

プリフェレンスによる経路制御と政策的経路制御

村山 優子
WIDE プロジェクト

松尾英普
日立電線株式会社

通信における各終端サイトには、網レベルの経路についての要望があり、その実現をプリフェレンスによる経路制御とよぶ。本論文では、既存の政策的経路制御手順を用いてのプリフェレンスによる経路制御問題の解決を試みる。

ROUTING BY PREFERENCE AND POLICY ROUTING

Yuko Murayama Hideyuki Matuo
WIDE Project Hitachi Cable, Ltd.

Each one of the end sites of a communication flow has its own preference in network routing. Routing by Preference is concerned with how to enable those end sites to choose their preferred paths. This paper examines a possible solution for Routing by Preference by means of the existing policy routing protocols.

1 まえがき

政策的経路制御 (Policy Routing) [13][1] とは、今や地球規模に広がりつつあるインターネット網を、域という単位に分け、各域をノードとみなすような域間網における経路制御において、それぞれの域内の網資源についての使用方針を反映させることである。従来の経路制御のように単純にひとつのコストについての最適な経路を決定するものではなく、各ノードの異なる主張や方針を考慮に入れて経路は決定される。域とは、共通の運用方針に従う物理的に連続した網の集合である。「物理的に連続」とは、その網集合内では、内部経路だけで、ある地点から任意の地点まで到達できる範囲という意味である。

域内部の経路制御は、この域間経路制御とは独立の手順で行なわれる。階層化思考 [4] [7] からすると、ネットワーク層が2層に分かれ、域間経路転送は上位層、域内経路転送は下位の機能と考えられる。域内の通信は、域内経路制御による内部経路転送で行なわれ、域外への通信パケットは、域内経路制御により最寄りの域間経路転送用の中間ノード (router) に域外へ送られる。これとは逆に、域外からのパケットは、まず、域間経路転送用の中間ノードに受けとられ、次に内部経路制御によって宛先あるいは次の域間経路転送用中間ノードに届けられる。

さて、我々は昨年、新しい種類の域間経路制御の分野として通信の終端サイトの意思を考慮した経路制御、プリフェレンスによる経路制御 [12] を指摘した。本論文では、既存の政策的経路制御を使って、プリフェレンスによる経路制御が、どこまで実現可能かを探る。

以下、2. では既存の政策的経路制御を紹介し、3. ではプリフェレンスによる経路制御のための必要条件を論じ、4. では既存の政策的経路制御による解決を試み、5. はこれらの議論のまとめをおこなう。

2 既存の政策的経路制御

2.1 概要

政策的経路制御の背景には2つの流れがあり、それぞれの流れから異なるプロトコルが作成されてきた。ひとつは、インターネット網の増大問題であり、その規模の拡大と網の相互接続の複雑化から、Border Gateway Protocol (BGP) [8] が生まれた。もうひとつは、アクセス制御 (access control) の観点から生まれた Inter-Domain Policy Routing (IDPR) [11] である。

BGP は distance-vector 型で、IDPR は lin-state 型の経路情報交換を行なう [9]。前者では、各隣接ノードから得られた距離情報に自分とその隣接ノードの間の距離を加えたものの中から、宛先まで最も距離の短い隣接ノードを次の転送先とする。後者では、各ノードが隣接ノードのリストと、それら隣接ノードへのコストを網上のすべてのノードに配り、従って、各ノードでは網全体の構成図を持ち、それによって経路決定できるというものである。

日本のインターネット網では、現在ノードの増大問題に対処するため、米国で広く使われている BGP を使用する方向にある。また、International Organisation for Standardisation (ISO) の開放型システム間接続 (OSI) 環境での域間経路制御として出てきた Inter-Domain Routing Protocol (IDRP) [6] は、BGP のスーパーセットと考えることができる。

以下に BGP, IDRP, IDPR を紹介する。本論文では、インターネット網環境の経路制御における終端システムをホストと呼び、中間システムをルータと呼ぶ。ルータの内、特別な経路制御手順動作を行なうものをそのプロトコル名をつけた形で呼ぶ。すなわち、BGP の仕様に従うルータ (Border Gateway) を BGP ルータと呼び、IDPR の仕様に従うルータ (Policy Gateway) を IDPR ルータと呼ぶ。

