

研究論文

キャンパスネットワークにおける スイッチングハブの輻輳制御方式に関する考察

小野 悟^{†1} 萬代 雅紀^{†2} 渡辺 尚^{†1}

一般にキャンパスネットワークにおける幹線から支線への拡張手段として、スイッチングハブの多段接続が用いられることが多い。このような接続形態では、ハブ間の接続ポート部に通信データが集中し、輻輳が発生しやすいという問題点がある。しかし、輻輳の発生時には TCP ウィンドウ制御など通信の抑制による方法がほとんどであり、ネットワークのスループットが低下することは避けられない。一方で通信効率を向上させるために複数のパケットを排他的論理和で多重化するネットワークコーディングが注目されている。本研究では、スイッチングハブの多段接続による潜在的なボトルネック問題をコンシューマ向けの市販実機によって明らかにするとともに、ネットワークの輻輳制御におけるフレーム処理方法にネットワークコーディングを活用し、スループットの向上と輻輳の緩和を同時に可能とする輻輳制御方式を提案する。実験環境として提案方式をアプリケーション層に実装し、提案方式の優位性を立証する。さらにキャンパスネットワークにおける実際の通信環境を分析し、提案方式が有効に機能する環境であることを明らかにした。

Discussions on Congestion Control with Network Coding for Switches in Campus Network

SATORU ONO,^{†1} MASAKI BANDAI^{†2}
and TAKASHI WATANABE^{†1}

In general, LANs are enhanced by cascade connection of switching hubs. In a LAN with the cascade connection, the communication at the port of connected points, congestion is subject to occur at the point. For the problem TCP's window control has been already proposed for a congestion control in full duplex Ethernet. However, throughput degrades due to controlling the transmission rate in these methods. On the other hand, the network coding is attracted attention as a technology that improves the bandwidth efficiency. In this paper, we clarify a potential problem by the cascade connection of switching hubs,

and we propose a congestion control method that enables the throughput improvement and the congestion reduction at the same time. We implement the proposal method on application layer to prove the effectiveness of the proposal method. Furthermore, we analysis real communication data in a campus network and revealed that a proposal method functioned in this environment.

1. はじめに

大学という高等教育機関においては、その設置規模は様々なながらも情報伝達のためのコンピュータネットワークシステムが整備されている場合がほとんどである。一般にキャンパスネットワークと称されるこれらのネットワーク上には多くのノードが接続され、教育・研究だけでなく、様々な用途で利用されている。大学という組織は、その大部分が講座や研究室から構成されており、それぞれが独立した自治によって運営される傾向が強い。一般企業などとは異なりキャンパスネットワークの工事をともなわない物理的拡張などは自治レベルで日常的に行われるのが普通である。また、キャンパスネットワークとして敷設されるネットワークシステムの物理的トポロジとしては、中央の大規模伝送交換装置からスター型に接続されたレイヤ 2 (あるいは 3) スwitchングハブを介して、居室などの情報コンセントを収容する形態が一般的である。この物理トポロジイメージを図 1 に示す。

近年の国立大学法人等高等教育機関においては、情報漏洩の抑止に加え、情報資産の一元管理などの目的のため、シンククライアントシステムの導入がさかんである。シンククライアントシステムの実装方式の一種としてネットワークブート型方式があり、この方式のシステムは、ハードディスクなどの外部記憶装置を有さない DTE がネットワークを経由してオペレーティングシステムやアプリケーションの起動を行うため、ネットワークリソースを大量に消費する。さらに研究室などで、居室に敷設された情報コンセント数以上のノードを収容する必要が生じた際には、末端の情報コンセントにスイッチングハブを多段接続することによって拡張する場合が多い。このような状況でスイッチングハブを多段接続した場合、ハブ間の接続はそれぞれ単一ポートを通じて行われるため、他の複数ポートに接続されたノードからの通信が集中し、輻輳が発生しやすいという問題点がある。これらのスイッチングハブ

^{†1} 静岡大学創造科学技術大学院

Graduate School of Science and Technology, Shizuoka University

^{†2} 上智大学理工学部情報理工学科

Faculty of Science and Technology, Sophia University

で利用されている通信プロトコルには、データリンク層では Ethernet, その上位層においては TCP/IP が用いられることが多く, 前者における輻輳制御方式としては IEEE802.3x pause option¹⁾ が, 後者は TCP のウィンドウ制御²⁾ によって実装されている. さらにこれらの制御方法には様々な改良方式^{3)~5)} が検討されている. しかし, これらの輻輳制御方式は, 輻輳を緩和するために通信の抑制を行うため, ネットワークのスループットが低下することが避けられない.

本稿では, まず, スwitchングハブの多段接続時に発生することが予想される輻輳を, 実機を用いて再現し, その問題点を明らかにする. そのうえで, その問題点に対応するため, ネットワークのスループット向上と輻輳の緩和を同時に可能とする輻輳制御方式を提案する. これを実現するために提案方式では輻輳制御にネットワークコーディング^{6),7)} を活用する. ネットワークコーディングとは, 通信効率を向上させるために複数の送信フレームを排他的論理和 (exclusive OR) で多重化する技術である. フレームを多重化することによって, 送信回数が削減されるため, 輻輳が発生したノードのキュー内に滞留するフレームを迅速に処理することが可能となり, 輻輳の緩和とスループットの向上が同時に実現できる. これを評価するために, 提案方式を実装した実験環境を構築し, 単位時間あたりのデータ転送量であるネットワークスループットについて, 従来方式である IEEE802.3x pause option との比較を行った. その結果, 提案方式がネットワークスループットについて良好な結果を示すことを確かめた. さらに実稼働する通信環境を調査し, その通信環境において提案方式が有効に機能することを確認した.

