

## VLAN イーサネットを用いた PC クラスタ向け 大規模ネットワーク構築法

大塚 智 宏<sup>†1</sup> 鯉 淵 道 紘<sup>†2,†3</sup>  
工 藤 知 宏<sup>†4</sup> 天 野 英 晴<sup>†1</sup>

本稿では、イーサネットを用いて大規模 PC クラスタのネットワークを構築するための手法である VLAN リネーミングを提案する。VLAN リネーミングは、既存のイーサネットスイッチの VLAN 機能を制御することで実現でき、1) 必要となる VLAN 数がトポロジの次数以下と少数で、2) システムソフトウェアが VLAN に対応していない場合にも使用可能、という 2 つの特徴を持つ。商用スイッチを用いた評価の結果、VLAN リネーミングは導入によるオーバーヘッドがほとんどなく、ネットワーク資源を効率良く使用した大規模クラスタネットワークの構築に適していることが分かった。

### A Methodology for Building Large-scale Interconnects for PC Clusters Using VLAN Ethernet

TOMOHIRO OTSUKA,<sup>†1</sup> MICHIMIRO KOIBUCHI,<sup>†2,†3</sup>  
TOMOHIRO KUDOH<sup>†4</sup> and HIDEHARU AMANO<sup>†1</sup>

This paper presents a technique called VLAN renaming in order to build interconnects of large-scale PC clusters using Ethernet. The VLAN renaming is managed by control of existing VLAN functions of Ethernet switches. It has following advantages of (1) using the smaller number of VLANs than the topology degree, (2) can be used with system software that accepts no VLAN tags. Evaluation results using a commercial switch show that the VLAN renaming introduces almost no bandwidth degradation and overhead of latency. The VLAN renaming is thus suitable for building interconnects of large-scale PC clusters.

### 1. はじめに

イーサネット (Ethernet) は、管理の容易さ、高い耐故障性、安価なハードウェア等の利点から、ローカルエリアネットワーク (LAN) だけでなく、広域ネットワークや PC クラスタのインタコネクタとしても幅広く採用されている。特に、Gigabit Ethernet (GbE) の普及、ツイストペアケーブルを用いる 10GBASE-T の標準化 (IEEE 802.3an-2006) 等により、イーサネットはハイパフォーマンスコンピューティング (HPC) 分野において、Myrinet 等の高価なシステムエリアネットワーク (SAN) に迫るインタコネクタとして注目を集めている。初期の Beowulf 型クラスタと違い、最近のイーサネットを用いた PC クラスタでは、クラスタ向けの SAN で採用されてきたゼロコピー通信やワンコピー通信を実現したシステムソフトウェアを利用することができ、低遅延のノード間通信が可能である。TOP500 スーパーコンピュータのサイト<sup>1)</sup> によると、2007 年 11 月のランキングにおいて、上位 500 台の中で GbE を用いたシステムが 54% と過半数になっている。

一方で、SAN とは異なり、イーサネットを用いた PC クラスタの多くは単純なツリー状トポロジを採用している。これは、基本的にイーサネットがループ構造を含むトポロジを許していないためである。ツリー状ネットワークにはトラフィックがツリーのルート付近に偏りやすいという欠点があるため、リンク集約化 (IEEE 802.3ad) 等によってルート付近のリンクを強化するのが一般的である。しかし、クラスタが大規模になると、リンク集約化だけではツリー状ネットワークの欠点を補いきれなくなる。また、リンク集約化のためにスイッチのポートを多数占有してしまうため、次数が低く直径の大きいトポロジしか構築できないという問題も生じる。これらのことから、ユーザやアプリケーションの要求に応じたトポロジ・ルーティングを採用している SAN や並列計算機の相互結合網に比べて、イーサネットを用いた PC クラスタは大規模化には向かないとされてきた。

そこで、リンク集約化以外にも、スイッチ間に複数リンクを接続することでバンド幅を向

†1 慶應義塾大学理工学部

Faculty of Science and Technology, Keio University

†2 国立情報学研究所

National Institute of Informatics

†3 総合研究大学院大学

The Graduate University of Advanced Studies (Sokendai)

†4 産業技術総合研究所

National Institute of Advanced Industrial Science and Technology (AIST)

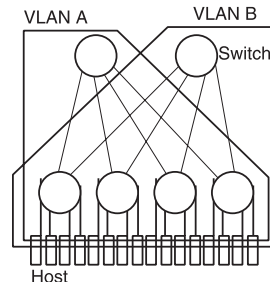


図 1 VLAN ルーティング法  
Fig. 1 VLAN-based routing method.

上させる方法が提案されている<sup>2),3)</sup>。これらの方法では、ループを解消してツリー構造を維持するスパニングツリープロトコル (STP, IEEE 802.1D) を用いないため、Fat ツリーやトラス等のループ構造を含むトポロジを構築することができる。STP を用いずに大規模クラスタシステムを構築する場合、ホストの追加やスイッチの故障、操作ミス等によるブロードキャストストームの発生を回避するために、ホストの MAC アドレスの管理が 1 つの課題となる。この点において、IEEE 802.1Q 標準のタグ VLAN 技術を応用した VLAN ルーティング法<sup>2)</sup> が既存の方法の中で有力である。

VLAN 技術は本来、同じ物理ネットワークに接続されたホストの集合を、複数の論理的なグループに分割するために用いられるが、VLAN ルーティング法ではこれをネットワークのスループット向上のために用いる。図 1 のように、各ホストが複数の VLAN グループのメンバになるようにしておき、各 VLAN にそれぞれ異なるリンク集合を割り当てる。ここで、各 VLAN ネットワークのトポロジはツリー構造となっているため、ブロードキャストストームは発生しない。このようにすることで、すべてのホストがどの VLAN を用いても互いに通信でき、VLAN を選択することで複数の経路を切り替えて使うことができるようになる。

しかし、既存の VLAN ルーティング法を適用して大規模な PC クラスタを構築するのは、1) 通信ライブラリの VLAN への対応、2) VLAN 数および MAC アドレステーブルのエントリ数の制限、の 2 つの問題により困難である。そこで本稿では、これらの問題を解決し、イーサネットを用いて大規模 PC クラスタのインタコネクタを構築するための手法として、既存の VLAN ルーティング法を改良した VLAN リネーミングを提案する。

VLAN リネーミングでは、1 つのフレームを各中継スイッチにおいて異なる VLAN ID

に乗せ換えながら転送する。各スイッチが、IEEE 802.1Q 標準の VLAN 技術を利用してそのスイッチ内の転送のみに有効な VLAN タグをフレームに割り当てることにより、スイッチの次数 (使用ポート数) 以下の VLAN 数で任意の規模のトポロジを構築することができるようになる。さらに、VLAN リネーミングは、ホストのシステムソフトウェアが VLAN に対応していない場合にも適用でき、様々なルーティングアルゴリズムを実装することができる。

以下、2 章で提案手法の詳細を述べ、3 章で評価を行う。4 章において本研究の関連研究を紹介し、最後に 5 章でまとめを述べる。

## 2. VLAN リネーミング

本章では、提案手法である VLAN リネーミングについてその詳細を述べる。まず、既存の VLAN ルーティング法の問題点を明らかにしたうえで、それを解決する手法として VLAN リネーミングを提案し、そのアルゴリズムを説明する。さらに、VLAN リネーミングを適用する際のスイッチ設定について例を示し、設計どおりに動作することを示す。最後に、VLAN リネーミングの適用範囲について述べる。

