

検査パケットによる帯域圧迫を抑制可能な L2ループ検出手法

野呂 正明^{1,a)} 高野 陽介¹ 小口 直樹¹ 阿部 俊二²

受付日 2021年5月5日, 採録日 2021年11月2日

概要: クラウドにおける IaaS や広域 LAN 接続など, 仮想 L2 ネットワークを顧客に提供するサービスが増加している. こうしたサービスでは, 顧客が構成した L2 ネットワークがループを起こした場合でも, サービス提供事業者は顧客所有の回線の遮断を行わず, 顧客にループの発生を通知する機会が多い. 従来のループ検出手法では, 回線遮断を前提に, ブロードキャストストームの発生をループの検出手段としている. しかし, ループが発生した場合に顧客所有回線を遮断せず, 顧客に問題解決を促す環境では, ブロードキャストストームを前提にした従来のループ検出手法は顧客の通信に大きな影響が出る. そこで本研究では, ループ発生時に顧客のネットワークを圧迫させることなく, ループを検出することが可能な手法を用い, 顧客にループ解消を促す方法を提案する. 本論文ではこれを評価し, ループ検出のために必要な帯域消費量, 所要時間ともに実用上問題ないことを確認した. また, 本研究で想定するネットワークでは, ループするパケットが回線を満たすまでの所要時間は長く, ネットワーク管理者がループを調査して除去する時間的余裕が存在することを確かめることができた.

キーワード: extended Berkeley packet filter, eXpress data path, MAC 層ループ

MAC Layer Loop Detection Method which Reduces Bandwidth Loss

MASAAKI NORO^{1,a)} YOSUKE TAKANO¹ NAOKI OGUCHI¹ SHUNJI ABE²

Received: May 5, 2021, Accepted: November 2, 2021

Abstract: Today Layer 2 overlay networks such as wide area ethernet services and virtual networks sited in the IaaS of cloud services are increasing more and more. In such networks, service providers of overlay networks do not block customer networks even when they detect layer 2 network loops and only report to their customers the existence of loops. However, as most of the conventional methods detect layer 2 network loop when broadcast storm is occurred, the quality of customer network is influenced by broadcast storm in this situation. Therefore, the authors propose layer 2 loop detection method that uses two kinds of unicast packet (probe packet and clear packet) and consume far less bandwidth. In this paper, the authors also evaluate and verify the efficiency of the proposed method. We found that the necessary bandwidth of the proposed method was very small and the duration where test packets retain in the network was also very short. The duration from network loop creation to network link depletion is enough long for network administrators to fix layer 2 network loop.

Keywords: extended Berkeley packet filter, eXpress data path, MAC layer loop

¹ 富士通株式会社
Fujitsu, Kawasaki, Kanagawa 211-8588, Japan

² 国立情報学研究所
National Institute of Informatics, Chiyoda, Tokyo 101-8430,
Japan

a) noro@fujitsu.com

1. 背景

近年のクラウド技術の普及にともない, 仮想化されたネットワークと VM を組み合わせて, 顧客自身で計算環境を構築することが容易となった. また, 複数組織間の仮想

化された MAC 層のネットワークを相互接続するケースも増加しており、ユーザが作成した Virtual Private Network (VPN) が原因でループが発生する可能性が増大している。従来、計算機環境の構築作業はネットワークや計算機の専門的な知識のある人だけが担当してきたが、クラウド環境における仮想計算環境の構築や、構築された計算環境に VPN などで接続するエンドユーザは専門的な知識を持たない人も多い。従来多く用いられてきた Spanning Tree Protocol (STP) [1] は、1つのレイヤ 2 (L2) ネットワークに接続されるすべてのスイッチが STP を有効にしており、ループが検出された場合の切断のためにスイッチ間に優先順位を付与することが求められる。そのため、ネットワークの知識のないエンドユーザに STP を正しく運用することを望むことは難しい。さらに、仮想ネットワーク技術を用いて複数組織間を接続する場合、責任分担の関係上、STP を用いないことにしている広域ネットワークサービスもある。

また、LAN 機器に搭載されたループ検出機能は Virtual eXtensible Local Area Network (VXLAN) [2] や QinQ (IEEE 802.1ad) [3] といった VPN 技術の増加に追従できていないだけでなく、クラウド環境内の仮想ネットワークは STP 以外のループ検出の機能を持たないため、ループ検出の仕組みを搭載した仮想計算機を接続する、もしくは、クラウド上の VM にループ検出用アプリケーションを搭載して運用することが必要な状況にある。

さらに、本研究の想定ユーザである、クラウド事業者、通信事業者や組織のシステム管理者は、自らが用意した計算機環境に接続してきた顧客がループを作成してしまった場合でも、顧客の回線を勝手に切断することができないことが多い。

一方、ネットワークの知識や構築経験が十分でない顧客が接続する環境であっても、ルータ、重要なサービスを動作させる VM や計算機と、それらが直接接続する LAN 機器は、本研究の想定ユーザである技術者に権限がある場合が多く、対策のためのソフトウェアのインストールや、管理用機器の接続が可能であることは珍しくない。このような対策をすることで、重要なサービスを実行する環境にアクセスするために VPN などで接続する顧客が作ってしまった L2 ネットワークのループを早期に発見し、それを顧客に連絡して対策してもらうことができれば、重要なサービスが停止する確率や時間を削減することができる。著者らは、ネットワークにループ検出のための計算機を外付けすることで、ループを検出・切断する手法を提案 [4], [5] してきた。本論文では、ループを検出しても切断できない運用ポリシーのもとでも有効な手法 [6] と、想定するネットワーク環境における有効性について述べる。

