

## tagged-VLANとマルチリンクに基づくPCクラスタ向け 高性能・耐故障ネットワークの実装と評価

三 浦 信 一<sup>†</sup> 岡 本 高 幸<sup>†</sup> 朴 泰 祐<sup>†</sup>  
佐 藤 三 久<sup>†</sup> 高 橋 大 介<sup>†</sup>

安価なコモディティネットワークである Ethernet は、その高いコストパフォーマンスから、多くの PC クラスタで使用されている。しかしながら、大規模な HPC クラスタに用いるには、バンド幅・耐故障性および拡張性が低く問題があった。我々は、この問題を解決する RI2N/UDP および VFREC-Net をこれまで開発してきた。これらのシステムは複数リンクをノード入出力あるいは中間スイッチに多用することにより、バンド幅の増強と耐故障性を実現する。この二つの技術は直交しており、同時に使用することで既存の Ethernet が持つ問題点を解決することが可能になる。加えて我々は 1U の筐体内に独立した 16 プロセッサを実装した省電力・高密度クラスタユニット MegaProto/E を開発している。この MegaProto/E 上に RI2N と VFREC-Net を構築しベンチマークを行ったところ、高いバイセクションバンド幅と耐故障性が得られることがわかった。これにより、RI2N と VFREC-Net を同時に用いることが、ネットワークシステムに対して有用であることがわかった。

### High-bandwidth and Fault-tolerant Network for PC Clusters based on Tagged-VLAN and Multi-link Ethernet Technologies

SHIN'ICHI MIURA,<sup>†</sup> TAKAYUKI OKAMOTO,<sup>†</sup> TAISUKE BOKU,<sup>†</sup>  
MITSUHIISA SATO<sup>†</sup> and DAISUKE TAKAHASHI<sup>†</sup>

Ethernet is the most popular interconnection network to be used on various PC clusters. Especially, Gigabit Ethernet provides a very high cost/performance ratio as the most used commodity network. However, it is difficult to utilize it on a large scale HPC clusters because its absolute bandwidth and scalability is limited. To solve this problem, we have been developing both RI2N/UDP and VFREC-Net systems for wide bandwidth, high scalability and high dependability for large scale HPC clusters. These two technologies are individual and orthogonal, therefore it is possible to couple them to provide a powerful solution over traditional Ethernet-based PC clusters. We combine these technologies on an ultra-compact cluster unit named MegaProto/E, which consists of 16 node processors fully connected by dual Gigabit Ethernet in a 1U form factor. Applying both RI2N/UDP and VFREC-Net on MegaProto/E cluster system, we confirmed they realize a high bisection bandwidth and dependability against link or switch failure. As a result, it is shown that the combination of these techniques is useful to build large scale HPC cluster systems.

#### 1. はじめに

高性能 PC サーバを相互結合した PC クラスタは、HPC 分野の様々な局面で多用されている。これらの HPC 向け PC クラスタは、比較的安価でありながら高性能であるため幅広い分野で利用されている。しかし HPC 分野ではノード間ネットワークに対し、バンド幅・耐故障性そして拡張性に関して要求が厳しい。そのため主要部品の多くにコモディティ製品を活用し

つつも、ネットワークだけは専用ネットワークを選択することが多い。特に現在、クラスタの規模は大きくなり、ノード数の増加と高密度化が進んでいる。ノード数の増加とともに相対的に部品点数の割合が増大するため、ネットワークは耐故障性に優れていなければならない。これに加えてクラスタの規模に合わせて全体性能を向上させるためには、より柔軟なネットワークトポロジを許容し、高いバイセクションバンド幅を確保するネットワークシステムが必要になる。このような環境で、コモディティネットワークである既存の Ethernet を用いた場合、耐故障性が低くなり、またネットワークトポロジは原則 Tree 形状に制限されるため、それによりバンド幅ボトルネックが生じやす

<sup>†</sup> 筑波大学大学院 システム情報工学研究科  
Graduate School of Systems and Information Engineering, University of Tsukuba

い。これらの問題が計算性能のボトルネックを引き起こす可能性がある。そのためクラスタ規模が大きくなるに伴い、これらの問題を解決する専用ネットワークである Myrinet<sup>1)</sup> や InfiniBand<sup>2)</sup> といった SAN (System Area Network) が多くの HPC クラスタで採用されている。しかし SAN は、コモディティネットワークである Ethernet と比較して高価である。今後クラスタがより大規模した場合、ネットワークの導入コストが問題になる。そのため既存の Ethernet を用いつつ、拡張性・耐故障性をもつネットワークシステムが必要とされている。

