

光リングネットワークにおける 帯域利用効率向上のためのパケットスケジューリング手法

岡村 悠貴[†] 鈴木 誠[†] 今泉 英明[†] 関谷 勇司^{††} 森川 博之[†]

[†] 東京大学 〒153-8904 東京都目黒区駒場 4-6-1

^{††} 東京大学 〒113-8658 東京都文京区弥生 2-11-16

E-mail: †{okamura,makoto,imaq,mori}@mlab.t.u-tokyo.ac.jp, ††sekiya@wide.ad.jp

あらまし 非同期可変長光パケットリングネットワーク技術を用いるとネットワークを低コストで構築できるが、高トラフィック時にはパケットを送出できない短い無信号区間が頻発し、帯域利用効率が低下するという問題がある。本稿では、この問題に対して、各ノードにおいて同じ宛先のパケットが連続するようにパケットを挿入するパケットスケジューリング手法が帯域利用効率向上に有効であることを示す。また、リング上のノードが協調することで効率的にパケットを連続させる協調型パケットスケジューリング方式を示し、シミュレーションにより本手法の性能評価を行う。
キーワード リングネットワーク, 光パケット交換, 可変長パケット, パケットスケジューリング

A Packet Scheduling Technique on Optical Ring Networks for Increasing Utilization

Yuki OKAMURA[†], Makoto SUZUKI[†], Hideaki IMAIZUMI[†], Yuji SEKIYA^{††}, and Hiroyuki MORIKAWA[†]

[†] The University of Tokyo Komaba 4-6-1, Meguro-ku, Tokyo, 153-8904 Japan

^{††} The University of Tokyo Yayoi 2-11-16, Bunkyo-ku, Tokyo, 113-8658 Japan

E-mail: †{okamura,makoto,imaq,mori}@mlab.t.u-tokyo.ac.jp, ††sekiya@wide.ad.jp

Abstract Variable-length asynchronous optical ring network technology is more appropriate for the Metropolitan Area Network (MAN) due to its simplicity, flexibility and cost-efficiency. However, when the network is heavily loaded, ring nodes can not find continuous void large enough to insert long packets. This property causes low bandwidth utilization. In order to increase the bandwidth utilization, this paper proposes a novel packet scheduling algorithm which intends to send packets towards same destination as continuously as possible. Moreover, to increase its performance this paper also proposes a cooperative packet scheduling mechanism. We evaluate these mechanisms through simulation.

Key words asynchronous Optical Ring, Optical Packet Switching, Variable-Length Packets, Packet Scheduling

1. はじめに

FTTH (Fiber To The Home) などの大容量なアクセスネットワークの普及に伴い、それらを収容するメトロポリタンエリアネットワーク (MAN: Metropolitan Area Network) の大容量化が求められている。MAN には光回線交換技術である ROADM が適用されているが、ROADM は各リンクへ静的に波長を割り当てるため帯域利用効率が悪いという問題がある [1]。

これに対して、光電変換を省略して光領域でパケットを交換する光パケットリングネットワークでは、スイッチング粒度を

小さくし、全波長帯域を全てのノードで共有することにより、統計多重効果が生まれるため帯域利用効率の大幅な改善が期待される [3]。光パケットリングネットワークは、固定長パケットを処理する同期型光パケットリングネットワーク方式と可変長パケットを処理する非同期型光パケットリングネットワーク方式の2方式に分類される。

これら2つは以下のような特徴を持つ。同期型光パケットリングネットワーク方式では、時間スロットを用意し、各スロット内に収まる固定長パケットを送受信することで通信する [2]。スロット割り当てによる帯域制御ができるため、非同期型の帯

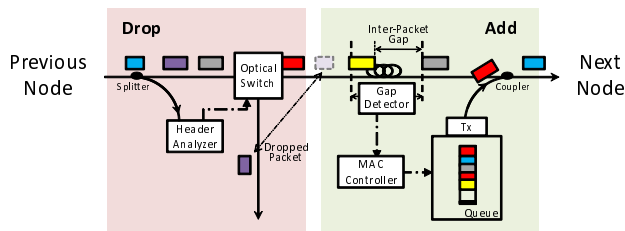


図 1 非同期光パケットリングノードの構成図

域割り当てに比べて帯域制御が容易になる。しかしながら、IP パケットなどの可変長パケットをスロットに收容するためには、スロット長を最大パケット長 (MTU:Maximum Transmission Unit) と等しくするか、パケットの分割/再構成が必要となり、帯域利用効率の低下に繋がる。また、ネットワーク全体でスロット同期を取るための同期機構が必要となるため複雑になりやすい。

