

超高速高品質次世代IP/WDMシステム

An IP/ATM over WDM system for ultra high speed and quality guaranteed network

中 平 佳 裕
Yoshihiro Nakahira

大 柴 小 枝 子
Saeko Oshiba

川 原 正 人
Masato Kawahara

要 旨

超高速で高品質な通信サービスを提供可能なIP・ATM over WDMシステムの構成方法を提案している。提案システムは、光XC (ADM) に対して、トラフィックの集中するノード間に光パスを動的に設定するよう指示し、中継ノードではIPを光レイヤでカットスルーさせることで、スループットや遅延時間特性等を向上させる。また、QoSの保証やクラス別サービスを実現するために、IntServやDiffServの実装方法を検討する。

1. ま え が き

近年の通信における大きなトピックスとして、IP (Internet Protocol) サービスの普及とWDM (Wavelength Division Multiplex) 技術^{1),2)}の実用化を挙げることができる。現在、IPルータに波長多重伝送装置を接続したシステムが稼動しつつあり、光クロスコネクタや光ADM (Add/Drop Multiplex) といった装置³⁾も研究されている。しかしIPレイヤとWDMのレイヤとで協調動作を行いIPのトラフィックに応じて動的に光パスを割り当て、QoS (Quality of Service) の向上に役立てる仕組みは実現されていない。

我々は、IP・ATM (Asynchronous Transfer Mode) レイヤと光レイヤのレイヤ間協調動作による運用が極めて重要になると考えている⁴⁾。これを上手に行うことができれば、低コストで効率が良く、高いQoSを実現することが可能である。

本稿では、今後必要となる網の要求条件を定め、これを満たすネットワークを実現するためのシステム構成を検討する。

2. 提案システムの構成

2.1 次世代バックボーン網に要求されること
21世紀初頭のバックボーンやメトロの通信網のインフラに求められる機能を以下のように規定する。

- (a) ノードのスループットは1 Tb/s ~ 10Tb/s以上 (1 Tb/s=10Mb/s × 10万加入同時アクセスを想定)
- (b) QoSの提供が可能であること。レイヤ3サービスは主にIPだが他サービスも可能であること
- (c) 過去に構築されたシステムを継承可能であり、かつ低コスト化が可能な構成であること

上記の条件を実現するための網の構成を、プロトコルスタックの観点から検討する。

図1(1)は現在使用されている典型的なプロトコルスタックである。しかし、この構成ではシステムが重

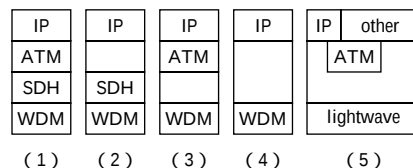


図1
プロトコルスタック
Fig. 1
Protocol stack



中平佳裕
ネットワークシステム
ビジネスグループ
ネットワークシステム
開発センタ
次世代光通信プロジェ
クト



大柴小枝子
ネットワークシステム
ビジネスグループ
ネットワークシステム
開発センタ
次世代光通信プロジェ
クト



川原正人
ネットワークシステム
ビジネスグループ
ネットワークシステム
開発センタ

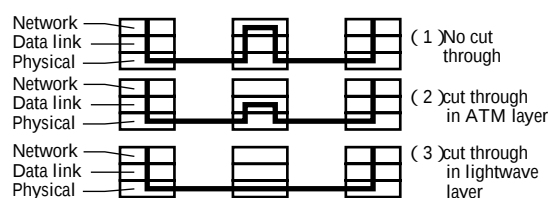


図2 カットスルー機能
Fig. 2 Cut-through function

く、また各階層で重複する機能が存在するためコストが大きいと言われる。このため、図1(2)および図1(3)に示す構成をとるべきであるとの意見や、さらには図1(4)に示す構成で十分であるとの意見もある。

