

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6801468号
(P6801468)

(45) 発行日 令和2年12月16日(2020.12.16)

(24) 登録日 令和2年11月30日(2020.11.30)

(51) Int. Cl. F I
G 0 6 F 9/455 (2006.01) G O 6 F 9/455 1 5 O
G 0 6 F 9/50 (2006.01) G O 6 F 9/50 1 2 O Z

請求項の数 6 (全 37 頁)

(21) 出願番号	特願2017-7695 (P2017-7695)	(73) 特許権者	000005223
(22) 出願日	平成29年1月19日 (2017.1.19)		富士通株式会社
(65) 公開番号	特開2018-116556 (P2018-116556A)		神奈川県川崎市中原区上小田中4丁目1番1号
(43) 公開日	平成30年7月26日 (2018.7.26)	(74) 代理人	100104190
審査請求日	令和1年10月8日 (2019.10.8)		弁理士 酒井 昭徳
		(72) 発明者	鈴木 成人
			神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
		(72) 発明者	児玉 宏喜
			神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
		(72) 発明者	稲野 聡
			神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

最終頁に続く

(54) 【発明の名称】 管理装置、制御方法、および管理プログラム

(57) 【特許請求の範囲】

【請求項1】

複数の計算資源をそれぞれ有する複数の情報処理装置を備える情報処理システムを管理する管理装置であって、

前記複数の情報処理装置がそれぞれ属し、所定のサービスをそれぞれ提供する複数のグループの各々がそれぞれ実行するホストOSから取得した、各情報処理装置の構成情報と性能情報とを、前記複数のグループ毎に記憶部に記憶する取得部と、

前記記憶部が前記複数のグループ毎に記憶する構成情報と性能情報とに基づき、選択対象の仮想マシンが属する対象グループから他のグループへの移動対象となる移動対象仮想マシンを前記複数のグループ毎に選択する選択部と、

前記移動対象仮想マシンを前記他のグループに移動させた場合、前記移動対象仮想マシンに対応するサービスの性能値が予め定められた基準値を満たすか否かを判断し、前記基準値を満たすとき、前記対象グループから前記他のグループへの前記移動対象仮想マシンの移動を決定する決定部と、

を有することを特徴とする管理装置。

【請求項2】

前記取得部は、各情報処理装置の構成情報に応じた仮想マシン種別を前記複数のグループ毎に記憶部に記憶し、

前記選択部は、前記記憶部が前記複数のグループ毎に記憶する仮想マシン種別と性能情報とに基づき、前記移動対象仮想マシンを複数のグループ毎に選択する、

ことを特徴とする請求項 1 に記載の管理装置。

【請求項 3】

前記決定部は、前記移動対象仮想マシンを前記他のグループに移動させた場合、前記移動対象仮想マシンに対応するサービスの性能値が、前記基準値を満たすか否か、および、移動前の前記移動対象仮想マシンに対応するサービスの性能値よりも良好であることを示すか否かを判断し、前記基準値を満たし、かつ、良好であることを示すとき、前記対象グループから前記他のグループへの前記移動対象仮想マシンの移動を決定する、ことを特徴とする請求項 1 または 2 に記載の管理装置。

【請求項 4】

前記決定部は、前記対象グループの計算資源より前記他のグループの計算資源が少ないとき、前記対象グループから前記他のグループへの前記移動対象仮想マシンの移動を決定する、ことを特徴とする請求項 1 ~ 3 のいずれか一つに記載の管理装置。

10

【請求項 5】

複数の計算資源をそれぞれ有する複数の情報処理装置を備える情報処理システムの制御方法であって、

前記情報処理システムを管理する管理装置が、

前記複数の情報処理装置がそれぞれ属し、所定のサービスをそれぞれ提供する複数のグループの各々がそれぞれ実行するホスト OS から取得した、各情報処理装置の構成情報と性能情報とを、前記複数のグループ毎に記憶部に記憶し、

前記記憶部が前記複数のグループ毎に記憶する構成情報と性能情報とに基づき、選択対象の仮想マシンが属する対象グループから他のグループへの移動対象となる移動対象仮想マシンを前記複数のグループ毎に選択し、

20

前記移動対象仮想マシンを前記他のグループに移動させた場合、前記移動対象仮想マシンに対応するサービスの性能値が予め定められた基準値を満たすか否かを判断し、前記基準値を満たすとき、前記対象グループから前記他のグループへの前記移動対象仮想マシンの移動を決定する、

処理を実行することを特徴とする制御方法。

【請求項 6】

複数の計算資源をそれぞれ有する複数の情報処理装置を備える情報処理システムを管理する管理装置の管理プログラムであって、

30

前記複数の情報処理装置がそれぞれ属し、所定のサービスをそれぞれ提供する複数のグループの各々がそれぞれ実行するホスト OS から取得した、各情報処理装置の構成情報と性能情報とを、前記複数のグループ毎に記憶部に記憶し、

前記記憶部が前記複数のグループ毎に記憶する構成情報と性能情報とに基づき、選択対象の仮想マシンが属する対象グループから他のグループへの移動対象となる移動対象仮想マシンを前記複数のグループ毎に選択し、

前記移動対象仮想マシンを前記他のグループに移動させた場合、前記移動対象仮想マシンに対応するサービスの性能値が予め定められた基準値を満たすか否かを判断し、前記基準値を満たすとき、前記対象グループから前記他のグループへの前記移動対象仮想マシンの移動を決定する、

40

処理を前記管理装置に実行させることを特徴とする管理プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、管理装置、制御方法、および管理プログラムに関する。

【背景技術】

【0002】

従来、データセンター業者が顧客に対して約定する、サービス内容やサービス品質等に関する要求水準である SLA (Service Level Agreement) を満たすよう、データセンター内の情報処理装置等の計算資源を管理し、データセンターにお

50

いてVM (Virtual Machine : 仮想マシン) に割り当てる計算資源を自律的に制御する技術がある。

【0003】

先行技術としては、例えば、異なる物理マシンで実行されている2つの仮想マシン間の双方向間における通信量について、通信量閾値を超えた回数が規定値を超えたとき、2つの仮想マシンが同じ物理マシンで実行されるようにするものがある。また、例えば、新規の仮想マシンを配置可能な1以上のサーバ装置について、新規の仮想マシンをそのサーバ装置に配置する適切度を示す指標値を計算し、計算した指標値を参照して新規の仮想マシンの配置先を決定する技術がある。また、例えば、複数の物理サーバ上に設けられた複数の仮想サーバについて、互いに異なる物理サーバに配置すべきか同一の物理サーバに配置すべきかを示すグループ条件が保たれることを条件に、仮想サーバの再配置を行う技術がある。

10

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2010-140134号公報

【特許文献2】国際公開第2007/136021号

【特許文献3】特開2009-199395号公報

【発明の概要】

【発明が解決しようとする課題】

20

【0005】

しかしながら、従来技術では、複数の仮想マシンを用いて1つのサービスを提供する情報処理システムにおいて用いられる計算資源を、そのサービスのSLAを満たす範囲で最小化することは難しい。例えば、1つ1つの仮想マシンに着目してSLAを満たすように計算資源を最小化しようとしても、情報処理システム全体について計算資源を最小化することができるとは限らない。

【0006】

1つの実施態様によれば、本発明は、情報処理システムにおいて用いられる計算資源の最小化を図ることができる管理装置、制御方法、および管理プログラムを提供することを目的とする。

30

【課題を解決するための手段】

【0007】

1つの実施態様によれば、複数の計算資源をそれぞれ有する複数の情報処理装置を備える情報処理システムを管理し、前記複数の情報処理装置がそれぞれ属し、所定のサービスをそれぞれ提供する複数のグループの各々がそれぞれ実行するホストOSから取得した、各情報処理装置の構成情報と性能情報とを、前記複数のグループ毎に記憶部に記憶する取得部と、前記記憶部が前記複数のグループ毎に記憶する構成情報と性能情報とに基づき、選択対象の仮想マシンが属する対象グループから他のグループへの移動対象となる移動対象仮想マシンを前記複数のグループ毎に選択する選択部と、前記移動対象仮想マシンを前記他のグループに移動させた場合、前記移動対象仮想マシンに対応するサービスの性能値が予め定められた基準値を満たすか否かを判断し、前記基準値を満たすとき、前記対象グループから前記他のグループへの前記移動対象仮想マシンの移動を決定する管理装置、制御方法、および管理プログラムが提案される。

40

【発明の効果】

【0008】

本発明の一態様によれば、情報処理システムにおいて用いられる計算資源の最小化を図ることができるという効果を奏する。

【図面の簡単な説明】

【0009】

【図1】図1は、実施の形態にかかる情報処理システム110の制御方法の一実施例を示

50

す説明図である。

【図 2】図 2 は、情報処理システム 110 の一例を示す説明図である。

【図 3】図 3 は、情報処理システム 110 の使用例を示す説明図である。

【図 4】図 4 は、管理装置 100 のハードウェア構成例を示すブロック図である。

【図 5】図 5 は、計算資源管理テーブル 500 の記憶内容の一例を示す説明図である。

【図 6】図 6 は、VM 種別テーブル 600 の記憶内容の一例を示す説明図である。

【図 7】図 7 は、VM 稼働状況シート 700 の記憶内容の一例を示す説明図である。

【図 8】図 8 は、管理装置 100 の機能的構成例を示すブロック図である。

【図 9】図 9 は、管理装置 100 の機能的構成の具体例を示すブロック図である。

【図 10】図 10 は、管理装置 100 の動作例 1 を示す説明図（その 1）である。

10

【図 11】図 11 は、管理装置 100 の動作例 1 を示す説明図（その 2）である。

【図 12】図 12 は、管理装置 100 の動作例 1 を示す説明図（その 3）である。

【図 13】図 13 は、管理装置 100 の動作例 1 を示す説明図（その 4）である。

【図 14】図 14 は、管理装置 100 の動作例 1 を示す説明図（その 5）である。

【図 15】図 15 は、管理装置 100 の動作例 1 を示す説明図（その 6）である。

【図 16】図 16 は、管理装置 100 の動作例 1 を示す説明図（その 7）である。

【図 17】図 17 は、管理装置 100 の動作例 1 を示す説明図（その 8）である。

【図 18】図 18 は、実施例 1 における管理処理手順の一例を示すフローチャート（その 1）である。

【図 19】図 19 は、実施例 1 における管理処理手順の一例を示すフローチャート（その 2）である。

20

【図 20】図 20 は、実施例 1 における管理処理手順の一例を示すフローチャート（その 3）である。

【図 21】図 21 は、実施例 1 における管理処理手順の一例を示すフローチャート（その 4）である。

【図 22】図 22 は、実施例 1 における管理処理手順の一例を示すフローチャート（その 5）である。

【図 23】図 23 は、管理装置 100 の動作例 2 を示す説明図（その 1）である。

【図 24】図 24 は、管理装置 100 の動作例 2 を示す説明図（その 2）である。

【図 25】図 25 は、管理装置 100 の動作例 2 を示す説明図（その 3）である。

30

【図 26】図 26 は、管理装置 100 の動作例 2 を示す説明図（その 4）である。

【図 27】図 27 は、管理装置 100 の動作例 2 を示す説明図（その 5）である。

【図 28】図 28 は、実施例 2 における管理処理手順の一例を示すフローチャート（その 1）である。

【図 29】図 29 は、実施例 2 における管理処理手順の一例を示すフローチャート（その 2）である。

【図 30】図 30 は、実施例 2 における管理処理手順の一例を示すフローチャート（その 3）である。

【図 31】図 31 は、実施例 2 における管理処理手順の一例を示すフローチャート（その 4）である。

40

【発明を実施するための形態】

【0010】

以下に、図面を参照して、本発明にかかる管理装置、制御方法、および管理プログラムの実施の形態を詳細に説明する。

【0011】

（実施の形態にかかる情報処理システム 110 の制御方法の一実施例）

図 1 は、実施の形態にかかる情報処理システム 110 の制御方法の一実施例を示す説明図である。管理装置 100 は、複数の情報処理装置 101 を備える情報処理システム 110 を管理するコンピュータである。情報処理装置 101 は、複数の計算資源を有し、仮想マシンを実行可能なコンピュータである。情報処理装置 101 は、例えば、データセンタ

50

ーにあるサーバである。

【0012】

計算資源は、例えば、ハードウェアリソースである。ハードウェアリソースは、例えば、サーバのCPU(Central Processing Unit)リソースなどである。CPUリソースは、例えば、仮想マシンに割り当てるCPUの数、CPU内のコアの数、または、仮想CPUの数などによって表現される。ハードウェアリソースは、例えば、サーバの記憶容量、ネットワーク帯域などであってもよい。

【0013】

以下の説明では、仮想マシンを「VM」と表記する場合がある。また、以下の説明では、ハードウェアリソースを「HWリソース」と表記する場合がある。

10

【0014】

データセンターを利用させるサービスでは、サービスの提供者とサービスの利用者との間で取り決められたSLAを満たすようにして、サービスの利用者が望んでいるユーザビリティを確保することが求められる傾向がある。また、データセンターの消費電力を低減すること、または、データセンターで提供可能なサービスの数を増大することなどのために、サービスに用いられるHWリソースを最小化することも求められる傾向がある。このように、データセンターを利用させるサービスでは、SLAを満たしつつ、そのサービスに用いられるHWリソースを低減することが好ましい。