以下, 2 章ではスイッチングハブの実機を用いて, これを多段接続し, ネットワーク機器

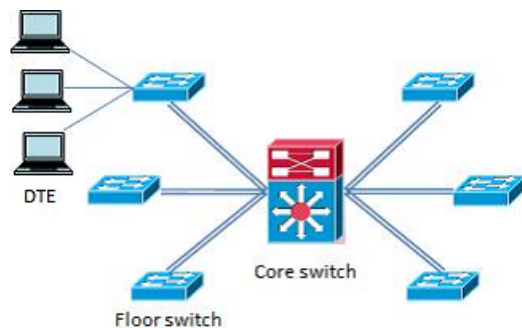


図 1 スター型トポロジ

Fig. 1 Star-type topology.

性能測定装置から疑似トラフィックを発生させることによって輻輳を再現し, 多段接続時に発生する問題点を明らかにする. 3 章ではネットワークコーディングの概要について説明する. 続く 4 章では, 提案方式であるネットワークコーディングを活用した輻輳制御方式について, 中継ノードとエンドノードにおける各機能を説明する. 次に本方式の評価を行うための実験環境に関する説明を 5 章で行うと同時に, 同環境を用いた性能評価によって本方式の優位性を示す. 6 章では, 実働する通信環境をサーベイし, 導出された通信特性より, 本方式が実環境において有効に機能することを確認する. 最後に 7 章で今後の検討課題と本稿のまとめを述べる.

2. 多段接続の問題点

本章では, スwitchングハブの多段接続時における問題点について, 実機を用いた実験結果に基づき説明する. 図 2 に実験環境における機器構成を示す. 図中 SMB で示すネットワーク性能評価試験装置には国内外で業界標準として利用されている Sprinet Communications 社製の SmartBits を利用した. シャシには 2 スロットを有する SMB600B を用い, 10/100/1000BASE-T Ethernet ラインカードを 2 枚実装した. 各ラインカードはそれぞれ 2 ポートを有しているため, 合計 4 ポートを用い, ツイストペアケーブルによる 4 リンクが利用可能である.

R1 が性能測定ターゲットデバイスであり, 本実験では, アライドテレシス株式会社製 SW08-TXR と FS716TX V2 の 2 機種を評価した. これらの 2 機種はいわゆるコンシューマ向け低価格帯デバイスとして一般に入手可能な製品群として分類されるスイッチングハブである. R1 と R2 は単一ポートで多段接続され, SMB が有する 4 ポートからそれぞれ R1 および R2 に向けて疑似トラフィックを発生させることで実験を行う. トラフィックの

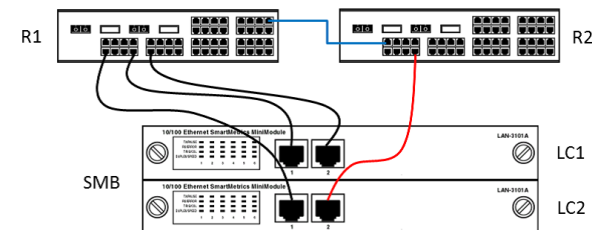


図 2 基礎実験の環境

Fig. 2 Configuration of fundamental experiment.

40 キャンパスネットワークにおけるスイッチングハブの輻輳制御方式に関する考察

フローは R1 から R2, R2 から R1 に向けて行われる双方向通信とした．実験における双方向通信の組合せと通信パラメータを表 1 に示す．R1 および R2 に接続された各リンクへの通信負荷を段階的に増加させることによって, R1 上の多段接続する通信ポートに対して輻輳を再現することが可能である．リンクへの通信負荷は, ワイヤレート比を基準として 10%~100%まで 20%ごとに増加させる．したがって, R1 から R2 への通信負荷が 50%に設定された時点で, R1 の多段接続部のポートには $50\% \times 3$ リンク, つまりワイヤレート比 150%の通信負荷が与えられることとなり, 輻輳が発生する．なお, 本実験では, 輻輳時における装置の輻輳状況を客観的に分析する必要があるため, 輻輳時に発生する R1 や R2 からの IEEE802.3x pause option によるフロー制御は行わないこととしている．

表 2 に各通信負荷における測定デバイスごとのフレームロス率を示す．フレームロス率は送信フレーム数を Tf , 受信フレーム数を Rf とした場合, 次式 (1) で計算される．

$$\frac{Tf - Rf}{Tf} \times 100 \quad (1)$$

ここで重要なことは, R2 → R1 で示す通信負荷はワイヤレート比 100%であり, 本来輻輳が発生しない状況であるにもかかわらず, 5%~7%程度のフレームロスが観測されている

表 1 通信の組合せとパラメータ

Table 1 Combinations of communication and parameters.

Flow	Source port	Destination port	Link speed	Frame size
R1 → R2	LC1-port1, LC1-port2, LC2-port1	LC2-port2	100Base-TX	64 Byte
R2 → R1	LC2-port2	LC1-port1		

表 2 フレームロス率

Table 2 Frame loss rate.

