(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7417812号
(P7417812)

(45)発行日　令和6年1月19日(2024.1.19)

(24)登録日　令和6年1月11日(2024.1.11)

(51)国際特許分類
H04L 45/76 (2022.01)

FI
H04L 45/76

請求項の数　36　（全104頁）

(21)出願番号　特願2022-503990(P2022-503990)
(86)(22)出願日　令和2年7月25日(2020.7.25)
(65)公表番号　特表2022-545608(P2022-545608A)
(43)公表日　令和4年10月28日(2022.10.28)
(86)国際出願番号　PCT/US2020/043635
(87)国際公開番号　WO2021/040934
(87)国際公開日　令和3年3月4日(2021.3.4)
審査請求日　令和4年3月18日(2022.3.18)
(31)優先権主張番号　16/662,379
(32)優先日　令和1年10月24日(2019.10.24)
(33)優先権主張国・地域又は機関
米国(US)
(31)優先権主張番号　62/892,473
(32)優先日　令和1年8月27日(2019.8.27)
最終頁に続く

(73)特許権者　514097912
ヴィーエムウェア エルエルシー
ＶＭｗａｒｅ ＬＬＣ
アメリカ合衆国 カリフォルニア州 ９４３０４，パロアルト，ヒルビュー アベニュー ３４０１
３４０１ Ｈｉｌｌｖｉｅｗ Ａｖｅ．，Ｐａｌｏ Ａｌｔｏ，ＣＡ ９４３０３，Ｕ．Ｓ．Ａ
(74)代理人　100087642
弁理士　古谷 聡
(74)代理人　100082946
弁理士　大西 昭広
(74)代理人　100195693
弁理士　細井 玲
最終頁に続く

(54)【発明の名称】　仮想ネットワークを実施するためのリコメンデーションの提供

(57)【特許請求の範囲】

【請求項１】

２つ以上のパブリッククラウドデータセンタのセットにわたって定義されるエンティティの仮想ネットワークに関するリコメンデーションを提供する方法であって、前記方法は、

エンティティから、前記エンティティのための仮想ネットワークが定義されるべきパブリッククラウドデータセンタの前記セットの識別を受信することと、

前記エンティティのための前記仮想ネットワークを実施するためのパブリッククラウドデータセンタの前記セットで動作する転送要素のセットをデプロイし、設定することと、

前記仮想ネットワークの性能メトリックを改善するために、少なくとも１つの転送要素を識別して、転送要素の前記セットに追加することであって、前記識別される転送要素は、前記エンティティによって識別されたパブリッククラウドデータセンタの前記セットにない特定のパブリッククラウド内にある、ことと、

前記仮想ネットワークの前記性能メトリックを改善するために、前記特定のパブリッククラウドをパブリッククラウドデータセンタの前記セットに追加するリコメンデーションを提供することと、を含む方法。

【請求項２】

請求項１に記載の方法であって、更に、

前記特定のパブリッククラウドをパブリッククラウドデータセンタの前記セットに追加する要求を、前記エンティティから受信することと、

前記特定のパブリッククラウドに少なくとも１つの特定の転送要素をデプロイすること

と、
前記特定の転送要素と転送要素の前記セットが前記仮想ネットワークを実施するように、前記特定の転送要素を設定し且つ転送要素の前記セットのうちの少なくとも１つの転送要素を再設定することと、を含む方法。
【請求項３】
請求項１に記載の方法であって、転送要素の前記セットを設定することは、
前記仮想ネットワークを介して前記エンティティのマシンを接続するパスを識別するように、パス探索プロセスを実行することと、
前記識別されたパスに基づいて転送記録を生成することと、
前記転送要素の前記セットを設定するために前記転送記録を使用することと、を含む方法。
【請求項４】
請求項３に記載の方法であって、前記エンティティの前記マシンは、前記エンティティのモバイルマシン、オフィスマシン、またはプライベートデータセンタマシンを含むパブリッククラウドデータセンタの前記セットの外部のマシンを含む、方法。
【請求項５】
請求項３に記載の方法であって、前記転送要素はルータを含み、前記転送記録はルーティング記録である、方法。
【請求項６】
請求項５に記載の方法であって、前記ルーティング記録は次ホップ記録である、方法。
【請求項７】
請求項５に記載の方法であって、前記ルータは、パブリッククラウドデータセンタの前記セットのホストコンピュータ上で動作するマシンとして実施される、方法。
【請求項８】
請求項１に記載の方法であって、パブリッククラウドデータセンタの前記セットは、パブリッククラウドデータセンタの第１のセットであり、前記方法は、更に、
パブリッククラウドデータセンタの前記第１のセットとパブリッククラウドデータセンタの前記第１のセットの外側にあるパブリッククラウドデータセンタとを含むパブリッククラウドデータセンタの第２のセット内に、前記パブリッククラウドデータセンタ内の測定マシンをデプロイすることと、
パブリッククラウドデータセンタの異なるペア間のネットワーク接続を評価するための測定値の受信することと、
前記仮想ネットワークを実施する転送要素の前記セットを設定するために、前記受信された測定値を用いて、パブリッククラウドデータセンタの前記第１のセットを介するパスの第１のセットを識別するためのパス探索の第１のセットを実行することと、
パブリッククラウドデータセンタの前記第１のセットの外部の前記特定のパブリッククラウドを使用し且つパスの前記第１のセットの１つのパスより良い少なくとも１つの特定のパスを識別するために、前記受信された測定値を用いて、パブリッククラウドデータセンタの前記第２のセットを介するパスの第２のセットを識別するためのパス探索の第２のセットを実行することと、を含む方法。
【請求項９】
請求項８に記載の方法であって、前記受信された測定値は、
(i) パブリッククラウドデータセンタのペア間の接続で発生した損失を定量化する測定値と、
(ii) パブリッククラウドデータセンタのペア間の接続で発生した遅延を定量化する測定値と、
(iii) パブリッククラウドデータセンタのペア間の接続で発生したジッタを定量化する測定値と、
(iv) パブリッククラウドデータセンタのペア間の接続の信頼性を表す測定値との少なくとも１つを含む、方法。

【請求項１０】
請求項１に記載の方法であって、パブリッククラウドデータセンタの少なくとも１つのペアが同じ領域内の２つのパブリッククラウドデータセンタであり、パブリッククラウドデータセンタの別のペアが２つの異なる領域内の２つのパブリッククラウドデータセンタである、方法。
【請求項１１】
２つ以上のパブリッククラウドデータセンタのセットにわたって定義され、エンティティのマシンのセットを接続する前記エンティティの仮想ネットワークに関するリコメンデーションを提供する方法であって、前記方法は、
エンティティから、前記エンティティのための仮想ネットワークが定義されるべきパブリッククラウドデータセンタの第１のセットの識別を受信することと、
パブリッククラウドデータセンタの前記第１のセットで動作する転送要素の第１のセットを設定して前記エンティティのための前記仮想ネットワークを実施するために、パス探索の第１のセットを実行して、転送要素の前記第１のセットを介するパスの第１のセットを識別することと、
前記エンティティのために、前記エンティティによって識別されたパブリッククラウドデータセンタの前記第１のセットには無い少なくとも１つの特定のパブリッククラウドを含むパブリッククラウドデータセンタの第２のセットを介するパス探索の第２のセットを実行することと、
パス探索の前記第２のセットが、特定のパブリッククラウドを使用し且つパブリッククラウドデータセンタの前記第１のセットを介する、識別されたパスの前記第１のセットのパスよりも優れている特定のパスを少なくとも１つ識別した場合、前記特定のパブリッククラウドをパブリッククラウドデータセンタの前記第１のセットに追加するリコメンデーションを提供する、方法。
【請求項１２】
請求項１１に記載の方法であって、更に、
前記エンティティが前記リコメンデーションを受け入れない場合、パスの前記第１のセットを有する転送要素の前記第１のセットを設定し、
前記エンティティが前記リコメンデーションを受け入れる場合、
転送要素の前記第１のセットに、前記特定のパブリッククラウドの特定の転送要素を追加することと、
前記特定の転送要素を有する転送要素の前記第１のセットを、パスの前記識別された第１のセットと前記特定のパスであって、前記特定のパスが置き換える、パスの前記第１のセットの任意のパスでないパスを用いて設定することと、を含む方法。
【請求項１３】
請求項１１に記載の方法であって、前記転送要素を設定する前に、それぞれのパブリッククラウドデータセンタに前記転送要素をデプロイすることをさらに含む、方法。
【請求項１４】
請求項１３に記載の方法であって、前記転送要素は、前記パブリッククラウドデータセンタ内のホストコンピュータ上で動作するマシンとして実施される転送要素を含む、方法。
【請求項１５】
請求項１４に記載の方法であって、マシンとして実施される前記転送要素は、前記エンティティとは異なる前記パブリッククラウドデータセンタのテナントの他の仮想マシンと共にホストコンピュータ上で実行される仮想マシンを含む、方法。
【請求項１６】
請求項１３に記載の方法であって、前記転送要素はルータを含み、前記転送要素を設定することは、前記識別されたパスに基づいて前記ルータにルーティング記録を提供することを含む、方法。
【請求項１７】
請求項１６に記載の方法であって、前記ルーティング記録は次ホップ記録である、方法。

【請求項１８】
請求項１１に記載の方法であって、パブリッククラウドデータセンタの前記第１のセットは、異なるパブリッククラウドプロバイダによって運用されるパブリッククラウドデータセンタを含む、方法。
【請求項１９】
請求項１１に記載の方法であって、パブリッククラウドデータセンタの前記第１のセットは、第１のパブリッククラウドプロバイダによって運用されるパブリッククラウドデータセンタを含み、一方、パブリッククラウドデータセンタの前記第２のセットは、第２のパブリッククラウドプロバイダによって運用される少なくとも１つのパブリッククラウドデータセンタを含む、方法。
【請求項２０】
請求項１１に記載の方法であって、前記エンティティのマシンの前記セットは、パブリッククラウドデータセンタの前記セットの外部のマシンを含み、且つ、前記エンティティのモバイルマシン、オフィスマシンまたはプライベートデータセンタマシンを含む、方法。
【請求項２１】
２つ以上のパブリッククラウドデータセンタにわたって定義されるエンティティの仮想ネットワークに関するリコメンデーションを提供する方法であって、前記方法は、
エンティティから、前記エンティティのための前記仮想ネットワークを定義するためのパブリッククラウドデータセンタのセットの識別の受信することと、
前記エンティティのための前記仮想ネットワークを実施するためのパブリッククラウドデータセンタの前記セットで動作する転送要素のセットをデプロイし、設定することと、
パブリッククラウドデータセンタの前記セットから削除する特定のパブリッククラウドデータセンタの少なくとも１つを識別するために前記仮想ネットワークの動作を監視することと、
前記仮想ネットワークの性能メトリックを改善するために、パブリッククラウドデータセンタの前記セットから前記特定のパブリッククラウドデータセンタを削除するリコメンデーションを提供することと、を含む方法。
【請求項２２】
請求項２１に記載の方法であって、前記仮想ネットワークの動作を監視することは、
転送要素の前記セットのうちの前記転送要素を通過するデータメッセージフローに関する統計値を収集することと、
閾値未満で使用されている前記特定のパブリッククラウドデータセンタの転送要素を識別するために前記統計値を使用することと、
前記識別に基づいて、前記エンティティのためにデプロイされ、設定されている転送要素の前記セットから前記転送要素を削除すべきかを判定することと、を含む方法。
【請求項２３】
請求項２１に記載の方法であって、さらに、
前記エンティティが前記リコメンデーションを受け入れる場合、
転送要素の前記セットから、前記特定のパブリッククラウドデータセンタの特定の転送要素を削除して、転送要素の修正されたセットを生成することと、
転送要素の前記修正されたセットのうちの前記転送要素を再設定して、前記特定の転送要素を用いない前記仮想ネットワークを実施することと、を含む方法。
【請求項２４】
請求項２３に記載の方法であって、前記特定の転送要素を削除することは、前記特定のパブリッククラウドデータセンタの前記特定の転送要素の動作を終了することを含む、方法。
【請求項２５】
請求項２４に記載の方法であって、
前記特定の転送要素は、前記特定のパブリッククラウドデータセンタのホストコンピュータにデプロイされ且つ転送要素動作を実行するように設定される特定のマシンインスタ

ンスであり、
前記特定の転送要素を削除することは、前記特定のパブリッククラウドデータセンタのAPIを用いて前記特定のマシンインスタンスをシャットダウンすることを含む、方法。
【請求項２６】
請求項２３に記載の方法であって、
前記監視することの前に転送要素の前記セットを設定することは、パス探索の第１のセットを実行して、転送要素の前記セットを介するパスの第１のセットを識別することと、パスの前記識別された第１のセットを用いて、前記仮想ネットワークを実施する転送要素の前記セットを設定することとを含み、
転送要素の前記修正されたセットを再設定することは、パス探索の第２のセットを実行して、前記仮想ネットワークを実施する転送要素の前記修正されたセットを介するパスの第２のセットを識別することを含む、方法。
【請求項２７】
請求項２６に記載の方法であって、前記転送要素はルータを含み、前記転送要素を設定することは、前記識別されたパスに基づいて前記ルータにルーティング記録を提供することを含む、方法。
【請求項２８】
請求項２７に記載の方法であって、前記ルーティング記録は次ホップ記録である、方法。
【請求項２９】
請求項２１に記載の方法であって、前記転送要素は、前記パブリッククラウドデータセンタのホストコンピュータで動作するマシンとして実施される転送要素を含み、少なくとも１つのホストコンピュータは、前記エンティティとは異なるパブリッククラウドデータセンタの少なくとも１つのテナントの少なくとも１つのマシンインスタンスを実行する、方法。
【請求項３０】
請求項２１に記載の方法であって、前記性能メトリックはコスト値である、方法。
【請求項３１】
請求項３０に記載の方法であって、前記リコメンデーションを提供することは、前記リコメンデーションが受け入れられ、前記特定のパブリッククラウドデータセンタが、前記仮想ネットワークが広がるパブリッククラウドデータセンタの前記セットから削除された場合に、前記仮想ネットワークに対して期待される調整された運用メトリック値と共に改善されたコスト値を提供することを含む、方法。
【請求項３２】
請求項２１に記載の方法であって、パブリッククラウドデータセンタの前記セットは、異なるパブリッククラウドプロバイダによって運用されるパブリッククラウドデータセンタを含む、方法。
【請求項３３】
請求項２１に記載の方法であって、前記仮想ネットワークは、パブリッククラウドデータセンタの前記セットの外部のマシンを含み、前記エンティティのモバイルマシン、オフィスマシンまたはプライベートデータセンタマシンを含む、前記エンティティのマシンのセットを接続する、方法。
【請求項３４】
少なくとも１つの処理ユニットによって実行されるプログラムを記憶する機械可読媒体であって、前記プログラムは、請求項１から３３のいずれか１項に記載の方法を実施するための命令のセットを含む、機械可読媒体。
【請求項３５】
少なくとも１つの処理ユニットによって実行されると、請求項１から３３のいずれか１項に記載の方法を実行するコンピュータプログラム。
【請求項３６】
コンピューティングデバイスであって、

処理ユニットのセットと、
少なくとも１つの前記処理ユニットによって実行されると、請求項１から３３のいずれか１項に記載の方法を実行するプログラムを記憶する機械可読媒体とを含む、コンピューティングデバイス。
【発明の詳細な説明】
【背景技術】
【０００１】
[0001] 現在、企業エンタープライズネットワークは、企業のさまざまなオフィスや部門を安全に接続する通信バックボーンである。このネットワークは、通常、広域ネットワーク（WAN）で、（1）支店および地域キャンパスのユーザ、（2）ビジネスアプリケーション、イントラネット、およびそれらに対応するデータをホストする企業データセンタ、（3）企業ファイアウォールおよびDMZ(非武装地帯）を介してグローバルインターネットを接続する。エンタープライズネットワークには、フレームリレーやMPLS (マルチプロトコル・ラベル・スイッチング）などの高価なリース回線によって相互接続されたスイッチ、ルータ、ミドルボックスアプライアンスなどの特殊なハードウェアが含まれる。
【０００２】
[0002] ここ数年、企業がコミュニケーションサービスを提供し、利用する方法にパラダイムシフトがあった。まず、モビリティ革命により、ユーザはいつでも、スマートフォンを中心としたモバイルデバイスを使用して、どこからでもサービスにアクセスできるようになった。このようなユーザは、パブリックインターネットおよびセルラネットワークを介してビジネスサービスにアクセスする。同時に、サードパーティのSaaS (Software as a Service) ベンダ（例えばSalesforce、Workday、Zendesk など）が従来のオンプレミスアプリケーションに取って代わり、プライベートデータセンタでホストされている他のアプリケーションがパブリッククラウドに再配置されてきた。このトラフィックはまだエンタープライズネットワーク内で伝送されているが、その大部分は企業ネットワーク境界の外側で発信および終端され、パブリックインターネットと企業ネットワークの両方を（1度または2度）通過する必要がある。最近の研究では、企業ネットワークの40％が、バックホールされたトラフィック（すなわち、企業ネットワークで観測されたインターネットトラフィック）の割合が80％以上であることを報告している。これは、企業トラフィックの大半が高価な専用線と消費者用インターネットの両方を介して伝送されることを意味する。
【０００３】
[0003] 消費者中心のサービスとしては、インターネット自体はビジネストラフィックに対して劣悪な媒体である。これは、重要なビジネスアプリケーションにおいて期待される、信頼性、QoS(Quality of Service)保証およびセキュリティを欠いている。さらに、絶えず増加する消費者トラフィック需要、ネット中立性規制、及びメジャーなプレーヤ（例えば、Netflix、Google、パブリッククラウド）によるインターネットバイパスの創出により、トラフィックの単位あたりの収益は低下してきた。このような傾向により、サービスプロバイダが消費者の要求に迅速に対応し、適切なビジネスサービスを提供するインセンティブが減少している。
【０００４】
[0004] パブリッククラウドの成長により、企業はより多くのコンピューティング基盤をパブリッククラウドデータセンタに移行している。パブリッククラウドプロバイダは、コンピューティングおよびネットワーク基盤への投資の先頭に立っている。これらのクラウドサービスは、2016 年にはAzure、AWS、IBMおよびGoogle を用いてそれぞれ世界中の領域の38、16、25および14 にまで拡大している、多くのデータセンタを世界中に構築してきた。各パブリッククラウドプロバイダは、潜水船によって配備された暗いファイバケーブルと海底ケーブルを使用する高価な高速ネットワークを使用して、独自のデータセンタを相互接続している。
【０００５】

[0005] 今日では、このような変化にもかかわらず、企業のネットワークポリシによって、すべての企業トラフィックがセキュアなWAN ゲートウェイを通過するよう強制されることがよくある。ユーザがモバイル化し、アプリケーションがSaaS やパブリッククラウドに移行するにつれて、企業WAN は全ての企業通信を低速化させるコストのかかる回り道となっている。企業WAN のトラフィックのほとんどは、インターネットから送信されるか、インターネットに送信される。インターネットを介してこのトラフィックを送信する代替的なセキュアなソリューションは、パフォーマンスが低く信頼性が低いため、適切ではない。

【発明の概要】

【０００６】

[0006] いくつかの実施形態は、1つ以上の領域（例えば、複数の都市、州、国など）の1つ以上のパブリッククラウドプロバイダの複数のパブリッククラウドデータセンタを介して、エンティティに対する仮想ネットワークを確立する。このような仮想ネットワークを確立することができるエンティティの例には、ビジネスエンティティ（例えば、企業）、非営利エンティティ（例えば、病院、研究組織など）、教育エンティティ（例えば、総合大学、カレッジなど）、または他の任意のタイプのエンティティが含まれ得る。パブリッククラウドプロバイダーの例としては、Amazon Web Services (AWS)、Google Cloud Platform (GCP)、Microsoft Azure などが含まれる。

【０００７】

[0007] いくつかの実施形態では、高速で信頼性の高いプライベートネットワークは、2つ以上のパブリッククラウドデータセンタ（パブリッククラウド）を相互接続する。いくつかの実施形態は、仮想ネットワークを、１つ以上のプライベートネットワーク（例えば、支社、部門、エンティティの部門、またはそれらに関連するデータセンタ内のネットワーク）、モバイルユーザ、SaaS(Software as a Service)プロバイダマシン、パブリッククラウド内のマシンおよび／またはサービス、およびその他のウェブアプリケーションを相互接続するための、複数のパブリッククラウドにまたがるオーバーレイネットワークとして定義する。

【０００８】

[0008] いくつかの実施形態は、パブリッククラウドコンポーネントを構成する論理的に集中化されたコントローラクラスタ（例えば、１つ以上のコントローラサーバのセット）を利用して、複数のパブリッククラウド上で仮想ネットワークを実装する。いくつかの実施形態では、冗長性および高可用性を改善するために、このクラスタ内のコントローラは、様々な異なるロケーション（例えば、異なるパブリッククラウドデータセンタ内）に存在する。いくつかの実施形態では、コントローラクラスタは、仮想ネットワークを確立するために使用されるパブリッククラウドコンポーネントの数、またはこれらのコンポーネントに割り当てられるコンピューティングまたはネットワークリソースを、スケールアップまたはスケールダウンする。

【０００９】

[0009] いくつかの実施形態では、同じパブリッククラウドプロバイダの同じパブリッククラウドのセット、および／または同じパブリッククラウドプロバイダまたは異なるパブリッククラウドプロバイダの異なるパブリッククラウドのセット上に、異なるエンティティのための異なる仮想ネットワークを確立する。いくつかの実施形態では、仮想ネットワークプロバイダは、異なるテナントが同じまたは異なるパブリッククラウド上で異なる仮想ネットワークを定義することを可能にするソフトウェアおよびサービスを提供する。いくつかの実施形態では、同じコントローラクラスタまたは異なるコントローラクラスタを使用して、パブリッククラウドコンポーネントを構成し、複数の異なるエンティティに対して同じまたは異なるパブリッククラウドのセット上の異なる仮想ネットワークを実装することができる。

【００１０】

[0010] 1つ以上のパブリッククラウド上にテナントの仮想ネットワークをデプロイする

ため、コントローラクラスタ は、（1）テナントの支店、データセンタ、モバイルユーザおよびSaaS プロバイダのロケーションに基づいて、テナントの仮想ネットワークに出入りする可能性のある入出力ルータを識別し、（2） 仮想ネットワークを実装する他の中間パブリッククラウドルータを介して、識別された入力ルータから識別された出力ルータへ通過する経路を識別する。これらの経路を識別した後、コントローラクラスタはこれらの経路を（１つ以上の）パブリッククラウド内の仮想ネットワークルータの転送テーブルに伝播する。Open vSwitch(OVS)ベースの仮想ネットワークルータを使用する実施形態では、コントローラはOpenFlowを使用して経路を配信する。

【００１１】

[0011] いくつかの実施形態は、異なるエンティティのためにいくつかのパブリッククラウドデータセンタにわたって異なる仮想ネットワークをデプロイするための新しい方法を提供する。各エンティティに対して、方法(1)は、エンティティのマシンのセットを接続するための1つ以上のパブリッククラウドプロバイダのパブリッククラウドデータセンタのセットを識別し、(2)パブリッククラウドデータセンタの識別されたセット内にエンティティの管理転送ノード(MFN)をデプロイし、(3)パブリッククラウドデータセンタのセットにわたってエンティティのマシンのセットを接続する仮想ネットワークを実施するようにMFNを構成する。

【００１２】

[0012] いくつかの実施形態における管理された転送ノードは、転送動作を実行するために、1つ以上のホストコンピュータのセット上で実行される1つ以上のモジュールを含む。いくつかの実施形態におけるMFNは、サービス動作、例えば、NAT動作、ロードバランシング動作などの他の動作も実行する。いくつかの実施形態では、方法がエンティティのためにデプロイするMFNは、単にそのエンティティのマシンのためのデータメッセージフローを処理するために使用される。例えば、いくつかの実施形態では、エンティティのためのデプロイされたMFNは、そのエンティティのためのデータメッセージフローのみを運ぶため、専用のMFNである。

【００１３】

[0013] いくつかの実施形態では、異なるエンティティに対して方法が定義する異なる仮想ネットワークは、それらがまたがるパブリッククラウドデータセンタ、それらが使用する異なるパブリッククラウドプロバイダのパブリッククラウド、および/またはそれらが定義されている異なるパブリッククラウド領域で異なる可能性がある。いくつかの実施形態では、その仮想ネットワークによって接続されるエンティティのマシンは、任意のパブリッククラウドの外部にあるマシンである。他の実施形態では、エンティティのマシンのいくつかはパブリッククラウド内にあり、他のマシンはパブリッククラウド外に存在する。また、いくつかの実施形態では、エンティティのマシンは、エンティティが特定のSaaS動作のために使用するSaaSプロバイダマシンを含む。

【００１４】

[0014] いくつかの実施形態では、本方法は、エンティティのネットワーク管理者からの入力を受信することによって、エンティティに対するパブリッククラウドデータセンタのセットを識別する。いくつかの実施形態では、この入力は、仮想ネットワークが定義されるべき、使用するパブリッククラウドプロバイダおよび/またはパブリッククラウド領域を指定する。併せて、または代替的に、いくつかの実施形態では、この入力は、使用する実際のパブリッククラウドデータセンタを指定する。上述のアプローチでは、エンティティが所望のパブリッククラウドプロバイダ、領域および/またはデータセンタに関して異なる入力を提供することが多いため、異なるエンティティはしばしば非常に異なる仮想ネットワークとなる。いくつかの実施形態の方法は、エンティティの入力を介してエンティティに対して識別されたパブリッククラウドデータセンタのセットを、以下にさらに説明するように、方法がエンティティのデータセンタのセットに追加するために望ましいデータセンタとして識別する1つ以上のパブリッククラウドデータセンタで補完する。

【００１５】

[0015] MFNを構成するために、この方法は、パブリッククラウドデータセンタ(PCD)に測定エージェントをデプロイし、これらのエージェントは、異なるPCDのペアまたは異なるPCDグループのペア(例えば、異なるパブリッククラウド領域または可用性ゾーン間)間のネットワーク接続の品質を定量化するネットワーク測定値を生成するためにメッセージを交換する。2つのPCDまたはPCDグループ間の接続に対してメソッドが生成するネットワーク測定の例には、この接続で発生する損失、遅延、ジッタ、および接続の信頼性とコストが含まれる。

【００１６】

[0016] いくつかの実施形態では、方法は、エンティティについて識別するPCDのセットについてそれが生成するネットワーク測定値に基づいて、エンティティのためにデプロイされるMFNを設定する。例えば、いくつかの実施形態では、本方法は、(1)エンティティの識別されたPCDのセットに対して生成された測定値を使用して、識別されたPCDのセットにわたってエンティティのマシンを接続するパスのセットを識別するパス識別プロセスを実行することによって、および(2)識別されたパスを使用して、異なるパスに沿ってデータメッセージフローを転送するようにMFNを設定する次ホップ記録を定義することによって、エンティティのデプロイされたMFNを設定する。MFNを設定するために、いくつかの実施形態における方法は、次ホップ記録を、それらをMFNに配布するコントローラのセットに提供する。

【００１７】

[0017] パスを識別するために、方法は、最小コスト(例えば、最短)パス探索を実行するために、生成された測定値を使用する。いくつかの実施形態は、異なるエンティティが、そのパス探索を実行する際に、異なるタイプの測定を使用するように方法を指示することを可能にする。例えば、あるエンティティは、メッセージ遅延を最小化するように方法を指示することができ、別のエンティティは、メッセージジッタを最小化するように方法を指示することができ、さらに別のエンティティは、損失および遅延を最小化するように方法を指示することができる。各エンティティに対して、いくつかの実施形態における方法は、そのパス探索動作を設定して、エンティティによって指定された基準のセットを最適化する。

【００１８】

[0018] 上述したように、いくつかの実施形態の方法は、エンティティの入力を介してエンティティに対して識別するパブリッククラウドデータセンタのセットを、方法がエンティティのデータセンタのセットに追加するために望ましいデータセンタとして識別する1つ以上のPCDまたはPCDグループで補完する。いくつかの実施形態では、方法は、エンティティのための任意のMFNをデプロイする前に、エンティティに、特定のパブリッククラウド、PCDグループまたはPCDを、エンティティによって指定されたパブリッククラウドまたはPCDのリストに追加するためのリコメンデーションを提供する。併せて、または代替的に、いくつかの実施形態における方法は、エンティティのMFNをデプロイし、それらの使用に関する統計値を収集し、統計値を分析して、パブリッククラウドまたはPCDを追加することが望ましいことを決定した後に、1つ以上のパブリッククラウドまたはPCDを追加するためのリコメンデーションを提供する。収集された統計値の分析に基づいて、いくつかの実施形態における方法はまた、未利用のパブリッククラウドまたはPCDの削除をリコメンドする。

【００１９】

[0019] いくつかの実施形態の方法は、エンティティが、特定の状況下で、そのエンティティのために特にデプロイされていない他のMFNを使用することを可能にする。いくつかの実施形態におけるこれらの他のMFNは、単一のエンティティに対して特別にデプロイされる専用MFNとは対照的に、複数のエンティティがそれらを使用することができる共有MFNである。いくつかの実施形態では、1つのエンティティに対して特別にデプロイされる専用MFNは、複数のエンティティによってデプロイされ使用される共有MFNと同じ属性を持ち、同じ動作(例えば、その転送動作を実行する際のテナント識別子のチェック)を実

行する。これらの実施形態における専用MFNと共有MFNとの唯一の違いは、専用MFNは、ただ1つのテナントのためのデータメッセージを処理するために使用され、一方、共有MFNは、複数のエンティティのためのデータメッセージを処理するために使用されることである。

【００２０】

[0020] エンティティが、そのエンティティのために特別にデプロイされていない共有MFNを使用することができる例の1つは、それらのロケーションのエンティティのモバイルデバイスによるリモートロケーションでのMFNの使用を伴う。たとえば、主に1つの領域で運営されているエンティティ(北米のオフィスのみなど)もあるが、国際的な移動を行うユーザがいる可能性があり、パブリッククラウドを介してデプロイされている仮想ネットワークを介してエンティティのネットワークにアクセスする必要がある場合がある。このような状況では、いくつかの実施形態の方法は、これらの移動ユーザのモバイルデバイス(例えば、電話、タブレット、ラップトップなど)が、その方法が外国の1つ以上のPCD(例えば、ヨーロッパまたはアジアのパブリッククラウド)にデプロイするMFNを介してその仮想ネットワークにアクセスすることを可能にする。

【００２１】

[0021] 別の例は、インターネットバックボーンの近くのデータセンタにデプロイされるか、インターネットバックボーンの一部として機能するMFNの使用を含む。インターネットバックボーンの近くまたはその位置にマシンを設置することは困難であり、1つのエンティティにとってコストがかかる可能性がある。したがって、通常、エンティティは、仮想ネットワークがインターネットバックボーンの近くまたはその位置にデプロイされた専用のMFNを持つことを要求しない可能性がある。このような専用MFNをデプロイすることさえできない場合がある。

【００２２】

[0022] エンティティのためにMFNをデプロイする前または後に、方法のパス探索は、エンティティの特定のコンピューティングノードのために、インターネットバックボーンの近くまたはそこにデプロイされた1つ以上の共有MFNを経由するパスを使用することが望ましい(すなわち、インターネットバックボーンの近くまたはそこにある1つ以上のMFNを使用するために、エンティティの専用仮想ネットワークから離れるパスを使用すること)。このような場合、いくつかの実施形態の方法は、インターネットバックボーンの近くまたはそこにデプロイされた共有MFNが使用されるべきであるというリコメンデーションをエンティティに提供する。エンティティがこのリコメンデーションを受け入れると、方法は、インターネットバックボーンの近くまたはそこにある共有MFNを使用する識別されたパスを使用するようにMFNを設定する。

【００２３】

[0023] いくつかの実施形態の方法は、エンティティの専用仮想ネットワークが輻輳しているように見えるか、またはエンティティのために特別にデプロイされた1つ以上のMFNで輻輳していると予想されるとき、共有MFNの一時的な使用を可能にする。いくつかの実施形態では、方法は、特定のエンティティに対してデプロイされた専用MFNの使用に関する統計値を収集し、分析する。この解析に基づいて、方法は、輻輳している1つ以上のMFNを識別し、それに応答して、データメッセージフローのいくつかを共有MFNにリダイレクトするために1つ以上のネットワーク要素(例えば、ロードバランサ)を再設定し、それらの共有MFNを再設定して、それらが宛先ノードに到達するか、または別の入力ノードに到達するまで、エンティティのデータメッセージフローをエンティティの仮想ネットワークに転送する。

【００２４】

[0024] いくつかの実施形態における方法は、仮想ネットワークを実施するためのMFNをデプロイ及び設定しない。例えば、いくつかの実施形態では、方法は、MFNをデプロイし設定する他のプロセスへのPCDまたはPCDグループ間の接続を定量化する測定値を提供する。他の実施形態では、方法は、パブリッククラウドにアプリケーションマシンをデプ

ロイし、測定値を使用してそのようなデプロイに最適なロケーションを識別するプロセスなど、他のクラウドベースの動作を実行する他のプロセスに、これらの測定値を提供する。さらに他の実施形態の方法は、これらの測定値を使用して、他のクラウドベースの動作自体を実行する(例えば、パブリッククラウドにアプリケーションマシンをデプロイし、測定値を使用して、そのようなデプロイに最適なロケーションを識別する)。

【００２５】

[0025] 前述の概要は、本発明のいくつかの実施形態の簡単な紹介として役立つように意図されている。本書において開示されているすべての発明の主題の紹介または概要を意味するものではない。以下の詳細な説明および詳細な説明で参照される図面は、概要ならびに他の実施形態で説明される実施形態をさらに説明する。したがって、本書面によって説明されるすべての実施形態を理解するために、概要、詳細な説明、図面、および特許請求の範囲の全体的な検討が必要である。さらに、特許請求の範囲に記載される主題は、概要、詳細な説明、および図面における例示的な詳細によって限定されるべきではない。

【図面の簡単な説明】

【００２６】

[0026] 本発明の新規な特徴は、添付の特許請求の範囲に記載されている。しかしながら、説明のために、本発明のいくつかの実施形態を以下の図に示す。

【図１Ａ】[0027] 図1A は、2つのパブリッククラウドプロバイダの複数のパブリッククラウドデータセンタ上に、１つの企業に対して定義された仮想ネットワークを示している。

【図１Ｂ】[0028] 図1B は、パブリッククラウド上にデプロイされた、2つの企業テナントに対する2つの仮想ネットワークの例を示している。

【図１Ｃ】[0029] 図1C は、パブリッククラウド上にデプロイされた1つのネットワークと、パブリッククラウドの別のペア上にデプロイされたもう1つの仮想ネットワークの、2つの仮想ネットワークの例を代替的に示している。

【図２】[0030] 図2は、本発明のいくつかの実施形態の管理転送ノードおよびコントローラクラスタの例を示す。

【図３】[0031] 図3は、いくつかの実施形態において、コントローラ測定処理レイヤが生成する測定グラフの例を示す。

【図４Ａ】[0032] 図4Aは、いくつかの実施形態において、測定グラフからコントローラパス識別レイヤが生成するルーティンググラフの例を示す。

【図４Ｂ】[0033] 図4Bは、2つのSaaSプロバイダの既知のIPを、これらのSaaSプロバイダのデータセンタに最も近いデータセンタにあるルーティンググラフの2つのノードに追加する例を示している。

【図４Ｃ】[0034] 図4Cは、2つのSaaSプロバイダを表すために2つのノードを追加することによって生成されるルーティンググラフを示す。

【図４Ｄ】[0035] 図4Dは、2つのパブリッククラウドにそれぞれ接続する既知のIPアドレスを持つ支店とデータセンタを表すために追加されたノードを含むルーティンググラフを示している。

【図５】[0036] 図5は、コントローラ測定レイヤから受信した測定グラフからルーティンググラフを生成するために、コントローラパス識別レイヤが使用するプロセスを示す。

【図６】[0037] 図6は、いくつかの実施形態のIPsecデータメッセージフォーマットを示す。

【図７】、

【図８】[0038] 図7は、いくつかの実施形態の2つのカプセル化ヘッダの例を示し、図8は、いくつかの実施形態においてこれら2つのヘッダがどのように使用されるかを説明する例を示す。

【図９】、

【図１０】、

【図１１】[0039] 図9-11は、2つの異なるブランチオフィスの2つのコンピュートデバイス間で送信されるメッセージを受信したときに、入力、中間、出力の管理転送ノード(M

FN)によってそれぞれ実行されるメッセージ処理を示している。
【図１２】[0040] 図12は、入力MFNと出力MFNの間に中間MFNを含まない例を示す。
【図１３】[0041] 図13は、支社の企業コンピューティングデバイスから別の支社またはSaaS プロバイダデータセンタの別のデバイスに送信されるメッセージを受信したときに、入力MFN のクラウド転送要素(CFE) によって実行されるメッセージ処理プロセスを示している。
【図１４】[0042] 図14は、1つ以上のパブリッククラウドプロバイダのN個のパブリッククラウドにネットワーク基盤とコントローラクラスタとを有する、仮想ネットワークプロバイダのM個のテナントに対するM個の仮想企業WANの例を示している。
【図１５】[0043] 図15は、仮想ネットワークプロバイダのコントローラクラスタによって実行される、特定のテナントに対して仮想WAN をデプロイおよび管理するプロセスを概念的に示している。
【図１６】[0044] 図16は、いくつかの実施形態の3レイヤSaaSデプロイモデルを示す。
【図１７】[0045] 図17は、いくつかの実施形態の2レイヤSaaSデプロイモデルを示す。
【図１８】[0046] 図18は、マルチホームのマルチマシンコンピューティングノード(MMCN)の経路を定義するためにいくつかの実施形態の中央コントローラクラスタによって使用されるプロセスを示す。
【図１９】[0047] 図19は、2つのMMCNとSaaSデータセンタの2つのブランチノードの例を示している。
【図２０】[0048] 図20は、マルチホームSaaSプロバイダの経路を定義するために、いくつかの実施形態の中央コントローラクラスタによって使用されるプロセスを示す。
【図２１】[0049] 図21は、VNPがいくつかの実施形態において、パブリッククラウドプロバイダの特定のセット、パブリッククラウド領域の特定のセット、および/またはパブリッククラウドデータセンタの特定のセット上にそのようなネットワークをデプロイすることを要求するエンティティのための専用仮想ネットワークを確立するために、専用MFNをデプロイおよび構成するために使用するプロセスを示す。
【図２２】[0050] 図22は、3つの異なる仮想ネットワークが3つの異なる企業のために米国の複数のパブリッククラウド上にデプロイされている例を示している。
【図２３】[0051] 図23は、収集された統計と新しい測定に基づいてMFNを追加するためのリコメンデーションを生成するVNP処理を概念的に示している。
【図２４】[0052] 図24は、パフォーマンスを向上させるために新しいPCDを仮想ネットワークに追加する例を示している。
【図２５】[0053] 図25は、使用率の低い1つ以上のMFNを削除するためのリコメンデーションを生成するVNPプロセスを概念的に示している。
【図２６】[0054] 図26は、このPCD内の使用率の低いMFNを削除するために、仮想ネットワークからPCDを削除する例を示している。
【図２７】[0055] 図27は、データメッセージトラフィックの一部を、エンティティの専用仮想ネットワークの専用MFN から、少なくとも1つのマシンエンドポイントのペア間の少なくとも1つのパスの期間中、1つ以上の共有MFN にオフロードするためのリコメンデーションを生成するVNP プロセスを概念的に示している。
【図２８】、
【図２９】、
【図３０】[0056] 図28-30は、専用仮想ネットワークの輻輳した専用MFNから共有MFNに、一部のデータメッセージが流れる例を示している。
【図３１】[0057] 図31は、専用MFNから離れたデータトラフィック負荷の一部のリダイレクトを停止するプロセスを示している。
【図３２】[0058] 図32は、1つ以上のバックボーンMFNを使用するエンティティの専用仮想ネットワークのリコメンデーションを生成するプロセスを概念的に示している。
【図３３】、
【図３４】[0059] 図33および図34は、データメッセージの一部が、エンティティの専用

仮想ネットワークの1つ以上の専用MFNから、その宛先共有バックボーンMFNに向けられる例を示す。
【図３５】[0060] 図35は、いくつかの実施形態が、VNPの共有MFNを通じて提供される遠隔地からのモバイルデバイスアクセスサポートをもつエンティティのための専用仮想ネットワークをデプロイするために使用するプロセスを示す。
【図３６】[0061] 図36は、北米で定義された専用の仮想ネットワークに接続された中国のモバイルデバイスの例を示している。
【図３７】[0062] 図37は、VNPが専用の仮想ネットワークを確立するために専用のMFNをデプロイし、構成するエンティティのマシンのセットから、いくつかの実施形態がSaaSプロバイダへの最適パスを識別するために使用するプロセスを示している。
【図３８】[0063] 図38は、エンティティが仮想ネットワークをデプロイするために選択するPCDに基づいてSaaSデータセンタが選択される例を示している。
【図３９】[0064] 図39は、MMCN の位置(オフィス、データセンタなど) にある複数のコンピュートデバイスを複数のリンクを介してインターネットに接続するエッジノードの例を示している。
【図４０】[0065] 図40は、SD-WAN エッジノードとインターネット間の可用性の異なるリンクを考慮し、エッジノードが接続する1つ以上のMFN のセットで各MFN に最適なリンクを使用するようにSD-WAN エッジノードを設定するプロセスを示している。
【図４１】[0066] 図41は、仮想ネットワーク内の宛先エンドポイントへの最適パスに最適なリンク/MFN ペアを使用するようにSD-WAN エッジノードを構成する例を示している。
【図４２】[0067] 図42は、いくつかの実施形態の測定インフラストラクチャによって生成された測定値に基づいて、1つ以上のリソースのセットをパブリッククラウドの最適なロケーションにデプロイするために使用するプロセスを示す。
【図４３】[0068] 図43いくつかの実施形態がパブリッククラウドにおける最適なロケーションを識別するために使用するいくつかの構成要素。
【図４４】[0069] 図44は、マシンにアクセスする必要がある特定のエッジノードセットを持つ特定のエンティティのマシンをデプロイするための最適なロケーションを示す例を示している。
【図４５】[0070] 図45は、本発明のいくつかの実施形態が実施されるコンピュータシステムを概念的に示している。
【発明を実施するための形態】
【００２７】
[0071] 本発明の以下の詳細な説明では、本発明の多数の詳細、例および実施形態を説明する。しかしながら、当業者であれば、本発明は、記載された実施形態に限定されず、また、本発明は、説明された特定の詳細および実施例の一部を伴わずに実施することができることは明らか且つ自明である。
【００２８】
[0072] いくつかの実施形態は、1つ以上の領域（例えば、複数の都市、州、国など）の1つ以上のパブリッククラウドプロバイダの複数のパブリッククラウドデータセンタを介して、エンティティに対する仮想ネットワークを確立する。このような仮想ネットワークを確立することができるエンティティの例には、ビジネスエンティティ（例えば、企業）、非営利エンティティ（例えば、病院、研究組織など）、教育エンティティ（例えば、総合大学、カレッジなど）、または他の任意のタイプのエンティティが含まれ得る。パブリッククラウドプロバイダーの例としては、Amazon Web Services (AWS)、Google Cloud Platform (GCP)、Microsoft Azure などが含まれる。
【００２９】
[0073] いくつかの実施形態は、仮想ネットワークを、複数のパブリッククラウドデータセンタ（パブリッククラウド）にまたがって、１つ以上のプライベートネットワーク（例えば、支社、部門、部署、またはそれらに関連するデータセンタ内のネットワーク）、

モバイルユーザ、SaaS (Software as a Service)プロバイダマシン、パブリッククラウド内のマシンおよび／またはサービス、およびその他のウェブアプリケーションを相互接続するオーバーレイネットワークとして定義する。いくつかの実施形態では、高速で信頼性の高いプライベートネットワークは、2つ以上のパブリッククラウドデータセンタを相互接続する。

【００３０】

[0074] いくつかの実施形態における仮想ネットワークは、インターネットを介したこのトラフィックのルーティングを最小化しようとする一方で、最適なエンドツーエンドの性能、信頼性、およびセキュリティのために、エンティティのデータメッセージの宛先へのルーティングを最適化するように構成することができる。また、いくつかの実施形態における仮想ネットワークは、ネットワークを通過するデータメッセージフローのレイヤ4処理を最適化するように構成することができる。例えば、いくつかの実施形態では、仮想ネットワークは、接続パスにわたってレート制御機構を分割することによって、TCP（トランスポート制御プロトコル）接続のエンドツーエンドのレートを最適化する。

【００３１】

[0075] いくつかの実施形態では、複数のパブリッククラウドに展開される複数のコンポーネントを構成することによって、仮想ネットワークを確立する。これらのコンポーネントには、いくつかの実施形態において、ソフトウェアベースの測定エージェント、ソフトウェア転送要素（例えば、ソフトウェアルータ、スイッチ、ゲートウェイ等）、レイヤ4接続プロキシ及びミドルボックスサービスマシン（例えば、アプライアンス、VM、コンテナ等）が含まれる。

【００３２】

[0076] いくつかの実施形態は、パブリッククラウドコンポーネントを構成する論理的に集中化されたコントローラクラスタ（例えば、１つ以上のコントローラサーバのセット）を利用して、複数のパブリッククラウド上で仮想ネットワークを実装する。いくつかの実施形態では、冗長性および高可用性を改善するために、このクラスタ内のコントローラは、様々な異なるロケーション（例えば、異なるパブリッククラウドデータセンタ内）に存在する。コントローラクラスタ内の異なるコントローラが異なるパブリッククラウドデータセンタに配置されている場合、いくつかの実施形態のコントローラは、それらの状態（例えば、テナントを識別するためにそれらが生成する設定データ、仮想ネットワークを通る経路など）を共有する。いくつかの実施形態では、コントローラクラスタは、仮想ネットワークを確立するために使用されるパブリッククラウドコンポーネントの数、またはこれらのコンポーネントに割り当てられるコンピューティングまたはネットワークリソースを、スケールアップまたはスケールダウンする。

【００３３】

[0077] いくつかの実施形態では、同じパブリッククラウドプロバイダの同じパブリッククラウドのセット、および／または同じパブリッククラウドプロバイダまたは異なるパブリッククラウドプロバイダの異なるパブリッククラウドのセット上に、異なるエンティティのための異なる仮想ネットワークを確立する。いくつかの実施形態では、仮想ネットワークプロバイダは、異なるテナントが同じまたは異なるパブリッククラウド上で異なる仮想ネットワークを定義することを可能にするソフトウェアおよびサービスを提供する。いくつかの実施形態では、同じコントローラクラスタまたは異なるコントローラクラスタを使用して、パブリッククラウドコンポーネントを構成し、複数の異なるエンティティに対して同じまたは異なるパブリッククラウドのセット上の異なる仮想ネットワークを実装することができる。

【００３４】

[0078] 以下の説明では、企業の仮想ネットワークのいくつかの例を示す。しかしながら、ある実施形態は、他の事業体、非営利組織、教育組織などの他のタイプのエンティティのための仮想ネットワークを定義することを、当業者は理解するであろう。また、本明細書で使用されているように、データメッセージは、ネットワークを介して送信される特

定の形式のビットの集合を指す。当業者は、データメッセージという用語が、ネットワークを介して送信されるビットの様々な定式化された集合を指すために本書で使用されることを認識するであろう。これらのビットのフォーマットは、標準化されたプロトコルまたは標準化されていないプロトコルによって指定することができる。標準化されたプロトコルに従ったデータメッセージの例としては、イーサネットフレーム、IPパケット、TCPセグメント、UDPデータグラムなどがある。また、本明細書で使用するL2、L3、L4、L7レイヤ（またはレイヤ2、レイヤ3、レイヤ4、レイヤ7）は、OSI (Open System Interconnection)レイヤモデルの第2データリンクレイヤ、第3ネットワークレイヤ、第4トランスポートレイヤ、第7レイヤプリケーションレイヤをそれぞれ参照している。

【００３５】

[0079] 図1Aは、2つのパブリッククラウドプロバイダAおよびBの複数のパブリッククラウドデータセンタ105および110にわたる、企業に対して定義される仮想ネットワーク100を示す。図示するように、仮想ネットワーク100は、異なるパブリッククラウドに異なる管理転送ノード150を配置し、オーバレイトンネル152を介して管理転送ノード（MFN）を相互に接続することによって確立されるセキュアなオーバレイネットワークである。いくつかの実施形態では、MFNは、パブリッククラウドデータセンタ内のいくつかの異なるコンポーネントを概念的にグループ化したものであり、他のパブリッククラウドデータセンタ内の他のMFN（他のコンポーネントのグループと共に）が、1つ以上のエンティティに対して1つ以上のオーバレイ仮想ネットワークを確立する。

