

Layer2 ネットワーク構成情報表示システムにおける STP 及び隣接機器探索プロトコルの活用について

藤田 俊輔[†] 吉田 和幸[‡]

[†]大分大学工学研究科 [‡]大分大学学術情報拠点

抄録 我々は LAN スイッチを主として構成される Layer2 ネットワークのトポロジを推測するアルゴリズムを提案し、機器間の接続関係を推測・表示するシステムを作成し利用してきた。本システムは FDB(Forwarding DataBase)を基にしてインタフェース間の接続を推測し、物理的な接続関係を実際にはケーブルを辿ることなく決定するものである。一方、LAN スイッチには FDB の他、ループ構造において実際に利用するリンク以外をブロックし、使用しているリンクに障害が起きたときにブロックしてあるリンクを有効にして冗長化を行う STP(Spanning Tree Protocol)という機能もあり、この情報から隣接機器を知ることが可能となる。そのほかにも標準規格ではないが、隣接機器探索のためのプロトコルも存在し、隣接機器情報を直接知ることができる。本論文では STP を動作させることにより得られる隣接機器の情報と、隣接機器探索プロトコルの本システムにおける有効性と問題点について述べる。

Use of STP and Neighbor Discovery Protocol for Layer2 Network Topology Visualizing System

Shunsuke FUJITA[†] Kazuyuki YOSHIDA[‡]

[†]Faculty of Engineering, Oita University

[‡]Center for Academic Information and Library Services, Oita University

Abstract We proposed the algorithm which infer Layer2 Network Topology. And we make and use the system which discovers LAN Topology using the algorithm, and displays it visually. This System infers connections between LAN interfaces based on FDB(Forwarding Database), and decides physical connection without actually tracing the cable. On the other hand, LAN Switches have the function of STP(Spanning Tree Protocol). Using STP's information enables knowing neighbor instrument. In addition, there are non Standard protocols that discover the neighbor instrument. In this paper, we describe effectiveness and the problem of STP and neighbor discovery protocol.

1. はじめに

近年、ネットワークの社会的有用性が高まり、ネットワークの肥大化・複雑化が進む中、ネットワークには可用性が求められている。ネットワークを管理し、可用性を保つ上でネットワーク構成の把握が重要となる。ネットワークの構成を正確に把握しておくことにより、ネットワークの管理者は問題に対して素早い対処と予防ができる。また、Layer2 レベル - イーサネットワーク - で問題となるブロードキャストドメインを分割するために多く利用されている

VLAN 技術を適用するためには、ネットワーク機器同士がどのインタフェースで接続しているかを知必要がある。そのため新たな VLAN を設定する際にもネットワーク構成の把握は重要となる。

しかしながら、ネットワークの物理的構成を反映する Layer2 レベルでの構成把握は困難である。我々は LAN スイッチから収集した FDB(Forwarding Database)を用いて、LAN スイッチ間の接続関係を推測するアルゴリズムを提案し、このアルゴリズムを用いて推測された

LAN スイッチ間の接続関係を表示するシステム[1-14]を作成し、利用してきた。このシステムではベンダや機種に依存する機能から得られる情報を利用することなく、ネットワーク全体の90%から95%の接続関係を決定することができる。また、LAN スイッチ間の物理的な接続を冗長化する STP(Spanning Tree Protocol)では、どの LAN スイッチのどのインタフェースと接続しているかという情報が必要となるため、STP が有効な LAN スイッチとそのインタフェースの情報をやり取りしている。この情報を得ることで隣接機器とそのインタフェースを直接決定できる。一方、LAN スイッチやルータには、どのインタフェースの先にどのような機器が直接接続されているかということを探索するための、ベンダ独自のプロトコルが実装されていることがある。このプロトコルはベンダ独自のものであり、ベンダごとに異なるため同一ベンダの機器同士でなければその効果を得ることはできないが、STP と同様に隣接機器とそのインタフェース情報を直接取得することが可能である。

現在の推測システムでは管理者による補正を行うことで、ネットワーク全体の構成をほぼ100%取得可能である。しかし、先に述べた STP や隣接機器探索プロトコルから得られる隣接機器情報を活用することで、システム全体の精度の向上と、推測速度の向上が可能となる。

