

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4677155号
(P4677155)

(45) 発行日 平成23年4月27日 (2011. 4. 27)

(24) 登録日 平成23年2月4日 (2011. 2. 4)

(51) Int. Cl. F I
H O 4 L 12/56 (2006. 01) H O 4 L 12/56 I O O C

請求項の数 29 (全 19 頁)

(21) 出願番号	特願2001-504633 (P2001-504633)	(73) 特許権者	510252438
(86) (22) 出願日	平成12年6月16日 (2000. 6. 16)		レベル・スリー・シーディーエヌ・インターナショナル・インコーポレーテッド
(65) 公表番号	特表2003-502941 (P2003-502941A)		Level 3 CDN International, Incorporated
(43) 公表日	平成15年1月21日 (2003. 1. 21)		アメリカ合衆国 コロラド州80021
(86) 国際出願番号	PCT/US2000/016732		ブルームフィールド, エルドラド・ブルバード, 1025
(87) 国際公開番号	W02000/079730	(74) 代理人	110000028
(87) 国際公開日	平成12年12月28日 (2000. 12. 28)		特許業務法人明成国際特許事務所
審査請求日	平成19年6月5日 (2007. 6. 5)	(72) 発明者	リッチウリ・リビオ
(31) 優先権主張番号	09/336, 487		アメリカ合衆国 カリフォルニア州95032
(32) 優先日	平成11年6月18日 (1999. 6. 18)		ロス・ガトス, エル・ガトー・レーン, 15635
(33) 優先権主張国	米国 (US)		
前置審査			

最終頁に続く

(54) 【発明の名称】 コンピュータ通信ネットワークのためのオンデマンドなオーバーレイルーティング

(57) 【特許請求の範囲】

【請求項 1】

パケット交換型のコンピュータ通信ネットワーク内において、ソースから宛先へとメッセージを伝送するための最適化パスを決定するための方法であって、

a) 前記メッセージを伝送するリクエストに応じ、前記通信ネットワークの1つまたはそれ以上の既存のルーティングメカニズムによって得られるデフォルトパスに沿って、前記ソースから前記宛先へのコストを測定する工程と、

b) 前記コンピュータ通信ネットワーク上に仮想のトポロジーを規定する複数の中間ノードの内、前記デフォルトパス上ではない1つまたはそれ以上の中間ノードを通る複数の代替パスに沿って、前記ソースから前記宛先へと前記メッセージを伝送する際の代替コストをそれぞれ測定する工程であって、

各前記中間ノードについて、他の前記中間ノードと間接的に結ばれる間接リンクおよび直接的に結ばれる直接リンクについて通信コストを測定し、測定された通信コストに基づき、間接リンクより通信コストが高い直接リンクを削除する第1の工程と、

第1の工程後に、前記各中間ノードの直接リンクの数が所定数を上回る場合には、前記各中間ノードについて、前記測定された通信コストが最も高い直接リンクを削除する第2の工程と、

直接リンクが所定数以下の中間ノードを通る複数の代替パスを生成する工程とを含み、前記生成された複数の代替パスに対応する複数の前記代替コストを測定する工程と、

c) 前記デフォルトコストと複数の前記代替コストとを比較することによって、前記最

10

20

適化パスを決定する工程と
を備える方法。

【請求項 2】

請求項 1 記載の方法であって、

前記方法の各工程は、前記メッセージを伝送する前記リクエストに応じてリアルタイムで、かつユーザとの明示的な相互作用を要しない方法で実施される方法。

【請求項 3】

請求項 1 記載の方法であって、

前記メッセージは、前記宛先から前記ソースへと受信された前のリクエストに応じて動的に生成される方法。

10

【請求項 4】

請求項 1 記載の方法であって、

前記工程 (b) は、指定された閾値コストを前記デフォルトコストが上回る場合にのみ実施される方法。

【請求項 5】

請求項 1 記載の方法であって、さらに、

d) 前記メッセージを、前記ソースから前記複数の中間ノードのうち、第 1 のノードへと伝送する工程と、

e) 前記メッセージを、前記複数の中間ノード中のうち、最後のノードから前記宛先へと伝送する工程と

20

を備えることによって、前記メッセージを、前記最適化パスによって前記ソースから前記宛先へと伝送する方法。

【請求項 6】

請求項 5 記載の方法であって、

前記第 1 の中間ノードおよび前記最後の中間ノードは 1 つであり同じである方法。

【請求項 7】

請求項 5 記載の方法であって、

前記第 1 の中間ノードおよび前記最後の中間ノードは異なる方法。

【請求項 8】

請求項 5 記載の方法であって、

工程 (d) および工程 (e) を実施するために、前記通信ネットワークの 1 つまたはそれ以上の既存のルーティングメカニズムが利用される方法。

30

【請求項 9】

請求項 8 記載の方法であって、

工程 (d) は、さらに、前記メッセージを変更することによって、前記メッセージを前記第 1 の中間ノードへとアドレッシングし、前記宛先のアドレスを前記メッセージ内にカプセル化する工程を含む方法。

【請求項 10】

請求項 8 記載の方法であって、

工程 (e) は、さらに、前記メッセージが応答メッセージを要求する場合に、前記メッセージに関するソースアドレス情報を変更することによって、前記ソースの識別を前記最後の中間モードの識別で置き換える工程を含む方法。

40

【請求項 11】

請求項 1 記載の方法であって、さらに、

前記最適化パスに関する情報を、前記複数の中間ノードにアクセス可能なストレージに格納する工程を備える方法。

【請求項 12】

請求項 11 記載の方法であって、

前記最適化パスに関する情報を格納する前記工程は、さらに、前記最適化パス上の前記 1 つまたはそれ以上の中間ノードのそれぞれに関し、前記メッセージを転送する次のアド

50

レスを格納する工程を含む方法。

【請求項 13】

請求項 1 記載の方法であって、

前記通信ネットワークは、1つまたはそれ以上の確立された通信プロトコルによって特徴付けられ、前記方法は、前記確立された通信プロトコルを変更する必要なく実施される方法。

【請求項 14】

請求項 13 記載の方法であって、

前記確立された通信プロトコルは、{インターネットプロトコル、http、ftp、SSL、およびTCP}を含むグループから選択される1つまたはそれ以上のプロトコルを含む方法。

10

【請求項 15】

請求項 1 記載の方法であって、

前記パケット交換型のコンピュータ通信ネットワークは静的トポロジを有する方法。

【請求項 16】

請求項 1 記載の方法であって、

前記パケット交換型のコンピュータ通信ネットワークは動的トポロジを有する方法。

【請求項 17】

請求項 1 記載の方法であって、

前記デフォルトコストおよび前記代替コストは、{遅延、帯域幅、ジッタ、ロス、およびセキュリティ}を含むグループから選択される1つまたはそれ以上の測定基準から得られる方法。

20

【請求項 18】

請求項 1 記載の方法であって、

前記複数の中間ノードはコンピュータハードウェアを備える方法。

【請求項 19】

請求項 1 記載の方法であって、

前記複数の中間ノードはコンピュータソフトウェアを備える方法。

【請求項 20】

請求項 1 記載の方法であって、

前記複数の中間ノードは、確立された通信プロトコルに適合してデータを送受信するように動作可能である方法。

30

【請求項 21】

請求項 20 記載の方法であって、

前記複数の中間ノードは、さらに、前記確立された通信プロトコル以外のプロトコルに適合してデータを送受信するように動作可能である方法。

【請求項 22】

静的トポロジを有したコンピュータ通信ネットワーク内において、ソースから宛先へとメッセージを伝送するための最適化パスを決定するための方法であって、

a) 前記メッセージを伝送するリクエストに応じ、前記通信ネットワークの1つまたはそれ以上の既存のルーティングメカニズムによって得られるデフォルトパスに沿って、前記ソースから前記宛先へのコストを測定する工程と、

