

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7445740号  
(P7445740)

(45)発行日 令和6年3月7日(2024.3.7)

(24)登録日 令和6年2月28日(2024.2.28)

(51)国際特許分類		F I	
G 0 6 F	11/34 (2006.01)	G 0 6 F	11/34 1 3 3
G 0 6 F	11/30 (2006.01)	G 0 6 F	11/34 1 4 7
G 0 6 F	16/13 (2019.01)	G 0 6 F	11/30 1 4 0 A
		G 0 6 F	16/13 1 0 0

請求項の数 13 (全24頁)

(21)出願番号	特願2022-501296(P2022-501296)	(73)特許権者	502303739 オラクル・インターナショナル・コーポ レイション アメリカ合衆国カリフォルニア州 9 4 0 6 5 レッドウッド・シティー, オラクル ・パークウェイ 5 0 0
(86)(22)出願日	令和2年7月15日(2020.7.15)	(74)代理人	110001195 弁理士法人深見特許事務所
(65)公表番号	特表2022-541424(P2022-541424 A)	(72)発明者	ソンダー, サンジープ・アール アメリカ合衆国、9 4 0 6 5 カリフォ ルニア州、レッドウッド・シティー、オ ラクル・パークウェイ、5 0 0
(43)公表日	令和4年9月26日(2022.9.26)	(72)発明者	ミルズ, マーク・エバン アメリカ合衆国、9 4 0 6 5 カリフォ ルニア州、レッドウッド・シティー、オ ラクル・パークウェイ、5 0 0 最終頁に続く
(86)国際出願番号	PCT/US2020/042153		
(87)国際公開番号	WO2021/016019		
(87)国際公開日	令和3年1月28日(2021.1.28)		
審査請求日	令和5年2月17日(2023.2.17)		
(31)優先権主張番号	16/517,236		
(32)優先日	令和1年7月19日(2019.7.19)		
(33)優先権主張国・地域又は機関	米国(US)		

(54)【発明の名称】 クラウドストレージソフトウェアアプライアンスを構成するための方法および装置

(57)【特許請求の範囲】

【請求項 1】

方法であって、

構成パラメータ値と、作業負荷値と、性能レベルを表わす性能値との間の相関関係を認識するように機械学習分析モデルを訓練することによって、機械訓練済み分析モデルを作成するステップと、

構成パラメータ値の第1のセットで構成されるシステムによって、指定された性能レベルで実行され得る作業の量を表わす作業負荷値の推奨されるセットについての第1の要求を構成アドバイザにおいて受信するステップと、

前記システムによって実行されるべき作業の第1の量を表わす作業負荷値の第1のセットを前記構成アドバイザによって選択するステップと、

前記構成アドバイザによって、前記構成パラメータ値の第1のセットおよび前記作業負荷値の第1のセットを含む第1の入力を前記機械訓練済み分析モデルにサブミットするステップと、

前記構成アドバイザによって、前記構成パラメータ値の第1のセットおよび前記作業負荷値の第1のセットに対応する性能値の第1のセットを含む出力を前記機械訓練済み分析モデルから受信するステップと、

前記構成アドバイザによって、前記性能値の第1のセットが前記指定された性能レベルを満たさないことを判断するステップと、

前記システムによって実行されるべき作業の第2の量を表わす作業負荷値の第2のセッ

10

20

トを前記構成アドバイザーによって選択するステップと、

前記構成パラメータ値の第 1 のセットおよび前記作業負荷値の第 2 のセットを含む第 2 の入力を前記機械訓練済み分析モデルにサブミットするステップと、

前記構成パラメータ値の第 1 のセットおよび前記作業負荷値の第 2 のセットに対応する性能値の第 2 のセットを前記機械訓練済み分析モデルからの出力として受信するステップと、

前記構成アドバイザーによって、前記性能値の第 2 のセットが前記指定された性能レベルを満たすかまたは超えることを判断するステップと、

前記作業負荷値の第 2 のセットを含む前記作業負荷値の推奨されるセットを含む前記第 1 の要求に応答するステップとを含み、

前記作業負荷値の推奨されるセットは、少なくとも複数の

- ( a ) 単位時間当たりの同時ユーザアクセスの数、
- ( b ) ファイルシステム当たりのファイルの数、
- ( c ) 読出し動作に対する書込み動作の割合、
- ( d ) ファイルメタデータ読出し動作の数、
- ( e ) ファイルメタデータ書込み動作の数、および
- ( f ) ファイル付加動作の数、

を含む、方法。

#### 【請求項 2】

前記構成パラメータ値の第 1 のセットは、第 2 のネットワーク上のリモートアプリケーションサーバにアプリケーション要求を送信するアプリケーションゲートウェイを第 1 のネットワーク上に構成するために用いられ、前記第 1 のネットワークは前記第 2 のネットワークとは物理的および仮想的に異なっている、請求項 1 に記載の方法。

#### 【請求項 3】

前記アプリケーションゲートウェイは、前記第 1 のネットワークと同じ仮想ネットワーク内で動作するアプリケーションにファイルサービスを提供する、請求項 2 に記載の方法。

#### 【請求項 4】

前記リモートアプリケーションサーバはストレージオブジェクトインターフェイスを前記アプリケーションゲートウェイに提供し、前記アプリケーションゲートウェイは、前記リモートアプリケーションサーバによって提供される前記ストレージオブジェクトインターフェイスを呼出すことによってファイルサービス要求をサブする、請求項 2 に記載の方法。

#### 【請求項 5】

アプリケーションゲートウェイについての前記構成パラメータ値は、少なくとも複数の

- ( a ) データキャッシュサイズ、
  - ( b ) メタデータサイズ、
  - ( c ) ログファイルサイズ、
  - ( d ) スレッドの数、
  - ( e ) コアの数および対応するコア速度、
  - ( f ) メモリのサイズ、
  - ( g ) ディスクコントローラモデル、
  - ( h ) ネットワークカードモデル、および、
  - ( i ) 前記アプリケーションゲートウェイとリモートサーバとの間の接続の帯域幅、
- についての値を含む、請求項 1 に記載の方法。

#### 【請求項 6】

前記指定された性能レベルは、

- ( a ) I / O レイテンシ、
- ( b ) I / O スループット、および
- ( c ) 秒当たりの I / O 動作、ならびに
- ( d ) I O タイムアウトの数、

10

20

30

40

50

のうち少なくとも1つの性能値によって表わされる、請求項1～5のいずれか1項に記載の方法。

【請求項7】

前記機械学習分析モデルを訓練することは、入力値の複数のセットを前記機械学習分析モデルにサブミットすることを含み、入力値の各セットは、構成パラメータ値、作業負荷値および性能値を含み、前記性能値は、それぞれの前記構成パラメータ値で構成される、アプリケーションゲートウェイ上のそれぞれの前記作業負荷値によって表わされる量の作業を実行しながら経験的に観察されたものである、請求項1に記載の方法。

【請求項8】

前記方法はさらに、

性能値の第3のセットによって表わされる指定された性能レベルで、指定された作業負荷値の第3のセットによって表わされる所望の量の作業を実行することができる推奨された構成についての第2の要求を受信するステップと、

前記作業負荷値の第3のセットおよび前記性能値の第3のセットを前記機械訓練済み分析モデルへの入力として提供するステップと、

前記作業負荷値の第3のセットおよび前記性能値の第3のセットに対応する構成パラメータ値の第2のセットを前記機械訓練済み分析モデルからの出力として受信するステップと、

前記構成パラメータ値の第2のセットを含む推奨された構成を含む前記第2の要求に回答するステップとを含む、請求項1～7のいずれか1項に記載の方法。

【請求項9】

前記方法はさらに、

作業負荷値の第3のセットによって表わされる作業の第3の量を、構成パラメータ値の第3のセットで構成されたアプリケーションゲートウェイによって実行するときに、達成可能な性能レベルの推定値についての第3の要求を受信するステップと、

前記構成パラメータ値の第3のセットおよび前記作業負荷値の第3のセットを前記機械訓練済み分析モデルへの入力として提供するステップと、

前記構成パラメータ値の第3のセットおよび前記作業負荷値の第3のセットに対応する性能値の第3のセットを前記機械訓練済み分析モデルからの出力として受信するステップと、

前記性能値の第3のセットを含む達成可能な性能レベルの推定値を含む前記第3の要求に回答するステップとを含む、請求項1に記載の方法。

【請求項10】

前記方法はさらに、

前記作業負荷値の第2のセットを含む前記作業負荷値の推奨されるセットを含む前記第1の要求に回答する前に、

前記システムによって実行されるべき作業の第3の量を表わす作業負荷値の第3のセットを、前記構成アドバイザーによって選択するステップと、

前記構成パラメータ値の第1のセットおよび前記作業負荷値の第3のセットを含む第3の入力を前記機械訓練済み分析モデルにサブミットするステップと、

前記構成パラメータ値の第1のセットおよび前記作業負荷値の第3のセットに対応する性能値の第3のセットを、前記機械訓練済み分析モデルからの出力として受信するステップと、

前記構成アドバイザーによって、前記性能値の第3のセットが前記指定された性能レベルを満たすかまたは超えることを判断するステップと、

前記作業負荷値の第2のセットによって表わされる前記作業の第2の量が、前記作業負荷値の第3のセットによって表わされる前記作業の第3の量よりも大きいと判断するステップとを含む、請求項1～7のいずれか1項に記載の方法。

【請求項11】

前記構成パラメータ値の第1のセットは、第2のネットワーク上のリモートアプリケー

10

20

30

40

50

ションサーバにアプリケーション要求を送信するアプリケーションゲートウェイを第1のネットワーク上に構成するために用いられ、前記第1のネットワークは前記第2のネットワークとは物理的および仮想的に異なっており、

前記リモートアプリケーションサーバは、ストレージオブジェクトインターフェイスを前記アプリケーションゲートウェイに提供し、前記アプリケーションゲートウェイは、前記リモートアプリケーションサーバによって提供される前記ストレージオブジェクトインターフェイスを呼出すことによってファイルサービス要求をサブスクリプションし、

