

Title	Split-Phase : 細粒度QoS指定指向型MPLS LSP設定手法の設計と実装
Author(s)	宇夫, 陽次郎; 宇多, 仁 ; 小柏, 伸夫; 篠田, 陽一
Citation	電子情報通信学会論文誌 B, J85-B(8): 1191-1198
Issue Date	2002-08-01
Type	Journal Article
Text version	publisher
URL	<a href="http://hdl.handle.net/10119/4667">http://hdl.handle.net/10119/4667</a>
Rights	Copyright (C)2002 IEICE. 宇夫 陽次郎, 宇多 仁, 小柏 伸夫, 篠田 陽一, 電子情報通信学会論文誌 B, J85-B(8), 2002, 1191-1198. <a href="http://www.ieice.org/jpn/trans_online/">http://www.ieice.org/jpn/trans_online/</a>
Description	

## Split-Phase : 細粒度 QoS 指定指向型 MPLS LSP 設定手法の設計と実装

宇夫陽次朗<sup>†\*a)</sup> 宇多 仁<sup>†</sup> 小柏 伸夫<sup>†</sup> 篠田 陽一<sup>†</sup>

Split-Phase : Design and Implementation of Multi-phase MPLS LSP Setup Method for Fine-grained QoS

Yojiro UO<sup>†\*a)</sup>, Satoshi UDA<sup>†</sup>, Nobuo OGASHIWA<sup>†</sup>, and Yoichi SHINODA<sup>†</sup>

あらまし 一般の利用者が接続されるネットワークにおいては、利用者ごとに非常に粒度の細かな要求が発生しやすく、通信フローの生成の時間間隔とほぼ等価、若しくはそれより細かな粒度での制御が必要となる可能性が高い。しかし既存の IP 層によるパケット制御機構では個別のフローに基づいた細粒度制御の導入は困難である。本論文では、筆者らが設計及び実装を行った *Split-Phase* について論じる。*Split-Phase* 機構は、IP 層の直下に位置するマルチプロトコル・ラベルスイッチング (MPLS) 技術における細粒度 QoS 制御を実現するための LSP 設定機構である。LSP を構成する要素のトポロジー依存部分と非依存部分に分解し非依存部分のパラメータの遅延評価を導入することで、フロー特性の変動頻度が大きな要求に対応可能な動的 LSP 制御を実現する。そのうえで、利用者からの要求を動的に反映させるネットワーク制御システムに対する *Split-Phase* 機構の反映について論じる。

キーワード 細粒度 QoS, Sub-IP, MPLS

## 1. インターネットアプリケーション

インターネットが設計されてから既に 30 年近い年月が経過した。その間にインターネットは広域に配備され、その利用者数は爆発的に増大した。この大きな成功の結果、インターネットはコンピュータネットワークの社会基盤となり、更に多様な要求を実現する圧力にさらされるようになった。

計算機技術及び伝送技術の発展の結果、計算機アプリケーションが実現可能な領域は拡大し続けている。それに伴い、インターネットが提供するサービスに対する要求も拡大しつつある。特に、通信品質に関する扱いに関しては様々な議論が行われている。

ネットワークのサービス品質 (QoS) には、対象となる問題や求められる特性に応じて様々なクラスが存在する。ネットワーク内の扱いをトラフィック若しくは顧客ごとに指定する技術はトラフィックエンジニアリン

グ (TE) と呼ばれ、現時点でも ISP (インターネットサービスプロバイダ) や IX (インターネットエクスチェンジ) は、顧客との契約に応じてサービスの区分化を行う場合がある。

サービスの区分化項目には最低帯域や最高遅延時間などがあるが、トラフィックの集約点である IX や ISP では処理しなければならない情報が膨大なため、フロー単位といった小さな粒度の処理が困難である。そのため現時点では、適用の粒度も顧客や装置のインタフェースを単位とすることが多い。しかし、ネットワークと利用者との接合部分においては、ネットワークの末端になるほど提供帯域が狭くなるため、通信品質及び通信制御に対して時間的/帯域的に細かな制御が必要になる。そのため、フロー単位若しくはそれ以上の細かな粒度での QoS 制御に対する需要は大きく、利用者若しくはその他のエンティティからの要求に応じてネットワークに反映させる動的 QoS 制御は多くの利用者に有効であろう。

本論文ではマルチプロトコル・ラベルスイッチング (MPLS) [1] ネットワークにおける、細粒度動的 QoS 制御方式 *Split-Phase* を提案する。*Split-Phase* は、MPLS における制御単位であるラベルスイッチン

