

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2016-510929

(P2016-510929A)

(43) 公表日 平成28年4月11日(2016.4.11)

(51) Int.Cl.

G06F 9/50 (2006.01)

F I

G06F 9/46 4 6 5 Z

テーマコード (参考)

審査請求 未請求 予備審査請求 有 (全 19 頁)

(21) 出願番号 特願2016-500677 (P2016-500677)
 (86) (22) 出願日 平成26年3月5日 (2014.3.5)
 (85) 翻訳文提出日 平成27年10月9日 (2015.10.9)
 (86) 国際出願番号 PCT/US2014/020834
 (87) 国際公開番号 W02014/158909
 (87) 国際公開日 平成26年10月2日 (2014.10.2)
 (31) 優先権主張番号 13/801,964
 (32) 優先日 平成25年3月13日 (2013.3.13)
 (33) 優先権主張国 米国 (US)

(71) 出願人 314015767
 マイクロソフト テクノロジー ライセン
 シング, エルエルシー
 アメリカ合衆国 ワシントン州 9805
 2 レッドモンド ワン マイクロソフト
 ウェイ
 (74) 代理人 100140109
 弁理士 小野 新次郎
 (74) 代理人 100075270
 弁理士 小林 泰
 (74) 代理人 100101373
 弁理士 竹内 茂雄
 (74) 代理人 100118902
 弁理士 山本 修

最終頁に続く

(54) 【発明の名称】 分散データセンター技術

(57) 【要約】

世界の特定地域に対応する地域データセンターシステム。データセンターシステムは、サービスへの地域固有の要求がデータセンターのいずれかによって満たされてもよいようにそれぞれ構成される、複数のデータセンターを含む。データセンターの1つまたは複数のデータセンターは、受信される地域固有の要求について負荷分散を行ってもよい。負荷分散を行うために、受信データセンターは、要求をそれ自体が処理してもよいが、しかし場合によっては、要求が地域データセンターシステム内の別のデータセンターによって処理されるべきであると判定してもよい。

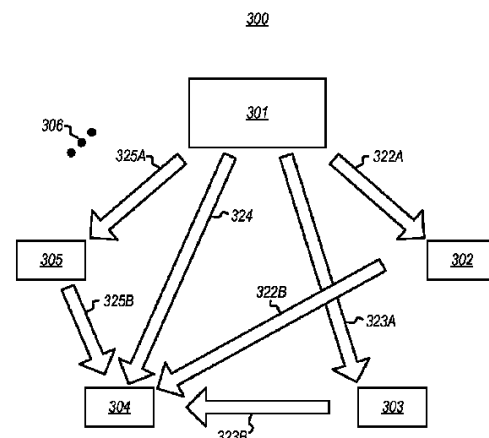


Figure 3B

【特許請求の範囲】**【請求項 1】**

複数のデータセンターを備えるシステムであって、

前記複数のデータセンターは、特定地域に対応するサービスへの地域固有の要求が、前記複数のデータセンターのいずれかによって満たすことができるように構成され、前記複数のデータセンターは、少なくとも第 1、第 2 および第 3 のデータセンターを含み、

前記第 1 のデータセンターは、前記特定地域に対応し、前記第 1 のデータセンターによって受信される複数の地域固有の要求について、少なくとも、

前記第 1 のデータセンターによって受信される前記地域固有の要求の少なくともいくつかについて、前記地域固有の要求が前記第 2 のデータセンターによって処理されるべきであると判定する前記第 1 のデータセンターのアクトを行うことによって、負荷分散を行うように構成される、システム。

10

【請求項 2】

請求項 1 に記載のシステムであって、前記特定地域に対応し、前記第 1 のデータセンターによって受信される前記複数の地域固有の要求の少なくともいくつかについて、前記第 1 のデータセンターはさらに、

前記第 2 のデータセンターへの反映のために前記地域固有の要求を前記第 3 のデータセンターに転送するアクトを行うように構成される、システム。

20

【請求項 3】

請求項 2 に記載のシステムであって、前記第 2 のデータセンターによって処理されるべきと前記第 1 のデータセンターが判定する前記地域固有の要求の少なくともいくつかについて、前記負荷分散はさらに、最初に前記第 3 のデータセンターから反映されることなく前記地域固有の要求を前記第 2 のデータセンターに転送するアクトを行うことによって行われる、システム。

【請求項 4】

請求項 1 に記載のシステムであって、前記第 1 のデータセンターによって受信される前記地域固有の要求の少なくともいくつかは、前記第 2 または第 3 のデータセンターを使用することなく前記第 1 のデータセンターによって処理される、システム。

【請求項 5】

請求項 1 に記載のシステムであって、前記第 1 のデータセンターはさらに、少なくとも、

前記第 1 のデータセンターによって受信される地域固有の要求の少なくともいくつかについて、前記地域固有の要求が前記第 3 のデータセンターによって処理されるべきであると判定する前記第 1 のデータセンターのアクト、および

前記第 3 のデータセンターによって処理されるべきと前記第 1 のデータセンターが判定する前記地域固有の要求の少なくともいくつかについて、前記第 3 のデータセンターへの反映のために前記地域固有の要求を前記第 2 のデータセンターに転送するアクトを行うことによって前記負荷分散を行うように構成される、システム。

40

【請求項 6】

請求項 1 に記載のシステムであって、前記第 1、第 2 および第 3 のデータセンターのそれぞれは、レイテンシエンベロープ内で動作することを検証されており、前記レイテンシエンベロープにおいて、前記レイテンシエンベロープ内の前記データセンターのいずれかの間で通信する際のレイテンシは、特定のしきい値を下回ることが検証されている、システム。

【請求項 7】

請求項 1 に記載のシステムであって、前記負荷分散は、プロトコルスタックでのネットワーク層で行われる、システム。

【請求項 8】

特定地域に対応するサービスへの複数の地域固有の要求を第 1 のデータセンターで受信

50

するアクトであって、前記複数の地域固有のサービス要求は、前記第 1 のデータセンター、第 2 のデータセンターおよび第 3 のデータセンターを含む複数のデータセンターのいずれかによって満たすことができる、アクトと、

受信される前記複数の地域固有のサービス要求の少なくともいくつかについて、

負荷分散される前記複数の地域固有のサービス要求の少なくともいくつかについて、前記第 2 のデータセンターが前記要求を処理すべきであると判定するアクトを行うことによって、受信される前記複数の地域固有のサービス要求の前記少なくともいくつかについて負荷分散を行う前記第 1 のデータセンターのアクトとを含む方法。

