

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2012-216961

(P2012-216961A)

(43) 公開日 平成24年11月8日(2012.11.8)

(51) Int.Cl. F I テーマコード (参考)  
 H04L 12/56 (2006.01) H04L 12/56 200Z 5K030

審査請求 未請求 請求項の数 10 O L (全 15 頁)

(21) 出願番号 特願2011-80163 (P2011-80163)  
 (22) 出願日 平成23年3月31日 (2011. 3. 31)

(71) 出願人 000004226  
 日本電信電話株式会社  
 東京都千代田区大手町二丁目3番1号  
 (71) 出願人 504137912  
 国立大学法人 東京大学  
 東京都文京区本郷七丁目3番1号  
 (74) 代理人 100070150  
 弁理士 伊東 忠彦  
 (74) 代理人 100124844  
 弁理士 石原 隆治  
 (72) 発明者 川原 亮一  
 東京都千代田区大手町二丁目3番1号 日  
 本電信電話株式会社内

最終頁に続く

(54) 【発明の名称】 トラフィック制御方法及びゲートウェイ仮想化ノード装置

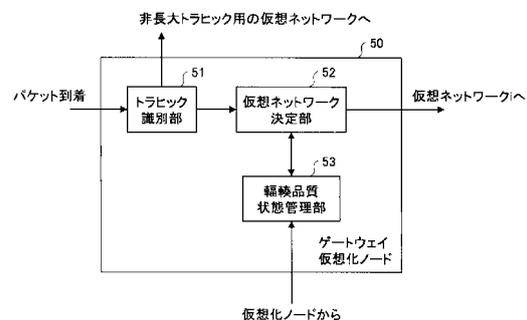
(57) 【要約】

【課題】 単一ドメインに閉じず、トラフィックの変動に対応して輻輳を回避して品質を維持できるようなトラフィック制御を行う。

【解決手段】 本発明は、エンドホストまたはサーバから、あるフローからのパケットが該ゲートウェイ仮想化ノードに到着すると、ゲートウェイ仮想化ノードが、予め定めた測定期間Tの間に到着したトラフィック量(パケット数またはバイト数)をフロー毎にカウントしておき、該トラフィック量と、予め定めた閾値を比較して、トラフィックが閾値を超えていれば、長大トラフィックと判定し、トラフィック識別情報と、各仮想ネットワーク内のトラフィック情報を用いて、複数存在する仮想ネットワークの中から適切な仮想ネットワークを選択してパケットを転送することでトラフィックを制御する。

【選択図】 図2

本発明の第1の実施の形態におけるゲートウェイ仮想化ノードの構成例



## 【特許請求の範囲】

## 【請求項 1】

ある物理ネットワークに接続する仮想化ノードによって構築される論理網である仮想ネットワークがあるとし、複数の仮想ネットワークが構成されているとし、該仮想ネットワークの境界においてどの仮想ネットワークへパケットを送信するかを決定するゲートウェイ仮想化ノードが存在し、エンドホストまたはサーバは該ゲートウェイ仮想化ノードを介して仮想ネットワークに接続するようなネットワークにおいて、

エンドホストまたはサーバから、あるフローからのパケットが該ゲートウェイ仮想化ノードに到着すると、

前記ゲートウェイ仮想化ノードは、

トラフィック測定により得られた該フローのトラフィック識別情報と、各仮想ネットワーク内のトラフィック情報を用いて、複数存在する仮想ネットワークの中から適切な仮想ネットワークを選択してパケットを転送することでトラフィックを制御する仮想ネットワーク決定ステップを行う

ことを特徴とするトラフィック制御方法。

## 【請求項 2】

前記仮想ネットワーク決定ステップにおいて、

前記トラフィック識別情報として、予め定めた測定期間 $T$ の間に到着したトラフィック量（パケット数またはバイト数）をフロー毎にカウントしておき、該トラフィック量と、予め定めた閾値を比較して、トラフィックが閾値を超えていれば、長大トラフィックと識別し、識別されたトラフィックに適した仮想ネットワークへ転送する第 1 のトラフィック識別ステップ、

または、

前記トラフィック識別情報として、パケット間隔およびパケットサイズの分布を測定し、リアルタイム通信トラフィックか否かを判定し、リアルタイム通信トラフィックであると識別された場合は、該トラフィックに適した仮想ネットワークへ転送する第 2 のトラフィック識別ステップ

のいずれかを行う

請求項 1 記載のトラフィック制御方法。

## 【請求項 3】

ある事業者がサービス提供のための物理ネットワークを複数の仮想ネットワークにスライスしてサービスを行うとし、その通信事業者内で利用されるアプリケーションに対して、ポート番号をサービス種別毎に割り当てられている場合に、

前記トラフィック識別情報として、前記ポート番号を参照して、リアルタイム性重視トラフィックかスループット重視トラフィックに分類する第 3 のトラフィック識別ステップを更に行い、

前記仮想ネットワーク決定ステップにおいて、

前記第 3 のトラフィック識別ステップにおける分類に基づいて、それぞれのQoS要求に適した仮想ネットワークへトラフィックを振り分ける

請求項 1 記載のトラフィック制御方法。

## 【請求項 4】

ある入側ゲートウェイ仮想化ノードから出側ゲートウェイ仮想化ノードへ向かうフローのうち、前記第 1 のトラフィック識別ステップで前記長大トラフィックと判定された、または前記第 3 のトラフィック識別ステップで前記スループット重視トラフィックと判定されたフローに対し、入側 - 出側ゲートウェイ仮想化ノードペア $X$ を接続する仮想ネットワークのうち、該フローが利用可能な仮想ネットワークが $M$ 個あるとし、また、各仮想ネットワーク $i$  ( $i=1$ から $M$ )において、測定区間 $t$ におけるネットワークの輻輳品質状態を表すメトリック $r(i, t)$ を測定する第 1 の輻輳品質状態測定ステップを更に行い、