<sup>†</sup> 北陸先端科学技術大学院大学, 石川県

School of Information Science, Japan Advanced Institute of Science and Technology, Asahidai 1-1, Tatsunokuchi-machi, Ishikawa-ken, 923-1292 Japan

\* 現在, 株式会社インターネットイニシアティブ技術研究所

a) E-mail: yuo@ayame.org

グパス (LSP) の制御をトポロジー依存処理とそれ以外の処理要素に分解し、トポロジー非依存部分の遅延反映を導入する。遅延反映はフローの生成や消失が頻繁に行われたり、フロー特性の変動頻度が大きいアプリケーションなど、時間粒度が細かな要求を満たすネットワークにおいて有効である。そのうえで、WIDE プロジェクトの研究会で構築した仮設ネットワークにおける、Split-Phase 方式の適用に関して論じる。

## 2. サービス品質制御技術

インターネット向け通信品質制御技術は今までも多くの提案が存在する。広く知られている技術は、IETF の標準化過程で検討されている統合サービス (Int-serv) [2] と区分化サービス (Diff-serv) [3] である。Int-serv はフローが通過するネットワークノードの資源を予約することで、フローの品質を保証する技術である。一方 Diff-serv は、各パケットに対して扱いのクラスを指定することで、ノードにおけるパケットの取扱い挙動を制御する技術である。Int-serv がある程度厳密な端点間の品質制御を目的とする一方で、Diff-serv はポリシーを共有する領域内におけるパケット挙動の区分化を実現する機構として設計されている。

一方で、今まで第3層によって提供されてきた網内の経路制御とパケットの転送機能を分離することで、通常のパケット処理よりも複雑な制御 (例えば TE など) の実現をめざすモデルとして “転送/制御分離モデル” (図1) が議論されている。MPLS は第3層の直下におけるネットワーク制御セマンティクスを提供する技術であり、このパラダイムを実現する基本技術として注目されている。

### 2.1 品質制御の分類

広域ネットワークに適用される制御技術は、制御の適用範囲の粒度が大きいほうが一般に対規模性能に優

れる。そのため現時点では、フロー単位の制御のような細粒度の制御は好まれず、静的な (始点アドレスの範囲や顧客先ユニットによるタグ付けに基づいた包括的な) 制御が一般的である。更に制御頻度に関して以下のように定義する。

(1) 品質制御が静的: 制御内容の変更間隔がネットワークのフローの生成及び消失間隔と比べ十分無視できるほど長い。

(2) 品質制御が動的: 制御内容の変更間隔がフローの生成及び消失間隔と同等若しくは短い [4], [5]。

制御粒度が大きい場合には必然的に静的な制御が適用されるが、ネットワークの状態や各利用者の必要に応じたネットワークの制御を想定する場合には、制御頻度を向上させて動的な制御を導入する必要がある。

### 2.2 細粒度品質制御

細粒度のネットワーク品質制御とは、フローの生成間隔とネットワーク内の品質制御の更新間隔が同じオーダとなる領域における制御を指す。このような場合、ネットワークの制御機構の更新オーバーヘッドの抑制が実現条件となる。更新オーバーヘッドの削減には、制御対象の更新を伝達する対象を削減する方法が有効である。

一般に第3層 (IP) ネットワークは各ノードが自律分散型に動作するため、パケットが通過する領域全体に対する制御が必要となり、領域の大きさに対する対規模性を維持できない。また、空間的な拡大はシグナリング等のプロトコルの伝達遅延を拡大する要因となり、前提となる細かな制御間隔の維持が困難である。そこで、第3層と別の領域での制御をラベルスイッチング技術によって導入し、対規模性能を維持しながら細粒度の制御が可能な方式を考案した。

## 3. マルチプロトコル・ラベルスイッチング

MPLS は IETF によって標準化が進められているラベルスイッチング技術である。

IP 層 (第3層) のパケット転送処理は各ルータで独立した動作として表現できる。各ルータはパケット転送時にパケットごとに以下の一連の処理を実行する。

- (1) パケットヘッダの解析
- (2) 転送方法の決定
- (3) 次転送先の決定と転送

これに対して、MPLS では 1~2 の処理を入口ラベルスイッチングルータ (LSR) のみで行う (同一転送クラス (FEC) 決定処理)。入口 LSR で決定された FEC

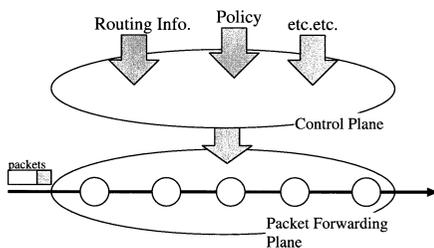


図1 転送/制御分離モデル

Fig. 1 The model of forwarding/control-separation.