アプリケーションゲートウェイについての前記構成パラメータ値は、少なくとも複数の

- (a) データキャッシュサイズ、
- (b) メタデータサイズ、
- (c) ログファイルサイズ、
- (d) スレッドの数、
- (e) コアの数および対応するコア速度、
- (f) メモリのサイズ、
- (g) ディスクコントローラモデル、
- (h) ネットワークカードモデル、ならびに、
- (i) 前記アプリケーションゲートウェイとリモートサーバとの間の接続の帯域幅、

についての値を含み、

前記指定された性能レベルは、

- (a) I/Oレイテンシ、
- (b) I/Oスループット、
- (c) 秒当たりのI/O動作、および、
- (d) I/Oタイムアウトの数、

のうち少なくとも1つの性能値によって表わされ、

機械訓練済み分析モデルを作成するように前記機械学習分析モデルを訓練することは、入力値の複数のセットを前記機械学習分析モデルにサブミットすることを含み、入力値の各セットは、構成パラメータ値、作業負荷値、および性能値を含み、前記性能値は、それぞれの前記構成パラメータ値で構成される、アプリケーションゲートウェイ上のそれぞれの前記作業負荷値によって表わされる量の作業を実行しながら経験的に観察されたものであり、

前記方法はさらに、

性能値の第3のセットによって表わされる指定された性能レベルで、指定された作業負荷値の第3のセットによって表わされる所望の量の作業を実行することができる推奨された構成についての第2の要求を受信するステップと、

前記作業負荷値の第3のセットおよび前記性能値の第3のセットを前記機械訓練済み分析モデルへの入力として提供するステップと、

前記作業負荷値の第3のセットおよび前記性能値の第3のセットに対応する構成パラメータ値の第2のセットを前記機械訓練済み分析モデルからの出力として受信するステップと、

前記構成パラメータ値の第2のセットを含む推奨された構成を含む前記第2の要求に応答するステップと、

作業負荷値の第4のセットによって表わされる作業の第3の量を、構成パラメータ値の第3のセットで構成されたアプリケーションゲートウェイによって実行するときに、達成可能な性能レベルの推定値についての第3の要求を受信するステップと、

前記構成パラメータ値の第3のセットおよび前記作業負荷値の第3のセットを前記機械訓練済み分析モデルへの入力として提供するステップと、

前記構成パラメータ値の第3のセットおよび前記作業負荷値の第4のセットに対応する性能値の第4のセットを前記機械訓練済み分析モデルからの出力として受信するステップと、

前記性能値の第4のセットを含む達成可能な性能レベルの推定値を含む前記第3の要求

10

20

30

40

50

に回答するステップと、

前記作業負荷値の第2のセットを含む前記作業負荷値の推奨されるセットを含む前記第1の要求に回答する前に、

前記システムによって実行されるべき作業の第5の量を表わす作業負荷値の第5のセットを、前記構成アドバイザーによって選択するステップと、

前記構成パラメータ値の第1のセットおよび前記作業負荷値の第5のセットを含む第3の入力を前記機械訓練済み分析モデルにサブミットするステップと、

前記構成パラメータ値の第1のセットおよび前記作業負荷値の第5のセットに対応する性能値の第5のセットを、前記機械訓練済み分析モデルからの出力として受信するステップと、

10

前記構成アドバイザーによって、前記性能値の第5のセットが前記指定された性能レベルを満たすかまたは超えることを判断するステップと、

前記作業負荷値の第2のセットによって表わされる作業の第2の量が、前記作業負荷値の第5のセットによって表わされる前記作業の第5の量よりも大きいと判断するステップとを含む、請求項1に記載の方法。

【請求項12】

請求項1～11のいずれか1項に記載の方法を1つ以上のプロセッサに実行させるためのプログラム。

【請求項13】

システムであって、

少なくとも1つのハードウェアプロセッサと、

命令を格納したメモリとを含み、

前記システムは以下のステップを実行するように構成され、前記以下のステップは、

構成パラメータ値と、作業負荷値と、性能レベルを表わす性能値との間の相関関係を認識するように機械学習分析モデルを訓練することによって、機械訓練済み分析モデルを作成するステップと、

構成パラメータ値の第1のセットで構成されるシステムによって、指定された性能レベルで実行され得る作業の量を表わす作業負荷値の推奨されるセットについての第1の要求を構成アドバイザーにおいて受信するステップと、

前記システムによって実行されるべき作業の第1の量を表わす作業負荷値の第1のセットを前記構成アドバイザーによって選択するステップと、

30

前記構成アドバイザーによって、前記構成パラメータ値の第1のセットおよび前記作業負荷値の第1のセットを含む第1の入力を前記機械訓練済み分析モデルにサブミットするステップと、

前記構成アドバイザーによって、前記構成パラメータ値の第1のセットおよび前記作業負荷値の第1のセットに対応する性能値の第1のセットを含む出力を前記機械訓練済み分析モデルから受信するステップと、

前記構成アドバイザーによって、前記性能値の第1のセットが前記指定された性能レベルを満たさないことを判断するステップと、

前記システムによって実行されるべき作業の第2の量を表わす作業負荷値の第2のセットを前記構成アドバイザーによって選択するステップと、

40

前記構成パラメータ値の第1のセットおよび前記作業負荷値の第2のセットを含む第2の入力を前記機械訓練済み分析モデルにサブミットするステップと、

前記構成パラメータ値の第1のセットおよび前記作業負荷値の第2のセットに対応する性能値の第2のセットを前記機械訓練済み分析モデルからの出力として受信するステップと、

前記構成アドバイザーによって、前記性能値の第2のセットが前記指定された性能レベルを満たすかまたは超えることを判断するステップと、

前記作業負荷値の第2のセットを含む前記作業負荷値の推奨されるセットを含む前記第1の要求に回答するステップとを含み、

50

前記作業負荷値の推奨されるセットは、少なくとも複数の  
 ( a ) 単位時間当たりの同時ユーザアクセスの数、  
 ( b ) ファイルシステム当たりのファイルの数、  
 ( c ) 読出し動作に対する書込み動作の割合、  
 ( d ) ファイルメタデータ読出し動作の数、  
 ( e ) ファイルメタデータ書込み動作の数、および  
 ( f ) ファイル付加動作の数、  
 を含む、システム。

【発明の詳細な説明】

【技術分野】

10

【0001】

本開示は構成管理に関する。具体的には、クラウドゲートウェイの性能調整用のモデルを訓練するために非線形ノンパラメトリック (nonlinear nonparametric : N L N P ) 回帰が用いられる。

【背景技術】

【0002】

背景

クラウドサービスのコンポーネントは、性能向上のために調整され得る多くの構成パラメータを有する。システムは、顧客の環境、クラウド環境、および、さまざまな環境間のネットワークング特徴に基づいて構成パラメータについての値を選択し得る。人はいくつかの構成パラメータを選択し得る。しかしながら、人にとって、性能を向上させるためにパラメータ値を構成するための無数の要因のすべてを考慮することは実用的ではない。

20

【0003】

いくつかのシステムは、自己調節動作を実行することにより、特定のリソースの消費が限度を超えるためにサービスが利用不可能になることを防止する。一例として、或るシステムはリソースの消費をリアルタイムで監視し得る。リソースの消費が閾値を超えるかまたは閾値に近づく場合、当該システムは補正動作を実行する。この補正動作は、リソースを追加する動作、または特定のプロセスへのリソース割当てを制限する動作を含み得る。

【0004】

実施形態は、添付の図面の図中において限定としてではなく例示のために示される。本開示において「ある」または「1つの」実施形態に言及する場合、これは必ずしも同じ実施形態を言及するものではなく、少なくとも1つを意味することに留意されたい。

30

【図面の簡単な説明】

【0005】

【図1】 1つ以上の実施形態に従った、リモートオブジェクトストレージサービスを用いるファイルサーバのコンポーネントを示すブロック図である。

【図2A】 1つ以上の実施形態に従った、推奨された作業負荷を決定するステップを示すフロー図である。

【図2B】 1つ以上の実施形態に従った、推奨された作業負荷を決定するステップを示すフロー図である。

40

【図3A】 1つ以上の実施形態に従った、機械学習分析モデルの訓練を示すブロック図である。

【図3B】 1つ以上の実施形態に従った、入力として受信された構成パラメータおよび所要の性能値のセットから作業負荷値のセットを推奨する構成アドバイザを示すブロック図である。

【図3C】 1つ以上の実施形態に従った、入力として受信された作業負荷値および所要の性能値のセットから構成パラメータのセットを推奨する構成アドバイザを示すブロック図である。

【図3D】 1つ以上の実施形態に従った、入力として受信された作業負荷値のセットおよび構成パラメータのセットから予想性能値を推定する構成アドバイザを示すブロック図で

50

ある。

【図 4】 1 つ以上の実施形態に従ったコンピュータシステムを示すブロック図である。

【発明を実施するための形態】

【0006】

図に示されるコンポーネントおよび動作は追加されてもよく、削除されてもよく、修正されてもよく、または組合わされてもよい。1 つのコンポーネント / 動作に関連付けて説明される機能は別のコンポーネントによって代わりに実現されてもよい。したがって、本明細書中に図示および / または説明される特定のコンポーネント / 動作は請求項のいずれかの範囲を限定するものとして解釈されるべきではない。

【0007】

詳細な説明

以下の記載では、完全な理解を提供するために多数の具体的な詳細が説明の目的で述べられている。1 つ以上の実施形態はこれらの具体的な詳細なしでも実施され得る。一実施形態において説明される特徴は別の実施形態において説明される特徴と組合わされてもよい。いくつかの例では、本発明を不必要に不明瞭にすることを避けるために、ブロック図を参照しながら周知の構造およびデバイスを説明する。

【0008】

1. 序論

クラウドサービスシステムについての構成パラメータ値を調整することは非常に困難な問題である。なぜなら、当該システムが複雑であるとともに、多数のパラメータが個々におよび組合わさって性能に影響を及ぼす作用を相関させることを必要とする可能性があるからである。