【0015】

これに対し、データセンターにあるサーバによって実行され、利用者によって利用されるVM毎に、そのVMに割り当てられるHWリソースをSLAを満たす範囲で調整し、そのVMに割り当てられるHWリソースを最小化しようとするのが考えられる。しかしながら、VM毎にSLAを満たす範囲でHWリソースを最小化しようとしても、データセンター全体でサービスのために用いられるHWリソースを最小化することができないことがある。

20

【0016】

例えば、複数のVMを用いて1つのサービスを提供する場合、あるVMに割り当てるHWリソースを増減すると、他のVMの動作にも影響し、他のVMに最小でどのくらいのHWリソースを割り当てればよいかも変動してしまう傾向がある。具体的には、あるVMがSLAに対するボトルネックである場合、そのVMに割り当てるHWリソースを調整せずに固定したまま、他のVMに割り当てるHWリソースを低減しようとしても、他のVMに割り当てるHWリソースを最小化することは難しい。

30

【0017】

そこで、本実施の形態では、サービスを提供する複数のVMに割り当てるHWリソースと性能情報とをマッピングすることにより、そのサービスを提供する情報処理システム110全体で用いられる計算資源の最小化を図ることができる制御方法について説明する。

【0018】

図1において、アプリケーションaが動作するVMA、アプリケーションbが動作するVMB、アプリケーションcが動作するVMCがある。そして、複数の情報処理装置101がそれぞれ属し、アプリケーションa~cによって所定のサービスをそれぞれ提供する複数のグループがある。

40

【0019】

図1の例では、グループG1~G3がある。グループG1は、例えば、VMAを実行する情報処理装置101Aと、VMBを実行する情報処理装置101Bと、VMCを実行する情報処理装置101Cとが属するグループである。グループG2は、例えば、VMAを実行する情報処理装置101Aと、VMBを実行する情報処理装置101Dと、VMCを実行する情報処理装置101Eとが属するグループである。グループG1、G2のように、異なるグループに、同じ情報処理装置101が属する場合があってもよい。

【0020】

グループG3は、例えば、VMAを実行する情報処理装置101Aと、VMBを実行す

50

る情報処理装置101Bと、VMCを実行する情報処理装置101Cとが属するグループである。グループG3では、情報処理装置101Aは、グループG1に属する場合とは異なるHWリソースを、VMAに割り当てる。グループG1, G3のように、異なるグループでVMに割り当てられるHWリソースの大きさが異なれば、異なるグループに属する情報処理装置101が同じ組み合わせである場合があってもよい。

【0021】

管理装置100は、複数のグループの各々がそれぞれ実行するホストOS (Operating System) から、各情報処理装置101の構成情報と性能情報とを取得し、複数のグループ毎に記憶部に記憶する。構成情報は、例えば、CPU数、記憶容量、ネットワーク帯域などである。性能情報は、例えば、情報処理装置101のレイテンシやスループットなどの状態を示す性能値である。性能情報は、例えば、ベンチマークの実行によって求められる。

10

【0022】

管理装置100は、例えば、それぞれのグループに属する各情報処理装置101においてVMA~VMCに用いられるCPU数、記憶容量、ネットワーク帯域などを示す構成情報を取得する。また、管理装置100は、例えば、それぞれのグループにVMA~VMCを実行させ、そのグループに属する情報処理装置101が実行するホストOS上のパフォーマンスモニタから、情報処理装置101のレイテンシやスループットなどを示す性能情報を取得する。

【0023】

これにより、管理装置100は、それぞれのグループについて、そのグループに属する各情報処理装置101においてVMA~VMCに用いられるHWリソースの大きさと、そのグループで提供するサービスの性能情報とをマッピングすることができる。このため、管理装置100は、1つ1つのVMに着目してSLAを満たすようにHWリソースを最小化しようとする場合に比べて、VMにHWリソースを割り当てるパリエーションをより多く考慮して、HWリソースの最小化を検討することができる。

20

【0024】

管理装置100は、記憶部が複数のグループ毎に記憶する構成情報と性能情報とに基づき、選択対象のVMが属する対象グループから他のグループへの移動対象となる移動対象VMを複数のグループ毎に選択する。選択対象のVMは、サービスを提供する複数のVMである。対象グループは、現在、選択対象のVMを実行中の情報処理装置101が属するグループである。管理装置100は、例えば、それぞれのグループについて、そのグループに属する各情報処理装置101にVMA~VMCを配置する場合、HWリソースの割り当てが変更されるVMを選択する。

30

【0025】

管理装置100は、移動対象VMを他のグループに移動させた場合、移動対象VMに対応するサービスの性能値が予め定められた基準値を満たすか否かを判断する。移動対象のVMに対応するサービスは、他のグループに属する各情報処理装置101で実行される複数のVMで提供されるサービスであり、対象グループに属する各情報処理装置101で実行していた複数のVMが提供するサービスと同じサービスである。性能値は、例えば、レイテンシやスループットである。管理装置100は、他のグループに属する各情報処理装置101の性能情報に基づいて、VMA~VMCで提供するサービスの性能値を算出し、その性能値がSLAを満たすか否かを判定する。

40

【0026】

管理装置100は、基準値を満たすと判断したとき、対象グループから他のグループへの移動対象VMの移動を決定する。管理装置100は、例えば、VMA~VMCを、他のグループに属する各情報処理装置101に配置する。これにより、管理装置100は、SLAを満たす範囲で、対象グループより、サービスのために用いられるHWリソースが小さい可能性があるグループを特定することができる。そして、管理装置100は、VMA~VMCを、SLAを満たす他のグループに移動して、HWリソースの最小化を図ること

50

ができる。

【0027】

ここでは、グループが、各情報処理装置101がそれぞれ属し、各情報処理装置101が複数のVMに割り当て可能であるHWリソースのバリエーションを示している場合について説明したが、これに限らない。例えば、グループが、それぞれのVMにVM種別が設定され、各情報処理装置101がそれぞれのVMに設定されたVM種別に対応したHWリソースを割り当てるバリエーションを示している場合があってもよい。これによれば、管理装置100は、考慮するバリエーションの数を低減して、HWリソースの最小化を検討する際にかかる負担の低減化を図ることができる。

【0028】

ここでは、予めグループがある場合について説明したが、これに限らない。例えば、管理装置100が、グループを作成する場合があってもよい。また、情報処理システム110の管理者が、グループを管理装置100に入力する場合があってもよい。

【0029】

(情報処理システム110の一例)

次に、図2を用いて、図1に示した管理装置100を適用した、情報処理システム110の一例について説明する。

【0030】

図2は、情報処理システム110の一例を示す説明図である。図2において、情報処理システム110は、管理装置100と、情報処理装置101とを含む。情報処理システム110において、管理装置100と情報処理装置101とは、有線または無線のネットワーク210を介して接続される。ネットワーク210は、例えば、LAN(Local Area Network)、WAN(Wide Area Network)、インターネットなどである。

【0031】

管理装置100は、情報処理装置101を管理するコンピュータである。管理装置100は、具体的には、サーバである。管理装置100は、データセンターにあるサーバである。管理装置100は、データセンターの外にあるサーバであってもよい。情報処理装置101は、計算資源を有するコンピュータである。情報処理装置101は、VMを実行することができる。情報処理装置101は、例えば、ホストOSによってVMを実行する。情報処理装置101は、例えば、ハイパーバイザによってVMを実行してもよい。情報処理装置101は、具体的には、サーバである。管理装置100は、いずれかの情報処理装置101と一体であってもよい。

【0032】

(情報処理システム110の使用例)

次に、図3を用いて、情報処理システム110の使用例について説明する。

【0033】

図3は、情報処理システム110の使用例を示す説明図である。図3において、情報処理システム110は、フロントエンドサーバの機能を実現するVMとアプリケーションサーバの機能を実現するVMとバックエンドサーバの機能を実現するVMとの連携により、クライアント装置303を介して利用者にサービスを提供する。

【0034】

情報処理システム110は、フロントエンドサーバの機能を実現するVMを、いずれかの情報処理装置101に配置し、フロントエンドサーバを実現している。フロントエンドサーバは、例えば、インターネット301とファイアウォール302とを介して、クライアント装置303と通信するサーバである。情報処理システム110は、アプリケーションサーバの機能を実現するVMを、いずれかの情報処理装置101に配置し、アプリケーションサーバを実現している。アプリケーションサーバは、例えば、クライアント装置303からの通信内容に応じて、所定の処理を行うサーバである。情報処理システム110は、バックエンドサーバの機能を実現するVMを、いずれかの情報処理装置101に配置

10

20

30

40

50

し、バックエンドサーバを実現している。バックエンドサーバは、例えば、DB (Data Base)などを管理するサーバである。

【0035】

管理装置100は、それぞれの情報処理装置101がVMに割り当てるHWリソースを制御し、または、複数の情報処理装置101のうちのVMの配置を制御する。管理装置100は、VMに割り当てるHWリソースを制御し、または、複数の情報処理装置101のうちのVMの配置を制御しつつ、それぞれの情報処理装置101がVMに割り当てたHWリソースの情報と、サービスの性能情報とを取得する。管理装置100は、取得したHWリソースの情報と、サービスの性能情報とに基づいて、SLAを満たす範囲で、それぞれの情報処理装置101がVMに割り当てるHWリソースを制御し、または、複数の情報処理装置101のうちのVMの配置を制御する。

10

【0036】

(管理装置100のハードウェア構成例)

次に、図4を用いて、管理装置100のハードウェア構成例について説明する。

【0037】

図4は、管理装置100のハードウェア構成例を示すブロック図である。図4において、管理装置100は、CPU401と、メモリ402と、ネットワークI/F (Interface) 403と、記録媒体I/F 404と、記録媒体405とを有する。また、各構成部は、バス400によってそれぞれ接続される。

【0038】

ここで、CPU401は、管理装置100の全体の制御を司る。メモリ402は、例えば、ROM (Read Only Memory)、RAM (Random Access Memory)およびフラッシュROMなどを有する。具体的には、例えば、フラッシュROMやROMが各種プログラムを記憶し、RAMがCPU401のワークエリアとして使用される。メモリ402に記憶されるプログラムは、CPU401にロードされることで、コーディングされている処理をCPU401に実行させる。

20

【0039】

ネットワークI/F 403は、通信回線を通じてネットワーク210に接続され、ネットワーク210を介して他のコンピュータに接続される。そして、ネットワークI/F 403は、ネットワーク210と内部のインターフェースを司り、他のコンピュータからのデータの入出力を制御する。ネットワークI/F 403には、例えば、モデムやLANアダプタなどを採用することができる。

30

【0040】

記録媒体I/F 404は、CPU401の制御にしたがって記録媒体405に対するデータのリード/ライトを制御する。記録媒体I/F 404は、例えば、ディスクドライブ、SSD (Solid State Drive)、USB (Universal Serial Bus)ポートなどである。記録媒体405は、記録媒体I/F 404の制御で書き込まれたデータを記憶する不揮発メモリである。記録媒体405は、例えば、ディスク、半導体メモリ、USBメモリなどである。記録媒体405は、管理装置100から着脱可能であってもよい。

40

【0041】

管理装置100は、上述した構成部のほか、例えば、キーボード、マウス、ディスプレイなどを有してもよい。また、管理装置100は、記録媒体I/F 404や記録媒体405を有していなくてもよい。また、情報処理装置101のハードウェア構成例は、管理装置100のハードウェア構成例と同様であるため、説明を省略する。

【0042】

(計算資源管理テーブル500の記憶内容)

次に、図5を用いて、計算資源管理テーブル500の記憶内容について説明する。計算資源管理テーブル500は、例えば、図4に示した管理装置100のメモリ402や記録媒体405などの記憶領域により実現される。

50

【 0 0 4 3 】

図5は、計算資源管理テーブル500の記憶内容の一例を示す説明図である。図5に示すように、計算資源管理テーブル500は、物理ホストと、CPUと、MEMとのフィールドを有する。計算資源管理テーブル500は、物理ホスト毎に各フィールドに情報を設定することにより、計算資源管理情報がレコードとして記憶される。

【 0 0 4 4 】

物理ホストのフィールドは、VMを実行可能な情報処理装置101のアドレスが設定される。CPUのフィールドは、その情報処理装置101が有するCPUの数が設定される。MEMのフィールドは、その情報処理装置101が有するメモリ402の大きさが設定される。これにより、計算資源管理テーブル500は、情報処理装置101にどのくらいのHWリソースを用いるVMであれば配置することができるかを判定可能にすることができる。

10

【 0 0 4 5 】

(VM種別テーブル600の記憶内容)

次に、図6を用いて、VM種別テーブル600の記憶内容について説明する。VM種別テーブル600は、例えば、図4に示した管理装置100のメモリ402や記録媒体405などの記憶領域により実現される。

【 0 0 4 6 】

図6は、VM種別テーブル600の記憶内容の一例を示す説明図である。図6に示すように、VM種別テーブル600は、VM種別と、CPUと、MEMと、NETとのフィールドを有する。VM種別テーブル600は、VM種別毎に各フィールドに情報を設定することにより、VM種別情報がレコードとして記憶される。

