

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5762643号
(P5762643)

(45) 発行日 平成27年8月12日(2015.8.12)

(24) 登録日 平成27年6月19日(2015.6.19)

(51) Int.Cl.	F 1
G06F 12/00	(2006.01) G06F 12/00 545B
G06F 13/10	(2006.01) G06F 13/10 340A
G06F 3/06	(2006.01) G06F 3/06 301J
G06F 9/46	(2006.01) G06F 12/00 501A G06F 9/46 350

請求項の数 14 (全 17 頁)

(21) 出願番号	特願2014-539947 (P2014-539947)	(73) 特許権者	514112215 ザダーラ ストレージ リミテッド イスラエル ネッシャー バー イエフダ ストリート 63
(86) (22) 出願日	平成24年8月30日(2012.8.30)	(74) 代理人	100092093 弁理士 辻居 幸一
(65) 公表番号	特表2015-501494 (P2015-501494A)	(74) 代理人	100082005 弁理士 熊倉 賢男
(43) 公表日	平成27年1月15日(2015.1.15)	(74) 代理人	100067013 弁理士 大塚 文昭
(86) 國際出願番号	PCT/US2012/052949	(74) 代理人	100086771 弁理士 西島 孝喜
(87) 國際公開番号	W02013/066482	(74) 代理人	100109070 弁理士 須田 洋之
(87) 國際公開日	平成25年5月10日(2013.5.10)		
審査請求日	平成26年12月1日(2014.12.1)		
(31) 優先権主張番号	13/290,084		
(32) 優先日	平成23年11月5日(2011.11.5)		
(33) 優先権主張国	米国(US)		
早期審査対象出願			

最終頁に続く

(54) 【発明の名称】クラウドサーバーのためのバーチャルプライベートストレージアレイサービス

(57) 【特許請求の範囲】

【請求項 1】

コンピュータネットワークを経てバーチャルプライベートストレージアレイをサービスとして提供するための方法において、

コンピュータネットワークを経てバーチャルプライベートストレージアレイのためのパラメータを受け取る段階を含み、前記バーチャルプライベートストレージアレイのためのパラメータは、前記バーチャルプライベートストレージアレイの各バーチャルコントローラのためのバーチャルコントローラハードウェアモデルと、前記バーチャルプライベートストレージアレイのためのドライブ特性と、前記バーチャルプライベートストレージアレイのためのドライブ量とを含むものであり；及び

プロセッサ/メモリ複合体、並びにサーバーコンピュータから利用できる物理的ドライブから前記バーチャルプライベートストレージアレイを生成する段階を含み、その各サーバーコンピュータは、ストレージノード及び計算エージェントの少なくとも1つを含むソフトウェアを実行するものであり、前記バーチャルプライベートストレージアレイを生成する段階は、

前記ドライブ特性及びドライブ量を満足するか又はそれを越える1組の物理的ドライブを選択することを含み、その選択された物理的ドライブは、1組のサーバーコンピュータからのものであり、更に、

前記1組のサーバーコンピュータ上のストレージノードに、前記バーチャルプライベートストレージアレイにバーチャルドライブを割り当てるよう命令することを含み、前記

ストレージノードは、

前記選択された物理的ドライブから前記バーチャルドライブを生成し、各バーチャルドライブは、1つの選択された物理的ドライブ全体であるか又は1つの選択された物理的ドライブの一部分である区画であり、及び

前記バーチャルプライベートストレージアレイのバーチャルコントローラに前記バーチャルドライブを露出する、

ように構成され；更に、

前記バーチャルコントローラハードウェアモデルを各々満足するか又はそれを越える1組のプロセッサ/メモリ複合体を選択することを含み、その選択されたプロセッサ/メモリ複合体は、他の1組のサーバーコンピュータからのものであり；更に、

10

前記他の1組のサーバーコンピュータ上の計算エージェントに、前記バーチャルコントローラのためのバーチャルマシンをスポーンするよう命令することを含み、その計算エージェントは、

前記バーチャルマシン専用の各選択されたプロセッサ/メモリ複合体の少なくとも一部分上の1つのバーチャルマシンをスポーンし、及び

前記バーチャルマシン当たり1つのバーチャルコントローラを始動して、前記バーチャルプライベートストレージアレイの各バーチャルコントローラが、異なるサーバーコンピュータにおいて実行されるようとする、

よう構成され、前記バーチャルコントローラの1つ以上は、

前記露出されたバーチャルドライブを発見し、

20

前記露出されたバーチャルドライブを含む1つ以上のバーチャルプールを生成し、前記1つ以上のバーチャルプールにおいてデータ保護スキームを実施し、

前記1つ以上のバーチャルプールから1つ以上のボリュームを生成し、

前記コンピュータネットワークを経て1つ以上のカスタマーコンピュータへ前記1つ以上のボリュームを露出させ、及び

前記1つ以上のカスタマーコンピュータから前記コンピュータネットワークを経て前記露出された1つ以上のボリュームへアクセス要求をハンドリングする、
ように構成される、方法。

【請求項2】

ウェブサーバが前記コンピュータネットワークを経てウェブフォームを送信し、ウェブフォームは、前記パラメータを返送するように構成され、

30

前記ウェブサーバが前記コンピュータネットワークを経てウェブページを送信し、当該ウェブページは、サービス料金を含み、当該ウェブページは、前記バーチャルプライベートストレージアレイを生成するための確認を返送するように構成され、及び

前記ウェブサーバが前記コンピュータネットワークを経て前記確認を受け取る、
ことを更に含む、請求項1に記載の方法。

【請求項3】

前記バーチャルコントローラハードウェアモデルは、中央処理ユニットモデル、中央処理ユニット特徴、中央処理ユニット量、ランダムアクセスメモリ容量、及びネットワーク帯域幅、の少なくとも1つを明示し、及び

40

前記ドライブ特性は、ドライブ形式、ドライブ容量、及びドライブ暗号、の少なくとも1つを明示する、請求項1に記載の方法。

【請求項4】

各バーチャルドライブは、物理的ドライブ全体又は物理的ドライブの一部分である区画から生成された論理的ボリュームであり、及び

前記ストレージノードは、前記バーチャルドライブをネットワークブロック装置として前記バーチャルコントローラに露出させることにより前記バーチャルドライブを前記バーチャルコントローラに露出するように構成される、請求項1に記載の方法。

【請求項5】

前記パラメータは、更に、前記バーチャルプライベートストレージアレイに対するデー

50

タ保護の R A I D レベル、及び前記バーチャルプライベートストレージアレイにおけるバーチャルコントローラ量の一方を含む、請求項 1 に記載の方法。

【請求項 6】

前記ストレージノードは、更に、
前記物理的ドライブを発見し；
前記物理的ドライブにドライブ形式及びドライブ容量について問合せし；及び
前記ドライブ形式及びドライブ容量を利用ゾーンコントローラへ送信する；
よう構成され、前記 1 組の物理的ドライブを選択することは、
前記ドライブ特性を満足するか又はそれを越える物理的ドライブを伴う候補サーバーコンピュータのリストを生成し；
1 つ以上の分類基準に基づいて前記リストを分類し；及び
トップランクの候補サーバーコンピュータの物理的ドライブを選択する；
ことを含む、請求項 1 に記載の方法。