40

b) 前記コンピュータ通信ネットワーク上に仮想のトポロジを規定する複数の中間ノードの内、前記デフォルトパス上ではない1つまたはそれ以上の中間ノードを通る複数の代替パスに沿って、前記ソースから前記宛先へと前記メッセージを伝送する際の代替のコストをそれぞれ測定する工程であって、

各前記中間ノードについて、他の前記中間ノードと間接的に結ばれる間接リンクおよび直接的に結ばれる直接リンクについて通信コストを測定し、測定された通信コストに基づき、間接リンクより通信コストが高い直接リンクを削除する第1の工程と、

第1の工程後に、前記各中間ノードの直接リンクの数が所定数を上回る場合には、前

50

記各中間ノードについて、前記測定された通信コストが最も高い直接リンクを削除する第2の工程と、

直接リンクが所定数以下の中間ノードを通る複数の代替パスを生成する工程とを含み、前記生成された複数の代替パスに対応する複数の前記代替コストを測定する工程と、

c) 前記デフォルトコストと複数の前記代替コストとを比較することによって、前記最適化パスを決定する工程と
を備える方法。

【請求項23】

請求項22記載の方法であって、

前記コンピュータ通信ネットワークはコネクション型である方法。

10

【請求項24】

パケット交換型のコンピュータ通信ネットワーク内において、ソースから宛先へとメッセージを伝送するための最適化パスを決定するためのオーバーレイネットワーク装置であって、前記通信ネットワークは、1つまたはそれ以上の確立された通信プロトコルによって特徴付けられ、前記装置は、

a) 前記確立された通信プロトコルに適合してデータを送受信するように動作可能であり、前記コンピュータ通信ネットワーク上に仮想のトポロジを規定する一群の複数の中間ノードと、

b) 前記ソースから前記宛先へと前記メッセージを伝送するリクエストに応じ、デフォルトパス上ではない1つまたはそれ以上の前記中間ノードを通り且つ前記ソースと前記宛先とを繋ぐ代替パスを発見するように動作可能な代替パス発見手段であって、前記中間ノードについて、他の前記中間ノードと間接的に結ばれる間接リンクおよび直接的に結ばれる直接リンクについて通信コストを測定し、測定された通信コストに基づき、間接リンクより通信コストが高い直接リンクを削除し、削除後における、前記各中間ノードの直接リンクの数が所定数を上回る場合には、前記各中間ノードについて、前記測定された通信コストが最も高い直接リンクを削除し、直接リンクが所定数以下の中間ノードを通る複数の代替パスを生成し、前記生成された複数の代替パスの中から前記代替パスを発見する、代替パス発見手段と、前記デフォルトパスは、前記通信ネットワークにおける1つまたはそれ以上の既存のルーティングメカニズムによって得られ、

20

c) 前記確立された通信プロトコルを変更する必要なしに、前記代替パスによって前記ソースから前記宛先へと前記メッセージを転送するための転送手段と
を備える装置。

30

【請求項25】

第1者と第2者との間で電子商取引を実施するための装置であって、前記第1者と前記第2者とは、第1および第2のネットワークノードによってコンピュータ通信ネットワークにそれぞれ接続され、前記通信ネットワークは、1つまたはそれ以上の確立された通信プロトコルによって特徴付けられ、前記装置は、

a) 前記確立された通信プロトコルに適合してデータを送受信するように動作可能であり、前記コンピュータ通信ネットワーク上に仮想のトポロジを規定する一群の複数の中間ノードと、

40

b) 前記電子商取引の一環として、前記第1のノードから前記第2のノードへと通信を伝送するリクエストに応じる代替パス発見手段であって、デフォルトパス上ではない1つまたはそれ以上の前記中間ノードを通り且つ前記第1のノードと前記第2のノードとを繋ぐ代替パスを発見するように動作可能である代替パス発見手段であって、前記中間ノードについて、他の前記中間ノードと間接的に結ばれる間接リンクおよび直接的に結ばれる直接リンクについて通信コストを測定し、測定された通信コストに基づき、間接リンクより通信コストが高い直接リンクを削除し、削除後における、前記各中間ノードの直接リンクの数が所定数を上回る場合には、前記各中間ノードについて、前記測定された通信コストが最も高い直接リンクを削除し、直接リンクが所定数以下の中間ノードを通る複数の代替パスを生成し、前記生成された複数の代替パスの中から前記代替パスを発見する、代替パ

50

ス発見手段と、前記デフォルトパスは、前記通信ネットワークにおける1つまたはそれ以上の既存のルーティングメカニズムによって得られ、

c) 前記確立された通信プロトコルを変更する必要なしに、前記代替パスによって前記第1のノードから前記第2のノードへと前記メッセージを転送するための転送手段とを備える装置。

【請求項26】

請求項1記載の方法であって、前記代替パスはオーバーレイノードを含む、方法。

【請求項27】

請求項22記載の方法であって、前記代替パスはオーバーレイノードを含む、方法。

【請求項28】

請求項24記載の方法であって、前記代替パスはオーバーレイノードを含む、方法。

【請求項29】

請求項25記載の装置であって、前記代替パスはオーバーレイノードを含む、装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、コンピュータネットワークに関し、特に、異種コンピュータネットワークを通る改良通信パスを、オーバーレイネットワークによって発見するためのメカニズムに関する。

【0002】

【従来の技術】

今日のインターネットは、相互接続された管理可能ドメインの集合からなる。各ドメインは、どのハードウェア装置を利用するか、そしてどの制御ソフトウェアを実行するか、を決定する1つまたはそれ以上の管理機能の制御下にある。管理可能ドメインをこのようにローカライズ制御するためには、分離された複数の異種システムを共存させる必要がある。異なるドメイン間の通信は、IP（インターネットプロトコル）という通信プロトコルを使用して行われる。インターネットによるデータ通信は、各種サイズのIPプロトコルパケットからなるシーケンスに変換されて伝送される。これらのパケットは、1つまたはそれ以上の管理可能ドメイン内を伝送されて、最終的な宛先へと到達する。

【0003】

IPでは、データの送信者が指定するのは宛先のアドレスのみであり、IPパケットが通る経路を選択するのはネットワーク（複数の管理可能ドメインに分散されたハードウェアシステムおよびソフトウェアシステムの集合）である。ルーティングプロトコルは、異なる管理可能ドメイン間で、様々なソフトウェア技術およびハードウェア技術を介して協力し合うことによって、インターネットにおける相互接続の存在および／または状態をモニタリングする。インターネットは規模が大きく、異種のものが混在しているという特性を有しており、ルーティングタスクは複雑であるので、これらのルーティングプロトコルは最小化されているのが通常であり、接続を確保することを主目的とし、パフォーマンスを最適化するよりもルーティングホップを最小化することを主目的とする傾向にある。その結果、使用量が特に多い時間帯にワールド・ワイド・ウェブ（WWW）を介してアクセスしようとする場合などは、遅延が長くてユーザがイライラするというおなじみの問題が生じる。

【0004】

インターネットの利用を改善するための提案は、数多くなされており、それぞれ幅広い範囲で転送パスを変更して最適なルーティングパフォーマンスを提供する際の、高度化（sophistication）および／または応答時間についての特徴を有している。これらのプロトコルは、インターネット内での実装が非常に困難であるので、実際にはほとんど使用されていない。異なる管理可能ドメイン間の相互協調処理は重要であり、プロトコルに対して転送パスに影響する何らかの変更がなされると、接続の可用性に大きく影響する可能性がある。このため、実際には、非常に単純なコスト関数および／またはヒューリスティックお

