

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2019-523621

(P2019-523621A)

(43) 公表日 令和1年8月22日(2019.8.22)

(51) Int.Cl.
H04L 12/803 (2013.01)F I
H04L 12/803テーマコード (参考)
5K030

審査請求 未請求 予備審査請求 未請求 (全 40 頁)

(21) 出願番号 特願2019-518182 (P2019-518182)
(86) (22) 出願日 平成29年6月18日 (2017.6.18)
(85) 翻訳文提出日 平成31年2月15日 (2019.2.15)
(86) 国際出願番号 PCT/US2017/038053
(87) 国際公開番号 WO2017/219009
(87) 国際公開日 平成29年12月21日 (2017.12.21)
(31) 優先権主張番号 62/351, 953
(32) 優先日 平成28年6月18日 (2016.6.18)
(33) 優先権主張国・地域又は機関
米国 (US)

(71) 出願人 518450511
クレバーネット、インコーポレイテッド
CLEVERNET, INC.
アメリカ合衆国 カリフォルニア州 94
159、サンフランシスコ、チャイナ
イスン ストリート 325、ナンバー1
512
325 China Basin St.
#1512, San Francis
co, CA 94159 (US)
(74) 代理人 100109634
弁理士 舩谷 威志
(74) 代理人 100129263
弁理士 中尾 洋之

最終頁に続く

(54) 【発明の名称】 複数のチャネルを使用してパフォーマンスを向上するインテリジェントアダプティブトランスポートレイヤー

(57) 【要約】

より効率的かつコントロールした方法で、情報フローを伝達する際に使用するために、1組の接続を確立し、継続的に評価し、コンピュータネットワーク上のホスト間で維持する。より良いおよび/または異なるパフォーマンス特性を有する接続を継続的に探索することにより、新しい接続が確立され、既存の接続が終了する。各接続は、ネットワークを通して、同じまたは異なるパスを利用することができ、時間とともに変化するパフォーマンス特性を有することができる。スループット、トランザクション時間、データの一貫性、レイテンシ、およびパケット損失を含む、ネットワークメトリックを向上するために、特定の情報フローに対して複数のパスを同時に使用できる。情報フローは1つまたは複数のサブフローに分割することができ、サブフローは1つまたは複数のアクティブな接続に割り当てることができる。さらに、ネットワークの状態に応じて、フローの分割方法および接続への割り当て方法に関する動的な決定を行うことができる。これらの接続を使用することで、コストを削減し、アプリケーションのQoS/QoEを保証して、公

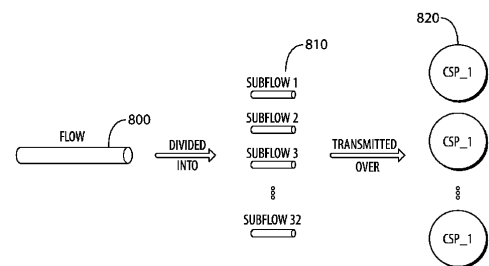


FIG. 8

【特許請求の範囲】**【請求項 1】**

ネットワークホストを結合し、ネットワークホスト間で情報フローを通信することができるネットワークを含み、システムにおいてネットワークホスト間の複数の接続を利用する方法であって：

第 1 のネットワークホストと第 2 のネットワークホストとの間の複数の接続を維持することであって、前記複数の接続のそれぞれに対して；

前記第 1 のネットワークホストと、前記第 2 のネットワークホストとの間の接続を確立すること；

前記接続を繰り返し評価して、前記接続に関連する少なくとも 1 つのパフォーマンスメトリクスを決定すること；

前記接続に関する少なくとも 1 つのパフォーマンスメトリクスに基づき、前記複数の状態に関連付けることであって、前記複数の状態のうち、1 つまたは複数の第 1 の状態が準備完了状態を表し、前記複数の状態のうち、1 つまたは複数の第 2 の状態が準備のできていない状態を表すこと；

前記第 1 のネットワークホストと前記第 2 のネットワークホストとの間で情報フローを移転する要求を受理すること；のサブステップを含み、

1 つまたは複数の前記第 1 の状態に関連付けられた、1 つまたは複数の前記接続に前記情報フローを割り当てること；

前記割り当てのステップに基づく前記複数の接続の 1 つまたは複数の前記情報フローからのデータを通信すること、を含む方法。

【請求項 2】

前記複数の接続が VPN トンネルである、請求項 1 に記載の方法。

【請求項 3】

前記少なくとも 1 つのパフォーマンスメトリクスが、レイテンシ、スループット、およびパケット損失からなるセットから取得される、1 つまたは複数のメトリクスを含む、請求項 1 に記載の方法。

【請求項 4】

前記複数の接続のうちの 1 つについて、前記ネットワークを通るパスが、前記複数の接続のうちの第 2 の接続について、前記ネットワークを通るパスとは異なる、請求項 1 に記載の方法。

【請求項 5】

前記第 1 のホストが複数のエンドポイントを有し、前記複数の接続のうちの少なくとも 1 つが、前記複数の接続のうちの第 2 の接続とは異なる 1 つの前記エンドポイントを使用する、請求項 1 に記載の方法。

【請求項 6】

前記繰り返し評価するステップが、定期的間隔で周期的な前記接続に関連する、パフォーマンスメトリクスの評価を構成する、請求項 1 に記載の方法。

【請求項 7】

前記 1 つまたは複数の第 2 の状態が、スタンバイ、待機中、降格中、および探索中の状態を表す状態を含む、請求項 1 に記載の方法。

【請求項 8】

ネットワークの少なくとも一部が公衆インターネットの一部を構成する、請求項 1 に記載の方法。

【請求項 9】

前記複数の接続のそれぞれについて、前記接続の前記ネットワークを通るパスを決定するサブステップをさらに含む、請求項 1 に記載の方法。

【請求項 10】

前記第 1 の状態のうちの 1 つまたは複数に関連付けられた、前記接続のうちの異なる 1 つまたは複数に前記情報フローを再割り当てし、前記再割り当てのステップが前記複数の接続の少なくとも 1 つに関連した少なくとも 1 つのパフォーマンスメトリックに基づくものであり；

前記再割り当てのステップに基づく前記複数の接続の 1 つまたは複数を通して、前記情報フローからのデータを通信するステップをさらに含む、請求項 1 の方法。

【請求項 1 1】

前記接続が、前記接続に関連する前記少なくとも 1 つのパフォーマンスメトリックに基づく基準を満たす場合に、前記接続を終了させ、前記基準は、前記接続に関連する少なくとも 1 つのパフォーマンスメトリックと、複数の接続のうちの他の接続に関連する 1 つまたは複数の他のメトリクスとの比較に基づく、請求項 1 に記載の方法。

10

【請求項 1 2】

ネットワークホストを結合するネットワークを通して、ネットワークホスト間で情報フローを通信する装置であって、

プロセッサ読み取り可能命令を記憶する、少なくとも 1 つのメモリと通信する少なくとも 1 つのプロセッサを含む第 1 のネットワークホストであって、前記少なくとも 1 つのプロセッサが、前記プロセッサ可読な命令により、

第 2 のネットワークホストとの間の複数の接続を維持することであって、前記複数の接続のそれぞれに対して；

前記第 2 のネットワークホストとの間の接続を確立すること；

20

前記接続を繰り返し評価して、前記接続に関連する少なくとも 1 つのパフォーマンスメトリクスを決定すること；

前記接続に関する少なくとも 1 つのパフォーマンスメトリックに基づき、前記複数の状態に関連付けることであって、前記複数の状態のうち、1 つまたは複数の第 1 の状態が準備完了状態を表し、前記複数の状態のうち、1 つまたは複数の第 2 の状態が準備のできていない状態を表すこと；

前記第 2 のネットワークホストに情報フローを移転する要求を受理すること；のサブステップを含み、

1 つまたは複数の前記第 1 の状態に関連付けられた、1 つまたは複数の前記接続に前記情報フローを割り当て；

30

前記割り当てのステップに基づく前記複数の接続の 1 つまたは複数の前記情報フローからのデータを通信する、

ように構成された装置。

【請求項 1 3】

前記第 1 のホストが複数のエンドポイントを有し、前記複数の接続のうちの少なくとも 1 つが、前記複数の接続のうちの第 2 の接続とは異なる 1 つの前記エンドポイントを使用する、請求項 1 2 に記載の装置。

【請求項 1 4】

繰り返しの評価は、規則的な間隔で周期的に前記接続に関連するメトリクスを評価することを含む、請求項 1 2 に記載の装置。

40

【請求項 1 5】

前記少なくとも 1 つのプロセッサは、プロセッサ可読命令により、さらに、

前記第 1 の状態のうちの 1 つまたは複数に関連付けられた前記接続のうちの異なる 1 つまたは複数のものに前記情報フローを再割り当てし、前記再割り当てステップは、前記複数の接続のうちの少なくとも 1 つに関連する少なくとも 1 つのパフォーマンスメトリクスに基づくものであり、

前記再割り当てステップに基づいて、前記複数の接続のうちの 1 つまたは複数を通して前記情報フローからのデータを通信するように動作可能に構成される、請求項 1 2 に記載の装置。

【請求項 1 6】

50

前記少なくとも１つのプロセッサは、プロセッサ可読命令により、さらに前記複数の接続について、

前記接続が、前記接続に関連する前記少なくとも１つのパフォーマンスメトリックに基づく基準を満たす場合に、前記接続を終了させ、前記基準が、前記接続に関連する少なくとも１つのパフォーマンスメトリック基準と、複数の接続のうちの他の接続に関連する、１つまたは複数の他のメトリクスとの比較に基づくものである、請求項１２に記載の装置。

【請求項１７】

請求項１に記載の方法を実行するように、少なくとも１つのプロセッサに指示するためのプログラムコードで符号化された、非一時的コンピュータ可読媒体。

【請求項１８】

ホスト間で情報を通信することができるネットワークに結合された、複数のネットワークホストを含むシステムにおいて、複数の利用可能な接続を介して、複数のパケットを含む情報のフローを通信する方法であって、システムは、ネットワークホスト間の複数の接続を利用し：

第１のネットワークホストと第２のネットワークホストとの間で情報フローを転送する要求を受理し、前記要求はパフォーマンス要件およびアプリケーション特性に関連付けられること；

情報フローを複数のサブフローに分割し、各サブフローはパケットのシーケンスを含み、分割するステップはパフォーマンス要件およびアプリケーション特性のうちの１つまたは複数のに基づくものであること；

複数のサブフローのそれぞれを、前記第１のネットワークホストと前記第２のネットワークホストとの間の複数の接続の中からの接続に割り当てること；

前記割当てに基づいて、前記複数の接続のうちの１つまたは複数の介して、前記複数のサブフローからのデータを通信することを含む方法。

【請求項１９】

前記複数の接続は、ＶＰＮトンネルである、請求項１８に記載の方法。

【請求項２０】

前記アプリケーション特性が、パケットサイズ、フローサイズ、フロー期間、レイテンシ要求、および優先度からなるセットから取得される１つまたは複数の特性を含む、請求項１８に記載の方法。

【請求項２１】

前記性能要件が、１つまたは複数のサービスレベル契約（ＳＬＡ）に基づく、請求項１８に記載の方法。

【請求項２２】

前記分割するステップはさらに、ネットワーク状態、ユーザ情報および履歴上の決定からなる組から取得される１つまたは複数の項目を含む情報に基づく、請求項１８に記載の方法。

【請求項２３】

前記分割するステップは、前記情報フローの転送中に、情報フローを前記複数のサブフローに動的に分割する方法を変更する、請求項１８に記載の方法。

【請求項２４】

前記割当てのステップが、前記情報フローの転送中に、前記複数の接続のうちの異なる接続への１つまたは複数のサブフローの割り当てを変更する、請求項１８に記載の方法。

【請求項２５】

さらに、

前記第１のネットワークホストと前記第２のネットワークホストとの間で、第２の情報フローを転送するための第２の要求を受理すること；

第２の情報フローを、それぞれがパケットのシーケンスを含む、第２の複数のサブフローに分割すること；

前記第１のネットワークホストと前記第２のネットワークホストとの間の、前記複数の

10

20

30

40

50

接続の中から、第 2 の複数のサブフローのそれぞれを接続に割り当てること；

第 2 の複数のサブフローのそれぞれを割り当てる前記ステップに基づいて、前記第 2 の複数の接続のうちの 1 つまたは複数を介して、前記第 2 の複数のサブフローのそれぞれからのデータを通信することをさらに含む、請求項 18 に記載の方法。

【請求項 26】

前記複数の接続のうちの少なくとも 1 つが、前記情報フローおよび前記第 2 の情報フローの両方からのデータを通信する、請求項 25 に記載の方法。

【請求項 27】

前記ネットワークの少なくとも一部が、前記公衆インターネットの一部を構成する、請求項 18 に記載の方法。

10

【請求項 28】

ネットワークホストを結合するネットワークを介して、ネットワークホスト間で、情報フローを通信するための装置であって、

プロセッサ可読命令を記憶する、少なくとも 1 つのメモリと通信する少なくとも 1 つのプロセッサを含む、第 1 のネットワークホストを含む装置であって、少なくとも 1 つのプロセッサは、プロセッサ可読な命令により、

パフォーマンスフローを第 2 のネットワークホストに転送する要求を受理し、ここで前記要求は、パフォーマンス要件およびアプリケーション特性に関連するものであり；

情報フローを、それぞれがパケットのシーケンスを含む複数のサブフローに分割し、分割するステップは、パフォーマンス要件およびアプリケーション特性のうちの 1 つまたは複수에基づくものであり；

20

前記複数のサブフローのそれぞれを、前記第 1 のネットワークホストと前記第 2 のネットワークホストとの間の複数の接続の中からの接続に割り当て；

前記割り当てに基づいて、前記複数の接続のうちの 1 つまたは複数を通して、前記複数のサブフローからのデータを通信する、プロセッサ可読命令によって動作可能に構成される、装置。

【請求項 29】

前記アプリケーション特性が、パケットサイズ、フローサイズ、フロー期間、レイテンシ要求、および優先度からなるセットから取られる、1 つまたは複数の特性を含む、請求項 28 に記載の装置。

30

【請求項 30】

前記プロセッサは、前記情報フローの転送中に情報フローが前記複数のサブフローに動的にどのように分割するかを変更するように動作可能に構成される、請求項 28 に記載の装置。

【請求項 31】

前記プロセッサは、前記情報フローの転送中に、前記複数の接続のうちの異なる接続への、1 つまたは複数のサブフローの割り当てを変更するように、動作可能に構成される、請求項 28 に記載の装置。

