

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号
特表2025-504816
(P2025-504816A)

(43)公表日 令和7年2月19日(2025.2.19)

(51)国際特許分類

F I

H O 4 L 41/085(2022.01) H O 4 L 41/085

H O 4 L 41/0895(2022.01) H O 4 L 41/0895

G O 6 F 9/50 (2006.01) G O 6 F 9/50 1 5 0 A

審査請求 未請求 予備審査請求 未請求 (全65頁)

(21)出願番号	特願2024-541887(P2024-541887)	(71)出願人	502303739
(86)(22)出願日	令和4年12月20日(2022.12.20)		オラクル・インターナショナル・コーポレーション
(85)翻訳文提出日	令和6年8月26日(2024.8.26)		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 ショアーズ, メール ストップ 5 オーピー 7 オラクル パークウェイ 5 0 0
(86)国際出願番号	PCT/US2022/082075		
(87)国際公開番号	WO2023/136965	(74)代理人	110001195
(87)国際公開日	令和5年7月20日(2023.7.20)		弁理士法人深見特許事務所
(31)優先権主張番号	63/298,683	(72)発明者	ブラール, ジャグウィンダー・シン
(32)優先日	令和4年1月12日(2022.1.12)		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、オラクル・パークウェイ、5 0 0、オラクル・インターナショナル・コーポレーション内
(33)優先権主張国・地域又は機関	米国(US)		
(31)優先権主張番号	17/935,711		
(32)優先日	令和4年9月27日(2022.9.27)		
(33)優先権主張国・地域又は機関	米国(US)		
(81)指定国・地域	AP(BW,CV,GH,GM,KE,LR,LS,MW,MZ		
	最終頁に続く		最終頁に続く

(54)【発明の名称】 一般的なワークロードの物理トポロジーネットワーク局所性の公開

(57)【要約】

本明細書では、一般的なワークロードの実行のために、クラスタネットワークに含まれるホストマシンの階層的局所性情報を活用する技術について論じられる。複数のホストマシンの各ホストマシンの階層的局所性情報が記憶される。複数の局所性の各局所性について、その局所性の位置情報を識別するホストマシンの階層的局所性情報。ワークロードの実行を要求する要求を受信したことに応答して、複数のホストマシンの階層的局所性情報が取得され、要求に応じて（例えば、顧客に）提供される。

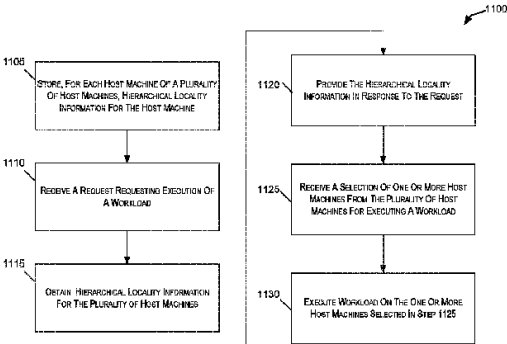


FIG. 11A

【特許請求の範囲】**【請求項 1】**

複数のホストマシンの各ホストマシンについて、前記ホストマシンの階層的局所性情報を記憶することを含み、ホストマシンの前記階層的局所性情報は、複数の階層レベルの各々について、前記ホストマシンの位置情報を識別し、

ワークロードの実行を要求する要求を受信したことに応答して、

前記複数のホストマシンの前記階層的局所性情報を取得することと、

前記要求に対する応答として、前記複数のホストマシンの前記階層的局所性情報を提供することと、を含む、方法。

【請求項 2】

前記複数の階層レベルは、ラック局所性、ブロック局所性、建物局所性、および利用可能なドメイン局所性を含む、請求項 1 に記載の方法。

【請求項 3】

前記複数の階層レベルの各階層レベルの前記位置情報は、前記階層レベルの識別子を含む、請求項 2 に記載の方法。

【請求項 4】

前記ホストマシンの前記階層的局所性情報を含む符号化される配列を生成することを含み、前記符号化される配列は複数のハッシュブロックを含み、各ハッシュブロックは特定の局所性の符号化される位置情報を含む、請求項 1 に記載の方法。

【請求項 5】

前記符号化される配列は、(i) 前記ホストマシンを含むラックのラック識別子のハッシュと顧客識別子を含む第 1 のハッシュブロック、(ii) 前記ラックを含むブロックのブロック識別子のハッシュと前記顧客識別子を含む第 2 のハッシュブロック、(iii) 前記ブロックを含む建物の建物識別子のハッシュと前記顧客識別子を含む第 3 のハッシュブロック、および (iv) 前記建物を含むドメインのドメイン識別子のハッシュと前記顧客識別子を含む第 4 のハッシュブロックを含む、請求項 4 に記載の方法。

【請求項 6】

顧客が、前記複数のホストマシンの前記階層的局所性情報を受信することと、

ワークロードを実行するために前記複数のホストマシンから 1 つまたは複数のホストマシンを選択することと、をさらに含み、前記選択は前記ワークロードに関連付けられる 1 つまたは複数の制約に基づいて実行され、

前記 1 つまたは複数のホストマシン上で前記ワークロードを実行することをさらに含む、請求項 1 に記載の方法。

【請求項 7】

1 つまたは複数の制約は、前記ワークロードに関連付けられるレイテンシ閾値に対応する第 1 の制約と、前記ワークロードを実行する際の所望の可用性の程度に関連付けられる第 2 の制約とを含む、請求項 6 に記載の方法。

【請求項 8】

前記ホストマシンの前記階層的局所性情報はインスタンスメタデータサービスから取得され、前記階層的局所性情報は、前記ホストマシンに関連付けられるネットワーク仮想化装置を介して前記インスタンスメタデータサービスによって前記ホストマシンに記憶される、請求項 1 に記載の方法。

【請求項 9】

コンピュータ実行可能命令を記憶するコンピュータ可読媒体であって、前記コンピュータ実行可能命令は、1 つまたは複数のプロセッサによって実行されると、

複数のホストマシンの各ホストマシンについて、前記ホストマシンの階層的局所性情報を記憶することを行わせ、ホストマシンの前記階層的局所性情報は、複数の階層レベルの各々について、前記ホストマシンの位置情報を識別し、

前記コンピュータ実行可能命令は、

ワークロードの実行を要求する要求を受信したことに応答して、

10

20

30

40

50

前記複数のホストマシンの前記階層的局所性情報を取得し、

前記要求に対する応答として、前記複数のホストマシンの前記階層的局所性情報を提供することを行わせる、コンピュータ可読媒体。

【請求項 10】

前記複数の階層レベルは、ラック局所性、ブロック局所性、建物局所性、および利用可能なドメイン局所性を含む、請求項 9 に記載のコンピュータ実行可能命令を記憶するコンピュータ可読媒体。

【請求項 11】

前記複数の階層レベルの各階層レベルの前記位置情報は、前記階層レベルの識別子を含む、請求項 10 に記載のコンピュータ実行可能命令を記憶するコンピュータ可読媒体。

【請求項 12】

前記命令は、1 つまたは複数のプロセッサによって実行されると、

前記ホストマシンの前記階層的局所性情報を含む符号化される配列を生成させる命令をさらに含み、前記符号化される配列は複数のハッシュブロックを含み、各ハッシュブロックは特定の局所性の符号化される位置情報を含む、請求項 9 に記載のコンピュータ実行可能命令を記憶するコンピュータ可読媒体。

【請求項 13】

前記符号化される配列は、(i) 前記ホストマシンを含むラックのラック識別子のハッシュと顧客識別子とを含む第 1 のハッシュブロック、(ii) 前記ラックを含むブロックのブロック識別子のハッシュと前記顧客識別子とを含む第 2 のハッシュブロック、(iii) 前記ブロックを含む建物の建物識別子のハッシュと前記顧客識別子とを含む第 3 のハッシュブロック、および(iv) 前記建物を含むドメインのドメイン識別子のハッシュと前記顧客識別子とを含む第 4 のハッシュブロックを含む、請求項 12 に記載のコンピュータ実行可能命令を記憶するコンピュータ可読媒体。

【請求項 14】

前記命令は、1 つまたは複数のプロセッサによって実行されると、

顧客によって、前記複数のホストマシンの前記階層的局所性情報を受信させ、

ワークロードを実行するために前記複数のホストマシンから 1 つまたは複数のホストマシンを選択させる命令をさらに含み、前記選択は前記ワークロードに関連付けられる 1 つまたは複数の制約に基づいて実行され、

前記 1 つまたは複数のホストマシン上で前記ワークロードを実行させる命令をさらに含む、請求項 9 に記載のコンピュータ実行可能命令を記憶するコンピュータ可読媒体。

【請求項 15】

1 つまたは複数の制約は、前記ワークロードに関連付けられるレイテンシ閾値に対応する第 1 の制約と、前記ワークロードを実行する際の所望の可用性の程度に関連付けられる第 2 の制約とを含む、請求項 14 に記載のコンピュータ実行可能命令を記憶するコンピュータ可読媒体。

【請求項 16】

前記ホストマシンの前記階層的局所性情報はインスタンスメタデータサービスから取得され、前記階層的局所性情報は、前記ホストマシンに関連付けられるネットワーク仮想化装置を介して前記インスタンスメタデータサービスによって前記ホストマシンに記憶される、請求項 9 に記載のコンピュータ実行可能命令を記憶するコンピュータ可読媒体。

【請求項 17】

プロセッサと、

命令を含むメモリと、を含むコンピューティング装置であって、

前記命令は、前記プロセッサで実行されると、前記コンピューティング装置に少なくとも

複数のホストマシンの各ホストマシンについて、前記ホストマシンの階層的局所性情報を記憶することを行わせ、ホストマシンの前記階層的局所性情報は、複数の階層レベルの各々について、前記ホストマシンの位置情報を識別し、

10

20

30

40

50

前記命令は、前記プロセッサで実行されると、前記コンピューティング装置に、ワークロードの実行を要求する要求を受信したことに応答して、前記複数のホストマシンの前記階層的局所性情報を取得させ、前記要求に応じて前記複数のホストマシンの前記階層的局所性情報を提供させる、コンピューティング装置。

【請求項 18】

前記複数の階層レベルは、ラック局所性、ブロック局所性、建物局所性、および利用可能なドメイン局所性を含む、請求項 17 に記載のコンピューティング装置。

【請求項 19】

前記複数の階層レベルの各階層レベルの前記位置情報は、前記階層レベルの識別子を含む、請求項 17 に記載のコンピューティング装置。

10

【請求項 20】

前記ホストマシンの前記階層的局所性情報を含む符号化される配列を生成するようにさらに構成され、前記符号化される配列は複数のハッシュブロックを含み、各ハッシュブロックは特定の局所性の符号化される位置情報を含む、請求項 17 に記載のコンピューティング装置。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

20

本出願は、2022年1月12日に提出された米国仮出願第63／298,683号および2022年9月27日に提出された米国非仮出願第17／935,711号に対する優先権を主張する。前述の出願の全内容は、あらゆる目的のために、その全体が参照により本明細書に組み込まれる。

【0002】

分野

本開示は、クラウドインフラストラクチャに含まれるホストマシンの局所性情報を利用してワークロードを実行することに関する。

【背景技術】

【0003】

30

背景

組織は、オンプレミスのハードウェアとソフトウェアの購入、更新、保守にかかるコストを削減するために、ビジネスアプリケーションとデータベースをクラウドに移行し続けている。高性能コンピューティングアプリケーションは、特定の成果または結果を達成するために、利用可能なすべてのコンピューティング能力を一貫して消費する。このようなアプリケーションには、専用のネットワーク性能、高速記憶、高いコンピューティング能力、大量のメモリが必要であるが、これらは、今日のコモディティクラウドを構成する仮想化インフラストラクチャでは不足しているリソースである。

【0004】

クラウドインフラストラクチャサービスプロバイダは、これらのアプリケーションの要件に対応するために、より新しく高速なグラフィック処理装置（GPU）を提供している。GPUワークロードは通常、1つまたは複数のホストマシン上で実行される。通常、このようなワークロードでは、期待されるレベルのスループットを達成できない。この問題の要因の1つは、フローエントロピー（例えば、等コストマルチパス（ECMP）フローエントロピー）の欠如である。さらに、ホストマシン（つまり、ホスト）が局所ネットワーク近隣の他のホストを考慮せずにトラフィックを交換するという事実によって、問題は悪化する。他のタイプのワークロードは通常、インフラストラクチャから1つまたは複数のホストマシンをランダム（つまり、任意）の方法で選択することによって実行される。つまり、ワークロードは、局所性情報（例えば、ホストマシンの物理的な場所）を考慮せずに実行される。そのため、これらのワークロードのスループットは低くなる。本明細書

40

50

で説明する実施形態は、これらの問題およびその他の問題に対処する。

【発明の概要】

【0005】

概要

本開示は、一般的には、クラウドインフラストラクチャに含まれるホストマシンの局所性情報を利用してワークロードを実行することに関する。本明細書では、方法、システム、1つまたは複数のプロセッサによって実行可能なプログラム、コード、または命令を記憶する非一時的なコンピュータ可読媒体などを含む、さまざまな実施形態が説明される。これらの例示的な実施形態は、開示を制限または定義するためではなく、理解を助けるための例を示すために言及されている。追加の実施形態は詳細な説明セクションで説明されており、そこでさらに詳しく説明されている。

10

【0006】

本開示の一実施形態は、複数のホストマシンの各ホストマシンについて、ホストマシンの階層的局所性情報を記憶することであって、ホストマシンの階層的局所性情報は、複数の階層レベルの各々について、ホストマシンの位置情報を識別する、ことと、ワークロードの実行を要求する要求を受信したことに応答して、複数のホストマシンの階層的局所性情報を取得し、複数のホストマシンの階層的局所性情報を要求への応答として提供することを含む方法に関する。

【0007】

本開示の一態様は、1つまたは複数のデータプロセッサと、1つまたは複数のデータプロセッサ上で実行されると、1つまたは複数のデータプロセッサに本明細書に開示される1つまたは複数の方法の一部または全部を実行させる命令を含む非一時的なコンピュータ可読記憶媒体とを含むシステムを提供する。

20

【0008】

本開示の別の態様は、1つまたは複数のデータプロセッサに本明細書に開示される1つまたは複数の方法の一部または全部を実行させるように構成される命令を含む、非一時的なマシン可読記憶媒体に有形に具体化されるコンピュータプログラム製品を提供する。

【0009】

前述の内容、ならびに他の特徴および実施形態は、以下の明細書、特許請求の範囲、および添付の図面を参照することでより明らかになるであろう。

30

【0010】

本開示の特徴、実施形態、および利点は、添付の図面を参照して以下の詳細な説明を読むことによってよりよく理解される。

【図面の簡単な説明】

【0011】

【図1】 特定の実施形態によるクラウドサービスプロバイダインフラストラクチャによってホストされる仮想またはオーバーレイクラウドネットワークを示す分散環境の高レベル図である。

【図2】 特定の実施形態によるC S P I内の物理ネットワーク内の物理構成要素の簡略化されるアーキテクチャ図を示す図である。

40

【図3】 特定の実施形態による、ホストマシンが複数のネットワーク仮想化装置（N V D）に接続されるC S P I内の配置例を示す図である。

【図4】 特定の実施形態によるマルチテナントをサポートするためのI / O仮想化を提供するためのホストマシンとN V Dとの間の接続を示す図である。

【図5】 特定の実施形態によるC S P Iによって提供される物理ネットワークの簡略化されるブロック図を示す図である。

【図6】 特定の実施形態による、C L O Sネットワーク配置を組み込んだクラウドインフラストラクチャの簡略化されるブロック図を示す図である。

【図7】 特定の実施形態による、クラウドインフラストラクチャに含まれるラックの例示的な構成を示す図である。

50

【図 8】 特定の実施形態による、局所性情報の階層を示す概略図である。

【図 9】 特定の実施形態による、異なるレベルの局所性情報を示す例示的なツリー図を示す図である。

【図 10】 特定の実施形態による、ホストマシンの局所性情報の例示的な符号化される配列を示す図である。

【図 11 A】 特定の実施形態による、要求をプロビジョニングする際に実行されるステップを示す例示的なフローチャートを示す図である。

【図 11 B】 特定の実施形態による、顧客のワークロード要求をプロビジョニングする際に実行されるステップを示す別の例示的なフローチャートを示す図である。

【図 12】 少なくとも 1 つの実施形態による、クラウドインフラストラクチャをサービスシステムとして実装するための 1 つのパターンを示すブロック図である。

【図 13】 少なくとも 1 つの実施形態による、クラウドインフラストラクチャをサービスシステムとして実装するための別のパターンを示すブロック図である。

【図 14】 少なくとも 1 つの実施形態による、クラウドインフラストラクチャをサービスシステムとして実装するための別のパターンを示すブロック図である。

【図 15】 少なくとも 1 つの実施形態による、クラウドインフラストラクチャをサービスシステムとして実装するための別のパターンを示すブロック図である。

【図 16】 少なくとも 1 つの実施形態による例示的な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】

【0012】

詳細な説明

以下の説明では、説明の目的で、特定の実施形態を完全に理解できるように具体的な詳細が記載されている。しかし、これらの具体的な詳細がなくても、さまざまな実施形態を実践できることは明らかである。図および説明は制限的なものではない。本明細書で使用されている「例示的」という語は、「例、事例、または説明として役立つ」という意味である。本明細書で「例示的」として説明される実施形態または設計は、必ずしも他の実施形態または設計よりも好ましいまたは有利であると解釈されるものではない。

【0013】

クラウドインフラストラクチャのアーキテクチャ例

クラウドサービスという用語は、一般的に、クラウドサービスプロバイダ（CSP）が提供するシステムおよびインフラストラクチャ（クラウドインフラストラクチャ）を使用して、CSPによってユーザまたは顧客にオンデマンド（例えば、加入モデルを介して）で提供されるサービスを指すために使用される。通常、CSPのインフラストラクチャを構成するサーバとシステムは、顧客独自のオンプレミスサーバおよびシステムとは別である。したがって、顧客はサービスのために別途ハードウェアおよびソフトウェアリソースを購入することなく、CSPが提供するクラウドサービス自体を利用できる。クラウドサービスは、サービス提供に使用されるインフラストラクチャの調達に顧客が投資することなく、アプリケーションやコンピューティングリソースへの簡単かつスケーラブルなアクセスを加入顧客に提供するように設計されている。

【0014】

さまざまなタイプのクラウドサービスを提供するクラウドサービスプロバイダが複数存在する。クラウドサービスには、Software-as-a-Service（SaaS）、Platform-as-a-Service（PaaS）、Infrastructure-as-a-Service（IaaS）など、さまざまなタイプやモデルがある。

【0015】

顧客は、CSPが提供する 1 つまたは複数のクラウドサービスに加入することができる。顧客は、個人、組織、企業など、あらゆるエンティティになり得る。CSPが提供するサービスに顧客が加入または登録すると、その顧客のテナントまたはアカウントが作成される。顧客は、次いでこのアカウントを介して、アカウントに関連付けられている 1 つま

10

20

30

40

50

たは複数の加入済みクラウドリソースにアクセスできるようになる。

【0016】

前述のように、*infrastructure as a service (IaaS)* は、クラウドコンピューティングサービスの特定のタイプである。*IaaS* モデルでは、*CSP* は、顧客が独自のカスタマイズ可能なネットワークを構築し、顧客リソースを展開するために使用できるインフラストラクチャ（クラウドサービスプロバイダインフラストラクチャまたは *CSP I* と呼ばれる）を提供する。したがって、顧客のリソースとネットワークは、*CSP* によって提供されるインフラストラクチャによって分散環境でホストされる。これは、顧客のリソースとネットワークが顧客が提供するインフラストラクチャによってホストされる従来のコンピューティングとは異なる。

10

【0017】

CSP I は、さまざまなホストマシン、メモリリソース、および物理ネットワークを形成するネットワークリソースを含む相互接続される高性能コンピューティングリソースを含み得、これはサブストレートネットワークまたはアンダーレイネットワークとも呼ばれる。*CSP I* のリソースは、1つまたは複数のデータセンターに分散し得、そのデータセンターは1つまたは複数の地理的リージョンにまたがって地理的に分散し得る。仮想化ソフトウェアは、これらの物理リソースによって実行され、仮想化される分散環境を提供し得る。仮想化により、物理ネットワーク上にオーバーレイネットワーク（ソフトウェアベースネットワーク、ソフトウェア定義ネットワーク、または仮想ネットワークとしても知られる）が作成される。*CSP I* 物理ネットワークは、物理ネットワーク上に1つまたは複数のオーバーレイネットワークまたは仮想ネットワークを作成するための基盤を提供する。仮想ネットワークまたはオーバーレイネットワークは、1つまたは複数の仮想クラウドネットワーク（*VCN*）を含むことができる。仮想ネットワークは、ソフトウェア仮想化技術（例えば、ハイパーバイザ、ネットワーク仮想化装置（*NVD*）によって実行される機能（例えば、スマート *NIC*）、トッポブブラック（*TOR*）スイッチ、*NVD* によって実行される1つまたは複数の機能を実装するスマート *TOR*、およびその他のメカニズム）を使用して実装され、物理ネットワーク上で実行できるネットワーク抽象化の層が作成される。仮想ネットワークには、ピアツーピアネットワーク、*IP* ネットワークなど、さまざまな形式をとり得る。仮想ネットワークは通常、層3 *IP* ネットワークまたは層2 *VLAN* のいずれかである。この仮想またはオーバーレイネットワークの方法は、仮想またはオーバーレイ層3ネットワークと呼ばれることがよくある。仮想ネットワーク用に開発されるプロトコルの例には、*IP-in-IP*（または *Generic Routing Encapsulation (GRE)*）、*Virtual Extensible LAN (VXLAN-IEEE RFC 7348)*、*Virtual Private Networks (VPN)*（例えば、*MPLS Layer-3 Virtual Private Networks (RFC 4364)*）、*VMware* の *NSX*、*GENEVE (Generic Network Virtualization Encapsulation)* などがある。

20

30

【0018】

IaaS の場合、*CSP* によって提供されるインフラストラクチャ（*CSP I*）は、パブリックネットワーク（例えば、インターネット）を介して仮想化されるコンピューティングリソースを提供するように構成することができる。*IaaS* モデルでは、クラウドコンピューティングサービスプロバイダがインフラストラクチャ構成要素（例えば、サーバ、記憶装置、ネットワークノード（例えば、ハードウェア）、展開ソフトウェア、プラットフォーム仮想化（例えば、ハイパーバイザ層）など）をホストできる。場合によっては、*IaaS* プロバイダが、これらのインフラストラクチャ構成要素に付随するさまざまなサービス（例えば、課金、監視、ログ記録、セキュリティ、負荷分散、クラスターリングなど）も提供することがある。したがって、これらのサービスはポリシー駆動型であり得るため、*IaaS* ユーザは、アプリケーションの可用性と性能を維持するために負荷分散を駆動するポリシーを実装できる可能性がある。*CSP I* は、顧客が可用性の高いホスト型分散環境で幅広いアプリケーションとサービスを構築および実行できるようにするインフラストラクチャと補完的なクラウドサービスのセットを提供する。*CSP I* は、顧客のオ

40

50

ンプレミスネットワークなどのさまざまなネットワークの場所から安全にアクセスできる柔軟な仮想ネットワークで、高性能なコンピューティングリソースと機能、および記憶容量を提供する。顧客がCSPが提供するIaaSサービスに加入または登録すると、その顧客用に作成されるテナントはCSP I内の安全で分離されるパーティションとなり、顧客はここで自分のクラウドリソースを作成、編成、管理できる。

【0019】

顧客は、CSP Iが提供するコンピューティング、メモリ、およびネットワークリソースを使用して独自の仮想ネットワークを構築できる。これらの仮想ネットワークには、コンピューティングインスタンスなどの1つまたは複数の顧客リソースまたはワークロードを展開できる。例えば、顧客はCSP Iが提供するリソースを使用して、仮想クラウドネットワーク（VCN）と呼ばれる1つまたは複数のカスタマイズ可能なプライベート仮想ネットワークを構築できる。顧客は、コンピューティングインスタンスなどの1つまたは複数の顧客リソースを顧客VCNに展開できる。コンピューティングインスタンスは、仮想マシン、ベアメタルインスタンスなどの形式をとることができる。したがって、CSP Iは、顧客が可用性の高い仮想ホスト環境で幅広いアプリケーションとサービスを構築および実行できるようにするインフラストラクチャと補完的なクラウドサービスのセットを提供する。顧客は、CSP Iによって提供される基礎となる物理リソースを管理または制御しないが、オペレーティングシステム、記憶、および展開されるアプリケーションを制御し、選択したネットワーク構成要素（例えば、ファイアウォール）を限定的に制御することもできる。

【0020】

CSPは、顧客およびネットワーク管理者がCSP Iリソースを使用してクラウドに展開されるリソースを構成、アクセス、および管理できるようにするコンソールを提供することができる。特定の実施形態では、コンソールは、CSP Iにアクセスして管理するために使用できるWebベースのユーザインターフェースを提供する。いくつかの実装では、コンソールはCSPによって提供されるWebベースのアプリケーションである。

【0021】

CSP Iは、単一テナントまたはマルチテナントアーキテクチャをサポートすることができる。単一テナントアーキテクチャでは、ソフトウェア（例えば、アプリケーション、データベース）またはハードウェア構成要素（例えば、ホストマシン、サーバ）が単一の顧客またはテナントにサービスを提供する。マルチテナントアーキテクチャでは、ソフトウェアまたはハードウェア構成要素が複数の顧客またはテナントにサービスを提供する。したがって、マルチテナントアーキテクチャでは、CSP Iリソースは複数の顧客またはテナント間で共有される。マルチテナントの状況では、各テナントのデータが分離され、他のテナントから見えないようにするために、CSP I内で予防措置が講じられ、安全対策が講じられる。

【0022】

物理ネットワークにおいて、ネットワークエンドポイント（「エンドポイント」）とは、物理ネットワークに接続され、接続先のネットワークと往復して通信するコンピューティング装置またはシステムを指す。物理ネットワーク内のネットワークエンドポイントは、局所エリアネットワーク（LAN）、広域ネットワーク（WAN）、またはその他のタイプの物理ネットワークに接続され得る。物理ネットワーク内の従来のエンドポイントの例には、モデム、ハブ、ブリッジ、スイッチ、ルータ、その他のネットワーク装置、物理コンピュータ（またはホストマシン）などが含まれる。物理ネットワーク内の各物理装置は、装置との通信に使用できる固定のネットワークアドレスを有する。この固定ネットワークアドレスは、層2アドレス（例えば、MACアドレス）、固定層3アドレス（例えば、IPアドレス）などになり得る。仮想化環境または仮想ネットワークでは、エンドポイントには、物理ネットワークの構成要素によってホストされる（例えば、物理ホストマシンによってホストされる）仮想マシンなどのさまざまな仮想エンドポイントを含めることができる。仮想ネットワーク内のこれらのエンドポイントは、オーバーレイ層2アドレス

(例えば、オーバーレイMACアドレス)およびオーバーレイ層3アドレス(例えば、オーバーレイIPアドレス)などのオーバーレイアドレスによってアドレス指定される。ネットワークオーバーレイは、ネットワーク管理者がソフトウェア管理(例えば、仮想ネットワークの制御プレーンを実装するソフトウェア経由)を使用して、ネットワークエンドポイントに関連付けられているオーバーレイアドレスを移動できるようにすることで柔軟性を実現する。したがって、物理ネットワークとは異なり、仮想ネットワークでは、ネットワーク管理ソフトウェアを使用して、オーバーレイアドレス(例えば、オーバーレイIPアドレス)を1つのエンドポイントから別のエンドポイントに移動できる。仮想ネットワークは物理ネットワーク上に構築されるため、仮想ネットワーク内の構成要素間の通信には、仮想ネットワークと基盤となる物理ネットワークの両方が関与する。このような通信を容易にするために、CSPの構成要素は、仮想ネットワーク内のオーバーレイアドレスをサブストレートネットワーク内の実際の物理アドレスにマッピングするマッピング、およびその逆のマッピングを学習して記憶するように構成されている。これらのマッピングは、次いで、通信を容易にするために使用される。顧客トラフィックは、仮想ネットワークでのルーティングを容易にするためにカプセル化される。

10

【0023】

したがって、物理アドレス(例えば、物理IPアドレス)は物理ネットワーク内の構成要素に関連付けられ、オーバーレイアドレス(例えば、オーバーレイIPアドレス)は仮想ネットワーク内のエンティティに関連付けられる。物理IPアドレスとオーバーレイIPアドレスはどちらも実際のIPアドレスの一種である。これらは、仮想IPアドレスが複数の実際のIPアドレスにマップされる仮想IPアドレスとは異なる。仮想IPアドレスは、仮想IPアドレスと複数の実際のIPアドレス間の1対多のマッピングを提供する。