前記仮想ネットワーク決定ステップにおいて、

予め定めた閾値 $Th_r$ に対して、前記メトリック $r(i, t)$ が $r(i, t) < Th_r$ を満たす仮想ネットワークのうち、 $r(i, t)$ が最大となる仮想ネットワークを選択する

10

20

30

40

50

請求項 2 または 3 記載のトラフィック制御方法。

【請求項 5】

前記仮想ネットワーク決定ステップにおいて、

前記メトリックに対し、重み付け選択確率を用いて仮想ネットワークを選択する

請求項 4 記載のトラフィック制御方法。

【請求項 6】

ある物理ネットワークに接続する仮想化ノードによって構築される論理網である仮想ネットワークがあるとし、複数の仮想ネットワークが構成されているとし、該仮想ネットワークの境界においてどの仮想ネットワークへパケットを送信するかを決定するゲートウェイ仮想化ノードが存在し、エンドホストまたはサーバは該ゲートウェイ仮想化ノードを介して仮想ネットワークに接続するようなネットワークにおけるゲートウェイ仮想化ノード装置であって、

エンドホストまたはサーバから、あるフローからのパケットが該ゲートウェイ仮想化ノードに到着すると、トラフィック測定により、トラフィック量、パケット間隔およびパケットサイズの分布、ポート番号のいずれかによりトラフィックを識別するトラフィック識別手段と

測定区間におけるネットワークの輻輳品質状態を表すメトリックを測定する輻輳品質状態管理手段と、

前記トラフィック識別情報、各仮想ネットワーク内のトラフィック情報、前記メトリックを用いて、複数存在する仮想ネットワークの中から適切な仮想ネットワークを選択してパケットを転送することでトラフィックを制御する仮想ネットワーク決定手段と、

を有することを特徴とするゲートウェイ仮想化ノード装置。

【請求項 7】

前記トラフィック識別手段は、

前記トラフィック識別情報として、予め定めた測定期間Tの間に到着したトラフィック量（パケット数またはバイト数）をフロー毎にカウントしておき、該トラフィック量と、予め定めた閾値を比較して、トラフィックが閾値を超えていれば、長大トラフィックと判定する第1のトラフィック識別手段、

または、

前記トラフィック識別情報として、パケット間隔およびパケットサイズの分布を測定し、リアルタイム通信トラフィックか否かを判定する第2のトラフィック識別手段、のいずれかを含む

請求項 6 記載のゲートウェイ仮想化ノード装置。

【請求項 8】

前記トラフィック識別手段は、

ある事業者がサービス提供のための物理ネットワークを複数の仮想ネットワークにスライスしてサービスを行うとし、その通信事業者内で利用されるアプリケーションに対して、ポート番号をサービス種別毎に割り当てられている場合に、

前記トラフィック識別情報として、前記ポート番号を参照して、リアルタイム性重視トラフィックかスループット重視トラフィックに分類する第3のトラフィック識別手段を含み、

前記仮想ネットワーク決定手段は、

前記第3のトラフィック識別手段における分類に基づいて、それぞれのQoS要求に適した仮想ネットワークへトラフィックを振り分ける手段を含む

請求項 6 記載のゲートウェイ仮想化ノード装置。

【請求項 9】

前記輻輳品質測定手段は、

ある入側ゲートウェイ仮想化ノードから出側ゲートウェイ仮想化ノードへ向かうフローのうち、前記第1のトラフィック識別手段で前記長大トラフィックと判定された、または前記第3のトラフィック識別ステップで前記スループット重視トラフィックと判定されたフローに対し、入側 - 出側ゲートウェイ仮想化ノードペアXを接続する仮想ネットワークのうち、

該フローが利用可能な仮想ネットワークがM個あるとし、また、各仮想ネットワーク $i$  ( $i=1\sim M$ )において、測定区間 $t$ におけるネットワークの輻輳品質状態を表すメトリック $r(i, t)$ を測定する手段を含み、

前記仮想ネットワーク決定手段は、

予め定めた閾値 $Th_r$ に対して、前記輻輳品質測定手段で測定された前記メトリック $r(i, t)$ が $r(i, t) < Th_r$ を満たす仮想ネットワークのうち、 $r(i, t)$ が最大となる仮想ネットワークを選択する手段を含む

請求項6または7記載のゲートウェイ仮想化ノード装置。

【請求項10】

前記仮想ネットワーク決定手段は、

前記メトリックに対し、重み付け選択確率を用いて仮想ネットワークを選択する手段を含む

請求項6記載のゲートウェイ仮想化ノード装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、トラフィック制御方法及びゲートウェイ仮想化ノード装置に係り、特に、仮想ネットワーク技術を用いてネットワークを流れるトラフィックの転送効率を向上させるためのトラフィック制御方法及びゲートウェイ仮想化ノード装置に関する。

【背景技術】

【0002】

IPネットワークを代表するインターネットは、多様なアプリケーションの収容を可能とすべく発展・普及してきており、その一方で、QoS (Quality of Service) に対する要求も高まってきている。

