

リフレクタモデルに基づくインターネット管理アーキテクチャ

2Aa-4

村上 健一郎

NTT ソフトウェア研究所

1 はじめに

インターネットのバックボーンでは、ネットワークの管理単位である自治システム AS (Autonomous System) が、BGP (Border Gateway Protocol) によって隣接する AS から得た経路情報を自分の経路制御のポリシーに従って操作して次の AS へ伝える。エンドノードからのパケットは、この経路情報をもとに各 AS で経路が決定されて中継されるが、その経路は経路情報の伝わる向きとは逆である。このため、自 AS へのパケットの中継に問題があっても自 AS ではそれがわからず、自 AS の経路情報が他 AS でどのように認識されているかを知らなければならない。ところが、各 AS は独立した組織によって管理されているため、現状のアーキテクチャでは、他 AS の中の自 AS に関する経路情報を覗くことができない。そこで、本論文では、各 AS が他 AS から経路情報をアクセスできるリフレクタを持つ経路制御管理モデルを提案し、これを実現するアーキテクチャの概要を説明する。

2 現状の経路管理の問題

世界のインターネットバックボーンでは、経路数が3万近くにもものぼっている。しかも、インターネットが現在成長期にあるため、日々かなりの経路情報が変わる。これには、まず長期にわたる成長過程に関する変動がある。また、回線の障害、ルータの障害、ルータの設定誤り等によって短期の大きな変動も発生する。これに加え、ポリシーの設計の誤りもある。

このような経路制御には世界レベルの多数のプロバイダが関与するため、経路制御の現状を把握したり、障害の原因を特定して回復作業を行うことが困難になっている。更に、障害からの回復にあたっては、複数のプロバイダのオペレータ間でコーディネーションが必要になる。それは、自分の経路の情報がどのように伝わっているかが自分で確認できないからである。ここでのコーディネーションとは、相互に、自分の姿を写す鏡の役割を相手のプロバイダへ求めることと等価である。つまり、これまででは、障害の原因を特定するためには、他 AS のオペレータに、そこにあるルータの BGP の経路表から自分に関する経路情報をダンプしてメールや電話で教えてもらう等の作業を行っていた。しかし、これでは、インターネットが世界レベルのネットワークであることも手伝って、相手が必ず必要な時に作業を行ってくれる保証がなく、障害からの回復に時間がかかってしまう。

3 リフレクタモデル

本論文では、これまで説明してきた問題を解決できるリフレクタモデルを提案する。リフレクタは、いくつかの主要な AS やプロバイダが相互接続する接続点 NAP (Network Access Point) などの完全な経路情報を保持するルータがある地点に設置され、BGP メッセージの履歴や経路情報を保持するエージェントである。そして、自 AS の経路情報が他 AS でどのように認識されているかを見るための鏡の役目をする。複数のリフレクタを置けば、例えば、経路制御の障害によって直接アクセスできない場所でも、いくつかのリフレクタで反射させることによって間接的に情報を得ることができる (図1)。また、多地点から反射させた情報 (マルチプルビュー) を総合することによって、障害の原因などを絞り込むこともできる。リフレクタは経路情報の更新履歴や経路表のスナップショットをデータベースとして持っており、問い合わせ処理によってしばらくこんだ情報だけをオペレータへ送る。これは、ダイナミックに変化する大量のデータを転送することなしに処理を行うとともに、リアルタイムに近い情報を処理できるようにするためである。

各リフレクタはオペレータを通じて協調し、経路を把握するための機能を提供する。その意味で、これはマルチエージェントシステムとなっている。複数のリフレクタを使用することにより、自 AS に関する経路の状態を世界レベルで把握することができる。しかし、完全な状態の把握には、複数のリフレクタを使いこなす知識や、経路制御のさまざまな情報が必要となる。このため、オペレータの手元には、リフレクタとのインタラクションを行ったり、ユーザのインタフェースを提供するアシスタントエージェントが必要となる。

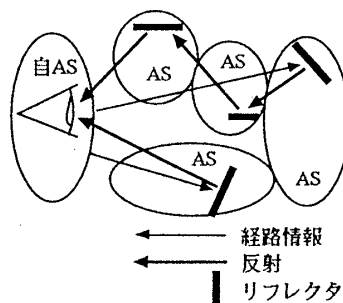


図1: 多重リフレクタ

Internet Routing Management based on Reflector Model

Ken-ichiro Murakami

NTT Software Laboratories

3-9-11, Midori-cho, Musashino-shi

Tokyo, 180 Japan

4 多重リフレクタアーキテクチャ

ここでは、リフレクタに基づく管理アーキテクチャの構成要素および機能を説明する。

4.1 オープンリフレクタエージェント

リフレクタは、NAP などにある完全な経路 (full route) を得ることができる BGP スピーカとセッションを張り経路情報を取得する。しかし、必ずしも物理的な場所は制限されない。例えば、専用線などを使用してレイヤ 2 で別の場所から接続することも考えられる。また、信頼性は下がるものの、マルチホップ BGP や iBGP (Internal BGP) を使用して、BGP スピーカから離れた場所に置くことができる。

