

オーバレイルーティング手法を用いた 大規模ネットワーク障害からの回復

A new method of proactive recovery mechanisms for large-scale network failures

堀江拓郎¹
Takuro Horie

長谷川剛¹
Go Hasegawa

亀井聡²
Satoshi Kamei

村田正幸¹
Masayuki Murata

大阪大学 大学院情報科学研究科¹
Graduate School of Information Science and Technology, Osaka University
NTT サービスインテグレーション基盤研究所²
NTT Service Integration Laboratories

1 はじめに

ネットワークの高信頼化を目指した既存研究の大半は、障害発生モデルとして単一障害を想定している。一方、大規模災害、テロ、ルータソフトウェア (OS) の不具合などによって発生すると考えられる、複数のネットワーク構成要素が同時に故障するような大規模ネットワーク障害に関する研究はほとんど行われていない。

本稿では、現状の IP ネットワークにおける種々の問題を克服する短中期的な手法として近年注目されているオーバレイネットワーク技術を用いて、大規模ネットワーク障害が発生した際に BGP による IP ルーティングでは回復に長時間を要する AS 間通信を、短時間で復旧する手法の提案を行う。提案手法はプロアクティブ型オーバレイルーティングの一種であり、障害発生後、事前に算出しておいた代替設定を用いて即座に経路変更を行うことができる。

2 提案手法

2.1 既存研究とその問題点

Resilient Routing Layers (RRL) [1] は、元のネットワークトポロジに加えて、障害発生時に用いる Routing Layer (RL) と呼ばれる仮想的なネットワークトポロジを事前に複数用意し、それらに対応した経路表を用いることで障害発生時の経路制御を行う。各 RL では障害ノードを想定し、そのノードのことを隔離ノード (safe ノード) と呼ぶ。また、作成した RL 群を Routing Layers Set (RLS) と呼ぶ。障害発生時に用いる RL は、想定障害が発生した際に障害箇所を経由する経路が使われないように、隔離ノードに接続するリンクの重みを最大値にする。障害検知時には、障害ノードが隔離されている RL を用いることで、他のノード間の通信が切断されることなく即座に障害を回避した経路を用いることができる。

RRL は全てのノードをいずれかの RL において隔離することで、単一障害に対して 100% の復旧を可能にする [1]。さらに、1 つの RL で同時に隔離されている複数のノードに対しては、それらに同時に障害が発生した場合においても復旧が可能である。しかし、1 つの RL において多くのノードを隔離すると使用可能なリンクが減少するため、平均経路長が大きくなるという問題がある。また、多数の RL を用意することでより多くの障害パターンに対応できるが、用意できる RL の個数はノードの記憶領域やパケットヘッダに RL を指定するために追加するビット数などにより制限を受ける。つまり、隔離するノードの選び方、用意する RL の個数には、複雑なトレードオフが存在する。そのため、同時に障害が発生しやすいノードを同一の RL で隔離することによって、複数障害に対して効果的な RLS を作成する必要がある。

RRL は、様々なレイヤにおける既存の経路制御に適用可能であるが、既存プロトコルの変更が必要となるた

め、IP ネットワークなどの広域ネットワークで利用するためには標準化作業が必要となる。そこで本稿では、アプリケーションレベルの制御手法であるため、プロトコルの標準化作業が不要となり、サービスの IP ネットワーク上での早期展開が可能なオーバレイルーティング技術を用いて RRL を実現することを前提とする。また、別の利点として、通常は BGP によって利用可能な AS が制限されているピアリングリンクを、全ての AS 間通信に用いることが可能となる点が挙げられる。これにより、IP ルーティングでは不可能であった経路を用いた代替経路の利用が可能である。さらに、元のネットワークトポロジを用いてルーティングを行う場合よりも、平均経路長が短くなることも期待される。

2.2 RLS 生成法

本稿においては、発生確率が高い複数同時障害として、隣接する複数ノードの障害および同じ性質を持つノード群の障害を想定し、それぞれに効果的と考えられる RLS 生成法として、ハブ法および属性法を提案する。また、比較のために簡易な方法で RLS を生成可能なランダム法、各ノードの次数を考慮して作成する次数法を合わせて説明する。以下、1 つのノードが複数の RL で隔離されることを重複と呼び、各 RLS 生成法について説明する。なお、紙面の制約上、各 RLS 生成法の具体的なアルゴリズムの表記は割愛する。