20

【0024】

クラウドインフラストラクチャまたはCSPは、世界中の1つまたは複数のリージョンにある1つまたは複数のデータセンターで物理的にホストされる。CSPには、物理ネットワークまたはサブストレートネットワーク内の構成要素と、物理ネットワーク構成要素上に構築される仮想ネットワーク内の仮想化構成要素(例えば、仮想ネットワーク、コンピューティングインスタンス、仮想マシン)が含まれ得る。特定の実施形態では、CSPは、レルム、リージョン、および可用性ドメインに編成され、ホストされる。リージョンは通常、1つまたは複数のデータセンターを含むローカライズされる地理的領域である。リージョンは一般に互いに独立しており、例えば、国や大陸をまたいで非常に離れた距離で隔てられている場合もある。例えば、第1のリージョンはオーストラリアにあり、別のリージョンは日本にあり、さらに別のリージョンはインドにあり得る、といった具合である。CSPリソースはリージョン間で分割されるため、各リージョンには独自の独立したCSPリソースのサブセットが存在する。各リージョンは、コンピューティングリソース(例えば、ベアメタルサーバ、仮想マシン、コンテナ、関連インフラストラクチャ)、記憶リソース(例えば、ブロックボリューム記憶、ファイル記憶、オブジェクト記憶、アーカイブ記憶)、ネットワークリソース(例えば、仮想クラウドネットワーク(VCN)、負荷分散リソース、オンプレミスネットワークへの接続)、データベースリソース、エッジネットワークリソース(例えば、DNS)、およびアクセス管理および監視リソースなど、コアインフラストラクチャサービスのセットとリソースを提供し得る。通常、各リージョンには、レルム内の他のリージョンに接続する複数のパスがある。

30

40

【0025】

一般に、アプリケーションは、最も頻繁に使用されるリージョン(すなわち、そのリージョンに関連付けられるインフラストラクチャ上に展開)に展開される。これは、近くのリソースを使用する方が遠くのリソースを使用するよりも高速であるためである。アプリケーションは、大規模な気象システムや地震などのリージョン全体のイベントのリスクを軽減するための冗長性、法的管轄区域、課税ドメイン、その他のビジネス基準や社会基準などのさまざまな要件を満たすためなど、さまざまな理由で異なるリージョンに展開され

50

ることもある。

【0026】

リージョン内のデータセンターは、さらに編成され、可用性ドメイン（AD）に細分化されることができる。可用性ドメインは、リージョン内に配置される1つまたは複数のデータセンターに対応し得る。リージョンは1つまたは複数の可用性ドメインで構成できる。このような分散環境では、CSPリソースは、仮想クラウドネットワーク（VCN）などのリージョン固有、またはコンピューティングインスタンスなどの可用性ドメイン固有になる。

【0027】

リージョン内のADは互いに分離されており、耐障害性であり、同時に障害が発生する可能性が非常に低いように構成されている。これは、ADがネットワーク、物理ケーブル、ケーブルバス、ケーブルエントリポイントなどの重要なインフラストラクチャリソースを共有しないことで実現されるため、あるリージョン内の1つのADで障害が発生しても、同じリージョン内の他のADの可用性に影響が及ぶ可能性は低くなる。同じリージョン内のADは、低レイテンシ、高帯域幅のネットワークで相互に接続できるため、他のネットワーク（例えば、インターネット、顧客のオンプレミスネットワークなど）への高可用性接続を提供し、高可用性と災害復旧の両方のために複数のADに複製されるシステムを構築できる。クラウドサービスでは、高可用性を確保し、リソース障害から保護するために複数のADを使用する。IaaSプロバイダが提供するインフラストラクチャが拡大するにつれて、追加の容量とともに、より多くのリージョンとADが追加され得る。可用性ドメイン間のトラフィックは通常、暗号化される。

10

20

【0028】

特定の実施形態では、リージョンはレルムにグループ化される。レルムはリージョンの論理的な集合である。レルムは互いに分離されており、データを共有しない。同じレルム内のリージョンは相互に通信できるが、異なるレルム内のリージョンは通信できない。CSPで顧客のテナントまたはアカウントは単一のレルムに存在し、そのレルムに属する1つまたは複数のリージョンに分散できる。通常、顧客がIaaSサービスに加入すると、レルム内の顧客指定のリージョン（「ホーム」リージョンと呼ばれる）にその顧客のテナントまたはアカウントが作成される。顧客は、レルム内の1つまたは複数の他のリージョンにわたって顧客のテナントを拡張できる。顧客は、顧客のテナントが存在するレルム外のリージョンにアクセスすることはできない。

30

【0029】

IaaSプロバイダは、複数のレルムを提供することができ、各レルムは特定の顧客またはユーザのセットに対応する。例えば、商業顧客向けに商業レルムが提供され得る。別の例として、特定の国内の顧客に対して、その国専用のレルムが提供され得る。さらに別の例として、政府などに政府レルムが提供され得る。例えば、政府レルムは特定の政府向けに作成され、商業レルムよりもセキュリティレベルが高くなっている場合がある。例えば、Oracle Cloud Infrastructure（OCI）は現在、商用リージョン用のレルムと、政府クラウドリージョン用の2つのレルム（例えば、FedRAMP認定およびIL5認定）を提供している。

40

【0030】

特定の実施形態では、ADは1つまたは複数の障害ドメインに分割できる。障害ドメインは、反親和性を提供するためにAD内のインフラストラクチャリソースをグループ化したものである。障害ドメインを使用すると、コンピューティングインスタンスを分散して、インスタンスが単一のAD内の同じ物理ハードウェア上に存在しないようにすることができる。これは反親和性として知られている。障害ドメインとは、単一の障害点を共有するハードウェア構成要素（例えば、コンピュータ、スイッチ）のセットを指す。コンピューティングプールは、論理的に障害ドメインに分割される。このため、1つの障害ドメインに影響するハードウェア障害またはコンピューティングハードウェアのメンテナンスイベントは、他の障害ドメインのインスタンスには影響しない。実施形態に応じて、各AD

50

の障害ドメインの数は異なり得る。例えば、特定の実施形態では、各 A D には 3 つの障害ドメインが含まれる。障害ドメインは、A D 内の論理データセンターとして機能する。

【0031】

顧客が I a a S サービスに加入すると、C S P I からのリソースが顧客向けにプロビジョニングされ、顧客のテナントに関連付けられる。顧客は、プロビジョニングされるこれらのリソースを使用してプライベートネットワークを構築し、これらのネットワーク上にリソースを展開できる。C S P I によってクラウドでホストされる顧客ネットワークは、仮想クラウドネットワーク (V C N) と呼ばれる。顧客は、顧客に割り当てられる C S P I リソースを使用して、1 つまたは複数の仮想クラウドネットワーク (V C N) を設定できる。V C N は、仮想またはソフトウェア定義のプライベートネットワークである。顧客の V C N に展開される顧客リソースには、コンピューティングインスタンス (例えば、仮想マシン、ベアメタルインスタンス) およびその他のリソースが含まれ得る。これらのコンピューティングインスタンスは、アプリケーション、負荷分散装置、データベースなどのさまざまな顧客ワークロードを表し得る。V C N に展開されるコンピューティングインスタンスは、インターネットなどのパブリックネットワークを介してパブリックにアクセス可能なエンドポイント (「パブリックエンドポイント」)、同じ V C N または他の V C N (例えば、顧客の他の V C N、顧客に属さない V C N) 内の他のインスタンス、顧客のオンプレミスデータセンターまたはネットワーク、サービスエンドポイント、およびその他のタイプのエンドポイントと通信できる。

10

【0032】

C S P は、C S P I を使用してさまざまなサービスを提供することができる。場合によっては、C S P I の顧客自身がサービスプロバイダのように行動し、C S P I リソースを使用してサービスを提供できる。サービスプロバイダは、識別情報 (例えば、I P アドレス、D N S 名、ポート) によって特徴付けられるサービスエンドポイントを公開し得る。顧客のリソース (例えば、コンピューティングインスタンス) は、特定のサービスに対してサービスによって公開されているサービスエンドポイントにアクセスすることで、特定のサービスを利用できる。これらのサービスエンドポイントは通常、インターネットなどのパブリック通信ネットワークを介してエンドポイントに関連付けられているパブリック I P アドレスを使用してユーザがパブリックにアクセスできるエンドポイントである。パブリックにアクセス可能なネットワークエンドポイントは、パブリックエンドポイントと呼ばれることもある。

20

30

【0033】

特定の実施形態では、サービスプロバイダは、サービスのエンドポイント (サービスエンドポイントと呼ばれることもある) を介してサービスを公開することができる。サービスの顧客は、次いで、このサービスエンドポイントを使用してサービスにアクセスできるようになる。特定の実装では、サービスに提供されるサービスエンドポイントに、そのサービスを利用しようとする複数の顧客がアクセスできる。他の実装では、専用のサービスエンドポイントが顧客に提供され得、その顧客のみがその専用サービスエンドポイントを使用してサービスにアクセスできるようになる。

【0034】

特定の実施形態では、V C N が作成されると、V C N に割り当てられるプライベートオーバーレイ I P アドレスの範囲 (例えば、10.0/16) であるプライベートオーバーレイクラスレスインタードメインルーティング (C I D R) アドレス空間に関連付けられる。V C N には、関連のサブネット、ルートテーブル、ゲートウェイが含まれる。V C N は単一のリージョン内に存在するが、リージョンの可用性ドメインの 1 つまたは複数またはすべてにまたがることができる。ゲートウェイは、V C N 用に構成される仮想インターフェースであり、V C N と V C N 外部の 1 つまたは複数のエンドポイントとの間のトラフィックの通信を可能にする。異なるタイプのエンドポイントとの通信を可能にするために、V C N に対して 1 つまたは複数の異なるタイプのゲートウェイを構成できる。

40

【0035】

50

V C Nは、1つまたは複数のサブネットなどの1つまたは複数のサブネットワークに分割することができる。したがって、サブネットは、V C N内で作成できる構成単位またはサブディビジョンである。V C Nには1つまたは複数のサブネットを含めることができる。V C N内の各サブネットは、そのV C N内の他のサブネットと重複せず、V C Nのアドレス空間内のアドレス空間サブセットを表す、連続した範囲のオーバーレイIPアドレス（例えば、10.0.0.0/24および10.0.1.0/24）に関連付けられる。

【0036】

各コンピューティングインスタンスは、コンピューティングインスタンスがV C Nのサブネットに参加できるようにする仮想ネットワークインターフェースカード（V N I C）に関連付けられている。V N I Cは、物理的なネットワークインターフェースカード（N I C）の論理表現である。一般に、V N I Cはエンティティ（例えば、コンピューティングインスタンス、サービス）と仮想ネットワーク間のインターフェースである。V N I Cはサブネット内に存在し、1つまたは複数の関連のIPアドレスと、関連のセキュリティルールまたはポリシーを持つ。V N I Cはスイッチ上の層2ポートに相当する。V N I Cは、コンピューティングインスタンスとV C N内のサブネットに接続される。コンピューティングインスタンスに関連付けられているV N I Cにより、コンピューティングインスタンスはV C Nのサブネットの一部になることができ、コンピューティングインスタンスは、コンピューティングインスタンスと同じサブネット上のエンドポイント、V C N内の異なるサブネットのエンドポイント、またはV C N外のエンドポイントと通信（例えば、パケットの送受信）できるようになる。したがって、コンピューティングインスタンスに関連付けられているV N I Cによって、コンピューティングインスタンスがV C Nの内外のエンドポイントに接続する方法が決定する。コンピューティングインスタンスが作成され、V C N内のサブネットに追加されると、コンピューティングインスタンスのV N I Cが作成され、そのコンピューティングインスタンスに関連付けられる。コンピューティングインスタンスのセットで構成されるサブネットの場合、サブネットにはコンピューティングインスタンスのセットに対応するV N I Cが含まれ、各V N I Cはコンピューティングインスタンスのセット内のコンピューティングインスタンスに接続される。

【0037】

各コンピューティングインスタンスには、コンピューティングインスタンスに関連付けられているV N I Cを介してプライベートオーバーレイIPアドレスが割り当てられる。このプライベートオーバーレイIPアドレスは、コンピューティングインスタンスの作成時にコンピューティングインスタンスに関連付けられているV N I Cに割り当てられ、コンピューティングインスタンスとの間のトラフィックのルーティングに使用される。特定のサブネット内のすべてのV N I Cは、同じルートテーブル、セキュリティリスト、およびD H C P任意選択肢を使用する。前述のように、V C N内の各サブネットは、そのV C N内の他のサブネットと重複せず、V C Nのアドレス空間内のアドレス空間サブセットを表す、連続した範囲のオーバーレイIPアドレス（例えば、10.0.0.0/24および10.0.1.0/24）に関連付けられている。V C Nの特定のサブネット上のV N I Cの場合、V N I Cに割り当てられるプライベートオーバーレイIPアドレスは、サブネットに割り当てられるオーバーレイIPアドレスの連続した範囲からのアドレスである。

【0038】

特定の実施形態では、コンピューティングインスタンスには、任意選択で、プライベートオーバーレイIPアドレスに加えて、例えばパブリックサブネット内にある場合は1つまたは複数のパブリックIPアドレスなどの追加のオーバーレイIPアドレスが割り当てられ得る。これらの複数のアドレスは、同じV N I C上、またはコンピューティングインスタンスに関連付けられている複数のV N I C上に割り当てられる。しかし、各インスタンスには、インスタンスの起動時に作成され、インスタンスに割り当てられるオーバーレイプライベートIPアドレスに関連付けられているプライマリV N I Cがあり、このプライマリV N I Cは削除できない。セカンダリV N I Cと呼ばれる追加のV N I Cは、プラ

イマリ V N I C と同じ可用性ドメイン内の既存のインスタンスに追加できる。すべての V N I C はインスタンスと同じ可用性ドメイン内にある。セカンダリ V N I C は、プライマリ V N I C と同じ V C N 内のサブネット、または同じ V C N 内または別の V C N 内の別のサブネットにあり得る。

【 0 0 3 9 】

コンピューティングインスタンスがパブリックサブネット内にある場合、任意選択でパブリック IP アドレスを割り当てることができる。サブネットは、作成時にパブリックサブネットまたはプライベートサブネットのいずれかとして指定できる。プライベートサブネットとは、サブネット内のリソース（例えば、コンピューティングインスタンス）と関連する V N I C にパブリックオーバーレイ IP アドレスを持たせることができないことを意味する。パブリックサブネットとは、サブネット内のリソースと関連する V N I C にパブリック IP アドレスを持たせることができることを意味する。顧客は、サブネットを単一の可用性ドメイン内、またはリージョンもしくはレム内の複数の可用性ドメイン間に存在するように指定できる。

【 0 0 4 0 】

上述のように、V C N は 1 つまたは複数のサブネットに分割され得る。特定の実施形態では、V C N 用に構成される仮想ルータ（V R）（V C N V R または単に V R と呼ばれる）によって、V C N のサブネット間の通信が可能になる。V C N 内のサブネットの場合、V R はそのサブネットの論理ゲートウェイを表し、サブネット（つまり、そのサブネット上のコンピューティングインスタンス）が V C N 内の他のサブネット上のエンドポイントおよび V C N 外の他のエンドポイントと通信できるようにする。V C N V R は、V C N 内の V N I C と V C N に関連付けられる仮想ゲートウェイ（「ゲートウェイ」）間のトラフィックをルーティングするように構成される論理エンティティである。ゲートウェイについては、図 1 を参照して以下でさらに詳しく説明する。V C N V R は、層 3 / IP 層の概念である。一実施形態では、V C N に対して 1 つの V C N V R があり、V C N V R には IP アドレスでアドレス指定されるポートが無制限に存在する可能性があり、V C N の各サブネットに対して 1 つのポートがある。このように、V C N V R は、V C N V R が接続されている V C N 内のサブネットごとに異なる IP アドレスを持つ。V R は、V C N 用に構成されるさまざまなゲートウェイにも接続される。特定の実施形態では、サブネットのオーバーレイ IP アドレス範囲からの特定のオーバーレイ IP アドレスが、そのサブネットの V C N V R のポート用に予約される。例えば、それぞれ 1 0 . 0 / 1 6 と 1 0 . 1 / 1 6 のアドレス範囲が関連付けられる 2 つのサブネットを持つ V C N を考える。アドレス範囲が 1 0 . 0 / 1 6 である V C N 内の第 1 のサブネットの場合、この範囲のアドレスがそのサブネットの V C N V R のポート用に予約される。場合によっては、範囲の第 1 の IP アドレスが V C N V R 用に予約され得る。例えば、オーバーレイ IP アドレス範囲が 1 0 . 0 / 1 6 のサブネットの場合、IP アドレス 1 0 . 0 . 0 . 1 がそのサブネットの V C N V R のポート用に予約され得る。同じ V C N 内のアドレス範囲が 1 0 . 1 / 1 6 である第 2 のサブネットの場合、V C N V R には IP アドレスが 1 0 . 1 . 0 . 1 であるその第 2 のサブネット用のポートが存在し得る。V C N V R には、V C N 内のサブネットごとに異なる IP アドレスがある。

【 0 0 4 1 】

いくつかの他の実施形態では、V C N 内の各サブネットには、V R に関連付けられる予約済みまたはデフォルトの IP アドレスを使用してサブネットによってアドレス指定可能な独自の関連 V R が存在し得る。予約済みまたはデフォルトの IP アドレスは、例えば、そのサブネットに関連付けられている IP アドレスの範囲の第 1 の IP アドレスであり得る。サブネット内の V N I C は、このデフォルトまたは予約済みの IP アドレスを使用して、サブネットに関連付けられている V R と通信（例えば、パケットの送受信）できる。このような実施形態では、V R はそのサブネットの入口 / 出口ポイントになる。V C N 内のサブネットに関連付けられる V R は、V C N 内の他のサブネットに関連付けられている他の V R と通信できる。V R は、V C N に関連付けられているゲートウェイとも通信でき

る。サブネットのVR機能は、サブネット内のVNICに対してVNIC機能を実行する1つまたは複数のNVD上で動作されているか、またはNVDによって実行される。

【0042】

VCNに対してルートテーブル、セキュリティルール、およびDHCP任意選択肢を構成できる。ルートテーブルはVCNの仮想ルートテーブルであり、ゲートウェイまたは特別に構成されたインスタンスを介して、VCN内のサブネットからVCN外部の宛先にトラフィックをルーティングするためのルールが含まれている。VCNのルートテーブルをカスタマイズして、VCNとの間でパケットを転送／ルーティングする方法を制御できる。DHCP任意選択肢とは、インスタンスの起動時に自動的に提供される構成情報を指す。

10

【0043】

VCNに対して構成されるセキュリティルールは、VCNのオーバーレイファイアウォールルールを表す。セキュリティルールには、入口ルールと出口ルールを含めることができ、VCN内のインスタンスに出入りできるトラフィックのタイプ（例えば、プロトコルおよびポートに基づく）を指定できる。顧客は、特定のルールがステートフルかステートレスかを選択できる。例えば、顧客は、ソースCIDR0.0.0.0/0、宛先TCPポート22でステートフル入口ルールを設定することで、どこからでもインスタンスのセットにSSHトラフィックを着信させることを許可できる。セキュリティルールは、ネットワークセキュリティグループまたはセキュリティリストを使用して実装できる。ネットワークセキュリティグループは、そのグループ内のリソースにのみ適用されるセキュリティ

20

【0044】

特定の実施形態では、VCNの構成情報は、VCN制御プレーンによって決定され、記憶される。VCNの構成情報には、例えば、VCNに関連付けられているアドレス範囲、VCN内のサブネットおよび関連情報、VCNに関連付けられている1つまたは複数のVR、VCN内のコンピューティングインスタンスおよび関連のVNIC、VCNに関連付けられるさまざまな仮想化ネットワーク機能（例えば、VNIC、VR、ゲートウェイ）を実行するNVD、VCNの状態情報、およびその他のVCN関連情報に関する情報が含まれ得る。特定の実施形態では、VCN配信サービスは、VCN制御プレーンによって記憶される構成情報またはその一部をNVDに公開する。配信される情報は、NVDによって記憶され、VCN内のコンピューティングインスタンスとの間でパケットを転送するために使用する情報（例えば、転送テーブル、ルーティングテーブル）を更新するために使用され得る。

30

【0045】

特定の実施形態では、VCNおよびサブネットの作成はVCN制御プレーン（CP）によって処理され、コンピューティングインスタンスの起動はコンピューティング制御プレーンによって処理される。コンピューティング制御プレーンは、コンピューティングインスタンスの物理リソースを割り当ての役割を担い、次いで、VCN制御プレーンを呼び出してVNICを作成し、コンピューティングインスタンスに接続する。VCN CPは、パケット転送およびルーティング機能を実行するように構成されるVCNデータプレーンにVCNデータマッピングも送信する。特定の実施形態では、VCN CPは、VCNデータプレーンに更新を提供する役割を担う配信サービスを提供する。VCN制御プレーンの例は、図12、13、14、および15（参照1216、1316、1416、および1516を参照）にも示されており、以下で説明する。

40

【0046】

顧客は、CSP Iによってホストされるリソースを使用して1つまたは複数のVCNを

50

作成することができる。顧客のV C Nに展開されるコンピューティングインスタンスは、さまざまなエンドポイントと通信し得る。これらのエンドポイントには、C S P IによってホストされるエンドポイントとC S P I外部のエンドポイントが含まれ得る。

【0047】

C S P Iを使用してクラウドベースのサービスを実装するためのさまざまな異なるアーキテクチャが図1、2、3、4、5、12、13、14、および15に示されており、以下で説明する。図1は、特定の実施形態によるC S P Iによってホストされるオーバーレイまたは顧客V C Nを示す分散環境100の高レベル図である。図1に示されている分散環境100には、オーバーレイネットワーク内の複数の構成要素が含まれている。図1に示す分散環境100は単なる例であり、請求される実施形態の範囲を不当に制限することを意図するものではない。多くのバリエーション、代替、変更が可能である。例えば、いくつかの実装では、図1に示されている分散環境には、図1に示されているものよりも多いまたは少ないシステムまたは構成要素が含まれ得、2つ以上のシステムを組み合わせ得るか、システムの構成または配置が異なり得る。

【0048】

図1に示す例に示すように、分散環境100は、顧客が加入して仮想クラウドネットワーク(V C N)を構築するために使用できるサービスおよびリソースを提供するC S P I 101を含む。特定の実施形態では、C S P I 101は加入顧客にI a a Sサービスを提供する。C S P I 101内のデータセンターは、1つまたは複数のリージョンに編成され得る。一例としてのリージョン「R e g i o n U S」102を図1に示す。顧客はリージョン102に顧客V C N 104を構成した。顧客はV C N 104上にさまざまなコンピューティングインスタンスを展開できる。コンピューティングインスタンスには、仮想マシンまたはベアメタルインスタンスが含まれ得る。インスタンスの例としては、アプリケーション、データベース、負荷分散装置などが含まれる。

【0049】

図1に示す実施形態では、顧客V C N 104は2つのサブネット、すなわち「サブネット1」と「サブネット2」で構成され、各サブネットは独自のC I D R I Pアドレス範囲を有する。図1では、サブネット1のオーバーレイI Pアドレス範囲は10.0/16、サブネット2のアドレス範囲は10.1/16である。V C N仮想ルータ105は、V C N 104のサブネット間およびV C N外部の他のエンドポイントとの通信を可能にするV C Nの論理ゲートウェイを表す。V C N V R 105は、V C N 104内のV N I CとV C N 104に関連付けられるゲートウェイ間のトラフィックをルーティングするように構成されている。V C N V R 105は、V C N 104の各サブネットにポートを提供する。例えば、V R 105は、サブネット1にI Pアドレス10.0.0.1のポートを提供し、サブネット2にI Pアドレス10.1.0.1のポートを提供できる。

【0050】

各サブネットには複数のコンピューティングインスタンスが展開され得、コンピューティングインスタンスは仮想マシンインスタンスおよび/またはベアメタルインスタンスであり得る。サブネット内のコンピューティングインスタンスは、C S P I 101内の1つまたは複数のホストマシンによってホストされ得る。コンピューティングインスタンスは、コンピューティングインスタンスに関連付けられているV N I Cを介してサブネットに参加する。例えば、図1に示すように、コンピューティングインスタンスC 1は、コンピューティングインスタンスに関連付けられているV N I Cを介してサブネット1の一部になる。同様に、コンピューティングインスタンスC 2は、C 2に関連付けられているV N I Cを介してサブネット1の一部になる。同様に、仮想マシンインスタンスまたはベアメタルインスタンスである可能性のある複数のコンピューティングインスタンスがサブネット1の一部になり得る。各コンピューティングインスタンスには、関連のV N I Cを介して、プライベートオーバーレイI PアドレスとM A Cアドレスが割り当てられる。例えば、図1では、コンピューティングインスタンスC 1のオーバーレイI Pアドレスは10.0.0.2、M A CアドレスはM 1であるが、コンピューティングインスタンスC 2のプ

プライベートオーバーレイIPアドレスは10.0.0.3、MACアドレスはM2である。コンピューティングインスタンスC1およびC2を含むサブネット1の各コンピューティングインスタンスには、IPアドレス10.0.0.1を使用してVCN VR105へのデフォルトルートがある。これは、サブネット1のVCN VR105のポートのIPアドレスである。

【0051】

サブネット2には、仮想マシンインスタンスおよび／またはベアメタルインスタンスなど、複数のコンピューティングインスタンスを展開できる。例えば、図1に示すように、コンピューティングインスタンスD1とD2は、それぞれのコンピューティングインスタンスに関連付けられているVNICを介してサブネット2の一部になる。図1に示す実施形態では、コンピューティングインスタンスD1はオーバーレイIPアドレス10.1.0.2とMACアドレスMM1を持ち、コンピューティングインスタンスD2はプライベートオーバーレイIPアドレス10.1.0.3とMACアドレスMM2を持つ。サブネット2内の各コンピューティングインスタンス（コンピューティングインスタンスD1およびD2を含む）には、IPアドレス10.1.0.1を使用するVCN VR105へのデフォルトルートがある。これは、サブネット2のVCN VR105のポートのIPアドレスである。

10

【0052】

VCN A104には、1つまたは複数の負荷分散装置も含まれ得る。例えば、サブネットに負荷分散装置が提供され、サブネット上の複数のコンピューティングインスタンス間でトラフィックを負荷分散するように構成され得る。VCN内のサブネット間でトラフィックの負荷を分散するために、負荷分散装置が提供され得る。

20

【0053】

VCN104上に展開される特定のコンピューティングインスタンスは、さまざまな異なるエンドポイントと通信できる。これらのエンドポイントには、CSP1200によってホストされるエンドポイントとCSP1200外部のエンドポイントが含まれ得る。CSP101によってホストされるエンドポイントには、特定のコンピューティングインスタンスと同じサブネット上のエンドポイント（例えば、サブネット1内の2つのコンピューティングインスタンス間の通信）、異なるサブネット上にあるが同じVCN内のエンドポイント（例えば、サブネット1内のコンピューティングインスタンスとサブネット2内のコンピューティングインスタンス間の通信）、同じリージョン内の異なるVCN内のエンドポイント（例えば、サブネット1内のコンピューティングインスタンスと同じリージョン106または110内のVCN内のエンドポイント間の通信、サブネット1内のコンピューティングインスタンスと同じリージョン内のサービスネットワーク110内のエンドポイント間の通信）、または異なるリージョン内のVCN内のエンドポイント（例えば、サブネット1内のコンピューティングインスタンスと異なるリージョン108内のVCN内のエンドポイント間の通信）が含まれ得る。CSP101によってホストされているサブネット内のコンピューティングインスタンスは、CSP101によってホストされていない（つまり、CSP101の外部にある）エンドポイントとも通信できる。これらの外部エンドポイントには、顧客のオンプレミスネットワーク116内のエンドポイント、他のリモートクラウドホストネットワーク118内のエンドポイント、インターネットなどのパブリックネットワーク経由でアクセス可能なパブリックエンドポイント114、およびその他のエンドポイントが含まれる。

30

40

【0054】

同じサブネット上のコンピューティングインスタンス間の通信は、ソースコンピューティングインスタンスと宛先コンピューティングインスタンスに関連付けられているVNICを使用して容易になる。例えば、サブネット1のコンピューティングインスタンスC1は、サブネット1のコンピューティングインスタンスC2にパケットを送信したい場合がある。ソースコンピューティングインスタンスから発信され、宛先が同じサブネット内の別のコンピューティングインスタンスであるパケットの場合、パケットは最初にソースコ

