第 7 部

ポリシールーティング

第 1 章

はじめに

米国を中心とするインターネットは年々拡大を続け、初期の ARPA-NET と地域ネットワークという構成から、複数のネットワークの相互接続という形に変化した。相互接続した各々のネットワークは単一な組織や均一な技術によって作られたものではないため、使用目的や、スピード、バンド幅、課金方法など、さまざまな点で違いがある。このような環境では、発信者から宛先への経路が複数存在しうるが、従来の EGP/BGP/RIP などでは、一つのコストについて最適な経路を選択するだけであった。

これに対し、使用し得る経路が複数あれば、ユーザの都合のよい経路を選択できるようにしたい。そのための経路制御手法が最近注目されてきている政策的経路制御 (Policy Routing) [88] である。それは、従来のようにひとつのコストついての最適な経路を計算するものではなく、各ノードの異なる主張や方針を考慮に入れて経路を決定するものである。

この政策的経路制御には 2 つの流れがあり、それぞれの流れから 2 つの種類のプロトコルが作成された [89]。

ひとつは、インターネットの規模の拡大による情報量の増加と、網の相互接続の複雑化に対応するための、 Border Gateway Protocol [90] である。もうひとつは、 使用制御 (access control) の観点から生まれた Inter-Domain Policy Routing [91] である。

日本のインターネットでは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。 RIP や EGP、BGP は、中間システム経路決定型 (hop-by-hop) の経路制御のための情報交換手順である。経路情報は、distance vector と呼ばれる形式のもので、 情報提供ルータは、自分が到達可能な網のリストと各網への次のノードを配布する。そのため、これらのプロトコルは、ネットワークプロバイダ間の通信や複数のプロバイダに接続している大きな組織間で、トポロジと組み合わせて政策を含めた経路制御を行なうために使われている。

これに対して、IDPR は、発信元で通過経路を指定する (source routing) 経路制御手順体系である。パケットの発信元域から宛先域までの道があらかじめ準備され、その道に沿って、パケットが送られる。経路情報は link state 形式で、各域の情報が flooding で他の域に届けられる。

この他、BGP と IDPR の二つの流れの妥協案 [92] として、Source Demand Routing Protocol (SDRP) [93] という手順が指定された。これは BGP を使用する環境下で、発信元が宛先までの経路を域単位で指定する経路制御手順である。 SDRP のための情報交換

手順などはまだ検討中である。

以上を踏まえて、第二章では、政策的経路制御で要求されているポリシの具体的な事例を示す。実際のインターネットにおいてどのような要求があり、またその要求が実際にどの様に実現されているかを見ることによって政策的経路制御で実現しなければならないことが明確になる。また、1993 年度末の WIDE 合宿で行ったポリシについてのアンケートの結果についてのまとめを行う。

第三章では、第二章で議論したポリシのいくつかを解決する経路制御手法を検討する。政策的経路制御には、網の政策、通信の始点、終点の意思という三種類の要求があり、いくつかの通信の始点、終点の意思についての事例を得ることができた。「宛先までいくのにできるだけ速い網を使いたい」「自分あてのトラフィック (メールなど) はできるだけ速く自分の域に入ってほしい」などである。これらは、速い域を通ってほしい、ある域を通ってほしいということを表している。それは、「嫌いな域ではコストを抑えたい」という意思が発信元と宛先にそれぞれ存在し、それらが網レベルの経路制御でどのように実現できるかということである。本章ではこの問題を改めて「複数経路制御問題」と呼び、その解決法としてプリフェレンスすなわち終端サイトの意思による経路制御 (Routing by Preference) を提案する。

最後に、昨年度に引続いて IDPR や SDRP といった既存の政策的経路制御手法について調査を行なった。これを第四章にまとめる。

第 2 章

ポリシの実例

2.1 はじめに

政策的経路制御 (ポリシルーティング) に関する検討を行う上で、ポリシの実例に注目することは重要である。実際のインターネットにおいてどのような要求があり、またその要求が実際にどの様に実現されているかを見ることによって政策的経路制御で実現できなければならないことが明確になる。

本章ではポリシの現われている具体的な事例を提示し、さらに各組織のポリシの実体調査のため 1993 年度末の WIDE 合宿で行ったアンケートの結果についてのまとめを行う。

2.2 具体的な事例

政策的経路制御が必要とされる具体的な事例として以下のものを示す。これらの中には 1992 年度の報告書において詳細に検討されているものもあるが、再び簡単に整理してまとめておく。特に、最初の項目である WIDE–JAIN 複数地点相互接続問題はポリシルーティング WG の活動の発端的問題である。

- WIDE–JAIN 複数地点相互接続問題
- バックボーン共同運用問題
- 国際プライベートリンク活用問題
- 国内プライベートリンク活用問題
- マルチホームにおける経路選択問題

2.2.1 WIDE–JAIN 複数地点相互接続問題

1992 年頃の国内主要ネットワークは WIDE, TISN/GENOME および JAIN であった。これらのネットワークは図 2.1に示すように九州、大阪、京都、東京の複数ポイントにおいて相互接続を行っていた。

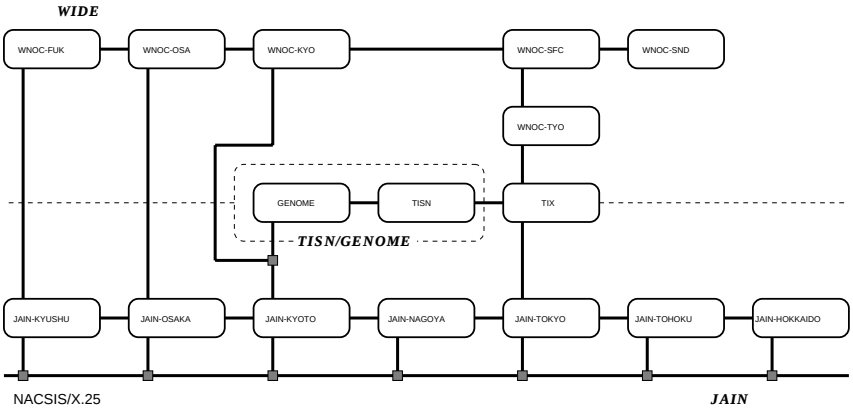

図 2.1: WIDE–JAIN の広域相互接続時代

ネットワークが複数ポイントで相互接続を行う場合にはそれぞれの接続点に関して様々な運用方針が考えられるが、特に回線の用途等に制限を持たせたりしない場合は、通信速度や回線容量などを考慮して、ネットワークに対する負荷がうまく分散するように経路選択が行われるようにすることが一般的である。WIDE と JAIN の相互接続にあたっては実際にそのような経路選択を実装する努力が行われていた。

例えば図2.1において、JAIN-KYUSYU と WNOC-FUK、JAIN-OSAKA と WNOC-OSA、JAIN-KYOTO と WNOC-KYO と GENOME、JAIN-TOKYO と WNOC-TYO と TISN のそれぞれ相互の通信に対しては、それぞれの直結リンクを利用したいという要求がある。また、相互接続をおこなっているリンクの両端同士以外の通信の場合は、どちらかの、あるいは両方のネットワークのバックボーンを経由することになるが、その場合はそれぞれのバックボーンのことをも考慮した適切な経路が選択されるようにすることが必要である。

WIDE と JAIN が相互接続する際に問題となったのは、回線の不均一な速度とそれぞれのネットワークの特性である。まず回線速度であるが、WIDE ではトラフィックの統計情報を取得して回線計画を立てており、ネットワークに対する需要が高い東京方面は他方面に対して太い回線が敷設されてきた。一方 JAIN は X.25 網の上に IP ネットワーク網を乗せていたという制約から 48K あるいは 64K の回線速度が最高であった。また、ネットワークの特性に関していうと、デジタル専用回線を利用していた WIDE に対して、X.25 網を利用していた JAIN では X.25 網によるオーバヘッドがかなり顕著に現われていた。このように、回線速度と性質に大きな違いがあるため、どの経路を選択するかによって、レスポンスが大きく異なる。結論として WIDE あるいは JAIN のどちらかのバックボーンを経由しなければいけない場合は、WIDE のバックボーンを通った方がレスポンスが良くなる。

このようなことから、経路情報交換プロトコルを操作して、WIDE と JAIN のどちらかのバックボーンを通る必要がある場合はできるだけ WIDE を利用するようにしたいという要求が出てきた。これが、いわゆるポリシの一形態であると考えられる。

最近では OSPF[11] や BGP[15] などの新しい経路情報交換プロトコルが研究開発され利用できるようになってきているが、当時はまだ RIP[9] を国内全域的に利用することしかできなかった。RIP では、ネットワーク的距離を示す値であるメトリックを隣接するゲートウェイ同士が交換することによってネットワークの到達性情報を広報する。したがって、選択されるべき経路の操作には単なるスカラ値であるメトリックの値を増減するしか方法がない。また、RIP のメトリック値の範囲は 1 から 15 までであり 16 は到達不能を示すため、非常に制約が大きい。特に、当時は RIP を国内全域で用いていたため、ネットワーク上の最遠地点間の距離が 16 に近くなると、メトリックの加算による操作が困難となる。さらに、RIP は目的ネットワークに到達するために経由する経路に関する情報を持たないため、得られた経路情報が複数あった場合に、利用する経路情報の選択が柔軟にできないという問題がある。

