(19)日本国特許庁(JP)　(12)公表特許公報(A)　(11)公表番号
特表2025-504416
(P2025-504416A)
(43)公表日　令和7年2月12日(2025.2.12)

(51)国際特許分類　FI
H04L 41/085 (2022.01)　H04L 41/085
H04L 41/0895 (2022.01)　H04L 41/0895
G06F 9/50 (2006.01)　G06F 9/50 150Z

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2024-541878(P2024-541878)
(86)(22)出願日　令和4年12月20日(2022.12.20)
(85)翻訳文提出日　令和6年8月19日(2024.8.19)
(86)国際出願番号　PCT/US2022/082073
(87)国際公開番号　WO2023/136964
(87)国際公開日　令和5年7月20日(2023.7.20)
(31)優先権主張番号　63/298,685
(32)優先日　令和4年1月12日(2022.1.12)
(33)優先権主張国・地域又は機関
米国(US)
(31)優先権主張番号　18/050,392
(32)優先日　令和4年10月27日(2022.10.27)
(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

(71)出願人　502303739
オラクル・インターナショナル・コーポレイション
アメリカ合衆国 ９４０６５ 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 ５００
(74)代理人　110001195
弁理士法人深見特許事務所
(72)発明者　ブラール，ジャグウィンダー
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・パークウェイ、５００、オラクル・インターナショナル・コーポレイション内
最終頁に続く

(54)【発明の名称】　グラフィック処理装置ワークロードの物理トポロジネットワーク局所性情報の公開

(57)【要約】
本明細書では、グラフィカル処理装置ベースのワークロードを実行するために、クラスタネットワークに含まれるホストマシンの局所性情報を利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について、ホストマシンの局所性情報がそこに記憶される。ホストマシンの局所性情報は、ホストマシンを含むラックを識別する。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うちの１つまたは複数のホストマシンが、ワークロードを実行するために使用可能であると識別される。１つまたは複数のホストマシンの各々について、ホストマシンの局所性情報が取得される。さらに、１つまたは複数のホストマシンのリンク情報が識別される。１つまたは複数のホストマシンの局所性情報とリンク情報は、要求に応じて提供される。

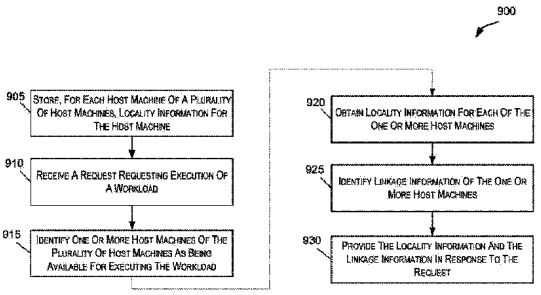

FIG. 9

【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、前記ホストマシンを含むラックを識別する前記ホストマシンの局所性情報を記憶することと、
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、
前記複数のホストマシンのうちの１つまたは複数のホストマシンを、前記ワークロードを実行するために使用可能であると識別することと、
前記１つまたは複数のホストマシンの各々について、前記ホストマシンの前記局所性情報を取得することと、
前記１つまたは複数のホストマシンのリンク情報を識別することと、
前記要求に応じて、前記１つまたは複数のホストマシンの前記局所性情報および前記リンク情報を提供することと、を含む、方法。
【請求項２】
前記ホストマシンの前記局所性情報は、前記ホストマシンを含む前記ラックのラック識別子を含む情報を含む、請求項１に記載の方法。
【請求項３】
前記ホストマシンの前記局所性情報は、前記ラックに関連付けられているトップオブラックスイッチの識別子を含む、請求項１に記載の方法。
【請求項４】
前記１つまたは複数のホストマシンの前記リンク情報は、前記１つまたは複数のホストマシンによって形成される論理トポロジに対応する、請求項１に記載の方法。
【請求項５】
前記１つまたは複数のホストマシンによって形成される前記論理トポロジはリングトポロジである、請求項４に記載の方法。
【請求項６】
前記ホストマシンの前記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７】
顧客が、前記１つまたは複数のホストマシンの前記局所性情報および前記リンク情報を受信することと、
前記顧客が、前記ワークロードを実行するための前記１つまたは複数のホストマシンのサブセットを選択することと、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の前記サブセット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８】
１つまたは複数の制約は、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７に記載の方法。
【請求項９】
前記１つまたは複数のホストマシンの前記サブセットを前記選択することは、前記１つまたは複数のホストマシンの前記リンク情報にさらに基づく、請求項８に記載の方法。
【請求項１０】
コンピュータ実行可能命令を記憶する１つまたは複数のコンピュータ可読非一時的な媒体であって、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、前記ホストマシンを含むラックを識別する前記ホストマシンの局所性情報を記憶することと、

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、
前記複数のホストマシンのうちの１つまたは複数のホストマシンを、前記ワークロードを実行するために使用可能であると識別することと、
前記１つまたは複数のホストマシンの各々について、前記ホストマシンの前記局所性情報を取得することと、
前記１つまたは複数のホストマシンのリンク情報を識別することと、
前記要求に応じて、前記１つまたは複数のホストマシンの前記局所性情報および前記リンク情報を提供することと、を行わせる、１つまたは複数のコンピュータ可読非一時的な媒体。

【請求項１１】
前記ホストマシンの前記局所性情報は、前記ホストマシンを含む前記ラックのラック識別子を含む情報を含む、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１つまたは複数のコンピュータ可読非一時的な媒体。

【請求項１２】
前記ホストマシンの前記局所性情報は、前記ラックに関連付けられているトップオブラックスイッチの識別子を含む、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１つまたは複数のコンピュータ可読非一時的な媒体。

【請求項１３】
前記１つまたは複数のホストマシンの前記リンク情報は、前記１つまたは複数のホストマシンによって形成される論理トポロジに対応する、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１つまたは複数のコンピュータ可読非一時的な媒体。

【請求項１４】
前記１つまたは複数のホストマシンによって形成される前記論理トポロジは、リングトポロジである、請求項１３に記載のコンピュータ実行可能命令を記憶する１つまたは複数のコンピュータ可読非一時的な媒体。

【請求項１５】
前記ホストマシンの前記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１０に記載のコンピュータ実行可能命令を記憶する１つまたは複数のコンピュータ可読非一時的な媒体。

【請求項１６】
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１つまたは複数のコンピュータ可読非一時的な媒体であって、１つまたは複数のプロセッサによって実行されると、
顧客が、前記１つまたは複数のホストマシンの前記局所性情報および前記リンク情報を受信することと、
前記ワークロードを実行するための前記１つまたは複数のホストマシンのサブセットを選択することと、を行わ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の前記サブセット上で前記ワークロードを実行すること、を行わせる命令をさらに含む、１つまたは複数のコンピュータ可読非一時的な媒体。

【請求項１７】
１つまたは複数の制約は、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１６に記載のコンピュータ実行可能命令を記憶する１つまたは複数のコンピュータ可読非一時的な媒体。

【請求項１８】
１つまたは複数のプロセッサと、
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１つまたは複数の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、前記ホストマシンを含むラックを識別する前記ホストマシンの局所性情報を記憶することと、

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、

前記複数のホストマシンのうちの１つまたは複数のホストマシンを、前記ワークロードを実行するために使用可能であると識別することと、

前記１つまたは複数のホストマシンの各々について、前記ホストマシンの前記局所性情報を取得することと、

前記１つまたは複数のホストマシンのリンク情報を識別することと、

前記要求に応じて、前記１つまたは複数のホストマシンの前記局所性情報および前記リンク情報を提供することと、を行わせる、コンピューティング装置。

【請求項１９】

前記ホストマシンの前記局所性情報は、前記ホストマシンを含む前記ラックのラック識別子を含む情報を含む、請求項１８に記載のコンピューティング装置。

【請求項２０】

前記ホストマシンの前記局所性情報は、前記ラックに関連付けられているトップオブラックスイッチの識別子を含む、請求項１８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

本出願は、２０２２年１月１２日に出願された米国仮出願第６３／２９８，６８５号および２０２２年１０月２７日に出願された米国非仮出願第１８／０５０，３９２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のリモートダイレクトメモリアクセス（ＲＤＭＡ）対応クラスタネットワークに含まれるホストマシンの局所性情報を、グラフィカル処理装置ベースのワークロードの実行に利用することに関する。

【背景技術】

【０００３】

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。高性能コンピュータプリケーションは、特定の成果または結果を達成するために、使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に他のホストが存在するかどうかに関係なくトラフィックを交換するという事実によって、問題はさらに悪化する。本明細書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、グラフィカル処理装置ベースのワークロードの実行において、クラウドインフラストラクチャのＲＤＭＡ対応クラスタネットワークに含まれるホストマシンの局所性情報を利用することに関する。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記憶媒体などを含む、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。
【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、ホストマシンを含むラックを識別するホストマシンの局所性情報を記憶することと、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うちの１つまたは複数のホストマシンをワークロードの実行に使用可能であると識別することと、１つまたは複数のホストマシンの各々について、ホストマシンの局所性情報を取得することと、１つまたは複数のホストマシンのリンク情報を識別することと、要求に応じて１つまたは複数のホストマシンの局所性情報およびリンク情報を提供することと、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コンピューティング装置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コンピューティング装置を提供する。
【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。
【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ＣＳＰＩ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。
【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（ＮＶＤ）に接続されるＣＳＰＩ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ＮＶＤ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ＣＳＰＩ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、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。
【図８Ａ】特定の実施形態による、局所性情報なしで構築される論理トポロジを示す図である。
【図８Ｂ】特定の実施形態による、局所性情報を用いて構築される別の論理トポロジを示す図である。
【図９】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１０】特定の実施形態による、顧客の要求を処理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】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ＣＳＰＩ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。

【００１７】

ＣＳＰＩ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。ＣＳＰＩ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。ＣＳＰＩ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）、トップオブラック（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３ＩＰネットワークまたは層２ＶＬＡＮ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－ＩＥＴＦ　ＲＦＣ７３４８）、Virtual Private Networks（ＶＰＮ）（例えば、MPLS Layer-3 Virtual Private Networks（ＲＦＣ４３６４））、ＶＭｗａｒｅのＮＳＸ、ＧＥＮＥＶＥ（Generic Network Virtualization Encapsulation）などがある。