2. 想定環境

1 章でも述べたように、本論文で想定するユーザは、クラウドサービスや広域 LAN サービスを提供する事業者や、大きな組織のシステム管理者であり、これらの管理者は、顧客回線を切断することが許されていない場合は珍しくないため、顧客がループを作成してしまった場合に、早期発見・顧客への連絡を行い、事故が大きくなる前に対処することが重要である。

ネットワークにループを作成した場合に大きな事故（ブロードキャストやマルチキャストパケットで回線が埋まる）に至るまでの時間であるが、ループを作成した回線に流れるパケットのうち、ブロードキャストやマルチキャストパケットが占める割合が少なければ、ループが作成されてから事故が大きくなるまでの時間的な余裕がある。

広域 LAN サービスの管理者や大きな組織のシステム管理者が運用を担当する、分散した拠点のネットワークを接続してイントラネットを構築する場合、拠点のルータを VPN や専用回線で広域 LAN サービスが提供する 1つの仮想 L2 ネットワークに接続する。ルータ間接続の L2 ネットワークで、ブロードキャストやマルチキャストパケットが用いられる通信は、ARP やルーティングプロトコルに限定される。また、ルータ間の接続を行うネットワークでは、スタティックな経路情報設定で運用される例も珍しくなく、ブロードキャストやマルチキャストパケットが全体の通信に占める割合は低い。

クラウドの場合、多数のエンドユーザ向けに提供するサービスを動作させることが多いため、VM で動作させる OS は uPnP や Bonjour, CIFS のようなブロードキャストやマルチキャストを多用するサービスをセキュリティ上の理由から停止させていることが多い。そのため、これら VM が接続されている仮想 L2 ネットワークは、ブロードキャストやマルチキャストパケットの占める割合が小さいと考えられる。

以上のような理由から、本研究のユーザ環境ではネットワークのループが発生しても、大事故に至るまでの時間は LAN 環境より長くなることが想定され、ループが発生した場合に、早期にそれを発見して管理者に通知することで大事故に至る前に修復することが期待できる。

3. 従来のループ検出手法

多くの LAN 機器に搭載されている従来手法は、あるポートで受信する単位時間あたりのブロードキャストやマルチキャストパケット数が、ユーザが設定した閾値を超えた場合にループありと判定する方法（パッシブな方法）である。この方法は、ループが存在するネットワークにおいて、ブロードキャストやマルチキャストパケットが大量にループし、回線が圧迫されるような状況にならないと異常ありと

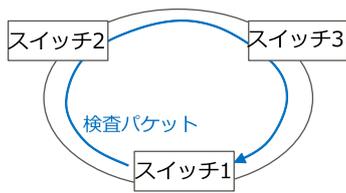


図 1 従来手法

Fig. 1 The conventional active method.

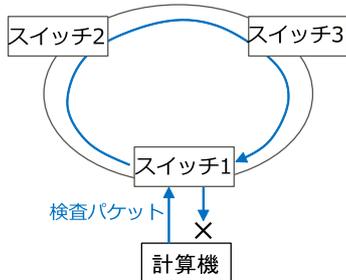


図 2 従来手法を計算機に適用した場合の問題

Fig. 2 The problem of computer with conventional active method.

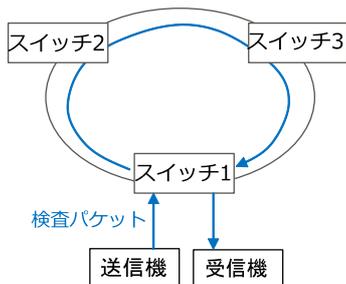


図 3 従来手法を組み合わせた場合

Fig. 3 The combination of two conventional methods.

判定しないため、早めにループを発見して対処の時間を稼ぐという本研究の目的にはそぐわない。

方法 [7] は、図 1 のようにスイッチに内蔵し、周期的にブロードキャスト（もしくはマルチキャスト）パケットを検査パケットとして送信し、自分が送信した検査パケットを受信した場合にループと判定するものであり、早期にループを検出することが可能である。

ただし、1 章でも述べたように、スイッチにこの機能が搭載されていない、もしくは、スイッチに搭載された本機能が利用している仮想ネットワーク技術に対応していないため、構築された仮想ネットワークに接続した計算機や VM にループ検出機能を搭載して利用することを想定している。もし、方法 [7] を計算機に搭載した場合（図 2）、計算機が送信したパケットはソースアドレスが発元計算機のものであるため、発元計算機には配送されない。

また方法 [8] もスイッチに内蔵し、既知のパケットを受信した場合にループが存在すると判断する。図 3 のように、手法 [7] のように検査パケットを送信する送信機と、手法 [8] のように既知のパケットを受信した場合にループあ

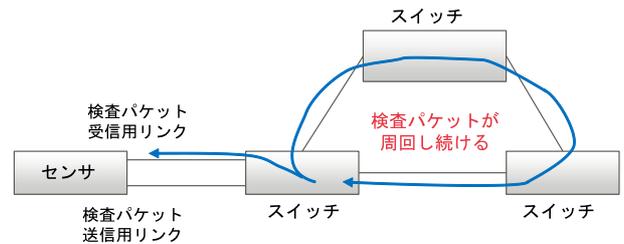


図 4 検査パケットのループ

Fig. 4 Loop of test packet.