【請求項 9】

請求項 8 に記載の方法であって、負荷分散を行う前記第 1 のデータセンターの前記アクトはさらに、

前記第 2 のデータセンターによって処理されるべきと前記第 1 のデータセンターが判定する前記複数の地域固有のサービス要求の少なくともいくつかについて、前記第 2 のデータセンターへの反映のために前記地域固有の要求を前記第 3 のデータセンターに転送するアクトを含む、方法。

【請求項 10】

請求項 9 に記載の方法であって、前記の第 1、第 2 および第 3 のデータセンターのそれぞれは、レイテンシエンベロープ内で動作することを検証されており、前記レイテンシエンベロープにおいて、前記レイテンシエンベロープ内の前記データセンターのいずれかの間で通信する際のレイテンシは、特定のしきい値を下回ることが検証されている、方法。

【発明の詳細な説明】

【背景技術】

【0001】

[0001]「クラウドコンピューティング」は、遍在する、便利なオンデマンドネットワークが、構成可能なコンピューティングリソース（例えば、ネットワーク、サーバー、記憶装置、アプリケーション、およびサービス）の共有プールにアクセスすることを可能にするためのモデルである。構成可能なコンピューティングリソースの共有プールは、仮想化を介して迅速に準備され、低管理努力またはサービスプロバイダー相互作用でリリースされ、次いでそれに応じて拡大縮小されることもある。クラウドコンピューティングモデルは、様々な特徴（例えば、オンデマンドセルフサービス、広域ネットワークアクセス、リソースプール化、高速融通性、計数サービス、その他）、サービスモデル（例えば、サービス型ソフトウェア（「SaaS」）、サービス型プラットフォーム（「PaaS」）、サービス型インフラストラクチャー（「IaaS」））、および展開モデル（例えば、プライベートクラウド、コミュニティクラウド、パブリッククラウド、ハイブリッドクラウド、その他）から構成されることもある。クラウドコンピューティングモデルを実施する環境は、しばしばクラウドコンピューティング環境と呼ばれる。

【0002】

[0002]クラウドコンピューティング環境は、処理能力、メモリ、記憶装置、帯域幅、その他などの、膨大な量のコンピューティングリソースをそれぞれ有する、いくつかのデータセンターを含むこともある。データセンターは例えば、何百もの個人を雇い、おそらく何百もしくは何千ものサーバーまたはサーバーのラックを含む高機能の事業拠点のこともある。各データセンターは、世界の特定地域内に、地域またはレイテンシに敏感でかつその特定地域で提示されるクラウドサービスを支援する傾向があるクラウドサービスのために置かれる。クラウドコンピューティングに対する需要が、その地域で増加すると、データセンターは、かなり大きいこともあるが、データセンターの利用もまた、データセンターの大部分が利用される点まで増加する可能性がある。

【発明の概要】

【0003】

[0003]本明細書で述べられる少なくとも 1 つの実施形態は、世界の特定地域に対応する

10

20

30

40

50

地域データセンターシステムに関する。データセンターシステムは、サービスへの地域固有の要求がデータセンターのいずれかによって満たされてもよいようにそれぞれ構成される、複数のデータセンターを含む。データセンターの1つまたは複数の他のデータセンターによって処理されるべき地域固有の要求について負荷分散を行ってもよい。負荷分散を行うために、受信データセンターは、要求をそれ自体が処理してもよいが、しかし場合によっては、要求が、地域データセンターシステム内の1つまたは複数の他のデータセンターによって処理されるべきであると判定してもよい。そのような地域データセンターシステムの使用は、データセンターのいずれか1つがコンピューティングリソースを使い果たすことになるというリスクを低減しながら、システムが単一データセンターの役割を果たすことを可能にする。

【0004】

[0004]この概要は、「詳細な説明」において以下でさらに述べられる概念の選択を簡略化された形で導入するために提供される。この概要は、特許請求される主題の重要な特徴または本質的特徴を識別することを意図されておらず、また特許請求される主題の範囲を決定する目的として使用されることも意図されていない。

【0005】

[0005]上記の利点および特徴ならびに他の利点および特徴が得られてもよい方法を述べるために、様々な実施形態のより詳しい説明が、添付の図面を参照することによって与えられることになる。これらの図面が、見本の実施形態だけを描写し、従って本発明の範囲を限定すると考えられるべきでないということを理解して、実施形態は、付随する図面の使用を通じて具体性および詳細を加えて述べられ、説明されることになる。

【図面の簡単な説明】

【0006】

【図1】[0006]本明細書で述べられるいくつかの実施形態が用いられてもよいコンピューティングシステムを抽象的に例示する図である。

【図2】[0007]複数のクライアントが、複数のデータセンターを有するクラウドコンピューティング環境と相互作用する環境を抽象的に例示する図である。

【図3A】[0008]特定地域に対応するサービスへの要求が、地域データセンターシステム内のデータセンターのいずれかによって処理され、満たされてもよいようにそれぞれ構成される複数のデータセンターを含む地域データセンターシステムを例示する図である。

【図3B】[0009]地域固有のサービス要求を処理するために含まれる関連フローを例示する図であり、受信データセンターは、そのフローの中で処理データセンターを選択する。

【図3C】[0010]地域データセンターシステムおよび地域固有のサービス要求のために含まれる関連フローを例示する図であり、受信データセンターは、そのサービス要求のための処理データセンターとして別のデータセンターを選択する。

【図4】[0011]地域データセンターシステム内のデータセンターのいずれかが、特定地域に対応する地域固有の要求について負荷分散を行うための方法のフローチャートを例示する図である。

【図5】[0012]追加のデータセンターを地域データセンターシステムに加えるための方法のフローチャートを例示する図である。

【発明を実施するための形態】

【0007】

[0013]本明細書で述べられる少なくとも1つの実施形態に従って、地域データセンターシステムが、述べられる。地域データセンターは、各構成データセンターがサービスへの地域固有の要求を処理することができるように、世界の特定地域に対応する複数のデータセンターを含む。データセンターの1つまたは複数の他のデータセンターによって処理されるべき地域固有の要求について負荷分散を行ってもよい。負荷分散を行うために、受信データセンターは、要求をそれ自体が処理してもよいが、しかし場合によっては、要求が地域データセンターシステム内の1つまたは複数の他のデータセンターによって処理されるべきであると判定してもよい。

【0008】

[0014] コンピューティングシステムのいくつかの導入的議論が、図 1 に関して述べられることになる。次いで、地域データセンターの原理が、図 2 から図 5 に関して述べられることになる。

