

平成 21 年 3 月 27 日現在

研究種目：基盤研究（C）
 研究期間：2007～2008
 課題番号：19500054
 研究課題名（和文）
 分散仮想マルチレイヤルーティング技術の研究
 研究課題名（英文）
 Research on Distributed Virtual Multi-layer Routing
 研究代表者
 廣津 登志夫（Toshio Hirotsu）
 豊橋技術科学大学・工学部・准教授
 研究者番号：10378268

研究成果の概要：

現在のネットワーク基盤で広く使われている仮想 LAN(VLAN)のような技術では、一つの物理的なネットワークインフラストラクチャ上に複数の仮想的なネットワークがソフトウェアにより構成される。しかし、仮想化されているかどうかは上位層からは区別が付かないため、冗長なトラフィックが発生する可能性がある。本研究では、複数レイヤの連携により仮想化されたネットワーク全体の利用効率を高め効率的な通信を実現する通信基盤の基礎技術を確立した。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	2,100,000	630,000	2,730,000
2008年度	1,400,000	420,000	1,820,000
年度			
年度			
年度			
総計	3,500,000	1,050,000	4,550,000

研究分野： 総合領域
 科研費の分科・細目： 情報学 計算機システム・ネットワーク
 キーワード： 情報通信システム

1. 研究開始当初の背景

『ネットワーク仮想化技術』は単一の物理ネットワーク上に複数の論理的なネットワークを実現するもので、物理的な場所の制約を受けずにネットワークを構築でき、ネットワークの設計・構成の自由度が格段に向上することから、現在では広く使われている。例えば、大学のキャンパス内で離れた場所にある複数の部屋を一つの研究室や学科のための単一の論理ネットワークとして構成したり、企業において遠隔地の事務所との間を

VPN(Virtual Private Network)で接続することであたかも閉じた内部ネットワークであるかのような環境を構築したりするのに使われる。

しかし、ネットワークを仮想化すると、従来データリンク層(2層,L2)、ネットワーク(3層,L3)、トランスポート層(4層,L4)の順に重なったプロトコル階層が崩れ、L2の上に別のL2が存在したり、L3、L4の上にL2が存在したりすることがある。プロトコル階層は下位になるほど物理的な性質を反映するが、仮想

化により物理構造と論理構造が乖離すると、論理ネットワークの中継点であるルータの配置によっては、ネットワーク全体の効率を低下させるという問題が生じる。このような冗長トラフィックは、ネットワーク機器により構成された組織内ネットワークだけでなく、オーバーレイネットワークやVPNによる仮想ネットワークでも生じる問題であり、今後のインターネット利用の場面で重要な問題となってくると考えられる。そこで、これらの問題を発見・解消するようなネットワーク基盤の制御技術が必要であった。

2. 研究の目的

本研究では、このような物理構造と論理構造の乖離による効率の低下を回避し、仮想ネットワークによるネットワーク構成の自由度とその上を流れる通信の効率の双方を両立させるようなネットワーク基盤の実現を目指している。本研究課題においては、対象を Ethernet VLAN(仮想 LAN)による組織内のネットワーク基盤に絞り、そこでのネットワーク制御技術の確立を目的とした。

具体的な研究課題は以下の二点である、

- (1) 観測した実際のトラフィック状況からの冗長トラフィックの発見と、状況に応じた通信トポロジの制御アルゴリズム
- (2) 実際のネットワーク機器を対象とした制御のためのメカニズムの実現

3. 研究の方法

最初に、典型的な組織内ネットワークを模式化したトポロジ(図 1)を用いて、VLAN 環境における問題を簡単に説明する。ここでは 4 台のスイッチ(SW-C, SW-E1, SW-E2, SW-E3)で構成された一つの物理ネットワーク上に、3 つの仮想 LAN(VLAN A, VLAN B, VLAN C)が構成されている場合を考える。VLAN はデータリンク層のネットワークであり、各々の VLAN に接続されている機器が相互に通信するた

