

## ■3群 (コンピュータネットワーク) - 5編 (通信品質)

# 4章 品質制御の実際

(執筆著: 長谷川亨) [2011年1月 受領]

### ■概要■

ルータなどのネットワーク機器はトランスポート層以上の品質を制御できないため、実際のISP (Internet Service Provider) ネットワークなどでは、3章で述べた品質制御技術を単独で、あるいは組み合わせて使用して、ネットワーク層以下で品質制御することが一般的である。また品質制御技術は、ISP や通信事業者にとって、他社との差別化に有望であり、積極的に導入されている。

近年の普及が目覚ましい、インターネットアクセス、音声、映像を同時に提供するトリプルプレイサービスでは、音声、映像とインターネットアクセスの品質差別化するため、DiffServを採用してISP ネットワークを構築することが多い。DiffServ ではルータにおけるパケットのキューイング制御と、音声、映像といったアプリケーションを単位とした粒度の荒い帯域割当てを組み合わせ、ネットワーク層でアプリケーション単位の差別化を実現している。RSVP のような複雑なフローごとのシグナリングを用いない簡便さが普及の原因の一つである。

一方、ISP ネットワークのバックボーンや、企業向けのVPN (Virtual Private Network) で採用されているMPLSは、ネットワーク層とデータリンク層の中間に位置しており、ネットワーク層全体の品質向上に寄与する。一般にOSPF (Open Shortest Path First) などの経路制御プロトコルを用いると、全パケットは最短長の経路を通過するため、最短経路へのトラヒック集中による輻輳を引き起こすだけでなく、ネットワーク全体の利用効率を低下させる。これに対して、トラヒックエンジニアリング技術を用いてトラヒック分散だけでなく、Fast Rerouteを用いたプロテクションを可能としている。現在は、単一のISP ネットワークだけでなく、インターネット環境での使用に向けた改良が進められている。

更にNGN (Next Generation Network) では、通信フロー単位での品質差別化を実現するため、SIP (Session Initiation Protocol) などで要求されるトランスポート層のシグナリング情報を利用している。具体的には、RACF (Resource and Admission Control Functions) を用いてネットワーク層の資源を管理し、SIP を用いて行われる通信フロー要求の受付制御を行う。例えば、十分な資源がない場合は、新たなSIPによる音声通話の要求は拒否される。一方、ネットワーク層でのパケットの制御はDiffServと同様なキューイング制御が用いられている。

### 【本章の構成】

品質制御技術はインターネットで広く採用されるまでには至っていないが、イントラネットでの採用が目覚ましい。4章では、3章の品質制御技術の応用例として、DiffServを用いたISP (Internet Service Provider) ネットワークでの複数クラスのQoS提供 (4-2節)、MPLSを用いたISP ネットワークのトラヒックエンジニアリング (4-3節)、NGN (Next Generation Network) への品質制御技術の適用 (4-4節) について述べる。

## ■3群 - 5編 - 4章

### 4-1 Diffserv (DS) の実際

(執筆著：村瀬 勉) [2009年5月 受領]

#### 【概要】

Differentiated Service (DS) は、IETF で規定されたものであるが、実装や運用に関しては、規定されていないところも多く、いろいろなバリエーションがある。また、規定がないため、実運用上で問題となる場合やそもそもの DS の本質的な問題が障害となるような場合もあり、それらを紹介する。以下では、DS の概要を説明してから、DS を使うに当たっての優先クラス割当て単位、DS の実現に当たっての実装方法、ほかの QoS 制御との関係に関する議論、アドミッションコントロール、運用上の課題を順に紹介する。

#### 4-1-1 DiffServ の概要

RFC273 で規定されているルータで構成される IP 網では、パケットの遅延、廃棄といった QoS に関する制御が全く行われないため、QoS を要求するサービス、特にマルチメディア通信において満足な品質が得られないことが、問題となっている。DiffServ は QoS に関して定量的な保証はしないが、QoS に差を付けることで、品質要求を満たそうという目的で開発された。先に開発された IntServ に比較して簡単な制御で実現できる機能を定義しているため IP 網での現実的な QoS 制御の解として多くの関心を集めた。

DS ではエッジルータとコアルータに機能が振り分けられている。IPv4 ヘッダの TOS ビット (DS ビット/バイトとも呼ぶ) に示された codepoint (DSCP) に応じたキューイングとフォワーディング処理(これらを Per Hop Behavior: PHB と呼ぶ)を行うのがコアルータである。またエッジルータは、パケットバイパケットに DSCP の作成、ポリシング、シェイピング、パケット廃棄などを行い、かつ PHB を実行する。PHB (図 4・1 参照) には、EF, AF, BE の 3 カテゴリーがあり、更に、AF は、12 に細分化されているため、14 種類の PHB が存在する。

AF では、四つのクラスと三つの廃棄レベル (Drop Precedence) が想定されている。四つのクラスは、リソースの割当てのレベルを意味する。廃棄レベルはクラスの中でそれぞれ意味を持ち、廃棄確率のレベルを異ならせる behavior を規定する。一方、遅延については、規定しない。EF は、常に入力トラヒックの合計レートよりも出力トラヒックレートを高くしておく behavior を要求する。EF に属するパケットのフォワーディングレートは、ほかのどんなトラヒックの影響も受けない。また、前記のようにレートを保証することで、キューイング遅延をほとんど受けないので、遅延変動も最少におさえることができる。キューイングがほとんどないことを前提とする behavior であるので、廃棄レベルは規定されない。

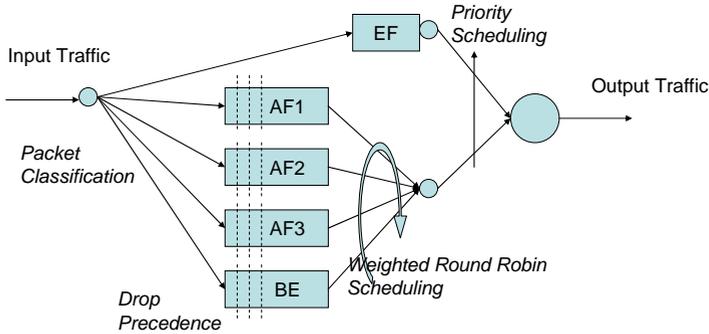


図 4・1 DiffServ の PHB を実現するキューイングの例

#### 4-1-2 優先クラスの割当て単位

##### (1) どの単位で優先制御をかけるべきか

DS においては、ルータでは DSCP によりパケットの PHB 制御が行われるため、この DSCP をどの単位で設定するかで、各種のサービスが実現できる。DSCP 設定単位としては、様々な単位があるが、典型的な単位を次に挙げる。それらは、ネットワークレイヤ、トランスポートレイヤ、アプリケーションレイヤというレイヤごとの単位と、ユーザプレーン、制御プレーン、管理プレーンというプレーン単位の 2 種類である。以下ではレイヤ単位のマッピングのみを紹介する。