【0003】

一方で、端末の多様化・高機能化及びアプリケーションの多様化に伴い、ネットワークを流れるトラフィックは年々増加し、またそのトラフィックパターン（交流トラフィックパターンや時間変動パターン）も大きく変動すると予想される。今後、トラフィックの増加に加えて、アクセス帯域や端末の高速化に伴い、ネットワークにトラフィックが集中して輻輳や品質劣化が生じることが問題になってくると考えられる。

【0004】

現在のIPネットワークでは、トラフィックはOSPF (Open Shortest Path First) のように最短ホップ経路に従うため、トラフィックが集中すると品質が劣化してしまうという問題がある。このような問題を回避する方法として、単一キャリアに閉じてトラフィックエンジニアリングを行う技術として、MPLS (Multi Protocol Label Switching) を用いることが考えられる。しかしながら、MPLS を用いてサービス毎のトラフィックエンジニアリングを実現するには、エッジルータにおいて他の全エッジルータへのLSP (Label Switching Path) を確立する必要がある。この場合、コアルータでは、エッジルータ数の二乗のオーダのさらにサービス数倍のオーダのエントリが必要となる。また、MPLS によるトラフィックエンジニアリングでは、主に予めLSP を張ることを前提としているため、ネットワークの規模が大きくなったり、サービス数が増えるとスケールしないという問題があった。

【0005】

例えば、交流トラフィックが与えられた/測定できたときに、ネットワーク内の各リンク使用率が均一となるようにLSP をエッジルータ間で張る方法が検討されている（例えば、非特許文献1～3参照）。しかしながら、予めトラフィック交流を予測する必要があるという問題点があった。

【0006】

また、ingress-egress エッジルータペアの間に複数のLSP を張っておき、ネットワーク内部の各リンク使用率を一定周期毎に監視して、入側エッジルータに新規フローが到着したら、コストの低いパスを高い確率で選択するようにして負荷分散を図る方法がある（

10

20

30

40

50

例えば、非特許文献4参照)。ネットワーク内部のリンク使用率を監視する代わりに、エッジノードペアの間で試験パケットを用いて各パスの品質を観測する方法もある(例えば、非特許文献5~10参照)。いずれの方法もLSPを予め静的に設定することを前提としているため、サービス毎のトラフィックエンジニアリングを実現する場合には、スケラビリティに問題があった。

【0007】

また単一ドメインに閉じたケースのみにしか適用できないという課題もあった。

【0008】

一方、利用者の多様な要求に柔軟に対応できるネットワーク技術としてネットワークの仮想化技術が近年注目されている。ネットワークの仮想化技術は、物理ネットワークを複数  
10  
の論理ネットワークにスライスして仮想ネットワークを構築する技術である(例えば、非特許文献11, 12参照)。これら仮想ネットワークの利用形態としては、

- ・各研究者が独自のプロトコルを検証するために、各仮想ネットワークを提供する(例えば、非特許文献11, 12参照)；

- ・物理ネットワークリソースを借りて、仮想ネットワークサービスを提供する仮想ネットワークサービス事業者向けにネットワークを提供する(例えば、非特許文献13, 14参照)；

- ・ユーザ毎にスライスを割り当てて、契約に応じてスライス数を変えることで差別化サービスを提供する(例えば、非特許文献15参照)；

20  
というものが考えられている。このような仮想化技術を使うことで、単一ドメインに閉じずに複数ドメインに跨って仮想ネットワークを構築することも可能となる。

【先行技術文献】

【非特許文献】

【0009】

【非特許文献1】Y.Wang, Z.Wang, and L. Zhang, "Internet traffic engineering with out full mesh overlaying," IEEE INFOCOM2001, vol.1, pp.565-571, 2001.

【非特許文献2】Y. Wang, and Z. Wang, "Explicit routing algorithms for Internet traffic engineering," IEEE ICCN99, pp.582-588, 1999.

【非特許文献3】X. Xiao, A. Hannan, and B. Bailey, and L.M.Ni, "Traffic Engineering with MPLS in the Internet," IEEE Netw., vol.14, no.2, pp.28-33, March/April 2  
30  
000.

【非特許文献4】I. Stoica, and H Zhang, "LIRA: An approach for service differentiation in the Internet," Proc. NOSSDAV'98, June 1998.

【非特許文献5】A. Elwalid, C. Jin, S. Low, and I. Widjaja, "MATE: MPLS adaptive traffic engineering," IEEE INFOCOM2001, pp.1300-1309, 2001.

【非特許文献6】D. Gao, Y. Shu, S. Liu, and O.W.W. Yang, "Delaybased adaptive load balancing in MPLS networks," IEEE ICC2002, pp.1184-1188, 2002.

【非特許文献7】T. Saad, T. Yang, D. Makrakis, and V. Groza, "Diffserveenabled adaptive traffic engineering over MPLS," IEEE ICII 2001, vol.2, pp.128-133, 2001.

【非特許文献8】E. Dinan, D.O. Awduche, and B. Jabbari, "Analytical framework for dynamic traffic partitioning in MPLS," IEEE ICC2000, pp.1604-1608, 2000.  
40

【非特許文献9】S.D. Patek, R. Venkateswaran, and J. Liebeherr, "Enhancing aggregate QoS through alternate routing," IEEE GLOBECOM'00, vol.1, pp.611-615, 2000.

【非特許文献10】R. Kawahara, "An adaptive load balancing method for multiple paths using flow statistics and its performance analysis," IEICE Trans. Commun., Vol. E87-B, No. 7, July 2004.

【非特許文献11】Andy Bavier, Nick Feamster, Mark Huang, Larry Peterson, and Jennifer Rexford, "In VINI Veritas: Realistic and controlled network experimentation," Proc. ACM SIGCOMM, September 2006.