【００３６】

[0080] 以下にさらに説明するように、MFNを形成するコンポーネントのグループは、いくつかの実施形態において、（1）パブリッククラウドデータセンタの外にある外部マシンロケーションである、エンティティのコンピューティングノード（例えば、オフィス、プライベートデータセンタ、リモートユーザなど）とVPN接続を確立する1つ以上のVPNゲートウェイ、（2）共有パブリッククラウドネットワークファブリック上でオーバレイ仮想ネットワークを定義するために、カプセル化されたデータメッセージを相互に転送する1つ以上の転送要素、（3）ミドルボックスサービス動作およびL4－L7最適化を実行する1つ以上のサービスマシン、および（4）パブリッククラウドデータセンタを通る所望のパスを識別するために、パブリッククラウドデータセンタ間のネットワーク接続品質に関する測定値を取得する1つ以上の測定エージェントを含む。いくつかの実施形態では、異なるMFNは、そのようなコンポーネントを異なる配置及び異なる数で有することができ、1つのMFNは、冗長性とスケーラビリティの理由により、そのようなコンポーネントを異なる数で有することができる。

【００３７】

[0081] また、いくつかの実施形態では、各MFNのコンポーネントのグループは、MFNのパブリッククラウドデータセンタ内の異なるコンピュータ上で実行される。いくつかの実施形態では、MFNのコンポーネントのいくつかまたはすべてをパブリッククラウドデータセンタの1つのコンピュータ上で実行することができる。いくつかの実施形態におけるMFNのコンポーネントは、他のテナントの他のマシンも実行するホストコンピュータ上で実行される。これらの他のマシンは、他のテナントの他のMFN の他のマシンであっても、他のテナントの関連のないマシン（コンピューティング仮想マシンやコンテナなど） であってもよい。

【００３８】

[0082] いくつかの実施形態における仮想ネットワーク100は、異なるエンティティ（例えば、仮想ネットワークプロバイダの異なる企業顧客／テナント）のために、同一または異なるパブリッククラウドデータセンタ上に異なる仮想ネットワークをデプロイする仮想ネットワークプロバイダによってデプロイ（配備）される。いくつかの実施形態における仮想ネットワークプロバイダは、MFNをデプロイし、これらのMFNを構成および管理するためのコントローラクラスタを提供するエンティティである。

【００３９】

[0083] 仮想ネットワーク100は、企業のコンピューティングエンドポイント（データセンタ、支店、モバイルユーザなど）を、相互に及び、パブリッククラウド内に存在するか、またはインターネットを介してアクセス可能なプライベートデータセンタに常駐する外部サービス（パブリックウェブサービス、またはOffice365やSalesforceなどのSaaSサービス）に、接続する。更に後述するように、いくつかの実施形態におけるSaasはサードパーティプロバイダがアプリケーションをホストし、顧客にインターネットを介してそれらを利用可能にする、ソフトウェア配信モデルである。

【００４０】

[0084] この仮想ネットワーク100は、異なるパブリッククラウドのさまざまなロケーションを利用して、異なる企業のコンピューティングエンドポイント（異なるプライベートネットワークや企業の異なるモバイルユーザなど）をその周辺のパブリッククラウドに接続する。企業のコンピューティングエンドポイントは、以下の説明では企業コンピューティングノードとも呼ばれる。いくつかの実施形態では、仮想ネットワーク100は、これらのパブリッククラウドを相互接続する高速ネットワークを利用して、パブリッククラウドを通るデータメッセージをその宛先に転送するか、またはその宛先にできるだけ近づけながら、インターネットを通るデータメッセージの横断を低減する。企業のコンピューティングエンドポイントが、仮想ネットワークがまたがるパブリッククラウドデータセンタの外部にある場合、これらのエンドポイントは外部マシンロケーションと呼ばれる。これは、企業の支店、プライベートデータセンタ、およびリモートユーザのデバイスについての場合である。

【００４１】

[0085] 図1Aに示す例では、仮想ネットワーク100は、パブリッククラウドプロバイダAの6つのデータセンタ105aから105fと、パブリッククラウドプロバイダBの4つのデータセンタ110aから110dにまたがっている。これらのパブリッククラウドにまたがって、この仮想ネットワークは、異なる地理的領域に位置する企業テナントのいくつかの支店、企業データセンタ、SaaSプロバイダ、およびモバイルユーザを接続する。具体的には、仮想ネットワーク100は、2つの異なる都市（例えば、カリフォルニア州サンフランシスコおよびインドのプネー(Pune)）の2つの支店130aおよび130b、別の都市（例えば、ワシントン州シアトル）の1つの企業データセンタ134、別の2つの都市（ワシントン州レッドモンドおよびフランスのパリ）の2つのSaaSプロバイダデータセンタ136aおよび136b、および世界の様々な場所のモバイルユーザ140を接続する。そのため、この仮想ネットワークは仮想企業WAN として見ることができる。

【００４２】

[0086] いくつかの実施形態では、支店130aおよび130bは、支店の場所にあるコンピュータと、パブリッククラウドの外部にある支店のプライベートデータセンタとを接続する自身のプライベートネットワーク（例えば、ローカルエリアネットワーク）を有する。同様に、いくつかの実施形態における企業データセンタ134は、自身のプライベートネットワークを有し、パブリッククラウドデータセンタの外部に存在する。しかしながら、他の実施形態では、企業データセンタ134または支店130aおよび130bのデータセンタはパブリッククラウド内にあってもよいが、仮想ネットワークは、企業データセンタまたは支店データセンタが仮想ネットワーク100のエッジに接続するので、このパブリッククラウドにはまたがらない。

【００４３】

[0087] 上述のように、仮想ネットワーク100は、オーバレイトンネル152を介して、異なるパブリッククラウド内の、デプロイされている異なる管理転送ノード150を接続することによって確立される。各管理転送ノード150は、いくつかの構成可能なコンポーネントを含む。上述し、さらに以下に説明するように、MFNコンポーネントは、いくつかの実施形態において、ソフトウェアベースの測定エージェント、ソフトウェア転送要素（例えば、ソフトウェアルータ、スイッチ、ゲートウェイなど）、レイヤ4プロキシ（例えば、TCPプロキシ）、およびミドルボックスサービスマシン（例えば、VM、コンテナなど

）を含む。いくつかの実施形態におけるこれらのコンポーネントのうちの1つ以上は、Open vSwitch、OpenVPN、strongSwanなどの標準化された、または一般的に利用可能なソリューションを使用する。

【００４４】

[0088] いくつかの実施形態では、各MFN (すなわち、概念的にMFNを形成するコンポーネントのグループ）は、パブリッククラウドデータセンタにMFNをデプロイおよび構成する仮想ネットワークプロバイダの異なるテナントによって共有され得る。結合的にまたは代替的に、いくつかの実施形態における仮想ネットワークプロバイダは、特定のテナントのために、1つ以上のパブリッククラウドデータセンタに一意のMFNのセットをデプロイすることができる。例えば、セキュリティ上の理由やサービス品質上の理由から、特定のテナントが別のテナントとMFN リソースを共有したくない場合がある。このようなテナントの場合、仮想ネットワークプロバイダは、複数のパブリッククラウドデータセンタに自身のMFN のセットをデプロイすることができる。

【００４５】

[0089] いくつかの実施形態では、論理的に集中化されたコントローラクラスタ160（例えば、１つ以上のコントローラサーバのセット）は、パブリッククラウド105および110の１つ以上の内外で動作し、管理転送ノード150のパブリッククラウドコンポーネントを構成して、パブリッククラウド105および110上に仮想ネットワークを実装する。いくつかの実施形態では、冗長性および高可用性を改善するために、このクラスタ内のコントローラは、様々なロケーション（例えば、異なるパブリッククラウドデータセンタ内）に存在する。いくつかの実施形態では、コントローラクラスタは、仮想ネットワークを確立するために使用されるパブリッククラウドコンポーネントの数、またはこれらのコンポーネントに割り当てられるコンピューティングまたはネットワークリソースを、スケールアップまたはスケールダウンする。

【００４６】

[0090] いくつかの実施形態では、コントローラクラスタ160、または仮想ネットワークプロバイダの別のコントローラクラスタは、同じパブリッククラウド105および110上に、および／または異なるパブリッククラウドプロバイダの異なるパブリッククラウド上に、別の企業テナントのための別の仮想ネットワークを確立する。コントローラクラスタに加えて、他の実施形態の仮想ネットワークプロバイダは、異なるテナントが同一または異なるパブリッククラウド上に異なる仮想ネットワークをデプロイすることを可能にする転送要素およびサービスマシンをパブリッククラウド内にデプロイする。図1Bは、パブリッククラウド105および110上にデプロイされた2つの企業テナントの2つの仮想ネットワーク100および180の例を示している。図1Cは、代替的に、パブリッククラウド105および110上にデプロイされた1つのネットワーク100と、パブリッククラウド110および115の別のペア上にデプロイされた他の仮想ネットワーク182とを備える、2つの仮想ネットワーク100および182の例を示す。

【００４７】

[0091] MFNの構成されたコンポーネントを介して、図1Aの仮想ネットワーク100は、企業テナントの異なるプライベートネットワークおよび／または異なるモバイルユーザが、これらのプライベートネットワークおよび／またはモバイルユーザに関して、（例えば、接続速度、損失、遅延および／またはコストに関して、物理的距離に関して、および／または、ネットワーク接続信頼性等に関して測定されるような）最適なロケーションにある異なるパブリッククラウドに接続することを可能にする。これらのコンポーネントはまた、いくつかの実施形態における仮想ネットワーク100が、パブリッククラウドを相互接続する高速ネットワークを使用して、パブリッククラウドを通過するデータメッセージを、インターネットを通過するそれらのトラバーサルを低減しながら、それらの宛先に転送することを可能にする。

【００４８】

[0092] いくつかの実施形態では、MFNコンポーネントはまた、エンドツーエンドのパ

フォーマンス、信頼性およびセキュリティを最適化する、ネットワークレイヤ、トランスポートレイヤおよびアプリケーションレイヤでの新規なプロセスを実行するように構成される。いくつかの実施形態では、これらのプロセスのうちの1つ以上は、現在のネットワークプロトコルの硬直化から解放された独自の高性能ネットワーキングプロトコルを実装する。したがって、いくつかの実施形態における仮想ネットワーク100は、インターネット自律システム、ルーティングプロトコル、またはエンドツーエンドのトランスポート機構によって制限されない。

【００４９】

[0093] 例えば、いくつかの実施形態では、MFN150のコンポーネントは、（1）最適化されたマルチパスおよび適応的な集中化ルーティングを生成し、（2）強力なQoSを提供し、（3）中間TCP分割および／または終端を通してエンドツーエンドのTCPレートを最適化し、（4）スケーラブルなアプリケーションレベルのミドルボックスサービス（例えば、ファイアウォール、侵入検知システム（IDS）、侵入防御システム（IPS）、WAN最適化など）を、グローバルネットワーク機能の仮想化（NFV）におけるクラウドのコンピューティング部分に再配置する。従って、既存のネットワークプロトコルに拘束されることなく、企業のカスタマイズされ且つ変化する要求に適合するように、仮想ネットワークを最適化することができる。また、いくつかの実施形態では、仮想ネットワークは、継続的な要求の変化に応じて、性能能力と地理的広がりの両方において、動的かつ弾力的にスケールアップ／ダウンすることができる「従量制(pay as you go)」基盤として構成することができる。

【００５０】

[0094] 仮想ネットワーク100を実装するためには、仮想ネットワークがまたがった各パブリッククラウドデータセンタ105aから105fおよび110aから110d内の少なくとも1つの管理転送ノード150を、コントローラのセットによって構成されなければならない。図2は、本発明のいくつかの実施形態の管理転送ノード150およびコントローラクラスタ160の例を示す。いくつかの実施形態では、各管理転送ノード150は、パブリッククラウドデータセンタ内のホストコンピュータ上で動作するマシン（例えば、ＶＭ又はコンテナ）である。他の実施形態では、各管理転送ノード150は、1つのパブリッククラウドデータセンタ内の同じホストコンピュータ上で動作する複数のマシン（例えば、複数のVMまたはコンテナ）によって実現される。さらに他の実施形態では、1つのMFNの2つ以上のコンポーネントは、1つ以上のパブリッククラウドデータセンタ内の2つ以上のホストコンピュータ上で動作する2つ以上のマシンによって実現され得る。

【００５１】

[0095] 図示されるように、管理転送ノード150は、測定エージェント205、ファイアウォールおよびNATミドルボックスサービスエンジン210および215、1つ以上の最適化エンジン220、エッジゲートウェイ225および230、およびクラウド転送要素235（例えば、クラウドルータ）を含む。いくつかの実施形態では、これらのコンポーネント205から235の各々は、2つ以上のコンポーネントのクラスタとして実装され得る。

【００５２】

[0096] いくつかの実施形態におけるコントローラクラスタ160は、各コンポーネントクラスタを動的にスケールアップまたはスケールダウンして、（1）各コンポーネントの機能を実装するためのマシン（例えば、VMまたはコンテナ）を追加または削除すること、および／または（2）そのクラスタのコンポーネントを実装する以前に配置されたマシンにコンピューティングおよび／またはネットワークリソースを追加または削除することができる。したがって、パブリッククラウドデータセンタ内に配置された各MFN 150を、MFNのクラスタとして見る、或いは、MFNの異なる動作を実行する複数の異なるコンポーネントクラスタを含むノードとして見ることができる。

【００５３】

[0097] また、いくつかの実施形態では、コントローラクラスタは、パブリッククラウドデータセンタ上に仮想ネットワークを定義する異なるテナントに対して、コントローラ

クラスタがパブリッククラウドデータセンタに異なるMFNのセットをデプロイする。このアプローチでは、任意の2つのテナントの仮想ネットワークは、MFNを共有しない。ただし、以下に説明する実施形態では、各MFNを用いて、異なるテナントに対して異なる仮想ネットワークを実装することができる。他の実施形態では、コントローラクラスタ160は、デプロイされた自身専用のMFNのセットを有するテナントの第1のセットの各テナントの仮想ネットワークを実装する一方で、デプロイされたMFNの共有セットを有するテナントの第2のセットの各テナントの仮想ネットワークを実装することができることを、当業者は理解するであろう。

【００５４】

[0098] いくつかの実施形態では、図2に示すように、支店ゲートウェイ225およびリモートデバイスゲートウェイ230は、MFN150に接続する１つ以上の支店130およびリモートデバイス（例えば、モバイルデバイス140）とそれぞれセキュアVPN接続を確立する。このようなVPN接続の一例は、以下でさらに説明するIPsec接続である。しかしながら、当業者は、他の実施形態では、このようなゲートウェイ225および／または230が異なるタイプのVPN接続を確立することを理解するであろう。

【００５５】

[0099] いくつかの実施形態におけるMFN150は、ファイアウォール動作、NAT動作、IPS動作、IDS動作、負荷分散動作、WAN最適化動作など、1つ以上のミドルボックスサービス動作を実行する1つ以上のミドルボックスエンジンを含む。これらのミドルボックス動作（例えば、ファイアウォール動作、WAN最適化動作など）を、パブリッククラウドに配置されたMFNに組み込むことにより、仮想ネットワーク100は、企業のデータセンタおよび／または支店において従来企業のWAN基盤によって実行されていた機能の多くをパブリッククラウドに実装する。

【００５６】

[00100] したがって、多くのミドルボックスサービスに対して、企業のコンピューティングノード（例えば、リモートデバイス、支店、データセンタ）は、プライベートデータセンタまたは支店における企業のWAN基盤にアクセスする必要がなくなり、これらのサービスの多くがパブリッククラウドにデプロイされるようになる。このアプローチにより、企業のコンピューティングノード（リモートデバイス、支店、データセンタなど）からこれらのサービスへのアクセスが高速化され、そうでなければこのようなサービスの提供専用になるプライベートデータセンタでのコストのかかる輻輳ネットワークのボトルネックが回避される。

【００５７】

[00101] このアプローチは、パブリッククラウドデータセンタにおける各種MFN にWAN ゲートウェイの機能を効果的に分散させる。例えば、いくつかの実施形態の仮想ネットワーク100では、従来の企業WANゲートウェイセキュリティ機能（例えば、ファイアウォール動作、侵入検知動作、侵入防御動作など）の大部分またはすべてが、パブリッククラウドMFN（例えば、コンピューティングエンドポイントからの仮想ネットワークへのデータが受信される入力MFN）に移動される。これにより、効果的に、仮想ネットワーク100は、仮想ネットワーク100を実装する多くの異なるMFNで実現される分散WANゲートウェイを有することができる。

【００５８】

[00102] 図2に示す例では、MFN150は、ファイアウォールエンジン210、NATエンジン215、および1つ以上のL4からL7の最適化エンジンを含むように示されている。当業者は、他の実施形態では、MFN 150が他のミドルボックス動作を実行するための他のミドルボックスエンジンを含むことを理解するであろう。いくつかの実施形態では、ファイアウォールエンジン210は、（1）仮想ネットワークへの入力パス上のデータメッセージフロー（例えば、ゲートウェイ225および230が支店130およびモバイルデバイス140から受信して処理するデータメッセージフロー）、および（2）仮想ネットワークからの出力パス上のデータメッセージフロー（例えば、NATエンジン215およびインターネット20

2を介してSaaSプロバイダデータセンタに送信されるデータメッセージフロー）にファイアウォールルールを実施する。

【００５９】

[00103] いくつかの実施形態におけるMFN 150のファイアウォールエンジン210は、ファイアウォールエンジンが、データメッセージフローが仮想ネットワークに入る入力MFNと、データメッセージフローが仮想ネットワークから出る出力MFNとの間の中間ホップであるMFNに属する場合にも、ファイアウォールルールを実施する。他の実施形態では、ファイアウォールエンジン210は、データメッセージフローの入力MFNおよび／または出力MFNの一部である場合にのみ、ファイアウォールルールを実施する。

【００６０】

[00104] いくつかの実施形態では、NATエンジン215は、ネットワークアドレス変換を実行して、インターネット202を介して仮想ネットワークから第三者デバイス（例えば、SaaSプロバイダマシン）への出力パス上のデータメッセージフローの送信元ネットワークアドレスを変更する。このようなネットワークアドレス変換により、アドレス変換なしでテナントやパブリッククラウドプロバイダのプライベートネットワークアドレスを指定するデータメッセージフローを処理するように、第三者マシン（SaaSマシンなど）を適切に構成することを可能にする。これは、異なるテナントやクラウドプロバイダのプライベートネットワークアドレスが重複する可能性があるため、特に問題になる。アドレス変換により、第三者デバイス（SaaSマシンなど）からの応答メッセージが、仮想ネットワーク（メッセージを仮想ネットワークから出すMFNのNATエンジンなど）によって適切に受信されることも可能にする。

【００６１】

[00105] いくつかの実施形態におけるMFNのNATエンジン215は、第三者マシンに届くように仮想ネットワークを出るか、または第三者マシンから仮想ネットワークに入る、各データメッセージフロー上でダブルNAT動作を行う。（参照により本明細書に組み込まれる）米国公開特許出願第2019-0103990A1号において更に記載されるように、2つのNAT動作のうちの1つのNAT動作は、仮想ネットワークに入るときにその入力MFNにおけるそのようなデータメッセージフロー上で実行され、もう1つのNAT動作は、仮想ネットワークから出るときにその出力MFNにおけるデータメッセージフロー上で実行される。

【００６２】

[00106] このダブルNAT アプローチは、より多くのテナントプライベートネットワークを、パブリッククラウドプロバイダのネットワークにマッピングすることを可能にする。このアプローチはまた、テナントプライベートネットワークへの変更に関するデータをMFNへ配信する負荷を低減する。入力または出力NAT動作の前に、いくつかの実施形態は、テナント識別子を使用して、テナントの送信元ネットワークアドレスを、NAT動作によってさらに別の送信元ネットワークアドレスにマッピングされる別の送信元ネットワークアドレスに最初にマッピングするテナントマッピング動作を実行する。ダブルNAT動作を実行すると、テナントプライベートネットワークへの変更に関するデータを配信するためのデータ配信負荷が軽減される。

【００６３】

[00107] 最適化エンジン220は、最良のエンドツーエンド性能と信頼性のために、エンティティのデータメッセージの宛先への転送を最適化する新規のプロセスを実行する。これらのプロセスのいくつかは、現在のネットワークプロトコルの硬直化から解放される、独自の高性能ネットワークプロトコルを実現する。例えば、いくつかの実施形態では、最適化エンジン220は、中間TCP分割および／または終端を通して、エンドツーエンドのTCPレートを最適化する。

【００６４】

[00108] クラウド転送要素235は、データメッセージフローが宛先に到達するために別のパブリッククラウドを通過する必要がある場合に、次ホップMFNのクラウド転送要素（CFE）にデータメッセージフローを転送し、またはデータメッセージフローが同じパブリ

ッククラウドを介して宛先に到達できる場合に、同じパブリッククラウド内の出力ルータにデータメッセージフローを転送する、MFNエンジンである。いくつかの実施形態では、MFN 150のCFE235はソフトウェアルータである。

【００６５】

[00109] データメッセージを転送するために、CFEはトンネルヘッダを用いてメッセージをカプセル化する。異なる実施形態は、トンネルヘッダを用いてデータメッセージをカプセル化するために異なるアプローチを用いる。以下に説明するいくつかの実施形態は、仮想ネットワークに出入りするためのネットワーク入力／出力アドレスを識別するために1つのトンネルヘッダを使用し、データメッセージが、出力MFNに到達するために1つ以上の中間MFNを通過しなければならない場合には、次ホップMFNを識別するために別のトンネルヘッダを使用する。

【００６６】

[00110] 具体的には、いくつかの実施形態では、CFEは、2つのトンネルヘッダである、（1）仮想ネットワークに出入りするための入力CFEおよび出力CFEを識別する内部ヘッダ、および（2）次ホップCFEを識別する外部ヘッダ、を有するデータメッセージを送信する。いくつかの実施形態における内部トンネルヘッダは、仮想ネットワークプロバイダの複数の異なるテナントが仮想ネットワークプロバイダのMFNのCFEの共通セットを使用することを可能にするために、テナント識別子も含む。他の実施形態は、オーバーレイ仮想ネットワークを定義するために、異なるトンネルヘッダを定義する。

【００６７】

[00111] 1つ以上のパブリッククラウド上にテナントの仮想ネットワークをデプロイするために、コントローラクラスタ は、（1）テナント企業のコンピューティングノード（支店、データセンタ、モバイルユーザ、SaaS プロバイダなど）のロケーションに基づいて、テナントの仮想ネットワークに出入りする可能性のある入出力ルータを識別し、（2）仮想ネットワークを実装する他の中間パブリッククラウドルータを介して、識別された入力ルータから識別された出力ルータへ通過する経路を識別する。これらの経路を識別した後、コントローラクラスタは、パブリッククラウド内のMFNのCFE235の転送テーブルにこれらの経路を伝播する。OVSベースの仮想ネットワークルータを使用する実施形態では、コントローラはOpenFlowを使用して経路を配信する。

【００６８】

[00112] いくつかの実施形態では、コントローラクラスタ160は、最良のエンドツーエンドの性能、信頼性、およびセキュリティを達成するために、いくつかのネットワーク処理レイヤを最適化するために仮想ネットワークを実装する各MFN150のコンポーネント205から235を構成することもできる。例えば、いくつかの実施形態では、これらのコンポーネントは、（1）レイヤ3トラフィックのルーティング（例えば、最短経路、パケット重複）を最適化し、（2）レイヤ4TCP輻輳制御（例えば、セグメンテーション、レート制御）を最適化し、（3）セキュリティ機能（例えば、暗号化、精密パケット検査、ファイアウォール）を実装し、および（4）アプリケーションレイヤの圧縮機能（例えば、重複排除、キャッシュ）を実装するために、構成される。仮想ネットワーク内では、企業トラフィックはセキュリティ保護され、検査され、そしてログが記録される。

【００６９】

[00113] いくつかの実施形態では、パブリッククラウドデータセンタ内の各MFNに対して、1つの測定エージェントがデプロイされる。他の実施形態では、パブリッククラウドデータセンタ内またはデータセンタの集合内（例えば、1つの利用可能ゾーン内のデータセンタのような、近くに関連するデータセンタの集合内）の複数のMFNは、1つの測定エージェントを共有する。レイヤ3および4の処理を最適化するために、各管理転送ノード150に関連付けられた測定エージェント205は、そのノードと他のいくつかの「隣接」ノードとの間のネットワーク接続の品質を定量化した測定値を、繰り返し生成する。

【００７０】

[00114] 異なる実施形態は、隣接ノードを異ならせるように定義する。特定のパブリッ

ククラウドプロバイダの1つのパブリッククラウドデータセンタにおける特定のMFNについて、いくつかの実施形態における隣接ノードは、（1）特定のパブリッククラウドプロバイダーの任意のパブリッククラウドデータセンタで動作する任意の他のMFN、および（2）特定のMFNと同じ「領域」内にある別のパブリッククラウドプロバイダーのデータセンタで動作する任意の他のMFNを含む。

【００７１】

[00115] 異なる実施形態は、同一領域を異ならせるように定義する。例えば、いくつかの実施形態は、特定の管理転送ノードの周りの境界形状を指定する距離に関して領域を定義する。他の実施形態は、カリフォルニア北部、カリフォルニア南部など、都市、州、または地域の観点から地域を定義する。このアプローチの前提は、同じパブリッククラウドプロバイダの異なるデータセンタが非常に高速なネットワーク接続で接続されているのに対し、異なるパブリッククラウドプロバイダのデータセンタ間のネットワーク接続は、データセンタが同じ領域内にある場合は高速である可能性が高く、データセンタが異なる領域にある場合は高速ではない可能性が高いということである。異なるパブリッククラウドプロバイダのデータセンタ間の接続では、データセンタが異なる領域にある場合、パブリックインターネットを経由して長い距離を通過する必要がある場合がある。

【００７２】

[00116] 測定エージェント205は、異なる実施形態において異なる方法で測定値を生成する。いくつかの実施形態では、測定エージェントは、周期的に（例えば、毎秒、毎N秒、毎分、毎M分等）pingメッセージを、その隣接する管理転送ノードの測定エージェントのそれぞれに送信する。pingメッセージのサイズが小さい場合、大きなネットワーク接続料金は発生しない。たとえば、各ノードが10 秒ごとに互いのノードにping を送信する100 ノードの場合、各ノードに対して約10Kb／s の入出力測定トラフィックが生成され、これにより、現在のパブリッククラウドの価格が与えられた場合、ノードあたり1 年あたり数ドル（たとえば＄5） のネットワーク消費料金が発生する。

【００７３】

[00117] 測定エージェント205は、受信する応答メッセージの速度に基づいて、ネットワーク接続スループット速度、遅延、損失、およびリンク信頼性などの測定メトリック値を計算し、更新する。これらの動作を繰り返し行うことによって、測定エージェント205は、隣接するノードへのネットワーク接続の品質を表す測定結果のマトリックスを定義し、更新する。エージェント205は、その隣接ノードの測定エージェントと相互作用するので、その測定マトリックスは、ノードのローカルクリークへの接続の品質のみを定量化する。

【００７４】

[00118] 異なる管理転送ノードの測定エージェントは、それらの測定マトリックスをコントローラクラスタ160に送り、その後、すべての異なるクリーク接続データを集約して、管理転送ノードの異なるペア間の接続の集合メッシュビューを得る。コントローラクラスタ160が、2つのペアの転送ノード間のリンクについて異なる測定値（例えば、異なる時間に1つのノードによって取られた測定値）を収集すると、コントローラクラスタは、異なる測定値からブレンドされた値を生成する（例えば、測定値の平均または加重平均を生成する）。いくつかの実施形態における集約メッシュビューは、管理転送ノードの各ペア間のすべてのネットワーク接続のフルメッシュビューであり、他の実施形態では、個々の管理転送ノードの測定エージェントによって生成されたものよりも完全なビューである。

【００７５】

[00119] 図2に示すように、コントローラクラスタ160は、1つ以上の測定処理エンジン280、1つ以上のパス識別エンジン282、および1つ以上の管理インタフェース284のクラスタを含む。不必要な詳細を用いて説明を不明瞭にしないために、これらのクラスタの各々は、単一のエンジンまたはインタフェースレイヤ、即ち、測定処理レイヤ280、パス識別レイヤ282、および管理インタフェースレイヤ284に関して、以下に参照される。

【００７６】

[00120] 測定処理レイヤ280は、管理転送ノードの測定エージェント205から測定マトリックスを受信し、これらの測定マトリックスを処理して、異なるペアの管理転送ノード間の接続品質を表す集約メッシュマトリックスを生成する。測定処理レイヤ280は、集合メッシュマトリックスをパス識別レイヤ282に提供する。集合メッシュマトリックスに基づいて、パス識別レイヤ282は、異なる企業データエンドポイント（例えば、異なる支店、企業データセンタ、SaaSプロバイダデータセンタおよび／またはリモートデバイス）を接続するために、仮想ネットワークを通る異なる所望のルーティングパスを識別する。次いで、このレイヤ282は、管理転送ノード150のクラウド転送要素235に配信される経路テーブル内のこれらのルーティングパスを提供する。

【００７７】

[00121] いくつかの実施形態では、データメッセージエンドポイントの各ペアのために識別されたルーティングパスは、最適化基準のセットに基づいて最適とみなされたルーティングパスである。例えば、それは、最速のルーティングパス、最短のルーティングパス、または、インターネットを最小限に使用するパスである。他の実施形態では、パス識別エンジンは、同じ2つのエンドポイント間の複数の異なるルーティングパスを（ルーティングテーブル内で）識別し、提供することができる。これらの実施形態では、管理転送ノード150のクラウド転送要素235は、次に、それらが実施しているQoS基準または他のランタイム基準に基づいて、パスの1つを選択する。いくつかの実施形態における各CFE235は、CFEから仮想ネットワークの出力ポイントまでのルーティングパス全体を受信せず、むしろパスの次ホップを受信する。

【００７８】

[00122] いくつかの実施形態では、パス識別レイヤ282は、集合メッシュマトリックス内の測定値を、それが実行するルーティングアルゴリズムへの入力として使用して、グローバルルーティンググラフを構築する。このグローバルルーティンググラフは、測定処理レイヤ280がいくつかの実施形態で生成する測定グラフの、集約され且つ最適化されたバージョンである。図3は、いくつかの実施形態において、コントローラ測定処理レイヤ280が生成する測定グラフ300の例を示す。このグラフは、AWS内のさまざまな管理転送ノード150 とGCP パブリッククラウド310 および320（すなわち、AWSおよびGCPのデータセンタ）の間のネットワーク接続を示している。図4Aは、いくつかの実施形態において、測定グラフ300からコントローラパス識別レイヤ282が生成するルーティンググラフ400の例を示す。

【００７９】

[00123] 図5は、コントローラパス識別レイヤが、コントローラ測定レイヤから受信した測定グラフからルーティンググラフを生成するために使用する処理500を示す。パス識別レイヤ282は、コントローラ測定レイヤから更新された測定グラフを繰り返し受信すると、この処理500を繰り返し実行する（例えば、新しい測定グラフを受信するたびに処理500を実行する、または新しい測定グラフを受信するたびにそれぞれのN回を実行する）。他の実施形態では、パス識別レイヤ282は、このプロセスを定期的に（例えば、12時間毎または24時間毎に）実行する。

【００８０】

[00124] 図示したように、パス識別レイヤは、最初に（505において）ルーティンググラフを測定グラフと同一であるように定義する（すなわち、管理転送ノードの同一ペア間に同一のリンクを有するようにする）。510において、プロセスは、測定グラフ300から不良リンクを除去する。不良リンクの例としては、過剰なメッセージ損失や信頼性の低いリンク（たとえば、過去15分間にメッセージ損失が2％を超えたリンクや、過去2分間にメッセージ損失が10％を超えたリンク）がある。図4Aは、測定グラフ300内のリンク302、304および306が、ルーティンググラフ400内で除外されることを示す。この図は、これらのリンクを破線で表すことによって、これらのリンクを除外する方法を示している。

【００８１】

[00125] 次に、処理500は、515において、いくつかの計算された値とプロバイダ固有

の値との加重された組合せとして、リンク重みスコア（コストスコア）を計算する。いくつかの実施形態では、重みスコアは、リンクの（1）計算された遅延値、（2）計算された損失値、（3）プロバイダネットワーク接続コスト、および（4）プロバイダコンピューティングコスト、の加重された組み合わせである。いくつかの実施形態では、プロバイダコンピューティングコストは、リンクによって接続された管理転送ノードがパブリッククラウドデータセンタ内のホストコンピュータ上で実行されるマシン（例えば、VMまたはコンテナ）であるとして考慮される。

【００８２】

[00126] 520において、プロセスは、仮想ネットワーク内のデータメッセージフローのための既知の送信元および宛先IPアドレス（例えば、企業エンティティによって使用されるSaaSプロバイダの既知のIP）をルーティンググラフに追加する。いくつかの実施形態では、プロセスは、可能性のあるメッセージフローのエンドポイントの各既知のIPアドレスを、そのエンドポイントに最も近いルーティンググラフのノード（例えば、MFNを表すノード）に追加する。そのような場合、いくつかの実施形態での処理は、遅延コストがゼロで損失コストがゼロのリンクを介して、そのような各エンドポイントが仮想ネットワークに接続されることを想定している。図4Bは、2つのSaaSプロバイダのための既知のIPを、これらのSaaSプロバイダのデータセンタに最も近いデータセンタにあるルーティンググラフの2つのノード402および404（2つのMFNを表す）に追加する例を示している。この例では、1つのノードがAWS パブリッククラウドにあり、もう1つのノードがGCP パブリッククラウドにある。

【００８３】

[00127] 代わりに、または、いくつかの実施形態では、処理500は、送信元および宛先エンドポイントを表すためにノードをこのグラフに追加し、これらのノードにIPアドレスを割り当て、これらの追加されたノードをルーティンググラフの他のノードに接続するリンク（例えば、パブリッククラウドのMFNを表すルーティンググラフのノード）に重み値を割り当てることによって、既知の送信元および宛先IPアドレスをルーティンググラフに追加する。フローの送信元および宛先エンドポイントがノードとして追加されると、パス識別エンジン282は、異なる送信元および宛先エンドポイント間の仮想ネットワークを通る異なる経路を識別するときに、これらのノードに到達するコスト（例えば、距離コスト、遅延コスト、および／または財務コストなど）を考慮することができる。

【００８４】

[00128] 図4Cは、2つのSaaSプロバイダを表すために、2つのノード412および414を図4Aのノードグラフ400に追加することによって生成されるルーティンググラフ410を示す。この例では、既知のIPアドレスがノード412および414に割り当てられ、これらのノードは、重みW1およびW2が割り当てられたリンク416および418を介してノード402および404（2つのMFNを表す）に接続される。このアプローチは、図4Bに示すアプローチに2つのSaaSプロバイダの既知のIPアドレスを追加するための代替アプローチである。

【００８５】

[00129] 図4Dは、より詳細なルーティンググラフ415を示す。このより詳細なルーティンググラフでは、追加のノード422および424が追加されて、AWSおよびGCPパブリッククラウド310および320にそれぞれ接続する既知のIPアドレスを持つ外部の企業コンピューティングノード（例えば、支店およびデータセンタ）を表す。これらのノード422／424の各々は、MFNを表すルーティンググラフノードの少なくとも1つに関連付けられる重み値Wiを有する、少なくとも1つのリンク426によって接続される。これらのノードのいくつか（例えば、いくつかの支店）は、同じMFNまたは異なるMFNへの複数のリンクで接続されている。

【００８６】

[00130] 次に、処理500は、525において、各MFNと、企業エンティティのデータメッセージフローのための仮想ネットワークの出力ロケーションとして機能することができる、それぞれの他のMFNとの間の最低コストパス（例えば、最短パスなど）を計算する。

いくつかの実施形態における出力MFNは、モバイルデバイス接続および出力インターネット接続のための候補ロケーションであるMFNと、外部の企業コンピューティングノード（例えば、支店、企業データセンタ、およびSaaSプロバイダデータセンタ）に接続されたMFNとを含む。いくつかの実施形態では、この計算は、異なるMFNのペア間の最短経路を識別する従来の最低コスト（例えば、最短経路）識別プロセスを使用する。

【００８７】

[00131] 各候補MFNペアに対して、最低コスト識別プロセスは、計算された重みスコア（すなわち、510で計算されたスコア）を使用して、MFNペアの間に複数のそのようなパスが存在するとき、最低スコアをもつパスを識別する。最低コストのパスを計算するためのいくつかの方法について、以下でさらに説明する。上述のように、パス識別レイヤ282は、いくつかの実施形態において、2つのMFNペア間の複数のパスを識別する。これは、異なる状況下で、クラウド転送要素235が異なるパスを使用することを可能にするためである。したがって、これらの実施形態では、処理500は、2つのMFNペア間の複数のパスを識別することができる。

【００８８】

[00132] 530において、プロセスは、525で識別される最低コストパスのいずれによっても使用されないMFNペア間のリンクをルーティンググラフから削除する。次に、535において、プロセスは、ルーティンググラフからクラウド転送要素235のためのルーティングテーブルを生成する。535において、プロセスは、これらのルーティングテーブルを管理転送ノードのクラウド転送要素235に配信する。535の後、プロセスは終了する。

【００８９】

[00133] いくつかの実施形態では、仮想ネットワークは、（1）エンティティのコンピューティングノード（例えば、支店、データセンタ、モバイルユーザなど）との外部セキュア接続と、（2）インターネットを介した第三者のコンピュータ（例えば、SaaSプロバイダサーバ）への外部接続との2種類の外部接続を有する。いくつかの実施形態では、仮想ネットワークの外部の送信元ノードおよび宛先ノードで終端する各データパスについて、最適な仮想ネットワークの入出力ロケーションを発見することによって、仮想ネットワークを最適化する。例えば、支店をSaaSプロバイダサーバ（例えば、salesforce.comサーバ）に接続するために、いくつかの実施形態は、支店を最適エッジMFN (例えば、支店または支店に最も近いネットワーク接続を有するMFN）に接続し、最適に配置されたSaaSプロバイダサーバ（例えば、支店のエッジMFNに最も近いか、SaaSプロバイダサーバに接続されたエッジMFNを介して支店のエッジMFNへの最速パスを有するSaaS）への最適なエッジMFNを識別する。

【００９０】

[00134] VPN接続を介してエンティティの各コンピューティングノード（支店、モバイルユーザなど）を最も近いMFNに関連付けるために、いくつかの実施形態では、仮想ネットワークプロバイダは、コンピューティングノードが接続するために、パブリッククラウドに1つ以上の権威(authoritative)ドメインネームサーバ（DNS）をデプロイする。いくつかの実施形態では、ある実施形態の企業コンピューティングノードが、仮想ネットワークプロバイダのMFNへのVPN接続を確立する（すなわち、VPN接続を初期化または再初期化する）必要があるたびに、コンピューティングノードは、まず、権威DNSサーバが企業コンピューティングノードに最も近いMFNとして識別するMFNの識別情報をこのサーバから得るために、このサーバを用いてその仮想ネットワークと関連付けられたアドレス（例えば、virtualnetworkX.net）を解決する。このMFNを識別するために、権威DNSサーバは、いくつかの実施形態において、MFN識別子（例えば、MFNのIPアドレス）を提供する。その後、企業コンピューティングノードは、この管理転送ノードへのVPN 接続を確立する。

【００９１】

[00135] 他の実施形態では、企業コンピューティングノードは、VNPのMFNへのVPN接続を確立する必要があるたびに、DNS解決を最初に実行しない（すなわち、特定のドメ

インのためのネットワークアドレスを最初に解決しない）。例えば、いくつかの実施形態では、企業コンピューティングノードは、対象のMFNが接続すべき最適なものであるかどうかを決定するために別のDNS解決を実行する前に、特定の期間（例えば、1日、1週間など）の間、DNS解決されたMFNを用いる。

【００９２】

[00136] DNS要求の送信元IPアドレスが、ノード自体ではなく、企業コンピューティングノードのローカルDNSサーバの送信元IPアドレスである場合、いくつかの実施形態では、権威DNSサーバは、企業コンピューティングノードに最も近いMFNではなく、ローカルDNSサーバに最も近いMFNを識別する。これに対処するために、いくつかの実施形態におけるDNS要求は、ドットによって連結および区切られた1つ以上の部分（ラベル）を含むドメイン名を用いて企業コンピューティングノードを識別する。ここで、これらの部分の1つは企業を識別し、他の部分は企業のコンピューティングノードを識別する。

【００９３】

[00137] いくつかの実施形態では、このドメイン名は、ドメイン名内の右ラベルから左ラベルへと階層が下がるドメインおよびサブドメインを規定する。右端の最初のラベルは特定のドメインを識別し、最初のラベルの左側の2 番目のラベルは企業エンティティを識別し、2 番目のラベルの左側の3 番目のラベルは、エンティティに複数の外部マシンロケーションがある場合のエンティティの外部マシンロケーションを識別する。例えば、いくつかの実施形態では、DNS要求は、企業コンピューティングノードを企業myCompanyのmyNodeとして識別し、アドレスmyNode.myCompany.virtualnetwork.netの解像度を求める。次いで、DNSサーバは、myNode識別子を使用して、企業コンピューティングノードがVPN接続を確立すべき入力MFNをより良く選択する。異なる実施形態では、myNode識別子は異なる方法で表現される。例えば、これは、IPアドレス、ロケーションの緯度／経度説明、GPS（地球測位システム）ロケーション、ストリートアドレスなどとしてアドレス指定されてもよい。

【００９４】

[00138] IPアドレスが適切にロケーションを反映している場合であっても、複数の潜在的な入力ルータが存在する可能性があり、たとえば、同じクラウド内の異なるデータセンタに属している場合や、同じ領域内の異なるクラウドに属している場合などである。このような場合、いくつかの実施形態における仮想ネットワークの権威サーバは、潜在的なMFNのCFE(例えば、C5、C8、C12）のIPのリストを送り返す。いくつかの実施形態における企業コンピューティングノードは、次に、リスト内の異なるCFEにpingして測定値（例えば、距離または速度測定値）を生成し、CFE候補のセットのなかで測定値を比較することによって、最も近いものを選択する。

【００９５】

[00139] さらに、企業コンピューティングノードは、企業エンティティの他のコンピューティングノードによって現在使用されているMFN を識別することによって、この選択の基礎とすることができる。例えば、いくつかの実施形態では、企業のコンピューティングノードは、企業の支店の多くがすでに所定のクラウドに接続されている場合、新しいコンピューティングノードが同じクラウドに接続するインセンティブを有し、これにより、処理、レイテンシー、およびドルに関してクラウド間コストを最小限に抑えるように、各MFNに接続コストを追加する。

【００９６】

[00140] 他の実施形態は、他のDNS解決技術を使用する。例えば、企業コンピューティングノード（例えば、支店、データセンタ、モバイルユーザなど）がDNS解決を行う必要があるたびに、企業コンピューティングノード（例えば、支店またはデータセンタのモバイルデバイスまたはローカルDNSリゾルバ）は、多数のエンティティの権威DNSリゾルバとして機能するDNSサービスプロバイダと通信する。いくつかの実施形態では、このDNSサービスプロバイダは、１つ以上のプライベートデータセンタに配置されたDNS解決マシンを有する一方、他の実施形態では、１つ以上のパブリッククラウドデータセンタの一

部である。

【００９７】

[00141] インターネットに直接接続するN個の管理転送ノードのうち、どれがSaaSプロバイダサーバに到達するために使用されるべきかを識別するために、いくつかの実施形態では、仮想ネットワーク（例えば、MFNを構成する入力MFNまたはコントローラクラスタ）は、N個の管理転送ノードから1つ以上の候補エッジMFNのセットを識別する。以下にさらに説明するように、いくつかの実施形態における各候補エッジMFNは、SaaSプロバイダサーバへの距離、ネットワーク接続速度、コスト、遅延および／または損失、ネットワーク計算コストなどの基準のセットに基づいて最適であるとみなされるエッジMFNである。

【００９８】

[00142] 最適なエッジポイントを識別することを支援するために、いくつかの実施形態のコントローラクラスタは、エンティティに対して、最も一般的なSaaSプロバイダおよび消費者ウェブの宛先およびそれらのIPアドレスサブネットのリストを保持する。このような宛先ごとに、コントローラクラスタは、１つ以上の最適なMFN（物理的な距離、ネットワーク接続速度、コスト、損失および／または遅延、計算コストなどによって再び判断される）を候補出力ノードとして割り当てる。次に、各出力MFN候補に対して、コントローラクラスタは、各可能性のある入力MFNから候補MFNまでの最良経路を計算し、それに応じて、インターネットSaaSプロバイダまたはウェブ宛先が正しい仮想ネットワークの次ホップノードに関連付けられるように、結果であるMFN内の次ホップテーブルを設定する。

【００９９】

[00143] サービス宛先には、（権威DNS サーバによって提供されるように） 複数のロケーションの複数のIP サブネットを介して到達できることが多いため、遅延を最小限に抑え、且つ負荷分散を提供するための、いくつかの潜在的な出力ノードが存在する。したがって、いくつかの実施形態では、コントローラクラスタは、各MFNについて最良のロケーションと出力ノードを計算し、それに応じて次ホップを更新する。また、SaaSプロバイダ（例えば、office365.com）に到達するための最良の出力ノードは、1つのパブリッククラウドプロバイダ（例えば、Microsoft Azure)を経由することもできるが、純粋に距離または接続速度からの最良の入力MFNは、別のパブリッククラウドプロバイダ（例えば、AWS）に存在するかもしれない。このような状況では、仮想ネットワークを出る前に別のクラウド（つまり、出力MFN が最適なパブリッククラウド）に移動するための遅延、処理、コストの点で最適ではないかもしれない。複数の候補エッジノードを提供することは、そのような状況において、最適エッジMFNおよび選択されたエッジMFNへの最適経路の選択を可能にする。

【０１００】

[00144] インターネットに接続するか、または企業エンティティの企業コンピューティングノードに接続する出力MFN への仮想ネットワークを介した最適なパスを識別するために、コントローラクラスタは、MFN 間の最適なルーティングパスを識別する。上述のように、いくつかの実施形態におけるコントローラクラスタは、直接接続されたMFNのペア間の各リンクにコストをかけることによって、例えば、推定された待ち時間と財務コストの加重和を反映するメトリックスコアに基づいて、任意の2つのMFN間の最良経路を識別する。レイテンシと財務コストは、いくつかの実施形態において、（1）リンク遅延測定、（2）推定されたメッセージ処理レイテンシ、（3）同一パブリッククラウドプロバイダの別のデータセンタへ、またはパブリッククラウドプロバイダのクラウド（例えば、別のパブリッククラウドプロバイダの別のパブリッククラウドデータセンタまたはインターネットへ）を出るために、特定のデータセンタからの発信トラフィックに対するクラウド料金、および（4）パブリッククラウド内のホストコンピュータ上で実行されるMFNに関連する推定メッセージ処理コストを含む。

【０１０１】

[00145] これらのペアワイズリンクの計算コストを使用して、コントローラクラスタは

、ルーティングパスによって使用される個々のペアワイズリンクのコストを集計することによって、これらのペアワイズリンクのうちの1つ以上を使用する各ルーティングパスのコストを計算することができる。上述のように、コントローラクラスタは、次に、ルーティングパスの計算コストに基づいてそのルーティンググラフを定義し、定義されたルーティンググラフに基づいて、MFNのクラウドルータの転送テーブルを生成する。また、上述のように、コントローラクラスタは、これらのコスト計算、グラフ生成、および転送テーブルの更新および配信動作を繰り返し、定期的に（例えば、12時間毎、24時間毎など）またはMFNの測定エージェントから測定更新を受信すると実行する。

【０１０２】

[00146] MFNのCFE C_iの転送テーブルが次ホップMFNのCFE C_jを指すときはいつでも、CFE C_iはC_jを近いとみなす。いくつかの実施形態では、CFE C_jは、CFE C_jへのセキュアでアクティブに維持されたVPNトンネルを確立する。いくつかの実施形態における安全なトンネルは、カプセル化されたデータメッセージのペイロードを暗号化することを要するトンネルである。また、いくつかの実施形態では、トンネルは、他のエンドポイントにキープアライブ(keep alive)信号を送信するトンネルの一方または両方のエンドポイントによって能動的に維持される。