本論文では、まず2章で我々が作成してきた Layer2 ネットワーク構成情報表示システム(以下、本システム)の概要について述べる。3章では現在の推測アルゴリズムとその問題点について説明する。4章では STP 及び隣接機器探索プロトコルとコルから得られる隣接機器の情報の活用について述べる。5章では STP の特徴について説明し、6章では隣接機器探索プロトコルの概要及びベンダ間の差異について述べる。7章では STP と隣接機器探索プロトコルの問題点について述べ、8章で結論としてまとめ、本システムの今後の課題について述べる。

2. Layer2 ネットワーク構成情報表示システムの概要

本章では本システムの概要について述べる。本システムは FDB やインタフェースの情報を収集する「収集部」、収集した情報を整理し接続関係を推測する「推測部」、推測された接続関係

を見やすく表示する「表示部」の3つからなる。図1に本システムの概要図を示す。

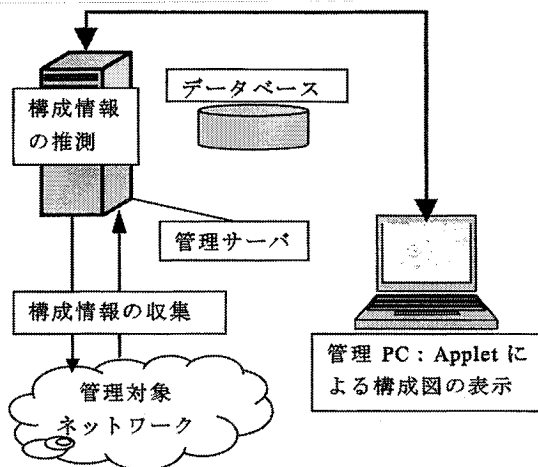


図1 ネットワーク構成表示システム概要図

2.1. 収集部

本システムではベンダ間の差異をできる限りなくし、異なるベンダの機器が混在する環境でも同様の手続きで情報を収集する必要がある。そこで収集部では管理対象であるネットワーク内に管理用のサーバを設置し、SNMP(Simple Network Management Protocol)[17]を用いてネッ

表1 現在収集している MIB 情報

ifType	インタフェースの種類
ifPhysAddress	インタフェースに対応した MAC アドレス
dot1dBaseBridgeAddress	ブリッジとしての MAC アドレス
dot1dBasePortIfIndex	ポート番号とインタフェース番号の対応表
dot1dTpFdbPort	FDB 情報
dot1qTpFdbPort	VLAN ごとの FDB 情報
dot1dStpPortDesignatedBridge	STP でのポートの代表ブリッジ
dot1dStpPortDesignatedPort	STP での代表ブリッジ側のポート
axsVBBasePortIfIndex (AlaxalA 社 privateMIB)	VLAN が設定されているポートの ifIndex 値

トワーク内の機器から構成情報の収集を行う。SNMP を用いることで LAN スイッチだけではなく、ルータやサーバなどの情報も収集可能である。表 1 は現在本システムにおいて収集している MIB 情報[18-21]である。

2.2. 推測部

収集部で得られた情報を整理し、機器同士がお互いにどのポートで接続しているかを推測し、機器間の直接の接続関係を推測する。機器間の接続を決定するために推測を必要とするのは、Layer2 ネットワークの主な構成要素である LAN スイッチには十分な経路情報が存在しないためである。詳しいアルゴリズムについては 3 章で述べる。