このように、RIP では制約事項が多く存在するため、ここで必要としている程度の政策でさえ実際の経路制御に反映することが困難となっている。

RIP を用いた政策的経路制御の検討については 1992 年度の報告書を参照されたい。

2.2.2　バックボーン共同運用問題

現在 WIDE と GENOME/TISN は京都–九州間のリンクを共同運用している。京都以東では WIDE と GENOME/TISN は独立したネットワークであるが、九州内部では特に明示的な区別は行われていない。

もしここで、九州側にも WIDE と GENOME/TISN の区別が存在すると仮定すると、九州で WIDE に属する部分は京都以東においても WIDE 部分を通り、九州で GENOME/TISN に属する部分は京都以東においても GENOME/TISN 部分を通るようにしたいという要求が出る可能性がある。この要求は、経路制御におけるポリシである。京都以東部分が完全に分離されているならばこの要求は容易に満たされるであろうが、京都以東部分において再び相互に接続されている部分があるとすると、その部分に関する経路の選択に関して問題が発生しうる。

この問題を抽象化すると、図2.2のようなモデルになる。この図の形状が金魚に似ていることから、このモデルは金魚モデルと呼ばれている。現時点における、このモデルに対する解としては、以下のものが考えられる。

1. ソースルーティングを用いる

2. 一つの物理回線上に複数の論理回線を構築し、論理的に分割して運用する

3. トンネリング技術を利用する

これらの解に対する検討は 1992 年度報告書に述べられているので参照されたい。また、トンネリング技術に関しては ddt-WG において議論が行われている。

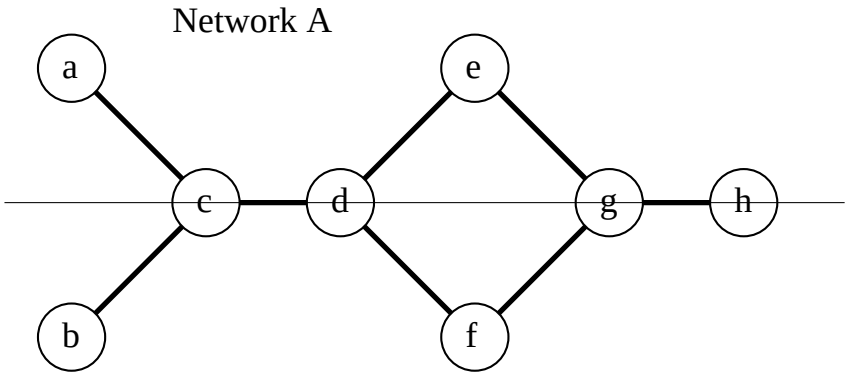

図 2.2: 金魚モデル

2.2.3 国際プライベートリンク活用問題

国外に関連会社を持つ企業では、独自にプライベート IP リンクを設置して業務に利用する場合がある。それぞれの会社はそれぞれの位置する場所でインターネットへの接続も行っていると、2 つの会社の間には 2 つの経路が存在することになり、2 つの経路の利用に関してポリシが発生する。

例えば富士通 (fujitsu.co.jp) の場合、アメリカの関係会社である Fujitsu America (fai.com) との間にプライベート IP リンクを持っているが、それぞれの会社は WIDE および BARRNET によってインターネットにも接続されている (図 2.3)。

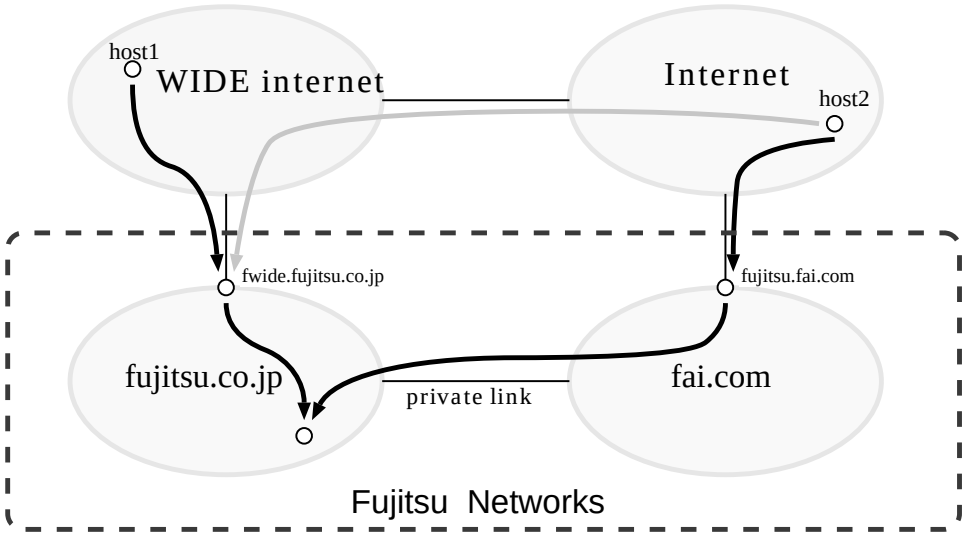

図 2.3: 国際プライベートリンクを持つ組織がインターネットに複数の接続を持つ場合

ここでインターネットで運用されている一般的な DNS の設定をそのまま用いたメール

の配送を考えると、fujitsu.co.jp 宛のメールは WIDE を経由して富士通のインターネットに対するゲートウェイに送られ、fai.com 宛のメールは BARRNET を経由して Fujitsu America のインターネットに対するゲートウェイに送られる。この状況下では、日本国内から発信された fai.com 宛のメールや、海外から発信された fujitsu.co.jp 宛のメールは共に WIDE の国際線を通過することになる。

ここで、これらの WIDE 国際線を通過するメールトラフィックをすべて富士通のプライベートリンクに振るということを考えてみる。これは 1 つのポリシの表現であるが、このポリシを実現する方法として次の 2 つの方法がある。

1. 日本国内、国外それぞれにおいてネームサーバの系列を 2 重化し、WIDE 国際線を当該メールが通過しないように DNS を設定する。

2. 経路情報を操作し、富士通/FAI へのパケットが WIDE の国際線を通過しないようにする。

前者のネームサーバの系列を 2 重化するという方法は管理上の手間が大きく、あまり現実的でない。日本では国際リンクの利用権にからむ特殊事情のためネームサーバの系列を複数設定しているが、これは本来は必要ないものであり、一企業の事情だけで系列を複数に分けることは容易に受け入れられることではなかろう。

後者はメールだけでなく全ての通信に対してポリシを実装することになり、根本的な解決方法である。この場合インタネット内に要求通りの経路選択ポリシを実現することも重要であるが、プライベートリンクも組織外から利用可能でなければならないという制約が発生する。また、2 つのインターネットの接続点から、それぞれ両組織の経路情報を配布しなければならないので、それぞれのプロバイダの承認も必要である。

2.2.4　国内プライベートリンク活用問題

前節と同様に関連企業がプライベートリンクによる接続を行っている場合ではあるが、さらにインターネットへの接続はそれぞれの企業が同一のネットワークプロバイダに接続している場合を考える。同一のネットワークプロバイダへの接続といっても様々な形態が考えられるが、実際の事例としてあったのは隣接した NOC にそれぞれ接続している場合である (図 2.4)。

この条件下で前節と同様にプライベートリンクを極力活用し、NOC 間のリンクに当該組織に関するパケットを通さなくしたいという要求 (ポリシ) を考えることができる。

これを満たすための要件は前節とほぼ同様であるが、DNS を同一プロバイダの西域と東域とで別系列にするという解はナンセンスであり、経路制御で解決するのが望ましいであろう。

このポリシを RIP を用いて実現する具体的な方法については、1992 年度の報告書で述べているので参照されたい。

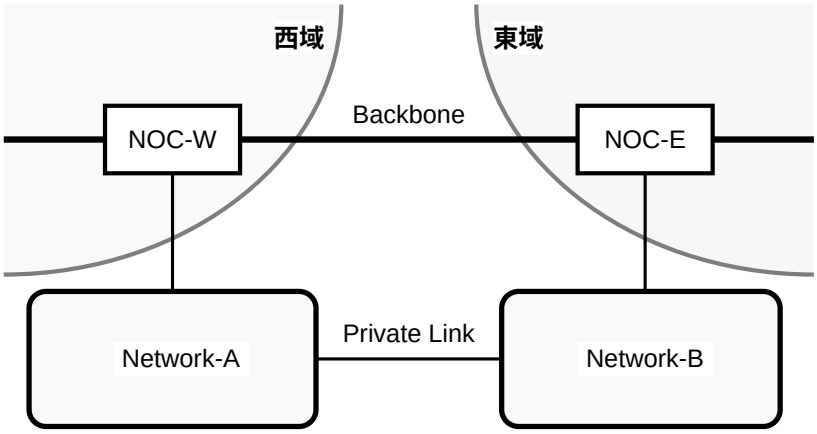

図 2.4: プライベートリンクを持つ組織の隣接 NOC への接続

2.2.5 マルチホームにおける経路選択問題

複数のネットワークに接続している組織で、当該組織に無関係なトラフィックの通過 (トランジット) を認めていないところはマルチホームと呼ばれる (図 2.5)。当該マルチホーム組織は複数の経路の中から一つの経路を選択する権利を持つ。

インターネットにおいては通常、到達性を最大化させることが望ましいので、一方のリンクからしか到達できない組織との通信である場合にはそのリンクを選ぶ。したがって、どちらの経路からでも到達できる場合にどちらのリンクを選択するかということが注目すべきポリシの要素である。