非同期型光パケットリングネットワーク方式では、各ノードがリング上のトラフィックを監視し、自ノード宛のパケットを抜き出し、パケットの挿入可能な無信号区間にパケットを挿入することで通信する。非同期型光パケットリングネットワーク方式は可変長パケットを分割/再構成せずに收容できるため IP データ通信との親和性が高く、各ノードの同期が不要となるため低コストにネットワークを構築できる [3]。しかしながら、短い無信号区間に対してはパケットを挿入できず、短い無信号区間が増加すると帯域利用効率が低下するという問題がある [4]。

この問題に対しては、大きく 2 つのアプローチがある。K.S.Kim ら [7]、T.Eiod ら [8] や W.S.Hwang ら [9] は、無信号区間よりも短いパケットを優先的に挿入して、無信号区間をできるかぎり利用することで帯域利用効率を向上させている。本方式では、従来手法より無信号区間を利用できるが、長いパケットはまとまった無信号区間が生じない限り挿入できない。また、短いパケットを優先させることでパケットの順序逆転が発生するため、上位プロトコルに悪影響を及ぼす。

一方、V.H.Nguyen ら [5] や T.Kitamura ら [6] は、各ノードがパケットの挿入を抑制することで、長い無信号区間を生み出している。本方式では高トラフィック時に長いパケットを挿入できるようになるが、まとまった無信号区間を生み出すためにパケットの挿入を抑制させるため、帯域を最大限まで利用することはできない。

本稿では、各ノードにおいて同じ宛先のパケットが連続するようにパケットを挿入することで、パケットの順序逆転や挿入抑制を発生させずに帯域利用効率を向上させる方式を示す。続いて、シミュレーション評価により本手法の有効性を示す。従来の手法では最大帯域利用効率が 75 % であるのに対して、本方式では帯域利用効率が 95 % まで向上する。

本稿の構成は以下の通りである。まず、2. において非同期式光パケットリングネットワークの特徴を明らかにし、関連研究とその問題点を示す。続いて、3. で同じ宛先を連続して挿入するパケットスケジューリング方式を示し、4. においてシミュレーションにより性能評価を行う。最後に、5. でまとめとする。

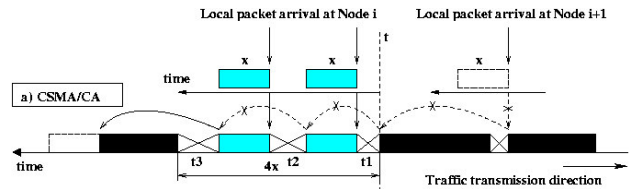


図 2 Bandwidth Fragmentation

2. 非同期式光パケットリングネットワーク

図 1 に非同期光パケットリングネットワークにおける交換ノードの構成図 [10] と、交換ノードを通過するパケットの様子を示す。交換ノードは受信部 (Drop) と送信部 (Add) から構成される。

受信部は、リング上を流れるパケットから自ノード宛のパケットを判別し、パケットのヘッダに含まれる宛先情報を用いて光スイッチを制御し、前ノードから到着したパケットを次ノードに通過させるか抜き出すかを決定する。光パケットが抜き出されるとその区間は無信号区間となり、現在のノード以降で帯域を利用可能になる。

送信部は、リングネットワーク上を流れるパケットと衝突しないようにローカルパケットを送出する。Gap Detector は無信号区間のサイズと通過タイミングを検出して送信用キューを制御する。送信用キューは無信号区間にキュー内の先頭パケットが収まる場合、無信号区間の通過タイミングに合わせてキュー内の先頭パケットを挿入する。

しかしながら、キュー内の先頭パケットより長い無信号区間が検出されない場合、パケットは挿入できない。高トラフィック時には、図 2 のようにリング上のパケット密度が高まって無信号区間が短くなる。IP パケットは 40Byte 付近の短いパケットと 1500Byte 付近の長いパケットに二極化しており、長いパケットが先頭にいると、短いパケットを挿入できる無信号区間も利用できなくなり、帯域利用効率の低下に繋がる [4]。