2.3. 表示部

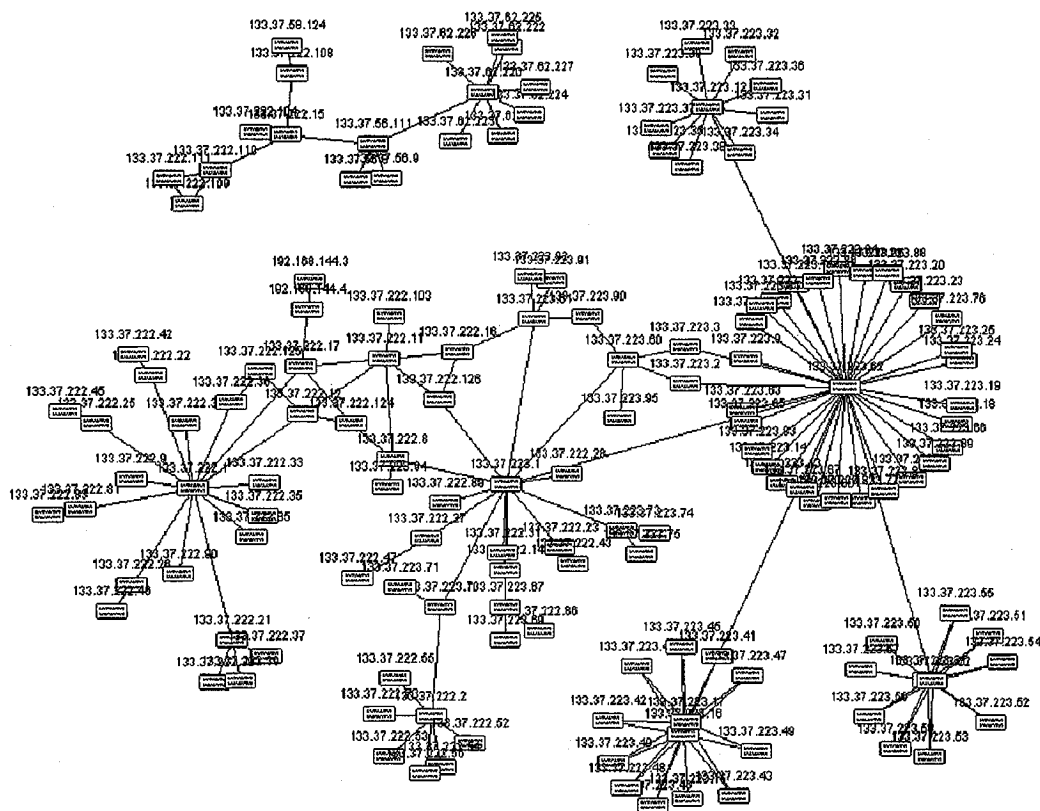
表示部では推測された接続関係を見やすく表示する。表示は機器を表すノードと、機器間

の接続関係を表すリンクからなり、リンクには接続情報としてインタフェース番号を表示することができる。また、ノードやリンクが重ならないように自動配置を行うアルゴリズムも実装している。他にもスケーリングや回転、差分表示が可能である。図 2 は構成情報表示プログラムに表示される構成図の画像である。

3. 接続の推測について

3.1. 推測アルゴリズム

本アルゴリズムではまず、機器間の「接続」について、推測を行う。ここでいう「接続」は機器のインタフェースと別の機器のインタフェースの間にネットワークが存在し、イーサフレームが到達可能であることで指す。一方、インタフェース間には一本の物理線のみが存在する場合を「直接接続」と呼ぶ。次に収集した情報を基に、ネットワーク構成を推測する手順を説明と共に示す。



① FDB はそのインタフェースの先に、FDB に格納された MAC アドレスを持つ機器が存在することを示している。つまり、インタフェースと機器の接続を示している。LAN スイッチにも MAC アドレスが存在するため、推測をせずに接続を決定できる。また、LAN スイッチによっては表 1 の ifPhysAddress の説明にあるように、インタフェースごとに異なる MAC アドレスを持つ場合があり、このときには接続先の機器のインタフェースも決定することができる。

② ①で取得できない接続については FDB の表れ方により推測を行う。機器 A の FDB において、インタフェース a に関わる MAC アドレスが、機器 B の FDB において複数のインタフェースで分散して現れるとき、機器 A のインタフェース a から機器 B への接続があると決定できる。図 3 のように、機器 A のインタフェース a に [h1, h2] が関連付けられている場合に、a から B への接続が推測される。

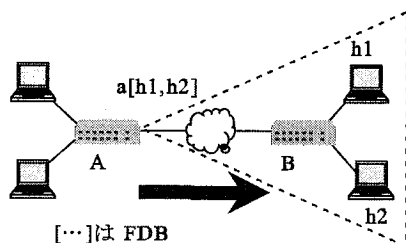


図 3 接続の推測例 1

③ ②の段階では機器 B 側のインタフェースが不明であるので、機器 B のインタフェースのうち、機器 A のインタフェース a 以外のインタフェースと同じ MAC アドレスが FDB に存在するインタフェースを機器 B 側のインタフェースとする。図 4 にこの時の接続例を示す。

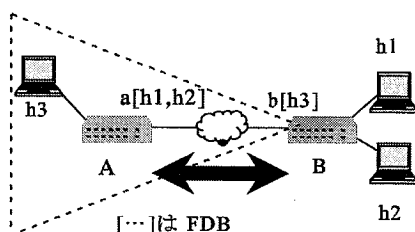


図 4 接続の推測例 2

④ 推測された接続の多くは直接接続ではない。このような直接接続でない接続を除外するために、非直接接続の推測を行う。接続関係にある 2つのインタフェース a・b が、機器 C と接続している場合、a・b の間には機器 C が存在し、a・b 間の接続は非直接接続であると言える。図 5 はこの時の接続例である。

