

(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

(43) 国際公開日
2020年5月7日(07.05.2020)



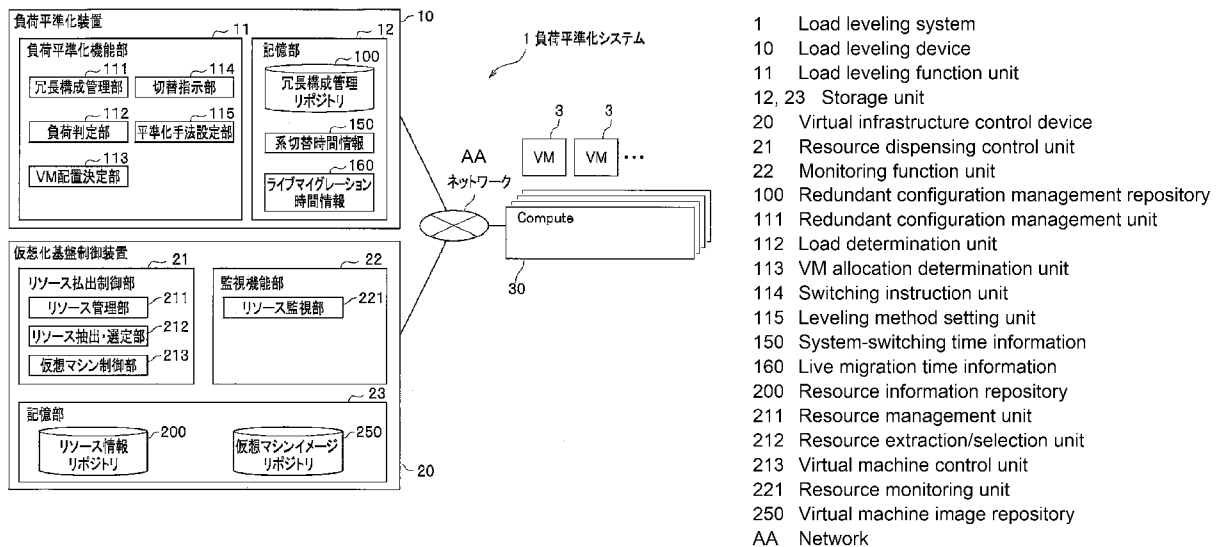
(10) 国際公開番号

WO 2020/090777 A1

- (51) 国際特許分類:
G06F 9/50 (2006.01) G06F 11/20 (2006.01)
G06F 9/455 (2018.01)
- (21) 国際出願番号: PCT/JP2019/042265
- (22) 国際出願日: 2019年10月29日(29.10.2019)
- (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 2018-207710 2018年11月2日(02.11.2018) JP
- (71) 出願人: 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目5番1号 Tokyo (JP).
- (72) 発明者: 岩佐 絵里子(IWASA, Eriko); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 Tokyo (JP). 浜田 信(HAMADA, Makoto); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 Tokyo (JP).
- (74) 代理人: 特許業務法人磯野国際特許商標事務所 (ISONO INTERNATIONAL PATENT OFFICE, P.C.); 〒1050001 東京都港区虎ノ門一丁目1番18号 ヒューリック虎ノ門ビル 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, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,

(54) Title: LOAD LEVELING DEVICE, LOAD LEVELING METHOD, AND LOAD LEVELING PROGRAM

(54) 発明の名称: 負荷平準化装置、負荷平準化方法および負荷平準化プログラム



(57) Abstract: [Problem] To realize leveling of a load between hosts (computes) by reducing an influence on services in a virtual infrastructure irrespective of limitation on application of a virtualization method. [Solution] A load leveling device 10 is provided with: a load determination unit 112 that acquires load information on computes 30 of VMs 3, calculates loads of each of the computes 30 to which the respective VMs 3 are allocated, by performing adding up for the compute 30, calculates the standard deviation of the loads of the compute 30, and determines that load bias is large when the calculated standard deviation is larger than a prescribed reference standard deviation; a VM allocation determination unit 113 that extracts a compute 30 of the highest load, from among the loads of the computes 30, selects a VM 3 that is allocated to the extracted compute 30, and determines a pair of VMs for system-switching by identifying a VM 3 which is to be paired

HN, HR, HU, ID, IL, IN, IR, IS, JO, KE, KG, KH,
KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ,
NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,
QA, RO, RS, RU, RW, SA, 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, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, 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, KM, ML, MR, NE, SN, TD, TG).

添付公開書類：

- 一 国際調査報告（条約第21条(3)）

up with the selected VM 3; and a switching instruction unit 114 that outputs instruction information for system-switching of the determined pair of VMs.

(57) 要約：【課題】仮想化基盤において、仮想化手法の適用制限に関わりなく、かつ、サービスへの影響を低減して、ホスト（コンピュータ）間の負荷の平準化を実現する。【解決手段】負荷平準化装置 10 は、各 VM 3 のコンピュータ 30 に対する負荷情報を取得し、当該 VM 3 が配置されているコンピュータ 30 毎に合計して各コンピュータ 30 の負荷を算出して各コンピュータ 30 の負荷の標準偏差を算出し、算出した標準偏差が所定の基準標準偏差より大きい場合に、負荷の偏りが大きいと判定する負荷判定部 112 と、各コンピュータ 30 の負荷のうち最も高い負荷のコンピュータ 30 を抽出し、当該コンピュータ 30 に配置される VM 3 を選択し、選択した VM 3 とペアとなる VM 3 を特定して系切り替えをするペアの VM として決定する VM 配置決定部 113 と、決定したペアの VM について系切り替えする指示情報を出力する切替指示部 114 と、を備える。

明 細 書

発明の名称：

負荷平準化装置、負荷平準化方法および負荷平準化プログラム

技術分野

[0001] 本発明は、仮想化基盤におけるVM（Virtual Machine：仮想マシン）間の系切り替えを利用した負荷平準化装置、負荷平準化方法および負荷平準化プログラムに関する。

背景技術

[0002] 近年、通信サービスにおいて、仮想化技術を適用したネットワーク機能の仮想化（NFV：Network Functions Virtualization）の採用が進展している。ネットワークの仮想化は、従来専用のハードウェアを用いて実現されていたネットワーク機能をソフトウェア化して汎用サーバ上で動作させるように構成する。このようなネットワーク機能の仮想化技術をキャリア網に適用することで、経済的にスケーラビリティと信頼性を両立させることや、迅速なサービスの提供、サービス毎の需要に応じた柔軟なリソース割当、ハードウェアの寿命に縛られないサービス展開等が実現できると期待されている。

[0003] この仮想化技術を用いてサーバやネットワークといった物理資源を抽象化・隠蔽し、複数のアプリケーションやサービスに対して共通基盤を提供するシステムを仮想化基盤という。

[0004] 仮想化基盤を実現する技術として、例えば、VMware vSphere（「VMware vSphere」は登録商標）と、OpenStack（「OpenStack」は登録商標）とが知られている。

VMware vSphereは、既存の仮想化基盤ソフトウェアの1つであり、そのうち、柔軟なリソース制御と稼働ホストの分散を実現する技術としてDRS（Distributed Resource Scheduler）がある（非特許文献1参照）。DRSでは、コンピュータ毎に負荷を測定し、偏りが発生した場合にはライブマイグレーションにより負荷の高いコンピュータから低いコンピュータに対してVM

を移行することで負荷を平準化させることが可能である。

[0005] OpenStackは、クラウド環境構築用のオープンソースソフトウェア群であり、そのうち、Watcherは、リソース最適化を目的とするコンポーネントである（非特許文献2参照）。このWatcherにより、コンピュータのロードアベレージや電力最適化の目標指標値を設定し、その達成に向けてマイグレーション等を自動で実行する。

[0006] 一方、通信ソフトウェアの多くは、A C T（Active）／S B Y（Standby）構成やN＋M構成のようなH A（High Availability）構成を採用することにより可用性を高めている。例えば、A C T／S B Y構成では、動作中（Active）コンポーネントと、待機中（Standby）コンポーネントとが存在し、Activeコンポーネントに障害が発生した場合には、Standbyコンポーネントが処理を引き継ぐことで、サービスの停止を防いだり、停止時間を極めて短い時間にしたりすることが可能である。

先行技術文献

非特許文献

[0007] 非特許文献1：“vSphere DRS,” [online], 2018/10/3, VMware, Inc., [平成30年10月25日検索], インターネット<URL : <https://blogs.vmware.com/jp-cim/2014/02/vspheredrs.html>>

非特許文献2：“Watcher,” [online], [平成30年10月25日検索], インターネット<URL : <https://wiki.openstack.org/wiki/Watcher>>

非特許文献3：“SR-IOV,” OpenStack Docs [online], [平成30年10月25日検索], インターネット<URL : <https://docs.openstack.org/newton/net-working-guide/config-sriov.html>>

発明の概要

発明が解決しようとする課題

[0008] 上記した仮想化基盤によるシステムにおいて、複数のアプリケーションを搭載した場合、VMの生成や削除、マイグレーション等が繰り返されること

により、仮想化基盤を構成するコンピュータ間で負荷が偏るケースが発生することがある。

特に、A C T / S B Y 構成を採用するアプリケーションを仮想化基盤に搭載すると、Active VMはリクエスト処理を担当するため負荷が高いが、Standby VMは負荷が低い。そのため、負荷の偏りが発生し易い。このような負荷の偏りが発生すると、負荷の高いコンピュータに搭載されたアプリケーションに性能劣化等の影響がでる可能性がある。

[0009] 上記した非特許文献1に記載の技術には、コンピュータ毎に負荷を測定し、偏りが発生した場合にはライブマイグレーションにより負荷の高いコンピュータから低いコンピュータに対してVMを移行することにより負荷を平準化させる仕組みが存在する。

[0010] しかしながら、性能要件の高いネットワーク機器の仮想化時にネットワークI/O (Input/Output) の高速化と効率化を達成するため、例えばS R - I O V (Single Root I/O Virtualization) (非特許文献3参照) といった技術を利用することがあるが、仮想化基盤によっては、その際にライブマイグレーションを適用できない、といった制限(以下、「仮想化手法の適用制限」と称する。)が生じるケースが存在する。