【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（ＣＳＰＩ）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。ＣＳＰＩ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。ＣＳＰＩは、顧客のオ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ＣＳＰＩ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、ＣＳＰＩ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ＣＳＰＩ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、ＣＳＰＩ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、ＣＳＰＩ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ＣＳＰＩ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、ＣＳＰＩ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】

ＣＳＰＩ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、ＣＳＰＩ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、ＣＳＰＩ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、ＣＳＰＩ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。

【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。

【００２４】

クラウドインフラストラクチャまたはＣＳＰＩ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。ＣＳＰＩ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、ＣＳＰＩ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。ＣＳＰＩ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ＣＳＰＩ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。

【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、ＣＳＰＩ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、ケーブルパス、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。

【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。

【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。

【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ＡＤには３つの障害ドメインが含まれる。障害ドメインは、ＡＤ内の論理データセンターとして機能する。

【００３１】

顧客がＩａａＳサービスに加入すると、ＣＳＰＩ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。ＣＳＰＩ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（ＶＣＮ）と呼ばれる。顧客は、顧客に割り当てられるＣＳＰＩ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（ＶＣＮ）を設定できる。ＶＣＮ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ＶＣＮ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。ＶＣＮ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ＶＣＮまたは他のＶＣＮ（例えば、顧客の他のＶＣＮ、顧客に属さないＶＣＮ）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。

【００３２】

ＣＳＰは、ＣＳＰＩ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、ＣＳＰＩの顧客自身がサービスプロバイダのように行動し、ＣＳＰＩ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、ＩＰアドレス、ＤＮＳ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ＩＰ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。

【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、ＶＣＮが作成されると、ＶＣＮ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（ＣＩＤＲ）アドレス空間に関連付けられる。ＶＣＮ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。ＶＣＮ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、ＶＣＮ用に構成される仮想インターフェースであり、ＶＣＮとＶＣＮ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、ＶＣＮ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。

【００３５】

ＶＣＮ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、ＶＣＮ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。ＶＣＮ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。ＶＣＮ内の各サブネットは、そのＶＣＮ内の他のサブネットと重複せず、ＶＣＮ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ＶＣＮ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（ＶＮＩＣ）に関連付けられている。ＶＮＩＣは、物理的なネットワークインターフェースカード（ＮＩＣ）の論理表現である。一般に、ＶＮＩＣ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。ＶＮＩＣ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。ＶＮＩＣはスイッチ上の層２ポートに相当する。ＶＮＩＣは、コンピューティングインスタンスとＶＣＮ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ＶＮＩＣにより、コンピューティングインスタンスはＶＣＮ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、ＶＣＮ内の異なるサブネットのエンドポイント、またはＶＣＮ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ＶＮＩＣによって、コンピューティングインスタンスがＶＣＮ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、ＶＣＮ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ＶＮＩＣ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ＶＮＩＣが含まれ、各ＶＮＩＣ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ＶＮＩＣ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ＶＮＩＣ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ＶＮＩＣは、同じルートテーブル、セキュリティリスト、およびＤＨＣＰ任意選択肢を使用する。前述のように、ＶＣＮ内の各サブネットは、そのＶＣＮ内の他のサブネットと重複せず、ＶＣＮ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。ＶＣＮの特定のサブネット上のＶＮＩＣの場合、ＶＮＩＣ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ＶＮＩＣ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ＶＮＩＣ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ＶＮＩＣがあり、このプライマリＶＮＩＣは削除できない。セカンダリＶＮＩＣと呼ばれる追加のＶＮＩＣは、プライマリＶＮＩＣ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ＶＮＩＣ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ＶＮＩＣは、プライマリＶＮＩＣと同じＶＣＮ内のサブネット、または同じＶＣＮ内または別のＶＣＮ内の別のサブネットにあり得る。

【００３９】

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ＶＮＩＣ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ＶＮＩＣ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ル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。

【００４０】

上述のように、ＶＣＮ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、ＶＣＮ用に構成される仮想ルータ（ＶＲ）（ＶＣＮ　ＶＲまたは単にＶＲと呼ばれる）によって、ＶＣＮのサブネット間の通信が可能になる。ＶＣＮ内のサブネットの場合、ＶＲ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ＶＣＮ内の他のサブネット上のエンドポイントおよびＶＣＮ外の他のエンドポイントと通信できるようにする。ＶＣＮ　ＶＲは、ＶＣＮ内のＶＮＩＣとＶＣＮ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。ＶＣＮ　ＶＲは、層３／ＩＰ層の概念である。一実施形態では、ＶＣＮに対して１つのＶＣＮ　ＶＲがあり、ＶＣＮ　ＶＲ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、ＶＣＮ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、ＶＣＮ　ＶＲは、ＶＣＮ　ＶＲが接続されているＶＣＮ内のサブネットごとに異なるＩＰアドレスを持つ。ＶＲは、ＶＣＮ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ＶＣＮ　ＶＲのポート用に予約される。例えば、それぞれ１０．０／１６と１０．１／１６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ＶＣＮを考える。アドレス範囲が１０．０／１６であるＶＣＮ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ＶＣＮ　ＶＲ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ＶＣＮ　ＶＲ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１０．０／１６のサブネットの場合、ＩＰアドレス１０．０．０．１がそのサブネットのＶＣＮ　ＶＲのポート用に予約され得る。同じＶＣＮ内のアドレス範囲が１０．１／１６である第２のサブネットの場合、ＶＣＮ　ＶＲにはＩＰアドレスが１０．１．０．１であるその第２のサブネット用のポートが存在し得る。ＶＣＮ　ＶＲには、ＶＣＮ内のサブネットごとに異なるＩＰアドレスがある。

【００４１】

いくつかの他の実施形態では、ＶＣＮ内の各サブネットには、ＶＲ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ＶＲ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ＶＮＩＣ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ＶＲ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、ＶＲはそのサブネットの入口／出口ポイントになる。ＶＣＮ内のサブネットに関連付けられるＶＲは、ＶＣＮ内の他のサブネットに関連付けられている他のＶＲと通信できる。ＶＲは、ＶＣＮ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。

【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ルールのセットで構成される。一方、セキュリティリストには、セキュリティリストを使用するサブネット内のすべてのリソースに適用されるルールが含まれる。ＶＣＮには、デフォルトのセキュリティルールを含むデフォルトのセキュリティリストが提供され得る。ＶＣＮに構成されるＤＨＣＰ任意選択肢は、インスタンスの起動時にＶＣＮ内のインスタンスに自動的に提供される構成情報を提供する。

【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。

【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。
【００４６】
顧客は、ＣＳＰＩ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を作成することができる。顧客のＶＣＮ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、ＣＳＰＩによってホストされるエンドポイントとＣＳＰＩ外部のエンドポイントが含まれ得る。
【００４７】
ＣＳＰＩ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ＣＳＰＩによってホストされるオーバーレイまたは顧客ＶＣＮ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（ＶＣＮ）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ＣＳＰＩ１０１を含む。特定の実施形態では、ＣＳＰＩ１０１は加入顧客にＩａａＳサービスを提供する。ＣＳＰＩ１０１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「ＲｅｇｉｏｎＵＳ」１０２を図１に示す。顧客はリージョン１０２に顧客ＶＣＮ１０４を構成した。顧客はＶＣＮ１０４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ＶＣＮ１０４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ＣＩＤＲ　ＩＰ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ＩＰ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。ＶＣＮ仮想ルータ１０５は、ＶＣＮ１０４のサブネット間およびＶＣＮ外部の他のエンドポイントとの通信を可能にするＶＣＮの論理ゲートウェイを表す。ＶＣＮ　ＶＲ１０５は、ＶＣＮ１０４内のＶＮＩＣとＶＣＮ１０４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。ＶＣＮ　ＶＲ１０５は、ＶＣＮ１０４の各サブネットにポートを提供する。例えば、ＶＲ１０５は、サブネット１にＩＰアドレス１０．０．０．１のポートを提供し、サブネット２にＩＰ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、ＣＳＰＩ１０１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ＶＮＩＣ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Ｃ１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Ｃ２は、Ｃ２に関連付けられているＶＮＩＣ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ＶＮＩＣを介し

て、プライベートオーバーレイＩＰアドレスとＭＡＣ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Ｃ１のオーバーレイＩＰアドレスは１０．０．０．２、ＭＡＣアドレスはＭ１であるが、コンピューティングインスタンスＣ２の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。

【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。

【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、ＣＳＰＩ２００によってホストされるエンドポイントとＣＳＰＩ２００外部のエンドポイントが含まれ得る。ＣＳＰＩ１０１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。ＣＳＰＩ１０１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、ＣＳＰＩ１０１によってホストされていない（つまり、ＣＳＰＩ１０１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。

【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。

【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　ＶＲ１０５にルーティングするように構成されている。ＶＣＮ　ＶＲ１０５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。

【００５６】

ＶＣＮ１０４内のコンピューティングインスタンスからＶＣＮ１０４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　ＶＲ１０５、およびＶＣＮ１０４に関連付けられるゲートウェイによって容易になる。ＶＣＮ１０４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。

【００５７】

例えば、コンピューティングインスタンスＣ１は、ＶＣＮ１０４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ＶＣＮ１０４のＶＣＮ　ＶＲ１０５に転送し得る。次に、ＶＣＮ　ＶＲ１０５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　ＶＲ１０５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、ＣＳＰＩ１０１の別のリージョン内のＶＣＮ１０８、またはＣＳＰＩ１０１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、ＣＳＰＩ１０１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、ＣＳＰＩ１０１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、通信チャネル１２４が設定され、チャネルの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ＣＳＰＩ１０１にあり、顧客のＶＣＮ１０４に接続されている。通信チャネル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。ＣＳＰＩ１０１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、ＣＳＰＩ１０１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。ＩＧＷ１１２０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て、サービスは顧客のＶＣＮ内のプライベートサブネット内のプライベートＩＰ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（ＰＥ）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ＶＣＮ内のサービスを表す。各ＰＥは、顧客のＶＣＮ内で顧客が選択したサブネット内のＶＮＩＣ（１つまたは複数のプライベートＩＰを持つＰＥ－ＶＮＩＣと呼ばれる）として現れる。したがって、ＰＥは、ＶＮＩＣを使用してプライベート顧客ＶＣＮ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ＶＮＩＣ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、ＶＮＩＣに関連付けられているすべての機能がＰＥ　ＶＮＩＣで使用できるようになる。