50

ンピューティングインスタンスに関連付けられているVNICによって処理される。ソースコンピューティングインスタンスに関連付けられているVNICによって実行される処理には、パケットヘッダーからパケットの宛先情報を決定すること、ソースコンピューティングインスタンスに関連付けられているVNICに構成されているポリシー（例えば、セキュリティリスト）を識別すること、パケットの次のホップを決定すること、必要に応じてパケットのカプセル化／カプセル化解除機能を実行すること、そして、パケットを次のホップに転送／ルーティングして、パケットを目的の宛先に容易に通信するようにすることが含まれ得る。宛先コンピューティングインスタンスがソースコンピューティングインスタンスと同じサブネットにある場合、ソースコンピューティングインスタンスに関連付けられているVNICは、宛先コンピューティングインスタンスに関連付けられているVNICを識別し、パケットをそのVNICに転送して処理するように構成される。次に、宛先コンピューティングインスタンスに関連付けられているVNICが実行され、パケットが宛先コンピューティングインスタンスに転送される。

10

【0055】

サブネット内のコンピューティングインスタンスから同じVCN内の別のサブネット内のエンドポイントにパケットを通信する場合、その通信は、ソースおよび宛先のコンピューティングインスタンスとVCN VRに関連付けられているVNICによって容易になる。例えば、図1のサブネット1内のコンピューティングインスタンスC1がサブネット2内のコンピューティングインスタンスD1にパケットを送信したい場合、パケットは、まずコンピューティングインスタンスC1に関連付けられているVNICによって処理される。コンピューティングインスタンスC1に関連付けられているVNICは、デフォルトルートまたはVCN VRのポート10.0.0.1を使用してパケットをVCN VR 105にルーティングするように構成されている。VCN VR 105は、ポート10.1.0.1を使用してパケットをサブネット2にルーティングするように構成されている。次に、パケットはD1に関連付けられているVNICによって受信および処理され、VNICはパケットをコンピューティングインスタンスD1に転送する。

20

【0056】

VCN 104内のコンピューティングインスタンスからVCN 104外のエンドポイントにパケットを通信する場合、その通信は、ソースコンピューティングインスタンスに関連付けられているVNIC、VCN VR 105、およびVCN 104に関連付けられるゲートウェイによって容易になる。VCN 104には、1つまたは複数のタイプのゲートウェイが関連付けられ得る。ゲートウェイは、VCNと別のエンドポイント間のインターフェースであり、別のエンドポイントはVCNの外部にある。ゲートウェイは層3／IP層の概念であり、VCNがVCN外部のエンドポイントと通信できるようにする。したがって、ゲートウェイは、VCNと他のVCNまたはネットワーク間のトラフィックフローを容易にする。さまざまなタイプのエンドポイントとのさまざまなタイプの通信を容易にするために、VCNにさまざまなタイプのゲートウェイを構成できる。ゲートウェイに応じて、通信はパブリックネットワーク（例えば、インターネット）経由またはプライベートネットワーク経由で行われ得る。これらの通信にはさまざまな通信プロトコルが使用され得る。

30

40

【0057】

例えば、コンピューティングインスタンスC1は、VCN 104の外部のエンドポイントと通信したい場合がある。パケットは、まずソースコンピューティングインスタンスC1に関連付けられているVNICによって処理され得る。VNIC処理により、パケットの宛先がC1のサブネット1の外部にあることが決定される。C1に関連付けられているVNICは、パケットをVCN 104のVCN VR 105に転送し得る。次に、VCN VR 105はパケットを処理し、処理の一環として、パケットの宛先に基づいて、VCN 104に関連付けられる特定のゲートウェイをパケットの次のホップとして決定する。その後、VCN VR 105は、パケットを特定の識別されるゲートウェイに転送し得る。例えば、宛先が顧客のオンプレミスネットワーク内のエンドポイントである場合、パケッ

50

トはVCN VR105によって、VCN104用に構成される動的ルーティングゲートウェイ(DRG)ゲートウェイ122に転送され得る。その後、パケットはゲートウェイから次のホップに転送され、最終的な目的の宛先へのパケットの通信が容易になり得る。

【0058】

VCNには、さまざまなタイプのゲートウェイを構成できる。VCN用に構成できるゲートウェイの例を図1に示し、以下に説明する。VCNに関連付けられるゲートウェイの例も、図12、13、14、および15に示されており(例えば、参照番号1234、1236、1238、1334、1336、1338、1434、1436、1438、1534、1536、および1538で参照されるゲートウェイ)、以下で説明する。図1に示す実施形態に示すように、動的ルーティングゲートウェイ(DRG)122は、顧客VCN104に追加されるか、または顧客VCN104に関連付けられ得、顧客VCN104と別のエンドポイントとの間のプライベートネットワークトラフィック通信のパスを提供する。ここで、別のエンドポイントは、顧客のオンプレミスネットワーク116、CSP101の別のリージョン内のVCN108、またはCSP101によってホストされていない他のリモートクラウドネットワーク118であり得る。顧客オンプレミスネットワーク116は、顧客のリソースを使用して構築される顧客ネットワークまたは顧客データセンターであり得る。顧客のオンプレミスネットワーク116へのアクセスは、通常、非常に制限されている。顧客オンプレミスネットワーク116と、CSP101によってクラウドに展開またはホストされている1つまたは複数のVCN104の両方を持つ顧客の場合、顧客はオンプレミスネットワーク116とクラウドベースのVCN104が相互に通信できるようにしたい場合がある。これにより、顧客は、CSP101によってホストされる顧客のVCN104とオンプレミスネットワーク116を含む拡張ハイブリッド環境を構築できるようになる。DRG122はこの通信を可能にする。このような通信を可能にするために、通信チャンネル124が設定され、チャンネルの1つのエンドポイントは顧客のオンプレミスネットワーク116にあり、もう1つのエンドポイントはCSP101にあり、顧客のVCN104に接続されている。通信チャンネル124は、インターネットなどのパブリック通信ネットワークまたはプライベート通信ネットワークを介することができる。インターネットなどのパブリック通信ネットワークを介したIPsecVPN技術や、パブリックネットワークの代わりにプライベートネットワークを使用するOracleのFastConnect技術など、さまざまな通信プロトコルを使用できる。通信チャンネル124の1つのエンドポイントを形成する顧客オンプレミスネットワーク116内の装置または機器は、図1に示すCPE126などの顧客構内機器(CPE)と呼ばれる。CSP101側では、エンドポイントはDRG122を実行するホストマシンであり得る。特定の実施形態では、リモートピアリング接続(RPC)をDRGに追加して、顧客が1つのVCNを別のリージョンの別のVCNとピアリングできるようにすることができる。このようなRPCを使用すると、顧客VCN104はDRG122を使用して別のリージョンのVCN108に接続できる。DRG122は、Microsoft Azureクラウド、Amazon AWSクラウドなど、CSP101でホストされていない他のリモートクラウドネットワーク118との通信にも使用できる。

【0059】

図1に示すように、顧客VCN104に対してインターネットゲートウェイ(IGW)120が構成され、VCN104上のコンピューティングインスタンスがインターネットなどのパブリックネットワークを介してアクセス可能なパブリックエンドポイント114と通信できるようになる。IGW120は、VCNをインターネットなどのパブリックネットワークに接続するゲートウェイである。IGW120は、VCN104などのVCN内のパブリックサブネット(パブリックサブネット内のリソースがパブリックオーバーレイIPアドレスを持つ)が、インターネットなどのパブリックネットワーク114上のパブリックエンドポイント112に直接アクセスできるようにする。IGW120を使用すると、VCN104内のサブネットまたはインターネットから接続を開始できる。

【0060】

ネットワークアドレス変換（NAT）ゲートウェイ128は、顧客のVCN104用に構成することができ、専用のパブリックオーバーレイIPアドレスを持たない顧客のVCN内のクラウドリソースがインターネットにアクセスできるようにし、これらのリソースを直接の着信インターネット接続（例えば、L4-L7接続）に公開することなく、これを実現する。これにより、VCN104のプライベートサブネット1など、VCN内のプライベートサブネットがインターネット上のパブリックエンドポイントにプライベートアクセスできるようになる。NATゲートウェイでは、プライベートサブネットからパブリックインターネットへの接続のみを開始でき、インターネットからプライベートサブネットへの接続は開始できない。

【0061】

特定の実施形態では、サービスゲートウェイ（SGW）126を顧客VCN104用に構成することができ、VCN104とサービスネットワーク110内のサポートされているサービスエンドポイントとの間のプライベートネットワークトラフィックのパスを提供する。特定の実施形態では、サービスネットワーク110はCSPによって提供され、さまざまなサービスを提供し得る。このようなサービスネットワークの例としては、顧客が利用できるさまざまなサービスを提供するOracleのサービスネットワークがある。例えば、顧客VCN104のプライベートサブネット内のコンピューティングインスタンス（例えば、データベースシステム）は、パブリックIPアドレスやインターネットへのアクセスを必要とせずに、サービスエンドポイント（例えば、オブジェクト記憶装置）にデータをバックアップできる。特定の実施形態では、VCNは1つのSGWのみを持つことができ、接続はサービスネットワーク110からではなく、VCN内のサブネットからのみ開始できる。VCNが別のVCNとピアリングされている場合、通常、他のVCNのリソースはSGWにアクセスできない。FastConnectまたはVPN接続を使用してVCNに接続されているオンプレミスネットワーク内のリソースも、そのVCN用に構成されるサービスゲートウェイを使用できる。

【0062】

特定の実装では、SGW126は、サービスクラスレスインタードメインルーティング（CIDR）ラベルの概念を使用する。これは、対象のサービスまたはサービスグループのすべてのリージョンのパブリックIPアドレス範囲を表す文字列である。顧客は、サービスへのトラフィックを制御するためにSGWおよび関連ルートルールを構成するときに、サービスCIDRラベルを使用する。顧客は、将来サービスのパブリックIPアドレスが変更される場合でもセキュリティルールを調整する必要なく、セキュリティルールを構成するときに任意選択でこれを利用できる。

【0063】

局所ピアリングゲートウェイ（LPG）132は、顧客VCN104に追加できるゲートウェイであり、VCN104が同じリージョン内の別のVCNとピアリングできるようにする。ピアリングとは、トラフィックがインターネットなどのパブリックネットワークを通過したり、顧客のオンプレミスネットワーク116を経由してトラフィックをルーティングしたりすることなく、VCNがプライベートIPアドレスを使用して通信することを意味する。好ましい実施形態では、VCNは、確立するピアリングごとに個別のLPGを持つ。局所ピアリングまたはVCNピアリングは、異なるアプリケーションまたはインフラストラクチャ管理機能間のネットワーク接続を確立するために使用される一般的な方法である。

【0064】

サービスネットワーク110内のサービスのプロバイダなどのサービスプロバイダは、異なるアクセスモデルを使用してサービスへのアクセスを提供することができる。パブリックアクセスモデルに従って、サービスは、インターネットなどのパブリックネットワークを介して顧客VCN内のコンピューティングインスタンスによってパブリックにアクセス可能なパブリックエンドポイントとして公開される場合もあれば、SGW126を介してプライベートにアクセス可能な場合もある。特定のプライベートアクセスモデルに従っ

10

20

30

40

50

て、サービスは顧客の V C N 内のプライベートサブネット内のプライベート I P エンドポイントとしてアクセス可能になる。これはプライベートエンドポイント (P E) アクセスと呼ばれ、サービスプロバイダが顧客のプライベートネットワーク内のインスタンスとしてサービスを公開できるようにする。プライベートエンドポイントリソースは、顧客の V C N 内のサービスを表す。各 P E は、顧客の V C N 内で顧客が選択したサブネット内の V N I C (1 つまたは複数のプライベート I P を持つ P E - V N I C と呼ばれる) として現れる。したがって、 P E は、 V N I C を使用してプライベート顧客 V C N サブネット内でサービスを提供する方法を提供する。エンドポイントは V N I C として公開されるため、ルーティングルール、セキュリティリストなど、 V N I C に関連付けられているすべての機能が P E V N I C で使用できるようになる。

10

【 0 0 6 5 】

サービスプロバイダは、 P E を介したアクセスを可能にするためにサービスを登録することができる。プロバイダは、サービスの可視性を顧客テナントに制限するポリシーをサービスに関連付けることができる。プロバイダは、特にマルチテナントサービスの場合、単一の仮想 I P アドレス (V I P) の下に複数のサービスを登録できる。同じサービスを表すこのようなプライベートエンドポイントが (複数の V C N に) 複数存在し得る。

【 0 0 6 6 】

プライベートサブネット内のコンピューティングインスタンスは、次いで、 P E V N I C のプライベート I P アドレスまたはサービス D N S 名を使用してサービスにアクセスできる。顧客 V C N 内のコンピューティングインスタンスは、顧客 V C N 内の P E のプライベート I P アドレスにトラフィックを送信することでサービスにアクセスできる。プライベートアクセスゲートウェイ (P A G W) 1 3 0 は、サービスプロバイダ V C N (例えば、サービスネットワーク 1 1 0 内の V C N) に接続できるゲートウェイリソースであり、顧客サブネットのプライベートエンドポイントとの間のすべてのトラフィックの入口／出口ポイントとして機能する。 P A G W 1 3 0 を使用すると、プロバイダは内部 I P アドレスリソースを利用せずに P E 接続の数を拡張できる。プロバイダは、単一の V C N に登録されている任意の数のサービスに対して 1 つの P A G W のみを構成する必要がある。プロバイダは、1 人または複数の顧客の複数の V C N 内のプライベートエンドポイントとしてサービスを表すことができる。顧客の観点から見ると、 P E V N I C は顧客のインスタンスに接続されているのではなく、顧客が対話したいサービスに接続されているように見える。プライベートエンドポイント宛てのトラフィックは、 P A G W 1 3 0 経由でサービスにルーティングされる。これらは、顧客からサービスへのプライベート接続 (C 2 S 接続) と呼ばれる。

20

30

【 0 0 6 7 】

P E 概念は、トラフィックが FastConnect / I P s e c リンクと顧客 V C N 内のプライベートエンドポイントを通過できるようにすることで、サービスのプライベートアクセスを顧客のオンプレミスネットワークとデータセンターに拡張するためにも使用できる。 L P G 1 3 2 と顧客の V C N 内の P E の間でトラフィックを流すことによって、サービスのプライベートアクセスを顧客のピア V C N に拡張することもできる。

【 0 0 6 8 】

顧客はサブネットレベルで V C N 内のルーティングを制御できるため、 V C N 1 0 4 などの顧客の V C N 内のどのサブネットが各ゲートウェイを使用するかを指定できる。 V C N のルートテーブルは、特定のゲートウェイを介して V C N からのトラフィックを許可するかどうかを決定するために使用される。例えば、特定のインスタンスでは、顧客 V C N 1 0 4 内のパブリックサブネットのルートテーブルは、 I G W 1 2 0 を介して非局所トラフィックを送信し得る。同じ顧客 V C N 1 0 4 内のプライベートサブネットのルートテーブルは、 C S P サービス宛てのトラフィックを S G W 1 2 6 を介して送信し得る。残りのトラフィックはすべて N A T ゲートウェイ 1 2 8 経由で送信され得る。ルートテーブルは、 V C N から送信されるトラフィックのみを制御する。

40

【 0 0 6 9 】

50

V C Nに関連付けられるセキュリティリストは、ゲートウェイを介して受信接続でV C Nに入るトラフィックを制御するために使用される。サブネット内のすべてのリソースは同じルートテーブルとセキュリティリストを使用する。セキュリティリストは、V C Nのサブネット内のインスタンスに出入りできる特定のタイプのトラフィックを制御するために使用できる。セキュリティリストルールは、入口（受信）ルールと出口（送信）ルールを含み得る。例えば、入口ルールでは許可されるソースアドレス範囲を指定し、出口ルールでは許可される宛先アドレス範囲を指定することができる。セキュリティルールでは、特定のプロトコル（例えば、T C P、I C M P）、特定のポート（例えば、S S Hの場合は2 2、Windows RDPの場合は3 3 8 9）などを指定できる。特定の実装では、インスタンスのオペレーティングシステムが、セキュリティリストルールに準拠した独自のファイアウォールルールを適用し得る。ルールはステートフル（例えば、接続が追跡され、応答トラフィックに対する明示的なセキュリティリストルールがなくても応答が自動的に許可される）またはステートレスになり得る。

10

【0 0 7 0】

顧客V C Nからのアクセス（すなわち、V C N 1 0 4上に展開されるリソースまたはコンピューティングインスタンスによるアクセス）は、パブリックアクセス、プライベートアクセス、または専用アクセスに分類できる。パブリックアクセスとは、パブリックIPアドレスまたはN A Tを使用してパブリックエンドポイントにアクセスするアクセスモデルを指す。プライベートアクセスにより、プライベートIPアドレス（例えば、プライベートサブネット内のリソース）を持つV C N 1 0 4内の顧客ワークロードは、インターネットなどのパブリックネットワークを経由せずにサービスにアクセスできるようになる。特定の実施形態では、C S P I 1 0 1により、プライベートIPアドレスを持つ顧客V C Nワークロードがサービスゲートウェイを使用してサービス（のパブリックサービスエンドポイント）にアクセスできるようになる。したがって、サービスゲートウェイは、顧客のV C Nと顧客のプライベートネットワークの外部にあるサービスのパブリックエンドポイントとの間に仮想リンクを確立することにより、プライベートアクセスモデルを提供する。

20

【0 0 7 1】

さらに、C S P Iは、FastConnectパブリックピアリングなどの技術を使用して専用のパブリックアクセスを提供し得る。顧客のオンプレミスインスタンスは、FastConnect接続を使用して、インターネットなどのパブリックネットワークを経由せずに、顧客V C N内の1つまたは複数のサービスにアクセスできる。C S P Iは、FastConnectプライベートピアリングを使用した専用プライベートアクセスも提供し得、プライベートIPアドレスを持つ顧客のオンプレミスインスタンスがFastConnect接続を使用して顧客のV C Nワークロードにアクセスできるようになる。FastConnectは、パブリックインターネットを使用して顧客のオンプレミスネットワークをC S P Iおよびそのサービスに接続するためのネットワーク接続の代替手段である。FastConnectは、インターネットベースの接続と比較して、より高い帯域幅任意選択肢と、より信頼性が高く一貫性のあるネットワークエクスペリエンスを備えた専用のプライベート接続を作成するための、簡単、柔軟、かつ経済的な方法を提供する。

30

40

【0 0 7 2】

図1および上記の付随する説明は、例示的な仮想ネットワーク内のさまざまな仮想化される構成要素について説明している。前述のように、仮想ネットワークは、基盤となる物理ネットワークまたはサブストレートネットワーク上に構築される。図2は、特定の実施形態による仮想ネットワークのアンダーレイを提供するC S P I 2 0 0内の物理ネットワーク内の物理構成要素の簡略化されるアーキテクチャ図を示している。図に示すように、C S P I 2 0 0は、クラウドサービスプロバイダ（C S P）によって提供される構成要素とリソース（例えば、コンピューティング、メモリ、ネットワークリソース）で構成される分散環境を提供する。これらの構成要素とリソースは、加入している顧客、つまりC S Pが提供する1つまたは複数のサービスに加入している顧客にクラウドサービス（例えば

50

、IaaSサービス）を提供するために使用される。顧客が加入しているサービスに基づいて、CSP I200のリソースのサブセット（例えば、コンピューティング、メモリ、ネットワークリソース）が顧客に対してプロビジョニングされる。顧客は、次いで、CSP I200が提供する物理的なコンピューティング、メモリ、およびネットワークリソースを使用して、独自のクラウドベース（つまり、CSP Iホスト）のカスタマイズ可能なプライベート仮想ネットワークを構築できる。前述のように、これらの顧客ネットワークは仮想クラウドネットワーク（VCN）と呼ばれる。顧客は、コンピューティングインスタンスなどの1つまたは複数の顧客リソースをこれらの顧客VCNに展開できる。コンピューティングインスタンスは、仮想マシン、ベアメタルインスタンスなどの形式になり得る。CSP I200は、顧客が可用性の高いホスト環境で幅広いアプリケーションとサービスを構築および実行できるようにするインフラストラクチャと補完的なクラウドサービスのセットを提供する。

10

【0073】

図2に示す実施例では、CSP I200の物理構成要素には、1つまたは複数の物理ホストマシンまたは物理サーバ（例えば、202、206、208）、ネットワーク仮想化装置（NVD）（例えば、210、212）、トップオブブラック（TOR）スイッチ（例えば、214、216）、物理ネットワーク（例えば、218）、および物理ネットワーク218内のスイッチが含まれる。物理ホストマシンまたはサーバは、VCNの1つまたは複数のサブネットに参加するさまざまなコンピューティングインスタンスをホストおよび実行できる。コンピューティングインスタンスには、仮想マシンインスタンスおよびベアメタルインスタンスが含まれ得る。例えば、図1に示されているさまざまなコンピューティングインスタンスは、図2に示されている物理ホストマシンによってホストされ得る。VCN内の仮想マシンコンピューティングインスタンスは、1台のホストマシンまたは複数の異なるホストマシンによって実行され得る。物理ホストマシンは、仮想ホストマシン、コンテナベースのホストまたは機能などもホストできる。図1に示されているVNICおよびVCN VRは、図2に示されているNVDによって実行され得る。図1に示すゲートウェイは、図2に示すホストマシンおよび／またはNVDによって実行され得る。

20

【0074】

ホストマシンまたはサーバは、ホストマシン上に仮想化環境を作成して有効にするハイパーバイザ（仮想マシンモニタまたはVMMとも呼ばれる）を実行することができる。仮想化または仮想化環境により、クラウドベースのコンピューティングが容易になる。ホストマシン上のハイパーバイザによって、ホストマシン上で1つまたは複数のコンピューティングインスタンスが作成、実行、管理され得る。ホストマシン上のハイパーバイザにより、ホストマシンの物理的なコンピューティングリソース（例えば、コンピューティング、メモリ、ネットワークリソース）を、ホストマシンによって実行されるさまざまなコンピューティングインスタンス間で共有できるようになる。

30

【0075】

例えば、図2に示すように、ホストマシン202および208は、それぞれハイパーバイザ260および266を実行する。これらのハイパーバイザは、ソフトウェア、ファームウェア、ハードウェア、またはそれらの組み合わせを使用して実装できる。通常、ハイパーバイザは、ホストマシンのオペレーティングシステム（OS）の上に配置され、次いで、ホストマシンのハードウェアプロセッサ上で実行されるプロセスまたはソフトウェア層である。ハイパーバイザは、ホストマシンの物理コンピューティングリソース（例えば、プロセッサ／コアなどの処理リソース、メモリリソース、ネットワークリソース）を、ホストマシンによって実行されるさまざまな仮想マシンコンピューティングインスタンス間で共有できるようにすることで、仮想化環境を提供する。例えば、図2では、ハイパーバイザ260はホストマシン202のOSの上に配置され、ホストマシン202のコンピューティングリソース（例えば、処理、メモリ、およびネットワークリソース）を、ホストマシン202によって実行されるコンピューティングインスタンス（例えば、仮想マシン）間で共有できるようにする。仮想マシンには独自のオペレーティングシステム（ゲス

40

50

トオペレーティングシステムと呼ばれる)を持たせることができ、これはホストマシンのOSと同じでも異なってもかまわない。ホストマシンによって実行される仮想マシンのオペレーティングシステムは、同じホストマシンによって実行される別の仮想マシンのオペレーティングシステムと同じでも異なってもかまわない。したがって、ハイパーバイザを使用すると、ホストマシンの同じコンピューティングリソースを共有しながら、複数のオペレーティングシステムを相互に並行して実行できるようになる。図2に示されているホストマシンには、同じタイプのハイパーバイザが搭載されている場合も、異なるタイプのハイパーバイザが搭載されている場合もある。

【0076】

コンピューティングインスタンスは、仮想マシンインスタンスまたはベアメタルインスタンスであり得る。図2では、ホストマシン202上のコンピューティングインスタンス268とホストマシン208上のコンピューティングインスタンス274が仮想マシンインスタンスの例である。ホストマシン206は、顧客に提供されるベアメタルインスタンスの例である。

【0077】

場合によっては、ホストマシン全体が単一の顧客にプロビジョニングされ得、そのホストマシンによってホストされる1つまたは複数のコンピューティングインスタンス(仮想マシンまたはベアメタルインスタンス)のすべてが同じ顧客に属する。他の場合には、ホストマシンが複数の顧客(つまり、複数のテナント)間で共有され得る。このようなマルチテナントシナリオでは、ホストマシンは異なる顧客に属する仮想マシンコンピューティングインスタンスをホストし得る。これらのコンピューティングインスタンスは、異なる顧客の異なるVCNのメンバーであり得る。特定の実施形態では、ベアメタルコンピューティングインスタンスは、ハイパーバイザのないベアメタルサーバによってホストされる。ベアメタルコンピューティングインスタンスがプロビジョニングされると、単一の顧客またはテナントが、ベアメタルインスタンスをホストするホストマシンの物理CPU、メモリ、およびネットワークインターフェースの制御を維持し、ホストマシンは他の顧客またはテナントと共有されない。

【0078】

前述のように、VCNの一部である各コンピューティングインスタンスは、コンピューティングインスタンスがVCNのサブネットのメンバーになることを可能にするVNICに関連付けられている。コンピューティングインスタンスに関連付けられているVNICは、コンピューティングインスタンスとの間のパケットまたはフレームの通信を容易にする。VNICは、コンピューティングインスタンスが作成されるときにコンピューティングインスタンスに関連付けられる。特定の実施形態では、ホストマシンによって実行されるコンピューティングインスタンスの場合、そのコンピューティングインスタンスに関連付けられているVNICは、ホストマシンに接続されるNVDによって実行される。例えば、図2では、ホストマシン202は、VNIC276に関連付けられる仮想マシンコンピューティングインスタンス268を実行し、VNIC276は、ホストマシン202に接続されるNVD210によって実行される。別の例として、ホストマシン206によってホストされるベアメタルインスタンス272は、ホストマシン206に接続されるNVD212によって実行されるVNIC280に関連付けられる。さらに別の例として、VNIC284はホストマシン208によって実行されるコンピューティングインスタンス274に関連付けられ、VNIC284はホストマシン208に接続されるNVD212によって実行される。

【0079】

ホストマシンによってホストされるコンピューティングインスタンスの場合、そのホストマシンに接続されるNVDは、コンピューティングインスタンスがメンバーであるVCNに対応するVCN VRも実行する。例えば、図2に示す実施形態では、NVD210は、コンピューティングインスタンス268がメンバーであるVCNに対応するVCN VR277を実行する。NVD212は、ホストマシン206および208によってホス

10

20

30

40

50

トされるコンピューティングインスタンスに対応する V C N に対応する 1 つまたは複数の V C N V R 2 8 3 を実行することもできる。

【0080】

ホストマシンには、ホストマシンを他の装置に接続できるようにする 1 つまたは複数のネットワークインターフェースカード (N I C) が含まれ得る。ホストマシン上の N I C は、ホストマシンを別の装置に通信可能に接続できるようにする 1 つまたは複数のポート (またはインターフェース) を提供し得る。例えば、ホストマシンは、ホストマシンと N V D に提供されている 1 つまたは複数のポート (またはインターフェース) を使用して N V D に接続できる。ホストマシンも、別のホストマシンなどの他の装置に接続され得る。

【0081】

例えば、図 2 では、ホストマシン 2 0 2 は、ホストマシン 2 0 2 の N I C 2 3 2 によって提供されるポート 2 3 4 と N V D 2 1 0 のポート 2 3 6 との間に延びるリンク 2 2 0 を使用して N V D 2 1 0 に接続されている。ホストマシン 2 0 6 は、ホストマシン 2 0 6 の N I C 2 4 4 によって提供されるポート 2 4 6 と N V D 2 1 2 のポート 2 4 8 との間に延びるリンク 2 2 4 を使用して N V D 2 1 2 に接続される。ホストマシン 2 0 8 は、ホストマシン 2 0 8 の N I C 2 5 0 によって提供されるポート 2 5 2 と N V D 2 1 2 のポート 2 5 4 との間に延びるリンク 2 2 6 を使用して N V D 2 1 2 に接続される。

【0082】

N V D は、次いで、通信リンクを介してトップオブザラック (T O R) スイッチに接続され、トップオブザラックスイッチは物理ネットワーク 2 1 8 (スイッチファブリックとも呼ばれる) に接続される。特定の実施形態では、ホストマシンと N V D 間のリンク、および N V D と T O R スイッチ間のリンクはイーサネット (登録商標) リンクである。例えば、図 2 では、N V D 2 1 0 および 2 1 2 は、リンク 2 2 8 および 2 3 0 を使用して、それぞれ T O R スイッチ 2 1 4 および 2 1 6 に接続されている。特定の実施形態では、リンク 2 2 0、2 2 4、2 2 6、2 2 8、および 2 3 0 はイーサネットリンクである。T O R に接続されるホストマシンと N V D の集合は、ラックと呼ばれることもある。

