

## 高性能・耐故障マルチリンク Ethernet 結合システムの NFS への適用と評価

三浦 信一<sup>†1</sup> 埜 敏博<sup>†1</sup>  
朴 泰祐<sup>†1,†2</sup> 佐藤 三久<sup>†1,†2</sup>

コストパフォーマンスが高い Gigabit Ethernet は、比較的中規模な PC クラスタで多く用いられている。この Gigabit Ethernet を高性能化を実現するために IEEE 802.3ad や Linux Channel Bonding などの Ethernet Trunking 技術が多く用いられている。しかし、これらの技術は既存の TCP/IP を用いる場合、遅延時間やバンド幅などの性能が低い問題があった。また、これらの技術は耐故障機能も持っているが、それらは実用十分ではなかった。本問題を解決するためには、我々は Linux Channel Bonding と非常に近い実装で RI2N/DRV (Redundant Interconnection with Inexpensive Network with Driver) を開発している。RI2N/DRV は TCP/IP と親和性を高くし、遅延時間やバンド幅の性能を高めている。RI2N/DRV はその実装方法により高い汎用性を持ち、初期に想定した PC クラスタ向けの通信だけではなく、既存の UNIX のサービスにも適用することが可能である。本稿では既存の UNIX サービスに RI2N/DRV を適用する場合の問題点を整理し、実際に RI2N/DRV を UNIX サービスのネットワーク環境として構築した。この環境を用いて、UNIX のサービスの 1 つである NFS サービスを評価したところ、システムに一切の変更なしで、高いバンド幅と耐故障性を得ることが確認できた。

### Evaluation of NFS on Ethernet Multilink Bonding System for High performance and Fault-tolerance

SHIN'ICHI MIURA,<sup>†1</sup> TOSHIHIRO HANAWA,<sup>†1</sup> TAISUKE BOKU<sup>†1,†2</sup>  
and MITSUHISA SATO<sup>†1,†2</sup>

Although recent high-end interconnection network devices and switches provide a high performance/cost ratio, most of the small to medium sized PC clusters are still built on the commodity network, Ethernet. To enhance performance on commonly used Gigabit Ethernet networks, link aggregation or bonding technology is used. Currently, a Linux kernel is equipped with a software solution named Linux Channel Bonding, which is based on IEEE802.3ad Link Aggregation technology. However, standard LCB has the problem of mismatching with the commonly used TCP protocol, which consequently implies several problems of both large latency and instability on bandwidth improvement. The fault-tolerant feature is also supported, but the usability is not sufficient. We have developed a new implementation similar to LCB named RI2N/DRV (Redundant Interconnection with Inexpensive Network with Driver) for use on a Gigabit Ethernet with a complete software stack that is very compatible with the TCP protocol. Our algorithm suppresses unnecessary ACK packets and retransmission of packets even in imbalanced network traffic and link failures on multiple links. It provides both high-bandwidth and fault-tolerant communication on multi-link Gigabit Ethernet. We confirmed that this system improves the performance and reliability of the network, and our system can be applied to ordinary UNIX services such as NFS, without any modification of other modules.

#### 1. はじめに

近年の IP 技術の大幅な進歩により、多くの機器が IP を用いたネットワークに接続されている。特に一般的な PC では様々なサービスを提供するネットワークへ

の接続が必要不可欠になっている。一方でこれらのクライアントにサービスを提供するサーバシステムなどは、接続されるクライアントが多くなるに従って、トラフィックが急増するため、より高性能なネットワークが求められている。またこれらのシステムは、1台の計算機のみで完結して動作するものではなく、それぞれの計算機が互いに協調して動作するものも少なくない。そのため、これらのシステムの持続的な安定動作のためには、それらの計算機のみならずネットワークを含めたシステム全体の耐故障性も同時に高める必

<sup>†1</sup> 筑波大学 計算科学研究センター  
Center for Computational Sciences, University of Tsukuba  
<sup>†2</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