りと判断する仕組みを組み合わせることで、スイッチに組み込むのではなく、ネットワークに接続する計算機（もしくは VM）でループ検出を実現することができる。

しかし、この 2 つの手法の組合せでは、検査パケットの宛先がブロードキャストもしくはマルチキャストアドレスであるため、ループを検出してもし回線を切断できない環境では、検査パケットがループし続ける（図 4）。検査パケットを周期的に送信すると、検査パケットが蓄積していき、最終的に回線容量が検査パケットによって占拠される。

ループ発生のある可能性があるネットワークに計算機（もしくは仮想計算機）を接続し、そのうえでループを検出するプログラムを動作させて、早期にループを検出することを目的とする場合、検査パケット自身がループし続けることを防止する必要がある。本研究の提案方式では、ループを検出するための検査パケットがループし続けることを止めることができる別のパケットを送信する。

4. 提案手法

本研究の提案手法では、3 章で説明した 2 つの従来手法（文献 [7] と [8]）を組み合わせた場合の問題である、検査パケットの永続的なループを防止するため、検査パケットにユニキャストパケットを用い、検査パケットのループを止めるループ解消パケットを別途送信することで、検査パケットによるブロードキャストストームの影響を軽減させる。

本研究の提案では、図 5 のようにループを構成する可能性があるスイッチのうちのいずれか 1 カ所に、ループ検出用の検査パケット送信用のリンクと、ループ解消パケット送信用リンクの、合計 2 つのリンクでセンサを取り付ける。このセンサはループ検出時に、検出結果を管理者に通知するため、第 3 のリンクで管理用のネットワークとも接続する。

センサで 2 つのネットワークインタフェースを検査対象のネットワークに接続し、片方のインタフェースは検査パケット送信専用とし、もう一方のインタフェースは検査パケット受信とループ解消パケットを送信するために用いる。また、検査パケットの宛先にはユニキャスト（ネットワーク上に該当アドレスを持つ機器が存在しないアドレス b)）を使う。この検査パケットの宛先アドレスはループを

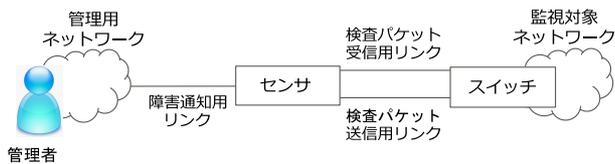


図 5 提案手法で用いるセンサ
Fig. 5 Loop detection sensor.

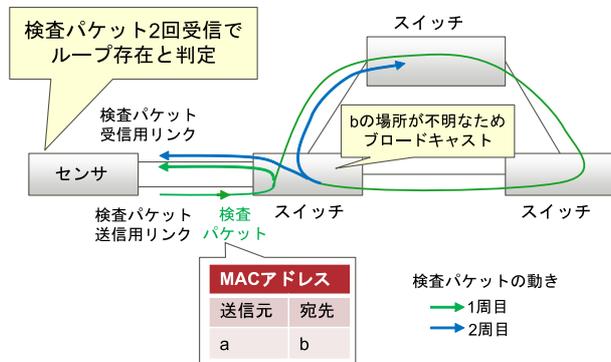


図 6 検査パケットのループ検出
Fig. 6 Loop detection by rounding test packet.

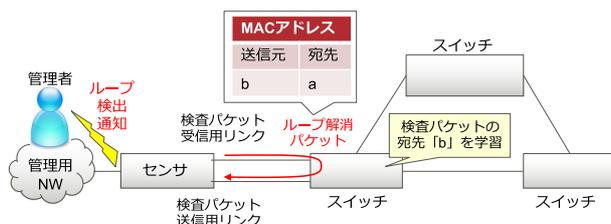


図 7 ループ解消パケットの送信
Fig. 7 Sending clear packet.

構成する経路上のスイッチすべてに学習されておらず、送信された検査パケットは各スイッチでブロードキャストされる。

ネットワークがループを構成している場合は、センサが送信した検査パケットはループし、検査パケット受信用インタフェースに複数回到着する (図 6)。

センサは、これをもってループが存在すると判定し、検査パケットのループを止める目的で、発信元が検査パケットの宛先アドレス b で宛先が検査パケット送信用インタフェースのアドレス a となっているループ解消パケットを送信する (図 7)。

MAC アドレス b を学習したスイッチは、ループする検査パケットをセンサに送信する (図 8) ため、ループする検査パケットは最終的にネットワークからなくなる (図 9)。

L2 ネットワークが木構造となっており、その一部に閉区間が作成され、センサ VM がループを構成するスイッチに接続されている状況 (図 10) では、ループする検査パケットが閉区間の外に出た場合、木構造の末端に到着するとそのまま配送が止まり、戻ることはない。また、検査パケットのループが検出されると、センサ VM がスイッチ x

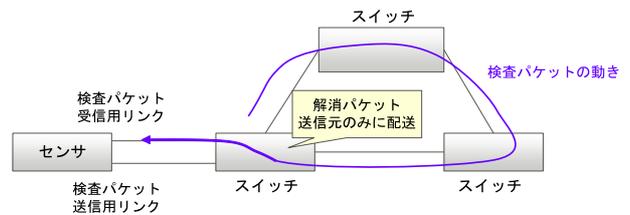


図 8 ループ解消パケット送信後の検査パケットの流れ
Fig. 8 Flow of test packet.