ここでは、上記の条件(a)～(c)を満たす構成として、図1(5)に示す構成の可能性を検討する。

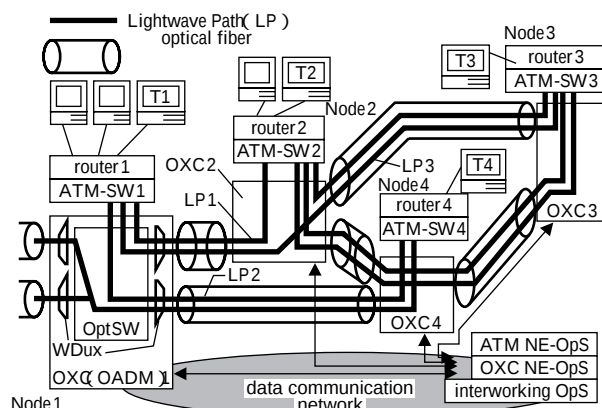
まず、各レイヤの機能について検討する。通常、ネットワーク層(L3)の機能はエンド・エンドでの情報転送機能をトランスポート層(L4)に提供することであり、データリンク層(L2)の機能は隣接ノード間に接続を設定し、これをL3に提供することであり、物理層(L1)の機能は伝送媒体をL2に提供することである。しかし、近年、ネットワーク層からの指示を受けて、トラフィックの多い2ノード間にデータリンクを設定し、中継ノードではIPレイヤをカットスルーさせるシステムが登場している。図2(2)がこれに相当する。このシステムはマルチレイヤスイッチまたはMPLSと呼ばれ、IPルータが転送すべき中継パケットの量を大幅に削減することを可能にする。

我々は、さらにスループットやQoSを向上させるため、L1である光レイヤとL3のIPレイヤとのインターワーキングが重要であると考え、図2(3)に示すような光レイヤでのカットスルーを提案した⁴⁾。IPトラフィック観測と光パス配置を連携させる機能を持つ装置を、ここではIP/lightwaveマルチレイヤスイッチと呼ぶ。以下では、これについて詳細に検討する。

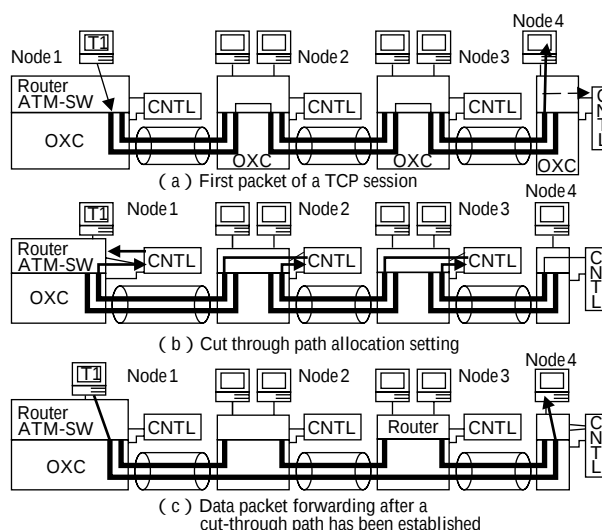
2.2 ネットワーク構成

IP/Lightwaveマルチレイヤスイッチを実現するために、つぎのような重要な検討項目がある。

- (a) L1とL3の協調方式
- (b) トラフィック観測技術
- (c) シグナリングプロトコルとその手順
- (d) IntServ (Integrated Service) やDiff-Serv (Differentiated Service) とのマイグレーション



(1) IP-ATM/lightwave multilayer switch (central control type)



(2) IP-ATM/lightwave multilayer switch (distributed control type)

図3 提案するシステムの構成

Fig. 3 Proposed system configurations

(e) カットスルーによるQoS向上の効果

(f) ハードウェアの実現方法およびデバイス技術
本稿では主として、(a)、(b)について検討する。

図3に提案するシステムの構成の例を示す。図3(1)は全体を管理するOpS (Operation System) を持つ構成を、図3(2)は分散制御を行う場合の構成を示す。

つぎに、動作について説明する。図3(1)において、たとえば端末T2から端末T1にパケットを転送する場合、端末T2から送出されたパケットはルータ2からルータ1へATM-SW2とATM-SW1によって設定されたVC (Virtual Channel) を介して転送される。このVCは、光パスLP (Lightwave path) 1の上に設定されている。LP1はATM-SW2とATM-SW1の間に設定された光パスであり、光クロスコネクタOXC (Optical Cross

Connect) 2とOXC 1によって提供されている。パケットがT 1からT 3に転送される場合には、LP 3が使われる。このとき、OXC 2において光レイヤでのIPカットスルーが行われる。

