IDは、空調設備111の識別情報を示す。電力情報は、この空調IDに対応する空調設備111の停止（又は調節）時に削減できる電力量（wH）を示す。稼動状態情報は、この空調IDに対応する空調設備111の空調グループに属するラック12（12-i、i=1~n）に収納された電子機器の現在の稼動状態（稼動／非稼動）を示す。関連ラック情報は、ラックIDに関する情報を含む。このラックIDは、この空調IDに対応する空調設備111の空調グループに属するラック12（12-i、i=1~n）の識別情報を示す。

[0076] 照明情報は、照明IDと、電力情報と、稼動状態情報と、関連ラック情報を含む。照明IDは、照明設備112の識別情報を示す。電力情報は、この照明IDに対応する照明設備112の停止（又は調節）時に削減できる電力量（wH）を示す。稼動状態情報は、この照明IDに対応する照明設備112の照明グループに属するラック12（12-i、i=1~n）に収納された電子機器の現在の稼動状態（稼動／非稼動）を示す。関連ラック情報は、ラックIDに関する情報を含む。このラックIDは、この照明IDに対応する照明設備112の照明グループに属するラック12（12-i、i=1~n）の識別情報を示す。

[0077] 電源情報は、電源IDと、電力情報と、稼動状態情報と、関連ラック情報を含む。電源IDは、電源設備113の識別情報を示す。電力情報は、この電源IDに対応する電源設備113の停止（又は調節）時に削減できる電力量（wH）を示す。稼動状態情報は、この電源IDに対応する電源設備113の電源グループに属するラック12（12-i、i=1~n）に収納された電子機器の現在の稼動状態（稼動／非稼動）を示す。関連ラック情報は、ラックIDに関する情報を含む。このラックIDは、この電源IDに対応する電源設備113の電源グループに属するラック12（12-i、i=1~n）の識別情報を示す。

[0078] なお、ここでは、説明の簡略化のため、空調設備111、照明設備112、電源設備113は、いずれもサーバールーム10内における同一の領域を対

象としているものとする。すなわち、空調グループ、照明グループ、電源グループは、同一のグループとする。但し、実際には、この例に限定されない。

[0079] ラック情報は、同一のラックに搭載された電子機器の情報を示す。ここでは、ラック情報は、ラックIDと、稼動状態情報と、組み込みノード情報を含む。このラックIDは、ラック12（12-i、i=1~n）の識別情報を示す。稼動状態情報は、このラックIDに対応するラック12（12-i、i=1~n）に収納された電子機器の現在の稼動状態（稼動/非稼動）を示す。組み込みノード情報は、ノードタイプと、ノードIDと、稼動状態情報を含む。ノードタイプは、当該ラック12（12-i、i=1~n）に収納された電子機器の種別（スイッチ/サーバ/ストレージ等）を示す。ノードIDは、当該ラック12（12-i、i=1~n）に収納された電子機器の識別情報を示す。例えば、ノードIDは、スイッチIDと、サーバIDと、ストレージIDのいずれかを示す。稼動状態情報は、当該ラック12（12-i、i=1~n）に収納された電子機器の稼動状態（稼動/非稼動）を示す。

[0080] ストレージ接続情報は、ストレージと通信可能なサーバ群の情報である。ここでは、ストレージ接続情報は、ストレージIDと、接続可能サーバ情報を含む。ストレージIDは、ストレージ123（123-i、i=1~z）の識別情報を示す。接続可能サーバ情報は、当該ストレージIDに対応するストレージと通信可能なサーバの識別情報（MACアドレス等）を示す。

[0081] サービス論理構成情報は、サービスを提供するために必要なノードの構成情報である。例えば、サービス論理構成情報は、接続する契機となる初めての packets（first packet）のルールと、構成するノード量に関する情報と、使用するストレージの情報（ストレージID）と、使用するノード群の情報（ノードID：VMが稼動しているサーバのID）に関する情報を含む。例えば、構成するノード量に関する情報は、ノードの詳細（種別：VM、ノード数）と、VMあたりの提供サービス量と、VMあたりの占

有物理リソース量（メモリ、CPU等）に関する情報を含む。

- [0082] OFC23は、トポロジー情報を保持する。ここで、トポロジー情報は、ネットワークのリンクのつながりを表す地図情報であり、フロー定義情報に該当する。
- [0083] トポロジー情報は、OFS群、サーバ群、外部ネットワーク等の接続状況に関する情報を含む。トポロジー情報は、監視対象であるスイッチ（OFS、仮想スイッチ等）と接続されるもののセット群で構成され、監視対象であるスイッチ毎に保持される情報である。
- [0084] ここでは、トポロジー情報は、スイッチIDと、稼動状態情報と、ポート数と、各ポートの接続情報に関する情報を含む。スイッチIDは、監視対象であるスイッチの識別情報を示す。稼動状態情報は、監視対象であるスイッチの稼動状態（稼動／非稼動）を示す。ポート数は、監視対象であるスイッチの保有するポートの数を示す。各ポートの接続情報は、接続種別情報と、接続先情報を含む。接続種別情報は、監視対象であるスイッチの接続先の種別（スイッチ／サーバ／ストレージ／外部ネットワーク等）に関する情報を含む。接続先情報は、監視対象であるスイッチの接続先の識別情報を示す。ここでは、接続先情報は、接続種別がスイッチの場合にはスイッチIDを示す。また、接続種別がサーバやストレージの場合にはMACアドレスを示す。また、接続種別が外部ネットワークの場合には外部ネットワークIDを示す。
- [0085] サービス管理部24は、サービス数情報を保持する。
- [0086] サービス数情報は、提供数と、応答時間に関する情報を含む。提供数は、各サービスの単位時間当たりの提供数を示す。応答時間は、各サービスの応答時間の平均値や最大値を示す。
- [0087] サーバ管理部25は、サーバ負荷情報を保持する。
- [0088] サーバ負荷情報は、負荷状態と、使用可能リソース数に関する情報を含む。負荷状態は、各サーバの単位時間当たりの負荷状態を示す。使用可能リソース数は、各サーバの使用可能リソース数を示す。ここでは、「サーバの使

用可能リソース数 = サーバのリソース数 - 予約されたリソース数」
とする。

[0089] <サービスの基本処理>

図2を参照して、サービスの基本処理について説明する。

[0090] (1) ステップS1

外部ネットワークからOFS121(121-i、i=1~x)にパケットが到着する。すなわち、OFS121(121-i、i=1~x)は、外部ネットワークからパケットを受信する。

[0091] (2) ステップS2

OFS121(121-i、i=1~x)は、受信されたパケットが初めてのパケット(first packet)である場合、OFC23に、当該パケットを初めてのパケット(first packet)として通知する。

[0092] (3) ステップS3

OFC23は、サービス管理部24に、初めてのパケット(first packet)の到着を通知する。

[0093] (4) ステップS4

サービス管理部24は、通知された初めてのパケット(first packet)に該当するサービスを検索する。

[0094] (5) ステップS5

サービス管理部24は、OFC23に、初めてのパケット(first packet)を検出したOFS121(121-i、i=1~x)と発見されたサービスが使用するVMを端点として指定し、指定された端点の通信経路の作成を指示する。

[0095] (6) ステップS6

OFC23は、指定された端点の通信経路をトポロジー情報から算出する。このとき、未稼働状態のOFS121(121-i、i=1~x)は算出に用いない。

[0096] (7) ステップ S 7

OFC 23 は、算出された通信経路の各 OFS 121 (121-i、i=1~x) 用のフローの定義情報を作成する。

[0097] (8) ステップ S 8

OFC 23 は、作成されたフローの定義情報を、該当する OFS 121 (121-i、i=1~x) に設定する。

[0098] <情報の収集及び構成変更の契機>

図3を参照して、フローの定義情報を設定した後の情報の収集及び構成変更の契機について説明する。

[0099] (1) ステップ C 1

OFC 23 は、OFS 121 (121-i、i=1~x) から情報を収集し、フロー数をカウントし、各フローをサービス管理部 24 に通知する。

[0100] (2) ステップ C 2

サービス管理部 24 は、該当フローをサービス毎に分類し、サービス毎の単位時間当たりのサービス数として情報を保持する。

[0101] (3) ステップ C 3

サービス管理部 24 は、サービス数を監視し、サービス数の変化を契機として、以下のようにサービスの論理構成・物理構成の変更を行うように、仮想マシン管理部 27 に指示する。

[0102] (4) ステップ C 4

仮想マシン管理部 27 は、サービス数が所定の閾値以上となった場合、所定の閾値より小さくなるように当該サービス数を軽減するための処理を行う。なお、ここでいう所定の閾値は、上限閾値でも良い。例えば、「(サービス数) \geq (1つのVMで負担できるサービス数) \times (VM数)」となった場合、仮想マシン管理部 27 は、当該サーバにおいてサービスへ割り当てるVMを増加させる。すなわち、当該サーバ上で当該サービス用のVMの稼働処理を行う。

[0103] 或いは、仮想マシン管理部 27 は、サービスの必要CPUを増加させるた

め、サーバ中のVM数を減少させる。このとき、当該サーバから他のサーバへVMを分散させ、他のサーバでVMの稼働処理を行う。

[0104] 更に、必要であれば、仮想マシン管理部27は、他のサーバと比較して、当該サーバ上で稼働中のVMの数が最も少なくなった場合、当該サーバ上で稼働中の残りのVMを、他のサーバへ片寄せする。サーバ管理部25は、この片寄により当該サーバ上で稼働中のVMがなくなった場合、当該サーバを停止する。サービス数が均等になるまで、この処理を繰り返しても良い。