【００６５】

サービスプロバイダは、ＰＥ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ＩＰアドレス（ＶＩＰ）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ＶＣＮに）複数存在し得る。

【００６６】

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、ＰＥ　ＶＮＩＣのプライベートＩＰアドレスまたはサービスＤＮＳ名を使用してサービスにアクセスできる。顧客ＶＣＮ内のコンピューティングインスタンスは、顧客ＶＣＮ内のＰＥのプライベートＩＰ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（ＰＡＧＷ）１３０は、サービスプロバイダＶＣＮ（例えば、サービスネットワーク１１０内のＶＣＮ）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。ＰＡＧＷ１３０を使用すると、プロバイダは内部ＩＰアドレスリソースを利用せずにＰＥ接続の数を拡張できる。プロバイダは、単一のＶＣＮに登録されている任意の数のサービスに対して１つのＰＡＧＷ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ＶＣＮ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、ＰＥ　ＶＮＩＣ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、ＰＡＧＷ１３０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（Ｃ２Ｓ接続）と呼ばれる。

【００６７】

ＰＥ概念は、トラフィックがFastConnect／ＩＰｓｅｃリンクと顧客ＶＣＮ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。ＬＰＧ１３２と顧客のＶＣＮ内のＰＥ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ＶＣＮに拡張することもできる。

【００６８】

顧客はサブネットレベルでＶＣＮ内のルーティングを制御できるため、ＶＣＮ１０４などの顧客のＶＣＮ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。ＶＣＮのルートテーブルは、特定のゲートウェイを介してＶＣＮ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ＶＣＮ１０４内のパブリックサブネットのルートテーブルは、ＩＧＷ１２０を介して非局所トラフィックを送信し得る。同じ顧客ＶＣＮ１０４内のプライベートサブネットのルートテーブルは、ＣＳＰサービス宛てのトラフィックをＳＧＷ１２６を介して送信し得る。残りの

トラフィックはすべてＮＡＴゲートウェイ１２８経由で送信され得る。ルートテーブルは、ＶＣＮから送信されるトラフィックのみを制御する。

【００６９】

ＶＣＮ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ＶＣＮ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、ＶＣＮ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、ＴＣＰ、ＩＣＭＰ）、特定のポート（例えば、ＳＳＨの場合は２２、Windows RDPの場合は３３８９）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。

【００７０】

顧客ＶＣＮからのアクセス（すなわち、ＶＣＮ１０４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ＮＡＴ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ＶＣＮ１０４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、ＣＳＰＩ１０１により、プライベートＩＰアドレスを持つ顧客ＶＣＮ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ＶＣＮ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。

【００７１】

さらに、ＣＳＰＩ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ＶＣＮ内の１つまたは複数のサービスにアクセスできる。ＣＳＰＩ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ＶＣＮ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ＣＳＰＩ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。

【００７２】

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ＣＳＰＩ２００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、ＣＳＰＩ２００は、クラウドサービスプロバイダ（ＣＳＰ）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ＣＳＰ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、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、ＣＳＰＩ２００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、ＣＳＰＩ２００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、ＣＳＰＩ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。ＣＳＰＩ２００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。

【００７３】

図２に示す実施例では、ＣＳＰＩ２００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。

【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。

【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いてもかまわない。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いてもかまわない。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ＶＣ

Ｎ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によってホストされるコンピューティングインスタンスに対応するＶＣＮに対応する１つまたは複数のＶＣＮ　ＶＲ２８３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（ＮＩＣ）が含まれ得る。ホストマシン上のＮＩＣ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ＮＶＤ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ＮＶＤ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２０２は、ホストマシン２０２のＮＩＣ２３２によって提供されるポート２３４とＮＶＤ２１０のポート２３６との間に延びるリンク２２０を使用してＮＶＤ２１０に接続されている。ホストマシン２０６は、ホストマシン２０６のＮＩＣ２４４によって提供されるポート２４６とＮＶＤ２１２のポート２４８との間に延びるリンク２２４を使用してＮＶＤ２１２に接続される。ホストマシン２０８は、ホストマシン２０８のＮＩＣ２５０によって提供されるポート２５２とＮＶＤ２１２のポート２５４との間に延びるリンク２２６を使用してＮＶＤ２１２に接続される。

【００８２】

ＮＶＤは、次いで、通信リンクを介してトップオブザラック（ＴＯＲ）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２１８（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ＮＶＤ間のリンク、およびＮＶＤとＴＯＲ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、ＮＶＤ２１０および２１２は、リンク２２８および２３０を使用して、それぞれＴＯＲスイッチ２１４および２１６に接続されている。特定の実施形態では、リンク２２０、２２４、２２６、２２８、および２３０はイーサネットリンクである。ＴＯＲに接続されるホストマシンとＮＶＤ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２１８は、ＴＯＲ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２１８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２１８はスイッチの多層Ｃｌｏｓネットワークであり、ＴＯＲスイッチ２１４および２１６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２１８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Ｃｌｏｓネットワーク構成が可能である。Ｃｌｏｓ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ＮＶＤ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ＮＶＤに接続される。例えば、図２では、ホストマシン２０２は、ホストマシン２０２のＮＩＣ２３２を介してＮＶＤ２１０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ＮＶＤに接続される。例えば、図２では、ホストマシン２０６および２０８は、それぞれＮＩＣ２４４および２５０を介して同じＮＶＤ２１２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ＮＶＤに接続される。図３は、ホストマシンが複数のＮＶＤに接続されているＣＳＰＩ３００内の例を示している。図３に示すように、ホストマシン３０２は、複数のポート３０６および３０８を含むネットワークインターフェースカード（ＮＩＣ）３０４を備えている。ホストマシン３００は、ポート３０

６およびリンク３２０を介して第１のＮＶＤ３１０に接続され、ポート３０８およびリンク３２２を介して第２のＮＶＤ３１２に接続される。ポート３０６および３０８はイーサネットポートであり得、ホストマシン３０２とＮＶＤ３１０および３１２間のリンク３２０および３２２はイーサネットリンクであり得る。ＮＶＤ３１０は、次いで、第１のＴＯＲスイッチ３１４に接続され、ＮＶＤ３１２は第２のＴＯＲスイッチ３１６に接続される。ＮＶＤ３１０と３１２およびＴＯＲスイッチ３１４と３１６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３１４および３１６は、多層物理ネットワーク３１８内の層０スイッチング装置を表す。

【００８６】

図３に示す配置は、物理スイッチネットワーク３１８からホストマシン３０２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３１４を経由してＮＶＤ３１０を経由してホストマシン３０２へ、第２のパスはＴＯＲスイッチ３１６を経由してＮＶＤ３１２を経由してホストマシン３０２へである。別々のパスにより、ホストマシン３０２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３０２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、ＮＶＤ２１０および２１２は、それぞれホストマシン２０２、およびホストマシン２０６および２０８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ＣＳＰＩ２００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。

【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。

【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。

【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て新しいパイプラインを構築したり、新しい段階を作成して既存のパイプラインに挿入することで新しい機能を追加したりできる。

【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。

【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インスタンス２７４に関連付けられているＶＮＩＣ２８４を実行する。ホストマシンは、異なる顧客に属する異なるＶＣＮに属するコンピューティングインスタンスをホストすることができ、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に対応するＶＮＩＣを実行する（つまり、ＶＮＩＣ関連機能を実行する）ことができる。

【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、ＮＶＤ２１０はパケット処理構成要素２８６を含み、ＮＶＤ２１２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ＶＣＮ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】

ＶＣＮ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ＶＮＩＣ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ＮＶＤに転送され得る。ＮＶＤ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ＶＮＩＣ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ＶＮＩＣ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ＶＮＩＣを実行するＮＶＤに転送され、ＮＶＤ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ＶＮＩＣは、同じＮＶＤ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ＮＶＤ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ＮＶＤ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。ＶＮＩＣは、ＮＶＤ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。

【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ＶＣＮ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ＮＶＤに通信される。ＮＶＤ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ＶＮＩＣの実行と、ＶＣＮに関連付けられているＶＲ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、ＮＶＤ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に対応する機能を実行または呼び出す（ＶＮＩＣを実行するとも呼ばれる）。ＶＮＩＣによって実行される機能には、パケット上のＶＬＡＮ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ＶＣＮ　ＶＲ機能が呼び出され、ＮＶＤによって実行される。次に、ＶＣＮ　ＶＲ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を実行するＮＶＤ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ＶＮＩＣ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ＶＮＩＣは、同じＮＶＤ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ＮＶＤ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ＮＶＤ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。

【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ＶＣＮ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ＮＶＤに通信される。ＮＶＤ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を実行する。パケットの宛先エンドポイントはＶＣＮ外にあるため、パケ

ットはそのＶＣＮのＶＣＮ　ＶＲによって処理される。ＮＶＤはＶＣＮ　ＶＲ機能を呼び出し、その結果、パケットがＶＣＮに関連付けられる適切なゲートウェイを実行するＮＶＤ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である場合、パケットはＶＣＮ　ＶＲによって、ＶＣＮ用に構成されるＤＲＧゲートウェイを実行するＮＶＤに転送され得る。ＶＣＮ　ＶＲ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を実行するＮＶＤと同じＮＶＤ上で実行されることも、別のＮＶＤ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ＮＩＣ、ホストマシン、またはその他のＮＶＤ実装であり得るＮＶＤ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（ＮＩＣ２３２を使用して）ホストマシン２０２からＮＶＤ２１０に通信され得る。ＮＶＤ２１０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ＶＮＩＣであるため、ＶＮＩＣ２７６が呼び出される。ＶＮＩＣ２７６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。