【 0 0 0 9 】

[0015] コンピューティングシステムは、今ではますます多種多様な形を取っている。コンピューティングシステムは例えば、ハンドヘルドデバイス、アプライアンス、ラップトップコンピューター、デスクトップコンピューター、メインフレーム、分散コンピューティングシステム、または従来コンピューティングシステムと考えられていなかったデバイスさえであってもよい。この説明および請求項では、用語「コンピューティングシステム」は、少なくとも 1 つの物理的かつ有形のプロセッサ、およびプロセッサによって実行されてもよいコンピューター実行可能命令をその上に有することができる物理的かつ有形のメモリを含む任意のデバイスまたはシステム（またはそれらの組み合わせ）を含むとして広く定義される。メモリは、任意の形を取ってもよく、コンピューティングシステムの性質および形に依存してもよい。コンピューティングシステムは、ネットワーク環境にわたって分散されてもよく、複数の構成コンピューティングシステムを含んでもよい。

【 0 0 1 0 】

[0016] 図 1 で例示されるように、その最も基本的な構成では、コンピューティングシステム 1 0 0 は典型的には、少なくとも 1 つの処理ユニット 1 0 2 およびメモリ 1 0 4 を含む。メモリ 1 0 4 は、物理的システムメモリであってもよく、それは、揮発性、不揮発性、またはその 2 つのいくつかの組み合わせであってもよい。用語「メモリ」はまた、物理的記憶媒体などの不揮発性大容量記憶装置を指すために本明細書で使用されてもよい。もしコンピューティングシステムが、分散されるならば、処理、メモリおよび / または記憶能力も、同様に分散されてもよい。本明細書で使用される場合、用語「モジュール」または「コンポーネント」は、コンピューティングシステムで実行されるソフトウェアオブジェクトまたはルーチンを指すことができる。本明細書で述べられる異なるコンポーネント、モジュール、エンジン、およびサービスは、コンピューティングシステムで実行されるオブジェクトまたはプロセスとして（例えば、個別のスレッドとして）実装されてもよい。

【 0 0 1 1 】

[0017] 次に来る説明では、実施形態は、1 つまたは複数のコンピューティングシステムによって行われるアクトを参照して述べられる。もしそのようなアクトが、ソフトウェアで実装されるならば、アクトを行う関連コンピューティングシステムの 1 つまたは複数のプロセッサは、実行されるコンピューター実行可能命令を有することに応答してコンピューティングシステムの動作を監督する。そのような動作の例は、データの操作を含む。コンピューター実行可能命令（および操作されるデータ）は、コンピューティングシステム 1 0 0 のメモリ 1 0 4 に記憶されてもよい。コンピューティングシステム 1 0 0 はまた、例えばネットワーク 1 1 0 を通じてコンピューティングシステム 1 0 0 が他のメッセージプロセッサと通信することを可能にする通信チャネル 1 0 8 を含有してもよい。

【 0 0 1 2 】

[0018] 本明細書で述べられる実施形態は、以下でさらに詳細に論じられるように、例えば 1 つまたは複数のプロセッサおよびシステムメモリなどのコンピューターハードウェアを含む、専用または汎用コンピューターを備えるまたは利用してもよい。本明細書で述べられる実施形態はまた、コンピューター実行可能命令および / またはデータ構造を運ぶまたは記憶するための物理的コンピューター可読媒体および他のコンピューター可読媒体も含む。そのようなコンピューター可読媒体は、汎用または専用コンピューターシステムによってアクセスされてもよい任意の入手可能な媒体とすることができる。コンピューター実行可能命令を記憶するコンピューター可読媒体は、物理的記憶媒体である。コンピューター実行可能命令を運ぶコンピューター可読媒体は、伝送媒体である。それ故に、例としてであって、限定でなく、本発明の実施形態は、少なくとも 2 つの明らかに異なる種類のコンピューター可読媒体、すなわちコンピューター記憶媒体および伝送媒体を備えるこ

とができる。

【 0 0 1 3 】

[0019] コンピューター記憶媒体は、RAM、ROM、EEPROM、CD-ROMもしくは他の光ディスク記憶装置、磁気ディスク記憶装置もしくは他の磁気記憶装置、または所望のプログラムコード手段をコンピューター実行可能命令もしくはデータ構造の形で記憶するために使用されてもよく、汎用もしくは専用コンピューターによってアクセスされてもよい任意の他の媒体を含む。

【 0 0 1 4 】

[0020] 「ネットワーク」は、コンピューターシステムおよび/またはモジュールおよび/または他の電子デバイス間で電子データの移送を可能にする1つまたは複数のデータリンクとして定義される。情報が、ネットワークまたは他の通信接続（有線か、無線、または有線もしくは無線の組み合わせ）を通じてコンピューターに転送されるまたは提供されるとき、コンピューターは、その接続を伝送媒体として適切に見る。伝送媒体は、所望のプログラムコード手段をコンピューター実行可能命令またはデータ構造の形で運ぶために使用されてもよく、汎用または専用コンピューターによってアクセスされてもよいネットワークおよび/またはデータリンクを含むことができる。上記の組み合わせもまた、コンピューター可読媒体の範囲内に含まれるべきである。

【 0 0 1 5 】

[0021] さらに、様々なコンピューターシステムコンポーネントに達すると、コンピューター実行可能命令またはデータ構造の形のプログラムコード手段は、伝送媒体からコンピューター記憶媒体に（またはその逆に）自動的に転送されてもよい。例えば、ネットワークまたはデータリンクを通じて受信されるコンピューター実行可能命令またはデータ構造は、ネットワークインターフェースモジュール（例えば、「NIC」）内のRAMにバッファリングされ、次いで最終的にコンピューターシステムRAMおよび/またはコンピューターシステムでの揮発性のより少ないコンピューター記憶媒体に転送されてもよい。それ故に、コンピューター記憶媒体は、また伝送媒体も（または主としてさえ）利用するコンピューターシステムコンポーネントに含まれてもよいことを理解すべきである。

【 0 0 1 6 】

[0022] コンピューター実行可能命令は例えば、プロセッサで実行されるとき、汎用コンピューター、専用コンピューター、または専用処理デバイスにある機能またはある群の機能を行わせる命令およびデータを備える。コンピューター実行可能命令は例えば、二進数、アセンブリ言語などの中間フォーマット命令、またはソースコードさえであってもよい。主題は、構造的特徴および/または方法論的アクトに固有の言語で述べられているけれども、添付の請求項で定義される主題は、必ずしも述べられる特徴または上述のアクトに限定されるとは限らないことを理解すべきである。むしろ、述べられる特徴およびアクトは、請求項を実装する例示的形として開示される。

