

大規模ネットワーク障害に対応可能なオーバレイルーティング手法

堀江 拓郎[†] 長谷川 剛^{††} 亀井 聡^{†††} 村田 正幸[†]

[†] 大阪大学 大学院情報科学研究科 〒560-0871 大阪府吹田市山田丘 1-5

^{††} 大阪大学 サイバーメディアセンター 〒560-0043 大阪府豊中市待兼山町 1-32

^{†††} NTT サービスインテグレーション基盤研究所 〒180-8585 東京都武蔵野市緑町 3-9-11

E-mail: †{t-horie,hasegawa,murata}@ist.osaka-u.ac.jp, ††kamei.satoshi@lab.ntt.co.jp

あらまし 本稿では、現状の IP 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、大規模なネットワーク障害が発生した際に従来の BGP による IP ルーティングでは到達不可能、あるいは到達可能な状態までの回復に長い時間を要する 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障害回復手法をオーバレイルーティングに適用したものであり、障害発生後、即座に代替設定を用いて経路制御を行うことができる。提案手法の有効性の評価は数値計算によって行う。その結果、提案手法を用いることにより、全体の 8% のノードに同時障害が発生した場合に、平均経路長を平常時とほぼ変えることなく、ネットワークの到達性を 87% から 98% まで改善できることがわかった。

キーワード オーバレイネットワーク、経路制御、ネットワーク障害、プロアクティブ型障害回復

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

Takuro HORIE[†], Go HASEGAWA^{††}, Satoshi KAMEI^{†††}, and Masayuki MURATA[†]

[†] Graduate School of Information Science and Technology, Osaka University

Yamadaoka 1-5, Suita, 565-0871 Japan

^{††} Cybermedia Center, Osaka University

Machikaneyama 1-32, Toyonaka, Osaka, 560-0043 Japan

^{†††} NTT Service Integration Laboratories, NTT Corporation

Midori-cho 3-9-11, Musashino, Tokyo, 560-0043 Japan

E-mail: †{t-horie,hasegawa,murata}@ist.osaka-u.ac.jp, ††kamei.satoshi@lab.ntt.co.jp

Abstract This paper proposes a recovery mechanism from large-scale failure based on overlay networking technologies. Our method is based on the proactive recovery scheme, which pre-calculates multiple routing tables for possible network failures and select one of them immediately after detecting the failures. Through numerical calculation results in terms of reachability and average path length, we show that our method improves the network reachability after the failures from 87% to 98%, while keeping the path length to be enough small, when 8% of the network nodes are down simultaneously.

Key words Overlay network, routing, large-scale network failures, proactive failure recovery

1. はじめに

従来、信頼性の高いネットワークは冗長性に優れた構成を組むことにより実現され、現用面の装置に故障が発生した場合には、予備面を用いて転送を行うことにより通信を復旧することを可能とする。しかし、インターネットの経路制御をつかさどる IP ルーティングプロトコルでの転送経路切替は、ネットワークを構成する各装置が故障箇所を把握し、最適な経路を再計算した後に新しい経路として採択するという方法であるため、短時間での転送経路の切替は困難である [1]。

ネットワーク機器やリンク、およびそれらを制御するソフト

ウェアを冗長化し、障害発生時に冗長系へ切り替えるような手法において重要となるのは制御コストと性能のトレードオフであり、既存研究の多くはこの点に着目している。そのため、大規模災害やテロ、大規模停電などによって引き起こされる大規模なネットワーク障害に対しては発生確率が小さいにも関わらず準備すべき冗長系の規模が大きく、コストが大幅に増大するため、対応が極めて困難となる。また、これまでのネットワークの高信頼化を目指した研究の大半は、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、

かつ、面的に発生する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。さらに、IP ネットワークに対する同種の研究も少ないといえる。この理由として、IP そのものが軽度の障害発生に対しては代替経路の発見が比較的短時間に行われることが挙げられる。しかし、インターネットの AS 間経路制御を行っている Border Gateway Protocol (BGP) [2] は障害が大規模である場合や、ある特定のトポロジ環境下においては、代替経路の発見および経路の収束に非常に長い時間(数分~数時間)を必要とすることが指摘されている [3, 4]。また、そもそも BGP には Count-to-Infinity 問題などが存在し、経路収束にかかる理論上限は存在しない [5]。そのため、BGP の経路収束時間を改善するための様々な手法が提案されている [6-8] が、そのほとんどは BGP や TCP/IP などのプロトコルそのものの改変を必要とするため、導入には標準化作業が必要となり、現在のインターネットへの適用は困難であると考えられる。さらに、AS 間リンクにはトランジットリンクやピアリングリンクなどのコスト構造が異なるリンクが存在し、各 Internet Service Provider (ISP) はそれらの経済的コストや政治的思惑を考慮してトラフィックの経路制御を行っている。このことは大規模ネットワーク障害の発生などの非常時におけるネットワーク接続性にも影響を与えると考えられる。