リンクが複数存在する場合のポリシとしては、リンクに順位付けをして順位が高い方のリンクを選択するようにする方法と、通信相手によって選択するリンクを振り分けるという方法、さらには TOS (Type Of Service) や負荷分散についても考慮する方法などが考えられるが、特に一般性のある要求として、通信相手がリンクが直接接続されているネットワークのいずれかに属する組織である場合には、他方から到達可能であっても、通常はその属しているネットワークへのリンクを利用したいというものがある。

この要求が実現できるためには、相手の組織がいずれのネットワークに属しているかが経路情報の一部として得られなければならない。当然人手を介入し、それぞれのネットワークに所属している組織のリストを静的に定義すれば明確に区別できるようになるが、人手の介入が必要なシステムは情報更新に遅れが出ることが目に見えているので、極力考えないこととする。

現在の技術をもとに考えた場合、経路情報交換プロトコルとして IGP (Internal Gateway Protocol) である RIP や OSPF を用いるとすると、現時点では以下のような状況でありあまり満足のいく設定ができない。

- RIP のみを利用
  経路情報にはどのネットワークに属しているかという情報が含まれないので、メト

リックの値に意味付けを行い値を恣意的に操作しないかぎり、リンク毎の組織の区別は難しい。

- OSPF のみを利用
  1 つのゲートウェイが、複数のネットワークに対してそれぞれ OSPF で別個に経路情報を交換することは、現在の gated では困難である。

- RIP と OSPF を併用
  OSPF ともう一つ別の経路情報交換プロトコルを利用すれば、プロトコルによってどちらのネットワークに属しているかが一応識別できるので、とりあえずなんとかなる。

マルチホーム組織に AS 番号を 1 つ割り当てて BGP などの External Gateway Protocol を用いることにすれば、相手組織の位置を明確に区別することができるようになるが、16 ビットしかない AS 番号をマルチホーム組織単位に割り当てることは現実的ではない。
以上のように、現時点では満足のいくスマートな解は見当たらない。

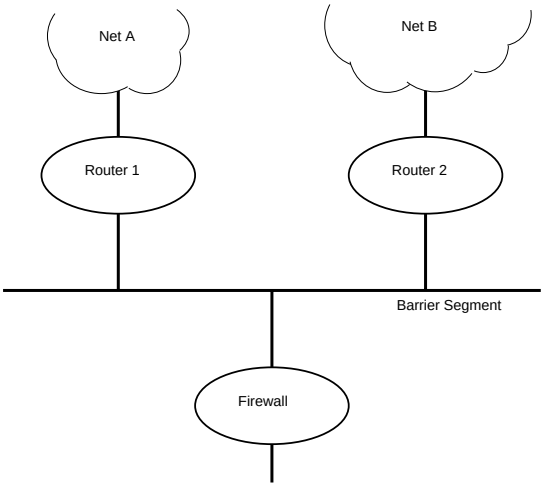

図 2.5: マルチホーム接続

2.3　WIDE 合宿でのアンケート調査

1993 年度末の WIDE 合宿において以下の項目に関するアンケート調査を行った。

- 組織の区分 [AC/CO/OR/GO/AD(ネットワークプロバイダと NIC)]
- 複数接続を持つか [yes/no]
- 通過 (トランジット) の許可 [yes/no]

- 経路制御の方針 (ポリシ)
- 経路制御の方針の実現の方法

2.3.1 アンケートの結果

各項目に対する回答を整理すると以下のような結果が得られた (表2.1)。

表 2.1: 回答のあった組織の区分と、うち複数接続を持つ組織の内訳

区分	回答総数	複数経路有
AC	15	6
CO	16	12
OR,GO	2	0
AD	9	–

回答があった AC 組織 (大学) で複数経路を持つ組織は、全てバリアセグメントを設けておりトランジットは禁止している。また、4 組織は経路制御の方針として SINET より WIDE 等を優先的に利用すると答えており、その実現方法としては gated などで RIP/BGP に関する設定部分で調整するという手段を採っているという回答が得られた。

回答があった CO 組織 (ネットワークプロバイダを除く企業等) で複数経路を持つ組織は WIDE と商用ネットワークへの接続を持っているが、さらに海外や SINET との接続をもつところもある。全ての組織はトランジットを禁止している。WIDE は研究ネットワークであり、IIJ/SPIN 等は商用ネットワークであって、利用用途が異なることから、企業ではネットワークの利用目的による使い分けが必要とされている。この使い分けを実現する手段としては、複数の firewall を設けてアプリケーションレベルで制御するという方法をとっているという回答が多く寄せられた。

OR,GO 組織で回答があった中には複数リンクを持つ組織はなかった。回答のあった 2 組織のうち一方からは、利用方針として、特定の研究目的利用について回線の優先的利用をさせることがあるという回答があった。

回答があった AD 組織 (ネットワークプロバイダおよび NIC) の内訳は表2.2の通りである。広域ネットワークプロバイダはトランジットに関して特に制限は行っていないが、その他の特殊なポリシを持つものとして、次のような回答があった。

JAIN: 回線の提供者 (学情) のポリシに従う
KARRN: バックボーン提供者 (WIDE, SINET) のポリシに従う
JPNIC: すべてのサイトとの通信

表 2.2: AD 組織の内訳

組織形態	回答数
広域商用	1
広域研究	1
地域	5
NIC	1

2.3.2　マルチホーム企業における経路選択の例

企業は、商用利用のできない研究ネットワークと利用に制限のない商用ネットワークの両方にリンクを持っている場合があり、リンクの使い分けが重要な問題となっている。

また、企業ではセキュリティを重要視するため firewall (防火壁: 外部からの自由な侵入を防ぐ機能をもつゲートウェイ) を用いているところが多くある。この場合、内部からインターネットへのアクセスにもある程度制限が加わるため、firewall において集中的に経路選択を行うことができる。

このようなことから、企業では以下のような解決方法を採ることによってリンクの選択をおこなっている。

- トポロジーによる解決
- firewall の連携による解決

トポロジーによる解決とは、図2.6のようにゲートウェイを論理的に離れた場所に置き、一つの組織を別々のアドレスを持つ二つのネットワークに分離して、リンクを使い分けようというものである。例えば、企業内部を研究と営業の 2 つの枠組に分割し、研究部門は研究ネットワークを、営業部門は商用ネットワークを利用するという形態をとる場合、トポロジーによる解決が可能である。

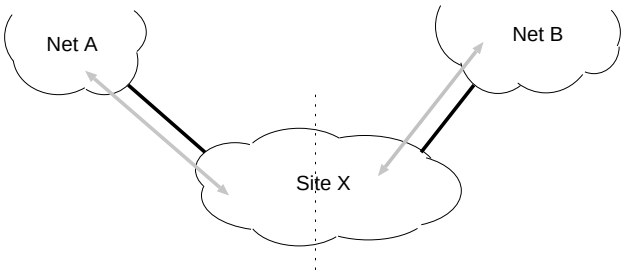

図 2.6: マルチホームでの経路選択 (1)

一方、firewall の連携による解決とは、Internet に直接接続できるマシンが firewall だけである場合に効果的で、図2.7のように、ネットワークプロバイダごとに firewall を用意し、各々の firewall には別々の class C のアドレスを与える。ネットワークプロバイダや経路を選択するためには、firewall を選べばよい。

この場合、firewall で発信および受信アドレスを参照して、経路の利用資格の判定を行なうことになるであろう。

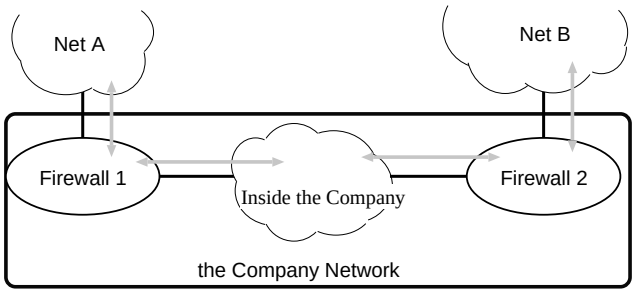

図 2.7: マルチホームでの経路選択 (2)

2.4 まとめ

本章ではインターネットにおいて求められている経路制御への要求と、実際に行われているポリシに基づく経路選択について具体的な事例を紹介した。これらの事例は、政策的経路制御に関して考察する上でおおいに参考となるものであり、これらの要求を満足する経路制御プロトコルの実現が望まれている。

第 3 章

プリファレンスによるルーティング

WIDE ポリシ・ルーティングワーキンググループでは、「日本における各域の政策とは?」ということについて議論していたが、日本のインターネットの現状では、「つながれば良い」という以外に具体的な方針を見い出すことができなかった。しかし、網制御の諸問題を検討していくうちに、ある基本的な要求が存在することを発見した。それは、「宛先まで行くのに、できるだけ速い網を通したい」ということである。これは、1992 年ごろの WIDE インターネットと JAIN の相互接続に関する問題であるので「WIDE–JAIN 相互接続問題」と呼んだ。同時に、2 つの国に網を持ち、それらが内部の国際リンクでつながっている組織では、「自分あてのトラヒックはできるだけ速く自分の域に入ってほしい」という要求があることが判明した。この 2 つの要求は、インターネット上での域間経路制御における興味深い問題に展開することができた。それは、「嫌いな域ではコストを抑えたい」という意思が発信元と宛先にそれぞれ存在し、それらが網レベルの経路制御でどのように実現できるかということである。本章では、この問題を改めて「複数経路選択問題」と呼び、その解決法としてプリフェレンスすなわち終端サイトの意思による経路制御 (Routing by Preference) を提案する。