【0083】

物理ネットワーク 2 1 8 は、T O R スイッチが相互に通信できるようにする通信ファブリックを提供する。物理ネットワーク 2 1 8 は、多層ネットワークにすることができる。特定の実装では、物理ネットワーク 2 1 8 はスイッチの多層 C l o s ネットワークであり、T O R スイッチ 2 1 4 および 2 1 6 は多層およびマルチノードの物理スイッチングネットワーク 2 1 8 のリーフレベルノードを表す。2 層ネットワーク、3 層ネットワーク、4 層ネットワーク、5 層ネットワーク、および一般に「n」層ネットワークを含むがこれらに限定されない、さまざまな C l o s ネットワーク構成が可能である。C l o s ネットワークの例を図 5 に示し、以下に説明する。

【0084】

ホストマシンと N V D の間では、1 対 1 構成、多対 1 構成、1 対多構成など、さまざまな接続構成が可能である。1 対 1 構成の実装では、各ホストマシンは独自の個別の N V D に接続される。例えば、図 2 では、ホストマシン 2 0 2 は、ホストマシン 2 0 2 の N I C 2 3 2 を介して N V D 2 1 0 に接続されている。多対 1 構成では、複数のホストマシンが 1 つの N V D に接続される。例えば、図 2 では、ホストマシン 2 0 6 および 2 0 8 は、それぞれ N I C 2 4 4 および 2 5 0 を介して同じ N V D 2 1 2 に接続されている。

【0085】

1 対多構成では、1 台のホストマシンが複数の N V D に接続される。図 3 は、ホストマシンが複数の N V D に接続されている C S P I 3 0 0 内の例を示している。図 3 に示すように、ホストマシン 3 0 2 は、複数のポート 3 0 6 および 3 0 8 を含むネットワークインターフェースカード (N I C) 3 0 4 を備えている。ホストマシン 3 0 0 は、ポート 3 0 6 およびリンク 3 2 0 を介して第 1 の N V D 3 1 0 に接続され、ポート 3 0 8 およびリンク 3 2 2 を介して第 2 の N V D 3 1 2 に接続される。ポート 3 0 6 および 3 0 8 はイーサネットポートであり得、ホストマシン 3 0 2 と N V D 3 1 0 および 3 1 2 間のリンク 3 2

10

20

30

40

50

0 および 3 2 2 はイーサネットリンクであり得る。NVD 3 1 0 は、次いで、第 1 の TOR スイッチ 3 1 4 に接続され、NVD 3 1 2 は第 2 の TOR スイッチ 3 1 6 に接続される。NVD 3 1 0 と 3 1 2 および TOR スイッチ 3 1 4 と 3 1 6 間のリンクは、イーサネットリンクであり得る。TOR スイッチ 3 1 4 および 3 1 6 は、多層物理ネットワーク 3 1 8 内の層 0 スwitchング装置を表す。

【0086】

図 3 に示す配置は、物理スイッチネットワーク 3 1 8 からホストマシン 3 0 2 への 2 つの別個の物理ネットワークパスを提供する。第 1 のパスは、TOR スイッチ 3 1 4 を経由して NVD 3 1 0 を経由してホストマシン 3 0 2 へ、第 2 のパスは TOR スイッチ 3 1 6 を経由して NVD 3 1 2 を経由してホストマシン 3 0 2 へである。別々のパスにより、ホストマシン 3 0 2 の可用性が向上する（高可用性と呼ばれる）。いずれかのパスに問題がある場合（例えば、いずれかのパスのリンクがダウンしている場合）または装置に問題がある場合（例えば、特定の NVD が機能していない場合）、ホストマシン 3 0 2 との通信に他のパスが使用され得る。

【0087】

図 3 に示す構成では、ホストマシンは、ホストマシンの NIC によって提供される 2 つの異なるポートを使用して 2 つの異なる NVD に接続される。他の実施形態では、ホストマシンには、ホストマシンを複数の NVD に接続できるようにする複数の NIC が含まれ得る。

【0088】

図 2 に戻ると、NVD は、1 つまたは複数のネットワークおよび／または記憶仮想化機能を実行する物理装置または構成要素である。NVD は、1 つまたは複数の処理装置（例えば、CPU、ネットワーク処理装置（NPU）、FPGA、パケット処理パイプライン）、キャッシュを含むメモリ、およびポートを備えた任意の装置であり得る。さまざまな仮想化機能は、NVD の 1 つまたは複数の処理装置によって実行されるソフトウェア／ファームウェアによって実行され得る。

【0089】

NVD は、さまざまな異なる形式で実装することができる。例えば、特定の実施形態では、NVD は、オンボードに組み込みプロセッサを備えたスマート NIC またはインテリジェント NIC と呼ばれるインターフェースカードとして実装される。スマート NIC は、ホストマシン上の NIC とは別の装置である。図 2 では、NVD 2 1 0 および 2 1 2 は、それぞれホストマシン 2 0 2、およびホストマシン 2 0 6 および 2 0 8 に接続されるスマート NIC として実装され得る。

【0090】

しかし、スマート NIC は NVD 実装の一例にすぎない。他にもさまざまな実装が可能である。例えば、他のいくつかの実装では、NVD または NVD によって実行される 1 つまたは複数の機能は、1 つまたは複数のホストマシン、1 つまたは複数の TOR スイッチ、および C S P I 2 0 0 の他の構成要素に組み込まれるか、またはそれらによって実行され得る。例えば、NVD はホストマシンに実装され得、NVD によって実行される機能はホストマシンによって実行される。別の例として、NVD は TOR スイッチの一部であり得るか、TOR スイッチが NVD によって実行される機能を実行するように構成され、TOR スイッチがパブリッククラウドに使用されるさまざまな複雑なパケット変換を実行できるようにし得る。NVD の機能を実行する TOR は、スマート TOR と呼ばれることもある。さらに他の実装では、ベアメタル（BM）インスタンスではなく仮想マシン（VM）インスタンスが顧客に提供され、NVD によって実行される機能は、ホストマシンのハイパーバイザ内に実装され得る。いくつかの他の実装では、NVD の機能の一部が、一連のホストマシン上で実行される集中型サービスにオフロードされ得る。

【0091】

図 2 に示すようにスマート NIC として実装される場合など、特定の実施形態では、NVD は、1 つまたは複数のホストマシンおよび 1 つまたは複数の TOR スイッチに接続で

きるようにする複数の物理ポートを備えることができる。NVD上のポートは、ホスト側ポート（「サウスポート」とも呼ばれる）、またはネットワーク側またはTOR側ポート（「ノースポート」とも呼ばれる）に分類できる。NVDのホスト側ポートは、NVDをホストマシンに接続するために使用されるポートである。図2のホスト側ポートの例には、NVD210のポート236、NVD212のポート248および254が含まれる。NVDのネットワーク側ポートは、NVDをTORスイッチに接続するために使用されるポートである。図2のネットワーク側ポートの例には、NVD210のポート256およびNVD212のポート258が含まれる。図2に示すように、NVD210は、NVD210のポート256からTORスイッチ214まで延びるリンク228を使用してTORスイッチ214に接続される。同様に、NVD212は、NVD212のポート258からTORスイッチ216まで延びるリンク230を使用してTORスイッチ216に接続される。

10

【0092】

NVDは、ホスト側ポートを介してホストマシンからパケットおよびフレーム（例えば、ホストマシンによってホストされるコンピューティングインスタンスによって生成されるパケットおよびフレーム）を受信し、必要なパケット処理を実行した後、NVDのネットワーク側ポートを介してパケットおよびフレームをTORスイッチに転送することができる。NVDは、NVDのネットワーク側ポートを介してTORスイッチからパケットとフレームを受信し、必要なパケット処理を実行した後、NVDのホスト側ポートを介してパケットとフレームをホストマシンに転送し得る。

20

【0093】

特定の実施形態では、NVDとTORスイッチの間に複数のポートと関連するリンクが存在し得る。これらのポートとリンクは集約され、複数のポートまたはリンクのリンクアグリゲータグループ（LAGと呼ばれる）を形成し得る。リンク集約により、2つのエンドポイント間（例えば、NVDとTORスイッチ間）の複数の物理リンクを1つの論理リンクとして扱うことができる。特定のLAG内のすべての物理リンクは、同じ速度で全二重モードで動作し得る。LAGは、2つのエンドポイント間の接続の帯域幅と信頼性を向上させるのに役立つ。LAG内の物理リンクの1つがダウンした場合、トラフィックはLAG内の他の物理リンクの1つに動的かつ透過的に再割り当てされる。集約される物理リンクは、個々のリンクよりも高い帯域幅を提供する。LAGに関連付けられる複数のポートは、単一の論理ポートとして扱われる。トラフィックは、LAGの複数の物理リンク間で負荷分散できる。2つのエンドポイント間に1つまたは複数のLAGを構成できる。2つのエンドポイントは、NVDとTORスイッチの間、ホストマシンとNVDの間などにあり得る。

30

【0094】

NVDは、ネットワーク仮想化機能を実装または実行する。これらの機能は、NVDによって実行されるソフトウェア／ファームウェアによって実行される。ネットワーク仮想化機能の例には、パケットのカプセル化およびカプセル化解除機能、VCNネットワークを作成するための機能、VCNセキュリティリスト（ファイアウォール）機能などのネットワークポリシーを実装するための機能、VCN内のコンピューティングインスタンスとの間のパケットのルーティングおよび転送を容易にする機能などが含まれるが、これらに限定されない。特定の実施形態では、パケットを受信すると、NVDは、パケットを処理し、パケットをどのように転送またはルーティングするかを決定するためのパケット処理パイプラインを実行するように構成される。このパケット処理パイプラインの一部として、NVDは、VCN内のcisに関連付けられているVNICの実行、VCNに関連付けられる仮想ルーター（VR）の実行、仮想ネットワーク内での転送またはルーティングを容易にするためのパケットのカプセル化とカプセル化解除、特定のゲートウェイ（例えば、局所ピアリングゲートウェイ）の実行、セキュリティリスト、ネットワークセキュリティグループの実装、ネットワークアドレス変換（NAT）機能（例えば、ホストごとにパブリックIPをプライベートIPに変換する）、スロットル機能、およびその他の機能な

40

50

ど、オーバーレイネットワークに関連付けられる1つまたは複数の仮想機能を実行し得る。

【0095】

特定の実施形態では、NVD内のパケット処理データパスは、各々が一連のパケット変換段階から構成される複数のパケットパイプラインを含み得る。特定の実装では、パケットを受信すると、パケットは解析され、単一のパイプラインに分類される。その後、パケットはドロップされるか、NVDのインターフェースを介して送信されるまで、段階的に線形に処理される。これらの段階は、基本的な機能パケット処理構築ブロック（例えば、ヘッダーの検証、スロットルの適用、新しい層2ヘッダーの挿入、L4ファイアウォールの適用、VCNカプセル化／カプセル化解除など）を提供するため、既存の段階を組み合わせ

10

【0096】

NVDは、VCNの制御プレーンとデータプレーンに対応する制御プレーン機能とデータプレーン機能の両方を実行できる。VCN制御プレーンの例は、図12、13、14、および15（参照1216、1316、1416、および1516を参照）にも示されており、以下で説明する。VCNデータプレーンの例は、図12、13、14、および15（参照1218、1318、1418、および1518を参照）に示されており、以下で説明する。制御プレーン機能には、データの転送方法を制御するネットワークの構成（例えば、ルートとルートテーブルの設定、VNICの構成）に使用される機能が含まれる。特定の実施形態では、オーバーレイからサブストレートへのすべてのマッピングを一元的に計算し、それらをNVDおよびDRG、SGW、IGWなどのさまざまなゲートウェイなどの仮想ネットワークエッジ装置に公開するVCN制御プレーンが提供される。ファイアウォールルールも同じメカニズムを使用して公開できる。特定の実施形態では、NVDはそのNVDに関連するマッピングのみを取得する。データプレーン機能には、制御プレーンを使用して設定される構成に基づいてパケットを実際にルーティング／転送する機能が含まれる。VCNデータプレーンは、顧客のネットワークパケットが基盤ネットワークを通過する前にカプセル化することによって実装される。カプセル化／カプセル化解除機能はNVDに実装されている。特定の実施形態では、NVDは、ホストマシンに出入りするすべてのネットワークパケットを傍受し、ネットワーク仮想化機能を実行するように構成される。

20

30

【0097】

前述のように、NVDはVNICやVCN VRなどのさまざまな仮想化機能を実行する。NVDは、VNICに接続される1つまたは複数のホストマシンによってホストされるコンピューティングインスタンスに関連付けられているVNICを実行し得る。例えば、図2に示すように、NVD210は、NVD210に接続されるホストマシン202によってホストされるコンピューティングインスタンス268に関連付けられているVNIC276の機能を実行する。別の例として、NVD212は、ホストマシン206によってホストされるベアメタルコンピューティングインスタンス272に関連付けられているVNIC280を実行し、ホストマシン208によってホストされるコンピューティング

40

【0098】

NVDは、コンピューティングインスタンスのVCNに対応するVCN仮想ルータも実行する。例えば、図2に示す実施形態では、NVD210は、コンピューティングインスタンス268が属するVCNに対応するVCN VR277を実行する。NVD212は、ホストマシン206および208によってホストされるコンピューティングインスタンスが属する1つまたは複数のVCNに対応する1つまたは複数のVCN VR283を実

50

行する。特定の実施形態では、そのVCNに対応するVCN VRは、そのVCNに属する少なくとも1つのコンピューティングインスタンスをホストするホストマシンに接続されるすべてのNVDによって実行される。ホストマシンが異なるVCNに属するコンピューティングインスタンスをホストしている場合、そのホストマシンに接続されるNVDは、それらの異なるVCNに対応するVCN VRを実行し得る。

【0099】

VNICおよびVCN VRに加えて、NVDはさまざまなソフトウェア（例えば、デーモン）を実行し、NVDによって実行されるさまざまなネットワーク仮想化機能を容易にする1つまたは複数のハードウェア構成要素を含み得る。簡潔にするために、これらのさまざまな構成要素は、図2に示すように「パケット処理構成要素」としてグループ化されている。例えば、NVD 210はパケット処理構成要素286を含み、NVD 212はパケット処理構成要素288を含む。例えば、NVDのパケット処理構成要素には、NVDのポートおよびハードウェアインターフェースと対話して、NVDによって受信され、NVDを使用して通信されるすべてのパケットを監視し、ネットワーク情報を記憶するように構成されるパケットプロセッサが含まれ得る。ネットワーク情報には、例えば、NVDによって処理されるさまざまなネットワークフローを識別するネットワークフロー情報と、フローごとの情報（例えば、フローごとの統計）が含まれ得る。特定の実施形態では、ネットワークフロー情報はVNICごとに記憶され得る。パケットプロセッサは、パケットごとの操作を実行できるほか、ステートフルNATおよびL4ファイアウォール（FW）を実装することもできる。別の例として、パケット処理構成要素には、NVDによって記憶される情報を1つまたは複数の異なるレプリケーションターゲットストアに複製するように構成されるレプリケーションエージェントが含まれ得る。さらに別の例として、パケット処理構成要素には、NVDのログ機能を実行するように構成されるログエージェントが含まれ得る。パケット処理構成要素には、NVDの性能と健全性を監視するソフトウェアや、NVDに接続される他の構成要素の状態と健全性を監視するソフトウェアも含まれ得る。

【0100】

図1は、VCN、VCN内のサブネット、サブネット上に展開されるコンピューティングインスタンス、コンピューティングインスタンスに関連付けられているVNIC、VCNのVR、およびVCN用に構成されるゲートウェイのセットを含む、仮想ネットワークまたはオーバーレイネットワークの例の構成要素を示している。図1に示されているオーバーレイ構成要素は、図2に示されている1つまたは複数の物理構成要素によって実行またはホストされ得る。例えば、VCN内のコンピューティングインスタンスは、図2に示す1つまたは複数のホストマシンによって実行またはホストされ得る。ホストマシンによってホストされるコンピューティングインスタンスの場合、そのコンピューティングインスタンスに関連付けられているVNICは通常、そのホストマシンに接続されるNVDによって実行される（つまり、VNIC機能は、そのホストマシンに接続されるNVDによって提供される）。VCNのVCN VR機能は、そのVCNの一部であるコンピューティングインスタンスをホストまたは実行するホストマシンに接続されているすべてのNVDによって実行される。VCNに関連付けられるゲートウェイは、1つまたは複数の異なるタイプのNVDによって実行され得る。例えば、特定のゲートウェイはスマートNICによって実行され得る一方、他のゲートウェイは1つまたは複数のホストマシンまたはNVDの他の実装によって実行され得る。

【0101】

前述のように、顧客VCN内のコンピューティングインスタンスは、さまざまな異なるエンドポイントと通信できる。エンドポイントは、ソースコンピューティングインスタンスと同じサブネット内にある場合もあれば、ソースコンピューティングインスタンスと同じVCN内にあるが別のサブネット内にある場合もあり、あるいはソースコンピューティングインスタンスのVCN外にあるエンドポイントと通信できる。これらの通信は、コンピューティングインスタンスに関連付けられているVNIC、VCN VR、およびVC

Nに関連付けられるゲートウェイを使用して容易になる。

【0102】

V C N内の同じサブネット上の2つのコンピューティングインスタンス間の通信では、ソースコンピューティングインスタンスと宛先コンピューティングインスタンスに関連付けられているV N I Cを使用して通信が容易になる。ソースと宛先のコンピューティングインスタンスは、同じホストマシンでホストすることも、異なるホストマシンでホストすることもできる。ソースコンピューティングインスタンスから発信されるパケットは、ソースコンピューティングインスタンスをホストしているホストマシンから、そのホストマシンに接続されるN V Dに転送され得る。N V Dでは、パケットはパケット処理パイプラインを使用して処理される。これには、ソースコンピューティングインスタンスに関連付けられているV N I Cの実行が含まれ得る。パケットの宛先エンドポイントは同じサブネット内にあるため、ソースコンピューティングインスタンスに関連付けられているV N I Cを実行すると、パケットは宛先コンピューティングインスタンスに関連付けられているV N I Cを実行するN V Dに転送され、N V Dはパケットを処理して宛先コンピューティングインスタンスに転送する。ソースおよび宛先コンピューティングインスタンスに関連付けられているV N I Cは、同じN V D上で実行される場合（例えば、ソースおよび宛先コンピューティングインスタンスの両方が同じホストマシンによってホストされている場合）と、異なるN V D上で実行される場合（例えば、ソースおよび宛先コンピューティングインスタンスが異なるN V Dに接続される異なるホストマシンによってホストされている場合）がある。V N I Cは、N V Dによって記憶されるルーティング／転送テーブルを使用して、パケットの次のホップを決定し得る。

10

20

【0103】

サブネット内のコンピューティングインスタンスから同じV C N内の異なるサブネット内のエンドポイントにパケットを通信する場合、ソースコンピューティングインスタンスから発信されるパケットは、ソースコンピューティングインスタンスをホストしているホストマシンから、そのホストマシンに接続されるN V Dに通信される。N V Dでは、パケットはパケット処理パイプラインを使用して処理される。これには、1つまたは複数のV N I Cの実行と、V C Nに関連付けられているV Rが含まれ得る。例えば、パケット処理パイプラインの一部として、N V Dはソースコンピューティングインスタンスに関連付けられているV N I Cに対応する機能を実行または呼び出す（V N I Cを実行するとも呼ばれる）。V N I Cによって実行される機能には、パケット上のV L A Nタグの確認が含まれ得る。パケットの宛先がサブネット外にあるため、次にV C N V R機能が呼び出され、N V Dによって実行される。次に、V C N V Rは、宛先コンピューティングインスタンスに関連付けられているV N I Cを実行するN V Dにパケットをルーティングする。次に、宛先コンピューティングインスタンスに関連付けられているV N I Cがパケットを処理し、パケットを宛先コンピューティングインスタンスに転送する。ソースおよび宛先コンピューティングインスタンスに関連付けられているV N I Cは、同じN V D上で実行される場合（例えば、ソースおよび宛先コンピューティングインスタンスの両方が同じホストマシンによってホストされている場合）と、異なるN V D上で実行される場合（例えば、ソースおよび宛先コンピューティングインスタンスが異なるN V Dに接続される異なるホストマシンによってホストされている場合）がある。

30

40

【0104】

パケットの宛先がソースコンピューティングインスタンスのV C Nの外部にある場合、ソースコンピューティングインスタンスから発信されるパケットは、ソースコンピューティングインスタンスをホストしているホストマシンから、そのホストマシンに接続されるN V Dに通信される。N V Dは、ソースコンピューティングインスタンスに関連付けられているV N I Cを実行する。パケットの宛先エンドポイントはV C N外にあるため、パケットはそのV C NのV C N V Rによって処理される。N V DはV C N V R機能を呼び出し、その結果、パケットがV C Nに関連付けられる適切なゲートウェイを実行するN V Dに転送され得る。例えば、宛先が顧客のオンプレミスネットワーク内のエンドポイント

50

である場合、パケットはV C N V Rによって、V C N用に構成されるD R Gゲートウェイを実行するN V Dに転送され得る。V C N V Rは、ソースコンピューティングインスタンスに関連付けられているV N I Cを実行するN V Dと同じN V D上で実行されることも、別のN V Dによって実行されることもある。ゲートウェイは、スマートN I C、ホストマシン、またはその他のN V D実装であり得るN V Dによって実行され得る。その後、パケットはゲートウェイによって処理され、次のホップに転送され、パケットが目的の宛先エンドポイントに通信しやすくする。例えば、図2に示す実施形態では、コンピューティングインスタンス268から発信されるパケットは、リンク220を介して(N I C 232を使用して)ホストマシン202からN V D 210に通信され得る。N V D 210では、ソースコンピューティングインスタンス268に関連付けられているV N I Cであるため、V N I C 276が呼び出される。V N I C 276は、パケット内のカプセル化される情報を調べ、パケットを転送するための次のホップを決定して、パケットを目的の宛先エンドポイントに通信しやすくし、決定した次のホップにパケットを転送するように構成されている。

10

20

30

40

50

【0105】

V C Nに展開されるコンピューティングインスタンスは、さまざまなエンドポイントと通信できる。これらのエンドポイントには、C S P I 200によってホストされるエンドポイントとC S P I 200外部のエンドポイントが含まれ得る。C S P I 200によってホストされるエンドポイントには、同じV C Nまたは他のV C N内のインスタンスが含まれ得る。これらのV C Nは、顧客のV C Nである場合もあれば、顧客に属さないV C Nである場合もある。C S P I 200によってホストされるエンドポイント間の通信は、物理ネットワーク218を介して実行され得る。コンピューティングインスタンスは、C S P I 200によってホストされていないエンドポイント、またはC S P I 200の外部にあるエンドポイントと通信する場合もある。これらのエンドポイントの例には、顧客のオンプレミスネットワークまたはデータセンター内のエンドポイント、またはインターネットなどのパブリックネットワーク経由でアクセス可能なパブリックエンドポイントが含まれる。C S P I 200外部のエンドポイントとの通信は、さまざまな通信プロトコルを使用して、パブリックネットワーク(例えば、インターネット)(図2には示されていない)またはプライベートネットワーク(図2には示されていない)を介して実行できる。

【0106】

図2に示すC S P I 200のアーキテクチャは単なる一例であり、限定するものではない。代替実施形態では、変形、代替、および修正が可能である。例えば、いくつかの実装では、C S P I 200は、図2に示されているものよりも多いか少ないシステムまたは構成要素を備えていてもよく、2つ以上のシステムを組み合わせてもよく、またはシステムの構成または配置が異なってもよい。図2に示されているシステム、サブシステム、およびその他の構成要素は、ハードウェア、またはそれらの組み合わせを使用して、それぞれのシステムの1つまたは複数の処理装置(例えば、プロセッサ、コア)によって実行されるソフトウェア(例えば、コード、命令、プログラム)で実装され得る。ソフトウェアは、非一時的な記憶媒体(例えば、メモリ装置)に記憶され得る。

【0107】

図4は、特定の実施形態によるマルチテナントをサポートするためのI/O仮想化を提供するためのホストマシンとN V Dとの間の接続を示す。図4に示すように、ホストマシン402は仮想化環境を提供するハイパーバイザ404を実行する。ホストマシン402は、顧客/テナント#1に属するV M 1 406と顧客/テナント#2に属するV M 2 408の2つの仮想マシンインスタンスを実行する。ホストマシン402は、リンク414を介してN V D 412に接続される物理N I C 410を含む。各コンピューティングインスタンスは、N V D 412によって実行されるV N I Cに接続される。図4の実施形態では、V M 1 406はV N I C - V M 1 420に接続され、V M 2 408はV N I C - V M 2 422に接続されている。

【0108】

図 4 に示すように、NIC 410 は、論理 NIC A 416 と論理 NIC B 418 の 2 つの論理 NIC を含む。各仮想マシンは、独自の論理 NIC に接続され、それを使用して動作するように構成される。例えば、VM 1 406 は論理 NIC A 416 に接続され、VM 2 408 は論理 NIC B 418 に接続される。ホストマシン 402 は複数のテナントによって共有される 1 つの物理 NIC 410 のみを含むが、論理 NIC があるため、各テナントの仮想マシンは独自のホストマシンと NIC があるものと認識する。

【0109】

特定の実施形態では、各論理 NIC に独自の VLAN ID が割り当てられる。したがって、テナント #1 の論理 NIC A 416 には特定の VLAN ID が割り当てられ、テナント #2 の論理 NIC B 418 には別の VLAN ID が割り当てられる。VM 1 406 からパケットが通信されると、テナント #1 に割り当てられるタグがハイパーバイザによってパケットに接続され、その後、パケットはリンク 414 を介してホストマシン 402 から NVD 412 に通信される。同様に、VM 2 408 からパケットが通信されると、テナント #2 に割り当てられるタグがハイパーバイザによってパケットに接続され、その後、パケットはリンク 414 を介してホストマシン 402 から NVD 412 に通信される。したがって、ホストマシン 402 から NVD 412 に通信されるパケット 424 には、特定のテナントおよび関連する VM を識別する関連タグ 426 が含まれる。NVD では、ホストマシン 402 から受信したパケット 424 について、パケットに関連付けられているタグ 426 を使用して、パケットが VNIC-VM 1 420 によって処理されるか、VNIC-VM 2 422 によって処理されるかが決定される。その後、パケットは対応する VNIC によって処理される。図 4 に示す構成により、各テナントのコンピューティングインスタンスは、独自のホストマシンと NIC を所有していると認識できるようになる。図 4 に示すセットアップでは、マルチテナントをサポートするための I/O 仮想化が提供される。

【0110】

図 5 は、特定の実施形態による物理ネットワーク 500 の簡略化されるブロック図を示す。図 5 に示す実施形態は、Clos ネットワークとして構成されている。Clos ネットワークは、高い二分帯域幅と最大限のリソース使用率を維持しながら接続の冗長性を提供するように設計される特定のタイプのネットワークトポロジである。Clos ネットワークは、非ブロッキング、マルチ段階、またはマルチ層のスイッチングネットワークの一種であり、段階または層の数は 2、3、4、5 などになり得る。図 5 に示す実施形態は、層 1、層 2、および層 3 を含む 3 層ネットワークである。TOR スイッチ 504 は、Clos ネットワーク内の層 0 スイッチを表す。1 つまたは複数の NVD が TOR スイッチに接続される。層 0 スイッチは、物理ネットワークのエッジ装置とも呼ばれる。層 0 スイッチは、リーフスイッチとも呼ばれる層 1 スイッチに接続される。図 5 に示す実施形態では、セットの「n」個の層 0 TOR スイッチがセットの「n」個の層 1 スイッチに接続され、一緒にポッドを形成する。ポッド内の各層 0 スイッチはポッド内のすべての層 1 スイッチに相互接続されているが、ポッド間のスイッチの接続はない。特定の実装では、2 つのポッドがブロックと呼ばれる。各ブロックは、セットの「n」個の層 2 スイッチ（スパインスイッチと呼ばれることもある）によってサービスが提供されるか、またはこれらのスイッチに接続される。物理ネットワークトポロジには複数のブロックが存在し得る。層 2 スイッチは、次に「n」個の層 3 スイッチ（スーパースパインスイッチと呼ばれることもある）に接続される。物理ネットワーク 500 を介したパケットの通信は、通常、1 つまたは複数の層 3 通信プロトコルを使用して実行される。通常、TOR 層を除く物理ネットワークのすべての層は n 方向の冗長性を備えているため、高可用性が実現する。物理ネットワークのスケールリングを可能にするために、ポッドとブロックにポリシーを指定して、物理ネットワーク内のスイッチ間の可視性を制御することができる。

【0111】

Clos ネットワークの特徴は、1 つの層 0 スイッチから別の層 0 スイッチ（または層 0 スイッチに接続される NVD から層 0 スイッチに接続される別の NVD）に到達するま