SW08TXR		R1 → R2									
		10%		30%		50%		70%		100%	
		TX	RX	TX	RX	TX	RX	TX	RX	TX	RX
R2 → R1	50%	0	0	0	0	33.3	0	52.3	0	66.6	0
	100%	0	0	0	0	33.3	5.5	52.3	6.6	66.6	6.9

FS716TX		R1 → R2									
		10%		30%		50%		70%		100%	
		TX	RX	TX	RX	TX	RX	TX	RX	TX	RX
R2 → R1	50%	0	0	0	0	33.3	0	52.3	0	66.6	0
	100%	0	0	0	0	33.2	4.6	52.3	7.4	66.6	7.5

点である．すなわち, Ethernet 全二重通信が利用なスイッチングハブの多段接続時を行った場合, 以下の 2 つの問題点がある．

- (1) 多段接続ポートが単一ポートである場合, 他ポートからの通信データが集中して輻輳が発生しやすい．
- (2) 全二重通信における片方向の通信が輻輳した場合, 他方向の通信に影響が及ぶことがある．

特記すべき点として, (2) の問題点より, 本来非同期に行われている全二重通信が輻輳時には, その非同期性が確保されていないことが考えられる．また, 装置内に実装されたメモリバッファを全二重通信で共有していることなどが予想される．

3. ネットワークコーディング

本章ではネットワークコーディングの概要について述べる．ネットワークコーディングは, ネットワークの情報伝達能力であるスループットをその理論的限界値である最大フローまで高めることを可能とする技術として提案された⁸⁾．以下, ネットワークコーディングについて, 無線ネットワークにおける事例を説明する．通常インフラストラクチャモードの無線ネットワークでは, アクセスポイントを介して, DTE 間あるいは別ネットワークへの通信が行われる．したがって, 図 3 に示すとおり, DTE である DTE-a と DTE-b が AP を通じてノード間通信を行う場合, 合計 4 回の送信が必要となる．ところが, AP が DTE-a, DTE-b 双方のノードから送信されたお互いが通信するためのフレームをタイミングよくほぼ同時に受信できた場合, これらのフレームを多重化し, 単一フレームとして一括送信することで送信回数を 3 回に削減することが可能である．多重化は eXclusiveOR によるビット演算で行われるため, 多重化フレームの復号は, 自身が保有する多重化前のフレームを保持

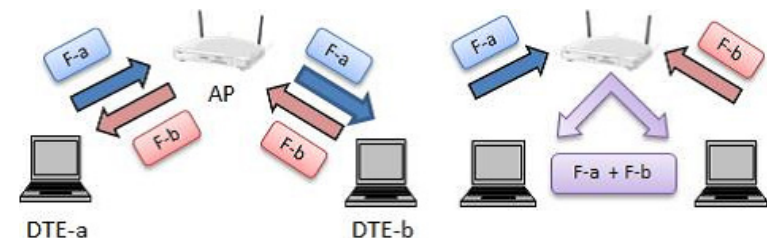


図 3 無線ネットワークにおけるネットワークコーディングの例

Fig. 3 Network coding in a wireless network.

していれば、そのフレームと多重化フレームを再度 XOR 演算することで所望の結果を得ることができる。これは、無線ネットワークで利用される電波の同報性をうまく利用した活用事例である。

このようにネットワークコーディングによるフレームの多重化処理を実現するためには、ネットワークを構成する機器に2つの機能が必要とされる。1つ目の機能は中継ノードであるAPにフレームを多重化する機能、もう1つは送信ノードであるDTEに自らが送信したフレームを一定期間保持する機能である。さらに効率良くフレームの多重化処理を行うためには、中継ノードに双方向からほぼ均等かつ適切な量のフレーム数が送られてくることが望ましい。なぜなら、通信方向が1方向のみに偏っていた場合や、統計的に均等であった場合でもそれぞれのフレームがタイミング良く中継ノードに到着しない場合には、多重化処理のためにどちらか一方のフレームの到着を待つことによる遅延が生じ、多重化一括配信によるスループット向上の恩恵が期待できないからである。このようにネットワークコーディングを効率的に活用するためには、通信環境に多くの制約が求められるため、効果的な実装には工夫が必要とされる。

4. 提案方式

本章では提案方式である FCNC (Flow Control with Network Coding) と BC (Buffering Coordination) の 2 つの機能について説明する．図 4 に FCNC/BC の機能概要を示す．図中，DTE である N-a1 と N-b1 が中継ノード R1 および R2 を介して通信が行われ

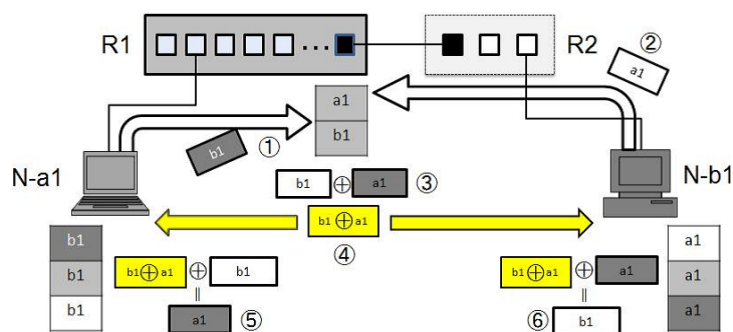


図 4 FCNC/BC の概要
Fig.4 Outline of FCNC/BC.