3. 宛先情報を利用したパケットスケジューリング方式

帯域利用効率向上のためには、まとまった無信号区間を生み出すことが重要となる。各ノードが自ノード宛のパケットを受信する際には、帯域の再利用を可能できるようにリング上からパケットを取り除くため、パケット受信後には無信号区間が生じる。このとき、同じ宛先のパケットが連続していると、パケットが取り除かれた後には、まとまった無信号区間が生まれることに着目する。

そこで、本稿では、リング上に同じ宛先のパケットが連続するようにパケットを挿入するためのパケットスケジューリング方式を示す。各ノードは宛先ごとの送信用キューを持ち、パケット挿入可能な無信号区間を検出した時に、1 つの無信号区間に対して同じ宛先のパケットを連続して挿入する。これによりリング上には同じ宛先のパケットが連続して並ぶ。

本手法では、無信号区間の開始時にキューを選択する必要が

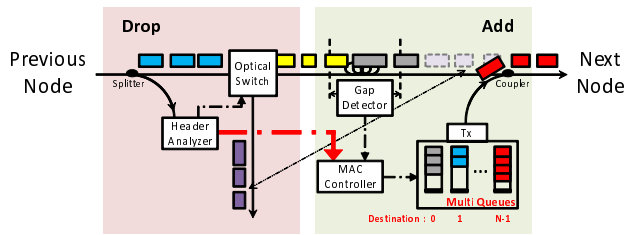


図3 宛先情報を利用した非同期光パケットリングノードの構成図

ある。キュー選択方式として、各ノードが独立して挙動する独立同一宛先パケット挿入方式（Local方式）と全ノードが協調して同じ宛先のパケットを連続させる協調同一宛先パケット挿入方式（Cooperative方式）の2方式を示す。

Local方式では、最も古いパケットを持つキューを選択する。この方式では遅延の増大が抑えられない。Cooperative方式では、リング上を通過するパケットと同じ宛先のパケットを優先して挿入する。これにより、各ノードが協調して同じ宛先のパケットを連続させるため、Local方式よりまとまった無信号区間を生み出せる。

図3に、提案する交換ノードの構成図と、交換ノードを通過するパケットの様子を示す。受信部は、従来の非同期光リングノードと同様に、前ノードから到着したパケットの宛先情報を取り出す。このとき、Cooperative方式ではキュー選択時にノードを通過する宛先情報を利用するために、宛先情報を送信部のMAC Controllerに伝達する。送信部は、宛先ごとにキューを用意する。

Local方式では、各キューの先頭パケットの到着時間から、最も早く到着したパケットを持つキューを選択する。そして、まとまった無信号区間に対して、選択したキュー内にパケットがある限り同じキューからパケットを挿入する。キュー内の無信号区間がなくなった場合は、最も早く到着したパケットを持つキューを再び選択して、同様に無信号区間が途切れるまで同じキューからパケットを連続して挿入する。

Cooperative方式では、キュー選択方法がLocal方式と異なり、パケットの到着時間ではなく、リング上を通過するパケットと同じ宛先のパケットを持つキューを選択する。選択したキュー内にパケットがなくなった場合、Local方式と同様に最も早く到着したパケットを持つキューを選択する。

本方式では、[5]のようなトラフィック抑制をせずに、まとまった無信号区間を生み出せる。また、本方式では送信用キューに到着したパケットを複数のキューに分けるためパケットの順序逆転が発生するが、各宛先ごとに分かれたキューにおいては先頭から順にパケットを挿入するため、[7][9]と異なり、宛先ノードにおいてはパケットの順序逆転が発生しない。

4. シミュレーション評価

同じ宛先のパケットを連続して挿入する方式の有効性を OM-Net++ [11] を用いたシミュレーションにより評価した。パケット長分布は [9][10] を元に、MCI社のバックボーンリンクのトレースデータを利用した。パケット長は図3のように40Byte、