10

20

30

40

50

よび／または手操作による介入によって、エンドポイント間の転送パスの決定が行われる。この方法では、ネットワークが十分に活用されない傾向がある。ローカル構成に依存していると、障害が生じた場合に、その障害が修復されるまで或いは手操作で転送パスが再構成されるまでの数分間または数時間のあいだ、その切断された転送パスを辿り続けることになる。Inktomi (Traffic Server product, <http://www.inktomi.com>) やAkamai (FreeFlow product, <http://www.akamai.com>) 等の会社によって、インターネットの通信速度を改善するための技術が提供および／または開発されている。これらの技術では、頻繁に参照されるインターネットのコンテンツを複製するか或いはそのコピーをキャッシュに入れ、その情報を必要とするエンドユーザのクライアントの近くにこれらのコピーを戦略的に分散させる方法をとっている。このようなアプローチは有用であるが、動的に生成されたコンテンツ (例えばCGIスクリプトされたウェブページ等) に対しては限界があったり効果的でなかったりする場合が多い。さらに、これらのアプローチでは、大量データの余分なコピーを物理的に格納する大容量のコンピュータサーバを使用するため、そして頻繁に更新されるコンテンツのリポジトリを同期に維持するために、必然的にかなりのコストが必要となる。

10

【0005】

したがって、動的に生成されたデータを含むデータを伝送する最適化ネットワークパスを、IP等の定着した通信プロトコルを何ら変更する必要なしに選択できるメカニズムが必要とされている。このようなメカニズムは、さらに、ネットワークのユーザに対して透過性があることが好ましい。

20

【0006】

【発明の概要】

本発明は、異種のコンピュータ通信ネットワーク上において、ソースと宛先ポイントとの間でデータ伝送を行うための最適化パスを発見するオンデマンドな方法およびシステムを提供する。本発明は、さらに、既存の通信プロトコルの変更を要しない透過性のある方法で、このような最適化パスに沿ってデータを伝送するための方法および装置を提供する。伝送されるデータは、例えばCGIスクリプトされたウェブページを獲得するためのhttpリクエストのような、宛先からソースへと受信された初期のリクエストに応じて動的に生成されたページを含むことができる (含まなくても良い)。

30

【0007】

簡単に言うと、本発明は、デフォルトパスに沿ってソースポイントから宛先ポイントへとデータを通信するコストを、データ伝送のリクエストに応じて測定するための手順および手段を開示する。デフォルトパスは、通信ネットワークの1つまたはそれ以上の既存のルーティングメカニズムによって獲得されるパスである。本発明は、さらに、特殊な一群の中間ノードを通る1つまたはそれ以上の代替のノンデフォルトパスに沿ってソースから宛先へとメッセージを伝送するコストを測定するための動作および手段を提供する。好ましい一実施形態において、これらの中間ノードは「オーバーレイネットワーク」と称される。最終的には、デフォルトコストと代替コストとを比較することによって、リクエストされた伝送を行うための最適化パスが選択される。

40

【0008】

さらに、本発明は、既存のネットワーク通信プロトコルを変更する必要なしに、オーバーレイノードを通るノンデフォルトな最適化パスに沿ってメッセージを伝送するための動作および手段を開示する。このようなメッセージの伝送には、ソースから第1の中間オーバーレイノードへと、そして最後の中間ノードから宛先へとメッセージを転送することが含まれる。別の一実施形態において、本発明は、伝送されたメッセージに対する応答を、ノンデフォルトな最適化パスに沿って宛先からソースへと戻すための動作および手段を含む。

【0009】

ここで、本明細書で言う「最適化 (optimized)」パスまたは「最適な (optimal)」パスとは、単に、選択されたコスト／パフォーマンスの基準および識別された一群の代替パス

50

に照らして好ましいと判断されたパスを意味する。したがって、グローバルな最適性 (global optimality) を意味しているのではない。

【0010】

【発明の実施の形態】

A. 概要

本発明の好ましい一実施形態は、インターネット等のコンピュータネットワーク上においてソースポイントと宛先ポイントとの間でデータを伝送するための最適化パスを、既存の通信プロトコルに何ら変更を加える必要のない方法で発見し、使用するためのメカニズムを提供する。このメカニズムは、オーバーレイのルーティングメカニズムをオンデマンドのルーティングメカニズムと組み合わせたものであることが好ましい。オンデマンドのルーティングメカニズムに関しては、以下で簡単に説明する。

【0011】

1) 利用可能なリソースを発見および活用するために、既存のインターネットルーティングメカニズムの上に、代替のルーティングメカニズムからなるオーバーレイネットワークを構築する。オーバーレイルーティングメカニズムは、完全に透過性を有するとともに、インターネットルーティングプロトコルから分離されており、分散ユーザアプリケーションとして、インターネット中で、小規模ではあっても広く分散された部分に実装されることが好ましい。図1は、この概念を例示した図である。ノード100、160は、インターネット等のネットワーク上において意図した通信を行うためのソースノードおよび宛先ノードである。これらのノードは、それぞれ伝送リンク110および170を介して基礎のネットワークに接続されている。ノード140a~140n (リンク145a~145nを介して基礎のネットワークに接続される) は、他のネットワークノードを表しており、ネットワークのルーティングメカニズムによっては、ノード100とノード160との間のデフォルト通信パスで利用されるノードである可能性もある。オーバーレイネットワークノード130a~130nは、ネットワークリンク135a~135nを介して既存のネットワーク伝送回線とインフラストラクチャーとを利用することによって仮想のトポロジー (接続形態) を形成する。オーバーレイネットワークは、ノード130a~130n等のような、互いに協力しあって基礎ネットワーク上に転送パスを提供する複数の計算装置を含むことが好ましい。オーバーレイネットワークノードは、すでに確立された既存のインターネットプロトコルを使用して通信することが好ましいので、現行の標準に何ら変更を加える必要はない。各オーバーレイノード130は、オーバーレイパスモジュール150を含むことが好ましく、ソースノードまたは宛先ノードのいずれかもまた、同様にオーバーレイパスモジュール120を含む。以下で詳述するように、これらのコンポーネントは、オーバーレイルーティングを活用することが望ましい場合に、利用可能なIPプロトコルを組み合わせることによってそのための追加機能を提供するようにプログラムされている。

【0012】

本発明の好ましい一実施形態において、オーバーレイネットワークは、主要なISPロケーション等のインターネットサイトに配置された一群の通常のコンピュータシステム (例えばPentium (登録商標) ベースのPC) を含む。これらの各PCは、追加のカスタムソフトウェアモジュールの形でオーバーレイパスモジュール150を備える。これらのモジュールは、本明細書で詳述するように、本発明の目的を達成するために、接続コストの情報と最適な転送パスの情報とを測定して記録するように動作することができる。オーバーレイネットワークは仮想ネットワークである。すなわち、新しい追加のハードウェア (同じ場所に配置された複数のボックス) を使用するものの、既存のネットワークとケーブルインフラストラクチャーとを使用して全ての通信を行うことが好ましい。このオーバーレイネットワークによって、通信コストを測定することが可能な一群の代替パスが、インターネットの周辺に効果的に提供される。所望の宛先へのパスとして、これらの代替パスの方がデフォルトパスよりも良いサービスを提供することが判明した場合に、クライアントはこれらの代替パスを使用することができる。