[0105] (5) ステップC5

仮想マシン管理部27は、サービス数が所定の閾値より小さくなった場合、所定の閾値以下の範囲内でサービス数を増大させるための処理を行う。なお、ここでいう所定の閾値は、下限閾値でも良い。例えば、「(サービス数) < (1つのVMで負担できるサービス数) × (VM数)」となった場合、仮想マシン管理部27は、当該サーバにおいてサービスからVMを削減させる。すなわち、当該サーバ上で当該サービス用のVMの停止処理を行う。

[0106] 或いは、仮想マシン管理部27は、サービスの必要CPUを削減させるため、サーバのVM数を増加させる。すなわち、他のサーバから当該サーバへVMを集約させ、他のサーバでVMの停止処理を行う。停止処理については、後述する「管理者による手動VM起動停止を契機とした省電力化」の停止処理を実施する。

[0107] 更に、必要であれば、仮想マシン管理部27は、稼働中のVMの数が最も少なくなったサーバ上で稼働中のVMを、当該サーバ又はリソースに余裕がある他のサーバへ片寄せする。サーバ管理部25は、この片寄により稼働中のVMがなくなったサーバを停止する。サービス数が均等になるまで、この処理を繰り返しても良い。

[0108] <稼働リソース最適化>

ここでは、VMを移動することにより更なる省電力化を目指す稼働リソース最適化処理を記述する。この処理は、VM停止処理の終了を契機として動

作する。なお、実際には、VMを停止せずに最適化のみを行っても構わない。その場合の契機としては、システム全体の負荷が閾値を超過した場合等が考えられる。なお、ここでいう閾値は、上限閾値と下限閾値で囲まれた一定の範囲に収まる値の集合でも良い。この場合、閾値を超過した場合とは、上限閾値と下限閾値で囲まれた一定の範囲に収まらない値であることを意味する。

[0109] 図4を参照して、稼動リソース最適化処理について説明する。

[0110] (1) ステップR1

統合管理部21は、以下の数式(1)を用いて、必要リソース量の算出を行う。

必要リソース量の算出

R_i : サービス*i*のVMの必要リソース量

V_i : サービス*i*のVM数

$$\sum_{i=1}^n (R_i \times V_i) \quad \dots(1)$$

[0111] (2) ステップR2

統合管理部21は、以下の数式(2)を用いて、必要サーバ数の算出を行う。

必要サーバ数の算出

R : サーバのリソース量

S : 必要サーバ数

$$S = \left\lceil \frac{\sum_{i=1}^n (R_i \times V_i)}{R} \right\rceil \quad \dots(2)$$

[0112] (3) ステップR3

統合管理部21は、停止サーバ数Tを算出する。停止サーバ数Tは、稼働サーバ数をCとすると、 $T=C-S$ となる。

[0113] (4) ステップR4

統合管理部21は、停止可能空調グループの検索を行う。説明の簡略化のため、便宜上、空調グループを対象とする場合を例に説明しているが、照明グループや電源グループを対象とする場合も基本的に同様である。すなわち、「空調」を、「照明」や「電源」と読み替えることも可能である。

[0114] 具体的には、統合管理部21は、T台以内のサーバを停止することで、空調グループ内の全ラックが停止する空調グループを検索する。従って、停止可能空調グループ内の全ラックで稼働中のサーバの合計台数は、T台以内であるものとする。統合管理部21は、停止可能空調グループがあった場合、当該空調グループを停止することを決定する。なお、統合管理部21は、停止可能空調グループがなかった場合、停止可能ラックグループの検索処理（ステップR6）へ移行する。

[0115] (5) ステップR5

統合管理部21は、サーバ管理部25及び仮想マシン管理部27に指示して、停止可能空調グループから停止しない他の空調グループへVMをマイグレーションする。統合管理部21は、停止可能空調グループの停止に伴い停止するサーバの台数をX台とした場合、現在の停止サーバ数TからXを減じ、残りの停止サーバ数を $T-X$ へ変更する。従って、残りの停止サーバ数 $T=$ 現在の停止サーバ数 T -停止するサーバの台数 X となる。

[0116] 具体的には、統合管理部21は、停止することを決定した空調グループ内のラックで稼働中のサーバ上の全VMについて、移動先サーバを決定する。移動先サーバの例としては、各VMが属するサービスの使用するストレージに接続できるサーバ、停止予定の空調グループのラックに属さないサーバ、及び、空調グループのOFSが停止してもサービスを構成するVM間及び外部ネットワークとの通信経路が確立できるVMが必要とするリソースを持つ

サーバ、のいずれかが考えられる。

[0117] 統合管理部 21 は、サーバ管理部 25 及び仮想マシン管理部 27 に指示して、決定した移動先サーバに各 VM をマイグレーションする。統合管理部 21 は、OFC 23 に指示して、ラック中の OFS を停止し、ラックグループ中の OFS を未稼働とする。統合管理部 21 は、ラックグループを未稼働とする。統合管理部 21 は、設備管理部 22 に指示して、空調を停止し、或いは、空調の程度（設定温度、風力、対象範囲、等）を調節し、空調グループを未稼働とする。

[0118] なお、統合管理部 21 は、停止可能空調グループの検索処理（ステップ R4）において停止可能空調グループを 1 つずつ検索する場合、停止可能空調グループの検索処理（ステップ R4）に戻り、他の停止可能空調グループの検索を行う。他の停止可能空調グループ内の全ラックで稼働中のサーバの合計台数は、残りの停止サーバ数以内であるものとする。

[0119] 以下、統合管理部 21 は、停止可能空調グループが発見できなくなるまで、同様の処理を繰り返す。統合管理部 21 は、停止可能空調グループが発見できなくなった後、まだ残りの停止サーバ数がある場合、停止可能ラックグループの検索処理（ステップ R6）へ移行する。

[0120] （6）ステップ R6

統合管理部 21 は、停止可能ラックグループの検索を行う。具体的には、統合管理部 21 は、T 台以内のサーバを停止することでラックが停止するラックグループを検索する。従って、停止可能ラックグループで稼働中のサーバの合計台数は、T 台以内であるものとする。統合管理部 21 は、停止可能ラックグループがあった場合、当該ラックグループを停止することを決定する。なお、統合管理部 21 は、停止するラックグループがなかった場合、停止可能サーバの検索処理（ステップ R8）へ移行する。

[0121] （7）ステップ R7

統合管理部 21 は、サーバ管理部 25 及び仮想マシン管理部 27 に指示して、停止対象ラックグループから停止しない他のラックグループへ VM をマ

イグレーションする。統合管理部 21 は、停止可能ラックグループの停止に伴い停止するサーバの台数を Y 台とした場合、現在の停止サーバ数 T から Y を減じ、残りの停止サーバ数を $T - Y$ へ変更する。従って、残りの停止サーバ数 $T =$ 現在の停止サーバ数 $T -$ 停止するサーバの台数 Y となる。

[0122] 具体的には、統合管理部 21 は、停止することを決定したラックグループで稼働中のサーバ上の全 VM について、移動先サーバを決定する。移動先サーバの例として、各 VM が属するサービスの使用するストレージに接続できるサーバ、停止予定のラックに属さないサーバ、ラックグループの O F S が停止してもサービスを構成する VM 間及び外部ネットワークとの通信経路が確立できる VM が必要とするリソースを持つサーバ、のいずれかが考えられる。

[0123] 統合管理部 21 は、サーバ管理部 25 及び仮想マシン管理部 27 に指示して、決定したサーバに各 VM をマイグレーションする。統合管理部 21 は、サーバ管理部 25 及び仮想マシン管理部 27 に指示して、サーバを停止し、ラックグループ中のサーバを未稼働とする。O F C 23 は、ラック中の O F S を停止し、ラックグループ中の O F S を未稼働とする。統合管理部 21 は、ラックグループを未稼働とする。

[0124] なお、統合管理部 21 は、停止可能ラックグループの検索処理（ステップ R6）において停止可能ラックグループを 1 つずつ検索する場合、停止可能ラックグループの検索処理（ステップ R6）に戻り、他の停止可能ラックグループの検索を行う。他の停止可能ラックグループ内の全ラックで稼働中のサーバの合計台数は、残りの停止サーバ数以内であるものとする。

[0125] 以下、統合管理部 21 は、停止可能ラックグループが発見できなくなるまで、同様の処理を繰り返す。統合管理部 21 は、停止可能ラックグループが発見できなくなった後、まだ残りの停止サーバ数がある場合、停止可能サーバの検索処理（ステップ R8）へ移行する。

[0126] （8）ステップ R8

統合管理部 21 は、停止可能サーバの検索を行う。具体的には、統合管理

部 2 1 は、ラック内の稼働サーバ数が最小となるラックグループを検索する。統合管理部 2 1 は、停止する T 台のサーバ群を決定する。従って、停止可能サーバの合計台数は、T 台以内であるものとする。統合管理部 2 1 は、停止可能サーバがあった場合、当該サーバを停止することを決定する。

[0127] なお、ラック内の稼働サーバ数が最小となるラックグループを検索する理由は、少ない台数のサーバを停止することで、ラック全体が停止可能となるラックを作り出し、その後（次回以降）の稼働リソース最適化処理で、ラック全体を停止できる可能性を高めるためである。

[0128] すなわち、この処理で発見された停止可能サーバを搭載したラックは、次回以降の稼働リソース最適化処理において、停止可能ラックグループの検索処理（ステップ R 6）で停止可能ラックグループとして選ばれる可能性が高くなる。