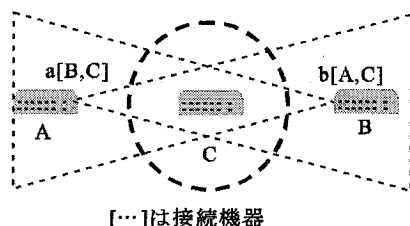


図 5 非直接接続の推測例

⑤ 非直接接続を除外し、機器 A のインタフェース a に残された接続が一つのみになった場合、これを直接接続と決定する。このとき、接続先の機器 B のインタフェースが不明でありこの機器のインタフェース b が、接続先のインタフェースは不明であるが、機器 A に接続している場合にはインタフェース a・b が直接接続していると推測する。

以上の手順のうち、手順①・②・③の接続の推測は 2つの機器の間だけで完結するため、各機器の組合せに対して一度だけ処理をすればよい。一方、手順④・⑤における直接接続の推測は、手順④において機器の間に存在するかも知れない機器が関わってくるため、直接接続が検出されなくなるまで繰り返す必要がある。

3.2. 本アルゴリズムの問題点

この推測アルゴリズムでは VLAN によるブロードキャストドメインの分割に起因するループ構造なども取得可能であり、複雑で規模の大きいイーサネットワークの構造を容易に把握できる。しかし、本アルゴリズムでは推測を主体としているため、当然接続を誤る可能性がある。また、LAN スイッチを中心としたイーサネットワークの構造は、多くの場合頻繁には変化しない。このような構造の場合、すべての接続を推測によって決定すると推測結果全体の精度の低下が生じ、低下した精度を補うため収集した構

成情報や機器の設定情報などを管理者が直接参照する必要が生じる。そこで LAN スイッチの持つ FDB 以外の機能や、ベンダ独自のプロトコルから直接接続している機器の情報を取得し、直接接続を決定する。その後、決定できない接続について推測を行うことにより精度の向上を図る。

4. 隣接機器情報の活用について

本章では本システムにおいて必要となる隣接機器の情報と、直接接続の決定を行うことにより精度が向上する理由について説明する。

本システムは 2.3 節及び 3.1 節で述べたように、ノードとしての LAN スイッチの識別子と、接続に使用されるインタフェース番号を必要とする。そこで、隣接機器の情報としてもこれらの組が必要となる。隣接機器の情報を取得可能な機能としては、LAN スイッチの多くに実装されており標準規格でもある STP と、ベンダごとに規格の異なる、隣接機器探索のためのプロトコルが存在する。それぞれのプロトコルと情報の取得方法については後の章で説明を行う。これらのプロトコルから隣接機器の情報を取得し、推測に先立ち直接接続を決定することで、決定された直接接続に使用されるインタフェースに関わる接続を、推測を必要とせず非直接接続とすることができる。これにより非直接接続とされた接続の、相手側のインタフェースの直接接続の候補が減少し、推測の精度が向上する。

5. STP の特徴について

Layer2 ネットワークでは、LAN スイッチは基本的にブロードキャストを行い、データフレームはすべてのインタフェースから送信される。そのため物理的に冗長化を行い、機器間にループが発生するとブロードキャストしたデータフレームが再び自機に戻り、それを再び送信するといったデータフレームの氾濫が起きる。これをブロードキャストストームと呼ぶ。STP では物理的な冗長化を行いながらもブロードキャストストームを回避するためのプロトコルで、木構造のルートとなる LAN スイッチを選出し、ルートからもっとも遠くにあるループの要因となる接続のインタフェースを不通状態にし、接続を論理的に遮断する。遮断されなかった接続に何らかの異常が発生した場合、不通状態にした

インタフェースを通常状態に復旧し、遮断した接続を回復することで冗長化を可能としている。STP では BPDU(Bridge Protocol Data Unit)と呼ばれる特殊なデータフレームがやりとりされる。このフレームの中にはプロトコルのバージョンやルートの ID、機器の MAC アドレスを基にしたブリッジ ID やインタフェースの ID が含まれる。STP ではブリッジ ID を基にルートを決定し、ルート側のインタフェースから BPDU を送信することで、あるインタフェースの最も近くにあるルート側の機器の MAC アドレスやインタフェースの情報を知ることができる。MIB は表 1 に示している。