また、VMのライブマイグレーションは、仮想ストレージ構成と割りリソース量によってはマイグレーションに時間がかかるため、負荷平準化のためにライブマイグレーションを頻度高く実行することに懸念がある場合がある。例えば、エフェメラルディスクを用いて仮想マシンインスタンスを起動しているときには、仮想ディスクイメージの内容を、移行前のインスタンスと移行後のインスタンス間で複製する処理(ブロックマイグレーション)が必要のため、時間がかかってしまう。

[0011] このような点を鑑みて本発明がなされたのであり、本発明は、仮想化基盤において、仮想化手法の適用制限に関わりなく、かつ、サービスへの影響を低減して、ホスト(コンピュータ)間の負荷の平準化を実現することができる、負荷平準化装置、負荷平準化方法および負荷平準化プログラムを提供す

ることを課題とする。

課題を解決するための手段

[0012] 前記した課題を解決するため、請求項 1 に記載の発明は、A C T (Active) / S B Y (Standby) 構成を採用するアプリケーションを、仮想化基盤の異なるコンピュータ上に構築された V M (Virtual Machine) に搭載し、Active V M と Standby V M とのペアの V M による冗長構成とした際に、仮想化基盤と連携動作することで複数の前記コンピュータの負荷平準化を行う負荷平準化装置であって、前記アプリケーションに対応付けた、前記ペアとなる V M それぞれの識別子を格納する冗長構成管理情報が記憶される記憶部と、各前記 V M の前記コンピュータに対する負荷を示す負荷情報を取得し、前記負荷情報を当該 V M が配置されている前記コンピュータ毎に合計して各コンピュータの負荷を算出し、算出した各コンピュータの負荷の標準偏差を算出し、算出した標準偏差が所定の基準標準偏差より大きい場合に、前記コンピュータ間の負荷の偏りが大きいと判定する負荷判定部と、前記コンピュータ間の負荷の偏りが大きいと判定された場合に、前記算出した各コンピュータの負荷のうち最も高い負荷のコンピュータを抽出し、抽出した当該コンピュータに配置される V M を選択し、前記選択した V M とペアとなる V M を前記冗長構成管理情報を参照して特定し、前記選択した V M と前記特定したそのペアとなる V M とを、Active V M と Standby V M との系切り替えをするペアの V M として決定する V M 配置決定部と、前記決定したペアの V M について系切り替えする指示情報を出力する切替指示部と、を備えることを特徴とする負荷平準化装置とした。

[0013] また、請求項 4 に記載の発明は、A C T (Active) / S B Y (Standby) 構成を採用するアプリケーションを、仮想化基盤の異なるコンピュータ上に構築された V M (Virtual Machine) に搭載し、Active V M と Standby V M とのペアの V M による冗長構成とした際に、仮想化基盤と連携動作することで複数の前記コンピュータの負荷平準化を行う負荷平準化装置の負荷平準化方法であって、前記負荷平準化装置が、前記アプリケーションに対応付けた、

前記ペアとなるVMそれぞれの識別子を格納する冗長構成管理情報が記憶される記憶部を備えており、各前記VMの前記コンピュータに対する負荷を示す負荷情報を取得し、前記負荷情報を当該VMが配置されている前記コンピュータ毎に合計して各コンピュータの負荷を算出し、算出した各コンピュータの負荷の標準偏差を算出し、算出した標準偏差が所定の基準標準偏差より大きい場合に、前記コンピュータ間の負荷の偏りが大きいと判定するステップと、前記コンピュータ間の負荷の偏りが大きいと判定された場合に、前記算出した各コンピュータの負荷のうち最も高い負荷のコンピュータを抽出し、抽出した当該コンピュータに配置されるVMを選択し、前記選択したVMとペアとなるVMを前記冗長構成管理情報を参照して特定し、前記選択したVMと前記特定したそのペアとなるVMとを、Active VMとStandby VMとの系切り替えをするペアのVMとして決定するステップと、前記決定したペアのVMについて系切り替えする指示情報を出力するステップと、を実行することを特徴とする負荷平準化方法とした。

[0014] このようにすることで、負荷平準化装置は、VMの負荷情報を取得して、各コンピュータの負荷を算出し、算出した各コンピュータの負荷の標準偏差が所定の基準標準偏差より大きい場合に、コンピュータ間の負荷の偏りが大きいと判定する。そして、負荷平準化装置は、負荷が最も高いコンピュータに配置されるVMを選択し、そのペアとなるVMとの間で系切り替えする指示情報を出力することができる。

よって、負荷平準化装置は、例えば、仮想化基盤において、SR-IOV等が設定されているような仮想化手法の適用制限がある場合であっても、コンピュータ間の負荷平準化を実現することができる。

[0015] 請求項2に記載の発明は、前記VM配置決定部が、前記算出した各コンピュータの負荷のうち最も高い負荷のコンピュータを抽出したときに、抽出した当該コンピュータに配置されるVMのうち負荷の高い順にVMを選択し、前記選択したVMとペアとなるVMを前記冗長構成管理情報を参照して特定し、前記選択したVMとそのペアとなるVMとを系切り替えした場合の前記

各コンピュータの負荷をシミュレーションし、前記シミュレーションした各コンピュータの負荷の標準偏差が前記所定の基準標準偏差以下になるまで前記抽出したコンピュータに配置されるVMの選択を繰り返して、前記系切り替えするペアのVMを決定すること、を特徴とする請求項1に記載の負荷平準化装置とした。

[0016] このようにすることで、負荷平準化装置は、所定の基準標準偏差以下になるように、系切り替えするペアのVMを決定するので、より確実に負荷の平準化を実現することができる。

[0017] 請求項3に記載の発明は、前記記憶部には、前記コンピュータに配置されるVMを、他のコンピュータにライブマイグレーションした場合にかかる時間を示すライブマイグレーション時間と、過去において前記指示情報に基づき系切り替えした場合にかかった時間を示す系切り替え時間とが記憶されており、所定期間の前記系切り替え時間の平均値を算出し、算出した前記系切り替え時間の平均値と、前記ライブマイグレーション時間とを比較し、前記ライブマイグレーション時間の方が短ければ、ライブマイグレーションによる負荷平準化機能を有効にする設定情報を出力し、前記系切り替え時間の平均値の方が短ければ、前記ライブマイグレーションによる負荷平準化機能を無効にする設定情報を出力する平準化手法設定部を、さらに備えることを特徴とする請求項1または請求項2に記載の負荷平準化装置とした。

[0018] このように、負荷平準化装置は、系切り替え時間の平均値とライブマイグレーション時間とを比較して、短い方の負荷平準化手法を設定させることができる。よって、サービスへの影響をより低減して、コンピュータ間の負荷の平準化を実現することができる。

[0019] 請求項5に記載の発明は、コンピュータを、請求項1乃至請求項3のいずれか1項に記載の負荷平準化装置として機能させるための負荷平準化プログラムとした。

[0020] このようにすることにより、一般的なコンピュータを用いて、請求項1乃至請求項3のいずれか1項に記載の負荷平準化装置の各機能を実現すること

ができる。

発明の効果

[0021] 本発明によれば、仮想化基盤において、仮想化手法の適用制限に関わりなく、かつ、サービスへの影響を低減して、ホスト（コンピュータ）間の負荷の平準化を実現する、負荷平準化装置、負荷平準化方法および負荷平準化プログラムを提供することができる。

図面の簡単な説明

[0022] [図1]本実施形態に係る負荷平準化装置の処理の概要を説明するための図である。

[図2]本実施形態に係る負荷平準化装置を含む負荷平準化システムの構成を示す図である。

[図3]本実施形態に係るリソース管理テーブル（リソース管理情報）のデータ構成の一例を示す図である。

[図4]本実施形態に係る冗長構成管理テーブル（冗長構成管理情報）のデータ構成の一例を示す図である。

[図5]本実施形態に係る負荷平準化システムにおけるVMの起動処理の流れを示すフローチャートである。

[図6]本実施形態に係る負荷平準化装置が実行する負荷偏り判定処理の流れを示すフローチャートである。

[図7]本実施形態に係る負荷平準化装置が実行する系切り替え指示出力処理の流れを示すフローチャートである。

[図8]本実施形態に係る負荷平準化装置が実行する平準化手法設定処理の流れ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

[0023] 次に、本発明を実施するための形態（以下、「本実施形態」と称する。）について説明する。まず、本実施形態に係る負荷平準化装置10（後記する図2参照）が実行する処理の概要を説明する。

[0024] <概要>

図1は、本実施形態に係る負荷平準化装置10の処理の概要を説明するための図である。

上記したように、仮想化基盤（後記する図2の「仮想化基盤制御装置20」）により、コンピュータ（Compute）に複数のアプリケーションを搭載した場合、VM生成や削除、マイグレーション等が繰り返されることで、仮想化基盤を構成するコンピュータ間で負荷が偏るケースが発生し得る。特に、ACT/SBY構成を採用するアプリケーションを仮想化基盤に搭載すると、Active VMは、リクエスト処理を担当するため負荷が高いが、Standby VMは負荷が低く負荷の偏りが発生しやすい。

そこで、本実施形態に係る負荷平準化装置10は、ACT/SBY構成を採用するアプリケーションを仮想化基盤のVMに搭載した際に、各宿主（コンピュータ）の負荷を監視し、負荷の偏りを検出する。そして、負荷平準化装置10は、負荷の偏りを検出した場合に、Active VM（ACT系のVM）とStandby VM（SBY系のVM）との系切り替えによって、負荷の偏りを低減することを特徴とする。

[0025] ここで、図1では、VM「#1」（ACT）とVM「#11」（SBY）が冗長構成のペアとなり、同様に、VM「#2」（ACT）とVM「#22」（SBY）、VM「#3」（ACT）とVM「#33」（SBY）、VM「#4」（ACT）とVM「#44」（SBY）が冗長構成のペアであるものとする。

この場合、図1（a）に示すように、コンピュータ「#1」は、ACT系のVMが多く配置されているため、負荷が高くなっている（符号A参照）。一方、コンピュータ「#4」は、SBY系のVMが多く配置されているため、負荷が低い（符号B参照）。このような負荷の偏りが発生すると、負荷の高いコンピュータに搭載されたアプリケーションに性能劣化といった影響が発生する可能性がある。

そこで、図1（b）に示すように、負荷の高いコンピュータ（コンピュータ「#1」）に配置されたVM（ここでは、VM「#1」）と、負荷が低い

コンピュータ（コンピュータ「#4」）に配置されているペアのVM（ここでは、VM「#11」）とにおいて、Active VM とStandby VMの系切り替えを行うことにより、負荷の平準化を実現する。