る場合を想定している。R1 と R2 を多段接続されたスイッチングハブと仮定する。ノード N-a1 は N-b1 にフレーム b1 を送信し (①), N-b1 は N-a1 に向けてフレーム a1 を送信する (②)。R1 では双方向から到着したフレームを XOR 演算によって単一のフレームに多重化する (③)。その後、R1 は多重化したフレームをブロードキャストによって一括配信する (④)。N-a1 および N-b1 では受信した多重化フレームを自身が保管する送信済みフレームを用いて復号する (⑤, ⑥)。以降、本図を用いて各提案方式の詳細な説明を行う。

4.1 FCNC (Flow Control with Network Coding)

FCNC は図 4 中 R1 に実装される機能である。このような中継ノードには輻輳の緩和やフレームロスを抑止するために、一般に一定量のメモリバッファが設けられている。機器の輻輳状態とはこのメモリバッファの収容限界を超えることを意味する。IEEE802.3x pause option によるフロー制御では、輻輳発生時、DTE に pause frame を送信し、DTE 側の送信を中断することで輻輳からの回復をはかるようにしているが、FCNC ではメモリバッファのしきい値として、図 5 に示す 3 つの値を設定し、フロー制御を行う。メモリバッファの占有量がしきい値 T_{cd} を超えたら、双方向からのフレームを XOR 演算で多重化し、ブロードキャストによって一括配信を行う。メモリバッファの占有率がなおも上昇し続け、ついにしきい値 T_{max} を超えた際には、pause frame を送信し、DTE からの送信抑制を行う。そして、メモリバッファのしきい値が T_{min} を下回った段階で、DTE からの送信を再開する。以上が、中継ノード R1 で実行される FCNC の一連の処理内容である。

R1 に実装されるメモリバッファには双方向からの通信を識別するために、それぞれ専用の論理キューが定義されており、各論理キューからフレームを FIFO に処理することで遅延のない適切な多重化処理を可能とする。なお、ここでいう双方向通信とは、R1 が R2 に対して多段接続する通信ポート上に学習された MAC アドレスを宛先とする通信と（図 4 中①）、R1 の他のポート上に学習された MAC アドレスを宛先とする通信（図 4 中②）のことである。多段接続ポートを除く R1 の他のポートに登録された MAC アドレスを有する

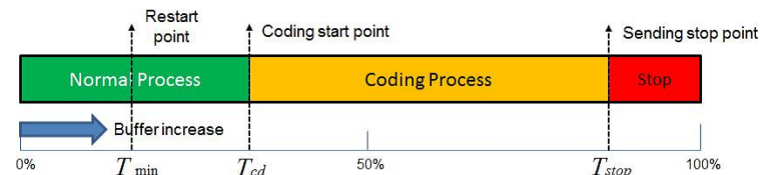


図5 しきい値 T_{max} , T_{cd} , T_{min}
Fig.5 Thresholds T_{min} , T_{cd} , T_{min} .

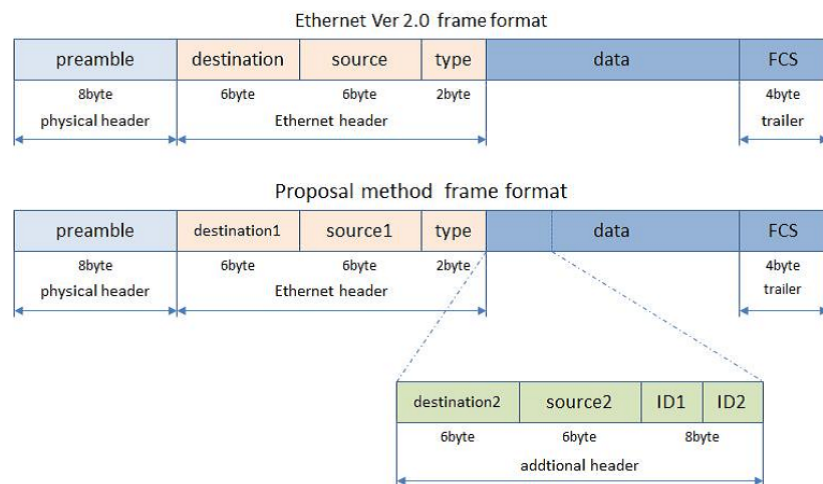


図 6 フレームフォーマットの比較
Fig. 6 Comparison of frame formats.

ノード間の通信に関しては、FCNC の処理対象にならない。

4.2 4.2 BC (Buffering Coordination)

中継ノード R1 からブロードキャスト一括配信された多重化フレームは、N-a1 および N-b1 で受信後、復号処理が行われる。そのため、受信した多重化フレームを通常のフレームと区別する必要がある。Ethernet フレームフォーマットを改変した。図 6 に TCP/IP で利用されている Ethernet Ver2.0 フォーマットと、提案方式で用いる拡張フォーマットを示す。多重化されたフレームは送信宛先を 2 つ有するため、ヘッダ情報として 2 つの送信宛先および送信元アドレスを定義した。DTE が多重化フレームを判別するために、type フィールドを利用する。復号処理は DTE が自ら送信したフレームと多重化フレームを XOR 演算することで行われるため、これらのフレームを一定期間保存する必要がある。保存されるフレームは一定期間経過後に削除される。

5. 実装実験環境

本章では提案方式の実装実験環境の説明を行う。図 7 に実装実験の機器構成を示す。提案方式は OSI 階層モデルにおける第 2 層に実装すべき機能であるが、実験環境としては、工

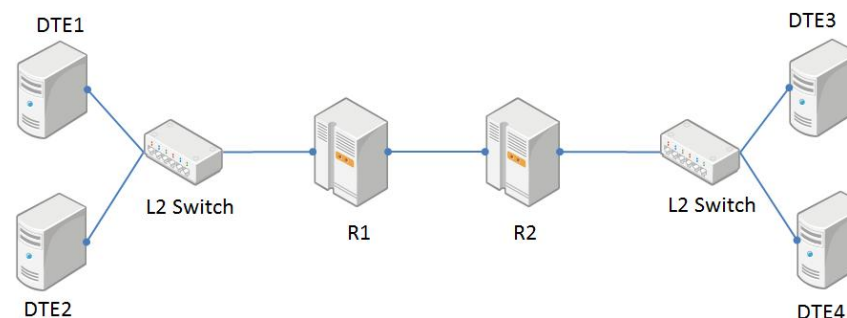


図 7 評価実験の環境
Fig. 7 Configuration of evaluation experiments.