[0129] 更に、このラックを含む空調グループは、次回以降の稼働リソース最適化処理において、停止可能空調グループの検索処理（ステップ R 4）で停止可能空調グループとして選ばれる可能性が高くなる。

[0130] （9）ステップ R 9

統合管理部 2 1 は、サーバ管理部 2 5 及び仮想マシン管理部 2 7 に指示して、停止対象サーバから VM をマイグレーションする。具体的には、統合管理部 2 1 は、停止することを決定したサーバ上の全 VM について、移動先サーバを決定する。移動先サーバの例として、各 VM が属するサービスの使用するストレージに接続できるサーバ、停止予定以外のサーバ、VM が必要とするリソースを持つサーバ、等が考えられる。統合管理部 2 1 は、サーバ管理部 2 5 及び仮想マシン管理部 2 7 に指示して、決定したサーバに各 VM をマイグレーションする。統合管理部 2 1 は、サーバ管理部 2 5 及び仮想マシン管理部 2 7 に指示して、サーバを停止し、ラックグループ中のサーバの状態を未稼働とする。

[0131] <ラック間の接続に関する条件>

なお、ラック間の接続に関しては、下記の制限を満たすことが望ましい。

- ・ ラック間の相互接続を可能な限りなくす
- ・ 空調グループ単位で相互接続用スイッチを設ける
- ・ 空調グループ間の接続を常時稼働のスイッチラックで行う

[0132] <稼働リソース最適化（拡張）>

統合管理部 2 1 は、O F C 2 3 に指示して、下記処理を追加することで、稼働 O F S 数を更に減少させることができる。

[0133] 統合管理部 2 1 は、O F C 2 3 に指示して、稼働中の O F S から任意の O F S を停止させた場合のトポロジー情報を作成し、全サービスについて使用している VM 間及び外部ネットワークとの通信経路が確立した場合、該当 O F S を停止、該当 O F S を含むラックグループの該当 O F S の状態を未稼働とし、トポロジー情報の該当 O F S を未稼働に変更する。

[0134] 統合管理部 2 1 は、O F C 2 3 に指示して、O F S が停止した結果、ラックグループ内で稼働しているノードがなくなった場合、ラックグループを未稼働とする。統合管理部 2 1 は、ラックグループが未稼働となった結果、空調グループで稼働しているラックがなくなった場合、設備管理部 2 2 に指示して、空調を停止し、或いは、空調の程度（設定温度、風力、対象範囲、等）を調節し、空調グループの空調を未稼働とする。

[0135] <移動先サーバの決定処理>

図 5 を参照して、移動先サーバの決定処理について説明する。なお、統合管理部 2 1 は、ラック停止後のトポロジー情報を経路算出に使用することを前提とする。

[0136] （１）ステップ D 1

統合管理部 2 1 は、未使用リソースがあるサーバをリストアップする。

[0137] （２）ステップ D 2

統合管理部 2 1 は、マイグレーション対象の VM をリストアップする。

[0138] （３）ステップ D 3

統合管理部 2 1 は、リストアップされた VM の中から、未移動の VM を 1 つ選択する。

[0139] (4) ステップD4

統合管理部21は、リストアップされたサーバの中から、所定の条件に適合するサーバを選択する。具体的には、統合管理部21は、選択された未移動のVMが含まれるサービスが使用するストレージと接続可能であり、サービスの他のVM及び外部ネットワークと通信経路が確立できるサーバを選択する。

[0140] (5) ステップD5

統合管理部21は、選択されたサーバにVMが使用する未使用リソースがあるか確認する。統合管理部21は、選択されたサーバにVMが使用する未使用リソースがある場合、選択された未稼働のVMを該当サーバに仮割り当てし、同時に未使用リソースを予約する。統合管理部21は、移動先が決定していないVMがあれば、再度、未移動のVMを1つ選択する。

[0141] (6) ステップD6

統合管理部21は、選択されたサーバに仮割り当てされたVMのうち、他の未使用リソースがあるサーバに移動可能なVMを検索する。統合管理部21は、他の未使用リソースがあるサーバに移動可能なVMがある場合、VMの仮割り当て先を変更し、同時に使用リソースの開放と新規仮割り当て先に予約する。統合管理部21は、選択された未稼働のVMが仮割り当てされたサーバに仮割り当てし、同時に未使用リソースを予約する。統合管理部21は、移動先が決定していないVMがあれば、再度、未移動のVMを1つ選択する。ここでは、VMの移動先を仮割り当てで一旦予約し、その後、未使用リソースが最小になるように変更することを考えている。

[0142] (7) ステップD7

統合管理部21は、動作中のVMのうち、他の未使用リソースがあるサーバに移動可能なVMを検索する。統合管理部21は、他の未使用リソースがあるサーバに移動可能なものがある場合、VMを移動対象とし、仮割り当て先を設定し、同時に使用リソースの開放と新規仮割り当て先に予約する。統合管理部21は、選択された未移動のVMが仮割り当てされたサーバに仮割

り当てし、同時に未使用リソースを予約する。統合管理部 21 は、移動先が決定していない VM があれば、再度、未移動の VM を 1 つ選択する。統合管理部 21 は、他の未使用リソースがあるサーバに移動可能な VM がない場合、選択された未移動の VM を保留とする。

[0143] (8) ステップ D8

統合管理部 21 は、移動先が決定していない VM があれば、再度、未移動の VM を 1 つ選択する。

[0144] (9) ステップ D9

統合管理部 21 は、各 VM の仮割り当てしたサーバをマイグレーション先として決定する。

[0145] <管理者による手動 VM 起動停止を契機とした省電力化>

次に、「管理者による手動 VM 起動停止を契機とした省電力化」に係る処理について、構成変更の契機、起動処理、停止処理の 3 つに分けて説明する。

[0146] <構成変更の契機>

統合管理部 21 は、管理者により、サービス毎の VM 数が変更された場合、次のように処理する。例えば、統合管理部 21 は、管理者によりサービスに VM が追加された場合、必要なサーバや VM の稼動処理を行う。また、統合管理部 21 は、管理者によりサービスから VM が削減された場合、不要なサーバや VM の停止処理を行う。

[0147] <稼動処理>

図 6 を参照して、稼動処理について説明する。

[0148] (1) ステップ A1

統合管理部 21 は、管理者により、サービスと追加 VM 数が指定され、稼動処理の開始が指示される。統合管理部 21 は、稼動中のサーバの中から、条件に適合するサーバを検索する処理を開始する。

[0149] (2) ステップ A2

統合管理部 21 は、稼動中のサーバの中に、条件に適合するサーバがある

か確認する。

[0150] 具体的には、統合管理部 2 1 は、統合管理部 2 1 のサービス構成情報とストレージ接続情報を用い、サービスが使用するストレージに接続可能なサーバを検索する。統合管理部 2 1 は、ラックグループ情報を用い、サービスが使用するストレージに接続可能なサーバのうち、稼働中のサーバ群を検索する。

[0151] 更に、統合管理部 2 1 は、発見されたサーバ群のうち、指定されたサービスの VM の必要リソースが提供できるサーバ（使用可能リソース量が必要リソース以上のサーバ）を検索する。統合管理部 2 1 は、指定されたサービスの VM の必要リソースが提供できるサーバのうち、指定されたサービスの VM が割り当て可能なサーバを検索する。

[0152] (3) ステップ A 3

統合管理部 2 1 は、稼働中のサーバの中に、条件に適合するサーバがない場合、稼働中のラックの中に、未稼働のサーバがあるか確認する。

[0153] (4) ステップ A 4

統合管理部 2 1 は、稼働中のラックの中に、未稼働のサーバがない場合、稼働中の空調グループの中に、未稼働のラックがあるか確認する。

[0154] (5) ステップ A 5

統合管理部 2 1 は、稼働中の空調グループの中に、未稼働のラックがない場合、未稼働の空調グループを稼働させる。

[0155] 具体的には、統合管理部 2 1 は、設備管理部 2 2 に指示し、未稼働の空調を起動し、稼働させる。また、統合管理部 2 1 は、空調情報を参照し、この空調グループの空調の状態を稼働に変更する。

[0156] ここでは、説明の簡略化のため、便宜上、空調グループを対象とする場合を例に説明しているが、照明グループや電源グループを対象とする場合も基本的に同様である。すなわち、「空調」を、「照明」や「電源」と読み替えることも可能である。

[0157] (6) ステップ A 6

統合管理部 2 1 は、稼働中の空調グループの中に、未稼働のラックがある場合、未稼働のラックを稼働させる。

[0158] 具体的には、統合管理部 2 1 は、O F C 2 3 に指示して、未稼働のラック内の O F S を起動し、稼働させる。また、統合管理部 2 1 は、O F C 2 3 に指示して、トポロジー情報を参照し、稼働状態にしたラックグループ内の O F S が未稼働である場合、該当 O F S を稼働し、ラックグループ内の稼働した O F S の状態を稼働に変更し、ラックグループ内の稼働した O F S の状態を稼働に変更し、トポロジー情報中でも当該 O F S を稼働に変更する。また、統合管理部 2 1 は、ラック情報を参照し、このラックグループを稼働状態に変更する。

[0159] (7) ステップ A 7

統合管理部 2 1 は、稼働中のラックの中に、未稼働のサーバがある場合、未稼働のサーバを稼働させる。その後、統合管理部 2 1 は、稼働中のサーバの中から条件に適合するサーバを検索する処理（ステップ A 1）に戻り、同様の処理を繰り返す。