### 2.1 問題提起

既存の VLAN ルーティング法を適用するために必要となる VLAN 数は、ネットワークの規模に応じて増加するため、VLAN 数が大規模クラスタを構築する場合の制限要因となりうる。VLAN 数は有限であり、IEEE 802.1Q の規定では 4,094 ( $2^{12} - 2$ ) 個となっているが、コストパフォーマンスに優れた安価な商用イーサネットスイッチでは数十～数百個程度の VLAN のサポートにとどまるものも多い。

さらに、VLAN 数が増加するにつれ、その管理も複雑となる。特に、スイッチ内で静的または (学習により) 動的に登録されるホストの MAC アドレスは、同じアドレスであっても VLAN ごとに別々に登録されることになる。そのため、VLAN 数が増加した場合、MAC アドレステーブルのエントリが不足するという状態になりかねない。

既存の VLAN ルーティング法を適用する際のもう 1 つの問題は、ホスト側のシステムソフトウェアが VLAN に対応していないことである。PC クラスタのノード間通信に用いられる軽量通信ライブラリは、イーサネットフレームの VLAN タグ付けに対応していない場合が多く、そのままでは VLAN ルーティング法を利用することができない。TCP/IP であれば VLAN に対応している場合がほとんどであるが、TCP はノード間通信に用いるには遅延が大きく、パフォーマンス上問題となる場合が多い。

また、IP を使用する場合、VLAN ごとの仮想インタフェースにそれぞれ IP アドレスを割り振る必要があるが、MPI 等の IP 上の通信ライブラリの実装では、ホストの指定に IP アドレス（またはそれに対応するホスト名）を用いることが多く、同一のホストに複数の IP アドレス（ホスト名）がある状態を扱うのが難しいという大きな問題がある。

以上より、安価な L2 イーサネットスイッチを用いて VLAN ルーティング法による大規模 PC クラスタを構築する場合、1) 必要となる VLAN 数がネットワーク内のスイッチ数によらず一定であること、2) システムソフトウェアが VLAN 技術に対応していない場合にも使用可能であること、という 2 つの条件を満たすことが重要となる。なお、我々がすでに提案しているスイッチでタグ付けを行う手法<sup>4)</sup> は 2) の条件は満たしているが、1) については対応していない。

## 2.2 提案手法

VLAN ルーティング法において、前節で述べた 2 つの条件を満たし、数百～数千台規模の PC クラスタを構築するために、VLAN の使用をチャネル循環依存を除去する目的に限定した VLAN リネーミングを提案する。VLAN リネーミングでは、ホストの MAC アドレスを、学習ではなく各 VLAN 内で静的にスイッチに登録する。

### 2.2.1 VLAN 認識型スイッチの動作

まず、準備として、IEEE 802.1Q に準拠した VLAN スwitch の動作を説明する。フレームがスイッチのポートに入力された際、そのフレームがどの VLAN に属するかは以下のように決定される。

- VLAN タグを持たない（タグなし）フレームの場合、そのポートのデフォルトの VLAN ID（ポート VLAN ID, PVID）でタグ付けされ、PVID で示される VLAN に属すると見なされる。
- VLAN タグが付加された（タグ付き）フレームの場合、タグはそのまま維持され、タグ内の VLAN ID フィールドで示される VLAN に属すると見なされる。

スイッチは、登録されている各 VLAN について、どのポートが VLAN のメンバとなっているかを示す情報（メンバセット）を保持しており、各フレームは、属している VLAN のメンバとなっているポートにのみ転送される。ここで、VLAN のメンバ情報には以下の 3 種類がある。

- “タグ付き”（tagged）メンバ
- “タグなし”（untagged）メンバ
- 非メンバ

属する VLAN の非メンバであるポートには決してフレームは転送されないが、メンバである場合、フレームがスイッチのポートから出力される際にタグ付けされているかどうかは以下のように決定される。

- ポートが VLAN の tagged メンバである場合、出力フレームはその VLAN ID でタグ付けされる。
- ポートが VLAN の untagged メンバである場合、出力フレームはタグなしフレームとなる（タグは除去される）。

なお、通常、ポートの PVID として登録される VLAN のメンバセットにはそのポートが含まれる。すなわち、各ポートは、PVID で示される VLAN の tagged または untagged のメンバとなっている必要がある\*1。

### 2.2.2 VLAN リネーミングアルゴリズム

VLAN リネーミングでは、多くの商用スイッチが提供している IEEE 802.1Q 標準の VLAN 機能を次のように利用する。なお、この機能は本来、VLAN に対応していない機器を VLAN イーサネットに収容するために加えられたものであるが、本手法では柔軟な経路設定のために利用する。

- 各スイッチのすべての出力ポートにおいて、フレームの VLAN タグをすべて除去（untagged）する。
- スwitch の入力ポートにはつねに VLAN タグなしフレームのみが到着し、それらは PVID に基づいて出力ポートに転送される。

上記により、各スイッチ間およびホスト–スイッチ間でやりとりされるフレームはすべてタグなしフレームであるため、各スイッチの設定および動作は VLAN とは関係なく独立である。そして、VLAN リネーミングは、各スイッチにおいて以下のように VLAN ID を割り当てて実現する。

- (1) 各入力ポート  $i$  に対して PVID  $v_i$  を割り当て、VLAN  $v_i$  の untagged メンバとして設定する（図 2 (a)）。
- (2) 入力ポート  $i$  に入力されたフレームが転送される出力ポートを、VLAN  $v_i$  の un-

\*1 PVID の VLAN を非メンバとする設定も可能ではあるが、その場合、入力フレームの VLAN ID をチェックするインGRESSフィルタを無効とする必要がある。入力フレームが属する VLAN のメンバセットに入力ポートが含まれていない（非メンバである）場合、インGRESSフィルタによってフレームは破棄される。タグなしフレームは PVID の VLAN に属すると見なされるため、インGRESSフィルタを無効にしない限り入力時に破棄されてしまう。

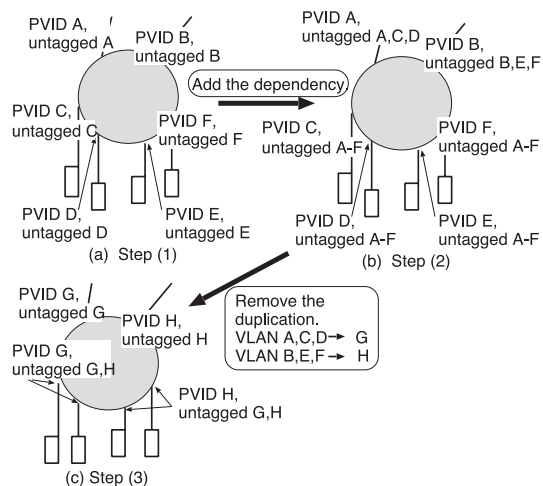


図2 VLAN リネーミングにおける VLAN ID の割当て  
Fig. 2 VLAN assignment of VLAN renaming.

tagged メンバとして設定する (図 2(b)).