[0026] このようにすることにより、本実施形態に係る負荷平準化装置10は、仮想化手法の適用制限に関わりなく、かつ、サービスへの影響を低減して、ホスト（コンピュータ）間の負荷の平準化を実現することができる。

[0027] <本実施形態>

次に、負荷平準化装置10を含む負荷平準化システム1について説明する。

図2は、本実施形態に係る負荷平準化装置10を含む負荷平準化システム1の構成を示す図である。

負荷平準化システム1は、複数のコンピュータ30（物理リソース：VMが動作するホスト（コンピュータ））と、仮想化基盤制御装置（仮想化基盤）20と、負荷平準化装置10とが、ネットワーク等を介して通信接続される。

[0028] コンピュータ30には、仮想化基盤制御装置20の制御に基づき、仮想化環境として1つ以上のVM3を構築することができる。また、VM3には、1つ以上のアプリケーションが配置（搭載）することができ、ユーザ端末（図示省略）からの要求に基づくサービスを提供する。

仮想化基盤制御装置20は、負荷平準化装置10、各コンピュータ30等に通信接続され、VM3の制御やリソースの割り当て等を行う。また、仮想化基盤制御装置20は、各コンピュータ30および各VM3のリソース監視を行う。

負荷平準化装置10は、仮想化基盤制御装置20等と通信接続され、各VM3の負荷の情報（負荷情報）を取得し、各コンピュータ30について負荷の偏りを検出した場合に、その負荷の偏りを低減するため、ペアとなるVM3のACT/SBY構成の系切り替えの指示情報を出力する。

以下、仮想化基盤制御装置20と、負荷平準化装置10について、詳細に

説明する。

[0029] ≪仮想化基盤制御装置≫

仮想化基盤制御装置（仮想化基盤）20は、制御部、入出力部（いずれも図示省略）および記憶部23を備えるコンピュータ等により実現される。

入出力部は、情報の送受信を行うための通信インタフェース、および、タッチパネルやキーボード等の入力装置や、モニタ等の出力装置との間で情報の送受信を行うための入出力インタフェースからなる。

[0030] 記憶部23は、フラッシュメモリやハードディスク、RAM（Random Access Memory）等により構成される。この仮想化基盤制御装置20の記憶部23には、図2に示すように、リソース情報リポジトリ200と仮想マシンイメージリポジトリ250とが記憶される。

リソース情報リポジトリ200には、各コンピュータ30に構築するVM3の情報（割当先の情報）を含むリソース管理テーブル210（リソース管理情報）が格納される（後記する図3参照）。また、仮想マシンイメージリポジトリ250には、インストール済みのゲストOSを含む仮想インスタンス（VM）を起動するためのテンプレートであるVMイメージが格納される。

[0031] また、制御部は、図2に示すように、リソース払出制御部21と、監視機能部22とを含んで構成される。なお、この制御部には、仮想化基盤を実現するために必要となる他の既存機能も備える。

リソース払出制御部21は、リソースの管理とVM3へのリソースの割り当てを行い、リソース管理部211、リソース抽出・選定部212、仮想マシン制御部213を備える。

[0032] リソース管理部211は、各コンピュータ30（ホスト）のリソース状況を、リソース管理テーブル210（図3参照）を用いて管理する。

[0033] 図3は、本実施形態に係るリソース管理テーブル210（リソース管理情報）のデータ構成の一例を示す図である。

リソース管理テーブル210には、コンピュータ30（物理リソース）の

識別情報である「HW I d」に対応付けて、VM 3の「割当先」の情報が格納される。

例えば、図3に示すように、HW I d「# 0」のコンピュータ30に、VM「# 1」とVM「# 2」が割り当てられていることを示している。

[0034] また、このリソース管理テーブル210には、「HW I d」に対応付けて、そのコンピュータ30の「多重率」や「使用率」を格納するようにしてもよい。

「多重率」は、基準となるCPUのCPU性能と各コンピュータ30のCPU性能の比率（CPU性能比率）であり、割当可能なCPUの「使用率」の上限を表わすものである。この多重率の値は、図3に示すように、100%が上限値ではなく、コンピュータ30の性能比率に応じて値が設定される。高性能なコンピュータ30では、100%より大きな値（例えば、「150%」）を設定する。これにより、性能の異なるコンピュータ30（物理サーバ）で構成されるシステムにおいても、一つのリソース要求情報に基づいて、リソースを割り当てることができる。

「使用率」は、割当済みの仮想CPUについて累積したCPU使用率を示す。

[0035] リソース抽出・選定部212は、リソース要求情報（VMの起動要求）を受け付けた場合に、割当可能なリソース（割当先物理リソース候補）を抽出し、その候補の中から、所定の割当ポリシーにより、割当先物理リソースを選定する。

[0036] 仮想マシン制御部213は、リソース抽出・選定部212が選定したコンピュータ30について、仮想マシンイメージリポジトリ250に記憶されたVMイメージを用いて、VM3を起動する。

[0037] 監視機能部22は、コンピュータ30およびVM3の負荷を監視するリソース監視部221を備える。

リソース監視部221は、例えば、各コンピュータ30やVM3のCPU使用率、メモリ使用率等を監視し、負荷情報として取得して記憶部23に記

憶する。また、リソース監視部 221 は、各コンピュータ 30 の負荷の監視を予め設定した所定の時間間隔（ここでは、「監視間隔 α 」とする。）で実行しているものとする。

[0038] この仮想化基盤制御装置 20 は、既存の仮想化基盤が備える機能として、他に、ライブマイグレーションによる負荷平準化機能（図示省略）を備えている。また、仮想化基盤制御装置 20 は、負荷平準化装置 10 から各コンピュータ 30 のコンフィグ (Configuration) 等の設定情報の問合せがあった場合に、その各コンピュータ 30 のコンフィグ等の設定情報を返信する機能も備えている。

なお、本実施形態に係る負荷平準化装置 10 を含む負荷平準化システム 1 では、既存の仮想化基盤（仮想化基盤制御装置 20）のアプリケーションや I/F（インタフェース）に修正を加える必要がなく、そのまま利用して本発明の作用・機能を実現することができる。よって、本発明を実行する際のシステム管理者等によるシステム改変の負担を軽減することができる。

[0039] ≪負荷平準化装置≫

負荷平準化装置 10 は、仮想化基盤制御装置 20 等から VM 3 のリソース監視の結果情報（負荷情報）を取得して、コンピュータ 30 の負荷（後記する「ホスト負荷」）を算出し、負荷の偏りを検出した場合に、その負荷の偏りを低減するため、ペアとなる VM 3 の ACT/SBY 構成の系切り替えの指示情報を出力する。

なお、負荷平準化装置 10 は、仮想化基盤制御装置 20 に対し、対象とする各コンピュータ 30 のコンフィグ (Configuration) 等の設定情報を問い合わせ取得し、ライブマイグレーションが可能な構成か否かを判定する。そして、負荷平準化装置 10 は、ライブマイグレーションが可能な場合は、系切り替え時間（の平均値）とマイグレーション時間とを比較し、より時間の短い方の手法を採用して負荷平準化を実行させる（詳細は後記）。

[0040] 負荷平準化装置 10 は、制御部、入出力部（いずれも図示省略）および記憶部 12 を備えるコンピュータ等により実現される。

入出力部は、情報の送受信を行うための通信インタフェース、および、タッチパネルやキーボード等の入力装置や、モニタ等の出力装置との間で情報の送受信を行うための入出力インタフェースからなる。

[0041] 記憶部12は、フラッシュメモリやハードディスク、RAM等により構成される。この負荷平準化装置10の記憶部12には、図2に示すように、冗長構成管理リポジトリ100と、系切替時間情報150と、ライブマイグレーション時間情報160とが記憶される（詳細は後記）。また、この記憶部12には、制御部の各機能を実行させるためのプログラム（負荷平準化プログラム）や、制御部の処理に必要な情報が一時的に記憶される。

[0042] また、制御部は、図2に示すように、負荷平準化機能部11を備えており、各コンピュータ30の負荷平準化処理を実行する。

この負荷平準化機能部11は、冗長構成管理部111と、負荷判定部112と、VM配置決定部113と、切替指示部114と、平準化手法設定部115とを含んで構成される。

[0043] 冗長構成管理部111は、仮想化基盤制御装置20の記憶部23（リソース情報リポジトリ200）に格納されたリソース管理テーブル210（図3）の割当先の情報、つまり、コンピュータ30に構築されるVM3の情報が更新された場合に、この更新情報を取得し、自身の記憶部12（冗長構成管理リポジトリ100）に格納された冗長構成管理テーブル110（冗長構成管理情報）を更新する。なお、冗長構成管理部111は、所定の時間間隔（例えば、監視間隔 α ）で、リソース管理テーブル210が更新されたか否かを仮想化基盤制御装置20に問い合わせることにより、更新情報を取得してもよい。また、仮想化基盤制御装置20のリソース管理部211が、リソース管理テーブル210を更新した場合に、その更新情報を、負荷平準化装置10へ通知するように設定してもよい。

[0044] 図4は、本実施形態に係る冗長構成管理テーブル110（冗長構成管理情報）のデータ構成の一例を示す図である。

冗長構成管理テーブル110には、アプリケーション名（App）および

アプリケーションID (App ID) に対応付けて、冗長構成のペアとなる VM3 の ID (識別子) (VM ID) と、そのアプリケーション (App) に対応付けた IP アドレス (Virtual IP) とが格納される。

図4に示すように、例えば、アプリケーション名 (App) が「VNF1」でアプリケーションID (App ID) が「#01」のアプリケーションに対応する、VM3の冗長構成 (ACT/SBY構成) のペアが、VM ID 「#1」「#11」のVM3であり、そのVirtual IPが「192.168.xx.xx」であることを示している。

[0045] なお、Active VMとStandby VMとの系切り替えは、コンピュータ30やVM3の状態のみならず、アプリケーションの個別要件（メンテナンスやソフトウェアアップグレード等の目的）にも依存して実行されることがある。そのため、アプリケーションの個別要件を意識しない仮想化基盤制御装置20（仮想化基盤）が備えるリソース管理テーブル210（図3）等では、どちらのVM3がACT系であるかSBY系であるかの状態（App状態）を管理していない。これに伴い、負荷平準化装置10側においても、冗長構成管理テーブル110等においてApp状態の管理を行わない構成とし、アプリケーションの個別要件を実行する機能を備えないことにより、様々なアプリケーションに共通的に利用可能なシステムを実現可能としている。