【非特許文献12】GENI <http://www.geni.net/wp-content/uploads/2009/04/geni-at-a->  
50

glance-final.pdf

【非特許文献 1 3】Cabo: Concurrent Architectures are Better than One. Nick Feamster, Georgia Tech Lixin Gao, UMass Amherst Jennifer Rexford, Princeton.

【非特許文献 1 4】G. Schaffrath et al., "Network Virtualization Architecture: Proposal and Initial Prototype," ACM SIGCOMM 2009 Workshop VISA, 2009.

【非特許文献 1 5】Katsushi Kobayashi, "Flexible Arrays of Inexpensive Network (FAIN): Toward global parallelism in the Internet to satisfy future traffic growth," ACM CoNEXT 2008 ReArch, 2008.

【発明の概要】

【発明が解決しようとする課題】

10

【0010】

ネットワークの仮想化技術を使うことで、単一ドメインに閉じずに複数ドメインに跨って仮想ネットワークを構築することも可能となる。しかしながら、上述したように、これら先行検討では、ネットワークを仮想化する方法、または仮想化されたネットワークをユーザやサービスに応じて仮想ネットワークを分けて使う方法についての検討であり、トラフィックエンジニアリングを具現化する方法ではない、という問題点があった。

【0011】

本発明は、上記の点に鑑みなされたもので、単一ドメインに閉じず、トラフィックの変動に対応して輻輳を回避して品質を維持できるようなトラフィック制御方法及びゲートウェイ仮想化ノード装置を提供することを目的とする。

20

【課題を解決するための手段】

【0012】

上記の課題を解決するため、本発明は、ある物理ネットワークに接続する仮想化ノードによって構築される論理網である仮想ネットワークがあるとし、複数の仮想ネットワークが構成されているとし、該仮想ネットワークの境界においてどの仮想ネットワークへパケットを送信するかを決定するゲートウェイ仮想化ノードが存在し、エンドホストまたはサーバは該ゲートウェイ仮想化ノードを介して仮想ネットワークに接続するようなネットワークにおいて、

前記ゲートウェイ仮想化ノードは、

エンドホストまたはサーバから、あるフローからのパケットが当該ゲートウェイ仮想化ノードに到着すると、

30

トラフィック測定により得られた該フローのトラフィック識別情報と、各仮想ネットワーク内のトラフィック情報を用いて、複数存在する仮想ネットワークの中から適切な仮想ネットワークを選択してパケットを転送することでトラフィックを制御する仮想ネットワーク決定ステップを行うことを特徴とする。

【発明の効果】

【0013】

上述のように、本発明によれば、単一ドメインに閉じず、トラフィックの変動に対応して輻輳を回避して品質を維持できるようなトラフィック制御の方法およびその装置を提供することができる。

40

【図面の簡単な説明】

【0014】

【図 1】本発明の第 1 の実施の形態における仮想ネットワークと物理ネットワークの基本構成を示す図である。

【図 2】本発明の第 1 の実施の形態におけるゲートウェイ仮想ノードの構成図である。

【図 3】本発明の第 1 の実施の形態における物理ネットワークサーバの構成例である。

【発明を実施するための形態】

【0015】

以下図面と共に、本発明の実施の形態を説明する。

【0016】

50

[ 第 1 の実施の形態 ]

図 1 は、本発明の第 1 の実施の形態における仮想ネットワークと物理ネットワークの基本構成を示す。

【 0 0 1 7 】

仮想ネットワーク 1 0 にある仮想化ノード 2 0 は、論理的に接続しているとする。つまり、論理的に接続されているとは、例えば、物理ネットワーク 3 0 が IP ネットワークである場合には、その接続先仮想化ノードの IP アドレスを知っており、通信可能な状態にある。また、物理ネットワーク 3 0 はいくつかの物理ルータ 4 0 で構成されている。また、各仮想ネットワーク 1 0 および物理ネットワーク 3 0 には管理サーバ 1 0 0 , 2 0 0 が存在する。なお、図 1 では、長大トラフィックのために用意された仮想ネットワーク M 個を記載

10

【 0 0 1 8 】

本図において、各仮想ネットワーク  $i$  ( $i=1$  から  $M$ ) を構成する仮想化ノード 2 0 は、物理ルータ 4 0 内に存在する。仮想ネットワーク  $i$  の仮想化ノード 2 0 は、仮想ネットワーク  $i$  に属するパケットに対するルーティング処理 (どのノードへパケットを転送すべきか決定) を行う。また、物理ネットワーク 3 0 と物理的に接続されたゲートウェイ仮想化ノード 5 0 が存在する。

【 0 0 1 9 】

図 2 は、本発明の第 1 の実施の形態におけるゲートウェイ仮想化ノードの構成例を示す

20

【 0 0 2 0 】

同図に示すゲートウェイ仮想化ノード 5 0 は、トラフィック識別部 5 1、仮想ネットワーク決定部 5 2、輻輳品質状態管理部 5 3 から構成される。

【 0 0 2 1 】

トラフィック識別部 5 1 は、エンドホストまたはサーバからパケットが転送されてきたら、予め定めた測定期間  $T$  の間に到着したトラフィック量 (パケット数またはバイト数) をフロー毎にカウントし、メモリ (図示せず) に格納しておき、該トラフィック量と、予め定めた閾値と比較してトラフィックが閾値を超えていれば、長大トラフィックと判定する。長大トラフィックと判定されなかったパケットは、非長大トラフィック用の仮想ネットワークへパケットを転送する。長大トラフィックと判定されたパケットは、仮想ネットワーク決定部 5 2 へ転送する。ここで、「フロー」とは、例えば、下位の物理ネットワークが IP ネットワークであれば、{ 発信元 IP アドレス ( $srcIP$ )、着信先 IP アドレス ( $dstIP$ )、発信元ポート番号 ( $srcPort$ )、着信先ポート番号 ( $dstPort$ )、プロトコル (Protocol) } の 5 つ組みを同じくするパケット群と定義してもよいし、{  $srcIP$ 、 $dstIP$  } の 2 つ組みで定義してもよい。