6. 隣接機器探索プロトコルについて

6.1. プロトコルの概要

隣接機器探索プロトコルでは自機と送信元となるインタフェースに関する情報をデータフレームにまとめ、送信する。このデータフレームを受信したインタフェースは、データフレームを送信した機器と隣接していることを知ることができる。図 6 に概要図を示す。

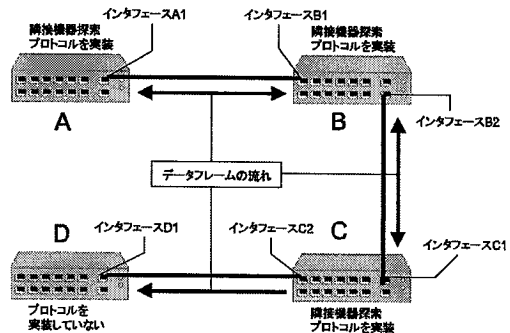


図 6 隣接機器探索プロトコル概要図

この図では、機器 A のインタフェース a1 には機器 B のインタフェース b1 から送信された b1 に関わる情報が保存される。一方、a1 から送信されたデータフレームは b1 で受信・保存された後、他のインタフェースからのフラグディングを行うことなく破棄されるため、機器 C は機器 A の情報を知り得ない。同様に機器 B・C 間でもデータフレームのやり取りが行われるが、実際に接続関係にあるインタフェース以外がこれらのデータフレームを取得することはない。機器 C・D 間であるが、機器 D が対応するプロトコルを実装していないため、インタフェース

c1 から送信されたデータフレームはインタフェース d1 では処理されず、機器 D に接続している他の機器に向けてフラディングされる。

6.2. 隣接機器情報の取得方法

このプロトコルは STP のような標準規格ではなくベンダごとに実装が異なるため、同じプロトコルを実装した機器同士が隣接している必要がある。また、隣接機器情報の収集方法についてもベンダごとのプライベート MIB を参照し、推測部で同様に処理できるように統一した形式に変換する必要がある。ここでは隣接機器探索プロトコルの例として、大分大学学内 LAN で運用している機器が実装している CDP(Cisco Discovery Protocol)、EDP(Extreme Discovery Protocol)、OADP (Octpower Auto Discovery Protocol) の 3 つのプロトコルと、これらのプロトコルの情報を参照するための MIB について説明する。CDP, EDP, OADP はそれぞれ Cisco, Extreme Networks, NEC の機器で用いられるプロトコルである。なお本学では NEC の機器ではなく、OADP を実装している AlaxalA の機器を運用している。

それぞれのプロトコルでは、何度も述べるが隣接機器を特定するための情報やインタフェースの情報を取得することができる。しかし、これらの情報の表現はプロトコルにより異なる。それぞれの表現方法について Cisco, Extreme Networks, AlaxalA の機器における SNMP で参照すべきプライベート MIB と交えて表 2, 表 3, 表 4 に示す。それぞれ OID の先頭の 1.3.6.1.4.1 (enterprises) は省略している。また、それぞれの OID は機器自身のインタフェース番号をインデックスに持つテーブルである。

表 2 Cisco プライベート MIB

cdpCacheAddress 9.9.23.1.2.1.1.4	隣接機器の Layer3 アドレス
cdpCacheDevicePort 9.9.23.1.2.1.1.7	隣接機器の送信インタフェースの ifDescr の値

表 3 Extreme Networks プライベート MIB

1916.1.13.2.1.3	隣接機器のシステム名
1916.1.13.2.1.5	隣接機器のインタフェースの スロット番号
1916.1.13.2.1.6	隣接機器のインタフェース番号

表 4 AlaxalA プライベート MIB

axsOadpNeighbor SNMPAgentAddress 21839.2.2.1.7.1.3.1.1.7	隣接機器の IPv4 または IPv6 アドレス
axsOadpNeighbor IfIndex 21839.2.2.1.7.1.3.1.1.11	隣接機器の送信インタフェースの ifIndex の値