【 0 0 1 7 】

[0023] 当業者は、本発明が、パーソナルコンピューター、デスクトップコンピューター、ラップトップコンピューター、メッセージプロセッサ、ハンドヘルドデバイス、マルチプロセッサシステム、マイクロプロセッサに基づくまたはプログラム可能な消費者向け電気機器、ネットワークPC、ミニコンピューター、メインフレームコンピューター、携帯電話、PDA、ポケットベル、ルーター、スイッチ、および同様のものを含む、多くの種類のコンピューターシステム構成を有するネットワークコンピューティング環境で実施されてもよいことを認識するであろう。本発明はまた、ネットワークを通じてリンクされる（有線データリンクか、無線データリンクによって、または有線および無線データリンクの組み合わせによって）ローカルコンピューターシステムおよびリモートコンピューターシステムが、両方ともタスクを行う分散システム環境で実施されてもよい。分散システム環境では、プログラムモジュールは、ローカルメモリ記憶デバイスおよびリモートメモリ記憶デバイスの両方に置かれてもよい。

【 0 0 1 8 】

10

20

30

40

50

[0024] 図 2 は、本明細書で述べられる原理が用いられてもよい環境 200 を抽象的に例示する。環境 200 は、インターフェース 202 を使用してクラウドコンピューティング環境 210 と相互作用する複数のクライアント 201 を含む。環境 200 は、3 つのクライアント 201 A、201 B および 201 C を有するとして例示されるが、省略記号 201 D は、本明細書で述べられる原理が、インターフェース 202 を通じてクラウドコンピューティング環境 210 とインターフェースで接続するクライアントの数に限定されないことを表す。クラウドコンピューティング環境 210 は、サービスをクライアント 201 にオンデマンドで提供してもよく、それ故にクラウドコンピューティング環境 210 からサービスを受信するクライアント 201 の数は、時間とともに変化してもよい。

【0019】

10

[0025] 各クライアント 201 は例えば、図 1 のコンピューティングシステム 100 について上で述べられたように構造化されてもよい。別法としてまたは加えて、クライアントは、インターフェース 202 を通じてクラウドコンピューティング環境 210 とインターフェースで接続するアプリケーションまたは他のソフトウェアモジュールであってもよい。インターフェース 202 は、アプリケーションプログラムインターフェースを使用することができる任意のコンピューティングシステムまたはソフトウェアエンティティがクラウドコンピューティング環境 210 と通信してもよいような方法で定義されるアプリケーションプログラムインターフェースであってもよい。

【0020】

[0026] クラウドコンピューティング環境は、分散されてもよく、国際的に分散されかつ / または複数の組織を横断して保有されるコンポーネントを有してさえよい。この説明および次に来る請求項では、「クラウドコンピューティング」は、構成可能なコンピューティングリソース（例えば、ネットワーク、サーバー、記憶装置、アプリケーション、およびサービス）の共有プールへのオンデマンドネットワークアクセスを可能にするためのモデルとして定義される。「クラウドコンピューティング」の定義は、適切に展開されるときにそのようなモデルから得ることができる他の多数の利点のいずれにも限定されない。

20

【0021】

[0027] 例えば、クラウドコンピューティングは現在、構成可能なコンピューティングリソースへの遍在する、便利なオンデマンドアクセスを提示するために市場で用いられている。さらに、構成可能なコンピューティングリソースの共有プールは、仮想化を介して迅速に準備され、低管理努力またはサービスプロバイダー相互作用でリリースされ、次いでそれに応じて拡大縮小されてもよい。

30

【0022】

[0028] クラウドコンピューティングモデルは、オンデマンドセルフサービス、広域ネットワークアクセス、リソースプール化、高速融通性、計数サービス、その他などの様々な特徴から構成されてもよい。クラウドコンピューティングモデルはまた、例えばサービス型ソフトウェア（「SaaS」）、サービス型プラットフォーム（「PaaS」）、およびサービス型インフラストラクチャー（「IaaS」）などの様々なサービスモデルの形になってもよい。クラウドコンピューティングモデルはまた、プライベートクラウド、コミュニティクラウド、パブリッククラウド、ハイブリッドクラウド、その他などの異なる展開モデルを使用して展開されてもよい。この説明および請求項では、「クラウドコンピューティング環境」は、クラウドコンピューティングが用いられる環境である。

40

【0023】

[0029] システム 210 は、処理、メモリ、記憶装置、帯域幅、その他などの、対応するコンピューティングリソースをそれぞれ含む複数のデータセンター 211 を含む。データセンター 211 は、より大きいデータセンター 211 A、211 B および 211 C を含むが、省略記号 211 D は、データセンター群 211 内の最初のデータセンターの数に関して制約がないことを表す。また、データセンター 211 は、より小さいデータセンター 211 a から 211 i も含むが、省略記号 211 j は、データセンター群 211 内のより小さいデータセンターの数に関して制約がないことを表す。データセンター 211 のそれぞ

50

れはおそらく、図 1 のコンピューティングシステム 100 について上で述べられたようにそれぞれ構造化されてもよい非常に多くのホストコンピューティングシステムを含んでもよい。データセンター 211 は、地理的に分散されてもよく、もしクラウドコンピューティング環境 200 が地球に及ぶならば、おそらく世界全体にわたってさえ分散されてもよい。

【0024】

[0030]クラウドコンピューティング環境 200 はまた、サービス 212 も含む。示される例では、サービス 200 は、5 つの個別サービス 212 A、212 B、212 C、212 D および 212 E を含むが、省略記号 212 F は、本明細書で述べられる原理が、システム 210 でのサービスの数に限定されないことを表す。サービス調整システム 213 は、データセンター 211 およびサービス 212 と通信し、それによってクライアント 201 によって要求されるサービス、および要求されるサービスの前提条件のこともある他のサービス（認証、課金、その他などの）を提供する。

【0025】

[0031]図 3 A は、特定地域に対応するサービスへの要求（以下ではまた「地域固有の要求」または「地域固有のサービス要求」とも呼ばれる）が地域データセンターシステム内のデータセンターのいずれかによって処理され、満たされてもよいようにそれぞれ構成される複数のデータセンターを含む地域データセンターシステム 300 を例示する。例示されるシステム 300 は、5 つのデータセンター 301 から 305 を含むとして示される。しかしながら、これは、例示目的のためだけである。省略記号 306 は、ここで述べられる原理が、そのような構成データセンターの数にかかわらず、複数の構成データセンターを含む任意のデータセンターシステムに当てはまることを表す。