30

【 0 0 2 2 】

仮想ネットワーク決定部 5 2 は、輻輳品質状態管理部 5 3 から、各仮想ネットワーク  $i$  ( $i=1 \sim M$ ) における測定区間  $t$  でのネットワークの輻輳品質状態を表すメトリック  $r(i, t)$  を読み出す。  $r(i, t)$  と予め定めた閾値  $Th_r$  を比較し、  $r(i, t) < Th_r$  を満たす仮想ネットワークのうち、  $r(i, t)$  が最大となる仮想ネットワークを選択する。なお、閾値を低めに設定することで、長大トラフィックのようなネットワークに対してインパクトの大きいフローを特定の仮想ネットワーク集約させることができる。その結果生じる空き帯域を他の種別のトラフィック (リアルタイム性重視トラフィック) へと割り当てることが可能となる。

40

【 0 0 2 3 】

輻輳品質状態管理部 5 3 は、各仮想化ノードにおいて一定周期毎に測定している測定情報 (仮想ネットワーク  $i$  における測定区間  $t$  での各仮想化ノード間に張られた仮想リンク  $j$  の使用帯域  $vu_{ij}(t)$ ) を収集し、蓄積している。一方、仮想リンク  $j$  に割り当てられた仮想リンク帯域を  $VC_{ij}$  とする。このとき、測定情報を用いて仮想ネットワークのメトリックを

【 0 0 2 4 】

【数 1】

$$r(i,t) = \min_{j \in S_{i(X)}} (VC_{ij} - vu_{ij}(t))$$

として計算する。但し、 $S_{i(X)}$  は、仮想ネットワーク*i*において該入側 - 出側ゲートウェイノードペア*X*の通信経路上に存在する仮想リンクの集合である。 10

【0025】

この方法は、各仮想ネットワーク10を構成する仮想リンク20に対して、使用可能な仮想リンク帯域 $VC_{ij}$ を割り当てられていることを前提としている。つまり、他の仮想ネットワークのリンクを流れるトラフィックの影響を受けることなく、割り当て帯域 $VC_{ij}$ は仮想ネットワーク*i*が利用可能としている。

【0026】

上記のゲートウェイ仮想化ノード50は、上記の手順でいずれかの仮想ネットワーク10を選択し、その仮想ネットワーク10にパケットを転送する。仮想ネットワーク10上でパケットを転送する方法としては、例えば、物理ネットワーク30がIPネットワークだとすると、仮想化ノード間をIP上でトンネルを設定し、それを仮想リンクとして使用する。例えば、ゲートウェイ仮想化ノード50で、このパケットは仮想ネットワークAに向けて転送すると決定した場合は、パケットを該仮想ネットワークAを構成する仮想リンク(トンネル)を通過させる。 20

【0027】

次に、物理ネットワーク管理サーバ200について説明する。

【0028】

図3は、本発明の第1の実施の形態における物理ネットワーク管理サーバの構成例を示す。

【0029】

同図に示す物理ネットワーク管理サーバ200は、物理リンク状態管理部210、仮想リンク状態管理部220、仮想リンクウェイト変更部230から構成される。 30

【0030】

物理リンク状態管理部210では、測定区間*t*における物理リンク*j*の使用帯域 $u_j(t)$ を測定している。また物理リンク*j*の帯域を $C_j$ として、メトリック( $C_j - u_j(t)$ )を計算し、その値が予め定めた閾値を下回ったら、該リンク*j\**が輻輳していると判断し、その旨を仮想リンク状態管理部220へ通知する。

【0031】

仮想リンク状態管理部220では、各仮想ネットワーク*i*の仮想化ノードから仮想リンク*j*の使用帯域 $vu_{ij}(t)$ を収集し管理している。物理リンク状態管理部210からの通知(物理リンク*j\**が輻輳している旨の通知)を受信したら、該物理リンクを経由する仮想リンク*j\**の使用帯域 $vu_{ij^*}(t)$ を読み出して、仮想リンクウェイト変更部230へ通知する。 40

【0032】

仮想リンクウェイト変更部230では、使用帯域 $vu_{ij^*}(t)$ に関する上位*N*個の各仮想ネットワーク*i*に対して、該仮想ネットワーク*i*の仮想リンク*j*のリンクウェイトを $V_{ij}$ として、 $V_{ij^*}$ を*X*倍( $X > 1$ )とし、 $V_{ij}(j = j^*)$ を*Y*倍( $Y < 1$ )する。但し、*X*は、予め定めた増加係数( $> 1$ )と、リンク使用帯域*x*に関する単調増加関数*f(x)*を用いて、

【0033】

## 【数 2】

$$X = \min\{\alpha, f(vu_{ij^*}(t))\}$$

とする。またYについては、ウェイト変更後の $V_{ij}$ の和が変更前と等しくなるように、Yを設定する。以上の設定変更後のウェイトの値を、仮想化ノード20へ通知する。 10