【請求項 7】

前記 1 つ以上の分類基準は、候補サーバーコンピュータの利用率を含む、請求項 6 に記載の方法。

【請求項 8】

前記 1 組のプロセッサ / メモリ複合体を選択することは、
前記サーバーコンピュータにプロセッサ / メモリ複合体について問合せし、
前記バーチャルコントローラハードウェアモデルを満足するか又はそれを越えるプロセッサ / メモリ複合体を伴う候補サーバーコンピュータのリストを生成し、
1 つ以上の分類基準に基づいて前記リストを分類し、及び
トップランクの候補サーバーコンピュータのプロセッサ / メモリ複合体を選択する、
ことを含む、請求項 1 に記載の方法。

【請求項 9】

前記 1 つ以上の分類基準は、候補サーバーコンピュータの利用率を含む、請求項 8 に記載の方法。

【請求項 10】

前記バーチャルコントローラは、負荷分担及び冗長性のために構成される、請求項 1 に記載の方法。

【請求項 11】

前記 1 組のサーバーコンピュータは、前記サーバーコンピュータの一部分又は全部を含む、請求項 1 に記載の方法。

【請求項 12】

前記 1 組のサーバーコンピュータ及び前記他の 1 組のサーバーコンピュータは、オーバーラップする、請求項 1 に記載の方法。

【請求項 13】

前記物理的ドライブは、磁気ハードディスクドライブ、ソリッドステートドライブ及び相変化メモリドライブの 1 つ以上を含む、請求項 1 に記載の方法。

【請求項 14】

1 つ以上の前記バーチャルコントローラは、データプライバシーを保証するためにアクセス要求を認証することによりアクセス要求をハンドリングするように構成される、請求項 1 に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、データストレージシステムに関するもので、より特定すれば、プライベート及びパブリッククラウドのユーザへのサービスとしてバーチャルストレージアレイをプロビジョニングするためのプラットホーム及び技術に関する。

【背景技術】

10

20

30

40

50

【0002】

既存のデータストレージアレイは、ストレージ及びリソースのプロビジョニングに関して融通性が不充分であり、且つそれらは、マルチテナントではなく、又、ストレージシステムのサブセットに対してそれらをクラウド環境に配備できるようにする性能を保証するものではない。更に、ストレージアレイは、プロビジョニングすることができず、ユーザコントロール型のサービスとしてクライドユーザに提供される。従って、必要なレベルのデータプライバシー、欠陥分離、及び予想し得る性能を慣習的なストレージシステムとして提供しながら、サービスとして構築し利用できる融通性のあるバーチャルストレージアレイが要望される。

【発明の概要】

10

【0003】

本発明の1つ以上の実施形態において、コンピュータネットワークを経てバーチャルプライベートストレージアレイ（VPSA）をサービスとして提供する方法は、ネットワークを経てVPSAのためのパラメータを受け取り、そしてサーバーコンピュータのリソースからVPSAを生成することを含む。VPSAを生成することは、指定のドライブ特性、ドライブ量、及びアレイ冗長性基準を満足するか又はそれを越えるドライブを、VPSAのバーチャルコントローラ（VC）に割り当て及び露出させ、そして指定のバーチャルコントローラハードウェアモデルを各々満足するか又はそれを越えるプロセッサ/メモリ複合体の各部分をVC専用とすることを含む。VCは、独立したサーバーコンピュータ上のプロセッサ/メモリ複合体の専用部分におけるバーチャルマシンにおいて実行される。VCは、露出されたドライブを発見し、露出された物理的ドライブからバーチャルプールを生成し、バーチャルプールにおいてデータ保護を実施し、バーチャルプールからボリュームを生成し、ネットワークを経てカスタマーコンピュータ（例えば、プライベート又はパブリッククラウドにおいて実行されるカスタマーアプリケーションサーバー）へそのボリュームを露出させ、そしてカスタマーコンピュータからボリュームへのアクセス要求をハンドリングする。各VPSAは、専用のリソース（例えば、中央処理ユニット、ランダムアクセスメモリ、ネットワークインターフェイスカード、及びディスクドライブ）と、クラウドのユーザによりコントロールされる専用のマネージメントグラフィックユーザインターフェイスと、を有する。このように、クラウドの各ユーザにストレージの一貫した性能、セキュリティ、及びコントロールを与えることができる。

20

【図面の簡単な説明】

【0004】

30

【図1】クラウドにおけるサービスとしてバーチャルプライベートストレージアレイ（VPSA）を動的にプロビジョニングする規範的なソフトウェアシステムのブロック図である。

【図2】図1のソフトウェアシステムを実施するための規範的なハードウェアシステムのブロック図である。

【図3】図2の規範的なストレージノードコンピュータのブロック図である。

【図4】図1のソフトウェアシステムを実施するための規範的な別のハードウェアシステムのブロック図である。

40

【図5】図1のシステムで新たなVPSAをスローンするための規範的な方法のフローチャートである。

【図6A】図1のシステムでVPSAにおけるバーチャルコントローラ（VC）にバーチャルドライブを割り当てるための規範的な方法のフローチャートである。

【図6B】図1のシステムでVPSAにおけるバーチャルコントローラ（VC）にバーチャルドライブを割り当てるための規範的な方法のフローチャートである。

【図7A】図1のシステムで、本発明の実施形態に基づいて全て構成されたVCをVPSAに生成するための規範的な方法のフローチャートである。

【図7B】図1のシステムで、本発明の実施形態に基づいて全て構成されたVCをVPSAに生成するための規範的な方法のフローチャートである。

50

【発明を実施するための形態】**【0005】**

種々の図面において同じ又は同様の要素は同じ参照番号を使用して示されている。

【0006】

図1は、本発明の1つ以上の実施形態において、クラウドにおけるサービスとしてバーチャルプライベートストレージアレイ(VPSA)を動的にプロビジョニングし管理する規範的なソフトウェアシステム100のブロック図である。このシステム100は、標準的なサーバーコンピュータ及びデータストレージ媒体環境からVPSAを生成し、プロビジョニングしそして管理することができる。システム100は、コンピュータネットワーク102を経て、カスタマーコンピュータで実行されるアプリケーションへ、例えば、別のサービスプロバイダーによりカスタマーに提供されるクラウドサーバーへVPSAを接続する。クラウドサーバーとは、プロバイダーによりカスタマーへサービスとして提供されるバーチャルサーバーである。サービスプロバイダーは、ストレージを付加的なサービスとして提供するが、カスタマーは、提供されるストレージの構成を何らコントロールすることができない。

【0007】

VPSAは、1つ以上のバーチャルコントローラ(VC)104、及びこのVC104に露出される1つ以上のバーチャルドライブ106から構成される。バーチャルドライブ106は、全物理的ドライブ108或いは物理的ドライブ108の一部分である区画から生成された論理的ボリュームである。VC104及び物理的ドライブ108は、利用ゾーン(AZ)を作り上げる標準的サーバーコンピュータ間に分散される。