【0013】

2) オーバーレイネットワークノードを通る改良された代替の転送パスを、オンデマンドで発見する。このプロセスは、図2のフローチャートにまとめられている。このプロセスは、ステップ200において、ソースA（すなわち図1のノード100）に接続されたユーザが宛先Bへのメッセージの通信を（直接的にまたはネットワークゲートウェイによって間接的に）リクエストすると開始される。ステップ210において、リクエストされた通信をデフォルトパスで行うコストが閾値を上回る場合は、ステップ220～225（後ほど詳述するように）において、オーバーレイパスモジュール120が、1つまたはそれ以上のオーバーレイノード130a～130nによって受信および処理された照会をブロードキャストし、オーバーレイパスモジュール150が、互いに協力しあってオーバーレイノード130を通る代替のパスを発見し、このような代替パスのコストが測定される。所定のパスに沿ったデータ伝送コストの測定（ステップ210, 220）は、一般に、コストの測定基準に依存して従来技術を使用して行われる。例えば、対象とするソースノードから対象とする宛先を「Pingする」ことによって、接続の遅延時間を容易に測定することができる。ステップ230～235において、これらの代替パスは、既存のインターネットルートまたは事前に選択されたオーバーレイルートと比較される。

10

【0014】

3) ステップ230～235において、デフォルトパスより優れたパフォーマンスを有し且つオーバーレイネットワークノードを通る新しいパスが発見されたら、ステップ240において、この改良パスに関する情報を、この新しいパスに含まれるオーバーレイネットワークノード130のそれぞれに送信することが好ましい。ステップ250において、本発明の好ましい一実施形態は、デフォルトのインターネット通信メカニズムに何ら変更を加える必要のない透過性の方法でパケットを新しいパスに経路付けるためのメカニズムを利用する。後で詳述するように、IPカプセル化（IP encapsulation）の形を使用することが好ましい。デフォルトパスよりも良いパフォーマンスを有し且つオーバーレイネットワークノードを通る新しいパスが発見されない場合は、ステップ260においてデフォルトパスを使用することが好ましい。

20

【0015】

以下では、本発明の好ましい一実施形態による様々なプロセスおよび構造を個々に詳述する。

30

【0016】

B. 改良されたオーバーレイ転送パスをオンデマンドで発見する

1. モニタリング

本発明は、オーバーレイネットワークを通る有用な転送パスを必要な場合にのみ発見して追加するオンデマンドなルーティングを提供することが好ましい。こうすると、可能な全ての転送パスを事前に計算して記録する手間が省かれるので、ブートストラッピングおよびデフォルトの動作では、デフォルトのインターネットルーティングメカニズムを使用することが可能になる。特に、本発明の好ましい実施形態では、エンドポイントAからエンドポイントBへの新しい転送パスを、（1）AとBとの間でエンドツーエンド通信がリクエストされた場合（図2のステップ200）、および（2）デフォルトインターネットルートよりも良いパフォーマンスを提供するパスのIDが、オーバーレイルーティングネットワークを通して発見された場合（図2のステップ210～215）にのみ作成する。

40

【0017】

したがって、オーバーレイ転送パスの発見は、デフォルトインターネットルーティングパス上で実施されるデータ通信に関し、対象とする1つまたはそれ以上のコスト／パフォーマンス測定基準をモニタリングすることからスタートすることが好ましい。このようなモニタリングは、ゲートウェイルータまたはソースエンドポイントすなわちノード100において実施されるのが最も典型的である。モジュール110は、モニタリングされた測定基準を組み合わせた所定のコスト関数を利用し、特定の所定の要請に見合わないエンドツーエンド通信を検知する。このような通信に関し、検知プロセスは、（1）ソースアドレ

50

スA、(2)宛先アドレスB、および(3)AからBへのデータ通信のコストを、モニタリング動作から抽出する。コスト情報の計算に関しては、後ほど詳述する。次いで、以下で述べるように、オンデマンドで転送パスを発見するために、この情報を使用する。

【0018】

2. 改良パスの発見

ソースノード100（およびデフォルトインターネット転送パス上の任意のルータ）は、特定の要請に見合わないエンドツーエンド通信を発見できる可能性がある。その際には、ステップ220～225を開始するために、モジュール120が、オーバーレイネットワークノード130に照会を送信し、そのオーバーレイネットワークがより良い転送パスを提供できるか否かを決定する。この照会は、その構成に依存して、指定個数（「q」）のオーバーレイネットワークルータ130に送信されることが好ましい。簡単な一実施形態において、q個の転送パス照会メッセージのそれぞれは、(1)宛先アドレスB、(2)ソースアドレスA、および(3)事前定義されたコスト関数Fのための識別子を含むことが好ましい。図1に示す例において、ソースAはノード100であり、宛先Bはノード160である。コスト関数Fは、実行者の優先順位やニーズにしたがって、遅延、スループット、ジッタ、ロスなど、一群のネットワーク通信パフォーマンスの測定基準から導かれることが好ましい。

【0019】

q個のオーバーレイネットワークノード130iが、それぞれ転送パス照会を受信すると、ステップ220が実行され、オーバーレイノードから宛先アドレスBに伝送された通信に関して、割り当てられたコスト関数Fの測定が行われ、値F(B, i)が得られる。F(B, i)は、第i番目のオーバーレイノードから宛先Bへのデフォルトネットワークパスに関して測定される。この簡単な一実施形態において、照会ノードモジュール120は、照会されたq個のオーバーレイネットワークルータのそれぞれから応答を受信する。照会ノードは、特定の転送パスを選択することを応答受信中の任意の時点で決定し、残りの応答を全て無視しても良い。照会ノードモジュール120は、各応答に含まれるF(B)の値をコスト関数F(i, A)と組み合わせることによって、最適化転送パスを選択することが好ましい。F(i, A)は、再びデフォルトネットワークパスに沿って照会ノードからオーバーレイノード130iへの通信コストを測定する関数である。当業者なら理解できるように、コスト関数を組み合わせると、必然的に値が加算される（コストの測定基準が遅延である場合など）或いは最小値が算出される（帯域幅である場合）か、または、コスト関数をパラメータ化して複素結合したものを一般に含んでも良い。いずれの場合であっても、ステップ230～235において、モジュール120は、代替のオーバーレイパスおよびデフォルトパスに関して算出された総コストを使用し、ソースノード100(A)と宛先ノード160(B)との間で通信を行うための最適化パスを選択することが好ましい。

【0020】

3. 複数のオーバーレイノードを通る改良されたパスを発見する

複数のオーバーレイノード130a～130nを有したオーバーレイネットワークに対しては、一般に、2つ以上のオーバーレイノードを通る代替のオーバーレイパス（「マルチホップパス」）を検討すると良い。ここでは、マルチホップパスの発見を含む代替の好ましい一実施形態に関して説明する。

【0021】

基本的に、オーバーレイネットワークは、基礎となるデフォルトインターネットルーティングメカニズムを通して論理的に完全な接続を行うことができる（つまり、どのオーバーレイノードも他の全てのノードと通信することができる）。しかしながら、マルチホップパスを発見するためには、可能なマルチホップパスの全てを徹底的に評価しても計算的な価値はない。したがって、本発明の好ましい実施形態では、オーバーレイネットワークのトポロジー（接続形態）表示をプルーニング（枝取り）することによって、所定のオーバーレイノードから他のオーバーレイノードへの可能な転送パスの数を、比較的小さい固定

数（「N」）まで減少させる。例えば、図3Aは、完全な接続性を有した8つのノードからなるオーバーレイネットワークの初期のトポロジー表示であり、図3Bは、同じオーバーレイネットワークをN=2までプルーニングした状態のトポロジー表示である。