これまで我々はこれらの問題を解決するべく、ノード間の接続に対して高バンド幅と耐故障性を提供する RI2N<sup>3)</sup>、および、スイッチ間の接続性に柔軟性を持たせ、クラスタの規模に拡張性を持たせる VFREC-Net<sup>4)</sup> をそれぞれ開発してきた。本稿ではこれら二つのシステムを統合し PC クラスタで活用することを提案する。加えて、今後のクラスタシステムのより高密度な実装を目的として、我々が開発してきた MegaProto<sup>5)</sup> にこのネットワークシステムを適用しその性能を評価する。

## 2. 提案システム

大規模で高密度なクラスタシステムでは、ネットワークに耐故障性と拡張性を与えることが重要になる。これに加え、ネットワークの導入コストについても意識しなければならない。しかし、ネットワークの導入コストに優れる Ethernet でこれらの問題を解決することは難しい。我々はこの問題を解決するために、コモディティネットワークである Ethernet を用いつつ、高性能 PC クラスタ向けのネットワークとして耐故障性と拡張性を提供する RI2N<sup>3)</sup> と VFREC-Net<sup>4)</sup> の 2 種類のシステムを提案・開発してきた。RI2N はノード間のネットワークに対して、複数リンクの適用により高バンド幅・耐故障性をもつネットワークを提供する。一方 VFREC-Net はスイッチ間の接続に複数経路を許す柔軟なトポロジを実現することで、スイッチ間に高いバンド幅と高いシステム拡張性を提供する。

まず始めにこの 2 種類のネットワークについて簡単に述べ、その後提案手法について述べる。なお、それぞれのネットワークの詳細については文献<sup>3),4)</sup> を参照されたい。

### 2.1 RI2N

RI2N (Redundant Interconnection with Inexpensive Network) はノード間のネットワークに耐故障性と高バンド幅を同時に提供するためのものである。RI2N では、各ノードを複数の異なるネットワークに所属させ、これらのネットワークリンクを同時に利用することで目的を達成する。すべてのネットワークが正常な場合、これらの複数のネットワークを同時に利用することでユーザに高いネットワークスループットを提供す

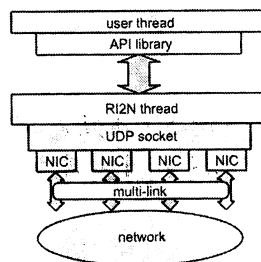


図 1 RI2N/UDP の実装イメージ

る。また一部のネットワークが故障などの理由により利用できなくなった場合、そのネットワークを送受信の対象から除外し、残りの正常なネットワークだけを用いて送受信を継続する。

現在の RI2N の実装として RI2N/UDP<sup>3)</sup> がある。RI2N/UDP の実装イメージを図 1 に示す。RI2N/UDP では実装の簡便性と移植性を得るため、一般的な socket API を用いたユーザレベル実装とした。RI2N/UDP は下位プロトコルとして、一般的な TCP/IP ではなく UDP/IP を使用する。コネクションレス指向型の UDP/IP を利用することで、故障発生時においても安定動作することが可能になる。RI2N/UDP は、マルチリンク上で UDP/IP プロトコルによるパケット制御を行うことにより、TCP/IP 相当のストリーム通信機能を提供している。このため、複数リンクの監視機能、パケットの順序制御機能、ロスパケットの再送機能、フロー制御機能等がユーザレベルライブラリ上に実装されている。RI2N/UDP は移植性を高めるためユーザに TCP/IP に類似した API を提供する。このことから、TCP/IP を用いる Open MPI<sup>6)</sup> など現在 PC クラスタで利用されている多くの MPI 実装に容易に移植が可能である。