【0008】

VPSAにおいて、VC104は、とりわけ、バーチャルドライブ106から1つ以上のバーチャルプールを生成し、バーチャルドライブ106のバーチャルプールにおいてデータ保護を実施し、バーチャルプールにおいて同じデータ保護を行うようにしてバーチャルプールから1つ以上のバーチャルボリュームをカーブアウトし、ターゲットドライブを経て1つ以上のカスタマーコンピュータ112へバーチャルボリューム110をエクスポートし、そしてカスタマーコンピュータ112(例えば、クラウドサーバー)からバーチャルボリューム110へ標準的入力/出力(I/O)要求をハンドリングするという役割を果たすソフトウェアコンポーネントである。データ保護は、独立ディスク(RAID)スキームの冗長アレイである。I/O要求は、インターネット小型コンピュータシステムインターフェイス(ISC SI)又はインフィニバンド(Infiband)I/O要求である。VC104は、認証メカニズムを実施して、VPSAへの許可されたI/Oアクセスを検証する。例えば、VC104は、I/O要求に埋め込まれるユーザ名及びパスワードのようなクレデンシャルを検証する。VC104は、VPSAのバーチャル及び物理的エンティティの持続性データベースを維持し、VPSAのためのマネージメントインターフェイス124をカスタマーに与え、そしてVPSAのための監視及び診断ツールをカスタマーに与える。VC104によりサービスされるウェブ形態のようなマネージメントインターフェイス124を使用して、カスタマーは、クラウドサーバー及びアプリケーションへのVPSAストレージ及びボリュームプロビジョニングを操作することができる。各VC104は、専用メモリ、処理及びネットワークリソースを伴う個別のバーチャルマシン(VM)において実行される。

【0009】

物理的ドライブ108は、ハードディスクドライブ(HDD)、ソリッドステートドライブ(SSD)、相変化メモリ(PCM)ドライブ、及び他の形式の持続性ストレージドライブの任意の組み合わせを含む。物理的ドライブ108は、AZにおける標準的サーバーコンピュータ間に分散されたストレージノード(SN)114にアタッチされる。SN114は、とりわけ、物理的ドライブ108にそれらのドライブ特性について問合せし、同様のドライブ特性をもつ物理的ドライブ108をサービス品質(QoS)グループへとグループ化し、そしてドライブインベントリーを利用ゾーンコントローラ(AZC)11

6に報告するという役割を果たす（例えば、QoSグループAには6つのドライブ及びQoSグループBには2つのドライブ）。SN114は、追加及び除去される物理的ドライブ108を動的に発見し、そしてAZC116を現在ドライブインベントリーで更新する。

【0010】

SN114は、物理的ドライブ108を区画化し、そしてAZC116からのインストラクションごとに区画からバーチャルドライブ106を生成する。上述したように、バーチャルドライブ106は、全物理的ドライブ108又は物理的ドライブ108の一部分である区画から生成された論理的ボリュームである。全物理的ドライブ108からバーチャルドライブ106を生成することは、物理的ドライブベースでの完全なデータ分離及びプライバシー、並びに物理的ドライブ自己暗号化能力のレバレッジを含めて、クラウドユーザに利益を与える。VC104及び完全な物理的ドライブ108は、単一カスタマーのコントロールのもとにあるので、カスタマーは、VPSAサービスプロバイダー又はVPSAサービスプロバイダーの他のカスタマーと共有されないプライベートな暗号キーを使用する。

【0011】

SN114は、ターゲットドライバーを経てVC104へバーチャルドライブ106を露出させて、VC104からバーチャルドライブ106への標準的I/O要求（例えば、SCSI又はインフィニバンド）を許す。又、SN114は、各バーチャルドライブ内に小さなセットアップ区画109を生成する。以下に述べるように、VC104は、セットアップ区画を使用してVPSAのメタデータを記憶する。又、SN114は、I/O計測及びエラー統計値を収集し及び維持し、VC104のインターフェイスを与えて、バーチャルドライブ106のドライブ特性及びドライブ統計値を問合せ、AZC116のインターフェイスを与えて、物理的ドライブ108のドライブインベントリー、ドライブ特性、及びドライブ統計値を問合せる。

【0012】

AZC116は、とりわけ、カスタマーからのサービス要求に基づいてAZのサーバーコンピュータにVCを生成及び配置するという役割を果たす。VPSAを生成するとき、AZC116は、サーバーコンピュータにおける利用可能な中央処理ユニット(CPU)及びランダムアクセスメモリ(RAM)、特殊なネットワークインターフェイスカード(NIC)アダプタ(SR-I0VイネーブルNICのような)の存在、及びI/O負荷バランスのような種々のネットワークトポロジー及びリソース制約を考慮する。又、AZC116は、もはや必要でないか又は支払が済んだVC104の削除も行う。AZC116は、カスタマーからのサービス要求を満足させるためにサーバーコンピュータからバーチャルドライブ106を割り当てる。AZC116は、物理的ドライブ108のドライブ特性と、サービス要求を満足させるために物理的ドライブ108がアタッチされるSN114とを考慮する。又、AZC116は、VC104がコントロール及びデータ通信のために当該SN114と通信でき、同じVPSAのVC104がVCクラスターマネージメントのために互いに通信でき、且つカスタマーがコントロール及びデータ通信のためにVC104と通信してVPSAへのアクセスのプライバシーを保証するための認証を行えるように、AZにおいてネットワークを構成する。

【0013】

ウェブサーバー118は、ウェブフォーム120をカスタマーコンピュータ112の1つへ送信する。ウェブフォーム120は、新たなVPSAのサービス要求のパラメータをカスタマーから返送するように構成される。それらのパラメータは、(1)VPSAにおける各VCのVCハードウェアモデル、(2)VPSAのドライブ特性、及び(3)VPSAのドライブ量、を明示する。VCハードウェアモデルは、CPUモデル、1つ以上のCPU特徴、CPU量、RAM容量、及びネットワーク帯域幅を明示する。ドライブ特性は、ドライブ形式(HDD、SSD又はPCM)、ドライブ容量、ドライブ暗号化、及びドライブインターフェイス（例えば、SCSI又はインフィニバンド）を明示する。それ

10

20

30

40

50

らのパラメータは、更に、VPSAへのI/Oアクセスを認証するためのユーザ名及びパスワードのようなクレデンシャルを含み、それらは、後で、AZC116によりVPSAのVC104へ送られる。

【0014】

それらのパラメータに基づいて、ウェブサーバー118は、VPSAに対してサービス料金を伴うウェブページ122をカスタマーコンピュータ112の1つへ送信する。サービス料金は、単位時間当たりのコストである。ウェブページ122は、カスタマーからVPSAを生成するための確認を返送するように構成される。ウェブサーバー118は、それが確認を受け取ると、パラメータをAZC116へ送信する。

【0015】

マルチテナント性を説明するために、システム100は、VPSA126及びVPSA128を含むように示されている。VPSA126は、VC104-1、104-2と、物理的ドライブ108-1、108-2、108-3から発生されたバーチャルドライブ106-1、106-2、106-3とを含む。物理的ドライブ108-1、108-2、108-3は、SN114-1、114-2及び114-3の間に分散される。VC104-1及び104-2は、バーチャルドライブ106-1、106-2、106-3から発生されたデータボリューム110-1を、コンピュータネットワーク102を経て、カスタマーコンピュータ112へ露出させる。

【0016】

VPSA128は、VC104-3、104-4と、物理的ドライブ108-4、108-5から発生されたバーチャルドライブ106-4、106-5とを含む。物理的ドライブ108-4及び108-5は、SN114-2及び114-3の間に分散される。VC104-3及び104-4は、バーチャルドライブ106-4及び106-5から発生されたデータボリューム110-2を、コンピュータネットワーク102を経て、カスタマーコンピュータ127へ露出させる。

【0017】

VPSAの冗長性のために、VCは、異なるサーバーコンピュータに配置され、そして異なるサーバーコンピュータに配置されたSNに物理的ドライブがアタッチされる。しかしながら、1つのサーバーコンピュータは、同じVPSAからのVC及びSNと、異なるVPSAからのVCとを実行する。性能を高め又はコストを下げるために、物理的ドライブは、同じSNにアタッチされる。