そこで本稿では、現状の IP 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ーレイネットワーク技術を用いて、大規模なネットワーク障害が発生した際に従来の BGP による IP ルーティングでは到達不可能、あるいは到達可能な状態までの回復に長い時間を要する 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ー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。提案手法の有効性は、インターネットにおける AS レベルのネットワークトポロジを用いて検証し、提案手法が単一 AS の障害のみに留まらず、複数 AS の同時障害発生時においても、BGP による IP ルーティングに比べて高い到達性を実現し、かつ障害発生後の平均経路長が理想的な場合に比べてほぼ増加しないことを示す。

2. 関連研究

近年、IP ネットワーク層の上のアプリケーション層に、独自の論理ネットワークを構築するオーバーレイネットワークが検討されている。例えば、P2P ネットワーク、Grid ネットワーク、IP-VPN サービスなどが挙げられる。これらのアプリケーションはある特定のサービスを前提として論理ネットワークを構築し、それぞれのアプリケーションのポリシーに従ってアプリケーショントラフィックの制御を行う。さらに、特定のアプリケーションを前提とせず、トラフィックのルーティングそのものを目的とする、オーバーレイルーティングと呼ばれる技術も登場しつつある [9, 10]。

また、ネットワーク障害回復手法として、リアクティブ型およびプロアクティブ型方式の回復手法が検討されている。リアクティブ型障害回復手法とは、ネットワーク障害を検出した後に、障害に応じた復旧処理内容を算出し、実行することによって障害から回復を行う障害回復技術である。具体的には、ネットワークから定期的にルーティング情報や到達性情報を取得することによって障害を検出し、その情報をもとに代替経路の算出を行い、新しい経路をネットワーク中に伝播させることにより、障害からの回復を行う技術である。リアクティブ型障害回復手法は、発生した障害に応じて動的に代替経路の計算や経路情報の伝播などを行うことを前提としているため、あらかじめ障害箇所を予測する必要がなく、様々な障害に柔軟に対応できる。一方、より早く障害を検出するためにはネットワークに関する情報をより多く、かつ頻りにノード間で交換する必要があるため、大規模なネットワークに適用することは難しい。また、障害発生後の新たな経路計算結果をネットワーク全体へ伝播させる必要があるため、経路が安定するまでに時間がかかるという欠点を持つ。リアクティブ型障害回復手法を適用したオーバーレイルーティング手法の例として、Resilient Overlay Network

(RON) [9] が挙げられる。しかしながら、RON が構成するオーバーレイネットワークはフルメッシュ接続を前提とし、総当りによって情報を取得しているため、参加ノード数の 2 乗に比例した量、および頻度で情報交換が行われる。すなわち、RON は情報交換のオーバーヘッドと引き換えに障害の早期検出、早期復旧を可能にする方法であり、小規模ネットワークや利用帯域に余裕のあるネットワークなど、特定の条件化においては非常に有効な方法である。しかし、一般的なネットワークはコストの面から必要最小限の構成が成されており、既存の、特に大規模なインターネット環境に適用するには問題点が多いと考えられる。

一方、プロアクティブ型障害回復手法とは、リアクティブ型障害回復手法とは対照的に、障害をあらかじめ予想し、事前に準備した復旧処理を行うことにより、障害からの早期復旧を目指す技術である。一般的なルーティングプロトコルでは、障害を検出したノードはそれを広告することになっており、プロアクティブ型障害回復手法はこのような情報を基に、どのノードに障害が発生したかを知り、復旧処理を適用する。プロアクティブ型手法はリアクティブ型手法のように平常時にはネットワークの観測を行わないため、障害が発生していない状況におけるコストを最小限に留めることができる。

そこで本稿で提案する障害回復手法は、大規模なインターネット環境に適用可能であり、かつ障害検出時にただちに障害回復を行うことができる、プロアクティブ型障害検出手法を用いるものとする。具体的には、簡潔、かつ拡張性に優れたプロアクティブ型経路制御手法の 1 つである Resilient Routing Layers (RRL) [11] を基として、大規模なネットワーク障害に対応可能な手法に拡張する。また、RRL には障害発生後のネットワークにおいても負荷分散を行えるように拡張した Multiple Routing Configurations (MRC) [12] があり、これと提案手法を統合することによって、性能のさらなる向上も可能であると考えられる。

3. 提案手法

3.1 Resilient Routing Layers (RRL) [11]

3.1.1 概要

RRL 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、それらに対応したルーティングテーブルを用いることで障害発生時の経路制御を行う。障害発生時に用いるネットワークトポロジは、そのトポロジにおいて想定している障害が発生した際に、障害発生箇所を利用した経路が使われないように、元のネットワークトポロジから適宜リンクを削除(リンクの重みを最大値に設定)したトポロジとなる。すなわち、障害が発生した際に、ネットワークの他の部分は障害の影響を受けずに通信を継続することが可能となる。