【0026】

[0032]データセンターのそれぞれの間の通信は、データセンター間の通信がいくらかの確率で特定のしきい値を下回るレイテンシで動作するような特定のレイテンシエンベロープ内で動作することを検証される。特定のレイテンシしきい値は好ましくは、地域固有のサービス要求を提出する平均的顧客が、単一データセンターとは対照的に、複数の分散データセンターから成る地域データセンターシステムを有意に区別することができないほど十分に小さい。例えば、レイテンシしきい値は、数ミリ秒以下の程度のことでもある。それ故に、データセンター 301 から 306 はおそらく、互いに何マイルも離れて分散されることもあるが、それらは、顧客の観点からすれば単一データセンターとして動作するように見えることもある。

【0027】

[0033]データセンターは、同じサイズである必要がないが、しかし地域データセンターシステム 300 の所有者および / または管理者によって望まれる任意の仕方でサイズを決められてもよい。例えば、一実施形態では、地域は、ほんの単一のより大きいデータセンターだけから始めてもよい。次いで、需要がその特定地域で増えると、追加のデータセンターが加えられてもよい。その場合、おそらく追加のデータセンターは、より大きい最初のデータセンターよりも計算能力が小さくてもよい。例えば、図 3 では、データセンター 301 は、より小さいデータセンター 302 から 305 よりも大きいとして示される。おそらくデータセンター 301 は、地域の最初のデータセンターであり、データセンター 302 から 305 は、需要が特定地域で増加したとき、その後に加えられた。一実施形態では、データセンター 302 から 305 の少なくとも 1 つは、最初のデータセンター 301 よりも半分未満の計算能力を有することもある。追加のデータセンターを地域データセンターシステムに加えるための方法は、図 5 を参照して以下で述べられることになる。例えば、データセンター 301 は、図 2 のデータセンター 211 A こともあり、一方データセンター 302 から 305 は、例えば図 2 のデータセンター 211 a、211 b、211 c および 211 d のこともある。別法として、データセンターシステム 300 は、サービス調整システム 213 の観点からすれば単一データセンターであると概念的に考えられてもよい。

10

20

30

40

50

【 0 0 2 8 】

[0034] 図 4 は、地域データセンターシステム内のデータセンターのいずれかが、特定地域に対応する地域固有の要求について負荷分散を行うための方法 4 0 0 のフローチャートを例示する。方法 4 0 0 は、図 3 A の地域データセンターシステム 3 0 0 内で行われてもよいので、方法 4 0 0 は、図 3 A をしばしば参照して今から述べられることになる。

【 0 0 2 9 】

[0035] 例えば、地域データセンター 3 0 0 のデータセンターは、方法 4 0 0 を行う。図 3 A では、データセンター 3 0 1 は、データセンター 3 0 1 が受信する地域固有のサービス要求に関して本方法を行う。しかしながら、本明細書で述べられる原理は、地域データセンターシステム 3 0 0 によって受け持たれる特定地域のための地域固有のサービス要求を受信するように構成されるデータセンターの特定のサブセットに限定されない。例えば、おそらくデータセンター 3 0 1 は、地域データセンターシステム 3 0 0 内で満たされるべき地域固有のサービス要求を受信する地域データセンターシステム 3 0 0 の唯一のデータセンターである。その一方で、おそらく地域データセンターシステム 3 0 0 のすべてのデータセンターは、地域データセンターシステム 3 0 0 内で満たされるべき地域固有のサービス要求を受信してもよい。その間で、おそらく地域データセンターシステム 3 0 0 のデータセンターのサブセットだけが、地域データセンターシステム 3 0 0 内で満たされるべき地域固有のサービス要求を受信してもよい。とにかく、方法 4 0 0 は、地域データセンターシステム 3 0 0 内のそのような要求を受信するように構成される任意のデータセンターによって同時に行われてもよい。

【 0 0 3 0 】

[0036] 図 3 A および図 3 B の具体的な例を参照すると、データセンター 3 0 1 は、矢印 3 1 0 によって表されるように地域データセンターシステム 3 0 0 によって受け持たれる特定地域のための地域固有のサービス要求を受信する（アクト 4 0 1）。これらの地域固有のサービス要求 3 1 0 は、地域データセンターシステム 3 0 0 内のデータセンターのいずれかによって満たされてもよい。

【 0 0 3 1 】

[0037] 受信データセンター（すなわち、地域固有のサービス要求を受信するデータセンター）は次いで、受信される地域固有のサービス要求の少なくともいくつかの処理について負荷分散を行う（アクト 4 1 0）。例えば、いくつかの地域固有のサービス要求については負荷分散を行わないなんらかの理由があってもよい。例えば、いくつかの要求は、状態に敏感なこともあり、データセンターの特定の 1 つ内に存在するデータまたは他の状態に依存することもある。しかしながら、データセンターによって受信される地域固有のサービス要求の他のものについては、データセンターは、地域固有のサービス要求の処理について負荷分散を行う（アクト 4 1 0）。例えば、図 3 A を参照すると、データセンター 3 0 1 は、入ってくる地域固有のサービス要求 3 1 0 の処理の負荷分散を行う。

【 0 0 3 2 】

[0038] 図 4 を参照すると、アクト 4 1 0 の内容は、地域固有のサービス要求の処理の負荷分散を行うことの一部であるアクトを含む（アクト 4 1 0）。例えば、この負荷分散の一部として、受信データセンターは、地域固有のサービス要求の処理を実際に行うためにデータセンターの 1 つまたは複数を選択する（アクト 4 1 1）。選択されるデータセンターはまた、本明細書では「処理データセンター」とも呼ばれることになる。

【 0 0 3 3 】

[0039] 例えば、図 3 A を参照すると、受信データセンター 3 0 1 は、矢印 3 1 1 によって表されるように、地域固有のサービス要求のいくつかのための処理データセンターとしてそれ自体を選択することもある。別法としてまたは加えて、受信データセンター 3 0 1 はまた、矢印 3 1 2 によって表されるように、地域固有のサービス要求のいくつかのための処理データセンターとしてデータセンター 3 0 2 を選択し、そのような地域固有のサービス要求をデータセンター 3 0 2 に転送することもある。受信データセンター 3 0 1 はまた、矢印 3 1 3 によって表されるように、地域固有のサービス要求のいくつかのための処