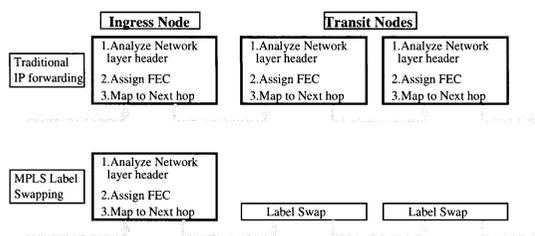


図 2 第 3 層転送と第 2 層転送 (ラベル置換)  
Fig. 2 L3 forwarding and L2/L2.5 label swapping.

情報はパケットにラベルとして付加されるだけでなく、ラベル値と FEC のマッピング情報として MPLS シグナリング機構を用いて MPLS ドメイン内に伝播される。途中 LSR はパケットに付加されているラベルから FEC 情報を、FEC から転送方法を導出してパケットを転送する ( 図 2 )。

MPLS では、このような境界ルータ及びドメイン内部ルータの集合を前提としたドメイン単位の制御を行うことで、転送処理と制御処理を分離するアーキテクチャを実現している。現在の IP 層では転送処理と制御処理が非分化なため、パケットはあらかじめ決められた仕様 ( 経路制御及び転送機構 ) で扱う必要があるが、MPLS ではこれらの処理を分離して扱えるので、複数の制御仕様を同時に実現できる。これは MPLS が実現する様々な特性のなかで最も着目される部分である。

#### 4. Split-Phase

MPLS 技術を用いて構成されるインターネット領域において、動的な品質制御を導入する技術として、多重フェーズ LSP 設定 ( Split-Phase ) 機構を設計した。Split-Phase は MPLS ( 第 2.5 層 ) におけるパス設定に必要な要素を分類したうえで、各要素の反映を最適化する方式である。

##### 4.1 MPLS パス構成要素

MPLS ドメイン内のフローの経路は LSP 単位で規定される。各パケット転送クラス ( FEC ) はその経路が LSP に束縛されるため、FEC の制御は実際には LSP の制御への反映を伴う。そのため、細粒度制御には LSP の制御の最適化が必要となる。

LSP を構成するパラメータは以下の 3 種類である。

( 1 ) パストポロジ : LSP の経路的な特性を示すパラメータ。制御対象となるネットワーク内のパスを構成するエンティティで必須。最悪値では制御対象領

域の全域で共有する必要がある。

( 2 ) LSP QoS 特性 : LSP の QoS 特性を示すパラメータ ( トポロジー関連の特性は含まない )。対象となる LSP に属するエンティティで必要である。

( 3 ) フロー特性 : LSP が扱わなければならないフローの特性を示すパラメータ。該当 LSP の入口 LSR で必須である。

通常 MPLS では、上記の全パラメータがそろわなければ LSP 確立を行わない。そのため LSP のパラメータと LSP 自体が不可分となり、LSP パラメータを部分更新する場合には、LSP の破棄、生成といった広域に及ぶ制御を含むため非効率である。LSP のパストポロジ及びパス全体に要求される QoS 特性の設定や反映には、必然的に LSP 全体へのシグナリング及びその反映が必要となる。一方で、LSP が扱うフローの特性の大部分は、MPLS ドメインの入口 LSR での反映処理で十分実現できる。

##### 4.2 Split-Phase : 多重フェーズ LSP 設定

LSP を構成するパラメータは、トポロジー特性関連のパラメータとそれ以外に分割できる。LSP パラメータの伝達の必要性はパス中の LSR の位置によって規定されるため、シグナリングされる LSR 集合を反映の必要性に応じて最適化できればパラメータの更新のオーバーヘッドを削減できる。また、LSP の特性パラメータを LSP 生成のタイミングと非同期に更新できれば動的 QoS 制御によって生じる制御の対象を削減可能となる。

LSP の更新には、

- ( 1 ) パストポロジの変更を伴う更新
- ( 2 ) パストポロジは保存したままパス特性の大域パラメータ ( QoS 特性等 ) を変更する更新
- ( 3 ) パストポロジ及び大域パラメータを保存したままフロー特性を変更する更新

の 3 種類が存在する。

( 1 ) は LSP の再設定と等価であり、設定コストはそれ以下にはなりえない。( 2 ) に関しては、制御対象は既存の LSP のパストポロジ上に存在する LSR に限定されるため、LSP に対する更新メッセージを用いることで最適化できる。この方式は RSVP のようなパス指向制御で利用される。( 3 ) は MPLS を前提とした場合には制御対象を LSP の入口だけに限定可能である。