での最大ホップ数が固定されていることである。例えば、3層のClosネットワークでは、ソースNVDとターゲットNVDがClosネットワークのリーフ層に接続されている場合、パケットが1つのNVDから別のNVDに到達するには最大7ホップが必要である。同様に、4層のClosネットワークでは、ソースNVDとターゲットNVDがClosネットワークのリーフ層に接続されている場合、パケットが1つのNVDから別のNVDに到達するには最大9ホップが必要である。したがって、Closネットワークアーキテクチャは、ネットワーク全体で一貫したレイテンシを維持する。これは、データセンター内およびデータセンター間の通信にとって重要である。Closトポロジは水平方向に拡張でき、コスト効率に優れている。ネットワークの帯域幅／スループット容量は、さまざまな層にスイッチを追加し（例えば、リーフスイッチとスパインスイッチを追加）、隣接する層のスイッチ間のリンク数を増やすことで簡単に増やすことができる。

10

【0112】

特定の実施形態では、CSP I内の各リソースには、クラウド識別子（CID）と呼ばれる固有の識別子が割り当てられる。この識別子はリソースの情報の一部として含まれており、例えば、コンソールやAPIなどを通じてリソースを管理するために使用できる。CIDの構文の例は次のとおりである。

ocid1.〈RESOURCE TYPE〉.〈REALM〉.[REGION][.FUTURE USE].〈UNIQUE ID〉

ocid1: CIDのバージョンを示すリテラル文字列。

resource type: リソースのタイプ（例えば、インスタンス、ボリューム、VCN、サブネット、ユーザ、グループなど）。

realm: リソースが存在するレルム。値の例としては、商用レルムの場合は「c1」、政府クラウドレルムの場合は「c2」、連邦政府クラウドレルムの場合は「c3」などがある。各レルムには独自のドメイン名があり得る。

region: リソースが存在するリージョン。リージョンがリソースに適用されない場合、この部分は空白になることがある。

future use: 将来の使用のために予約されている。

unique ID: IDの固有の部分。形式は、リソースまたはサービスのタイプによって異なり得る。

20

【0113】

図6は、特定の実施形態による、CLOSネットワーク配置を組み込んだクラウドインフラストラクチャ600のブロック図を示す。クラウドインフラストラクチャ600には、複数のラック（例えば、ラック1 610、ラック2、620）が含まれる。各ラックには複数のホストマシン（以下、ホストとも呼ばれる）が含まれる。ラック1 610には、つまり、ホスト1-A 612とホスト1-B 614の2つのホストマシンが含まれていることが示され、ラック2 620には、つまり、ホスト2-A 622とホスト2-B 624の2つのホストマシンが含まれていることが示されている。図6の説明（すなわち、各ラックに2台のホストマシンが含まれる）は、例示を目的としたものであり、限定するものではないことが理解される。例えば、クラウドインフラストラクチャには2台以上のラックが含まれ得、各ラックには2台以上のホストマシンが含まれ得る。さらに、各ラックには同じ数のホストが存在するという制限がないことに留意されたい。むしろ、ラックは、別のラックに含まれるホストマシンの数よりも多い場合も少ない場合もあり得る。

30

40

【0114】

各ホストマシンには、複数のグラフィック処理装置（GPU）が含まれる。例えば、ホストマシン1-A 612には、N個のGPU（例えば、GPU1、613）が含まれる。さらに、図6の説明では、各ホストマシンに同じ数のGPU、つまりN個のGPUが含まれているが、これは例示を目的としたものであり、限定するものではなく、つまり、各ホストマシンには異なる数のGPUを含めることができることが理解される。各ラックには、ラック内のホストマシンでホストされているGPUと通信可能に結合されるトップオブラック（TOR）スイッチが含まれている。例えば、ラック1 610には、ホストマシ

50

ンホスト１－Ａ、６１２およびホスト１－Ｂ、６１４に通信可能に結合されるＴＯＲスイッチ（つまり、ＴＯＲ１）６１６が含まれるが、ラック２ ６２０には、ホストマシンホスト２－Ａ、６２２およびホスト２－Ｋ、６２４に通信可能に結合されるＴＯＲスイッチ（つまり、ＴＯＲ２）６２６が含まれる。図６に示されているＴＯＲスイッチ（つまり、ＴＯＲ１ ６１６およびＴＯＲＭ６２６）には各々、ラックに含まれる各ホストマシンでホストされているＮ個のＧＰＵにＴＯＲスイッチを通信可能に結合するために使用されるＮ個のポートが含まれていることが理解される。図６に示されているＴＯＲスイッチとＧＰＵの結合は、例示を目的としており、限定するものではない。例えば、いくつかの実施形態では、ＴＯＲスイッチは複数のポートを有し得、各ポートは各ホストマシン上のＧＰＵに対応し、つまり、ホストマシン上のＧＰＵは通信リンクを介してＴＯＲの固有のポートに接続され得る。

10

【０１１５】

各ラックのＴＯＲスイッチは、複数のスパインスイッチ、例えばスパインスイッチ１、６３０およびスパインスイッチＰ６４０に通信可能に結合されている。例えば、図６に示すように、ＴＯＲ１ ６１６は、２つのリンクを介してスパインスイッチ１ ６３０に接続され、さらに別の２つのリンクを介してスパインスイッチＰ６４０にそれぞれ接続される。特定のＴＯＲスイッチからスパインスイッチに送信される情報は、本明細書ではアップリンクを介して行われる通信と呼ばれ、一方、スパインスイッチからＴＯＲスイッチに送信される情報は、本明細書ではダウンリンクを介して行われる通信と呼ばれる。いくつかの実施形態によれば、ＴＯＲスイッチとスパインスイッチはＣＬＯＳネットワーク配置（例えば、多段スイッチングネットワーク）で接続され、各ＴＯＲスイッチはＣＬＯＳネットワーク内の「リーフ」ノードを形成する。

20

【０１１６】

いくつかの実施形態によれば、ホストマシンに含まれるＧＰＵはマシン学習に関連するタスクを実行する。このような設定では、単一のタスクが、複数のホストマシンおよび複数のラックにまたがる可能性のある多数のＧＰＵ（例えば、６４個のＧＰＵ）にわたって実行／分散され得る。これらすべてのＧＰＵは同じタスク（つまり、ワークロード）で動作しているため、すべてが時間同期される方法で相互に通信する必要がある。さらに、どの時点においても、ＧＰＵはコンピューティングモードまたは通信モードのいずれかの状態にあり、つまり、ＧＰＵはほぼ同時に相互に通信する。ワークロードの速度は、最も遅いＧＰＵの速度によって決定する。

30

【０１１７】

通常、ソースＧＰＵから宛先ＧＰＵにパケットをルーティングするために、等コストマルチパス（ＥＣＭＰ）ルーティングが利用される。ＥＣＭＰルーティングでは、送信者から受信者へのトラフィックのルーティングに使用できる等コストパスが複数ある場合、選択技術を使用して特定のパスを選択する。したがって、トラフィックを受信するネットワーク装置（例えば、ＴＯＲスイッチやスパインスイッチ）では、選択アルゴリズムを使用して、ネットワーク装置から後続の装置にトラフィックを転送するために使用される送信リンクが選択される。この送信リンクの選択は、送信者から受信者へのパス内の各ネットワーク装置で行われる。ハッシュベースの選択は、広く使用されているＥＣＭＰ選択技術であり、ハッシュは、例えば、パケットの４組（例えば、ソースポート、宛先ポート、ソースＩＰ、宛先ＩＰ）に基づき得る。

40

【０１１８】

ＥＣＭＰルーティングはフローを考慮したルーティング技術であり、各フロー（つまり、データパケットのストリーム）はフローの持続期間中、同じパスにハッシュされる。したがって、フロー内のパケットは、特定の送信ポート／リンクを使用してネットワーク装置から転送される。これは通常、フロー内のパケットが順番に到着することを保証するため、つまりパケットの順序変更が不要であることを確認するために行われる。しかし、ＥＣＭＰルーティングは帯域幅（またはスループット）を認識しない。つまり、ＴＯＲスイッチとスパインスイッチは、並列リンク上のフローの統計的なフロー認識型（スループッ

50

ト非認識型) E C M P 負荷分散を実行する。

【0119】

標準の E C M P ルーティング (つまり、フローのみを認識するルーティング) では、ネットワーク装置が 2 つの別々の着信リンクを介して受信したフローが同じ送信リンクにハッシュされ、それによりフロー衝突が発生し得るという問題がある。例えば、2 つのフローが 2 つの別々の 100 G 着信リンクを介して着信し、各フローが同じ 100 G 送信リンクにハッシュされる状況を考慮する。このような状況では、着信帯域幅が 200 G であるのに対し、発信帯域幅が 100 G であるため、輻輳 (つまり、フロー衝突) が発生し、パケットがドロップされる。図 6 に示すように、2 つのフローがある。ホストマシンホスト 1-A、612 の第 1 の G P U から T O R スイッチ 616 に向けられるフロー 1 641 と、ホストマシン 614 上の別の G P U から T O R スイッチ 616 に向けられるフロー 2 643 である。2 つのフローは別々のリンクで T O R スイッチに向けられることに留意されたい。図 6 に示されているすべてのリンクの容量 (帯域幅) は 100 G であると想定されている。T O R スイッチ 616 が E C M P ルーティングアルゴリズムを実行する場合、2 つのフローがハッシュされて、T O R の同じ送信リンク (例えば、T O R スイッチ 616 をスパインスイッチ 630 に接続するリンク 650) を使用し得る。この場合、2 つのフロー間に衝突が発生し (「X」マークで表される)、パケットがドロップされる。

10

【0120】

このような衝突シナリオは、プロトコルに関係なく、すべてのタイプのトラフィックで一般的に問題になる。例えば、T C P は、パケットがドロップされ、送信者がドロップされるパケットの確認応答を受信しなかった場合に、パケットが再送信されるという点でインテリジェントである。しかし、リモートダイレクトメモリアクセス (R D M A) タイプのトラフィックの場合、状況は悪化する。R D M A ネットワークでは、さまざまな理由から T C P を使用しない (例えば、T C P は複雑な論理を持ち、低レイテンシと高性能に適していない)。R D M A ネットワークでは、インフィニバンド上の R D M A や統合イーサネット上の R D M A (R o C E) などのプロトコルが使用される。R o C E には輻輳制御アルゴリズムがあり、送信者が輻輳の発生やパケットのドロップを識別すると、送信者はパケットの送信速度を低下させる。ドロップされたパケットの場合、ドロップされたパケットだけでなく、ドロップされたパケットの周囲のいくつかのパケットも再送信されるため、使用可能な帯域幅がさらに消費され、性能が低下する。

20

30

【0121】

フロー衝突の問題は、厳格な時間同期要件のため、例えば、G P U ワークロードなどのワークロード実行にとって重大な問題である。例えば、G P U は、すべての G P U が時間同期して相互に通信する必要があるマシン学習タスク (つまり、ワークロード) を実行し得る。さらに、顧客のワークロードはさまざまなタイプがあり得、各々に異なる要件があり得る。例えば、第 1 のワークロード (例えば、G P U ワークロード) は帯域幅に敏感である一方、第 2 のワークロード (例えば、高性能コンピューティング (H P C) ワークロード) は遅延またはレイテンシに敏感であり得る。さらに、顧客はアプリケーションに基づいてワークロードを実行するホストマシンを選択することを所望し得る。例えば、遅延 (またはレイテンシ) に敏感であるワークロードの場合、顧客は相互に物理的に近接した (例えば、同じラックに含まれている) ホストマシンを選択することを所望し得る。別の例として、顧客は、ネットワーク障害発生時により高い可用性を実現するために、ワークロードを反親和性グループに配置すること、つまりホストマシンを異なるラックに分散することを所望する場合がある。通常、ホストマシンは、(それぞれのワークロードを実行するために) 任意 (ランダム) の方法で顧客に割り当てられる。そのため、トラフィックの輻輳が増加する可能性が高まり、ワークロードのスループットが低下する。さらに、ホストマシンの局所性情報を顧客に提供する使用可能なメカニズムは現在存在しない。

40

【0122】

以下に、上記の問題を克服するための技術について説明する。具体的には、本明細書で説明する技術は、ホストマシンの局所性情報の階層を顧客に提供し、顧客が (提供される

50

局所性情報に基づいて) 特定のホストマシンを選択してワークロードを実行できるようにする。例えば、本開示のいくつかの実施形態は、顧客がワークロードを近くのホストマシンに「配置」することによって、アプリケーションからサービスへのレイテンシを削減することを可能にする。さらに、顧客は局所性情報を使用して、より高い反親和性が得られるようにワークロードを配置するため、リソースの共有結果を減らすことでより高い回復力を得ることができる。

【0123】

図7を参照すると、特定の実施形態によるラック700の例示的な構成が示されている。図7に示すように、ラック700には、つまり、ホスト1-A710とホスト1-B720の2つのホストマシンが含まれている。ラック700には2台のホストマシンが含まれているように示されているが、ラック700にはさらに多くのホストマシンが含まれていてもよいことは理解される。各ホストマシンは、さらに複数のGPUと複数のCPUを含み得る。

10

【0124】

ホストマシン(すなわち、ホスト1-A710およびホスト1-B720)は、トップオブラックスイッチ、すなわちTOR1 750を介してネットワークファブリックに通信可能に結合される。このネットワークファブリックは、本明細書ではラック700のフロントエンドネットワークとも呼ばれ、外部ネットワークに対応する場合もある。ホストマシン、すなわちホスト1-A710は、ネットワークインターフェースカード(NIC)730およびTOR1 750に結合されるNVD735(すなわち、ネットワーク仮想化装置)を介してフロントエンドネットワークに接続される。ホストマシン、すなわちホスト1-B720は、ネットワークインターフェースカード(NIC)740およびネットワーク仮想化装置745を介してフロントエンドネットワークに接続され、ネットワーク仮想化装置745はTOR1 750に結合される。ホストマシンホスト1-A710とホスト1-B720は、反対側のQoS対応バックエンドネットワークに接続されている。QoS対応のバックエンドネットワークは、図6に示すようにGPUクラスタネットワークに対応し得る。ホストマシン1-A710は、別のNIC765を介してTOR2スイッチ760に接続され、TOR2スイッチ760はホストマシンをバックエンドネットワークに通信可能に結合する。同様に、ホストマシン1-B720はNIC780を介してTOR2スイッチ760に接続され、TOR2スイッチ760はホストマシンをバックエンドネットワークに通信可能に結合する。

20

30

【0125】

いくつかの実施形態によれば、ホストマシンはネットワークの物理的なトポロジを認識しない、すなわち、特定のホストマシンはネットワーク内の他のホストマシンの物理的な場所/位置を認識しない。例えば、図6を参照すると、ホストマシン1-A612は、ホストマシン1-B614が実際には同じラック(つまり、ラック1 610)に含まれており、同じTORスイッチ(つまり、TOR1 616)の背後に配置されていることを認識していない。しかし、ネットワーク制御プレーンは、ホストマシンの全体的な物理トポロジを認識する。1つの実装では、ネットワーク制御プレーンは、トラフィックの局所性を実現し、不必要なトラフィックの輻輳を回避するために、このような局所性情報をホストマシンに公開する。そうすることで、ワークロードの性能に大きな影響を与える。

40

【0126】

いくつかの実施形態では、ネットワーク制御プレーンは、インスタンスメタデータサービス(IMDS)を利用して、メタデータ情報(例えば、局所性情報)をホストマシンに公開(および記憶)する。このようなメタデータ情報は、ホストマシンに関連付けられているそれぞれのNVDを介してホストマシンに公開され得る。例えば、図7を参照すると、局所性情報はNVD735を介してホスト1-A710に(IMDS経由で)公開され、一方、局所性情報はNVD745を介してホスト1-B720に(IMDS経由で)公開される。次に図8および図10を参照して説明されるように、局所性情報には、ホストマシンの局所性情報の階層を示すメタデータが含まれ得ることが理解される。各ホストマ

50

シンはIMDSにクエリを実行して、ホストマシンに関連付けられるメタデータ情報を取得できることが理解される。

【0127】

図8は、特定の実施形態による、ホストマシンの局所性情報850の階層を示す概略図である。説明のために、図8の各ラックには2台のホストマシンが含まれているように示されている。例えば、ラック1 801には2台のホストマシンH1とH2（TORスイッチ、つまりTOR1に関連付けられている）が含まれ、ラック2 803には2台のホストマシンH3とH4（TORスイッチ、つまりTOR2に関連付けられている）が含まれ、ラック3 811には2台のホストマシンH5とH6（TORスイッチ、つまりTOR3に関連付けられている）が含まれ、ラック4 813には2台のホストマシンH7とH8（TORスイッチ、つまりTOR4に関連付けられている）が含まれる。同様に、ラック5 821～ラック8 833にはそれぞれ2台のホストマシンが含まれるように示されている。各ラックには2台のホストマシンが含まれることに制限されるわけではないことが理解される。むしろ、各ラックにはより多数のホストマシンが含まれ得、さらに、ラックに含まれるホストマシンの数は、別のラックに含まれるホストマシンの数と比較して異なり得る。

10

【0128】

いくつかの実施形態によれば、ラックの集合がブロックを構成し得る。例えば、ラック1（801）とラック2（803）はブロック1 805を構成し、一方、ラック3（811）とラック4（813）はブロック2（815）を構成する。同様に、ラック5（821）とラック6（823）はブロック3（825）を構成し、ラック7（831）とラック8（833）はブロック4（835）を構成する。さらに、ブロックの集合が建物を構成することができる。例えば、ブロック1 805とブロック2 815は建物1 820を構成し得、一方、ブロック3 825とブロック4 835は建物2 840を構成し得る。建物の集合が可用性ドメインを構成し得る。例えば、図8に示すように、建物1 820および建物2 840は、可用性ドメイン1 850に含まれ得る。

20

【0129】

同じラックに含まれるホストマシン、例えばラック1 801に含まれるH1とH2は、ラックに関連付けられているトップオブラック（つまり、TOR）スイッチを共有することが理解される。さらに、ブロックは、互いに離れた位置に配置されているラック、例えば、ラック1 801（ブロック1 805に含まれる）とラック4（ブロック2に含まれる）とは対照的に、互いに近く（つまり、地理的に近い）に配置されるラックに対応する（例えば、ブロック1 805に含まれるラック1 801とラック2 803は、互いに近接して配置される）。さらに、建物の集合は可用性ドメイン（例えば、データセンター）に対応し得、つまり、可用性ドメイン1 850に含まれる建物1 820と建物2 840である。したがって、図9を参照して以下に説明するように、ホストマシンの局所性情報には、ラックの位置、ブロックの位置、建物の位置、可用性ドメインの位置など、複数のレベルの階層情報がある。

30

【0130】

図9は、特定の実施形態による、ホストマシンの局所性情報のレベル（すなわち、階層）を示す例示的なツリー図を示す。階層的局所性情報は、レイテンシの降順で表示される。図9に示すように、階層的局所性情報のレベル1には、複数の可用性ドメイン（AD）、例えばAD1、AD2、...、ADKが含まれる。レベル2では、各可用性ドメインには、可用性ドメイン2に含まれる複数の建物、例えば建物1、建物2、...、建物Lが含まれ得る。レベル3では、各建物には、複数のブロック、例えば、建物2に含まれるブロック1、ブロック2、...、ブロックMが含まれ得る。最後に、レベル4（つまり、レイテンシ量が最小のレベル）では、各ブロックに1つまたは複数のラックが含まれ得、各ラックには1つまたは複数のホストマシンが含まれる。例えば、図9に示すように、ブロック2には複数のラック、すなわちラック1、ラック2、...、ラックNが含まれる。ラック2にある特定のホストの階層的局所性情報は、図9に示す局所性ツリー内の、ルートノー

40

50

ドとホストの場所の間のノードを識別する情報を含む（例えば、ホスト A の場合、階層的局所性情報は、Availability__domain__2→Building__2→Block__2→Rack__2を含む）。

【0131】

いくつかの実施形態によれば、各ホストマシンはインスタンスメタデータサービス（IMDS）と通信して、ホストマシンに関連する階層的局所性情報を取得できる。例えば、各ホストマシンはNVDに関連付けられており、その中でアプリケーションを実行してプライベートIPアドレス（例えば、IMDSのアドレス）と通信し、階層的局所性情報を取得できる。1つの実装では、ホストマシンはIMDSから階層的局所性情報の符号化される配列を取得できる。さらに、図10を参照して後述するように、符号化される階層的局所性情報は、顧客ごとに固有になり得る。

10

【0132】

図10は、特定の実施形態による、ホストマシンの階層的局所性情報の例示的な符号化される配列を示す。階層的局所性情報の符号化される配列1000には、例えば、ハッシュブロック1 1010、ハッシュブロック2 1020、ハッシュブロック3 1030、およびハッシュブロック4 1040などのハッシュされる情報の複数のブロックが含まれる。連続する2つのハッシュブロックは、ハイフン、セミコロン、カンマなどに対応し得る識別子1015によって区切られ得ることが理解される。

【0133】

一実施形態では、ハッシュブロック1 1010には、ラック識別子（すなわち、ホストマシンを含むラックのID）のハッシュとテナントID（すなわち、顧客ID）が含まれ得る。そうすることで、符号化される配列は顧客ごとに固有になる。具体的には、ハッシュブロックに顧客情報（例えば、テナントID）を含めることで、偶然同じラックにホストマシンがある2人の顧客が、同じラックにあることに気付かないようにすることができる。しかし、これにより、特定の顧客は、同じラック上で実行されている顧客のインスタンスを知ることができる。言い換えれば、符号化される配列は、他の顧客の配置を明らかにすることなく、特定の顧客の階層的局所性情報を提供する。同様に、ハッシュブロック1020にはブロック識別子のハッシュとテナントIDが含まれ、一方、ハッシュブロック1030には建物識別子のハッシュとテナントIDが含まれ、ハッシュブロック1040には可用性ドメイン識別子のハッシュとテナントIDが含まれ得る。

20

30

【0134】

いくつかの実施形態では、ホストマシンの階層的局所性情報がクラウドインフラストラクチャの顧客に提供され得る。これにより、顧客はホストマシンの階層的局所性情報を活用して、ワークロードに関連付けられるアプリケーションのタイプに基づいてワークロードをプロビジョニングできる。例えば、低レイテンシを要求するワークロードの場合、ユーザはハッシュブロック1（1010）に含まれる情報、つまりラック情報を利用して、同じラックまたは隣接するラック上のホストマシンを選択できる。別の例として、GPUワークロードの場合、局所性情報を使用して、帯域幅効率の高い通信パターンを構成することができる（例えば、大量のデータを迅速に移動する）。高可用性を必要とするアプリケーション、またはラックのフロントエンドネットワーク側で実行されるアプリケーション（図7を参照）に関しては、ユーザは、ハッシュブロック（つまり、図10のハッシュブロック1010、1020、1030、および／または1040）の1つまたは複数に含まれる情報を活用して、例えば、異なるラック、ブロック、または建物に属するホストマシンを選択し、一定レベルの反親和性（つまり、高可用性）を達成することができる。ホストマシンの階層的局所性情報は、分散Hadoopワークロードおよび／または高頻度取引アプリケーション（アプリケーションをサポートするために1つまたは複数のホストマシンの選択を必要とする）などのアプリケーションによって、目的のタスクを実行するために1つまたは複数のホストマシンを選択する際に利用され得ることが理解されている。

40

【0135】

50

別の例として、ファイルシステムの場合を考えると、ネットワーク局所性情報を利用して、クライアントをファイルシステムサーバの近くに配置して低レイテンシを実現することができる。さらに別のアプリケーションは、3つの層、つまり、クライアント層、中間層（例えば、アプリケーションサーバ層）、およびデータベース層で構成されるデータベースアプリケーションである。このようなアプリケーションでは、顧客はデータベースワークロードの低レイテンシを所望する場合がある。顧客は、ネットワーク局所性情報を活用して、クライアント、中間層サーバ、データベースを適切に配置し、低レイテンシを実現できる。このシナリオでは、顧客はネットワーク局所性情報を利用することで、中間層とデータベース層で同時に高可用性を実現できることが理解される。

【0136】

図11Aは、特定の実施形態による、要求をプロビジョニングする際に実行されるステップを示す例示的なフローチャートを示している。図11に示されている処理は、それぞれのシステム、ハードウェア、またはそれらの組み合わせの1つまたは複数の処理装置（例えば、プロセッサ、コア）によって実行されるソフトウェア（例えば、コード、命令、プログラム）で実装され得る。ソフトウェアは、非一時的な記憶媒体（例えば、メモリ装置）に記憶され得る。図11に示され、以下で説明される方法1100は、例示を目的としており、限定するものではない。図11は、特定のシーケンスまたは順序で発生するさまざまな処理ステップを示しているが、これは限定するものではない。特定の代替実施形態では、ステップは異なる順序で実行され得、また、一部のステップは並行して実行され得る。

【0137】

プロセスはステップ1105で開始され、複数のホストマシン（すなわち、クラスタに含まれるホストマシン）の各ホストマシンについて、そのホストマシンの階層的局所性情報が記憶される。ホストマシンの階層的局所性情報には、複数の局所性の各局所性について、その局所性の位置情報を識別する情報が含まれる。いくつかの実施形態では、インスタンスメタデータサービスを利用して、ホストマシンに関連付けられるネットワーク仮想化装置（NVD）を介して、ホストマシンに階層的局所性情報を記憶することができる。図9に示すように、階層的局所性情報は、ホストマシンを含むラックの識別子（すなわち、ラックID）、ラックを含むブロックの識別子、ブロックを含む建物の識別子、および建物を含むドメインの識別子を示す情報に対応することができることに留意されたい。

【0138】

ステップ1110では、制御プレーンは、ワークロードの実行を要求する要求（例えば、顧客からの要求）を受信する。一実施形態では、ワークロードは、ホストマシンに関連付けられるGPUを使用して実行される1つまたは複数のプロセスに対応することに留意されたい。次に、プロセスはステップ1115に移動し、複数のホストマシンの各々について階層的局所性情報が取得される。階層的局所性情報（各ホストマシンの）は、対応するホストマシンのインスタンスメタデータサービスによって記憶され得ることに留意されたい。

【0139】

その後、プロセスはステップ1120に進み、要求に応じて複数のホストマシンの階層的局所性情報が提供される。いくつかの実施形態によれば、顧客は、複数のホストマシンの階層的局所性情報を取得すると、ワークロードを実行するための1つまたは複数のホストマシンを選択できる。1つまたは複数のホストマシンの選択は、ワークロードに関連付けられる1つまたは複数の制約に基づいて行われる。制約に関する詳細については、図11Bを参照して後で説明する。ステップ1125では、制御プレーンは1つまたは複数のホストマシンの選択を受信する。ステップ1130では、プロセスは選択される1つまたは複数のホストマシン上でワークロードの実行に進む。

【0140】

ホストマシンの階層的局所性情報を利用してワークロードを実行する上記のステップは、本開示の範囲を決して制限するものではないことは理解される。むしろ、ステップに対

10

20

30

40

50

する特定の変更は、本開示の範囲内に十分含まれる。例えば、一実施形態では、図 1 1 B を参照して次に説明するように、顧客がワークロードを実行するためのホストマシンの数を選択するのではなく、制御プレーンが顧客のホストマシンの数を選択し得る（つまり、顧客のワークロードに関連付けられる 1 つまたは複数の制約に基づいて）。

【0141】

図 1 1 B は、特定の実施形態に従って、顧客のワークロード要求をプロビジョニングする際に実行されるステップを示す別の例示的なフローチャートを示す。図 1 1 に示される処理は、それぞれのシステム、ハードウェア、またはそれらの組み合わせの 1 つまたは複数の処理装置（例えば、プロセッサ、コア）によって実行されるソフトウェア（例えば、コード、命令、プログラム）で実装され得る。ソフトウェアは、非一時的な記憶媒体（例えば、メモリ装置）に記憶され得る。図 1 1 に示され、以下で説明される方法 1 1 5 0 は、例示を目的としており、限定するものではない。図 1 1 は、特定のシーケンスまたは順序で発生するさまざまな処理ステップを示しているが、これは限定するものではない。特定の代替実施形態では、ステップは異なる順序で実行され得、また、一部のステップは並行して実行され得る。

10

【0142】

プロセスはステップ 1 1 5 5 で開始され、複数のホストマシン（すなわち、クラスタに含まれるホストマシン）の各ホストマシンについて、そのホストマシンの階層的局所性情報が記憶される。ホストマシンの階層的局所性情報には、複数の局所性の各局所性について、その局所性の位置情報を識別する情報が含まれる。いくつかの実施形態では、インスタンスメタデータサービスを利用して、ホストマシンに関連付けられるネットワーク仮想化装置（NVD）を介して、ホストマシンに階層的局所性情報を記憶することができる。図 9 に示すように、階層的局所性情報は、ホストマシンを含むラックの識別子（すなわち、ラック ID）、ラックを含むブロックの識別子、ブロックを含む建物の識別子、および建物を含むドメインの識別子を示す情報に対応することができることに留意されたい。

