

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5632010号  
(P5632010)

(45) 発行日 平成26年11月26日 (2014.11.26)

(24) 登録日 平成26年10月17日 (2014.10.17)

(51) Int. Cl.	F I
<b>G06F 12/00 (2006.01)</b>	G06F 12/00 545A
<b>G06F 13/10 (2006.01)</b>	G06F 12/00 514E
	G06F 13/10 330C
	G06F 13/10 340A

請求項の数 14 (全 18 頁)

(21) 出願番号	特願2012-539922 (P2012-539922)	(73) 特許権者	500046438
(86) (22) 出願日	平成22年10月29日 (2010.10.29)		マイクロソフト コーポレーション
(65) 公表番号	特表2013-511104 (P2013-511104A)		アメリカ合衆国 ワシントン州 9805
(43) 公表日	平成25年3月28日 (2013.3.28)		2-6399 レッドモンド ワン マイ
(86) 国際出願番号	PCT/US2010/054701		クロソフト ウェイ
(87) 国際公開番号	W02011/059811	(74) 代理人	100107766
(87) 国際公開日	平成23年5月19日 (2011.5.19)		弁理士 伊東 忠重
審査請求日	平成25年10月8日 (2013.10.8)	(74) 代理人	100070150
(31) 優先権主張番号	12/619,307		弁理士 伊東 忠彦
(32) 優先日	平成21年11月16日 (2009.11.16)	(74) 代理人	100091214
(33) 優先権主張国	米国 (US)		弁理士 大貫 進介

最終頁に続く

(54) 【発明の名称】 プロブとしての仮想ハードドライブ管理

(57) 【特許請求の範囲】

【請求項 1】

クラウドコンピューティングプラットフォームにおいてプロブとしての仮想ハードドライブを管理するコンピュータ実装方法であって：

一つまたはそれ以上のアプリケーションに対して仮想ハードドライブとして公開されるプロブをプロブストア内に作成するステップと、

前記仮想ハードドライブをマウントするステップであり、前記一つまたはそれ以上のアプリケーションから前記仮想ハードドライブに対する読み出し及び書き込み要求は、前記プロブに対してコミットされる、ステップと、

前記仮想ハードドライブへの読み出し及び書き込み要求を管理するステップと、を含み

10

前記管理は：

アプリケーションプログラミングインターフェースを使用してアプリケーションによって生成された要求をインターセプトする段階であり、前記アプリケーションプログラミングインターフェースは、前記仮想ハードドライブへのアクセスをサポートする、段階と

前記アプリケーションによって生成された前記要求をプロブ指令に変換する段階であり、前記プロブ指令は、プロブインターフェースを介した前記プロブへのアクセスをサポートする、段階と、

前記プロブ指令が前記プロブストアで実行されるように、前記プロブ指令を前記プロ

20

ブインターフェースに対して通信する段階と、  
に基づいて行われる、  
 ことを特徴とするコンピュータ実装方法。

【請求項 2】

前記一つまたはそれ以上のアプリケーションは、仮想マシン上で実行され、かつ、  
 前記読み出し及び書き込み要求は、レガシーアプリケーションである、  
 請求項 1 に記載のコンピュータ実装方法。

【請求項 3】

前記レガシーアプリケーションは、  
 N - T - F - S ファイルシステムまたは F - A - T ファイルシステムの内の少なくとも  
一つを含んでいる、  
 請求項 2 に記載のコンピュータ実装方法。

10

【請求項 4】

プロブは、前記プロブストアに対する前記プロブインターフェースを介してアクセスさ  
れ、かつ、操作され、  
前記プロブストアはレプリケーションをサポートする耐久性クラウドストレージシステ  
ムである、  
 請求項 1 に記載のコンピュータ実装方法。

【請求項 5】

前記一つまたはそれ以上のアプリケーションから選択された少なくとも一つのアプリケ  
ーションは、前記アプリケーションインターフェースと前記プロブインターフェースの両  
方を介してプロブにアクセスし、前記プロブは、前記プロブインターフェースを介して、  
アップロード、ダウンロード、又は、デバッグされる、  
 請求項 1 に記載のコンピュータ実装方法。

20

【請求項 6】

所定の期間後に終了するプロブリースを介して、アプリケーションのためのプロブに対  
して排他的書き込みアクセスが提供され、かつ、  
 他のアプリケーションが前記プロブから読み出すことが許可されている間、前記アプリ  
 ケーションは、前記リースの有効性に基いて条件的に前記プロブに書き込む、  
 請求項 1 に記載のコンピュータ実装方法。

30

【請求項 7】

クラウドコンピューティングプラットフォーム内で仮想ハードドライブをプロブとして  
 管理する方法を実行する命令を格納した一つまたはそれ以上のコンピュータで読取り可能  
 な媒体であって、前記方法は：

前記クラウドコンピューティングプラットフォーム内の仮想マシン上で実行しているア  
 プリケーションから、仮想ハードドライブに対する I / O 要求を受け取るステップと、  
 前記仮想ハードドライブに対する I / O 要求をドライバによって管理するステップと、  
 を含み、

前記管理するステップは、前記仮想ハードドライブに関連するローカルキャッシュに  
アクセスする段階であり、前記ローカルキャッシュはプロブストア内のデータに対応する  
データを保管している段階に基づき、

40

読み出し要求が前記ローカルキャッシュに基づいて実行され、かつ、書き込み要求が  
、

アプリケーションプログラミングインターフェースを使用してアプリケーションに  
よって生成された前記書き込み要求をインターセプトする段階であり、前記アプリケーシ  
ョンプログラミングインターフェースは、前記仮想ハードドライブへのアクセスをサポート  
する、段階と、

前記アプリケーションによって生成された前記書き込み要求をプロブ指令に変換す  
る段階であり、前記プロブ指令は、プロブインターフェースを介した前記プロブへのアク  
セスをサポートする、段階と、

50

前記プロブ指令が前記プロブストアで実行されるように、前記プロブ指令を前記プロブインターフェースに対して通信する段階と、

に基づいて実行される、

ことを特徴とする媒体。

【請求項 8】

以前の書き込みを、前記クラウドコンピューティングプラットフォーム内に依然として残っている同一のデータに対する古い書き込みによって上書きすることを防止するために、シーケンス番号が前記 I/O 要求と関連付けられており、

前記ドライバーは、前記プロブストアによりアクノリッジされていない以前の書き込みを無視するために、前記シーケンス番号を増加させる、

請求項 7 に記載のコンピュータで読取り可能な媒体。

【請求項 9】

前記ドライバーは、前記プロブストア内の前記プロブに対して重なった書き込み要求を命令し、かつ、並行して前記プロブストア内の前記プロブに対して重ならない書き込み要求を送付する、

請求項 8 に記載のコンピュータで読取り可能な媒体。

【請求項 10】

前記ローカルキャッシュの更新は、前記プロブストアからアクノリッジが受信された後、又は、前記プロブストアへの書き込み要求と並行して、のいずれかに発生し、前記プロブストアからアクノリッジが受信された後に前記更新が可能である、

請求項 7 に記載のコンピュータで読取り可能な媒体。

【請求項 11】

仮想ハードドライブをプロブとして管理するクラウドコンピューティングプラットフォームであって：

データを読み出し、及び、書き込むために仮想ハードドライブにアクセスするアプリケーションを実装する仮想マシンを実行するように構成されたファブリックコンピュータと；

仮想ハードドライブとして前記アプリケーションに公開されたプロブを格納するように構成されたプロブストアと、

を含み、

アプリケーションインターフェースとプロブインターフェースが、プロブと仮想ハードドライブに対するアクセスを可能にし、

プロブにアクセスすることは、

アプリケーションプログラミングインターフェースを使用してアプリケーションによって生成された要求をインターセプトする段階であり、前記アプリケーションプログラミングインターフェースは、前記仮想ハードドライブへのアクセスをサポートする、段階と

、

前記アプリケーションによって生成された前記要求をプロブ指令に変換する段階であり、前記プロブ指令は、プロブインターフェースを介した前記プロブへのアクセスをサポートする、段階と、