【請求項 32】

少なくとも 1 つのプロセッサは、プロセッサ可読命令によって、さらに、

40

前記第 1 のネットワークホストと前記第 2 のネットワークホストとの間で、第 2 の情報フローを転送するための第 2 の要求を受理し；

第 2 の情報フローを、それぞれがパケットのシーケンスを含む第 2 の複数のサブフローに分割し；

前記第 1 のネットワークホストと前記第 2 のネットワークホストとの間の前記複数の接続の中から、第 2 の複数のサブフローのそれぞれを接続に割り当て；

第 2 の複数のサブフローのそれぞれを割り当てることに基づいて、前記第 2 の複数のサブフローのそれぞれからのデータを、前記第 2 の複数の接続のうちの 1 つまたは複数を通して通信するように動作可能に構成される、請求項 28 に記載の装置。

【請求項 33】

50

前記複数の接続のうちの少なくとも１つが、前記第１の情報フローおよび前記第２の情報フローの両方からのデータを通信する、請求項３３に記載の装置。

【請求項３４】

請求項１８に記載の方法を実行するように少なくとも１つのプロセッサに指示するためのプログラムコードで暗号化された非一時的コンピュータ可読媒体。

【発明の詳細な説明】

【技術分野】

【０００１】

10

関連出願

本出願は、２０１６年６月１８日に提出した米国仮出願６２／３５１，９５３によって優先権を主張し、参照により本明細書に組み込まれる。

【背景技術】

【０００２】

背景

本発明は、コンピュータネットワーキングの分野に関し、より詳細には、パケット交換ネットワークのエンドポイント間のレイテンシ、フロー完了時間（ＦＣＴ）およびスループットなどのネットワークメトリクスを制御することに関する。これは、パブリックインターネット、プライベートインターネット、および３Ｇ／４Ｇ／５Ｇモバイルネットワークなどのネットワークを含む。

20

【０００３】

インターネットは優れた接続性を提供しながら、回復力、分散化、およびベストエフォート型パケット配信などの特性を保証する。しかし、これらの特性により、トラフィックのピークを処理するためのインターネットコアの使用率が低くなる。加えて、インターネットは、一般的に決定論的なものではなく、重要なアプリケーションへの利用を妨げることは事実である。企業は通常、スループットやレイテンシなど、関心のある指標を確保するためにプライベートネットワークを利用する。しかし、プライベートネットワークには、大きな運用経費（ＯＰＥＸ）と設備投資（ＣＡＰＥＸ）があり、すべての企業がそれらを購入できるわけではない。仮想プライベートネットワーク（ＶＰＮ）は、プライベートネットワークの利点をエミュレートするための技術を基礎として、インターネットを使用する。ＶＰＮは、公衆ネットワーク上で、セキュリティとパフォーマンスを確保するために、トンネリング技術に依存する。しかし、ＶＰＮは通常、単一のトンネルを使用してネットワークメトリクスを保証せずに情報を送信することは、企業にとって重要な側面である。いくつかのシステムは、トラフィックに優先順位を付けるためにマルチプロトコルラベルシステム（ＭＰＬＳ）のようなパケット交換レイヤを適用するが、これは特定のキャリアのネットワーク内でのみ機能する。必要とするものは、トラフィックが、公衆インターネットを横断するときに、サービス品質（ＱｏＳ）または品質経験（ＱｏＥ）を保証しながら、スループットなどのメトリクスを制御するために改善した方法である。

30

【０００４】

40

インターネット・プロトコル・スイート（ＴＣＰ／ＩＰ）は、データのあるポイントから別のポイントに送信する方法を指定して、エンドツーエンドの通信フレームワークを提供する。このモデルは、通常、ＯＳＩの７層アーキテクチャまたは４層スキーム（リンク、インターネット、トランスポート、アプリケーション）を通じて提示される。このシステムでは、コストを低く抑えながらエンドポイントの数が急増したため、インターネットを拡張できる。今日、公衆インターネットは、莫大な量のサービスおよびアプリケーションによる、１つのプライマリーシステムである。多くの企業が、サービスを提供し、インフラストラクチャを管理するために、インターネットを使用する。

【０００５】

公衆インターネットの大きな欠点は、レイテンシやスループットなどのネットワークメ

50

トリクスを保証する確定的サービスを提供することが、一般的に不可能なことである。この事実は、いくつかの機関や企業が特定のサービス品質（QoS）を保証するために、プライベートネットワークの構築を引き起こした。このようなネットワークは、公衆インターネットにも接続するが、最低限の目的のため、またはファイアウォールを介した、重要でないサービスのためだけに接続されている。これらのシステムは、利用と、保守とが大規模な民間企業によって処理されるため、高価な利用となる。対照的に、公衆インターネットは、統合されたプロトコルの下で動作する、世界中の共有インフラを持つネットワークの、ネットワークである。

【0006】

これらのプライベートネットワークは、物理的なものであれ仮想的なものであれ、エンドポイントを相互接続するために、さまざまなネットワークプロトコルや技術に依存する。例えば、参照により本明細書に組み込まれるインターネット技術特別調査委員会（IETF）のコメント勧告（RFC）3031に記載されているマルチプロトコル・ラベルスイッチング（MPLS）は、仮想専用通信チャネルのおかげで、標準インターネット接続と比較して、制御可能な性能および信頼性を提供する。他の最適化技術の中でも、これは、アプリケーションタイプなどの、様々なパラメータに基づき、トラフィックに優先順位を付ける。しかし、MPLS利用は通常、1メガビット当たり100倍多いオーダーで、はるかに高いコストを伴う。“What is the cost of MPLS?” マッシュルームネットワークブログ（Mushroom Networks Blog）2015年8月20日の例を参照。これらは参照により本明細書に組み込まれる。

【0007】

近年では、クラウドコンピューティングの出現により、相互接続問題がいっそう悪化した。“A View of Cloud Computing” M. Armbrustらを参照。これらは参照により本明細書に組み込まれる。クラウドコンピューティングは、オンデマンドサービスとして提供される柔軟なインフラストラクチャの外部化を提供し、実際に使用されるリソースに対して支払いする。しかし、クラウドコンピューティングモデルには、データが生成および消費される企業間の通信と、データを処理し保存することがあるデータセンターとにおいて重要なポイントがある。この事実は、クラウドコンピューティングパラダイムの統合を妨げた。一部の業界リーダーによると、公衆クラウドにはわずか5%のワークロードしかない。主な理由は、セキュリティ、ロックインコスト、データのプライバシー、およびネットワークのコストであり、これらは、ITチームの躊躇によりいっそう悪化する。

【0008】

重要なアプリケーションを管理することができるネットワークの使用、および重要ではないアプリケーションに対して特定の境界を確保する必要性など、クラウドコンピューティングの可能性を達成するために必要とされる欠落した部分がある。最近では、大企業のみが、プライベート広域ネットワーク（WAN）を必要とする、決められたリンクのために、支払いをおこなうことができる。さらに、これらのネットワークは、パケット交換ネットワークではなく、ポイントツーポイント接続に依存しているため、スケーラビリティの問題があり、公衆インターネットなどのネットワークの主な利点を放棄する。後になって、最大のクラウドサービスプロバイダは、プライベートデータセンターを公衆クラウドに接続するための解決策を提供することにより、この問題に取り組んできた。例えば、アマゾンウェブサービスはダイレクトコネクトを提供し、マイクロソフトアズールは、エクスプレスルートを提供する。しかし、これらの解決策は、これらのクラウドへのプライベート接続を使用して、問題を処理する。ゆえに、低コストでのネットワークコネクションの問題および信頼性を維持するエクセレントスケーラビリティは、未解決のままである。

【図面の簡単な説明】

【0009】

図1は、エンタープライズシナリオの例を示す。

【0010】

図 2 は、エンタープライズシナリオの例を示す。

【 0 0 1 1 】

図 3 は、クラウドシナリオの例を示す。

【 0 0 1 2 】

図 4 は、クラウドシナリオの例を示す。

【 0 0 1 3 】

図 5 は、グローバルシナリオの例を示す。

【 0 0 1 4 】

図 6 A は、M P T C P シナリオの例を示す。

【 0 0 1 5 】

図 6 B は、標準的な T C P と M P T C P プロトコルの比較を示す。

【 0 0 1 6 】

図 7 は、企業本部へ通じる支社の例示的な実施形態を示す。

【 0 0 1 7 】

図 8 は、複数の C S P 内のフローの分割を示す。

【 0 0 1 8 】

図 9 A は、異なる I S P とインターネットのコアとを横断する C S P を示す。

【 0 0 1 9 】

図 9 B は、インターネット上の異なる経路を示す。

【 0 0 2 0 】

図 1 0 は、C S P の状態のダイアグラムを示す。

【 0 0 2 1 】

図 1 1 は、3 層 C S P 管理アーキテクチャを示す。

【 0 0 2 2 】

図 1 2 は、単一の企業の観点から見たシステムアーキテクチャを示す。

【 0 0 2 3 】

図 1 3 は、2 つの異なるエンタープライズを伴うグローバルの観点から見たシステムアーキテクチャを示す。

【 0 0 2 4 】

図 1 4 は、企業間中間層を含む、2 つの異なるエンタープライズを伴うグローバルの観点から見たシステムアーキテクチャを示す。

【 0 0 2 5 】

図 1 4 B は、各企業の特定の副層と結合副層とを有するより複雑な中間層を示す。

【 0 0 2 6 】

図 1 5 は、アーキテクチャ層内の学習構成要素を示す。

【 0 0 2 7 】

図 1 6 は、各レベルに対する入力および出力を含む階層的学習アーキテクチャを示す。

【 発明を実施するための形態 】

【 0 0 2 8 】

詳細な明細

本発明の実施形態は、その現在のインフラストラクチャを使用して、アプリケーションのサービス品質 (Q o S) または品質経験 (Q o E) を保証するために、公衆インターネットを賢明に使用する。サービス品質 (Q o S) は、客観的にサービスパラメータ (パケットロス率や平均スループットなど) を測定するが、品質経験 (Q o E) は顧客のサービス体験 (例、W e b 閲覧、携帯電話、テレビ放送、またはコールセンターへの電話) を測定する別の関連概念である。主な前提は、保証した Q o S / Q o E を提供すると同時に、低価格とそのスケーラビリティを維持することである。

【 0 0 2 9 】

本発明の実施形態を使用して、より効率的で制御された方法で情報を送信するために複

10

20

30

40

50

数のエンドポイント間に一組の接続が確立される。これらの接続を通じて、スループット、パケット損失およびレイテンシなどのネットワークメトリクスに対して、低コストを実現し、アプリケーションのQoS/QoEを保証することができる。本発明の実施形態を採用することは、公衆インターネットのような既存のネットワークが企業クラス接続を提供することを可能にし、それは現在のインターネットインフラストラクチャを修正することなく、企業クラウド採用を加速するために使用できる。

【0030】

本発明の実施形態は、ほとんどのネットワーク、特にインターネットにおいて未使用の容量があるという観察に基づく。トラフィックのピークを吸収するために、コア内のルータは過剰にプロビジョニングされているため、低い使用率レベルで稼働する。つまり、輻輳時を除いて、未使用の帯域幅が十分にある。この追加の帯域幅を使用するために、利用するツールおよび技術は、マルチパスプロトコルを使用することができる。スループット、トランザクション時間、データの一貫性、レイテンシ、およびパケット損失など、さまざまなネットワークメトリクを改善するために複数のパスが同時に使用される。さらに、動的な決定は、ネットワークの状況に応じて行われる。

【0031】

以下の頭字語はここで使われる：

【0032】

CSP サーティファイドスターフロー（商標）パス（Certified Star flow（商標）Path）

【0033】

QoE 品質経験

【0034】

ASP 悪化安全経路（Aggregated Secured Paths）

【0035】

BW 帯域幅

【0036】

QA 品質保証

【0037】

LCI ローカルコンテキスト情報

【0038】

GCI グローバルコンテキスト情報

【0039】

本明細書では、インターネットという用語は、公衆または私的を問わず、公衆インターネット、またはインターネット・プロトコルを利用する他のパケット交換ネットワークを意味するために使用される。本発明の実施形態は、異なるシナリオの多様性の下で用いられることができる。下記に記載されたものは、関心のあるシナリオのうちのいくつかのものである。

【0040】

シナリオ1：オフィス相互接続

【0041】

本シナリオでは、図1の描写では、支社100、110、120、130）は、本部HQ（140）への接続を確立する。公衆インターネット（150）上の複数のルーター（例、160）は、接続を確立するために使用される。近年では、使用したプライベートリンクは、もはや必要ないが、なぜなら本発明は、QoEの補償を提供するからである。プライベートリンクへの依存よりも、インテリジェントポリシーは、マルチパス技術を使用して、インターネット、3G/4G/5Gモバイルネットワーク、および/またはプライベートネットワークの既存の過剰容量を利用する。この解決策により、企業はOPEXとCAPEXを削減しながら、高いパフォーマンスを発揮する、仮想プライベートWANを実現する。

10

20

30

40

50

【 0 0 4 2 】

図 2 では、エンタープライズシナリオは、モバイル端末と共に示される。支社（ 2 0 0 、 2 1 0 ）に加えてモバイル端末（ 2 2 0 、 2 3 0 、 2 4 0 ）は、ルータ（ 2 6 0 ）および本部（ H Q ）（ 2 7 0 ）を通して、公衆インターネット（ 2 5 0 ）に結合される。モバイル端末は、オフィス相互接続シナリオに新たな次元を追加する。本ケースでは、端末は、物理的な場所の境界内に制限されていないが、異なるモビリティパターンを有する。例えば、モバイル端末は、H Q との接続を確立するが、特定の支社に滞在する。次いで、従業員は、その端末を持って、別の場所にいる顧客を訪問できる。本発明は、接続を切ることなく、または Q o E を劣化することなく、それらの状況を扱う。さらに、企業の敷地内で、追加の制約が適用される場合がある。企業ポリシーは、すべての端末が、ファイアウォールを使用するために、1 つオフィスに最初に接続することを必要とすることを特定できる。他の企業は、H Q に直接的に接続するモバイル端末を許可できる。これらは、ハイレベルな状況であるが、モバイル端末の数およびそれらのモビリティパターンが、本発明によって提示される解決策に影響を与えることは明らかである。

10

【 0 0 4 3 】

シナリオ 2 : クラウド : クライアントを伴うクラウドデータセンターの接続 :

【 0 0 4 4 】

図 3 は、クラウドを含むシナリオを示す。支社と本部を接続する以外に、複数のクライアント（ 3 0 0 、 3 1 0 、 3 2 0 、 3 3 0 ）は、ルータ（ 3 6 0 ）を通した公衆インターネット（ 3 5 0 ）を通して、公衆クラウドデータセンター（ 3 4 0 ）へ結合する。それらは、クラウドスペースを通した 2 つの主な学校である。それらのリードする企業（例、アマゾンおよびグーグル）は、排他的に公衆クラウドにフォーカスし、すべての処理を実行し、データを保存する。対照的に、他の企業（例、マイクロソフト）は、ハイブリッドクラウドモデルを提唱し、ワークロードの分割を可能とし、公衆クラウドおよびプライベートエンタープライズクラウド間に保存する。