20

【0143】

ステップ 1 1 6 0 では、顧客からの要求が受信される。この要求では、顧客がワークロードを実行するために所望するホストマシンの数を要求する。要求には、ワークロードに関連付けられる 1 つまたは複数の制約に対応するメタデータが含まれ得ることが理解される。いくつかの実施形態では、1 つまたは複数の制約には、レイテンシ閾値に関連付けられる第 1 の制約が含まれ得、つまり、顧客は、レイテンシが所定の閾値よりも低くなるようにワークロードを実行することを所望する可能性がある。第 2 の制約は、反親和性制約に対応し得る。このような制約は、ホストマシンの可用性をある程度確保したいという顧客の要望に対応する。つまり、ワークロードを実行するために選択されるホストマシンの少なくとも一部は、異なるラックに配置する必要がある。このような制約は通常、ラックの障害の問題に対処するために顧客によって組み込まれる。

30

【0144】

ステップ 1 1 6 5 では、ステップ 1 1 6 0 で受信した要求から 1 つまたは複数の制約が抽出される。その後、プロセスはステップ 1 1 7 0 に移動し、制御プレーンは、顧客のワークロードを実行するために（複数のホストマシンから）1 つまたは複数のホストマシンを割り当てる。ワークロードを実行するために割り当てられるホストマシンは、ステップ 1 1 1 5 で識別される 1 つまたは複数の制約および／またはホストマシンの階層的局所性情報に従って決定され得ることが理解される。いくつかの実施形態によれば、ステップ 1 1 7 0 のプロセスは、ネットワーククラスタのトポロジ（例えば、ラックに含まれるホストマシンの数、システムで使用可能なラックの数）に従って、ワークロードを実行するためにホストマシンを割り当てる。各ラックに合計 N 台のホストマシンが含まれており、ユーザが要求するホストマシンの数が K 台であると考慮する。K が N より小さい場合（つまり、 $K < N$ ）、1 つの実装では、K 台のホストマシンすべてを 1 つのラックから割り当てることができる。つまり、K 台のホストマシンすべてを 1 つの TOR スイッチの背後に配置することができる。このような割り当てにより、レイテンシが最小限に抑えられること

40

50

に留意されたい。しかし、ユーザが特定のレベルの反親和性、つまり可用性を実現することを所望する場合もあり、その場合、ユーザに割り当てられるホストマシンの一部は、ネットワーククラスタの他のラックから選択され得るため、このような割り当てはユーザの承認を待って提供され得ることが理解される。別の例としては、 K が N より大きい（ $K > N$ ）場合に対応し得る。つまり、ユーザが要求するホストマシンの数が、ラックに含まれるホストマシンの数よりも多い場合である。この場合、1つの実装では、最小数のTORスイッチが使用されるように、すなわちラック内ルーティングが最小化されるように割り当てを実行することができる。その後、プロセスはステップ1175に移動し、割り当てられるホストマシンの階層的局所性情報が顧客に提供される。プロセスはステップ1180に移動し、顧客のワークロードは顧客に割り当てられる1つまたは複数のホストマシン上で実行される。

10

【0145】

クラウドインフラストラクチャの実施形態の例

前述のように、*infrastructure as a service*（*IaaS*）は、クラウドコンピューティングの特定のタイプである。*IaaS*は、パブリックネットワーク（例えば、インターネット）経由で仮想化されるコンピューティングリソースを提供するように構成できる。*IaaS*モデルでは、クラウドコンピューティングプロバイダがインフラストラクチャ構成要素（例えば、サーバ、記憶装置、ネットワークノード（例えば、ハードウェア）、展開ソフトウェア、プラットフォーム仮想化（例えば、ハイパーバイザ層）など）をホストできる。場合によっては、*IaaS*プロバイダが、これらのインフラストラクチャ構成要素に付随するさまざまなサービス（例えば、課金、監視、ログ記録、セキュリティ、負荷分散、クラスタリングなど）も提供することができる。したがって、これらのサービスはポリシー駆動型であり得るため、*IaaS*ユーザは、アプリケーションの可用性と性能を維持するために負荷分散を駆動するポリシーを実装できる可能性がある。

20

【0146】

場合によっては、*IaaS*顧客はインターネットなどの広域ネットワーク（WAN）を介してリソースおよびサービスにアクセスし、クラウドプロバイダのサービスを使用してアプリケーションスタックの残りの要素をインストールできる。例えば、ユーザは*IaaS*プラットフォームにログインして仮想マシン（VM）を作成し、各VMにオペレーティングシステム（OS）をインストールし、データベースなどのミドルウェアを展開し、ワークロードとバックアップ用の記憶バケットを作成し、そのVMにエンタープライズソフトウェアをインストールすることもできる。顧客は、次いでプロバイダのサービスを使用して、ネットワークトラフィックの分散、アプリケーションの問題のトラブルシューティング、性能の監視、災害復旧の管理など、さまざまな機能を実行できる。

30

【0147】

ほとんどの場合、クラウドコンピューティングモデルではクラウドプロバイダの参加が必要になる。クラウドプロバイダは、*IaaS*の提供（例えば、提供、レンタル、販売）を専門とするサードパーティサービスである場合もあるが、そうである必要はない。エンティティはプライベートクラウドを展開し、独自のインフラストラクチャサービスのプロバイダになることを選択することもできる。

40

【0148】

いくつかの例では、*IaaS*展開は、新しいアプリケーション、またはアプリケーションの新しいバージョンを準備されるアプリケーションサーバなどに配置するプロセスである。また、サーバの準備プロセス（例えば、ライブラリ、デーモンなどのインストール）も含まれ得る。これは多くの場合、ハイパーバイザ層（例えば、サーバ、記憶装置、ネットワークハードウェア、仮想化）の下でクラウドプロバイダによって管理される。したがって、顧客は、（OS）、ミドルウェア、および／またはアプリケーションの展開（例えば、オンデマンドで起動できるセルフサービス仮想マシン）の処理を担当し得る。

【0149】

いくつかの例では、*IaaS*プロビジョニングは、使用するためにコンピュータまたは

50

仮想ホストを取得し、必要なライブラリまたはサービスをそれらにインストールすることを指し得る。ほとんどの場合、展開にはプロビジョニングは含まれず、最初にプロビジョニングを実行する必要があると得る。

【0150】

場合によっては、IaaSプロビジョニングには2つの異なる課題がある。まず、何かを実行する前に、インフラストラクチャの初期セットをプロビジョニングするという最初の課題がある。第2に、すべてがプロビジョニングされた後、既存のインフラストラクチャを進化させる（例えば、新しいサービスの追加、サービスの変更、サービスの削除など）という課題がある。場合によっては、インフラストラクチャの構成を宣言的に定義できるようにすることで、これら2つの課題に対処できることがある。つまり、インフラストラクチャ（例えば、必要な構成要素やそれらの対話方法）は、1つまたは複数の構成ファイルによって定義できる。したがって、インフラストラクチャの全体的なトポロジ（例えば、どのリソースがどのリソースに依存しているか、また、各々のリソースがどのように連携するか）を宣言的に記述できる。場合によっては、トポロジが定義されると、構成ファイルに記述されているさまざまな構成要素を作成および／または管理するワークフローを生成できる。

10

【0151】

いくつかの例では、インフラストラクチャは相互接続される多くの要素を持ち得る。例えば、コアネットワークとしても知られる、1つまたは複数の仮想プライベートクラウド（VPC）（例えば、潜在的なオンデマンドで構成可能なおよび／または共有コンピューティングリソースのプール）が存在し得る。いくつかの例では、ネットワークのセキュリティの設定方法と1つまたは複数の仮想マシン（VM）を定義するためにプロビジョニングされる1つまたは複数のセキュリティグループルールが存在する場合もある。負荷分散装置、データベースなどの他のインフラストラクチャ要素もプロビジョニングされ得る。より多くのインフラストラクチャ要素が所望されたりおよび／または、追加されたりするにつれて、インフラストラクチャは段階的に進化し得る。

20

【0152】

場合によっては、継続的展開技術を使用して、さまざまな仮想コンピューティング環境にわたってインフラストラクチャコードの展開を可能にすることができる。さらに、説明した技術により、これらの環境内でのインフラストラクチャ管理を可能にすることができる。いくつかの例では、サービスチームは、1つまたは複数の、多くの場合は多数の異なる運用環境（例えば、さまざまな地理的な場所、場合によっては世界中）に展開する必要があるコードを作成できる。しかし、一部の例では、コードが展開されるインフラストラクチャを最初にセットアップする必要がある。場合によっては、プロビジョニングを手動で実行したり、プロビジョニングツールを使用してリソースをプロビジョニングしたり、および／またはインフラストラクチャがプロビジョニングされた展開ツールを使用してコードを展開したりすることもできる。

30

【0153】

図12は、少なくとも1つの実施形態によるIaaSアーキテクチャの例示的なパターンを示すブロック図1200である。サービスオペレータ1202は、仮想クラウドネットワーク（VCN）1206およびセキュアホストサブネット1208を含み得るセキュアホストテナント1204に通信可能に結合できる。いくつかの例では、サービスオペレータ1202は、1つまたは複数のクライアントコンピューティング装置を使用し得る。これらの装置は、ポータブルハンドヘルド装置（例えば、iPhone（登録商標）、携帯電話、iPad（登録商標）、コンピューティングタブレット、パーソナルデジタルアシスタント（PDA））またはウェアラブル装置（例えば、Google Glass（登録商標）ヘッドマウントディスプレイ）であり得、Microsoft Windows Mobile（登録商標）などのソフトウェア、および／またはiOS、Windows Phone、Android、BlackBerry 8、Palm OSなどのさまざまなモバイルオペレーティングシステムを実行し、インターネット、電子メール、ショートメッセージサービス（SMS）、BlackBerry（登録商標）、

40

50

またはその他の通信プロトコルが有効になっている。あるいは、クライアントコンピューティング装置は、例えば、Microsoft Windows（登録商標）、Apple Macintosh（登録商標）、および／またはLinux（登録商標）オペレーティングシステムのさまざまなバージョンを実行するパーソナルコンピュータおよび／またはラップトップコンピュータを含む、汎用パーソナルコンピュータにすることができる。クライアントコンピューティング装置は、さまざまな市販のUNIX（登録商標）またはUNIXライクなオペレーティングシステム（Google Chrome OSなどのさまざまなGNU/Linuxオペレーティングシステムを含むがこれに限定されない）のいずれかを実行するワークステーションコンピュータであり得る。代替的に、または追加的に、クライアントコンピューティング装置は、シンクライアントコンピュータ、インターネット対応ゲームシステム（例えば、Kinect（登録商標）ジェスチャ入力装置の有無にかかわらずMicrosoft Xboxゲームコンソール）、および／または個人用メッセージング装置など、VCN1206および／またはインターネットにアクセスできるネットワークを介して通信できるその他の電子装置であり得る。

10

【0154】

VCN1206には、SSH VCN1212に含まれるLPG1210を介してセキュアシェル（SSH）VCN1212に通信可能に結合できる局所ピアリングゲートウェイ（LPG）1210を含めることができる。SSH VCN1212にはSSHサブネット1214が含まれ得、SSH VCN1212は、制御プレーンVCN1216に含まれるLPG1210を介して制御プレーンVCN1216に通信可能に結合できる。また、SSH VCN1212は、LPG1210を介してデータプレーンVCN1218に通信可能に結合できる。制御プレーンVCN1216とデータプレーンVCN1218は、IaaSプロバイダが所有および／または動作できるサーステナント1219に含めることができる。

20

【0155】

制御プレーンVCN1216には、境界ネットワーク（例えば、企業イントラネットと外部ネットワーク間の企業ネットワークの一部）として機能する制御プレーン非武装地帯（DMZ）層1220が含まれることができる。DMZベースのサーバは責任範囲が制限されており、セキュリティ侵害の抑制に役立ち得る。さらに、DMZ層1220には、1つまたは複数の負荷分散装置（LB）サブネット1222、アプリサブネット1226を含むことができる制御プレーンアプリ層1224、データベース（DB）サブネット1230（例えば、フロントエンドDBサブネットおよび／またはバックエンドDBサブネット）を含むことができる制御プレーンデータ層1228を含めることができる。制御プレーンDMZ層1220に含まれるLBサブネット1222は、制御プレーンアプリ層1224に含まれるアプリサブネット1226および制御プレーンVCN1216に含まれ得るインターネットゲートウェイ1234と通信可能に結合することができ、アプリサブネット1226は、制御プレーンデータ層1228に含まれるDBサブネット1230およびサービスゲートウェイ1236およびネットワークアドレス変換（NAT）ゲートウェイ1238と通信可能に結合することができる。制御プレーンVCN1216には、サービスゲートウェイ1236とNATゲートウェイ1238を含めることができる。

30

40

【0156】

制御プレーンVCN1216には、アプリサブネット1226を含むことができるデータプレーンミラーアプリ層1240を含めることができる。データプレーンミラーアプリ層1240に含まれるアプリサブネット1226には、コンピューティングインスタンス1244を実行できる仮想ネットワークインターフェースコントローラ（VNIC）1242を含めることができる。コンピューティングインスタンス1244は、データプレーンミラーアプリ層1240のアプリサブネット1226を、データプレーンアプリ層1246に含めることができるアプリサブネット1226に通信可能に結合できる。

【0157】

データプレーンVCN1218には、データプレーンアプリ層1246、データプレーン

50

ンDMZ層1248、およびデータプレーンデータ層1250が含まれ得る。データプレーンDMZ層1248には、データプレーンアプリ層1246のアプリサブネット1226およびデータプレーンVCN1218のインターネットゲートウェイ1234に通信可能に結合できるLBサブネット1222を含めることができる。アプリサブネット1226は、データプレーンVCN1218のサービスゲートウェイ1236およびデータプレーンVCN1218のNATゲートウェイ1238に通信可能に結合できる。データプレーンデータ層1250には、データプレーンアプリ層1246のアプリサブネット1226と通信可能に結合できるDBサブネット1230も含まれ得る。

【0158】

制御プレーンVCN1216およびデータプレーンVCN1218のインターネットゲートウェイ1234は、パブリックインターネット1254に通信可能に結合できるメタデータ管理サービス1252に通信可能に結合できる。パブリックインターネット1254は、制御プレーンVCN1216およびデータプレーンVCN1218のNATゲートウェイ1238に通信可能に結合できる。制御プレーンVCN1216およびデータプレーンVCN1218のサービスゲートウェイ1236は、クラウドサービス1256と通信可能に結合できる。

【0159】

いくつかの例では、制御プレーンVCN1216またはデータプレーンVCN1218のサービスゲートウェイ1236は、パブリックインターネット1254を経由せずに、クラウドサービス1256へのアプリケーションプログラミングインターフェース（API）呼び出しを行うことができる。サービスゲートウェイ1236からクラウドサービス1256へのAPI呼び出しは一方方向にすることができる。つまり、サービスゲートウェイ1236はクラウドサービス1256へのAPI呼び出しを行うことができ、クラウドサービス1256は要求されるデータをサービスゲートウェイ1236に送信することができる。しかし、クラウドサービス1256は、サービスゲートウェイ1236へのAPI呼び出しを開始できない場合がある。

【0160】

いくつかの例では、セキュアホストテナント1204は、そうでなければ分離され得るサービステナント1219に直接接続することができる。セキュアホストサブネット1208は、LPG1210を介してSSHサブネット1214と通信できる。LPG1210により、分離されるシステム上で双方向通信が可能になる。セキュアホストサブネット1208をSSHサブネット1214に接続すると、セキュアホストサブネット1208はサービステナント1219内の他のエンティティにアクセスできるようになる。

【0161】

制御プレーンVCN1216は、サービステナント1219のユーザが所望するリソースを設定またはプロビジョニングすることを可能にする。制御プレーンVCN1216でプロビジョニングされる必要なリソースは、データプレーンVCN1218に展開されるか、または使用され得る。いくつかの例では、制御プレーンVCN1216はデータプレーンVCN1218から分離され得、制御プレーンVCN1216のデータプレーンミラーアプリ層1240は、データプレーンミラーアプリ層1240およびデータプレーンアプリ層1246に含まれ得るVNIC1242を介して、データプレーンVCN1218のデータプレーンアプリ層1246と通信できる。

【0162】

いくつかの例では、システムのユーザ、または顧客は、メタデータ管理サービス1252に要求を通信できるパブリックインターネット1254を介して、作成、読み取り、更新、または削除（CRUD）動作などの要求を行うことができる。メタデータ管理サービス1252は、インターネットゲートウェイ1234を介して制御プレーンVCN1216に要求を通信できる。要求は、制御プレーンDMZ層1220に含まれるLBサブネット1222によって受信され得る。LBサブネット1222は、要求が有効であると決定でき、この決定に応じて、制御プレーンアプリ層1224に含まれるアプリサブネット1

10

20

30

40

50

226に要求を送信できる。要求が検証され、パブリックインターネット1254への呼び出しが必要な場合、パブリックインターネット1254への呼び出しは、パブリックインターネット1254への呼び出しを行うことができるNATゲートウェイ1238に送信され得る。要求によって記憶が望まれる可能性のあるメモリは、DBサブネット1230に記憶できる。

【0163】

いくつかの例では、データプレーンミラーアプリ層1240は、制御プレーンVCN1216とデータプレーンVCN1218との間の直接通信を容易にすることができる。例えば、データプレーンVCN1218に含まれるリソースに、構成の変更、更新、またはその他の適切な修正を適用することが望まれ得る。制御プレーンVCN1216は、VNIC1242を介して、データプレーンVCN1218に含まれるリソースと直接通信し、それによって、構成の変更、更新、またはその他の適切な修正を実行できる。

【0164】

いくつかの実施形態では、制御プレーンVCN1216およびデータプレーンVCN1218は、サービステナント1219内に含まれることができる。この場合、システムのユーザまたは顧客は、制御プレーンVCN1216またはデータプレーンVCN1218のいずれかを所有または動作することはできない。代わりに、IaaSプロバイダは、制御プレーンVCN1216とデータプレーンVCN1218を所有または動作することができ、その両方がサービステナント1219に含まれ得る。この実施形態は、ユーザまたは顧客が他のユーザまたは他の顧客のリソースと対話することを妨げる可能性のあるネットワークの分離を可能にすることができる。また、この実施形態により、システムのユーザまたは顧客は、記憶のために、望ましいレベルのセキュリティを備えていない可能性のあるパブリックインターネット1254に依存することなく、データベースを非公開に記憶できるようになる。

【0165】

他の実施形態では、制御プレーンVCN1216に含まれるLBサブネット1222は、サービスゲートウェイ1236からの信号を受信するように構成することができる。この実施形態では、制御プレーンVCN1216およびデータプレーンVCN1218は、パブリックインターネット1254を呼び出すことなく、IaaSプロバイダの顧客によって呼び出されるように構成できる。IaaSプロバイダの顧客は、顧客が使用するデータベースがIaaSプロバイダによって制御され、パブリックインターネット1254から分離され得るサービステナント1219に記憶され得るため、この実施形態を所望し得る。

【0166】

図13は、少なくとも1つの実施形態による、IaaSアーキテクチャの別の例のパターンを示すブロック図1300である。サービスオペレータ1302（例えば、図12のサービスオペレータ1202）は、仮想クラウドネットワーク（VCN）1306（例えば、図12のVCN1206）およびセキュアホストサブネット1308（例えば、図12のセキュアホストサブネット1208）を含み得るセキュアホストテナント1304（例えば、図12のセキュアホストテナント1204）に通信可能に結合できる。VCN1306には、SSH VCN1312に含まれるLPG1310を介してセキュアシェル（SSH）VCN1312（例えば、図12のSSH VCN1212）に通信可能に結合できる局所ピアリングゲートウェイ（LPG）1310（例えば、図12のLPG1210）を含めることができる。SSH VCN1312には、SSHサブネット1314（例えば、図12のSSHサブネット1214）が含まれ得、SSH VCN1312は、制御プレーンVCN1316に含まれるLPG1310を介して制御プレーンVCN1316（例えば、図12の制御プレーンVCN1216）に通信可能に結合されることができる。制御プレーンVCN1316は、サービステナント1319（例えば、図12のサービステナント1219）に含めることができ、データプレーンVCN1318（例えば、図12のデータプレーンVCN1218）は、システムのユーザまたは顧客が所有ま

たは動作する可能性のある顧客テナント1321に含めることができる。

【0167】

制御プレーンVCN1316は、LBサブネット1322（例えば、図12のLBサブネット1222）を含むことができる制御プレーンDMZ層1320（例えば、図12の制御プレーンDMZ層1220）、アプリサブネット1326（例えば、図12のアプリサブネット1226）を含むことができる制御プレーンアプリ層1324（例えば、図12の制御プレーンアプリ層1224）、データベース（DB）サブネット1330（例えば、図12のDBサブネット1230と同様）を含むことができる制御プレーンデータ層1328（例えば、図12の制御プレーンデータ層1228）を含むことができる。制御プレーンDMZ層1320に含まれるLBサブネット1322は、制御プレーンアプリ層1324に含まれるアプリサブネット1326および制御プレーンVCN1316に含まれ得るインターネットゲートウェイ1334（例えば、図12のインターネットゲートウェイ1234）と通信可能に結合することができ、アプリサブネット1326は、制御プレーンデータ層1328に含まれるDBサブネット1330およびサービスゲートウェイ1336（例えば、図12のサービスゲートウェイ）およびネットワークアドレス変換（NAT）ゲートウェイ1338（例えば、図12のNATゲートウェイ1238）と通信可能に結合することができる。制御プレーンVCN1316には、サービスゲートウェイ1336とNATゲートウェイ1338を含めることができる。

10

【0168】

制御プレーンVCN1316には、アプリサブネット1326を含むことができるデータプレーンミラーアプリ層1340（例えば、図12のデータプレーンミラーアプリ層1240）を含めることができる。データプレーンミラーアプリ層1340に含まれるアプリサブネット1326には、コンピューティングインスタンス1344（例えば、図12のコンピューティングインスタンス1244と同様）を実行できる仮想ネットワークインターフェースコントローラ（VNIC）1342（例えば、1242のVNIC）を含めることができる。コンピューティングインスタンス1344は、データプレーンミラーアプリ層1340に含まれるVNIC1342とデータプレーンアプリ層1346に含まれるVNIC1342を介して、データプレーンミラーアプリ層1340のアプリサブネット1326と、データプレーンアプリ層1346（例えば、図12のデータプレーンアプリ層1246）に含まれ得るアプリサブネット1326との間の通信を容易にすることができる。

20

30

【0169】

制御プレーンVCN1316に含まれるインターネットゲートウェイ1334は、パブリックインターネット1354（例えば、図12のパブリックインターネット1254）に通信可能に結合できるメタデータ管理サービス1352（例えば、図12のメタデータ管理サービス1252）に通信可能に結合できる。パブリックインターネット1354は、制御プレーンVCN1316に含まれるNATゲートウェイ1338と通信可能に結合できる。制御プレーンVCN1316に含まれるサービスゲートウェイ1336は、クラウドサービス1356（例えば、図12のクラウドサービス1256）と通信可能に結合できる。

40

【0170】

いくつかの例では、データプレーンVCN1318は顧客テナント1321に含まれることができる。この場合、IaaSプロバイダは、各顧客に制御プレーンVCN1316を提供することができ、また、IaaSプロバイダは、各顧客に対して、サービステナント1319に含まれる固有のコンピューティングインスタンス1344を設定することができる。各コンピューティングインスタンス1344は、サービステナント1319に含まれる制御プレーンVCN1316と、顧客テナント1321に含まれるデータプレーンVCN1318との間の通信を可能にすることができる。コンピューティングインスタンス1344は、サービステナント1319に含まれる制御プレーンVCN1316でプロビジョニングされるリソースを、顧客テナント1321に含まれるデータプレーンVCN

50

1 3 1 8に展開したり、その他の方法で使用したりすることを可能にする。

【0 1 7 1】

他の例では、I a a Sプロバイダの顧客は、顧客テナント1 3 2 1内に存在するデータベースを有し得る。この例では、制御プレーンV C N 1 3 1 6には、アプリサブネット1 3 2 6を含めることができるデータプレーンミラーアプリ層1 3 4 0を含めることができる。データプレーンミラーアプリ層1 3 4 0はデータプレーンV C N 1 3 1 8内に存在できるが、データプレーンミラーアプリ層1 3 4 0はデータプレーンV C N 1 3 1 8内に存在できない場合がある。つまり、データプレーンミラーアプリ層1 3 4 0は顧客テナント1 3 2 1にアクセスできる可能性があるが、データプレーンミラーアプリ層1 3 4 0はデータプレーンV C N 1 3 1 8に存在しないか、I a a Sプロバイダの顧客によって所有または動作されていない可能性がある。データプレーンミラーアプリ層1 3 4 0は、データプレーンV C N 1 3 1 8への呼び出しを行うように構成できるが、制御プレーンV C N 1 3 1 6に含まれるエンティティへの呼び出しを行うように構成することはできない。顧客は、制御プレーンV C N 1 3 1 6にプロビジョニングされるリソースをデータプレーンV C N 1 3 1 8に展開したり、その他の方法で使用したりすることを所望する場合があり、データプレーンミラーアプリ層1 3 4 0は、顧客の所望する展開、またはリソースのその他の使用を容易にすることができる。

10

【0 1 7 2】

いくつかの実施形態では、I a a Sプロバイダの顧客は、データプレーンV C N 1 3 1 8にフィルタを適用することができる。この実施形態では、顧客はデータプレーンV C N 1 3 1 8がアクセスできるものを決定でき、顧客はデータプレーンV C N 1 3 1 8からパブリックインターネット1 3 5 4へのアクセスを制限できる。I a a Sプロバイダは、フィルタを適用したり、データプレーンV C N 1 3 1 8から外部のネットワークやデータベースへのアクセスを制御できない場合がある。顧客テナント1 3 2 1に含まれるデータプレーンV C N 1 3 1 8に顧客によるフィルタと制御を適用すると、データプレーンV C N 1 3 1 8を他の顧客およびパブリックインターネット1 3 5 4から分離するのに役立ち得る。

20

【0 1 7 3】

いくつかの実施形態では、クラウドサービス1 3 5 6は、サービスゲートウェイ1 3 3 6によって呼び出され、パブリックインターネット1 3 5 4、制御プレーンV C N 1 3 1 6、またはデータプレーンV C N 1 3 1 8に存在しない可能性があるサービスにアクセスすることができる。クラウドサービス1 3 5 6と制御プレーンV C N 1 3 1 6またはデータプレーンV C N 1 3 1 8間の接続は、存在または継続的ではない場合がある。クラウドサービス1 3 5 6は、I a a Sプロバイダが所有または動作する別のネットワーク上に存在し得る。クラウドサービス1 3 5 6は、サービスゲートウェイ1 3 3 6からの呼び出しを受信するように構成され、パブリックインターネット1 3 5 4からの呼び出しを受信しないように構成され得る。一部のクラウドサービス1 3 5 6は他のクラウドサービス1 3 5 6から分離され得、制御プレーンV C N 1 3 1 6は、制御プレーンV C N 1 3 1 6と同じリージョンにない場合があるクラウドサービス1 3 5 6から分離され得る。例えば、制御プレーンV C N 1 3 1 6は「リージョン1」に配置され得、クラウドサービス「展開1 2」はリージョン1と「リージョン2」に配置され得る。リージョン1にある制御プレーンV C N 1 3 1 6に含まれるサービスゲートウェイ1 3 3 6によって展開1 2への呼び出しが行われた場合、その呼び出しはリージョン1の展開1 2に送信され得る。この例では、制御プレーンV C N 1 3 1 6、つまりリージョン1の展開1 2は、リージョン2の展開1 2と通信可能に結合されていないか、または通信していない可能性がある。

30

40

【0 1 7 4】

図1 4は、少なくとも1つの実施形態による、I a a Sアーキテクチャの別の例のパターンを示すブロック図1 4 0 0である。サービスオペレータ1 4 0 2（例えば、図1 2のサービスオペレータ1 2 0 2）は、仮想クラウドネットワーク（V C N）1 4 0 6（例えば、図1 2のV C N 1 2 0 6）およびセキュアホストサブネット1 4 0 8（例えば、図1