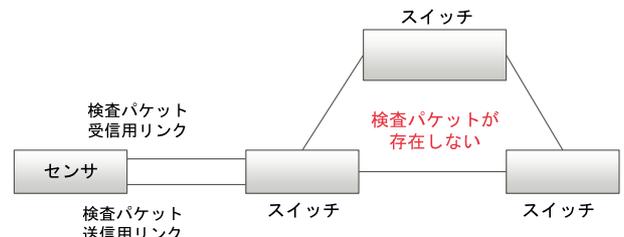


図 9 一定時間経過後のネットワーク
Fig. 9 Loop network after sending clear packet.

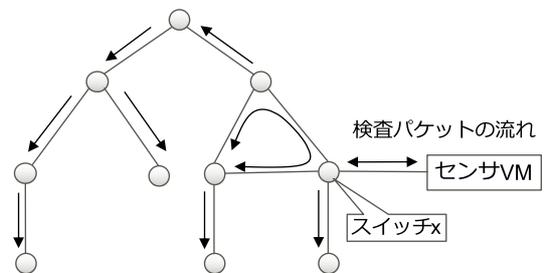


図 10 ネットワークトポロジ A
Fig. 10 Type A network topology.

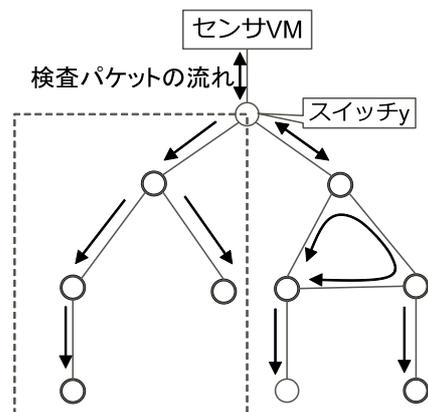


図 11 ネットワークトポロジ B
Fig. 11 Type B network topology.

に MAC アドレスを学習させる。これにより、検査パケットのループが止まることから、ネットワーク全体から検査パケットがなくなる。

一方、センサ VM が閉区間の外側に存在した場合 (図 11)、検査パケットのアドレスが学習されるのは、スイッチ y であるため、検査パケットはループし続ける。そのため、図 11 の点線内の範囲のみ検査パケットの配送が止まり、

それ以外の範囲のネットワークでは検査パケットが流れ続ける。スイッチyに限定すると、ループ向けのリンクとセンサVMが接続しているリンクには検査パケットが流れ続けるが、それ以外のリンクでは検査パケットの配送は止まる。この状況をセンサVMから見ると、検査パケットが届き続けるため、自分が接続しているスイッチ以外の場所でループが発生していることが分かる。本研究が想定する適用先では、クラウド上でサービスを実行しているVMが接続している仮想スイッチや、顧客に提供している仮想L2ネットワークを構成するスイッチにセンサVMを配置しておくことで、業務上重要な部分は守ることができる。

5. 評価

本評価では、最初にループが存在するネットワークで提案手法を実施した場合、検査パケットがループを一定時間ループすることから、検査パケットがループする時間とその期間に消費する帯域を評価し、実用上問題ないことを確認する。次に、本研究で想定する環境ではループ発生から回線がブロードキャストパケットやマルチキャストパケットで埋まるまでの時間をシミュレーションで求め、自動的なループの切断の代わりに管理者への通知でも実用上意味があることを示す。

5.1 プロトタイプ

図12は評価用に試作したセンサの実装であるが、ループ検出センサはPythonとextended Berkeley Packet Filter (eBPF) [9]で作成し、検査パケット送信モジュールとループ解消パケットモジュールはPythonのパケット操作ライブラリであるscapy [10]を用いて開発した。このプロトタイプにおいて、検査パケット、ループ解消パケットともに、arpを模擬したフォーマットのパケット (MACフレーム

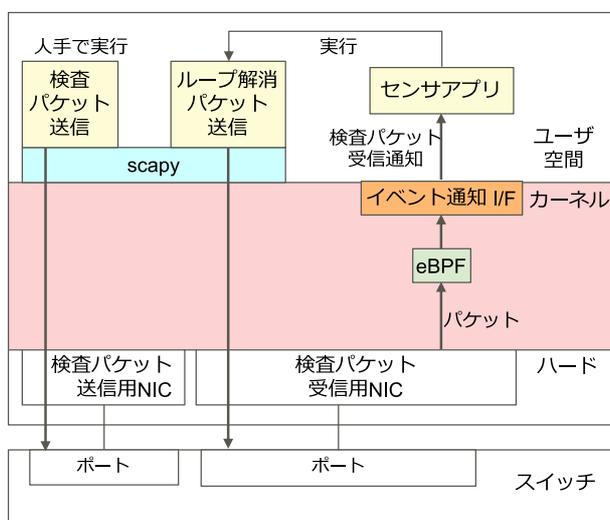


図12 センサ用計算機のシステム構成
Fig. 12 Architecture of loop detection sensor.