20

【 0 0 4 5 】

本発明の実施形態は、エンタープライズシナリオとしてのクラウドシナリオで、同様の根底にある技術を使用する。各エンドポイント（クライアントまたはクラウド）は、マルチパス技術を使用して複数のパスを同時に有効にし、それによって必要なアクセス Q o E を確保する。

30

【 0 0 4 6 】

モバイル端末も、クラウド・クライアントとして考慮することができる。図 4 は、モバイル端末を併合するクラウドシナリオを示す。クライアント（ 4 0 0 、 4 1 0 、 4 2 0 ）およびモバイルクライアント（ 4 3 0 、 4 4 0 ）は、ルータ（ 4 6 0 ）を介した公衆インターネット（ 4 5 0 ）の使用と共に、クラウドデータセンター（ 4 7 0 ）に結合される。クラウドおよびエンタープライズシナリオ間の主な相違は、多種多様なアプリケーションの種類、トラフィックパターン、リソースの共有、とりわけその他のリソースの管理を含む。

【 0 0 4 7 】

シナリオ 3 : グローバルシナリオ

40

【 0 0 4 8 】

図 5 は、グローバルまたはオールツーオール・シナリオを示す。このシナリオは、モバイル端末を含む、エンタープライズおよびクラウドシナリオの結合による結果である。クライアント（ 5 0 0 、 5 2 0 、 5 3 0 、 5 4 0 ）およびモバイルクライアント（ 5 1 0 、 5 5 0 ）に伴うエンタープライズ本部（ 5 7 0 ）およびクラウドデータセンター（ 5 8 0 ）は、ルータ（ 5 6 0 ）を通す公衆インターネット（ 5 9 0 ）に結合する。本シナリオでの本発明の実施形態は、クライアントのニーズに基づいて異なる Q o E を保証しながら、エンドポイント（例、オフィス、H Q、モバイル機器、公衆クラウドなど）間の異なる接続をサポートする。

【 0 0 4 9 】

50

イントロダクション

【0050】

インターネット移送プロトコルはもともと回復力、ロバスト性、および安定性を念頭に置いてデザインされた。加えて、各ルータの可視性は隣接ルータに限定されるため、これらはリンクごとに動作する。ゆえに、ルーティングの決定は、ネットワーク全体の状況を考慮するのではなく、ローカルの条件に基づく。このアーキテクチャ上の決定は、インターネットのスケラビリティに大きく貢献するが、スループットやレイテンシなどの他の測定基準では、そのパフォーマンスを低下させる。

【0051】

クラウドのようなモデルは、これらのサービスを提供する公衆インターネットに依存するため、ネットワークメトリクスは、全体的なパフォーマンスに大きな影響を有する。新たな問題は、データ上のコントロールの欠如などが挙げられる。それらの影響にも関わらず、企業は、インストールの容易さ、柔軟なインスタンス、より優れたコスト対効果、可用性などの要素を悪用するために、公衆クラウドを使用するよう動機付けられている。

10

【0052】

本発明の実施形態は、公衆インターネットインフラストラクチャの現存を有する最高のパフォーマンスを提供する、ネットワークの解決策を使用する。この目的のために、最適化されたルーティング技術が最終接続者に体感品質（QoE）を保証するために、異なる接続タイプ上のローカルおよびグローバルコンテキスト情報を利用するのに使用される。

20

【0053】

使用する主な技術は、優れた方法で情報パケットを伝達する複数のパスの利用である。集約パスの機能には、スループットの向上、セキュリティの強化、レイテンシの短縮、パケット損失の低減、および信頼性の向上が含まれる。

【0054】

プロトコル

【0055】

上述したシナリオに直面する企業の所望の解決策は、プライベートネットワークの利点を得ることであるが、インフラストラクチャを基礎とした公衆インターネットの使用である。このアプローチは、展開および管理費の削減をもたらすであろう。この達成を可能にする技術は、トンネルの実装に部分的に依存する。それらは、公衆インターネット上で転送するトラフィックにロバスト性、完全性およびセキュリティを追加する。

30

【0056】

本発明の実施形態は、トンネルを実装するために、仮想プライベートネットワーク（VPN）を使用する。VPNは、インターネット技術特別調査委員会（IETF）のコメント勧告（RFC）2764で説明され、参照により組み込まれる。本発明の一実施形態では、オープンVPN（<https://openvpn.net/>）を使用する。代替の実施形態では、IPSEC実装を使用する。IPSECは、インターネット技術特別調査委員会（IETF）のコメント勧告（RFC）6071で説明され、参照により組み込まれる。

【0057】

本発明の実施形態は、暗号化を使用した機密方法でデータを移送するVPNトンネルを使用する。代替の実施形態では、トンネルは、暗号化したパケットを生成して、パフォーマンスを向上するのに必要な計算時間を短縮するのに安全ではない。

40

【0058】

本発明の実施形態は、その根底にある公衆インターネットインフラストラクチャを利用する。インターネットは、パケット交換ネットワークとしてデザインし機能し、それに従って送信元と宛先との間に潜在的に多数のパスが存在する。本発明の実施形態は、複数のパスを発見し維持するためにトンネリングを利用するが、代替の実施形態は接続プールのような他の技術を利用する。接続プーリングは、接続プールを常にアクティブに維持し、それらをアプリケーションデータの転送に再利用することで構成する。接続プーリングとト

50

ンネリングの違いは、接続プーリングはパケットのペイロードを転送するのに対し、トンネルは変更していないパケットをヘッダとともに転送するため、帯域外またはカスタムプロトコルを介して、追加の詳細情報を送信する必要があることである。本発明の一実施形態では、TCPプーリングは、インターネット上でTCPデータを移送するために使用する。

【0059】

マルチパス・トランスポートは、ネットワーク機器の変更または再設定を必要とせずに、一連の異なるパスを使用できるようにする。このことは、マルチパス・トランスポートは、ルータに対して透過的である。本発明の実施形態は、マルチパス・トランスポートの利点とVPNを組み合わせて、事前に確立し特徴付けられたトンネルのセットを介して、データを同時に転送する。OSIデータリンク層または上記のイーサネット、ARP、ICMP、IP、TCP、またはUDPといった特徴に適合するプロトコルのパケットを伝送することができる。

10

【0060】

多くの技術は、複製や分散など、複数のパスを活用できる手法を定義できる。複製は、情報フローを複製して、最高のパフォーマンスを得るために、異なるパスを通してそれらを送信する。信頼の向上は、コストを生じさせる。この技術は、複数回送信された同じトラフィックが、互いに干渉してルータを輻輳する、大きなオーバーヘッドを有する。対照的に、分散は、情報フローをサブフローに分割し、それらを異なるパスで送信する。この技術は混雑を減らす、潜在的に順不同問題を悪化させることができる。転送するトラフィックの特定の要件に従って、パスを賢く使用する場合は、パフォーマンスが向上する。両方の技術は、複数のパスを作成するのに、共通のオーバーヘッドを共有する。

20

【0061】

インターネットのマルチコネクティビティを利用することに加えて、近年の発明の実施形態はまた、一般にそのスループットおよび性能を改善するためにVPNおよびTCPプロトコルパラメータを微調整する。例えば、一実施形態は、巨大な最大伝送ユニット(MTU)値(イーサネット9KBジャンボフレーム、例えば48KB MTUよりもさらに大きい)を使用して、LANをシミュレートするためにVPNトンネルによって提供する抽象化を利用することができる; TCP最大セグメントサイズ(MSS)は、この調整に適応し、トンネル仮想インターフェイスに大きなフレームを挿入することができる; これらの巨大なフレームは、パス上のルータまたはそれらを生成する同じ物理/仮想ホストによって、複数のIPフラグメントに断片化される。このようにして、送信側のTCP輻輳制御アルゴリズムは送信ウィンドウ値のより速い成長を示し、その結果、一貫してスループットが向上する。

30

【0062】

これらの利点を活用するために、TCPとUDPが主な指数であるが、異なるプロトコルを使用できる。近年の発明の実施形態は、トンネルを通してパケットを転送する。トンネルの選択とパケットのスケジューリングは、パケットの優先順位付け要件に一致する、トンネルメトリックに基づく。

【0063】

MPTCP

40

【0064】

マルチパス・トランスポートへの1つのアプローチ、マルチパスTCPは、インターネット技術特別調査委員会(IETF)のコメント勧告(RFC)6182に記載されている、マルチパスTCP開発のためのアーキテクチャガイドラインに記載されており、参照により明細書に組み込まれる。マルチパスTCP(MPTCP)は、複数のパスを可能とするTCPの延長である。図6Aは、MPTCPの単純な例を示す。2つの接続(A1、A2)は、ホストA(600)をインターネット(610)と結合し、2つの接続(B1、B2)は、ソフトB(620)をインターネット(610)に結合する。ホストAからホストBへの各パスは、IP送信元アドレスと宛先アドレス、IP送信元ポートと宛先

50

ポート番号、および使用するプロトコルで構成する 5 タプルによって、一意に識別される。

【 0 0 6 5 】

上述の通り、インターネットを通したパスは、分断される必要はない。例えば、A 1 - B 1 および A 1 - B 2 は、ネットワーク内の共通のリンクを共有し、巧妙な選択は、輻輳と、相互干渉とを防止する。

【 0 0 6 6 】

M P T C P は、2 つの主な目的を有し：(i) 複数のパスを同時に使用したスループットの向上および(i i) いくつかのパスを通してセグメントが送付できることによる回復性の向上。ノード障害は、M P T C P の回復力が重要になる極端な場合である。いくつかの状態では、M P T C P は、T C P をアウトパフォームできる。実際には、実装の詳細は、考慮しなければならない。例えば、M P T C P のオーバーヘッドコストは、小さなファイルの転送で、その利点を相殺する可能性がある。

10

【 0 0 6 7 】

M P T C P の使用が一般化する場合、それはまた予備容量のより良い使用法を通して混雑したボトルネックからトラフィックを移動することによって、インターネット全体の混雑の減少をもたらす可能性がある。

【 0 0 6 8 】

M P T C P 建設用スタックのオーバービューは、図 6 B に描写される。標準的な T C P (6 3 0) は左側に示され、M P T C P (6 4 0) は右側に示される。M P T C P 層は、パス管理、パケットスケジューリング、サブフローインターフェイス、および輻輳制御を扱う。サブフローは、転送の基礎となる各パスを与える標準的な T C P セッションである。これらのすべての詳細は、アプリケーションに対して透過的である。

20

【 0 0 6 9 】

L i n u x 用の参照 M P T C P 実装は、カーネル空間で実行する。本発明の実施形態は、参照 L i n u x 実装の特定の、M P T C P カーネルモジュールに対する修正を含む。これらのモジュールは、ユーザスペースで実装されるプロセスからの M P T C P 接続を最適化するネットリンクを通して、コマンドを受信する。

【 0 0 7 0 】

一実施形態では、デフォルトのフルメッシュトポロジを使用する代わりに、クライアント（例えば支社）と、サーバ（例えば H Q ）との間のポイントツーポイント接続を改善する新しいカーネルモジュールが利用される。このモジュールは、不使用の接続の排除を通して、マルチパス転送のオーバーヘッドを削減する（例、同じオフィスの同じ端末の 2 つの I P ポート間）。一実施形態では、参照 M P T C P ・ L i n u x 実装のインターフェイス制限が増大された。この修正と共に、フローは、元の実装で課せられている制限である 8 つではなく、最大 3 2 のインターフェイスを使用できる。代替の実施形態では、インターフェイスとサブフローの最大数を、増大することができる。

30

【 0 0 7 1 】

代替の実施形態では、フルユーザスペース実装と共に、ユーザスペース I / O （例、D P D K ）は、利用できる。本実施形態は、ハイブリッド・カーネルユーザスペースの解決策よりも効果的である。これは、カーネルからユーザスペースへの通信および、その逆の通信に関連するオーバーヘッドを削減しながら、プロトコル実装の制御およびパフォーマンスの最適化を可能にする。もう一つの代替実施形態では、フル・カーネルスペースを実装することができる。

40

【 0 0 7 2 】

本発明の実施形態では、2 つの M P T C P サブモジュールの新バージョンは、利用可能であり：パスマネージャーとパケットスケジューラである。パスマネージャーは、確立するために、サブフローの数を決める。新バージョンは、サブフローを動的に作成して閉じると共に、他の修正を以下に記す。スケジューラは、パケットをサブフローに割り当てる。新しいスケジューラアルゴリズムは、サブフローとアプリケーション間で、負荷を分散

50

する。

【0073】

マルチパスUDP (MPUDP)

【0074】

UDPは、他のメジャーなトランスポートプロトコルである。UDPは、まだ標準化されていないMPUDPをサポートするために拡張するが、基本的な考え方は同じであり：UDP接続で複数のパスを同時に利用できるようにする。TCPとは異なり、UDPは信頼できるトランスポートプロトコルではなく、例えば、それはアプリケーションに欠落、順不同および重複パケットという課題を残す。

【0075】

MPUDPは、信頼できないトランスポート、輻輳制御およびパケット順序付けの保証の欠如のUDP属性を引き継ぐ。近年、3つの標準的なトランスポート層のプロトコルのみが、信頼できないトランスポートを実装する：UDP、DCCP およびSCTPである。UDPは輻輳制御を持たないため、主に低スループット要件の転送で使われる。アプリケーションの要件に応じて、異なる主なメトリクスは最適化を要する。例えば、いくつかのアプリケーションでは、パケット損失を減らすことは非常に重要であるが、一方では、一方向の遅延を減らすことがより重要である。マルチパスは、このトピックに非常に役立つ。例えば、パケット損失を減らすために、UDPトラフィックは、混雑の少ないパスに移行できる。一方の遅延を減らすために、UDPトラフィックを複製し、同時に複数のパスを通して送信できる。後者の場合、受信側で複製を排除することが重要である。これは、UDP転送を、DCCP転送に変更することで実現できる。

【0076】

データグラム輻輳制御プロトコル(DCCP)は、TCPのような輻輳制御(セッションハンドシェイクとシーケンス番号を使用)で信頼性の低い転送を実装する。DCCPは、参照により本明細書に組み込まれるインターネット技術標準化委員会(IETF)のコメント勧告(RFC)4340に記載され、参照により明細書に組み込まれる。

【0077】

ストリーム制御伝送プロトコル(SCTP)には、部分信頼性(PR)の拡張機能を有し、期限切れデータの再送信を防止する。SCTPは、インターネット技術特別調査委員会(IETF)のコメント勧告(RFC)4960に記載され、参照により明細書に組み込まれる。