各 RL において障害発生が想定されているノードのことを隔離ノード (safe ノード)、作成した RL 群を Routing Layers Set (RLS) と呼ぶ。平常時の RL (元のネットワークトポロジをそのまま利用)を除く、全ての RL には 1 つ以上の隔離ノードが存在する。また、各 RL においては、隔離ノードへのリンクの重みを最大値にすることで、隔離ノードを途中経路として用いないルーティングテーブルを持つことができる。このようにリンク重みを与えることによって、送受信ノードが隔離ノードである場合にのみ、隔離ノードへのリンクが用いられる。障害発生時には、障害箇所を隣接ノードが知るだけで事前に準備しておいた RL を直ちに適用することができるため、そのノードが隔離ノードとなっている RL を選択して経路を切り替えることで、他のノード間の通信が切断されることなく即座に障害回復後の経路を用いることができる。

RRL が作成する RLS の例を図 1 に示す。各 RL においては、隔離ノードを送受信ノードとしない通信のためには、通常のリンクである実線のリンクのみが使用できる。図 1(a) は、元のネットワークトポロジを示しており、障害が発生していない状況ではこの RL_0 を用いてルーティングを行う。同様に、図 1(b) はノード 1, 2, 3, 4 を隔離ノードにした RL_1 を、図 1(c)

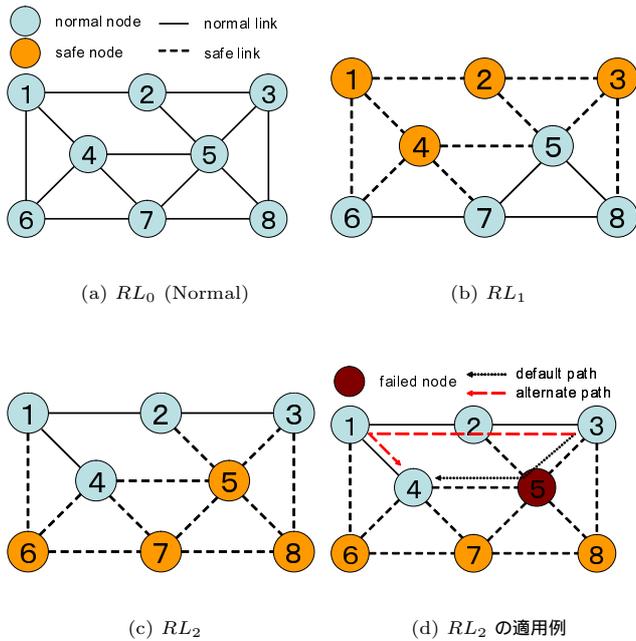


図 1 RLS の例

はノード 5, 6, 7, 8 を隔離ノードにした RL_2 をそれぞれ示している。このトポロジにおいてノード 5 に障害が発生した場合に、ノード 3 からノード 4 へ伝送を行うことを考える。元のトポロジによる通常のルーティングでは、ノード 3-5-4 という経路で伝送され、障害ノードを経由するためノード 4 まで伝送されない。しかし、ノード 3 がノード 5 の障害を検知することによって、図 1(d) のように RL_2 を用いることで、ノード 3-2-1-4 という経路が選択されるため、経路は長くなるもののノード 4 まで伝送が行われる。このように、RRL は送信ノードが障害を検知することによって、残りのノード間で障害の影響を受けることなく通信を行うことができる。

3.1.2 複数障害への対応

RRL はネットワーク中の全てのノードを RLS のいずれかの RL において隔離ノードにしておくことで、単一障害に対する 100% の復旧を可能にしている [11]。全てのノードを隔離ノードにするために必要な RL の個数は、千ノード程度のトポロジでは高々数十～数十であり、記憶量に制限のあるような環境下でも実用に耐える。さらに、1 つの RL 内で隔離ノードになっている複数のノードに対しては、それらのノードに同時に障害が発生した場合も、復旧が可能である。一方、1 つの RL において多くのノードを隔離ノードにすると、その RL において使うことができないリンクが増えるため、平均経路長が大きくなるという問題がある。また、多くの RL を用意することによって、より多くのパターンの障害に対応可能となるが、用意できる RL の個数は、ノードの記憶領域やルーティングを適用する際に必要となるパケットヘッダのビット数などにより制限を受ける。つまり、RL を何個用意し、それぞれの RL でどのノードをどれだけ隔離ノードにするか、という問題には、上述のような複雑なトレードオフが存在する。このため、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が、これまでに RRL を利用して複数障害に対応するための具体的な手法は提案されていない。そこで、3.2 節にて、複数障害に対して効果的な RLS 生成法を複数提案する。

3.1.3 オーバレイルーティングへの適用

RRL は様々なレイヤにおける経路制御に適用可能である [13] では、MPLS ネットワークにおける経路制御に適用する方法について検討されている。また、IP ネットワークにおいても、ルーティングに使用する RL を指定するために IP パケットヘッ

ダの未使用部分を利用することで、IP ルーティングに適用することができると考えられる。しかし、これらの方法は、既存プロトコルの変更が必要になるため、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り、既存の IP ネットワーク上での展開が可能となる、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。オーバレイルーティングを用いることによって、上述したようなプロトコルの標準化作業などを避けることができるため、サービスを早期に展開することが可能となる。また、通常は BGP によって利用可能な AS が限定されているピアリングリンクを、その他の AS 間通信における経路途中のリンクとして用いることが可能となる。これにより代替経路の選択肢が広がり、IP ルーティングでは不可能であった経路を用いることも可能である。さらに、ピアリングリンクを活用することによって、元のネットワークポロジを用いてルーティングを行うよりも平均経路長が短くなることも期待される。

