マルチメディア通信と分散処理 85－28
（1997. 11. 7）

自律系における経路情報の監視方法の提案

森島 直人† 小林 和真† 山口 英† 尾家 祐二‡

†奈良先端科学技術大学院大学
‡九州工業大学/奈良先端科学技術大学院大学

概要

急激なインターネットの拡大にともない，経路の管理にかかる負荷が問題となっている．ネットワークの巨大化および構成の複雑化により制御しなければならない経路が増加したため，問題が発生してもその原因となる箇所の特定が困難になってきたからである．そこで本研究ではOSPF(*Open Shortest Path First*)において交換される経路情報を監視することにより，自律系における経路の異常およびその原因となる情報の発生源を推測し，経路管理にかかる負荷を軽減することを目的としている．

本稿では自律系における経路を監視する手法を提案する．さらにその手法に基づいたシステムを実装して実際のインターネットで運用することにより，提案した手法の有効性について評価を行なった．

A proposal of the method to monitor the routing information in the autonomous system

Naoto Morishima† Kazumasa Kobayashi† Suguru Yamaguchi† Yuji Oie‡

†Nara Institute of Science and Technology
‡Kyushu Institute of Technology/Nara Institute of Science and Technology

Abstract

One of major issues arising in the Internet is related to routing management. Rapid growth of the Internet increases routing information much quickly, thereby making the routing management a tough task. In fact, it will be very difficult to identify which part of the network is troubled if the trouble occurs.

In this paper, our study focuses on an effective way for the routing management. In particular, we propose a way to monitor the routing information exchanged by OSPF (Open Shortest Path First) in an autonomous system, which enables us to detect some trouble in terms of routing on the Internet. In addition, we can identify which router causes it. Furthermore, we implement our method, and then perform experiments on the Internet. Through the experiments, we examine how our method works well, and discuss its effectiveness.

1 はじめに

WWW による情報発信や電子メイルなどの著しい普及にしたがって, これらの技術を有効に利用するため多くの大学や研究機関・企業がインターネットに接続するようになった. これらの接続組織の受け口として多くの研究ネットワークや ISP(Internet Service Provider) が構成されている. また, 以上のようなネットワークの管理単位である自律系 (Autonomous System; 以下 AS) 同士が IXP(Inter eXchange Point) で相互接続されることも盛んに行なわれている. このため AS 内のネットワーク構成や AS 間の関係も複雑になっている. これにともない, 管理者が制御しなければならない経路数が膨大になっている. インターネット全体で交換される経路数は 4 万 6000 を越え, WIDE Project[1] 内部においても 2800 を越えている.

膨大な経路から障害の発生を検出することは難しい. また, 経路のふらつき (route flapping) のようにその検出自体が難しい障害もある [2]. さらに膨大な経路情報からその障害の原因を突き止めることは容易なことではない. このような状況であるにも関わらず現状では経験を積んだ管理者が対処している. 経路制御の経験が豊富な管理者は少数であるため, すべての障害を早期に検出して解決することは極めて困難となっている.

本研究では AS 内の経路制御に幅広く用いられている OSPF (Open Shortest Path First) [3] において交換される経路情報を監視する. さらに経路情報から計算される経路の変動を監視する. これによって経路の細かなふらつきを含む異常を検出することが可能となる. また異常の原因となる経路情報の生成源を推測する. 本研究は以上の仕組みをもって経路管理にかかる負荷の軽減を目的とするものである.

本稿では OSPF 経路情報の監視手法と監視対象について述べる. 次に提案手法を実現するための経路監視システムの設計について述べ, それに基づいた実装について述べる. さらに本システムの運用を行ない, その有効性について考察する.

2 経路情報の監視方法

OSPF は AS 内の経路制御のための代表的なプロトコルであり, 多くの場合 AS を複数の領域に分割して運用される.

本章では, OSPF において交換される経路情報 (以下, 経路情報) およびそれらから計算される経路 (以下, 経路) の本研究での監視方法について述べる.