20

【 0 0 4 7 】

VM種別のフィールドは、VMに要求されるスペックに対応するVM種別が設定される。CPUのフィールドは、そのスペックを実現するために、そのVMに割り当てるCPUの条件が設定される。CPUの条件は、コアの数であってもよい。MEMのフィールドは、そのスペックを実現するために、そのVMに割り当てるメモリ402の大きさの条件が設定される。NETのフィールドは、そのスペックを実現するために、そのVMに割り当てるネットワーク帯域の大きさの条件が設定される。これにより、VM種別テーブル600は、あるVM種別のVMに用いられるHWリソースの大きさを特定可能にすることができる。

30

【 0 0 4 8 】

(VM稼働状況シート700の記憶内容)

次に、図7を用いて、VM稼働状況シート700の記憶内容について説明する。VM稼働状況シート700は、例えば、図4に示した管理装置100のメモリ402や記録媒体405などの記憶領域により実現される。

【 0 0 4 9 】

図7は、VM稼働状況シート700の記憶内容の一例を示す説明図である。図7に示すように、VM稼働状況シート700は、1以上のVMと、合計HWリソースと、サービス性能と、目標サービス性能とのフィールドを有する。VM稼働状況シート700は、サービスを実現する複数のVMに割り当てられるHWリソースのバリエーション毎に、各フィールドに情報を設定することにより、VM稼働状況情報がレコードとして記憶される。

40

【 0 0 5 0 】

VMのフィールドは、サービスを実現するそれぞれのVMについて設けられる。VMのフィールドは、さらに、VM種別と、CPUと、MEMと、NETとのフィールドを有する。VM種別のフィールドは、VMに要求されるスペックに対応するVM種別が設定される。CPUのフィールドは、そのスペックを実現するために、そのVMに割り当てるCPUの条件が設定される。MEMのフィールドは、そのスペックを実現するために、そのVMに割り当てるメモリ402の大きさの条件が設定される。NETのフィールドは、そのスペックを実現するために、そのVMに割り当てるネットワーク帯域の大きさの条件が設

50

定される。

【 0 0 5 1 】

合計HWリソースのフィールドは、さらに、CPUと、MEMと、NETとのフィールドを有する。CPUのフィールドは、サービスを実現する複数のVMを実行するために、各情報処理装置101で用いられるCPUの数の合計値が設定される。MEMのフィールドは、サービスを実現する複数のVMを実行するために、各情報処理装置101で用いられるメモリ402の大きさの合計値が設定される。NETのフィールドは、サービスを実現する複数のVMを実行するために、各情報処理装置101で用いられるネットワーク帯域の大きさの合計値が設定される。これにより、VM稼働状況シート700は、あるサービスを提供する複数のVMにHWリソースを割り当てるバリエーションをより多く考慮可能にすることができる。VM稼働状況シート700の具体例は、図10～図17や図23～図27に後述する。

10

【 0 0 5 2 】

(管理装置100の機能的構成例)

次に、図8を用いて、管理装置100の機能的構成例について説明する。

【 0 0 5 3 】

図8は、管理装置100の機能的構成例を示すブロック図である。図8に示すように、管理装置100は、記憶部801と、取得部802と、選択部803と、決定部804と、出力部805とを含む。

【 0 0 5 4 】

記憶部801は、例えば、図4に示したメモリ402や記録媒体405などの記憶領域によって実現される。取得部802～出力部805は、制御部となる機能である。取得部802～出力部805は、具体的には、例えば、図3に示したメモリ402や記録媒体405などの記憶領域に記憶されたプログラムをCPU401に実行させることにより、または、ネットワークI/F403により、その機能を実現する。各機能部の処理結果は、例えば、図3に示したメモリ402や記録媒体405などの記憶領域に記憶される。

20

【 0 0 5 5 】

記憶部801は、複数のグループ毎に、各情報処理装置101の構成情報と性能情報とを記憶する。情報処理装置101は、複数の計算資源を有し、情報処理システム110に備えられる。グループは、複数の情報処理装置101がそれぞれ属し、所定のサービスを提供する。記憶部801は、例えば、VM稼働状況シート700である。構成情報は、例えば、CPU数、記憶容量、ネットワーク帯域などである。性能情報は、例えば、情報処理装置101のレイテンシやスループットなどの状態を示す性能値である。性能情報は、例えば、ベンチマークの実行によって求められる。計算資源は、HWリソースである。

30

【 0 0 5 6 】

これにより、記憶部801は、あるサービスを提供する複数のVMにHWリソースを割り当てるバリエーションを、比較的多く記憶しておくことができる。このため、記憶部801は、1つ1つのVMに着目してSLAを満たすようにHWリソースを最小化しようとする場合に比べて、VMにHWリソースを割り当てるバリエーションをより多く考慮可能にすることができる。

40

【 0 0 5 7 】

取得部802は、複数のグループの各々がそれぞれ実行するホストOSから、各情報処理装置101の構成情報と性能情報とを取得し、複数のグループ毎に記憶部801に記憶する。取得部802は、例えば、それぞれのグループに属する各情報処理装置101において複数のVMに用いられるCPU数、記憶容量、ネットワーク帯域などを示す構成情報を取得する。また、取得部802は、例えば、それぞれのグループに複数のVMを実行させ、そのグループに属する情報処理装置101が実行するホストOS上のパフォーマンスモニタから、情報処理装置101のレイテンシやスループットなどの状態を示す性能情報を取得する。

【 0 0 5 8 】

50

これにより、取得部 802 は、それぞれのグループについて、そのグループに属する各情報処理装置 101 において複数の VM に用いられる HW リソースの大きさと、そのグループで提供するサービスの性能情報とをマッピングすることができる。このため、取得部 802 は、どのグループが SLA を満たすか、どのグループの HW リソースが比較的小さいかを特定可能にすることができる。

【0059】

取得部 802 は、取得済みの構成情報と性能情報とに基づき、性能情報を未取得であるグループについて性能情報を取得するか否かを判定し、取得すると判定したグループについて性能情報を取得して記憶部 801 に記憶してもよい。

【0060】

取得部 802 は、例えば、複数の VM のうち、ある VM に割り当てる HW リソースを可変にし、他の VM に割り当てる HW リソースを固定した場合についての 1 以上のグループを特定する。取得部 802 は、特定した 1 以上のグループのうち、最も HW リソースが小さいグループから、最も HW リソースが大きいグループに変更した場合に、10% 以上性能値が良好になるか否かを判定する。取得部 802 は、良好にならなければ、特定した 1 以上のグループの残りのグループについては性能情報の取得を省略する。これにより、取得部 802 は、情報処理装置 101 に VM を配置してベンチマークを実行して性能情報を取得する際にかかる時間を削減し、管理装置 100 にかかる負担を低減することができる。

【0061】

取得部 802 は、例えば、複数の VM のうち、ある VM に割り当てる HW リソースを可変にし、他の VM に割り当てる HW リソースを固定した場合についての 1 以上のグループを特定する。取得部 802 は、特定した 1 以上のグループのうち、最も HW リソースが小さいグループから順に SLA を満たすまで性能情報を取得し、SLA を満たした時点で性能情報が未取得のグループについては性能情報の取得を省略する。これにより、取得部 802 は、情報処理装置 101 に VM を配置してベンチマークを実行して性能情報を取得する際にかかる時間を削減し、管理装置 100 にかかる負担を低減することができる。

【0062】

取得部 802 は、例えば、複数の VM のうち、ある VM に割り当てる HW リソースを可変にし、他の VM に割り当てる HW リソースを固定した場合についての 1 以上のグループを特定する。取得部 802 は、特定した 1 以上のグループのうち、最も HW リソースが大きいグループから順に SLA を満たさなくなるまで性能情報を取得し、SLA を満たさなくなった時点で性能情報が未取得のグループについては性能情報の取得を省略する。これにより、取得部 802 は、情報処理装置 101 に VM を配置してベンチマークを実行して性能情報を取得する際にかかる時間を削減し、管理装置 100 にかかる負担を低減することができる。

【0063】

取得部 802 は、VM 種別テーブル 600 に基づいて、各情報処理装置 101 の構成情報に応じた VM 種別を複数のグループ毎に記憶部 801 に記憶してもよい。VM 種別は、VM のスペックを示す種別である。取得部 802 は、例えば、VM 種別の単位で規定される複数のグループにおける複数の VM のそれぞれの VM についての VM 種別の組み合わせを取得し、記憶部 801 に保存する。

【0064】

これにより、取得部 802 は、それぞれのグループについて、そのグループに属する各情報処理装置 101 において複数の VM に用いられる HW リソースの大きさと、そのグループで提供するサービスの性能情報とをマッピングすることができる。このため、取得部 802 は、どのグループが SLA を満たすか、どのグループの HW リソースが比較的小さいかを特定可能にすることができる。取得部 802 は、VM 種別の単位で複数の VM に割り当てる HW リソースの組み合わせのバリエーションを考慮するため、考慮するバリエーションの数を低減することができ、HW リソースの最小化を検討する際にかかる負担の低

10

20

30

40

50

減化を図ることができる。

【 0 0 6 5 】

選択部 8 0 3 は、記憶部 8 0 1 が複数のグループ毎に記憶する構成情報と性能情報とに基づき、選択対象の VM が属する対象グループから他のグループへの移動対象となる移動対象 VM を複数のグループ毎に選択する。選択対象の VM は、サービスを提供する複数の VM である。対象グループは、現在、選択対象の VM を実行中の情報処理装置 1 0 1 が属するグループである。移動対象 VM は、複数の VM をグループ間で移動する際に、複数の VM のうちで HW リソースの割り当てが変更される VM である。

【 0 0 6 6 】

選択部 8 0 3 は、例えば、計算資源管理テーブル 5 0 0 に基づいて、それぞれのグループについて、そのグループに属する各情報処理装置 1 0 1 に複数の VM を配置する場合、HW リソースの割り当てが変更される VM を選択する。これにより、選択部 8 0 3 は、それぞれのグループで複数の VM を実行するために、どの VM の HW リソースの割り当てを変更すればよいかを特定することができる。

【 0 0 6 7 】

選択部 8 0 3 は、記憶部 8 0 1 が複数のグループ毎に記憶する VM 種別と性能情報とに基づき、移動対象 VM を複数のグループ毎に選択してもよい。選択部 8 0 3 は、例えば、計算資源管理テーブル 5 0 0 と、VM 種別テーブル 6 0 0 とに基づいて、あるグループに規定された VM 種別で複数の VM を実行する際に、HW リソースの割り当てが変更される VM を選択する。これにより、選択部 8 0 3 は、それぞれのグループでの VM 種別の組み合わせに対応して、サービスを提供する複数の VM を実行するためには、どの VM の HW リソースの割り当てを変更すればよいかを特定することができる。

【 0 0 6 8 】

決定部 8 0 4 は、移動対象 VM を他のグループに移動させた場合、移動対象 VM に対応するサービスの性能値が予め定められた基準値を満たすか否かを判断する。そして、決定部 8 0 4 は、基準値を満たすと判断したとき、対象グループから他のグループへの移動対象 VM の移動を決定する。移動対象の VM に対応するサービスは、他のグループに属する各情報処理装置 1 0 1 で実行される複数の VM で提供されるサービスであり、対象グループに属する各情報処理装置 1 0 1 で実行していた複数の VM が提供するサービスと同じサービスである。サービスの性能値は、例えば、レイテンシやスループットである。サービスの性能値は、例えば、各情報処理装置 1 0 1 のレイテンシやスループットなどから求められる、そのサービスの利用者が用いるクライアント装置に対するレイテンシやスループットなどである。

【 0 0 6 9 】

決定部 8 0 4 は、例えば、他のグループに属する各情報処理装置 1 0 1 の性能情報に基づいて、複数の VM で提供するサービスの性能値を算出し、その性能値が S L A を満たすか否かを判断する。そして、決定部 8 0 4 は、例えば、複数の VM を、他のグループに属する各情報処理装置 1 0 1 に配置すると決定する。これにより、決定部 8 0 4 は、S L A を満たす範囲で、対象グループより、サービスのために用いられる HW リソースが小さくなる可能性があるグループを特定することができる。そして、管理装置 1 0 0 は、V M A ~ V M C を、S L A を満たす他のグループに移動して、HW リソースの最小化を図ることができる。

【 0 0 7 0 】

決定部 8 0 4 は、さらに、移動対象 VM を他のグループに移動させた場合、移動対象 VM に対応するサービスの性能値が、移動前の移動対象 VM に対応するサービスの性能値よりも良好であることを示すか否かを判断してもよい。そして、決定部 8 0 4 は、基準値を満たし、かつ、良好であることを示すとき、対象グループから他のグループへの移動対象 VM の移動を決定する。性能値は、例えば、レイテンシであれば、小さい値の方が良好である。性能値は、例えば、スループットであれば、大きい値の方が良好である。これにより、決定部 8 0 4 は、現在よりもサービスの性能値が良好になる場合に VM を移動するた

10

20

30

40

50

め、効率よくサービスの性能値の向上を図ることができる。

【0071】

決定部804は、さらに、対象グループの計算資源より他のグループの計算資源が少ないか否かを判断してもよい。そして、決定部804は、少ないとき、対象グループから他のグループへの移動対象VMの移動を決定する。これにより、決定部804は、現在よりもサービスに用いられるHWリソースが小さくなる場合にVMを移動するため、効率よくサービスに用いられる計算資源の最小化を図ることができる。