[0160] (8) ステップ A 8

統合管理部 2 1 は、稼働中のサーバの中に、条件に適合するサーバがある場合、条件に適合するサーバに、指定されたサービスの VM を割り当てる。

[0161] 具体的には、統合管理部 2 1 は、仮想マシン管理部 2 7 に指示して、条件に適合するサーバで、指定されたサービスの VM を稼働する。すなわち、条件に適合するサーバに、指定されたサービスの VM が追加される。統合管理部 2 1 は、仮想マシン管理部 2 7 に指示して、追加された VM の情報をサービス構成情報に記録し保持する。統合管理部 2 1 は、仮想マシン管理部 2 7 に指示して、条件に適合するサーバの使用可能リソース量を、稼働させた VM が使用する分だけ削減する。

[0162] (9) ステップ A 9

統合管理部 2 1 は、未割り当ての VM があるか確認する。すなわち、統合管理部 2 1 は、条件に適合するサーバに、指定された数の VM を全て割り当

て、条件に適合するサーバ上で指定された数のVMを稼動することができたか確認する。

[0163] 具体的には、統合管理部21は、仮想マシン管理部27に指示して、条件に適合するサーバ上で指定された数のVMを稼動できた場合、指定された数のVMを稼動できた場合、VM間の通信が可能かどうか確認する。統合管理部21は、OFC23に指示して、稼動しているOFSとサービスを構成するVMのみのトポロジー情報を作成して、経路探索を行う。統合管理部21は、OFC23に指示して、経路が確立できた場合、一連の処理を終了する。また、統合管理部21は、仮想マシン管理部27に指示して、条件に適合するサーバ上で指定された数のVMを稼動できていない場合、稼動中のサーバの中から条件に適合するサーバを検索する処理（ステップA1）に戻り、同様の処理を繰り返す。

[0164] <停止>

図7を参照して、停止処理について説明する。

[0165] (1) ステップP1

管理者より、統合管理部21に、サービスと停止VM数が指定され、稼動処理の開始が指示される。これにより、統合管理部21は、処理を開始する。

[0166] (2) ステップP2

統合管理部21は、サービス構成情報とラック情報と空調情報から指定されたVM群のうち、停止するVMを選択し、決定する。

[0167] (選択の優先度)

VMの選択の優先度の例として、以下の順序に従って選択することが考えられる。

1. 当該VMを停止することで、サーバが停止でき、ラックが停止でき、空調も停止できるVM

2. 当該VMを停止することで、サーバが停止でき、ラックが停止できるVM

3. 当該VMを停止することで、サーバが停止できるVM

4. 当該VMを停止することで、サーバ中の稼動VM数が最小となるVM

[0168] (3) ステップP3

統合管理部21は、仮想マシン管理部27に指示して、片寄を行い停止するVMに処理がこないようにする。

[0169] (4) ステップP4

統合管理部21は、仮想マシン管理部27に指示して、VMで全処理が終了したら、VMを停止し、サービスの稼動VM数を減少させる。

[0170] (5) ステップP5

統合管理部21は、サーバ管理部25に指示して、該当サーバの削除したVMの分使用可能リソースを増加させる。

[0171] (6) ステップP6

統合管理部21は、サーバ管理部25に指示して、移動VMがなくなったサーバが発生した場合、サーバを停止し、該当ラックグループの該当サーバの状態を未稼働とする。

[0172] (7) ステップP7

統合管理部21は、サーバ停止の波及効果を調査する。

[0173] 具体的には、統合管理部21は、OFC23に指示して、サーバが停止した結果、ラックグループ内に移動サーバがなくなった場合、ラック内のOFSを停止しても稼働中のサービスに影響がないか（通信経路が算出できないパターンが存在しないか）を確認し、影響がない場合OFSを停止し、該当ラックグループの該当OFSの状態を未稼働とし、トポロジー情報から該当OFSを未稼働に変更する。

[0174] 統合管理部21は、OFC23に指示して、OFSが停止した結果、ラックグループ内で稼働しているノードがなくなった場合、ラックグループを未稼働とする。

[0175] 設備管理部22は、ラックグループが未稼働となった結果、空調グループで稼働しているラックがなくなった場合、空調を停止し、或いは、空調の程

度（設定温度、風力、対象範囲、等）を調節し、空調グループの空調を未稼働とする。

[0176] <停止処理（拡張）>

更に、他の停止処理について検討する。

統合管理部 21 は、OFC 23 に指示して、停止処理の終了後に他の O F S を未稼働にできないかを調査することで、未稼働リソースを増やすことも可能である。例えば、以下の 3 つの方法が考えられる。

[0177] 1. 統合管理部 21 は、OFC 23 に指示して、稼働中の O F S から任意の O F S を停止させた場合のトポロジー情報を作成し、全サービスについて私用している VM 間及び外部ネットワークとの通信経路が確立した場合、該当 O F S を停止、該当 O F S を含むラックグループの該当 O F S の状態を未稼働とし、トポロジー情報から該当 O F S を未稼働に変更する。

[0178] 2. 統合管理部 21 は、OFC 23 に指示して、O F S が停止した結果、ラックグループ内で稼働しているノードがなくなった場合、ラックグループを未稼働とする。

[0179] 3. 設備管理部 22 は、ラックグループが未稼働となった結果、空調グループで稼働しているラックがなくなった場合、空調を停止し、或いは、空調の程度（設定温度、風力、対象範囲、等）を調節し、空調グループの空調を未稼働とする。

[0180] <ストレージのアクセス制限>

なお、本発明において、ストレージのアクセス制限を考慮すると、統合管理部 21 は、VM 起動（稼働）時、或いは、VM マイグレーション時に、ストレージ管理部 26 と連携し、ストレージに適切なアクセス制限を設定する必要があると考えられる。

[0181] <負荷判断（拡張）>

また、本発明において、上記のサービス数の代わりに、サービスに応じたネットワークにおける通信負荷やサーバにおける処理負荷をサービス負荷とし、このサービス負荷の変化を契機として、サーバの論理構成・物理構成の

変更を行うようにすることも考えられる。

[0182] 例えば、統合管理部 21 は、以下の数式 (3) を用いて、この場合のサービス負荷の算出を行う。ここでは、当該サーバ上では、 n 台の仮想マシン (VM) が稼動しており、1 台の仮想マシン (VM) につき、1 つのサービスを提供しているものとする。

サービス負荷の算出

A_i : サービス i を行う VM_i についてのフロー数

B_i : サービス i を行う VM_i についての CPU 処理負荷

C_i : サービス i を行う VM_i についての

フロー単位のサービスによる通信負荷

D : サーバの最大通信能力

E : 最大 CPU 処理能力

F : サービス負荷

$$F = \max \left(\frac{\sum_{i=1}^n C_i A_i}{D}, \frac{\sum_{i=1}^n B_i}{E} \right) \quad \dots (3)$$

[0183] C_i に関しては、各サービス毎のフローの統計情報から算出する方法がある。

[0184] 上記の数式 (3) では、 $F = 1$ になると、サーバの能力限界まで使用したことになる。

[0185] なお、通信負荷と CPU 負荷の両方が 1 に近づくように VM を片寄してサーバに配置すると、リソースの効率が良くなると考えられる。

[0186] <付記>

上記の実施形態の一部又は全部は、以下の付記のように記載することも可能である。但し、実際には、以下の記載例に限定されない。

[0187] (付記 1)

複数のサーバ上でそれぞれ稼動する仮想マシンにより提供されるサービスの通信情報を監視するステップと、

前記通信情報から求められた仮想マシン毎のサービス負荷を基に、仮想マシンを他のサーバ上へ移動させ、稼動する仮想マシンが存在しなくなったサーバを未稼動状態にするステップと、

前記サービスの通信パケットの転送を行うスイッチに対してフロー単位の経路情報を設定するステップと、

仮想マシンの移動に従って前記スイッチに対して設定した経路情報を変更するステップと

をコンピュータに実行させるための省電力化プログラムを記憶した記憶媒体。

[0188] (付記2)

付記1に記載の記憶媒体であって、

前記サービスに応じたネットワークにおける通信負荷及び前記サービスに応じたサーバにおける処理負荷に基づいて、前記サービス負荷を求めるステップ

を更にコンピュータに実行させるための省電力化プログラムを記憶した記憶媒体。

[0189] (付記3)

付記1又は2に記載の記憶媒体であって、

前記仮想マシンの移動に伴って前記スイッチの経路情報が変更された結果、稼動不要なスイッチが存在すれば、当該スイッチを未稼動状態にするステップ

を更にコンピュータに実行させるための省電力化プログラムを記憶した記憶媒体。

[0190] (付記4)

付記1乃至3のいずれかに記載の記憶媒体であって、

前記スイッチ及び前記複数のサーバを格納したラック内のサーバ上で稼動

する仮想マシンを他のラック内のサーバ上へ移動させるステップと、

当該ラック外のスイッチを経由するように前記スイッチに設定された経路情報を変更するステップと、

前記スイッチ及び前記複数のサーバをコンピュータに実行させるための省電力化プログラムを記憶した当該ラック全体を未稼働状態にするステップとを更にコンピュータに実行させるための省電力化プログラムを記憶した記憶媒体。

[0191] (付記5)

付記4に記載の記憶媒体であって、

ラックの周辺環境状態を調整する設備を制御し、前記未稼働状態となったラックの周辺環境状態の調整を変化または停止させるステップ

を更にコンピュータに実行させるための省電力化プログラムを記憶した記憶媒体。

[0192] (付記6)