【0018】

図2は、本発明の1つ以上の実施形態においてソフトウェアシステム100(図1)を実施するための規範的なハードウェアシステム200のブロック図である。システム200において、SN114(図1)は、以下「SNコンピュータ」と称されるサーバーコンピュータ214に位置される。SNコンピュータ214におけるSN114が充分なハードウェアリソースを有するよう保証するために、SNコンピュータ214におけるVC104の数が制限される。システム200は、SNコンピュータ214-1、214-2、214-3、任意のAZCコンピュータ216、及びウェブサーバーコンピュータ218を備えている。SNコンピュータ214-1、214-2及び214-3は、VPSA126及び128(図1)を構成するためにプロセッサ/メモリ複合体、NIC、及び物理的ドライブの物理的プールを形成する。AZCコンピュータ216は、AZC116(図1)を実行して、VPSAのVC104及びバーチャルドライブ106をSNコンピュータ214-1、214-2及び214-3間に分散させる。例えば、VC104-1及びSN114-1(図1)は、物理的ドライブ108-1と共にSNコンピュータ214-1に配置され、VC104-2、104-3及びSN114-2(図1)は、物理的ドライブ108-2及び108-4と共にSNコンピュータ214-2に配置され、そしてVC104-4及びSN114-3(図1)は、物理的ドライブ108-3及び108-5と共にSNコンピュータ214-3に配置される。或いは又、AZC116は、専用のAZCコンピュータ216ではなく、SNコンピュータ214-1、214-2及び214

10

20

30

40

50

- 3 の 1 つにおいて実行される。

【 0 0 1 9 】

ウェブサーバーコンピュータ 218 は、ウェブサーバー 118 (図 1) を実行して、ウェブフォーム 120 及びウェブページ 122 を生成する。SN コンピュータ 214 及びウェブサーバー 218 は、1 つ以上のパブリックスイッチ 232 によってコンピュータネットワーク 102 へ接続される。SN コンピュータ 214、AZC コンピュータ 216 及びウェブサーバー 218 は、1 つ以上のプライベートスイッチ 234 及び 236 により互いに接続される。

【 0 0 2 0 】

図 3 は、本発明の 1 つ以上の実施形態における規範的な SN コンピュータ 214-2 の 10 ブロック図である。SN コンピュータ 214-2 のハードウェアは、CPU 304 及び RAM 306 を伴うプロセッサ / メモリ複合体 302 と、NIC 308 と、物理的ドライブ 108-2 及び 108-4 とを含む。SN コンピュータ 214-2 のソフトウェアは、SN 114-2 と、VM 310-2、310-3 で実行される VC 104-2、104-3 とを含む、VM は、次いで、ハイパーバイザ 312 で実行される。VM 310-2 及び 310-3 は、SN コンピュータ 214-2 の専用及び / 又は共有ハードウェアから生成されたバーチャル CPU、RAM、及び NIC を有する。SN コンピュータ 214-2 のソフトウェアは、更に、専用の CPU、RAM 及び NIC で VM 310-2 及び 310-3 をスプーンし且つ VM 310-2、VM 310-3 において VC 104-2、104-3 をスタートする計算エージェント 314 を含む。例えば、計算エージェント 314 は、専用の CPU 及び RAM から生成されたバーチャル CPU 及び RAM を伴うが NIC からの利用可能なネットワーク帯域幅の一部分で生成されたバーチャル NIC (VNIC) も伴う VM を生成する。

【 0 0 2 1 】

図 4 は、本発明の 1 つ以上の実施形態においてソフトウェアシステム 100 (図 1) を実施するための規範的な別のハードウェアシステム 400 のブロック図である。システム 400 において、VC 104 (図 1) は、以下「計算ノードコンピュータ」又は「CN コンピュータ」412 と称されるサーバーコンピュータに位置され、そして SN 114 (図 1) は、SN コンピュータ 414 に位置される。システム 400 は、CN コンピュータ 412-1、412-2、412-3、SN コンピュータ 414-1、414-2、414-3、任意の AZC コンピュータ 216、及びウェブサーバーコンピュータ 218 を備えている。

【 0 0 2 2 】

CN コンピュータ 412 は、VC 104 を実施するためにプロセッサ / メモリ複合体及び NIC の物理的プールを形成し、そして SN コンピュータ 414 は、物理的ドライブ 108 の物理的プールを形成する。例えば、VC 104-1 (図 1) は、CN コンピュータ 412-1 に配置され、VC 104-2、104-3 (図 1) は、CN コンピュータ 412-2 に配置され、そして VC 104-4 (図 1) は、CN コンピュータ 412-3 に配置される。SN 114-1 (図 1) は、物理的ドライブ 108-1 と共に SN コンピュータ 414-1 に配置され、SN 114-2 (図 1) は、物理的ドライブ 108-2 及び 108-4 と共に SN コンピュータ 414-2 に配置され、そして SN 114-3 (図 1) は、物理的ドライブ 108-3 及び 108-5 と共に SN コンピュータ 414-3 に配置される。

【 0 0 2 3 】

各 CN コンピュータ 412 は、図 3 の SN コンピュータ 214-2 と同様に実施されるが、ストレージノード 114-2 をもたず、VC 104 の数も異なる。各 SN コンピュータ 414 は、SN コンピュータ 214-2 と同様に実施されるが、VC 104、計算エージェント 314 及びハイパーバイザ 312 (図 3) をもたない。

【 0 0 2 4 】

CN コンピュータ 412 及びウェブサーバー 218 は、1 つ以上のパブリックスイッチ 50

232によりコンピュータネットワーク102に接続される。CNコンピュータ412、SNコンピュータ414、AZCコンピュータ216、及びウェブサーバー218は、1つ以上のプライベートスイッチ234により互いに接続される。

【0025】

図5は、本発明の1つ以上の実施形態においてシステム100で新たなVPSAをスポーンする方法500のフローチャートである。この方法500及びここに述べる他の方法は、1つ以上のブロックで示された1つ以上のオペレーション、ファンクション又はアクションを含む。ブロックは、順次に示されているが、並列に遂行されてもよいし、及び／又はここに述べるものとは異なる順序で遂行されてもよい。又、種々のブロックは、望ましい具現化に基づいて、より少数のブロックへ合成されてもよいし、付加的なブロックへ分割されてもよいし、及び／又は排除されてもよい。方法500は、ブロック502でスタートする。

【0026】

ブロック502において、ウェブサーバー118は、ウェブフォーム120をカスタマーコンピュータ112へ送信して、カスタマーがVPSAのパラメータを与えることができるようとする。上述したように、パラメータは、(1)VPSAにおける各VCのVCハードウェアモデル、(2)VPSAのドライブ特性、及び(3)VPSAのドライブ量、を明示する。パラメータは、更に、VPSAへのI/Oアクセスを認証するためのクレデンシャルを含む。パラメータは、ウェブサーバー118へ返送される。ブロック502に続いて、ブロック504が行われる。

【0027】

ブロック504において、ウェブサーバー118は、ウェブページ122をVPSAの料金と共にカスタマーコンピュータ112へ送信する。方法500の説明上、VPSAを生成するための確認がウェブサーバー118へ返送されると仮定する。ブロック504に続いて、ブロック506が行われる。

【0028】

ブロック506において、ウェブサーバー118は、サービス要求と、VPSAのパラメータをAZC116へ送信する。ブロック506に続いて、ブロック508が行われる。

【0029】

ブロック508において、システム100は、VPSAにおける更に生成されるべきVC104のためのプレースホルダにバーチャルドライブ106を割り当てる。ブロック508は、以下に詳細に述べる図6A及び図6Bの方法600で実施される。ブロック508に続いて、ブロック510が行われる。