- (3) まったく同一の出力ポート群 (同一の untagged メンバセット) で構成される VLAN ID 群を, VLAN  $v_j$  に統一する (図 2(c)).

Fat ツリートポロジに VLAN リネーミングを適用した例を図 3 に示す. この図において, 各スイッチの PVID A のポートに接続されたホスト (1, 2, 5, 6, 9, 10, 13, 14) から送信されたフレームはスイッチ s5 を経由して転送される. 一方, PVID B のポートに接続されたホスト (3, 4, 7, 8, 11, 12, 15, 16) から送信されたフレームはスイッチ s6 を経由して転送される.

VLAN リネーミングにより, フレームは各スイッチ内において異なる VLAN ID に乗せ換えられながら転送されることになる. 図 3 のスイッチ s1 に着目すると, たとえばホスト 1 と接続しているポート (PVID A) から入力されたフレームは, 目的地に応じて, スイッチ s5 およびホスト 2, 3, 4 に接続されたポート (いずれも VLAN A の untagged メンバ) のいずれかへと転送される. 一方, スイッチ s6 へのポートは VLAN A のメンバではないため, ホスト 1 から入力されたフレームが転送されることはない. また, ホスト 3 からホスト 11 へ送信されるフレームは, スイッチ s1, s3 では VLAN B, スイッチ s6 では VLAN

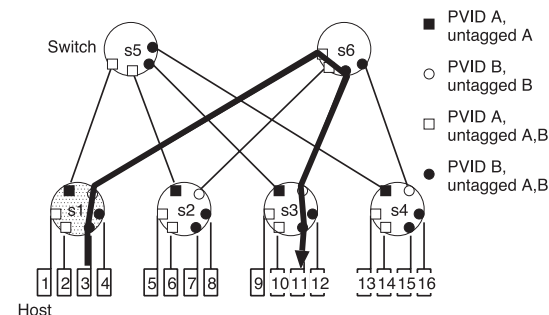


図3 Fat ツリーにおける VLAN リネーミング  
Fig. 3 VLAN renaming on fat tree.

A に属するフレームとして, 矢印で示す経路で転送される.

定理 1 VLAN リネーミングにおいてデッドロックフリー (固定) ルーティングアルゴリズムを用いた場合, ブロードキャストストームは生じない.

証明 1 デッドロックフリー (固定) ルーティングアルゴリズムではチャンネル循環依存が存在しない. よって, フラディング等によりブロードキャストが生じた場合でも, フレームが 1 度通過したスイッチに再度到達することはない. よって, ブロードキャストストームは発生しない. ■

また, VLAN リネーミングでは, VLAN ごとにホストの MAC アドレスをスイッチに静的に登録することにより, 様々なルーティングアルゴリズムを実装することができる.

定理 2 VLAN リネーミングでは, トポロジの次数以下の VLAN 数で  $C \times N \mapsto C$  ルーティングアルゴリズム<sup>\*1</sup>を実装することができる. ここで,  $C$  はチャンネル集合,  $N$  はノード (スイッチ) 集合を表す.

証明 2 VLAN リネーミングでは, 入力ポートに割り当てられた PVID と目的地の MAC アドレスにより出力ポートが決定される. よって, VLAN リネーミングにおいて  $C \times N \mapsto C$  ルーティングアルゴリズムを実装することができる. また, 各スイッチで使用する VLAN ID は, 最大でも各入力ポートに割り当てる PVID 数以下である. よって, VLAN リネーミングはトポロジの次数以下の VLAN 数により実装することができる. ■

\*1 各スイッチにおいて, 入力ポートとフレームの目的地の 2 つの情報からフレームの出力ポートを決定する, ごく一般的な種類のルーティングアルゴリズム<sup>5)</sup>.

### 2.3 スイッチの設定例

VLAN リネーミングを適用する場合に必要なスイッチの設定は以下である。

- スパニングツリープロトコル (STP) を無効にする。
- 使用する VLAN ID を登録する。
- 各ポートについて、所属する VLAN のメンバ設定を行い、PVID を設定する。
- 各 VLAN について、各ホストの MAC アドレスを登録する。
- 性能向上のためにリンクレベルフロー制御を有効にする (オプション)

図 3 に示した Fat ツリートポロジにおいて、スイッチとして Dell PowerConnect 5324 を使用した場合のスイッチ s1 の設定ファイルの記述例を図 4 に示す。ここで、図 3 において、スイッチ s1 のポート 1~6 に、ホスト 1, 2, 3, 4 およびスイッチ s5, s6 がそれぞれ接続されているものとし、VLAN ID は A = 101, B = 102 を割り当てる。なお、図 4 の各行の左端の数字は説明のための行番号であり、実際には記述されない。具体的なコマンドはベンダおよびスイッチの種類ごとに異なるが、設定内容はほぼ同じである。

例として、図 3 のホスト 3 からホスト 11 へフレームを送信する場合 (矢印の経路)、スイッチ s1 では以下のように処理が行われる。

- (1) ホスト 3 から送信された (タグなし) フレームがポート 3 で受信される。この際、図 4 の 19 行目の PVID の設定により、フレームは VLAN 102 (VLAN B) に属すると見なされる。
- (2) フレームの宛先 MAC アドレス (ホスト 11 の MAC アドレス) を、VLAN 102 の MAC アドレステーブル (44~51 行目で設定) 内で検索し、出力すべきポートがポート 6 (スイッチ s6 へのポート) であることを得る (48~50 行目。ただし当該行は省略してある)。
- (3) ポート 6 の VLAN メンバ設定 (26~30 行目) において、VLAN 102 がメンバとして登録されているかどうかを調べ、“タグなし”メンバであることを得る (28 行目)。
- (4) フレームをポート 6 から出力する。このとき、“タグなし”メンバであるため出力フレームの VLAN タグを除去する。