のサイズ 60 Bytes) を用いている。なお、検査対象のネットワークに接続した2つのインタフェースは、レイヤ3以上のプロトコルに関する設定 (IP アドレスの付与など) は行っていない。

なお、図12のプロトタイプは、評価用の実装であるため、ループ検査用のパケットは、評価の内容に合わせて評価の作業者が発信する。これに対して、実用に供するための実装では、ユーザが設定した周期で自動的に送信する仕組みとする必要がある。この周期を短くするとループが作成された場合の検出が早くなるものの、4章で説明したように、センサが接続されていない部分でループが発生した場合にループする検査パケットが蓄積していく。以上の理由から、実用的には数分間隔で検査パケットを送信し、ループ検査を行うのが妥当な利用シーンが多い。

図13が評価用ネットワークであり、すべてのリンクや機器のネットワークインタフェースの速度は1G (bps) である。ループ検出センサを搭載した仮想計算機とスイッチの代わりとなるブリッジを構成するLinux VMのスペックは表1のとおりである。また、図13のネットワークは、表2の環境で動作するGraphical Network Simulator-3 (GNS3) [11]の上に構築した。

本評価では、検査パケットがループ開始から消えるまでの時間を測定するため、ループ上の1つの経路をLinuxで構成したブリッジとしている。このブリッジのインタフェー

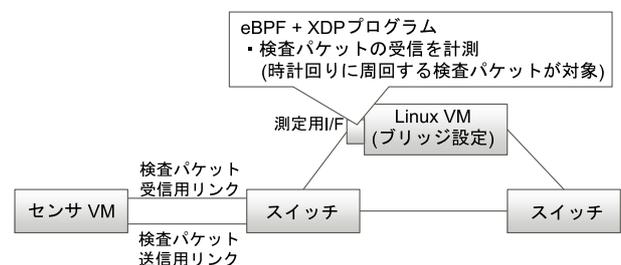


図13 評価環境
Fig. 13 Performance evaluation environment.

表1 各VMの仕様
Table 1 Specification of virtual machines.

VM	CPU数	メモリ	OS
センサ	4	8 G	Ubuntu20.04
ブリッジ用 Linux	2	8 G	Ubuntu20.04

表2 評価用マシンスペック
Table 2 Machine spec of evaluation environment.

項目	値など
CPU数	1
コア数	6
仮想CPU数	12
メモリ容量	32 G
OS	Windows 10 pro (20H2)

スの1つで eXpress Data Path (XDP) [12], [13], [14] を用いた観測用プログラムを動作させ、ループするパケットを監視する。なお、図 13 の環境では検査パケットが時計回りと反時計回りの両方向でループするが、XDP は受信パケットしか監視できないため、観測用プログラムは時計回りのみを取り扱う。また、パケットがループする時間を変化させるため、ブリッジを構成する Linux は tc コマンドを用いて、配送に遅延を追加する。

検査パケットの宛先アドレスがどこかのスイッチの MAC テーブルに学習済みとなっていなければ、同じ検査パケットを再利用可能であるため、ループが発生していない状況では同じアドレスの検査パケットを何度も利用可能である。これに対して、ループが検出され、ループ解消パケットが送信されると、センサ VM が直接つながっているスイッチに検査パケットの送信先アドレス（ループ解消パケットの送信元アドレス）が学習されてしまうため、スイッチの MAC アドレス学習テーブルの寿命（多くの機種で 300 秒）を過ぎるまでは同じアドレスを利用することができない。そのため、その期間はループの検査を止めるか、検査用の MAC アドレスをプールしておき、ローテーションで MAC アドレス学習テーブルの寿命より長い時間同じアドレスが使われないように実装で調整する必要がある。

5.2 ループする検査パケットが消費する帯域

本提案手法では、一定時間検査パケットネットワークをループする際に消費する帯域を求める。

フレームのサイズが P (bit) のパケットが τ (秒) でループする場合、1 秒でループする回数は $1/\tau$ となる。すると、このパケットで消費される帯域 B (bps) は、以下の式 (1) で計算できる。

$$B = P \times \frac{1}{\tau} \text{(bps)} \quad (1)$$

図 13 のネットワークにおいて、ブリッジとなっている Linux マシンの遅延を 0 に設定し、検査パケットを止めずにループさせた場合、ループする時間の平均値 τ は約 0.38 ms (0.00038 秒) であった。

検査パケットは 60 Bytes (480 bit) であるため、ループする検査パケットが消費する帯域 B とループ時間 τ の関係を示す式 (1) をグラフにすると、図 14 となる。 $\tau = 0.3$ ms (0.0003 秒) の場合でループするパケットが消費する帯域は 1.6 Mbps, $\tau = 500$ ms (0.5 秒) では 1 kbps 未満となり、検査パケットがループした場合に消費される帯域は非常に少なく、同じネットワークに同居する利用者に問題となる量ではない。

今回の評価では、ループするパケットのループ時間は最大で 500 ms を想定しており、これは有線回線でパケットが地球を 1 周するより長い時間であるため、通常のネットワーク構成ではループ時間としては十分な時間である。

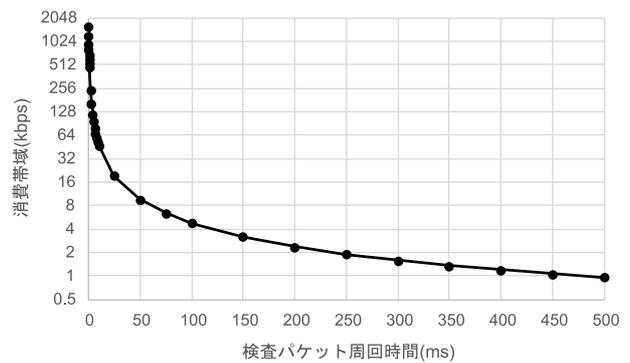


図 14 検査パケットの消費帯域

Fig. 14 Bandwidth consumed by rounding test packet.

