知的ポリシ制御機構AISLEの提案

明石 修 光来 健一† 福田 健介 廣津 登志夫
佐藤 孝治 丸山 充 菅原 俊治
{akashi,kourai,fukuda,hirotsu,koji,mitsuru,sugawara}@core.ecl.net
NTT 未来ねっと研究所, †東京工業大学

概要

インターネットはASと呼ばれる独立したネットワークの集合体であり，その運用はポリシと呼ばれる各AS毎の管理運用基準に基づいて行う．しかしオープンな系であるインターネットでは，接続形態，トラフィック管理などが複雑化するにつれ，その変化に迅速かつ柔軟に対応することが困難になっている．このような環境に対応するために，より抽象度の高いポリシ記述に基づき，ネットワークから取得する情報や他のASと共有する情報を用いながら知的に自己制御を行うネットワーク環境AISLEを提案する．本論文では，特にAS間の経路制御に焦点をあて，BGP属性情報を利用した協調制御に関して述べる．

AISLE: an Intelligent Routing-policy Control Mechanism

Osamu Akashi, Kenichi Kourai†, Kensuke Fukuda, Toshio Hirotsu,
Koji Sato, Mitsuru Maruyama and Toshiharu Sugawara
NTT Network Innovation Laboratories, †Tokyo Institute of Technology

abstract

Inter-AS routing is difficult to control since advertised BGP routing information changes as it spreads through ASes that are independently managed by each organization based on its policy. For operating the Internet reliably considering environmental changes, we propose flexible and autonomous policy control mechanism AISLE. AISLE analyzes network status information and uses cooperative actions among distributed agents to control inter-AS routing behavior by utilizing BGP attributes.

1 はじめに

インターネットはAS(Autonomous System)と呼ばれる，ISP，企業，大学などのそれぞれ独立した組織が運用するネットワークが相互接続した巨大な分散システムである．インターネット全体の運用は，その構成要素である各AS毎の制御の相互作用の結果である．管理の単位であるASは，ポリシと呼ばれる管理運用基準に基づき，各AS毎のオペレータにより運用される．

ポリシとは，隣接ASとの接続形態や，いかにパケットをフォワードするかなどのトラフィック管理を含む．ポリシは最終的にはルータの設定primitiveに翻訳し，場合によっては複数のルータを制御して，ASとしての挙動を決定する．すなわち，抽象的な表現である管理ポリシに基づいて，宣言的記述である単一ノード上で動作するルータ制御primitiveを用いた制御文を作成し，それに基づいて経路制御を行ないトラフィックを交換するという枠組みである．

しかし，オープンな系であるインターネットが，接続AS数を増やし，接続形態やトラフィック管理が複雑化するにつれ，その管理運用ポリシ自体が複雑化，あるいはポリシ自体も頻繁に変更する必要があり，迅速かつ柔軟に変化に対応することが困難になってきた．例えば，より高度な知的判断処理や，他の観測した環境情報によって，初めて決定できるものの例として，従量制課金によるト

ラフィックの制限や変更が挙げられる．これはあらかじめルータの中に定義することはできないため，アプリケーション層との相互作用による制御が必須となる．

本研究では，トラフィック制御と，管理者によるポリシの間で制御を行う層を定義し，更にその機構が協調して複数の AS で動作することにより自律的に知的自己制御を行うネットワーク環境 (AISLE: Autonomous and Intelligent Self-control Environment)[1] を提案する．AISLE は，独立に運用される AS の集合に適用されるため，AS 間経路障害診断のためのシステム ENCORE[2] と同様のアプローチを採用し，各 AS での制御を担当するエージェントと，そのエージェント間協調により機能を実現する．本稿では，AISLE における AS 間の経路制御に焦点をあて，BGP[3] 情報を用いた協調観測・制御によるネットワークの状態に応じたネットワーク運用と制御に関して述べる．

