

経路制御におけるネットワーク可用性向上に関する研究

筑波大学大学院 ビジネス科学研究科 企業科学専攻 システムマネジメントコース

鈴木一哉

主指導教員：吉田健一教授 副指導教員：久野靖教授、大木敦雄准教授、倉橋節也准教授

2011/5/30

1 はじめに

近年インターネットに代表される Internet Protocol を用いたネットワークは、我々の生活に欠かせない社会インフラとして、その存在意義を増している。特に VoIP に代表されるリアルタイム系のアプリケーションが広く用いられるようになっており、IP ネットワークには従来以上の高い可用性が求められている。しかし、インターネットの可用性は 98.1 ~ 99.3% と言われており、社会インフラとして求められる可用性 (99.99%) との差は大きい。そのため、IP ネットワークの可用性の向上が急務となっている。

このような背景を鑑み、本研究ではまず到達不能ノードの発生要因の分析 [1]を行う。この分析結果を用いた局所化アルゴリズムの提案を行う。さらに、この局所化アルゴリズムを用いた二つの高速復旧手法である事前計算型経路更新手法 [2]と故障箇所高速迂回手法 [3]の提案を行う。これらの二つの手法により故障発生時における復旧時間を短縮することで、ネットワーク可用性の向上を実現する。

2 関連研究と本研究の位置づけ

2.1 可用性向上技術の分類

可用性を向上するためには、(1) 障害を起りにくくする、もしくは (2) すばやく障害から復旧するという二つのアプローチが考えられる。前者に対しては、ノードやリンクをそれぞれ二重化し、一方の故障を他方がカバーすることにより、障害を防ぐというアプローチで、さまざまな研究が行われている。一方で後者に関しては、予めノードやリンクを冗長に構成しておき、故障発生時にはこれらを切り替えて使用するというアプローチで、経路制御技術と呼ばれる技術が確立されている。これらの技術をまとめたのが、図 1 である。

2.2 経路制御による高速な復旧

経路制御による切替動作は、故障の検知、故障の通知、経路表の再構築という各フェーズから構成されている。故障検知を高速に行う手法として、Bidirectional Forwarding Detection (BFD) や下位レイヤの情報を用いて検知を行う手法が提案され、実用化されている。また経路表の再構築時間を短縮するために、故障前後の最短パスツリーにおける差分だけを再計算する手法が提案されている。

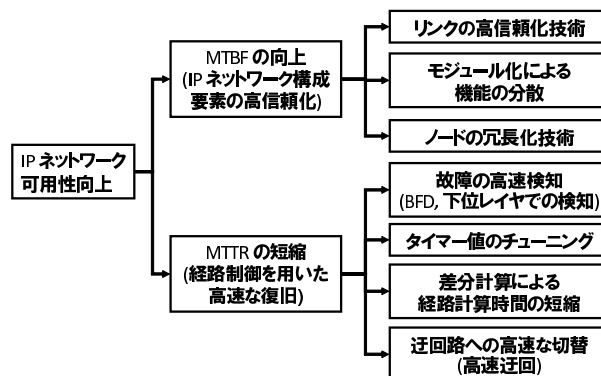


図 1: 可用性向上技術の分類

また、これら復旧の各フェーズには、安定化を目的とした各種タイマー値が設定されており、完全に切替が完了するまでに数百ミリ秒から数秒の時間がかかる。これに対し、Francoisらは、これらタイマー値をチューニングすることで、切替時間の短縮を実現している [4]。

これらの手法を活用することにより、復旧時間の短縮をある程度実現可能であるが、依然として故障の通知時間や経路の再計算時間には数百ミリ秒の時間を必要とする。

2.3 高速迂回手法

故障通知や経路表再構築にかかる時間の短縮を目的として、高速迂回 (Fast Reroute) と呼ばれる手法 [5, 6, 7] が提案されている。これらの手法では、故障を検知したノードが即座に、パケットを代替パスへ迂回させる。この代替パスを事前に設定しておくことで、故障通知やパスの再計算などの動作を故障後に行うことなく、パケットの到達性を復旧させることが可能となる。しかしこれらの手法では、(1) 転送方式の拡張もしくは (2) 管理サーバによる集中計算の、いずれかもしくは両方を必要とする。自律分散で動作を行っている既存の IP ネットワークへの適用を考えた場合、これらのことが導入への制限となっていた。