前記プロブ指令が前記プロブストアで実行されるように、前記プロブ指令を前記プロブインターフェースに対して通信する段階と、

を含む、

ことを特徴とするクラウドコンピューティングプラットフォーム。

【請求項 12】

前記ファブリックコンピュータは、

プロブデータをキャッシュするローカルキャッシュと、

プロブ指令を格納するプロブライブラリーと、

仮想ハードドライブ I/O 命令をプロブ指令に翻訳するドライバと、を含む、

請求項 11 に記載のクラウドコンピューティングプラットフォーム。

**【請求項 1 3】**

前記ローカルキャッシュは、各データブロックに対して前記データブロックを使用する前に前記データブロック内にエラーが存在するか否かを判別する巡回冗長検査コードを格納する、

請求項 1 2 に記載のクラウドコンピューティングプラットフォーム。

**【請求項 1 4】**

前記アプリケーションが、前記ローカルキャッシュ内に格納されたデータを使用して要求を完結することが可能なように、前記ローカルキャッシュと前記プロブが同期される、

請求項 1 3 に記載のクラウドコンピューティングプラットフォーム。

**【発明の詳細な説明】**

10

**【技術分野】****【0001】**

本発明は、プロブとしての仮想ハードドライブ管理に関する。

**【背景技術】****【0002】**

従来より、クラウドコンピューティングプラットフォームはインターネット接続可能な仮想環境においてソフトウェアアプリケーションをホストしている。クラウドコンピューティングプラットフォームはサードパーティーによって設計されて維持されているデータセンターを団体が利用することを許している。従来の仮想環境は要求されたハードウェアリソース、ソフトウェアアプリケーションリソース、ネットワークリソース、及びストレージリソースを小さな又は大きな団体に提供する。仮想環境はまた、アプリケーションのセキュリティ、アプリケーションの信頼性、アプリケーションのスケラビリティ及び可用性も提供する。

20

**【発明の概要】****【0003】**

従来のデータセンターは物理コンピューティングリソース、物理ストレージリソース及び物理ネットワークリソースを提供する。データセンター内の物理リソース仮想化され、一組のアプリケーションプログラミングインターフェースとして団体に公開される。団体は自身のハードウェアリソース又はソフトウェアリソースを維持する必要が無く、又は、信頼性が有りスケラブルなデータセンターを維持することは必要でない。

30

**【0004】**

団体は仮想環境を介してソフトウェア又は根底にある物理ハードウェアの詳細を知ること無しにこれらの物理リソースに効率的にアクセスすることができる。

**【0005】**

発明の実施形態は、一面において、クラウドコンピューティングプラットフォーム、コンピュータにより読取り可能な媒体、及びクラウドコンピューティングプラットフォーム内の仮想ハードドライブを管理するコンピュータ実装方法に関している。クラウドコンピューティングプラットフォームはファブリックコンピュータと仮想ハードドライブにアクセスするために用いられるプロブストアとを含んでいる。格納されたプロブが有効で耐久性があることを確実にするために各プロブの幾つかのコピーを保持する複製を用いる耐久性クラウドストレージシステムである。

40

**【0006】**

ファブリックコンピュータは複数のアプリケーションを実装している仮想マシンを実行するように構成されている。ファブリックコンピュータは仮想ハードドライブにアクセスして入力/出力(I/O)要求を介してデータを読み出し書き込む。ファブリックコンピュータはドライブライブラリー、複数のドライバ、ローカルキャッシュ及びプロブストアへのインターフェースを含む。アプリケーションはI/O要求を生成する。一方、I/O要求はI/O要求をプロブ指令(コマンド)に翻訳するドライバに送られる。ドライバはI/O要求をローカルキャッシュに格納されたデータで完結させても良く、或いはドライバはI/O要求によってアクセスされた仮想ハードドライブに関連したプロブ内に格納さ

50

れたデータを取得するためにプロブストアにアクセスしても良い。プロブ指令はプロブストアにアクセスするために用いられる。プロブストアはアプリケーションに仮想ハードドライブとして公開されるプロブを格納するように構成されている。プロブにアクセスすることを可能にするアプリケーションプログラミングインターフェース及びセマンティックは仮想ハードドライブにアクセスするためにアプリケーションが実行するアプリケーションプログラミングインターフェース及びセマンティックとは異なっている。

【 0 0 0 7 】

この概要は以下の詳細な説明で更に記述された概念の1つの選択を簡単な形式で導入するために提供されたものである。この概要は請求項に記載された主題の主要な特徴又は必須の特徴を特定することを意図したものでなく、請求項に記載された主題の範囲を決定する際の補助として独立して使用されることを意図したものでない。

10

【図面の簡単な説明】

【 0 0 0 8 】

【図1】発明の実施形態に応じた模範的なクラウドコンピューティングプラットフォームを図示したネットワーク図である。

【図2】模範的なクラウドコンピューティングプラットフォームにおける模範的なファブリックコンピュータと模範的なプロブストアを図示したブロック図である。

【図3】仮想ハードドライブとして公開されるプロブを作成するための模範的な方法を図示した論理図である。

【図4】仮想ハードドライブへのI/O要求を管理する模範的な方法を図示した論理図である。

20

【発明を実施するための形態】

【 0 0 0 9 】

この特許は特許を受ける主題を法定の要件を満たすための詳細に記述したものである。しかしながら、記述自体はこの特許の範囲を限定することを意図したものではない。寧ろ、発明者は請求項に記載された主題は、他の現在又は未来の技術と共に、異なるステップ又はこの特許において記述されたものに類似したステップの組み合わせを含めることによって他の方法でも実施され得ることを考慮にしている。更に、本明細書において「ステップ」及び「ブロック」という用語は採用された方法の異なる要素を暗示しているように用いられているけれども、個々のステップの順序が明示的に記述されない限り、本明細書において開示された種々のステップ間の特定の順序を暗示していると解釈されるべきではない。更に、実施形態は添付された図面を参照しつつ以下に詳述されており、それらの図面の全体が本明細書に援用により包含される。

30

【 0 0 1 0 】

本明細書で用いられているように、用語「プロブ」はバイナリーラージオブジェクトを意味している。

【 0 0 1 1 】

幾つかの実施形態において、クラウドコンピューティングプラットフォーム内で動作しているアプリケーションは耐久性が有り有効な仮想ハードドライブ内に格納されたファイルシステム内のデータにアクセスする。仮想ハードドライブ内のデータはソフトウェア又はハードウェアが故障した場合（例えば、ドライブの故障、ノードの故障、ラックの故障、ビット(bit rot)等の場合)にも利用可能状態を維持する。アプリケーションは仮想ハードドライブ内に格納されたデータにアクセスするために、ウインドウズ(登録商標)NTFS A P I等のアプリケーションプログラミングインターフェース(A P I)及びセマンティックを使用しても良い。

40

【 0 0 1 2 】

一実施形態において、仮想ハードドライブはマウント可能なブロックレベルの仮想装置である。仮想ハードドライブはクラウドコンピューティングプラットフォーム内の仮想マシン上で動作しているアプリケーションが、アプリケーションのために記憶装置とのインターフェースを行うファイルシステムA P Iを使用して仮想ハードドライブ内のデータに

50

アクセスすることを可能にする。仮想ハードドライブは耐久性が有り、プロブストアへの書き込みを確定することによって全ての書き込みについてデータの消失を防止している。仮想ハードドライブはプロブである。仮想ハードドライブ及び仮想ハードドライブに関連したメタデータの内容はプロブに格納される。プロブはプロブインターフェースを介してプロブストアからアップロード又はプロブストアにダウンロードすることができる。加えて、プロブインターフェースはアプリケーションがプロブに関連した仮想ハードドライブをマウントし、仮想ハードドライブ内に格納されたデータにアクセスすることを許可することができる。