また、プリファレンスによる経路制御を考慮したより一般的な経路制御の概要といくつかの問題点について考察し、最適経路の導出についての一つのモデルを提案する。

以下、 1 節では「速い網を通りたい」という WIDE–JAIN 相互接続問題や内部国際リンクを持つ組織の要求を解説し、 それらを域間環境の問題として展開した形で「複数経路選択問題」を再定義する。2 節では解決の条件を議論し、3 節ではより一般的な経路制御のアイデアを示し、4 節でそのためのモデルの一つを示し、5 節でこれらの議論のまとめをおこなう。

3.1 複数経路選択問題の定義

以下に WIDE–JAIN 相互接続問題、組織内リンク活用問題、そしてこれらを域間の環境において展開させた**複数経路選択問題**を論じる。

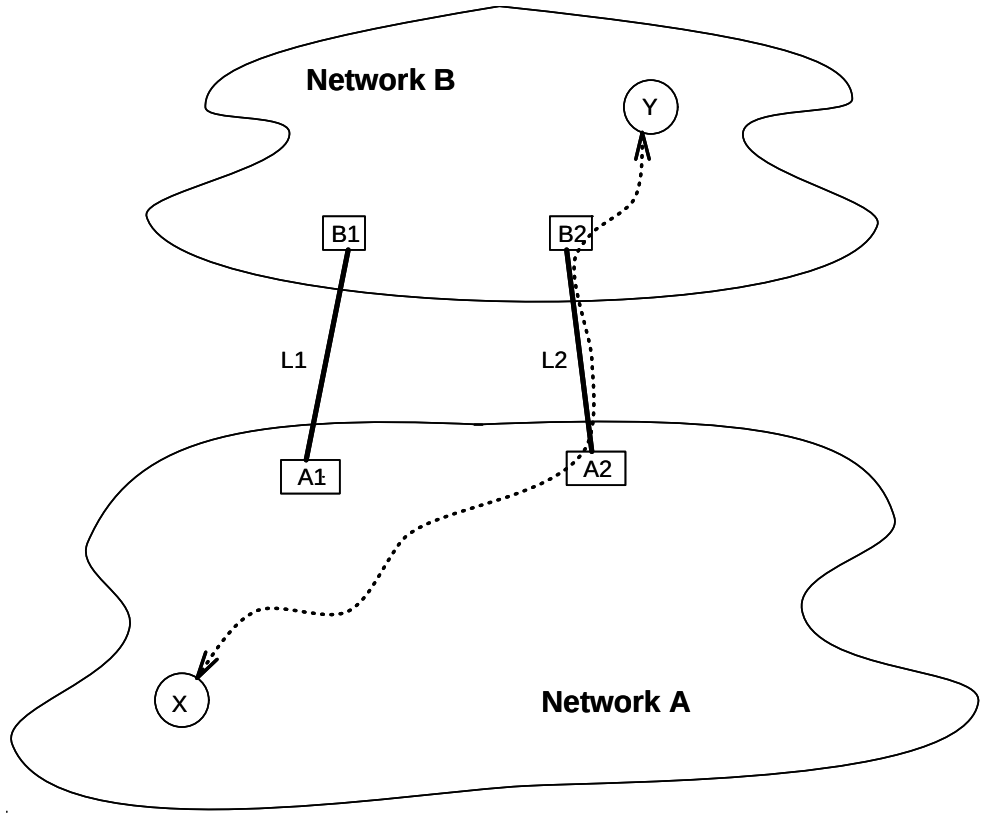

図 3.1: WIDE–JAIN 相互接続問題

3.1.1 WIDE–JAIN 相互接続問題

2 つの網を接続するリンクが複数ある場合、2 つの網間を流れるトラヒックについて、複数存在するリンクをそれぞれどの様に利用するかということに関して何らかの政策的要求が発生する。リンクが複数あると経路も複数できるため、経路の性質の違いによる得失を考慮した経路設定が望まれる訳である。例えば、 両方の網を接続するそれぞれのリンクの特性 (回線速度の違いなど) およびその位置付け (バックアップ的利用が目的であるなど)、さらに各網自身の持つ特性 (一方の網を構成しているバックボーンの回線速度が非常に遅い、 あるいはパケット交換に非常に時間がかかる) 等が経路の性質の違いとなる。

ここでは、 網を接続する各リンクの性質はほぼ等しいものと考え、一方の網の遅延が他方に比べて大きい場合について考えてみる。この状況下で考えられる要求としては、まず会話的処理が快適に行なえるようにするというものがある。

例えば、図 3.1のように L1、L2 の 2 本のリンクによって接続された 2 つの網 A、B があり、A、B を結ぶ 2 本のリンク L1、L2 の性質は同じものであるとする。図中の A1、A2 は網 A のゲートウェイであり、B1、B2 は網 B のゲートウェイである。さらに、網 A と B では遅延が大きく異なるものとし、網 A の方が B より遅延が少ないと仮定する。

ここで、網 A に属するサイト X と網 B に属するサイト Y の 2 つのホストの間の通信を考えた場合、図 3.1のようにサイト X がゲートウェイ A2 より A1 に時間的に近い位置にあり、 サイト Y がゲートウェイ B1 より B2 に時間的に近い位置にある場合、遅延を少なくするためには、一般に網 B を通過する時間を極力小さくするような経路が選択されることが要求される。具体的には、以下のような要求が満たされることが必要となる。

1. 網 A に属するサイト X から出たパケットは、サイト Y に最も近いリンク L2 を経由して網 B に移りサイト Y に到達する。

2. 網 B に属するサイト Y から出たパケットは、最も近い網 A へのリンク L2 経由して網 A に移りサイト X に到達する。

明示的に通信チャネルを設定するような経路制御を用いない限りは、 上記 1 と 2 の要求は個別に実装する必要がある。この要求を RIP(Routing Information Protocol)[9] を用いて経路情報を交換し、 最小コストにより経路決定を行なっている網において実装することを考えたが、 いずれかのリンクに障害が発生した場合にも最適な経路制御が行なわれるような一般的な制御は容易ではないという結論に達した [94]。

3.1.2　組織内リンク活用問題

海外にオフィスを持つ企業などでは、それらのサイトとの間で企業内リンクを持っていて、図 3.2のように二ヶ所でインターネットに接続している場合がある。ここで、その企業の全体の網を A、A の日本側部分を A1、海外部分を A2、A1 と接続しているインターネットの部分を B、A2 と接続しているインターネットの部分を C とする。

このとき企業 A は、インターネットの国際リンク L3 のトラヒックを軽減するため、あるいは、安全性の理由から、自分あてや自分からのトラヒックを、できるだけリンク L4 を通したいと考える。この方針は具体的には以下のようになる。

1. A1 中のサイト X からインターネットの C 中のサイト Y へ情報を転送する場合、なるべく長く企業 A の網を通したいので、 リンク L4、 L2 を通す。

2. C 中のサイト Y から A1 中のサイト X へ情報を送る場合、なるべく早く A2 に入れたいので、リンク L2、L4 を通す。

3. A2 中のサイトからインターネットの B 中のサイトへのトラヒックは、なるべく長く企業 A の網を通したいので、 リンク L4、L1 を通す。

4. B 中のサイトから A2 中のサイトへのトラヒックは、なるべく早く A2 に入れたいので、リンク L1、 L4 を通す。

3 と 4 は 1 と 2 にそれぞれ同等である。1 と 2 は前節で述べた WIDE–JAIN 相互接続問題と同じである。従って、 海外との内部リンクを持つサイトのメールの経路問題は、

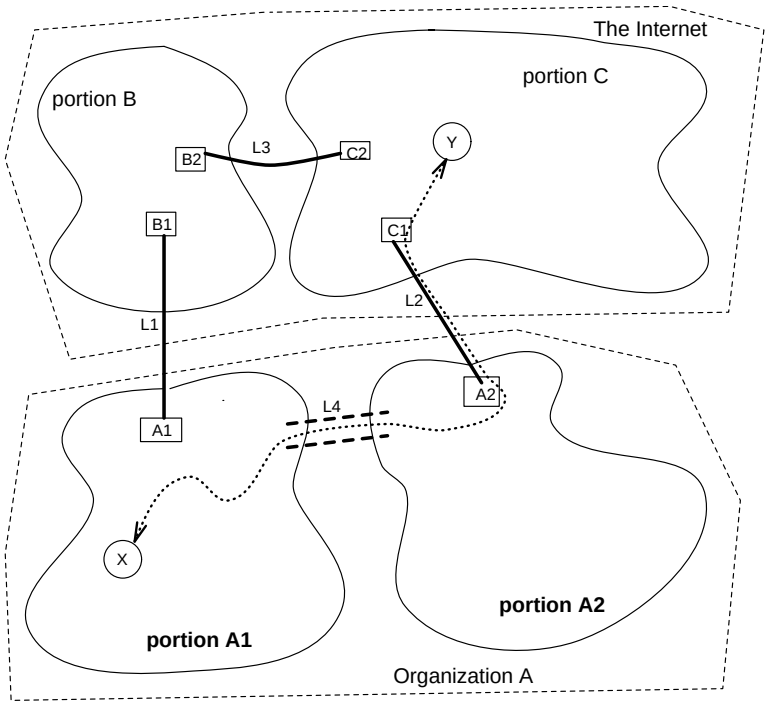

図 3.2: 組織内リンク活用問題

WIDE–JAIN 相互接続問題とほぼ同じであるように見える。しかし、WIDE–JAIN 相互接続問題は、 発信元の意思による問題であったのに対し、 ここで解説した内部リンク活用問題は、 宛先の意思による問題も含んでいる。