## 【0034】

仮想化ノード20では、各仮想ネットワークiでの経路制御アルゴリズムとしてOSPFを用い、OSPFで使用される仮想リンクのリンクウェイトを物理ネットワーク管理サーバ20からの指示に従って変更する。これにより、特定の物理リンクにトラヒック集中しないようにする。

## 【0035】

以上の手順を、一定周期毎に実施することで、特定リンクへの負荷集中を回避する。

## 【0036】

また、一定周期毎に実施する代わりに、以下の手順で実施してもよい。 20

## 【0037】

測定区間tにおける、入側 - 出側ゲートウェイ仮想化ノードペア間の仮想ネットワークiでの交流トラヒックが測定できているとする。その場合、上記の手順でリンクウェイトを調整した後に、該交流トラヒックがどのルートにどれだけ流れるかを推定する（OSPFを前提としているので、リンクウェイトが決定すれば、ダイクストラ法により各入側 - 出側ゲートウェイ仮想化ノードペアのルートが決定し、そのルートに沿って、交流トラヒックが流れるとして、各リンクへのトラヒック量を算出可能である）。つまり、現在輻輳しているリンク $j^*$ の使用帯域がどの程度削減されるかを推定でき、制御後の使用帯域推定値 $u_j(t)_{est}$ を用いて、メトリック( $C_j - u_j(t)_{est}$ )を計算し、その値が予め定められた閾値を下回っていたら、リンクウェイトを再度同じ手順で増加させる。以上を、メトリック( $C_j - u_j(t)_{est}$ )が予め定めた閾値を上回るまで繰り返し、その最終的なリンクウェイトを次測定区間t+1でのリンクウェイトとして決定する。 30

## 【0038】

上記の方法において、メトリックが、ネットワークの空きリンク帯域、あるいは、スループットを表しているため、その値が最も大きいネットワークを選択することで、高スループットを実現しようとしている。

## 【0039】

なお、この方法で仮想ネットワークを選定するタイミングとしては、新規に発生したフローについて仮想ネットワークを決定し、そのフローは通信が終了するまで同じ仮想ネットワークを使用する方法、あるいは、一定周期毎に仮想ネットワークを選択し直す方法等がある。 40

## 【0040】

[第2の実施の形態]

第1の実施の形態では、長大トラヒックまたはスループット重視トラヒックを対象にしていたのに対して、本実施の形態では、音声通信サービスのようリアルタイム性重視トラヒックを対象にした場合について述べる。

## 【0041】

本実施の形態における第1の実施の形態からの変更点は以下の通りである。

## 【0042】

・ゲートウェイ仮想ノード50内のトラヒック識別部51において、パケット間隔およ 50

びパケットサイズの分布を測定し、例えば、文献「太井，阿多，岡，"環境非依存型オンラインリアルタイムトラフィック識別手法，"信学技報NS2007-34，2007年6月。」の技術を用いてリアルタイム通信トラフィックが否かを判定する。

【0043】

・ゲートウェイ仮想ノード50内の輻輳品質状態推定部53において、各仮想ネットワーク*i* (*i*=1~*M*)における測定区間*t*でのネットワークの品質状態を表すメトリック*r*(*i*,*t*)として、試験パケットあるいは通信中のフローを測定することで得られる遅延時間やパケット損失率とする。

【0044】

具体的には、仮想ネットワーク*i*において該入側 - 出側ゲートウェイ仮想化ノードペア*X*の間における遅延時間またはパケット損失率を*r*(*i*,*t*)、あるいは、そのようにゲートウェイペア毎ではなく仮想ネットワーク全体としての遅延時間またはパケット損失率の平均あるいは*X*パーセンタイル値を*r*(*i*,*t*)とする。

10

【0045】

・仮想ネットワーク決定部52において、まず、測定区間*t*における仮想ネットワーク*i*に対する重みを*w*(*i*,*t*)とおく。その初期値を*w*(*i*,1)=1(全ての*i*について)とする。そして、区間*t*=1においては、仮想ネットワーク*i*を確率*w*(*i*,1)/ $\sum w(i,1)$ で選択する。区間*t*においてメトリック*r*(*i*,*t*)を測定したら、次周期で用いる重みを

【0046】

【数3】

20

$$w(i,t+1)=w(i,t) \beta^{\{r(i,t)/r_{max}\}}$$

と更新する。ただし、 $\beta$ は、 $0 < \beta < 1$ を満たす予め定めるパラメータであり、*r*<sub>max</sub>も予め設定するパラメータで*r*(*i*,*t*)の取りうる値の最大値に設定する。つまり、

30

$$0 \leq r(i,t) \leq r_{max}$$

となるようにする。そして、区間*t*(>1)においては、仮想ネットワーク*i*を確率*w*(*i*,*t*)/ $\sum w(i,t)$ で選択する。

【0047】

[第3の実施の形態]

第1の実施の形態では、ゲートウェイ仮想ノード50内のトラフィック識別部51において、フロー毎のトラフィック量を測定して長大トラフィックを識別していたのに対して、通信事業者がサービス提供のための物理ネットワーク30を複数の仮想ネットワークにスライスしてサービスを行うとし、その通信事業者内で利用されるアプリケーションに対して、ポート番号をサービス種別毎に割り当てて、ゲートウェイ仮想化ノード50では、ポート番号をみて、リアルタイム性重視トラフィックかスループット重視トラフィックかといった分類を行ってもよい。

40

【0048】

上記の具体的サービス例としては、背景技術の項で述べたように、仮想ネットワークサービス事業者が複数の仮想ネットワークを用意してサービスする形態が考えられる。あるいは、別の例として、物理ネットワークを有する通信事業者が自身の提供するサービスの効率的な提供を目的として仮想ネットワークを利用する形態も考えられる。