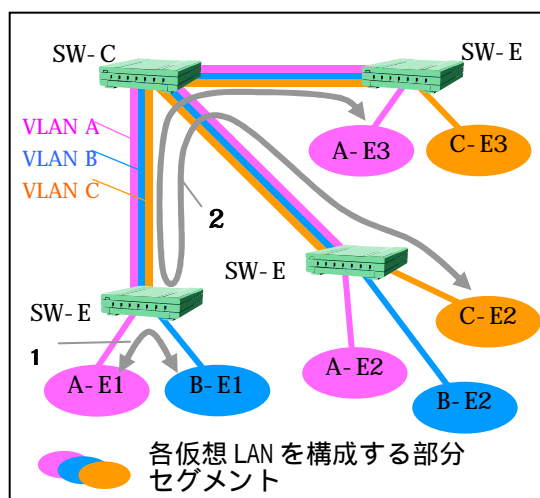


図 1 冗長トラフィックの発生

めには、ネットワーク層の中継機能であるルータにより接続する必要がある。高機能なスイッチ(L3 スイッチ)にはこのネットワーク層の中継機能をもつものもあり、ネットワーク層のインターフェースを設定することで中継点を設定することができる(以下、ルーティングポイントの設定と呼ぶ)。

このネットワークで VLAN A のルーティングポイントを SW-E1 に設置した場合、A-E1 と B-E1 の間の通信は効率的であるが(図中矢印 1)、A-E3 と C-E2 の間の通信は同じデータリンクを 2 回流れるという冗長が生じネットワークの利用効率が悪くなる(図中矢印 2)。しかし、SW-C にルーティングポイントを設置すると、全てのネットワーク間通信がスイッチ間の接続を複数回跨ぐことになり、さらに全体の効率低下をまねく。

ここで、ルーティングポイントを何処に配置するのが効率的であるかは、実際のネットワーク上のトラフィックの状況によって変わってくる。そのため、ネットワーク機器があたかも一つのルータであるかのように相互に協調的に動作して、各 VLAN の中継を適切な所で行う「分散仮想ルーティング」の機能が重要になってくる。Ethernet VLAN を対象とした場合には、VLAN レベルでの通信の状況を把握しルーティングを行う中継点(ルーティングポイント)を適切に設定することにより、より効率の良い構成へと動的に変えていくことになる。本研究課題においては、以下の順に研究を進めた。

- (1) 実ネットワークのトラフィック観測
Ethernet のパケットキャプチャによるデータ収集と、パケット解析による冗長トラフィックの発見。
Netflow data export(NDE)によるデータ収集に基づいたルーティングポイント制御アルゴリズムの開発
- (2) Virtual Router Redundancy Protocol (VRRP) の応用によるルーティングポイント移送機構の実現

4. 研究成果

研究の各段階毎に内容と成果をまとめる。

- (1) 実ネットワークのトラフィック観測
豊橋技術科学大学の学内ネットワークに複数の観測点を設置し、パケットキャプチャと netflow の二つの手法で、実ネットワークのトラフィックデータを収集した。パケットキャプチャは、センタースイッチ(左図の SW-C に相当)とエッジスイッチ(左図の SW-E1 ~ E3 に相当)の間の通信線に対して行い、libpcap を用いて解析を行った。netflow は、センタースイッチで中継されているトラフィックの情報を収集した。