3.1.3 域間環境における複数経路選択問題

ここでは、「できるだけ速い (遅延の小さい) 通信をしたい」という問題と、「できるだけ組織内のリンクを使いたい」という要求をまとめ、域間 (inter-domain) における経路制御の問題として定義を行なう。以下、 この問題を**複数経路選択問題**と呼ぶ。

問題を域間環境に適用させるため、発信元と宛先を独立した域 — Autonomous System(AS)— に属し、その途中に様々な域が介在すると仮定する。事実、 複雑な相互接続のインターネットでは、 発信元や宛先が主要網などの通過のための中間網とは別である方が一般的であろう。図 3.3のように、 発信元は X という名の域 (AS) に属し、 宛先は AS Y という域に属する。

「できるだけ速い (遅延の小さい) 通信をしたい」という問題は次のように拡張できる。X と Y が通信する際、2 つの域、 AS A と AS B を通過しなければならないとする。ここで、2 つの域の通過コストが異なる場合、すなわち、A は B よりも速くパケットを通すとすると、問題は次の 2 つの方針を実現することである:

1. X に属する X0 から Y に属する Y0 に情報を送る時は、できるだけ A を通し、B に

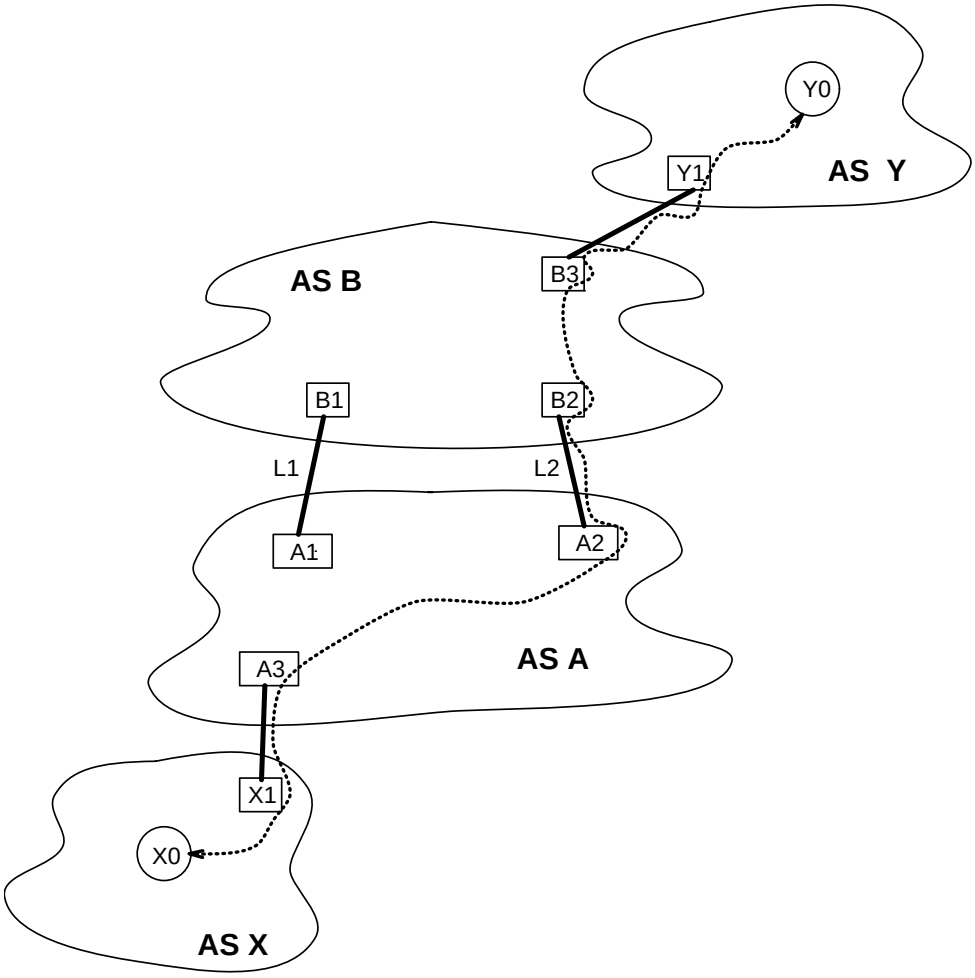

図 3.3: 複数経路選択問題

おける通過コストをできるだけ少なくする。

2. 逆に、 Y から X へ情報を送る時は、出来るだけ早く B から出し、A に入れる。

もう少し一般化すると、発信元から宛先までの経路を構成する n 個の AS のうち、 自分が望む m 個の AS がある場合、望まない AS ではできるだけコストを低く通したいという方針を実現することである。複数経路選択問題における発信元の要求は、 嫌いな AS ではできるだけ通過コストを抑えることで達成されるのである。すなわち、 好きな AS ではどんなにコストがかかっても気にしないが、嫌いな AS ではあるコスト面での最短経路を通したいという要求である。

さて、「できるだけ組織内の国際リンクを使いたい」という要求のうち、 その組織から外部へのトラヒックについては、 できるだけ自分のところに通って外に出るという発信元の要求は、 先に述べた WIDE–JAIN 相互接続問題と同じである。これに対して、 外

部からのトラヒックについての要求は、 できるだけ早く自分の組織域にはいって欲しいということで、宛先の要求である。

以下に、宛先 AS の受け入れトラヒックに対する要望を、 域間環境（図 3.3 参照）において展開していく。本来は、できるだけ早く自分の AS に入って欲しいということであるが、これを一般化すると、できるだけ早く**自分の好きな** AS に入って欲しいという要求に展開できる。それは、好きな AS まではコストを抑えて来て欲しいということであり、「望まない AS ではできるだけコストを低くおさえたい」ということに帰結する。

これらをまとめると、「**望まない AS では できるだけコストを低くおさえたい**」という意思が発信元にも宛先にも存在するということである。 これを**複数経路選択問題**と呼び、 その解決策を Routing by Preference と呼ぶ。

Routing by Preference では、発信元と宛先それぞれの意思 (preference)、また中間域の意志 (policy) の調整によって経路が決定されなければならない。ここで、 **意思**とは、 通信における終端サイトが発信元から宛先の間に介在する中間域全体に対する「できれば、かなえて欲しい」という要求であり、これに対して、 中間域の**意志**とは、 中間域が通過トラヒックに対して絶対行使されるアクセス制御の方針である。 すなわち、 Routing by Preference は、 様々な質の網サービスが提供され、 様々なアクセス制御の行使された域が混在する環境下で、 如何に通信の終端サイトの「**望まない AS では できるだけコストを低く通したい**」という意思を行使していくかの問題である。

3.2 解決の条件

問題の解決のためには、 Routing by Preference に必要な情報と、その情報がいつ誰により提供されるか、また実際にいつ使われるかを考えなければならない。これらは今後の研究課題であるが、ここでは現在考えられる可能性を論じる。第 1 項では中間域の情報について、第 2 項では始点・終点の要求について議論し、第 3 項ではその情報がどのように流れ、使用されて行くかを述べる。

3.2.1 中間域の情報

中間 AS の通過パケットに対する諸条件は、 経路制御情報として何らかの形で関係 AS に配布される。これには、AS 間の接続関係、各リンクの質の情報、 AS のパケット中継政策などが含まれる。パケット中継政策の実現の例としては、パケットの IP ヘッダをみてのフィルタリングという方法がある。この場合、パケット中継政策を各終端ノードに教えないと通信ができなくなる場合がある。

さらに、これらに加えて、Routing by Preference では次のような情報が必要となる。

発信元や宛先などの終端サイトの意思は「避けたい AS」と「低く抑えたいコスト」という 2 つのパラメータで表される。

避けたい AS は、 AS レベルの属性やその AS の網レベルのサービスの質 (Quality of Service(QoS))、AS 名などで指定することができる。例えば次のようなもので指定できる：

- AS レベル属性
  例）特定のプロジェクトのメンバーである／ない、学術系、研究系、商業系など。

- 網レベル属性
  例）AS 内の網の最大パケット長 (Maximum Transmission Unit (MTU))

- 網レベル QoS 基準
  例）遅延時間、 速度、 安全性の度合い、 課金コストなど

低く抑えたいコストは、 網サービスの品質（Quality of Services) のパラメータを指す。これは、 距離に比例するものと、 独立なものに分けられる。複数経路選択問題では、 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は可能である。発信元は、 避けたい AS で低くしたいパフォーマンスや安全性のレベル、 課金のコストなどを指定する。

3.2.2　始点・終点の意思

始点の意思

始点の意思には、パケット発信者の意思、サイトの管理者の意思、始点の網の意思のように階層的なグラニュラリティが存在すると考えられる。パケット発信者の意思にしても、使うアプリケーションや、時間により変化すると考えられる。また、一番基本的なグラニュラリティはコネクションであると考えられる。

原始的な意思はユーザにより与えられ、それを階層的に認証し、その階層での要求を付け加え、最終的にネットワークレベルの意思を作る。要求だけではなく、支払えるコストを与える。それには以下のようなものが考えられる。

- 遅延 (latency) を抑えたい
- (自分の組織からみた)security の重視
- 支払い可能なコストの最大値 (allowance)
- 経路の希望、通って欲しい網の指定、通って欲しくない網の指定

これ以外に、ユーザの意思を認証するために、ユーザ情報、アプリケーション (ポート番号) や、ユーザの要求の正当性より、実際に経路のもととなるネットワークレベルの意思を作成する。