【0049】

後者の具体例として、例えば、音声通信サービス*X*(リアルタイム性重視)とデータ通信サービス*Y*(スループット重視)を提供するとし、各サービス向けの仮想ネットワーク

50

を各々複数面用意し、まずフローが発生するとXかYのいずれのサービスを判断する。例えばXだったとすると、次にX用に用意された複数の仮想ネットワークのうち、各仮想ネットワークのトラヒック状況に応じて、適切な仮想ネットワークを選択する。

【0050】

なお具体的な仮想ネットワーク選択方法については、後述する。

【0051】

[第4の実施の形態]

第1の実施の形態で用いたメトリックの代わりとして以下を用いてもよい。

【0052】

仮想ネットワーク*i*において、測定区間*t*における各仮想化ノード間に張られた仮想リンク*j*の使用帯域 $vu_{ij}(t)$ を測定しているとし、また仮想リンク*j*に割り当てられた仮想リンク帯域を $VC_{ij}$ とする。このとき、

10

【0053】

【数4】

$$r(i,t) = \min_{j \in S_i} (VC_{ij} - vu_{ij}(t))$$

20

をメトリックとする。ただし、 $S_i$ は、仮想ネットワーク*i*上の全ての仮想リンクの集合である。

【0054】

第1の実施の形態では、入側 - 出側仮想化ゲートウェイノードペア毎にメトリック計算しているのに対し、本実施の形態では、ペア毎に情報管理せず、仮想ネットワーク全体でメトリックを計算している。こうすることで管理すべき情報の量を削減している。

【0055】

[第5の実施の形態]

第1の実施の形態で用いたメトリックの代わりとして、以下を用いてもよい。

【0056】

各仮想ネットワーク*i*に対して、試験パケットを用いて測定区間*t*における該入側 - 出側ゲートウェイノードペア*X*のスループットを測定し、それをメトリック $r(i,t)$ とする。あるいは、試験パケットを用いる代わりに測定区間*t*において実際に通信中のフローのスループットを測定してもよい。

30

【0057】

[第6の実施の形態]

第1の実施の形態で用いたメトリックの代わりとして以下を用いてもよい。

【0058】

各仮想ネットワーク*i* ( $i=1 \sim M$ )に対して、孤立した仮想リンク帯域を割り当てる代わりに、*M*個の仮想ネットワークで物理リンク帯域（あるいはその*M*個の仮想ネットワークで共有して利用可能な仮想リンク帯域）を共有しているとする。このとき、測定区間*t*における物理リンク*j*の使用帯域 $u_j$ を測定しているとし、また物理リンク*j*のリンク帯域を $C_j$ とする。このとき、

40

【0059】

【数5】

$$r(i,t) = \min_{j \in S_{i(X)}} (C_j - u_j(t))$$

をメトリックとする。但し、 $S_{i(X)}$  は仮想ネットワーク  $i$  において、該入側 - 出側ゲートウェイノードペア  $X$  の通信経路上に存在する仮想リンクの集合である。

10

【0060】

仮想ネットワーク決定部52では、上記の  $r(i, t)$  をメトリックとして仮想ネットワーク1~Mの中から仮想ネットワークを選択する。

【0061】

[第7の実施の形態]

第1の実施の形態での仮想ネットワーク決定部52の動作を以下に置き換えてもよい。

【0062】

測定区間  $t$  における仮想ネットワーク  $i$  に対する重みを  $w(i, t)$  とおく。まず、仮想ネットワーク  $i$  の重みの初期値を  $w(i, 1) = 1$  (全ての  $i$  について) とする。そして、区間  $t = 1$  においては、メトリック  $r(i, t)$  が  $r(i, t) < Th_r$  を満たす仮想ネットワークの中から仮想ネットワーク  $i$  を確率  $w(i, 1) / \sum w(i, 1)$  で選択する。 $Th_r$  は予め定めた閾値である。区間  $t$  においてメトリック  $r(i, t)$  を測定したら、次周期で用いる重みを

20

【0063】

【数6】

$$w(i, t+1) = w(i, t) \beta^{\{r(i, t)/r_{max}\}}$$

30

と更新する。ただし、 $\beta$  は、 $0 < \beta < 1$  を満たす予め定めるパラメータであり、 $r_{max}$  も予め設定するパラメータで  $r(i, t)$  の取りうる値の最大値に設定する。つまり、

$$0 < \beta^{r(i, t)/r_{max}} < 1$$

となるようにする。そして、区間  $t (> 1)$  においては、メトリック  $r(i, t)$  が  $r(i, t) < Th_r$  を満たす仮想ネットワークの中から仮想ネットワーク  $i$  を確率  $w(i, t) / \sum w(i, t)$  で選択する。

【0064】

第1の実施の形態では、最もメトリック値が大きいネットワークを選択していたが、本実施の形態は、第1, 4, 5, 6の実施の形態のいずれかで定義されたメトリックについて、上記の重み選択確率を用いて仮想ネットワークを選択する。このようにする理由について述べる。単純にメトリック値が最大なネットワークを選んだ場合、いくつかの入側 - 出側ゲートウェイ仮想化ノードペアが同時に独立に仮想ネットワークを選択する場合、空き帯域の大きい仮想ネットワークにトラヒックが集中してしまい制御が安定しない、あるいは一期前に空いていたネットワークが却って輻輳してしまうという可能性が考えられる。そこで、このように確率的に選択させることで、一部のネットワークへのトラヒック集中を回避しつつ、空いているネットワークを効率的に利用できるようにしている。

