

仮想化ネットワーク基盤におけるトラフィック制御技術

廣津 登志夫 (情報メディア基盤センター, 情報環境コア)

1 はじめに

インターネットに代表されるネットワーク基盤は、メディア情報の交換・流通に欠かせないものとなっている。近年、このネットワーク基盤の運用・構築においては仮想化技術が一般的になっており、一つのネットワーク基盤をさまざまな利用環境の要求に応じて柔軟に提供できるようになってきている。一方、ネットワーク特有のレイヤ化アーキテクチャに対して、高い透過性をもつ仮想化技術が導入されたため、複数レイヤを通して全体的にみると、ネットワークの利用効率を落とすような状況が起きうる。本稿では、この問題に関する実ネットワークの状況の調査と、それに対する解決のアプローチについて報告を行う。

2 仮想化ネットワークの現状と問題

現在、大学や企業などのある程度大きなネットワークでは、仮想 LAN (VLAN) 技術を用いてネットワークを構成することが一般的になっている。VLAN によるネットワークの仮想化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、ネットワーク機器内でブロードキャストドメインを分離したり、物理的な場所の制約を受けずにネットワークを構築したりといった利点がある。一方、現在広く利用されている Ethernet による VLAN 技術では、データリンク層で仮想化が実現されており、ネットワーク層 (IP 層) より上位の通信プロトコルからは、従来のデータリンク層と全く同じに見える¹。

図 1 に典型的な VLAN ネットワークの構成例を示す。ここでは、中心となる大型の L3 スイッチ (図中 SW-C、以下センタースイッチと称する) に、複数の L3 スイッチ (図中 SW-E1 ~ SW-E3、以下エッジスイッチと称する) が接続されており、センター・エッジのいずれかの場所で IP 層の packets 中継 (すなわちルーティング) が行われる。このように、VLAN ネットワーク中にルーティング機能を持つ機器が複数存在するネットワークでは、各論理ネットワークの中継点であるルータの配置 (以下、ルーティングポイントの設定と呼ぶ) によっては、冗長トラフィックを多発させ、ネットワーク全体の効率を落とす可能性がある。図 1 では、1 台のセンタースイッチと 3 台のエッジスイッチで構成された一つの物理ネットワーク上に、3 つの VLAN (VLAN A, VLAN B, VLAN C) がセンタース

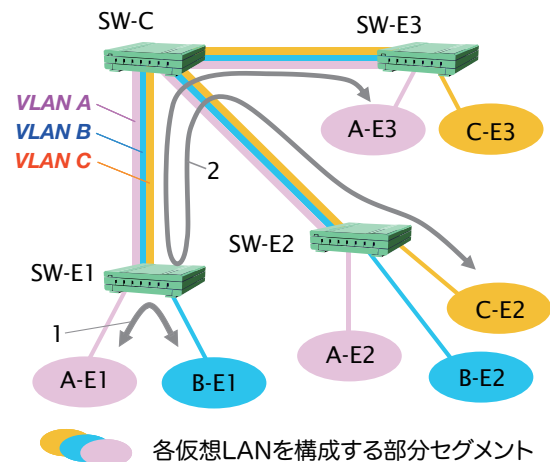


Fig. 1 VLAN トポロジーの例

イッチを跨ぐ形に構成されている。ここで、SW-E1 に VLAN A のインタフェースを設定しルーティングさせる場合、A-E1 と B-E1 の間の通信はエッジスイッチ内で効率的に行われる (図中矢印 1)。これに対して、A-E3 と C-E2 の間の通信について考えてみる。A-E3 のノードはルーティング情報に従って SW-E1 に IP パケットを送信し、SW-E1 は VLAN C にルーティングする、そこから C-E2 に存在するノードに転送するために SW-E2 に向かって送られる。これにより、一つの IP データグラムが SW-C と SW-E1 の間の線を往復することになる (図中矢印 2)。

そこで、センタースイッチを跨ぐ VLAN のルーティングポイントは SW-C に設定するという戦略をとってみる。すると、異なるエッジスイッチを通るような通信 (例えば前述の A-E3 と C-E2 の通信) に関しては冗長を回避できるが、本来ならスイッチ内で閉じていた VLAN 間通信 (例えば A-E1 と B-E1) がセンタースイッチ経由になってしまう。

3 実ネットワークの観測

このような問題がどの程度発生しているかを知るために、実際の VLAN ネットワークのトラフィック状況の観測と解析を行った。使用したデータは、豊橋技術科学大学のキャンパスネットワークのネットワーク機器で収集したトラフィックデータである。観測環境のネットワークは、図 1 に類似した構造を取っており、学内の各ネットワークにはセンタースイッチでルーティングされているものと、エッジスイッチに接続されたルータ (学科や研究室毎に管理) でルーティ