障害が発生した際には、障害を検知したノードから知らされる障害ノードの情報を基に適用可能な RL を選択し、その RL の識別子をパケットヘッダに加えることで、他のノードに対して用いる RL を指定する。これにより、全てのノードに障害情報が伝わる以前から、RL を用いた経路制御が可能となる。また、RL のために必要となる識別子の大きさは、 $\log_2(n)$ [bits] (n : RL の個数) で抑えられるため、特にアプリケーションレベルでオーバレイルーティングを行う場合には、実用上問題にはならないと考えられる。

3.2 RLS 生成法

本稿においては、起こりうる可能性が高い複数同時障害として、隣接する複数ノードの障害、および同じ地域もしくは同じルータソフトウェア (OS) の同じバージョンなどの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として、充填ランダム法、一様ランダム法、および重複ランダム法、各ノードの次数を考慮して作成する次数降順法、および次数昇順法をあわせて説明する。なお、各アルゴリズムで用いるパラメータを表 1 に示す。

各手法で共通する隔離条件を示す。その RL においてノード n が以下の全ての条件を満たすとき、 n は隔離できる。

- (1) n が 1 つ以上の通常ノードに隣接している
- (2) 隣接する隔離ノードが n 以外の通常ノードに隣接している
- (3) n 以外の通常ノードからなるトポロジは連結グラフである

以下、各 RLS 生成法について説明する。なお、紙面の制約上、具体的なアルゴリズムの表記は省略する。

3.2.1 ハブ法

ハブ法は、ネットワーク全体の到達性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、ランダムにノードを隔離した RL を作成する。ハブ法は、ノードが RLS 中のいずれかの RL で高々 1 回しか隔離しない。

重複ハブ法は、ハブ法の特徴を残しつつランダム障害に対する適応力を高めた手法である。重複ハブ法は、ハブ法と同様にハブノードとその隣接ノードを隔離する RL と、ランダムにノードを隔離した RL を作成する。ただし、ハブ法とは異なり、隔離するノードが RLS 中の異なる RL で重複して隔離される。

3.2.2 属性法

属性法と重複属性法は、同じ地域にある、同じ OS の同じバージョン、および同じ管理ポリシーなどの属性が等しいノード群の障害を想定している。具体的には、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。属性法は、ノードが RLS 中のいずれかの RL で高々 1 回

しか隔離しない。重複属性法は、ノードが RLS 中の異なる RL で重複して隔離される。

3.2.3 ランダム法

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、ヒューマンエラー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。また、この生成法は計算量が小さく、短時間に RLS の生成が可能である。ランダムに選択したノードを隔離する RL の選び方によって、以下の 3 種類の生成法を用いる。

充填ランダム法は、記憶量に制限のあるネットワークでの適用を想定している。具体的には、ランダムに選択したノードをできるだけ同じ RL で隔離することを試みる。また、充填ランダム法は、ノードを RLS 中のいずれかの RL で高々 1 回隔離する。

一様ランダム法は、低頻度でランダムに障害が発生するネットワークでの適用を想定している。具体的には、ランダムに選択したノードを、RL 1 個当たりの隔離ノード数が $safe_{max}^{uni}$ 以下となるように隔離する RL を作成する。また、一様ランダム法は、ノードを RLS 中のいずれかの RL で高々 1 回隔離する。

重複ランダム法は、高頻度でランダムに障害が発生するネットワークでの適用を想定している。具体的には、ランダムに選択したノードを、隔離ノード数が $safe_{max}^{OL-rnd}$ 以下となる RL を、 L_{OL-rnd} 個作成する。このとき、ノードは RLS 中の RL で重複して隔離される。

3.2.4 次数法

次数法は、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、この生成法は、ランダム法と同様に計算量が小さく、短時間に RLS の生成が可能である。次数の昇順と降順の 2 種類の生成法を用いる。

次数降順法は、リンク数に比例して障害発生確率が増すネットワークでの適用を想定し、次数昇順法は、末端ノードほど障害発生確率が増すネットワークでの適用を想定している。また、次数法は、ノードを RLS 中のいずれかの RL で高々 1 回隔離する。

3.3 RL 選択法

3.3.1 静的 RL 選択

3.2 節で作成した RL を、実際のルーティングでどう用いるかについて、提案方式では 2 つの方法を対象とする。1 つ目の選択法として、送信ノードが選択する方法（静的 RL 選択）が考えられる。静的 RL 選択では、送信ノードが RL を選択するために、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保証されない。なお、この場合には、到達性を高めるためにより多くの障害ノードを隔離している RL を選択する。この方法を取ることによって、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。