【0072】

出力部805は、VMへのHWリソースの割り当ての変更指示、または、VMの配置指示などを出力する。出力部805は、VMへのHWリソースの割り当ての変更の結果、または、VMの配置の結果などを出力してもよい。出力部805は、各情報処理装置101の性能情報、または、サービスの性能値を出力してもよい。出力形式は、例えば、ディスプレイへの表示、プリンタへの印刷出力、ネットワークI/F403による外部装置への送信、または、メモリ402や記録媒体405などの記憶領域への記憶である。

【0073】

(管理装置100の機能的構成の具体例)

次に、図9を用いて、管理装置100の機能的構成の具体例について説明する。

【0074】

図9は、管理装置100の機能的構成の具体例を示すブロック図である。図9に示すように、管理装置100は、VM管理ソフトウェア900によって、DB901と、グループ定義902と、サービス性能情報903と、タイマー部904と、データ取得部905と、移動対象選択部906と、制御方針決定部907とを管理する。また、管理装置100は、インフラ管理ソフトウェア910によって、制御受け取り部911と、制御指示部912とを管理する。

【0075】

DB901は、HWリソース管理テーブル、VM種別テーブル600、および、VM稼働状況シート700などを記憶する記憶領域である。グループ定義902は、情報処理システム110の管理者によって入力された、サービスを提供する複数のVMが含まれるVMグループを定義する情報である。サービス性能情報903は、そのサービスの現在のサービス性能を示す情報である。サービス性能情報903は、例えば、データ取得部905によって取得される。

【0076】

タイマー部904は、一定時間毎に、要求入力をデータ取得部905に出力する。データ取得部905は、DB901に基づいて、VM種別の組み合わせと、その組み合わせで提供するサービスのサービス性能を示すサービス性能情報903とを取得する。データ取得部905は、取得した、VM種別の組み合わせと、その組み合わせで提供するサービスのサービス性能を示すサービス性能情報903とに基づいて、DB901を更新する。

【0077】

移動対象選択部906は、DB901に基づいて、移動対象のVMを選択する。制御方針決定部907は、選択された移動対象のVMに割り当てるHWリソースについての制御方針を決定する。制御方針決定部907は、例えば、選択されたVMに割り当てるCPUリソースの大きさを決定し、または、選択されたVMを移動させる情報処理装置101を決定する。制御方針決定部907は、選択したVMと、決定した制御方針とを、制御受け取り部911に出力する。

【0078】

制御受け取り部911は、制御方針決定部907から制御対象のVMと制御方針とを受け付け、制御指示部912に出力する。制御指示部912は、制御対象のVMを実行中の情報処理装置101に対して、制御方針にしたがって制御信号を出力する。情報処理装置101上で動作する仮想化ソフトウェア921、922などは、制御信号を受信し、その制御信号にしたがって情報処理装置101の仮想化環境を制御し、VMへのHWリソース

10

20

30

40

50

の割り当てなどを制御する。これにより、制御指示部 912 は、情報処理装置 101 を制御し、VM に割り当てる HW リソースを制御することができる。

【0079】

(管理装置 100 の動作例 1)

次に、図 10 ~ 図 17 を用いて、管理装置 100 の動作例 1 について説明する。

【0080】

図 10 ~ 図 17 は、管理装置 100 の動作例 1 を示す説明図である。動作例 1 は、管理装置 100 が、性能情報としてレイテンシ [usec] を用いる場合についての動作例である。

【0081】

図 10 において、管理装置 100 は、VM 種別の単位で、HW リソースの組み合わせを複数作成し、VM 稼働状況シート 700 を作成する。管理装置 100 は、例えば、VM 種別テーブル 600 に基づいて、VMA に関する 5 つの VM 種別と、VMB に関する 5 つの VM 種別と、VMC に関する 5 つの VM 種別とを組み合わせ、125 の組み合わせを作成する。VM 種別は、例えば、Low Spec、CPU Rich、MEM Rich、NET Rich、High Spec などである。管理装置 100 は、作成したそれぞれの組み合わせを、VM 稼働状況シート 700 のそれぞれのレコードの VMA ~ VMC のフィールドに設定することにより記憶する。

【0082】

これにより、管理装置 100 は、あるサービスを提供する複数の VM に HW リソースを割り当てるバリエーションを、比較的多く考慮可能にすることができる。また、管理装置 100 は、VM 種別の単位で複数の VM に HW リソースを割り当てるバリエーションを考慮するため、考慮するバリエーションの数を低減し、HW リソースの最小化を検討する際にかかる負担の低減化を図ることができる。次に、図 11 の説明に移行する。

【0083】

図 11 において、管理装置 100 は、VMA の VM 種別を可変にし、VMB や VMC の VM 種別を固定した 1 以上の組み合わせのうち、最も HW リソースが小さい組み合わせを選択する。図 11 の例では、管理装置 100 は、VM 稼働状況シート 700 のレコード 1100 を選択する。管理装置 100 は、選択したレコード 1100 に基づいて、VMA ~ VMC の VM 種別が変更されるように、VMA ~ VMC を複数の情報処理装置 101 に配置し、ベンチマークを実行する。

【0084】

その後、管理装置 100 は、サービスに用いられる HW リソースの大きさの情報として、6 CPU、24 GB、1 Gbit/s x 3 などを取得し、VM 稼働状況シート 700 の合計 HW リソースのフィールドに設定することにより記憶する。管理装置 100 は、サービスのレイテンシ 48 usec を取得し、VM 稼働状況シート 700 のサービス性能のフィールドに設定することにより記憶する。次に、図 12 の説明に移行する。

【0085】

図 12 において、管理装置 100 は、VMA の VM 種別を可変にし、VMB や VMC の VM 種別を固定した 1 以上の組み合わせのうち、最も HW リソースが大きい組み合わせを選択する。図 12 の例では、管理装置 100 は、VM 稼働状況シート 700 のレコード 1200 を選択する。管理装置 100 は、選択したレコード 1200 に基づいて、VMA ~ VMC の VM 種別が変更されるように、VMA ~ VMC を複数の情報処理装置 101 に配置し、ベンチマークを実行する。

【0086】

その後、管理装置 100 は、サービスに用いられる HW リソースの大きさの情報として、8 CPU、32 GB、1 Gbit/s x 2 + 10 Gbit/s x 1 などを取得し、VM 稼働状況シート 700 の合計 HW リソースのフィールドに設定することにより記憶する。管理装置 100 は、サービスのレイテンシ 38 usec を取得し、VM 稼働状況シート 700 のサービス性能のフィールドに設定することにより記憶する。次に、図 13 の説明に

10

20

30

40

50

移行する。

【0087】

図13において、管理装置100は、レコード1100に対応する最もHWリソースが小さい組み合わせから、レコード1200に対応する最もHWリソースが大きい組み合わせに変更した場合に、サービスのレイテンシが10%以上改善されるか否かを判定する。管理装置100は、10%以上改善されるため、レコード1100に対応する最もHWリソースが小さい組み合わせと、レコード1200に対応する最もHWリソースが大きい組み合わせとの間にある、他の組み合わせについての処理を省略しない。

【0088】

図13の例では、管理装置100は、VM稼働状況シート700のレコード1300～1302を順次選択する。まず、管理装置100は、選択したレコード1300に基づいて、VMA～VMCのVM種別が変更されるように、VMA～VMCを複数の情報処理装置101に配置し、ベンチマークを実行する。

10

【0089】

その後、管理装置100は、サービスに用いられるHWリソースの大きさの情報として、8CPU、24GB、1Gbit/s×3などを取得し、VM稼働状況シート700の合計HWリソースのフィールドに設定することにより記憶する。管理装置100は、サービスのレイテンシ48usecを取得し、VM稼働状況シート700のサービス性能のフィールドに設定することにより記憶する。

【0090】

20

また、管理装置100は、選択したレコード1301に基づいて、VMA～VMCのVM種別が変更されるように、VMA～VMCを複数の情報処理装置101に配置し、ベンチマークを実行する。

【0091】

その後、管理装置100は、サービスに用いられるHWリソースの大きさの情報として、6CPU、32GB、1Gbit/s×3などを取得し、VM稼働状況シート700の合計HWリソースのフィールドに設定することにより記憶する。管理装置100は、サービスのレイテンシ46usecを取得し、VM稼働状況シート700のサービス性能のフィールドに設定することにより記憶する。

【0092】

30

また、管理装置100は、選択したレコード1302に基づいて、VMA～VMCのVM種別が変更されるように、VMA～VMCを複数の情報処理装置101に配置し、ベンチマークを実行する。

【0093】

その後、管理装置100は、サービスに用いられるHWリソースの大きさの情報として、6CPU、24GB、1Gbit/s×2+10Gbit/s×1などを取得し、VM稼働状況シート700の合計HWリソースのフィールドに設定することにより記憶する。管理装置100は、サービスのレイテンシ45usecを取得し、VM稼働状況シート700のサービス性能のフィールドに設定することにより記憶する。次に、図14の説明に移行する。

40

【0094】

図14において、管理装置100は、VMBのVM種別を可変にし、VMAやVMCのVM種別を固定した1以上の組み合わせのうち、最もHWリソースが小さい組み合わせを選択する。図14の例では、管理装置100は、VM稼働状況シート700のレコード1400を選択する。管理装置100は、選択したレコード1400に基づいて、VMA～VMCのVM種別が変更されるように、VMA～VMCを複数の情報処理装置101に配置し、ベンチマークを実行する。

【0095】

その後、管理装置100は、サービスに用いられるHWリソースの大きさの情報として、6CPU、24GB、1Gbit/s×3などを取得し、VM稼働状況シート700の

50

合計HWリソースのフィールドに設定することにより記憶する。管理装置100は、サービスのレイテンシ48usecを取得し、VM稼働状況シート700のサービス性能のフィールドに設定することにより記憶する。次に、図15の説明に移行する。

【0096】

図15において、管理装置100は、VMBのVM種別を可変にし、VMAやVMCのVM種別を固定した1以上の組み合わせのうち、最もHWリソースが大きい組み合わせを選択する。図15の例では、管理装置100は、VM稼働状況シート700のレコード1500を選択する。管理装置100は、選択したレコード1500に基づいて、VMA~VMCのVM種別が変更されるように、VMA~VMCを複数の情報処理装置101に配置し、ベンチマークを実行する。

10

【0097】

その後、管理装置100は、サービスに用いられるHWリソースの大きさの情報として、8CPU、32GB、1Gbit/s×2+10Gbit/s×1などを取得し、VM稼働状況シート700の合計HWリソースのフィールドに設定することにより記憶する。管理装置100は、サービスのレイテンシ45usecを取得し、VM稼働状況シート700のサービス性能のフィールドに設定することにより記憶する。次に、図16の説明に移行する。

【0098】

図16において、管理装置100は、レコード1400に対応する最もHWリソースが小さい組み合わせから、レコード1500に対応する最もHWリソースが大きい組み合わせに変更した場合に、サービスのレイテンシが10%以上改善されるか否かを判定する。管理装置100は、10%以上改善されないため、レコード1400に対応する最もHWリソースが小さい組み合わせと、レコード1500に対応する最もHWリソースが大きい組み合わせとの間にある、他の組み合わせについての処理を省略する。

20

【0099】

図16の例では、管理装置100は、VM稼働状況シート700のレコード1600~1602の、合計HWリソースのフィールド、および、サービス性能のフィールドには、「ベンチマーク実施しない」を設定する。これにより、管理装置100は、情報処理装置101にVMを配置してベンチマークを実行してレイテンシを取得する際にかかる時間を削減し、管理装置100にかかる負担を低減することができる。その後、管理装置100は、同様に、VM稼働状況シート700の、合計HWリソースのフィールド、および、サービス性能のフィールドを設定する。次に、図17の説明に移行する。

30

【0100】

図17において、管理装置100は、VM稼働状況シート700において、サービス性能が目標サービス性能を満たすVM種別の組み合わせのうち、合計HWリソースが最小になる組み合わせを選択する。目標サービス性能は、SLAに対応する性能値である。そして、管理装置100は、選択したVM種別の組み合わせで、VMA~VMCを複数の情報処理装置101に配置する。

【0101】

図17の例では、管理装置100は、具体的には、サービス性能のレイテンシが、目標サービス性能のレイテンシよりも小さくなり、SLAを満たすVM種別の組み合わせのうち、レコード1700に対応するVM種別の組み合わせを選択する。そして、管理装置100は、そのVM種別の組み合わせで、VMA~VMCを複数の情報処理装置101に配置する。

40

【0102】

これにより、管理装置100は、サービスに用いられるHWリソースを、SLAを満たす範囲で最小化することができる。このため、管理装置100は、サービスにかかる消費電力の低減化を図ることができる。また、管理装置100は、他のサービスに利用可能なHWリソースを確保し、情報処理システム110上でより多くのサービスを提供可能にすることができる。また、管理装置100は、あるサービスの負荷が急上昇した際に、その

50

サービスに割り当て可能な、空きHWリソースを確保しておき、情報処理システム110の安定性の向上を図ることができる。

【0103】

ここで、1つ1つのVMに着目してSLAを満たすようにHWリソースを最小化しようとする場合が考えられる。例えば、フロントエンドサーバの機能を実現するVMに割り当てられるHWリソースを考慮せず、アプリケーションサーバの機能を実現するVMに着目し、そのVMに割り当てられるHWリソースを最小化しようとしてしまう場合が考えられる。