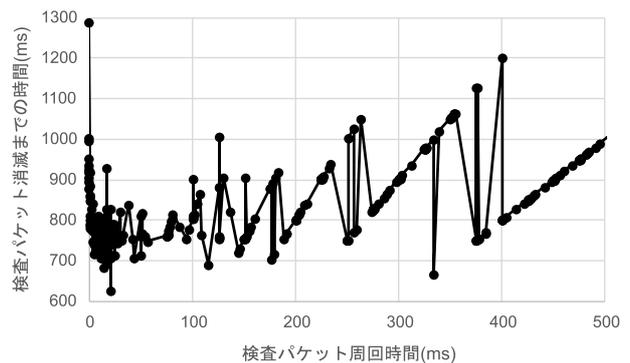


図 15 検査パケットが消えるまでの時間

Fig. 15 Duration until to clear test packet loop.

5.3 ループする検査パケットがなくなるまでの所要時間

図 15 は、図 13 における観測用のネットワークインタフェースを通過した検査パケットの通過時刻から計算したループ時間の平均値（ミリ秒単位）を横軸に、最初の検査パケットが同インタフェースで受信された時刻と最後に検査パケットを受信した時刻の差分（検査パケットのループ終了までの経過時間）をミリ秒単位で計算したものを縦軸にしたものとなっている。図 15 のグラフから分かるように、本評価で用いた実装では、1.3 秒未満の時間でループする検査パケットをネットワークからなくすることができている。

ループする検査パケットがネットワークからなくなる時刻は、ループ解消パケットの送信処理が行われ、スイッチにアドレスが学習された後、その時点でループしているパケットがアドレス学習済みのスイッチに届く時間が加わる。そのため、MAC アドレスの学習が行われたタイミングがループする検査パケット通過直前の場合と、通過直後の場合でループ 1 周分の時間の差がでる。

そのため、ループ一周の時間が長くなると、ループする検査パケットが観測されなくなるまでの最短時間と最長の時間の差分は大きくなる（図 15 で検査パケットループ時間が 400 ms 以下の範囲）。また、ループ 1 周の時間が非常に長くなると（図 15 のグラフで 400 ms より大きい範囲）、ループする検査パケットを検出してから、検査パケットが

戻ってくるまでの時間的余裕が大きくなる。スイッチに MAC アドレスを学習させるために必要な時間と、検査パケットが1周して戻ってくる時間を比べた場合に、学習に必要な時間の方が短くなることから、検査パケットの追加の1周が発生せず、リニアな特性となっている。ループ時間が長くなると、検査パケットが観測されなくなるまでの時間も長くなるが、5.2 節で述べたように、ループ時間の最大値を 500 ms と想定していることから、想定するループ時間の範囲内では最大 1.3 秒と判断した。一方、検査パケットのループ時間が非常に短い場合（横軸が 0 付近）、検査パケットが最後に観測されるまでの時間が非常に長くなっている。これは、大量の検査パケットがセンサ VM に押し寄せるため VM の負荷が上昇し、ループ解消パケットの送信処理に時間がかかるためである。

以上のことから、センサ VM やセンサ VM が接続されている仮想スイッチを実行する計算機環境が資源不足で、センサ VM が過負荷で止まるなど、パケットロスが多発する環境でもない限り、本評価の結果と大きな差は発生しない。

ネットワークに輻輳が発生した場合や、パケット廃棄率が大きくなった場合でも、エンドユーザが大きな影響を受けるのは、コネクションが切断される場合、ルータ間のルーティング情報の交換に失敗するような場合である。TCP のコネクションタイムアウトやルーティング情報交換の間隔は 30 秒であることが多いため、1.3 秒であればそれほど大きな問題にはならない。

5.4 ループ発生から大事故までの所要時間

L2 ネットワークのループを自動的に切断しない場合、ループした回線に各種機器が発生させたブロードキャストパケットとマルチキャストパケットがループを続け、最終的には回線がループするパケットで満たされる。そのため、本提案手法を利用する環境では、ネットワーク管理者にループが発生したネットワークを早期に通知し、調査と修復に用いる時間をかせぐ必要がある。

本研究で想定する環境でループ発生から帯域が満杯になるまでの所要時間がどの程度になるかを実際に稼働しているネットワークの観測データに基づいたシミュレーションによって明確にする。

5.4.1 ループするパケットで回線が埋まるまでの時間

まず、ループが発生してからどの程度の時間で回線がループするブロードキャストやマルチキャストパケットで満たされるかを式で表す。

表 3 は計算に必要なパラメータを表している。このパラメータを用いると、 t 秒間で発生するブロードキャスト/マルチキャストパケットがループで消費する帯域 $B(t)$ (bps) とループするブロードキャストパケットやマルチキャストパケットが消費する帯域と回線の帯域の比率の r は次の式で定義することができる。

表 3 パラメータの定義

Table 3 Definition of parameters.