【0022】

この代替の好ましい実施形態では、各オーバーレイノードモジュール150が、他のオーバーレイルータノードまでの（基礎ネットワークのデフォルト通信パスを介した場合の）通信コストを測定する。あるノード130iに関し、そのオーバーレイノード130iから隣接するオーバーレイノード130jまでのデフォルト通信コストを測定するコスト関数 $L(j, i)$ が、ネットワークのプルーニングおよび続くパスの発見のために使用される。プルーニングのプロセスが開始する時点において、各オーバーレイネットワークノード130iは、他の各オーバーレイネットワークノードに対する接続性を網羅したメッシュと、それら他のオーバーレイノード130jへのリンクに関連した一群のコスト $L(j, i)$ とを有する。プルーニングの動作は、2段階からなることが好ましい。第1の段階では、間接リンクよりも高価な直接リンクを削除する。任意の所定オーバーレイノードからの直接リンクの数が、未だNを上回る場合は、第2の段階において、コスト関数が最も悪い直接リンクをプルーニングする。これら2つの段階の終了時には、各オーバーレイネットワークルータが有する直接リンク接続の数がN以下になるので、好ましい制約を満たすことができる。結果として得られるオーバーレイネットワーク全体のトポロジーは、完全に接続されている場合もそうでない場合もある。こうして得られたトポロジーは、代替パスを発見するための次のプロセスにおいて使用される。

【0023】

初期のトポロジーの形成と、最終的な論理的トポロジーおよびその関連のリンクコストの決定とは、デフォルトのインターネット通信メカニズムを使用して実施されることが好ましい。オーバーレイノードが特殊な通信メカニズム、または拡張された通信メカニズムによって互いに接続されている場合は（必ずしもそうである必要ではない）、オーバーレイリンクコスト $L(j, i)$ の計算はそのような形で行われる。

【0024】

図2のステップ220と連動するステップ225は、マルチホップ形式の実施形態におけるパス発見のプロセスに対応している。この実施形態において、パスの照会は、上記した簡単な実施形態で上述したようにオーバーレイネットワークに送信されるが、このような照会は、さらに、ノードで転送パスの照会を受信する度に減分される「存続時間」フィールド(TTL)を含むことが好ましい。上記した簡単な実施形態と同様で、パスの照会を受信する各オーバーレイノードは、ステップ220において値 $F(B, i)$ を計算する。また、照会を受信する各ノードは、TTLフィールドがゼロでない場合にTTLフィールドを減分し、その照会を、プルーニングされたオーバーレイネットワークの論理的トポロジー表示中の各隣接ノードへと転送する（上述したように、隣接するノード数は最大でN個であることが好ましい）。これらの各オーバーレイノードは、測定された自身の $F(B, i)$ と、トポロジー表示中の各隣接ノードへのリンクに関する一群のコスト $L(i, j)$ とを、照会ノードに対して応答する。

【0025】

マルチホップ形式のこの好ましい実施形態では、可能な各オーバーレイ転送パスに関する関連のコストを、ステップ230~235において計算および比較する目的で、照会ノードのモジュール110が、適切なコスト関数を全て組み合わせる。したがって、図4Aの実施形態において、対象となる最終的な通信パスは、ゲートウェイノードG（ソースノードAはこのノードを介してネットワークにアクセスする）と宛先ノードBとを結ぶパスである。GからBへのデフォルトネットワークパスによるコストは、ノードRとしてまとめられており、総コスト $C(A, B)$ を有することが示される。図4Aは、また、図3Bと同様で、オーバーレイネットワーク400をプルーニングした後のトポロジー表示を示した図であり、8つのオーバーレイノードおよびN=2の接続性によって特徴付けられる。図4Aに示すように、この実施例では、パスの発見に関する照会をノードGからオーバー

レイノードO1に1つ送信すると、合計3つの代替オーバーレイパスに関する応答が得られる。図4Bでは、このような代替オーバーレイパスの1つとして、ゲートウェイGからオーバーレイノードO1およびオーバーレイノードO2を経て宛先Bに至るマルチホップ形式の転送パスが示されている。この実施例では、この代替オーバーレイパスに関連する総コストを、F(O1, G)、L(O2, O1)、F(B, O2)を組み合わせることによって得ることができる。

【0026】

C. オーバーレイ転送パスの構築および使用

ステップ230～235において、ユーザがリクエストした通信を送信するための最適化パスとして1オーバーレイ転送パスが選択された場合、本発明の好ましい実施形態は、オーバーレイネットワークを通る転送パスを動的に構築する方法で既存のインターネットプロトコルを組み合わせる。

10

【0027】

再び図2に関し、ステップ230～235において改良オーバーレイ転送パスが検知されたら、続くステップ240において、パス照会の発信者（例えばソースノード100のモジュール110であるのが代表的）は、改良転送パスに関する情報を、その改良転送パスに含まれる各オーバーレイノードに格納する。パス情報は、テーブル等の形で各オーバーレイノード130に格納されることが好ましいので、宛先エンドポイントが指定されると、転送パス上の1オーバーレイノードは、メッセージの転送先であるノンデフォルトパス上において、次のノードのアドレスを検索することができる。本発明の好ましい一実施形態では、AとBとの間で所定の時間以上エンドツーエンド通信が生じない場合、またはオーバーレイ転送パスのパフォーマンスがデフォルトインターネットパスのそれを下回ると確信される場合、または単に所定の時間が経過した場合に、オーバーレイノードによってこの情報を削除して良いので、後続の通信リクエストでは、ネットワークの最新の状態に基づいてオーバーレイパスの発見および測定が行われる。

20

【0028】

図5は、本発明の好ましい一実施形態において、ノンデフォルトのオーバーレイパスを通してメッセージデータを送信するために実施される方法（図2のステップ250）を、詳細に示した図である。従来技術によると、IPプロトコル等のプロトコルにしたがったメッセージパケットは、一般に「ソース」フィールドと「宛先」フィールドとを含む。「宛先」フィールドは、意図する宛先へとパケットを転送するために、基礎通信ネットワーク（例えばインターネット）のノードにおいてルーティングメカニズムによって使用される。本発明の好ましい一実施形態では、デフォルトのインターネットルーティングパスを効果的にバイパスし、オーバーレイノードを通る代替の最適化転送パスを利用するために、本「カプセル化」と称される技術が、発明のシステムにおける主要なメカニズムとして使用される。カプセル化は、もとの宛先IPアドレスを別の宛先IPアドレスで置き換え、もとの宛先アドレスをエンコード化カプセル内に記録する技術である。

30

【0029】

例として、メッセージが一方通行（応答なし）で且つ代替のオーバーレイパスがワンホップパスである（例えばオーバーレイノードを1つだけ通る）場合を取り上げる。この実施例では、ノード100におけるクライアント（またはゲートウェイノード100を通してネットワークに接続されたクライアント）が、インターネットのようなネットワーク上で、宛先ノード160にメッセージを送信しようと試みる。本発明の好ましい一実施形態では、160との通信のための最適化オーバーレイパスを発見するために、ステップ210～240が実施される。ここで、メッセージを160に送信するための最適化パス（転送率によらずデフォルトネットワークパスよりも適している）は、パケットを100からオーバーレイノード130aへと送信し、さらに130aから160へと転送するパスであることが決定されたとする。すなわち、100→130aのデフォルトネットワークパスを使用して100から130aにパケットを送信し、130a→160のデフォルトネットワークパスを使用して130aから160にパケットを転送するパスが望ましいと仮