#### 【 0 0 1 3 】

幾つかの実施形態において、仮想ハードドライブはアプリケーションによってボリュームとしてマウントされる。アプリケーション I / O 要求プロブストアに格納されたプロブにリダイレクトされる。一実施形態において、プロブへの書き込みは書き込まれたデータをプロブストアに格納することによって耐久性の有るものとされ、プロブストアはプロブストアにおいて書き込みが耐久性の有るものとされた後にアプリケーションに書き込みが成功したことをアクノリッジ（注：応答すること）する。仮想ハードドライブがマウント解除されるか又はドライブをマウントした仮想マシンが故障したときには、仮想ハードドライブに関連したプロブが耐久性の有るプロブストア内に残る。従って、他の仮想マシンが仮想ハードドライブに関連した同一のプロブをマウントし、プロブに格納されたデータにアクセスしても良い。

#### 【 0 0 1 4 】

クラウドコンピューティングプラットフォームは特に、記憶装置のリソースを使用する可能性のあるリレーショナルデータベースを含むレガシーアプリケーションの移動と実行、故障や遮断時の仮想マシン上のデータの保存、他の仮想マシンによって使用されるためのデータの記憶、及び多数の仮想マシン間での同時のデータ共有を可能にする。レガシーアプリケーションの移動と実行は、仮想マシンがウインドウズ（R）NTFS ファイルシステムを介した I / O 要求を開始するように構成されたアプリケーションを実行することを可能にする。仮想マシンはまた、mySQL または Oracle 等の構造化問い合わせ言語（SQL）を使用するリレーショナルデータベースを追加的なデータベースミラーリングを実行する必要無しに実行する。仮想マシンは仮想ハードドライブにデータを格納し、仮想ハードドライブに格納されたデータは、もし現在の仮想マシンが故障したときには追加の仮想マシンにより利用可能である。スナップショットは仮想ハードドライブに関連したプロブの現在の状態を記憶する。スナップショットは仮想マシンによってアクセス可能な読み出し専用の仮想ハードドライブを作成するために用いられても良い。仮想ハードドライブに格納されたデータは多数の仮想マシンに同時に利用可能であっても良い。一実施形態において、クラウドコンピューティングプラットフォームは仮想ハードドライブを更新し、スナップショットを作成し、その後スナップショットを他の仮想マシンと共用する。

#### 【 0 0 1 5 】

一実施形態において、クラウドコンピューティングプラットフォームは物理マシンを仮想マシンとして公開させてもよい。物理マシンは仮想マシンによって使用される指令を介してアクセスされても良い。

#### 【 0 0 1 6 】

当業者によって理解されるように、クラウドコンピューティングプラットフォームはハードウェア、ソフトウェア、又はハードウェアとソフトウェアの組み合わせを含んでも良い。ハードウェアはメモリに格納された命令を実行するように構成されたプロセッサとメモリを含む。一実施形態において、メモリはコンピュータ実装方法のためにコンピュータが使用することが可能な命令を有するコンピュータプログラム製品を格納したコンピュータにより読取り可能な媒体を含む。コンピュータにより読取り可能な媒体は揮発性及び不揮発性媒体の両方、取り外し可能及び固定の媒体、及びデータベース、スイッチ、及び種々の他のネットワーク装置によって読み出し可能な媒体を含む。ネットワークスイッチ

10

20

30

40

50

、ルーター、及び関連したコンポーネントは従来の性質のものであり、それと通信するための手段である。限定ではない例として、コンピュータにより読取り可能な媒体にはコンピュータ記憶媒体及び通信媒体が含まれる。コンピュータ記憶媒体、又はマシンによりマシン能な媒体は、情報を格納する方法又は技術によって実現される媒体を含む。格納された情報の例にはコンピュータにより利用可能な命令、データ構造、プログラムモジュール、及び他のデータ表現が含まれる。コンピュータ記憶媒体にはランダムアクセスメモリー（ＲＡＭ）、リードオンリーメモリー（ＲＯＭ）、電氣的消去可能なプログラマブルリードオンリーメモリー（ＥＥＰＲＯＭ）、フラッシュメモリー又は他のメモリ術、コンパクトディスクリードオンリーメモリー（ＣＤ－ＲＯＭ）、デジタル多目的ディスク（ＤＶＤ）、ホログラム媒体又は他の光学式ディスク記憶装置、磁気カセット、磁気テープ、磁気ディスク記憶装置及び他の磁気記憶装置が含まれるがこれらに限定されない。これらのメモリ術はデータを瞬時的、一時的、又は永久に格納することが可能である。

10

#### 【００１７】

一実施形態において、クラウドコンピューティングプラットフォームはクライアント装置が利用できるクラウドアプリケーションを含んでいる。クライアント装置はクラウドコンピューティングプラットフォームにアクセスしてクラウドアプリケーションを実行する。クラウドアプリケーションはクラウドコンピューティングプラットフォームにおいて利用できるストレージとプロセッシングリソースを使用して実装される。

#### 【００１８】

図１は本発明の実施形態による模範的なコンピューティングシステム１００を図示したネットワーク図である。図１に示されたコンピューティングシステム１００は単なる例示であって、範囲と機能に関して如何なる限界を示唆することも意図していない。本発明の実施形態は多数の他の構成と共に動作することが可能である。図１を参照すると、コンピューティングシステム１００はクラウドコンピューティングプラットフォーム１１０、クラウドアプリケーション１２０、及びクライアント装置１３０を含んでいる。

20

#### 【００１９】

クラウドコンピューティングプラットフォーム１１０はクライアント装置１３０によって要求されたクラウドアプリケーション１２０を実行するように構成されている。クラウドコンピューティングプラットフォーム１１０はプロブストアを維持している。プロブストアはクラウドアプリケーション１２０によってアクセスされたデータを格納したプロブを提供する。クラウドコンピューティングプラットフォーム１１０は、無線ネットワーク、ローカルエリアネットワーク、有線ネットワーク、又はインターネット等の通信ネットワークを介してクライアント装置１３０に接続する。

30

#### 【００２０】

クラウドアプリケーション１２０はクライアント装置１３０が利用可能である。クラウドコンピューティングプラットフォーム１１０上で実行されたソフトウェアがクラウドアプリケーション１２０を実装する。一実施形態において、クラウドコンピューティングプラットフォーム１１０内の仮想マシンがクラウドアプリケーション１２０を実行する。クラウドアプリケーション１２０は編集アプリケーション、ネットワーク管理アプリケーション、財務アプリケーション、又はクライアント装置１３０によって要求又は開発されたどのようなアプリケーションをも含むことができる。或る実施形態において、クラウドアプリケーション１３０の幾つかの機能はクライアント装置１３０上で実行されても良い。

40

#### 【００２１】

クライアント装置１３０はユーザーがクラウドコンピューティングプラットフォーム１１０によって提供されたクラウドアプリケーション１２０と対話するために使用される。幾つかの実施形態においてクライアント装置１３０はクラウドアプリケーション１２０にアクセスするために、クラウドコンピューティングプラットフォーム１１０に登録しなければならない。クラウドコンピューティングプラットフォーム１１０からのアカウントを有するクライアント装置１３０がクラウドアプリケーション１２０とクラウドコンピューティングプラットフォーム１１０内に提供された他のリソースにアクセスすることが可能

50

である。クライアント装置 130 にはパーソナルデジタルアシスタント、スマートフォン、ラップトップ、パーソナルコンピュータ、ゲームシステム、セットトップボックス、又は他の適当なクライアントコンピューティング装置が含まれるが、これらに限定されない。クライアント装置 130 はクライアント装置 130 上のユーザー及びシステム情報を格納するユーザー及びシステム情報記憶装置を含む。ユーザー情報はサーチ履歴、クッキー、パスワードを含んでも良い。システム情報はインターネットプロトコルアドレス、キャッシュされた Web ページ及びシステム利用率を含んでも良い。クライアント装置 130 はクラウドアプリケーション 120 から結果を受け取るためにクラウドコンピューティングプラットフォーム 110 と通信する。

#### 【0022】

従って、コンピューティングシステム 100 はクライアント装置 130 にクラウドアプリケーション 120 を提供するクラウドコンピューティングプラットフォーム 110 が構成されている。クラウドアプリケーション 120 はクライアント装置 130 上の複数のローカルクライアントアプリケーションを更新し管理する負担を除いている。