リフレクタの機能は、主に次の 3 つに分類することができる。まず、データベース機能では、一定時間ごとの経路表のスナップショット、BGP の更新メッセージ等をその発生時間と共にデータベースに記録する。これらは、オペレータからの要求によって検索され、時間変動などの履歴や統計情報を提供する。2 番目の機能は、リアルタイムインスペクションである。これは、現在の経路表をリアルタイムに検索したり、traceroute や reverse-traceroute (traceroute と異なり、逆経路、即ち、自 AS への経路がわかるツールである。この詳細は別の機会に譲る。) などによって現状の経路を把握する機能である。また、経路の変動などのイベントを検出し、メールや SNMP (Simple Network Management Protocol) のトラップ機能を使用してオペレータへイベントを通知する機能も持つ。最後は、コマンド中継機能である。リフレクタは、障害などによって直接通信できない別のリフレクタへ検索コマンド等を中継する。即ち、代理リフレクタとして動作する。

リフレクタは一般に開放されるが、セキュリティ確保のために 2 つのレベルを用意している。1 つは、バックボーンのエベレータを想定したレベルであり、すべての機能を開放する。これは、各 AS が経路や制御ポリシーを登録する IRR (Internet Routing Registry) に登録されたオペレータを想定しており、デジタル署名による認証を行う。もう一つのレベルは、バックボーンの加入者側が自分の経路情報の状態を把握するためのものである。この場合には、高度な認証は行わないものの限定された機能だけを開放する。

4.2 アシスタントエージェント

オペレータ側には、ユーザインターフェースを持つアシスタントエージェントがあり、リフレクタとのインタラクションだけでなく、SNMP などを使用した情報の集収、さまざまな知識源からの情報の集収などを行う。その情報源には、Web サイトから提供される統計情報、メーリングリスト、ftp サーバから提供される各種統計情報、トラブルチケットの情報なども含まれる。そして、集収したデータから障害の場所や原因を推論する。

アシスタントエージェントは、情報の差分によって経路制御の誤りを見つける。例えば、現状の経路の把握にあたっては、リフレクタへコマンドを送り、現在の BGP の経路表、BGP セッションの統計情報、各経路の AS パスや安定時間、経路の変化などを知る。一方、理想的な経路は IRR から取得する。そして、IRR と現状の経路制御の差分を見れば、オペレーションの誤りを見つけることができる。その障害を発生したネッ

トワークの所有者や管理者などは whois データベースや IRR のデータベースから見つける。

差分に着目する方法は、経路の異常の程度を判断するためにも使用される。つまり、通常の定常状態の経路を監視していれば、各 AS の通常の性質がわかり、現在の状態が許容範囲であるのか、あるいは異常状態にあるのかを、現状との差分で認識することができる。また、データベース処理によって長期的な変動と短期的な変動を分離し、通常の成長過程における一時的な不安定状態と異常とを分離することもできる。なお、これらの性質の把握にあたっては、ネットワークアドレス、AS 番号、経路情報全体といった、マクロ、メソ、ミクロの異なる視点からネットワークの性質を把握することが必要である。これらの検出は、経路制御の知識をもった推論エンジンによって行われる。

5 制御プロトコルとデータベース

リフレクタを利用するためには、アシスタントエージェントとリフレクタ、リフレクタとリフレクタの間のプロトコルが規定されていなければならない。データベースの問い合わせ処理などでは、インタラクティブな操作で状態を持つことが必要となるため、この制御プロトコルではトランスポート層のプロトコルに TCP (Transmission Control Protocol) を使用する。アシスタントあるいはリフレクタは、相手リフレクタへ TCP の確立を行ったあと、コマンドを送る。コマンドは、デバッグの容易さなどからキャラクタベースである。但し、セッションの確立直後に、ユーザの認証を行い、認証に失敗した場合にはコネクションを切断する。なお、アシスタントから直接通信できない相手へのメッセージ中継については、パラメータとして、相手のリフレクタのアドレスも必要である。

リフレクタ内のデータベースには BGP のメッセージが時刻とともに記録される。これで、各経路の up / down というイベントが記録されることになる。しかし、これは差分ベースのメッセージであり、メッセージの矛盾なども想定される。また、初期の経路表も必要である。そこで、一定時間ごとに、BGP の経路表のスナップショットも記録しておく。データベースのタプルは、BGP のメッセージ形式に準じている。即ち、経路アドレス、次のホップ、メトリック、ローカルプリファレンスやウェイト等のタグ、AS パス、経路情報源などである。

アシスタントエージェントは、特定経路や経路全体の変動の統計を要求したり、ある変動の着目した一連のレコードを要求するので、リフレクタはリアルタイムに処理を行って、その結果を返す。なお、アシスタントエージェントは、データベースの検索だけでなく、現在の経路表の統計や検索もリアルタイムに要求する。いくつかのリフレクタへ同時に検索を要求できるので、その結果を総合して、ネットワーク全体でのある時間の経路の様子や変動を把握することも可能である。

6 おわりに

本論文では、各 AS が自 AS の経路をオペレータのコーディネーションなしに把握できるリフレクタモデルを提案した。また、そのアーキテクチャの構成要素と機能を説明した。この方式によって、これまで人手で行ってきた経路情報のオペレーションがマルチエージェントによって自動化され、経路情報の的確な把握と障害からの回復が可能となる。