【０１０５】

ＶＣＮ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、ＣＳＰＩ２００によってホストされるエンドポイントとＣＳＰＩ２００外部のエンドポイントが含まれ得る。ＣＳＰＩ２００によってホストされるエンドポイントには、同じＶＣＮまたは他のＶＣＮ内のインスタンスが含まれ得る。これらのＶＣＮは、顧客のＶＣＮである場合もあれば、顧客に属さないＶＣＮである場合もある。ＣＳＰＩ２００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、ＣＳＰＩ２００によってホストされていないエンドポイント、またはＣＳＰＩ２００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。ＣＳＰＩ２００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ＣＳＰＩ２００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、ＣＳＰＩ２００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いてもよく、またはシステムの構成または配置が異なっていてもよい。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ＮＶＤ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ＶＭ１　４０６と顧客／テナント＃２に属するＶＭ２　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ＮＶＤ４１２に接続される物理ＮＩＣ４１０を含む。各コンピューティングインスタンスは、ＮＶＤ４１２によって実行されるＶＮＩＣに接続される。図４の実施形態

では、ＶＭ１　４０６はＶＮＩＣ－ＶＭ１　４２０に接続され、ＶＭ２　４０８はＶＮＩＣ－ＶＭ２　４２２に接続されている。

【０１０８】

図４に示すように、ＮＩＣ４１０は、論理ＮＩＣ　Ａ４１６と論理ＮＩＣ　Ｂ４１８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、ＶＭ１　４０６は論理ＮＩＣ　Ａ４１６に接続され、ＶＭ２　４０８は論理ＮＩＣ　Ｂ４１８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ＮＩＣ４１０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　Ａ４１６には特定のＶＬＡＮ　ＩＤが割り当てられ、テナント＃２の論理ＮＩＣ　Ｂ４１８には別のＶＬＡＮ　ＩＤが割り当てられる。ＶＭ１　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ＮＶＤ４１２に通信される。同様に、ＶＭ２　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ＮＶＤ４１２に通信される。したがって、ホストマシン４０２からＮＶＤ４１２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－ＶＭ１　４２０によって処理されるか、ＶＮＩＣ－ＶＭ２　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０ＴＯＲ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。

【０１１２】
特定の実施形態では、ＣＳＰＩ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。

【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－Ａ６１２とホスト１－Ｂ６１４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。ラック２　６２０には、つまり、ホスト２－Ａ６２２とホスト２－Ｂ６２４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約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。

【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－Ａ６１２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。

【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。

【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。

【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。

【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（スループット非認識型）ＥＣＭＰ負荷分散を実行する。

【０１１９】

標準のＥＣＭＰ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１００Ｇ着信リンクを介して着信し、各フローが同じ１００Ｇ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２００Ｇであるのに対し、発信帯域幅が１００Ｇ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ＧＰＵからＴＯＲ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ＧＰＵからＴＯＲ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ＴＯＲ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１００Ｇであると想定されている。ＴＯＲスイッチ６１６がＥＣＭＰ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、ＴＯＲの同じ送信リンク（例えば、ＴＯＲ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。

【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、ＴＣＰ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（ＲＤＭＡ）タイプのトラフィックの場合、状況は悪化する。ＲＤＭＡネットワークでは、さまざまな理由によりＴＣＰが使用されない（例えば、ＴＣＰの性能が高くない）。ＲＤＭＡネットワークでは、インフィニバンド上のＲＤＭＡや統合イーサネット上のＲＤＭＡ（ＲｏＣＥ）などのプロトコルが使用される。ＲｏＣＥ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。ドロップされるパケットの場合、ドロップされるパケットだけでなく、ドロップされる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。

【０１２１】

フロー衝突問題は、厳格な時間同期要件のため、ＧＰＵにとって重大な問題である。例えば、前述のように、ＧＰＵは、すべてのＧＰＵが時間同期した方法で相互に通信するマシン学習タスク（つまり、ワークロード）を実行し得る。マシン学習タスクやその他のタイプのタスクの場合、ＧＰＵ間の通信を可能にするために、ホストマシンの論理トポロジ（例えば、リングトポロジ、ツリートポロジなど）が構築される。ＧＰＵは、複数レベルまたは多次元であり得る論理トポロジを使用して相互に接続する。いくつかの実装では、ワークロードを実行するために、アプリケーションはＧＰＵを相互接続するための仮想（または論理）トポロジを構築する。通常、このようなアプリケーションは、ホストマシンの基礎となる物理トポロジを認識しないため、ランダム（つまり、任意の）方法で論理トポロジを構築しようとする。このようにランダムに構築される論理トポロジにより、ＧＰＵホストマシンは、局所ネットワーク近隣に他のＧＰＵホストが存在するかどうかに関係なく、トラフィックを交換することになる。そのため、トラフィックの輻輳が発生する可能性が高まり、ＧＰＵスループットが低下する。例えば、特定のマシン学習タスクを実行するためにホストマシンのペアが必要な場合を考慮する。この場合、ホストマシンのペアのランダム選択が実行され、ホストマシンの１つが第１の局所近隣（例えば、第１ラック

）にあり、もう１つのホストマシンが別の局所近隣（第１の局所近隣とは異なる）、例えば第２ラックにある場合、マシン学習タスクの実行によって、一定のレイテンシ（例えば、第１のホストマシンと第２のホストマシン間の通信で発生する遅延）が発生し、トラフィックの輻輳の可能性も高まる可能性がある。対照的に、マシン学習タスクを実行するために選択されるホストマシンのペアが同じ局所近隣に存在する場合（例えば、同じラックに存在する場合）、ホストマシン間の通信によるレイテンシが最小限に抑えられ、トラフィックの輻輳を回避する可能性が高くなることが理解される。

【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、論理トポロジを構築するプロセスでＧＰＵの局所性情報を活用し、それによって不必要なトラフィックの輻輳を回避する。さらに、本開示の実施形態は、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まで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。

【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。

【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、図８Ａおよび８Ｂを参照して以下に示すように

、ＧＰＵワークロードの性能に大きな影響を与える。

【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スマートＮＩＣ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、局所性情報はスマートＮＩＣ７３０を介してホスト１－Ａ７１０に（ＩＭＤＳ経由で）公開され、一方、局所性情報はスマートＮＩＣ７４０を介してホスト１－Ｂ７２０に（ＩＭＤＳ経由で）公開される。局所性情報には、ホストマシンのラック情報を示すメタデータが含まれ得ることが理解される。例えば、ラック情報は、特定のホストマシンが属するラックＩＤ（例えば、ラック１　６１０）を示すだけでなく、ラックに関連付けられているＴＯＲスイッチ（例えば、ＴＯＲ１スイッチ６１６）を示す場合もある。各ホストマシンはＩＭＤＳにクエリを実行して、ホストマシンに関連付けられているメタデータ情報を取得できることが理解される。以下に説明するように、公開される局所性情報は、より高いＧＰＵワークロードスループットを実現するための最適な論理トポロジを構築するために使用される。

【０１２７】

図８Ａを参照すると、特定の実施形態による、局所性情報なしで構築される論理トポロジが示される。図８Ａに示されている論理トポロジは、図６の４つのホストマシン、すなわちホストマシン１－Ａ６１２、ホストマシン１－Ｂ６１４、ホストマシン２－Ａ６２２、およびホストマシン２－Ｂ６２４に対応している。図６に示すように、ホストマシン１－Ａと１－Ｂはラック１　６１０に属し（つまり、同じＴＯＲスイッチＴＯＲ１　６１６の後ろに配置されている）、一方、ホストマシン２－Ａと２－Ｂはラック２　６２０に属し（つまり、同じＴＯＲスイッチＴＯＲ２　６２６の後ろに配置されている）ことに留意されたい。

【０１２８】

図８Ａに示す論理トポロジは、局所性情報なしで構築されるリングトポロジであり、すなわち、リングトポロジはランダム（任意）な方法で構築される。図８Ａに示すように、リングは、８０５というラベルの付いたリンクを介して、ホスト１－Ａがホスト２－Ｂに直接接続される（つまり、ホスト１－Ａの８つのＧＰＵが、ホスト２－Ｂの対応する８つのＧＰＵに接続される）ように構築されている。さらに、ホスト２－Ｂは８１０というラベルの付いたリンクを介してホスト１－Ｂに接続され、ホスト１－Ｂは８１５というラベルの付いたリンクを介してホスト２－Ａに接続され、ホスト２－Ａは８２０というラベルの付いたリンクを介してホスト１－Ａに接続される。

【０１２９】

図８Ａに示すように構築される論理トポロジでは、ＥＣＭＰトラフィック分散によりネットワークフローの衝突が発生しやすくなる。ホスト１－Ａ６１２とホスト１－Ｂ６１４は同じＴＯＲスイッチ、つまりＴＯＲ１　６１６の背後に配置されているが、ホスト１－Ｂから発信されホスト１－Ａ宛てのトラフィックは、次のルートを通過することに留意されたい。トラフィックは、最初にホスト１－Ｂからホスト２－Ａ（つまり、仮想リンク８１５）にルーティングされ、次にホスト２－Ａからホスト１－Ａ（つまり、仮想リンク８２０）にルーティングされる。したがって、発信元ホスト（ホスト１－Ｂ）と同じラック（ラック１）に配置されている宛先ホスト（ホスト１－Ａ）宛てのトラフィックは、ＴＯＲスイッチを介してラック外のホストマシン（ホスト２－Ａ）に不必要にルーティングされ、さらにラック内の宛先ホストマシンにリダイレクトされる。トラフィックをホストマシン２－Ｂからホストマシン２－Ａにルーティングする場合にも、同様の状況が発生する。個々のホストマシンは自身の局所性情報を認識していないため、リングトポロジが任意に構築されることに留意されたい。図８Ａの任意に構築される論理トポロジにより、フロー衝突の可能性が高まり、ＧＰＵワークロードのスループットが低下する（例えば、レイテンシの増加、ジッター損失など）。