2.1 経路情報の解析

本システムは AS 内の経路制御に幅広く用いられている OSPF において交換される経路情報を直接解析し, その変動を観測することによって経路の異常およびその原因となる情報の生成源を推測するものである.

既存の経路監視システムは各々のルータから情報を定期的に収集し, それらを解析することによって経路の異常などを発見するものであった. しかしこの手法ではある収集から次の収集までの間に起こる経路のふらつきなどは検出することができない.

さらに, ルータから収集した情報には多くの場合経路情報の生成源などは含まれていない. そのため異常を検出しても収集した情報から原因となる経路情報の生成源を特定することは困難である. また Cisco Systems 社製のルータなどのように交換されている経路情報を直接見ることができるルータも存在するが, 経路と経路情報の関係が不明であるため経路のふらつきの原因となる経路情報を特定することはできない.

以上の理由から, 経路のふらつきを引き起こす経路情報の生成源は管理者が手作業で特定するしかなかった. これらの問題は既存の経路監視システムが経路情報を直接解析するのではなくルータが解析した結果を収集・解析するために起こるものである.

それに対し, 本システムでは交換されている経路情報を直接解析するため経路情報の細かな変動までを解析することが可能である. さらに取得した経路情報から経路を計算することにより経路のふらつきの原因となる経路情報を特定することが可能である. また OSPF の経路情報には生成したルータの識別子が含まれるため, 異常があるルータを推測することが容易となる.

2.2 監視対象となる情報

本システムでは OSPF によって交換される経路情報の変動を監視する. さらに経路情報から SPF (Shortest Path First) アルゴリズムによって最小木を計算する際に選択される枝の移り変わり (以下, 最小木の変動) を監視する.

2.2.1 経路情報の監視

OSPF では 5 種類の情報を交換する．そのうち実際の経路情報は *Link State Update* によって交換される．*Link State Update* はさらに 5 種類に分類される．本システムではこれらの 5 種類に含まれる情報のうち以下にあげるものを抽出して監視の対象とする．

- *Router Links*

  この経路情報ではあるルータに接続されているネットワークまたはルータの一覧を広告 (advertize) する．接続されているネットワークまたはルータの変化は，そのネットワークに接続されているインターフェイスまたはネットワーク自体の障害を示していると考えられる．

- *Network Links*

  この経路情報ではあるネットワークに接続されているルータの一覧を広告する．接続されているルータの変化は該当するルータのインターフェイスの障害やネットワーク自体の障害を示していると考えられる．

  またこの情報の識別子は DR (*Designated Router*) のインターフェイスアドレスとなる．DR とは該当するネットワークに関して責任を持つルータである．ネットワークのアドレスがほとんど不変であることを利用し，これに着目して経路情報の識別子を監視することにより DR の変動を検出することができる．DR の変動は該当するネットワークに接続されているルータの設定の誤りを示していると考えられる．

- *Network Summary Links*

  この経路情報では監視している領域に隣接する他の領域の経路情報の要約を広告する．これに含まれるメトリックの変化は要約を生成しているルータおよび該当する領域内のルータの障害を示していると考えられる．

- *ASBR Summary Links*

  この経路情報では AS の境界のルータ (AS Boundary Router) に関する経路情報を広告する．これに含まれるメトリックの変化は要約を生成しているルータの障害を示していると考えられる．

