

ラベルスイッチングネットワークにおける QoS 保証の実現

松本勝之[†] 松古典夫[†] 西村浩二[‡] 相原玲二[‡]

[†]広島大学 大学院 工学研究科 [‡]広島大学 総合情報処理センター

近年、QoS を要求するアプリケーションが登場し、またそれらが利用される機会が増えていることから、ネットワークには QoS の保証が求められるようになった。しかし、既存のネットワークでは必ずしもアプリケーションごとに十分な QoS を提供できていない。その原因として、帯域の不足、ルータの負荷による遅延の発生などがあげられる。それらの問題に取り組むため、本稿では CSR (Cell Switch Router) をベースとするラベルスイッチングネットワークを構築した。また、そのラベルスイッチングネットワークの性能評価を行い、CSR が高いパフォーマンスを達成できることを示した。さらに、その上でルータの負荷分散、帯域の分散などを実現し、効率的に QoS を提供する方式の検討を行った。

QoS Guarantee in Label Switching Networks

Katsuyuki MATSUMOTO[†], Norio MATSUFURU[†], Kouji NISHIMURA[‡], and Reiji AIBARA[‡]

[†]Graduate School of Engineering, Hiroshima University

[‡]Information Processing Center, Hiroshima University

Recently, network applications that require QoS (Quality of Service) are increasing, therefore it is important to guarantee QoS in the networks. However, current networks don't necessarily provide enough QoS such as bandwidth, delay, packet loss ratio, and so on. To solve the problems, we built label switching networks using CSRs (Cell Switch Routers). We evaluated the performance of a CSR network and indicated that the CSR network can achieve high performance. Also, we consider methods of realizing load sharing, using bandwidth efficiently, and guaranteeing QoS.

1 はじめに

近年、インターネットは様々な用途に利用されるようになり、それに伴いネットワークのトラフィックは年々増加傾向にある。最近では、インターネットを使った動画配信や音声配信なども行われるようになり、今後トラフィックのさらなる増加が見込まれている。

最近普及してきた映像や音声を扱うアプリケーションは、遅延やパケットロス率などの QoS (Quality of Service) を要求する。しかし、既存のネットワーク環境では十分な QoS を保証できていない。既存のネットワークは Best Effort 型と呼ばれる処理を行っているため、帯域、低遅延などを特定のアプリケーションに保証することは難しい。今後トラフィックが増加する中、QoS を要求するアプリケーションに対して、どのように QoS を保証していくのが今大きな問題になっている。

本稿では、MPLS (Multi Protocol Label Switching) [1] の実装の 1 つである CSR (Cell Switch Router) [2] [3] をベースとするラベルスイッチングネットワークを構築した。MPLS では、フロー

(同じ始点・終点・ポート番号を持つ一連のパケットの流れ) に固定長の短いラベルを割り当て、ラベルを交換することによって高速にパケットの転送を行う。どのようなラベルを用いるかは、データリンク層に依存する。例えば、データリンク層に ATM を用いる場合、VPI/VCI (Virtual Path Identifier/Virtual Connection Identifier) がラベルとして使われる。CSR は、データリンク層に ATM を用いることで、ATM の特徴である高速な転送や QoS の保証を実現している。

本稿では、CSR をベースとするラベルスイッチングネットワークの性能評価を行った。また、その上でルータの負荷分散、帯域の分散などを実現し、効率的に QoS を提供する方式の検討を行った。

2 CSR

2.1 概要

図 1 に CSR の概要を示す。CSR はデータの転送に 2 種類のパスを使う。1 つは隣接 CSR との間に

確立される Default Path と呼ばれるパスで、もう1つは、ラベルスイッチングネットワークの ingress CSR から egress CSR までを結ぶ Dedicated Path と呼ばれるパスである。

Default Path は、CSR が通常のルータと同じようにデータを転送する場合に用いられる。パケットはルーティング処理を行った後、隣接 CSR へ Hop by Hop で転送される。一方、Dedicated Path は高速に転送したいフロー、あるいは QoS を提供したいフローを転送する場合に用いられる。図 1 に示すように、Dedicated Path では Edge から Edge まで専用の VCC (Virtual Channel Connection) を使ってパケットの転送を行う。ATM スイッチを介して Edge と Edge が直接接続されているため、中間 CSR でのルーティング処理が不要になる。それにより、通常の転送よりも高速に転送できる。また、中間 CSR の負荷も軽減することができる。