##### (2) IP アドレス

IP アドレス単位を用いることで、IP アドレスごとにパケットの転送を制御できる。例えば、192.168.1.60 というソースアドレスとするパケットを EF クラスにマッピングすること（EF クラスの DSCP を割り当てること）で、192.168.1.60 という IP アドレスを持つ端末からの送信データを高優先にすることが可能である。この制御単位では、例えば、個人別に優先度を設定したり、サーバやクライアントといった機能ごとの端末単位で優先制御を行いたいときが有用である。また、もちろん、192.168.1.\*、という集合アドレスでのマッピングも可能であり、組織別あるいは AS（自律システム）別の優先度付けにも使うことができる。NAT（Network Address Translation）などで、複数の端末が、この 192.168.1.60 というアドレスを使っている場合には、複数の端末の送信データを高優先にすることになる。

##### (3) TCP/UDP フロー

TCP/UDP フロー単位を用いることで、フローごとにパケットの転送を制御できる。すなわち、IP アドレスという端末ごとよりも更に細かい粒度での制御が可能となる。例えば、TCP ポートの 80 番をソースアドレスとするパケットを EF クラスにマッピングすることで、80 番の送信ポートを持つ通信、web のコンテンツ送信、を高優先にすることができる。ただし、実際に用いられる wellknown ポート番号とアプリケーションの対応は必ずしも IETF で規定されている組合せとは限らない。例えば、web の wellknown ポートとされている 80 番や 443 番を、ファイアウォールを超えるためのポートとして多様な用途で用いたりすることがある。

この TCP/UDP ポート番号を IP アドレスと組み合わせることで、特定の端末の特定のフローだけに限定して、優先度を与えることができる。例えば、データバックアップのトラヒックには、最低の優先度を与えるために、あるサーバからあるサーバへの ftp 通信を、つまり 192.168.1.60 が送信元 IP で、なおかつポート 21 番のトラヒックを最低優先度に割り当てるといった使い方である。

#### (4) アプリケーションレイヤ

アプリケーション単位を用いることで、アプリケーションごとあるいはメディアごとの制御が可能である。メディアには、遅延の品質要求の厳しいモノや、スループットの要求品質の厳しいモノなど、メディアにより各種の QoS 要求がある。そのため、メディアごとにマッピングクラスを割り当てることで、それぞれの QoS 要求をうまく満たせる場合がある。例えば、RTP/UDP のトラヒックには、遅延の小さいクラス、例えば EF クラスを割り当てるといった割り当てである。

なお、RFC 4594 (Guidelines for DiffServ Service Classes : August 2006) には、アプリケーションと DSCP の割当ての例が述べられている。

#### (5) 実際の通信内容とプロトコルヘッダとのミスマッチの解決法

DS では、ユーザではなくネットワークのエッジルータがトラヒックのヘッダ情報などを参考にして、DSCP を割り当てる (つまりクラスを決定する)。近年、プロトコルの偽装やカプセル化が盛んになっており、通常のヘッダを見るだけでは、実際に用いられているアプリケーションを判断できないことがある。例えば、RTP/UDP のストリーミングが HTTP/TCP でカプセル化されているような場合である。ファイアウォールを超えるためによく使われるカプセル化である。このとき、通常のヘッダ位置には、HTTP/TCP があるため、遅延要求の緩い web アプリケーションと判断してしまうかもしれないが、更にそのヘッダの奥に RTP/UDP ヘッダがあり、ストリーミングの動画データが入っている場合もある。また、VoIP が使用することになっているポート番号を高優先にした場合、そのポート番号が不正にデータ通信に使用される場合がある。このような偽装を防ぐために、各種の方式が提案されているが、完全に偽装を防御するのはかなり困難であると思われる。

### 4-1-3 DS 実現 (実装) 方法

#### (1) 実現に関する規定と実装

DS には、優先指定の規定はある、つまりルータに入力されたパケットがどう制御するか (PHB) は規定されているが、ルータの実機にてどう扱われるかは、実装依存である。規定を実現するような実装方法は、いくつかあると思われるが、直感的には、スケジューリングとキューイングを用いれば、実現できる。例えば、図 4・1 に示したような完全優先 priority キューと Weighted-Round-Robin (WRR) キューを組み合わせられた方法が考えられる。EF クラスは、完全優先 priority キューとし、ほかのクラスは、非優先キューとする。更に非優先キューを四つの細分キューに分ける。細分キューは、WRR キューとし、運用で決められた weight (重み) でスケジューリングする。更に、各 AF の三つの廃棄クラスの実現には、各キューの入り口において待ち行列長での閾値によるパケット廃棄を行う。

#### (2) クラス間の優先度とクラス内の廃棄優先度の実現

DS では、14 のクラスが定義されているが、DS として明確に品質に言及している EF と BE

クラスに比べて、AF クラスはどのような品質として使うかは規定されておらず、ユーザの運用に任されている。四つの AF クラスは、パケットの扱いにおいて確率的に相対的な優先を規定している。運用においては、各 AF クラスに対して割当てトラヒックに応じた帯域やバッファ量といったリソースを確保する。その確保リソース量とそこに割り当てるトラヒック量に応じて品質が決まる。したがって、網に收容すべきトラヒック量が決まっているときには、リソース量で品質が決まり、逆に、目的とする品質とリソース量が決まっているときには、收容可能トラヒック量が決まる。收容可能トラヒック量を管理するためには、後述のアドミッション制御のような品質定義単位のトラヒックごと（例えばフローごと）の入力制限が必要になる。

### (3) レイヤ2の優先キューイング制御との関連

DS は、IP レイヤでの QoS 制御機能であるが、そのほかのレイヤにも、QoS 制御機能が備わっている。エンドエンドの QoS のために、DS を用いるにあたっては、ほかの QoS 制御との連携も考えるべきであろう。複数の DS ルータ間がイーサスイッチで接続されている網は、当然あり得る。イーサスイッチのようなレイヤ2のスイッチにおいては、DS の制御は期待できない。したがって、イーサスイッチがボトルネックになるような網では、DS での QoS 制御の効果はあまり現れない。しかしながら、DS の優先制御情報をほかのレイヤの QoS 制御にマッピングすることで、ある程度は、DS の QoS 効果を維持できる可能性はある。標準的な QoS 制御として、イーサネットでは、IEEE802.1p が、ATM・MPLS では、コネクションベース・ラベルベースの QoS-TE がある。ただし、802.1p では、クラスは7クラスしか定義されておらず、DS の14クラスを802.1pのどのクラスにマッピングするかを決めることが必要になる。

また、802.1p では、どの DS のエッジルータのようなインGRESSルータがコアルータと役割を分担するような機能定義がされておらず、802.1p の優先値の設定・変更がどのルータ・イーサスイッチでも可能であるところに、運用上の課題がある。ATM・MPLS では、基本的には、フローベースの QoS 制御になるため、DS クラスを ATM・MPLS の優先度フローに割り当てることが可能である。