項目	変数
回線の帯域 (bps)	W
回線利用率 (0~1)	u
ブロードキャスト/マルチキャストパケットの比率 (0~1)	b
ループ一周の時間 (s)	τ
ループするパケットが消費する帯域と回線帯域の比 (0~1)	r

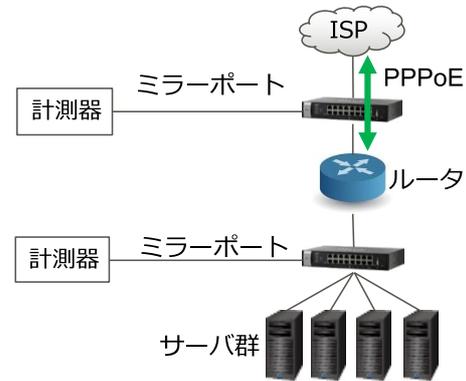


図 16 データ取得環境

Fig. 16 System configuration of measurement.

$$B(t) = \frac{buWt}{\tau} \tag{2}$$

$$r = \frac{B(t)}{W} = \frac{but}{\tau} \tag{3}$$

ここで、回線にループが発生して、ブロードキャストやマルチキャストパケットで回線が埋め尽くされる ($r = 1$ となる) 時刻 t' は次の式で求めることができる。

$$t' = \frac{\tau}{bu} \tag{4}$$

この t' はループを発見できた場合に、回線切断などを行わなくてもよい時間となる。この値が、人によるループの調査や修復に必要なと思われる時間より長ければよい。本評価における b は本研究の想定するユーザ環境に近い構成の実際のネットワークで測定した結果を用いる。

5.4.2 b の測定

回線がループするパケットで満たされるまでの所要時間は、計算機などが発生させるブロードキャスト/マルチキャストパケットが全トラフィックに占める割合 b の値が必要であり、この値 b は環境によって異なる。一方、1 章および 2 章で述べたように、提案手法の想定ユーザおよび適用先のネットワークは外部ユーザにサービスを提供する VM が接続する仮想 L2 ネットワークや、組織のある拠点のネットワーク出口のルータ間の相互接続用の L2 であるため、著者らの所属する組織でそれに近い環境 (図 16) である、通信事業者と組織内のネットワークの接続点のルータの通信、その配下に接続されているサービス実行用のサーバ群が接続されている L2 ネットワークについて測定を行った。

本研究では、広域 LAN サービスに接続する LAN 機器

表 4 ブロードキャストパケットが消費する帯域の比率

Table 4 Bandwidth ratio of broadcast/multicast packets.

測定場所	比率 b
ルータ, ISP 間	0
サーバ	5.0×10^{-8}

やクラウドとオンプレミスのクラウド環境間を VPN で接続する環境をループが発生する場所として想定しているため、図 16 のルータと ISP との間の回線、サーバが接続されているサブネットの 2 カ所で 4 週間ずつ測定した。

表 4 は測定結果である。ISP との接続回線については、PPPoE で運用されていることもあり、ブロードキャスト/マルチキャストパケットは非常に少ないことが予想されたが、4 週間の間送受信されたブロードキャストやマルチキャストパケットは 0 であった。また、サーバ側のネットワークはサーバしかつながっておらず、各サーバは連続運転が前提となっているため、こちらも非常に低い値であった。

5.4.3 ループによる帯域枯渇までの時間

まず、広域 LAN サービスのリーフとなる組織の出口ルータを模擬するルータではブロードキャストやマルチキャストパケットが 4 週間で 0 であったため、ループが発生してもループで回線がパンクするために必要な所要時間は無限大である。ただし、実際のネットワーク環境では一定時間通信が発生しない場合や、LAN 機器などが再起動した直後は ARP パケットなどが飛ぶため、きわめて低い確率でブロードキャストパケットやマルチキャストパケットが発生するだけにとどまる。そのため、ループを検出した場合に人が調査する時間は十分に存在する。

次に、クラウドを想定したサーバ環境は、測定値を用いてシミュレーションを行う。帯域が枯渇するまでの時間 t のパラメータはループが発生していない状態での回線利用率 u と、ループが発生した場合にパケットがループするために必要な所要時間 τ の 2 つである。

この 2 つのパラメータのうち、ループ時間 τ を横軸に、回線がループするパケットで一杯になるまでの時間を縦軸にとり、 u については代表的な値として 25%, 50%, 75%, 100% についてグラフ化したものが図 17 である。

この図から分かるように、回線利用率が高いほどブロードキャストやマルチキャストを装置が送信する量が多くなるため、回線が埋まるまでの時間が短くなるが、回線利用率が 100% であっても、ループ時間 τ が 1.5 ms 以上で回線容量枯渇までに 8 時間以上かかる。

広域網を介して複数拠点のサーバを VPN 接続した場合、ユーザが誤って作成したリンクをパケットがループする時間が 1.5 ms 未満となる確率は非常に低いと考えられるため、一般的なユースケースではループを発見した場合に人が調査・修復する時間は十分ある。

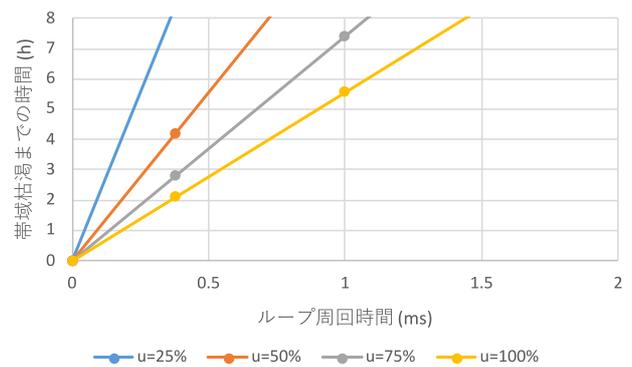


図 17 帯域枯渇までの時間

Fig. 17 Time period until bandwidth exhausting.