2.2 BGP

インターネット網においては、初期には、Exterior Gateway Protocol (EGP) [10] が、域間経路制御手順として生み出された。当時は、まだ、ARPANET 時代から相変わらず、網管理は一貫して BBN で行なうという状況 [3] で、EGP では、域間 (inter-domain) 経路制御の階層も、最上階は BBN 管理下のノードがコア・システムを形成し、その下に各サイト管理下の網が、木構造 (tree structure) でつながるように決められていた。ループ問題を回避するための、この強制された物理的な木構造 (engineered tree structure) は、その後発展を続けるインターネット網の任意な相互接続の構造にそぐわなくなっていた。それまでのインターネット網は、米国の国防省の援助によって構築されていたが、学術系の The National Science Foundation (NSF) による主要網 (backbone) が構築されるに至り、異なる運用方針を持つ主要網や地域網間の経路制御は単純に階層化できず、対等な域間の制御

が必要となった。こうした背景から、EGP を発展させた形で、Border Gateway Protocol(BGP) [8] は生まれた。

BGP は、各ノードにおける経路決定 (hop-by-hop routing) のための情報交換システムである。各 BGP ルータは、自分から到達可能な網や網群までの距離を中間ノード数などで表し、隣接する域の経路情報サーバ(多くの場合これは BGP ルータに内蔵されていると考えられる) に配る。配られる経路情報には、距離に加えて、その距離が算出された経路についての次のような属性が含まれる:

- 目的地までの域単位の経路
- 今流している距離情報は、どのプロトコルによって学習したものか
- その経路における次のホップである BGP ルータ
- 以前可能だった経路が、現在到達不可能である
- 自分と隣接域との間に複数のリンクが存在する場合、それぞれのリンクの重み付け

特に最初の属性である各経路についての域単位経路の提示により、域間接続模様を示すグラフを構築でき、同じ地点を2度以上通ってしまうというループ問題を回避している。これが、物理的に木構造に接続しなければならなかった EGP から大きく逸脱している点である。

この他、各域にはローカルな方針・政策を指定した Policy Information Base (PIB) が存在し、どの情報をどの隣接域に配布するかについての制御を行なうことができる。また、各域は、自分自身が使用する経路についての情報だけを流すことになっている。これにより、確実に使用可能な経路だけが公開される。しかし、他の域からのパケットが、途中の域において、そこで使用されていない経路を選ぶことはできない。

EGP は、ネットワーク層の上に位置し、EGP 情報はそのままひとつの網レベル・パケットに含まれて送られていた。しかし、網の数が増加し、経路情報が一つの IP 層のパケットでは運び切れなくなり、一時は網情報が一部欠けてしまう事態となった。これは、各網の運べるパケットの最大値 (Maximum Transmission Unit — MTU) がそれぞれ異なるインターネット網では、経路情報を運ぶ IP パケットが、MTU が小さい網において、細分化 (fragmentation) されてしまい、受け取りのルータでは、細分化されたものを元に戻すこと (reassembly)

はできなかったからである。インターネット網では、ホストと呼ばれる終端システムでなければ、reassembly の機能はなかったのである。BGP は、トランスポート層の Transmission Control Protocol (TCP) の上に位置する。従って、細分化、再送などの制御は、TCP で解決される。しかし、そのためには、隣接する 2 BGP ルータ間すべてに、TCP 接続が必要である。

この他、最近では、発信元のルータで、域単位の経路を決める形式の Source Demand Routing Protocol (SDRP) が、この BGP を基盤とするプロトコル体系の一部として、出てきている。

2.3 IDR P

IDRP は OSI のネットワーク層における域間の経路情報交換を行なうプロトコルであり、BGP の機能を包含すると考えられる。域は、同じ経路制御手順によって動作するホストとルータのグループであり、単一の組織によって管理される。