40

50

定する。ステップ250では、実際にデータの伝送が行われる。図5は、この様子を詳しく示したフローチャートである。ステップ500では、ノード100におけるオーバーレイソフトウェア120が、160ではなく130aへとパケットをアドレッシングするが、さらに、160のアドレスを、そのメッセージに取り込まれた所定のフォーマットに「カプセル化する」またはエンコードする。次いで、ステップ510では、好ましくはデフォルトネットワークルーティングメカニズムによって、オーバーレイノード130aへとメッセージが送信される。130aでパケットが受信されたら、オーバーレイソフトウェア150aによって、カプセル化データがデコードされるまたは非カプセル化され、エンコードされた「160」のアドレスが発見される。ステップ520では、ノード130aのモジュール150aが、ステップ240で事前に格納されたオーバーレイパス情報をチェックし、オーバーレイ転送パス上の次のノードを識別する。この実施例では、転送パス上に他にオーバーレイノードが存在しないで、ソフトウェア150aはステップ530に進み、ノード160にリセットされた宛先アドレスを使用して、もとのメッセージを復元する。繰り返すが、この実施例には応答メッセージが含まれないので、ソフトウェア150aはステップ580に進み、ノード160にある最終的な宛先へと単純にパケットを転送する。この方法によると、もとのメッセージは、オーバーレイノード130aを通るノンデフォルトの最適化パスに沿って、クライアント（またはゲートウェイ）100から宛先ノード160へと伝送される。これは、基礎のネットワークで確立された通信プロトコル（例えばIP）を何ら変更する必要なく、そして宛先ノード160を何ら変更する（或いは認識することすら）必要なく達成される。

【0030】

次に、マルチホップオーバーレイパスを含む場合を取り上げる。この実施例でも、やはり一方通行の通信を取り扱うことにする。この実施例では、100から160へとメッセージを伝送するためのオーバーレイノード130a、130bを通る最適化パスが、ステップ210～240のプロセスにおいて発見されたとする。すなわち、この実施例では、100→130aのデフォルトネットワークパスを使用して100から130aへとパケットを送信し、130a→130bのデフォルトネットワークパスを使用して130aから130bへとパケットを転送し、最後に130b→160のデフォルトネットワークパスを使用して130bから160へとパケットを転送するパスが望ましいと仮定する。繰り返すが、ステップ500では、ノード100のオーバーレイソフトウェア120が、パケットを130aへとアドレッシングし、さらに160のアドレスをカプセル化する。そして、ステップ510において、オーバーレイノード130aにメッセージが送信される。ノード130aでパケットが受信されると、ノード130aのオーバーレイソフトウェア150aは、エンコードされた「160」のアドレスを発見し、続くステップ520において、ステップ240で事前に格納されたオーバーレイパス情報をチェックし、オーバーレイ転送パス上の次のノードとしてオーバーレイノード130bを識別する。図5のフローチャートにしたがって、モジュール150aはステップ510にループバックし、モジュール150bが類似の機能を果たしているオーバーレイノード130bへとメッセージを転送する。今度のステップ520では、モジュール150bは、転送パス上にもう他のオーバーレイノードが存在しないことを決定し、ノード160にリセットされた宛先アドレスを使用して、（ステップ530において）もとのメッセージを復元する。この実施例にも応答メッセージは含まれないので、ソフトウェア150bはステップ580に進み、ノード160にある最終的な宛先へとパケットを転送する。この方法によると、もとのメッセージは、オーバーレイノード130a、130bを通るノンデフォルトの最適化パスに沿って、クライアント（またはゲートウェイ）100から宛先ノード160へと伝送される。繰り返すが、これは、基礎のネットワークで確立された通信プロトコルを何ら変更する必要なく達成される。

【0031】

第3の例として、応答を要求するメッセージ（ファイルをゲットするためのhttpリクエストなど）を取り上げる。この実施例でも、上記した実施例と同様に、オーバーレイノ

ード130a, 130bを通るマルチホップの転送パスを扱うことにする。この場合の本発明の好ましい実施形態は、モジュール150bがステップ535に到達し、メッセージが宛先ノード160からの応答を要求していることが決定されるまで、上記した実施例と同じ方法で動作する。図5のフローチャートにしたがって、モジュール150bは、ステップ540においてパケットに関するソース情報を「マスカレード」する。本発明の好ましい実施形態では、転送パス上の最後のオーバーレイノードでマスカレードのタスクを実行することによって、オーバーレイ転送パスの双方向使用を可能にする。マスカレードが実行されない場合、ノード160からノード100に送信される応答は、デフォルトのネットワークルーティングを使用した戻りパスを通るのが通常である。一般に、マスカレードのタスクでは、IPパケットのソースアドレスを、マスカレードするノードのアドレスで置き換え、十分な情報を局所的に記録することによって、再生IPパケットが戻された際に、もとのソースアドレスを復元することを可能にする。インターネット等のネットワークに関連した好ましい実施形態において、マスカレードを行うノードのモジュール150bは、もとのソースアドレスと、パケットの送信元であるポート（あるノードが他のどのネットワークノードと接続しているかを固有に識別するポート）とを局所的に格納する。ステップ550において、オーバーレイノード135bは、マスカレードされたメッセージを宛先ノード160に送信する。ソース情報がマスカレードされているので、応答パケットがノード160から送信されると、これらのパケットは、オーバーレイノード135bへとアドレッシングされる。ノード135bの適切なポートに応答が戻されたら、ステップ570において、モジュール150bは、ステップ540で事前に格納されたノード100のもとのソースアドレス、すなわち扱っている応答メッセージが真に意図する宛先を検索し、この意図するノード100の宛先アドレスがカプセル化されている応答メッセージを構築する。ステップ510に戻り、モジュール150bは、ステップ240で格納されたパス情報にアクセスすることによって（この場合のパス情報は、もちろん、ソース100から宛先160への通信用の最適化オーバーレイパスのそれを、単に逆転させたものである）、ノード100に至る最適化パス上の次のオーバーレイノードへと、カプセル化されたメッセージを転送する。

【0032】

このようにして、ネットワーク上のソースおよび宛先のエンドポイント間で交換される制御パケットおよびデータパケットを、（ソースにモジュール110を追加する以外は）ソースノードまたは宛先ノードを何ら変更することなく、そして基礎ネットワークで確立された通信プロトコルを何ら変更することなく、オーバーレイネットワークを通るパスを辿るように経路変更することができる。既存のネットワークアプリケーションを現実的に解決する案としての本発明による方法は、この透過性によって適応性が向上されるので、これは、本発明の好ましい実施形態によって得られる重要な利点の1つである。

【0033】

D. コストの測定基準、アプリケーション

本発明は、企業、個人、および他の事業体で見られる現実的な問題に対処する方法として、多くの可能なアプリケーションを提供することができる。コスト／パフォーマンスの測定基準は、所望のアプリケーションにとって有利になるように自由に選択してよい。好ましい測定基準には、例えば、遅延、スループット、ジッタ、およびロスが含まれる。以下では、幾つかのアプリケーションに関して説明する。

【0034】

1. 信頼性

オンライントレーディングは、今日の世界経済において重要な役割を担うインターネットアプリケーションである。リアルタイムのトレーディングでは、売りの時にアクセスが割り込まれない必要がある。数分間の注文の遅れが、オンライントレーダーに多額の損失をもたらす可能性がある。実際、最良で且つ最も信頼できるインターネットサービスを有するトレーダーは、他のトレーダーに対して明らかに優位だと考えられる。本発明によるこのアプリケーションでは、動的でオンデマンドなオーバーレイルーティングが、どのよう