2 ネットワーク運用モデルの問題

現在のネットワーク管理パラダイムにおける問題は，以下の 2 つの点にある．

1. 制御記述の抽象度
2. 制御対象の範囲

制御記述の抽象度の問題は，ルータの制御記述と抽象的な実体であるポリシの間の隔たりが大きいため，管理運用ポリシと実際のトラフィック情報を考慮してオペレータがルータの primitive に翻訳する際に誤りが発生しやすく，言語記述の上からも，システム上からも，その検証は困難である．更にネットワークの状態に応じた記述方法が存在せず，測定した値に基づた操作や時間的に変動する状況に対応した操作を記述できない．そのためネットワークの状態変化に対応して，動的にトラフィック制御を変更するような柔軟な管理運用も困難である．

制御対象の範囲の問題はネットワークがしばしば “雲” として書かれるように，複数のルータ，LAN，より広域では複数の AS から成る広がりを持つ実体として考えられるが，実際の制御はその内部の構成要素の点であるルータに対してのみ行なわれることである．ルータには経路制御プロトコルやフォワード動作の制御の記述を通して，他のルータとの相互作用に関する動作を記述するが，その動作自体はプロトコルとして定義された範囲をルータの制御 primitive を通じて操作するだけで，複数の AS 間での協調動作を記述することはできない．

2.1 Inter-AS レベルの BGP 運用における問題と要求条件

具体的に，AS 間経路制御のデファクトスタンダードである BGP [3] による運用に注目する．経路制御は，AS 毎の管理ポリシに基づき，AS 間で到達性情報を交換して実現する．管理ポリシとは，自 AS に関するパケットの振舞を定義する抽象的な表現であり，最終的に経路制御を行う IP 層では，BGP レベルの情報として変換されている必要がある．図 1 に広報される BGP 経路情報とパケットの流れを示す．なお，内向きのトラフィック制御に関しては，図 1 に示すように経路情報到達を検証する問題があるが，[2] のアプローチがあり本稿では議論しない．

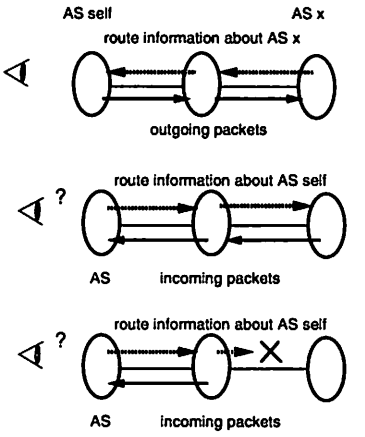

図 1: 広報される BGP 経路情報とパケットの流れ

AS 間経路制御の観点からは，現在は BGP のさまざまな属性パラメータの値を設定して，外向きのトラフィックを静的に制御するが，ネットワークのトラフィックの状態や時間による制御の変更，複数の AS で協調しての，内向を含めたトラフィク制御を行う事ができない．

例えば，国内向けの外向きトラフィックは ISP A 経由を用い，その他は ISP B を用いるというポリシを実現するために，正規表現等を用いた記述で国内から広報された AS 番号をオリジンに持つ BGP 経路情報に対して，優先度が高くなるように BGP 属性情報を操作する．BGP 経路選択ルールは RFC で厳密にあらかじめ定義されているが，AS を管理するオペレータが意図するポリシから，ルータの設定，あるいは，BGP レベルの表現に変換する際には，さまざまな誤りが混入する可能性があるという問題がある．

また，これらの BGP レベルの操作が定義された後では，与えられたポリシ条件を満たす範囲で，環

境の変化に応じて，ルータの設定やBGP属性などを適応的には変更できない．言い換えると，BGPレベルの属性情報を充分に使いこなせていないと言える．これは，ポリシは抽象的な表現であり，実際の制御を記述するルータの設定記述との間には大きな隔たりがあることに由来する．

一方，内向きのトラフィックに関しては，自ASの外側のASでどのように扱われるかは外部ASのポリシに依存し，図2に示すようなポリシの不整合が生じることがある．この例では，AS_xは，帯域の大きなリンクを用いてAS_iと，それより帯域の小さなリンクを用いてAS_jとマルチホーム接続をしているとする．AS_xの広報した経路情報は，マルチホーム接続先であるAS_iとAS_jをそれぞれ経由して，トランジットASであるAS_kに到達する．AS_kでは，AS_xに対して，それぞれ(AS_i, AS_x)と(AS_j, AS_x)というASパスを持つ2つのBGPエントリがある．この例では，AS_kは自分自身のポリシに基づき，(AS_j, AS_x)をベストパスとして選択したとする．その結果，AS_yではAS_xに対して，(AS_k, AS_j, AS_x)のASパスを持つBGPエントリが受信される．