【０１３０】
図８Ｂは、特定の実施形態による、局所性情報を使用して構築される別の論理トポロジを示す。図８Ｂに示す論理トポロジは、図６の４つのホストマシン、すなわちホストマシン１－Ａ６１２、ホストマシン１－Ｂ６１４、ホストマシン２－Ａ６２２、およびホストマシン２－Ｂ６２４に対応し、局所性情報に基づいて構築される。ホストマシン１－Ａと１－Ｂはラック１　６１０に属し（つまり、同じＴＯＲスイッチＴＯＲ１　６１６の背後に配置されている）、ホストマシン２－Ａと２－Ｂはラック２　６２０に属し（つまり、同じＴＯＲスイッチＴＯＲ２　６２６の背後に配置されている）ことに留意されたい。
【０１３１】
図８Ｂに示す論理トポロジは、局所性情報を用いて構築されるリングトポロジであり、すなわち、リングトポロジは、例えばＩＭＤＳから取得される局所性情報に基づいて構築される。いくつかの実施形態では、論理トポロジは、ラックに含まれるホストマシンの１つであり得る構成ホストマシンによって構築され得る。図８Ｂに示すように、リングは、８６５というラベルの付いたリンクを介して、ホスト１－Ｂがホスト１－Ａに直接接続される（つまり、ホスト１－Ｂの８つのＧＰＵが、ホスト１－Ａの対応する８つのＧＰＵに接続される）ように構築されている。さらに、ホスト１－Ａは８５０というラベルの付いたリンクを介してホスト２－Ｂに接続され、ホスト２－Ｂは８５５というラベルの付いたリンクを介してホスト２－Ａに接続され、ホスト２－Ａは８６０というラベルの付いたリンクを介してホスト１－Ｂに接続される。
【０１３２】
図８Ｂに示す論理トポロジは、同じラック内にある（つまり、同じＴＯＲスイッチの背後に配置されている）２台のホストマシン間のトラフィックがラックの外部に不必要にルーティングされることを回避する。例えば、ホスト１－Ｂからホスト１－Ａにトラフィックをルーティングする場合を考慮すると、トラフィックはホストマシン１－Ｂからホストマシン１－Ａに直接送信できる（仮想リンク８６５経由）。これは、図８Ａに関して実行されるルーティングとは対照的である。図８Ａでは、ホスト１－Ｂで発信され、ホスト１－Ａ宛てのトラフィックが最初にホスト２－Ａに向けられ（つまり、ＴＯＲスイッチ経由でラックの外部にルーティングされ）、さらに宛先ホストマシン、つまりホスト１－Ａに送信される。このように、局所性情報に基づいて論理トポロジを構築すると、フロー衝突の可能性が低減され、ＧＰＵワークロードのスループットが向上する。
【０１３３】
図９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９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９に示され、以下で説明される方法は、例示を目的としており、限定するものではない。図９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。
【０１３４】
プロセスはステップ９０５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局所性情報が記憶される。ホストマシンの局所性情報には、ホストマシンを含むラック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局所性情報を記憶することができる。局所性情報は、ホストマシンを含むラックの識別子（ラックＩＤ）、ラックに関連付けられているＴＯＲスイッチの識別子などを示す情報に対応し得ることに留意されたい。
【０１３５】

ステップ９１０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。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９１５に移動し、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がワークロードの実行に使用可能であると識別される。使用可能なホストマシンの識別は、いくつかの方法で実行できる。例えば、一実施形態では、制御プレーンは、ホストマシンによって処理される現在の負荷（つまり、ワークロード）を維持することができる。制御プレーンは、各ホストマシンの容量と、ホストマシンによって処理される現在のワークロードの量に基づいて、顧客のワークロードを実行するために使用可能な１つまたは複数のホストマシンを選択できる。さらに、別の実施形態では、制御プレーンは、顧客ごとに一定数のホストマシンを事前に割り当てることができる。さらに、この一定数のホストマシンから、使用可能なホストマシンを決定することもできる。

【０１３６】

次に、プロセスはステップ９２０に進み、ステップ９１５で識別される１つまたは複数のホストマシンの各々について、局所性情報が取得される。（各ホストマシンの）局所性情報は、例えば、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。ステップ９２５では、プロセスは１つまたは複数のホストマシンのリンク情報を識別する。１つまたは複数のホストマシンのリンク情報は、１つまたは複数のホストマシンによって形成される論理トポロジ（例えば、図８Ｂに示す論理トポロジ）に対応することが理解される。その後、プロセスはステップ９３０に移動し、要求に応じて１つまたは複数のホストマシンの局所性情報とリンク情報が提供される。いくつかの実施形態によれば、顧客は、局所性情報およびリンク情報を取得すると、ワークロードを実行するために１つまたは複数のホストマシンのサブセットを選択できる。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実行され得ることに留意されたい。制約に関する詳細については、図１０および１１を参照して後で説明する。ホストマシンのサブセットを選択すると、選択したホストマシン上でワークロードが実行され得る。

【０１３７】

図１０は、特定の実施形態による、顧客の要求を処理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１０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１０に示され、以下で説明される方法は、例示を目的としており、限定するものではない。図１０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、ステップは一部の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。

【０１３８】

プロセスはステップ１００５で開始され、複数のホストマシン（つまり、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局所性情報が記憶される。ホストマシンの局所性情報には、ホストマシンを含むラック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局所性情報を記憶することができる。局所性情報は、ホストマシンを含むラックの識別子（ラックＩＤ）、ラックに関連付けられているＴＯＲスイッチの識別子などを示す情報に対応し得ることに留意されたい。

【０１３９】

ステップ１０１０では、制御プレーンは顧客からの要求を受信し、その要求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の情報を要求する。ステップ１０１５では、制御プレーンは、顧客に割り当てられる１つまたは複数のホストマシンを識別する。いく

つかの実施形態では、制御プレーンは、異なる顧客に割り当てられるホストマシンのマッピングを維持することができる。このように、制御プレーンはマッピングを利用して、顧客に割り当てられる１つまたは複数のホストマシンを識別できる。

【０１４０】

次に、プロセスはステップ１０２０に進み、ステップ１０１５で識別される１つまたは複数のホストマシンの局所性情報が取得される。（各ホストマシンの）局所性情報は、例えば、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。ステップ１０２５では、プロセスは１つまたは複数のホストマシンのリンク情報を識別する。１つまたは複数のホストマシンのリンク情報は、１つまたは複数のホストマシンによって形成される論理トポロジ（例えば、図８Ｂに示す論理トポロジ）に対応することが理解される。その後、プロセスはステップ１０３０に移動し、ここで、要求に応じて、１つまたは複数のホストマシンの局所性情報とリンク情報が（例えば、顧客に）提供される。前述のように、顧客はステップ１０３０で提供される情報を利用して、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従ってワークロードを実行するためのホストマシンを選択できる。

【０１４１】

図１１は、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例示的なフローチャートを示している。図１１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１１に示され、以下で説明される方法は、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。

【０１４２】

プロセスはステップ１１０５で開始され、複数のホストマシン（つまり、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局所性情報が記憶される。ホストマシンの局所性情報には、ホストマシンを含むラック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局所性情報を記憶することができる。局所性情報は、ホストマシンを含むラックの識別子（ラックＩＤ）、ラックに関連付けられているＴＯＲスイッチの識別子などを示す情報に対応し得ることに留意されたい。

【０１４３】

ステップ１１１０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。

【０１４４】

ステップ１１１５では、ステップ１１１０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１１２０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１つまたは複数のホストマシン（複数のホストマシンから）

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１１１５で識別される１つまたは複数の制約に従って決定でき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１１２０のプロセスは、１つまたは複数の制約および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など）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。例えば、図６および８Ｂを参照し、制約が最小のレイテンシを達成することである場合を考慮すると、この場合、マシン学習タスクを実行するために２台のホストマシンが必要な場合、１つの実装では、ホスト１－Ａとホスト１－Ｂ（両方とも図６に示すようにラック１、６１０に存在する）がマシン学習タスクを実行するために割り当てられ得る。このようなタスクの実行では、トラフィックがラック１内に局所に封じ込められ、それにより低レイテンシが実現されることに留意されたい。対照的に、別の例では、制約が高可用性の達成である場合、この場合は、ホスト１－Ａ（ラック１内）とホスト２－Ｂ（ラック２内）がマシン学習タスクの実行用に割り当てられ得る。

【０１４５】

次に、プロセスはステップ１１２５に進み、ステップ１１２０で割り当てられる１つまたは複数のホストマシンの各々について、局所性情報が取得される。（各ホストマシンの）局所性情報は、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。このプロセスでは、さらに、１つまたは複数のホストマシンのリンク情報を識別する。１つまたは複数のホストマシンのリンク情報は、１つまたは複数のホストマシンによって形成される論理トポロジ（例えば、図８Ｂに示す論理トポロジ）に対応することが理解される。１つの実装では、アプリケーションはホストマシンの局所性情報に基づいてホストマシンの論理トポロジを構築でき、各ホストマシンはインスタンスメタデータサービスから局所性情報を取得することに留意されたい。取得される局所性情報および顧客に割り当てられる１つまたは複数のホストマシンのリンク情報は、ステップ１１２５で顧客に提供される。その後、プロセスはステップ１１３０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。

【０１４６】

したがって、本開示の技術は、顧客がワークロードを例えば、ラック内の近隣のホストマシンの近くに「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プロビジョニングする。さらに、顧客は局所性情報を使用して、より高い反親和性を実現し、そのためより高い回復力を獲得するような方法でワークロードを配置できる。一実施形態によれば、ＧＰＵ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プロセスで局所性情報（例えば、ラック情報）を利用する機能により、ホストマシンをランダムな方法でユーザ要求に割り当てる、つまりホストマシンの局所性情報を考慮せずまたは顧客に提供しない単純な技術の性能と比較して、ＧＰＵワークロードの性能（例えば、スループット）が６倍向上する。

【０１４７】

図９～１１に示されるフローチャートに対する修正は、本開示の範囲内であることが理解される。例えば、１つの実装では、ワークロードを実行するように構成される１つまたは複数のホストマシンの局所性情報が顧客に提供され得る。顧客は、１つまたは複数のホストマシンの局所性情報を利用して、それぞれのユースケースのリンクトポロジを作成できる（例えば、１つまたは複数のホストマシンに関連付けられる親和性、反親和性、またはその他の制約を取得できる）。