## 4-1-4 アドミッション制御

BE クラスを除くどのクラスにおいても、品質設計を行うためには、アドミッション制御が必要である。アドミッション制御により入力トラヒックを適正に管理することで、許容量の遅延といった QoS が達成できる。アドミッション制御とは、通信開始前に、その通信を許可していいかどうかを判断する制御になる。その通信を許可することで、その通信及び既存の通信の QoS が許容値以下になる可能性がある場合には、通信を許可しないという制御である。この判断においては、通信要求者は、通信するトラヒック特性、通信ペア及び所望の QoS 値の申告が必要である。網側は、通信ペアを結ぶ経路におけるリソース使用（あるいは確保）状況とそこに該トラヒックを追加した場合に、観測され得る QoS 値が許容値を満足するかを推測し、判断を行う。

アドミッション制御を行う機能として、BB (Bandwidth Broker) という名称の機能が知られている。BB は、リソース管理を行う機能であり、リソースの使用状況を考慮して、新しい DiffServ トラヒックを網に入れても良いかどうか判断する。この判断には、リソースの有

無、クラスへの割当ての可否といった制御ポリシーが考慮される。これは、各ノードにおけるリソースの有無のみならず、エンドエンドでの、つまりパス上のすべてのルータにおけるリソース調達を考慮して、判断される。また、パスがドメイン間をまたがるような場合にも、各ドメインの BB が、隣どうしで了解を取り合って、行われる。クラスへの割当ての可否においては、あらかじめ決められたポリシーに基づいて、そもそも通信する資格があるか、通信トラヒック量が基準値以内か、などを考慮して判断される。

またアドミッション制御の導入には、規定量以上のトラヒックの流入を防止するためのトラヒック流量制御が必須である。このトラヒック流量制御(あるいは traffic Enforcer)として、トラヒックシェイパーなどが用いられる。トラヒックシェイパーでは、規定量以上の流入を防ぐために、トラヒックの平滑化バッファが用いられるが、このバッファも有限であるため、バッファ量以上のトラヒックに対しては、「規定量以上のトラヒックである」ということをマーキングして、網内での廃棄制御 (Drop precedence) に任せることになる。

一般的には、上記トラヒック特性の申告及び QoS 値の推測は、困難である。将来的に発生するトラヒックを事前に予測するのは困難であるため、ピークレートや平均レートやバースト持続時間といったトラヒック特性の申告は難しい。トラヒック特性の申告値が大きすぎると、網リソースが無駄になるし、申告値が小さすぎると、シェイパーにより、パケットが遅延させられたり、廃棄させられたり、マーキングされて、廃棄確率が高くなったりと、不利なことが起こる。

#### 4-1-5 Ds の運用上の課題

DiffServ は、QoS 制御を目的に導入された。しかしながら、実際には DS で規定しているのは、ノードでのパケットの扱いのみであり、陽に QoS の制御を規定しているわけではない。このため、制御の結果として、ある程度の QoS は期待できるが、所望の QoS を得られる保証はない。例えば、パケット遅延に関しては、EF クラスでは、最優先でパケットが転送されるものの、経由するルータの段数や、ルートの変更や、ルータのインタフェース速度などで、決まるため、あらかじめ QoS 値を予測することは困難である。また、通信中に変化することもある。AF クラスも同様に最低帯域は保証するような制御が施されるが、遅延や廃棄率といった QoS 値を予測することは非常に困難である。結果として、帯域以外の QoS を保証あるいは予測することはかなり困難である。

課金は、DiffServ を含めて、公衆網においては、優先制御一般の課題である。公衆網では、通信リソース (網) とトラヒック発生者 (ユーザ) が、利害関係にあるためである。網は、所定のリソース内でできるだけ多くのトラヒックを多重して収益を増やそうと試み、ユーザは、同じ料金でできるだけ混雑が少ない状況で通信したいと望むからである。網は、QoS 制御のために、ユーザの要求に応じたリソースを用意する必要がある。ユーザが、多くの帯域を要求し、多くの時間を通信に費やせば、それだけリソースが多く必要になる。したがって、課金は、QoS のクラス、トラヒック量、通信時間といったパラメータで決まるリソース量に応じて行う必要がある。このパラメータから課金を決める方法は、特に定められた方法があるわけではなく、また網の運用単位でも異なるのが普通である。そのため、特に運用管理の異なる複数のドメインをまたがる通信では、課金が難しくなる。更に、通常のインターネットの利用料金は、電話のような距離性ではなく、通信相手までのルートや経由ルータ数には

依存せず料金一定であることが普通である（もちろん QoS はなんら考慮されない）。ところが、上述のような課金を導入すると、通信先が、同じ ISP であるのか、複数の ISP を経由する必要があるのかによって、異なるのが自然である。さらに、BGP などの ISP 間のルーティングがダイナミックに変更される場合に、その変動を ISP 間でどのように吸収するのかといった課題も残されている。

キャリア間連携の課題には、ルーティング制御、課金、受付制御ポリシーといった制御・運用の統一性の課題がある。ルーティング制御では、BGP を用いている ISP 間のルーティングと DS の親和性が課題となる。課金については、前述したとおりであるため、ここでは省略する。受付制御ポリシーとしては、異なるドメイン間で異なるポリシーが設定されている場合に、受付制御における受付可否判断が困難になる。

#### ■参考文献

- 1) ietf RFC273, [www.ietf.org/rfc/rfc273.txt](http://www.ietf.org/rfc/rfc273.txt)
- 2) DiffServe全般の情報, [www.ietf.org/WG](http://www.ietf.org/WG)

## ■3群 - 5編 - 4章

### 4-2 MPLS トラヒックエンジニアリング

(執筆者：熊木健二) [2009年3月 受領]

近年、個人向けインターネットサービス及び法人向け VPN サービスにおいて、インターネットアクセス、音声、映像を同時に提供するトリプルプレイサービスが普及している。そのサービスの普及に伴い、ISP (Internet Service Provider) はバックボーンネットワークの高品質・高信頼性を求められるようになってきた。そのため、ISP はネットワークの高品質・高信頼性を確保する技術の一つとして MPLS (Multi Protocol Label Switching)<sup>1)</sup> 技術を導入してきた。その結果、ISP は MPLS 技術を用いて①帯域確保、②帯域の有効利用、③高速障害迂回を容易に行うことが可能となった。

そこで、最初に ISP における MPLS トラヒックエンジニアリングの現状を説明し、次に単一 ISP ネットワークの大規模化に伴うマルチエリアネットワークに対する MPLS トラヒックエンジニアリングの改良、および ISP ネットワーク間での MPLS トラヒックエンジニアリングの改良を説明し、最後にパス計算サーバ (PCE (Path Computation Element))<sup>2)</sup> を用いたトラヒックエンジニアリングを説明する。