50

2のセキュアホストサブネット1208)を含み得るセキュアホストテナント1404(例えば、図12のセキュアホストテナント1204)に通信可能に結合できる。VCN1406には、SSH VCN1412に含まれるLPG1410を介してSSH VCN1412(例えば、図12のSSH VCN1212)に通信可能に結合できるLPG1410(例えば、図12のLPG1210)を含めることができる。SSH VCN1412には、SSHサブネット1414(例えば、図12のSSHサブネット1214)が含まれ得、SSH VCN1412は、制御プレーンVCN1416(例えば、図12の制御プレーンVCN1216)に含まれるLPG1410を介して制御プレーンVCN1416に通信可能に結合され得、データプレーンVCN1418(例えば、図12のデータプレーン1218)に含まれるLPG1410を介してデータプレーンVCN1418に通信可能に結合され得る。制御プレーンVCN1416およびデータプレーンVCN1418は、サービステナント1419(例えば、図12のサービステナント1219)に含めることができる。

10

【0175】

制御プレーンVCN1416には、負荷分散装置(LB)サブネット1422(例えば、図12のLBサブネット1222)を含むことができる制御プレーンDMZ層1420(例えば、図12の制御プレーンDMZ層1220)、アプリサブネット1426(図12のアプリサブネット1226と同様)を含むことができる制御プレーンアプリ層1424(例えば、図12の制御プレーンアプリ層1224)、DBサブネット1430を含むことができる制御プレーンデータ層1428(例えば、図12の制御プレーンデータ層1228)が含まれ得る。制御プレーンDMZ層1420に含まれるLBサブネット1422は、制御プレーンアプリ層1424に含まれるアプリサブネット1426および制御プレーンVCN1416に含まれ得るインターネットゲートウェイ1434(例えば、図12のインターネットゲートウェイ1234)に通信可能に結合することができ、アプリサブネット1426は、制御プレーンデータ層1428に含まれるDBサブネット1430およびサービスゲートウェイ1436(例えば、図12のサービスゲートウェイ)およびネットワークアドレス変換(NAT)ゲートウェイ1438(例えば、図12のNATゲートウェイ1238)に通信可能に結合することができる。制御プレーンVCN1416には、サービスゲートウェイ1436とNATゲートウェイ1438を含めることができる。

20

30

【0176】

データプレーンVCN1418は、データプレーンアプリ層1446(例えば、図12のデータプレーンアプリ層1246)、データプレーンDMZ層1448(例えば、図12のデータプレーンDMZ層1248)、およびデータプレーンデータ層1450(例えば、図12のデータプレーンデータ層1250)を含むことができる。データプレーンDMZ層1448には、データプレーンアプリ層1446の信頼できるアプリサブネット1460および信頼できないアプリサブネット1462、およびデータプレーンVCN1418に含まれるインターネットゲートウェイ1434に通信可能に結合できるLBサブネット1422を含めることができる。信頼できるアプリサブネット1460は、データプレーンVCN1418に含まれるサービスゲートウェイ1436、データプレーンVCN1418に含まれるNATゲートウェイ1438、およびデータプレーンデータ層1450に含まれるDBサブネット1430に通信可能に結合できる。信頼できないアプリサブネット1462は、データプレーンVCN1418に含まれるサービスゲートウェイ1436およびデータプレーンデータ層1450に含まれるDBサブネット1430に通信可能に結合できる。データプレーンデータ層1450には、データプレーンVCN1418に含まれるサービスゲートウェイ1436に通信可能に結合できるDBサブネット1430を含めることができる。

40

【0177】

信頼できないアプリサブネット1462には、テナント仮想マシン(VM)1466(1)-(N)に通信可能に結合できる1つまたは複数のプライマリVNIC1464(

50

1) - (N)を含めることができる。各テナントVM1 466 (1) - (N)は、それぞれの顧客テナント1470 (1) - (N)に含まれ得るそれぞれのコンテナ出口VCN1468 (1) - (N)に含まれ得るそれぞれのアプリサブネット1467 (1) - (N)に通信可能に結合することができる。それぞれのセカンダリVNIC1472 (1) - (N)は、データプレーンVCN1418に含まれる信頼できないアプリサブネット1462とコンテナ出口VCN1468 (1) - (N)に含まれるアプリサブネット間の通信を容易にすることができる。各コンテナ出口VCN1468 (1) - (N)には、パブリックインターネット1454 (例えば、図12のパブリックインターネット1254)に通信可能に結合できるNATゲートウェイ1438を含めることができる。

【0178】

10

制御プレーンVCN1416に含まれ、データプレーンVCN1418に含まれるインターネットゲートウェイ1434は、パブリックインターネット1454に通信可能に結合できるメタデータ管理サービス1452 (例えば、図12のメタデータ管理システム1252)に通信可能に結合できる。パブリックインターネット1454は、制御プレーンVCN1416に含まれるNATゲートウェイ1438およびデータプレーンVCN1418に含まれるNATゲートウェイ1438に通信可能に結合できる。制御プレーンVCN1416に含まれ、データプレーンVCN1418に含まれるサービスゲートウェイ1436は、クラウドサービス1456と通信可能に結合できる。

【0179】

20

いくつかの実施形態では、データプレーンVCN1418は顧客テナント1470と統合することができる。この統合は、コード実行時にサポートを所望する場合など、いくつかの場合で、IaaSプロバイダの顧客にとって役立つ、または望ましい場合がある。顧客は、破壊的であったり、他の顧客リソースと通信したり、その他の望ましくない影響を引き起こす可能性のあるコードを実行するように提供し得る。これに応じて、IaaSプロバイダは、顧客からIaaSプロバイダに提供されるコードを実行するかどうかを決定できる。

【0180】

30

いくつかの例では、IaaSプロバイダの顧客は、IaaSプロバイダに一時的なネットワークアクセスを許可し、データプレーン層アプリ1446に接続する機能を要求することができる。関数を実行するコードはVM1 466 (1) - (N)で実行され得、コードはデータプレーンVCN1418上の他の場所で実行されるように構成されない可能性がある。各VM1 466 (1) - (N)は1つの顧客テナント1470に接続できる。VM1 466 (1) - (N)に含まれるそれぞれのコンテナ1471 (1) - (N)は、コードを実行するように構成されてもよい。この場合、二重の分離 (例えば、コンテナ1471 (1) - (N)がコードを実行し、コンテナ1471 (1) - (N)が少なくとも信頼できないアプリサブネット1462に含まれるVM1 466 (1) - (N)に含まれ得る)が存在し得、これにより、誤ったコードまたは望ましくないコードがIaaSプロバイダのネットワークに損害を与えたり、別の顧客のネットワークに損害を与えたりするのを防ぐことに役立つことができる。コンテナ1471 (1) - (N)は、顧客テナント1470に通信可能に結合され、顧客テナント1470からデータを送信または受信するように構成され得る。コンテナ1471 (1) - (N)は、データプレーンVCN1418内の他のエンティティからデータを送信または受信するように構成されない場合がある。コードの実行が完了すると、IaaSプロバイダはコンテナ1471 (1) - (N)を強制終了するか、その他の方法で処分することができる。

40

【0181】

いくつかの実施形態では、信頼できるアプリサブネット1460は、IaaSプロバイダが所有または動作し得るコードを実行し得る。この実施形態では、信頼できるアプリサブネット1460は、DBサブネット1430と通信可能に結合され、DBサブネット1430でCRUD動作を実行するように構成され得る。信頼できないアプリサブネット1462はDBサブネット1430と通信可能に結合され得るが、この実施形態では、信頼

50

できないアプリサブネットはDBサブネット1430で読み取り動作を実行するように構成され得る。各顧客のVM1466(1) - (N)に含まれ、顧客からのコードを実行する可能性のあるコンテナ1471(1) - (N)は、DBサブネット1430と通信可能に結合されていない可能性がある。

【0182】

他の実施形態では、制御プレーンVCN1416とデータプレーンVCN1418は、通信可能に直接結合されない場合がある。この実施形態では、制御プレーンVCN1416とデータプレーンVCN1418の間に直接通信が行われない場合がある。しかし、少なくとも1つの方法を通じて間接的に通信が行われ得る。IaaSプロバイダによって、制御プレーンVCN1416とデータプレーンVCN1418間の通信を容易にできるLPG1410が確立され得る。別の例では、制御プレーンVCN1416またはデータプレーンVCN1418は、サービスゲートウェイ1436を介してクラウドサービス1456を呼び出すことができる。例えば、制御プレーンVCN1416からクラウドサービス1456への呼び出しには、データプレーンVCN1418と通信できるサービスの要求が含まれ得る。

【0183】

図15は、少なくとも1つの実施形態による、IaaSアーキテクチャの別の例のパターンを示すブロック図1500である。サービスオペレータ1502(例えば、図12のサービスオペレータ1202)は、仮想クラウドネットワーク(VCN)1506(例えば、図12のVCN1206)およびセキュアホストサブネット1508(例えば、図12のセキュアホストサブネット1208)を含み得るセキュアホストテナント1504(例えば、図12のセキュアホストテナント1204)に通信可能に結合できる。VCN1506には、SSHVCN1512に含まれるLPG1510を介してSSHVCN1512(例えば、図12のSSHVCN1212)に通信可能に結合できるLPG1510(例えば、図12のLPG1210)を含めることができる。SSHVCN1512には、SSHサブネット1514(例えば、図12のSSHサブネット1214)が含まれ得、SSHVCN1512は、制御プレーンVCN1516(例えば、図12の制御プレーンVCN1216)に含まれるLPG1510を介して制御プレーンVCN1516に通信可能に結合され得、データプレーンVCN1518(例えば、図12のデータプレーン1218)に含まれるLPG1510を介してデータプレーンVCN1518に通信可能に結合され得る。制御プレーンVCN1516およびデータプレーンVCN1518は、サービステナント1519(例えば、図12のサービステナント1219)に含めることができる。

【0184】

制御プレーンVCN1516は、LBサブネット1522(例えば、図12のLBサブネット1222)を含むことができる制御プレーンDMZ層1520(例えば、図12の制御プレーンDMZ層1220)、アプリサブネット1526(例えば、図12のアプリサブネット1226)を含むことができる制御プレーンアプリ層1524(例えば、図12の制御プレーンアプリ層1224)、DBサブネット1530(例えば、図14のDBサブネット1430)を含むことができる制御プレーンデータ層1528(例えば、図12の制御プレーンデータ層1228)を含むことができる。制御プレーンDMZ層1520に含まれるLBサブネット1522は、制御プレーンアプリ層1524に含まれるアプリサブネット1526および制御プレーンVCN1516に含まれ得るインターネットゲートウェイ1534(例えば、図12のインターネットゲートウェイ1234)に通信可能に結合することができる。アプリサブネット1526は、制御プレーンデータ層1528に含まれるDBサブネット1530およびサービスゲートウェイ1536(例えば、図12のサービスゲートウェイ)およびネットワークアドレス変換(NAT)ゲートウェイ1538(例えば、図12のNATゲートウェイ1238)に通信可能に結合することができる。制御プレーンVCN1516には、サービスゲートウェイ1536とNATゲートウェイ1538を含めることができる。

10

20

30

40

50

【0185】

データプレーンVCN1518は、データプレーンアプリ層1546（例えば、図12のデータプレーンアプリ層1246）、データプレーンDMZ層1548（例えば、図12のデータプレーンDMZ層1248）、およびデータプレーンデータ層1550（例えば、図12のデータプレーンデータ層1250）を含むことができる。データプレーンDMZ層1548には、データプレーンアプリ層1546の信頼できるアプリサブネット1560（例えば、図14の信頼できるアプリサブネット1460）および信頼できないアプリサブネット1562（例えば、図14の信頼できないアプリサブネット1462）、およびデータプレーンVCN1518に含まれるインターネットゲートウェイ1534に通信可能に結合できるLBサブネット1522を含めることができる。信頼できるアプリサブネット1560は、データプレーンVCN1518に含まれるサービスゲートウェイ1536、データプレーンVCN1518に含まれるNATゲートウェイ1538、およびデータプレーンデータ層1550に含まれるDBサブネット1530に通信可能に結合できる。信頼できないアプリサブネット1562は、データプレーンVCN1518に含まれるサービスゲートウェイ1536およびデータプレーンデータ層1550に含まれるDBサブネット1530に通信可能に結合できる。データプレーンデータ層1550には、データプレーンVCN1518に含まれるサービスゲートウェイ1536に通信可能に結合できるDBサブネット1530を含めることができる。

10

【0186】

信頼できないアプリサブネット1562には、信頼できないアプリサブネット1562内に存在するテナント仮想マシン（VM）1566（1）－（N）と通信可能に結合できるプライマリVNIC1564（1）－（N）を含めることができる。各テナントVM1566（1）－（N）は、それぞれのコンテナ1567（1）－（N）内でコードを実行し、コンテナ出口VCN1568内に含まれ得るデータプレーンアプリ層1546内に含まれ得るアプリサブネット1526に通信可能に結合される。それぞれのセカンダリVNIC1572（1）－（N）は、データプレーンVCN1518に含まれる信頼できないアプリサブネット1562とコンテナ出口VCN1568に含まれるアプリサブネット間の通信を容易にすることができる。コンテナ出口VCNには、パブリックインターネット1554（例えば、図12のパブリックインターネット1254）に通信可能に結合できるNATゲートウェイ1538を含めることができる。

20

30

【0187】

制御プレーンVCN1516に含まれ、データプレーンVCN1518に含まれるインターネットゲートウェイ1534は、パブリックインターネット1554に通信可能に結合できるメタデータ管理サービス1552（例えば、図12のメタデータ管理システム1252）に通信可能に結合できる。パブリックインターネット1554は、制御プレーンVCN1516に含まれるNATゲートウェイ1538およびデータプレーンVCN1518に含まれるNATゲートウェイ1538に通信可能に結合できる。制御プレーンVCN1516に含まれ、データプレーンVCN1518に含まれるサービスゲートウェイ1536は、クラウドサービス1556と通信可能に結合できる。

【0188】

いくつかの例では、図15のブロック図1500のアーキテクチャによって示されるパターンは、図14のブロック図1400のアーキテクチャによって示されるパターンの例外とみなされる場合があり、IaaSプロバイダが顧客と直接通信できない場合（例えば、切断されるリージョン）には、IaaSプロバイダの顧客にとって望ましい場合がある。各顧客のVM1566（1）－（N）に含まれるそれぞれのコンテナ1567（1）－（N）は、顧客がリアルタイムでアクセスできる。コンテナ1567（1）－（N）は、コンテナ出口VCN1568に含まれ得るデータプレーンアプリ層1546のアプリサブネット1526に含まれるそれぞれのセカンダリVNIC1572（1）－（N）を呼び出すように構成できる。セカンダリVNIC1572（1）－（N）は、NATゲートウェイ1538に呼び出しを送信することができ、NATゲートウェイ1538は、呼び

40

50

出しをパブリックインターネット 1554 に送信することができる。この例では、顧客がリアルタイムでアクセスできるコンテナ 1567 (1) - (N) は、制御プレーン VCN 1516 から分離でき、データプレーン VCN 1518 に含まれる他のエンティティからも分離できる。また、コンテナ 1567 (1) - (N) は、他の顧客のリソースから分離され得る。

【0189】

他の例では、顧客はコンテナ 1567 (1) - (N) を使用してクラウドサービス 1556 を呼び出すことができる。この例では、顧客はコンテナ 1567 (1) - (N) 内でクラウドサービス 1556 からサービスを要求するコードを実行し得る。コンテナ 1567 (1) - (N) は、この要求をセカンダリ VNIC 1572 (1) - (N) に送信でき、セカンダリ VNIC 1572 (1) - (N) は、要求を NAT ゲートウェイに送信でき、NAT ゲートウェイは、要求をパブリックインターネット 1554 に送信できる。パブリックインターネット 1554 は、インターネットゲートウェイ 1534 を介して、制御プレーン VCN 1516 に含まれる LB サブネット 1522 に要求を送信できる。要求が有効であると決定される場合、LB サブネットは、サービスゲートウェイ 1536 を介してクラウドサービス 1556 に要求を送信できるアプリサブネット 1526 に要求を送信できる。

10

【0190】

図に示される IaaS アーキテクチャ 1200、1300、1400、1500 は、図に示されるもの以外の構成要素を有し得ることに留意されたい。さらに、図に示される実施形態は、本開示の実施形態を組み込むことができるクラウドインフラストラクチャシステムのいくつかの例にすぎない。いくつかの他の実施形態では、IaaS システムは、図に示されているよりも多くの構成要素またはより少ない構成要素を有し、2 つ以上の構成要素を組み合わせたり、構成要素の構成または配置が異なったりし得る。

20

【0191】

特定の実施形態では、本明細書で説明する IaaS システムには、セルフサービス、加入ベース、弾力的にスケラブル、信頼性、可用性が高く、安全な方法で顧客に提供されるアプリケーション、ミドルウェア、およびデータベースサービスオフリングのスイートが含まれ得る。このような IaaS システムの一例としては、本譲受人が提供する Oracle Cloud Infrastructure (OCI) が挙げられる。

30

【0192】

図 16 は、さまざまな実施形態が実装され得る例示的なコンピュータシステム 1600 を示す。システム 1600 は、上記のいずれかのコンピュータシステムを実装するために使用できる。図に示すように、コンピュータシステム 1600 には、バスサブシステム 1602 を介して多数の周辺サブシステムと通信する処理装置 1604 が含まれている。これらの周辺サブシステムには、処理加速装置 1606、I/O サブシステム 1608、記憶サブシステム 1618、および通信サブシステム 1624 が含まれ得る。記憶サブシステム 1618 には、有形のコンピュータ可読記憶媒体 1622 とシステムメモリ 1610 が含まれる。

【0193】

バスサブシステム 1602 は、コンピュータシステム 1600 のさまざまな構成要素とサブシステムが意図したとおりに相互に通信できるようにするメカニズムを提供する。バスサブシステム 1602 は単一のバスとして概略的に示されているが、バスサブシステムの代替の実施形態では複数のバスを利用し得る。バスサブシステム 1602 は、メモリバスまたはメモリコントローラ、周辺バス、およびさまざまなバスアーキテクチャのいずれかを使用する局所バスを含む、いくつかのタイプのバス構造のいずれかになり得る。例えば、このようなアーキテクチャには、業界標準アーキテクチャ (ISA) バス、マイクロチャネルアーキテクチャ (MCA) バス、拡張 ISA (EISA) バス、ビデオエレクトロニクス標準協会 (VESA) 局所バス、および周辺機器相互接続 (PCI) バスが含まれ得、これらは IEEE P1386.1 標準に従って製造されるメザニンバスとして実装

40

50

できる。

【0194】

処理装置1604は、1つまたは複数の集積回路（例えば、従来のマイクロプロセッサまたはマイクロコントローラ）として実装することができ、コンピュータシステム1600の動作を制御する。処理装置1604には1つまたは複数のプロセッサが含まれ得る。これらのプロセッサには、単一コアプロセッサまたはマルチコアプロセッサが含まれ得る。特定の実施形態では、処理装置1604は、各処理装置に単一またはマルチコアプロセッサが含まれる1つまたは複数の独立した処理装置1632および／または1634として実装され得る。他の実施形態では、処理装置1604は、2つのデュアルコアプロセッサを1つのチップに統合して形成されるクアッドコア処理装置として実装されることもできる。

10

【0195】

さまざまな実施形態において、処理装置1604は、プログラムコードに応じてさまざまなプログラムを実行することができ、複数の同時実行プログラムまたはプロセスを維持することができる。任意の時点で、実行されるプログラムコードの一部またはすべてが、プロセッサ1604および／または記憶サブシステム1618に存在し得る。適切なプログラミングにより、プロセッサ1604は上記のさまざまな機能を提供できる。コンピュータシステム1600には、さらに、デジタル信号プロセッサ（DSP）、専用プロセッサなどを含むことができる処理加速装置1606が含まれ得る。

【0196】

I/Oサブシステム1608には、ユーザインターフェース入力装置とユーザインターフェース出力装置が含まれ得る。ユーザインターフェース入力装置には、キーボード、マウスやトラックボールなどのポインティング装置、ディスプレイに組み込まれたタッチパッドまたはタッチスクリーン、スクロールホイール、クリックホイール、ダイヤル、ボタン、スイッチ、キーパッド、音声コマンド認識システムを備えたオーディオ入力装置、マイク、その他のタイプの入力装置が含まれ得る。ユーザインターフェース入力装置には、例えば、Microsoft Kinect（登録商標）モーションセンサーなどのモーションセンシング装置および／またはジェスチャ認識装置が含まれ得る。Microsoft Kinect（登録商標）モーションセンサーを使用すると、ユーザはジェスチャや音声コマンドを使用した自然なユーザインターフェースを通じて、Microsoft Xbox（登録商標）360ゲームコントローラなどの入力装置を制御および入力装置と対話できる。ユーザインターフェース入力装置には、ユーザの目の動き（例えば、写真を撮影しているときおよび／またはメニューを選択しているときの「まばたき」）を検出し、その目のジェスチャを入力装置（例えば、Google Glass（登録商標））への入力として変換するGoogle Glass（登録商標）まばたき検出器などの目のジェスチャ認識装置も含まれ得る。さらに、ユーザインターフェース入力装置には、ユーザが音声コマンドを通じて音声認識システム（例えば、Siri（登録商標）ナビゲータ）と対話できるようにする音声認識センシング装置が含まれ得る。

20

30

【0197】

ユーザインターフェース入力装置には、3次元（3D）マウス、ジョイスティックまたはポインティングスティック、ゲームパッドおよびグラフィックタブレット、ならびにスピーカ、デジタルカメラ、デジタルビデオカメラ、ポータブルメディアプレーヤ、ウェブカメラ、画像スキャナ、指紋スキャナ、バーコードリーダ3Dスキャナ、3Dプリンタ、レーザ距離計、視線追跡装置などのオーディオ／ビジュアル装置も含まれ得るが、これらに限定されない。さらに、ユーザインターフェース入力装置には、例えば、コンピュータ断層撮影、磁気共鳴画像化、位置放射断層撮影、医療用超音波検査装置などの医療用画像化入力装置が含まれ得る。ユーザインターフェース入力装置には、例えば、MIDIキーボード、デジタル楽器などのオーディオ入力装置も含まれ得る。

40

【0198】

ユーザインターフェース出力装置には、ディスプレイサブシステム、インジケータライ

50

ト、またはオーディオ出力装置などの非視覚ディスプレイなどが含まれ得る。表示サブシステムは、ブラウン管（CRT）、液晶ディスプレイ（LCD）やプラズマディスプレイを使用するフラットパネル装置、投影装置、タッチスクリーンなどであり得る。一般に、「出力装置」という用語の使用は、コンピュータシステム1600からユーザまたは他のコンピュータに情報を出力するためのあらゆるタイプの装置およびメカニズムを含むことを意図している。例えば、ユーザインターフェース出力装置には、モニタ、プリンタ、スピーカ、ヘッドフォン、自動車ナビゲーションシステム、プロッタ、音声出力装置、モデムなど、テキスト、グラフィック、オーディオ／ビデオ情報を視覚的に伝えるさまざまなディスプレイ装置が含まれ得るが、これらに限定されない。

【0199】

コンピュータシステム1600は、現在システムメモリ1610内に配置されているように示されるソフトウェア要素を含む記憶サブシステム1618を含むことができる。システムメモリ1610には、処理装置1604でロードおよび実行可能なプログラム命令、およびこれらのプログラムの実行中に生成されるデータが記憶され得る。

【0200】

コンピュータシステム1600の構成およびタイプに応じて、システムメモリ1610は揮発性（例えば、ランダムアクセスメモリ（RAM））および／または不揮発性（例えば、読み取り専用メモリ（ROM）、フラッシュメモリなど）であってもよい。RAMには通常、処理装置1604によってすぐにアクセス可能であり、または現在処理装置1604によって動作および実行されているデータおよび／またはプログラムモジュールが含まれている。いくつかの実装では、システムメモリ1610には、静的ランダムアクセスメモリ（SRAM）や動的ランダムアクセスメモリ（DRAM）などの複数の異なるタイプのメモリが含まれ得る。いくつかの実装では、起動時などコンピュータシステム1600内の要素間で情報を転送するのに役立つ基本ルーチンを含む基本入出力システム（BIOS）が、通常、ROMに記憶され得る。限定ではなく例として、システムメモリ1610には、クライアントアプリケーション、Webブラウザ、中間層アプリケーション、リレーショナルデータベース管理システム（RDBMS）などを含み得るアプリケーションプログラム1612、プログラムデータ1614、およびオペレーティングシステム1616も示されている。一例として、オペレーティングシステム1616には、Microsoft Windows（登録商標）、Apple Macintosh（登録商標）、および／またはLinuxオペレーティングシステムのさまざまなバージョン、市販のさまざまなUNIX（登録商標）またはUNIXライクなオペレーティングシステム（さまざまなGNU/Linuxオペレーティングシステム、Google Chrome（登録商標）OSなどを含むがこれらに限定されない）、および／またはiOS、Windows（登録商標）Phone、Android（登録商標）OS、BlackBerry（登録商標）16 OS、Palm（登録商標）OSオペレーティングシステムなどのモバイルオペレーティングシステムが含まれ得る。

【0201】

記憶サブシステム1618は、いくつかの実施形態の機能を提供する基本的なプログラミングおよびデータ構造を記憶するための有形のコンピュータ可読記憶媒体も提供することができる。プロセッサによって実行されると上記の機能を提供するソフトウェア（プログラム、コードモジュール、命令）は、記憶サブシステム1618に記憶され得る。これらのソフトウェアモジュールまたは命令は、処理装置1604によって実行され得る。記憶サブシステム1618は、本開示に従って使用されるデータを記憶するためのリポジトリも提供することができる。

【0202】

記憶サブシステム1600には、コンピュータ可読記憶媒体1622にさらに接続できるコンピュータ可読記憶媒体リーダ1620も含まれ得る。コンピュータ可読記憶媒体1622は、システムメモリ1610と一緒に、また任意選択で組み合わせて、リモート、局所、固定、および／または取り外し可能な記憶装置に加えて、コンピュータ可読情報を一時的および／またはより永続的に含有、記憶、送信、および取得するための記憶媒体を

10

20

30

40

50

包括的に表すことができる。非一時的なコンピュータ可読記憶媒体には、揮発性メモリ記憶装置および／または不揮発性記憶装置を含む物理的に有形のメモリまたは記憶装置が含まれ得る。非一時的なコンピュータ可読記憶媒体の例には、磁気記憶媒体（例えば、ディスクやテープ）、光学記憶媒体（例えば、DVD、CD）、さまざまなタイプのRAM、ROM、またはフラッシュメモリ、ハードドライブ、フロッピー（登録商標）ドライブ、取り外し可能なメモリドライブ（例えば、USBドライブ）、その他のタイプの記憶装置が含まれる。

【0203】

コードまたはコードの一部を含むコンピュータ可読記憶媒体1622には、情報の記憶および／または送信のための任意の方法または技術で実装される揮発性および不揮発性、取り外し可能および取り外し不可能な媒体など、記憶媒体および通信媒体を含むがこれらに限定されない、当該技術分野で既知または使用されている任意の適切な媒体も含まれ得る。これには、RAM、ROM、電子的に消去可能なプログラマブルROM（EEPROM）、フラッシュメモリまたはその他のメモリ技術、CD-ROM、デジタル多用途ディスク（DVD）、またはその他の光学式記憶装置、磁気カセット、磁気テープ、磁気ディスク記憶装置またはその他の磁気記憶装置、またはその他の有形のコンピュータ可読媒体などの有形のコンピュータ可読記憶媒体が含まれ得る。

【0204】

一例として、コンピュータ可読記憶媒体1622には、取り外し不可能な不揮発性磁気媒体から読み取りまたは書き込みを行うハードディスクドライブ、取り外し可能な不揮発性磁気ディスクから読み取りまたは書き込みを行う磁気ディスクドライブ、およびCDROM、DVD、Blu-Ray（登録商標）ディスクなどの取り外し可能な不揮発性光ディスク、またはその他の光媒体から読み取りまたは書き込みを行う光ディスクドライブが含まれ得る。コンピュータ可読記憶媒体1622には、Zip（登録商標）ドライブ、フラッシュメモリカード、ユニバーサルシリアルバス（USB）フラッシュドライブ、セキュアデジタル（SD）カード、DVDディスク、デジタルビデオテープなどが含まれ得るが、これらに限定されない。コンピュータ可読記憶媒体1622には、フラッシュメモリベースのSSD、エンタープライズフラッシュドライブ、ソリッドステートROMなどの不揮発性メモリに基づくソリッドステートドライブ（SSD）、ソリッドステートRAM、動的RAM、静的RAM、DRAMベースのSSD、磁気抵抗RAM（MRAM）SSD、およびDRAMとフラッシュメモリベースのSSDの組み合わせを使用するハイブリッドSSDなど、揮発性メモリに基づくSSDも含まれ得る。ディスクドライブおよびそれに関連するコンピュータ可読媒体は、コンピュータシステム1600のコンピュータ可読命令、データ構造、プログラムモジュール、およびその他のデータの不揮発性記憶を提供し得る。