【0030】

ブロック510において、AZC116は、バーチャルドライブ106がVPSAに首尾良く割り当たされたかどうか決定する。もしそうでなければ、ブロック510に続いて、ブロック512が行われる。物理的ドライブ108が首尾良く割り当たられた場合には、ブロック510に続いて、ブロック514が行われる。

【0031】

ブロック512において、AZC116は、VPSAのドライブ必要性を満足するバーチャルドライブ106が不充分であるときに、エラーが発生したと決定する。AZC116は、ウェブサーバー118がエラーメッセージをカスタマーコンピュータ112へ送信するようにさせ、そして方法500を終了させる。

【0032】

ブロック514において、システム100は、以下に詳細に述べる図7A及び図7Bの方法700によりVPSAのVC104を生成する。ブロック514に続いて、ブロック516が行われる。

【0033】

ブロック516において、AZC116は、VPSAのVC104が首尾良く生成され

10

20

30

40

50

たかどうか決定する。もしそうでなければ、ブロック516に続いて、ブロック518が行われる。VPSAのVC104が首尾良く生成された場合には、ブロック516に続いて、ブロック522が行われる。

【0034】

ブロック518において、AZC116は、ブロック508でVPSAに既に割り当てられたバーチャルドライブ106を解放する。ブロック518に続いて、ブロック520が行われる。

【0035】

ブロック520において、AZC116は、VPSAに対する明示されたVCハードウェアモデルを伴うVC104が不充分であるときに、エラーが発生したと決定する。AZC116は、ウェブサーバー118がエラーメッセージをカスタマーコンピュータ112へ送信するようにさせ、そして方法500を終了させる。10

【0036】

ブロック522において、VC104は、互いにクラスター化ハンドシェークを確立して、VCの役割を確立する。例えば、あるVC104が一次として働く一方、別のVC104がスタンバイであるか、或いは両VCがアクティブに負荷を分担する。ブロック522に続いて、ブロック524が行われる。

【0037】

ブロック524において、VC104は、バーチャルドライブ106からSN114により生成されたセットアップ区画を発見するように試みる。上述したように、セットアップ区画がVC104により使用されて、VPSAシステム情報及びメタデータを記憶するためのセットアップボリュームを生成する。ブロック524に続いて、ブロック526が行われる。20

【0038】

ブロック526において、VC104は、それらがセットアップ区画を発見したかどうか決定する。もしそうでなければ、ブロック526に続いて、ブロック528が行われる。VC104がセットアップ区画を発見した場合には、ブロック526に続いて、ブロック530が行われる。

【0039】

ブロック528において、VC104は、エラーが生じたことをAZC116に通知する。AZC116は、ウェブサーバー118がエラーメッセージをカスタマーコンピュータ112へ送信するようにさせ、そして方法500を終了させる。30

【0040】

ブロック530において、VC104は、そのセットアップ区画の冗長セットから、保護されたセットアップボリュームを生成する。そのセットアップボリュームは、これに限定されないが、物理的及びバーチャルオブジェクトメタデータ、計測統計値、並びにロギング及びトレーシングを含むVPSAシステムデータの持続ストレージを与えるのに使用される。

【0041】

図6A及び図6Bは、本発明の1つ以上の実施形態においてシステム100でバーチャルドライブ106をVPSAのVC104に割り当てるための規範的な方法600のフローチャートである。この方法600は、ブロック602で始まる。40

【0042】

ブロック602において、ウェブサーバー118は、ウェブフォーム120をカスタマーコンピュータ112へ送信し、カスタマーがVPSAのパラメータを与えることができるようになる。上述したように、パラメータは、ドライブ特性と、VPSAに対する1組のバーチャルドライブ106のドライブ量とを含む。ドライブ特性は、ドライブ形式、ドライブ容量、及びドライブ暗号を明示する。パラメータは、ウェブサーバー118へ返送される。ブロック602に続いて、ブロック604が行われる。

【0043】

ブロック 604において、ウェブフォーム 120 は、カスタマーが別の1組のバーチャルドライブ 106 を追加したいかどうか決定するように構成される。例えば、ウェブフォーム 120 は、別の1組のバーチャルドライブ 106 を追加するためにカスタマーが選択する「より多くのドライブを追加する」ボタンを含む。カスタマーが別の1組のバーチャルドライブ 106 を追加したい場合には、ブロック 604 は、ブロック 602 ヘループバックする。さもなければ、ブロック 604 に続いて、ブロック 606 が行われる。

【0044】

ブロック 606において、AZC116 は、利用可能な物理的ドライブ 108 及びそれらのドライブ特性のリストを全てのVC104 から検索する。このリストは、SN114 によりレポートされるドライブインベントリーから発生される。ブロック 606 に続いて、ブロック 608 が行われる。10

【0045】

ブロック 608 は、要求された全ドライブ形式を通るループのスタートである。要求された各組のバーチャルドライブ 106 について、AZC116 は、要求された組に対して明示されたドライブ特性を満足するか又はそれを越える利用可能な物理的ドライブ 108 を有する SN114 の候補 SN リストを生成する。ブロック 608 に続いて、ブロック 610 が行われる。

【0046】

ブロック 610において、AZC116 は、候補 SN リストが空であるかどうか決定する。もしそうであれば、ブロック 610 に続いて、ブロック 612 が行われる。さもなければ、ブロック 610 に続いて、ブロック 614 が行われる。20

【0047】

ブロック 612において、AZC116 は、利用可能な物理的ドライブ 108 がないときにエラーが発生したと決定する。AZC116 は、ウェブサーバー 118 がエラーメッセージをカスタマーコンピュータ 112 へ送信するようにさせ、そして方法 600 を終了させる。

【0048】

ブロック 614において、AZC116 は、1つ以上の分類基準に基づいて候補 SN リストを分類する。分類基準は、基礎となるサーバーコンピュータの利用率である。ブロック 614 に続いて、ブロック 616 が行われる。30

【0049】

ブロック 616において、AZC116 は、1つ以上のドライブ分布及び RAID 保護基準を満足する候補 SN リストにおけるトップランク SN を選択する。例えば、2 方向 RAID - 1 は、2 つの SN を要求し、一方、RAID - 5 は、できるだけ多数の SN の間での分布を要求する。ブロック 616 に続いて、ブロック 618 が行われる。

【0050】

ブロック 618において、AZC116 は、充分な SN があるかどうか決定する。もしそうでなければ、ブロック 618 に続いて、ブロック 620 が行われる。さもなければ、ブロック 618 に続いて、ブロック 630 が行われる。

【0051】

ブロック 630において、AZC116 は、追加要求された1組以上のバーチャルドライブ 106 がループに留まるかどうか決定する。もしそうであれば、ブロック 630 は、ブロック 608 ヘループバックし、新たに要求される組に対して新たな候補 SN リストを生成し、そして上述したプロセスを繰り返す。さもなければ、ブロック 630 に続いて、ブロック 632 が行われる。40

【0052】

ブロック 632において、AZC116 は、選択された SN リストにおける選択された各 SN114 へメッセージを送信し、ドライブ特性及びドライブ量を満足するか又はそれを越えるバーチャルドライブ 106 を VPSA に割り当てる。ブロック 632 に続いて、ブロック 634 が行われる。50