#### 4-2-1 MPLS トラヒックエンジニアリングの概要

MPLS では、ルータがデータを転送するため IP パケットにラベル<sup>3)</sup>を付加し、そのラベルに従いデータが宛先に転送される。そのラベルの配送手法として LDP (Label Distribution Protocol)<sup>4)</sup> 及び RSVP-TE (Resource Reservation Protocol - Traffic Engineering)<sup>5)</sup> が使用されている。具体的なラベル配送手法として、明示的な経路制御によるラベル配送及び動的な経路制御によるラベル配送がある。前者は、運用者がルータに対して明示的に経路を設定し、RSVP-TE によるシグナリングによりラベルが配送される。また、後者は、ルータが IGP 及び IGP の拡張<sup>6,7)</sup> による経路情報を用いて経路決定し、LDP 及び RSVP-TE によるシグナリングによりラベルが配送される。このように、ISP はネットワーク内において明示的な経路制御および動的な経路制御により MPLS を用いたトラヒックエンジニアリングを可能とする。

また、ISP はネットワークの高品質・高信頼性を確保するために MPLS 技術を導入し、冒頭にて述べた①～③を実現している。左記①～③を満たすラベル配送手法として RSVP-TE が最適であるため、以下 RSVP-TE を用いたトラヒックエンジニアリングとして冒頭に述べた①～③を中心に説明を行う。

#### 4-2-2 ISP における MPLS イントラエリアトラヒックエンジニアリング

ISP では、ネットワークの高品質・高信頼性を確保するため MPLS 技術を導入してきた。その結果、①帯域確保、②帯域の有効利用、③高速障害迂回を容易に行うことが可能となった。ここでは、単一エリアにおける RSVP-TE を用いたトラヒックエンジニアリングとして上記①～③に関する説明を行う。

##### (1) 帯域確保

実際に帯域確保する手法として、ここでは DS-TE (Diffserv-aware Traffic Engineering)<sup>8)</sup> 技術に関する説明を行う。具体的には、表 4・1 に示すとおり DS-TE では Preemption と TE-Class

の関係を定義する。また、コントロールプレーンに関して RDM (Russian Doll Model) モデル<sup>9)</sup>を使用する例を図 4・2 に示す。更に、データプレーンに関するキューイングの例を表 4・2 に示す。その結果、EF (Expedited Forwarding) クラスのトラヒックである音声と映像がネットワーク上で保有する帯域 (MPLS パスで予約される帯域) と連携しながらデータプレーン上で帯域確保を行うことができる。

表 4・1 Preemption と TE-Class の関係

PreemptionとTE-Class	トラヒック
TE-Class[0] <-> <CT1 , preemption 0>	音声 (EXP=5)
TE-Class[1] <-> <CT1 , preemption 1>	映像 (EXP=4)
TE-Class[2] <-> <CT0 , preemption 7>	その他 (EXP=0,1,2,3)

表 4・2 キューイング (データプレーン) の例

トラヒック	キューイングの種類
音声, 映像 (EF)	PQ (200M確保)
その他 (AF+BE)	WRR+RED

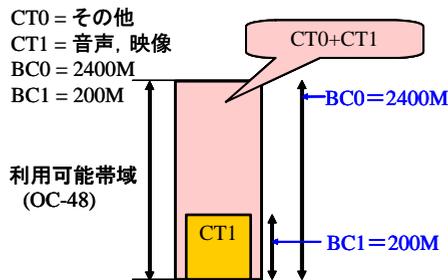


図 4・2 RSVP シグナリングによる帯域確保の例

## (2) 帯域有効利用

帯域を有効する手法として、単一 ISP におけるシングルエリアのネットワークでは、始点ルータ (Head-end ルータ) は終点ルータ (Tail-end ルータ) に対して動的な経路制御を行うため OSPF (Open Shortest Path First), IS-IS (Intermediate System to Intermediate System) の拡張として OSPF Opaque-area LSA (Link State Advertisement)<sup>6)</sup>, The extended IS reachability TLV (Type Length Value)<sup>7)</sup> (リンクに対する TE メトリック, 最大帯域, 最大予約帯域, 未使用帯域情報等) を利用し, ダイクストラアルゴリズムなどを使用して図 4・3 に示すとおり, ルータ R1 が動的な MPLS パス計算を行っている。この場合, 物理インタフェースはすべて 10G [bps] とし, 残り帯域から 5G [bps] を満たしかつ最短経路 (10G [bps] インタフェースのコストを 1 とする) (ERO (Explicit Route Object)<sup>4)</sup>: R1-R3-R5-R6) を選択する。また, 同様に (1) において述べた DS-TE 技術を使用して, 始点ルータが動的な MPLS パス計算を行っている。しかし, 動的に帯域を有効利用する場合, 始点ルータが OSPF や IS-IS を基盤にダイクストラアルゴリズムなどを使用して MPLS パス計算しているため, 基本的には一つの制約 (コストなど) に従った最短経路を求めている。その結果, MPLS パス計算に限界があるた

め、MPLS では運用者が明示的に MPLS パス経路を指定して確立することが可能である。

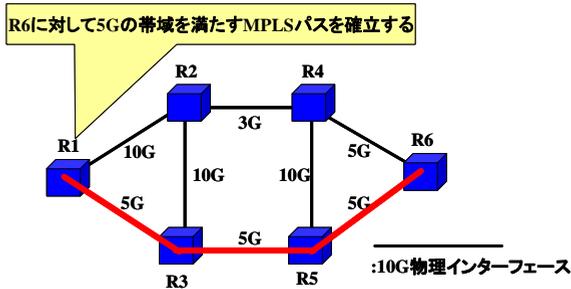


図 4・3 帯域制約に従った最短経路

### (3) 高速障害迂回

最後に障害の迂回を実現する手法として、MPLS ネットワークにはローカルリカバリとグローバルリカバリがある<sup>10)</sup>。前者は、障害部分だけを高速に迂回するために、局所的にバックアップ LSP を確立し、障害時にそのバックアップ LSP に迂回する。後者は、障害を高速に迂回するために、始点ルータと終点ルータ間にバックアップ LSP を確立し、障害時にそのバックアップ LSP に迂回する。それぞれ図 4・4、図 4・5 に示す。図 4・4 はローカルリカバリの一実施形態である 1 対 N で高速迂回を実現する Link Protection LSP、Node Protection LSP<sup>11)</sup> を示しており、PLR は Point of Local Repair<sup>11)</sup>、MP は Merge Point<sup>11)</sup>、Link Protection LSP はルータ R1-R2 のリンク障害用バックアップパス、Node Protection LSP はルータ R2 のノード障害用バックアップパスを示す。図 4・5 はグローバルリカバリの一実施形態である 1 対 1 で高速迂回を実現する Secondary/Backup Path を示している。

FRR (Fast ReRoute)<sup>11)</sup> はローカルリカバリの一実施形態であり、1 対 1 で高速迂回を実現する Detour LSP と 1 対 N で高速迂回を実現する Link Protection LSP、Node Protection LSP に分類される。ここで、実装の多い後者の説明を行う。Link Protection は、ルータ間のリンク障害 (例えば、ファイバ障害、ルータインターフェース障害など) を高速に迂回する技術である。ルータが高速にその障害を検知し、あらかじめ設定された Link Protection LSP にデータ転送を行い、そのリンク障害を高速迂回する。