付記5に記載の記憶媒体であって、

前記設備による周辺環境状態の制御対象となっているラック群単位で未稼働状態にするステップが可能な対象を探索し、次に、ラック単位、サーバ単位の順に未稼働状態にするステップが可能な対象を順に探索するステップ

を更にコンピュータに実行させるための省電力化プログラムを記憶した記憶媒体。

[0193] (付記7)

付記4又は5に記載の記憶媒体であって、

空調の対象となる空調グループに属するラックを冷却する空調設備を制御し、前記空調グループに属する全てのラックが未稼働状態になった場合、前記空調設備による冷却を調節または停止させるステップと、

照明の対象となる照明グループに属するラックを照らす照明設備を制御し、前記照明グループに属する全てのラックが未稼働状態になった場合、前記照明設備による照明を調節または停止させるステップと、

電源の対象となる電源グループに属するラックに配電する電源設備を制御し、前記電源グループに属する全てのラックが未稼働状態になった場合、前記電源設備による配電を調節または停止させるステップとを更にコンピュータに実行させるための省電力化プログラムを記憶した記憶媒体。

[0194] 以上、本発明の実施形態を詳述してきたが、実際には、上記の実施形態に限られるものではなく、本発明の要旨を逸脱しない範囲の変更があっても本発明に含まれる。

[0195] なお、本出願は、日本出願番号2009-233366に基づく優先権を主張するものであり、日本出願番号2009-233366における開示内容は引用により本出願に組み込まれる。

請求の範囲

- [請求項1] 複数のサーバ上でそれぞれ稼動する仮想マシンにより提供されるサービスの通信情報を監視し、前記通信情報から求められた仮想マシン毎のサービス負荷を基に、仮想マシンを他のサーバ上へ移動させ、稼動する仮想マシンが存在しなくなったサーバを未稼動状態にする管理装置と、
- 前記サービスの通信パケットの転送を行うスイッチに対してフロー単位の経路情報を設定するコントローラと
- を具備し、
- 前記コントローラは、仮想マシンの移動に従って前記スイッチに対して設定した経路情報を変更する
- 省電力化システム。
- [請求項2] 請求項1に記載の省電力化システムであって、
- 前記管理装置は、前記サービスに応じたネットワークにおける通信負荷及び前記サービスに応じたサーバにおける処理負荷に基づいて、前記サービス負荷を求める
- 省電力化システム。
- [請求項3] 請求項1又は2に記載の省電力化システムであって、
- 前記管理装置は、仮想マシンの移動に伴って前記スイッチの経路情報が変更された結果、稼動不要なスイッチが存在すれば、当該スイッチを未稼動状態にする
- 省電力化システム。
- [請求項4] 請求項1乃至3のいずれか一項に記載の省電力化システムであって、
- 前記スイッチ及び前記複数のサーバは、ラック内に格納され、
- 前記管理装置は、当該ラック内のサーバ上で稼動する仮想マシンを他のラック内のサーバ上へ移動させ、
- 前記コントローラは、当該ラック外のスイッチを経由するように前

記スイッチに設定された経路情報を変更し、

前記管理装置は、前記スイッチ及び前記複数のサーバを含む当該ラック全体を未稼働状態にする

省電力化システム。

[請求項5]

請求項4に記載の省電力化システムであって、

前記管理装置は、ラックの周辺環境状態を調整する設備を制御し、前記未稼働状態となったラックの周辺環境状態の調整を変化または停止させる

省電力化システム。

[請求項6]

請求項5に記載の省電力化システムであって、

前記管理装置は、前記設備による周辺環境状態の制御対象となっているラック群単位で未稼働状態にすることが可能な対象を探索し、次に、ラック単位、サーバ単位の順に未稼働状態にすることが可能な対象を順に探索する

省電力化システム。

[請求項7]

請求項5又は6に記載の省電力化システムであって、

前記管理装置は、

空調の対象となる空調グループに属するラックを冷却する空調設備を制御し、前記空調グループに属する全てのラックが未稼働状態になった場合、前記空調設備による冷却を調節または停止させる空調設備管理部と、

照明の対象となる照明グループに属するラックを照らす照明設備を制御し、前記照明グループに属する全てのラックが未稼働状態になった場合、前記照明設備による照明を調節または停止させる照明設備管理部と、

電源の対象となる電源グループに属するラックに配電する電源設備を制御し、前記電源グループに属する全てのラックが未稼働状態になった場合、前記電源設備による配電を調節または停止させる電源設備

管理部と

を具備する

省電力化システム。

[請求項8] 請求項 1 乃至 7 のいずれか一項に記載の省電力化システムで、管理装置及びコントローラのうち少なくとも一方として使用されるコンピュータ。

[請求項9] コンピュータにより実施される省電力化方法であって、
複数のサーバ上でそれぞれ稼動する仮想マシンにより提供されるサービスの通信情報を監視することと、

前記通信情報から求められた仮想マシン毎のサービス負荷を基に、仮想マシンを他のサーバ上へ移動させ、稼動する仮想マシンが存在しなくなったサーバを未稼動状態にすることと、

前記サービスの通信パケットの転送を行うスイッチに対してフロー単位の経路情報を設定することと、

仮想マシンの移動に従って前記スイッチに対して設定した経路情報を変更することと

を含む

省電力化方法。

[請求項10] 請求項 9 に記載の省電力化方法であって、
前記サービスに応じたネットワークにおける通信負荷及び前記サービスに応じたサーバにおける処理負荷に基づいて、前記サービス負荷を求めること

を更に含む

省電力化方法。

[請求項11] 請求項 9 又は 10 に記載の省電力化方法であって、
前記仮想マシンの移動に伴って前記スイッチの経路情報が変更された結果、稼動不要なスイッチが存在すれば、当該スイッチを未稼動状態にすること

を更に含む

省電力化方法。

[請求項12]

請求項9乃至11のいずれか一項に記載の省電力化方法であって、
前記スイッチ及び前記複数のサーバを格納したラック内のサーバ上で稼動する仮想マシンを他のラック内のサーバ上へ移動させることと、

当該ラック外のスイッチを経由するように前記スイッチに設定された経路情報を変更することと、

前記スイッチ及び前記複数のサーバを含む当該ラック全体を未稼動状態にすることと

を更に含む

省電力化方法。

[請求項13]

請求項12に記載の省電力化方法であって、

ラックの周辺環境状態を調整する設備を制御し、前記未稼動状態となったラックの周辺環境状態の調整を変化または停止させること

を更に含む

省電力化方法。

[請求項14]

請求項13に記載の省電力化方法であって、

前記設備による周辺環境状態の制御対象となっているラック群単位で未稼動状態にすることが可能な対象を探索し、次に、ラック単位、サーバ単位の順に未稼動状態にすることが可能な対象を順に探索すること

を更に含む

省電力化方法。

[請求項15]

請求項13又は14に記載の省電力化方法であって、

空調の対象となる空調グループに属するラックを冷却する空調設備を制御し、前記空調グループに属する全てのラックが未稼働状態になった場合、前記空調設備による冷却を調節または停止させることと、

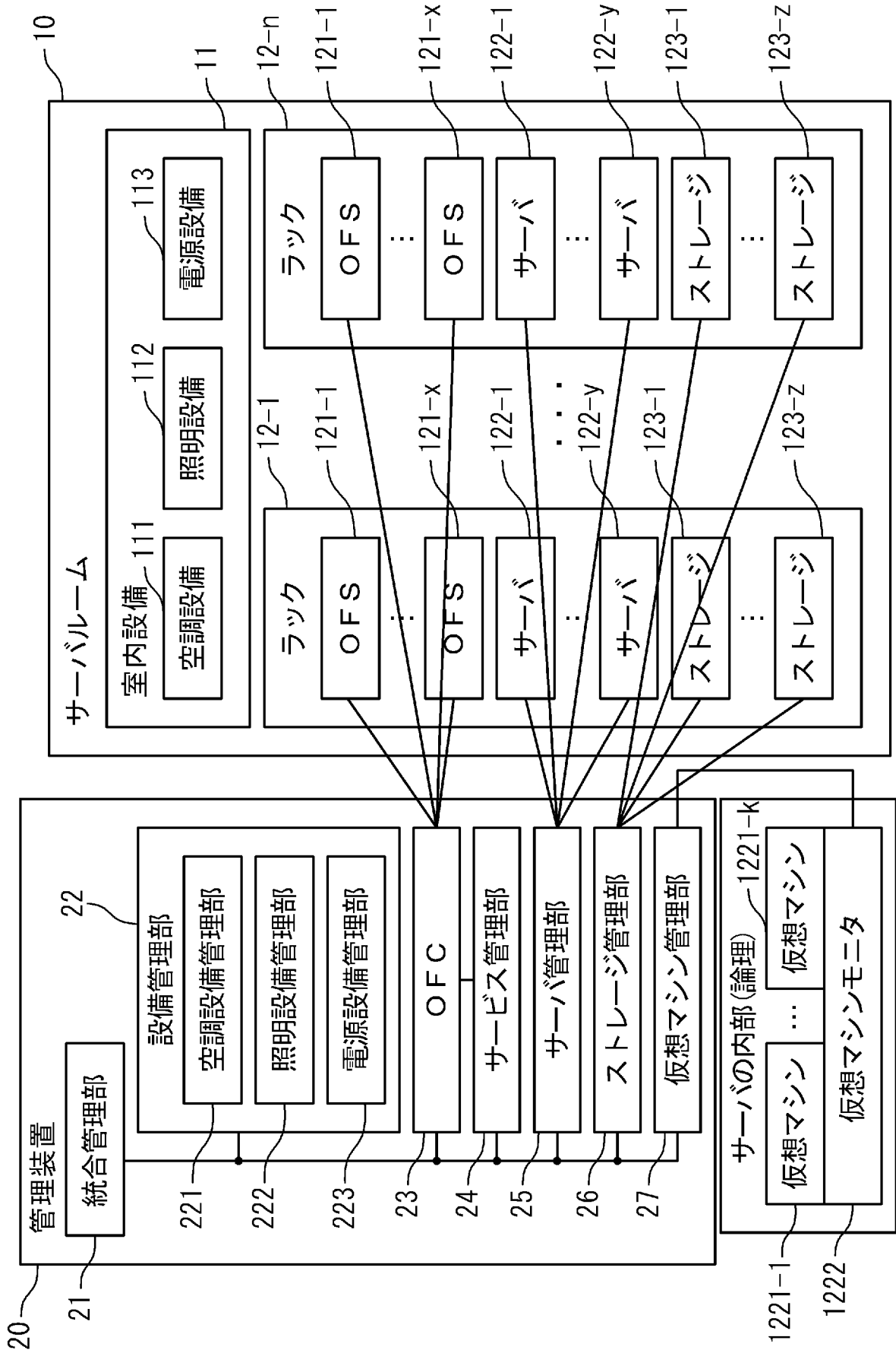
照明の対象となる照明グループに属するラックを照らす照明設備を制御し、前記照明グループに属する全てのラックが未稼働状態になった場合、前記照明設備による照明を調節または停止させることと、