パケット解析による冗長トラフィック

の発見

収集したキャプチャデータを解析し冗長トラフィックの発生を発見できるようにした。トポロジ的な折り返しによる冗長トラフィックは、パケットの送受信 IP アドレスは変わらないのに、データリンク層の MAC アドレスだけが変わるので、IP アドレス対に対する MAC アドレス対を調べれば冗長の発生が判定できる。実際に収集・解析したパケットデータにおいて、10%程度のトラフィックが冗長を発生させていることが判明した。

netflow によるルーティングポイント制御アルゴリズム

まず、図 3 VLAN ネットワークの例 3 のような構成を単純化した VLAN のトポロジを

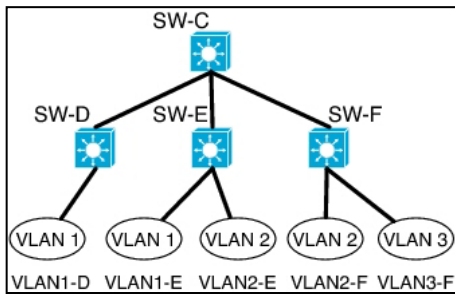


図 3 VLAN ネットワークの例

考える。

このトポロジではコアスイッチ (SW-C) とエッジスイッチ (SW-D, SW-E, SW-F) の両方が L3 スイッチで、ルーティング機能を持っているとする。VLAN は 3 つであるが、説明の便宜上、それぞれのエッジスイッチの配下にある VLAN の断片に名前を付与した (VLAN1-D, VLAN1-E など)。

図中の VLAN3 のように、ある VLAN のセグメントが一つしか無い場合はルーティングポイントをコア側 (SW-C) に配置しようがエッジ側 (SW-F) に配置しようが IP データグラムの転送コストはさほど変わらない。SW-F に配置した場合は他スイッチ配下の VLAN との通信において、基幹 VLAN を通って SW-C を経由することになるので、IP 層で見たホップ数が一段増えるが、物理的なパケットの配送経路は同じである。ただし、SW-F 内に他の VLAN のルーティングポイントがある場合、トラフィックがエッジ内で収束できる可能性があるため、単一の VLAN セグメントの場合は基本的にエッジ側に配置することとする。この配置はネットワークの拡張などにより VLAN3 が他のスイッチに新設されるような構成変更を除けば、静的なトポロジ情報から実現することができる。

一方、複数のスイッチに跨る VLAN の場合は、状況がもう少し複雑になる。上図において、

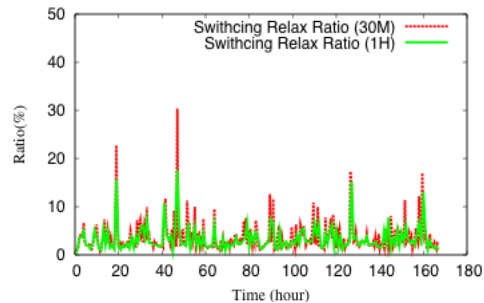
VLAN1 と VLAN2 の間のルーティングについて考える。ここで、VLAN1 のルーティングポイントを SW-D に、VLAN2 のルーティングポイントを SW-F に配置すると、SW-E の配下にある VLAN1-E と VLAN2-E の通信は、SW-C との間を 3 往復する (VLAN1-E SW-D SW-F VLAN2-E と流れる)。

しかし、VLAN2-F と VLAN3-F の間の通信が頻繁であれば、全体としては VLAN2 のルーティングポイントを SW-F に配置する方が得策である。したがって、このような場合はネットワークの動的な状態に依存して決定する必要がある。これを制御するためには、全ての VLAN セグメント間の通信状況のマトリックスから最適になるような推定を行う。ここでは、各 VLAN セグメントの使われ方の傾向はある程度の期間続くと考え、一定期間 (周期) の通信状況マトリックスから次の周期の配置を決定することにする。ある VLAN のルーティングポイントをコアからあるエッジスイッチに移動するかどうかの決定は、そのエッジスイッチ配下で折り返した周期中のトラフィック量が、そのエッジスイッチ以外の全てのエッジスイッチ配下にある当該 VLAN の部分セグメントが他のスイッチとの間で通信した量より多ければ、そのエッジスイッチに配置するという戦略をとる。例えば、VLAN2 のルーティングポイントの設定を SW-E に移動するかどうかの判断は、VLAN2-E の SW-E 内の通信量と VLAN2-F の SW-E 以外のスイッチ向けの通信量を比べて、SW-E 内の通信量が多い場合にだけ移動する。

