

AS 内部のネットワークの安定性に関する研究

松井佑馬¹ 江崎浩²

東京大学工学部¹ 東京大学大学院情報理工学系研究科²

1. はじめに

インターネットは AS(Autonomous System)とよばれる自律的に構築運用されているネットワークを単位として、これらが経路制御によって相互接続されグローバルスケールの大規模分散システムを構成している。それぞれの AS の中身も、自律的に構築運用されているより細かな単位のネットワークの集合体となっており、その大きさは実に様々である。最近インターネットの急速な普及とともに、ネットワークの肥大化と複雑化の傾向が急激に加速している。AS と呼ばれる自律的に運用されているネットワークは、例えばネットワークプロバイダなどであり、そのネットワークの構成(トポロジー)はスター型のような単純なものではなく、網の目状の複雑なものとなってきている。

AS 内での IP パケットの転送経路を制御し、IP パケットが良好にネットワーク上で分散して目的の計算機に到達するために、OSPF(Open Shortest Path First)([1],[2],[3])のような経路制御プロトコルが動作している。AS のように地理的にも、ネットワークを構成する要素(ルータ、スイッチ、デジタル回線など)の数としても大規模なネットワークにおいては、障害やメンテナンスなどの様々な原因により、その内部で経路が不安定となることが深刻な問題の一つとなってきている。経路が不安定になった場合、現状では、その都度ネットワークのオペレータが暫定的な対応を手作業で行うということが日常的である。しかし、障害を調査するにあたって確立された手法というものは存在していない。

そこで本研究では大規模な AS の内部における障害状況を実測データ(OSPF の流す情報)から把握し、ネットワークのオペレーションを支援するための手法を提案・実装するとともに、経路の不安定要因を探る。

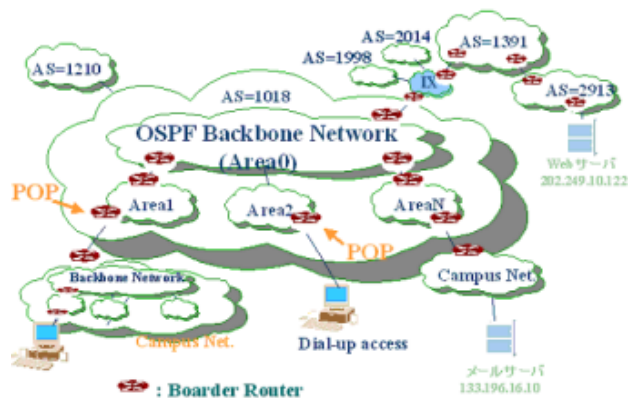


図1: インターネットの構造の例

2. AS 内部の経路制御

2.1. インターネットの構造と AS

インターネットにおけるデータの転送単位である IP パケットを目的のノードに配送するために、経路制御機能がある。この実現手段が経路制御プロトコル(Routing Protocol)である。経路制御機能により、インターネットにおける「ネットワークのネットワークング」が実現されている。経路制御は自律的に運用されるネットワークを単位として動作し、階層的かつ再帰的に運用される(図1)。すなわち、複数のエンドホストの集合体が最も下位層のネットワークであり、これらのネットワークの集合が次のレベルのネットワークとして定義される。このようなネットワークされたネットワークを上位のネットワークと定義することができ、上位のネットワークから見ると、再帰的にネットワークがネットワークされた構造となっている。各ネットワークは、そのネットワークの大きさと経路制御ポリシーに適した経路制御方式を個別に選択することができるようになっている。同一の経路制御方式を適用しているネットワークのことを、ルーティングドメインと呼ぶ。

AS とはインターネットの構成単位であり、自律的に運用されるネットワークのことである。各 AS は 16 ビットで表現されるグローバルユニ

Research on Stability of Network in AS

¹ Yuma Matsui, School of Engineering, The University of Tokyo

² Hiroshi Esaki, Graduate School of Information Science and Technology, The University of Tokyo

ークな AS 番号を持っている。プロバイダ(ISP) 1社で複数の AS 番号を持つこともあるが、一般的には ISP が AS に相当する。今現在で、全世界に約 1 万の AS が存在している。

2.2. 経路制御 (ルーティング)