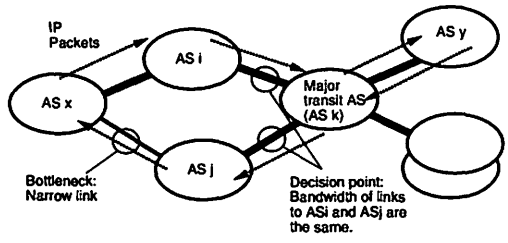

図 2: ポリシ不整合の例

AS_xがAS_yにアクセスする場合，AS_x自身のポリシにしたがって，AS_iにパケットをフォワードすることは可能である．しかし，戻りのパケットはAS_kのポリシに従い，帯域の小さなリンクを経由するAS_jを通る．このような不整合の解消のためには，ASの外側でBGP情報を操作する必要があり，現在の枠組では困難である．ASは管理の単位であるため，外部ASが特定の意図を持って設定したポリシは優先されるのが自然であるが，特に外部AS側で特定の目的も持たずに設定したポリシと自ASの意図との不整合は，外部ASに情報を与え，協調作業により不整合を解消するのが望ましい．

また，自らが広報した経路が相手ASでどのようなポリシで扱われて，経路選択ルール適用後にインターネット中に広報されるかは，あらかじめ調べることは困難である．それらは，実際の運用環境でパケットを流してみないと検証ができない．例えば，あるマルチホームするASが，新たに大きなトランジットASに接続を予定している場合，BGPのセッションを実際に確立した後でないと，その他のBGP peerとの関係で，自らの意図通りにパケットが流れるかどうか検証できない．更に，BGPセッション確立後でも，他から広報した経路が優先していると，その経路はそのAS外からは観測不能であり，障害時に意図通りの経路の切り替えが行なわれるのかどうかも外部からは検証できない．

3 ポリシ制御機構 AISLE

3.1 アーキテクチャ

2章で述べた問題を解決するために，ポリシ制御機構AISLEを提案する．AISLEは図3に示すように，AS内での制御を行う層と，AS間における制御を行う層から構成する[1]．前者はオペレータの意図するポリシとルータprimitiveの中間に位置し，与えられたポリシ記述に基づき，ネットワークの状態を観測しAS内のボーダルータを制御する機能から成り，後者はAS内の制御機構間の協調動作を通じて構成する．このようなアーキテクチャを用いるのは，AS間診断システムENCORE[2]と同様に，ASが独立した管理の単位であり，集中制御の枠組みの適用は困難であることによる．

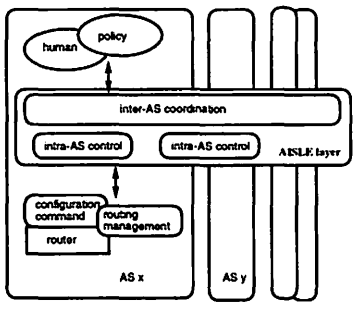

図 3: AISLE システムの構成

システム構成としては，図4に示すように複数ASに分散配置するマルチエージェントモデルを用いる．エージェントは，環境に応じた適用やポリシの時間的な変化などの，より高度な記述を解釈実行する主体であり，必要に応じてマルチエージェント間協調を用いることにより，自ASや外部ASの情報を得て設定を動的に変える枠組みを与え，内向き/外向きのトラフィックの制御を可能とする．