Split-Phase は、LSP の制御パラメータをそれぞれ独立して扱う機構である。Split-Phase の LSP 生成処

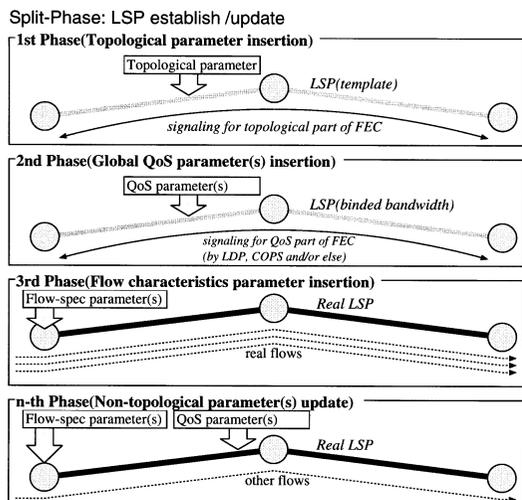


図 3 Split-Phase LSP 設定

Fig. 3 Establishment of an LSP using Split-Phase.

理は、

- トポロジー依存パラメータを用いた LSP の雛形の生成 ( 図 3 1st Phase )
- トポロジー非依存パラメータによる LSP の実体化処理 ( 図 3 2nd/3rd Phase )

に分解され、その後の運用に関しても 2 段目の制御であるトポロジー非依存パラメータを任意のタイミングで適用/更新可能である ( 図 3 n-th Phase ) .

通常の LSP 確立動作では、すべてのパラメータがそろっている必要があるが、Split-Phase ではパラメータの遅延反映ができるため、パス確立の事前処理や事後の変更処理を低コストで実現可能である .

### 4.3 Split-Phase の特性

IP 層の制御モデルでは、パケット転送はホップごとに独立して行われるため、フローごとに挙動を変更するためにはパスを構成するすべてのルータにフローに関する情報を伝播し、各ルータがその情報に基づいて動作する必要がある . そのため Int-serv モデルでは、RSVP プロトコルを用いてフローに束縛されるノードごとに資源予約を行うことでフローの確立を行う . つまり、パスの確立、帯域の予約、パスに対するフローの設定を、フローの生成時にパスを構成するルータすべてに対して行わなければならない . この操作はフローに関するパラメータを更新する際にも必要であり、フローの生成時及び更新時にはネットワークの広範囲にまたがる比較的高い処理が必要となる .

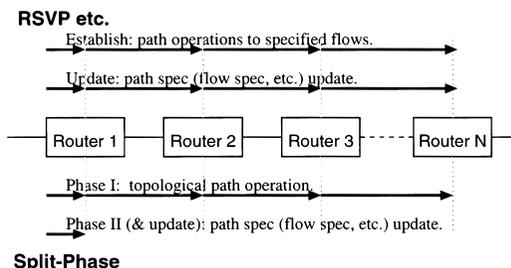


図 4 Split-Phase と他方式の比較  
Fig. 4 compare Split-Phase to another method.

方で Diff-serv モデルでは、Diff-serv ドメイン内の全ルータに対して小さな挙動セットを事前に共有しておく . この方式は処理コストを削減できる一方でフローごとに状態をもつような処理は実現できない .

Split-Phase では、まず任意のパスを確立すると同時にそのパスに対して資源の予約を行う . その後、帯域を予約したパスに対するフローを必要に応じてパスに対して設定していく . パスの確立処理はフローの生成処理から独立であるため、ネットワークの広範囲にわたって大きなインパクトを与える可能性のあるパス確立処理を時間的にフローの生成処理と分離できる . また Split-Phase では、LSP をトンネルとして用いることで、フローに対する予約した帯域の利用を、入口ルータに対する設定だけで開始できる . これも、フロー生成時のネットワークに対する影響の削減につながる . これは、パス指向のパラダイムに基づいた手法であり、MPLS のパラダイムと非常に親和性が高い .

Split-Phase は、パスの確立、パスにおける帯域の予約、パスに対するフローの追加、のそれぞれの操作を分離できるため、フローの生成処理とネットワークに対する反映処理を時間的に分離できる . 同時に、反映処理の分離によって処理の時間的な集中を回避できる . これらの特性から Split-Phase は、フローとそれに対する帯域の予約が動的かつ高頻度で変化する状況で、ネットワークに対する影響が最小になるよう処理を最適化できる .