電源の対象となる電源グループに属するラックに配電する電源設備を制御し、前記電源グループに属する全てのラックが未稼働状態になった場合、前記電源設備による配電を調節または停止させることとを更に含む

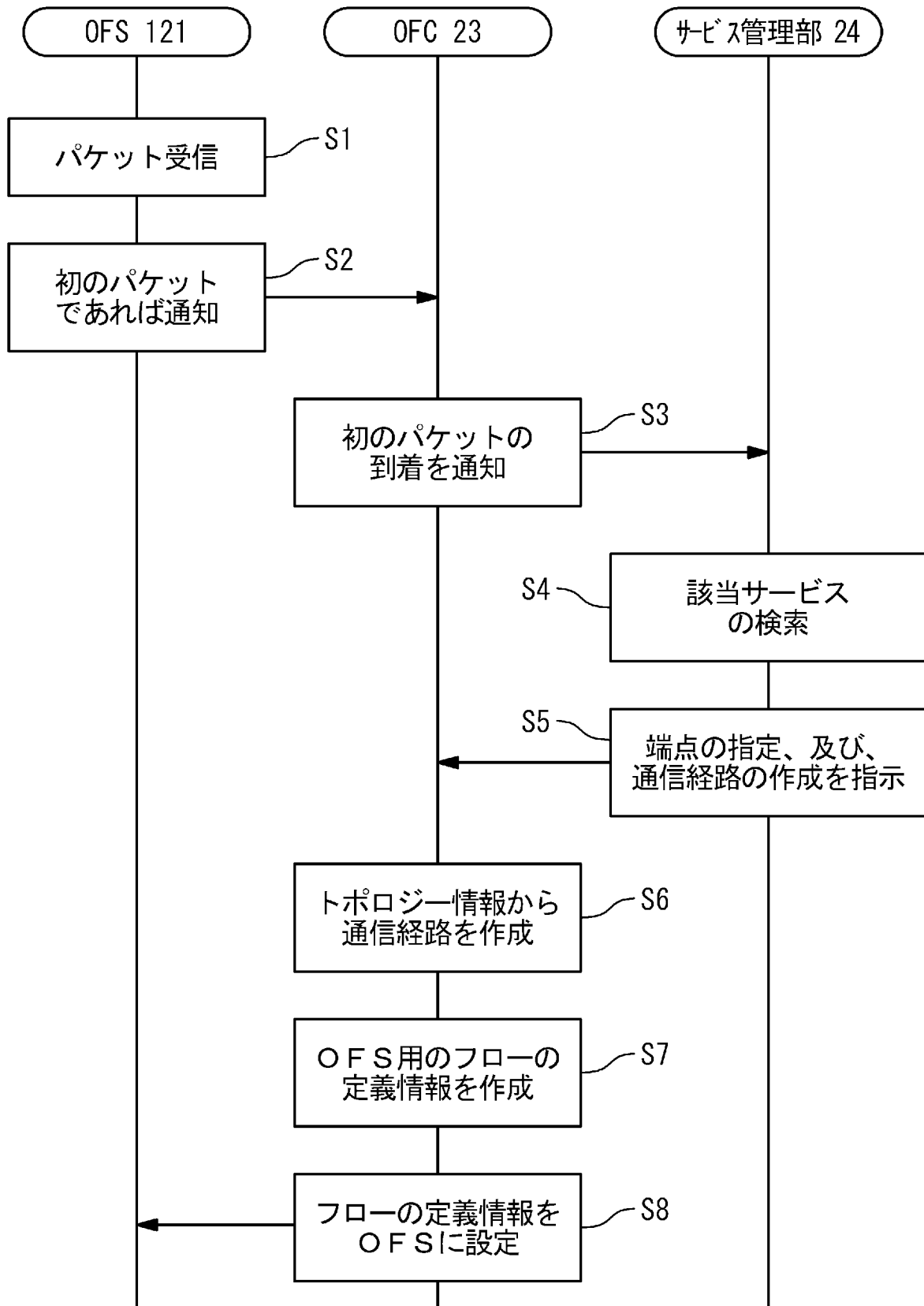
省電力化方法。

[請求項16] 請求項9乃至15のいずれか一項に記載の省電力化方法をコンピュータに実行させるための省電力化プログラムを記憶した記憶媒体。

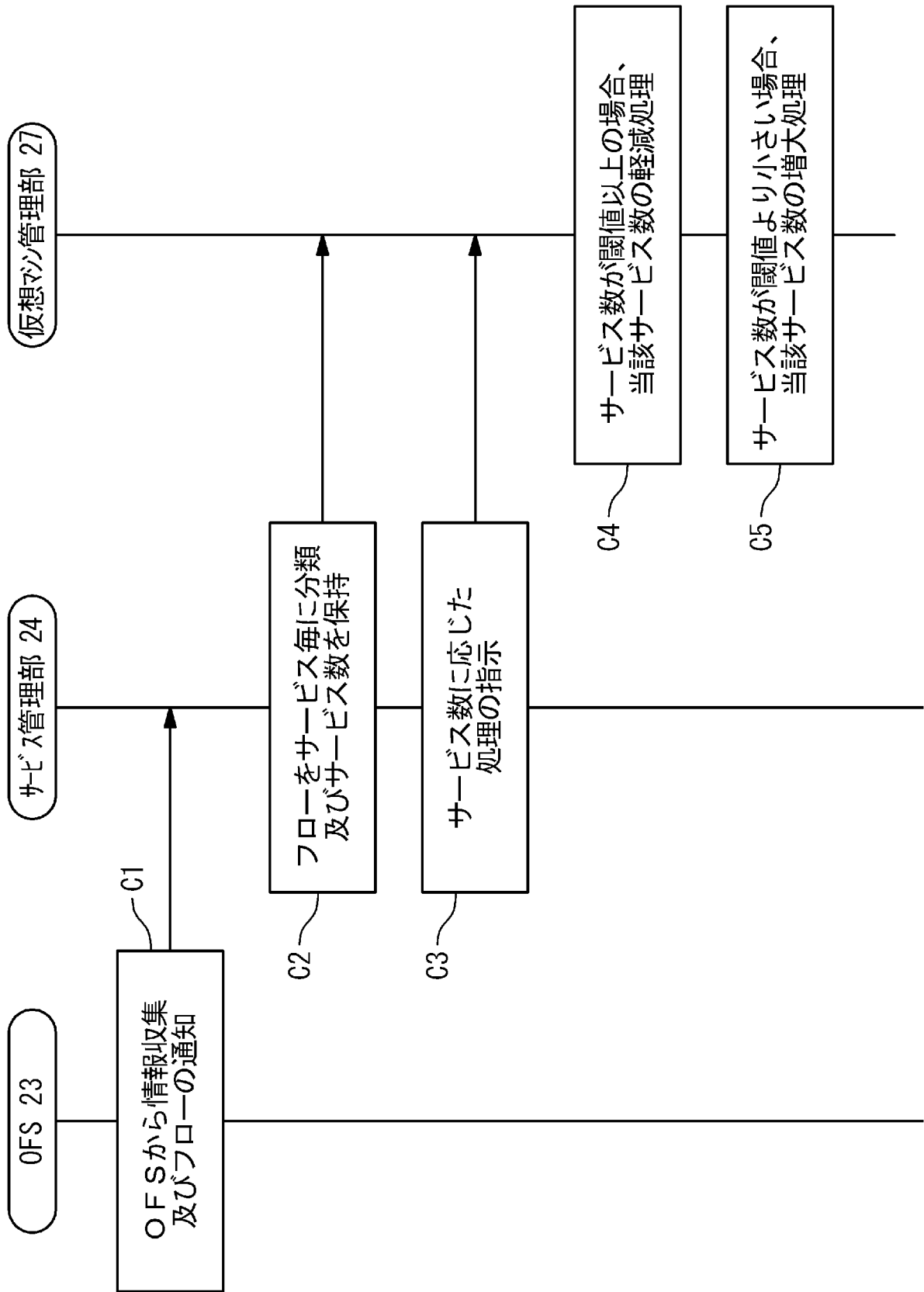
[図1]