【０１０３】

[00147] 他の実施形態では、CFEは、セキュアでアクティブに維持されているVPNトンネルを確立しない。例えば、いくつかの実施形態では、CFE間のトンネルは、キープアライブ信号の送信を通じて能動的に監視されない静的トンネルである。また、いくつかの実施形態では、CFE間のこれらのトンネルは、それらのペイロードを暗号化しない。いくつかの実施形態では、CFEペアの間のトンネルは、2つのカプセル化ヘッダを含み、データメッセージが仮想ネットワークに出入りする（すなわち、パブリッククラウドに出入りする）ためのテナントIDと入力CFEを識別する内部ヘッダと、入力CFEから出力CFEまでゼロ以上のCFEを通過するための送信元および宛先ネットワークアドレス（例えば、IPアドレス）を特定する外部カプセル化ヘッダとを有する。

【０１０４】

[00148] いくつかの実施形態では、内部トンネルに加えて、仮想ネットワークは、上述のように、VPNトンネルを使用して企業コンピューティングノードをエッジMFNに接続する。したがって、セキュアトンネルを使用してCFEと接続する実施形態では、データメッセージは、完全にセキュアなVPNパスを使用して仮想ネットワークを通過する。

【０１０５】

[00149] 仮想ネットワークデータメッセージは、仮想ネットワーク内のカプセル化を使用して転送されるため、いくつかの実施形態では、仮想ネットワークは、テナントの異なるプライベートネットワークによって使用されるプライベートアドレスとは異なる自身のユニークなネットワークアドレスを使用する。他の実施形態では、仮想ネットワークは、それが定義されているパブリッククラウドのプライベートおよびパブリックネットワークアドレス空間を使用する。さらに別の実施形態では、仮想ネットワークは、そのコンポーネント（例えば、そのMFN、CFE、および／またはサービスの一部）のために、パブリッククラウドのプライベートおよびパブリックネットワークアドレス空間をそのコンポーネントの他のために使用する一方で、その自身のユニークなネットワークアドレスのいくつかを使用する。

【０１０６】

[00150] また、いくつかの実施形態では、仮想ネットワークは、独自のプロトコルを備えたクリーンスレート通信プラットフォームを使用する。データメッセージが全体としてソフトウェアMFNルータ（例えば、ソフトウェアCFE）を介して転送される実施形態では、仮想ネットワークは、長距離のエンドツーエンド接続のための最適化されたレート制御を提供することができる。これは、いくつかの実施形態では、各MFN 150でTCP最適化プロキシエンジン220を動作させることによって達成される。（例えば、HTTPSを用いる）TCP自体が破られない他の実施形態では、これは、プロキシエンジン220が、TCP受信

ウィンドウおよびACK操作を共に有するフロー毎の中間バッファリングを用いて、レート制御をセグメント化することによって達成される。

【０１０７】

[00151] クリーンスレート(clean-slate)の性質により、いくつかの実施形態では、仮想ネットワークは、より良いサービスを提供するために、そのコンポーネントの多くを最適化する。例えば、いくつかの実施形態では、仮想ネットワークは、仮想ネットワークを介してルーティングされるプレミアム帯域幅保証のVPNセットアップをサポートするために、マルチパスルーティングを用いる。いくつかの実施形態では、このようなVPNは、ATM／MPLSルーティングに類似した各MFN内の状態データを含み、それらの確立および除去は、集中的に制御される。いくつかの実施形態は、（パケットペアまたは類似のプロセスを介して）直接測定することにより、または、リンクの所与の容量を有することにより、そして既にこのリンクを介して送信されたトラフィックをこの容量から減少させることにより、送信リンク当たりの利用可能な帯域幅を識別する。

【０１０８】

[00152] いくつかの実施形態は、制約としてリンクの残留帯域幅を使用する。例えば、リンクが利用可能な帯域幅の少なくとも2 Mbpsを有していない場合、いくつかの実施形態のコントローラクラスタは、任意の宛先への最低コストパス（例えば、最短パス）を計算するために使用されるリンクのセットからリンクを除去する（例えば、グラフ400のようなルーティンググラフからリンクを除去する）。このリンクの除去後もエンドツーエンド経路が未だ使用可能である場合、新らたなVPN はこの新らたな経路を介してルーティングされる。VPN を除去すると、使用可能なキャパシティを特定のリンクに戻すことができる。これにより、このリンクを最低コストパス（最短パスなど） の計算に含めることができる。いくつかの実施形態は、例えば、MPTCP (マルチパスTCP）を使用するなど、複数パスにわたるトラフィックのロードバランシングのような、複数パスルーティングのための他のオプションを用いる。

【０１０９】

[00153] いくつかの実施形態は、仮想ネットワーク内の2つの分離したパス（例えば、最大分離したパス）を介して、入力MFNから出力MFNへのトラフィックを複製するために、パスの並列性および安価なクラウドリンクを利用することにより、プレミアム顧客に対してより良いサービスを提供する。このアプローチでは、到着した最も早いメッセージは受け入れられ、後のメッセージは破棄される。このアプローチは、出力処理の複雑さを増加させる代わりに、仮想ネットワークの信頼性を増加させ、遅延を低減する。いくつかのそのような実施形態では、順方向誤り訂正（FEC）技法を使用して、信頼性を高める一方で、重複トラフィックを低減する。クリーンスレートの性質により、いくつかの実施形態の仮想ネットワークは、アプリケーションレイヤの最適化（例えば、重複排除およびキャッシュ動作）およびセキュリティ最適化（例えば、暗号化、DPI (ディープパケット検査）およびファイアウォール処理の追加）のような、他の上位レイヤの最適化を行う。

【０１１０】

[00154] いくつかの実施形態の仮想ネットワークは、エニーキャストメッセージングを使用することによって仮想ネットワーク設定をさらに改善するための、クラウドプロバイダとのコラボレーションについて説明する。例えば、全てのMFNが同じ外部IPアドレスを取得する場合、いくつかの実施形態では、エニーキャスト接続を使用して、任意の新しい企業コンピューティングノードを最適エッジノード（例えば、最も近いエッジノード）に接続することは容易である。同様に、SaaSプロバイダは、このIPアドレスを取得し、最適なMFN (例えば、最も近いMFN）に接続することができる。

【０１１１】

[00155] 上述したように、異なる実施形態は、異なるタイプのVPN接続を使用して、企業コンピューティングノード（例えば、支店およびモバイルデバイス）を、企業エンティティの仮想ネットワークを確立するMFNに接続する。いくつかの実施形態は、これらのVPN接続を設定するためにIPsecを使用する。図6は、いくつかの実施形態のIPsecデータ

メッセージフォーマットを示す。具体的には、この図は、企業のコンピューティングノードのマシンによって生成されたデータメッセージ605の元のフォーマットを示し、データメッセージ605がカプセル化された後に（例えば、企業のコンピューティングノードまたはMFN）、IPsecトンネル（例えば、MFNまたは企業のコンピューティングノード）を介して送信するためのIPsecカプセル化データメッセージ610を示す。

【０１１２】

[00156] この例では、IPsecトンネルは、ESPトンネルモード（ポート50）を用いて設定される。図示するように、このモードは、この例においてIPヘッダ内のTCPプロトコル識別子をESPプロトコル識別子で置き換えることによって設定される。ESPヘッダは、メッセージ615（すなわち、ヘッダ620およびペイロード625）の開始を識別する。メッセージ615は、IPsecカプセル化データメッセージの受信装置によって（例えば、MFNのIPsecゲートウェイによって）認証されなければならない。ペイロード625の開始は、メッセージ615の次のフィールド622の値によって識別される。また、ペイロード625は暗号化される。このペイロードは、元のデータメッセージ605のIPヘッダ、TCPヘッダおよびペイロード、ならびに次のフィールド622を含むパディングフィールド630を含む。

【０１１３】

[00157] いくつかの実施形態では、MFNのIPsecゲートウェイのそれぞれは、同一または異なる仮想ネットワークテナント（例えば、同一企業または異なる企業）のための複数のIPsec接続を扱うことができる。従って、いくつかの実施形態におけるMFNのIPsecゲートウェイ（例えば、ゲートウェイ230）は、トンネルID、テナントID(TID)、および企業コンピューティングノードのサブネットに関して各IPsec接続を識別する。いくつかの実施形態では、テナントの異なる企業ノード（例えば、異なる支店）は、重複するIPサブネットを持たない（RFC 1579による）。いくつかの実施形態におけるIPsecゲートウェイは、各IPsecトンネルID（IPsecトンネルヘッダに含まれる）をテナントIDにマッピングするテーブルを有する。IPsecゲートウェイが処理するように設定されている特定のテナントの場合、IPsecゲートウェイは、MFN とそのクラウド転送要素によって確立された仮想ネットワークに接続する、そのテナントのすべてのサブネットのマッピングを有する。

【０１１４】

[00158] 第１のパブリッククラウドデータセンタの入力側の第１のMFN が、テナントID に関連付けられた、第２のパブリッククラウドデータセンタの出力側の第２のMFN に接続する宛先（ブランチやデータセンタのサブネット、SaaSプロバイダなど）に向けられたデータメッセージをIPsecトンネルを介して受信すると、第１のMFNのIPsec ゲートウェイはIPsecトンネルヘッダを削除する。いくつかの実施形態では、第1のMFNのCFEは、メッセージが、入力側の第1のMFNから出力側の第2のMFNまで、直接または1つ以上の他の中間MFNを介して、パスを通過することを可能にする2つのカプセル化ヘッダでメッセージをカプセル化する。第１のMFN のCFE は、コントローラが構成したルーティングテーブルを使用して、このパスを識別する。

【０１１５】

[00159] 上述のように、いくつかの実施形態における2つのカプセル化ヘッダは、（1）カプセル化されたデータメッセージが仮想ネットワークのMFNを通過して出力MFNのCFEに到達することを可能にする次ホップMFNのCFEを指定する外部ヘッダと、（2）テナントIDと、仮想ネットワークに出入りするデータメッセージのためのMFNを識別する入力及び出力MFNのCFEを指定する内部ヘッダとを含む。

【０１１６】

[00160] 具体的には、いくつかの実施形態では、内部カプセル化ヘッダは、第２の出力MFNのCFEの宛先IPアドレスと、第１の入力MFNのCFEの送信元IPアドレスとを有する有効なIPヘッダを含む。このアプローチは、標準IPルータソフトウェアをMFNのすべてのCFEで使用することを可能にする。カプセル化には、テナントID（たとえば、顧客CID）がさらに含まれる。メッセージが第２の出力MFN のCFE に到着すると、メッセージはカプセル化解除され、第２のMFN によって宛先に送信される（たとえば、第２のMFN のIPse

cゲートウェイによって、テナントIDとメッセージの宛先サブネットに関連付けられた別のIPsecトンネルを介して宛先に送信される）。

【０１１７】

[00161] クラウドプロバイダによっては、マシンが送信元IP を「なりすまし」することを禁止したり、TCPおよびUDPトラフィックに他の制限を課したりすることがある。このような制限に対処するために、いくつかの実施形態は、1つ以上の経路によって使用されるMFNの隣接するペアを接続するために、外部ヘッダを使用する。いくつかの実施形態におけるこのヘッダは、送信元および宛先IPアドレスおよびUDPプロトコルパラメータを指定するUDPヘッダである。いくつかの実施形態では、入力MFNのCFEは、そのIPアドレスを外部ヘッダの送信元IPアドレスとして指定し、一方、次のMFNのCFEホップのIPアドレスを外部ヘッダの宛先IPアドレスとして指定する。

【０１１８】

[00162] 出力MFNのCFEへのパスに1つ以上の中間MFNのCFE が含まれる場合、中間CFEは、受信したダブルカプセル化メッセージの外部ヘッダの送信元IPアドレスをそのIPアドレスに置き換える。また、内部ヘッダの宛先IPアドレスを使用して、ルーティングテーブルで経路ルックアップを実行し、内部ヘッダの宛先IPアドレスへのパス上にある次ホップのMFNのCFEの宛先IPアドレスを識別する。次に、中間CFEは、外部ヘッダ内の宛先IPアドレスを、その経路テーブルルックアップを介して識別したIPアドレスで置き換える。

【０１１９】

[00163] ダブルカプセル化されたデータメッセージが出力MFN のCFE に到達すると、CFEは、内部ヘッダで宛先IPアドレスを取得した場合に、この宛先IPアドレスがデータメッセージの出力ノードであると判断し、この宛先IPアドレスがそれに属していると判断する。このCFEは、次に、データメッセージから2つのカプセル化ヘッダを削除し、それを宛先に送信する（例えば、データメッセージの元のヘッダのテナントIDと宛先IPアドレスまたはサブネットに関連付けられた別のIPsecトンネルを介して、MFNのIPsecゲートウェイを通って宛先に送信する）。

【０１２０】

[00164] 図7は、いくつかの実施形態の2つのカプセル化ヘッダの例を示し、図8は、いくつかの実施形態においてこれら2つのヘッダがどのように使用されるかを説明する例を示す。以下の説明では、内部ヘッダはテナントヘッダと呼ばれる。これには、テナントの企業コンピューティングエンドノードに接続されている仮想ネットワークの入出力ノードのID とともにテナントID が含まれるためである。外部ヘッダは、データメッセージが入力および出力MFNの CFEsの間の仮想ネットワークを通るパスを移動するときに、仮想ネットワークを通る次ホップを識別するために使用されるため、以下のVNホップトンネルヘッダと呼ばれる。

【０１２１】

[00165] 図7は、元のヘッダ755およびペイロード760を有する元のデータメッセージ750をカプセル化する、VNホップトンネルヘッダ705およびテナントトンネルヘッダ720を示す。図示されるように、いくつかの実施形態におけるVNホップトンネルヘッダ705は、UDPヘッダ710およびIPヘッダ715を含む。いくつかの実施形態におけるUDPヘッダは、UDPプロトコルに従って定義される。いくつかの実施形態において、VNホップトンネルは標準のUDPトンネルであるが、他の実施形態では、このトンネルは独自のUDPトンネルである。さらに他の実施形態では、このトンネルは、標準または独自のTCPトンネルである。いくつかの実施形態におけるトンネルヘッダ705は、そのペイロードを暗号化した暗号化済みのものであり、他の実施形態においては、それは、暗号化されていないトンネルである。

【０１２２】

[00166] 以下にさらに説明するように、いくつかの実施形態におけるトンネルヘッダ705は、オーバーレイVNPネットワークを定義するために使用され、各MFNのCFEによって、アンダーレイパブリッククラウドネットワークを介して次ホップMFNのCFEに到達す

るために使用される。したがって、トンネルヘッダ705のIPヘッダ715は、VNPトンネルによって接続された第1および第2の隣接MFNの第1および第2のCFEの送信元および宛先IPアドレスを識別する。場合によっては（たとえば、次ホップの宛先MFN が、送信元MFNとは異なるパブリッククラウドベンダーの別のパブリッククラウドにある場合）、送信元IPアドレスと宛先IPアドレスは、MFNを含むパブリッククラウドデータセンタによって使用されるパブリックIPアドレスである。また、送信元と宛先のMFNのCFE が同じパブリッククラウドに属している場合、送信元と宛先のIPアドレスは、パブリッククラウドでのみ使用されるプライベートIPアドレスにすることができる。或いは、このような場合でも、送信元と宛先のIPアドレスがパブリッククラウドベンダのパブリックIPアドレスであり得る。

【０１２３】

［00167］図7に示すように、テナントトンネルヘッダ720は、IPヘッダ725、テナントIDフィールド730および仮想回線ラベル735を含む。テナントトンネルヘッダ720は、データメッセージを出力MFNの出力CFEに転送するための次ホップを識別するために、入力ホップCFEの後の各ホップCFEによって使用される。したがって、IPヘッダ725は、入力CFEのIPアドレスである送信元IPアドレスと、出力CFEのIPアドレスである宛先IPアドレスとを含む。VNホップヘッダ705の送信元および宛先IPアドレスと同様に、テナントヘッダ720の送信元および宛先IPアドレスは、1つのパブリッククラウドプロバイダのプライベートIPアドレス（データメッセージが1つのパブリッククラウドプロバイダのデータセンタのみを通過する経路を通過する場合）または1つ以上のパブリッククラウドプロバイダのパブリックIPアドレス（たとえば、データメッセージが2つ以上のパブリッククラウドプロバイダのデータセンタを通過する経路を通過する場合）のいずれかであることができる。

【０１２４】

［00168］テナントヘッダ720のIPヘッダは、いくつかの実施形態では、任意の標準ソフトウェアルータおよびIPルーティングテーブルを使用してルーティングされることができる。テナントIDフィールド730は、テナントIDを含み、これは、テナントをユニークに識別するために入出力MFNで使用することができる、ユニークなテナント識別子である。いくつかの実施形態における仮想ネットワークプロバイダは、プロバイダのテナントである異なる企業エンティティに対して異なるテナントIDを定義する。VCLフィールド735は、いくつかの実施形態が、ネットワークを通してメッセージを転送するための代替の方法（非IPベースの方法）を提供するために使用する、オプションのルーティングフィールドである。いくつかの実施形態では、テナントトンネルヘッダ720は、GUE（汎用UDPカプセル化）ヘッダである。

【０１２５】

［00169］図8は、いくつかの実施形態において、これら2つのトンネルヘッダ705および710がどのように使用されるかを示す例を表している。この例では、データメッセージ800は、企業の第1の支店805内の第1のマシン802（例えば、第1のVM）から、企業の第2の支店810内の第2のマシン804（例えば、第2のVM）に送信される。2つのマシンは、2つの異なるサブネット、10．1．0．0 と10．2．0．0 にある。第１のマシンのIP アドレスは10．1．0．17、第２のマシンのIP アドレスは10．2．0．22 である。この例では、第1の支店805は、第1のパブリッククラウドデータセンタ830内の入力MFN 850に接続し、第2の支店810は、第2のパブリッククラウドデータセンタ838内の出力MFN85に接続する。また、この例では、第1および第2のパブリッククラウドデータセンタの入力MFN850および出力MFN855は、第3のパブリッククラウドデータセンタ836の中間MFN857を介して間接的に接続されている。

【０１２６】

［00170］図示されるように、マシン802からのデータメッセージ800は、最初の支店805を入力MFN850に接続するIPsecトンネル870に沿って入力MFN850に送られる。このIPsecトンネル870は、第1の支店のIPsecゲートウェイ848と、入力MFN850のIPsec

ゲートウェイ852との間に確立される。このトンネルは、データメッセージ800をIPsecトンネルヘッダ806でカプセル化することによって確立される。

【０１２７】

[00171] MFN850のIPsecゲートウェイ852は、データメッセージをカプセル化解除し（すなわち、IPsecトンネルヘッダ806を除去し）、カプセル化解除されたメッセージをこのMFNのCFE832に直接、または1つ以上のミドルボックスサービスマシン（例えば、図2のマシン210のようなファイアウォールマシン）を介して渡す。このメッセージを渡す際に、いくつかの実施形態では、IPsecゲートウェイまたはMFN850の他のモジュールは、メッセージをIPsecトンネルのトンネルIDおよび企業のテナントIDと関連付ける。このテナントIDは、仮想ネットワークプロバイダのレコード内の企業を識別する。

【０１２８】

[00172] 関連付けられたテナントIDおよび／またはIPsecトンネルIDに基づいて、入力MFN850のCFE832は、異なるパブリッククラウドデータセンタ内のMFNによって確立された仮想ネットワークを介して、メッセージの宛先のマシンのサブネット（すなわち、第2の支店810）への経路を識別する。たとえば、CFE832は、テナントIDおよび／またはIPsecトンネルIDを使用して、企業のルーティングテーブルを識別する。このルーティングテーブルでは、CFE832は、受信メッセージの宛先IPアドレス10．2．0．22を使用して、パブリッククラウドデータセンタ838の出力MFN 855のCFE 853を、データメッセージ800の宛先出力転送ノードとして識別するレコードを識別する。いくつかの実施形態では、識別されたレコードは、第2の支店810のサブネット10．2．0．0／16全体をMFN855のCFE 853にマップする。

【０１２９】

[00173] 出力CFE853を識別した後、入力MFN850のCFE832は、受信したデータメッセージを、そのIPヘッダ725内に入力CFE 832の送信元IPおよび出力CFE853の宛先IPを含むテナントトンネルヘッダ860でカプセル化する。いくつかの実施形態では、これらのIPアドレスは、パブリックIPアドレス空間で定義される。トンネルヘッダ860はまた、入力MFN850においてデータメッセージに関連付けられたテナントIDを含む。上述したように、このトンネルヘッダは、いくつかの実施形態において、VCLヘッダ値も含む。

【０１３０】

[00174] いくつかの実施形態では、入力CFE832は、出力CFE853への所望のCFEルーティングパス上にある次ホップMFNも識別する。いくつかの実施形態では、入力CFE832は、出力CFE 853の宛先IPアドレスを使用して、そのルーティングテーブル内のこの次ホップCFEを識別する。この例の次ホップMFNのCFE は、第３のパブリッククラウドデータセンタ836の第３のMFN857のCFE856である。

【０１３１】

[00175] 次ホップMFNのCFEを識別した後、入力MFNのCFEは、カプセル化されたデータメッセージ800をVNホップ、第2のトンネルヘッダ862でカプセル化する。このトンネルヘッダは、メッセージを次ホップCFE856にルーティングできるようにする。この外部ヘッダ862のIPヘッダ715において、入力MFN のCFE832は、送信元および宛先IPアドレスを、入力CFE832の送信元IPおよび中間CFE856の宛先IPとして指定する。また、いくつかの実施形態において、そのレイヤ4プロトコルをUDPとして指定する。

【０１３２】

[00176] 第３のMFN857のCFE856がダブルカプセル化されたデータメッセージを受信すると、CFE856は、VNホップ、第２のトンネルヘッダ862を取り除き、MFN855のCFE853の宛先IPアドレスをテナントヘッダ860から取り出す。このIPアドレスはCFE856に関連付けられていないため、データメッセージは宛先に到達するために別のMFNを通過しなければならない。したがって、CFE856は、取り出された宛先IPアドレスを使用して、次ホップMFNのCFE853を識別するルーティングテーブル内のレコードを識別する。次に、この変更は外部ヘッダ705を用いたデータメッセージを再カプセル化し、そのIPヘッダ715内の送信元および宛先IPアドレスを、それ自身のIPアドレスおよびMFNのCFE 85

3の宛先IPアドレスとして指定する。次に、CFE856は、パブリッククラウドデータセンタ836および838の中間ルーティングファブリックを介して、ダブルカプセル化データメッセージ800を出力CFE853に転送する。

【０１３３】

[00177] カプセル化データメッセージを受信した後、出力CFE 853は、内部ヘッダ860内の宛先IPアドレスを検索して、この宛先IPアドレスがそれに属していると判定した場合、カプセル化されたメッセージがそれに向けたものであると判定する。出力CFE853は、データメッセージ800からカプセル化ヘッダ860および862の両方を除去し、データメッセージの元のヘッダ内の宛先IPアドレスを取り出す。この宛先IPアドレスは、第２の支店のサブネット内の第２のマシン804のIPアドレスを示す。

【０１３４】

[00178] 除去されたテナントトンネルヘッダ860のテナントIDを使用して、出力CFE853 は、検索する正しいルーティングテーブルを識別し、受信データメッセージの元のヘッダ値から抽出された宛先IP アドレスに基づいてこのルーティングテーブルを検索する。この検索から、出力CFE853は、データメッセージを宛先に転送するために使用するIPsec接続を識別するレコードを識別する。次に、IPsecコネクション識別子とともにデータメッセージを第2のMFNのIPsecゲートウェイ858に提供し、次にこのメッセージをIPsecトンネルヘッダ859でカプセル化し、それを第2の支店810のIPsecゲートウェイ854に転送する。次に、ゲートウェイ854は、IPsecトンネルヘッダを除去し、データメッセージを宛先マシン804に転送する。

【０１３５】

[00179] ここで、図9から図15を参照して、より詳細なメッセージ処理の例について説明する。これらの例では、各テナントIPsec インタフェースがVNPトンネルと同じローカルパブリックIPアドレス上にあると想定されている。したがって、いくつかの実施形態のインタフェースは、単一のVRF(仮想ルーティングおよび転送）ネームスペースに接続される。このVRFネームスペースは、以下ではVNPネームスペースと呼ぶ。

【０１３６】

[00180] 図9から図11は、テナントの2つの異なる外部マシンのロケーション（例えば、支店、データセンタなど）にある2つのコンピューティングデバイス間で送信されるメッセージを受信したときに、入力MFN、中間MFN、および出力MFNによってそれぞれ実行されるメッセージ処理プロセス900-1100を示す。いくつかの実施形態では、コントローラクラスタ160は、各CFEがテナントの異なるデータメッセージフローの入力、中間、および出力CFEとして動作するように各MFNのCFEを構成するが、そのような各CFEは、テナントの異なるデータメッセージフローのための入力、中間、および出力CFEとしてサービスを提供する候補である。

【０１３７】

[00181] プロセス900‐1100は、図8および図12の2つの例を参照して、以下に説明される。上述のように、図8は、データメッセージが中間MFNを通過して出力MFNに到達する場合の例を示している。図12は、入力MFNと出力MFNの間に中間MFNを含まない例を示す。具体的には、2つの支店が直接接続された2つのMFN1250および1255を有する2つのパブリッククラウドデータセンタ1230および1238に接続される場合の、第1の支店1205内の第1のデバイス1202から第2の支店1220内の第2のデバイス1210に送信されるデータメッセージ1200を示す。図示されるように、これらの例のMFNのCFE1232および1253は、各MFNに関連付けられたルーティング動作を実行する。

【０１３８】

[00182] いくつかの実施形態では、入力MFN850および1250の入力CFE (例えば、入力CFE832または1232）は、プロセス900を実行する。 図9に示すように、入力プロセス900は、受信データメッセージ内のIPsecトンネル（例えば、806または1206）の識別子に基づいて、（905において）最初にテナントルーティングコンテキストを識別することによって開始する。いくつかの実施形態では、IPsecゲートウェイまたは他のMFNモジ

ュールは、マッピングテーブルにIPsecトンネルIDに対するテナントIDを格納する。データメッセージが特定のIPsecトンネルに沿って受信されると、IPsecゲートウェイはIPsecトンネルIDを抽出する。このIPsecトンネルID は、このゲートウェイまたは別のMFNモジュールが、そのマッピングテーブルを参照することによって、関連するテナントIDを識別するために使用する。テナントID を識別することで、プロセスは、使用するテナントルーティングテーブルまたはVRFネームスペースのテナント部分を識別する。

【０１３９】

[00183] 910で、プロセスは、このデータメッセージを受信することを考慮して、識別されたIPsecトンネルのRX (受信）カウンタをインクリメントする。次に、915において、プロセスは、識別されたテナントルーティングコンテキスト（例えば、VRF名前空間のテナントの部分）において、パブリッククラウドデータセンタを介して構築されたテナントの仮想ネットワークを出るための出力インタフェースのIPアドレスを識別するために、経路ルックアップ（例えば、最長のプレフィックスの一致、LPM、ルックアップ）を実行する。ブランチ・ツー・ブランチの例では、出力インタフェースは、宛先ブランチに接続されたMFNの出力CFE (CFE 853 または1253など） のIPアドレスである。

【０１４０】

[00184] 920において、プロセスは、受信したデータメッセージにテナントトンネルヘッダ（例えば、ヘッダ860または1260）を追加し、このトンネルヘッダの送信元および宛先IPアドレスとして、入力CFEの送信元IPアドレス（例えば、入力CFE832または1252）および出力CFEの宛先IPアドレス（例えば、出力CFE853または1253）を埋め込む。テナントヘッダでは、テナントID(905 で識別）もテナントヘッダに格納される。920で、プロセスはテナントヘッダの外部にVNホップトンネルヘッダ（例えばヘッダ862または1262）を追加し、そのIPアドレスをこのヘッダの送信元IPアドレスとして格納する。このプロセスはまた、（920で）VNPトンネルヘッダのUDPパラメータ（例えばUDPポート）を規定する。

【０１４１】

[00185] 次に、925において、プロセスは、このデータメッセージの送信を考慮するために、テナントのためのVN送信カウンタをインクリメントする。930において、プロセスは、このデータメッセージの次ホップインタフェースを識別するために、識別されたVNPルーティングコンテキスト（例えば、VRF名前空間のVNPの部分）において、経路ルックアップ（例えば、LPMルックアップ）を実行する。いくつかの実施形態では、この経路ルックアップは、（例えば、VRF名前空間のVNPの部分における）出力CFEの宛先IPに少なくとも部分的に基づいたLPMルックアップである。

【０１４２】

[00186] 935で、プロセスは、次ホップ出力インタフェースが入力CFEのローカルインタフェース（例えば、物理または仮想ポート）であるかどうかを判定する。もしそうなら、プロセスは（937で）VNホップ外部トンネルヘッダの宛先IPアドレスを915で識別された出力インタフェースIPアドレスとして定義する。次に、940において、プロセスは、宛先の出力CFEに転送できるように、そのローカルインタフェースにダブルカプセル化データメッセージを提供する。940の後、プロセス900は終了する。

【０１４３】

[00187] 図12は、入力CFE1232が第1の支店1205のデバイス1202から受信するデータメッセージ1200の動作905から940の例を示す。図示されるように、このCFEのMFN1250は、IPsecゲートウェイ1252において、このデータメッセージを、第1の支店1205のIPsecゲートウェイ1248からのIPsecカプセル化メッセージとして受信する。受信CFE1232は、受信したメッセージ1200を（IPsecヘッダがIPsecゲートウェイ1252によって除去された後に）VNホップトンネルヘッダ1262およびテナントトンネルヘッダ1260でカプセル化し、このダブルカプセル化されたメッセージをパブリッククラウド1238のMFN1255の出力CFE1253に転送する。図示されるように、この例では、トンネルヘッダ1260および1262の両方の送信元および宛先IPアドレスは同一である。これらの2セットのI

Pアドレスが同一である場合、CFE 856のような、介在するCFEを介してデータメッセージがルーティングされない場合、いくつかの実施形態は、外部IPヘッダ1262を使用しない。

【０１４４】

[00188] プロセスが（935で）次ホップ出力インタフェースが入力CFEのローカルインタフェースではなく、むしろ別のルータの宛先IPアドレスであると判定する場合、プロセスは（945で）VNホップトンネルヘッダに、次ホップ中間CFE (例えば、中間CFE856）の宛先IPアドレスをVNホップトンネルヘッダの宛先IPアドレスとして埋め込む。

【０１４５】

[00189] 次に、950において、プロセスは、識別されたVNPルーティングコンテキスト（例えば、VRFネームスペースのVNPの部分）において、別の経路ルックアップ（例えば、LPMルックアップ）を実行する。今回、ルックアップは、VNPトンネルヘッダで識別される中間CFEのIPアドレスに基づいている。中間CFE (例えば、CFE856）は、入力CFE (例えば、CFE832）のための仮想ネットワークにおける次ホップCFEであるので、ルーティングテーブルは、中間CFEに送られるデータメッセージのためのローカルインタフェース（例えば、ローカルポート）を識別する。したがって、VNPルーティングコンテキストにおけるこのルックアップは、受信CFEが（950において）ダブルカプセル化されたメッセージを提供するローカルインタフェースを識別する。その後、プロセスは、このデータメッセージの送信を考慮して、VN中間カウンタを（955で）インクリメントする。955の後、プロセスは終了する。

【０１４６】

[00190] 図10は、出力MFNのCFE(例えば、CFE853または1253）が、MFNに接続された企業コンピューティングノード（例えば、支店、データセンタ、リモートユーザのロケーション）に転送されるべきデータメッセージを受信したときに、いくつかの実施形態において実行するプロセス1000を示す。図示されるように、プロセスは、最初に、仮想ネットワークに関連付けられたインタフェース上でデータメッセージを受信する（1005で）。このメッセージは、VNホップトンネルヘッダ（ヘッダ862または1262など）とテナントトンネルヘッダ（ヘッダ860または1260など）でカプセル化される。

【０１４７】

[00191] 1010で、プロセスは、VNホップトンネルヘッダの宛先IPアドレスがそのCFEの宛先IPアドレスであると判定する（CFE853または1253のIPアドレスなど）。次に、1015で、プロセスは2つのトンネルヘッダを除去する。その後、プロセスは、除去されたテナントトンネルヘッダからテナントIDを（1020で）取得する。受信したデータメッセージを考慮するために、CFEは、抽出されたテナントID で指定されたテナントに対して保持しているRX (受信）カウンタを（1025で）インクリメントする。

【０１４８】

[00192] 次に、1030において、プロセスは、このデータメッセージの次ホップインタフェースを識別するために、識別されたテナントルーティングコンテキスト（すなわち、1020で抽出されたテナントIDによって識別されたテナントのルーティングコンテキスト）において、経路ルックアップ（例えば、LPMルックアップ）を実行する。プロセスは、いくつかの実施形態において、受信データメッセージの元のヘッダ（例えば、ヘッダ755）内の宛先IPアドレスに基づいて、このルックアップを実行する。このルックアップによって識別されるレコードから、プロセス1000は、データメッセージが宛先に送られる必要があるIPsecインタフェースを識別する。従って、プロセス1000は、カプセル化解除され受信されたデータメッセージを、そのMFNのIPsecゲートウェイ（例えば、ゲートウェイ858または1258）に送信する。

【０１４９】

[00193] 次に、このゲートウェイは、IPsecトンネルヘッダ（例えば、トンネルヘッダ859または1259）でデータメッセージをカプセル化し、それを宛先の企業コンピューティングノード（例えば、宛先支店）のゲートウェイ（例えば、ゲートウェイ854または12

54）に転送する。データメッセージはそこでカプセル化解除され、宛先に転送される。1030の後、CFEまたはそのMFNは、宛先の企業コンピューティングノード（例えばゲートウェイ854と858の間のIPsec接続、またはゲートウェイ1254と1258の間のIPsec接続）へのIPsec接続に沿ったメッセージの送信を管理するカウンタを（1035で）インクリメントする。

【０１５０】

[00194] 図11は、いくつかの実施形態において、中間MFNのCFE(例えば、CFE856)が、別のMFNの別のCFEに転送されるべきデータメッセージを受信したときに実行するプロセス1100を示す。図示されるように、プロセスは、最初に、仮想ネットワークに関連付けられたインタフェース上で（1105において）データメッセージを受信する。いくつかの実施形態では、このメッセージは、2つのトンネルヘッダである、VNトンネルヘッダ（例えば、ヘッダ862）およびテナントトンネルヘッダ（例えば、ヘッダ860）でカプセル化されている。

【０１５１】

[00195] 1110 で、プロセスは、このトンネルヘッダの宛先IPアドレスがそのCFEの宛先IPアドレスであると判定されるため、VNホップトンネルを終了する（例えば、CFE856の宛先IPアドレス）。次に、1115 で、プロセスはVNホップトンネルヘッダが正しいUDPポートを指定するかどうかを判定する。そうでない場合、プロセスは終了する。それ以外の場合、1120 で、プロセスはVNホップトンネルヘッダを削除する。受信したデータメッセージを考慮するために、CFEは、中間ホップCFEとして受信したメッセージの数を定量化するために、保持するRX (受信）カウンタを（1125において）インクリメントする。

【０１５２】

[00196] 1130において、プロセスは、このデータメッセージの次ホップインタフェースを識別するために、識別されたVNPルーティングコンテキスト（例えば、VRF名前空間のVNPの部分）において、経路ルックアップ（例えば、LPMルックアップ）を実行する。いくつかの実施形態では、この経路ルックアップは、内部テナントトンネルヘッダで識別される出力CFEの宛先IPに少なくとも部分的に基づいた（例えば、VRF名前空間のVNPの部分における）LPMルックアップである。

【０１５３】

[00197] 次に、プロセスは、（1135において）次ホップ出力インタフェースが中間CFEのローカルインタフェースであるかどうかを決定する。その場合、プロセスは（1140において） テナントトンネルヘッダを用いて既にカプセル化されているデータメッセージにVNホップトンネルヘッダを追加する。プロセスは、VNホップトンネルヘッダの宛先IPアドレスを、テナントトンネルヘッダで指定されている出力CFE の宛先IPアドレスに（1142において） 設定する。また、VNホップトンネルヘッダの送信元IPアドレスを（1142において) CFE のIPアドレスに設定する。このトンネルヘッダでは、プロセスはUDP属性（例えばUDPポートなど）も設定する。

【０１５４】

[00198] 次に、1144において、プロセスは、データメッセージを宛先の出力CFEに転送できるように、（1130で識別される）そのローカルインタフェースに、ダブルカプセル化されたデータメッセージを提供する。このVNホップトンネルのカプセル化解除および転送の一例は、図8のCFE 856の動作を参照して上述した。受信したデータメッセージを考慮するために、CFEは、中間ホップCFEとして送信したメッセージの数を定量化するために保持するTX (送信）カウンタを（1146で）インクリメントする。1146の後、プロセス1100は終了する。

【０１５５】

[00199] 一方、プロセスは、（1135において）次ホップ出力インタフェースがそのCFEのローカルインタフェースではなく、むしろ別のルータの宛先IPアドレスであると判定した場合、プロセスは、以前にVNホップトンネルヘッダを削除したデータメッセージに（1150において）VNホップトンネルヘッダを追加する。新しいVNホップトンネルヘッダで

は、処理1100がそのCFE の送信元IPアドレスと次ホップ中間CFEの宛先IPアドレス（1130 で特定）を、VNホップトンネルヘッダの送信元および宛先IPアドレスとして（1150において）埋め込む。このVNPトンネルヘッダは、UDPの宛先ポートを持つUDPレイヤ4プロトコルも指定する。

【０１５６】

[00200] 次に、1155において、プロセスは、識別されたVNPルーティングコンテキスト（例えば、VRFネームスペースのVNPの部分）において、別の経路ルックアップ（例えば、LPMルックアップ）を実行する。今回、ルックアップは、新しいVNホップトンネルヘッダで識別される次ホップ中間CFEのIPアドレスに基づく。この中間CFEは、仮想ネットワーク内の現在の中間CFEの次ホップであるため、ルーティングテーブルは、次ホップ中間CFEに送信されるデータメッセージのローカルインタフェースを識別する。したがって、VNPルーティングコンテキストにおけるこのルックアップは、現在の中間CFEがダブルカプセル化されたメッセージを提供するためのローカルインタフェースを識別する。次に、プロセスは、このデータメッセージの送信を考慮するための（1160において）VN中間TX (送信）カウンタをインクリメントする。1160の後、プロセスは終了する。

【０１５７】

[00201] 図13は、テナントの企業コンピューティングデバイス（例えば、支店内）から別のテナントマシン（例えば、別の支店、テナントデータセンタまたはSaaSプロバイダデータセンタ内）に送信されるテナントのメッセージを受信したときに、入力MFNのCFEによって実行されるメッセージ処理プロセス1300を示す。図9のプロセス900は、後述するように、このプロセス1300のサブセットである。図13に示すように、プロセス1300は、入ってくるIPsecトンネルの識別子に基づいて、テナントルーティングコンテキストを（905において）最初に識別することによって、開始される。

【０１５８】

[00202] 1310において、プロセスは、受信データメッセージのヘッダ内の送信元IPアドレスと宛先IPアドレスの両方がパブリックIPアドレスであるかどうかを判定する。そうである場合、プロセス（1315）は、データメッセージをドロップし、受信データメッセージのIPsecトンネルのために保持しているドロップカウンタをインクリメントする。1315 において、プロセスがカウンタをドロップするのは、テナントのIPsecトンネルを介してメッセージを受信する場合には、パブリックIPアドレスとの間でアドレス指定されたメッセージを受信すべきではないためである。いくつかの実施形態では、プロセス1300はまた、ICMPエラーメッセージを送信元の企業コンピューティングマシンへ送り返す。

【０１５９】

[00203] 一方、（1310において、)データメッセージがパブリックIPアドレスからのものではなく、他のパブリックIPアドレスへのものでもないと判定した場合、（1320において）受信データメッセージのヘッダ内の宛先IPアドレスがパブリックIPアドレスであるか否かを判定する。その場合、プロセスは、プロセス1300の開始時に実行された動作905を除き、図9のプロセス900を実行するために1325に遷移する。1325の後、プロセス1300は終了する。一方、プロセス1300は、（1320において）受信データメッセージのヘッダ内の宛先IPアドレスがパブリックIPアドレスでないと判定した場合、このデータメッセージを受信するために、識別されたIPsecトンネルのRX (受信）カウンタを（1330で）インクリメントする。

【０１６０】

[00204] 次に、プロセス1300は、識別されたテナントルーティングコンテキスト（例えば、VRFネームスペースのテナントの部分）において、（1335において）経路ルックアップ（例えば、LPMルックアップ）を実行する。このルックアップは、パブリッククラウドデータセンタを介して構築されたテナントの仮想ネットワークを出るための出力インタフェースのIPアドレスを識別する。図13に示す例では、プロセス1300は、データメッセージがSaaSプロバイダデータセンタ内のマシンに向けたものである場合、ルックアップ動作1335に到達する。したがって、このルックアップは、テナントの仮想ネットワーク

を出てSaaSプロバイダマシンに到達するための出力ルータのIPアドレスを識別する。いくつかの実施形態では、すべてのSaaSプロバイダ経路は、1つの経路テーブルまたはVRFネームスペースの1つの部分にインストールされ、他の実施形態では、異なるSaaSプロバイダの経路は、異なる経路テーブルまたは異なるVRFネームスペース部分に格納される。

【０１６１】

[00205] 1340において、プロセスは、受信データメッセージにテナントトンネルヘッダを追加し、このトンネルヘッダに送信元および宛先IPアドレスとして、入力CFEの送信元IPアドレスおよび出力ルータの宛先IPアドレスを埋め込む。次に、1345において、プロセスは、このデータメッセージの送信を考慮するために、テナントのためのVN送信カウンタをインクリメントする。1350において、プロセスは、このデータメッセージの次ホップインタフェースとしてそのローカルインタフェースの一つを識別するために、VNPルーティングコンテキスト（例えば、VNPネームスペースのVNPの部分）において、経路ルックアップ（例えば、LPMルックアップ）を実行する。次ホップが別のCFEである場合（他のパブリッククラウドデータセンタでの場合など）、いくつかの実施形態でのプロセスは、さらにVNホップヘッダでデータメッセージをカプセル化し、VNホップヘッダの送信元および宛先としてCFEのIPアドレスと他のCFEのIPアドレスを組み込む。1355において、プロセスは、データメッセージをその出力ルータに転送できるようにするために、カプセル化されたデータメッセージを識別されたローカルインタフェースに提供する。1355の後、プロセス1300は終了する。

【０１６２】

[00206] いくつかの場合には、入力MFNは、そのCFEが別のMFNのCFEを経由せずにデータメッセージの宛先マシンに直接転送することができる、テナントのためのデータメッセージを受信することができる。このようないくつかの場合において、CFEがテナント固有の情報を他の後続のVN処理モジュールに中継する必要がない場合、または必要な情報を他のメカニズムを介して後続のVN処理モジュールに提供可能である場合、データメッセージをテナントヘッダまたはVNホップヘッダでカプセル化する必要がない。

【０１６３】

[00207] たとえば、テナントのデータメッセージを外部のSaaSプロバイダデータセンタに直接転送するには、後でさらに説明するように、入力MFNのNATエンジン215は、テナント識別子に基づいてNAT動作を実行しなければならない。入力MFN内の入力CFEまたは別のモジュールは、入力MFNの関連するNATエンジン215にテナント識別子を提供しなければならない。入力CFEおよびNATエンジンが同じコンピュータ上で実行される場合、いくつかの実施形態は、この情報を共有メモリの場所に格納することによって、これら2つのモジュール間でこの情報を共有する。一方、CFEおよびNATエンジンが同じコンピュータ上で実行されない場合、いくつかの実施形態は、他のメカニズム（例えば、帯域外通信）を使用して、入力CFEおよびNATエンジン間でテナントIDを共有する。しかしながら、このような場合、他の実施形態は、入力MFNの異なるモジュール間で保存IDを記憶し、共有するためにカプセル化ヘッダを使用する（すなわち、帯域内通信を使用する）。

【０１６４】

[00208] いくつかの実施形態では、仮想ネットワークプロバイダは、上記のプロセス、システム、およびコンポーネントを使用して、同一または異なるパブリッククラウドプロバイダの複数のパブリッククラウド上で、複数の異なるテナント（たとえば、複数の企業の複数の異なる企業WAN）に複数の仮想WANを提供する。図14 は、1つ以上のパブリッククラウドプロバイダの第Nのパブリッククラウド1405 で、ネットワーク基盤およびコントローラクラスタ1410を持つ仮想ネットワークプロバイダの第Mのテナントの第Mの仮想企業WAN1415の例を示している。

【０１６５】

[00209] 各テナントの仮想WAN1415は、第Nのパブリッククラウド1405の全て又はこれらのパブリッククラウドのサブセットにまたがることができる。各テナントの仮想WAN1415は、1つ以上の支店1420、データセンタ1425、SaaSプロバイダデータセンタ1

430、およびテナントのリモートデバイスを接続する。いくつかの実施形態では、各テナントの仮想WANは、VNPのコントローラクラスタがテナントの異なるコンピューティングノード1420から1435の間でデータメッセージを効率的に転送するために必要であるとみなすパブリッククラウド1405にまたがる。パブリッククラウドの選択において、いくつかの実施形態におけるコントローラクラスタは、テナントが選択するパブリッククラウド、および／またはテナントまたはテナントの少なくとも1つのSaaSプロバイダが1つ以上のマシンを持つパブリッククラウドも考慮する。

【０１６６】

[00210] 各テナントの仮想WAN1415は、テナントのリモートデバイス1435（例えば、モバイルデバイスまたはリモートコンピュータ）が、SaaSプロバイダサービスにアクセスするために（すなわち、SaaSプロバイダマシンまたはマシンクラスタにアクセスするために）、任意の支店またはテナントデータセンタのテナントのWANゲートウェイとのやり取りを回避することを可能にする。いくつかの実施形態におけるテナントの仮想WANは、これらのWANゲートウェイ（例えば、WANセキュリティゲートウェイ）の機能を仮想WANがまたがるパブリッククラウド内の１つ以上のマシンに移動することによって、リモートデバイスが支店およびテナントデータセンタのWANゲートウェイを回避することを可能にする。

【０１６７】

[00211] たとえば、リモートデバイスがテナントまたはそのSaaSプロバイダサービスのコンピューティングリソースにアクセスできるようにするには、いくつかの実施形態のWANゲートウェイが、リモートデバイスがテナントのコンピュータリソースまたはそのSaaSプロバイダサービスにアクセスする方法を制御するファイアウォールルールを適用する必要がある。テナントのブランチまたはデータセンタWANゲートウェイを回避するために、テナントのファイアウォールエンジン210は、テナントの仮想WANがまたがる１つ以上のパブリッククラウドの仮想ネットワークMFNに配置される。

【０１６８】

[00212] これらのMFN内のファイアウォールエンジン210は、リモートデバイスとの間のデータメッセージフローに対してファイアウォールサービス動作を実行する。1つ以上のパブリッククラウドにデプロイされた仮想ネットワークでこれらの動作を実行することで、テナントのリモートデバイスに関連付けられたデータメッセージトラフィックが、ファイアウォールルール処理を受けるためにテナントのデータセンタまたは支店を介して不必要にルーティングされる必要がなくなる。これにより、テナントデータセンタおよび支店でのトラフィックの輻輳が緩和され、これらのロケーションでコンピューティングリソースに送信されないトラフィックを処理するために、これらのロケーションでコストの高い受信／送信ネットワーク帯域幅を消費することが回避される。また、このアプローチは、データメッセージフローが宛先（例えば、入力MFN、出力MFN、または中間ホップMFN）を通過するときに、介在するファイアウォールルール処理が仮想ネットワーク内で発生することを可能にするため、リモートデバイスとの間でのデータメッセージトラフィックの転送を高速化するのにも役立つ。

【０１６９】

[00213] いくつかの実施形態では、MFNのファイアウォールを実施するエンジン210（例えば、ファイアウォールサービスVM）は、VNP中央コントローラ160からファイアウォールルールを受信する。いくつかの実施形態におけるファイアウォールルールは、ルール識別子およびアクションを含む。いくつかの実施形態におけるルール識別子は、ファイアウォールルールがデータメッセージに一致するかどうかを判定するために、レイヤ2属性（例えば、MACアドレス）、レイヤ3属性（例えば、5つのタプル識別子など）、テナントID、ロケーションID（例えば、オフィスロケーションID、データセンタID、リモートユーザIDなど）などのデータメッセージ属性と比較される1つ以上の一致用の値(match value)を含む。