同様の設定をスイッチ s2~s6 に対しても行うことで、図 3 の Fat ツリーにおける VLAN リネーミングの設定が完了する。この際、スイッチ s2~s4 においては、それぞれホスト 5~8, 9~12, 13~16 をポート 1~4 に接続し、スイッチ s5, s6 をポート 5, 6 に接続する。また、スイッチ s5, s6 においては、スイッチ s1~s4 をポート 1~4 に接続する。これにより、スイッチ s2~s6 の設定は、MAC アドレスの登録部分以外の設定はスイッチ s1 とまったく

```

1: // スパニングツリープロトコルを無効にする
2: no spanning-tree
3:
4: // VLAN ID の登録
5: vlan database
6: vlan 101,102 // VLAN ID 101,102 を登録
7: exit
8:
9: // 各ポートの VLAN の設定
10: interface range ethernet g(1,2) // ポート 1,2 (ホスト 1,2 へのポート) を同時に設定
11: switchport mode general // ポートの VLAN モードの設定 (802.1Q フルサポートモード)
12: // VLAN 101,102 のタグなしメンバとして登録
13: switchport general allowed vlan add 101,102 untagged
14: switchport general pvid 101 // PVID を 101 に設定
15: exit
16: interface range ethernet g(3,4) // ポート 3,4 (ホスト 3,4 へのポート) を同時に設定
17: switchport mode general
18: switchport general allowed vlan add 101,102 untagged
19: switchport general pvid 102
20: exit
21: interface ethernet g5 // ポート 5 (スイッチ s5 へのポート) の設定
22: switchport mode general
23: switchport general allowed vlan add 101 untagged
24: switchport general pvid 101
25: exit
26: interface ethernet g6 // ポート 6 (スイッチ s6 へのポート) の設定
27: switchport mode general
28: switchport general allowed vlan add 102 untagged
29: switchport general pvid 102
30: exit
31:
32: // 各 VLAN ごとの MAC アドレスの登録
33: interface vlan 101 // VLAN 101 に関する設定
34: // ホスト 1~4 の MAC アドレスを登録 (ホスト 1~4 はポート 1~4 にそれぞれ接続されている)
35: bridge address MAC_Host1 ethernet g1 delete-on-reset
36: bridge address MAC_Host2 ethernet g2 delete-on-reset
37: bridge address MAC_Host3 ethernet g3 delete-on-reset
38: bridge address MAC_Host4 ethernet g4 delete-on-reset
39: // ホスト 5~16 の MAC アドレスを登録 (VLAN 101 では、ホスト 5~16 はポート 5 (スイッチ s5) の先にある)
40: bridge address MAC_Host5 ethernet g5 delete-on-reset
41: bridge address MAC_Host6 ethernet g5 delete-on-reset
42: ... // 以下同様
43: exit
44: interface vlan 102 // VLAN 102 に関する設定
45: // ホスト 1~4 の MAC アドレスを登録
46: ... (VLAN 101 と同様なので省略)
47: // ホスト 5~16 の MAC アドレスを登録 (VLAN 102 では、ホスト 5~16 はポート 6 (スイッチ s6) の先にある)
48: bridge address MAC_Host5 ethernet g6 delete-on-reset
49: bridge address MAC_Host6 ethernet g6 delete-on-reset
50: ... // 以下同様
51: exit
52:
53: // フローコントロールを全ポートで有効にする (オプション)
54: interface range ethernet all
55: flowcontrol on
56: exit

```

図 4 Fat ツリー (2,4,2) における VLAN リネーミングのスイッチ設定  
Fig. 4 Switch configuration for VLAN renaming on fat tree (2,4,2).

同じとなる。

なお、これらの設定ファイルの記述は、現段階ではクラスタ設計者が手動で行う必要があるため、スイッチ数やホスト数が増加した場合にはその分工数がかかることになる。スクリプトファイル等を記述することにより半自動化することは可能だが、あらゆるトポロジに対応するのは困難である。また、記述した設定ファイルは、構築する結合網とは別のネットワーク（別システムのイーサネットや、シリアルインタフェース等）を経由して各スイッチにアップロードする必要がある。ただし、以上の手順は、従来の SAN を用いたクラスタや、多くの並列分散システムにおいても同様である。

#### 2.4 適用範囲および制限事項

VLAN リネーミングを利用するには、適用可能なイーサネットスイッチ、およびそのスイッチの仕様により、構築可能なネットワークの規模の上限が定まる。

##### 2.4.1 適用可能なイーサネットスイッチ

現在、イーサネットスイッチは、1000BASE-T に限定したとしても数千円～百万円前後のものまで多岐にわたる。しかし、きわめて安価なスイッチは VLAN 技術にすら対応していないものが多く、ごく一部の VLAN 機能のみを提供している安価なイーサネットスイッチでは VLAN リネーミングを利用できない場合がある。つまり、VLAN リネーミングは、1) 2.2.1 項で述べた IEEE 802.1Q 準拠の VLAN タグ操作と、2) 静的な MAC アドレスの登録に対応したスイッチにおいてのみ利用可能である。

##### 2.4.2 構築可能なシステム規模

VLAN リネーミングにより構築可能なシステムの規模は、使用するスイッチの仕様（使用可能な VLAN 数、および登録可能な MAC アドレス数）により上限が定まる。

IEEE 802.1Q の規定では VLAN 数は  $4,094 (2^{12} - 2)$  個と有限であり、コストパフォーマンスの高い安価な商用スイッチでは数十～数百個程度しかサポートしていないものも多い。ただし、VLAN リネーミングで必要となる VLAN 数は、3.2 節の評価結果で述べるとおり、既存の多くのスイッチがサポートしている VLAN 数に比べて小さい。そのため、スイッチがサポートする VLAN 数がシステム規模を制限することはない。

一方、スイッチ内で静的または（学習により）動的に登録されるホストの MAC アドレスは、同じアドレスであっても VLAN ごとに別々に登録されることになる。このため、登録可能な MAC アドレス数、つまりシステムに接続可能なホスト数  $H$  は、使用する VLAN 数  $V$ 、スイッチに（静的に）登録できる MAC アドレス数  $M$  により以下となる。

$$H = \frac{M}{V} \quad (1)$$

この式は、既存の VLAN ルーティング法において MAC アドレスを静的に登録する場合にもあてはまる。しかし、VLAN リネーミングでは、従来の VLAN ルーティング法に比べて  $V$  の値が小さくなるため、より大規模なシステムが構築可能である。

また、VLAN リネーミングは、単純なツリー構造以外のトポロジ、つまりループを含むトポロジに適用した場合に初めて効果が生じる。そのため、適用可能なシステム規模の下限は、2 スイッチで構成され、その 2 スイッチ間が複数リンクで接続されたトポロジである。ただし、複数リンクが独立に制御される状況、つまりリンク集約化を用いない場合に限られる。

### 3. 評価

本章では、提案した VLAN リネーミングの評価を行う。まず、VLAN リネーミングと既存の VLAN ルーティング法（スイッチでタグ付けを行う手法等）のオーバーヘッドにほとんど差がないことを示す。次に、各トポロジにおいて必要となる VLAN 数を明示することにより、2.4.2 項で述べたシステム規模の上限をトポロジごとに算出できるようにする。以上の 2 点により、VLAN リネーミングが大規模な PC クラスタに適用可能であることを示す。

なお、VLAN リネーミングと既存の VLAN ルーティング法のオーバーヘッドがほぼ同じであることを示すことで、VLAN リネーミングを導入することにより実現した並列計算機や SAN のトポロジ・デッドロックフリールーティングと、イーサネットで一般に用いられているツリートポロジとの性能比較については、既存の VLAN ルーティング法を用いた場合の結果<sup>2),4),6)</sup> と同等となる。そのため、我々が保持する評価環境（16 スイッチ、32 ホストの PC クラスタ）での様々なトポロジを用いたシステム評価は既知であると見なし、省略する。

#### 3.1 VLAN リネーミングのオーバーヘッド

コストパフォーマンスに優れる一般的な商用イーサネットスイッチにおいて、VLAN リネーミングのオーバーヘッドを測定した結果を表 1 に示す。数値は、2 ホスト間の ping (ICMP echo メッセージ) を用いて、ノンブロッキング L2 スイッチである Dell PowerConnect 5324 におけるフレーム通過時間を GtrcNET-1<sup>7)</sup> で各 300 回測定した際の最小値 (Min)・平均値 (Ave)・最大値 (Max) をそれぞれとったものである。ICMP echo メッセージのサイズはヘッダを含めて 64 byte とした。よって、イーサネットフレームのデータサイズは IP ヘッダ 20 byte を含めて 84 byte である。ここで、GtrcNET-1 は、搭載する FPGA により機能

表 1 PowerConnect 5324 におけるフレーム通過遅延 ( $\mu\text{sec}$ )Table 1 Latency of PowerConnect 5324 ( $\mu\text{sec}$ ).