【0053】

ブロック634において、選択されたSN114は、各バーチャルドライブ106においてセットアップ区画及びデータ区画を生成する。上述したように、セットアップ区画は、VPSAシステム情報及びメタデータを記憶するためのセットアップボリュームを生成するようにVC104により使用される。ブロック634に続いて、ブロック636が行われる。

【0054】

ブロック636において、選択されたSNは、セットアップ及びデータ区画をVPSAのVC104へ露出させる。ブロック636に続いて、ブロック638が行われる。

【0055】

ブロック638において、選択されたSN114は、更新されたドライブインベントリーをAZC116へ報告し、方法600を終了させる。

【0056】

図7A及び図7Bは、本発明の1つ以上の実施形態においてシステム100でVPSAにVCを生成するための規範的な方法700のフローチャートである。この方法700は、ブロック702で開始する。

【0057】

ブロック702において、ウェブサーバー118は、ウェブフォーム120をカスタマーコンピュータ112へ送信して、カスタマーがVPSAのパラメータを与えることができるようとする。上述したように、パラメータは、VPSAの各VCに対するVCハードウェアモデルと、VPSAへのI/Oアクセスを認証するためのクレデンシャルとを含む。VCハードウェアモデルは、CPUモデル、1つ以上のCPU特徴、CPU量、RAM容量、及びネットワーク帯域幅を明示する。パラメータは、ウェブサーバー118へ返送され、ウェブサーバーは、それらをAZC116へ送信する。ブロック702に続いて、ブロック704が行われる。

【0058】

ブロック704において、AZC116は、全てのサーバーコンピュータにおける計算エージェント314から、CPU、メモリ、及びNICのネットワーク帯域幅の利用状態を検索する。ブロック704に続いて、ブロック706が行われる。

【0059】

ブロック706において、AZC116は、明示されたVCハードウェアモデルを満足するか又はそれを越える利用可能なネットワーク帯域幅をもつ利用可能なプロセッサ/メモリ複合体及びNICを有するAZ内のサーバーコンピュータの候補サーバコンピュータリストを生成する。ブロック706に続いて、ブロック708が行われる。

【0060】

ブロック708において、AZC116は、候補サーバコンピュータリストがVC104の要求された数より少ないかどうか決定する。もしそうであれば、ブロック708に続いて、ブロック710が行われる。さもなければ、ブロック708に続いて、ブロック712が行われる。

【0061】

ブロック710において、AZC116は、要求されたVCハードウェアモデルを満足するか又はそれを越えるプロセッサ/メモリ複合体及びNICが不充分であるときにエラーが生じたと決定する。AZC116は、ウェブサーバー118がエラーメッセージをカスタマーコンピュータ112へ送信するようにさせ、そして方法700を終了させる。

【0062】

ブロック712において、AZC116は、1つ以上の分類基準に基づいて候補サーバコンピュータリストを分類する。分類基準は、サーバーコンピュータの利用率である。ブロック712に続いて、ブロック714が行われる。

【0063】

ブロック714において、AZC116は、利用可能なネットワークリソースをチェック

10

20

30

40

50

クし、そしてVCネットワーク構成を定義する。ある範囲の利用可能なパブリックIPアドレスから、AZC116は、カスタマーコンピュータと通信するためにVC104のVNICへパブリックIPアドレスを割り当てる。ある範囲の利用可能なプライベートIPアドレスから、AZC116は、キャッシングVC間、及びVCとSNとの間で通信するためにVC104のVNICにIPアドレスを割り当てる。ブロック714に続いて、ブロック716が行われる。

【0064】

ブロック716において、AZC116は、VC104及びSN114に割り当てるための充分なパブリック及び/又はプライベートIPアドレスがあるので、ネットワーク構成が首尾良くセットされたかどうか決定する。もしそうでなければ、ブロック716に続いて、ブロック718が行われる。さもなければ、ブロック716に続いて、ブロック720が行われる。

10

【0065】

ブロック718において、AZC116は、パブリック及び/又はプライベートIPアドレスが不充分であるときにエラーが生じたと決定する。AZC116は、ウェブサーバー118がエラーメッセージをカスタマーコンピュータ112へ送信するようにさせ、そして方法700を終了させる。

【0066】

ブロック720において、AZC116は、候補サーバーコンピュータリストにおいてトップランクサーバーコンピュータを選択し、そしてそれに基づいてVCソフトウェア(画像)でVMをスポーンするための要求を、それに基づいて計算エージェント314(図3)へ送信する。ブロック720に続いて、ブロック722が行われる。

20

【0067】

ブロック722において、選択されたサーバーコンピュータにおける計算エージェント314は、その要求を受信して、新たなVMをスポーンする。ブロック722に続いて、ブロック724が行われる。

【0068】

ブロック724において、計算エージェント314は、VMが首尾良くスポーンされたかどうか決定する。もしそうでなければ、ブロック724に続いて、ブロック726が行われる。さもなければ、ブロック724に続いて、ブロック728が行われる。

30

【0069】

ブロック726において、AZC116は、VMをスポーンするときにエラーが生じたかどうか決定する。AZC116は、ウェブサーバー118がエラーメッセージをカスタマーコンピュータ112へ送信するようにさせ、そして方法700を終了させる。

【0070】

ブロック728において、VCは、VMにおいて開始し、そしてAZC116からVPSA及びVC情報を検索し、方法700を終了させる。AZC116から検索される情報は、VPSA情報(名前及びID)、VC情報(各VCは、AZ内に独特のインスタンスIDを有する)、ネットワーク情報(VCのMAC及びIPアドレス、並びにプライベート及びパブリックネットワークに対するVMのVNICの関連性)、及びVPSAへのI/Oアクセスを認証するためのクレデンシャルを含む。そのような情報は、VCが持続的識別を維持し、そのネットワークを適切にセットアップし、及び1つ以上のキャッシングVCとのクラスター化ハンドシェークを確立するために必要である。

40

【0071】

上述した実施形態の種々の他の適応及び組み合わせも本発明の範囲内である。多数の実施形態が特許請求の範囲によって包含される。

【符号の説明】

【0072】

100：ソフトウェアシステム

102：コンピュータネットワーク

50

- 104 : パーチャルコントローラ (V C)
 106 : パーチャルドライブ
 108 : 物理的ドライブ
 110 : パーチャルボリューム
 112 : カスタマーコンピュータ
 114 : ストレージノード (S N)
 116 : 利用ゾーンコントローラ (A Z C)
 118 : ウェブサーバー
 120 : ウェブフォーム
 122 : ウェブページ
 124 : マネージメントインターフェイス
 126, 128 : V P S A
 200 : ハードウェアシステム
 214 : サーバーコンピュータ
 216 : A Z C コンピュータ
 218 : ウェブサーバーコンピュータ
 302 : プロセッサ / メモリ複合体
 304 : C P U
 306 : R A M
 308 : N I C
 310 : V M
 312 : ハイパーバイザ
 400 : システム
 412 : C N コンピュータ
 414 : S N コンピュータ