【0104】

しかしながら、フロントエンドサーバのネットワーク帯域がボトルネックになり、アプリケーションサーバの機能を実現するVMに割り当てられるHWリソースを増大させても、サービスの性能値が良好になりにくいことがある。このため、SLAを満たすために、アプリケーションサーバの機能を実現するVMに、比較的大きいHWリソースを割り当てなければならないと判断してしまう可能性がある。また、サービスがSLAを満たさないと判断してしまう可能性もある。結果として、サービスに用いられるHWリソースを最小化することができない可能性がある。

10

【0105】

これに対し、管理装置100は、サービスを1つのVMで提供することのほか、サービスを複数のVMで提供することについても考慮することができる。管理装置100は、例えば、1つ1つのVMに着目してSLAを満たすようにHWリソースを最小化しようとする場合に比べて、VMにHWリソースを割り当てるバリエーションをより多く考慮することができる。このため、管理装置100は、サービスに用いられるHWリソースを、SLAを満たす範囲で、より小さくすることができる。

20

【0106】

ここでは、管理装置100が、VM種別の単位で、HWリソースの組み合わせを複数作成する場合について説明したが、これに限らない。例えば、管理装置100が、計算資源管理テーブル500に基づいて、それぞれのVMに割り当て可能なHWリソースの大きさを特定し、それぞれのVMに割り当て可能なHWリソースの組み合わせを複数作成する場合があってもよい。

【0107】

例えば、管理装置100は、それぞれのVMに割り当て可能なCPUの最小数と最大数とを特定し、最小数と最大数の間を所定単位で区切り、それぞれのVMに割り当てられるCPUの数になる複数の候補を特定する。また、管理装置100は、同様に、それぞれのVMに割り当てられる記憶容量になる複数の候補、ネットワーク帯域になる複数の候補を特定する。そして、管理装置100は、それぞれの候補を組み合わせ、それぞれのVMに割り当てられるHWリソースの組み合わせを複数作成し、VM稼働状況シート700に設定する。

30

【0108】

(実施例1における管理処理手順の一例)

次に、図18～図22を用いて、実施例1における管理処理手順の一例について説明する。

【0109】

図18～図22は、実施例1における管理処理手順の一例を示すフローチャートである。図18において、データ取得部905は、タイマー部904が発生する要求入力を検出したことに応じて、ユーザ定義されたサービスを運用するVMグループにおけるHWリソースの上限値を、OSから取得し、DB901に保存する(ステップS1801)。

40

【0110】

データ取得部905は、要求入力に応じて、ユーザ定義されたサービスを運用するVMグループに対して、サービス性能となるレイテンシと各VMのHWリソースの使用状況を取得して、DB901に保存する(ステップS1802)。

【0111】

データ取得部905は、移動対象となる物理ホストの使用状況を取得して、DB901

50

に保存する（ステップS1803）。データ取得部905は、VM種別の組み合わせを特定し、VM稼働状況シート700を作成し、DB901に保存する（ステップS1804）。管理装置100は、図19のステップS1901の処理に移行する。

【0112】

図19において、データ取得部905は、VMAのVM種別を可変にし、VMBとVMCとのVM種別を固定した、1以上の組み合わせのうち、VM稼働状況シート700上でサービス性能が未設定である組み合わせがあるか否かを判定する（ステップS1901）。ここで、未設定である組み合わせがない場合（ステップS1901：No）、管理装置100は、図20のステップS2001の処理に移行する。

【0113】

一方で、未設定である組み合わせがある場合（ステップS1901：Yes）、データ取得部905は、未設定である組み合わせのうち、VM稼働状況シート700上で最も上にある組み合わせを選択する（ステップS1902）。

【0114】

制御方針決定部907は、選択した組み合わせのVMAとVMBとVMCとのVM種別で、VMAとVMBとVMCとを配置可能な物理ホストが存在するか否かを判定する（ステップS1903）。ここで、存在しない場合（ステップS1903：No）、制御方針決定部907は、選択した組み合わせと同じVMBとVMCとのVM種別を含む組み合わせについて、サービス性能をベンチマーク実施しないに設定する（ステップS1904）。管理装置100は、図20のステップS2001の処理に移行する。

【0115】

一方で、存在する場合（ステップS1903：Yes）、制御方針決定部907は、ステップS1905の処理に移行する。制御方針決定部907は、選択した組み合わせのVMAとVMBとVMCとのVM種別で、VMAとVMBとVMCとを配置し、その組み合わせについて、レイテンシを取得してサービス性能を設定する（ステップS1905）。

【0116】

制御方針決定部907は、最もHWリソースが大きいVM種別で、VMAを配置可能な物理ホストが存在するか否かを判定する（ステップS1906）。ここで、存在しない場合（ステップS1906：No）、制御方針決定部907は、ステップS1907の処理に移行する。制御方針決定部907は、選択した組み合わせと同じVMBとVMCとのVM種別を含み、サービス性能が未設定である組み合わせについて、サービス性能をベンチマーク実施しないに設定する（ステップS1907）。管理装置100は、図20のステップS2001の処理に移行する。

【0117】

一方で、存在する場合（ステップS1906：Yes）、制御方針決定部907は、VMAが取りうる最もHWリソースが大きいVM種別で、VMAを配置可能な物理ホストにVMAを配置し、レイテンシを取得してサービス性能を設定する（ステップS1908）。

【0118】

制御方針決定部907は、VMAを配置した結果、サービス性能が10%以上改善したか否かを判定する（ステップS1909）。ここで、改善しない場合（ステップS1909：No）、制御方針決定部907は、ステップS1910の処理に移行する。制御方針決定部907は、選択した組み合わせと同じVMBとVMCとのVM種別を含み、サービス性能が未設定である組み合わせについて、サービス性能をベンチマーク実施しないに設定する（ステップS1910）。管理装置100は、図20のステップS2001の処理に移行する。

【0119】

一方で、改善している場合（ステップS1909：Yes）、制御方針決定部907は、ステップS1911の処理に移行する。制御方針決定部907は、選択した組み合わせと同じVMBとVMCとのVM種別を含み、サービス性能が未設定である組み合わせにつ

10

20

30

40

50

いて、レイテンシを取得してサービス性能を設定する（ステップS 1 9 1 1）。管理装置 1 0 0 は、図 2 0 のステップS 2 0 0 1 の処理に移行する。

【 0 1 2 0 】

図 2 0 において、データ取得部 9 0 5 は、V M B の V M 種別を可変にし、V M A と V M C との V M 種別を固定した、1 以上の組み合わせのうち、V M 稼働状況シート 7 0 0 上でサービス性能が未設定である組み合わせがあるか否かを判定する（ステップS 2 0 0 1）。ここで、未設定である組み合わせがない場合（ステップS 2 0 0 1 : N o）、管理装置 1 0 0 は、図 2 1 のステップS 2 1 0 1 の処理に移行する。

【 0 1 2 1 】

一方で、未設定である組み合わせがある場合（ステップS 2 0 0 1 : Y e s）、データ取得部 9 0 5 は、未設定である組み合わせのうち、V M 稼働状況シート 7 0 0 上で最も上にある組み合わせを選択する（ステップS 2 0 0 2）。

【 0 1 2 2 】

制御方針決定部 9 0 7 は、選択した組み合わせの V M A と V M B と V M C との V M 種別で、V M A と V M B と V M C とを配置可能な物理ホストが存在するか否かを判定する（ステップS 2 0 0 3）。ここで、存在しない場合（ステップS 2 0 0 3 : N o）、制御方針決定部 9 0 7 は、選択した組み合わせと同じ V M A と V M C との V M 種別を含む組み合わせについて、サービス性能をベンチマーク実施しないに設定する（ステップS 2 0 0 4）。管理装置 1 0 0 は、図 2 1 のステップS 2 1 0 1 の処理に移行する。

【 0 1 2 3 】

一方で、存在する場合（ステップS 2 0 0 3 : Y e s）、制御方針決定部 9 0 7 は、ステップS 2 0 0 5 の処理に移行する。制御方針決定部 9 0 7 は、選択した組み合わせの V M A と V M B と V M C との V M 種別で、V M A と V M B と V M C とを配置し、その組み合わせについて、レイテンシを取得してサービス性能を設定する（ステップS 2 0 0 5）。

【 0 1 2 4 】

制御方針決定部 9 0 7 は、最も H W リソースが大きい V M 種別で、V M B を配置可能な物理ホストが存在するか否かを判定する（ステップS 2 0 0 6）。ここで、存在しない場合（ステップS 2 0 0 6 : N o）、制御方針決定部 9 0 7 は、ステップS 2 0 0 7 の処理に移行する。制御方針決定部 9 0 7 は、選択した組み合わせと同じ V M A と V M C との V M 種別を含み、サービス性能が未設定である組み合わせについて、サービス性能をベンチマーク実施しないに設定する（ステップS 2 0 0 7）。管理装置 1 0 0 は、図 2 1 のステップS 2 1 0 1 の処理に移行する。

【 0 1 2 5 】

一方で、存在する場合（ステップS 2 0 0 6 : Y e s）、制御方針決定部 9 0 7 は、V M B が取りうる最も H W リソースが大きい V M 種別で、V M B を配置可能な物理ホストに V M B を配置し、レイテンシを取得してサービス性能を設定する（ステップS 2 0 0 8）。

【 0 1 2 6 】

制御方針決定部 9 0 7 は、V M B を配置した結果、サービス性能が 1 0 % 以上改善したか否かを判定する（ステップS 2 0 0 9）。ここで、改善しない場合（ステップS 2 0 0 9 : N o）、制御方針決定部 9 0 7 は、ステップS 2 0 1 0 の処理に移行する。制御方針決定部 9 0 7 は、選択した組み合わせと同じ V M A と V M C との V M 種別を含み、サービス性能が未設定である組み合わせについて、サービス性能をベンチマーク実施しないに設定する（ステップS 2 0 1 0）。管理装置 1 0 0 は、図 2 1 のステップS 2 1 0 1 の処理に移行する。

【 0 1 2 7 】

一方で、改善している場合（ステップS 2 0 0 9 : Y e s）、制御方針決定部 9 0 7 は、ステップS 2 0 1 1 の処理に移行する。制御方針決定部 9 0 7 は、選択した組み合わせと同じ V M A と V M C との V M 種別を含み、サービス性能が未設定である組み合わせについて、レイテンシを取得してサービス性能を設定する（ステップS 2 0 1 1）。管理装置

10

20

30

40

50

100は、図21のステップS2101の処理に移行する。

【0128】

図21において、データ取得部905は、VMCのVM種別を可変にし、VMAとVMBとのVM種別を固定した、1以上の組み合わせのうち、VM稼働状況シート700上でサービス性能が未設定である組み合わせがあるか否かを判定する(ステップS2101)。ここで、未設定である組み合わせがない場合(ステップS2101:No)、管理装置100は、図22のステップS2201の処理に移行する。

【0129】

一方で、未設定である組み合わせがある場合(ステップS2101:Yes)、データ取得部905は、未設定である組み合わせのうち、VM稼働状況シート700上で最も上にある組み合わせを選択する(ステップS2102)。

10

【0130】

制御方針決定部907は、選択した組み合わせのVMAとVMBとVMCとのVM種別で、VMAとVMBとVMCとを配置可能な物理ホストが存在するか否かを判定する(ステップS2103)。ここで、存在しない場合(ステップS2103:No)、制御方針決定部907は、選択した組み合わせと同じVMAとVMBとのVM種別を含む組み合わせについて、サービス性能をベンチマーク実施しないに設定する(ステップS2104)。管理装置100は、図22のステップS2201の処理に移行する。

【0131】

一方で、存在する場合(ステップS2103:Yes)、制御方針決定部907は、ステップS2105の処理に移行する。制御方針決定部907は、選択した組み合わせのVMAとVMBとVMCとのVM種別で、VMAとVMBとVMCとを配置し、その組み合わせについて、レイテンシを取得してサービス性能を設定する(ステップS2105)。

20

【0132】

制御方針決定部907は、最もHWリソースが大きいVM種別で、VMCを配置可能な物理ホストが存在するか否かを判定する(ステップS2106)。ここで、存在しない場合(ステップS2106:No)、制御方針決定部907は、ステップS2107の処理に移行する。制御方針決定部907は、選択した組み合わせと同じVMAとVMBとのVM種別を含み、サービス性能が未設定である組み合わせについて、サービス性能をベンチマーク実施しないに設定する(ステップS2107)。管理装置100は、図22のステップS2201の処理に移行する。

30

【0133】

一方で、存在する場合(ステップS2106:Yes)、制御方針決定部907は、VMCが取りうる最もHWリソースが大きいVM種別で、VMCを配置可能な物理ホストにVMCを配置し、レイテンシを取得してサービス性能を設定する(ステップS2108)。

【0134】

制御方針決定部907は、VMCを配置した結果、サービス性能が10%以上改善したか否かを判定する(ステップS2109)。ここで、改善しない場合(ステップS2109:No)、制御方針決定部907は、ステップS2110の処理に移行する。制御方針決定部907は、選択した組み合わせと同じVMAとVMBとのVM種別を含み、サービス性能が未設定である組み合わせについて、サービス性能をベンチマーク実施しないに設定する(ステップS2110)。管理装置100は、図22のステップS2201の処理に移行する。