#### 【0023】

或る実施形態において、クラウドコンピューティングプラットフォームはファブリックコンピュータ及びプロブストアを提供する。ファブリックコンピュータは 1 以上の仮想マシンにおいてクラウドアプリケーションを実行する。プロブストアはクラウドアプリケーションによって使用されるデータを格納する。一方、クラウドアプリケーションはクラウドアプリケーション I/O 要求をプロブ指令に翻訳するドライバを介して仮想ハードドライブとしてのプロブとインターフェースで繋がる。

#### 【0024】

図 2 は模範的なファブリックコンピュータ 210 及び模範的なクラウドコンピューティングプラットフォームにおける模範的なプロブストア 230 を図示したブロック図である。ファブリックコンピュータ 210 がクラウドアプリケーション 221 を実装する仮想マシン 220 を実行する。プロブストア 230 はインターネットプロトコル (IP) アドレスを介してファブリックコンピュータ 210 によってアクセスすることができる。一実施形態において、クラウドアプリケーション 221 はファブリックコンピュータ 210 によって実行されても良く、後者は、クラウドアプリケーション 221 によって使用される仮想ハードドライブを実装したプロブを取得するためにプロブストアにアクセスする。

#### 【0025】

仮想マシン 220 はドライブライブラリー 222、I/O リダイレクション 223、ドライバ 224、ローカルキャッシュ 225 を含む。ドライブライブラリー 222 はドライブ管理を提供しても良く、I/O リダイレクション 223 はドライブ指令をドライブ指令をプロブ指令に翻訳するドライバ 224 にリダイレクトすることによって仮想ハードドライブに格納されたコンテンツにアクセスすることを可能にしても良い。一実施形態において、仮想マシン 220 はドライブライブラリー 222 に格納されたクリエイティブ (CreateDrive)、デリートドライブ (DeleteDrive)、リストドライブ (ListDrive)、マウントドライブ (MountDrive)、アンマウントドライブ (UnmountDrive) 及びスナップショットドライブ (SnapshotDrive) 等のドライブ指令を使用して仮想ハードドライブ (VHD) に対応したプロブ内のデータにアクセスしても良い。一実施形態において、ドライブライブラリー 222 は仮想マシン 220 にドライブ指令を検索するためのプロブ名前空間を提供しても良い。例えば、仮想マシン 220 はプロブ名前空間を <http://<accountname>.blob.core.windows.net/<containername>/<blobname>> 等のドライブライブラリー 222 における URL にアクセスすることによって検索しても良く、上記の URL において、accountname はクラウドコンピューティングプラットフォームのユーザーによって登録されたアカウントの名前であり、containername はプロブストアの名前であり、blobname はプロブの名前である。

#### 【0026】

一方、仮想マシン 220 は VHD に I/O 要求を送出しても良い。I/O 要求はドライ

10

20

30

40

50



バ 2 2 4 によってプロブ指令に翻訳される。プロブストア 2 3 0 内のプロブは仮想マシン 2 2 0 上で実行されているクラウドアプリケーション 2 2 1 に仮想ハードドライブ (V H D) として公開されても良い。一実施形態において、V H D はクラウドアプリケーション 2 2 1 によって時限的なリースでマウントされることが可能な特別なタイプのプロブ (例えばページプロブ) であっても良い。V H D は例えばウインドウズ (R) N T F S セマンティックなどのファイルシステムセマンティック又はドライブライブラリー 2 2 2 に含まれるドライブ指令を使用して如何なるクラウドアプリケーション 2 2 1 によってもアクセスすることができる。

【 0 0 2 7 】

V H D によって送出された I / O 要求は I / O リダイレクション 2 2 3 によって受取られる。I / O リダイレクション 2 2 3 は V H D に対してクラウドアプリケーション I / O 要求をドライバ 2 2 4 に送り出すように構成されたコンポーネントである。

【 0 0 2 8 】

或る実施形態において、ドライバ 2 2 4 はクラウドアプリケーションから受取ったドライブ指令をリースプロブ (LeaseBlob)、スナップショットプロブ (SnapshotBlob)、プットページ (PutPage)、クリアページ (ClearPage)、ゲットプロブ (GetBlob)、等のプロブ指令に翻訳するためにクラウドアプリケーションによって使用されるインターフェースを提供及び / 又は実現する。ドライバ 2 2 4 はまたマウントされた仮想ハードドライブに向けられたクラウドアプリケーション I / O 要求をプロブストア 2 3 0 に対して送出され、処理されたプロブ指令に翻訳する。ドライバ 2 2 4 は仮想ハードドライブのためのローカルキャッシュ 2 2 5 がプロブ内において仮想ハードドライブに対応して格納されたデータを格納するのを管理しても良い。ドライバ 2 2 4 はクラウドアプリケーション 2 2 1 からの読み出し動作に対応したデータを取り出すためにローカルキャッシュ 2 2 5 を用いても良い。クラウドアプリケーション 2 2 1 からの書き込み動作はローカルキャッシュ 2 2 5 及びプロブストア 2 3 0 内のプロブの両方に送出されても良い。他の実施形態において、ドライバ 2 2 4 がクラウドアプリケーションがリースが終了したときにローカルキャッシュ 2 2 5 内又はプロブストア内のデータにアクセスすることを防止するために仮想ハードドライブに対応したリースを管理するようにしても良い。

【 0 0 2 9 】

各ローカルキャッシュ 2 2 5 は単一の V H D と関連付けられている。マウントされた V H D のためのローカルキャッシュ 2 2 5 は同一のファブリックコンピュータ 2 1 0 上に配置されても良く、又はローカルキャッシュ 2 2 5 は V H D をマウントしたファブリックコンピュータ 2 1 0 の同一のラック内のコンピューティングリソース上に配置されても良い。この構成はネットワーク帯域幅を保持しても良い。或る実施形態において、仮想マシン 2 2 0 は複数の V H D を要求しても良い。各 V H D のために、仮想マシンはローカルキャッシュ 2 2 5 として使用するファブリックコンピュータ上のディスクドライブの容量を特定しても良い。ローカルキャッシュ 2 2 5 に格納されたデータを、仮想マシンが誤りの有るデータを使用することを防止するために、サイクリックリダンダンシーチェック (C R チェックされ) してもよい。

【 0 0 3 0 】

或る実施形態において、ローカルキャッシュ 2 2 5 は V H D の対応するプロブと同期している。ドライバ 2 2 4 はローカルキャッシュ 2 2 5 を管理しても良い。ドライバ 2 2 4 はプロブストア 2 3 0 内のプロブからアクノリッジが受信されるまでローカルキャッシュ 2 2 5 への書き込みを停止するように構成されても良い。あるいは、ドライバ 2 2 4 はローカルキャッシュ 2 2 5 とプロブストア 2 3 0 の両方に平行して並行して書き込んでも良い。もしプロブストア 2 3 0 に並行して書き込んでいる間に仮想マシン 2 2 0 が破損してリセットすると、仮想マシン 2 2 0 は仮想ハードドライブを再度マウントして対応するローカルキャッシュ 2 2 5 を再度使用することを試みても良い。仮想マシン 2 2 0 はどのデータブロックがプロブストア 2 3 0 からのアクノリッジ無しにローカルキャッシュ 2 2 5 に書き込まれたと推測されるか判別するためにローカルキャッシュ 2 2 5 をチェックして

10

20

30

40

50

も良い。一方、仮想マシン 220 はアクノリッジが返されないデータブロックのローカルキャッシュ 225 をクリアしても良い。他の実施形態において、仮想マシン 220 はローカルキャッシュ 225 の全体又はローカルキャッシュ 225 の選択された部分をクリアしても良い。ローカルキャッシュ 225 はファブリックコンピュータ 210 上の仮想マシン 220 によってしばしば使用されるデータを格納する。ローカルキャッシュ 225 はプロブストア 230 から要求された読み出しの数を減少させ、それによって、ネットワーク帯域幅が節約され、ネットワークの競合が減少する。加えて、仮想マシン 220 はローカルキャッシュ 225 から読み取りデータを取り込むことによって得られる待ち時間の減少による性能の改善が得られる可能性が有る。幾つかの実施形態において、プロブストア 230 への高い接続性（例えば 10 Gps）を有する仮想マシンはローカルキャッシュ 225 を使用しない可能性が有る。