[0046] 図2に戻り、負荷判定部112は、各コンピュータ30の負荷として、「ホスト負荷」を算出する。負荷判定部112は、ホスト負荷を次の（式1）により算出する。

$$\text{ホスト負荷} = (\text{ホストに搭載されているVMの負荷の合計値}) / (\text{ホストの容量}) \quad \cdots (\text{式1})$$

ここで、例えば、CPU使用率を指標とした場合は、VM3の負荷は各VM3のCPU使用率となり、ホストの容量は100となる。

[0047] 負荷判定部112は、各VM3の負荷情報を、直接各VM3から取得するようにしてもよいし、仮想化基盤制御装置20から取得するようにしてもよい。

また、負荷判定部 112 は、ホスト負荷の算出を予め設定した所定の時間間隔（例えば、算出間隔 β ）で実行する。なお、算出間隔 β は、監視間隔 α の定数倍とする。

負荷判定部 112 は、各コンピュート 30 の負荷（ホスト負荷）について、負荷の算出時から最も近い時刻に取得した VM 3 の負荷情報の値を利用するようにしてもよいし、過去所定期間の平均値や重み付き平均値等を利用するようにしてもよい。この重み付き平均値の重みは、例えば、古い順から時刻 t_1 、 t_2 、 t_3 のデータがある場合、最も新しい t_3 に、より高い重みを付けることにより行う。

[0048] 負荷判定部 112 は、各ホスト（コンピュート 30）のホスト負荷を算出すると、全ホストのホスト負荷の標準偏差を算出する。そして、負荷判定部 112 は、算出したホスト負荷の標準偏差が、予め設定した（所定の）基準標準偏差 γ より大きい場合、ホスト間の負荷の偏りが大きいと判定し、その旨の情報（後記する「ホスト負荷偏り通知」）を VM 配置決定部 113 に出力する。

[0049] なお、負荷判定部 112 によるコンピュート 30（ホスト）の負荷の偏りの判定手法は、上記の手法に限定されない。例えば、Active VM、Standby VM の個数を負荷とし、この個数の比較を行うなどして負荷の偏りを判定してもよい。

[0050] VM 配置決定部 113 は、所定期間のホスト負荷の平均値が最も大きいホスト（コンピュート 30）を抽出し、そのホスト内で最も負荷の高い VM 3 を選択する。そして、VM 配置決定部 113 は、冗長構成管理テーブル 110（図 4）を参照し、その VM（Active VM）とペアになる VM（Standby VM）を切替候補として決定する。VM 配置決定部 113 は、切替候補となったペアを切り替えた場合に、ホスト負荷の平均値が指標値 δ 以下になるかをシミュレーションして判定する。指標値 δ 以下にならない場合、VM 配置決定部 113 は、当該ホスト内で次に負荷の高い VM 3 とそのペアの VM 3 を選択して切替候補とし、ホスト負荷の平均値が指標値 δ 以下になるまで、この

処理を繰り返す。そして、VM配置決定部113は、ホスト負荷の平均値が指標値 δ 以下になった時点での切替候補のペアを、系切り替えのペアとして決定する。なお、詳細は、後記する図7を用いて説明する。

[0051] なお、VM配置決定部113による、系切り替えの対象となるVM3のペアの決定手法は、上記手法に限定されない。例えば、既存の組み合わせ最適化問題の解法を利用して最適／準最適な系切り替え案を算出するようにしてもよい。

[0052] 切替指示部114は、VM配置決定部113が決定したペアのVM3についての系切り替えを、そのVM3等に対して指示する。

系切り替えの実行は、自動で行われるようにしてもよいし、システム管理者が介在して手動で行われるようにしてもよい。また、系切り替えの指示は、VM3に直接してもよいし、アプリケーションを管理するEMS (Element Management System) を経由してVM3に指示するようにしてもよい。

[0053] 平準化手法設定部115は、対象とする各コンピュータ30のコンフィグ等の設定情報を仮想化基盤制御装置20に問い合わせ取得し、ライブマイグレーションが可能な構成か否かを判定する。なお、この判定に関する技術は、例えば、非特許文献4 (<https://docs.openstack.org/nova/queens/admin/configuring-migrations.html#section-configuring-compute-migrations>) に詳しい。そして、平準化手法設定部115は、ライブマイグレーションが可能である場合に、負荷平準化装置10の系切り替えの指示により行われたそれ以前（過去）の系切り替えで計測された時間を示す系切り替え時間と、ライブマイグレーションにかかる時間（以下、「ライブマイグレーション時間」と称する。）とを比較し、短い時間の方の手法を負荷平準化手法として採用する。

[0054] 具体的には、平準化手法設定部115は、切替指示部114の指示に基づき、ACT/SBY構成のペアのVM3間で系切り替えが行われた場合に、その系切り替えにかかった時間を取得し、系切替時間情報150として、記憶部12に記憶しておく。例えば、切替指示部114が切り替え完了応答と

ともに、系切り替え時間を取得するようにしてもよい。

この系切替時間情報 150 は、VM3 で実行されるアプリケーションにより測定した情報を取得してもよいし、EMS が測定した情報を取得するようにしてもよい。切替指示部 114 からアプリケーションに対して所定の時間間隔でポーリングすることにより状態を監視し、切替時間を測定するようにしてもよい。

[0055] また、ライブマイグレーション時間に関しては、平準化手法設定部 115 が、既存のライブマイグレーションによる負荷平準化手法により実行されたライブマイグレーション時間を、仮想化基盤制御装置 20 等から取得し、記憶部 12 に記憶しておく。または、システム管理者等がライブマイグレーション時間について事前の検証をしておき、平準化手法設定部 115 は、その結果から固定値を設定し、ライブマイグレーション時間情報 160 として記憶部 12 に記憶しておくようにしてもよい。

[0056] 平準化手法設定部 115 は、記憶部 12 を参照し、系切替時間情報 150 とライブマイグレーション時間情報 160 とを用いて、所定時間（期間）の系切替時間の平均値を算出し、その算出した所定時間の系切替時間の平均値とライブマイグレーション時間とを比較し、時間の短い方の負荷平準化手法を採用（設定）する。

[0057] なお、負荷平準化手法を、ライブマイグレーションに切り替える際には、仮想化基盤制御装置 20 のライブマイグレーションによる負荷平準化機能を有効にする設定指示を、平準化手法設定部 115 が送信する。また、平準化手法設定部 115 は、一度仮想化基盤制御装置 20 のライブマイグレーションによる負荷平準化機能を有効にした後において、系切り替えによる負荷平準化手法を採用した場合には、仮想化基盤制御装置 20 のライブマイグレーションによる負荷平準化機能を無効にする設定指示を送信する。

このようにすることにより、負荷平準化に伴うアプリケーションの停止等によるサービスへの影響を、より低減する負荷平準化手法を設定することができる。

[0058] なお、図2においては、負荷平準化装置10と仮想化基盤制御装置20とを別装置（別筐体）として説明したが、上記した負荷平準化装置10の各構成を、仮想化基盤制御装置20に組み込んで一つの装置（筐体）として構成するようにしてもよい。

[0059] <処理の流れ>

次に、負荷平準化装置10が実行する処理の流れについて、図5～図8を参照して説明する。まず、図5を参照して、仮想化基盤制御装置20によるVM3の起動処理について説明する。例えば、このVM3の起動処理を契機として、負荷平準化装置10の記憶部12（冗長構成管理リポジトリ100）内の冗長構成管理テーブル110（図4）が更新される。

[0060] （VMの起動処理に伴う冗長構成管理テーブルの更新）

図5は、本実施形態に係る負荷平準化システム1におけるVM3の起動処理の流れを示すフローチャートである。

まず、仮想化基盤制御装置20のリソース抽出制御部21が新たなVM3の起動要求を受け付ける（ステップS1）。なお、ここでは、ACT/SBY構成としてペアとなる2つのVM3の起動要求がなされたものとして説明する。

[0061] ステップS1において、起動要求を受け付けると、リソース抽出制御部21のリソース抽出・選定部212は、記憶部23のリソース情報リポジトリ200内のリソース管理テーブル210（図3）を参照し、割当可能なリソース（割当先物理リソース候補）を抽出する（ステップS2）。そして、リソース抽出・選定部212が、その抽出された候補の中から、所定の割当ポリシーにより、割当先となるリソース（コンピュータ30）を選定する（ステップS3）。

[0062] 所定の割当ポリシーは、例えば、AffinityルールやAnti-Affinityルールを適用したり、CPUの固定割当（CPU Pinning）を適用したり、既に割り当てられているVM3の使用リソース量の少ないコンピュータ30を選択する等のルールである。この所定の割当ポリシーは、既存のVMの割当先選定手法

であり、例えば非特許文献5 (<https://docs.openstack.org/ocata/config-reference/compute/schedulers.html>) に詳しい。

[0063] ステップS3において起動するVM3の割当先となるコンピュータ30が選定されると、リソース管理部211は、リソース情報リポジトリ200内のリソース管理テーブル210（図3）を更新する（ステップS4）。具体的には、リソース管理部211が、選定したコンピュータ30のHWIDに対応する「割当先」欄に、新たに起動させるVM3のVMIDを追加する。ここでは、ACT/SBY構成としてペアとなる2つのVM3のVMIDが登録される。

そして、リソース管理部211は、リソース管理テーブル210が更新されたことを示す更新情報を、負荷平準化装置10に通知する。

[0064] 負荷平準化装置10の冗長構成管理部111は、仮想化基盤制御装置20から更新情報を受信すると、冗長構成管理リポジトリ100内の冗長構成管理テーブル110（図4）を更新する（ステップS5）。

[0065] 続いて、仮想化基盤制御装置20の仮想マシン制御部213は、記憶部23内の仮想マシンイメージリポジトリ250を参照し、ステップS4において選択されたコンピュータ30に対応するVMイメージを用いて、新たなVM3を起動する（ステップS6）。

そして、リソース管理部211が、新たなVM3が起動されたことを示す応答情報を出し（ステップS7）、処理を終える。

[0066] このようにすることで、仮想化基盤制御装置20により、新たなVM3が起動され設定が更新される度に、負荷平準化装置10はその更新情報を取得して、冗長構成管理テーブル110を、最新の状態に更新することが可能となる。