理データセンターとしてデータセンター 303 を選択し、そのような地域固有のサービス要求をデータセンター 303 に転送することもある。受信データセンター 301 はまた、矢印 314 によって表されるように、地域固有のサービス要求のいくつかのための処理データセンターとしてデータセンター 304 を選択し、そのような地域固有のサービス要求をデータセンター 304 に転送することもある。受信データセンター 301 はまた、矢印 315 によって表されるように、地域固有のサービス要求のいくつかのための処理データセンターとしてデータセンター 305 を選択し、そのような地域固有のサービス要求をデータセンター 305 に転送することもある。

【0034】

[0040] 受信データセンターは次いで、地域固有のサービス要求が処理センターに転送される仕方では負荷分散を行う（アクト 420）。処理データセンターによって扱われるべきそれらの要求について、受信データセンターと処理データセンターとの間に直接の通信チャンネルがあるけれども、要求のいくつかは、地域データセンターのなお別のデータセンターを介してディスパッチされてもよい。例えば、受信データセンターは、要求を処理データセンターに伝達するときにランダム化またはラウンドロビン二相通信を使用してもよい。そのようなことは、地域データセンターシステム内のすべてのデータセンター間での帯域幅のより効率的な利用を可能にする。

【0035】

[0041] 図 4 を参照すると、アクト 420 の内容は、処理データセンターへの地域固有のサービス要求の転送の負荷分散を行うことの一部であるアクトを含む（アクト 420）。負荷分散のレベルは、例えばプロトコルスタックでのネットワークレベルで生じてもよい。アクト 420 の内容は、地域固有のサービス要求の処理に寄与すべき 1 つまたは複数のデータセンターのそれぞれについて行われてもよい。

【0036】

[0042] もし処理データセンターが、受信データセンターと同じであるならば（決定ブロック 421 で「イエス」）、その時受信データセンターは、地域固有のサービス要求のすべてまたは一部をそれ自体が処理して満たすことになる（アクト 430）。その場合、要求のその部分についての転送プロセスの実質的な負荷分散は、含まれるデータセンター間通信がないので、行われないことになる。

【0037】

[0043] もし処理データセンターが、受信データセンターと異なるならば（決定ブロック 421 で「ノー」）、受信データセンターはその時、中間データセンターが、地域固有のサービス要求を転送するのに使用されるべきかどうかを判定する（決定ブロック 422）。もし否であるならば（決定ブロック 422 での「ノー」）、その時負荷分散は、最初に中間データセンターから反映されることなく地域固有の要求を処理データセンターに転送すること（アクト 423）によってさらに行われる。もし中間データセンターが、地域固有のサービス要求を転送するのに使用されるべきであるならば（決定ブロック 422 での「イエス」）、受信データセンターは、中間データセンターを選択し（アクト 424）、地域固有のサービス要求が処理される（アクト 430）ところの処理データセンターへの反映（アクト 423）のために地域固有の要求を中間データセンターに転送する（アクト 425）。

【0038】

[0044] 図 3 B は、地域データセンターシステム 300 および地域固有のサービス要求のために含まれる関連フローを例示し、受信データセンター 301 は、そのサービス要求のための処理データセンターとしてデータセンター 304 を選択する。結果として生じる転送動作は、図 3 A で矢印 314 によって抽象的に表される。しかしながら、図 3 B は、プロトコルスタックのネットワークレベルにおいて受信データセンター 301 から処理データセンター 304 への地域固有のサービス要求の転送を負荷分散するために様々な方法があることを例示する。図 3 B での様々なフローは、図 4 の方法 400 を参照して今から説明されることになる。それぞれの場合に、受信データセンター 301 は、地域固有のサー

10

20

30

40

50

ビス要求を処理データセンター 304 に転送することを思い出されたい。

【0039】

[0045] 矢印 322A および 322B は、データセンター 302 が中間のものとして選択される (アクト 424) 場合を表し、その場合サービス要求は、矢印 322A によって表されるように中間データセンター 302 に転送され (アクト 425)、矢印 322B によって表されるように処理データセンター 304 に反映される (アクト 423)。

【0040】

[0046] 矢印 323A および 323B は、データセンター 303 が中間のものとして選択される (アクト 424) 場合を表し、その場合サービス要求は、矢印 323A によって表されるように中間データセンター 303 に転送され (アクト 425)、矢印 323B によって表されるように処理データセンター 304 に反映される (アクト 423)。

10

【0041】

[0047] 矢印 324 は、中間データセンターが選択されない (決定ブロック 422 での「ノー」) 場合を表し、その場合地域固有のサービス要求は、中間のものを使用することなく処理データセンター 304 に転送される (アクト 423)。

【0042】

[0048] 矢印 325A および 325B は、データセンター 305 が中間のものとして選択される (アクト 424) 場合を表し、その場合サービス要求は、矢印 325A によって表されるように中間データセンター 305 に転送され (アクト 425)、矢印 325B によって表されるように処理データセンター 304 に反映される (アクト 423)。

20

【0043】

[0049] 図 3C は、地域データセンターシステム 300 および地域固有のサービス要求のために含まれる関連フローを例示し、受信データセンター 301 は、そのサービス要求のための処理データセンターとしてデータセンター 303 を選択する。結果として生じる転送動作は、図 3A で矢印 313 によって抽象的に表される。しかしながら、図 3C は、プロトコルスタックのネットワークレベルにおいて受信データセンター 301 から処理データセンター 303 への地域固有のサービス要求の転送を負荷分散するためにいろいろな方法があることを例示する。図 3C での様々なフローは、図 400 の方法 400 を参照して今から説明されることになる。

【0044】

30

[0050] 矢印 332A および 332B は、データセンター 302 が中間のものとして選択される (アクト 424) 場合を表し、その場合サービス要求は、矢印 332A によって表されるように中間データセンター 302 に転送され (アクト 425)、矢印 332B によって表されるように処理データセンター 303 に反映される (アクト 423)。

【0045】

[0051] 矢印 333 は、中間データセンターが選択されない (決定ブロック 422 での「ノー」) 場合を表し、その場合地域固有のサービス要求は、中間のものを使用することなく処理データセンター 303 に転送される (アクト 423)。

【0046】

[0052] 矢印 334A および 334B は、データセンター 304 が中間のものとして選択される (アクト 424) 場合を表し、その場合サービス要求は、矢印 334A によって表されるように中間データセンター 304 に転送され (アクト 425)、矢印 334B によって表されるように処理データセンター 303 に反映される (アクト 423)。