【0078】

これら3つのプロトコル間では、DCCPおよびSCTPのみが輻輳制御を有し、SCTPのみがマルチストリームおよびマルチホーミング(マルチパス)のサポートを有する。ゆえに、効率的なMPUDPのための2つの可能な候補は、マルチパスDCCP(手動でパスマネージャーとパケットスケジューラを実装する)か、または部分信頼性とマルチホーミングを持つSCTPになり得る。前者のようなアプローチは、C. Huang, Y. ChenおよびS. Linによる“マルチパスデータグラム輻輳制御プロトコルのためのパケットスケジューリングおよび輻輳制御方式”に提案され、参照により明細書に組み込まれる。後者のようなアプローチは、C. HuangおよびM. Linによる“マルチホームネットワークのための部分的信頼性同時マルチパス転送(PR-CMT)”に提案され、参照により明細書に組み込まれる。

【0079】

加えて、いくつかのUDPベースのアプリケーションは、マルチパス配信によって悪化する可能性があるパケットの順序に敏感である。DCCPおよびSCTPは、この問題を阻止するために、パケットの並び替えをサポートする。

【0080】

メトリクス

【0081】

本発明の実施形態で利用するトンネルは、ネットワークメトリクスの用語で評価する。

10

20

30

40

50

それらのいくつかは以下である：

【 0 0 8 2 】

帯域幅（BW）：帯域幅は、一定時間内に伝達できる理論上の最大データ量として定義する。したがって、帯域幅は、データ転送をサポートするためのネットワーク接続の容量を表す。帯域幅は、1秒ごとのビットでしばしば表現される（bps、Kbps、Mbps、Gbps）。

【 0 0 8 3 】

パケット損失：パケットスイッチシステムでは、パケット損失とは、目的の宛先に到着できなかったパケットの数のことである。パケット損失の原因となる主な要因は、リンク輻輳、端末性能（ルータ、スイッチなど）であり、バッファの過負荷、ネットワークデバイスのソフトウェアの問題、およびハードウェアの障害である。破棄は、パケットの意図的な破棄である。

【 0 0 8 4 】

信頼性：信頼性とは、指定された条件下で、特定の期間システムまたはコンポーネントが機能する能力を示す。

【 0 0 8 5 】

スループット：スループットは、実際にどのくらいの量のデータが実際に“チャンネル”を通過するかを示す。これは、レイテンシ、パケット損失、および使用されるプロトコルなど、さまざまなことによって制限される可能性がある。スループットは通常、1秒あたりのビット数（bps、Kbps、Mbps、Gbps）で測定する。

【 0 0 8 6 】

レイテンシ：レイテンシは、アプリケーションが送信するデータを生成してから送信先のアプリケーションに到着して処理されるまでの時間として定義する。パケット交換ネットワークにおけるレイテンシは、特に処理遅延、バッファ満杯および待ち行列遅延などの長距離ネットワークの動作環境において、多くの異なる要因によって影響を受ける可能性がある。

【 0 0 8 7 】

ジッタ：ジッタは、同じストリームに属する、2つの連続した受信パケットの転送遅延間の差の絶対値である。ジッタは、ネットワークの輻輳、タイミングドリフト、およびルート変更に起因する。インターネット技術特別調査委員会（IETF）のコメント勧告（RFC）3393の報告では、参照により明細書に組み込まれ、パケット遅延の変動を識別するためのパケット交換ネットワークにおけるジッタという用語は、完全には正しくない。パケット遅延の変動（PDV）は、この文脈で使用するには、より良い用語であり得る。

【 0 0 8 8 】

往復時間（RTT）：往復遅延時間とも呼ばれる往復時間は、パケットが特定の送信元から特定の宛先に送信され、返信パケットが送信元に送信されるまでに必要な時間である。

【 0 0 8 9 】

インターパケットタイム（IPT）：インターパケットタイムは、フロー内で2つの連続するパケット間で経過する時間である。インターパケット到達時間（IPAT）は、インターパケット発行時間（IPE T）と比較する場合に、ジッタを計算するのに、便利で効率的な方法を提供する。

【 0 0 9 0 】

フロー完了時間（FCT）：フロー完了時間は、ネットワークフローを使用して成功した、トランザクションを実行するために必要な時間である。トランザクションの種類とその正確性はアプリケーションによって異なる。

【 0 0 9 1 】

観察可能接続経路：観察可能接続経路は、フローの構成要素パケット（例、2つのエンドポイント間の持続的トンネル接続）が通過する測定可能なノードの集合である。

10

20

30

40

50

【 0 0 9 2 】

サーティファイドスタートフロー（商標）パス（Certified Starflow（商標）Paths）

【 0 0 9 3 】

本発明の実施形態の重要なコンセプトは、“Certified Starflow（商標）パス”、またはC S Pである。C S Pは、ネットワークの2つのエンドポイント間で開かれた持続的な接続である。発明の実施形態では、この持続的な接続は、V P Nトンネルの機能により実装する。接続が開かれて生存していると、前述のようなメトリックが監視され始める。トンネルが所望の間隔の間に或るしきい値を通過すると、トンネルは“Certified Starflow（商標）パス”、またはC S Pの状態へと誘導する。同様に、C S Pは、所望の間隔の或るしきい値を維持できない場合に、降格される。異なるアルゴリズムは、所望のQ o E s要求と共に他の要因に従って、昇格および降格プロセスを支配する。例えば、企業が、1 0 M b p s以上のスループットを必要とするQ o Eを有する場合、このしきい値を超えるV P Nトンネルのみが、促進したC S Pを得る。代替の実施形態では、C S Pは、セキュリティまたは暗号化を使用しない。これらは、トンネリング技術で実装できる。

【 0 0 9 4 】

インターネットなどのパケット交換ネットワーク内のパスは、2つのエンドポイント間で交換されるパケットが通過する一連のホップとして定義できる。そのセットは、各ホップでの次第である。インターネット内のルータは、それらを一意に識別し（例えば、フロー識別子としての5つのタプル）、アクティブである間それらの対応するルーティング状態を維持することによって、同じフローに属するパケットを転送する傾向があり、これはフロースティックネスと呼ばれる。

【 0 0 9 5 】

発明の実施形態は、インターネットによる動的な複数のパスを活用する。本発明の実施形態では、I S Pは2つのエンドポイントを接続するために使用でき、それらは別々の管理ドメインを横断するため、それらの間に複数のパスを暗黙的に提供する。エンドポイントごとに単一のI S Pであっても、2つのエンドポイント間のトラフィックフローイングは別々のパスを通過する。これは、パス上の一部のルータに同じ宛先への複数の発信ルートがあるからである。この多数のネクストホップは、トラフィックを負荷分散するためにルータによって利用され、それによってパフォーマンスを向上させ、輻輳を軽減する。マルチパスルーティングの状況で、接続指向プロトコル（T C Pなど）を破壊しないようにするために、ルータは同じフローに属するパケットを同じネクストホップに送信する傾向があり、フローのスティッキ性を強化する。本発明の実施形態は、“パスフィッシング”として定義するプロセスにおいて、これらの接続を確立したエージェントに関係なく、エンドポイント間の接続を積極的に維持する。

【 0 0 9 6 】

同じ2つのエンドポイント間に異なるルートを持つC S Pを検出することは、ルータがフローごとのロードバランシングをどのように実施するかによって異なる。特定のフローに属するパケットを識別し、それらを同じネクストホップにルーティングするために、通常、フローごとの不変フィールドに対してハッシュ値を計算する。一例は、U D P / T C Pフローパケットの5タプルであり：I P送信元、I P宛先、ポート送信元、ポート宛先、および転送プロトコルI Dである。本発明の実施形態では、U D Pトンネルを使用する。他の実施形態では、I P s e cまたは他のトンネリング手法を使用できる。いくつかのケースでは、送信元または宛先ポートは存在しない。（例、E S Pヘッダを使用してトンネリングモードで動作するI P s e cパケット内であり、これはトンネルモードでのI P s e cの最も一般的な設定である。）この場合、同じI P s e cトンネルに属するすべてのパケットに対して安定したパスを維持するために、パス上のルータは、ヘッダの他のバイトを送信元I Pおよび宛先I Pとともにハッシュ値として使用できる。

【 0 0 9 7 】

発見されたC S Pは、関連するメトリックの集合に基づいて特徴付けられ、パスはそれらのうちの1つである。特徴付けは能動的または受動的のいずれかである。前者はアクテ

ィブプローブを使用して測定値を取得するが、後者は実際に転送されるトラフィックから測定値を収集する。同等のパスを持つ複数のＣＳＰが検出された場合でも、他のメトリクスに関しては、異なる動作ができる。２つのサイト間のＣＰＳの数は動的であり、必要なＱｏＥに応じてインテリジェントで管理する。

【００９８】

発明の一実施形態では、ＣＳＰの数は、唯一のトンネル識別子の数の部分で制限する。例えば、５タプル技術を使用する場合、特定の送信元から特定の送信先へのＣＳＰの数は、エンドポイント間で確立できる、一意の５タプルの数によって決まる。

【００９９】

図７は、本部のエンタープライズに接続する支社の、例示的な実施形態を示す。本実施例では、支社は２つの異なるＩＳＰ接続を有し、単一のＩＰを伴う１つのＩＳＰ（７００）は利用可能な２つの転送ポートを指定し、２つのＩＰを有する他のＩＳＰ（７１０）は３および２つの転送ポートをそれぞれ指定する。接続の他の側面では、本部は、１つのＩＰアドレスと２つの転送ポートを伴う単一のＩＳＰ（７２０）を有する。

【０１００】

ネットワーク・セットアップのトップでは、ローテーション間の全ての可能なＣＳＰ（７３０）も示される。支社内の全てのＣＳＰの注釈は、表示されない。これは単純なネットワークトポロジであるが、エンタープライズはそれらのオフィスを相互接続するのにより複雑なトポロジを必要とし得る。その結果、トポロジは、ＣＳＰを決定づける。

【０１０１】

本例では、ＩＰ転送ポート制限は１４のＣＳＰが可能となり、支社で可能なＩＰ転送ポートのペアの数の合計と、本部で可能なＩＰ転送ポートのペアの数の合計として表現できる。したがって本発明の実施形態は、これら１４以上のＣＳＰのトラフィックを潜在的に転送できる。

【０１０２】

ネットワークの構成がＣＳＰの数を制限するので、伝達下のフロー数は、ユーザやアプリケーションによる。また、フロー数はＣＳＰの容量によって制限され、各接続の帯域幅である。さらに利用可能なＣＳＰにトラフィックを分散しながら、システムの粒度を高めるために、フローはサブフローに分割される。

【０１０３】

本発明の一実施形態では、フローは最大３２の異なるＣＳＰを同時に使用でき、各ＣＳＰはそのフローのサブフローの総数のサブセットを処理できる。代替実施形態では、使用できるＣＰＳの異なる数は、３２以上とすることができる。これは図８に示される。単一のフロー（８００）は、最大３２個のサブフロー（８１０）に分割され、最大６個のＣＳＰ（８２０）を通して伝達する。

【０１０４】

本発明の一実施形態では、オリジナルＭＰＴＣＰプロトコルを利用し、すべてのエンドポイント間でフルメッシュ接続を確立する。代替実施形態では、未使用トンネルのオーバーヘッドを排除し、可能なサブフローの数を増やし、ＣＳＰ割り当てのフローの制限を克服するために、ポイントツーポイント・トポロジを使用する。

【０１０５】

ＣＳＰ方向性

【０１０６】

本発明の一実施形態では、ＣＳＰは、トンネルから継承された特性として双方向である。両方のＣＳＰ方向が同じネットワークパスを使用するかどうかは、ネットワークインフラストラクチャによって異なる。例えば、両方向のＩＰベーストンネルは、同じネットワークパスを通過する必要はない。それは、異なる特性（例、レイテンシ、または帯域幅の観点）を潜在的に有するトンネルの各方向となる。

【０１０７】

しかしながら、いくつかのアプリケーションおよび／またはプロトコルは、両方のフロ

10

20

30

40

50

一方向の特徴で、本非対称性により影響を受けることがある。例えば、新しいデータを転送する前のプロトコルの確認応答パケット（ACK）への依存性のため、TCP接続はRTT（双方向の合成遅延）に大きく依存し、グローバルなエンドツーエンドのQoSを保証する。

【0108】

本発明の一実施形態では、各データパケットは、その時点でのアプリケーションの特性またはシステムのポリシーにより、よりよく適合するCSPを通して転送できる。例えば、TCP接続は、与えられたCSPを通して最初に確立することもあるが、その後、データパケットは、より高い帯域幅などのより良い特性で同じ宛先に到達する、代替CSPを通して送信できる。さらに、対応するACKパケットは、レイテンシなどの他の特性に基づいて、必ずしもデータパケットと同じではない、これらのCSPのうちのいずれかを通して戻ることができる。これらの決定は動的であり、時間とともにパケットごとに変わる可能性がある。

10

【0109】

本発明の一実施形態では、CSPは、上述したメトリクス的手段によって、識別される。加えて、フローは、それらの同じメトリックに加えて、それらのフローを開くアプリケーションによって定義された特定の特性（例えば、トラフィックの種類、ファイル転送サイズなど）に関して、一連の要件を有する。システムは、フローのパケットを転送するためにアプリケーションのニーズを最適化するCSPを選択できる（VoIPや、大量のデータ転送に使用可能な帯域幅が広いCSPなど、レイテンシに敏感なアプリケーション用の低レイテンシCSP）。

20

【0110】

CSPを分類するために、トンネルが何らかの基準に違反したときに、動的に反応するトンネルの継続的な監視がある。例えば、CSPはスループットに従い分類できる。分類の実装では、CSPは、パケット損失メトリクスに従い分類できる。例えば、カテゴリAには特定のしきい値より高いスループットを持つCSPが含まれ、カテゴリBにはそのしきい値を下回るがパケット損失が少ないCSPがある。

【0111】

加えて、分類パターンは、ある種類のトラフィックを伴うCSPを随伴する。例えば、大きいファイルを送信する場合、所望のCSPはエクセレント・スループットを有するべきである。代替的に、データ受理確認を送信する場合、レイテンシの短縮はスループットよりも好ましい。さらに、フローはユーザおよびアプリケーションの種類に従い、異なる優先を持つことができる。例えば、分類は、リアルタイムやQoSなどの特性を損なうことなく、これらの優先順位を保証するのに役立つ。

30

【0112】

利用

【0113】

本発明の実施形態は、任意の装置またはネットワーク機器に利用することができる仮想マシン（VM）内で動作する“エージェント”として配置される。VMの実行場所によって、CSPの送信元と送信先が決まる。VM展開オプションは3つあり、システムパフォーマンスの衝撃を有する：1．WANを伴うLANを接続するネットワーク要件でのVMの実行；2．LAN内の或る集約ポイントでのVMの実行；および3．LAN内での各端末でのVMの実行である。