### 2.2 VFREC-Net

クラスタの性能をノード数に合わせて向上させるためには、スイッチ間のバンド幅を出来る限り増強しなくてはならない。たとえばノード-スイッチ間に Gigabit Ethernet (以下 GbE) を用いる場合、スイッチ-スイッチ間の接続には GbE よりも高速なリンク、たとえば 10GbE 等を用いるべきである。しかし、それらの高速なリンクをサポートするスイッチは GbE のそれよりも高価であり、コストパフォーマンスのために GbE を利用している以上、上位リンクでの 10GbE の利用はそのメリットを乏しくする。バンド幅を高めるためにスイッチ間に複数のリンクを用意するという方法も考えられるが、安価な Layer-2(以後 L2) スイッチを用いる場合、それを実現することは困難となる。L2 スイッチを用いる場合、ネットワークにループができる構造はブロードキャストストームの原因となるため利用できず、そのため複数のリンクをスイッチ間に用意できない。IEEE 802.3ad<sup>7)</sup> を用いることでスイッチ間に複数のリンクを用意することが可能であ

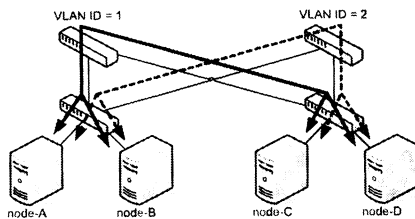


図2 VLAN ルーティング法イメージ図

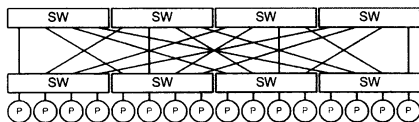


図3 VLAN Based Fat Tree ネットワーク

るが、その規格制約上、やはりバンド幅は制限される。そのため、現在まで多くのPCクラスタでは単純Tree型のネットワークのみが用いられてきた。そこで、tagged-VLAN<sup>8)</sup>を用い、各ノードから送信するパケットのVLAN IDを制御することで、Ethernetを用いた場合でも柔軟なトポロジを可能にするVLANルーティング法が提案されている<sup>9)</sup>。VLANルーティング法を用いた場合のイメージを図2に示す。VLANルーティング法を用いることで既存のEthernetでもスイッチ間のバンド幅ボトルネックを解決できる。我々は、このVLANルーティング法を改良し、システムレベルで実装したうえで、よりPCクラスタ向けに利用しやすいVFREC-Net(VLAN-based Flexible, Reliable and Expandable Commodity Network)<sup>4)</sup>を開発した。

VFREC-Netは仮想Ethernetデバイスのドライバとしてシステムレベル実装されている。そのためTCP/IPやUDP/IPといったネットワークプロトコルを一切変更することなく使用することが可能になる。このVFREC-Netを用いることで図3に示すようなVBFT(VLAN Based Fat Tree)<sup>9)</sup>の構成をとることが可能になり、ノード数に応じて高いバイセクションバンド幅を得ることが可能になる。

### 2.3 ネットワークシステムの統合

RI2N/UDPは、前述のように通常のUDP/IPを用いたユーザレベル実装となっている。一方でVFREC-Netはドライバを用いたシステムレベル実装である。このように、現状のRI2N/UDPとVFREC-Netは、実装の階層が明確に分離されているため、この2つのシステムを同時に利用することが可能となっている。VFREC-Netでスイッチ間のバンド幅を高めたネットワークを、そしてそれらのネットワークを複数用意し、そのネットワークをRI2N/UDPで同時に利用することで、高バンド幅・耐故障性のあるネットワークを実現できる。

表1 MegaProto/E プロセッサカード仕様

|          |                                                                                                         |
|----------|---------------------------------------------------------------------------------------------------------|
| CPU      | Transmeta Efficeon 1.0GHz (TM8820)                                                                      |
| Memory   | DDR/333MHz 512Mbytes (Single Channel)                                                                   |
| NIC      | ALI Uli1563 100 Base-TX [管理用]<br>Broadcom BCM5704S 1000Base-T × 2<br>PCI-X 64bit/66MHz (MTU=1500) [計算用] |
| OS       | Linux Kernel 2.6.16                                                                                     |
| Compiler | GCC ver.4.0                                                                                             |

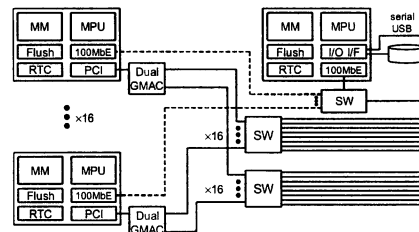


図4 MegaProto/E Unit 内ネットワーク

## 3. MegaProto

我々は現在、高性能計算のための低電力・高密度クラスタMegaProto<sup>5)</sup>を開発している。現在、最終プロトタイプとして稼動しているMegaProto/Eは、プロセッサにTransmeta社のEfficeonを用いている。MegaProto/Eは、16個の計算ノードと1台の管理ノード、およびそれらのノードを接続するEthernetスイッチ3台が1Uの筐体内に収められている。計算ノードはそれぞれ完全に独立したPCとして動作し、この1Uの筐体のみで16ノードのPCクラスタとして動作可能になっている。表1にそれぞれのMegaProto/Eでの各計算ノードの詳細と、図4に筐体内の内部ネットワークの詳細を示す。