数をかけずに比較的短期に容易に構築することができるアプリケーション層に本機能を実装し評価を行うこととした。提案方式の基本機能はネットワークコーディングを用いたデータ多重化による一括配信であり、基本的な挙動を確認するにあたっては、アプリケーションを用いたエミュレーションで十分可能であると判断したためである。アプリケーションの開発言語には Microsoft Visual Basic Version6.0 SP6 を用いた。アプリケーションで用いる通信プロトコルには UDP を利用した。これは提案方式の効果を基礎的に評価するにあたり、利用する通信プロトコルが有するフロー制御の影響を除外し、実験データの正確性を担保するためである。データ到着性の保証はアプリケーションの機能により実装する。図中、R1 および R2 は中継ノードであり、多段接続されたスイッチングハブに相当する。R1 には FCNC が搭載されるため、フレーム多重化機能に加え、輻輳緩和のために実装されるメモリバッファとしてデータ格納変数配列を静的に 400 個分定義した。この数値の根拠として、2 章の実験で用いたスイッチングハブ (FS716TX V2) には 512 Kbyte のメモリバッファが実装されており、Ethernet の MTU 値である 1,518 byte 長のフレームを格納した場合、約 345 個のフレームを保持することが可能であるためである。DTE は双方向に 2 台ずつ設置し (DTE1, DTE2, DTE3, DTE4)、中継装置である R1 および R2 とはスイッチングハブ (L2 Switch) を介して接続される。通信は各 DTE 間で双方向にランダムで行われる。DTE に搭載するアプリケーションは、データ送受信機能と BC を有する。DTE は受信した多重化フレームを復号するために必要となるキーフレームを、自らが送信したフレームの中から選択する必要がある。本実験においては多重化フレームのペイロード部にこれを選択するための 32 bit の識別子を定義した (図 6 中 ID1, ID2)。データの多重化と復号処

理はアプリケーションで行うため、送信するデータに図 6 で示したヘッダ情報を付加した。中継ノードである R1 に搭載するアプリケーションは、データ中継機能と FCNC を有する。FCNC によるデータの多重化は DTE から送信されたファイル情報部分のみに対して実施される。R2 には単にデータ中継機能のみを有するアプリケーションをインストールした。DTE から送信されるデータ量を逐次増加させ、あらかじめ設定された R1 の中継処理能力を超えることで疑似的に輻輳状態を再現することが可能である。R1 は DTE から送信されたデータを一定量保持するためのバッファを有している。このバッファの全体容量に対して図 5 で示した 3 つのしきい値を定義し、提案方式による輻輳制御をエミュレートした。性能評価のため、従来方式である 802.3x pause option によるフロー制御との比較を行った。

5.1 性能評価

R1 に設定したバッファの各しきい値について、 T_{min} を 20%、 T_{cd} を 40%、 T_{max} を 90% とした。提案方式では、R1 のバッファ占有率が 40% を超えた段階で双方向通信データの多重化一括配信が行われる。一方、比較対象である従来方式の 802.3x では、 T_{min} を 20%、 T_{max} を 90% とし、2 つのしきい値を用いて評価を行った。各 DTE において送信処理に利用されるファイルサイズは 100 Kbyte とした。このファイルの受信を完了する総時間を測定し、ファイルサイズを総送信時間で除算することによって単位時間あたりにおける通信量を算出した。図 8 に提案方式と従来方式の性能比較を示す。グラフの横軸はデータ送信間隔であり、値が小さくなるに従い、単位時間あたりに送出されるデータ量が大きくなることを表している。

結果より、提案方式が従来方式と比較して単位時間あたりの通信量が大きくなることに加え、輻輳が発生する送信レートが高くなっていることが分かる。単位時間あたりの通信量は従来方式と比較するとほぼ 2 倍の値が観察できる。これは R1 のバッファ内に格納された通信データを多重化して送信することで、データの送信処理が効率的に行われたためである。次に、提案方式と従来方式の輻輳時通信特性を確認するため、通信データをキャプチャし、スループットの変動を確認した。これを図 9 に示す。縦軸に単位時間あたりの送信量、横軸は経過時間である。提案方式、従来方式いずれも鋸刃状のグラフを示すが、提案方式がより大きな送信データ量を長時間継続できており（図 9 の区間①）、輻輳状態からの復帰（図 9 の区間②）が迅速に行われていることが確認できる。

5.2 通信量の比率

これまでの実装実験では、双方向からの通信量を同率として行った。提案方式では双方向からの通信を多重化して一括送信するため、それぞれの通信量がほぼ等しくなる環境が求め