40

【0114】

代替の実施形態では、エージェントはVMではなく、コンテナの仕様で展開する。コンテナは、オペレーションシステムによって提供される軽い仮想化メカニズムであり、プロセスのグループまたはシステムリソースのグループを分離することを可能にする。例えば、Linuxでは、アプリケーションとネットワーク端末は、ネームスペースで分離できる。

【0115】

50

いくつかのケースでは、C S Pは異なるネットワークドメインを横切ることがあり（例、異なる自律システム、A S）、インターネットコアのルータを含む。加えて、上記の3番目のオプションを選択する場合、C S Pは、L A Nも横断する。図9 Aはこの状態を示す。2つのI P S接続（9 1 0）を有する1つのエンドポイント（9 0 0）は、インターネットコア（9 2 0）を通した2つのI S Pをも有するもう1つのエンドポイント（9 4 0）に接続する。各C S Pは、エンドポイントの1つのI P / 転送ポートのペアでスタートし、他のI P / 転送ポートのペアで終着する。

【0 1 1 6】

エージェントが配置する場所は、C S Pの総数に影響する。本決定は、接続するエンドポイントの数を決定づける。例えば、上記の1番目のオプションを使用し、L A Nの入出力でV Mを展開する場合、C S Pの数は、ポイントの入出力で利用可能なC S P識別子の数を制限するために、削減する。開かれたポートに関するセキュリティ上の懸念があるため、この制限はエンドポイントによって公開する利用可能な転送ポートのポート数に起因する。対照的に、V MがL A N内の全ての端末で展開する場合、より多くのC S P識別子が典型的に利用可能となるため、C P Sの数はより多くなることができる。この選択は、計算能力、セキュリティ、および会社の方針の観点からも影響を及ぼす。各端末でのV M作動は、いくつかの企業では望まれないだろう。なぜなら彼らはL A Nを保護するためにファイアウォールを使用することが好ましいからである。

【0 1 1 7】

すべての端末上でV Mを作動することはエンドツーエンドQ o Eの結果になり得るが、L A Nの出入り口のポイントでこれらを用いることは、Q o Eがこれら2つのポイント間のみで補償されるため、パフォーマンスの不確実性となり得る。さらに、内部のルーティングオプションは、典型的に制限するので、企業内のC S Pは、これらのリンクの大部分を共有しがちである。この事実は、相互干渉と輻輳が原因で、パフォーマンスに衝撃を与えることができる。

【0 1 1 8】

3つ目の選択肢は、前2つ間で中級の解決策である。すべてのL A N端末でV Mを作動するのでなく、特定の重要な集約のみがV Mを行う。この手法は2番目の選択肢ほど積極的ではなく、L A N内でもQ o Eを認証できる。

【0 1 1 9】

C S Pアサイメント

【0 1 2 0】

公衆インターネットを通した2つのエンドポイントを接続するパスの数は、特に大きくなることができる。エージェントは、C S Pとして昇格および利用するトンネルの適切なセットの決定に責任的である。パケットをスケジュールし転送するC S Pの選択は、動的である。本決定は、2倍の基準、トラフィック要件およびC S Pの特徴に基づく。前者は転送されたトラフィックの種類の分類からなるが、後者はC S Pの予防的モニタリングの結果となる。これは、動的にトラフィックを適合し、ネットワークで変更するために反応する。これは、エージトラフィック要件と一致するより良いメトリクスを伴うC S Pを可能にし、これ等の1つとなる測定可能なパスとなる；例えば、異なるC S Pは、異なるネットワークパス（完全に観測可能）を横断でき、システムは信頼性を高めるために、これらの異なるC S P上でパケットを複製することが決定できる。

【0 1 2 1】

V Mの場所は独立的に作動するが、図9 Bは、C S Pのトンネルおよびプールのセット間の関係を示す。エンドポイントA（9 5 0）および エンドポイントB（9 7 0）は、中間のルータ（9 6 0）を通してインターネット上で伝達する。本シナリオでは、4つのユニークなパスがあり下の表の最初の4行に示されている。

【0 1 2 2】

加えて、V Mを作動する異なる選択肢が、示される。一実施形態では、V Mは、エンドポイント要件（9 8 0）で作動するが、もう1つの実施形態では、V Mは、企業（9 5 0

10

20

30

40

50

) のルータの出口でのみ、行われる。

【 0 1 2 3 】

以下の表は、5つの可能なパスとそれぞれが通過するノードを示す。

	交差ノード						
パス 1	0	1	2	3	5	6	7
パス 2	0	1	2	3	6	7	
パス 3	0	1	2	4	6	7	
パス 4	0	1	2	z	4	6	7
パス 5	0	1	2	3	6	7	
パス 6	0	1	2	3	5	6	7

10

【 0 1 2 4 】

2つの異なるC S Pが同じルータを通過することは可能であり、追加のトンネルが他のトンネルの1つと同じノードを交差することも可能であることに留意されたい。例えば、トンネル5はトンネル2と同じノードを通過し、トンネル6はトンネル1と同じノードを通過する。たとえ2つの異なるC S Pが同じノードを通過しても、これらは異なるパフォーマンスの特徴を有する。

20

【 0 1 2 5 】

本例では、エージェントには、1つのアクティブC S P (トンネル1)、スタンバイC S P (トンネル2)、およびプロービングC S P (トンネル3)を有することがある。考えられるC S Pの状態および状態遷移の1つの可能なセットは、下記で詳細に説明する。C S Pの積極的な管理とモニタリングは、それぞれを特定のしきい値に対して評価し、それに応じて行動を起こす(例、成果の低いC S Pを削除する)。プロービングC S PがアクティブC S PまたはスタンバイC S P、あるいはその両方のC S Pよりも優れていない場合は、破棄し、そのC S P ID (5タプル)が別のトンネルに再利用される。ゆえに、新しいトンネルは、他の2つの使用が促進した場合に、プロービングC S Pを分析できるようにする。これらの間のC S Pの状態と転移は、以下でさらに詳細に述べる。

30

【 0 1 2 6 】

本発明の一実施形態では、C S Pのプールは、それらの1つ以上がプロービングC S Pとなる場合に使用される。ゆえに、パフォーマンスは周期的に、例えば5分ごとに評価する。代替の実施形態では、測定間の時間は、他の論理または行動パターンに基づいてもよい。測定のオーバーヘッドを減らすために、実際のトラフィックが伝達される場合はパッシブ測定を使用でき、これらのC S Pにトラフィックがない場合はアクティブ測定を使用

40

【 0 1 2 7 】

いくつかの実施形態では、他のC S Pのどれよりも優れていないプロービングC S P / トンネルの複数の連続テストは、システムにC S Pを廃棄させ、代わりにより良いものを選択させる。C S Pを廃棄するための試験基準は、時間内に分散された連続試験からなり得る。このケースでは、C S P上のアクティブフローがあり、アクティブフローが終了するまで維持されるが、新しいフローはこれに割り当てられない。これは新しい5タプルを伴うC S P / トンネルの再起動の原因となり、前記5タプルと同じになり得る。新しいC S Pは、転送ポートの送信元で異なり、現在は送信元で利用可能なポートよりランダムに選択される。本ポートの修正は、破棄した同じパスの再利用の可能性を最小限にするため

50

に、行われる。この変更にも関わらず、“新しい”トンネルが破棄されたまたはアクティブなトンネルとことなく特性を有することは、保証されない。この理由により、エージェントは、これが異なる特性を有するか否かをチェックする。

【0128】

代替の実施形態では、同じ基本パスを使用する複数のCSPが、維持される。これは、例えば、ネットワークパスが同じであっても、異なるパフォーマンスが2つのCSPを生じることができるという事実によるものである。これらの違いは、ルータとそのキューイング、ルーティング、および/または内部ロードバランシング戦略によって発生できる。

【0129】

いくつかの実施形態では、エージェントは、CSPをオープンに維持するためにVPNソフトウェアの選択肢を通してキープ・アライブを送信する。本手法は、周期的にネットワークメッセージを発生し、例えば、トンネルを開いたままにしておくために、10秒ごとにICMP “ping” パケットを送信する。

【0130】

代替的な実施形態では、ポリシーは、新しいトンネルをプローブし、動的アルゴリズムを使用して昇格/降格プロセスを誘導するタイミングを決定するために使用する。また、CSPをテストするためのサンプリング周波数、およびネットワークの状態に適切に適応するために必要な粒度も影響を受ける。

【0131】

ルートの識別性

【0132】

CSPを区別する1つの方法は、CSPが通過するパスを通ることである。このケースで、本情報は、利用可能であり公衆インターネットにおける、もう一つの大きな関心事を満たし、可視性を獲得する。ユーザは、どのノード（例えば、IPアドレスのセット）が自分のトラフィックを通過するかについての制御、または少なくとも知識を得たいと考えている。本発明の実施形態は、この可視性を提供するために、2つの異なる手法を利用する。

【0133】

最初の方法は、“トレースルート”手法である。この手法は、Time To Live (TTL)を増分して、一連のパケットを送信する。各パケットが或るルータで所望の最大数に達する場合、それらのルータはパケットを送信元に返し、どのノードに到達したかを示す。

【0134】

例えば、最初のパケットは、TTL = 1を有する。一度、最初のルータが本パケットを受理し、1によるTTLを減少する。結果的な値がゼロになるため、ルータはそれ自身のアドレスまたは識別性を示すパケットを返す。

【0135】

トレースルート方法は、完全に正確というわけではない。第一に、ルータは回答する責任を持たず、このケースでは、これらはそれら自身の正しい情報を送信しないかもしれない。それらは、それらのISPを示す一般的な回答を送信するか、あるいは間違ったIPアドレスさえ提供するかもしれない。これは意図的に行うことができるが、一般的にはそうではない。トレースルートの異常性は、B. Augustinらによって、“パリのトレースルートでのトレースルートの異常回避”で分析し、説明されており、参照によって明細書に組み込まれる。トレースルートの異常は、一般にトポロジー自体またはルータのロードバランシングポリシーに関連する。

【0136】

また、トレースルート手法では、パケット交換ネットワークでは、必ずしも満たされないという仮定がある。1つのパケットは或るパスを使用できるが、次のパケットは異なるパスを使用するかもしれない。本ケースでは、インクリメンタルTTLは、同じパス内でルータの識別性を提供しない。この事実は、潜在的に不正確なリンク識別性を招くことも

10

20

30

40

50

ある。

【 0 1 3 7 】

一実施形態では、代替の T T L 値を有するプローブが、トレースしたい接続フローの実際のトラフィックとして偽装する、トレースルート技法を修正する代替方法を使用する。パケットは、接続フローパケットの 1 つと、ルータの応答を特定のプローブで照合するためにフィンガープリントされたランダムなペイロードに等しい、5 タブルを有するように偽造する。この修正したトレースルート手法は、プローブによる周知またはランダムなポートを使用せずに、実際のトラフィックフローをシミュレートするので、従来のトレースルート手法よりも信頼性がある。

【 0 1 3 8 】

フロースティッキ性の影響（ロードバランスをまたぐ場合でも一般的に保証されている）と共に、経時的なルート変更の検出を含むルートの発見という点で、再現性のある結果を達成できる。しかしながら、本手法は、全てのトレースルーティング手法に影響する欠点を有する：（ i ） N A T は、プライベートネットワークでの書き換え、（ i i ） T T L が満期になると、 I S P が応答しなくなることに対処することである。

【 0 1 3 9 】

修正したトレースルート手法に特有の特別な欠点は、5 タブルがそのポート上のリスニングサービス（例えば、O p e n V P N サービス）に転送される実際の有効なタブルであるので、ラストホップから何の回答も受け取らない可能性である。トンネルサービスは、内部プロトコルの要件に従わないため、パケットをドロップするのに十分スマートであるが：その特定のプローブの I C M P 時間超過メッセージは生成されない。本目的のため、ツールは、終了して接続が終了したと判断する前にしばらくの間 T T L 値を増分する応答を受信しなかった後も、プローブを送信し続ける。ルート発見の代替の手法は可能であり、いくつかの実施形態では、他のルート発見器を使用する。

【 0 1 4 0 】

トンネルに沿って通過したネットワークノードに関する情報を提供することができる手法を確立すると、ノードのセットを含む情報を維持することができ、各パスについて関心のある他のメトリックを収集することができる。ゆえに、大きな C S P プールを管理するのでなく、削減した C S P セットは利用できる。この最適化は、C S P 管理の簡略化による、オーバーヘッドの削減となる。

【 0 1 4 1 】

C S P ステータス

【 0 1 4 2 】

ネットワークの変動は C S P のパフォーマンスに影響を与え、エージェントが初期化されたときの静的設定が妨げられる。むしろ、C S P は、絶えず変化するインターネットの状況に適応する必要がある。この問題に対処するために、各 C S P は、C S P の役割およびそれらの測定基準（例えば、スループット、レイテンシなど）に基づいて異なる状況を反映する一組の可能な状態を有する。

【 0 1 4 3 】

最終ユーザにとってシステムは安定しているが、使用される C S P のセットは所望の Q o E を維持するために動的に変更される。図 1 0 は、C S P の可能な状態の例を示す。一実施形態は 5 つの状態を使用する：アクティブ（ 1 3 1 ）、スタンバイ（ 1 1 1 ）、ウェイティング（ 1 4 1 ）、降格（ 1 2 1 ）、プロービング（ 1 0 1 ）である。他の実施形態はこれらの状態を修正でき、例えば、ウェイティング状態を除く。図 1 0 に示される状態のダイアグラムは、これらの状態間の推移を表す。これらの状態は、下記で詳細に示される。

【 0 1 4 4 】

アクティブ：これは送信可能な C S P 、または現在トラフィックを送信している C S P 。

【 0 1 4 5 】

スタンバイ：このC S P状態は、アクティブに昇格した場合に、使用する資格があることを意味する。その昇格は、C S Pメトリクスに基づく（帯域幅、レイテンシなど）。昇格後、スタンバイC S Pはアクティブになる。品質が特定の基準に違反する場合、スタンバイC S Pも削除できる。

【0146】

ウェイティング：オープンのみで、パフォーマンスが一定の制限を超えた場合、またはより多くのリソースが必要な場合に、C S Pをスタンバイに昇格することができる。例えば、ウェイティングC S Pは、時々高いパフォーマンスを有し、他の周期で低いパフォーマンスを有することがある。

【0147】

降格：アクティブC S Pが低下した後、降格状態になる。この状態は、システムが、そのC S Pに、新しいフローまたはトラフィックを割り当てることができないことを意味する。さらに、降格C S Pを通した進行中のトラフィックを扱うための、異なる選択肢がある。例えば、積極的なオプションはC S Pをカットして、（送信されていたかバッファの上のパケットのために）T C Pにパケットの損失を処理させることからなるだろう。もう1つの選択肢は、パケット損失なしに、C S P間でソフトな推移を実装するだろう。この選択肢は、進行中のパケットを送付し、C S Pを完全に降格する。