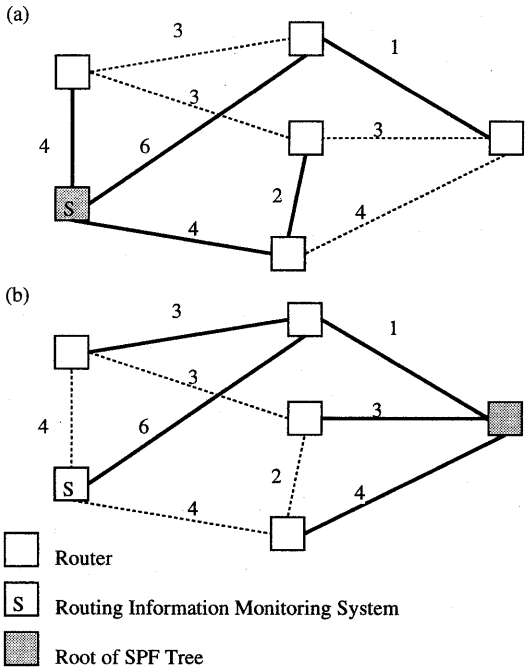

図 1: OSPF における領域内のルータの持つデータと SPF による計算

- *AS External Links*

  この経路情報では AS 外部から BGP4[4] 等で与えられた経路情報を広告する．この情報には外部のネットワークに対する AS 内の出口となるルータが含まれる．また AS 外のメトリックと AS 内でのメトリックの扱い方も含まれる．AS 外への経路は以上の情報および AS 内の経路より計算される．これらの情報の変化は情報を生成しているルータの障害または BGP4 などによる AS 間における経路制御の障害を示していると考えられる．

2.2.2 最小木の監視

OSPF におけるルータは自らを根とした最小木を計算することによって経路制御を行なう (図 1(a))．しかし OSPF では領域内のすべてのルータが同一のネットワーク構成に対する情報を持つため，領域内の任意のルータを根にした最小木を計算することができる (図 1(b))．これは領域内の任意のルータを始

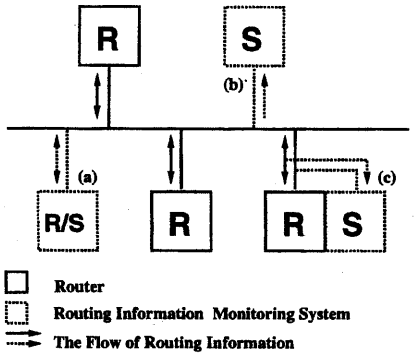

図 2: 経路情報の取得方法

点とするすべての経路を計算できることを意味する. 本システムではこれを利用してすべてのルータを根とした最小木の変動を観測することにより, 領域内の任意のルータを始点としたすべての経路のふらつきを監視する.

3 経路監視システムの設計

本章では, 2章で述べた手法を用いた経路監視システムの設計について述べる. さらに条件による障害の抽出や記録・報告の方法についても述べる.

3.1 経路情報の取得方法

前にも述べたように, 経路情報の取得はネットワーク上で交換される OSPF の情報を収集することによっておこなう. この情報を取得するには次の 3 つの方法がある (図 2).

(a) OSPF の経路制御に加わる.

(b) bpf(Berkeley Packet Filter) などを用い, データリンク層から直接該当する情報を抽出する.

(c) すでに OSPF で経路制御を行なっているプログラムと同じソケットを開き, 情報の複製を得る.

このうち, (a) を実現するためには接続するネットワークの DR と *adjacency* と呼ばれる関係を形成しなければならない. この場合, そのホストは OSPF でルータとして広告される. OSPF では領域内に存在するルータが増えると各ルータにかかる負荷が急激に大きくなるという性質があるためこの方法はあまり好ましくない.

また (b) を採用した場合, 本システムの起動時にすべての経路情報を得ることができない. したがって経路の変動の検出が遅くなる. さらにすべての経路を得たという判定が不可能であるため, SPF によって計算される最小木の変動が障害によるものかどうか判断できない.

以上に対し, (c) では新たに OSPF の経路制御に加わるわけではないので, 同じ領域の他のルータに負荷をかけることはない. また *adjacency* を形成する過程ですべての経路情報が交換される. したがって本システムと OSPF 経路制御プログラムを同時に起動することにより, 短時間にすべての経路情報を得ることができる. これによって最小木の変動の監視を開始することが可能となる.

以上の考察より, 本システムでは (c) の方法を採用する.

3.2 変化の検出条件および処理の記述

変化の検出は同一の識別子を持ったルータまたはネットワークに関する経路情報が変化する頻度を計算することによって行なう. また最小木の変化に関してもその頻度を計算することによって検出する.

経路情報や経路の変化は常に起こっているものである. どのような変化がどのくらいの頻度で起こった場合を異常と判断するか, また異常と判断した場合にどのような手順によって異常の記録・報告や情報収集などの処理を行なうかは管理者や AS によって異なり, また管理者の豊富な知識と経験によるところが大きい.