Cisco では cdpCacheAddress の値として Hex-STRING が用いられ、値としてネットワーク層のアドレス (IP アドレス) が格納される。本システムは機器を表すために IP アドレスを用いているため、この情報はそのまま利用可能である。また、cdpCacheDevicePort では隣接機器側のインタフェース情報として、ifDescr が用いられている。これは機器のインタフェースを文字列で表現したものであり、実際のインタフェース番号を取得するため ifDescr を参照し、対応するインタフェース番号に変換する必要がある。

次に Extreme Networks では MIB の値としては機器を表す正確な情報を取得することができない。その代りに OID の末端 6 桁が隣接機器の MAC アドレスとなっており、MIB の値と容易に関連付けることができる。インタフェースの表現は、比較的大きな LAN スイッチなどではインタフェースをスロット番号とスロット内ポート番号とで管理している。EDP ではそのためスロットを特定し、そのスロットの何番のインタフェースであるかをインタフェース情報の一部としている。スロット番号とスロット内のインタフェース番号から、ifIndex の値に変換するため、ここでも ifDescr を参照する。スロットごとに管理されている機器では、ifDescr の最後に「スロット番号:スロット内インタフェース番号」の情報があり、これと対応させることでイン

タフェース番号に変換する。

AlaxalA では axsOadpNeighborSNMPAgentAddress に隣接機器の IP アドレスが文字列として格納されている。これは Hex-STRING で格納されていた Cisco とは異なる。axsOadpNeighborIndex から得られる値は ifIndex で得られるインタフェース番号そのものであるため、特に変換を必要としない。そのほかに隣接機器のブリッジアドレスも取得可能であるが、IP アドレスにより機器の特定できるため、収集する必要はない。

7. STP 及び隣接機器探索プロトコルを利用する上での問題点

STP 及び隣接機器探索プロトコルは共に Layer2 レベルで動作するため、ルータなどを通して行うことはない。逆に、どちらのプロトコルも機器がプロトコルに対応していなければそのまま透過するため、実際には隣接関係のない機器の情報を取得する可能性がある。そこでこれらの機能を利用する際には、構成推測の手順③の直後、つまり機器間の接続関係が分かった時点で、これらのプロトコルから得られた情報が正確であるかどうか、手順④の非直接接続の推測により判断する。

また、STP はその特徴からネットワークの構造が変化するたびに再計算を行い、場合によってはネットワークを一時不通にする恐れがある。そのためより高速にスパニングツリーを構築できるように RSTP(Rapid STP)というプロトコルが存在するが、対応していない機器も多く、必ずしも有用ではない。

今回挙げた隣接機器探索プロトコルはベンダ独自の規格であり、実際に効率よく運用するためには同じベンダの機器を用いてネットワークを構築する必要がある。しかし 2005 年 4 月、異なるベンダ間でも隣接機器探索プロトコルが利用できるように IEEE802.1AB[22](Station and Media Access Control Connectivity Discovery)として LLDP(Link Layer Discovery Protocol)が標準化された。今回は対象が大分大学学内 LAN であり、LLDP を実装していない機器のほうが多かったため直接触れてはいないが、このプロトコルを用いることで SNMP による管理のできない IP 電話機などとの隣接関係も取得できるため、より広範な構成情報を取得できるようになる。

8. 終わりに

Layer2 ネットワーク構成情報表示システムにおける構成情報推測アルゴリズムの補助として、STP と隣接機器探索プロトコルから得られる隣接機器情報を利用した際の精度向上と、それぞれのプロトコルについて述べた。すでに構築されている Layer2 レベルのネットワークは一度にすべての機器を交換するといったことが難しい。そのため LLDP のような比較的新しいプロトコルをすべての機器が実装しているとは限らない。このような新旧の機器が混在するようなネットワークの構成を正確に把握するため、推測アルゴリズム自体の改良が必要となる。また、現在は Layer2 で動作する LAN スイッチと Layer3 以上で動作するルータやサーバを区別しておらず、すべて同等の機器として扱っている。今後はこれらを区別し、推測が必要な接続を LAN スイッチ間のみに絞ることで、推測精度が向上するか検証する必要がある。