【図 1】

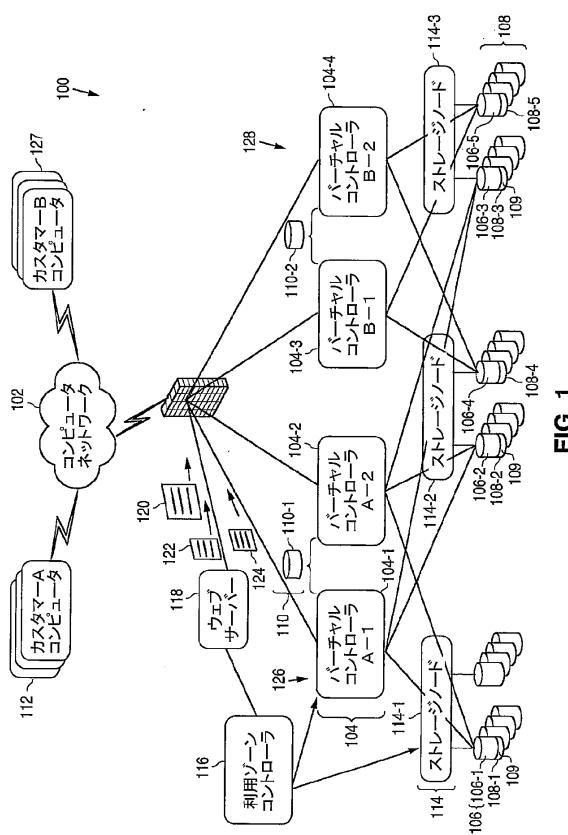


FIG. 1

【図 2】

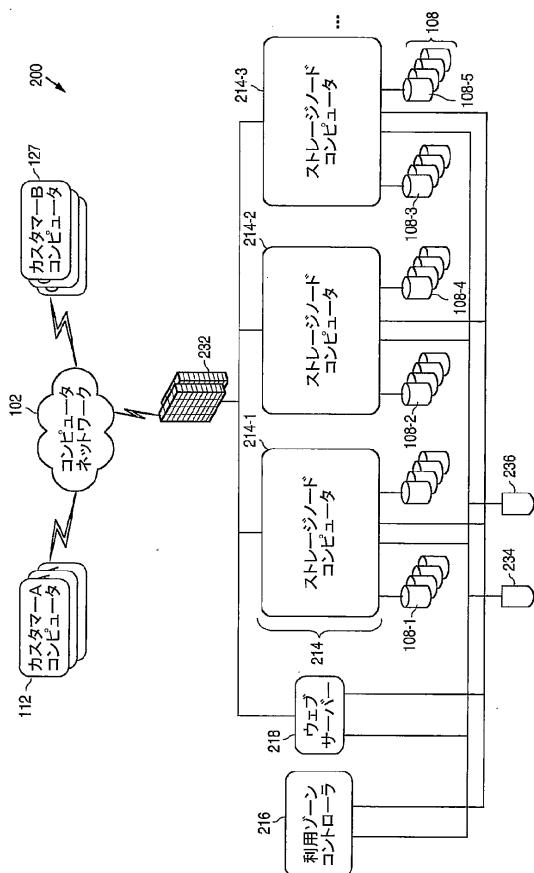


FIG. 2

【図3】

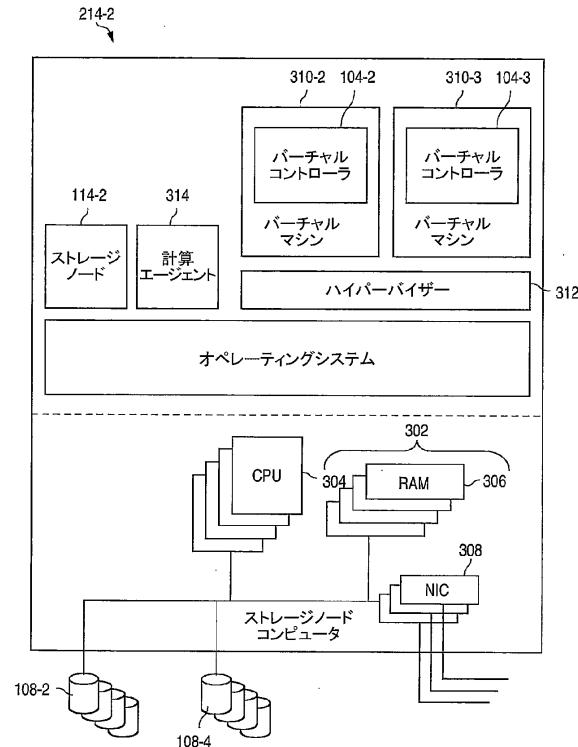


FIG. 3

【図4】

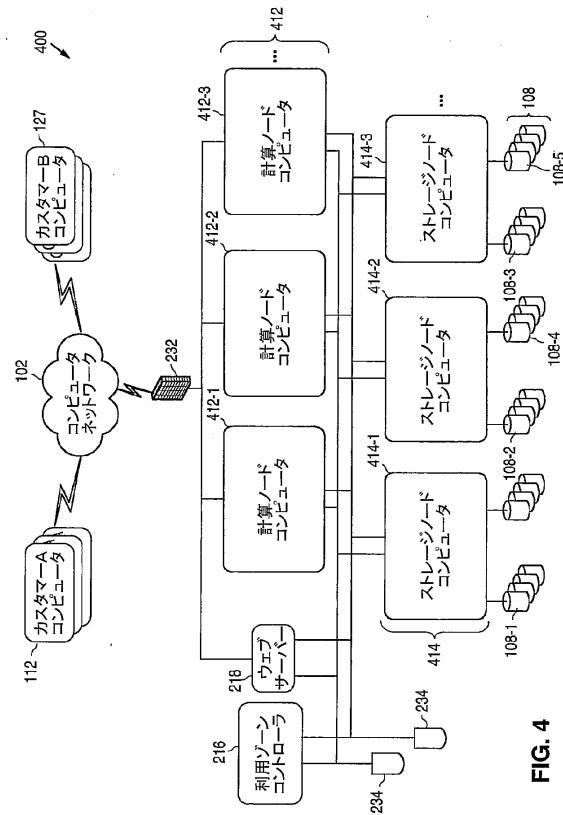


FIG. 4

【図5】

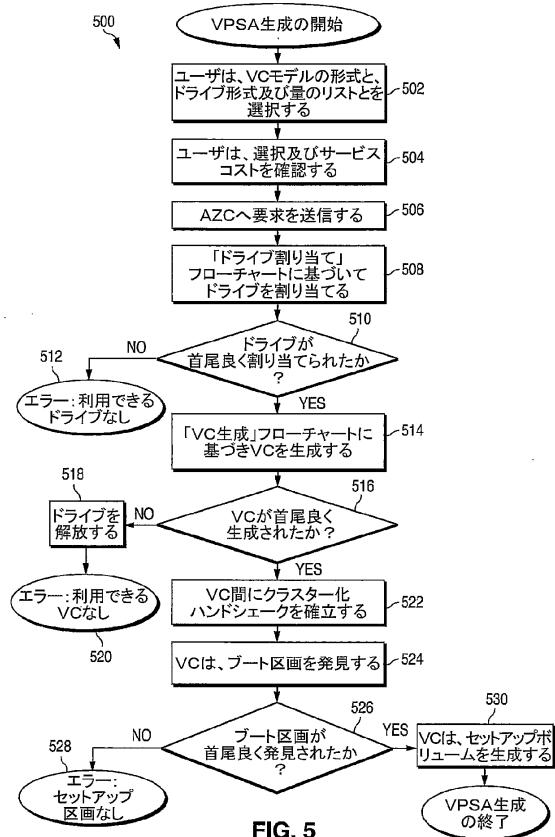


FIG. 5

【図6A】

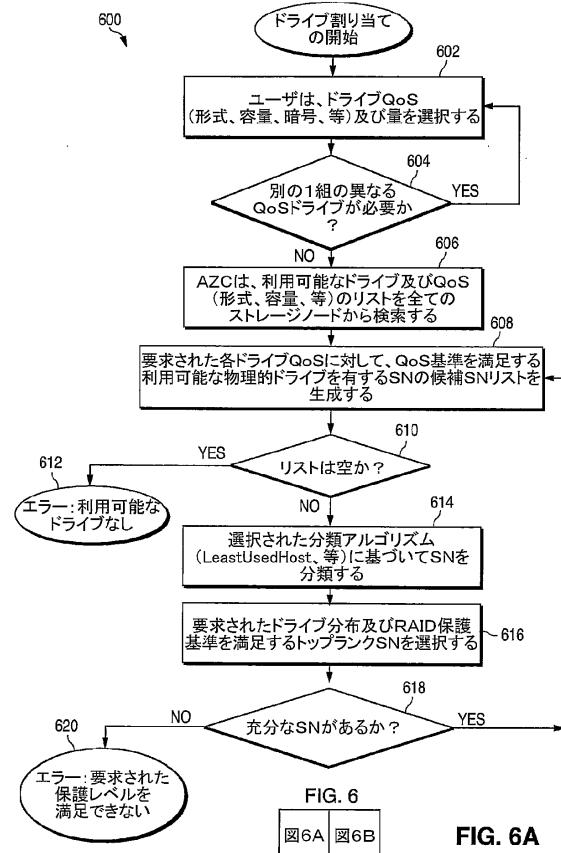


FIG. 6A

FIG. 6B

【図 6 B】

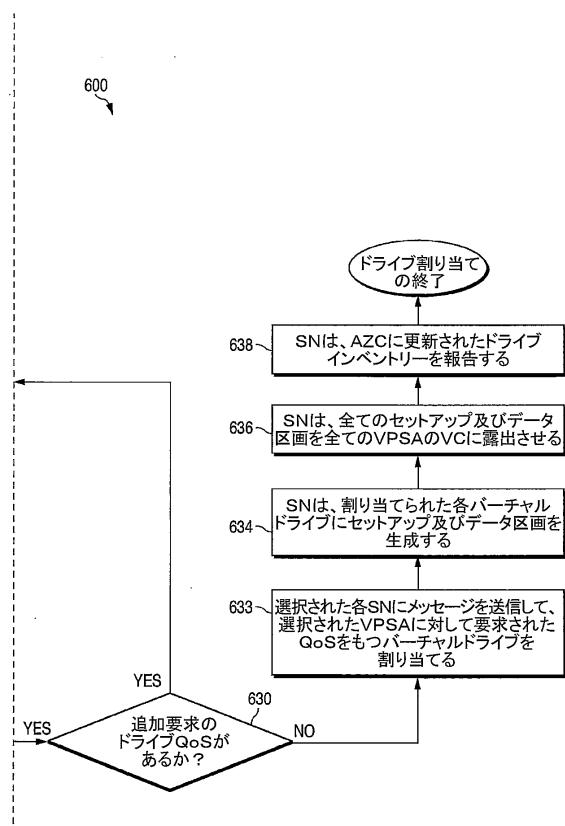


FIG. 6B

【図 7 A】

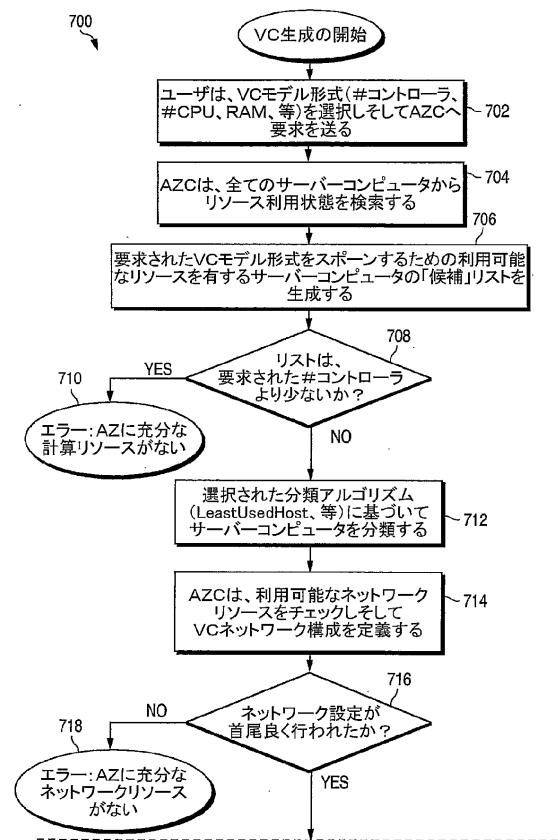
FIG. 7
図7A
図7B

FIG. 7A

【図 7 B】

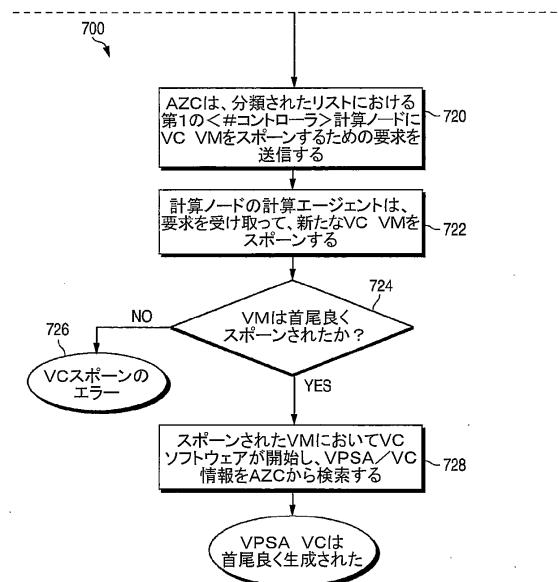


FIG. 7B

フロントページの続き

(74)代理人 100109335
弁理士 上杉 浩

(74)代理人 100120525
弁理士 近藤 直樹

(74)代理人 100139712
弁理士 那須 威夫

(72)発明者 ナホム ネルソン
アメリカ合衆国 カリフォルニア州 92782 タスティン カミリア アベニュー 1660