例として、ある発信元から終点に到達する経路として、遅延の小さい学術目的に制限された回線と、遅延は大きいが目的に制限のない回線がある場合の経路選択について考える。発信元は、遅延の少ない速いネットワークが使いたいと主張するだろうから、どちらの経路を使ってもいい場合は遅延の小さい方を使いたい。そのとき自己申告の形で商用目的か学術目的かを与える。このような方針で経路を決めると、

商用目的の場合は 遅延は大きいが目的に制限のない回線、学術目的の場合は学術目的に限られているが遅延の小さい回線を使うような経路制御を行なうことになる。

終点の意思

終点の意思であるが、クライアント・サーバモデルでの終点はサーバ機能となり、外からのトラフィックに対する要求、すなわち以下のようなものを例として挙げることができる。

- セキュリティ問題などで、mail はできるだけ早く自分の信頼している網に入ってほしい
- anonymous ftp や Gopher などを公開してる場合に、パケットを発信する側が課金される網を経由する場合は拒否したい
- 遅延を抑えたい
- 経路の希望

始点の意思と同様に階層的に認証する必要がある。

これが、「国際リンク活用問題」のできるだけはやく自分の AS にはいってほしい場合に該当する。

終点の意思の階層化

次に、終点の意思の階層化について議論する。さきほど議論した域の政策のうち、ユーザ・クラスによる利用制御や、課金の管理などは、ユーザ個人を把握し、認証する必要から、終端域の各組織の網や、そのなかのサブグループ内の網レベルで解決されることになるだろう。

経路制御のグラニュラリティとしては、始点・終点の意思 (プリファレンス) をどのように与えるか、そしてそれをだれが認めるかという問題がある。なぜなら、域としてはその中の組織の網の 1 ユーザのことは知らないだろうからである。また、各組織でも、ユーザの所属部所でしか、各ユーザの事情は把握できないであろう。また、ユーザの意思以外に、所属部所での方針や、組織の方針、域の方針のため、終端の意思は図 3.4 のようにあらわせる。

ユーザがあるサービスを受ける時、ある意思 (プリファレンス) を持つとすると、一つ上の階層のユーザの所属部所でユーザ個人の認証を行ない、そして所属部所の政策/意思を付加されてさらに上位に送られる。そこで所属組織レベルの政策/意思を付加され、ネットワークレベルの意思 (プリファレンス) となる。この意思と各域の政策をもとに、経路を得る。

ここで例をあげる。ある組織では個別課金されるネットワークの利用は教員に限るような制限をしているとする。すると、部所として、学部は教員リストを持っているため、あるユーザが教員であり、課金されるネットワークを使ってでも速いネッ

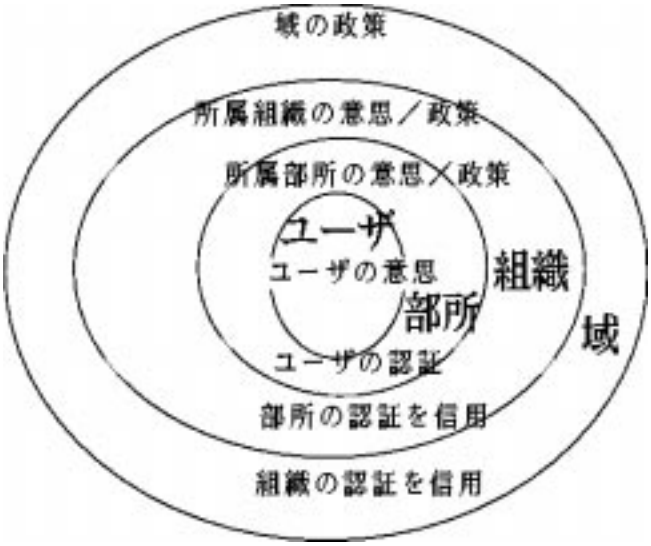

図 3.4: 終端の意思とその認証

トワークを使いたいという意思を持ったとする。すると、ユーザの所属部所である学部がユーザが教員であることを認証し、それを上位の組織レベルに伝え、組織の意思/政策を満たすか調べ、そして域の政策を満足したものがネットワークレベルの意思となる。

ユーザの分散管理という面から、ユーザ管理のドメインごとにこのようなシステムが必要であるので、DNS(Domain Name System)[95] の各ドメインなどに一つづつ程度、このような仕組みが必要になるだろう。

3.2.3　必要情報の流れと使われ方

終点の意思は、前項で説明した過程により発生し、網レベルでのプリフェレンス (意思) を生み出すと考えられる。ただし、 どのようなトラヒックについても、網レベルでは一様な意思が存在している場合もある。

終点の意思は、経路決定の際に使われる。域間経路決定では、 AS 単位での全体経路 (route) とルータ単位での部分経路 (path) の決定を行なう。これらは、 hop-by-hop などの方式のように同時に同じ場所で決定される場合や、それぞれ異なる時に異なるところで決定されることもできる。例えば、hop-by-hop を基盤とするインターネットにおいて発信元指定の AS 単位の経路を用いれば、 AS 単位の経路は発信元で決められるが、それぞれの AS 内での経路は、 hop-by-hop で各ルータごとに決定される。

発信元の意思情報、「避けたい AS」と「低く抑えたいコスト」は、hop-by-hop の場合、各関係ルータに知らせなければならない。この場合は、 おそらく、各パケットに発信元の意思を付加情報として挿入した形で運ばれて行くと思われる。発信元指経路定の場合は、 その発信元 AS 内の経路決定者 (例： 経路サーバ) に経路要求情報を内部で伝えるで

あろう。

また、「低く抑えたいコスト」要求を満たすために、 部分経路決定者は、中間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。

宛先の「避けたい AS」と「低く抑えたいコスト」の意思は、発信元からのトラヒックが発生する以前に存在すると思われる。この意思情報は、 発信元からの問い合わせに応じて発表されることもあり、 また、中間域の方針のように、経路情報として流すことができる。さらに、 この宛先域の意思の背景には、公的な主要網の一部のリンクの使用軽減につながる場合もあるので、 他の域も協力した形での経路情報制御が可能であろう。

中間域の方針（意志）も、 基本的には経路情報として他の域に配布されていく。既存の政策的経路制御のプロトコルで使用されている経路情報配布の方法には、distance vector における隣の ASs への配布や link state での flooding 方式などが上げられる。我研究グループでは、このような経路情報は網レベルで経路制御プロトコルの一部の機能を利用して配布するという従来の手法から離れ、もっと自由な発想の情報配布方法を検討中である。すなわち、 情報の配布自体を経路制御から独立した機能として見ている。現段階では、 次のようなものが考えられる。

- 新聞（マスメディア）方式
- セールスマン方式
- 問い合わせ方式

新聞方式は、本質的に flooding で、すべてのノードが一様な情報を配布する。X.500 ディレクトリの使用なども、 情報がだれにでも公開されるという観点から見て、 この方式の一種であると考えられる。

セールスマン方式は、 異なるノード群あるいは個別のノードに、それぞれ、異なる内容の情報を配布する。すなわち、伝える相手によって異なる情報を異なる暗号で流すことができる。このように、アクセス制御の面からは非常に有用な方法であるが、 域間網全体としての相互接続性の一貫性の維持に留意しなければならない。

問い合わせ方式とは、Request-Reply の方式で、必要な時に各ノードが相手ノードに直接問い合わせて情報を得る方式である。

この他にも数々の手法があると思うが、それらの検討とともに、その組合せた形の配布手法を考えてみたい。

ただし、一般的な経路政策情報を解くためには、終点の意思とすべてのネットワークの政策情報、接続情報、すべてのリンクのパラメータを持っている必要がある。少なくとも域に一つはこれらの情報を集めているノードが必要である。また情報量が大きいため、リンクステート型としたほうがいいだろう。基本形としては、すべての網で同じデータを持ち、同じ条件で経路を決定すべきであろう。

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。終端サイトの意思が、どの程度中間域の意志の上に実現されるかで経路が決定される。これらの方針や意志の承認手続きも必要で、 これらは今後の重要な研究課題である。

3.3　プリファレンスを考慮したより一般的な経路制御について

現在、わりと細かく政策的経路制御を行なうプロトコルとして IDPR が提案され、テストされている。IDPR では、網の政策、通信の始点の意思を実現するような経路制御を実現しようとしている。また、始点の意思については、接続組織などの大きめの単位でやろうとしているようである。

そこで、この節では、プリファレンスを考慮した上で、より一般的で実現可能そうな経路制御について考える。

経路制御に必要な網の政策、終点の意志などの情報は前節で説明した内容とする。その場合、次のような手順により、経路情報を計算する。

- 各域は政策や接続情報を全ての域に広報する
- 各域や各組織のルーティングサーバはすべての域の情報を収集する
- ユーザが新しいセッションを張るとき、ユーザの意思を上位の意思サーバに送る
- 中間意思サーバでは、ユーザを認証し、ユーザの意思に自分のレベルでの意思/政策を追加し、ユーザの意思を上位の意思サーバに送る
- 各組織の意思サーバでは、中間サーバから受けとった意思に自分のレベルでの意思/政策を追加し、ネットワークレベルの意思をつくる
- ポリシルーティングサーバは、発信元と宛先の意思を得、持っている各域の政策、接続情報をもとに最適経路を計算する
- その結果をもとに、通信を行なう

