

拡張 ARP による複数 NIC を用いた広帯域通信*

東京電機大学 理工学部 情報システム工学科

斉藤 輪 小野 真和 桧垣 博章[§]

1 背景と目的

近年、コンピュータネットワーク技術の発達にともない、高品質な音声動画像を含むマルチメディアデータのリアルタイム配送などの広帯域通信に対する要求が高まっている。例えば、LAN 内における非圧縮 D1 や HDTV といった高品質なマルチメディアデータのリアルタイム配送システム [4] や医用画像管理システム [5] 等の研究開発が行われている。非圧縮 D1 と HDTV データの配送にはそれぞれ約 270Mbps、約 1.2Gbps の帯域を必要とする。

一方、スイッチネットワーク技術の普及によって、各端末に装着した複数の NIC(Network Interface Card) を同一 LAN に接続することにより、複数経路を用いた広帯域通信が可能となっている。複数経路を利用する手法として Bonding Device [3] や IEEE802.3ad [2] がある。しかし、Bonding Device は送信フレーム群を複数 NIC に分散できるが、単一 NIC ですべてのフレームを受信しなければならないため、通信帯域幅は受信 NIC の帯域幅を越えることはできない。一方、IEEE802.3ad では送信側と受信側の双方で複数の NIC にフレーム群を分散することができる。しかし、各通信には 1 対の NIC を対応させなければならないため、個々の通信に対する広帯域化が実現されない。

そこで、通信相手が装着している複数の NIC を動的に調査し、送信先への複数経路を検出する手法と、単一の通信要求から発生するフレームを複数経路を使用して送受信する手法が必要とされる。ただし、現在運用されている LAN へ導入するためには、以下の条件を満たすことが求められる。

[要求条件]

1. 提案する機構を導入しても、既存のアプリケーションへの変更を行うことなくそのまま利用できる。これを実現するために同一のコンピュータに装着される複数の NIC には同一の IP アドレスを与える。(これらの MAC アドレスは異なる)
2. すべてのコンピュータに提案する機構が導入されていることを前提としない。提案機構が導入されているコンピュータと導入されていないコンピュータとが混在していても、TCP/IP による通信が正しく行われる。

本論文では、これらの要求条件のもとに IP を用いたイーサネット LAN において複数経路を利用した広帯域通信を実現する拡張 ARP プロトコルを提案し、UNIX 系オペレーティングシステムに実装する。

2 ARP

イーサネット LAN において、各 NIC には MAC アドレスと IP アドレスがそれぞれ割り当てられている。通信データを含むフレームは MAC アドレス宛に送信

されるため、送信元は送信先 MAC アドレスを何らかの方法で取得しておく必要がある。そこで、IP アドレスをもとに MAC アドレスを取得し、対応付けを行う方法として ARP [1] が主に用いられている。ARP は送信先の MAC アドレスを取得するため、送信先の IP アドレスを格納した ARP リクエストを LAN 内にブロードキャストする。ARP リクエストを受信した端末の中で要求された IP アドレスをもつ送信先の端末は、対応する MAC アドレスを格納した ARP リプライを送信元へ返送する。

Linux カーネル 2.6.17 および 2.4.31 に実装された ARP においては、各 NIC に与えられた IP アドレスに関わらず、送信先は ARP リクエストを受信した NIC を用いて ARP リプライを送信する。このため、送信先に同一の LAN に接続する複数の NIC が装着されている場合には、それぞれの NIC の MAC アドレスを含む複数の ARP リプライが返送される(図 1)。このため、送信先において単一の ARP リクエストに対する複数回の応答処理によるオーバーヘッドを要することに加え、LAN のトラフィック増が問題となる。また、送信元は送信先の NIC 枚数をあらかじめ入手することができないため、適切な受信待ちタイムアウトを設定することができず、ARP 処理時間の増加や送信先 MAC アドレスの一部欠落と言った問題が発生する。