【0009】

クラウドベースのソリューションのサイジングおよび効率的構成は、顧客らにとって、自身の特定の作業負荷に関するストレージ I/O スループット（または性能）を向上させるために非常に重要である。顧客の作業負荷をサポートするように構成されていないシステムは、長いレイテンシ、低いスループット、I/O タイムアウト、サービスレベル合意（service level agreement: SLA）違反などの不満足な性能をもたらすおそれがある。加えて、システムが構成されて動作状態にされると、構成パラメータ値の変更に関連するコストが追加される可能性がある。たとえば、性能を向上させるためにプロセッサまたはメモリなどのより多くのハードウェアリソースを追加するには法外な費用がかかる可能性がある。性能向上のためにリソースを追加するコスト対利益比は継続的または定期的に評価される。

【0010】

本明細書中で説明する 1 つ以上の例は、クラウドベースのストレージソリューション、および、ストレージサービスに関連する構成パラメータを参照している。開示される技術は、システムが動作可能になる前に初期構成を推奨するために、動作中のシステムの性能問題を診断するために、または、性能調整のために動作システムの構成に対する変更を推奨するために、用いられ得る。これらの技術は他のクラウドベースのサービスまたはデータセンタベースのサービスにも適用され得るものであり、構成パラメータのタイプはさまざまであり得るとともに他の対応するサービスタイプに基づき得る。

【0011】

2. 概要

実施形態は、性能が 1 つ以上の性能基準を満たすかまたは上回ることを確実にした状態でリソースのセットについての特定のレベルの作業負荷を推奨する。多変量状態推定技術（multivariate state estimation technique: MSET）モデルは、観察された作業負荷または推定された作業負荷を処理する際にシステムの性能に影響を及ぼす多くの構成パラメータ間の相関関係のパターンを学習するように訓練される。モデルの訓練は、さまざまな構成および作業負荷で実行されるかまたは匿名化された履歴データのリポジトリを用いて実行された実験から蓄積されたデータを用いてオフラインで実行されてもよい。当該

10

20

30

40

50

モデルは、訓練されると、構成されているシステムの性能要件およびリソース制約に基づいて、当該システムによって実行され得る作業負荷を推奨するために用いられ得る。代替的には、当該モデルはまた、指定された目標作業負荷を処理するのに必要なシステム構成を推奨するために用いられてもよい。

【0012】

本明細書および/または特許請求の範囲において記載されるいくつかの実施形態はこの概要の項目に含まれない可能性もある。

【0013】

3. 構成アドバイザーアーキテクチャ

図1は、1つ以上の実施形態に従った、リモートオブジェクトストレージサービスを用いるファイルサーバのコンポーネントを示すブロック図である。コンポーネントは、追加されてもよく、削除されてもよく、修正されてもよく、または組合わされてもよい。1つのコンポーネントに関連付けて説明される機能が別のコンポーネントによって代わりに実現されてもよい。したがって、本明細書中において例示および/または説明される特定のコンポーネントは請求項のいずれかの範囲を限定するものとして解釈されるべきではない。

【0014】

図1のシステムは、ネットワーク115を介してリモートデータセンタ160内のストレージサービス190にオブジェクトを要求するファイルサーバ110を含むシステム100を示す。

【0015】

ファイルサーバ110は、新しいファイルを作成するか、または既存のファイルを検索するか、修正するかもしくは削除するなどのファイルを管理するための要求を受信する。ある実施形態では、ファイルサーバ110はNFSサーバである。ファイルサーバ110はハードウェアとソフトウェアとの組合わせである。たとえば、ファイルサーバは、ストレージソフトウェアアプライアンス120およびオペレーティングシステム150を実行させる。

【0016】

ストレージソフトウェアアプライアンス120は、ファイルシステム130およびファイルゲートウェイ140を含む。ファイルシステム130は、ファイルサーバ上にローカルに格納されるとともに当該ファイルサーバによって管理されるファイルの一部を格納する。本明細書で用いられる語「ローカル」および「リモート」は、仮想的な近接性および物理的な近接性の両方を指し得る。すなわち、自身の構内でファイルサーバを実行する企業はその企業にとって物理的にローカルであり得る。企業のファイルサーバが企業の仮想プライベートネットワーク上にある物理的に離れたデータセンタにおいて動作している場合、ファイルサーバは企業にとって仮想的にローカルである。同様に、リモートデータセンタ160は、ファイルサーバ110から少なくとも仮想的にリモートであるとともに、場合によってはファイルサーバ110から物理的にもリモートである。リモートデータセンタ160を表す破線ボックス内に示すコンポーネントは、企業の仮想ネットワーク上になく(すなわち、仮想的にリモートであり)、企業のファイルサーバの位置から物理的に離れていてもよい。ネットワーク115はローカルシステムとリモートシステムとを接続する。ローカルシステムおよびリモートシステムが地理的に遠く離れている場合、ネットワークはレイテンシを導入し得る。ネットワークの速度は、リモートストレージサービスとの対話を含むファイル要求に対する応答の全体的な性能に影響を及ぼし得る。

【0017】

ファイルキャッシュ136に格納されたファイルに対して動作するようにとの要求が受信されると、ファイルがファイルキャッシュ136から取出され、取出されたファイルに対して当該要求が実行される。しかしながら、ファイルキャッシュ136内に存在しないファイルが要求されると、ファイルゲートウェイ140は、ストレージサービス190にストレージオブジェクトを要求し、取出されたストレージオブジェクトからファイルを作成し、作成されたファイルをファイルキャッシュ136内に格納し得る。ストレージサー

10

20

30

40

50

ビス190は、オブジェクトストア195内のオブジェクトに対するオブジェクトインターフェイスであってファイルサーバによって提供されるファイルインターフェイスとは異なるオブジェクトインターフェイスを提供し得るとともに、ファイルゲートウェイ140は、ファイルインターフェイスとオブジェクトインターフェイスとの間のプロトコル変換を実行し得る。ある実施形態では、新しいファイルの作成はファイルサーバ上でローカルに実行され得るとともに、新しいファイルはファイルキャッシュ内に格納され得る。ファイルゲートウェイは、次いで、ストレージサービスを呼出して、新しいファイルのための1つ以上のストレージオブジェクトを格納し得る。同様に、キャッシュ内のファイルの修正は、ストレージソフトウェアアプライアンスによってローカルに実行され得る。次いで、当該修正がストレージオブジェクトに対する変更としてファイルゲートウェイによってストレージサービスに送信されて、ファイルの状態を同期させ得る。

10

#### 【0018】

ストレージソフトウェアアプライアンス120の構成は本開示の主題である。構成推奨は、予想される作業負荷を実行して、指定された性能目標を満たすようにアプライアンスを構成するために行なわれ得る。指定された性能目標はサービス品質契約に基づいて必要とされてもよく、または、性能目標は想起的であってもよい。

#### 【0019】

本明細書で用いられる「作業負荷」は、システムによって実行されるとシステムに対する負荷を生成するタスクのセットである。たとえば、ファイルサーバの場合、タスクのセットは、ファイルの作成、修正または削除などのファイルに対して実行すべき動作のセットである。作業負荷値は、対応する作業負荷によって表わされる作業の量を定量化したものである。作業負荷値は、本明細書では、作業の量を表わす作業負荷値の省略表現として用いられることもある。

20

#### 【0020】

性能目標は、サービス品質(Quality of Service: QoS)レベルを表わし得る。QoSレベルはタスクまたはサブタスクごとに定義され得る。一例として、QoSレベルは、書込み動作を完了するための時間の量に対応し得る。QoSレベルは、出荷、請求書発行、および在庫更新のためのコマンドをサブミットすることを含む、販売注文の実行を完了するための時間の量に対応し得る。QoSレベルは、単一の動作に対するディスクレイテンシまたはネットワークレイテンシなどのより粒度の細かい動作を指し得る。QoSレベルは、1秒あたりに転送される数メガバイト(たとえば、100MB)の平均スループットを指し得る。QoSレベルは、所与の期間内で実行される動作の数を指し得る。

30

#### 【0021】

構成パラメータ133は、ファイルゲートウェイ140と、ファイルゲートウェイが動作する環境とについての構成パラメータ値を格納する。たとえば、構成パラメータ値は、ファイルサーバが動作するハードウェアを特徴付け得る。「ゲートウェイを構成する」という語句は、本明細書中では、ゲートウェイが動作する環境を構成することと、ゲートウェイ自体を構成することとを意味するために用いられている。構成パラメータの例は、データキャッシュサイズ、メタデータサイズ、ログファイルサイズ、およびスレッドの数などの動作パラメータと、ファイルサーバアーキテクチャ(コアの数およびコア速度、メモリ、ディスクコントローラ、ネットワークカードなどと表わされる)およびネットワーク115の帯域幅などの非動作パラメータとを含み得る。

40

#### 【0022】

作業負荷値は、ファイルゲートウェイによって実行される作業の量を示すとともに、ファイルサーバ要求を満たすことによってシステムに課されるであろう計算、ネットワーク送信、メモリトランザクション、およびI/O負荷の量を示す。作業負荷値は、ゲートウェイのファイル要求の量およびそれらの組合せ、たとえば、同時ユーザアクセスの数、単位時間あたりにアクセスされるファイルの数、ファイルの平均サイズ、ディレクトリ階層構造、および単位時間あたりの書込み/読出し/付加といったI/O動作の数などを特徴付け得る。ファイルサーバが動作可能であり、ファイルサービス要求を受信すると、ファイ

50

ルサーバデータが経験的に被る作業負荷についての値は、性能監視を行なうことによって取得され得る。ある実施形態では、遠隔測定を用いて、ファイルゲートウェイおよび下層のオペレーティングシステムから作業負荷値に関するデータを収集し得る。すなわち、運用可能なファイルゲートウェイを再構成するための推定値としては、経験的に観察された作業負荷値を用い得る。

【0023】

構成パラメータおよび作業負荷に加えて、構成パラメータのセット下で作業負荷を実行した結果を表わす性能データが収集され得る。例示的な性能データは、ファイルサーバにおいて受信されるファイル要求に応答してファイルを検索または格納するための平均時間、(レイテンシ)平均CPU使用率、およびメモリ使用率を含み得る。