3.3.2 動的 RL 選択

静的に RL を選択する場合、全ての障害ノードを隔離している RL がない場合には、回避できない障害ノードがあり、到達性が低下する。そこで、2 つ目の選択法として、より多くの障害ノードを回避するために実際のパケットフォワーディングを想定し、各中継ノードが RL を動的に選択する方法（動的 RL 選択）を用いる。各ホップでは、ルーティングに暫定的に用いる RL を決定し、パケットを転送する。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 はそのパケットの転送においては再度用いないようにする。そのため、各 RLS の RL を隔離ノード数順にソートし、隔離ノード数が少ない RL から順に暫定 RL として適用する。

4. 数値計算による評価

4.1 評価方法

評価に用いるネットワークトポロジは、CAIDA [14] が BGP トラフィックの計測を行い公開している AS のネットワークトポ

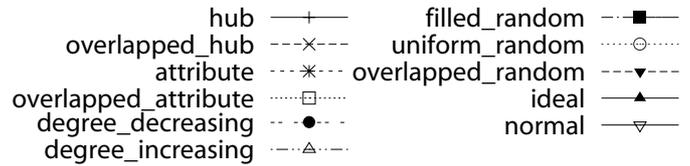


図 2 各 RLS 生成法の凡例

ロジのデータベースから、JPNIC が管轄する AS のみをノードとして抽出し作成する。この際に、リンクを複数持たないノードは代替経路を持つことができないため、次数が 1 のノードとその隣接ノード、および直列に繋がっているノード群と繋がっている先のノードをグループ化して 1 つのノードと見なす。その結果、対象となるネットワークポロジは、ノード数が 259、リンク数が 1162、平均次数が 4.4 となる。

発生させる障害は、ランダムに選択した複数のノードに障害を発生させるランダム障害 (F_RND)、隣接する複数のノードに障害を発生させる隣接ノード障害 (F_ADJ)、特定の属性を持つ複数のノードに障害が発生する属性ノード障害 (F_ATR)、および Internet eXchange (IX) の障害を想定したリンク障害 (F_LNK) である。

なお、各ノードに属性を与えるために、ネットワークを分割グラフをまたぐリンクの数（カットサイズ）が小さくなるようにグラフを分割し、分割結果に応じて各ノードに属性を持たせる。そのため、作成された分割グラフはリンクが密な集合となるため、隣接ノード障害に対して属性法が有効となることが予想される。

表 1 に、3.2 節で説明した RLS 生成法が用いるパラメータと各 RLS の RL の個数を示す。CASE 1 は、各ノードの記憶領域に制限がある場合を想定しており、各 RLS の RL の個数が約 10 個となるようにパラメータを調整している。CASE 2 は、各ノードの記憶領域が十分にある場合を想定しており、RLS の RL の個数に制限を設けていない。CASE 1 では動的 RL 選択について、CASE 2 では静的および動的 RL 選択について評価を行う。

評価指標として、障害発生時に、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、および到達可能なすべてのノードペア間の経路長の平均を用いる。これらの評価基準として、性能下限を意味する元のトポロジを用いた場合 (normal) と、上限を意味する障害ノードを除き経路の再計算を行う、理想的な場合 (ideal) を導入する。図 2 に以降の評価結果を示すグラフ中で用いる各手法の凡例を示す。

4.2 静的 RL 選択

図 3 に、CASE 2 における、障害発生ノード数に対する静的 RL 選択時の到達性の変化を示す。この図より、RL 1 個当たりの隔離ノードが多く、RL の個数が最も多い重複ランダム法が総じて到達性が高いことがわかる。特にランダムノード障害 (図 3(a)) では、2 ノード障害発生時に 98% の到達性を 99.99% まで回復している。また、属性ノード障害 (図 3(c)) においては、重複属性法が最も到達性が高く、こちらは 2 ノード障害発生時に、98% の到達性を 99.9% まで回復している。これは、属性に応じたノードの複数同時障害に対して効果的な RLS を生成していることが要因であると考えられる。リンク障害 (図 3(d)) においては、2 ノード間の IX 障害では、その他の障害発生方法とは異なり、RL 1 個当たりの隔離ノード数が少ない次数法などの方が到達性の向上割合が大きい。これは、本稿において想定したリンク障害は、次数の高いノードが持つリンクに発生しやすく、リンクの両端のノードを同一の RL で隔離する必要があり、および RL 1 個当たりの隔離ノード数が少ないと次数の高いノードを同一の RL で隔離しやすいことが要因と考えられる。また、障害の発生方法に関わらず、障害ノード数が増えると、障害の発生パターンが爆発的に増加するため、到達性の大きな改善は見られないことがわかる。