ところで、すべてのノードの組を光パスで結ぶことが可能ならば、すべてのパケットを1ホップで転送することが可能である。しかし、現状では使用可能な波長数に制限があり、光パスで全ノードをフルメッシュ接続するのは困難である。図3(1)の例で、端末T 2から端末T 4にパケットを転送する場合、一旦、LP 1を介してルータ1に転送してから、LP 2を介してルータ4に転送する。これはノード2とノード4を直接接続する光パスが存在しないためである。この間に光パスを設定しないのは、このノード間のトラフィックが他と比べて少ないということが前提となっており、もし、この間のトラフィックが多いならば、この間を直接接続する光パスを設定すべきである。

なお、図3に示す構成ではATM-SWが使われているが、厳密なQoSの保証が不要ならば省略できる。

本提案方式の最大のメリットは、光レイヤで大量のIPフローをカットスルーするため、IPルータやATMスイッチ等の上位レイヤ装置の実装容量を大きく削減できる点にある。

2.3 光波パス切替えの方法の提案

IP/Lightwaveマルチレイヤスイッチを動作させるには、IPのトラフィックを把握し、光波パスを再配置するためのインタワーキングの仕組みが必要である。以下では、この仕組みについて5種類の方法を提案し、議論する。

第1の方法<定時ポーリングによる方法>

この方法は、レイヤ間協調オペレーションシステム(Inter-working OpS)が一定時間間隔でATM網のオペレーションシステムATM NE(Network Element)-OpSに対してトラフィックを問い合わせるというものである。ATM網のOpSは、各ATM交換機のMIB(Management Information Base)にSNMP(Simple Network Management Protocol)等を用いてトラフィック情報を収集する。この情報は解析され、OXCとATM-SWに対して光波パスとVCを最適な形に再配置するようにレイヤ間協調OpSから命令が下される。また、VCのコネクションを守るために、VCが設定されている光パスを移動させる際は、あらかじめそのVCを別の光パス上に移動させる制御が必要となる。ATMが実装され

ない場合には、流量情報はATM交換機でなくIPルータから得ることになる。

第2の方法<ノードのトラップによる方法>

この方法は、IPルータやATM交換機にトラップを設け、特定の経路の流量が閾値を超えると、自動的にレイヤ間協調OpSに通知し、これを受けて混雑を緩和するように光波パスを設定する方法である。この方法では、第1の方法より早いタイミングで輻輳情報を入手できるが、反面、網全体が混雑した際、オペレーションシステムは多数のIPルータやATM交換機から押し寄せる情報を混乱なく処理する必要がある。

第3の方法<pingとtracerouteによる方法>

これは、各ノードが定期的に他のノードに対してpingやtracerouteをかけて自ノードを中心とした網の輻輳状態を把握し、それがある閾値を超えた場合に改善要求を出す方法である。

第4の方法<アプリケーションレベルでの監視による方法>

この方法は、各端末に大容量のトラフィックが流れそうなコマンドが投入されないかを監視する機能を持たせ、大きなトラフィックが流れそうな場合、自動的にレイヤ間協調OpSに知らせ、これに対応したパスのリアレンジを行う方法である。

第5の方法<フロードリブによる方法>

この方法は、図3(2)に示す方法である。ここでは、あるTCP(Transmission Control Protocol)セッションの最初のパケットが送信側のエッジルータに届いた際、そのパケットはホップバイホップで宛先に転送される。宛先側のエッジルータをパケットが通る際、これをトリガとしてカットスルー用の光パスを設定すべきか否かの判断が行われる。設定すべきと判断された場合、このエッジノードから上流側のノードにカットスルー用の光パス資源情報を転送する。上流側でカットスルーが可能であれば、その光パスを延長してその情報をさらに上流に伝える。この動作を送信側のエッジノードまで続け、カットスルー用光パスを設定する。

この方法は分散制御が可能である。ただし、どのようにして複数のTCPのセッションを1本の光パス上に多重するかが課題となる。

2.4 QoS保証とIntServやDiffServのサポート方法

本システムでは、QoS保証はATMの機能を用いて実現する。もし、QoS保証が必要なTCPのセッションが発生した場合、AAL(ATM Adaptive Layer)タイプ1またはタイプ2VCを一つ、そのTCPセッションのため