なお、図5においては、ACT/SBY構成としてペアとなる2つのVM3の起動要求が一度になされたものとして説明したが、ACT/SBY構成としてペアとなるVM3を別々に起動させるようにしてもよい。その場合、各VM3それぞれについてリソース管理テーブル210（図3）を更新した

後、仮想化基盤制御装置 20 のリソース管理部 211 が、リソース管理テーブル 210 が更新されたことを示す更新情報を、負荷平準化装置 10 に通知する。

[0067] (負荷偏り判定処理)

次に、負荷平準化装置 10 の負荷判定部 112 による負荷偏り判定処理について説明する。

図 6 は、本実施形態に係る負荷平準化装置 10 が実行する負荷偏り判定処理の流れを示すフローチャートである。

負荷判定部 112 は、各コンピュート 30 の負荷（ホスト負荷）を算出し、その標準偏差を算出する。そして、負荷判定部 112 は、算出したホスト負荷に基づき、負荷の偏りが予め設定した（所定の）基準標準偏差 γ より大きいか否かを判定し、大きい場合にホスト間の負荷の偏りが大きいと判定する。

なお、ここでは、仮想化基盤制御装置 20 のリソース監視部 221 が、監視間隔 α でホスト（コンピュート 30）と VM3 の負荷を監視しているものとする。以下、具体的に説明する。

[0068] まず、負荷判定部 112 は、負荷の監視結果である負荷情報を取得する回数を制御するための定数 X に 0 を設定 ($X = 0$) する（ステップ S10）。

[0069] 次に、負荷判定部 112 は、各 VM3 の負荷情報（ここでは、CPU 利用率とする。）を取得する（ステップ S11）。負荷判定部 112 は、例えば、仮想化基盤制御装置 20 から、所定の時間間隔（例えば、監視間隔 α ）毎に、負荷情報を取得しておく。そして、負荷判定部 112 は、記憶部 12 内に記憶されたその負荷情報を用いて負荷の偏りを判定してもよいし、所定の時間間隔毎に、仮想化基盤制御装置 20 に要求情報を送信して負荷情報を取得してもよい。

[0070] 続いて、負荷判定部 112 は、監視間隔 α が経過 ($Wait = \alpha$) するまで待つ（ステップ S12）。そして、負荷判定部 112 は、 X に 1 を加算 ($X = X + 1$) する（ステップ S13）。

[0071] 次に、負荷判定部 112 は、ホスト負荷算出のために事前に設定した算出間隔 β (β は、 α の定数倍) が経過したか否かを判定する。具体的には、負荷判定部 112 は、 $X = \beta / \alpha$ が成り立つか否かにより算出間隔 β が経過したか否かを判定する (ステップ S 14)。

ここで、 $X = \beta / \alpha$ が成り立たなければ (ステップ S 14 → No)、ステップ S 11 に戻り、負荷情報の取得の処理を続ける。

一方、 $X = \beta / \alpha$ が成り立てば (ステップ S 14 → Yes)、次のステップ S 15 に進む。

[0072] ステップ S 15 において、負荷判定部 112 は、各ホスト (コンピュータ 30) のホスト負荷を上記した (式 1) により算出する。

[0073] ここで、CPU 使用率を負荷の指標とした場合、VM 3 の負荷は、各 VM 3 の CPU 使用率であり、ホストの容量は 100 として算出する。

なお、負荷は、算出時から最も近い時刻に取得した負荷の値を用いてもよいし、過去の所定期間の平均値を用いてもよい。また、重み付き平均値を用いてもよい。

[0074] 続いて、負荷判定部 112 は、各ホスト (コンピュータ 30) のホスト負荷の標準偏差を算出する (ステップ S 16)。

そして、負荷判定部 112 は、算出したホスト負荷の標準偏差が、予め設定した (所定の) 基準標準偏差 γ より大きいか否かを判定する (ステップ S 17)。

ここで、ホスト負荷の標準偏差が、基準標準偏差 γ 以下の場合 (ステップ S 17 → No)、負荷判定部 112 は、各ホストの負荷の偏りが大きくないと判定し、ステップ S 10 へ戻る。

一方、ホスト負荷の標準偏差が、基準標準偏差 γ より大きい場合 (ステップ S 17 → Yes)、負荷判定部 112 は、各ホストの負荷の偏りが大きいと判定し、次のステップ S 18 へ進む。

[0075] ステップ S 18 において、負荷判定部 112 は、各ホスト 3 の負荷の偏りが大きいと判定した旨の情報 (ホスト負荷偏り通知) を、VM 配置決定部 1

13に出力し、処理を終える。

[0076] このようにすることで、負荷平準化装置10（負荷判定部112）は、各ホストの負荷が所定の基準よりも大きくなっているか否かを判定することができる。

[0077] （系切り替え指示出力処理）

次に、負荷平準化装置10のVM配置決定部113と切替指示部114による系切り替え指示出力処理について説明する。

図7は、本実施形態に係る負荷平準化装置10が実行する系切り替え指示出力処理の流れを示すフローチャートである。

VM配置決定部113は、負荷判定部112により、各ホスト（コンピュータ30）の負荷の偏りが大きい旨の情報（ホスト負荷偏り通知）を取得すると、冗長構成管理テーブル110（図4）を参照し、ACT/SBY構成となるペアのVM3の切替候補を算出する。そして、切替指示部114が、系切り替えの対象として算出したペアのVM3に対し、系切り替えの指示情報を出力する。以下、具体的に説明する。

[0078] まず、負荷平準化装置10（負荷平準化機能部11）のVM配置決定部113は、負荷判定部112から各ホスト（コンピュータ30）の負荷の偏りが大きいと判定した旨の情報（ホスト負荷偏り通知）を受信する（ステップS20）。

[0079] 続いて、VM配置決定部113は、各ホスト（コンピュータ30）における所定期間のホスト負荷の情報をを用いて、各ホストのホスト負荷の平均値（以下、「ホスト負荷平均値」と称する。）を算出し、ホスト負荷平均値が最も大きいホストを抽出する（ステップS21）。

[0080] 次に、VM配置決定部113は、抽出したホスト負荷平均値が最も大きいホスト内で、各VM3の負荷（以下、「VM負荷」と称する。）を降順に並べる（ステップS22）。

なお、このとき利用する各VM3のVM負荷は、算出時から最も近い時刻に取得した負荷の値を用いてもよいし、過去の所定期間の平均値を用いても

よい。また、重み付き平均値を用いてもよい。

[0081] 続いて、VM配置決定部113は、VM負荷の高いVM3を抽出する回数を制御するための定数Yに0を設定($Y=0$)する(ステップS23)。そして、VM配置決定部113は、Yに1を加算($Y=Y+1$)する(ステップS24)。

[0082] 続いて、VM配置決定部113は、各VM3のうちVM負荷が高い順にY番目のVM3を抽出し、冗長構成管理テーブル110(図4)を参照して、抽出したVM3とペアとなるVM3、つまり、抽出したVM3について、ACT/SBY構成となるペアのVM3を特定する。

なお、 $Y=1$ の場合、VM配置決定部113は、各VM3のうちVM負荷が最も高いVM3を抽出し、そのペアのVM3を特定する。 $Y=2$ の場合、VM配置決定部113は、VM負荷が2番目に高いVM3を抽出し、そのペアのVM3を特定する。以下同様である。

そして、VM配置決定部113は、Y番目に特定したVM3のペアを系切り替えしたと仮定した場合のホスト負荷および所定期間のホスト負荷平均値を算出する(ステップS25)。

[0083] 次に、VM配置決定部113は、算出したホスト負荷平均値が、予め設定したホスト負荷平均値の指標値 δ より大きいかな否かを判定する(ステップS26)。

ここで、ホスト負荷平均値が指標値 δ より大きい場合(ステップS26→Yes)、ステップS27に進み、Yがホスト上のVM数と等しいかな否かを判定する。そして、Yがホスト上のVM数と等しい場合には(ステップS27→Yes)、負荷平準化処理を実行できない旨(負荷平準化NG)をシステム管理者等に通知する(ステップS28)。一方、Yがホスト上のVM数と等しくない場合には(ステップS27→No)、ステップS24に戻り処理を続ける。

[0084] また、ステップS26において、ホスト負荷平均値が指標値 δ 以下である場合(ステップS26→No)、切替指示部114が、ステップS25にお

いてホスト負荷平均値を算出したすべてのVM3のペアのうちの一つのVM3のペアを選択し、そのVM3のペアについて系切り替え指示を出力する（ステップS30）。なお、切替指示部114は、VM配置決定部113がステップS25において、ホスト負荷平均値を算出したVM3のペア分、ステップS30の系切り替え指示の出力を繰り返し（ステップS29～S31）、処理を終える。

[0085] このようにすることで、負荷平準化装置10は、仮想化手法の適用制限に関わりなく、つまり、ライブマイグレーションが適用できない場合であっても、系切り替え処理により負荷平準化を実現することができる。

[0086] （平準化手法設定処理）

次に、負荷平準化装置10の平準化手法設定部115が実行する平準化手法設定処理について説明する。

図8は、本実施形態に係る負荷平準化装置10が実行する平準化手法設定処理の流れを示すフローチャートである。

この平準化手法設定処理では、まず、平準化手法設定部115が、コンピュータ30の設定がライブマイグレーション可能な構成か否かを判定する。そして、平準化手法設定部115は、ライブマイグレーションが可能と判定した場合に、VM3の系切り替えによる負荷平準化に係る系切り替え時間の平均値と、ライブマイグレーションによる負荷平準化に係るライブマイグレーション時間とを比較し、短い方の負荷平準化手法を選択して設定（有効化）する処理を行う。なお、この平準化手法設定処理は、上記した、負荷偏り判定処理（図6）および系切り替え指示出力処理（図7）とは独立して所定の時間間隔で実行される。

[0087] まず、負荷平準化装置10の平準化手法設定部115は、対象とするコンピュータ30のコンフィグ等の設定情報を仮想化基盤制御装置20に問い合わせ取得し、ライブマイグレーションが可能な構成か否かを判定する（ステップS40）。

ここで、ライブマイグレーション可能な構成でないと判定した場合（ステ

ップS 4 0→N o)、平準化手法設定部 1 1 5 は、処理を終了する。この場合、負荷平準化装置 1 0 は、VM 3 のペアの系切り替えによる負荷平準化を実行する。