10

#### 【0031】

プロブストア 230 はファブリックコンピュータ 210 に接続されている。ファブリックコンピュータ 210 は IP アドレスを介してプロブストア 230 にアクセスする。プロブストア 230 はプロブサーバー 233 を含んでいる。プロブサーバー 233 はファブリックコンピュータ 210 とインターフェースで繋がり、プロブへのアクセスを制御し、プロブ上のプロブ指令を実装する。プロブサーバー 233 は複製されたプロブストレージシステムの一部であり、そこでは格納されたプロブは幾つかのサーバーを介して複製され、ドライブ、ノード、又はラックにおいて故障があった場合にも別々に維持されるプロブのコピーが作られる。これによって故障が有った場合にもプロブが利用でき、耐久性を有することが確実化される。

20

#### 【0032】

プロブにアクセスしたときにドライバ 224、又はクラウドアプリケーション 221 から受取ったプロブ指令には、リースプロブ（Lease Blob）、スナップショットプロブ（Snapshot Blob）、プットページ（Put Page）、クリアページ（Clear Page）及びゲットプロブ（Get Blob）が含まれる。

#### 【0033】

リースプロブ指令はプロブストア 230 にプロブストア 230 上に格納されたプロブにリースを作成させる。幾つかの実施形態において、リースは排他的書き込みリースであり、一つのクラウドアプリケーション 221 のみにプロブに書き込むことを許す。排他的書き込みリースを取得するとき、リース識別子（ID）がプロブサーバー 233 によって作成され、クラウドアプリケーション 221 に返される。ドライバ 224 はリース ID を格納し、リース ID をプロブストア 230 に送出されたどのような書き込みと共に含める。リースプロブ指令はまた排他的読み出し書き込みリース及び他のタイプのリースをサポートする。幾つかの実施形態においてリースプロブ指令はドライバ 224 がリースの持続時間を指定することを許可する。リースを更新し又はリースを終了するためにドライバ 224 によってリースプロブ指令が使用されても良い。

30

#### 【0034】

スナップショットプロブ指令はプロブストアにプロブ内のデータのスナップショットを作成させる。スナップショットは読み出し専用である。ドライバ 224 はプロブの内容に同時にアクセスすることを可能にするためにこの指令を用いても良い。同一のプロブの 1 以上のスナップショットが 2 以上のクラウドアプリケーションによって同時にアクセスされても良い。スナップショットはプロブのためにバージョン化を提供しても良い。プロブのためのスナップショットはプロブの名前と（複数のバージョンの場合）プロブサーバー 233 によって自動作成されたバージョンタイムスタンプに基づいてアクセスされる。換言すれば、プロブが更新されたときにプロブの更新の前と後に生成されたスナップショットは異なるバージョンタイムスタンプを含む。或る実施形態において、クラウドアプリケーションはドライバ 224 にスナップショットのためにプロブストアにメタデータを送出させても良い。メタデータはスナップショットを記述するため又は瞬時にスナップショットの位置を決定するために使用されても良い。

40

50

## 【 0 0 3 5 】

プットページ指令はプロブサーバー 2 3 3 にプロブストア 2 3 0 内の指定されたアドレスにおけるページの範囲を格納させる。各ページはプライマリストレージ装置内のメモリーブロック範囲によって構成されても良い。プットページ指令はプロブストアにデータを書き込むためにドライバ 2 2 4 によって使用されても良い。プロブストア 2 3 0 内の有効なアドレス内に格納することを可能にするためにオフセットが指定されても良い。ページは順番に書き込まれる必要はなく、プロブのアドレス空間内でギャップ（例えば空のブロック）が有っても良い。例えばドライバ 2 2 4 はプロブストア 2 3 0 内で 4 K B のページをオフセット 0 でプットし、他のページをオフセット 4 K B \* 1 百万のオフセットでプットしても良い。プロブストア 2 3 0 はこれらの 2 つのページを指定されたアドレスに格納しても良い。これらのページはプットページ指令が成功した後にプロブストア 2 3 0 に深く関係付られる。一方、プロブストア 2 3 0 はドライバ 2 2 4 に成功したことのアクノリッジを行う。プットページ指令はプロブサーバー 2 3 3 にプロブを更新する前にリース I D をチェックさせる。プロブのために排他的書き込みリースが指定されたとき、ドライバによってプロブに更新を要求している間に有効なリース I D が指定されたときにプットページ指令は成功となる。リースが終了したとき、プットページ指令は不成功になり、プロブサーバー 2 3 3 は終了したリース I D を除去しても良い。

10

## 【 0 0 3 6 】

クリアページ指令はプロブサーバー 2 3 3 にプロブストア 2 3 0 から特定のページ又はページ範囲をクリアさせる。ドライバ 2 2 4 はクラウドアプリケーション 2 2 1 の仮想ハードドライブからデータを消去する要求に回答してプロブからページを除去するクリアページ指令を送出する。クリアページ指令はプロブサーバー 2 3 3 に消去要求の対象であるデータを有するページを除去する前に V H D に関係したプロブのために有効なリース I D が指定されていることを確認させる。リースが終了したとき、クリアページ指令は不成功になり、プロブサーバー 2 3 3 は終了したリース I D を除去しても良い。

20

## 【 0 0 3 7 】

ゲットプロブ指令はプロブサーバー 2 3 3 に指定されたプロブのためにプロブ全体又はページの範囲（例えばバイト）を取り込ませる。ドライバ 2 2 4 はクラウドアプリケーション 2 2 1 の仮想ハードドライブからデータを取得する要求に回答してゲットプロブ指令を送出する。ゲットプロブ指令はプロブサーバー 2 3 3 にページを取得する前に V H D に関連づけられたプロブに有効なリース I D が指定されていることを確認させる。リースが終了したとき、ゲットプロブ指令は不成功となり、プロブサーバー 2 3 3 は終了したリース I D を除去しても良い。

30

## 【 0 0 3 8 】

ドライバ 2 2 4 はクラウドアプリケーション 2 2 1 からドライブ指令及び I / O 要求を受け取る。一方、ドライバ 2 2 4 はクラウド指令及び I / O 要求をプロブ指令に翻訳する。ドライブ指令はクラウドアプリケーションによって V H D を管理するために使用され、I / O 要求は V H D に格納されたデータにアクセスするために用いられる。ドライブ指令にはとりわけクリエイイトドライブ（Create Drive）、マウントドライブ（Mount Drive）、アンマウントドライブ（UnMount Drive）、デリートドライブ（Delete Drive）、リストドライブ（List Drive）、及びスナップショットドライブ（Snapshot Drive）が含まれる。

40

## 【 0 0 3 9 】

クリエイイトドライブ指令はクラウドアプリケーション 2 2 1 によって仮想ハードドライブを作成するために使用される。クラウドアプリケーション 2 2 1 は仮想ハードドライブのためにサイズとフォーマットを指定する。例えば、クラウドアプリケーション 2 2 1 は仮想ハードドライブをウインドウズ（R）NTFS ファイルシステムを使用して単一パーティション単一ボリュームとしてフォーマットしても良い。一方、ドライバ 2 2 4 はドライブ指令をプロブ名前空間で利用できる適切なプロブ指令に翻訳する。ドライバは次にプロブストア 2 3 0 に、プロブ指令、例えば V H D にプロブを作成するプットページ、リー

50

スプロブ等を送出しても良い。

【 0 0 4 0 】

マウントドライブ指令はクラウドアプリケーション 2 2 1 によって V H D をマウントするために使用される。V H D をマウントするとき、クラウドアプリケーション 2 2 1 はマウントされた V H D のためのローカルキャッシュ 2 2 5 として使用するローカルディスクスペースの量を指定しても良い。クラウドアプリケーション 2 2 1 はまた排他的書き込み V H D、書き込み V H D、共用読み出し専用 V H D、等を要求しても良い。