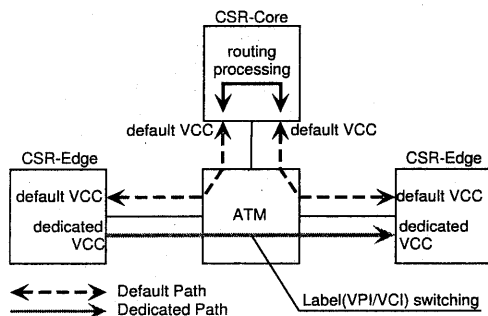


図 1: CSR 概略図

すべてのフローに対し Dedicated Path が確立できれば理想的だが、残念ながらラベルリソースは有限であり、すべてのフローにパスを確立することはできない。そこで、高速に転送する必要のあるフロー、データ量の多いフロー、QoS を要求するフローなどに限って、Dedicated Path を確立することが重要である。

2.2 パス確立方式

Dedicated Path は隣接 CSR 間でネゴシエーションが行われることにより確立される。パス確立方式は、確立の仕方によって次の3つに分けられる。

1. データドリブン方式[4]

2. RSVP (Resource ReSerVation Protocol) ドリブン方式[4] [5]

3. トポロジードリブン方式[6]

データドリブン方式の場合、始点・終点アドレス・ポート番号など、特定の条件を満たすパケットの到着がパス確立のトリガーとなる。そのパケットが CSR に到着したとき、CSR は隣接 CSR にパス確立リクエストメッセージを出し、ネゴシエーションを行ってパスを確立する。RSVP ドリブン方式では RSVP メッセージがトリガーとなる。他はデータドリブン方式とほぼ同様である。トポロジードリブン方式はトポロジー情報に基づき、パケットが実際に到着する前にパスをあらかじめ確立しておく方式である。

データドリブン方式の場合、実際のパケット到着がトリガーとなるので、パス確立に遅延が発生する。パスが確立されるまでの間は、Default Path を使って転送を行わなければならない。しかし、仙波ら[7]はパス確立遅延を軽減する方式を提案しており、遅延の軽減は可能である。また、中間 CSR に高い負荷がかかっている状態では、パスが確立できない場合があることを確認している。この方式では、異なる始点・終点アドレスを持つフローに対し排他的にパスを確立できるが、ラベルリソース上の制約を受ける。

トポロジードリブン方式の場合、事前にパスが確立されているため、パス確立に遅延が発生することはない。また、1つのパスを複数のフローで共用する方式のため、使用するラベルリソースを抑えることができる。しかし、複数のフローが1つのパスを使うので、フローごとに十分な QoS を保証できない場合もある。

3 評価

図 2 のようなネットワークを実際に構築し、性能測定を行った。CSR はいずれも、BSD/OS 3.1 + KAME + CSR で構成されており、IPv4 だけでなく IPv6 にも対応している。ATM の帯域は 35Mbps である。ただし、AAL5 を使用していること、ATM セルは 5 オクテットのヘッダと 48 オクテットのペイロード部分から成り立っていること、さらに MTU=1512 に設定していることを考慮に入れば、実際は 30.4Mbps 程度の転送が限界である。

CSR、および ATM は東広島市にある広島大学と、広島市にある広島市立大学に設置した。CSR のハードウェアのスペックは、表 1 の通りである。ATM には Hitachi AN1000-5 を使用した。また CSR-Edge や CSR-Core のネットワークインターフェースカードに、155Mbps の ATM NIC を使用した。

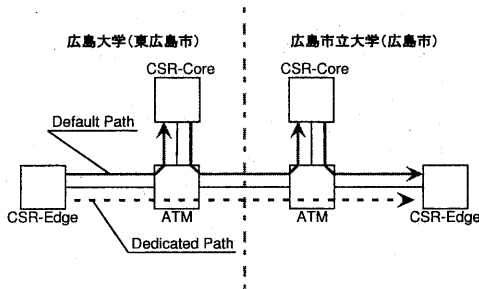


図 2: CSR ネットワーク

		CPU	Memory
広島大学側	CSR-Core	Pentium III 450MHz	128MB
	CSR-Edge	Pentium II 333MHz	64MB
広島市立大学側	CSR-Core	Pentium II 333MHz	64MB
	CSR-Edge	Pentium Pro 180MHz	128MB

表 1: ハードウェアスペック

3.1 測定条件

パス確立はフロードリブン方式で行った。今回は特定のポート番号を持つフローが到着した際、トリガーがかかり Dedicated Path が確立されるように設定した。性能測定に利用したアプリケーションは、Netperf である。Netperf を使って、メッセージサイズ 8K のデータを 60 秒間送りだし、スループットを測定した。また、Socket Size を変化させることで送り出すデータ量を調節した。