10

【0024】

オペレーティングシステム150は、ファイルサーバ110のハードウェア上で動作するオペレーティングシステムである。Linux(登録商標)およびWindows(登録商標)はオペレーティングシステムの例である。オペレーティングシステムは、CPU使用率などの性能情報を取得するための遠隔測定を実行する。加えて、オペレーティングシステムを構成するために、キャッシュサイズなどの構成パラメータ133に格納された値のうちいくつかがいられ得る。

【0025】

ストレージサービス190はオブジェクトストア195内のストレージオブジェクトを管理する。オブジェクトストア195内の1つ以上のストレージオブジェクトは顧客のローカル環境内のファイルに対応する。

20

【0026】

機械学習分析モデル172は、構成パラメータ値と作業負荷値と性能値との間の相関関係を判断する数学的回帰モデルである。当該モデルは、テストシステム上で相関することが経験的に観察された値のセットを用いて訓練され得る。モデルを訓練するために、複数回のテストを実行するための構成値、作業負荷値および性能値のいくつかのセットがいられ得る。

【0027】

訓練データは、さまざまな構成下のテスト環境において、テストシステムに負荷をかけるファイル要求のセットを実行することによって取得され得る。テストシステムがファイル要求のセットを処理する各テストごとに、(a)テストシステムによって実行される作業の量を定量化する作業負荷値と、(b)訓練システムを構成するために用いられる構成パラメータ値と、(c)テストシステムの観察された性能を表わす値と、を含む値が記録される。各パラメータは、製造システムを構成するために典型的に用いられる妥当な値の範囲に関連付けられ得る。性能テストは、値の妥当な範囲から選択されるパラメータ値の組み合わせを用いて構築され得る。

30

【0028】

性能調整のための別の技術は、システムが実際の作業を実行している間に、特定のリソースのリソース消費を監視することを含む。リソースの使用率が閾値を超えると、そのタイプの追加のリソースが割当てられ得る。当該リソースについての構成パラメータの値は、リソースの新しい量を反映するように変更される必要があるかもしれない。しかしながら、この技術は、構成パラメータを互いとは無関係に監視および調整する。たとえば、構成パラメータ間の相互作用がこのシステムに認識されていないので、仮想メモリの構成量を変更することなく、当該システムに物理メモリが追加されてもよい。別の例として、メモリ帯域幅およびI/O帯域幅に関連付けられたパラメータ値に影響を及ぼす可能性のある専用CPUコアまたは「CPU+GPU形状」を当該構成に対して追加することもあるだろう。「形状」とは、高性能コンピューティングおよび他の低レイテンシストリーミングの分析作業負荷に適した新しいGPUを備えた従来のCPUについての認可済みかつ事前テスト済みの基準構成である。1つの構成パラメータの値を変更すると、別の構成変数値の値との一貫性が取れなくなる可能性がある。分析モデルからの推奨は、相互に一貫性が

40

50

あるとともにより優れた全体的なシステム性能調整を可能にし得る構成パラメータ値のセットを含み得る。一例として、利用可能なローカルストレージを増やすことなく、または、他の場合には、キャッシュされたデータの増加に対応する容量を確保することなく、複数のゲートウェイファイルシステムをホストしているサーバ上のローカルファイルキャッシュ閾値を増やすこともあるだろう。別個に見た場合、サーバは、より大規模なキャッシュを使用可能にするのに十分に利用可能なローカル容量を有するよう見える可能性があるが、ローカルストレージは共有リソースであるため、リソース競合を引起こすことによって他のゲートウェイファイルシステムに悪影響を及ぼす可能性がある。

#### 【 0 0 2 9 】

訓練されたモデルは、本明細書では、機械訓練済み分析モデル 1 7 5 (「モデル 1 7 5」) と称され得る。モデル 1 7 5 は、値の入力ベクトルを受信し得るとともに、当該入力ベクトルに最も類似していると推定される値であって当該値が関連していることが分かっている値のベクトルを返し得る。

10

#### 【 0 0 3 0 】

構成ポータル 1 8 0 は、訓練データを機械学習分析モデル 1 7 2 に送信するために用いられ得る。ある実施形態では、構成ポータル 1 8 0 は、構成アドバイザー 1 7 0 に対して構成についての推奨を要求するために用いられ得る。ある実施形態では、ストレージサービスに関連付けられた専門のサービス、販売またはサポートの担当者は、リモートデータセンタ 1 6 0 内の構成ポータルを用いて、顧客に推奨すべき構成パラメータのセットを決定し得る。別の実施形態では、構成アドバイザーは、構成推奨を要求するために顧客が直接使用可能でありネットワーク 1 1 5 を介してアクセス可能なサービスとして提供され得る。

20

#### 【 0 0 3 1 】

構成アドバイザー 1 7 0 は、入力値と、所望の出力値に関する情報とを受信し得る。たとえば、構成アドバイザーは、作業負荷値を含む要求と、性能値と、要求に応じて構成パラメータが要求される指示とを受信し得る。構成アドバイザーは、モデル 1 7 5 によって提供されるインターフェイスに応じて要求内にデータをパッケージ化し得る。構成アドバイザーは、すべての要求に関して明確に供給される必要がない値をモデル 1 7 5 に供給するために用いられ得るとともにリモートデータセンタに登録および格納された顧客データにアクセスでき得る。たとえば、リモートデータセンタは顧客 S L A にアクセスできる可能性もある。付加的または代替的には、構成アドバイザーは、1 つ以上のパラメータについての「妥当な」値の範囲を格納してもよい。たとえば、キャッシュサイズについての妥当な値の範囲は、顧客に配備した中で用いられることが分かっているキャッシュサイズを反映して、2 G B ~ 1 0 G B であり得る。供給されない要求の値、既知でない要求の値および/もしくは一定ではない要求の値、または妥当な範囲外の値を有する要求の値は、構成アドバイザーによって、パラメータ値についての妥当な範囲内の値と置換えられてもよい。ある実施形態では、妥当な範囲内の置換値がランダムに選択され得る。別の実施形態では、妥当な範囲内にある事前構成されたデフォルト値を用いて、欠落している値または範囲外の値を置換または供給してもよい。

30

#### 【 0 0 3 2 】

構成アドバイザーは、モデルから相関値のセットを受信すると、推奨についての要求に応答する前に出力を再パッケージ化し得る。一実施形態では、構成アドバイザーは、モデルから受信した値のベクトルを送信することによって要求に応答し得る。別の実施形態では、構成アドバイザーは、推奨が要求された値のみを返してもよい。

40

#### 【 0 0 3 3 】

#### 4 . 構成アドバイザーの動作

図 2 A および図 2 B は共に、1 つ以上の実施形態に従った、推奨された作業負荷を決定するステップを示すフロー図である。図 2 A および図 2 B を参照して本明細書で説明される動作は省略されてもよく、修正されてもよく、または並べ替えられてもよい。図 2 A および図 2 B に示されていない動作は、図示された動作の代わりに、または図示された動作に加えて実行されてもよい。したがって、本明細書で説明される動作は請求項のいずれか

50

の範囲を限定するものと解釈されるべきではない。

【0034】

ここで図2Aを参照すると、機械学習分析モデル172は、関連させるべきと経験的に判断される受信した値のセットに対して回帰を実行することによって訓練される(動作205)。訓練データは、QoSレベルを表わし得る構成パラメータについての値、作業負荷値、および性能値を含み得る。上述したように、訓練データは、結果として得られるシステム性能を経験的にテストして観察することによって収集されてもよい。システムは、当該システムが構成パラメータのセットで構成されている間に作業負荷を実行し、さらに、結果として得られる性能が観察される。構成、作業負荷、および性能に関する値について観察された各々の組み合わせは単一の一貫したシステム状態を表わし得る。多くのシステム状態がモデル訓練プロセスに入力される。モデルは、一貫したシステム状態のセットを表わし得るモデルを作成するためにベクトル類似性評価技術を採用し得る。特定の機械学習解析モデルが訓練されて用いられる実施形態のさらなる詳細について、多変量状態推定モデル(MULTIVARIATE STATE ESTIMATION (MSET) MODEL)と題された以下の段落を参照されたい。

10

【0035】

別の実施形態では、機械学習モデルを訓練することは、パラメータ値間の関係を表わす多項式を構築することを含み得る。多項式の各項は、訓練データセットにおいて表わされる特定のパラメータに対応し得る。係数の項として用いられる変数は、対応するパラメータの値を表わし得る。モデルにデータを適用することは、各パラメータの値をその対応する項変数に代入することと、未知のパラメータ値を解くこととを含み得る。

20

【0036】

機械訓練済み分析モデル175は、訓練されると、作業負荷値および性能値を入力構成として取る。しかしながら、入力の2つのセットが既知である可能性があり、第3のセットが推定される可能性もある。推定値の第3のセットに関して推奨値が出力され得る。ある実施形態では、入力の第3のセットはヌル値を有し得る。

【0037】

ある実施形態では、構成アドバイザー170は、作業に関する推定された作業負荷値についての要求を受信する(動作210)。当該要求は構成パラメータ値のセットを指定する。

【0038】

サービスプロバイダの専門家が構成推奨を要求するために構成ポータルと対話する実施形態では、当該要求は、顧客が初めてファイルゲートウェイをインストールするのに備えてサブミットされ得る。推奨された値は新しいインストールを構成するために用いられ得る。加えて、専門家は、不良な性能についての苦情を受信したことに応答して要求をサブミットしてもよい。

30

【0039】

構成アドバイザーがネットワークを介してアクセス可能なサービスとして提供される実施形態では、構成推奨についての要求は、システムリソースの変更を検出すると、または性能が低下したとの判断に応答して、動的に生成され得る。

【0040】

この例示的なフローでは、作業負荷値はモデルからの出力として要求される。構成アドバイザーは作業負荷値の初期候補セットを選択する(動作220)。構成アドバイザーは、モデルへの入力として提供するために妥当な値の範囲から候補作業負荷値を選択し得る。候補作業負荷値の妥当な範囲は、上述したように訓練データを収集するために用いられる値に基づいていてもよい。