IDRP は、インターネットにおける BGP を OSI の環境に適合させ、そこからさらに発展させたものである。この IDR P を逆にインターネット環境で動かし、政策的経路制御を実現するための研究も行なわれている [2]。

IDRP は、BGP と同様、経路に属性をつけ、メッセージに付加される属性と、ローカルに定義された方針・政策 (PIB) を用いて経路を選択する。

IDRP の特徴は次のようである:

- 特定の宛先に対して複数の網サービスの質 (Qualities of Service), に従った経路を持つことができる。
- 域のグループである域連合を構成でき、これによって大規模ネットワークへの適応や方針・政策の形成が容易にできる。

BGP は、TCP 上で動作するが、IDRP は、それ自身が、再送やフロー制御の機能を持っており、ネットワーク層 (Connectionless Network Service — CLNS [5]) の上で動く。

2.4 IDPR

IDPR は、網資源へのアクセス制御 (access control) の観点から、それまでの経路制御手順から全く独立に生まれた経路制御手順体系である。各域はその域を通過するパケットに対する、域の網資源のアクセス制御の方針

(transit policy) を基本的に flooding により他の域へ配布する。flooding 以外にも、request-reply 形式の経路情報収集方法も装備されている。域は、自分の通過パケットに対する方針と隣接域への接続(connectivity) 情報からなる経路制御情報を生成し、配布する。各域は経路制御情報の配布先を制限することにより、自域の網資源を利用する域を制限できる。

各域が主張できる通過パケットに対する方針の種類は BGP よりもかなり多い。例えば、ユーザ別(研究、商用)や発信元・宛先 AS による使用制御、提供できる網サービスの質、バイト、メッセージ、単位時間当りの課金等の金銭コストよりなる。その域内から外へ出ていくパケットに対する方針(source policy) は、経路として望ましい、または避けるべき域、許容できる遅延、速度、信頼性等の網サービスの質、許容できるセッションコスト等の金銭コストなどがある。これらは他の域には公開されない。

パケット転送は、発信元経路指定で行なわれる。域内からのパケットが外の目的地へ送り出される際には、1) そのパケットの種類、2) その域内の方針、3) 目的地までの通過域の方針などを考慮して、経路サーバが経路を通過すべき域と域間リンクを指定した域単位での経路を設定する。その域単位での経路に沿って各ホップとなる IDPR ルータが予め設定される。この経路は閉鎖されるまで、発信域から宛先域までのパケット配送に使われる。いつ閉鎖されるかは、各域の動作方針によるが、一般にこれは一回ごとのパケット転送とは独立に行なわれる。すなわち、あるパケット転送時には、目的地までの経路がすでに存在していれば、それを使用することができ、これによって、経路の開閉のためのオーバーヘッドが少なくなる。発信元ホストからでたパケットは、発信元 IDPR ルータで、経路 ID などのヘッダをつけた IDPR データメッセージとして宛先域まで送られる。

経路制御情報の配布や、経路設定など、網経路管理用のメッセージは、IDPR 用のトランスポート層プロトコル、Control Message Transport Protocol (CMTP) を使って送られる。CMTP は、connectionless で、request-reply 形式のプロトコルで、再送機能がある。

3 プリフェレンスによる経路制御のための条件

プリフェレンスによる経路制御とは、域間網環境において、通信の発信元と宛先それぞれの「望まない域ではできるだけコストを低く通り抜きたい。」という意思をいかに実現できるかということである。問題はどのようにして、発信元から宛先までの経路を構成する域の内、嫌いなところではできるだけ通過コストを抑えるかということである。すなわち、好きな域ではどんなにコストがかかっても気にしないが、嫌いな域ではあるコスト面での最短経路を通りたいという要求である。

このためには、次のようなことが必要である：

1. 発信元と宛先のそれぞれの意思情報の配布
2. 中間域の通過パケットに対する経路制御情報
3. 発信元と宛先のそれぞれの意思の調整および中間域の方針との調整

終端サイトの意思是、2つのパラメータ「避けたい域」と「そこで低く抑えたいコスト」で表される。「避けたい域」は、域の属性やその域内網のサービスの質(QoS)で指定することができる。例えば、域の属性には、特定のプロジェクトのメンバーである/ない、学術系域、研究系域、商業系域などがある。また、域内網のQoSや網レベル属性には、遅延時間、速度、安全性の度合い、課金コスト、域内の網の最大パケット長(Maximum Transmission Unit (MTU))などが考えられる。