測定は、Default Path を使った転送の場合と Dedicated Path を使った転送の場合の 2 通り行った。また、HOP 数は、1HOP (隣接間)、2HOP、3HOP と変化させた。さらに、IPv4、IPv6 の両環境で測定を行った。

3.2 測定結果

測定結果を図 3、および 4 に示す。図 3 は IPv6 環境において、Default Path を使って転送を行ったときのスループットである。HOP 数が増えるにつれて、スループットは減少している。Socket Size が大きいところでは差がないが、これは ATM ヘッダや AAL5 のオーバーヘッドにより 30.4Mbps 以上のスループットがでないため、帯域が飽和状態にあるためである。

図 4 は同じく IPv6 において測定を行ったものであるが、こちらはフロードリブンによって Dedicated Path を確立し、それを使って高速転送を行ったときの結果である。ホップ数が増えても、図 3 の時とは違い、スループットはさほど減少していない。ホップ数が増えているにもかかわらずスループットがあまり減少していないのは、中間 CSR でルーティング処理が行われなくなったことや、パケットからセルへの分解、逆にセルからパケットへの再構成が行われなくなったことによるものである。今回の実験データは、CSR はホップ数が増えた場合においても Dedicated Path を確立することにより高速転送を実現できる、ということを示している。なお、IPv4 においてもほぼ同様の結果が得られている。

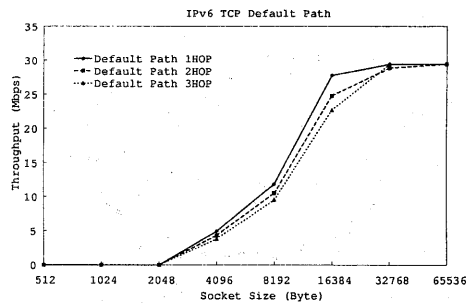


図 3: IPv6、Default Path における TCP スループット測定結果

4 既存方式の問題点

前述したように、パス確立には 3 方式があるが、いずれの方式でもパス確立のリクエストメッセージはルーティング情報に従って Next Hop CSR に出される。そのため、Dedicated Path は RIP [8] や OSPF

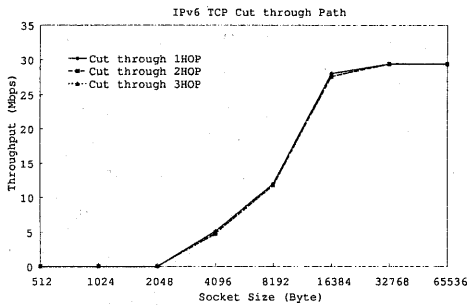


図 4: IPv6、Dedicated Path における TCP スループット測定結果

[9]などのルーティングプロトコルの提供するルートにそって確立される。従って、同じ終点アドレスを持つフローに対して ingress CSR から egress CSR へ確立されるパスの経路は 1 つである。

図 5 に示すように、1 つのルートに多くのパスが確立され帯域に余裕がない状況でも、それ以外のルートを使って Dedicated Path が確立されることはない。そのような状況下では、パスが確立できず十分な QoS を提供できない可能性がある。しかし、迂回ルートを使うことができればパスを確立できる場合もある。

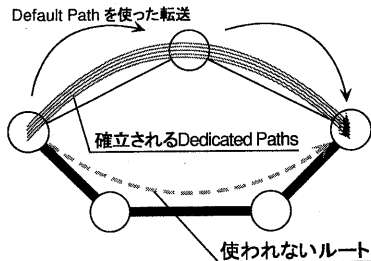


図 5: 2 通りのルートがある場合

5 提案方式

以上の問題点を解決するため、同一終点アドレスを持つフローに対し、複数のルートを提供する方式を提案する。複数のルートを提供することで負荷分散、帯域分散を実現し、また QoS の保証できるルートを選択できるようにする。

複数のルートを提供するには、レイヤ 3 により提供する方法と、レイヤ 2 により提供する方法の 2 通りの方法がある。それぞれの方法について検討する。

1. レイヤ 3 で迂回ルートを提供する。カーネルのルーティングテーブルを変更し、これまでは終点アドレスのみから決まっていた Next HOP を、始点・終点アドレス・ポート番号・プロトコル・QoS 関連情報の組によって決めるようにする。
2. レイヤ 2 で迂回ルートを提供する。ATM スイッチを設定し、VCI の値によって異なるルートを通るようにする。それにより複数のルートを提供する。