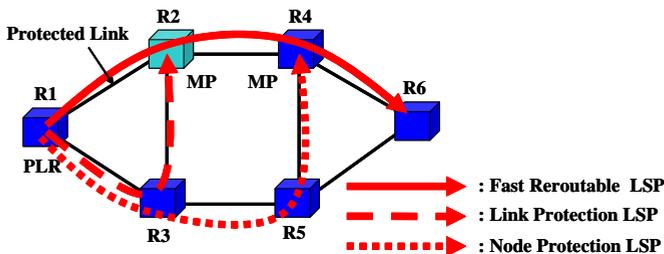


図 4・4 ローカルリカバリ

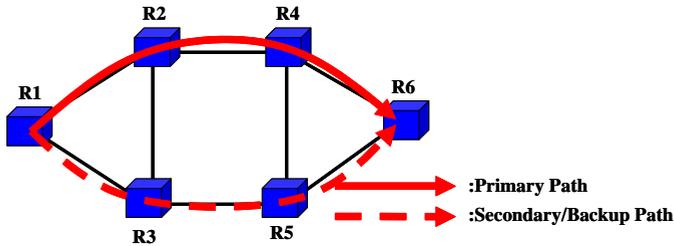


図 4・5 グローバルリカバリ

### 4-2-3 ISPにおけるMPLSインターエリトラフィックエンジニアリング

ISPはMPLS技術を基盤としたネットワーク上でインターネット、L2VPN、L3VPNなどの様々なサービスを展開してきた。そのサービス展開に伴い、ISPはネットワークの拡張を行う必要が出てきたため、OSPFなどのIGPのスケラビリティを考慮しIGPドメインを複数のエリアに分割（マルチエリア化）してきた。しかし、MPLSネットワークにおいては、マルチエリア化することにより、従来MPLSパスを計算するために用いられる情報伝達が制限される。具体的には、始点ルータがエリアを跨いで終点ルータに対してMPLSパスを計算する場合、OSPF Opaque-area LSAがABR（Area Border Router）を越えてほかのエリアに到達しないため、始点ルータはエンド・エンドで帯域やTEメトリックなどの制約に従った最適なMPLSパスを計算できない。これらを解決するために、IETFのTEWGにおいてインターエリトラフィックエンジニアリングに関する要求仕様<sup>12)</sup>の標準化を行ってきた。要求仕様に対し、図4・6に示すように、インターエリアの環境においてエリアの境界ルータがエリアごとに最適なMPLSパス計算及び確立する手法として文献13)が提案された。更に、ネットワークトポロジの変化に伴う再最適化を行う手法として文献14)が提案された。しかし、エンド・エンドでMPLSパスを計算する場合、エリアごとには最適なMPLSパスを計算できるが、エリアを跨いで最適なMPLSパスを計算できないことがある。これを解決する手法は、4-2-5節で説明する。

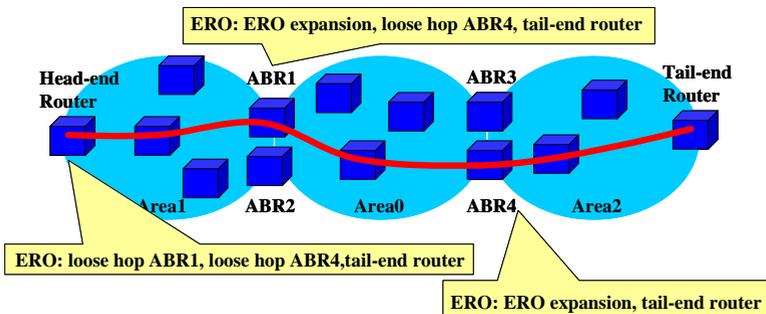


図 4・6 インターエリア環境におけるMPLSパス確立手法

また、高速障害迂回に関して、インターエリアの環境においては ABR のノードプロテクションが課題となる。文献 11)にて定義した手法では ABR を越えて FRR を動作させることができないことから FRR の拡張を行うために文献 15)が提案されている。図 4・7 に示すように、コントロールプレーンでは、PLR は MP から RSVP により RRO (Record Route Object) node-id subobject を用いてプッシュするラベルを受信し、障害時、データプレーンでは、PLR は MP から受信したラベルをプッシュしてパケット転送をする。現在、インターエリア環境の MPLS ネットワークにおいて上記の提案手法を用いて高速障害迂回をすることが可能である。

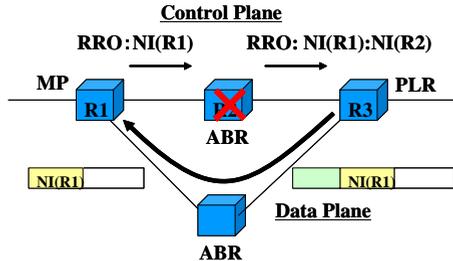


図 4・7 ABR ノードプロテクション

#### 4-2-4 ISP における MPLS インターAS トラフィックエンジニアリング

4-2-3 節と同様に、ISP は L2VPN や L3VPN のサービス展開拡張に伴い、他 ISP との MPLS ネットワークによる相互接続を行ってきた。相互接続にあたり、AS (Autonomous System) 間において MPLS パスを計算するために用いられる情報伝達が 4-2-3 節と同様に制限される。また、AS 間においてこれらの情報をすべて広報することによる IGP のスケールラビリティに関しても懸念された。これらを解決するために、IETF の TE WG においてインターAS トラフィックエンジニアリングに関する要求仕様<sup>16)</sup>の標準化を行ってきた。要求仕様に対し、図 4・8 に示すように、インターAS の環境において AS の境界ルータが AS ごとに最適な MPLS パス計算及び確立する手法として文献 13) が提案された。更に、ネットワークトポロジの変化に伴う再最適化を行う手法として文献 14) が提案された。しかし、エンド・エンドで MPLS パスを計算する場合、AS ごとには最適な MPLS パスを計算できるが、AS を跨いで最適な MPLS パスを計算できないことがある。これを解決する手法は、4-2-5 節で説明する。



図 4・8 インターAS 環境における MPLS パス確立手法

また、高速障害迂回に関して、文献<sup>11)</sup>にて定義した手法ではASBRを越えてFRRを動作させることができないことからFRRの拡張を行うために文献<sup>15)</sup>が提案されている。図4・9に示すように、コントロールプレーンでは、PLRはMPからRSVPによりRRO node-id subobjectを用いてプッシュするラベルを受信し、障害時、データプレーンでは、PLRはMPから受信したラベルをプッシュしてパケット転送をする。現在、インターAS環境のMPLSネットワークにおいて上記の提案手法を用いて高速障害迂回をすることが可能である。