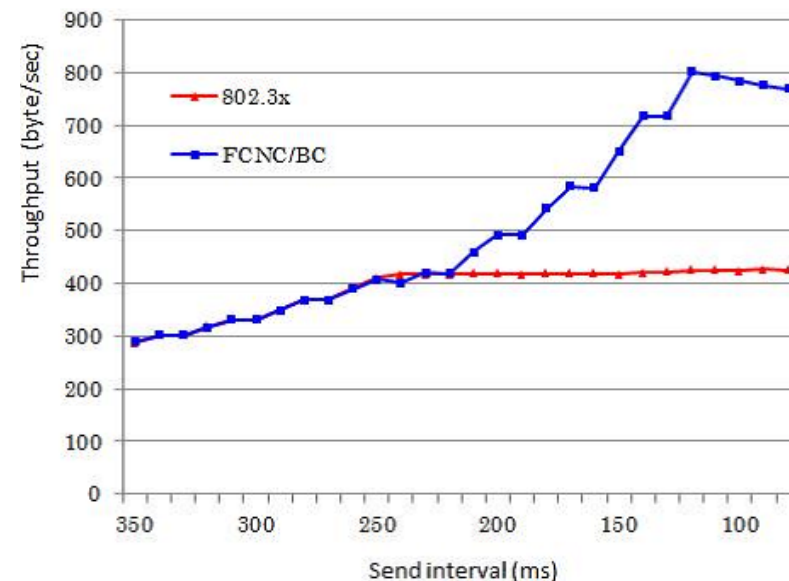


図 8 送信間隔に対するスループットの評価

Fig. 8 Evaluation of throughput vs. sending interval.

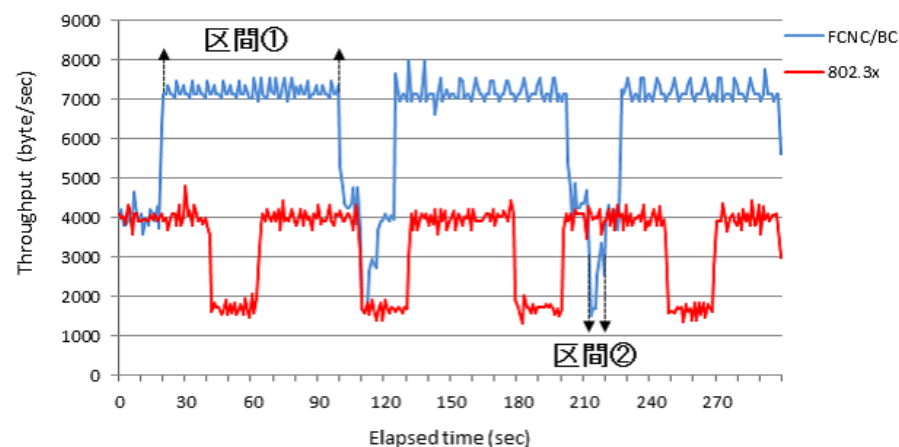


図 9 スループットの変動

Fig. 9 Fluctuation of throughput.

られるためであるが、現実に移働する通信環境において、必ずしも双方向からの通信量が同率となることは想定しにくい。ため、提案方式が有効に機能することができる通信比率を、実験環境を用いて調査した。なお、ここでいう通信量とは、単に DTE から送信されるフレーム数のことを示しており、送信される絶対情報量ではないことに留意されたい。

DTE からの通信比率の調整は、それぞれの DTE が送信するデータ送信間隔をアプリケーションによって制御することで行った。たとえば、通信比率が 7:3 の場合には、R1 と R2 を挟む一方の DTE からのデータ送信間隔は 126 ms ごとであり、もう一方では 54 ms ごとに送信するという具合である。各通信比率における提案方式と従来方式の比較を図 10 に示す。比率 5:5 と 6:4 では単位時間あたりの通信量の差はほとんど認められない。これは図 5 に示すしきい値 T_{cd} からしきい値 T_{max} の間で、各 DTE から送信されたデータのほとんどが多重化可能であったことを示している。ところが、比率が 7:3 になると単位時間あたりの通信量は大幅に低下し、以降通信比率の隔たりが大きくなるにつれて、提案方式の優位性が薄れていく。このことから、提案方式が有効に機能することができる通信比率の最適値は 5:5 から 6:4 の範囲にあるといえる。

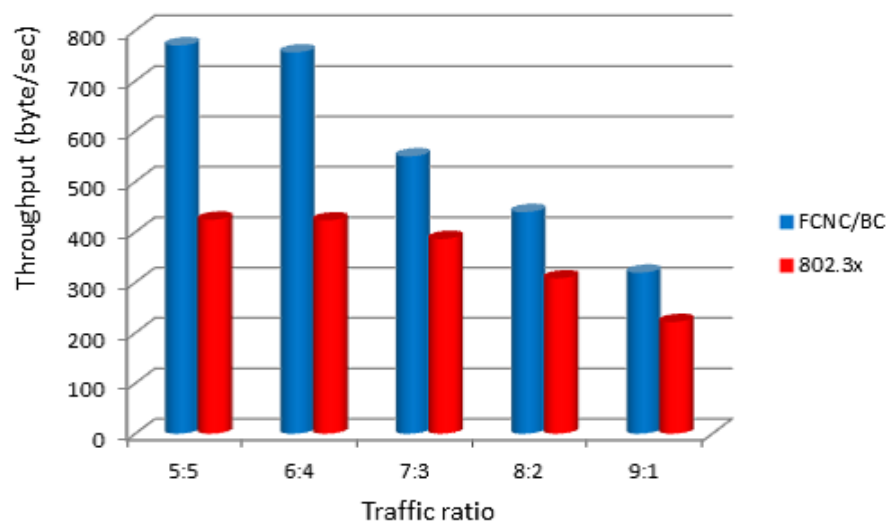


図 10 通信比率に対するスループットの評価
Fig. 10 Evaluation of throughput vs. traffic ratio.