このようにすると、QoS ベースの経路制御の特徴と政治色を盛り込んだプリファレンスによる経路制御を組み合わせた経路制御をおこなうことができる。

3.4　接点モデル

次に、前節で述べた要求を満たすような経路制御を取り扱うモデルを考える。

一般的には、網と網はその間を二点間リンクや IX と呼ばれる接続点によって接続される。複数の経路があるネットワークの例を図 3.5の左側に示す。

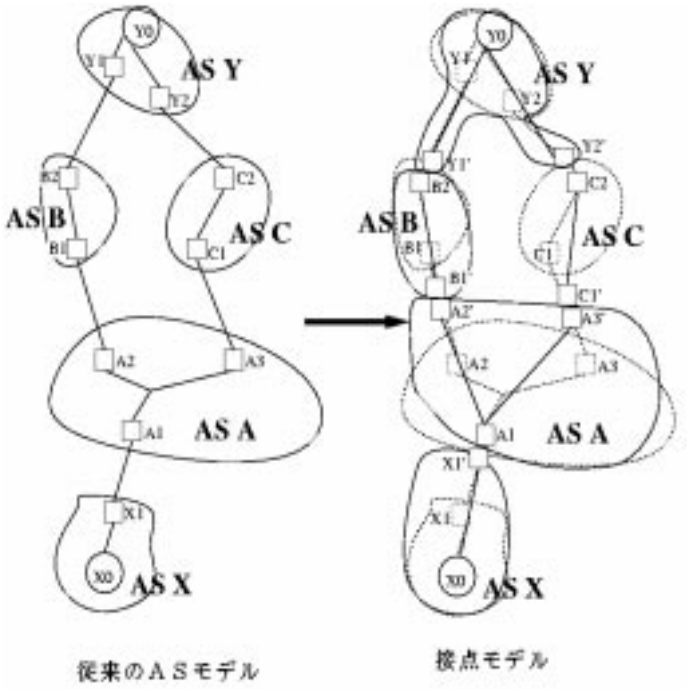

図 3.5: 接点モデル

3.4.1 接点モデル

ここで、網と網の間の二点間リンクにはコストがないというモデルを考える。つまり、図3.5左の A1-B1 リンクのどこかに網の境界があり、そこで網が接していると考える。そして、その接点を新しく網の出入口と考える。

なぜなら、遠くの組織からみるとある組織と別の組織の間のリンクはほとんど見えず、組織と組織が接続していることが主に意味を持つからである。

そのようにして簡略化すると、図3.5の右側のような簡略化されたグラフとなる。各網 (AS) は網を通過するときの情報を、任意の接点間の情報としてだけ与えればよい。網単位に経路制御を行なうため、各網は一度のみ通る。

3.4.2 接点モデルでの IX

一つのサブネットに複数の網が接続し、網間の接続を行なっている場所を IX(Internet eXchange) という。

IX で物理的に接続していても、経路情報をやりとりしていない場合は接続していないのと同じであるので、網と網の接続が複数あると考えるのが妥当である。そのため、IX のモデルは経路情報交換を行なっている網間の接点複数に分割される。

図 3.6に 3 つの網が相互接続している IX のモデル化の例を示す。

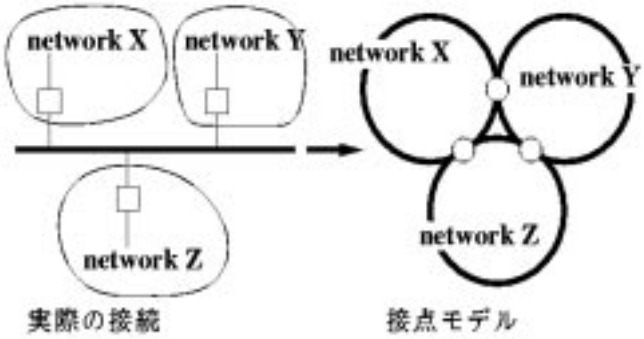

図 3.6: 接点モデル (IX)

3.4.3　最適経路の導出とコスト モデル

最適経路を求めるには、可能な経路の中からコストが最小のものを選ぶという方法をとる。つまり、リンク i のコストを Ci とし、可能な経路の集合 Routes の要素のうちで、全体のコストを最小にするものを選ぶ。

$$\text{最適経路} = R: \ \underset{R \in Routes}{\text{minimum}} \sum_{i \in R} C_i$$

接点モデルで与えられる網の経路の政策、各リンク QoS 情報を A とすると、リンク i のパラメータ A_iは次のように与えられる。

$$A_i = QoS_i + Policy_{AS(i)}$$

QoS パラメータの部分には、遅延、スループット、MTU、セキュリティパラメータ、金銭的コストなどが含まれる。 次に、始点・終点の意思であるが、全体として抑えたいコストを Bとし、プリファレンスによる経路制御でのプリファレンスパラメータを β とし、通したい網、通したくない網を指定する。

そうすると、各リンクのコストは以下の式で与えられる。

$$Ci = f(A_i, B, \beta_i)$$

このように定義すると、最適経路制御問題を最短経路問題に帰結できる。函数 f にて発信元、宛先アドレスやユーザクラスなどによる政策のチェックを行なっていると考えている。

3.5 むすび

本章では、 経路制御を使用制御の観点から考えた政策的経路制御において実施されたい具体的な方針のひとつを定義した。それは、「ある点からみて好ましい域をできるだけ通って宛先に行ってほしい/来て欲しい」という発信元と宛先のそれぞれの要求を、 その間に介在する中間域の通過トラヒックに対する経路制御の方針の下に、 如何に実現できるかである。 この問題を「複数経路選択問題」と呼び、 解決のための経路制御を Routing by Preference と呼んだ。

解決のためには、発信元、宛先の要求を関係域に伝えなければならない。これは、パケット転送時や経路情報配布時に行なわれるであろう。

また、できるだけ一般的な政策的経路制御について述べ、政策的経路制御を行なうのに適したモデル化と、その場合の経路制御について述べた。

今後は、経路情報として交換する情報の決定、情報交換の方法、始点・終点の意思の認証、始点・終点間での意思の統一などの問題を解決し、経路制御に必要なコストの算出を行ない、実際に動くモデルを作る必要がある。

そのために、まず、政策 (ポリシ) 記述言語を設計する必要があるだろう。

第 4 章

関連プロトコル

4.1 IDPR

4.1.1 IDPR の概要

IDPR(Inter-Domain Routing Protocol) は、ネットワークを AD という単位に分割し、Source AD のポリシと Transit AD のポリシを合わせて相手先までの経路を生成し、その経路を使用して通信を行うというプロトコルである。

この **AD**(Administrative Domain) とは、ある特定のポリシの元に単一の管理主体が一元的に網資源の運用・管理を行っているネットワークの集合体を指している。IDPR は基本的にこの AD を単位に経路情報の収集や伝搬、Source AD から Destination AD までの経路情報 (PR:Policy Route) の生成、実際の経路 (Path) の生成などを行っている。

IDPR による経路制御は、各 AD が管理・運用する様々な網資源をどのような形で他の AD に利用させるのか、という事をベースに考えられている。この事は他の AD が自 AD の所有する網資源を利用する際のアクセス制御をその網資源を所有する AD が積極的に行うということを意味する。そこで各 AD は自 AD の網資源利用に関するポリシを相互に交換する事によって、経路設定のベースとなる情報をそれぞれに持つ。通常各 AD が交換するのは自分が Transit AD となる場合のポリシ (Transit Policy) である。IDPR では自分が経路の終端になる場合のポリシは特に考慮されておらず、自分が Source AD になる場合のポリシと Transit AD になる場合のポリシについてだけ注目されている。自 AD のポリシを反映させた網資源へのアクセス制御の具体的な手法としては、自 AD の Transit Policy や隣接 AD とのリンクの状態などから経路制御の基盤となる経路情報を生成し、他 AD に配布する。この際に生成された経路情報を配布する AD を制限する事によって、自 AD の網資源を使用する AD を制限できるのである。隣接 AD 同士は VG(Virutal Gateway) によって接続されている。この VG は隣接 AD 同士を結ぶ仮想的なリンクで、PG(Policy Gateway) の集合である。PG は VG 内に存在する物理的ゲートウェイで、IDPR の処理を実際に担当する。

そして IDPR の様々な機能を処理するのが下記のモジュールである。

- Path Agent
  経路情報の選択、PR の情報に基づく Path(実際の経路) の生成および管理などを

行う。

- Route Server
  収集した経路情報と Source AD のポリシ情報から PR を生成する。また既に生成された PR のキャッシュを維持・管理する。

- Mapping Server
  IP address と FQDN を AD の ID にマッピングする。将来的には DNS の技術と統合されることが考えられている。

- Configuration Server
  IDPR で使用される様々な情報のデータベース。

IDPR において「ポリシ」という言葉の持つ意味は、「自 AD の網資源を他の AD が利用する際の制限、方針」という事になる。そのためネットワークにおける網資源予約 (Resource Reservation)、フロー制御 (Flow Control) といった事を含めた網資源割当 (Resource Allocation) のメカニズムを提供することも IDPR の役割の一つとなっている。

4.1.2 IDPR の最近の動向

IDPR の元々のアイデアは、IETF の Open Routing ワーキンググループによるものであり、後にこれが IDPR ワーキンググループへと名前を変え、現在に至っている。IDPR ワーキンググループでは IDPR のプロトコルの設計や各種 Internet Draft や RFC の作成、gated への実装といったことを中心に活動を行ってきていたが、1993 年 11 月にヒューストンで行われた IETF におけるセッションを最後にワーキンググループとしての活動は中断されている[1]。これは IDPR Version1 の設計及びワーキンググループスタート時に想定されていた gated への IDPR プロトタイプの実装がある程度完成した事から、当初の目的は達成されたという判断がなされた事による。今後は IDPR Version1 の実装と IDPR Version2 の Protocol Standard が発表され次第、再度ワーキンググループの活動を再開する予定になっている。