実際のインターネットでは目的の宛先に到達するまでに、経由するルータの組み合わせが何通りも存在する場合がある。このようなときに送信元から目的の宛先までの最適な経路を選択する仕事をすることを経路制御と呼ぶ。具体的には、パケットが最適な経路を通るようにルーティングテーブルを作成し、それにもとづいてルータに入ってきたパケットをどのインターフェイスに送り出すか決定する、ということである。

次にルーティングテーブルの作成方法というのを考えると、これはスタティックルーティングとダイナミックルーティングの 2 種類に大別できる。スタティックルーティングとは人が直接、ルータにデータを入力してルーティング情報を固定的に設定する方法である。手で思い通りに設定できる分、ネットワークの規模によっては設定の作業量が膨大になってしまう。それに対してダイナミックルーティングはルータ同士がネットワークの経路情報をやり取りし、自動的にルータでルーティングテーブルを作成する方法である。この場合、手間は減るが、計算に基づき経路を動的に決定していくのでルータに負荷がかかる。

ダイナミックルーティングにおけるプロトコルは大きく IGP(Interior Gateway Protocol)と EGP(Exterior Gateway Protocol)に分けられる。IGP は AS の内部で用いられるルーティングプロトコルのことであり、一般的に小規模なネットワークでは RIP が、AS のような大規模なネットワークでは OSPF が用いられる。EGP は AS 間で用いられるルーティングプロトコルのことであり、現在使用されているものはほとんどが BGP である。

2.3. OSPF の概要

OSPF は IGP の 1 つであり、AS 内部での経路制御に用いられる。OSPF は Open Shortest Path First の略であり、これは Shortest Path First(または Dijkstra)アルゴリズムに由来する。

OSPF の大きな特徴は、リンクステートアルゴリズムを採用し素早い収束性を実現していることである。各ルータはリンクステートと呼ばれ

るメッセージを流し、ここにはルータが接続しているリンクの状態や、そのリンクのネットワーク、コストなどの情報が入っている。リンクステートを受け取った各ルータでは、その情報にもとづいてネットワーク構成を把握し、構成を表す表を構築する。つまり、各ルータが全ルータとリンクの構成を把握した地図を持っているようになる。このネットワーク構成を表す表をリンクステートデータベースという。そして各ルータはこのデータベースから Shortest Path First アルゴリズムを用いて、自身を始点とした最短パスツリーを作成し、そこからルーティングテーブルを作成する。

リンクステートアルゴリズムのメリットとして、各ルータがネットワーク構成を把握していることにより、ネットワーク構成が変化した場合に素早くルーティングテーブルを再構築することができるが挙げられる。このことから OSPF は大規模なネットワークにおいても運用可能であり、スケーラビリティに優れていると言える。

2.4. LSA(Link State Advertisement)の概要

実際のリンクステートのやり取りを担っているのは Link State Update パケットと呼ばれるパケットである。この中には LSA と呼ばれる、ルータやネットワークの状態を記述したデータが含まれている。LSA には以下に示すような、主に 5 種類が存在している。

- LSA Type1 (Router LSA)
ルータがどのようにリンクと接続しているのかという情報。
- LSA Type2 (Network LSA)
ネットワークにどのようなルータが接続しているかという情報。
- LSA Type3 (Summary LSA - IP network)
Summary LSA はエリア外の情報。Type 3 はエリア外の OSPF で生成されたネットワークの情報。
- LSA Type4 (Summary LSA - ASBR)
こちらも Summary LSA だが、Type 4 は ASBR(AS Boarder Router・AS 境界ルータ)の情報。
- LSA Type5 (AS-external LSA)
ASBR によって作成される、AS-external 経路のネットワークに関する情報。

3. ネットワークの安定性要因

3.1. 経路の不安定性

OSPF を適用して動的な経路制御を行っている AS 内部において、経路制御的な観点でネットワークが不安定となる要因は、リンクやノードのアップダウン(障害)として現れる。なお OSPF では、リンクとして、単純にデジタル専用線のようないわゆるリンクと呼ばれる通信回線(ルータが収容しているリンク)だけではなく、リンクに接続しているルータや境界ルータが広告する経路情報も含まれる。AS 内のノードやリンクが障害などの原因で落ちる(リンクダウンおよびノードダウン)と、そのリンク・ノードを通過するはずだったトラフィックが、他のノードやリンクを通らざるを得なくなる。その結果、トラフィックに偏りが生じ、特定のリンクやノードが過負荷の状態になってしまうことも起こりうる。