に設定する。

IntServの提供方法には二通りの方法が考えられる。第1の方法は、光パスによるカットスルーとIntServを直接には関係させない方法である。光パスは、VCの要求に先立って、トラフィックの情報に応じてマクロ的に最適と思われる配置をとる。もし、ある経路に関して容量が不足すれば、資源が完全に枯渇する前に新たなVCの設定が可能となるように光パスを再配置する。これは2.3節の方法1～4で有効な方法である。第2の方法は、TCPのセッションごとに光パスの再配置を行うべきか否かの判断を行い、L1, L2, L3を直接連動させるというものである。これは、2.3節の方法5で有効な方法である。

DiffServのサポートに関しては、エッジノードにDiffServプロトコルを理解させ、求められたTCPセッションに適したVCを与えるように実装する。宛先側のエッジノードで、次のネットワークに対してDiffServのメッセージを転送する機能も必要である。たとえば上記実装によりQoSのサポートが可能である。

3. テストベッドの試作

提案システムの実現性の確認のため、光ADM装置を試作した。装置の外観を写真1に、実験装置の構成を図4に示す。現状では、図中のelectric partにはO/E, E/O変換器およびSDHアナライザが実装されている。本実験系において、光ADMのNE-OpSから光パスの切替え指示を出し、命令送出から切替え完了までの時間を計測した。光パスの切替えを指令するTCPのセッションのIPパケットの先頭ビットがNE-OpSより送出されてから、光パスの切替えが完了しO/E変換器でビット同期が確立するまでの時間をオシロスコープで測定した結果、19ミリ秒という値を得た。IPカットスルーを主目的とした光パスの動的配置の実験は、恐らく世界でも初めてではな



写真1 実験装置の外観
Photo 1 Test-bed system

本論文はForum Proceedings of World TELECOM 99+Interactive 99に掲載される論文を日本語に翻訳したものである。本論文の内容は、すべて著者の見解であり、国際電気通信連合(ITU)の見解とはまったく関係を持たない。

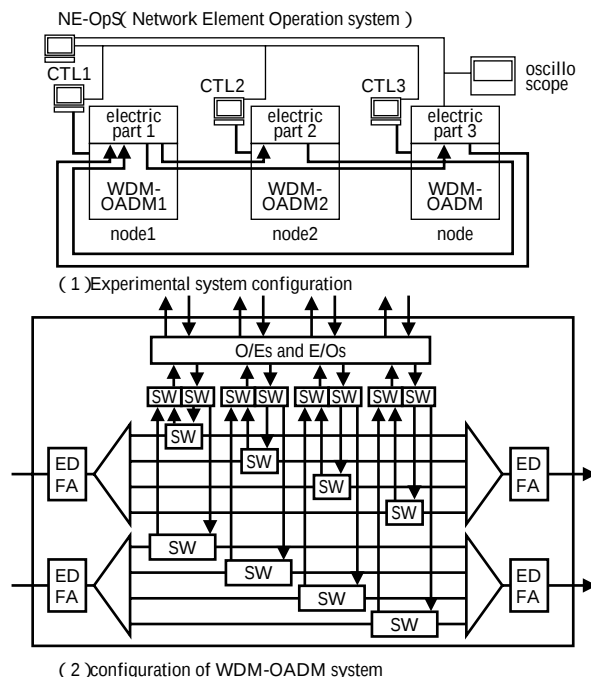


図4 実験装置の構成

Fig. 4 System configuration of test-bed system

いかと思われる。本実験は、2.3節の第1と第3の方法の実現に向けたものである。今後、切替時間19ミリ秒という実力で適用可能なシステムの運用方法について検討を進める。

4. あとがき

超高速で高品質な通信サービスを提供可能な次世代IP over WDMシステムの構成方法を提案し、基本実験の成果を述べた。今後、IP装置との協調動作部の実装を進め、実用化・製品化に向けて研究開発に取り組んでいく。

5. 参考文献

- (1) Yano *et al.*: "2.6 Tera bit/sec WDM.." ECOC96 post-deadline, 1996
- (2) ITU-T Recomnd G.872
- (3) Nakahira *et al.*: "An optical crossconnect..." ICC97, No. 36.3, June 1997
- (4) 中平, 他: "波長多重光XC網でのIP・ATMサービス提供方式の一検討" 信学総大B-12-13, pp. 570 ~ 571, Mar. 1999