6. 実環境の調査

本章では現実に移働する通信環境における多段接続されたスイッチングハブ上の多段接続ポート部の通信状況を調査し、実環境での双方向通信の比率を明らかにする。調査環境の概要を図 11 示す。調査対象の機器は多段接続接続された 48 ポートを持つレイヤ 2 スwitchングハブ (図中 R1) である。各ポートのリンク速度は 100Base-TX であるが、輻輳対策のため多段接続ポートのみ 1000Base-SX インタフェースを装備している。また、多段接続ポートの通信内容を収集するために GBIC モジュールを用いて 1000Base-T ポートを 1 つ増設し、同ポートのミラーリングポートとして設定した。測定対象のスイッチングハブには 75 台の業務用パソコンが接続されている。一方、多段接続先のスイッチ (図中 R2) には業務用サーバが 40 台接続され、上流であるインターネットへの接続起点が收容される。したがって、多段接続ポートを通じて想定される通信内容としては、Web 閲覧や電子メールの送受信、およびクライアントサーバ型業務用アプリケーションの利用などが主であり、きわめて一般的な通信環境である。図中 R1 から R2 への通信を上り通信 (upload)、R2 から R1 への通信を下り通信 (download) と定義した。多段接続ポート上を通過するすべての通信をミラーリングポートからサンプリングし、分析を行った。収集と分析にはパケットアナライザとして著名である Wireshark⁹⁾ を用いた。データのサンプリング期間は月曜日から金曜日の業務が行われている平日 5 日間とし、それぞれ午前と午後 1 回ずつ実施した。サンプリング時間は 180 秒である。1 回あたりのサンプリングで約 150 MB の通信データが収集される。これらの通信データに対して、MAC フレームのヘッダ情報に記録された宛先 MAC アドレスと送信元 MAC アドレスを解析し、登り通信と下り通信の区別をした。R1

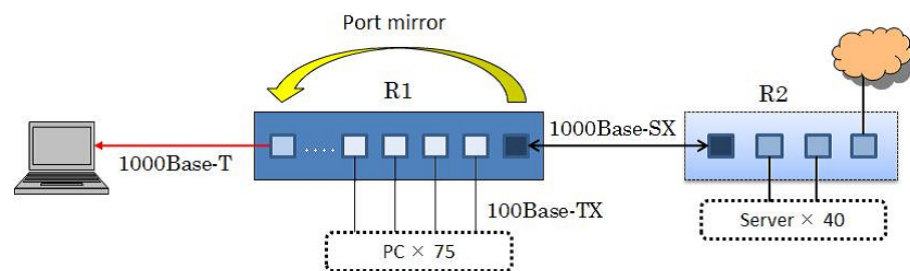


図 11 調査実験の環境
Fig. 11 Configuration of investigation experiments.

45 キャンパスネットワークにおけるスイッチングハブの輻輳制御方式に関する考察

表 3 通信比率の調査結果
Table 3 Investigated traffic ratio.

Mon.				Tues.				Wed.				Thus.				Fri.			
AM		PM		AM		PM		AM		PM		AM		PM		AM		PM	
up	dwn	up	dwn	up	dwn	up	dwn	up	dwn	up	dwn	up	dwn	up	dwn	up	dwn	up	dwn
38	61	62	36.1	50	49.1	68.8	29.7	43	56.2	69.4	28.6	39.2	60.5	66.2	32.2	65.1	32.9	76.7	22.1

95% confidence intervals

	average	median	upper confidence	lower confidence
upload traffic	57.84	63.55	67.36	48.32
download traffic	40.84	34.5	50.65	31.03

から R2 への通信を upload とし、R2 から R1 への通信を download として定義している。

なお、通信量の算出は 5.2 節で述べたとおり、フレーム数を単位として計算している。表 3 に通信比率の調査結果を示す。upload の平均値は 57.9%，download の平均値は 40.9%と上下通信の比率に大きな隔たりは観測されないが、計測した数値群の母集団に対して標準偏差を求めたところ、upload で 13.31，download で 13.71 とかなり大きいものであったため、ネットワークを流通するデータのバースト的特性に鑑み、それぞれの中央値を算出した。その結果、upload は 63.55%，download は 34.5%であることから、フレーム数の比較という観点からは、一般によく知られているような下り通信に対して通信量が偏っているような状況でないといえる。また、フレーム数の平均値という観点からは、双方向通信量はおおむね等しい状況であることが確認できた。さらにこれらの数値が確率的に示す範囲を明らかにするために、信頼区間を求め、その平均値が 95%の確率で存在する範囲を算出した。表 3 に母数となるサンプリングデータの一覧と 95%信頼区間算出結果を示す。5.2 節での調査結果より、提案方式が有効に機能することができる通信比率は 5 : 5 から 6 : 4 程度であり、算出した 95%信頼区間の範囲より、実働する通信環境における定常的な通信比率はおおむね提案方式が有効に稼働することができる範囲内であるといえることができる。

7. 検討課題とまとめ

7.1 検討課題

今回の実験では、提案方式をアプリケーション層に実装することにより、基礎評価を行った。この評価によって、提案方式を用いた輻輳制御の基本的挙動が確認できたものと考えられるが、今回の実験環境はソフトウェアおよびハードウェアの実装に依存する部分が多いため、

提案方式の正確な評価を行うために本来実装すべきデータリンク層における提案方式の挙動を確認する必要がある。このためにはよりハードウェアに近い環境を構築し、実験することが求められる。近年のスイッチングハブには、その制御部に汎用プロセッサではなく、カスタマイズされた専用 ASIC (Application Specific Integrated Circuit) が用いられていることが多い。このような専用集積回路をユーザが自由に設計することが可能な FPGA (Field Programmable Gate Array)¹⁰⁾ を用いることにより、任意の論理機能を集積回路上に設定・実装し、評価することが可能である。今後は、FPGA を用いて提案方式をデータリンク層に実装・評価することを予定している。また、本稿では提案方式の実装を性能が低い低価格なスイッチングハブを対象として検討・評価を行ったが、よりいっそうの高帯域および高性能が求められるミドルレンジ以上のスイッチングハブに対する提案方式の実装を検討すべきと考える。さらに、提案方式の実装に係るコストの問題や、DTE 側の改造手法など(たとえば、デバイスドライバのアップデートなど)、提案方式の実現方法の詳細に関しては、今後十分な検討が必要である。