3.2. 不安定性の要因

経路を不安定にする要因というのは様々であり一概に言えるのものではないが、現状では主に次のようなものを可能性として考えることができる。

- ルータ上のソフトウェアの更新
- ルータの再起動
- インターフェイスのダウン
- 外部 AS から広告される経路のフラッピング

これらの中でもソフトウェアの更新やルータの再起動などはメンテナンスの一環として意図的に行われることが多いだろう。しかし再起動に関して言えば、ルータに高い負荷がかかった際に突発的に起こるということもありうる。オペレータが、ネットワークが不安定になっているのを見たときには、それが自分の把握しているメンテナンスの範囲ならば問題は小さいが、そうでない場合には上に挙げたような各要因の特徴を考えに入れつつ、様々な角度から分析をする必要がある。

4. LSA を用いた障害の検出手法

4.1. 従来技術

2.4. で述べたように LSA にはネットワークの状態が記述されているという特徴があるので、これをネットワークの障害情報を知るために用

いることができる。そのための手法として、既存の技術では次のようなものが考えられる。

- ルータのログ
単純にルータのログを調べる。ただし、そこには必ずしも自分の必要としている情報があるとは限らない。ルーティングをつかさどるソフトウェアの種類によってログの形式も違ってくる。GNU Zebra[4]のようにオープンソースなプログラムにおいては、ソースを書き換えることにより自分の望むようにログの形式を変えることも可能だろうが、これはすべてのルータに関して適用できる方法ではない。
- SNMP [5]
SNMP はネットワークに接続された機器類をネットワーク越しにて監視するためのプロトコルである。SNMP エージェントの搭載された機器の情報は、手元の SNMP マネージャからリモートで取得することが可能になる。この仕組みを用いて OSPF の情報を取得することもできる。しかし、この手法でネットワークの障害情報を取得しようとする、情報を知りたいルータ全てにたいしてリクエストを送らなければならないので手間が大きいうえ、ネットワークにたいしても負荷がかかるという問題がある。

4.2. 提案システムの概要と特徴

4.1. で述べたような技術よりも汎用的な情報の取得方法として Link State Update パケットを直接キャプチャして、中の LSA を覗くという方法が考えられる。この方法の利点として、パケットのフォーマットが規定されているためにルーティングをつかさどるソフトウェアの種類に関わらず用いることができる点と、AS の中のある 1 台のノードに設置すればそれだけで必要な情報が全て得られるという点が挙げられる。

パケットをキャプチャする方法として、一般的には tcpdump[6]というプログラムがよく用いられる。しかしここでは障害の情報を知りたいという目的にかなった、OSPF のパケットをキャプチャすることに特化したプログラム(永見³による制作)を用いることにした。このプログラムの動作は、OSPF の Link State Update パケットをキャプチャして、その中に格納された LSA を解析し、結果をログとして出力するというもの

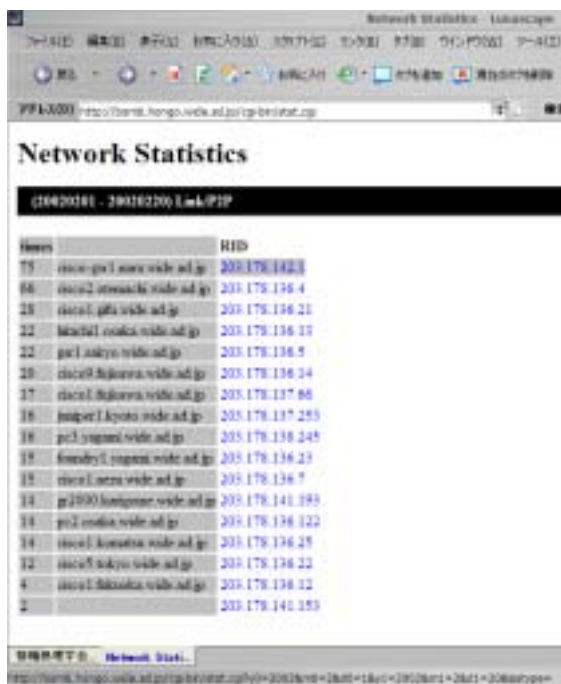
³ WIDE Project

である。

次に、そのログをどのような形でオペレータの役に立つよう整理するのかということの説明する。何も加工しないログそのままでは、ある特定の期間、特定のリンクについてのアップダウンの状況を知りたい、などという要求に対して、毎回自分で適切な処理を行う必要が生じてしまう。そこで、ユーザからはパラメータだけを与えてもらい、それに応じたログの処理を行い、結果を出力して手間を軽減してくれるようなプログラムを実装した。