2.4 本研究の位置づけ

本研究では、故障からの復旧時間の短縮を目的とする。そのため、まず故障発生時における到達不能ノードの発生要因の分析を行う [1] さらに、この分析結果を用いて、パケット到達性復旧のために経路更新が

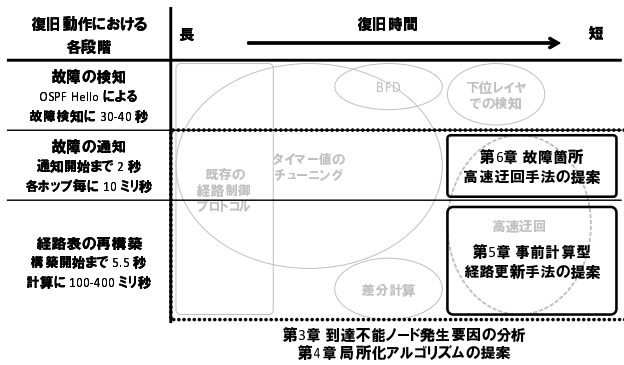


図 2: 関連研究と本研究の位置づけ

必要なノードを特定する局所化アルゴリズムの提案を行う。このアルゴリズムを用いて、リンク故障時に使用する代替経路表を事前に計算しておくことで、経路表の計算時間を短縮する事前計算型経路更新手法の提案を行う [2]。また、故障を検知したノードがトンネルを用いてパケットを迂回させることで、故障通知時間を不要とする故障箇所高速迂回手法を提案する [3]。これらの手法により、従来技術で課題となっていた故障通知および経路更新の時間を実現する。

3 到達不能ノードの発生要因の分析

故障発生時その故障は、経路制御プロトコルにより、ネットワーク中の全ノードに通知され、各ノードはそれぞれ経路表の更新を行う。しかし、パケット到達性の維持という観点から見ると、故障が影響を与えるのはその近傍のノードのみであり、必ずしもネットワーク中の全ノードが経路更新を行う必要はないと考えられる。本研究では、この仮説を検証すべく、ネットワークのモデル化を行い、そのモデルを用いてシミュレーションを行った。

ここで、あるノード Y のみが故障に対して経路表の更新を行わなかった場合、パケットの到達性にどのように影響を与えるかをシミュレーションを用いて調べた。1000 ノードのネットワークにおいて、ノード X に故障が発生したと仮定する。このとき、ノード Y から到達不能になるノードの数を縦軸に、ノード X, Y 間の距離を横軸にあらわしたのが、図 3 である。この図が示すように、故障距離が近いときにより多くのノードへ到達不能となり、距離が遠くなるにつれて到達不能ノード数が減ることが確認できた。

また、故障発生を行ったときに到達不能となる原因について分析を行った結果、以下の四つに分類できることが分かった。

1. 宛先ノード自体の故障
2. 故障により宛先ノードまでのパスの喪失
3. 隣接ノードの故障
4. ループの発生

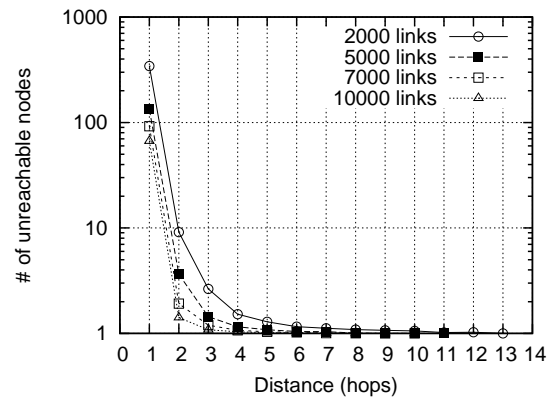


図 3: 到達不能ノード数と故障ノードまでのホップ数の関係

このうち、はじめの二つは経路制御以外で解決すべき問題である。一方で 3, 4 の要因による故障に関しては、パケット到達性を維持するためには経路更新を行う必要がある。しかし、隣接ノードでなくかつループの発生する可能性がない場合には、経路更新を行わずともパケット到達性を維持できると考えられる。この結果は、次章の局所化アルゴリズムの基本的アイデアとなっている。