【 0 0 4 1 】

排他的書き込み V H D はクラウドアプリケーション 2 2 1 のみが V H D を更新しても良いことを意味している。書き込み V H D はクラウドアプリケーション 2 2 1 が V H D を更新してもよく、また順次他のクラウドアプリケーションが V H D を更新しても良いことを意味している。共用読み出し専用 V H D は V H D が読み出し専用であり、他の仮想マシン 2 2 0 が同一の V H D から同時に読み出ししても良いことを意味している。ドライバ 2 2 4 はリース I D が V H D の対応したプロブに関連していないかも知れないため読み出し専用 V H D に書き込むことを阻止しても良い。一方、ドライバ 2 2 4 はマウントドライブ指令を適切なプロブ指令に翻訳する。例えば、ドライバ 2 2 4 プロブストア 2 3 0 に対して V H D に対応したプロブ上のリースを取得し維持するリースプロブ指令を送出しても良い。加えて、ドライバは V H D へのアクセスを有するクラウドアプリケーション 2 2 1 を提供し続けるためにプロブ上のリースを日常的に更新しても良い。もしリースが更新されなければ、クラウドアプリケーション書き込み要求 2 2 1 は不成功になる可能性がある。

【 0 0 4 2 】

アンマウントドライブ指令は指定された V H D をマウント解除するためにクラウドアプリケーション 2 2 1 によって用いられる。アンマウントドライブ指令はドライバ 2 2 4 に、指定された V H D に関連したプロブのためのリースを終了させることができる。ファブリックコンピュータ 2 1 0 又は仮想マシン 2 2 0 が故障したとき、ドライバ 2 2 4 はファブリックコンピュータ 2 1 0 又は仮想マシン 2 2 0 によって使用されているプロブのための如何なるリースも終了させるアンマウントドライブ指令を自動的に送しても良い。

【 0 0 4 3 】

デリートドライブ指令は指定された V H D を削除するためにクラウドアプリケーション 2 2 1 によって用いられる。デリートドライブ指令はドライバ 2 2 4 に指定された V H D に関連したプロブのためのリースを終了させるプロブ指令を送出させるようにしても良い。幾つかの実施形態において、ドライバ 2 2 4 は指定された V H D に関連したプロブのための全てのページを除去するクリアページ指令を送出しても良い。V H D を除去した後、プロブ内のデータはクラウドコンピューティングプラットフォーム又はクラウドアプリケーション 2 2 1 には利用できない。

【 0 0 4 4 】

リストドライブ指令はクラウドアプリケーション 2 2 1 に関連したすべての V H D をリストするためにクラウドアプリケーション 2 2 1 によって使用される。リストドライブ指令はドライバ 2 2 4 にクラウドアプリケーション 2 2 1 に関連した各リース I D の場所を検索させても良い。幾つかの実施形態において、ドライバ 2 2 4 は、クラウドアプリケーション 2 2 1 が利用可能な各リース I D に対応した文字又は数字をクラウドアプリケーション 2 2 1 送しても良い。ドライバ 2 2 4 はまた、スナップショット又はリース I D の無い他のドライブに関連した数字又は文字を受け取っても良い。

【 0 0 4 5 】

スナップショットドライブ指令はクラウドアプリケーション 2 2 1 に関連した各 V H D のスナップショットを得るためにクラウドアプリケーション 2 2 1 によって使用される。スナップショットドライブ指令はドライバ 2 2 4 に V H D に関連したプロブを検索させ、プロブストアにスナップショットプロブ指令を送出させても良い。一方、ドライバ 2 2 4 はクラウドアプリケーション 2 2 1 にスナップショットの文字、数字、又は例えば日付及び時刻などのタイムスタンプを返す。プロブのスナップショットはクラウドアプリケーシ

ョン 2 2 1 によって V H D としてアクセスされても良い。

【 0 0 4 6 】

或る実施形態において、クラウドアプリケーション 2 2 1 からの I / O 要求はドライバ 2 2 4 によって処理される。クラウドアプリケーション 2 2 1 によってマウントされた V H D は格納されたデータに I / O 要求を実行するためにアクセスされる。I / O 要求は、特に、読み出し及び書き込み要求を含んでも良い。

【 0 0 4 7 】

例えば、V H D はクラウドアプリケーション 2 2 1 から読み出し要求を受け取ることができる。I / O リダイレクション 2 2 3 が読み出し要求をドライバ 2 2 4 に迂回させるようにしても良い。一方、ドライバ 2 2 1 は読み出し要求をクラウドアプリケーション 2 2 1 から受け取った重なった前の書き込み要求とともに整理する。これによって、書き込み要求がプロブストア 2 3 0 内に格納された後にのみ、読み出しが処理中の更新を返すことが確実にされる。ドライバ 2 2 4 はバッファにおいて要求されたデータについて最近書き込まれたデータをチェックする。もしバッファが要求されたデータを含んでいなければ、ローカルキャッシュ 2 2 5 がチェックされる。データがローカルキャッシュ 2 2 5 で利用でき、ローカルキャッシュ 2 2 5 から読み出されたとき、ブロックのための C R C がチェックされて、データの完全性が確認される。もし C R C の不一致があれば、データがキャッシュから除去され、データはプロブストア 2 3 0 から取り出される。

【 0 0 4 8 】

もしローカルキャッシュ 2 2 5 においてデータが見つからなければ、ドライバ 2 2 4 はプロブストア 2 3 0 からデータを読み出す。プロブストア 2 3 0 からデータを読み出すためには、ドライバ 2 2 4 は要求されたデータに対応したページにアクセスするためにゲットプロブ指令を送出する。或る実施形態において、プロブストア 2 3 0 からのデータ読み出しは有効なリース I D を有することが条件とされる。データが読み出されて有効にされた後にデータはクラウドアプリケーション 2 2 1 に返される。データがプロブストア 2 3 0 から取り出されたとき、ドライバー 2 2 4 はローカルキャッシュ 2 2 5 のサイズ、キャッシュ置換ポリシー等に基づいてデータをキャッシュするか否かを決定する。もしデータがローカルキャッシュ 2 2 5 に格納すべきであれば、データが格納され、C R C コードが計算されて格納される。C R C はデータと共に格納することができ、又は後にファブリックコンピュータ 2 1 0 上の異なる位置に書き込んでも良い。

【 0 0 4 9 】

クラウドアプリケーション 2 2 1 は V H D にデータを書き込む。I / O リダイレクションは書き込み要求ドライバ 2 2 4 を迂回させても良い。一方、ドライバ 2 2 4 は入力された I / O 要求を処理中の読み出し及び書き込み要求に対して順序付けても良い。ドライバ 2 2 4 は書き込み要求をプロブストア 2 3 0 に送出的。一実施形態において、ドライバ 2 2 4 はプロブストアへの全ての重なった書き込み要求を整理し、重ならない書き込み要求のみをプロブストア 2 3 0 に送出的。処理中の重なった読み出し又は書き込み要求があるとき、入力された書き込み要求は整理され、重なった I / O 要求がプロブストア 2 3 0 に送出的されてアクノリッジが返されるまで待たねばならない。

【 0 0 5 0 】

重なった要求が整理された後、入力された書き込みがプロブストア 2 3 0 に送出的される。書き込みはマウントされた V H D に対応したプロブのための有効なリース I D に基づいた条件付きである。もしリースが終了していれば書き込み要求は失敗する。この場合、ドライバ 2 2 4 はリースの再取得を試みても良く、もしできなければ書き込みは依然として失敗する。書き込み要求が成功すると、書き込み要求をプロブストア 2 3 0 に格納すること及びクラウドアプリケーション 2 2 1 に成功のアクノリッジを返す前に書き込みを複製することによって耐久性がもたされる。

【 0 0 5 1 】

もし書き込み要求がタイムアウトしたとき、(例えば、プロブストア 2 3 0 からアクノリッジが受け取られなかったとき)、ドライバ 2 2 4 は書き込み要求を取り込む。ドライ