【0148】

プロービング：プロービングC S Pは、有効なC S Pとして昇格したり排除したりする前の、（対象となるネットワークメトリクスに従った）分析中のトンネルである。

【0149】

削除したC S Pは、それらに関連するすべてのものが消去されるため、状態とは見なされない。図10の交差は、C S Pの排除を表す。

【0150】

異なる状態の推移を決定付けるコンディションは、スループット、レイテンシ、およびパケット損害を含む、ネットワークメトリクスに依存する。本発明のいくつかの実施形態では、それらの推移は静的で手動的である。代替実施形態では、より多くの測定基準を考慮して、プロセスは自動化される。

【0151】

エージェントは、C S P I Dを含む状態ごとにキューを管理する。エージェントが伝達する新C S Pを必要とする場合、アクティブキューにし、所望のメトリクを伴うC S P I Dを選択する。加えて、優先は、C S P選択プロセスを最適化する、これらのキューに適応できる。

【0152】

C S Pアサイメントのフロー

【0153】

ポリシーは、フロー由来のサブフローの最大数および使用可能なC S Pの数を特定する。それらのパラメータのトップではC S P割り当てのフローは、パフォーマンスを最適化するために実装する。

【0154】

一実施形態では、新しいフローがV Mに到達する場合、アクティブC S Pの数と同様のサブフローに分割する。ゆえに、1つのサブフローは、各アクティブC S Pを通して、送付する。このプロセスは静的で、新しいフローがエージェントに到着するたびに行われる。

【0155】

代替の実施形態では、トラフィックの特徴、およびネットワークコンディションに基づいて分割した、動的フロー/サブフローを利用する。動的な解決策は、分割した1つのフローおよびC S P割り当てはある瞬間に良いパフォーマンスを有するが、もう1つの時間に乏しいパフォーマンスを有するので、利点を有する。いくつかの実施形態では、サブフローの数は変動可能であり、C S P割り当ては動的に変化できる。C S Pは同様に動的で

10

20

30

40

50

あり、動的な割り当ては、アプリケーションタイプ、時刻、ユーザのタイプなどに依存できる

【0156】

アーキテクチャ

【0157】

異なるエンドポイント間で多数のCSPを管理し、CSPの割り当てへのフローをネットワーク状態に適応させ、エンドツーエンドのQoEを保証することは、スケーラビリティと複雑性の観点から重要な作業である。この管理は、とりわけネットワーク状態（例えば、輻輳、失敗など）、クロストラフィック干渉、および公衆インターネットの確率的性質への対処などのリアルタイム情報に基づいて、行動することを含む。これらの問題を取り組むために、階層アーキテクチャを定義した。図11は、このアーキテクチャを示す。一実施形態は3層の階層システムを利用し、最上層は、グローバル層（151）、中間層は、中間層（161）、および最下層はデータ層（171）である。アーキテクチャの下部には、情報、測定値およびパケットを管理するデータラベルがある。上記のデータレベルは中間レベルであり、異なるサブレベルを集計の複数のレベルに分解できる。これらの範囲は、エンタープライズ要件（例、部署、ユーザポリシーなどおよびネットワークポリシー（例えば、IPS、ポートの数など）を含むローカルコンテスト情報（LCI）を包含する。このレベルは、LCIに基づくCSP割り当てのフローを含むポリシーを扱う。アーキテクチャのトップには、社会政治的状況、壊滅的な災害、スポーツゲームなどの世界規模のイベントを扱う、グローバルレベルがある。このグローバルコンテスト情報（GCI）は、より低いレベルをガイドするポリシーを発生するために、使用する。インターフェイスは通信を扱い、これらのレベル間の情報を交換する重要な特性は、半独立した運用体制である。データ層が上位層から切断した場合でも、利用可能な情報に基づいて動作を継続できる。

【0158】

本発明の実施形態は、それらの異なる文脈情報を利用して、それらの機能性を最適化するために各層に、分散学習アルゴリズムを有する。各層が異なる種類の情報を扱うという事実は、公衆インターネット上でQoEを保証するための、スケーラブルな解決策の提供に貢献する。

【0159】

レベルが高いほど、システムの抽象度と範囲は高くなるが、粒度は粗くなる。それらのレベルは、異なるタイムスケールも反映する。データレベルはリアルタイムで作動できるが、中間およびグローバルレベルは作動しない。これらのさまざまな運用体制は、エンドポイントで作動するエージェント（データレベル）と、オンプレミスまたはクラウドで実行する、中間レベルおよびグローバルレベルの分散型アーキテクチャに起因する。

【0160】

単一エンタープライズの視点

【0161】

エンタープライズが本発明の実施形態を使用するエンドツーエンドQoEを保証したい場合、それらのオフィスでエージェントを展開することは最初のアクションとなる。それらのエージェントの組み合わせは、データ層の実装となる。各エージェントは、CSPの管理に責任的であり、フローをポリシーのセットに基づいて割り当てる。

【0162】

それらのポリシーは、LCIに基づいて発生させる中間層のアウトプットである。同様に、中間層は、GCIを考慮するグローバル層よりポリシーを受理する。単一のエンタープライズのケースでは、中間層とグローバル層は、各QoEを確保するのに貢献する。図12は、この状態を示す。エンタープライズ・エンドポイントでデータ層（221、261、251）を扱うエージェントは、ISP（231）を通して、公衆インターネット（241）に接続する。中間層（211）およびグローバル層（201）は、公衆インターネット（241）並びに相互の通信およびデータ層（221、261、251）にも接

続する。

【 0 1 6 3 】

グローバルな視点

【 0 1 6 4 】

データ層に関する本発明の実施形態は、中間レベル内のサブ層の一部と共に、各クライアントについて複製される。グローバルレベルは、共同分析およびトラフィックを開発する、全て顧客間で共有する。情報を集約するにも関わらず、各エンタープライズのドメインは、データや通信のセキュリティおよびプライバシーを確定するために分離する。

【 0 1 6 5 】

アーキテクチャのグローバルな視点は、図 1 3 に示される。1 番目のエンタープライズ “企業 A” (3 2 1、3 3 1、3 7 1)、および 2 番目のエンタープライズ “企業 B” (3 4 1、3 6 1) を扱うエージェントは、ISP (3 5 1) を通して 公衆インターネット (3 9 1) に接続する。“企業 A” (3 1 1) および “企業 B” (3 8 1) の中間層並びにグローバル層 (3 0 1) も、公衆インターネット (3 9 1) 並びにそれぞれの通信および データ層 (3 2 1、3 3 1、3 7 1、3 4 1、3 6 1) に結合する。

10

【 0 1 6 6 】

本ケースでは、2 つのエンタープライズがある。各オフィスは、他のオフィスの接続を確立し扱うエージェントを有する。これらのエージェントは、それぞれの中間層からポリシーやコマンドを受理し、クラウド上で作動する。同時に、中間層ソフトウェアは G C I を扱うグローバル層よりポリシーを受理する。グローバル層は、世界中で分配されるクラウド上で作動することもある。

20

【 0 1 6 7 】

本実例では、同じエンタープライズ内でのオフィスのみが、これらの間で接続できる。代替実施形態では、異なるエンタープライズよりオフィスを相互接続するより複雑な解決策を実装する。本トポロジーは、ポリシーの上昇した複雑性になり得る。中間層内の異なるサブ層は、特定のシナリオに適応するポリシーを発生するこの複雑性を扱うために、構成できる。

【 0 1 6 8 】

代替実施形態では、企業間接続を扱うより高いサブ層を、各エンタープライズの特定のミッドレベルサブ層のトップで使用する。この構成は、図 1 4 A および 1 4 B で示される。企業 A (4 2 1、4 3 1、4 7 1)、および企業 B (4 4 1、4 6 1) でデータ層を扱うエージェントは、ISP (4 5 1) を通して、公衆インターネット (4 9 1) に結合する。企業 A (4 1 1) および企業 B (4 8 1) の中間層に加えて、企業間エンタープライズ中間層 (4 8 3) と同様にグローバル層 (4 0 1) も、公衆インターネット (4 9 1) に接続し、互いにデータ層 (4 2 1、4 3 1、3 7 1、4 4 1、4 6 1) と通じる。概念的に、3 つの中間層は、企業 A 中間層 (4 1 3) および 企業 B 中間層 (4 2 3) のトップで、A - B 企業間層 (4 0 3) として見ることができる。

30

【 0 1 6 9 】

中間層内の異なるサブ層を創造する 1 つの利点は、エンタープライズ間でのローカルな意識と、連携特徴を提供することである。異なるエンタープライズからの L C I の組み合わせでデザインしたポリシーは、促進したパフォーマンスの結果となることができる。階層化したアーキテクチャは、エンドユーザーからは完全に透過的でありながら、複雑さに対処する。

40

【 0 1 7 0 】

データ層

【 0 1 7 1 】

データ層は、パケットと、フローとを扱い、ユーザデータにアクセスして作用する。この層は 2 つの主要要素、データプレーンおよびコントロールプレーンを有する。データプレーンは、一連のスイッチングおよびルーティングポリシーに基づき、パケットを転送する役割を担う。このプレーンは、フローおよび対応するサブフローを扱う。また、データ

50

プレーンは、トラフィックシェーピングおよび優先のための場所である。対照的に、コントロールプレーンは、C S Pの管理、C S P割り当てへのフロー、サブフローポリシーの適用、および学習アルゴリズムに重点を置く。

【0172】

主なタスクは、C S Pメトリクスの測定および収集である。本発明の実施形態は、帯域幅、パケット損失、およびレイテンシを含む上述したメトリクスのセットを収集する。これらの分析の目的は、システムのパフォーマンスをリアルタイムで監視し、Q o Eを維持し改善するための決定を下すことにある。

【0173】

重要な機能は、いくつかのネットワークメトリクスおよび主なパフォーマンス・インディケ이터（K P I）の抽出、変換、およびロード（E T L）である。発明の一実施形態では、このデータはすべてキュレーションされ、その粒状性に基づいて、中間層に送信される。理想的には、厳密に必要な情報だけが、1つの層から他の層に行く。

10

【0174】

これらの層間通信の頻度は、周期的分布に従うことができる。例えば、データのセットの各秒は、データ層から中間層へ送ることができる。この間隔の値は、応答システムを有効にし、そのスケーラビリティを危険にさらさないために、非常に重要となる。間隔を短くすると、処理と伝達の両方で莫大なオーバーヘッドのコストがかかるが、より細かい粒度が得られる。長い間隔はそのコストを避けるが、反応時間を短縮する。ゆえに、精度とリソース使用量との間には、システムの粒度をコントロールするためのトレードオフがある。本実施形態では、データ層から中間層への通信の間隔は、およそ1秒である。代替の実施形態では、他の間隔が実装される。

20

【0175】

一実施形態では、Z e r o M Qは、高性能非同期メッセージングライブラリであり、メッセージパックは、コンピュータデータ交換フォーマットであり、この層間通信を可能にする。加えて、データプレーンの実施形態は、メッセージキューテレメトリトランスポート（M Q T T）をサポートし、これは、T C P / I Pのトップで使用するためのI S O規格のパブリッシュ・サブスクライブベースの軽量メッセージングプロトコルである。

【0176】

例外（イベントベース）メカニズムは、反応時間を改善するため利用される。データ層がC S Pの突然の変化を検出すると、インターバルベースの通信の外側で、中間層に例外を送信する。本例外は、このようなバリエーションをどのように反応するかを決める中間層での、異なるイベントの引き金となるだろう。この手法は、システムのスケーラビリティを改善し、同時に、わずかなオーバーヘッドで予期しない変更を管理するための反応時間を改善する。この機能は、エンドユーザーとアプリケーションに対する透過性を強化する。

30

【0177】

トンネルとC S Pの監視に加えて、データ層は、異なるC S Pにフローを分散するために、必要なタスクの実行も担当する。この機能には、パケットのカプセル化、オプションの暗号化、およびマルチパス・トランスポートを有効にするためのフローに対する、f o r k / j o i n操作の実行が含まれる。

40

【0178】

加えて、データ層は、利用可能な帯域幅、パケット損失、およびレイテンシのような物理的または仮想的なインターフェイスの特徴を測定する。並行して、データ層は、トンネルおよびC S P上の基本的な操作を実行する。例えば、データ層は、C S Pの状態とは独立に、C S P状態をオープンに維持するキープアライブメッセージを担当する。同じ原則を、トンネルにも適用する。トラフィックがそれらに送信されない場合、トンネルはC S Pになるか、またはトンネルは破棄される。

【0179】

上述は、データ層の異なる機能性および中間層に送るものに関係する。逆方向では、中

50

間層は、データ層を強化するためのポリシーを送信する。ポリシーは、情報の状態からアクションへのマッピングである。システムの状態を考慮して、ポリシーは、どのCSPを使用してどのタイプのトラフィックを送信できるか、またトラフィックを他の機能とともにサブフローに分割する方法を決定付ける。ポリシーはまた、異なるCSP状態間の遷移（例えば、CSPが、アクティブからスタンバイへ移行すること）と共に、CSPへのトンネルの昇格／降格の引き金となる。ポリシーは、データの取得方法など、異なるイベントもコントロールする。

【0180】

本発明の実施形態では、CSPは、フォワードリターンペアになる。往路と復路は正確に合致する必要はないが、それらは独立して選択されない。ゆえに、CSPの選択は、ヘッドポイント（送信元）と、テールポイント（送信先）との共同責任であるべきである。重要な情報は、依然としてローカルなものであるが（ほとんどの場合、入力リンクと、出力リンクでのアクセス輻輳に関連する）、これは2つのエンドポイント（または、この通信が他の送信先を含む場合は、それ以上）ではローカルである。これは、CSPの選択が、中間層で行われる理由を正当化する。両方のケースでは、中間層はブローカーとして機能し、さまざまなエージェントに影響を与える決定を下す。

【0181】

ポリシーに加え、中間層は、データ層が送信した分析に基づいて、または特定のメトリックを要求するために、CSPへの特定のトンネルを促進するコマンドを送信することもできる。コマンドを受信すると、データ層は、所望のアクションを行う。もう一つのコマンド例は、トンネルのブループである。ゆえに、データ層は、その状態に関係なくすべてのCSPを観測するが、それらに基づいて行動するには、中間層からのポリシーが必要である。この場合のコントロールは、スケーラビリティの問題により、データ層から消去される。