一方、平準化手法設定部 1 1 5 は、ライブマイグレーション可能な構成であると判定した場合には(ステップS 4 0→Y e s)、次のステップS 4 1に進む。

[0088] ステップS 4 1において、平準化手法設定部 1 1 5 は、所定の時間が経過したか否かを判定する。そして、所定の時間が経過していない場合には(ステップS 4 1→N o)、所定の時間が経過するまで待つ。一方、所定の時間が経過した場合は(ステップS 4 1→Y e s)、次のステップS 4 2に進む。

[0089] ステップS 4 2において、平準化手法設定部 1 1 5 は、アプリケーション側で計測している、Active VMとStandby VMとの系切り替え時間を取得し、記憶部 1 2 に系切替時間情報 1 5 0 として格納する。そして、平準化手法設定部 1 1 5 は、系切り替え時間について所定時間(期間)での平均値を算出する。

[0090] 続いて、平準化手法設定部 1 1 5 は、ライブマイグレーション時間を取得する(ステップS 4 3)。このライブマイグレーション時間は、記憶部 1 2 に固定値として予め設定された情報を取得するようにしてもよいし、ライブマイグレーションにかかる時間を、既存のライブマイグレーションによる負荷平準化機能により測定することにより取得してもよい。

[0091] 次に、平準化手法設定部 1 1 5 は、系切り替え時間の平均値とライブマイグレーション時間とを比較し、系切り替え時間の平均値の方が長いかなかを判定する(ステップS 4 4)。

ここで、系切り替え時間の平均値の方が長ければ(ステップS 4 4→Y e s)、平準化手法設定部 1 5 は、平準化手法をライブマイグレーションに設定する。具体的には、平準化手法設定部 1 5 は、平準化手法を系切り替えからライブマイグレーションに切り替える。即ち、平準化手法設定部 1 1 5 は

、仮想化基盤制御装置 20 のライブマイグレーションによる負荷平準化機能を有効化する指示情報を出力する（ステップ S 4 5）。そして、ステップ S 4 1 に戻り処理を続ける。

一方、系切り替え時間の平均値の方がライブマイグレーション時間より長くなければ（同じも含む）（ステップ S 4 4 → N o）、平準化手法設定部 1 5 は、平準化手法を VM 3 のペアの系切り替えに設定する。具体的には、平準化手法設定部 1 1 5 は、仮想化基盤制御装置 20 のライブマイグレーションによる負荷平準化機能が有効であるかを判定し、有効であれば、ライブマイグレーションによる負荷平準化機能を無効化する（ステップ S 4 6）。これにより、負荷平準化機能部 1 1 による系切り替えによる負荷平準化処理が実行される設定となり、ステップ S 4 1 に戻り処理を続ける。

[0092] このように、負荷平準化装置 10 は、系切り替え時間の平均値とライブマイグレーション時間とを比較し短い方の負荷平準化手法を採用するため、負荷平準化に伴うアプリケーションの停止等によるサービスへの影響をより低減することができる。

[0093] 以上説明したように、本実施形態に係る、負荷平準化装置 10、負荷平準化方法および負荷平準化プログラムによれば、仮想化基盤において、仮想化手法の適用制限に関わりなく、かつ、サービスへの影響を低減して、ホスト（コンピュータ）間の負荷の平準化を実現することができる。また、既存の仮想化基盤制御装置 20（仮想化基盤）の修正を加えることなく利用可能となるため、システム管理者等によるシステム改変の負担を軽減することもできる。

符号の説明

- [0094]
- | | |
|----|-----------|
| 1 | 負荷平準化システム |
| 3 | VM（仮想マシン） |
| 10 | 負荷平準化装置 |
| 11 | 負荷平準化機能部 |
| 12 | 記憶部 |

- 2 0 仮想化基盤制御装置
- 2 1 リソース払出制御部
- 2 2 監視機能部
- 2 3 記憶部
- 3 0 コンピュート（ホスト）
- 1 0 0 冗長構成管理リポジトリ
- 1 1 0 冗長構成管理テーブル（冗長構成管理情報）
- 1 1 1 冗長構成管理部
- 1 1 2 負荷判定部
- 1 1 3 VM配置決定部
- 1 1 4 切替指示部
- 1 1 5 平準化手法設定部
- 1 5 0 系切替時間情報
- 1 6 0 ライブマイグレーション時間情報
- 2 0 0 リソース情報リポジトリ
- 2 1 0 リソース管理テーブル（リソース管理情報）
- 2 1 1 リソース管理部
- 2 1 2 リソース抽出・選定部
- 2 1 3 仮想マシン制御部
- 2 2 1 リソース監視部
- 2 5 0 仮想マシンイメージリポジトリ

請求の範囲

[請求項1]

A C T (Active) / S B Y (Standby) 構成を採用するアプリケーションを、仮想化基盤の異なるコンピュータ上に構築されたVM (Virtual Machine) に搭載し、Active VMとStandby VMとのペアのVMによる冗長構成とした際に、仮想化基盤と連携動作することで複数の前記コンピュータの負荷平準化を行う負荷平準化装置であって、

前記アプリケーションに対応付けた、前記ペアとなるVMそれぞれの識別子を格納する冗長構成管理情報が記憶される記憶部と、

各前記VMの前記コンピュータに対する負荷を示す負荷情報を取得し、前記負荷情報を当該VMが配置されている前記コンピュータ毎に合計して各コンピュータの負荷を算出し、算出した各コンピュータの負荷の標準偏差を算出し、算出した標準偏差が所定の基準標準偏差より大きい場合に、前記コンピュータ間の負荷の偏りが大きいと判定する負荷判定部と、

前記コンピュータ間の負荷の偏りが大きいと判定された場合に、前記算出した各コンピュータの負荷のうち最も高い負荷のコンピュータを抽出し、抽出した当該コンピュータに配置されるVMを選択し、前記選択したVMとペアとなるVMを前記冗長構成管理情報を参照して特定し、前記選択したVMと前記特定したそのペアとなるVMとを、Active VMとStandby VMとの系切り替えをするペアのVMとして決定するVM配置決定部と、

前記決定したペアのVMについて系切り替えする指示情報を出力する切替指示部と、

を備えることを特徴とする負荷平準化装置。

[請求項2]

前記VM配置決定部は、

前記算出した各コンピュータの負荷のうち最も高い負荷のコンピュータを抽出したときに、抽出した当該コンピュータに配置されるVMのうち負荷の高い順にVMを選択し、前記選択したVMとペアとなる

VMを前記冗長構成管理情報を参照して特定し、前記選択したVMとそのペアとなるVMとを系切り替えした場合の前記各コンピュータの負荷をシミュレーションし、前記シミュレーションした各コンピュータの負荷の標準偏差が前記所定の基準標準偏差以下になるまで前記抽出したコンピュータに配置されるVMの選択を繰り返して、前記系切り替えするペアのVMを決定すること、

を特徴とする請求項1に記載の負荷平準化装置。

[請求項3]

前記記憶部には、前記コンピュータに配置されるVMを、他のコンピュータにライブマイグレーションした場合にかかる時間を示すライブマイグレーション時間と、過去において前記指示情報に基づき系切り替えした場合にかかった時間を示す系切り替え時間とが記憶されており、

所定期間の前記系切り替え時間の平均値を算出し、算出した前記系切り替え時間の平均値と、前記ライブマイグレーション時間とを比較し、前記ライブマイグレーション時間の方が短ければ、ライブマイグレーションによる負荷平準化機能を有効にする設定情報を出力し、前記系切り替え時間の平均値の方が短ければ、前記ライブマイグレーションによる負荷平準化機能を無効にする設定情報を出力する平準化手法設定部を、さらに備えること

を特徴とする請求項1または請求項2に記載の負荷平準化装置。

[請求項4]

ACT (Active) / SBY (Standby) 構成を採用するアプリケーションを、仮想化基盤の異なるコンピュータ上に構築されたVM (Virtual Machine) に搭載し、Active VMとStandby VMとのペアのVMによる冗長構成とした際に、仮想化基盤と連携動作することで複数の前記コンピュータの負荷平準化を行う負荷平準化装置の負荷平準化方法であって、

前記負荷平準化装置は、

前記アプリケーションに対応付けた、前記ペアとなるVMそれぞれ

の識別子を格納する冗長構成管理情報が記憶される記憶部を備えており、

各前記VMの前記コンピュータに対する負荷を示す負荷情報を取得し、前記負荷情報を当該VMが配置されている前記コンピュータ毎に合計して各コンピュータの負荷を算出し、算出した各コンピュータの負荷の標準偏差を算出し、算出した標準偏差が所定の基準標準偏差より大きい場合に、前記コンピュータ間の負荷の偏りが大きいと判定するステップと、