要がある。

現在、最も多くのシステムで用いられているネットワーク環境は Ethernet である。特に多くのサーバシステムでは、そのコストパフォーマンスの高さから Gigabit Ethernet (以下、GbE) が幅広く用いられている。しかしより大規模な環境においては、Gigabit Ethernet よりも高バンド幅を持つ 10 Gigabit Ethernet (以下、10GbE) など既存の Ethernet 技術を拡張した高速ネットワークも用いられる。また、InfiniBand<sup>1)</sup> や Myrinet<sup>2)</sup>、もしくは Fiber Channel など、比較的高速なネットワークも、ストレージやデータベース等のネットワークに用いられていることも多い。しかし、このようなバンド幅が高いネットワークは、コストパフォーマンスが良くなりつつあるが、現在最もコストパフォーマンスが高い GbE と比較して高価であり、ネットワーク構築のコストが問題となる。加えて、これらのネットワークシステムは単純に構成しただけでは、故障に対して脆弱であり、何らかの冗長性を確保する必要がある。

このような中、我々は高いバンド幅と耐故障性を持つ高性能クラスタ向けネットワーク RI2N<sup>3)</sup> を研究・開発してきた。RI2N はマルチリンクの Ethernet を用いて、平常時にはそれらの複数の Ethernet を同時に利用することにより高いバンド幅を実現する。また、ネットワークの異常時においては、ネットワークの冗長性を利用して正常なネットワークのみを用いて通信が継続することを可能にする。我々は、この RI2N のコンセプトを元にユーザ透過に本機能を提供可能な RI2N/DRV を開発している<sup>4)</sup>。本来 PC クラスタ向けの通信機構として開発した RI2N/DRV は、ユーザ透過性を向上させるために、TCP/IP などの環境をそのまま適用可能であり汎用性が非常に高い。現在までの RI2N/DRV は、PC クラスタでの MPI などの利用を想定して開発してきたが、RI2N/DRV の持つ汎用性の高さから、より一般的な Linux ネットワークシステムにも適用可能である。本稿では、高性能クラスタ向けのネットワークとして研究・開発してきた RI2N/DRV を、より一般的な UNIX 環境のネットワークに適用し、その有効性について検証を行う。具体的な応用例として、一般的なネットワークファイルシステムである NFS のネットワークに適用し、RI2N/DRV によって得られるバンド幅と耐故障性について評価する。

次章では、RI2N の概要と既存研究について説明し、Linux ドライバによる実装である RI2N/DRV について説明する。3 章では本システムを既存の UNIX 環境のネットワークに応用した例について示し、NFS への適用について説明する。その後 4 章の評価では、RI2N/DRV を NFS サービスに適用しその性能について評価する。

## 2. RI2N

### 2.1 概 要

我々は高性能クラスタ向けに複数リンクの Ethernet を同時に利用することによって高いバンド幅と耐故障性を同時に実現する RI2N (Redundant Interconnection with Inexpensive Network) というコンセプトを提唱し、それを実現するシステムを提案・実装している<sup>3)5)</sup>。RI2N とは、安価な複数リンクの Ethernet とソフトウェアの拡張のみで高バンド幅化と信頼性の向上を同時に実現を目指すものである。具体的には各ノード間に複数リンクの Ethernet ネットワークを設置し、正常時にはデータのストライピングによってスループットを向上させる。そしてリンクが故障した場合には、冗長なリンクを利用して通信を継続させる。

このような機能は一般的には Ethernet トランッキングと呼ばれるものであり、いくつかの先行研究が挙げられる。まず高性能クラスタでの利用に特化した軽量通信ライブラリ PM/Ethernet<sup>6)7)</sup> がある。PM/Ethernet は複数のネットワークを同時に利用する機能として、PM/Ethernet Network Trunking を持っており、遅延時間、スループットにおいて高い性能を得ることができる。一方で PM/Ethernet は既存の UNIX Socket とは互換性のない専用の通信体系と API を用いるため、プログラムの可搬性や相互運用性に大きな問題がある。ハードウェアもしくはソフトウェア機能により、マルチリンク Ethernet を用いるシステムとして、すでに IEEE 802.3ad<sup>8)</sup> によって規格されている Link Aggregation Control Protocol がある。本技術は主に 2 台のスイッチ間もしくは、ノードとスイッチ間に 2 つ以上のネットワークリンクを用意し、高バンド幅と耐故障性を同時に実現する。しかし、本技術はスイッチの冗長化は考慮にいれておらず、スイッチ自体の故障に対処することは難しい。また、専用スイッチの導入も必要とする。最も RI2N コンセプトに近いものとして、Ethernet におけるマルチリンクの利用をノード間にも適用するドライバソフトウェアとして Linux Channel Bonding<sup>9)</sup> (以後、LCB) がある。LCB はいくつかの通信モードがあり、経路の多重化との 2 ノード間の高バンド幅を同時に実現するモードとして balance-rr モードがある。しかし本機能はパケット到着順序の入れ替わりが頻繁に発生し、TCP/IP では性能を十分に発揮することが難しいことが示されている<sup>9)</sup>。また、LCB では ARP 機能を用いてリンクの故障判断を行っているが、検出可能な範囲が限定され、耐故障性については十分に配慮されていない。