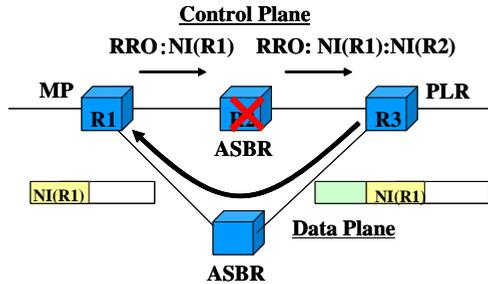


図4・9 ASBR リンク/ノードプロテクション

#### 4-2-5 パス計算サーバ (PCE) を用いたトラフィックエンジニアリング

4-2-3 節、4-2-4 節において述べたように、ある制約条件下において MPLS パスを計算する場合、インターエリア/AS の環境では、MPLS パスを計算するために用いられる経路情報伝達に関して制限がある。前者に関しては、OSPF Opaque-area LSA が ABR を越えてほかのエリアに到達しないため、帯域や TE メトリックなどの情報伝達ができない。また、後者に関しては、BGP (Border Gateway Protocol) での経路集約により、帯域や TE メトリックなどの情報伝達ができない。それらの情報を補うために、PCE は各エリア/AS における帯域や TE メトリックを含んだ経路情報を保持し、MPLS パス計算を行う場合、その経路情報を保持している PCE に MPLS パス計算するためのパス計算要求を行う。その結果、エンド・エンドで最適な MPLS パスを計算することが可能である。

表4・3 PCE アーキテクチャモデル

計算モデル	特徴
集中型計算モデル	<ul style="list-style-type: none"> <li>全てのエリア/ASの経路情報を1つのPCE(冗長化を含む)で保持する</li> <li>全てのPCReqメッセージは1つのPCEに送信される</li> </ul>
分散型計算モデル	<ul style="list-style-type: none"> <li>各エリア/ASの経路情報を複数のPCEで保持する</li> <li>PCReqメッセージは分散された1つのPCEもしくは複数のPCEに送信される</li> </ul>

表4・4 PCE 計算モデル

PCEモデル	特徴
Composite PCE	<ul style="list-style-type: none"> <li>ルータ内部にPCEが存在</li> <li>ルータ間で連携</li> </ul>
External PCE	<ul style="list-style-type: none"> <li>ルータ外部にPCEが存在</li> <li>ルータとPCEおよびPCE間で連携</li> </ul>

ここでは、PCE 技術を用いて具体的にインターエリア/AS 環境で適用する手法に関して説明する。適用例に関しては、文献 17) にて提案されている。PCE 技術をインターエリア/AS 環境に適用するにあたり、インターエリア PCE に関する要求仕様<sup>18)</sup>及びインターAS PCE に関する要求仕様<sup>19)</sup>が提案された。これらの要求に基づき、PCE を用いたインターエリア/AS 環境におけるパス計算手法の実施形態として、BRPC (Backward Recursive PCE-based Computation) 手法<sup>20)</sup>が提案された。図 4・10 に示すように AS1 の始点ルータが AS2 の終点ルータに対して MPLS パスを確立することを仮定する。この場合、PCC が PCE1 を文献 21) を用いて検出し PCReq メッセージを送出する。PCE1 は AS2 の終点ルータに関する経路を知らないで、明示的に設定された PCE2 に対して PCReq メッセージを転送する。その後、PCE2 は ASBR3 及び 4 から終点ルータまでの最短経路を計算し VSPT (Virtual Shortest Path Tree) として PCE1 に送付する。PCE1 は AS1 のトポロジと受信した VSPT を使用してエンド・エンドで最短の経路を計算する。その結果、4-2-3 節、4-2-4 節にて解決できなかったエンド・エンドで最適な MPLS パスを確立することが可能となる。

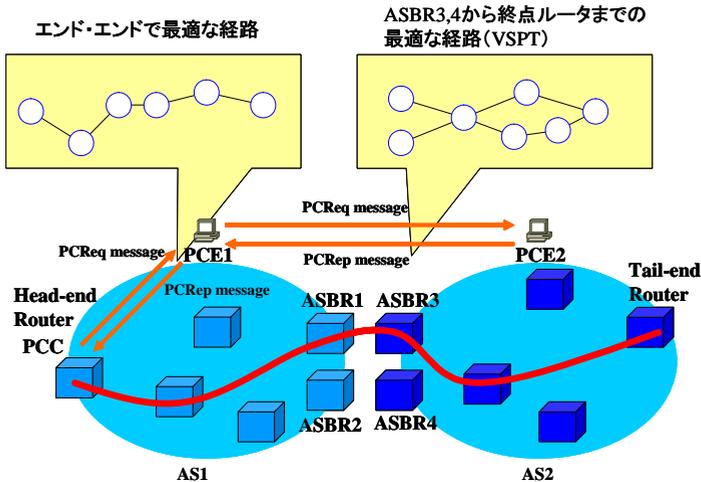


図 4・10 BRPC 計算手法

■参考文献

- 1) E. Rosen, A. Viswanathan and R. Callon, "Multiprotocol Label Switching Architecture," IETF RFC3031, Jan. 2001.
- 2) A. Farrel, J.P. Vasseur and J. Ash, "A Path Computation Element (PCE)-Based Architecture," IETF RFC4655, Aug. 2006.
- 3) E. Rosen, D. Tappan, G. Fedorkow, Y. Rekhter, D. Farinacci, T. Li and A. Conta, "MPLS Label Stack Encoding," IETF RFC3032, Jan. 2001.
- 4) L. Andersson, I. Minei and B. Thomas, "LDP Specification," IETF RFC5036, Oct. 2007.
- 5) D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels," IETF RFC3209, Dec. 2001.
- 6) D. Katz, K. Kompella and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2," IETF RFC3630, Sep.

- 7) H. Smit and T. Li, "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)," IETF RFC3784, June 2004.
- 8) F. Le Faucheur, J. Boyle, W. Townsend, D. Skalecki, K. Kompella and T.D. Nadeau, "Protocol Extensions for Support of Diffserv-aware MPLS Traffic Engineering," IETF RFC4124, June 2005.
- 9) F. Le Faucheur, J. Boyle, K. Kompella, W. Townsend, T.D. Nadeau, D. Skalecki, "Russian Dolls Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering," IETF RFC4127, June 2005.
- 10) J.P. Vasseur, M. Pickavet and P. Demeester, "Network Recovery," pp.309-310, pp.314-324, p.336, p.341, Morgan Kaufmann Pub, Aug. 2004.
- 11) P. Pan, G. Swallow and A. Atlas, "Fast Reroute Extensions to RSVP-TE for LSP Tunnels," IETF RFC4090, May 2005.
- 12) J.L. Le Roux, J.P. Vasseur, J. Boyle, T. Chung, R. Zhang, K. Kumaki, Y. Ikejiri and P. Lahiri, "Requirements for Inter-Area MPLS Traffic Engineering," IETF RFC4105, June 2005.
- 13) J.P. Vasseur, A. Ayyangar and R. Zhang, "A Per-domain path computation method for establishing Inter-domain Traffic Engineering (TE) Label Switched Paths (LSPs)," IETF RFC5152, Feb. 2008.
- 14) J.P. Vasseur, Y. Ikejiri and R. Zhang, "Reoptimization of Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Loosely Routed Label Switched Path (LSP)," IETF RFC4736, Nov. 2006.
- 15) J.P. Vasseur, Z. Ali and S. Sivabalan, "Definition of a Record Route Object (RRO) Node-Id Sub-Object," IETF RFC4561, June 2006.
- 16) R. Zhang, J.P. Vasseur, K. Kenji, P. Mabey, N. Constantine, P. Merck, T.W. Chung, J.L. Le Roux and Y. Kim, "MPLS Inter-Autonomous System (AS) Traffic Engineering (TE) Requirements," IETF RFC4216, Nov. 2005.
- 17) 熊木, "PCE applicability," MPLS JAPAN 2005, [http://www.mpls.jp/2005/presentations/051121\\_11.pdf](http://www.mpls.jp/2005/presentations/051121_11.pdf), 2005 年 11 月.
- 18) J.L. Le Roux, J. Ash, N. Bitar, D. Cheng, K. Kumaki, E. Oki, R. Zhang, R. Zhang, "Path Computation Element Communication Protocol (PCECP) Specific Requirements for Inter-Area MPLS and GMPLS Traffic Engineering," IETF RFC4927, June 2007.
- 19) N. Bitar, R. Zhang, K. Kumaki, "Inter-AS Requirements for the Path Computation Element Communication Protocol (PCECP)," IETF RFC5376, Nov. 2008.
- 20) J.P. Vasseur, R. Zhang, N. Bitar and J.L. Le Roux, "A Backward Recursive PCE-based Computation (BRPC) Procedure to Compute Shortest Constrained Inter-domain Traffic Engineering Label Switched Paths," IETF RFC5441, April 2009.
- 21) J.L. Le Roux, et al., "OSPF protocol extensions for Path Computation Element (PCE) Discovery," IETF RFC5088, Jan. 2008.