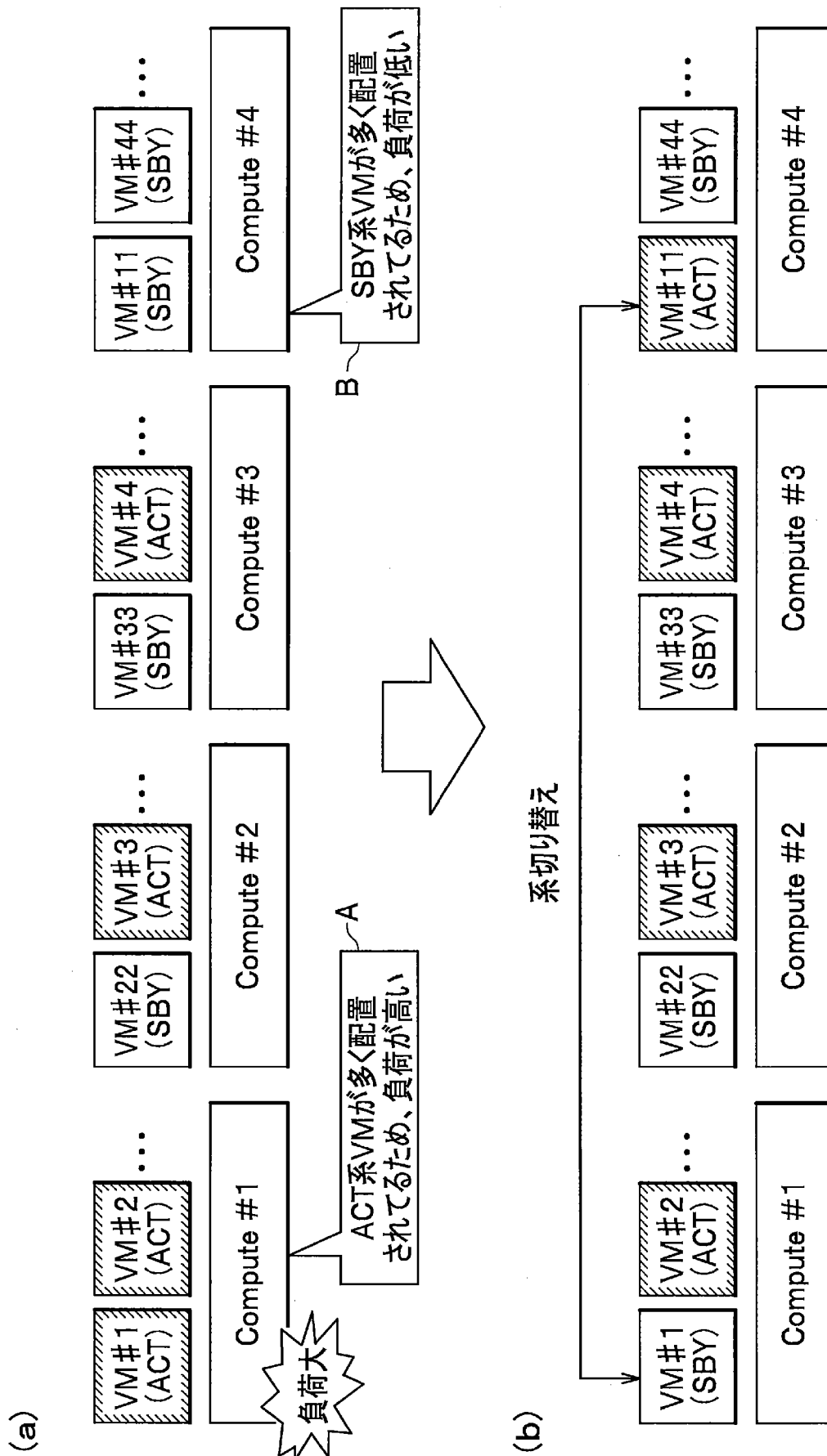
前記コンピュータ間の負荷の偏りが大きいと判定された場合に、前記算出した各コンピュータの負荷のうち最も高い負荷のコンピュータを抽出し、抽出した当該コンピュータに配置されるVMを選択し、前記選択したVMとペアとなるVMを前記冗長構成管理情報を参照して特定し、前記選択したVMと前記特定したそのペアとなるVMとを、Active VMとStandby VMとの系切り替えをするペアのVMとして決定するステップと、

前記決定したペアのVMについて系切り替えする指示情報を出力するステップと、

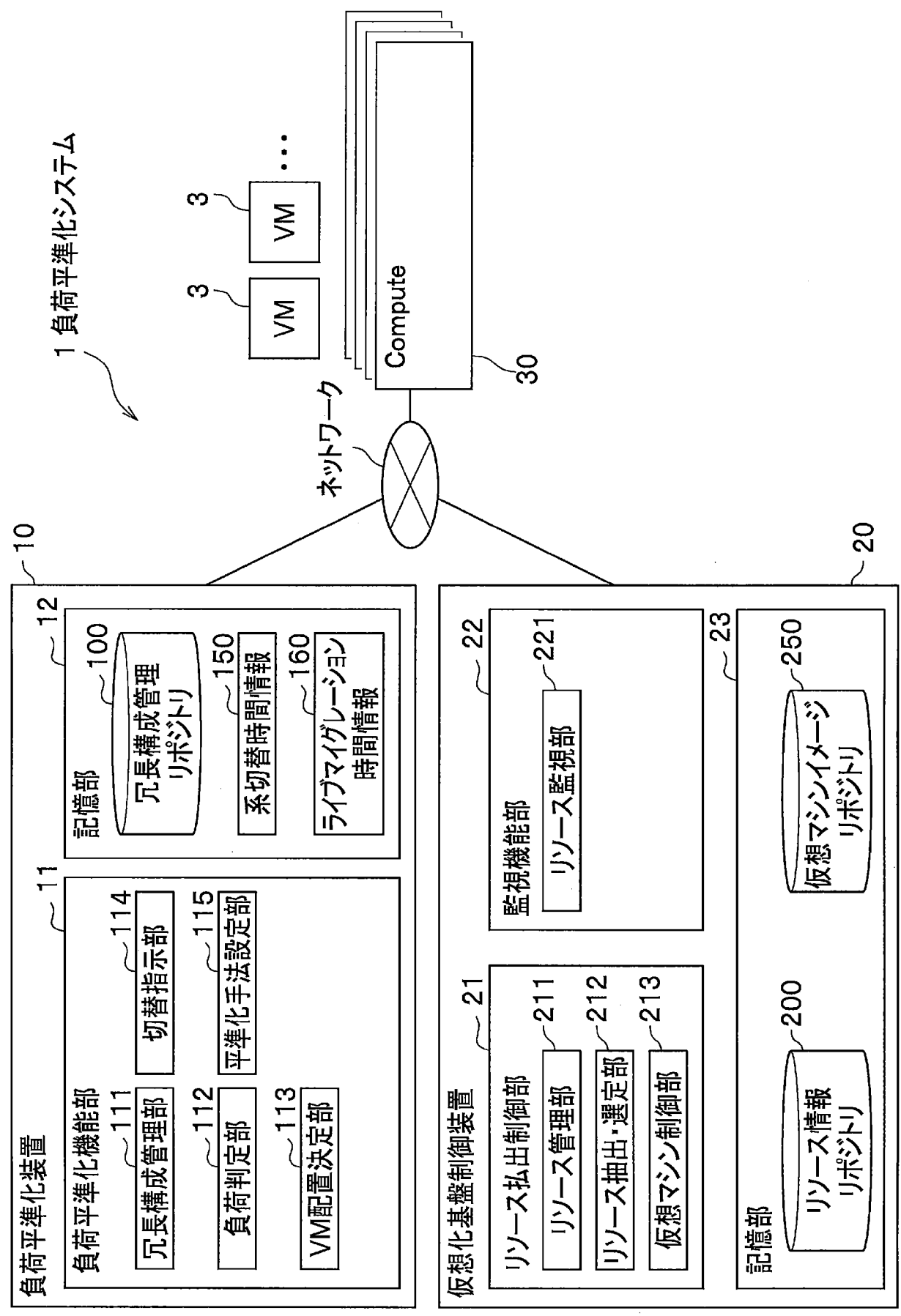
を実行することを特徴とする負荷平準化方法。

[請求項5] コンピュータを、請求項1乃至請求項3のいずれか1項に記載の負荷平準化装置として機能させるための負荷平準化プログラム。

[図1]



[図2]



[図3]

210 リソース管理テーブル



HW Id	多重率	使用率	割当先
#0	100%	50%	VM#1, VM#2
#1	100%	70%	VM#3, VM#10, VM#22
#2	150%	20%	VM#4, VM#33
#3	150%	50%	VM#11, VM#44
⋮	⋮	⋮	⋮

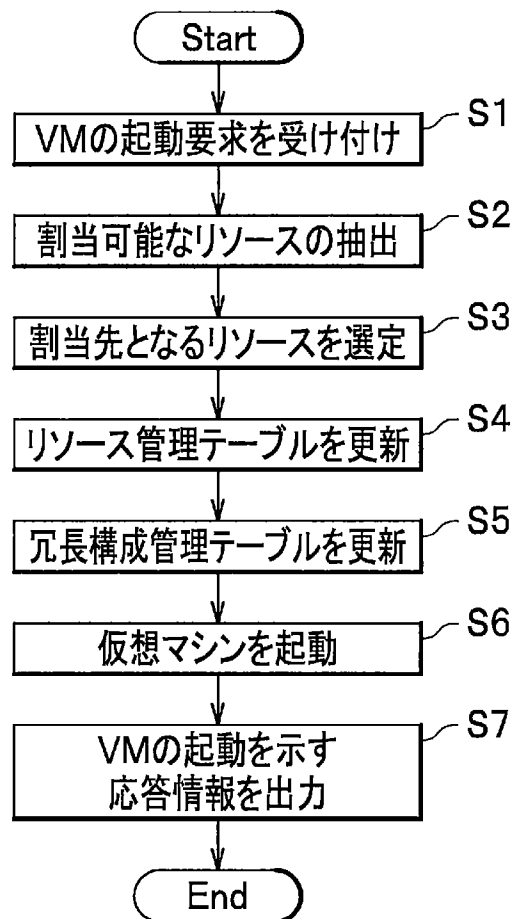
[図4]

110 冗長構成管理テーブル

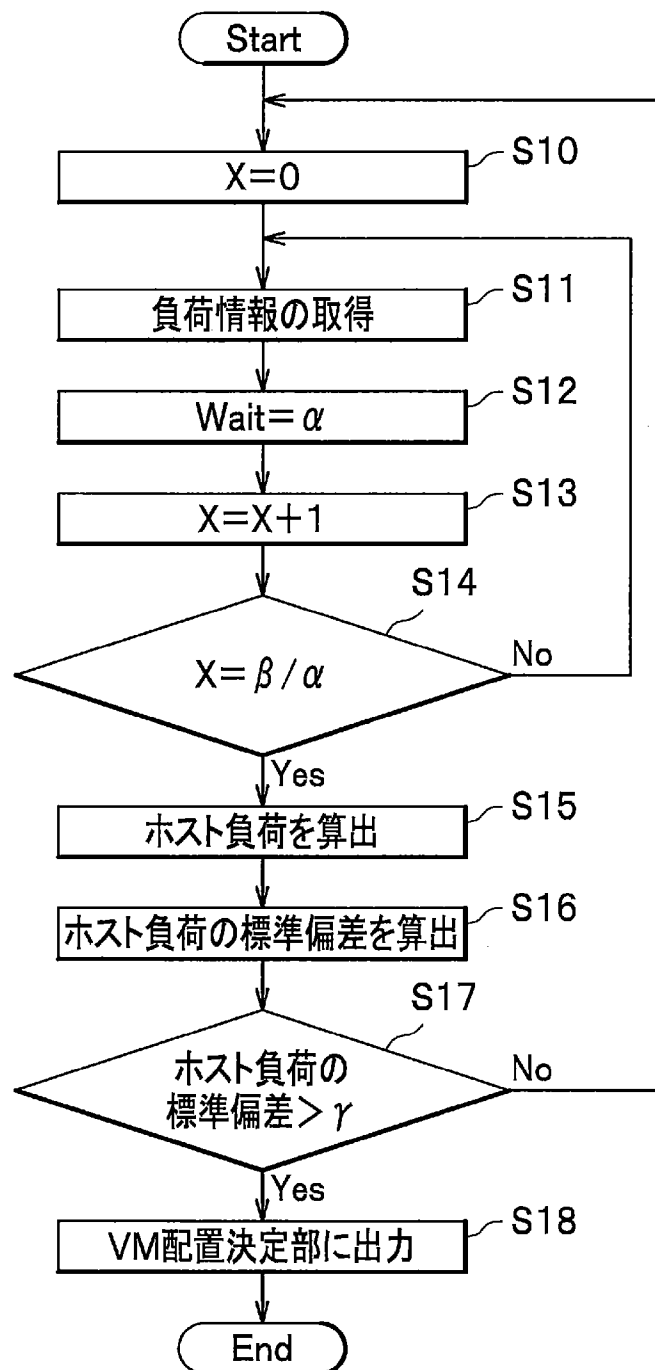


App	App ID	VM ID	Virtual IP
VNF1	#01	#1, #11	192.168.xx.xx
VNF2	#02	#3, #33	192.168.xx.xx
⋮	⋮	⋮	⋮
VNF10	#010	#2, #22	192.168.xx.xx