40

【0047】

[0053] 矢印 335A および 335B は、データセンター 305 が中間のものとして選択される (アクト 424) 場合を表し、その場合サービス要求は、矢印 335A によって表されるように中間データセンター 305 に転送され (アクト 425)、矢印 335B によって表されるように処理データセンター 303 に反映される (アクト 423)。

【0048】

[0054] 前に述べられたように、図 3A から図 3C の例では、おそらくデータセンター 3

50

01は、特定地域にサービスを提供する最初のデータセンターであり、次いで後に、データセンター302から305が、需要を満たすために加えられた。図5は、特定地域に対応する地域固有のサービス要求を受け持つ地域データセンターシステムでいくつかのデータセンターを増強するための方法500のフローチャートを例示する。

【0049】

[0055]特定地域で容認されるべきレイテンシしきい値が、(アクト501)で決定される。例えば、このレイテンシしきい値は、地域データセンターシステムの所有者または管理者にとって許容できる任意の値であってもよいが、しかし実施形態では、数ミリ秒以下である。

【0050】

[0056]方法300はまた、1つまたは複数の既存のデータセンターが利用率を上回ると判定すること(アクト502)も含む。この判定は好ましくは、既存のデータセンターが完全に利用されるより前に方法500の残りをを行うための準備ができるほど十分に早く行われるべきである。このアクト502は、アクト501とアクト502との間に時間的依存関係がないので、レイテンシしきい値の決定(アクト501)と並列に示される。

【0051】

[0057]オプションとして、方法300はまた、追加のデータセンターを構築すること(アクト503)も含む。この構築(アクト503)は、アクト501、502、および503の間に厳密な時間的依存関係がないので、レイテンシしきい値の決定(アクト501)および利用率の判定と並列に示されるが、ただし既存のデータセンターの完全な利用が達せられる前に、アクト503が必要ならば行われてもよいように、アクト502は、十分前もって生じることが好ましいという、すでに記されていることを除く。

【0052】

[0058]方法500はまた、特定地域に対応する地域固有のサービス要求を処理するように追加のデータセンターを構成すること(アクト505)も含む。これは、追加のデータセンターがそのような要求を実際に受信することを意味せず、追加のデータセンターが、もしそのような要求が受信されるならば、それら进行处理する能力があるということだけを意味する。

【0053】

[0059]オプションとして、方法500はまた、追加のデータセンターと地域データセンターシステムの1つまたは複数の既存のデータセンターとの間に1つまたは複数の通信チャネルの少なくとも一部分を実際に確立すること(アクト504)も含む。この場合もやはり、このアクト504と他のアクト501から503との間に厳密な時間関係はなく、それ故にこのアクトは、再び並列に示される。

【0054】

[0060]方法500はまた、追加のデータセンターのレイテンシがレイテンシしきい値を下回ることを確実にするために、追加のデータセンターと地域データセンターシステムの1つまたは複数の既存のデータセンターとの間のネットワークレイテンシを測定すること(アクト506)も含む。いったんこれが、確保されると、全体としての地域データセンターは、地域データセンターシステムによって受信される地域固有のサービス要求の少なくともいくつか、処理のために追加のデータセンターに送られるように構成されてもよい(アクト507)。

【0055】

[0061]それに応じて、システムが外部からは単一データセンターのように見えてもよい、地域データセンターシステムが、述べられる。さらに、データセンターは、需要が増加すると、追加のデータセンターで動的に増強されてもよい。本発明は、その趣旨または本質的特徴から逸脱することなく他の特定の形で具体化されてもよい。述べられる実施形態は、あらゆる点で説明に役立つだけで、制約するものではないと考えるべきである。従って、本発明の範囲は、前述の説明によってよりもむしろ添付の請求項によって示される。請求項の等価性の意味および範囲内に入るすべての変形は、請求項の範囲内に包含される