これらの問題を解決するために、我々は RI2N のコンセプトを実現する 1 つの実装として RI2N/DRV を開発している。

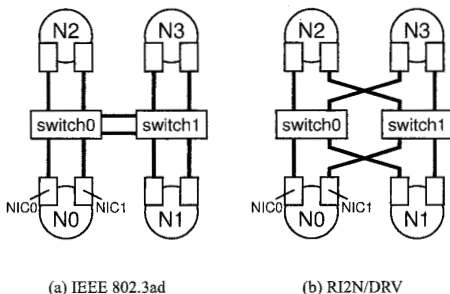


図 1 IEEE802.3ad でのネットワーク構成と RI2N でのネットワーク構成

## 2.2 RI2N/DRV の実装

RI2N/DRV は RI2N のコンセプトに基づき、LCB と同様に仮想的な Ethernet デバイスとして実装されている。RI2N/DRV は Linux のロードブルモジュールで実装されており、OS に対して既存の Ethernet デバイスドライバのように振る舞う。そして既存の Ethernet 環境の上位プロトコルである TCP/IP、UDP/IP もしくは ARP などのプロトコルを、一切の変更なしで利用することを可能にする。

図 1(a) に示すような IEEE 802.3ad を用いたネットワーク構成と異なり、RI2N/DRV では、スイッチの故障にも対応するために、図 1(b) のようにスイッチについても多重化し、2 ノード間で全く異なる 2 つ以上の経路を用意することで 1 台のスイッチが故障した場合にも通信が継続できるようにする。

システム全体の基本的な仕組みは Linux Channel Bonding で `balance-rr` モードを用いた場合と類似する。しかし、LCB や IEEE802.3ad でのリンク結合のように非常に単純な処理のみを行うのではなく、RI2N/DRV では順序入れ替えのような比較的複雑な処理を行い、ある程度のオーバーヘッドを許容した上で性能改善を目指している。RI2N/DRV は、基本的に上位レイヤにどのようなプロトコルが使われても対応可能であるが、現在最も利用されている TCP/IP の挙動を想定した実装となっている。以後の説明では、TCP/IP の使用を想定し説明する。

RI2N/DRV の重要な機能として、**パケット到着順序制御機構**と**故障/回復検出機構**の 2 つがある。次にこの 2 つの機構について説明する。

### パケット到着順序制御機構

マルチリンクのネットワークを用いる場合、物理的な距離やネットワークの混雑状態により、パケット到着順序に不整合が生じる場合がある。下位レイヤで比較的パケットが順序正しく到着することを期待する上位プロトコルでは、これは性能低下の原因になる。現在よく用いられている通信プロトコルである TCP/IP は、パケット到着の順序の狂いを許容できる仕組みを