具体的には、ユーザから与えられるパラメータとして、まず期間(何年何月何日から何年何月何日までという情報)と LSA の種類(による知りたい対象のタイプの情報)の2つを指定してもらう。その時点で、その期間においてアップダウンの回数が多かった順にルータの RID(Router ID)を一覧で出力する。(図 2)さらにそこから RID を指定してやれば、詳細なログとアップダウンを視覚化したグラフが出力される(図 3)。

このグラフは横軸を時間、縦軸をアップなら 1、ダウンなら 0 としてプロットしたものであり、状態変化の状況を特定期間における時系列で図示したものとと言える。グラフによる利点は直感的な把握が容易になることに加えて、状態変化回数の多かったルータ同士のグラフを比較することで関連性を発見でき、障害の原因究明に役立つといったことも考えられる。



Items	RID
75 r100-g0/1/area-wide-ad.jp	203.178.136.3
66 r100-2/area-wide-ad.jp	203.178.136.4
28 r100-1/p1/area-wide-ad.jp	203.178.136.21
22 h100-1/area-wide-ad.jp	203.178.136.10
22 g0/1/area-wide-ad.jp	203.178.136.5
20 r100-9/area-wide-ad.jp	203.178.136.14
17 r100-1/area-wide-ad.jp	203.178.137.86
16 h100-1/area-wide-ad.jp	203.178.137.255
16 g0/1/area-wide-ad.jp	203.178.136.245
15 h100-1/area-wide-ad.jp	203.178.136.23
15 r100-1/area-wide-ad.jp	203.178.136.7
14 g2/100/area-wide-ad.jp	203.178.141.193
14 g0/2/area-wide-ad.jp	203.178.136.122
14 r100-1/area-wide-ad.jp	203.178.136.25
12 r100-5/area-wide-ad.jp	203.178.136.22
4 r100-1/area-wide-ad.jp	203.178.136.12
2	203.178.140.153

図 2: 状態変化回数の多かったルータ一覧の例

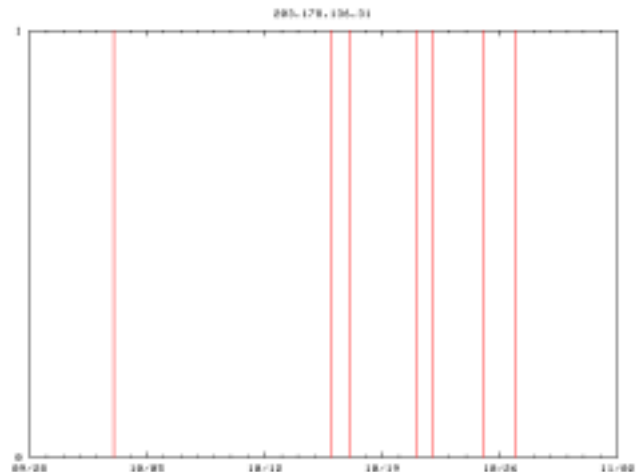


図 3: 状態変化の視覚化の例

5. おわりに

本論文ではまず 2 章で AS 内部の経路制御に関する技術を紹介した。次に 3 章では、ときに深刻な問題となる、ネットワークにおける経路の安定性について説明と簡単な考察を行った。そして 4 章では LSA を用いた障害の検出手法に関して、従来技術との比較のうえで、OSPF のパケットキャプチャプログラムをもとに、オペレーションに役に立つという観点から障害情報の見せ方について提案・実装を行った。

残された課題としては、今回実装したツールをもとに広くデータの収集・分析を行い、またある程度はネットワークの不安定要因の切り分けを行うことができるような機能を追加するといったことが考えられる。

参考文献

- [1] J. Moy. *RFC2328 OSPF version2*. 1998.
- [2] J. Moy. *OSPF Anatomy of an Internet Routing Protocol*. 1998.
- [3] 友近剛史, 池尻雄一, 小早川知昭. *インターネットルーティング入門*. 2001.
- [4] *GNU Zebra routing software*
<http://www.zebra.org/>
- [5] J. Case. *RFC1157 SNMP*. 1990.
- [6] *TCPDUMP public repository*
<http://www.tcpdump.org/>