【０１７０】

[00214] ファイアウォールルールのいくつかの実施形態におけるアクションは、ファイアーウォールルールがデータメッセージの属性と一致する場合に、ファイアウォール実施エンジン210がデータメッセージに対して取るべきアクション（例えば、許可、ドロップ、再ダイレクトなど）を特定している。複数のファイアウォールルールがデータメッセージと一致する可能性に対処するために、ファイアウォール実施エンジン210は、あるファイアウォールルールが別のファイアウォールルールよりも高い優先順位を持つことができるように、ファイアウォールルールデータストレージにファイアウォールルール（コントローラクラスタ160から受信する）を階層的に保存する。データメッセージが2つのファイアウォールルールに一致する場合、ファイアウォール実施エンジンは、いくつかの実施形態において、より高い優先度をもつルールを適用する。他の実施形態では、ファイアウォール実施エンジンは、その階層に従ってファイアウォールルールを検査し（すなわち、より高い優先順位のルールをより低い優先順位のルールの前に検査する）、別のより低い優先順位のルールがデータメッセージと一致する場合に、それが最初により高い優先順位のルールと一致することを確実にする。

【０１７１】

[00215] いくつかの実施形態では、コントローラクラスタは、データメッセージが仮想ネットワークに入るときに入力ノード（例えば、ノード850）において、仮想ネットワーク上の中間ノード（例えば、ノード857）において、または仮想ネットワークを出るときに出力ノード（例えば、ノード855）において、ファイアウォールサービスがデータメッセージを検査するようにMFNコンポーネントを構成することができる。これらのノードのそれぞれにおいて、いくつかの実施形態におけるCFE(例えば、832、856、または858）は、その関連するファイアウォールサービスエンジン210を呼び出して、CFEが受信するデータメッセージ上でファイアウォールサービス動作を実行する。いくつかの実施形態では、ファイアウォールサービスエンジンは、このモジュールがデータメッセージに対してファイアウォールのアクションを実行することができるように、ファイアウォールサービスエンジンがそれを呼び出したモジュール（例えば、CFE）にその決定を返すのに対し、他の実施形態では、ファイアウォールサービスエンジンが、データメッセージに対してそのファイアウォールのアクションを実行することができるようにする。

【０１７２】

[00216] いくつかの実施形態において、他のMFNコンポーネントは、ファイアウォールサービスエンジンにその動作を実行させるよう指示する。例えば、いくつかの実施形態におけるVPNゲートウェイ（例えば、225または230）は、データメッセージが入力ノードのCFEに渡されるべきかどうかを判定するため、関連するファイアウォールサービスエンジンにその動作を実行するように指示する。また、いくつかの実施形態のCFEは、データメッセージをその関連するファイアウォールサービスエンジンに渡し、それは、データメッセージを通過させることを決定した場合、外部ネットワーク（例えば、インターネット）を介してその宛先にデータメッセージを渡すか、または、外部ネットワークを介してその宛先にデータメッセージを渡す前に、そのデータメッセージをその関連するNATエンジン215に渡してそのNAT動作を実行する。

【０１７３】

[00217] いくつかの実施形態の仮想ネットワークプロバイダは、パブリッククラウドで定義されているテナントのWANセキュリティゲートウェイが、ファイアウォールサービスに加えて、またはファイアウォールサービスの代わりに、他のセキュリティサービスを実装することを可能にする。たとえば、テナントの分散型WANセキュリティゲートウェイ（いくつかの実施形態では、テナントの仮想ネットワークがまたがる各パブリッククラウドデータセンタに分散される）には、ファイアウォールサービスエンジンだけでなく、侵入検知エンジンや侵入防御エンジンも含まれる。いくつかの実施形態では、侵入検知エンジンおよび侵入防御エンジンは、MFN150にアーキテクチャ的に組み込まれ、ファイアウォールサービスエンジン210と同様の位置を占める。

【０１７４】

[00218] いくつかの実施形態におけるこれらのエンジンの各々は、中央コントローラクラスタ160によって配信される侵入検知／防御ポリシーを格納する1つ以上のストレージを含む。いくつかの実施形態では、これらのポリシーは、テナントの仮想ネットワーク（複数のパブリッククラウドデータセンタにデプロイされる）への不正な侵入を検出／防止し、検出された侵入イベント（ログの生成、通知の送信、サービスまたはマシンのシャットダウンなど）に応じてアクションを実行するようにエンジンを構成する。ファイアウォールルールと同様に、侵入検知／防御ポリシーは、仮想ネットワークが定義されている、それぞれの管理転送ノード（たとえば、入力MFN、中間MFN、および／またはデータメッセージフローの出力MFN）で実施することができる。

【０１７５】

[00219] 前述のように、仮想ネットワークプロバイダは、仮想WANがまたがる各パブリッククラウドに少なくとも１つのMFNをデプロイして、テナントのメッセージフローが仮想WANに出入りできるようにMFN間の経路を定義するようにデプロイされたMFNを構成する。また、上述したように、各MFNは、いくつかの実施形態では異なるテナントによって共有されることができ、他の実施形態では、各MFNは、ただ1つの特定のテナントのためにデプロイされる。

【０１７６】

[00220] いくつかの実施形態では、各テナントの仮想WANは、オーバレイトンネルを介してそのWANによって使用されるMFNを接続することによって確立されるセキュアな仮想WANである。いくつかの実施形態におけるこのオーバレイトンネルアプローチは、各テナントのデータメッセージフローを、各テナントに固有のトンネルヘッダでカプセル化する。例えば、テナントを一意に識別するテナント識別子を含む。テナントの場合、いくつかの実施形態では、仮想ネットワークプロバイダのCFEは、テナントの仮想WANに出入りするための入力／出力転送要素を識別するために1つのトンネルヘッダを使用し、仮想ネットワークの介在する転送要素を通過するために別のトンネルヘッダを使用する。仮想WANのCFEは、他の実施形態において、異なるオーバレイカプセル化機構を使用する。

【０１７７】

[00221] 1つ以上のパブリッククラウド上にテナントの仮想WAN をデプロイするために、VNP のコントローラクラスタ は、（1）テナントの会社のコンピューティングノード（支店、データセンタ、モバイルユーザ、SaaS プロバイダなど） のロケーションに基づいて、テナントが利用可能なエッジMFN (異なるデータメッセージフローの入力または出力MFN として機能できる） を識別し、（2） 利用可能なすべてのエッジMFN 間の経路を識別する。これらの経路が識別されると、CFEの転送テーブルに伝播される（たとえば、OpenFlowを使用して、異なるOVSベースの仮想ネットワークルータに伝播される）。具体的には、テナントの仮想WAN を通過する最適な経路を識別するために、このWAN に関連付けられたMFN は、それらと隣接するMFN との間のネットワーク接続の品質を定量化する測定値を生成し、その測定値をVNPのコントローラクラスタに定期的に提供する。

【０１７８】

[00222] 前述のように、コントローラクラスタは、異なるMFNからの測定値を集約し、これらの測定値に基づいてルーティンググラフを生成し、テナントの仮想WANを通る経路を定義し、これらの経路をMFNのCFEの転送要素に配信する。テナントの仮想WAN の定義済み経路を動的に更新するために、このWAN に関連付けられているMFN は、定期的に測定値を生成し、これらの測定値をコントローラクラスタに提供する。コントローラクラスタは、受信した更新された測定値に基づいて、測定集計、経路グラフ生成、経路識別、および経路配信を定期的に繰り返す。

【０１７９】

[00223] テナントの仮想WAN を経由する経路を定義する場合、VNP のコントローラクラスタは、必要なエンドツーエンドのパフォーマンス、信頼性、およびセキュリティのために経路を最適化し、テナントのメッセージフローのルーティングを最小限に抑えよう

とする。また、コントローラクラスタは、ネットワークを通過するデータメッセージフローのレイヤ4処理を最適化するように、MFNコンポーネントを構成する（たとえば、接続パス全体でレート制御機構を分割することによってTCP接続のエンドツーエンドレートを最適化する）。

【０１８０】

[00224] パブリッククラウドの普及により、企業の各支店に近い主要なパブリッククラウドデータセンタを見つけることが非常に容易になった。同様に、SaaSベンダは、パブリッククラウド内、または、パブリッククラウドデータセンタの近くに同様に配置されるアプリケーションのホスティングを増加させている。したがって、仮想企業WAN1415は、パブリッククラウド1405を、企業のコンピューティングノード（例えば、支店、データセンタ、リモートデバイス、SaaSプロバイダ）の近傍に存在する企業ネットワーク基盤として安全に使用する。

【０１８１】

[00225] 企業WAN では、ビジネスクリティカルなアプリケーションを常に許容可能なパフォーマンスで提供するために、帯域幅の保証が必要である。このようなアプリケーションは、例えば、ERP、財務または調達、期限指向アプリケーション（例えば、工業またはIoT制御）、リアルタイムアプリケーション（例えば、VoIPまたはビデオ会議）のような、対話型データアプリケーションであり得る。したがって、従来のWAN基盤（フレームリレーやMPLSなど）は、このような保証を提供する。

【０１８２】

[00226] マルチテナントネットワークで帯域幅保証を提供する際の主な障害(例えば複数の仮想ネットワークで定義される共有仮想ネットワーク)は、特定の顧客のために1つ以上のパスで帯域幅を確保する必要があることである。いくつかの実施形態では、VNPはQoSサービスを提供し、入力認定レート(Ingress Committed Rate、ICR)保証および出力認定レート(Egress Committed Rate、ECR)保証を提供する。ICRは仮想ネットワークに入るトラフィックレートを表し、ECRは仮想ネットワークからテナントサイトに出るトラフィックレートを表す。

【０１８３】

[00227] トラフィックがICRおよびECRの制限を超えない限り、いくつかの実施形態における仮想ネットワークは、帯域幅および遅延の保証を提供する。たとえば、HTTPの入力または出力トラフィックが1 Mbpsを超えない限り、帯域幅と低遅延が保証される。これは、ポイント・ツー・クラウドモデルであり、なぜなら、QoSの目的のためには、VNPは、その宛先がICR／ECRの境界内にある限り、トラフィックの宛先を追跡する必要がないためである。このモデルは、ホースモデルと呼ばれることもある。

【０１８４】

[00228] より厳格なアプリケーションでは、顧客がポイント・ツー・ポイント保証を希望する場合、非常に重要なトラフィックを配信するために仮想データパイプを構築する必要がある。たとえば、エンタープライズでは、サービスレベルアグリーメントの保証が高い2つのハブサイトまたはデータセンタが必要になる場合がある。そのため、VNPルーティングでは、各顧客の帯域幅制約を満たすルーティングパスが自動的に選択される。これは、ポイント・ツー・ポイントモデルまたはパイプモデルと呼ばれる。

【０１８５】

[00229] エンドユーザに保証された帯域幅を提供するVNPの主な利点は、変化する帯域幅要求に応じてVNP基盤を調整することが可能なことである。ほとんどのパブリッククラウドでは、同じクラウドの異なる領域に配置された2つのインスタンス間の最低帯域幅保証が提供される。現在のネットワークに、新しい要求に対して、保証された帯域幅を提供するのに十分な未使用キャパシティがない場合、VNPはそのファシリティに新しいリソースを追加する。たとえば、VNPは高需要領域に新しいCFEを追加することができる。

【０１８６】

[00230] 1つの課題は、経路を計画し、基盤をスケールアップおよびスケールダウンす

る際に、この新しい次元のパフォーマンスとコストを最適化することである。アルゴリズムおよび帯域幅アカウンティングを容易にするために、いくつかの実施形態は、エンド・ツー・エンドの帯域幅確保が分割されないことを想定する。他の方法として、ある特定の帯域幅（例えば、10Mbps）が、特定のテナントのブランチAとブランチBの間に確保されている場合、帯域幅は、ブランチAが接続する入力CFEから始まる単一のパス上に割り当てられ、次に、ブランチBに接続されている出力CFEに到達するために、ゼロ以上の中間CFEのセットを通過する。また、いくつかの実施形態は、帯域幅の保証されるパスは単一のパブリッククラウドを通過するだけであると想定する。

【０１８７】

[00231] ネットワークトポロジ上で交差する様々な帯域幅の確保を考慮するために、いくつかの実施形態では、VNPは、確保された帯域幅のパス上のルーティングを静的に定義し、データメッセージフローが、帯域幅要件のために、確保された同じ経路を常に通過するようにする。いくつかの実施形態では、各経路は、その経路によって通過する各CFEが、この経路に関連付けられた単一の発信インタフェースと一致する単一のタグで識別される。具体的には、各CFE は、ヘッダにこのタグがあり、特定の着信インタフェースから到着する各データメッセージに対して、単一の発信インタフェースを照合する。

【０１８８】

[00232] いくつかの実施形態では、コントローラクラスタは、いくつかの相互接続されたノードによって形成されるネットワークグラフを維持する。グラフの各ノードnには、このノードに割り当てられた保証帯域幅の合計（TBW_n） と、このノードによってすでに確保（特定の予約パスに割り当て） されている帯域幅の合計（RBW_n） がある。さらに、各ノードについて、グラフは、ギガバイトあたりのコスト（C_{ij}）と、このノードとグラフ内の他のすべてのノードとの間の送信トラフィックに関連するミリ秒単位の遅延（D_{ij}）とを含む。ノードiとノードjの間の送信トラフィックに関連する重みは$W_{ij}＝a＊C_{ij}＋D_{ij}$であり、ここでaは、典型的には1から10の間のシステムパラメータである。

【０１８９】

[00233] ブランチAとブランチBの間の値BWの帯域幅確保の要求が受け入れられると、コントローラクラスタは、まず、ブランチAとブランチBにそれぞれバインドされている特定の入力および出力ルータnとmに要求をマッピングする。次に、コントローラクラスタは、nとmの間の2つの最低コスト（例えば、最短経路）を計算するルーティングプロセスを実行する。最初の経路は、計算された経路に沿った利用可能な帯域幅に関係なく、nとmの間の最低コスト（例えば最短経路）経路である。この経路の全重みは、W1として計算される。

【０１９０】

[00234] 第2の最低コスト（例えば、最短経路）計算は、最初にBW ＞ TBW_i-RBW_iの全てのノードiを排除することによってグラフを修正する。修正されたグラフは、トリミング済グラフと呼ばれる。次に、コントローラクラスタは、トリミング済みのグラフ上で第２の最低コスト（例えば、最短経路）経路計算を実行する。第２の経路の重みがK％以下（K は通常、第１の経路よりも10％～30％高い） の場合、第２の経路が優先パスとして選択される。一方、この要件を満たさない場合、コントローラクラスタは、第１の経路に、TBW_i-RBW_iの価値が最も小さいノードiを追加し、次に、2つの最低コスト（例えば、最短経路）計算を繰り返す。コントローラクラスタは、条件が満たされるまで、さらにルータを追加し続ける。その時点で、確保された帯域幅BWがすべてのRBW_iに追加され、ここで$_i$は選択された経路上のルータである。

【０１９１】

[00235] すでに帯域幅が確保されている経路の追加帯域幅を要求する特別な場合、コントローラクラスタは、まずノードAとBの間の現在の帯域幅確保を削除し、これらのノード間の合計帯域幅要求のパスを計算する。これを行うために、いくつかの実施形態では、各ノードのために保持される情報は、各タグ、または各送信元および宛先ブランチのために確保された帯域幅も含み、確保された帯域幅全体だけではない。帯域幅確保がネットワー

クに追加された後に、仮想ネットワーク経由の測定されたネットワーク遅延またはコストに大きな変化がない限り、いくつかの実施形態は、経路を再考しない。しかしながら、測定および／またはコストが変更した場合、これらの実施形態は、帯域幅予約および経路計算処理を繰り返す。

【０１９２】

[00236] 図15は、仮想ネットワークプロバイダのコントローラクラスタ160によって実行され、特定のテナントのために仮想WANをデプロイし管理するプロセス1500を概念的に示している。いくつかの実施形態では、プロセス1500は、コントローラクラスタ160上で実行されるいくつかの異なるコントローラプログラムによって実行される。このプロセスの動作は、図15に示すシーケンスに従う必要はない。なぜなら、これらの動作は、異なるプログラムによって並列または異なるシーケンスで実行することができるからである。したがって、これらの動作は、コントローラクラスタによって実行される動作の1つの例示的なシーケンスを説明するためだけに、この図に示されている。

【０１９３】

[00237] 図示されているように、コントローラクラスタは、最初に、いくつかの異なるパブリッククラウドプロバイダ(Amazon AWS、Google GCP など）の複数のパブリッククラウドデータセンタに（1505 で） 複数のMFN をデプロイする。いくつかの実施形態におけるコントローラクラスタは、プロセス1500が示されている特定のテナントとは異なる1つ以上の他のテナントに対して、これらのデプロイされたMFNを（1505において）構成する。

【０１９４】

[00238] 1510において、コントローラクラスタは、特定のテナントの外部マシン属性およびロケーションに関する特定のテナントデータを受信する。いくつかの実施形態では、このデータは、特定のテナントによって使用されるプライベートサブネットと、特定のテナントが外部マシンを有する1つ以上のテナントオフィスおよびデータセンタの識別子とを含む。いくつかの実施形態では、コントローラクラスタは、APIを介して、またはコントローラクラスタが提供するユーザインタフェースを介してテナントデータを受信することができる。

【０１９５】

[00239] 次に、1515において、コントローラクラスタは、特定のテナントのための仮想ネットワークを確立するために使用する候補MFNであるMFN150の測定エージェント205によって収集された測定値から、特定のテナントのためのルーティンググラフを生成する。上述のように、ルーティンググラフは、MFNを表すノードと、MFN間のネットワーク接続を表すノード間のリンクを有する。リンクには、関連付けられた重みがある。これは、リンクによって表されるネットワーク接続を用いる際の品質やコストを定量化するコスト値である。前述のように、コントローラクラスタは、まず、収集された測定値から測定グラフを生成し、次に、最適でない（例えば、大きな遅延またはドロップ率を有する）測定グラフからリンクを除去することによって、ルーティンググラフを生成する。

【０１９６】

[00240] ルーティンググラフを構築した後、コントローラクラスタは（1520で）パス検索を実行して、テナントの外部マシンが（MFNによって配置される）仮想ネットワークにデータメッセージを送信し、そして仮想ネットワークからデータメッセージを受信するために使用することができる、入出力ノード（つまり、MFN）の候補の異なるペア間の利用可能な経路を識別する。いくつかの実施形態では、コントローラクラスタは、既知のパス探索アルゴリズムを使用して、各候補ノードの入出力ペア間の異なるパスを識別する。このようなペアの各パスは、入力ノードから出力ノードへ、ゼロ個以上の中間ノードを介して連結されたときに通過する、1つ以上のリンクを使用する。

【０１９７】

[00241] いくつかの実施形態では、任意の2つのMFN間のコストは、2つのMFN間の接続リンクのための推定遅延と財務コストの加重和を含む。待ち時間と財務コストは、いく

つかの実施形態において、（1）リンク遅延測定、（2）推定メッセージ処理待ち時間、（3）同一パブリッククラウドプロバイダの別のデータセンタからパブリッククラウドプロバイダの別のデータセンタへ、またはパブリッククラウドプロバイダのクラウド（例えば、別のパブリッククラウドプロバイダの別のパブリッククラウドデータセンタへ、またはインターネットへ）を出るために、特定のデータセンタからの発信トラフィックに対するクラウド料金、および（4）パブリッククラウド内のホストコンピュータ上で実行されるMFNに関連する推定メッセージ処理コストのうちの1つ以上を含む。

【０１９８】

[00242] いくつかの実施形態は、可能な限りこのような横断を最小化するために、公衆インターネットを通過する2つのMFN間の接続リンクのペナルティを評価する。また、いくつかの実施形態は、そのような接続を使用するために経路生成をバイアスするために、2つのデータセンタ間のプライベートネットワーク接続の使用を促進する（例えば、接続リンクコストを削減することによって）。これらのペアワイズリンクの計算コストを使用して、コントローラクラスタは、ルーティングパスによって使用される個々のペアワイズリンクのコストを集計することによって、これらのペアワイズリンクのうちの1つ以上を使用する各ルーティングパスのコストを計算することができる。

【０１９９】

[00243] 次に、コントローラクラスタは、ノードの各候補の入出力ペア間の識別された候補パスの計算されたコスト（例えば、最低の集約コスト）に基づいて、1つまたは最大N個の識別されたパス（ここで、Nは、1より大きい整数）を（1520において）選択する。いくつかの実施形態では、各パスについて計算されたコストは、上述のように、パスによって使用される各リンクの重みコストに基づく（例えば、各リンクの関連する重み値の合計である）。コントローラクラスタは、2つのMFN 間で複数の経路が必要な場合に、入力MFN または中間MFN がマルチパス動作を実行できるように、入力／出力ノードのペア間で複数のパスを選択することができる。

【０２００】

[00244] （1520で）入力／出力ノードの候補ペアごとに1つまたはN個のパスを選択した後、コントローラクラスタは選択されたパスに基づいて1つまたはN個の経路を定義し、特定のテナントの仮想ネットワークを実装するMFNの経路テーブルまたは経路テーブルの一部を生成する。生成された経路レコードは、特定のテナントの異なるサブネットに到達するエッジMFNを識別し、入力MFNから出力MFN への経路を通過するための次ホップMFNを識別する。

【０２０１】

[00245] 1525において、コントローラクラスタは、特定のテナントのための仮想ネットワークを実装するように、これらのMFNの転送要素235を構成するために、経路レコードをMFNに配信する。いくつかの実施形態では、コントローラクラスタは、ソフトウェアで定義されるマルチテナントデータセンタで現在使用されている通信プロトコルを使用して転送要素と通信し、ホストコンピュータ上で実行されるソフトウェアルータを構成して、ホストコンピュータにまたがる論理ネットワークを実装することにより、経路レコードを通過させる。

【０２０２】

[00246] MFN が構成され、仮想ネットワークが特定のテナントに対して動作するようになると、エッジMFNはテナントの外部マシン（つまり、仮想ネットワーク外のマシン）からデータメッセージを受信し、これらのデータメッセージを仮想ネットワーク内のエッジMFNに転送する。このエッジMFNは、テナントの他の外部マシンにデータメッセージを転送する。このような転送動作を実行している間、入力MFN、中間MFN、および出力MFNは、その転送動作に関する統計を収集する。また、いくつかの実施形態では、いくつかの実施形態では、各MFN上の1つ以上のモジュールが、パブリッククラウドデータセンタにおけるネットワークまたはコンピューティング消費に関する他の統計を収集する。いくつかの実施形態では、パブリッククラウドプロバイダは、そのような消費データを収集し、

収集されたデータを仮想ネットワークプロバイダに渡す。

【０２０３】

[00247] 課金サイクルに近づくと、コントローラクラスタは、MFNによって収集された（例えば、1530において）統計、および／またはMFNによって収集された、またはパブリッククラウドプロバイダによって提供されたネットワーク／コンピューティング消費データを収集する。収集された統計、および／または提供されたネットワーク／コンピューティング消費データに基づいて、コントローラクラスタは（1530で）請求レポートを生成し、特定のテナントに請求レポートを送信する。

【０２０４】

[00248] 前述のように、請求レポートで請求される金額は、コントローラクラスタが受信する統計およびネットワーク／消費データ（たとえば、1530）を示す。また、いくつかの実施形態では、この請求は、仮想ネットワークプロバイダが（特定のテナントのために仮想ネットワークを実装する）MFNを運用するために費やしたコストと、収益率（例えば、10%の増加）を加算したコストを考慮している。特定のテナントは、テナントの仮想ネットワークがデプロイされている複数の異なるパブリッククラウドプロバイダからの請求を処理する必要がないため、この課金スキームは特定のテナントにとって便利である。いくつかの実施形態におけるVNPの発生費用には、パブリッククラウドプロバイダがVNPに請求する費用が含まれる。1530において、コントローラクラスタはまた、クレジットカードを請求するか、または請求書に反映される料金のために銀行口座から電子的に資金を引き出す。

【０２０５】

[00249] 1535において、コントローラクラスタは、測定エージェント205から新しい測定値を受信したか否かを判定する。そうでない場合、プロセスは1545に遷移する。これについては後述する。一方、コントローラクラスタは、測定エージェントから新しい測定値を受信したと判定した場合、新しい測定値に基づいて、特定のテナントのルーティンググラフを再検査する必要があるかどうかを（1540 で）判定する。MFNの障害がない場合、いくつかの実施形態のコントローラクラスタは、受信され、更新された測定値に基づいて、特定の期間（例えば、24時間毎または週毎）に、各テナントのルーティンググラフを最大で1回更新する。

【０２０６】

[00250] コントローラクラスタが、受信した新しい測定に基づいてルーティンググラフを再検査する必要があると判定した場合（1540で）、プロセスは、新たに受信した測定に基づいて新しい測定グラフを（1545において）生成する。いくつかの実施形態では、新しい測定セットが受信されるたびに、測定グラフのリンクに関連する測定値が劇的に変動しないことを確実にするために、コントローラクラスタは、重み付け和を使用して、各新しい測定を以前の測定と混合する。

【０２０７】

[00251] 1545において、コントローラクラスタはまた、調整された測定グラフに基づいてルーティンググラフを調整する必要があるかどうか（例えば、ルーティンググラフのリンクの重み値を調整する必要があるかどうか、またはリンクに関連する調整された測定値のためにルーティンググラフ内のリンクを追加または削除する必要があるかどうか）を判定する。その場合、コントローラクラスタ（1545）は、ルーティンググラフを調整し、パス探索動作（動作1520など）を実行して、入出力ノードのペア間の経路を識別し、識別された経路に基づいて経路レコードを生成し、経路レコードをMFNに配信する。1545から、プロセスは1550に遷移する。

【０２０８】

[00252] また、コントローラクラスタが（1540において）ルーティンググラフを再検査する必要がないと判定した場合、プロセスは1550に遷移する。2150で、コントローラクラスタは、処理されたデータメッセージおよび消費されたネットワーク／コンピューティングリソースに関する統計を収集する必要があるかを判定する。そうでない場合、プロ

セスは1535に戻り、MFN測定エージェントから新しい測定値を受信したかどうかを判定する。それ以外の場合、プロセスは1530に戻り、統計、ネットワーク／コンピューティング消費データの収集、および請求レポートの生成と送信を行う。いくつかの実施形態では、コントローラクラスタは、特定のテナントがパブリッククラウドデータセンタ全体にデプロイされる仮想ネットワークを必要としなくなるまで、処理1500の動作を繰り返し実行する。

【０２０９】

[00253] いくつかの実施形態では、コントローラクラスタは、パブリッククラウドデータセンタにテナント用の仮想ネットワークをデプロイするだけでなく、パブリッククラウドデータセンタにコンピューティングノードマシンおよびサービスマシンをデプロイし、構成する際にテナントを支援する。デプロイされたサービスマシンは、MFNのサービスマシンとは別のマシンにすることができる。いくつかの実施形態では、コントローラクラスタの請求レポートは、配置されたコンピューティングおよびサービスマシンによって消費されるコンピューティングリソースも考慮する。繰り返しになるが、複数のパブリッククラウドプロバイダの複数のパブリッククラウドデータセンタで消費されるネットワークおよびコンピューティングリソースのために、1つの仮想ネットワークプロバイダから1つの請求書を発行する方が、複数のパブリッククラウドプロバイダから複数の請求書を受信するよりもテナントにとって望ましい。

【０２１０】

[00254] 他の実施形態は、専用または共有MFNをデプロイするために他のデプロイモデルを使用して、ネットワーク上に専用または共有仮想ネットワークを実施し、2つ以上のパブリッククラウドプロバイダのインフラストラクチャを計算する。例えば、いくつかの実施形態では、仮想ネットワークプロバイダは、1つ以上のクラウドサービス再販業者が、1つ以上の顧客のために専用または共有MFNをデプロイすることを可能にする。いくつかの実施形態では、専用仮想ネットワークは、1つのエンティティに対してデプロイされる専用MFNで確立されるものである。一方、これらの実施形態のいくつかでは、共有仮想ネットワークは、複数のエンティティによる使用のためにデプロイされる共有MFNを用いて確立されるものである。いくつかの実施形態では、複数のエンティティに対する複数の分離された仮想ネットワークを、共有された仮想ネットワーク上で定義することができる。

【０２１１】

[00255] 図16は、このようなデプロイモデルの1つを示している。図に示すように、このデプロイモデルでは、3つのSaaSサービスセットを提供する3つのレベルのSaaSプロバイダ1605-1615を使用する。第1のSaaSレイヤ1605は、複数の異なるパブリッククラウド1650において、コンピューティングおよびネットワークインフラストラクチャ(例えば、ホストコンピュータ、VMおよび/またはコンテナなど)と、コンピューティング要素を接続するネットワーク要素(ハードウェアまたはソフトウェアスイッチ、ルータ、ミドルボックス要素など)とを提供する1つ以上のパブリッククラウドプロバイダ1620によって提供される。第2のSaaSレイヤ1610は、クラウド再販業者のために複数のパブリッククラウド1650にわたって仮想ネットワークをデプロイするためのツールを提供するVNP1625によって提供される。クラウド再販業者1630は、第3のSaaSレイヤ1615をその客先1640に提供し、第3のSaaSレイヤ1615は、クラウド再販業者ツールを使用して、1つ以上のパブリッククラウド1650にわたってデプロイするためのコンピューティング要素およびネットワークインフラストラクチャ(例えば、仮想ネットワーク)を定義する。

【０２１２】

[00256] いくつかの実施形態におけるクラウド再販業者1630は、パブリッククラウドプロバイダの各々と独自の顧客アカウントを有する。いくつかの実施形態では、クラウド再販業者は、各パブリッククラウドプロバイダとの顧客アカウントを直接確立し(破線のアーク1652によって識別される)、次いで、この顧客アカウントのセキュリティクレデンシャルをVNPプロバイダ1625に提供する。他の実施形態では、クラウド再販業者は、VNP

プロバイダのマシン1625を介してそのパブリッククラウド顧客アカウントを確立する。パブリッククラウドインフラストラクチャを介してそのテナントの1つの仮想ネットワークをデプロイするようにVNPに指示するために、クラウド再販業者マシン1630は、クラウド再販業者を認証するために、最初にクラウド再販業者のVNPセキュリティクレデンシャル(例えば、ユーザ名、パスワード、証明書など)をVNPマシン1625に提供する。

【０２１３】

[00257] 図に示すように、クラウド再販業者は、VNP1625に所望のネットワーク設定も提供する。この設定では、テナントにデプロイする必要がある仮想ネットワークの属性について記載する。いくつかの実施形態では、この設定データは、クラウド再販業者1630がVNP1625に仮想ネットワークをデプロイするように命令する顧客のために使用するテナント識別子も含む。いくつかの実施形態におけるこのテナント識別子は、VNPから得られる。他の実施形態では、設定データは、クラウド再販業者が自身のテナントのための自身のテナント識別子にマップする、VNPによって提供されるテナント識別子を含む。

【０２１４】

[00258] いくつかの実施形態では、VNPマシン1625は、再販業者の顧客のために仮想ネットワークをデプロイするとき(例えば、MFNをデプロイするとき、適切なルーティング記録を用いてMFN CFEを構成するときなど)に、クラウド再販業者のセキュリティクレデンシャルをパブリッククラウドマシン1620に提供する。これらのセキュリティクレデンシャル(例えば、ユーザ名、パスワード、セキュリティ証明書等)は、VNPマシンが、クラウド再販業者の最上位で動作するマシンとして自身を認証することを可能にする(例えば、クラウド再販業者マシン1630がパブリッククラウドマシン1620にログオンしているかのように、パブリッククラウドマシン1620にログオンする)。

【０２１５】

[00259] クラウド再販業者の各テナントのために仮想ネットワークをデプロイするために、VNPマシンは、クラウド再販業者のテナントのためのテナント識別子と共に、仮想ネットワークを形成するコンポーネント(例えば、MFN、MFNゲートウェイ/サービスボックス/CFEなど)も、いくつかの実施形態において構成する。いくつかの実施形態では、これらの構成された構成要素は、次いで、それらが収集する統計(例えば、ルーティング統計およびデプロイマシン計算統計)をテナント識別子と関連づけ、これにより、クラウド再販業者の顧客は、以下にさらに述べるように、これらの統計に基づいて適切に課金されることができる。

【０２１６】

[00260] クラウド再販業者がその顧客のテナント識別子をVNPマシン1625に提供する実施形態のいくつかにおいて、VNPマシン1625は、クラウド再販業者1630によって提供される各テナント識別子をVNPの固有テナント識別子にマッピングし、収集された統計のVNPテナント識別子を、クラウド再販業者に統計を提供する前にクラウド再販業者のテナント識別子に変換する。他の実施形態では、クラウド再販業者マシン1630および/またはVNPプロバイダマシン1625は、パブリッククラウドプロバイダマシン1620によって提供されるテナント識別子を使用する。

【０２１７】

[00261] VNPマシン1625は、クラウド再販業者1630によって提供されるテナントネットワーク設定データを使用して、クラウド再販業者の顧客ごとに仮想ネットワークの仮想ネットワーク構成要素(例えば、MFNなど)をデプロイおよび/または構成する。いくつかの実施形態では、複数の異なる顧客が、同じ仮想ネットワークおよび/または仮想ネットワークコンポーネントを共有する。統合的に、または代替的に、いくつかの実施形態におけるVNPマシン1625は、クラウド再販業者の任意の個々の顧客のための専用仮想ネットワークを定義することができる。図に示すように、VNPマシン1625は、パブリッククラウドプロバイダマシン1620を介して仮想ネットワークおよび/または仮想ネットワークコンポーネントをデプロイし、構成する。いくつかの実施形態では、再販業者、VNPおよびパブリッククラウドマシン1620～1630は、それぞれのAPIインタフェースおよび介在す

るネットワークファブリック(例えば、インターネット)を介して互いに通信する。

【０２１８】

[00262] VNPマシン1625は、デプロイ仮想ネットワークコンポーネントからテナントごとの統計情報(例えば、CFEおよびゲートウェイからのルーティング統計情報など)を収集し、統計情報を集約し、集約した統計情報をクラウド再販業者に転送する。VNPマシンが再販業者顧客識別子をVNPテナント識別子にマッピングする実施形態では、VNPマシンは、変換された顧客識別子と共に集計される統計を供給する前に、テナント識別子を顧客識別子に変換する。また、VNPマシンがパブリッククラウド識別子をVNPテナント識別子にマッピングする実施形態では、VNPマシンはパブリッククラウド識別子をVNP識別子に変換する。

【０２１９】

[00263] 示されるように、VNPマシン1625は、クラウド再販業者1630に定期的に請求レポートを送信して、VNPによって実行されるサービスの料金を収集する。いくつかの実施形態では、これらの請求レポートは、クラウド再販業者のための仮想ネットワークを2つ以上のパブリッククラウド上にデプロイするためのVNPのサービス料金に対して、クラウド再販業者に請求する。これらのデプロイ費用には、パブリッククラウド内のMFN間、およびテナントの外部マシンのロケーション間のリンクの品質および/またはコストを定量化する測定を生成する測定動作など、そのようなデプロイをサポートする補助的な動作を実行するための料金が含まれる。

【０２２０】

[00264] また、いくつかの実施形態におけるVNPマシン1625は、クラウド再販業者のために1つ以上のパブリッククラウドプロバイダから課金データを受信する。この請求データは、いくつかの実施形態において、クラウド再販業者の顧客クレデンシャル(例えば、クラウド再販業者のためのPCプロバイダ顧客番号)に関連付けられる。これらの実施形態におけるVNPマシン1625は、請求データをクラウド再販業者に渡す(例えば、マークアップ調整を伴うか、マークアップ調整を伴わない)。他の実施形態では、パブリッククラウドプロバイダは、破線1652によって示されるように、請求レポートをクラウド再販業者マシン1630に直接送信する。

【０２２１】

[00265] いくつかの実施形態におけるクラウド再販業者マシン1630は、VNPマシン1625によって提供される使用統計を使用して、その顧客に仮想ネットワークの料金を請求する。いくつかの実施形態では、VNPマシン1625は、ネットワークインフラストラクチャをデプロイするだけでなく、クラウド再販業者マシン1630のコンピューティングインフラストラクチャをデプロイする。これらの実施形態のいくつかでは、使用統計は、使用されたコンピューティングリソースを反映し、クラウド再販業者マシン1630は、これらの統計を使用して再販業者の顧客に課金する。一部の実施形態では、クラウド再販業者は、収集された統計を使用してその顧客に課金するのではなく、顧客がデプロイを要求したコンピューティングおよび/またはネットワーク構成に基づいてその顧客に課金する。

【０２２２】

[00266] 図16の3レイヤSaaSデプロイモデルの違いをさらに説明するために、図17は、前述した2レイヤSaaSデプロイモデルの同様の図を示している。図17のこの2層モデルには、クラウド再販業者1615は含まれない。むしろ、この2層モデルでは、VNPの顧客は、VNPが専用の仮想ネットワークをデプロイするエンティティであるか、VNP が複数の顧客にデプロイした共有仮想ネットワークを使用する。前述したように、各顧客のネットワークトラフィックが他の顧客のネットワークトラフィックから安全に分離されるため、異なる顧客がコンポーネントを安全に共有して、1つ以上のパブリッククラウドを介して共有仮想ネットワークを定義できる。いくつかの実施形態では、上述したように、テナント識別子を使用して異なる顧客のトラフィックを分離することによって、複数の仮想ネットワークが共有仮想ネットワーク上の複数の顧客に対して定義される。

【０２２３】

[00267] 図17の2レイヤモデルでは、第1のSaaSレイヤ1605は、複数の異なるパブリッククラウド1650においてコンピューティングおよびネットワークインフラストラクチャを提供する1つ以上のパブリッククラウドプロバイダ1620によって提供され、一方、第2のSaaSレイヤ1610は、そのいくつかの顧客のために複数のパブリッククラウド1650にわたって仮想ネットワークをデプロイするためのツールを提供するVNP 1625によって提供される。図に示すように、VNPマシン1625は、パブリッククラウドプロバイダに、パブリッククラウドのVNPセキュリティクレデンシャルを提供する。

【０２２４】

[00268] VNPマシンは、(テナント識別子に関連付けられる)顧客ごとにテナントネットワーク設定データを受信し、このデータに基づいて、顧客ごとに仮想ネットワークの仮想ネットワークコンポーネント(MFNなど)をデプロイおよび/または構成する。図に示すように、VNPマシン1625は、パブリッククラウドプロバイダマシン1620を介して仮想ネットワークおよび/または仮想ネットワークコンポーネントをデプロイし、構成する。いくつかの実施形態では、VNPおよびパブリッククラウドマシン1620および1625は、それぞれのAPIインターフェースおよび介在するネットワークファブリック(例えば、インターネット)を介して互いに通信する。

【０２２５】

[00269] さらに示されるように、VNPマシン1625は、デプロイ仮想ネットワークコンポーネントからテナントごとの統計情報(例えば、CFEおよびゲートウェイからのルーティング統計情報など)を収集し、収集された統計情報を集約する。VNPマシン1625は、VNPによって実行されるサービスの料金を収集するために、各VNP顧客に定期的に課金レポートを送信する。上述したように、いくつかの実施形態におけるこれらの手数料は、パブリッククラウドプロバイダが顧客の仮想ネットワークによって消費されたリソース(例えば、コンピューティングおよび/またはネットワークリソース)のためにVNPを請求した手数料に、特定のマークアップ割合を加えたものを含む。VNPマシンは、これらのマシンが収集し、顧客の関連付けられた識別子について集約する統計に基づいて、各顧客の仮想ネットワークによるリソースの量を識別する。他の実施形態では、VNPマシンは、顧客の仮想ネットワークによって消費されるリソースに対する各パブリッククラウドプロバイダの料金に加えて、各顧客のVNPリソースの使用に対する料金を各顧客に送信する。

【０２２６】

[00270] いくつかの実施形態は、1つ以上のパブリッククラウドプロバイダの複数のパブリッククラウド(例えば、複数のパブリッククラウドデータセンタ)への複数の接続リンクを介して、テナントのマルチマシンコンピューティングノード(例えば、支社またはデータセンタ)をテナントのパブリッククラウド仮想ネットワークに接続する。マルチマシンコンピューティングノード(MMCN) とパブリッククラウド仮想ネットワークの間に複数のリンクがあると、1つ以上の接続リンクの障害に耐えることができるため、仮想ネットワークの可用性が高くなる。いくつかの実施形態では、1つのリンクはプライマリリンクであり、他方のリンクはフェイルオーバーリンクである。このアプローチでは、仮想ネットワークを経由する最適なルーティングパスのMMCN から、仮想ネットワークから別のMMCNまたはSaaSプロバイダデータセンタに出るための出口ノードへの最適なルーティングパスを選択することによって、各MMCN から他のMMCNまたはSaaSプロバイダデータセンタへの最適な経路を確立することもできる。

【０２２７】

[00271] 以下の説明では、マルチホームMMCNという用語を使用して、1つ以上のパブリッククラウドプロバイダの複数のパブリッククラウドへの複数の接続リンクを介してテナントのパブリッククラウド仮想ネットワークに接続するテナントのマルチマシンコンピューティングノードを参照する。以下の説明では、マルチホームSaaSデータセンタという用語も使用している。これは、仮想ネットワークが、1つ以上のパブリッククラウドプロバイダの1つ以上のパブリッククラウド(複数のパブリッククラウドデータセンタなど) 内の複数のMFNを関連付けるSaaSデータセンタを指す。いくつかの実施形態におけるこれら

のMFNは、SaaSプロバイダに到達するために仮想ネットワークを通過する経路の候補出力ノードである。SaaSデータセンタの仮想ネットワークを接続するために2つ以上の出力ノードを使用することは、リンクフェイルオーバーサポートを可能にし、異なるペアの外部コンピュータノード(例えばリモートコンピュータまたはMMCN)とSaaSプロバイダデータセンタ間の最適経路の使用を可能にするという点でも利点がある。

【０２２８】

[00272] いくつかの実施形態では、仮想ネットワークコントローラ160が複数のMFNをSaaSデータセンタに関連付けている場合でも、SaaSデータセンタは、複数のパブリッククラウドデータセンタ内の仮想ネットワークの複数のMFNへの経路を開始する必要はない。一方、いくつかの実施形態におけるマルチホームMMCNは、仮想ネットワークへの異なるリンクを介する経路を積極的に開始する必要がある。これを行うために、フェイルオーバー機能を備えた適切なルータを持つことによって(たとえば、Cisco2800 シリーズを使用することによって)、マルチホームMMCNでフォールバック機能を提供することが容易になる。

【０２２９】

[00273] 最適なルーティングのために、マルチホームMMCNは、MMCNと、MMCNが接続可能な異なるパブリッククラウドデータセンタとの間の性能(遅延、損失など)を測定するために測定プロセスを実行する1つ以上のコンピュータまたは装置を、いくつかの実施形態において含む。さらに、いくつかの実施形態におけるMMCNは、エンティティ(例えば、テナントのため)のための仮想ネットワークを定義する中央コントローラクラスタ(例えば、コントローラクラスタ160)によって定義される経路に基づいて、その全体的なルーティング動作を実行する。これを実現するために、マルチホームMMCNには、仮想ネットワークをデプロイするための中央コントロールプレーンの一部として動作するSD-WAN機能(Velocloud やViptelaアプライアンスなど) が搭載されている。上述したように、中央コントロールプレーンは、いくつかの実施形態において、2つ以上のコントローラのクラスタによって実現される。

【０２３０】

[00274] 図18は、特定のマルチホームMMCNのための経路を定義するために、いくつかの実施形態の中央コントローラクラスタによって使用されるプロセス1800を示す。このプロセスでは、特定のマルチホームMMCN 内の特殊なルータを使用して、複数の接続リンクを介してMMCN から仮想ネットワークにデータメッセージを転送するルーティング動作を実行するために定義された経路を使用する。特殊化されたルータは、いくつかの実施形態ではソフトウェアルータまたはソフトウェアルータのクラスタであり、一方、他の実施形態ではルーティング装置(例えば、SD-WAN装置)である。MMCN内の特殊なルータまたはルータクラスタは、以下の説明ではMMCN のエッジノードと呼ばれる。いくつかの実施形態では、中央コントローラクラスタは、インターネットを介してMMCNのエッジノードを遠隔制御する。

【０２３１】

[00275] 図に示すように、中央コントローラクラスタは、最初に、DNSサーバサービスおよびエッジノードIPアドレスに従って、特定のMMCNエッジノードに最も近いN個の異なるクラウド領域からN個のMFN(例えば、10から12個)のサブセットを(1805において)識別する。いくつかの実施形態におけるN個のMFNは、特定のMMCNの特定の距離内に、各クラウドプロバイダから少なくとも1つの候補MFNを有する。また、いくつかの実施形態では、N個のMFNのそれぞれは、特定のMMCNエッジノードとの安全な接続リンクを確立するためのゲートウェイ(例えば、ブランチゲートウェイ225)を含む。

【０２３２】

[00276] また、上述したように、いくつかの実施形態におけるDNSサーバサービスは、1つ以上のパブリッククラウドで動作し、エンティティのMMCNのDNSサーバにDNS情報を提供するサービスマシンまたはいくつかのサービスマシンのクラスタである。これらの実施形態のいくつかでは、DNSサーバは、仮想ネットワークプロバイダによって、また

は別のエンティティによって操作される。他の実施形態では、DNSサーバサービスは、仮想ネットワークを実施するパブリッククラウドの外部に存在し、処理1800が実行される特定のMMCNの近くにあるパブリッククラウド内のエッジノードを識別できる(例えば、サードパーティの)ジオIPサービス(geo-IP service)である。

【０２３３】

[00277] 次に、1810において、コントローラクラスタはN個のノードへの識別されたリストを特定のMMCNのエッジノードにダウンロードして、MMCNのエッジノードが、リスト内の各N個のMFNへのコネクションの品質を定量化する測定を行うことができるようにする。いくつかの実施形態では、各MMCNエッジノードは、そのような測定を生成する測定エージェント(例えば、MMCNコンピュータの1つで実行されるプロセス)を有する。この測定エージェントは、様々な実施形態において異なる方法で測定値を生成する。いくつかの実施形態では、測定エージェントは、周期的に(例えば、毎秒、N秒毎、毎分、M分毎など)pingメッセージ(例えば、UDPエコーメッセージ)を、リスト内のN個のMFNの測定エージェントの各々に送信する。MMCN測定エージェントは、受信した応答メッセージの速度に基づいて、ネットワーク接続スループット速度、遅延、損失、リンク信頼性などの測定基準値を計算し、更新する。いくつかの実施形態において、複数のMFNは、1つの測定エージェントを共有する(例えば、MFNをホストするパブリッククラウドプロバイダの同じデータセンタまたは近くのデータセンタ内で)。

【０２３４】

[00278] いくつかの実施形態では、特定のMMCNの測定エージェントは、周期的にこれらの測定を実行し、コントローラクラスタがその重み演算および経路生成を更新できるように、1820から1835を参照して後述するように、新しい測定値をコントローラクラスタに周期的に送信する。また、新たに追加されたか、または以前に使用されたパブリッククラウドデータセンタにおいて新しいMFNが追加されるたびに、いくつかの実施形態におけるコントローラクラスタは、N個の候補MFNの更新リストを生成する。

【０２３５】

[00279] 1815で、コントローラクラスタは、特定のMMCNのエッジノードによって取得された測定値を受信する。これらの測定に基づいて、中央コントローラクラスタは、特定のMMCNのエッジノードをN個のMFNの各々に接続する各接続リンクに対して(1820において)リンク重みを計算する。例えば、いくつかの実施形態では、中央制御装置は、遅延測定値に指数フィルタを使用し、重み乗数として損失パラメータを使用する(例えば、損失の各1%に対して重みを2倍にする)ことによって、各リンクの重みを計算する。

【０２３６】

[00280] 次に、計算された加重に基づいて、中央コントローラは、エッジノードが接続される「ホーム」ノードとしてM(例えば、5または6) MFNのサブセットを(1825において)識別する。いくつかの実施形態において、Mノードは、最小の重み値を有するノードである。他の実施形態では、Mノードは、最も低い重み値を有するノードであるが、各クラウドプロバイダ内の少なくとも1つの代表MFNがMノードに含まれる。Mノードのリストは時間と共に変化することがあり、新しいMFNが追加されると、および/または特定のMMCNエッジノードから新しい測定値が受信されると、MFNをドロップしてリストに追加することができる。いくつかの実施形態におけるコントローラクラスタは、M個のMFNのリストにおける頻繁な変更を回避するために「ヒステリシス」プロセスを使用する。いくつかの実施形態におけるヒステリシスプロセスは、MFNリストへのメンバの追加/削除の速度を低減するために、MFNリストの以前の状態(すなわち、以前のメンバー)を使用する。また、いくつかの実施形態では、コントローラクラスタは、別のMFNが測定のウィンドウ(例えば、期間)に対して10%小さい平均重みを有しない限り、リストからMFNをドロップしない。