RSVP 及び Split-Phase のパス制御コストを比較することで Split-Phase の有効性を示す . ここでのパス制御コストとは、ノードにおける計算量及び系に必要な通信総量である . 双方の方式のパス確立、帯域予約、フロー設定の比較を図 4 に示す . パスを構成するルータの総数を  $N_r$ 、各ノードのフィルタ設定のコストを  $C_f$ 、パス設定のコストを  $C_p$  とする . 各コストの最大

値を  $C_{p_{\max}}, C_{f_{\max}}$  とする . RSVP 及び Split-Phase における各パス設定コスト  $C_{RSVP}$  及び  $C_{SP}$  は ,

$$\begin{aligned} C_{RSVP} &= C(RSVP_{\text{path}}) + C(RSVP_{\text{filt}}) \\ &= \sum_{N_r} C_p + \sum_{N_r} C_f \\ &\leq N_r(C_{p_{\max}} + C_{f_{\max}}) \\ C_{SP} &= C(SP_{\text{path}}) + C(SP_{\text{filt}}) \\ &= \sum_{N_r} C_p + C_f \\ &\leq N_r C_{p_{\max}} + C_{f_{\max}} \end{aligned}$$

となる . RSVP では全ルータでパスの確立処理を行わなければならないのに対し , Split-Phase ではフィルタ設定に要するコスト  $C(SP_{\text{filt}})$  は入口ルータに対するフィルタ設定に要するコスト , すなわち単一のルータにフィルタ設定するコスト  $C_f$  のみとなる .

ここで , 単位時間当りのパス確立頻度を  $F_c$  , フローのフィルタ更新頻度を  $F_u$  とした場合のパス構成要素全体のコスト総量を比較する . それぞれの場合のコスト総量  $C_{RSVP_{\text{all}}}, C_{SP_{\text{all}}}$  は ,

$$\begin{aligned} C_{RSVP_{\text{all}}} &= F_c(F_u + 1)C(RSVP_{\text{path}}) \\ &\quad + F_c(F_u + 1)C(RSVP_{\text{filt}}) \\ &= F_c(F_u + 1) \left\{ \sum_{N_r} C_p + \sum_{N_r} C_f \right\} \\ C_{SP_{\text{all}}} &= F_c C(SP_{\text{path}}) + F_c(F_u + 1)C(SP_{\text{filt}}) \\ &= F_c \sum_{N_r} C_p + F_c(F_u + 1)C_f \\ &= F_c(F_u + 1) \left\{ \left( 1/(F_u + 1) \sum_{N_r} C_p \right) + C_f \right\} \end{aligned}$$

となる .

ルータでのパス設定コスト及びフィルタ設定コストが RSVP と Split-Phase でほぼ等価であるとすれば , 上記の式より , 同一のネットワークを対象した場合には Split-Phase は RSVP と比べて , 既存パスに対するフローに対するフィルタの特性更新の頻度が高いほど相対的に低いコストで同等の処理ができることがわかる . また , ネットワークが大規模であるほど , 単一のフローに対する処理の差が大きくなることもわかる .

## 5. Split-Phase の実装

筆者らが設計及び実装を行っている研究用 MPLS 実装 AYAME [6] ~ [8] に対して Split-Phase 機構を実装した .

LSR において Split-Phase 方式を実現する場合には以下の要素を検討しなければならない .

- 同一転送クラス ( FEC ) の扱い
- 転送機構の構成
- 制御機構の構成

### 5.1 パケット転送クラスの決定

MPLS LSR のパケットの転送処理は対象となるパケットに属する FEC の識別結果に基づいており , 転送に用いられる LSP は FEC から決定される . IP 層の動作をエミュレートする MPLS 網では , パケットの FEC はあて先ネットワークから規定される . この場合 , IP 層の経路表のエントリーと FEC が 1 対 1 対応しており , FEC の内容の変更は FEC 自体の変更を意味する . したがって , 各 FEC を単位として考える場合には FEC の同一性を維持している限り FEC の構成要素が変化することはない . そのため FEC 自体を他の要素を参照する静的な識別子として転用できる . 例えば , FEC と対応付けられている LSP の識別子として FEC 自体を用いる場合が考えられるが , これは制御処理を FEC と強く束縛するモデルとみなせる .

しかし , Split-Phase 方式のように FEC の構成要素を部分的に変えられる場合には , FEC の同一性を FEC の構成要素から示すことはできないため , このモデルは適用できない . 更に , 本来は操作的に依存関係のない LSP の操作と FEC の要素の間に制約が生じるため , LSP 操作の識別子として FEC の要素を直接的に用いることは好ましくない .