4.1.3 IDPR の設計とその実装例について

ここでは gated へ IDPR を実装した例を紹介する。IETF での議論によって設計された IDPR プロトタイプの gated への実装を DCA[2]や BBN などのメンバーが中心となって行ったもので、これが現在唯一公式に入手できる IDPR の実装として知られている。IETF の IDPR ワーキンググループによる Internet Draft、"*Inter-Domain Policy Routing Protocol Specification*[3]" をベースとしているが、この Internet Draft で求められている仕

[1] 1994 年 4 月現在
[2] Defense Communication Agency
[3] 現在では RFC1479 として公開されている。

様と実際の実装ではいくつか異なる点がある。まずこの Internet Draft に記述されている IDPR の仕様の主なものを列挙する。

- CMTP (Control Message Transport Protocol)
  IDPR で扱われる全てのコントロールメッセージを扱うためのプロトコルで、この CMTP を使って VGP や RID などのプロトコルが送受される。CMTP はこれらのプロトコルの下位層のプロトコルとして機能する。

- VGP (Virtual Gateway Protocol)
  AD 内の各 PG が VG 間あるいは VG 内に存在する PG に対して経路情報の収集・配布を行うときに使われるプロトコルである。各 PG はこの VGP によって収集された VG に対する経路情報を元にして隣接 AD への到達性を検証する。

- RID (Routing Information Distribution)
  IDPR に関係する各 AD が Route Server に対して自 AD の経路情報を生成・配布する仕組みのことである。IDPR の経路情報には指定された AD を経由する際の Transit Policy や隣接 AD に対する VG の到達性の情報が含まれている。Route Server はこれらの情報を元に PR の生成を行う。

- RSQP (Route Server Query Protocol)
  同じ AD に属する Policy Gateway や他の AD の Route Server がその Route Server が保持している経路情報のデータベースに対して問い合わせを行うためのプロトコルである。各 AD の Policy Gateway はこの RSQP による問い合わせの結果を元に経路の選択を行う。

- RTGEN (Route Generation)
  与えられた発信元 AD と中継 AD の情報を元に Policy Server が PR を生成するための仕組みで、Source AD のポリシと要求する UC (User Class) や TOS (Type Of Service)、Transit AD のポリシと提供される TOS などに関する情報を元に PR の生成を行う。

- PCP (Path Control Protocol)
  PR 情報を元に生成された Path のセットアップや維持管理を行うためのプロトコルである。Path Agent と Route Server で PR と Path に関する情報を交換するために使われる。

- IDPR message の encapsulation
  IDPR message にはユーザデータと PG や Route Server によって生成され IDPR entity の間で交換されるコントロールデータの 2 種類がある。IDPR では IDPR と関係のないゲートウェイやホストの経路制御に影響を与えないように、IDPR のコントロールデータの送受信には IP encapsulation(トンネリング) を用いている。

ここで紹介する gated への IDPR の実装では、概ねこれらの各機能に対する設計を満たす形で行われているが、VGP がリンクの Up/Down をチェックするために用いている sliding window の実装や VGP が各 PG の保持する Policy 関連データを扱う部分の実装が省略ないしは簡略化されている。これは可能な限り機能をシンプルにすることによって本質的ではない部分の実装の手間を省く為であるということが付属のドキュメントやソースコードの中に記述されている。

その他 CMTP における ACK/NAK の扱いや RID における経路情報の更新を行う部分などについても実装を行ったスタッフ独自の判断により、プロトコルの仕様を変更する形での実装となっている。

基本的な実装は gated に対して独立した各モジュールを付加することによって実現されており、IDPR 独自の設定ファイルを解釈するためのモジュールや IDPR に関わる様々な情報を維持・管理する為のデータベースなども実装されている。

この他にも gated に対しての実装と共に UNIX のカーネルに対しても IDPR をサポートするための実装が行われている。主な特徴を列記しておく。なお gated とカーネルモジュールの間では ioctl のインタフェイスを使った制御が数多く行われている。

- IDPR コントロールメッセージの送受信には IP over IP のトンネリングを用いる
- IDPR に関する各種操作を行うため関数や PR や host table を制御するための関数を system call としてカーネル中に実装している
- 各種データの情報を検索するために radix tree のアルゴリズムを用いた検索ルーチンが実装されている。
  - セットアップされた Path にユニークに付けられた ID の検索
  - PR 中に存在するホスト情報の検索 (source, destination アドレスのみ)

4.2 SDR(Source Demand Routing)

昨年度の報告書では SDRP と書いたが、IETF のワーキンググループ名が SDR なので、今年度は SDR とする。この節では、IETF の SDR ワーキンググループの目的や活動予定を紹介し、プロトコルの昨年度からの差を説明したあと、インプリメンテーションを紹介する。

4.2.1 IETF SDR ワーキンググループ

SDR ワーキンググループは、IDRP や BGP というドメイン間ルーティングプロトコルと関連したドメイン間のルーティングプロトコルとしての SDRP(Source Demand Routing Protocol) の指定と利用推進のために認可された。SDR の目的は BGP/IDRP で提供されたノード選択を補い、パケット発信者による経路選択を支援することである。

SDR ワーキンググループの目的は、IETF プロトタイプとしての SDR を公表し、Internet での運用経験を得ることである。SDR について十分な経験ができれば、SDR を標準化のために IESG に提出する。

SDR は 4 つの要素からなる。

- プロトコル制御メッセージとユーザデータグラムのカプセル化のパケットフォーマット
- ユーザのデータと制御メッセージの処理と転送
- ルーティング情報の配布と収集
- 設定と使用

ワーキンググループは以下のような予定で活動する。

1. できるだけ早い時期にあらかじめきめられたルーティングを行うのに使えるように、パケットフォーマット、パケットの生成、処理、転送アルゴリズムを決める。ドメイン間の経路は静的に決めるが、SDR 経路間のドメイン内の経路はドメイン内の経路制御プロトコルに依存する。カプセル化による MTU や ICMP、パフォーマンスなどの問題については、広めるまでに十分に評価しておく。
2. 動的なドメイン間の接続情報の収集と、その情報から経路を組み立てる簡単な案を開発
3. 1,2 と並行して、使用法についての文書の作成と、静的経路の SDR、簡単な動的な経路変化に対応した SDR の有用性を説明するプロトタイプの作成
4. 簡単なモデルの評価が終ったら、第二段階の経路情報交換、経路生成法の開発
5. SDR に帯域予約機構を組み込み multicast と組み合わせた用途の研究
6. SDR のセキュリティオプションに関する研究
7. SDR-WG と BGP/IDRP-WG との調整

93 年にパケットフォーマット、転送アルゴリズム、プロトコルについてのドラフトを公開した。[93]

93 年に Route Setup に関するドラフトを提出した。

95 年末にはプロトコルを IESG に提出する予定らしい。

4.2.2　SDR パケットフォーマットとユーザデータの処理

SDR のパケットフォーマット、パケットフォーワーディングについては、昨年度の報告書での調査と比べ、ほとんど変わっていない。ソースルートとして IP address を書くことができるようになったこと、Route Setup のパケットフォーマット、フォーワーディングについて規定されたことぐらいである。[93]

4.2.3 SDR インプリメンテーション

ftp://catarina.usc.edu/pub/sdrp に SDR の SunOS4.1.3 へのインプリメンテーションがある。kernel の ip_input, ip_output に処理を追加し、ソースルートテーブル、DFIB 変更のシステムコールを追加している。

このコードを使って、次の節で説明するようなトンネリングを実現することができる。

4.3 BGP/IDRP の拡張

IBM の Yakov Rekhter により、hop by hop routing のままで第二章で説明した金魚モデル問題を解決する提案が行なわれている。[96] [97]

第一は、ある組織が図 2.6のように複数のネットワークと接続しているときに、組織内の複数のマシンがそれぞれ別の運用方針を持つ場合の解決法で、組織の一般的な運用方針と異なるマシンの場合はトンネリングを使用するという方法である。その場合の経路のバックアップなども考慮されている。基本的にはその組織の界面ルータが賢く、組織内のマシンは界面ルータまで SDRP などのフォーマットでトンネリングを行なうという方法である。片方のリンクが切れた場合は別のリンクを使うといった政策も考慮されている。[96]

第二は、バックボーン共同運用の場合と似ているが、図 4.1 のようなネットワーク構成で、C B E がサービスプロバイダ、D F A がサービスサブスクライバとする。通常の BGP/IDRP では、D と F から A への経路はそれらへサービスを提供している C の経路に従うが、それ以外の政策がある場合はトンネリングを使って解決するという方法をとる。ユーザから見た場合にトンネリングを隠すため、仮想ルータや仮想リンクにより、従来の経路制御を拡張する方法も提案されている。

どちらの場合でも、往路の経路指定だけであり、hop by hop routing の問題で復路の経路指定はできないのが問題であるが、SDR や IDPR と違い、現在の実装からの変更点は少なくてもすむ。

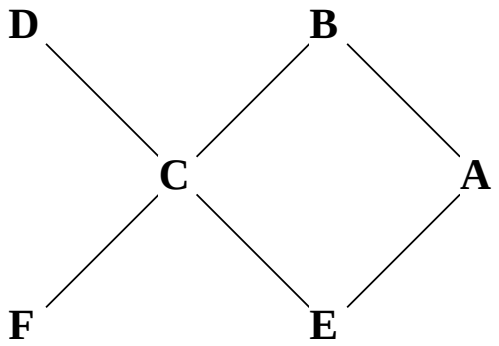

図 4.1: 金魚モデル問題