表 2 に、CASE 2 における静的 RL 選択時の平均経路長を示す。表中の各欄には、同時障害ノード数がそれぞれ 2/10/20

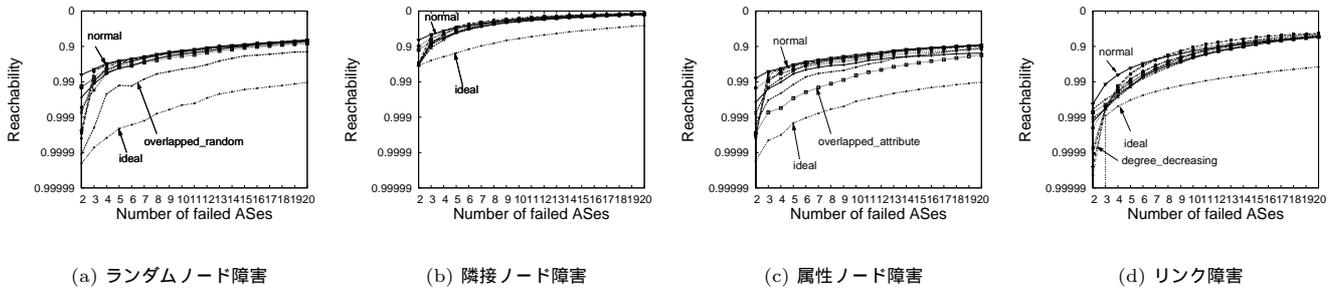


図 3 静的 RL 選択法における到達性 (CASE 2)

表 1 RLS 生成法が用いるパラメータと各 RLS の RL の個数

| | | CASE 1 | CASE 2 |
|-----------------------|--------------------------------|-----------|-----------|
| パラメータ名 | 用途 (仕様 RLS 生成法) | 値 | |
| L | 作成する RL の個数 (ハブ法) | 10 | 259 |
| L_{random} | ランダム部分の RL の個数 (ハブ法, 重複ハブ法) | 3 | 3 |
| Deg_{min} | ハブノードの最低次数 (重複ハブ法) | 40 | 20 |
| L_{hub} | ハブノード 1 つから生成する RL の個数 (重複ハブ法) | 2 | 38 |
| A | ノードの属性集合 (属性法, 重複属性法) | $ A = 2$ | $ A = 4$ |
| L_{attr} | 1 属性から作成する RL の個数 (重複属性法) | 2 | 60 |
| $safe_{max}^{uni}$ | RL の隔離ノード数の上限 (一様ランダム法) | 26 | 26 |
| L_{OL-rnd} | 作成する RL の個数 (重複ランダム法) | 10 | 2000 |
| $safe_{max}^{OL-rnd}$ | RL の隔離ノード数の上限 (重複ランダム法) | 259 | 259 |
| RLS 生成法 | | RL の個数 | |
| ハブ法 | | 11 | 260 |
| 重複ハブ法 | | 11 | 269 |
| 属性法 | | 13 | 254 |
| 重複属性法 | | 13 | 254 |
| 充填ランダム法 | | 7 | 7 |
| 一様ランダム法 | | 12 | 12 |
| 重複ランダム法 | | 10 | 2000 |
| 次数昇順法 | | 12 | 12 |
| 次数降順法 | | 12 | 11 |

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している。表より、重複ランダム法の平均経路長が総じて長いことがわかる。これは、RL 1 個当たりの隔離ノードが多いため、利用可能なリンクが限られるためであると考えられる。一方、他の RLS 生成法では、平均経路長は元のトポロジを用いる場合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、通常の IP ルーティングでは用いることができないピアリングリンクを途中の経路として用いることができるためだと考えられる。

4.3 動的 RL 選択

図 4 に、CASE 1 における、障害発生ノード数に対する動

表 2 静的 RL 選択法における平均経路長 (CASE 2)

| 障害:2/10/20 | F_RND | F_ADJ | F_ATR | F_LNK |
|------------|----------------|----------------|----------------|----------------|
| ハブ法 | 2.78/2.91/2.92 | 2.80/2.77/2.68 | 2.79/2.88/2.86 | 2.74/2.88/2.80 |
| 重複ハブ法 | 2.83/2.74/2.70 | 2.87/2.58/2.20 | 2.84/2.74/2.71 | 2.89/2.84/2.71 |
| 属性法 | 2.73/2.71/2.68 | 2.79/2.61/2.48 | 2.73/2.71/2.69 | 2.81/2.79/2.78 |
| 重複属性法 | 2.79/2.81/2.77 | 2.80/2.72/2.57 | 2.83/2.76/2.72 | 2.86/2.85/2.79 |
| 充填ランダム法 | 2.78/2.73/2.70 | 2.88/2.46/2.15 | 2.78/2.73/2.71 | 2.90/2.79/2.71 |
| 一様ランダム法 | 2.75/2.74/2.72 | 2.76/2.69/2.56 | 2.75/2.75/2.73 | 2.79/2.86/2.78 |
| 重複ランダム法 | 2.99/2.96/2.89 | 2.99/2.75/2.47 | 2.99/2.97/2.92 | 2.98/3.01/2.93 |
| 次数降順法 | 2.72/2.69/2.66 | 2.79/2.53/2.36 | 2.71/2.69/2.67 | 2.79/2.68/2.73 |
| 次数昇順法 | 2.72/2.68/2.66 | 2.80/2.58/2.45 | 2.72/2.68/2.66 | 2.80/2.78/2.78 |
| 元のトポロジ | 2.84/2.82/2.80 | 2.82/2.73/2.60 | 2.84/2.83/2.81 | 2.85/2.82/2.78 |
| 理想経路 | 2.70/2.73/2.77 | 2.78/3.30/4.02 | 2.70/2.72/2.73 | 2.70/2.72/2.80 |