6. まとめ

本論文では、クラウド環境における仮想ネットワークや広域ネットワークに計算機もしくは仮想計算機を取り付け、ネットワークに発生したループを検出するのに適した仕組みを提案し、その性能を評価した。本提案方式では、ループが存在するネットワークでは一定時間検査パケットがネットワークをループしてしまうが、想定するネットワーク環境では 1.3 秒未満で検査パケットは消え、ループする検査パケットが消費する帯域も最大でも 1.6 Mbps であり、実用上問題がない範囲である。

また、ループが発生時に自動で切断せず、管理者が人手で調査・復旧を行うことを想定した場合、その時間は十分にすることも確認することができた。

参考文献

- [1] IEEE: IEEE Standard for Local and metropolitan area networks: Media Access Control (MAC) Bridges, IEEE Std 802.1D-2004 (Revision of IEEE Std 802.1D-1998), pp.1–281 (online), DOI: 10.1109/IEEESTD.2004.94569 (2004).
- [2] Mahalingam, M., Dutt, D., Duda, K., Agarwal, P., Kreeger, L., Sridhar, T., Bursell, M. and Wright, C.: Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks, RFC 7348 (2014).
- [3] IEEE: IEEE Standard for Local and Metropolitan Area Networks – Virtual Bridged Local Area Networks – Amendment 4: Provider Bridges, IEEE Std 802.1ad-2005 (Amendment to IEEE Std 802.1Q-2005), pp.1–74 (online), DOI: 10.1109/IEEESTD.2006.6044678 (2006).
- [4] 野呂正明, 高野陽介, 小口直樹, 阿部俊二: eBPF による MAC 層ループ対策, 情報処理学会研究報告, マルチメディア通信と分散処理研究会報告, Vol.2020, No.62, pp.1–7 (2020).
- [5] 野呂正明, 高野陽介, 小口直樹, 阿部俊二: 広域ネットワークで複数拠点を接続する環境での MAC 層ループ対策の評価, 第 28 回マルチメディア通信と分散処理ワークショップ論文集, pp.209–214 (2020).
- [6] 野呂正明, 高野陽介, 小口直樹, 阿部俊二: オーバレイネットワークにおける帯域消費が軽微な L2 ループ検出手法の提案, 情報処理学会研究報告, マルチメディア通信と分散処理研究会報告, Vol.2021, No.55, pp.1–6 (2020).

- [7] Tzeng, S.: LOOP DETECTION FOR A NETWORK DEVICE, US Patent 0285.499A1 (2006).
- [8] 武藤亮一, 杉谷樹一: ループフレーム検知装置およびループフレーム検知方法, 特開 2006-33275 (2006).
- [9] Gregg, B.: Performance Superpowers with Enhanced BPF, *USENIX*, USENIX Association (2017).
- [10] Biondi, P. et al.: Scapy Packet crafting for Python2 and Python3, available from (<https://scapy.net/>) (accessed 2021-01-20).
- [11] Galaxy Technologies: Graphical Network Simulator-3, available from (<https://www.gns3.com/>) (accessed 2021-01-20).
- [12] Høiland-Jørgensen, T., Brouer, J.D., Borkmann, D., Fastabend, J., Herbert, T., Ahern, D. and Miller, D.: The EXpress Data Path: Fast Programmable Packet Processing in the Operating System Kernel, *Proc. 14th International Conference on Emerging Networking EXperiments and Technologies*, Vol.CoNEXT, No.18, pp.54-66, ACM (online), DOI: 10.1145/3281411.3281443 (2018).
- [13] Choudhury, D.G.: XDP-Programmable Data Path in the Linux Kernel, ; *login.*, Vol.43, No.1 (2018).
- [14] IO-Visor: XDP eXpress Data Path, available from (<https://www.iovisor.org/technology/xdp>) (accessed 2021-01-20).



野呂 正明 (正会員)

1988年名古屋大学工学部電気工学科卒業。1990年同大学大学院工学系研究科情報工学専攻博士課程前期課程修了。同年株式会社富士通研究所入社。2002年よりTAOおよびNICTに出向、2008年大阪大学大学院情報科学研究科博士課程後期課程修了。2021年より富士通株式会社に勤務。ネットワークの品質制御、データ転送プロトコル、OS、クラウド環境の研究開発に従事。2014年情報処理学会山下記念研究賞受賞。本会シニア会員。



高野 陽介

2010年大阪大学基礎工学部システム科学科卒業。2012年同大学大学院基礎工学系研究科システム創成専攻博士課程前期課程修了。同年株式会社富士通研究所入社。仮想ネットワーク、クラウド・マイクロサービス運用の研究開発に従事。



小口 直樹 (正会員)

1992年東北大学工学部電子工学科卒業。2016年総合研究大学院大学より博士号授与。1992年株式会社富士通研究所入社。フレームリレー、ATM交換機、IPルータ、WiMAX無線通信、クラウドの研究に従事する。2021年より富士通株式会社でLinuxやクラウドOSの開発を進めている。



阿部 俊二 (正会員)

1980年3月豊橋技術科学大学工学部情報工学課程卒業。1982年3月同大学大学院修士課程修了。1996年5月博士(工学)取得(東京大学)。1982年4月(株)富士通研究所入社、1995年6月学術情報センター、2000年4月から国立情報学研究所に勤務。通信ネットワークの性能評価/性能改善方式、トラフィック解析/トラフィック制御方式等の研究開発およびSINET構築/運用/利用促進活動等に従事。