40

【0041】

推奨要求とともに受信された構成パラメータおよび作業負荷値の初期候補セットはモデルにサブミットされる。構成アドバイザーは、性能値のセットをモデルからの出力として受信する(動作230)。性能値のセットは、作業負荷値の候補セットに従ってロードされるとともに入力構成パラメータ値に従って構成されるシステムについての推定性能を表わ

50

す。

【 0 0 4 2 】

先に説明したように、モデルは、既知の構成および作業負荷で実行することから収集されるデータを用いて訓練され得る。訓練されたモデルによって所望の構成が表わされない場合、シミュレーションを実行することによってデータが収集され得る。一例として、システムは構成パラメータに従って構成され得る。現実世界の環境を模倣する動作が実行され得る。これらの動作は、特定の日の時間および/または特定のクライアントシステムに典型的であろう計算の作業負荷、読出し、書込み、およびネットワーク I / O を含み得る。現実世界の環境において動作を実際に完了させるのにかかる時間の量は推定された性能値として記録され得る。

10

【 0 0 4 3 】

構成アドバイザーは、推定された性能値が所要の性能値を満たすかどうかを判断する（動作 2 4 0）。推定された性能値が所要の性能値を満たすかまたは超える場合、モデルにサブミットされた作業負荷値の候補セットが推奨された作業負荷値として返され得る（動作 2 4 5）。ある実施形態では、返された推奨は構成ポータルのディスプレイ上に提示され得る。別の実施形態では、返された推奨は、要求を生成したシステムに送信されてもよい。

【 0 0 4 4 】

推定された性能値が所要の性能値を満たさない場合、作業負荷値の新しい候補セットが妥当な値の範囲から作業負荷値の現在の候補セットとして選択される（図 2 B の動作 2 5 0）。ある実施形態では、現在の候補作業負荷値がランダムに選択され得る。別の実施形態では、性能値が作業負荷値の以前の候補セットに関する所要の性能値を満たさなかった場合、作業負荷値の以前の候補セットよりも少ない作業を表わす作業負荷値の現在の候補セットが選択され得る。

20

【 0 0 4 5 】

作業負荷値の現在の候補セットおよび当初に提供された構成パラメータ値がモデルに送信される（動作 2 6 0）。更新された性能値のセットがモデルからの出力として取得される（動作 2 7 0）。更新された性能値のセットが所要の性能値と比較される（動作 2 8 0）。たとえば、観察されたレイテンシが 7 5 m s であり、所要の性能目標が 5 0 m s のレイテンシである場合、推定された性能値は所要の値を満たさない。更新された性能値のセットが所要の性能値を満たすかまたは超える場合、構成アドバイザーは、作業負荷値の現在の候補セットを推奨された作業負荷値として返し得る（動作 2 9 0）。再び、更新された性能値のセットが所要の性能値を満たさない場合、作業負荷値の新しい候補セットが性能値の現在の候補セットとして選択されるプロセスが繰り返され得る。モデルが妥当な範囲内の作業負荷値の任意の候補セットに関する許容可能な性能値を返さない場合、フローはエラーを出して停止し得る。

30

【 0 0 4 6 】

別の実施形態では、構成アドバイザーは、所要の性能値を満たすかまたは超える性能値と関連する推奨された作業負荷値の第 1 のセットを返すのではなく、モデル 1 7 5 に対して、作業負荷値の多数の候補セットについての性能推定値を要求し、所要の性能値と一致する推奨された最大作業負荷値を返してもよい。

40

【 0 0 4 7 】

図 3 A は、1 つ以上の実施形態に従った、機械学習分析モデルの訓練を示すブロック図である。機械学習分析モデル 1 7 2 は、データの複数のセットを受信することによって訓練される。この場合、データの各セット内の値はテスト環境 / 実験環境内で関連させたものであり、データの各セットは、構成パラメータについての値、作業負荷値、および観察された性能値を含む。

【 0 0 4 8 】

図 3 B は、1 つ以上の実施形態に従った、入力として受信された構成パラメータ値および所要の性能値のセットと関連する作業負荷値のセットを推奨する構成アドバイザー 1 7 0 を示すブロック図である。構成アドバイザーは、要求とともに受信された所要の性能値およ

50

び構成を含む入力値のベクトルと、推定された作業負荷値のセットとをモデル 175 に提供し得る。モデル 175 は、入力ベクトルにおいて指定された構成値および性能値と相関する改訂済み作業負荷値のセットを含む出力ベクトルを生成し得る。構成アドバイザー 170 は、単に、モデル 175 によって指定された作業負荷値を要求エンティティに返すだけでもよい。

【0049】

図 3C は、1 つ以上の実施形態に従った、入力として受信された作業負荷値および性能値のセットから構成パラメータのセットを推奨する構成アドバイザーを示すブロック図である。構成アドバイザーは、要求とともに受信された作業負荷値および性能値のセットを含む入力値のベクトルと、推定された構成値のセットとをモデル 175 に提供し得る。当該モデルは、入力ベクトルにおいて供給された作業負荷値および性能値と相関する改訂済み構成値のセットを含む出力ベクトルを生成し得る。

10

【0050】

図 3D は、1 つ以上の実施形態に従った、入力として受信された作業負荷値のセットおよび構成パラメータのセットを想定して予想される性能値を推定する構成アドバイザーを示すブロック図である。構成アドバイザーは、要求とともに受信された作業負荷値および構成パラメータ値のセットを含む入力値のベクトルと推定された性能値のセットとをモデル 175 に提供し得る。当該モデルは、入力ベクトルにおいて指定された作業負荷値および構成パラメータ値と相関する改訂済み性能値のセットを含む出力ベクトルを生成し得る。

20

【0051】

#### 5. 多変量状態推定 (MSET) モデル

MSET モデルは、非線形ノンパラメトリック (NLNP) 回帰と総称されるパターン認識数学的処理のクラスから予後アルゴリズムを用いて相関パラメータのセットを推定する。MSET モデルは、ベクトル類似性評価技術を採用し得る機械学習モデルである。MSET モデルは、 $n$  個の観察ベクトルを格納する  $m \times n$  行列を含み得る。この場合、各ベクトルは、キャッシュサイズ、平均ファイルサイズおよびスループットなどの  $m$  個のパラメータの各々についての値を含む。 $n$  列の各行は、システムの既知の状態を表わすベクトルである。MSET モデルの訓練中、テストシステムは、構成値および作業負荷値の多くのさまざまな組合わせでテストされる。各テストは、構成パラメータ値の特定のセットで構成されるテストシステムを用いて実行されるとともに、作業負荷値のセットによって特徴付けられる特定の作業負荷を処理する。結果として得られたシステムの性能が観察されて記録される。このような各々のテストごとに、特定の構成と、作業負荷と、時間に対して結果として得られた性能値とを含むベクトルの集合が訓練データとして MSET モデルに入力される。

30

【0052】

訓練されたモデルは、入力ベクトルおよび既知の状態の行列を用いてパラメータ値の出力ベクトルを生成する。当該モデルは、行列によって表わされる既知の状態に対する入力ベクトルの類似性を判断し、これらの類似性値を用いて出力ベクトルについての新しい推定値を計算する。モデルがどのように機能するかに関するより詳細な説明については、K. C. Gross、R. M. Singer、S. Wegerich および J. Mott による「多変量状態推定技術 (MSET) ベースの監視システム (Multivariate State Estimation Technique (MSET) Based Surveillance System)」(事例 ID ANL-IN-94-135: 米国特許第 5,764,509 号 (1998 年 6 月 2 日)) を参照されたい。この特許および他の刊行物に記載されているように、MSET モデルは異常の検出に用いられてきた。構成パラメータ値を推定するために MSET モデルを用いることは新しい応用例である。

40

【0053】

#### 6. 実施例

このセクションでは、過去に回答することができなかったが当該モデルが回答することができるいくつかの例示的な質問について検討する。或る顧客が、10000 個のファイ

50

ルにアクセスしている2つのCPUを実行させる場合がある。当該顧客は、ファイルの数が20000にまで増加する場合、何個のCPUの追加が必要になり得るか、同等の性能を維持するためにどれだけのキャッシュおよびローカルディスク空間の割当てが必要になり得るかを知ることがを所望する場合がある。特定の環境における特定の作業負荷のためのファイルの数とCPUとの間の関係は分かっていない。

#### 【0054】

モデルを用いることにより、数学的に決定された関係を用いて特定の顧客環境を表わす相関値を推定することが可能となり、経験的にテストするために顧客固有の値を用いて実験を行なう必要がなくなる。実験室環境における経験的テストは、全てではないが、いくつかの組み合わせをテストし得る。加えて、研究室でのテストは特定の作業負荷で実行されなくてもよい。訓練データからの推論を用いて、経験的にテストされなかったパラメータ値の組み合わせについての結果を予測することができる。

10

#### 【0055】

##### 7. コンピュータネットワークおよびクラウドネットワーク

1つ以上の実施形態では、コンピュータネットワークは、1組のノード間の接続性を提供する。これらのノードは、互いにローカルおよび/またはリモートであってもよい。これらのノードは1セットのリンクによって接続される。リンクの例は、同軸ケーブル、無被覆ツイストケーブル、銅ケーブル、光ファイバ、および仮想リンクを含む。

#### 【0056】

ノードのサブセットはコンピュータネットワークを実装する。そのようなノードの例は、スイッチ、ルータ、ファイアウォール、およびネットワークアドレストランスレータ (network address translator: NAT) を含む。ノードの他のサブセットはコンピュータネットワークを用いる。そのようなノード (「ホスト」とも呼ばれる) は、クライアントプロセスおよび/またはサーバプロセスを実行し得る。クライアントプロセスは、コンピューティングサービス (たとえば、特定のアプリケーションの実行、および/または特定の量のデータの格納など) についての要求を行なう。サーバプロセスは、要求されたサービスを実行することによって、および/または対応するデータを返すことによって応答する。