	Min	Ave	Max
U-U	2.47	2.74	2.79
T-T	2.47	2.76	2.79
RENAME	2.47	2.75	2.79

表 2 TCP/UDP 転送のバンド幅 (Mbps)

Table 2 Bandwidth of TCP/UDP transfer (Mbps).

U-U (TCP)	941.1
T-T (TCP)	936.9
RENAME(TCP)	941.1
U-U (UDP)	957.0
T-T (UDP)	954.4
RENAME(UDP)	957.0

をプログラム可能なネットワーク試験用装置であり、ネットワークの遅延やバンド幅の計測、流量制御等を Gigabit Ethernet のワイヤレートで行うことができる。

表 1 において、U-U は VLAN をいっさい用いない場合、T-T はホストにおいて VLAN タグ付きフレームを送受信した場合（従来の VLAN ルーティング法に相当）、RENAME はスイッチ内でのみ PVID に基づく転送を行う場合（VLAN リネーミングに相当）をそれぞれ示す。結果より、VLAN リネーミングの導入による遅延はほとんどないことが分かる。なお、VLAN リネーミングでは、ホストでの VLAN 処理はいっさい行わないため、ホストにおける VLAN リネーミング導入のオーバーヘッドはない。

次に、Tperf 1.5<sup>8)</sup> を用いた TCP/UDP 転送のバンド幅の測定結果を表 2 に示す。結果より、VLAN リネーミングの導入によるバンド幅の低下はないことが分かる。一方で、ホストが VLAN タグ付きフレームを送受信する場合（表の T-T）、数 Mbps のバンド幅低下が見られる。これは、タグ付きフレームでは、フレーム全体に占めるペイロードの割合が 4 byte の VLAN タグの分だけ少なくなるためである。ここで、VLAN リネーミングでは、フレームはリンク上においては VLAN タグを含まない。つまり、VLAN リネーミングでは、VLAN の処理をスイッチ内でのみ行うことにより、従来の VLAN ルーティング法と違って VLAN タグ導入によるバンド幅低下を生じない。

### 3.2 VLAN 数

本節では、並列計算機や SAN で採用されている典型的なトポロジにおいて、ホストで

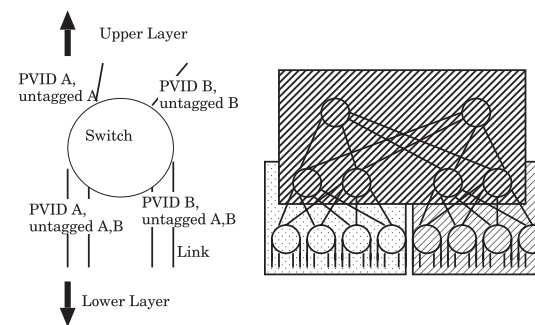


図 5 Fat ツリー (2,4,2) における VLAN リネーミング

Fig. 5 VLAN renaming on fat tree (2,4,2).

VLAN タグ付けを行う従来の VLAN ルーティング法と VLAN リネーミングとでそれぞれ必要になる VLAN 数を比較する。トポロジとしては、Fat ツリー、メッシュおよびトーラス ( $k$ -ary  $n$ -cube)、不規則トポロジを取り上げ、まず、各トポロジにおける VLAN リネーミングおよび従来の VLAN ルーティング法の適用例を示してから、必要となる VLAN 数の比較を行う。

#### 3.2.1 Fat ツリー

2.2.2 項の例で述べた方法（図 3）を階層的に適用することで、Fat ツリートポロジに VLAN リネーミングを適用することができる。図 5 に示す Fat ツリーでは、VLAN A および B の 2 個の VLAN を必要とする。一般に、スイッチの上位リンク数  $u$ 、下位リンク数  $d$ 、レイヤ数  $r$  の Fat ツリー ( $u, d, r$ ) の場合には、 $u$  個の VLAN が必要となる。

一方、従来の VLAN ルーティング法では、図 5 の Fat ツリー (2,4,2) 上で VLAN リネーミングと同じ経路集合をとる場合、図 6 に示す 4 つのトポロジそれぞれに VLAN を割り当てる必要がある。Fat ツリー ( $u, d, r$ ) の場合には、 $u^r$  個の VLAN が必要となる。

#### 3.2.2 $k$ -ary $n$ -cube メッシュ

$k$ -ary  $n$ -cube メッシュトポロジにおいて、デッドロックを回避しつつ分散された経路集合を確保するために、次元順ルーティング<sup>5)</sup> が用いられる。次元順ルーティングは、各フレームを X 方向、Y 方向、Z 方向と次元順にそれぞれ必要ホップ数転送することにより、最短経路をとることができるデッドロックフリールーティングアルゴリズムである。ここでは、メッシュトポロジに VLAN リネーミングを適用して次元順ルーティングに従う経路集合を

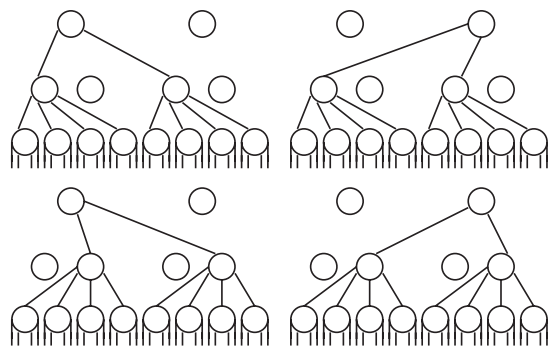


図 6 Fat ツリー (2,4,2) における VLAN ルーティング  
Fig. 6 VLAN-based routing on fat tree (2,4,2).

とる場合の VLAN 割当てを示す。

図 7 は、3次元メッシュ ( $k$ -ary 3-cube) に VLAN リネーミングを適用した場合の各スイッチにおける VLAN 割当てである。図に示すとおり、Z 次元入力ポートに入力されたフレームは、Z 次元出力ポートもしくはホストに転送される。同様に、Y 次元入力ポートに入力されたフレームは、Y、Z 次元出力ポートもしくはホストに転送される。また、X 次元入力ポートまたはホストと接続されたポートに入力されたフレームは、X、Y、Z 次元出力ポートもしくはホストに転送される。このように、 $k$ -ary  $n$ -cube メッシュトポロジにおける次元順ルーティングでは、VLAN は各次元に対して 1 つ、つまり  $n$  個必要となる。

一方、従来の VLAN ルーティング法では、以下のように VLAN 割当てを行う<sup>9)</sup>。図 7 の 3次元メッシュの矢印で示されるスイッチ (A とする) に接続されたホストは、図 8(a) に示すツリートポロジを用いることにより、どの宛先ホストに対しても次元順ルーティングに従う最短経路でフレームを送信することができる。このトポロジは、スイッチ A を通る X 次元方向のリンク (集合)  $L_X$  (図 8(b)),  $L_X$  と交差する Y 次元方向のリンク集合  $L_Y$  (図 8(c)),  $L_Y$  と交差する Z 次元方向のリンク集合  $L_Z$  (図 8(d))<sup>\*1</sup> で構成される。このようなトポロジはスイッチの Y 座標と Z 座標の組  $(y, z)$  に対し 1 つ定まり、各トポロジに VLAN を割り当てることにより、全ホスト間で次元順ルーティングに従う経路をとることができる。よって、図 7 の 3次元メッシュ (4-ary 3-cube) では  $4^2 = 16$  個の VLAN が必