ハブ法および重複ハブ法は、ネットワーク全体に大きな影響を与える次数の高いハブノードと、その隣接ノードの障害を想定している。具体的には、ハブノードと、その隣接ノードを隔離した RL を作成する。また、ランダム障害対策として全てのノードを隔離するために、ランダムにノードを隔離した RL も作成する。

属性法および重複属性法は、同じ地域、同じ OS の同じバージョン、および同じ管理ポリシーといった属性が等しいノード群の障害を想定している。RLS を生成する際には、ノードを与えられた属性でグループに分割し、そのグループを隔離した RL を作成する。

充填ランダム法、一様ランダム法、および重複ランダム法は、隔離するノードをランダムに選択する。この生成法は、機器の経年劣化などのランダムに発生する障害に対して有効であると考えられる。次数降順法および次数昇順法は、ノード次数に従って隔離するノードを選択する。この生成法は、人為的な攻撃による影響が大きい箇所の障害などに対して有効であると考えられる。また、ランダム法および次数法は計算量が小さく、短時間に RLS の生成が可能である。

2.3 RL 選択法

RL を用いてルーティングする際の最も単純な方法として、送信ノードが RL を選択する方法 (静的 RL 選択) が挙げられる。静的 RL 選択では送信ノードが RL を選

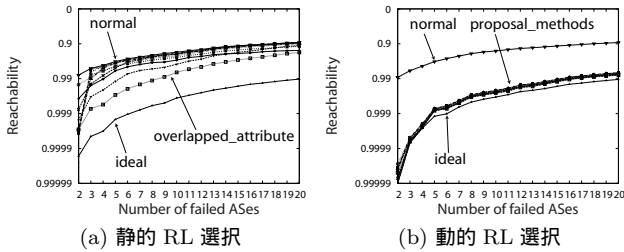


図 1 障害ノード数に対する到達性の変化

表 1 平均経路長 (障害ノード数: 2 / 10 / 20)

	静的 RL 選択	動的 RL 選択
重複属性法	2.83 / 2.76 / 2.72	2.79 / 2.80 / 2.83
重複ランダム法	2.99 / 2.97 / 2.92	2.99 / 2.95 / 2.91
元のトポロジ	2.84 / 2.83 / 2.81	2.84 / 2.82 / 2.80
理想トポロジ	2.70 / 2.72 / 2.73	2.70 / 2.72 / 2.74

択するために、全ての障害ノードを隔離している RL が存在しなければ、残りのノード間の到達性は完全に保証されない。しかし、静的 RL 選択では中継ノードでの処理が簡略化されるため、低遅延ネットワークに適している方法であるといえる。

また、より多くの障害ノードを回避するために、各中継ノードが RL を動的に選択する方法 (動的 RL 選択) が考えられる。このとき、パケット転送がループに陥ることを防ぐため、一度暫定 RL として用いられた RL は、そのパケット転送においては再度用いないようにする必要がある。

3 数値計算による評価

3.1 評価方法

評価トポロジは、CAIDA [2] が公開する AS トポロジから JPNIC が管轄する AS のみを抽出した、ノード数が 259、リンク数が 1162、平均次数が 4.4 のものを用いる。発生障害は、同じ属性のノード群に障害が発生する、属性ノード障害を用いる。ここでは、与えられたネットワークを分割グラフをまたぐリンク数が小さくなるように分割し、分割グラフに応じて各ノードに属性を与える。また、各ノードの記憶領域が十分な環境を想定しており、各 RLS において生成する RL 数に制限を設けていない。評価指標として、障害ノードを送受信ノードに含まないすべてのノードペア間の到達性、到達可能なすべてのノードペア間の平均経路長、および平常時に対する障害発生時のリンクの使用回数の比 (トラヒック比) を用いる。評価基準として、障害の回復がされていない元のトポロジを用いた場合 (normal) と、障害ノードを除いて経路の再計算を行った場合である理想トポロジ (ideal) を導入する。