【０２３７】

[00281] 上述したように、いくつかの実施形態における特定のMMCNエッジノードは、M個のMFNのそれぞれの仮想ネットワークゲートウェイへの安全な接続(例えば、IPsec

接続)を維持する。いくつかの実施形態では、コントローラクラスタは、(1825において)特定のMMCNエッジノードに、M個のMFNの各々との安全な接続を確立するように指示する。1830で、コントローラクラスタは、選択されたM個のMFNの計算された重みを使用して、特定のMMCNエッジノードと、仮想ネットワークを介して他のMMCNエッジノードまたはSaaSプロバイダデータセンタとの間を横断するデータメッセージフローのために、特定のMMCNエッジノードと、互いに可能なノードとを接続するための最適経路およびフェイルオーバー経路を識別する。このような経路を生成するために、いくつかの実施形態におけるコントローラクラスタは、上述したように、最短パス経路識別プロセスを使用する。

【０２３８】

[00282] いくつかの実施形態では、コントローラクラスタは、その経路識別処理を、周期的に、または、計算された重み値が変化するたびに(例えば、M個のMFNのリスト内のMFNの新しい測定または追加/削除に基づいて)、繰り返す。いくつかの実施形態では、コントローラクラスタは、他のMMCNへの複数の接続リンクおよびSaaSプロバイダへの複数の接続リンクが特定のMMCNへおよび特定のMMCNからの最適な経路を識別するのに関連するので、他のマルチホームMMCNおよび/またはマルチホームSaaSプロバイダへの経路識別動作と共に、特定のMMCNエッジノードに対する経路識別動作を(1830において)実行する。これらの計算された経路は、リモートデバイス(リモートラップトップ、デスクトップ、またはスマートフォン、タブレットなどのモバイルデバイス)に接続するための候補である仮想ネットワークMFN との間の経路も考慮する。

【０２３９】

[00283] これらの経路を識別した後、コントローラクラスタは、特定のMMCN エッジノードとMFN に1つ以上の経路の転送記録を(1835 で) 供給する。例えば、いくつかの実施形態では、コントローラクラスタは、特定のMMCNエッジノードおよびMFN CFEへの転送記録(例えば、次ホップを指定するルーティング記録、仮想ネットワーク出力ノードを指定するルーティング記録など)を提供する。これらの転送記録を使用してルーティング動作を実行することにより、特定のMMCN エッジノードとMFN CFEは、コントローラクラスタによって(1830 で)定義された最適およびフェイルオーバー経路を実施する。いくつかの実施形態では、コントローラクラスタは、新しい最適またはフェイルオーバー経路を識別するたびに、特定のMMCNエッジノードおよびMFN CFEに新しいルーティング記録を供給する。

【０２４０】

[00284] このようにして、図18のプロセス1800は、特定のMMCNと仮想ネットワークへのいくつかの接続のそれぞれとの間のコネクションの品質を表す計算された重み値に基づいて、そのルーティング計算を行う。このアプローチでは、特定のMMCN エッジノードと異なるMMCN、異なるSaaSノード、および/または異なるリモートデバイスに対して、異なる仮想ネットワーク入力/出力ノードペアを選択できる。このため、いくつかの実施形態におけるコントローラクラスタは、上述したように、1つ以上のマルチホームMMCNおよび/またはマルチホームSaaSプロバイダのための経路識別動作(すなわち、動作1830)を共に実行する。

【０２４１】

[00285] 図19は、2つのMMCN1905および1910とSaaSデータセンタ1915の2つのブランチノードEN1およびEN2の例を示している。各ブランチノードは、2つまたは3つのパブリッククラウドプロバイダの3つのパブリッククラウドデータセンタで定義された仮想ネットワークMFN 1925 から1935を介して仮想ネットワーク1920に接続する。一方、SaaSデータセンタ1915は、仮想ネットワークMFN1940および1945を介してアクセスすることができる。関連する分岐ノードEN1とEN2、MFN 1925から2545 とSaaSデータセンタ1915 の間で測定された重みは、これらのノード間のリンク上に描かれている。この例では、ノード1925と1935の間のような他の重みがはるかに高く(例えば10)、最短パスルーティングアルゴリズムが最適コストパスでそれらを使用しないであろうと想定

される。

【０２４２】

[00286] この例から理解されるように、EN1からSaaSデータセンタへの最適パスは、ノード1925 と1940を通過する。これは、この経路の合計の重みが14 であり、他のパスの他の重みコストよりも小さいためである。たとえば、ノード1930を通過すると、最初のホップではウェイトが小さくなるが、合計の重みは最小で15になる。ブランチノードEN2からの最適な経路は、ルータ1935と1945を経由し、合計の重みは15である。その結果、2つのブランチは2つの異なる経路を使用してSaaSデータセンタに到達する。EN1とEN2の間で通信するには、MFN1930を経由するのが最善の経路で、合計の重みは13 である。

【０２４３】

[00287] 上述したように、いくつかの実施形態は、2つ以上の仮想ネットワークMFNを各SaaSプロバイダのデータセンタに関連付ける。SaaSは、サードパーティプロバイダがアプリケーションをホストし、インターネット経由で顧客が利用できるようにする、ソフトウェア配布モデルである。SaaSを使用すると、組織は、独自のコンピュータまたは独自のデータセンタにアプリケーションをインストールして実行する必要がなくなる。これにより、ソフトウェアのライセンス、インストール、サポートだけでなく、ハードウェアの取得、プロビジョニング、保守にかかる費用を排除できる。また、インストールするソフトウェアや、それをサポートするための追加のハードウェアを購入するのではなく、顧客はSaaSオファリング(offering)に加入する。一般的に、従量課金モデルを使用して、このサービスの料金を月単位で支払う。コストを定期的な運用経費に移行することで、多くの企業はより優れた予測可能な予算を行使できる。また、ユーザはSaaSオファリングをいつでも終了して、それらの反復コストを停止することができる。

【０２４４】

[00288] SaaSは高い拡張性を提供する。これにより、より多くのコンピュータをプロビジョニングしたり購入したりすることなく、より多くの、またはより少ない数のサービスや機能に顧客がアクセスできるようになる。新しいバージョンを購入したり、顧客が所有するバージョンを更新したりするのではなく、ソフトウェアを更新する必要がある場合、顧客はSaaSプロバイダに頼って、更新とパッチ管理を自動的に実行することができる。これにより、社内のITスタッフの負担をさらに軽減することができる。SaaSアプリケーションはインターネット経由で配信されるため、ユーザはインターネット対応の任意のデバイスとロケーションからアクセスできる。これらの利点により、SaaSは、顧客のハードウェアを使用して顧客構内にインストールされるパッケージ化されたソフトウェアの非常に一般的な代替手段となった。SaaSプロバイダは、プライベートデータセンタ内の1つ以上のサーバ、またはパブリッククラウド内の1つ以上の領域に存在する1つ以上のサーバでサービスをホストできる。

【０２４５】

[00289] 通常、SaaSプロバイダは、そのサービスのドメイン名(www.myworkday.comなど) によって識別される。多くの場合、SaaSプロバイダのパブリックWebページ(www.workday.com)を実行するサーバと、SaaSアプリケーションを実行するサーバ(www.myworkday.com)とは異なるドメイン名が関連付けられている。このドメイン名は、SaaSアプリケーションサーバのIP アドレスを提供するDNSクエリを介して解決できる。

【０２４６】

[00290] 複数のサーバがある場合、DNSサーバは異なるサーバに関連付けることができる2つの異なる要求に異なるIPアドレスを返すことがある。異なる名前のロジックは、ロケーションベースのものである。SaaSプロバイダがサーバを所有する世界中に複数の領域がある場合、各リクエスタはより近いIPサーバを取得する。同じ領域内では、DNSサービスはロードバランシングの観点から異なるサーバを選択することができる。返されるIPアドレスは、同じ領域内の異なるサーバに関連付けられる。この後者の場合、通常は同じIPサブネットを共有する異なるサーバのIPアドレスが返される。

【０２４７】
[00291] いくつかの実施形態における仮想ネットワークのコントローラクラスタは、既知のIP SaaSアドレスのテーブルを保持する。仮想ネットワークが顧客からパケットを取得する場合、宛先IPは3つの異なるタイプにすることができる。第1に、パケットはエンティティのプライベートなロケーションに関連付けることができる(すなわち、エンティティのプライベートIP空間に宛先アドレスを持っている)。この状況では、いくつかの実施形態における仮想ネットワークは、パケットの宛先アドレスに関連するエンティティの対応する計算ノードにパケットをルーティングする。
【０２４８】
[00292] 第2に、パケットは、仮想ネットワークに知られていないパブリック(プライベートではない) IPアドレスである宛先アドレスを持っている。これらのIP アドレスは、汎用パブリックIPアドレスと呼ばれる。いくつかの実施形態における仮想ネットワークは、このようなパケットを入口仮想ネットワークノードからインターネットに送る。第3に、パケットは、SaaSプロバイダのIPアドレスであることが仮想ネットワークに知られているパブリック(プライベートではない) IPアドレスである宛先アドレスを有する。このようなIP アドレスは、SaaS IPアドレスと呼ばれる。いくつかの実施形態では、このようなパケットは、第1の仮想ネットワークノード(例えば、第1のMFNの第1のCFE)から第2の仮想ネットワークノード(例えば、第2のMFNの第2のCFE)に経路指定され、そこから、いくつかの実施形態では可能な限り最短の方法でSaaS IPに提供される。
【０２４９】
[00293] 図20は、マルチホームSaaSプロバイダに対する経路を定義するために、いくつかの実施形態の中央コントローラクラスタによって使用される処理2000を示す。このプロセスは、SaaS サービスに関連付けられているさまざまなIP アドレスを識別し、異なるコンピュートエンドノードから1つ以上のSaaS プロバイダデータセンタへの最短の経路を識別する。図に示すように、一部の実施形態では、プロセス2000は、コントローラクラスタがSaaSドメイン名リストを受信すると、(2005において)開始する。いくつかの実施形態では、SaaSドメインリストは、パブリッククラウド仮想ネットワークプロバイダの管理者によって提供されるが、他の実施形態では、このリストは、パブリッククラウド仮想ネットワークが仮想ネットワークプロバイダによって定義されるエンティティの管理者によって提供される。次の表に、このようなSaaS リストの例を示す。

	名前
1.	login.adaptiveinsights.com
2.	adobeid-na1.services.adobe.com
3.	athenanet.athenahealth.com
4.	login.bws.birst.com
5.	account.box.com
6.	centrify.com
7.	identity.citrix.com
8.	login.constantcontact.com
9.	account.docusign.com
10.	login.github.com
11.	secure.gooddata.com
12.	app.hubspot.com
13.	login.huddle.net
14.	hub.insidesales.com
15.	login.marketo.com
16.	system.netsuite.com
17.	login.newrelic.com
18.	login.microsoftonline.com
19.	login.okta.com
20.	login.oracle.com
21.	myapps.paychex.com
22.	login.salesforce.com
23.	servicemax.cloudforce.com
24.	hi.service-now.com
25.	auth.tableausoftware.com
26.	login.ultimatesoftware.com
27.	support.veeva.com
28.	login.xero.com
29.	www.zendesk.com

【０２５０】

[00294] 2010において、コントローラクラスタはSaaS ドメインリストをデータベースに保存する。いくつかの実施形態では、このデータベースは、仮想ネットワークプロバイダの管理者および/または仮想ネットワークがデプロイされるエンティティ(例えばテナント)の管理者に、1つ以上のインターフェース(例えば、ウェブサーバインターフェースおよび/またはAPIインターフェース)を介してアクセス可能である。このインターフェースを介して、管理者はSaaSプロバイダおよび/または関連するドメイン名をリストに追加または削除できる。

【０２５１】

[00295] 次に、2015において、コントローラクラスタは、ドメイン名のリスト上で、ドメインに関連付けられたSaaSサーバの可能なIPアドレスをできるだけ多く学習する。そのために、いくつかの実施形態におけるコントローラクラスタは、パブリッククラウド(異なるパブリッククラウド上にデプロイされた1つ以上の仮想ネットワークのためにVNPによってデプロイされる)内の異なる測定エージェント205に、リスト上の各ドメイン名に対するDNSクエリを実行するように指示する。このようなクエリは定期的に(例: 30分ごとに) 繰り返される。測定エージェント205は、各ドメイン名について学習したIPアドレスを(2015において)コントローラクラスタに戻す。

【０２５２】

[00296] SaaSが地理的DNSサービスを使用してクライアントと隣接するサーバを照合している場合、測定エージェントによって異なるIP アドレスが返されることがある。SaaSプロバイダは通常、SaaSサーバとそのロケーションのリストを持つ信頼されたDNSサーバを使用する。このようなDNSサーバは、DNS要求を取得すると、測定エージェントのIP アドレスを受信し、このIP アドレスを使用して地域IP マップを取得して測定エージェントのロケーションを識別し、測定エージェントに最適なサーバのIP アドレスを返す。いくつかの実施形態では、測定エージェントは、仮想ネットワークのエンド計算ノードのIPアドレスも提供し、SaaSプロバイダによって使用されるDNSサーバは、エンドコンピューティングノードのIPアドレスに基づいてIPアドレスを提供する。

【０２５３】

[00297] コントローラクラスタは、返されたIP アドレスとそれに関連付けられたドメイン名を(2020において)データベースに保存する。少なくともいくつかのIP(例えば5つ)が同じIPサブネット(例えば、255以下の異なるアドレスを含むクラスCサブネット)に属する場合、コントローラクラスタはサブネットそれ自体をデータベースに追加する。いくつかの実施形態では、このデータベースは、仮想ネットワークプロバイダの管理者および/または仮想ネットワークがデプロイされるエンティティ(例えばテナント)の管理者に、1つ以上のインターフェース(例えば、ウェブサーバインターフェースおよび/またはAPIインターフェース)を介してアクセス可能である。このインターフェースを介して、管理者はIPアドレスを追加または削除できる。このインターフェースでは、管理者によって追加/削除されるドメイン名に関連付けられたレコードの追加/削除も可能になる。また、いくつかの実施形態では、コントローラクラスタは、ある期間(例えば、毎日、数日ごと、毎週または数週間ごとなど)使用されていると報告されていないIPアドレスをパージする。

【０２５４】

[00298] 2020の後、中央コントローラは、1つ以上のパブリッククラウド(レポーティング領域)におけるレポーティング測定エージェントから受信したレポートされた各IPアドレスについて、レポーティング領域に近い(すなわち、閾値距離以内の)パブリッククラウドのセットを(2025において)識別する。いくつかの実施形態では、2つの領域の近さは、領域間で別々に測定されるネットワーク距離に関して判定される。いくつかの実施形態では、処理2000は、サードパーティのDNSサービスを使用して、各IPアドレスのおよそのロケーションを識別し、次に、識別されたIPアドレスのロケーションを使用して、2つのIPアドレス間の距離を定量化する。報告されたすべてのIP アドレスに対して識別される領域のセットのリストは、IP近傍レポートと呼ばれる。このような動作が行われない場合、IPの近傍レポートでは、すべての仮想ネットワーク領域が各IPアドレスの近くにあると定義される。

【０２５５】

[00299] 2030において、中央コントローラは、異なるパブリッククラウド上にデプロイされた1つ以上の仮想ネットワークのためにVNPによってデプロイされた、デプロイ測定エージェント205に、IP近傍レポートを提供する。その後、各測定エージェントは、測定エージェントと各SaaS プロバイダIPアドレスとの間の距離を定期的に測定する(数分または数時間ごとに1 回など)。この距離は、IP近傍レポートで測定エージェントの近くにあると識別される。いくつかの実施形態では、測定エージェントは、このIPアドレスにある

サーバとのTCP接続を開始するための遅延に関して、このIPアドレスまでの距離を計算する。このIPアドレスを持つサーバが応答すると、その応答までの時間が測定される。最初の応答が計算されると、測定エージェントは、いくつかの実施形態において、TCPコネクションを能動的に終了する。いくつかの実施形態では、測定エージェントは、成功したTCP接続イベントおよび/または損失したパケットの個数もカウントする。他の実施形態における測定エージェントは、上述したいずれかの測定技法のような他の測定技法を使用する。

【０２５６】

[00300] 2035で、コントローラクラスタは、測定エージェントから距離測定値を受信する。次に、2040で、コントローラクラスタは、返された測定値(例えば、各測定エージェントから報告された遅延および損失数)を使用して、各可能な入力MFNおよび/または各可能なMMCNから各SaaSプロバイダ(例えば、各SaaS IPアドレス)への経路を識別する。経路を識別するために、コントローラクラスタは、異なるSaaS IPアドレス、異なるMFN間、および異なるMMCNへの測定に基づいて計算される重み値に依存する、いくつかの実施形態における最短パス経路識別プロセスを実行する。

【０２５７】

[00301] いくつかの実施形態では、コントローラクラスタは、その経路識別処理を周期的に、または、計算された重み値が変化するたびに(例えば、新しい測定またはMFNおよびSaaS IPアドレスの追加/削除に基づいて)、繰り返す。いくつかの実施形態では、コントローラクラスタは、MMCNおよびSaaSプロバイダに関連する複数の出口ノードが、任意の1つのSaaSプロバイダへの最適経路を識別する際に関連することになるので、複数のMMCNおよびSaaS IPアドレスのために(2040において)経路識別動作を実行する。

【０２５８】

[00302] これらの経路を識別した後、コントローラクラスタは(2045で)これらの経路をMMCNエッジノードとMFNに供給する。例えば、いくつかの実施形態では、コントローラクラスタは、転送記録(例えば、次ホップを指定するルーティング記録、仮想ネットワーク出口ノードを指定するルーティング記録など)をMMCNエッジノードおよびMFN CFEに提供する。これらの転送記録を使用してルーティング動作を実行することにより、特定のMMCNエッジノードおよびMFN CFEは、コントローラクラスタによって(2040で)定義された最適経路を実施する。いくつかの実施形態では、コントローラクラスタは、新しい経路を識別するたびに、MMCNエッジノードおよびMFN CFEに新しいルーティング記録を供給する。

【０２５９】

[00303] いくつかの実施形態では、上記のプロセスによって発見されたSaaS IPアドレスは、それらが接続する仮想ネットワークノードへのゼロルーティング距離を有すると想定される(すなわち、仮想ネットワークのパブリッククラウド領域に仮想的に位置すると想定される)。他の実施形態では、パブリッククラウド領域とSaaS IPアドレスとの間のルーティングリンクは、(図19の例に反映されるように)それらに関連する重みを有し、これらの重みは、それらのパブリッククラウド領域からSaaS IPアドレスへのパスに関連するコスト(例えば、測定された遅延および/または損失)を反映する。このアプローチでは、特定のIPアドレスに接続するための最良の領域は、計算された重み値(すなわち、パケット遅延および損失に関して測定されたコスト)が小さい領域である。

【０２６０】

[00304] 複数のパブリッククラウド領域でSaaS IPアドレスを複数のMFN CFEに関連付ける1つの合理的な理由は、SaaSサーバから複数の領域までの距離が、領域間の一般的な距離よりもはるかに小さいことである。さらに、1つのパブリッククラウドで発信されるトラフィックをルーティングするコストが低くなるため、同じクラウド内の出力ノードまで維持される可能性がある。この場合、いくつかの実施形態におけるコントローラクラスタは、最も近い領域からのコスト(例えば、遅延および損失)が何らかのコスト(例えば、遅延および損失)閾値を下回る限り、各SaaS IPを各パブリッククラウド内の少なくとも1つの領域にバインドする。経路識別プロセスは、特定のIPアドレスへの最短パスを計算す

る必要がある場合、まずこのIPアドレスがバインドされている領域を探し、次に各出力ノードからバインドされている領域までの最短パスを計算する。いくつかの実施形態では、ルータ内のルーティングテーブル自体は、データメッセージが宛先出力ノードまでトンネル内で運ばれ、その後トンネル内のIPアドレスまでルックアップするので、外部IPアドレスを含む必要はない。

【０２６１】

[00305] 上述したように、いくつかの実施形態における計算された重み値は、2つのパブリッククラウド領域間、パブリッククラウド領域とSaaSプロバイダデータセンタ間、またはパブリッククラウド領域とテナントのコンピューティングエンドノード間のリンクにおけるパケット遅延および/または損失のコストを考慮する。他の実施形態では、このようなリンクごとに計算された重み値は、パブリッククラウドプロバイダによって課金されるデータスループットコスト、および/またはパブリッククラウドプロバイダによって課金されるコンピューティングコスト(パブリッククラウドでMFN構成要素を実施するために使用されるコンピューティング要素の)など、他のタイプのパラメータに関して計算される。

【０２６２】

[00306] いくつかの実施形態では、エンティティは、エンティティの仮想ネットワークをデプロイするために使用するパブリッククラウドに関する入力をVNPに提供することができる。いくつかの実施形態におけるこの入力はまた、VNPが、1つ以上のパブリッククラウドプロバイダの1つ以上のパブリッククラウド上に、エンティティ(例えば、VNPのテナントのため)のための専用仮想ネットワークを作成する要求を伴う。いくつかの実施形態では、エンティティの入力は、仮想ネットワークが定義されるべきパブリッククラウドプロバイダおよび/またはパブリッククラウド領域を使用するように指定する。併せて、または代替的に、いくつかの実施形態では、この入力は、使用する実際のパブリッククラウドデータセンタ(PCD)を指定する。いくつかの実施形態におけるVNPは、エンティティの入力を通じてエンティティに対して識別されたパブリッククラウドデータセンタのセットを、VNPがエンティティのデータセンタのセットに追加するために望ましいデータセンタとして識別する1つ以上のPCDで補完する。

【０２６３】

[00307] 図21は、特定のセットのパブリッククラウドプロバイダ、特定のセットのパブリッククラウド領域、および/または特定のセットのパブリッククラウドデータセンタ上にこのようなネットワークをデプロイするように要求するエンティティのための専用仮想ネットワークを実施するために、VNPがいくつかの実施形態で専用MFNをデプロイおよび構成するために使用するプロセス2100を示す。エンティティの場合、このプロセス(1) エンティティのマシンセットを接続するために、1つ以上のパブリッククラウドプロバイダのパブリッククラウドデータセンタのセットを識別し、(2) 識別されたパブリッククラウドデータセンタのセット内のエンティティの専用MFNをデプロイし、そして、(3) 識別されたパブリッククラウドデータセンタのセット全体でエンティティのマシンセットを接続する専用仮想ネットワークを実施するように、これらのMFNを構成する。

【０２６４】

[00308] この専用仮想ネットワークは、1つ以上のパブリッククラウドデータセンタのコンピューティングおよびネットワークインフラストラクチャを介して定義される。いくつかの実施形態では、1つのエンティティに対して特別にデプロイされる専用MFNは、複数のエンティティによってデプロイされ使用される共有MFNと同じ属性を持ち、同じ動作(例えば、その転送動作を実行する際のテナント識別子のチェック)を実行する。これらの実施形態における専用MFNと共有MFNとの唯一の違いは、専用MFNは、ただ1つのテナントのためのデータメッセージを処理するために使用され、一方、共有MFNは、複数のエンティティのためのデータメッセージを処理するために使用されることである。

【０２６５】

[00309] 上述したように、いくつかの実施形態におけるMFNは、パブリッククラウドデータセンタにおけるいくつかの異なる構成要素の概念的なグループであり、他のパブリ

ッククラウドデータセンタにおける他のMFN(構成要素の他のグループと共に)と、1つ以上のエンティティに対する1つ以上のオーバレイ仮想ネットワークを確立する。これらの構成要素の1つは、仮想ネットワークを介してパスを実施するために必要な次ホップ転送を実行するクラウド転送要素(CFE) である。また、いくつかの実施形態では、各MFNの構成要素のグループは、MFNのパブリッククラウドデータセンタ内の異なるコンピュータ上で実行され、一方、他の実施形態では、MFNの構成要素のいくつかまたはすべてが、パブリッククラウドデータセンタの1台のコンピュータ上で実行される。いくつかの実施形態におけるMFNの構成要素は、VNPのテナントの仮想マシンまたはコンテナ、あるいはパブリッククラウドプロバイダの他のテナントなど、他のテナントの他のマシンも動作するホストコンピュータ上で実行される。

【０２６６】

[00310] いくつかの実施形態では、処理2100によってデプロイされる仮想ネットワークによって接続されるエンティティのマシンは、任意のパブリッククラウドの外部にあるマシンである。他の実施形態では、エンティティのマシンのいくつかはパブリッククラウド内にあり、他のマシンはパブリッククラウド外に存在する。また、いくつかの実施形態では、エンティティのマシンは、エンティティが特定のSaaS動作のために使用するSaaSプロバイダマシンを含む。

【０２６７】

[00311] エンティティの入力を受け取る前に、プロセス2100は、(2105において)測定エージェントを1つ以上のPCDおよび/またはPCDグループにデプロイし、これらのエージェントは、異なるPCDのペア、異なるPCDグループのペア(例えば、異なるパブリッククラウド領域または利用可能性ゾーン間)、および異なるPCD/PCDグループのペア間のネットワークコネクションの品質を定量化するネットワーク測定を生成するために、メッセージを交換する。2つのPCD、2つのPCDグループ、および/またはPCDとPCDグループ間の接続に対してプロセスが生成するネットワーク測定の例には、この接続で発生する損失、遅延、ジッタ、ならびに接続の信頼性とコストが含まれる。

【０２６８】

[00312] 上述したように、いくつかの実施形態は、2つのPCDおよび/またはPCDグループ内の2つの測定エージェント間で交換されるメッセージの往復時間(RTT)を測定することによって、2つのPCDおよび/または2つのPCDグループ間の接続リンクにおける遅延を測定する。いくつかの実施形態におけるジッタは、遅延(例えば、RTT)の導関数である。いくつかの実施形態における損失は、特定の時間フレームにおいてドロップされたパケットの数から導出される。いくつかの実施形態における信頼性は、損失から導出される、例えば、非可逆接続は、低い信頼性スコアを有する。

【０２６９】

[00313] エンティティのネットワーク管理者から、プロセスは専用の仮想ネットワークをデプロイする要求(2110 で)を受信する。いくつかの実施形態におけるこの要求は、仮想ネットワークを介して接続される必要があるパブリッククラウドの外および内に位置するエンティティのマシンに関連するロケーション(例えば、ネットワークアドレス)に関する情報を提供する。このロケーション情報には、複数のマシンを含むオフィスやデータセンタの属性(名前、ネットワークアドレスなど)が含まれる。

【０２７０】

[00314] オフィスおよびデータセンタにおける個々のエンドコンピューティングノードマシンのロケーションは、いくつかの実施形態において具体的には識別されず、むしろオフィスおよびデータセンタとのそれらの関連付けによって識別される。いくつかの実施形態における提供されたロケーション情報はまた、パブリッククラウドにおける1つ以上のマシンの1つ以上のグループのロケーション、ならびにエンティティによって使用されるSaaSプロバイダのアイデンティティ(および場合によってはロケーション)を識別することができる。

【０２７１】

[00315] いくつかの実施形態では、2110で受信された要求は、仮想ネットワークが定義されるべきパブリッククラウドプロバイダ、および/またはパブリッククラウド領域も指定する。併せて、または代替的に、いくつかの実施形態では、この入力は、使用する実際のパブリッククラウドデータセンタを指定する。この入力は、場合によっては、複数の異なるパブリッククラウドプロバイダ、および/または異なるパブリッククラウドプロバイダのPCD および/またはPCDグループを指定できる。

【０２７２】

[00316] 2110で受信された入力に基づいて、プロセスは、(2115において)仮想ネットワークが定義され得る一組の候補PCDおよび/またはPCDグループを識別する。いくつかの実施形態では、識別されたPCDのセットおよび/またはPCDグループは、2110で受信された入力で明示的に識別されたPCDおよび/またはPCDグループを含み、および/または、この入力で明示的に識別されたパブリッククラウドプロバイダのPCDおよびPCDグループである。プロセスは、いくつかの実施形態において、これらの明示的に識別されたPCDおよび/またはPCDグループのみを含む。

【０２７３】

[00317] しかしながら、他の実施形態では、プロセスは、(2115において)エンティティの仮想ネットワークのために望ましいデータセンタである可能性がある他のPCDおよび/またはPCDグループの識別されたセットを候補PCDおよび/またはPCDグループに含む。具体的には、エンティティにリコメンデーションを提供するために、いくつかの実施形態では、処理2100は、エンティティの入力を通じてエンティティに対して識別されたパブリッククラウドデータセンタのセットを、処理がデータセンタのセットに追加するために望ましいPCDおよび/またはPCDグループとして識別される1つ以上の他のPCDおよび/またはPCDグループで補完する。

【０２７４】

[00318] この他のPCDおよび/またはPCDグループを識別するために、いくつかの実施形態におけるプロセスは、特定のロケーションおよび/または領域(例えば、エンティティおよび/またはエンティティのオフィスおよびデータセンタが存在するロケーションおよび/または領域)のための望ましいPCDおよび/またはPCDグループを識別する、事前に構成されたルールを使用する。これらのルールは、他のエンティティのために以前にデプロイされた仮想ネットワークに基づいて、および/またはVNPが異なるPCDおよび/またはPCDグループにデプロイする測定エージェントによって継続的に取得される測定値に基づいて、いくつかの実施形態で定義される。

【０２７５】

[00319] デプロイされているすべてのパブリッククラウドの測定エージェントによって生成されたネットワーク測定の現在のセットから、プロセスは、2115で識別されたPCDおよび/またはPCDグループのためのネットワーク測定のサブセットを(2120で)選択する。いくつかの実施形態では、任意の所与の時間における接続リンクの任意の1つのネットワーク測定(例えば、遅延)は、過去の時間(例えば、過去数秒、分、時間、日など)にわたってその接続リンクに対して生成された測定値(例えば、遅延)の加重混合平均(weight-blended average)である。このような平均により、測定値が平滑化され、使用される測定値が特定の特定の時点での過渡的なネットワーク条件に依存しすぎないようにする。

【０２７６】

[00320] 次に、識別されたPCDおよびPCDグループに基づいて、プロセスは、図3から5を参照して上述した技術の1つ以上を使用することによって、(2125において)ルーティンググラフを生成する。2125で、プロセスは、ルーティンググラフを使用して、2120で選択された測定値を使用して、識別されたPCDのセット全体にわたってエンティティのマシン(使用するSaaSマシンを含む)を接続するパスのセットを識別するパス探索を実行する。モバイルデバイス接続を考慮するために、パス探索では、専用仮想ネットワーク全体のパスの入力ノードおよび出力ノードとして機能できるすべてのPCD ペア間のパスも識別される。

【０２７７】
[00321] パスを識別するために、いくつかの実施形態におけるプロセスは、2120で選択された測定値を使用して、最短パス探索を実行する。いくつかの実施形態では、プロセス2100は、2組のパス探索を実行する。パス探索の第１のセットは、2110で受信したエンティティの要求で、またはエンティティによって明示的に識別されたPCDおよびPCDグループのみを考慮する。これらの実施形態におけるパス探索の別のセットは、明示的に識別されたPCDおよびPCDグループを考慮するだけでなく、プロセスが2115において潜在的に望ましいと識別するPCDおよびPCDグループも含む。
【０２７８】
[00322] 上述したように、最短パス探索は、生成されたネットワーク測定から重みが導出される、ルーティンググラフのリンクに割り当てられた重みに基づいている。異なるネットワーク測定が異なる実施形態で使用されることを前提として、最短パス探索は、異なる基準(遅延、ジッタ、損失、信頼性など)のいずれか1つ以上に基づいて、仮想ネットワーク内のマシンエンドポイント間(またはエッジノード間)の最適パスを識別する、最小コスト探索である。上述したように、いくつかの実施形態におけるコスト探索基準は、パブリッククラウドプロバイダによって課金される1つ以上の金銭的コスト(例えば、コンピューティングコスト、ストレージコスト、ネットワーキングコスト)も含む。
【０２７９】
[00323] いくつかの実施形態は、異なるエンティティが、プロセスが、異なるタイプの基準の異なる組合せを使用して、そのパス探索を実行する間に、そのパスにコストをかけるように指示することを可能にする。例えば、あるエンティティは、メッセージ遅延を最小化するようにプロセスに指示することができ、別のエンティティは、メッセージジッタを最小化するようにプロセスに指示することができ、さらに別のエンティティは、損失および遅延を最小化するようにプロセスに指示することができる、などである。
【０２８０】
[00324] 2130において、プロセスは、(2110において受信したエンティティ要求を通じて明示的に識別されたPCD/PCDグループと、2115においてプロセスが潜在的に望ましいと識別された他のPCD/PCDグループの両方を使用した)第2のパス探索セットが、(2110において受信したエンティティ要求を通じて明示的に識別されたPCD/PCDグループのみを使用した)第1のパス探索セットによって生成された任意のパスよりも、エンティティのマシンエンドポイントの2つ(モバイルデバイスに接続するためのエンドポイントを含む)を接続するためにより良い任意のパスをもたらしたか否かを判定する。
【０２８１】
[00325] もたらしていない場合、プロセスは、パス探索の最初のセットによって生成されたパスによって使用されるPCD およびPCDグループ内のエンティティの専用MFNを(2135 で) デプロイする。一部の実施形態におけるこれらの専用MFNは、そのエンティティに対するデータメッセージフローのみを転送する。いくつかの実施形態では、専用MFNは、プロセス2100が実行されるエンティティ以外の任意の他のエンティティに対するデータメッセージフローを転送しない。他の実施形態では、これらのMFNはまた、特定のエンティティのための仮想ネットワークを管理するためにVNPが送受信する必要があるVNPデータメッセージフローを運ぶ。いくつかの実施形態では、プロセスは、パス探索の第1のセットによって生成されたパスによって使用される各PCDまたはPCDグループ内のエンティティに対して1つのMFNをデプロイする。他の実施形態では、プロセスは、トラフィック負荷を処理するために追加のMFNが必要な場合に、PCDまたはPCDグループ内のエンティティに対して複数のMFNをデプロイすることができる。
【０２８２】
[00326] パブリッククラウドプロバイダのPCD/PCDグループにMFNをデプロイするために、いくつかの実施形態では、プロセス2100は、PCDのパブリッククラウドプロバイダのAPIを使用して、プロバイダのサーバに、MFNの異なる構成要素(例えば、そのCFEなど)を実施する1つ以上のマシンを追加するように指示する。同様に、本明細書に記載する

他のプロセスは、このようなAPIを使用することによって、PCD/PCDグループにMFNをデプロイする。

【０２８３】

[00327] プロセス2100はまた、これらのMFN(例えば、これらのMFNの転送要素)を、(1)異なるパスに沿ってデータメッセージフローを転送するようにMFNおよびエッジノード(エンティティのマルチマシンコンピューティングノード内)を設定する転送ルール(例えば、次ホップ記録)を定義するために識別されたパスを使用することによって、これらのパスを実施するように(2135において)構成し、(2)これらの転送ルールをMFNおよびエッジノードに配信する。転送ルールを配信するために、いくつかの実施形態におけるプロセスは、ルールを一組のコントローラに提供し、コントローラは、それらをMFNおよびエッジノードに配信する。いくつかの実施形態では、プロセスは、エッジノードがVNPマネージャ/コントローラによって直接アクセス可能でないとき、これらのエッジノードと通信することによって(例えば、サービスAPIまたは何らかの他のメカニズムを介して)、またはボーダーゲートウェイプロトコル(BGP)を使用することによって、エッジノードを設定する。2135の後、プロセスは終了する。

【０２８４】

[00328] プロセス2100が、第2のパス探索セットが、プロセスが2115で潜在的に望ましいと識別する他のPCD/PCDグループの1つを使用する少なくとも1つのより良いパスをもたらしたと(2130で)判定すると、プロセスは、UI (ユーザインターフェース)管理コンソールまたは他のインターフェース(例えば、電子メールなど)を介して表示するためのレポートを生成し(2140で)、エンティティ管理者に、要求内のエンティティによって明示的に識別されたパブリッククラウド、PCDまたはPCDグループのリストに、他のPCDおよび/またはPCDグループを追加するためのリコメンデーションを提供する。いくつかの実施形態では、リコメンドされるPCD/PCDグループは、エンティティによって指定されるものとは異なるパブリッククラウドプロバイダに属することができ、および/または、エンティティによって指定されるPCD/PCDグループとは異なる領域に属することができる。

【０２８５】

[00329] いくつかの実施形態におけるプロセス2100はまた、第2の検索セットがいかなるより良いパスも生成しないときに2135においてMFNをデプロイする前であっても、ユーザインタフェースにおける表示のためのレポートを生成するので、エンティティ管理者は、デプロイされるべき仮想ネットワークに関する情報をレビューし、修正を提供することができる。他の実施形態では、プロセス2100は、2125で生成されたパスに基づいて識別する仮想ネットワークをデプロイする前に、このようなレポートを生成しないが、仮想ネットワークがデプロイされた後にこのレポートを提供する。また、いくつかの実施形態では、処理2100は、専用仮想ネットワークをデプロイする前にリコメンデーションを提供せず、専用仮想ネットワークをデプロイした後に、2110で受信した入力を通じて識別されたPCDおよび/またはPCDグループにまたがるリコメンデーションのみを提供する。

【０２８６】

[00330] 2145において、プロセスは、エンティティの管理者が、レポートが2140において提示した何らかのリコメンデーションを受け入れたか否かを判定する。受け入れていない場合、プロセスは2135 に遷移して、第1のパス探索セットを介して生成されたパスを実施するための専用MFNをデプロイおよび構成する。一方、エンティティの管理者が1つ以上のPCD/PCDグループを追加するための1つ以上のリコメンデーションを受け入れると、プロセスは、第2のパス検索セットによって生成されたパスによって使用されるPCDおよびPCDグループ内のエンティティの専用MFNを(2150 で) デプロイする。いくつかの実施形態では、プロセスは、第2のパス探索のセットを通して識別されるパスによって通過する各PCDまたはPCDグループに1つのMFNをデプロイする。他の実施形態では、プロセスは、トラフィック負荷を処理するために追加のMFNが必要な場合に、このようなPCDまたはPCDグループに複数のMFNをデプロイすることができる。

【０２８７】

[00331] プロセス2100は、デプロイされたMFN(例えば、MFNのCFE)を(2150において)構成し、第2のパス探索のセットによって生成されたパスを実施する。これらのMFNを構成するために、プロセス(1)は、異なるパスに沿ってデータメッセージフローを転送するようにMFNとエッジノード(エンティティのマルチマシンコンピューティングノード内)を設定する転送ルール(例えば次ホップ記録)を定義するために識別された第2の探索セットパスを使用し、(2)これらの転送ルールをMFN(例えばMFNのCFE)とエッジノード(例えばこれらのMFNとエッジノードの転送要素)に配信する。

【０２８８】

[00332] 転送記録を分配するために、いくつかの実施形態におけるプロセスは、一組のコントローラに転送レコードを提供し、コントローラは、次に、それらをMFNおよびエッジノードに分配する。いくつかの実施形態では、プロセスは、エッジノードを、これらのエッジノードと通信することによって(例えば、サービスAPIまたは他の何らかのメカニズムを介して)構成するが、他の実施形態では、エッジノードがVNPマネージャ/コントローラによって直接アクセス可能でないときに、BGPまたは他のルーティングプロトコルを使用する。2150の後、プロセスは終了する。

【０２８９】

[00333] いくつかの実施形態では、VNPプロセスは、それらがデプロイする専用および共有仮想ネットワークの両方に対して同じ転送方式を使用する。例えば、いくつかの実施形態では、これらのプロセスは、専用および共有仮想ネットワークの両方に対して同じ二重カプセル化アプローチを使用する。しかしながら、他の実施形態は、専用仮想ネットワークに対して異なる転送スキームを使用する。例えば、1つのパブリッククラウドプロバイダのPCD/PCDグループに張られた空間専用仮想ネットワークをデプロイする場合、一部の実施形態は、上述のIPベースの二重カプセル化アプローチを使用せず、むしろ、テナント識別子がいくつかの実施形態では専用MFNのためにもはや使用されないので、単一のカプセル化ヘッダを使用してデータメッセージをルーティングする。

【０２９０】

[00334] 異なるエンティティに対して実行される場合、プロセス2100は、このプロセスの開始入力として、パブリッククラウドプロバイダの特定のセット、PCDの特定のセットまたはPCDのグループを使用するための各エンティティの要求であるため、これらのエンティティに対して大きく異なる仮想ネットワークを生成することができる。図22は、3つの異なる企業のために米国のいくつかのパブリッククラウド上にデプロイされた3つの異なる仮想ネットワーク2202、2204、2206を示す例を示している。この例では、最初の2つの企業はサンフランシスコ、パロアルト、ロサンゼルスにのみオフィスを設置しているが、2つの非常に異なる仮想ネットワーク2202 と2204 があり、1つはAmazonのパブリッククラウドインフラストラクチャを選択し、もう1つはGoogleのパブリッククラウドインフラストラクチャを選択している。この例の第3の会社は、サンフランシスコとシカゴのGoogle PCD、およびニューヨークのAmazon PCD にまたがる専用の仮想ネットワーク2206を有する。

【０２９１】

[00335] エンティティのための処理2100を完了した後でも、1つ以上のVNP処理は、エンティティのためにデプロイされたネットワーク条件と仮想ネットワークを監視し、必要とされるかもしれない仮想ネットワークに対する調整を調整またはリコメンドする。例えば、いくつかの実施形態では、これらのプロセスは、測定エージェントが反復的に生成するネットワーク測定値を繰り返しチェックする。これらのネットワーク測定は、パブリッククラウド間およびパブリッククラウドへのコネクションの品質が時間とともに変化する可能性があるときに必要とされる。

【０２９２】

[00336] 新しい測定では、VNP処理は新しいパス探索を実行し、その結果、パブリッククラウドを通る新しい優先パス、および/またはエンティティのための新しい専用MFNのデプロイをもたらす可能性がある。いくつかの実施形態では、VNPプロセスは、これら

の変更を自動的に行う。他の実施形態では、これらのプロセスは、これらの新しいパスを使用するためのリコメンデーションを生成し、および/または新しいMFNをデプロイし、UI管理コンソールまたは他のインタフェース(例えば、電子メールなど)を介してエンティティのネットワーク管理者に提示するためのこれらのリコメンデーションを提供する。

【０２９３】

[00337] また、VNP プロセスは、新しいパス探索を実行したり、仮想ネットワークを調整したり、エンティティ用にデプロイされた専用MFN から収集した統計に基づいて新しいリコメンデーションを生成したりする。これらの実施形態のいくつかでは、VNP処理は、MFNから収集された統計と測定エージェントからの新しい測定の両方に基づいて、新しいパス探索を実行する。これらの新しいパス探索に基づいて、いくつかの実施形態におけるVNPプロセスは、1つ以上のパブリッククラウド、PCDおよび/またはPCDグループを追加するためのリコメンデーションを提供する。収集された統計値および/または新しい測定値の分析に基づいて、いくつかの実施形態におけるVNP処理はまた、使用率の低いパブリッククラウド、PCDおよび/またはPCDグループの削除をリコメンドする。

【０２９４】

[00338] 図23は、収集された統計および新しい測定に基づいてMFNを追加するためのリコメンデーションを生成するVNPプロセス2300を概念的に示す。いくつかの実施形態では、このプロセスは、各エンティティのデプロイされた仮想ネットワークごとに反復的に(例えば、1時間ごと、数時間ごと、1日ごとなど)実行される。図に示すように、プロセス2300は、最初に、エンティティの仮想ネットワークを実施するMFNおよびエッジノードから(2305において)統計を収集する。収集された統計は、MFNおよびエッジノードによるデータメッセージフロー(例えば、データメッセージの数、バイト数、データメッセージの伝送の遅延など)の転送に関する。

【０２９５】

[00339] 次に、2310において、プロセスは、仮想ネットワークのMFNがデプロイされているPCDまたはPCDグループ(すなわち、仮想ネットワークが現在スパンしているPCDまたはPCDグループ)にデプロイされた測定エージェントによって生成された新しい測定値を収集する。これらの測定は、これらのPCDおよび/またはPCDグループ間の接続のためのネットワーク測定である。上述したように、いくつかの実施形態における任意の所与の時間における接続リンクに対する任意の1つのネットワーク測定(例えば、遅延)は、過去の時間(例えば、過去数秒、分、時間、日など)にわたってその接続リンクに対して生成された測定値(例えば、遅延)の加重混合平均である。

【０２９６】

[00340] 2315において、プロセスは、仮想ネットワークを実現するために使用される現在のPCDおよび/またはPCDグループに追加する1つ以上の候補PCDおよび/またはPCDグループを識別し、測定エージェントがこれらの新しい候補PCDおよび/またはPCDグループに対して生成したネットワーク測定値を収集する。これらの測定値は、これらのPCDおよび/またはPCDグループ間の接続のためのネットワーク測定値である。

【０２９７】

[00341] いくつかの実施形態では、候補PCDおよび/またはPCDグループは、仮想ネットワークを実施するために現在使用されているPCDおよび/またはPCDグループの近くまたは間にあるPCDおよび/またはPCDグループである。いくつかの実施形態における候補PCDおよび/またはPCDグループは、各PCDまたはPCDグループ(例えば、仮想ネットワークによって現在使用されているPCDまたはPCDグループ)を他の候補PCDまたはPCDグループと関連付けるルックアップテーブル、またはPCDおよび/またはPCDグループ(ペア内の各PCD/PCDグループは、ペア内の他のPCD/PCDグループ内のMFNとMFN交換データメッセージフローを有する)の各接続されたペアと、他の候補PCDまたはPCDグループとを関連付けるルックアップテーブルから識別される。これらまたは他の実施形態では、候補PCDおよび/またはPCDグループは、仮想ネットワークによって現在使用されているPCDおよび/またはPCDグループのしきい値距離内にあるすべてのPCDおよび/またはPCDグループ

を識別することによって識別される。

【０２９８】

[00342] いくつかの実施形態では、プロセス2300は、仮想ネットワークによって使用される各現在のPCD/PCDグループの新しい候補PCD/PCDグループを識別しない。むしろ、これらの実施形態では、プロセス2300は、輻輳を超えているか、性能が低いか、または性能が低いパス上にある既存のPCD/PCDグループについて、置換候補PCD/PCDグループのみを識別する。 いくつかの実施形態における処理2300は、エンティティが以前に仮想ネットワークに対して明確に拒絶したPCDまたはPCDグループの候補を識別しない。

【０２９９】

[00343] 既存および候補PCD/PCDグループに基づいて、プロセスは、図3から5を参照した上述のアプローチの1つを使用することによって、(2320において)ルーティンググラフを作成する。次に、2325において、プロセスは、エンティティマシンエンドポイント(例えば、支社、データセンタ、モバイルデバイスアクセスロケーションなど)を接続するためのパスを識別するために、パス探索の新しいセットを実行する。これらのパスを識別するために、プロセス2300は、上述したパス探索技術を使用する。

【０３００】

[00344] 2330において、プロセス2300は、新しいパス探索のセットが、仮想ネットワークによって現在使用されている任意のパスよりも、エンティティの2つのマシンエンドポイントを接続するのに適した任意のパスをもたらしたか否かを決定する。あるパスは、それがより良いコスト(例えば、より小さい遅延、より少ないドロップされたパケット、より少ないジッタ、より信頼性、金銭コストなど)を有する場合、いくつかの実施形態において、別のパスより良い。いくつかの実施形態におけるコストは、複数の基準値(例えば、遅延、ドロップされたパケット、ジッタ、信頼性、財務コストなど)をブレンドすることによって生成されるブレンドされたコストである。いくつかの実施形態で使用される基準値は、VNP処理によって選択されるが、他の実施形態では、それらはエンティティによって指定される。また、いくつかの実施形態は、2305で収集された統計を基準値として使用するか、または現在のパスの基準値を生成する一方で、測定エージェントによって生成された測定を使用して、新しく識別されたパスの1つ以上の部分の基準値を生成する。いくつかの実施形態はまた、測定エージェントの測定値を使用して、現在のパスの基準値を少なくとも部分的に生成する。

【０３０１】

[00345] 処理2300が、新たに識別された各パスが、同じ2つのエンティティマシンエンドポイントを接続するための対応する現在のパスより悪いと判定した場合(2330において)、処理は終了する。そうでなければ、プロセスは、UI (ユーザインターフェース)管理コンソールまたは他のインタフェース(例えば、電子メールなど)を介して表示するためのレポートを(2335において)生成し、エンティティ管理者に、評価された基準スコアによって測定される仮想ネットワーク性能を改善する1つ以上の候補PCDおよび/またはPCDグループを追加するためのリコメンデーションを提供する。