AYAME では , FEC に対しその要素とは独立した識別子 (  $FECid$  ) を割り当てており ,  $FECid$  を LSP 操作の識別子として用いる設計となっている .  $FECid$  は FEC の抽象的な指示子であり ,  $FECid$  で指示される FEC の実体への変更に影響を受けずに FEC の同一性を示すことができる . FEC 構成要素の変更は  $FECid$  を仲介して FEC の実体に対して処理 ( 要素の追加及び削除 ) することで実現している . これにより , FEC の要素に対する操作と LSP の操作を概念的に切り放すモデルでの運用が可能となる .

### 5.2 Split-Phase 型転送機構

上記で論じた方式を実現するシステムでは ,

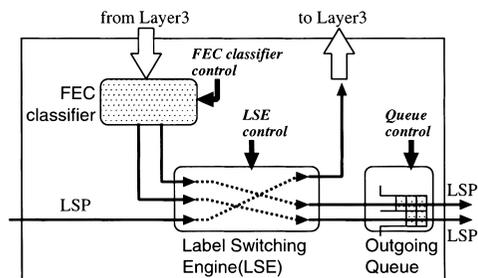


図5 AYAME LSRにおける Split-Phase 機構  
Fig. 5 Structure of the AYAME LSR for processing the Split-Phase.

- 第1段階：LSPの確立及び破棄
  - 第2段階：LSPへのFECの要素特性の適用
- のそれぞれの制御を独立して扱う必要がある。そこで、転送機構において、第1段階にかかわる“ラベルスイッチング機構(LSE)”と第2段階にかかわる“パケット区分器(FEC classifier)”を独立して制御可能な構成とした。AYAMEのLSRの構成を図5に示す。

## 6. Split-Phaseの適用

Split-Phase方式の動作検証を目的としてWIDEプロジェクトの仮設ネットワークにおいて実証実験を行った。この仮設ネットワークは、WIDEプロジェクトが年2回4日間にわたって開催している合宿形式の研究會に設置され、250人程度のネットワーク研究者が日常ネットワークとして利用した。

### 6.1 仮設ネットワーク構成

仮設ネットワークは、一般的な外部アクセス用ネットワークとして利用される一方で、WIDEプロジェクト内の様々なアクティビティを検証する場としても利用される。仮設ネットワークのSplit-Phaseに関連する部分の構成を図6に示す。

仮設ネットワークには、以下に示すように有線による地上回線及び無線による衛星回線を用いて外部接続性が提供されている。

- 地上専用線 640 kbps (128 kbps×5 を集束利用)
- 衛星回線<sup>注1)</sup> (768 kbps の双方向対称リンク)

地上回線だけではなく衛星を用いる構成になっている理由は、仮設ネットワークの特性上、物理的に有線接続が提供できない、若しくは極めて低い帯域のみしか提供できない場合を考慮しているためである。

### 6.2 実験内容

仮設ネットワークは外部と地上回線 640 kbps, 衛星

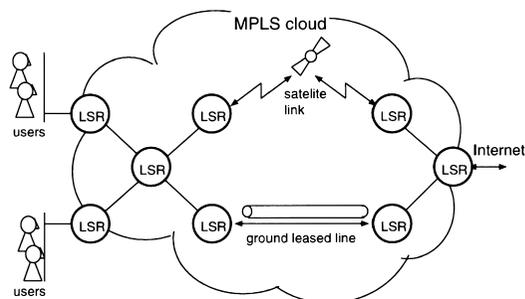


図6 仮設ネットワーク構成  
Fig. 6 Topological view of the WIDE project camp network.

回線 768 kbps の帯域で接続されているが、この程度では250名程度のユーザにサービスを提供することは困難である。また、帯域の半分以上が衛星回線経路であるが、回線特性に合わせた制御を行わなければ利便性を損なう可能性が高い。このようなネットワークにおける利用者からの要求事項を以下に示す。

#### (a) インタラクティブトラヒックの差別化

衛星回線経路のRTTは500ms程度であり、リモートシェルアプリケーションなどのインタラクティブトラヒックの使用感を著しく損なう。そのため、トラヒックの用途に応じて地上回線及び衛星回線を使い分けられる経路制御が要求される。

#### (b) フェアな帯域利用

狭い帯域幅を多数で共有するため、公正な帯域利用が要求される。

本実験ではこれらの要求に対する解として、Split-Phaseを用いた「利用者の要求に応じて動的にフロー単位に経路制御及び帯域制御を行う要求駆動型QoS制御機構」を実装し利用者に適用することで(a)を解決した。(b)の解決に関しては、帯域を含むネットワーク資源の予約システムに疑似通貨を用いた制限を導入して通貨の配布範囲内で資源を消化する手法を用いた。仮設ネットワークでは複数のサービス品質を提供できるようになっており、通常トラヒックは帯域をベストエフォート、経路を衛星経路で外部へ接続される。利用者がそれ以上の帯域幅やレイテンシ特性が必要な場合には、疑似通貨を用いて<帯域幅、経路>を要求できるようになっている。