【０１４８】

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。

【０１４９】

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。

【０１５０】

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。

【０１５１】

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。

【０１５２】

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り得る。

【０１５３】

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。

【０１５４】

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。

【０１５５】

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。

【０１５６】

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）、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などのさまざまなクラウド/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。

【０１５７】

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ビステナント１２１９に含めることができる。

【０１５８】

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。

【０１５９】

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。

【０１６０】

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。

【０１６１】

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。

【０１６２】

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。

【０１６３】

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。

【０１６４】

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。

【０１６５】

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１２２６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。

【０１６６】
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。
【０１６７】
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。
【０１６８】
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。
【０１６９】
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。
【０１７０】
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。

【０１７１】

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。

【０１７２】

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。

【０１７３】

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ＶＣＮ１３１８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。

【０１７４】

他の例では、ＩａａＳプロバイダの顧客は、顧客テナント１３２１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ＶＣＮ１３１６には、アプリサブネット１３２６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はデータプレーンＶＣＮ１３１８内に存在できるが、データプレーンミラーアプリ層１３４０はデータプレーンＶＣＮ１３１８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１３４０は顧客テナント１３２１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１３４０はデータプレーンＶＣＮ１３１８に存在しないか、ＩａａＳ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１３４０は、データプレーンＶＣＮ１３１８への呼び出しを行うように構成できるが、制御プレーンＶＣＮ１３１６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ＶＣＮ１３１６にプロビジョニングされるリソースをデータプレーンＶＣＮ１３１８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１３４０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。

【０１７５】

いくつかの実施形態では、ＩａａＳプロバイダの顧客は、データプレーンＶＣＮ１３１８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ＶＣＮ１３１８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ＶＣＮ１３１８からパブリックインターネット１３５４へのアクセスを制限できる。ＩａａＳプロバイダは、フィルタを適用したり、データプレーンＶＣＮ１３１８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１３２１に含まれるデータプレーンＶＣＮ１３１８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ＶＣＮ１３１８を他の顧客およびパブリックインターネット１３５４から分離するのに役立ち得る。

【０１７６】

いくつかの実施形態では、クラウドサービス１３５６は、サービスゲートウェイ１３３６によって呼び出され、パブリックインターネット１３５４、制御プレーンＶＣＮ１３１６、またはデータプレーンＶＣＮ１３１８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１３５６と制御プレーンＶＣＮ１３１６またはデータプレーンＶＣＮ１３１８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１３５６は、ＩａａＳ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１３５６は、サービスゲートウェイ１３３６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１３５４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１３５６は他のクラウドサービス１３５６から分離され得、制御プレーンＶＣＮ１３１６は、制御プレーンＶＣＮ１３１６と同じリージョンにない場合があるクラウドサービス１３５６から分離され得る。例えば、制御プレーンＶＣＮ１３１６は「リージョン１」に配置され得、クラウドサービス「展開１２」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ＶＣＮ１３１６に含まれるサービスゲートウェイ１３３６によって展開１２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１２に送信され得る。この例では、制御プレーンＶＣＮ１３１６、つまりリージョン１の展開１２は、リージョン２の展開１２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。

【０１７７】

図１４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１４００である。サービスオペレータ１４０２（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）１４０６（例えば、図１２のＶＣＮ１２０６）およびセキュアホストサブネット１４０８（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　ＶＣＮ１４

１２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。

【０１７８】

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。

【０１７９】

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。

【０１８０】

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）１　４６６（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（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１４

６２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。

【０１８１】

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。

【０１８２】

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。

【０１８３】

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。

【０１８４】

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ＶＭ１　４６６（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。

【０１８５】

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。

【０１８６】

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、ＳＳＨ　ＶＣＮ１５１２に含まれるＬＰＧ１５１０を介してＳＳＨ　ＶＣＮ１５１２（例えば、図１２のＳＳＨ　ＶＣＮ１２１２）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。ＳＳＨ　ＶＣＮ１５１２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、ＳＳＨ　ＶＣＮ１５１２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。

【０１８７】

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、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。

【０１８８】

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Ｄ

ＭＺ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。

【０１８９】

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）１　５６６（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ＶＭ１　５６６（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。

【０１９０】

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。

【０１９１】

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ＶＭ１　５６６（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ＶＣＮ１５１６から分離でき、データプレーンＶＣＮ１５１８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。

【０１９２】
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ＶＮＩＣ１５７２（１）－（Ｎ）に送信でき、セカンダリＶＮＩＣ１５７２（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ＶＣＮ１５１６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。

【０１９３】
図に示されるＩａａＳアーキテクチャ１２００、１３００、１４００、１５００は、図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。

【０１９４】
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、弾力的にスケーラブル、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ァ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。

【０１９５】
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。

【０１９６】
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ＩＥＥＥＰ１３８６．１標準に従って製造されるメザニンバスとして実装できる。

【０１９７】
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。

【０１９８】

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。

【０１９９】

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ー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。

【０２００】

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。

【０２０１】

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。

【０２０２】

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。

【０２０３】

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ブラウザー、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。

【０２０４】

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。

【０２０５】

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。

【０２０６】

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。

【０２０７】

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。

【０２０８】

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（ＩＥＥＥ８０２．１１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。

【０２０９】

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。

【０２１０】
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。
【０２１１】
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。
【０２１２】
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。
【０２１３】
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。
【０２１４】
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。
【０２１５】
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て使用することができる。
【０２１６】
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。

【０２１７】

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。

【０２１８】

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。

【０２１９】

「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。

【０２２０】

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、本明細書に添付される請求項に記載される主題のすべての修正および同等物が含まれる。さらに、本明細書に別途記載がない限り、上記要素のあらゆる可能なバリエーションの任意の組み合わせが本開示に包含される。

【０２２１】

本明細書に引用されている出版物、特許出願、特許を含むすべての参考文献は、各参考

文献が個別に具体的に参照により組み込まれることが示され、本明細書にその全体が記載された場合と同程度に、参照により本明細書に組み込まれる。前述の明細書では、本開示の態様は、その特定の実施形態を参照して説明されているが、当業者であれば、本開示はそれに限定されないことは理解されるであろう。上述の開示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て使用することができる。さらに、実施形態は、本明細書のより広い趣旨および範囲から逸脱することなく、本明細書で説明されているもの以外の任意の数の環境および用途で利用することができる。したがって、明細書および図面は、制限的なものではなく、例示的なものとしてみなされる必要がある。