表 3 動的 RL 選択法における平均経路長 (CASE 1)

| 障害:2/10/20 | F_RND | F_ADJ | F_ATR | F_LNK |
|------------|----------------|----------------|----------------|----------------|
| ハブ法 | 2.88/2.88/2.91 | 2.90/3.13/3.15 | 2.87/2.88/2.89 | 2.82/2.85/2.85 |
| 重複ハブ法 | 2.85/2.83/2.89 | 2.89/3.13/3.15 | 2.85/2.84/2.88 | 2.89/2.85/2.86 |
| 属性法 | 2.84/2.88/2.91 | 2.91/3.12/3.12 | 2.76/2.83/2.88 | 2.83/2.85/2.85 |
| 重複属性法 | 2.89/2.88/2.91 | 2.92/3.12/3.11 | 2.88/2.84/2.86 | 2.98/2.87/2.87 |
| 充填ランダム法 | 2.77/2.84/2.89 | 2.92/3.11/3.11 | 2.86/2.88/2.89 | 2.86/2.85/2.87 |
| 一様ランダム法 | 2.85/2.88/2.91 | 2.93/3.12/3.12 | 2.85/2.88/2.89 | 2.83/2.86/2.86 |
| 重複ランダム法 | 3.10/2.93/2.91 | 2.96/3.05/2.94 | 2.96/2.92/2.89 | 2.95/2.91/2.86 |
| 次数降順法 | 2.74/2.84/2.89 | 2.92/3.11/3.11 | 2.74/2.83/2.88 | 2.80/2.85/2.85 |
| 次数昇順法 | 2.79/2.88/2.91 | 2.93/3.11/3.12 | 2.78/2.88/2.89 | 2.80/2.85/2.85 |
| 元のトポロジ | 2.84/2.83/2.80 | 2.83/2.74/2.60 | 2.84/2.83/2.80 | 2.85/2.82/2.78 |
| 理想経路 | 2.70/2.73/2.77 | 2.79/3.34/4.10 | 2.70/2.73/2.76 | 2.70/2.73/2.80 |

的 RL 選択時の到達性の変化を示す。図から、隣接ノード障害 (図 4(b)) を除き、理想的なルーティングを行った場合とほぼ同等の到達性を実現していることがわかる。属性ノード障害 (図 4(c)) においては、20 ノードの同時障害発生時にも、87%の到達性を 98%まで回復している。しかし、隣接ノード障害発生時には到達性があまり改善されていない。これは、隣接ノード障害は次数の高いノードほど障害が発生しやすいため、他の障害発生方法と比べて複数のハブノードに同時に障害が発生しやすいことが要因であると考えられる。

また、動的 RL 選択を用いた場合においては、各 RLS 生成法で到達性に大きな差は見られないことがわかる。これは、いずれの RLS でも動的 RL 選択を用いることで、障害発生規模、および障害発生パターンに関わらず、到達性を確保できるためである。また、複数のハブノードに障害が発生した場合には、どの手法においても到達性を向上させることができないことも要因の 1 つであると考えられる。

図 5 に、CASE 2 における、障害発生ノード数に対する動的 RL 選択時の到達性の変化を示す。図より、CASE 2 においては、各 RLS の RL の個数が多いため、ハブノードを同時に隔離している RL が多くなり、全ての障害発生方法に対して、理想的なルーティングを行った場合の到達性に CASE 1 よりも近づいていることがわかる。

表 3 に CASE 1 における動的 RL 選択時の平均経路長を示す。表中の各欄には、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している。表より、静的 RL 選択時と同様に、提案手法は動的 RL 選択時にも平均経路長をほぼ増加させないことがわかる。

5. おわりに

本稿では、現状の IP 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

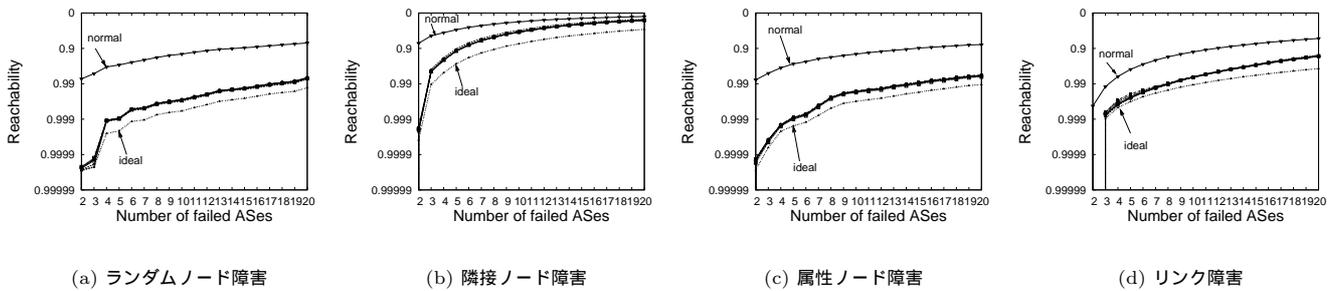


図 4 動的 RL 選択法における到達性 (CASE 1)