#### 6.2.1 要求駆動型QoS制御

仮設ネットワークでは利用者はWebを利用した予

(注1): (株)JSATとWIDEの共同研究による利用。

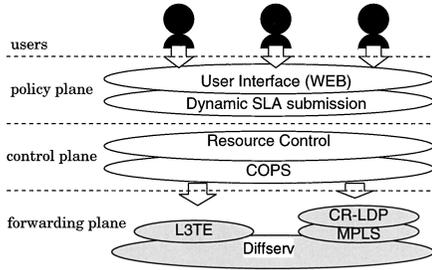


図 7 仮設ネットワークで利用した技術の関係図  
Fig. 7 Technological relations on the WIDE camp network.

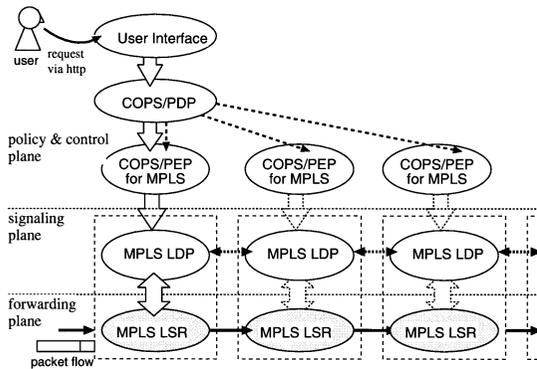


図 8 シグナリングの流れ  
Fig. 8 Signaling flow.

約システムを用いて、必要に応じて利用しているフローに対して品質パラメータを動的に指定可能である。利用者が選択可能な品質パラメータを以下に示す。

- 回線種類 ( 衛星回線, 地上回線 )
- 帯域幅 ( 4kbps 単位 )

パラメータが未指定のフローは < 衛星回線経由, ベストエフォート > として扱われる。

これらの要求をパラメータとして扱えるシステムとして図 7 に示すシステムを構築した。本システムでは利用者からの要求を、COPS プロトコルを用いたパラメータ決定層を経由して Split-Phase 方式を実装した MPLS 網に投入する。パラメータ情報の流れを図 8 に示す。利用者からの要求は COPS のパラメータ決定層において、ネットワーク中の利用可能帯域、課金情報の正当性が検証され、網に適用するパラメータが決定される。その後、Split-Phase の第 1 段階として、

(1a) CR-LDP プロトコルを用いたパス設定機構にパストロジーの設定を要求

(1b) 生成された LSP に対して要求帯域を束縛し、要求された帯域と経路を満たすパスが規定される。続いて、Split-Phase の第 2 段階として、

(2) 生成された LSP に対してフローを束縛するために境界ルータにフロー特性情報を指定する。Split-Phase では、(2) で指定するフロー特定情報の置換や追加ができるため、利用者は一度予約した帯域及び経路に対して新たに発生したフローを追加できる。

### 6.2.2 NAT 空間内部からの制御

仮設ネットワークでは NAT ( ネットワークアドレス変換器 ) によってプライベートアドレスも利用していたため、NAT を仲介した場合でも帯域/経路制御サービスを提供するための機構を設計し運用した。

一般にフローの識別には、始点及び終点の IP アドレス及び付加的な情報 ( 例えば TCP port など ) が用いられる。しかし、NAT を経由したフローではそれらの情報が変更されてしまいフローの連続性を認識することが阻害されてしまう。更に、NAT ではフローの生成に応じて動的に写像処理が行われるため、NAT の対応関係を取得できるのは、実際にフローが生成された後になることが問題を複雑にしている。

そこで本実験では、NAT の内部情報からもととのフローと NAT 後のフローの対応関係を動的に取得する機構 ( natmapd ) [9] と、Split-Phase のパス設定機構を協調させるシステムを設計した。本方式のフロー処理手順を以下に示す。

- 予約要求時には、フロー特定情報は未定状態のまま帯域及び経路のみを予約する。
- 実フローが生成され NAT による写像処理が行われてフロー特定情報が確定し次第、natmapd が COPS によるパラメータ決定層に通知し、その結果、網側にフロー特定情報が伝達される。

予約時のフロー指定は、< 始点アドレス, 終点アドレス, プロトコル > の組が用いられるが、実際には < 始点アドレス, \*,\* > といったワイルドカード指定が用いられる場合が多い。NAT を経由すると、各 TCP セッションが複数のフローに展開されるため、各フローの生成に対応してフロー識別情報を予約されているパスに対して追加しなければならない。