【0182】

ポリシーには、データ層が実行できる、異なる自由度を有することができる。例えば、ポリシーは、10Mbps以上のスループットを伴うCSPのみが利用できることを示すことがある。この状態では、データ層は、この条件を満たすCSPのうちどれを使用するかを決定できる。いくつかの実施形態では、条件付きの状態に基づく制御アルゴリズムとして機能する厳密なポリシーを利用する。もう1つの実施形態では、より高い柔軟性は、中間層のポリシーによって可能となる。インテリジェントアルゴリズムは、パフォーマンスを最適化するために、これらの自由度を利用することができる。例えば、機械学習アルゴリズムを使用して、中間層ポリシーを維持しながら、どのCSPを介してデータが送信されるかを決定するために、リアクティブ情報（すなわち、各CSPでのテスト）を利用することができる。

【0183】

本発明の一実施形態では、データ層のすべての機能は、各エンドポイントにあるエージェントの仮想マシン（VM）上で実行する。1つのVMは、1つのエンドポイントの物理的なポートのセットを管理できる。データ層は、アーキテクチャ全体の基盤を構成するため重要である。マルチパスの最適化およびアルゴリズムの学習は、データ層の上位に構築される。

【0184】

要約すると、好ましい実施形態では、データ層は、以下の機能を行う：1．CSP上のメトリクスおよび間隔を基礎としたトンネルの測定；2．キュレート分析；3．CSPライブの維持；4．中間層ポリシーおよびコマンドの実行；5．リアルタイムのコントロールオペレーション；および6．アプリケーションの解読および分類（ポリシーによりコントロール可能である）。

【0185】

好ましい実施形態では、データ層は、以下の入力を受取る：1．中間層ポリシー；および2．トンネル、CPS、およびフロー管理のコマンド；並びに以下の出力：1．RT

10

20

30

40

50

T、片道遅延、スループット、容量、およびトレースルート、などのキュレート分析（CPSおよびトンネル）；並びに2．極端な状況の場合に例外を処理する（例、利用可能なBWなどの或る容量の突然の紛失）。

【0186】

データ層のいくつかの実施形態は、以下のモジュールを構成する：1．サブフローマネージャー；2．アプリケーションの分類；3．ネットワークコントローラー；および4．レポートエージェント。代替の実施形態では、データ層はまた、中間層ポリシーによって許容する自由度を行うための機械学習アルゴリズムを実装し、そして新しいネットワークメトリクスおよび互いの関係を発見する。

【0187】

中間層

【0188】

中間層は、各企業それぞれのシステムを理解しコントロールする層である。本発明の実施形態では、本情報は、ISP、物理的なインターフェイス、IPアドレス、転送ポート、トポロジーなどの詳細を含む。この情報は、ローカルコンテキスト情報（LCI）と呼ばれるものを含む。加えて、LCIは、異なるアプリケーション間の優先順位、新しいポリシーを導くパターンを創造するためのアプリケーションおよびフローのカテゴリ、とりわけクロストラフィックに関する情報を含む。ゆえに、この層の範囲は、各エンタープライズの内側のシステムの写真を創造する。

【0189】

中間層はまた、トンネル識別子を通過する異なるトンネルを探索するように、データ層に指示する。データ層がネットワーク測定値を送り返すと、中間層は、ポリシーまたはコマンドを通して、それらをCSPに昇格するか降格するかを決定できる。ゆえに、それらの順番は、所望のポリシーまたはコマンドを行うデータ層へ、送り返される。他のコマンドは、利用可能な帯域幅、レイテンシCSP状態間の推移の測定を含むことができ、データ層でETLを行う頻度を修正でき、ローカルコンテキストの意識（エンタープライズ・レベル）を利用するポリシーを定義できる。

【0190】

加えて、中間層は、入力として、キュレーション分析と組み合わせてポリシーを誘導する入力アプリケーションおよびユーザの要件を有する。中間層は、論理層であるため、フローまたはアプリケーションデータを見ない。いくつかの実施形態では、ネットワーク構成（転送ポート、IPアドレス、インターフェイスなど）は、各エンタープライズのIT担当者によるGUIを通して、この層に入り込む。代替の実施形態では、この情報は、自己発見プロセスの結果である。

【0191】

上位層への出力として、中間層は、（すべての中間レベルのサブ層からの）予期しない状況および接続の状態による、パフォーマンス上のボトルネックについての情報をグローバル層に送信する。より良いスケーラビリティを有するめに抽象化を得ながら、粒度を減らすために、より高いレベルに進む前に、ステータステストおよびデータ分析が再度キュレーションされる。逆方向では、中間層は、グローバル層からのポリシーおよびコマンドを受理する。これらの規則はより高い層から来るので、それらはより抽象的であり：政治状況を避けるために、或る国を横断するネットワークパスを使用しない。データ層、中間層およびグローバル層間の関係は、軍事組織に類似化することができる。兵士（データ層）は、キャプテン（中間層）から注文を受け、同時に将軍（グローバル層）からの注文に従う。ポリシーを作成する層が高いほど、その細分性は低くなるが、低いレベルで保証する優先順位は高くなる。

【0192】

機械学習アルゴリズムは、ローカルのコンテキスト情報（LCI）を活用して、解決策のパフォーマンスを向上させる新しいポリシーを動的に調整または作成できる。グローバル層ポリシーによって残された自由度は、中間層の改善分野を決定する。これは、学習ア

10

20

30

40

50

ーキテクチャも各層で利用可能な情報に従い、階層構造に従うことを明らかにする。

【0193】

好ましい実施形態では、中間層は、エージェントではなく、クラウド上で作動する。異なるサブ層は、中間層を形成でき、階層方法でも整理される。サブ層の数は、扱うコンテキスト情報、実装するインテリジェンスの量、とりわけエンタープライズの関係によるであろう。

【0194】

サマリーでは、好ましい実施形態では、中間層は下記の機能を実行する：1．システムのビュー（全社的）の処理；2．CSPおよびトンネルの管理；3．データ層へのポリシーおよびコマンドの送信；4．グローバルデータ層からのポリシーおよびコマンドの実行；並びに5．データプレーンおよびより低い中間層からの測定を収集し、処理する。

【0195】

好ましい実施形態では、中間層は、下記の入力を有する：1．グローバル層からのポリシー；2．データ層からのキュレートデータ分析；並びに下記の出力を有する：1．グローバル層へのキュレートデータ分析；2．データ層へのポリシーおよびコマンド3．LCI；および4．GUIユーザデータ。

【0196】

代替の実施形態では、中間層は、トラフィックパターンのポリシーおよび分類を動的に創造する、機械学習アルゴリズムを利用する。

【0197】

グローバル層

【0198】

グローバル層は、すべてのエンタープライズ、グローバル公衆インターネット、および社会政治的出来事、ニュースなどの外部要因を含むシステムの全体像を有する。ゆえに、このレベル内のコンテキスト情報は、中間層に関して単一エンタープライズの焦点を超えているため、グローバルコンテキスト情報（GCI）である。

【0199】

グローバル層は、中間層より受理したキュレート分析に従い、そのポリシーと決定をも誘導する。これらのポリシーは、そのルールとその自由度の制約により、下位層で最適化の余地を残す。グローバルポリシーの例は、或る国のリンクを使用しない、スポーツイベントによる混雑を避けるか、もしくは休日または夜間による低い利用を有する地域を通すCSPを利用するであろう。

【0200】

要約すると、好ましい実施形態では、グローバル層は以下の入力を利用する：1．中間層からのキュレート分析；2．社会政治的出来事に関する外部情報；および3．情報のクロール、並びに以下の出力を有する：1．中間層へのポリシーおよびコマンド

【0201】

代替の実施形態では、グローバル層は、動的なポリシーを創造する機械学習に適応し、世界的な出来事およびネットワークメトリクス間の関係を確立する。

【0202】

学習

【0203】

異なるレベルで構成したシステムアーキテクチャを、上述した。また、いくつかの実施形態では、各レベルにおいて、利用可能な情報を活用して、より良い決定を下し、システム全体のパフォーマンスを最適化する学習アルゴリズムを採用するという事実も上述した。以下に、各レベルでの学習の下に、その主なタスク、必要な入力および出力、およびそれらの最終目的を含めて説明する。

【0204】

本実施形態の学習ソリューションは、3つの異なる層に適用される：(i) データ層、(ii) 中間層、および(iii) グローバル層。これらの3つの層は、学習の3つの異なる

10

20

30

40

50

ステージを翻訳する。データ層から始めて、伝達するC S Pとパケットに関する情報に基づく、反応フェーズがある。中間層では、エンタープライズ内の交差トラフィックなどのLCIによるローカルフェーズがある。最後に、グローバル層では、より高い抽象化（例、一国からのトラフィック）を伴うトラフィックフローを扱う、グローバルフェーズがある。

【0205】

より高い層は、より大きなレイテンシを犠牲にしてより良い可視性を持ち、それらのリアルタイム能力を低下させる。図15は、アーキテクチャのフレームワーク内のこれらの異なるフェーズを示す。学習モジュール(531)を有するグローバル層(501)は、学習モジュール(541、551)を有する中間層(511)に結合され、学習モジュール(561)を有するデータ層(521)に結合する。

10

【0206】

これら3つのレベルで学習を分解することの利点は、異なる時間および抽象化の体制を通して、インターネット規模の解決策を提供することにある。本発明の実施形態は、最下層のエージェントから中間層およびグローバル層のクラウドまで、各アーキテクチャ層における、異なるインフラストラクチャ機能から利益を得る。ゆえに、システムがリアルタイムの制約に直面する場合、データ層でワークロードを実行しながら、容量の問題が少ないクラウドで、大量の重要でないデータを処理できる。

【0207】

本発明の実施形態のマルチレベル学習アーキテクチャは、階層学習(HL)として示され、さらに図16に示される。グローバル層(601)は、中間層(611)に結合され、中間層(611)はデータ層(621)に結合される。

20

【0208】

HLアーキテクチャの実施形態は、各レベルで、異なる機械学習手法を使用する。例えば、1つのレベルは再帰的ニューラルネットワーク(RNN)であり、次のレベルはディープラーニング(DL)を実装できる。各レベルは、下位レベルの出力が上位レベルの入力の1つになるという点で、互いに接続される。加えて、各層は、層の相互接続を補完する独自のデータセットを有する。

【0209】

例えば、データ層は、入力としての反応的な情報および中間層からの入力を有する（例えば、ポリシー）。結果として、学習システムの適切な機能を保証するために決定を下す可能性があるため、各レベルは独立する。例えば、データ層およびその反応的な学習アルゴリズムがクラウドより切断されたケースでは、データ層内のローカル学習は、層が利用可能であるという情報に従い、フローを最良のC S Pに割り当てことを決断できる。このケースでは、情報の深度はクラウドと同様ではないが、システムは働き続ける。その後、各層が接続され適切に機能している場合、各レベルからの抽象化およびそれらが持つ異なるシステムビジョンを活用することができる。

30

【0210】

インフラストラクチャ機能は、HLの重要なパラメータになる。エージェントは機能の異なる異種端末でホストされるため、各端末で実行する学習アルゴリズムは実行プラットフォームによって大きく影響する。例えば、いくつかの実施形態では、VMはハイエンドサーバーで展開するが、他のケースでは、携帯電話などのリソースに制約のある端末で実行される最適化されたエージェントを検討する。

40

【0211】

異種端末で実行する際の問題を避けるために、インフラストラクチャの能力は、各学習レベルでの入力として用いられる。そして、能力は、あるレベルから別のレベルへのフィードバックを通して、MLベースの決定に影響を与える。この設計の特徴により、基盤となるインフラストラクチャに応じて、階層学習の自律的最適化が行われる。データ層のノードがより多くの計算を引き受けることができる場合、それはこの事実を中間層に伝達する。中間層はかくして、そのポリシーにさらに自由度を残す。逆のケースでは、エージェ

50

ントはその限られたリソースを通信することができるので、ポリシーはより制約され、したがってより少ない計算リソースを必要とする。

【0212】

各HLレベルで使用するML手法は、上位層からのポリシーによって課される制約に基づいて決定する。それらのポリシーは、システムを管理する各レベルで利用可能な情報に従って、動的に定義される。例えば、それらは利用可能な異なるCSPへのフロー割り当てを指定する。それぞれのポリシーは異なる自由度を残しており、それらのポリシーに違反することなく、機械学習手法を適用するために、下のレベルで行われる。

【0213】

いくつかの実施形態は、(ニューラルネットワークにおける重み、またはベイズネットにおける確率表などの)機械学習パラメータが、層間で交換することを可能にし、すなわち、すでに訓練した機械学習エンジンを複製することによって得られた知識が交換される。このアプローチにより、各層がポリシーとコントロールコマンドを自身で取得できるようになり、クラウドから切断された場合でも、層の自律性が保証される。このアプローチはまた、より迅速な意思決定を意味する。言い換えれば、最下層は上層の可視性と知識を獲得し、逆もまた同様である。各層は、機械学習パラメータを更新するように努める。

【0214】

例えば、中間層はデータ層にポリシーを送信して、10Mbpsを超えるスループットを伴うCSPのみを使用できるように指定する。反応的な情報に従って、データ層MLアルゴリズムは、ポリシーおよびそれが有する自由度に基づく、パフォーマンスを最適化できる。このケースでは、データ層は、その条件を満たすすべてのアクティブCSPから選択できる。もう1つのより制限的な状況は、特定のアクティブCSPのみを使用することである。ここでは、ポリシーによって利用可能な動作が制約しているため、データ層ではほとんど最適化できない。

【0215】

これらの自由度は、システム変数(エージェントの機能、コンテキスト情報、ネットワーク設定、トラフィック干渉など)を考慮して、各階層レベルで動的に最適化される。HLアーキテクチャは、概念的にはNN上の異なる反復を連結する、リカレントニューラルネットワーク(RNN)と類似する。さらに、NNを常に同様にするのではなく、それらの間のフィードバックを維持しながら、異なるML手法が各レベルの新しい入力と連結する。

【0216】

本発明の実施形態では、機械学習は、異なる分野で適用する。HLの内側のアルゴリズムは、分野の広範囲に適用できる、例えば：1.トンネル発見(IPアドレス、ディスカバリ時間などに基づいて通過したネットワークノードとのパス情報)；2.CSP昇格および降格；3.CSP状態推移；4.SLA、QoEなどのアプリケーションタイプに従ったCSP分類；5.フローおよびトラフィック分類；6.CSPへのフロー割り当て(コンテキスト情報、インフラストラクチャ機能、ネットワーク条件などに基づくフロー分割(サブフロー数による))7.履歴や出来事などに従ったネットワーク状態の予想；並びに8.ポリシー生成(中間およびグローバル層内)。これらの分野は、より高いレベルからのポリシーによって課される各建築レベルにおける自由度により調整される。

【0217】

ML手法を適用する主なエリアは、利用可能なCSPへのフロー割り当てである。効果的な解決策を提供するため、動的である必要があり、ネットワーク変動に適應する。考慮すべきデータ量は膨大であり、時間とともに大きく変動し、アーキテクチャにとって大きな課題となる。これは、より静的なCSP/トンネル検出とは対照的である。