「低く抑えたいコスト」は、距離に比例するものと、独立なものに分けられる。もとのり問題では、距離関係のコストの方が分かり易いが、独立なものでも論じることが可能である。避けたい域でおさえたい時間、安全性、課金などのコストを指定する。

これら終端サイトの意思是、通信の際考慮されたいのだが、そのためには、なんらかの形でこれらの意思情報が経路決定サイトに伝わっていなければならない。域間経路決定では、域単位での全体経路(route)やルータ単位での部分経路(path)の決定を行なう。BGPやIDRPのような、hop-by-hopなどの方式では、ルータごとに宛先が参照され全体経路と部分経路が同時に決定され、IDPRでは、発信元で全体経路が指定され、通信の初期に全体経路に沿った部分経路が各域で設定される。

また、「低く抑えたいコスト」要求を満たすために、部分経路決定サイトは、関係域における通過コスト情報を知る必要がある。これは経路情報の配布の問題である。

最後に、これら発信元、宛先、中間域のそれぞれの意向は調整が必要である。

4 既存の政策的経路制御での解決

ここでは、2. で紹介した既存の政策的経路制御手順を使用して、どこまでプリフェレンスによる経路制御が可能かを探る。以下、BGP、IDRP、IDPR、それぞれを用いた場合について論ずる。

これらのプロトコルはすべて、前節で述べた必要条件の2の中間域の経路情報については、問題なく既存の手順をそのまま使える。また、条件3の意思の調整については、通信の終端サイトの異なる意思を考えること自体が新しい試みであるので、いずれのプロトコルも調整機能はない。従って、以下では、条件1、すなわち、意思情報の配布について考える。

4.1 BGP を使用する場合

BGP も IDRP も各ルータにおいて経路が決定される (hop-by-hop) ので、終端サイトの意思は、各関係ルータに知らせなければならない。

発信元の意思は、おそらく、各パケットに付加情報として挿入した形で運ばれて行くことができる。これについては現在の網レベルのパケットに付加情報が必要となる。

宛先の意思については、各域が自分に宛てられたパケットの通る経路に対する要望を、前もって、経路情報の一部として、他に流すことができるであろう。ただし、部分経路についての要望は、ひとつの宛先については一様となり、発信元や通信内容によって変えることはできない。

4.2 IDRP を使用する場合

BGP 同様、各ルータにおいて経路が決定される (hop-by-hop) ので、終端サイトの意思は、各関係ルータに知らせなければならない。また、発信元の意思については、現在の網レベルのパケットに付加情報が必要となる。

IDRP では、宛先の域が、好きな域や嫌いな域を指定することができる。これは、経路情報の配布の際、その

情報を配布して良い、域あるいは配布して欲しくない域のリストを付加して送る。この情報を受けとった各域は、配布してほしくないというリストにある域にはこの経路情報を配布しない。その結果、避けたい域にはその宛先への経路がないことになり、この域にパケットは転送されることはない。

配布して欲しい域のリストを指定することにより、同様に情報配布の制御ができる。このように、IDRP では、情報配布の制御により、避けたい域を指定できる。発信元の意思についてはこのような指定方法はない。

IDRP では、このように異なる域を経由する複数の経路がある場合、宛先の域が経由して欲しい域または経由して欲しくない域を指定することができる。しかしコストを抑えて欲しい特定の域を指定することはできない。また、IDRP が扱う経路は域間の経路のみであり、域間における2つのリンクについての選択は指定できない。これについてはBGP 同様、情報配布の際に、2つのリンクに重み付けして違いをひきだせる。しかし、これでは、ある宛先について一様に部分経路が選ばれてしまい、経路についての意思が、発信元によって変えたり、通信の内容によって変えるわけにはいかない。