【0205】

通信サブシステム1624は、他のコンピュータシステムおよびネットワークへのインターフェースを提供する。通信サブシステム1624は、コンピュータシステム1600から他のシステムとの間でデータを送受信するためのインターフェースとして機能する。例えば、通信サブシステム1624は、コンピュータシステム1600がインターネットを介して1つまたは複数の装置に接続できるようにする。いくつかの実施形態では、通信サブシステム1624には、無線音声および／またはデータネットワーク（例えば、携帯電話技術、3G、4G、EDGE（地球規模の進化のためのデータレートの向上）などの高度なデータネットワーク技術、WiFi（IEEE 802.11ファミリ標準、またはその他のモバイル通信技術、またはそれらの任意の組み合わせ）、グローバルポジショニングシステム（GPS）受信機構成要素、および／またはその他の構成要素を使用）にアクセスするための無線周波数（RF）トランシーバー構成要素が含まれ得る。いくつかの実施形態では、通信サブシステム1624は、無線インターフェースに加えて、または無線インターフェースの代わりに、有線ネットワーク接続（例えば、イーサネット）を提供できる。

10

20

30

40

50

【0206】

いくつかの実施形態では、通信サブシステム1624は、コンピュータシステム1600を使用する可能性のある1人または複数のユーザに代わって、構造化データフィードおよび／または非構造化データフィード1626、イベントストリーム1628、イベント更新1630などの形式で入力通信を受信することもできる。

【0207】

一例として、通信サブシステム1624は、Twitter（登録商標）フィード、Facebook（登録商標）アップデート、Rich Site Summary（RSS）フィードなどのWebフィード、および／または1つまたは複数のサードパーティ情報ソースからのリアルタイム更新など、ソーシャルネットワークおよび／またはその他の通信サービスのユーザからデータフィード1626をリアルタイムで受信するように構成できる。

10

【0208】

さらに、通信サブシステム1624は、連続データストリームの形式でデータを受信するように構成することもできる。連続データストリームには、明示的な終了がなく、本質的に連続的または無制限である可能性のあるリアルタイムイベントおよび／またはイベント更新1630のイベントストリーム1628が含まれ得る。連続データを生成するアプリケーションの例としては、例えば、センサーデータアプリケーション、金融ティッカー、ネットワーク性能測定ツール（例えば、ネットワーク監視およびトラフィック管理アプリケーション）、クリックストリーム分析ツール、自動車交通監視などが含まれ得る。

【0209】

通信サブシステム1624は、構造化データフィードおよび／または非構造化データフィード1626、イベントストリーム1628、イベント更新1630などを、コンピュータシステム1600に結合される1つまたは複数のストリーミングデータソースコンピュータと通信している可能性のある1つまたは複数のデータベースに出力するように構成することもできる。

20

【0210】

コンピュータシステム1600は、ハンドヘルドポータブル装置（例えば、iPhone（登録商標）携帯電話、iPad（登録商標）コンピューティングタブレット、PDA）、ウェアラブル装置（例えば、Google Glass（登録商標）ヘッドマウントディスプレイ）、PC、ワークステーション、メインフレーム、キオスク、サーバラック、またはその他のデータ処理システムを含む、さまざまなタイプのいずれかであり得る。

30

【0211】

コンピュータとネットワークは常に変化する性質のため、図に示されているコンピュータシステム1600の説明は、特定の例としてのみ意図されている。図に示されているシステムよりも多くの構成要素またはより少ない構成要素を持つ他の多くの構成が可能である。例えば、カスタマイズされるハードウェアも使用され得、および／または特定の要素がハードウェア、ファームウェア、ソフトウェア（アプレットを含む）、またはそれらの組み合わせで実装され得る。さらに、ネットワーク入出力装置などの他のコンピューティング装置への接続も使用できる。本明細書で提供される開示および教示に基づいて、当業者は、さまざまな実施形態を実装するための他の方法および／または手法を理解するであろう。

40

【0212】

特定の実施形態について説明したが、さまざまな修正、変更、代替構成、および同等のものも本開示の範囲内に含まれる。実施形態は、特定のデータ処理環境内での動作に限定されず、複数のデータ処理環境内で自由に動作することができる。さらに、実施形態は特定の一連のトランザクションおよびステップを使用して説明されているが、本開示の範囲は、説明されている一連のトランザクションおよびステップに限定されないことは、当業者には明らかであるはずである。上述の実施形態のさまざまな特徴および態様は、個別にまたは組み合わせで使用することができる。

【0213】

50

さらに、実施形態は特定のハードウェアとソフトウェアの組み合わせを使用して説明されているが、他のハードウェアとソフトウェアの組み合わせも本開示の範囲内にあることを認識すべきである。実施形態は、ハードウェアのみ、ソフトウェアのみ、またはそれらの組み合わせを使用して実装することができる。本明細書で説明するさまざまなプロセスは、同じプロセッサまたは異なるプロセッサの任意の組み合わせで実装できる。したがって、構成要素またはモジュールが特定の動作を実行するように構成されていると説明されている場合、そのような構成は、例えば、動作を実行するための電子回路を設計することによって、動作を実行するためのプログラマブル電子回路（例えば、マイクロプロセッサ）をプログラミングすることによって、またはそれらの任意の組み合わせによって実現できる。プロセスは、プロセス間通信の従来の技術を含むがこれに限定されないさまざまな技術を使用して通信することができ、異なるプロセスのペアが異なる技術を使用することも、同じプロセスのペアが異なる時間に異なる技術を使用することもできる。

10

【0214】

したがって、本明細書および図面は、限定的な意味ではなく、例示的な意味としてみなされるべきである。しかし、請求項に規定されるより広い趣旨および範囲から逸脱することなく、追加、控除、削除、およびその他の修正および変更が可能であることは明らかである。したがって、特定の開示実施形態が説明されているが、これらは限定することを意図するものではない。さまざまな変更および同等のものが、以下の請求項の範囲内にある。

【0215】

開示される実施形態を説明する文脈（特に以下の請求項の文脈）における用語「a」および「an」および「the」および類似の参照対象の使用は、本明細書で別途記載されていない限り、または文脈によって明らかに矛盾しない限り、単数形と複数形の両方を含むものと解釈される。「含む」、「有する」、「含む」、および「含有する」という用語は、特に明記されていない限り、オープンエンドの用語（つまり、「含むが、これらに限定されない」という意味）として解釈される。「接続」という用語は、何かが介在している場合でも、部分的または全体的に含まれている、接続されている、または結合されていると解釈される。本明細書における値の範囲の記載は、本明細書で特に明記されていない限り、その範囲内に含まれる各個別の値を個別に参照するための簡略な方法として機能することをただ目的としており、各個別の値は、本明細書で個別に記載されるかのように明細書に組み込まれる。本書で説明するすべての方法は、本書で別途記載されている場合、または文脈によって明らかに矛盾している場合を除き、任意の適切な順序で実行できる。本明細書で提供されるあらゆる例、または例示的な言語（例えば、「など」）の使用は、実施形態をよりわかりやすく説明することのみを意図しており、特に断りがない限り、開示の範囲を制限するものではない。明細書中のいかなる文言も、請求されていない要素が開示の実施に必須であることを示すものとして解釈されるべきではない。

20

30

【0216】

「X、Y、またはZのうちの少なくとも1つ」という語句などの離接語は、特に明記されていない限り、一般的に、項目、用語などがX、Y、またはZのいずれか、またはそれらの任意の組み合わせ（例えば、X、Y、および／またはZ）であり得ることを示すために使用される文脈内で理解されることが意図されている。したがって、このような離接語は、一般的に、特定の実施形態がXの少なくとも1つ、Yの少なくとも1つ、またはZの少なくとも1つを各々存在させる必要があることを暗示することを意図しておらず、また暗示すべきではない。

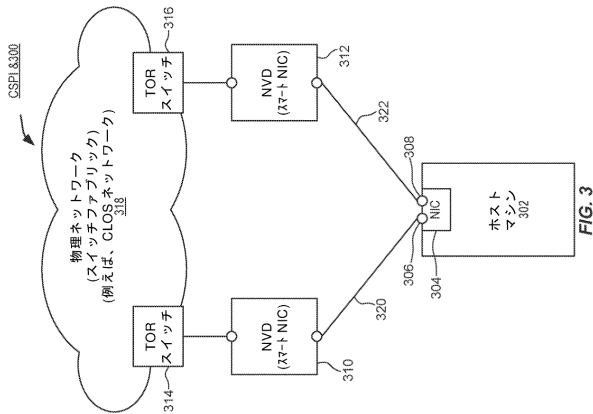
40

【0217】

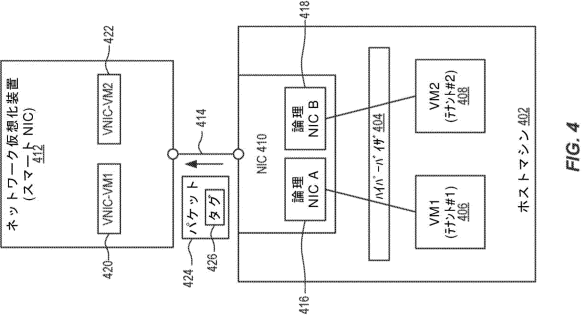
本開示の好ましい実施形態は、本開示を実施するために知られている最良のモードを含めて、本明細書に記載されている。これらの好ましい実施形態の変形は、前述の説明を読めば、当業者には明らかになるであろう。当業者であれば、必要に応じてこのような変形を採用することができ、本開示は、本明細書に具体的に記載されるものとは別の方法で実施されてもよい。したがって、本開示には、適用法によって許可される限り、本明細書に

50

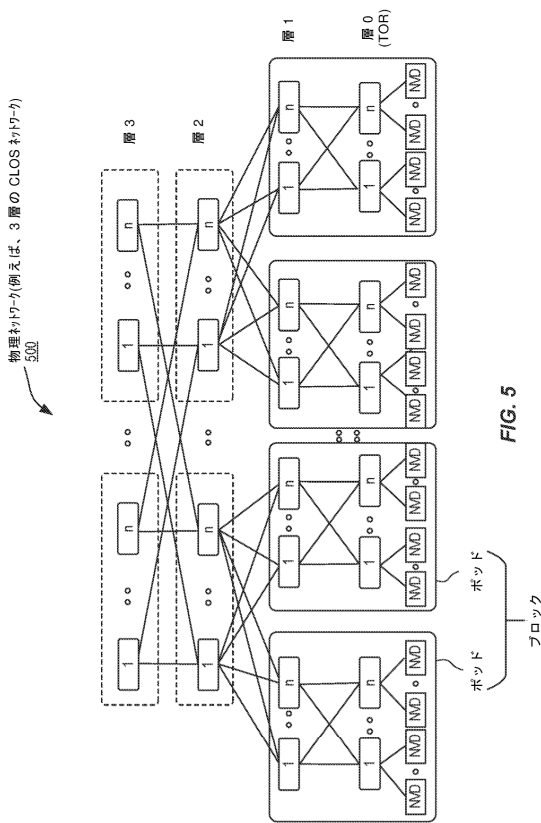
【図 3】



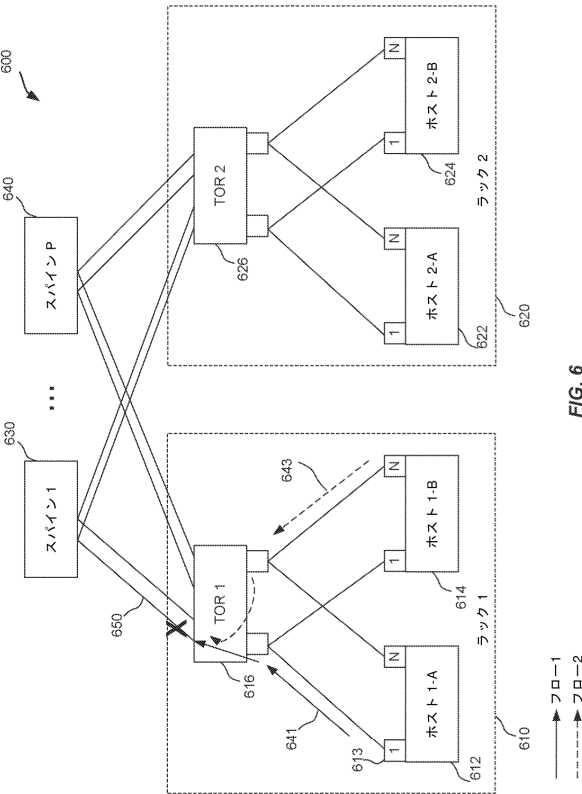
【図 4】



【図 5】



【図 6】



10

20

30

40

50

【図 7】

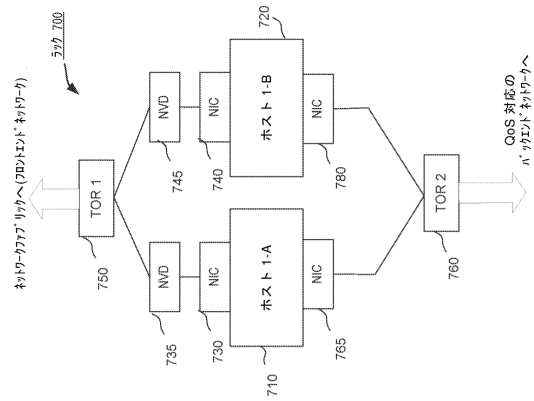


FIG. 7

【図 8】

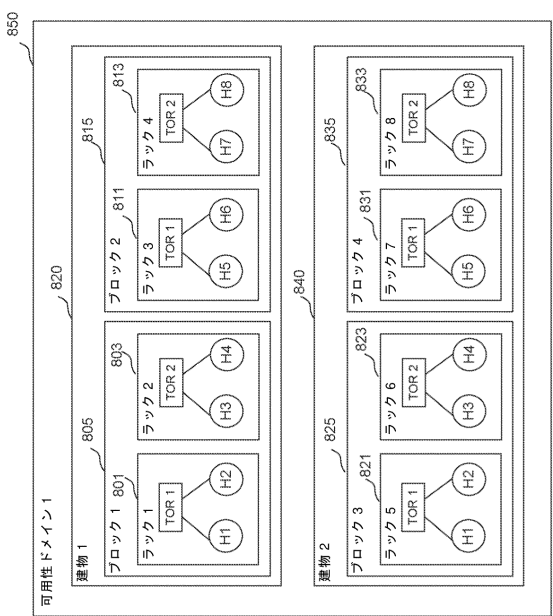


FIG. 8

【図 9】

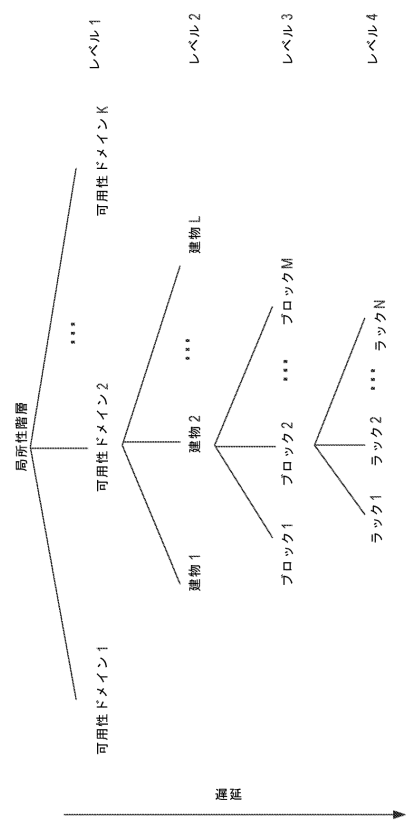


FIG. 9

【図 10】

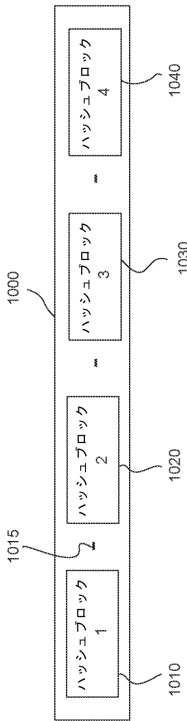


FIG. 10

【図 1 1 A】

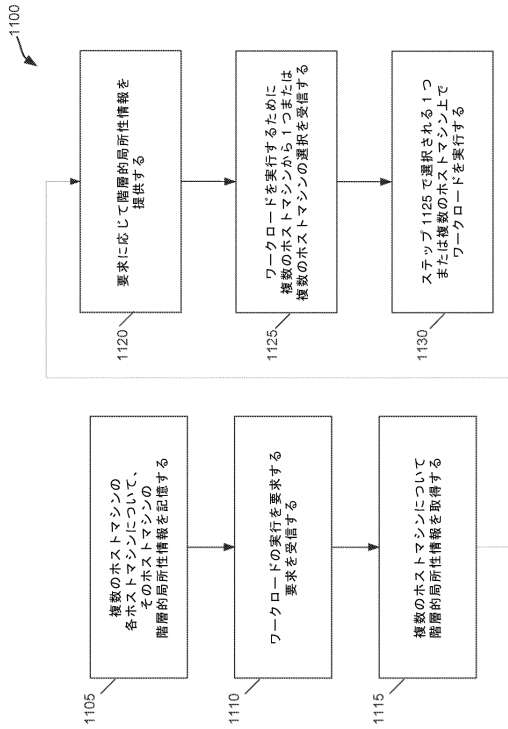


FIG. 11A

【図 1 1 B】

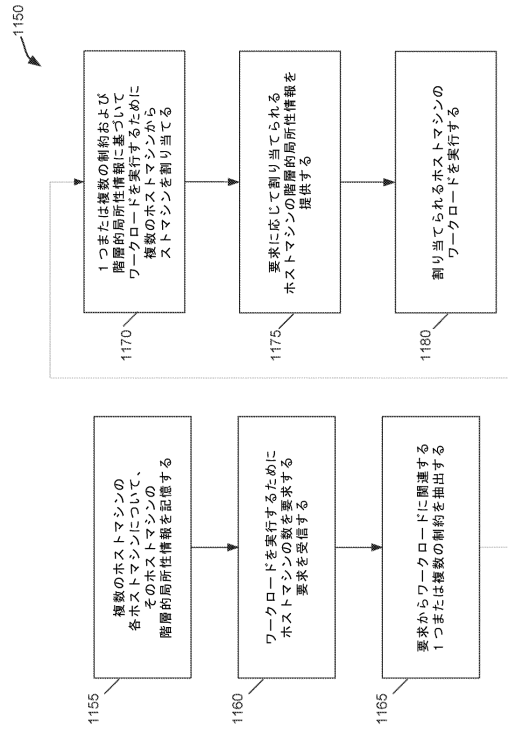


FIG. 11B

【図 12】

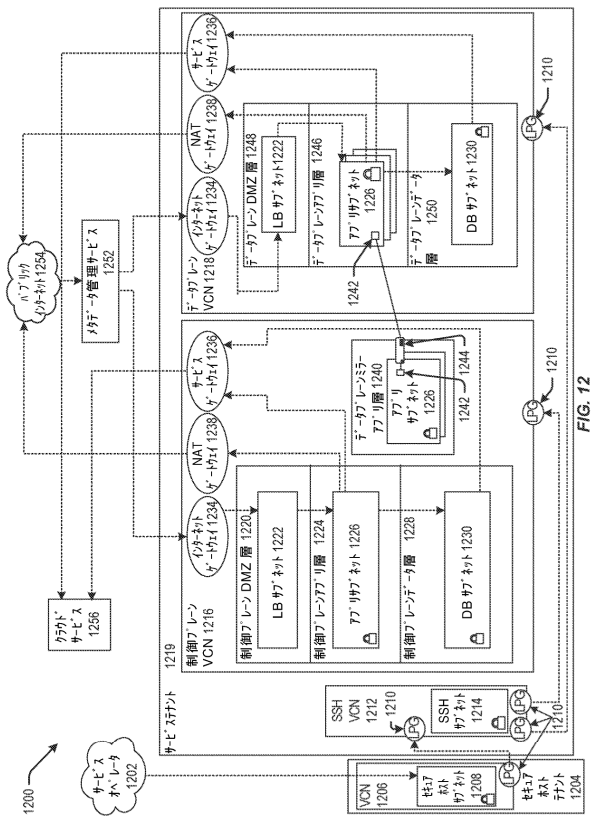


FIG. 12

【図 13】

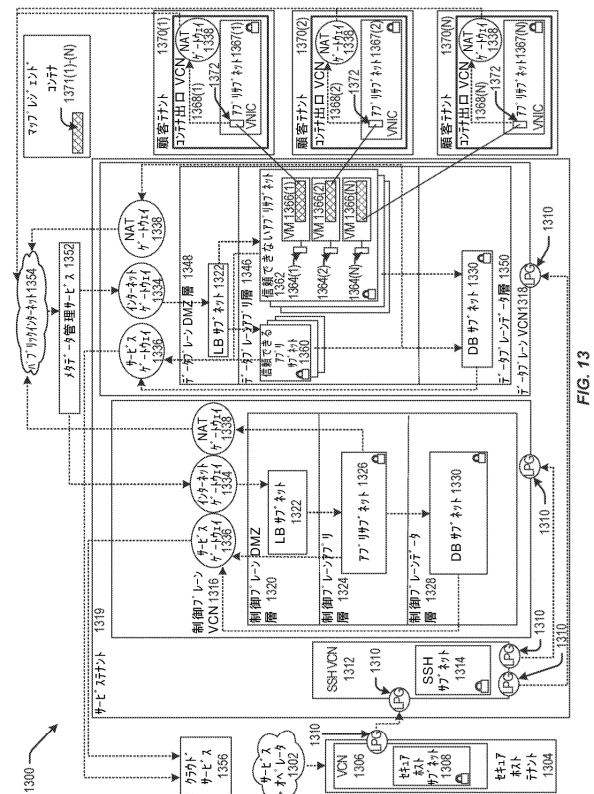


FIG. 13

【図 14】

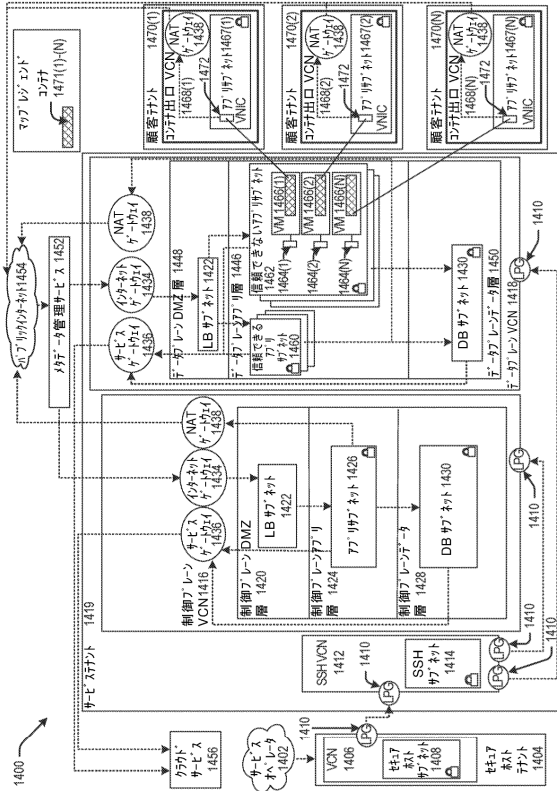


FIG. 14

【図 15】

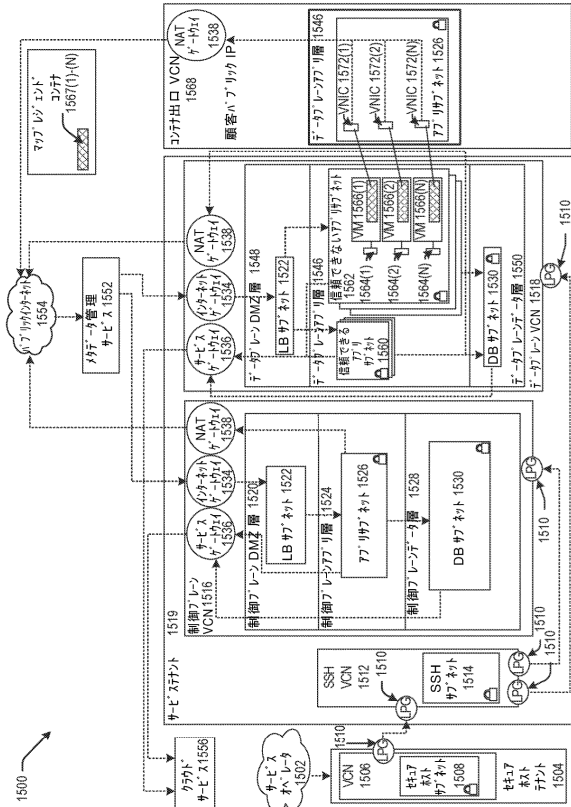


FIG. 15

【図 16】

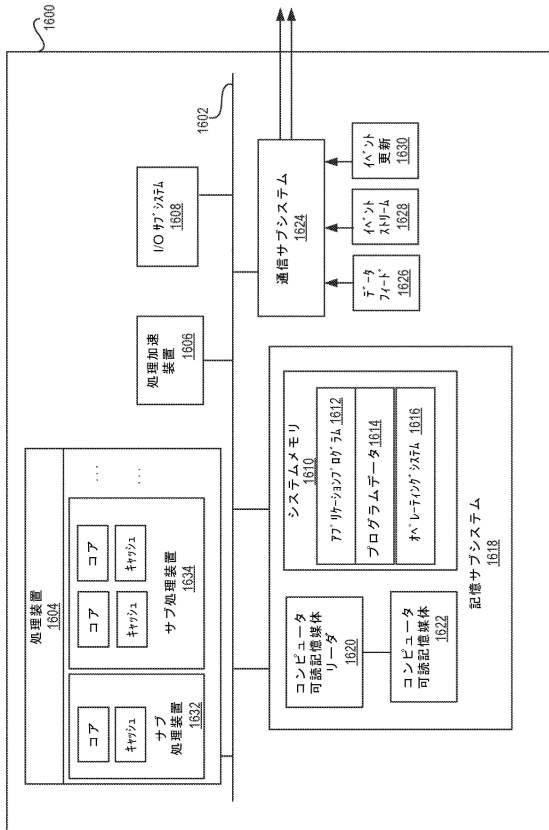


FIG. 16

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2022/082075

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 9 451 393 B1 (CULLEN PATRICK BRIGHAM [US] ET AL) 20 September 2016 (2016-09-20) column 4, line 43 - line 67 column 5, line 4 - line 7 column 9, line 22 - line 67 column 10, line 1 - line 46 column 11, line 5 - line 10 column 13, line 5 - line 12 column 6, line 1 - line 42 column 5, line 63 - line 67 -----	1-20
A	US 8 099 380 B1 (SHAHABI CYRUS [US] ET AL) 17 January 2012 (2012-01-17) column 6, line 47 - line 67 column 7, line 1 - line 28 ----- -/--	1-20

☒ 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

"&" document member of the same patent family

Date of the actual completion of the international search

31 March 2023

Date of mailing of the international search report

12/04/2023

Name and mailing address of the ISA/
European Patent Office, P.B. 5618 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Buzgan, C

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2022/082075

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2021/108358 A1 (AMAZON TECH INC [US]) 3 June 2021 (2021-06-03) paragraphs [0047] - [0049], [0098] - [0104], [0112], [0113] -----	1-20

1

10

20

30

40

50

INTERNATIONAL SEARCH REPORT				International application No PCT/US2022/082075	
Information on patent family members					
Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US 9451393	B1	20-09-2016	NONE		

US 8099380	B1	17-01-2012	NONE		

WO 2021108358	A1	03-06-2021	CN	114902183 A	12-08-2022
			EP	4049139 A1	31-08-2022
			US	10979534 B1	13-04-2021
			WO	2021108358 A1	03-06-2021

10

20

30

40

50

フロントページの続き

,NA,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,D
E,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,ME,MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,S
M,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,CV,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,
ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IQ,IR,IS,IT,JM,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,L
A,LC,LK,LR,LS,LU,LY,MA,MD,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,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,V
N,WS,ZA,ZM,ZW

(特許庁注：以下のものは登録商標)

1. iOS

- (72)発明者 ベッカー，デイビッド・デイル
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、オラクル・インターナショナル・コーポレーション内
- (72)発明者 コクマード，ハロルド・ドネル
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、オラクル・インターナショナル・コーポレーション内
- (72)発明者 ウィルフォード，アダム・フランクリン
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、オラクル・インターナショナル・コーポレーション内
- (72)発明者 ロビー，イルディコ・バルガ
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、オラクル・インターナショナル・コーポレーション内
- (72)発明者 ボッター，ジョシュア・スコット
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、オラクル・インターナショナル・コーポレーション内