## ■3群 - 5編 - 4章

### 4-3 NGN への品質制御技術の適用

(執筆著：山田秀昭) [2011年1月 受領]

サービス品質については、例えば電話のような広域サービスの場合、社会基盤としての重責から、災害時の通信輻輳対策も含め、従来のベストエフォート型のサービスでは実現が困難な、高い信頼性と安定した品質が要求される。

NGN (Next Generation Network) では、情報転送に必要な帯域や同時接続数などの品質要件を、サービス利用時に端末からネットワーク側システムへ要求し、ネットワーク側で品質制御を行う仕組みが適用されている。このような仕組みは、音声通話や映像配信、SNS (Social Network Service) チャット、ファイル転送など、通信の多様化にかかわらず、全てのユーザが安心してサービスを利用するために必要である。

#### 4-3-1 品質制御の仕組み

IP ベースのネットワークを基本とする NGN では、サービス提供機能 (サービス・ストラタム) とパケット転送機能 (トランスポート・ストラタム) がそれぞれ分離独立した 2 層 (二つのストラタム) 構成を特徴としている。ここで、ITU-T で定義されている NGN の基本構成を図 4・11 に示す。



図 4・11 ITU-T NGN の基本構成

サービス・ストラタムは、サービス提供用のインタフェースを提供するアプリケーション / サービス・サポート機能と、電話サービスや IPTV サービスといった、サービス単位の接続制御などを実現するサービス制御機能から構成される。

一方、トランスポート・ストラタムは、IP アドレスの割当てやユーザの認証を行うネットワーク接続制御機能、サービスの品質を測定管理する MPM (Management of Performance Measurement) 機能、端末エンド・ツー・エンドのネットワーク資源予約やアドミッション制御を行う RACF (Resource and Admission Control Functions) 機能、及び、トランスポート機能から構成される。ここで、トランスポート・ストラタムの機能構成を図 4・12 に示す。

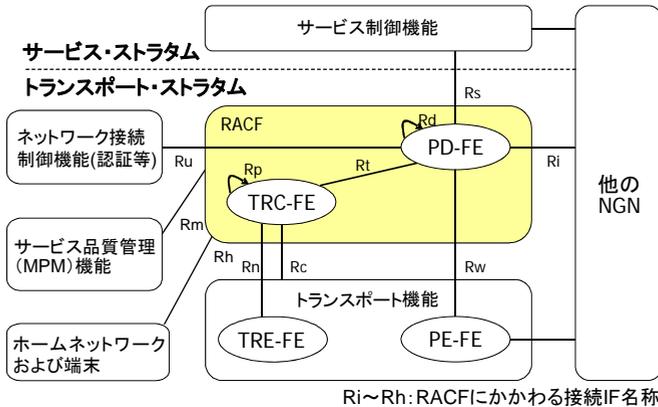


図 4-12 トランスポート・ストラタムの機能構成

#### (1) RACF (Resource and Admission Control Functions) 機能

RACFを構成する機能要素 (FE; Functional Entity) は、前述のサービス制御機能からの指示や端末からの要求を受信し、トランスポート機能からのネットワーク資源情報に基づいてサービスの実行に必要な帯域幅やパケット転送の優先度設定を行う PD-FE (Policy Decision-FE) と、トランスポート技術に依存したネットワーク資源予約/アドミッション制御機能を指示する TRC-FE (Transport Resource Control-FE) から構成される。これら二つのFEは、トランスポート機能内にある、PE-FE (Policy Enforcement-FE) 及び TRE-FE (Transport Resource Enforcement-FE) と通信することにより、品質制御を実現する。ここでPE-FEは、PD-FEの指定したパケット転送条件に従い、トラヒックの分類やパケット転送の優先順位のマーキング、ポリシング/シェーピングと共に、NAPT/NAT (IPアドレスやアプリケーションが使用するIPレベルのポート番号の変換) 制御を行う。一方、TRE-FEは、品質制御対象のネットワーク・トポロジーやリソース情報を収集し、TRC-FEへ通知する機能を提供する。

#### (2) トランスポート機能

トランスポート機能は、RACFを経由したサービス制御機能からの要求に基づき、サービスに応じた品質制御を実現する。更に、固定 (FTTH/ADSL など) や移動 (無線 LAN, WiMAX など) といったアクセスネットワークとの接続を実現し、パケット転送の優先度設定や NAT (Network Address Translation) 設定などのネットワークエッジ機能、大容量なルータやスイッチ群からなるコアネットワーク機能から構成される。

### 4-3-2 品質制御機能

RACFは、サービス制御機能やトランスポート機能と通信を行い、主に、リソース/受付制御機能、サービス・トラヒックの制御機能、転送品質クラスに応じたパケット転送機能などに基づく品質制御を実現する。それぞれの機能について以下に示す。

#### (1) ネットワーク資源予約/アドミッション制御機能

回線交換型の電話サービスでは、物理的な接続回線数があらかじめ決まっており、受け付



行われる。なお、送信スケジュールについては、PQ (Priority Queuing) や WFQ (Weighted Fair Queuing) などの転送パケットのキューイング・アルゴリズムが適用されている。

なお、現状の RACF の PD-FE は、サービス開始時における静的な品質制御をサポートしており、通信の途中でパケット転送の優先度を変更するといった、動的な品質制御については、今後の課題である。