図 1 複数 ARP リプライの返送

3 拡張 ARP

イーサネット LAN における複数経路通信を実現するためには、送信先の単一 IP アドレスに対応する複数の NIC の MAC アドレスを送信元が取得しなければならない。しかし、ARP は IP アドレスと MAC アドレスを 1 対 1 で対応付けるため、送信先の 1 つの IP アドレスからは 1 つの MAC アドレスしか検出することができない。そこで、送信先の IP アドレスに対応する複数の MAC アドレスを格納した単一の ARP リプライを単一の NIC からのみ返送する拡張 ARP を提案する。

拡張 ARP では、ARP リクエストと ARP リプライに装着 NIC 枚数、2 枚目以降の NIC の MAC アドレスとその帯域幅等の拡張データを追加する(図 2)。この追加部分は、拡張 ARP に対応した端末が受信した場合のみ解釈され、拡張 ARP に対応していない端末は拡張部分を無視してアドレス解決を行うため、従来どおりの通信が可能である。ただし、拡張 ARP に対応していない端末が拡張 ARP メッセージを受信した場合には ARP メッセージ部分のみが解釈可能であることから、ここに格納された 1 枚目の NIC の MAC アドレスが通信データを含むフレームの配送に用いられることになる。そこで、拡張 ARP メッセージの 1 枚目の NIC には、これを送信する端末に装着された最大帯域幅の NIC を選択することとする。このような NIC が複数存在する場合に

* Extended ARP for high Throughput LAN of Computers with Multiple NICs

† Tokyo Denki University

‡ Rin Saito, Masakazu Ono, Hiroaki Higaki

§ {rin, masa, hig}@higlab.net

は、ラウンドロビンなどのアルゴリズムによって選択するものとする。

同一 LAN に接続する複数の NIC を装着した端末が拡張 ARP によって ARP リプライを送信する場合、いずれの NIC を用いるかが問題となる。ARP の応答時間を短縮する目的として最初に ARP リクエストを受信した NIC を用いる方法が考えられるが、異なる NIC で受信された ARP リクエストの同一性を確認する手続きとこの手続きのために受信した ARP リクエストを一定時間キャッシュに保存することが必要となり実装オーバーヘッドが大きい。そこで、拡張 ARP の実装においては、ARP に用いる NIC を固定することとする。ここでは、端末に装着された使用可能な NIC の中で最大帯域幅であるもののひとつを用いることとする。従来の Linux では、カーネルが NIC を認識した順にルーティングテーブルが作成され、ルーティングテーブルの上位に登録された NIC が ARP メッセージの交換に用いられている。拡張 ARP では、NIC の帯域幅を検出し、ルーティングテーブルの送信インタフェースを最大帯域幅の NIC に書き換えることで、この NIC からの ARP リプライの送信を実装する。なお、NIC の使用可能性を判定するために `ifconfig` の実行結果を用いる場合、ネットワークケーブルの切断等を正しく反映することができない。これによって選択された NIC から ARP リプライが返送されない問題が発生し得る。そこで、Linux における NIC のデバイスドライバから物理ポートの接続状況を取得する機能である Link-Beat Detection を利用する `ifplugd` によって NIC の使用可能性を判定する。

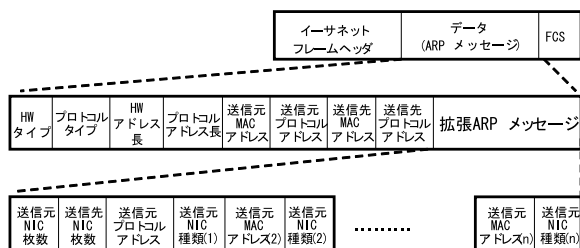


図2 拡張 ARP メッセージのフォーマット

4 複数 NIC による広帯域通信

3章で述べた拡張 ARP を用いることにより、送信元は送信先に装着された複数の NIC の MAC アドレスとその帯域幅を得ることができる。そこで、各送信元 NIC をいずれかの送信先 NIC に対応づけることによって複数経路を構成し、広帯域通信を実現することができる。ただし、各端末に装着されている NIC の枚数は必ずしも同一ではなく、それぞれの帯域幅も異なることが考えられる。