このような問題を解決するため, 本システムでは以下のような設定方法を提供することにより管理者の経験や知識をシステムに反映し, その AS の特質に応じた異常の検出やその後の処理を行なうことを可能としている.

変化の検出条件の記述

監視できる情報の種類ごとに変化を保持しておく時間とその時間内における情報の変化のしきい値を記述する. この条件を満たした場合, 次に述べる処理を行なう. ここではおもに異常が発生している可能性があると判断するための条件を記述する. しかし次に述べる処理を行なうために任意の条件を記述

することができる.

処理の記述

上記の条件を満たした場合の処理および各処理のために必要となる情報の形式を記述する. 処理の記述は任意のプログラム名を指定することによって行なう. 上記の条件だけでは異常であると断定できない場合でも, さらなる情報の収集を行なうような処理を記述することによって検出できる. またすべての変化に対して処理を行なう条件記述と HTML を生成する処理を組み合わせることによって障害の記録を WWW で表示するといったことも可能となる.

既存のシステムでは経路の管理を主眼においていたため, 記録・報告やその後の処理を決まった方法でしか行なうことができなかった. しかし本システムでは経路情報および経路の監視部分と障害の検出およびその後の処理を完全に分離したため, 経路情報および経路の任意の変化に対して任意の処理を行なうことができる.

4 実装

本システムは C 言語によりデーモンとして実装した. また監視装置としては OSPF を利用可能とした GateD の稼働する UNIX マシンで構成し, OSPF によって経路制御が行なわれている領域の任意の位置で運用することを想定している. 現状では経路情報の取得およびその解析による異常の検出, 記録および報告の処理部を実装している.

経路情報は OSPF による経路制御に参加している GateD が OSPF のために使用している生ソケット (raw socket) を利用することによってその複製を取得している.

異常の検出は同一の識別子を持ったルータまたはネットワークに関する情報が変化する頻度を計算することによって行なう. この値が管理者の設定した値を越えた時, 記録・報告などを行なう. 記録および報告の方法としてあらかじめ syslog による記録および電子メイルによる報告を用意している. さらに管理者の設定によってある条件を満たした時に任意のプログラムを起動することができる.

変化の検出および処理の記述例を図 3に示す.

なお領域内の各ルータを根にした最小木をもとめることによる経路のふらつきの監視に関しては未実装である.

```
# General Configuration
general {
    procdef 1 {
        proc "/home/is/naoto-m/work/isalive";
        format "%a";
    }
}
# Router Advertisement
rtr_adv {
    link {
        up_down {
            interval 900;
            rule {
                cnt > 0;
                proc 1;
            }
            rule {
                cnt > 9;
                proc syslog;
            }
            rule {
                cnt = 20;
                proc mail;
            }
        }
    }
}
```

図 3: 条件および動作の記述例

```
NET_LA: ls_id = 203.178.137.160, adv_rtr = 203.178.137.167
 att_rtr = 203.178.137.181, down, stat.up_down = 20
```

図 4: 奈良 NOC および京都 NOC での報告内容

5 評価

まず本稿における提案システムのうち実装できている部分への実験を行なった. イーサネットで構成された WIDE Project 奈良 NOC のバックボーンに (α) BSD/OS3.0 および (β)BSD/OS2.1, 京都 NOC のバックボーンに (γ)BSD/OS3.0 の稼働する PC をそれぞれ接続した. (α)(β)(γ) ともに GateD を用いて OSPF により経路制御を行なった. また (β)(γ) において本システムを稼働し, (α) のインターフェイスの up/down を繰り返しながら本システムの挙動を観察した. 本実験では (β)(γ) において *Network Links* の変動が観測されると予想されたが, インターフェイスの変化から約 30 秒遅れて (β)(γ) ともに変化が観測された (図 4). この遅延は *OSPF Hello* の間隔の影響と考えられる. この報告では識別子が *203.178.137.160* であるネットワークの *Network Links* を識別子 *203.178.137.167* のルータが広告し, そのネットワークに接続されている識別子 *203.178.137.181* のルータが up/down を 20 回繰り返して現在の状態が down であることがわかる.