### 4-3-3 品質制御

RACF による品質制御については、端末からサービス制御機能への品質要求に対し、サービス制御機能がリソース/受付制御の完了までをサポートする“プッシュ型制御”と、サービス制御機能がトランスポート機能へのアクセス許可を端末に与え、端末がトランスポート機能と直接リソース設定などを行う“プル型制御”の2種類のシナリオがある。

プッシュ型制御シナリオは、端末が品質設定のネゴシエーション機能を具備しない場合や、トランスポート機能と端末間でのリソース設定が不許可/不可能な場合のシナリオである。他方、プル型制御シナリオは、RSVP といったリソース予約プロトコルを活用する場合など、端末がトランスポート機能に対する品質設定のネゴシエーション機能を具備する場合のシナリオである。これらの品質制御シナリオを図 4・14 に示す。このように、二つの制御シナリオは、端末が具備する品質確保の機能に応じて使い分けられる。

#### (1) プッシュ型制御

- ① 帯域や同時接続数など、品質にかかわる要求を含むサービスのマルチメディア情報を転送するセッションの設定要求 (SIP INVITE, HTTP GET などの要求) を、端末からサービス制御機能へ送信する。
- ② サービス制御機能は、サービスセッションの設定要求から品質にかかわる要求条件を抽出し、リソースへの要求を RACF へ送信する。
- ③ RACF は、サービス運用方針に基づくリソースへの要求を確認するとともに、リソース情報、トランスポート機能へのアクセス権限情報などに基づいて、サービスの受付判断を行う。その後、リソース制御を行うべく、前述のサービス・トラヒックの制御や転送品質クラスに応じたパケット転送の実行指示を、トランスポート機能に指示する。

#### (2) プル型制御

- ①と②はプッシュ型制御と同じ。
- ③ ユーザのサービス利用条件を確認後、ネットワーク資源への要求を転送する。
- ④ 端末は、サービスセッションの設定要求に対するサービス制御機能からの返信を受け、返信メッセージに確認トークンが含まれている場合は、それと共に、RACF が返送した要求条件をトランスポート機能に送信する。
- ⑤ トランスポート機能は、端末から送信された要求条件を RACF へ送信する。
- ⑥ リソース情報などに基づいて、サービスの受付判断を行う。その後、リソース制御を行うべく、前述のサービス・トラヒックの制御や転送品質クラスに応じたパケット転送の実行指示を、トランスポート機能に指示する。

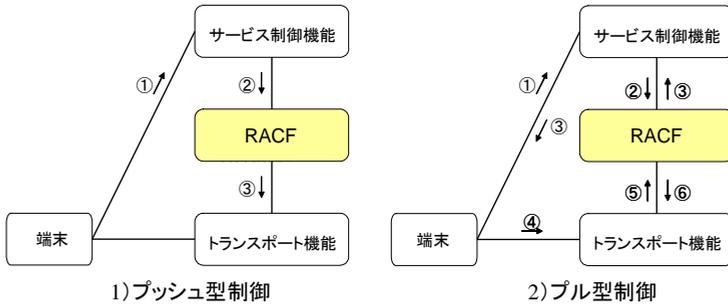


図 4・14 RACF 制御シナリオ

#### 4-3-4 事業者ネットワーク間における品質制御の連携

通常のサービス運用においては、無線／有線などを含む複数の異なる事業者ネットワークを介した環境が一般的であり、その際にも RACF は、図 4・15 に示すような  $R_i$  と呼ばれるインタフェース (IF) を介した通信を行い、RACF 間で連携した品質制御を行う。

また、RACF 間の連携動作においては、異なる事業者ネットワーク間でのアドレス隠蔽やプライベート・アドレスの利用を実現し、アクセスネットワークとコアネットワークや、コアネットワーク間のアドレス変換を行う。更に、変換したアドレス情報と一致するように、サービスのマルチメディア情報を制御するシグナリング・メッセージ内のアドレス情報を書き換えるなど、ゲートウェイ制御機能を提供する。

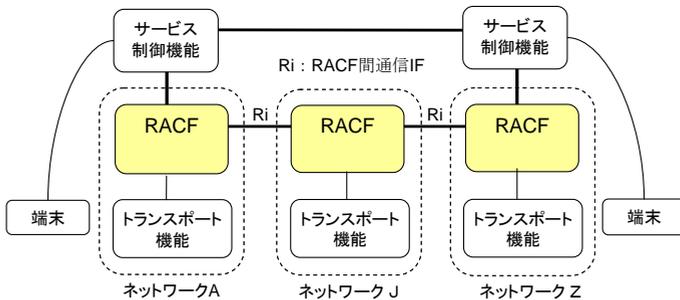


図 4・15 RACF 間連携

#### 4-3-5 端末モビリティを考慮した制御

NGN の多様なアクセスネットワークを移動しながらサービスを継続する場合、サービス端末の環境変化に連動した端末エンド・トゥ・エンドの品質維持にかかわる課題がある。図 4・16 の例では、端末とネットワーク側のシステムが連携して品質情報を収集し、これを活用した RACF 制御を実現する。

具体的には、品質変動の大きな FMC (Fixed and Mobile Convergence) ネットワーク環境に

における端末エンド・トゥ・エンドの品質を、MPM (Management of Performance Measurement) 機能が把握し、通信中の品質変動を考慮した動的なサービスのセッションの切り替えやマルチメディア情報の転送レートの制御を行う。ここでは、音声と映像を問わず、ほとんどのマルチメディア情報の転送に利用されている RTP (Real-time Transport Protocol) とその制御プロトコルの RTCP (RTP Control Protocol) により、MPM 機能が、端末側パケット受信バッファを含む端末エンド・トゥ・エンドのロスや遅延などの品質情報を、周期的に測定して管理している。

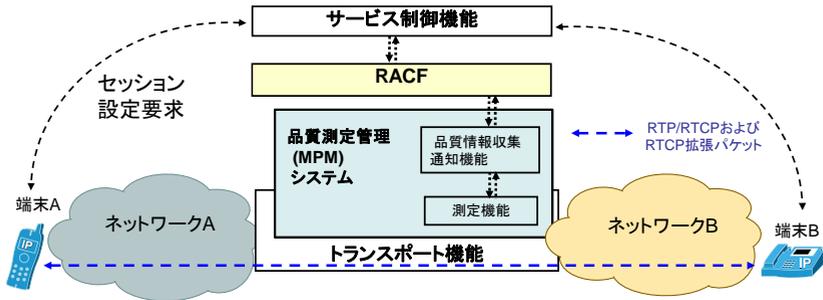


図 4・16 端末モビリティを考慮した RACF 制御

#### ■参考文献

- 1) 今中秀郎, 鎌谷修, 山田秀昭, 原田崇, “NGN リリース 1 のサービス要求条件と機能・制御アーキテクチャ,” 電子情報通信学会誌, vol.89, no.12, pp. 1051-1056, 2006.
- 2) ITU-T 勧告 Y.2111, “Resource and admission control functions in Next Generation Networks,” 2006.9.
- 3) 井上祐二 監修, “最新技術 NGN 入門,” インプレス R&D, 2007.2.