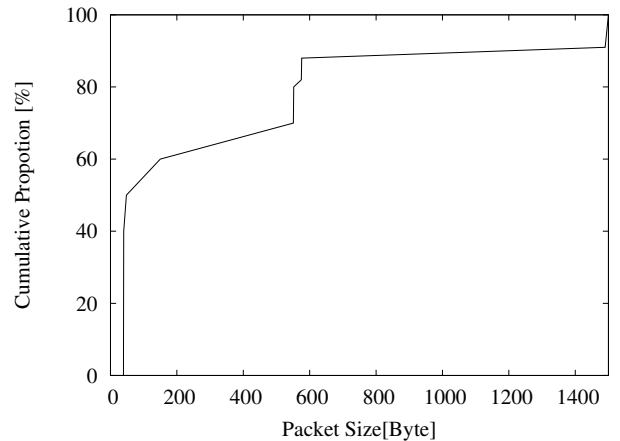


図3 パケット長分布

552Byte、1500Byteに集中する不均一な分布であり、平均パケット長は332Byteである。到着間隔はポアソン到着、宛先は各ノードに対して一様分布とした。リングネットワークポロジは、ノード数が10、波長数が10波長、ビットレートが1波長あたり40Gb/sの合計400Gb/sである。各ノードへの送出トラフィック量を4Gb/sから80Gb/sまで変化させ、キューの先頭から順にパケットを送出するFIFO (First-In First-Out)方式と、ローカル内で同じ宛先のパケットが連続するようにパケットを挿入するローカル同一宛先パケット連続方式 (Local)、リング上を流れるパケットの宛先情報を用いて同じ宛先のパケットが連続するようにパケットを挿入するグローバル同一宛先パケット連続方式 (Global) の3方式を比較した。

図4は、各ノードのトラフィック量が72Gbpsの時に、送信部で検出される無信号区間の、400Gb/sのリンク帯域に対するビットレートの分布を示している。FIFO方式では、552Byte以下の短い無信号区間が多く発生し、無信号区間のうち1500Byte以上の無信号区間は31%である。一方で2種類の同一宛先パケット連続方式では、多くが1500Byte以上の無信号区間であり、無信号区間のうち55%以上を占めている。また、本方式では無信号区間を利用できるため、送信部で検出される無信号区間の全体量がFIFO方式に比べて少ない。この結果から、本手法によりまとまった無信号区間を生み出し、帯域利用効率を向上できていることがわかる。

図5に各リンクの帯域利用効率を示す。FIFO方式では75%以上帯域を利用できていないが、同一宛先パケット連続方式 (Address-Based) では帯域利用効率が94%に達する。これは、同一宛先パケット連続方式ではまとまった無信号区間が生じているため、パケットを効率的に挿入できるためである。FIFO方式では60Gb/s以上になると短い無信号区間が増加し、リングネットワーク上の帯域の25%を利用できない無信号区間を占める。

図6に各ノードでの平均遅延時間を示す。両方式ともトラフィック量増加に対して遅延時間が増加する。FIFO方式ではトラフィック量が60Gb/sを超える時点で、遅延時間が急激に増加する。一方で、同一宛先パケット連続方式では、FIFO方式に

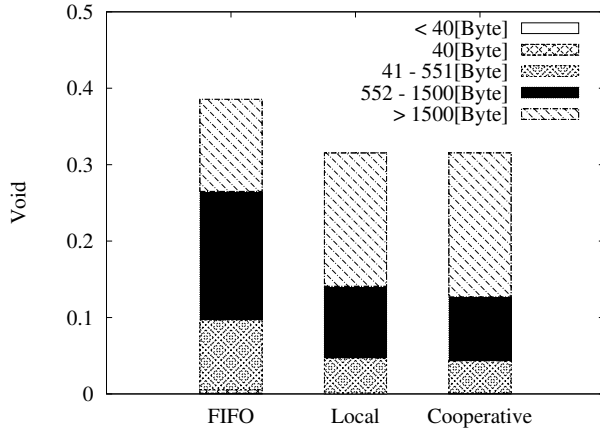


図 4 送信部における無信号区間のビットレート分布

比べて高トラフィック時にも遅延時間が小さい。また、Global 方式は Local 方式に比べても遅延時間が小さくなる。同一宛先パケット連続方式では宛先間でのパケットの順序逆転が発生するが、低トラフィック時には帯域が逼迫しておらず速やかにパケットを挿入できるため、遅延に影響を及ぼさないことがわかる。また、高トラフィック時には、同一宛先パケット連続手法を用いることでパケットの挿入可能な無信号区間が到着する頻度が増し、パケット挿入までの遅延時間が抑えられていることがわかる。

5. まとめ

非同期光パケットリングネットワークの帯域利用効率低下問題を解決するために、同じ宛先のパケットを連続して挿入する帯域利用効率を向上させる方式を示した。従来の FIFO 方式での最大帯域利用効率は 75% であるのに対し、本手法では帯域利用効率が 94% まで達した。また、各ノードのトラフィックレートが 75Gbps の時、Cooperative 方式はキュー内での遅延時間が Local 方式に比べて 20% に減少した。現在は、宛先に偏りがあるトラフィックパターンに対する本手法の適用を進めている。