【図面】

【図１】

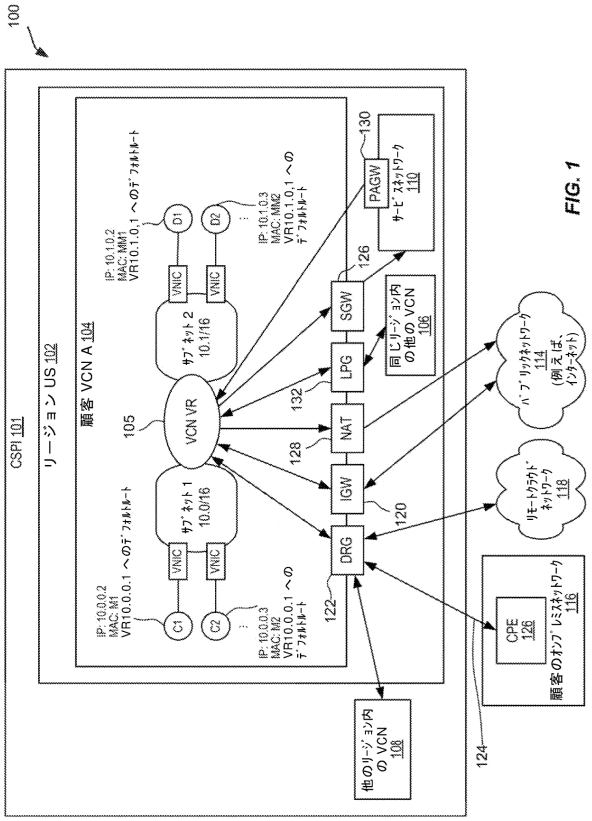

FIG. 1

【図２】

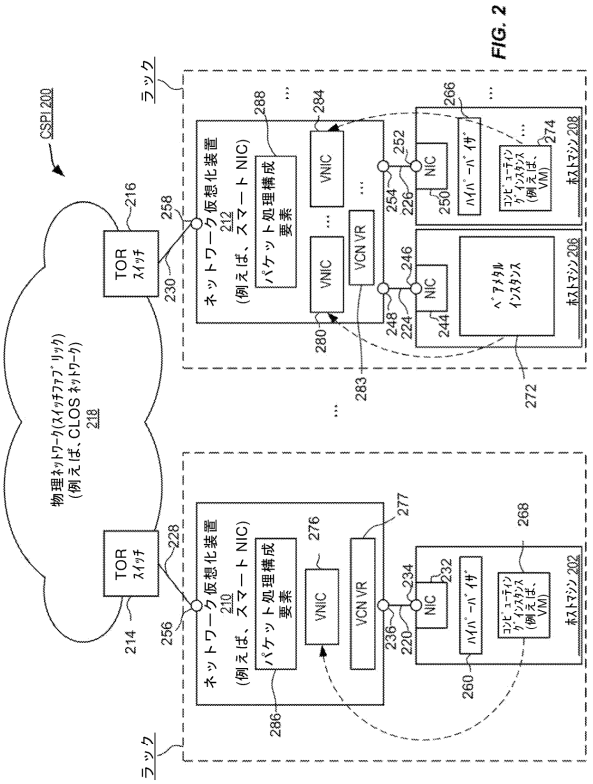

FIG. 2

【図３】

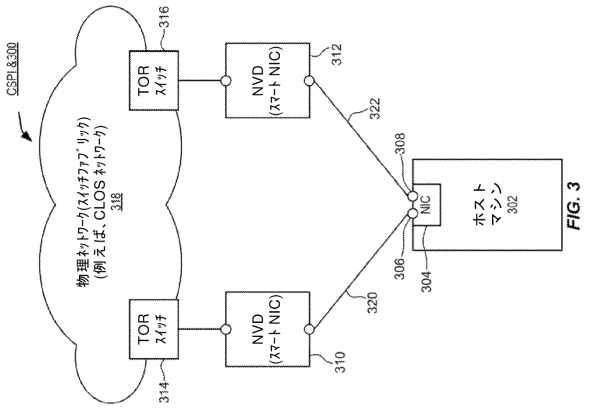

FIG. 3

【図４】

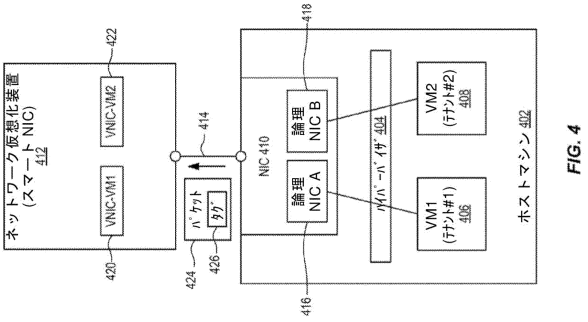

FIG. 4

【図５】

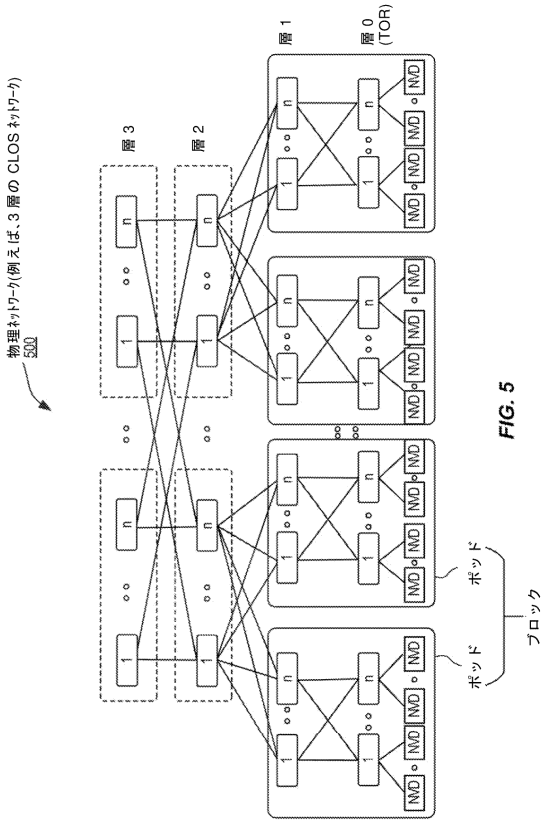

FIG. 5

【図６】

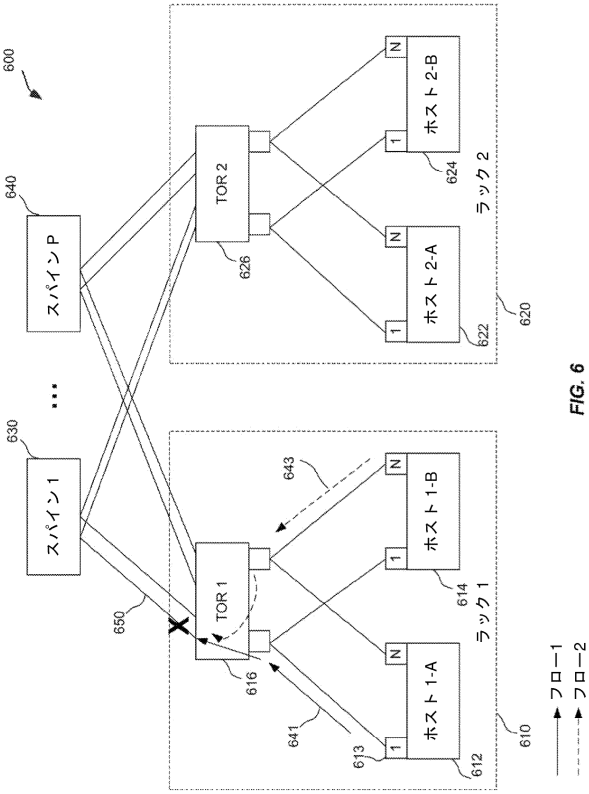

FIG. 6

【図７】

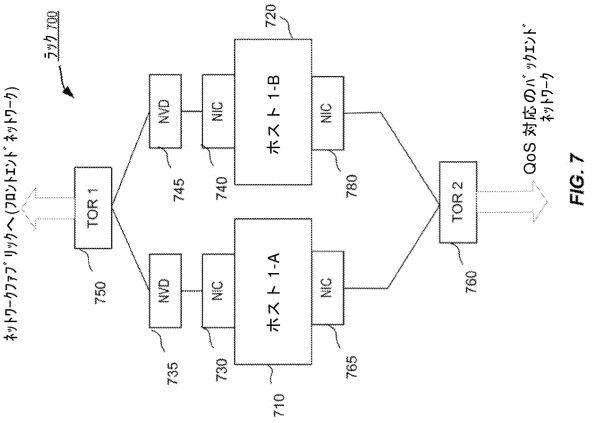

FIG. 7

【図８Ａ】

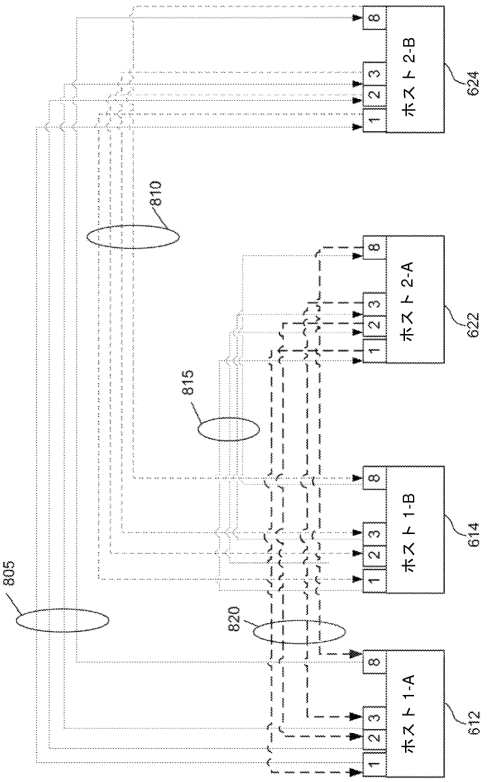

FIG. 8A

【図８Ｂ】

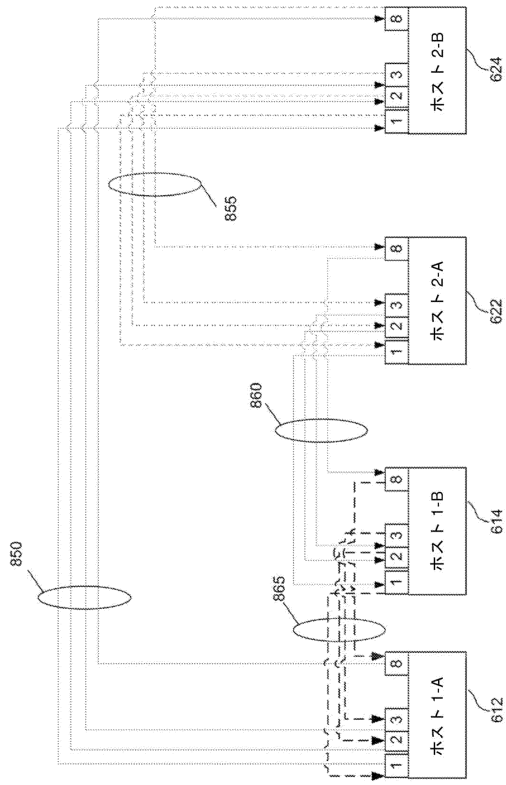

FIG. 8B

【図９】

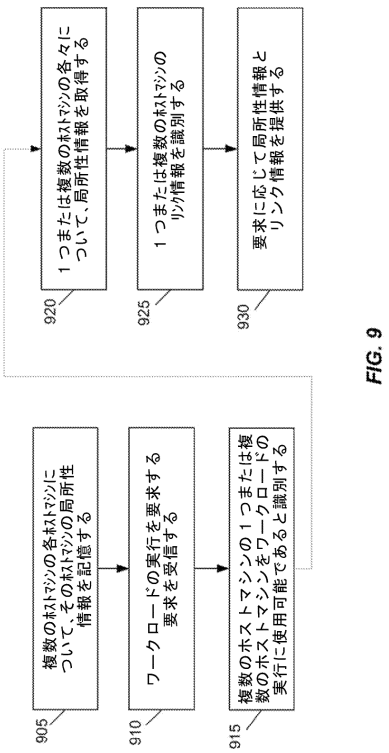

FIG. 9

【図１０】

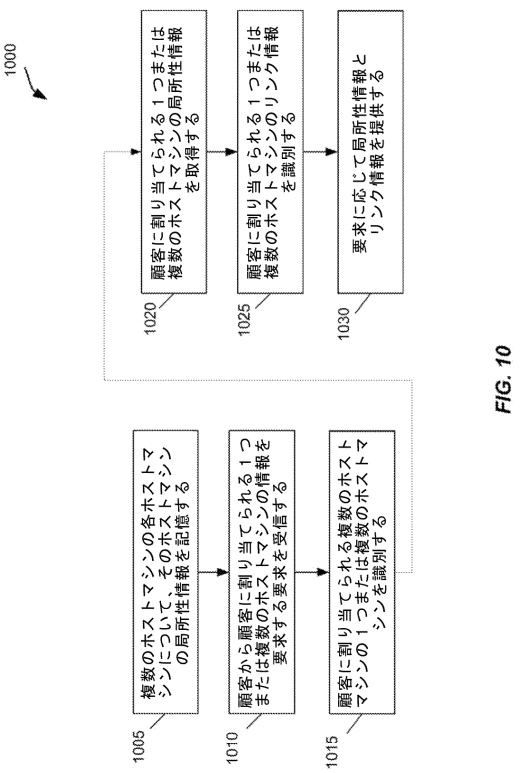

FIG. 10

【図１１】

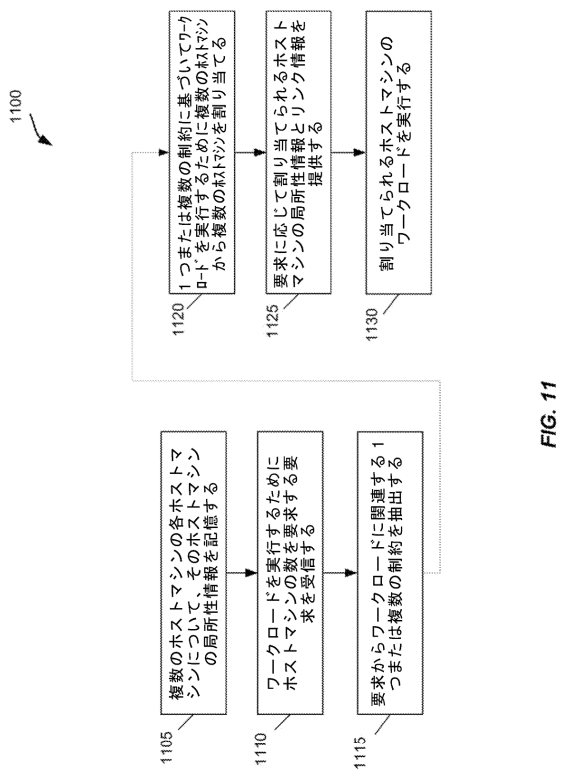

FIG. 11

【図１２】

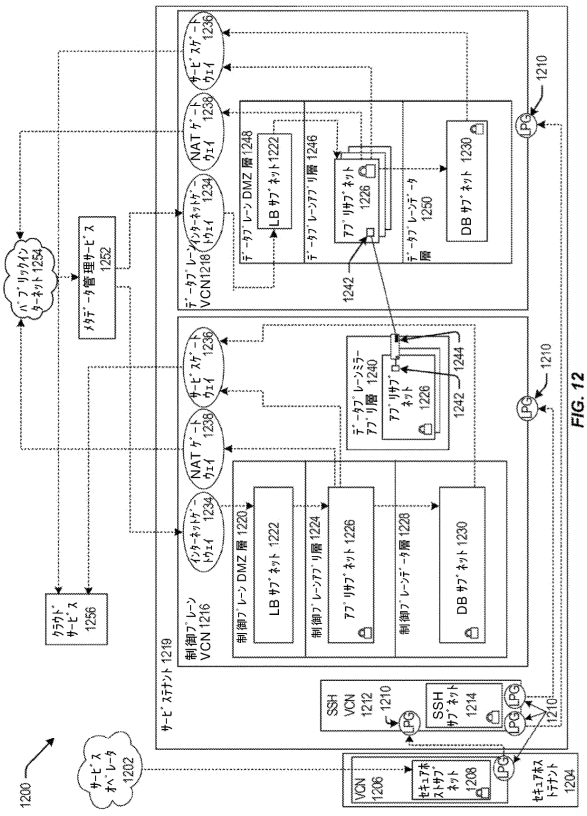

FIG. 12

【図１３】

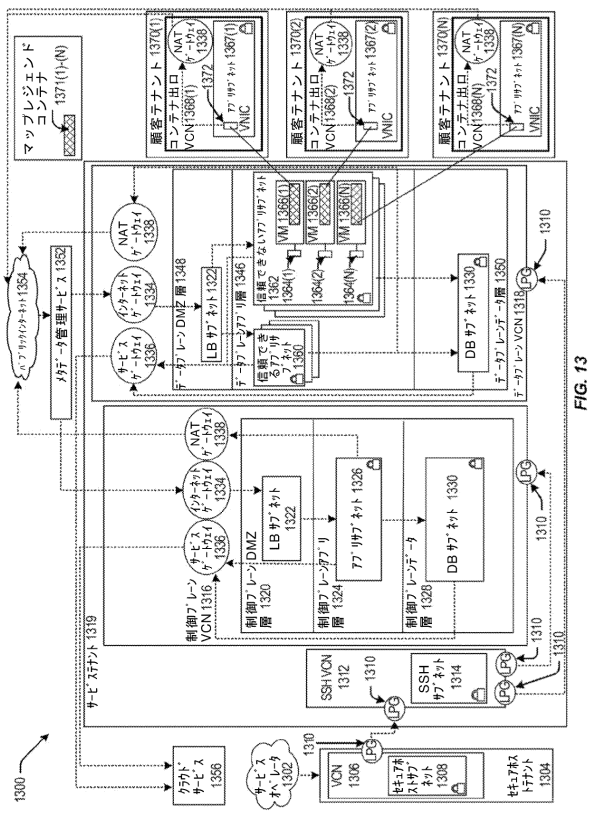

FIG. 13

【図１４】

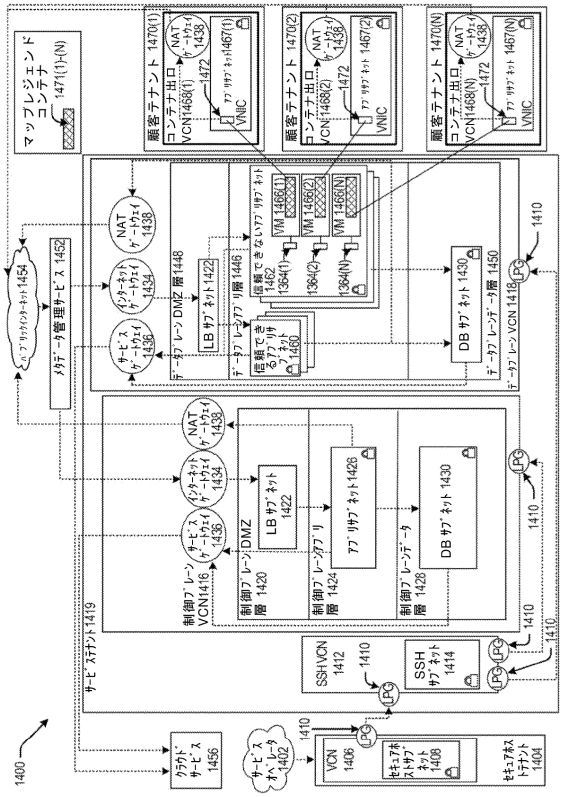

FIG. 14

【図１５】

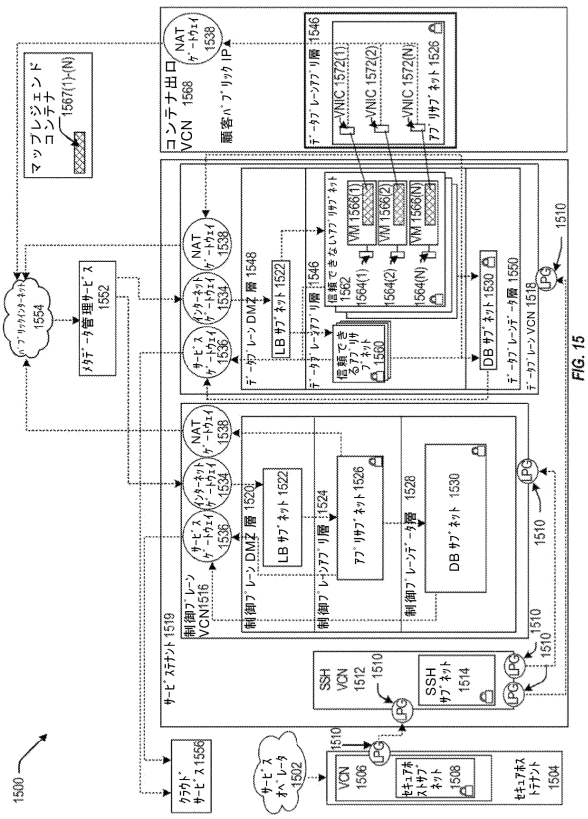

FIG. 15

【図１６】

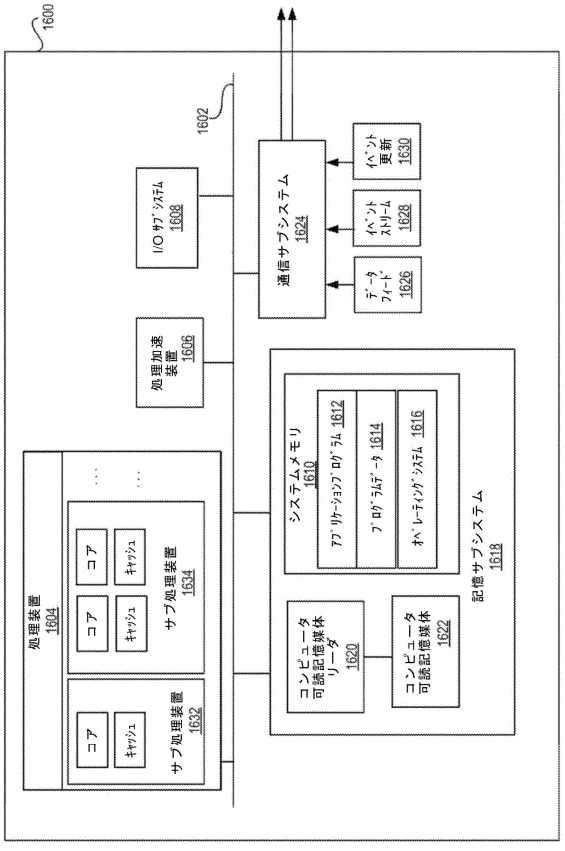

FIG. 16

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2022/082073

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	WO 2021/108358 A1 (AMAZON TECH INC [US]) 3 June 2021 (2021-06-03) paragraphs [0047] - [0049], [0098] - [0104], [0112], [0113]	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	Date of mailing of the international search report
30 March 2023	11/04/2023

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Buzgan, C

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2022/082073

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 9451393 B1	20-09-2016	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

Form PCT/ISA/210 (patent family annex) (April 2005)

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,DE,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,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,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,LA,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,VN,WS,ZA,ZM,ZW

（特許庁注：以下のものは登録商標）

１．ＷＩＮＤＯＷＳ ＰＨＯＮＥ

２．ｉＯＳ

(72)発明者 ベッカー，デイビッド
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・パークウェイ、５００、オラクル・インターナショナル・コーポレイション内

(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・パークウェイ、５００、オラクル・インターナショナル・コーポレイション内