4 局所化アルゴリズムを用いた事前計算型経路更新手法の提案

本章では、局所化アルゴリズムを用いた事前計算型経路更新手法の提案を行う。

本手法の目的は、リンク故障発生時の切替時間のうち、経路計算時間を短縮することにある。そのためには、故障に対応した経路計算を事前に行い、故障時に使用する代替経路表として保持しておけばよい。しかし、想定されるすべての故障に対して、それぞれ事前に経路計算を行うことは現実的ではない。そのため本手法では、あるリンク故障に対し、経路更新を行わなければループが発生するノード (影響ノード) の特定を行い、これらのノードのみがそのリンクに対する代替経路表を計算する。このことにより、それぞれのノードが用意すべき代替経路表数を、一般的なルータ装置に実装可能まで削減する。

4.1 局所化アルゴリズム

ここでは、影響ノードを特定するためのアルゴリズムについて、説明を行う。このアルゴリズムでは、以下の条件を判定することで、影響ノードを特定する。

1. 故障前の最短パスツリーに故障リンクを含む
2. 故障リンクに接続するノードからの故障後の最短パスツリーに含まれる

条件 1 を満たさないノードは、故障前後で宛先までの最短パスに変化がない。つまり、代替経路表を用意する必要のあるノードは、すべて条件 1 を満たしていると考えられる。さらに、条件 2 の故障発生後に故障リンクに接続されるノードからの最短パスツリーに含まれるということは、この条件を満たすノード上にリンク故障に伴う迂回パスが構成されることを意味する。つまり、これらの二つの条件を満たすノードのみが、故障リンクに対する代替経路表を用意すればよい(図 4)。

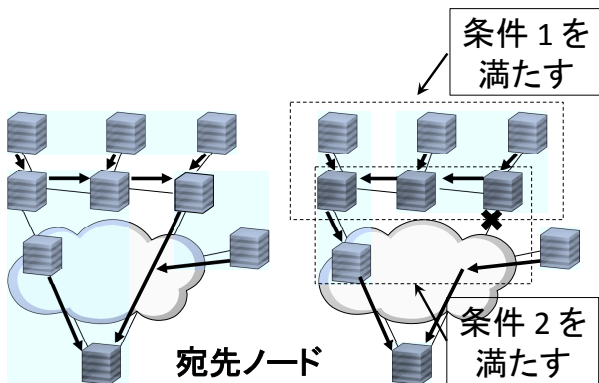


図 4: 提案手法で代替経路表を用意するノード

4.2 評価

提案手法における代替経路表を用意すべきリンク数について、シミュレーションを用いた評価を行った。評価では、ノード数 1000 とし、リンク数はそれぞれ異なる各ネットワークに対し、従来手法と提案手法をそれぞれ適用した場合について調査を行った。その結果をまとめたのが、表 1 である。表 1 を参照すると、提案手法は、従来手法と比べ $1/60 - 1/100$ となっている。つまり、提案手法を用いることで、代替経路表を必要とするリンク数を大幅に削減できていることが確認できた。

表 1: 代替経路表を用意すべきリンク数

全リンク数	1997	2994	3990	4985
従来手法 1	1997	2994	3990	4985
従来手法 2	999	999	999	999
提案手法	10.6	12.6	14.7	16.8