今回提案したRI2N/UDPとVFREC-Netを組み合わせたシステムを、このMegaProto/Eで利用する。16台のノードは、1系統の管理ネットワークと2系統の計算用ネットワークの合計3系統のネットワークに接続されている。それぞれのネットワークは、MegaProto内に用意した3台のスイッチで接続されている。それぞれのスイッチは相互に接続されておらず、独立したネットワークとなる。2系統の計算用ネットワークスイッチは24 portのL2 GbEスイッチに相当し、ユニット内部の計算ノードの接続に16 portが割り当てられ、残り8 portがMegaProto/E筐体外への接続として利用できる。

まず、複数のMegaProto/Eユニットを接続するために筐体外への接続8 portをVFREC-Netを用いて同時に利用する。これによってMegaProto 筐体間のバンド幅は1ネットワークあたり1 GB/sec(全二重2 GB/sec)となり、スイッチ間で発生するバンド幅ボトルネックは理論的には小さくなる。図5にMegaProto/Eを4 unit

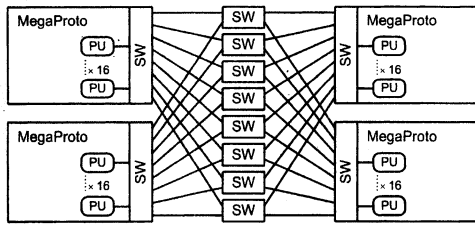


図5 VFREC-Netを用いた MegaProto/E の接続

表2 Ita クラスタ環境

|          |                                                              |
|----------|--------------------------------------------------------------|
| CPU      | Intel Xeon 3.0GHz 1-way (Hyper Thread ON)                    |
| Memory   | 1.0 Gbytes DDR2/400MHz (Dual Channel)                        |
| NIC      | Intel PRO/1000MT Dual Ports (MTU=6000)<br>PCI-X 64bit/133MHz |
| OS       | Linux Kernel 2.6.17                                          |
| Compiler | GCC ver.4.0                                                  |

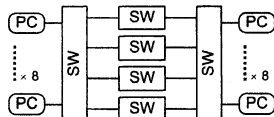


図6 Ita クラスタの接続

用いた場合に、計算ネットワークをどのように接続するかの一例（VBFT）を示す。実際には MegaProto/E には 2 系統の計算ネットワークがあるため、この接続とまったく同じ構成を 2 つ用意することになる。そして RI2N/UDP を用いて各々のノードに接続された 2 系統のネットワークを同時に利用する。この場合、各ノードを結ぶネットワーク性能の理論ピークは 250 MB/sec(全二重 500 MB/sec) となり、また 1 系統のリンクが故障した場合においても残りの 1 系統のリンクを用いて通信を継続可能となる。

#### 4. 性能評価

##### 4.1 評価環境

RI2N/UDP と VFREC-Net を組み合わせた場合の性能を評価する。評価環境として前述の MegaProto/E を 4 Unit (合計 64 ノード) 用いる。ところで MegaProto/E には、16 プロセッサを 1U の筐体に収めるために、使用するプロセッサに厳しい電力制限を必要とした。そのため、それらのプロセッサが GbE の処理能力を十分に引き出せないことが考えられる。そこで提案するネットワークシステムの実性能を測定するため、表 2 に示す 16 ノードのクラスタを併せて構築・評価する。以後これを Ita クラスタと呼ぶ。それぞれトポロジとして VBFT 構成を用い、MegaProto には前述の図 5、Ita クラスタには図 6 にそれぞれ示すネットワークを構築した。これらの環境を用いてスループット、耐故障性、遅延時間、バイセクションバンド幅を評価する。