40

【0135】

一方で、改善している場合(ステップS2109:Yes)、制御方針決定部907は、ステップS2111の処理に移行する。制御方針決定部907は、選択した組み合わせと同じVMAとVMBとのVM種別を含み、サービス性能が未設定である組み合わせについて、レイテンシを取得してサービス性能を設定する(ステップS2111)。管理装置100は、図22のステップS2201の処理に移行する。

50

【 0 1 3 6 】

図 2 2 において、データ取得部 9 0 5 は、合計 HW リソースを算出し、VM稼働状況シート 7 0 0 に設定する(ステップ S 2 2 0 1)。データ取得部 9 0 5 は、DB 9 0 1 を参照して、VM稼働状況シート 7 0 0 上でサービス性能が未設定である組み合わせがあるか否かを判定する(ステップ S 2 2 0 2)。ここで、未設定である組み合わせがある場合(ステップ S 2 2 0 2 : Y e s)、管理装置 1 0 0 は、図 1 9 のステップ S 1 9 0 1 の処理に移行する。

【 0 1 3 7 】

一方で、未設定である組み合わせがない場合(ステップ S 2 2 0 2 : N o)、移動対象選択部 9 0 6 は、DB 9 0 1 を参照して、VM稼働状況シート 7 0 0 上で、SLA を満たすサービス性能があるか否かを判定する(ステップ S 2 2 0 3)。ここで、SLA を満たすサービス性能がない場合(ステップ S 2 2 0 3 : N o)、管理装置 1 0 0 は、管理処理を終了する。

10

【 0 1 3 8 】

一方で、SLA を満たすサービス性能がある場合(ステップ S 2 2 0 3 : Y e s)、制御方針決定部 9 0 7 は、SLA を満たし、合計 HW リソースが最も小さい組み合わせを選択する(ステップ S 2 2 0 4)。制御方針決定部 9 0 7 は、選択した組み合わせになるように、VM を配置する(ステップ S 2 2 0 5)。管理装置 1 0 0 は、管理処理を終了する。これにより、管理装置 1 0 0 は、サービスに用いられる HW リソースを、SLA を満たす範囲で最小化することができる。

20

【 0 1 3 9 】

(管理装置 1 0 0 の動作例 2)

次に、図 2 3 ~ 図 2 7 を用いて、管理装置 1 0 0 の動作例 2 について説明する。

【 0 1 4 0 】

図 2 3 ~ 図 2 7 は、管理装置 1 0 0 の動作例 2 を示す説明図である。動作例 2 は、管理装置 1 0 0 が、性能情報として処理量 [o p s (o p e r a t i o n s)] またはスループット [o p s / s e c] を用いる場合についての動作例である。

【 0 1 4 1 】

図 2 3 において、管理装置 1 0 0 は、VM 種別の単位で、HW リソースの組み合わせを複数作成し、VM稼働状況シート 7 0 0 を作成する。管理装置 1 0 0 は、例えば、VM 種別テーブル 6 0 0 に基づいて、VMD に関する 5 つの VM 種別と、VME に関する 5 つの VM 種別とを組み合わせた、2 5 の組み合わせを作成する。VM 種別は、例えば、Low Spec、CPU Rich、MEM Rich、NET Rich、High Spec などである。管理装置 1 0 0 は、作成したそれぞれの組み合わせを、VM稼働状況シート 7 0 0 のそれぞれのレコードの VMD や VME のフィールドに設定することにより記憶する。

30

【 0 1 4 2 】

これにより、管理装置 1 0 0 は、あるサービスを提供する複数の VM に HW リソースを割り当てるバリエーションを、比較的多く考慮可能にすることができる。また、管理装置 1 0 0 は、VM 種別の単位で複数の VM に HW リソースを割り当てるバリエーションを考慮するため、考慮するバリエーションの数を低減し、HW リソースの最小化を検討する際にかかる負担の低減化を図ることができる。次に、図 2 4 の説明に移行する。

40

【 0 1 4 3 】

図 2 4 において、管理装置 1 0 0 は、VMD の VM 種別を可変にし、VME の VM 種別を固定した 1 以上の組み合わせのうち、最も HW リソースが小さい組み合わせを選択する。図 2 4 の例では、管理装置 1 0 0 は、VM稼働状況シート 7 0 0 のレコード 2 4 0 0 を選択する。管理装置 1 0 0 は、選択したレコード 2 4 0 0 に基づいて、VMD や VME の VM 種別が変更されるように、VMD や VME を複数の情報処理装置 1 0 1 に配置し、ベンチマークを実行する。

【 0 1 4 4 】

50

その後、管理装置100は、サービスに用いられるHWリソースの大きさの情報として、4CPU、16GB、1Gbit/s×2などを取得し、VM稼働状況シート700の合計HWリソースのフィールドに設定することにより記憶する。管理装置100は、サービスの処理量90opsを取得し、VM稼働状況シート700のサービス性能のフィールドに設定することにより記憶する。次に、図25の説明に移行する。

【0145】

図25において、管理装置100は、VMDのVM種別を可変にし、VMEのVM種別を固定した1以上の組み合わせのうち、最もHWリソースが大きい組み合わせを選択する。図25の例では、管理装置100は、VM稼働状況シート700のレコード2500を選択する。管理装置100は、選択したレコード2500に基づいて、VMDやVMEのVM種別が変更されるように、VMDやVMEを複数の情報処理装置101に配置し、ベンチマークを実行する。

10

【0146】

その後、管理装置100は、サービスに用いられるHWリソースの大きさの情報として、6CPU、24GB、1Gbit/s×1+10Gbit/s×1などを取得し、VM稼働状況シート700の合計HWリソースのフィールドに設定することにより記憶する。管理装置100は、サービスの処理量104opsを取得し、VM稼働状況シート700のサービス性能のフィールドに設定することにより記憶する。次に、図26の説明に移行する。

【0147】

20

図26において、管理装置100は、レコード2400に対応する最もHWリソースが小さい組み合わせから、レコード2500に対応する最もHWリソースが大きい組み合わせに変更した場合に、サービスの処理量が10%以上改善されるか否かを判定する。管理装置100は、10%以上改善されるため、レコード2400に対応する最もHWリソースが小さい組み合わせと、レコード2500に対応する最もHWリソースが大きい組み合わせとの間にある、他の組み合わせについての処理を省略しない。

【0148】

図26の例では、管理装置100は、VM稼働状況シート700のレコード2600~2602を順次選択する。まず、管理装置100は、選択したレコード2600に基づいて、VMDやVMEのVM種別が変更されるように、VMDやVMEを複数の情報処理装置101に配置し、ベンチマークを実行する。

30

【0149】

その後、管理装置100は、サービスに用いられるHWリソースの大きさの情報として、6CPU、16GB、1Gbit/s×2などを取得し、VM稼働状況シート700の合計HWリソースのフィールドに設定することにより記憶する。管理装置100は、サービスの処理量90opsを取得し、VM稼働状況シート700のサービス性能のフィールドに設定することにより記憶する。

【0150】

また、管理装置100は、選択したレコード2601に基づいて、VMDやVMEのVM種別が変更されるように、VMDやVMEを複数の情報処理装置101に配置し、ベンチマークを実行する。

40

【0151】

その後、管理装置100は、サービスに用いられるHWリソースの大きさの情報として、4CPU、24GB、1Gbit/s×2などを取得し、VM稼働状況シート700の合計HWリソースのフィールドに設定することにより記憶する。管理装置100は、サービスの処理量96opsを取得し、VM稼働状況シート700のサービス性能のフィールドに設定することにより記憶する。

【0152】

また、管理装置100は、選択したレコード2602に基づいて、VMDやVMEのVM種別が変更されるように、VMDやVMEを複数の情報処理装置101に配置し、ベン

50

チマークを実行する。

【0153】

その後、管理装置100は、サービスに用いられるHWリソースの大きさの情報として、4CPU、16GB、1Gbit/s×1+10Gbit/s×1などを取得し、VM稼働状況シート700の合計HWリソースのフィールドに設定することにより記憶する。管理装置100は、サービスの処理量94opsを取得し、VM稼働状況シート700のサービス性能のフィールドに設定することにより記憶する。その後、管理装置100は、同様に、VM稼働状況シート700の、合計HWリソースのフィールド、および、サービス性能のフィールドを設定する。次に、図27の説明に移行する。

【0154】

図27において、管理装置100は、VM稼働状況シート700において、サービス性能が目標サービス性能を満たすVM種別の組み合わせのうち、合計HWリソースが最小になる組み合わせを選択する。目標サービス性能は、SLAに対応する性能値である。そして、管理装置100は、選択したVM種別の組み合わせで、VMDやVMEを複数の情報処理装置101に配置する。

【0155】

図27の例では、管理装置100は、具体的には、サービス性能の処理量が、目標サービス性能の処理量よりも大きくなり、SLAを満たすVM種別の組み合わせのうち、レコード2700に対応するVM種別の組み合わせを選択する。そして、管理装置100は、そのVM種別の組み合わせで、VMDやVMEを複数の情報処理装置101に配置する。

【0156】

これにより、管理装置100は、サービスに用いられるHWリソースを、SLAを満たす範囲で最小化することができる。このため、管理装置100は、サービスにかかる消費電力の低減化を図ることができる。また、管理装置100は、他のサービスに利用可能なHWリソースを確保し、情報処理システム110上でより多くのサービスを提供可能にすることができる。また、管理装置100は、あるサービスの負荷が急上昇した際に、そのサービスに割り当て可能な、空きHWリソースを確保しておき、情報処理システム110の安定性の向上を図ることができる。

【0157】

ここでは、管理装置100が、VM種別の単位で、HWリソースの組み合わせを複数作成する場合について説明したが、これに限らない。例えば、管理装置100が、計算資源管理テーブル500に基づいて、それぞれのVMに割り当て可能なHWリソースの大きさを特定し、それぞれのVMに割り当て可能なHWリソースの組み合わせを複数作成する場合があってもよい。

【0158】

例えば、管理装置100は、それぞれのVMに割り当て可能なCPUの最小数と最大数とを特定し、最小数と最大数の間を所定単位で区切り、それぞれのVMに割り当てるCPUの数になる複数の候補を特定する。また、管理装置100は、同様に、それぞれのVMに割り当てる記憶容量になる複数の候補、ネットワーク帯域になる複数の候補を特定する。そして、管理装置100は、それぞれの候補を組み合わせ、それぞれのVMに割り当てるHWリソースの組み合わせを複数作成し、VM稼働状況シート700に設定する。

【0159】

(実施例2における管理処理手順の一例)

次に、図28～図31を用いて、実施例2における管理処理手順の一例について説明する。

【0160】

図28～図31は、実施例2における管理処理手順の一例を示すフローチャートである。図28において、データ取得部905は、タイマー部904が発生する要求入力を検出したことに応じて、ユーザ定義されたサービスを運用するVMグループにおけるHWリソースの上限値を、OSから取得し、DB901に保存する(ステップS2801)。

10

20

30

40

50

【 0 1 6 1 】

データ取得部 9 0 5 は、要求入力に応じて、ユーザ定義されたサービスを運用する VM グループに対して、サービス性能となる処理量と各 VM の HW リソースの使用状況を取得して、DB 9 0 1 に保存する（ステップ S 2 8 0 2 ）。

【 0 1 6 2 】

データ取得部 9 0 5 は、移動対象となる物理ホストの使用状況を取得して、DB 9 0 1 に保存する（ステップ S 2 8 0 3 ）。データ取得部 9 0 5 は、VM 種別の組み合わせを特定し、VM 稼働状況シート 7 0 0 を作成し、DB 9 0 1 に保存する（ステップ S 2 8 0 4 ）。管理装置 1 0 0 は、図 2 9 のステップ S 2 9 0 1 の処理に移行する。

【 0 1 6 3 】

図 2 9 において、データ取得部 9 0 5 は、VMD の VM 種別を可変にし、VME の VM 種別を固定した、1 以上の組み合わせのうち、VM 稼働状況シート 7 0 0 上でサービス性能が未設定である組み合わせがあるか否かを判定する（ステップ S 2 9 0 1 ）。ここで、未設定である組み合わせがない場合（ステップ S 2 9 0 1 : No ）、管理装置 1 0 0 は、図 3 0 のステップ S 3 0 0 1 の処理に移行する。

【 0 1 6 4 】

一方で、未設定である組み合わせがある場合（ステップ S 2 9 0 1 : Yes ）、データ取得部 9 0 5 は、未設定である組み合わせのうち、VM 稼働状況シート 7 0 0 上で最も上にある組み合わせを選択する（ステップ S 2 9 0 2 ）。

【 0 1 6 5 】

制御方針決定部 9 0 7 は、選択した組み合わせの VMD と VME との VM 種別で、VMD と VME とを配置可能な物理ホストが存在するか否かを判定する（ステップ S 2 9 0 3 ）。ここで、存在しない場合（ステップ S 2 9 0 3 : No ）、制御方針決定部 9 0 7 は、選択した組み合わせと同じ VME の VM 種別を含む組み合わせについて、サービス性能をベンチマーク実施しないに設定する（ステップ S 2 9 0 4 ）。管理装置 1 0 0 は、図 3 0 のステップ S 3 0 0 1 の処理に移行する。

【 0 1 6 6 】

一方で、存在する場合（ステップ S 2 9 0 3 : Yes ）、制御方針決定部 9 0 7 は、選択した組み合わせの VMD と VME との VM 種別で、VMD と VME とを配置し、その組み合わせについて、処理量を取得してサービス性能を設定する（ステップ S 2 9 0 5 ）。