\*1  $L_Z$  は Z 次元方向の全リンクで構成される。

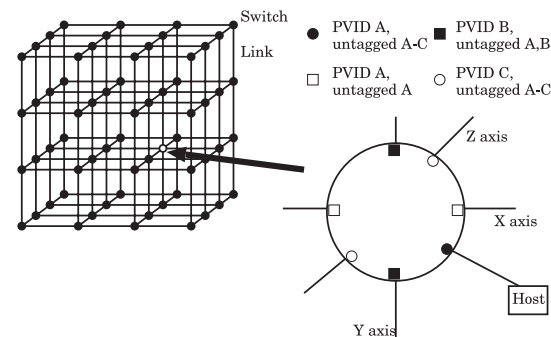


図 7 3次元メッシュにおける VLAN リネーミング  
Fig. 7 VLAN renaming on 3-D mesh.

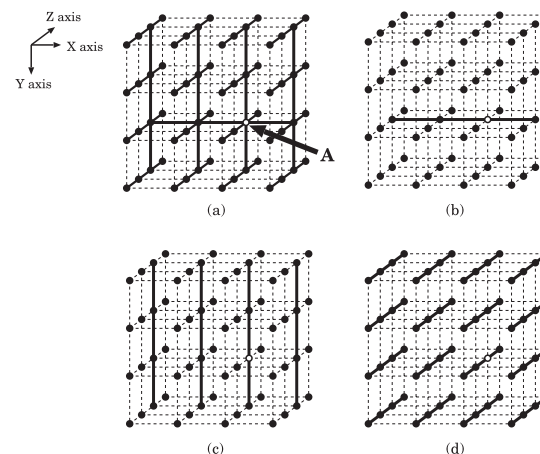


図 8 3次元メッシュにおける VLAN ルーティング  
Fig. 8 VLAN-based routing on 3-D mesh.

要となる。一般に  $k$ -ary  $n$ -cube メッシュの場合では、 $k^{n-1}$  個の VLAN が必要となる。

### 3.2.3 $k$ -ary $n$ -cube トーラス

一般に、並列計算機の相互結合網において次元順ルーティングをトーラストポロジに適用する場合、デッドロックフリーを保証するために仮想チャンネルが必要になる。ここで、イー



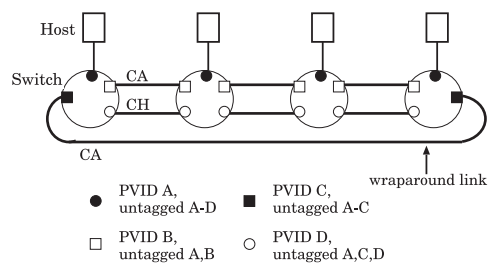


図 9 1次元トーラスにおける VLAN リネーミング  
 Fig. 9 VLAN renaming on 1-D torus.

サネットには仮想チャンネルがないため、VLAN リネーミングを適用する場合、スイッチ間に複数リンク（それぞれ CA リンク、CH リンクと呼ぶ）を設ける。そして、それぞれのリンクに異なる VLAN ID を割り当てる。

図 9 の 1 次元トーラス（リングトポロジ）において、wraparound リンクを通過する前のフレームは CA リンクを、wraparound リンクを 1 度通過した（今後通過しない）フレームは CH リンクを利用して転送する。この制御によりチャンネル循環依存は除去され、デッドロックフリーが保証される<sup>10)</sup>。このように、 $k$ -ary  $n$ -cube トーラストポロジにおける次元順ルーティングでは、VLAN は各次元に対して 3 つ（図 9 の B, C, D）必要であるが、メッシュの場合と違いホストに接続されたポートに別の PVID を割り当てなければならないため（図 9 の A）、必要な VLAN 数は  $3n + 1$  個となる。

一方、従来の VLAN ルーティング法をたとえば  $4 \times 4$  2 次元トーラス（4-ary 2-cube）に適用する場合、図 10 の (a) ~ (h) に示す 8 つのトポロジにそれぞれ VLAN を割り当てる<sup>9)</sup>。図では wraparound リンクは途中で切れているが、上下および左右で 1 本につながっているものとする。また、VLAN リネーミングの場合（図 9）と同様、循環依存を除去するためにスイッチ間に複数リンクを設ける必要があるが、図では省略している。

図 10 において、たとえば矢印で示すスイッチ A に接続されたホストは、宛先に応じて VLAN (b) と (f) のいずれかを選択することにより、すべての宛先ホストに対して次元順ルーティングに従う最短経路でフレームを送信することができる。一般に  $k$ -ary  $n$ -cube トーラスの場合では、 $2k^{n-1}$  個の VLAN が必要となる。

### 3.2.4 不規則トポロジ

不規則なトポロジ上でも使用することができるルーティングアルゴリズムとして、

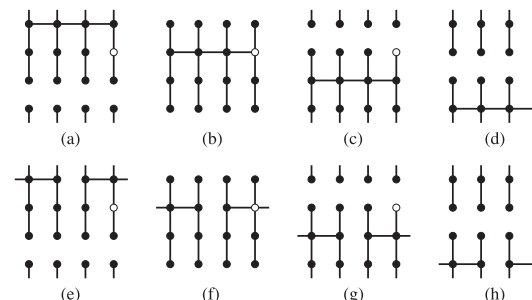
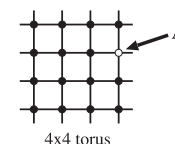


図 10 2次元トーラスにおける VLAN ルーティング  
 Fig. 10 VLAN-based routing on 2-D torus.

Up\*/Down\*ルーティング<sup>11)</sup>がある。Up\*/Down\*ルーティングは、スパニングツリーに基づいた有向グラフを用いることで、任意のトポロジに適用することができる典型的なデッドロックフリールーティングアルゴリズムである。各リンクは単方向チャンネル 2 本で構成されており、図 11 のように、ツリーのルート方向へ向かうチャンネルに Up 方向、リーフ方向へ向かうチャンネルに Down 方向を割り当てる。すべてのフレームは、0 回以上 Up 方向に転送された後に、0 回以上 Down 方向に転送されることで宛先ホストまで到達する。Down 方向から Up 方向へのターンを行うことができないため、チャンネル間の循環依存が除去され、デッドロックフリーが保証される。

VLAN リネーミングを Up\*/Down\*ルーティングに適用する場合、Down 方向から Up 方向へのフレームの転送を防ぐために、各スイッチにおいて必要となる VLAN 数は Down リンク数 + 1 となる。

一方、従来の VLAN ルーティング法の場合は、不規則なトポロジにおいて Up\*/Down\*ルーティングを実現するための VLAN 割当てを一般化するのには困難である。

### 3.2.5 VLAN 数の比較

表 3 に、各トポロジにおいて、ホストで VLAN タグ付けを行う従来の VLAN ルーティング法（Host Tagged）と VLAN リネーミングのそれぞれで必要となる VLAN 数を示す。

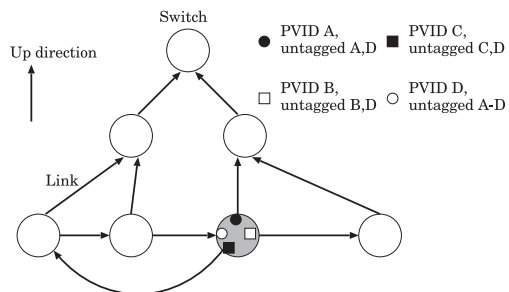


図 11 Up\*/Down\*ルーティングにおける VLAN リネーミング  
Fig. 11 VLAN renaming on Up\*/Down\* routing.