40

【0065】

[第8の実施の形態]

50

本実施の形態では、ある入側ゲートウェイ仮想化ノードから出側ゲートウェイ仮想化ノードへ向かうフローのうち、第2の実施の形態及び第3の実施の形態では、リアルタイム性重視トラヒックと判定されたフローに対し、以下の手順で仮想ネットワークを選択する。

【0066】

入側 - 出側ゲートウェイ仮想化ノードペアXを接続する仮想ネットワークのうち、該フローが利用可能な仮想ネットワークがM個あるとし、また、各仮想ネットワーク $i$  ( $i=1$ から $M$ )において、測定区間 $t$ におけるネットワークの品質状態を表すメトリック $r(i, t)$ を測定しておき、このメトリックを用いて転送すべき仮想ネットワークを決定する。

【0067】

なお、品質状態としては、試験パケットあるいは通信中のフローを測定することで得られる遅延時間やパケット損失率とする。具体的には、仮想ネットワーク $i$ において該入側 - 出側ゲートウェイ仮想化ノードペアXの間における遅延時間またはパケット損失率を $r(i, t)$ 、あるいは、そのようにゲートウェイペア毎ではなく仮想ネットワーク全体としての遅延時間またはパケット損失率の平均あるいはXパーセンタイル値を $r(i, t)$ とする。

【0068】

なお、図2に示すゲートウェイ仮想化ノードの各構成要素の動作をプログラムとして構築し、ゲートウェイ仮想化ノードとして利用されるコンピュータにインストールして実行させる、または、ネットワークを介して流通させることが可能である。

【0069】

本発明は、上記の実施の形態に限定されることなく、特許請求の範囲内において種々変更・応用が可能である。

【符号の説明】

【0070】

- 10 仮想ネットワーク
- 20 仮想化ノード
- 30 物理ネットワーク
- 50 ゲートウェイ仮想化ノード
- 51 トラヒック識別部
- 52 仮想ネットワーク決定部
- 53 輻輳品質状態管理部
- 100 仮想ネットワーク管理サーバ
- 200 物理ネットワーク管理サーバ
- 210 物理リンク状態管理部
- 220 仮想リンク状態管理部
- 230 仮想リンクウェイト変更部

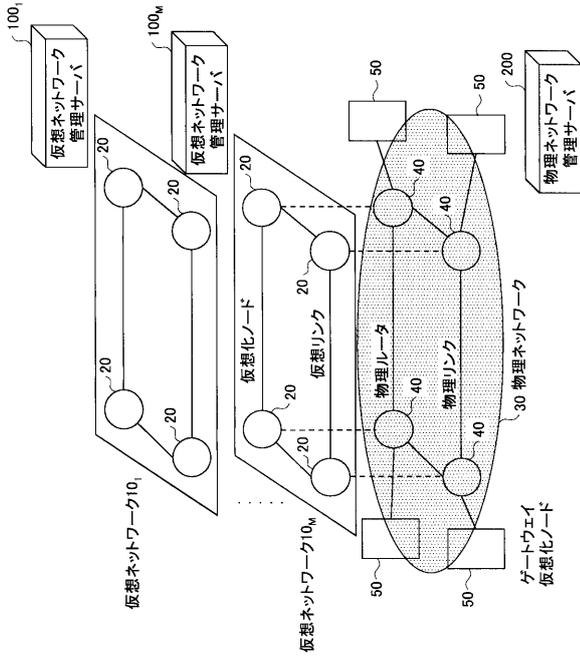
10

20

30

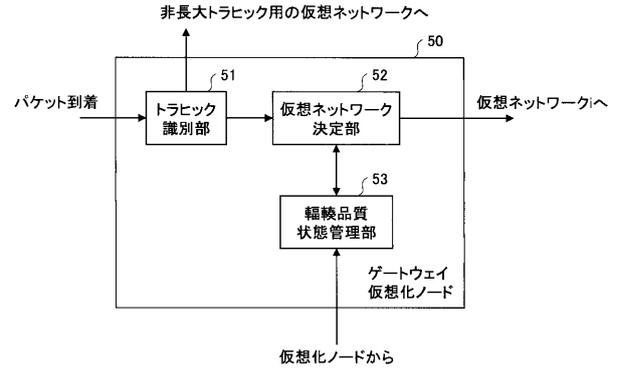
【 図 1 】

本発明の第1の実施の形態における  
仮想ネットワークと物理ネットワークの基本構成を示す図



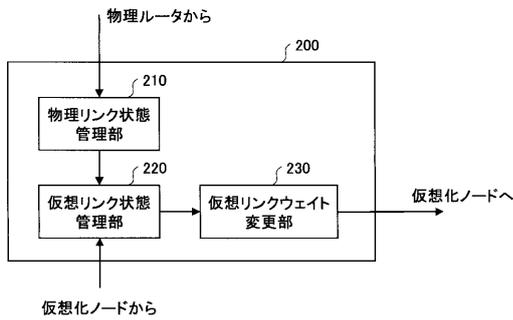
【 図 2 】

本発明の第1の実施の形態におけるゲートウェイ仮想化ノードの構成例



【 図 3 】

本発明の第1の実施の形態における物理ネットワークサーバの構成例



---

フロントページの続き

- (72)発明者 上山 憲昭  
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
- (72)発明者 森 達哉  
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
- (72)発明者 長谷川 治久  
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
- (72)発明者 中尾 彰宏  
東京都文京区本郷七丁目3番1号 国立大学法人東京大学内
- Fターム(参考) 5K030 GA11 HD03 LC11 MB09 MD07