【 0 1 6 7 】

制御方針決定部 9 0 7 は、最も HW リソースが大きい VM 種別で、VMD を配置可能な物理ホストが存在するか否かを判定する（ステップ S 2 9 0 6 ）。ここで、存在しない場合（ステップ S 2 9 0 6 : No ）、制御方針決定部 9 0 7 は、選択した組み合わせと同じ VME の VM 種別を含み、サービス性能が未設定である組み合わせについて、サービス性能をベンチマーク実施しないに設定する（ステップ S 2 9 0 7 ）。管理装置 1 0 0 は、図 3 0 のステップ S 3 0 0 1 の処理に移行する。

【 0 1 6 8 】

一方で、存在する場合（ステップ S 2 9 0 6 : Yes ）、制御方針決定部 9 0 7 は、VMD が取りうる最も HW リソースが大きい VM 種別で、VMD を配置可能な物理ホストに VMD を配置し、処理量を取得してサービス性能を設定する（ステップ S 2 9 0 8 ）。

【 0 1 6 9 】

制御方針決定部 9 0 7 は、VMD を配置した結果、サービス性能が 1 0 % 以上改善したか否かを判定する（ステップ S 2 9 0 9 ）。ここで、改善しない場合（ステップ S 2 9 0 9 : No ）、制御方針決定部 9 0 7 は、選択した組み合わせと同じ VME の VM 種別を含み、サービス性能が未設定である組み合わせについて、サービス性能をベンチマーク実施しないに設定する（ステップ S 2 9 1 0 ）。管理装置 1 0 0 は、図 3 0 のステップ S 3 0 0 1 の処理に移行する。

【 0 1 7 0 】

一方で、改善している場合（ステップ S 2 9 0 9 : Yes ）、制御方針決定部 9 0 7 は

10

20

30

40

50

、選択した組み合わせと同じVMEのVM種別を含み、サービス性能が未設定である組み合わせについて、処理量を取得してサービス性能を設定する(ステップS2911)。管理装置100は、図30のステップS3001の処理に移行する。

【0171】

図30において、データ取得部905は、VMEのVM種別を可変にし、VMDのVM種別を固定した、1以上の組み合わせのうち、VM稼働状況シート700上でサービス性能が未設定である組み合わせがあるか否かを判定する(ステップS3001)。ここで、未設定である組み合わせがない場合(ステップS3001:No)、管理装置100は、図31のステップS3101の処理に移行する。

【0172】

一方で、未設定である組み合わせがある場合(ステップS3001:Yes)、データ取得部905は、未設定である組み合わせのうち、VM稼働状況シート700上で最も上にある組み合わせを選択する(ステップS3002)。

【0173】

制御方針決定部907は、選択した組み合わせのVMDとVMEとのVM種別で、VMDとVMEとを配置可能な物理ホストが存在するか否かを判定する(ステップS3003)。ここで、存在しない場合(ステップS3003:No)、制御方針決定部907は、選択した組み合わせと同じVMDのVM種別を含む組み合わせについて、サービス性能をベンチマーク実施しないに設定する(ステップS3004)。管理装置100は、図31

【0174】

一方で、存在する場合(ステップS3003:Yes)、制御方針決定部907は、選択した組み合わせのVMDとVMEとのVM種別で、VMDとVMEとを配置し、その組み合わせについて、処理量を取得してサービス性能を設定する(ステップS3005)。

【0175】

制御方針決定部907は、最もHWリソースが大きいVM種別で、VMEを配置可能な物理ホストが存在するか否かを判定する(ステップS3006)。ここで、存在しない場合(ステップS3006:No)、制御方針決定部907は、選択した組み合わせと同じVMDのVM種別を含み、サービス性能が未設定である組み合わせについて、サービス性能をベンチマーク実施しないに設定する(ステップS3007)。管理装置100は、図

【0176】

一方で、存在する場合(ステップS3006:Yes)、制御方針決定部907は、VMEが取りうる最もHWリソースが大きいVM種別で、VMEを配置可能な物理ホストにVMEを配置し、処理量を取得してサービス性能を設定する(ステップS3008)。

【0177】

制御方針決定部907は、VMEを配置した結果、サービス性能が10%以上改善したか否かを判定する(ステップS3009)。ここで、改善しない場合(ステップS3009:No)、制御方針決定部907は、選択した組み合わせと同じVMDのVM種別を含み、サービス性能が未設定である組み合わせについて、サービス性能をベンチマーク実施

【0178】

一方で、改善している場合(ステップS3009:Yes)、制御方針決定部907は、選択した組み合わせと同じVMDのVM種別を含み、サービス性能が未設定である組み合わせについて、処理量を取得してサービス性能を設定する(ステップS3011)。管理装置100は、図31のステップS3101の処理に移行する。

【0179】

図31において、データ取得部905は、合計HWリソースを算出し、VM稼働状況シート700に設定する(ステップS3101)。データ取得部905は、DB901を参

10

20

30

40

50

照して、VM稼働状況シート700上でサービス性能が未設定である組み合わせがあるかを判定する(ステップS3102)。ここで、未設定である組み合わせがある場合(ステップS3102:Yes)、管理装置100は、図29のステップS2901の処理に移行する。

【0180】

一方で、未設定である組み合わせがない場合(ステップS3102:No)、移動対象選択部906は、DB901を参照して、VM稼働状況シート700上で、SLAを満たすサービス性能があるかを判定する(ステップS3103)。ここで、SLAを満たすサービス性能がない場合(ステップS3103:No)、管理装置100は、管理処理を終了する。

10

【0181】

一方で、SLAを満たすサービス性能がある場合(ステップS3103:Yes)、制御方針決定部907は、SLAを満たし、合計HWリソースが最も小さい組み合わせを選択する(ステップS3104)。制御方針決定部907は、選択した組み合わせになるように、VMを配置する(ステップS3105)。管理装置100は、管理処理を終了する。これにより、管理装置100は、サービスに用いられるHWリソースを、SLAを満たす範囲で最小化することができる。

【0182】

以上説明したように、管理装置100によれば、複数の情報処理装置101がそれぞれ属し、所定のサービスをそれぞれ提供する複数のグループ毎に、各情報処理装置101の構成情報と性能情報とを取得し、記憶部に記憶することができる。また、管理装置100によれば、記憶部が複数のグループ毎に記憶する構成情報と性能情報とに基づき、選択対象の仮想マシンが属する対象グループから他のグループへの移動対象となる移動対象仮想マシンを複数のグループ毎に選択することができる。また、管理装置100によれば、移動対象仮想マシンを他のグループに移動させた場合、移動対象仮想マシンに対応するサービスの性能値が予め定められた基準値を満たすとき、対象グループから他のグループへの移動対象仮想マシンの移動を決定することができる。これにより、管理装置100は、SLAを満たす範囲で、対象グループより、サービスのために用いられるHWリソースが小さい可能性があるグループを特定することができる。そして、管理装置100は、VMA~VMCを、SLAを満たす他のグループに移動して、HWリソースの最小化を図ることができる。

20

30

【0183】

管理装置100によれば、各情報処理装置101の構成情報に応じた仮想マシン種別を複数のグループ毎に記憶部に記憶することができる。また、管理装置100によれば、記憶部が複数のグループ毎に記憶する仮想マシン種別と性能情報とに基づき、移動対象仮想マシンを複数のグループ毎に選択することができる。これにより、管理装置100は、VM種別の単位で、複数のVMに割り当てるHWリソースの組み合わせのバリエーションを考慮することができる。このため、管理装置100は、考慮するバリエーションの数を低減することができる。HWリソースの最小化を検討する際にかかる負担の低減化を図ることができる。

40

【0184】

また、管理装置100によれば、さらに、移動対象仮想マシンを他のグループに移動させた場合、移動対象仮想マシンに対応するサービスの性能値が、移動前の移動対象仮想マシンに対応するサービスの性能値よりも良好であるかを判断することができる。また、管理装置100によれば、基準値を満たし、かつ、良好であるとき、対象グループから他のグループへの移動対象仮想マシンの移動を決定することができる。これにより、管理装置100は、現在よりもサービスの性能値が良好になる場合にVMを移動するため、効率よくサービスの性能値の向上を図ることができる。

【0185】

また、管理装置100によれば、対象グループの計算資源より他のグループの計算資源

50

が少ないとき、対象グループから他のグループへの移動対象仮想マシンの移動を決定することができる。これにより、管理装置100は、現在よりもサービスに用いられるHWリソースが小さくなる場合にVMを移動するため、効率よくサービスに用いられる計算資源の最小化を図ることができる。

【0186】

また、管理装置100によれば、取得済みの構成情報と性能情報とに基づき、性能情報を未取得であるグループについて性能情報を取得するか否かを判定し、取得すると判定したグループについて性能情報を取得して記憶部に記憶することができる。これにより、管理装置100は、情報処理装置101にVMを配置してベンチマークを実行して性能情報を取得する際にかかる時間を削減し、管理装置100にかかる負担を低減することができる。

10

【0187】

なお、本実施の形態で説明した情報処理システムの制御方法は、予め用意されたプログラムをパーソナル・コンピュータやワークステーション等のコンピュータで実行することにより実現することができる。本実施の形態で説明した管理装置の管理プログラムは、ハードディスク、フレキシブルディスク、CD-ROM、MO、DVD等のコンピュータで読み取り可能な記録媒体に記録され、コンピュータによって記録媒体から読み出されることによって実行される。また、本実施の形態で説明した管理装置の管理プログラムは、インターネット等のネットワークを介して配布してもよい。

【0188】

上述した実施の形態に関し、さらに以下の付記を開示する。

20

【0189】

(付記1) 複数の計算資源をそれぞれ有する複数の情報処理装置を備える情報処理システムを管理する管理装置であって、

前記複数の情報処理装置がそれぞれ属し、所定のサービスをそれぞれ提供する複数のグループの各々がそれぞれ実行するホストOSから取得した、各情報処理装置の構成情報と性能情報とを、前記複数のグループ毎に記憶部に記憶する取得部と、

前記記憶部が前記複数のグループ毎に記憶する構成情報と性能情報とに基づき、選択対象の仮想マシンが属する対象グループから他のグループへの移動対象となる移動対象仮想マシンを前記複数のグループ毎に選択する選択部と、

30

前記移動対象仮想マシンを前記他のグループに移動させた場合、前記移動対象仮想マシンに対応するサービスの性能値が予め定められた基準値を満たすか否かを判断し、前記基準値を満たすとき、前記対象グループから前記他のグループへの前記移動対象仮想マシンの移動を決定する決定部と、

を有することを特徴とする管理装置。

【0190】

(付記2) 前記取得部は、各情報処理装置の構成情報に応じた仮想マシン種別を前記複数のグループ毎に記憶部に記憶し、

前記選択部は、前記記憶部が前記複数のグループ毎に記憶する仮想マシン種別と性能情報とに基づき、前記移動対象仮想マシンを複数のグループ毎に選択する、

40

ことを特徴とする付記1に記載の管理装置。

【0191】

(付記3) 前記決定部は、前記移動対象仮想マシンを前記他のグループに移動させた場合、前記移動対象仮想マシンに対応するサービスの性能値が、前記基準値を満たすか否か、および、移動前の前記移動対象仮想マシンに対応するサービスの性能値よりも良好であることを示すか否かを判断し、前記基準値を満たし、かつ、良好であることを示すとき、前記対象グループから前記他のグループへの前記移動対象仮想マシンの移動を決定する、ことを特徴とする付記1または2に記載の管理装置。

【0192】

(付記4) 前記決定部は、前記対象グループの計算資源より前記他のグループの計算資源

50

が少ないとき、前記対象グループから前記他のグループへの前記移動対象仮想マシンの移動を決定する、ことを特徴とする付記 1 ~ 3 のいずれか一つに記載の管理装置。

【 0 1 9 3 】

(付記 5) 前記取得部は、取得済みの構成情報と性能情報とに基づき、性能情報を未取得であるグループについて性能情報を取得するか否かを判定し、取得すると判定したグループについて性能情報を取得して前記記憶部に記憶する、ことを特徴とする付記 1 ~ 4 のいずれか一つに記載の管理装置。

【 0 1 9 4 】

(付記 6) 複数の計算資源をそれぞれ有する複数の情報処理装置を備える情報処理システムの制御方法であって、

前記情報処理システムを管理する管理装置が、

前記複数の情報処理装置がそれぞれ属し、所定のサービスをそれぞれ提供する複数のグループの各々がそれぞれ実行するホスト OS から取得した、各情報処理装置の構成情報と性能情報とを、前記複数のグループ毎に記憶部に記憶し、

前記記憶部が前記複数のグループ毎に記憶する構成情報と性能情報とに基づき、選択対象の仮想マシンが属する対象グループから他のグループへの移動対象となる移動対象仮想マシンを前記複数のグループ毎に選択し、

前記移動対象仮想マシンを前記他のグループに移動させた場合、前記移動対象仮想マシンに対応するサービスの性能値が予め定められた基準値を満たすか否かを判断し、前記基準値を満たすとき、前記対象グループから前記他のグループへの前記移動対象仮想マシンの移動を決定する、