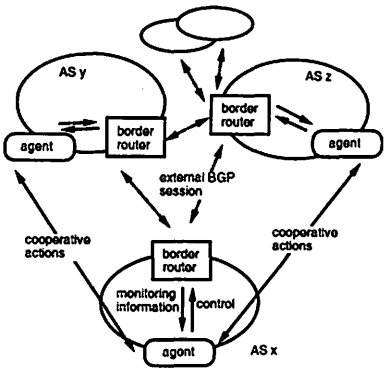

図 4: AISLE エージェントモデル

3.2 エージェント機能に対する要求条件

以下に，AISLE の機能を実現するための，エージェントが持つ機能に対する要求条件を示す．

- ネットワーク層からの情報の取得
  - ルータ (snmp, rsh)，解析ツールによるネットワーク情報の取得．既存の解析ツールに加えて，プログラマブルなパケット解析システム [4] から情報を得ることにより，ネットワーク上のアプリケーション層の機能に依存した情報を得る．
  - エージェント間協調による他の AS からの情報を取得，共有する．
- ネットワークの状態に応じた制御
  - ボーダルータの持つ BGP 経路情報の属性値を変え，AS 間経路の切替を行う．特殊なハードウェアを仮定せずに実装する (4 章参照)．
  - ネットワーク管理者の視点からの制御のみでなく，ネットワークとアプリケーションの協調動作のフレームワーク [5] に組み入れ，より柔軟な環境を構築する．
- オペレータからの情報の反映
  - 単一 AS の視点から見た AS 間制御ポリシの記述．今回は，同一 AS 内の複数ボーダルータ間の制御を必要とするような，AS 内で複数エージェントを用いる協調は扱わない．
  - 複数 AS 間でのポリシ記述．自分の視点から記述する他の AS への要求ベースとする．複数の AS をまたがる大域的な視点での記述は今回は扱わない．
- オペレータへのネットワーク情報の提示
  - オペレータが現在のネットワークの状況を理解し，エージェントに与える制御ポリシの設定，修正に用いるため，エージェントから，自 AS 内，AS 間の情報を提示．

ネットワークの状態変化に応じた，動的かつ柔軟な管理運用の観点からは，ネットワークの状態取得が必要である．ネットワークのモニタに関しては，ボーダルータからの，SNMP，あるいは rsh を用いてのデータ取得や，解析ツールを用いて行う．しかし DNS などのネットワーク層の機能と密接なつながりがあるが，アプリケーション層の機能として実現されているものは，ツールをその度に作成・修正しながら対応していく必要があるが，これはプログラマブルなパケット解析システム [4] と組合わせて対応する．

また，管理ポリシからそれを実際のトラフィック管理に反映させるために，抽象的な表現であるポリシを表現し，収集したネットワークの状態に応じて動的に制御を変更することが必要である．ポリシの記述には，従来のルータのトラフィック制御の部分に加えて，動的に変更する部分の記述が必要である．

3.3 ポリシと制御

ネットワークの状態に応じたポリシを記述するためには，必要な機能を抜き出す必要がある．そのためには実ネットワークでの適用実験を行う．今回の実装では，専用のポリシ記述言語を設計するのではなく，基本機能を Lisp で実装した ENCORE の機能を拡張する形で，S 式の表現を用いて記述するアプローチを取る．

ポリシ記述の中に，以下を実装する．

- 動的なパラメータは，ポリシ記述の変数として記述し，参照時に実値を束縛する．if 文などの条件分岐構造と組合せることにより，ポリシーを変更可能にする．またこの制御記述は，ENCORE の観測戦略記述を利用し実行する時間間隔を直接指定するか，他の観測動作からのイベントにより間接的に起動する．
- ポリシ記述が与えられた動作中のエージェントに対する問い合わせ機能
  - ローカルなチェック機構．例えば，ある source IP address と port の組を与えて，特定のインタフェースから入力した場合のフィルタの動作状況．
  - AS 間のエージェント協調を通じた検証機構．例えば，自 AS 宛の内向きのトラフィックがどの AS 経由であるかを問い合わせる機構．

- ポリシ記述の分離．すなわち自 AS のポリシとして変更不可能な部分と，外部に対して開放する部分を分けて記述する．

アクティブネットなどのプログラムを送り込んだり，遠隔手続き呼出しなどを用いる集中制御システムとの大きな違いとしては，preference 交換を用いるマルチエージェントによる制御であることであり，あくまで主権は受信側エージェントにある．