20

#### 【0057】

コンピュータネットワークは、物理的リンクによって接続された物理的ノードを含む物理的ネットワークであってもよい。物理的ノードは任意のデジタルデバイスである。物理的ノードは、ハードウェアスイッチ、ハードウェアルータ、ハードウェアファイアウォール、およびハードウェアNATなどの特定機能のハードウェアデバイスであってもよい。付加的または代替的には、物理的ノードは、それぞれの機能を実行するさまざまな仮想マシンおよび/またはアプリケーションを実行するように構成された汎用マシンであってもよい。物理的リンクとは、2つ以上の物理的ノードを接続する物理的媒体である。リンクの例は、同軸ケーブル、無被覆ツイストケーブル、銅ケーブル、および光ファイバを含む。

30

#### 【0058】

コンピュータネットワークはオーバーレイネットワークであってもよい。オーバーレイネットワークとは、別のネットワーク (物理的ネットワークなど) の上に実装された論理的ネットワークである。オーバーレイネットワークにおける各ノードは、下位ネットワークにおけるそれぞれのノードに対応する。よって、オーバーレイネットワークにおける各ノードは、(オーバーレイノードをアドレス指定するための) オーバーレイアドレスと、(オーバーレイノードを実現するアンダーレイノードをアドレス指定するための) アンダーレイアドレスとの双方に関連付けられる。オーバーレイノードは、デジタルデバイスおよび/またはソフトウェアプロセス (仮想マシン、アプリケーションインスタンス、またはスレッドなど) であってもよい。オーバーレイノード同士を接続するリンクは、下位ネットワークを通るトンネルとして実現される。トンネルの両端のオーバーレイノードは、当該オーバーレイノード同士の間の下位マルチホップ経路を、単一の論理的リンクとして扱う。カプセル化および非カプセル化によってトンネリングが実行される。

40

50

## 【 0 0 5 9 】

ある実施形態では、クライアントはコンピュータネットワークに対してローカルおよび/またはリモートであってもよい。クライアントは、コンピュータネットワークに、プライベートネットワークまたはインターネットなどの他のコンピュータネットワークを介してアクセスしてもよい。クライアントは、ハイパーテキスト・トランスファ・プロトコル (Hypertext Transfer Protocol : H T T P ) などの通信プロトコルを用いてコンピュータネットワークに要求を通信してもよい。要求は、クライアントインターフェイス (ウェブブラウザなど)、プログラムインターフェイス、またはアプリケーションプログラミングインターフェイス (application programming interface : A P I ) などのインターフェイスを通して通信される。

10

## 【 0 0 6 0 】

ある実施形態では、コンピュータネットワークは、クライアントとネットワークリソースとの間に接続性を提供する。ネットワークリソースは、サーバプロセスを実行するように構成されたハードウェアおよび/またはソフトウェアを含む。ネットワークリソースの例は、プロセッサ、データ記憶装置、仮想マシン、コンテナ、および/またはソフトウェアアプリケーションを含む。ネットワークリソースは複数のクライアント間で共有される。クライアントは互いから独立して、コンピュータネットワークにコンピューティングサービスを要求する。ネットワークリソースは、オンデマンドベースで要求および/またはクライアントに動的に割り当てられる。各要求および/またはクライアントに割り当てられたネットワークリソースは、たとえば、( a ) 特定のクライアントによって要求されたコンピューティングサービス、( b ) 特定のテナントによって要求された、集約されたコンピューティングサービス、および/または、( c ) コンピュータネットワークに要求された、集約されたコンピューティングサービス、に基づいて拡大または縮小されてもよい。そのようなコンピュータネットワークは「クラウドネットワーク」とも称され得る。

20

## 【 0 0 6 1 】

ある実施形態では、サービスプロバイダは、クラウドネットワークを1人以上のエンドユーザに提供する。サービスとしてのソフトウェア (Software-as-a-Service : S a a S )、サービスとしてのプラットフォーム (Platform-as-a-Service : P a a S )、およびサービスとしてのインフラストラクチャ (Infrastructure-as-a-Service : I a a S ) を含むもののそれらに限定されない、さまざまなサービスモデルがクラウドネットワークによって実現され得る。S a a S では、サービスプロバイダは、ネットワークリソース上で実行されるサービスプロバイダのアプリケーションを使用する能力をエンドユーザに提供する。P a a S では、サービスプロバイダは、ネットワークリソース上にカスタムアプリケーションを配備する能力をエンドユーザに提供する。カスタムアプリケーションは、サービスプロバイダによってサポートされるプログラミング言語、ライブラリ、サービス、およびツールを用いて作成され得る。I a a S では、サービスプロバイダは、ネットワークリソースによって提供される処理、記憶、ネットワーク、および他の基本的コンピューティングリソースをプロビジョニングする能力をエンドユーザに提供する。オペレーティングシステムを含む任意のアプリケーションがネットワークリソース上に配備され得る。

30

## 【 0 0 6 2 】

ある実施形態では、プライベートクラウド、パブリッククラウド、およびハイブリッドクラウドを含むもののそれらに限定されない、さまざまな配備モデルがコンピュータネットワークによって実現され得る。プライベートクラウドでは、ネットワークリソースは、1つ以上のエンティティからなる特定のグループによる独占的使用のためにプロビジョニングされる (ここで用いられる「エンティティ」という語は、企業、組織、人または他のエンティティを指す)。ネットワークリソースは、エンティティからなる特定のグループの構内に対してローカルおよび/またはリモートであってもよい。パブリッククラウドでは、クラウドリソースは、互いに独立した複数のエンティティ (「テナント」または「顧客」とも称される) のためにプロビジョニングされる。コンピュータネットワークおよびそのネットワークリソースは、異なるテナントに対応するクライアントによってアクセス

40

50

される。そのようなコンピュータネットワークは「マルチテナントコンピュータネットワーク」と称されてもよい。いくつかのテナントは、異なる時間に、および/または同じ時間に、同じ特定のネットワークリソースを用いてもよい。ネットワークリソースは、テナントの構内に対してローカルおよび/またはリモートであってもよい。ハイブリッドクラウドでは、コンピュータネットワークはプライベートクラウドおよびパブリッククラウドを含む。プライベートクラウドとパブリッククラウドとの間のインターフェイスが、データおよびアプリケーションの可搬性を可能にする。プライベートクラウドに格納されたデータと、パブリッククラウドに格納されたデータとは、インターフェイスを通して交換されてもよい。プライベートクラウドで実現されるアプリケーションと、パブリッククラウドで実現されるアプリケーションとは、互いに対する従属性を有していてもよい。プライベートクラウドにおけるアプリケーションからパブリッククラウドにおけるアプリケーションへの（逆もまた同様）呼出しがインターフェイスを通して実行されてもよい。

10

**【0063】**

ある実施形態では、マルチテナントコンピュータネットワークのテナントは互いに独立している。たとえば、1つのテナントのビジネスまたは事業は、別のテナントのビジネスまたは事業とは別のものであり得る。さまざまなテナントが、コンピュータネットワークに対してさまざまなネットワーク要件を要求し得る。ネットワーク要件の例は、処理速度、データ記憶量、セキュリティ要件、性能要件、スループット要件、レイテンシ要件、復元力要件、サービス品質（Quality of Service：QoS）要件、テナント分離、および/または一貫性を含む。同じコンピュータネットワークが、さまざまなテナントによって要求されるさまざまなネットワーク要件を実現することが必要になる可能性もある。

20

**【0064】**

1つ以上の実施形態では、マルチテナントコンピュータネットワークにおいて、さまざまなテナントのアプリケーションおよび/またはデータが互いに共有されないことを確実にするために、テナント分離が実現される。さまざまなテナント分離アプローチが用いられてもよい。

**【0065】**

ある実施形態では、各テナントは、テナントIDに関連付けられる。マルチテナントコンピュータネットワークの各ネットワークリソースが、テナントIDでラベル付けされる。テナントと特定のネットワークリソースとが同じテナントIDに関連付けられている場合のみ、テナントは特定のネットワークリソースへのアクセスが許可される。

30

**【0066】**

ある実施形態では、各テナントはテナントIDに関連付けられている。コンピュータネットワークによって実現される各アプリケーションがテナントIDでラベル付けされている。付加的または代替的には、コンピュータネットワークによって格納された各データ構造および/またはデータセットがテナントIDでラベル付けされている。テナントと特定のアプリケーション、データ構造、および/またはデータセットとが同じテナントIDに関連付けられている場合のみ、テナントは、特定のアプリケーション、データ構造、および/またはデータセットへのアクセスが許可される。

**【0067】**

一例として、マルチテナントコンピュータネットワークによって実現される各データベースはテナントIDでラベル付けされてもよい。対応するテナントIDに関連付けられたテナントのみが、特定のデータベースのデータにアクセスしてもよい。別の例として、マルチテナントコンピュータネットワークによって実現されるデータベースにおける各エントリが、テナントIDでラベル付けされてもよい。対応するテナントIDに関連付けられたテナントのみが、特定のエントリのデータにアクセスしてもよい。しかしながら、データベースは複数のテナントによって共有されてもよい。

40

**【0068】**

ある実施形態では、サブスクリプションリストは、どのテナントがどのアプリケーションにアクセスするための認証を有するかを示す。アプリケーションごとに、アプリケーシ

50

ョンへのアクセスを認証されたテナントのテナントIDのリストが格納されている。テナントのテナントIDが、特定のアプリケーションに対応するサブスクリプションリストに含まれる場合のみ、テナントは、特定のアプリケーションへのアクセスが許可される。