処理を実行することを特徴とする制御方法。

【 0 1 9 5 】

(付記 7) 複数の計算資源をそれぞれ有する複数の情報処理装置を備える情報処理システムを管理する管理装置の管理プログラムであって、

前記複数の情報処理装置がそれぞれ属し、所定のサービスをそれぞれ提供する複数のグループの各々がそれぞれ実行するホスト OS から取得した、各情報処理装置の構成情報と性能情報とを、前記複数のグループ毎に記憶部に記憶し、

前記記憶部が前記複数のグループ毎に記憶する構成情報と性能情報とに基づき、選択対象の仮想マシンが属する対象グループから他のグループへの移動対象となる移動対象仮想マシンを前記複数のグループ毎に選択し、

前記移動対象仮想マシンを前記他のグループに移動させた場合、前記移動対象仮想マシンに対応するサービスの性能値が予め定められた基準値を満たすか否かを判断し、前記基準値を満たすとき、前記対象グループから前記他のグループへの前記移動対象仮想マシンの移動を決定する、

処理を前記管理装置に実行させることを特徴とする管理プログラム。

【 符号の説明 】

【 0 1 9 6 】

- 1 0 0 管理装置
- 1 0 1 情報処理装置
- 1 1 0 情報処理システム
- 2 1 0 ネットワーク
- 3 0 1 インターネット
- 3 0 2 ファイアウォール
- 3 0 3 クライアント装置
- 4 0 0 バス
- 4 0 1 CPU
- 4 0 2 メモリ
- 4 0 3 ネットワーク I / F
- 4 0 4 記録媒体 I / F

10

20

30

40

50

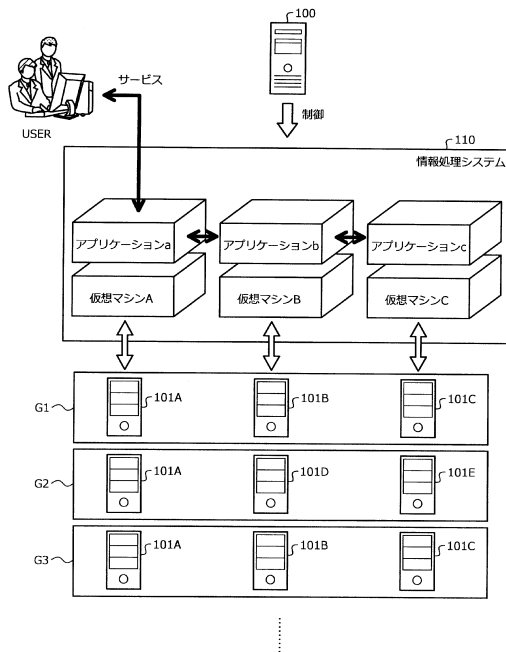
- 4 0 5 記録媒体
- 5 0 0 計算資源管理テーブル
- 6 0 0 V M 種別テーブル
- 7 0 0 V M 稼働状況シート
- 8 0 1 記憶部
- 8 0 2 取得部
- 8 0 3 選択部
- 8 0 4 決定部
- 8 0 5 出力部
- 9 0 0 V M 管理ソフトウェア
- 9 0 1 D B
- 9 0 2 グループ定義
- 9 0 3 サービス性能情報
- 9 0 4 タイマー部
- 9 0 5 データ取得部
- 9 0 6 移動対象選択部
- 9 0 7 制御方針決定部
- 9 1 0 インフラ管理ソフトウェア
- 9 1 1 制御受け取り部
- 9 1 2 制御指示部
- 9 2 1 , 9 2 2 仮想化ソフトウェア

10

20

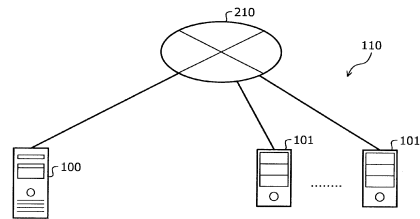
【図 1】

実施の形態にかかる情報処理システム 110 の制御方法の一実施例を示す説明図



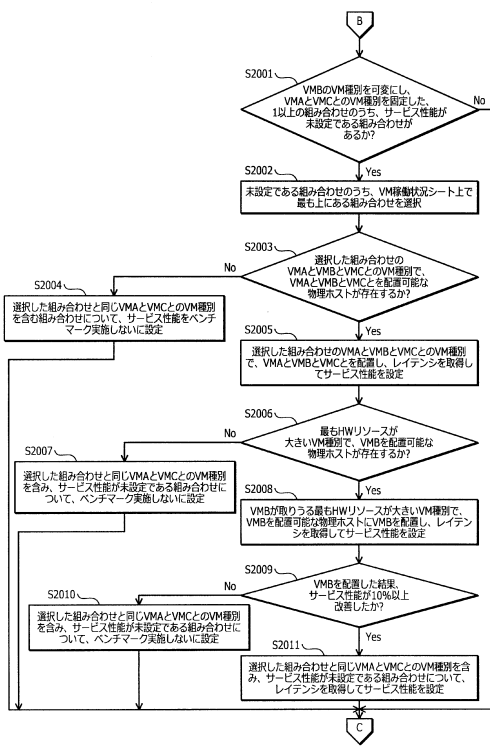
【図 2】

情報処理システム 110 の一例を示す説明図



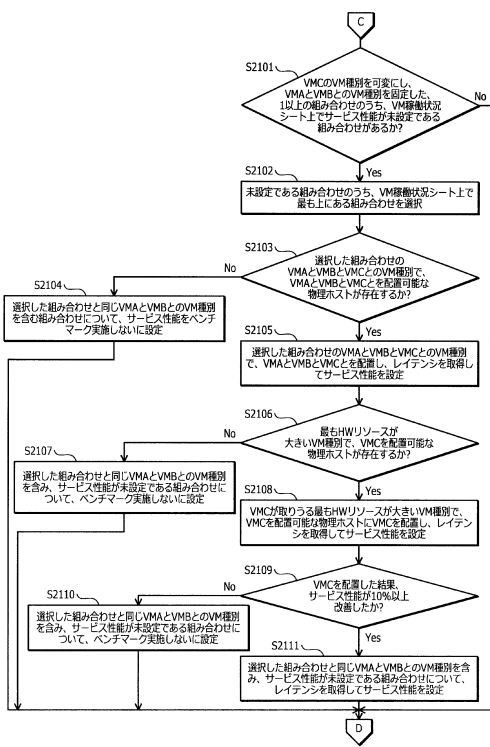
【図 20】

実施例 1 における管理処理手順の一例を示すフローチャート (その 3)



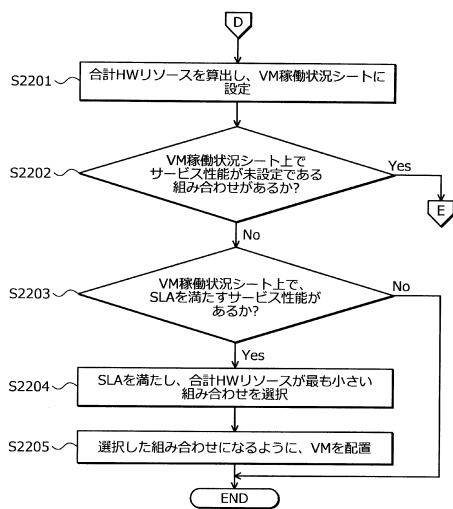
【図 21】

実施例 1 における管理処理手順の一例を示すフローチャート (その 4)



【図 22】

実施例 1 における管理処理手順の一例を示すフローチャート (その 5)



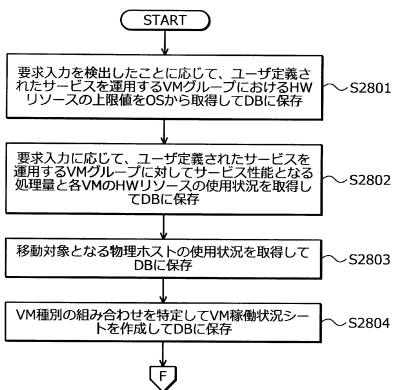
【図 23】

管理装置 100 の動作例 2 を示す説明図 (その 1)

VM種別	仮想マシンD			仮想マシンE			合計ハードウェアリソース			サービス性能	目標サービス性能
	CPU	MEM	NET	CPU	MEM	NET	CPU	MEM	NET		
Low Spec	2CPU	8GB	1Gbit/s	2CPU	8GB	1Gbit/s	4CPU	16GB	2Gbit/s	...	1000ops
CPU Rich	4CPU	8GB	1Gbit/s	"	"	"	4CPU	8GB	1Gbit/s	...	
MEM Rich	2CPU	16GB	1Gbit/s	"	16GB	"	2CPU	16GB	1Gbit/s	...	
NET Rich	2CPU	8GB	10Gbit/s	"	"	10Gbit/s	2CPU	8GB	10Gbit/s	...	
High Spec	4CPU	16GB	10Gbit/s	"	"	"	4CPU	16GB	10Gbit/s	...	
Low Spec	2CPU	8GB	1Gbit/s	2CPU	8GB	1Gbit/s	4CPU	16GB	2Gbit/s	...	
"	"	"	"	4CPU	8GB	1Gbit/s	4CPU	8GB	1Gbit/s	...	
"	"	"	"	2CPU	16GB	1Gbit/s	2CPU	16GB	1Gbit/s	...	
"	"	"	"	2CPU	8GB	10Gbit/s	2CPU	8GB	10Gbit/s	...	
"	"	"	"	4CPU	16GB	10Gbit/s	4CPU	16GB	10Gbit/s	...	

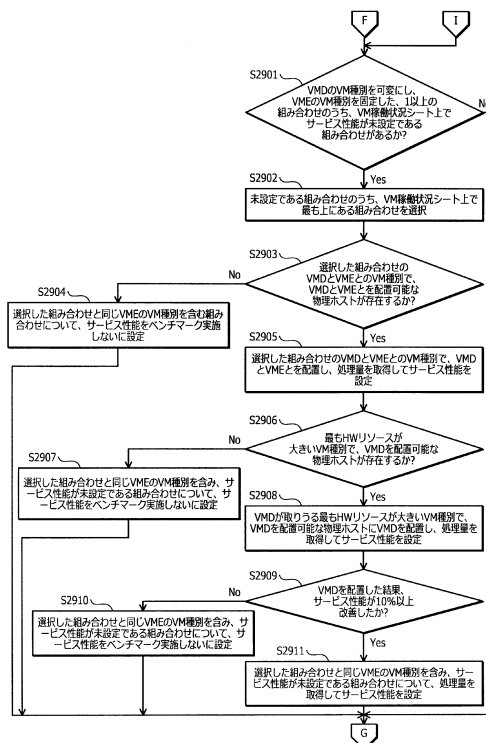
【図 28】

実施例 2 における管理処理手順の一例を示すフローチャート (その 1)



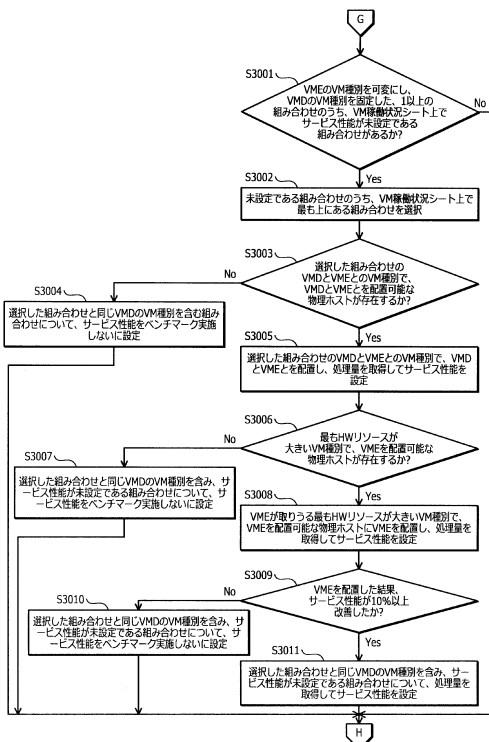
【図 29】

実施例 2 における管理処理手順の一例を示すフローチャート (その 2)



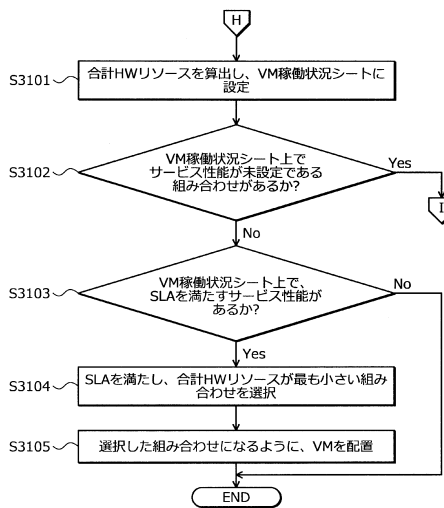
【図 30】

実施例 2 における管理処理手順の一例を示すフローチャート (その 3)



【図 31】

実施例 2 における管理処理手順の一例を示すフローチャート (その 4)



フロントページの続き

(72)発明者 福田 裕幸
神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

審査官 桜井 茂行

(56)参考文献 国際公開第2010/024027(WO, A1)
米国特許出願公開第2011/0161483(US, A1)
国際公開第2012/066640(WO, A1)
米国特許出願公開第2013/0238804(US, A1)
特開2010-113677(JP, A)
特開2016-110248(JP, A)

(58)調査した分野(Int.Cl., DB名)
G06F 9/455 - 9/54