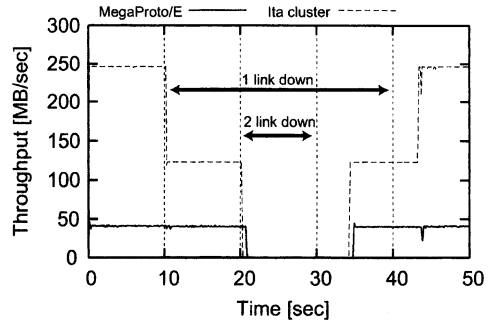


図7 スループットと耐故障性測定結果

##### 4.2 スループットと耐故障性

まず、単一ノード間のスループットの評価を行う。本評価では、純粋に RI2N による性能向上・耐故障性の評価を行うため、スイッチ間をまたがる通信ストリームを 1 つとし、VFREC-Net の効果は生じないようにした。通信性能の評価を行うと同時に、リンクまたはスイッチに故障が発生した場合の耐故障性のテストも行う。スループット計測中、意図的に時刻 10 秒の時点で 1 リンク、20 秒の時点で 2 リンク目を切断する。その後開始から 30 秒で 1 リンクを復帰させ、40 秒ですべての残りの 1 リンクを復帰させる。そのときのスループットの変化について計測する。測定結果を図 7 に示す。

まず、始めに Ita クラスタでの結果について考察する。全リンクが正常な時には、ノード間の最大スループットは約 247 MB/sec となった。GbE 2 リンク分の理論最大性能が 250 MB/sec であるため、十分に GbE の性能を引き出しており、RI2N/UDP の高バンド幅化が十分に機能していることが確認できた。また、同時に行った耐故障性の評価では、1 リンク故障した場合においても残りの正常なリンクを利用することで通信を継続し、その場合で約 123 MB/sec の性能を示している。このことから、RI2N/UDP は高バンド幅・耐故障性をユーザに提供できていると言える。

次に MegaProto/E での結果を考察する。結果では正常時のスループットが約 42 MB/sec となっている。この結果は前述の Ita クラスタでの結果と比較して非常に低い。この問題を考えるため本実験と併せて、RI2N/UDP を用いず通常の TCP/IP を用いたスループットを計測した。結果は Ita クラスタが約 123 MB/sec と GbE のほぼ理論最大性能を示しているのに対し、MegaProto/E では約 55 MB/sec の性能しか得られていない。この原因は、プロセッサの処理能力の差、およびメモリバンド幅であると我々は結論付けた。CPU やメモリの能力によって単一リンクの GbE を最大限に活用できていないならば、2 リンクを用いたとしても性能向上は見込めない。特に RI2N/UDP はユーザレベル実装であるため、システムレベル実装の TCP/IP の性能と比較し

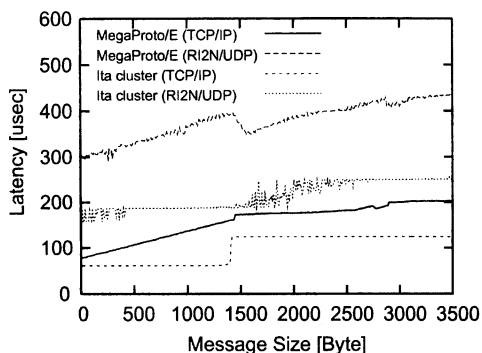


図8 遅延時間 測定結果

て性能は低くなることが容易に予想できる。このようなことから、MegaProto/E で RI2N/UDP を用いた場合のスループットの結果は妥当な値であると結論付けられる。しかし、MegaProto/E において RI2N/UDP を用いることは無駄ではなく、1 リンク故障時にも正しく通信を継続できており、耐故障性があることが確認できた。

#### 4.3 遅延時間

次に遅延時間について評価する。二つのクラスタ環境において、RI2N/UDP を用いた場合、および TCP/IP を用いた場合それぞれについて評価する。評価結果を図8に示す。