3.4 適用例

時間的に変動する値を用いた動的ポリシ制御の適用例を示す．AS 内での外向きトラフィックの制御に関しては，以下の例が挙げられる．

- リンク間のロードバランス ... IGP コスト値を利用したパケット単位の細かい分散ではなく，より大きなタイムスケールを対象とする．例えば，一日の中での，あるいは週の中での時間帯といった周期も扱う．
- 帯域の狭いバックアップ経路の障害時の広報 ... ルータの他のインタフェースの up/down 情報，トラフィックの状況を考慮して，経路制御プロトコルの制御パラメータを変える．
- 従量制課金の制約による，トラフィックの変更 ... 月内での観測したトラフィック平均累積値による，通信プロトコルとは関係しない課金制約に応じた振り分け．

次に，AS 間での協調が必要なポリシへの適用例を挙げる．

- 内向きパケットの経路の制御 ... 外向きトラフィックの制御と類似であるが，エージェント間の協調を通じて行う．外向きより，より大きなタイムスケールで行うことが想定される．
- 上の例は，自分の上流がどの AS であるかの大域的なトポロジ情報を得た上で行うが，自 AS からの観測情報ではトラフィックの大域的な振舞が不明な場合も，エージェント間の協調を用いることで適応可能となる．例えば，隣接する AS に対して多段で，altenative path があれば自 AS に対する next_hop を分散するように依頼を送ればよい．ただし，全 AS に拡散しないように，major な transit AS で止めるロジックを入れる，あるいは同時に hop 数制限をつけるなどの制御が必要である．
- 攻撃に対する防御 ... AS 内で行うシステムはあるが，エージェント間で情報を共有することにより，動的にフィルタを設定することが可能となる．
- AS 間の動作検証
  - BGP 情報の共有やポリシ記述を用いた，経路の広報前やベストパスでない経路の検証
  - 定常的な相互ポリシ監視

以下に，内向き，外向きのトラフィックを例にした制御記述を示す．外向きトラフィックの例では，エージェントが自 AS 内のボーダルータに対して，観測動作を行ない，その結果を基づきパケットフォワード先を調整する．

```
;;; Distribute outbound traffic evenly
;;; in number of destination routes.
(def-strategy check-out-traffic
  (:interval (* 60 10)) ;;; 10 [min]
  (:inhibit-interval (* 60 20))
   ;;; Inhibit interval against any changes
  (rule distribute-out-bound) )

(def-rule distribute-out-bound
  (acq get-balance-info)
  (eval balance-next-hop-in-number
        (acq-result) ))
```

一方，図 2 のような内向きトラフィック制御の場合，その調整は自 AS 外部で行う必要があり，外部 AS で自 AS に対するフォワード先操作が必要である．しかし，外部 AS は他の組織による運営であるため，直接制御することできない．このため外部の AS(図 2 の例では AS_k) に対して要求を送る必要がある．このやりとりは，RPC のような手続き呼び出しでなく，あくまで自 AS の送る経路制御に関する希望という抽象的要求の送付とし，受け取る側でそれに相反するポリシ記述がない場合に実行される，という枠組みで要求される．

```
;;; Distribute inbound traffic
(def-sp-var up-stream-AS-list
  (get-initial-list) )
(def-strategy check-up-agents ...)
;;;
(def-sp-var in-pref '(ASi ASj))
(def-strategy modify-in-pref-list
  (:interval 3600) ... )
;;;
(def-strategy check-in-traffic
  (:interval (* 3600 6)) ;;; 6[hour]
  (rule change-in-bound
     :every target-AS up-stream-AS-list ))

(def-rule change-in-bounnd
  (acq set-next-AS-preference-if-possible
    (in-pref)
    :cooperative target-AS)
  (eval report-result (acq-result)) )
```

この例では，自 AS が送った要求と，外部 AS の行動指針の間で不整合が起きないが，AS_k が AS_j をフォワード先としてメインに使うというポリシが陽に定義されていた場合は実行されない．

また，AS_x と AS_j にマルチホームする AS_z があり，帯域の大きな AS_j 経由でパケットを受け取りたいとする．この時，AS_x が下流へのトランジットに対しても同様の preference を送っていた場合，双方のポリシが相反する．このような場合，(要求元 AS に関するポリシ) > (要求元 AS の下流 AS に関するポリシ) > (その他の AS に関するポリシ)，と優先度を定めれば解決可能である．