【0218】

CSP割り当てに最適化されたフローを提供するために、異なるML技術をLアーキテクチャで実行することが考えられる。いくつかの実施形態では、データ層では、リアルタイムの反応的情報を扱う、リカレントニューラルネットワーク(RNN)を使用する。中

10

20

30

40

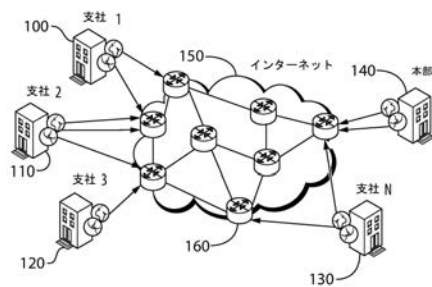
50

間およびグローバル層では、ディープラーニング手法の利点は、それぞれローカルおよびグローバルのコンテキスト情報をカバーするために動作する。

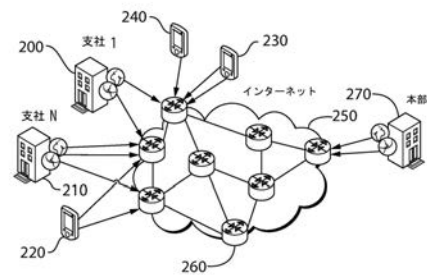
【 0 2 1 9 】

本発明は、いくつかの好ましい実施形態について上述された。これは説明の目的のみになされたものであり、そして本発明のバリエーションは、当業者には容易に明らかであり、そしてまた本発明の範囲内に入るものである。

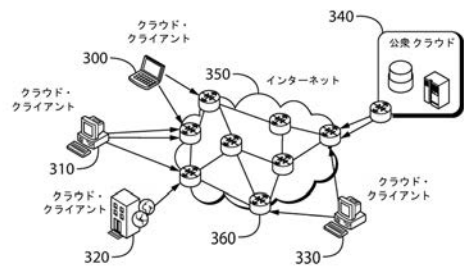
【 図 1 】



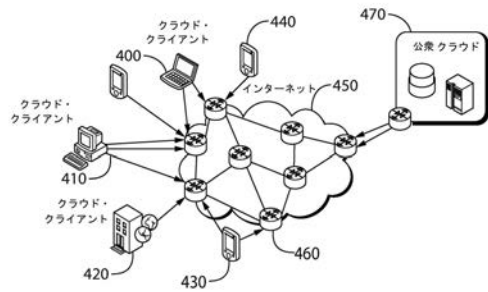
【 図 2 】



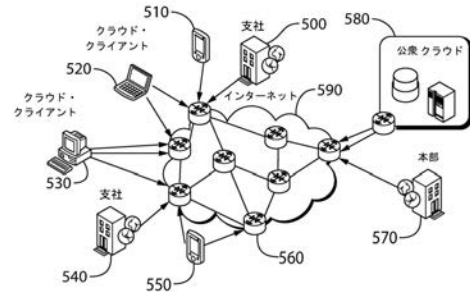
【 図 3 】



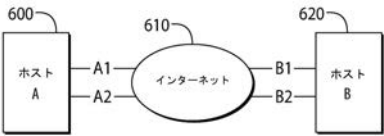
【 図 4 】



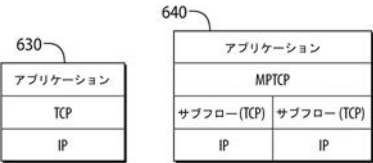
【 図 5 】



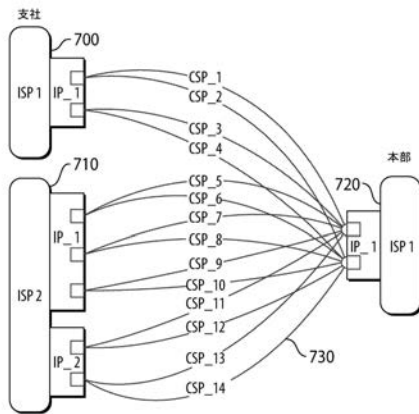
【 図 6 A 】



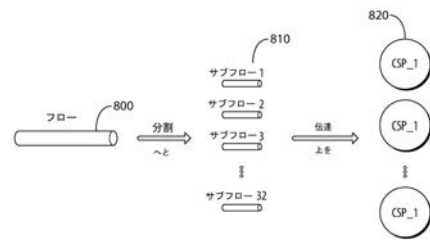
【 図 6 B 】



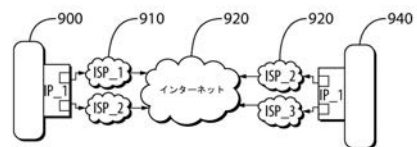
【図 7】



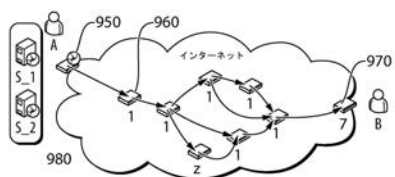
【図 8】



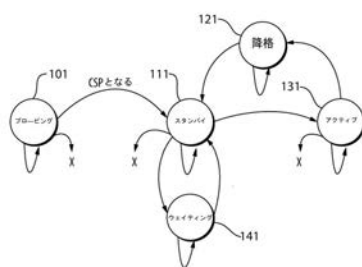
【図 9 A】



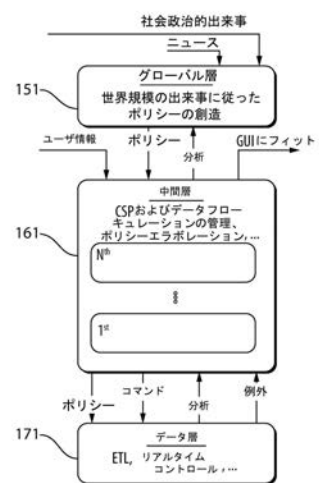
【図 9 B】



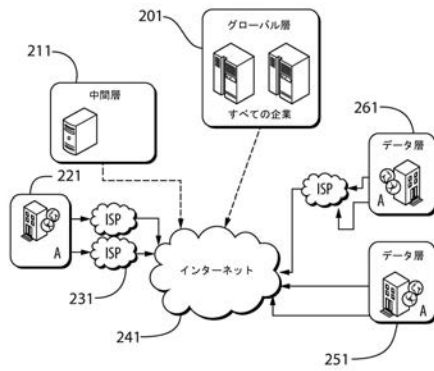
【図 10】



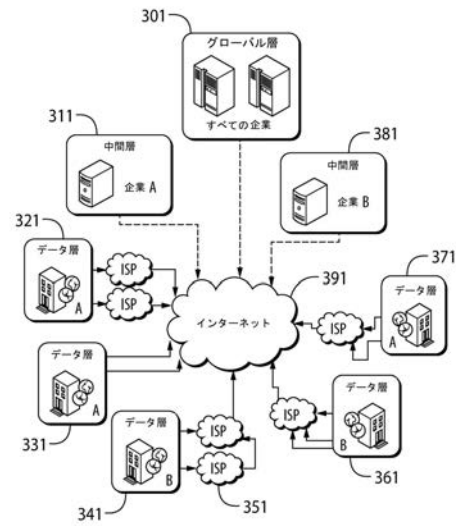
【図 11】



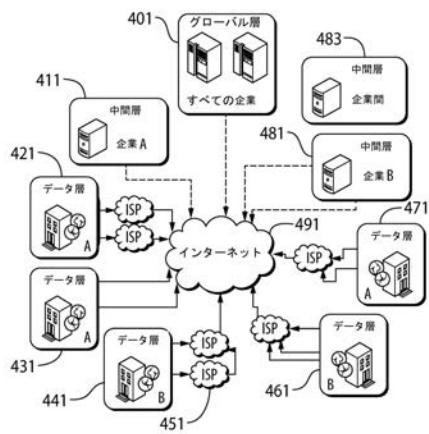
【図 1 2】



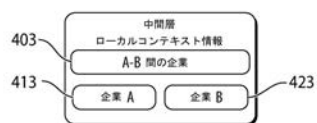
【図 1 3】



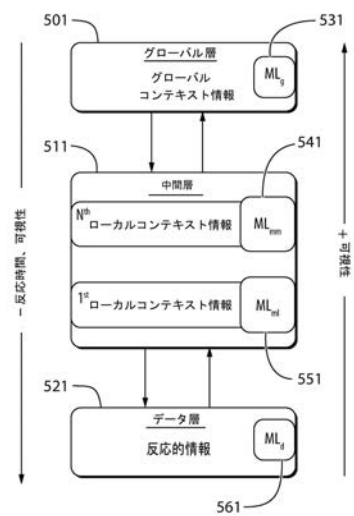
【図 1 4 A】



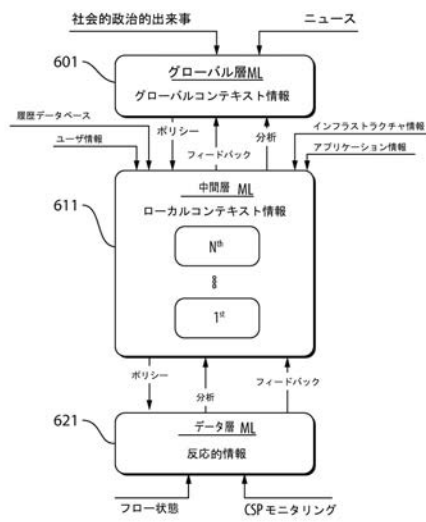
【図 1 4 B】



【図 1 5】



【図 16】



【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US17/38053

A. CLASSIFICATION OF SUBJECT MATTER IPC - H04W 76/02, H04W 76/04, H04L 29/02 (2017.01) CPC - H04L41/08, H04W76/022, H04L67/148		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) See Search History document		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched See Search History document		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) See Search History document		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 2014/0310417 A1 (AMAZON TECHNOLOGIES, INC.) 16 October 2014; figures 15, 23; paragraphs [0043]-[0050], [0063], [0198], [0239], [0254], [0255], [0282], [0297]	1-20, 22-34 — 21
Y	US 6,816,905 B1 (SHEETS, K et al.) 09 November 2004; figure 12; column 19, lines 1-32	21
A	EP 0,912,015 A2 (LUCENT TECHNOLOGIES INC.) 28 April 1999; entire document	1-34
A	US 2014/0098685 A1 (SHATTIL, S) 10 April 2014; entire document	1-34
A	US 2011/0047291 A1 (ISHII, T et al.) 24 February 2011; entire document	1-34
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "Z" document member of the same patent family		
Date of the actual completion of the international search 25 September 2017 (25.09.2017)		Date of mailing of the international search report 18 OCT 2017
Name and mailing address of the ISA/ Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300		Authorized officer Shane Thomas PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US17/38053

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Group I: Claims 1-17; Group II: Claims 18-34

-***-Continued in extra sheet-***-

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US17/38053

---Continued from Box No. III - Observations where unity of invention is lacking---

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fee must be paid.

Group I: Claims 1-17 are directed towards repeatedly evaluating a connection to determine at least one performance metric and associating a state for said connection.

Group II: Claims 18-34 are directed towards breaking the information flow into sub-flows based on one or more of performance requirements and application characteristics.

The inventions listed as Groups I-II do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The special technical features of Group I include at least maintaining a plurality of connections between a first network host and a second network host by performing, for each of said plurality of connections, the sub-steps of: establishing a connection between the first network host and the second network host; repeatedly evaluating said connection to determine at least one performance metric associated with said connection; associating a state for said connection, based on the at least one performance metric associated with said connection, from among a plurality of states, wherein one or more first states from among said plurality of states represents a ready condition and one or more second states from among said plurality of states represents a not ready condition, which are not present in Group II.

The special technical features of Group II include at least wherein said request is associated with performance requirements and application characteristics; breaking the information flow into a plurality of sub-flows, each comprising a sequence of packets, wherein the step of breaking is based on one or more of the performance requirements and application characteristics; assigning each of the plurality of sub-flows to a connection from among a plurality of connections between said first network host and said second network host; communicating data from said plurality of sub-flows over one or more of said plurality of connections based on said step of assigning, which are not present in Group I.

The common technical features shared by Groups I-II are a method comprising, and apparatus for communicating information flows over a plurality of connections between network hosts over a network coupling the network hosts, the apparatus comprising: a first network host comprising at least one processor in communication with least one memory storing processor readable instructions, wherein the at least one processor is operably configured by the processor readable instructions to: receive a request to transfer an information flow between said first network host and said second network host.

However, these common features are previously disclosed by US 5,058,110 to BEACH et al. (hereinafter "Beach"). Beach discloses a method comprising, and apparatus for communicating information flows over a plurality of connections between network hosts over a network coupling the network hosts (apparatus and method for communication (information flows) between a plurality of computing devices such as host computers in a network using hubs networked together (connections between network hosts over a network); column 2, lines 59-66), the apparatus comprising: a first network host comprising at least one processor in communication with least one memory storing processor readable instructions, wherein the at least one processor is operably configured by the processor readable instructions (a computing device (first network host) is coupled with a protocol processor 305 co-located with logic 304 and 306 (processor readable instructions) in a module implemented in hardware or software (memory storing processor readable instructions); figure 3A; column 3, lines 3-12; column 6, line 56 to column 7, line 2) to: receive a request to transfer an information flow between said first network host and said second network host (data arrives with header information that indicates the request is for communication with another host (information flow between said first network host and said second network host); figure 7(a); column 7, lines 35-39).

Since the common technical features are previously disclosed by the Beach reference, these common features are not special and so Groups I-II lack unity.

フロントページの続き

(81)指定国・地域 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT

(特許庁注：以下のものは登録商標)

1. L i n u x

(72)発明者 ネミロフスキー, マリオ
アメリカ合衆国 カリフォルニア州 95030、ロス ガトス、 サウス サンタ クルーズ
アベニュー 20、ナンバー105

(72)発明者 セラル - グラシア, ルネ
アメリカ合衆国 カリフォルニア州 95030、ロス ガトス、 サウス サンタ クルーズ
アベニュー 20、ナンバー105

(72)発明者 チアチャ, フランシスコ
アメリカ合衆国 カリフォルニア州 95030、ロス ガトス、 サウス サンタ クルーズ
アベニュー 20、ナンバー105

(72)発明者 ロメロ ルイス, アイヴァン
アメリカ合衆国 カリフォルニア州 95030、ロス ガトス、 サウス サンタ クルーズ
アベニュー 20、ナンバー105

Fターム(参考) 5K030 GA14 GA19 LB06 LC01 MB06 MB09 MB13

【要約の続き】

衆インターネットなどの既存のネットワークがエンタープライズクラスの接続を提供でき、本インターネットインフラストラクチャを修正せずに、エンタープライズクラウドの採用を加速するために、利用できる。