表 3 VLAN 数の比較  
Table 3 Comparison of number of VLANs.

	Host Tagged	RENAME
Fat ツリー ( $u, d, r$ )	$u^r$	$u$
$k$ -ary $n$ -cube メッシュ	$k^{n-1}$	$n$
$k$ -ary $n$ -cube トーラス	$2k^{n-1}$	$3n + 1$
不規則 (UD ルーティング)	-	$D + 1$

ここで、Up\*/Down\*ルーティング (UD ルーティング) において、 $D$  は対象トポロジ中で 1 つのスイッチが持つ Down 方向のチャンネル数の最大値を表す。

表 3 より、ホストにおいてフレームの VLAN タグ付けを行う従来の方法と異なり、VLAN リネーミングでは少数の VLAN で結合網を構築できることが分かる。ここで、2.4.2 項の式 (1) に示したとおり、結合網に接続できるホスト数  $H$  は、必要となる VLAN 数  $V$  に反比例する。たとえば QsNET<sup>12)</sup> で用いられる 4 進 Fat ツリー (Fat ツリー ( $4, 4, r$ )) を構築する場合、従来の VLAN ルーティング法では、レイヤ数  $r = 2$  で  $V = 4^2 = 16$  個、 $r = 3$  で  $V = 4^3 = 64$  個となる。スイッチの MAC アドレステーブルのエントリ数  $M = 8,000$  と仮定すると、最大でもそれぞれ  $H = 8,000/16 = 500$  ホスト、 $H = 8,000/64 = 125$  ホストしか接続できない。一方、VLAN リネーミングでは、レイヤ数  $r$  によらず  $V = 4$  個の VLAN で済むため、 $H = 8,000/4 = 2,000$  ホストまで接続することが可能である。

以上より、VLAN リネーミングを用いることにより、従来の VLAN ルーティング法に比べ多数のスイッチを含む大規模なネットワークを構築可能であることが分かる。

#### 4. 関連研究

イーサネットにおいて VLAN を用いてホスト間に複数の経路を設定し、ループ構造を含む様々なトポロジを利用できるようにするルーティング技術は国内外でほぼ同時期に提案された<sup>2),3)</sup>。工藤らが提案した VLAN ルーティング法<sup>2)</sup>では、図 1 のようにループを含まない各リンク集合にそれぞれ異なる VLAN を割り当てることで、ブロードキャストストームを回避しつつ同一スイッチ間に複数の経路を実現する。

VLAN 技術を利用して PC クラスタのインタコネクトを構築する手法はその後国内を中心に活発に議論され、三浦らの研究<sup>6),13)</sup>では、MAC アドレスから VLAN ID を決定しタグ付けを行うための Linux 用デバイスドライバを開発し、TCP/IP を用いた VLAN ルーティング法の実現している。この手法では、MAC アドレスに基づいた VLAN ID の制御を行うことで、送信先に応じた VLAN の選択をドライバに任せられるようになるため、上位レイヤのソフトウェア環境に手を加えることなく VLAN ルーティング法を実現できる。

これに対し我々は、様々なトポロジにおける VLAN の割当て方法<sup>9)</sup>や、スイッチにおいて VLAN タグ付けを行うことで、システムソフトウェアが VLAN 機能をサポートしていない場合にも VLAN ルーティング法を利用できるようにする手法<sup>4)</sup>を提案している。さらに、ループを含むトポロジにおいて、チャンネル循環依存が生じる経路でフレームの転送を行う場合に転送バンド幅が著しく低下することを検証し、リンクレベルフロー制御 (IEEE 802.3x) の有無にかかわらずデッドロックフリールーティングが効果的であることを示した<sup>4)</sup>。

一方、VLAN 技術を用いずに、静的にホストの MAC アドレスを登録することでルーティングを行う方法も検討されているが、ブロードキャストストームが発生した場合の対処方法、ならびに使用可能なルーティングアルゴリズムが限定される。たとえば、3.2.4 項で評価を行った Up\*/Down\*ルーティング<sup>11)</sup>では、図 12 において S2 から D へは点線の経路により 1 ホップで転送されるが、S1 から D へは実線の経路により 3 ホップで転送される。ここで、VLAN を用いない方法では、S2 において同一宛先のフレームの出力ポートを入力ポートごとに設定することができないため、Up\*/Down\*ルーティングは実装できない。一方、VLAN を用いた場合、入力ポートごとに異なる VLAN を割り当てることで、様々なルーティングアルゴリズムを実装することができる。

このほか、朴らの研究<sup>14),15)</sup>では、VLAN ルーティング法とは異なるが、大規模クラスタ

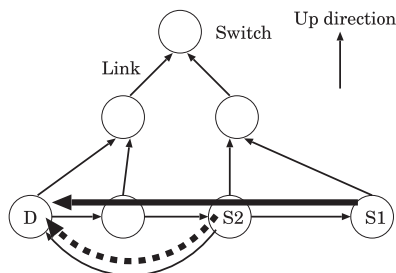


図 12 Up\*/Down\*ルーティングの例  
Fig.12 Example of Up\*/Down\* routing.

システム用のネットワークとしてイーサネットによる 3 次元ハイパークロスバ網を実現し、そのための専用通信ライブラリ PM/Ethernet-HXB を新たに開発した。VLAN ルーティング法によるハイパークロスバ網<sup>2)</sup>では節点において各次元方向のスイッチを直接接続するが、この方式では、ホスト上の PM ドライバがフレームのルーティングを担当する点異なる。また、イーサネットにおいて、スパンニングツリープロトコル (STP) を用いずに循環除去を行う高性能なルーティングアルゴリズムについての議論も行われている<sup>16),17)</sup>。

これらの関連研究に対し、我々の提案する VLAN リネーミングは、VLAN 数およびスイッチのアドレステーブルに登録する MAC アドレス数を大きく削減可能であり、ホストのシステムソフトウェアが VLAN タグ付けに対応していない場合にも適用可能である。既存の VLAN ルーティング法を用いて大規模クラスタを構築する場合のこれらの問題点を解決している点で、本手法は柔軟性が高く、イーサネットを用いた大規模クラスタの構築法として有力であるといえる。