10

20

30

40

50

べきである。

【 図 1 】

コンピューティングシステム
100

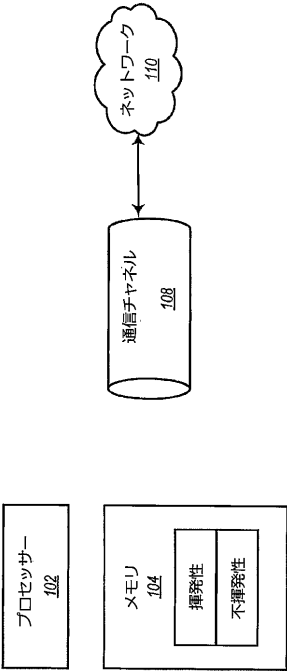


Figure 1

【 図 2 】

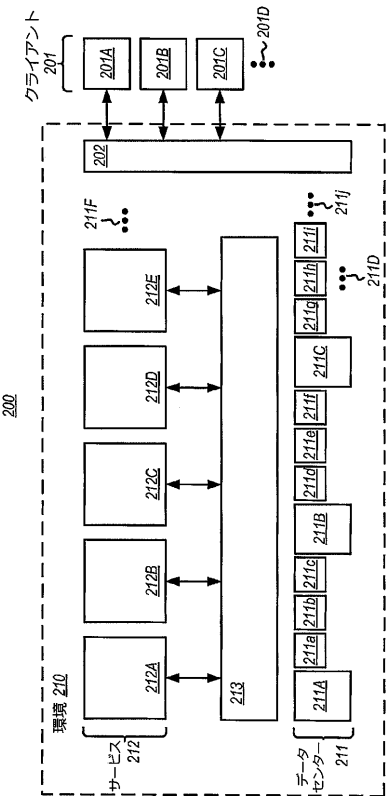


Figure 2

【図 3 A】

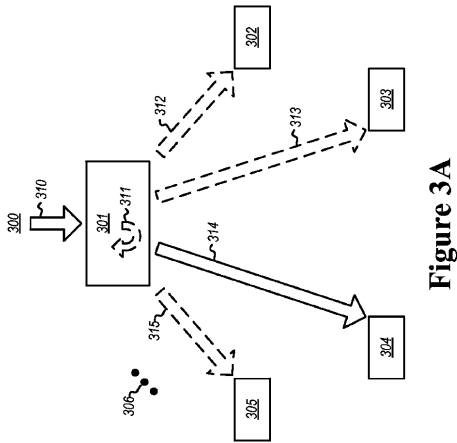


Figure 3A

【図 3 B】

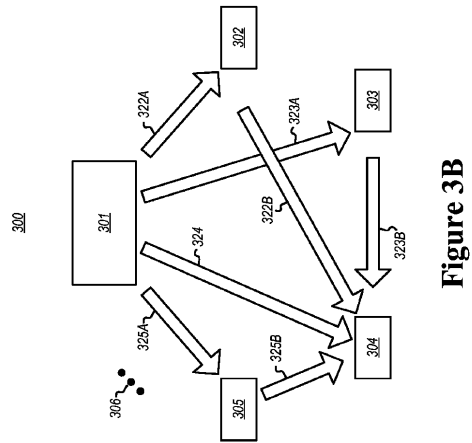


Figure 3B

【図 3 C】

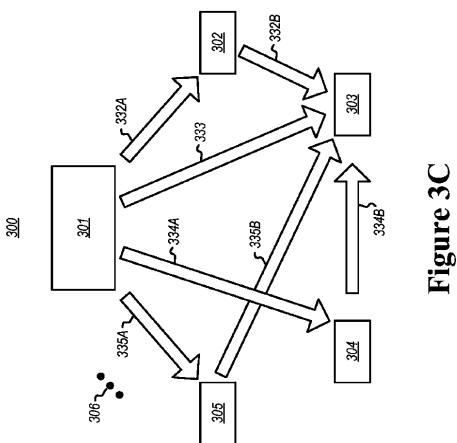


Figure 3C

【図 4】

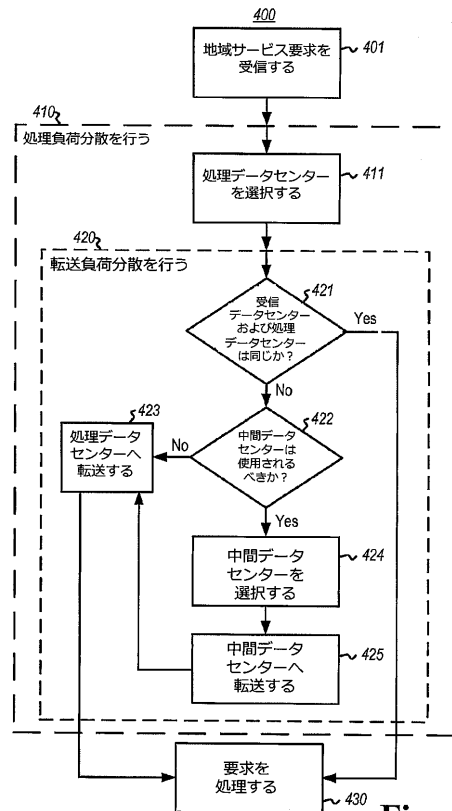


Figure 4

【図 5】

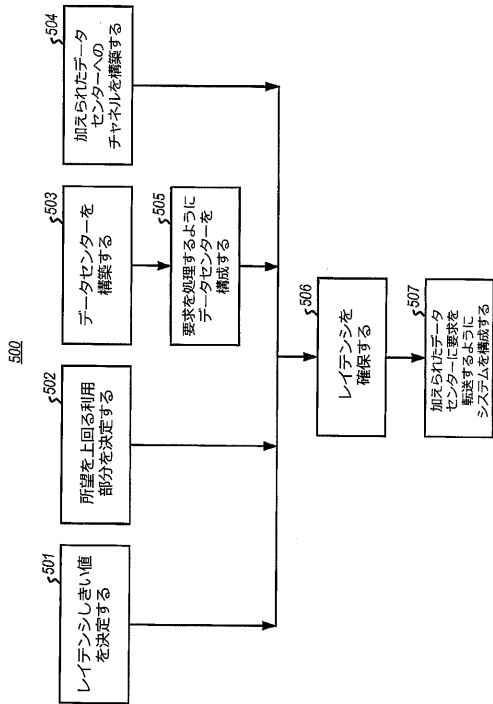


Figure 5

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2014/020834

A. CLASSIFICATION OF SUBJECT MATTER

INV. G06F9/50
ADD.

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)

G06F

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

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

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6 249 801 B1 (ZISAPEL ROY [IL] ET AL) 19 June 2001 (2001-06-19)	1,3-8,10
Y	column 6, line 30 - column 7, line 60; figures 2A-2F	2,9
Y	----- US 2010/036956 A1 (NISHIKAWA YASUHIRO [JP]) 11 February 2010 (2010-02-11) paragraph [0006] - paragraph [0007]; figure 2	2,9
A	----- US 6 647 408 B1 (RICART GLENN [US] ET AL) 11 November 2003 (2003-11-11) the whole document	1-10
	----- -/--	

☒ 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

"Z" document member of the same patent family

Date of the actual completion of the international search

27 May 2014

Date of mailing of the international search report

03/06/2014

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Mühlenbrock, Martin

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2014/020834

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>K P Bubendorfer: "Resource Based Policies for Load Distribution",</p> <p>31 August 1996 (1996-08-31), XP055036862, Victoria University of Wellington</p> <p>Retrieved from the Internet: URL:http://www.mcs.vuw.ac.nz/comp/graduate s/archives/msc/KrisBubendorfer.ps.gz [retrieved on 2012-08-30] the whole document</p> <p>-----</p>	1-10

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2014/020834

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6249801	B1	19-06-2001	US 6249801 B1 19-06-2001
			US 2002103846 A1 01-08-2002
			US 2005022203 A1 27-01-2005
US 2010036956	A1	11-02-2010	JP 5246157 B2 24-07-2013
			US 2010036956 A1 11-02-2010
			WO 2008129597 A1 30-10-2008
US 6647408	B1	11-11-2003	US 6647408 B1 11-11-2003
			US 7356819 B1 08-04-2008

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(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), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), 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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, 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, 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

(74)代理人 100153028

弁理士 上田 忠

(72)発明者 キム, チャンフーン

アメリカ合衆国ワシントン州 9 8 0 5 2 - 6 3 9 9 , レッドモンド, マイクロソフト コーポレーション, エルシーエイ - インターナショナル・パテンツ(8 / 1 1 7 2)

(72)発明者 カリディ, ユーセフ・エイ

アメリカ合衆国ワシントン州 9 8 0 5 2 - 6 3 9 9 , レッドモンド, マイクロソフト コーポレーション, エルシーエイ - インターナショナル・パテンツ(8 / 1 1 7 2)

(72)発明者 スリヤナラヤナン, グハン

アメリカ合衆国ワシントン州 9 8 0 5 2 - 6 3 9 9 , レッドモンド, マイクロソフト コーポレーション, エルシーエイ - インターナショナル・パテンツ(8 / 1 1 7 2)