結果が示すように通常の TCP/IP を用いた場合と比較して、RI2N/UDP を用いた場合では遅延時間が大きく増加した。特に MegaProto/E で RI2N/UDP を用いた場合に悪く、通常の TCP/IP を用いた場合と比較して3倍以上となっている。RI2N/UDP では下位のプロトコルとして UDP/IP を用いているが、UDP/IP の遅延時間は TCP/IP のそれにほぼ等しい。RI2N/UDP では UDP/IP にさらに手を加えて通信を行うため、TCP/IP の遅延時間と比較して悪化することが容易に想像できる。しかし、この遅延時間の差は予想以上に大きい。この原因として考えられることが、RI2N/UDP のプロトコル設計が重いことに加え、MegaProto/E に搭載しているプロセッサ処理能力が不足していることである。RI2N/UDP はユーザレベル実装になっており、通常の TCP/IP と比較してプロセッサとメモリの処理能力を多く必要としている。そのために低速なプロセッサを使用する場合、性能が低下する恐れがある。それを示すように、MegaProto/E と比較して十分に高速なプロセッサとメモリを持つ Ita クラスタでは、遅延時間が MegaProto/E の半分程度になっている。この MegaProto/E と Ita クラスタの差は、プロセッサ・メモリ性能によるものである。一方で Ita クラスタの RI2N/UDP と TCP/IP の遅延時間の差は、プロトコル設計やユーザレベル実装による性能低下が原因であると考えている。今後、より低遅延なネットワークシ

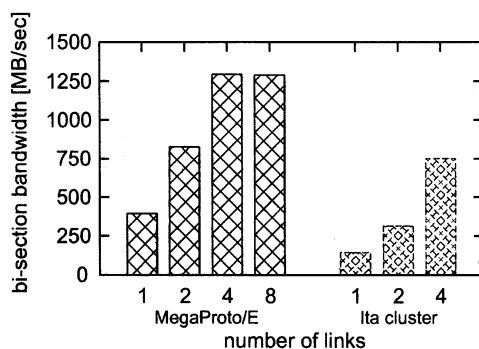


図9 バイセクションバンド幅 測定結果

テムを提供するためには、RI2N/UDP の設計方針やプロトコル設計について見直す必要がある。

#### 4.4 バイセクションバンド幅

最後にバイセクションバンド幅を評価する。評価では図5、図6に示した VFREC-Net で構築されたネットワークを用いて、すべての通信が左から右へと必ず最上部のスイッチを経由する通信とする。そして、スイッチ間に用意された複数の経路を偏りなく利用するように、VFREC-Net で制御する。VFREC-Net を利用することでスイッチ間バンド幅が正しく増加することを確認するため、MegaProto/E ユニット間を接続するスイッチ（リンク）数、または Ita クラスタのノードを接続するスイッチを接続するスイッチ（リンク）数を変化させた場合のバイセクションバンド幅を計測した。結果を図9に示す。

まず Ita クラスタの結果を考察する。結果が示すようにリンク数が増加することで、バイセクションバンド幅の性能が改善していくことがわかる。しかしその絶対値は必ずしも理論値どおりではない。各ノードで RI2N/UDP を利用していることから、本来“リンク数  $\times 2 \times 125$ ” MB/sec の性能があるはずである。前述のスループット計測で Ita クラスタが GbE の性能を最大限利用できているので、理論ピークに近いバンド幅が期待できるが、ここで示されている値はそれよりも低い。これは RI2N/UDP の実装がもっとも大きな原因だと我々は推測している。RI2N/UDP はフロー制御の判断材料を通信先ノードのバッファ空き容量のみにしている。バイセクションバンド幅の計測のようにネットワークが混雑することによって発生するパケット消失等を現在の RI2N/UDP では考慮していない。そのため、送信ノードはネットワーク状況を考慮にいかず常に最大限の送信能力を使い送信する。これが更なるネットワークの混雑を引き起こし、結果としてバイセクションバンド幅は低下していると推測した。スイッチ間リンク数が増加するとこの問題は徐々に緩和され、それが利用効率の改善につながる。この問題は、今後 RI2N/UDP の改良により改善できると期

待できる。MegaProto/Eでもリンク数が増加するにしたいが、バイセクションバンド幅が改善している。ここでもItaクラスタとおなじ問題が発生すると考えられるが、MegaProto/EのネットワークスループットがItaクラスタのノード間スループットが低いと、影響はItaクラスタよりも小さい。リンク数が4から8へと変化した際に、バイセクションバンド幅が改善していないが、これは各ノードがすでに最大スループットに達しているためである。

このようにVFREC-Netを用いることでバイセクションバンド幅の改善に効果があることがいえる。

## 5. 関連研究

Ethernetのバンド幅や耐故障性を高める技術として、IEEE 802.3ad<sup>7)</sup>がある。この技術は、すでに多くのOSやスイッチで実装されているため、ユーザは既存環境を維持したまま使用することが可能になっている。しかしIEEE 802.3adはHPCクラスタ向けに設計されておらず、接続形態やネットワークポロジに制約が多い。特にスイッチ間並列接続については同一スイッチ間での直接接続のみ許されている点が最大の問題点である。並列リンクもラウンドロビンによる送信といった単純な方法でしか実装されていない。このことから、HPCクラスタではIEEE 802.3adが最適に機能できないと思われる。