前章で論じたように、Split-Phase 方式ではパスに対するフロー識別情報の更新コストが低いため、NAT を仲介することで発生する数倍 ~ 10 数倍 ( 利用者のフロー数に比例する ) の処理頻度の増大に対しても、実

用的なサービスを提供可能であった。

## 7. む す び

本論文では、インターネットにおける細粒度な制御に着目し、MPLS 技術を前提とした細粒度ネットワーク制御機構 Split-Phase 方式を提案し、設計と実装に関して論じた。Split-Phase 技術は、配送経路を第 2.5 層で構築しながら制御層をドメイン単位で実現する MPLS 技術の特性を生かし、制御対象となる要素を時間的及び空間的に分割することで、細粒度の制御処理を実現する。この技術によって、フローの生成頻度と同等の速度で変更が必要なアプリケーションを実現できるようになる可能性が高い。

本論文ではそのうえで、Split-Phase 方式を用いて、利用者からの動的な要求を直接遂行するネットワークを設計し実際に稼働させた経験に関して論じた。フロー単位で QoS パラメータを要求するだけでなく、フローが利用されている途中で動的に LSP のパラメータを変更できる機構を導入することで、利用者にとって直感的な制御を提供できた。今後は、本方式を利用できるプラットフォームを広域に配置し、広域環境における有効性を検証する予定である。

## 文 献

- [1] E. Rosen, A. Viswanathan, and R. Callon, "Multi-protocol Label Switching Architecture," RFC 3031, IETF, Jan. 2001.
- [2] J. Wroclawski, "The Use of RSVP with IETF Integrated Services," RFC 2210, IETF, Sept. 1997.
- [3] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An Architecture for Differentiated Service," RFC 2475, IETF, Dec. 1998.
- [4] N. Morishima, A. Ogawa, K. Ehara, and Y. Kadobayashi, "Field-trial of dynamic sla in diffserv-capable network," First COST 263 International Workshop (QoSIS 2000), Lecture Notes in Computer Science, vol.1922, pp.117-128, Berlin, Germany, Sept. 2000.
- [5] N. Morishima, A. Ogawa, H. Esaki, O. Nakamura, S. Yamaguchi, and J. Murai, "Preliminary field-trial for qos routing and dynamic sla," IEICE Special Issue on Internet Technology (SIIT 2001), vol.E84-B, no.8, pp.2039-2047, Jan. 2001.
- [6] Y. Uo, S. Uda, N. Ogashiwa, S. Ohta, and Y. Shinoda, "Ayame: A design and implementation of the cos capable mpls layer for bsd network stack," INET 2000, no.1h-2, Yokohama, Japan, July 2000.
- [7] 宇多 仁, 宇夫陽次郎, 篠田陽一, "Mpls 実装 ayame におけるバケット転送機構の設計及び実装," 情処学シンポジウムシリーズ, vol.2000, no.15, pp.127-132, Dec. 2000.

[8] "AYAME Project," <http://www.ayame.org/>.

[9] 横山輝明, 森島直人, 小川晃通, 宇夫陽次郎, 宇多 仁, 江崎 浩, 山口 英, "Nat の内部情報通知による end-to-end 性の確保," 情処研報, vol.2001, no.59, pp.95-100, June 2001.

(平成 13 年 11 月 30 日受付, 14 年 3 月 18 日再受付)



宇夫陽次郎 (正員)

平 7 東工大・工卒。平 9 北陸先端科技大学院大情報科学研究科博士前期課程了。平 14 同博士後期課程単位取得退学。現在、(株)インターネットイニシアティブ技術研究所。副 IP 層をベースとしたインターネットアーキテクチャの構造に関する研究及びあやめプロジェクトにおける MPLS ルータ開発に従事。



宇多 仁

平 9 東京理科大・理工卒。平 11 北陸先端科技大学院大情報科学研究科博士前期課程了。現在、同博士後期課程在学中。高性能高速ルータの研究及びあやめプロジェクトにおける MPLS ルータ開発に従事。



小柏 伸夫

平 11 芝浦工大・システム工卒。平 13 北陸先端科技大学院大情報科学研究科博士前期課程了。現在、同博士後期課程在学中。統合シグナリング層及びマルチキャストシステムの研究、あやめプロジェクトにおける MPLS ルータ開発に従事。



篠田 陽一 (正員)

昭 58 東工大・工卒。昭 60 同大学院工学部博士前期課程了。昭 63 同大工学部助手。平 3 北陸先端科技大学院大情報科学研究科助教授。現在、同大学情報科学センター教授。工博。