10

20

30

40

50

バ 2 2 4 はまた、プロブストアが「タイムアウト」又は「サーバービジー状態」を送出したときに書き込み要求を再試行しても良い。ドライバ 2 2 4 はプロブストア 2 3 0 において一旦再試行が成功すると成功を返す。他の実施形態において、ドライバ期限が過ぎた (state) 書き込み (例えば、プロブストア 2 3 0 に送出されたアクノリッジされていない書き込み要求) が後の再試行を繰り返さないことを確実にする。期限が過ぎた書き込みはプロブサーバータイムアウトメッセージを取得することを待つか、又はプロブサーバータイムアウト期間を待つことによって廃棄することができる。従って、重なった書き込み要求を書き込み再試行の範囲で実行する前に、ドライバ 2 2 4 はそれらのプロブサーバータイムアウト期間が経過するのを待つことによって、システムを介して期限が過ぎた書き込みがクリアされることを確実にする。プロブサーバー 2 3 3 は所定の時間が経過した後に期限が過ぎた書き込みを廃棄する。或いは、ドライバ 2 2 4 はプロブに関連したリース又はシーケンス番号をリセットしても良い。リース ID の変化によって期限が切れたリース ID を有する期限が過ぎた書き込みがプロブストア 2 3 0 を更新することを防止することもできる。任意選択で、プロブに関連したシーケンス番号は、各々の成功した書き込み要求の後に増加させても良い。ドライバ 2 2 4 がプロブストア 2 3 0 からアクノリッジを取得しないとき、ドライバ 2 2 4 はシーケンス番号を増加させてプロブストア 2 3 0 に若いシーケンス番号を有する前の書き込みは廃棄すべきであること通知する。従って、プロブストア 2 3 0 は期限が過ぎた書き込みを無視し、ドライバ 2 2 4 からの将来の書き込みは新しいリース ID 又はシーケンス番号を使用する。

10

#### 【 0 0 5 2 】

20

プロブサーバー 2 3 3 はプロブストア 2 3 0 内のプロブのためにシーケンス番号を格納する。シーケンス番号はプロブサーバー 2 3 3 によってプロブサーバーが期限が過ぎた書き込みを受け取ったときにプロブへのアクセスを拒否するために使用される。シーケンス番号を使用する際、各ページブット又はページクリアー指令はシーケンス番号を渡し、指令に含められたシーケンス番号はプロブストア内にプロブと共に格納されたシーケンス番号以上であるか否かがチェックされる。指令に含められたシーケンス番号がプロブストア内にプロブと共に格納されたシーケンス番号以上であるとき、指令が成功するようにすることができる。そうでなければ指令は失敗する。

#### 【 0 0 5 3 】

一実施形態において、VHD をマウントする際に対応したプロブのためのシーケンス番号は 0 に設定される。VHD 及びその対応したプロブへの全ての書き込みがシーケンス番号 0 に渡る。一方、プロブサーバー 2 3 3 は渡されたシーケンス番号がプロブのために格納されたものと一致するため書き込みを許容する。もし書き込みの一つが期限切れであるとき、ドライバ 2 2 4 は書き込みが実行の待ち行列内に有るか否かを知らない。例えば 10 秒のタイムアウト期間の後、ドライバ 2 2 4 は処理中の書き込みを有するプロブに関連したシーケンス番号を増加させる要求をプロブサーバー 2 3 3 に送出する。プロブサーバー 2 3 3 はプロブのシーケンス番号を 1 に増加させる。その後、ドライバ 2 2 4 はシーケンス番号が 1 の VHD のための書き込みをプロブサーバー 2 3 3 に送出し、プロブサーバーは受け取った書き込みをプロブ内で実現する。シーケンス番号が更新された後にシーケンス番号が 0 の期限が過ぎた書き込みがプロブサーバー 2 3 3 に到着した場合、渡されたシーケンス番号がプロブのために格納された番号より小であるため、プロブサーバー 2 3 3 は書き込みを廃棄する。

30

40

#### 【 0 0 5 4 】

一実施形態において、クラウドコンピューティングプラットフォームは VHD を作成したクラウドアプリケーションを実行する。VHD はクラウドアプリケーションによって使用されるデータを格納する。一方、VHD を代表するプロブストアは本来のクラウドアプリケーション指令及び I/O 要求を VHD に対応したプロブの為のプロブ指令として解釈するためにドライバによって管理される。

#### 【 0 0 5 5 】

図 3 は仮想ハードドライブ (VHD) として公開されるプロブを作成するための模範的

50

な方法を図示した論理図である。この方法ステップ310で初期化される。クラウドコンピューティングプラットフォームはステップ320においてVHDとして公開されるプロブをプロブストア内に作成する。一実施形態においてクラウドコンピューティングプラットフォームはVHDをN-T-F-Sファイルシステム又はF-A-Tファイルシステムの内の一つとしてフォーマットしても良い。一方、ステップ330においてVHDはクラウドコンピューティングプラットフォーム内の仮想マシン上で実行している1以上のアプリケーションによってマウントされる。ステップ340において、ドライバVHDへの読み出し及び書き込み要求を管理する。ドライバVHDへの全てのI/Oを遮り、I/Oをプロブストア内のプロブ宛に変更する。I/O要求内の書き込みはアクノリッジをアプリケーションに返す前にプロブストアに渡される。指定された期間の後に終了するプロブリースを介してクラウドアプリケーションのためのプロブに排他的書き込みアクセスが提供されても良い。ドライバはアプリケーションが動作していてそれを使用することを欲している限りVHDがマウントされ続けるためにリースを継続的に更新するスレッドを実行しても良い。他の実施形態において、プロブに対する成功したアップデートが存在する毎にリースが更新するようにしても良い。一方、クラウドアプリケーションは他のクラウドアプリケーションがプロブからの読み出しを許可されている間、リースの有効性に基いて条件的にプロブへの書き込みを行う。

#### 【0056】

プロブはドライバに利用できるプロブインターフェースを介してアクセスされ、操作される。プロブのスナップショットが追加のアプリケーションがプロブに格納されたデータを同時に読み出すことを許可するためにプロブインターフェースを介して作成されても良い。或る実施形態において、プロブはまた、クラウドコンピューティングプラットフォームによって実行されている幾つかのクラウドアプリケーションが利用可能なプロブインターフェースを介して、アップロードダウンロード、又はデバッグされても良い。クラウドアプリケーションはデータの整合性のある格納についてチェックするためにデバッガーを実行しても良い。ステップ350においてこの方法は終了する。

#### 【0057】

幾つかの実施形態において、ドライバはクラウドアプリケーションのためのI/O要求を管理する。ドライバはVHDに対するI/O要求をプロブストア内の対応したプロブに送出される適当なプロブ指令に翻訳する。一方、プロブストアはプロブ指令を実現するためにプロブにアクセスする。

#### 【0058】

図4は仮想ハードドライブへのI/O要求を管理する模範的な方法を図示した論理図である。この方法はステップ410において初期化する。ステップ420において、クラウドコンピューティングプラットフォーム内の仮想マシン上で実行しているクラウドアプリケーションからI/O要求がVHDのために受け取られる。一実施形態において、ドライバは以前の書き込み要求が同一データへの最近の書き込み要求で上書きされるのを防止するためにシーケンス番号をI/O要求に関連づける。シーケンス番号はまた、期限が過ぎた書き込みが最近の書き込み要求を上書きすることを防止する。幾つかの実施形態において、ドライバはプロブのためにシーケンス番号を増加させ、シーケンス番号がプロブ指令に渡されて以前の書き込み要求がプロブストアによってアクノリッジされていないとき古いシーケンスを有する以前の書き込み要求を無視するようにしても良い。ステップ430において、ドライバがVHDへのI/O要求を管理する。ドライバはプロブストア内のプロブへの全ての重なったI/O要求を整理し、重ならないI/O要求をプロブストア内のVHDに対応したプロブに並行して送出する。他の実施形態において、仮想マシン上のローカルキャッシュは書き込み要求に応じてドライバにより更新され、ローカルキャッシュの更新はアクノリッジがプロブストアによって受け取られた後又はプロブストアへの更新と並行して行われる。ローカルキャッシュ内の更新はアクノリッジがプロブストアから受信された後にのみ利用できる。クラウドアプリケーション又は仮想マシンを再起動させたときにキャッシュとプロブが不一致になったとき、ローカルキャッシュの全ての部分はク