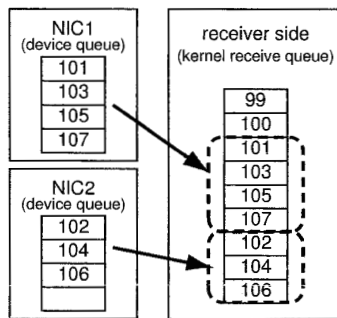


図 2 複数の NIC を用いた場合のバッファによるパケット到着順序の不整合

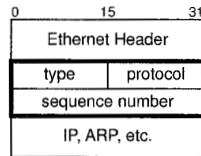


図 3 RI2N ヘッダ

持つ。しかし、定期的な到着の順序の狂いが発生するマルチリンクの環境では、この仕組みを用いてもスループットの低下を招き、またこれを回避するためにプロセッサに高負荷が生じる。この順序の不整合の最も大きな原因は、図 2 に示すような、NIC が提供する **interrupt coalescing** 機能<sup>10)</sup> による受信バッファ蓄積である。基本的にこの順不整合を、送信や受信の工夫のみで取り除くことは難しい。RI2N/DRV ではこの問題を取り除くために、新たに図 3 に示す RI2N ヘッダを導入し、受信側で順序制御を行うことでこの不整合を取り除く。このヘッダは Ethernet ヘッダ以降のペイロード部に格納されるため、使用できるスイッチや NIC などのハードウェアに制限がない。RI2N/DRV では、このヘッダを用いて簡略された順序制御を行うが、再送制御や輻輳制御などは行わない。これらの機能は、上位層のプロトコルである TCP/IP などに期待する。

図 4 に RI2N/DRV の基本的な送受信の流れを示す。TCP/IP のパケットが RI2N/DRV によって作成された仮想的なデバイス (以下 RI2N デバイス) に送られる。RI2N デバイスではこのパケットを **round-robin** で実在する複数の NIC を用いて送信する。受信側では、各 NIC で受信されたパケットを RI2N デバイスで集約し、送信された順番にパケットを並べ替えた後に IP 層などの上位層にパケットを渡す。

### 故障/回復の検出機構

マルチリンクのネットワークで 1 系統のネットワークに故障が発生すると  $1/n$  ( $n$  はネットワークの冗長度) の確率でパケットが損失される。TCP/IP では、パ

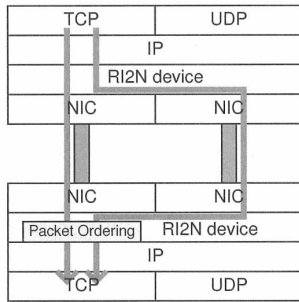


図 4 RI2N/DRV の仕組み

ケット損失が確認されると混雑が原因であると推測し、congestion control により window サイズが低下する。しかし、パケット損出の原因が混雑ではなく故障である場合には、window サイズが低下したとしても一定の確率でパケット損出が継続する。それによって、いつかは window サイズがきわめて 0 に近づきネットワークが停止と同様の状態になる。これを回避するために、なるべく早期にネットワークの故障を検出する必要がある。そこで RI2N/DRV では、バースト転送時にはパケット到着数の偏りを用いて故障の検出を高速化している。また、バースト転送状態以外においても故障を検出する機構として、ハートビートも併用している。RI2N/DRV では、このパケット到着の偏りとハートビートの 2 種類の故障検出機構を利用することで、様々な故障のパターンに対応できる。

故障の検出に加えて RI2N では、故障からの回復も検出する。回復の検出にはハートビートを用いる。回復には一般的に人手を要するため、検出に比較的時間を要するハートビートで十分である。

### 3. RI2N/DRV の応用検討

前述のように RI2N/DRV では、上位層のプロトコルとは関係なく、どのようなプロトコルにも理論的には対応可能である。とくに TCP/IP は、2.2 節に示したように、設計段階で想定されているプロトコルである。これにより RI2N/DRV は既存の並列計算で用いられている MPI の通信に限らず、一般的な UNIX のサービスなどの通信にも使用することが可能である

現実的な運用では、すべてのノードが冗長にネットワークに接続されているとは限らず、システムの規模や重要度に応じて、通信性能と冗長度を考慮し、各ノード毎のネットワークを構成できることが望ましい。RI2N/DRV では、各々の通信ノード毎に、使用可能なネットワークを管理している。そのため図 5 に示すような、複数リンク接続と 1 リンクのみ接続が混在するようなネットワーク構成（以後、『単数/複数リンク混合ネットワーク』と呼ぶ）であった場合でも正常な