に、株式の売却または購入を不可能にする障害をよけたルーティングを可能にするかを説明する。

【0035】

現在のインターネットルーティングプロトコルは、いずれもルーティングパスを再構築するための組み込みメカニズムを有しているので、最終的には、ネットワークの障害を適切な方法で検知して対応することが可能である。このような場合、ストックトレーダーがリンクの障害時に注文をさばけるかどうかは、そのISPのリルーティング機能に依存して非常に様々である。現行のルーティング動作では、局所的な随時ソリューションのために最短でも45秒の時間が必要であり、特定のISPによっては最長で数時間が必要である。本発明によって実現される高速、軽量、オンデマンドなルーティング機能を使用すれば、適切なリルーティングに必要な時間を数秒未満に短縮できる可能性がある。好ましい一実施形態では、ストックトレーダーが注文ボタンを押すと、そのトレーダーのインターネットネットワークで実行されるリルーティングデーモンによって、TCP接続の遅延が測定される。この測定は、1秒のタイムアウトで実施される。リンク障害がある場合は、TCP接続の測定はタイムアウトする。この測定のすぐ後に、本発明によるリルーティングデーモンは、測定された接続時間（リンク障害の場合は1秒）を指定する照会をオーバーレイネットワークに送信する。宛先への代替パスを発見できる場合、オーバーレイネットワークは、その新しいパスを数ミリ秒以内にリルーティングデーモンに戻す。この新しいパスは、ストックトレーダーのローカルネットワークに直ちにインストールされ、後続のTCI接続リクエストは全て、障害のあるパスではなく新しく発見されたこのパスを通して転送される。このリルーティングメカニズムによって、このストックトレーダーは、（1）新しいパスを発見するための反応時間が、数分間から2秒未満まで短縮される、（2）リルーティングメカニズムは、トレーダーのISP構成から完全に分離されているので、独立して管理することができる、の2つの利点を得ることができる。

【0036】

2. 帯域幅

インターネットを通じたマルチメディアデータの拡散は、ホームへの帯域幅の増加に伴って、ますます顕著になる傾向がある。現行の帯域幅の可用性は、音声の伝送や低品質のビデオに対しては十分であるが、近い将来には、帯域幅の増大によってフルモーションビデオの伝送が可能になるだろう。生の帯域幅が増大されるだけでなく、サービス品質ルーティングや他のリソース割り当てプロトコル（RSVP等）によって、ユーザアプリケーションが、予測可能な帯域幅を獲得してサービスを適切なレベルで保証することも可能になる。これらの新しいプロトコルを使用できるようにインターネットを移行させるには、多くの課題が残されており、数年間にわたって普及の試みがなされてきた。

【0037】

この分野において、本発明は、インターネット標準によって提供されるQOSサービスの上層に制御層を追加することを可能にする。このアプリケーションでは、オーバーレイネットワークは、結果として得られるシステムおよびそのプロトコルのパフォーマンスを制御する役割を果たす。こうすると、デフォルトインターネットメカニズムから必要な品質のサービスが送られてこない場合に、本発明にしたがったオーバーレイシステムをフォールバックとして考え付くことが可能になる。本発明にしたがったオーバーレイルーティングシステムは、（現行のまたは将来の）デフォルトのインターネットプロトコルによって提供される帯域幅よりも優れた帯域幅を提供するパスを発見するために使用することができる。このアプリケーションでは、ビデオを受信しているエンドシステムによって、受信されているビデオの品質をモニタリングすることができる。すなわち、ユーザのローカルネットワーク上で実行されるリルーティングデーモンが、ビデオの受信で消費される平均のビットレートを測定し、ビデオの品質が良くないことを検知する。このビットレートが一定の閾値を下回る場合、ルーティングデーモンは、より良いパスがないかどうかをオーバーレイネットワークに照会することができる。すると、オーバーレイネットワークが、代替のパスで使用可能なビットレートを測定し、（もしあるならば）適切なオーバーレイ

転送パスをユーザに戻す。オーバーレイノードは、ビデオソースに対してビデオをリクエストし、そのリンクの品質を測定することによって、代替パスで使用可能なビットレートを測定する。このようなリクエストは、デフォルトのインターネットプロトコルを使用して実施されるので、説明は不要である。ビデオをリクエストするオーバーレイノードは、リルーティングを試みているユーザと同じクレデンシャル (credentials) を保持するように構成されるべきである。オーバーレイノードは、ビデオソースからの有効スループットを測定するだけでなく、テストパケットの伝送を通して、ユーザへのオーバーレイリンクの品質も測定する。より良い総帯域幅を有したオーバーレイ転送パスが発見されたら、リルーティングデーモンに通知し、オーバーレイネットワークを通るようにビデオ受信の経路を変更する。この移行時には同期ロスが生じるので、ユーザは、表示の非連続性を体験する可能性がある。この欠点に関しては、帯域幅のロスが原因で生じる信号劣化の不快加減が、再同期化による非連続性よりも大きいと考えられる場合にのみ、リルーティングするにあたって考慮すれば良い。

10

【0038】

3. 遅延

ハイパースペースは、数年前には考えられなかった方法で人々の距離を縮めつつある。インターネットによって可能になった多数対多数の通信パラダイムは、人々のコミュニケーションのあり方、ビジネスのやり方、そしてレジャーの過ごし方に、革命を起こしつつある。これらの革命的なアプリケーションの多くにおいて、ユーザは、他のユーザの属するコミュニティとほぼリアルタイムに対話しているような錯覚に捕らわれる。本発明によるシステムの使用例として、ユーザがインターネットを通してビデオゲームに参加するビデオゲームのアプリケーションを取り上げる。このアプリケーションでは、伝送の遅延を小さくすることによってゲームの質を高めることができる。また、インターネット上の様々な場所にいる複数の個人がゲームに参加している場合は、遅延が小さくなると、より速い応答時間を得ることができるので、より有利になる。

20

【0039】

本発明による好ましい一実施形態は、伝送の総遅延を短縮できるようなオーバーレイ転送パスを発見することができる。このタイプのアプリケーションでは、エンドツーエンドデフォルト転送パスの品質がアクティブに測定される。このタイプのアプリケーションは、往復遅延の測定に使用できる制御信号メッセージを欠いたUDP上で実行される可能性が最も高い。さらに、ビデオストリーミングのアプリケーションと異なり、リンクの品質を評価するに当たって定義済みのサービス品質を使用することはできない。

30

【0040】

このアプリケーションにおいて、応答を生じるテストパケット（例えばICMPエコーパケット）は、（ユーザの私設ネットワーク上に設けられた）ルーティングデーモンによって宛先へと送信される。したがって、ルーティングデーモンによって往復遅延を繰り返し測定することによって、宛先への遅延が評価される。測定された遅延が大きすぎる（事前定義された閾値を上回る）場合、ルーティングデーモンは、遅延をより小さくできるパスがあるかどうかをオーバーレイネットワークに照会する。遅延をより小さくできるパスが発見された場合、ルーティングデーモンは、そのパスを直ちにオーバーレイに切り替えることによって、伝送の品質向上を図る。この場合、ユーザがゲーム中に非連続性を体験するか否かは、ゲームがどのように実装されているかに依存する。

40

【0041】

E. 結論、本発明の範囲

以上から、本発明にしたがったオンデマンドオーバーレイルーティングの好ましい実施形態によって、ネットワークパフォーマンスを向上させるための実用的で軽量で経済的なメカニズムを提供できることがわかる。このメカニズムは高透過性であるので、IP等の基礎ネットワークのプロトコルや、ネットワーク通信に依存するクライアントアプリケーションのプログラムを変更する必要はない。