10

20

30

40

50

ラウドアプリケーションの故障又は仮想マシンの故障の後にドライバによってクリアされても良い。ステップ440においてこの方法が終了する。

【0059】

以上を要約すると、クラウドコンピューティングプラットフォームはプロブストア内のプロブを使用してVHDを提供する。例えば、クラウドコンピューティングプラットフォーム上でデータベースアプリケーションを実行することができる。データベースアプリケーションにおいて、時系列的にログが書き込まれる。ログは、例えばやり直し(Redo)/取り消し(Undo)記録、BTree変更記録、及びコミット記録等のかなり多数の書き込みについて用いられる。専用の「ログドライブ」がログアクセスとデータアクセスの競合を防止する。専用の「データベースドライブ」がデータベース内のデータにアクセス(書き込み及び読み出し)するために用いられる。「ログドライブ」の読み出しの割合は非常に低く、データベースドライブが読み出し及び書き込みの両方について高い割合を有している。

10

【0060】

クラウドコンピューティングプラットフォームを実行するデータベースアプリケーションするは2つのVHD:「ログドライブ」と「データベースドライブ」を作成しても良い。一方、プロブストアは「ログドライブ」及び「データベースドライブ」に夫々対応する2つのプロブを作成する。データベースアプリケーションは「ログドライブ」を"マウントし、トラフィックの多数はログ書き込みであるためキャッシングを設定しない。一方、クラウドアプリケーションは「データベースドライブ」をマウントし、100%のキャッシングを設定し、読み出し要求の殆どがローカルキャッシュによって出されるようにする。

20

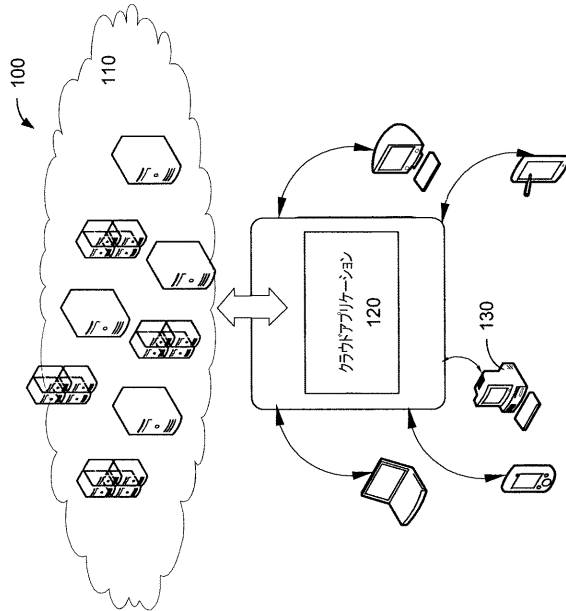
【0061】

以上の発明の実施形態の記述は説明的なものであり、構成及び実装の変更はこの記述の範囲内のものである。例えば、発明の実施形態は図1-4に関連して記述されてきたけれども、それらの記述は、例示的なものである。主題は構造的な特徴又は方法論的な作用に固有の言語によって記述されてきたけれども、添付の請求項によって画定された主題は必ずしも上述の特定の特徴又は作用に限定されないことが理解される。寧ろ上述の特定の特徴及び作用は請求の範囲を実現する形式の例として開示されたものである。従って、発明の実施形態の範囲は以下の請求項によってのみ限定される。

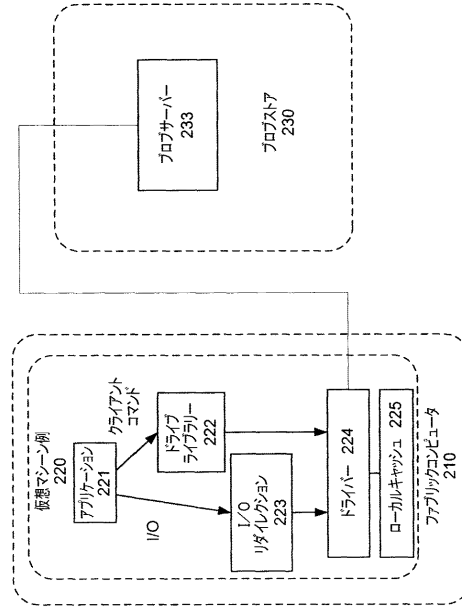
30



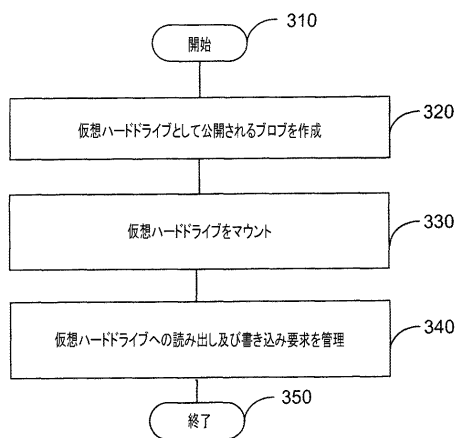
【図 1】



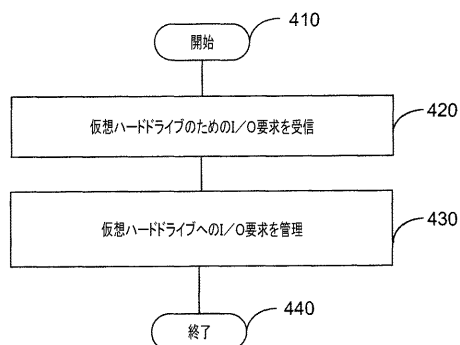
【図 2】



【図 3】



【図 4】



## フロントページの続き

- (72)発明者 ブラッドリー ジーン カルダー  
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ  
イクロソフト コーポレーション エルシーエー - インターナショナル パテンツ内
- (72)発明者 アンドリュー ジェームス エドワーズ  
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ  
イクロソフト コーポレーション エルシーエー - インターナショナル パテンツ内
- (72)発明者 ジュー ワン  
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ  
イクロソフト コーポレーション エルシーエー - インターナショナル パテンツ内
- (72)発明者 サメール アラフェーフ  
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ  
イクロソフト コーポレーション エルシーエー - インターナショナル パテンツ内
- (72)発明者 アヌ エンジンニア  
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ  
イクロソフト コーポレーション エルシーエー - インターナショナル パテンツ内
- (72)発明者 ユエ ズオ  
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ  
イクロソフト コーポレーション エルシーエー - インターナショナル パテンツ内

審査官 原 秀人

- (56)参考文献 特表2003-527659(JP,A)  
特開2003-256256(JP,A)  
特開2005-070965(JP,A)  
特開2007-219609(JP,A)  
特開2009-134601(JP,A)  
特開平10-124437(JP,A)  
特開2002-222110(JP,A)  
特開2005-346426(JP,A)  
特開平08-255103(JP,A)  
国際公開第2009/111495(WO,A1)  
松岡 清一, マイクロソフトのクラウドプラットフォーム Windows Azureの衝撃  
とマイクロソフトの戦略, 月刊アスキーdotテクノロジー, 日本, 株式会社アスキー・メデ  
ィアワークス, 2009年7月1日, Vol.14 #7 (通巻159号), p.72-7  
9  
田中 慎司, 1000万ユーザーを支える裏側を大公開! はてなに学ぶWeb-DBシステム  
仮想化の実践, DB Magazine, 日本, 株式会社翔泳社, 2009年10月27日, 第  
19巻 第8号, pp.39-42

## (58)調査した分野(Int.Cl., DB名)

G06F 12/00  
G06F 13/10  
G06F 9/46