5 局所化アルゴリズムを用いた故障箇所高速迂回手法の提案

本章では、局所化アルゴリズムを用いた故障箇所高速迂回手法の提案を行う。

前章の事前計算型経路更新手法では、故障復旧時間のうち、経路計算時間の短縮を実現していたが、故障通知時間に関しては考慮されていなかった。本手法は、故障を検知したノードが、パケットを即座にトンネルへと迂回させることにより、中断なくパケット転送を継続させることができる。他のノードへ故障を通知する必要がないため、事前計算型経路更新手法と比較し、より短い時間で故障からの復旧が可能である。

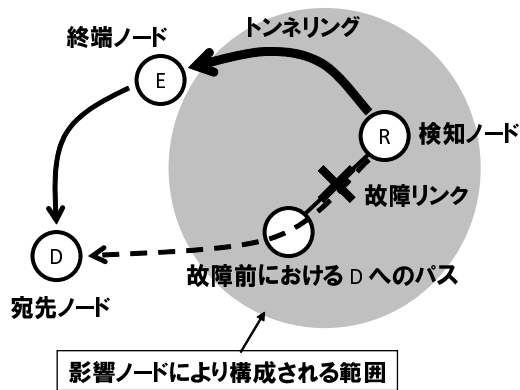


図 5: 故障箇所高速迂回手法

5.1 評価

本章で提案した高速迂回手法と、前章提案の経路更新手法との比較評価を行った。

まず、それぞれの手法において必要となる代替経路表を、計算するためのコストおよび格納するためのメモリ量をシミュレーションを用いて調べた。ノード数 1000、平均次数 4 のネットワークに対してシミュレーションを行い、その結果をまとめたのが表 2 である。1000 ノードのネットワークにおける IPv6 の通常経路表のサイズは、おおよそ 35 キロバイト程度になる。表 2 の結果から、経路更新手法および高速迂回手法それぞれにおいて代替経路表を格納するのに必要なメモリ量は、最大で 184 キロバイトおよび 35 キロバイト程度となる。これらの代替経路表を格納するのに DRAM を使用した場合、両手法とも実装上の問題となるサイズではない。両手法を比較した場合、高速迂回手法の方が必要とする計算コストおよびメモリ量共に少ない。

表 2: 代替経路表のために必要なリソース (通常経路表に必要なリソースを 1 とした場合)

	経路更新手法		高速迂回手法	
	平均	最大	平均	最大
計算コスト	1.29	2.92	1.01	1.06
メモリ量	2.31	5.27	1.0	

次に、各手法における迂回パス長のオーバーヘッドについて評価を行った。図 6 中の R-D 間のリンクに

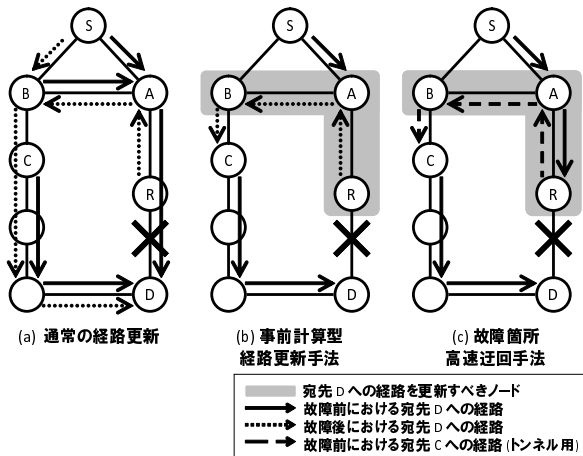


図 6: 各手法において生成される迂回パス

故障が発生した場合における S-D 間の迂回パス長は、通常の経路更新を行った場合 5 になる。経路更新手法では、ノード S は影響ノードではなく経路更新を行わないため、迂回パス長は 6 となる。さらに高速迂回手法では、ノード D 宛ての packets は一旦ノード R まで転送された後トンネルを使って終端ノード C へと転送されるため、迂回パス長は 8 となる。200 ノードのネットワークにおいて、両手法における迂回パス長が、通常の経路更新の場合と比較して何 % 長くなるかを調べた結果を、図 7 に示す。この結果によれば、両手法とも通常の経路更新に対してオーバーヘッドが存在するが、その割合は 0.2 % 以下であり、実用上大きな問題にはならないと考えられる。