ARP リプライを受信することによって送信先に装着されている NIC の枚数と各 NIC の帯域幅を取得した送信元は、端末間の総帯域幅がより広くなるように送信元 NIC と送信先 NIC を対応付け、これに基づいてイーサネットフレームを送信する (図3)。NIC の対応付けは、同一帯域幅であるものを対応づけることを基本とし、これで対応付かない NIC については異なる帯域幅での対応付けを行う。さらに、NIC 枚数が異なる場合には、複数の狭帯域 NIC を単一の広帯域 NIC に対応付けることにより、より広帯域な通信が可能とする。なお、NIC の対応付けは送信側が受信側とは独立的に行うため、双方向通信時には各方向で異なる対応付けを行うこともあり得る。

この NIC の対応付けに従って OS が送信フレームを各 NIC に分配し、各 NIC が送信先 NIC にフレームを送信する。しかし、フレームを各 NIC に等分配すると

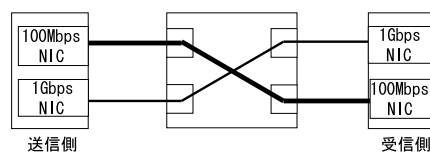


図3 帯域幅情報による NIC の対応付け

それぞれの通信経路が同一帯域幅であるとは限らないために、図4のように狭帯域 NIC からのフレーム送信時間と広帯域 NIC からのフレーム送信時間の差により広帯域経路に空き時間が発生し、帯域幅が縮小する。そこで、各経路の帯域幅に応じてフレーム配分率を変化させる機構を追加した。広帯域の経路により多くのフレームを分配することで、図5のように、広帯域経路の空き時間の発生を抑制し、送信先において全フレームをより短時間で受信することが可能となる。

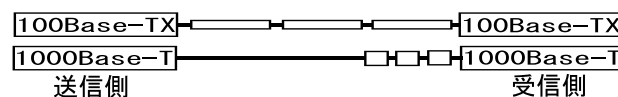


図4 フレーム等分配による総帯域幅の縮小

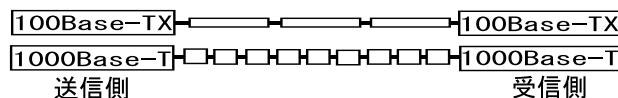


図5 経路帯域幅に応じたフレーム分配による総帯域幅の拡大

3章で提案した拡張 ARP プロトコルと4章で提案した広帯域通信プロトコルを Fedora Core Linux カーネル 2.6.17 に実装した。

5 まとめと今後の課題

イーサネット LAN において複数経路を用いた広帯域通信を実現するための拡張 ARP プロトコルを提案し、Linux カーネルに実装した。拡張 ARP によって取得した複数の送信先 MAC アドレスを用いた広帯域通信手法は、送信元端末が単一の送信先端末と通信することを前提としている。今後は、n 端末が相互にフレームを交換するネットワークアプリケーションを対象とするための適切なフレーム分配手法を考案する。また、FreeBSD カーネルへの実装を行う。

参考文献

- [1] David C. Plummer, "An Ethernet Address Resolution Protocol" RFC826 (1981).
- [2] IEEE Project 802.3 Working Group, "Aggregation of Multiple Link Segments," <http://grouper.ieee.org/groups/802/3/ad/> (2000).
- [3] Thomas Davis, "Linux Kernel Documentation," <http://www.mjmwired.net/kernel/Documentation/networking/bonding.txt>.
- [4] 勝本, 木俣, 櫻田, 北口, 杉浦, 中川, "超高画質画像配信システムにおけるコンテンツ管理の提案," 情報処理, Vol. 2001, No. 70, pp. 55-62 (2001).
- [5] 加藤子, 武藤, 江本, 岩田, "医療情報共有のための医用画像管理システムの開発" <http://mars.elcom.nitech.ac.jp/dicom/paper/VBLabst.html>.