【０３０２】

[00346] いくつかの実施形態における基準スコアは、仮想ネットワークデプロイのコストを含む。他の実施形態では、基準スコアはコストを含まない。しかしながら、このようないくつかの実施形態においても、プロセス2300は、仮想ネットワークのコスト見積もりを、それがエンティティのネットワーク管理者に提供するというリコメンデーションと共に(2335において)提供する。2335において、リコメンドされるPCD/PCDグループは、エンティティによって指定されたものとは異なるパブリッククラウドプロバイダに属することができ、および/または、エンティティによって指定されたPCD/PCDグループとは異なる領域に属することができる。

【０３０３】

[00347] 2340において、プロセスは、エンティティの管理者が、レポートが2335において提示した何らかのリコメンデーションを受け入れたか否かを判定する。そうでない

場合、プロセスは終了する。一方、エンティティの管理者が1つ以上のPCD/PCD グループを追加するための1つ以上のリコメンデーションを受け入れると、プロセスは、新しく追加されたPCD および/またはPCD グループ内のエンティティの新しい専用MFNを(2345で) デプロイする。いくつかの実施形態では、プロセスは、そのような各PCDまたはPCDグループに1つのMFNをデプロイする。他の実施形態では、プロセスは、トラフィック負荷を処理するために追加のMFNが必要な場合に、PCDまたはPCDグループにさらに多くのMFNをデプロイすることができる。

【０３０４】

[00348] プロセス2300は、新たにデプロイされたMFN(例えば、新たにデプロイされたMFNのCFE)を(2345において)設定して、1つ以上の以前に使用されたパスよりも良い、新たに識別されたパスの1つ以上を実施する。また、以前にデプロイされたMFN (例えば、 以前にデプロイされたMFN のCFE)を(2345 で) 再設定して、置き換えられる古いパスを削除する。新たにデプロイされたMFNを設定し、以前にデプロイされたMFNを再設定するために、プロセス(1)は、データメッセージフローを所望の古いパスと新しいパスに沿って転送するようにMFNとエッジノード(エンティティのマルチマシンコンピューティングノード内)を設定する転送ルール(例えば、次ホップ記録)を定義し、(2)これらの転送ルールを古いと新しいMFNとエッジノードに配信する。

【０３０５】

[00349] 転送ルール(例えば、次ホップ記録)を配布するために、いくつかの実施形態におけるプロセスは、一組のコントローラに転送ルールを提供し、コントローラは、それをMFNおよびエッジノードに配信する。1つ以上の古いパスが削除されると、以前にデプロイされた1つ以上のMFN が不要になる場合がある。このような場合、プロセスは不必要なMFN の動作を終了する。2345の後、プロセスは終了する。

【０３０６】

[00350] 図24は、その性能を改善するために、仮想ネットワーク2400に新しいPCDを追加する例を示す。この図は、仮想ネットワーク2400の2つの動作ステージ2405および2410を示す。第1の運用ステージ2405は、第1のデプロイ後、仮想ネットワーク2400が、サンフランシスコ、ニューヨーク、およびフィラデルフィアの3つの都市にある2つのパブリッククラウドプロバイダの3つのPCDにまたがることを示している。この仮想ネットワークは、2つの支社2422および2424と、企業の1つのデータセンタ2426とを接続する。

【０３０７】

[00351] 第2の動作ステージ2410は、シカゴにおける第4のPCDを追加するように修正された後の仮想ネットワーク2400を示す。このPCDは、サンフランシスコブランチ2422とニューヨークブランチ2424との間のパスの性能を改善するために追加された。第1のステージ2405に示すように、サンフランシスコとニューヨークのPCD間のポイントツーポイントパスは重み値(例えば、遅延値)12を有し、一方、サンフランシスコからシカゴのPCDへのパスおよびシカゴからニューヨークのPCDへのパスは、統合された重み値10を有する。したがって、この例では、San FranciscoとNew Yorkのオフィス間の接続を改善するために、仮想ネットワークのデプロイ後にシカゴPCDが追加される。しかしながら、この例では、仮想ネットワーク2400は、ニューヨークからサンフランシスコに流れるために、ニューヨークとサンフランシスコのオフィス間の直接接続を使用し続ける。

【０３０８】

[00352] 上述したように、いくつかの実施形態におけるVNPプロセスはまた、収集された統計値および/または新しい測定値の分析に基づいて、利用率の低いパブリッククラウド、PCDおよび/またはPCDグループの削除をリコメンドする。図25は、使用率の低い1つ以上のMFNを削除するためのリコメンデーションを生成するVNPプロセス2500を概念的に示す。いくつかの実施形態では、このプロセスは、周期的に(例えば、1時間ごと、数時間ごと、1日ごとなど)、または各エンティティのデプロイされた仮想ネットワークごとに反復的に実行される。いくつかの実施形態では、プロセス2500は、他のイベント、例

えば、データトラフィックの重大な変化、PCDの動作停止などに基づいて、繰り返し実行される。

【０３０９】

[00353] 図に示すように、プロセス2500は、最初に、エンティティの仮想ネットワークを実施するMFNから(2505において)統計情報を収集する。収集された統計は、特定の期間中のMFNによるデータメッセージフロー(例えば、データメッセージの数、バイト数、データメッセージの伝送の遅延など)の転送に関する。2510において、プロセス2500は、収集された統計を分析して、仮想ネットワークの活用されていないMFNを識別する。いくつかの実施形態では、活用されていないMFNは、特定の期間において、特定の閾値を下回る数のデータメッセージフローを処理するMFNである。他の実施形態では、活用されていないMFNは、特定の期間において、閾値量(例えば、毎秒ビットの閾値量)未満のデータを転送するMFNである。

【０３１０】

[00354] 次に、2515において、プロセスは、2510において、活用されていないMFNを識別したか否かを判定する。そうでない場合、プロセスは終了する。さもなければ、プロセスは、(2520において)識別されたMFNなしでパス探索を実行することによって、識別された活用されていないMFNを削除することを審査する。複数の使用率の低いMFNが識別されると、いくつかの実施形態におけるプロセスは、活用されていないMFNをすべて削除し、削除されたMFNの無い最適パスのセットを識別することによって、ただ1セットのパス探索を実行する。複数の活用されていないMFNを有するような場合、他の実施形態におけるプロセス2500は、1つ以上の利用度の高いMFNの異なる組合せを削除した異なるパス探索のセットを実行する。これらの異なるパス探索により、プロセスは、削除される活用されていないMFNの異なる組合せに基づいて異なるリコメンデーションを提供することができる。

【０３１１】

[00355] 次に、2520で実行されたパス探索に基づいて、プロセスは(2525において)UI(ユーザインターフェース)管理コンソールまたは他のインタフェース(例えば、電子メールなど)を介して表示するためのレポートを生成し、仮想ネットワーク性能を改善した(評価された基準スコアによって測定される)1つ以上の既存のPCDおよび/またはPCDグループを削除するか、または仮想ネットワークコストを改善するための1つ以上のリコメンデーションをエンティティ管理者に提供する。

【０３１２】

[00356] 各PCD/PCDグループに1つのMFNをデプロイする実施形態では、各リコメンデーションで識別されるPCDおよび/またはPCDグループは、活用されていないMFNがデプロイされるPCDおよび/またはPCDグループである。他の実施形態では、いくつかの実施形態における各リコメンデーションは、特定のMFNを削除するための参照によって提供される(例えば、複数のMFNを各PCD/PCDグループにデプロイできる場合)。いくつかの実施形態では、処理2500は、活用されていないMFNの削除をリコメンドするだけでなく、他の基準に基づいて他のMFNの削除をリコメンドする。例えば、いくつかの実施形態では、中間MFNの削除は、中間MFNがその性能を低下させるにつれて、仮想ネットワークのより良い速度または遅延性能をもたらす可能性がある。

【０３１３】

[00357] いくつかの実施形態では、提供されたそれぞれのリコメンデーション(2525で)は、リコメンデーションが受け入れられた場合、仮想ネットワークのためのコスト推定および/または1つ以上の期待性能メトリックを伴う。例えば、リコメンデーションが受け入れられ、仮想ネットワークがこのリコメンデーションに基づいてデプロイされる場合、仮想ネットワークのコストは月あたり$X 未満になるが、マシンエンドポイントの1つ以上のペア間の1つ以上のパスは、1つの基準によって測定される特定のパーセンテージ(たとえば、Y% 以上遅くなる) だけ悪化する。複数のパスが、特定のリコメンデーションにおける1つ以上のMFNの特定の組み合わせの削除の影響を受ける場合、リコメンデーション

は、これらのパスの各々がどのように影響されるかも表す(例えば、各パスについて、そのパスがどれだけ悪いか、あるいは、どれくらい良くなるかを表す基準スコアパーセンテージを提供することによって)。

【０３１４】

[00358] 2530において、プロセスは、エンティティの管理者が、レポートが2525において提示した何らかのリコメンデーションを受け入れたか否かを判定する。受け入れていない場合、プロセスは終了する。一方、エンティティの管理者が1つ以上のPCD/PCDグループを削除するための1つ以上のリコメンデーションを受け入れると、選択したリコメンデーションで削除対象として指定されているMFN が(2535 で) 削除される。削除されたMFNがPCDにデプロイされている1つ以上のマシンによって実施されると、処理2500は、PCDのパブリッククラウドプロバイダのAPIを使用してプロバイダのサーバにMFNを削除するように指示することによって、MFNを削除する。

【０３１５】

[00359] 2535において、プロセスはまた、残りのデプロイされたMFN(例えば、これらのMFNのCFE)および/またはエッジノードを再設定して、1つ以上の新しいパス(2520で識別される)を実施し、それぞれが以前に削除されたMFNを使用した以前のパスによって接続されていた1つ以上のマシンエンドポイントを接続する。上述した他のプロセスと同様に、いくつかの実施形態におけるプロセス2500は、1つ以上の残りのMFNに次ホップ転送ルール(例えば、次ホップ記録)を定義し、配信することによって、MFNおよび/またはエッジノードを再設定する。2535の後、プロセスは終了する。

【０３１６】

[00360] 図26は、このPCD内の活用されていないMFN 2620を削除するために、仮想ネットワーク2600からPCD 2616を削除する例を示す。この図は、仮想ネットワーク2600の2つの動作ステージ2605および2610を示す。第1の運用ステージ2605は、最初のデプロイ後、仮想ネットワーク2600が、サンフランシスコ、ロサンゼルス、サンディエゴの3つの都市にある2つのパブリッククラウドプロバイダの3つのPCD2612、2614、および2616にまたがることを示している。この仮想ネットワークは、企業の3つの支社2622、2624、2626を接続する。

【０３１７】

[00361] 第2の動作ステージ2610は、サンディエゴのPCD2616を削除するように修正された後の仮想ネットワーク2600を示す。このPCDは、MFN2620が活用されていないため、削除されている。このPCDが削除されると、サンディエゴのオフィス2626はロサンゼルスのPCD 2614に接続する。第2のステージ2610に示すように、ロサンゼルスとサンディエゴのオフィス2624と2626の間のパスは10%遅くなっているが、仮想ネットワークの月次コストは30%削減されている。これは、サンディエゴオフィスとの間でやり取りされる少数のデータメッセージフローを考慮すると、許容可能なトレードオフである。

【０３１８】

[00362] いくつかの実施形態のVNPプロセスは、専用仮想ネットワークを有するエンティティが、特定の状況下で、そのエンティティに対して特別にデプロイされていない他のMFNを使用することを可能にする。いくつかの実施形態におけるこれらの他のMFNは、単一エンティティのために特別にデプロイされ、そのエンティティのための専用仮想ネットワークのみをデプロイする専用MFNとは対照的に、異なるエンティティのための異なる仮想ネットワークをデプロイするために使用される、共有MFNである。専用仮想ネットワークを持つエンティティに共有MFNを使用する例の1つは、専用仮想ネットワーク内の特定のロケーションでの輻輳を緩和することである。

【０３１９】

[00363] 具体的には、いくつかの実施形態のVNPプロセスは、エンティティの専用仮想ネットワークの一部(例えば、エンティティの1つ以上の専用MFN)が輻輳しているように見えるか、または輻輳していると思われるとき、1つ以上の共有MFNの一時的な使用を可能にする。いくつかの実施形態では、これらの共有MFNは、VNPプロバイダがデプロイ

したすべての専用仮想ネットワークのトラフィックオーバーフローを処理するためにデプロイする共有仮想ネットワークの一部である。他の実施形態では、共有MFNは、専用の仮想ネットワークを有する1つ以上のエンティティのオーバーフローデータトラフィックを処理するだけでなく、複数の他のエンティティのために複数の個別の仮想ネットワークを定義するためにも使用され、これらの仮想ネットワークのそれぞれは、その関連エンティティのプライマリデータメッセージトラフィックを通過する。このような共有仮想ネットワークの例は、図1から20を参照して上述した。

【０３２０】

[00364] 図27は、データメッセージトラフィックの一部を、エンティティの専用仮想ネットワークから、少なくとも1 組のマシンエンドポイント間の少なくとも1つのパスの少なくとも一部の1つの共有MFN にオフロードするためのリコメンデーションを生成するVNPプロセス2700を概念的に示している。いくつかの実施形態では、このプロセスは、各エンティティのデプロイされた仮想ネットワークごとに反復的に(例えば、1時間ごと、数時間ごと、1日ごとなど)実行される。

【０３２１】

[00365] 図に示すように、プロセス2700は、最初に、エンティティの仮想ネットワークを実施するMFNから(2705において)統計情報を収集する。収集された統計は、特定の期間中のMFNによるデータメッセージフロー(例えば、データメッセージの数、バイト数、データメッセージの伝送の遅延など)の転送に関する。2710で、プロセス2700は、収集された統計を分析して、仮想ネットワークの1つ以上の混雑したMFNを識別する。いくつかの実施形態では、混雑したMFNは、特定の期間において、特定の閾値よりも大きい数のデータメッセージまたは量のデータメッセージを転送するMFNである。

【０３２２】

[00366] 2715で、プロセスは2710で混雑したMFNを識別したかどうかを判定する。識別していない場合、プロセスは終了する。そうでなければ、プロセスは(2720で)パス探索を実行して、混雑したMFNへの負荷を減らすために、識別された各混雑した専用MFNから1つ以上の共有MFN(すなわち、複数のエンティティのために複数の仮想ネットワークをデプロイするために使用されるMFN)にデータメッセージフローを転送する1つ以上のパスを追加することを審査する。データメッセージフローの一部のみを混雑したMFNから流用するために、いくつかの実施形態におけるパス探索は、既存の混雑したMFNから候補共有MFNにリダイレクトされるべきデータメッセージフローの量を制御するための重み(すなわち、閾値など)を定義する。

【０３２３】

[00367] 各転送パスを識別するために、この方法は、エンティティの識別された、混雑した各専用MFNのトラフィックの一部を転送するために、1つ以上の候補共有MFNを識別する。いくつかの実施形態では、候補、共有MFNは、そのような候補、共有MFNが利用可能なときに識別される、混雑した専用MFNと同じPCDまたはPCDグループにある。このような共有MFN候補が利用可能でない場合、VNPプロセス2700は、混雑した、専用MFNのPCDまたはPCDグループにこのようなMFNを1つデプロイするか、または混雑した、専用MFNのPCDまたはPCDグループの近くにあるPCDまたはPCDグループに既存の共有MFN (またはこのようなMFN上にデプロイする)を選択する。

【０３２４】

[00368] 次に、2720で実行されたパス探索に基づいて、プロセスは、UI (ユーザインターフェース)管理コンソールまたは他のインタフェース(例えば、電子メールなど)を介して表示するためのレポートを(2725で)生成し、エンティティ管理者に1つ以上のリコメンデーションを提供して、一部のデータメッセージフローを1つ以上の混雑した専用MFNから共有仮想ネットワークにリダイレクトする。いくつかの実施形態では、提供されたそれぞれのリコメンデーション(2725で)は、そのリコメンデーションが受け入れられた場合、コスト推定および/または仮想ネットワークに対するパフォーマンス期待値を伴う。たとえば、リコメンデーションが受け入れられた場合、仮想ネットワークは月あたり$X 以下の

コストになり、少なくとも1つの特定のエンドポイントマシンのペア間の接続(複数の基準の1つ以上で表される) がY% 改善されると言うことができる。調整により、2つ以上のエンドポイントとマシンのペア間の2つ以上のコネクションのパフォーマンスが向上する場合、いくつかの実施形態におけるリコメンデーションは、各ペアについての改善を別々に表している。

【０３２５】

[00369] 2730 では、プロセスは、レポートが2725で提示したリコメンデーションをエンティティの管理者が受け入れたかどうかを判定する。受け入れていない場合、プロセスは終了する。一方、エンティティの管理者が、データメッセージフローの一部を共有仮想ネットワークに流用するためのリコメンデーションの1つ以上を受け入れると、プロセスは(2735で)専用仮想ネットワークを定義して、データメッセージフローの一部を共有MFNに流用するように再設定し、(2)リダイレクトされたデータメッセージフローをその宛先に転送するように共有MFN(例えば、共有MFNのCFE)を設定したり、専用仮想ネットワークの別のエッジノードにある専用仮想ネットワークに戻したりする。

【０３２６】

[00370] いくつかの実施形態では、プロセス2700は、複数のエンティティのために複数の共有仮想ネットワークをデプロイするために使用される共有MFNにデータメッセージフローのいくつかを流用するように1つ以上のフロントエンドロードバランサを構成することによって、専用仮想ネットワークを再構成する。いくつかの実施形態におけるこのプロセスは、新たに定義された経路を使用して転送ルール(例えば、次ホップ転送ルール)を定義することによって共有MFN(例えば、共有MFNのCFE)を構成し、その後、それらのMFNを構成するために共有MFNに(コントローラを通して)配信する。

【０３２７】

[00371] いくつかの実施形態では、プロセスは、ロードバランサが共有MFNに転送すべきデータメッセージフローの量を指定する重み値と共に、それらの動作を設定するために、フロントエンドロードバランサにそのような転送ルールを提供する。他の実施形態では、プロセス2700は、混雑した専用MFNの負荷を低減するために、これらのMFNが混雑した専用MFNおよび代替共有MFNにフローを分散する方法を制御する新しいWCMP (重みコストマルチパス動作)パラメータを用いて1つ以上の専用MFN(例えば、MFNのCFE)を設定することによって、専用仮想ネットワークを再設定する。2735の後、プロセスは終了する。

【０３２８】

[00372] 図28は、専用仮想ネットワーク2800内の輻輳した専用MFN 2820を介して、一部のデータメッセージフローを共有仮想ネットワーク2850に向ける例を示す。この図は、ロサンゼルス、シカゴ、ニューヨークのエンティティの3つのオフィス2880、2882、2884を接続する仮想ネットワーク2800の2つの運用ステージ2805と2810を示している。第1の動作ステージ2805は、シカゴPCD 2814内の専用MFN 2820が過度に混雑していることを示している。このMFNは、Los Angles PCD 2812とNew York PCD 2816で専用のMFN2822と2824を接続する。

【０３２９】

[00373] 第2のステージ2810は、データメッセージをリダイレクトするためにロサンゼルスPCD 2812内に再構成された専用MFN 2822が、混雑した専用MFN 2820と同じシカゴPCD 2814内の共有MFN 2830を介して、ロサンゼルス専用MFN 2822からニューヨーク専用MFN2824に流れることを示す。共有MFN 2830は、共有仮想ネットワーク2850 の一部である。このネットワークは、第2の企業がシカゴとニューヨークのオフィス2870と2872を接続するために使用される。

【０３３０】

[00374] この例では、シカゴのこの共有MFN 2830は、これらのデータメッセージをニューヨークのMFN 2832に転送し、ニューヨークの専用MFN 2824に提供する。他の例では、シカゴの共有MFN 2830は、これらのデータメッセージをニューヨークの専用M

FN 2824 に転送できる。また、図28の例では、シカゴPCD 2814は、依然として、LAとシカゴオフィス2880と2882の間、NYとシカゴオフィス2884と2882の間、およびニューヨークオフィス2884からLAオフィス2880の間で、他のメッセージを交換するために使用されている。

【０３３１】

[00375] いくつかの実施形態では、データメッセージは、LAからニューヨークに流れ、ロードバランサによってシカゴの共有MFN 2830に転送され、これは、LA MFN 2822、またはこのMFN 2822の部分と分離することができる。さらに別の実施形態では、このロードバランシング動作は、MFN 2822のクラウド転送要素が実行するWCMP動作の一部である。さらに別の実施形態では、このロードバランサをシカゴPCD 2814にデプロイして、専用MFN2820と共有MFN2830との間で負荷の一部を分散させる。他の実施形態は、LAからニューヨークへのすべてのフローを共有仮想ネットワークを介して転送するだけでなく、他の技術を用いて、シカゴの専用MFN 2820と共有MFN 2830との間で負荷を分散する。例えば、LANとニューヨークのMFNは、これら2つのMFNの間でシカゴに送るフローを分配するためにWCMPを実行する。すなわち、重みを使用して、これらのMFNの各々に向けるフローのパーセンテージを指定する。

【０３３２】

[00376] 図28では、共有MFN 2830は中間ノードであり、フローの一部を専用MFN 2820 から離れた位置にリダイレクトするが、これらのフローを別の専用MFN 2824の専用仮想ネットワークに戻す。他の例では、共有仮想ネットワークは、リダイレクトされたデータメッセージフローをエンティティのマシンエンドポイント、例えば、エンティティのオフィスまたはブランチロケーションに配信する。図29は、このような例を示している。この例は、図28の例と似ているが、ニューヨークの共有MFN 2832は、最初にニューヨークの専用MFN 2824を経由せずに、LA オフィスからニューヨークのオフィス2884に直接フローを提供する点が異なる。

【０３３３】

[00377] さらに他の例では、1つ以上の共有仮想ネットワークの1つ以上の共有MFNが、専用仮想ネットワークを通るパスに沿った複数の異なるロケーションでのフローの輻輳を緩和する方法として、いくつかの実施形態で使用される。図30は、このような例を示している。この図の第1のステージ3005は、ニューヨーク州ロサンゼルス、バルセロナ、およびテレアビブの5つのPCD 3011、3012、3014、3016、および3018を使用して、東京とテレアビブの2つのオフィス3080と3082を接続する、専用の仮想ネットワーク3000に沿って5つのホップを作成するパスを示している。また、LAとバルセロナのMFN3020と3024は、あまりに多くのフローを処理するために使用されていることを示している。

【０３３４】

[00378] 第2のステージ3010は、東京オフィスとテレアビブオフィス間で交換されるデータメッセージフローの一部が、LA PCD 3012およびバルセロナPCD3016で共有仮想ネットワーク3090にリダイレクトされることを示している。これらのフローのリダイレクトでは、東部MFN 3021は、特定の数のフローをLA PCD 3012の共有MFN 3060に転送するように設定されている。New Yorkの共有MFN 3062は、これらのリダイレクトされたフローを受信し、New YorkのPCD 3014の専用MFN 3022の専用仮想ネットワーク3000にリダイレクトする。

【０３３５】

[00379] 専用の仮想ネットワーク3000は、東京からテレアビブへのパスに沿ってバルセロナで輻輳しているので、バルセロナの専用MFN 3024は、これらのフローをバルセロナPCD 3016の共有MFN 3064に向けることによって、データメッセージフローの一部を再び共有仮想ネットワーク3090にリダイレクトする。バルセロナ共有MFN 3064は、これらのリダイレクトされたフローを受信してテレアビブ共有MFN3066 に転送し、テレアビブ専用MFN 3026に転送することによって専用仮想ネットワーク3000 にリダイレクトする。

【０３３６】
[00380] 図30の例では、リダイレクションは、ロードバランサを使用することによっていくつかの実施形態で達成されるが、他の実施形態では、MFNのクラウド転送要素が、データメッセージフローのためのWCMP動作を、東京オフィス3080からテレアビブオフィス3082に実行するために使用する重みのセットを通じて達成される。また、この例では、共有仮想ネットワークのオフロード共有MFNは、対応する専用MFN と同じデータセンタにある。前述したように、これは該当するケースはない(つまり、データメッセージフローは、近くのデータセンタなど、他のデータセンタ内の他のMFN にリダイレクトでくる)。
【０３３７】
[00381] エンティティの専用仮想ネットワークのデータメッセージトラフィックの一部を処理するように1つ以上の共有MFNを設定した後、いくつかの実施形態のVNPプロセスは、専用仮想ネットワークのMFNを監視し続けて、以前に輻輳したMFN上の負荷が、1つ以上の共有MFNがもはや不要になるように十分に低下したかどうかを判定する。図31は、このようなプロセス3100を示す。このプロセスは、プロセス2700がエンティティの少なくとも1つの専用MFN が輻輳しており、トラフィックの一部またはすべてを1つ以上の共有MFN にリダイレクトする必要があることを検出した後、VNPマシンによって反復的に実行される。
【０３３８】
[00382] 図に示すように、プロセス3100は、最初に、エンティティの専用仮想ネットワークを実施するMFNから(3105で)統計情報を収集する。収集された統計は、特定の期間中のMFNによるデータメッセージフロー(例えば、データメッセージの数、バイト数、データメッセージの伝送の遅延など)の転送に関する。3110において、プロセス3100は、収集された統計を分析して、プロセス2700が輻輳しているとして以前に識別された専用MFNを混雑していないものとして再分類すべきか否かを判定する。以前に識別された混雑したMFNのこの再分類のために、いくつかの実施形態におけるプロセス3100は、ある時間の間(例えば、プロセス3100の複数の反復に対して)、MFN上のデータメッセージロード(例えば、処理されたデータメッセージフローの数、またはMFNによって処理されたバイトの全体の数)が特定の閾値を下回ることを必要とする。これは、1つのMFN が混雑した状態と混雑していない状態を常に切り替えないようにするために行われる。
【０３３９】
[00383] 3115において、プロセスは、任意のMFNの状態を混雑した状態から混雑していない状態に変化させたか(3110において)を判定する。変化していない場合、プロセスは終了する。そうでなければ、プロセスは、専用仮想ネットワークを(3120 で) 再設定して、そのステータスが(3110 で) 輻輳から非輻輳に変更された専用MFN からいかなるデータメッセージトラフィックもリダイレクトしなくなる。プロセス2700がフロントエンドロードバランサを使用して、データメッセージフローの一部を混雑した専用MFNからリダイレクトする実施形態では、プロセス3100は、プロセス2700が任意の再分類されたMFN(3110)のためにデプロイした任意のフロントエンドロードバランサに新しいロードバランシング基準を配信することによって、その再設定を行う。
【０３４０】
[00384] 一方、プロセス2700が、WCMPパラメータをMFNおよびエッジノードの転送要素に配信することによって、データメッセージフローの一部を混雑した専用MFNから離れてリダイレクトする実施形態では、プロセス3100は、データメッセージを再分類されたMFNに転送するこれらのクラウド転送要素およびエッジノードに新しいWCMPパラメータを配信することによって(3120で)そのリコンフィギュレーションを行う(3110で)。これらの新しいWCMPパラメータにより、これらの転送要素は、再分類されたMFN から離れたデータメッセージフローをリダイレクトしない。 3120では、プロセス3100は、また、専用仮想ネットワークのリダイレクトされたトラフィックを転送するために不要になったVNP共有仮想ネットワークの任意の共有MFN (例えば、任意の共有MFNのCFE)か

ら転送記録を削除する。3120の後、プロセスは終了する。

【０３４１】

[00385] 専用仮想ネットワークが1つ以上の共有MFNを使用することを可能にするいくつかの実施形態のVNPプロセスの別の例は、これらのプロセスが、専用仮想ネットワークが、そのようなMFNの使用がエンティティの全体的な仮想ネットワークのパフォーマンスを改善する場合に、インターネットバックボーン(以下ではバックボーンMFNと呼ぶ)において又は近くで共有MFNを使用することを可能にするときに発生する。一般に理解されているように、インターネットバックボーンは、商用、政府、学術、その他の大容量ネットワークセンターなどを含む多数の企業が所有する複数の冗長ネットワークの集合体である。バックボーンは、通常、コアネットワークとこれらのセンターのルータ間の主要なデータ経路によって定義される。

【０３４２】

[00386] 一部のデータセンタ(例えばEquinuxデータセンタ)は、インターネットバックボーンの一部であるデータセンタの近くにあり、その場所により、インターネットへの非常に高速な接続を提供する。これらのデータセンタと、インターネットバックボーンにサービスを提供する他のいくつかのデータセンタは、限定的なホスティングサービスを提供する。しかしながら、インターネットのバックボーンの位置又はその近くにマシンを設置することは困難であり、1つのエンティティにとってコストがかかる可能性がある。したがって、通常、エンティティは、仮想ネットワークがインターネットバックボーンの近くまたはそこの位置にデプロイされた専用のMFNを持つことを要求しない可能性がある。このような専用MFNを個々のエンティティにデプロイすることさえできない場合がある。しかしながら、いくつかの実施形態におけるVNPは、1つ以上のバックボーンPCDの各々またはバックボーンPCDの近くに1つ以上のMFNをデプロイし、その結果、これらのMFNは、VNPによってデプロイされた共有および専用の仮想ネットワークによって使用可能である。

【０３４３】

[00387] エンティティのためのMFNをデプロイする前または後に、VNP処理は、エンティティの特定の計算ノードを決定する可能性があるパス検索を実行する。インターネットバックボーンの近くまたはそこにデプロイされた1つ以上の共有MFNを経由するパスを使用することが望ましい(すなわち、インターネットバックボーンの近くまたはそこにあるMFNを使用するために、エンティティの専用仮想ネットワークから少なくとも1つ以上のパスの一部分を離れるパスを使用すること)。このような場合、いくつかの実施形態のプロセスは、インターネットバックボーンの位置または近くにデプロイされた共有MFNが使用されるべきであるというリコメンデーションをエンティティに提供する。エンティティがこのリコメンデーションを受け入れると、プロセスは、インターネットバックボーンの位置または近くにある共有MFNを使用する識別されたパスを使用するようにMFNを設定する。

【０３４４】

[00388] 図32は、1つ以上のバックボーンMFNを使用するための専用仮想ネットワークのリコメンデーションを生成するVNPプロセス3200を概念的に示している。いくつかの実施形態では、このプロセスは、各エンティティのデプロイされた仮想ネットワークごとに反復的に(例えば、毎週1回、毎月1回など)実行される。いくつかの実施形態では、プロセス3200は、特にバックボーンMFNの使用を一般的に拒絶した、またはプロセス3200の以前の反復で識別されたすべての実行可能なバックボーンMFNを拒絶したエンティティに対して実行されない。

【０３４５】

[00389] 図に示すように、プロセス3200は、最初に、エンティティのデータメッセージトラフィックの一部を転送するための候補である1つ以上の共有バックボーンMFNを(3205において)識別する。いくつかの実施形態では、プロセス3200は、エンティティの仮想ネットワークが定義されているPCDまたはPCDグループの閾値物理距離内および/また

は閾値ネットワーク距離内(例えば、ある接続遅延内)にある各バックボーンMFNを識別することによって、候補バックボーンMFNを(3205で)識別する。

【０３４６】

[00390] 3210で、プロセスは、ルーティンググラフノードとして、エンティティの仮想ネットワークによって現在スパニングされているPCDおよび/またはPCDグループのMFNを含むだけでなく、3205でルーティンググループノードとして識別される任意の共有バックボーンMFN(すなわち、パス探索に使用されるルーティンググラフ内のノードとして)を含むパス探索を実行する。バックボーンMFNをパス探索にルーティンググラフノードとして含めることにより、プロセスは、バックボーン共有MFNを使用してデータメッセージフローを専用仮想ネットワーク内のあるロケーションからフロー宛先へ、または専用仮想ネットワーク内の別のロケーションへ戻す最適パスを識別しようとする。エンティティの1組のマシンエンドポイントの場合、プロセスのパス探索では、これらの探索によって、専用仮想ネットワークのみを通過することによって同じエンドポイントペアを接続するより良いパスを識別する場合（遅延、損失、ジッタ、信頼性などの1つ以上の基準のセットのいずれかに従って測定された場合）、共有バックボーンMFNを使用するパスは識別されない。

【０３４７】

[00391] 3215で、プロセスは、専用仮想ネットワークのみを通過することによって、同じエンドポイントペアを接続するために識別された他のすべてのパスよりも、マシンエンドポイントの1つのペアに適したバックボーン使用パスを(3210で)識別したかどうかを判定する。識別していない場合、プロセスは終了する。そうでなければ、3220において、プロセスは、UI (ユーザインターフェース)管理コンソールまたは他のインタフェース(例えば、電子メールなど)を介して表示するためのレポートを生成し、エンティティ管理者に、エンティティのデータメッセージフローの少なくとも一部を転送するために1つ以上のバックボーンMFNを使用するための1つ以上のリコメンデーションを提供する。

【０３４８】

[00392] いくつかの実施形態では、提供されたそれぞれのリコメンデーション(3220で)は、そのリコメンデーションが受け入れられた場合、コスト推定および/または仮想ネットワークに対するパフォーマンス期待値を伴う。例えば、リコメンデーションが受け入れられた場合、仮想ネットワークは月に$Xを減らし、特定のマシンエンドポイントペア(いくつかの基準のいずれか1つで表される)の接続速度をY%改善すると言うことができる。調整により、2つ以上のエンドポイントとマシンのペア間の2つ以上のコネクションのパフォーマンスが向上する場合、いくつかの実施形態におけるリコメンデーションは、各ペアについての改善を別々に表している。

【０３４９】

[00393] 3225 では、プロセスは、レポートが3220で提示したリコメンデーションをエンティティの管理者が受け入れたかどうかを判定する。受け入れていない場合、プロセスは終了する。一方、エンティティの管理者が、提示されたリコメンデーションで特定された1つ以上のバックボーンMFNにデータメッセージフローの一部を流用するためのリコメンデーションを1つ以上受け入れると、プロセスは(3230で)新しい設定データ(1)専用仮想ネットワークを再設定してデータメッセージフローの一部をVNPの共有仮想ネットワークとその共有バックボーンMFNに流用し、(2)共有仮想ネットワークとその共有バックボーンMFN (例えば、共有バックボーンMFNのCFE)を設定して、リダイレクトされたデータメッセージフローをその宛先に転送するか、この専用仮想ネットワークの別のエッジノードにある専用仮想ネットワークに戻すように定義する。

【０３５０】

[00394] いくつかの実施形態では、プロセス3200は、MFNおよび/またはエッジノードの1つ以上の転送要素を設定して、データメッセージフローの一部を共有バックボーンMFNに転送することによって専用仮想ネットワークを再設定する。いくつかの実施形態では、このプロセスは、新たに定義されたパスを使用して転送ルール(例えば、次ホップ転送ル

ール)を定義することによって共有仮想ネットワークを設定し、その後、これらのMFNを設定するために共有仮想ネットワークの共有MFNに(コントローラを介して)配信する。3230の後、プロセスは終了する。

【０３５１】

[00395] 図33は、データメッセージの一部が、専用仮想ネットワーク3300から、共有仮想ネットワーク3350およびその共有バックボーンMFN 3355を介してその宛先に向けられる例を示す。この図は、ロサンゼルス、シカゴ、ニューヨークのエンティティの3つのオフィス3380、3382、3384を接続する仮想ネットワーク3300の2つの運用ステージ3305と3310を示している。

【０３５２】

[00396] 第1の運用ステージ3305は、LAおよびシカゴオフィス3380および3382を接続するシカゴPCD 3314内の専用MFN 3320を示す。このMFN 3320は、また、LAオフィス3380からNYオフィス3384にデータメッセージを転送するために、LAおよびNY PCD 3312および3316内の専用MFN 3322および3324を接続する。NYからLAへのデータメッセージの場合、NY専用MFN 3324は、NYとLA PCD3316と3312との間の接続リンクを使用して、データメッセージを直接LA専用MFN3322に転送する。このダイレクトパスは、LAからNYへのダイレクトパスがLAからシカゴ、シカゴからNYへのパスよりも遅いにもかかわらず、NYからLAへのフローの方が高速である。

【０３５３】

[00397] 第2のステージ3310は、共有バックボーンMFN3355およびVNPの共有仮想ネットワーク3350を使用して、LAオフィス3380からNYオフィス3384にフローを転送することを示す。図示のように、LA MFN 3322は、これらのフローをバックボーンPCD 3360内のMFN 3355に提供し、その後、フローを別のバックボーンPCD 3362内の別の共有MFN 3357に転送する。この他のMFN 3357は、NY PCD 3316 の共有MFN 3332にフローを提供し、その後、このPCD のMFN 3324 にフローを転送する。このパスは、LAからNYへのフローに対してシカゴPCDを通過するパスよりも高速であるため、使用される。しかし、シカゴPCDとそのMFNは、LAとシカゴの間のフロー、およびシカゴとニューヨークの間のフローに引き続き使用されている。

【０３５４】

[00398] 図33では、共有バックボーンMFN3355は、フローの一部をLA PCD 3312の専用仮想ネットワークから離れるようにリダイレクトする中間ノードであるが、これらのフローをNY PCD 3316の専用仮想ネットワークに戻す。他の例では、共有バックボーンMFNは、エンティティエッジノード(例えば、ブランチオフィスまたはデータセンタ)に直接接続して、エンティティのマシンエンドポイント、例えば、エンティティのオフィスまたはブランチロケーションから直接フローを受信し、および/またはそれに直接フローを提供する。図34は、このような例を示している。この例は、図33の例と類似しているが、LAオフィスエッジノードは、そのフローをバックボーンMFN 3355に直接提供するのに対し、バックボーンMFN 3357は、これらのフローをNYオフィス3384のエッジノードに直接提供する点が異なる。他の例では、LAオフィスとNYオフィスの一方がバックボーンとの間で直接データメッセージフローを受信し、他方がPCD内のMFNを通してデータメッセージフローを受信する。

【０３５５】

[00399] エンティティが共有MFNを使用できる別の例には、それらのロケーションのエンティティのモバイルデバイスによるリモートロケーションでの共有MFN の使用が含まれる。たとえば、主に1つの領域で運営されているエンティティ(北米のオフィスなどのみ)もあるが、国際的に旅行を行うユーザがいて、それらの遠隔地にある仮想ネットワークを介してエンティティの専用仮想ネットワークにアクセスする必要がある場合がある。このような状況では、いくつかの実施形態のプロセスは、これらの旅行ユーザのモバイルデバイス(例えば、フォン、タブレット、ラップトップなど)が、プロセスが外国の1つ以上のPCD(例えば、ヨーロッパまたはアジアのパブリッククラウド)にデプロイされる1つ以上の

共有MFNを介してその仮想ネットワークにアクセスすることを可能にする。

【０３５６】

[00400] 図35は、いくつかの実施形態が、VNPの共有MFNを介して提供される遠隔地からのモバイルデバイスアクセスサポートを有するエンティティのための専用仮想ネットワークをデプロイするために使用するプロセス3500を示す。図に示すように、プロセス3500は、エンティティの専用仮想ネットワークが定義されるべき領域の第１のセットを(3505で)受信する。いくつかの実施形態では、プロセスは、都市(例えば、サンフランシスコ、LAN、NYなど)、または地域エリア(例えば、北カリフォルニア、南カリフォルニアなど)のリストとしてこれらのエリアを受け取る。他の実施形態では、プロセスは、1つ以上のパブリッククラウドプロバイダのPCDまたはPCDグループのリストとして、これらの領域を受信する。

【０３５７】

[00401] 次に、3510で、プロセスは、エンティティがモバイルデバイスサポートを領域の第1のセットの外部で必要とする領域の第2のセットを受信する。例えば、エンティティが北米にオフィスを持っていても、中国と日本に定期的に移動する従業員がいる場合、専用の仮想ネットワークを複数の米国の都市または地理的領域にある複数のパブリッククラウドデータセンタにデプロイするように指定できるが、モバイルデバイスに、中国および日本の特定の都市または地理的領域からこの専用の仮想ネットワークへのアクセスを要求することもできる。

【０３５８】

[00402] 識別された第1および第2の領域のセットにおいて、プロセス(3515)は、次に、専用仮想ネットワークがまたがるべき、または遠隔地でモバイルデバイスアクセスが利用可能であるべきPCDまたはPCDグループを識別する。いくつかの実施形態では、PCDおよび/またはPCDグループは、3505および3510で提供されるユーザ入力で具体的に識別されるが、他の実施形態では、PCDおよび/またはPCDグループのいくつかまたはすべては、エンティティによって識別されない。

【０３５９】

[00403] 次に、3520において、プロセスは、専用仮想ネットワークおよびパスを定義するためのパスと、モバイルデバイスの遠隔ロケーションアクセスをサポートする共有MFNからこのネットワークへのパスとを識別するために一連のパス探索を実行する。これらのパス探索を実行するために、プロセスはルーティンググラフを生成し、上述のパス探索技術(例えば、最短パス技術)を使用する。VNPの共有MFNを介して遠隔地からこのネットワークへの専用仮想ネットワークおよび遠隔アクセスを実施するためのパスを識別した後、プロセス3500は、識別された第1の領域のセットにおけるPCDおよび/またはPCDグループの識別されたセット内に専用MFNをエンティティに対して(3525で)デプロイする。

【０３６０】

[00404] プロセスはまた、これらのPCDおよび/またはPCDグループが以前にそのようなMFNを持たなかった場合、識別された第2の領域セット内における識別されたPCDおよび/またはPCDグループ内に(3525で)共有MFNをデプロイする。共有MFNが、識別された第2の領域セットの識別されたPCDおよび/またはPCDグループに既にデプロイされている場合、プロセスは(3525で)これらの共有MFNを識別する。3525で、プロセスは、また、専用仮想ネットワークを実施するために識別されたパス(3520で)を実施するために、デプロイされた専用MFN(例えば、これらの専用MFNのCFE)に転送ルール(例えば、次ホップ転送ルール)を配信する。また、VNP共有MFN からエンティティの専用仮想ネットワークへのパスを実施するために、(3525 で)共有MFN(3525 でデプロイおよび/または識別された)転送ルール(次ホップ転送ルールなど)にも配布する。いくつかの実施形態では、これらの転送規則は、これらの共有MFNのCFEを、3520で識別される1つ以上のパスを実施するように設定する。

【０３６１】

[00405] 専用および共有MFN が適切な転送ルールで設定されると、プロセスは、識別

された第2の領域のセットの1つでPCDまたはPCDグループに接続するモバイルデバイスから専用の仮想ネットワークにアクセスする(3530 で) 要求を受信できる。接続するために、モバイルデバイスは最初にDNS 動作を実行して、専用仮想ネットワークに接続するための最適なロケーションを識別する。モバイルデバイスは、専用仮想ネットワークがスパンしない第2の領域セットの1つにあるため、DNS サーバは、モバイルデバイスが接続する第2のセットの領域で共有MFN のネットワークアドレスを提供する。

【０３６２】

[00406] その後、モバイルデバイスはこの共有MFN に接続し、認証/承認プロセスを経由して専用の仮想ネットワークにアクセスする。モバイルデバイスの認証/認可を処理した後(例えば、モバイルデバイスの認証/認可クレデンシャルを認証/認可サーバに転送する)、共有MFNは(3535で)データメッセージを、その設定された転送ルールに基づいて、モバイルデバイスから専用の仮想ネットワークに転送する。この共有MFNは、専用仮想ネットワークがスパンするPCDおよび/またはPCDグループにデータメッセージフローを直接提供することができ、またはVNPの1つ以上の他の共有MFNを通してこれらのフローを間接的に提供しなければならないかもしれない。3535 の後、プロセス3500は終了する。

【０３６３】

[00407] いくつかの実施形態では、共有MFNは、専用MFNとフローを交換する前に、同じPCDまたはPCDグループ内になければならない。他の実施形態では、共有MFNは、異なるPCDまたはPCDグループ内の専用MFNとフローを交換することができる。これらの他の実施形態のいくつかでは、共有MFNは、両方のMFNが同じパブリッククラウドプロバイダの施設で動作する限り、別のPCDまたはPCDグループ内の専用MFNとのみフローを交換することができる。

【０３６４】

[00408] モバイルデバイスからのデータメッセージがエンティティの専用仮想ネットワークに到達すると(つまり、エンティティにデプロイされた専用MFN に到達すると)、このネットワークを介してマシンエンドポイントの宛先に転送される。これは、ブランチオフィスまたはデータセンタマシン(ブランチオフィスまたはデータセンタエッジノードを介して)、SaaS マシン、またはパブリッククラウドで動作するエンティティのマシンのいずれかになる。モバイルデバイスが接続するリモート共有MFN と、VNPの中間共有MFN も、専用仮想ネットワークに接続されているエンティティマシンからモバイルデバイスにデータメッセージフローを渡す。

【０３６５】

[00409] いくつかの実施形態では、VNPは、専用仮想ネットワークへの直接マシンアクセス(モバイルまたはその他)とは異なる方法で、遠隔地から専用仮想ネットワークへのモバイルデバイスアクセスに課金する。例えば、いくつかの実施形態では、専用および共有MFNは、データメッセージフローに関する統計を維持し、これらの統計をVNPのサーバセットに提供する。次に、このサーバーセットは、共有MFNを使用しないデータメッセージフローの請求を計算するために1セットの請求を使用し、共有MFNを使用するデータメッセージフローの請求を計算するために別の異なる請求セットを使用する。いくつかの実施形態では、VNPは、専用仮想ネットワークへの直接マシンアクセスとは異なる方法で、遠隔地から専用仮想ネットワークへのモバイルデバイスアクセスを課金する。

【０３６６】

[00410] 図36は、北米で定義された専用の仮想ネットワーク3600に接続された中国のモバイルデバイス3605の例を示している。図示のように、モバイルデバイス3605は、日本のPCD 3614内の別の共有MFN 3615およびLAのPCD 3616内の共有MFN 3620を介して専用仮想ネットワーク3600に接続するために、中国のPCD 3612内のVNP共有MFN 3610に接続する。PCD 3616内の共有MFN 3620は、LA内の同じPCD 3616内の仮想ネットワーク3600の専用MFN 3625に接続する。そこから、モバイルデバイス3605は、専用の仮想ネットワーク3600に接続される。他の実施形態では、共有MFN 3610は、モバイルデバイス3605から専用MFN 3625または共有MFN 3620にデータメッセージを直

接転送するように構成される。具体的には、いくつかの実施形態は、エンティティの仮想ネットワークを実施する領域外の外部領域内の1つの共有MFNのみを設定して、外部領域内またはその付近のモバイルデバイスからエンティティの仮想ネットワークにデータメッセージフローを転送することができる。

【０３６７】

[00411] 図37は、VNPが専用の仮想ネットワークをデプロイするエンティティの一組のマシンからSaaSプロバイダへの最適パスを識別するために使用するプロセス3700を示す。エンティティから、プロセス3700は、エンティティの専用仮想ネットワークが定義されるべきPCDのセットおよび/またはPCDグループを(3705で)最初に受信する。たとえば、LA、サンフランシスコ、ニューヨークにオフィスがある企業では、これら3つの都市にあるAmazonのPCDを使用して仮想ネットワークを定義するように指定する場合がある。

【０３６８】

[00412] 3710 では、プロセスは、これらのエンドポイントのマシンエンドポイントとネットワークアドレスのリストも受信する(使用可能な場合)。例えば、上記の例では、いくつかの実施形態におけるプロセスは、LA、サンフランシスコおよびニューヨークの支社および/またはデータセンタのネットワークアドレスを受信する。3710において、いくつかの実施形態におけるプロセスは、エンティティがモバイルデバイスアクセスサポートを望む領域の識別も受信する。いくつかの実施形態におけるエンティティは、さらに、エンティティによって使用されるSaaSプロバイダのリストを(3715において)提供する。いくつかの実施形態では、このリスト上のSaaSプロバイダは、それらのドメイン名を通じて、または上述したSaaSプロバイダを識別するための他の方法のいずれかを通じて識別される。

【０３６９】

[00413] 次に、3720において、プロセス3700は、専用仮想ネットワークを定義するためのパスを識別するために、一連のパス探索を実行する。これらの識別されたパスは、異なるロケーションにあるエンティティのマシンエンドポイントを接続するだけでなく、これらのマシンをSaaS プロバイダマシンに接続する。これらのパス探索を実行するために、プロセスはルーティンググラフを生成し、上述のパス探索技術(例えば、最短パス技術)を使用する。