5.1 レイヤ 3 迂回ルート

まず、レイヤ 3 で迂回ルートを提供する仕組みを示す。通常、カーネルのルーティングテーブルは、終点アドレスとそれに対応する Next HOP の組で構成されている (表 2-A)。終点アドレスのみを Next HOP 検索のキーとするため、同じ終点アドレスを持つパケットはたとえ始点アドレスが異なっても、同じ Next HOP に送られていた。このテーブルを変更し、終点アドレスだけでなく始点アドレス、始点ポート番号、終点ポート番号、プロトコル (TCP/UDP)、QoS 関連情報も Next HOP 検索のキーとして扱えるようにする (表 2-B)。これにより、同じ終点アドレスを持つフローに対し、複数の経路を提供する。

A:変更前

Destination Address	Next Hop
192.168.4.0	192.168.2.1

B:変更後

Source Address	Destination Address	Source Port	Destination Port	Protocol	QoS	Next Hop
192.168.1.1	192.168.4.0	*	*	TCP	1	192.168.2.1
192.168.1.2	192.168.4.0	*	*	UDP	*	192.168.3.1

表 2: ルーティングテーブルの変更

新しいルーティングテーブルを用いてルーティングを行った例を、図 6 に示す。今、左の 2 つのノードが右の同じ終点に対して、パケットを送信しようとしている。そのとき、CSR Network の ingress CSR のルーティングテーブルが、表 2 の B であったとす

ると、それぞれの始点から出たパケットは図の矢印の向きに転送される。パケットは異なったルートを通り、パスは異なったルートで確立される。

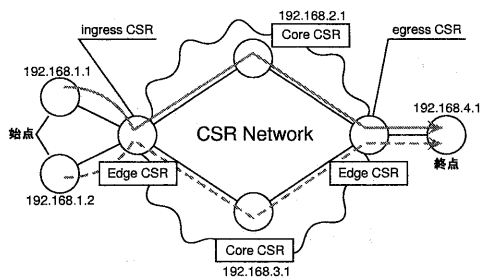


図 6: 異なったルートを通る Dedicated Path

現在、この方法の実装を行っている。現段階で、図 6 に示したような、終点アドレスは同じだが始点アドレスの異なるフローに対し、異なったルートを使って転送を行うことができています。また、そのルート上に Dedicated Path を確立することができています。実装では、通常のルーティングテーブル (表 2-A) と変更したルーティングテーブル (表 2-B) を併用している。上流からパケットが送られてきたとき、変更したテーブルにそのパケットに関するエントリーがあれば、そのテーブルで指定された Next HOP に転送する。エントリーがなければ、通常のテーブルを使った転送を行う。エントリーは現在、手動で追加している。

5.2 レイヤ 2 迂回ルート

次に、レイヤ 2 により迂回ルートを提供する方法について述べる。この方法では、レイヤ 3 の方式と違い、ルーティングテーブルの変更は必要ない。

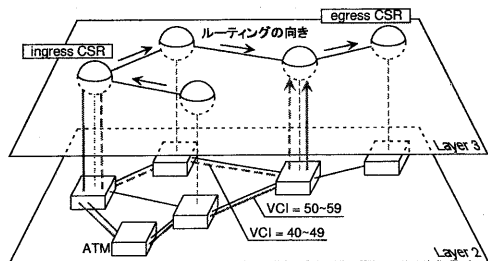


図 7: レイヤ 2 による迂回ルート

ATM インターフェースに出力されたフローは、そのフローに対応する VCI を使って Next HOP へ送り出される。ここで ATM スイッチを設定することにより、VCI の値によって異なったルートを通す。図 7 に例を示す。この例では、VCI = 40~49 のフローは上の、VCI = 50~59 のフローは下のルートを通るように設定してある。このようにして、迂回ルートを提供する。

レイヤ 2 で迂回ルートを提供する場合、ループの発生に注意しなければならない。図 8 に正しく迂回ルートを設定した例と、誤って迂回ルートを設定した例を示す。ループを回避するためには、迂回先ノードが、迂回させたいフローをルーティング情報に従って転送を行った場合の経路上にあってはならない。また、迂回先ノードが自分自身であってはならない。図 8 の場合では、迂回先ノードとして指定できるのは 1,2,3 のいずれかである。正しく迂回させた場合、経路上にあるノードにパケットを送るので、ループが発生することはない。