3.2 評価結果

図 1 に、静的 RL 選択および動的 RL 選択時の障害ノード数に対する到達性の変化を示す。図 1(a) より、静的 RL 選択時には重複属性法が最も到達性が高いことがわかる。具体的には、2 ノード障害発生時に、98% の到達性を 99.9% まで回復する。これは、重複属性法が想定した障害が発生したために効果的な回復が行われたと考えられる。

一方、図 1(b) では、動的 RL 選択時には障害ノードが多い場合にも、全ての RLS 生成法が理想トポロジを用いた場合とほぼ同等の到達性を実現することがわかる。特に 20 ノード障害発生時には、87% の到達性を 98% まで回復している。また、RLS 生成法ごとの到達性の大きな差は見られないことがわかる。これらは、いずれの

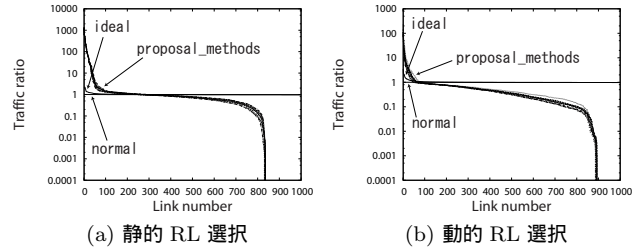


図 2 2 ノード障害発生時におけるトラヒック比

RLS も動的に RL を切り替えることによって、障害の規模および RLS 生成法に関わらずほとんどの障害を回避できたためと考えられる。

表 1 に、到達性の改善度合いが大きい重複属性法と、最も多くのノードを隔離する重複ランダム法の静的 RL 選択および動的 RL 選択時の平均経路長を示す。表中は、同時障害ノード数がそれぞれ 2/10/20 ノードの時の、各手法において実現される平均経路長を示す。表 1 より、提案手法の平均経路長は、静的 RL 選択と動的 RL 選択時のいずれも元のトポロジからほぼ増加しないことがわかる。また、平均経路長が元のトポロジの経路よりも減少している場合がある。これは、オーバレイルーティングを用いることによって、IP ルーティングでは全てのノード間で用いることができないピアリングリンクを、用いられるようになったことが要因と考えられる。

図 2 に、静的 RL 選択および動的 RL 選択時の 2 ノード障害発生時におけるトラヒック比を示す。平常時に比べて、静的 RL 選択時には約 1000 倍 (図 2(a))、動的 RL 選択時には約 100 倍 (図 2(b)) にトラヒックが増大するリンクがあることがわかる。これは、RL の切り替えによるトポロジの大きな変化、およびリンクの減少による平均経路長の増加が要因と考えられる。

4 おわりに

本稿では、オーバレイネットワーク技術を用いて、大規模ネットワーク障害が発生した際に BGP による IP ルーティングでは回復に長時間を要する AS 間通信を、短時間で復旧する手法の提案を行った。数値計算による評価の結果、静的 RL 選択時には、同時障害が発生しやすいノード群を隔離した RLS を用いることで、平均経路長をほとんど増加させることなく、到達性を大幅に改善できることを確認した。動的 RL 選択時には、平均経路長をほとんど増加させることなく、理想経路とほぼ同等の到達性を実現できることを確認した。また、動的 RL 選択の場合においては、RLS 生成法の違いによる性能差はほとんど確認できないため、ランダム生成法などの簡易な方法で十分であると予想される。一方、提案手法の回復処理によってトラヒックが急激に増大するリンクが存在することが明らかとなった。そのため、障害発生時には重要な通信を優先的に転送する等の制御が必要となると考えられる。

謝辞

本研究の一部は文部科学省科学技術振興調整費「先端融合領域イノベーション創出拠点の形成: ゆらぎプロジェクト」の研究助成によるものである。ここに記して謝意を表す。

参考文献

- [1] A. Hansen, A. Kvalbein, T. Čičić, and S. Gjessing, "Resilient routing layers for network disaster planning," *Lecture notes in computer science*, vol. 3421, pp. 1097–1105, Apr. 2005.
- [2] The CAIDA Web Site, available at <http://www.caida.org/home/>.