【０３７０】

[00414] また、SaaSプロバイダマシンエンドポイントへの最適パスを識別するために、プロセス3700は、上述の方法論を使用する。しかしながら、プロセス3700は、エンティティの専用仮想ネットワークのためにエンティティによって指定されたPCDおよび/またはPCDグループの文脈でこの方法論を使用するので、この方法論は、VNPプロセスが特定のPCDおよび/またはPCDグループのセットを選択しなかったエンティティを有するSaaSプロバイダマシンエンドポイントのセットを識別する、またはVNPプロセスが、エンティティの指定されたPCDおよび/またはPCDグループのセットと同じ領域で定義された共有仮想ネットワークのためにSaaSプロバイダマシンエンドポイントのセットを識別するよりも、エンティティのための異なるSaaSプロバイダマシンエンドポイントのセットを識別することができる。PCDおよび/またはPCDグループの所望のセットに関する特定のエンティティ入力が、別のセットのSaaSプロバイダマシンエンドポイントの使用をもたらし得る方法の一例を、図38を参照して説明する。

【０３７１】

[00415] このネットワークを介して専用仮想ネットワークおよびSaaSプロバイダアクセスを実施するためのパスを識別した後、プロセス3700は、エンティティに対して、指定されたPCDおよび/またはPCDグループ内の専用MFNを(3725で)デプロイする。3725で、プロセスは、専用仮想ネットワークを実施するために識別されたパス(3725で)を実施するために、デプロイされた専用MFN(例えば、これらの専用MFNのCFE)転送ルール(例えば、次ホップ転送ルール)を配信する。3725 の後、プロセス3700は終了する。

【０３７２】

[00416] 図38は、エンティティが仮想ネットワークをデプロイするために選択するPCDに基づいてSaaSデータセンタが選択される例を示している。この例では、4つのPCDが示されている。エンティティによって選択された2つのPCD 3802 と3804、およびエンティティによって選択されていない2つのPCD 3806と3808である。いくつかの実施形態において、選択されたPCDは、エンティティによって選択された1つのパブリッククラウドプロバイダに属し、一方、選択されなかったPCDは、エンティティによって選択されなかった別のパブリッククラウドプロバイダに属する。エンティティがPCD 3806および3808を選択した場合、SaaS プロバイダへの最適パスは、オフィス3803のエッジノード3801からSaaS データセンタ2へのこれら2つのPCD 3806 および3808を介したものになる。これは、このパスのコストが4 しかないためである。しかしながら、これら2つのPCDが選択されなかったので、SaaSデータセンタ1が選択され、それへの最適パスはPCD3802と3804のMFN3812と3814を経由する。このパスは合計コストが5であり、他のパスはコストが6と7である。

【０３７３】

[00417] いくつかの実施形態では、マルチマシンコンピューティングノード(MMCN)のエッジノードは、複数の異なるネットワーク接続(リンクと呼ばれる)を介して、それをインターネットに接続する1つ以上の外部ネットワークに接続するSD-WANエッジノードである。図39は、MMCNのロケーション3905(オフィス、データセンタなど)の複数の計算デバイス3910を複数のリンク3902、3904、および3906を介してインターネット3920に接続する、このようなSD-WANエッジノード3900の例を示している。異なるインターネットサービスプロバイダが各リンクをインターネットに接続することができ、例えば、MPLSリンクはMPLSプロバイダを介して接続し、高速インターネット接続リンク(例えば、ケーブルモデム)は高速インターネットプロバイダのネットワーク(ケーブルプロバイダのネットワーク)を介して接続し、無線リンクは無線電気通信ネットワークプロバイダ等を介して接続する。

【０３７４】

[00418] いくつかの実施形態におけるVNPプロセスは、これらの異なるリンクの利用可能性を考慮し、エッジノードが接続する1つ以上のMFNのセット内の各MFNに対して最良のリンクを使用するように各SD-WANエッジノードを構成する。図40は、いくつかの実施形態のそのようなプロセス4000の1つを示す。各SD-WANエッジノードに対して、プロセス4000は、最初に、近くの候補PCDおよび/またはPCDグループ内の測定エージェントが、SD-WANの各リンクに対するエッジノードへのネットワークコネクション測定(例えば、RTT遅延を測定)を行う(4005において)。

【０３７５】

[00419] いくつかの実施形態における近くの候補PCDおよび/またはPCDグループは、実行可能なPCDおよび/またはPCDグループである各SD-WANエッジノードへの閾値物理的または接続距離内のPCDおよび/またはPCDグループである(すなわち、拒絶されていない、および/または、その専用仮想ネットワークのためのエンティティによって受け入れられていないPCD/PCDグループの一部ではない)。リンクの測定を行うために、測定エージェントには、各リンクに関連付けられた1つ以上のネットワークアドレス(静的IP アドレスなど) が提供される。いくつかの実施形態におけるこれらの測定エージェントは、それぞれのアドレスを使用して、pingメッセージ(例えば、UDPエコーメッセージ)を送信し、これらのメッセージのRTT遅延および他の特性を測定する。

【０３７６】

[00420] 次に、4010で、プロセスは、各SD-WANエッジノードに対して1つ以上のリンクのセットを有するルーティンググラフを生成する。近隣の候補PCDおよび/またはPCDグループ(すなわち、各候補PCD/PCDグループ内のMFN)ごとに、プロセス4000は、ルーティンググラフのために各SD-WANエッジノードの1つのリンクを(4010で)指定する。各PCD/ノードペアの指定されたリンクは、単一の基準値(例えば、遅延、損失など)または2つ以上の基準のブレンドされた基準値に基づいて測定される場合、ペアにとって最善の

リンクである。他の実施形態では、プロセスは、ルーティンググラフに、各エッジノードとその候補PCD/PCDグループとの間の複数のリンクを含む。

【０３７７】

[00421] 4015で、プロセスは、測定エージェントによって生成された測定値(各SD-WANエッジノードの異なるリンクに対する異なる測定値を含む)を使用するパス探索を実行して、PCDおよび/またはPCDグループのセットにわたってエンティティのマシン(使用するSaaSマシンを含む)を接続するパスのセットを識別する。モバイルデバイス接続を考慮するために、パス探索では、専用仮想ネットワーク全体のパスの入力ノードおよび出力ノードとして機能できるすべてのPCD/PCDグループペア間のパスも識別される。これらのパス探索は、2つのマシンエンドポイント間の最小コストパスを見つけようとする最短探索パス探索であり、ここで、パスコストは、1つ以上の基準値(例えば、遅延、ジッタ、損失、信頼性など)で表現される。

【０３７８】

[00422] 仮想ネットワークの異なるマシンエンドポイントを接続するための最適パスを識別した後、プロセス4000(4020で)は、識別されたパスによってスパニングされたPCDおよび/またはPCDグループ内のエンティティの専用MFN(例えば、これらのMFNのCFE)をデプロイし、設定する。これらのMFNを設定するために、プロセス(1)は、パスに沿ってデータメッセージフローを転送するために使用する、MFNの転送ルール(例えば、次ホップ記録)を定義するために識別されたパスを使用し、(2)これらの転送ルールをMFNに(例えば、これらのMFNのCFEに)配信する。転送ルールを配信するために、いくつかの実施形態におけるプロセスは、ルールを一組のコントローラに提供し、コントローラは、それらをMFNに配信する。プロセス4000は、いくつかの実施形態において、エンティティのための共有仮想ネットワークをデプロイするために使用される。このような実施形態では、プロセス4000は、4015で識別されるパスを実施するために(4020で)共有MFNを構成する。

【０３７９】

[00423] また、プロセス4000は、データメッセージフローを、4015で識別されるパスによってスパニングされるPCD/PCDグループ内の特定のMFNへの特定のリンクに沿って特定の宛先に転送するように、SD-WANエッジノードを構成する。これを行うために、プロセスには、MFN のクラウド転送要素が、使用する転送記録のセットをエッジノードに提供するためにBGPプロトコルを使用して、これらのエッジノードを設定する。BGPプロトコルを使用して、MFNはエッジノードに経路と経路属性を提供し、これによりエッジノードは(例えば、重み属性、ローカル優先属性などに基づいて)それらが使用するための最良の経路を識別することができる。

【０３８０】

[00424] 図41は、仮想ネットワーク内の宛先エンドポイントへの最適パスに最適なリンク/MFN ペアを使用するようにSD-WAN エッジノードを構成する例を示している。この例では、メキシコシティのオフィス4102は、2つのパブリッククラウドプロバイダAとBの2つのデータセンタセットを介してオレゴンのデータセンタ4104にアクセスする必要がある。オフィス4102には、2つの異なる接続リンク4112 と4114を持つエッジノード4105があり、これはMPLSリンクとケーブルモデムリンクにすることができる。

【０３８１】

[00425] 示されているように、これらのリンクの各々は、それに関連するコスト(すなわち、リンク4112については1のコスト、リンク4114については2のコスト)を有する。これらのコストは、いくつかの実施形態において、PCP AおよびBのメキシコシティPCD4122および4124内の測定エージェントによって生成され、さらに示されるように、メキシコシティオフィス4102とオレゴンデータセンタ4104との間の3つの異なるパスが探索される。これらのパスのうち2つは、PCP A の場合はより適切なリンクであるリンク4112を通過し、1つのパスはPCP B の場合はより適切なリンクであるリンク4114を通過する。

【０３８２】
[00426] これら3つのパスのうち、PCP A のメキシコシティPCD 4122 とPCP A のオレゴンPCD 4126を通過するパス4150は、メキシコシティオフィスからオレゴンデータセンタ4104へのデータメッセージフローの最適なパスである。示されているように、このパスは、1つ以上の基準値のセットに基づいて計算される場合、4のコストを有する。PCP AのメキシコシティPCD 4122 とPCP AのSF PCD 4128を通過するパスのコストは5であるが、PCP BのメキシコシティPCD 4124とPCP BのオレゴンPCD 4132を通過するパスのコストは7である。したがって、いくつかの実施形態におけるメキシコシティPCD 4122のMFN 4130は、BGPプロトコルを通じて、リンク4112およびMFN 4130を介してオレゴンデータセンタ4104にフローを転送するようにエッジノード4105を設定する。
【０３８３】
[00427] 多くの実施形態について、仮想ネットワークプロバイダのプロセスおよび動作を参照することによって説明した。当業者は、上述のプロセスおよび動作の多くが、そのPCDおよび/またはPCDグループの複数にまたがる仮想ネットワークを定義するために、1つのパブリッククラウドプロバイダによって他の実施形態で実行されることを認識するであろう。あるいは、いくつかの実施形態では、各エンティティは、これらのプロセスおよび動作を実行するネットワークマネージャ/コントローラを操作して、1つ以上のPCPの1つ以上のPCD/PCDグループ上にエンティティの仮想ネットワークをデプロイする。
【０３８４】
[00428] このようなネットワークマネージャ/コントローラに対して、いくつかの実施形態は、仮想ネットワークを実施するためにMFNをデプロイ及び設定しない。代わりに、これらの実施形態は、MFNをデプロイおよび設定する他のプロセスに、PCDまたはPCDグループ間の接続を定量化する測定値を提供する。いくつかの実施形態は、これらの測定値を、パブリッククラウドにアプリケーションマシンをデプロイし、その測定値を使用して、そのようなデプロイに最適なロケーションを識別するプロセスなど、他のクラウドベースの動作を実行する他のプロセスに提供する。他のプロセス(例えば、エンティティのネットワークマネージャ/コントローラ)は、他のクラウドベースの動作自体(例えば、パブリッククラウドにアプリケーションマシンをデプロイし、そのようなデプロイに最適なロケーションを識別するために測定を使用する)を実行するために、これらの測定値を使用する。
【０３８５】
[00429] 図42は、いくつかの実施形態の測定インフラストラクチャによって生成された測定に基づいて、1つ以上のリソースのセットをパブリッククラウド内の最適なロケーションにデプロイするために使用するプロセス4200を示す。デプロイされるリソースのセットは、コンピューティングリソース(例えば、VM、コンテナ、サーバなど)、ネットワークリソース(例えば、ルータ、スイッチ、ゲートウェイなど)またはサービスリソース(例えば、任意のミドルボックスサービスマシンまたはアプライアンス、例えば、ファイアウォール、ロードバランサなど)のような、任意のタイプのリソースを含むことができる。ただし、以下に示すいくつかの例では、デプロイされたリソースは、アプリケーションまたはアプリケーションのセットを実行するマシンである。いくつかの実施形態において、プロセス4200は、図43に示されているデプロイ位置アイデンティファイア4300によって実行される。
【０３８６】
[00430] 図42に示すように、様々な可能性のあるPCD候補4330の中の1つのPCD内の特定のアプリケーションまたはアプリケーションのセットを実行するために、マシン(例えば、VMまたはコンテナ)をデプロイするための最適なロケーションのために、サーバセット4305内のサーバに対する要求を(4205で)デプロイ位置アイデンティファイア4300が受信したときに、プロセス4200が開始する。いくつかの実施形態におけるサーバセット4305は、1つ以上のコンピューティングマネージャまたはコントローラのセットであり、ネットワーク4350(例えば、LAN、WANまたはインターネット)を介してデプロイ位置アイデンティファイア4300と通信する。

【０３８７】
[00431] いくつかの実施形態における受信された要求は、(1)候補PCD、PCDグループおよび/またはPCPプロバイダのリスト、および/または(2)デプロイされるべきマシンにアクセスするエンティティのエッジノード4345のロケーション(例えば、ネットワークアドレス)を提供する。いくつかの実施形態において、この情報の全部または部分(すなわち、候補PCD、PCDグループおよび/またはPCPプロバイダ、および/またはエッジノードのロケーションのリスト)は、要求と共に提供されない。なぜならば、情報は、以前にエンティティに対して提供されており、したがって、エンティティのための構成データベースに格納されているからである。モバイルデバイスのアクセスロケーションのためのマシンロケーションデプロイ(machine location deployment)を最適化するために、いくつかの実施形態は、エッジノードロケーションを、モバイルデバイスが典型的に接続するPCDまたはPCDグループのロケーションとして定義する。また、いくつかの実施形態におけるエッジノードは、複数のPCDおよび/またはPCDグループにまたがるように定義される仮想ネットワークを介して通信可能に接続される。他の実施形態では、このような仮想ネットワークは、エッジノードを接続するために定義されない。
【０３８８】
[00432] 4210で、デプロイ位置アイデンティファイア4300は、(1)要求されたリソースセットをデプロイするための、または(2)デプロイされたリソースセットに到達するために通過するための、候補PCDおよび/またはPCDグループであるPCDおよび/またはPCDグループの異なるペア間のネットワーク接続測定のセットに対して、ネットワーク4350を介して測定アグリゲータ4310に要求を送信する。4205での要求が候補PCDおよび/またはPCDグループのリストを提供しない実施形態(例えば、要求は単にPCDに接続するためのPCPプロバイダおよび/またはノードのネットワークアドレスのリストを提供する)では、プロセス4200は、4210での測定を要求する前に、まず候補PCDおよび/またはPCDグループのリストを識別しなければならない。
【０３８９】
[00433] このような状況でPCDおよび/またはPCDグループの候補を識別するために、いくつかの実施形態では、プロセス4200は、4205で要求と共に受信する入力のセットを使用して、利用可能なPCDおよび/またはPCDグループのリストをフィルタリングし、適切でないPCDおよび/またはPCDグループを破棄する。例えば、いくつかの実施形態におけるサーバセット4305は、受け入れ可能なPCPとしての2つのPCPと、デプロイされたリソースにアクセスするための北カリフォルニアのオフィスに関連するプロバイダのネットワークアドレスを提供する。このような場合、デプロイ位置アイデンティファイアは、2つの許容可能なPCPによって操作されていないすべてのPCDおよび/またはPCDグループを破棄し、その後、北カリフォルニア以外のこれら2つの許容可能なPCPのPCDおよび/またはPCDグループを破棄する。
【０３９０】
[00434] 4210において、デプロイ位置アイデンティファイア4300が測定アグリゲータ4310に送信する要求は、いくつかの実施形態において、異なるPCD/PCDグループとエンティティのエッジノードロケーション(例えば、エンティティのＭＭＣＮロケーション)との間のネットワークコネクションに関連するネットワークコネクション測定のセットも要求する。これらの測定値は、いくつかの実施形態において、PCD/PCDグループとエンティティエッジノードとの間のリンクのコストを表現するために使用される。これらの実施形態のいくつかでは、測定エージェントは、他のPCD/PCDグループ内の他の測定エージェントにpingするだけでなく、エンティティエッジノード内の転送要素にpingして、PCD/PCDグループとエンティティエッジノード間の測定値を生成する。
【０３９１】
[00435] 4215で、デプロイ位置アイデンティファイア4300は、候補PCDおよび/またはPCDグループに対して要求された測定値を、測定アグリゲータ4310から受信する。上述したように、いくつかの実施形態の測定装置は、多数のPCDおよび/またはPCDグルー

プ4330のそれぞれに測定エージェント4335を含み、そのような各測定エージェントは、そのそれぞれのPCDまたはPCDグループと他の測定エージェントのPCD/PCDグループとの間のネットワークコネクションの品質を定量化するために、ネットワーク測定値を連続的に生成する。そのようなネットワーク測定値(例えば、遅延測定、ジッタ測定、損失測定、信頼性測定など)の生成は、上述されている。いくつかの実施形態では、測定エージェントは、測定アグリゲータによってデプロイされる。他の実施形態では、測定エージェントは、仮想ネットワークプロバイダまたは別のエンティティのコントローラおよび/またはマネージャによってデプロイされる。

【０３９２】

[00436] 測定エージェント4335は、生成したネットワーク接続測定を測定アグリゲータ4310 に提供する。いくつかの実施形態では、測定エージェント4335は、これらの測定値をアグリゲータ4310に周期的にプッシュ(転送)するが、他の実施形態では、エージェント4335からこれらの測定値を周期的にプル(取得)する。報告期間は、いくつかの実施形態において構成可能であるが、典型的には、毎時、数時間(例えば、6または12時間)、毎日等に設定される。

【０３９３】

[00437] いくつかの実施形態では、測定アグリゲータは、新しい測定がペア間のネットワークコネクションの品質を単に定義しないことを保証するために、PCD/PCDグループペア(例えば、2つのPCD、1つのPCDと1つのPCDグループ、または2つのPCDグループ)間のネットワークコネクションのための新しい測定をブレンドするために、加重アグリゲーション定式(例えば、加重和)を使用する。測定アグリゲータ4310は、他の実施形態において同様に、例えば、複数のPCD/PCDグループにまたがる仮想ネットワークをエンティティに対して定義するためにMFNをデプロイおよび設定する実施形態において使用される。

【０３９４】

[00438] ネットワーク測定値を受信した後(4215で)、プロセスは受信した測定値を(4220で)使用して、要求された一連のリソースをデプロイすべき最適なPCDまたはPCDグループを識別する。これを行うために、プロセス4200は、一連のパス探索を実行して、デプロイされた一連のリソースにアクセスできるすべてのロケーションから、デプロイされたロケーションへのパスの最適な全体セットになるデプロイロケーションを識別する。パスの最良の全体的なセットを識別するために、プロセス4200は、いくつかの実施形態では、1つのタイプの測定値(例えば、接続速度)に基づいてパスのコストを求め、一方、他の実施形態では、複数のタイプの測定(例えば、接続速度およびパケット損失)に基づいてパスのコストを求める。

【０３９５】

[00439] これらのパス探索を実行するために、いくつかの実施形態におけるプロセス4200は、PCD/PCDグループ間のネットワーク接続に関する測定値を使用するだけでなく、PCD/PCDグループ間のネットワーク接続およびデプロイされたリソースにアクセスするエンティティのロケーション(例えば、ネットワークアドレス)に関する測定値も使用する。例えば、図43の例では、デプロイ位置アイデンティファイア4300は、受信した測定のセットとエッジノード4345のロケーションを使用して、エッジノード4356に対して要求されたマシン(アプリケーションまたはアプリケーションセットを実行するための)がデプロイされるべき最適なPCDまたはPCDグループを識別するための一連のパス探索を実行する。

【０３９６】

[00440] 図44は、マシンにアクセスする必要がある特定のエッジノードのセットを有する特定のエンティティのマシンをデプロイするための最適なロケーションをさらに説明する例を示している。この例では、2つのエッジノード4402 および4404 の2台のマシンによるアクセスのためにマシンをデプロイできるPCD候補である8つのPCD 4410から4426がある。また、この例では、エッジノードのマシンとデプロイされたマシンとの間でデ

ータメッセージフローが移動できるように、各PCD に少なくとも1つのMFN 4130をデプロイできることを前提としている。

【０３９７】

[00441] 図44は、PCDの複数のペア間の接続リンクと、リンクの横にある各リンクのコストも表示している。いくつかの実施形態では、2つのPCD間の各接続リンクのコストは、これら2つのPCD内の測定エージェントが2つのPCD間のネットワークコネクションの品質を定量化するために生成するネットワークコネクション測定値から導出される。例えば、PCD4422と4424の間の3のコストは、これら2つのPCD間のネットワーク接続がPCD4410と4412の間の1のコストよりもずっと悪いことを意味する。

【０３９８】

[00442] この例では、デプロイ位置アイデンティファイア4300は、PCD の1つに対するエッジノード4402または4402 ごとにさまざまなパスを探索する。次に、デプロイ位置アイデンティファイアは、2つのエッジノードからの2つのパスのコスト合計が最小になるPCDを識別する。図44において、この識別されたPCDは、エッジノード4402からの経路4440のコストの合計としてPCD 4412であり、エッジノード4404からのパス4445は、エッジノード4402および4404から他の任意のPCDへの任意の他のパスのペアの最小の全体コストである。すべてのパスペアの全体的なコストを以下の表1に示す:

表1

候補 PCD	4402 からのコスト	4404 からのコスト	合計
4410	1	3	4
4416	2	4	6
4422	2	4	6
4412	2	2	4
4418	3	3	6
4414	5	1	6
4420	5	1	6
4424	5	1	6

この例では、PCD 4410および4412は、表1に示されたものと同じ全体コストを有するが、PCD 4412は、他の基準(例えば、コストに基づいて、両エッジノード4402および4404への物理的近接性に基づいて等)に基づいて選択される。

【０３９９】

[00443] 図44に示す例では、各エッジノードは他のエッジノードと同じように扱われる。実際には、一方のエッジノードは、他方のエッジノードよりもはるかに頻繁にデプロイされたマシンにアクセスする必要がある場合がある。これを容易にするために、いくつかの実施形態は、重みを使用して、異なるエッジノードのアクセスコストを区別する。例えば、第1のエッジノードがアクセスの80%を有し、第2のエッジノードがアクセスの20%を有する場合、いくつかの実施形態は、結果のパスコスト値を加算する前に、第2のエッジノードのパスコストを0.2倍し、一方で、第1のエッジノードのパスコストを0.8倍する。

【０４００】

[00444] リソースセットをデプロイするための最適なPCDまたはPCDグループを識別した後、プロセス4225は、処理を呼び出したサーバに最適なPCDまたはPCDグループのI

Dを提供する。このリソースは、その後、識別されたPCD/PCDグループにデプロイすることができる(例えばコンピューティングマネージャ/コントローラによってデプロイおよび設定される)。最適なロケーションを提供した後、プロセス4200は終了する。

【０４０１】

[00445] 上述の特徴およびアプリケーションの多くは、コンピュータ可読記憶媒体（コンピュータ可読媒体とも呼ばれる）に記録された命令のセットとして規定されるソフトウェアプロセスとして実行される。これらの命令が、１つ以上の処理部（例えば、一つ以上のプロセッサ、プロセッサのコア、または他の処理部）によって実行されるとき、それらは、処理部に命令に示された動作を実行させる。コンピュータ可読媒体の例としては、ＣＤ－ＲＯＭ、フラッシュドライブ、ＲＡＭチップ、ハードドライブ、ＥＰＲＯＭ等があるが、これらに限定されない。コンピュータ可読媒体は、有線で又は有線接続を介して通過する搬送波及び電気信号を含まない。

【０４０２】

[00446] 本明細書では、用語「ソフトウェア」は、読出専用メモリに存在するファームウェア、または磁気ストレージに記憶され、プロセッサによって処理するためにメモリに読み取ることができるアプリケーションを含むことを意味する。また、いくつかの実施形態では、複数のソフトウェア発明は、個別のソフトウェア発明を維持する一方で、より大きなプログラムのサブ部分として実施されてもよい。いくつかの実施形態において、複数のソフトウェア発明は、別個のプログラムとして実施することもできる。最後に、ここに記載されるソフトウェア発明を共に実施する別個のプログラムの任意の組合せは、本発明の範囲内である。いくつかの実施形態では、ソフトウェアプログラムは、１つ以上の電子システム上で動作するようにインストールされると、ソフトウェアプログラムの動作を実行し実行する１つ以上の特定のマシン実装を定義する。

【０４０３】

[00447] 上記の説明は、主に、ソフトウェアを実行するマイクロプロセッサまたはマルチコアプロセッサを指すが、いくつかの実施形態は、特定用途向け集積回路(ASIC)またはフィールドプログラマブルゲートアレイ(FPGA)などの1つ以上の集積回路によって実行される。いくつかの実施形態では、そのような集積回路は、回路自体に格納された命令を実行する。

【０４０４】

[00448] 本明細書で使用される「コンピュータ」、「サーバ」、「プロセッサ」、「メモリ」という用語は、すべて電子デバイスまたはその他の技術的なデバイスを指す。これらの用語は、人または人々のグループを除外する。明細書では、用語の表示または表示は、電子デバイス上での表示を意味する。本明細書で使用されるように、「コンピュータ可読媒体」（複数）、「コンピュータ可読媒体」、および「マシン可読媒体」（複数）という用語は、コンピュータが読み取り可能な形式で情報を格納する有形の物理オブジェクトに全体として限定される。これらの用語は、ワイヤレス信号、有線ダウンロード信号、およびその他の一時的または一時的な信号を除外する。

【０４０５】

[00449] 図45は、本発明のいくつかの実施形態が実施されるコンピュータシステム4500を概念的に示している。コンピュータシステム4500は、上述のホスト、コントローラ、およびマネージャのいずれかを実装するために使用することができる。したがって、上記のプロセスのいずれかを実行するために使用することができる。このコンピュータシステムは、様々なタイプの一時的でない機械可読媒体と、様々な他のタイプの機械可読媒体のためのインタフェースとを含む。コンピュータシステム4500は、バス4505、処理ユニット4510、システムメモリ4525、読出専用メモリ4530、永続的記憶装置4535、入力デバイス4540、および出力デバイス4545を含む。

【０４０６】

[00450] バス4505は、集合的に、コンピュータシステム4500の多数の内部デバイスを通信的に接続するすべてのシステム、周辺デバイス、およびチップセットバスを表す。

例えば、バス4505は、処理ユニット4510を読出専用メモリ4530、システムメモリ4525、および永続的記憶デバイス4535と通信可能に接続する。

【０４０７】

[00451] これらの様々なメモリユニットから、処理ユニット4510は、本発明のプロセスを実行するために、処理するための命令を検索し、処理するためのデータを検索する。処理ユニットは、異なる実施形態では、単一プロセッサまたはマルチコアプロセッサであってもよい。読出専用メモリ（ＲＯＭ）4530は、処理ユニット4510およびコンピュータシステムの他のモジュールによって必要とされる静的データおよび命令を格納する。一方、永続的記憶デバイス4535は、読書可能な記憶デバイスである。このデバイスは、コンピュータシステム4500がオフであっても命令及びデータを記憶する不揮発性メモリユニットである。本発明のいくつかの実施形態は、永続的記憶デバイス4535として大容量記憶デバイス（磁気または光ディスクおよびその対応するディスクドライブなど）を使用する。

【０４０８】

[00452] 他の実施形態では、取り外し可能ストレージデバイス（フロッピーディスク、フラッシュドライブなど）を永続ストレージデバイスとして使用する。永続的記憶デバイス4535と同様に、システムメモリ4525は、読書可能な記憶デバイスである。しかしながら、記憶デバイス4535とは異なり、システムメモリは、ランダムアクセスメモリのような揮発性の読み出し／書き込みメモリである。システムメモリには、プロセッサが実行時に必要とする命令とデータの一部が格納される。いくつかの実施形態では、本発明のプロセスは、システムメモリ4525、永続的記憶デバイス4535、および／または読出専用メモリ4530に記憶される。これらの様々なメモリユニットから、処理ユニット4510は、いくつかの実施形態のプロセスを実行するために、実行する命令および処理するデータを検索する。

【０４０９】

[00453] バス4505はまた、入出力デバイス4540および4545に接続する。入力デバイスは、ユーザが情報を通信し、コマンドをコンピュータシステムに選択することを可能にする。入力デバイス4540は、英数字キーボード及びポインティングデバイス（「カーソル制御デバイス」とも呼ばれる）を含む。出力デバイス4545は、コンピュータシステムによって生成された画像を表示する。出力デバイスには、陰極線管（ＣＲＴ）または液晶ディスプレイ（ＬＣＤ）などのプリンタおよび表示デバイスが含まれる。いくつかの実施形態は、入力デバイスおよび出力デバイスの両方として機能するタッチスクリーンなどのデバイスを含む。

【０４１０】

[00454] 最後に、図45に示すように、バス4505は、ネットワークアダプタ（図示せず）を介して、コンピュータシステム4500をネットワーク4565に結合する。このようにして、コンピュータは、（ローカルエリアネットワーク（「LAN」）、ワイドエリアネットワーク（「WAN」）、イントラネット、インターネットなどのネットワークのネットワーク、などのコンピュータのネットワークの一部になることができる。コンピュータシステム4500の任意のまたはすべての構成要素を、本発明と併せて使用することができる。

【０４１１】

[00455] いくつかの実施形態は、機械可読媒体またはコンピュータ可読媒体（代替として、コンピュータ可読ストレージ媒体、機械可読媒体、または機械可読ストレージ媒体）にコンピュータプログラム命令を記憶する、マイクロプロセッサ、ストレージデバイスおよびメモリなどの電子コンポーネントを含む。そのようなコンピュータ可読媒体のいくつかの例には、RAM、ROM、読出専用のコンパクトディスク（CD-ROM）、記録可能コンパクトディスク（CD-R）、書き換え可能コンパクトディスク（CD-RW）、読出専用のデジタル多用途ディスク（例えば、DVD-ROM、デュアルレイヤDVD-ROM）、様々な記録可能／書き換え可能DVD（例えば、DVD-RAM、DVD-RW、DVD＋RWなど）、フラッシュメモリ（例えば、SDカード、ミニSDカード、マイクロSDカードなど）、磁気および／

またはソリッドステートハードドライブ、読出専用のおよび記録可能ブルーレイ（登録商標）ディスク、超密度光ディスク、任意の他の光または磁気媒体、およびフロッピー（登録商標）ディスクが含まれる。コンピュータ可読媒体は、少なくとも1つの処理ユニットによって実行可能であり、種々の動作を実行するための命令のセットを含むコンピュータプログラムを記憶することができる。コンピュータプログラムまたはコンピュータコードの例としては、コンパイラによって生成されるようなマシンコードと、インタープリタを使用してコンピュータ、電子コンポーネント、またはマイクロプロセッサによって実行される上位レベルのコードを含むファイルがある。

【０４１２】

[00456] 本発明は、数多くの特定の詳細を参照して説明されてきたが、当業者は、本発明が本発明の精神から逸脱することなく、他の特定の形態で実施可能であることを認識するであろう。たとえば、上記の例のいくつかは、仮想ネットワークプロバイダの企業テナントの仮想企業WAN を示している。当業者は、ある実施形態では、仮想ネットワークプロバイダが、企業以外のテナント（例えば、学校、大学、非営利団体など）のために、1つ以上のパブリッククラウドプロバイダの複数のパブリッククラウドデータセンタ上に仮想ネットワークをデプロイすることを理解するであろう。これらの仮想ネットワークは、企業以外のエンティティの複数のコンピューティングエンドポイント（オフィス、データセンタ、コンピュータ、リモートユーザのデバイスなど）を接続する仮想WAN である。

【０４１３】

[00457] 上述のいくつかの実施形態は、オーバレイカプセル化ヘッダ内の各種データの断片を含む。当業者の1人は、他の実施形態が、このデータのすべてを中継するためのカプセル化ヘッダを使用しない可能性があることを理解するであろう。例えば、オーバレイカプセル化ヘッダにテナント識別子を含める代わりに、他の実施形態は、データメッセージを転送するCFEの宛先からテナント識別子を導出する。例えば、異なるテナントがパブリッククラウドにデプロイされた独自のMFNを有するいくつかの実施形態では、テナント識別子は、テナントメッセージを処理するMFNに関連付けられる。

【０４１４】

[00458] また、いくつかの図は、本発明のいくつかの実施形態のプロセスを概念的に示す。他の実施形態では、これらのプロセスの特定の動作は、これらの図に示されて説明されている正確な順序では実行されない。特定の動作は、1つの連続した一連の動作では実行されず、異なる特定の動作は、異なる実施形態で実行され得る。例えば、エンティティに対して、いくつかの実施形態は、パスがエンティティに対して生成される前に、PCD/PCDグループ内のMFNのいくつかまたはすべてをデプロイするが、上記の図のいくつかは、パス探索後のMFNのデプロイを示す。さらに、プロセスは、いくつかのサブプロセスを使用して、またはより大きなマクロプロセスの一部として実施することができる。

【０４１５】

[00459] 上述のいくつかの実施形態は、エンティティの専用仮想ネットワークの一部(例えば、エンティティの1つ以上の専用MFN)が輻輳しているように見えるか、または輻輳していると想定されるとき、1つのエンティティに対してデプロイされる専用仮想ネットワークが1つ以上の共有MFNを使用することを可能にする。他の実施形態は、このような輻輳を異なる方法で扱う。例えば、いくつかの実施形態は、追加の専用MFN(すなわち、冗長MFN)を、輻輳しているか、または輻輳していると予測される専用仮想ネットワークの部分にデプロイする。これらの実施形態のいくつかは、次に、PCD内の転送要素(例えば、フロントエンドロードバランサ)、MMCNにおけるエッジノード、および他の専用MFNを構成して、複製(例えば、冗長) MFNのクラスタ間で負荷を分散する。したがって、当業者であれば、本発明は、上述の例示的な詳細によって限定されるものではなく、むしろ、添付のクレームによって定義されるものであることが理解されるであろう。