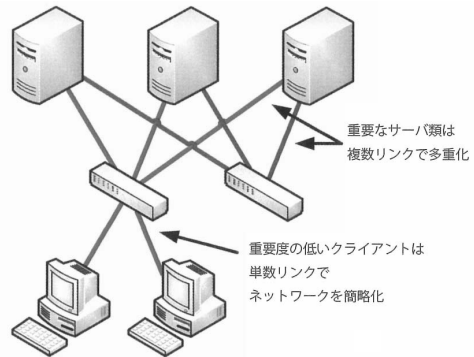


図 5 単数/複数リンク混合ネットワーク構成

通信を行うことが可能である。LCB を用いた場合でもネットワークに耐故障性を持たせることが出来るが、LCB の balance-rr モードでは 16 ノード以上の環境で故障を検出することができない。また、既存の LCB の balance-rr モードでは、使用可能なネットワークをノード毎に管理しておらず、いずれか 1 ノードのネットワークが 1 系統でも故障すると、正常なノードとの通信の場合でも、そのネットワークは使用されない。その結果 LCB では単数/複数リンク混合ネットワークに対応することは難しい。

これに対し、RI2N/DRV では単数/複数リンク混合ネットワークに対応できるが、以下のような制限が伴う。

- 単数/複数リンク混合ネットワーク構成をとる場合、少なくとも 1 つのネットワークには必ず全ノードを結合している必要がある。
- RI2N/DRV 独自のヘッダを利用するため、すべてのノードは RI2N/DRV の機構を介して通信する必要がある。

以上の制限を満たすことで、既存の Ethernet を用いたネットワークの代替として、RI2N/DRV を高バンド幅と耐故障の持つネットワークとして利用することが可能となる。RI2N ヘッダを導入するために、既存の Ethernet ヘッダ中のプロトコル番号を RI2N/DRV 独自の番号に変更している。そのため RI2N/DRV で構築したネットワークは、一種の VLAN<sup>11)</sup> を用いたネットワークのように振る舞う。これを用いることで、RI2N/DRV 独自の機能を用いたネットワークと、RI2N/DRV を利用しない既存の Ethernet ネットワークを、2 つの異なる IP サブネットのネットワークとして共存させることも可能である。

#### NFS への適用

以上のように、RI2N/DRV を既存の UNIX 環境のネットワークに適用することは問題ない。しかし、RI2N/DRV が提供するマルチリンクの Ethernet の持つ、高いバンド幅を有効に活用できるネットワークサービスは限られる。我々は現状のネットワークサー

表 1 評価環境

Item	Specification
CPU	Intel Xeon 5110 1.6GHz dual core
Memory	DDR2 2048MB
OS	Linux kernel 2.6.22
GbE NIC	Intel PRO1000PT dual port 1000base-T
GbE NIC driver	Intel PRO/1000 Network Driver 7.3.20-k2-NAPI
10 GbE NIC	Myri-10G PCI Express NIC
10 GbE NIC driver	Myricom 10G driver 1.3.2-1.287
Bonding driver	Ethernet Channel Bonding Driver, v3.1.3

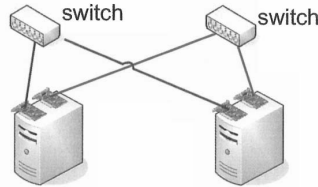


図 6 評価システム構成

ビスで最も RI2N/DRV の持つ性能を活用できるのは、ネットワークファイルシステムである NFS であると考えた。現在、最も利用されている NFSv3 では、サーバ・クライアント間の接続に TCP/IP が利用されている。そのため、現状の NFS への RI2N/DRV の適用については技術的な問題はない。RI2N/DRV を用いることにより、既存の 1 リンクの Ethernet 環境を用いた場合よりも、高バンド幅かつ耐故障性のあるネットワークを構築することが可能になる。また既存技術である LCB を利用した場合と比較しても、少ないオーバーヘッドで冗長化したネットワークを有効利用できることが期待できる。

#### 4. 評価

評価として、RI2N/DRV を用いて NFS サーバ・クライアント環境を構築し、同システム上での性能を評価する。評価環境として表 1 に示す 2 ノードの計算機を用意しシステムを構築した。評価に用いたネットワーク構成を図 6 に示す。ノード間はスイッチ (DELL PowerConnect 5224 Gigabit Ethernet) で接続した 2 系統のネットワークを用意し、通常の GbE では片方のネットワーク、LCB と RI2N/DRV の環境ではこの 2 系統のネットワークを同時に用いる。また、ノード間を直接 10GbE で接続した環境についても評価を行った。なお本評価では、Intel Pro1000PT のドライバ機能である遅延割込を有効に設定し評価する。今回評価を行ったマルチリンクの Ethernet では 2 系統の Gigabit Ethernet で構築されたネットワークで接続されたネットワークを用いる。そのため、最大性能で 200MB/sec 以上の性能が得られる。一般に、これに見合う十分な