## 5. ま と め

本稿では、比較的安価な L2 イーサネットスイッチを用いて大規模クラスタを構築するために、従来の VLAN ルーティング法の問題点を解決する VLAN リネーミングを提案した。VLAN リネーミングは、1) 必要となる VLAN 数がスイッチの次数以内と少数であり、2) ホストのシステムソフトウェアが VLAN 技術に対応していない場合にも適用可能、という 2 つの利点を持ち、イーサネット上で様々なトポロジおよび (デッドロックフリー) ルーティングアルゴリズムを実装することができるようになる。評価結果より、VLAN リネーミング導入によるオーバーヘッドはほとんどなく、並列計算機や SAN で用いられている典型的な

トポロジを構築する場合に、必要となる VLAN が数個ときわめて少ないことが分かった。

VLAN リネーミングは、多くの商用 L2 イーサネットスイッチにおいてサポートされている VLAN 機能を制御することにより実現できる点、集約化したリンク群に対して VLAN ID を割り当てることによりリンク集約化技術と併用することができる点において、高い実用性と汎用性を持つ技術であるといえる。また、VLAN リネーミングはクラスタ内部のイーサネットに閉じた手法ではなく、柔軟な経路を設定するためにグリッドや LAN 技術の一部として利用することも可能である。

謝辞 本研究の一部は、科学技術振興機構「JST」の戦略的創造研究推進事業「CREST」の支援による。

## 参 考 文 献

- 1) TOP500 Supercomputing Sites. <http://www.top500.org/>
- 2) 工藤知宏, 松田元彦, 手塚宏史, 児玉祐悦, 建部修見, 関口智嗣: VLAN を用いた複数パスを持つクラスタ向き L2 Ethernet ネットワーク, 情報処理学会論文誌: コンピューティングシステム, Vol.45, No.SIG 6(ACS 6), pp.35-43 (2004).
- 3) Sharma, S., Gopalan, K., Nanda, S. and cker Chiueh, T.: Viking: A Multi-Spanning-Tree Ethernet Architecture for Metropolitan Area and Cluster Networks, *Proc. 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (Infocom 2004)*, pp.2283-2294 (2004).
- 4) 大塚智宏, 鯉淵道紘, 工藤知宏, 天野英晴: スイッチでタグ付けを行う VLAN ルーティング法, 情報処理学会論文誌: コンピューティングシステム, Vol.47, No.SIG 12(ACS 15), pp.46-58 (2006).
- 5) Dally, W.D. and Towles, B.: *Principles and Practices of Interconnection Networks*, Morgan Kaufmann (2003).
- 6) 三浦信一, 岡本高幸, 朴 泰祐, 佐藤三久, 高橋大介: tagged-VLAN に基づく PC クラスタ向け高バンド幅ツリーネットワークの開発, 情報処理学会研究報告 2005-HPC-104, pp.13-18 (2005).
- 7) Kodama, Y., Kudoh, T., Takano, R., Sato, H., Tatebe, O. and Sekiguchi, S.: GNET-1: Gigabit Ethernet Network Testbed, *Proc. 2004 IEEE International Conference on Cluster Computing (Cluster2004)* (2004).
- 8) Tperf. <http://www.am.ics.keio.ac.jp/~terry/tperf/>
- 9) Otsuka, T., Koibuchi, M., Jouraku, A. and Amano, H.: VLAN-based Minimal Paths in PC Cluster with Ethernet on Mesh and Torus, *Proc. 2005 International Conference on Parallel Processing (ICPP-05)*, pp.567-576 (2005).
- 10) Duato, J., Yalamanchili, S. and Ni, L.: *Interconnection Networks: An engineering approach*, Morgan Kaufmann (2002).

- 11) Schroeder, M.D., Birrell, A.D., Burrows, M., Murray, H., Needham, R.M. and Rodeheffer, T.L.: Autonet: A High-speed, Self-configuring Local Area Network Using Point-to-point Links, *IEEE Journal on Selected Areas in Communications*, Vol.9, No.8, pp.1318-1335 (1991).
- 12) Petrini, F., Feng, W.C., Hoisie, A., Coll, S. and Frachtenberg, E.: The Quadrics Network (QsNet): High-Performance Clustering Technology, *Proc. Hot Interconnects 9*, pp.125-130 (2001).
- 13) Miura, S., Okamoto, T., Boku, T., Sato, M. and Takahashi, D.: Low-cost High-bandwidth Tree Network for PC Clusters based on Tagged-VLAN Technology, *Proc. 8th International Symposium on Parallel Architectures, Algorithms and Networks (I-SPAN 2005)*, pp.84-93 (2005).
- 14) 朴 泰祐, 佐藤三久, 宇川 彰: 計算科学のための超並列クラスタ PACS-CS の概要, 情報処理学会研究報告 2005-HPC-103, pp.133-138 (2005).
- 15) 住元真司, 久門耕一, 朴 泰祐, 佐藤三久, 宇川 彰: PACS-CS のための Ethernet を用いた高性能通信機構の設計, 情報処理学会研究報告 2005-HPC-103, pp.139-144 (2005).
- 16) Pellegrini, F.D., Starobinski, D., Karpovsky, M.G. and Levitin, L.B.: Scalable Cycle-Breaking Algorithms for Gigabit Ethernet Backbones, *Proc. 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (Infocom 2004)*, pp.2175-2184 (2004).
- 17) Reinemo, S.-A. and Skeie, T.: Ethernet as a Lossless Deadlock Free System Area Network, *Proc. 3rd International Symposium on Parallel and Distributed Processing and Applications (ISPA'05)*, pp.901-914 (2005).

(平成 20 年 5 月 9 日受付)

(平成 20 年 8 月 14 日採録)



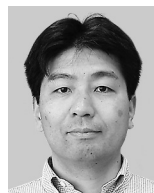
大塚 智宏 (学生会員)

2003 年慶應義塾大学大学院理工学研究科開放環境科学専攻前期博士課程修了。現在, 同後期博士課程に在学。2006 年度より慶應義塾インフォメーションテクノロジーセンター本部助教。計算機アーキテクチャ, 並列分散処理に関する研究に従事。



鯉淵 道紘 (正会員)

平成 12 年慶應義塾大学理工学部情報工学科卒業。平成 15 年同大学大学院理工学研究科開放環境科学専攻博士課程修了。博士 (工学)。平成 14 年度より 16 年度まで日本学術振興会特別研究員。現在, 国立情報学研究所助教, 総合研究大学院大学複合科学研究科情報学専攻助教 (兼任)。ハイパフォーマンスコンピューティングとインタコネクトに関する研究に従事。IEEE Computer Society Japan Chapter Young Author Award 2007, 平成 19 年度情報処理学会論文賞受賞。IEEE, 電子情報通信学会各会員。



工藤 知宏 (正会員)

1991 年慶應義塾大学大学院理工学研究科博士課程単位取得退学。東京工科大学助手, 講師, 助教授を経て, 1997 年より新情報処理開発機構並列分散システムアーキテクチャつくば研究室長, 2002 年より産業技術総合研究所グリッド研究センタークラスタ技術チーム長。2008 年より同所情報技術研究部門インフラウェア研究グループ長。博士 (工学)。並列処理, 通信アーキテクチャに関する研究に従事。IEEE CS 会員。



天野 英晴 (正会員)

1986 年慶應義塾大学大学院理工学研究科電気工学専攻博士課程修了。工学博士。現在, 同大学理工学部情報工学科教授。計算機アーキテクチャ, リンコンフィギャラブルシステムの研究に従事。電子情報通信学会, IEEE 各会員。