以上のアルゴリズムを、実際の netflow のデータを元に評価を行った。

図 2 に転送パケット数の削減量をシミュレーションにより求めた結果を示す。それぞれのグラフは、トラフィック状況の学習に用いた周期の違いを示しており、30 分学習して次の 30 分の制御に適用するものから 24 時間単位のものまでの結果を示した。対象のデータには最大で 40%、平均で 4%程度の冗長が含まれており、最大で 30%程度の

Effects of Routing Point Migration (Switch Reduction Ratio)



Effects of Routing Point Migration (Switch Reduction Ratio)

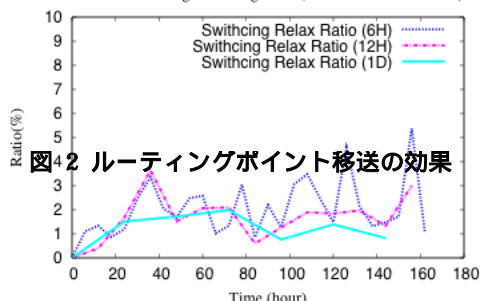


図 42 ルーティングポイント移送の効果

転送量の削減が実現できことが明らかになった。

(2) VRRP によるルーティングポイント移送機構

数ヶ月間隔の稀な設定変更であれば、定期的に手動で行うという手法も考えられるが、もっと頻繁に発生する変動に対する適応を考慮すると自動的にルーティングポイントを移動させるための仕組みが必要となる。そこで、VRRP のようなルータ冗長化のプロトコルを応用することを考える。

VRRP はネットワークの出口ルータを冗長化し、障害時にスタンバイしている代替ルータに自動的に切り替える仕組みである。VRRP はデフォルトゲートウェイを設定している端末が接続されたネットワークにも対応できるように、IP 層とデータリンク層の境界で稼働する。具体的には、個々の機材に付与された実 VLAN インタフェースアドレス以外にゲートウェイアドレスとなる仮想アドレスを保持しており、そのネットワーク中の機器は中継が必要なトラフィックを仮想アドレス向けに送出する。MAC アドレスも仮想的なものを保持するため、クライアントに気付かずにルータを切り替えることが可能となっている。ある仮想アドレスを担う機器が複数稼働している時には、それぞれに設定された優先度情報を元にアクティブな機器以外はスタンバイ状態となり、その VLAN についてのルーティングは一台のみが処理する。

現在多く見られるスター型のトポロジにおいては、VRRP のルータの組はネットワークポロジ上の同一メトリックに水平に配置することが多い。分散仮想ルーティングで必要となる物理的なメトリックが異なる L3 スイッチ間でのルーティングポイントの移動を実現するには、コアとエッジの L3 スイッチ間で冗長化構成となるようにする。具体的には、各 VLAN では全ての L3 スイッチに実 VLAN インタフェースアドレスを付与した上で、デフォルトゲートウェイとなる仮想インタフェースアドレスを L3 スイッチ間で冗長化して制御する。そして、各 VLAN に各スイッチの実 VLAN インタフェースアドレスの優先度を変更することで、動的なルーティングポイントの移動が実現される。

この機構を実際の汎用な L3 機能を持ったネットワーク機器(Yamaha RTX-1100) で実験環境を試作し、動作を確認した。このメカニズムを前述のアルゴリズムと組み合わせることで、ネットワークのトラフィック状況に応じて、冗長トラフィックを抑制する制御が可能となり、分散協調ルーティングの基礎技術が確立された。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計4件)

廣津登志夫, 福田健介, 栗原聡, 明石修, 菅原俊治, “分散仮想ルータのための動的な中継点制御機構”, 情報処理学会論文誌: コンピューティングシステム (ACS), Vol. 2, No.1, pp.123~132, 2009, 査読有。
O. Akashi, K. Fukuda, T. Hirotsu and T. Sugawara, “Policy-based BGP-control Architecture for Inter-AS Routing Adjustment Computer Communications”, Computer Communications, Vol. 31, pp. 2996-3002, Elsevier, Science Direct 2008.