表 2 NFS v3 (TCP) Sequential Output

Condition	rwsize=8 KB		rwsize=32 KB	
	MB/sec	%CPU	MB/sec	%CPU
10GbE	297.0	29	316.4	24
GbE	107.9	8	109.8	7
LCB	153.2	16	179.9	15
RI2N/DRV	131.6	15	157.2	15

表 3 TCP/IP 片方向バースト転送時の最大スループット<sup>4)</sup>

Condition	スループット [MB/sec]
GbE	112
LCB	222
RI2N/DRV	222

I/O 性能を備えたストレージは高機能な RAID システムとなる。本評価ではそのようなストレージを用意できないため、これを NFS サーバ上にメモリファイルシステムである tmpfs を用いて 1.5 GB のファイルシステムを構築し、NFS によって export する領域として用いる。本環境でファイルシステムの直接 I/O 性能を調べたところ、約 1.4 GB/sec の write 性能が得られ、すべてのネットワーク環境下で十分な I/O 性能を得ることができ、ストレージの性能が性能低下の原因にならないことが考えられる。

このような NFS の環境において、NFS write 性能と耐故障性能の 2 つを評価する。

##### 4.1 NFS write 性能

NFS クライアントの write block size として 8 KB, 32 KB をオプションとして指定した 2 つの場合について評価する。評価では bonnie++ v1.03<sup>12)</sup> を用いて 1.4 GB のサイズを持つファイルの sequential write の性能を評価する。表 2 に性能評価の結果を示す。

LCB, RI2N/DRV の環境を GbE の環境と比較すると、22%から 64%性能向上を得られた。先行研究<sup>4)</sup>では TCP/IP を用いた片方向バースト転送時の最大スループットで表 3 に示す評価結果が得られている。この結果と比較すると、GbE の性能に対し、2 つの環境の性能向上の割合が小さい。しかし、LCB, RI2N/DRV いずれの結果でも、block size が 8 KB から 32 KB に増加することで性能向上することから、原因は NFS の block size であることが推測できる。これは、GbE の性能の 10 倍のバンド幅を持つ 10GbE ですら、最大性能の 30%の性能しか得られていないことから推測できる。10GbE と比較すると LCB, RI2N/DRV の 2 つの環境は、43%~56%低い性能を示すが、10GbE の現在の導入コストを考えると、マルチリンク Ethernet のコストパフォーマンスは非常に高いことがわかる。

LCB と RI2N/DRV の性能を比較すると、block size が 8 KB の場合 12%、32 KB の場合 14%性能が低い。現在の TCP/IP では、順序の不整合が続くとそれを学習する仕組みがある。この NFS のテストでは 1.4 GB という比較的大きい書き込みを行うため、一方の

表 4 故障前後の NFS の性能

	正常時性能 [MB/sec]	故障時性能 [MB/sec]
LCB	179.9	108.7
RI2N/DRV	157.2	109.0

バースト転送にほぼ同じとなる。そのため、順序の不整合による LCB のデメリットが小さくなる。一方で、RI2N/DRV では順序制御や RI2N ヘッダの処理により LCB よりも、明らかに処理量が多くなる。これらの RI2N/DRV が行う処理のオーバーヘッドが、NFS の処理に何らかの影響を与えている可能性がある。しかし、本性能評価で得られた LCB と RI2N/DRV の性能差については、まだ原因が明らかではない。今後、本問題を解決するために、より詳細な評価を行う必要がある。

#### 4.2 NFS 耐故障性能

NFS サーバに対してクライアントが write リクエストを継続中に、2 系統あるネットワークのうち、1 系統のネットワークを物理的に切断し故障を発生させる。その後 10 秒後に切断させたネットワークを再度接続しネットワーク状態を復帰させる。表 4 に故障前後における NFS クライアントが得られたスループットの結果を示す。