【0069】

ある実施形態では、異なるテナントに対応するネットワークリソース（デジタルデバイス、仮想マシン、アプリケーションインスタンス、およびスレッドなど）同士は、マルチテナントコンピュータネットワークによって維持されるテナント特有オーバーレイネットワークに分離される。一例として、テナントオーバーレイネットワークにおける任意の発信元デバイスからのパケットは、同じテナントオーバーレイネットワーク内の他のデバイスにのみ送信され得る。テナントオーバーレイネットワーク上の発信元デバイスから他のテナントオーバーレイネットワークにおけるデバイスへの送信をいずれも禁止するために、カプセル化トンネルが使用される。具体的には、発信元デバイスから受信されたパケットは外部パケット内でカプセル化される。外部パケットは、（テナントオーバーレイネットワークにおける発信元デバイスと通信する）第1のカプセル化トンネルエンドポイントから、（テナントオーバーレイネットワークにおける宛先デバイスと通信する）第2のカプセル化トンネルエンドポイントへ送信される。第2のカプセル化トンネルエンドポイントは、発信元デバイスによって送信される元のパケットを取得するために、外部パケットを非カプセル化する。元のパケットは、第2のカプセル化トンネルエンドポイントから、同じ特定のオーバーレイネットワークにおける宛先デバイスへ送信される。

【0070】

#### 8. ハードウェア概要

一実施形態に従うと、ここに説明される手法は、1つ以上の専用コンピューティングデバイスによって実現される。これらの専用コンピューティングデバイスは、これらの手法を実行するためにハードワイヤードされてもよく、もしくは、これらの手法を実行するように永続的にプログラムされた1つ以上の特定用途向け集積回路（application-specific integrated circuit：ASIC）、フィールドプログラマブルゲートアレイ（field programmable gate array：FPGA）、もしくはネットワーク処理ユニット（network processing unit：NPU）といったデジタル電子デバイスを含んでいてもよく、または、ファームウェア、メモリ、他の記憶装置、もしくはそれらの組合せにおいてプログラム命令に従ってこれらの手法を実行するようにプログラムされた1つ以上の汎用ハードウェアプロセッサを含んでいてもよい。このような専用コンピューティングデバイスはまた、これらの手法を遂行するために、カスタムハードワイヤードロジック、ASIC、FPGA、またはNPUをカスタムプログラミングと組み合わせてもよい。当該専用コンピューティングデバイスは、デスクトップコンピュータシステム、ポータブルコンピュータシステム、手持ち型デバイス、ネットワークデバイス、または、これらの手法を実現するためにハードワイヤードロジックおよび/もしくはプログラムロジックを組込んだ任意の他のデバイスであってもよい。

【0071】

たとえば、図4は、本発明の一実施形態が実現され得るコンピュータシステム400を示すブロック図である。コンピュータシステム400は、情報を通信するためのバス402または他の通信機構と、情報を処理するためにバス402と結合されるハードウェアプロセッサ404を含む。ハードウェアプロセッサ404は、たとえば、汎用マイクロプロセッサであってもよい。

【0072】

コンピュータシステム400はまた、プロセッサ404によって実行されるべき命令および情報を格納するためにバス402に結合された、ランダムアクセスメモリ（random access memory：RAM）または他の動的記憶装置などのメインメモリ406を含む。メインメモリ406はまた、プロセッサ404によって実行されるべき命令の実行中に一時的な変数または他の中間情報を格納するために使用されてもよい。そのような命令は、プロセッサ404にアクセス可能な非一時的記憶媒体に格納されると、コンピュータシス

テム 4 0 0 を、命令で特定された動作を実行するようにカスタマイズされた専用マシンにする。

【 0 0 7 3 】

コンピュータシステム 4 0 0 はさらに、プロセッサ 4 0 4 のための静的情報および命令を格納するためにバス 4 0 2 に結合された読出専用メモリ (read only memory : R O M ) 4 0 8 または他の静的記憶装置を含む。磁気ディスクまたは光ディスクなどの記憶装置 4 1 0 が設けられ、情報および命令を格納するためにバス 4 0 2 に結合される。

【 0 0 7 4 】

コンピュータシステム 4 0 0 は、情報をコンピュータユーザへ表示するための陰極線管 (cathode ray tube : C R T )、液晶ディスプレイ (liquid crystal display : L C D )、または発光ダイオード (light emitting diode : L E D ) などのディスプレイ 4 1 2 に、バス 4 0 2 を介して結合され得る。英数字キーおよび他のキーを含む入力装置 4 1 4 が、情報およびコマンド選択をプロセッサ 4 0 4 へ通信するためにバス 4 0 2 に結合される。別のタイプのユーザ入力装置は、方向情報およびコマンド選択をプロセッサ 4 0 4 へ通信するための、ならびに、ディスプレイ 4 1 2 上のカーソル移動を制御するための、マウス、トラックボール、またはカーソル方向キーなどのカーソル制御装置 4 1 6 である。この入力装置は典型的には、第 1 の軸 (たとえば x ) および第 2 の軸 (たとえば y ) という 2 つの軸において 2 つの自由度を有しており、これにより、当該装置が平面における位置を特定することを可能にする。

【 0 0 7 5 】

コンピュータシステム 4 0 0 は、コンピュータシステムと組合されることでコンピュータシステム 4 0 0 を専用マシンにするかまたは専用マシンになるようプログラムする、カスタマイズされたハードワイヤードロジック、1 つ以上の A S I C もしくは F P G A、ファームウェアおよび/またはプログラムロジックを用いて、ここに説明された手法を実現し得る。一実施形態に従うと、本明細書中の手法は、メインメモリ 4 0 6 に含まれる 1 つ以上の命令の 1 つ以上のシーケンスをプロセッサ 4 0 4 が実行することに対応して、コンピュータシステム 4 0 0 によって実行される。このような命令は記憶装置 4 1 0 などの別の記憶媒体からメインメモリ 4 0 6 に読込まれてもよい。メインメモリ 4 0 6 に含まれる命令のシーケンスを実行することにより、ここに説明されたプロセスステップをプロセッサ 4 0 4 に実行させる。代替的な実施形態では、ハードワイヤード回路が、ソフトウェア命令の代わりに、またはソフトウェア命令と組合わせて使用されてもよい。

【 0 0 7 6 】

本明細書中において用いられる「記憶媒体」という語は、マシンを特定の態様で動作させるデータおよび/または命令を格納する任意の非一時的媒体を指す。そのような記憶媒体は不揮発性媒体および/または揮発性媒体を含み得る。不揮発性媒体は、たとえば、記憶装置 4 1 0 などの光ディスク、磁気ディスク、またはソリッドステートディスクを含む。揮発性媒体はメインメモリ 4 0 6 などのダイナミックメモリを含む。記憶媒体の一般的な形態は、たとえば、フロッピー (登録商標) ディスク、フレキシブルディスク、ハードディスク、ソリッドステートドライブ、磁気テープ、または任意の他の磁気データ記憶媒体、C D - R O M、任意の他の光学データ記憶媒体、穴のパターンを有する任意の物理的媒体、R A M、P R O M、E P R O M、F L A S H - E P R O M、N V R A M、任意の他のメモリチップまたはカートリッジ、連想メモリ (content-addressable memory : C A M )、および 3 値連想メモリ (ternary content-addressable memory : T C A M ) を含む。

【 0 0 7 7 】

記憶媒体は伝送媒体とは異なるが、伝送媒体とともに用いられてもよい。伝送媒体は、記憶媒体間の情報の転送に参与する。たとえば、伝送媒体は、バス 4 0 2 を含むワイヤを含む、同軸ケーブル、銅線、および光ファイバを含む。伝送媒体はまた、電波通信および赤外線データ通信時に生成されるような音波または光波の形態を取ることもできる。

【 0 0 7 8 】

10

20

30

40

50

さまざまな形態の媒体が、1つ以上の命令の1つ以上のシーケンスを実行のためにプロセッサ404へ搬送することに関与し得る。たとえば、命令は最初、リモートコンピュータの磁気ディスクまたはソリッドステートドライブ上で搬送され得る。リモートコンピュータは、そのダイナミックメモリに命令をロードし、ネットワークインターフェイスカード(network interface card: NIC)を用いてネットワーク回線を介して、またはモデムを用いて電話回線を介して命令を送信し得る。コンピュータシステム400に対してローカルなNICは、ネットワークライン上でデータを受信し、バス402上にデータを配置することができる。コンピュータシステム400に対してローカルなモデムが、電話回線を介してデータを受信し、赤外線送信機を用いてデータを赤外線信号に変換し得る。赤外線検出器が、赤外線信号で搬送されたデータを受信し得るとともに、適切な回路がデータをバス402上に配置し得る。バス402はデータをメインメモリ406へ搬送し、そこから、プロセッサ404が命令を取出して実行する。メインメモリ406が受信した命令は、任意には、プロセッサ404による実行の前または後に、記憶装置410上に格納されてもよい。

10

**【0079】**

コンピュータシステム400はまた、バス402に結合された通信インターフェイス418を含む。通信インターフェイス418は、ローカルネットワーク422に接続されたネットワークリンク420への双方向データ通信結合を提供する。たとえば、通信インターフェイス418は、総合サービスデジタル網(integrated services digital network: ISDN)カード、ケーブルモデム、衛星モデム、または、対応するタイプの電話回線へのデータ通信接続を提供するためのモデムであってもよい。別の例として、通信インターフェイス418は、互換性のあるLANへのデータ通信接続を提供するためのローカルエリアネットワーク(local area network: LAN)カードであってもよい。無線リンクが実現されてもよい。任意のそのような実現例では、通信インターフェイス418は、さまざまなタイプの情報を表わすデジタルデータストリームを搬送する電気信号、電磁信号、または光信号を送受信する。

20

**【0080】**

ネットワークリンク420は典型的には、1つ以上のネットワークを通して他のデータデバイスにデータ通信を提供する。たとえば、ネットワークリンク420は、ローカルネットワーク422を介して、ホストコンピュータ424に、またはインターネットサービスプロバイダ(Internet Service Provider: ISP)426によって操作されるデータ機器に接続し得る。ISP426は次に、現在一般的に「インターネット」428と称されている全世界的パケットデータ通信ネットワークを通してデータ通信サービスを提供する。ローカルネットワーク422およびインターネット428はともに、デジタルデータストリームを搬送する電気信号、電磁信号または光信号を用いる。さまざまなネットワークを通る信号、および、通信インターフェイス418を通るネットワークリンク420上の信号はコンピュータシステム400との間でデジタルデータを搬送するものであって、伝送媒体の例示的な形態である。