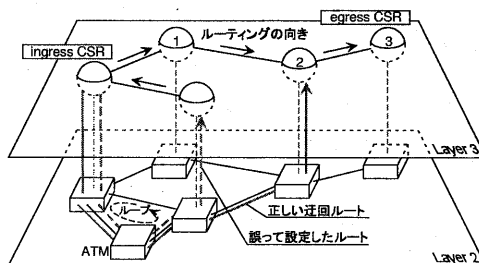


図 8: ループの防止

現在、レイヤ 3 による方法と同様、レイヤ 2 による方法の実装も行っている。現在までのところ、特定のフローに対して特定の VCI を設定し、迂回ルートを使って転送を行うことができています。ただし、迂回ルートの経路などは事前に設定している。

5.3 考察

2つの方法には、それぞれ問題点がある。レイヤ 3 のみのアプローチでは、ATM の帯域情報、パスの情報などがわからない。そのため、レイヤ 3 のみの情報からフローに QoS の保証できるルートを提供することは難しい。また、フローを迂回させる場合、迂回させるフローごとにルーティングテーブルにエントリーを追加しなければならない。従って、

ルーティングテーブルは大きくなりがちである。

一方、レイヤ2による方法ではATMの帯域やパスの情報を利用することができるが、ルーティングが問題になる。レイヤ2のみの情報でルーティングを行うことは難しい。ループを起こさないようにするためにはレイヤ3の情報が必要である。

以上のように、どちらか一方のみでQoSを保証するルートを提供することは難しい。従って、2つの方法をうまく組み合わせる必要がある。現在、レイヤ3のみ、レイヤ2のみの両面から実装を行いつつ、その2つの接点を探っているところである。例えば、レイヤ3によりルーティング情報を得、それをもとにレイヤ2でパスを確立する方法を検討している。

現在のところレイヤ3、レイヤ2いずれも、ルーティングテーブルやVCI、迂回ルートを手動で設定している。それを動的に行うためには、新たなルーティングプロトコルの開発が必要である。新たなルーティングプロトコルでは、帯域分散、負荷分散を考慮に入れたルーティングを行う。現在、OSPFのようなネットワークのトポロジー情報を持つことのできるプロトコルの上に、ネットワーク上の負荷、帯域などの情報を加え、QoSを提供できるルートを選択できるプロトコルを検討している。

6 おわりに

本稿では、実際にCSRネットワークを構築し、その性能評価を行った。それにより、CSRがHOP数の増加した状態でも高いスループットを維持できることを示した。

しかし、トラフィックの多い状態では、QoSアプリケーションに十分なQoSを保証できない場合があると考え、新たな方式を提案した。提案方式では、同一終点アドレスをもつフローに複数のルートを用意し、始点アドレスやポート番号により異なったルートを通るDedicated Pathの確立を可能にする。それを実現する方法は、レイヤ3を利用する方法とレイヤ2を利用する方法の2種類あるが、両方をうまく組み合わせることが必要であることを示した。提案方式では、複数のルートを通るパスを確立することで負荷分散を行い、またより多くの帯域確保を行う。さらに、トラフィックの多い状態でもQoSを提供できるようにする。

本研究の目的は、QoSを提供できるパスを動的に

確立するための新しいルーティングプロトコルの開発とその評価であるが、ルーティングプロトコルの開発は今後の課題である。

参考文献

- [1] Eric C. Rosen., et. al. Draft: Multiprotocol Label Switching Architecture, April 1999.
- [2] Yasuhiro Katsube, et. al. Draft: Cell Switch Router, December 1997.
- [3] 永見健一, 今泉英明, 中村修, 江崎浩. ラベルスイッチ技術 ~次世代高速光機能インターネットバックボーンの基盤技術~. 情報処理, Vol. 40, No. 7, pp. 718-722, July 1999.
- [4] Ken-ichi Nagami, et. al. Draft: Flow Attribute Notification Protocol Version 2 (FANPv2) Distributed Control Mode, December 1997.
- [5] A. Mankin, Ed., et. al. RFC 2208: Resource ReSerVation Protocol (RSVP), September 1997.
- [6] Yoshihiro Ohba, et. al. Draft: Flow Attribute Notification Protocol Version 2 (FANPv2) Ingress Control Mode, December 1997.
- [7] 仙波伸広, 西村浩二, 相原玲二. ラベルスイッチングにおけるラベルディストリビューションプロトコルの評価. 情報処理学会研究報告, 97-DSM-8 1997.
- [8] G. Malkin. RFC 1723: RIP Version 2, November 1994.
- [9] J. Moy. RFC 2328: OSPF Version 2, April 1998.