結果として LCB および RI2N/DRV 共に故障中の性能は一時的に低下したが、停止することなくファイルの I/O 処理を継続することが確認できた。先行研究<sup>4)</sup>では、ネットワーク故障の瞬間には数秒の停止時間が見られたが、数秒の停止時間であれば TCP はタイムアウトすることなく、かつ NFS クライアント側のバッファリングの効果により接続が維持される。また、ネットワークの切断回復後は、NFS の処理性能が回復することも確認できた。

このように LCB も RI2N/DRV も耐故障機能が機能することを確認できた。しかし、LCB の故障検出は 16 ノードという制限がある。一方で RI2N/DRV は故障検出にノード数の制限はない。また、LCB は故障情報をノード毎に管理しないため、一部ノード間のネットワークに故障を検出すると、正常な他のノード間の通信でも、故障と検出されたネットワークを用いない。しかし、RI2N/DRV ではそれらの故障情報はノード毎に管理しているため、正常なノード間の通信では高いスループットで通信し続けることができる。このことから、耐故障性については RI2N は LCB よりも高いと考えられる。

## 5. おわりに

本稿では、PC クラスタ向けに開発されたマルチリンク Ethernet 環境を用いた高バンド幅かつ耐故障性を持つネットワークシステムである RI2N/DRV を、UNIX のネットワークサービスへの応用について検討した。RI2N/DRV は既存技術である LCB と比較して、ネッ

トワークの拡張性が高く、現実的なネットワークシステムとして既存の UNIX サービスに適用することが可能である。本システムをネットワークファイルシステムの 1 つである NFS に RI2N 適用したところ、既存のシングルリンクの GbE をそのまま利用する場合と比較して、高いスループットと耐故障性を得られることが確認できた。一部の性能評価では、LCB よりも性能が低い問題点があったが、今後は詳細な評価を行い、より一層の性能向上を目指す。

**謝辞** 本研究の一部は、科学技術振興機構戦略的創造研究推進事業 (CREST) 研究領域「実用化を目指した組込みシステム用ディペンダブル・オペレーティングシステム」、研究課題「省電力高信頼組込み並列プラットフォーム」による。

## 参考文献

- 1) InfiniBand Trade Association: InfiniBand.
- 2) Myricom: Myrinet.
- 3) Miura, S., Boku, T., Sato, M. and Takahashi, D.: RI2N - Interconnection Network System for Clusters with Wide-Bandwidth and Fault-Tolerance Based on Multiple Links, *ISHPC-V*, Lecture Notes in Computer Science, Vol. 2858, Springer, pp. 342–351 (2003).
- 4) 岡本高幸, 三浦信一, 朴 泰祐, 埴 敏博, 佐藤三久: ユーザ透過に利用可能な高性能・耐故障マルチリンク Ethernet 結合システム, 情報処理学会論文誌. コンピューティングシステム, Vol. Vol.1 No.1, No.8, pp.12–27 (2008).
- 5) 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 クラスタ向け高バンド幅・耐故障ネットワーク RI2N/UDP, 情報処理学会論文誌. コンピューティングシステム, Vol.48, No.8, pp.153–164 (2007).
- 6) Sumimoto, S. and Kumon, K.: PM/Ethernet-kRMA: A High Performance Remote Memory Access Facility Using Multiple Gigabit Ethernet Cards, *CCGrid 2003*, pp.326–333 (2003).
- 7) Sumimoto, S., Ooe, K., Kumon, K., Boku, T., Sato, M. and Ukawa, A.: A scalable communication layer for multi-dimensional hyper crossbar network using multiple gigabit ethernet, *ICS '06: Proceedings of the 20th annual international conference on Supercomputing*, pp.107–115 (2006).
- 8) IEEE: IEEE 802.3ad – Link Aggregation (2000).
- 9) Davis, T.: Linux Ethernet Bonding Driver.
- 10) Intel Corporation: *Intel PRO Network Connections User Guides*.
- 11) IEEE: IEEE 802.1q – Virtual LANs.
- 12) Coker, R.: Bonnie++ : benchmark suite of hard drive and file system performance.