¹すなわち、仮想化は透過的である

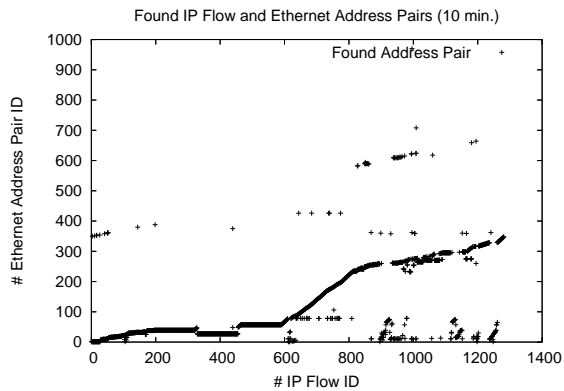


Fig. 2 観測されたL2/L3のアドレス対の対応(10分)

ングされているものがある。

この環境下でセンタースイッチとエッジスイッチの間のトラフィックを観測し、冗長トラフィックがどの程度含まれているかを調べた。観測した各パケットについて、各パケットのIPアドレス対(送信者、受信者)とEthernetアドレス対(送信者、受信者)を抽出し順に識別子を付与すると、その対応が複数あるものが冗長トラフィックと判定できる。ここで、トポロジ上学外のネットワーク向けの接続は一方に限定されるので、外部アドレスは全て0.0.0.0に対応させた。あるIPアドレス対に対して出現したEthernetアドレス対について、10分間に出現した識別子の対応をグラフにプロットしたものが図2である。グラフでは可視性を高めるために同一IPアドレス対に対して出現した二つ目以降のMACアドレス対は、上方にずらして表示してある。このグラフでは350より上方にあるプロットが冗長トラフィックに相当する²。さらに、1日分のデータについても解析した結果、出現したIPアドレス対6250組(学外アドレスは全て0.0.0.0という一つのアドレスに変換)に対して、複数のEthernetアドレス対が確認されたもの546組で1割弱のノードへの通信は冗長トラフィックの可能性を持っていることがわかった。

4 解決に向けて

このような問題の根本的な原因は、TCP/IPを中心とした現在のネットワーク技術が設計・開発された当初は、仮想LANのような技術が存在しなかったため、ネットワークの仮想化技術のことが考慮されていない点にある。当初のネットワークは??章で述べたように、物理的な構造と論理的な構造が一致していた。つまり、端末やルータなどの通信ノードを「点」、通

信ノードが接続されたネットワークを「線」と考えると、「線と線が点で繋がる」形態をとっていた。これに対して、現在の仮想化技術を利用したネットワークは、単一の物理ネットワーク全域に複数の論理ネットワークが「面」的に広がり、個々のデータリンク層のネットワークが層状に重なるような形となっている。つまり、「面と面が接する」形に変化しているのである。そのため、現在の仮想LANによるネットワークは、多数の面を接続する中継(ルーティング)機能を持った複数の中継点が、ネットワーク中に分散して存在する「面と面が多数の点で繋がる」形態になっていると言える。

このような仮想ネットワーク環境においては、ネットワーク全体に分散した中継点(ルータ)が協調して、あたかも単一のルータのように振る舞い適切な中継を行う分散仮想ルーティングの考え方が必要であると考えている。この分散仮想ルーティングが目指すのは、仮想化技術が一般的になった現在のネットワークに適合したネットワーク制御・運用技術の確立および制御機構の実現である。現在、その実現を目指して次のようなアプローチで研究を進めている。

ルーティングポイントの設定支援

ネットワークトラフィックの状況から、静的なルーティングポイントの設定を予測する。

動的なルーティングポイントの移動

ネットワークのトラフィック状況の変動に対して、動的にルーティングポイントを移動する。

多点・並行的なルーティング

ネットワーク中の多点で並行的にルーティングし、常に最適な転送を可能とする。

参考文献

- [1] *IEEE Standard 802.1D-1998, Information technology— Telecommunications and information exchange between systems—Local and metropolitan area networks— Common specifications*, May 1998.

発表論文

- [1] 廣津登志夫、福田健介、菅原俊治, “VLAN環境における分散仮想ルーティングに関する一考察,” 情報処理学会研究報告 (ISSN 0919-6072), Vol. 2006, No. 15, 2006-OS-101, pp. 17-24, Feb. 2006.

²実際には設定ミスによると思われる冗長も存在したが僅かであった(右端の3つめのアドレス対が出現しているもののみ)。