1

(72)発明者 ハーシュコ ヤイル
イスラエル 36005 アロネイ アバ ピーオーボックス 190

(72)発明者 オシア ヨーブ
イスラエル 34353 ハイファ ピンスキーハイテク ストリート 39エイ

(72)発明者 カウシイク ヴイ シヤム
インド カルナータカ 560101 バンガロール エイチエスアール レイアウト セカンド
セクター トゥウェンティセブンス クロス ナインティーンス メイン ナガルジュナ グリ
ーン リッジ アパートメント ジー-602

(72)発明者 ヴァインブラット レヴ
イスラエル キリアット ヤム ピンハス ラボン 17/21

(72)発明者 ラカス アレックス
イスラエル 34790 ハイファ ハサーラヴ ストリート 35/1

(72)発明者 メルニコフ モシェ
イスラエル 30095 ラマット イシャイ ハエゴス ストリート 33 ピーオーボックス
1218

(72)発明者 ポポフスキー ウラジーミル
アメリカ合衆国 カリフォルニア州 92606 アーバイン デル ビンセント 8

審査官 加内 慎也

(56)参考文献 特開2008-257572 (JP, A)
特開2013-041569 (JP, A)
米国特許出願公開第2011/0022812 (US, A1)
米国特許出願公開第2011/0185355 (US, A1)
米国特許出願公開第2012/0096149 (US, A1)

(58)調査した分野(Int.Cl., DB名)

G 06 F 12/00
G 06 F 3/06
G 06 F 9/46
G 06 F 13/10