外部に送付するポリシには有効期限をつけ，この継続には，エージェント間での更新処理を必要とする．これは厳密な状態を維持するモデルは，接続性の保証のない環境のため仮定しないが，受理された preference はファイルに保持しておき，再立ち上げ時には，有効期限内であれば再度そのポリシの適用を行い，不必要な振動を避け継続性を保つことができる．

4 実装

ボーダルータは，受信した BGP の経路情報に対して，定義される BGP 属性情報 (`next_hop`, `local_pref`, `weight`(Cisco ルータ) を操作し，その結果を BGP 経路選択ルールにあてはめる事により，経路選択を行う．AISLE エージェントは，図 5 に示すように，対応する属性値を変更し，外向きのトラフィックを制御する．操作対象は BGP 属性値であり，BGP のベストパス選択ルールは既存の枠組みで行うため，従来と同様ルーティングループが起こらない実装である．

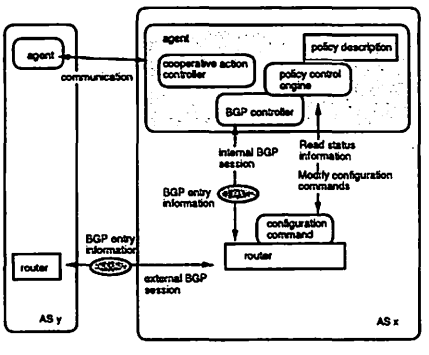

図 5: BGP 属性値の操作

`bgp-controller` は，ボーダルータと iBGP 接続して BGP 経路を受信し，AISLE ポリシに従って `next_hop` や `local_pref` 等の属性値を調整した updata メッセージを送る．ただし従来の iBGP と同様に段数が増える分，経路の変更に対して収束が遅くなり，ボーダルータと一貫性が保たれない期間が発生する問題がある．例えば，ボーダルータがある経路の `withdraw` を受信した時に即時対応ができず，収束するまで経路の矛盾が起きてしまう．

`policy-control-engine` は，直接ルータの設定 primitive を rsh で送り込み，設定を変更することが可能であるが，rsh ではトランザクションとしての実行が厳密には保証されないため，マスタファイル更新後，ルータの `reload` コマンド実行する方法と使い分ける必要がある．

他の AS のエージェントから送られたポリシ変更の要求は，`policy-control-engine` がチェックを行い，それを受理した場合に，外向きのトラフィック制御と同様の処理を `bgp-controller` を通じて行う．

5 おわりに

ネットワークから取得する情報や他の AS と共有する情報を用いながら知的に自己制御を行うネットワーク環境 AISLE を提案した．今後は実ネットワークでの実験を通じた機能の検証，修正，拡張を行う予定である．また AS 内で複数のボーダルータ間の協調を行うシステムとの連係や，コンパクトなポリシ記述言語の設計を進めていく予定である．

参考文献

[1] O. Akashi, K. Kourai, K. Sato, T. Hirotsu, M. Maruyama, and T. Sugawara. "Agents Support for Flexible Inter-AS Policy Control". In *SAINT'03 Workshops / AI on the Internet*, pages 294–298. IEEE / IPSJ, Jan 2003.

[2] O. Akashi, T. Sugawara, K. Murakami, M. Maruyama, and N. Takahashi. "Multiagent-based Cooperative Inter-AS Diagnosis in ENCORE". In *IEEE/IFIP Network Operations and Management Symposium*, pages 521 – 534, Apr 2000.

[3] Y. Rekhter and T. Li. "A Border Gateway Protocol 4 (BGP-4)", 1995. RFC1771.

[4] 菅原, 明石, 廣津, and 佐藤. "プログラマブルなパケット解析システムの提案". In 第 *90* 回システムソフトウェアとオペレーティングシステム研究会. 情報処理学会, June 2002.

[5] 佐藤, 廣津, 福田, 明石, 山崎, and 菅原. "Clea におけるアプリケーションとネットワークのための協調メカニズム". In プログラミングおよび応用のシステムに関するワークショップ *(SPA02)*. 日本ソフトウェア科学会, Mar 2002.