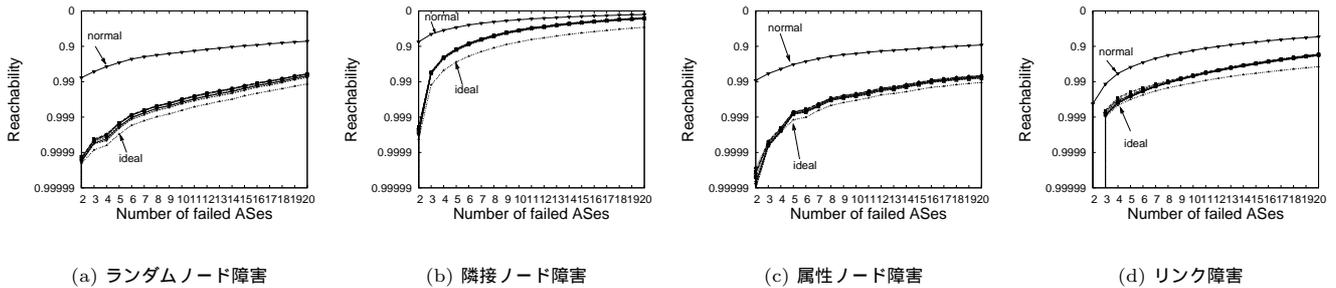


図 5 動的 RL 選択法における到達性 (CASE 2)

術を用いて、大規模なネットワーク障害が発生した際に従来の BGP による IP ルーティングでは到達不可能、あるいは到達可能な状態までの回復に長い時間を要する 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 選択を用いた場合においては、同時障害を発生しやすいノード群を考慮した 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 選択を用いた場合においては、平均経路長をほとんど増加させることなく、理想的な場合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。これらの結果から、提案手法は大規模なネットワーク障害に対して有効な手法であるといえる。

今後の課題としては、複数のハブノードに同時に障害が発生した場合における効果的な RLS の生成法、およびオーバーレイに参加するノードの割合を考慮した評価の検討が挙げられる。また、障害発生後のトラフィック負荷を分散させることを考慮した RLS の生成法を検討したい。

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。

文 献

- [1] 岡田忠信, “次世代情報通信ネットワークの高信頼化,” 電子情報通信学会総合大会講演論文集, vol. 2005, no. 2, pp. SS-9-SS-10, Mar. 2005.
- [2] Y. Rekhter and T. Li, “A border gateway protocol 4 (BGP-4),” *RFC 1771*, 1995.
- [3] C. Labovitz, A. Ahuja, A. Abose, and F. Jahanian, “Delayed internet routing convergence,” in *ACM SIGCOMM 2000*, vol. 9, no. 3, Aug. 2000, pp. 293-306.

- [4] B. Zhang, D. massey, and L. Zhang, “Destination reachability and BGP convergence time,” in *GLOBECOM 2004*, vol. 3, Apr. 2004, pp. 1383-1389.
- [5] A. S. Tanenbaum, *COMPUTER NETWORKS*, 3rd ed. Prentice-Hall International, Inc., 1996.
- [6] C. Labovits, A. Ahuja, R. Wattenhofer, and S. Venkatachry, “The impact of Internet policy and topology on delayed routing convergence,” in *Proceedings of INFOCOM 2001*, Dec. 2001, pp. 537-546.
- [7] Z. M. Mao, R. Govindan, G. Varghese, and R. H. Katz, “Route flap damping exacerbates Internet routing convergence,” *ACM SIGCOMM Computer Communication Review*, vol. 32, no. 4, pp. 221-233, Oct. 2002.
- [8] D. Pei, M. Azuma, D. Massey, and L. Zhang, “BGP-RCN: Improving BGP convergence through root cause notification,” UCLA CSD, Tech. Rep. CO80523-1873, Dec. 2004.
- [9] D. Andersen, H. Balakrishnan, M. Kaashoek, and R. Morris, “Resilient overlay networks,” in *Proceedings of the 18th ACM Symposium on Operating Systems Principles*, Oct. 2001.
- [10] Z. Xu, M. Mahalingam, and M. Karlsson, “Turning heterogeneity into an advantage in overlay routing,” in *INFOCOM 2003*, vol. 2, Apr. 2003, pp. 1499-1509.
- [11] 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.
- [12] A. Kvalbein, T. Čičić, and S. Gjessing, “Post-failure routing performance with multiple routing configurations,” in *Proceedings of INFOCOM 2007*, May 2007, pp. 98-106.
- [13] A. Hansen, A. Kvalbein, T. Čičić, S. Gjessing, and O. Lysne, “Resilient routing layers for recovery in packet networks,” in *Proceedings of the 2005 International Conference on Dependable Systems and Networks*, Jul. 2005, pp. 238-247.
- [14] The CAIDA Web Site, available at <http://www.caida.org/home/>.