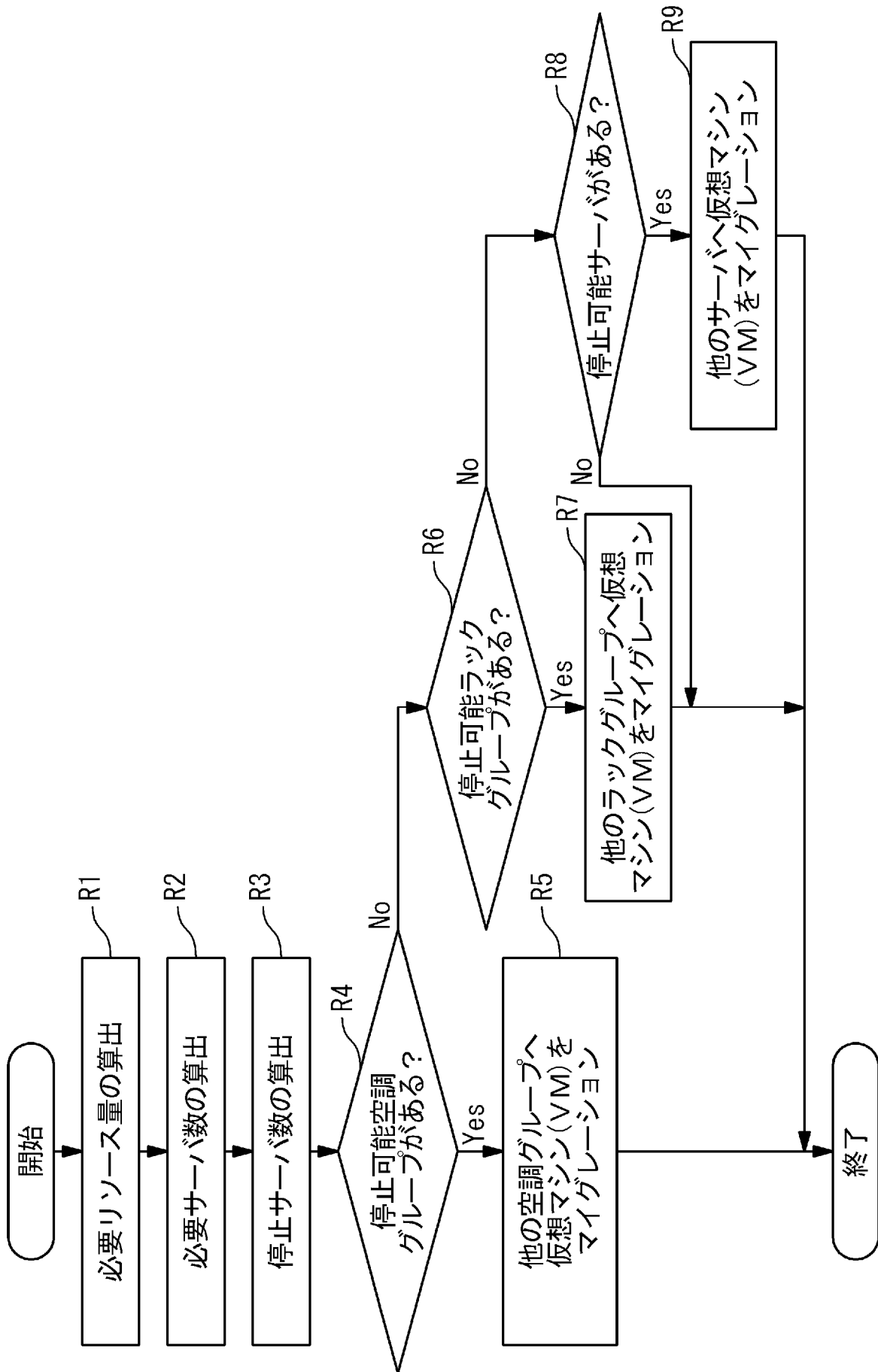
[図2]



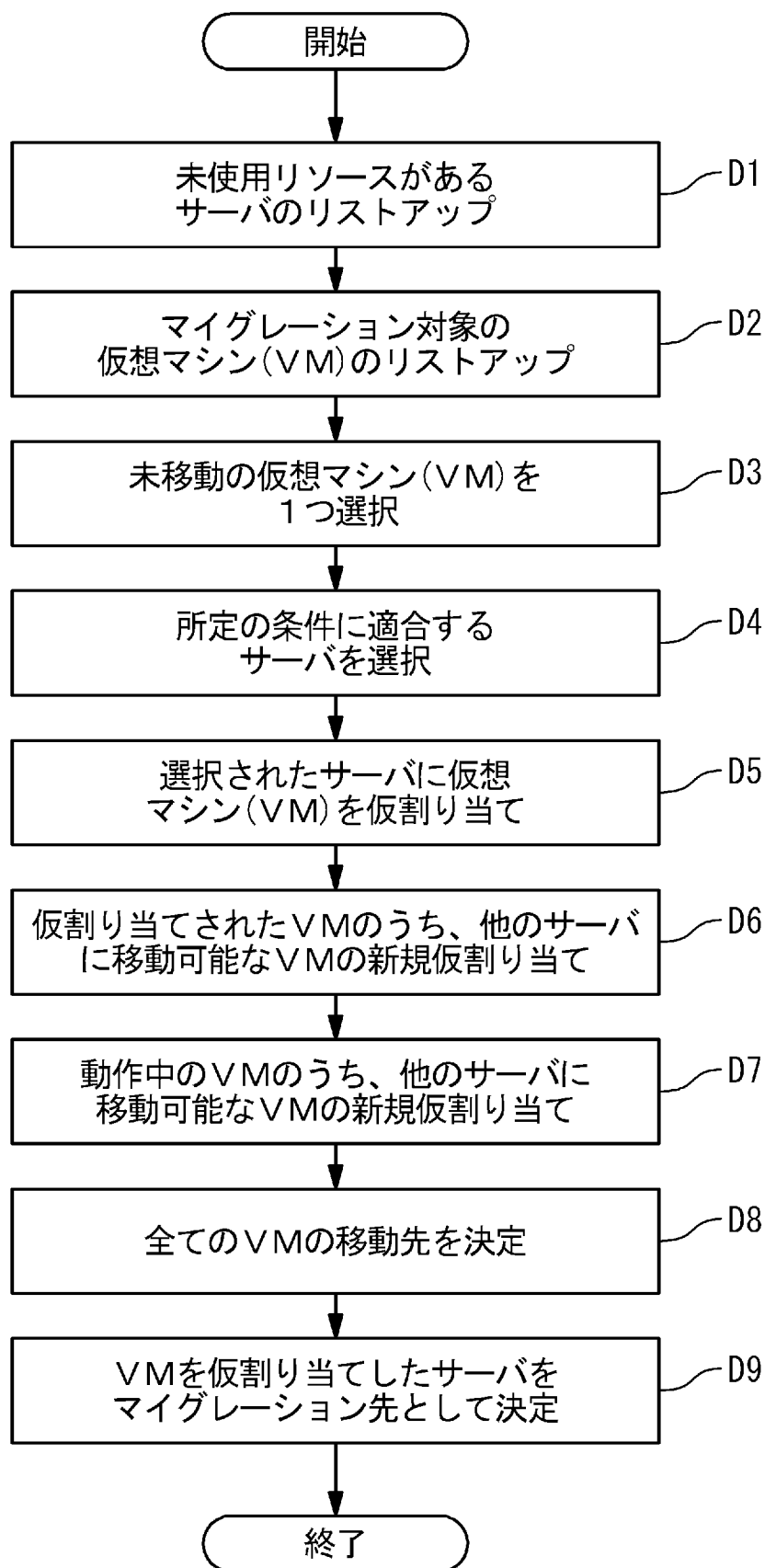
[図3]