7.2 ま と め

本稿ではネットワークコーディングを活用し、多段接続されたスイッチングハブにおける輻輳制御方式を提案し、実験環境を構築してその評価を行った。実験結果より、提案方式は従来方式よりも高いスループットを実現できると同時に、輻輳からの迅速な復帰が可能であることが確認できた。さらに提案方式が有効に機能することができる通信比率を考察したうえで、現実に稼働する通信環境のデータをサンプリングし、多重化可能な双方向通信量の比率を収集した通信データによって分析した。分析結果より実働環境における双方向の通信比率は、おおむね提案方式が有効に機能することができる範囲であることが明らかになった。

参 考 文 献

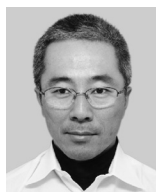
- 1) Feuser, O. and Wenzel, A.: On the effects of the IEEE802.3x flow control in full-duplex Ethernet LANs, *Proc. 24th Conference on Local Computer Networks (LCN'99)* (1999).
- 2) Mo, J., La, R.J., Anantharam, V. and Walrand, J.: Analysis and comparison of TCP reno and vegas, *Proc. INFOCOM'99* (1999).
- 3) Floyd, S. and Jacobson, V.: Random early detection gateways for congestion avoidance, *IEEE/ACM Trans. Networking*, Vol.1, pp.397–413 (1993).
- 4) 谷川浩一，関根佳明，稲井 寛：スイッチングハブにおけるポーズフレーム送信方法に関する検討，*信学論*，Vol.J86-B, No.3, pp.590–594 (2003).
- 5) 谷川浩一，稲井 寛：Ethernet スイッチングハブにおける Selective-Back-Pressure

非対応ステーションの救済方法, 信学技法, Vol.102, No.384 (IN2002-102), pp.71-76 (2002).

- 6) Fragouli, C., Le Boudec, J.-Y. and Widmer, J.: Network Coding: An Instant Primer, *ACM SIGCOMM Computer Communication Review*, Vol.36, No.1, pp.63-68 (2006).
- 7) Li, S.-Y.R., Yeung, R.W. and Cai, N.: Linear Network Coding, *IEEE Trans. Information Theory*, Vol.49, No.1, pp.2371-381 (2003).
- 8) Ahlswede, R., Cai, N., Li, S.-Y.R. and Yeung, R.W.: Network information flow, *IEEE Trans. Information Theory*, Vol.46, No.4, pp.1204-1216 (2003).
- 9) available from (<http://www.wireshark.org/>).
- 10) Alachiotis, N., Berger, S.A. and Stamatakis, A.: EFFICIENT PC-FPGA COMMUNICATION OVER GIGABIT ETHERNET, *10th IEEE International Conference on Computer and Information Technology (CIT 2010)*, pp.1727-1734 (2010).

(平成 23 年 3 月 19 日受付)

(平成 23 年 9 月 6 日採録)



小野 悟

昭和 39 年生。平成 22 年静岡大学大学院情報学研究科修士課程修了。現在、静岡大学創造科学技術大学院自然科学教育部情報科学専攻博士後期課程に在学中。修士(情報学)。国立大学法人浜松医科大学に勤務。情報システムおよびネットワークシステムの管理運営に従事。



萬代 雅希(正会員)

平成 8 年慶應義塾大学理工学部電気工学科卒業。平成 10 年慶應義塾大学大学院修士課程修了。平成 16 年同大学院博士課程修了。同年静岡大学情報学部情報科学科助手。同助教, 講師を経て, 平成 22 年上智大学理工学部情報理工学科准教授。平成 10~12 年ソニー(株)勤務。平成 13~15 年日本学術振興会特別研究員(DC2)。平成 18~19 年ブリティッシュ・コロンビア大学訪問研究員。博士(工学)。主として, 通信ネットワークに関する研究に従事。平成 19 年情報処理学会山下記念研究賞, 平成 19, 21 年電子情報通信学会通信ソサイエティ活動功労賞, 平成 20 年船井情報科学奨励賞受賞。



渡辺 尚(正会員)

昭和 57 年大阪大学工学部通信工学科卒業。昭和 59 年大阪大学大学院工学研究科通信工学専攻博士前期課程修了。昭和 62 年大阪大学大学院工学研究科通信工学専攻博士後期課程修了。博士(工学)。同年徳島大学工学部助手。平成 2 年静岡大学工学部情報知識工学科助教授。平成 8 年静岡大学情報学部情報科学科教授。現在, 静岡大学創造科学技術大学院教授。静岡大学情報学部情報科学科教授兼任。平成 7 年文部省在外研究員(カリフォルニア大学アーバイン校)。計算機ネットワーク, 分散システムに関する研究に従事。平成 17 年より情報処理学会モバイルコンピューティングとユビキタス通信研究会主査。訳書『計算機設計技法』, 『802.11 無線ネットワーク管理』等。IEEE 会員。

正 誤 表

ページ	訂正箇所	誤	正
38	第二著者名	萬 代 雅 紀	萬 代 雅 希