HPCクラスタ向けに開発されたネットワークシステムとしてPM Ethernet<sup>10)</sup>がある。PM EthernetはEthernet上でマルチリンクを使用した上で低遅延・高バンド幅を実現している。しかし、PM Ethernetは直接デバイスドライバを変更するため、使用できるNIC等に制限が生じるなど、システムへの依存が大きい。また、耐故障機能についても現在実装されていない。これに加えて、使用できるAPIもTCP/IPなどと違い独自プロトコルPMを用いるため既存のsocket APIを用いたプログラムが流用できないといった問題もある。

## 6. おわりに

本稿では、RI2N/UDPとVFREC-Netを併用することで高バンド幅・耐故障性を持ちつつ、かつ高い拡張性を有するPCクラスタ向けネットワークを提案した。このネットワーク環境を実現するクラスタシステムとしてMegaProto/Eを用意し評価を行った。その結果、MegaProto/Eでは、そのシステム能力の限界からRI2N/UDPの高バンド幅機能は十分に発揮できないが、耐故障性に関しては当初設計通りの機能が実現できていることが確認された。この問題はRI2N/UDPを改良することで改善できると期待している。またVFREC-Netを用いることで複数台のMegaProto/Eユニットを、ユニット間バイセクションバンド幅を維持しつつ拡張することが可能であることがわかった。これらの技術

を応用することで高いバイセクションバンド幅を維持しつつ数万台のノード規模にも拡張することが可能になると考えている。

**謝辞** 本研究を行うにあたり、貴重な助言・提言をいただいたCREST「メガスケールクラスタ研究チーム」のメンバに深く感謝します。本研究の一部は、科学技術振興事業団「戦略的創造研究推進事業(CREST)-情報社会を支える新しい高性能情報処理技術-『超低電力技術によるディペンダブルメガスケールコンピューティング』」および文部科学省科学研究費補助(基盤研究(C)17500031)による。

## 参考文献

- 1) Myricom, inc.: Myrinet.  
<http://www.myri.com/>.
- 2) InfiniBand Trade Association: InfiniBand,  
<http://www.infinibandta.org/>.
- 3) 岡本高幸, 三浦信一, 朴泰祐, 佐藤三久, 高橋大介: Ethernet マルチリンクによる PC クラスタ向け耐故障ネットワーク RI2N / UDP, 情報処理学会ハイパフォーマンコンピューティング研究会(HOKKE2006), pp. 85-90 (2006).
- 4) 三浦信一, 岡本高幸, 朴泰祐, 佐藤三久, 高橋大介: VFREC-Net: ドライバ制御による tagged-VLAN を用いた PC クラスタ向けマルチパスネットワーク, 先進的計算基盤シンポジウム (SACIS2006), pp. 117-125 (2006).
- 5) 中島浩, 中村宏, 佐藤三久, 朴泰祐, 松岡聡, 高橋大介, 堀田義彦: 高性能計算のための低電力・高密度クラスタ MegaProto, 情報処理学会論文誌コンピューティングシステム, Vol. 46, No. SIG12(ACS11), pp. 46-61 (2005).
- 6) Graham, R. L., Woodall, T. S. and Squyres, J. M.: Open MPI: A Flexible High Performance MPI, *Proceedings, 6th Annual International Conference on Parallel Processing and Applied Mathematics*, Poznan, Poland (2005).
- 7) IEEE: P802.3ad Link Aggregation Task Force.  
<http://grouper.ieee.org/groups/802/3/ad/>.
- 8) IEEE: 802.1Q - Virtual LANs.  
<http://www.ieee802.org/1/pages/802.1Q.html>.
- 9) 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. SIG06(ACS6), pp. 35-44 (2004).
- 10) Sumimoto, S. and Kumon, K.: PM/Ethernet-kRMA: A High Performance Remote Memory Access Facility Using Multiple Gigabit Ethernet Cards, *3rd IEEE International Symposium on Cluster Computing and the Grid (CCGrid 2003)*, IEEE Computer Society, pp. 326-333 (2003).