【図面】

【図１Ａ】

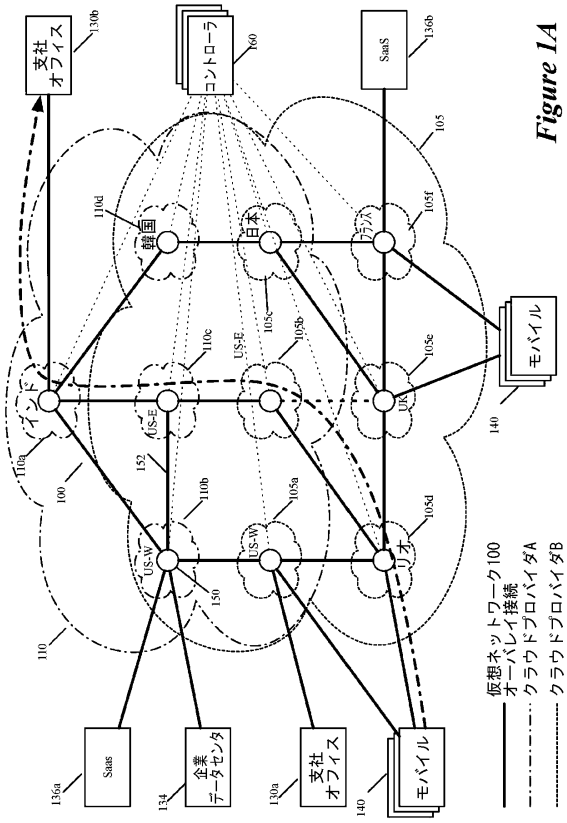

Figure 1A

【図１Ｂ】

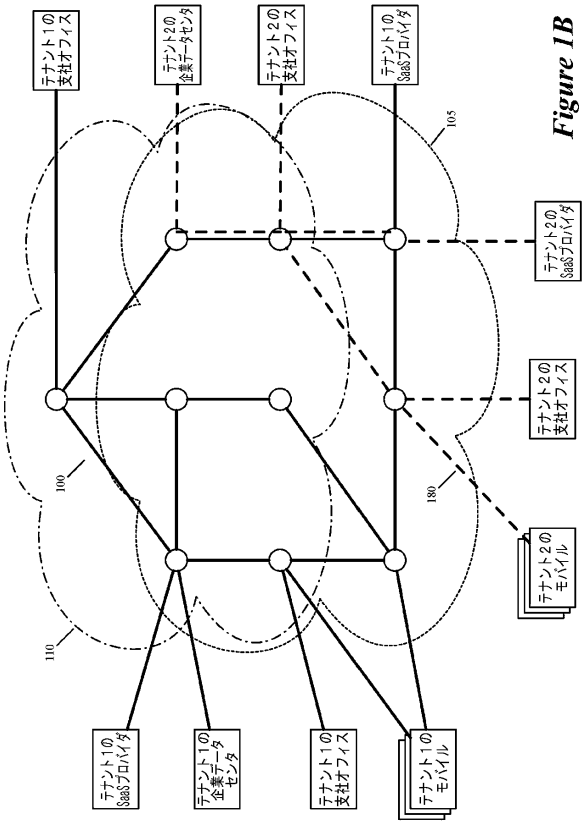

Figure 1B

【図１Ｃ】

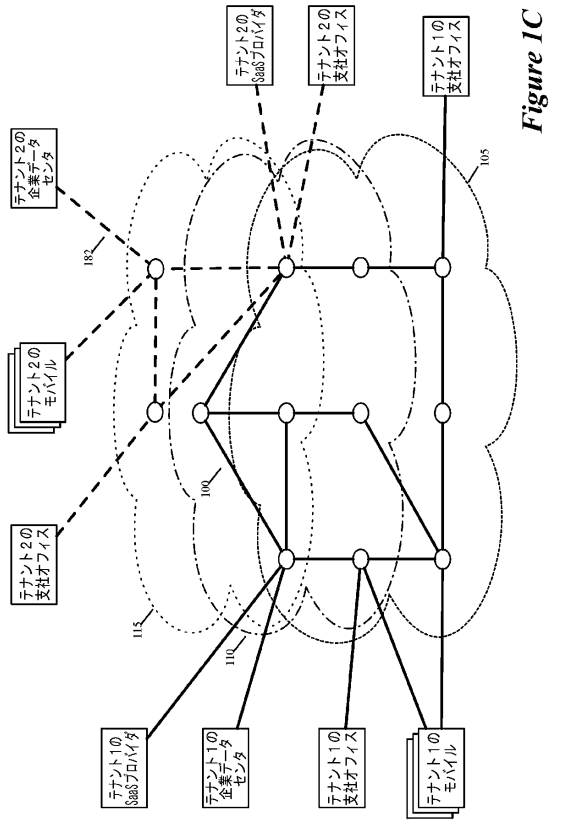

Figure 1C

【図２】

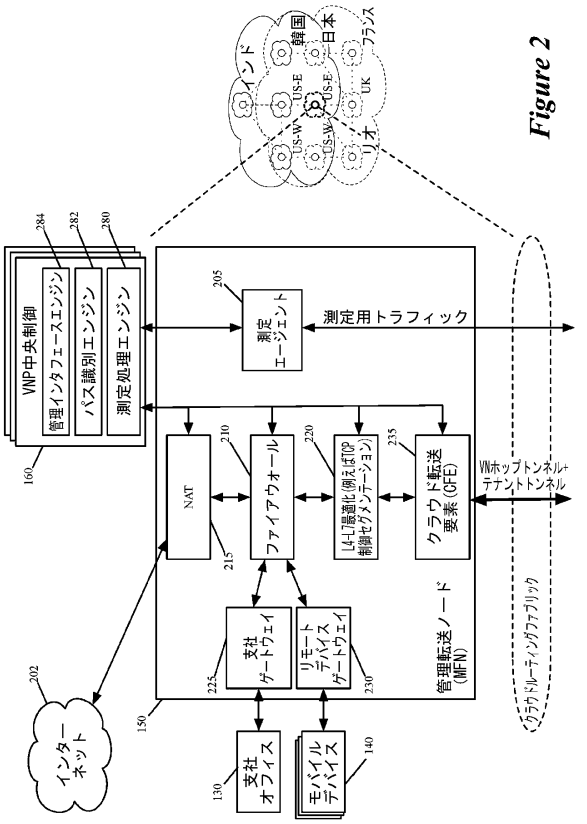

Figure 2

【図３】

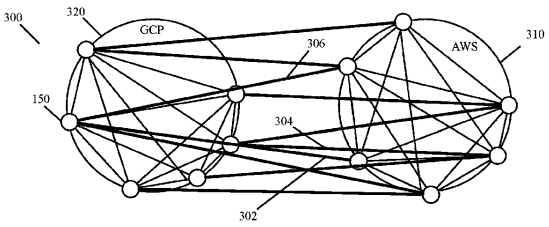

Figure 3

【図４Ａ】

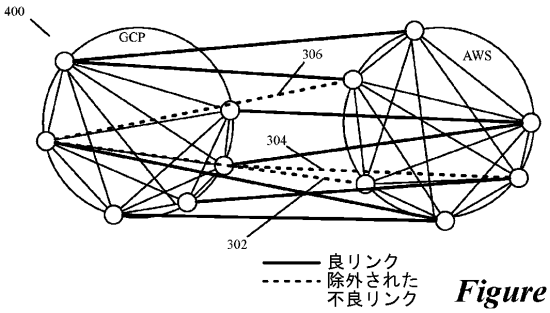

Figure 4A

【図４Ｂ】

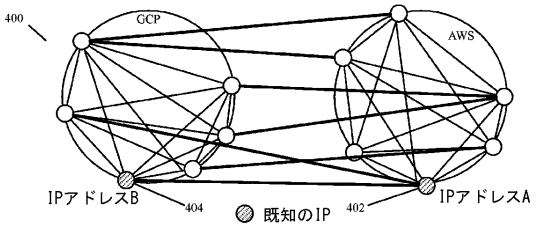

Figure 4B

【図４Ｃ】

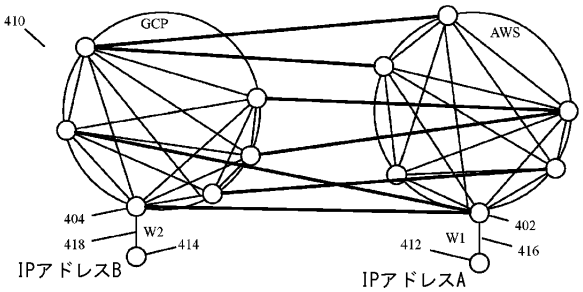

Figure 4C

【図４Ｄ】

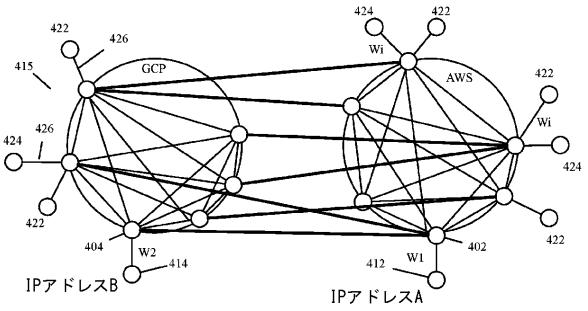

Figure 4D

【図５】

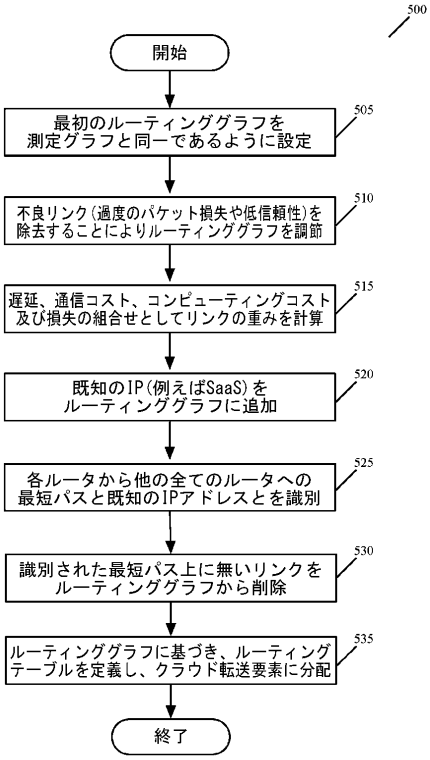

Figure 5

【図６】

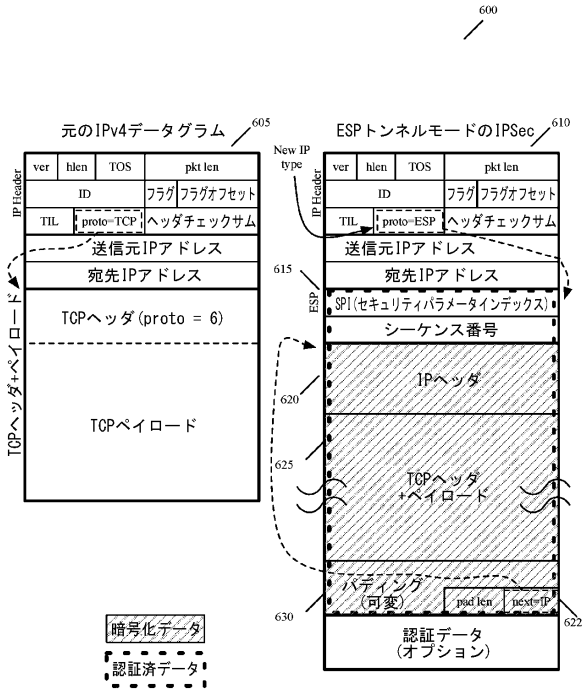

Figure 6

【図７】

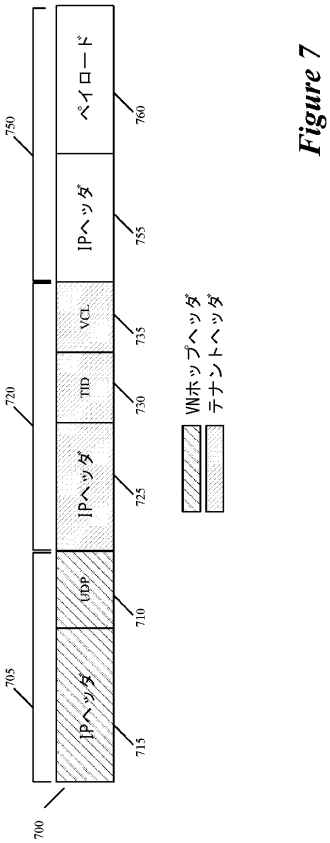

Figure 7

【図８】

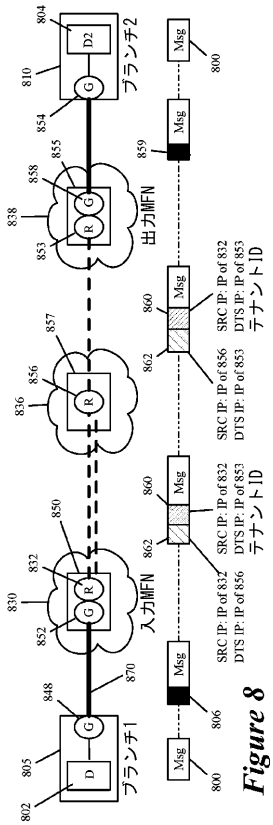

Figure 8

【図９】

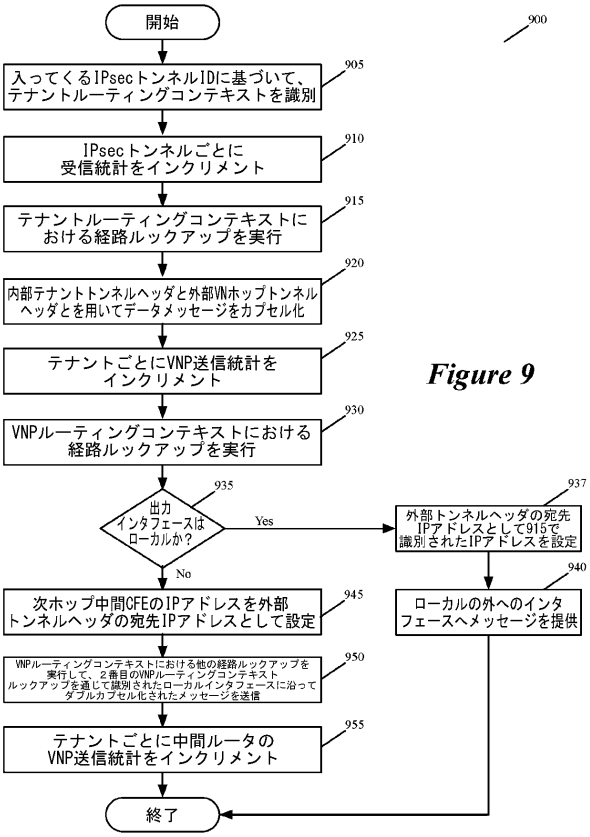

Figure 9

【図１０】

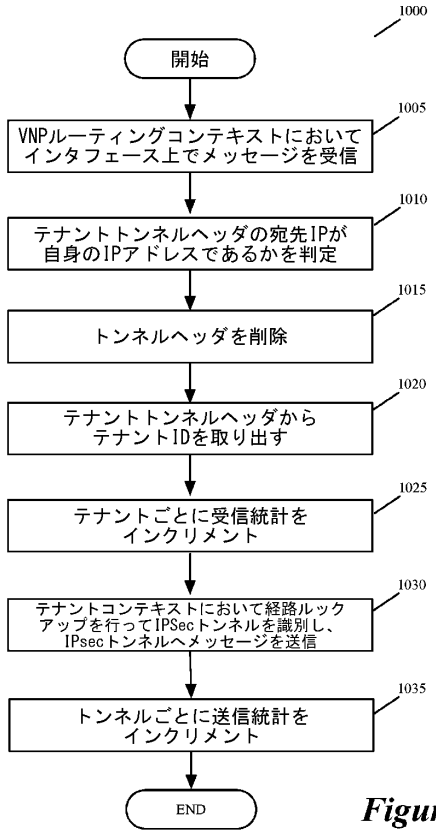

Figure 10

【図１１】

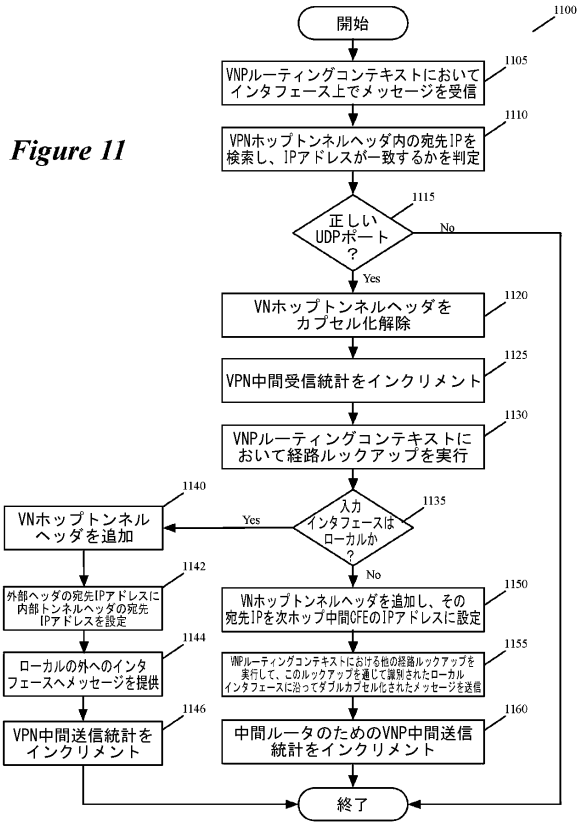

Figure 11

【図１２】

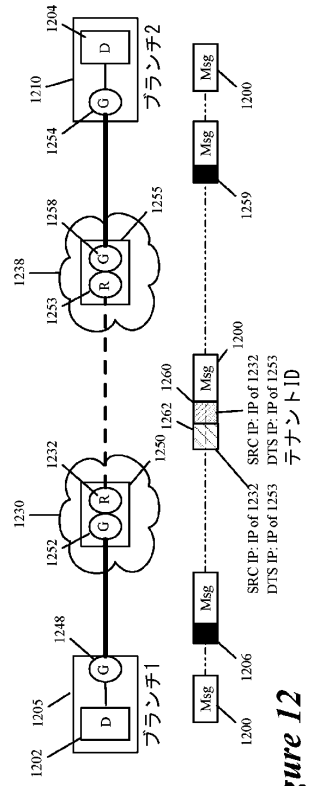

Figure 12

【図１３】

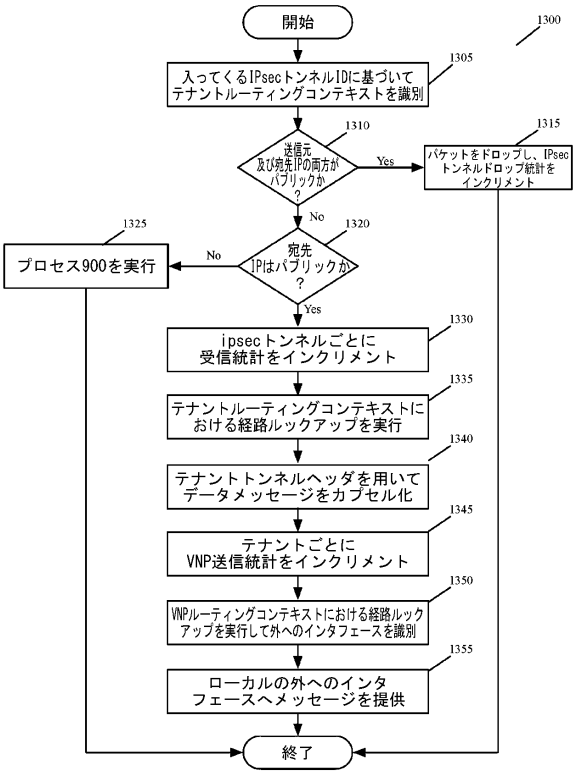

Figure 13

【図１４】

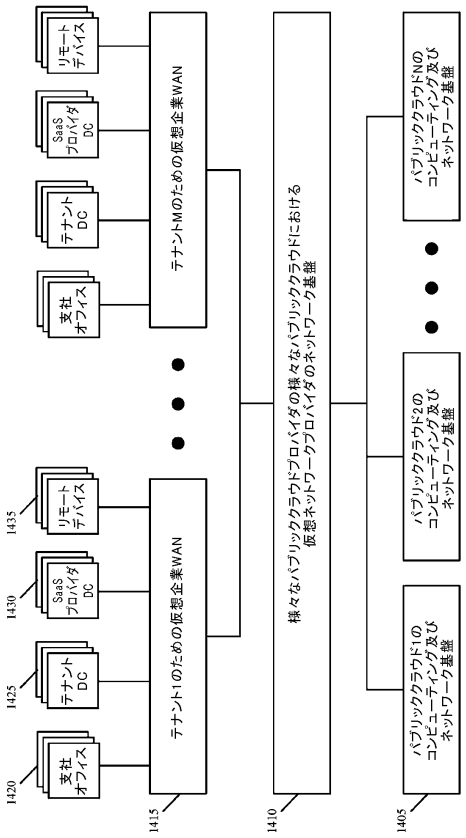

Figure 14

【図１５】

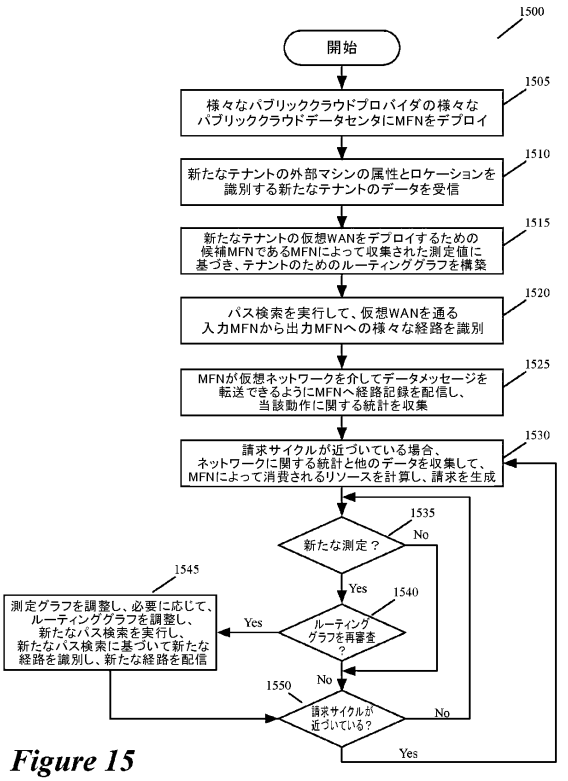

Figure 15

【図１６】

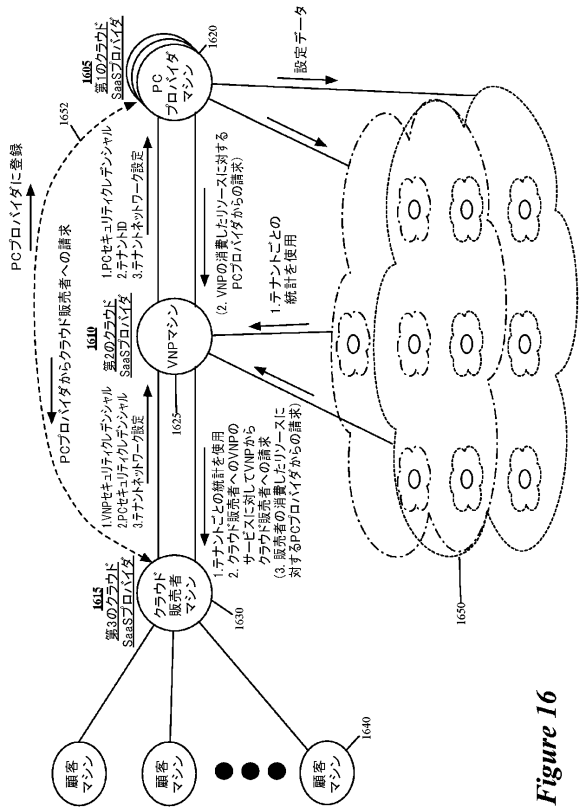

Figure 16

【図１７】

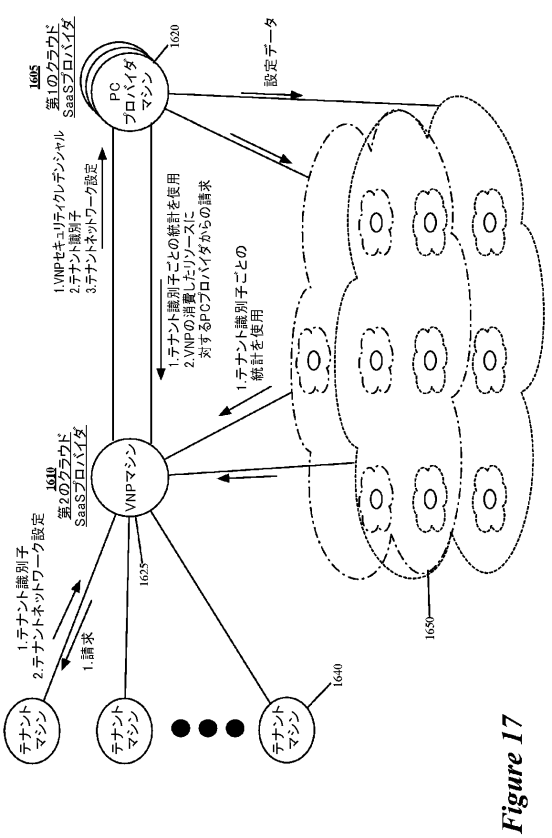

Figure 17

【図１８】

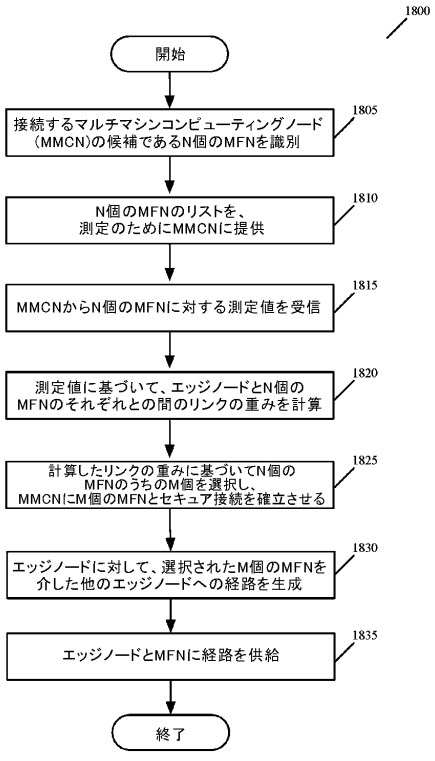

Figure 18

【図１９】

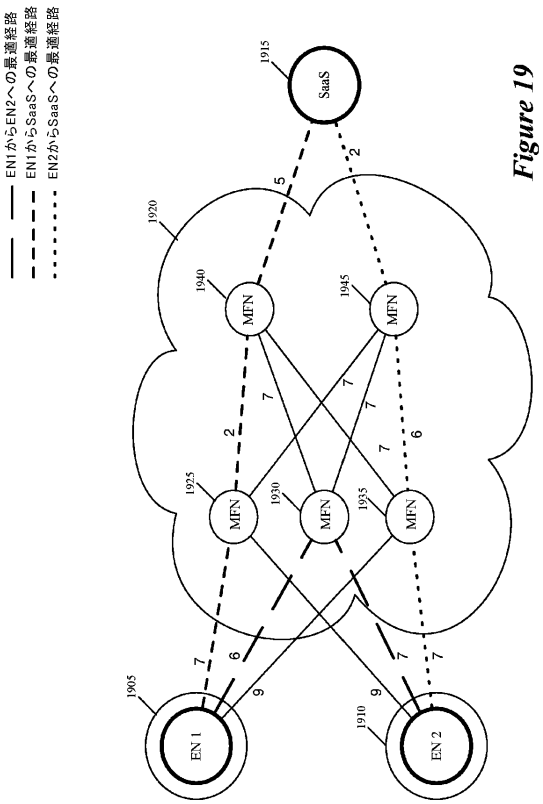

Figure 19

【図２０】

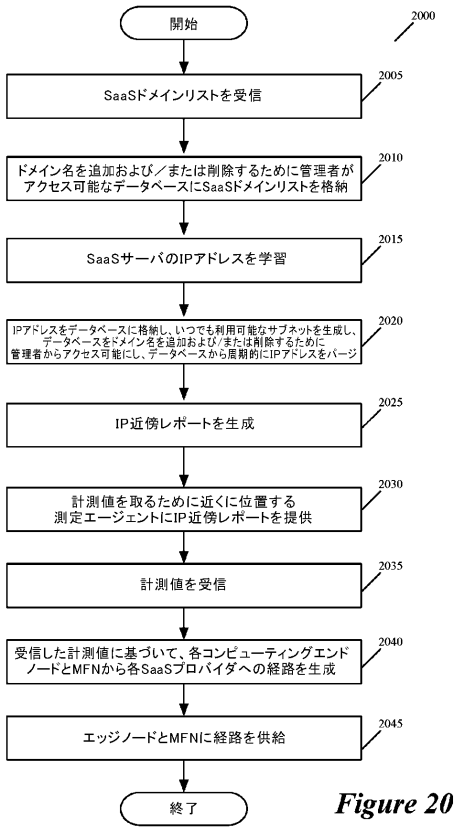

Figure 20

【図２１】

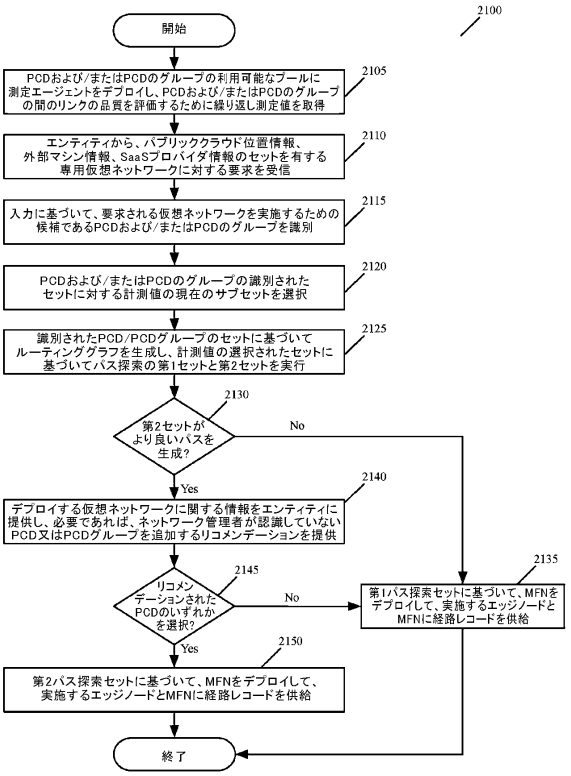

Figure 21

【図２２】

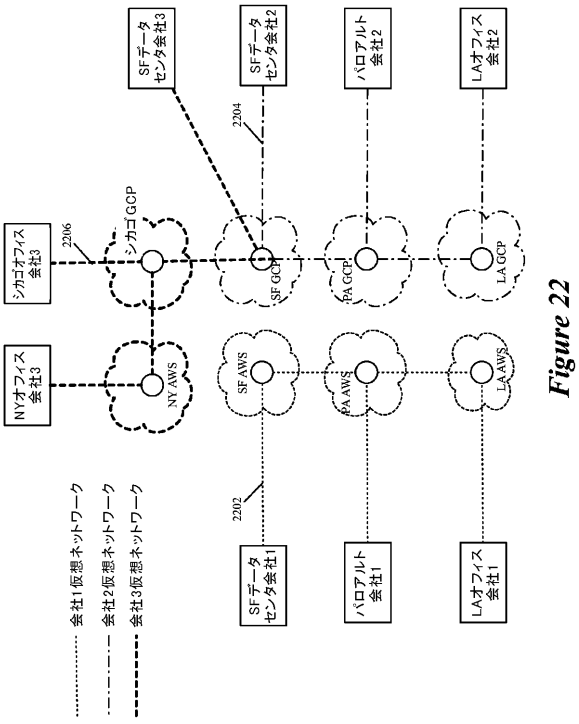

Figure 22

【図２３】

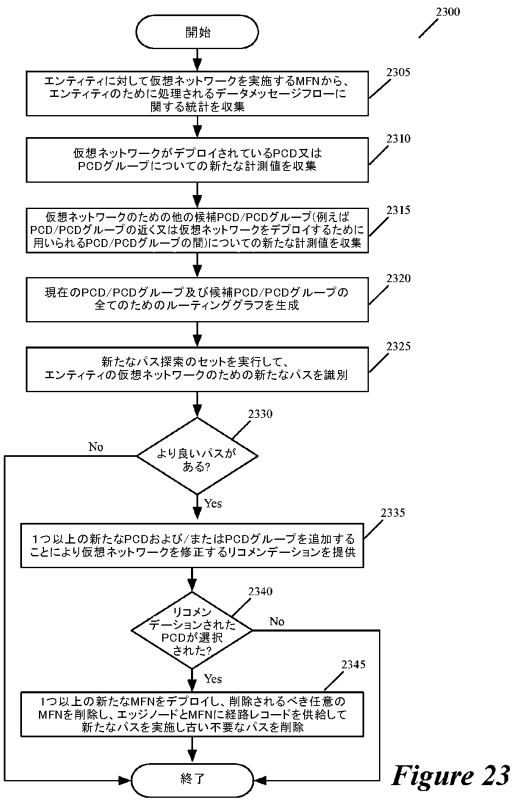

Figure 23

【図２４】

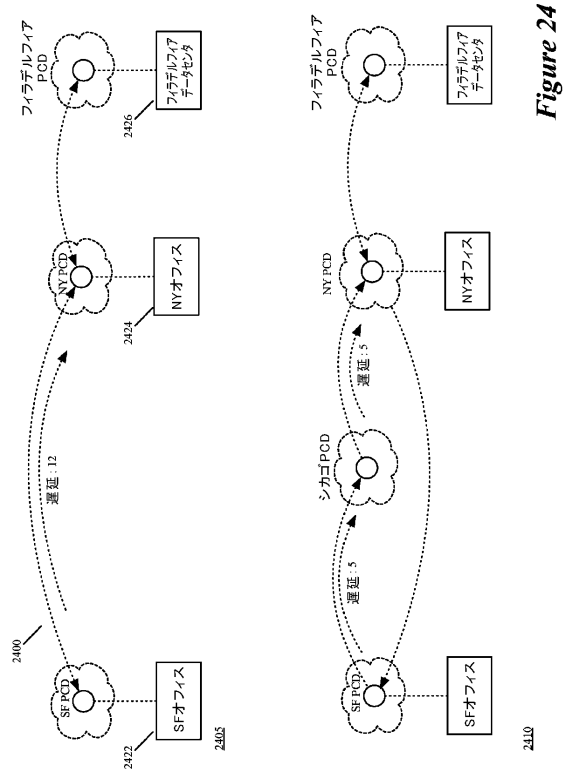

Figure 24

【図２５】

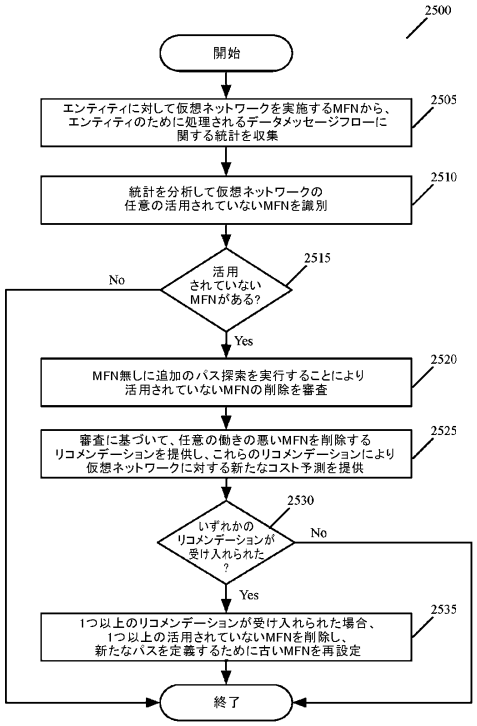

Figure 25

【図２６】

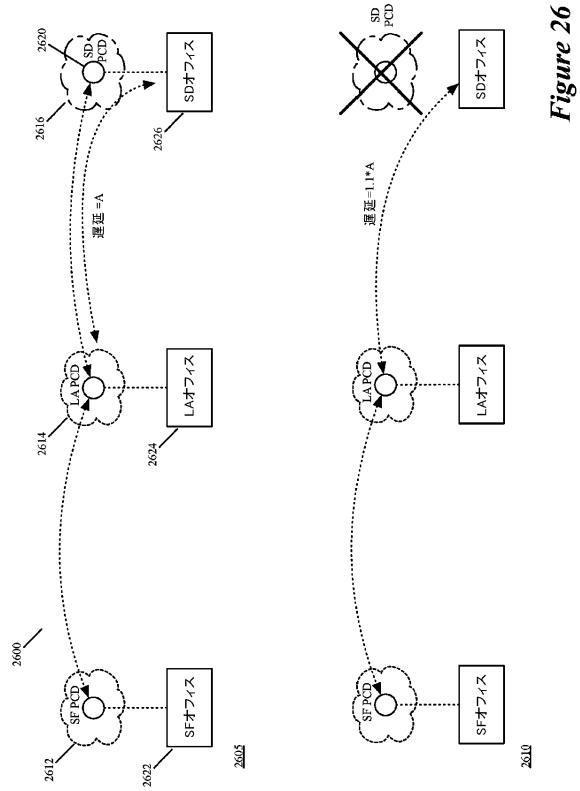

Figure 26

【図２７】

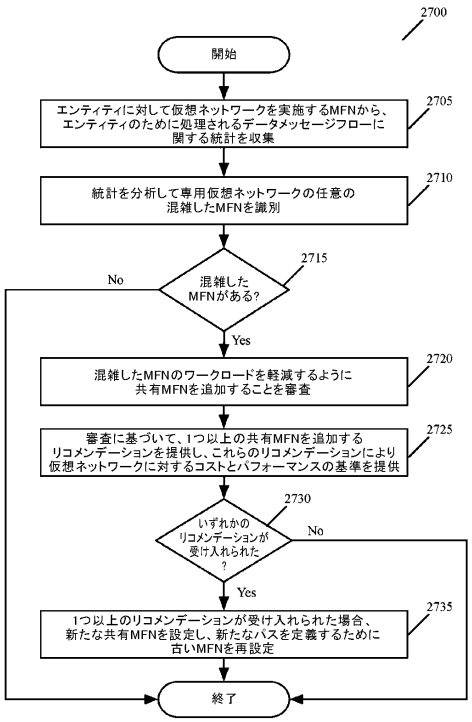

Figure 27

【図２８】

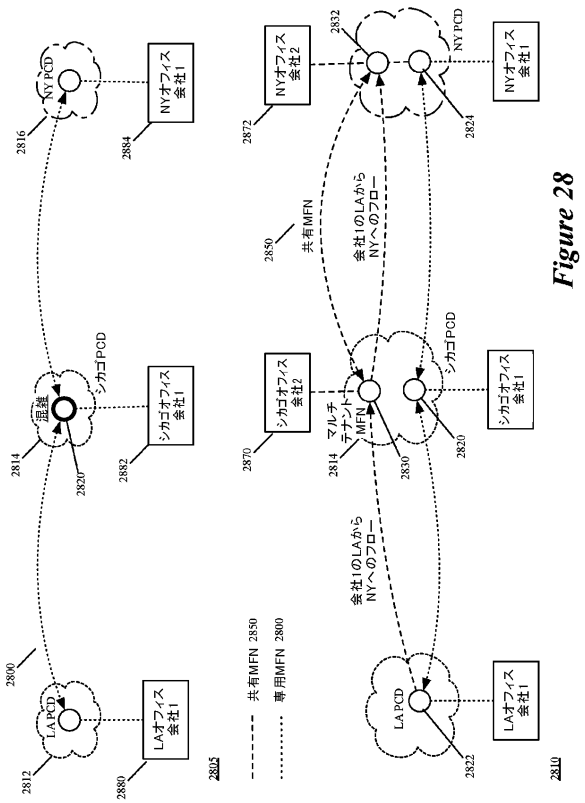

Figure 28

【図２９】

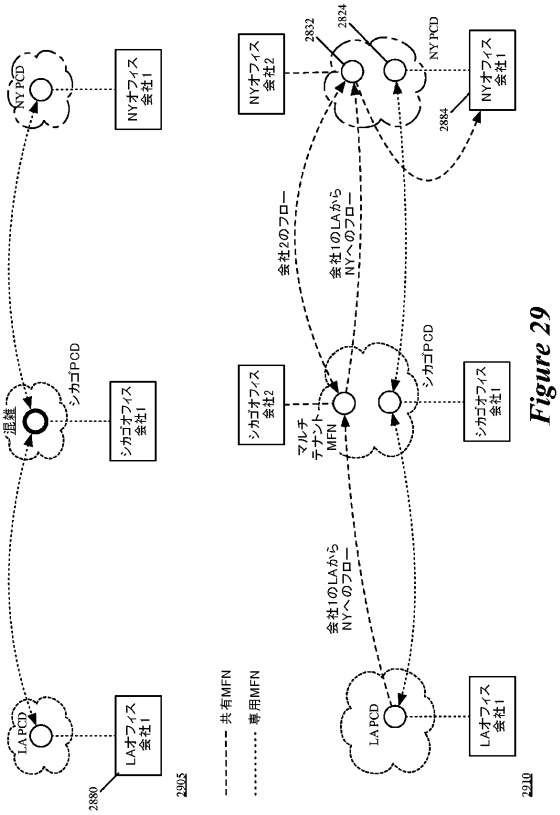

Figure 29

【図３０】

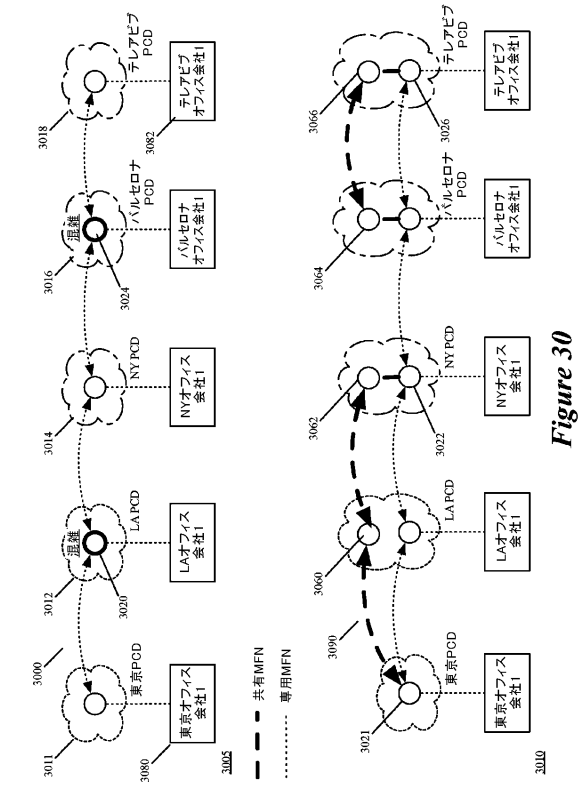

Figure 30

【図３１】

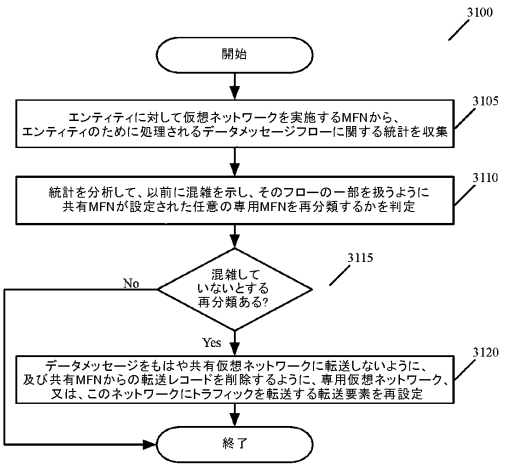

Figure 31

【図３２】

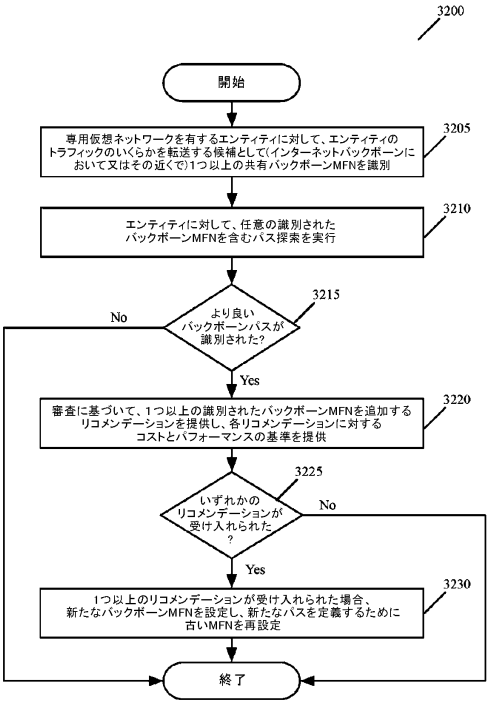

Figure 32

【図３３】

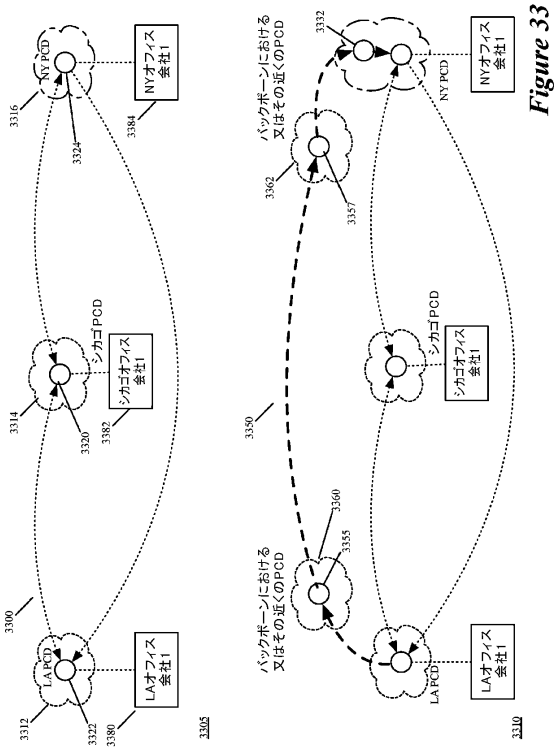

Figure 33

【図３４】

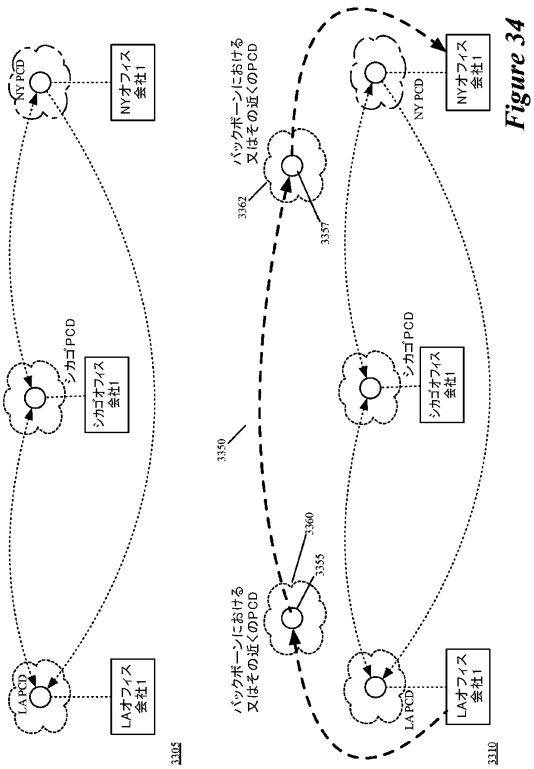

Figure 34

【図３５】

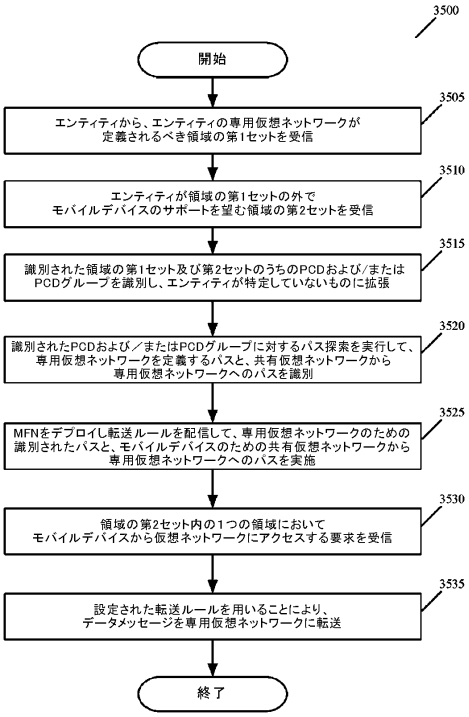

Figure 35

【図３６】

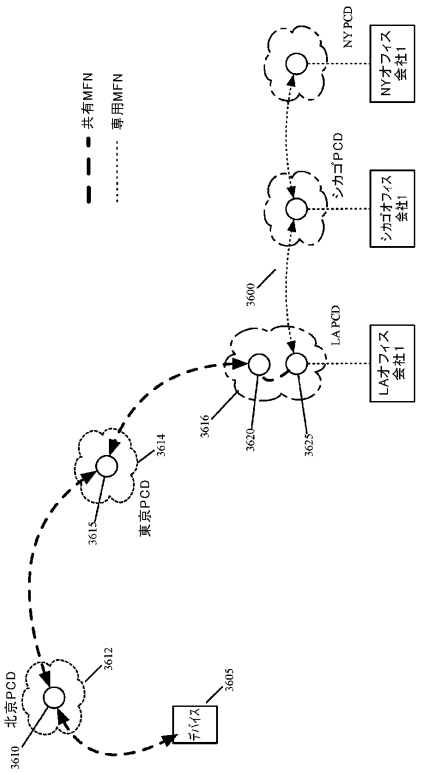

Figure 36

【図３７】

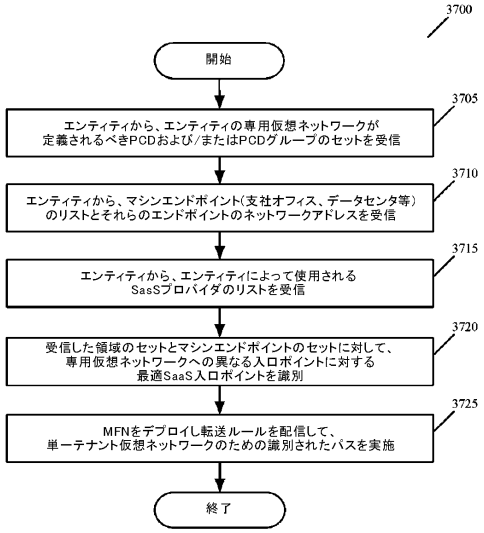

Figure 37

【図３８】

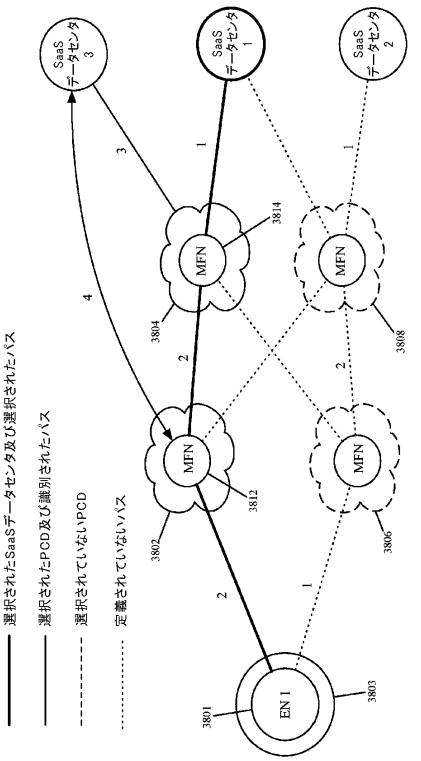

Figure 38

【図３９】

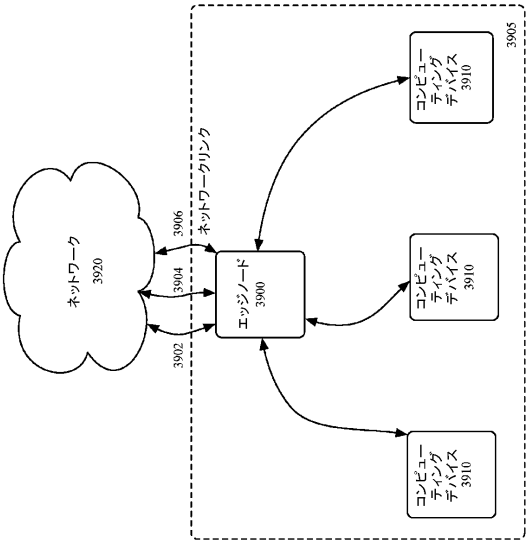

Figure 39

【図４０】

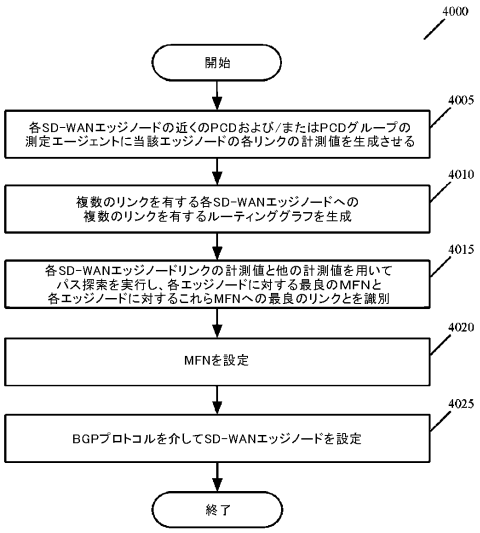

Figure 40

【図４１】

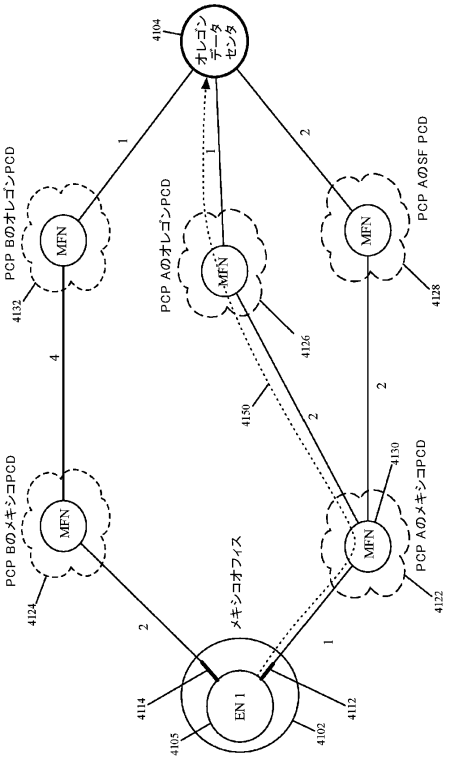

Figure 41

【図４２】

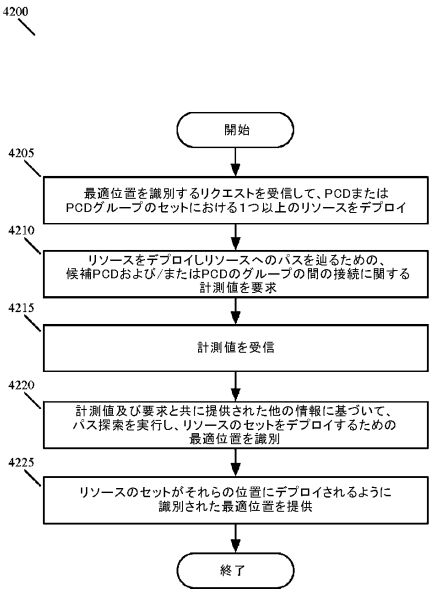

Figure 42

【図４３】

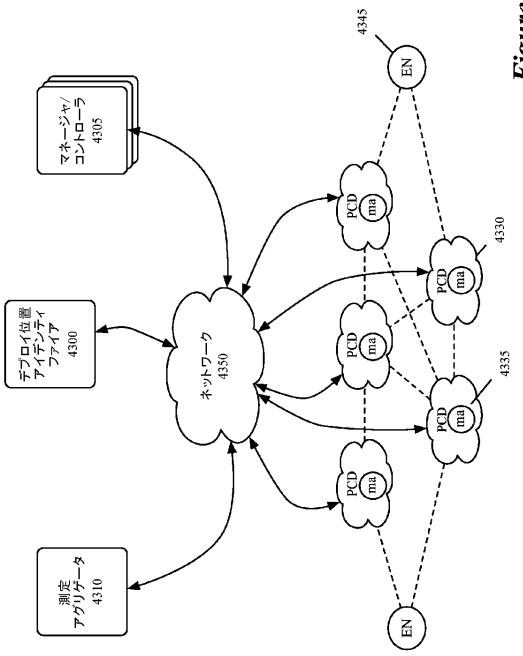

Figure 43

【図４４】

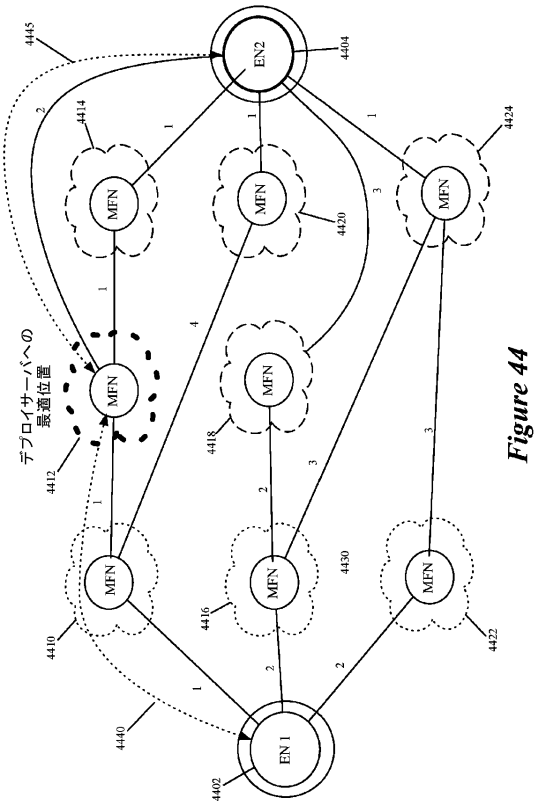

Figure 44

【図４５】

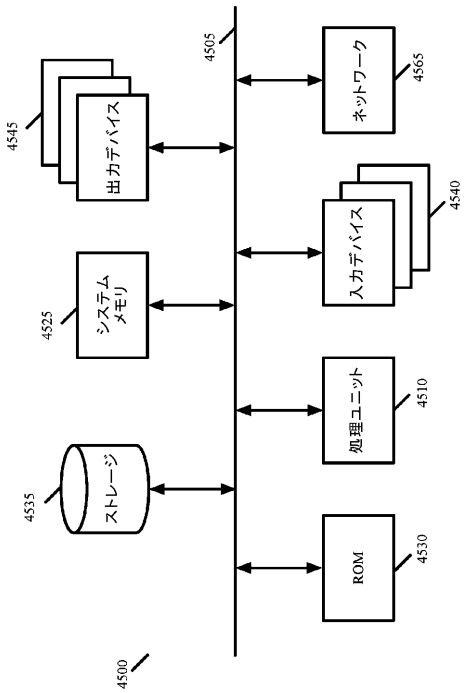

Figure 45

フロントページの続き

(33)優先権主張国・地域又は機関
米国(US)
(31)優先権主張番号　16/662,363
(32)優先日　令和1年10月24日(2019.10.24)
(33)優先権主張国・地域又は機関
米国(US)
(31)優先権主張番号　16/662,427
(32)優先日　令和1年10月24日(2019.10.24)
(33)優先権主張国・地域又は機関
米国(US)
(31)優先権主張番号　16/662,531
(32)優先日　令和1年10月24日(2019.10.24)
(33)優先権主張国・地域又は機関
米国(US)
(31)優先権主張番号　16/662,591
(32)優先日　令和1年10月24日(2019.10.24)
(33)優先権主張国・地域又は機関
米国(US)
(31)優先権主張番号　16/662,587
(32)優先日　令和1年10月24日(2019.10.24)
(33)優先権主張国・地域又は機関
米国(US)
(31)優先権主張番号　16/662,402
(32)優先日　令和1年10月24日(2019.10.24)
(33)優先権主張国・地域又は機関
米国(US)
(31)優先権主張番号　16/662,489
(32)優先日　令和1年10月24日(2019.10.24)
(33)優先権主張国・地域又は機関
米国(US)
(31)優先権主張番号　16/662,510
(32)優先日　令和1年10月24日(2019.10.24)
(33)優先権主張国・地域又は機関
米国(US)
(31)優先権主張番号　16/662,570
(32)優先日　令和1年10月24日(2019.10.24)
(33)優先権主張国・地域又は機関
米国(US)
(74)代理人　100203242
弁理士　河戸 春樹
(74)代理人　100212657
弁理士　塚原 一久
(72)発明者　サイドン， イスラエル
アメリカ合衆国 カリフォルニア州 ９４３０４， パロアルト， ヒルビュー アベニュー ３４０１
(72)発明者　ベヌゴパル， プラシャンス
アメリカ合衆国 カリフォルニア州 ９４３０４， パロアルト， ヒルビュー アベニュー ３４０１
(72)発明者　バーグマン， アラン
イスラエル国 ギヴァタイム ５３３９５１３， カレン ケイメット レイスラエル ４４Ａ
(72)発明者　ダール， チェン

イスラエル国 マグシミム ５６９１０００， ロテムストリート ４
(72)発明者 マルクゼ， アレックス
イスラエル国 ラマト ガン ５２５２３２８， エーピー ３３， ビアリク １３４
(72)発明者 ゾハール， イヤル
イスラエル国 シムシット １７９０６， ピーオーボックス ３４５
審査官 宮島 郁美
(56)参考文献 国際公開第２０１９／０７０６１１（WO，Ａ１）
米国特許出願公開第２００８／０１７５１５０（ＵＳ，Ａ１）
特開２００２－３６８７９２（ＪＰ，Ａ）
特表２０１７－５２４２９０（ＪＰ，Ａ）
(58)調査した分野 (Int.Cl.，ＤＢ名)
Ｈ０４Ｌ１２／００－１２／６６，１３／００，４１／００－４９／９０５７，６１／００－６５／８０，６９／００－６９／４０