謝 辞

本研究は情報通信研究機構 (NICT) の委託研究の一環である。

文 献

- [1] M. Herzog *et al.*: "Metropolitan Area Packet-Switched WDM Networks: A Survey on Ring Systems," IEEE Commun. Surv. Tut., vol.6, no.2, 2004.
- [2] M.C.Yuang *et al.*: "HOPSMAN: An Experimental Optical Packet-Switched Metro WDM Ring Network With High-Performance Medium Access Control," J. Opt. Commun. Netw., vol. 2, no. 2, 2010.
- [3] H.Castel *et al.*: "Performance Analysis of an Optical MAN Ring for Asynchronous Variable Length Packets," Telecommunications and Networking — ICT 2004, 2004.
- [4] V.H.Nguyen *et al.*: "L-MAC: A fair access protocol for passive asynchronous optical ring networks," IEEE ICC., 2006.
- [5] V.H.Nguyen *et al.*: "Dynamic Intelligent MAC Protocol for Metropolitan Optical Packet Switching Ring Networks,"

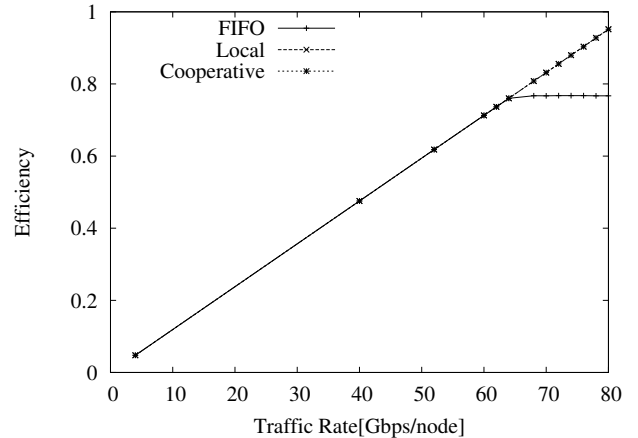


図 5 帯域利用効率

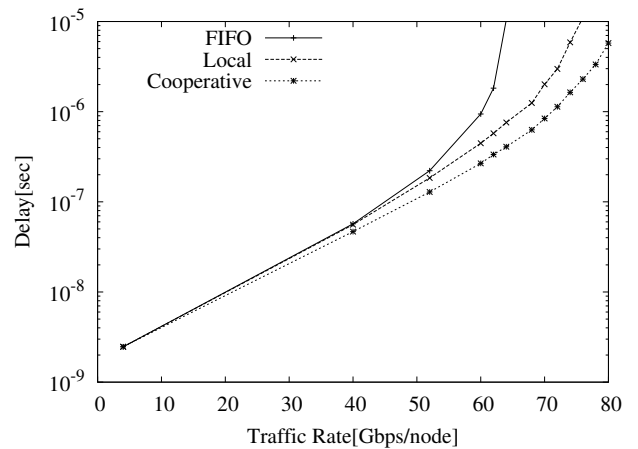


図 6 遅延時間

- IEEE ICC., 2006.
- [6] T.Kitamura *et al.*: "A MAC Scheme of Variable Length Packet in HORNET," IEICE IEIC Tech. Report, vol.102, no.514, 2002.
- [7] K.S.Kim *et al.*: "Design and performance evaluation of scheduling algorithms for unslotted CSMA/CA with back-off MAC protocol in multiple-access WDM ring networks," Information Sciences, 2003.
- [8] T.Eido *et al.*: "Performance of Optical Ring Architectures with Variable-Size Packets In-Line Buffers vs Semi-synchronous and Asynchronous Transparent MAC Protocols," FITraMen 2008, 2009.
- [9] T. Eido *et al.*: "Performance of Optical Ring Architectures with Variable-Size Packets: In-Line Buffers vs Semi-synchronous and Asynchronous Transparent MAC Protocols," HET-NETs'05, 2005.
- [10] D. Wonglumsom *et al.*: "Experimental Demonstration of an Access Point for HORNET—A Packet-Over-WDM Multiple-Access MAN," J. LightWave Technol., vol.18, no.12, 2000.
- [11] OMNeT++ Community Site, <http://www.omnetpp.org/>