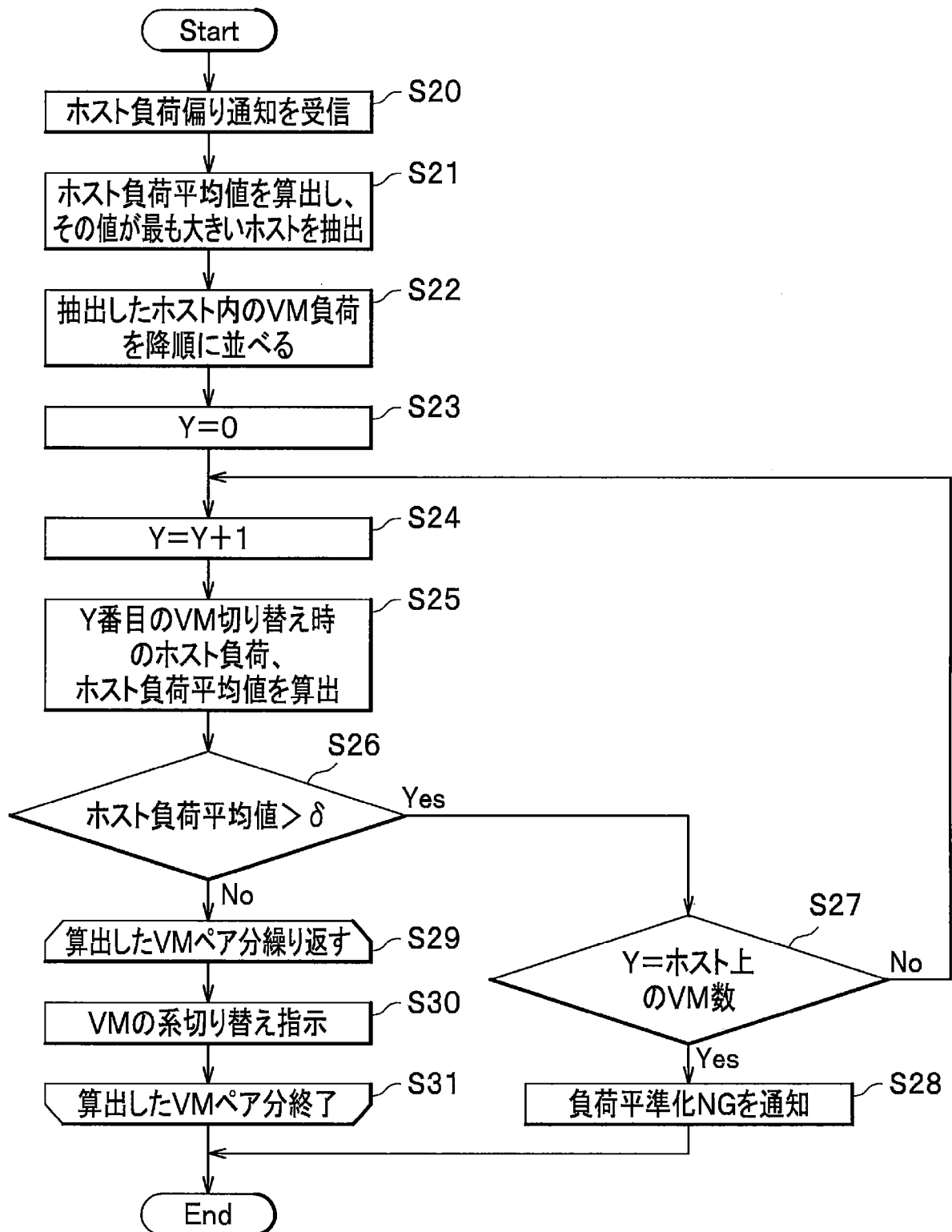
[図5]



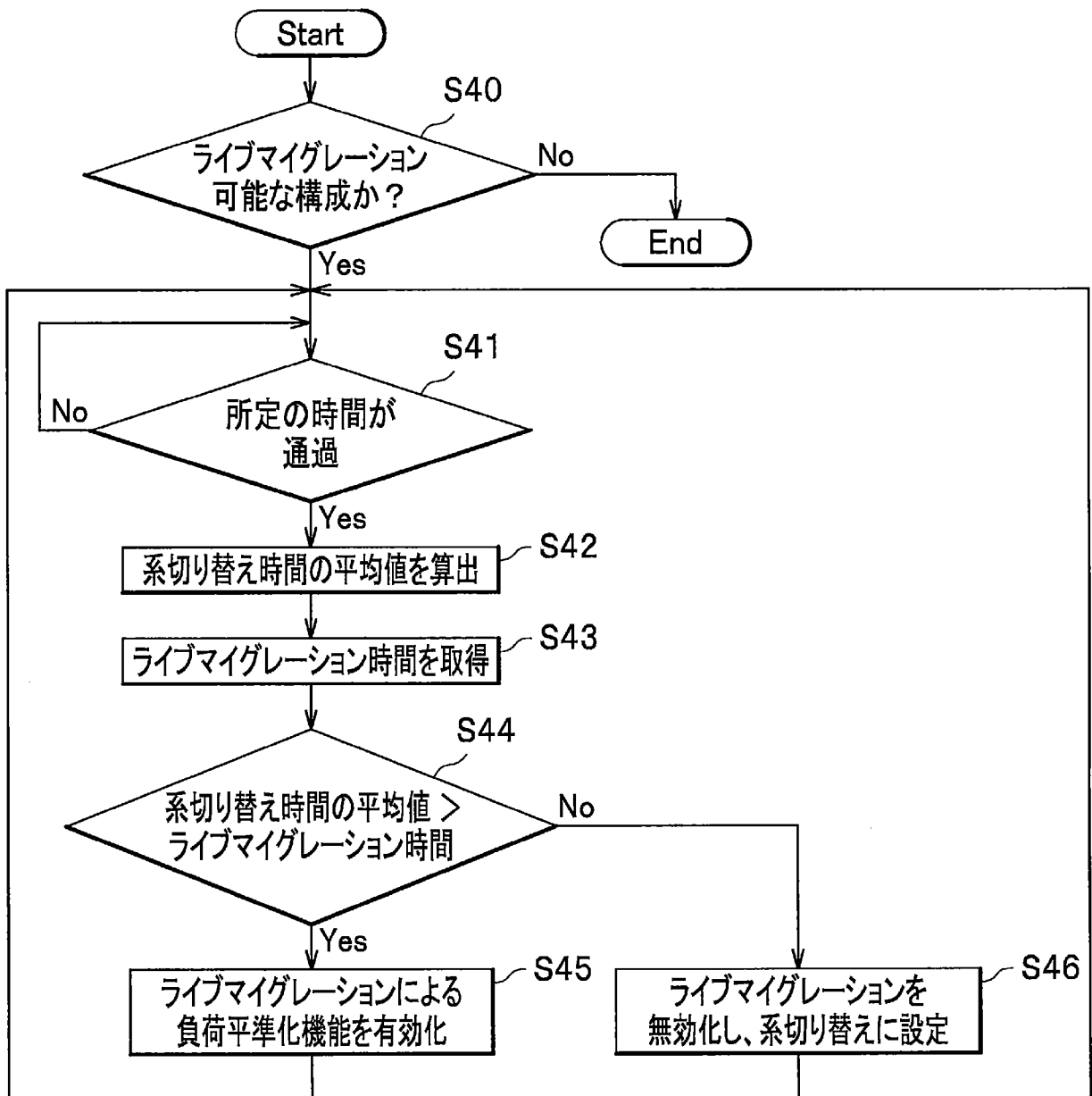
[図6]



[図7]



[図8]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/042265

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. G06F9/50(2006.01)i, G06F9/455(2006.01)i, G06F11/20(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)

Int.Cl. G06F9/50, G06F9/455, G06F11/20

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

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2019

Registered utility model specifications of Japan 1996-2019

Published registered utility model applications of Japan 1994-2019

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.
A	JP 2016-153929 A (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) 25 August 2016 (Family: none)	1-5
A	WO 2016/162916 A1 (HITACHI, LTD.) 13 October 2016 & US 2018/0011728 A1	1-5

☐ 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
13.12.2019

Date of mailing of the international search report
24.12.2019

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G06F9/50(2006.01)i, G06F9/455(2006.01)i, G06F11/20(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. G06F9/50, G06F9/455, G06F11/20

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	1971-2019年
日本国実用新案登録公報	1996-2019年
日本国登録実用新案公報	1994-2019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
A	JP 2016-153929 A（日本電信電話株式会社）2016.08.25 （ファミリーなし）	1-5
A	WO 2016/162916 A1（株式会社日立製作所）2016.10.13 & US 2018/0011728 A1	1-5

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

「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

13.12.2019

国際調査報告の発送日

24.12.2019

国際調査機関の名称及びあて先

日本国特許庁（ISA/J P）

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）

大桃 由紀雄

5B

5583

電話番号 03-3581-1101 内線 3545