【0042】

50

以上の説明では多くの細目および具体例が挙げられているが、これらは本発明の範囲を限定することを意図するのではなく、むしろ、本発明の好ましい実施形態を例示することを意図するものである。したがって、多くの変更形態が可能である。例えば、前述した実施例は、インターネット等のIPネットワークに関するが、本発明は、IP以外のネットワークキングプロトコルにも、他のネットワーク層にも、そしてhttp、ftp、TCP、SSLを含むがこれらに限定されない通信プロトコルにも、適用することが可能である。本発明は、また、インターネット以外のパケット交換ネットワークにも、他の静的トポロジーネットワーク（パケット交換型またはコネクション型のいずれか）にも適用可能である。また、同様のメカニズムを、他のネットワークプロパティ（例えばセキュリティ）を最適化するために使用することもでき、アクティブなネットワークインフラストラクチャまたは可用性の他の計算リソースを使用する等によって、ソフトウェアオンリーの実装で実現することもできる。

10

【0043】

したがって、本発明の範囲は、例示された実施形態によって決定されるのではなく、添付した特許請求の範囲およびそれらの法的な等価物によって決定されるべきである。

【図面の簡単な説明】

【図1】 インターネット等の異種のクライアントサーバネットワークに組み込まれたオーバーレイネットワークの構造的な概観を、本発明の好ましい一実施形態にしたがって示した図である。

【図2】 1つまたはそれ以上のオーバーレイノードを通るデータ伝送用の改良転送パスを発見し、その改良パスを通して意図する宛先へとデータを転送するために、本発明の好ましい一実施形態において実施される、一連の基本動作を示した図である。

20

【図3A】 全ての接続性を網羅したオーバーレイネットワークトポロジーの一例を示した図である。

【図3B】 本発明の好ましい一実施形態にしたがってプルーニングした後の、オーバーレイネットワークの一例を示した図である。

【図4A】 オーバーレイノードを通り且つソースから宛先へとネットワークデータを伝送するための最適化パスの発見を示した図である。

【図4B】 オーバーレイノードを通り且つソースから宛先へとネットワークデータを伝送するための最適化パスのコストの関連性を示した図である。

30

【図5】 既存のネットワークルーティングプロトコルを変更する必要のない方法で、1つまたはそれ以上のオーバーレイネットワークを通る最適化転送パスに沿ってデータを伝送するために、本発明の好ましい一実施形態において実施される動作を示した図である。

【符号の説明】

100…ソースノード

110…伝送リンク

120…オーバーレイパスモジュール

130…オーバーレイネットワークノード

135…ネットワークリンク

140…ネットワークノード

40

145…リンク

150…オーバーレイパスモジュール

160…宛先ノード

170…伝送リンク

【図 1】

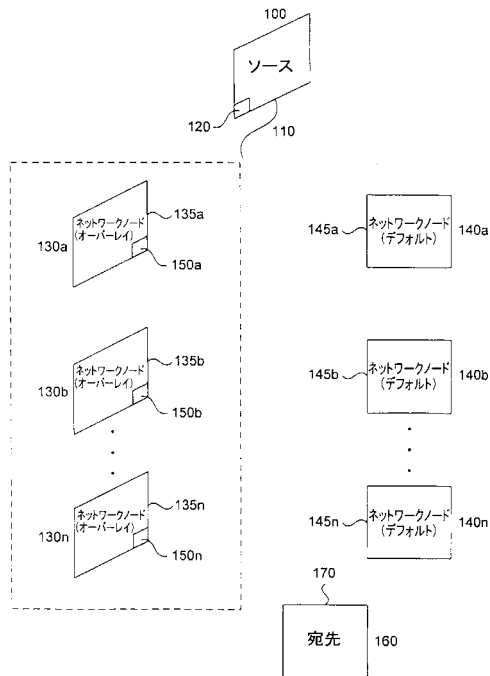


FIG. 1

【図 2】

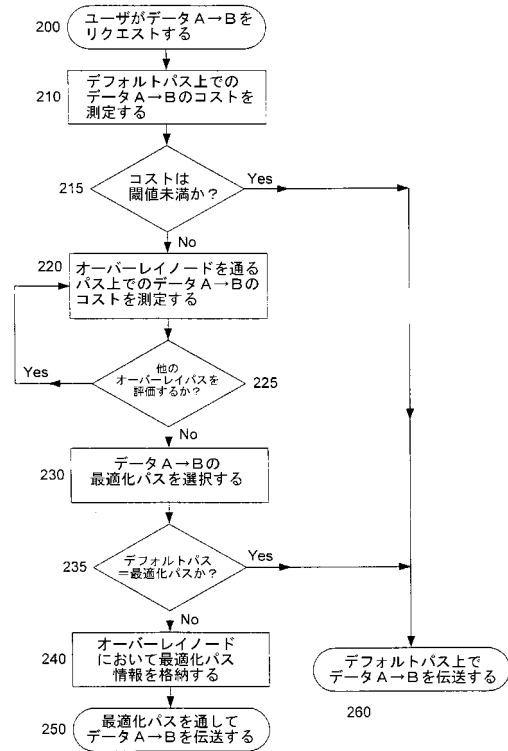


FIG. 2

【図 3 A】

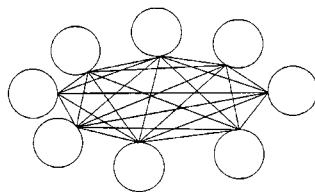


FIG. 3A

【図 4 A】

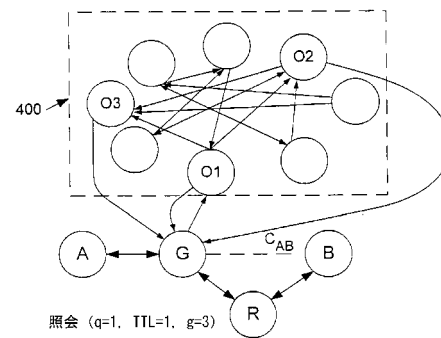


FIG. 4A

【図 3 B】

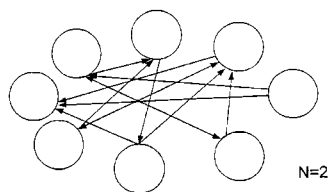


FIG. 3B

【図 4 B】

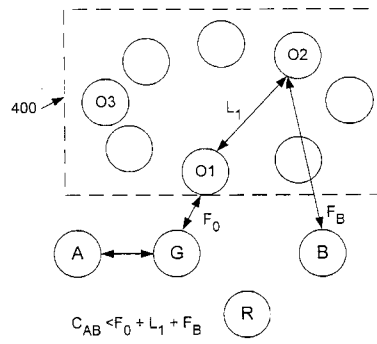


FIG. 4B

【図 5】

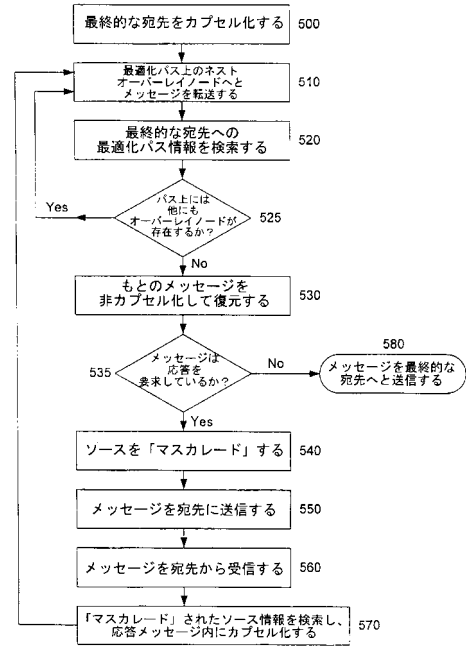


FIG. 5

フロントページの続き

審査官 玉木 宏治

(56)参考文献 特開平10-070571 (JP, A)
特開平05-130144 (JP, A)

(58)調査した分野(Int.Cl., DB名)
H04L 12/00-66