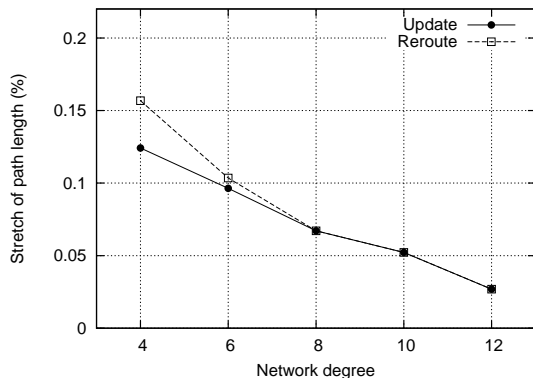


図 7: パス長のオーバーヘッド

6 おわりに

まず、到達不能ノードの発生要因について解析を行った。故障に対し経路更新を行わなかったとき、パケット到達性にどのような影響がでるかをシミュレーションを用いて調査し、以下の事実が確認できた。

- 故障距離が近いときにより多くのノードへ到達不

能となり、距離が遠くなるにつれて到達不能ノード数が減る。

- 到達不能要因は四つのいずれかに分類可能である。

また、この解析結果を用いて局所化アルゴリズムを提案し、このアルゴリズムを用いた事前計算型経路更新手法の提案を行った。この手法では、故障発生時にパケット到達性の復旧に必要なノードのみが、その故障に対する代替経路表を事前に計算しておくことで、事前計算が必要な代替経路表の数を従来手法と比べ、平均で 1/60 - 1/100 に抑えることができた。

さらに、局所化アルゴリズムを用いた故障箇所高速迂回手法の提案を行った。この手法では、故障を検知したノードがトンネルを用いてパケットを即座に迂回させる。つまり故障通知が不要であるため、故障発生後、より短時間で復旧を実現できる。また代替経路表の計算に必要なコストおよび格納するためのメモリ量は、共に通常経路表で必要なそれとほぼ同等であり、一般的なルータ装置への実装に対して問題とならない。また、迂回パス長に関してはオーバーヘッドが存在するが、実用上大きな問題とならないサイズである。

これらの成果を活用することにより、IP ネットワークの可用性を従来よりも飛躍的に向上させることが可能であると考えている。今後 IP ネットワークの適用範囲がさまざまな分野に広がり、多くの人々がいつでもその恩恵を受けられるようになれば幸いである。

参考文献

- [1] K. Suzuki, et al. : "Formalization and Analysis of Routing Loops by Inconsistencies in IP Forwarding Tables", IEICE Trans. Comm., vol.E90-B, no.10, pp.2755-2763, 2007.
- [2] K. Suzuki, et al. : "Selective Precomputation of Alternate Routes using Link-State Information for IP Fast Restoration", IEICE Trans. Comm., vol.E93-B, no.05, 2010.
- [3] K. Suzuki, et al. : "Comparison of Proactive and Reactive Methods for IP Fast Restoration using Localization Algorithm", Proceedings of the ICSPCS, IEEE, 2010.
- [4] P. Francois, et al. : "Achieving sub-second IGP convergence in large IP networks", ACM SIGCOMM CCR, Vol. 35, No. 3, pp. 35-44, 2005.
- [5] J. Wang, et al. : "IP fast reroute with failure inferring", Proceedings of the ACM SIGCOMM Workshop on INM, pp. 268-273, 2007.
- [6] A. Kvalbein, et al. : "Fast IP Network Recovery Using Multiple Routing Configurations", Proceedings of the IEEE INFOCOM, pp. 1-11, 2006.
- [7] M. Shand, et al. : "IP Fast Reroute Using Not-via Addresses", draft-ietf-rtgwg-ipfrr-notvia-addresses-05, IETF, 2010.