30

**【0081】**

コンピュータシステム400は、ネットワーク、ネットワークリンク420、および通信インターフェイス418を通して、メッセージを送信し、プログラムコードを含むデータを受信し得る。インターネットの例では、サーバ430が、インターネット428、ISP426、ローカルネットワーク422、および通信インターフェイス418を通して、アプリケーションプログラムのための要求されたコードを送信し得る。

40

**【0082】**

受信されたコードは、受信時にプロセッサ404によって実行されてもよく、および/または、後で実行するために記憶装置410または他の不揮発性記憶装置に格納されてもよい。

**【0083】**

実施形態は、ハードウェアプロセッサを含むとともに本明細書に記載された動作および

50

／または添付の特許請求の範囲のいずれかに列挙される動作のいずれかを実行するように構成された、1つ以上のデバイスを含むシステムを対象とする。

【0084】

ある実施形態では、非一時的なコンピュータ可読記憶媒体は、1つ以上のハードウェアプロセッサによって実行されると、本明細書に記載される動作および／または特許請求の範囲のいずれかに列挙される動作のいずれかを実行させる命令を備える。

【0085】

ここに説明された特徴と機能との任意の組合せが、1つ以上の実施形態に従って用いられてもよい。前述の明細書では、実施形態は、実現例ごとに異なり得る多くの具体的詳細を参照して説明されてきた。したがって、明細書および添付の図面は、限定的な意味ではなく例示的な意味で考慮されるべきである。本発明の範囲の唯一かつ排他的な指標、および、出願人によって本発明の範囲であることが意図されるものは、任意の後の補正を含めて、そのような特許請求の範囲が生じる特定の形態での、本願から生じる特許請求の範囲のセットの文字通りかつ同等の範囲である。

10

20

30

40

50

【図面】  
【図 1】

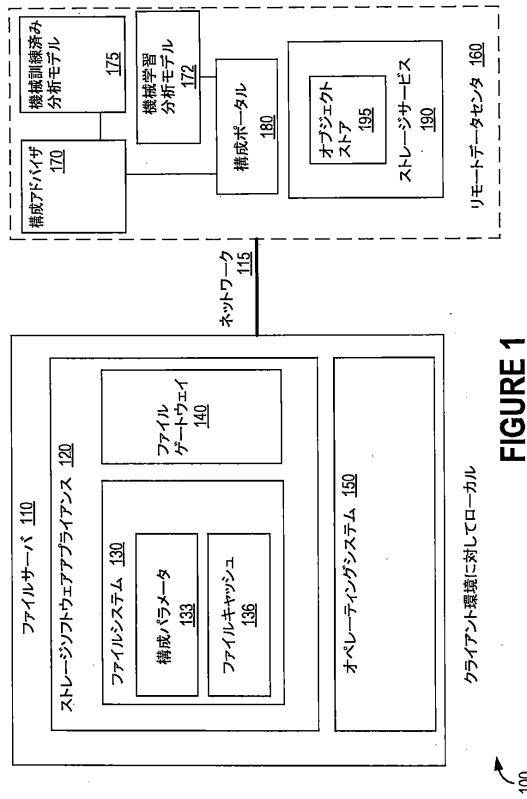


FIGURE 1

【図 2 A】

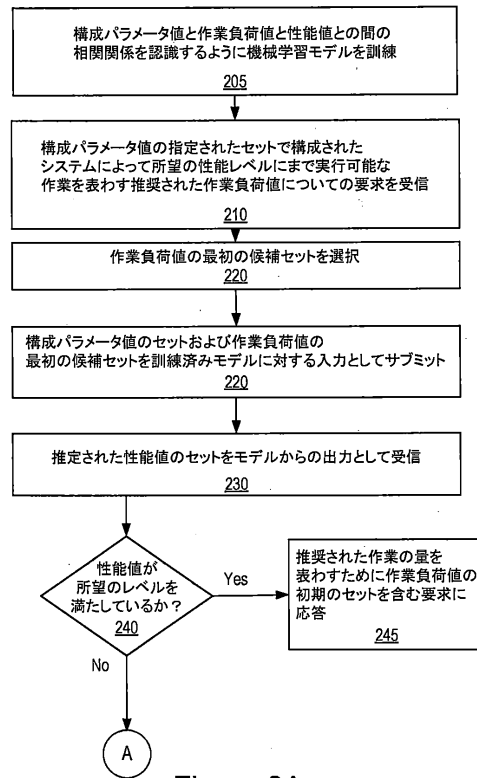


Figure 2A

【図 2 B】

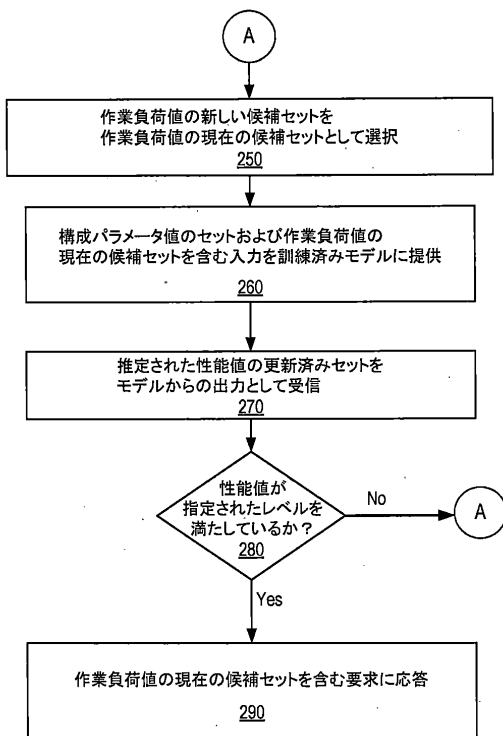


Figure 2B

【図 3 A】

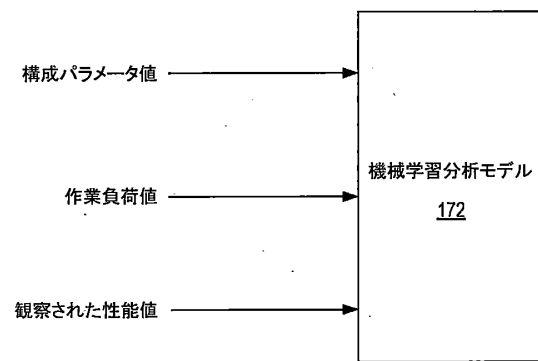


Figure 3A

10

20

30

40

50

【図 3 B】

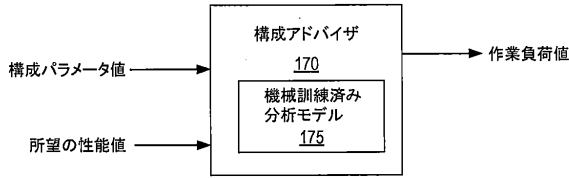


Figure 3B

【図 3 C】

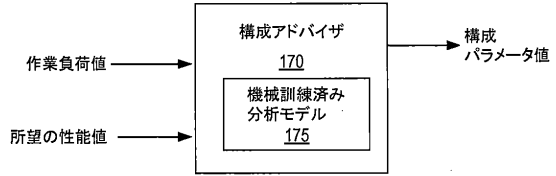


Figure 3C

【図 3 D】

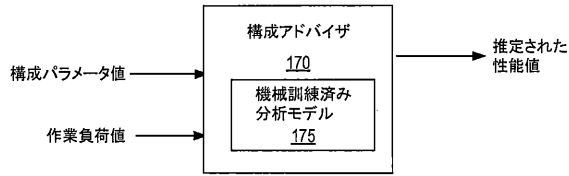
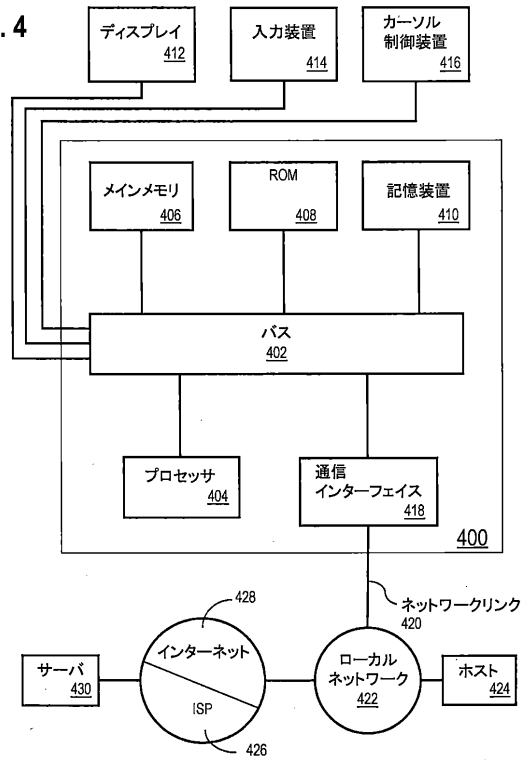


Figure 3D

【図 4】

FIG. 4



10

20

30

40

50

## フロントページの続き

- ラクル・パークウェイ、500
- (72)発明者 バイヤーズ, ブランドン・パトリック  
アメリカ合衆国、94065 カリフォルニア州、レッドウッド・シティー、オラクル・パークウ  
エイ、500
- (72)発明者 ダッシュ, プラサンタ・ランジャン  
アメリカ合衆国、94065 カリフォルニア州、レッドウッド・シティー、オラクル・パークウ  
エイ、500
- (72)発明者 グロス, ケニー・クレイトン  
アメリカ合衆国、94065 カリフォルニア州、レッドウッド・シティー、オラクル・パークウ  
エイ、500
- 審査官 円子 英紀
- (56)参考文献 米国特許出願公開第2015/0379420 (US, A1)  
特開2010-237901 (JP, A)  
国際公開第2014/054230 (WO, A1)  
特開2014-174609 (JP, A)
- (58)調査した分野 (Int.Cl., DB名)  
G06F 11/30  
G06F 11/34  
G06F 16/13