次に本システムを用いて WIDE Project のバックボーンを監視した. 今回の約 3 週間の運用では特別なプログラムを指定することなく経路情報のふらつきのみを監視し, syslog での記録および電子メイルでの報告を行なった.

その結果, GateDが稼働しているルータを発生源とする *Router Links* に含まれる情報の一つが交互にstubあるいはtransit networkになるという異常な挙動をしていることが観測された. バックボーンのOSPFルータは新しい経路情報を受信する度に最小木を再計算していたため負荷を増大させていたと考えられる. 今までこの現象が検知されなかったのはこのルータが比較的重要でない場所に使用されていたこと, さらに通信が不可能になるような障害ではなかったことが理由であると推測される. この障害を取り除いたことによりバックボーンのOSPFルータ全体の負荷が減少したと考えることができる.

以上の実験および運用結果から, 本システムによって経路情報の変化をとらえることが可能であることが確認できた. また管理者の目の届かない異常の検出に有効である可能性が確認された. さらに, 管理者がこの異常によるルータの負荷に気がついていたとしても, その原因箇所を特定することが困難であるために大きな負担がともなうと考えられる. 今回の運用では障害の原因となるルータを特定できたことから, 本システムが管理負荷の軽減に対しても有効であることが期待される.

6 今後の課題と考察

本章では今後の課題について考察する.

6.1 経路情報の矛盾の検出

現在の設計では単一のサーバのみで領域のすべての経路情報を監視することができる. しかし一つのサーバだけでは領域の分断・再生などによるルータ間での経路情報の矛盾に関しては検出できない. また観測点と異なる領域の経路と観測点を含む領域になにがれる要約との矛盾に関しても検出できない.

この問題を解決するためには複数のサーバを領域内の別のネットワークに接続し, サーバ間で保持している経路情報の比較を行なう必要がある.

6.2 条件書式の多様化

現状では異常の検出に用いられる設定はごく単純なものとなっている. しかし管理者が手作業で障害の発生を判断するための条件は複雑なものである. このような条件を記述できるようにするため制御構造などを採り入れた条件書式の多様化についても検討する必要がある.

6.3 GateDによる経路制御の少数化

現在, 経路制御にGateDを用いることは少なくなってきた. このため, 本システムの「経路情報の複製を取得できるシステム」という前提を満たすことは困難となりつつある. これを解決するための1つの方法としてOSPFに情報取得用のグループを追加することがあげられる. このようなプロトコルの変更は困難ではあるが不可能ではない. また, 現在のOSPFの仕様ではルータ以外のネットワークに接続されている機器がOSPFによって経路を得ようとすると, 3.1章で述べたように他のルータに負担をかける形とならざるを得ない. このような問題に対しても上で述べたプロトコルの変更は有効である.

7 おわりに

本研究では経路情報を直接監視することによりその変動から異常およびその原因となるルータを推測するシステムを提案した. 本システムはAS内部の経路制御にもちいられるOSPFの経路情報を監視し, 管理者によって設定された条件によって設定された動作を行なうものである. 本システムによって経路のふらつきやその発生源, ルータ自身の異常などを推測することが可能となる.

今後は経路の計算による経路のふらつきの監視機能の強化を含めた実装の詳細化を行なっていく. またルータ間での経路情報の矛盾を発見するためシステムの分散化についても検討を行ない, AS内の経路情報監視システムとして一般化していく予定である.

参考文献

[1] WIDE Project. http://www.wide.ad.jp/, 1997.

[2] The IPMA Project. Internet Routing Recommendations. http://www.merit.edu/ipma/-docs/help.html, 1997.

[3] J. Moy. OSPF version 2. RFC2178, Internet Acrivities Borad, 1997.

[4] Y. Rekhter and T. Li. A border gateway protocol 4 (BGP-4). RFC1771, Internet Acrivities Borad, 1995.