O. Akashi, K. Fukuda, S. Kurihara, T. Hirotsu and T. Sugawara, “Application of Massively Multi-agent System to Internet Routing Management”, Massively Multi-Agent Technology, LNCS 5043, Springer, pp. 131-145, 2008, 査読有。

福田健介, 佐藤進也, 明石修, 廣津登志夫, 栗原聡, 菅原俊治, “ネットワークポロジの次数情報に着目したサーバ・クライアント負荷分散方式の提案と評価”, コンピュータソフトウェア, No. 24, Vol. 4, pp. 78~87, 2007, 査読有。

[学会発表](計7件)

S. Nomoto, K. Fukuda, M. Uehara, H. Mori, “A Visualization of Internet AS Topology with Valley-free Rules”, Proceedings of International Conference on Complex, Intelligent and Software Intensive Systems (CISIS 2009), pp.279-286, 2009/03/16-19, Fukuoka, Japan, 査読有。

廣津登志夫, 福田健介, 栗原聡, 明石修, 菅原俊治, “分散仮想ルータのための動的な中継点制御機構”, コンピュータシステムシンポジウム (ComSys 2008), 情報処理学会, 2008/11/12-13, 東京, 査読有。
片山忠和, 廣津登志夫, 福田健介, 明石修, 菅原俊治, 村上健一郎, “フリースケールネットワーク方式におけるアドレス変換の解析”, 情報学ワークショップ (WiNF 2008), pp. 161-166, 2008/09/25-26, 名古屋, 査読無。

片山忠和, 廣津登志夫, 福田健介, 明石修, 菅原俊治, 村上健一郎, “フリースケールネットワーク方式の予備評価-仮想アドレス使用量の予測”, 第9回インターネットテクノロジーワークショップ (WIT2008), 日本ソフトウェア科学会, 2008/06/26-27, 小樽, 査読有。

廣瀬博昭, 廣津登志夫, 福田健介, 明石修, 栗原聡, 菅原俊治, “仮想化ネットワークにおけるトラフィック分布の推定”, 第70回 情報処理学会大会, 2008/03/13-15, 筑波, 査読無.

O. Akashi, T. Hirotsu, K. Fukuda and T. Sugawara, “Analysis of Diagnostic Capability for Hijacked Route Problem”, Proceedings of the 7th IEEE International Workshop on IP Operations and Management (IPOM 2007), LNCS 4786, pp. 37-48, 2007/10/31-11/2, San Jose, 査読有.

J. Phuritatkul, K. Nguyen, M. Koibuchi, Y. Ji, K. Fukuda, et. al., “Investigating QoS Performance on a Testbed Network”, International Workshop on Performance Modeling and Evaluation in Computer and Telecommunication Network (PMECT07), 2007/08/13-16, Hawaii, 査読有.

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

取得状況(計 0 件)

〔その他〕

6. 研究組織

(1) 研究代表者

廣津 登志夫 (HIROTSU TOSHIO)
豊橋技術科学大学・工学部・准教授
研究者番号: 10378268

(2) 研究分担者

(3) 連携研究者

福田 健介 (FUKUDA KENSUKE)
国立情報学研究所・アークテクチャ研究系・准教授
研究者番号: 90435503
菅原 俊治 (SUGAWARA TOSHIHARU)
早稲田大学・理工学術院・教授
研究者番号: 70396133
栗原 聡 (KURIHARA SATOSHI)
大阪大学・産業科学研究所・准教授
研究者番号: 30397658
村上 健一郎 (MURAKAMI KEN-ICHIROU)
法政大学・ビジネススクールイノベーションマネジメント研究科・教授
研究者番号: 00409351
明石 修 (AKASHI OSAMU)
NTT 未来ねっと研究所・特別研究員
研究者番号: 無し