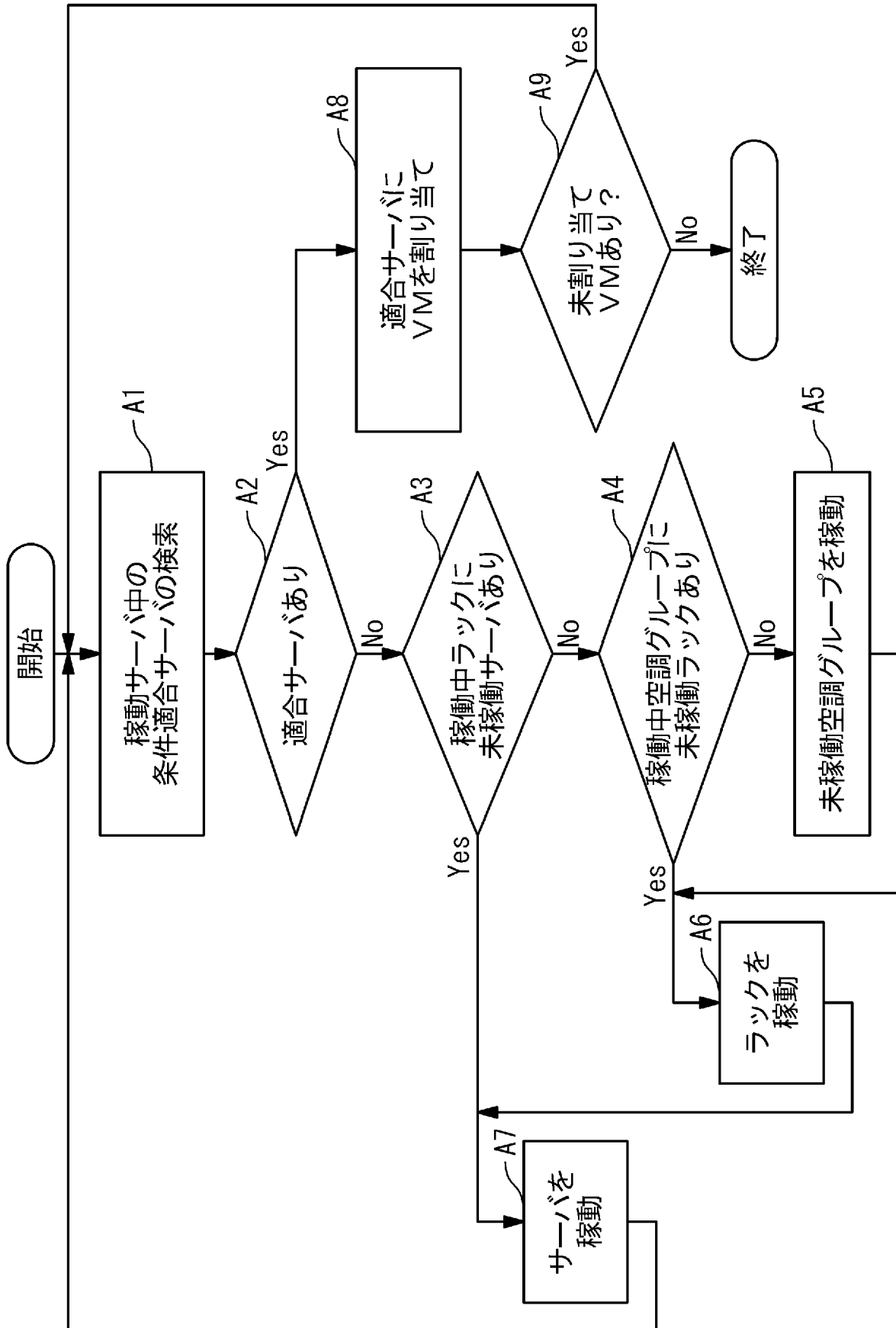
[図4]



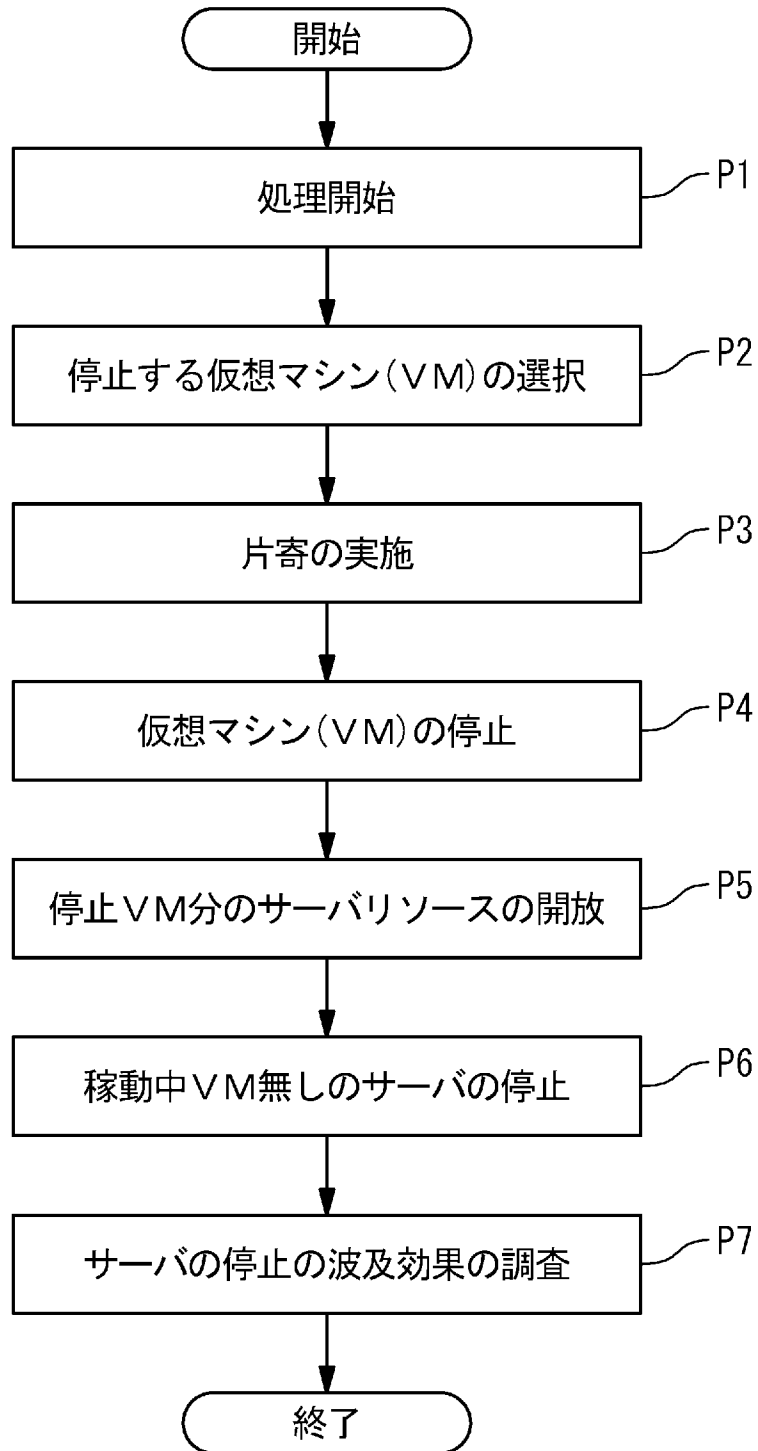
[図5]



[図6]



[図7]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2010/067414

A. CLASSIFICATION OF SUBJECT MATTER

H04L12/28(2006.01)i, G06F1/32(2006.01)i, G06F9/46(2006.01)i, G06F9/50(2006.01)i

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)

H04L12/28, G06F1/32, G06F9/46, G06F9/50

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2010
Kokai Jitsuyo Shinan Koho	1971-2010	Toroku Jitsuyo Shinan Koho	1994-2010

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	JP 2009-169858 A (NEC Corp.), 30 July 2009 (30.07.2009), entire text & US 2009/0187775 A1	1-5, 7-13, 15-16 6, 14
Y A	NEC Corp., "Shin Sedai NW ni Muketa Programmable Flow Switch o Shisaku shi Nichibei-kan deno Jissho Jikken ni Seiko, 29 October 2008 (29.10.2008): Press Release NEC", [online], 29 October 2008 (29.10.2008), [retrieval date 17 December 2010 (17.12.2010)], Internet <URL:http://www.nec.co.jp/press/ja/ 0810/2903.html>	1-5, 7-13, 15-16 6, 14

 Further documents are listed in the continuation of Box C. 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 but 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
17 December, 2010 (17.12.10)Date of mailing of the international search report
28 December, 2010 (28.12.10)Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2010/067414

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	JP 2003-281008 A (Toshiba Corp.), 03 October 2003 (03.10.2003), paragraphs [0016] to [0018]; fig. 1 to 2, 4 & US 2003/0225904 A1	1-5, 7-13, 15-16 6, 14
Y A	JP 2007-189848 A (Hitachi, Ltd.), 26 July 2007 (26.07.2007), paragraphs [0012] to [0014]; fig. 1 (Family: none)	5, 7, 13, 15 6, 14
A	JP 2008-276320 A (NEC Corp.), 13 November 2008 (13.11.2008), entire text (Family: none)	6, 14

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. H04L12/28(2006.01)i, G06F1/32(2006.01)i, G06F9/46(2006.01)i, G06F9/50(2006.01)i

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int.Cl. H04L12/28, G06F1/32, G06F9/46, G06F9/50

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2010年
日本国実用新案登録公報	1996-2010年
日本国登録実用新案公報	1994-2010年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y A	JP 2009-169858 A (日本電気株式会社) 2009.07.30, 全文 & US 2009/0187775 A1	1-5, 7-13, 15-16 6, 14
Y A	日本電気株式会社, “新世代 NW に向けたプログラマブルフロッソスイッチを試作し日米間での実証実験に成功(2008年10月29日): プレスリリース NEC”, [online], 2008.10.29, [検索日 2010.12.17], インターネット <URL : http://www.nec.co.jp/press/ja/0810/2903.html >	1-5, 7-13, 15-16 6, 14

C欄の続きにも文献が列挙されている。

パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー	の日の後に公表された文献
「A」特に関連のある文献ではなく、一般的技術水準を示すもの	「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)	「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「O」口頭による開示、使用、展示等に言及する文献	「&」同一パテントファミリー文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願	

国際調査を完了した日 17.12.2010	国際調査報告の発送日 28.12.2010
国際調査機関の名称及びあて先 日本国特許庁 (ISA/JP) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官 (権限のある職員) 三浦 みちる 電話番号 03-3581-1101 内線 3521

5E 4442

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y A	JP 2003-281008 A (株式会社東芝) 2003.10.03, 段落【0016】－【0018】, 図1－2, 4 & US 2003/0225904 A1	1-5, 7-13, 15-16 6, 14
Y A	JP 2007-189848 A (株式会社日立製作所) 2007.07.26, 段落【0012】－【0014】, 図1 (ファミリーなし)	5, 7, 13, 15 6, 14
A	JP 2008-276320 A (日本電気株式会社) 2008.11.13, 全文 (ファミリーなし)	6, 14