参考文献

- [1] 藤田俊輔, 飯田隆義, 兒玉清幸, 吉田和幸, Layer2 ネットワーク構成情報推測・表示システムにおける VLAN 情報の収集と表示の提案, “マルチメディア, 分散, 協調とモバイル(DICOMO 2008)シンポジウム”, pp.1190-1197, 2008
- [2] 飯田隆義, 兒玉清幸, 有田敏充, 藤田俊輔, 吉田和幸, “ネットワーク構成情報表示システムのための自動配置アルゴリズムの改良と評価”, “マルチメディア, 分散, 協調とモバイル(DICOMO 2008)シンポジウム”, pp.11908-1205, 2008
- [3] 藤田俊輔, “Layer2 ネットワーク構成情報推測システムの改良について”, “情報処理学会火の国情報シンポジウム 2008 論文集”, C-2-5, pp.1-7(published by CD-ROM), 2008
- [4] 飯田隆義, “ネットワーク構成情報表示システムのための自動配置アルゴリズムの改良について”, “情報処理学会火の国情報シンポジウム 2008 論文集”, C-6-1, pp.1-8(published by CD-ROM), 2008
- [5] 飯田隆義, “ネットワーク構成情報表示システムのための自動配置アルゴリズムの改良について”, “大分大学学士論文”, 2008.
- [6] 吉田和幸, “ネットワークトポロジの

発見とその表示について”，“大学情報システム環境研究”，VOL.10，pp.68-74，2007.

[7] 近藤純平，“Layer2 ネットワーク構成の推測システムについて—情報収集できない LAN スイッチの存在推測—”，“大分大学学士論文”，2007.

[8] 河野優，“Layer2 ネットワーク構成の推測システムについて”，“大分大学修士論文”，2007.

[9] 兒玉清幸，釜崎正吾，吉田和幸，“ネットワーク構成情報表示システムのための自動配置アルゴリズムの評価”，“情報処理学会マルチメディア，分散，協調とモバイル(DICOMO 2007)論文集”，pp.1754-1761，2007.

[10] 兒玉清幸，加来徹，釜崎正吾，吉田和幸，“ネットワーク構成情報表示システム--自動配置アルゴリズムの改善--”，“情報処理学会火の国情報シンポジウム 2007 論文集”，C-8-3，pp.1-8(published by CD-ROM)，2007.

[11] 河野優，釜崎正吾，大浦昇，吉田和幸，“ループを考慮した Layer2 ネットワーク構成情報の推測アルゴリズムについて”，“分散システム/インターネット運用技術シンポジウム 2006 論文集”，pp.7-12，2006.

[12] 大浦昇，河野優，釜崎正吾，吉田和幸，“VLAN を考慮した Layer2 ネットワーク構成情報推測アルゴリズムについて”，“マルチメディア，分散，協調とモバイル(DICOMO2006)論文集”，pp.629-632，2006.

[13] 河野優，釜崎正吾，平川龍，大浦昇，吉田和幸，“Layer2 ネットワーク構成情報の推測アルゴリズムの改良について”，“分散システム/インターネット運用技術シンポジウム 2005 論文集”，pp.61-66，2005.

[14] 河野優，釜崎正吾，吉田和幸，“Layer2 におけるネットワーク構成図表示システム”，“マルチメディア，分散，協調とモバイル(DICOMO2005)論文集”，pp. 757-760，2005.

[15] B.Lowe kamp, D.R.O'Hallaron, T.R.Gross, “Topology Discovery for Large Ethernet Networks”, “Procs. of Applications, Technologies, Architectures, and Protocols for Computer Communications”, pp.27-31, 2001.

[16] Y.Breit bart, M.Garofalakis, C.Martin, R.Rastogi, S.Seshadri, A.Silberschatz, “Topology discovery in heterogeneous IP networks”, “Procs. of INFOCOM2000”, 2000.

[17] J.Case, M.Fedor, M.Schoffstall, J.Davin, “A Simple Network Management Protocol”, RFC 1157, <http://www.ietf.org>, 1990.

[18] K.McCloghrie, “Management Information Base for Network Management of TCP/IP-based internets”, RFC 1156, <http://www.ietf.org>, 1990.

[19] K.McCloghrie, “Management Information Base for Network Management of TCP/IP-based internets:MIB-2”, RFC 1213, 1991.

[20] E. Decker, P. Langille, A. Rijsinghani, K. McCloghrie, “Definitions of Managed Objects for Bridges”, RFC 1493, <http://www.ietf.org>, 1991.

[21] D.Levi, D.Harrington, “Definitions of Managed Objects for Bridges with Traffic Classes, Multicast Filtering and Virtual LAN Extensions”, RFC 4363, <http://www.ietf.org>, 2006.
“IEEE Standard for Local and metropolitan area networks Station and Media Access Control Connectivity Discovery”, 802.1AB-2005