4.3 IDPR を使用する場合

IDPR は発信元指経路定であるので、発信元域内の経路決定者である経路サーバに、発信元の意思情報を内部で伝えることが必要だが、これは既存の手順に含まれている。

宛先の意思情報は、経路情報として発信元となる可能性のある域の経路サーバに配布することが可能である。しかし、既存の経路情報の形式に、ある域について、宛先としての意思を指定できる欄を付け加える必要がある。

条件3の調整は、発信域から宛先域への通信初期の経路設定時に実現できる。この場合、既存の経路設定プロトコルに、各域が設定拒否する場合の理由を具体的に示すことができるようにし、終端域同士の交渉が可能になれば、調整は完全に実現できるであろう。

5 むすび

本論文では、新しい経路制御の分野、プリフェレンスによる経路制御を数種の既存の政策的経路制御プロトコルを使用した解決を探ってみた。既存のプロトコルにいく

つかの機能を付加することによって、ある程度のことはできるであろう。

これら政策的経路制御や域間経路制御のプロトコル群は、元来、網サービスの提供者の要求から生まれた手順である。これに反し、我々のめざすプリフェレンスによる経路制御の基である、終端サイトの意思とは、網サービスの利用の際に、利用者とその達成したい仕事の内容や種類から、発生するものである。従って、本論文で考えた可能な修正を施したとしても、それは本当の解決とはなるとは考えられない。

我々はこの新しい形の経路制御のモデル化を進めている。現在、藤原による政策や意思も含めた一般的な経路制御モデル [14] が出来つつあるが、今後、その延長上で、終端サイトからみた経路制御要求のモデル化を進めたいと考えている。

謝辞

WIDEプロジェクトのPreferenceのグループのメンバーの諸氏の協力に感謝します。

References

- [1] D. Estrin. Policy requirements for inter-administrative domain routing. *Computer Networks and ISDN Systems*, No. 22, pp. 179-191, 1991.
- [2] S. Hares. Component of osi: Inter-domain routing protocol(idrp). *CONNEXIONS: The Interoperability Report*, Vol. 6, No. 5, May 1992.
- [3] R. Hinden, J. Haverty, and A. Sheltzer. The darpa internet: Interconnecting heterogeneous computer networks with gateways. *COMPUTER (IEEE)*, Vol. 16, No. 9, pp. pp.38-48, September 1983.
- [4] ISO. Iso 7498 information processing systems - open systems interconnection - basic reference model. International Standard ISO 7498, 1984.
- [5] ISO. Protocol for providing the connectionless-mode network service. International Standard ISO 8473, 1986.
- [6] ISO. Information processing systems - telecommunications and information exchange between systems - protocol for exchange of inter-domain routing information among intermediate systems to support forwarding of iso 8473 pdus. Draft Final Text ISO/IEC 10747:1993, ISO/IEC, July 1993.
- [7] L. Kleinrock and F. Kamoun. Hierarchical routing for large networks: Performance evaluation and optimization. *Computer Networks*, Vol. 1, No. 3, pp. 155-174, 1977.
- [8] K. Lougheed and Y. Rekhter. Border gateway protocol (bgp). RFC 1163, June 1990.
- [9] R. Perlman. *Interconnections: Bridges and Routers*. Addison-Wesley Publishing Company, 1992. ISBN 0-201-56332-0.
- [10] E. C. Rosen. Exterior gateway protocol (egp). RFC 827, October 1982.
- [11] M. Steenstrup. Inter-domain policy routing protocol specification: Version 1. Internet draft, May 1992.
- [12] 村山優子, 中村素典, 清水亮博, 藤原和典, 相川秀幸. 域間網経路制御における終端サイトの意思について. 情報処理学会研究報告, November 1993.
- [13] 村山優子, 曾根文樹. 政策的経路制御とその動向. 情報処理学会マルチメディア通信と分散処理研究会報告書, May 1993.
- [14] 藤原和典, 村山優子, 吉田茂樹. プリフェレンスによる経路制御のモデル. 情報処理学会研究報告, January 1994.